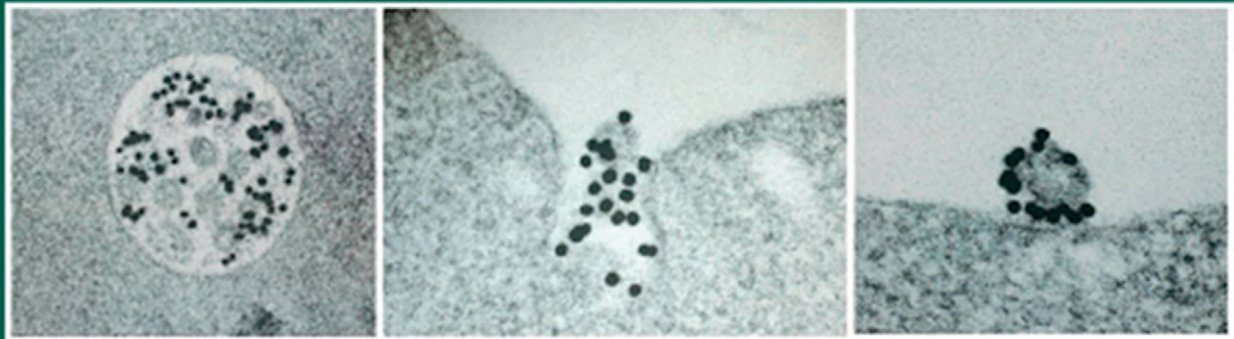


Encyclopedia of Cell Biology



EDITORS IN CHIEF

Ralph A. Bradshaw and Philip D. Stahl



How to go to your page

This eBook contains Four Volumes. It contains an Index at the end of each volume.

The front matter and content of each book have their own page numbering scheme, consisting of a volume number and a page number, separated by a Colon.

For example, to go to page 18 of Volume 1, type "V1:18" in the "page #" box at the top of the screen and click "Go."

To go to page "123" of Volume 2, type "V2:123" or to go to page "vii" of Volume 3, type "V3:vii"and so forth.

Please refer to the eTOC for further clarification.

ENCYCLOPEDIA OF CELL BIOLOGY

ENCYCLOPEDIA OF CELL BIOLOGY

EDITORS-IN-CHIEF

RALPH A BRADSHAW

University of California, Irvine, CA, USA

PHILIP D STAHL

Washington University School of Medicine, St. Louis, MO, USA

VOLUME 1

MOLECULAR CELL BIOLOGY

VOLUME EDITOR

GERALD W HART

Johns Hopkins University School of Medicine, Baltimore, MD, USA

SECTION EDITORS

KENNETH E NEET

Rosalind Franklin University of Medicine and Science, North Chicago, IL, USA

SARAH A WOODSON

Johns Hopkins University School of Medicine, Baltimore, MD, USA

JUDITH S BOND

Penn State University College of Medicine, Hershey, PA, USA



AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
225 Wyman Street, Waltham, MA 02451, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

Copyright © 2016 Elsevier Inc. All rights reserved

Cover background image created by M W Goldberg and T D Allen.

Foreground images © Clifford Harding, John Heuser and Philip D Stahl, 1983, originally published in *The Journal of Cell Biology*, Volume 97, 329–339.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-394447-4

For information on all publications visit our
website at <http://store.elsevier.com>

Printed and bound in the United States of America

15 16 17 18 19 20 10 9 8 7 6 5 4 3 2 1



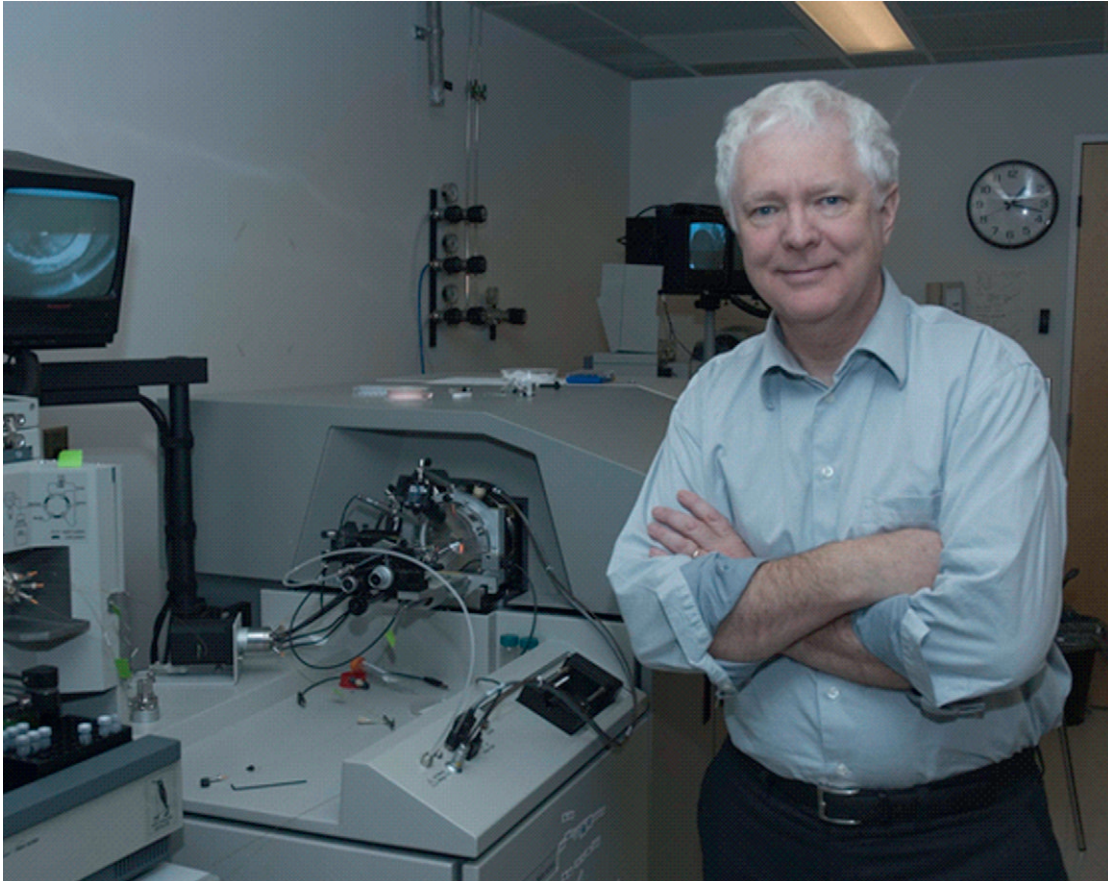
Working together
to grow libraries in
developing countries

www.elsevier.com • www.bookaid.org

Publisher: Adam Gilbert
Acquisition Editor: Ginny Mills
Content Project Manager: Blerina Osmanaj
Associate Content Project Manager: Victoria Dyer
Cover Designer: Maria Inês Cruz

DEDICATION

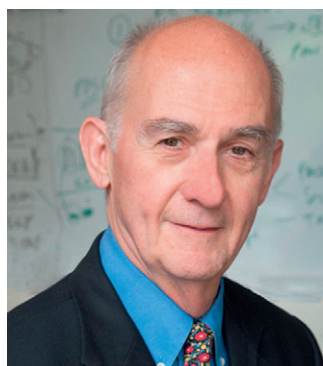
Dedicated to Anthony James (Tony) Pawson (1952–2013), in recognition of his outstanding contributions to cell biology and our understanding of intracellular communication mechanisms...



EDITORS-IN-CHIEF



Ralph A Bradshaw is Professor Emeritus in the Department of Physiology and Biophysics at the University of California, Irvine. Prior to that he was on the faculty of the Department of Biological Chemistry, Washington University School of Medicine in St. Louis, MO and was Professor and Chair of the Department of Biological Chemistry at University of California, Irvine. From 2006 to 2015, he was a member of the Mass Spectrometry Facility and Professor of Pharmaceutical Chemistry at the University of California, San Francisco. He holds degrees from Colby College and Duke University and was a post-doctoral fellow at Indiana University and the University of Washington. He has served as president of FASEB, was the founding president of the Protein Society and was the treasurer of the American Society for Biochemistry and Molecular Biology. His research has focused on protein chemistry and proteomics, with emphasis on the structure and function of growth factors and their receptors, particularly nerve growth factor and fibroblast growth factor, and the involvement of receptor tyrosine kinases in cell signaling. He has also studied the role of proteolytic processing and N-terminal modification in protein stability and turnover.



Philip D Stahl is E. Mallinckrodt Jr. Professor Emeritus at Washington University School of Medicine in St. Louis, MO. He was educated at West Virginia University with post-doctoral work at Vanderbilt University. He served as Head of the Department of Cell Biology and Physiology and Director of the Division of Biology and Biomedical Sciences at Washington University. He has been the recipient of many awards including a MERIT award from the NIH and the WICB Senior Recognition award given by the American Society for Cell Biology in recognition of his work supporting the advancement of women in science. Among StahlLab contributions are the discovery of the lysosomal enzyme clearance pathway now implemented in the treatment of lysosomal storage disease, discovery of the innate immune receptor, the mannose receptor, discovery of the exosome secretion pathway, and the role of Rab5 and Arf6 in endocytosis.

Currently his research focuses on endocytosis, signal transduction, and exosome biogenesis and secretion.

VOLUME EDITORS



Gerald Hart is the Director and DeLamar Professor of Biological Chemistry at Johns Hopkins Medical School. He began his research on glycoconjugates about 40 years ago as a graduate student, and he has been active in the field of Glycobiology ever since. Among his research accomplishments, he determined the minimal sequence requirements for N-linked glycosylation while a postdoc in William Lennarz's lab, and he collaborated with Paul Englund's group at Johns Hopkins to elucidate the biosynthetic pathway for GPI-anchors. In the early 1980s, while probing cells with glycosyltransferases, Hart's laboratory discovered cytoplasmic and nuclear protein glycosylation by O-linked *N*-acetylglucosamine (O-GlcNAc) (e.g., *Journal of Biological Chemistry* 259: 3308; *Journal of Biological Chemistry* 261: 8049). Since that time, the Hart laboratory has published nearly 200 papers on O-GlcNAcylation, identifying and cloning the enzymes controlling cycling, characterizing O-GlcNAcylation and its interplay with phosphorylation on hundreds of proteins, and they have developed many of the tools and methods in use today to study this modification. In

1989, Hart founded the leading journal in the field, *Glycobiology*, serving as Editor-in-Chief until 2001. Hart received the first International Glycoconjugate Organization (IGO) Award in 1997, the Karl Meyer Award from the Society for Glycobiology in 2006, and served as the 2009–2011 president of the IGO. Hart is currently an Associate Editor for *The Journal of Biological Chemistry* and an Associate Editor for *Molecular and Cellular Proteomics*. To date, Hart has published about 263 papers, all in the area of glycosciences. H factor=84.



Bruno Goud studied biochemistry and immunology at the Ecole Normale Supérieure de Cachan and the University of Paris. He received his PhD in 1981 under the mentorship of Jean-Claude Antoine and Stratis Avrameas at Institut Pasteur and did postdoctoral work under the mentorship of Peter Novick at Yale University. In 1995, he established an independent research group in the Department of Cell Biology at the Institut Curie in Paris. The main focus of his studies is the regulation of intracellular transport and membrane trafficking in eukaryotic cells. In particular, his group is working on the functions of Rab GTPases, the mechanisms that sustain the global organization of intracellular compartments and the functions of myosins in membrane traffic. Since 2000, he has also developed original *in vitro* approaches to unravel physical parameters such as membrane tension or membrane curvature that underlie transport processes.

Bruno Goud is a recipient of an ERC (European Research Council) Advanced grant (2013). He is a member of the European Molecular Biology Organization (EMBO). He has been the Head of the Department of Cell Biology of Institut Curie since 2003.



Graça Raposo received her PhD in 1989 in Membrane Biology and Immunology at the University of Paris VII where she specialized in electron microscopy and membrane biology. From 1990 to 1995 she was a postdoctoral fellow at the Immunology Center in Marseille and then in the Department of Cell Biology, Utrecht University, The Netherlands. She is the deputy Director of the Department of Cell Biology at Institut Curie and Director of the Training unit. Her major research interests focus on the biogenesis and functions of exosomes and lysosome related organelles with implications in neurodegenerative disorders, lysosomal diseases, and cancer. Her group have started to unravel the cellular and molecular mechanisms regulating the biogenesis of melanosomes, the lysosome related organelles of epidermal melanocytes, studies that open a new avenue to modulate pigmentation in health and disease. In 2012 she was awarded the CNRS Silver Medal and in 2013 the Descartes Huygens Price from the Royal Dutch Academy of Sciences. She is a member of the European Molecular Biology Organization (EMBO).



Alan Ezekowitz, MBChB, DPhil, FAAP, is Co-Founder, President, and CEO of Abide Therapeutics, a company cofounded by Professors Dale Boger and Ben Cravatt from the Scripps Research Institute. He was previously Senior Vice President and Franchise Head, Bone, Respiratory, Immunology, Endocrine, Dermatology and Urology at Merck Research Laboratories. At Merck, he was responsible for the overall scientific direction of the drug discovery and development process. Prior to Merck, Dr. Ezekowitz was Chief of Pediatric Services at the Massachusetts General Hospital for Children and the Partners Healthcare System and the Charles Wilder Professor of Pediatrics at the Harvard Medical School. He chaired the committee that led to the establishment of the Academy at the Harvard Medical School where he served as a Scholar and Founding member. He served on the Board of Directors of the Partners Healthcare System and Massachusetts General Physicians Organization. In 2008 he was honored with the establishment R. Alan Ezekowitz Professorship in Pediatrics at the Harvard Medical School.

He is a member of the American Society of Clinical Investigation, the Association of American Physicians, the American Pediatric Society, and a Fellow of AAAS. He served on NIH Subcommittees on Biodefense and Vaccine Adjuvants; NAS panels on antibiotic resistance and academic pharmaceutical partnerships; and the ARISE 2 task force of American Academy of Arts & Science that is evaluating the impact of federal and industrial funding of science, engineering, and medicine on American universities.

Dr. Ezekowitz received his MBChB and DPhil from University of Cape Town and Oxford University, respectively. At Children's Hospital in Boston, he was a postdoctoral fellow in the Division of Hematology and Oncology with clinical training in pediatrics. He was on the staff of the Children's Hospital of Boston prior to moving to the Massachusetts General Hospital.

He is a pioneer in the field of innate immunity and has over 150 publications. In particular his group played a major role in defining the structure function of the mannose binding lectin and the macrophage mannose receptor.



Douglas A. Lauffenburger is Ford Professor of Bioengineering, and cofounder and Head of the Department of Biological Engineering at MIT, with affiliate appointment in the Department of Biology; he was a founding codirector of the MIT Computational and Systems Biology Initiative in 2002. His major research interests are in cell engineering, with central focus on cell-cell communication and intracellular signal transduction, emphasizing predictive computational models derived from quantitative experiment. Lauffenburger has coauthored a monograph entitled *Receptors: Models for Binding, Trafficking & Signaling* (Oxford University Press, 1993), has coedited the book entitled *Systems Biomedicine: Concepts and Perspectives* (Elsevier, 2010), and has supervised more than 100 doctoral and post-doctoral students. He is a member of the National Academy of Engineering and the American Academy of Arts and Sciences, and has served as President of the Biomedical Engineering Society, Chair of the AIMBE College of Fellows, and on the NIH NIGMS Advisory Council, and coauthored the NRC report on *A New Biology for the 21st Century*.

SECTION EDITORS



Kenneth E Neet received his PhD in Biochemistry in 1965 (with Dr. Frank W. Putnam) from the University of Florida. He was a postdoctoral fellow at the University of California, Berkeley (with Dr. Daniel E. Koshland, Jr.), and joined the faculty of Case Western Reserve University (CWRU) in 1967 as an Assistant Professor of Biochemistry.

Dr. Neet received a Faculty Research Award of the American Cancer Society from 1968 to 1973 and was on sabbatical leave at the National Institute for Medical Research, Mill Hill, England (with Dr. N. Michael Green) 1973–1974. From 1978 to 1990, he was Professor of Biochemistry at CWRU. During 1980–1981, he took a sabbatical year as a Josiah Macy Faculty Scholar in the Department of Neurobiology, Stanford University (with Dr. Eric M. Shooter).

Dr. Neet moved to Rosalind Franklin University of Medicine and Science/The Chicago Medical School in 1990 and was Professor and Chair of Biochemistry & Molecular Biology until 2005. Dr. Neet became Associate Dean for Research of the Chicago Medical School, RFUMS in 2004.

Dr. Neet has served on the Editorial Boards of the *Journal of Biological Chemistry*, *Protein Science*, and *Molecular & Cellular Proteomics* and as a member of study sections for the National Institutes of Health and the National Science Foundation. He was an Associate Editor of the *Journal of Biological Chemistry* (1996–2013).

Dr. Neet's research interests have been in the general areas of protein–protein interactions, allosteric interactions, protein conformational changes, and cell signaling. In his earlier years he studied theoretical and experimental aspects of slow transitions in enzymes, particularly the cooperativity of the monomeric enzyme glucokinase. Later, he studied ligand–receptor interactions of nerve growth factor, establishing a mutational system to study the interactions with its receptors (TrkA and p75), and ultimately the complex signaling emanating from these receptors within neuronal systems.

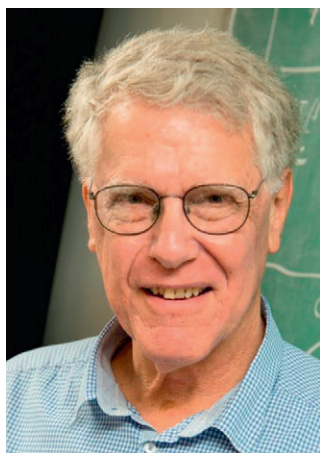


Sarah A Woodson is a biophysical chemist interested in how RNA molecules fold into specific three-dimensional structures and how they interact with proteins to turn genes on and off in the cell. She was born in Michigan, USA, and attended Kalamazoo College before receiving her PhD in Biophysical Chemistry from Yale University in 1987. After postdoctoral research in the laboratory of Tom Cech at the University of Colorado Boulder from 1987 to 1990, she joined the faculty of the University of Maryland in 1990 and the faculty of Johns Hopkins University in 1999, where she is currently the T.C. Jenkins Professor of Biophysics. Together with Mark Chance and Michael Brenowitz, she pioneered methods for visualizing how RNA molecules change shape in real time. She served on the board of the RNA Society and was elected an AAAS Fellow in 2011 and a Pew Scholar in the Biomedical Sciences in 1993. She is currently a reviewing editor of *Biopolymers* and *Journal of Molecular Biology* and serves on the editorial boards of *Nucleic Acids Research* and *RNA*.



Judith Bond, Evan Pugh Emeritus Professor of Biochemistry and Molecular Biology at the Pennsylvania State University College of Medicine, was President of the Federation of American Societies for Experimental Biology (2012–13) and an Associate Editor of the *Journal of Biological Chemistry* (1999–2013). She received her BS degree from Bennington College in Vermont, her PhD from Rutgers University, and did postdoctoral research at Vanderbilt University. She rose through the academic ranks at the Medical College of Virginia, Virginia Commonwealth University, became Professor and Head of Biochemistry and Nutrition at Virginia Polytechnic Institute and State University, and served at Penn State University College of Medicine as Professor and Chair of Biochemistry and Molecular Biology from 1992 to 2012. At Penn State, she also served as assistant dean for graduate studies, codirector of the All-University Interdisciplinary Biological Sciences Program and founding director of the Medical Scientist Training Program. She recently directed NIH-funded summer research programs for high school students and teachers and for undergraduate students (particularly underrepresented minority groups). Her work on metalloproteases has been funded continuously by the NIH for over 35 years. She has

trained 5 Master's, 21 PhD students, and 19 postdoctoral trainees. Her professional service included member and chair National Institutes of Health Study Sections, member of the National Institute of Diabetes, Digestive and Kidney Diseases Advisory Council, member of the American Association of Medical Colleges – Howard Hughes Medical Institute Committee on the Scientific Foundations for Future Physicians, and President of the American Society for Biochemistry and Molecular Biology.



Elliot Elson received his AB at Harvard, his PhD at Stanford under the mentorship of R.L. Baldwin and did postdoctoral work at University of California, San Diego, under Bruno Zimm. His first independent faculty position was in the Department of Chemistry, Cornell University, where he stayed for 11 years. He then moved to the Department of Biological Chemistry (now the Department of Biochemistry and Molecular Biophysics) at Washington University in St. Louis, School of Medicine. He has pursued research in several biophysical areas. One of these is the development of Fluorescence Correlation Spectroscopy and Fluorescence Photobleaching Recovery and their application to studies of diffusion in cell membranes and of phase separation in model membrane bilayers as well as of other cellular and noncellular phenomena. He has also worked in the areas of cell mechanics, cell motility, and the mechanical properties of engineered tissues. His laboratory has developed approaches for measuring mechanical properties of cells in monolayer culture and for determining the contributions of cells and extracellular matrix to the mechanics of engineered heart and connective tissue constructs.



Tamotsu Yoshimori received his PhD degree in 1989 at Osaka University. After working at several places including European Molecular Biology Laboratory (Prof. Kai Simons' lab) and National Institute of Basic Biology (Prof. Yoshinori Ohsumi's lab), he is now a distinguished professor of Osaka University (Graduate School of Medicine and of Frontier Biosciences). His research interests are focused on intracellular membrane trafficking, and especially for the last 18 years, on autophagy. He identified LC3 as an autophagosome-binding protein, which has been widely used as the gold standard in autophagy assays. The paper has been cited over 3000 times. He also provided new insights into membrane biogenesis in autophagy and the role of autophagy in pathogen defense and suppression of various diseases. He authored or coauthored over 220 journal articles and book chapters. He is an editor of *Journal Cell Science*, and on the editorial board of *Journal of Cell Biology*, *Molecular Biology of the Cell*, and so on. He is a member of the Faculty of 1000 and a vice president of Japan Society for Cell Biology. He was awarded Osaka University Presidential Award for 2012 and 2013, Prize for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology in 2013, Kakiuchi Saburo Memorial Prize by the Japanese Biochemical Society in 2014, and selected as Highly Cited Researchers 2014 by Thomson Reuters.



Paul A Gleeson is currently Head of the Department of Biochemistry and Molecular Biology at the University of Melbourne, located at Bio21 Institute. His research interests include the molecular mechanisms of intracellular membrane transport and the molecular basis of organ-specific autoimmune diseases. Paul Gleeson obtained his PhD in 1980 from the University of Melbourne and did postdoctoral research in the biosynthesis and function of glycoproteins at the Hospital for Sick Children, Toronto, National Institute for Medical research, Mill Hill London and Department of Biochemistry, La Trobe University, Melbourne. He established an independent laboratory at Monash University in 1986 where he defined the targeting signals of Golgi glycosyltransferases, identified golgins of the *trans*-Golgi network and along with his colleagues developed mouse models of autoimmune gastritis. In 2001 he moved to Department of Biochemistry and Molecular Biology at the University of Melbourne and has been Head of the Department since 2006. He has a number of international research collaborations and has been a visiting scientist at the EMBL, Heidelberg, and the Institut Curie, Paris.

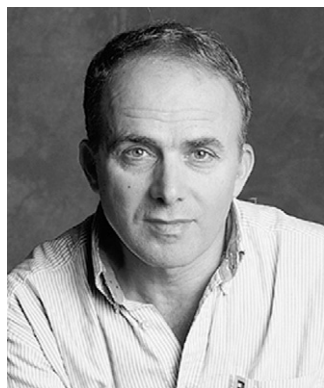


Anna Akhmanova studied biochemistry and molecular biology at the Moscow State University. She received her PhD in 1997 at the University of Nijmegen, the Netherlands. She worked as a postdoc at the Department of Microbiology and Evolutionary Biology at the University of Nijmegen and at the Department of Cell Biology at the Erasmus Medical Center in Rotterdam, the Netherlands. In 2001, she started her own research group at the Erasmus Medical Center. Since 2011, Anna Akhmanova is professor of Cell Biology at Utrecht University, the Netherlands.

Akhmanova studies cytoskeletal organization and trafficking processes, which contribute to cell polarization, differentiation, vertebrate development, and human disease. The main focus of her studies is the microtubule cytoskeleton. Her group has described key interactions between proteins associated with microtubule ends and determined their role in the regulation of microtubule organization and dynamics. She has also characterized mechanisms of bidirectional microtubule-based motility of membrane organelles such as cell nuclei and exocytotic vesicles. Her research relies on combining high-resolution live cell imaging and quantitative analysis of cytoskeletal remodeling, measurement of protein dy-

namics using advanced microscopic assays, *in vitro* reconstitution of dynamic cytoskeleton-based processes and different methods of identification of protein–protein interactions.

Anna Akhmanova is a recipient of the ALW Vernieuwingsimpuls VIDI (2001) and VICI awards (2007) (Netherlands Organization for Scientific Research), and an ERC Synergy grant (2013). She is the Chair of the board of the Netherlands Microscopy Society and a member of the European Molecular Biology Organization (EMBO).



Yosef (Yossi) Yarden is a Professor in the Department of Biological Regulation at Weizmann Institute of Science (Rehovot, Israel). Dr. Yarden obtained his PhD from Weizmann Institute (under the mentoring of Dr. Joseph Schlessinger) and later trained at both Genentech Inc., with Dr. Axel Ullrich, and at the Whitehead Institute (MIT), with Dr. Robert A. Weinberg. His group studies the roles played by growth factors and their receptors in cancer progression. They also explore strategies able to intercept growth factor signals in tumors.



Jason D Weber is an Associate Professor in the Departments of Internal Medicine and Cell Biology and Physiology at Washington University and is the Coleader of the Breast Cancer Research Program in the Siteman Comprehensive Cancer Center. Dr. Weber obtained his PhD from Saint Louis University and received postdoctoral training under the mentoring of Dr. Charles J. Sherr at St. Jude Children's Research Hospital in Memphis, TN. His group studies the molecular interplay between oncogenes and tumor suppressors, particularly focused on their role in regulating cellular growth processes.



Michael Dustin received his PhD in 1990 in Cell and Developmental Biology from Harvard University, where he worked in the laboratory of Timothy A. Springer, PhD. During his graduate work he was involved in the identification and cloning of ICAM-1 and ICAM-2, key ligands for the integrin LFA-1. He further demonstrated that LFA-1 dependent adhesion was stimulated by T cell receptor signaling. He did his postdoctoral training with Stuart Kornfeld at Washington University School of Medicine in St. Louis, focusing on mechanisms of lysosome biogenesis. He started his independent lab also at Washington University School of Medicine, and used supported lipid bilayers to define the first protein to modulate organization of the immunological synapse- CD2 associated protein- and the dynamics of immunological synapse formation. He moved to NYU School of Medicine in 2001 where he applied *in vivo* microscopy to tolerance induction and immune responses in effector sites including the liver, brain, and spleen. Continued work with the immunological synapse model resulted in discovery of signaling microclusters, the molecular basis of immunological synapse stability, and the direct budding of extracellular microvesicle enriched in T cell receptors in the immunological synapse. Professor Dustin recently took a position at the

University of Oxford with support of a Wellcome Trust Principal Research Fellowship. The focus of his new post is the translation of the immunological synapse. He received a Presidential Early Career Award in Science and Engineering and the DART-NYU Biotechnology Achievement Award.



David Sassoon directs a research team at Paris Sorbonne. He received his PhD from Columbia University (NYC, USA) in 1986 in the biological sciences and was a professor at Boston University Medical School and Mt. Sinai Medical School before relocating to Paris in 2006. His long-standing interests are in developmental and stem cell biology in a number of tissues including skeletal muscle, skin, CNS, and heart. Dr. Sassoon recently concluded the coordination of a EC-funded 15 partner international consortium (Endostem) designed to identify novel therapeutics for mobilizing endogenous stem/progenitor cells (<http://www.endostem.eu/>) and is a recent recipient as coordinator of a multipartner Transatlantic Network of Excellence grant from the Fondation Leducq focused on cardiac stem cell biology (<http://www.fondationleducq.org/nivel2.aspx?idsec=1360>).

A major focus of his research is on adult progenitor/stem cell biology and how these cells respond to stress. His team has identified a parentally imprinted gene, PW1/Peg3, which is involved in both p53 and inflammatory responses. PW1 is expressed in adult stem cells in all tissues identified to date. Loss of PW1 function in stem cells results in a loss of stem cell competence including a reduced capacity to undergo self-renewal and respond to hypoxic stress. We are currently evaluating a role for PW1 in postnatal and adult heart with a particular focus on its role in vessel precursors using both myocardial infarction and stress-induced myocardial hypertrophy.



Jason M Haugh is a Professor of Chemical and Biomolecular Engineering at North Carolina State University in Raleigh, NC, USA, where he has been on the faculty since 2000. His laboratory has been among those to pioneer the synthesis of quantitative experiments and modeling to study signal transduction in mammalian cells. Since the lab's inception, their approach has combined biochemical measurements, live-cell fluorescence microscopy, and computational modeling to elucidate signaling mechanisms by analyzing their kinetics and spatial organization in cells. The systems studied by the Haugh Laboratory include: regulation of the phosphoinositide 3-kinase, Ras/extracellular signal-regulated kinase, and phospholipase C pathways mediated by receptor tyrosine kinases, and cross talk between these pathways; dynamic organization of multimolecular complexes at cell membranes; signaling mediated by cytokine and chemokine receptors in immune cells; and integration of adhesion, signaling, and cytoskeletal dynamics that direct cell migration.



After graduating with an MA in Mathematics from the University of Cambridge, Helen Mary Byrne received her Master of Science and DPhil from the University of Oxford. She worked as a postdoctoral researcher at the Universities of Oxford and Bath, before taking up a lectureship at the UMIST in Manchester. She moved to the University of Nottingham in 1998 and was awarded a prestigious Advanced Research Fellowship from the United Kingdom's Engineering and Physical Sciences Research Council (2000–2006). She was promoted to Reader in 2002 and Professor in 2003. While in Nottingham, she established and then led the Centre for Mathematical Medicine and Biology until she returned to Oxford in 2011 where she is now based in the Mathematical Institute at the University of Oxford. She has over 20 years' experience of developing, analyzing, and simulating continuum and multiscale models of biomedical systems, and a particular interest in studying the growth and response to treatment of solid tumors.



Rune Linding completed his PhD at the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany, followed by postdoctoral training at EMBL. He then jointly trained with professors Tony Pawson and Mike Yaffe at the Lunenfeld at Mount Sinai Hospital in Toronto, Canada, and the Massachusetts Institute of Technology (MIT) in Cambridge, US, respectively. Dr. Linding then established his own laboratory of Cellular and Molecular Logic at the Institute of Cancer Research (ICR) in London, UK, before returning to Denmark to take a position as professor of cellular signal integration at the Technical University of Denmark. In 2014, Dr. Linding moved his laboratory to the Biotech Research and Innovation Centre (BRIC) at University of Copenhagen where he is currently professor of cellular signaling. His research group focuses on big data network biology, exploring biological systems by developing and deploying algorithms aimed to predict cell behavior, in particular looking at cellular signal processing and decision making. A strategic focus is to continue to develop computational tools (such as KinomeExplorer, NetworKIN, and NetPhorest) and to deploy these on genome-scale quantitative data obtained by, for example, mass spectrometry, genomic, and phenotypic screens to understand the principles of how spatio and temporal assembly of mammalian signaling networks transmit and process information at a systems level in order to alter cell behavior. His overarching aim is to advance network

medicine by identifying and targeting signaling networks associated with complex diseases. To this end Dr. Linding is currently leading high-level, strategic, multidisciplinary studies of signaling network dynamics driving cancer metastasis in collaboration with other labs at Harvard, Yale, The Jackson Laboratory, Memorial Sloan Kettering Cancer Center, MIT, and BRIC.

CONTRIBUTORS TO VOLUME 1

TM Antalis
University of Maryland School of Medicine, Baltimore, MD, USA

B Antonny
Université de Nice Sophia Antipolis et CNRS, Valbonne, France

SS Apte
Cleveland Clinic, Cleveland, OH, USA

RS Aquino
Harvard Medical School, Boston, MA, USA

BJ Bailus
The Buck Institute for Research on Aging, Novato, CA, USA

SL Beaucage
US Food and Drug Administration, Silver Spring, MD, USA

C Becker-Pauly
University of Kiel, Kiel, Germany

M Blaber
Florida State University, Tallahassee, FL, USA

CP Blobel
Hospital for Special Surgery, New York, NY, USA

JS Bond
Penn State University College of Medicine, Hershey, PA, USA

P Boon Chock
National Institutes of Health, Bethesda, MD, USA

NA Bradbury
Chicago Medical School, North Chicago, IL, USA

SG Bradley
Virginia Commonwealth University, Richmond, VA, USA

RA Bradshaw
University of California, San Francisco, CA, USA

K Brew
Florida Atlantic University, Boca Raton, FL, USA

JL Brodsky
University of Pittsburgh, Pittsburgh, PA, USA

TP Burris
Saint Louis University School of Medicine, St. Louis, MO, USA

MS Buzza
University of Maryland School of Medicine, Baltimore, MD, USA

JJ Caramelo
Fundación Instituto Leloir, Buenos Aires, Argentina

JC Cardenas
University of North Carolina School of Medicine, Chapel Hill, NC, USA and University of Texas Health Science Center at Houston, Houston, TX, USA

TJ Carrocci
University of Wisconsin–Madison, Madison, WI, USA

M Cascio
Duquesne University, Pittsburgh, PA, USA

AM Case
Abbott Molecular Inc., Des Plaines, IL, USA

S Chakraborty
University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

Z Chang
Peking University, Beijing, China

JJ-L Chen
Arizona State University, Tempe, AZ, USA

FC Church
University of North Carolina School of Medicine, Chapel Hill, NC, USA

NE Clarke
University of Leeds, Leeds, UK

WA Cramer
Purdue University, West Lafayette, IN, USA

M Datto
Duke University Medical Center, Durham, NC, USA

J Doherty
The Scripps Research Institute, Jupiter, FL, USA

RJ Dohmen
University of Cologne, Köln, Germany

AV Domashevskiy
John Jay College of Criminal Justice, New York, NY, USA

W Dowhan
University of Texas Medical School at Houston, Houston, TX, USA

M Dunder
*Chicago Medical School, Rosalind Franklin University of
Medicine and Science, North Chicago, IL, USA*

MJ Edmonds
University of Liverpool, Liverpool, UK

LM Ellerby
*The Buck Institute for Research on Aging, Novato, CA,
USA*

AR Ferré-D'Amaré
*National Heart, Lung and Blood Institute, Bethesda,
MD, USA*

DI Fisher
AstraZeneca, Cambridge, UK

PA Frey
*The University of Wisconsin–Madison, Madison, WI,
USA*

R Fukunaga
*Johns Hopkins University School of Medicine, Baltimore,
MD, USA*

DJ Goss
Hunter College CUNY, New York, NY, USA

GA Grant
*Washington University School of Medicine, St. Louis,
MO, USA*

M Guo
The Scripps Research Institute, Jupiter, FL, USA

M Hafner Česen
Jožef Stefan Institute, Ljubljana, Slovenia

RA Harris
*Indiana University School of Medicine, Indianapolis,
IN, USA*

GW Hart
*Johns Hopkins University School of Medicine, Baltimore,
MD, USA*

VC Hascall
Cleveland Clinic, Cleveland, OH, USA

ML Hastings
*Chicago Medical School, Rosalind Franklin University of
Medicine and Science, North Chicago, IL, USA*

K Hayashida
Harvard Medical School, Boston, MA, USA

W-D Heyer
University of California, Davis, CA, USA

DM Hinton
National Institutes of Health, Bethesda, MD, USA

AA Hoskins
University of Wisconsin–Madison, Madison, WI, USA

W Hu
*Whitehead Institute for Biomedical Research,
Cambridge, MA, USA*

PJ Hudson
Avipep Pty Ltd., Parkville, VIC, Australia

JM Huibregtse
The University of Texas at Austin, Austin, TX, USA

HV Jain
*US Food and Drug Administration, Silver Spring, MD,
USA*

M Jaskolski
*Adam Mickiewicz University, Poznan, Poland and
Polish Academy of Sciences, Poznan, Poland*

SS Jenkins
University of California, Davis, CA, USA

JM Jez
*Washington University in St. Louis, St. Louis, MO,
USA*

K Ji
*Wayne State University School of Medicine, Detroit,
MI, USA*

A Jinno
Harvard Medical School, Boston, MA, USA

S Joseph
University of California, San Diego, La Jolla, CA, USA

BM Kenwood
Georgia Institute of Technology, Atlanta, GA, USA

YH Ko
University of Maryland BioPark, Baltimore, MD, USA

AA Komar
Cleveland State University, Cleveland, OH, USA

M Kong
University of Pittsburgh, Pittsburgh, PA, USA

DA Korasick
*Washington University in St. Louis, St. Louis, MO,
USA*

SE Kotermanski
Duquesne University, Pittsburgh, PA, USA

O Kulaeva
Fox Chase Cancer Center, Philadelphia, PA, USA

MWL Lau
*National Heart, Lung and Blood Institute, Bethesda,
MD, USA*

BA Lewis
*Center for Cancer Research, National Cancer Institute,
 National Institutes of Health, Bethesda, MD, USA*

I Lindberg
University of Maryland-Baltimore, MD, USA

M Lizotte-Waniewski
Florida Atlantic University, Boca Raton, FL, USA

E London
Stony Brook University, Stony Brook, NY, USA

LM Longo
Florida State University, Tallahassee, FL, USA

RL Lundblad
Chapel Hill, NC, USA

Y Luo
National Institute of Infectious Diseases, Tokyo, Japan

G Marius Clore
National Institutes of Health, Bethesda, MD, USA

LM Mayr
AstraZeneca, Cheshire, UK

AA McCracken
University of Nevada, Reno, NV, USA

WC Merrick
Case Western Reserve University, Cleveland, OH, USA

AH Merrill Jr.
Georgia Institute of Technology, Atlanta, GA, USA

AJ Moorhouse
UNSW Australia, Sydney, NSW, Australia

Amitava Mukherjee
*University of Pittsburgh School of Medicine, Pittsburgh,
 PA, USA*

S Mukherjee
University of California, Davis, CA, USA

G Murphy
University of Cambridge, Cambridge, UK

KE Neet
*Rosalind Franklin University of Medicine and Science,
 North Chicago, IL, USA*

IP O'Carroll
*National Cancer Institute at Frederick, Frederick, MD,
 USA*

ME O'Donnell
The Rockefeller University, New York, NY, USA

K Okamoto
Nagasaki University, Nagasaki, Japan

H Oubrahim
National Institutes of Health, Bethesda, MD, USA

CM Overall
University of British Columbia, Vancouver, BC, Canada

PW Park
Harvard Medical School, Boston, MA, USA

JL Parsons
University of Liverpool, Liverpool, UK

MS Patel
*University at Buffalo, The State University of New York,
 Buffalo, NY, USA*

PL Pedersen
*Johns Hopkins University School of Medicine, Baltimore,
 MD, USA*

JR Peinado
University of Castilla-La Mancha, Ciudad Real, Spain

DH Perlmutter
*University of Pittsburgh School of Medicine, Pittsburgh,
 PA, USA and Children's Hospital of Pittsburgh of
 UPMC, Pittsburgh, PA, USA*

RN Pike
Monash University, Clayton, VIC, Australia

JD Podlevsky
Arizona State University, Tempe, AZ, USA

DM Raben
*Johns Hopkins School of Medicine, Baltimore, MD,
 USA*

SC Raghavan
Indian Institute of Science, Bangalore, India

B Ramos-Molina
University of Maryland-Baltimore, MD, USA

GH Reed
*The University of Wisconsin-Madison, Madison, WI,
 USA*

A Rein
*National Cancer Institute at Frederick, Frederick, MD,
 USA*

CM Rein-Smith
*University of North Carolina School of Medicine,
 Chapel Hill, NC, USA and Whitsell Innovations Inc.,
 Chapel Hill, NC, USA*

PJ Roach
*Indiana University School of Medicine, Indianapolis,
 IN, USA*

RG Roth
AstraZeneca, Mölndal, Sweden

- BM Ryan
Laboratory of Human Carcinogenesis Center for Cancer Research, Bethesda, MD, USA
- IA Sawyer
Chicago Medical School, Rosalind Franklin University of Medicine and Science, North Chicago, IL, USA
- M Scheffner
University of Konstanz, Konstanz, Germany
- DF Sellitti
Uniformed Services University of the Health Sciences, Bethesda, MD, USA
- S Sharma
Indian Institute of Science, Bangalore, India and FIRC Institute of Molecular Oncology, Milan, Italy
- SK Singh
Purdue University, West Lafayette, IN, USA
- S Sitaula
The Scripps Research Institute, Jupiter, FL, USA
- BF Sloane
Wayne State University School of Medicine, Detroit, MI, USA
- N Solis
University of British Columbia, Vancouver, BC, Canada
- LL Spremulli
University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
- SA Stewart
Washington University School of Medicine, Saint Louis, MO, USA
- V Stoka
Jožef Stefan Institute, Ljubljana, Slovenia and International Postgraduate School Jozef Stefan, Ljubljana, Slovenia
- VM Studitsky
Fox Chase Cancer Center, Philadelphia, PA, USA and Russian Academy of Sciences, Moscow, Russia
- K Suzuki
National Institute of Infectious Diseases, Tokyo, Japan
- DC Teasley
Washington University School of Medicine, Saint Louis, MO, USA
- BP Toole
Medical University of South Carolina, Charleston, SC, USA
- C Tredup
University of Kiel, Kiel, Germany
- T Tsukuba
Nagasaki University, Nagasaki, Japan
- B Turk
Jožef Stefan Institute, Ljubljana, Slovenia and Centre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins, Ljubljana, Slovenia
- AJ Turner
University of Leeds, Leeds, UK
- B Tu-Sekine
Johns Hopkins School of Medicine, Baltimore, MD, USA
- B Van Houten
University of Pittsburgh, Pittsburgh, PA, USA
- MN Vu
Georgia Institute of Technology, Atlanta, GA, USA
- KW Walker
Amgen Inc., Thousand Oaks, CA, USA
- PH Weigel
University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA
- LC Wijeyewickrema
Monash University, Clayton, VIC, Australia
- A Wlodawer
National Cancer Institute, Frederick, MD, USA
- MS Wolfe
Harvard Medical School, Boston, MA, USA
- MJ Wolfgang
Johns Hopkins University School of Medicine, Baltimore, MD, USA
- SA Woodson
Johns Hopkins University School of Medicine, Baltimore, MD, USA
- K Yamamoto
Kyushu University, Fukuoka, Japan
- NY Yao
The Rockefeller University, New York, NY, USA
- LF Zacchi
Fundación Instituto Leloir, Buenos Aires, Argentina
- SC Zeeman
ETH Zürich, Zürich, Switzerland

CONTENTS OF VOLUME 1

Preface	xxv
Cell Biology: An Overview <i>RA Bradshaw and PD Stahl</i>	xxvii

Molecular Cell Biology

Molecular Cell Biology: An Overview <i>KE Neet, SA Woodson, JS Bond, and GW Hart</i>	1
--	---

Molecular Principles, Components, Technology, and Concepts

Basic Principles

Chemical and Physical Principles <i>H Oubrahim and P Boon Chock</i>	5
Biocatalysis <i>PA Frey and GH Reed</i>	14

Nucleic Acids

DNA, RNA Chemical Properties (Including Sequencing and Next-Generation Sequencing) <i>M Datto and RL Lundblad</i>	24
The Chemical Synthesis of DNA and RNA Oligonucleotides for Drug Development and Synthetic Biology Applications <i>SL Beaucage and HV Jain</i>	36

Proteins

Expression Systems <i>DI Fisher, LM Mayr, and RG Roth</i>	54
Isolation/Purification of Proteins <i>GA Grant</i>	66
Protein Sequence Determination: Methodology and Evolutionary Implications <i>RA Bradshaw</i>	75
Posttranslational Modifications: Key Players in Health and Disease <i>I Lindberg and JR Peinado</i>	84
Protein Domains: Structure, Function, and Methods <i>DA Korasick and JM Jez</i>	91
NMR in Structural and Cell Biology <i>G Marius Clore</i>	98
Folding, Misfolding, Disordered Proteins, and Related Diseases <i>LM Longo and M Blaber</i>	108
Diseases of Protein Folding: Huntington's Disease and Amyotrophic Lateral Sclerosis <i>BJ Bailus and LM Ellerby</i>	115
Site-Directed Mutagenesis <i>KW Walker</i>	122
Chemical Biology <i>RL Lundblad</i>	128
Drug Design <i>RL Lundblad</i>	135
Antibodies and Improved Engineered Formats (as Reagents) <i>PJ Hudson</i>	141

Lipids

Lipidomics <i>BM Kenwood and AH Merrill Jr.</i>	147
Synthesis and Structure of Glycerolipids <i>W Dowhan</i>	160
Cholesterol and Other Steroids <i>S Sitaula and TP Burris</i>	173

Glycolipids	<i>AH Merrill Jr. and MN Vu</i>	180
Lipid Signaling	<i>B Tu-Sekine and DM Raben</i>	194

Membranes

Composition, Physical Properties, and Curvature	<i>B Antonny</i>	201
Lipid Rafts/Membrane Rafts	<i>E London</i>	208
Membrane Potential: Concepts	<i>AJ Moorhouse</i>	218
The Outer Mitochondrial Membrane, a Smooth 'Coat' with Many Holes and Many Roles: Preparation, Protein Components, Interactions with Other Membranes, Involvement in Health, Disease, and as a Drug Target	<i>PL Pedersen and YH Ko</i>	237
Neuronal Action Potentials and Ion Channel Allostery	<i>SE Kotermanski and M Cascio</i>	244
Cystic Fibrosis	<i>NA Bradbury</i>	252

Carbohydrates

Glycogen and Starch	<i>PJ Roach and SC Zeeman</i>	263
Proteoglycans	<i>PW Park, K Hayashida, RS Aquino, and A Jinno</i>	271
Hyaluronan	<i>VC Hascall, PH Weigel, and BP Toole</i>	279

Metabolism

Metabolic Regulation	<i>MS Patel and RA Harris</i>	288
A Structure Perspective on Organelle Bioenergetics	<i>WA Cramer and SK Singh</i>	298

Nucleic Acid Synthesis/Breakdown

RNA Synthesis/Function

Transfer RNA	<i>J Doherty and M Guo</i>	309
Messenger RNA (mRNA): The Link between DNA and Protein	<i>DJ Goss and AV Domashevskiy</i>	341
The Interplay between Eukaryotic mRNA Degradation and Translation	<i>W Hu</i>	346
miRNAs/Small Noncoding RNAs	<i>R Fukunaga</i>	354
Small RNAs/Cancer	<i>BM Ryan</i>	364
Riboswitches and Ribozymes	<i>MWL Lau and AR Ferré-D'Amaré</i>	375
Ribosomal RNAs and Protein Synthesis	<i>S Joseph</i>	384
Comparison of Bacterial and Eukaryotic Replisome Components	<i>NY Yao and ME O'Donnell</i>	396

DNA Synthesis/Repair

Telomeres and Telomerase	<i>JJ-L Chen and JD Podlevsky</i>	418
Telomere Biology	<i>DC Teasley and SA Stewart</i>	426
Eukaryotic Nucleotide Excision Repair	<i>B Van Houten and M Kong</i>	435
The Base Excision Repair Pathway	<i>JL Parsons and MJ Edmonds</i>	442

Nonhomologous DNA End Joining	<i>S Sharma and SC Raghavan</i>	451
DNA Repair by Homologous Recombination	<i>SS Jenkins, S Mukherjee, and W-D Heyer</i>	456
Transcription		
Prokaryotic Transcription	<i>DM Hinton</i>	468
Eukaryotic Transcriptional Regulation	<i>BA Lewis</i>	481
Distant Activation of Transcription by Enhancers	<i>O Kulaeva and VM Studitsky</i>	488
The Spliceosome and Pre-mRNA Splicing	<i>TJ Carrocci and AA Hoskins</i>	495
Pre-mRNA Splicing: Function and Dysfunction	<i>AM Case, IA Sawyer, M Dundr, and ML Hastings</i>	503
Nucleic Acid Technology		
Transgenesis and Gene Replacement	<i>MJ Wolfgang</i>	512
Viral Nucleic Acids	<i>IP O'Carroll and A Rein</i>	517
Protein Synthesis/Degradation		
Translation		
Components, Initiation, Elongation, Termination, and Regulation	<i>AA Komar and WC Merrick</i>	525
Biogenesis of Secretory Proteins	<i>Z Chang</i>	535
The Protein Biosynthetic Machinery of Mitochondria	<i>LL Spremulli</i>	545
Regulated Proteolysis of Signaling Molecules: The Proprotein Convertases	<i>B Ramos-Molina, I Lindberg, and JR Peinado</i>	555
Protein Degradation – General		
Mass Spectrometry-based Methodologies for Studying Proteolytic Networks and the Degradome	<i>N Solis and CM Overall</i>	568
Protein Degradation – Intracellular		
Ubiquitin, Ubiquitin-Like Proteins, and Proteasome-Mediated Degradation	<i>RJ Dohmen, JM Huibregtse, and M Scheffner</i>	582
Endoplasmic Reticulum-Associated Degradation and Protein Quality Control	<i>LF Zacchi, JJ Caramelo, AA McCracken, and JL Brodsky</i>	596
Role of Lysosomes in Intracellular Degradation	<i>M Hafner Česen, V Stoka, and B Turk</i>	612
Protein Degradation – Protease Classes		
Matrix Metalloproteinases	<i>G Murphy</i>	621
ADAMTS Proteases: Mediators of Physiological and Pathogenic Extracellular Proteolysis	<i>SS Apte</i>	630
ADAMs Regulate Cell–Cell Interactions by Controlling the Function of the EGF-Receptor, TNF α and Notch	<i>CP Blobel</i>	639
Extracellular: Plasma Membrane Proteases – Serine Proteases	<i>TM Antalis and MS Buzza</i>	650
Aspartic Proteases of Alzheimer's Disease: β - and γ -Secretases	<i>MS Wolfe</i>	661

The Calpain Proteolytic System	<i>Y Luo, DF Sellitti, and K Suzuki</i>	670
Cathepsin E: An Aspartic Protease with Diverse Functions and Biomedical Implications	<i>K Yamamoto, K Okamoto, and T Tsukuba</i>	681
Metalloproteases Meprin α and Meprin β in Health and Disease	<i>C Tredup and C Becker-Pauly</i>	691
Kallikrein	<i>RA Bradshaw and RL Lundblad</i>	699
Naturally-Occurring Polypeptide Inhibitors: Cystatins/Stefins, Inhibitors of Apoptosis (IAPs), Serpins, and Tissue Inhibitors of Metalloproteinases (TIMPs)	<i>K Brew and M Lizotte-Waniewski</i>	706
Proteolytic Pathways		
Overview of Blood Coagulation and the Pathophysiology of Blood Coagulation Disorders	<i>JC Cardenas, CM Rein-Smith, and FC Church</i>	714
Molecular Mechanisms Underlying the Actions of the Complement System	<i>RN Pike and LC Wijeyewickrema</i>	723
Digestive Proteases: Roles in the Human Alimentary Tract	<i>SG Bradley</i>	732
Protein Degradation – Pathological Aspects		
Inhibitors of HIV Protease	<i>A Wlodawer and M Jaskolski</i>	738
Blood Pressure, Proteases and Inhibitors	<i>NE Clarke and AJ Turner</i>	746
Cancer – Proteases in the Progression and Metastasis	<i>K Ji and BF Sloane</i>	753
Lysosomal Diseases	<i>TM Cox</i>	763
Alpha-1-Antitrypsin Deficiency: A Misfolded Secretory Glycoprotein Damages the Liver by Proteotoxicity and Its Reduced Secretion Predisposes to Emphysematous Lung Disease Because of Protease-Inhibitor Imbalance	<i>S Chakraborty, DH Perlmutter, and Amitava Mukherjee</i>	789

PREFACE

Cell biology is the study of cells, the integral unit that is the basis of all living tissues. In its earliest phases, this area of investigation was largely defined by microscopic-based observations of structure and organization that were, by their nature, largely descriptive. This also served to delineate cell biology from its nearest neighbor, biochemistry, which, dating from its post WWII renaissance, was driven in large part by reductionist approaches, where tissues, cells, and organelles were broken down in order to isolate and then characterize individual components. The melding force that bridged these two areas was molecular biology that allowed the cell biologist to transform pictures into functions and allowed the biochemist to reassemble their components into ensembles and subcellular entities. Not surprisingly the boundary between these two key biological sciences has become blurred beyond recognition. Today a considerable part of what passes as cell biology is by any other name biochemistry (or molecular biology – the distinction between these two areas being equally indistinct) and vice versa. Indeed, one of the first challenges we faced in organizing this Encyclopedia was how much molecular detail to include. It was our conclusion that for the sake of completeness, and to appropriately introduce the volumes Organizational Cell Biology and Functional Cell Biology, we needed a description of the molecular components of the cell, particularly with respect to the synthesis and degradation of proteins and nucleic acids, along with outlines of important chemical principles and key methodology. We were guided to some degree in these decisions by input from undergraduate and graduate teachers and the contents of their introductory courses that they shared with us and by discussions with our volume and section editors. The result was the first volume which we recognize is very 'biochemical,' but certainly lacking the more detailed coverage of the *Encyclopedia of Biological Chemistry* (Lennarz and Lane, 2004) or the many excellent biochemical texts that dot the educational landscape. Readers who find 'gaps' or incomplete treatments in this section will likely be able to find the additional detail they seek from these sources.

The second major decision that we had to confront was whether we would attempt to cover the full range of living organisms or whether we would place some limits on the scope of the material we included. In the final result, it was decided to place the emphasis on higher eukaryotes with only very prescribed treatments of lower eukaryotes and prokaryotes. Volumes 2 and 3 thus deal with material that has perhaps been more traditionally thought of as cell biology. Again we were guided in our decisions about what to include by teaching considerations, brainstorming sessions with our editors, and a perceived need to separate cellular organization from cellular function. However, as even a quick perusal will reveal, this is not a sharp delineation either. Volume 2 begins with a methodological treatment of imaging, which is basically divided between light and electron microscopic applications, and is followed by sections on organelles, interorganellar communication, and the cytoskeleton (including motors). The

last part of volume 2 is devoted to intracellular infection and covers a variety of external agents and pathogens that impact cell function as well as human pathology. Volume 3, in contrast, largely deals with major cell functions: signaling, cell-cycle control, and apoptosis. The last two parts are devoted to cells of the immune system and their responses and stem cells – two highly specialized niches characterized by unique cellular involvement.

The last part of the Encyclopedia deals with systems biology. This is really the newest area of cell biology and considers cellular involvement as part of a phenotypic response. As such, it strives to integrate all of the levels of the first three volumes in a consistent manner to produce a seamless description of cellular function. This is clearly a newly evolving area and the most rapidly changing part of cell biology; indeed it is where we expect to see the greatest change in the next few years.

Assembling a treatment of a topic as large as cell biology had multiple challenges. Coverage was a significant problem: on the one hand it was almost impossible to avoid some redundancy in places while, on the other, there were inevitable gaps. Some of these arose from late cancellations; others from oversights on our part. We can only promise to fill these in future editions. We also note that as can be expected for a large multiauthor compilation the individual articles do differ in detail and treatment. We felt it was more important to allow our experts substantial latitude in deciding how to present their topics than to apply rigid guidelines. We did try to limit the number of references (to make it easier for people not familiar with a topic to make the transition to more extensive articles) but there is admittedly some considerable variation in the bibliographic material as well.

When we were well into the project, we were approached by Graham Nisbet (Elsevier) about including the Encyclopedia in a larger collection entitled the Reference Module in Biomedical Sciences. The Reference Module plan was grouped in single integrated work information from several other compendia and it was our view that this would greatly leverage the value of our efforts (and those of our contributors). While we were the 'new kid on the block' in that we were still compiling the Encyclopedia and the other collections to be included were already extant, we felt that it was too good an opportunity to pass up. The Reference Module has since been launched and although it does not yet contain the complete content from the Encyclopedia, this will be achieved in the not too distant future. We consider that this will be a particularly valuable resource for researchers that are just entering (or changing into) this field.

No work of this size could be completed without the help and advice of many people. In this regard we would like to thank a number of people who were instrumental in bringing this project to fruition. First and foremost we would like to recognize our volume editors: Jerry Hart, Graça Raposo, Bruno Goud, Alan Ezekowitz, and Doug Lauffenburger; and our section editors: Ken Neet, Sarah Woodson, Judy Bond,

Elliot Elson, Tamotsu Yoshimori, Paul Gleeson, Anna Akhmanova, Yosef Yarden, Jason Weber, Michael Dustin, David Sassoon, Jason Haugh, Helen Byrne, and Rune Linding; and thank them for their invaluable input, enthusiasm, and hard work. They are truly a global group and have been a great team to work with. We also wish to thank the staff at Elsevier including Janice Audet, Peter Labella, Nicky Carter, Karen Maloney, Erin Hill-Parks, Rashmi Phadnis, Blerina Osmanaj, and Ginny Mills for their many efforts, perseverance (particularly with our idiosyncrasies), and skillful management of every aspect of this project.

In the course of planning this work, we gathered information from a number of sources. At the outset the publishers polled some 167 faculty from both undergraduate and graduate schools to ascertain their perceived interest in such a project, which also garnered considerable input about potential topics. The response was highly favorable. We also talked with many colleagues about experiences at their institutions and would like to particularly thank Catherine Bevier (Colby College), Robert Simoni (Stanford University), Larry Marsh (UC Irvine), and John Cooper (Washington University) for providing syllabi of germane courses taught in their institutions. Special thanks also go to Ruedi Abersold (ETH, Zurich) and Sergio Grinstein (University of Toronto) for their expertise that they readily shared with us, which was invaluable in planning certain sections.

Finally we would like to recognize all of our authors. We have been extraordinarily fortunate in attracting individuals from all around the globe to take time from their busy schedules to prepare this splendid set of contributions. Without these efforts there would be no Encyclopedia. Under ordinary circumstances it would be appropriate to dedicate the work to them and all the colleagues whose works they wrote about. However, during the preparation of this work, we sadly lost one of the most important contributors to cell biology, Tony Pawson. Tony was a true pioneer in elucidating basic mechanisms of how cells transmit information intracellularly, a fundamental knowledge that permeates all of cell biology. In an effort that was initiated by one of the section editors, Rune Linding, it was decided to dedicate the entire work to his memory as recognition of his contributions and a lasting tribute to a man who left us too soon.

We hope that the Encyclopedia will be of value to students and researchers alike and will become a useful tool in the training of future generations of scientists.

Ralph A Bradshaw and Philip D Stahl

References

Lennarz, W.J., Lane, M.D. (Eds.), 2004. Encyclopedia of Biological Chemistry. Oxford: Elsevier.

Cell Biology: An Overview

RA Bradshaw, University of California, Irvine, CA, USA

PD Stahl, Washington University School of Medicine, St. Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

All living organisms can be divided into three principal categories: archaeobacteria, eubacteria, and eukaryotes; although they differ in structure and organization, they are all composed of cells as the fundamental life unit. At the molecular level, there is also a great deal of similarity in the basic materials that make up these entities because they use the same kinds of molecules to store and reproduce information, to run the cellular metabolism and machinery and to provide the structural framework. Thus nucleic acids, proteins, lipid membranes, and carbohydrates – alone and in various combinations – are universally present, albeit in distinguishable forms, along with innumerable metabolites and ions. There are components that are apparently essential for life and are found in one form or another in all species and there are many unique moieties and associated activities that are highly specialized and are found in relatively few organisms. Indeed, the similarities have underpinned the development of our understanding of cellular function at a rudimentary level and the differences, basically engendered by evolution, have illustrated and delineated the complexity that speciation has introduced. Perhaps the largest of these differences is that which separates single cell organisms from multicellular organisms. The latter are exclusively eukaryotes while the former are composed of both eukaryotes and prokaryotes. The cellular organization and architecture that distinguishes these two major life forms is striking; although cell biology correctly embraces both, traditionally prokaryotic organisms have been the province of the microbiologists and the majority of cell biology research has been devoted to the eukaryotic world. In practical terms this translates for the most part into the study of human cells and those of easily maintained laboratory animals and selected paradigms, for example, fruit flies, worms, and zebra fish.

Human and animal cell biology is not a tightly proscribed science with well-defined borders. Basically it serves at the interface between biochemistry, molecular biology, and genetics, on the one hand, and anatomy and physiology, on the other. The continuum of these disciplines forms the core of the biomedical sciences, which also include the related but separate fields of pharmacology, microbiology, immunology, and pathology that provide the connections to disease and health. Cell biology has strong connections to all of these. There are also specialized areas, for example, neuroscience, that are of such importance that they warrant their own category and the cell biology associated with them is also highly specialized. Thus, cell biology is as complex as the enormous variety of cells that exist and achieving an accurate description of all of them in terms of their components and functions has long been a major part of the research in this field.

Imaging and Organelle Organization

Among the developments that propelled cell biology into the modern era are the introduction of the ultracentrifuge and

the rapid advances in microscopy both electron and light microscopy – the former allowing investigators to fractionate and characterize the components of the cell and the latter to literally see them – either *in situ* or in isolated form. This progress is illustrated by a series of Nobel Prizes starting in the early 1970s that chronicle the grand confluence of classical biochemistry and general physiology that created modern cell biology (Claude *et al.*, 1974) or the discovery and characterization of intracellular organelles – lysosomes and endoplasmic reticulum (ER) among others (Mitchell, 1978), for the chemiosmotic hypothesis (Brown and Goldstein, 1985), for endocytosis (Ciechanover *et al.*, 2004), for ubiquitination and protein degradation, and just recently (Rothman *et al.*, 2013), for vesicle trafficking, and (Betzig *et al.*, 2014) for advances in light microscopy. The early cell biologists insisted on quantitative application of these new techniques and laid the groundwork for modern cell biology. As with biochemistry, reductionism has been the keystone for the amazing success over these past several decades. Each organelle has its own research history – the nucleus, mitochondria, peroxisomes, the proteasome, the ER, and cytoskeleton. The Golgi apparatus/secretory pathway and the endosome–lysosome network have all been examined in detail both in isolation and in their relationships with each other. A new member, the exosome (aka extracellular vesicle) has emerged more recently and stimulated the imagination of aspiring young investigators (Harding *et al.*, 2013). The rise of genomics, the development of spectacular imaging modalities, and evolving biochemical techniques (proteomics) have led to the discovery of molecular motors, the identification of macromolecular complexes such as the exocyst, ESCRT, GARP, SNAREs among others that choreograph complex intracellular pathways, including cell motility and cytokinesis, have brought cell biology into this new era where the whole is now seen as larger than the individual components. This creates the segue from Parts 1 and 2 of the *Encyclopedia (Molecular Cell Biology and Organizational Cell Biology)* to Part 3 (*Functional Cell Biology*) where signaling modalities will be seen as an integrative force for regulation and control.

Signaling

While all organisms can sense their environment and respond to cues from it, multicellular organisms must in addition coordinate their responses, which require intercellular communication at a sophisticated level (Bradshaw and Dennis, 2009). The higher the development, the more complex these communication systems become. Thus the cell biologist must focus not only on how molecular function is translated into cell organization and how these functions are coordinated from organelle to organelle but also on the external interactions and signals that control the larger functional responses of organs and ultimately organisms.

Intercellular communication is afforded by stimuli that can be transmitted across the cell membrane, either directly or via membrane bound molecules that in turn pass signals to the intracellular compartment. The origins of these stimuli may be quite diverse. Among other things, they can include cell–cell contacts, soluble factors/messengers, or foreign agents. Lipophilic molecules can cross the membrane by diffusion or by facilitated transport to recognize and bind to intracellular entities but most substances bind to membrane bound proteins and induce their signal by ‘activating’ them instead. These so-called receptors are capable of producing a variety of responses that invariably involve the generation of posttranslational modifications of preexisting proteins. Protein phosphorylation, which in eukaryotes mostly occurs on tyrosine, serine, and threonine residues, is highly prevalent and appears to be directly or indirectly involved in almost all transmembrane signaling (Gnad *et al.*, 2011). However, it is by no means the only vehicle for transmitting information as acetylation, methylation, monoglycosylation, and ubiquitination, to name a few, are both important and widespread. There are over three hundred different covalent modifications (Khoury *et al.*, 2011) that are introduced into proteins (and many more that have not yet been chemically defined) of both a transient and stable nature and most are likely to be involved to some extent in signal transduction processes. In addition, limited proteolysis can also be an important part of a pathway and this activity is a major player in cellular responses (Turk *et al.*, 2012).

The induction of an intracellular signal is basically perpetrated by the formation of new interactions between the modified proteins and other entities, which can range from small molecules to macromolecular protein complexes. One of the most important advances in understanding how cells process information induced from external stimuli was the appreciation that the newly formed sites produced by the protein modifications were recognized by other proteins with modules specifically designed for this purpose (Pawson, 1995). Thus, for example, phosphotyrosine residues could be recognized by other proteins containing a domain, termed SH2 for its relatedness to a similar domain found in the Src protein kinase, which could then be further modified allowing for additional interactions to occur. By these ‘docking’ events, signaling complexes could be assembled, often in multiple steps, that ultimately lead to the activation of key effectors. The end point of many of the pathways is the activation of transcription factors that lead to the modulation of gene expression of that cell by ultimately changing its protein expression profile (Bradshaw and Dennis, 2009). However, many molecules are usually modified and activated ‘along the way’ and these ‘new’ activities also add to the overall response to the original stimuli. Short term responses indeed require that the protein effectors necessary for it are already present; long term responses per force require new protein synthesis.

Signal transduction is thus dependent on two phenomena: posttranslational modifications, particularly of the readily reversible type, and protein–protein interactions. The extent to which these two activities take place, even in resting, unstimulated cells was greatly underestimated before the advent of proteomics and the introduction of mass spectrometric methodology into cell biological research (Bradshaw and

Burlingame, 2005). These high-throughput unbiased analyses of very complex mixtures, basically derived directly from cell lysates, revealed that essentially every protein had multiple interacting partners and that posttranslational modifications affected a very significant proportion of the proteins present. These were profound differences from what had been the prevailing wisdom and amounted to a paradigm shift in thinking about cellular organization and structure. As noted above, these same tools have gone on to cast new light on organelle organization in terms of resident proteins and have helped to disclose the structure of important cellular machines, such as the proteasome (Voges *et al.*, 1999), the nuclear pore (Routa *et al.*, 2000), and transcription complexes (Kornberg, 2007). They have also begun to elucidate the complexity of epigenetic regulation of gene transcription at the histone level (Cosgrove, 2012), which will also be essential in completing our understanding of signaling processes.

Concluding Remarks

Two of the most fundamental aspects of cells are their ability to reproduce themselves and, in higher organisms, to undergo changes that lead to new cell types and functions. These developmental processes leading to differentiation are what allow a single fertilized egg to form a complex adult organism and require the synthesis of all aspects of cell biology to understand. Knowing how cells go from one state to another in a timed and regulated fashion forms the core of developmental biology, which can be viewed as a sub discipline of cell biology. Also of major importance is the maintenance of viability and the associated turnover processes. The control of cell death is not only an essential part of development it is also key to managing situations that have gone awry and underlies many serious disease conditions.

It is clear that science is still a long way from understanding how even simple cells work. There is a long list of functions and structures that remain to be elucidated and integrating the various experimental approaches and the data they produce still lags far behind. This is the ‘Era of Big Data’ and vast amounts of new information that impact on our understanding of cells and their processes are collected every day. Genomic information is a good example – despite the rapidity that this sort of information can now be obtained, it still remains to be effectively mined in terms of what it can tell us about basic principles of how cells work. There has been an understandable pressure to apply this information to managing disease (Collins and Varmus, 2015), particularly those that are life threatening, but there certainly is information that is yet to be formulated that would apply to fundamental problems as well. The same can be said for transcriptomics, proteomics, and metabolomics. In addition to these powerful new technologies, singular advances in analyzing single cells promises to augment these molecular approaches significantly (Di Palma and Bodenmiller, 2015).

One aspect of cell biology that is looking to address the challenges of integrating these relatively newly minted studies is systems biology (Part 4). It is far too early to assess the value of these efforts in coordinating this flood of information but it certainly looks promising. It will be an elusive goal for some

time to come but also a stimulus for new generations of cell biologists that will be forthcoming.

References

- Betzig, E., Hell, S.W., Moerner, W.E., 2014. The Nobel Prize in chemistry — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2014/press.html (accessed 03.04.15).
- Bradshaw, R.A., Burlingame, A.L., 2005. From proteins to proteomics. *IUBMB Life* 57, 267–272.
- Bradshaw, R.A., Dennis, E.A. (Eds.), 2009. *Handbook of Cell Signaling*, second ed., vol. 1–3. San Diego, CA: Elsevier Academic Press, pp. 1–3047.
- Brown, M.S., Goldstein, J.L., 1985. Physiology or medicine — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1985/press.html (accessed 03.04.15).
- Ciechanover, A., Hershko, A., Rose, I., 2004. The Nobel Prize in chemistry — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2004/press.html (accessed 03.04.15).
- Claude, A., de Duve, C., Palade, G.E., 1974. Physiology or medicine — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1974/press.html (accessed 03.04.15).
- Collins, F.S., Varmus, H.A., 2015. New initiative on precision medicine. *New England Journal of Medicine* 372, 793–795.
- Cosgrove, M.S., 2012. Writers and readers: Deconvoluting the harmonic complexity of the histone code. *Nature Structural & Molecular Biology* 19, 739–740.
- Di Palma, S., Bodenmiller, B., 2015. Unravelling cell populations in tumors by single-cell mass cytometry. *Current Opinion in Biotechnology* 31, 122–129.
- Gnad, F., Gunawardena, J., Mann, M., 2011. PHOSIDA 2011: The posttranslational modification database. *Nucleic Acids Research* 39 (suppl. 1), D253–D260.
- Harding, C.V., Heuser, J.E., Stahl, P., 2013. Exosomes: Looking back three decades and into the future. *Journal of Cell Biology* 200, 367–371.
- Khoury, G.A., Baliban, R.C., Floudas, C.A., 2011. Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database. *Scientific Reports* 1, 90. doi:10.1038/srep00090.
- Kornberg, R.D., 2007. The molecular basis of eukaryotic transcription. *Proceedings of the National Academy of Sciences of the United States of America* 104, 12955–12961.
- Mitchell, P., 1978. The Nobel Prize in chemistry — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1978/press.html (accessed 03.04.15).
- Pawson, T., 1995. Protein modules and signaling networks. *Nature* 373, 573–580.
- Rothman, J.E., Schekman, R.W., Südhof, T.C., 2013. The 2013 Nobel Prize in physiology or medicine — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/2013/press.html (accessed 03.04.15).
- Rout, M.P., Aitchison, J.D., Supratka, A., *et al.*, 2000. The yeast nuclear pore complex — Composition, architecture, and transport mechanism. *Journal of Cell Biology* 148, 635–652.
- Türk, B., Türk, D., Türk, V., 2012. Protease signalling: The cutting edge. *EMBO Journal* 31, 1630–1643.
- Voges, D., Zwickl, P., Baumeister, W., 1999. The 26S proteasome: A molecular machine designed for controlled proteolysis. *Annual Review of Biochemistry* 68, 1015–1068.

Molecular Cell Biology: An Overview

KE Neet, Rosalind Franklin University of Medicine and Science, North Chicago, IL, USA

SA Woodson, Johns Hopkins University School of Medicine, Baltimore, MD, USA

JS Bond, Penn State University College of Medicine, Hershey, PA, USA

GW Hart, Johns Hopkins University School of Medicine, Baltimore, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Basic Molecular Components and Technology

The history of cell biology is rooted in the related sciences of biochemistry and molecular biology. Biochemistry grew from physiology and chemistry early in the twentieth century and gave fruit to molecular biology in mid-century. The term 'cell biology' first appeared in PubMed in 1917, occurred sparsely for many years, crossed the 10 paper per year mark in 1948, did not reach 1000 papers per year until the mid-1980s, and currently appears in more than 25 000 papers per year. The driving force of this explosion of information is the inherent interest in biological systems, i.e., ourselves, as well as the intrinsic importance to medicine, i.e., ourselves in a pathological state.

As these sciences moved into the twentieth century, the dividing lines among these disciplines became blurred so that an *Encyclopedia of Cell Biology* must begin with a presentation of the essentials of biochemistry and molecular biology that are needed to understand cell biology. Currently, these disciplines are nearly indistinguishable with a considerable overlapping continuum. Good investigators follow significant research questions regardless of what discipline is involved in the solutions. Biology follows a continuum from atom to molecule to system to organism. Advances in these related scientific areas could not have been made without parallel development of experimental techniques capable of asking and answering the appropriate questions. This first subsection **Basic Molecular Components and Technology** leads to the remainder of the first major section of the Encyclopedia, **Molecular Aspects/Molecular Principles, Components, Technology, and Concepts Important for Understanding Cell Biology** and deals with evolving techniques that have helped elucidate the nature of the molecules, complexes, organelles, cells, and interactions involved in living organisms. The emphasis is not on the techniques themselves, but rather on what the technology reveals about the basic components of biological cells and systems.

The early 'central dogma' of 'DNA makes RNA makes protein' has become outdated since the discovery of reverse transcription in certain RNA viruses and the discovery of catalytic RNA. The concept of one gene for one protein had to be expanded and modified due both to splicing and post-translational modifications. The terminology of noncoding DNA as 'junk' DNA has been made passé by microRNA and other types of noncoding RNA. Thus, the field of biochemistry/molecular biology/cell biology, like with all science, advances. Indeed, as the evolution of the understanding of cell biology has progressed, we have found broader concepts to replace

simpler ones resulting in an ever-increasing appreciation of the complexity of structure, function, regulation, and growth of living organisms. Many of these principles and molecules fundamental to cell biology have also been treated from a somewhat different perspective and format in the related Elsevier publication '*Encyclopedia of Biological Chemistry*,' second edition, and should be consulted for additional information.

Since life is based on chemistry, the *Encyclopedia of Cell Biology* starts with certain chemical principles of equilibria, bonding, and catalysis before moving into structural considerations. The biological structures discussed are the four usual major classes of biomolecules: nucleic acids, proteins, lipids, and carbohydrates. More complex molecules/systems encompass more than one of these groups, for example, glycoproteins glycolipids, lipoproteins, and membranes. Several articles are devoted to metabolites and metabolism ('classical' biochemistry and essential for functioning of cells). Certain diseases that are linked to molecular function or dysfunction, for example, cystic fibrosis (see [Cystic Fibrosis](#)) and Huntington's disease (see [Diseases of Protein Folding: Huntington's Disease and Amyotrophic Lateral Sclerosis](#)), are specifically presented.

Topics making up the first section, in order, are basic chemical principles (including equilibria) (see [Chemical and Physical Principles](#)), bonding (see [Chemical Biology](#)), and catalysis (see [Biocatalysis](#)); nucleic acids (including chemical properties and sequencing) (see [DNA, RNA Chemical Properties \(Including Sequencing and Next-Generation Sequencing\)](#)), chemical synthesis (see [The Chemical Synthesis of DNA and RNA Oligonucleotides for Drug Development and Synthetic Biology Applications](#)), cloning and expression systems (see [Expression Systems](#)) and site-directed mutagenesis (see [Site-Directed Mutagenesis](#)); Proteins (including structure and domains) (see [Protein Domains: Structure, Function, and Methods](#)), folding and misfolding (see [Folding, Misfolding, Disordered Proteins, and Related Diseases; Diseases of Protein Folding: Huntington's Disease and Amyotrophic Lateral Sclerosis](#)), posttranslational modifications (see [Posttranslational Modifications: Key Players in Health and Disease](#)), drug design (see [Drug Design](#)), antibodies (see [Antibodies and Improved Engineered Formats \(as Reagents\)](#)), sequencing (see [Protein Sequence Determination: Methodology and Evolutionary Implications](#)), purification (see [Isolation/Purification of Proteins](#)) and spectroscopy (see [NMR in Structural and Cell Biology](#)); Lipids (including cholesterol) (see [Cholesterol and Other Steroids](#)), complex lipids (see [Synthesis and Structure of Glycerolipids; Glycolipids](#)), signaling (see [Lipid Signaling](#)) and

lipidomics (see [Lipidomics](#)); Membranes (including properties and biogenesis) (see [Composition, Physical Properties, and Curvature](#)), rafts (see [Lipid Rafts/Membrane Rafts](#)), electrochemical potential (see [Membrane Potential: Concepts](#)), nerves (see [Neuronal Action Potentials and Ion Channel Allostery](#)) and mitochondria (see [The Outer Mitochondrial Membrane, a Smooth 'Coat' with Many Holes and Many Roles: Preparation, Protein Components, Interactions with Other Membranes, Involvement in Health, Disease, and as a Drug Target](#)); Complex carbohydrates (including glycogen) (see [Glycogen and Starch](#)), proteoglycans (see [Proteoglycans](#)) and hyaluronan (see [Hyaluronan](#)); and Metabolites (including regulation) (see [Metabolic Regulation](#)), bioenergetics and organelles (see [A Structure Perspective on Organelle Bioenergetics](#)).

Nucleic Acid Synthesis/Breakdown

Deoxyribose nucleic acid (DNA) was first discovered in 1869 by the Swiss scientist, Friedrich Miescher. Nucleic acids are biopolymers comprised of nucleotide monomers that are composed of three moieties, a five-carbon sugar, a phosphate group, and a nitrogenous base. DNA contains deoxyribose as the sugar component and RNA contains the sugar ribose. Polynucleotides are formed by covalent linkages between the phosphate of one nucleotide and the sugar of another, resulting in phosphodiester linkages. Nucleic acids are the major information molecules of all known forms of life by encoding, transmitting, and expressing genetic information. Elucidation of the structure of DNA by Watson and Crick in 1953 immediately suggested the semiconservative mechanism by which DNA is accurately reproduced and later demonstrated by Meselson and Stahl. The ability of DNA sequences to be copied into RNA or into copies of DNA with high fidelity in a template-dependent fashion is one of the most fundamentally important processes in living organisms. The transfer of genetic information from DNA to RNA to protein is considered the fundamental process of all living systems. However, as occurs in certain viruses, it is possible to transfer sequence-encoded information from RNA to DNA by reverse transcription. Topics in this section of the *Encyclopedia of Cell Biology* cover all major roles of nucleic acids in information transfer, including the mechanisms by which genetic information is regulated. This section additionally covers how RNA molecules transmit, edit, splice, and regulate the expression of genetic information – processes that greatly increase the number of combinations in which protein-coding sequences can be used.

Topics in this section, in order, are: [Comparison of Bacterial and Eukaryotic Replisome Components](#), which describes the structures and mechanisms of the protein machinery that replicates DNA in bacteria and eukaryotes; [Transfer RNA](#), which not only describes how tRNAs serve as adaptor molecules to translate mRNA sequences into protein sequence but also describes tRNA's roles in many other cellular processes; [Messenger RNA \(mRNA\): The Link between DNA and Protein](#) compares mRNA's structure/functions in bacteria and eukaryotes and how mutations in untranslated

regions of mRNAs contribute to human disease; [Telomeres and Telomerase](#) describes the end-capping structures, called telomeres, on eukaryotic chromosomes that are critical for chromosome stability, and discusses the importance of telomerase enzymes in maintenance of the telomere; [Telomere Biology](#) describes how telomeres maintain genome stability and prevent cellular senescence, and how cancer cells bypass the normal limits on telomere elongation to gain immortality; [Small RNAs/Cancer](#) reviews the discovery, biogenesis, and roles of microRNAs in cancer etiology; [Eukaryotic Nucleotide Excision Repair](#) describes highly conserved DNA repair mechanisms that remove and repair a wide variety of chemical lesions in DNA structure, which in some cases contribute to the onset of cancer; [The Base Excision Repair Pathway](#) reviews a pathway that is critical to genomic stability by correcting small DNA base lesions, first by excising the damaged locus, and then by replacing damaged nucleotides with the correct ones; [Nonhomologous DNA End Joining](#) describes and compares how cells repair double-strand breaks in DNA by either nonhomologous DNA end joining or by homologous recombination (HR) mechanisms; [DNA Repair by Homologous Recombination](#) reviews mechanisms and proteins involved in repairing complex DNA damage by HR and how specific HR proteins protect stalled replication forks; [Prokaryotic Transcription](#) reviews the current understanding of transcription in bacteria from the recognition of DNA to the termination of RNA synthesis; [Eukaryotic Transcriptional Regulation](#) describes current models of transcriptional regulation in eukaryotes; [Distant Activation of Transcription by Enhancers](#) reviews how distantly-acting DNA enhancers activate transcription by chromatin looping; [miRNAs/Small Noncoding RNAs](#) describes how precursor mRNAs in eukaryotes are spliced to remove intervening sequences by a complex ribonucleoprotein complex called the spliceosome; [Pre-mRNA Splicing: Function and Dysfunction](#) further expands on the process of splicing pre-mRNA and describes human diseases resulting from dysregulation of the splicing machinery; [Ribosomal RNAs and Protein Synthesis](#) discusses the structure of ribosomes and how the ribosomal RNA functions in the synthesis of proteins; [miRNAs/Small Noncoding RNAs](#) reviews our current knowledge about the biological synthesis and processing of microRNAs, which are typically 22 nucleotides long and function in translational repression and other regulatory processes; [The Interplay between Eukaryotic mRNA Degradation and Translation](#) summarizes current views on how eukaryotic mRNA degradation interconnects with mRNA translation to regulate gene expression; [Riboswitches and Ribozymes](#) describes the biological roles and potential as tools of riboswitches and ribozymes, which are regions of mRNA that regulate gene expression by binding small molecules, or are RNA that catalyze chemical reactions, respectively; [Transgenesis and Gene Replacement](#) presents an overview of the most widely used methods for experimentally manipulating the genomes of mammals to understand gene functions; and see [Viral Nucleic Acids](#) provides an overview of viral genomes, which can be either RNA or DNA, and how they are replicated in host cells. This collection of articles on nucleic acids not only provide the reader a comprehensive understanding of the state of the science in nucleic acid research, but also provides

a solid foundation for more focused investigations into the topic.

Protein Synthesis and Degradation

This section focuses on molecular mechanisms underlying the regulation of protein concentration and active forms of proteins both inside and outside cells and highlights diseases/pathologies that can result in dysregulation of proteins. Just as the knowledge/understanding of RNA and DNA and other fundamental components of cells and tissues have evolved, so have those of the proteins that make up cells, the extracellular environment, and fluids of living organisms. The complexity of the interactions of proteins has become apparent and the importance of networks of proteins (death pathways, blood coagulation, complement), interacting proteases (the 'protease web'), extracellular matrix, intracellular scaffolds, and membrane, organelle, and factor interactions with intracellular and extracellular proteins are now clearly realized.

Many of the structures involved in the synthesis of proteins have been solved (including the ribosome, translation factors) and there has been considerable progress clarifying the mechanisms that control initiation, elongation, and termination in eukaryotic cells (see [Components, Initiation, Elongation, Termination, and Regulation](#)). Similarly the components of the mitochondrial protein biosynthetic machinery and mechanisms of transcription and translation have been elucidated (see [The Protein Biosynthetic Machinery of Mitochondria](#)). Many proteins are synthesized as inactive or latent proteins (proproteins) and have to be activated by enzymes such as proprotein convertases and kallikreins (see [Regulated Proteolysis of Signaling Molecules: The Proprotein Convertases; Kallikrein](#)). Biosynthesis of secretory proteins that takes place in the endoplasmic reticulum (ER) involves multiple factors (signalases, chaperones, isomerases) to form a mature correctly folded protein (see [Biogenesis of Secretory Proteins](#)). Factors are in place for quality control in the ER to recognize and shuttle incorrectly folded proteins into the ER-associated degradation pathway (see [Endoplasmic Reticulum-Associated Degradation and Protein Quality Control](#)).

It was once believed one protease could degrade many proteins in a relatively unregulated manner (somewhat like trypsin in the intestinal tract). It is now known that there are multiple proteases in all cells and fluids that are highly regulated. There are ~600 genes for proteases in the human genome and if there were no regulation there would be widespread necrosis, cell death, and destruction. It is clear that proteolytic systems are highly regulated by localization, activation/inhibition, synthesis of latent proenzymes, interactions with multiple components such as cofactors, carbohydrates, lipids, membranes, organelles, and the pH of the environment. The complement of all proteases and their substrates is known as 'the degradome' and high-throughput methods have recently been developed to study the network of protease and substrate interactions (see [Mass Spectrometry-based Methodologies for Studying Proteolytic Networks and the Degradome](#)).

Proteolytic systems exist both intracellularly and extracellularly as well as at cell membranes. The mammalian intestinal system serves as a good example of coordinated digestion of food proteins that involves many different types of extracellular proteases (see [Digestive Proteases: Roles in the Human Alimentary Tract](#)). Blood coagulation represents another extracellular system that is critical to host defense and hemostasis and involves cells (e.g., platelets), a host of plasma proteins (most of which are proteases) and protease inhibitors, which are highly controlled to form and degrade blood clots appropriately (see [Overview of Blood Coagulation and the Pathophysiology of Blood Coagulation Disorders](#)). The complement system involves highly regulated proteolytic enzymes critical to our immune systems that remove targeted pathogens (see [Molecular Mechanisms Underlying the Actions of the Complement System](#)). The proteasome is key to intracellular proteolytic systems and works in coordination with ubiquitin-tagged proteins for protein quality control and several signaling systems (see [Ubiquitin, Ubiquitin-Like Proteins, and Proteasome-Mediated Degradation](#)). In addition, the intracellular lysosomal system containing multiple hydrolases (e.g., proteases, glycosidases) that interact with cellular and extracellular components to degrade and maintain protein/amino acid homeostasis (see [Role of Lysosomes in Intracellular Degradation](#)). Dysregulation of all of these systems is associated with disease, and there are specific articles on lysosomal storage diseases (see [Lysosomal Diseases](#)) and the proteases involved in the progression of cancer (see [Cancer – Proteases in the Progression and Metastasis](#)).

Individual proteases are discussed in many articles and some that are highlighted are matrix metalloproteinases (see [Matrix Metalloproteinases](#)), calpain (see [The Calpain Proteolytic System](#)), ADAMS and ADAMTS (see [ADAMTS Proteases: Mediators of Physiological and Pathogenic Extracellular Proteolysis; ADAMs Regulate Cell–Cell Interactions by Controlling the Function of the EGF-Receptor, TNF \$\alpha\$ and Notch](#)), membrane-anchored serine proteases (see [Extracellular: Plasma Membrane Proteases – Serine Proteases](#)), meprins (see [Metalloproteases Meprin \$\alpha\$ and Meprin \$\beta\$ in Health and Disease](#)), kallikreins (see [Kallikrein](#)), and aspartic proteases (see [Aspartic Proteases of Alzheimer's Disease: \$\beta\$ - and \$\gamma\$ -Secretases; Cathepsin E: An Aspartic Protease with Diverse Functions and Biomedical Implications](#)). Protease inhibitors are also critical to the regulation of protein degradation and disease. Endogenous polypeptide protease inhibitors exist for many of the proteolytic systems (e.g., apoptosis, blood coagulation; [Naturally-Occurring Polypeptide Inhibitors: Cystatins/Stefins, Inhibitors of Apoptosis \(IAPs\), Serpins, and Tissue Inhibitors of Metalloproteinases \(TIMPs\)](#)). Alpha-1-antitrypsin deficiency is an example of how important endogenous protein inhibitors are for preventing disease, as this deficiency leads to liver damage and emphysema (see [Alpha-1-Antitrypsin Deficiency: A Misfolded Secretory Glycoprotein Damages the Liver by Proteotoxicity and Its Reduced Secretion Predisposes to Emphysematous Lung Disease Because of Protease-Inhibitor Imbalance](#)). Examples of synthetic protease inhibitors used to control disease process are blood pressure inhibitors (ACE inhibitors; [Blood Pressure, Proteases and Inhibitors](#)) and HIV-protease

inhibitors that have been designed to manage AIDS (*see* [Inhibitors of HIV Protease](#)).

In summary, the purpose of part I of the Encyclopedia is to describe the chemical principles and the molecular components/organization of the cell and its environment in

sufficient detail to allow a clearer understanding of how these components are then assembled, function, and are regulated at a higher level – the themes of the next three parts of the Encyclopedia.

MOLECULAR PRINCIPLES, COMPONENTS, TECHNOLOGY, AND CONCEPTS: BASIC PRINCIPLES

Contents

Chemical and Physical Principles

Biocatalysis

Chemical and Physical Principles

H Oubrahim and P Boon Chock, National Institutes of Health, Bethesda, MD, USA

Published by Elsevier Inc.

Introduction

Cells of all living organisms contain four classes of macromolecules or polymers consisting mainly of carbon, oxygen, nitrogen, hydrogen, and a small quantity of sulfur and phosphate atoms. Proteins are polymers of amino acids, which serve as enzymes, regulators, structure elements, and receptors; DNA and RNA are polymers of nucleotides, which store and transmit genetic information, and polysaccharides are polymers of simple sugars, which serve as energy-rich fuel stores. To investigate the molecular processes that make life possible, it is essential that investigators are well versed in fundamental chemical and physical principles, since they govern both the thermodynamics and dynamics of biochemical processes.

Living cells must perform work to stay alive, grow, replicate, and evolve, while maintaining themselves in a dynamic steady state, far from equilibrium with their environment. To understand how cells accomplish these processes, within the topic limitation of this article, we focus our discussion on the physical and chemical principles that govern interactions between cellular molecules. They include the energetic aspect that determines whether a molecular process can occur spontaneously, and the kinetic and regulatory aspects of biological processes, as well as effects of the crowded cellular environment.

The Laws of Thermodynamics and Living Cells

Living cells have developed highly efficient mechanisms to utilize the energy obtained from chemical fuels and light to carry out numerous energy-requiring processes in order to maintain themselves in dynamic steady states. When a cell fails to obtain energy, it will die and decay toward equilibrium with its surroundings. To understand how the energy is extracted, stored and channeled into useful work in living cells, we address cellular energy conversions in context of the law of thermodynamics and the quantitative relationships among free energy, enthalpy and entropy.

The laws of thermodynamics are general principles that provide the quantitative description of heat and energy changes and chemical equilibria. These laws apply to all chemical and physical processes, including biochemical reactions ([van Holde *et al.*, 1998](#); [Edsall and Gutfreund, 1983](#); [Alberty, 2003](#)). Their importance resides in the fact that they determine the conditions in which a biochemical reaction can proceed.

In thermodynamics, the field of observation is divided into two conceptual regions: the system and the surroundings. The 'system' refers to everything within a defined region of space, including all the constituent reactants, products, solvent of the reaction, and the immediate atmosphere; while the system and its surroundings together constitute the universe. When the system does not exchange either matter or energy with its surroundings, it is considered isolated. If the system exchanges only energy and not matter with its surroundings, it is defined as a closed system. An open system is one that exchanges both matter and energy with its surroundings. The first law of thermodynamics describes the principle of the conservation of energy. In any physical or chemical change, the total energy, E , of a system and its surroundings is constant, although the form of the energy may vary. In other words, the first law states that E can be changed only by the flow of energy as heat or by work. Consequently, energy can neither be created nor destroyed, it can only be changed from one form to another as shown in the mathematical expression below,

$$\Delta E = E_B - E_A = Q - W \quad [1]$$

where E_A and E_B is the energy of a system at the beginning and the end of the transformation, respectively, Q is the heat absorbed, and W is the work done by the system. Note the change in energy of a system depends only on the initial and final states, independent of the transformation pathway.

When a given chemical reaction occurs under constant pressure, the amount of heat released or absorbed reflects the nature and number of chemical bonds altered during the course of the reaction. This heat of reaction is referred to as enthalpy, H , expressed as joules/mol or calories/mol

(1 cal = 4.184 J). Because the total enthalpy of a system cannot be measured directly, only the change of enthalpy, ΔH , is evaluated. If heat is being absorbed by the reaction, its ΔH is positive and the reaction is endothermic. On the other hand, if heat is generated by the reaction, the reaction is exothermic and its ΔH is negative. However, the first law of thermodynamics is insufficient to predict whether a reaction can occur spontaneously since some endothermic reactions do occur spontaneously. Thus, a function other than ΔH is necessary to account for this observation. One such function is the entropy, S , expressed in unit of J/mol K. Note that entropy is a quantitative expression for the randomness or disorder in a system. When the products of a reaction are less complex and more disordered than the reactants, the reaction proceeds with a gain in entropy. It is worth mentioning that entropy is a central concept in biochemistry since life requires continual maintenance of order while increased randomness is the natural tendency. The second law of thermodynamics states that a process can occur spontaneously, if and only if, the sum of the entropies of the system and its surroundings is > 0 . This indicates that the entropy of a system can decrease during a spontaneous reaction, if the entropy of the surroundings increases such that their sum is positive. However, the entropy changes of chemical reactions are not readily determined and the second law indicates that to determine whether the reaction can occur spontaneously requires one to know the value of the entropy changes for both the surroundings and the system of interest. At constant temperature and pressure, a condition fulfilled by most biological systems, this constraint imposed by the second law can be obviated by using a different thermodynamic state function termed free energy (G) or Gibbs' free energy, derived from the combining of the first and second law of thermodynamics by Gibbs (1876–1878, 1878).

The basic equation is:

$$\Delta G = \Delta H - T\Delta S \quad [2]$$

where ΔG is the change in Gibbs free energy of a reaction under constant pressure, P , and temperature, T , and ΔS and ΔH is the change in entropy and in enthalpy of the reaction, respectively.

Gibbs Free Energy Always Decreases for a Spontaneous Process at Constant Temperature and Pressure

All reactions are generally affected by two forces: The tendency to achieve the most stable chemical bond, indicated by ΔH , and the tendency to achieve the highest degree of randomness, expressed by ΔS . The net effect of these two factors is summed up by the change of Gibbs free energy described by eqn [2]. Thus, ΔG provides a valuable criterion for determining whether a reaction can occur spontaneously. By convention, ΔS is positive when entropy increases and ΔH is negative when heat is released by the system to its surroundings. When either of these conditions is energetically favorable the reaction tends to yield a negative ΔG , a condition for a spontaneous reaction. To determine the actual free energy change, ΔG , for the reaction [3], one needs to take into account the nature of the reactants and products as well as their concentrations as shown

in the eqn [4].



where a , b , c , and d represent the stoichiometry of the indicated components.

$$\Delta G = \Delta G^0 + RT \ln \left(\frac{[C]^c [D]^d}{[A]^a [B]^b} \right) \quad [4]$$

where ΔG^0 is the standard free energy change, a constant that is characteristic of each reaction, R is the gas constant, T is the absolute temperature, and $[A]$, $[B]$, $[C]$ and $[D]$ are molar concentrations of reactants and products. (More precisely, the activity, defined as a thermodynamic function that correlates changes in chemical potential with changes in concentration, through relations formally equivalent to those for ideal systems. For practical reasons, molar concentration is used in biochemical literature.) ΔG^0 is the change in free energy under standard conditions (298 K and 1 atm) when reactants and products are initially present at 1 M, or for gas at 1 atm. Equation [4] indicates that ΔG of a reaction depends on the nature of the reactants, expressed by ΔG^0 , and their concentrations, expressed by the second term of the equation.

With this definition, the standard state for reactions involving hydrogen ions is $[H^+] = 1$ M or pH 0. However, most biochemical reactions occur in relatively well buffered aqueous media at a pH around 7, such that both the pH and the concentration of water (55.5 M) are essentially constant. To simplify ΔG^0 calculation, biochemists adopted a convention that in the biochemical standard state, $[H^+]$ is 10^{-7} M (pH7), $[H_2O]$ is 55.5 M, and when the reactions involve Mg^{2+} (when ATP is a reactant), $[Mg^{2+}]$ is 1 mM, a standard state different from that used in chemistry and physics. With this convention, when H_2O , H^+ , or Mg^{2+} are reactants or products, their concentrations are not included in the K_{eq} expression, but are incorporated, with a value of 1 for each of their activities, into the constants K'_{eq} and $\Delta G'^0$. These constants are referred to as standard transformed constants written with a prime, to distinguish them from the standard constants used by chemists and physicists. Thus, the standard transformed free energy change at pH 7 is denoted as $\Delta G'^0$, which can be calculated using the equilibrium constant, K'_{eq} , and the equation $\Delta G'^0 = -RT \ln K'_{eq}$ where $K'_{eq} = \frac{[C]_{eq}^c [D]_{eq}^d}{[A]_{eq}^a [B]_{eq}^b}$. Therefore, the criterion for the spontaneity of a biochemical reaction depends on the value of ΔG , expressed as

$$\Delta G = \Delta G'^0 + RT \ln ([products]/[reactants]) \quad [5]$$

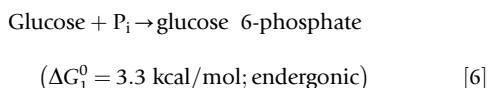
Gibbs Free Energy Changes are Additive

The free energy change of a reaction is independent of its reaction pathway. However, when a reaction consists of two or more successive reactions, and each of the two successive reactions shares a common intermediate, the Gibbs free energy change of the net reaction is equal to the sum of the ΔG of the individual reactions. In other words, the Gibbs free energy changes are additive. This property allows one to determine the free energy of the formation of complex molecules needed to sustain living cells, from ΔG of the individual reaction steps that lead to the formation of the final complex molecules.

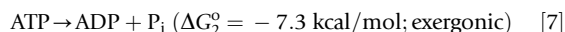
Consequently, a thermodynamically unfavorable reaction, with a positive ΔG , can be driven by a thermodynamically favorable reaction to yield a negative ΔG for the sum of free energy changes of the two reactions. Utilizing this energy-coupling strategy, biological systems are able to synthesize and maintain the information-rich polymers required to sustain living cells, and to formulate their metabolic pathways by coupling enzyme-catalyzed reactions such that the overall ΔG of the pathway is negative. In biological systems, this principle is frequently utilized to couple the energy of ATP hydrolysis, a highly exergonic reaction, to otherwise unfavorable reactions, such as those involved in the biosynthetic pathways.

Coupling of ATP Hydrolysis to Drive Thermodynamically Unfavorable Reactions

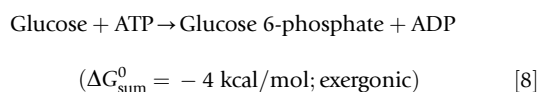
ATP, maintained at a relatively high and constant level in the cell, has been considered as the molecular currency of intracellular energy transfer. It has been extensively utilized to drive thermodynamically unfavorable biochemical pathways. To illustrate how this principle works, we consider the coupling of ATP hydrolysis to the synthesis of glucose 6-phosphate, an endergonic first reaction step in glucose oxidation pathway:



where P_i is inorganic phosphate. $K_{\text{eq}1} = [\text{Glucose 6-phosphate}] / ([\text{Glucose}][\text{P}_i]) = 10^{(-\Delta G_1^0/1.36)} = 3.98 \times 10^{-3} \text{ M}^{-1}$ at 25°C (Note $2.3 \times RT = 1.36 \text{ kcal mol}^{-1}$). This indicates that there is no net conversion if the molar ratio is equal to or larger than $3.98 \times 10^{-3} \text{ M}^{-1}$. When this reaction is coupled to:



where $K_{\text{eq}2} = ([\text{ADP}][\text{P}_i]) / [\text{ATP}] = 2.18 \times 10^5 \text{ M}$. The sum of these two reactions becomes



The equilibrium constant of this coupled reaction becomes

$$K_{\text{eq}} = K_{\text{eq}1} \times K_{\text{eq}2} \\ = ([\text{Glucose 6-phosphate}][\text{ADP}]) / ([\text{Glucose}][\text{ATP}]) \\ = 10^{(-\Delta G_{\text{sum}}^0/1.36)} = 8.7 \times 10^2.$$

At equilibrium, the ratio of $[\text{Glucose 6-phosphate}] / [\text{Glucose}][\text{P}_i] = K_{\text{eq}} \times [\text{ATP}]_{\text{eq}} / ([\text{ADP}]_{\text{eq}}[\text{P}_i]_{\text{eq}})$.

The ATP generating system in cells maintains the $[\text{ATP}]_{\text{eq}} / ([\text{ADP}]_{\text{eq}}[\text{P}_i]_{\text{eq}})$ at a high level, typically on the order of 500; thus

$$[\text{Glucose 6-phosphate}] / ([\text{Glucose}][\text{P}_i]) = 8.7 \times 10^2 \\ \times 500 = 4.35 \times 10^5 \text{ M}^{-1}$$

Together, this indicates that by coupling the conversion of glucose to glucose 6-phosphate to ATP hydrolysis in the cell,

under standard conditions would shift the equilibrium ratio of $[\text{Glucose 6-phosphate}]$ to $[\text{Glucose}] \times [\text{P}_i]$ by a factor of 10^8 .

This example demonstrates the thermodynamic essence of ATP's action as an energy-coupling agent. Since all living cells maintain a concentration of ATP much higher than its equilibrium concentration, coupling of a cellular reaction with the hydrolysis of an ATP molecule can change the equilibrium ratio of products to reactants by a huge factor, for example, by a factor of 10^8 . Thus, ATP functions as a major carrier of chemical energy in cells to convert thermodynamically unfavorable reaction sequences into favorable ones.

Reaction Rate and Rate Constant

Note that ΔG of a reaction depends only on the nature of the reaction and the concentration of both reactants and products, but ΔG is independent of the mechanistic pathway. Furthermore, ΔG does not provide information on the rate of a chemical reaction. In fact some thermodynamically favorable reactions fail to take place at measurable rates due to the high activation energy required for the reactions. To overcome this problem, biological systems utilize enzymes to catalyze slow reactions by lowering the activation energy via an alternative reaction pathway to facilitate the formation of the transition state. Like all catalysts, enzymes cannot alter the equilibrium constants of a reaction. They only increase the rate by which a reaction proceeds in the direction governed by thermodynamics. In other words, thermodynamics cannot provide information about intervening states of the system.

While thermodynamics predicts whether a reaction can occur, and how much energy can be derived from them, the concentration of most molecules in living cells is maintained in dynamic steady states, and not by equilibrium constants. Therefore, rates of biochemical reactions, mostly catalyzed by enzymes, play an important role in shaping the metabolic pathways in living cells. The fact that about a quarter of the protein-encoding genes in the human genome encode enzymes and their regulatory proteins points to the importance of understanding the kinetics of biochemical reactions in cellular regulation and metabolism (For further reading on kinetics, see references [Moore and Pearson, 1981](#); [Purich, 2010](#); [Connors, 1990](#)).

The rate, $d[\text{P}]/dt$, for the formation of a reaction product, P , is determined by the concentration of the reactant(s) and its rate constant, k . In general, a rate equation has the form

$$d[\text{P}]/dt = k[\text{S}_1]^a[\text{S}_2]^b \dots \quad [9]$$

The value of k is independent of the concentration of S_1 , S_2 , ..., but dependent on the nature of reaction investigated, and the environmental conditions of the reaction, such as temperature and chemical properties of the reaction medium. The power variables a and b represent the order of reaction with respect to S_1 and S_2 , respectively, and their sum ($a + b + \dots$) represents the overall order of the reaction. When eqn [9] is in a form of simple rate equations, it represents the rate expression of elementary (single-step) reactions. Some complex (multiple step) reactions may also possess simple rate equations. Nevertheless, complicated rate equations are required for

kinetic analysis of complex reactions such as those involving parallel or consecutive reactions.

When a reaction is unimolecular, eqn [9] is reduced to $d[P]/dt = k[S_1]$. In this case, the rate of the reaction depends only on k and the concentration of S_1 . This reaction is defined as a first-order reaction, and k is a first order-rate constant that has a unit of reciprocal time, for example, s^{-1} . A first-order reaction proceeds via an exponential function and its half-life, $t_{1/2} = \ln 2/k$ or $0.693/k$, can be used to calculate its rate constant. Note that the $t_{1/2}$ of a first-order reaction is constant, independent of the initial time point used for analysis. In the case of a bimolecular reaction, the reaction rate depends on the concentration of two different reactants, or two molecules of the same reactant, such that $d[P]/dt = k[S_1]^2$ or $k[S_1][S_2]$, where k is the second-order rate constant, with an unit of $M^{-1}s^{-1}$. The time course of the second-order reaction is qualitatively similar to that of the first-order reaction, except the consumption of the reactant(s) proceeds faster initially and slower toward the end of the reaction due to the nature of second-order dependency. A third-order reaction, whose rate depends on the product of three concentrations, is relatively rare. In the case of a zero-order reaction, its rate is independent of the concentration of the reactant, and the unit of its rate constant is $M s^{-1}$. A zero-order reaction proceeds with a linear time course. This type of reaction is not commonly observed except in heterogeneous systems and in catalyzed reactions when the catalyst is present at a significantly lower concentration than that of the reactant, such that the catalyst is saturated with the reactant. Under this condition, the concentration of the reactant will not change significantly during the course of the initial rate measurement, and the reaction rate depends only on the concentration of the catalyst used.

Reaction Rate-Limiting Step

When the conversion of reactant S to product P proceeds via formation of a number of intermediates, for example, I_1 and I_2 , and if the formation of I_1 , I_2 , or P is very much slower than the rest of the reaction steps, then the rate of P formation will be limited by the rate of the slowest step, a reaction that possesses the highest activation energy, $\Delta G^\ddagger$. This slowest step is called a rate-limiting or rate-determining step since it is the 'bottleneck' of the overall reaction. Strictly speaking, this flow analogy is valid only for consecutive and irreversible reactions, and it can be misleading if the reverse reaction is significant. In fact even for irreversible reactions, the rate-determining step concept is meaningful only if one of the reactions is much slower than the others. Note when the overall reaction includes more than two elementary steps, the situation may not be easy to analyze, since the product of the n th step is the reactant of the $(n+1)$ st step. For these two states to be represented by the same free energy they must have the same composition. This means that the stoichiometric composition must be constant throughout the entire successive reactions (Boyd, 1978).

Rate Constant and Activation Energy

To co-relate the magnitude of a rate constant to Gibbs free energy of activation, $\Delta G^\ddagger$, a transition-state theory was derived.

This theory assumes that the transition state is in equilibrium with reactants, such that the population of the transition state is governed by Gibbs free energy of activation, $\Delta G^\ddagger$. Thus, the magnitude of a rate constant can be expressed as a function of Gibbs free energy of activation and temperature:

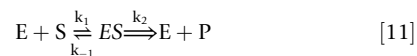
$$k = (k_B T/h)(e^{-\Delta G^\ddagger/RT}) \quad [10]$$

where k_B is the Boltzmann constant, and h is the Planck's constant. The important point to emphasize here is that the relationship between the rate constant and the activation energy, $\Delta G^\ddagger$, is inverse and exponential. In other words, this is the basis for the statement that lower activation energy yields a faster reaction rate.

Enzyme Kinetics

Equation [10] shows that the function of catalysts is to lower the activation energy, $\Delta G^\ddagger$, of a reaction and thereby speed up the reaction rate. Most of the catalysts in biological systems are enzymes, and the majority of enzymes are proteins. The majority of enzymes are known to catalyze only one particular type of reaction under very limited chemical and physical conditions, and they are capable of enhancing a reaction rate up to factors of 10^6 or more. Enzyme-catalyzed reactions are characterized by the formation of an enzyme-substrate, ES, complex. The catalytic specificity and capacity of an enzyme are derived mainly from multiple weak interactions between the enzyme and its substrates, mediated by hydrogen bond formation, hydrophobic, and ionic interactions. Furthermore, the enzyme active site tends to be structured such that some of these interactions occurred specifically to stabilize the transition state. The need for multiple interactions could be one of the reasons that an enzyme is a relatively large size molecule. The enzyme-substrate binding energy could be utilized to offset the energy required for activation, as well as to induce protein conformational change at the active site to properly position catalytic functional groups to facilitate the cleavage and formation of chemical bonds by a variety of mechanisms, including general acid-base catalysis, covalent catalysis, and metal ion catalysis. These processes could lead to transient covalent interactions with a substrate or group transfer to or from the enzyme, to provide a new, lower-energy reaction pathway. Once the catalytic action is completed, including the release of the tightly bound product(s), the enzyme reverts to its unbound state.

Multiple techniques have been employed to elucidate the catalytic mechanism of purified enzymes. They include three-dimensional structural analysis of enzymes, coupled with chemical modification and site-directed mutagenesis to examine the role of individual amino acid residues in enzyme structure and action. However, kinetic study of enzyme action remains the most important method for elucidating the mechanistic action of enzymes. Kinetic properties of many enzymes can be described by the Michaelis-Menten model. In this model, an enzyme, E , first forms an ES complex, where S is the substrate, prior to product, P , formation.



The initial rate of product generation can be described by the Michaelis–Menten equation:

$$v = V_{\max}([S]/([S] + K_M)) \quad [12]$$

where $V_{\max}$ is the reaction rate when the enzyme is fully saturated with the substrate and K_M is the Michaelis constant, $K_M = (k_{-1} + k_2)/k_1$. When k_2 is rate-limiting such that $k_2 \ll k_{-1}$, K_M is reduced to k_{-1}/k_1 , the dissociation constant of ES. It should be pointed out that the validity of the steady-state method does not depend on the assumption that $d[ES]/dt = 0$. This assumption is not valid at the beginning of the reaction when [ES] is being built up, and toward the end of the reaction when [S] is too low to maintain a constant [ES]. This problem was resolved by Wong in 1975 (Wong, 1975) as follow:

Based on the reaction scheme shown in eqn [11],

$$d[ES]/dt = k_1[S]([E]_0 - [ES]) - (k_{-1} + k_2)[ES] \quad [13]$$

The initial rate of product formation is,

$$v = k_2[ES] = k_2(k_1[S][E]_0 - d[ES]/dt)/(k_1[S] + k_{-1} + k_2) \quad [14]$$

When $d[ES]/dt$ is small relative to $k_1[S][E]_0$ where $[E]_0$ is the total concentration of the enzyme used, then

$$v = k_2[ES] = k_1k_2[S][E]_0/(k_1[S] + k_{-1} + k_2) \quad [15]$$

During the early phase of the reaction, if $[S] \gg [E]_0$, the rate of change in [ES] due to diminishing [S] will be relatively slow. Thus, the validity of steady state is closely tied to the high substrate to enzyme ratio. The maximal rate, $V_{\max} = k_2[E]_0$, and k_2 , also known as k_{cat} , or turnover number, is the number of substrate molecules being converted into product per unit time at a single catalytic site when the enzyme is fully saturated with substrate. The ratio k_{cat}/K_M provides a good indication of the catalytic efficiency of an enzyme. Note that the Michaelis–Menten equation is also applicable to bisubstrate reactions, which proceed via a ternary complex or double-displacement pathways. A short-hand method of expressing two or more substrates or products has been proposed by Cleland to differentiate order, random, Theorell–Chance, and Ping–Pong mechanisms (Cleland, 1963). It should be pointed out that while the kinetic constants can be accurately determined from *in vitro* study using a high substrate to enzyme ratio, this unique condition required for validating the Michaelis–Menten equation may not be met under cellular conditions. Therefore, one needs to be cautious in drawing conclusions from steady-state kinetic study under such circumstances because it can lead to intolerably high errors if analyzed with the unmodified Michaelis–Menten expression (Cha, 1970).

Allosteric Enzymes

All enzymes that exhibit hyperbolic kinetics with respect to increasing substrate concentration follow Michaelis–Menten kinetics, and their K_M represents the [S] required to achieve a rate equal to half $V_{\max}$. However, the Michaelis–Menten model fails to account for the kinetic properties of many enzymes. An important group of enzymes that do not obey Michaelis–Menten kinetics are the allosteric enzymes. An Allosteric

enzyme/protein is one in which binding of one substrate or ligand to one site affects the binding affinity of another site in the same molecule. When binding of one ligand impairs the subsequent binding of other ligands, its kinetics exhibits a negative cooperativity. In contrast, the activity of most allosteric enzymes displays a sigmoidal kinetics due to substrate binding at one site that promotes the subsequent binding affinity at a distinctly different site. This group of enzymes is generally, but need not, consists of multiple subunits or multiple active sites. They play a crucial role in many fundamental biological processes, including, but not limited to, cell signaling and metabolism regulation since they are susceptible to be regulated by signaling molecules. With the sigmoidal saturation response, the value of [S] at half-maximal rate cannot be designated as K_M because the enzymatic activity does not follow a hyperbolic function. Thus, the symbol $S_{0.5}$ or $K_{0.5}$ is often used to represent the substrate concentration required to achieve half-maximal velocity of the allosteric enzyme. It should be pointed out that this class of enzyme has important regulatory properties. On one hand, sigmoidal kinetics provides a mean for allosteric enzymes to greatly alter their catalytic output in response to a relatively small change in substrate/effector concentration. On the other hand, an enzyme with a strong negative cooperativity can provide a means for a rapid surge or decrease in enzymatic activity through changes in the concentration of an effector that desensitizes the negative cooperativity (Huang *et al.*, 1982). Interestingly, to date, among the known allosteric proteins, oxygen binding to hemoglobin, an oxygen carrying protein consists of four oxygen binding sites, is the most thoroughly studied. Consequently, hemoglobin is sometimes being referred to as an ‘honorary allosteric enzyme.’

From a kinetic standpoint a multisubunit enzyme is not required to achieve a cooperative type of kinetics, since it can be generated by a monomeric enzyme that can exist in several conformations at steady state or utilize alternative pathways in a multisubstrate reaction. The fact that regulatory enzymes, almost without exception, are polymeric, argues strongly for the role of subunit interaction in the cooperativity. The potential for achieving sophisticated control is undoubtedly greater with a polymeric, than with a monomeric enzyme. However, most of prevailing models for explaining the cooperative effects of subunits are based on ligand-promoted conformational equilibria. While they adequately describe the change of enzymatic activity in response to metabolite concentration, they fail to link intersubunit cooperation to the catalytic mechanism *per se*.

Concerted and Sequential Models

Two major models, the concerted model of Monod *et al.* (1965), and the sequential model of Koshland *et al.* (1966), have been proposed to explain the cooperative binding of ligands to multisubunit proteins. In the concerted model, Monod *et al.* assumed that an enzyme having identical and noninteracting subunits arranged in a symmetrical manner can exist in two conformational states, T and R, with different ligand binding affinities. A key feature of the Monod model is the conservation of symmetry, i.e., the T and R transition is a

concerted one such that all subunits in a given state are equivalent. The cooperative phenomenon, however, arises from the preexisting $T \rightleftharpoons R$ equilibrium in an apparent cooperativity. This two-state Monod model can be extended to accommodate more complicated situations. For instance, the original Monod model based on rapid equilibria cannot generate negative cooperativity. However, a proposed pseudo-conservative transition has been demonstrated as a modified Monod model that allows the two-state model to accommodate both positive and negative cooperativities (Viratelle and Seydoux, 1975). The sequential model of Koshland *et al.* is based on the induced-fit theory. Mathematically, this model is similar to those developed by Adair (Adair, 1925) and Pauling (Pauling, 1935). In this case the subunits are initially identical and binding of a ligand to one subunit affects the subsequent binding of ligands to the remaining subunits. This model is often referred to as the sequential model as opposed to the concerted model of Monod *et al.* The sequential model can be readily adapted to the case where the subunits are initially nonidentical (preexisting asymmetry). Since the binding of a ligand to one subunit can either improve or impair the binding affinities for the remaining subunits, positive, negative, and 'mixed' cooperativities may result. Importantly, both concerted and sequential models have been successfully applied, for example, by Schachman's and Koshland's groups, to the analysis of cooperative systems (Huang *et al.*, 1982).

Water as Life's Aether

Water, representing about 70% of cell's weight, plays a vital role in sustaining all forms of life. Therefore many aspects of life are affected by the physical and chemical properties of water, and many of these properties are derived from different electronegativities of H and O atoms that make water a highly polar molecule. As a consequence, most water molecules are in contact with their neighboring molecules through hydrogen bonding with itself and with solutes (Liu *et al.*, 1996). The hydrogen bonds are rapidly broken and reformed, with an average lifetime of $\sim 10^{-12}$ s. The capability of hydrogen bonding and its polarity make water a highly interacting molecule, as well as an excellent solvent for polar solutes by weakening electrostatic forces and forming hydrogen bonds between polar molecules. However, this property of water also poses a problem for living cells because it weakens interactions between polar molecules. To overcome this problem, biochemical systems generate hydrophobic microenvironments to maintain polar interactions at their maximal strength and specificity where needed. Furthermore, since water binds strongly to itself, it induces self-aggregation of nonpolar molecules such as lipids in an aqueous medium. This capacity of water facilitates the formation of cellular membranes that define the boundaries of cells and their internal components.

Another unique property of water is that it can ionize into H^+ and OH^- where H^+ exists as hydronium ions, H_3O^+ , in aqueous medium. For simplicity H^+ is used instead of the actual species present. On average, 1 out of 10^7 water molecules is ionized in its pure liquid. In biochemistry the concentration of H^+ is expressed as pH, defined as $pH = -\log$

$[H^+]$, where $[H^+]$ is in units of molarity. Thus, the greater the acidity of a solution, the lower its pH. A pH 7.0 solution contains $[H^+] = 1.0 \times 10^{-7}$ M. Since the concentration of water is 55.5 M, it does not change much under most biological conditions. Thus, the equilibrium constant for water can be simplified to

$$K_w = [H^+][OH^-] = 1.0 \times 10^{-14} M^2 \quad [16]$$

at 25 °C. This indicates that the ionization of water at 25 °C is highly unfavorable since its ΔG^0 is 19.1 kcal mol⁻¹. However, the extent of water ionization can be altered by the presence of other species that can bind either H^+ or OH^- . These species include proteins, DNA, RNA as well as cellular metal ions. From the K_w expression, one can obtain the $[OH^-]$ in aqueous solution knowing the pH value. For example, if $[H^+]$ is 10^{-3} M, then $[OH^-] = 10^{-14}/10^{-3} = 10^{-11}$ M. In essence, $[H^+]$ and $[OH^-]$ exhibit a reciprocal relationship.

Acid-Base Reactions Play a Central Role in Most Biochemical Processes

The equilibrium constant for the ionization of a weak acid, HA, can be described as $K = ([H^+][A^-])/[HA]$. The pK of this acid is defined as $pK = -\log K$. Thus the pH of a solution can be calculated using the eqn [17], known as Henderson-Hasselbalch equation, if the molar ratio of A- to HA and the pK of HA is known. Conversely, the pK of an acid can be calculated if the molar ratio of A- to HA and the pH of the solution is known.

$$pH = pK + \log([A^-]/[HA]) \quad [17]$$

When a solution contains a weak acid-base conjugate pair, for example, acetic acid and acetate, it can serve as a buffer with the capacity to prevent a significant change in its pH due to the addition of a small quantity of either strong acid or base. In general, the best buffer capacity of a given buffer system occurs in a range of one pH unit on either side of its pK. Interestingly, cells and organisms maintain a specific and constant cytosolic pH to keep biomolecules in their optimal ionic state. Furthermore, a significant change in pH can potentially lead to protonation or deprotonation of key functional groups of biomacromolecules that cause disruption of their molecular structures and lead to harmful biological effects. Thus, nature has evolved to minimize pH changes in biological systems. To this end, biological systems make use of a number of weak acids as their buffer systems to maintain a relatively constant physiological pH, typically around pH 7.4. Since a buffer functions best close to its pK value, among the biological buffers, phosphoric acid, which exists primarily in a near equal mixture of $H_2PO_4^-$ and HPO_4^{2-} at about pH 7.4, plays a major role to maintain physiological pH. Consistent with this notion, inorganic phosphate is present at about 1 mM in blood for maintaining its pH at 7.4.

Noncovalent Interactions Play Key Roles in Mediating Functions of Biomacromolecules

Weak noncovalent interactions exert a decisive role in maintaining the structure and function of macromolecules,

although all biomacromolecules are formed mainly by covalently linked carbon bonds (bond energy $\sim 85 \text{ kcal mol}^{-1}$), and covalent interconversion of enzyme cascades, which possess an enormous capacity for signal and rate amplification, play important roles in regulating cell signaling and metabolism (Chock and Stadtman, 1996). Specifically, noncovalent interactions are involved in stabilizing the double helix structure of DNA, in orchestrating RNAs interactions to exert their transcriptional and translational regulation, and in the folding of proteins/enzymes into intricate three-dimensional structures to accommodate their enzymatic activities and substrate specificity. The four extensively utilized noncovalent interactions are electrostatic interactions, hydrogen bonding interactions, van der Waals interaction, and hydrophobic interactions. They differ in their nature of their interactions, bond strength and specificity as follows: (1) Electrostatic interactions, resulting from the Coulombic interaction between two opposite atomic charges located on two molecules. The energy of this interaction is governed by Coulomb's law, namely, $E = kq_1q_2/Dr$, where E is the energy, q_1 and q_2 are the charges on the two atoms, at r angstroms apart, D is the dielectric constant of the medium and k is a proportionality constant, which has a value of 332 when the energy is expressed in kcal mol^{-1} . By convention, an attractive interaction would have a negative energy. In an aqueous medium, $D=80$, when $r=3 \text{ \AA}$ between the two ions bear single opposite charges, the electrostatic interaction energy is $-1.4 \text{ kcal mol}^{-1}$. However, this interaction is significantly strengthened when it occurs on the protein surface ($\sim -4.8 \text{ kcal mol}^{-1}$) or in the interior of the protein ($\sim -60 \text{ kcal mol}^{-1}$) due to changes in the dielectric constant of the reaction medium. (2) Hydrogen bonds, where the hydrogen atom is partially shared by two electronegative atoms such as nitrogen or oxygen. This leads to a favorable dipole-dipole interaction. The strength of this interaction falls off quickly with distance, or when the angle between the dipole is far from linear. Hydrogen bonds between the protein back bone amide nitrogen and carboxyl oxygen play a major role in stabilizing the α -helix and β -sheet structure of proteins, as well as determining the conformation of proteins. The energy for hydrogen bonds is in the range of $\sim 1\text{--}5 \text{ kcal mol}^{-1}$. (3) van der Waals interactions, derived from weak attractions that occur between atoms in close proximity to each other. The basis of these interactions is the attraction of the positively charged nucleus and negatively charged electron clouds between different atoms. For atoms commonly found in biological molecules, van der Waals attractions are optimal at distances between 3 and 4 $\text{\AA}$, and become negligible beyond 5 $\text{\AA}$. The van der Waals repulsion prevents atoms getting much closer than $\sim 3 \text{ \AA}$ apart. The energy associated with the van der Waals interaction is $\sim 0.5\text{--}1 \text{ kcal mol}^{-1}$, depending on the van der Waals distance. (4) Hydrophobic interactions are due to the tendency of nonpolar molecules to stick together in an aqueous medium. Nonpolar molecules do not easily dissolve in water, in part due to their inability to participate in hydrogen bonding or ionic interactions with water and to restructure the hydrogen bonding among water molecules. The poor solubility is governed by a large entropy reduction. A model, generally known as the iceberg model, has been proposed to provide a molecular interpretation for the large entropy loss to the structural

enhancement of water molecules near the vicinity of a nonpolar solute. However, the validity of this model is still a matter for debate, in view of recent experimental and theoretical analysis that reveals that water does not form a structure around the nonpolar solute at room temperature and a large part of negative entropy of solvation can be attributed to the small size of water molecules such that the nonpolar solute can interact with high number of water molecules and leads to a large decrease in entropy. For details on this issue, please see reference (Blokzijl and Engberts, 1993; Lee, 1985). Nevertheless, hydrophobic interactions are well accepted as the dominant energetic factor to mediate the formation of protein tertiary structure, enzyme-substrate/effector interaction and the stability of biological membranes.

While each of these noncovalent interactions is relatively weak, collectively they determine the biological structures and functions of proteins, nucleic acids, lipids, and carbohydrates. It is essential to note that every ion in an aqueous medium is surrounded by a shell of oriented water molecules maintained by the attraction of water dipoles to the charged ion. Thus, hydration of ions has a major influence on all aspects of electrostatic interactions for which the strength of acids and bases plays an important role. Since proteins contain multiple acidic and basic groups, it is reasonable to expect both the conformation and activity of enzymes to be altered as a function of the concentration of hydrogen ions.

Effect of Molecular Crowding in Living Cells

The biochemical and biophysical principles discussed in preceding sections were derived mainly from *in vitro* studies using pure reactants, often small molecules, or in relatively low concentrations, for example, less than 1 mg ml^{-1} of total macromolecule such as protein, nucleic acids, or polysaccharides. However, all cells contain various biomacromolecules at high concentration (Hoppert and Mayer, 1999). Biochemical reactions in living cells occur in media crowded with other soluble or structured macromolecules, resulting in nonspecific interactions between individual macromolecules and their immediate surroundings in the cytosol. These background interactions lead to three different phenomena: (1) macromolecular crowding, attributed to volume excluded by one soluble macromolecule to another; (2) macromolecular confinement, attributed to steric-repulsive interactions between the macromolecule of interest and its static boundaries; and (3) macromolecular adsorption, due to reversible association of a macromolecule to the surface of a nearby fiber or membrane (Minton, 2006). Nonspecific interactions may either be repulsive, leading to preferential size and shape dependent exclusion, or attractive, leading to nonspecific binding or adsorption. Predominantly repulsive background interactions tend to enhance the rate and extent of macromolecular association in solution, while predominantly attractive background interactions tend to enhance the tendency of macromolecules to cluster nonspecifically or adsorb onto surfaces. However, in a complex and heterogeneous medium of the cytoplasm, it is a challenge to discern whether the locally dominant background interactions are likely to be

attractive or repulsive and to identify their effects on any specific reaction.

Molecular crowding, in principle, can markedly slow down the diffusion rate. Consequently crowding plays a role in all biological processes mediated by noncovalent associations or conformational changes of the macromolecules, such as those involved in the synthesis of nucleic acids and proteins, intermediary metabolism and cell signaling, as well as the functioning of dynamic motile systems. In general, macromolecular crowding nonspecifically enhances reactions leading to the reduction of total excluded volume, independent of hydrogen bondings, van der Waals forces or charges. These reactions include the formation of macromolecular complexes in the medium, binding of macromolecules to surface binding sites, formation of insoluble aggregates, as well as compaction or folding of proteins. Simple statistical-thermodynamic modeling studies reveal that the 'passive crowding macromolecules' could exert order-of-magnitude or greater changes in both the rates and equilibria of numerous reactions tested. To this end, one should also recognize that system studies via simulation are still models instead of the real thing. Not all idiosyncratic details of the model system are of general value for understanding the real cellular system. Biological systems are more complex than theoretical or *in vitro* experimental studies because of enhanced heterogeneity and the presence of nonspecific repulsive and attractive intermolecular interactions in addition to volume exclusion. Model studies also show that the magnitude of the effects is strongly dependent on the relative sizes and shapes of the concentrated crowding species used and on the nature of the macromolecular reactants and products. However, to date, the results obtained via model simulation studies have provided important new insights for understanding the subject. In view of the complexity and heterogeneity of the intracellular fluids, results from simplified model studies can only partially address the complex problems encountered with the *in vivo* system (Zhou *et al.*, 2008; Ellis, 2001).

The densely packed environment in the cytosol appears to impede the folding of relatively large polypeptides since their diffusion rates would be more drastically reduced relative to those of smaller polypeptides. Furthermore, the presence of a large number of crowding macromolecules would increase the probability for a newly synthesized polypeptide to interact with other macromolecules before it can properly fold. To overcome these problems, nature makes use of a class of molecular chaperones as well as a number of protein disulfide isomerases to facilitate proper folding of nascent proteins, including those mediated by cysteine disulfide bond formation, to yield functional proteins. As a result, proteins are found to be folded very efficiently when synthesized inside the cell. Anfinsen showed that it took several hours for ribonuclease-A to fold in the test tube, a rate much slower than the rate at which functional ribonuclease-A is produced in cells (about 2 min.). A similar rationale was adopted in cell signaling. To facilitate cell signaling processes inside crowded cytosols, scaffold or anchorage proteins are adopted to generate signalsomes to process cell signaling. To this end, formation of intracellular Dishevelled-based signalsomes has been demonstrated to occur during the activation of Wnt signaling (Yokoyama *et al.*, 2010).

Concluding Remarks

To understand the molecular mechanisms of biological processes through which living cells stay alive, grow, reproduce, and evolve, we must understand how fundamental chemical and physical principles govern these reactions. In this article, a brief discussion is provided of (1) the chemical nature of cellular macromolecules, (2) the roles of water molecules and noncovalent interactions in stabilizing the reactive conformations of these macromolecules, (3) thermodynamic principles which determine whether reactions can occur spontaneously, and (4) principles of reaction kinetics in maintaining reaction processes at dynamic steady states, away from reaction equilibrium. For a more in-depth understanding of these principles, readers are referred to additional literature (Edsall and Gutfreund, 1983; Alberty, 2003; Moore and Pearson, 1981; Purich, 2010; Connors, 1990; Zhou *et al.*, 2008; Ellis, 2001; Yokoyama *et al.*, 2010; Atkins and de Paula, 2012). Finally, the potential effects of molecular crowding inside the cell were discussed in terms of principles derived from *in vitro* studies with dilute purified macromolecules and high substrate concentrations.

References

- Adair, G.S., 1925. The hemoglobin system: VI. The oxygen dissociation curve of hemoglobin. *Journal of Biological Chemistry* 63, 529–545.
- Alberty, R.A., 2003. *Thermodynamics of Biochemical Reactions*. Hoboken, NJ: Wiley Interscience.
- Atkins, P.W., de Paula, J., 2012. *Physical Chemistry for the Life Sciences*, second ed. New York, NY: W. H. Freeman & Company.
- Blockzijl, W., Engberts, J.B.F.N., 1993. Hydrophobic effects. *Opinions and facts. Angewandte Chemie International Edition* 32, 1545–1579.
- Boyd, R.K., 1978. Some common oversimplifications in teaching chemical kinetics. *Journal of Chemical Education* 55, 84–89.
- Cha, S., 1970. Kinetics behavior at high enzyme concentrations: Magnitude of errors of Michaelis-Menten and other approximations. *Journal of Biological Chemistry* 245, 4814–4818.
- Chock, P., Stadtman, Earl R., 1996. Cyclic cascades in cellular regulation. In: Bittar, E.E., Bittar, N. (Eds.), *Principles of Medical Biology*, Vol. 4, Cell Chemistry and Physiology: Part II. Greenwich, CT: Jai Press, pp. 201–220.
- Cleland, W.W., 1963. The kinetics of enzyme-catalyzed reactions with two or more substrates or products: I. Nomenclature and rate equations. *Biochimica et Biophysica Acta* 67, 104–137. II. Inhibition: Nomenclature and theory. *Biochimica et Biophysica Acta* 67, 173–187; III. Prediction of initial velocity and inhibition patterns by inspection. *Biochimica et Biophysica Acta* 67, 188–196.
- Connors, K.A., 1990. *Chemical Kinetics. The Study of Reaction Rates in Solution*. New York, NY: VCH Publishers, Inc.
- Edsall, J.T., Gutfreund, H., 1983. *Biothermodynamics: The Study of Biochemical Processes at Equilibrium*. New York, NY: John Wiley & Sons, Inc.
- Ellis, J., 2001. Macromolecular crowding: Obvious but underappreciated. *Trends in Biochemical Sciences* 26, 597–604.
- Gibbs, J.W., 1876–1878. On the equilibrium of heterogeneous substances. *Transaction of the Connecticut Academy of Arts & Sciences* vol III, 108–248, 343–524.
- Gibbs, J.W., 1878. Abstract on the "equilibrium of heterogeneous substances". *American Journal of Sciences Ser vol XVI* (3), 441–458.
- van Holde, K.E., Johnson, C., Ho, P.S., 1998. *Principles of Physical Biochemistry*. Upper Saddle River, NJ: Prentice-Hall, Inc.
- Hoppert, M., Mayer, F., 1999. Prokaryotes. *American Scientist* 87, 518–525.
- Huang, C.Y., Rhee, S.G., Chock, P.B., 1982. Subunit cooperation and enzymatic catalysis. In: Snell, E.E., Boyer, P.D., Meister, A., Richardson, C.C. (Eds.), *Annual Review of Biochemistry*, vol. 51. Palo Alto, CA: Annual Reviews Inc., pp. 935–971.
- Koshland Jr., D.E., Nemethy, G., Filmer, D., 1966. Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochemistry* 5, 365–385.

- Lee, B., 1985. The physical origin of the low solubility of nonpolar solutes in water. *Biopolymers* 24, 813–823.
- Liu, K., Cruzan, J.D., Saykally, R.J., 1996. Water clusters. *Science* 271, 929–933.
- Minton, A.P., 2006. How do biochemical reactions within cells differ from those in test tubes? *Journal of Cell Science* 119, 2863–2869.
- Monod, J., Wyman, J., Changeux, J.-P., 1965. On the nature of allosteric transition: A plausible model. *Journal of Molecular Biology* 12, 88–118.
- Moore, J.W., Pearson, R.G., 1981. *Kinetics and Mechanism*, third ed. New York, NY: John Wiley.
- Pauling, L., 1935. The oxygen equilibrium of hemoglobin and its structural interpretation. *Proceedings of the National Academy of Sciences* 21, 186–191.
- Purich, D.L., 2010. *Enzyme kinetics: Catalysis & Control. A Reference of Theory and Best-Practice Methods*. London: Academic Press.
- Viratelle, O.M., Seydoux, F.J., 1975. Pseudoconservative transition: A two-state model for the co-operative behavior of oligomeric proteins. *Journal Molecular Biology* 92, 193–205.
- Wong, J.T., 1975. *Kinetics of Enzyme Mechanisms*. New York, NY: Academic Press.
- Yokoyama, N., Golebiewska, U., Wang, H.-Y., Malbon, C.C., 2010. Wnt-dependent assembly of supermolecular dishevelled-3-based complexes. *Journal of Cell Science* 123, 3693–3702.
- Zhou, H.-X., Rivas, G., Minton, A.P., 2008. Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annual Review of Biophysics* 37, 375–397.

Biocatalysis

PA Frey and GH Reed, The University of Wisconsin–Madison, Madison, WI, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Biological Catalysts

Enzymes and ribozymes are essential to life. These macromolecules catalyze a vast array of chemical reactions that, in the absence of the biocatalyst, would take place at very low intrinsic and uncoordinated rates incompatible with life. Biocatalysts control the complex chemistry of thousands of life processes through acceleration and regulation of the rates of virtually every chemical reaction important to life. Prevention of unwanted chemical reactions relies on keeping reactive molecules apart through compartmentalization or on binding of the chemicals to macromolecules that stabilize their desired forms.

The magnitudes of rate enhancement brought about by biocatalysts approach astronomical values. For example, at 25 °C uncatalyzed hydrolysis of phosphodiester linkages between nucleotides in DNA proceeds with a half-life ($t_{1/2}$) of 30 000 000 years (Schroeder *et al.*, 2006). The refractory nature of phosphodiester linkages between nucleotides in DNA permits preservation of genetic information for hundreds of thousands, if not millions, of years. Yet biological processes at times require certain phosphodiester linkages in DNA to be cleaved. Staphylococcal nuclease, one of the enzymes that catalyzes DNA-hydrolysis, functions with a turnover number of 95 s^{-1} , corresponding to a $t_{1/2}$ of 7 ms. This enormous rate difference, or enzymatic enhancement factor, of 1.4×10^{17} corresponds to typical values for enzymes, and is neither the minimum nor the maximum. Nonenzymatic counterparts of certain enzymatic reactions are too slow to measure, so that only lower limits of rate enhancements for those reactions can be estimated.

In life, individual biocatalytic rates must be coordinated, and they must change under varying cellular developmental, metabolic, and environmental conditions. Coordination arises through several mechanisms. In genetic regulation, the relative amounts of enzymes produced by gene transcription and translation can be regulated to produce the appropriate amounts of metabolically or developmentally related enzymes. In metabolic control, metabolites regulate the activities of key enzymes through allosteric effects upon binding to regulatory sites of key enzymes. For example, in end-product inhibition the enzyme catalyzing the first committed step in a metabolic or biosynthetic pathway is often inhibited by the end product of that pathway, thereby shutting it down once the end product rises to an optimal concentration. The intra-organellar microenvironments in cells can activate or inhibit certain enzymes. For example, the acidic environments in lysosomes activate pH-dependent proteolytic enzymes. Posttranslational modifications of enzymes such as reversible phosphorylation of specific amino acid side chains are also important control mechanisms. These regulatory processes orchestrate the best balance of enzymatic activities to support the life of an organism.

Enzymes

Composition

Most biocatalysts are proteins known as enzymes. As proteins, they are linear polypeptides composed of the 20 common α -amino acids (Table 1) linked by peptide amide bonds. They range in molecular weights from ~ 8000 to $> 150\,000$. The polypeptide chains of enzymes are folded into definite globular structures, many of which have been determined by X-ray crystallography or by nuclear magnetic resonance spectroscopy. The folded structures include secondary, periodic structures of α -helix and β -sheet segments, together with non-periodic loop-segments. More than 5000 'different' enzymes are currently covered by the Enzyme Commission, and more are discovered each year.

Figure 1 depicts the protein chain fold of homodimeric enolase, showing the α -helical, β -sheet segments, and loop segments. Enolase catalyzes the dehydration of 2-phosphoglycerate to phosphoenolpyruvate. The chain fold of the core of enolase is widely known as the TIM barrel, named for the first enzyme found to display this fold, triosephosphate isomerase. It is also known as the β -barrel. While many protein chain folds are known, the TIM barrel is highly versatile and serves many purposes in enzymology. About 10% of known enzyme structures are based on the TIM barrel.

Enzymes exist as single polypeptides or as aggregates of identical or different subunits. Multisubunit enzymes can be dimers, trimers, tetramers, pentamers, hexamers, etc. of identical subunits. Some enzymes have nonidentical subunits, in which the subunits are either different enzymes catalyzing related reactions or regulatory entities that control catalytic rates. The largest aggregated enzymes include several enzymes acting in concert to catalyze processive reactions. Examples include α -ketoacid dehydrogenase complexes, which contain three enzymes plus regulatory subunits and are approximately the size of ribosomes. Other examples are the fatty acid synthase complexes, which contain six enzymes and produce fatty acids from acetyl coenzyme A in assembly-line fashion. Polyketide synthases and nonribosomal polypeptide synthases are analogous, aggregated enzymes that produce antibiotics such as erythromycin and tyrocidine.

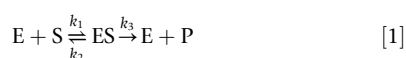
Kinetics

Activities of typical enzymes can be measured as initial rates under specified conditions. Dependence of initial rates produced by a fixed amount of an enzyme, acting on varying concentrations of the substrate, produce hyperbolic plots of rate against substrate concentrations. The plots are consistent with the Briggs–Haldane revision of the Michaelis–Menten kinetic mechanism and the corresponding rate law, eqns [1] and [2a], where E, S, and P represent enzyme, substrate, and product. In eqn [2a], $K_m = (k_2 + k_3)/k_1$ and $V_m = k_3[E_0]$. Note

Table 1 The α -amino acid structures

	R–		R–
Glycine	H–	Serine	HOCH ₂ –
Alanine	CH ₃ –	Cysteine	HSCH ₂ –
Valine	(CH ₃) ₂ CH–	Aspartic acid	HO ₂ C–CH ₂ –
Leucine	(CH ₃) ₂ CH ₂ CH–	Glutamic acid	HO ₂ C–CH ₂ CH ₂ –
Isoleucine	CH ₃ (CH ₃)CHCH ₂ –	Histidine	
Phenylalanine		Tyrosine	
Tryptophan		Lysine	H ₂ NCH ₂ CH ₂ CH ₂ CH ₂ –
Proline		Arginine	
Asparagine	H ₂ NCO–CH ₂ –	Cystine	
Glutamine	H ₂ NCO–CH ₂ CH ₂ –		
Threonine	CH ₃ (OH)CH–		

that the dissociation constant of the substrate, K_d , is defined as k_2/k_1 .



$$v = \frac{V_m[S]}{K_m + [S]} \quad [2a]$$

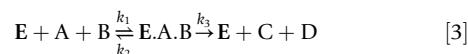
$$\frac{1}{v} = \frac{1}{V_m} + \frac{K_m}{V_m} \cdot \frac{1}{[S]} \quad [2b]$$

Equation [2a] predicts that when $[S] \ll K_m$, the rate is proportional to $[S]$, and when $[S] \gg K_m$ the rate approaches V_m , the maximum as shown in **Figure 2(a)**. The maximum rate in any experiment is proportional to the enzyme concentration $[E_0]$. The reciprocal form of eqn [2a] (eqn [2b]; **Figure 2(b)**) predicts a linear relationship between $1/v$ and $1/[S]$ having an intercept on the ordinate of $1/V_m$ and a slope of K_m/V_m . Manual fitting of kinetic data is facilitated by the linear form, whereas computer programs can be used to fit data directly to eqn [2a].

Equation [1] refers to an enzyme with one substrate and one product. Many complex enzymes act on two, three, or more substrates. In such cases, more complicated kinetic mechanisms apply, and the rate laws are correspondingly more complex, with more terms in the numerators and denominators than in eqn [2a]. However, for enzymes not subject to

special regulatory effects, the complex equations are reducible to the form of the Michaelis–Menten equation when all co-substrates are held constant and only one is varied. In these cases, the measured parameters are apparent values V_m^{app} and K_m^{app} , from which limiting values can be obtained by varying the concentrations of co-substrates in further analysis.

A full discussion of multisubstrate kinetics lies beyond the scope of this article. However, the complications can be appreciated by considering the general patterns for two-substrate, two-product reactions, the case of substrates A and B reacting to form products C and D. The most common case is that of eqn [3], where the two substrates become associated with the enzyme in stepwise fashion, either in random order or in compulsory order, to form the ternary complex E.A.B in eqn [3]. The substrates react within the ternary complex to form the products C and D, which dissociate in either random or ordered steps. Kinetic mechanisms of this type are known as sequential mechanisms. The rate eqn [4] describes the initial rate kinetics of the rapid equilibrium version of the sequential mechanism.



$$v = \frac{V_m[A][B]}{K_{iA}K_B + K_B[A] + K_A[B] + [A][B]} \quad [4]$$

In eqn [4], $K_{iA} = (E)(A)/(EA)$ is the dissociation constant of A, $K_B = (EA)(B)/(EAB)$, and $K_A = (EB)(A)/(EAB)$. In an initial rate

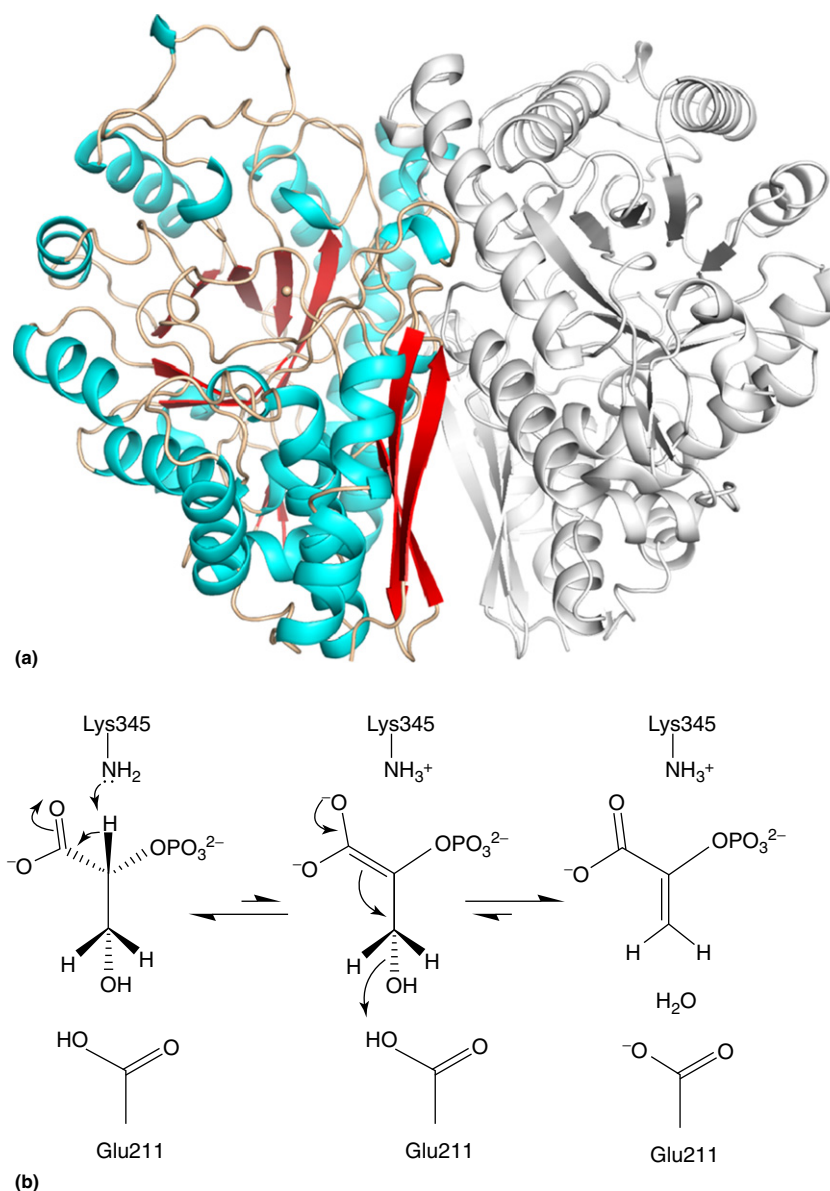
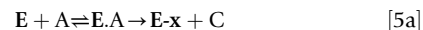


Figure 1 The structure and mechanism of enolase. (a) Shows a cartoon representation of the molecular structure of homodimeric enolase from baker's yeast. Secondary structural elements are color coded in one subunit of the dimer: cyan α -helical; red β -sheet; wheat loops. The tightly bound Mg^{2+} in each subunit appears as a sphere. The eight-stranded barrel is atypical in having one strand pointed in the opposite direction. The structure was obtained from PDB file 1ONE (Larsen *et al.*, 1996). (b) Shows the stepwise mechanism by which enolase catalyzes the dehydration of (*R*)-2-phosphoglycerate, showing the acid–base catalysis by Glu211 and Lys345 and the intermediate formation of the enolate (Poyner *et al.*, 1996).

study with $[\text{B}]$ held constant and $[\text{A}]$ varied, the data can be treated as in the one-substrate case.

Another kinetic mechanism frequently observed is that described in eqns [5a] and [5b]. In this mechanism substrate A binds and reacts chemically with the active site, transforming it in a distinctive way to a form designated as E-x and releasing product C. The transformation can frequently be ligation of a group ($-\text{x}$) from the substrate to the enzyme. Alternatively, reducing equivalents can be transferred to a cofactor (see below) at the active site. The chemically modified enzyme then reacts with the second substrate B, transferring the group ($-\text{x}$) to form product D. Kinetic schemes of this type are frequently

called ping pong mechanisms.



$$v = \frac{V_m [\text{A}][\text{B}]}{K_A [\text{B}] + K_B [\text{A}] + [\text{A}][\text{B}]} \quad [6]$$

The rate equation for the ping pong mechanism (eqn [6]) lacks the constant term, $K_A K_B$, in the denominator of eqn [4].

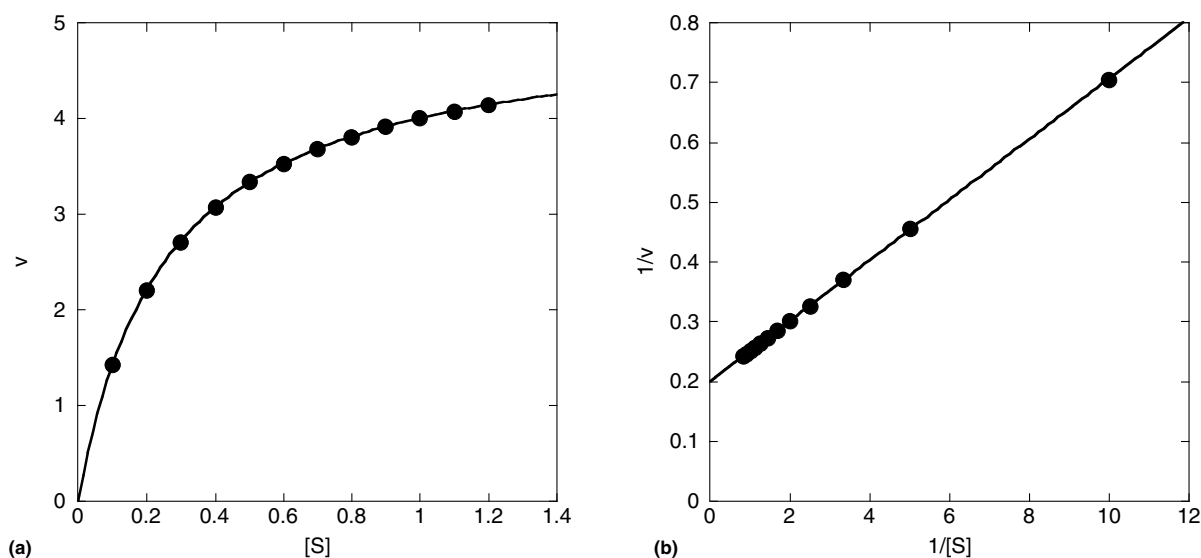


Figure 2 Reaction kinetics for an enzymatic reaction having a single substrate and single product. (a) Shows the direct plot of velocity, v , versus substrate concentration, $[S]$, according to eqn [2a] (see text) using V_m of 5 (arbitrary units) and a K_m of 0.25 (arbitrary units). (b) Shows the same data plotted in double reciprocal form according to eqn [2b], where the ordinate intercept is $1/V_m$ and the slope is K_m/V_m .

Double reciprocal plots of data, $1/v$ versus $1/[B]$ at different concentrations of A appear as parallel lines, having a common slope and different intercepts. A particular property of this mechanism is that the intermediate E-x can be generated, isolated, and chemically characterized by excluding substrate B. This provides chemical information about the reaction mechanism and the function of the active site. A compilation of kinetic expressions and details of their applications are available (Segel, 1993; Cook and Cleland, 2007).

Whether a multisubstrate enzyme functions by a sequential or ping pong mechanism depends on several factors, including the type of chemical reaction. Comparisons of the mechanisms of chemically similar phosphotransfer reaction, where -x is a phosphoryl or nucleotidyl group, have led to the principle of economy in the evolution of binding sites as a governing factor (Frey, 1992). The central fact that phosphotransferring enzymes function by ping pong mechanisms when the phospho-accepting substrates are sterically and electrostatically related inspired the postulation of this principle. The kinetic and chemical mechanisms of adenylate kinase (AK) and nucleoside diphosphate kinase (NucDipK) exemplify this principle.



AK functions by a sequential mechanism, whereas NucDipK functions by a ping pong mechanism, in which -x is the phosphoryl group (PO_3) bonded to a histidine residue. Overall, the reactions are chemically similar, phosphoryl transfer between phosphogroups. In the action of NucDipK, the sterically and electrostatically similar phosphoacceptors, NDP going forward and ADP in reverse, occupy the same binding site. Thus, the enzyme needs just one binding site, incorporating the phospho-accepting histidine residue.

In the reaction of AK two binding sites are required. The phosphoacceptors, AMP going forward and ADP in reverse, are sterically and electrostatically incompatible with a single binding site. Therefore, AK has a phosphodonor site and a phosphoacceptor sites. This allows ternary complex formation (E.ATP.AMP) and avoids the need for an E- PO_3 to bind the phosphoryl group during the interchange of phosphoacceptors.

The principle of economy in the evolution of binding sites appears to govern a number of other classes of enzymatic reaction mechanisms, where single binding sites function in double-duty fashion (Frey, 1992; Grove et al., 2011).

Substrate Selectivity

The specificity of enzymes for their substrates is frequently overstated. 'Highly selective' is a more apt description of their capacity to discriminate among similar molecules. Most enzymes will act on molecules related to their natural substrates, but at lower rates. An enzyme often functions on alternative substrates at 1/100th or 1/1000th the rate for the natural substrate, and such a rate could still correspond to a rate enhancement of, for example, 10^8 – 10^{12} . Metabolism of drugs and other xenobiotic compounds typically exploits the capacity of enzymes to function with alternative substrates.

Cooperativity and Allosteric Regulation

In classical behavior, each active site in a multi-subunit enzyme acts independently of neighboring subunits. Multi-subunit enzymes may also exhibit cooperative behavior among subunits in binding of substrates and in catalysis such that binding of a substrate to one subunit influences the subsequent binding of substrate to neighboring subunits. This cooperative behavior, either negative or positive, may also be modulated by effector molecules that bind to distinct regulatory sites in the oligomer termed allosteric sites.

Cooperativity and associated allosteric control provide a means to ‘fine tuning’ of catalytic activity to meet changing demands of cells and organisms.

Chemical Mechanisms

Rates of chemical reactions can be described in terms of transition state theory also known as absolute reaction rate theory. This formalism accounts for the temperature dependence of thermally activated chemical reactions. Descriptions of photochemical and electron transfer reactions are slightly different. For thermally activated reactions, a tiny fraction of the reactant(s) (determined by the Boltzmann distribution law) initially in their ground vibrational state(s), acquire from their surrounding, sufficient excess internal (vibrational) energy to ascend transiently to a high energy state, called the transition state or activated complex. In a reaction coordinate diagram, the transition state is located at the apex – a position at which the molecule or complex can descend either in the reverse direction to reactant(s) or in the forward direction to product(s). **Figure 3** illustrates the relevant energies for the reaction of a substrate S to a product P by way of $S^\ddagger$, the transition state or activated complex. The activation energy for the spontaneous reaction is E_a^S .

In the corresponding enzymatic reaction, **Figure 3** shows the energy of activation E_a^{ES} for the $E.S$ complex. The ground state for the $E.S$ complex is shown as slightly lower than for the spontaneous reaction. The difference $E_a^S - E_a^{ES} = \Delta E_a$ represents the decrease in activation energy for the enzymatic relative to the nonenzymatic reaction. This difference represents the magnitude of enzymatic catalysis, or the rate enhancement. Because activation energies have an inverse exponential influence on rates (Boltzmann distribution of activated complexes)

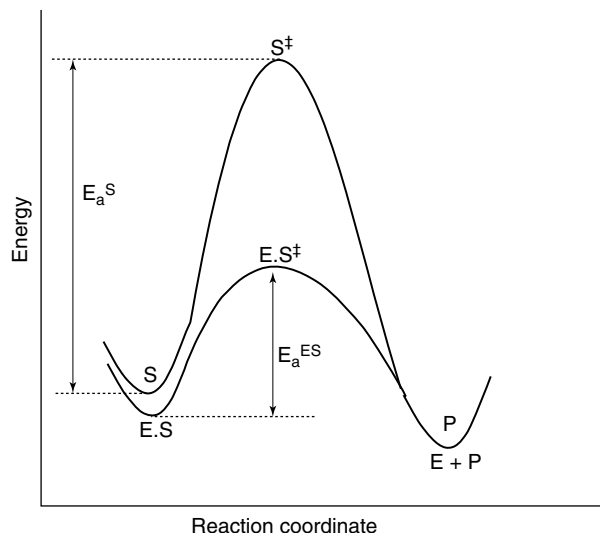


Figure 3 Relative potential energy curves along the reaction path for chemical reactions. The upper curve shows the potential energy along a reaction coordinate for a nonenzymatic reaction, showing the energy of the transition state or activated complex $S^\ddagger$ at the maximum. The activation energy is E_a^S . The lower curve depicts the potential-reaction coordinate for the same reaction catalyzed by an enzyme. The activation energy is E_a^{ES} .

modest changes in ΔE_a translate to large changes in rate. Although kinetic measurements can detect small changes in the activation energy, the general rule is that the smallest free energy difference that can be conceptualized is ~ 1 kcal, which corresponds to a rate difference of approximately sevenfold. When functioning as catalysts at concentrations much less than those of their respective substrates, enzymes do not alter the thermodynamic equilibria of the reactions that they catalyze but rather alter only the rate at which equilibrium is attained.

Rate enhancement by enzymes begins with the binding of substrate(s) at the pre-organized active site, where the substrate(s) is or are bound in close proximity to catalytic groups within the active site. For enzymes that do not require a cofactor or coenzyme, the catalytic groups are one or more of the reactive amino acid side chains in the right-hand column of **Table 1**. Each of these groups, with the exception of cystine, displays both acid/base and nucleophilic reactivities. In different enzymes, each group functions either as a nucleophilic catalyst or as an acid/base catalyst. For example, in the reaction of enolase (**Figure 1**) the elimination of water from 2-phosphoglycerate is facilitated by acid–base catalysis by Lys345 and Glu211. In the action of NucDipK, phosphotransfer is facilitated by a histidine residue and nucleophilic catalysis (see above).

Transition State Analogs and Catalytic Antibodies

A prediction of transition state theory is that a substantial part of the catalytic rate enhancement of enzymes is due to a tight binding of the transition state of the respective reaction. Molecules that mimic the transition state of an enzyme-catalyzed reaction (transition state analogs) are expected to bind to the active site with high affinity (Wolfenden, 1972). A substantial fraction of drugs are molecules that bind to enzymes in place of the normal substrates and thereby inhibit the activity of the target enzyme. The transition state analog approach to design of selective tight binding inhibitors of target enzymes has been practiced with substantial success (Schramm, 2007).

Another practical application of transition state mimics is in the generation of antibodies that have catalytic activities for specific reactions. Antibodies arise in the blood serum of animals in the process of immunization, which entails exposure to foreign proteins, antigens, for example, the coat proteins of bacteria or viruses. Antibodies are produced in specialized mammalian cells and released into the blood stream. They recognize and mark antigen molecules for destruction by binding them very tightly ($K_d \sim 10^{-14}$ M). Antibodies are large protein molecules, in which a smaller domain, the Fab fragment, encompasses the antigen-binding site. An antigen-binding site recognizes and binds segments known as epitopes of antigenic proteins. An epitope might be a decapeptidyl unit within a foreign protein. However, antigen-binding sites are not limited to polypeptides as ligands and may bind other molecules as well. Thus, immunization of an animal with a foreign protein chemically linked to a dinitrophenyl (DNP) group, a hapten, will lead to the production of anti-DNP antibodies. The tight binding between an antigen and an antibody can be exploited to create a catalytic antibody. In this method, a molecule embodying a stable,

structural analog of the transition state for a desired reaction can be used as a hapten to elicit the production of an antibody that binds the transition state analog tightly. In theory, such tight binding should lead to catalysis of a reaction with a transition state structurally similar to that of the hapten. The antibody is then a catalytic antibody. The emergence of methods to generate homogeneous populations (clones) of antibody molecules has enabled the production and characterization of monoclonal antibodies having homogeneous antigen-binding sites and properties (Shokat and Schultz, 1990). Thus, haptens mimicking transition states of chemical reactions have been used to generate monoclonal antibodies having high affinities for these transition state mimics (Padioulet-Lefevre *et al.*, 2014). Rate enhancements achieved with catalytic antibodies are frequently modest (10^2 – 10^4); however, rate enhancements of 10^{10} have been reported (Xu *et al.*, 2004). Several classes of chemical reactions have been investigated using the catalytic antibody approach.

Coenzymes and Cofactors

Enzymatic reactions are so diverse that the side-chain functional groups are not chemically sufficient, by themselves, to catalyze all enzymatic reactions. Processes such as electron transfer, oxidation/reduction, oxygenation, decarboxylation of α -amino acids and α -ketoacids, dehydrogenation, complex group transfers, and free radical reactions, among others, require special chemistry. Thus, the activities of many enzymes require co-catalysts known as cofactors or coenzymes.

The coenzymes and cofactors include metal ions, activated vitamins, and small organic molecules shown in Figure 4(a). These molecules exhibit special chemical properties that are essential in some enzymatic reactions. A few of the co-catalysts, including TTQ, TPQ, and MIO arise from posttranslational modifications of amino acid side chains of the enzyme itself. Figure 4(a) includes activated forms of vitamins such as niacin (NAD^+), thiamine (ThDP), riboflavin (FAD, FMN), vitamin B₆ (PLP), pantothenic acid (phosphopantetheine), and biotin. The vitamin coenzymes play essential roles in many oxidation–reduction, decarboxylation, carboxylation, racemization, elimination, and carbon–carbon bond cleavage reactions. The npn-vitamin derived cofactors PQQ, TTQ, TPQ, lipoamide, and MIO play essential chemical roles in dehydrogenases and oxidoreductases.

Figure 4(b) shows the structures of a number of metallic cofactors, including heme, which functions in oxygenation and electron transfer reactions. AdoCbl, the coenzyme form of vitamin B₁₂, and AdoMet-[4Fe-4S] function similarly to initiate free radical chemistry in many biological reactions. Iron–sulfur clusters are important in many electron transfer roles. The oxo-Fe₂ complexes facilitate oxygenation and dehydrogenation in several enzymes. Molybdopterin and tungstopterin complexes are essential in a number of oxidases and reductases. In dinitrogenase, the FeMo coenzyme reduces N₂ to 2NH₃ in an ATP-dependent process. In certain organisms, vanadium replaces molybdenum whenever the latter is unavailable. The nickel complex F₄₃₀ anchors the reductive process of methanogenesis.

Metal ions involved in enzymatic catalysis include magnesium, iron, copper, manganese, zinc, cobalt, calcium, nickel,

and potassium. Mononuclear complexes of iron or copper in many enzymes react with O₂ and introduce hydroxyl groups into organic substrate molecules. In enolase two Mg²⁺ ligate the carboxylate group of 2-phosphoglycerate. The resultant positive charge increases the acidity 2-phosphoglycerate and allows Lys345 to abstract C2(H) to give the enolate intermediate in the dehydration reaction (Larsen *et al.*, 1996). Mg²⁺ complexes of nucleoside di- and triphosphates are the actual substrates of the multitude of reactions of enzymes operating on these molecules. Metalloproteases use either mononuclear or dinuclear complexes of Zn²⁺ or Co²⁺ in catalyzing hydrolysis of peptide bonds. Many phosphatases use ions of metals such as Zn, Mg, Ca, Fe singly or in combinations in their hydrolytic reactions.

Ribozymes

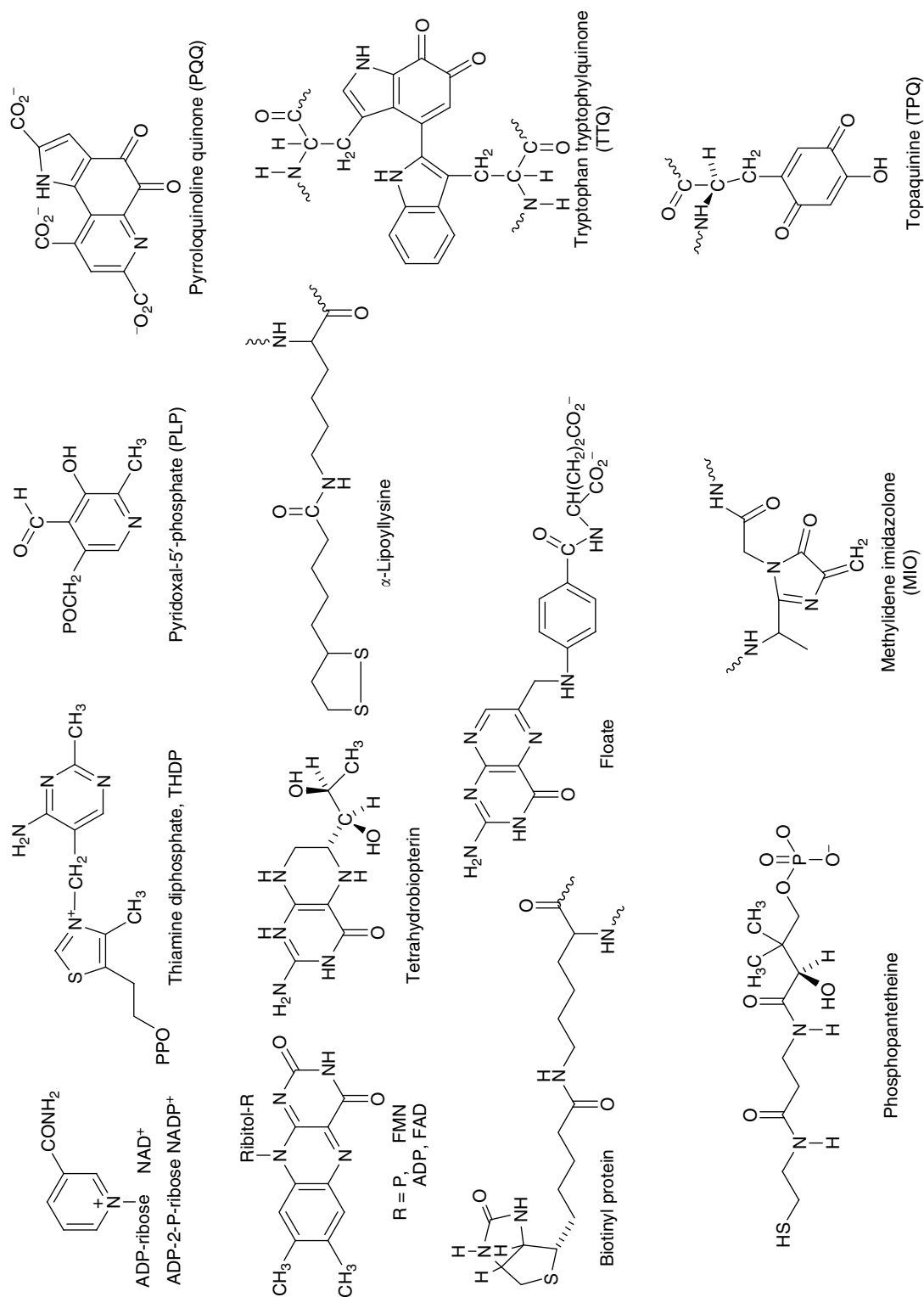
Ribozymes are RNA enzymes. They catalyze the chemical processing of RNA and of the information contained within the nucleotide sequences of RNA. Ribozyme function includes the excision of introns from the initial mRNA transcripts of genes encoded by DNA. Ribozyme function also includes peptide bond formation – the final step in the translation of DNA in protein biosynthesis.

Hammerhead

The ribozymes range in size from about 50 ribonucleotides to the species of RNA in the small and large subunits of the ribosome that comprise up to a few thousand ribonucleotides. The hammerhead ribozyme belongs to the small ribozyme family. It catalyzes the self-cleavage of a precursor RNA and the joining of resultant RNA fragments to produce the hammerhead ribozyme. The name hammerhead arises from the appearance of an earlier secondary structural representation that resembled the outline of a hammerhead shark. Figure 5(a) shows the secondary structure of a generic hammerhead ribozyme. The loops and helices fold into a tertiary structure. The hammerhead functions in rolling circle replication.

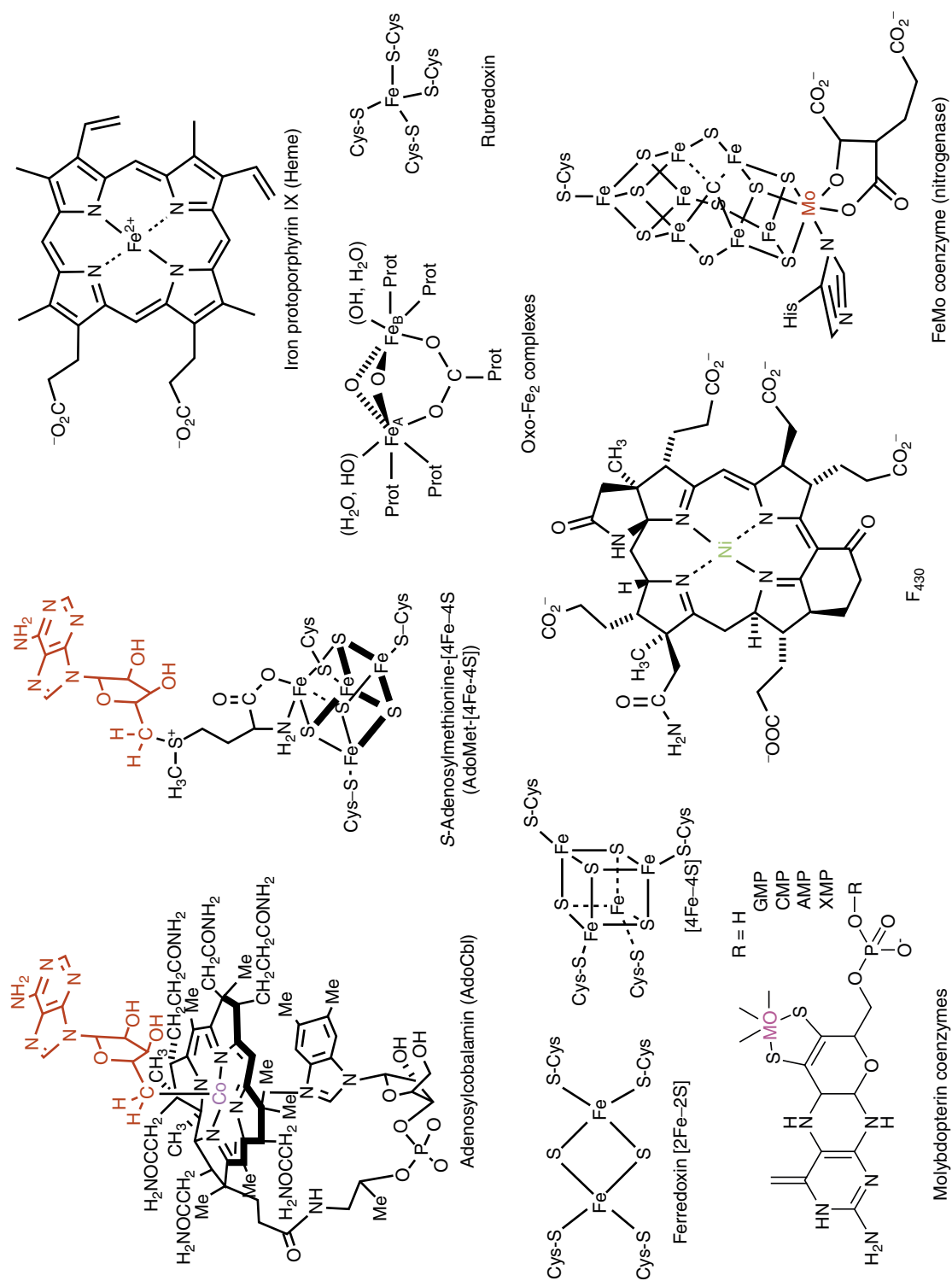
The hammerhead, like many ribozymes, catalyzes the cleavage of specific 3',5'-phosphodiester linkages within RNA. The initial cleavage produces smaller species of RNA in which one product contains, 2',3'-cyclic-CMP at the 3'-end, as shown in Figure 5(b). The other cleavage product has a 5'-OH end. The reaction proceeds with a turnover number of $\sim 1 \text{ min}^{-1}$. As indicated in Figure 5(b), the reaction is reversible because the 2',3'-cyclic-phosphodiester is an activated species that reacts favorably with the 5'-OH group to generate a 3',5'-phosphodiester. This reversibility allows rejoining activity in RNA processing.

The RNA cleavage reaction in Figure 5(b) involves an acid, A–H, to donate a proton to the 5'-O leaving group and a base, :B, to abstract a proton from the C2'(OH) group of the cytidyl nucleotide. This mechanism takes advantage of the most biologically and chemically significant property of RNA, the susceptibility of RNA to base catalyzed hydrolysis. The hydroxide ion alone is sufficient to cleave RNA by abstraction of the proton from C2'(OH). Even the low concentration of hydroxide in neutral solutions (10^{-7} M) makes RNA too unstable to serve as the repository of genetic information, and



(a)

Figure 4 Structures of representative coenzymes and cofactors. (a) Shows structures of representative organic coenzymes and cofactors. (b) Shows structures of representative metallocofactors. Functions of these cofactors can be found in [Frey and Hegeman \(2007\)](#).



(b)

Figure 4 Continued.

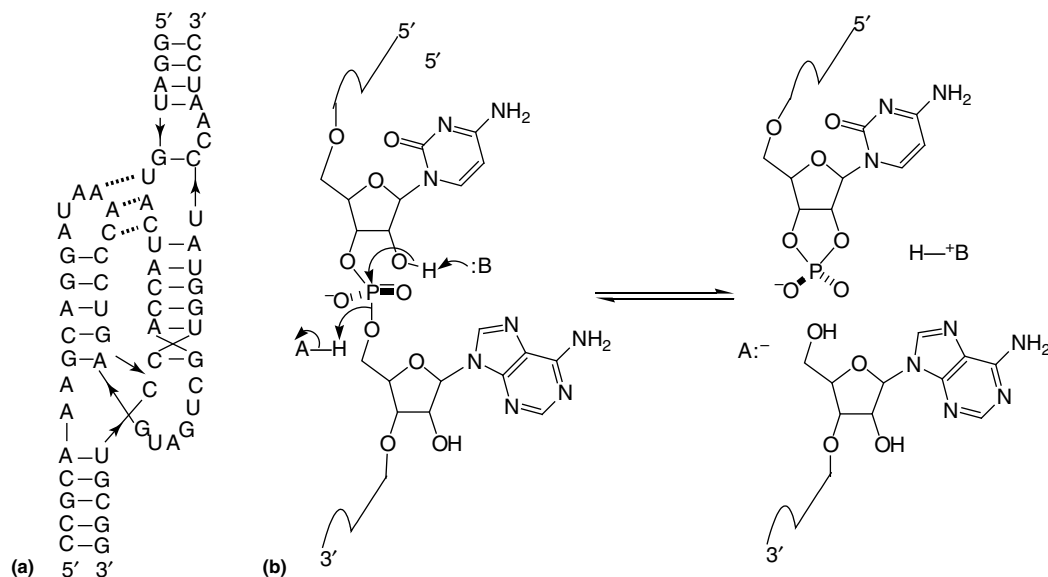


Figure 5 The hammerhead ribozyme. (a) Shows the secondary structure of a typical hammerhead ribozyme (Martick and Scott, 2006). The large arrow indicates the cleavage site. Small arrows indicate chain connectivity. Solid lines indicate Watson–Crick pairs. Dashed lines indicate non Watson–Crick tertiary interactions. (b) Shows the chemistry of the strand cleavage catalyzed by the hammerhead ribozyme, where A–H represents an acid and :B represents a base.

this process also contributes to the high mutability of the RNA genome in RNA viruses. In ribozymes, acid and base groups within the structure accelerate the cleavage. In the hammerhead, G12 and G8 are believed to function as acid and base. Other small ribozymes function similarly to the hammerhead.

Self-Splicing Introns

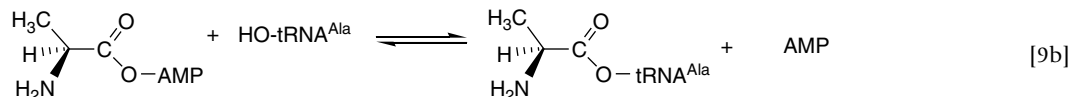
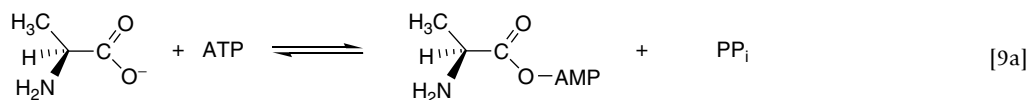
Ribozymes also include the self-splicing introns in mRNA transcripts. These are much larger than small ribozymes and carry out the essential function of excising introns from transcribed mRNA prior to translation in protein biosynthesis. The essential chemistry is that shown in Figure 5(b). The intron internally cleaves itself at each end, forming two 2',3'-cyclic-nucleotide ends and two 5'-OH ends. Cleavages are sequential and not simultaneous, but overall they lead to a 2',3'-cyclic nucleotide at the 3'-terminus of one exon, and a 5'-OH at the 5'-end of the neighboring exon. Then, the reversibility shown in Figure 5(b) allows these two ends to anneal. The intron itself can also anneal to a circular structure. The overall process is highly ordered and proceeds with complex tertiary structural rearrangements.

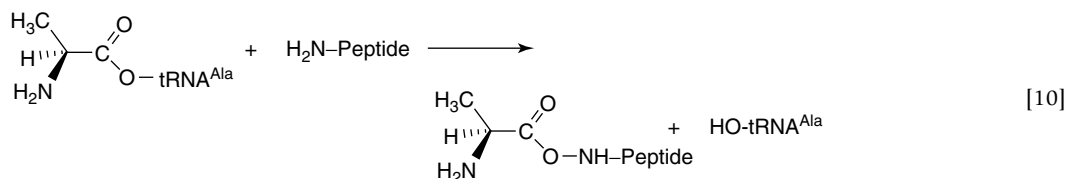
RNaseP

RNaseP is a ribonucleoprotein, in which the RNA component catalyzes the hydrolysis of RNA. The mechanism of 3',5'-phosphodiester cleavage is similar to Figure 5(b), but the 2',3'-phosphodiester product is further cleaved by water to the phosphomonoester. This process is analogous to the mechanism by which the protein enzyme ribonuclease catalyzes the hydrolysis of RNA. This ribozyme functions catalytically, like the protein enzyme.

The Ribosome

Translation of a mature mRNA produces a polypeptide chain that folds into a protein. The protein is either functional or undergoes posttranslational modification to become functional. In most cells, translation requires 20 amino acids, 20 aminoacyl-tRNA synthetase enzymes, 20 species of tRNA, ancillary protein based factors, and the ribosome. The ribosome catalyzes the final step of peptide bond formation. The process for a single amino acid is illustrated for alanine in eqns [9] and [10]. Alanyl-tRNA synthetase catalyzes the reactions shown in eqns [9a] and [9b], activating alanine for peptide bond for-





mation by initially forming alanyl adenylate (eqn [9a]), a carboxylic-phosphoric anhydride. In the second step, eqn [9b], alanyl adenylate reacts with the 2'-OH group of alanyl-tRNA (HO-tRNA^{Ala}) to form alanyl-tRNA^{Ala} and releasing AMP. The activated alanyl group in alanyl-tRNA^{Ala} is selected by the complex of mRNA-ribosome to react with the amino-terminal end of a growing polypeptide chain extending it and releasing HO-tRNA^{Ala}. The complementation of the anticodon in the tRNA of alanyl-tRNA^{Ala} with the codon for Ala in mRNA governs the selection alanyl-tRNA^{Ala} in eqn [10], while the ribosome catalyzes the chemistry in eqn [10]. The codons in mRNA govern the selection of the other nineteen species of aminoacyl-tRNA^{AA} to generate an accurate, translated amino acid sequence.

A ribosome consists of small and large species of ribosomal RNA (rRNA), as well as more than 20 protein subunits that stabilize the structure. The rRNA functions as a ribozyme (Cech, 2000) in catalyzing polypeptide biosynthesis by step-wise transfer of aminoacyl groups from the 20 species of aminoacyl-tRNA to the amino-terminal end of the growing polypeptide chain in the exact order specified by the codons in mature mRNA. This word picture is fully supported by the structure of the 70 S ribosome (Selmer *et al.*, 2006).

See also: Molecular Principles, Components, Technology, and Concepts: Basic Principles: Chemical and Physical Principles

References

- Cech, T., 2000. Structural biology. The ribosome is a ribozyme. *Science* 289, 878–879.
- Cook, P.F., Cleland, W.W., 2007. *Enzyme Kinetics and Mechanism*. London and New York: Garland Science.
- Frey, P.A., 1992. Nucleotidyltransferases and phosphotransferases: Stereochemistry and covalent intermediates. In: Boyer, P.D. (Ed.), *The Enzymes*, third ed., vol XX. New York, NY: Academic Press, pp. 141–186.

- Frey, P.A., Hegeman, A.D., 2007. *Enzymatic Reaction Mechanisms*. New York, NY: Oxford University Press (Chapters 3 and 4).
- Grove, T.L., Randle, M.I., Krebs, C., Booker, S.J., 2011. Cfr and RlmN contain a single [4Fe–4S] cluster, which directs two distinct reactivities for *S*-adenosylmethionine: Methyl transfer by SN2 displacement and radical generation. *Journal of the American Chemical Society* 133, 19586–19589.
- Larsen, T.M., Wedekind, J.E., Rayment, I., Reed, G.H., 1996. A carboxylate oxygen of the substrate bridges the magnesium ions at the active site of enolase: Structure of the yeast enzyme complexed with the equilibrium mixture 2-phosphoglycerate and phosphoenolpyruvate at 1.8 Å resolution. *Biochemistry* 35, 4349–4358.
- Martick, M., Scott, W.G., 2006. Tertiary contacts distant from the active site prime a ribozyme for catalysis. *Cell* 126, 309–320.
- Padioulet-Lefevre, S., Naya, R.B., Shahsavarian, M.A., Friboulet, A., Avallé, B., 2014. Catalytic antibodies and their applications in biotechnology: State of the art. *Biotechnology Letters* 36, 1369–1379.
- Poyner, R.R., Laughlin, L.T., Sowa, G., Reed, G.H., 1996. Towards identification of acid/base catalysts in the active site of enolase: Comparison of the properties of K345A, E168Q and E211Q variants. *Biochemistry* 35, 1692–1699.
- Schramm, V.L., 2007. Enzymatic transition state theory and, transition state analogue design. *Journal of Biological Chemistry* 282, 28297–28300.
- Schroeder, G.K., Lad, C., Wyman, P., Williams, N.H., Wolfenden, R., 2006. The time required for water attack at the phosphorus atom of simple phosphodiester and of DNA. *Proceedings of the National Academy of Sciences of the USA* 103, 4052–4055.
- Segel, I.H., 1993. *Enzyme Kinetics*. New York, NY: Wiley & Sons.
- Selmer, M., Dunham, C.M., Murphy 4th, R.V., *et al.*, 2006. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* 313, 1935–1942.
- Shokat, K.M., Schultz, P.G., 1990. Catalytic Antibodies. *Annual Review of Immunology* 8, 335–363.
- Wolfenden, R., 1972. Analog approaches to the structure of the transition state in enzyme reactions. *Accounts of Chemical Research* 5, 10–18.
- Xu, Y., Yamamoto, N., Janda, K.D., 2004. Catalytic antibodies: Hapten design strategies and screening methods. *Bioorganic & Medicinal Chemistry* 12, 5247–5268.

Relevant Websites

- www.brenda-enzymes.org/
BRENDA Enzyme Data Base.
- enzyme.expasy.org/
Enzyme Nomenclature Data Base.
- www.ebi.ac.uk/thornton-srv/databases/enzymes/
Enzyme Structure Data Base.

MOLECULAR PRINCIPLES, COMPONENTS, TECHNOLOGY, AND CONCEPTS: NUCLEIC ACIDS

Contents

DNA, RNA Chemical Properties (Including Sequencing and Next-Generation Sequencing)

The Chemical Synthesis of DNA and RNA Oligonucleotides for Drug Development and Synthetic Biology Applications

DNA, RNA Chemical Properties (Including Sequencing and Next-Generation Sequencing)

M Datto, Duke University Medical Center, Durham, NC, USA

RL Lundblad, Chapel Hill, NC, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Base-pairing The interaction between nucleobases responsible for the specificity of nucleic acid strand binding to other nucleic acid strands. Base-pairing occurs between adenine and thymine or cytosine with guanine in DNA; uridine replaces thymine in RNA and is base-paired with adenine.

Bridge amplification A method of preparing a polony based on the ligation of two adapter sequences surrounding a library insert. The adapter insert will hybridize to a oligonucleotide on a surface in flow cell. The adapter serves as primary for the synthesis of a daughter strand which will hybridize to another oligonucleotide via the second adapter which contains a cleavage site. Denaturation will remove the parent strand leaving the daughter strand which can then be amplified to yield a polony.

cDNA A single-stranded DNA complementary to an RNA synthesized *in vitro* by reverse transcription using a DNA polymerase usually obtained from a retrovirus.

Contig A major sequence of DNA which has been assembled from the sequences of many smaller DNA fragments; assumed to be a contraction of contiguous.

DNA distortions A conformational change in the regulator helical structure of DNA which can result from a chemical inserting into the helical structure, a process referred to as intercalation.

Hyperchromicity An increase in optical absorbance such as that seen with the denaturation of DNA.

Library A collection of cloned fragments which represent the genome. DNA sequence is obtained from fragments obtained from the library.

Long read A sequence of 500–1000 bases.

Long-noncoding RNA Large RNA molecules which are mostly bound to protein and serve a regulatory function.

Long read A sequence of 500–1000 bases.

MicroRNA Small RNA molecules which can modulate gene expression by influencing mRNA stability functioning as small interfering RNA (siRNA); can be referred to as small-noncoding RNA.

Nucleic acid footprinting A method of identifying regions of DNA which are exposed, not bound to protein, or not available because of conformation. This process can involve chemical modification or availability of the DNA to enzymatic cleavage.

Nucleobase An aromatic heterocyclic compound, a purine or pyrimidine, that is responsible for specificity of recognition.

Nucleoside The compound, a glycoside, formed by a covalent bond from a nucleobase to a ribose or deoxyribose ring at the C₁ position. Nucleoside analogues are useful drugs.

Nucleotide A nucleoside with a covalently bound phosphate group which is the monomer unit in nucleic acids. Nucleotide structures are also found in coenzymes such as nicotinamide adenine dinucleotide (NAD).

Polony A 'colony' of DNA molecules obtained by the replication of a single DNA template under conditions where the PCR product remains physically close to the DNA template.

Polymerase chain reaction A method by which a single strand of DNA is amplified to provide a large, homogeneous DNA sample. The method use a repetitive temperature cycle using a thermostable DNA polymerase to extend primer which binds to a known DNA sequence flanking the region of interest. Twenty cycles has the potential to provide 2²⁰ amplification. The polymerase chain reaction has a variety of applications including the preparation of samples for DNA sequence analysis.

Primer A short DNA or RNA oligonucleotide sequence with a free 3'-hydroxyl group that binds to a region on the DNA and serves as the nidus to start copy of a complementary strand.

Primer extension A technique used to detect modifications in RNA polymer. Primer extension is based on the reverse transcriptase extension of a synthetic DNA oligonucleotide complementary to the target RNA. Extension of the cDNA is compromised by the presence of a cleavage in the target RNA or a modified nucleobase.

Short reads Base sequence data of 25–100 bases which is most frequently obtained with second-generation sequencers.

Reverse transcriptase An enzyme which catalyzes the formation of a DNA strand from an RNA template.

RNAseq The use of next-generation DNA sequence technology to determine RNA structure.

Tautomerism A form of isomerization where the two forms are directly interconvertible as in the case of enol–keto isomerization. This differs from isomers such as propyl and isopropyl which are not interconvertible.

Template A single-stranded DNA or RNA, usually with a primer region attached, which serves to direct the sequence of synthesis of a complementary polynucleotide.

Introduction

The term nucleic acid was used to by R. Altmann in 1889 (cited in [Esanu, 1988](#)) to describe material obtained from salmon sperm and calf thymus. Clear recognition that there are two types of nucleic acids dates to the work of P.A. Levine and others more than 90 years ago ([Levine and London, 1929](#)). This work demonstrated that sugar content of thymonucleic acid was deoxyribose (deoxypentose) with adenine, guanine, cytosine, and thymine as the nucleobase components ([Figure 1](#)). Hammarsten ([Hammarsten, 1894](#)) had previously defined ribonucleic acid as containing pentose (ribose) with adenine, guanine, cytosine, and uracil as nucleobase

components ([Figure 1](#)). It has been subsequently shown that 5-methyl cytosine is present at a level of 1% ([Tost, 2009](#)) reflecting its importance in the regulation of gene expression ([Kulis and Esteller, 2010](#)).

As an indication of its large size, genomic DNA is expressed in weight (pg) and the size of the human genome (7 pg) is estimated to be 3.1 gigabases (3.1×10^9 bases) ([Doležal and Greilhuber, 2010](#)). Messenger RNA is smaller and is composed of a noncoding leader sequence (a 5'-UTR), a region coding for the primary structure of the protein and a 3'-UTR ([Mortimer et al., 2014](#)). Long-noncoding RNA are longer than 200 base pairs (bp) and structurally similar to messenger RNA ([Yan and Wang, 2012](#)). There are small RNA species which have been the

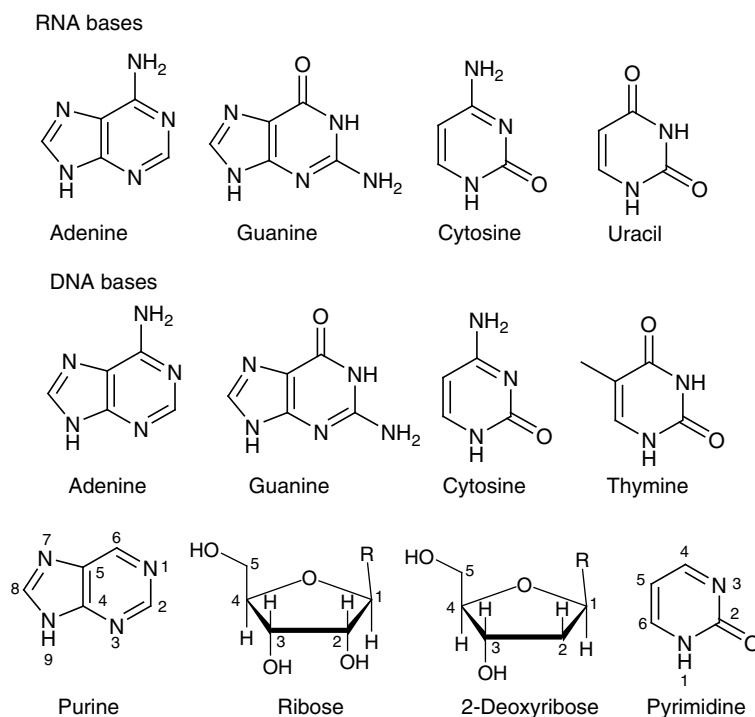


Figure 1 The structure of the common bases found in nucleic acids. Also shown is ribose (RNA) and deoxyribose (DNA). The ring numbering system for purine bases, pyrimidine bases, and the ribose moiety is shown.

subject of study in the past two decades (Cech and Steitz, 2014). Both DNA and RNA are polyanions reflecting the presence of a phosphoryl group. The negative charge on DNA is part of the basis for interaction with basic proteins such as histones although other factors are involved (Hadnagy *et al.*, 2008).

The function of nucleic acids is based on the complementary interaction of bases via hydrogen bonding as established by the Watson–Crick Model for DNA (Watson and Crick, 1953). Nucleobase are electron-rich heterocyclic aromatics which may contain exocyclic functional groups such as amino groups and hydroxyl groups which can exist in tautomeric forms (Figure 2). As a point of historical interest, it was when work on base-pairing with the enol forms was discarded in favor of the keto forms that it was possible for Watson to visualize the structure (Watson, 1968). Tautomerization can cause mistakes in base-pairing creating mutations (Goodman, 1995).

Nucleobases are joined to either ribose (RNA) or deoxyribose to form nucleosides (Figure 1). Nucleosides are converted to nucleotides by phosphorylation of the 5'-hydroxyl group. Nucleotide triphosphates are the substrates for the formation of DNA and RNA with the release of pyrophosphate and the formation of a phosphodiester bond through the 3' hydroxyl group on one ribose/deoxyribose and the 5' hydroxyl group on the adjacent ribose/deoxyribose ring. The development of nucleoside analogues has resulted in useful therapeutics for a variety of diseases including cancer, hepatitis B, and HIV (Figure 3). Nucleoside analogous can be cytotoxic, antiviral, or immunosuppressive in their action by several different mechanisms including direct enzyme inhibition, incorporation into either DNA or RNA resulting in loss of function, or inhibition of DNA synthesis.

Physical Structure of Nucleic Acid

The solution structure of nucleic acids is dominated by interaction with the various nucleobases. The double-helix of DNA is of great cultural fame (Watson, 1968) and of great biological importance. The genome is composed of double-stranded DNA which is held together by hydrogen bond and stacking of the nucleobases. The strands of DNA can be dissociated by denaturation (heat or base). The resulting single-stranded material obtained by heat denaturation will reform the original double-helical configuration if allowed to slowly cool; rapid cooling results in the formation of single-stranded random coils. The process of DNA denaturation/strand separation by heating is referred to as melting and is associated with an increase in UV absorbance at 260 nm. This increase in absorbance results from the 'destacking' of the nucleobases in the DNA. A plot of absorbance at 260 nm versus temperature, the DNA melting curve, is usually sigmoidal and the midpoint is designated T_m and is considered an attribute of the particular nucleic acid species (Figure 4). Fluorophores can bind to DNA and insert into the helical region in a process called intercalation; this binding is associated with an increase in fluorescence which is lost on strand separation (Figure 4). The T_m and the shape of the melting curve are related to DNA composition and the T_m increases with increasing GC content. Single-stranded DNA can renature with other DNA molecules in a process called hybridization. Hybridization can be used to measure the complementarity of single DNA molecules with the sensitivity of detecting single nucleotides polymorphisms (Lipsky *et al.*, 2001).

There are higher-order DNA complexes with triplexes (Fox and Brown, 2011) and quadriplexes (Tarsounas and Tijsterman, 2013). DNA molecules with catalytic activity,

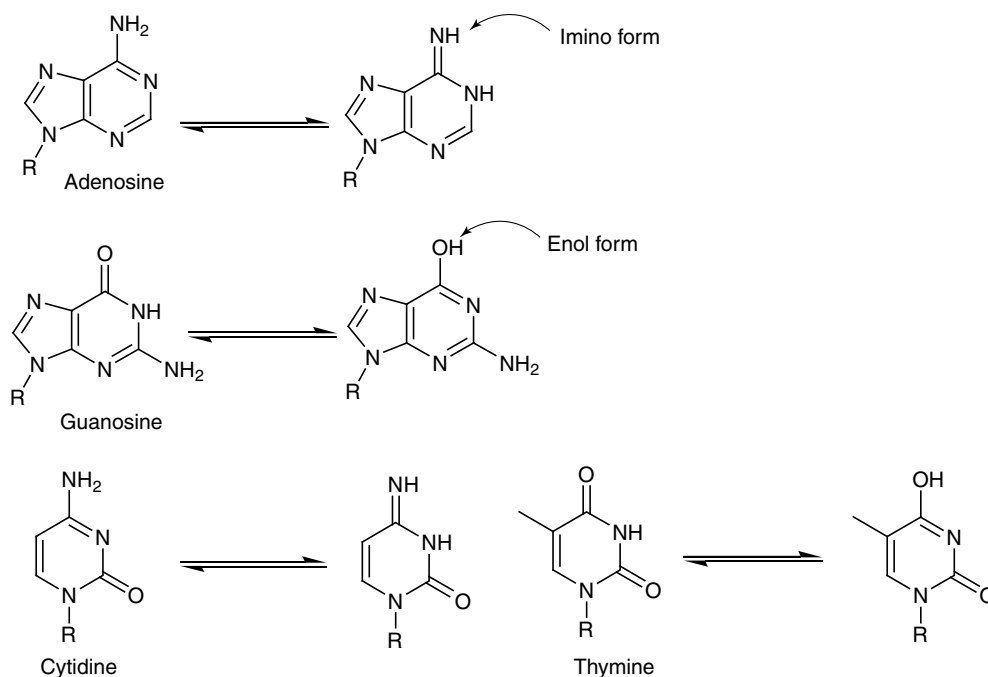


Figure 2 Tautomerization of nucleobases.

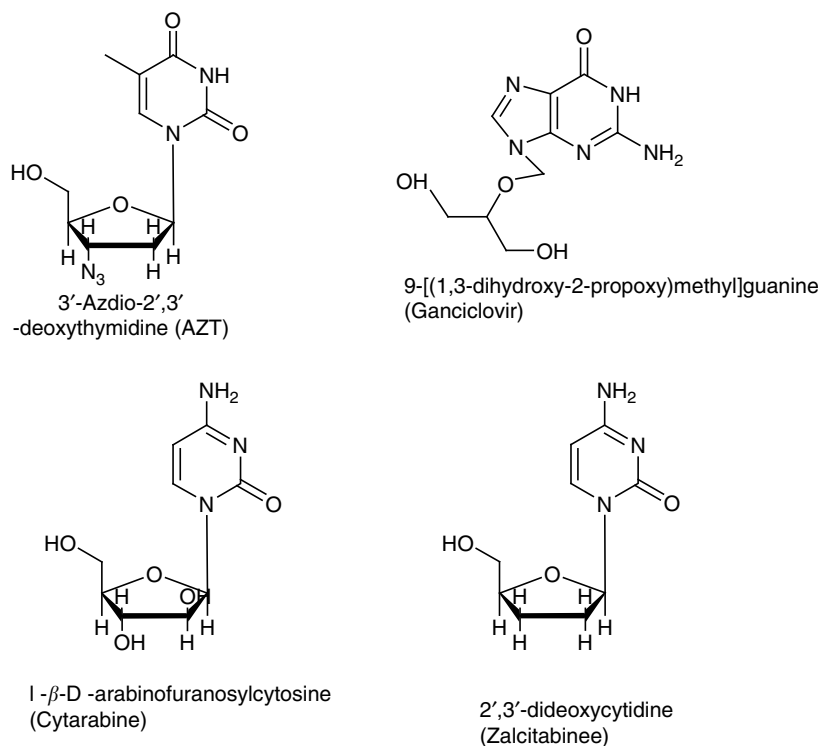


Figure 3 Nucleoside analogues as drugs.

deoxyribozymes, have been described (Faulhammer and Famulok, 1997). DNA aptamers are short, synthetic DNA molecules that bind to specific targets; aptamers are obtained by Systematic Evolution of Ligands by Exponential Enrichment (SELEX) techniques (Marimuthu *et al.*, 2012). DNA aptamers can be single-stranded or more complex such as G-quadruplex DNA aptamers (Tucker *et al.*, 2012).

RNA exists in a variety of forms including messenger RNA, transfer RNA (t-RNA) which carries amino acids to be incorporated in the growing polypeptide chain on ribosomes, and microRNA which are involved in the regulation of transcription. MicroRNA species are also used to 'knock down' gene expression (Cullen, 2006) and may well have therapeutic utility (Akagi *et al.*, 2014). The microRNA species are known as small interfering RNA (siRNA). These are relatively small, double-stranded RNA which binds to the target mRNA initiating degradation of the mRNA resulting in the termination of translation (Snead and Rossi, 2012).

Chemical Modification of Nucleic Acids

Early use of chemical modification formed the basis for the histochemical identification of nucleic acid components in tissues and the Feulgen reaction continues to be used for the histochemical analysis of DNA (Fleskens *et al.*, 2010). The Feulgen reaction is based on the acid-catalyzed removal of purine bases from the deoxyribose resulting in the formation of an aldehyde which can be detected with Schiff's reagent. The chemical modification of nucleic acids is not as complex as that of proteins since there are fewer monomer units and less

diversity of nucleophilic reactive group. Reaction at the primary amine groups of adenine is referred to as exocyclic modification, whereas reaction at the imine nitrogens or at carbon atoms in pyrimidine or purine rings is referred to as an endocyclic modification. There are also ring-opening reactions and cross-linking reactions involving the nucleic acid bases. Other than cleavage of the glycosidic bond, reaction at the ribose ring does not usually occur. Mortimer and Weeks (2007) did show that reaction at the 2'-hydroxyl group could occur (Figure 5) and used this modification to study RNA conformation by SHAPE (selective 2'-hydroxyl acylation analyzed by primer extension) chemistry.

Modification of nucleic acid with dimethyl sulfate (DMS) (Figure 5) was important for early DNA sequencing methods (Maxam and Gilbert, 1977). *N*-Methyl *N'*-nitro-*N*-nitrosoguanidine, DMS, and methyl methanesulfonate (Figure 5) are used to modify DNA for the study of DNA repair mechanisms (Wyatt and Pittman, 2006).

Diethylpyrocarbonate (Figure 6) reacts with adenine and guanine bases in nucleic acids resulting in rupture of the purine ring and subsequent cleavage of the glycosidic bond. Diethylpyrocarbonate was introduced into nucleic acid purification protocols as a general nuclease inhibitor. The finding that there was a reaction between nucleic acid and diethylpyrocarbonate prompted an investigation of the mechanism(s) of a possible reaction (Solymosy *et al.*, 1971). The conformational sensitivity of the reaction of diethylpyrocarbonate with nucleic acids is useful in determining DNA distortions (Kahl and Paule, 2009). Kethoxal (3-ethoxy-1,1-dihydroxy-3-butanone) (Figure 6) is specific for the modification of guanine bases in RNA (Litt, 1971) and is used for the study of the conformation of RNA in

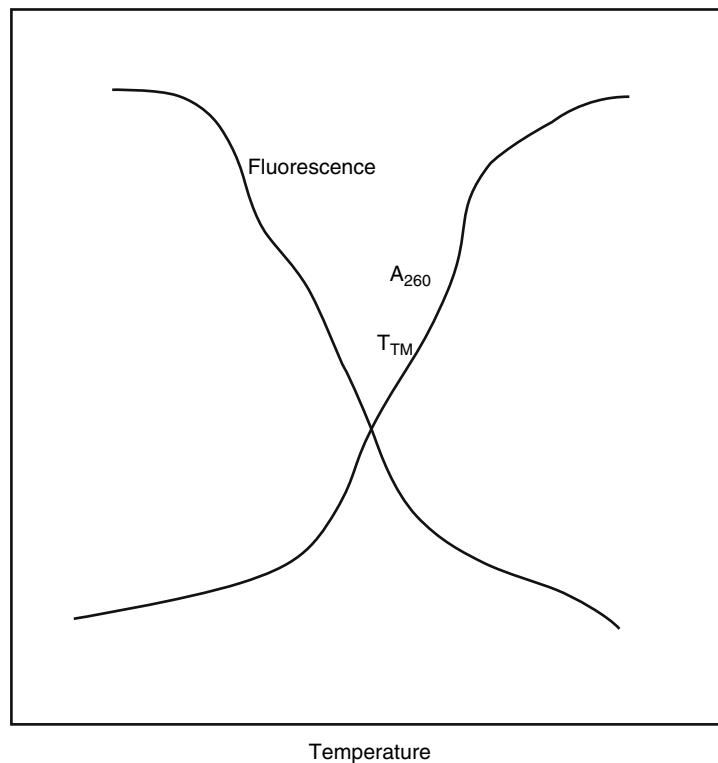
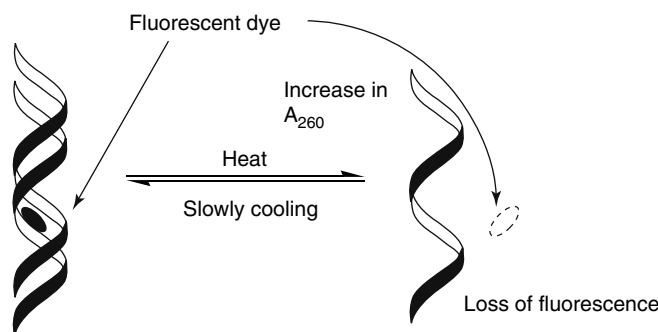


Figure 4 DNA melting curve.

macromolecular complexes (Kostopoulou *et al.*, 2013). Mortimer *et al.* (2009) suggest that reaction with (*N,N*-dimethylamino) dimethylchlorosilane (Figure 6) is a more accurate measure of the accessibility of guanine. Environmentally significant aldehydes such as acrolein and the closely related crotonaldehyde, acetaldehyde, and 4-hydroxy-2-nonenal react with nucleic acids. The products of the reaction of exogenous aldehydes with DNA are referred to as DNA adducts (Minko *et al.*, 2009). Permanganate oxidation has also been used to study DNA mutation by detection of mismatches in DNA (Kahl and Paule, 2009). The key to these various applications is the reaction of permanganate with an unpaired or mismatched thymine; in other words, the fully complementary site is resistant under mild reaction condition while the exposed or reactive site is highly susceptible to oxidation. Tabone *et al.* (2006) used the absorbance of the hypomanganese diester to develop a microplate assay for mutation scanning based on mismatching in base-pairing.

DNA Modification by Radiation

DNA is subject to modification by radiation. This damage included chemical modification of nucleobases and DNA strand cross-linking resulting from the absorption of light by the various nucleobases. Ionizing radiation can damage DNA by multiple mechanisms including the generation of free radicals from the ionization of water resulting in DNA strand cleavage (Peak and Peak, 1990). The production of hydroxyl radicals from water is also used for DNA footprinting (Franchet-Beuzit *et al.*, 1993). Unlike UV radiation, ionizing radiation can cause DNA damage at the sugar ring (von Sonntag, 2014). Irradiation of nucleic acids with UV light (254 nm) results in the hydration of pyrimidine rings. There is interest in the use of natural antioxidants in protecting against radiation, visible and UV, damage (Malhomme de la Roche *et al.*, 2010).

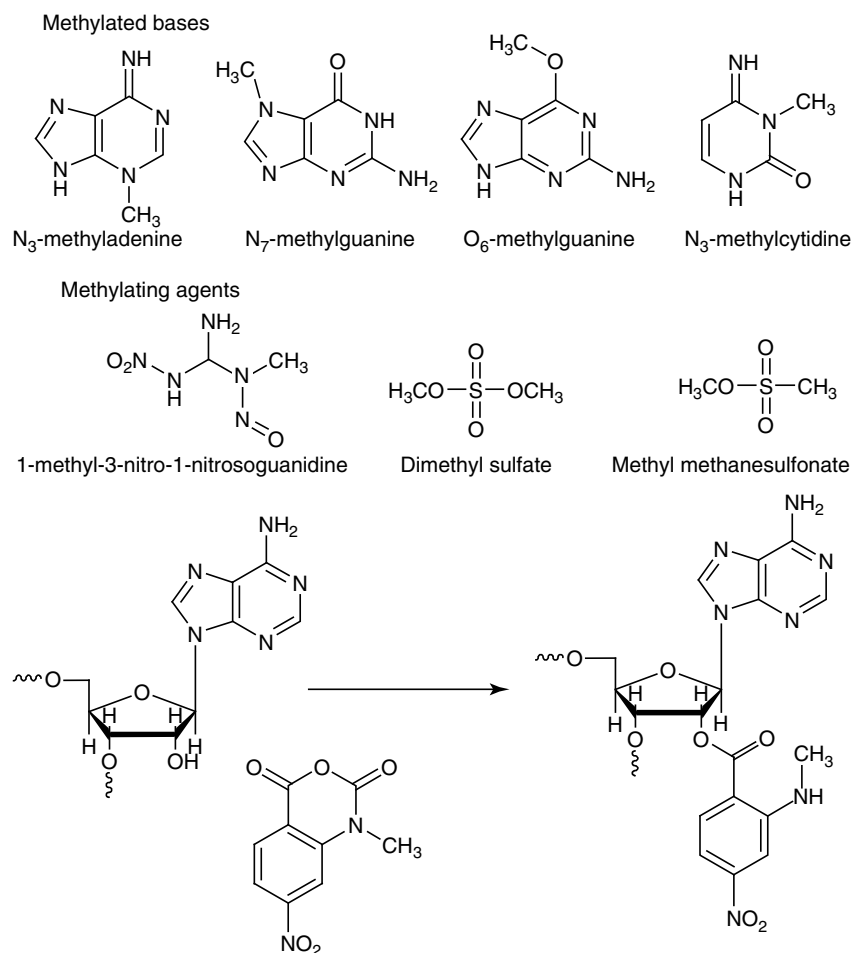


Figure 5 Chemical modification of nucleobases and nucleosides.

DNA Mapping

5-Methylcytosine (Adams, 1995) and its hydroxymethyl derivative (Pfeifer and Szabo, 2009; Figure 7) may be the best known epigenetic modification. Sodium bisulfite converts cytosine to uracil in a deamination reaction while the deamination reaction with 5-methylcytosine is two orders of magnitude slower permitting the identification of cytosine methylation sites in DNA. Sodium bisulfite does not distinguish between 5-methylcytosine and 5-hydroxymethylcytosine (Jin *et al.*, 2010) and other approaches are required (Nestor *et al.*, 2014). DMS is used for DNA mapping studies where there is preferential modification of guanine. DNA is susceptible to alkali-induced cleavage at methylated purine residues which permits use for genomic footprinting of DNA (Toth *et al.*, 1990). DMS is also used for footprinting of RNA where modification at adenine and cytidine can be determined by reverse transcriptase (RT) (primer extension) (Tijerina *et al.*, 2007). Footprinting with DNaseI or hydroxyl radical is used to map sites of protein interaction with DNA (Hampshire *et al.*, 2007).

Nucleic Acid Sequencing

Modern personalized medicine is built on our ability to determine the sequence of bases in a patient's DNA. Variants in

DNA sequence from the population 'normal' are at the root of many inherited and acquired disease processes. The diagnosis of inherited diseases by DNA sequencing is now considered a routine clinical practice. Our ability to interrogate a patient's genetic makeup has evolved over the past three decades from the sequencing of individual regions of selected genes to the routine sequencing of all protein coding regions in a patient's entire genome (Miller *et al.*, 2014). The use of DNA sequencing techniques, however, is not restricted to inherited disease diagnosis. DNA sequencing is used to understand how a patient may respond to a specific drug, what dose of a drug is most appropriate, or what side effects could be expected. Collectively, this is the field of pharmacogenomics. DNA sequencing is also routinely used in the context of cancer to find mutations in key tumor suppressor genes or oncogenes. These data are used not only to subtype or diagnose different neoplastic diseases, but also to develop patient specific prognoses. Finally, chemotherapeutics have been developed which target oncogenes (primarily tyrosine kinases) with specific activating DNA mutations. Sequencing of a tumor's DNA to detect these mutations to help guide therapy is now the standard of care. All of these advances stem from two key innovations: (1) development of the dideoxy method for DNA sequencing, referred to as Sanger sequencing and (2)

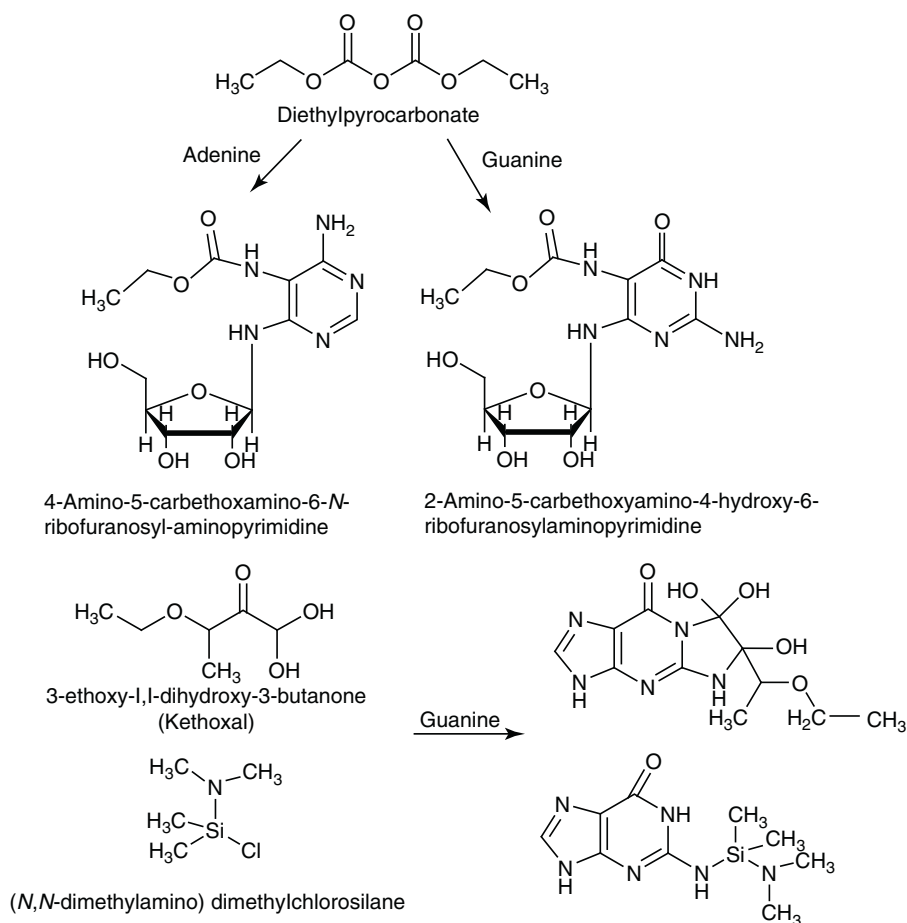


Figure 6 Diethylpyrocarbonate and Kethoxal modification of nucleobases.

development of massively parallel DNA sequencing through synthesis, a group of techniques collectively known as 'next-generation' sequencing (NGS). The current techniques used for DNA sequencing (both Sanger and NGS) exploit the machinery that exists in nature for DNA replication, specifically, DNA polymerases.

Sanger DNA Sequencing

Fred Sanger had already been awarded a Nobel Prize for work on the determination of protein sequence when he developed what is known as the dideoxy method or Sanger sequencing (Sanger *et al.*, 1977). This method is based on the ability of modified dideoxynucleotide (e.g., dideoxyadenine; Figure 8) to terminate DNA extension by a DNA polymerase.

The first step in Sanger sequencing is the selection of the DNA region to be sequenced. This could be a specific coding exon in the gene that causes cystic fibrosis (CFTR). This could be a DNA sequence within a bacterium that causes antibiotic resistance. The options are only limited to things with nucleic acids. In case of large DNA, the DNA fragment is broken into smaller pieces which are assembled into plasmids flanked by

known DNA sequences. Polymerase chain reaction (PCR) primers are designed to target the region to be sequenced. These PCR primers are short oligonucleotides (20–30 bp) and are the complement to DNA sequences that flank the 5' and 3' sides of the target DNA. Typically, PCR primers are designed to amplify several hundred base pairs in the target DNA sequence. PCR amplification is then performed to provide sufficient homogeneous sample for sequence analysis. Excess PCR primers and deoxynucleotides are removed. Sanger sequencing of the purified PCR amplification product is then performed. The Sanger method begins with a third oligonucleotide primer that is designed to be the complement to one end of the DNA region to be sequenced. This primer is referred to as the 'sequencing primer.' Like the PCR primers, the sequencing primer is a short synthetic oligonucleotide (20 or so base pairs long). This primer is added to the PCR amplification product, along with a heat stable polymerase, deoxynucleotides (dATP, dCTP, dGTP, dTTP) and a lower concentration of fluorescently labeled 3' dideoxynucleotides (ddATP, ddCTP, ddGTP, ddTTP). Each dideoxynucleotide is labeled with one of four different fluorophores. The mixture is heated and cooled to allow the sequencing primer to anneal to its target sequence. The polymerase then incorporates deoxynucleotides using the PCR

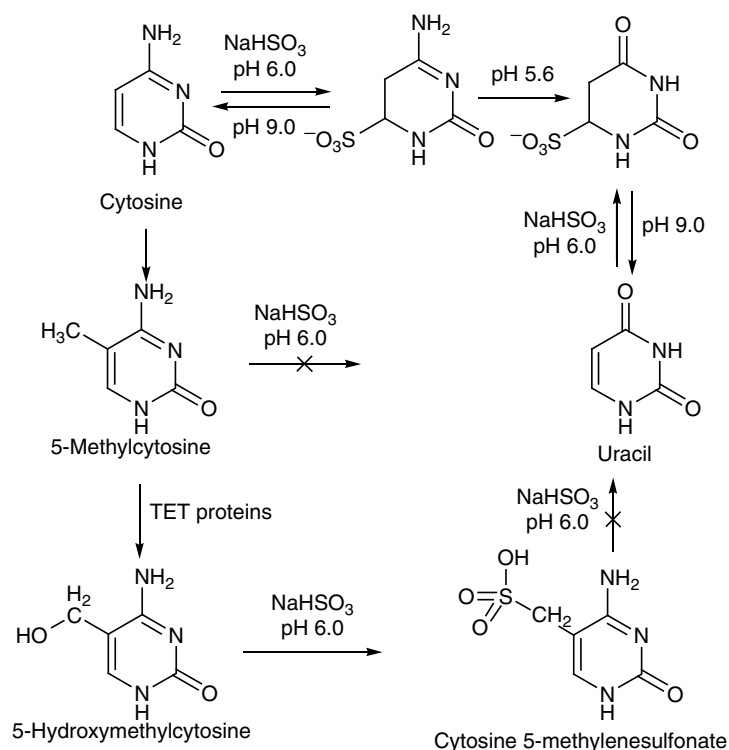
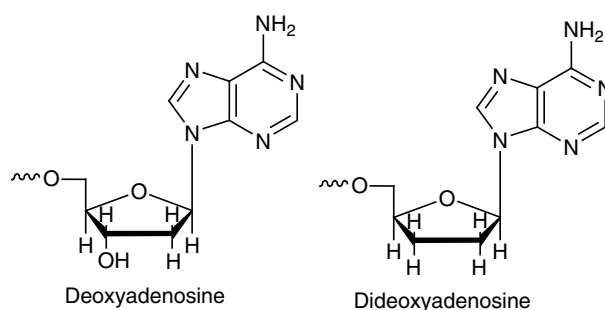


Figure 7 Reaction of methylcytosine with sodium bisulfite.

amplification product as a template until a dideoxynucleotide is incorporated. Without a 3' OH on the ribose group, further extension is not possible. The result is multiple DNA oligonucleotides, each ending with a dideoxynucleotide, which are the complement to the target sequence. Each of these resulting products is fluorescently labeled with one of four fluorophores, depending on which dideoxynucleotides terminated the growth of the nascent DNA chain. Using fluorescent capillary electrophoresis, the sequence of the original DNA can be inferred (**Figure 8**; Huang *et al.*, 1992).

NGS

NGS can be broadly defined as any DNA sequencing technology beyond Sanger sequencing. An excellent review of NGS sequencing methods is available (Metzker, 2010). NGS approaches have many similarities with Sanger sequencing. Both exploit the natural DNA replication machinery to derive DNA sequence (e.g., a DNA polymerase using a 'sequencing' primer and target sequence as a template). In addition, PCR amplification plays an important role in the preparation of the DNA target regions to be sequenced in both approaches. Two features are unique to NGS approaches: (1) they are applied to single DNA molecules, while Sanger sequencing looks at all DNA molecules in a 'sequencing reaction' in aggregate; (2) they generate real-time sequence data as each new nucleotide is incorporated into a nascent DNA strand. The steps involved in NGS can be broadly grouped into



Sanger sequencing

	5' → 3'							
	A	G	T	C	A	G	C	T
ddT	ddT				ddT			
ddC		ddC				ddC		
ddA			ddA				ddA	
ddG				ddG			ddG	
								Primer

Figure 8 Sanger or dideoxy sequencing.

three main parts: (1) library preparation, (2) sequencing through DNA synthesis, and (3) data analysis.

Library Preparation

As with Sanger sequencing, the first step in NGS is the selection and amplification of the target DNA regions to be

sequenced. In Sanger sequencing, a defined region of less than 500–600 bases is selected. In next-generation approaches, there is no limit to the size or complexity of the target DNA regions. One could select a single exon of a single gene to PCR amplification and purification, as in the Sanger sequencing (example above). Alternatively one could develop a multiplex PCR approach and amplify all exons for a particular gene, or all exons for multiple genes. Instead of selecting regions by PCR, one could employ a physical capture method. In these approaches, short oligonucleotide sequences are physically immobilized onto a solid support surface. These oligonucleotides are the complement to the target sequences to be ‘captured.’ Extracted DNA is physically sheared or enzymatically cleaved and added to these immobilized capture probes. Unbound DNA is eluted and what remains is the target DNA of interest. Using these approaches, one could capture many genes at once, or even all coding sequences within the human genome. Finally, one could simply start with the entire human genome as the target sequence. The selected, amplified, captured, or targeted DNA sequences are either enzymatically digested or physically sheared to multiple diverse sequences of one to two hundred base pairs in length. Alternatively, using PCR-based selection approaches, amplicons are intentionally designed to be small (100–200 bp). These small fragments of DNA are end-ligated to an ‘adapter’ of defined/known sequence. This adapter is a small oligonucleotide that will be used to capture individual pieces of DNA and also used as a known sequence for further PCR amplification. At this point, library creation is complete (Head *et al.*, 2014).

Sequencing through DNA Synthesis

The next step in the sequencing process is the physical separation and PCR amplification of individual DNA molecules in the DNA fragment library prepared above. This is done through the use of the ligated adapters. There are two main methods for physical separation and amplification. The first uses very small beads which are coated with the complement oligonucleotide to the ligated adapter sequence. Beads and denatured library DNA are mixed in a defined ratio to ideally achieve a single DNA molecule hybridized to each bead. Excess DNA is washed away and the beads are added to an oil-aqueous emulsion in such a way as to have a single bead in each aqueous bubble. Also contained in this mixture are a heat stable polymerase, dNTPs, and primers that hybridize to the library adapter sequence. Each aqueous bubble can be thought of as an individual PCR reaction chamber. During PCR, copies are made of the individual sequence. These are contained within the aqueous bubble and re-hybridize to the target bead. After PCR is performed, each bead contains many identical copies of the individual DNA molecule that was originally hybridized to the bead. The second approach uses a single solid support coated with the complement to the ligated library DNA adapter. Library DNA is denatured and hybridized to this solid support. An amplification reaction is then performed. When new DNA molecules are synthesized from their captured template they hybridize back onto the solid support adjacent to their parent. Through multiple iteration of

amplification, tiny islands of DNA, all derived from the same sequence, are created.

The next step in this process is sequencing of the DNA bound to each individual bead or from each isolated DNA island (Figure 9). Pyrosequencing is ideally suited to the bead-based approaches (Harrington *et al.*, 2013). The beads are loaded into a honeycomb-welled plate. Each well is loaded to contain an individual bead; each bead is coated with a unique clonal DNA sequence. A deoxynucleotide (e.g., dATP), ADP, sulfurylase, luciferin, firefly luciferase, and a DNA polymerase are added to the plate. If the nucleotide (e.g., dATP) is incorporated by the DNA polymerase into the sequence attached to the bead, pyrophosphate is released. Pyrophosphate is converted to ATP by sulfurylase using ADP as a substrate. Luciferase then takes the substrates luciferin and ATP to produce light. In short, if the deoxynucleotide is incorporated into the DNA attached to a bead, the well containing that bead will glow. From this, it can be inferred that the sequence starts with the complement to the nucleotide added (i.e., if dATP was added, the sequence starts with a T). Next dCTP is added using the same approach, with the same result; a glowing well indicates that the next nucleotide in the target sequence is a G. Next comes dGTP and then dTTP. This cycle is repeated hundreds of times. Each time, a glowing well indicates the incorporation of the added nucleotide. Raw data are literally the amount of light produced by each well in each cycle. From this, the sequence of each bead-attached clonal DNA population can be inferred. In a single plate, millions of individual beads can be sequenced, resulting in hundreds of millions of sequenced nucleotides from the short 100–200 bp fragment sequence library. The analysis of millions of individual ‘reactors’ is why this approach is often referred to as ‘massively parallel’ DNA sequencing. Newer bead-based sequencing approaches can detect a change in pH in each reactor well that occurs with the incorporation of a nucleotide and associated release of pyrophosphate. This simplifies the chemistry behind the sequencing process.

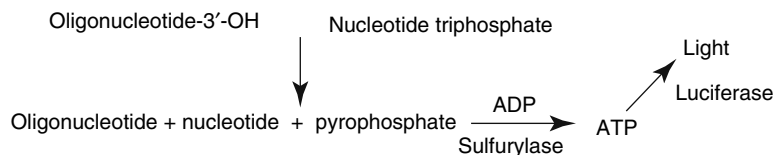
For the solid support approach, modified dNTPs are used in the sequencing reaction. These modifications include a 3’ OH blocking group and a fluorophore. The cycle of sequencing is conceptually very similar to the bead-based approach. To the solid support, a DNA polymerase and a single modified dNTP (e.g., dATP) is added. If the dATP is incorporated the tiny DNA island will fluoresce. A picture is taken and the blocking agent and fluorophore are chemically removed. Each dNTP is added in turn (ddAPT, ddCTP, ddGTP, ddTTP) and the process is cycled hundreds of times. The end result is an inferred sequence based on the fluorescence, or lack thereof, in each cycle for each clonal DNA island. As with the bead-based approach, this can produce millions of individual 100–200 bp long sequence reads, one for each DNA island.

Data Analysis

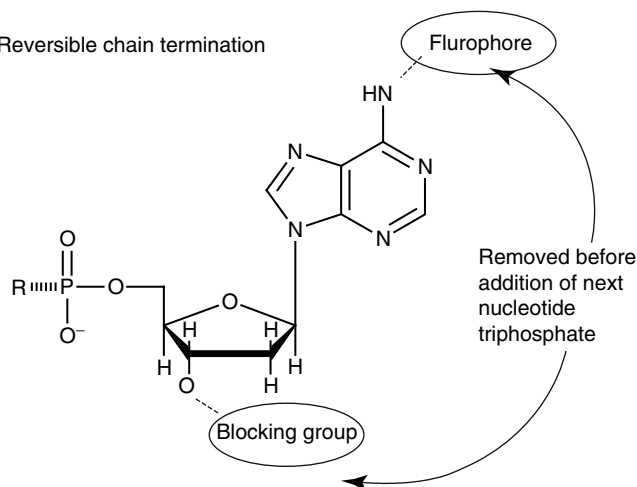
The end result of the wet work above is hundreds of thousands to millions of individual short DNA ‘reads.’ If a single DNA amplicon was sequenced, all of the millions of reads would be of this one short sequence. The number of times that an individual sequence is seen as a DNA island or on a DNA-coated bead is referred to as the ‘read depth.’ A read depth of one

Second-generation sequencing technologies

Pyrosequencing



Reversible chain termination



Real time sequencing

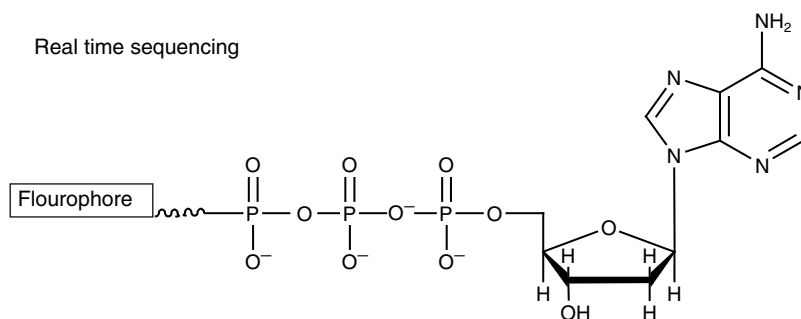


Figure 9 Methods for measuring DNA sequence analysis by synthesis.

million for a single DNA sequence in almost every situation would be excessive. If the entire coding sequence from a human genome was used to create the library, each individual read could be from any of the originally captured exons. This would lead to a smaller read depth for any one sequence but a much broader coverage across multiple individual DNA regions.

There are two challenges in the analysis of NGS data: the alignment of these individual short reads to the target genome (Schbath *et al.*, 2012) and the identification and annotation of base pair changes from the reference sequence (Nielsen *et al.*, 2011). Sequence alignment is computationally intense. When sequencing from the entire human genome, the target reference sequence is approximately 3 billion bases. To find where a single 100–200 bp fragment best fits within this reference sequence, particularly when the match might not be exact, takes special annotation of the target genome and alignment algorithms that are well beyond the scope of this article. Suffice it to say that aligning sequence from even a subset of the

human genome (e.g., all coding sequences, also referred to as the 'exome') takes many, many hours on even the fastest multiprocessor servers. The next step is a bit more straightforward; if a change from the reference sequence is found, it must be annotated as (1) a normal sequence variant of no clinical significance, (2) a disease causing or disease associated sequence variant, or (3) a variant of unknown clinical significance. While this can be a time-consuming process for the scientist or physician, databases exist which annotate all currently known sequence variants in the human genome. These are, by no means, complete, but are commonly used as a 'first pass' for sequence annotation (Nielsen *et al.*, 2011).

Future Directions

These NGS approaches have moved from the research laboratory to the clinic. Many laboratories routinely use NGS to

perform testing on patients. The great majority of these tests are developed by individual clinical labs. However, the Food and Drug Administration (FDA) has recently approved its first NGS platform for the full sequencing of the cystic fibrosis transmembrane conductance regulator (CFTR) gene for the diagnosis of CF in symptomatic patients.

The field of DNA sequencing continues to evolve and develop. On the analysis side, we can determine much more than sequence variants. Newer analysis algorithms allow for the determination of DNA copy number. This is very useful in detecting gene deletions and duplications. It has also been used on the research side to quantitate the expression of RNA transcripts, referred to as 'RNAseq' (Fox-Walsh *et al.*, 2011). Newer instrument methods for DNA sequencing focus on longer reads, faster data acquisition, and reduced costs with the goal of facile and economical acquisition of genomic sequence for the purpose of personalized medicine (Miller *et al.*, 2014). The majority of these technologies are still in the development phase. One approach is based on the 'real-time' analysis of the sequencing of a single DNA molecule in real time (Bashir *et al.*, 2012). This technique uses a unique device (zero-mode waveguide, ZMW) to measure the incorporation of a single individual deoxyribonucleotide triphosphates into a single DNA molecule. This approach has the advantage of 1000 base reads (English *et al.*, 2012). One other new approach that has received much notice is nanopore sequencing. Nanopore sequencing is based on the measurement of ion current levels as a DNA molecule is moved through a protein nanopore (Lazlo *et al.*, 2014). A portable device for nanopore sequencing has been developed (Quick *et al.*, 2014) raising the possibility of point of care NGS. While this still feels like a far off dream, 20 years ago, the idea that we could sequence the entire human genome if a few days for a few thousand dollars would have sounded crazy.

See also: Molecular Principles, Components, Technology, and Concepts: Nucleic Acids: The Chemical Synthesis of DNA and RNA Oligonucleotides for Drug Development and Synthetic Biology Applications

References

- Adams, R.L., 1995. Eukaryotic DNA methyltransferases — structure and function. *Bioessays* 17, 139–145.
- Akagi, R., Saho, T., Saito, M., *et al.*, 2014. Effective knock down of matrix metalloproteinase-13 by an intra-articular injection of small interfering RNA (siRNA) in a murine surgically-induced osteoarthritis model. *Journal of Orthopedic Research* 32, 1175–1180.
- Bashir, A., Klammer, A.A., Robins, W.P., *et al.*, 2012. A hybrid approach for the automated finishing of bacterial genomes. *Nature Biotechnology* 30, 701–707.
- Cech, T.R., Steitz, J.A., 2014. The noncoding RNA revolution—trashing old rules to forge new ones. *Cell* 157, 77–94.
- Cullen, B.R., 2006. Induction of stable RNA interference in mammalian cells. *Gene Therapy* 13, 503–508.
- Doležel, J., Greilhuber, J., 2010. Nuclear genome size: Are we getting closer? *Cytometry A* 77, 635–642.
- English, A.C., Richards, S., Han, Y., *et al.*, 2012. Mind the gap: Upgrading genomes with Pacific Biosciences RS long-read sequencing technologies. *PLoS One* 7 (11), e47768.
- Esanu, Y., 1988. Quelques moments du l'histoire des acides nucléiques, I. Chimie des acides nucléiques. *Virologie* 39, 221–234.
- Faulhammer, D., Famulok, M., 1997. Characterization and divalent metal-ion dependence of *in vitro* selected deoxyribozymes which cleave DNA/RNA chimeric oligonucleotides. *Journal of Molecular Biology* 269, 188–202.
- Fleskens, S.J., Takes, R.P., Otte-Höller, I., *et al.*, 2010. Simultaneous assessment of DNA ploidy and biomarker expression in paraffin-embedded tissue sections. *Histopathology* 57, 14–26.
- Fox, K.R., Brown, T., 2011. Formation of stable DNA triplexes. *Biochemical Society Transactions* 39, 634–639.
- Fox-Walsh, K., Davis-Turak, J., Zhou, Y., Li, H., Fu, X.D., 2011. A multiplex RNA-seq strategy to profile poly (A+) RNA: Application to analysis of transcription response and 3' end formation. *Genomics* 98, 266–271.
- Franchet-Beuzit, J., Spothem-Maurizot, M., Sabattier, R., Blazy-Baudras, B., Chaliel, M., 1993. Radiolytic footprinting. Beta rays, gamma photons, and fast neutrons probe DNA-protein interactions. *Biochemistry* 32, 2104–2110.
- Goodman, M.F., 1995. DNA models. Mutations caught in the act. *Nature* 378, 237–238.
- Hadnagy, A., Beaulieu, R., Balicki, D., 2008. Histone tail modifications and noncanonical functions of histones: Perspectives in cancer epigenetics. *Molecular Cancer Therapeutics* 7, 740–748.
- Hammarsten, O., 1894. Zur Kenntniss der nucleoproteide Zeitschrift für. *Physiologische Chemie* 19, 9–37.
- Hampshire, A.J., Rusling, D.A., Broughton-Head, V.J., Fox, K.R., 2007. Footprinting method for determining the sequence selectivity, affinity and kinetics of DNA-binding ligands. *Methods* 42, 128–140.
- Harrington, C.T., Lin, E.T., Olson, M.T., Eshleman, J.R., 2013. Fundamentals of pyrosequencing. *Archives of Pathology and Laboratory Medicine* 137, 1296–1303.
- Head, S.R., Komori, H.K., LaMere, S.A., *et al.*, 2014. Library construction for next-generation sequencing and challenges. *Biotechniques* 56, 61–68.
- Huang, X.C., Quesada, M.A., Mathies, R.A., 1992. DNA sequencing using capillary array electrophoresis. *Analytical Chemistry* 64, 2149–2154.
- Jin, S.G., Kadam, S., Pfeifer, G.P., 2010. Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Research* 38, e125.
- Kahl, B.F., Paule, M.R., 2009. The use of diethyl pyrocarbonate and potassium permanganate as probes for strand separation and structural distortions in DNA. *Methods in Molecular Biology* 543, 73–85.
- Kostopoulou, O.N., Papadopoulos, G., Kouvela, E.C., Kalpaxis, D.L., 2013. Clindamycin binding to ribosomes revisited: Foot printing and computation detection of two binding sites within the peptidyl transferase center. *Pharmazie* 68, 616–621.
- Kulis, M., Esteller, M., 2010. DNA methylation and cancer. *Advances in Genetics* 70, 27–56.
- Lazlo, A.H., Derrington, I.M., Ross, B.C., *et al.*, 2014. Decoding long nanopore sequencing reads of natural DNA. *Nature Biotechnology* 32, 829–833.
- Levine, P.A., London, E.S., 1929. The structure of thymonucleic acid. *Journal of Biological Chemistry* 83, 793–802.
- Lipsky, R.H., Mazzanti, C.M., Rudolph, J.G., *et al.*, 2001. DNA melting analysis for detection of single nucleotide polymorphisms. *Clinical Chemistry* 47, 635–644.
- Litt, M., 1971. Inactivation of yeast phenylalanine transfer ribonucleic acid by kethoxal. *Biochemistry* 10, 2223–2227.
- Malhomme de la Roche, H., Seagrove, S., Mehta, A., *et al.*, 2010. Using natural dietary sources of antioxidants to protect against ultraviolet and visible radiation-induced DNA damage: An investigation of human green tea ingestion. *Journal of Photochemistry and Photobiology* 101, 169–173.
- Marimuthu, C., Tang, T.H., Tominaga, J., Tan, S.C., Gopinath, S.C., 2012. Single-stranded DNA (ssDNA) production in DNA aptamer generation. *Analyst* 137, 1307–1315.
- Maxam, A.M., Gilbert, W., 1977. A new method for sequencing DNA. *Proceedings of the National Academy of Sciences United States of America* 74, 560–564.
- Metzker, M.L., 2010. Sequencing technologies — The next generation. *Nature Reviews Genetics* 11, 31–46.
- Miller, F.A., Hayeems, R.Z., Bytautas, J.P., *et al.*, 2014. Testing personalized medicine: Patient and physician expectations of next-generation genomic sequencing in late-stage cancer care. *European Journal of Human Genetics* 22, 391–395.
- Minko, I.G., Kozekov, I.D., Harris, T.M., *et al.*, 2009. Chemistry and biology of DNA containing 1,N²-deoxyguanosine adducts of the $\alpha\beta$ -unsaturated aldehydes acrolein, crotonaldehyde and 4-hydroxynonenal. *Chemical Research in Toxicology* 22, 759–778.
- Mortimer, S.A., Johnson, J.S., Weeks, K.M., 2009. Quantitative analysis of RNA solvent accessibility by N-silylation of guanosine. *Biochemistry* 48, 2109–2114.

- Mortimer, S.A., Kidwell, M.A., Doudna, J.A., 2014. Insights into RNA structure and function from genome-wide studies. *Nature Reviews in Genetics* 15, 469–479.
- Mortimer, S.A., Weeks, K.M., 2007. A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *Journal of the American Chemical Society* 129, 4144–4145.
- Nestor, C.E., Reddington, J.P., Benson, M., Meehan, R.R., 2014. Investigating 5-hydroxymethylcytosine (5hmC): The state of the art. *Methods in Molecular Biology* 1094, 243–258.
- Nielsen, R., Paul, J.S., Albrechtsen, A., Song, Y.S., 2011. Genotype and SNP calling from next-generation data. *Nature Reviews Genetics* 12, 443–451.
- Peak, M.J., Peak, J.G., 1990. Hydroxyl radical quenching agents protect against DNA breakage caused by both 365-nm UVA and by gamma radiation. *Photochemistry and Photobiology* 51, 649–652.
- Pfeifer, G.P., Szabo, P.E., 2009. 5-Hydroxymethylcytosine, a modified mammalian DNA base with a potential regulatory role. *Epigenomics* 1, 21–22.
- Quick, J., Quinlan, A.R., Loman, N.J., 2014. A reference bacterial genome dataset generated on the MinION™ portable single-molecule nanopore sequencer. *Gigascience* 3, 22.
- Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* 74, 5463–5467.
- Schbath, S., Martin, V., Zytynski, M., *et al.*, 2012. Mapping reads on a genomic sequence: An algorithmic overview and a practical comparative analysis. *Journal of Computational Biology* 19, 796–813.
- Snead, N.M., Rossi, J.J., 2012. RNA interference trigger variants: Getting the most out of RNA for RNA interference-based therapeutics. *Nucleic Acid Therapeutics* 22, 139–146.
- Solymosy, F., Hívösm, O., Gulyás, A., *et al.*, 1971. Diethyl pyrocarbonate, a new tool in the chemical modification of nucleic acids? *Biochimica et Biophysica Acta* 248, 406–416.
- von Sonntag, C., 2014. Carbohydrate radicals: From ethylene glycol to DNA strand breakage. *International Journal of Radiation Biology* 909, 416–422.
- Tabone, T., Sallman, G., Chiotis, M., *et al.*, 2006. Chemical cleavage of mismatch (CCM) to locate base mismatches in heteroduplex DNA. *Nature Protocols* 1, 2297–2304.
- Tarsounas, M., Tijsterman, M., 2013. Genomes and G-quadruplexes: For better or for worse. *Journal of Molecular Biology* 425, 4782–4789.
- Tijerina, P., Mohr, S., Russell, R., 2007. DMS footprinting of structured RNAs and RNA-protein complexes. *Nature Protocols* 2, 2008–2023.
- Tost, J., 2009. DNA methylation: An introduction to the biology and the disease-associated changes of a promising biomarker. *Methods in Molecular Biology* 507, 3–20.
- Toth, M., Müller, U., Doerfler, W., 1990. Establishment of de novo DNA methylation patterns. Transcription factor binding and deoxycytidine methylation at CpG and non-CpG sequences in an integrated adenovirus promoter. *Journal of Molecular Biology* 214, 673–683.
- Tucker, W.O., Shum, K.T., Tanner, J.A., 2012. G-quadruplex DNA aptamers and their ligands: Structure, function and application. *Current Pharmaceutical Design* 18, 2014–2026.
- Watson, J.D., 1968. *The Double Helix*. New York: Atheneum.
- Watson, J.D., Crick, F.H.C., 1953. A structure for deoxyribose nucleic acid. *Nature* 171, 737–738.
- Wyatt, M.D., Pittman, D.L., 2006. Methylating agents and DNA repair responses: Methylated bases and sources of strand breaks. *Chemical Research in Toxicology* 19, 1580–1594.
- Yan, B., Wang, Z., 2012. Long noncoding RNA: Its physiological and pathological roles. *DNA and Cell Biology* 31 (Suppl. 1), S34–S41.

Further Reading

- Adams, R.L.P., Knowler, J.T., Leader, D.P., 1986. *The Biochemistry of the Nucleic Acids*, tenth ed. London: Chapman and Hall.
- Balbo, S., Tureksy, R.J., Vialta, P.W., 2014. DNA adductomics. *Chemical Research in Toxicology* 27, 356–366.
- Blackburn, G.M., Gait, M.J., Loakes, D., Williams, D.M. (Eds.), 2006. *Nucleic Acids in Chemistry and Biology*, third ed. Cambridge: RSC Publishing.
- Buermans, H.P.J., den Dunnen, J.T., 2014. Next generation sequencing techniques – Advances and applications. *Biochimica et Biophysica Acta* 1842, 1932–1941.
- Franca, L.T., Carrilho, E., Kist, T.B., 2002. A review of DNA sequencing techniques. *Quarterly Review of Biophysics* 35, 169–200.
- Fredlake, C.P., Hert, D.G., Kan, C.W., *et al.*, 2008. Ultrafast DNA sequencing on a microchip by a hybrid separation mechanism that gives 600 bases in 6.5 min. *Proceedings of the National Academy of Sciences USA* 105, 476–481.
- Hartman, R.K. (Ed.), 2014. *Handbook of RNA Biochemistry*. Weinheim: Wiley-VCH.
- Kendrew, J. (Ed.), 1994. *The Encyclopedia of Molecular Biology*. Oxford: Blackwell.
- Kladwang, W., Mann, T.H., Becka, A., *et al.*, 2014. Standardization of RNA chemical mapping experiments. *Biochemistry* 53, 3063–3065.
- Mayer, G., 2010. *The Chemical Biology of Nucleic Acids*. Chichester: Wiley.
- Neidle, S., 2008. *Principles of Nucleic Acid Structure*. Amsterdam: Elsevier.
- Niedringhaus, T.P., Milanova, D., Kerby, M.B., Snyder, M.P., Barron, A.E., 2011. Landscape of next generation sequencing technologies. *Analytical Chemistry* 83, 4327–4341.
- Pickrell, W.O., Rees, M.I., Cheung, S.K., 2012. Next generation methodologies –an overview. *Advances in Protein Chemistry and Structural Biology* 89, 1–26.
- Rice, P.A., Correll, C.C., 2008. *Protein-Nucleic Acid Interactions: Structural Biology*. Cambridge: Royal Society of Chemistry.
- Sanchez-Flores, A., Abreu-Goodger, C., 2014. A practical guide to sequencing genomes and transcriptomes. *Current Topics in Medicinal Chemistry* 14, 398–406.
- Zimin, A., Stevens, K.A., Crepeau, M.W., *et al.*, 2014. Sequencing and assembly of the 22-gb loblolly pine genome. *Genetics* 196, 875–890.

The Chemical Synthesis of DNA and RNA Oligonucleotides for Drug Development and Synthetic Biology Applications

SL Beaucage and HV Jain, US Food and Drug Administration, Silver Spring, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The ability to design and synthesize DNA and RNA sequences has had a massive impact on biotechnology particularly in the rapidly growing fields of synthetic biology and nucleic acid-based drug development. Indeed, the chemical synthesis of DNA primers for amplification of specific DNA sequences via the polymerase chain reaction or DNA probes for the detection of RNA and DNA through Watson–Crick base-pairing recognition, has revolutionized modern molecular biology. The use of synthetic DNA/RNA sequences and their analogs for recognition and binding to messenger RNAs encoding disease-causing proteins has led to the production of nucleic acid-based drugs capable of inhibiting the expression of these proteins through either an antisense or an RNA interference pathway for therapeutic applications (Crooke, 2004). Furthermore, the ability to construct entire genes through enzymatic assembly of synthetic DNA sequences has enabled the total synthesis of *Mycoplasma genitalium* mouse mitochondrial genomes and the genome of an entire microorganism (Gibson *et al.*, 2010), thereby underscoring the power of chemically synthesized DNA in the realm of synthetic biology, genomics, and gene therapy. In this context, the *in situ* synthesis of DNA sequences on microarrays has also become a versatile tool for widespread applications including biomarker detection, high-throughput genomics, drug discovery, and synthetic biology. Moreover, *in situ* synthesized oligonucleotide arrays can also

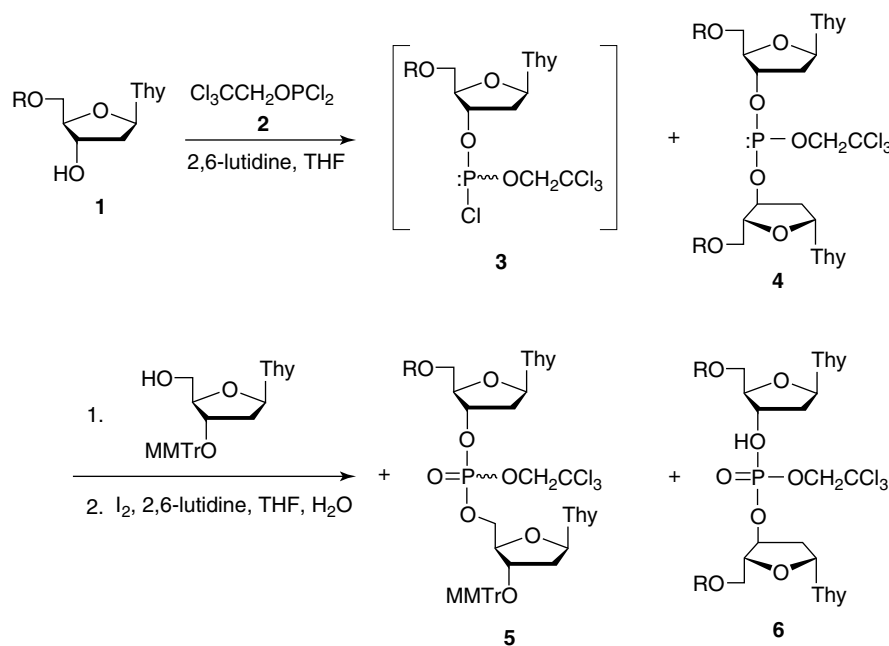
be used to generate pools of mixed DNA sequences for potential use in gene assembly (Tian *et al.*, 2004).

The chemical synthesis of DNA and RNA sequences has a long history; several DNA and RNA synthetic approaches based on the phosphate diester (Khorana *et al.*, 1972) and phosphate triester methods (Itakura *et al.*, 1984) have been reported for this purpose. Although these methods had significantly contributed to unraveling important biological problems, they are now of historical interest and will not be discussed in this article. More recent methodologies that are amenable to automated DNA/RNA synthesis on solid supports and on microarrays will be the focus of this report.

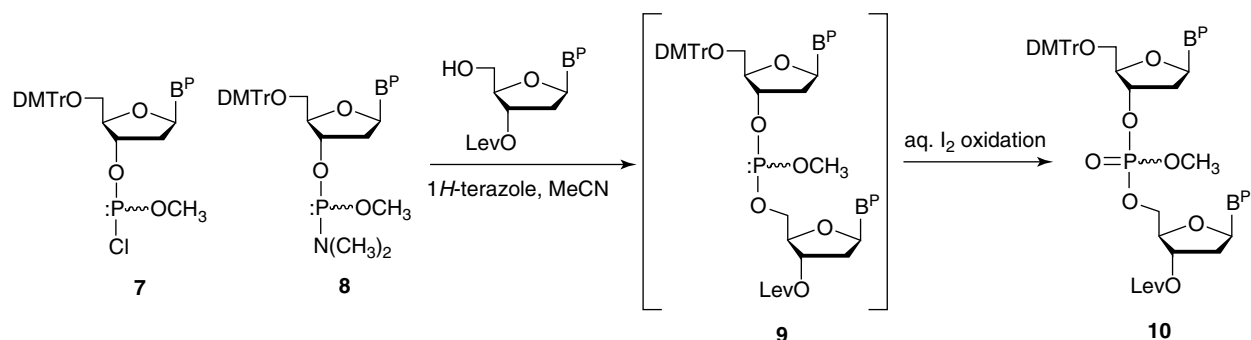
The Chemical Synthesis of Oligodeoxyribonucleotides

The Phosphoramidite Approach

The development of deoxyribonucleoside phosphoramidites is derived from the pioneering work of Letsinger and his coworkers on the use of deoxyribonucleoside chlorophosphites as synthetic intermediates (Letsinger and Lunsford, 1976). As shown in Scheme 1, the reaction of 5'-O-protected thymidine (1) with the bifunctional phosphitylating reagent 2 provides the deoxyribonucleoside chlorophosphite 3, within 5 min at -78°C , along with the symmetrical (3'→3')-dinucleoside phosphite triester 4. Addition of 3'-O-protected thymidine



Scheme 1 The phosphite triester approach to the synthesis of DNA sequences. R, phenoxyacetyl; Thy, thymine-1-yl; MMTr, 4-methoxytrityl.



Scheme 2 Replacement of deoxyribonucleoside chlorophosphites (**7**) with deoxyribonucleoside phosphoramidites (**8**) for the synthesis of DNA sequences. DMTr, 4,4'-dimethoxytrityl; B^P, protected pyrimidine and purine nucleobases; Lev, levulinyl.

followed by aqueous iodine oxidation results in the formation of dinucleoside phosphate triesters (**5** and **6**) within 1 h.

The formation of internucleoside linkages with such kinetics had been unprecedented in the mid-1970s. Because of the high reactivity and sensitivity of deoxyribonucleoside chlorophosphites to air oxidation and hydrolysis, these compounds had to be prepared immediately prior to use and stored under an inert atmosphere at -78°C . Furthermore, the production of variable amounts of **4** has made it difficult to consistently maintain the stoichiometry of **3** for optimal coupling efficiency. These limitations have compromised the automation of deoxyribonucleoside chlorophosphites for the synthesis of DNA sequences. Replacement of the chlorine atom of deoxyribonucleoside chlorophosphites **7** with a dimethylamino function has led to deoxyribonucleoside phosphoramidites **8** (Scheme 2), which have been introduced in the early 1980s as a novel class of reagents for the synthesis of DNA sequences (Beaucage and Caruthers, 1981).

The synthesis of these reagents consists of reacting suitably protected deoxyribonucleosides with chloro-(*N,N*-dimethylamino)methoxyphosphine in the presence of *N,N*-diisopropylethylamine. The phosphoramidites **8** are isolated as powders, stable to air oxidation and hydrolysis under normal laboratory conditions. The ability to convert these relatively stable phosphoramidites to highly reactive intermediates, under weakly acidic conditions, is the hallmark of the phosphoramidite approach to the synthesis of oligodeoxyribonucleotides. Typically, the reaction of **8** with 3'-*O*-levulinyl deoxythymidine in the presence of 1*H*-tetrazole produces, within minutes, the dinucleoside phosphite triesters **9**, which are ultimately oxidized to the phosphate triesters **10** (Scheme 2) in near quantitative yields based on ³¹P nuclear magnetic resonance (³¹P NMR) spectroscopy evidence (Beaucage and Caruthers, 1981). The phosphoramidite strategy has been successfully applied to the synthesis of various DNA sequences (Caruthers *et al.*, 1982) through the use of properly functionalized silica gel-based solid support (Köster, 1972; Matteucci and Caruthers, 1981). Over the years, the phosphoramidites **8** have been subjected to minor structural modifications in order to improve the stability of phosphoramidite monomers to the conditions used for solid-phase

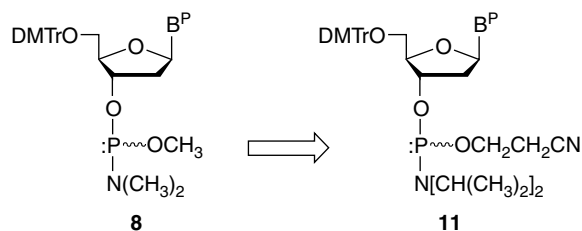
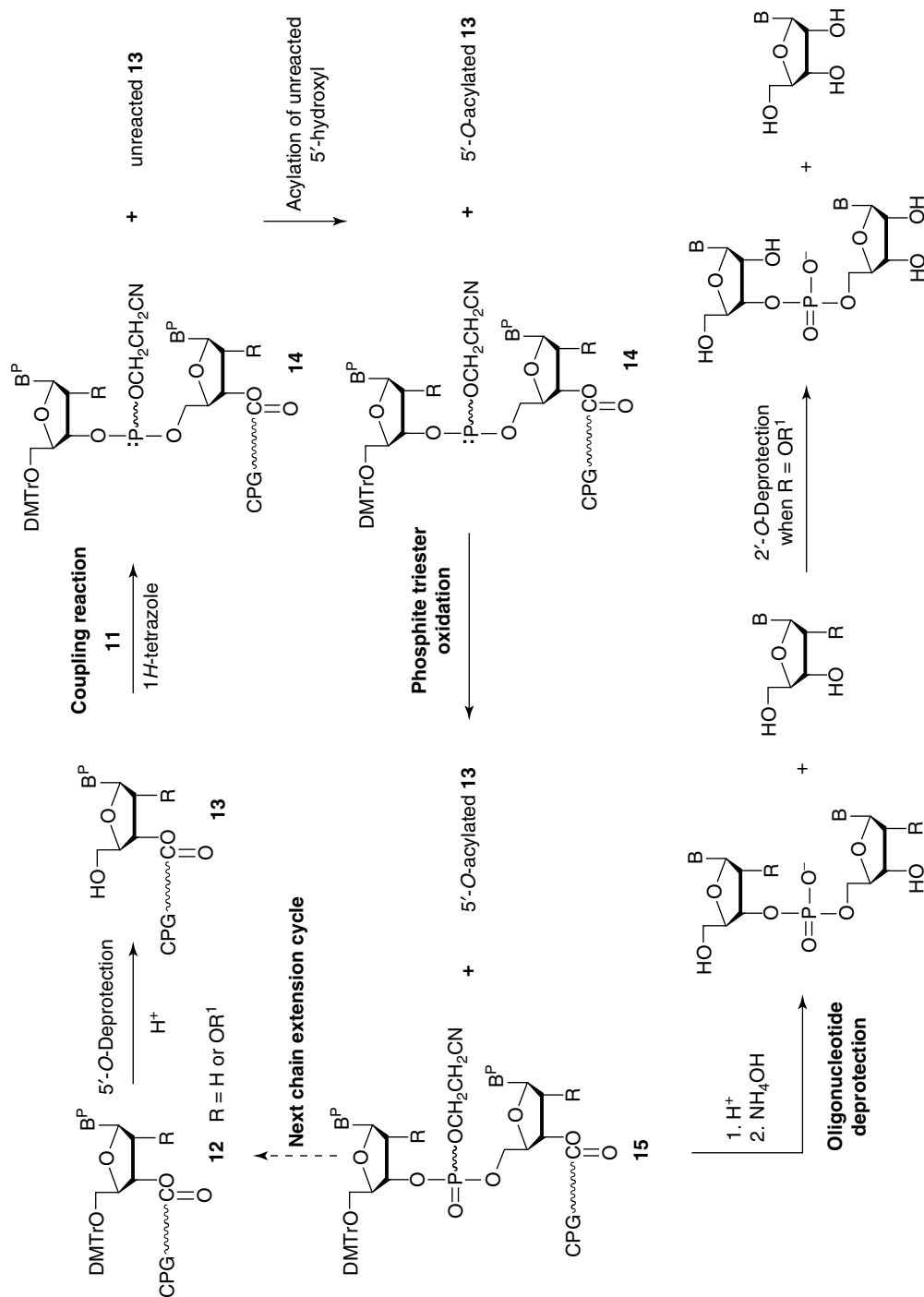


Figure 1 Structural modifications of the original deoxyribonucleoside phosphoramidite monomers (**8**) leading to their modern commercial versions (**11**). DMTr, 4,4'-dimethoxytrityl; B^P, protected pyrimidine and purine nucleobases.

DNA synthesis. These modifications include the replacement of the original dimethylamino function with other *N,N*-dialkylamino or *N,N*-cycloalkylamino groups and replacement of the methoxy P(III) protection with a variety of functionalized alkoxy groups (Beaucage and Iyer, 1992; Iyer and Beaucage, 1999) often at the expense of phosphoramidite reactivity. Out of all those proposed phosphoramidite modifications, the most popular deoxyribonucleoside phosphoramidites that have been used since the mid-1980s are the phosphoramidites **11** (Figure 1; Sinha *et al.*, 1984).

As shown in Scheme 3, the solid-phase synthesis protocol consists of: (1) cleaving the 5'-protecting group of a deoxyribonucleoside covalently attached to controlled-pore glass (CPG, **12**) under acidic conditions; (2) performing the reaction of **11** with the 5'-hydroxy function of **13** in the presence of 1*H*-tetrazole to produce the dinucleoside phosphite triester **14** along with some unreacted **13**; (3) capping the 5'-hydroxyl of unreacted **13** by acylating reagents to prevent further chain extension of this impurity; and (4) converting the phosphite triester **14** to its phosphate triester **15** using an appropriate oxidant. The synthesis protocol is then iteratively repeated until the desired chain length is obtained; the DNA sequence is finally exposed to aqueous or gaseous alkylamines to remove the amino protecting groups of the nucleobases while concomitantly cleaving the phosphate protecting groups and releasing the 5'-protected DNA sequence from the solid support.



Scheme 3 The phosphoramidite approach to the solid-phase synthesis of DNA or RNA sequences. DMTr, 4,4'-dimethoxytrityl; CPG, controlled-pore glass; R¹, protecting group; B^P, protected pyrimidine and purine nucleobases; B, thymine-1-yl/uracil-1-yl, cytosine-1-yl, adenine-9-yl, or guanine-9-yl.

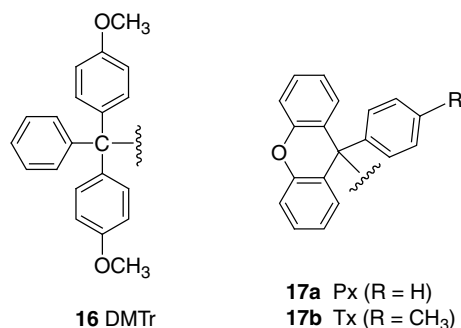


Figure 2 Chemical structure of the 4,4'-dimethoxytrityl (**16**), 9-phenylxanthen-9-yl (**17a**), and 9-(*p*-tolyl)xanthen-9-yl (**17b**) groups for 5'-hydroxy protection.

Hydrophobic 5'-protecting groups are often left intact in order to facilitate the separation of the DNA sequence from shorter than full-length sequence impurities using reversed-phase high-performance liquid chromatography (RP-HPLC) techniques (Iyer and Beaucage, 1999). The hydrophobic 5'-protecting group is removed after RP-HPLC purification of the DNA sequence is done.

Over the years, attempts have been made to improve individual steps of the solid-phase synthesis protocol described in Scheme 3. It is important to adequately protect the exocyclic amino groups of cytosine, adenine, and guanine moieties to prevent 'branching' of the DNA sequence during the course of its assembly. The base-labile acetyl, benzoyl, and isobutyryl groups for amino-protection of cytosine, cytosine or adenine and guanine, respectively, have been routinely used for this purpose, whereas the phenoxyacetyl and dimethylformamide groups have been recommended for amino-protection of cytosine/adenine and guanine, respectively, when milder basic deprotection conditions are required for specific applications. These nucleobase protecting groups are currently the most popular ones among those that have been proposed over the past 50 years (Iyer and Beaucage, 1999).

The acidic conditions used for the iterative cleavage of the 5'-O-dimethoxytrityl (DMTr) protecting group (**16**, Figure 2) during solid-phase synthesis of DNA sequences have been reassessed; the use of 3% trichloroacetic acid or 1.5% dichloroacetic acid in dry dichloromethane over a period of 20 or 60 s, respectively, has been found optimal for this purpose (Tram et al., 2011). Although the 9-phenylxanthen-9-yl group (Px, **17a**) and its 9-(*p*-tolyl)xanthen-9-yl derivative (**17b**) (Reese and Yan, 2004) have been proposed as attractive alternative 5'-O-protecting groups (Figure 2), the DMTr group remains the most popular 5'-hydroxy protective group for solid-phase synthesis of DNA sequences.

The development of acidic activators for deoxyribonucleoside phosphoramidite monomers has been the focus of intense scrutiny. Although 1*H*-tetrazole is still the most widely used phosphoramidite activator for solid-phase DNA synthesis, the coupling efficiency of deoxyribonucleoside phosphotetrazolide intermediates (Bernier et al., 1989) decreases when used in the synthesis of sterically demanding DNA or RNA sequences. In this regard, more acidic activators than 1*H*-tetrazole ($pK_a=4.9$), namely, 5-(4-nitrophenyl)-1*H*-tetrazole ($pK_a=3.7$), 5-(bis-3,5-trifluoromethylphenyl)-1*H*-

tetrazole ($pK_a=3.4$), 5-benzylthio-1*H*-tetrazole ($pK_a=4.1$), and 5-ethylthio-1*H*-tetrazole ($pK_a=4.3$) have been found more potent in terms of coupling kinetics and efficiencies for both solid-phase DNA and RNA synthesis (Wei, 2013). However, it should be noted that premature cleavage of the 5'-DMTr group leading to phosphoramidite double-couplings and degradation of the tricoordinated phosphite triester linkages are known to occur with the use of increasingly acidic activators. Several other classes of nucleoside phosphoramidite activators have also been investigated over the years (Wei, 2013); those include other azoles, azolium, pyridinium, saccharin and ammonium salt complexes, carboxylic acids, Lewis acids, trimethylchlorosilane, and 2,4-dinitrophenol. Currently, 1*H*-tetrazole, 5-benzylthio-1*H*-tetrazole, 5-ethylthio-1*H*-tetrazole, and 4,5-dicyanoimidazole are widely employed in the solid-phase synthesis of native and modified DNA and RNA sequences.

Although activated deoxyribonucleoside phosphoramidite monomers lead to high coupling efficiencies, coupling yields are not quantitative. Any unreacted DNA sequences should be deactivated through an acylation reaction in order to prevent further extension of those 'failure' sequences, which can complicate the purification of the desired DNA sequence from shorter than full length ones. The acylation reaction has also been shown to be effective in reducing the extent of guanine phosphorylation at O-6 during the coupling step of the solid-phase synthesis cycle (Pon et al., 1986). It is also recommended to perform this acylation reaction prior to the oxidation step, which would otherwise convert O⁶-phosphitylated guanine residues into more hydrolytically stable O⁶-phosphorylated congeners and result in the formation of an unwanted sequence ramification site at each modified guanine. The acylation of unreacted DNA sequences, is still performed nowadays in routine solid-phase DNA synthesis protocols.

Oxidation of the phosphite triester **14** to phosphate **15** is the last step of the solid-phase DNA synthesis protocol (Scheme 3). An aqueous solution of iodine and 2,6-lutidine (Letsinger and Lunsford, 1976) or pyridine (Usman et al., 1985) is generally used for this purpose primarily because this formulation is stable and provides rapid-reaction kinetics, usually without formation of side products. However, for specific applications, nonaqueous oxidizing reagents may advantageously offer an alternative to aqueous iodine for the oxidation of oligodeoxyribonucleoside phosphite triesters. These reagents (Beaucage and Iyer, 1992; Uzagare et al., 2003) include *m*-chloroperbenzoic acid, iodobenzene diacetate and tetra-*n*-butylammonium periodate, *tert*-butyl hydroperoxide and various hydroperoxides, dimethyldioxirane and ethyl(methyl)dioxirane, (1*S*)-(+)-(10-camphorsulfonyl)oxaziridine, and *N*-bromosuccinimide-dimethyl sulfoxide (DMSO). Out of these oxidants, an aqueous solution of iodine in pyridine, *tert*-butyl hydroperoxide and (1*S*)-(+)-(10-camphorsulfonyl)oxaziridine are most commonly used in the solid-phase synthesis of native and modified DNA sequences. The oxidation of phosphite triester **14** is particularly important, when it relates to the chemical synthesis of DNA sequences with internucleotide phosphorothioate diesters, which by virtue of their enhanced resistance to hydrolysis catalyzed by nucleases (Eckstein, 1985), have been extensively

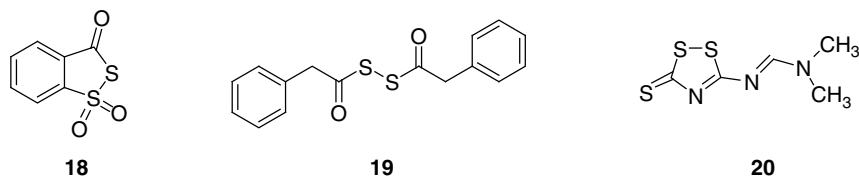


Figure 3 Potent sulfur-transfer reagents for converting internucleotidic phosphite triesters to phosphorothioate triester functions.

used as antisense sequences in the inhibition of gene expression. Automated synthesis of these modified oligonucleotides via the phosphoramidite method consists of replacing the aqueous iodine oxidation step with a sulfurization reaction that has originally been effected by elemental sulfur (Stec and Zon, 1984). The poor solubility of elemental sulfur in organic solvents has made its use, in automated systems, difficult. This problem has been resolved when phenylacetyl disulfide (Kamer *et al.*, 1989) and 3H-1,2-benzodithiol-3-one 1,1-dioxide (Iyer *et al.*, 1990) have been employed as sulfurization reagents. The success of these sulfur-transfer reagents in the automated synthesis of phosphorothioate DNA sequences has prompted the development of additional sulfurizing reagents. These sulfur-transfer agents (Beaucage and Caruthers, 2000) include *N,N,N',N'*-tetraethylthiuram disulfide, dibenzoyl tetrasulfide, bis-(*O,O*-diisopropoxyphosphinothioyl) disulfide, benzyltriethylammonium tetrathiomolybdate, bis-(*p*-toluenesulfonyl)disulfide, 3-ethoxy-1,2,4-dithiazoline-5-one, thiiranes, bis(ethoxythiocarbonyl)tetrasulfide, 3-methyl-1,2,4-dithiazoline-5-one, and 5-(dimethylaminomethylidene)amino-3H-1,2,4-dithiazole-3-thione (Guzaev, 2011). Out of those sulfur-transfer reagents, 3H-1,2-benzodithiol-3-one (18), phenylacetyl disulfide (19), and 5-(dimethylaminomethylidene)amino-3H-1,2,4-dithiazole-3-thione (20) are currently the most extensively used in solid-phase synthesis of oligonucleoside phosphorothioates (Figure 3).

An alternate two-step version of the phosphoramidite synthesis cycle has been reported to provide a simpler and more robust automation with significant cost reduction, when undertaking the large-scale synthesis of DNA sequences (Sierzchala *et al.*, 2003). The highpoints of this method are replacement of the 5'-DMTr blocking group with an aryloxycarbonyl protective group and the use of DMTr protection for the exocyclic amino functions of adenine and cytosine. A peroxy anion serves as a nucleophile to iteratively and simultaneously remove the 5'-O-carbonate protecting group and oxidize phosphite triesters to their tetracoordinated phosphate triester counterparts. However, for some reasons, this two-step synthesis cycle for the solid-phase synthesis of DNA sequences has not so far gained popularity among end users.

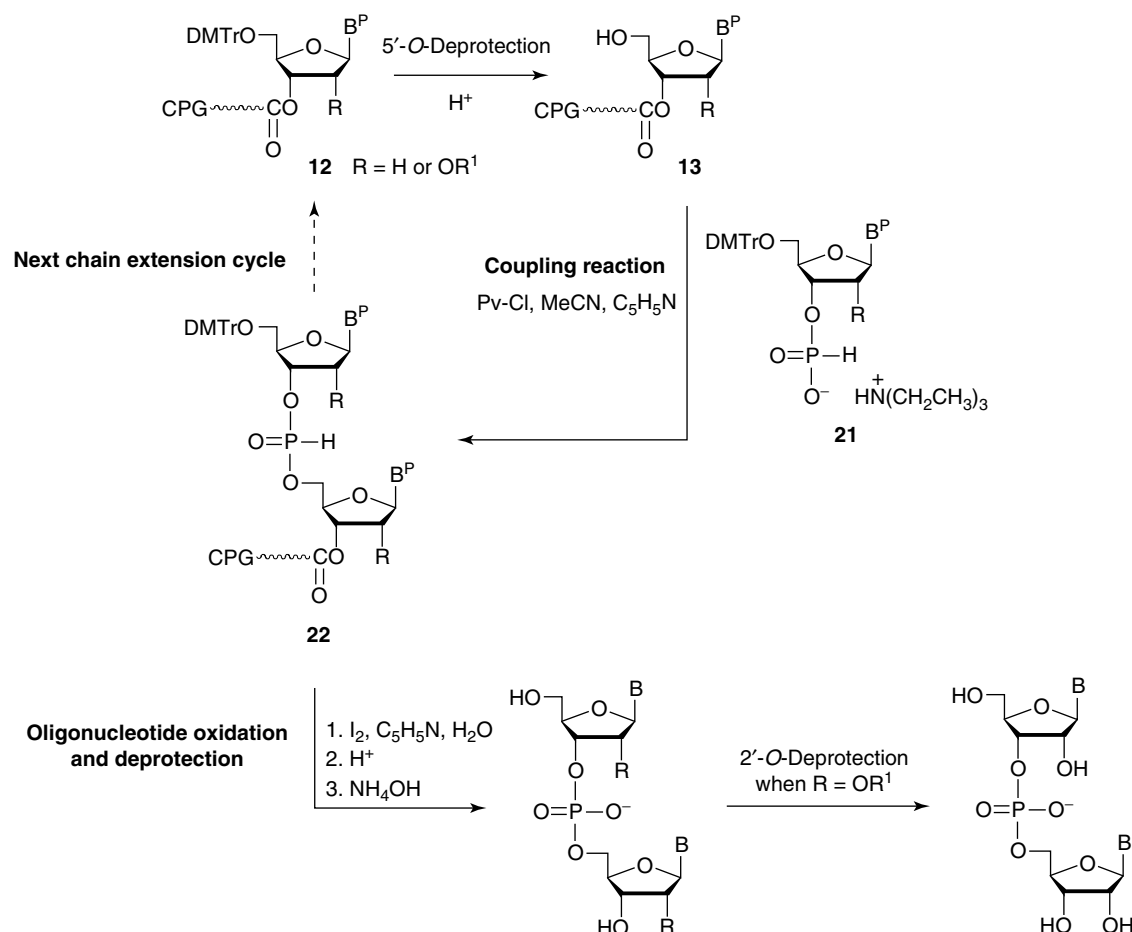
The *H*-Phosphonate Approach

H-Phosphonate monoesters are a distinct class of tetra-coordinated phosphorus compounds, which upon activation, are amenable to reaction with nucleophiles, thereby making them attractive for the preparation of oligonucleotides. The formation of *H*-phosphonate diesters from the reaction of *H*-phosphonate monoesters with diphenyl phosphorochloridate and a protected nucleoside had first been

demonstrated in the 1950s (Hall *et al.*, 1957). However, the use of *H*-phosphonate monoesters in the synthesis of oligonucleotides has been reevaluated in the mid-1980s. Deoxyribonucleoside *H*-phosphonates (21, Scheme 4) are key intermediates in the solid-phase synthesis of DNA sequences; these compounds are currently commercially available or can easily be prepared (Stawinski and Strömberg, 2001).

The solid-phase synthesis of DNA sequences via *H*-phosphonates consists of cleaving the 5'-DMTr group of a deoxyribonucleoside covalently attached to CPG (12) under acidic conditions and performing the reaction of 13 with 21 in the presence of pivaloyl chloride to produce the dinucleoside *H*-phosphonate diester 22. As shown in Scheme 4, the synthesis protocol can iteratively be repeated until the desired chain length is obtained; the DNA sequence is then subjected to an aqueous iodine treatment in the presence of pyridine or *N*-methylmorpholine to convert all phosphonate diester functions to native phosphate diesters. Given the sensitivity of *H*-phosphonate diesters to mildly alkaline conditions, a second aqueous iodine treatment in the presence of triethylamine is necessary to accelerate the oxidation reaction and prevent cleavage of the DNA chain (Strömberg and Stawinski, 2004). Upon completion of this critical oxidation step, the DNA sequence is exposed to concentrated aqueous ammonia in order to cleave all nucleobase protecting groups and release the native DNA sequence from the CPG support.

Over the years, several reagents have been proposed for the activation of protected deoxyribonucleoside *H*-phosphonate monoesters (Strömberg and Stawinski, 2004); pivaloyl chloride (Froehler and Matteucci, 1986) and adamantoyl chloride (Andrus *et al.*, 1988) have remained the most popular activators for solid-phase synthesis of DNA sequences via the *H*-phosphonate approach. Although a capping step appears unnecessary to prevent extension of unphosphorylated shortmers, the incorporation of a capping reaction into the *H*-phosphonate synthetic protocol has been shown to potentially improve the overall performance of the method. The capping procedure consists of performing an additional coupling reaction using isopropyl *H*-phosphonate after each nucleoside *H*-phosphonate coupling step (Andrus *et al.*, 1988). The solid-phase synthesis of DNA sequences, reaching 107 nucleobases in length, has been achieved using the *H*-phosphonate method (Froehler *et al.*, 1986). However, this approach to the synthesis of DNA sequences has limitations. As depicted in Scheme 5, premixing the deoxyribonucleoside *H*-phosphonate monoesters 21 with excess pivaloyl chloride leads to the formation of bis-acylphosphites (23) along with competing *O*-acylation of the 5'-hydroxyl of unphosphorylated DNA sequences, as a consequence of the slower phosphorylation kinetic of 23. Furthermore, because of the



Scheme 4 The *H*-phosphate approach to the solid-phase synthesis of DNA or RNA sequences. DMTr, 4,4'-dimethoxytrityl; CPG, controlled-pore glass; R¹, protecting group; Pv, trimethylacetyl; B^P, protected pyrimidine and purine nucleobases; B, thymine-1-yl/uracil-1-yl, cytosine-1-yl, adenine-9-yl, or guanine-9-yl.

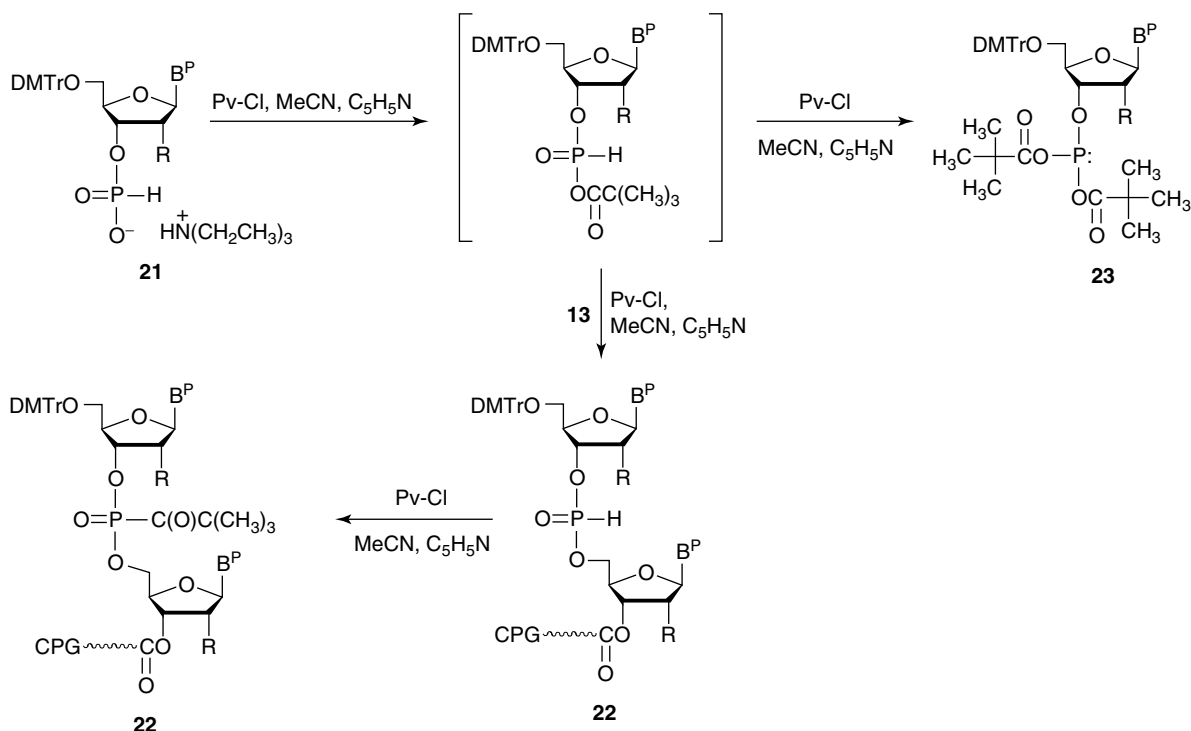
iterative exposure to pivaloyl chloride during the course of solid-phase DNA synthesis, *P*-acylation of *H*-phosphonate functions, to produce 24, occurs and could become problematic (Strömberg and Stawinski, 2004).

As mentioned above, the most damaging reaction encountered during the aqueous oxidation of polynucleoside *H*-phosphonates is the alkaline hydrolysis of *H*-phosphonate diester linkages, which leads to chain cleavage. When taken together, these drawbacks contribute to decrease the efficiency of the *H*-phosphonate approach to the synthesis of DNA sequences. Nonetheless, one of the advantages of the *H*-phosphonate method relates to the oxidation reaction, which can be carried out under nonaqueous conditions. Oxidation of polynucleoside *H*-phosphonates with elemental sulfur in pyridine/carbon disulfide or with 3*H*-1,2-benzothiaseleno1-3-one in pyridine/triethylamine affords phosphorothioate or phosphoroselenoate DNA sequences (Agrawal and Tang, 1990; Stawinski and Thelin, 1994). Alternatively, oxidation of polynucleoside *H*-phosphonates with carbon tetrachloride in the presence of primary and secondary amines or alcohols produced DNA phosphoramidate or phosphotriester sequences (Froehler, 1986; Froehler *et al.*, 1988).

Although the phosphoramidite approach continues to be preferred for the solid-phase synthesis DNA sequences on the basis of higher stepwise yields and fewer side-products, the *H*-phosphonate approach is particularly well-suited for the synthesis of RNA and modified DNA sequences (Stawinski and Krazewski, 1998; Strömberg and Stawinski, 2004).

The Chemical Synthesis of Oligoribonucleotides via the Phosphoramidite Approach

Given the recent developments in RNA interference as a means to silence the expression of specific genes, the availability of rapid and efficient methods for the chemical synthesis of RNA sequences in sufficient quantity for pharmaceutical applications has become an urgent issue. Although the various steps involved in the chemical synthesis of DNA and RNA oligonucleotides are similar (Scheme 2), the sequence of deprotection steps for RNA oligonucleotides differs in that the 2'-hydroxyl protecting group must be removed last, under conditions that would not result in the formation of (2'→5')-internucleotidic phosphodiester linkages and/or RNA chain cleavage.



Scheme 5 Formation of side products during solid-phase oligonucleotide synthesis via the *H*-phosphonate approach. DMTr, 4,4'-dimethoxytrityl; R, H, or OR¹; CPG, controlled-pore glass; Pv, trimethylacetyl; B^P, protected pyrimidine and purine nucleobases; B, thymine-1-yl/uracil-1-yl, cytosine-1-yl, adenine-9-yl, or guanine-9-yl.

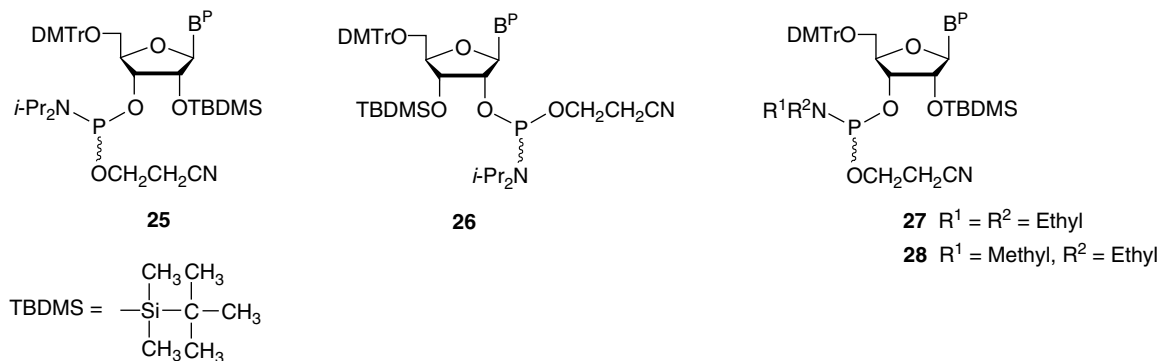


Figure 4 The *tert*-butyldimethylsilyl group for 2'-hydroxyl protection of ribonucleosides and their phosphoramidite derivatives. DMTr, 4,4'-dimethoxytrityl; B^P, protected pyrimidine and purine nucleobases.

The choice of a protecting group for the 2'-hydroxy function of ribonucleosides is clearly the most important decision to be made when undertaking the chemical synthesis of RNA sequences. The selected protecting group should optimally: (1) be easy to introduce; (2) remain completely stable until the final deblocking step of the fully assembled RNA sequence (Scheme 2); and (3) be totally removable under conditions that do not compromise the structural integrity of the RNA sequence. The main objective of this section is to report the most significant advances that have been made in the development of 2'-hydroxyl protecting groups for ribonucleosides as their ethers, acetals, and esters in the solid-phase synthesis of oligoribonucleotides.

Ether Protecting Groups

The tert-butyldimethylsilyl group

With the advent of the phosphoramidite method for solid-phase synthesis of oligodeoxyribonucleotides, the ribonucleoside phosphoramidites **25** ([Figure 4](#)) have been successfully employed in the total synthesis of a 77-nucleotide RNA sequence ([Ogilvie et al., 1988](#)). Although the *tert*-butyldimethylsilyl (TBDMS) group is still widely used for 2'-hydroxyl protection in the solid-phase synthesis of RNA sequences, it is not without limitations. The propensity of the TBDMS group to undergo (2' $\leftrightarrow$ 3')-isomerization during synthesis, purification, and isolation of 5'-O-DMTr-2'-O-

TBDMS-ribonucleosides (Damha and Ogilvie, 1993) is problematic, as it may lead to the production of isomeric 3'-O-TBDMS phosphoramidite impurities (26, Figure 4) during the preparation of 25. The incorporation of 26 into RNA sequences during solid-phase synthesis should be avoided as it would result in the formation of unwanted (2'→5')-internucleotidic linkages, which could negatively affect the efficiency of small interfering RNA (siRNA) sequences in silencing gene expression (Prakash *et al.*, 2006). The use of 2-cyanoethyl (*N,N*-diisopropyl)phosphoramidochloridite, 2,4,6-collidine, and *N*-methylimidazole has been found to not lead to detectable migration of the TBDMS group during preparation of the ribonucleoside phosphoramidites 25 (Scaringe *et al.*, 1990).

The removal of amino protecting groups from the nucleobases of 2'-O-TBDMS-protected oligoribonucleotides is also problematic given that concentrated aqueous NH₃ cannot be used at elevated temperature (55 °C) without causing premature deprotection of the 2'-O-TBDMS groups, which could result in internucleotide cleavages. The use of triethylamine trihydrofluoride (Gasparutto *et al.*, 1992; Westman and Strömberg, 1994) has been proposed for the removal of 2'-O-TBDMS groups from oligoribonucleotides. Because of its effectiveness and reliability, triethylamine trihydrofluoride is currently the most widely used reagent for this purpose. It should be noted that a 'one-pot' deprotection of 2'-O-TBDMS-protected RNA sequences using anhydrous methylamine and neat triethylamine trihydrofluoride (Bellon, 2000) has been developed to decrease the overall deprotection time of RNA sequences.

The coupling kinetics of the phosphoramidites 25 is notoriously sluggish, presumably because of the steric crowding created by the 2'-O-TBDMS group around the activated phosphoramidite function. More than 10 min is often required to produce satisfactory coupling efficiencies (> 98%). As for the activation of deoxyribonucleoside phosphoramidites, activators more powerful than 1*H*-tetrazole including 5-ethylthio-1*H*-tetrazole, 5-benzylthio-1*H*-tetrazole, and 4,5-dicyanoimidazole have been proposed for improving the coupling kinetics of ribonucleoside phosphoramidites. As discussed above in the previous section, the increased acidity of an activator may cause the premature cleavage of the 5'-O-DMTr group and double coupling of activated phosphoramidites within a given synthetic cycle could occur. This must be avoided in order to preserve the integrity of the RNA sequences. An alternate strategy for improving the coupling rates and coupling efficiencies of the ribonucleoside phosphoramidites 25 is to replace these monomers with the corresponding *N,N*-diethylphosphoramidites 27 (Figure 4) or other less sterically hindered phosphoramidites (28; Gasparutto *et al.*, 1992). Presumably because these ribonucleoside phosphoramidites are not commercially available, they have not been frequently used in solid-phase RNA synthesis.

The 2-cyanoethyl group

The 2-cyanoethyl group has been reported for the 2'-O-protection of suitably protected ribonucleosides and their 2'-O-(2-cyanoethyl) ribonucleoside phosphoramidite derivatives (29, Figure 5). The solid-phase synthesis of a 2'-O-cyanoethylated RNA sequence (21-mer) has been carried out using the

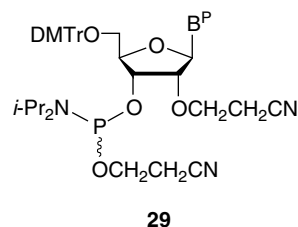


Figure 5 The 2-cyanoethyl group for 2'-hydroxyl protection of ribonucleosides and their phosphoramidite derivatives. DMTr, 4,4'-dimethoxytrityl; B^P, protected pyrimidine and purine nucleobases.

phosphoramidites 29 and 5-benzylthio-1*H*-tetrazole over a 10-min coupling time (Saneyoshi *et al.*, 2007).

Complete removal of the 2'-O-cyanoethyl groups from the RNA sequence is accomplished by treatment with *n*-Bu₄NF in THF/*n*-propylamine. Interestingly, fully or partially 2'-O-cyanoethylated RNA sequences induce the formation of relatively stable hybrids with complementary DNA or RNA sequences. Although 2'-O-cyanoethylated RNA sequences could have been suitable candidates for RNA interference applications, the results of such applications remain to be reported.

Acetal Protecting Groups

Unlike 2'-O-acyl or 2'-O-silyl-protecting groups, 2'-O-acetal functions do not migrate to vicinal hydroxy functions and are completely stable under the basic conditions usually used for the removal of nucleobase and internucleotide phosphate protecting groups. Such properties are attractive and have led to the widespread use of acetal groups for 2'-hydroxyl protection in the chemical synthesis of RNA sequences. A general limitation of these 2'-acetal protecting groups in RNA synthesis is their sensitivity to acidic conditions, which are routinely used for deblocking the 5'-O-DMTr or Px protective groups. Consequently, the development of 2'-acetal protecting groups should be carefully planned to minimize the inherent incompatibility with 5'-acid-labile protecting groups.

The 1-aryl-4-alkoxy piperidin-4-yl groups

The most noteworthy properties of the 1-(2-chloro-4-methylphenyl)-4-methoxypiperidin-4-yl (Ctmp), 1-(2-fluorophenyl)-4-methoxypiperidin-4-yl (Fpmp), and 1-(4-chlorophenyl)-4-methoxypiperidin-4-yl (Cpep) protecting groups (Lloyd *et al.*, 2000) are their stability under the relatively drastic acidic conditions required for the removal of 5'-DMTr or Px protecting groups and, paradoxically, their facile removal under mild acidic conditions in the final deblocking of RNA sequences. 2'-O-Ctmp-, 2'-O-Fpmp-, and 2'-O-Cpep-protected ribonucleosides have been prepared from suitably protected ribonucleosides and converted to their phosphoramidite derivatives 30, 31, and 32, respectively (Figure 6).

Activation of these phosphoramidites with 5-ethylthio-1*H*-tetrazole requires a 10–15 min coupling time; the need for such a long coupling reaction time underscores the importance of the steric bulk imparted by the 2'-hydroxyl protecting group on the coupling kinetics and coupling efficiencies of ribonucleoside phosphoramidites. The 2'-O-protected oligoribonucleotides provide stability against ubiquitous

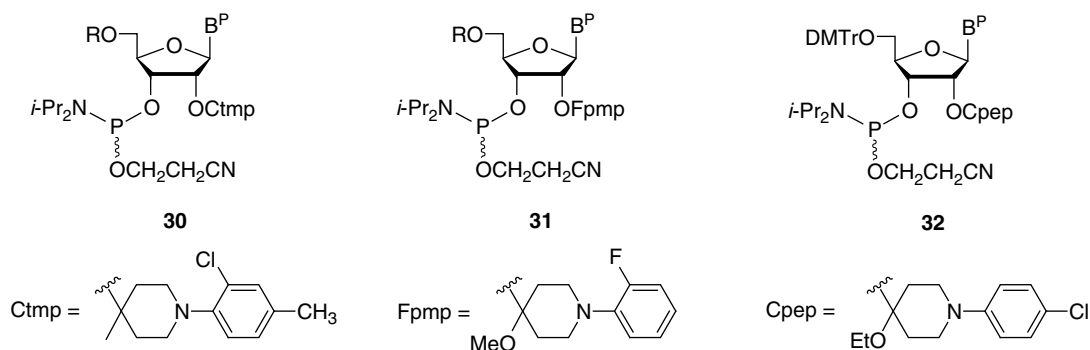


Figure 6 The 1-aryl-4-alkoxypiperidin-4-yl groups for 2'-hydroxyl protection of ribonucleosides and their phosphoramidite derivatives. R, 4,4'-dimethoxytrityl (DMTr) or 9-phenyloxanthren-9-yl; B^P, protected pyrimidine and purine nucleobases.

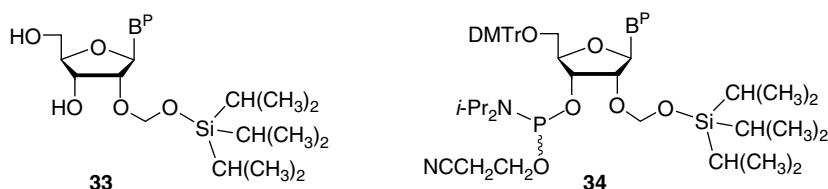


Figure 7 The TOM group for 2'-hydroxyl protection of ribonucleosides and their phosphoramidite derivatives. DMTr, 4,4'-dimethoxytrityl; B^P, protected pyrimidine and purine nucleobases.

ribonucleases throughout purification and storage. When required, the 2'-O-Ctmp, -Fpmp, and -Cpep groups are removed from RNA oligonucleotides under carefully controlled acidic conditions to reduce both (2' ↔ 3')-isomerization and the cleavage of internucleotidic phosphodiester linkages to a negligible extent. As eluted above, the Cpep group has been found to be more stable to acidic hydrolysis at pH 0.5 than either 2'-O-Ctmp or 2'-O-Fpmp and more labile to acidic hydrolysis at pH 3.75 than either of the latter, therefore allowing milder deprotection conditions (Lloyd *et al.*, 2000).

The large-scale synthesis of 2'-O-Cpep ribonucleosides and that of their phosphoramidite derivatives 32 has been reported as being cost-effective (Pon *et al.*, 2005) and therefore may offer a budgetary incentive over other ribonucleoside phosphoramidites when undertaking the large-scale preparation of siRNA sequences.

The triisopropylsilyloxymethyl group

The inherent flexibility of formaldehyde acetal linkers has been exploited to improve the coupling rates and coupling efficiencies of those ribonucleoside phosphoramidites, which have their 2'-hydroxy functions blocked with bulky groups. An example of this synthetic strategy relates to the development of the triisopropylsilyloxymethyl (TOM) group (Pitsch *et al.*, 2001), as an alternative to the TBDMS group for protection of the 2'-hydroxy functions of ribonucleosides and their phosphoramidite derivatives (Figure 7). One advantage of the 2'-O-TOM over the 2'-O-TBDMS group is that it does not undergo (2' ↔ 3')-isomerization during synthesis, purification, and isolation of 2'-O-TOM ribonucleosides (33) and their 5'-protected 2'-O-TOM phosphoramidite derivatives (34).

The high isomeric purity of 34 imparts correspondingly high integrity of the (3' → 5')-internucleotide linkages throughout RNA chain assembly. The coupling efficiency of the ribonucleoside phosphoramidites 34 is optimal when activated with 5-benzylthio-1H-tetrazole over a coupling time of 3–4 min. Under these conditions, coupling efficiencies >99% are achieved in the preparation of 40–84-nucleotide long RNA sequences on CPG supports. Removal of the nucleobase and phosphate protecting groups under basic conditions also releases the 2'-O-TOM-protected RNA sequences from the CPG supports. The 2'-O-TOM groups are removed from the RNA sequences by treatment with *n*-Bu₄NF•3H₂O or Et₄NF•3H₂O in DMSO, DMF or 1-methylpyrrolidin-2-one (Pitsch *et al.*, 2001). By virtue of the stability of the 2'-O-TOM group to a variety of reagents and reaction conditions and the commercial availability of 34, the use of these phosphoramidite monomers is still, nowadays, very popular for the preparation of native and modified RNA sequences (Porcher and Pitsch, 2005).

The 2-cyanoethoxymethyl group

A remarkable method for the solid-phase synthesis of oligoribonucleotides has emerged with the development of the 2'-O-(2-cyanoethoxymethyl) (CEM) ribonucleosides 35 (Figure 8) and their 5'-protected 2'-O-CEM phosphoramidite derivatives 36 (Shiba *et al.*, 2007).

These phosphoramidite monomers are particularly efficient at low concentration (0.05 M) in the solid-phase synthesis of oligoribonucleotides. Mixed-nucleobase RNA sequences of up to 170 nucleobases in length have been prepared on a CPG support with a pore size of 2000 Å (Nagata *et al.*, 2010). Activation of the phosphoramidites 36 is effected using 5-

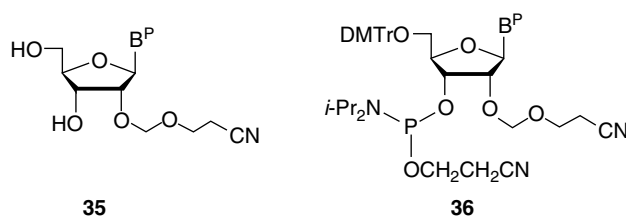


Figure 8 The 2-cyanoethoxymethyl group for 2'-hydroxyl protection of ribonucleosides and their phosphoramidite derivatives. DMTr, 4,4'-dimethoxytrityl; B^P, protected pyrimidine and purine nucleobases.

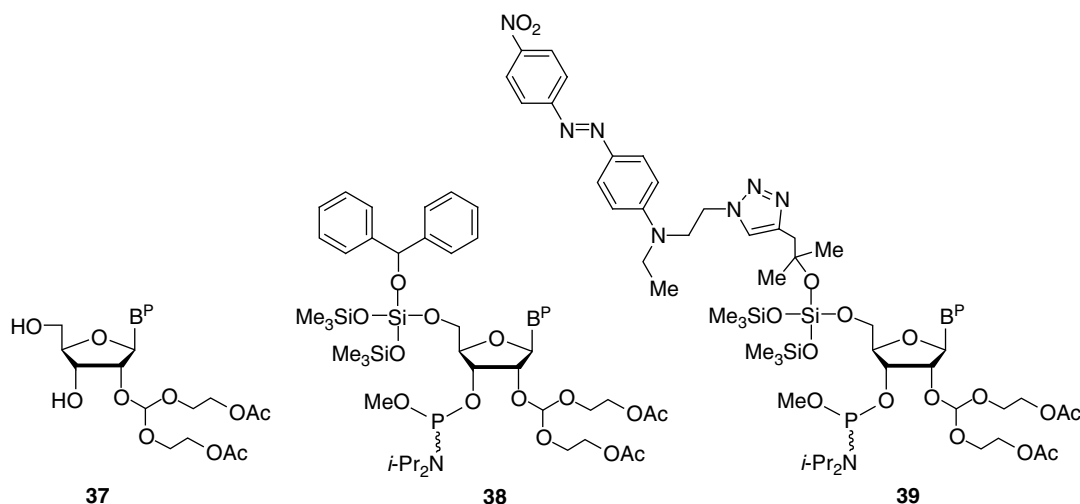


Figure 9 The bis-(2-acetoxyethoxy)methyl group for 2'-hydroxyl protection of ribonucleosides and their phosphoramidite derivatives. B^P, protected pyrimidine and purine nucleobases.

benzylthio-1*H*-tetrazole in acetonitrile over a period of 150 s. The modest sensitivity of the 2'-O-CEM group to the basic conditions used for nucleobase and phosphate deprotection and release of the 2'-O-CEM protected RNA sequences from the CPG support is noteworthy. Less than 5% of the 2'-O-CEM groups are removed on exposure to ethanolic ammonium hydroxide. Removal of the 2'-O-CEM group from the RNA sequence is effected by treatment with *n*-Bu₄NF in DMSO and nitromethane at ambient temperature. The use of 2'-O-CEM protection in the synthesis of a pyrophosphorylated 130-mer has led to, after enzymatic addition of a 5'-cap and a 3'-poly (A) tail, the unprecedented preparation of an artificial mRNA, which has been shown to support protein synthesis in a cell-free system (Nagata *et al.*, 2010).

Orthoester and Ester Protecting Groups

The bis-(2-acetoxyethoxy)methyl group

The use of the bis-(2-acetoxyethoxy)methyl (ACE) orthoester group for the protection of the 2'-hydroxy functions and the use of a silyl ether group for protection of the 5'-hydroxy function of ribonucleosides in solid-phase RNA synthesis has been reported (Scaringe, 2001, 2004). Conversion of the 2'-O-ACE-protected ribonucleosides 37 to 5'-silylated 2'-O-ACE-protected ribonucleoside phosphoramidites 38 (Figure 9) is achieved in two steps upon reaction with bis-(*N,N*-

diisopropylamino)methoxyphosphine in the presence of a catalytic amount of 1*H*-tetrazole. The solid-phase synthesis of 2'-O-ACE-protected RNA sequences is performed on a polystyrene support functionalized with an appropriate 2'-O-ACE-protected ribonucleoside 3'-O-hemisuccinate ester because CPG supports are not compatible with the iterative triethylamine trihydrofluoride-assisted 5'-O-desilylation reaction that is required for chain extension.

In this regard, commercial DNA/RNA synthesizers must be modified to accommodate the use of triethylamine trihydrofluoride formulations for 5'-O-desilylation reactions (Scaringe, 2001). The phosphoramidites 38 are activated with 5-ethylthio-1*H*-tetrazole and the coupling time is set to not exceed 90 s. Under these conditions, average coupling yields exceeding 99% are achieved; the oxidation of each phosphite triester linkage is effected by treatment with *tert*-butyl hydroperoxide. The solid-phase-linked RNA sequences are then exposed to disodium 2-carbamoyl-2-cyanoethylene-1,1-dithiolate (Dahl *et al.*, 1990) in DMF for complete removal of the methyl groups protecting the internucleotidic phosphate linkages while leaving the oligoribonucleotides attached to the polystyrene support. Aqueous MeNH₂ is then used to cleave the amino protecting groups of the nucleobases, deacylate the 2'-O-ACE-functions, and release the 2'-O-protected RNA sequences from the solid support. A particular advantage of the ACE-protecting group is that the removal of its two acetyl groups increases the water solubility of the still 2'-protected

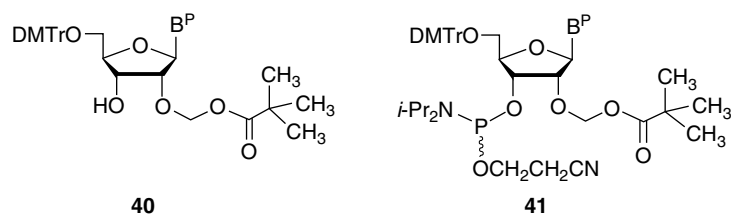


Figure 10 The pivaloyloxymethyl group for 2'-hydroxyl protection of ribonucleosides and their phosphoramidite derivatives. DMTr, 4,4'-dimethoxytrityl; B^P, protected pyrimidine and purine nucleobases.

RNA sequences. It should be noted that the resulting 2'-O-[bis-(2 hydroxyethyl)]methyl protecting groups are an order of magnitude more labile to acidic hydrolysis than the original ACE groups. Thus, complete cleavage of the 2'-O-protecting groups from the RNA sequences is achieved in a TEMED-acetate buffer solution (pH 3.8) at 60 °C (Scaringe, 2001).

The introduction of 5'-O-silyl-2'-O-ACE-protected ribonucleoside phosphoramidites **39** (Figure 9) with a visible chromophore appended to the 5'-silyl-protecting group allows for measurements of coupling efficiencies throughout the chain assembly process (Delaney *et al.*, 2008). It should be pointed out that the use of 5'-silylated 2'-O-ACE-protected ribonucleoside phosphoramidites in solid-phase RNA synthesis is one of the very best strategies for the preparation of high purity RNA sequences.

The pivaloyloxymethyl group

The base-labile pivaloyloxymethyl (PivOM) group has been proposed for the protection of 2'-hydroxy functions of ribonucleosides (Lavergne *et al.*, 2008, 2010). 5'-Protected 2'-O-PivOM ribonucleosides **40** (Figure 10) are prepared as a mixture of regioisomers. The chromatographically purified ribonucleosides **40** are converted to their phosphoramidite derivatives **41** under standard phosphitylation conditions.

These phosphoramidites have been employed in the solid-phase synthesis of chimeric RNA/DNA sequences of up to 21 nucleobases in length. The phosphoramidites **41** are activated with 5-benzylthio-1*H*-tetrazole and coupling efficiencies of >99% have been obtained within a coupling time of 3 min. The solid-phase bound chimeric RNA sequences are exposed to an acetonitrile solution of 1,8-diazabicyclo-[5.4.0]undec-7-ene to remove the 2-cyanoethyl phosphate protecting groups. In a second step, a 28% aqueous ammonia solution is used over a limited period of time (3–5 h) to remove the acyl groups protecting the nucleobases and the 2'-O-PivOM groups while releasing the chimeric RNA sequences from the CPG support. It has been claimed that the 2'-O-(hydroxymethyl) groups remain intact in aqueous ammonia solution, thereby preventing degradation of the fully deprotected RNA sequences under these basic conditions. It is however recommended that isopropylamine be added to these solutions, prior to evaporation to dryness, in order to avoid a reaction between the guanine residues and formaldehyde that is produced from the cleavage of the 2'-hemiacetal functions (Lavergne *et al.*, 2010). No chain cleavage, base modification and presence of unnatural internucleotidic phosphodiester linkages have been detected under these conditions. Crude chimeric RNA sequences are obtained in high yields and high purity.

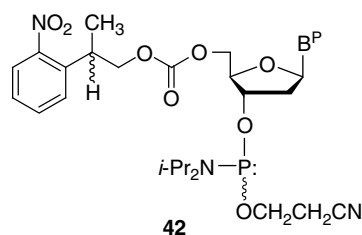


Figure 11 Deoxyribonucleoside phosphoramidites with 5'-photolabile protecting groups used in the synthesis of DNA sequences on glass surfaces. B^P, protected pyrimidine and purine nucleobases.

DNA and RNA Microarrays

DNA Microarrays

Although synthetic DNA oligonucleotides and genes have become widespread, the costs associated with the production of these biomolecules remain high with limited throughput. DNA microarrays have attracted considerable attention as they are recognized as abundant sources of inexpensive oligonucleotides for the construction of synthetic genes. Unlike conventional solid-phase oligonucleotide synthesis on CPG-filled columns, DNA microarray synthesis technologies utilize the surface of a silicon chip or a glass slide, with oligonucleotide growth confined to addressable surface locations. A variety of techniques including photolithography with physical or digital masks, inkjet printing, and electrochemical methods were used in the production of DNA microarrays for high-throughput identification and quantification of nucleic acids, monitoring gene expression or total gene assembly. These methods will be reviewed in this report.

Photolithography with physical or digital masks

The surface of a glass slide is commonly used for the covalent attachment of oligonucleotides; the glass surface is functionalized with a hydroxylalkylated linker (McGall *et al.*, 1996), which upon reaction with any deoxyribonucleoside phosphoramidites (42) bearing 5'-photolabile protecting groups (Figure 11), sets the stage for photolithographic DNA synthesis (Beier *et al.*, 2001).

A stack of masks needs to be prefabricated based on the oligonucleotide sequences to be synthesized on the arrays. These masks define the areas on the glass slide that need light exposure to remove the 5'-photolabile blocking groups and enable the free 5'-hydroxyl to react with an activated phosphoramidite monomer (Fodor *et al.*, 1991). Mask replacements and synthetic cycles are repeated until the DNA chain

assemblies are completed. This technology has been applied to the production of high-density probe arrays for gene expression and nucleic acid sequence analysis (McGall *et al.*, 1996). It should however be pointed out that the use of large numbers of prefabricated photomasks is costly. A relatively low-cost alternative to the use of physical masks is the maskless digital photolithographic method (Singh-Gasson *et al.*, 1999) using the Digital Light Processing technology (Richmond *et al.*, 2004). This method involves a reflective display device consisting of an electromechanically controlled array of micro-mirrors. This technology is easily programmable and very flexible; high resolution, precisely controllable light patterns can be generated in an automated manner. Scheme 6 illustrates a typical DNA synthesis cycle in maskless light-directed production of microarrays using the deoxyribonucleoside phosphoramidites 42 on an amine-derivatized glass slide that is functionalized with a leader nucleoside 43 through a base-labile linker. The DNA synthetic cycle follows the following order: (1) 5'-photodeprotection of 43 to 44, (2) coupling, (3) capping, and (4) 5'-photodeprotection.

An oxidation reaction is not required in each cycle because 5'-photodeprotection does not cleave oligonucleotides at the phosphite ester linkages, thereby allowing intermittent or a single final oxidation step. The coupling reaction consists of activating the phosphoramidites 42 with 4,5-dicyanoimidazole in acetonitrile over a period of 30–150 s. Under these conditions, the phosphoramidite coupling efficiency approaches 99%. A capping step, using an appropriate anhydride and 1-methylimidazole, is then performed and is followed by a drying step using helium, which has been found to improve the photodeprotection step by clearing fluid out of the delivery and waste lines. 5'-Photodeprotection of the 2-(2-nitrophenyl)propyloxycarbonyl group is performed using near ultraviolet light (365 nm) in the presence of an organic base. A single exposure to aqueous iodine is usually sufficient to convert all phosphite triester linkages to their corresponding phosphate triesters. Upon completion of microarray synthesis, the nucleobase and phosphate protecting groups are removed by immersing the array in an ethanolic solution of ethylenediamine or aqueous ammonia solution (Richmond *et al.*, 2004), which also releases the DNA sequences from the glass surface depending on the nature of the linker anchoring the DNA sequences to the surface of the array.

Maskless array synthesis is an efficient and versatile method for synthesizing high-density arrays of DNA sequences for hybridization and other molecular binding-based experiments. For applications requiring high DNA sequence purity, such as gene assembly, unintended deprotection due to light scattering and diffraction leads to sequence insertion errors, which contribute significantly to the overall error rate. However, these complications can be reduced with straightforward experimental strategies (Agbavwe *et al.*, 2011).

Inkjet printing

Piezoelectric inkjet oligoarray synthesis (Blanchard *et al.*, 1996) is a particularly versatile method that allows for rapid construction of oligonucleotide arrays containing any desired sequences. By dispensing deoxyribonucleoside phosphoramidite monomers from a multichannel inkjet print head,

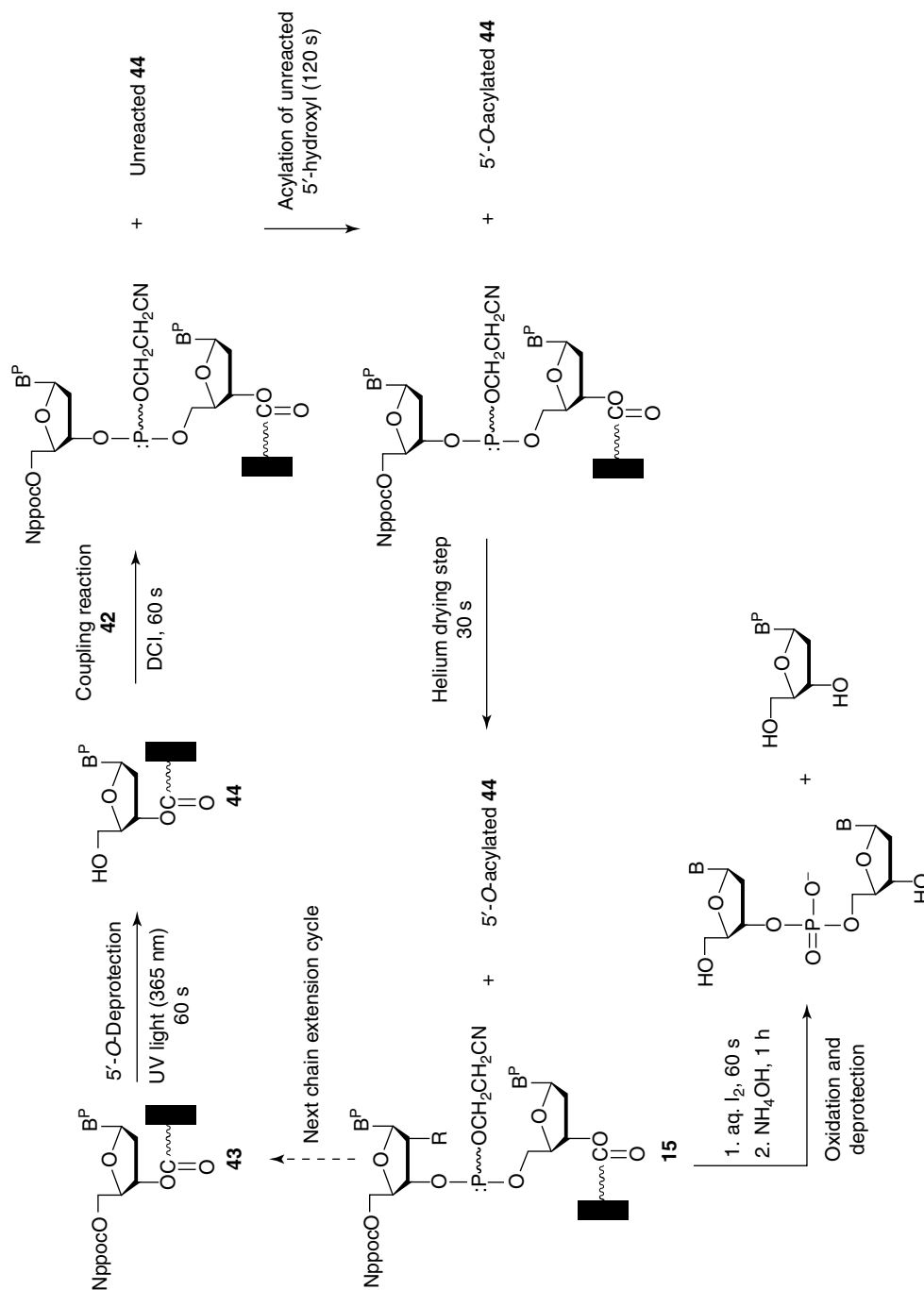
large numbers of sequences can be chemically synthesized in parallel. *De novo* oligonucleotide synthesis using a piezoelectric inkjet oligoarray synthesizer overcomes several of the problems inherent to conventional photolithographic mask arrays. DNA synthesis using the inkjet mechanism has been demonstrated with microfabricated piezoelectric inkjet pumps (Blanchard *et al.*, 1996); the inkjet head has six fluid channels, four of which delivering phosphoramidite precursors, one for an activator, and another one for an optional linker or modified nucleobase. The use of commercial deoxyribonucleoside phosphoramidites and reagents that are essentially the same as those employed for standard solid-phase DNA synthesis underscores the versatility of the method. Although, acetonitrile is the preferred solvent for deoxyribonucleoside phosphoramidites, this solvent is not suitable for inkjet printing due to its high volatility; droplets may even evaporate before contact with the slide surface. Furthermore, phosphoramidite precipitates can accumulate on the print head and clog the nozzles. To minimize evaporation of picolitre-sized droplets during printing, the selection of a suitable solvent is critical. A 1:1 solution of methyl glutaronitrile and 3-methoxypropionitrile had the most favorable combination of volatility, solubility, surface tension, and synthesis parameters. The solution readily dissolves phosphoramidite monomers and produces coupling efficiencies similar to those measured when using acetonitrile during routine solid-phase synthesis of DNA sequences up to 60 nucleobases in length; a 1×3 in glass slide contains about 185 000 DNA sequences. The inkjet printing method for microarray synthesis has mainly been developed and used for gene-expression profiling and comparative genomic hybridization applications (Wolber *et al.*, 2006).

Aside from the above technologies that have been used for the preparation of DNA microarrays, a number of other methods have been developed using either electrochemical (Egeland and Southern, 2005; Chow *et al.*, 2009) or microfluidic (Moorcroft *et al.*, 2005; Hua *et al.*, 2006; Huang *et al.*, 2007) devices for microscale parallel DNA synthesis.

Electrochemical arrays

Microelectrodes are employed to induce electrochemical reactions resulting in the production of an acid, which is required to iteratively deblock the 5'-DMTr groups of growing DNA chains (Egeland and Southern, 2005). The acid is produced only at specified sites by electrochemical oxidation of hydroquinone using individually addressable microelectrodes. As shown in Figure 12, the diffusion of the acid to its intended location on the DNA array is controlled by an adjacent cathode, which prevents the 5'-DMTr deblocking of neighboring sites by consuming any diffused acid.

Deoxyribonucleoside phosphoramidites are then added for DNA chain extension leading, after multiple synthesis cycles, to the construction of different DNA sequences on the array. The synthesis of only short DNA sequences has been successfully achieved on electrochemical arrays. A fundamental technical limitation of electrochemical DNA synthesis is the amount of acid produced after application of current to the microelectrodes; too much or too little acid-mediated deblocking, over multiple couplings, has a dramatic effect



Scheme 6 Synthesis of DNA oligonucleotides on glass surfaces using deoxyribonucleoside phosphoramidites with 5'-photolabile protecting groups. Nppoc, 2-(2-nitrophenyl)propyloxycarbonyl; DCI, 4,5-dicyanoimidazole; B^P, protected pyrimidine and purine nucleobases; B, thymine-1-yl, cytosine-1-yl, adenine-9-yl, or guanine-9-yl.

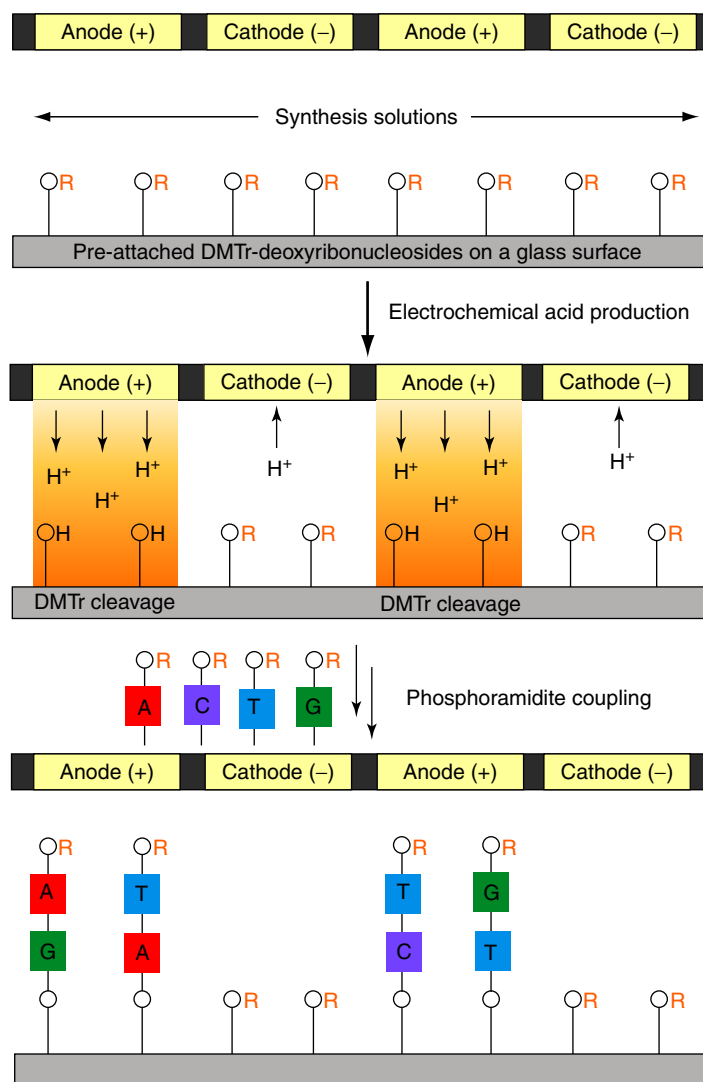


Figure 12 Illustration of an electrochemical DNA microarray. The synthesis solutions include a solution of hydroquinone and a solution of benzoquinone in acetonitrile. Anodic oxidation of hydroquinone produces acid, which migrates downward toward the array of pre-attached 5'-DMTr-protected deoxyribonucleosides and cleaves the DMTr groups (orange-colored areas). Acid that migrates sideways undergoes cathodic reduction to regenerate benzoquinone. Free 5'-hydroxyls are then exposed to activated deoxyribonucleoside phosphoramidites to extend the growing DNA chains. R, 4,4'-dimethoxytrityl (DMTr).

on the yields and integrity (depurination) of long DNA sequences.

Gene assembly from DNA microarrays

Successful attempts have been made to use oligonucleotides synthesized from DNA microarrays as building blocks for gene assembly. The use of cleavable linkers for anchoring the DNA sequences to the surface of the arrays are necessary to permit, when needed, the release of these DNA sequences from the arrays (Richmond *et al.*, 2004). For this purpose, a treatment with aqueous ammonia is adequate for complete deprotection and release of the DNA sequences from the surface of the arrays. The DNA sequences are then collected as a pool and used for downstream processing. Typically, 10^3 – 10^6 different DNA sequences can be synthesized on microarrays, albeit in

low yields ($\sim 10^6$ molecules of each DNA sequence); a pre-amplification of the DNA sequences is necessary to produce an optimal concentration of the sequences for a gene assembly reaction. In order to accommodate post-synthesis enzymatic amplification, the DNA sequences should have been designed to include short universal PCR primers, flanking each sequence. After amplification, the primers must be removed using appropriate restriction enzymes. Erroneous DNA sequences produced from array-synthesized and PCR-amplified oligonucleotides must be removed from the correct DNA sequences using hybridization arrays constructed with short DNA oligonucleotides complementary to the pool of DNA sequences earmarked for gene construction. After stringent hybridization and washes, the correct DNA sequences are sufficiently enriched for use in gene assembly (Tian *et al.*,

2004). A DNA microarrays with 10^3 – 10^6 different DNA sequences can lead to the construction of sequences equivalent to the 582 970 base pair genome of *M. genitalium* (Gibson *et al.*, 2008). Thus, the availability of synthetic DNA sequences has already and will continue to revolutionize molecular biology and bioengineering.

Microfluidics

Ultraviolet light-emitting diode-mediated DNA synthesis in glass capillaries

As an alternative to digital photolithography, the use of microfluidic devices for microscale parallel DNA synthesis has been proposed (Blair *et al.*, 2006). The synthesis of DNA strands inside hydroxyalkylated glass capillaries has been carried out using ultraviolet light-emitting diodes (LEDs) and photolabile deoxyribonucleoside phosphoramites (42). Through synchronization of fluid delivery and UV light emission from each LED, discrete DNA sequences are grown

on the inner surface of the capillary at the location of each LED (Figure 13).

A typical DNA synthesis cycle requires 180 s for photo-deprotection and 60 s for phosphoramidite coupling. After synthesis, the DNA sequences are deprotected by immersing the capillary in an ethanolic solution of ethylenediamine. Multiple oligonucleotides of up to 70 nucleobases in length have been synthesized in single capillaries and characterized by hybridization, sequencing, and gene synthesis. The DNA sequences have been demonstrated to be of high quality (up to 44% perfect sequences).

RNA microarrays

Over the past two decades, noncoding RNAs such as micro-RNAs, small interfering RNAs, and riboswitches have been reported to play important roles in various biological events by regulating gene expression. These events include development, differentiation, and metabolism. In this context, RNA aptamers and ribozymes have found applications in therapy and diagnostics, synthetic biology, and biosensors. With the objective to better understand the functions of these RNAs, it has become imperative to employ microarray technologies for high-throughput analysis of functional RNAs. The construction of RNA microarrays can be divided into three types. The first type involves the direct attachment of chemically modified RNA onto an appropriate surface; the immobilization of biotinylated RNA aptamers onto a streptavidin-coated glass slide (Collett *et al.*, 2005) provides an example of this type of microarrays for the detection of specific proteins. Alternatively, thiol-modified RNAs onto a gold-coated or maleimide-functionalized glass plates have been prepared for ultrasensitive detection and analysis of target DNA molecules using RNase H as an amplification tool (Goodrich *et al.*, 2004).

The second type of RNA microarrays relates to a method consisting of arraying various unmodified RNA sequences through ligation with the 5'-phosphate function of identical pre-arrayed DNA sequences on a gold surface using T4 RNA ligase. This method has been used to produce a five-component RNA microarray of potential aptamers for protein factor IXa (Figure 14). Surface plasmon resonance imaging (SPRI) measurements have been used to correctly identify the best aptamer to protein factor IXa; the structure of this aptamer has previously been determined from systematic evolution of ligands by exponential enrichment (SELEX) measurements (Li *et al.*, 2006).

The third type of RNA microarrays is an extension of the second type and is aimed at generating RNA sequences from the T7 RNA polymerase transcription of double-stranded DNAs attached to gold thin films in a microfluidic format (Figure 15). The RNA sequences diffuse, as they are synthesized in the microfluidic channel, and are captured by hybridization onto adjacently arrayed single-stranded DNA sequences complementary to the 5'-end of the RNA aptamer sequences (Chen *et al.*, 2012).

SPRI is then used to measure the bioaffinity of RNA aptamers for protein biomarkers. This method allows for the concurrent fabrication of multiple RNA sequences that can hybridize simultaneously to distinct anchored DNA sequences to produce a multiplexed RNA aptamer microarray.

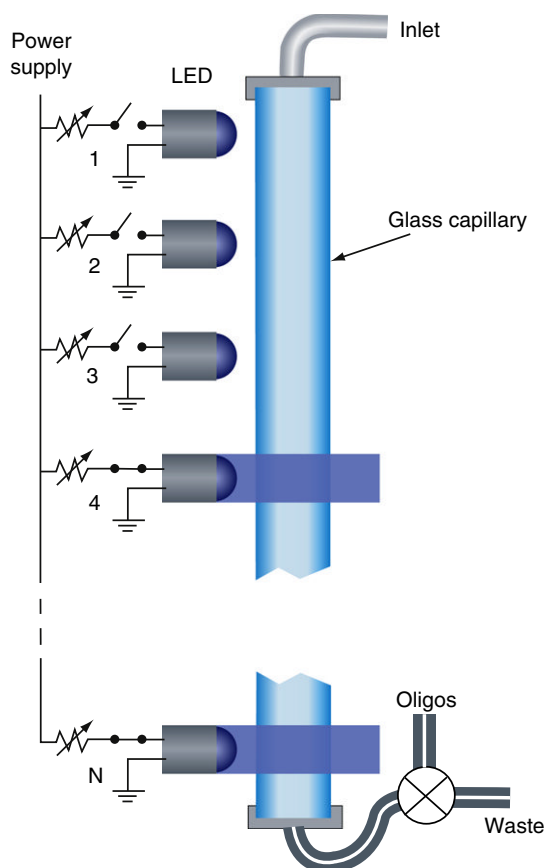


Figure 13 UV-LED-mediated DNA synthesis in glass capillaries. Selected regions of a glass capillary are illuminated by LEDs. The UV light emitted by each LED deprotects the 5'-hydroxyls of DNA strands covalently attached to the inside of the capillary through hydroxyalkylated linkers. Photolabile deoxyribonucleoside phosphoramites (42) are used for iterative DNA chain extension. LED, light-emitting diode. Reprinted from Blair, S., Richmond, K., Rodesch, M., Bassetti, M., Cerrina, F., 2006. A scalable method for multiplex LED-controlled synthesis of DNA in capillaries. *Nucleic Acids Research* 34, e110 with permission from Oxford University Press.

The *in situ* synthesis of RNA microarrays is considerably more problematic than that of DNA microarrays. However, in spite of challenging difficulties, rU₁₂ and rA₁₂ have been synthesized *in situ* on a glass surface using a maskless photolithographic array synthesizer and 5'-O-[2-(2-nitrophenyl)

propyloxycarbonyl] ribonucleoside phosphoramidites (45 and 46, Figure 16).

These phosphoramidites monomers are activated with 4,5-dicyanoimidazole in acetonitrile over a coupling time of 10–15 min. The standard capping reaction with acetic anhydride is performed and is followed by an aqueous iodine oxidation. The 5'-photolabile protecting group is iteratively cleaved upon exposure to UV light (365 nm). In order to determine coupling efficiencies, the RNA sequences are terminally labeled through a Cy3 phosphoramidite coupling reaction. On the basis of fluorescence intensity measurements, the coupling efficiency of 45 and 46 averaged 96% (Lackey *et al.*, 2009). Deprotection of rA₁₂ consisted of immersing the synthesized microarrays in an acetonitrile solution of triethylamine to remove the 2-cyanoethyl phosphate protecting groups; the levulinyl group protecting the nucleobase and 2'-O-acetal function of adenosine residues are subsequently cleaved under hydrazinolytic conditions. One important advantage, this method has over all other methods used for the fabrication of RNA microarrays, is that the removal of the

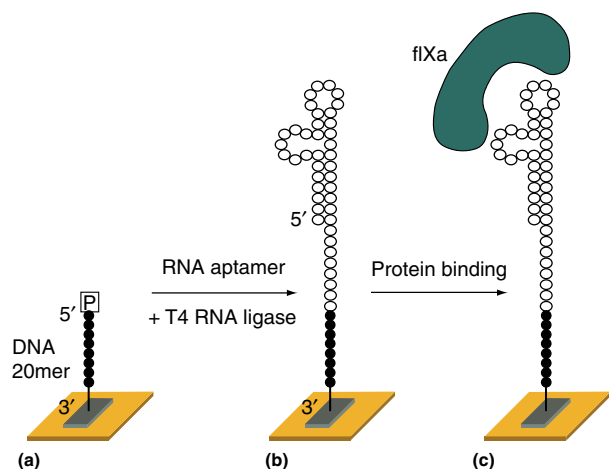


Figure 14 Simplified representation of an RNA aptamer microarray constructed using RNA–DNA ligation chemistry catalyzed by T4 RNA ligase. (a) The microarray consists of 5'-phosphate-terminated ssDNA molecules immobilized on a chemically modified gold surface via a 3'-thiol/surface maleimide reaction. (b) Ligation of the 3'-hydroxyl of an RNA aptamer to the 5'-phosphate of the DNA using T4 RNA ligase. (c) The RNA microarray can be used for the study of aptamer–protein affinity interactions. DNA bases and RNA bases are shown as closed circles and open circles, respectively; fIXa, protein factor IXa. Reprinted from Li, Y., Lee, H.J., Corn, R.M., 2006. Fabrication and characterization of RNA aptamer microarrays for the study of protein–aptamer interactions with SPR imaging. *Nucleic Acids Research* 34, 6416–6424 with permission from Oxford University Press.

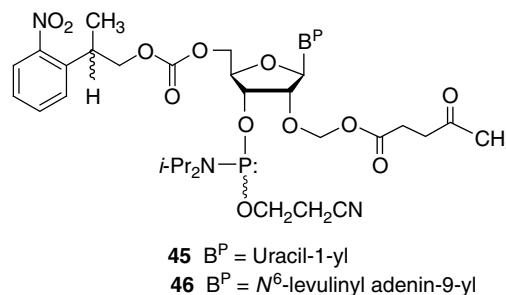


Figure 16 Ribonucleoside phosphoramidites with 5'-photolabile protecting groups used in the synthesis of RNA sequences on glass surfaces.

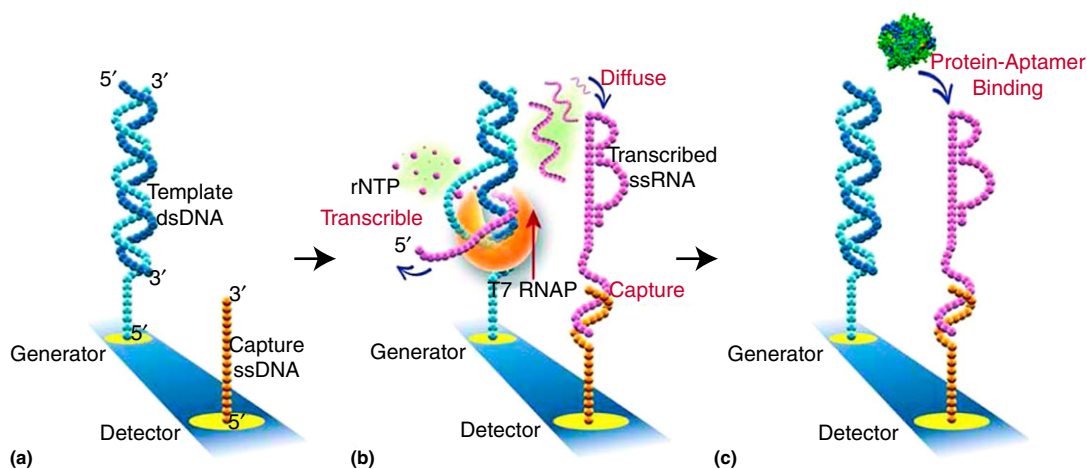


Figure 15 Schematic representation of on-chip synthesis of RNA aptamer microarrays. (a) DNA is covalently attached to the gold surface on the generator elements and then hybridized with a template DNA. (b) RNA is transcribed from dsDNA on the generator elements using T7 RNA polymerase and ribonucleoside 5'-triphosphates (rNTPs). The transcribed ssRNA is captured by ssDNA bound to the detector elements. (c) Injected protein is detected by the RNA aptamer on the detector elements. Reprinted with permission from Chen, Y., Nakamoto, K., Niwa, O., Corn R.M., 2012. On-chip synthesis of RNA aptamer microarrays for multiplexed protein biosensing with SPR imaging measurements. *Langmuir* 28, 8281–8285. © 2012, American Chemical Society.

2'-hydroxyl protecting group can be efficiently performed on the array, when needed. This feature minimizes the potential for degradation of the RNA sequences by ubiquitous RNases. Although this approach to the preparation of RNA microarrays has not as yet been extended to the *in situ* synthesis of RNA sequences with mixed nucleobases, this work represents a step in the right direction for the construction of RNA microarrays dedicated to the analysis of RNA aptamers as well as RNA-protein and RNA-RNA interactions of biological significance.

Concluding Remarks

The chemical synthesis of DNA and RNA sequences has had and will continue to have a positive influence on the development of nucleic acids-based drugs for the treatment of cancer and infectious diseases. Furthermore, with the use of array-derived oligonucleotides and the ability to synthesize longer DNA sequences, one might expect those technologies to have a wide range of applications in high-throughput genomics, biodetection, and synthetic biology in the context of gene assembly and genome constructs, which should benefit all areas of molecular biology.

See also: Molecular Principles, Components, Technology, and Concepts: Nucleic Acids: DNA, RNA Chemical Properties (Including Sequencing and Next-Generation Sequencing)

References

- Agbawue, C., Kim, C., Hong, D., *et al.*, 2011. Efficiency, error and yield in light-directed maskless synthesis of DNA microarrays. *Journal of Nanobiotechnology* 9, 57–74.
- Agrawal, S., Tang, J.-Y., 1990. Efficient synthesis of oligoribonucleotide and its phosphorothioate analog using H-phosphonate approach. *Tetrahedron Letters* 31, 7541–7544.
- Andrus, A., Efcavitch, J.W., McBride, L.J., Giusti, B., 1988. Novel activating and capping reagent for improved hydrogen-phosphonate DNA synthesis. *Tetrahedron Letters* 29, 861–864.
- Beaucage, S.L., Caruthers, M.H., 1981. Deoxynucleoside phosphoramidites — A new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Letters* 22, 1859–1862.
- Beaucage, S.L., Iyer, R.P., 1992. Advances in the synthesis of oligonucleotides by the phosphoramidite approach. *Tetrahedron* 48, 2223–2311.
- Beaucage, S.L., Caruthers, M.H., 2000. Synthetic strategies and parameters involved in the synthesis of oligodeoxyribonucleotides according to the phosphoramidite method. In: Beaucage, S.L., Bergstrom, D.E., Glick, G., Jones, R.A. (Eds.), *Current Protocols in Nucleic Acid Chemistry*, vol. 1. New York, NY: Wiley, pp. 3.3.1–3.3.20.
- Beier, M., Stephan, A., Hoheisel, J.D., 2001. Synthesis of photolabile 5'-O-phosphoramidites for the photolithographic production of microarrays of inversely oriented oligonucleotides. *Helvetica Chimica Acta* 84, 2089–2095.
- Bellon, L., 2000. Oligoribonucleotides with 2'-O-(*tert*-Butyldimethylsilyl) groups. In: Beaucage, S.L., Bergstrom, D.E., Glick, G., Jones, R.A. (Eds.), *Current Protocols in Nucleic Acid Chemistry*, vol. 1. New York, NY: Wiley, pp. 3.6.1–3.6.13.
- Berner, S., Mühlegger, K., Seliger, H., 1989. Studies on the role of tetrazole in the activation of phosphoramidites. *Nucleic Acids Research* 17, 853–864.
- Blair, S., Richmond, K., Rodesch, M., Bassetti, M., Cerrina, F., 2006. A scalable method for multiplex LED-controlled synthesis of DNA in capillaries. *Nucleic Acids Research* 34, e110.
- Blanchard, A.P., Kaiser, R.J., Hood, L.E., 1996. High-density oligonucleotide arrays. *Biosensors and Bioelectronics* 11, 687–690.
- Caruthers, M.H., Beaucage, S.L., Becker, C., *et al.*, 1982. New methods for synthesizing deoxyoligonucleotides. In: Setlow, J.K., Hollaender, A. (Eds.), *Genetic Engineering: Principles and Methods*, vol. 4. New York, NY: Plenum, pp. 1–17.
- Chen, Y., Nakamoto, K., Niwa, O., Corn, R.M., 2012. On-chip synthesis of RNA aptamer microarrays for multiplexed protein biosensing with SPR imaging measurements. *Langmuir* 28, 8281–8285.
- Chow, B.Y., Emig, C.J., Jacobson, J.M., 2009. Photoelectrochemical synthesis of DNA microarrays. *Proceedings of the National Academy of Sciences of the United States of America* 106, 15219–15224.
- Collett, J.R., Cho, E.J., Lee, J.F., Levy, M., Hood, A.J., 2005. Functional RNA microarrays for high-throughput screening of anti-protein aptamers. *Analytical Biochemistry* 338, 113–123.
- Crooke, S.T., 2004. Antisense strategies. *Current Molecular Medicine* 4, 465–487.
- Dahl, B.H., Bjergårde, K., Henriksen, L., Dahl, O., 1990. A highly reactive, odourless substitute for thiophenol/triethylamine as a deprotection reagent in the synthesis of oligonucleotides and their analogues. *Acta Chemica Scandinavica* 44, 639–641.
- Damha, M.J., Ogilvie, K.K., 1993. Oligoribonucleotide synthesis. In: Agrawal, S. (Ed.), *Methods in Molecular Biology: Protocols for Oligonucleotides and Analogs*, vol. 20. Totowa, NY: Humana Press, pp. 81–114.
- Delaney, M.O., Thomas, A., Ricketts, C., Kitchen, D.E., Kaiser, R.J., 2008. Chromophoric 5'-O-silyl protection of *N*-protected 2'-ACE ribonucleosides for solid-phase RNA synthesis. In: Beaucage, S.L., Herdewijn, P., Matsuda, A. (Eds.), *Current Protocols in Nucleic Acid Chemistry*, vol. 3. Hoboken, NJ: Wiley, pp. 2.14.1–2.14.26.
- Eckstein, F., 1985. Nucleoside phosphorothioates. *Annual Review of Biochemistry* 54, 367–402.
- Egeland, R.D., Southern, E.M., 2005. Electrochemically directed synthesis of oligonucleotides for DNA microarray fabrication. *Nucleic Acids Research* 33, e125.
- Fodor, S., Read, J., Pirrung, M., *et al.*, 1991. Light-directed, spatially addressable parallel chemical synthesis. *Science* 251, 767–773.
- Froehler, B.C., 1986. Deoxynucleoside H-phosphonate diester intermediates in the synthesis of internucleotide phosphate analogues. *Tetrahedron Letters* 27, 5575–5578.
- Froehler, B.C., Matteucci, M.D., 1986. Nucleoside H-phosphonates: Valuable intermediates in the synthesis of deoxyoligonucleotides. *Tetrahedron Letters* 27, 469–472.
- Froehler, B.C., Ng, P.G., Matteucci, M.D., 1986. Synthesis of DNA via deoxynucleoside H-phosphonate intermediates. *Nucleic Acids Research* 14, 5399–5407 and references therein.
- Froehler, B., Ng, P., Matteucci, M., 1988. Phosphoramidate analogs of DNA: Synthesis and thermal stability of heteroduplexes. *Nucleic Acids Research* 16, 4831–4839.
- Gasparutto, D., Livache, T., Bazin, H., *et al.*, 1992. Chemical synthesis of a biologically active natural tRNA with its minor bases. *Nucleic Acids Research* 20, 5159–5166.
- Gibson, D.G., Benders, G.A., Andrews-Pfannkoch, C., *et al.*, 2008. Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* 319, 1215–1220.
- Gibson, D.G., Smith, H.O., Hutchison III, C.A., Venter, J.C., Merryman, C., 2010. Chemical synthesis of the mouse mitochondrial genome. *Nature Methods* 7, 901–903 and references therein.
- Goodrich, T.T., Lee, H.J., Corn, R.M., 2004. Enzymatically amplified surface plasmon resonance imaging method using RNase H and RNA microarrays for the ultrasensitive detection of nucleic acids. *Analytical Chemistry* 76, 6173–6178.
- Guzavev, A.P., 2011. Reactivity of 3*H*-1,2,4-dithiazole-3-thiones and 3*H*-1,2-dithiole-3-thiones as sulfurizing agents for oligonucleotide synthesis. *Tetrahedron Letters* 52, 434–437.
- Hall, R.H., Todd, A., Webb, R.F., 1957. Nucleotides. Part XLI. Mixed anhydrides as intermediates in the synthesis of dinucleoside phosphates. *Journal of the Chemical Society*, 3291–3296.
- Hua, A., Xia, Y., Srivannavit, O., *et al.*, 2006. A versatile microreactor platform featuring a chemical-resistant microvalve array for addressable multiplex syntheses and assays. *Journal of Micromechanics and Microengineering* 16, 1433–1443.
- Huang, Y., Castrataro, P., Lee, C.-C., Quake, S.R., 2007. Solvent resistant microfluidic DNA synthesizer. *Lab on a Chip* 7, 24–26.
- Itakura, K., Rossi, J.J., Wallace, R.B., 1984. Synthesis and use of synthetic oligonucleotides. *Annual Review of Biochemistry* 53, 323–356.
- Iyer, R.P., Beaucage, S.L., 1999. Oligonucleotide synthesis. In: Barton, D., Nakanishi, K., Meth-Cohn, O., Kool, E.T. (Eds.), *Comprehensive Natural Product Chemistry*, vol. 7. New York, NY: Elsevier, pp. 105–152.

- Iyer, R.P., Phillips, L.R., Egan, W., Regan, J.B., Beaucage, S.L., 1990. The automated synthesis of sulfur-containing oligodeoxyribonucleotides using 3/*H*-1,2-benzodithiol-3-one 1,1-dioxide as a sulfur-transfer reagent. *Journal of Organic Chemistry* 55, 4693–4698.
- Kamer, P.C.J., Roelen, H.C.P.F., van den Elst, H., van der Marel, G.A., van Boom, J. H., 1989. An efficient approach toward the synthesis of phosphorothioate diesters via the Schönberg reaction. *Tetrahedron Letters* 30, 6757–6760.
- Khorana, H.G., Agarwal, K.L., Büchi, H., *et al.*, 1972. Studies on polynucleotides. 103. Total synthesis of the structural gene for an alanine transfer ribonucleic acid from yeast. *Journal of Molecular Biology* 72, 209–217.
- Köster, H., 1972. Polymer support oligonucleotide synthesis VI Use of inorganic carriers. *Tetrahedron Letters* 13 (16), 1527–1530.
- Lackey, J.G., Mitra, D., Somoza, M.M., Cerrina, F., Damha, M.J., 2009. Acetal levulinyl ester (ALE) groups for 2'-hydroxyl protection of ribonucleosides in the synthesis of oligoribonucleotides on glass and microarrays. *Journal of the American Chemical Society* 131, 8496–8502.
- Lavergne, T., Bertrand, J.-R., Vasseur, J.-J., Debart, F., 2008. A base-labile group for 2'-OH protection of ribonucleosides: A major challenge for RNA synthesis. *Chemistry – A European Journal* 14, 9135–9138.
- Lavergne, T., Janin, M., Dupouy, C., Vasseur, J.-J., Debart, F., 2010. Chemical synthesis of RNA with base-labile 2'-*O*-(pivaloyloxymethyl)-protected ribonucleoside phosphoramidites. In: Egli, M., Herdewijn, P., Matsuda, A., Sanghvi, Y.S. (Eds.), *Current Protocols in Nucleic Acid Chemistry*, vol. 3. Hoboken, NJ: Wiley, pp. 3.19.1–3.19.27.
- Letsinger, R.L., Lunsford, W.B., 1976. Synthesis of thymidine oligonucleotides by phosphite triester intermediates. *Journal of the American Chemical Society* 98, 3655–3661.
- Lloyd, W., Reese, C.B., Song, Q., *et al.*, 2000. Some observations relating to the use of 1-aryl-4-alkoxy-piperidin-4-yl groups for the protection of the 2'-hydroxy functions in the chemical synthesis of oligoribonucleotides. *Journal of the Chemical Society, Perkin Transactions 1*, 165–176.
- Li, Y., Lee, H.J., Corn, R.M., 2006. Fabrication and characterization of RNA aptamer microarrays for the study of protein-aptamer interactions with SPR imaging. *Nucleic Acids Research* 34, 6416–6424.
- Matteucci, M.D., Caruthers, M.H., 1981. Synthesis of deoxyoligonucleotides on a polymer support. *Journal of the American Chemical Society* 103, 3185–3191.
- McGall, G., Labadie, J., Brock, P., *et al.*, 1996. Light-directed synthesis of high-density oligonucleotide arrays using semiconductor photoresists. *Proceedings of the National Academy of Sciences of the United States of America* 93, 13555–13560.
- Moorcroft, M.J., Meuleman, W.R.A., Latham, S.G., *et al.*, 2005. In situ oligonucleotide synthesis on poly(dimethylsiloxane): A flexible substrate for microarray fabrication. *Nucleic Acids Research* 33, e75.
- Nagata, S., Hamasaki, T., Uetake, K., *et al.*, 2010. Synthesis and biological activity of artificial mRNA prepared with novel phosphorylating reagents. *Nucleic Acids Research* 38, 7845–7857.
- Ogilvie, K.K., Usman, N., Nicoghossian, K., Cedergren, R.J., 1988. Total chemical synthesis of a 77-nucleotide-long RNA sequence having methionine-acceptance activity. *Proceedings of the National Academy of Sciences of the United States of America* 85, 5764–5768.
- Pitsch, S., Weiss, P.A., Jenny, L., Stutz, A., Wu, X., 2001. Reliable chemical synthesis of oligoribonucleotides (RNA) with 2'-*O*-[(triisopropylsilyl)oxy]methyl (2'-*O*-tom)-protected phosphoramidites. *Helvetica Chimica Acta* 84, 3773–3795.
- Pon, R.T., Usman, N., Damha, M.J., Ogilvie, K.K., 1986. Prevention of guanine modification and chain cleavage during the solid phase synthesis of oligonucleotides using phosphoramidite derivatives. *Nucleic Acids Research* 14, 6453–6470.
- Pon, R.T., Yu, S., Prabhavalkar, T., *et al.*, 2005. Large scale synthesis of "Cpep" RNA monomers and their application in automated RNA synthesis. *Nucleosides & Nucleic Acids* 24, 777–781.
- Porcher, S., Pitsch, S., 2005. Synthesis of 2'-*O*-[(triisopropylsilyl)oxy]methyl (=tom)-protected ribonucleoside phosphoramidites containing various nucleobase analogues. *Helvetica Chimica Acta* 88, 2683–2704.
- Prakash, T.P., Kraynack, B., Baker, B.F., Swayze, E.E., Bhat, B., 2006. RNA interference by 2',5'-linked nucleic acid duplexes in mammalian cells. *Bioorganic & Medicinal Chemistry Letters* 16, 3238–3240.
- Reese, C.B., Yan, H., 2004. Alternatives to the 4,4'-dimethoxytrityl (DMTr) protecting group. *Tetrahedron Letters* 45, 2567–2570.
- Richmond, K.E., Li, M.-H., Rodesch, M.J., *et al.*, 2004. Amplification and assembly of chip-eluted DNA (AACED): A method for high-throughput gene synthesis. *Nucleic Acids Research* 32, 5011–5018.
- Saneyoshi, H., Ando, K., Seio, K., Sekine, M., 2007. Chemical synthesis of RNA via 2'-*O*-cyanoethylated intermediates. *Tetrahedron* 63, 11195–11203.
- Scaringe, S.A., 2001. RNA oligonucleotide synthesis via 5'-silyl-2'-orthoester chemistry. *Methods* 23, 206–217.
- Scaringe, S.A., Francklyn, C., Usman, N., 1990. Chemical synthesis of biologically active oligoribonucleotides using β -cyanoethyl protected ribonucleoside phosphoramidites. *Nucleic Acids Research* 18, 5433–5441.
- Scaringe, S.A., Kitchen, D., Kaiser, R.J., Marshall, W.S., 2004. Preparation of 5'-silyl-2'-orthoester ribonucleosides for use in oligoribonucleotide synthesis. In: Beaucage, S.L., Bergstrom, D.E., Glick, G., Jones, R.A. (Eds.), *Current Protocols in Nucleic Acid Chemistry*, vol. 1. New York, NY: Wiley, pp. 2.10.1–2.10.16.
- Shiba, Y., Masuda, H., Watanabe, N., *et al.*, 2007. Chemical synthesis of a very long oligoribonucleotide with 2-cyanomethoxymethyl (CEM) as the 2'-*O*-protecting group: Structural identification and biological activity of a synthetic 110mer precursor-microRNA candidate. *Nucleic Acids Research* 35, 3287–3296.
- Sierzchala, A.B., Dellinger, D.J., Betley, J.R., *et al.*, 2003. Solid-phase oligodeoxynucleotide synthesis: A two-step cycle using peroxy anion deprotection. *Journal of the American Chemical Society* 125, 13427–13441.
- Singh-Gasson, S., Green, R.D., Yue, Y., *et al.*, 1999. Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nature Biotechnology* 17, 974–978.
- Sinha, N.D., Biernat, J., McManus, J., Köster, H., 1984. Polymer support oligonucleotide synthesis XVIII: Use of (β -cyanoethyl-N,N-dialkylamino-/N-morpholino phosphoramidite of deoxynucleosides for the synthesis of DNA fragments simplifying deprotection and isolation of the final product. *Nucleic Acids Research* 12, 4539–4557.
- Stawinski, J., Thelin, M., 1994. Nucleoside H-phosphonates. 14. Synthesis of nucleoside phosphoroselenates and phosphorothioselenates via stereospecific selenization of the corresponding H-phosphonate and H-phosphonothioate diesters with the aid of new selenium transfer reagent, 3/*H*-1,2-benzothiaselenol-3-one. *Journal of Organic Chemistry* 59, 130–136.
- Stawinski, J., Krazewski, A., 1998. Some aspects of oligoribonucleotide synthesis via the H-phosphonate approach. *Acta Biochimica Polonica* 45, 907–915.
- Stawinski, J., Strömberg, R., 2001. Deoxyribo- and ribonucleoside H-phosphonates. In: Beaucage, S.L., Bergstrom, D.E., Glick, G., Jones, R.A. (Eds.), *Current Protocols in Nucleic Acid Chemistry*, vol. 1. New York, NY: Wiley, pp. 2.6.1–2.6.15.
- Stec, W.J., Zon, G., 1984. Stereochemical studies of the formation of chiral internucleotide linkages by phosphoramidite coupling in the synthesis of oligodeoxyribonucleotides. *Tetrahedron Letters* 25, 5279–5282.
- Strömberg, R., Stawinski, J., 2004. Synthesis of oligodeoxyribo- and oligoribonucleotides according to the H-phosphonate method. In: Beaucage, S.L., Bergstrom, D.E., Glick, G., Jones, R.A. (Eds.), *Current Protocols in Nucleic Acid Chemistry*, vol. 1. New York, NY: Wiley, pp. 3.4.1–3.4.15.
- Tian, J., Gong, H., Sheng, N., *et al.*, 2004. Accurate multiplex gene synthesis from programmable DNA microchips. *Nature* 432, 1050–1054.
- Tram, K., Sanghvi, Y.S., Yan, H., 2011. Further optimization of detritylation in solid-phase oligodeoxyribonucleotide synthesis. *Nucleosides, Nucleotides and Nucleic Acids* 30, 12–19.
- Usman, N., Pon, R.T., Ogilvie, K.K., 1985. Preparation of ribonucleoside 3'-*O*-phosphoramidites and their application to the automated solid phase synthesis of oligonucleotides. *Tetrahedron Letters* 26, 4567–4570.
- Uzagare, M.C., Padiya, K.J., Salunkhe, M.M., Sanghvi, Y.S., 2003. NBS-DMSO as a nonaqueous nonbasic oxidation reagent for the synthesis of oligonucleotides. *Bioorganic & Medicinal Chemistry Letters* 13, 3537–3540.
- Wei, X., 2013. Coupling activators for the oligonucleotide synthesis via phosphoramidite approach. *Tetrahedron* 69, 3615–3637.
- Westman, E., Strömberg, R., 1994. Removal of *t*-butyldimethylsilyl protection in RNA-synthesis. Triethylamine trihydrofluoride (TEA, 3HF) is a more reliable alternative to tetrabutylammonium fluoride (TBAF). *Nucleic Acids Research* 22, 2430–2431.
- Wolber, P.K., Collins, P.J., Lucas, A.B., De Witte, A., Shannon, K.W., 2006. The Agilent in situ-synthesized microarray platform. *Methods in Enzymology* 410, 28–57.

MOLECULAR PRINCIPLES, COMPONENTS, TECHNOLOGY, AND CONCEPTS: PROTEINS

Contents

Expression Systems

Isolation/Purification of Proteins

Protein Sequence Determination: Methodology and Evolutionary Implications

Posttranslational Modifications: Key Players in Health and Disease

Protein Domains: Structure, Function, and Methods

NMR in Structural and Cell Biology

Folding, Misfolding, Disordered Proteins, and Related Diseases

Diseases of Protein Folding: Huntington's Disease and Amyotrophic Lateral Sclerosis

Site-Directed Mutagenesis

Chemical Biology

Drug Design

Antibodies and Improved Engineered Formats (as Reagents)

Expression Systems

DI Fisher, AstraZeneca, Cambridge, UK

LM Mayr, AstraZeneca, Cheshire, UK

RG Roth, AstraZeneca, Mölndal, Sweden

© 2016 Elsevier Inc. All rights reserved.

Glossary

Baculovirus Member of a family of DNA viruses infectious to invertebrates. In the field of recombinant protein expression this usually refers to isolates of *Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV).

Episomal Referring to DNA fragments that are replicated and maintained extrachromosomally.

Heterologous protein Protein derived from an alternative organism or source to the chosen expression host.

Hypothermic shift A technique whereby the temperature of incubation is dropped below the standard used for growth. Reduction of the temperature has a negative effect on cell growth/division, but may have a positive effect on protein expression.

Overview of Expression Systems

Expression systems are used to express proteins that are used in a range of methods ranging from structural biology to *in vivo* studies or as therapeutic agents. An expression system essentially consists of three different components necessary for the production of recombinant proteins. Firstly, a biological environment, most often a cell, is used to provide energy and machinery for protein synthesis. It is also possible to use cellular extracts containing the necessary components, i.e., cell-free protein expression systems. Secondly, a vector that facilitates the introduction of genetic material into the cell; containing regulatory parts that provide for replication of the genetic material and usually selection markers for

maintenance. Lastly, the expression cassette, which is incorporated into the vector that contains the open reading frame encoding the amino acid sequence for the protein to be expressed. The expression cassette also contains all the components necessary for transcription and translation of the protein to be produced (**Figures 1 and 3**).

The expression host can be anything that has the necessary framework to produce protein from the genetic material provided. A large variety of organisms are used for the production of proteins, even transgenic animals or plants are used for protein production (Houdebine, 2009; Egelkrout *et al.*, 2012; Berlec and Strukelj, 2013). In this overview we focus on systems that are commonly used in a research setting, which are most commonly cellular systems (**Table 1**). A special

Table 1 Characteristics of common expression systems

Expression host		Transcription regulation (inducer)	Doubling time (h)	Vectors	Glycosylation	PTM	Common use
Bacteria (simple, robust, lowest cost)	<i>Escherichia coli</i>	PT7, Ptac (IPTG)	0.3–0.5	Plasmid (episomal)	No	No	Small proteins Single domains Antigens
	Other Enterobacteria	PBAD (arabinose) PL (switch to 42 °C) Ptac (IPTG) PalkBp (sodium salicylate) Pm/xyIS (toluinic/benzoic acid) PxylA (xylose) PT7 (IPTG)	1.0–2.0	Plasmid (episomal)	No	No	Specific receptors Fimbrial proteins Heme proteins
	Firmicutes	<i>Bacillus</i> spp.	0.6–1.0	Plasmid (episomal)	No	No	Secreted enzymes Virulence factors
	Unicellular fungi (yeast)	PAOX1 (methanol) PFLD (acetamide) PGAP (constitutive)	1.5–3.0	Plasmid (integrative)	High	Some	Secreted proteins Complement factors Membrane proteins Peroxisome proteins Kinases
Eukarya (glycosylation, PTMs)	Non-methylotrophic (<i>Saccharomyces cerevisiae</i> , <i>Kluyveromyces fragilis</i>)	PGAL1/PGAL10 (galactose) Acetamidase (acetamide) Glycolytic (constitutive)	1.3–2.0	Plasmid (episomal/integrative)	High	Some	Nuclear proteins
	Filamentous fungi	<i>Aspergillus</i> spp. <i>Hypocrea</i> <i>Trichoderma</i> spp.	3.0–4.0	Plasmid (episomal/integrative)	High	Some	Membrane proteins Antibodies Secreted enzymes Glycoenzymes
	Protozoa	<i>Leishmania tarentolae</i>	6.0–13	Plasmid (integrative)	Complex	High	General purpose
	Insect cells	Sf9, Sf21, High Five™ BTI-TN-5B1-4, S2	16–72	Plasmid (episomal) Virus	Complex	Yes	General purpose
Mammalian cells		pMT (copper/cadmium) pAc5 (constitutive) SV40, CMV (constitutive)	14–36	Plasmid (episomal/integrative) Virus	Complex	Yes	Secreted proteins Antibodies

Source: Adapted from Fernandez, F.J., Vega, M.C., 2013. Technologies to keep an eye on: Alternative hosts for protein production in structural biology. Current Opinion in Structural Biology 23 (3), 365–373.

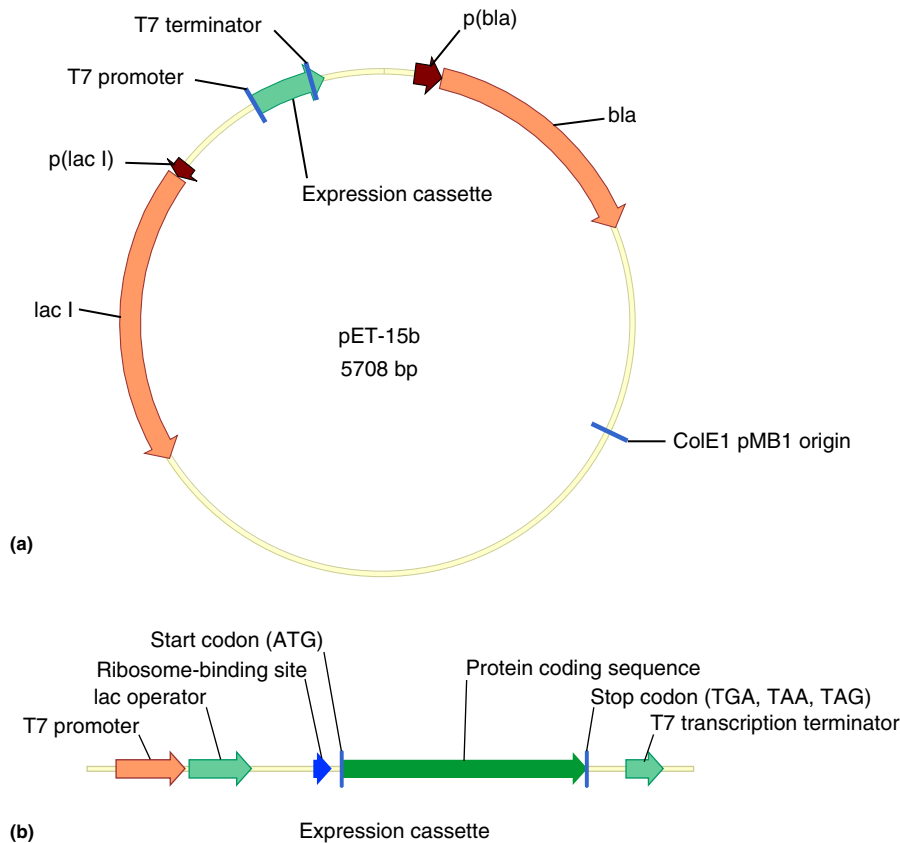


Figure 1 Schematic of pET-15b used as a vector for heterologous protein expression in *E. coli*. (a) The *lacI* gene provides repressor protein to titrate the *lac* operator controlling expression from the T7 promoter. For selection the plasmid contains the *bla* gene encoding beta-lactamase providing resistance to ampicillin. Also depicted is the position of the origin of replication (ColE1 pMB1) and the expression cassette. (b) Contents of the expression cassette in pET-15b. T7 RNA polymerase initiates transcription at the T7 promoter. Additional control of transcription is provided by the *lac* operator where binding of *lac* repressor prevents transcription from the T7 promoter. The ribosome-binding site is located upstream of the open reading frame containing the protein coding sequence between the start and stop codons. The expression cassette ends with the T7 transcription termination sequence where T7 RNA polymerase stops and dissociates from the plasmid DNA sequence.

case is the cell-free expression systems that are described in a separate section. Hosts are commonly divided into prokaryotic or eukaryotic since it broadly describes the characteristics of the expression system. For detailed properties of each system, read the subsequent sections detailing the different host systems.

The most commonly used vector for recombinant protein expression is plasmid DNA. Many organisms carry plasmids which are extrachromosomal, self-replicating, circular pieces of DNA. Plasmids are used as vectors to introduce genetic material since they can be readily purified, manipulated, and re-introduced into several host organisms. Other vectors that can be used to carry the genetic information required to guide protein production are viruses or linear DNA fragments.

Bacterial Expression Systems

Bacterial expression systems are used for their ease of genetic manipulation, rapid growth, and the ability to grow in low cost, complex, or chemically defined media. The gram-negative bacterium *Escherichia coli* is widely used for recombinant protein production. *E. coli* has been used as a model organism

for over 60 years and besides the complete characterization of its genome sequence there is also a wealth of knowledge about its biochemistry and molecular biology. A couple of years after the start of the molecular biology revolution in the 1970s it was the first host for heterologous protein expression (Cohen *et al.*, 1973) and it is still one of the most important organisms for recombinant protein production. In *E. coli* the recombinant gene is almost always expressed from a plasmid and not stably integrated into the host cell chromosome. The common features present in a plasmid used for recombinant protein expression in *E. coli* are shown in Figure 1.

The number of plasmid copies within the cell can range from one to several hundreds. The copy number of plasmids is a combined property of the replicon present in the plasmid, the cell, and its current metabolic state. The ColE1-family of replicons are used in popular vectors system such as pET (pMB1, 15–60 copies/cell) and pUC (pMB1 variant, 500–700 copies/cell). Gene expression can be dependent on gene dosage although it is much more common that the limiting factor for protein production is in processes downstream of transcription (Novoa and Ribas de Pouplana, 2012; Gustafsson *et al.*, 2012). A very high plasmid copy number and high level of transcription will impose a metabolic burden on the cells

Table 2 Solubility-enhancing protein domains

Tag	Protein and source organism	Mw (kDa)	IEP
PDI	Protein disulfide isomerase (<i>Homo sapiens</i>)	55.4	4.4
NusA	N-utilization substance (<i>E. coli</i>)	54.9	4.3
MBP ^a	Maltose-binding protein (<i>E. coli</i>)	40.8	5.0
T7PK	N-terminal domain of T7 protein kinase (Bacteriophage T7)	27.6	4.5
GST ^a	Glutathione-S-transferase (<i>Schistosoma japonicum</i>)	25.5	6.5
DsbC	Protein disulfide isomerase (<i>E. coli</i>)	23.5	6.3
IF2 I	N-terminal domain of initiation factor 2 (<i>E. coli</i>)	17.7	9.8
ZZ ^a	Protein A IgG binding double domain (<i>Staphylococcus aureus</i>)	14.5	4.6
SUMO	Small ubiquitin like modifier (<i>H. sapiens</i>)	10.9	5.0
Trx	Thioredoxin (<i>E. coli</i>)	11.8	4.6
Ub	Ubiquitin (<i>H. sapiens</i>)	8.6	7.5
GB1 ^a	Protein G Ig binding domain B1 (<i>Streptococcus</i> sp.)	6.5	4.4
Rbx	Rubredoxin (<i>Thermotoga maritima</i>)	6.0	3.9

^aThese tags can also be used for direct affinity purification without adding an additional affinity tag.

Note: Solubility-enhancing domains including their molecular weight (Mw) and isoelectrical point (IEP) calculated from the native tag sequence without any adjacent amino acids.

lowering the growth rate, rate of translation and hence negatively influence the production of the recombinant protein. Common alternative replicons that are present at lower copy number and do not belong to the same incompatibility group are p15A (~10 copies/cell) and pSC101 (~5 copies/cell).

Recombinant gene expression in *E. coli* is normally under the control of an inducible promoter that requires the presence of a specific metabolite or synthetic inducer to direct transcription of the gene of interest. Several different promoters are in use with the most common being lac/tac, lacUV5, araBAD, and rhaBAD (Lutz and Bujard, 1997; Guzman *et al.*, 1995; Giacalone *et al.*, 2006).

For the popular pET expression system in *E. coli*, the expression induced by the addition of inducer is not of the actual recombinant gene but instead of a viral RNA polymerase from bacteriophage T7 that in turn recognizes the promoter controlling the recombinant gene. The T7 RNA polymerase, present in the IDE3 prophage integrated into the bacterial chromosome of the commonly used BL21(DE3) strain, is under the control of a lacUV5 promoter. Upon induction with lactose or isopropyl- β -D-1-thiogalactopyranoside (IPTG; a synthetic, non-metabolizable lactose analogue) the expressed T7 RNA polymerase will recognize the strong T7 promoter which controls recombinant gene expression on the introduced plasmid (Studier, 2005). The main benefit of this system is very strong transcription of the recombinant gene, since the T7 RNA polymerase has very high processivity and only recognizes the T7 promoter present in the pET plasmid and not endogenous promoters in *E. coli* (Studier *et al.*, 1990).

Heterologous protein production of eukaryotic genes in *E. coli* (and other prokaryotes) has a potential drawback as many posttranslational modifications present in eukaryotes are not performed in these organisms (including glycosylation, myristylation, isoprenylation, and amidation). The bacteria have a reductive environment in the cytoplasm preventing formation of disulfides necessary for the structure of many extracellular proteins from eukaryotes. It is also frequently seen in bacteria that the protein is expressed in intracellular aggregates called inclusion bodies. These inclusion bodies can be isolated, denatured, and then refolded to produce protein that is active

and of the correct conformation. This process has been used extensively in the past to produce small, single domain proteins and also with some success for larger, disulfide-bonded proteins (Mayer and Buchner, 2004). But from an expression standpoint, inclusion body formation is most often to be avoided due to the extra effort required to obtain functional proteins in their native state. There are a number of strategies described to increase the amount of soluble recombinant protein produced (Correa and Opezzo, 2015). Methods to reduce inclusion body formation include techniques that reduce the rate of protein translation (reduced growth temperatures or weaker induction of expression) or improve the folding of the recombinant protein by co-expression of chaperones. Another very common strategy to both increase the overall yield but also improve the solubility of the expressed protein is to attach protein domains that are known to be soluble and express well in the selected expression system (Table 2; Walls and Loughran, 2011).

Many of the alternative prokaryotic expression systems available have been developed to circumvent deficiencies in the standard *E. coli* systems. The gram-negative *Pseudomonas fluorescens* has been engineered to secrete protein to the periplasmic and extracellular space with high efficiency and is suitable for high-density fermentations (Retallack *et al.*, 2012).

Looking at gram-positive bacteria there are several expression systems that have been used for heterologous protein production with success. Many of these bacteria secrete proteins naturally with high efficiency which enables high protein production and facilitates downstream protein purification. The most commonly used species is *Bacillus subtilis*, usually variants that have been modified to lack extracellular proteases to increase the stability of the produced proteins (Pohl and Harwood, 2010). A more recent addition to the available gram-positive expression systems is *Corynebacterium glutamicum* (commercialized by Ajinomoto as CorynexTM) that can secrete both folded proteins through the twin arginine translocation (TAT) pathway and proteins that fold in the extracellular space through the secretion translocase (SEC) pathway (Kikuchi *et al.*, 2009).

There are bacterial expression systems with more specific utility for certain protein classes. One such example is

Lactococcus lactis that has been shown to be well suited to secrete proteins but also to express membrane proteins (Villatoro-Hernandez *et al.*, 2012; Kunji *et al.*, 2003). Selection is facilitated by the use of auxotrophic strains and there is an inducible variant (P170) that is activated by the accumulation of lactose during fermentation. Other benefits of the gram-positive bacteria mentioned here are that they are GRAS (Generally Regarded As Safe) and in contrast to gram-negative bacteria do not produce endotoxins. Endotoxins are lipopolysaccharides that are synthesized by gram-negative bacteria as components of the outer cell membrane and elicit an immune response when used *in vivo* in animals and humans.

Yeast Expression Systems

Yeasts are single cell microbial eukaryotes that have been used in food production by humans for thousands of years. The classic Baker's yeast, *Saccharomyces cerevisiae* was the first yeast to be used for recombinant protein production, but in the last 10 years *Pichia pastoris* has rapidly gained popularity and is a common choice today (Mattanovich *et al.*, 2012). Other yeasts that are increasingly used for heterologous protein production are *Kluyveromyces lactis* and *Pichia angusta* (formerly known as *Hansenula polymorpha*) (van Ooyen *et al.*, 2006; Gellissen *et al.*, 2005).

For the expression of heterologous protein in *S. cerevisiae* there are options for both episomal and integrative vectors. In both cases the promoters that direct expression of the gene of interest is either constitutive (usually from the glycolytic pathway) or inducible (PGAL1/PGAL10). Episomal vectors in *S. cerevisiae* make use of 2 μ , which is a 6.3 kbp plasmid that can be extrachromosomally maintained at 50–100 copies in the nucleus.

Pichia pastoris and *P. angusta* are both methylotrophic yeasts and expression systems based on the strong promoters induced by methanol have been developed for both systems (Cregg *et al.*, 2000; Gellissen *et al.*, 2005). The PAOX1 promoter for alcohol oxidase 1 in *P. pastoris* is very strong and can drive expression of heterologous proteins up to several grams per liter. Using methanol for induction has some technical safety complications since it is flammable and toxic. Other gene regulatory sequences have been adapted for use in recombinant protein production in *P. pastoris* and both inducible (PFLD; acetamide) and constitutive (PGAP) promoters are available (Cos *et al.*, 2006).

Protein expression in the methylotrophic yeasts are always driven from vectors that have been integrated into chromosomal DNA by homologous recombination. Strains which have vectors stably integrated are commonly selected through antibiotic selection with Zeocin or with *ade2* auxotrophs (commercialized by Life Technologies as *Pichia Pink*TM). Cytoplasmic proteins can be expressed successfully in *P. pastoris* although it is more frequently used to secrete proteins. Complex eukaryotic proteins with multiple disulfide bridges have been expressed in *P. pastoris*. To enable secretion the pre-pro sequence from the α -mating factor of *S. cerevisiae* is introduced at the N-terminus of the protein to be expressed. The methylotrophic yeasts are especially well suited for high-density cultivation and, in bioreactors, can reach over 120 g of dry weight per liter.

As eukaryotes, yeasts are capable of most posttranslational modifications including glycosylation. Depending on the use of the protein product, care should be taken since there are significant differences in the type of glycans added to proteins by yeast compared to mammalian cells. Glycans added to the nitrogen of asparagine in the sequence Asn-Xaa-Ser/Thr (N-glycosylation) are similar to what is seen in mammalian cells. But in yeasts the produced N-glycosylation structures are different from what is seen in mammalian cells and are very rich in mannose, an effect termed hypermannosylation. In *S. cerevisiae* these chain are much longer (50–150) than what is usually seen in *P. pastoris* (8–17). Glycosylation of the O-type on serines/threonines also occur in *P. pastoris* but is less characterized.

Besides the yeasts described above, other unicellular microbial eukaryotes used for recombinant protein production are filamentous fungi such as *Aspergillus spp.* and *Trichoderma spp.* (Ward, 2012) and protozoans including *Leishmania tarentolae* (Niimi, 2012). Their use for heterologous protein production in research laboratories has so far been limited. The interested reader can find more information about these systems in the suggested further reading.

Cell-Free Expression Systems

The cell-free expression systems are based on cell extracts that contain all the necessary macromolecular components for translation of protein from messenger RNA (mRNA). Rabbit reticulocytes and wheat germ cells are two common eukaryotic sources for extracts that have high intrinsic capacity for protein production (Bernhard and Tozawa, 2013; Whittaker, 2013). When using extracts from *E. coli* the bacteria have been cultured at very high specific growth rates which increases the ribosome concentration and hence the translation capacity of the extract. During the preparation of the extract, most low molecular weight molecules are lost and must be supplemented back to enable efficient translation; this includes amino acids, energy sources, cofactors, and energy regenerating systems.

Eukaryotic extracts are isolated from the cytoplasmic fraction and lack endogenous RNA polymerase activity. Instead exogenous RNA polymerases (from bacteriophages T7 or SP6) are normally added, and the corresponding promoter is incorporated into the DNA template. For extracts from *E. coli*, the endogenous RNA polymerase activity can be utilized but often the extract is made from strains containing the DE3 prophage and hence can express T7 RNA polymerase. The DNA used as templates in the cell-free expression systems can be added either as plasmids or linear fragments; this enables rapid screening of protein variants directly from polymerase chain reaction products.

In the first generation of cell-free expression systems all necessary components were directly mixed in the reaction chamber and the production stopped once a required component was depleted. The current methodology is based on continuous-exchange cell-free systems where fresh precursors are provided and waste products are removed across a dialysis membrane to enable much longer reaction times (Stech *et al.*, 2014).

The unique characteristic of cell-free systems is that they are fast since they are not dependent on the growth characteristics

of any biological system. It is also an open system which means that it is possible to directly add to the reaction unnatural or isotopically labeled amino acids (Ozawa *et al.*, 2012). Since the reaction is independent of the viability or growth of any cell, the expression of toxic proteins or incorporation of unnatural amino acids can be more efficient. Since most enzymatic activities from the cell are suppressed and the reaction is fast, proteins that are sensitive to proteolytic degradation can be isolated intact. Also membrane proteins are frequently expressed in cell-free systems, either directly into detergent micelles or as protein precipitates that are easily isolated and subsequently solubilized (Nozawa *et al.*, 2010). The downside of cell-free systems is that they are expensive to run at scale and have therefore been limited to mostly screening applications.

Insect Expression Systems

Moving on in complexity from the single-celled microbial expressions systems discussed earlier, we reach higher order eukaryotic expression systems. Insect cells have long been considered an obvious choice for expression of proteins that are not suitable for expression in microbial systems. These cells have some important benefits over prokaryotes in terms of expression complexity. They are able to process posttranslational modifications that may be required by heterologous recombinant proteins, including signal peptide cleavage, glycosylation, myristylation, sulfation, amidation, γ -carboxylation, and β -hydroxylation. However, whilst insect cells are capable of expressing N-linked glycoproteins, the chains formed are significantly different from those seen in mammalian cells (Jarvis *et al.*, 1998; Kim *et al.*, 2013). As shown in Table 3, insect glycosylation chains are limited to high mannose glycans or glycans with trimannosyl cores as they lack sufficient glycosyltransferases needed to process complex chains. Engineered cell lines have been developed that attempt to mimic more mammalian-like glycosylation patterns (Jarvis *et al.*, 1998). Life Technologies have licensed a humanized glycosylation cell line from GlycoBac and have commercialized it as the Mimic™ Sf9 cell line. Insect cells are naturally able to secrete extracellular proteins through the endoplasmic reticulum (ER) and Golgi apparatus. However, typical yields of recombinant proteins are usually several-fold lower than the microbial systems.

Based on data available from the Worldwide Protein Data Bank, insect cells appear to be the preferred eukaryotic

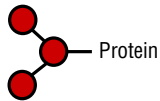
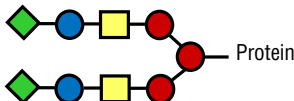
expression system (Figure 2), and second overall expression system after *E. coli*, for producing proteins for structural studies. This is at least partially due to the fact the system is amenable to fairly high-throughput expression screening. Of the immortalized insect cells available, the most commonly utilized lines are derived from *Spodoptera frugiperda* (fall armyworm) with Sf9 and Sf21 being the most common. Expression systems involving these lines usually utilize baculoviruses as the vector for delivery of recombinant genes. In addition to the *spodoptera* lines, cells derived from *Trichoplusia ni* (cabbage looper) are also used with baculoviral-mediated delivery (High Five™ BTI-TN-5B1-4 and Tn-368). These cells were originally developed and commercialized to solve a perceived failing of *spodoptera*-derived lines in expressing high levels of secreted proteins (Wickham *et al.*, 1992).

Generation of recombinant baculoviruses containing the heterologous gene of interest is commonly achieved by either *in vivo* transposition in *E. coli* or direct recombination in insect cells. The most common baculovirus isolates used for heterologous recombinant protein expression are from *Autographa californica* (alfalfa looper) multiple nuclear polyhedrosis virus (AcMNPV) and less commonly from *Bombyx mori* (silkworm) nuclear polyhedrosis virus (BmNPV). In almost all cases, the gene of interest is inserted downstream of the baculoviral polyhedrin promoter, in place of the endogenous polyhedrin gene; this gene is not required for baculoviral-mediated infection of cultured cells. If multiple heterologous genes are to be expressed from the same baculovirus, the p10 promoter is commonly used. There are many systems that have been commercialized to perform the modification of the baculoviral genome for the purpose of generating recombinant proteins, the most common are listed in Table 4. Once a high-titer virus has been established, this can be used to express recombinant proteins in addition to propagation of further virus generations.

Due to the lytic nature of the viral expression system, heterologous recombinant protein generation in these systems is relatively quick; commonly no longer than three days post infection. However, the lytic nature must also be taken into account when considering the heterologous protein of interest; intracellular proteins must be harvested before the lytic cycle has progressed too far, preventing harvesting of the whole cells, and any proteins susceptible to proteolysis require harvesting before they have been overexposed to proteases released during lysis.

As an alternative to the baculoviral-mediated methods, traditional transfection systems have been developed.

Table 3 Differences between insect and mammalian glycosylation

	Insect cells	Mammalian cells
Enzymes produced	Low levels of glycosylation enzymes required to produce complex glycoproteins	Working levels of glycosylation enzymes required to produce complex glycoproteins
Resulting glycoprotein sidechain	Paucimannose or high mannose glycan	Complex N-glycans with terminal sialic acid
Glycoprotein sidechain structure		

Note:  N-Acetylglucosamine;  Mannose;  Galactose;  Sialic acid.

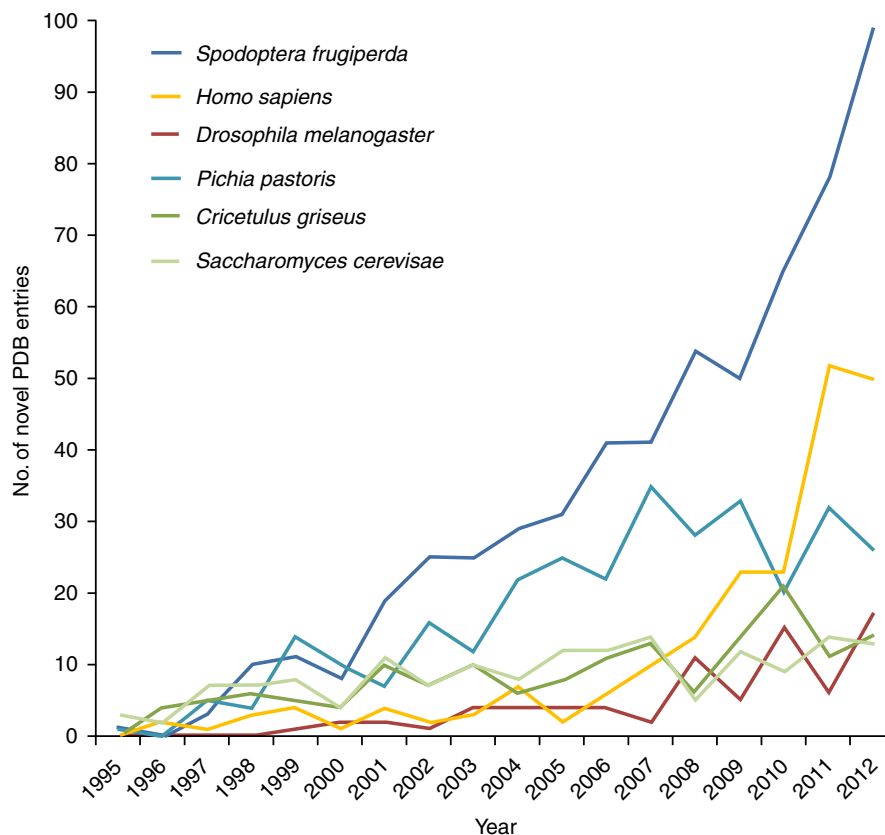


Figure 2 Impact of eukaryotic expression systems on structural biology. The graph shows the number of new, unique entries added to the Protein Data Bank (PDB) per year (unique is defined here as entries with <90% sequence identity). Note that *S. frugiperda* equates to Baculovirus expression, whilst *Homo sapiens* corresponds primarily to HEK293 cells (PDB was searched using '*Homo sapiens*' as a search term for expression system, but no attempt was made to quantify the exact number of entries using HEK293 cells specifically). Reproduced with permission from Assenberg, R., Wan, P.T., Geisse, S., Mayr, L.M., 2013. Advances in recombinant protein expression for use in pharmaceutical research. *Current Opinion in Structural Biology* 23 (3), 393–402.

Traditional transient transfection of *spodoptera* cells is used in the InsectDirect™ (BD Biosciences) transfection system. Additionally, systems using Schneider 2 (S2) cells derived from *Drosophila melanogaster* (fruit fly) embryos have been developed as both transient and stable expressions systems. Unlike the viral-mediated system described above, expression of heterologous protein is achieved via traditional transfection of the cells with a suitable plasmid vector. Stable lines may be developed using antibiotic resistance markers to enable longer term expression. Inducible expression can be achieved in the S2 system by utilizing an inducible promoter, such as the metallothionein promoter which can be induced by the presence of heavy metals (usually copper or cadmium). Alternatively, constitutive expression can be achieved by using a constitutive promoter such as the actin 5C promoter (Angelichio *et al.*, 1991).

For details of transient transfection methodologies, see Section Mammalian Expression Systems.

Mammalian Expression Systems

Mammalian expression systems provide for the most comprehensive extent of posttranslational modifications. However,

this is tempered by the fact that, until recently, yields from these systems have been much lower than any of the cellular alternatives already discussed. Improvements in this area have been boosted by the fact that many biopharmaceuticals are expressed in these systems (Wurm, 2004) giving rise to a strong driver to improve yields (Durocher *et al.*, 2002; Dar-amola *et al.*, 2014). Like the insect systems described in Section Insect Expression Systems, mammalian cells are also naturally able to cleave signal peptides and secrete extracellular proteins.

Delivery of recombinant genes into mammalian cells has traditionally been achieved through transfection of the cells with plasmid DNA (Figure 3), either through binding of the DNA to a suitable transfer molecule or electroporation. There are a variety of transfer molecules available to traffic DNA across the cell membrane to which it is naturally impermeable, and the most commonly used for recombinant protein expression are summarized in Table 5. Additional methods include the use of viral delivery systems, including lentivirus and a modified baculovirus system (BacMam). In all cases the vectors utilize a promoter suitable for over-expression in mammalian cells, often a constitutive viral promoter such as the cytomegalovirus (CMV) immediate early promoter.

Table 4 Common commercially available baculovirus expression systems

Baculoviral-mediated expression system	Notes
Bac-to-Bac™; Life Technologies (ThermoFisher)	Heterologous gene insertion achieved via <i>in vivo</i> transposition in DH10Bac™ <i>E. coli</i> . Purification of bacmid DNA via alkali lysis and subsequent transfection of insect cells yields budding virus particles
MultiBac™; Geneva Biotech	Specifically designed to allow for multiple heterologous genes to be expressed from a single viral genome. Heterologous gene insertion achieved via <i>in vivo</i> transposition in DH10MultiBac™ <i>urbo</i> or EmBacY <i>E. coli</i> . These strains contain bacmids with <i>v-cath</i> and <i>chiA</i> knockouts. Purification of Bacmid DNA via alkali lysis and subsequent transfection of insect cells yields budding virus particles
BacMagic™; Merck Millipore	Heterologous gene insertion achieved via <i>in vivo</i> recombination in insect cells. BacMagic1 is a <i>chiA</i> knockout, BacMagic2 is a <i>chiA</i> and <i>v-cath</i> knockout, whilst Bacmagic3 is a <i>chiA</i> , <i>v-cath</i> , p10, p24, and p26 knockout. Transfection of insect cells with a viral genome in addition to the appropriate donor vector leads to direct production of budding virus particles
BaculoGold™; BD Biosciences	Heterologous gene insertion achieved via <i>in vivo</i> recombination in insect cells. Transfection of insect cells with the viral genome in addition to the appropriate donor vector leads to direct production of budding virus particles
FlashBAC™; Oxford Expression Technologies	Heterologous gene insertion achieved via <i>in vivo</i> recombination in insect cells. FlashBACGOLD is a <i>chiA</i> and <i>v-cath</i> knockout. FlashBACULTRA is a <i>chiA</i> , <i>v-cath</i> , p10, p24, and p26 knockout. Transfection of insect cells with a viral genome in addition to the appropriate donor vector leads to direct production of budding virus particles
BaculoDirect™; Life Technologies (ThermoFisher)	Heterologous gene insertion achieved via <i>in vitro</i> recombination. Subsequent transfection of insect cells yields budding virus particles

Note: Table shows the major baculoviral-mediated expression systems available, including notes on the individual features of each.

The most commonly used mammalian expression cell lines are derived from *Homo sapiens* (human) and rodent species, notably *Cricetulus griseus* (Chinese hamster) and *Mus musculus* (house mouse). The most well-known mammalian expression line is the human-derived HEK293 cell. Developed in 1973 by transformation of an embryonic kidney culture with sheared adenovirus DNA, it became popular due to the ease at which it is cultivated and transfected by many of the traditional transfection methods. Increasingly, systems based on Chinese hamster ovary lines (CHO K1 or CHO DG44) are becoming more popular. These lines have traditionally been more difficult to transfect, but are able to reach higher cell densities than standard HEK293 cells and are amenable to a variety of cell culture manipulations, including hypothermic shift and treatment with histone deacetylase inhibitors, that contribute to higher protein yields (Wulhfard *et al.*, 2010). More recently, efforts have been made to derive new mammalian expression hosts, from novel progenitor cell lines. For instance, the CAP systems from CEVEC Pharmaceuticals use human amniocyte-derived cells from excess cells obtained during amniocentesis.

The simplest mammalian expression systems involve introducing plasmid DNA to the cells and allowing the cells to produce protein (Figure 4) until the plasmid is naturally lost. These systems are known as transient methodologies due to the limited expression window. As mentioned previously, plasmids are self-replicating, extrachromosomal, closed circular DNA fragments. Although plasmids can be replicated in mammalian cells, they are not automatically maintained with chromosomal DNA during cell division. This means that as the cells divide they are prone to lose the plasmid DNA as it requires re-trafficking to the nucleus. Since this results in a fairly short period of expression, usually peaking between 24 h and 72 h, several approaches have been used to extend the expression window with the intent of improving the recombinant protein expression yield.

The most obvious way to improve the expression window is to integrate the heterologous gene into the host cell's chromosomal DNA. At this point you are generating a stable cell line that will continue to express the protein of interest so long as the heterologous gene remains in a section of open chromatin accessible to RNA polymerase. The downside to this approach is that it usually takes a significant amount of upstream effort to prepare the cell line prior to expression of your heterologous protein. Generation of a stable cell line for expression can be carried out via a variety of methods, including viral delivery and linearized plasmid delivery, though the end result is essentially the same; incorporation of the heterologous gene along with a selective marker that can be used to enrich for your integrants.

Since stable cell-line generation is a lengthy process, much effort has continued into improving transient methodologies for research purposes. There have been four main areas of focus to this end: improve transfection efficiency to ensure a greater population of expressing cells, improve the productivity of your cells, improve the window of expression, and improve the density of your culture to increase the volumetric yield.

As described in Table 5, there are a number of commercially available reagents that have been designed and optimized to efficiently transport DNA across the cell membrane. Factorial experimental design approaches are routinely used to determine the conditions for which maximal transfection efficiency can be achieved by these reagents. However, often these reagents suffer from inherent toxicity issues and efforts have been made to investigate alternative delivery systems. The large-scale electroporation devices from MaxCyte®, such as the MaxCyte STX®, have been shown to provide for excellent transfection efficiencies and reports suggest this method can be used for recombinant protein generation. Additionally, BacMam has been successfully used for recombinant protein

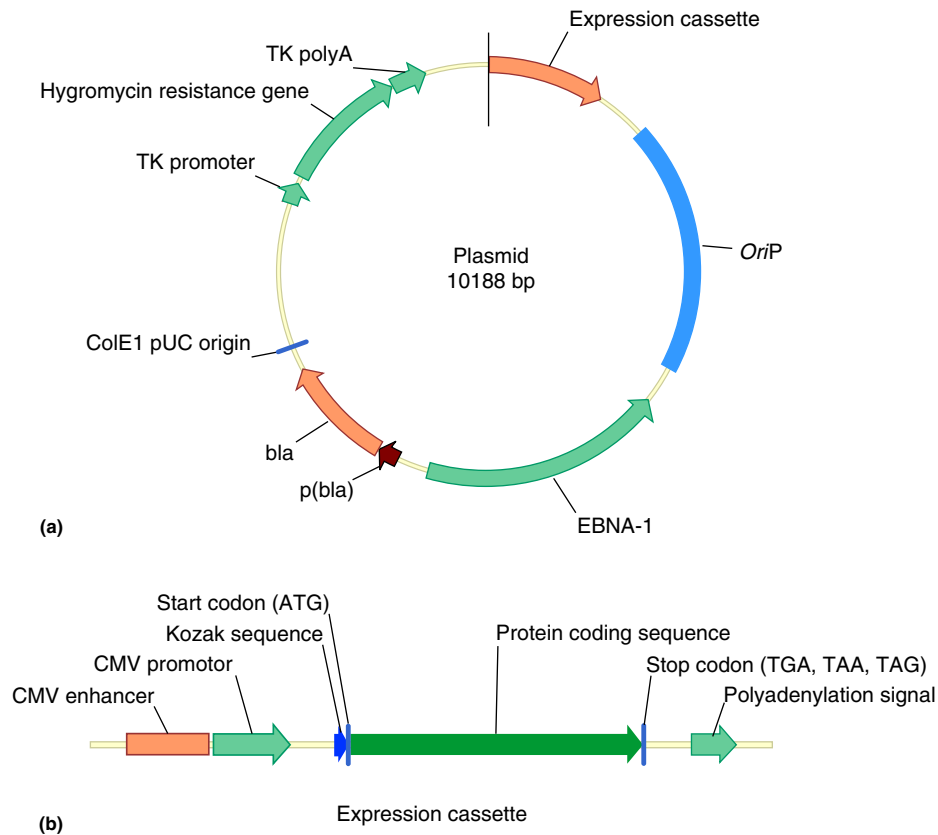


Figure 3 Schematic of a plasmid used as a vector for recombinant protein expression in mammalian cells. (a) The plasmid contains elements to enable episomal replication or stable cell-line generation. The EBNA-1 gene provides EBNA-1 protein to enable episomal replication of the plasmid via the *oriP* sequence. For stable cell-line generation, the plasmid contains a gene conferring resistance to hygromycin B. For selection in an *E. coli* propagation strain the plasmid contains the *bla* gene encoding beta-lactamase providing resistance to ampicillin. Also depicted is the position of the origin of replication (ColE1 pUC) and the expression cassette. (b) Contents of the expression cassette. The CMV enhancer and promoter region are required for initiation of transcription in mammalian cells. The Kozak sequence exists adjacent to the start codon (ATG) and assists in initiation of translation. The heterologous gene of interest lies in the protein coding sequence between the start and stop codons. After the gene comes the polyadenylation signal which codes for the mRNA poly(A) tail. The poly(A) tail is required for the nuclear export, translation, and stability of the transcript.

expression in mammalian cells (Ames *et al.*, 2007). BacMam uses the same baculoviral-mediated technology described in Section Insect Expression Systems and is a modification of the AcMNPV baculoviral isolate. Baculoviruses do not infect or replicate in mammalian cells, however, it was discovered that they are able to traffic their genetic content into mammalian cells under normal cell culture conditions via a method termed 'transduction.' By utilizing an appropriate mammalian-effective promoter in place of the polyhedrin promoter it is possible to generate virus particles that act as heterologous gene delivery vehicles into mammalian cells.

To improve the productivity of the cells, media, supplements, and nutrient feeds have been continually revised and optimized. There are a number of commercially available feed products that have been designed to boost expression yields by providing a variety of nutrients and energy sources, the most common of which are noted in Table 6. Additionally, work has focused on molecules that may affect cell biology in such a way as to boost expression levels. Aliphatic acid-type histone deacetylase inhibitors, such as butyric and valproic

acids, have been added to cultures as expression enhancers (Wulhfard *et al.*, 2010). The mechanism by which these molecules confer their performance-enhancing effect is not fully understood, but by preventing deacetylation of histones the molecules prevent the condensing of chromatin structures, thus keeping DNA more accessible. It is unclear, however, why this activity should confer enhancement to non-chromosomally located heterologous genes. Life Technologies' Expi293™ expression system makes use of proprietary enhancer supplements added post transfection to improve productivity.

Improving the expression window can be achieved by ensuring the plasmid containing the heterologous gene is maintained during cell division. Episomal expression systems have been developed and commercialized for this purpose. These systems make use of the machinery used by viruses in latent phases of their cycle. By including a viral origin and DNA-binding sequence on the expression plasmid and a source of the appropriate DNA-binding protein, for example the *oriP* origin and repeat binding sequence and EBNA-1 protein from

Table 5 Common transfection reagents for recombinant protein expression

Transfection reagent	Type	Notes
Calcium phosphate (usually as calcium chloride and HEPES buffered saline)	Salt precipitant	Co-precipitates of DNA and calcium phosphate are able to bind to the cell surface and enter via endocytosis. It is an extremely low-cost method, but very cell type dependent and time consuming to perform at scale
Polyethylenimine (PEI)	Cationic polymer	PEI condenses DNA into particles that can bind to the cell surface. Again DNA entry is via endocytosis. PEI is very low cost and therefore economical at scale. However, it is also extremely cytotoxic
JetPEI®; Polyplus transfection	Cationic polymer	Commercial, linear PEI derivative transfection reagent
Freestyle™MAX; Life Technologies (ThermoFisher)	Proprietary	Commercial reagent used in the Freestyle™ expression system. The system comprises transfection reagent, optimized media, and cell lines derived from HEK293 or CHO
Expiectamine™293; Life Technologies (ThermoFisher)	Proprietary	Commercial reagent used in the Expi293™ expression system. The system comprises transfection reagent, transfection enhancers, optimized media, and a cell line derived from HEK293F

Note: Table shows commonly utilized transfection reagents for recombinant protein expression from mammalian cell lines.

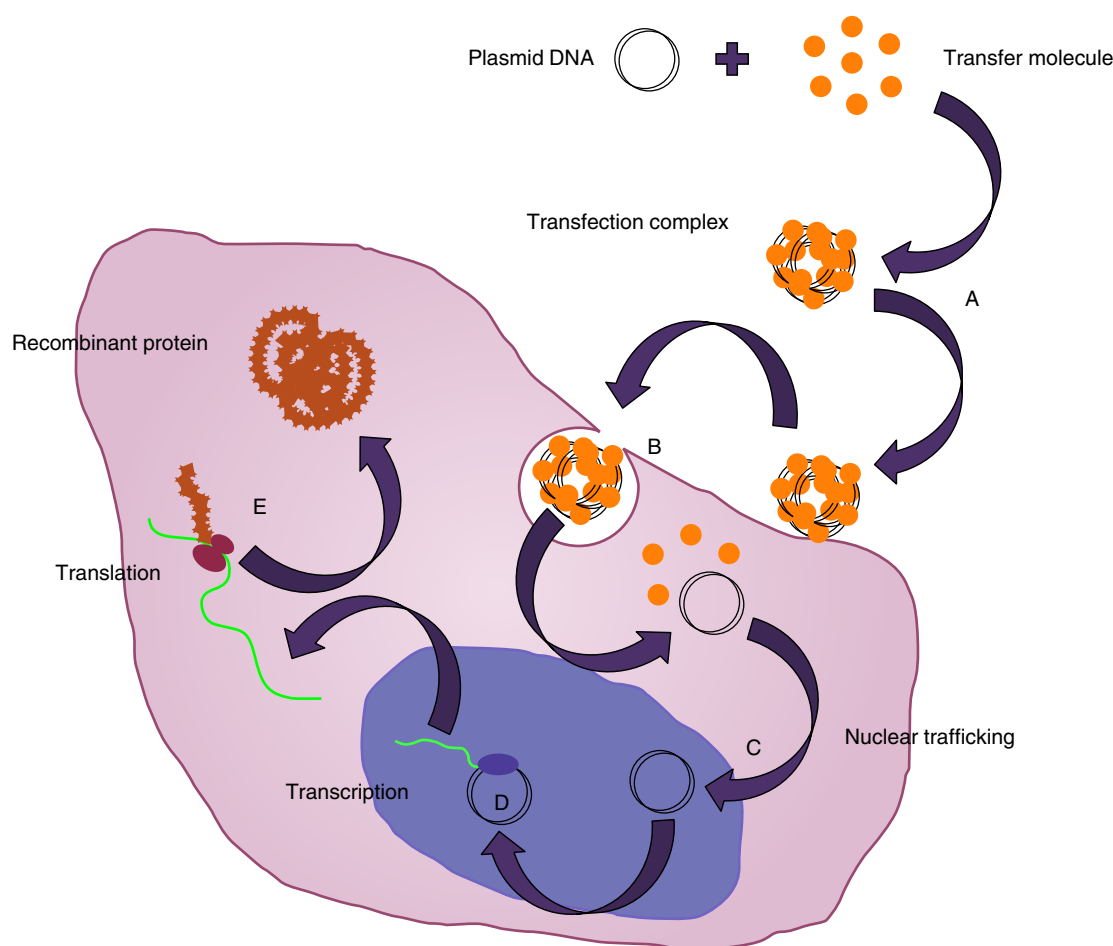


Figure 4 Diagram demonstrating transfection reagent-mediated plasmid DNA delivery and subsequent transgene expression. Figure illustrates (A) transfection complex formation, (B) DNA uptake, (C) DNA trafficking, (D) transcription, and (E) translation of recombinant protein.

Epstein Barr virus, the plasmid can be coupled to chromosomal DNA during cell division. This process greatly improves the retention of the plasmid DNA, which in turn significantly

improves the expression window. Stable cell lines expressing the EBNA-1 gene, or variants thereof, have been developed specifically for use with *oriP*-containing plasmids. The most

Table 6 Common commercially available media supplements for recombinant protein expression

Media supplement	Notes
BD Recharge™; BD Biosciences	Chemically defined supplement designed to mimic yeast peptones by supplying only the key nutrients
BD Resurge™; BD Biosciences	Family of five chemically defined feed supplements. Containing a proprietary mix of nutrients
CD EfficientFeed™; Life Technologies (ThermoFisher)	Family of chemically defined feed supplements in various formats. They contain a proprietary mix of nutrients including energy source, amino acids, and vitamins
HyClone Cell Boost; GE Healthcare	Family of six chemically defined feeds in increasing complexity. They contain a proprietary mix of nutrients starting with glucose and amino acids, through to lipids and growth factors
HyPep 1510; Kerry	Ultrafiltered soy hydrolysate containing amino acids and peptides
Tryptone N1; Organotechnie	Casein hydrolysate containing amino acids and peptides
Tryptone Plus; Organotechnie	Casein hydrolysate containing amino acids and peptides
Yeast extract	Prepared by autolysis of <i>S. cerevisiae</i> cultures, it is a rich, though unrefined, source of amino acids and peptides

Note: Table shows common media supplements designed to boost expression levels in mammalian expression systems.

common EBNA-1 lines are based on HEK293 cells, but CHO variants are also available. Taking this further, it is possible to supply selection pressures to cells with episomally replicating plasmids to ensure retention. This concept has been developed and commercialized by Icosagen in the QMCF system.

Recently, Life Technologies have been specifically looking at increasing the cell culture densities that can be achieved during transient expression. They have commercialized a HEK293-based system (Expi293™) that allows for high-density transfection and high-density production. By increasing the number of cells per liter, the amount of protein produced per liter of culture is increased proportionally. The system includes a HEK293 clone that showed excellent growth and heterologous expression properties, along with a new chemically defined media to support the higher densities.

Summary/Conclusion

The range of available expression systems have grown significantly over the past 15 years. Technical development and reduction in the complexity of setup required for performing work in the more complex systems, especially those based on mammalian cells, have increased their use even in less specialized laboratories. Today the choice of expression system has increasingly become a decision based on what is most appropriate for the protein to be expressed rather than availability of equipment and technical competence.

It is challenging to make general recommendations on the suitability of different expression systems. As has been described in the sections above the ability to express complex proteins with posttranslational modifications increases from the microbial to the eukaryotic systems, especially the mammalian systems. Looking at historical use most often *E. coli* is utilized for the expression of relatively small intracellular proteins not requiring disulfides or posttranslational modifications. Many of the gram-positive bacteria are used for secreted proteins and membrane proteins. The more advanced eukaryotic systems (insect and mammalian) are more all-purpose systems that can be very versatile. But an analysis of the most appropriate expression system for a given protein should be based on the characteristics of the expressed protein and the subsequent application (Brondyk, 2009).

See also: Molecular Principles, Components, Technology, and Concepts: Proteins: Isolation/Purification of Proteins

References

- Ames, R.S., Kost, T.A., Condreay, J.P., 2007. BacMam technology and its application to drug discovery. *Expert Opinion on Drug Discovery* 2 (12), 1669–1681.
- Angelichio, M.L., Beck, J.A., Johansen, H., Ivey-Hoyle, M., 1991. Comparison of several promoters and polyadenylation signals for use in heterologous gene expression in cultured *Drosophila* cells. *Nucleic Acids Research* 19 (18), 5037–5043.
- Berlec, A., Strukelj, B., 2013. Current state and recent advances in biopharmaceutical production in *Escherichia coli*, yeasts and mammalian cells. *Journal of Industrial Microbiology & Biotechnology* 40 (3–4), 257–274.
- Bernhard, F., Tozawa, Y., 2013. Cell-free expression – making a mark. *Current Opinion in structural biology* 23 (3), 374–380.
- Brondyk, W.H., 2009. Selecting an appropriate method for expressing a recombinant protein. *Methods in Enzymology* 463, 131–147.
- Cohen, S.N., Chang, A.C., Boyer, H.W., Helling, R.B., 1973. Construction of biologically functional bacterial plasmids in vitro. *Proceedings of the National Academy of Sciences of the United States of America* 70 (11), 3240–3244.
- Correa, A., Oppezzo, P., 2015. Overcoming the solubility problem in *E. coli*: Available approaches for recombinant protein production. *Methods in Molecular Biology* 1258, 27–44.
- Cos, O., Ramon, R., Montesinos, J.L., Valero, F., 2006. Operational strategies, monitoring and control of heterologous protein production in the methylotrophic yeast *Pichia pastoris* under different promoters: A review. *Microbial Cell Factories* 5, 17.
- Cregg, J.M., Cereghino, J.L., Shi, J., Higgins, D.R., 2000. Recombinant protein expression in *Pichia pastoris*. *Molecular Biotechnology* 16 (1), 23–52.
- Daramola, O., Stevenson, J., Dean, G., et al., 2014. A high-yielding CHO transient system: Coexpression of genes encoding EBNA-1 and GS enhances transient protein expression. *Biotechnology Progress* 30 (1), 132–141.
- Durocher, Y., Perret, S., Kamen, A., 2002. High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells. *Nucleic Acids Research* 30 (2), E9.
- Egelkrout, E., Rajan, V., Howard, J.A., 2012. Overproduction of recombinant proteins in plants. *Plant Science: An International Journal of Experimental Plant Biology* 184, 83–101.
- Gellissen, G., Kunze, G., Gaillardin, C., et al., 2005. New yeast expression platforms based on methylotrophic *Hansenula polymorpha* and *Pichia pastoris* and on dimorphic *Arxula adenivorans* and *Yarrowia lipolytica* – A comparison. *FEMS Yeast Research* 5 (11), 1079–1096.
- Giacalone, M.J., Gentile, A.M., Lovitt, B.T., et al., 2006. Toxic protein expression in *Escherichia coli* using a rhamnose-based tightly regulated and tunable promoter system. *BioTechniques* 40 (3), 355–364.
- Gustafsson, C., Minshull, J., Govindarajan, S., et al., 2012. Engineering genes for predictable protein expression. *Protein Expression and Purification* 83 (1), 37–46.

- Guzman, L.M., Belin, D., Carson, M.J., Beckwith, J., 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. *Journal of Bacteriology* 177 (14), 4121–4130.
- Houdebine, L.M., 2009. Production of pharmaceutical proteins by transgenic animals. *Comparative Immunology, Microbiology and Infectious Diseases* 32 (2), 107–121.
- Jarvis, D.L., Kwar, Z.S., Hollister, J.R., 1998. Engineering N-glycosylation pathways in the baculovirus-insect cell system. *Current Opinion in Biotechnology* 9 (5), 528–533.
- Kikuchi, Y., Itaya, H., Date, M., Matsui, K., Wu, L.F., 2009. TatABC overexpression improves *Corynebacterium glutamicum* Tat-dependent protein secretion. *Applied and Environmental Microbiology* 75 (3), 603–607.
- Kim, K.R., Kim, Y.K., Cheong, H., Kim, J.Y., Cha, H.J., 2013. Expression and N-glycan analysis of human 90 K glycoprotein in *Drosophila* S2 cells. *Enzyme and Microbial Technology* 53 (3), 170–173.
- Kunji, E.R., Slotboom, D.J., Poolman, B., 2003. *Lactococcus lactis* as host for overproduction of functional membrane proteins. *Biochimica et Biophysica Acta* 1610 (1), 97–108.
- Lutz, R., Bujard, H., 1997. Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/11-12 regulatory elements. *Nucleic Acids Research* 25 (6), 1203–1210.
- Mattanovich, D., Branduardi, P., Dato, L., et al., 2012. Recombinant protein production in yeasts. *Methods in Molecular Biology* 824, 329–358.
- Mayer, M., Buchner, J., 2004. Refolding of inclusion body proteins. *Methods in Molecular Medicine* 94, 239–254.
- Niimi, T., 2012. Recombinant protein production in the eukaryotic protozoan parasite *Leishmania tarentolae*: A review. *Methods in Molecular Biology* 824, 307–315.
- Novoa, E.M., Ribas de Pouplana, L., 2012. Speeding with control: Codon usage, tRNAs, and ribosomes. *Trends in Genetics* 28 (11), 574–581.
- Nozawa, A., Nanamiya, H., Tozawa, Y., 2010. Production of membrane proteins through the wheat-germ cell-free technology. *Methods in Molecular Biology* 607, 213–218.
- van Ooyen, A.J., Dekker, P., Huang, M., et al., 2006. Heterologous protein production in the yeast *Kluyveromyces fragilis*. *FEMS Yeast Research* 6 (3), 381–392.
- Ozawa, K., Loscha, K.V., Kuppan, K.V., et al., 2012. High-yield cell-free protein synthesis for site-specific incorporation of unnatural amino acids at two sites. *Biochemical and Biophysical Research Communications* 418 (4), 652–656.
- Pohl, S., Harwood, C.R., 2010. Heterologous protein secretion by bacillus species from the cradle to the grave. *Advances in Applied Microbiology* 73, 1–25.
- Retallack, D.M., Jin, H., Chew, L., 2012. Reliable protein production in a *Pseudomonas fluorescens* expression system. *Protein Expression and Purification* 81 (2), 157–165.
- Stech, M., Quast, R.B., Sachse, R., et al., 2014. A continuous-exchange cell-free protein synthesis system based on extracts from cultured insect cells. *PLoS One* 9 (5), e96635.
- Studier, F.W., 2005. Protein production by auto-induction in high density shaking cultures. *Protein Expression and Purification* 41 (1), 207–234.
- Studier, F.W., Rosenberg, A.H., Dunn, J.J., Dubendorff, J.W., 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods in Enzymology* 185, 60–89.
- Villatoro-Hernandez, J., Kuipers, O.P., Saucedo-Cardenas, O., Montes-de-Oca-Luna, R., 2012. Heterologous protein expression by *Lactococcus lactis*. *Methods in Molecular Biology* 824, 155–165.
- Walls, D., Loughran, S.T., 2011. Tagging recombinant proteins to enhance solubility and aid purification. *Methods in Molecular Biology* 681, 151–175.
- Ward, O.P., 2012. Production of recombinant proteins by filamentous fungi. *Biotechnology Advances* 30 (5), 1119–1139.
- Whittaker, J.W., 2013. Cell-free protein synthesis: The state of the art. *Biotechnology Letters* 35 (2), 143–152.
- Wickham, T.J., Davis, T., Granados, R.R., Shuler, M.L., Wood, H.A., 1992. Screening of insect cell lines for the production of recombinant proteins and infectious virus in the baculovirus expression system. *Biotechnology Progress* 8 (5), 391–396.
- Wulfhart, S., Baldi, L., Hacker, D.L., Wurm, F., 2010. Valproic acid enhances recombinant mRNA and protein levels in transiently transfected Chinese hamster ovary cells. *Journal of Biotechnology* 148 (2–3), 128–132.
- Wurm, F.M., 2004. Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature Biotechnology* 22 (11), 1393–1398.

Further Reading

- Almo, S.C., Love, J.D., 2014. Better and faster: Improvements and optimization for mammalian recombinant protein production. *Current Opinion in Structural Biology* 26, 39–43.
- Assenberg, R., Wan, P.T., Geisse, S., Mayr, L.M., 2013. Advances in recombinant protein expression for use in pharmaceutical research. *Current Opinion in Structural Biology* 23 (3), 393–402.
- Cuozzo, J.W., Soutter, H.H., 2014. Overview of recent progress in protein-expression technologies for small-molecule screening. *Journal of Biomolecular Screening* 19 (7), 1000–1013.
- Estes, S., Melville, M., 2014. Mammalian cell line developments in speed and efficiency. *Advances in Biochemical Engineering/Biotechnology* 139, 11–33.
- Fernandez, F.J., 2013. Technologies to keep an eye on: Alternative hosts for protein production in structural biology. *Current Opinion in Structural Biology* 23 (3), 365–373.
- Fernandez, F.J., Vega, M.C., 2013. Technologies to keep an eye on: Alternative hosts for protein production in structural biology. *Current Opinion in Structural Biology* 23 (3), 365–373.
- Koehn, J., Hunt, I., 2009. High-throughput protein production (HTPP): A review of enabling technologies to expedite protein production. *Methods in Molecular Biology* 498, 1–18.
- Palomares, L.A., Estrada-Mondaca, S., Ramirez, O.T., 2004. Production of recombinant proteins: Challenges and solutions. *Methods in Molecular Biology* 267, 15–52.
- Zerbs, S., Frank, A.M., Collart, F.R., 2009. Bacterial systems for production of heterologous proteins. *Methods in Enzymology* 463, 149–168.

Relevant Websites

- www.addgene.org/
Addgene.
- www.wwpdb.org/
Worldwide Protein Data Bank.

Isolation/Purification of Proteins

GA Grant, Washington University School of Medicine, St. Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Affinity chromatography Separation based on a biological affinity for a specific molecule such as an antibody, a sugar, or other bound moiety.

Anion exchange The exchange of one anion for another electrostatically bound to a cation immobilized on a matrix.

Cation exchange The exchange of one cation for another electrostatically bound to an anion immobilized on a matrix.

Chromatography A method used for the separation of mixtures of molecules that relies on the differential partitioning of molecules between phases. In column chromatography, the partitioning takes place between a solid, stationary phase and a mobile, liquid phase.

Density gradient centrifugation A centrifugal method where particles are separated in a preformed density gradient by the application of centrifugal force until they reach a zone equal to their own density.

Dialysis The process of separating molecules in solution by differences in their diffusion through a semipermeable membrane. Often used to remove or exchange salts from protein solutions.

Differential centrifugation A centrifugal method where larger particles are separated from smaller particles by subsequent increases in centrifugal force.

Electrophoresis A method of separating charged molecules based on their migration in an electrical field.

Isoelectric focusing An electrophoretic technique where a molecule will move in a pH gradient generated by an electric field to a pH equivalent to its isoelectric point.

Isoelectric point (pI) The pH of a solution at which the net charge on a molecule is zero, i.e., the sum of the positive charges is equal to the sum of the negative charges. At pH above the pI, the protein will be negatively charged. At pH below the pI, it will be positively charged.

Isotachopheresis An electrophoretic method where charged molecules are focused at the interface between two electrolytes with different mobility in an electric field.

Ligand A substance, usually a small molecule that forms a complex with a biomolecule, usually a larger molecule such as a protein.

Matrix A substance such as a chromatography gel within which something develops or is contained.

Membrane filtration A method to separate substances by passing them through a membrane with a fixed pore size that lets some pass through and retains others.

Organelles Subcellular structures including the mitochondria, nucleus, endoplasmic reticulum, golgi bodies, endosomes, lysosomes, phagosomes, synaptic vesicles, chloroplasts, and others that contain molecules and other structures separated from the cytoplasm by a membrane.

Polymerase chain reaction An automated, cyclic enzymatic method for the *in vitro* amplification of DNA molecules.

Recombinant DNA technology The isolation, manipulation, and study of single genes.

Sodium dodecyl sulfate-poly-acrylamide gel electrophoresis (SDS-PAGE) An electrophoretic technique for the separation of proteins based on their size in an insoluble polyacrylamide gel. Since the proteins are denatured and dissociated by the SDS, they migrate through the gel based on their covalent molecular weights.

Sonication Application of sound energy to agitate particles. In the context of protein isolation it is used for the disruption of structures such as cells by exposure to ultrasonic frequencies.

Ultrafiltration A form of membrane filtration where a solution is forced through a membrane by centrifugation or applying pressure.

Western blot Electrophoretic separation followed by detection of the protein with a labeled, specific antibody to the desired protein.

Introduction

Proteins normally exist *in vivo* and perform their function in a dense mixture of other proteins, carbohydrates, lipids, nucleic acids, salts, and other metabolites. They may exist free in solution, as components of protein complexes, adsorbed to cellular structures, or as integral parts of cellular components such as membranes. In all cases, they exist to perform a function. In order to be able to assign a particular function to a particular protein, it is usually necessary to isolate it so that it can be studied in the absence of other proteins. In isolation, their physical and chemical characteristics, and in the case of

enzymes their catalytic properties, can be determined so that one is reasonably sure that a particular function is being performed by a specific protein, or in some cases, a defined group of proteins.

Scientists have been engaged in the endeavor of protein isolation since it was first realized that proteins and enzymes were individual molecular entities that could be differentiated from their environment. Indeed, most of what we know about the function of individual proteins comes from studies of isolated homogeneous protein samples. It could be argued that not all proteins may act the same in isolation as they do in their native milieu, but at the very least, studying them in

isolation provides a detailed and invaluable starting point upon which to base further investigations and is really still the best way to pin down a particular protein's function.

Literally thousands of articles have been published that either deal with the isolation of a particular protein or contain sections that deal with it. The thing that one will note when reading a large number of these articles is that they often appear to be quite similar but hardly ever exactly the same. The protocols may look similar because there are a limited number of techniques from which to choose, but different because, while some techniques may be successful with many different proteins, they will not be effective for all proteins. If the same steps are used, the order in which they are performed may differ. The art of protein isolation comes in knowing what techniques are available and how to most efficiently apply them to the particular protein at hand. This article will hopefully assist in the former, but the latter comes with experience. With some exceptions, almost all published protocols for protein isolation did not work the first time and came from trial and error, with incremental successes and failures along the way.

The goal is to obtain an essentially homogeneous or 'pure' protein in sufficient quantity to produce the desired information. A purification protocol cannot be considered truly successful if there is not sufficient protein at the end of it to perform the required analyses. Two things to strive for in this regard are to find conditions where the protein is stable and to minimize losses by minimizing the number of steps in the procedure because each step comes with unavoidable loss of material.

Before the advent of gene cloning and the polymerase chain reaction (PCR), virtually all protein isolation started with a tissue or a culture of microorganisms followed by a multistep purification procedure. One of the great advances in protein isolation came with recombinant technology and the ability to over-express a particular protein. Recombinant technology also allowed the protein to be expressed with a 'tag' that could be used to pick it out of solution with a matrix that contained an entity with specific affinity for the tag. Such a technique is used very often today, but not all proteins can be successfully over-expressed so it is still quite necessary to be familiar with more 'conventional' isolation techniques. In addition, various isolation techniques can be used to 'clean up' protein preparations that are not rendered pure enough by the affinity tag technique. In all cases, their success relies on differences in the location, solubility, charge, size, and affinity of each protein.

This article will focus on laboratory-scale isolation. That is, procedures that can be used in a typical academic laboratory setting to produce modest amounts of protein for analytical purposes rather than large scale production processes. Although many of the same techniques are used in both settings, the scale of production-based isolation is such that many new challenges arise and the infrastructure required is very different.

A protein isolation procedure can be viewed as a combination of steps where the protein progresses in purity with each step: (1) identification and acquisition of a source, (2) extraction from the source, (3) separation from nonprotein components such as nucleic acids and lipids, (4) concentration

of the bulk protein and separation into fractions that contain different proteins, and (5) fine-resolution techniques such as chromatography (Figure 1) that provide the final product.

A few words about purity are in order. What does 'pure' mean and how do you know that a protein is pure? Conceptually, 'pure' means that no other component is present. Of course, this excludes things like buffer components that are intentionally included. For proteins, it usually means the presence of other proteins, but it could also mean the absence of other things like nucleic acids, carbohydrates or lipids that are not covalently attached to the protein. In reality, true purity is not often completely achievable or measurable. The best that you can do is to demonstrate that you are not able to detect the presence of other components or that they are present at less than a certain acceptable percentage. This depends on the sensitivity of the method that you use for detection and the type of component that method will detect.

Unfortunately, an article of this size does not allow an in-depth discussion of every aspect of protein isolation. Hopefully, this article will provide an overview of what is available for the task at hand and provide references where more detailed information can be found. There are many excellent reference sources dealing with protein isolation that anyone wanting to learn more about the subject will find very valuable. Notable among them are the *Methods in Enzymology* series and *Current Protocols in Protein Science*. These and others are listed in the Further Reading section.

The Protein Source

How a particular purification is designed and carried out is very dependent on the source of the protein. It used to be that the only source was from an animal or plant tissue or cultures of cells or microorganisms. More than likely, the protein of interest was not present at levels any higher than many other proteins and often less, thus requiring a significant enrichment along the way. With over-expressed proteins in recombinant systems, the protein of interest is already enriched since it is usually present at levels much higher than any other protein.

One potential drawback from over-expression of a protein from one species in another species, such as a protein from a eukaryote expressed in bacteria, is the formation of inclusion bodies. These are inactive, insoluble aggregates of protein formed when the expressed protein finds itself in a foreign environment at a very high concentration. In addition, mechanisms for such things as glycosylation, membrane transport, and specific protein folding are not generally found in prokaryotes so that the resultant protein may take up a nonnative conformation. One potential advantage of a protein forming inclusion bodies is that, if it can be reactivated from the aggregated form, the inclusion bodies themselves can serve as a form of purification (Palmer and Winfield, 2004; Winfield *et al.*, 2004; Singh and Panda, 2005). The formation of inclusion bodies may also be temperature sensitive. In many instances, failure to obtain soluble protein after expression in *Escherichia coli* at 37 °C, for example, can be at least partially overcome by expressing the protein at lower temperatures, such as 16 °C. While substantial amounts of protein may still

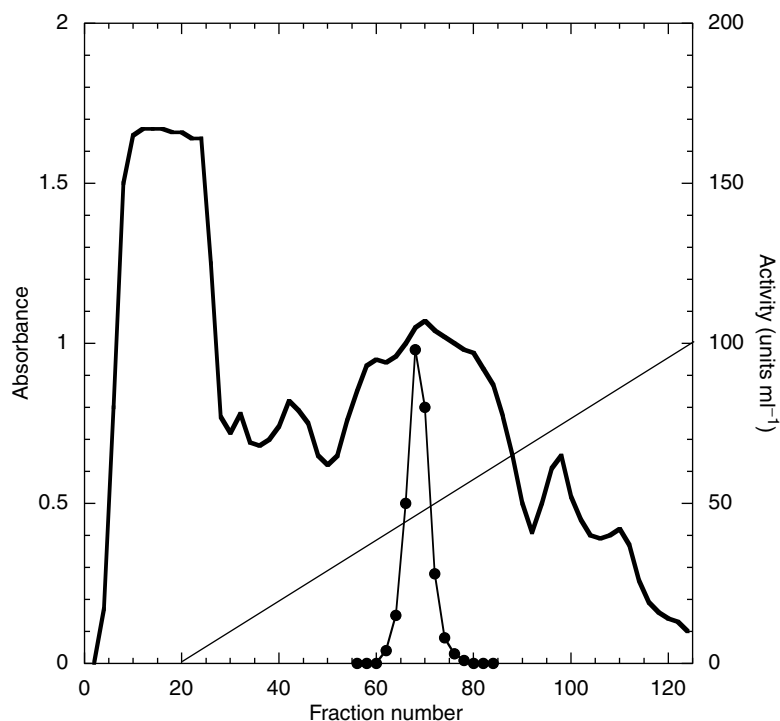


Figure 1 An example of a column chromatogram. A heterogeneous protein mixture was applied to an ion-exchange column and developed with an increasing salt gradient. Absorbance (bold solid line) and enzyme activity (•) are plotted versus the fraction number of the collected column eluate. The diagonal line indicates the gradient of increasing salt concentration. Pooling of fractions 63–75, for example, will yield a more pure protein than that applied to the column.

be found in the insoluble pellet, the supernatant may often contain sufficient amounts of active protein.

Obviously, one of the biggest purification steps one can make is to start with an organ or other tissue rather than the whole organism. But, which tissue to use can be just as important since enzymes can be differentially expressed depending on the tissue. For instance, a particular enzyme may be present in high amounts in the kidney but not in the liver. Therefore, although one may be able to isolate it from the liver, isolating it from the kidney may be easier because there is a lot more to start with and yield will be less of a concern. However, tissues can express specific isoforms differentially so this also should be taken into account when choosing a source.

Assays

Equally important as a good source of protein is a good assay for the protein of interest. It is usually not possible to purify a protein if you cannot identify or locate it as you proceed to separate it from other material. If it does not have enzymatic activity or there is no enzymatic assay available, it could be followed by western blotting or perhaps even simply by following a particular band on a sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gel. In the case of western blotting, it is possible to make antibodies to a protein without having it in pure form first if the sequence is known

from genome studies. This is done by making antibodies to synthetic peptides of hydrophilic stretches of amino acid sequence found in the target protein (Grant, 2002). But remember, you usually isolate what you assay for. If, for instance, your assay detects more than one protein because of cross reactivity, your end product may be different than intended or possibly a mixture of things.

As a protein purification proceeds, one is also interested in determining the total of all protein present in a fraction as a way to judge the efficacy of each step. There are many ways to assay for the total amount of protein, from something as simple as reading the absorbance of a solution (Grimsley and Pace, 2003) to chemical-based colorimetric assays (Olson and Markwell, 2007) and even quantitative amino acid analysis (Rutherford and Dunn, 2011).

Considerations

Accountability

The goal in protein purification is to remove other material from the fraction containing your target protein while retaining as much of your target protein as possible. A good way to monitor this is by constructing a protein purification table that records the specific activity and yield at each step (Table 1). The specific activity of your protein is simply some assessment of the level of your protein relative to all the

Table 1 Example of a protein purification table that documents the work-flow of a purification protocol. The final protein solution provided 15 mg of protein and represented a purification of over 3600 fold with a yield of 36%

Step	Total activity (units)	Total protein (mg)	Specific activity (units/mg)	Yield (% activity)	Purification (fold)
Homogenate/10 000 × g supernatant	50 000	150 000	0.33	100	1.0
40–60 % Ammonium sulfate fraction	40 500	90 000	0.45	81	1.4
Blue-sepharose	36 000	10 000	3.6	72	10.9
QAE-sepharose	20 000	100	200.0	40	606
Size exclusion-sepharose	18 000	15	1200.0	36	3636

protein in the sample. It may be expressed in terms of milligrams of target protein per gram or milligram of total protein or as units of activity per milligram of total protein. Another very useful method of following a purification procedure is to run an SDS-PAGE gel of your protein sample at each step of the purification. The goal is to arrive at a high and constant specific activity and a single band on an SDS gel (or the appropriate multiple bands if your protein is composed of different subunits). If your protein does not have a measurable activity, then a combination of western blots and protein stained SDS gels may suffice to monitor the progress. The key is that you must have some way to identify the target protein. One thing to keep in mind is that although the overall goal is to continually increase the specific activity of the target protein, sometimes a productive step can also be one in which the specific activity does not significantly increase by a large amount, but a contaminant is removed that may have been difficult to remove otherwise.

Accessibility

Once a source has been identified the next step is making the protein accessible. If it is not already soluble in a biological fluid such as blood, this usually involves breaking open the tissue or cell in which it is contained in a way that does not harm the protein. Many methods have been developed for this such as mechanical disruption, freeze-thawing, sonication, high pressure, enzymatic digestion of cell walls, and treatment with detergents (see Further Reading: Scopes, *Methods in Enzymology*). Detergents can be particularly useful for membrane or lipid-associated proteins that are not normally soluble in aqueous solution but they can also interfere with some techniques (Arnold and Linke, 2008; Coligan, 1998).

Stabilization

Once a tissue or cell is broke open, the protein is also subject to the outside environment and all of the elements of the tissue or cell that may have been compartmentalized and originally hidden from the protein. A successful isolation depends on the protein being stable long enough to be purified and studied. This can depend on finding the proper temperature, pH, and ionic strength in which to work. A reducing agent such as dithiothreitol (DTT) can be included if the protein is sensitive to redox conditions. If it is sensitive to

oxygen, an anaerobic chamber may be necessary. A metal chelator, such as EDTA, can be added to bind metal ions that may be detrimental or that promote proteolysis. Control of pH is critical for two reasons. First, since they contain ionizable groups, proteins have a range of pH over which they are stable or active. Note also, that the pH at which a protein is most stable may not be the pH at which it exhibits optimal activity. Secondly, most isolation procedures are sensitive to pH. So control of pH is critical to the reproducibility of an isolation procedure. There are many suitable buffers available to control pH (Blanchard, 1984; Good *et al.*, 1966). Also, proteins may need to be protected from proteases that were also liberated during cell disruption, in which case protease inhibitors may need to be added at least in the initial stages of purification. It is common practice to perform purification steps on ice or in a cold room since the reduced temperature may also serve to reduce any proteolytic damage or chemical reaction rates such as oxidation. However, unless the target protein is not stable at room temperature, most purification steps can be performed at ambient room temperature. In fact this can be an advantage because the equilibration needed for many techniques will occur faster and the operator will be more comfortable. The addition of certain substances, such as glycerol (10–30%) and sugar solutions, has also been shown to aid in the stabilization of proteins. Many proteins can be stabilized by freezing in solution or by lyophilization (freeze-drying). On the other hand, some proteins are irreversibly denatured by freezing. Storage as an ammonium sulfate precipitate can be very effective for stabilization (see below).

Repeatability

Unless you intend to never purify a particular protein again, an absolutely necessary criterion of a protein purification protocol is that it is repeatable. This requires that very good notes as to the conditions and processes of every step be kept so that each step can be reconstructed in the same way each time it is attempted. A purification protocol should be repeated at least two or three times with comparable results before it is published.

Initial Steps

Certain techniques can be very advantageous in the early stages of a purification protocol and some will be used repeatedly

throughout the purification. Differential precipitation and batch adsorption can often remove large amounts of material early on that may interfere with or overload more highly resolving techniques such as column chromatography steps.

Differential Precipitation

The inclusion of a differential precipitation step or steps relatively early in the process should always be considered because it can provide an effective and easy bulk separation, and it can also be used to concentrate the protein for subsequent steps. Since not every protein has the same characteristics, their differences can be exploited by altering their solubility by changing ionic strength, pH, or temperature, or by the addition of miscible organic solvents. Perhaps the most used and effective differential precipitation technique is ammonium sulfate precipitation. Proteins in general undergo what has been termed 'salting in' and 'salting out.' Some proteins such as globulins are insoluble at very low ionic strength and need increasing salt concentration to remain in solution. As the salt concentration increases, other proteins may precipitate. The most effective salt for this purpose is ammonium sulfate because it can effectively alter the solubility without harming the protein. In fact, ammonium sulfate is often very effective in stabilizing proteins and can be a good way to preserve a protein overnight between purification steps. The process involves adding ammonium sulfate in steps to incrementally increase its concentration. The precipitate that results is then removed by centrifugation after each addition. The choice of which concentrations to use is a matter of trial and error but a good initial concentration is 25% saturation since this is effective in removing pre-aggregated material and very high molecular weight protein. Increasing salt concentration beyond this in steps can differentially precipitate specific proteins. If the next step in the purification is sensitive to ionic strength, the salt can conveniently be removed by dialysis or ultrafiltration. Ammonium sulfate concentrations can be expressed as molar concentrations but are more often expressed as percent saturation. Since the volume of the solution will increase with added salt, the amount needed for a particular saturation is usually read from a table (Wingfield, 1998), but can also be calculated with the following equation to take the solution from an initial percent saturation level (S1) to the next percent saturation level (S2):

$$\begin{aligned} &\text{grams of ammonium sulfate to add/liter at } 20^{\circ}\text{C} \\ &= (533[S2 - S1]) / (100 - 0.3[S2]) \end{aligned}$$

Substances with specific adsorption characteristics or affinities can also be added to precipitate components such as nucleic acids. Streptomycin sulfate or polyethyleneimine (polymyxin P) are most commonly used for this purpose and recommended as one of the first steps after tissue or cell disruption (Burgess, 1991).

Batch Adsorption

The principle of this technique is similar to many of the types of adsorptive chromatography presented later and can be performed with the same types of materials, i.e., ion-exchange

gels, dye-ligand gels. The batch procedure can be performed relatively quickly compared to a chromatographic step. It simply entails mixing the protein solution in a beaker with the matrix, gently stirring or shaking for a period of time, removing the solid material by filtration or centrifugation, and either retaining the supernatant or eluting the protein from the matrix. Because it can be performed quickly and can be easily scaled up, it is particularly useful during the early stages of a purification procedure.

Centrifugation

This is one of the easiest and handiest tools available. A very early step in almost any isolation is separating particulate matter such as cellular debris from soluble components. It is also useful in separating precipitates from solution after differential precipitation procedures (see below). It can also be very useful in separating and enriching sub-cellular organelles or large protein complexes by differential centrifugation or density gradient centrifugation (Castle, 2004; Aniento and Gruenberg, 2003). It is also used to force liquid through a membrane for ultrafiltration.

Concentration

It is almost always desirable to work with smaller volumes of more concentrated material than the other way around. Since many separation techniques tend to produce fractions that are more dilute than the starting material, it is often desirable to include a concentration step before or between separation steps. A very good way to do this is to use ammonium sulfate precipitation. If the desire is simply to precipitate your protein rather than effect an additional purification, and depending on the protein's solubility, an ammonium sulfate concentration of 75–80% can be routinely employed.

Another very effective concentration technique is termed ultrafiltration. Centrifugation is used to force liquid through a membrane that will retain material above a particular size or molecular weight cut-off. Many such products are commercially available as disposable cartridges.

It is also possible to concentrate a protein by adsorbing it batch-wise to a solid matrix, such as an ion-exchange resin, and then eluting it with a small volume of buffer of sufficient ionic strength.

Chromatography

Chromatography consists of a set of techniques used to separate mixtures based on the partitioning of components between two phases. Of the different types of chromatography (i.e., Paper, Thin-layer, Column, etc.), the one most useful for protein purification is column chromatography. It consists of a solid matrix placed in a cylindrical column through which a liquid solution is passed. The differential interaction of the components in solution with the matrix results in a separation. Chromatography columns are developed by introducing liquid to the top of the column and collecting it from the bottom of the column. Collection is usually done in small

increments or volumes, called fractions, so that the separation of material accomplished by the column is retained. The fractions are then assayed individually for the presence of the target protein. The results of a chromatography can be represented by constructing a chromatogram, which is a plot of the amount of material exiting the column as a function of volume, time, or fraction (see [Figure 1](#)). Typically, total protein is measured by recording its absorbance, usually at 220 or 280 nm, and the target protein by its particular assay.

Column chromatography is very well suited for protein isolation because the conditions under which the separation occurs are usually very gentle to the protein and the resolution can be very high. Separation of protein by column chromatography relies on various chemical and physical principles and includes separations based on charge, size, selective adsorption, hydrophobic interaction, and specific affinity for certain molecules.

Ion-Exchange Chromatography

Ion-exchange chromatography (IEX) depends on the interaction of charge on the surface of a protein with an opposite charge on an insoluble matrix. Since the charge on the protein, and usually the matrix as well, is dependent on protonatable groups (amino acid side chains on the protein), the interaction between the two will depend on pH. Since most proteins do not tolerate extremes of pH, ion exchange is usually conducted at pH between 4 and 9. Most ion-exchange resins have a rather sharp transition between being charged and not, so the behavior of the chromatography is usually more sensitive to the ionization state of the protein. In a typical ion-exchange procedure, protein is loaded onto the column in a buffer that maintains a steady pH and contains a relatively low level of salt such as NaCl. Some salt is beneficial to keep the protein soluble but should be low enough so that it does not interfere with the ionic interaction between protein and matrix. Monovalent salts such as NaCl, KCl, or NH_4Cl are typically used and a typical buffer may be 20 mM Tris-HCl, pH 8.0, 100 mM NaCl. Common buffers may be Tris-HCl, ammonium bicarbonate, phosphate, or a series of buffers called the 'Good' buffers ([Good et al., 1966](#)). To maintain buffering capacity, the pH used should be within 0.5 pH units of the pK_a of the buffer. Proteins are usually eluted from the column by increasing the ionic strength of the buffer by increasing the salt concentration. This can be done in an incremental step-wise fashion but is more often accomplished with a gradient that gradually and smoothly increases the salt concentration. A typical gradient may be from 0.1 to 1 M salt. Elution from the column can also be accomplished by changing the pH of the buffer but this is not as common and usually does not provide as good of a resolution of components.

The most common charge groups on ion-exchange matrices are diethylaminoethyl (DEAE) and quaternary aminoethyl (QAE) for anion exchangers and carboxymethyl (CM) and sulfoethyl (SE) for cation exchangers. Phosphocellulose, whose phosphate group is another charged group, can be used effectively with proteins. In fact, this material can also exhibit characteristics of affinity chromatography (see below) for

proteins that bind phosphorylated compounds. The matrices themselves must also be made from materials that have pores large enough for proteins to enter. This increases the surface area accessible to proteins and thus the capacity of the matrix. Ion-exchange matrices such as Dowex, used for small molecules, are not suitable for proteins since they have very small pores and thus very low capacity for large molecules. The typical types of materials used for protein ion-exchange matrices are cellulose, agarose (Sephacrose), dextran (Sephadex), and acrylamide (Sephacryl). Other support material and ionic groups may be available depending on the manufacturer.

Determining which ion-exchange matrix to use is often accomplished by trial and error. The isoelectric point (pI) of a protein can be estimated from its amino acid composition if it is available using the protparam tool at the ExPASy website. This gives some guidance as to a starting point but the real pI may be altered by the protein structure. One general approach is to start with an anion exchange matrix like DEAE in 20 mM Tris buffer, pH 7.5–8.0, 100 mM NaCl. The pH and ionic strength of the protein solution should be matched to the column buffer by prior dialysis. The volume of sample applied to the column can be decreased by a concentration step, but a very high protein concentration should be avoided. It is actually better to load the column with a more dilute solution ($> 5\text{--}10$ mg protein/ml) to avoid local concentration and pH effects that may interfere with adsorption to the matrix. If the target protein does not 'stick' to the column, it will probably be retained by a matrix of the opposite charge. If the protein is not retained by either matrix, try adjusting the pH or lowering the salt concentration. Do not overlook the fact that just because the target protein isn't retained, it does not mean the column is not effective. Many other proteins may be retained and a good purification may have been achieved. In the end, the effectiveness is always judged by the yield and specific activity. Some proteins may be irreversibly adsorbed or even inactivated by the column matrix. In which case, the choice of matrix should be reconsidered. Always be mindful of any other buffer components. For instance, a high concentration of EDTA (a polyanionic reagent) may interfere with protein capacity on an anion-exchange resin.

The dimensions of the column are not as important for ion-exchange as they are for size exclusion chromatography where column length is a factor (see below). The major concern is for capacity and a good choice would be to use a column whose length is 4–5 times its diameter and whose diameter is large enough so that the protein adsorbs to the top 20–30% of the column matrix. In general, the wider the column the higher the flow rate that can be achieved. A higher flow rate will allow the protein to spend the least amount of time on the column and for the chromatography to be accomplished in a shorter amount of time.

Remember that when the protein is eluted from the ion-exchange column, it will be in a certain concentration of salt. If the next step is adversely affected by the salt, it must first be removed by a procedure such as dialysis or membrane filtration. As you will see next, size exclusion chromatography is also a good way to remove salt from a protein solution, and one often sees this as a step immediately following ion-exchange without prior removal of salt.

Size Exclusion Chromatography

Size exclusion chromatography has also been called gel filtration chromatography, gel permeation chromatography, and molecular sieve chromatography. This type of chromatography separates on the basis of size or more correctly hydrodynamic volume. It employs a gel matrix with pores of a specific size range and separates on the basis of whether or not a protein of a particular size can enter the pore or is excluded from the pore. Larger proteins that cannot enter the pores do not equilibrate with the interior volume of the gel bead and therefore move through the column faster than smaller proteins that can enter the pore and are momentarily retained because they are essentially removed from the main column flow for a short period of time. Molecules that are completely excluded from the beads elute in what is termed the void volume (V_o) of the column, which is approximately 30% of the bed volume. Molecules that can completely equilibrate inside the pores of the beads elute in the inclusion volume (V_i), typically around 80% of the bed volume. Since this process occurs repeatedly as a molecule moves down the column, longer columns tend to display better resolution of different size species. A typical column size may be $3\text{ cm} \times >100\text{ cm}$ and have a flow rate of 1–2 ml/min. Also, because larger molecules move faster, they elute first from the column while smaller species elute later. Because of this, size exclusion chromatography (SEC) can be used in an analytical mode, with the appropriate standards, to determine the molecular weight of proteins and other macromolecules in their native state. This is particularly good for analyzing the native molecular weight of multi-subunit proteins or protein complexes. Because gel beads cannot be made with a single uniform pore size, SEC gels separate over a range of molecular weights with different gels separating in different size ranges. Like ion-exchange gels, SEC gels are made of similar materials including cellulose, agarose (Sephacrose), dextran (Sephadex), and acrylamide (Sephacryl, Bio-gel P).

Adsorption Chromatography

The term 'adsorption' chromatography is generally used to refer to adsorptive processes that are not mainly due to ionic interactions or dependent on a 'biological' affinity. Chief among these are Dye-ligand chromatography and Inorganic Adsorbent chromatography.

Dye-ligand chromatography

Certain organic dyes that can be attached to a solid support have been found to bind to certain types of proteins. The most widely used of these are Cibacron Blue and Procion Red that have a good affinity for NAD and NADP dependent proteins such as dehydrogenases, with the latter being more selective for NADP enzymes. Other dyes are also available depending on the manufacturer. Elution of the protein is usually performed by increasing the ionic strength but if there is sufficient hydrophobic interaction this can increase the binding affinity. For dehydrogenases, the particular coenzyme, NAD(H) or NADP(H), can be used but these are relatively expensive.

Hydrophobic interaction chromatography

This type of chromatography relies on the interaction of hydrophobic areas on proteins with a matrix containing hydrophobic groups in the presence of high salt concentrations. The high salt promotes the hydrophobic interaction and the proteins are eluted by decreasing the salt concentration. As such, it is well suited as a step after ammonium sulfate precipitation or IEX where the target protein is eluted at a relatively high salt concentration.

This should not be confused with reversed phase chromatography (RPC) or hydrophilic interaction chromatography (HILIC). These techniques are usually not well suited for proteins because they employ water miscible organic solvents such as acetonitrile that can disrupt protein structure or cause aggregation.

Inorganic adsorbent chromatography

The most common and useful among these for proteins is hydroxyapatite, a crystalline form of calcium hydroxyphosphate. The mode of action is not well-defined and it is a matter of trial and error to determine if it is useful for a particular target protein. Other adsorbents that have been employed include silica gel and alumina. In practice, these adsorbents are most often used in industrial processes to remove contaminants and some unwanted proteins. They could also be used for batch adsorption early in a protocol.

Affinity Chromatography

Affinity chromatography is a general term referring to a separation based on a biological affinity for a specific molecule such as an antibody, a sugar, or some other particular molecular recognition process (Chapter 9 in *Current Protocols in Protein Science* has many useful articles).

Immunoaffinity chromatography

This chromatography entails the covalent attachment of an antibody for a particular protein to a solid support. As such, the column can be highly selective. However, because antibodies tend to have a very high affinity for their antigen, eluting the bound protein can be problematic.

Lectin affinity chromatography

Lectins are molecules that specifically bind sugars, and this type of chromatography is used to selectively bind glycoproteins. Immobilized concanavalin-A and wheat germ agglutinin are the most widely used because they have wide specificity and are relatively inexpensive. There are many other lectins that are commercially available but they are not generally used for protein purification because they are expensive. Also, even if a variety of lectins were used, they might not yield a pure protein since other proteins with similar sugars would tend to co-purify.

Metabolite affinity chromatography

If a particular metabolite that is a substrate, product, or cofactor for an enzyme can be immobilized on a matrix and still retain binding affinity for the enzyme, it can be effective and often very specific. One example is the use of 5'-AMP and 5'-ADP Sepharose (and similar materials) for proteins that

bind nucleotides. These materials may also exhibit ion-exchange characteristics since they contain charged groups.

Affinity tag chromatography

The development of recombinant DNA technology and the PCR led to a major breakthrough in protein purification by enabling the widespread use of affinity tags. It is safe to say that, today, many more purification protocols employ this technique than do not. Not only is it easy and relatively inexpensive, it often results in sufficient purification in just a single step. It has proven to be so effective that it should be considered a potential part of every protein purification procedure. This type of chromatography depends on a specific interaction between a molecule that has been covalently attached to a protein (usually as a polypeptide extension) and some specific solid support that has affinity for the 'tag.' The most common and widely used type is immobilized metal affinity chromatography (IMAC) through the addition of a sequence of 6 histidine residues to the amino- or carboxyl-terminus of a protein that specifically interact with metals such as nickel or cobalt that have been immobilized on a solid support. Other notable and commonly used tags include FLAG, Strep-Tactin, glutathione S-transferase (GST), and maltose-binding protein (MBP) (Kimple *et al.*, 2013). In addition, a variety of specific proteolytic cleavage sites can be placed between the tag region and the protein so that the tag can be removed after purification. This is not always necessary but there are instances where the tag can interfere with protein function. Many commercially available protein expression vectors have built-in proteolytic cleavage sites for removing affinity tags.

Affinity tags are generally of two types, those that function mainly in affinity recognition for purification and those that aid in keeping the expressed protein soluble. Often the tag can serve both purposes, such as MBP. A discussion of all the affinity tags that have been developed would take more space than is available here, but good review articles are available (Kimple *et al.*, 2013; Lichty *et al.*, 2005).

The use of tandem affinity tags have also proven useful for reducing the background of nonspecifically bound proteins that may survive a purification using only a single tag. Two affinity tags with different specificity are engineered onto the same protein that is then passed through the two appropriate affinity matrices in succession.

A variation on the tandem tag method has also been used successfully in the isolation of eukaryotic proteins from cell extracts where they are not expressed at very high levels and isolation by a single tag does not result in protein of sufficient purity (Günzl and Schimanski, 2009; Gloeckner *et al.*, 2009). While the technique does not usually produce very high levels of protein (sometimes only picomolar amounts) the protein is usually very pure and sufficient enough for some analytical techniques such as partial sequencing by mass spectrometry. This approach has proven quite useful for the isolation of protein complexes or proteins composed of nonidentical subunits.

Migration in an Electric Field

Proteins can be separated by their differential migration in an electric field. Three such methods that have been adapted for

preparative purposes are electrophoresis, isoelectric focusing and isotachopheresis. However, they are not widely used for preparative purposes due to their expense, complexity, and generally low yield. On the other hand, gel electrophoresis, especially SDS gel electrophoresis, has seen much success as an analytical method and is a major technique in assessing protein purity.

Protein Crystallization

Today, protein crystallization is usually performed for the purpose of structure determination by X-ray crystallography (McPherson, 2004). However, back when protein purification methods were less sophisticated, a crystallization step was sometimes employed for purification. The common method for crystallization is to form a supersaturated solution of the protein with precipitants such as ammonium sulfate or polyethylene glycol. Crystallization can be the ultimate purification step since the crystals are made up of a repeating array of a single protein. However, the down side is that many proteins need to be very pure already in order to induce them to crystallize. Nonetheless, some proteins can crystallize from fairly heterogeneous solutions and even micro-crystals will suffice. Crystallization is also a good method to stabilize proteins for storage. These days, one does not usually see a crystallization step in most modern purification procedures, but it bears keeping in mind for those times when it could be helpful.

Evaluating Purity

As mentioned at the beginning of this article, assessment of protein purity is really the demonstration that no other proteins are present. This of course is dependent on the sensitivity and resolving power of the analytical methods used. By far the most common method used is SDS gel electrophoresis and staining the proteins in the gel with dye or silver-based stains. Another good indication of purity is a constant specific activity across a chromatographic peak. Mass spectrometry, although very sensitive and useful, is not recommended by itself because not all proteins may be detectable since they may not all ionize successfully or equally. Likewise, the presence of additional protein signals can be misleading unless all the spectra can be compared quantitatively.

A Final Word

Traditionally, every protein purification procedure was tailored to the specific protein being purified and a successful purification often had an element of art as well as science. Nonetheless, there was probably always more than one way to purify any protein and most procedures could probably be improved in some way. As technology improves, the approach that one takes also tends to change. For instance, recombinant DNA technology has produced a kind of 'one size fits all' situation where a large number of diverse proteins can be purified with essentially the same protocol, such as the

hexa-histidine metal affinity tag. However, it would be a mistake to think that this is the answer to all purification challenges, especially those involving eukaryotic proteins, and especially when they are expressed in other than bacterial cells or must be isolated from tissues or cell cultures. Thus, one would do well to be familiar with the whole repertoire of purification techniques for those times when the 'art' of protein purification still has a role.

See also: Molecular Principles, Components, Technology, and Concepts: Basic Principles: Chemical and Physical Principles. Molecular Principles, Components, Technology, and Concepts: Proteins: Antibodies and Improved Engineered Formats (as Reagents); NMR in Structural and Cell Biology; Protein Domains: Structure, Function, and Methods

References

- Aniento, F., Gruenberg, J., 2003. Subcellular fractionation of tissue culture cells. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Section 4.3.1–4.3.12.
- Arnold, T., Linke, D., 2008. The use of detergents to purify membrane proteins. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Section 4.8.
- Blanchard, J.L., 1984. Buffers for enzymes. *Methods in Enzymology* 104, 404–414.
- Burgess, R.R., 1991. Use of polyethyleneimine in purification of DNA-binding proteins. *Methods in Enzymology* 208, 3–10.
- Castle, J.D., 2004. Purification of organelles from mammalian cells. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Section 4.2.
- Coligan, J., 1998. Commonly used detergents. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Appendix A.1B.1.
- Gloeckner, C.J., Boldt, K., Ueffing, M., 2009. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Section 19.20.
- Good, N.E., Winget, G.D., Winter, W., *et al.*, 1966. Hydrogen ion buffers for biological research. *Biochemistry* 5, 467–477.
- Grant, G.A., 2002. Synthetic peptides for production of antibodies that recognize intact proteins. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Section 18.3.
- Grimsley, G.R., Pace, C.N., 2003. Spectrophotometric determination of protein concentration. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Section 3.1.
- Günzl, A., Schimanski, B., 2009. Tandem affinity purification of proteins. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Section 19.19.
- Kimple, M.E., Brill, A.L., Pasker, R.L., 2013. Overview of affinity tags for protein purification. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Section 9.9.
- Lichty, J.J., Malecki, J.L., Agnew, H.D., Michelson-Horowitz, D.J., Tan, S., 2005. Comparison of affinity tags for protein purification. *Protein Expression and Purification* 41, 98–105.
- McPherson, A., 2004. Introduction to protein crystallization. *Methods* 34, 254–265.
- Olson, B., Markwell, J., 2007. Assays for determination of protein concentration. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Section 3.4.
- Palmer, I., Winfield, P.T., 2004. Preparation and extraction of insoluble (inclusion body) proteins from *Escherichia coli*. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Section 6.3.
- Rutherford, S.M., Dunn, B.M., 2011. Quantitative amino acid analysis. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Section 3.2.
- Singh, S.M., Panda, A.K., 2005. Solubilization and refolding of bacterial inclusion body proteins. *Journal of Bioscience and Bioengineering* 99, 303–310.
- Winfield, P.T., Palmer, I., Liang, S.-M., 2004. Preparation and extraction of insoluble (inclusion body) proteins from *Escherichia coli*. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Section 6.5.
- Wingfield, P., 1998. Protein precipitation using ammonium sulfate. In: Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc., Appendix A.3.F.1–A.3.F.7.

Further Reading

- Bergmeyer, H.U. (Ed.), 1990. *Methods of Enzymatic Analysis* (12 volume series). Weinheim: Wiley-VCH.
- Coligan, J.L., Dunn, B.M., Speicher, D.W., Wingfield, P.T. (Eds.), *Current Protocols in Protein Science*. John Wiley & Sons, Inc.
- Eisenthal, R., Danson, M.J. (Eds.), 2002. *Enzyme Assays: A Practical Approach*, second ed. Oxford: Oxford University Press.
- Hunte, C., von Jagow, G., Schagger, H. (Eds.), 2003. *Membrane Protein Purification and Crystallization: A Practical Guide*, second ed. San Diego, CA: Academic Press.
- Janson, J.-C. (Ed.), 2011. *Protein Purification: Principles, High Resolution Methods, and Applications*, third ed. Hoboken, NJ: Wiley.
- Methods in Enzymology*. Elsevier. Especially volumes 182, 270, and 463 (see Relevant Websites).
- Protein Expression and Purification*. Journal published by Elsevier. (See Relevant Websites).
- Protein purification protocols. In: Cutler, P. (Ed.), *Methods in Molecular Biology*, second ed., vol. 244. Totowa, NJ: Humana Press.
- Scopes, R.K., 1994. *Protein purification: Principles and practice*. In: Cantor, C.R. (Ed.), *Springer Advanced Texts in Chemistry*, third ed. New York, NY: Springer.
- Simpson, R.J., Adams, P.D., Golemis, E.A. (Eds.), 2009. *Basic Methods in Protein Purification and Analysis: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Walls, D., Loughran, S.T. (Eds.), 2011. *Methods in Molecular Biology, Protein Chromatography: Methods and Protocols*. New York, NY: Humana Press.

Relevant Websites

- www.expasy.org/resources
ExPASy.
- <http://web.expasy.org/protparam/>
ExPASy.
- www.gelifesciences.com
GE Healthcare Life Sciences.
- <http://www.sciencedirect.com/science/bookseries/00766879>
ScienceDirect. *Methods in Enzymology*.
- <http://www.sciencedirect.com/science/journal/10465928>
ScienceDirect. *Protein Expression and Purification*.

Protein Sequence Determination: Methodology and Evolutionary Implications

RA Bradshaw, University of California, Irvine, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Sequencing Proteins by Chemical Techniques

Early Studies

Life processes require the accurate translation of genetic information, which is basically linear in nature, into proteins, which generally only function correctly when they have folded to their proper conformation. Thus, proteins have both a linear component (amino acid sequence; referred to in the earlier literature as the primary structure) that is dictated by genomic DNA (via RNA copies) and a three-dimensional structure that is dictated by the sequence (although often helped by auxiliary proteins and covalent modifications, usually termed posttranslational modifications or PTMs) (Anfinsen, 1993). However, as described below, the elucidation of protein sequences has evolved into a relatively simple task during the last 75 years, it is still not possible to predict the exact structure that any given unknown sequence will adopt, although considerable progress has been made in achieving this goal. The realization, some 50 years ago, that proteins with recognizably related sequences also had related structures spurred the development of molecular evolution and, as the numbers of protein structures from both X-ray crystallography and two-dimensional NMR increased, so did the number of predicted structures from model building. Today, with the availability of vast amounts of predicted protein sequence data from an ever widening collection of organisms derived from high-throughput (or next Gen) nucleic acid sequencing, the challenges of understanding protein structure/function properties tend to be ones of identifying and understanding PTMs, particularly transient ones, and the dynamic nature of the multiplicity of protein-protein interactions that characterize cellular processes.

The elucidation of peptide and protein sequences began in earnest about the time of the Second World War. At that juncture, the covalent structure of peptides (the Hofmeister-Fischer hypothesis) was well accepted (Rosenfeld, 2012) but there was some uncertainty about whether they were composed of random or ordered sequences of amino acids. It was also uncertain whether they had 'open' N- and C-termini or occurred as closed loops that lacked free ends. These issues were definitively resolved by Sanger and his colleagues with the determination of the sequence of insulin (Sanger, 1964), which in its mature state is composed of two distinct polypeptide chains covalently linked by two interchain disulfide bonds. These studies, which required some 12 years to complete, depended on the earlier work of Martin and Synge (1941), who introduced chromatography (initially on paper) as a means of separating complex mixtures and the development of a method for tagging free amino groups with fluorodinitrobenzene (now known as Sanger's reagent). Using strong and weak (partial) acid hydrolysis and eventually proteolytic enzymes (pepsin, chymotrypsin, and trypsin) and expanding their fractionation techniques to include ionophoresis and ion-exchange chromatography, they completed the sequence determination and elucidated the pairing of the three disulfide

bonds (there was one intrachain linkage). They also introduced the strategy of deducing the full sequence from overlapping peptides that became the basis for protein sequencing for the next 25 years (Schroeder, 1968; Blackburn, 1970).

The Move to Larger Proteins

As important as these studies were, it was clear that insulin was a relatively small protein (in some estimations it was on the borderline between a peptide and protein) and the elucidation of its sequence had been greatly facilitated by the fact that it was isolated in two smaller pieces (after the disulfides were cleaved by performic acid oxidation) and that it lacked methionine and tryptophan, two amino acids that can present problems during analytical manipulations. Fortunately there were two additional technological advances 'in the wings' that proved to be essential in allowing the application of Sanger's methods and strategy to larger and more complicated proteins. The first of these was a highly reproducible analytical method for doing amino acid analyses (Moore and Stein, 1993) and the second was a different method for tagging and determining the N-terminal residue of a peptide that could be applied sequentially (Edman, 1950). Both of these key contributions were in fact already being developed during the time Sanger was sequencing insulin but neither were fully adopted until after he had completed that task.

The development of the automated amino acid analyzer was the work of Stein and Moore and their colleagues (Moore and Stein, 1993) and utilized ion-exchange chromatography for separation and the ninhydrin reagent for detection and quantification. The resin that was finally adopted for this purpose was a sulfonated cross-linked polystyrene known as Dowex 50. In addition to its ion-exchange properties, it had considerable hydrophobic character that allowed the separation of the nonpolar amino acids as well as the ones bearing polar side chains and provided quantitative identification on appropriately calibrated instruments. In the early stages of development, a single analysis required about 24 h but constant improvements eventually shortened the time to less than an hour. Coincidentally, as a part of this work, they also developed the fraction collector (Stein and Moore, 1948), which is still, in its various manifestations, a staple of any biochemistry laboratory. Both the analyzer and the fraction collector were subsequently developed commercially.

The new N-terminal method, first developed for this purpose by Edman (1950), was based on the reaction of the α -amino group of a peptide with phenyl isothiocyanate in dilute alkali, which produced a phenylthiocarbamyl derivative. The modified peptide would, on exposure to acid, cyclize to a substituted anilinothiazolinone that removed the first residue from the remainder of the peptide. Subsequent conversion of the unstable thiazolinone to the corresponding phenyl thiohydantoin (commonly referred to as a PTH-amino acid) allowed for its identification by a variety of analytical

techniques. Under conditions that limited the amount of acid cleavage of other peptide bonds in the remainder of the peptide during cyclization, the reaction could be repeated to identify the second amino acid (now the new N-terminus) and so on (Figure 1). When originally applied to isolated peptides derived from enzymatic digests (that were generally in the range of 2–20 amino acid residues), it could also be carried out in a subtractive mode (Hirs *et al.*, 1960) whereby the composition of the peptide was determined before and after each cycle of the Edman degradation to determine what amino acid had been removed (by difference). The availability of high-speed, quantitative amino acid analyses made this a particularly popular approach in the 1960s and early 1970s. The usefulness of the Edman degradation to sequencing studies was greatly enhanced by the adoption of (nearly) anhydrous trifluoroacetic acid (or similar compounds) that allowed cyclization to occur with only trace amounts of background hydrolysis (and minimal degradation of some amino acids) (Konigsberg and Hill, 1962). This was particularly true when this invaluable method was subsequently also automated (Begg and Edman, 1967).

The Role of Separations Methodology

Beside its modest size, insulin was also available to the Sanger group in relatively large amounts in homogenous form, a prerequisite for chemical sequencing studies. It is therefore not surprising that most of the studies that followed Sanger's pioneering effort were with proteins that were also relatively easy to purify and obtain in substantial amounts (see Schroeder, 1968, for a description of a representative listing of these early efforts). These tended to be extracellular moieties that were found in fluids or other conveniently obtained rich sources.

Nonetheless, the substantial improvements in separations technology that followed the Sanger work played a very significant role in making available more and more proteins for sequence analysis. Two of the most important of these were ion-exchange resins suitable for protein separations, such as substituted celluloses and gel filtration media. These latter resins, consisting of cross-linked dextrans (Sephadex) or polymerized and cross-linked acrylamide (Bio-gel), basically fractionated protein mixtures on the basis of size, and gave a new dimension to purification schemes that was largely independent of separations based on charge. Performed in tandem they were very effective. Of course there were many additional methods that were modified or introduced, such as gel electrophoresis, isoelectric focusing, reverse phase high-performance liquid chromatography (HPLC), and affinity chromatography among others, that substantially increased the arsenal of separations procedures. When any or all of these were used in conjunction with the older bulk methods, such as alcohol or ammonium sulfate precipitation, sophisticated multistep protocols to purify even relatively rare proteins began to be fairly commonplace by the mid-1960s. It was not till recombinant technology became available that these often laborious protocols became largely passé and the need to sequence proteins by chemical means (other than to obtained partial sequence data to construct oligonucleotide probes for cloning experiments), for the most part, disappeared.

The development of separations technology also played a very significant role in preparing peptides and fragments for sequencing. In early studies the fragments generated by enzymatic cleavage were generally small and soluble (although there was invariably an insoluble core that often contained longer fragments, whose further fractionation always presented significant challenges). In fact, until automated sequencing

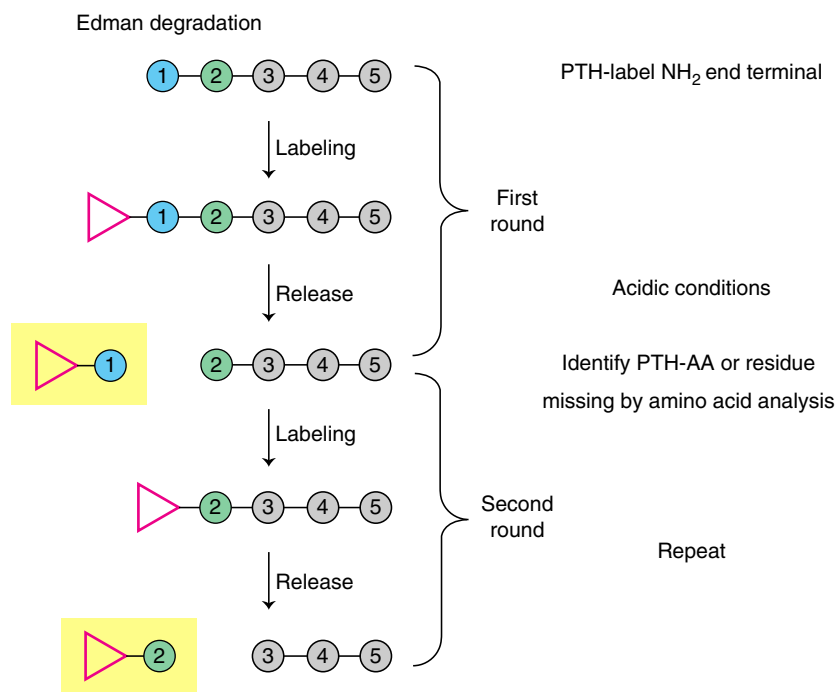


Figure 1 Schematic diagram of sequential cycles of Edman degradation.

became routinely available, peptides much longer than 10 residues would require some sort of further reduction in size since manual Edman degradation rarely went much further than that. Thus it was common to use several types of enzymes, with differing specificities, as a means to generate enough soluble peptides to get complete coverage of the sequence being determined. Since, after trypsin, the specificity of the proteases available was less precise and, therefore, less reliable this often proved to be a problem. For the most part, mixtures of soluble peptides from enzymatic digests were amenable to fractionation by the same resins used for separating amino acids (Dowex 50, its basic counterpart, Dowex 1, and Amberlite IRC 50, which was a carboxylate resin and hence a weak anionic exchanger). Gel filtration resins, with low-molecular weight cutoffs, and in later years, HPLC, were also very useful.

The wide-scale introduction of the automated protein sequencer in the early 1970s had a material impact on strategy and hence the technology employed. At first, it was most valuable in determining the N-terminal sequence of whole proteins but that soon led to the realization that by creating bigger fragments these could be analyzed as if they were proteins in their own right (and unlike the whole protein, which was often blocked by an acetyl group if it was a cytoplasmic protein from a eukaryotic cell, they had an accessible N-terminus). Employing cleavage strategies that produced fewer cuts and thus larger pieces had already appeared in more ambitious studies of significantly larger proteins. Many of the more successful methods for creating bigger fragments depended on chemical reagents (as opposed to enzymes). The most important and widely used of these was cyanogen bromide (CNBr) (Gross and Witkop, 1961), a highly reactive and toxic substance that in strongly acidic solutions, for example 70% formic acid, reacts only with the thioether of methionine side chains to produce a sulfonium salt that then results in cyclization leading to cleavage of the adjacent peptide bond (and leaves homoserine lactone in the place of the methionine in the C-terminal position of each fragment generated, except the one derived from the C-terminus of the protein). Generally the large fragments produced were due to the relatively rare occurrence of methionine. Various other methods, often depending on the introduction of one or more modifications first, were also developed (Blackburn, 1970).

The creation of big fragments that were easy to analyze by automated sequencing and were effective in subdividing large sequences into more manageable sized pieces also produced their own challenges. Being derived from denatured proteins and therefore essentially without structure themselves, they were often difficult to solubilize and purification schemes for them usually required more extreme solvent conditions than the fractionation of the smaller soluble mixtures. Gel filtration was particularly useful in these applications because it was better suited to tolerate such solvents (including detergents).

It is noteworthy that the spinning cup technology of Begg and Edman (1967) became much more amenable to the analysis of smaller peptides with the discovery that polybrene[®] added to the reaction mixture materially reduced losses of sample, thus greatly enhancing its value in large-scale sequence projects.

Challenges and Limitations

The strategy of determining a complete sequence by piecing together overlapping fragments depended almost exclusively on N-terminal sequencing. Determining C-terminal sequences was considerably more difficult and there was no sequential degradation method comparable to the Edman degradation for attacking peptides from the C-terminal end despite several efforts to develop one (Inglis, 1983). Hydrazinolysis (Akabori *et al.*, 1952), which converted every residue in a peptide or protein to a hydrazide except the C-terminal one, was a relatively complex procedure with a number of confounding problems associated with identifying the one free amino acid and was not practical for routine peptide sequencing. The only method that was widely used was timed hydrolysis by carboxypeptidases A and/or B (Blackburn, 1970). The latter was very efficient in removing lysine and arginine and hence worked well on most tryptic peptides but its counterpart, carboxypeptidase A, gave much more variable results as some residues were released so rapidly they could not be distinguished while others were completely intransigent. Since this method was very much sequence-dependent, it could rarely be depended on to give more than two or three reliable identifications and often produced nothing at all.

This one-sided approach to collecting data often undermined reliable peptide alignments and residue assignments. Many times overlaps were made on the basis of a single residue, which was risky at best, and the conclusion that there were no residues that had been overlooked in the final proposed sequence usually depended on the known molecular weight and the amino acid composition that had been determined (based on this number). It is illustrative that Sanger worked for many years under the illusion that insulin had a molecular mass of about 12 000 Da rather than the actual figure of about 6000 Da and it was certainly welcome news to him that he only had ~50 residues to determine rather than 100 he had originally thought (Sanger, 1964). This did become less of a problem when methodology for determining subunit structure and protomer molecular weights, particularly sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Weber and Osborn, 1969), became routine (and basically replaced much more sophisticated and expensive methods, such as analytical ultracentrifugation).

Good practice also dictated that it was important to establish the identity of each residue in two different ways, i.e., in two different peptides, that added to the challenge of assembling the final proposed sequence (that this rule was often violated which led to a significant number of incorrect assignments that were only subsequently uncovered with the advent of nucleic acid sequencing). Automated sequencer runs of whole proteins and large fragments could often identify 40–50 amino acids with few if any undetermined residues. The repetitive yield of the reactions of each cycle and the accumulated loss of material determined the limit of obtaining useful data. Using polybrene[®] added to the reaction vessel and a modified coupling program, Thomas *et al.* (1981) reported the identification of 79 residues out of 83 present (with only 4 unidentified sites, 1 of which was ultimately identified as asparagine with N-linked carbohydrate attached, and all of which occurred in the last 18 cycles of the run) using ~600

nmols of a naturally occurring fragment of the γ -subunit of 7S NGF, which probably represents the practical limit of this technology. As sequence determinations increasingly became dependent on sequencer data and relatively long stretches were achieved, the number of identifications determined only a single time also increased. To some degree this was offset by improved technology, such as the introduction of the gas-phase sequencer (Hunkapiller *et al.*, 1991) that largely displaced the spinning cup methodology of the original Begg-Edman design, and significantly increased accuracy and sensitivity. However, in the end, the development of nucleic acid sequencing effectively made the problem no longer moot.

The stability and reactivity of the 20 acids that make up most proteins in biology are not the same and accordingly presented variable issues in sequence studies. Several amino acids were known to be partially destroyed by the conditions of acid hydrolysis commonly used (6 N HCl at 110 °C for 24 h, *in vacuo*) such as serine, threonine, and tryptophan. In the last case, this was probably more related to contaminants than to any inherent instability. These, and other residues, are also partially destroyed in the cyclization and conversion of the thioazolinones to the hydantoins, which was a major problem in sequencer analyses. Cysteine and its oxidized form, cystine, were recognized early on as unstable amino acids (Sanger converted the 6 half-cystines in insulin to cysteic acid) and it was a routine step to modify them to a stable form. Until it was determined that disulfide bonds were introduced in the endoplasmic reticulum (ER) for proteins that were destined to be exported through the ER-Golgi continuum to the extracellular environment and that intracellular proteins, except for a few special cases, had only cysteine residues, newly isolated proteins were reduced and alkylated routinely without regard to what form they were in the isolated protein. Because this obscured the original state of the side chain, the number of Cys residues was usually reported as half-cystines, a term that did not distinguish their state in the native proteins. Since the conditions for strong oxidation also modified several other residues, milder reactions such as the introduction of carboxymethyl and carboxamidomethyl groups were favored. There are several other modifications also available, depending on the type of subsequent analysis envisioned. Converting half-cystines to trypsin-sensitive sites by converting them to *S*-aminoethyl derivatives is an example. The availability of the germane reagents in radioactive form was also a useful way to count half-cystine residues (as labeled peptides) following tryptic digestion and was an effective way to determine the true molecular weight based on the minimal molecular weight calculated from the independently determined half-cystine content (Angeletti *et al.*, 1971).

The determination of disulfide pairing was a commonly encountered problem in early studies because these tended to focus on proteins found in easily obtained materials, such as body fluids, that were by definition from extracellular environments. Although the half-cystine residues could be assigned from residues converted to stable derivatives, such as was done with insulin, the connectivity could only be ascertained from peptides derived from the unmodified protein. Because it was already recognized by Sanger (1964) that disulfide interchange could occur around pH 7 and higher, most chose pepsin as the enzyme of choice, which functions in

much more acidic conditions. The basic procedure employed was to isolate the cystine-containing peptides to homogeneity, cleave them by oxidation (or reduction and alkylation) and identify the two contributing halves (Spackman *et al.*, 1960). Since at this point the sequence was usually known, the peptides did not have to be sequenced extensively to make unequivocal identifications. The spacing of some half-cystine residues (not bonded to each other) was sometimes so close that pepsin cleavage between them was not achieved and then other enzymes or chemical methods were required to generate unique cystine peptides (see, e.g., Angeletti *et al.*, 1973). When a protein contained multiple disulfide bonds, the fractionation of these linked peptides to homogeneity was often challenging. One method that was relatively simple and proved to be quite effective was diagonal electrophoresis (Brown and Hartley, 1966). The digest containing the cystine peptides was separated by paper electrophoresis in one dimension, exposed to performic acid vapor (forming cysteic acid *in situ*) and then re-electrophoresed under the same conditions at 90° to the original direction. The pairs of peptides making up the original disulfide-linked entity moved off the diagonal because of their changed charge properties although all the non-cystine peptides did not. The cysteic acid containing peptides were then eluted and identified by partial sequence analysis (sometimes amino acid composition alone was sufficient). With the advent of the new ionization techniques in mass spectrometry, the identification of cystine peptides in unfractionated samples became possible saving a great deal of experimental manipulation and thus time (see, e.g., Raffioni *et al.*, 1988).

The presence of naturally occurring PTMs, with the exception of carbohydrate, was largely ignored and was not a point of focus until the combination of cell signaling studies and proteomic technology demonstrated the extent to which they are present. Despite the fact that it is now known that modifications, such as phosphorylation, occur extensively affecting a very high proportion of intracellular proteins (Gnad *et al.*, 2011) (and even some extracellular ones), these hardly ever were detected in chemical sequence analyses. This is probably mainly due to instability during handling and to low occupancy of many if not most of the modifications. One exception to this was the introduction of proteolytic cleavages, which were readily detected and accurately mapped in many proteins. However, distinguishing between cleavages introduced posttranslationally (see, e.g., Thomas *et al.*, 1981) from proteins synthesized as separate entities was not easily accomplished until genomic sequence data became readily available.

Other Sequencing Methodology

Mass Spectrometry

The applicability of mass spectrometry to determining amino acid sequences was recognized in the late 1950s to early 1960s by a number of laboratories, particularly those of Biemann (MIT) (Biemann *et al.*, 1959) and Lederer (Paris-Orsay) (Barber *et al.*, 1965). Although clearly there were advantages to this approach including the ability to determine a number of residues at the same time and in the presence of

contaminating peptides, the principal problem was the lack of volatility of peptidic material and this proved to be a major limitation until greatly improved desorption-ionization methodology was developed. Early workers were able to achieve sequence data by this approach by derivatizing various functional groups in relatively short peptides or taking advantage of natural modifications. Fortuitine, a mycobacterial peptidolipid of nine residues that contained equal amounts of eicosanoic and docosanoic acids in amide linkage with the N-terminal valine, two residues of *O*-acetyl threonine, two residues of *N*-methyl leucine, and a C-terminal methyl ester was an example of the latter (Barber *et al.*, 1965). *N*-Methylation of peptide bonds, among several other modifications, which increased volatility, was subsequently adopted for other sequence determinations (Blackburn, 1970). For the most part, early sequence studies with mass spectrometry were effective in studies with cyclic peptides and antibiotics or when used in conjunction with other chemical approaches as exemplified by the sequence analysis of the heptadecapeptide amide, gastrin (Agrawal *et al.*, 1969) and were generally limited to peptides of <10 residues in length. They were also laborious and insensitive, owing in part to the losses incurred during the modification reactions.

The value, and hence the use, of mass spectrometry for determining peptide and protein sequences dramatically changed with the introduction of new technology that overcame these problems. Beginning in the early 1980s, first fast atom bombardment (FAB) and then matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) allowed the analysis of peptides of ever increasing size to eventually include proteins (without any prior modification) to produce highly accurate masses (Siuzdak, 1996). When analyzed with tandem mass spectrometers, peptides could be isolated and then subjected to fragmentation (by collision with inert molecules) to produce a series of sequentially truncated ions that corresponded to the original sequence. Three different ion pairs (six different types of ions) can result from cleavage of the three different types of covalent bonds found in the polypeptide backbone but the most commonly encountered are *b* and *y* ions (see Figure 2; Roepstorff and Fohlman, 1984). A completely overlapping set (*b* ions are derived from the N-terminus and *y* ions are derived from the C-terminus) will yield the entire sequence. When this methodology was connected with chromatographic systems that were compatible with ESI, it was possible to continually analyze the effluent to produce a very large number of at least partially sequenced peptides (Burlingame, 1989).

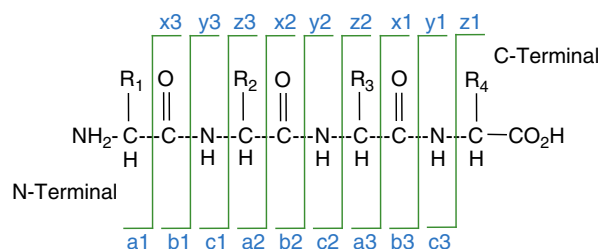


Figure 2 Fractionation pattern of peptides that occurs in tandem mass spectrometric analyses as proposed in Roepstorff and Fohlman (1984).

However, by the time of the new millennium when this methodology had already reached this level of sophistication (and beyond), the need for sequencing individual proteins had been long eclipsed by advances in nucleic sequencing (see below). Thus, mass spectrometric sequencing rather found its most useful applications in the unbiased identification of proteins (inferred from sequenced peptides), usually in unfractionated, complex samples (but also in smaller scale experiments) for the purposes of (1) analyzing expression, elaborating the composition of protein complexes/machines and elucidating partners in protein-protein interactions and (2) determining/localizing PTMs. These types of studies are at the core of the new field of proteomics (Bradshaw and Burlingame, 2005).

Because most proteomic measurements depend on pre-formed databases that have been mainly assembled from annotated genomic sequences or other nucleic acid based analyses, the use of this methodology (or of the chemical methodology described above) to determine new, unknown sequences is fairly limited. Even proteins from more exotic species can often be deduced from homologous proteins, whose sequence has already been determined, using programs such as BLAST (Altschul *et al.*, 1990). However, on occasion it is necessary to determine a sequence *de novo* and there are two basic approaches referred to as top-down and bottom-up analyses. The latter depends on the fragmentation of the proteins, usually by trypsin, and the identification and sequencing of the corresponding peptides, a procedure that is entirely analogous to the overlapping chemical approaches pioneered by Sanger and colleagues for insulin (Sanger, 1964). The large-scale protein identification (Washburn *et al.*, 2001) experiments also use this strategy but do not depend on achieving full coverage of the sequence since identification can virtually always be reliably accomplished with a single sequence of 6–7 amino acids in length although rigorous studies usually require at least two such peptides. Identifications based on sequenced peptides cannot distinguish splice variants unless the peptide covers a juncture corresponding to two exons. It can identify individual variants from unique peptides but the absence of peptides in a mass spectrometric analysis cannot be used to infer the absence of that variant. When the intention is to determine a new sequence (or identify modified residues in an already known protein), essentially complete coverage of the sequence is necessary and, as in the earlier chemical approaches, usually requires more than one set of peptides. Top-down analyses (Kelleher *et al.*, 1999) basically start with a pure protein (or a relatively simple mixture) and subject the entire sample to fragmentation in the mass spectrometer. This results in very complex spectra and the presence of a great number of fragment ions that must be resolved and identified. This challenging approach has been increasing successful with new advances in mass spectrometric analyses, mainly electron capture dissociation (ECD) and, more recently, electron-transfer dissociation (ETD), and proteins of over 200 residues have been extensively sequenced in a single experiment (Peng *et al.*, 2014). This type of analysis has also been applied to large fragments and is often termed ‘middle down.’ The main value of this methodology, which can only be expected to improve, is most likely to be in the identification and localization of PTMs, particularly in proteins, such

as histones, that carry multiple kinds and sites of modifications in complex and varying patterns (Garcia, 2010).

Inferring Protein Sequences from Nucleic Acid Sequences

The acquisition of protein sequence information by chemical methodology reached its peak in the late 1970s. However, even with many improvements in this technology, it remained a relatively labor-intensive and time-consuming endeavor. Thus, the codevelopment of cloning techniques around that time that allowed the preparation of DNA copies of mRNA (called complimentary or cDNA) and reliable DNA sequencing techniques that were faster and cheaper, which spawned the methodology of recombinant DNA (Berg, 1993), soon displaced chemical methodology as a source of protein sequence information. Two different methods for sequencing DNA were independently developed by Sanger (1993) and Gilbert (1993) for which they shared the Nobel Prize in Chemistry for 1980. Sanger's procedure eventually was adopted as the method of choice. Thus he established the methodology for determining the sequences of both proteins and nucleic acids, an achievement of inestimable importance to further studies in cell and molecular biology. Initially, knowledge of at least a portion of the amino acid sequence of a target protein was necessary in order to synthesize DNA probes to screen appropriate libraries to find the corresponding cDNA sequence and the techniques of protein sequencing were highly valuable in achieving this. Eventually, however, DNA manipulations, in general, and DNA sequencing, in particular, became sufficiently efficient that this information was no longer necessary and which finally culminated in the rapid determination of whole genomes, including that of *Homo sapiens*. The accumulation of genome sequence data on a massive scale continues unabated. As a result the overwhelming majority of protein sequence information that has been collected and is available today, spanning the entire range of living (and some that are extinct) organisms, is based not on direct measurements but on translated nucleic acid sequences. Indeed there is still between 10 and 20% of the predicted human proteome that has not yet been directly identified in an appropriate sample by any technique (Kim *et al.*, 2014; Wilhelm *et al.*, 2014). Although nucleic acid sequencing has provided a plethora of information (that would have required decades to obtain by direct protein sequencing, even with the advances in mass spectrometry), there are consequences that have presented new challenges, the most important of which is the detection of splice variants (as alternatively expressed forms of a given gene) and PTMs. While the number of genes predicted to make up the human genome plummeted during the process of elucidating it (from somewhere ~150 000 to the eventually determined number of ~20 000), the number of splice variants and particularly the extent of PTM in the actual expressed proteome rose dramatically. In the end it became clear that the concept of 'one gene, one protein' (Beadle and Tatum, 1941; Berg and Singer, 2003) was a vast oversimplification and that rather than depend on a huge multiplicity of unique genes to drive biological complexity, nature simply found ways to modify a more restricted group of chemically distinct proteins (and thus increase the

number of functions they could perform or participate in) through splicing and chemical modification instead. Unfortunately, simple DNA sequences are not reliable indicators of what splicing events will occur and under what conditions and consensus sequences for PTMs do not guarantee that a site will actually be modified. Furthermore, the majority of PTMs are governed by only very loose rules regarding site characteristics and in many cases there are no discernible rules at all. Thus the objectives of protein sequence analyses have shifted from determining the order of amino acids to determining exon usage and downstream covalent modifications. Mass spectrometry, in its various manifestations, is overwhelming the method of choice for such analyses.

Sequence Comparisons

One of the earliest observations to arise from chemical sequencing studies was the realization that proteins of the same function in different species or of similar but not identical function in the same (or other) species showed readily detectable similarities that could be deduced by simply aligning the two. Proteins that showed such a relationship were then said to be homologous, meaning that they had a common ancestry. The term homology came from biological studies and, before sequence analysis, was based primarily on the comparison of anatomical structures. Two things (structures) were said to be homologous if they had a common ancestry. Protein chemists co-opted this word to apply to related protein sequences (assuming, as they did, that they had evolved from common genes) and further manipulated the term to imply degrees of similarity. Thus they introduced qualifiers such as partial, weak, strong, etc. that, since homology as defined by the biologists was a yes or no condition, changed the meaning of the word. Despite an extensive polemic, this modified use of the term has remained in common usage. Sanger and colleagues (Brown *et al.*, 1955; Harris *et al.*, 1956) established that the same protein in different species were similar by studying the insulins from four additional species: pig, sheep, horse, and whale (his original sequence analysis was carried out with bovine insulin). He found them to be highly similar with only a three residue sequence, located in the intrachain disulfide loop of the A chain, showing substitutions. Thus the number of identical residues in the two chains was 48/51 (=94%). Human hemoglobin chains and myoglobin also showed similarities but in this case the number of identities was considerably lower (see Figure 3). However, it was also obvious that many of the substitutions in other positions were by residues of similar chemical character and it was concluded that these might be well tolerated if the proteins also showed three-dimensional similarity as well. These two proteins both function to transport oxygen through the agency of a heme prosthetic group, although in quite different cellular locales, and they have distinctly different subunit structures (often referred to as quaternary structure) as myoglobin is a monomer and hemoglobin is a heterotetramer ($\alpha_2\beta_2$), where the α and β chains also show the same kind of relatedness, albeit more extensive than that found in the myoglobin comparisons. Thus the myoglobin/hemoglobin comparison represents the evolution of function. These concepts were cemented by early



Figure 3 Comparison of the amino acid sequences of hemoglobin α chain and myoglobin. Boxed residues are identical; those highlighted in yellow are chemically similar.

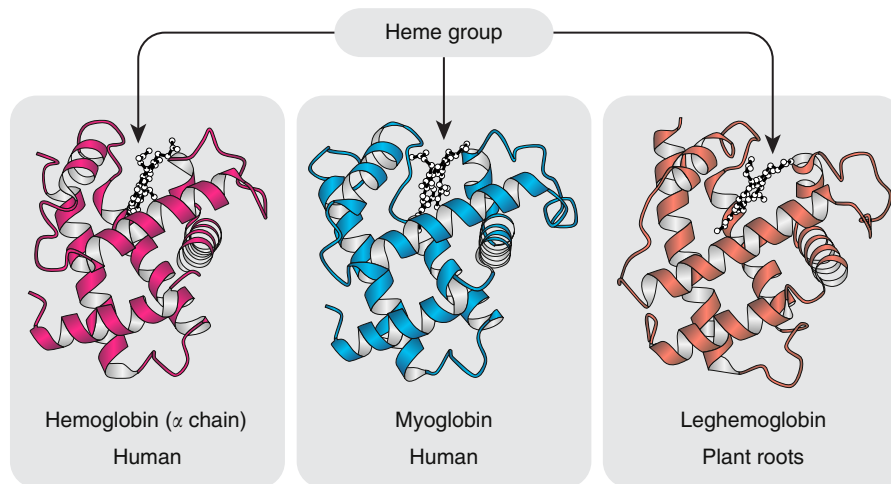


Figure 4 Three-dimensional structures of hemoglobin α chain, myoglobin, and leghemoglobin.

sequence studies on various serine proteases (trypsin, chymotrypsin, elastase, and thrombin) (Walsh and Neurath, 1964; Hartley *et al.*, 1965; Hartley, 1968) and the carboxypeptidases (Bradshaw *et al.*, 1969). All of these examples were also the subject of early crystallographic analyses using X-ray diffraction and it became quite clear that the similarities in amino acid sequence were also manifested in similarities in the folding of the polypeptide chain. Thus, for example, the folding of the polypeptide chain of myoglobin and either of the hemoglobin chains is obviously similar, even without further computerized analyses. Figure 4 compares the structure of these two and a third heme proteins found in plant roots termed leghemoglobin; all have essentially the identical three-dimensional structure. In fact it has ultimately been shown that the three-dimensional structure of a protein is a more sensitive indication of homology, particular with proteins that are distantly related. When sequence comparisons fall below 25% identities (a region Doolittle (Doolittle, 1981) rightfully termed the 'twilight zone'), the certainty that two proteins are truly homologous becomes considerably more difficult to ascertain by sequence data alone.

The evolution of species variants and of functional variants has since been defined in more specific terms: orthologous sequences are from proteins derived from the same ancestral sequence now found in different species and paralogous sequences are from those proteins still present in the same species that arose via a gene duplication event (see Figure 5). Once

two copies of the gene exist, one is free to accumulate point mutations leading to potential new function. When this function exerts some sort of advantage, the changes will become fixed in the population but will be recognized as a paralog of the original gene product as long as a sufficient amount of the original sequence remains unaltered. A paralogous group may constitute a family containing a large number of members, such as the protein kinases (Manning *et al.*, 2002), which retain the common capability of phosphorylating certain protein side chains utilizing ATP as donor but recognize different sequence motifs containing the target residue, or may be composed of only two members. In the case of the kinases, their catalytic function has not been altered but their substrate specificity has been. Enzymes (or other protein types) that basically retain their catalytic function but show differential expression in cells and organs are usually termed isozymes (isoforms) and can sometimes be exploited as a biomarker if they are elevated (or decreased) and are released into the circulation in a certain disease or condition (see, e.g., Peng *et al.*, 2014). Paralog groups may also be found in a variety of species if the speciation event in question occurred after the gene duplication.

Although the paralogous gene duplications cited above basically correspond to the generation of complete copies, partial duplications are also well known. Indeed many proteins consist of identifiable domains (which often coincide with exons) and these clearly can be copied and recopied during the gene duplication events that marked evolution,

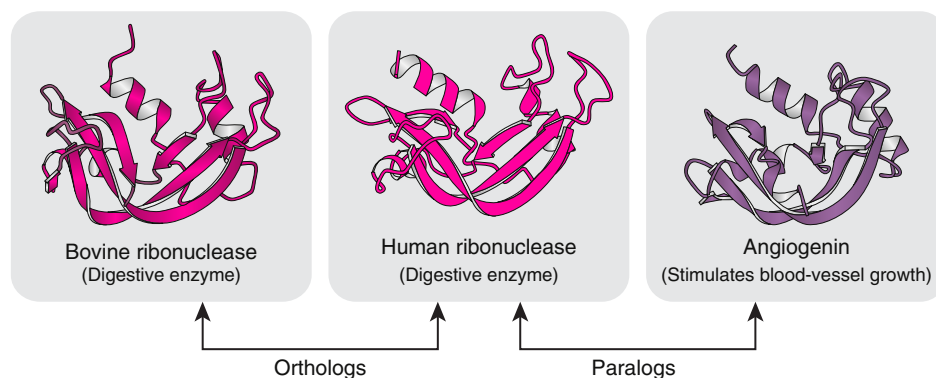


Figure 5 Examples of orthologous and paralogous relationships.

thus ending up in variety of proteins, sometimes totally unrelated (Tang *et al.*, 1978). This phenomenon is called ‘domain swapping’ and can occur multiple times in larger proteins. The ‘EF hand,’ a Ca^{++} binding domain composed of two α -helices on either side of a short loop containing the carboxylate ligands that bind the metal that was first identified in fish parvalbumin (Kretsinger *et al.*, 1971), is widely distributed in mammalian proteomes, being found in a large number of proteins of quite varied function that are regulated by Ca^{++} .

Concluding Remarks

The determination of linear sequence information, first from proteins and then nucleic acids, has paralleled the history of research in the biological sciences for the past 75 years. It can be viewed in three principal phases: the chemical determination of protein sequences, the determination of nucleic acid sequences (and the inferred sequences of many additional proteins) and the determination (and identification) of proteins from partial sequences, their modifications (PTMs) and their interacting partners in unfractionated samples by high-throughput mass spectrometry. Each phase basically represents a paradigm shift in our understanding of the fundamentals of cell biology and, importantly, each was dependent on the previous phase for its inception and development. Although sequence and structure will never by themselves completely explain the workings of the cell, they are indispensable information that will be required to reach this goal.

See also: Molecular Principles, Components, Technology, and Concepts: Proteins: Expression Systems; Isolation/Purification of Proteins; Posttranslational Modifications: Key Players in Health and Disease; Site-Directed Mutagenesis

References

- Agrawal, K.L., Kenner, G.W., Shepperd, R.C., 1969. Feline gastrin. An example of peptide sequence analysis by mass spectrometry. *Journal of the American Chemical Society* 91, 3096–3097.
- Akabori, S., Ohno, K., Narita, K., 1952. On the hydrazinolysis of proteins and peptides: A method for the characterization of carboxyl-terminal amino acids in proteins. *Bulletin of the Chemical Society of Japan* 25, 214–218.
- Altschul, S., Gish, W., Miller, W., Myers, E., Lipman, D., 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215, 403–410.
- Anfinsen, C.B., 1993. Studies on the principles that govern the folding of protein chains. In: Sture, F. (Ed.), *Nobel Lectures, Chemistry 1971–1980*. Singapore: World Scientific Publishing Co. (Nobel lecture, 1972) Tore Frängsmyr (Editor-in-Charge).
- Angeletti, R.H., Bradshaw, R.A., Wade, R.D., 1971. Subunit structure and amino acid composition of mouse submaxillary gland nerve growth factor. *Biochemistry* 10, 463–469.
- Angeletti, R.H., Hermansen, M.H., Bradshaw, R.A., 1973. Amino acid sequences of mouse 2.5 S nerve growth factor. II. Isolation and characterization of the thermolytic and peptic peptides and the complete covalent structure. *Biochemistry* 12, 100–115.
- Barber, M., Jollès, P., Vilkas, E., Lederer, E., 1965. Determination of amino acid sequences in oligopeptides by mass spectrometry. I. The structure of fortuitine, an acyl-nonapeptide methyl ester. *Biochemical and Biophysical Research Communications* 18, 469–473.
- Beadle, G.W., Tatum, E.L., 1941. Genetic control of biochemical reactions in *neurospora*. *Proceedings of the National Academy of Sciences of the United States* 27, 499–506.
- Begg, G., Edman, P., 1967. A protein sequenator. *European Journal of Biochemistry* 1, 80–91.
- Berg, P., 1993. Dissections and reconstructions of genes and chromosomes. In: Sture, F. (Ed.), *Nobel Lectures, Chemistry 1971–1980*. Singapore: World Scientific Publishing Co. (Nobel lecture, 1980) Tore Frängsmyr (Editor-in-Charge).
- Berg, P., Singer, M., 2003. *George Beadle, an Uncommon Farmer: The Emergence of Genetics in The 20th Century*. New York, NY: Cold Spring Harbor Press.
- Biemann, K., Gapp, G., Seibl, J., 1959. Application of mass spectrometry to structure problems. I. Amino acid sequence in peptides. *Journal of the American Chemical Society* 81, 2274–2275.
- Blackburn, S., 1970. *Protein Sequence Determination: Methods and Techniques*. New York, NY: Marcel Dekker, Inc.
- Bradshaw, R.A., Burlingame, A.L., 2005. From proteins to proteomics. *IUBMB Life* 57, 267–272.
- Bradshaw, R.A., Neurath, H., Walsh, K.A., 1969. Considerations of the concept of structural homology as applied to bovine carboxypeptidases A and B. *Proceedings of the National Academy of Sciences of the United States* 63, 406–411.
- Brown, H., Sanger, F., Kitai, R., 1955. The structure of pig and sheep insulins. *Biochemical Journal* 60, 556–565.
- Brown, J.R., Hartley, B.S., 1966. Location of disulphide bridges by diagonal paper electrophoresis. *Biochemical Journal* 101, 214–228.
- Burlingame, A.L., 1989. Peptide and protein characterization by mass spectrometry. In: Hugli, T.E. (Ed.), *Techniques in Protein Chemistry*. San Diego, CA: Academic Press, pp. 121–126.
- Doolittle, R.F., 1981. Similar amino acid sequences: Chance or common ancestry? *Science* 214, 149–159.
- Edman, P., 1950. Method for determination of the amino acid sequence in peptides. *Acta Chemica Scandinavica* 4, 283–293.
- Garcia, B.A., 2010. What does the future hold for top down mass spectrometry? *Journal of the American Society for Mass Spectrometry* 21, 193–202.

- Gilbert, W., 1993. DNA sequencing and gene structure. In: Sture, F. (Ed.), Nobel Laureates in Chemistry during the period 1971–1980. Singapore: World Scientific Publishing Co. (Nobel lecture, 1980), Tore Frängsmyr (Editor-in-Charge).
- Gnad, F., Gunawardena, J., Mann, M., 2011. PHOSIDA 2011: The posttranslational modification database. *Nucleic Acids Research* 39, D253–D260. **Database issue.**
- Gross, E., Witkop, B., 1961. Selective cleavage of the methionyl peptide bonds in ribonuclease with cyanogen bromide. *Journal of the American Chemical Society* 83, 1510–1511.
- Harris, J.I., Sanger, F., Naughton, M.A., 1956. Species differences in insulin. *Archives of Biochemistry and Biophysics* 65, 427–438.
- Hartley, B.S., 1968. Primary structures of serine proteinases. *Biochemical Journal* 110, 1P.
- Hartley, B.S., Brown, J.R., Kauffman, D.L., Smillie, L.B., 1965. Evolutionary similarities between pancreatic proteolytic enzymes. *Nature* 207, 1157–1159.
- Hirs, C.H.W., Moore, S., Stein, W.H., 1960. The sequence of the amino acid residues in performic acid-oxidized ribonuclease. *Journal of Biological Chemistry* 236, 633–647.
- Hunkapiller, M.W., Hewick, R.M., Dreyer, W.J., Hood, L.E., 1991. High-sensitivity sequencing with a gas-phase sequencer. *Methods in Enzymology* 91, 399–413.
- Inglis, A.S., 1983. Cleavage at aspartic acid. *Methods in Enzymology* 91, 324–332.
- Kelleher, N.L., Lin, H.Y., Valaskovic, G.A., *et al.*, 1999. Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. *Journal of the American Chemical Society* 121, 806–812.
- Kim, M.S., Pinto, S.M., Getnet, D., *et al.*, 2014. A draft map of the human proteome. *Nature* 509, 575–581.
- Konigsberg, W., Hill, R.J., 1962. The structure of human hemoglobin: III. The sequence of amino acids in the tryptic peptides of the α chain. *Journal of Biological Chemistry* 237, 2547–2561.
- Kretsinger, R.H., Nockolds, C.E., Coffee, C.J., Bradshaw, R.A., 1971. The structure of a calcium-binding protein from carp muscle. *Cold Spring Harbor Symposia on Quantitative Biology* 36, 217–220.
- Manning, G., Whyte, D.B., Martinez, R., Hunter, T., Sudarsanam, S., 2002. The protein kinase complement of the human genome. *Science* 298, 1912–1934.
- Martin, A.J.P., Synge, R.L.M., 1941. A new form of chromatogram employing two liquid phases. *Biochemical Journal* 35, 1358–1368.
- Moore, S., Stein, W.H., 1993. The chemical structures of pancreatic ribonuclease and deoxyribonuclease. In: Sture, F. (Ed.), Nobel Laureates in Chemistry during the period 1971–1980. Singapore: World Scientific Publishing Co. (Nobel lecture, 1972), Tore Frängsmyr (Editor-in-Charge).
- Peng, Y., Gregorich, Z.R., Valeja, S.G., *et al.*, 2014. Top-down proteomics reveals concerted reductions in myofilament and Z-disc protein phosphorylation after acute myocardial infarction. *Molecular & Cellular Proteomics* 13, 2752–2764.
- Raffioni, S., Luporini, P., Chait, B.T., Disper, S.S., Bradshaw, R.A., 1988. Primary structure of the mating pheromone Er-1 of the ciliate *Euplotes raikovi*. *Journal of Biological Chemistry* 263, 18152–18159.
- Roepstorff, P., Fohlman, J., 1984. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biological Mass Spectrometry* 11, 601.
- Rosenfeld, L., 2012. *Origins of Clinical Chemistry: The Evolution of Protein Analysis*. Amsterdam: Elsevier Publishing Co. p. 43.
- Sanger, F., 1964. The chemistry of insulin, *Nobel Lectures, Chemistry 1942–1962*. Amsterdam: Elsevier Publishing Co. (Nobel lecture, 1958).
- Sanger, F., 1993. Determination of nucleotide sequences in DNA. In: Sture, Forsen (Ed.), Nobel Laureates in Chemistry during the period 1971–1980. Singapore: World Scientific Publishing Co. (Nobel lecture, 1980), Tore Frängsmyr (Editor-in-Charge).
- Schroeder, W.A., 1968. *The Primary Structure of Proteins: Principles and Practices for the Determination of Amino Acid Sequence*. New York, NY: Harper & Row, Publishers.
- Siuzdak, G., 1996. *Mass Spectrometry for Biotechnology*. San Diego, CA: Academic Press.
- Spackman, D.H., Stein, W.H., Moore, S., 1960. The disulfide bonds of ribonuclease. *Journal of Biological Chemistry* 235, 648–659.
- Stein, W.H., Moore, S., 1948. Chromatography of amino acids on starch columns. Separation of phenylalanine, leucine, isoleucine, methionine, tyrosine, and valine. *Journal of Biological Chemistry* 176, 337–365.
- Tang, J., James, M.N., Hsu, I.N., Jenkins, J.A., Blundell, T.L., 1978. Structural evidence for gene duplication in the evolution of the acid proteases. *Nature* 271, 618–621.
- Thomas, K.A., Silverman, R.E., Jeng, I., Baglan, N.C., Bradshaw, R.A., 1981. Electrophoretic heterogeneity and polypeptide chain structure of the γ -subunit of 7S nerve growth factor. *Journal of Biological Chemistry* 256, 9147–9155.
- Walsh, K.A., Neurath, H., 1964. Trypsinogen and chymotrypsinogen as homologous proteins. *Proceedings of the National Academy of Sciences of the United States* 52, 884–889.
- Washburn, M.P., Wolters, D., Yates III, J.R., 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology* 19, 242–247.
- Weber, K., Osborn, M., 1969. The reliability of molecular weight determinations by dodecyl sulfate-polyacrylamide gel electrophoresis. *Journal of Biological Chemistry* 244, 4406–4412.
- Wilhelm, M., Schlegl, J., Hahne, H., *et al.*, 2014. Mass-spectrometry-based draft of the human proteome. *Nature* 509, 582–587.

Posttranslational Modifications: Key Players in Health and Disease

I Lindberg, University of Maryland-Baltimore, MD, USA

JR Peinado, University of Castilla-La Mancha, Ciudad Real, Spain

© 2016 Elsevier Inc. All rights reserved.

Introduction

The one gene–one enzyme hypothesis introduced in 1941 proposed that each gene was responsible for producing a single enzyme that in turn participates in a single step in a specific metabolic pathway (Beadle and Tatum, 1941). Although this was a necessary first step in our understanding of protein translation, since that moment, the concept of protein synthesis has evolved greatly. We now know that one gene can be transcribed into many mRNAs which in turn produce several proteins; and that, although sharing a core amino acid sequence, the resultant proteins may bear many different reversible or irreversible chemical alterations, including the covalent incorporation of lipids, carbohydrates, and other chemical groups. Collectively, the modifications that proteins undergo, either during synthesis or after translation are fully completed, and are designated ‘posttranslational modifications’ (PTMs).

Biologically, PTMs are involved in a variety of cellular activities, such as regulation of gene expression, activation/deactivation of enzymatic activity, protein stability or turnover, and mediation of protein–protein interactions. Therefore, the characterization of PTMs provides invaluable insight into the cellular functions underlying the origin of these processes. In the present article we provide a brief but comprehensive overview of the most important PTMs and explain their major contributions to protein and cellular function.

Proteolytic PTM of Proteins

Before presenting the PTMs that result from reversible or irreversible anchorage of molecules of differing nature to the amino acid core, we first discuss one of the most important, ubiquitous and irreversible protein modifications: the limited and highly specific hydrolysis of peptide and isopeptide bonds of a protein by proteases. Proteases exist in all orders of life and constitute one of the largest enzyme families in humans (Puentes *et al.*, 2003). Limited proteolysis is especially important as part of the maturation process for peptide hormone precursors. A clear example of the posttranslational processing of a protein is the cleavage of pro-opiomelanocortin (Bicknell, 2008), a precursor whose differential cleavage by specific proteases within the secretory pathway gives rise to various product peptide hormones with diverse biological activities. Interestingly, most of the proteases that carry out these maturation cleavages, as well as many other proteases, are synthesized as inactive precursors, termed zymogens, that also require proteolytic processing (often autocatalytic) in order to become proteolytically active.

In addition to involvement in prohormone maturation, proteolysis also regulates a wide array of cellular processes; for example, removal of the signal peptide, which is essential for correct secretory protein maturation. Pathological alterations

in regulated proteolysis result in the loss of many important regulatory processes (e.g., peptide hormone deficiency arising from nonfunctional maturation enzymes). There are a host of other proteolytic events important to essential processes that control cellular homeostasis, such as cell death and cell proliferation, as well as pathological processes such as inflammation, cancer, or arthritis (Barrett *et al.*, 1998). Advances in the identification of proteolytic cleavage sites have been powered by the development of techniques that permit the identification of N- and C-terminal peptides, and these methodologies are now increasingly used for identifying protease substrates and cleavage sites (Rogers and Overall, 2013).

Trimming of peptides by peptidases also falls under the category of proteolytic modification of proteins. Cells contain many enzymes which trim proteins and peptides by a small number of amino acids; these enzymes are termed aminopeptidases and carboxypeptidases. As two examples, the initiating N-terminal methionine of many proteins is removed by a cytosolic aminopeptidase (Bradshaw *et al.*, 1998); and secreted proteins require the action of specific carboxypeptidases that remove basic residues left on the C-terminus following convertase action (Canaff *et al.*, 1999).

Phosphorylation

Although there is currently no method to readily assess the occurrence of every known PTM on a specific protein, a recent study reported the relative abundances of many PTMs found experimentally from high-quality proteome-wide data (Khouri *et al.*, 2011), and deposited in an updated resource (see PTMCuration in ‘Relevant Websites’). Figure 1 depicts the relative abundance of the various PTMs, where phosphorylation appears as the most widespread PTM among proteins. The term phosphorylation appeared in 1954 when it was observed that an enzyme (a protein kinase) was able to transfer a phosphate onto another protein (Burnett and Kennedy, 1954). Immediately following this discovery, other researchers (Fischer and Krebs, 1955; Sutherland and Wosilait, 1955) showed that an enzyme involved in glycogen metabolism was regulated by the addition or removal of a phosphate, suggesting that reversible phosphorylation (dephosphorylation) could somehow regulate enzyme activity. The concept of reversible phosphorylation has stood the test of time, and it is now known that the simple addition of a phosphate group is able to greatly alter the conformation of a protein, modifying its structure and function, while dephosphorylation switches the protein back to its original conformation. The balance between phosphorylation and dephosphorylation provides a host of cellular proteins with varying conformations that directly modify their activities.

The transfer of phosphates onto proteins is catalyzed by a variety of enzymes in the cell, called protein kinases, which

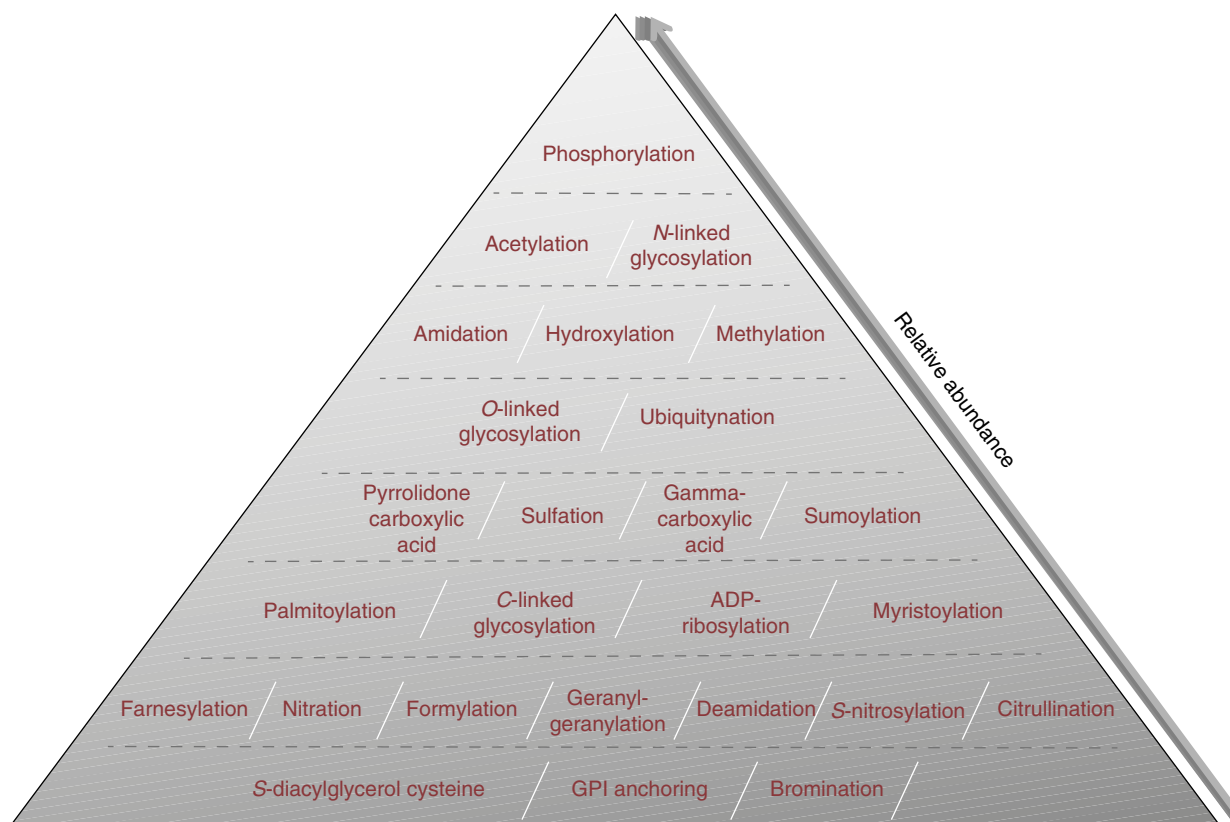


Figure 1 Relative abundance of the most widely distributed PTMs.

share certain biochemical characteristics (reviewed in Sefton, 2001). Kinases constitute 2% of the genome and fall into 10 different families. All kinases possess the ability to remove phosphate from the chemically energy-rich molecule ATP and place it onto the side chain of specific amino acids within proteins. The hydroxyl groups (-OH) of serine, threonine, and tyrosine side chains are by far the most common kinase targets. A second class of enzymes is responsible for the reverse reaction, in which phosphates are removed from a protein. These enzymes are termed protein phosphatases, and they are also frequently regulatory enzymes.

Often, cells use a sequential string of phosphorylation events, termed a 'phosphorylation cascade' as a response to an external signal in order to produce an integrated response to a cellular event. In a kinase cascade, a given kinase is activated by phosphorylation and in turn phosphorylates other kinases, amplifying the original signal. Several different kinase cascades can respond to a cellular signal and become integrated in the final phosphorylation events. For example, cellular insulin signaling is mediated by a phosphorylation cascade that begins with the tyrosine phosphorylation of the insulin receptor, which then induces the cytoplasmic binding of IRS-1 (insulin receptor substrate 1) to these receptors. The receptor then phosphorylates several tyrosine residues within IRS-1, which enables IRS-1 to activate several downstream signaling pathways. Discoveries such as this have uncovered the importance of phosphorylation in pathologies such as diabetes (Nandi *et al.*, 2004). Many other key cellular signaling processes depend directly upon the correct functioning of various

phosphorylation cascades. New kinases are still being discovered; for example, the secreted kinase that likely phosphorylates all secreted proteins, FAM20C, was discovered only in 2010 (reviewed in Tagliabracci *et al.*, 2013); this enzyme phosphorylates proteins at Ser-X-Asp/Glu motifs.

Glycosylation (*N*- and *O*-linked)

Glycosylation is one of the most common and complex forms of PTM of proteins (Ferguson *et al.*, 2009), and in the human genome, about 1–2% of the total genome corresponds to genes that are directly involved in glycan assembly (Schachter and Freeze, 2009). This process involves a covalent attachment of one or several carbohydrate chains to the amide group of the asparagine (*N*-glycoprotein), or, less frequently but still quite often, to the hydroxyl group of serine or threonine (*O*-glycoprotein). *O*-linked glycosylation can occur either in the cytosol or the secretory pathway, while *N*-linked glycosylation occurs only on proteins made within the secretory pathway.

N-linked glycosylation begins with assembly of a ubiquitous 14-residue high-mannose precursor on dolichol, a lipid in the membrane of the rough endoplasmic reticulum (ER). This preformed oligosaccharide then is transferred to specific asparagine (Asn) residues within a consensus sequence (Asn-X-Ser/Thr) on nascent polypeptide chains within the ER lumen. Further modifications to *N*-linked oligosaccharides occur in the ER and are completed in the Golgi complex, with terminal sialylation or 'capping' occurring in the trans-Golgi

network (Lodish, 2000). The co-translational addition of sugars is thought to stabilize the growing protein chain and block inappropriate side chain interactions. An elaborate system of quality control exists to ensure that secretory proteins are correctly folded, and carbohydrates play an important role in this process. N-linked chains are generally about 2–3 kDa in size. On the other hand, O-linked oligosaccharides, which are bound to serine, threonine, or hydroxylysine residues, are generally short, often containing only one to four sugar residues. Protein-bound oligosaccharides contribute heavily to folding, stability and the final bioactivity profile of glycoproteins; other glycosylation events such as the addition of such as mannose 6-phosphate also perform the important role of lysosomal targeting of these proteins.

Since correct glycosylation is essential for obtaining and supporting the normal biological activity of proteins, impairment of glycosylation will lead to the synthesis of glycoproteins with reduced or lost function. The ability of carbohydrates to modulate the activities of proteins has been previously assessed (Lis and Sharon, 1993). Many pathological states are characterized by changes in the carbohydrate structure of cellular glycoproteins, and most of these are due directly to defective activity of specific glycosidases or transferases. As a consequence of the widespread occurrence and variety of glycosylation events, this PTM is the primary cause of variability among proteins with similar amino acid cores.

In healthy people, glycoproteins have a normal protein isoform pattern that may change in a characteristic manner in different diseases. Disturbances of glycosylation may be either acquired or congenital (Schachter and Freeze, 2009). As one of many examples, altered N-glycosylation has been found in several pathologies linked to neurodegenerative diseases. Specifically, in Alzheimer's disease the tau and beta-amyloid peptide precursor proteins, which contribute to the major pathological hallmarks of neurofibrillary tangles and amyloid plaques, respectively, have both been found to be glycosylated (Schedin-Weiss *et al.*, 2013). Considering that the N-glycosylation of proteins was thought to occur within the secretory pathway, i.e., only on extracellular (secreted) proteins or on the extracellular domains of membrane-bound proteins, it is especially intriguing that the cytosolic protein tau was found to be N-glycosylated in Alzheimer's but not in control brains (Wang *et al.*, 1996). Another disease in which altered N-glycosylation occurs is cancer progression, for example, in gastric carcinogenesis (Pinho *et al.*, 2013). O-glycosylation may also be altered in cancer, and it is thought that the altered pattern of surface sugars on tumor cells may be involved in metastasis and/or immune evasion (Wang *et al.*, 1996).

Congenital disorders of glycosylation (CDG) mostly result from genetic defects leading to the deficiency or loss of a specific enzyme activity involved in glycan (sugar) synthesis and processing, or to a deficiency of specific transporters (Freeze *et al.*, 2009; Cylwik *et al.*, 2013a,b). At this moment, 30 CDG subtypes have been identified at a molecular and biochemical level. The clinical manifestations of CDGs are heterogeneous and may be highly variable within the same subtype. Novel insights into certain extremely complex glycosylation pathways have led researchers to a reclassification of the group of CDGs, and today CDGs comprise not only the formerly known multisystem glycosylation defects, but also

certain tissue-specific glycosylation defects. Since diagnosis and treatment depend heavily on the particular underlying glycosylation defects, progress in understanding glycosylation pathways will contribute to improved patient treatment (reviewed in Vodopietz and Bodamer, 2008).

N-Acetylation

Lysine acetylation is a PTM that refers to the addition of an acetyl group (CH_3CO) to the lysine side chain (N-acetylation). This modification was originally observed in abundant proteins such as tubulin and histones, conferring microtubule stability or an increase in transcriptionally-active chromatin, respectively (reviewed in Close *et al.*, 2010). Histone acetylation and deacetylation are the processes by which the histones on lysine residues within the N-terminal tail and on the surface of the nucleosome core are acetylated and deacetylated during normal gene regulation (Grunstein, 1997). Positively charged lysine residues in histone proteins interact electrostatically with negatively charged phosphate groups along the DNA backbone. Acetylation reduces these interactions and loosens the DNA, facilitating its transcription. Since histones were identified as the first substrates of proteins that catalyze the removal of an acetyl moiety, these enzymes were termed histone deacetylases (HDAC) – even though some do not actually target histones. In fact, lysine acetylation has also been found to regulate transcription factors and coactivators including p53 or PGC1 α (Mellert and McMahon, 2009; Yang and Seto, 2008), and this PTM is especially abundant in the mitochondria, where it has been found in 20% of mitochondrial proteins, including metabolically important enzymes and proteins whose activity is linked to life span regulation (Kim *et al.*, 2006).

Recent work indicates that lysine acetylation may be comparable in scope with that of other major PTMs. Although high-resolution mass spectrometry (MS) using a leukemia cell line allowed for the identification of more than 1500 acetylated proteins, a key feature of the acetylome is its cell-type specificity (Choudhary *et al.*, 2009). Notwithstanding this, protein acetylation has been identified in enzymes involved in glycolysis, gluconeogenesis, the tricarboxylic acid and urea cycles, fatty acid metabolism, and glycogen metabolism, as described in hepatocytes (Zhao *et al.*, 2010). This PTM is now seen as a major mechanism by which the activity of multiple substrates is regulated, with key consequences on aging and on diseases associated with aging (Wagner *et al.*, 2011).

Protein N-terminal acetylation is one of the most common covalent modifications of proteins. This modification can occur on the initiating methionine residue, which forms the main target of acetylation, or can take place after the removal of this N-terminal methionine residue. Together, these modifications occur on the vast majority of eukaryotic proteins; in the case of mammalian systems, it has been suggested that up to 90% of the proteins are N-acetylated (Helbig *et al.*, 2010).

Amidation

C-terminal α -amidation is a common secretory PTM mediated by the hydroxylase and lyase activities of a bifunctional

enzyme, peptidylglycine α -amidating monooxygenase (PAM) (Prigge *et al.*, 2000). Amidation serves as a rate-limiting process of the final steps in bioactive peptide synthesis for amidated peptides and therefore is subject to regulation (Eipper *et al.*, 1992). α -amidation has also been implicated in a variety of pathological processes such as neural dysfunction (Bousquet-Moore *et al.*, 2010) and hypertension (Shimosawa *et al.*, 2000).

Methylation

The recognition of protein methylation as an important regulator of protein interactions has grown in recent years. Methylation of proteins at arginine and lysine residues is a PTM involved in many different cellular processes, such as signal transduction, protein translocation, maturation of hnRNPs, or processing of mRNAs and transcriptional control (reviewed in Erce *et al.*, 2012). The functional impact of this PTM has been best characterized in histone proteins (reviewed in Lee *et al.*, 2005), although lysine methylation of other proteins has been emphasized as a prominent participant in non-histone protein bioactivities (Huang and Berger, 2008).

Lipid Modification (Lipidation)

Protein Prenylation

The hydrophobicity of a protein can be modified by the covalent addition of lipids. For example, several proteins, such as Ras and Ras-related GTP-binding proteins undergo prenylation (the addition of isoprenoid groups) in order to be able to attach to the particular cell membrane where they carry out their function. There are two major prenylation modifications: attachment of farnesyl (15-carbon) or geranylgeranyl (20-carbon) isoprenoids to conserved cysteine residues at or near the C-terminus (Zhang and Casey, 1996).

GPI Anchoring

Anchoring of a protein to the membrane can also occur by the binding of the glycolipid glycosylphosphatidylinositol (GPI); this is another PTM that, although not among the most abundant PTMs (Figure 1), is nonetheless ubiquitously found in all kingdoms (Ikezawa, 2002). GPI-linked proteins are thought to be preferentially located in specialized portions of the cell membrane known as lipid rafts, suggesting a high level of organization within plasma membrane microdomains.

Myristoylation and Palmitoylation

Lipidation by myristoylation (the attachment of a 14-carbon fatty acid to the N-terminal glycine of a protein) can have numerous effects, including influencing protein–protein interactions, enhancing interactions of the protein with either organelle or plasma membranes, and altering protein stability (Wright *et al.*, 2010). Palmitoylation, which represents the addition of a 16-carbon fatty acid to a C-terminally located cysteine, also greatly enhances the hydrophobicity of proteins,

contributes to their membrane association, and plays a significant role in subcellular trafficking of proteins between membrane compartments, as well as in modulating protein–protein interactions. In contrast to prenylation and myristoylation, palmitoylation is usually reversible (Charollais and Van Der Goot, 2009), providing cellular flexibility in membrane attachment.

Oxidative Stress-Related PTMs

During the conversion of the flux of nutrients into energy, mitochondria generate intermediates for biosynthesis and reactive oxygen species (ROS) (Cheng and Ristow, 2013). Although ROS may serve as second messengers in signal transduction cascades, any imbalance in ROS production causes oxidative damage to nucleic acids, lipids, and proteins (Cheng and Ristow, 2013). Indeed, ROS and oxidative stress have been consistently identified as important pathogenic components of metabolic diseases (Whaley-Connell *et al.*, 2011). Accordingly, several studies have been aimed at characterizing the oxidatively modified forms of proteins that ultimately contribute to the development of a pathological status and associated metabolic disorders (Evans *et al.*, 2003). Targeted proteomic approaches have been developed to investigate ROS-related PTMs, and these are referred to as ‘redox proteomics.’ Among the redox-related PTMs, hydroxylation, carbonylation, and nitrosylation have been particularly frequently described in adipose tissue, liver, and pancreas (Peinado *et al.*, 2014).

Hydroxylation

This modification consists of the introduction of a hydroxyl group (–OH) into a protein; the most abundant hydroxylated proteins are the structural proteins collagen and elastin, in which hydroxylation of prolines and lysines plays an important role in certain protein cross-linking reactions which contribute to tissue stability. Ascorbate (vitamin C) is a required cofactor for prolyl and lysyl hydroxylase, and a deficiency of this vitamin results in the collagen disease known as scurvy. Additional cross-linking of lysines – also required for tissue structural stability – is accomplished by lysyl oxidase, a copper-requiring enzyme that catalyzes the formation of bridging aldehydes between lysines within collagen and elastin. Aberrant protein hydroxylation as a consequence of increased ROS species has been associated with a higher apoptotic rate in various pathological processes; for example, hydroxylated proteins are especially abundant in mitochondria during the development of type 2 diabetes (Deng *et al.*, 2010).

Carbonylation

Protein carbonylation can also occur from ROS-induced oxidation reactions. Protein carbonyls can be generated through the incorporation of carbon monoxide into the side chains of lysine, arginine, proline, and threonine. Elevated protein carbonylation levels have been associated with several human metabolic diseases (Dalle-Donne *et al.*, 2003).

Nitrosylation

A growing body of experimental evidence has demonstrated that S-nitrosylation of cysteine residues within a broad functional spectrum of proteins is responsible for the ubiquitous influence of nitric oxide (NO) on cellular function (Hess *et al.*, 2005). The emerging recognition that protein S-nitrosylation is involved in a multiplicity of cellular signal transduction pathways points to the possibility that dysregulated S-nitrosylation could contribute to pathophysiologies characteristic of a wide range of disease states (Foster *et al.*, 2009). The relatively recent development of improved methods for the analysis of protein S-nitrosylation has permitted the identification of numerous S-nitrosylated proteins (SNO-proteins) whose levels of S-nitrosylation can be altered in disease, including prominent disorders of the cardiovascular, musculoskeletal, and nervous systems (Foster *et al.*, 2009).

Ubiquitin and Targeted Protein Degradation

When proteins become damaged or misfolded, they must be degraded to prevent the acquisition of potentially abnormal activities. One of the major mechanisms for the elimination of cellular proteins involves the protein complex known as the proteasome. In eukaryotic cells the proteasome comprises a catalytic core as well as regulatory complexes. Proteins destined to become degraded by the proteasome are first tagged by attachment of ubiquitin, a 76-amino acid polypeptide, to lysine residues. The process of ubiquitin attachment to the substrate protein involves multiple ubiquitin additions and thereby constitutes a PTM that drives the protein to the proteasome. Many proteins involved in the regulation of cell proliferation and differentiation, apoptosis, and DNA repair undergo regulated degradation via proteasomal action. In humans, deregulation of the proteasome can contribute to the pathogenesis of a variety of human diseases, such as cancer, myeloproliferative diseases, and neurodegenerative disease (Schmidt and Finley, 2014). More extensive information on ubiquitination can be found in the article 'Ubiquitination' (see 'See also' section).

γ -Carboxylation

γ -Carboxylation is a PTM carried out by the enzyme γ -glutamyl carboxylase during the biosynthesis of vitamin K-dependent proteins (Stenflo and Suttie, 1977). γ -Glutamyl carboxylase oxidizes vitamin K while simultaneously adding CO₂ to protein-bound glutamic acid to form γ -carboxyglutamic acid (known as 'Gla'), which allows these proteins to bind calcium (carboxylation). γ -Carboxylated proteins are involved in both bone formation and in blood coagulation. A deficiency of vitamin K – found in leafy green vegetables such as romaine lettuce, spinach, and kale – can result in problems with hemostasis, and drugs such as warfarin which interfere with vitamin K synthesis serve as powerful anti-coagulants. γ -Carboxyglutamic acid has been observed in both vertebrates and invertebrates, and γ -carboxylase activity is also found in nearly all mammalian tissues (Furie *et al.*, 1999). The

widespread distribution of both Gla and γ -carboxylase activity supports the idea that this PTM plays other, less well-understood biological roles (Kulman *et al.*, 2007).

Conclusions: Many More PTMs Exist, and There Is Cross Talk between Them

The importance of PTMs, both common and uncommon, is now emerging as an additional level of control in biological processes. It should be stressed that although protein phosphorylation and glycosylation are the most widely studied modifications, many other PTMs also occur *in vivo* that we have not discussed here; although several are well known (i.e., sulfation; Figure 1) the function of others is still under study (i.e., sumoylation). During the last years a number of reviews have been written that explore the importance of sulfation and sumoylation, as well as other PTMs such as methylation (Hickey *et al.*, 2012; Paul and Snyder, 2012; Chen *et al.*, 2011). Furthermore, it should be noted that many different PTMs (as well as occurrences of the same PTM) may occur on the same protein, which suggests that there is PTM cross talk. An interesting example is the case of the p53 tumor suppressor protein. The p53 protein may undergo up to six different PTMs, including phosphorylation, acetylation, and sumoylation, that likely modulate its interactions with other proteins and collectively affect its stability and activity (Xu, 2003). It is clear, however, that we are still far from a complete understanding of how the coexistence of several PTMs within a given protein can work together to impact its overall function, and how a given PTM can regulate a wide number of cellular processes, from cell proliferation to cell death.

Given that any covalent modification that occurs *in vivo* on a protein can be considered a PTM, there is a continuous flux of proposals for new PTMs. Mass spectroscopy will continue to play an important role in PTM identification. For example, a new PTM was recently discovered that appears to take place on numerous mammalian proteins; this PTM originates from the reaction of lysines with a primary glycolytic metabolite, 1,3-BPG (bisphosphoglycerate). This reaction both increases the size and inverts the charge potential of the modified residue from positive to negative (Moellering and Cravatt, 2013). Continuous updating of the panel of existing and new PTMs will be required in order to gather comprehensive information on all of the potential modifications that occur on a given protein, and their effects on function. This process has been assisted by the development of predictive algorithms which take advantage of rapid advances in data processing never imagined 20 years ago. A variety of new bioinformatics programs now permit the prediction of potential PTMs within a given protein; and several very useful internet nodes compile all of these predictive programs. Two representative nodes are the CBS Prediction Servers and the ExPASy servers (see Relevant Websites section).

See also: Protein Synthesis/Degradation: Protein Degradation – Intracellular: Ubiquitin, Ubiquitin-Like Proteins, and Proteasome-Mediated Degradation; Protein Synthesis/Degradation:

Translation: Regulated Proteolysis of Signaling Molecules: The Proprotein Convertases

References

- Barrett, A.J., Rawlings, N.D., Woessner, J.F., 1998. Handbook of Proteolytic Enzymes. San Diego, CA: Academic Press.
- Beadle, G.W., Tatum, E.L., 1941. Genetic control of biochemical reactions in neurospora. Proceedings of the National Academy of Sciences of the United States of America 27, 499–506.
- Bicknell, A.B., 2008. The tissue-specific processing of pro-opiomelanocortin. Journal of Neuroendocrinology 20, 692–699.
- Bousquet-Moore, D., Prohaska, J.R., Nillni, E.A., et al., 2010. Interactions of peptide amidation and copper: Novel biomarkers and mechanisms of neural dysfunction. Neurobiology of Disease 37, 130–140.
- Bradshaw, R.A., Brickey, W.W., Walker, K.W., 1998. N-terminal processing: The methionine aminopeptidase and N alpha-acetyl transferase families. Trends in Biochemical Sciences 23, 263–267.
- Burnett, G., Kennedy, E.P., 1954. The enzymatic phosphorylation of proteins. Journal of Biological Chemistry 211, 969–980.
- Canaff, L., Bennett, H.P., Hendy, G.N., 1999. Peptide hormone precursor processing: Getting sorted? Molecular and Cellular Endocrinology 156, 1–6.
- Charollais, J., Van Der Goot, F.G., 2009. Palmitoylation of membrane proteins (Review). Molecular Membrane Biology 26, 55–66.
- Chen, C., Nott, T.J., Jin, J., Pawson, T., 2011. Deciphering arginine methylation: Tudor tells the tale. Nature Reviews. Molecular Cell Biology 12, 629–642.
- Cheng, Z., Ristow, M., 2013. Mitochondria and metabolic homeostasis. Antioxidants & Redox Signaling 19, 240–242.
- Choudhary, C., Kumar, C., Gnad, F., et al., 2009. Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science 325, 834–840.
- Close, P., Creppe, C., Gillard, M., et al., 2010. The emerging role of lysine acetylation of non-nuclear proteins. Cellular and Molecular Life Sciences: CMLS 67, 1255–1264.
- Cylwik, B., Lipartowska, K., Chrostek, L., Gruszewska, E., 2013a. Congenital disorders of glycosylation. Part II. Defects of protein O-glycosylation. Acta Biochimica Polonica 60, 361–368.
- Cylwik, B., Naklicki, M., Chrostek, L., Gruszewska, E., 2013b. Congenital disorders of glycosylation. Part I. Defects of protein N-glycosylation. Acta Biochimica Polonica 60, 151–161.
- Dalle-Donne, I., Giustarini, D., Colombo, R., Rossi, R., Milzani, A., 2003. Protein carbonylation in human diseases. Trends in Molecular Medicine 9, 169–176.
- Deng, W.J., Nie, S., Dai, J., Wu, J.R., Zeng, R., 2010. Proteome, phosphoproteome, and hydroxyproteome of liver mitochondria in diabetic rats at early pathogenic stages. Molecular & Cellular Proteomics: MCP 9, 100–116.
- Eipper, B.A., Stoffers, D.A., Mains, R.E., 1992. The biosynthesis of neuropeptides: Peptide alpha-amidation. Annual Review of Neuroscience 15, 57–85.
- Erce, M.A., Pang, C.N., Hart-Smith, G., Wilkins, M.R., 2012. The methylproteome and the intracellular methylation network. Proteomics 12, 564–586.
- Evans, J.L., Goldfine, I.D., Maddux, B.A., Grodsky, G.M., 2003. Are oxidative stress-activated signaling pathways mediators of insulin resistance and beta-cell dysfunction? Diabetes 52, 1–8.
- Ferguson, M.A.J., Kinoshita, T., Hart, G.W., 2009. Glycosylphosphatidylinositol anchors. In: Varki, A., Cummings, R.D., Esko, J.D., et al. (Eds.), Essentials of Glycobiology. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (Chapter 11). Available at: http://www.cshlpress.com/default.tpl?cart=14199872082488378&fromlink=T&linkaction=full&linksortby=oop_title&eqSKUdataq=666 (accessed 01.01.15).
- Fischer, E.H., Krebs, E.G., 1955. Conversion of phosphorylase b to phosphorylase a in muscle extracts. Journal of Biological Chemistry 216, 121–132.
- Foster, M.W., Hess, D.T., Stamler, J.S., 2009. Protein S-nitrosylation in health and disease: A current perspective. Trends in Molecular Medicine 15, 391–404.
- Freeze, H.H., Esko, J.D., Parodi, A.J., 2009. Glycans in glycoprotein quality control. In: Varki, A., Cummings, R.D., Esko, J.D., et al. (Eds.), Essentials of Glycobiology. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (Chapter 36). Available at: http://www.cshlpress.com/default.tpl?cart=14199872082488378&fromlink=T&linkaction=full&linksortby=oop_title&eqSKUdataq=666 (accessed 01.01.15).
- Furie, B., Bouchard, B.A., Furie, B.C., 1999. Vitamin K-dependent biosynthesis of gamma-carboxyglutamic acid. Blood 93, 1798–1808.
- Grunstein, M., 1997. Histone acetylation in chromatin structure and transcription. Nature 389, 349–352.
- Helbig, A.O., Gauci, S., Raijmakers, R., et al., 2010. Profiling of N-acetylated protein termini provides in-depth insights into the N-terminal nature of the proteome. Molecular & Cellular Proteomics: MCP 9, 928–939.
- Hess, D.T., Matsumoto, A., Kim, S.O., Marshall, H.E., Stamler, J.S., 2005. Protein S-nitrosylation: Purview and parameters. Nature Reviews. Molecular Cell Biology 6, 150–166.
- Hickey, C.M., Wilson, N.R., Hochstrasser, M., 2012. Function and regulation of SUMO proteases. Nature Reviews. Molecular Cell Biology 13, 755–766.
- Huang, J., Berger, S.L., 2008. The emerging field of dynamic lysine methylation of non-histone proteins. Current Opinion In Genetics & Development 18, 152–158.
- Ikezawa, H., 2002. Glycosylphosphatidylinositol (GPI)-anchored proteins. Biological & Pharmaceutical Bulletin 25, 409–417.
- Khoury, G.A., Baliban, R.C., Floudas, C.A., 2011. Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database. Scientific Reports 1. pii: srep00090.
- Kim, S.C., Sprung, R., Chen, Y., et al., 2006. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. Molecular Cell 23, 607–618.
- Kulman, J.D., Harris, J.E., Xie, L., Davie, E.W., 2007. Proline-rich Gla protein 2 is a cell-surface vitamin K-dependent protein that binds to the transcriptional coactivator Yes-associated protein. Proceedings of the National Academy of Sciences of the United States of America 104, 8767–8772.
- Lee, D.Y., Teyssier, C., Strahl, B.D., Stallcup, M.R., 2005. Role of protein methylation in regulation of transcription. Endocrine Reviews 26, 147–170.
- Lis, H., Sharon, N., 1993. Protein glycosylation. Structural and functional aspects. European Journal of Biochemistry/FEBS 218, 1–27.
- Lodish, H., Berk, A., Zipursky, S.L., et al., 2000. Molecular Cell Biology, fourth ed. New York, NY: W.H. Freeman.
- Mellert, H.S., McMahon, S.B., 2009. Biochemical pathways that regulate acetyltransferase and deacetylase activity in mammalian cells. Trends in Biochemical Sciences 34, 571–578.
- Moellering, R.E., Cravatt, B.F., 2013. Functional lysine modification by an intrinsically reactive primary glycolytic metabolite. Science 341, 549–553.
- Nandi, A., Kitamura, Y., Kahn, C.R., Accili, D., 2004. Mouse models of insulin resistance. Physiological Reviews 84, 623–647.
- Paul, B.D., Snyder, S.H., 2012. H(2)S signalling through protein sulfhydration and beyond. Nature Reviews. Molecular Cell Biology 13, 499–507.
- Peinado, J.R., Ruiz, A.D., Fruhbeck, G., Malagon, M.M., 2014. Mitochondria in metabolic disease. Getting clues from proteomic studies. Proteomics 14, 452–466.
- Pinho, S.S., Carvalho, S., Marcos-Pinto, R., et al., 2013. Gastric cancer: Adding glycosylation to the equation. Trends in Molecular Medicine 19, 664–676.
- Prigge, S.T., Mains, R.E., Eipper, B.A., Amzel, L.M., 2000. New insights into copper monooxygenases and peptide amidation: Structure, mechanism and function. Cellular and Molecular Life Sciences: CMLS 57, 1236–1259.
- Puente, X.S., Sanchez, L.M., Overall, C.M., Lopez-Otin, C., 2003. Human and mouse proteases: A comparative genomic approach. Nature Reviews. Genetics 4, 544–558.
- Rogers, L.D., Overall, C.M., 2013. Proteolytic post-translational modification of proteins: Proteomic tools and methodology. Molecular & Cellular Proteomics: MCP 12, 3532–3542.
- Schachter, H., Freeze, H.H., 2009. Glycosylation diseases: Quo vadis? Biochimica et Biophysica Acta 1792, 925–930.
- Schedin-Weiss, S., Winblad, B., Tjernberg, L.O., 2014. The role of protein glycosylation in Alzheimer disease. FEBS Journal 281 (1), 46–62.
- Schmidt, M., Finley, D., 2014. Regulation of proteasome activity in health and disease. Biochimica et Biophysica Acta 1843, 13–25.
- Sefton, B.M., 2001. Overview of protein phosphorylation. In: Bonifacio, J.S., et al. (Eds.), Current Protocols in Cell Biology. Chapter 14, Unit 14.11.
- Shimosawa, T., Kanozawa, K., Nagasawa, R., et al., 2000. Adrenomedullin amidation enzyme activities in hypertensive patients. Hypertension Research: Official Journal of the Japanese Society of Hypertension 23, 167–171.
- Stenflo, J., Suttie, J.W., 1977. Vitamin K-dependent formation of gamma-carboxyglutamic acid. Annual Review of Biochemistry 46, 157–172.
- Sutherland Jr., E.W., Wosilait, W.D., 1955. Inactivation and activation of liver phosphorylase. Nature 175, 169–170.
- Tagliabracci, V.S., Pinna, L.A., Dixon, J.E., 2013. Secreted protein kinases. Trends in Biochemical Sciences 38, 121–130.

- Vodopivec, J., Bodamer, O.A., 2008. Congenital disorders of glycosylation — A challenging group of IEMs. *Journal of Inherited Metabolic Disease* 31, 267–269.
- Wagner, G.R., Payne, R.M., 2011. Mitochondrial acetylation and diseases of aging. *Journal of Aging Research* 2011, 234875.
- Wang, J.Z., Grundke-Iqbal, I., Iqbal, K., 1996. Glycosylation of microtubule-associated protein tau: An abnormal posttranslational modification in Alzheimer's disease. *Nature Medicine* 2, 871–875.
- Whaley-Connell, A., McCullough, P.A., Sowers, J.R., 2011. The role of oxidative stress in the metabolic syndrome. *Reviews in Cardiovascular Medicine* 12, 21–29.
- Wright, M.H., Heal, W.P., Mann, D.J., Tate, E.W., 2010. Protein myristoylation in health and disease. *Journal of Chemical Biology* 3, 19–35.
- Xu, Y., 2003. Regulation of p53 responses by post-translational modifications. *Cell Death and Differentiation* 10, 400–403.
- Yang, X.J., Seto, E., 2008. Lysine acetylation: Codified crosstalk with other posttranslational modifications. *Molecular Cell* 31, 449–461.
- Zhang, F.L., Casey, P.J., 1996. Protein prenylation: Molecular mechanisms and functional consequences. *Annual Review of Biochemistry* 65, 241–269.
- Zhao, S., Xu, W., Jiang, W., *et al.*, 2010. Regulation of cellular metabolism by protein lysine acetylation. *Science* 327, 1000–1004.

Relevant Websites

- <http://www.cbs.dtu.dk/services>
Center for Biological Sequence Analysis.
- <http://www.expasy.org/proteomics>
ExPASy.
- <http://selene.princeton.edu/PTMCuration>
PTMCuration.

Protein Domains: Structure, Function, and Methods

DA Korasick and JM Jez, Washington University in St. Louis, St. Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Within cells, proteins are the workhorses that both drive signal transduction and govern cellular responses. Although proteins perform a myriad of functions, the structure of a single, globular protein can be broken down into distinct structural levels: primary, secondary, tertiary, and quaternary structure. The primary structure of a protein describes the sequence of amino acids, with directionality from amino- to carboxy-terminus, which comprises a polypeptide chain. The secondary structure of a protein refers to the structural elements – for example, α -helices and β -strands – that form the protein into more than simply a string of amino acids. The tertiary structure of a protein describes how the primary and secondary structural elements fold into the overall three-dimensional arrangement of the polypeptide. Finally, the quaternary structure describes the higher order assembly of multiple polypeptides into oligomers. Although some proteins function as monomers of a single polypeptide chain, many are assemblies of copies of one (i.e., homo-oligomers) or multiple polypeptides (hetero-oligomers). The combination of these structural levels dictates the function of a single protein by influencing aspects of its biochemical reactivity, regulation, protein–protein interaction specificity, and posttranslational modifications.

These general classifications of protein structure are somewhat limited, as evolution promotes the formation of more complex proteins and structural elements for diverse biological and chemical functions. As the expanse of both sequence and structure space is explored, the classification of protein structure and function expands to include structural motifs and units that convey specific functionality within proteins. To this end, the building blocks of many complex proteins are often domains.

By definition, investigation of protein domains and their functions is guided by understanding protein structure at a molecular level. Examining the structural characteristics of proteins enable scientists to understand important facts about protein function and to synthesize what is known to make further predictions regarding the function of uncharacterized proteins. Further, elucidation of protein structure coupled with phylogenetic analysis allows for generation of informed hypotheses involving strategies and themes in protein evolution. Here we discuss protein structure and domains; common themes in domain functionality; advances in bioinformatic tools that improve prediction of domains and domain functionality; and applications of domain research within protein science and cell biology.

Protein Structure: Motifs, Folds, and Domains

The classic definition of a protein domain is that a polypeptide chain comprising the domain remains highly stable such that it retains its three-dimensional structure and biological

function independent of the rest of the larger protein (Richardson, 1981). It is important to distinguish a protein domain from other primary and secondary structural units. A domain refers to the structure and function of a portion of a protein, not simply the sequence. In other words, the amino acid sequence may suggest the presence of a specific domain, but the presence of a defined domain may only be confirmed by linking the structure and function of that region. To this end, primary sequence alignments of proteins can provide important clues in identifying both the presence of protein domains, and defining known protein domain boundaries. Standard biochemical, biophysical, and bioinformatic methods for assessing the presence of a domain and its boundaries are discussed later in this article.

Domains are independent functional units that contribute to the tertiary structure of a globular protein. Domains range in size from less than 40 amino acids to around 700 amino acids (Orengo *et al.*, 2002) with an average around 100 amino acids (Wheeler *et al.*, 2000). The energetics of protein folding influences the general limits of domain size, although other factors, including metals, disulfide bonds, other cofactors, and hydrophobic interactions, also constrain size of a domain (Xu and Nussinov, 1998).

The structure of individual domains and small proteins is defined by the organization of secondary structure and super-secondary structures, also called motifs. Motifs are defined by the composition and organization of α -helices, β -structures, and unstructured loops and often comprise repeating units within a protein region. Motifs may be simple – for example, helix-loop-helix motifs – or more complex – for example, α/β -barrels. Excellent references about protein structure motifs illustrate the range from simple to complex in the organization of protein secondary structure (Branden and Tooze, 1999).

Although the structure of domains is defined by secondary elements and motifs, the presence of a motif or series of motifs does not dictate the function of a domain or a protein. Rather, individual or multiple motifs organize to form a protein fold (Figure 1). A three-dimensional fold, however, is distinct from a protein domain. Although a domain preserves both function and structure, the term fold only describes the overall three-dimensional structure. Further, domains may be described and classified by their three-dimensional fold, but the presence of a fold does not necessarily dictate the presence of a specific domain. A series of proteins may share common structural motifs or folds, but possess unrelated functionality.

A classic example of a family of proteins that possess structural similarity due to the presence of a common fold and motif, but with low sequence similarity and a wide variety of functionalities are the α/β hydrolases (Nardini and Dijkstra, 1999). The α/β hydrolases are defined by the presence of a complex mixed α/β structural motif, but range in functionality from those lacking catalytic activity – for example, the *Arabidopsis thaliana* α/β hydrolase KARRIKIN INSENSITIVE2 (KAI2; Guo *et al.*, 2013) – to a host of lipases, esterases, and proteases

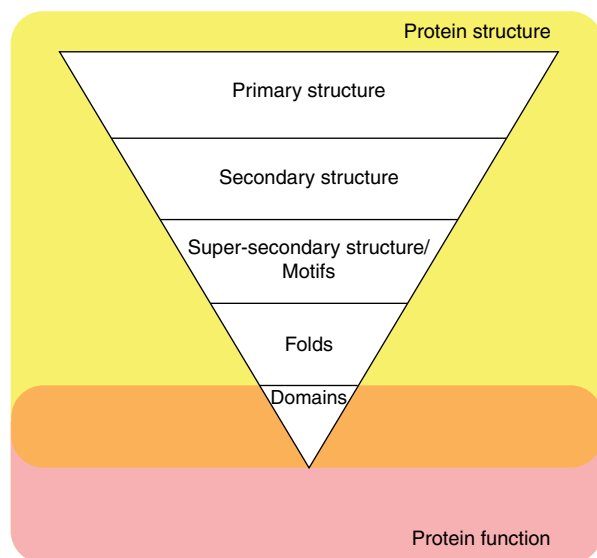


Figure 1 A hierarchical representation of protein structure. This graphic illustrates the various levels of protein structure from least structured (primary structure, top) to most structured (folds and domains). Because domains are functional–structural modules within proteins, the progression from purely structural units to functional units is illustrated.

across numerous species and kingdoms (Nardini and Dijkstra, 1999). Although structural variance exists within this protein family, this example outlines the importance of distinguishing between architectural similarity and functional redundancy.

The organization of protein motifs, folds, and domains has been laid out in a number of databases and resources. Some examples of such databases are outlined later in this article.

Protein Domain Functions

Protein domains are the smallest functional unit of a globular protein. It is important to acknowledge the vast number of structural domains that have been identified. Each of these structural domains conveys a specific cellular function. Because space limitations prevent in-depth analysis of all protein domains, below is a brief discussion of representative domains classified by general function: interaction and binding domains, DNA-binding domains, and enzymatic/catalytic domains.

Interaction and Binding Domains

The function of many protein domains is to facilitate protein–protein interactions by binding to specific recognition motifs or structural features in various ligand molecules (Table 1; Figure 2). These targets may include posttranslational modifications, such as phosphorylated tyrosine, serine, and threonine residues. For example, the Src Homology 2 (SH2) domain is highly conserved in tyrosine kinase signaling pathways and facilitates protein–protein interaction by recognizing phosphotyrosine-containing sequences (Liu *et al.*,

Table 1 Summary of representative interaction and binding domain recognition targets. Different interaction/binding domains target specific molecules and target sequences

Interaction and binding domain	Target
C2	Membrane lipids and Ca^{2+}
EH (Eps15 homology)	S/T-N-P-F-Z
EF-hand/calmodulin	Ca^{2+}
EVH1 (enabled/VASP homology 1)	D/E-Z-P-P-P
PB1 (Phox Bem1p)	Charged surfaces
PDZ	X-X-S/T-X-V-Cterm
PH (pleckstrin homology)	PIP2, PIP3
PTB (phosphotyrosine binding)	Z-X-N-P-X-pY
SH2 (Src homology 2)	pY-X-X-Z
SH3 (Src homology 3)	R/K-X-X-P-X-X-P
WW	P-P-X-Y
14-3-3	R-S-X-pS-X-P

Note: X = variable residue; Z = hydrophobic residue.

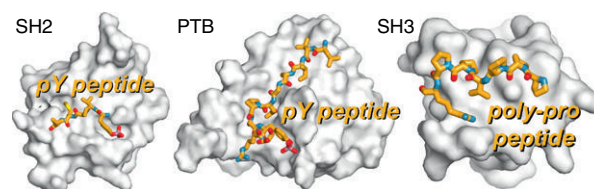


Figure 2 Representative protein–protein interaction domains. Protein surface representations of the v-Src SH2 domain (PDB: 1SHA; Waksman *et al.*, 1992), insulin receptor substrate-1 PTB domain (PDB: 1IRS; Zhou *et al.*, 1996), and the SEM-5 SH3 domain (PDB: 1SEM; Lim *et al.*, 1994) are shown with target peptides bound (gold stick). Phosphotyrosine (pY) peptides are bound to the SH2 and PTB domains and a poly-proline peptide is bound to the SH3 domain.

2012). Similarly, phosphotyrosine binding (PTB) domains also recognize phosphotyrosines, but with a different sequence context (Kafasla *et al.*, 2012). The identity of the phosphorylated residue can also confer interaction specificity. For example, 14-3-3 proteins target phosphorylated serine and threonine residues (Aitken *et al.*, 1995).

Distinctive amino acid sequence features also provide recognition sites for interaction domains. Src homology 3 (SH3), enabled/VASP homology 1 (EVH1), and WW domains recognize poly-proline repeats (Kay, 2012; Renfranz and Beckerle, 2002; Ilsley and Sudol, 2002), whereas, Eps15 homology (EH) and PDZ domains bind specific amino acid sequence motifs (Table 1; Confalonieri and Di Fiore, 2002; Ye and Zhang, 2013). Alternatively, surface features of macromolecules may mediate interactions. For example, the Phox Bem1p (PB1) domain uses concentrated patches of positive and negative charge located, respectively, on opposite faces of the domain for front-to-back interaction with other PB1 domain-containing proteins (Sumimoto *et al.*, 2007). Other domains, such as C2, EF-hand, and pleckstrin homology (PH) domains, facilitate protein interactions either with specific lipid molecules or ions (Corbalan-Garcia and Gomez-Fernandez, 2014; Kawasaki *et al.*, 1998; Lemmon *et al.*, 2007).

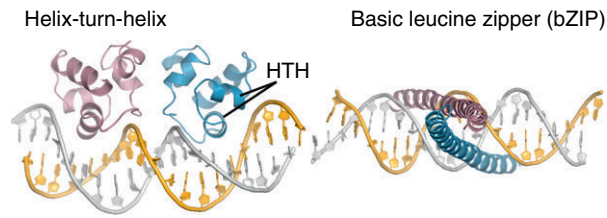


Figure 3 Representative protein–DNA interaction domains. The helix–turn–helix domain of the phage 434 repressor (PDB: 2OR1; Aggarwal *et al.*, 1998) and the basic leucine zipper domain of the cAMP responsive element-binding protein (CREB; PDB: 1DH3; Schumacher *et al.*, 2000) are shown as ribbon diagrams. The helix–turn–helix structural motif of phage 434 repressor is indicated by HTH.

DNA-Binding Domains

Although some interaction domains facilitate the interaction with other proteins, membrane lipids, or signaling molecules, a separate subsection of domains enable interaction of a protein with DNA. Such domains are common in transcription factors (Figure 3). Among the most prevalent of DNA-binding domains in prokaryotes are the helix–turn–helix and winged–helix structures that place an α -helix for interaction with bases in the major groove of the target DNA sequence (Harrison and Aggarwal, 1990; Brennan, 1993). In eukaryotes, the basic leucine zipper (bZIP) and zinc-finger domains are common transcription factor features. The bZIP domains share two characteristics – a hydrophobic, leucine-rich dimerization region and a lysine/arginine-rich DNA-binding region (Ellenberger, 1994). In contrast, zinc-finger transcription factors often consist of multiple ‘fingers’ and can recognize either DNA or RNA (Hall, 2005; Klug, 2010). Likewise, another common DNA-binding domain is the B3 domain, which is conserved in higher land plants (Romanel *et al.*, 2009). Some transcription factors also contain protein–protein interaction domains. For example, the signal transducers and activators of transcription (STAT) proteins are phosphorylated on a tyrosine to mediate homodimerization via an SH2 domain, which leads to formation of a DNA-binding site (Chen *et al.*, 1998).

Enzymatic and Catalytic Domains

Proteins domains are not limited to serving as mediators of interactions, but can also provide enzymatic functions. In eukaryotes, protein phosphorylation is a common posttranslational modification in up to 30% of cellular proteins (Ubersax and Ferrell, 2007). Although phosphorylation occurs in a diverse range of proteins and amino acid sequence contexts, all protein kinases share a conserved three-dimensional structure defined by a canonical catalytic domain of approximately 250 amino acids (Ubersax and Ferrell, 2007). The specificity for tyrosine versus serine/threonine is determined by the depth of the active site cleft with additional structure features, including various protein interaction domains, leading to proximal and distal determinants of substrate recognition (Ubersax and Ferrell, 2007). Although not as abundant as kinases, protein phosphatases also share a common catalytic domain fold with combinatorial interaction between shared catalytic subunits

and one of a large number of regulatory subunits to provide target specificity (Shi, 2009).

Intrinsically Disordered Domains

Although protein domains, by definition, are classified and described by their overall ordered structure, an increasing number of protein domains and proteins are characterized by their intrinsic lack of ordered tertiary structure. These domains, or proteins, are described as intrinsically disordered domains (IDDs). IDD are generally considered to be flexible and unfolded in physiologically relevant conditions (Fink, 2005; Uversky and Dunker, 2010). It is hypothesized that conformational flexibility provided by overall lack of structure allows for wider, heterogeneous interaction within cells. The alternate conformations of IDDs may be induced by post-translational modification, or interaction with a potential substrate (Dunker *et al.*, 2008; Wright and Dyson, 2009). The importance of IDDs in human disease – for example, α -synuclein (Dunker *et al.*, 2008) – and in plant developmental signaling – for example, Aux/IAA repressor proteins and SAUR proteins (Wang and Estelle, 2014), for example, is shifting our understanding of proteins and protein structure. In this way, understanding and characterizing the lack of structure of these IDDs is as important as the efforts imposed upon understanding protein structure. This is an emerging frontier area in protein structure and function.

Identifying and Classifying Protein Domains

The hallmark of identifying the presence of a protein domain is to understand its three-dimensional structure. Visualizing protein structure occurs largely through two separate techniques: X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy. Through the use of these two techniques, scientists are able to generate models of the three-dimensional structure of a protein and analyze the structure to understand the function. To bypass these specialized techniques; however, bioinformatics platforms can also predict the presence of domains based on a number of different factors. Overall, the combination of these techniques enables both direct assessment of protein functionality and an understanding of how proteins work within a cell.

Structural Determination by Protein Crystallography and NMR Spectroscopy

There are two main methodologies used for generating a three-dimensional model of a protein. The most commonly used method is X-ray crystallography, which involves six major steps: (1) protein purification, (2) growing protein crystals, (3) optimizing the crystals for uniform crystal packing, (4) obtaining X-ray diffraction data, (5) phasing the data to generate an electron density map, and (6) building and refining a structural model of the protein. A major limitation in acquiring a structural model by X-ray crystallography is the generation of diffraction-quality crystals. To achieve crystal packing, the protein must be highly ordered and, for the most

part, rigid. Extended regions of disordered polypeptide, or the existence of major conformational changes within a protein or a protein subunit or domain may prevent crystallization.

NMR spectroscopy is also a powerful tool for obtaining protein structural data. In contrast to X-ray crystallography, NMR spectroscopy only requires soluble protein – i.e. no crystal formation. Instead, any soluble, stable protein that is highly ordered can be labeled with ^{15}N and ^{13}C . Then, using solution NMR spectroscopy, one can obtain structural information from the two-dimensional chemical shift data (Wüthrich, 1990). Using NMR for obtaining protein structural data has two major limitations. The proteins must be relatively small (<70 kDa) and any disordered regions will not yield a single structural result; rather, conformationally flexible regions will appear in a number of possible states. Depending upon the final application of the structural data, this second limitation may either obstruct analysis or provide interesting insight into the dynamics of the protein structure.

Ultimately, both X-ray crystallography and NMR spectroscopy are extremely powerful and complementary tools for generating high resolution atomic models of globular macromolecules. Both methods allow for identifying domain boundaries visually: if the general expected structure of the domain is known, obtaining the structure of the overall protein will allow for definitive identification of the domain boundaries. In addition, both techniques are especially useful for investigating the structure of a single protein domain if the boundaries of the domain are either already known or inferred – by definition, protein domains are independent, stably folded structural units. Moreover, protein domains are usually much smaller than an entire protein, which eliminates one of the major limitations of NMR.

Despite the independently stable nature of protein domains, understanding where domain boundaries exist can be difficult and unclear. One of the major methods for experimentally determining domain boundaries is through limited proteolysis. In limited proteolysis, proteins are nonspecifically digested in solution by specific or nonspecific proteases, such as trypsin, chymotrypsin, or proteinase K. The general principle of limited proteolysis is that, in solution, the folded, globular protein will not be cleaved, whereas disordered regions flanking highly ordered globular domains are proteolyzed. Thus, incubation of a protein domain with a protease will, over a time course, trim the protein to its more rigid core, thereby affording the globular domain.

Proteins subjected to limited proteolysis may be further examined by a number of analytical methods, including mass spectrometry for identification of domain boundaries and preparation of the trimmed protein for structure determination by either X-ray crystallography or NMR spectroscopy. A classic experimental approach is to monitor limited proteolysis reactions over time by SDS-PAGE and then subjecting the final product to in-gel trypsin digest followed by mass spectrometry to elucidate the core structure of the domain in question (Stroh *et al.*, 2005; Gao *et al.*, 2005). Limited proteolysis has also proven to be a useful tool in crystallizing previously recalcitrant proteins (Dong *et al.*, 2007; Wernimont and Edwards, 2009).

Although these approaches work for studying protein structure, the study of IDD requires other approaches. For

example, X-ray crystallography cannot be used to determine the three-dimensional structure or conformation of IDDs due to the lack of rigidity. Likewise, NMR, while useful for understanding intramolecular spatial relativity, cannot provide an overall impression of the shape or structure of IDDs. Instead, other methods, such as small-angle X-ray scattering (SAXS), fluorescence spectroscopy, and circular dichroism (CD) can be used to characterize and monitor overall molecular shape, intramolecular interactions and positioning, and the presence and in-solution integrity of secondary structure (Schwalbe *et al.*, 2014). Furthermore, a host of other techniques have been developed and employed to characterize and study IDDs (Tomba, 2011).

Bioinformatic Platforms

The Protein Data Bank (PDB) is the definitive central repository for all protein structural data (Bernstein *et al.*, 1997). As of May 2014 there are ~100,000 individual structures deposited in the PDB. Because knowing the three-dimensional structure of proteins directly impacts our understanding of biological function and evolutionary mechanisms, a number of automated bioinformatic platforms can sort and classify this wealth of structural data. The general assumption employed by each of these databases focuses on identifying structural similarity and extending the evolution of individual domain structure to provide possible biological function (Hadley and Jones, 1999). In the following section, we summarize a number of databases and bioinformatic tools for classifying protein structures, inferring aspects of structural ancestry and evolution, and for predicting secondary and tertiary structure of proteins.

To understand the manner by which many of the bioinformatics databases classify protein domains, one must make distinctions between the levels of domain classification. Generally, protein domains can be described by class, fold, and superfamily (Hadley and Jones, 1999). The most general classification, class, refers to the secondary structure composition of the domain – for example, α -helical versus β -strand structures. Fold is a more specific term as it describes how the structural elements within a structural domain are connected and interact. The final classification is superfamily, which links three-dimensional fold with function. The latter classification is useful in evolutionary analysis, as proteins that share a three-dimensional fold and similar function are likely evolutionarily related. In addition, domains of high functional relatedness and high sequence identity are sometimes classified into families (Hadley and Jones, 1999). These classifications of domains provide for structural analysis and comparison using these databases.

The most utilized servers for domain classification and identification are the Structural Classification of Proteins (SCOP), the Class, Architecture, Topology, and Homologous superfamily (CATH), and Families of Structurally Similar Proteins (FSSP) databases (Hadley and Jones, 1999). The SCOP database (Murzin *et al.*, 1995) is a useful tool for understanding the contributions of different motifs on the structure of domains and multi-domain proteins. The SCOP database organizes proteins into five unique classes by fold

composition: (1) α proteins, whose structure is comprised of mostly α -helices; (2) β proteins, whose structure is comprised of mostly β -sheets; (3) α/β proteins, whose fold integrates both α -helices and β -strands; (4) $\alpha + \beta$ proteins, whose fold contains both α -helices and β -strands that are largely separated in the global domain or protein fold; and (5) multi-domain proteins, which may contain distinct domains from one or more than one of the four aforementioned classes. Additional classes of proteins, including membrane proteins, small proteins, designed proteins, and peptides, are also represented in this database.

In contrast to SCOP, the CATH database (Orengo *et al.*, 2002) employs automated computer comparisons to classify and identify protein domains. In essence, CATH generates a hierarchical structural lineage of a protein based on automated classification at the class level to further assign similarities in topology (or three-dimensional fold) and at the homologous superfamily level. The architecture level is assigned manually. Together, the hierarchically grouping of each queried protein sequence at each level allows for indexing and comparison to proteins of known fold and function. Overall, this database allows for comparison of fold and even evolutionary ancestry.

The FSSP database is also a fully automated platform that employs statistical Z-scores to compare structural homology (Holm and Sander, 1999). Unlike the SCOP and CATH databases, FSSP uses the Dali program to directly compare a queried structure to other structures in the PDB. Based on the Z-score, FSSP generates a ranking of three-dimensional similarity where a higher Z-score correlates to high structural homology. Matches with a Z-score less than two are considered to lack structural similarity. Comprehensive comparison of SCOP, CATH, and FSSP (Hadley and Jones, 1999; Getz *et al.*, 2002) reveals both inherent advantages and disadvantages of each individual platform.

Whereas these SCOP, CATH, and FSSP are important and powerful tools, implementing libraries of hidden Markov models (HMMs) in protein and nucleotide sequence analysis (Hughey and Krogh, 1996) allowed for improvement of domain identification and classification databases. For instance, the SUPERFAMILY database examines the evolutionary relationships of an individual protein fold or domain (Gough and Chothia, 2002; Wilson *et al.*, 2009). Users can subject a nucleotide or protein sequence to be grouped into SCOP superfamily classifications using a library of HMMs based on each of the SCOP superfamilies. SUPERFAMILY compares the queried sequence to HMMs of all proteins with a known structure and then allows comparison of the queried structure to all sequenced genomes to provide insight into the evolutionary lineage of domains. Overall, this level of analysis allows users to examine conservation of a domain regardless of the presence of amino acid sequence identity.

Additionally, implementing HMMs on multiple sequence alignments led to the development of databases like Simple Modular Architecture Research Tool (SMART; Letunic *et al.*, 2012; Schultz *et al.*, 1998) and Pfam (Finn *et al.*, 2014). Both databases are useful for domain identification and classification at the protein sequence level. Likewise, both the I-TASSER (Zhang, 2008) and Phyre2 (Kelley and Sternberg, 2009) servers provide similar services. Building upon previously mentioned databases, both I-TASSER and Phyre2 use

HMMs to classify proteins whose three-dimensional structures have not yet been elucidated. Both servers are also capable of generating homology models for use in understanding structure-function relationships, for informational purposes, or for use as search models for molecular replacement in X-ray crystallography.

Overall, the development of these bioinformatic resources enable structural, functional, and evolutionary analyses of many proteins without obtaining direct structural data. The value and power of these databases lies in the ease with which they generate new testable hypotheses.

Domains in Evolution: Modularity and Combinatorial Protein Structure

Development of biochemical, biophysical, and bioinformatic tools to identify and classify protein domains is a means of linking and understanding protein structure in the context of evolution. Another way of describing domains is as structural modules that lead to functional diversity in evolution (Moore *et al.*, 2008). The smallest proteins often possess only a single domain, although many proteins contain multiple domains. The fundamental idea that proteins evolved to have either repeated domains or multiple domains has been postulated for decades. Recent advances in bioinformatic resources, however, provide a clearer view of structural similarity and homology with functionality across genomes and over evolutionary timescales. This evidence, then, suggests evolution exploits the modularity of protein domains to create multi-domain proteins with more complex and functions.

Multi-domain protein families evolved to perform complex biochemical and cell signaling processes. A classic example of a multi-domain protein is pyruvate kinase, which catalyzes the conversion of phosphoenolpyruvate to pyruvate in glycolysis (Gupta and Bamezai, 2010). Pyruvate kinase contains three modular domains: an α/β -barrel domain, a β -barrel domain, and an $\alpha/\beta/\alpha$ -sandwich domain. These three domains work in concert to catalyze the transfer of a phosphate group from phosphoenolpyruvate to adenosine diphosphate (ADP) to generate pyruvate and adenosine triphosphate (ATP). Interestingly, this three-domain protein is retained throughout almost all organisms and the modular structure of the protein is also largely conserved.

Protein domains can also greatly expand the diversity of function through reorganization of functional modules in a combinatorial manner. For example, non-ribosomal peptide synthetases (NRPS) are large proteins that possess multiple enzymatic domains with each domain catalyzing a specialized reaction (Strieker *et al.*, 2010). These large proteins synthesize non-ribosomal peptides, which are often important natural products with antimicrobial or other qualities. Importantly, these proteins evolved as gene clusters that encode an ordered series of domains capable of performing intricate biosynthetic steps, including adenylation, cyclization, condensation, epimerization, methylation, and oxidation/reduction reactions for the synthesis of non-ribosomal peptides. Moreover, the diverse combination of NRPS domains directly results in chemical diversity in the resulting molecules.

The study of NRPS, as well as other multi-domain proteins, led to the idea that scientists can engineer multi-domain proteins to perform a custom biochemical process (Cane *et al.*, 1998). This science, often called combinatorial protein biochemistry or protein engineering, exploits the independent stability and modularity of protein domains to generate novel unnatural proteins by domain swapping. This engineering step has applications throughout many scientific fields including, but not limited to novel natural product biosynthesis, attenuation or perturbation of signaling cascades, and potentially the creation of proteins of novel function.

Conclusions

Proteins are the workhorses of the cells – they perform nearly all of the cellular processes that enable life. Central to their ability to manipulate biochemical processes, protein domains provide modular functionality to tailor macromolecular interactions, binding specificity, and chemistry. Advancements in elucidating the structures of domains, identifying domain functions, predicting domain boundaries, and inferring domain functionality, all contribute to improving our understanding of both protein function and evolution in cellular systems.

See also: Molecular Principles, Components, Technology, and Concepts: Proteins: Folding, Misfolding, Disordered Proteins, and Related Diseases; NMR in Structural and Cell Biology

References

- Aggarwal, A.K., Rodgers, D.W., Drott, M., Ptashne, M., Harrison, S.C., 1998. Recognition of a DNA operator by the repressor of phage 434: A view at high resolution. *Science* 242, 899–907.
- Aitken, A., Jones, D., Soneji, Y., Howell, S., 1995. 14-3-3 Proteins: Biological function and domain structure. *Biochemical Society Transactions* 23, 605–611.
- Bernstein, F.C., Koetzle, T.F., Williams, G.J., *et al.*, 1997. The Protein Data Bank: A computer-based archival file for macromolecular structures. *Journal of Molecular Biology* 112, 535–542.
- Branden, C., Tooze, J., 1999. *Introduction to Protein Structure*, second ed. New York, NY: Garland Science.
- Brennan, R.G., 1993. The winged-helix DNA-binding motif: Another helix-turn-helix takeoff. *Cell* 74, 773–776.
- Cane, D.E., Walsh, C.T., Khosla, C., 1998. Harnessing the biosynthetic code: Combinations, permutations, and mutations. *Science* 282, 63–68.
- Chen, X., Vinkemeier, U., Zhao, Y., *et al.*, 1998. Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA. *Cell* 93, 827–839.
- Confalonieri, S., Di Fiore, P.P., 2002. The Eps15 homology (EH) domain. *FEBS Letters* 513, 24–29.
- Corbalan-Garcia, S., Gómez-Fernández, J.C., 2014. Signaling through C2 domains: More than one lipid target. *Biochimica et Biophysica Acta* 1838, 1536–1547.
- Dong, A., Xu, X., Edwards, A.M., *et al.*, 2007. In situ proteolysis for protein crystallization and structure determination. *Nature Methods* 4, 1019–1021.
- Dunker, A.K., Silman, I., Uversky, V.N., Sussman, J.L., 2008. Function and structure of inherently disordered proteins. *Current Opinion in Structural Biology* 18, 756–764.
- Ellenberger, T., 1994. Getting a grip in DNA recognition: Structures of the basic region leucine zipper, and the basic region helix-loop-helix DNA-binding domains. *Current Opinion in Structural Biology* 4, 12–21.
- Fink, A.L., 2005. Natively unfolded proteins. *Current Opinion in Structural Biology* 15, 35–41.
- Finn, R.D., Bateman, A., Clements, J., *et al.*, 2014. Pfam: The protein families database. *Nucleic Acids Research* 42, D222–D230.
- Gao, X., Bain, K., Bonanno, J.B., *et al.*, 2005. High-throughput limited proteolysis/mass spectrometry for protein domain elucidation. *Journal of Structural and Functional Genomics* 6, 129–134.
- Getz, G., Vendruscolo, M., Sachs, D., Domany, E., 2002. Automated assignment of SCOP and CATH protein structure classifications from FSSP scores. *Proteins* 46, 405–415.
- Gough, J., Chothia, C., 2002. SUPERFAMILY: HMMs representing all proteins of known structure, SCOP sequence searches, alignments and genome assignments. *Nucleic Acids Research* 30, 268–272.
- Guo, Y., Zheng, Z., La Clair, J.J., Chory, J., Noel, J.P., 2013. Smoke-derived karrikin perception by the α/β -hydrolase KAI2 from *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America* 110, 8284–8289.
- Gupta, V., Bamezai, R.N., 2010. Human pyruvate kinase M2: A multifunctional protein. *Protein Science* 19, 2031–2044.
- Hadley, C., Jones, D.T., 1999. A systematic comparison of protein structure classifications: SCOP, CATH and FSSP. *Structure* 7, 1099–1112.
- Hall, T.M., 2005. Multiple modes of RNA recognition by zinc finger proteins. *Current Opinion in Structural Biology* 15, 367–373.
- Harrison, S.C., Aggarwal, A.K., 1990. DNA recognition by proteins with the helix-turn-helix motif. *Annual Review of Biochemistry* 59, 833–869.
- Holm, L., Sander, C., 1999. Protein folds and families: Sequence and structure alignments. *Nucleic Acids Research* 27, 244–247.
- Hughey, R., Krogh, A., 1996. Hidden Markov models for sequence analysis: Extension and analysis of the basic method. *Computer Applications in the Biosciences* 12, 95–107.
- Ilsley, J.L., Sudol, M., Winder, S.J., 2002. The WW domain: Linking cell signalling to the membrane cytoskeleton. *Cell Signalling* 14, 183–189.
- Kafasla, P., Mickleburgh, I., Llorian, M., *et al.*, 2012. Defining the roles and interactions of PTB. *Biochemical Society Transactions* 40, 815–820.
- Kawasaki, H., Nakayama, S., Kretsinger, R.H., 1998. Classification and evolution of EF-hand proteins. *Biomaterials* 11, 277–295.
- Kay, B.K., 2012. SH3 domains come of age. *FEBS Letters* 586, 2606–2608.
- Kelley, L.A., Sternberg, M.J., 2009. Protein structure prediction on the Web: A case study using the Phyre server. *Nature Protocols* 4, 363–371.
- Klug, A., 2010. The discovery of zinc fingers and their applications in gene regulation and genome manipulation. *Annual Review of Biochemistry* 79, 213–231.
- Leemans, M.A., 2007. Pleckstrin homology (PH) domains and phosphoinositides. *Biochemical Society Symposium* 74, 81–93.
- Letunic, I., Doerks, T., Bork, P., 2012. SMART 7: Recent updates to the protein domain annotation resource. *Nucleic Acids Research* 40, D302–D305.
- Lim, W.A., Richards, F.M., Fox, R.O., 1994. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* 372, 375–379.
- Liu, B.A., Engelmann, B.W., Nash, P.D., 2012. The language of SH2 domain interactions defines phosphotyrosine-mediated signal transduction. *FEBS Letters* 586, 2597–2605.
- Moore, A.D., Björklund, A.K., Ekman, D., Bornberg-Bauer, E., Elofsson, A., 2008. Arrangements in the modular evolution of proteins. *Trends in Biochemical Sciences* 33, 444–451.
- Murzin, A.G., Brenner, S.E., Hubbard, T., Chothia, C., 1995. SCOP: A structural classification of proteins database for the investigation of sequences and structures. *Journal of Molecular Biology* 247, 536–540.
- Nardini, M., Dijkstra, B.W., 1999. Alpha/beta hydrolase fold enzymes: The family keeps growing. *Current Opinion in Structural Biology* 9, 732–737.
- Orengo, C.A., Bray, J.E., Buchan, D.W., *et al.*, 2002. The CATH protein family database: A resource for structural and functional annotation of genomes. *Proteomics* 2, 11–21.
- Renfranz, P.J., Beckerle, M.C., 2002. Doing (F/L)PPPPs: EVH1 domains and their proline-rich partners in cell polarity and migration. *Current Opinion in Structural Biology* 14, 88–103.
- Richardson, J.S., 1981. The anatomy and taxonomy of protein structure. *Advances in Protein Chemistry* 34, 167–339.
- Romanel, E.A., Schrago, C.G., Couñago, R.M., Russo, C.A., Alves-Ferreira, M., 2009. Evolution of the B3 DNA binding superfamily: New insights into REM family gene diversification. *PLoS One* 4, e5791.
- Schumacher, M.A., Goodman, R.H., Brennan, R.G., 2000. The structure of a CREB bZIP-somatostatin CRE complex reveals the basis for selective dimerization and divalent cation-enhanced DNA binding. *Journal of Biological Chemistry* 275, 35242–35247.

- Schultz, J., Milpetz, F., Bork, P., Ponting, C.P., 1998. SMART, a simple modular architecture research tool: Identification of signaling domains. *Proceedings of the National Academy of Sciences of the United States of America* 95, 5857–5864.
- Schwalbe, M., Ozenne, V., Bibow, S., *et al.*, 2014. Predictive atomic resolution descriptions of intrinsically disordered hTau40 and α -synuclein in solution from NMR and small angle scattering. *Structure* 22, 238–249.
- Shi, Y., 2009. Serine/threonine phosphatases: Mechanism through structure. *Cell* 139, 468–484.
- Strieker, M., Tanović, A., Marahiel, M.A., 2010. Nonribosomal peptide synthetases: Structures and dynamics. *Current Opinion in Structural Biology* 20, 234–240.
- Stroh, J.G., Loulakis, P., Lanzetti, A.J., Xie, J., 2005. LC-mass spectrometry analysis of *N*- and *C*-terminal boundary sequences of polypeptide fragments by limited proteolysis. *Journal of the American Society for Mass Spectrometry* 16, 38–45.
- Sumimoto, H., Kamakura, S., Ito, T., 2007. Structure and function of the PB1 domain, a protein interaction module conserved in animals, fungi, amoebas, and plants. *Science's STKE* 2007.re6.
- Tomba, P., 2011. Unstructural biology coming of age. *Current Opinion in Structural Biology* 21, 419–425.
- Ubersax, J.A., Ferrell, J.E., 2007. Mechanisms of specificity in protein phosphorylation. *Nature Reviews Molecular Cell Biology* 8, 530–541.
- Uversky, V.N., Dunker, A.K., 2010. Understanding protein non-folding. *Biochimica et Biophysica Acta* 1804, 1231–1264.
- Waksman, G., Kominos, D., Robertson, S.C., *et al.*, 1992. Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. *Nature* 358, 646–653.
- Wang, R., Estelle, M., 2014. Diversity and specificity: Auxin perception and signaling through the TIR1/AFB pathway. *Current Opinion in Plant Biology* 21C, 51–58.
- Wernimont, A., Edwards, A., 2009. In situ proteolysis to generate crystals for structure determination: An update. *PLoS One* 4, e5094.
- Wheeler, S.J., Marchler-Bauer, A., Bryant, S.H., 2000. Domain size distributions can predict domain boundaries. *Bioinformatics* 16, 613–618.
- Wilson, D., Pethica, R., Zhou, Y., *et al.*, 2009. SUPERFAMILY — sophisticated comparative genomics, data mining, visualization and phylogeny. *Nucleic Acids Research* 37, D380–D386.
- Wright, P.E., Dyson, H.J., 2009. Linking folding and binding. *Current Opinion in Structural Biology* 19, 31–38.
- Wüthrich, K., 1990. Protein structure determination in solution by NMR spectroscopy. *Journal of Biological Chemistry* 265, 22059–22062.
- Xu, D., Nussinov, R., 1998. Favorable domain size in proteins. *Folding & Design* 3, 11–17.
- Ye, F., Zhang, M., 2013. Structures and target recognition modes of PDZ domains: Recurring themes and emerging pictures. *Biochemical Journal* 455, 1–14.
- Zhang, Y., 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9, 40.
- Zhou, M.M., Huang, B., Olejniczak, E.T., *et al.*, 1996. Structural basis for IL-4 receptor phosphopeptide recognition by the IRS-1 PTB domain. *Nature Structural Biology* 3, 388–393.

NMR in Structural and Cell Biology

G Marius Clore, National Institutes of Health, Bethesda, MD, USA

Published by Elsevier Inc.

Introduction

Nuclear magnetic resonance (NMR) is a powerful spectroscopic technique that permits the detailed study at atomic resolution of the three-dimensional structures and dynamics of macromolecules and their complexes in solution (Wüthrich, 1986; Clore and Gronenborn, 1989; Clore and Gronenborn, 1998a; Cavanagh *et al.*, 2007). The major source of structural information resides principally in a large number of short (<6 Å), approximate interproton distance restraints derived from nuclear Overhauser enhancement (NOE) measurements. These can be supplemented by torsion angle restraints derived from three-bond scalar couplings and backbone chemical shifts, orientational restraints in the form of residual dipolar couplings (RDCs) measured in weak alignment media, and long-range distance restraints (up to 35 Å) from paramagnetic relaxation enhancement (PRE) or paramagnetic pseudocontact shift (PCS) measurements. Dynamical information can be derived over a wide range of time scales ranging from picoseconds to seconds (Cavanagh *et al.*, 2007). Techniques include relaxation spectroscopy to measure dynamics of bond vectors in the picosecond to low nanosecond regime (faster than the rotational correlation time), relaxation dispersion spectroscopy in the microsecond to millisecond regime, RDCs potentially in the picosecond to millisecond regime, *z*-exchange spectroscopy for the millisecond to sub-second regime, and real time spectroscopy for the second regime upward. Recent interest in dynamics has largely focused on the application of relaxation dispersion spectroscopy to detect and characterize the kinetic properties of sparsely populated species (Korzhnev and Kay, 2008), PRE to detect and visualize such species (Clore *et al.*, 2007), and lifetime line broadening and dark-state exchange saturation transfer (DEST) spectroscopy to probe kinetics and dynamics of interactions of NMR visible molecules with large megadalton assemblies that are invisible to NMR (Fawzi *et al.*, 2011).

Brief Historical Perspective of NMR and Solution Structure Determination of Macromolecules

The development of two-dimensional ^1H -NMR spectroscopy (Ernst *et al.*, 1987) led to the first three-dimensional solution structure determinations of small proteins in the mid-1980s (Williamson *et al.*, 1985; Clore *et al.*, 1986). Subsequent work incorporating larger numbers of interproton distance and torsion angle restraints together with stereospecific assignments led to significant increase in both precision and accuracy (Clore and Gronenborn, 1998a,b; Clore *et al.*, 1986). Chemical shift overlap, however, limited the successful application of 2D ^1H -NMR techniques to proteins less than about 100 residues (~ 10 kDa). The late 1980s and early 1990s witnessed the development of 3D and 4D heteronuclear NMR spectroscopy which extended the range of applicability

of the NMR method to significantly larger systems (Clore and Gronenborn, 1991). The first structure determination of a protein larger than 150 residues using 3D and 4D NMR was interleukin- 1β (~ 18 kDa), which at the time was $\sim 50\%$ larger than any previous NMR protein structure determination (Clore *et al.*, 1991). Over the next few years, these methods were extended to a variety of protein-peptide and protein-DNA complexes (Clore and Gronenborn, 1998a). Hybrid approaches combining existing high-resolution structures determined either by crystallography or NMR with sparse experimental NMR data, solution X-ray scattering, and computational techniques involving the application of conjoined rigid body/torsion angle-simulated annealing saw the structure determination of a number of significantly larger protein-protein complexes and proteins (up to ~ 150 kDa) (Clore and Gronenborn, 1998a; Schwieters *et al.*, 2010). At the same time, deuteration combined with transverse relaxation optimized spectroscopy further increased the molecular weight range by significantly reducing linewidths (Peruvshin *et al.*, 1997; Kay, 2005). More recently, novel techniques, based on the application of PRE, have been developed to visualize sparsely populated, highly transient species that are undetectable by conventional biophysical and structural techniques, including crystallography, conventional NMR, cryoelectron microscopy, and single molecule spectroscopies (Clore *et al.*, 2007; Iwahara and Clore, 2006; Tang *et al.*, 2006, 2007; Clore and Iwahara, 2009).

Fundamentals of NMR Structure Determination

Macromolecular structure determination by NMR is intrinsically a highly specialized, labor-intensive and time-consuming technique. In addition, for a system of any reasonable size (say greater than about 70 residues) isotopic labeling with ^{15}N and ^{13}C is required. For even larger systems, additional labeling schemes are also necessary, including site specific isotope-labeling, deuteration and methyl protonation on a ^{13}C and ^2H background. Numerous reviews have been written on the subject detailing the experimental and computational methodologies involved (Wüthrich, 1986; Clore and Gronenborn, 1989, 1991, 1998a,b; Cavanagh *et al.*, 2007).

Determining the structure of a single protein by NMR can be broken down into essentially four steps: (1) sequential resonance assignment making use of a number of experiments to identify through-bond connectivities along the backbone and side chains (usually 3D triple resonance experiments); (2) assignment of cross-peaks in nuclear Overhauser enhancement spectra (usually 3D and 4D) to obtain short (≤ 6 Å) interproton distance restraints which provide the main source of geometrical information; (3) measurement of additional NMR observables that provide useful conformational information (these may include three-bond scalar couplings that are related to torsion angles by simple empirical equations; backbone

chemical shifts which are related empirically to backbone ϕ/ψ torsion angles; long-range orientational restraints, such as RDCs measured in dilute liquid crystalline media; and (4) calculation of the three-dimensional structure from the experimental NMR restraints using simulated annealing. Generally an iterative refinement strategy is employed (Clore and Gronenborn, 1989, 1991, 1998b): calculations are initially carried out with a limited set of interproton distance restraints corresponding to NOE cross-peaks with unambiguous assignments; further interproton distance restraints from the remaining NOE cross-peaks are subsequently added in an iterative manner on the basis of a successively calculated series of structures. While improvements in spectrometer technology (e.g., the advent of cryoprobe technology that increases the signal-to-noise ratio three- to fourfold; higher field magnets that increase spectral resolution, thereby reducing spectral overlap) has reduced the measurement time to some extent, collecting all the data necessary to solve an NMR structure at high accuracy may still require several months. Similarly, improvements in both spectral analysis software (Herrmann *et al.*, 2002a; Gerstein *et al.*, 2003; Yee *et al.*, 2003) and structure calculation algorithms (Schwieters and Clore, 2001; Linge *et al.*, 2003; Herrmann *et al.*, 2002b; Kuszewski *et al.*, 2004, 2008) has permitted the introduction of some degree of automation, but extensive human intervention is still necessary to fully and reliably interpret the data in all but the simplest of cases.

Place of NMR Spectroscopy in Structural and Cell Biology

In this light, what contribution can NMR make to structural and cell biology? There are two major methods for deriving high-resolution structural information at atomic resolution: NMR spectroscopy in solution and single crystal X-ray diffraction. In rare instances, cryoelectron microscopy is also capable of providing high-resolution information in the solid state. In addition, mass spectrometry in combination with cross-linking data is potentially capable of providing low-resolution structural information when combined with the computational techniques conventionally employed to derive structures from NMR data. If crystals can be rapidly obtained, there is little doubt that crystallography, particularly with the advent of synchrotron X-ray sources, offers the fastest route to high-resolution structure determination. However, complexes are generally more difficult to crystallize than isolated proteins, and it is usually the case that weak complexes (with K_D 's in the 1–100 μ M range) are extremely difficult to co-crystallize, while very weak complexes (K_D 's > 1 mM) are virtually impossible to crystallize. In the case of NMR, complexes are amenable to structural investigation providing exchange is either fast (weak binding) or slow (tight binding) on the chemical shift time scale. If exchange, however, is intermediate on the chemical shift time scale, the signals are broadened out precluding any detailed structural work.

A full structure determination of a protein–protein complex by NMR is extremely time consuming. For example, in the case of the 40 kDa EIN · HPr complex from the bacterial phosphotransferase system, the total NMR measurement time

alone was ~ 3500 h (or 4.8 months) (Garrett *et al.*, 1999). Clearly, therefore, the conventional approach is not suitable for high throughput. Fortunately, new developments have significantly shortened the amount of time required by making full use of prior knowledge in the form of existing high-resolution crystal or NMR structures of the free proteins (Schwieters and Clore, 2001; Clore, 2000; Clore and Bewley, 2002). Measurement of RDCs can quantitatively confirm that the structures of the components within the complex are either unchanged from that in the free-state or exhibit specific regions with identifiable structural changes. Similarly, chemical shift perturbations can also be used qualitatively in this regard since small chemical shift perturbations do not entail any significant structural changes. With this information in hand it is then possible to derive high-resolution structures of complexes using limited intermolecular NOE data to provide translation (as well as orientational) information and, if measurable, RDC data (Bax *et al.*, 2001; Prestegard *et al.*, 2000) to generate very accurate orientational information. In addition, strategies based on orientational information from RDCs have been developed whereby translational information from NOE data can be entirely replaced in suitable cases by highly ambiguous intermolecular distance restraints derived from $^{15}\text{N}/^1\text{H}_\text{N}$ chemical shift perturbation mapping (Clore and Schwieters, 2003). Long-range distance restraints from a paramagnetic label attached to an engineered surface cysteine on one partner to nuclei of the other partner can also be very helpful and can replace the measurement of intermolecular NOE data completely if multiple paramagnetic labels are used (Pintacuda *et al.*, 2007).

Experimental and Computational Considerations

Given that structure determination of proteins has been extensively reviewed, we will briefly review the main experimental restraints and computational techniques used in the structure determination of complexes.

Intermolecular Distance Restraints

As noted above, the NOE is the primary source of geometric information for NMR-based structure determination (Wüthrich, 1986; Clore and Gronenborn, 1989, 1998a,b). The NOE (in the initial rate approximation) is proportional to the sixth root of the distance between two protons. The upper limit for interproton distances that can be detected using the NOE is 5–6 Å. The key to deriving intermolecular NOE-derived interproton distance restraints lies in combining various isotope (^{15}N and ^{13}C)-labeling strategies with isotope-filtering experiments that permit one to detect NOEs on protons attached to specific isotopes of nitrogen and carbon (i.e., NMR active such as ^{15}N or ^{13}C , or NMR inactive such as ^{14}N and ^{12}C) (Clore and Gronenborn, 1998a). For example, in a complex comprising one protein labeled uniformly with ^{13}C and the other at natural isotopic abundance (i.e., ^{12}C), one can selectively detect NOEs from protons attached to ^{13}C to protons attached to ^{12}C .

Paramagnetic-Based Distance Restraints

It is possible to derive intermolecular distance restraints using another NMR-based approach which involves derivatizing (one at a time) suitable surface accessible cysteines (which may have to be introduced by site-directed mutagenesis) on one protein with either a nitroxide spin label or a metal-binding site (such as EDTA) and measuring the resulting PRE or PCS on the other protein to yield long-range (15–35 Å) distance restraints (Clore *et al.*, 2007; Pintacuda *et al.*, 2007). Because in most cases the paramagnetic label is attached to the protein by a linker involving several rotatable bonds, it is essential to consider the conformational space sampled by the paramagnetic label in order to obtain accurate results (Clore and Iwahara, 2009). Two types of measurement can be made with paramagnetic labels, depending on the nature of the label. For paramagnetic labels with an isotropic g -tensor (e.g. Mn^{2+} , Gd^{3+} , nitroxide radical), pseudocontact shifts (PCSs) are not observed and PRE measurements can be carried out to determine the PRE rates for each paramagnetic- ^1H interaction from the difference in relaxation rates (usually transverse) between the paramagnetic and diamagnetic samples. When the g -tensor is anisotropic (e.g., many lanthanide ions), PCSs are observed.

Both PRE and PCS effects arise from the dipole-dipole interaction between the unpaired electron of the paramagnetic spin and a nuclear spin. Because of the large magnetic moment of the electron these effects are very large and measurable over much larger distances (up to 35 Å in suitable cases) than the dipolar interaction between two protons that gives rise to the NOE. The magnitude of the PRE is related to the sixth root of the distance between the paramagnetic center and a proton and can be directly used in refinement. The PCS has the same functional form as RDCs and is dependent on both the cube root of the distance between the paramagnetic center and the nuclei of interest and the orientation that the paramagnetic center-nucleus vector makes relative to the χ paramagnetic tensor.

In general, paramagnetic effects arising from an extrinsic paramagnetic center (as opposed to one that is intrinsic as is the case for metalloproteins) can only be applied in a rational manner if one already has a good idea of the interaction surfaces involved in complex formation. Such information can be derived rather easily by either $^{15}\text{N}/^1\text{H}_\text{N}$ chemical shift perturbation mapping (Walters *et al.*, 2001) or cross-saturation experiments (Takahashi *et al.*, 2000). The latter experiment is far more challenging experimentally since it necessitates that one of the protein is not only ^{15}N -labeled but fully deuterated as well.

Paramagnetic effects, however, do have to be used with some caution. In the case of very tightly binding complexes where exchange is slow on the paramagnetic relaxation time scale, the paramagnetic effects will arise solely from the specific complex (Clore *et al.*, 2007; Clore and Iwahara, 2009). However, when exchange is fast, the footprints from paramagnetic effects arising from minor species and configurations are apparent. This is more marked for PRE measurements than PCS ones owing to the respective sixth versus cube root distance dependencies. It is precisely these effects that have permitted states that are undetectable by conventional structural and biophysical techniques to be studied.

Other Sources of Distance Information

NMR is not the only method that can be used to derive intermolecular distance restraints. It is also possible to derive distance restraints using a combination of cross-linking, proteolytic digestion, and mass spectrometry (Bennett *et al.*, 2000; Sinz and Wang, 2001; Schulz *et al.*, 2004). In many cases, however, the data will not yield unique cross-linking partners but multiple possibilities.

Fluorescence energy transfer (FRET) through non-radiative dipolar-dipolar coupling from the fluorophore, the energy donor, to a second chromophore, the energy acceptor, scales as the sixth root of the distance between the two chromophores and can probe separations ranging from 10 to 100 Å (Hillisch *et al.*, 2001). Likewise double nitroxide spin-labeling coupled with pulsed electron paramagnetic resonance (EPR) methods such as double electron-electron resonance (DEER) based on the magnitude of the magnetic dipolar coupling of the unpaired nitroxide electrons which scales as the cube root of the separation between the two nitroxide labels, can yield remarkably accurate distances in the 20–60 Å range (Altenbach *et al.*, 2008). FRET and EPR are not limited by the molecular weight of the system being studied, but suffer from a major drawback in so far that only a single pairwise distance can be measured per sample (i.e., each distance requires a new double spin-labeled or double chromophore labeled sample, with the labels in different positions). Thus, although FRET and EPR can yield very specific information they do not afford a practical approach for solving three-dimensional structures of proteins or their complexes.

Orientational Restraints

Long-range orientational restraints can be derived from the measurement of RDCs (Bax *et al.*, 2001; Prestegard *et al.*, 2000) and chemical shift anisotropy (Wu *et al.*, 2001; Tjandra *et al.*, 1997) in liquid crystalline media, and in suitable cases from heteronuclear T_1/T_2 data (Tjandra *et al.*, 1997). The characteristic feature of these various parameters is that they yield direct geometric information on the orientation of an interatomic vector(s) with respect to an external axis system (e.g., the alignment tensor in liquid crystalline media, the diffusion tensor for relaxation measurements) expressed in terms of two angles: θ , the angle between the interatomic vector and the z axis of the tensor, and ϕ , the angle which describes the position of the projection of the interatomic vector on the x - y plane of the tensor.

For most practical purposes, RDCs provide the easiest method for deriving orientational information. In an isotropic medium, the dipolar couplings average to zero. In the solid state, the maximum value of the ^{15}N - ^1H dipolar coupling is 20.7 kHz. To effectively measure dipolar couplings in solution, therefore, it is necessary to devise means of inducing only a small ($\sim 10^{-3}$) degree of order such that the ^{15}N - ^1H dipolar couplings lie in the ± 20 Hz range. Experimentally, this is achieved by dissolving the protein or protein complex of interest in a dilute, water soluble, liquid crystalline medium. Examples of such media include lipid bicelles, filamentous phages such as fd or pf1, rod-shaped viruses such as tobacco mosaic virus, polyethylene glycol/hexanol, and stretched polyacrylamide gels.

Some Computational Methods

In many instances, protein complex formation involves no significant changes in backbone conformation. Thus, if the structures of the individual proteins are already known at high resolution and it can be shown that the backbone conformation remains essentially unchanged upon complex formation (e.g., by comparison of dipolar coupling data measured on the complex with the X-ray structures of the free proteins), one can then make use of conjoined rigid body/torsion angle dynamics to rapidly solve the structure of the complex on the basis of intermolecular NOE data and backbone NH dipolar couplings (Schwieters and Clore, 2001; Clore, 2000). In this procedure, only the interfacial side chains are allowed to alter their conformation. The backbone and non-interfacial side chains of one protein are held fixed, while those of the second protein are only allowed to rotate and translate as a rigid body. This has been applied with considerable success in the case of the 30–70 kDa protein–protein complexes of the bacterial phosphotransferase system (Clore and Venditti, 2013), as well as to the 42 kDa ternary Oct1 · Sox2 · Hoxb1-DNA ternary transcription factor complex (Williams *et al.*, 2004).

It should be emphasized that conjoined rigid body/torsion angle dynamics can readily be extended to cases where significant changes in backbone conformation are localized to specific regions of the protein, such as the binding interface. In such cases, both the interfacial side chains and the relevant portions of the protein backbone would be given torsional degrees of freedom, and the experimental data would also have to include intramolecular NMR restraints (NOE, dipolar coupling, etc.) relating to that portion of the backbone. This, for example, is the strategy that was employed to solve the structure of the IIA^{Mtl} · HPr complex (Cornilescu *et al.*, 2002). This was necessitated because the crystal structure of IIA^{Mtl} (Van Montfort *et al.*, 1998) which contains multiple copies of IIA^{Mtl} in the unit cell, revealed alternate conformations for four loops in relatively close proximity to the putative interaction surface with HPr.

Providing the complex under study can be aligned in a suitable liquid crystalline medium, the measurement of dipolar couplings is straightforward and permits one to determine the relative orientation of two proteins in a complex. Dipolar couplings, however, do not yield any translational information which is essential for docking. Clearly, NOE-derived intermolecular interproton distance restraints provide the most useful and reliable source of translational information. However, intermolecular NOEs are not always easy to observe and their unambiguous assignment is still difficult and time consuming, particularly for larger complexes. Backbone ¹H_N and ¹⁵N chemical shifts, on the other hand, are highly sensitive to environment and have been extensively used to rapidly map interaction surfaces on proteins (Walters *et al.*, 2001). Not surprisingly, examination of the NMR literature reveals hundreds of examples of chemical shift mapping studies; to date, however, only a handful of structures of macromolecular complexes have been determined by NMR. Recently, it has been shown that it is possible to convert chemical shift perturbation maps into highly ambiguous intermolecular distance restraints which, in combination with orientational restraints from dipolar couplings, can reliably

and accurately dock the partner proteins in a complex by means of rigid body/torsion angle dynamics calculations (Clore and Schwieters, 2003). Clearly, this methodology provides a powerful tool for high throughput structural proteomics and, moreover, can greatly accelerate the determination of higher accuracy NMR structures of complexes (including the detailed placement of interfacial side chains) by providing a good starting point for the assignment of intermolecular NOE data.

Structural Proteomics of the Bacterial Phosphotransferase System

In bacteria, carbohydrate transport across the membrane is mediated by the phosphoenolpyruvate:sugar phosphotransferase system (PTS) which provides tight coupling of translocation and phosphorylation (Deutscher *et al.*, 2006). The PTS is a classical example of a signal transduction pathway involving phosphoryl transfer whereby a phosphoryl group originating on phosphoenolpyruvate is transferred to the translocated carbohydrate via a series of bimolecular protein–protein complexes. The first two steps of the PTS are common to all sugars: enzyme I (EI) is autophosphorylated by phosphoenolpyruvate and subsequently donates the phosphoryl group to the histidine phosphocarrier protein HPr. The proteins downstream from HPr comprise the sugar-specific enzymes II which fall into four distinct families: glucose (Glc), mannitol (Mtl), mannose (Man) and lactose/chitobiose (Chb). Although the four families bear no sequence or structural similarity, they do possess similar organizations consisting of two cytoplasmic domains A and B, and one or two membrane-bound domains, C and D, which may or may not be covalently linked to one another. The active site residue of the A domains is always a histidine which accepts the phosphoryl group from HPr on its Nε2 atom and donates a phosphoryl group to either a cysteine residue (in the case of IIB^{Glc}, IIB^{Mtl} and IIB^{Chb}) or to the Nδ1 atom of a histidine residue (in the case of IIB^{Man}). Subsequently the phosphoryl group is transferred onto the incoming sugar on the cytoplasmic side of the membrane-bound C domain (also known as the sugar permease).

The complexes in this pathway are rather weak with *K_D*'s ranging from 1 μM to 3–6 mM. The *K_D*'s in the millimolar range relate to complexes involving isolated domains that are connected by 20–30 residue long flexible linkers in the intact protein (Clore and Venditti, 2013). Although binding in such instances is very weak, it is in fact perfectly tuned to the system. In particularly, it can readily be calculated, based on the expected average end-to-end distance for the linkers, that these millimolar equilibrium dissociation constants correspond to 50–85% probabilities of the two linked domains interacting with one another at any given time. Although high-resolution crystal structures and NMR structures have been determined for many of the individual proteins of the PTS, crystallization of these protein–protein complexes has proven to be refractory, despite many years of trying. Thus, this system provides a showcase for the impact of NMR in structural proteomics.

Figure 1 shows ribbon diagrams of structures of all nine cytoplasmic PTS complexes solved in our laboratory

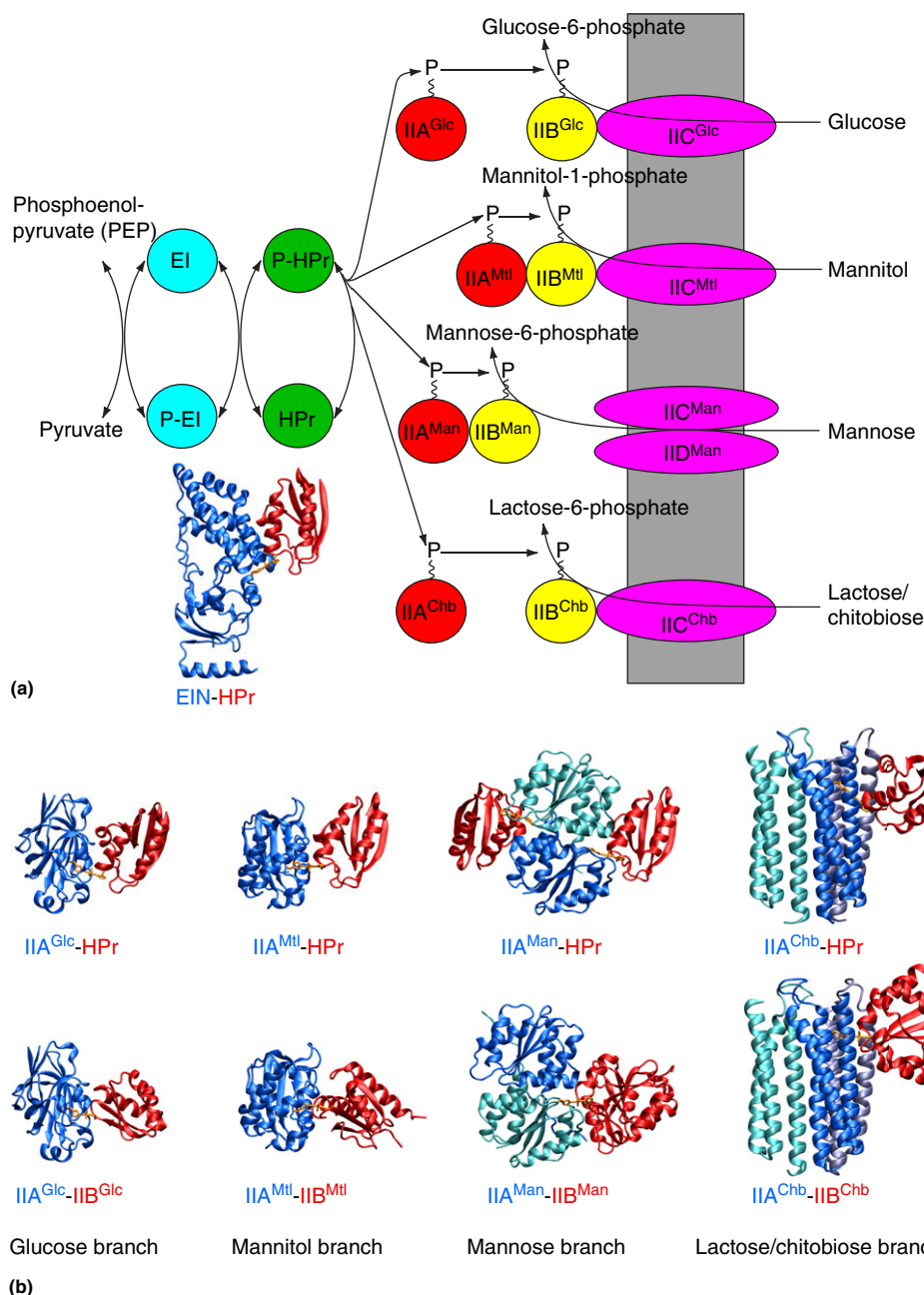


Figure 1 Structural biology of the bacterial phosphotransferase signal transduction pathway. (a) Diagrammatic representation of the pathway. The first two steps are common to all branches, and thereafter the pathway splits into four sugar-specific branches (glucose, mannitol, mannose, and lactose/chitobiose). (b) Ribbon diagrams of the nine protein-protein complexes of the *Escherichia coli* phosphotransferase system.

(Clore and Venditti, 2013). These complexes shed light on understanding fundamental aspects of protein-protein recognition, mechanisms for phosphoryl transfer between proteins, and the diversity of structural elements recognized by a single protein. Specificity of the protein-protein interaction surfaces is characterized by geometric and chemical complementarity, coupled with extensive redundancy to permit the effective recognition of multiple partners. There is little or no conformational change in the protein backbone before and after association. Some interfacial side chains, however, adopt different conformations (side chain conformational plasticity)

depending on the interacting partner so as to achieve optimal intermolecular interactions. A consequence of these properties is increased velocity in signal transduction by eliminating any unnecessary time delay required for significant conformational change.

The interaction surfaces for HPr on enzyme I and the four sugar-specific enzymes IIA are very similar despite the fact that their underlying structures are completely different in terms of linear sequence, secondary structure and topological arrangement of structural elements. HPr makes use of essentially the same surface to interact with both its upstream and

downstream partners (cf. [Figure 1](#)). Concomitantly, the binding sites for the sugar-specific enzymes IIB and HPr on the corresponding sugar-specific enzymes IIA overlap extensively. One might therefore anticipate that the enzymes IIB could also interact with EIN. However, NMR data indicate that there is absolutely no interaction between EIN and any of the sugar-specific enzymes IIB at millimolar concentrations. From a functional perspective this is important since it ensures that the PTS cascade is not bypassed. In addition, prevention of any potential shortcuts circumventing HPr and the sugar-specific enzymes IIA is also necessary since these proteins in different states of phosphorylation also regulate the functions of proteins in other pathways ([Deutscher et al., 2006](#)). The structural basis for specificity and discrimination lies in the different charge distributions on the interaction surfaces of HPr and the sugar-specific enzymes IIB such that binding of a sugar-specific enzyme IIB to EIN is precluded by electrostatic repulsion ([Clore and Venditti, 2013](#)).

Exploring Sparsely Populated States of Proteins and Their Complexes

Many biological processes proceed via sparsely populated species. Examples include the initial formation of encounter complexes in macromolecular association, target searching in specific protein–DNA recognition, conformational selection in ligand binding, conformational transitions associated with allostery, intermediates along the protein folding pathway or in the course of enzyme catalysis, and early events in self-assembly processes. In general the populations of these states at equilibrium are low and their lifetimes are short. Consequently, transient states arising from rare but rapid excursions between the global free energy minimum and higher free energy local minima are extremely challenging to study at atomic resolution under equilibrium conditions since they are effectively invisible to most structural and biophysical techniques including crystallography, conventional NMR spectroscopy, electron microscopy and single molecule spectroscopy. Recent developments in NMR have rendered short-lived, sparsely populated states accessible to spectroscopic analysis, yielding considerable insights into their kinetics, thermodynamics, and structures.

Principal NMR Methods to Probe Transient Sparsely Populated States

Three main NMR methods have been developed over the last few years to probe rare invisible states of macromolecules and their complexes at equilibrium: namely PRE, relaxation dispersion spectroscopy and lifetime line broadening coupled with DEST spectroscopy (see [Anthis and Clore \(2015\)](#) for a recent in-depth review).

The PRE requires that the distances between a paramagnetic label and the monitored spins (usually protons) are significantly shorter in the sparsely populated state than in the major species, and that the lifetime of the minor species is less than $\sim 250\text{--}500\ \mu\text{s}$ ([Iwahara and Clore, 2006](#); [Tang et al., 2006, 2007](#); [Clore and Iwahara, 2009](#)). In this exchange regime, the footprint of sparsely populated states can be observed on PRE

profiles measured on the resonances of the major species, thereby yielding structural information that is directly related to paramagnetic center–nuclei distances, from which it is possible, under suitable circumstances, to compute a structure or ensemble of structures for the minor species ([Tang et al., 2006, 2007](#)). Relaxation dispersion spectroscopy is dependent on the existence of significant chemical shift differences between the NMR active nuclei (^1H , ^{15}N or ^{13}C) in the various states, and in general can be used to probe events occurring on time scales ranging from about $50\ \mu\text{s}$ to $10\ \text{ms}$ ([Korzhnev and Kay, 2008](#); [Palmer et al., 2001](#); [Loria et al., 2008](#); [Baldwin and Kay, 2009](#)). Lastly, lifetime line broadening and DEST probe exchange dynamics at atomic resolution between NMR visible molecules and large (in excess of $1\ \text{MDa}$) NMR invisible ‘dark’ states on time scales ranging from $0.5\ \text{ms}$ to $1\ \text{s}$ ([Fawzi et al., 2011, 2014](#); [Libich et al., 2013](#)). The DEST experiment relies entirely on large differences in transverse relaxation rates between the NMR visible and invisible species. These three complementary techniques are capable of detecting states with populations as low as 0.5% .

Basis of PRE for the Study of Sparsely Populated States

The PRE yields structural information directly but cannot be used to obtain kinetic information (i.e., rate constants). The underlying theory of the PRE for static systems dates back to the late 1950s ([Solomon, 1955](#); [Bloembergen and Morgan, 1961](#)) and the PRE has long been used in the study of paramagnetic metalloproteins. The potential of the PRE for structure determination of single proteins, however, was first demonstrated in the mid-1980s ([Kosen, 1989](#)) but then largely neglected until about 10 years ago with the advent of straightforward biochemical methods for introducing paramagnetic labels at specific sites in proteins ([Battiste and Wagner, 2000](#)). Moreover, the quantitative use of the PRE for structure determination was thwarted until the introduction of the appropriate theoretical framework and computational methods to take into account the large conformational space sampled by a paramagnetic label attached to the protein via a linker with multiple rotatable bonds ([Iwahara et al., 2004](#)). By representing the paramagnetic label by an ensemble of states and taking care to calculate PRE order parameters from the coordinates during the course of structure refinement, it is possible to directly refine against the PRE relaxation rates and obtain accurate structures where agreement between the model and the experimental data is quantitatively assessed by a Q-factor analogous to a crystallographic R-factor ([Iwahara et al., 2004](#)).

The key insight into using the PRE to detect transient low-population species lies in rapid exchange phenomena whereby the transverse PRE observed on a major species is modulated by the presence of the minor species ([Iwahara and Clore, 2006](#)). In a two-site exchange system comprising two species A and B that interconvert on a time scale that is fast on the PRE time scale, the observed PRE measured on either resonance will be the population weighted average of the PRE rates for the two species. Thus, if one has, for example, a system where a particular paramagnetic center–proton distance is $30\ \text{\AA}$ for the major species and $8\ \text{\AA}$ for the minor species, the corresponding PRE rates (for a system $\sim 30\ \text{kDa}$ in size with Mn^{2+} as the

paramagnetic center) will have values of $\sim 2 \text{ s}^{-1}$ and $\sim 5600 \text{ s}^{-1}$, respectively. If the major and minor species are populated at 99% and 1%, respectively, the minor species will be invisible in the NMR spectrum. But in the fast exchange limit the observed PRE measured on the resonance of the major, NMR visible, species will be the population weighted average of the PREs for the major and minor species, in this instance $\sim 50 \text{ s}^{-1}$, much larger than that expected for the major species alone. Therefore, providing distances between the paramagnetic center and the protons of interest are significantly shorter in the minor species than the major one, and the interconversion rate between the two species is fast, the PRE profiles observed on the major species will reveal the footprint of the minor species. The PRE profiles can be analyzed quantitatively to derive structural information if the PRE profile for the major species is either known or can be calculated from a known structure (Tang *et al.*, 2006, 2007; Clore and Iwahara, 2009). As the exchange rate decreases, the influence of the minor species on the observed PRE profile for the major species will be reduced until in the slow exchange limit the PRE profile for the major species will be unaffected by the presence of the minor species. Thus, the use of the PRE to detect and characterize sparsely populated states is limited to rapidly exchanging systems, typically with lifetimes less than about 250–500 μs (Clore and Iwahara, 2009).

Basis of Relaxation Dispersion Spectroscopy

In an exchanging system between multiple states, the transverse relaxation rate (R_2) is given by the sum of the intrinsic transverse relaxation rate R_2^0 and an exchange contribution R_{ex} . The R_{ex} term is a function of the exchange rate k_{ex} (which in the case of a two-site exchange system is simply the sum of the forward and backward rate constants) and the chemical shift difference for the nucleus in the two distinct chemical environments. In the slow exchange regime, when k_{ex} is much (20-fold or more) smaller than the chemical shift difference measured in radians s^{-1} , two distinct resonances will be observed. In the fast exchange regime when k_{ex} is much (≥ 20) larger than the chemical shift difference, a single resonance will be observed at a position corresponding to the population weighted mean of the chemical shifts in the two states. In the extreme fast and slow exchange limits, the exchange contribution to the linewidth is negligible. In the intermediate regime, however, the R_{ex} term results in line broadening which is most marked when k_{ex} is equal to the chemical shift difference. Under these conditions, even the presence of a state populated at the 0.5% level can cause significant line broadening of the resonances of the major species. The key to relaxation dispersion spectroscopy lies in the use of special pulse sequences to progressively attenuate the R_{ex} contribution to the measured R_2 rate which can be achieved by applying a train of refocusing pulses while magnetization evolves under the influence of a chemical shift that varies stochastically as a result of the exchange process (Palmer *et al.*, 2001; Loria *et al.*, 2008; Baldwin and Kay, 2009). Because each nucleus follows a slightly different trajectory, dephasing of magnetization occurs resulting in larger R_{ex} rates and hence broader linewidths. By reducing the interval between the refocusing pulses (i.e. increasing the number of pulses during a fixed period T), dephasing is

decreased, the R_{ex} term is reduced and the linewidths become narrower. Plots of the observed R_2 rate as a function of the interval between the refocusing pulses yields what is known as a relaxation dispersion curve with large observed R_2 rates at low repetition rates and small R_2 rates at high repetition rates. The detailed shape of the relaxation dispersion curve is a complex function of the exchange rate k_{ex} , the populations of the states, the chemical shift difference between the states and the repetition rate. For data recorded at a single magnetic field, the contribution of population and chemical shift difference to R_{ex} cannot be separated *a priori*. Thus, unless the population of the species or the chemical shift difference between these species is already known, deconvolution of these two terms necessitates recording relaxation dispersion data at different magnetic field strengths since the species populations are independent of magnetic field while the chemical shift difference in frequency units is linearly proportional to the magnetic field. Typically relaxation dispersion experiments using refocusing pulses can probe exchange processes with lifetimes ranging from $\sim 50 \mu\text{s}$ to $\sim 10 \text{ ms}$ and occupancies for the minor species as low as 0.5–1%.

The key feature of NMR that distinguishes it from all other forms of spectroscopy is that interactions involving many sites of known identity can be probed simultaneously. The sites comprise NMR observable nuclei (^1H , ^{15}N , ^{13}C) whose resonance assignments in the major species are readily obtained using modern triple resonance NMR spectroscopy. Using global fitting procedures in which all the relaxation dispersion data at the observed sites are fitted simultaneously, it is possible to dissect kinetic pathways. The most commonly employed nucleus for relaxation dispersion experiments is the ^{15}N of the backbone amide groups. Yet, the structural information provided by ^{15}N chemical shifts alone is generally quite limited unless reference ^{15}N chemical shifts are already available for the various states being studied. More recent developments have extended relaxation dispersion measurements to all ^1H , ^{13}C and ^{15}N backbone atoms, as well as to side chain methyl groups (Sekhar and Kay, 2013). This opens the way to obtain highly reliable backbone ϕ/ψ torsion angle restraints for minor states derived from complete backbone chemical shifts, and even to compute 3D structures under suitable circumstances, thereby providing a potential avenue for obtaining full 3D structural information on sparsely populated states. More recent developments have shown that one can use relaxation dispersion experiments to obtain bond vector orientation information on minor states (Vallurupalli *et al.*, 2007). Two NMR observables are available, RDCs and residual chemical shift anisotropy (RCSA). Both involve the use of weakly aligned media (such as dilute solutions of bicelles and filamentous phages) to reintroduce anisotropic magnetic interactions that are otherwise averaged to zero in isotropic solution. Because these effects are small, highly accurate relaxation dispersion measurements are required and the experiments are extremely demanding.

Basis of Lifetime Line Broadening and DEST

Exchange dynamics between molecules free in solution and bound to the surface of a large supramolecular structure, a polymer, a membrane or solid support are important in many

phenomena in biology and material science. These interactions can be probed by lifetime line broadening and DEST spectroscopy to probe exchange dynamics at atomic resolution between NMR visible molecules and large (in excess of 1 MDa) NMR invisible 'dark' states on time scales ranging from 0.5 ms to 1 s (Fawzi *et al.*, 2010, 2011, 2014; Libich *et al.*, 2013).

Exchange line broadening can arise from either differences in chemical shifts (chemical exchange line broadening) or transverse relaxation rates (lifetime line broadening) between the free and bound states. When a molecule binds to a high (> 700 kDa) molecular weight entity, the reduced rate of molecular tumbling leads to a marked increase in transverse (R_2) relaxation rates (i.e., severe line broadening) which precludes direct observation of the bound state by standard NMR techniques. If the dissociation rate constant k_{off} is considerably smaller (by two orders of magnitude or more) than the R_2 in the bound state, the difference in R_2 values (ΔR_2) for the NMR visible species in the presence and absence of the large molecular weight species will be equal to the pseudo-first order association rate constant (Fawzi *et al.*, 2010, 2011). If k_{off} is comparable to or larger than R_2 in the bound state, ΔR_2 will be dependent upon both the association and dissociation rate constants as well as the R_2 in the bound state (Libich *et al.*, 2013).

The essence of the DEST experiment is that the large R_2 values in the bound state that preclude direct observation by NMR allow for efficient partial saturation of longitudinal magnetization of bound state resonances by a weak radio-frequency field, even at offsets where the magnetization of the free species is completely unaffected (Fawzi *et al.*, 2011). In other words, even though the bound resonances are completely broadened out beyond detection (i.e. they effectively lie in the baseline) they can be perturbed and partially saturated. Saturation of the bound resonances is then transferred back to the corresponding resonances of the free species by chemical exchange and subsequently measured as attenuation of the easily observed resonances of the NMR visible species. Operationally, the DEST experiment involves the creation of an action profile by applying weak saturation at set intervals from say + 35 kHz to - 35 kHz, and measuring the cross-peak intensities in a 2D-correlation spectrum as a function of the frequency offset of the saturation pulse. The resulting profiles are dependent upon the association and dissociation rate constants, and the transverse relaxation rates in the bound state.

Relevance to Cell Biology

The PRE has been used to study how the opposing constraints of speed and specificity are optimized in biological interactions. Examples include the first direct demonstration of intra- and intermolecular translocation of transcription factors to enhance specific site searching (Iwahara and Clore, 2006); the first experimental visualization of encounter complexes in protein-protein association (Tang *et al.*, 2006); the elucidation of conformational selection of a very low-population of correctly configured dimer in auto-processing of the HIV-1 protease precursor monomer (Tang *et al.*, 2008), a phenomenon of fundamental practical importance in the design of novel HIV protease inhibitors; and the dissection of the

complementary interplay between conformational selection and induced fit, exemplified by the characterization of a transient open-to-closed transitions in apo maltose binding protein (Tang *et al.*, 2007) and calmodulin that facilitate ligand-induced formation of the holo state (Anthis *et al.*, 2011).

Relaxation dispersion has shed fundamental insights into a range of biological problems of considerable significance (Korzhnev and Kay, 2008; Palmer *et al.*, 2001; Loria *et al.*, 2008; Baldwin and Kay, 2009; Sekhar and Kay, 2013). In the case of protein folding, this includes the structure determination of folding intermediates, the elucidation of the kinetics of their interconversion and the delineation of on- and off-pathway events. Relaxation dispersion has also been used to probe allosteric mechanisms associated with ligand binding in very large assemblies including chaperones, aspartate transcarbamoylase and the proteasome.

Lastly, lifetime line broadening and DEST have been used to probe exchange processes between monomer and protofibril-bound states of amyloid β (Fawzi *et al.*, 2010, 2011, 2014) and between intrinsically disordered proteins and the chaperonin GroEL (Libich *et al.*, 2013) on time scales of 0.5–100 ms, imprinting the residue-by-residue footprint of the NMR invisible protofibril-bound and GroEL-bound states on the easily observed monomer. The experiments on amyloid β shed light on protofibril formation, the structure and dynamics of protofibrils and exchange processes occurring on the surface and ends of protofibrils, which are of interest since the accumulation of toxic, aggregated forms of amyloid β are implicated in the etiology of Alzheimer's disease. The experiments on the interaction of intrinsically disordered proteins with GroEL unveil the complex molecular recognition process whereby GroEL recognizes a large array of sequences and structures. The demonstrated ability of the DEST technique to examine dynamics at single residue resolution of otherwise NMR invisible 'dark' states has the potential to revolutionize many areas of current interest in both biology and materials science.

Acknowledgments

This work was supported by the Intramural Program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health. This review is adapted from Clore, G.M., 2014. NMR in structural biology. In: Howard, G.C., Brown, W.E., Auer, M. (Eds.), *Imaging Life: Biological Systems from Atoms to Tissues*. Oxford: Oxford University Press, pp. 51–73.

See also: Molecular Principles, Components, Technology, and Concepts: Proteins: Diseases of Protein Folding: Huntington's Disease and Amyotrophic Lateral Sclerosis; Protein Domains: Structure, Function, and Methods

References

- Altenbach, C., Kusnetzow, A.K., Ernst, O.P., Hoffmann, K.P., Hubbell, W.L., 2008. High-resolution distance mapping in rhodopsin reveals the pattern of helix

- movement due to activation. *Proceedings of the National Academy of Sciences of the United States of America* 105, 7439–7444.
- Anthis, N.J., Clore, G.M., 2015. Visualizing transient dark states by NMR spectroscopy. *Quarterly Reviews of Biophysics* 48 (1), 35–116. doi:10.1017/S0033583514000122.
- Anthis, N.J., Doucleff, M., Clore, G.M., 2011. Transient sparsely-populated compact states of apo and calcium-loaded calmodulin probed by paramagnetic relaxation enhancement: interplay of conformational selection and induced fit. *Journal of the American Chemical Society* 133, 18974–18976.
- Baldwin, A.J., Kay, L.E., 2009. NMR spectroscopy brings invisible protein states into focus. *Nature Chemical Biology* 5, 808–814.
- Battiste, L., Wagner, G., 2000. Utilization of site-directed spin labeling and high-resolution heteronuclear nuclear magnetic resonance for global fold determination of large proteins with limited nuclear Overhauser effect data. *Biochemistry* 39, 5355–5365.
- Bax, A., Kontaxis, G., Tjandra, N., 2001. Dipolar couplings in macromolecular structure determination. *Methods in Enzymology* 339, 127–174.
- Bennett, K.L., Kussmann, M., Bjork, P., *et al.*, 2000. Chemical cross-linking with thiol-cleavable reagents combined with differential mass spectrometric peptide mapping: a novel approach to assess intermolecular protein contacts. *Protein Science* 9, 1503–1518.
- Bloembergen, N., Morgan, L.O., 1961. Proton relaxation times in paramagnetic solutions: effects of electron spin relaxation. *Journal of Chemical Physics* 34, 842–850.
- Cavanagh, J., Fairbrother, W.J., Palmer, A.G., Skelton, N.J., 2007. *Protein NMR Spectroscopy: Principles and Practice*, second ed. Amsterdam: Elsevier Academic Press.
- Clore, G.M., 2000. Accurate and rapid docking of protein-protein complexes on the basis of intermolecular nuclear Overhauser enhancement data and dipolar couplings by rigid body minimization. *Proceedings of the National Academy of Sciences of the United States of America* 97, 9021–9025.
- Clore, G.M., Bewley, C.A., 2002. Using conjoined rigid body/torsion angle simulated annealing to determine the relative orientation of covalently linked protein domains from dipolar couplings. *Journal of Magnetic Resonance* 143, 329–335.
- Clore, G.M., Gronenborn, A.M., 1989. Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. *CRC Critical Reviews in Biochemistry and Molecular Biology* 24, 479–564.
- Clore, G.M., Gronenborn, A.M., 1991. Structures of larger proteins in solution: Three- and four-dimensional heteronuclear NMR spectroscopy. *Science* 252, 1390–1399.
- Clore, G.M., Gronenborn, A.M., 1998a. Determining structures of larger proteins and protein complexes by NMR. *Trends in Biotechnology* 16, 22–34.
- Clore, G.M., Gronenborn, A.M., 1998b. New methods of structure refinement for macromolecular structure determination by NMR. *Proceedings of the National Academy of Sciences of the United States of America* 95, 5891–5898.
- Clore, G.M., Iwahara, J., 2009. Theory, practice and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chemical Reviews* 109, 4108–4139.
- Clore, G.M., Nilges, M., Sukuraman, D.K., *et al.*, 1986. The three-dimensional structure of α 1-purothionin in solution: Combined use of nuclear magnetic resonance, distance geometry and restrained molecular dynamics. *EMBO Journal* 5, 2729–2735.
- Clore, G.M., Schwieters, C.D., 2003. Docking of protein-protein complexes on the basis of highly ambiguous distance restraints derived from $^1\text{H}_\text{N}/^{15}\text{N}$ chemical shift mapping and backbone ^{15}N - ^1H residual dipolar couplings using conjoined rigid body/torsion angle dynamics. *Journal of the American Chemical Society* 125, 2902–2912.
- Clore, G.M., Tang, C., Iwahara, J., 2007. Elucidating transient macromolecular interactions using paramagnetic relaxation enhancement. *Current Opinion in Structural Biology* 17, 603–616.
- Clore, G.M., Venditti, V., 2013. Structure, dynamics and biophysics of the cytoplasmic protein-protein complexes of the bacterial phosphoenolpyruvate:sugar phosphotransferase system. *Trends in Biochemical Sciences* 38, 515–530.
- Clore, G.M., Wingfield, P.T., Gronenborn, A.M., 1991. High resolution three-dimensional structure of interleukin-1 β in solution by three- and four-dimensional nuclear magnetic resonance spectroscopy. *Biochemistry* 30, 2315–2323.
- Cornilescu, G., Lee, B.R., Cornilescu, C.C., *et al.*, 2002. Solution structure of the phosphoryl transfer complex between the cytoplasmic A domain of the mannitol transporter I $^\text{Mannitol}$ and HPr of the *Escherichia coli* phosphotransferase system. *Journal of Biological Chemistry* 277, 42289–42298.
- Deutscher, J., Francke, C., Postma, P.W., 2006. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. *Microbiology and Molecular Biology Reviews* 70, 939–1031.
- Ernst, R.R., Bodenhausen, G., Wokaun, A., 1987. *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*. Oxford: Clarendon Press.
- Fawzi, N., Ying, J., Torchia, D.A., Clore, G.M., 2010. Kinetics of amyloid β monomer to oligomer exchange by NMR relaxation. *Journal of the American Chemical Society* 132, 9948–9951.
- Fawzi, N.L., Libich, D.S., Ying, J., Tugarinov, V., Clore, G.M., 2014. Characterizing methyl-bearing side chain contacts and dynamics mediating amyloid beta protofibril interactions using $^{13}\text{C}_\text{methyl}$ -DEST and lifetime line broadening. *Angewandte Chemie* 53, 10345–10349.
- Fawzi, N.L., Ying, J., Ghirlando, R., Torchia, D.A., Clore, G.M., 2011. Atomic resolution dynamics on the surface of amyloid Ab protofibrils probed by solution NMR. *Nature* 480, 268–272.
- Garrett, D.S., Seok, Y.J., Peterkofsky, A., Gronenborn, A.M., Clore, G.M., 1999. Solution structure of the 40,000 Mr phosphoryl transfer complex between the N-terminal domain of enzyme I and HPr. *Nature Structural Biology* 6, 166–173.
- Gerstein, M., Edwards, A., Arrowsmith, C.H., Montelione, G.T., 2003. Structural genomics: current progress. *Science* 299, 1663.
- Herrmann, T., Güntert, P., Wüthrich, K., 2002a. Protein NMR structure determination with automated NOE-identification in the NOESY spectra using the new software ATNOS. *Journal of Biomolecular NMR* 24, 171–189.
- Herrmann, T., Güntert, P., Wüthrich, K., 2002b. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics program DYANA. *Journal of Biomolecular NMR* 319, 209–227.
- Hillisch, A., Lorenz, M., Diekmann, S., 2001. Recent advances in FRET: distance determination in protein-DNA complexes. *Current Opinion in Structural Biology* 11, 201–207.
- Iwahara, J., Clore, G.M., 2006. Detecting transient intermediates in macromolecular binding by paramagnetic NMR. *Nature* 440, 1227–1230.
- Iwahara, J., Schwieters, C.D., Clore, G.M., 2004. Ensemble approach for NMR structure refinement against ^1H paramagnetic relaxation enhancement data arising from a flexible paramagnetic group attached to a macromolecule. *Journal of the American Chemical Society* 126, 5879–5896.
- Kay, L.E., 2005. NMR studies of protein structure and dynamics. *Journal of Magnetic Resonance* 173, 192–207.
- Korzhnev, D.M., Kay, L.E., 2008. Probing invisible low-populated states of protein molecules by relaxation dispersion NMR spectroscopy: Application to protein folding. *Accounts of Chemical Research* 41, 442–451.
- Kosen, P.A., 1989. Spin labeling of proteins. *Methods in Enzymology* 177, 86–121.
- Kuszewski, J., Schwieters, C.D., Garrett, D.S., *et al.*, 2004. Completely automated, highly error-tolerant macromolecular structure determination from multidimensional nuclear Overhauser enhancement spectra and chemical shift assignments. *Journal of the American Chemical Society* 126, 6258–6273.
- Kuszewski, J., Thottungal, R.A., Clore, G.M., Schwieters, C.D., 2008. Automated error-tolerant macromolecular structure determination from multidimensional nuclear Overhauser enhancement spectra and chemical shift assignments: improved robustness and performance of the PASD algorithm. *Journal of Biomolecular NMR* 41, 221–239.
- Libich, D.S., Fawzi, N.L., Ying, J., Clore, G.M., 2013. Probing the transient 'dark-state' of substrate binding to GroEL by relaxation-based solution NMR. *Proceedings of the National Academy of Sciences of the United States of America* 110, 11361–11366.
- Linge, J.P., Habeck, M., Rieping, W., Nilges, M., 2003. ARIA: Automated NOE assignment and NMR structure calculation. *Bioinformatics* 19, 315–316.
- Loria, J.P., Berlow, R.B., Watt, E.D., 2008. Characterization of enzyme motions by solution NMR relaxation dispersion. *Accounts of Chemical Research* 41, 214–221.
- Palmer 3rd, A.G., Kroenke, C.D., Loria, J.P., 2001. Nuclear magnetic resonance methods for quantifying microsecond-to-millisecond motions in biological macromolecules. *Methods in Enzymology* 339, 204–238.
- Peruvshin, K., Riek, R., Wider, G., Wüthrich, K., 1997. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proceedings of the National Academy of Sciences of the United States of America* 94, 12366–12371.
- Pintacuda, G., John, M., Su, X.C., Otting, G., 2007. NMR structure determination of protein-ligand complexes by lanthanide labeling. *Accounts of Chemical Research* 40, 206–212.
- Prestegard, J.H., al-Hashimi, H.M., Tolman, J.R., 2000. NMR structures of biomolecules using field-oriented media and residual dipolar couplings. *Quarterly Reviews of Biophysics* 33, 371–424.
- Schulz, D.M., Ihling, C., Clore, G.M., Sinz, A., 2004. Mapping the topology of the calmodulin-melittin complex by chemical cross-linking and high resolution FTICR mass spectrometry. *Biochemistry* 43, 4703–4715.

- Schwieters, C.D., Clore, G.M., 2001. Internal coordinates for molecular dynamics and minimization in structure determination and refinement. *Journal of Magnetic Resonance* 152, 288–302.
- Schwieters, C.D., Suh, J.-Y., Grishaev, A., *et al.*, 2010. Solution structure of the 128 kDa Enzyme I dimer from *Escherichia coli* and its 146 kDa complex with HPr using residual dipolar couplings and small and wide angle X-ray scattering. *Journal of the American Chemical Society* 132, 13026–13045.
- Sekhar, A., Kay, L.E., 2013. NMR paves the way for atomic level descriptions of sparsely-populated, transiently formed biomolecular conformers. *Proceedings of the National Academy of Sciences of the United States of America* 110, 12867–12874.
- Sinz, A., Wang, K., 2001. Mapping protein interfaces with a fluorogenic cross-linker and mass spectrometry: Application to nebulin-calmodulin complexes. *Biochemistry* 40, 7913–7903.
- Solomon, I., 1955. Relaxation processes in a system of two spins. *Physical Review* 99, 559–565.
- Takahashi, H., Nakanishi, T., Kami, K., Arata, Y., Shimada, Y., 2000. A novel NMR method for determining the interfaces of large protein-protein complexes. *Nature Structural Biology* 7, 220–223.
- Tang, C., Iwahara, J., Clore, G.M., 2006. Visualization of transient encounter complexes in protein-protein association. *Nature* 444, 383–386.
- Tang, C., Louis, J.M., Aniana, A., Suh, J.-Y., Clore, G.M., 2008. Visualizing transient events in amino-terminal auto-processing of HIV-1 protease. *Nature* 455, 693–696.
- Tang, C., Schwieters, C.D., Clore, G.M., 2007. Open-to-closed transition in apo maltose-binding protein visualized by paramagnetic NMR. *Nature* 449, 1078–1082.
- Tjandra, N., Garrett, D.S., Gronenborn, A.M., Bax, A., Clore, G.M., 1997. Defining long-range order in NMR structure determination from the dependence of heteronuclear relaxation times on rotational diffusion anisotropy. *Nature Structural Biology* 4, 443–449.
- Vallurupalli, P., Hansen, D., Stollar, E., Meirovitch, E., Kay, L.E., 2007. Measurement of bond vector orientations in invisible excited states of proteins. *Proceedings of the National Academy of Sciences of the United States of America* 104, 18473–18477.
- Van Montfort, R.L., Pijning, T., Kalk, K.H., *et al.*, 1998. The structure of the *Escherichia coli* phosphotransferase IIA^{Mannitol} reveals two conformations of the active site. *Structure* 6, 377–388.
- Walters, K.J., Ferentz, A.E., Hare, N.J., *et al.*, 2001. Characterizing protein–protein complexes and oligomers by nuclear magnetic resonance spectroscopy. *Methods in Enzymology* 339, 238–258.
- Williams, D.C., Cai, M., Clore, G.M., 2004. Molecular basis for synergistic activation by Oct1 and Sox2 revealed from the solution structure of the 42 kDa Oct1·Sox2·Hoxb1-DNA ternary transcription factor complex. *Journal of Biological Chemistry* 279, 1449–1457.
- Williamson, M.P., Havel, T.F., Wüthrich, K., 1985. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ¹H nuclear magnetic resonance and distance geometry. *Journal of Molecular Biology* 182 (295), 315.
- Wu, Z., Tjandra, N., Bax, A., 2001. ³¹P chemical shift anisotropy as an aid in determining nucleic acid structure in liquid crystals. *Journal of the American Chemical Society* 123, 3617–3618.
- Wüthrich, K., 1986. *NMR of Proteins and Nucleic Acids*. New York, NY: John Wiley.
- Yee, A., Pardee, K., Christendat, D., *et al.*, 2003. Structural proteomics: Toward high-throughput structural biology as a tool in functional genomics. *Accounts of Chemical Research* 36, 183–189.

Folding, Misfolding, Disordered Proteins, and Related Diseases

LM Longo and M Blaber, Florida State University, Tallahassee, FL, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Amyloid Fibrous protein aggregates in which unfolded or misfolded polypeptides assemble to form an extended β -sheet structure. Amyloid accumulation is associated with a number of human pathologies.

Capture radius Greatest distance at which a protein can meaningfully interact with ('capture') its cognate ligand. Intrinsically unstructured proteins are thought to have a larger capture radius compared to structured proteins because they can occupy extended conformations.

Clathrate An ice-like water cage that surrounds solvent-exposed hydrophobic groups. Clathrate waters are less mobile than waters in bulk solvent; consequently, clathrate formation is entropically unfavorable. Their formation is enthalpically favored as their structure provides space for a hydrophobic solute while minimizing unsatisfied solvent H-bonding. See hydrophobic effect.

Coupled folding and binding Folding of an intrinsically unstructured protein into a well-defined conformation upon binding to a ligand. Coupled folding and binding is a common theme for intrinsically unstructured proteins.

Denatured State A heterogeneous ensemble of extended conformations occupied by a polypeptide; in most contexts, denatured state is synonymous with unfolded state. Compared to the native state, the denatured state is often nonfunctional, susceptible to proteolysis, and prone to aggregation.

Domain swapping The exchange of structural elements (e.g., a β -strand) between two independently folding domains. Often, domain swapping recapitulates the native 3° structure of the participating domains by swapping equivalent structural elements – but does so with nonnative connectivities. Domain swapping can be a form of protein misfolding.

Energy landscape A molecular analogy to the topographic map, the energy landscape is a plot of energy (e.g., Gibbs energy) as a function of molecular conformation. Although the energy landscape is frequently stylized as a three-dimensional surface, the folding reaction of even a simple 50 residue protein has hundreds of degrees of freedom.

Folding The conformational transition from the denatured state to the native state.

Folding frustration Formation of nonnative conformations with relatively low energy that kinetically block further productive folding.

Folding funnel The hypothesized shape of the folding energy landscape. A funneled energy landscape will efficiently guide a folding protein to the native conformation, proposed to be the energetic minimum, without needing to sample all possible conformations.

Folding mechanism The general features of folding that are shared by many proteins (e.g., nucleation followed by extension of native structure).

Folding nucleus The specific region of the polypeptide that initiates the folding reaction and templates the extension of native-like structure to adjacent sites.

Folding pathway The specific details of the folding reaction (e.g., which turn forms first). The folding pathway is sensitive to both the global architecture and the sequence of the protein under consideration.

Folding transition state The highest energy conformation, and therefore the highest energy barrier, along the folding trajectory. The energy of the folding transition state is related to the rate of folding, with higher energies resulting in slower folding processes.

Hydrophobic effect The propensity for hydrophobic substances to be excluded from aqueous solvent and coalesce (to minimize the hydrophobe–water interface). The hydrophobic effect is entropically driven by solvent properties; see clathrate.

Intrinsically unstructured proteins Proteins that do not adopt a well-defined ensemble of compact conformations.

Native state A homogeneous and limited ensemble of compact conformations adopted by a protein. The native state is often capable of specific chemical or biological function.

Roughness (of energy landscape) A measure of the number of conformational 'traps' that exist along the folding pathway.

Salt bridge An attractive interaction between positively and negatively charged amino acids comprised of both a charge interaction and hydrogen bonding. It is generally considered to be the strongest type of non-covalent interaction stabilizing protein structure.

Folding

General Principles

Proteins (polypeptides) are one of the three principle linear, unbranched biopolymers that define living systems, along with polyribonucleic acid (RNA) and polydeoxyribonucleic acid (DNA). Although all three classes of biopolymers are rich

in chemical information, it is proteins that serve as the key structural components, molecular machines, and catalysts that enable complex metabolism. Proteins are unique compared to other biopolymers because they must often 'fold' into a precise ensemble of closely related conformations, the 'native state,' to be functionally active. The structure of the native state is predominantly defined by the amino acid sequence of the protein, but can be modulated by environmental factors (e.g.,

temperature, pH, ionic strength, ligand concentration, voltage across a membrane, or mechanical stress) with functional consequences. To date, an analytical understanding of how amino acid sequence and chemical environment shape the native state structure remains elusive, though many general principles of protein structure and folding are known.

Protein Stability, A Delicate Balance between Enthalpy and Entropy

To a first approximation, the folding of a protein can be viewed as a transition between two thermodynamic states: the native state (an ensemble of closely related, compact conformations) and the 'denatured state' (a heterogeneous ensemble of extended, flexible conformations). The native state is most often associated with biological function whereas the denatured state tends to be nonfunctional, susceptible to degradation by proteases, and prone to aggregation. The overall stability of a protein – that is, the energetic difference between the native state and the denatured state – can be thought of as a 'tug of war' between various entropic and enthalpic considerations that favor one state over the other. In general, the native state is only marginally more stable than the denatured state, with an energy gap between the two states equivalent in magnitude to only a few hydrogen bonds.

The flexibility of the denatured state is a direct consequence of the molecular structure of the protein backbone, which traces the covalent bonds that 'string together' amino acids to form a polypeptide. For each peptide bond (the chemical link between amino acids) two freely-rotatable, single bonds are present in the protein backbone. In the denatured state, rotation about these single bonds is unimpeded and a vast number of conformations can be sampled. Upon folding, the conformational freedom of the backbone is constrained by the reinforcing intramolecular interactions that comprise the native state. Thus, from the perspective of the polypeptide chain, collapse into the native state structure is accompanied by a dramatic 'loss of entropy' (an unfavorable transition).

The native state and the denatured state make fundamentally different interactions with the surrounding solvent (water molecules) and, as a consequence, solvent structure must be considered to understand protein stability and folding. In the denatured state, nearly all residues – including both hydrophilic and hydrophobic side-chains – are solvent-exposed; thus, hydrogen bond donors and acceptors that reside on the protein backbone and side-chains are satisfied by water. The hydrophobic moieties of a denatured protein, however, are unable to form hydrogen bonds and, instead, induce a structural transition in the adjacent water molecules. To avoid the loss of a hydrogen bond (an unfavorable loss of enthalpy) water molecules form an ice-like cage, referred to as a 'clathrate,' around the hydrophobic group. The clathrate structure, however, is itself a trade-off by maximizing the number of hydrogen bonding partners, clathrate waters suffer a substantial loss of translational and rotational freedom (an unfavorable loss of entropy). The formation of clathrates is the basis of the 'hydrophobic effect' and is a major driving force for protein folding.

Upon folding into the native structure, the hydrophobic residues that were exposed to solvent in the denatured state are

buried into a solvent-excluded core, thereby releasing the clathrate waters. In addition, the hydrogen bonding of backbone donors and acceptors becomes largely self-satisfied by the native state structure, which serves as the basis for the commonly observed types of protein 2° structure (e.g., α -helix or β -strand). Thus, a large number of the waters that were H-bonded to polar groups in the denatured state are released when the protein adopts the native state structure. Taken together, collapse into the native state, from the perspective of the solvent, is accompanied by a dramatic 'gain of entropy' (a favorable transition).

In addition to the entropic benefit of sequestering hydrophobic residues into a solvent-excluded core, protein core formation is associated with favorable London dispersion forces, resulting in an enthalpic contribution to folding. To maximize the enthalpy of packing, protein cores often exhibit features of structural complementarity with minimal packing defects, resembling a sort of three-dimensional jigsaw puzzle. Other enthalpic contributions that favor folding result from 'salt bridges,' which are combined hydrogen bonding and electrostatic interactions between oppositely charged, hydrogen bonding-competent side-chains (e.g., lysine and glutamic acid).

Protein stability involves complex trade-offs between the configurational entropy of the protein backbone, the entropy of the solvent, and the favorable enthalpies of interaction between protein groups, such as backbone hydrogen-bonding within 2° structure, salt bridges, and London dispersion forces in the core (Figure 1). Enumeration of these terms typically identifies large opposing energetic contributions between enthalpy and entropy (e.g., ΔH and $|\Delta S|$ of folding are often $>400 \text{ kJ mol}^{-1}$ for small proteins) but with a small net balance (e.g., $20\text{--}30 \text{ kJ mol}^{-1}$) in favor of the native state structure. Thus, evolved proteins have a modest energetic preference for the folded conformation – a balance that can, in some cases, be disrupted by a single energetically unfavorable point mutation. For example, a single strong salt bridge can provide $10\text{--}15 \text{ kJ mol}^{-1}$ of favorable stability; thus, a mutation that eliminates one member of the salt bridge can cause protein unfolding (Anderson *et al.*, 1990).

Protein Folding Mechanisms and the Search for the Native State Conformation

The conformational space accessible to a protein is vast, and even a short protein of 100 amino acids can sample an astronomical number of conformations ($>10^{140}$). Thus, if a protein were to attain its correctly folded structure by sequentially sampling all of the possible conformations, it would require a time longer than the age of the universe to arrive at the correct native state – even if conformations are sampled at nanosecond or picosecond rates. Experimental studies of protein folding, in contrast, observe that the folding reaction occurs on the millisecond to second timescale, far faster than if the protein were sampling conformations at random (Levinthal, 1969). The above contradiction – first pointed out by Cyrus Levinthal and referred to as 'Levinthal's paradox' – suggests that for proteins to fold efficiently there must be a 'folding pathway' that biases the conformational search for the native state structure. Although the details of a

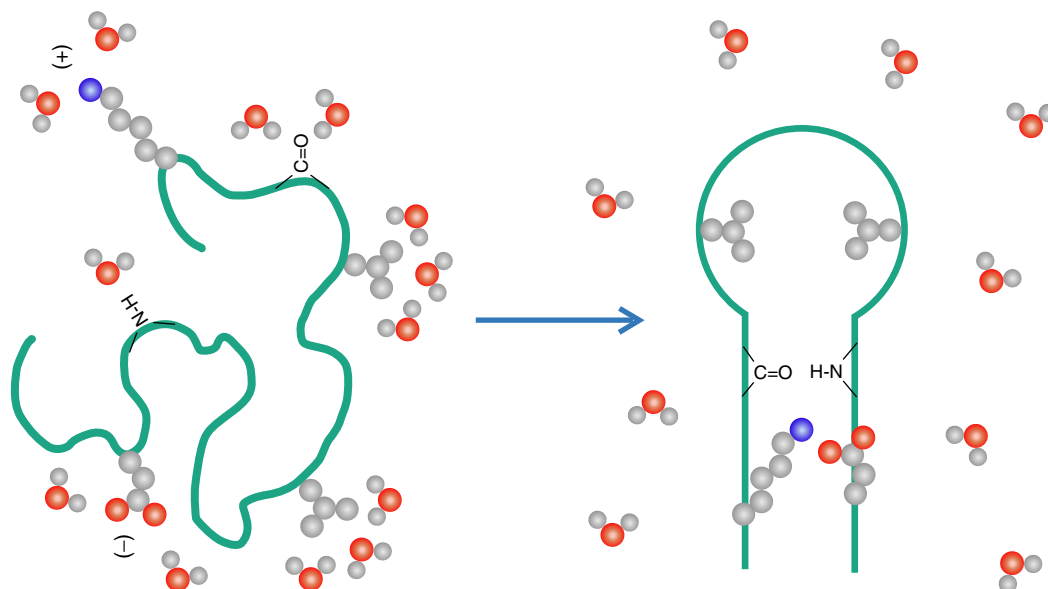


Figure 1 Basic structural and energetic principles of protein folding. Protein folding involves solvent entropy gain from the burial of hydrophobic groups (i.e., elimination of water clathrates), and enthalpy gain of favorable intra-chain charged, polar, and van der Waals interactions. The summation of these factors offset (slightly) the unfavorable protein conformational entropy penalty.

folding pathway are specific to both the amino acid sequence and native state structure of the protein, several general features have emerged leading to the belief that many proteins share a common ‘folding mechanism.’

Protein folding is nucleation-dependent, meaning that the folding reaction is initiated by a region of the polypeptide chain (the ‘folding nucleus’) that then templates the propagation of native structure to adjacent sites (Wetlaufer, 1973; Rackovsky, 1978). Early studies of protein folding noted that the common types of protein 2° structure (α -helix, β -hairpin, and β -strand) are examples of structures that, once nucleated, can be rapidly propagated to contiguous amino acids. Structurally this propensity for nucleation is due to correct positioning of the main chain amide and carbonyl hydrogen-bonding groups, which permits the addition of new partners that extend the 2° structure (Figure 2). Protein 2° structure is defined by either a repeating (α -helix, β -strand) or a specific sequential (e.g., reverse turn) backbone angle conformation, and the preference for a given type of 2° structure is heavily influenced by amino acid composition (Argos and Palau, 1982; Blaber *et al.*, 1993; Kim and Berg, 1993). Thus, ‘amino acid propensities’ for a given 2° structure can tune or control nucleation and propagation of 2° structure elements. The formation of protein 3° structure by nucleation is also feasible, and the two types of structure (2° and 3°) are free to form concurrently by establishing mutually reinforcing interactions.

A ‘molten globule’ is a term given to a polypeptide having defined 2° structure elements, a generally compact (i.e., collapsed) conformation, but no well-defined (i.e., tightly packed) hydrophobic core. It can exist in a protein whose stability has been perturbed by heat or chemical denaturants, and is also often observed as an undesirable feature of some *de novo* designed proteins.

Implicit in the terminology ‘folding pathway’ is the suggestion that protein folding proceeds via a well-ordered,

defined series of conformational transitions. Simulations of protein folding, however, challenge this view. Instead, the search for the native state likely proceeds with a high degree of variability early in the folding reaction, and becomes progressively more committed to an ordered pathway as the number of native-like interactions increases (referred to as zipping and assembly). Indeed, the progressive ordering of the folding pathway is analogous to the shape of the conformational ‘energy landscape’ upon which folding occurs, termed the ‘folding funnel’ (Leopold *et al.*, 1992; Figure 3). From the perspective of the folding funnel, native-like interactions are mutually reinforcing, and bias the subsequent conformational search toward productive pathways on the energy landscape. Because native interactions promote the formation of yet more native interactions, the need to sample huge numbers of conformations is avoided, thereby resolving Levinthal’s paradox (Levinthal, 1968; Dill and Chan, 1997). The folding funnel describes the efficient flux through a series of conformational states that become progressively less flexible (losing configurational entropy) but lower in energy.

Application of transition state theory to protein folding has permitted experimental characterization the ‘folding transition state,’ the key conformational ensemble that dictates the rate of the folding reaction (Dill, 1985). The folding transition state is the highest energy state along the pathway, and is envisioned as a saddle point along the energy surface. To date, a number of folding transition states have been characterized. Generally, the folding transition state is (1) structurally compact, (2) has limited regions of 2° and 3° structure with largely correct native interactions, and (3) shields a number of hydrophobic groups from solvent. The regions of the folding transition state that exhibit native-like structure are thought to be central to the folding process, and believed to be a key part of the folding nucleus. Loss of configurational chain entropy (discussed above) appears to be a central factor responsible for

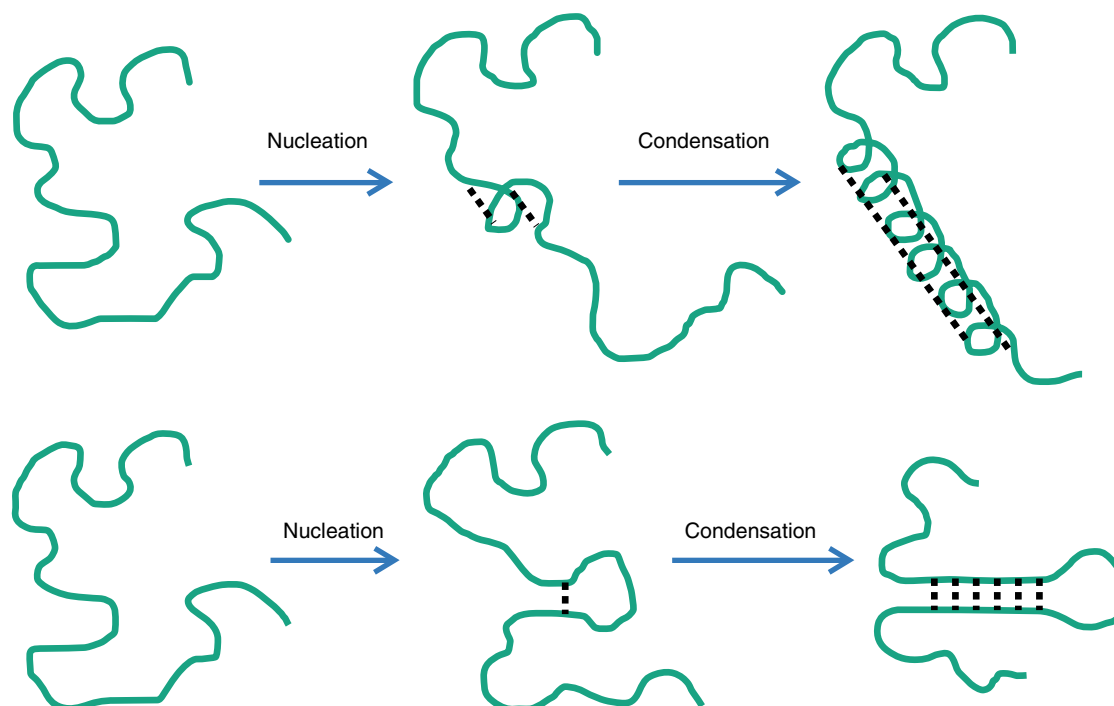


Figure 2 Example of folding nucleation by protein 2° structure. Nucleation is the stochastic rate-limiting step in which local interactions set up an initial short region of organized 2° structure (either α -helical (upper panel) or β -strand (lower panel)). This short region of 2° structure correctly positions main chain amide and carbonyl groups to template rapid subsequent extension ('condensation').

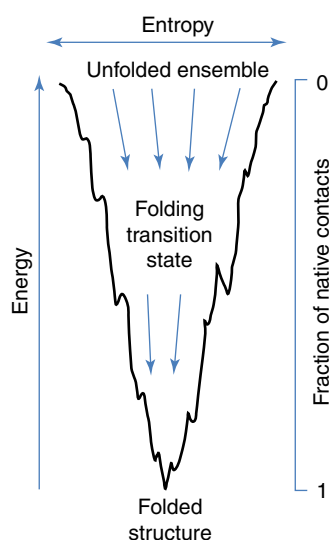


Figure 3 The energy landscape and 'folding funnel' theory of protein folding (Leopold *et al.*, 1992). The folded structure is accessible via multiple pathways, each involving a cumulative increase in native contacts, favorable enthalpic interactions, and loss of conformational entropy. A key event is formation of the folding transition state which establishes a nucleus of native structure (with limited entropic cost) able to promote formation of additional native contacts.

the height of the barrier to protein folding because measures of topological complexity (e.g., the relative closeness of interacting pairs within the sequence of a folded protein) correlate

well with the observed rates of folding across a number of different protein architectures (Plaxco *et al.*, 1998).

Misfolding

The energy landscape of folding can be described by a folding funnel (Figure 3), as discussed above. The 'roughness of the walls' of the folding funnel is a key aspect of the conformational energy landscape. The local valleys indicate 'kinetic traps' that can slow folding, including conformations that may satisfy a subset of native-like attractive interactions, but are not compatible with the final folded structure (i.e., a misfolded conformation). Such a situation is termed 'folding frustration' and may require back-tracking, or partial unfolding, to a higher energy conformation before productive folding can proceed (Nymeyer *et al.*, 1998; Capraro *et al.*, 2008). Proteins are thus hypothesized to have evolved to optimize the energy of the folded conformation as well as to avoid 1° or 3° structure features that contribute to folding frustration, to yield a minimally frustrated, smooth energy landscape (Onuchic and Wolynes, 2004).

From the earliest days of protein structure determination it was realized that many protein architectures exhibit some form of structural symmetry – postulated to be the consequence of gene duplication and fusion events in their evolution from simpler structural motifs (McLachlan, 1972; Ohno, 1970; Tang *et al.*, 1978). For example, in the extreme case of the muscle protein Titin there are 132 copies of fibronectin type III domain and 112 copies of an immunoglobulin domain. Such repeating

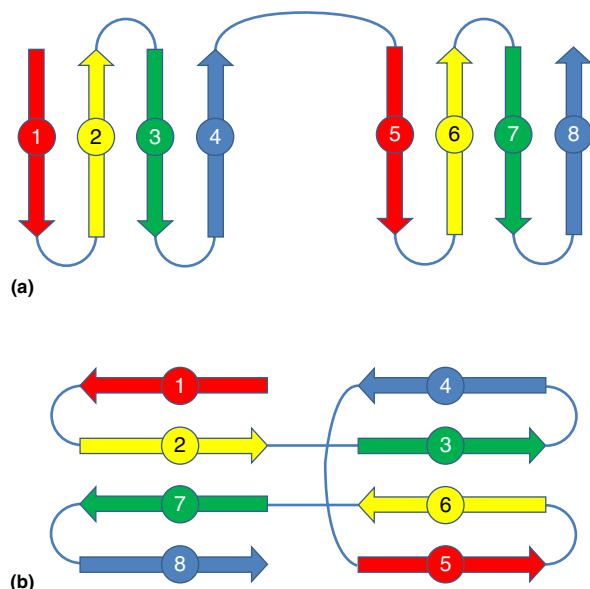


Figure 4 Domain swapping resulting from gene duplication and fusion. (a) Shows a protein architecture (e.g., a four-stranded anti-parallel β -sheet) that has undergone duplication and fusion of its coding gene. (b) Shows how exact repetition of such a motif can alternatively fold by domain swapping. This alternative folding has different connectivities (and is distinct from the native structure in panel a), but has essentially equivalent native-like contacts and is therefore likely to be kinetically trapped in its folding.

motifs offer the potential for ‘domain swapping’ during folding (Bennett *et al.*, 1994), in which repeated domains exchange structural elements resulting in significant native-like (i.e., stabilizing) interactions but with an overall incorrect structure (Figure 4). In the case of exact repeating domains (as would occur immediately subsequent to any gene duplication and fusion event) such domain swapping could potentially generate ‘native’ architecture albeit with nonnative connectivities. The folding of proteins having repetitive motifs is an area of significant interest as such studies can inform aspects of protein evolution as well as *de novo* design. Favorable local interactions, control over connecting loop length and dynamics, and sequence asymmetry between repeating motifs can each shift the folding pathway of such proteins between native and domain-swapped/misfolded forms (Hakansson *et al.*, 2001; Dyer *et al.*, 2004; Wright *et al.*, 2005). The compatibility of ‘pure symmetry’ (i.e., exact repeating 1° structure within a symmetric 3° structure) to yield an efficiently folding polypeptide has been a matter of substantial debate; however, recent successes in this area have demonstrated both the utility of symmetry in the design of efficiently folding globular proteins and the plausibility of the gene duplication and fusion hypothesis of symmetric protein evolution (Lee and Blaber, 2011; Lee *et al.*, 2011; Broom *et al.*, 2012; Voet *et al.*, 2014).

Disordered Proteins

The properties of ‘intrinsically unstructured proteins’ rest at the interface of polymer physics and structural biology. Nearly

80% of eukaryotic proteins, to varying degrees, are characterized by regions believed to lack well-defined structure (on the basis of sequence analysis). However, because the tools of molecular biology and protein chemistry are particularly well suited for the study of compact, globular proteins, an understanding of the functional significance of proteins with a high degree of conformational flexibility has lagged behind that of structured proteins. In light of recent advances, proteins are now envisioned to exist on an order–disorder continuum, with completely structured proteins at one end of the spectrum and completely unstructured proteins at the other.

Among the most common disordered structural elements is the flexible linker – a stretch of amino acids that does not adopt a defined conformation, the main function of which is to tether two (structured) protein domains together. Tethered domains are not free to move independently through space; thus, the search of a ligand or substrate molecule that must interact with both domains (either simultaneously or sequentially) is dramatically reduced. ‘Loops’ are another structural element often present within the context of a well-folded protein scaffold. Like the flexible linker, though to a lesser extent, loops are characterized by a relatively high degree of conformational disorder. Structurally, loops permit a doubling back of the protein backbone that is not possible with the linear α -helix or β -strand 2° structure elements, thereby enabling globular protein architectures to be built up from what are fundamentally linear (i.e., fibrous) structural elements. In many proteins, extended loops are hotspots for conformational changes associated with function, and are frequently observed near enzyme active sites and ligand binding sites – controlling access and specificity of substrates.

‘Coupled folding and binding’ occurs when an unstructured polypeptide binds to a ligand (e.g., RNA, DNA, or another polypeptide) and the energy of binding induces protein folding. For functional proteins that lack well-defined structure, coupled folding and binding is a common theme. In the absence of ligand, these proteins have conformational properties more akin to the denatured state of a structured protein; with ligand bound, these proteins essentially adopt a ‘native state’ conformation. Structural characterization of unliganded proteins known to undergo coupled folding and binding have revealed ‘flickering’ (i.e., transient) native-like 2° structure (Tompa, 2005). Consistent with this view, studies that modulate the amount of 2° structure flickering within an intrinsically unstructured protein have demonstrated specific functional changes (Kennedy *et al.*, 2013).

There are two principle advantages of coupled folding and binding: First, it represents an approach to encode high-specificity, low-affinity interactions (Dunker *et al.*, 2002; Uversky, 2002). In the case of structured proteins, improvements in specificity would likely come at the cost of greater affinity. Biological networks, however, often rely on transient (labile) interactions but demand remarkable specificity. Thus, by coupling folding to binding, the entropic penalty associated with structuring the polypeptide resists binding (depressing the affinity for the ligand) while the precise alignment of chemical groups (a feature of structured proteins) can still be leveraged to achieve exquisite binding specificity. Second, proteins with a high degree of structural flexibility will, by definition, be characterized by a larger radius of gyration. It has

been suggested (Shoemaker *et al.*, 2000) that a larger radius of gyration results in a greater 'capture radius,' and increases the chances that the unstructured protein will encounter a cognate ligand. The features described above – highly specific, labile interactions and an expanded capture radius – may find use at the heart of a regulatory pathway or metabolic 'hub.'

A developing interest in the field of intrinsically unstructured proteins is understanding the role that these proteins play in phase separations known as nonmembrane-bound organelles. Cellular phase separations (which often contain, but do not require, RNA) are found in a myriad of biological processes, including neuronal granules, germ granules, and nucleoli (Weber and Brangwynne, 2012). Physically, phase-separated granules behave as liquid droplets with high viscosity, are spherical in shape, and can merge upon collision. Like traditional organelles, nonmembrane-bound granules provide a mechanism to co-localize high concentrations of key molecules both spatially and temporally for biological function. It has recently been noted that unstructured, low-complexity sequences are associated with granule formation, and cellular phase separations in general, though the exact mechanisms are still poorly understood. The ability of proteins to mediate phase separations within the cell speaks to the remarkable versatility of the polypeptide.

Related Diseases

Mutations that destabilize a protein may reduce the concentration of native (and therefore functional) protein, increase susceptibility to proteolytic degradation (a process that requires partial unfolding to bind to the protease active site), and increase aggregation (due to exposure of hydrophobic groups). With sufficient depletion of active protein, metabolic dysregulation leading to pathogenic conditions can emerge. A well-known example of such a mutation is the $\Delta F508$ variant of the cystic fibrosis transmembrane conductance regulator (CFTR), in which the CFTR protein has suffered a deletion of phenylalanine at position 508. This results in substantial destabilization of the protein (Sharma, 2001) such that it does not exit the ER/Golgi quality control system. As a result, the correctly folded form of CFTR never makes it to the cell membrane at significant concentrations, thereby blocking proper chloride transport. Small molecules, whose interaction with $\Delta F508$ CFTR can enhance folding or stability, may prove useful as a protein folding-based therapy to treat this debilitating disease (Pedemonte, 2005).

Another form of misfolding disease is amyloidosis, where misfolded proteins aggregate in an ordered manner to yield insoluble, fibrous plaques that deposit in extracellular tissue. The fibrils of such plaques are the macroscopic manifestation of continuously extended β -pleated sheets, a structural feature referred to as an 'amyloid.' In 1951 Linus Pauling, Robert Corey, and Herman Branson discovered the α -helix and β -sheet 2° structure 'building blocks' of proteins. These are low-energy conformations of the peptide backbone that satisfy local backbone H-bond requirements. Once established within a small region of a polypeptide, these 2° structures can nucleate further growth (condensation). Both types of 2° structure are fundamentally linear and are terminated by different

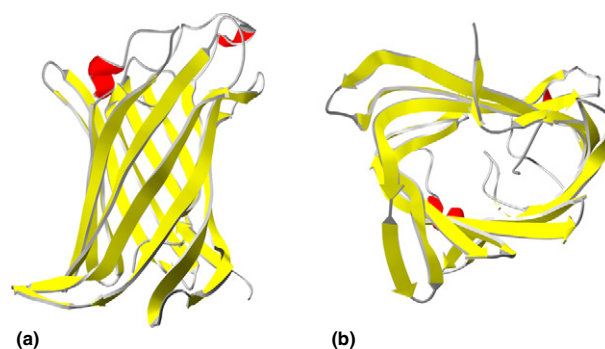


Figure 5 A β -barrel protein architecture. (a) Side view of a β -barrel (PDB accession 1QD6) where the β -strands comprising the β -sheet are oriented approximately vertically and the extension (i.e., growth) is consequently horizontal (i.e., normal to the β -strand axis); (b) Bottom view of the same structure illustrating how circularization of the sheet (i.e., barrel formation) effectively terminates the sheet extension. Adjacent sheets may be contained within the same polypeptide (as shown here) but are demarcated by reverse turn structure – thus, the β -sheet is effectively an inter- rather than intra-chain condensation growth. The structures are 'ribbon' diagrams colored by 2° structure (β -strand = yellow, α -helix = red, turn = gray).

types of 2° structure (typically reverse-turns) that are incompatible with further condensation. While the α -helix establishes local intra-chain H-bond interactions, the β -sheet can involve distal intra-chain or inter-chain H-bond interactions. In other words, an individual α -helix can extend only as far as the length of the polypeptide in which it is nucleated, whereas the β -sheet can continue to grow as long as new peptides condense onto the growing edges of the sheet. Thus, the β -sheet is much more capable of extensive, and potentially uncontrolled, growth – and it is believed that nearly any polypeptide sequence, under the right conditions, can form an extended β -sheet structure. Nature typically deals with this potential for out of control growth by forming a barrel (Figure 5) in which the β -sheet closes upon itself, neatly terminating its own growth. However, should β -sheets form peptides whose sequence disfavors curvature of the sheet, long insoluble fibrils can result (where the growing axis of the fibril is normal to the axis of the β -sheet). The gross assembly and consequent aggregation of such fibrils is associated with amyloid-based pathogenic diseases, most notably Alzheimer's disease.

One of the more fascinating amyloid-based diseases is transmissible spongiform encephalopathy (TSE), which includes bovine spongiform encephalopathy (BSE or mad cow disease) and Creutzfeldt–Jakob disease (CJD) in humans. The basis of the human disease is a misfolded form of a normal human protein found in the membrane of nerve cells called prion protein (PrP). In the native conformation, PrPc does not have a tendency to aggregate; PrPc, however, has an alternative, kinetically-trapped, low-energy conformation that favors aggregation into an amyloid structure. The amyloid form of PrP, (PrPsc) contains β -strand 2° structure that, once nucleated, can efficiently grow by the incorporation of additional PrPsc. The binding energy associated with incorporation is sufficient to drive the conformational conversion of normal PrPc into PrPsc. Thus, the infectious agent of CJD is a

misfolded protein, not nucleic acid, virus, or bacteria. This serves to highlight the fact that molecular complementarity (the basis of genetic information) can be found in molecules other than nucleic acids (indeed, yeast have been shown to use amyloid-based structures as a form of inheritance). Once the fundamental molecular complementarity of DNA bases was elucidated, interest in the potential of such properties in proteins essentially disappeared. Protein folding, misfolding, and amyloidosis, however, has prompted renewed interest in molecular complementarity within proteins.

See also: Molecular Principles, Components, Technology, and Concepts: Basic Principles: Chemical and Physical Principles. Molecular Principles, Components, Technology, and Concepts: Membranes: Cystic Fibrosis. Molecular Principles, Components, Technology, and Concepts: Proteins: Diseases of Protein Folding: Huntington's Disease and Amyotrophic Lateral Sclerosis; Protein Domains: Structure, Function, and Methods

References

- Anderson, D.E., Becktel, W.J., Dahlquist, F.W., 1990. pH-induced denaturation of proteins: A single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* 29, 2403-2408.
- Argos, P., Palau, J., 1982. Amino acid distribution in protein secondary structures. *International Journal of Peptide and Protein Research* 19, 380-393.
- Bennett, M.J., Choe, S., Eisenberg, D., 1994. Domain swapping: Entangling alliances between proteins. *Proceedings of the National Academy of Sciences of the United States of America* 91, 3127-3131.
- Blaber, M., Zhang, X.-J., Matthews, B.W., 1993. Structural basis of amino acid α helix propensity. *Science* 260, 1637-1640.
- Broom, A., Doxey, A.C., Lobsanov, Y.D., *et al.*, 2012. Modular evolution and the origins of symmetry: Reconstruction of a three-fold symmetric globular protein. *Structure* 20, 161-171.
- Capraro, D.T., Roy, M., Onuchic, J.N., Jennings, P.A., 2008. Backtracking on the folding landscape of the beta-trefoil protein interleukin-1beta? *Proceedings of the National Academy of Sciences of the United States of America* 105, 14844-14848.
- Dill, K.A., 1985. Theory for the folding and stability of globular proteins. *Biochemistry* 24, 1501-1509.
- Dill, K.A., Chan, H.S., 1997. From Levinthal to pathways to funnels. *Nature Structural Biology* 4, 10-19.
- Dunker, A.K., Brown, C.J., Lawson, J.D., Iakoucheva, L.M., Obradovic, Z., 2002. Intrinsic disorder and protein function. *Biochemistry* 41, 6573-6582.
- Dyer, R.B., Maness, S.J., Peterson, E.S., *et al.*, 2004. The mechanism of beta-hairpin formation. *Biochemistry* 43, 11560-11566.
- Hakansson, M., Svensson, A., Fast, J., Linse, S., 2001. An extended hydrophobic core induces EF-hand swapping. *Protein Science: A Publication of the Protein Society* 10, 927-933.
- Kennedy, J.A., Daughdrill, G.W., Schmidt, K.H., 2013. A transient alpha-helical molecular recognition element in the disordered N-terminus of the Sgs1 helicase is critical for chromosome stability and binding of Top3/Rmi1. *Nucleic Acids Research* 41, 10215-10227.
- Kim, C.A., Berg, J.M., 1993. Thermodynamic beta-sheet propensities measured using a zinc-finger host peptide. *Nature* 362, 267-270.
- Lee, J., Blaber, M., 2011. Experimental support for the evolution of symmetric protein architecture from a simple peptide motif. *Proceedings of the National Academy of Sciences of the United States of America* 108, 126-130.
- Lee, J., Blaber, S.I., Dubey, V.K., Blaber, M., 2011. A polypeptide "building block" for the B-trefoil fold identified by "top-down symmetric deconstruction". *Journal of Molecular Biology* 407, 744-763.
- Leopold, P.E., Montal, M., Onuchic, J.N., 1992. Protein folding funnels: A kinetic approach to the sequence-structure relationship. *Proceedings of the National Academy of Sciences of the USA* 89, 8721-8725.
- Levinthal, C., 1968. Are there pathways for protein folding? *Journal de Chimie Physique et de Physico-Chimie Biologique* 65, 44-45.
- Levinthal, C., 1969. How to fold graciously. In: DeBrunner, J.T.P., Munck, E. (Eds.), *Mossbauer Spectroscopy in Biological Systems*. Monticello, IL: Allerton House, pp. 22-24.
- McLachlan, A.D., 1972. Repeating sequences and gene duplication in proteins. *Journal of Molecular Biology* 64, 417-437.
- Nymeyer, H., Garcia, A.E., Onuchic, J.N., 1998. Folding funnels and frustration in off-lattice minimalist protein landscapes. *Proceedings of the National Academy of Sciences of the United States of America* 95, 5921-5928.
- Ohno, S., 1970. *Evolution by Gene Duplication*. New York, NY: Allen and Unwin.
- Onuchic, J.N., Wolynes, P.G., 2004. Theory of protein folding. *Current Opinion in Structural Biology* 14, 70-75.
- Pedemonte, N., 2005. Small-molecule correctors of defective DeltaF508-CFTR cellular processing identified by high-throughput screening. *Journal of Clinical Investigation* 115, 2564-2571.
- Plaxco, K.W., Simons, K.T., Baker, D., 1998. Contact order, transition state placement and the refolding rates of single domain proteins. *Journal of Molecular Biology* 277, 985-994.
- Rackovsky, S., 1978. Influence of ordered backbone structure on protein folding. A study of some simple models. *Macromolecules* 11, 1-8.
- Sharma, M., 2001. Conformational and temperature-sensitive stability defects of the delta F508 cystic fibrosis transmembrane conductance regulator in post-endoplasmic reticulum compartments. *Journal of Biological Chemistry* 276, 8942-8950.
- Shoemaker, B.A., Portman, J.J., Wolynes, P.G., 2000. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proceedings of the National Academy of Sciences of the United States of America* 97, 8868-8873.
- Tang, J., James, M.N., Hsu, I.N., Jenkins, J.A., Blundell, T.L., 1978. Structural evidence for gene duplication in the evolution of the acid proteases. *Nature* 271, 618-621.
- Tomba, P., 2005. The interplay between structure and function in intrinsically unstructured proteins. *FEBS Letters* 579, 3346-3354.
- Uversky, V.N., 2002. Natively unfolded proteins: A point where biology waits for physics. *Protein Science: A Publication of the Protein Society* 11, 739-756.
- Voet, A.R.D., Noguchi, H., Addy, C., *et al.*, 2014. Computational design of a self-assembling symmetrical β -propeller protein. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1412768111.
- Weber, S.C., Brangwynne, C.P., 2012. Getting RNA and protein in phase. *Cell* 149, 1188-1191.
- Wetlaufer, D.B., 1973. Nucleation, rapid folding, and globular intrachain regions in proteins. *Proceedings of the National Academy of Sciences of the United States of America* 70, 697-701.
- Wright, C.F., Teichmann, S.A., Clarke, J., Dobson, C.M., 2005. The importance of sequence diversity in the aggregation and evolution of proteins. *Nature* 438, 878-881.

Diseases of Protein Folding: Huntington's Disease and Amyotrophic Lateral Sclerosis

BJ Bailus and LM Ellerby, The Buck Institute for Research on Aging, Novato, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Huntington's disease (HD) and amyotrophic lateral sclerosis (ALS) are neurodegenerative diseases, which affect two different regions of the brain. Symptoms of HD include chorea, mental disturbances, mood changes, and dementia and are due to degeneration of cells in the striatum and cortex; while those with ALS exhibit difficulty moving limbs, muscle wasting, decline in cognitive function, and eventual respiratory arrest and are due to degeneration of upper and lower motor neurons of the cerebral cortex, brainstem, and spinal cord (Labbadia and Morimoto, 2013; Garbuzova-Davis *et al.*, 2011). Although there are no currently available treatments there is a large scientific community aggressively pursuing therapy options. Both HD and ALS are generally caused by a gain-of-function mutation in their respective genes that results in protein aggregation leading to disruption of protein homeostasis, altered ubiquitin-proteasome system (UPS) and cellular autophagy, eventually resulting in cellular death (Kiernan *et al.*, 2011; Ghavami *et al.*, 2014; Walker, 2007). The cellular dysfunction or death of the affected neuronal cells is the cause of the physiological symptoms. We will cover the genetic basis of each disease, the toxic mutant proteins, the

formation of the aggregates, the effect on protein homeostasis, and cellular autophagy. The theories on cellular transfer of aggregates and why neurons are specifically vulnerable to these toxic protein conformations will be explored. The conclusion will focus on the current trends in the field toward potential therapies and cures in the context of addressing protein homeostasis and autophagy.

Protein homeostasis, termed proteostasis, is defined as the correct maintenance of proteome in the cell, and the cellular activities associated with this, including folding, assembly, transport, repair, and degradation of proteins (Kikis *et al.*, 2010). It is imperative that damaged proteins are broken down and that the correct cellular balance of proteins is maintained for the cell to operate in an effective manner. A change in a protein conformation can alter the cellular function, concentration, and degradation of the mutated protein, impacting the cell's ability to maintain the correct protein homeostasis (see Figure 1). In a normal cell protein homeostasis is maintained through two main pathways, the ubiquitin dependent pathway and autophagy. The ubiquitin dependent pathway relies on a series of ubiquitin carrying proteins that transfer ubiquitin onto the target protein, usually tagging the protein for degradation by the proteasome. This pathway predominantly

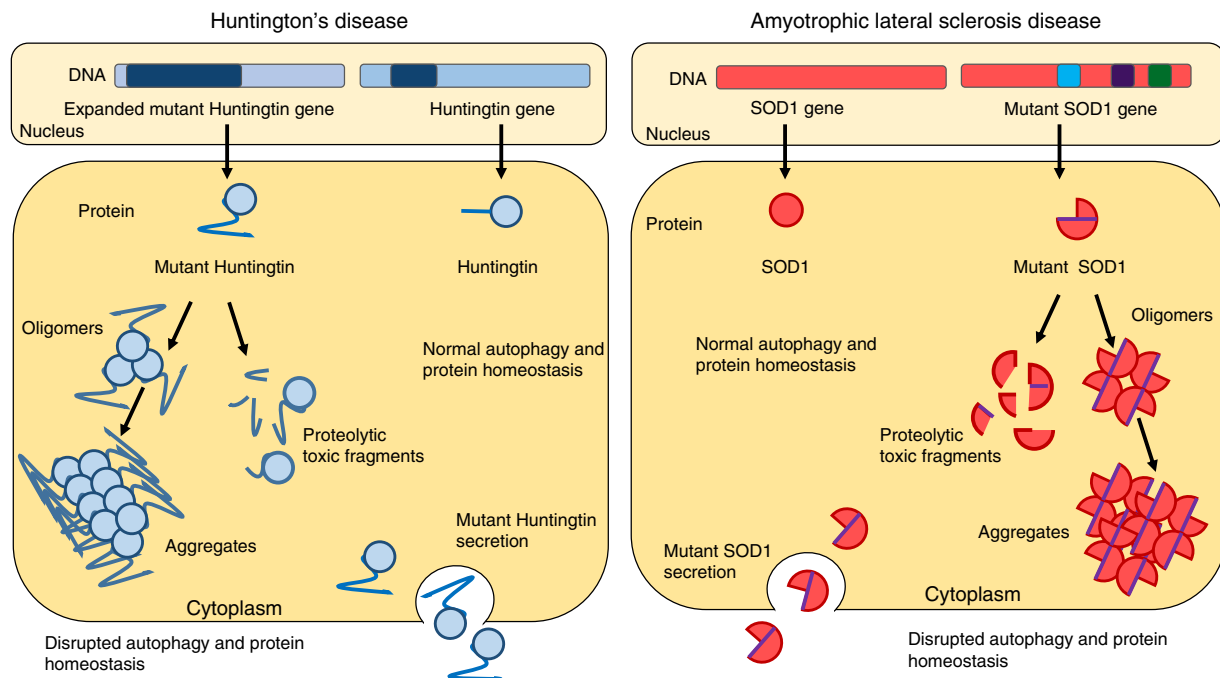


Figure 1 A depiction of HD and ALS at a genetic and protein level. The mutant DNA is shown with either the expanded CAG region for HD, or a variety of possible mutations in the SOD1 gene for ALS. The DNA codes for a mutant version of the respective proteins HTT and SOD1, which have a tendency to form oligomers and then aggregates. The mutant proteins have unique proteolytic sites which result in toxic fragments. The cells preferentially secrete the mutant protein into the extracellular matrix, which are then taken up by neighboring cells, propagating the spread of the disease.

focuses on short-lived proteins. The other pathway, autophagy degrades longer lived proteins and cellular organelles utilizing the lysosome pathway (Ghavami *et al.*, 2014; Lin *et al.*, 2013; Cuervo *et al.*, 2005). Working synergistically with these pathways is a group of cleavage enzymes that help direct which pathway the protein are broken down through (Wellington *et al.*, 2000, 2002; Gafni and Ellerby, 2002).

Genetic Basis of HD and ALS

Huntington and ALS are diseases caused by autosomal dominant mutations that result in a toxic gain of function for their respective proteins. In autosomal dominant diseases only one allele needs to be inherited for the individual to be affected. The huntingtin gene has a series of CAG repeats in exon 1, normal repeat length is less than 36 repeats, and the disease state is more than 40 repeats (Walker, 2007; Mangiarini *et al.*, 1996). The CAG trinucleotide repeats code for the amino acid glutamine. This increased polyglutamine tract results in a toxic gain of function for the huntingtin protein (HTT), causing a conformational change leading to toxic fragment production, accumulation of oligomers, fibrils, and inclusions. The CAG repeat is genetically stable when the repeat length is less than 36, but as the length increases the genetic stability decreases, with a tendency toward increased repeat length, due to slippage during DNA replication (Petruska *et al.*, 1998). This slippage during DNA replication results in an increase of CAG repeat length for each successive generation of offspring. The length of the CAG repeat region correlates to the onset of the disease, the more repeats the earlier the onset of symptoms occurs (Andrew *et al.*, 1993; Krawczak *et al.*, 1991; Snell *et al.*, 1993). Since 1993, a genetic test has been available for individuals, who chose to utilize it (de Die-Smulders *et al.*, 2013; Myers, 2004). Along with HD, there are 16 different diseases caused by trinucleotide expansion triplet repeat expansions, collectively called trinucleotide repeat expansion disorders.

The function of the HTT (350 kDa) is not fully elucidated but the protein is composed of so called HEAT repeats and acts as a scaffold protein forming a large number of protein:protein interactions that give clues to its function (Tourette *et al.*, 2014). The HTT protein is found associated with organelles, is ubiquitously expressed and knockout of this gene is embryonic lethal, suggesting a fundamental cellular function for HTT. The HEAT repeat domains of HTT are predicted to act in molecular motion (Grinthal *et al.*, 2010) and are involved in vesicular trafficking, secretory vesicle fusion, autophagy, and cilia function (Pla *et al.*, 2014).

In ALS about 90% of the cases are sporadic, in which there is no previously known familial cause of the disease (Hardiman *et al.*, 2011). There are many causes potentially implicated in the disease including but not limited to head trauma, diet, and environment. The remaining 10% of the cases are genetic; over 40 different genes have been implicated in the disease. One of the most common mutations is a mutation in the SOD1 gene. Mutations in the SOD1 gene account for 2% of all ALS cases (Robberecht and Philips, 2013; Andersen and Al-Chalabi, 2011). Over 100 different mutations in the SOD1 gene have been characterized and implicated in ALS, the specific mutation will influence the manifestation and onset of symptoms

(Robberecht and Philips, 2013). Genetic testing is available for individuals, for the common familial associated mutations.

The SOD1 gene codes for a copper and zinc-containing enzyme called superoxide dismutase that protects against free radical damage detoxification of superoxide. The protein is ubiquitously expressed in the body. The mutations in SOD1 can result a conformational change in the protein that results in a gain of function, and protein aggregation (Robberecht and Philips, 2013). The deficiency in copper in the various genetic forms may be a factor that leads to the propensity of these disease proteins to form fibrillar SOD1 aggregates (Pratt *et al.*, 2014). Metal homeostasis may in general be altered in all forms of ALS (Lovejoy and Guillemin, 2014) and lead to altered protein homeostasis and aggregation. Both HD and ALS belong to a group of neurodegenerative diseases including Alzheimer's disease and Parkinson's disease called proteinopathies due to the propensity of the proteins to form aggregates.

Effects of Protein Products on Protein Homeostasis

In both HD and ALS the genetic mutations result in a gain of function, with the mutated proteins forming protein inclusion bodies. These protein aggregates accumulate in the cells, and have been implicated in affecting protein homeostasis, both as a detriment and as a compensatory mechanism for maintaining protein homeostasis. In HD the predominant species of protein found in the aggregates is HTT, in ALS the predominant species is SOD1. The inclusion bodies have ubiquitin and proteasome components. Although originally thought to be the causative agent of the disease, the aggregates formed from these mutated proteins are now considered to be the cells compensatory mechanism, as a way to maintain a modified protein homeostasis (Arrasate *et al.*, 2004). The formation of these aggregates into inclusion bodies has been shown to positively predict cellular survival, and decreases the amount of diffuse toxic huntingtin present in the rest of the neuronal cell (Arrasate *et al.*, 2004). In ALS the protein aggregates are composed of a mixture of SOD1, TDP-43, and other proteins (Kiernan *et al.*, 2011).

Cellular toxicity in both ALS and HD is considered to be a result of the smaller fragments that are created from the disease protein, and proteolytic events (Ellerby and Orr, 2006; Gafni and Ellerby, 2002; Gafni *et al.*, 2012; Hermel *et al.*, 2004). These fragments aggregate more readily in the cells, and their localization within the cell may determine their neurotoxicity on both HD and ALS (Arrasate *et al.*, 2004). The fragments are generated when the expanded or mutated transcript results in alternative proteolytic sites on the protein, resulting in the generation of smaller fragments. These smaller fragments appear to be less prone toward aggregation, the cells coping mechanism for larger fragments. The smaller fragments diffuse throughout the cell, evading the normal degradation pathways, disrupting cellular processes, and protein homeostasis (Warby *et al.*, 2008; Labbadia and Morimoto, 2013).

Clearance Pathways of Aggregates – Cellular UPS

One pathway that helps maintain proteostasis by protein degradation is the UPS. It is utilized by cells to reduce the

number of damaged proteins and represents a major mechanism of defense for post mitotic neurons. In UPS system operates by (1) targeting the protein for degradation by modifying its C-terminus with ubiquitin and (2) then subsequent substrate proteolysis in the proteasome. In HD and ALS there are numerous studies evaluating how the UPS is involved in the disease process. It is known that specific proteasomal inhibitors when added to cells, results in increased aggregation of the mutant HTT or SOD1 protein which are commonly referred to as inclusion bodies. In cells and post-mortem human disease tissue inclusion bodies include ubiquitin and proteasomal subunits (Katsuno *et al.*, 2014). It is hypothesized that the UPS system may be impaired at the targeting or proteolysis stage. Numerous experiments have tested whether the proteasomal activity is inhibited by inclusion bodies, but this does not appear to be the case. Studies have looked at whether inclusion bodies or conformationally altered proteins obstructed the channel of the 20S core subunit by preventing ubiquitinated substrates from entering the channel; this does not appear to be the case as the UPS is able to degrade polyQ expanded HTT and inclusions (Hoffner and Djan, 2014). The recruitment of the proteasomal is dynamic and cell biology studies show it is reversible with respect to inclusion bodies. Further studies suggest that it is not the inclusion bodies but the fibrillar or possibly oligomeric forms that interfere with the 26S proteasome. These mechanisms suggest a direct interaction of mutant HTT with the UPS system but additional studies are still needed for confirmation. Another possibility that the field is exploring is that the global proteostasis network could be perturbed by aggregated disease proteins and this would delay the maturation of cellular chaperone clients, increasing the overall number of ubiquitinated substrates, disrupting the UPS. A recent example is the demonstration that aggregation prone proteins such as mutant HTT and superoxide dismutase effect a fundamental cellular process-clathrin-mediated endocytosis through the heat shock protein 70. Clathrin-mediated endocytosis is required for AMPA receptor internalization and dysregulation could lead to neurodegeneration (Yu *et al.*, 2014).

Clearance Pathways of Aggregates – Autophagy Pathway

The second clearance pathway used to maintain proteostasis is known as autophagy, a process by which proteins and organelles (mitochondria, endoplasmic reticulum, and ribosomes) are sequestered into autophagosomal vesicles and delivered to the lysosome for degradation. The details of this pathway were worked out in yeast and there are four major steps involved. The genes Atg1p-Atg13p-Atg17p are involved in the first steps of autophagosome formation and this is regulated by nutrient signaling protein, mTOR. The Vsp34p-Vsp30p facilitates the assembly of the autophagosome via phosphoinositide signaling. The ubiquitin-like protein (UBL) conjugation cascade (Atg7p, Atg10p, Atg3p, Atg8p, and Atg12p) is used for autophagosome maturation and cargo recruitment. The last step involves the recycling system, Atg9p, Atg2p, Atg18p, and Atg21p to transfer and recycle components from the isolated membrane. In HD and ALS, a number of

these steps have been shown to be disrupted by the mutated protein and effects the degradation of aggregated proteins. In both HD and ALS there is increase number of autophagosomes and inhibition of mTOR. Mutant HTT has been shown to interfere with the ability of the autophagic vacuole in cargo recruitment (Koga *et al.*, 2011). In addition, a specialized form of autophagy called chaperone-mediated autophagy is increased in HD and appears to be a compensatory mechanism for clearing proteins; however, even with the upregulation of different forms of autophagy the mutant HTT is not cleared. Generally, it is understood that mild activation of autophagy can provide benefit in clearing the disease containing proteins but over activation leads to cellular dysfunction and ultimately cell death. Another pathway that appears to be important in clearing mutant HTT is through the scaffolding protein called ALFY which clears aggregated proteins via the lysosome and the levels of the ALFY protein are decreased in the HD striatum (Filimonenko *et al.*, 2007, 2010). The HTT has recently been proposed to function as an Atg11 and AT23 homolog and control aspects of cellular autophagy (Ochaba *et al.*, 2014). This data would suggest HTT has a critical protein for cellular pathways mediating clearance of proteins or organelles.

Cellular Transfer of Aggregates on Toxic Fragments

In HD it has been shown that both the aggregates and the fragments can be propagated both within the cell and by cell-to-cell transmission. Transmission is the ability of HTT fragments that are expressed in one cell type to spontaneously transfer to an adjacent cell. The transmission of mutant HTT has been demonstrated in mouse models of HD, in HD patients grafted with fetal tissue, in HD patients and in HD stem cells implanted into mouse brains. Further, when healthy human neurons were transplanted into HD mouse models, these neurons began to accumulate mutant HTT and display altered neuronal morphology. The spread of aggregates could have implications on how the disease progresses during aging and which regions of the brain are most vulnerable to this transmission mechanism. The exact cellular mechanics for this type of transfer have not yet been elucidated but appears to require the synaptic vesicle fusion machinery or cell surface proteins (Trevino *et al.*, 2012). For Tau, which exhibits similar aggregation and fragment tendencies when mutated in Alzheimer's disease, the transmission mechanism requires heparin sulfate proteoglycans (Holmes *et al.*, 2013). Research is currently progressing to better understand the precise mechanism used for HTT transmission.

One of the hallmarks of ALS is the progression of the disease, the outward spread of muscle atrophy from a central starting point (Hardiman *et al.*, 2011; Robberecht and Philips, 2013). It has been hypothesized that the propagation of symptoms is due to the propagation of the mutant SOD1 aggregates, both within an individual cell and from one cell to another. The intracellular concept involves a prion-like transmission in which a mutant form of the SOD1 will upon contact with a normal form of SOD1 influence the normal SOD1 into changing its confirmation and folding into a mutant-like confirmation (Munch and Bertolotti, 2011). This is a seeding effect in which one mutant protein will affect many

normal proteins, which in turn will change their conformation and affect other normal proteins (Fernandez-Borges *et al.*, 2013; Munch and Bertolotti, 2011; Munch *et al.*, 2011). This cascade of events eventually proves toxic for the cell, however this only effects a single cell. A broader question examines if a cell with mutant aggregates can transmit these mutant fragments to a neighboring cell and cause the same cascade event. Several studies have shown that in cell culture the secretion of mutant fragments and aggregates occurs, the neighboring cells internalize these mutant fragments through endocytosis, which then act as 'seeds' converting normal folded proteins into misfolded proteins (Munch and Bertolotti, 2011). The mechanism for SOD1 secretion is specific for mutant SOD1, which exhibits a gain of function by interacting with chromogranins A and B which lead to its secretion, and eventual transmission to neighboring cells (Urushitani *et al.*, 2006). This progression from cell to cell helps explain the progression of symptoms outward from a local starting point in ALS.

Cell Specificity and Susceptibility

In HD and ALS, the particular vulnerability of neurons to altered proteostasis could be the cause of the selective neurodegeneration observed in these diseases. Unlike neurons, which do not divide, dividing cells can dilute the misfolded protein through uneven division and the younger cells eventually no longer carry the misfolded or aggregated proteins. Neurons and potentially other specific cell types in the brain are particularly vulnerable to these toxic proteins because of their different ability to employ UPS, autophagy, or the chaperone network. For example, the activity of UPS is lower in the striatal cells when compared to cortical cells. This would lead to a reduced ubiquitination and clearance of the mutant HTT in the striatum, causing cellular dysfunction and death. The protein chaperone network, which assists in folding proteins, has distinct expression levels in different regions of the brain (Margulis and Finkbeiner, 2014). By analogy similar changes in the proteostasis network may give clues to the selective neurodegeneration found in ALS.

Symptoms and Treatments of the Disease

The physical manifestations of both HD and ALS on an organismal level are profound, both diseases exhibit massive neuronal death, enough to be distinguished on autopsies of the deceased patients (Figure 2; Osborn *et al.*, 2004). This massive neuronal death disrupts the brain affecting the patient's ability to perform movements and in many cases cognitive functions. In HD the region with the highest incidence of neuronal death is the striatum, particularly the medium spiny neurons. Medium spiny neurons are GABAergic neurons, which have an inhibitory effect on the neurons they are in contact with, these neurons are essential in the coordinating movement (Tepper *et al.*, 2008; Walker, 2007). The progressive neurodegeneration leads to in chorea, a movement disorder associated with HD, in which the patient exhibits spastic, involuntary movements. Tetrabenazine, is currently one of USDA approved drug for treating chorea, it works by

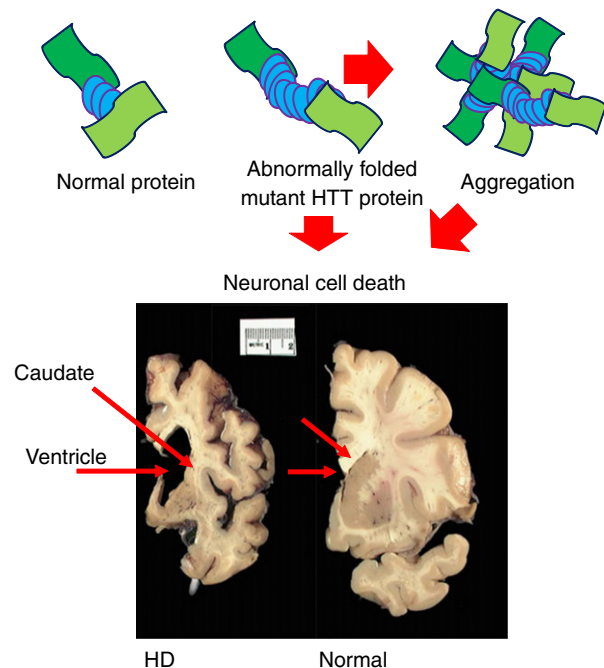


Figure 2 Huntington's disease protein, aggregation and pathology. The upper panel is a comparison of HTT protein with a normal CAG repeat size compared to a mutant HTT with an expanded CAG repeat size and the aggregation. Macroscopic image in which a slice of Huntington's brain (left) is put next to a slice from a normal control (right). Note that the striatum (the caudate nucleus, putamen, and nucleus accumbens), on the left, is severely atrophied in the Huntington's brain. Note also that the cerebral cortex is atrophied in the Huntington's brain. Photo courtesy of the Harvard Brain Tissue Resource Center.

inhibiting dopamine receptors (Zheng *et al.*, 2006; Huntington Study, 2006; Poon *et al.*, 2010). Beyond the physical symptoms, HD patients experience cognitive decline, depressive episodes, mood swings, and obsessive behavior (Montoya *et al.*, 2006; van Duijn *et al.*, 2007). The behavioral and cognitive changes increase as the disease progresses, treatment involves the standard psychiatric medications. The mutant HTT is expressed throughout the body tissues, leading to muscle atrophy, cardiac failure, impaired glucose tolerance, osteoporosis, and weight loss (van der Burg *et al.*, 2009). These symptoms can be treated with a combination of physical therapy, diet changes, and medications. There are no HD specific drugs or treatments and there is no cure, the eventual outcome of the disease is death.

ALS progresses at a much more rapid rate than HD, usually resulting in death within 3–5 years of the first symptoms (Robberecht and Philips, 2013). In ALS the patient first notices difficulties moving an extremity, this progresses toward muscle atrophy, whole body movement difficulty, and eventually respiratory failure (Kiernan *et al.*, 2011). These symptoms are due to axonal retraction, death of the neuron, and denervation of the muscles, resulting in paralysis (de Carvalho *et al.*, 2010). In a subpopulation of patients the prefrontal and temporal cortex are affected to a degree which results in dementia (Phukan *et al.*, 2007). A subsection of the patient population

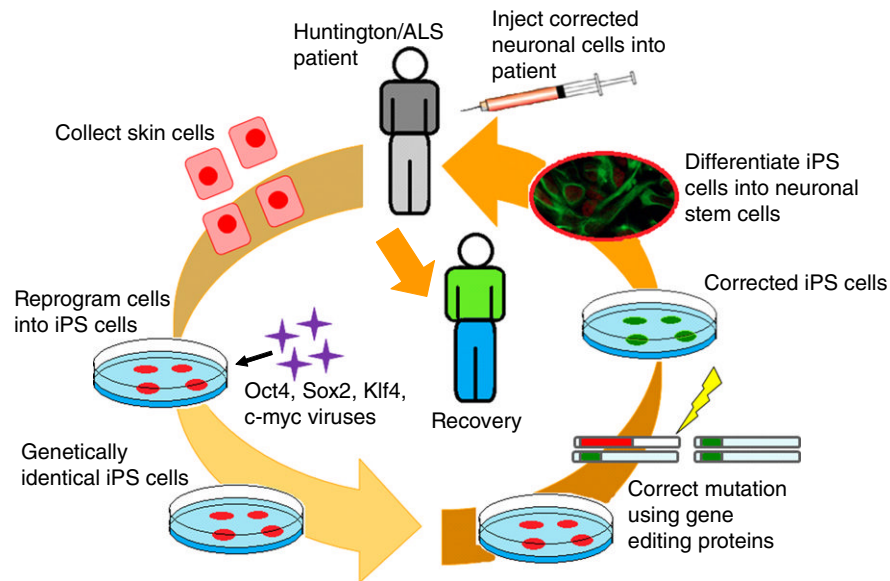


Figure 3 Use of induced pluripotent stem cells for therapeutics. Summarized are the steps for the use of induced pluripotent stem cells for autologous stem cell therapy in genetic diseases. (1) First skin cells are collected from patient; (2) then the cells are reprogrammed into pluripotent stem cells; (3) the stem cells are genetically corrected to remove the disease mutation; (4) the induced pluripotent stem cell is converted into the relevant cell type needed in the disease; and (5) finally the cells are transplanted into the patient.

does not experience any cognitive decline and is fully aware of their slow paralysis. The progression and severity of symptoms is dependent on both the type of genetic mutation and the age at onset of symptoms, creating a very heterogeneous population of patients in symptoms and survival time (Kiernan *et al.*, 2011; Hardiman *et al.*, 2011). No cure exists and medications are mostly focused on reducing symptom severity, Riluzole has shown some promise in increasing survival time, if given in the early stages of the disease (Carlesi *et al.*, 2011).

Potential Treatments and Therapies

There are three main treatment strategies being pursued for HD and ALS, drugs, stem cell therapies and genetic manipulation. The most conventional is drug screening for small molecules that will either delay the progression of the disease or ameliorate the symptoms of the disease. Several experimental drugs have been tried including Rapamycin in HD and Minocycline in ALS and HD (Ravikumar *et al.*, 2006; Berger *et al.*, 2006; Van Den Bosch *et al.*, 2002; Plane *et al.*, 2010). These drugs have side effects that have made them undesirable for long-term treatment. Rapamycin inhibits the TORC1 complex and has become a drug of interest for HD with the elucidation that the TORC1 complex of the mTOR pathway has been shown to interact with mutant HTT (Pryor *et al.*, 2014). Unfortunately Rapamycin has side effects including diabetic-like symptoms and lung toxicity (Lamming *et al.*, 2012; Pham *et al.*, 2004; Chhajed *et al.*, 2006). Minocycline, a candidate drug for both HD and ALS was found to be ineffective in HD and exacerbated the progress of ALS, perhaps through increased immunosuppression (Gordon *et al.*, 2007; Plane *et al.*, 2010; Huntington Study Group, 2010). Drug

screening for small molecules continues and there are several ongoing Phase II clinical trials for ALS and HD (Shefner *et al.*, 2013). As more information is discovered about HD and ALS more potential targets are revealed. The potential for stem cell therapies is being explored in both HD and ALS, in HD it has been shown in mice that injections of stem cells in the brain can have a beneficial effect on symptoms, for ALS there are ongoing human trials to evaluate the efficacy of this treatment in muscles (An *et al.*, 2012; Sadan *et al.*, 2009; Cohen, 2013). The main challenge for these experiments is the successful genetic modification of the stem cells and the delivery of them to the appropriate local within the patient (Figure 3). The third strategy involving genetic modification has advanced dramatically in the last 5 years as newer and more reliable tools have emerged in the field. Several group of DNA modification enzymes (Zinc Fingers, TALEs, and CRISPR) have the ability to replace the mutant gene with a modified nonmutant copy of the gene, or to regulate transcription levels (An *et al.*, 2012, 2014; Bailus and Segal, 2014; Segal and Meckler, 2013). The major challenge of this technology is efficiency and deliverability. Gene modifying tools are currently being utilized to modify stem cells, which are then injected into the patient in an effort to replace the damaged and dead cells with new nonmutant protein coding cells (An *et al.*, 2014; Hou *et al.*, 2013; Horii *et al.*, 2013; Yin *et al.*, 2014; Ye *et al.*, 2014). As more is understood about HD and ALS better therapies can be designed and delivered to patients in an effort toward increasing life span and survival rate in patients.

See also: Molecular Principles, Components, Technology, and Concepts: Proteins: Folding, Misfolding, Disordered

Proteins, and Related Diseases; Protein Domains: Structure, Function, and Methods

References

- An, M.C., O'Brien, R.N., Zhang, N., *et al.*, 2014. Polyglutamine disease modeling: Epitope based screen for homologous recombination using CRISPR/Cas9 system. *PLoS Currents* 6, PMC3994193, pii: ecurrents.hd.0242d2e7ad72225efa72f6964589369a. doi:10.1371/currents.hd.0242d2e7ad72225efa72f6964589369a.
- An, M.C., Zhang, N., Scott, G., *et al.*, 2012. Genetic correction of Huntington's disease phenotypes in induced pluripotent stem cells. *Cell Stem Cell* 11, 253–263.
- Andersen, P.M., Al-Chalabi, A., 2011. Clinical genetics of amyotrophic lateral sclerosis: What do we really know? *Nature Reviews Neurology* 7, 603–615.
- Andrew, S.E., Goldberg, Y.P., Kremer, B., *et al.*, 1993. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nature Genetics* 4, 398–403.
- Arrasate, M., Mitra, S., Schweitzer, E.S., Segal, M.R., Finkbeiner, S., 2004. Inclusion body formation reduces levels of mutant huntingtin and the risk of neuronal death. *Nature* 431, 805–810.
- Bailus, B.J., Segal, D.J., 2014. The prospect of molecular therapy for Angelman syndrome and other monogenic neurologic disorders. *BMC Neuroscience* 15, 76.
- Berger, Z., Ravikumar, B., Menzies, F.M., *et al.*, 2006. Rapamycin alleviates toxicity of different aggregate-prone proteins. *Human Molecular Genetics* 15, 433–442.
- van der Burg, J.M., Bjorkqvist, M., Brundin, P., 2009. Beyond the brain: Widespread pathology in Huntington's disease. *Lancet Neurology* 8, 765–774.
- Carlesi, C., Pasquali, L., Piazza, S., *et al.*, 2011. Strategies for clinical approach to neurodegeneration in amyotrophic lateral sclerosis. *Archives Italiennes de Biologie* 149, 151–167.
- de Carvalho, M., Pinto, S., Swash, M., 2010. Association of paraspinal and diaphragm denervation in ALS. *Amyotrophic Lateral Sclerosis* 11, 63–66.
- Chhajed, P.N., Dickenmann, M., Bubendorf, L., *et al.*, 2006. Patterns of pulmonary complications associated with sirolimus. *Respiration* 73, 367–374.
- Cohen, T., 2013. Mayo clinic to hold trial for brainstem's ALS stem cell therapy. *Reuters*. 18 Mar 2013.
- Cuervo, A.M., Bergamini, E., Brunk, U.T., *et al.*, 2005. Autophagy and aging: The importance of maintaining "clean" cells. *Autophagy* 1, 131–140.
- de Die-Smulders, C.E., de Wert, G.M., Liebaers, I., Tibben, A., Evers-Kiebooms, G., 2013. Reproductive options for prospective parents in families with Huntington's disease: Clinical, psychological and ethical reflections. *Human Reproduction Update* 19, 304–315.
- van Duijn, E., Kingma, E.M., van der Mast, R.C., 2007. Psychopathology in verified Huntington's disease gene carriers. *Journal of Neuropsychiatry and Clinical Neurosciences* 19, 441–448.
- Ellerby, L.M., Orr, H.T., 2006. Neurodegenerative disease: Cut to the chase. *Nature* 442, 641–642.
- Fernandez-Borges, N., Erana, H., Elezgarai, S.R., *et al.*, 2013. Infectivity versus seeding in neurodegenerative diseases sharing a prion-like mechanism. *International Journal of Cell Biology* 2013 9, 583498.
- Filimonenko, M., Isakson, P., Finley, K.D., *et al.*, 2010. The selective macroautophagic degradation of aggregated proteins requires the PI3P-binding protein Alf1. *Molecular Cell* 38, 265–279.
- Filimonenko, M., Stuffers, S., Raiborg, C., *et al.*, 2007. Functional multivesicular bodies are required for autophagic clearance of protein aggregates associated with neurodegenerative disease. *Journal of Cell Biology* 179, 485–500.
- Gafni, J., Ellerby, L.M., 2002. Calpain activation in Huntington's disease. *Journal of Neuroscience* 22, 4842–4849.
- Gafni, J., Papanikolaou, T., Degiacomo, F., *et al.*, 2012. Caspase-6 activity in a BACHD mouse modulates steady-state levels of mutant huntingtin protein but is not necessary for production of a 586 amino acid proteolytic fragment. *Journal of Neuroscience* 32, 7454–7465.
- Garbuzova-Davis, S., Rodrigues, M.C., Hernandez-Ontiveros, D.G., *et al.*, 2011. Amyotrophic lateral sclerosis: A neurovascular disease. *Brain Research* 1398, 113–125.
- Ghavami, S., Shojaei, S., Yeganeh, B., *et al.*, 2014. Autophagy and apoptosis dysfunction in neurodegenerative disorders. *Progress in Neurobiology* 112, 24–49.
- Gordon, P.H., Moore, D.H., Miller, R.G., *et al.*, 2007. Efficacy of minocycline in patients with amyotrophic lateral sclerosis: A phase III randomised trial. *Lancet Neurology* 6, 1045–1053.
- Grinthal, A., Adamovic, I., Weiner, B., Karplus, M., Kleckner, N., 2010. PR65, the HEAT-repeat scaffold of phosphatase PP2A, is an elastic connector that links force and catalysis. *Proceedings of the National Academy of Sciences of the United States of America* 107, 2467–2472.
- Hardiman, O., van den Berg, L.H., Kiernan, M.C., 2011. Clinical diagnosis and management of amyotrophic lateral sclerosis. *Nature Reviews Neurology* 7, 639–649.
- Hermel, E., Gafni, J., Propp, S.S., *et al.*, 2004. Specific caspase interactions and amplification are involved in selective neuronal vulnerability in Huntington's disease. *Cell Death & Differentiation* 11, 424–438.
- Hoffner, G., Djian, P., 2014. Polyglutamine aggregation in Huntington disease: Does structure determine toxicity? *Molecular Neurobiology*, published online October 2014, Springer US. doi:10.1007/s12035-014-8932-1.
- Holmes, B.B., Devos, S.L., Kfoury, N., *et al.*, 2013. Heparan sulfate proteoglycans mediate internalization and propagation of specific proteopathic seeds. *Proceedings of the National Academy of Sciences of the United States of America* 110, E3138–E3147.
- Horii, T., Tamura, D., Morita, S., Kimura, M., Hatada, I., 2013. Generation of an ICF syndrome model by efficient genome editing of human induced pluripotent stem cells using the CRISPR system. *International Journal of Molecular Sciences* 14, 19774–19781.
- Hou, Z., Zhang, Y., Propson, N.E., *et al.*, 2013. Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*. *Proceedings of the National Academy of Sciences of the United States of America* 110, 15644–15649.
- Huntington Study Group, D. I., 2010. A futility study of minocycline in Huntington's disease. *Movement Disorders* 25, 2219–2224.
- Huntington Study, G., 2006. Tetrabenazine as antichorea therapy in Huntington disease: A randomized controlled trial. *Neurology* 66, 366–372.
- Katsuno, M., Watanabe, H., Yamamoto, M., Sobue, G., 2014. Potential therapeutic targets in polyglutamine-mediated diseases. *Expert Review of Neurotherapeutics* 14, 1215–1228.
- Kiernan, M.C., Vucic, S., Cheah, B.C., *et al.*, 2011. Amyotrophic lateral sclerosis. *Lancet* 377, 942–955.
- Kikis, E.A., Gidalevitz, T., Morimoto, R.I., 2010. Protein homeostasis in models of aging and age-related conformational disease. *Advances in Experimental Medicine and Biology* 694, 138–159.
- Koga, H., Martinez-Vicente, M., Arias, E., *et al.*, 2011. Constitutive upregulation of chaperone-mediated autophagy in Huntington's disease. *Journal of Neuroscience* 31, 18492–18505.
- Krawczak, M., Bockel, B., Sandkuijl, L., *et al.*, 1991. Covariate-dependent age-at-onset distributions for Huntington disease. *American Journal of Human Genetics* 49, 735–745.
- Labbadia, J., Morimoto, R.I., 2013. Huntington's disease: Underlying molecular mechanisms and emerging concepts. *Trends in Biochemical Sciences* 38, 378–385.
- Lamming, D.W., Ye, L., Katajisto, P., *et al.*, 2012. Rapamycin-induced insulin resistance is mediated by mTORC2 loss and uncoupled from longevity. *Science* 335, 1638–1643.
- Lin, N.Y., Beyer, C., Giessel, A., *et al.*, 2013. Autophagy regulates TNF α -mediated joint destruction in experimental arthritis. *Annals of the Rheumatic Diseases* 72, 761–768.
- Lovejoy, D.B., Guillemin, G.J., 2014. The potential for transition metal-mediated neurodegeneration in amyotrophic lateral sclerosis. *Frontiers in Aging Neuroscience* 6, 173.
- Mangiarini, L., Sathasivam, K., Seller, M., *et al.*, 1996. Exon 1 of the HD gene with an expanded CAG repeat is sufficient to cause a progressive neurological phenotype in transgenic mice. *Cell* 87, 493–506.
- Margulis, J., Finkbeiner, S., 2014. Proteostasis in striatal cells and selective neurodegeneration in Huntington's disease. *Frontiers in Cellular Neuroscience* 8, 218.
- Montoya, A., Price, B.H., Menear, M., Lepage, M., 2006. Brain imaging and cognitive dysfunctions in Huntington's disease. *Journal of Psychiatry & Neuroscience* 31, 21–29.
- Munch, C., Bertolotti, A., 2011. Self-propagation and transmission of misfolded mutant SOD1: Prion or prion-like phenomenon? *Cell Cycle* 10, 1711.
- Munch, C., O'Brien, J., Bertolotti, A., 2011. Prion-like propagation of mutant superoxide dismutase-1 misfolding in neuronal cells. *Proceedings of the National Academy of Sciences of the United States of America* 108, 3548–3553.
- Myers, R.H., 2004. Huntington's disease genetics. *NeuroRx Research* 1, 255–262.

- Ochaba, J., Lukacsovich, T., Csikos, G., *et al.*, 2014. Potential function for the Huntingtin protein as a scaffold for selective autophagy. *Proceedings of the National Academy of Sciences of the United States of America* 111 (47), 16889–16894.
- Osborn, A.G., Salzman, K., Katzman, G., *et al.*, 2004. Diagnostic Imaging: Brain. Salt Lake City: Amirsys.
- Petruska, J., Hartenstine, M.J., Goodman, M.F., 1998. Analysis of strand slippage in DNA polymerase expansions of CAG/CTG triplet repeats associated with neurodegenerative disease. *Journal of Biological Chemistry* 273, 5204–5210.
- Pham, P.T., Pham, P.C., Danovitch, G.M., *et al.*, 2004. Sirolimus-associated pulmonary toxicity. *Transplantation* 77, 1215–1220.
- Phukan, J., Pender, N.P., Hardiman, O., 2007. Cognitive impairment in amyotrophic lateral sclerosis. *Lancet Neurology* 6, 994–1003.
- Pla, P., Orvoen, S., Saudou, F., David, D.J., Humbert, S., 2014. Mood disorders in Huntington's disease: From behavior to cellular and molecular mechanisms. *Frontiers in Behavioral Neuroscience* 8, 135.
- Plane, J.M., Shen, Y., Pleasure, D.E., Deng, W., 2010. Prospects for minocycline neuroprotection. *Archives of neurology* 67, 1442–1448.
- Poon, L.H., Kang, G.A., Lee, A.J., 2010. Role of tetraabenazine for Huntington's disease-associated chorea. *Annals of Pharmacotherapy* 44, 1080–1089.
- Pratt, A.J., Shin, D.S., Merz, G.E., *et al.*, 2014. Aggregation propensities of superoxide dismutase G93 hotspot mutants mirror ALS clinical phenotypes. *Proceedings of the National Academy of Sciences of the United States of America* 111, E4568–E4576.
- Pryor, W.M., Biagioli, M., Shahani, N., *et al.*, 2014. Huntingtin promotes mTORC1 signaling in the pathogenesis of Huntington's disease. *Science Signaling* 7, ra103.
- Ravikumar, B., Berger, Z., Vacher, C., O'Kane, C.J., Rubinshtein, D.C., 2006. Rapamycin pre-treatment protects against apoptosis. *Human Molecular Genetics* 15, 1209–1216.
- Robberecht, W., Philips, T., 2013. The changing scene of amyotrophic lateral sclerosis. *Nature Reviews Neuroscience* 14, 248–264.
- Sadan, O., Melamed, E., Offen, D., 2009. Bone-marrow-derived mesenchymal stem cell therapy for neurodegenerative diseases. *Expert Opinion on Biological Therapy* 9, 1487–1497.
- Segal, D.J., Meckler, J.F., 2013. Genome engineering at the dawn of the golden age. *Annual Review of Genomics and Human Genetics* 14, 135–158.
- Shefner, J.M., Watson, M.L., Meng, L., Wolff, A.A., Neals/Cytokinetics, S.T., 2013. A study to evaluate safety and tolerability of repeated doses of tirasemtiv in patients with amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* 14, 574–581.
- Snell, R.G., MacMillan, J.C., Cheadle, J.P., *et al.*, 1993. Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. *Nature Genetics* 4, 393–397.
- Tepper, J.M., Wilson, C.J., Koos, T., 2008. Feedforward and feedback inhibition in neostriatal GABAergic spiny neurons. *Brain Research Reviews* 58, 272–281.
- Tourette, C., Li, B., Bell, R., *et al.*, 2014. A large scale Huntingtin protein interaction network implicates Rho GTPase signaling pathways in Huntington disease. *Journal of Biological Chemistry* 289, 6709–6726.
- Trevino, R.S., Lauckner, J.E., Sourigues, Y., *et al.*, 2012. Fibrillar structure and charge determine the interaction of polyglutamine protein aggregates with the cell surface. *Journal of Biological Chemistry* 287, 29722–29728.
- Urushitani, M., Sik, A., Sakurai, T., *et al.*, 2006. Chromogranin-mediated secretion of mutant superoxide dismutase proteins linked to amyotrophic lateral sclerosis. *Nature Neuroscience* 9, 108–118.
- Van Den Bosch, L., Tilkin, P., Lemmens, G., Robberecht, W., 2002. Minocycline delays disease onset and mortality in a transgenic model of ALS. *NeuroReport* 13, 1067–1070.
- Walker, F.O., 2007. Huntington's disease. *Lancet* 369, 218–228.
- Warby, S.C., Doty, C.N., Graham, R.K., *et al.*, 2008. Activated caspase-6 and caspase-6-cleaved fragments of huntingtin specifically colocalize in the nucleus. *Human Molecular Genetics* 17, 2390–2404.
- Wellington, C.L., Ellerby, L.M., Gutekunst, C.A., *et al.*, 2002. Caspase cleavage of mutant huntingtin precedes neurodegeneration in Huntington's disease. *Journal of Neuroscience* 22, 7862–7872.
- Wellington, C.L., Singaraja, R., Ellerby, L., *et al.*, 2000. Inhibiting caspase cleavage of huntingtin reduces toxicity and aggregate formation in neuronal and nonneuronal cells. *Journal of Biological Chemistry* 275, 19831–19838.
- Ye, L., Wang, J., Beyer, A.I., *et al.*, 2014. Seamless modification of wild-type induced pluripotent stem cells to the natural CCR5Delta32 mutation confers resistance to HIV infection. *Proceedings of the National Academy of Sciences of the United States of America* 111, 9591–9596.
- Yin, H., Xue, W., Chen, S., *et al.*, 2014. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nature Biotechnology* 32, 551–553.
- Yu, A., Shibata, Y., Shah, B., *et al.*, 2014. Protein aggregation can inhibit clathrin-mediated endocytosis by chaperone competition. *Proceedings of the National Academy of Sciences of the United States of America* 111, E1481–E1490.
- Zheng, G., Dwoskin, L.P., Crooks, P.A., 2006. Vesicular monoamine transporter 2: Role as a novel target for drug development. *AAPS Journal* 8, E682–E692.

Further Reading

- Mardones, P., Martinez, G., Hetz, C., 2015. Control of systemic proteostasis by the nervous system. *Trends in Cell Biology* 25 (1), 1–10.

Site-Directed Mutagenesis

KW Walker, Amgen Inc., Thousand Oaks, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Background

Site-directed mutagenesis (SDM), sometimes referred to as site-specific or directed mutagenesis, is a method of altering the nucleotide sequence of a gene at a specified location, which stands in contrast to general mutagenesis that employs mutagenic compounds or high-energy radiation to randomly alter DNA. SDM is a highly versatile technology that can be used to probe gene function of an isolated gene product in the context of a complex biological system, and it can be used to engineer proteins and even entire organisms for new functions or capabilities. The magnitude of the impact and widespread use of this technology in the biological sciences can be partially appreciated by conducting a PubMed search of the term 'directed mutagenesis,' which identifies 68 832 citations (as of April 2014).

The revolutionary approach that allowed for flexible SDM employing DNA polymerase and the primer extension method was pioneered by Clyde Hutchinson and Michael Smith in 1978, for which Michael Smith shared the 1993 Nobel Prize in Chemistry with Kary Mullis who developed another key technology for SDM, polymerase chain reaction (PCR). While the techniques evolved substantially over time, the fundamental aspect of employing mutation containing primers and DNA polymerases is still the cornerstone of modern techniques.

Methods

The basic technique for SDM employs a short synthetic DNA primer containing the desired mutation (base changes, deletions, or insertions), which is then hybridized to the template DNA that contains the parental gene of interest (GOI). The hybridized primer is then extended using DNA polymerase to duplicate the remainder of the GOI, which is then inserted into a host vector to enable propagation and expression of the mutated gene. The final step requires sequencing of the GOI to select those clones that contain the desired mutation. Modern variations of this process include the QuikChange method (Figure 1), which employs a pair of complementary primers containing the desired mutation, which are employed to synthesize the entire plasmid using a high-fidelity non-strand displacing polymerase such as *pfu* (Zheng *et al.*, 2004). The template DNA is then eliminated by enzymatic digestion with a restriction enzyme specific for the methylated template DNA, which enables elimination of the parental copy of the GOI. For multisite mutagenesis, alternative strategies such as Golden GATEway (Kirchmaier *et al.*, 2013), which is a cassette mutagenesis strategy that does not require retention of the restriction site; multiple patch cloning (MUPAC) (Taniguchi *et al.*, 2013), which is designed for introducing up to nine point mutants in a single GIO; as well as sequence and ligation independent cloning (SLIC) (Li and

Elledge, 2007), which is a cassette mutagenesis strategy that also allows for multiple mutations in a single reaction, are all excellent options. Additionally, with the substantial reduction in the cost of synthetic DNA primers, it is also practical in many circumstances to synthetically produce a cassette that is bordered by two unique restriction sites, or even the entire GOI, with the desired mutations.

Applications

SDM can be employed to substitute single or multiple amino acids in a protein, remove or insert amino acids, and alter the noncoding regions of a gene such as regulatory elements (Figure 2). Ultimately these directed changes can be used to probe biological systems for structure–function relationships, such as determining critical residues for enzyme catalysis or

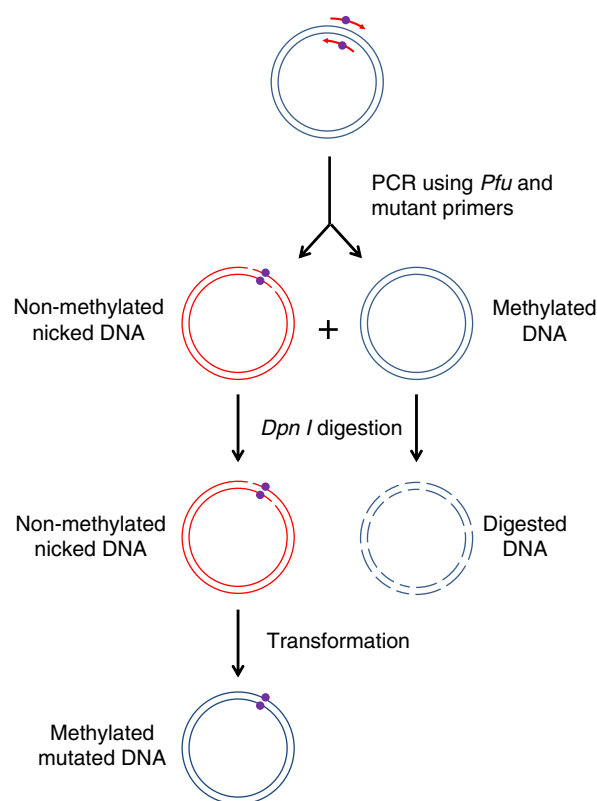


Figure 1 QuikChange method of SDM. The first step employs a template plasmid, methylated (blue), containing the GOI, the accurate *pfu* DNA polymerase, and complimentary primers (red arrows) containing the desired mutation (purple dot) which replicate the DNA with the mutation producing non-methylated (red) DNA. The methylated parental plasmid is then digested with *Dpn I*, while the non-methylated mutant containing plasmid is resistant to degradation. The nicked mutation containing plasmid is then transformed into a host cell where the nick is repaired, it is replicated and methylated.

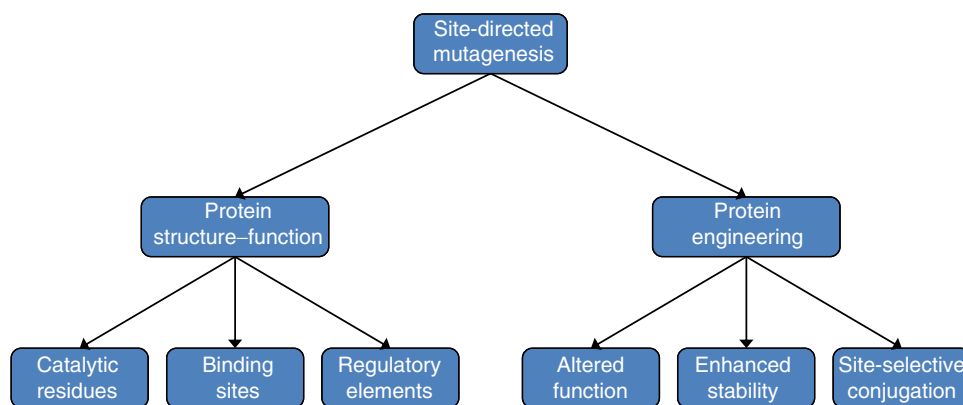


Figure 2 Applications of SDM. SDM is used to probe structure–function relationships including catalytic residues, binding-site residues, and regulatory elements. In addition, SDM is used to engineer proteins to enable enhanced function, increased stability, and site-selective conjugation.

substrate/cofactor binding, as well as identifying residues involved in regulating protein function. In addition, SDM can be employed to engineer new properties into proteins such as altered substrate specificity, thermal stability, and site-directed conjugation, all of which have applications in research and biotechnology. A powerful variant of SDM is alanine scanning mutagenesis, in which every residue in the entire protein, or region of interest, is individually changed to alanine and the functional properties of the variants are assessed (Cunningham and Wells, 1989). This allows for a systematic determination of the structure–function relationship of each amino acid in a protein without structural information or prior knowledge of the residue’s role. A notable advantage of alanine scanning mutagenesis is that it can identify allosteric effects and often leads to unexpected findings which broaden our knowledge of protein architecture.

Protein Structure–Function

Catalytic residues

Prior to the advent of SDM, identification of the catalytic residues in an enzyme was a painstaking process often involving chemical modification of the protein followed by peptide mapping and sequencing. However, SDM allows for the direct probing of the functional significance of residues suspected of being involved in enzyme-catalyzed reactions, and often enables subtle mechanistic determinations. For example, shortly after the development of SDM, it was employed to determine that Cys-35 in tyrosyl-tRNA synthetase, Glu-165 of triphosphate isomerase, and Asn-155 of subtilisin are all involved in stabilizing the transition state of their reaction intermediates. This was accomplished by changing these residues to other amino acids and observing the effect on enzyme function (Wilkinson *et al.*, 1983; Straus *et al.*, 1985; Bryan *et al.*, 1986). SDM can also be used to probe the catalytic roles of proximal and identical residues that cannot be segregated by chemical probes. For example, it was determined that chemical modification of Lys-83 or Lys-84 in aspartate transcarbamoylase inactivates the enzyme; however, it was not possible to determine if either or both residues were required. By mutating each residue individually it was demonstrated

that only Lys-84 is essential for catalysis, since substitution of Lys-83 with Gln does not significantly impact activity while this substitution abolishes activity at position 84 (Robey *et al.*, 1986). SDM was also used to dissect the novel mechanistic role of two nearby active site cysteines in protein disulfide isomerase, demonstrating that one is responsible for attacking the substrate disulfide bond creating a covalent complex with the substrate, while the other acts as a timer releasing the covalently bound enzyme, should the enzyme become kinetically trapped with the substrate (Walker and Gilbert, 1997).

SDM is also very useful for confirming homology model predictions by mutating residues predicted to have functional roles to determine if the expected outcome is achieved. This technique involves predicting the tertiary structure of a protein by sequence alignments to related proteins with known structures and then computationally generating a three-dimensional model of the target protein. Since the resulting model is only a prediction, SDM can be used to verify the accuracy of the model as demonstrated by Vasu *et al.* (2012) for *cis*-epoxysuccinate hydrolase where SDM was employed to validate the residues hypothesized to be part of a catalytic triad.

Binding sites

SDM is also a very powerful tool for mapping and verifying the residues predicted to be critical for substrate recognition, cofactor binding, and protein–protein interactions. For example, computational alanine scanning mutagenesis was used to identify potential substrate interactions sites in maltotriose trehalohydrolase. The *in silico* scanning mutagenesis predictions provided insight into residues potentially involved in trehalose binding, which were subsequently confirmed by actual mutagenesis studies (Fu *et al.*, 2013). SDM is also frequently used to probe the functional role of metal binding sites. For example, Kiraly *et al.* (2009) used SDM to alter residues of five Ca^{2+} binding sites in human transglutaminase 2, which allowed them to determine that while all five Ca^{2+} ions influence transglutaminase activity, only two of the sites are involved in GTPase activity and only one is responsible for determining the antigenicity associated with celiac disease.

Even when structural information is available to assist in identifying substrate or cofactor interacting residues, SDM is

critical for determining if the crystal structure identified interactions are significant in the solution phase and are not artifacts of crystallization. Ouyang *et al.* (2013) employed a combination of alanine scanning mutagenesis and salt-bridge residue substitutions to demonstrate that many of the side chains of HIV-1 nucleocapsid protein predicted by the crystal structure to interact with Zn^{2+} ions do not significantly contribute to the stability of the complex. In another example of SDM being used to clarify misleading crystal structure interactions, Koch *et al.* (2012) demonstrated that solution binding to a glutaminyl cyclase inhibitor in solution was different than that observed in the co-crystal structure.

In some cases, a combination of chemical modification and SDM can be used in concert to probe protein-protein interactions. For example, Bradshaw *et al.* (1994) employed this combination of techniques to demonstrate that aromatic residues are critical for stabilization of nerve growth factor dimer, while surface lysine residues from both subunits in the nerve growth factor dimer are responsible for receptor interactions.

Regulatory elements

SDM can be used as a powerful probe to identify and confirm the role of specific amino acids in the regulation of protein activity, and it is most frequently employed to probe the role of known and putative phosphorylation sites. Mutations that eliminate phosphorylation potential (i.e., alanine substitutions for serine and threonine and phenylalanine substitutions for tyrosine), when coupled with mass spectrometry, can be used to determine if the predicted site is actually phosphorylated, and it can be used to determine if phosphorylation of that site impacts protein function. For example, the role of tyrosine phosphorylation on fibroblast growth factor receptor 1 signaling was investigated by mutating the endodomain tyrosines to phenylalanine and observing the impact on cellular differentiation, resulting in the identification of two tyrosines critical for bioactivity (Foehr *et al.*, 2001). In another example, the functional role of multiple phosphorylation sites on the viral NS1 protein was dissected by individually mutating the phosphorylation sites and determining the impact on NS1 activity. This resulted in identification of two new *in vivo* phosphorylation sites and demonstrated that certain phosphorylation sites were responsible for switching the virus from productive to cytotoxic functions (Corbau *et al.*, 2000).

It is also possible to mimic phosphorylation of serine and threonine residues by converting them to aspartic or glutamic acid, which has a charge and structure similar to phosphoserine and phosphothreonine. For example, Woods and colleagues individually mutated three potentially phosphorylated serines and threonines to acidic residues in order to mimic phosphorylation and determine their role in AMP-activated protein kinase (AMPK), a protein responsible for energy metabolism regulation. Using this method, they were able to determine that only one of the three residues was involved in AMPK activation, thus extending the mechanistic understanding of an important target for developing therapeutics for treating metabolic disorders (Woods *et al.*, 2003). An alternative indirect SDM-based method used to identify phosphorylation sites employs incorporation of basic residues to

create nonnatural tryptic digest sites enabling cleavage of complex phosphoproteins into more manageable segments for mass spectroscopic analysis. This method was employed with the checkpoint protein kinase resulting in identification of 17 novel phosphorylation sites in this complex glycoprotein (King *et al.*, 2007).

Protein Engineering

Altered function

SDM can be employed to change the substrate specificity of an enzyme, which is a powerful tool for probing enzyme mechanism as well as enabling proteins to catalyze new chemistries of use for scientific and industrial applications. For example, Wiersma-Koch *et al.* (2013) were able to alter the substrate preference of the nucleotide pyrophosphatase/phosphodiesterase from a 10^6 -fold preference for diesters over monoesters to equal activity on both substrates with a 10^{11} -fold rate enhancement for the monoester. This was accomplished by using structural information to predict the substrate binding pocket followed by mutating four of those residues to generate the greatly enhanced enzyme, which is of use for biomedical and industrial applications. In another example of altering enzyme substrate specificity, mutation of a single residue in lysyl-tRNA synthetase, which converts ATP into diadenosine tri- and tetraphosphates, altered Zn^{2+} binding, resulting in greatly reduced diadenosine formation activity, ultimately shifting the protein to a substantially different function as a glycerol kinase (Chen *et al.*, 2013).

Homology modeling is a powerful method that can be employed to identify substrate binding residues in two related enzymes that have differing substrate specificities, allowing one to predict how the binding sites may be changed to alter substrate specificity for protein engineering applications. This was demonstrated by mutating two of the binding-site pocket residues of a purine-specific nucleoside hydrolase, converting it into a protein that is capable of hydrolyzing both purines and pyrimidines (Porcelli *et al.*, 2012). Homology modeling was also employed to predict the residues involved in the substrate specificity profile of a eukaryotic methionine aminopeptidase using the prokaryotic isozyme crystal structure. This information was used to create variants of the eukaryotic enzyme that have a broadly enhanced substrate profile, which can be used to produce native amino termini in recombinantly produced proteins for research and pharmaceutical applications (Walker and Bradshaw, 1999). In a unique approach to alter substrate specificity, Hui *et al.* (2013) employed both SDM and chemical conjugation to alter the substrate specificity of a lipase. They mutated an alanine to a reactive cysteine, and then chemically coupled that cysteine to thio-nitrobenzoic acid. This provided stronger steric exclusion and structural rigidity to the protein, which resulted in substantially enhanced diastereopreference with potential application in synthetic chemical production.

Stability

SDM has been widely employed to enhance the stability of proteins to thermal, pH, and oxidative stress. Increased protein stability is particularly useful for research, pharmaceutical, and

industrial chemistry applications, since reactions are often much more efficient at elevated temperatures, more extreme pH, and often occur under conditions that would put many proteins at risk for oxidative inactivation.

Thermal

Increases in the thermal stability of enzymes is often critical to enable the proteins to remain active at the high reaction temperatures often required for specific applications in both research and industry. For example, in order to develop a DNA polymerase that is active at the high temperatures required during PCR, while maintaining an exceptionally high fidelity, two point mutations were incorporated into *Thermococcus cel-ericrescens* DNA polymerase. These changes produced an enzyme with enhanced thermo stability, while maintaining its substantially lower error rate compared to *Taq* DNA polymerase, thus enabling long and accurate PCR of large gene segments (Kim *et al.*, 2011). In another example, to improve the production of D-amino acids for use in pharmaceuticals and the food industry, five point mutations were incorporated into a mesophilic D-amino acid dehydrogenase, enabling it to maintain complete activity at 65 °C and greatly enhancing its industrial utility (Akita *et al.*, 2012).

The food industry is a major user of enzymes, and improvements in thermal stability are often required to make the enzymes practical for this application. For example, pullanase can be used for starch debranching during the production of high glucose syrup; however, the wild type enzyme has relatively low stability and catalytic efficiency preventing its widespread application. To accomplish this, Duan *et al.* (2013) incorporated two point mutations in pullanase improving the half-life at 60 °C by 4.3-fold and increasing its catalytic efficiency by 100%, enabling industrial use of the enzyme. Another starch processing enzyme, α -amylase, achieved a melting temperature increase of 26 °C by employing two point mutations combined with a two amino acid deletion to produce an enzyme that could operate effectively at 70 °C (Kachan and Evtushenkov, 2013).

pH

Often the pH optimum of an enzyme does not match the pH of the environment for which it is needed, providing an opportunity for SDM to produce a better match between the protein and its operating environment. For example, bio bleaching in the pulp and paper industry requires both high temperatures and alkaline pH; however, known xylanases, which can catalyze bleaching reactions, do not function under those conditions. To address this, Zheng *et al.* (2014) combined three point mutations and one nonnative disulfide bond to increase the pH optimum of xylanase from 7.0 to 9.0 and increase the optimum temperature from 60 to 70 °C. In some circumstances, a more acidic pH optimum is required, such as for β -mannanase, which could be included in animal feed to increase the availability of manno-oligosaccharides, important as growth factors for beneficial intestinal microorganisms. Since the intestinal pH is low (5.5–6.0), a β -mannanase with a single point mutation was developed that dropped the pH optimum of the enzyme from 6.5 to 5.5 enabling its practical use as a feed additive (Xu *et al.*, 2013).

Oxidative

Many applications for enzymes expose them to an oxidative environment which can often result in inactivation of the protein. Although methionine, cysteine, and tryptophan are all susceptible to oxidation, methionine is frequently the most problematic residue, due to its higher oxidative vulnerability compared to tryptophan and its much wider prevalence, compared to reduced cysteine, in enzymes important for research and industrial applications. Generally only a sub-population of the oxidation vulnerable residues need to be changed to maintain function, since many are buried in the protein, and thus, not exposed to the oxidant, or they are not critical for activity. For example, *N*-acetyl-D-amino acid amidohydrolase is useful for the production of D-amino acids used in food and pharmaceutical products; however, this protein is susceptible to inactivation through methionine oxidation. Peng *et al.* (2012) were able to engineer this enzyme to increase its half-life by 6-fold and its catalytic efficiency 2.4-fold by mutating a single methionine to leucine. Enzymes are also frequently used in laundry detergent to enhance cleaning; however, oxidants are often present in the detergent, which can inactivate many enzymes. To address this, a single methionine in α -amylase from *Thermotoga maritima* was mutated to alanine enabling the enzyme to maintain 50% of its activity in the presence of 100 mM hydrogen peroxide, conditions under which the wild type enzyme is completely inactive (Ozturk *et al.*, 2013). Enzymes are often useful for bioremediation applications, which frequently require exposure to harsh environments. For example, the recalcitrant anthraquinone dyes often present in wastewaters are good targets for bioremediation using the unique *Anabaena* peroxidase; however, this enzyme is susceptible to oxidative inactivation in the presence of the required quantity of hydrogen peroxide. Mutation of a single methionine residue to phenylalanine enabled this enzyme to operate in the presence of 5 mM hydrogen peroxide enabling its use for wastewater management (Ogola *et al.*, 2010).

Site-selective conjugation

SDM can be employed to create specific positions on a protein for selective modification by creating a unique chemically reactive handle for immobilization or conjugation to a chemical moiety, or it can be used for directing cell mediated post-translational modifications such as glycosylation. For example, aequorin is a photoprotein that emits light upon calcium binding and is useful as a calcium sensor in some applications that require immobilization of the protein. In order to allow site-selective immobilization of aequorin, the native cysteines were selectively mutated to non-thiol containing residues, and then a cysteine was incorporated at a new site on the protein, which provides a unique thiol nucleophile that can be coupled to a wide variety of solid substrates (Lewis *et al.*, 2000). This allows for site-selective coupling of aequorin to the solid surface in a uniform manner that prevents inactivation often associated with random coupling.

Use of endogenous proteins as biotherapeutics often requires addition of large polymers such as polyethylene glycol (PEG) to improve pharmacokinetic performance and reduce injection frequency. While random PEG attachment has been employed in the past, it is highly preferable to produce a

uniform product in order to maximize bioactivity and reduce lot to lot variability. In the case of fibroblast growth factor 21, reactive cysteines residues were engineered on to the surface of this metabolic regulatory protein to allow for site-specific PEG conjugation. This enabled production of a uniform protein with substantially improved pharmacokinetic properties appropriate for potential application as a therapeutic for type II diabetes (Xu *et al.*, 2013).

An alternative to *in vitro* conjugation of synthetic molecules at engineered sites for serum half-life enhancement of therapeutic proteins involves the addition of sialated glycans during expression in the cell producing the recombinant protein. This approach has the advantage over synthetic conjugation in that it does not require an additional conjugation and subsequent purification step during the manufacturing process. Erythropoietin was modified in this manner by incorporating five mutations, creating two new consensus N-glycosylation sites, resulting in a protein, darbepoetin alfa, with a 3-fold improvement in the serum half-life, thus decreasing the number of injections required to achieve the therapeutic effect (Egrie and Browne, 2001).

Summary

SDM is a tremendously powerful technique that allows us to probe the function of highly complex protein molecules in a very precise manner, and it allows us to engineer proteins to enhance their functions. Even though SDM was developed over 35 years ago, application of this technology continues to accelerate due to the increasing need to investigate complex biological systems and the expanded demand for engineered proteins for research, food processing, organism engineering, bioremediation, high-value chemical production, and pharmaceutical applications. Increases in the efficiency of SDM, coupled with lowering costs, is enabling widespread application of this technology to enhance our understanding of biology and create novel proteins to address a wide variety of technological needs.

See also: Molecular Principles, Components, Technology, and Concepts: Basic Principles: Biocatalysis. Molecular Principles, Components, Technology, and Concepts: Nucleic Acids: DNA, RNA Chemical Properties (Including Sequencing and Next-Generation Sequencing). Molecular Principles, Components, Technology, and Concepts: Proteins: Drug Design; Expression Systems; Protein Domains: Structure, Function, and Methods; Protein Sequence Determination: Methodology and Evolutionary Implications

References

- Akita, H., Doi, K., Kawarabayasi, Y., Ohshima, T., 2012. Creation of a thermostable NADP⁺-dependent D-amino acid dehydrogenase from *Ureibacillus thermosphaericus* strain A1 meso-diaminopimelate dehydrogenase by site-directed mutagenesis. *Biotechnology Letters* 34, 1693–1699.
- Bradshaw, R., Murray-Rust, J., Ibanez, C., *et al.*, 1994. Nerve growth factor: Structure/function relationships. *Protein Science* 3, 1901–1913.
- Bryan, P., Pantoliano, M., Quill, S., Hsiao, H., Polous, T., 1986. Site-directed mutagenesis and the role of the oxyanion hole in subtilisin. *Proceedings of the National Academy of Sciences of the United States of America* 83, 3743–3745.
- Chen, X., Boonyalai, N., Lau, C., *et al.*, 2013. Multiple catalytic activities of *Escherichia coli* lysyl-tRNA synthetase (LysU) are dissected by site-directed mutagenesis. *FEBS Journal* 280, 102–114.
- Corbau, R., Duverger, V., Rommelaere, J., NuEesch, J., 2000. Regulation of MVM NS1 by protein kinase C: Impact of mutagenesis at consensus phosphorylation sites on replicative functions and cytopathic effects. *Virology* 278, 151–167.
- Cunningham, B., Wells, J., 1989. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* 244 (4908), 1081–1085.
- Duan, X., Chen, J., Wu, J., 2013. Improving the thermostability and catalytic efficiency of *Bacillus deramificans* pullulanase by site-directed mutagenesis. *Applied and Environmental Microbiology* 79, 4072–4077.
- Egrie, J., Browne, J., 2001. Development and characterization of novel erythropoiesis stimulating protein (NESP). *Nephrology Dialysis Transplantation* 16 (Suppl. 3), 3–13.
- Foehr, E., Raffioni, S., Murray-Rust, J., Bradshaw, R., 2001. The role of tyrosine residues in fibroblast growth factor receptor 1 signaling in PC12 cells. *Journal of Biological Chemistry* 276, 37529–37536.
- Fu, C., Wang, Y., Fang, T., Lin, T., 2013. Interaction between trehalose and MTHase from *Sulfolobus solfataricus* studied by theoretical computation and site-directed mutagenesis. *PLoS One* 8, 1–10.
- Hui, C., Jianping, W., Lirong, Y., Gang, X., 2013. A combination of site-directed mutagenesis and chemical modification to improve diastereopreference of *Pseudomonas alcaligenes* lipase. *Biochimica et Biophysica Acta* 1834, 2494–2501.
- Kachan, A., Evtushenkov, A., 2013. Thermostable mutant variants of *Bacillus* sp. 406 α -amylase generated by site-directed mutagenesis. *Central European Journal of Biology* 8 (4), 346–356.
- Kim, K., Cho, S., Lee, K., Youn, M., Kwon, S., 2011. Improved thermostability and PCR efficiency of *Thermococcus celericrescens* DNA polymerase via site-directed mutagenesis. *Journal of Biotechnology* 155, 156–163.
- King, J., Gross, J., Lovly, C., Pownica-Worms, H., Townsend, R., 2007. Identification of protein phosphorylation sites within Ser/Thr-rich cluster domains using site-directed mutagenesis and hybrid linear quadrupole ion trap Fourier transform ion cyclotron resonance mass spectrometry. *Rapid Communications in Mass Spectrometry* 21, 3443–3451.
- Kiraly, R., Csoz, E., Kurtan, T., *et al.*, 2009. Functional significance of five noncanonical Ca²⁺-binding sites of human transglutaminase 2 characterized by site-directed mutagenesis. *FEBS Journal* 276, 7083–7096.
- Kirchmaier, S., Lust, K., Wittbrodt, J., 2013. Golden GATEway cloning – A combinatorial approach to generate fusion and recombination constructs. *PLoS One* 8 (10), e76117.
- Koch, B., Buchholz, M., Wermann, M., *et al.*, 2012. Probing secondary glutamyl cyclase (QC) inhibitor interactions applying an in silico-modeling/site-directed mutagenesis approach: Implications for drug development. *Chemical Biology & Drug Design* 80, 937–946.
- Lewis, J., Lopez-Moya, J., Daunert, S., 2000. Bioluminescence and secondary structure properties of aequorin mutants produced for site-specific conjugation and immobilization. *Bioconjugate Chemistry* 11, 65–70.
- Li, M., Elledge, S., 2007. Harnessing homologous recombination *in vitro* to generate recombinant DNA via SLIC. *Nature Methods* 4 (3), 251–256.
- Ogola, H., Hashimoto, N., Miyabe, S., *et al.*, 2010. Enhancement of hydrogen peroxide stability of a novel *Anabaena* sp. DyP-type peroxidase by site-directed mutagenesis of methionine residues. *Applied Microbiology and Biotechnology* 87, 1727–1736.
- Ouyang, W., Okaine, S., McPike, M., Lin, Y., Borer, P., 2013. Probing the RNA binding surface of the HIV-1 nucleocapsid protein by site-directed mutagenesis. *Biochemistry* 52, 3358–3368.
- Ozturk, H., Ece, S., Gundeger, E., Evran, S., 2013. Site-directed mutagenesis of methionine residues for improving the oxidative stability of α -amylase from *Thermotoga maritima*. *Journal of Bioscience and Bioengineering* 116, 449–451.
- Peng, I., Lo, K., Hsu, C., Lee, C., 2012. Increasing the storage and oxidation stabilities of N-acyl-D-amino acid amidohydrolase by site-directed mutagenesis of critical methionine residues. *Process Biochemistry* 47, 1785–1790.
- Porcelli, M., De Leo, E., Marabotti, A., Cacciapuoti, G., 2012. Site-directed mutagenesis gives insights into substrate specificity of *Sulfolobus solfataricus* purine-specific nucleoside hydrolase. *Annals of Microbiology* 62, 881–887.
- Robey, E., Wente, S., Markby, D., *et al.*, 1986. Effect of amino acid substitutions on the catalytic and regulatory properties of aspartate transcarbamoylase. *Proceedings of the National Academy of Sciences of the United States of America* 83, 5934–5938.

- Straus, D., Raines, R., Kawashima, E., Knowles, J., Gilbert, W., 1985. Active site of triosephosphate isomerase: *In vitro* mutagenesis and characterization of an altered enzyme. *Proceedings of the National Academy of Sciences of the United States of America* 82, 2272–2276.
- Taniguchi, N., Nakayama, S., Kawakami, T., Murakami, H., 2013. Patch cloning method for multiple site-directed and saturation mutagenesis. *BMC Biotechnology* 13 (91), 1–8.
- Vasu, V., Kumaresan, J., Babu, M., Meenakshisundaram, S., 2012. Active site analysis of *cis*-epoxysuccinate hydrolase from *Nocardia tartaricans* using homology modeling and site-directed mutagenesis. *Applied Microbiology and Biotechnology* 93, 2377–2386.
- Walker, K., Bradshaw, R., 1999. Yeast methionine aminopeptidase I: Alteration of substrate specificity by site-directed mutagenesis. *Journal of Biological Chemistry* 274 (19), 13403–13409.
- Walker, K., Gilbert, H., 1997. Scanning and escape during protein-disulfide isomerase assisted protein folding. *Journal of Biological Chemistry* 272, 8845–8848.
- Wiersma-Koch, H., Sunden, F., Herschalag, D., 2013. Site-directed mutagenesis maps interactions that enhance cognate and limit promiscuous catalysis by an alkaline phosphatase superfamily phosphodiesterase. *Biochemistry* 52, 9167–9176.
- Wilkinson, A., Fersht, A., Blow, D., Winter, G., 1983. Site-directed mutagenesis as a probe of enzyme structure and catalysis: Tyrosyl-tRNA synthetase cysteine-35 mutation. *Biochemistry* 22, 3581–3586.
- Woods, A., Vertommen, D., Neumann, D., *et al.*, 2003. Identification of phosphorylation sites in AMP-activated protein kinase (AMPK) for upstream AMPK kinases and study of their roles by site-directed mutagenesis. *Journal of Biological Chemistry* 278 (31), 28434–28442.
- Xu, J., Bussiere, J., Yie, J., *et al.*, 2013. Polyethylene glycol modified FGF21 engineered to maximize potency and minimize vacuole formation. *Bioconjugate Chemistry* 24, 915–925.
- Zheng, H., Liu, Y., Sun, M., *et al.*, 2014. Improvement of alkali stability and thermostability of *Paenibacillus campinasensis* Family-11 xylanases by directed evolution and site-directed mutagenesis. *Journal of Industrial Microbiology and Biotechnology* 41, 153–162.
- Zheng, L., Baumann, U., Reymond, J., 2004. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Research* 32 (14), 1–5.

Chemical Biology

RL Lundblad, Chapel Hill, NC, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Biomacromolecule A polymer associated with biological systems. Examples include nucleic acids, proteins, and polysaccharides.

Bioorthogonal modification The process of covalently modifying a biomacromolecule in a complex biological environment such as cell lysate or intact cell in a selective manner without a negative effect on normal cell function.

Click Chemistry A general term describing a simple chemical reaction joining two molecules; the term is used most frequently to describe alkyne-azide cycloaddition.

Metabolic labeling An experimental approach where the metabolic pathways inside a cell are used to label a protein

or other macromolecule with a label or reactive groups.

Examples of metabolic labeling include stable isotope labeling of proteins or the insertion of a reactive group into a protein during the synthetic process. Much earlier examples would include the use of radioisotopes to trace metabolic pathways.

Structural biology The study of macromolecular structure and supramolecular structure. Objects of interest range from proteins to organelles such as mitochondria and ribosomes. Techniques include computational biology, crystallography, nuclear magnetic resonance, and electron microscopy.

Introduction

The term chemical biology was first used in a lecture by Professor J.B. Leathes before the Royal College of Physicians in 1930 (Leathes, 1930). Professor Leathes made the case that the role of chemistry in biology was seen by William Harvey and others in work stemming discovery of the circulation of the blood. Some years later, Maurice Florkin used the term chemical biology in the title of a rather nice book in 1960 (Florkin, 1960). An historical perspective of chemical biology can be obtained from a treatise on the inhibition of enzymes by chemicals (Webb, 1963–66). One could make an argument that Webb's book (Webb, 1963–66) which was entitled *Enzymes and Metabolic Inhibitors* should have contained the term chemical biology in the title. Iodoacetate was used to define the early processes of glycolysis by inhibition of zymase, the enzyme responsible for the breakdown of glucose (Webb, 1966). Iodoacetate is still used as a metabolic poison to cause energy depletion (Jackson *et al.*, 2014).

Chemical biology has evolved into a term referring to the study of biochemical processes occurring inside of a live cell. It is important to accept that there is considerable overlap between chemical biology and biological chemistry such that *BMC Chemical Biology* is incorporated as section within *BMC Biochemistry*. Another example is provided by the transition of *Journal of Peptide Research* to *Chemical Biology and Drug Design*. Chemical biology is a broad area of investigation as demonstrated by the diversity of titles in the various journals in this area. This article has focused on topics within chemical biology that distinguish chemical biology as a separate discipline from other biological sciences.

Imaging

In a sense, chemical biology is not so much a specific discipline such as biochemistry, chemistry, or physics but more of a framework to use current tools to ask interesting questions. If you define chemical biology as an approach to 'look' at the

chemistry of the living cell, then methods by which one could see inside the cell at the molecular level would be most useful. This would be a form of molecular imaging directed at the cellular level. Molecular imaging, though, is a term that has seen most use in application for diagnostic radiology for whole animal studies (Rudin, 2009). Molecular imaging in chemical biology utilizes major advances in optical microscopy (Eggrling and Heilemann, 2014) as well as developments in spectroscopy including near-infrared fluorescence (Hilderbrand and Weisleder, 2010) and Raman microscopy. The spectral characteristics of the alkyne functional group show strong Raman scattering in a region where cellular constituent have null scattering (Yamakoshi *et al.*, 2012). Alkyne-labeled coenzyme Q analogues were measured inside living cells using Raman spectroscopy. Raman spectroscopy has also been used to image the distribution of tyrosine kinase inhibitors such as imatinib and nilotinib inside a living cell (Fu *et al.*, 2014).

Investigators (Korbas *et al.*, 2011) used synchrotron-based X-ray fluorescence imaging to show different cell and tissue distribution of inorganic and organic mercury in zebrafish embryos. The imaging can be considered an extension of earlier work on histochemistry, immunohistochemistry, and enzyme histochemistry but with more sophisticated optics and reagents. An example of such progress can be seen from the study on the imaging of peroxynitrite in cells with a fluorescent probe (Peng *et al.*, 2014). Reaction of peroxynitrite with HKGreen-4 which is not fluorescent results in the loss of aryl substituent and the formation of *N*-methylrhodol which is fluorescent (Figure 1) permitting the identification of peroxynitrite in cells. The design of increasingly complex substrates has yielded substrates with useful specificity such as shown for caspase-3 (Figure 1; Vickers *et al.*, 2014) permitting intracellular localization.

Metabolic Labeling

Metabolic labeling is a technique which involves the development of reactive substrates which can act as reporter groups

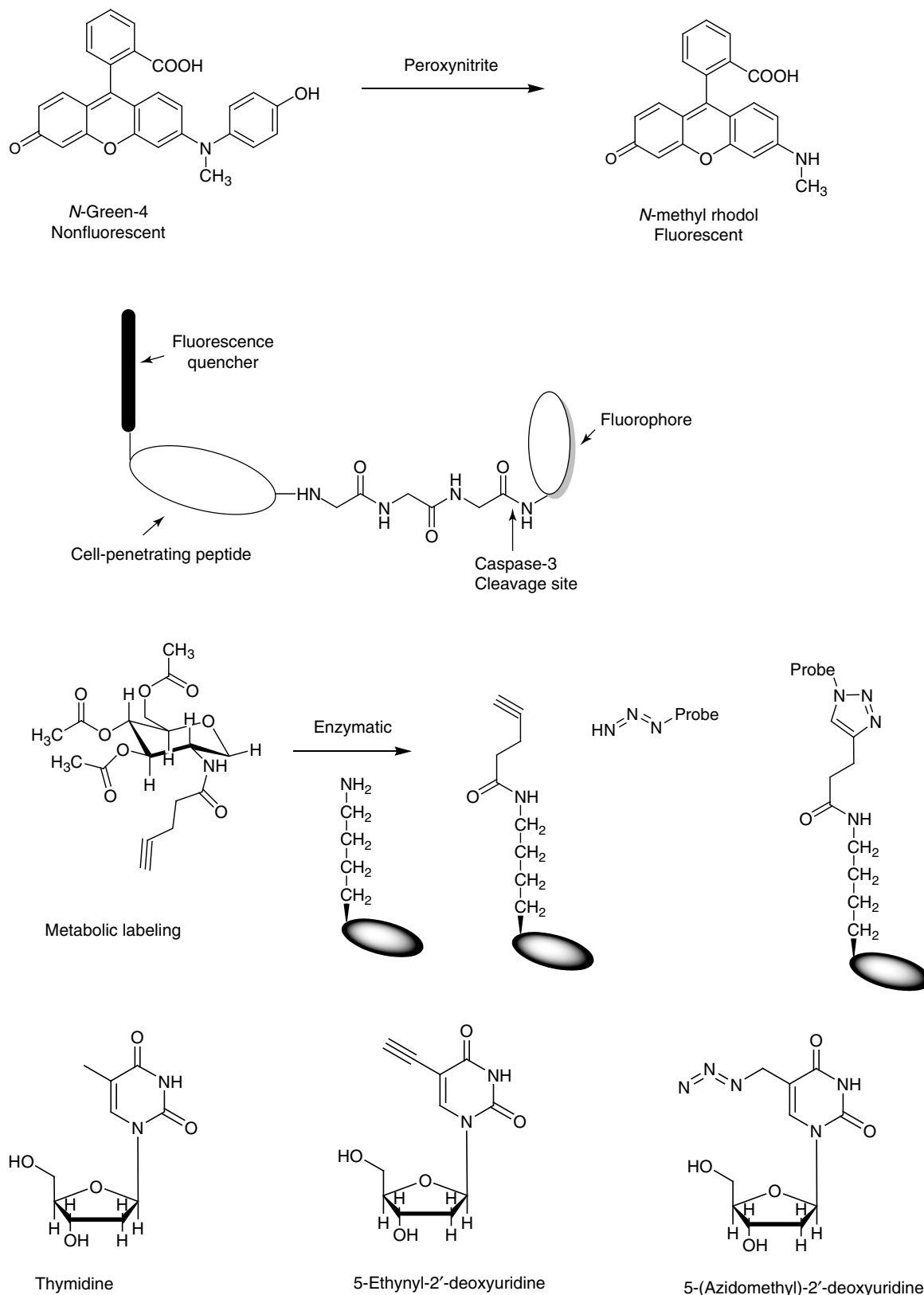


Figure 1 Some technical approaches in chemical biology. At the top is shown the visualization of peroxynitrite in the cell by reaction with *N*-Green-4. Below that is a fluorescence resonance energy transfer (FRET) substrate used for the intracellular visualization of Caspase-3. At the bottom is shown examples of metabolic labeling with an azide-hexamine used for labeling glycans and thymidine analogous for nucleic acids.

(Zurkin *et al.*, 1977) inside the living cell. The key to this approach is the development of a substrate analogue which contains a functional group which is unreactive to cellular constituents but which can be tagged with a reporter groups such as a fluorophore. An alkyne group is an example of such a functional group. *N*-acetyl hexosamines are precursor of glycoproteins and the acetyl function may also enter the pathway for carbohydrate metabolism. Although acetyl-coenzyme A derived from glycolysis serves as the acetyl group for protein acetylation, there is some evidence to suggest the acetyl function from *N*-acetyl hexosamines can also be a source of acetyl groups for protein acetylation. The *N*-acetylation of proteins is a posttranslational mechanism as important as a regulatory process in cell biology. The use of 1-deoxy-*N*-pentynyl glucosamine (Figure 1) suggested that there was a transfer of the pentynyl group to proteins (Zaro *et al.*, 2014). The pentynyl group could be labeled by an azide probe via click chemistry permitting the attachment of a fluorophore. A peracetylated glucosamine derivative was used to facilitate transport into the cell; the *O*-acetyl groups were removed by cytoplasmic esterases. Use of *N*-pentynyl glucosamine resulted in incorporation into *O*-glycans as well as the putative transfer of the pentynyl group to lysine residues on proteins. The use of the 1-deoxy-*N*-pentynyl glucosamine derivative resulted in the labeling of 99 proteins as compared to 367 proteins labeled

with the *N*-pentynyl glucosamine; 46 proteins were labeled with either reagent. This work supports the concept that the acetyl function from *N*-acetyl hexosamines can be transferred to a protein.

Metabolic labeling can be used for other biopolymers such as nucleic acids or proteins. DNA has been labeled with 5-ethynyl-2'-deoxyuridine (Lulai *et al.*, 2014) or 5-azidomethyl-2'-deoxyuridine (Neet and Luedtke, 2014); both of these compounds are analogues of thymidine (Figure 1) and can be coupled with fluorescent probes via click chemistry to permit identification. An approach related to metabolic labeling is the incorporation of unnatural amino acids with reactive groups (Liu and Schultz, 2010). The use of unnatural amino acids such as *p*-azidophenylalanine (Tian *et al.*, 2014) or alkene lysine analogues (Torres-Kolbus *et al.*, 2014) yield proteins with reactive site groups available for modification with probes.

Affinity Labeling

Both bioorthogonal labeling and activity-based probes are based, in part, on earlier work on affinity labeling of proteins (Wofsy *et al.*, 1962). The principle here is the specific binding of molecule to a protein which could, for example, be an enzyme or a receptor. The affinity label contains two parts: an

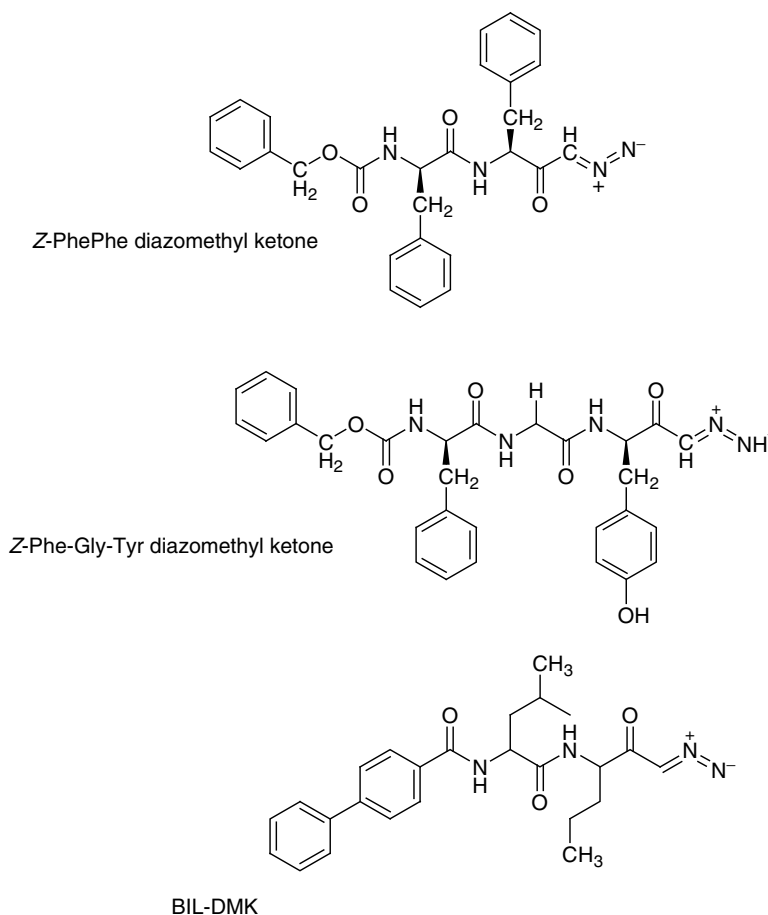


Figure 2 The evolution of affinity labels for sulfhydryl proteases.

affinity group which would bind to a specific site on a protein or other biopolymer and a reactive group such as an α -haloketo function. The reader is directed to a review on the design and use of affinity labels for additional information (Plapp, 1982). The reactive group should be sufficiently unreactive that reaction with a functional group such that proximity secondary to binding is the dominant factor. For example, consider the reaction of the cysteine residue in the enzyme active site of sulfhydryl proteinase with a peptide diazomethyl ketone where the rate of reaction can be as much as 10^{10} faster than reaction with free cysteine (Shaw and Green, 1981; Figure 2). Subsequent work showed that more complex peptidyl diazoethyl ketones could inactivate proteases (Savory *et al.*, 1993) but the reactions were quite slow. More recent work has used these observations to develop activity-based probes for identifying intracellular cathepsin activity (Falgout *et al.*, 2004). These investigators showed that BIL-DMK (biphenyl-leucine-diazomethylketone) inactivated purified cathepsins reacted with cathepsins in intact HepG2 cells. The insertion of a radiolabel into the reagent allowed identification of the modified proteins. Later work with this reagent demonstrated the distribution of the label in leukocytes populations in whole blood (Veilleux *et al.*, 2011).

Activity-Based Probes

Activity-based protein probes (Serim *et al.*, 2012) is a chemical biology approach based on covalent modifying extracellular and intracellular enzymes with probes which can be used to identify and isolate the active enzymes. The initial studies were performed on proteolytic enzymes and evolved from early work on affinity labels. The concept of activity-based probes depends on the relative inactivity of the zymogen or precursor forms of a proteolytic enzyme and the importance of functional group present at an enzyme active site. Diisopropylphosphorofluoridate (DFP) is a classic inhibitor of serine proteases such as trypsin and chymotrypsin. Liu *et al.* (1999) prepared a biotinylated derivative of DFP (Figure 3) which could be used to label and isolate serine proteases in crude tissue extracts. More sophisticated probes are diphenylphosphonate derivatives (Pan *et al.*, 2006) and di(chlorophenyl) phosphonate derivatives (Guarino *et al.*, 2014).

Activity-based probes have been developed for other enzymes such as kinases. Imbruvica[®] is a drug developed for the treatment of chronic lymphocytic leukemia (CLL) with the target being the covalent modification of Bruton's tyrosine kinase. Lanning *et al.* (2014) developed an alkyne derivative of ibrutinib (Imbruvica[®]) which could be used to detect Bruton's tyrosine kinase in cells (Figure 3). However, the first alkyne derivative was somewhat promiscuous as there are other protein sulfhydryl groups of sufficient reactivity and exposure to react with the Michael acceptor function. Decreasing the reactivity of the Michael acceptor improved the specificity of modification as there are undesirable side reactions observed with Imbruvica[®] (Levade *et al.*, 2014). This work provides an example of the use of chemical biology in drug development where quality is improved as a result of potential reduction of adverse reactions.

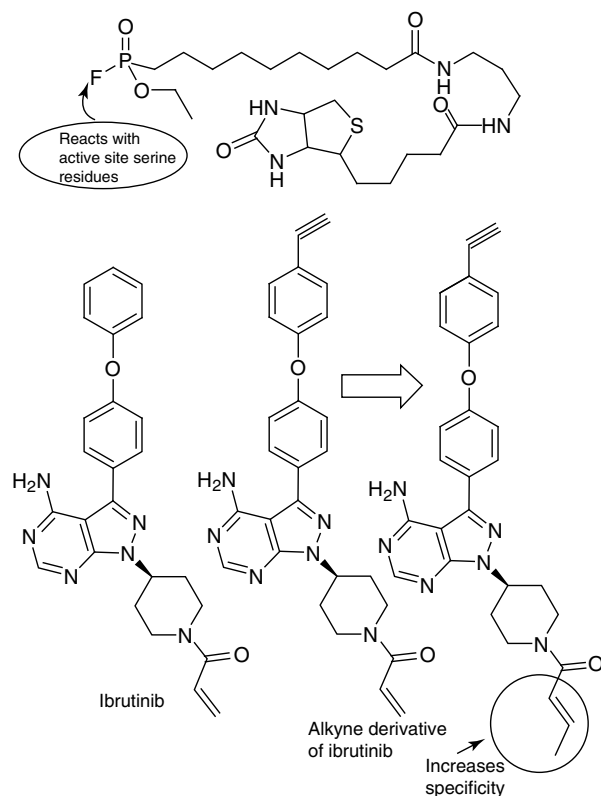


Figure 3 Activity-based probes. At the top is shown an activity-based probe for a serine protease while at the bottom is an activity-based probe for a tyrosine kinase.

Bioorthogonal Labeling

The development of the concept of bioorthogonal modification (Hang *et al.*, 2003; Bertozzi and Wu, 2013) has also been important in the maturation of chemical biology. A recent review (Patterson *et al.*, 2014) describes a number of chemistries available for bioorthogonal modification. The various techniques in chemical biology do overlap as the seminal paper on bioorthogonal labeling (Hang *et al.*, 2003) used metabolic labeling. The various tools in chemical biology do allow a study of proteins and other macromolecules inside the cell. The study of polysaccharides, both inside and outside the cell has been a particular challenge. It is therefore of no small importance that the signature paper (Hang *et al.*, 2003) on the use of bioorthogonal chemistry in cell biology involved the labeling of a O-linked mucin chain with an azide function (Figure 4) which could be subsequently tagged with a label such as a fluorophore. *N*-azidoacetylglucosamine was incorporated into the mucin chain in place of *N*-acetylglucosamine. Over the decade since this seminal work, there have been advances in chemistry with the use of a carbamate-linked cyclopropene function for a Diels-Alder condensation with a tag that, for example, contains biotin (Figure 4). The use of peracetylated derivatives enhances transport into the cell. This specific bioorthogonal labeling of polysaccharides permits imaging of glycans on cell surfaces (Stairs *et al.*, 2013); this study used orthogonal labels such as

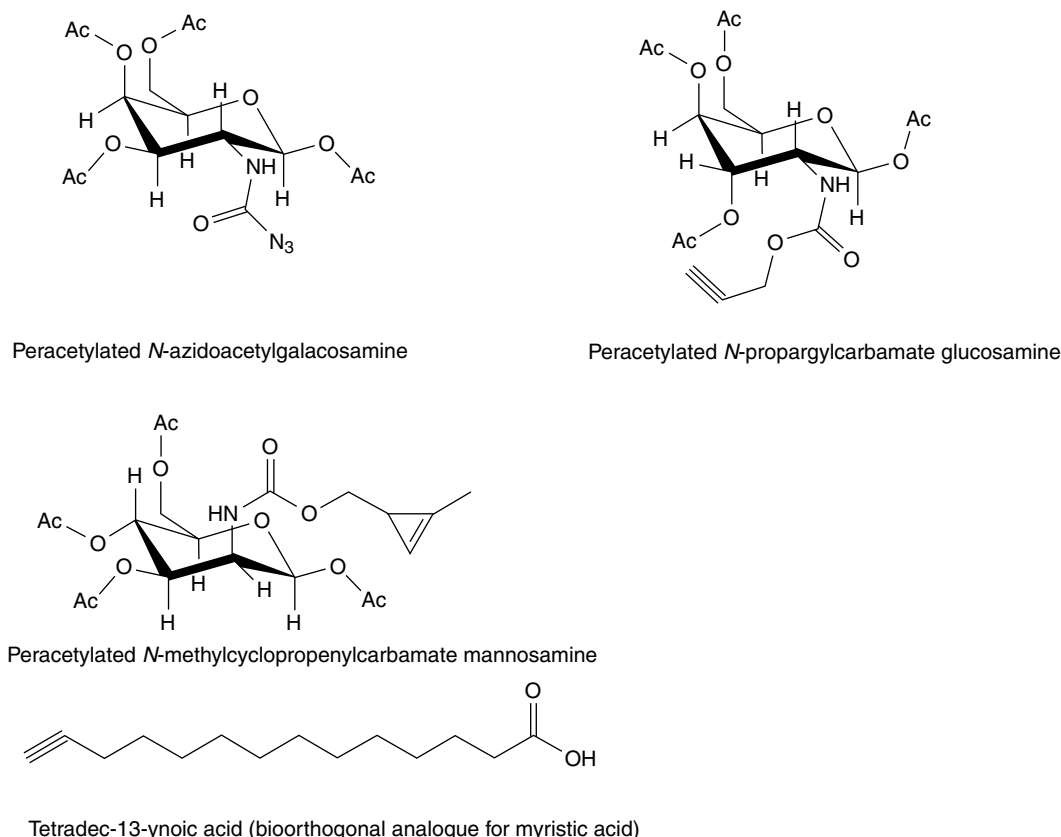


Figure 4 Some examples of bioorthogonal reagents.

isonitrile and azide or cyclopropene and azide. A recent novel approach, described as ligand-directed tosyl chemistry (Tsukiji and Hamachi, 2014), permits the attachment of a probe such as fluorophore with a single reagent.

ω -Alkyn fatty acid derivatives are useful for localizing intracellular lipid-binding proteins (Raghavan *et al.*, 2008; Figure 4). An alkyne analogue of myristic acid has been prepared and has been demonstrated to couple to the N-terminus of a protein as coenzyme A intermediate (Host *et al.*, 2008; Thinon *et al.*, 2014). Bioorthogonal labeling of RNA has been accomplished by the incorporation of nucleobase analogues (Lulai *et al.*, 2014; Neet and Luedtke, 2014) or by post-synthetic modification (McDonald *et al.*, 2014).

Chemical Biology and Drug Development

Although the identification of the target of a drug is an essential part of the current drug development process, this was not so with earlier drugs such as aspirin where various therapeutic effects were discovered through use. The effect of aspirin on platelet aggregation was known before the discovery of the effect of aspirin on thromboxanes. Chemical biology now allows the study of the mechanism of action of drugs like aspirin in a rational manner. It is then of no small interest that there are chemical biology studies designed to define the mechanism of action of aspirin (2-(acetyloxy)-benzoic acid;

acetylsalicylic acid). Bateman *et al.* (2013) used an alkyne-aspirin derivative to identify aspirin targets in intact cells. Both the salicylate moiety and the 'active' acetyl group were known to be involved in pharmacological effects in inflammation and thrombosis, respectively. Recent studies have suggested that aspirin may have a function in reducing cancer mortality (Fraser *et al.*, 2014). Bateman *et al.* (2013) used 'click chemistry' to label intracellular protein targets of aspirin with a fluorescent tag (Figure 5). These investigators were able to demonstrate the modification of 120 proteins with this probe. Identification of the modified proteins may provide insight into the therapeutic targets of aspirin in cancer (Alfonso *et al.*, 2014). The work by Bateman *et al.* (2013) is another example of a bioorthogonal modification.

Chemical biology continues to develop as a sophisticated approach for the study of cellular physiology. The increasing sophistication of the structure of the probes will increase specificity of labeling. The use of Raman spectroscopy and near-infrared spectroscopy will eliminate the necessity of attaching a reporter group. The most exciting potential is the elucidation of the mechanism of action of therapeutics. The aspirin story is discussed above and it is likely that the bioorthogonal labeling of target protein by aspirin-based reagents will elucidate the putative role of aspirin in cancer. Recalling that aspirin developed from the use of willow bark as an analgesic and anti-inflammatory therapeutic, it would be useful to prepare bioorthogonal reagents based on emerging

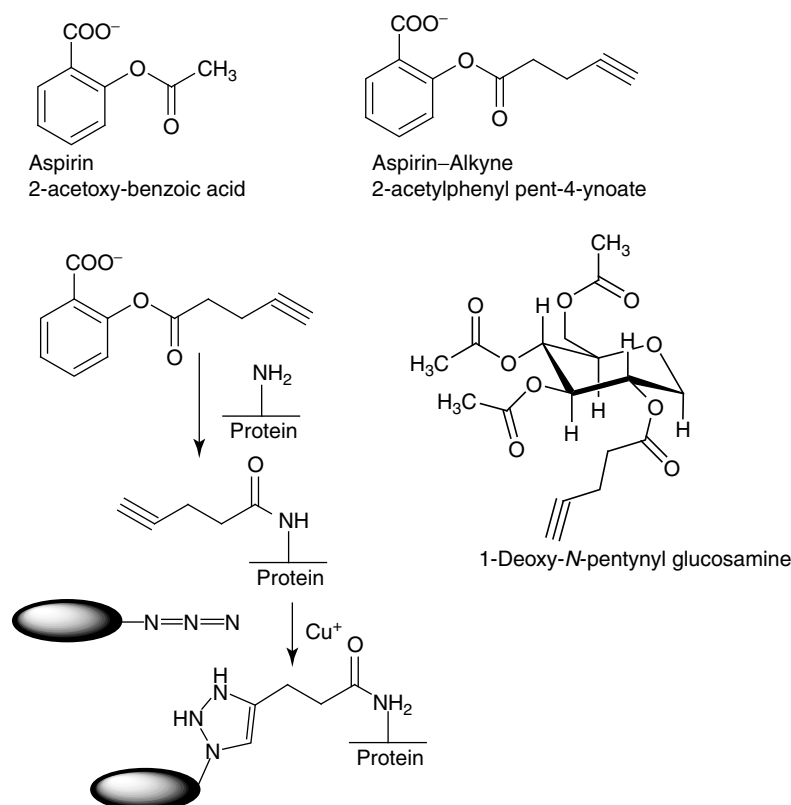


Figure 5 The development of a bioorthogonal reagent based on aspirin.

herbal therapeutics (Gannier *et al.*, 2007) to determine their target(s).

See also: Molecular Principles, Components, Technology, and Concepts: Proteins: Drug Design; Site-Directed Mutagenesis

References

- Alfonso, L., Ai, G., Spitale, R.C., Bhat, G.J., 2014. Molecular targets of aspirin and cancer prevention. *British Journal of Cancer* 111, 61–67.
- Bateman, L.A., Zaro, B.W., Miller, S.M., Pratt, M.R., 2013. An alkyne-aspirin chemical reporter for the detection of aspirin-dependent protein modification in living cells. *Journal of the American Chemical Society* 135, 14568–14573.
- Bertozzi, C.R., Wu, P., 2013. *In vivo* chemistry. *Current Opinion on Chemical Biology* 17, 717–718.
- Eggrling, C., Heilemann, M., 2014. Editorial overview: Molecular imaging. *Current Opinion on Chemical Biology* 20, 6–7.
- Falgout, J.P., Black, W.C., Cromlish, W., *et al.*, 2004. An activity-based probe for the determination of cysteine cathepsin protease activities in whole cells. *Analytical Biochemistry* 335, 218–227.
- Florkin, M., 1960. *Unity and Diversity in Biochemistry. An Introduction to Chemical Biology* (T. Wood, Trans.). Oxford: Pergamon Press.
- Fraser, D.M., Sullivan, F.M., Thompson, A.M., McCowan, C., 2014. Aspirin use and survival after the diagnosis of breast cancer: A population-based cohort study. *British Journal of Cancer* 111, 623–627.
- Fu, D., Zhou, J., Zhu, W.S., *et al.*, 2014. Imaging the intracellular distribution of tyrosine kinase inhibitors in living cells with quantitative hyperspectral stimulated Raman scattering. *Nature Chemistry* 6, 615–623.
- Gannier, J.J., van Tulder, M.W., Berman, B., Bombardier, C., 2007. Herbal medicine for low back: A Cochrane review. *Spine* 32, 82–92.
- Guarino, C., Legowska, M., Epinette, C., *et al.*, 2014. New selective peptidyl di (chlorophenyl)-phosphonate esters to visualize and block neutrophil proteinase 3 in human diseases. *Journal of Biological Chemistry* 289 (46), 31777–31791.
- Hang, H.L., Yu, C., Kato, D.J., Bertozzi, C.R., 2003. A metabolic labeling approach toward proteomic analysis of mucin-type O-linked glycosylation. *Proceedings of the National Academy of Sciences of the United States of America* 100, 14848–14851.
- Hilderbrand, S.A., Weissleder, R., 2010. Near-infrared fluorescence: Application to *in vivo* molecular imaging. *Current Opinion on Chemical Biology* 14, 71–79.
- Host, W.P., Wickramasinghe, S.R., Leatherbarrow, R.J., Tate, E.W., 2008. *N*-Myrisoyl transferase-mediated protein labelling *in vivo*. *Organic & Biomolecular Chemistry* 6, 2308–2315.
- Jackson, E.K., Cheng, D., Mi, Z., Gillespie, D.G., 2014. Guanosine regulates adenosine levels in the kidney. *Physiology Reports* 2 (5), e12028.
- Korbas, M., MacDonald, T.C., Pickering, I.J., George, G.N., Krone, P.H., 2011. Chemical form matters: Differential accumulation of mercury following inorganic and organic mercury exposures in zebrafish larvae. *ACS Chemical Biology* 7, 411–420.
- Lanning, B.R., Whitby, L.R., Dix, M.M., *et al.*, 2014. A road map to evaluate the proteome-wide selectivity of covalent kinase inhibitors. *Nature Chemical Biology* 10, 760–767.
- Leathes, J.B., 1930. The Harveian Oration on the birth of chemical biology. *British Medical Journal* 2 (3642), 671–676.
- Levade, M., David, E., Garcia, C., *et al.*, 2014. Ibrutinib treatment affects collagen and von Willebrand factor-dependent platelet functions. *Blood* 124 (26), 3991–3995.
- Liu, C.C., Schultz, P.G., 2010. Adding new chemistries to the genetic code. *Annual Review of Biochemistry* 79, 413–444.
- Liu, Y., Patricelli, M.P., Cravatt, B.F., 1999. Activity-based protein profiling: The serine hydrolases. *Proceedings of the National Academies of Science of the United States of America* 96, 14694–14699.
- Lulai, E.C., Neubauer, J.D., Suttle, J.C., 2014. Kinetics and localization of wound-induced DNA biosynthesis in potato tuber. *Journal of Plant Physiology* 171, 1571–1575.

- McDonald, R.I., Guilinger, J.P., Mukherji, S., *et al.*, 2014. Electrophilic activity-based RNA probes reveal a self-alkylating RNA for RNA labeling. *Nature Chemical Biology* 10, 1049–1054.
- Neet, A.B., Luedtke, N.W., 2014. An azide-modified nucleoside for metabolic labeling of DNA. *ChemBioChem* 15, 789–793.
- Pan, Z., Jeffery, D.A., Chehade, K., *et al.*, 2006. Development of activity-based probes for trypsin-family serine proteases. *Bioorganic and Medicinal Chemistry Letters* 16, 2882–2885.
- Patterson, D.M., Nazarova, L.A., Prescher, J.A., 2014. Finding the right (bioorthogonal) chemistry. *ACS Chemical Biology* 9, 592–605.
- Peng, T., Wong, N.-K., Chen, X., *et al.*, 2014. Molecular imaging of peroxynitrite with HKGreen-4 in live cells and tissues. *Journal of the American Chemical Society* 136, 11728–11734.
- Plapp, B.V., 1982. Application of affinity labeling for studying structure and function of enzymes. *Methods in Enzymology* 87, 469–499.
- Raghavan, A., Charron, G., Flexner, J., Hang, H.C., 2008. Chemical probes for profiling fatty-acid associated proteins in living cells. *Bioorganic & Medicinal Chemistry Letters* 18, 5982–5986.
- Rudin, M., 2009. Noninvasive structural, functional, and molecular imaging in drug development. *Current Opinion in Chemical Biology* 13, 360–371.
- Savory, P.J., Djballah, H., Anglikar, H., Shaw, E., Rivett, A.J., 1993. Reaction of proteasomes with peptidylchloromethanes and peptidyl diazomethanes. *Biochemical Journal* 296, 601–605.
- Serim, S., Haedke, U., Verhelst, S.H.L., 2012. Activity-based probes for the study of proteases: Recent advances and developments. *ChemMedChem* 7, 1146–1159.
- Shaw, E., Green, G.D.J., 1981. Inactivation of thiol proteases with peptidyl chloromethyl ketones. *Methods in Enzymology* 80, 820–828.
- Stairs, S., Neves, A.A., Stöckmann, H., 2013. Metabolic glycan imaging by isonitrile-tetrazine click chemistry. *ChemBioChem* 14, 1063–1067.
- Thinon, E., Serwa, R.A., Broncel, M., *et al.*, 2014. Global profiling of co- and post-translationally *N*-myristoylated proteomes in human cells. *Nature Communications* 5, 4919.
- Tian, H., Naganathan, S., Kazami, M.A., *et al.*, 2014. Bioorthogonal fluorescent labeling of functional G-protein-coupled receptors. *ChemBioChem* 15, 1820–1829.
- Torres-Kolbus, J., Chou, C., Liu, J., Deiters, A., 2014. Synthesis of non-linear protein dimers through a genetically encoded thiol-ene reaction. *PLoS One* 9, e105467.
- Tsukiji, S., Hamachi, I., 2014. Ligand-directed tosyl chemistry for *in situ* native protein labeling and engineering in living systems: From basic properties to applications. *Current Opinion in Chemical Biology* 21, 136–143.
- Veilleux, A., Black, W.C., Gauthier, J.Y., *et al.*, 2011. Probing cathepsin S activity in whole blood by the activity-based probe BIL-DMK: Cellular distribution in human leukocyte populations and evidence of diurnal modulation. *Analytical Biochemistry* 411, 43–49.
- Vickers, C.J., González-Páez, G.E., Wolan, D.W., 2014. Discovery of a highly selective caspase-3 substrate for imaging live cells. *ACS Chemical Biology* 9, 2199–2203.
- Webb, J.L., 1963–66. *Enzyme and Metabolic Inhibitors*. New York, NY: Academic Press.
- Webb, J.L., 1966. Iodoacetate and iodoacetamide, in *Enzyme and Metabolic Inhibitors*, Vol. III. New York, NY: Academic Press, pp. 1–283 (Chapter 1).
- Wofsy, L., Metzger, H., Singer, S.J., 1962. Affinity labeling — a general method for labeling the affinity sites of antibody and enzyme molecules. *Biochemistry* 1, 1031–1039.
- Yamakoshi, H., Dodo, K., Palonpon, A., *et al.*, 2012. Alkyne-tag Raman imaging for visualization of mobile small molecules in live cells. *Journal of the American Chemical Society* 134, 20681–20689.
- Zaro, B.W., Chuh, K.N., Pratt, M.R., 2014. Chemical reporter for visualizing metabolic cross-talk between carbohydrate metabolism and protein modification. *ACS Chemical Biology* 9, 1991–1996.
- Zurkin, R.S., Hartig, P.R., Koshland Jr., D.E., 1977. Use of a distant reporter group as evidenced for a conformation change in a sensory receptor. *Proceedings of the National Academy of Sciences of the United States of America* 74, 1932–1936.
- Journal of Chemical Biology. Berlin Heidelberg: Springer-Verlag.
- Nature Chemical Biology. New York, NY: Nature.
- Antonio, K.A., Schultz, Z.D., 2014. Advances in biomedicine Raman microscopy. *Analytical Chemistry* 86 (SI), 30–46.
- Best, M.D., 2009. Click chemistry and bioorthogonal reactions: Unprecedented selectivity in the labeling of biological molecules. *Biochemistry* 48, 6571–6584.
- Best, M.D., Rowland, M.M., Bostic, H., 2011. Exploiting bioorthogonal chemistry to elucidate protein-lipid binding interactions and other biological roles of phospholipids. *Accounts of Chemical Research* 44, 686–698.
- Boons, G.-J., 2010. Bioorthogonal chemical reporter methodology for visualization, isolation and analysis of glycoconjugates. *Carbohydrate Chemistry* 36, 152–167.
- Chen, X., Henschke, L., Wu, Q., *et al.*, 2013. Site-selective azide incorporation into endogenous RNA via a “chemistry” approach. *Organic and Biomolecular Chemistry* 11, 353–361.
- Chen, Z., Jing, C., Gallagher, S.S., Sheetz, M.P., Cornish, V.W., 2012. Second-generation covalent TMP-tags for live cell imaging. *Journal of the American Chemical Society* 134, 13692–13699.
- Dobson, C.M., Gerrard, J.A., Pratt, A.J., 2001. *Foundations of Chemical Biology*. Oxford: Oxford University Press.
- Erdmann, V.A., Markiewicz, W.T., Barciszewski, J. (Eds.), 2014. *Chemical Biology of Nucleic Acids*. Heidelberg: Springer-Verlag.
- Fraser-Reid, B.O., Tatsuta, K., Thiem, J. (Eds.), 2001. *Glycoscience: Chemistry and Chemical Biology*, I–III. Berlin/Heidelberg: Springer.
- Geacintov, N.E., Broyde, S. (Eds.), 2010. *The Chemical Biology of DNA Damage*. Weinheim: Wiley.
- King, M., Wagner, A., 2014. Developments in the field of bioorthogonal bond forming reactions—past and present trends. *Bioconjugate Chemistry* 25, 825–839.
- Kobayashi, H., Longmire, M.R., Ogawa, M., Choyke, P.L., 2011. Rational chemical design of the next generation of molecular imaging probes based on physics and biology: Mixing modalities, colors and signals. *Chemical Society Reviews* 40, 4626–4648.
- Kuerschner, L., Thiele, C., 2014. Multiple bonds for the lipid interest. *Biochimica Biophysica Acta* 1841, 1031–1037.
- Lavis, L.D., 2011. Histochemistry: Live and in color. *Journal of Histochemistry and Cytochemistry* 59, 139–145.
- McFarlane, H.E., Döring, A., Persson, S., 2014. The cell biology of cellulose synthesis. *Annual Review of Plant Biology* 65, 69–94.
- Meier-Ruge, W.A., Bruder, E., 2008. Current concepts of enzyme histochemistry in modern pathology. *Pathobiology* 75, 233–243.
- Neidle, S. (Ed.), 2009. *Cancer Drug Design and Discovery*. Amsterdam: Elsevier.
- O'Hare, H.M., Johnsson, K., Gautier, A., 2007. Chemical probes shed light on protein function. *Current Opinion in Structural Biology* 17, 488–494.
- Osada, H. (Ed.), 2009. *Protein Targeting with Small Molecules*. Hoboken: Wiley.
- Palonpon, A.F., Sodeoka, M., Fujita, K., 2013. Molecular imaging of live cells by Raman microscopy. *Current Opinion in Chemical Biology* 17, 708–715.
- Patterson, D.M., Jones, K.A., Preschner, J.A., 2014. Improved cyclopropene reporters for probing protein glycosylation. *Molecular BioSystems* 10, 1693–1697.
- Schulz, D., Rentmeister, A., 2014. Current approach for RNA labeling *in vitro* and in cells based on click reactions. *ChemBioChem* 15, 2342–2347.
- Shioh, P., Siegrist, M.S., Culler, A.J., Bertozzi, C.R., 2014. Imaging bacterial peptidoglycan with near-infrared fluorogenic azide probes. *Proceedings of the National Academy of Sciences of the United States of America* 111, 5456–5461.
- Späte, A.-K., Buszkamp, H., Niederwieser, A., *et al.*, 2014. Rapid labeling of metabolically engineered cell-surface glycoconjugates with a carbamate-linked cyclopropene reporter. *Bioconjugate Chemistry* 25, 147–154.
- Späte, A.-K., Scharf, V.F., Häfner, J., *et al.*, 2014. Expanding the scope of cyclopropene reporters for the detection of metabolically engineered glycoproteins by Diels-Alder reaction. *Beilstein Journal of Organic Chemistry* 10, 2235–2242.
- Tautermann, C.C., 2014. GPCR structures in drug design, emerging opportunities with new structures. *Bioorganic and Medicinal Chemistry Letters* 24, 4073–4079.
- Thome, N., Ingelse, J., Auld, D.S., 2010. Illuminating insights into firefly luciferase and other bioluminescent reagents used in chemical biology. *Chemical Biology* 17, 646–657.
- Wang, Z., Ding, X., Li, S., Shi, J., Li, Y., 2014. Engineered fluorescent tags for *in vivo* protein labeling. *RSC Advances* 4, 7235–7245.
- Warner, J.B., Muthusamy, A.K., Petersson, E.J., 2014. Specific modulation of protein activity by using a bioorthogonal reaction. *ChemBioChem* 15, 2508–2514.
- Whyte, G.F., Vilar, R., Woscholski, R., 2013. Molecular recognition with boronic acids — applications in chemical biology. *Journal of Chemical Biology* 6, 161–174.
- Xie, R., Hong, S., Chen, X., 2013. Cell-selective metabolic labeling of biomolecules with bioorthogonal functionalities. *Current Opinion in Chemical Biology* 17, 747–752.

Further Reading

ACS Chemical Biology. Washington, D.C: American Chemical Society.

BMC Biochemistry. London: BioMed Central.

ChemBioChem. A European Journal of Chemical Biology. Weinheim: Wiley.

Chemical Biology and Drug Design. Hoboken: Wiley-Blackwell.

Current Opinion in Chemical Biology. Amsterdam: Elsevier.

Drug Design

RL Lundblad, Chapel Hill, NC, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Active pharmaceutical ingredient (API) The final product in the manufacturing process prior to formulation. API may also be known as the active ingredient.

Click chemistry A general term describing a simple chemical reaction joining two molecules; the term is used most frequently to describe alkyne–azide cycloadditions.

Comparative molecular field analysis (CoMFA) Extends the quantitative structure activity relationship (QSAR) analysis into three dimensions by establishing a relationship between biological activity and the electronic properties of the chemical in three dimensions; a quantitative consideration of electronic and physical shape of a chemical. CoMFA can be an analytical tool to rationalize observed activity or to predict the activity of a new compound.

Complementary determining region The specific region in the Fab domain (Fv) of an antibody molecule which recognizes the antigen. The complementary determining region may also be referred to as the paratope.

Drug substance The formulated form of the active pharmaceutical ingredient which is the material given to a patient. The drug substance may also be known as the drug product or dosage form.

Epitope That portion of an antigen which is recognized and bound by the paratope on the antibody. An epitope may be continuous in that it is a linear sequence of amino acids or other monomer unit or discontinuous in that the epitope is contributed by residues from disparate regions of a polymer which is the antigen.

Formulation The process by which the active pharmaceutical ingredient is converted into the drug substance and may involve the addition of active or inactive ingredients. The addition of active ingredients may result in a product being designated as a combination product. Inactive ingredients generally referred to as excipients are added to improve the properties of the drug substance.

Fusion protein A protein where two different proteins or two distinct protein domains are expressed as a single protein encoded by the joining of the respective DNA sequences. Fusion proteins may also be referred to as chimeric proteins.

Hammett equation A metric developed by Louis Hammett to predict reaction rates based on two constants, σ which is unique to a chemical group such as chloro group or

nitro group and ρ which was a descriptor of the nature of the reaction, solvent, and temperature.

High-throughput screening The process by which a lead compound may be developed into an active pharmaceutical ingredient. Large numbers of structural variations of the lead compounds are evaluated for biological activity either in silico or by evaluation of physical binding to a target using solid-phase (microplate) technology.

Hit A term used in high-throughput screening to describe a compound with improved properties which will be further developed in the search for a lead compound.

Lead compound A chemical compound which has the desired pharmacological effect but is unsuitable to serve as a drug. The lead compound provides the base structure which is then modified by organic chemistry or enzymes to provide the active pharmaceutical ingredient.

Log *P* A partition coefficient derived from the ratio of concentration of an organic compound, frequently a drug, in water and an organic phase, usually *n*-octanol. Log *P* increased with the lipophilic nature (hydrophobicity of the compound).

Payload A term used to describe the drug component of an antibody–drug bioconjugate.

Quality-by-design A systematic approach to the development of a product with specific predefined objectives. An emphasis is placed on the understanding of the product, understanding of the process, and understanding of process control based on sound science and quality management.

Quantitative structure activity relationship (QSAR) An analytical technique which uses the Hammett constant σ in combination with a constant, π , which is a measure of the lipophilicity of the compound, and steric component. The success of QSAR depends on the quality of the descriptors used for the model building. The initial work on 2D-QSAR was extended to 3D-QSAR by CoMFA which allows the consideration of shape (electronic and physical). Structure-based 3D-QSAR can be used when the structures of the ligand and the receptor are known while ligand-based 3D-QSAR is used when only ligand structure is known.

Structural biology The study of macromolecular structure and supramolecular structure. Objects of interest range from proteins to organelles such as mitochondria and ribosomes. Techniques include computational biology, crystallography, nuclear magnetic resonance, and electron microscopy.

Introduction

The term drug is used broadly within this article. Regulatory agencies such as the Food and Drug Administration (FDA) separate drugs from biologics on a somewhat arbitrary basis.

Drugs are usually low-molecular weight compounds prepared by organic chemical synthesis. Drugs are classified on the basis of solubility and permeability using the biopharmaceutical classification system (BCS) (Benet, 2013). Biologics include gene therapy, cellular therapy products, and proteins that are

isolated from a biological source such as blood plasma. Recombinant deoxyribonucleic acid (DNA) technology yields proteins that are therapeutics manufactured by fermentation in bacterial cells or cell culture in mammalian cells. More recently drugs have been coupled to monoclonal antibodies to create conjugate drugs most commonly known as antibody–drug conjugates. Drugs can also be linked to other biologicals to create bioconjugate drugs (Nielsen *et al.*, 2014). Fusion proteins which combine a therapeutic protein with a protein which influences pharmacokinetic behavior are also within the purview of drug design; an example is provided by the fusion of the carboxyl terminal region of the immunoglobulin protein (Fc domain) with factor VIII to provide a product with an extended half-life (Shapiro, 2013). Products obtained from recombinant DNA technology are considered to be drugs for regulatory purposes. The term biopharmaceutical can be used to describe both drugs and biologics.

The term target is used to describe the site of action of a drug. Drugs are most commonly directed at the inhibition of enzymes or receptors and more recently have been directed at protein-binding sites with the goal of disrupting protein–protein interactions. Another class of drugs are replacements rather than inhibitors; two examples are insulin and growth hormone, two peptide hormones, which were approved as drugs and not biologics. More recently, advances in structural biology enable the visualization of binding sites on proteins permitting computer design of drugs based on fit into a space on the target protein as was done for HIV-1 (human immunodeficiency virus type 1) protease and haloperidol (DesJarlais *et al.*, 1990). These investigators used a computer to make a negative image of the enzyme active site much in the way one would make a plaster cast of a footprint in the ground. A database was then used to find a chemical which would best fit in the computer-derived enzyme active site cavity. The program identified bromperidol as a good steric fit and haloperidol, an approved antipsychotic drug, that served as lead compound for further testing.

Implicit in the term drug is the understanding that this is a substance that will be used to effect a physiological change in the recipient. Legal drugs required a regulatory pathway that has little to do with the science employing concepts such as quality-by-design (QbD) (Rathore and Winkler, 2009; Martin-Moe *et al.*, 2011). It is imperative that good records are maintained from the start of a project such that rational development of the final drug product can be validated.

History of Drug Development

The current process of drug design and development (Figure 1) is a complex process with considerable regulatory oversight. However, over the course of history, drugs were first identified by the observation of the effects that various natural products had on sick and healthy people. The use of willow bark to relieve pain dates back to thousands of years and was shown to be dependent on its content of salicylates (Vane, 2000). Identification of salicylate as the active component of willow bark provided one of the first ‘lead compounds’ (salicylic acid) for drug development resulting in the discovery of aspirin (Sneider, 2000). The fortuitous discovery of useful

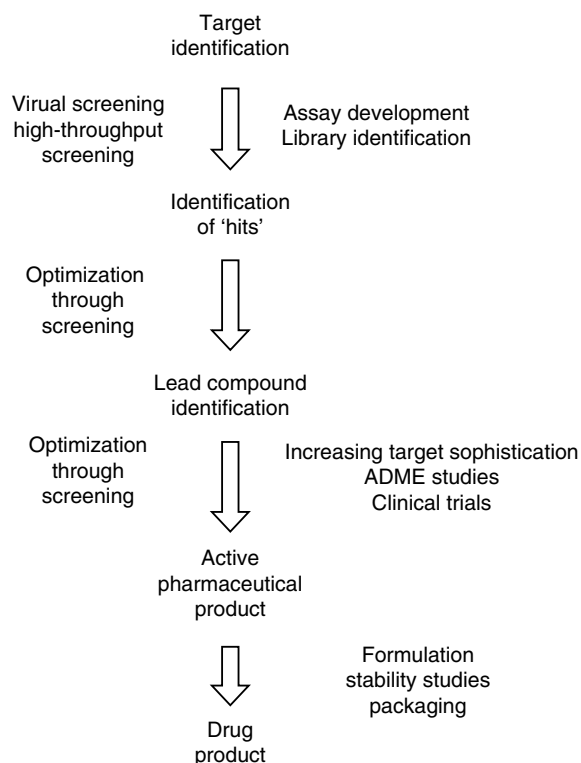


Figure 1 The process of drug development.

drugs remained valuable for the pharmaceutical industry over the centuries with more recent examples such as disulfiram (Kragh, 2008) and Viagra[®] (Campbell, 2000). In the case of disulfiram, it was careful consideration of laboratory data for unexpected results and with Viagra[®], it was the careful study of all results from an otherwise unsuccessful clinical study.

Evolution of Drug Design

Drug design is more complex than just the development of the pharmacologically active compound. A lead compound frequently needs to be chemically modified, for example, to eliminate undesirable side effects or improve pharmacokinetic behavior. In the case of aspirin (Figure 2), salicylic acid was considered the active ingredient for pain but caused gastric irritation; acetylation of the *o*-hydroxyl group provided the therapeutic derivative used today. While some aspects of the mechanism of action of aspirin are understood, such as acetylation of target proteins, some newly identified applications in cancer therapy remain poorly understood (Madhi *et al.*, 2006). Another example of the evolution of a natural product to a drug is provided by physostigmine (Figure 2). Physostigmine (eserine) was identified as the active component of the Calabar bean (Esere nut; *Physostigma venenosum*) found in tropical Africa. It was known to contain a potent poison. Work some 60 years ago (cf. Goldstein, 1945) demonstrated that physostigmine was a competitive inhibitor of acetylcholine esterase. The early work established physostigmine as a lead compound for a class of cholinesteric drugs resulting in novel

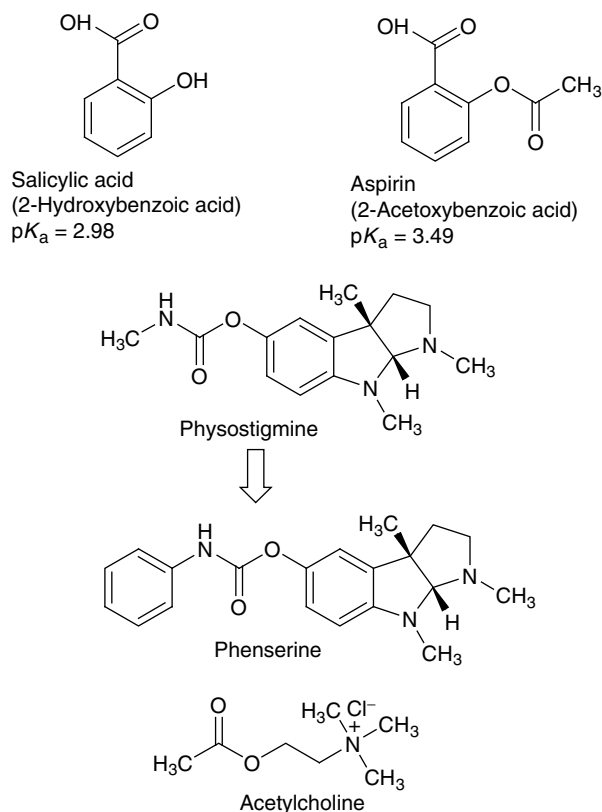


Figure 2 Lead compounds from nature to drug products.

derivatives such as phenserine which is under consideration for use in Alzheimer disease (Reale *et al.*, 2014; Figure 2). Another example is provided by thrombin, the final enzyme in the process of blood coagulation. Knowledge of the importance of guanidine group in substrates resulted in the preparation of a number of compounds which never became good drugs. Improvements in our understanding of thrombin have resulted in the development of new drugs which are in fact pro-drugs such as dabigatran etexilate which inhibit thrombin after hydrolysis by an esterase (Figure 3). As development has proceeded from lead compounds, specificity and avidity increased resulting in a more effective drug. Current drug development uses techniques of structural biology, high-throughput screening (HTS), and virtual screening, or computer modeling.

Structural Biology and Drug Development

The use of structural biology in the development of inhibitors for HIV protease was mentioned above. This is an example of rational drug design. Rational drug design is a method for identifying new drugs based on knowledge of the three-dimensional structure of the target protein (Hardy *et al.*, 1987). Knowledge of the three-dimensional structure provides information on the binding sites of substrates and inhibitors in the case of enzymes, agonists, and antagonists. Computer programs then optimize the binding characteristics of various

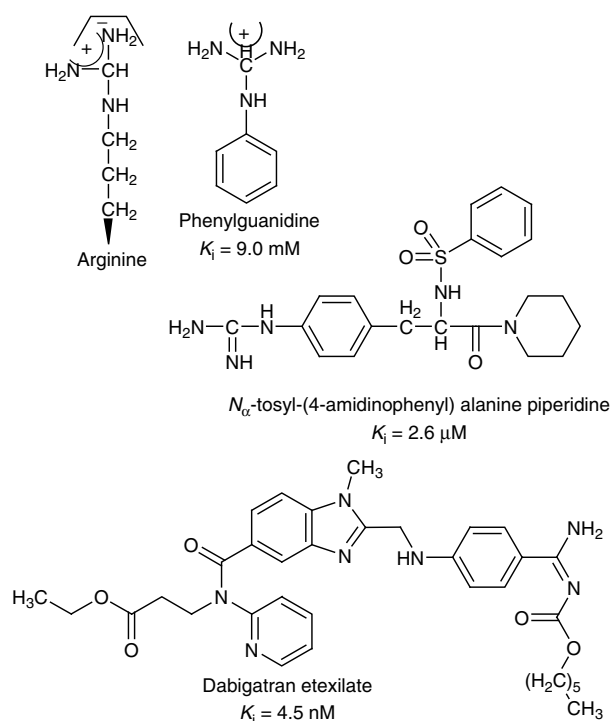


Figure 3 Aryl guanidines – from lead compounds to active pharmaceutical ingredient for use as thrombin inhibitors.

compounds yielding candidates which are then evaluated in *in vitro* assay systems for potential therapeutic effectiveness. Knowledge of the target structure is essential for docking studies; knowledge of protein structure is critical for protein modeling. Structural biology is frequently considered an activity within chemical biology.

HTS

HTS is another approach for identifying potential drug candidates. The success of HTS is dependent on the identification of meaningful assay systems. One example is the development of peptide substrates for the measurement of IgA1 protease activity in a 386-well microplate assay system (Choudary *et al.*, 2013). Microplate formats now range from 4 wells/plate to 1536 wells/plate; most HTS procedures use 96 well plates, 386 well plates, or 1536 well plates. The use of higher density plates greatly increases screening output; use of a 96-well plate allows 10 000 compounds to be screened per day, 40 000 compounds per day with a 384-well microplate, and 200 000 compounds per day with a 1536-well microplate.

Other solid phase assay formats can also be used, such as microfluidic chips for cell-based screening systems (Young, 2013). The process of HTS is costly and can be labor-intensive depending on the extent to which robotics are used in the process. A relatively new approach to HTS is fragment-based screening (Schulz and Hubbard, 2009) which involves a smaller library of small molecules which are evaluated for the binding to a target. Fragments which bind are then assembled

into larger compounds which are then evaluated for binding to a target.

Virtual Screening

The magnitude of potential drug compounds estimated to be 10^{17} (Valler and Green, 2000) requires the automation of drug design such that even with high density plates and combinatorial synthetic methods, screening of libraries in excess of a million compounds is a formidable task. The basis for virtual screening depends on the establishment of quantitative structure activity relationships (QSAR). QSAR is a modeling technique used to predict the activity of a chemical from molecular characteristics known as descriptors (Goodarzi *et al.*, 2012). QSAR models are also used to predict chemical toxicity (Benigni and Richard, 1998). QSAR modeling can be traced to the work of Hammett (1937). Hammett was able to separate the effect of a substituent on benzene ring on the rate of reaction into a constant, σ (sigma) which is a function of the substituent group and ρ (rho) which is the reaction constant dependent on the nature of the reaction, solvent, and temperature. Approximately 30 years later, Hansch, Fujita, and coworkers (Hansch *et al.*, 1962) took the Hammett constant, σ , and combined it with a measure of lipophilicity, π (pi), which is derived from $\log P$ ($\log P$ is the logarithm of the distribution of the compound in question between *n*-octanol and water) measurement and developed a new method for predicting correlation between biological activity and chemical structure. This work with that of others evolved into QSAR (Hansch, 1976). Comparative molecule field analysis (CoMFA) can be considered an extension of QSAR from 2-D to 3-D emphasizing the importance of molecular shape on biological activity (Cramer *et al.*, 1988). Virtual fragment-based screening is also being used to develop lead compounds through docking programs (Cavasotto and Orry, 2007). Regardless of the success of virtual screening, there will be an eventual need to document the value of a lead compound before proceeding to drug development (Funatsu *et al.*, 2011). Virtual screening can also be used to predict ADME (absorption, distribution, metabolism, excretion) properties (Hop *et al.*, 2008).

Classical HTS screening and computer-assisted drug design should result in the same candidate product as was demonstrated with an inhibitor for transforming growth factor β type I receptor kinase. One group used the HTS technique (Sawyer *et al.*, 2003) while the other group (Singh *et al.*, 2003) used 'shape-based' screening which uses docking programs and virtual screening for optimization of a starting chemical which had activity. A more complete description of the factors driving the congruence of these two is available (Shekhar, 2008).

Drug Design and Biologics

It has been possible to use drug design in the development of biologics, specifically antibodies. Haber (1983) was one of the first to consider antibodies as drugs which would be amenable to rational drug design. The use of polyclonal immunoglobulin as a therapeutic product dates to the 1940s with the

development of the Cohn fractionation scheme for blood plasma. Some 40 years later, it became possible to prepare monoclonal antibodies which would be specific for a specific epitope on a target. Various monoclonal antibodies have become therapeutic products such as Herceptin[®] or Avastin[®]. The classical development of monoclonal antibodies is a tedious process and there has been considerable interest in using similar techniques to those developed for small molecules for the development of specific monoclonal antibodies. The objective here is to model the site on the antibody, the complementary determining region (CDR) or paratope which combines with the epitope or antigenic determinant. So the challenge here is two-fold compared to the docking experiments described above for small molecules in that the structures of two large proteins are required.

An antibody such as immunoglobulin G (IgG) is a large protein which can, for the purpose of the current discussion, be divided into two parts (Figure 4). Antibody-drug design has focused on the amino terminal region (Fab domain) portion of the protein as it is the region of the protein that recognizes the antigen and it is possible to use limited proteolysis to obtain the Fab fragment suitable for study by structural biology. In particular, it has been possible to obtain a high-resolution crystal structure for Fab-antigen complexes such as that for the Fab derived from a monoclonal antibody and the α -toxin from *Staphylococcus aureus* (Oganesyan *et al.*, 2014). This monoclonal antibody (MED14893) is being developed as a therapeutic agent for diseases mediated by *S. aureus* such as soft tissue infections and sepsis. The structure demonstrated that the antibody recognized a novel epitope and prevented the α -toxin from forming a lytic conformation. Only a portion of the Fab region is thought to bind the antigen; this is the variable fragment (Fv) region which is composed of the amino terminal regions of the heavy and light chains and contains the CDR or paratope region. Work has progressed on the modeling of Fv fragments where the predicted structure of the Fv domain is compared to those obtained by crystallography (Teplyakov *et al.*, 2014). The rapid progress in this area suggests that ability to design antibodies as drugs will increase in the next decade; challenges will

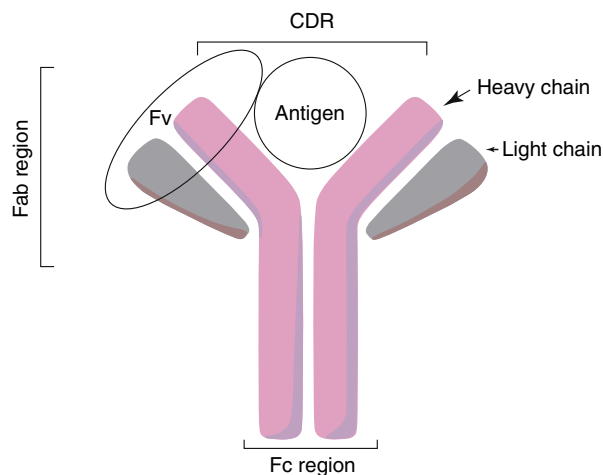


Figure 4 The structure of an immunoglobulin antibody showing the several functional domains.

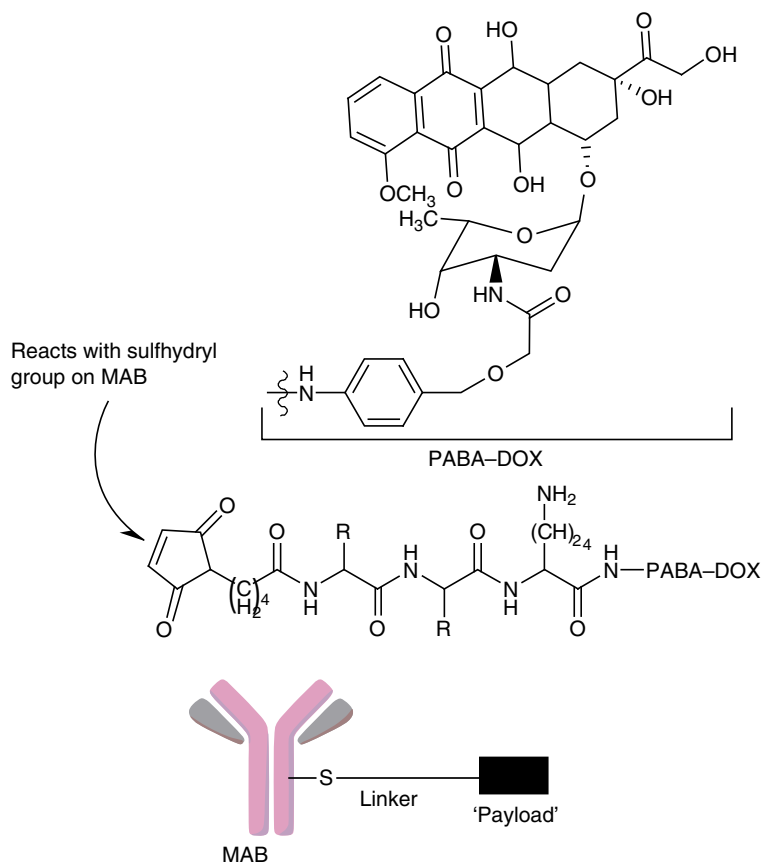


Figure 5 The structure of an antibody–drug conjugate.

remain driven, in part, by the discontinuous nature of epitopes (Kuroda *et al.*, 2012).

Antibody–Drug Conjugates

One of the several problems in drug design is the delivery of the drug to the site of action without collateral damage. Many anticancer drugs are cell poisons and do not discriminate between healthy cells and tumor cells. In addition, many drugs are complex chemicals requiring many synthetic steps and the manufacturing process is expensive. Coupling of the drug to a monoclonal antibody directed toward the target cell has evolved as a novel approach to this problem (Leal *et al.*, 2014). In principle this approach allows the delivery of an inactive drug to the target cell where the antibody–drug conjugate is internalized and transported to the lysosome where the active drug is released (Figure 5). The coupling of the drug to protein uses a linker usually attached to a maleimide which will selectively react with cysteine residues on the antibody protein. Conjugate vaccines consisting of a polysaccharide antigen linked to a protein carrier are also used. Coupling of polysaccharides to protein carriers such as tetanus toxoid protein usually used periodate oxidation of the carbohydrate with subsequent coupling to lysine residues in the protein (Frasch, 2009).

See also: Horizontal Integration: Disease: Drug Targeting. Molecular Principles, Components, Technology, and Concepts: Proteins: Chemical Biology

References

- Benet, L., 2013. The role of BCS (biopharmaceutics classification system) and BDDCS (biopharmaceutics drug disposition classification system) in drug development. *Journal of Pharmaceutical Sciences* 102, 34–42.
- Benigni, R., Richard, A.M., 1998. Quantitative structure-based modeling applied to characterization and prediction of chemical toxicity. *Methods* 14, 264–276.
- Campbell, S.F., 2000. Science, art and drug discovery: A personal perspective. *Clinical Science (London)* 99, 255–260.
- Cavasotto, C.N., Orry, A.J.W., 2007. Ligand docking and structure-based virtual screening in drug discovery. *Current Topics in Medicinal Chemistry* 7, 1006–1014.
- Choudary, S.K., Qiu, J., Plaut, A.G., Kritzer, J.A., 2013. Versatile substrates and probes for IgA1 protease activity. *ChemBioChem* 14, 2007–2012.
- Cramer III, R.D., Patterson, D.E., Bunce, J.D., 1988. Comparative molecular field analysis (CoMFA). I. Effect of shape on binding of steroids to carrier proteins. *Journal of the American Chemical Society* 110, 5959–5967.
- DesJarlais, R.L., Seibel, G.L., Kuntz, I.D., *et al.*, 1990. Structure-based design of nonpeptide inhibitors specific for the human immunodeficiency virus 1 protease. *Proceedings of the National Academy of Sciences* 87, 6644–6648.
- Frasch, C.E., 2009. Preparation of bacterial polysaccharide-protein conjugates: Analytical and manufacturing challenges. *Vaccine* 27, 6468–6470.
- Funatsu, K., Miyao, T., Arakawa, M., 2011. Systematic generation of chemical structures for rational drug design based on QSAR models. *Current Computer-Aided Drug Design* 7, 1–9.

Goldstein, A., 1945. The mechanism of enzyme-inhibitor-substrate reactions: Illustrated by the cholinesterase-physostigmine-acetylcholine system. *Journal of General Physiology* 27, 529–580.

Goodarzi, M., Dejaegher, B., Vander Heyden, Y., 2012. Feature selection methods in QSAR studies. *Journal of AOAC International* 95, 636–651.

Haber, E., 1983. Antibodies as models for rational drug design. *Biochemical Pharmacology* 32, 1967–1977.

Hammett, L.P., 1937. The effect of structure on the reactions of organic compounds. Benzene derivatives. *Journal of the American Chemical Society* 59, 96–103.

Hansch, C., 1976. On the structure of medicinal chemistry. *Journal of Medicinal Chemistry* 19, 1–6.

Hansch, C., Maloney, P.P., Fujita, T., Muir, R.M., 1962. Correlation of biological activity of phenoxyacetic acids with Hammett substituent coefficients and partition coefficients. *Nature* 194, 178–180.

Hardy, L.W., Finer-Moore, J.S., Montfort, W.R., *et al.*, 1987. Atomic structure of thymidylate synthase: Target for rational drug design. *Science* 235, 448–455.

Hop, C.E., Cole, M.J., Davidson, R.E., *et al.*, 2008. High throughput ADME screening: Practical considerations, impact on the portfolio and enabler of in silico ADME models. *Current Drug Metabolism* 9, 847–853.

Kragh, H., 2008. From disulfiram to Antabuse: The invention of a drug. *Bulletin of the History of Medicine* 33, 82–88.

Kuroda, D., Shirai, H., Jacobson, M.P., Nakamura, H., 2012. Computer-aided antibody design. *Protein Engineering, Design, and Selection* 25, 507–521.

Leal, M., Sapra, P., Hurvitz, S.A., *et al.*, 2014. Antibody-drug conjugates: An emerging modality for the treatment of cancer. *Annals of the New York Academy of Sciences* 1321, 41–54.

Mahdi, J.G., Mahdi, A.J., Mahdi, A.J., Bowen, I.D., 2006. The historical analysis of aspirin discovery, its relation to the willow tree and antiproliferative and anticancer potential. *Cell Proliferation* 39, 147–155.

Martin-Moe, S., Lim, F.J., Wong, R.L., *et al.*, 2011. A new roadmap for biopharmaceutical drug product development: Integrating development, validation, and quality by design. *Journal of Pharmaceutical Sciences* 100, 3031–3043.

Nielsen, C., Kjærns, J., Sørensen, K.R., Engelholm, L.H., Behrendt, N., 2014. Advances in targeted delivery of small interfering RNA using simple bioconjugates. *Expert Opinion on Drug Delivery* 11, 791–822.

Oganessian, V., Peng, L., Damschroder, M.M., *et al.*, 2014. Mechanisms of neutralization of a human anti- α -toxin antibody. *Journal of Biological Chemistry* 289, 29874–29880.

Rathore, A.S., Winkler, H., 2009. Quality by design for biopharmaceuticals. *Nature Biotechnology* 27, 26–34.

Reale, M., Di Nicola, M., Velluto, L., *et al.*, 2014. Selective acetyl- and butyrylcholinesterase inhibitors reduce amyloid- β ex vivo activation of peripheral chemo-cytokines from Alzheimer's disease subjects: Exploring the cholinergic anti-inflammatory pathway. *Current Alzheimer Research* 11, 608–622.

Sawyer, J.S., Anderson, B.D., Beight, D.W., *et al.*, 2003. Synthesis and activity of new aryl and heteroaryl substituted pyrazole inhibitors of the transforming growth factor β type I receptor kinase domain. *Journal of Medicinal Chemistry* 46, 3953–3956.

Schulz, M.N., Hubbard, R.E., 2009. Recent progress in fragment-based lead discovery. *Current Opinion in Pharmacology* 9, 615–621.

Shapiro, A., 2013. Development of long-acting recombinant FVIII and FIX Fc fusion proteins for the management of hemophilia. *Expert Opinion in Biological Therapy* 13, 1287–1297.

Shekhar, C., 2008. In silico pharmacology: Computer-aided methods could transform drug development. *Chemistry & Biology* 15, 413–414.

Singh, J., Chuaqui, C.E., Boriack-Sjodin, P.A., *et al.*, 2003. Successful shape-based virtual screening: The discovery of a potent inhibitor of the type I TGF β receptor kinase (T β RI). *Bioorganic & Medicinal Chemistry Letters* 13, 4355–4359.

Sneader, W., 2000. The discovery of aspirin: A reappraisal. *British Medical Journal* 321, 1591–1594.

Tepljakov, A., Luo, J., Obmolova, G., *et al.*, 2014. Antibody modeling assessment II. Structures and models. *Proteins* 82, 1563–1582.

Valler, M.J., Green, D., 2000. Diversity screening versus focused screening in drug discovery. *Drug Delivery Today* 5, 286–293.

Vane, J.R., 2000. The fight against rheumatism: From willow bark to COX-1 sparing drugs. *Journal of Physiology and Pharmacology* 51, 573–586.

Young, E.W., 2013. Advances in microfluidic cell culture systems for studying angiogenesis. *Journal of Laboratory Automation* 18, 427–436.

Further Reading

Agarotiis, D.K., Lobanov, V.S., Salemme, F.R., 2002. Combinational informatics in the post-genomic era. *Nature Review Drug Discovery* 1, 337–345.

Alban, S., 2008. Pharmacological strategies for inhibition of thrombin activity. *Current Pharmaceutical Design* 14, 1152–1175.

Alekseevich, A., 2012. *Glycobiology and Drug Design*. Washington: American Chemical Society.

Alley, S.C., Okeley, N.M., Senter, P.D., 2010. Antibody–Drug Conjugates: Targeted Drug Delivery for Cancer. *Current Opinion in Chemical Biology* 14, 529–537.

Aslam, M., Dent, A., 1998. Bioconjugation. *Protein Coupling Techniques for the Biomedical Sciences*. New York: Grove's Dictionaries.

Burger, A., 1988. *Drugs & People. Medications, Their History and Origins, and the Way They Act*, Revised edn. Charlottesville: University of Virginia Press.

Davis, A.M., Keeling, D.J., Steele, J., *et al.*, 2005. Components of successful lead generation. *Current Topics in Medicinal Chemistry* 5, 421–439. **Drug Development and Industrial Pharmacy**, Informa Healthcare, London.

Flygore, J.A., Pillow, T.H., Aristoff, P., 2013. Antibody–Drug conjugates for the treatment of cancer. *Chemical Biology and Drug Design* 81, 119–121.

Guo, Z., Wang, Q., 2009. Recent development in carbohydrate-based cancer vaccines. *Current Opinion in Chemical Biology* 13, 608–617.

Hansch, C., Fujita, T., 1995. *Classical and Three-Dimensional QSAR in Agrochemistry*. Washington: American Chemical Society.

Höltje, H.D., Sippl, W., Rognan, D., Folkes, G., 2008. *Molecular Modeling: Basic Principles and Applications*, second ed. Weinheim: Wiley-VCH.

Holmgren, E.B., 2014. *Theory of Drug Development*. Boca Raton, FL: Chapman and Hall/Taylor and Francis.

Huber, L., 2007. *Validation and Qualification in Analytical Laboratories*, second ed. New York: Informa.

Klebe, G., 2013. *Drug Design: Methodology, Concepts, and Mode-of Action*. Berlin: Springer.

Kubinyi, H., 1993. *QSAR: Hansch Analysis and Related Approaches*. Weinheim: VCH.

Kyranos, J.N. (Ed.), 2004. *High Throughput Analysis for Early Drug Discovery*. Amsterdam: Elsevier.

Leahy, D.E., Sykora, V., 2008. Automation of decision making in drug design. *Drug Discovery Today* 10, e437–e441.

Mann, J., 1992. *Murder, Magic, and Medicine*. Oxford: Oxford University Press.

Mayr, L.M., Fuerst, P., 2008. The future of high-throughput screening. *Journal of Biomolecular Screening* 13, 443–448.

Neidle, S. (Ed.), 2009. *Cancer Drug Design and Discovery*. Amsterdam: Elsevier.

Osada, H. (Ed.), 2009. *Protein Targeting with Small Molecules*. Hoboken: Wiley.

Otvos, L. (Ed.), 2008. *Peptide-based Drug Design*. Totowa: Humana.

Rankovic, Z., Murphy, R. (Eds.), 2010. *Lead Generation Approaches in Drug Discovery*. Hoboken: Wiley.

Reynolds, C.H., Holloway, M.K., Cox, H.K., 1995. *Computer-aided Molecular Design: Applications in Agrochemicals, Materials, and Pharmaceuticals*. Washington: American Chemical Society.

Schmidt, T., Bergner, A., Schwede, T., 2014. Modelling three-dimensional protein structures for applications in drug design. *Drug Discovery Today* 19, 890–897.

Silverman, R.B., 1992. *The Organic Chemistry of Drug Design and Drug Action*. San Diego: Academic Press.

Slivoski, G., Kothiwale, S., Meiler, J., Lowe Jr., E.W., 2014. Computational methods in drug discovery. *Pharmacology Reviews* 66, 334–395.

Tautermann, C.C., 2014. GPCR structures in drug design, emerging opportunities with new structures. *Bioorganic and Medicinal Chemistry Letters* 24, 4073–4079.

Thorne, N., Ingelse, J., Auld, D.S., 2010. Illuminating insights into firefly luciferase and other bioluminescent reagents used in chemical biology. *Chemistry & Biology* 17, 646–657.

Wang, P.G. (Ed.), 2009. *High-throughput Analysis in the Pharmaceutical Industry*. Boca Raton, FL: CRC Press.

Zanders, E.D., 2005. *Chemical Genomics*. Totowa: Humana.

Relevant Website

<http://www.fda.gov/drugs/informationondrugs/ucm079436.htm>
Food and Drug Administration.

Antibodies and Improved Engineered Formats (as Reagents)

PJ Hudson, Avipep Pty Ltd., Parkville, VIC, Australia

© 2016 Elsevier Inc. All rights reserved.

Introduction

The discovery of hybridoma technology by G. Kohler and C. Milstein in 1975 heralded a new era in antibody research and clinical development (Milstein, 2000). These murine hybridomas provided the first reliable source of monoclonal antibodies (IgG; Figure 1). Initially, antibodies were developed for *in vitro* diagnosis of blood components and disease-associated biomarkers (immunoassays and enzyme-linked immuno-sorbent assays (ELISA)). Later, antibodies became the core detection system for immunohistochemistry (IHC), predominantly for detection of cancer in tissue biopsies. Throughout the 1990s, innovative recombinant DNA technology (chimerization and humanization) successfully reduced the immunogenicity of injectable (*in vivo*) murine antibodies thereby enhancing clinical efficiency and leading to a number of therapeutic IgG and antibodies and their smaller fragments (Fab, scFv molecules). These developments have continued and since 2000 over 30 injectable antibodies against cancer, viral, and inflammatory diseases have been approved. Most recently, innovative structural designs have

further improved *in vivo* pharmacokinetics, expanded immune repertoires and enabled screening against refractory targets and complex proteome arrays, while novel molecular evolution strategies have enhanced affinity, stability, and expression levels. This article discusses the creation of a vast range of novel, engineered, antibody-based reagents which specifically target biomarkers of human health and disease. Beyond 2015, an international collaboration will make antibodies available against every human protein, analyzed for expression in normal and diseased tissue, for research, diagnosis and eventually for improved therapy (the Human Protein Atlas; Fagerberg *et al.*, 2014).

Antibodies for *In Vitro* Diagnosis, in Soluble (RIA), Surface-Bound (ELISA), Cell-Bound (Fluorescence-Activated Cell Sorting (FACS)), Tissue-Bound (IHC) and Modern E-Device Formats

Antibodies are the backbone of the immunodiagnostics industry, providing high-affinity reagents with precise target

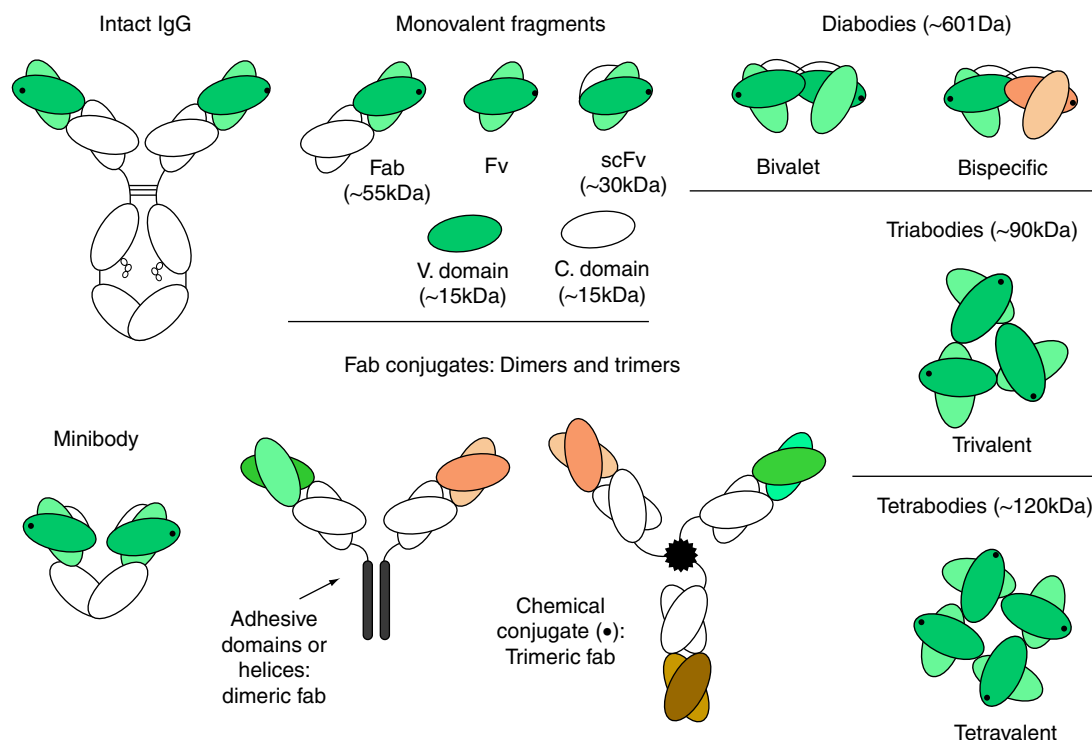


Figure 1 Intact antibody (IgG) and engineered formats. Schematic representation of an intact IgG together with Fab and Fv fragments and single V-domains and C-domains. V-domains are depicted as colored ovals with a dot representing the antigen-binding site and C-domains are uncolored ovals. Engineered recombinant antibodies are shown as scFv monomers, dimers (diabodies), trimers (triabodies), tetramers (tetrabodies) with linkers joining the V-domains represented as a black line. Minibodies are shown as two scFv modules joined by two C-domains. Also shown are Fab dimers (conjugates by adhesive polypeptide or protein domains) and Fab trimers (chemically conjugated). Colors denote different specificities for the bispecific scFv dimers (diabodies) and Fab dimers/trimers.

specificity for the detection of biological molecules (e.g., disease biomarkers) and chemical molecules (e.g., blood components and toxins). The detection output can range from a simple colorimetric Yes/No score (e.g., pregnancy testing in urine samples using dip-sticks) to accurate high-throughput quantitation (e.g., multi-channel analytical instrumentation for full-component blood analyses). Antibodies are extremely stable proteins and can be stored for years, either in solution or immobilized on diagnostic devices. Initially, antibodies were tagged with radioactivity to measure target binding (radioimmunoassay (RIA)) but this has been superseded by fluorescent and luminescent probes for accurate laser-driven quantitation, or enzyme amplification for improved sensitivity. For example, in the commonly used ELISA format, after the initial antibody binding to target, an enzyme is used to generate and amplify a color or fluorescent or electrochemical signal. For rapid, inexpensive, home-based testing, antibodies are immobilized on membrane devices and in 'sandwich' ELISA format, used to detect agents in saliva or blood (pin-prick) or, for example, in urine for pregnancy testing. The ELISA process can be scaled up in microarray (tray) formats or automated for high-throughput diagnosis with robotically driven spectrophotometric or laser detectors. In the USA alone, the \$80 billion diagnostic-lab industry performs nearly 10 billion tests annually and is estimated to provide the basis for about 70% of doctors' medical decisions. Speed and sample size is critical, and new technologies for complete blood diagnosis from a single blood pin-prick sample are in development (TheranosTM). In the most recent developments, antibodies can now be attached to multiple bead populations (or nanoparticles) differentiated by size and levels of fluorescence excitation to simultaneously quantitate multiple analytes in a single assay. In similar technology, DNA 'barcoded' antibodies have been developed with simple amplification systems for extreme-sensitivity, multitarget diagnosis (Agasti *et al.*, 2012).

One of the most widespread diagnostic applications has been the detection of target antigens in tissue sections (IHC; Figure 2). In this method, the initial antibody binding to the antigen is followed by an enzyme-driven color deposit and used for accurate diagnosis (Figure 2). In a further advance, fluorescently tagged antibodies can be used to selectively bind ('paint') a unique cell type in a complex cell mixture. Flow cytometers are then used to accurately assess the quantity and antigen-type displayed on the tagged cells in the mixture, using the unique surface fluorescence provided by the antibody. In cell-sorting (FACS) formats, the tagged cells can be separated from the untagged cells, for more detailed examination of these minor cell populations. Single cells can be isolated from extremely complex biological mixtures, providing exquisite diagnosis of cell physiology and phenotype.

Antibodies for *In Vivo* Applications: Immunotherapy with Humanization and Deimmunisation

Intact multivalent antibodies (IgG; Figure 1) provide high-specificity, high-affinity targeting agents for immunotherapy and currently provide over 40% of \$80 billion per annum biopharmaceutical industry (Weiner *et al.*, 2012; Reichert,

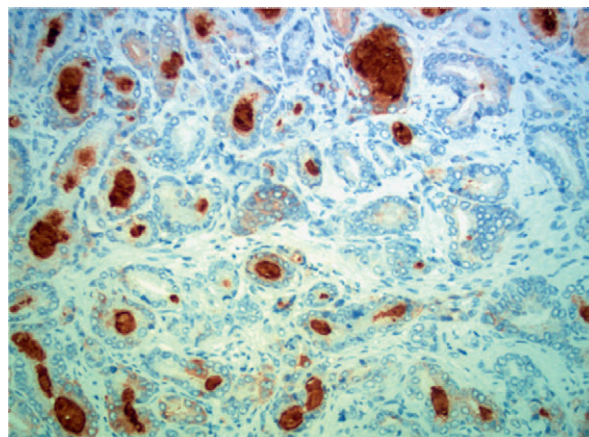


Figure 2 Cancer diagnosis by Immunohistochemistry (IHC). Antibodies are the core reagents for the detection of tumor-associated antigens in tissue sections of cancer biopsies. Shown below is the TAG-72 tumor antigen, expressed in adenocarcinoma of the prostate and identified by a murine CC49 antibody specific to TAG-72. For detection, a second antibody carrying peroxidase is used to bind to the CC49 murine antibody and the peroxidase used to develop a brown stain. Note the presence of secreted antigen (dark brown) and cell-associated antigen (light brown), against a background of clear benign (noncancerous) tissue. Prostate cancer tissue was positive for TAG72 (dark areas) while benign tissue was negative (light areas).

2014). Their simultaneous binding to two antigens increases functional affinity and confers high retention times on target molecules. In addition, intact antibodies comprise Fc domains which can be important for cancer immunotherapy by both recruiting cytotoxic effector functions and extending the serum half life, mediated by the Fc receptor. Unmodified murine monoclonal antibodies formed the first wave of approved immunotherapeutic reagents, although their *in vivo* applications were limited since repeated administrations provoked an anti-mouse immune response. Simple strategies have been developed to avoid, mask or redirect this human immune surveillance, including fusion of murine variable regions to human constant regions as 'chimeric' antibodies, or 'deimmunizing' by removal of T-cell epitopes or 'humanizing' by grafting murine surface residues onto human acceptor frameworks. Modern alternative strategies now allow selection of fully human antibodies directly from natural or synthetic repertoires, including live transgenic mice producing purely human antibodies.

Design of 'Antibody Fragments' for Unique Clinical Applications *In Vivo*

For cytokine inactivation, receptor blockade or viral neutralization, the Fc-induced effector functions are often unwanted and can be simply removed by proteolysis of intact antibodies to yield monovalent Fab fragments. Proteolysis, however, does not easily yield molecules smaller than a Fab fragment, and microbial expression of single-chain Fv (scFv) is currently the favored method of production (Figure 1). In scFvs, the variable (V_H and V_L) domains are stably tethered together with a flexible polypeptide linker (Figure 1; Holliger and Hudson, 2005).

In comparison with whole antibodies, small antibody fragments such as Fab or scFv exhibit improved pharmacokinetics for tissue penetration and also provide full binding specificity since the antigen-binding surface is unaltered. However, Fab and scFv are monovalent and often exhibit fast off-rates and poor retention time on the target. Therefore, Fab and scFv fragments have been engineered into dimeric, trimeric, or tetrameric conjugates to increase functional affinity, employing either chemical or genetic cross-links (Figure 1). Various methods have been devised to genetically encode multimeric scFv, with the most successful design being the simple reduction of scFv linker length to direct the formation of bivalent dimers (diabodies, 60 kDa), triabodies (90 kDa), or tetra-bodies (120 kDa) (Figure 1). Indeed, the first clinical imaging trials of scFv fragments are likely to be as multivalent reagents (diabodies and minibodies), since they exhibit high functional affinity and have been very successful in preclinical studies (Knowles and Wu, 2012; Li *et al.*, 2010, 2011).

Pharmacokinetics of Intact Antibodies versus Fragments

The efficiency of antibodies *in vivo*, for example, in cancer therapy, lies in their capacity to discriminate tumor-associated antigens at low levels. Immunotherapy has been more successful against circulating cancer cells than solid tumors due to the better cell accessibility, and delivery of therapeutic payload. The lymphoma-associated antigens CD30 (Hodgkin's) and CD20 (Non-Hodgkin's) have been successfully treated by therapeutic antibodies.

Radiolabelled antibodies are important clinical reagents for both tumor imaging and therapy. The choice of radionuclide dictates the application, for example, 124-Iodine is used for positron emission tomography (PET) imaging and 123-Iodine for single-photon emission computed tomography (SPECT). Therapeutic administration using for example 90-Yttrium requires a balance between long dissociation rates at the target site and slow blood clearance which can lead to liver accumulation and high radiation exposure of other tissues. Bio-distribution studies in solid tumors have also revealed that large size of IgG molecules (150 kDa) can inhibit rapid tumor penetration, whilst intermediate sized multivalent fragments providing rapid tissue penetration, high target retention, and rapid blood clearance (Knowles and Wu, 2012). For example, diabodies (60 kDa) are ideal with short-lived radioisotopes for clinical imaging due to the fast clearance rates. Larger molecules such as PEGylated diabodies (80 kDa), are used with long-lived radioisotopes and are ideal for tumor therapy since they achieve a higher total tumor 'load' (Figure 3; Li *et al.*, 2010, 2011). Minibodies (90 kDa) and Fab dimers (110 kDa) are also useful for tumor imaging (Knowles and Wu, 2012).

High-affinity antibodies ($K_d < 10^{-10}$ mol l⁻¹) may not be the best reagents for *in vivo* applications, since they bind tightly to first-contact tumor cells and can fail to fully penetrate solid tumors (Adams *et al.*, 2001). High-affinity antibodies can also form large immune complexes with soluble antigen leading to significant off-target toxicity. Intermediate-affinity antibodies ($K_d \sim 10^{-8}$ mol l⁻¹) are potent tumor-loading agents more

resistant to immune complex formation (Cao *et al.*, 2012). Note that affinity is an exceptionally important consideration when evaluating the design and efficacy of targeted therapeutics.

The short half life of antibody fragments can also be extended using modification with polyethylene glycol, which avoids both renal and hepatic localization. The improved functional affinity, tumor penetration, and biodistribution of these novel engineered antibody fragments herald the development of a new generation of reagents for both imaging and therapy.

Engineering Multiple Specificity in Antibody Fragments

Bispecific Antibodies

Bispecific antibodies contain two different binding specificities fused together and, in the simplest example, bind to two adjacent epitopes on a single target antigen, thereby increasing the avidity. Alternatively, bispecific antibodies can cross-link two different antigens and are powerful therapeutic reagents, particularly for recruitment of cytotoxic T cells for cancer treatment. For example, CD19 on a cancer cell can be cross-linked to CD3 on a T-cell, thereby recruiting T cells to attack and kill the cancer cell. Bispecific antibodies and also dimeric ScFv fragments and bispecific diabodies have been produced for this purpose (Figure 1; Maher and Adami, 2013).

Bifunctional Antibodies

The original 'Magic Bullet' concept has proven a success; antibodies have been fused to a vast range of molecules that ascribe important ancillary functions following target binding. These include radionuclides (discussed above) and also cytotoxic drugs, toxins, peptides, proteins, enzymes, and viruses, the latter for targeted gene therapy. For cancer therapy, bifunctional antibodies are engineered to effectively target tumor-associated antigens at low levels and then deliver a cytotoxic payload to tumor cells. The latest antibody-toxin conjugates are stable *in vivo*, minimally immunogenic and are proven cancer therapies; AdcetrisTM for Hodgkin's lymphoma and KadcylaTM for Her-2 positive breast cancer (Panowski *et al.*, 2014). These antibody-drug conjugates (ADCs) carry only a few drug molecules on each antibody. As nanoparticle conjugates and immunoliposomes, antibodies have been shown to deliver higher drug payloads than simple chemical fusion of the drugs to the antibody. Antibody-enzyme fusions have also been developed for prodrug activation, primarily for cancer therapy.

Antibody Libraries: Construction, Display, and Selection

Library display has superseded hybridoma technology through the creation of large natural and synthetic antibody repertoires *in vitro*. From these novel libraries, specific high-affinity antibodies can be selected by linking phenotype (binding affinity) to genotype, thereby allowing simultaneous recovery of the

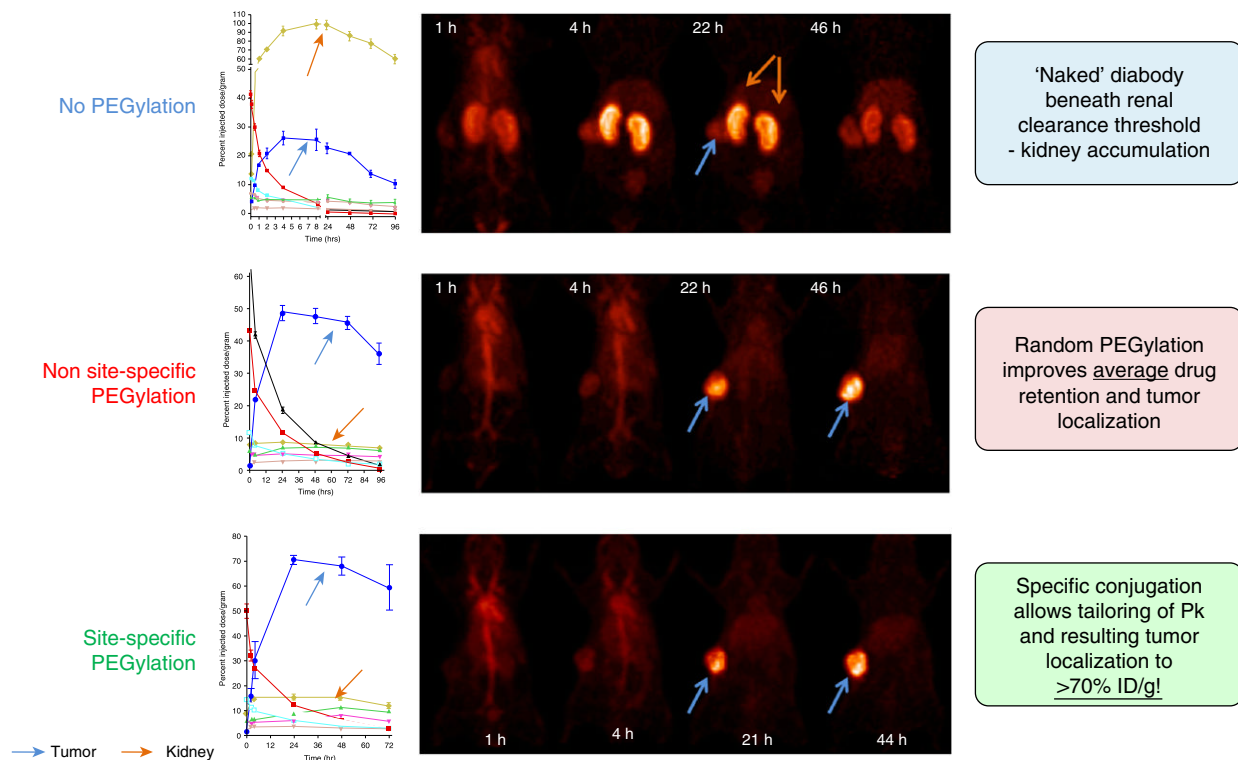


Figure 3 Molecular Imaging using antibodies targeting human tumors (xenografts, growing in mice). Diabodies with and without conjugated PEG, were labeled with 64-Copper and PET (positron emission tomography) images taken at 1, 4, 22, and 46 h. Without PEG the diabodies are filtered through the kidney (orange arrows), but with added PEG, kidney uptake is avoided and the diabody loads almost exclusively into the xenograft tumor (blue arrows), achieving high-contrast tumor imaging at a remarkable 70% ID/gram tumor uptake. Reproduced from Li, L., Turatti, F., Crow, D., *et al.*, 2011. Site-specific conjugation of monodispersed DOTA-PEGn to a thiolated diabody reveals the effect of increasing peg size on kidney clearance and tumor uptake with improved 64-copper PET imaging. *Bioconjugate Chemistry* 22, 709–716; and Li, L., Crow, D., Turatti, F., *et al.*, 2010. Monodispersed DOTA-PEG–conjugated anti-TAG-72 diabody has low kidney uptake and high tumor-to-blood ratios. *Journal of Nuclear Medicine* 51, 1139–1146.

gene encoding the selected antibody. Antibodies are usually displayed as monovalent Fab or scFv fragments and then, as required, reassembled into intact immunoglobulin (Ig) or multivalent variants following selection. If the repertoire is sufficiently large, a high-affinity Fab or scFv can be selected directly or, more frequently, the recovered gene is subjected to cycles of mutation and further selection to enhance affinity (Harel Inbar and Benhar, 2012).

Bacteriophage Display

Fd-phage and phagemid technologies are currently the most popular *in vitro* methods for the display of large repertoires and for the selection of high-affinity recombinant antibodies against a range of clinically important target molecules. Important improvements in selection technology have included array screening for high-avidity antibodies and recovery of internalized phage from live cells to select against internalizing (human) receptors, which are useful for ADC therapeutic designs.

Cell-Surface Libraries

Antibodies had been displayed on the surface of cells (yeast or bacterial cells) enabling efficient screening by high-speed flow cytometers to identify high-affinity antibodies.

Transgenic Mice

Transgenic mice have been produced that lack the native murine immune repertoire and instead harbor most of the human antibody repertoire in the germline. Injection of antigens into these humanized animals leads to the development of human-like antibodies which have undergone murine somatic hypermutation and selection to relatively high affinity. Recovery of antibodies can be by classic hybridoma technology or, for more efficient affinity enhancement, can be recovered by the *in vitro* display and selection technologies.

Maturation

Transgenic mice and display libraries typically produce human antibodies with affinities in the range from 10^{-7} to 10^{-9} M. Obtaining higher affinity antibodies is crucial for efficient binding to the antigenic target for *in vitro* diagnosis, viral neutralization, cell targeting, and *in vivo* imaging. To improve antibody affinity, various *in vitro* strategies have been developed to mimic the mammalian *in vivo* process of somatic hypermutation, focusing on several cycles of mutation, display, selection (recovery), and gene amplification. Affinity enhancement can be restricted to mutations in the

antigen-binding surface (complementarity determining regions (CDRs) loops) but even with the most detailed structural information, the techniques for design of precisely complementary surfaces via interface mutations remain in their infancy. Importantly, mutations in the underlying framework regions have frequently provided large increases in affinity, stability, and expression.

Production, Stability, and Expression Levels

Production of antibodies for preclinical and clinical trials has been evaluated in numerous expression systems including bacteria, yeast, plants, insect, and mammalian cells. Bacteria are favored for expression of small, non-glycosylated Fab and scFv fragments, usually with terminal polypeptides, such as c-myc, His, or FLAG for affinity purification. Mammalian cells are favored for intact antibodies and occasionally, also for expression of scFvs, diabodies, and minibodies.

High-Value Clinical Applications

Cancer targeting for delivery of radioisotope or cytotoxic drug payloads has been discussed earlier, and this section describes some other clinical applications of engineered antibodies for viral infection, autoimmune disease, stroke, neurodegenerative diseases, and image-guided surgery.

Pathogen Neutralization, Antiviral Therapy, and Vaccines

Antibody binding can directly and effectively block the activity of many pathogens without requiring Fc-mediated cytotoxicity. Indeed, this has always been the promise of antibody-mediated viral neutralization and the first monoclonal antibody (Palivizumab; SynagisTM) for the treatment of viral disease was launched in 1998. Palivizumab is a humanized antibody used in immunoprophylaxis for the prevention of severe respiratory syncytial virus (RSV) disease. Despite this success, and the huge range of antibodies available against human immunodeficiency type 1 (HIV) and hepatitis C (HepC), the use of recombinant antibodies as therapeutics against viral infection has been limited. Only a few rare antibodies have exhibited potent neutralization *in vitro* and antiviral efficacy in animal models (Koff *et al.*, 2013). This is probably due to viral efficiency both in producing escape mutants and in evolving antibody-resistant receptor-binding surfaces.

At the crux of immune modulation, vaccines have also been effective for producing effective, host-derived antibodies against a number of major diseases (polio and cholera). In response, many pathogens have evolved effective evasion mechanisms to avoid host immuno-surveillance. A promising new strategy is to use antibodies as Trojan horses to drive immunogenic epitopes and molecular motifs directly into dendritic cells and force the host to produce a more effective immune response (Sehgal *et al.*, 2014). This re-booting or boost to the natural immune effector mechanism has worked in animal models and is being applied to clinical diseases (Koff *et al.*, 2013).

T-Cell Targeting and Immune Checkpoint Inhibitors

Cancer cells can be destroyed by cell recruitment of cytotoxic T cells, natural-killer (NK) cells or macrophages which can themselves be targeted by encoding cell-surface antibodies (usually scFv). The latest strategies involve removal and then genetic engineering a patient's T-cells to produce chimeric antigen receptors (CARs), which then efficiently target tumor when re-infused into the patient. Alternative cell recruitment strategies include using bispecific antibodies to cross-link and activate effector cells at the tumor site, or the use of bifunctional antibodies fused to cytokines and chemokines to provide local T-cell stimulation and proliferation, also at the tumor site (Hess and Neri, 2014).

Immune checkpoint-blocking antibodies modulate T-cell pathways that regulate T cells and have the potential to re-invigorate an antitumor immune response (Kyi and Postow, 2014). Ipilimumab (YervoyTM) was the first US Food and Drug Administration (FDA)-approved immune checkpoint antibody licensed for the treatment of metastatic melanoma (MM) and blocks a checkpoint molecule CTLA-4 (cytotoxic T-lymphocyte antigen 4). The Programmed Death receptor (PD-1) has also been effectively targeted by systemically administered antibodies, reducing the effect of the tumor-associated PD-L1 and enhancing T-cell responses to mediate antitumor activity. While both CTLA-4 and PD-1 function as negative regulators, CTLA-4 attenuates the early activation of naïve and memory T cells whilst PD-1 is primarily involved in modulating T-cell activity in peripheral tissues via interaction with its ligands, PD-L1, and PD-L2. Unfortunately, not all patients respond to these therapies, and evaluation of the optimal clinical strategy continues. These immune checkpoint inhibitors may also have profound effects in ameliorating stroke (Bodhankar *et al.*, 2014).

Angiogenesis and Vascular Blockade/Transport

Blocking angiogenesis (blood vessel growth) to prevent the establishment and expansion of tumors is becoming an important strategy in the fight to reduce cancer spread. Bevacizumab (AvastinTM) targets VEGF, blocks angiogenesis and is an FDA-approved therapy for several metastatic cancers. Indeed, some of the anti-angiogenic antibodies act synergistically to enhance the efficacy of tumor-targeting antibodies and may become more highly efficient therapeutics when used in combination (Yasuda *et al.*, 2013).

To tackle neurodegenerative diseases, antibodies, and engineered fragments have been designed to transport through neural vasculature thus crossing the blood-brain barrier to target their activity within plaques to inhibit or ameliorate misfolded protein aggregates (Pardridge, 2014).

Image-Guided Surgery: From Radioactive Image-Guided Surgery (RIGS) to Fluorescence Image-Guided Surgery (FIGS)

Antibody-guided imaging can detect cancer in real-time during surgery, providing more accurate cancer treatment while sparing healthy tissue. Antibodies were first used as

radioimaging probes to detect and pinpoint tumor lesions to guide surgical excision. Relying on PET and SPECT whole-body cameras, this RIGS procedure was too lengthy for practical use and has now been replaced by antibodies conjugated to fluorescent or luminescent tracers to can effectively 'paint' tumors or diseased lesions (FIGS; Keereweer *et al.*, 2013). Using rapid detection by externally-guided optical-probe technology (otoscopes) which can be directly coupled to robotic surgical instruments, real-time guided dissection is now possible. Once detected by FIGS, tumors can also be ablated by new high-precision external beam radiotherapy.

Conclusion

Antibodies and their engineered fragments will continue to provide the most powerful range of reagents for our discovery of disease-associated biomarkers and to improve human health through their application as clinical diagnostic and therapeutic reagents. New technologies using antibodies are rapidly increasing, both in scientific and commercial application, stemming from the elucidation of the key elements required for antibody design, efficient expression and desired pharmacokinetics. By providing a highly stable, protease-resistant scaffold, engineered recombinant antibody fragments will continue to be the paradigm for selection of high-affinity clinical reagents for efficient diagnosis and delivery of therapeutic payloads.

See also: Molecular Principles, Components, Technology, and Concepts: Proteins: Drug Design

References

- Adams, G.P., Schier, R., McCall, A.M., *et al.*, 2001. High affinity restricts the localization and tumor penetration of single-chain fv antibody molecules. *Cancer Research* 61, 4750–4755.
- Agasti, S.S., Liong, M., Peterson, V.M., Lee, H., Weissleder, R., 2012. Photocleavable DNA barcode-antibody conjugates allow sensitive and multiplexed protein analysis in single cells. *Journal of the American Chemical Society* 134, 18499–18502.
- Bodhankar, S., Chen, Y., Lapato, A., *et al.*, 2014. Targeting immune co-stimulatory effects of PD-L1 and PD-L2 might represent an effective therapeutic strategy in stroke. *Frontiers in Cellular Neuroscience* 8, 228.
- Cao, Y., Marks, J.D., Huang, Q., *et al.*, 2012. Single-chain antibody-based immunotoxins targeting Her2/neu: Design optimization and impact of affinity on antitumor efficacy and off-target toxicity. *Molecular Cancer Therapeutics* 11, 143–153.
- Fagerberg, L., Hallström, B.M., Oksvold, P., *et al.*, 2014. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Molecular & Cellular Proteomics* 13, 397–406.
- Harel Inbar, N., Benhar, I., 2012. Selection of antibodies from synthetic antibody libraries. *Archives of Biochemistry and Biophysics* 526, 87–98.
- Hess, C., Neri, D., 2014. Evaluation of antibody-chemokine fusion proteins for tumor-targeting applications. *Experimental Biology and Medicine* (Maywood) 239, 842–852.
- Holliger, P., Hudson, P., 2005. Engineered antibody fragments. *Nature Biotechnology* 23, 1126–1130.
- Keereweer, S., Van Driel, P.B., Snoeks, T.J., *et al.*, 2013. Optical image-guided cancer surgery: Challenges and limitations. *Clinical Cancer Research* 19, 3745–3754.
- Knowles, S.M., Wu, A.M., 2012. Advances in immuno-positron emission tomography: Antibodies for molecular imaging in oncology. *Journal of Clinical Oncology* 30, 3884–3892.
- Koff, W.C., Burton, D.R., Johnson, P.R., *et al.*, 2013. Accelerating next-generation vaccine development for global disease prevention. *Science* 340, 1232910.
- Kyi, C., Postow, M.A., 2014. Checkpoint blocking antibodies in cancer immunotherapy. *FEBS Letters* 588, 368–376.
- Li, L., Crow, D., Turatti, F., *et al.*, 2010. Monodispersed DOTA-PEG-conjugated anti-TAG-72 diabody has low kidney uptake and high tumor-to-blood ratios. *Journal of Nuclear Medicine* 51, 1139–1146.
- Li, L., Turatti, F., Crow, D., *et al.*, 2011. Site-specific conjugation of monodispersed DOTA-PEGn to a thiolated diabody reveals the effect of increasing peg size on kidney clearance and tumor uptake with improved 64-copper PET imaging. *Bioconjugate Chemistry* 22, 709–716.
- Maier, J., Adami, A.A., 2013. Antitumor immunity: Easy as 1, 2, 3 with monoclonal bispecific trifunctional antibodies? *Cancer Research* 73, 5613–5617.
- Milstein, C., 2000. With the benefit of hindsight. *Immunology Today* 21, 359–364.
- Panowski, S., Bhakta, S., Raab, H., Polakis, P., Junutula, J.R., 2014. Site-specific antibody drug conjugates for cancer therapy. *MAbs* 6, 34–45.
- Pardridge, W.M., 2014. Blood-brain barrier drug delivery of IgG fusion proteins with a transferrin receptor monoclonal antibody. *Expert Opinion on Drug Delivery* 20, 1–16.
- Reichert, J.M., 2014. Antibodies to watch in 2014 (Annual review). *MAbs* 6, 5–14.
- Sehgal, K., Dhodapkar, K.M., Dhodapkar, M.V., 2014. Targeting human dendritic cells in situ to improve vaccines. *Immunology Letters* 2478, 144–148.
- Weiner, L.M., Murray, J.C., Shuptrine, C.W., 2012. Antibody-based immunotherapy of cancer. *Cell* 148, 1081–1084.
- Yasuda, S., Sho, M., Yamato, I., *et al.*, 2013. Simultaneous blockade of programmed death 1 and vascular endothelial growth factor receptor 2 (VEGFR2) induces synergistic anti-tumour effect *in vivo*. *Clinical and Experimental Immunology* 172, 500–506.

Relevant Websites

- <http://antibodies.cancer.gov/apps/site/default>
Antibody Portal via NIH.
- <http://www.antibodypedia.com>
Antibodypedia.
- <http://en.wikipedia.org/wiki/ELISA>
ELISA Techniques.
- http://en.wikipedia.org/wiki/Flow_cytometry
FACS and Flow Cytometry Techniques.
- <http://en.wikipedia.org/wiki/Immunohistochemistry>
IHC Techniques.
- www.proteinatlas.org
The Human Protein Atlas.

MOLECULAR PRINCIPLES, COMPONENTS, TECHNOLOGY, AND CONCEPTS: LIPIDS

Contents

Lipidomics

Synthesis and Structure of Glycerolipids

Cholesterol and Other Steroids

Glycolipids

Lipid Signaling

Lipidomics

BM Kenwood and AH Merrill Jr., Georgia Institute of Technology, Atlanta, GA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The lipidome is comprised of all of the biomolecules considered to be lipids, which "... have been loosely defined as biological substances that are generally hydrophobic in nature and in many cases soluble in organic solvents..." (Fahy *et al.*, 2005). This encompasses compounds that are so structurally diverse that they were divided into eight categories by LIPID MAPS, an NIH-funded Consortium that spent a decade exploring ways to facilitate lipidomic studies, and maintains an ongoing website with tools and information. The eight categories were fatty acyls, glycerolipids, glycerophospholipids, sphingolipids, sterol lipids, prenol lipids, saccharolipids, and polyketides (Fahy *et al.*, 2005, 2009).

Lipidomics has been defined as "the full characterization of lipid molecular species and of their biological roles with respect to expression of proteins involved in lipid metabolism and function, including gene regulation" (Spener *et al.*, 2003), but from the earliest days of use of the suffix '-ome' in reports on lipid analysis (Kishimoto *et al.*, 2001; Han and Gross, 2003) to today, it has only been technically feasible for a lipidomic analysis to encompass a fraction of the total lipid species – albeit sometimes large numbers of compounds – but certainly not the entire lipidome. Thus, some of the important questions to answer when examining a report of a lipidomic analysis are 'what was measured?' and 'for what purpose?' and 'was the appropriate method used to provide the desired data?' Some might consider the first question to be 'why lipidomics?' so that will be addressed, too.

Why Lipidomics?

An 'omic' analysis of any type (lipidomics, genomics, etc.) has two types of justifications that sometimes overlap but can

be quite different. One – the 'systems biology' perspective – is to obtain enough information about a biological system to derive equations that describe it and enable predictions about additional features, including future outcomes. This goal can be motivated by scientific curiosity or a desired application, such as to predict how perturbation of one step of a pathway by a drug might impact other metabolites. The other – sometimes called 'discovery science' – is to try to find something unexpected about a biological system that might challenge current hypotheses and/or enable new ones to be proposed. This foray into the realm of 'the unknown unknowns' can also have practical application if, for example, it enables discovery of a disease biomarker or an off-target effect of a drug.

The application of an 'omics' approach to lipids is clearly justified from both of these perspectives because their metabolic pathways are highly interconnected, beginning with a shared biosynthetic precursor and interwoven by downstream intermediates, as illustrated in Figure 1.

This diagram illustrates how acetyl-CoA is a common precursor of all eight lipid categories. After it has been utilized to make the distinguishing compounds in one category (e.g., fatty acids), these will be utilized to produce compounds in other lipid categories, as shown for the acyl-groups of the glycerol(phospho)lipids, the sphingoid base and amide-linked fatty acyls of sphingolipids, and sterol esters, to name just a few of the cases where compounds in one category are metabolically related to other categories.

In addition, the lipidome is comprised of not just compounds that are made *de novo* (colloquially 'from scratch') by an organism (e.g., mammals for the example in the inner box of Figure 1) but also from food and other sources, including gastrointestinal microflora (possible sources of saccharolipids and polyketoides as well as other lipids). When all of these are taken into account, the numbers of distinct molecular

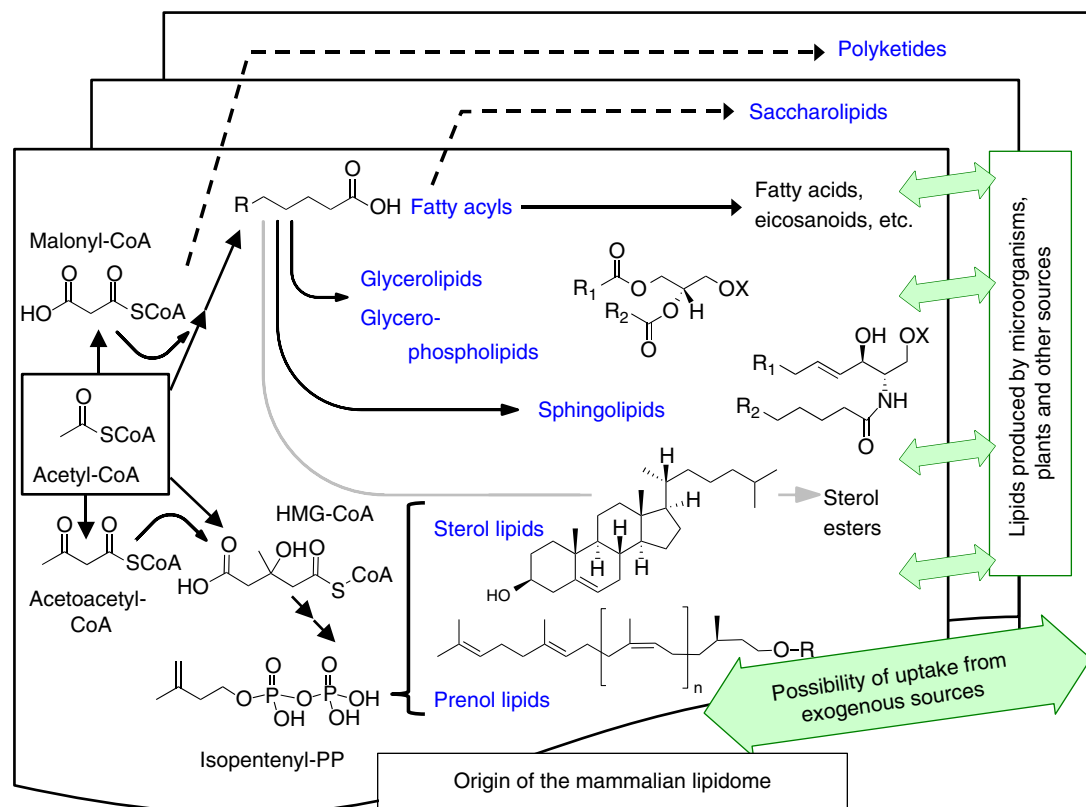


Figure 1 An overview of lipid categories and some of their metabolic interrelationships. This diagram is a modification of a LIPID MAPS summary illustrating how eight lipidome categories are fundamentally related by their precursors (acetyl-CoA and downstream metabolites). It also displays how later metabolites are also shared; for example, fatty acids for biosynthesis of the glycerol(phospho)lipids, sphingolipids, and sterol esters made by mammals (and saccharolipids by other organisms). The mammalian lipidome is also impacted by lipids encountered from food and other routes such as gastrointestinal microflora.

species is mind-boggling: the number of structural variations in lipids has been estimated to be $> 10^5$ (Dennis, 2009), and this is probably an underestimate because novel compounds continue to be found.

Description of How the Components of the Lipidome Are Usually Analyzed

The large number of lipids and the high degree of structural variation among, and even within, the eight categories makes their analysis much more challenging than for most other categories of metabolites. While it has been possible to profile some lipids by common structural features, such as glycolipids that can be captured by lectins that bind the carbohydrate moiety, for a more comprehensive lipidomic analysis it is necessary to use a method that is highly structure specific – and preferably can also be quantitative – such as mass spectrometry (MS).

The power of mass spectrometry is that compounds can be detected in very small amounts once they are ionized, and instruments are available to resolve ions by the mass to charge ratio (m/z), sometimes with high enough mass accuracy to match the data with the molecular formula of the corresponding lipid(s). The analysis might not need such high mass accuracy if the compounds of interest can be distinguished

otherwise (e.g., by their behavior on chromatography prior to mass spectrometry); however, the experimenter must always validate that the ions truly reflect the compound that it has been presumed to be. This is even important when the instrument has a high mass accuracy (such as a time-of-flight, ToF, or fourier transform ion cyclotron resonance, FT-ICR mass spectrometers) because many lipids are isomeric (such as glucosylceramides vs galactosylceramides, which vary only in the stereochemistry of one of the hydroxyls on the carbohydrate) and/or are isobaric (i.e., have the same mass but different composition) within the accuracy of the analysis.

Additional information can be acquired about the compound from other features of the mass spectrum, such as the isotopic distribution, evidence that the compound is present as multiply charged species, etc. In metabolism studies, one can also introduce compounds that are enriched in stable isotopes and track these into downstream metabolites.

Lipidomic analysis requires a myriad of choices that must be made, and these steps will be discussed here and in the following sections of this article. The critical choices are summarized in Figure 2 and additional information can be found in excellent reviews (Murphy and Axelsen, 2011) and monographs on this topic (Brown, 2007). The first depend on the goals of the study: Is quantitation of the lipids necessary? If so, then a targeted LC-MS/MS protocol with multiple extractions and internal standards will probably be required. Is

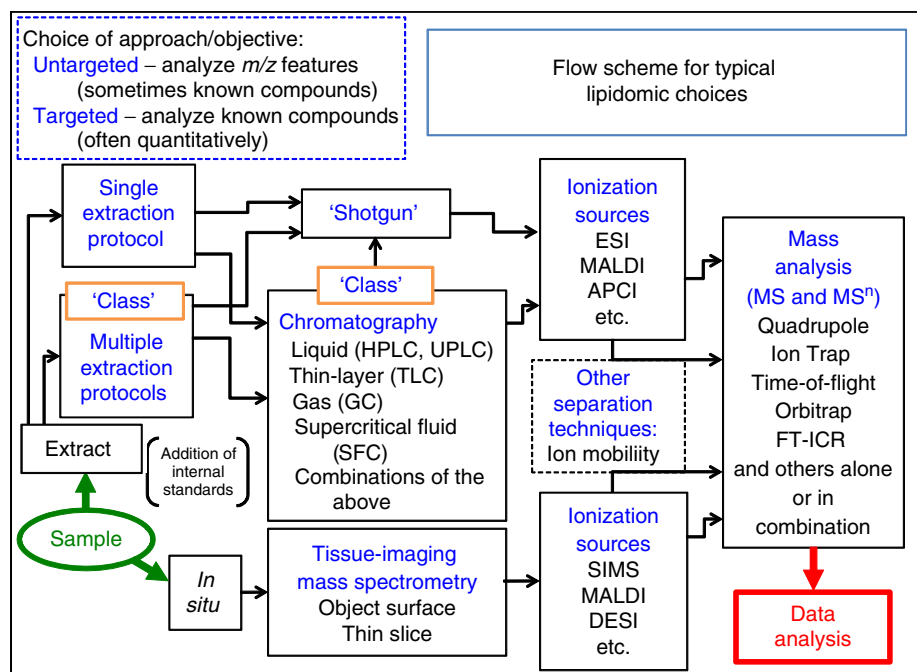


Figure 2 A flow diagram for the types of technical choices that can be used for a lipidomic analysis. Depending on the sample and the information desired, one might elect to extract the lipids or analyze them *in situ*. If extracted, the compounds can be analyzed using a 'one-size-fits-all' extraction protocol or with multiple extraction schemes that ensure high recoveries for all lipidomic categories (this has been called the CLASS approach for comprehensive lipidomics analysis by separation simplification). The extracts can be analyzed by a 'shotgun' approach, which often involves introducing the entire sample into the mass spectrometer and having it conduct a broad, but high resolution, scan to identify as many compounds as possible, or if comparing two samples, features in the spectra that differ and might be interesting to pursue. In the chromatography toolkit, a variety of methods can be used to partially separate the lipids to minimize ionization suppression of minor species as well as to separate compounds that cannot be distinguished by mass spectrometry alone (such as isomers). In the mass spectrometry toolkit, ion trap (which includes quadrupole and orbitrap), time-of-flight, and FT-ICR can be used alone or in combination with one another. If analysis *in situ* is desired, this can be conducted by methods that ablate ions from the surface of the material directly (such as DESI and SIMS) or by first impregnating the sample with a matrix material that can be excited by a laser to generate ions (MALDI).

the goal to discover potential biomarkers by surveying a large number of samples? If so, then an untargeted approach might be the best choice. Is it important to know the location of the analyte(s) in a sample such as a histological slice? If so, then an imaging MS method is usually selected.

Getting Started – Sample Preparation for Lipidomic Mass Spectrometry

With the exception of imaging MS methods, which generate ions from the surface of a material (SIMS and DESI) or within the sample with imbedded MALDI matrix compounds (for more information, see [Murphy and Merrill, 2011](#)), lipidomic studies begin with extraction of the lipids from the samples to be analyzed. This presents two of the most vexing problems of lipidomic analysis: (1) that lipids across multiple categories vary so widely in both their structures and chemical/biophysical properties; and (2) that lipids differ in amounts by many orders of magnitude. Thus, it is difficult – and perhaps even impossible – to find a single extraction protocol that will efficiently extract compounds that vary so much in: stability – from very chemically stable ceramides to easily oxidized polyunsaturated glycerophospholipids; hydrophobicity – from very hydrophobic triacylglycerols to highly polar gangliosides; and amount – from highly abundant sterols to trace amounts of signaling lipids.

There have been many approaches to lipid extraction, and to some degree, most of the protocols followed today are similar to the 'classic' Bligh and Dyer and Folch extraction methods with modifications to improve the recovery of particular categories of lipids. However, even within a single category of lipids, one protocol might not be able to extract all of the subspecies with high yield; therefore, multiple protocols are used to cover all categories in this type of lipidomic study. This has been referred to as CLASS lipidomics for 'comprehensive lipidomics analysis by separation simplification' ([Harkewicz and Dennis, 2011](#)).

The CLASS protocols of LIPID MAPS were optimized with 'cocktails' of internal standards by the LIPID MAPS Consortium and have been compiled in reviews published in *Methods in Enzymology* ([Brown, 2007](#)) which can be downloaded as a single PDF on the LIPID MAPS website. Some of the procedures were expanded in later reports ([Murphy and Merrill, 2011](#)). The 'LIPID MAPS' internal standards can be obtained from Avanti Polar Lipids (Alabaster, AL) ([Sims et al., 2014](#)) for most of the categories and from Cayman Chemicals (Ann Arbor, MI) for eicosanoids.

There have been a number of recent reports of single extraction protocols that seem to work well with several categories of lipids ([Lydic et al., 2014](#); [Reis et al., 2013](#)), but the number of compounds that have been examined has been

fewer than for the CLASS protocols, so it is not known if single-solvent protocols will become prevalent for broader-based lipidomics analysis. Single extraction protocols have been mainly used in untargeted lipidomic studies where the goal is to survey a wide spectrum of ions to look for noticeable differences between two biological samples (Harkewicz and Dennis, 2011). In this context, the term 'lipidomic' connotes that a wide profile has been examined rather than that all species have been analyzed quantitatively.

Lipidomic Mass Spectrometry

Most lipidomic analyses use some form of chromatography prior to mass spectrometry (as shown in Figure 2), but the mass spectrometric analysis will be discussed first. An excellent source for information about the analysis of each category of lipid by mass spectrometry has recently been published (Murphy, 2015) and is available in print and online.

Mass spectrometry

Mass spectrometry is quite simply the analysis of compounds using an instrument that can separate the ionized state of the compounds by their mass to charge ratio (m/z). Since the exact mass can be calculated for any compound, one can use the m/z to verify that a compound has its presumed molecular composition. *Vice versa*, when presented with a mass spectrum and a spreadsheet with the m/z of the detected ions, one can attempt to deduce the molecular composition of the ions. This might sound simple, but a number of factors complicate structural assignments from the m/z of lipids, such as: (a) for compounds to be seen, they must be ionized under the conditions of the analysis, and this varies considerably for different lipids; (b) even if a compound can be ionized, its behavior might be affected by the presence of other lipids, especially when there are large amounts of other compounds that might promote formation of aggregates; (c) some compounds are chemically unstable under the ionization conditions, so they might not be seen at all at the predicted m/z (this is not always a problem because the decomposition product might suffice as an ion for the compound); (d) depending on the mass accuracy of the instrument, there will probably be several to dozens of compounds that share the same m/z within experimental error; (e) even if the mass accuracy allows an exact molecular composition to be assigned, there are a large number of isomeric and isobaric lipids, which preclude absolute structural identification; and (f) each compound will give more than one m/z because it will also be comprised of high-mass natural abundance isotopes (e.g., ^{13}C), which can also result in the isotopomer of a high-abundance compound to obscure the signal from a lower abundance compound of interest with the same m/z .

Despite these limitations, a high-resolution mass spectroscopic analysis of biological samples without optimizing extraction protocols (such as in a 'shotgun' mode) can be very informative, especially if combined with chromatography or other techniques that help overcome some of these technical limitations (Wang and Han, 2014).

Tandem mass spectrometry

The method that is most often used to confirm the identity of the primary ion is tandem mass spectrometry, where ions of a selected m/z from the first mass analyzer are fragmented by a process termed collision-induced dissociation (CID) and information about the fragments is obtained using another mass analyzer. This is called 'tandem MS,' 'MS/MS,' or 'MS².' Multiple repetitions of this process (n times) is called 'MSⁿ.' Multiple MS analysis is usually conducted with an instrument with 'ion trapping' capability, which allows for the filtration of specific ions and fragments within the mass spectrometer. As described herein, the combination of MSⁿ and ion trap is a powerful technique that can be utilized to robustly measure specific lipids.

Since different categories of lipids typically undergo distinct fragmentations, the precursor-product pairs can distinguish between many of the isobaric compounds, but often does not distinguish isomers. The precursor-product pairs can also be used for a type of analysis termed 'multiple reaction monitoring' (MRM) where the first mass analyzer is set to pass a particular precursor m/z and the second mass analyzer to detect its partner product m/z . By programming the instrument to cycle among a selected set of ion pairs, the instrument can optimize the data collection time for compounds of interest, which is particularly useful when the sample is flowing through the instrument transiently, such as during a chromatographic separation. Some instruments also have the capacity perform more sophisticated operations, such as to have features activated when particular ions are detected (e.g., to scan a full spectrum, or analyze particular product m/z when a particular precursor m/z of interest is detected).

This process is summarized in Figure 3 using ceramides as an example of the main steps that are taken in setting up a protocol for analysis of lipids by liquid chromatography-tandem mass spectrometry (LC-MS/MS). In the first step, lipid standards are examined by MS to determine the types of ions that are detected by the instrument(s) available to the investigator (it is worth mentioning here that due to the high cost of these instruments, the type of analysis is sometimes influenced by the instrument that happens to be available to the investigator). In this figure, C16-ceramide (N-palmitoylsphingosine) has been infused into the mass spectrometer (an ABI 4000 Qtrap) in methanol with electrospray ionization (ESI). Note that the protonated ceramide (abbreviated $[\text{M} + \text{H}^+]$) is not the most abundant ion in this spectrum because there is a large amount of the Na^+ adduct (labeled $[\text{M} + \text{Na}^+]$). The appearance of multiple ions is common in mass spectrometry because salts are present in most biological samples and can arise from glass test tubes and solvents.

The next spectrum in Figure 3 shows the fragments detected when the $[\text{M} + \text{H}^+]$ ion for C16-ceramide (m/z 538.5) is collided with nitrogen (i.e., the process referred to as CID) and the product ions are scanned. Note that there happens to be one major fragment (m/z 264.3) and the structural diagram on the right of the spectrum shows where this fragmentation is thought to occur (i.e., loss of the fatty acyl- and both hydroxyls of the sphingoid base). This fragmentation is common for all chain-length ceramides so one can analyze a mixture of ceramides by varying the precursor m/z and monitoring which

Example of a typical LC-MS/MS analysis protocol

1. Optimize ionization and identify parent ion(s) for analytes of interest (shown here for a C16-ceramide, see structure on the right)

2. Identify major fragmentation product(s) and select candidate precursor/product ion pairs for MRM analysis

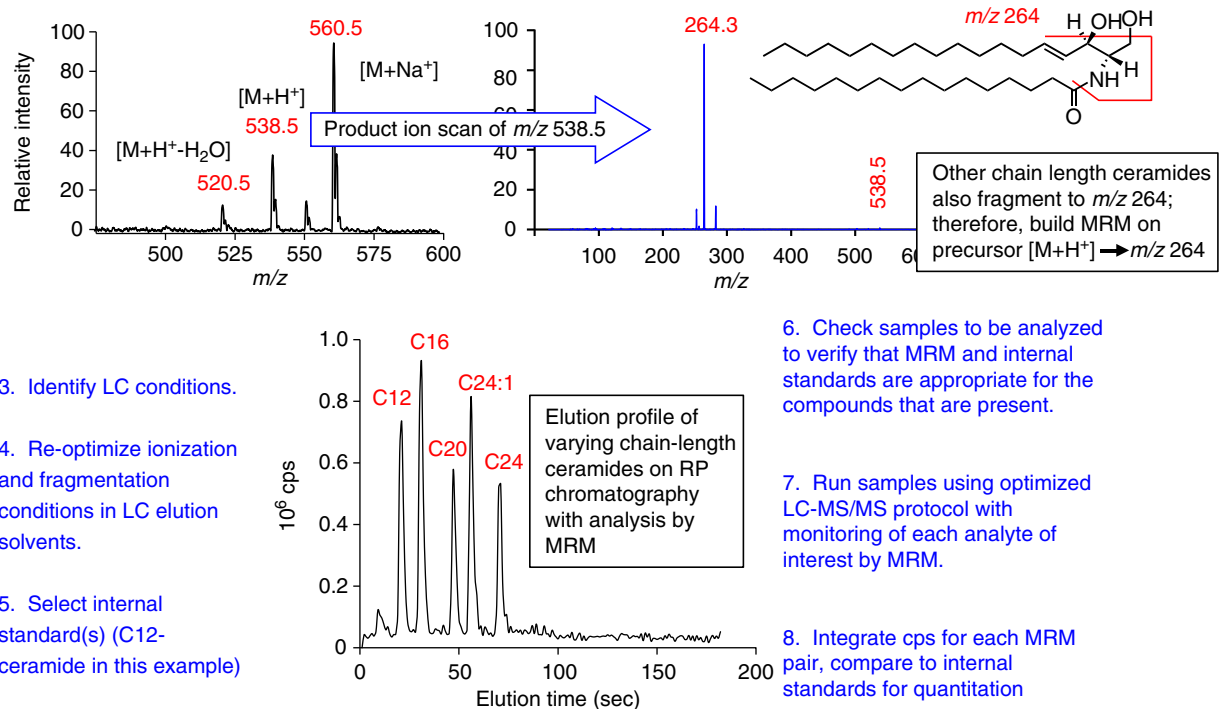


Figure 3 An example of a typical protocol for analysis of lipids by liquid chromatography-tandem mass spectrometry (LC-MS/MS). These spectra are for ceramides, but the principles apply to most lipid categories. In the first step, lipid standards are examined by MS to determine the types of ions that are detected by the instrument being used (in this case, electrospray ionization using an ABI 4000 QTrap), then identification of the major product ions upon collision-induced fragmentation of the $[M+H]^+$ precursor ion (see text for further discussion of the spectra, such as the origin of the large Na^+ adduct in the precursor ion spectrum). Having candidate precursor/product ion pair(s) for a MRM analysis (which is a technique to monitor specific precursor/product pairs), the liquid chromatography (LC), ionization and mass spectrometric analysis conditions are optimized. Samples can then be analyzed as they elute from LC (elution of different chain-length subspecies is labeled in red in the third figure). If the relationship between the ion yields for an internal standard (C12-ceramide in this data set) versus the other analytes, then the areas under the curves can be converted to pmol per sample amount.

deliver m/z 264.3 upon CID. Likewise, ceramides with other sphingoid base backbones (different chain lengths, number of double bonds, or hydroxyls) will give different product m/z (266.3, e.g., from dihydrosphingosine, also named sphinganine), so this can be used to monitor those subspecies of compounds. The efficiency of ionization and fragmentation will usually vary for different subspecies, so this must be taken into consideration if one plans to obtain quantitative information about a sample of interest by including adequate internal standards.

Chromatography in combination with mass spectrometry

To separate compounds that cannot be distinguished by fragmentation patterns alone, one often utilizes some form of chromatography to separate the isomers before MS: gas, GC; high performance-liquid, ultra-high performance liquid, UPLC; supercritical lipid, SFC; and others). Additional

advantages of chromatography are that it often removes interfering ions (such as the Na^+), and helps minimize a process termed 'ionization suppression,' wherein the presence of other lipids, salts, or other factors interfere with formation of ions from the compounds of interest.

Having candidate precursor/product ion pair(s) for a MRM analysis (MRM, which is a technique to monitor specific precursor/product pairs), the liquid chromatography (LC), ionization and mass spectrometric analysis conditions are optimized. Samples can then be analyzed as they elute from LC (elution of different chain-length subspecies is labeled in red in the third diagram of Figure 3). If the relationship between the ion yields for an internal standard (C12-ceramide in this data set) versus the other analytes, is known, then the areas under the curves can be converted to pmol per sample amount. This the most common type of mass spectrometry used for CLASS analysis.

'Shotgun' lipidomic mass spectrometry

In addition the techniques described above, there have been other approaches that provide quicker profiling by 'shotgun lipidomics' which sometimes refers to reliance on the mass spectrometer to provide the lipid 'separation' by identification of the species based on their m/z measured with high mass accuracy (Wang *et al.*, 2014). There have been considerable advances in 'shotgun' methodologies by addition of complementary methods for compound identification (including MS/MS scans that are triggered when ions of interest are detected, as well as pre-MS separation by LC, etc.) (Wang and Han, 2014). Another consideration is whether the analysis is targeted (as are most of the CLASS protocols described in the preceding sections) or untargeted, where all ions are collected and distinctions between two states (e.g., healthy versus disease) are determined by statistical analysis of the data sets and identification of the lipids using mass spectral and retention index libraries (Kind *et al.*, 2009).

Ion-mobility mass spectrometry

Another approach to distinguish isomers is the use of ion-mobility spectrometry (IMS) (Kliman *et al.*, 2011), which occurs after, rather than before, the ions are generated, as shown in Figure 2 in the box with dashed lines. In this technique, the ions are passed through a carrier buffer gas that retards compounds based on their collisional cross-section, which often differs for isomers. IMS has been applied, for example, to the localization of fatty acyl and double bond positions

(Castro-Perez *et al.*, 2011), and works well in combination with hydrophilic interaction liquid chromatography and reversed phase chromatography for increased resolution, specificity, and signal-to-noise ratio (Baker *et al.*, 2014).

Selection of instruments for lipidomic mass spectrometry

The box on the far right of Figure 2 illustrates that a wide variety of instruments are used for analysis of lipidomics samples. They differ in their features in many ways, but as a simple comparison, quadrupole and ion trap instruments tend to be workhorses for quantitative analysis of compounds by LC-MS/MS, and time-of-flight (ToF), orbitrap and FT-ICR instruments are slower but employed when greater mass accuracy is needed.

Tissue-imaging mass spectrometry

This analytical approach has been included in Figure 2, but is shown in more detail in Figure 4. Tissue-imaging mass spectrometry can be advantageous because the biological material is not homogenized and extracted (as symbolized by the blender in Figure 4), but rather ions are generated on the surface of the biomaterial – often tissue slices but sometimes intact organisms. The ions are generated by several methods. In secondary ion mass spectrometry (SIMS), the sample is bombarded with a focused primary ion beam and the ejected secondary ions are analyzed. The ion beams are often of such high energy that organic compounds decompose during ionization, but milder methods using a buckminsterfullerene

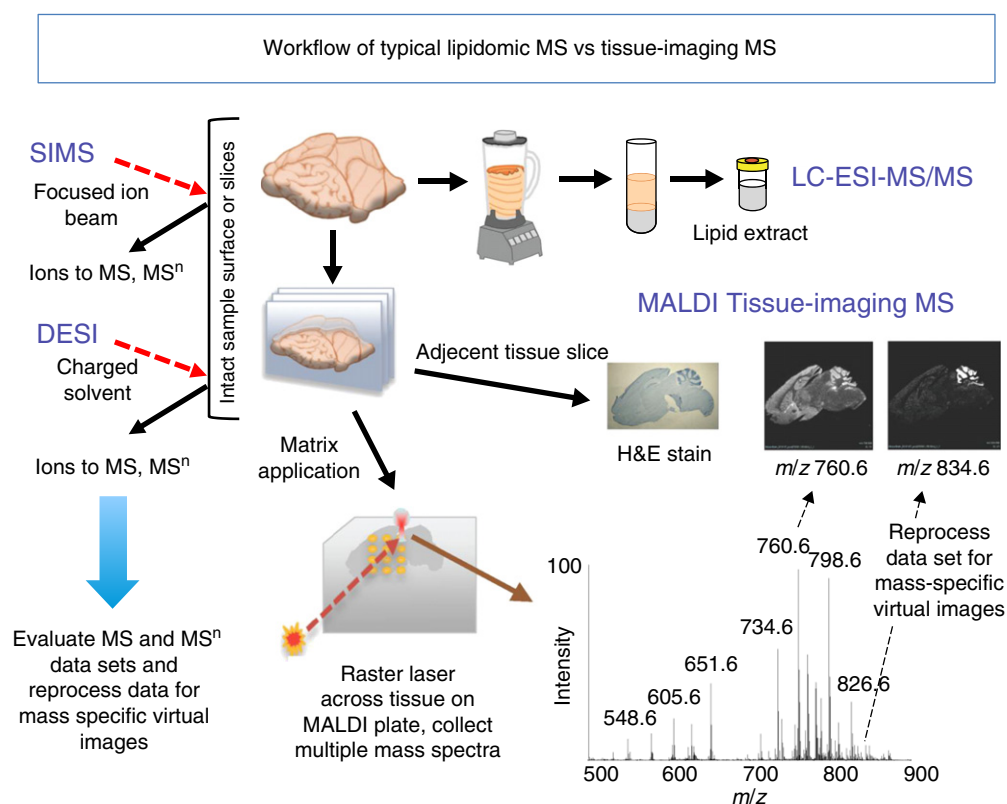


Figure 4 A diagrammatic representation of the differences between lipidomic analysis for samples that have been homogenized and extracted (upper portion of the diagram) versus analyzed *in situ* by SIMS, DESI, or MALDI Tissue-imaging mass spectrometry. Modified from Murphy, R., Merrill, A.H. (Eds.), 2011. Lipidomics and imaging mass spectrometry. *Biochimica et Biophysica Acta* 1811, 635–1000.

(C_{60}) cluster ion beam is less destructive to the lipids (Wino-grad and Garrison, 2010). In desorption electrospray ionization (DESI), an electrically charged solvent is sprayed onto the sample surface and the charged compounds are analyzed in the ejected droplets (Takats *et al.*, 2005). In the methods that use MALDI to generate ions, the sample is usually deposited on a chilled MALDI plate or a conductive glass slide, the MALDI matrix compound is imbedded in the sample as uniformly as possible with minimal disruption of the lipid localization of the tissue (usually by some type of spraying method), then a laser beam is moved incrementally across the sample (described by the term 'raster') to generate ions for analysis by mass spectrometry (Murphy *et al.*, 2009).

As illustrated for MALDI tissue-imaging MS in Figure 4, once the spectra have been collected, the location of any ion(s) of interest can be visually displayed for images that are much like a microscopic analysis of the sample (e.g., an H&E stain as shown in Figure 4), but with the images representing the distribution of that compound. A caveat for MALDI tissue-imaging MS is that the intensity of the ions can be affected by their local environment and other factors, however, this has nonetheless proven to be a powerful tool for gathering information about the location of lipids and other compounds in tissues.

In general, the region that is undergoing ionization is considerably smaller for SIMS (nanometers to sub-micrometers) than MALDI and DESI (typically, 10 to 100 μm in diameter). All of the procedures generate a huge number of individual spectra, but imaging software is available which allows the ions of interest to be displayed (often with a different

color for each ion displayed) in a manner equivalent to a molecular image.

Lipid Categories of the Lipidome

There are several online resources provide information about lipid structures and nomenclature, tools to facilitate lipid analysis, metabolic pathways and genes, and sometimes updates and notices of upcoming conferences. Of note are the American Oil Chemistry Society (AOCS) Lipid Library, LIPID MAPS, and LipidBank (the web addresses are given at the end of this article). In an attempt to obtain uniformity in nomenclature and the way to display the structures of these compounds, the LIPID MAPS Lipid Classification system (Fahy *et al.*, 2005, 2009) has been adopted by some journals, and the website provides downloadable structure diagrams for many different lipids. The following sections highlight the types of structures that distinguish each lipid category and briefly describe some of the features that are utilized in their analysis. The subsections also mention some of the features that are utilized in lipidomic mass spectrometry, but a full bibliography of those methods are beyond the scope of this article and readers are referred to the already cited monographs and reviews for that information.

Fatty acyls

The fatty acyls are comprised of carbon atoms arranged in chains of varying length, unsaturation, branching, and other features (e.g., presence of hydroxyl-groups, cyclopentene, peroxy, nitro, etc. moieties). Their structures are abbreviated with

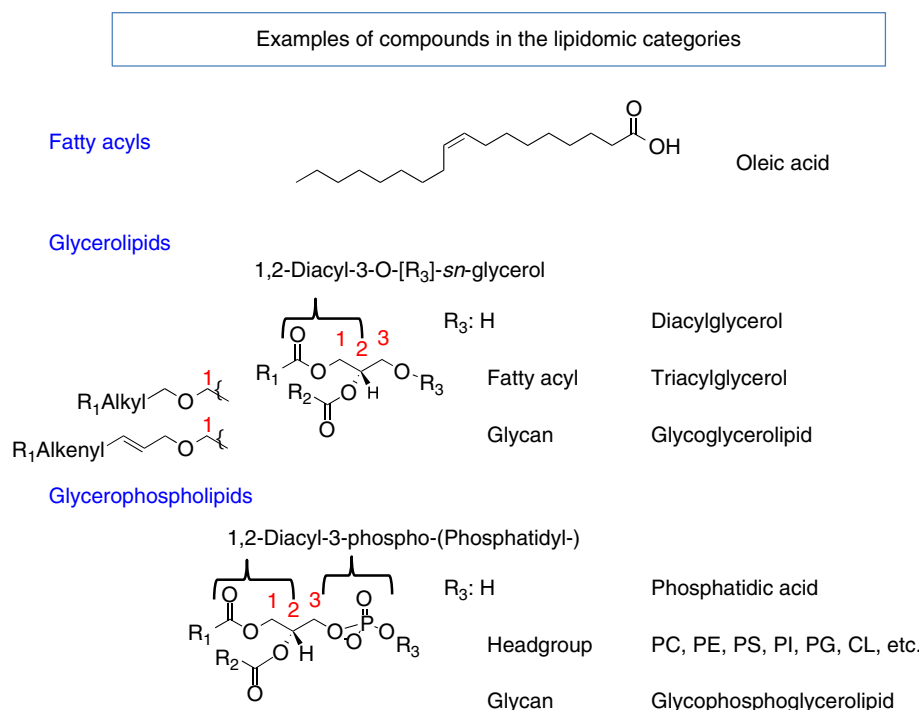


Figure 5 Representative features of three major lipid categories: Fatty acyls, glycerolipids, and glycerophospholipids. A general prototype for these compounds is shown in the upper diagram with R1 and R2 representing the alkyl chains of fatty acids attached to the glycerol backbone at these two positions and at position 3 is attached the glycan.

the number of carbon atoms followed by the number of double bonds and any other relevant information, such as the double bond position and whether *cis* (Z) or *trans* (E). Some of the better known members are the saturated fatty acids palmitic acid (C16:0) and stearic acid (C18:0), the mono-unsaturated fatty acid oleic acid (C18:1, 9Z) and poly-unsaturated fatty acids such as linoleic acid (C18:2, 9Z,12Z) and arachidonic acid (C20:4, 5Z,8Z,11Z,14Z), and the broad spectrum of highly bioactive metabolites that fall under the category of eicosanoids (prostaglandins, leukotrienes, etc.). The fatty acyl category also includes compounds such as fatty acyl-CoA's, fatty amides, etc.

Fatty acids have long been analyzed by derivatization and GC-MS, but the negatively charged carboxylate also allows them to be seen by ESI, with assistance in structural identification by using LC-MS/MS (Hellmuth *et al.*, 2012). They are not easily fragmented by CID in structurally informative ways, so many strategies have been developed to gather 'omic' scale information by shotgun lipidomic analysis (Yang *et al.*, 2011) and for quantitation, EI-GC/MS, ESI-LC/MS, and ESI-LC/MS/MS using stable isotope labeling (Murphy and Axelsen, 2011). Quantitation can also be performed without fragmentation using 'pseudo-molecular' MRM, in which the ions do not fragment (Schiesl *et al.*, 2010), but this is limited by

identifying fatty acid species by retention time and the precursor ion. This allows for identification and quantitation of fatty acids by utilizing two MRM transitions (Hellmuth *et al.*, 2012).

Glycerolipids

Glycerolipids typically have a 1,2-diacyl-*sn*-glycerol moiety (as shown in the prototype in Figure 5) where R1 and R2 are long-chain fatty acids in ester linkage to the glycerol backbone. In some cases, a fatty alcohol can be linked to the glycerol backbone by an ether linkage. Since glycerol per se does not have stereochemistry, the 'sn' designation of glycerolipids (standing for 'stereospecific numbering') takes into account that stereoisomers are generated once asymmetric substituents have been added to the hydroxyls. Glycerolipids are found as mono-, di-, and tri-glycerides and with other groups attached, including carbohydrates, which are commonly found in plants, algae, and bacteria (Holzl and Dormann, 2007; Zhang *et al.*, 2014). When phosphorylated, they fit in the next category, glycerophospholipids.

Glycerolipids are commonly measured by mass spectrometry neutral loss scanning wherein the data is collected for precursor and product ions that differ by the unit of what has been lost (in this case, due to fragmentation at the ester

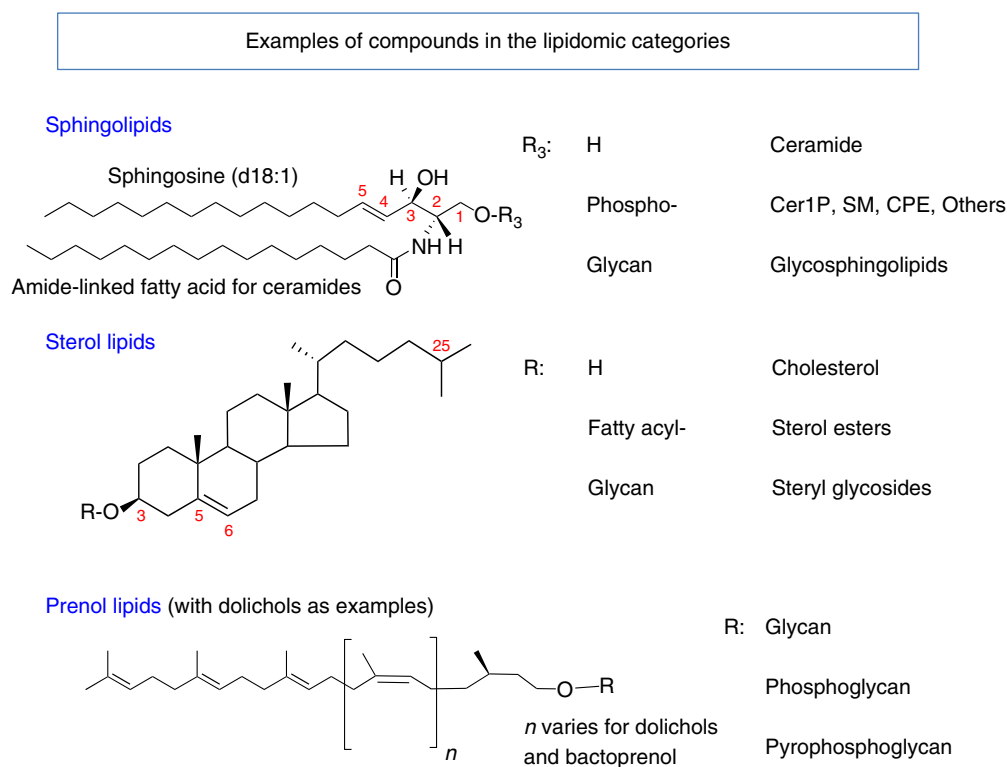


Figure 6 Representative features of three major lipid categories: sphingolipids, sterols, and prenyls. Sphingolipids are comprised of a sphingoid base (shown for sphingosine) that can be derivatized with an amide-linked fatty acid (for ceramide) and/or addition of a headgroup at the hydroxyl on carbon 1, with the simplest cases being sphingosine-1-phosphate and ceramide-1-phosphate (Cer1P). SM and CPE refer to sphingomyelins and ceramide phosphoethanolamines which have the phosphocholine and phosphoethanolamine headgroups, respectively. The glycan headgroups include simple monosaccharides (glucose or galactose) to very complex polysaccharides (such as those found in gangliosides). Sterols have variations of the 4-ring system shown for cholesterol as well as additional groups such as addition of a fatty acid or carbohydrate to the free hydroxyl at carbon 3. Prenols are comprised of isoprene units (shown in brackets) on some combination (the compounds shown are linear polymers, but they can also be part of rings, as in the case of retinol, ubiquinone, etc.). The dolichols and bactotrenols that are shown can be linked to glycans via phosphate or pyrophosphate diesters.

bonds) (Murphy, 2015). Na^+ is often added to facilitate detection of the positively charged adducts (Callender *et al.*, 2007; Murphy and Axelsen, 2011; Murphy *et al.*, 2011).

Glycerophospholipids

The glycerophospholipids have a so-called 'phosphatidic acid' backbone (i.e., 1,2-diacylglycerol 3-phosphate) as shown in Figure 5. If a headgroup is added in phosphodiester linkage, the compound is named phosphatidyl-'headgroup' (such as phosphatidylcholine, PC) unless one of the alkyl chains is a fatty alcohol attached in ether linkage, in which case it is commonly referred to as an alkyl ether- or alkenyl ether-phospholipid (the term 'phosphatidal' is often used when the ether linkage is with an O-alkenyl group). Some organisms, such as Archaea, have two isoprenoid chains attached to positions 2,3 of glycerol via ether bonds as their way of forming an ether lipid (in this case, 2,3-diphytanyl-*sn*-glycerol) (Mori *et al.*, 2014). If there is only one fatty acyl or ether-linked alkyl chain, the name is usually preceded with 'lyso.'

The most common headgroups for glycerophospholipids are choline (PC), ethanolamine (PE), serine (PS), inositol (and its phosphates) (PI, PIP, PIP₂, etc.), and glycerol (PG), which is also present in cardiolipid, 1,3-bis(phosphatidyl)-glycerol (Figure 5). There are also other glycerophospholipid linked glycans, such as CDP-diacylglycerol, a biosynthetic intermediate for phospholipid biosynthesis by many organisms, phosphatidylglucose (Ishibashi *et al.*, 2013), and the 'glycosylphosphatidylinositol anchor' (GPI) that is used to attach some categories of proteins to membranes (Paulick and Bertozzi, 2008).

Glycerophospholipids are often analyzed by loss of the headgroup since the phosphate bond is relatively labile during CID, which gives a backbone atomic composition (i.e., total number of carbons, oxygens, and double bonds) but not the specific fatty acyls and their position (Brown, 2007; Murphy, 2015). Therefore, these data are reported for the sum of the alkyl chain carbons instead of specific molecular species. There are also methods that produce fragments that give information about specific fatty acyls and position (Maccarone *et al.*, 2014).

Sphingolipids

Sphingolipids are defined by containing sphingosine (Figure 6) or another sphingoid base as part of their structure. There are many variations in this backbone lipid moiety, especially for microorganisms and invertebrates (Pruett *et al.*, 2008). For most sphingolipids (i.e., the so-called 'complex' sphingolipids), the amino group is connected to a fatty acid, and these compounds are generically referred to as 'ceramides.' There are also hundreds of different headgroups attached to the hydroxyl-group on the first carbon, which fall in the subcategories of phosphosphingolipids (sphingomyelins, ceramide phosphoethanolamines, ceramide-1-phosphates, and others), neutral glycosphingolipids (where the headgroup is comprised of uncharged sugars – glucose, galactose, glucosamine, galactosamine, and fucose, for mammals), and acidic glycosphingolipids, such as gangliosides (which have sialic acid) and sulfatides (sulfated glycans) and others (for overviews, see Yu *et al.*, 2007; Merrill, 2011 and the 'sphingomap' web site).

Lipidomic methods are available for analysis and quantitation of the sphingolipids that comprise the major components of most of mammalian plasma and cells, which include most of the compounds implicated in cell signaling (sphingosine-1-phosphate, ceramides, ceramide-1-phosphate, etc.) (e.g., see Shaner *et al.*, 2009). The more complex (and often less abundant) glycosphingolipids pose a bigger analytical challenge. The glycan headgroups are sometimes determined after removal of the ceramide (Parry *et al.*, 2007), or for the intact glycosphingolipids, using MSⁿ to identify both the carbohydrates and ceramide backbones (capitalizing on the characteristic fragmentation for ceramides that was illustrated in Figure 3) (Sullards *et al.*, 2011). For examples of these methods, see Farwanah and Kolter (2012) for a general overview of the lipidomics of glycosphingolipids; Ito *et al.* (2012) for a description of the use of atmospheric pressure matrix-assisted laser desorption/ionization (AP-MALDI) coupled to a quadrupole ion trap time-of-flight (QIT-TOF) mass spectrometer for soft ionization and MSⁿ analysis of gangliosides; and Kouzel *et al.* (2014) for an example of how glycosphingolipids have been analyzed in crude lipid extracts by treatment with phospholipase C before TLC to separate and prepare the samples for identification using immunostaining and infrared matrix-assisted laser desorption/ionization orthogonal time-of-flight mass spectrometry (IR-MALDI-o-TOF MS) in combination with CID).

Sterol lipids

The most often thought of sterol is cholesterol (Figure 6), its derivatives (sterol esters and steryl glycosides), and metabolites (bile acids, steroid hormones, corticosteroids, and others). However, there are many additional structural variants (sitosterol, ergosterol, campesterol, stigmasterol, brassicasterol, and others) that are produced by plants, fungi, and other organisms. There are probably more subspecies yet to be discovered considering that the presence of cholesterol glucoside in mammalian cells has only been recently appreciated (Ishibashi *et al.*, 2013).

Many sterols are very nonpolar and relatively difficult to ionize without derivatization (Griffiths *et al.*, 2006), and structural similarities often make them difficult to separate for analysis of isobaric and isomeric species. An additional challenge for mammalian systems is that cholesterol is present in such large amounts compared to the bioactive metabolites that the range in amounts is at least six orders of magnitude – from mg ml⁻¹ for cholesterol to ng ml⁻¹ for some trace oxysterols – but with the appropriate care, these barriers have been overcome (McDonald *et al.*, 2012). Indeed, considerable advances have been made in recent years in the separation and analysis of large families of sterols and sterol esters in an 'omic' context by taking advantage of LC and GC chromatography (John *et al.*, 2014; Matysik *et al.*, 2012; McDonald *et al.*, 2012; Murphy *et al.*, 2011).

Prenol lipids

Prenol lipids are synthesized from five-carbon isoprene units (illustrated in the bracket shown in Figure 6) that can be combined in a wide variety of polymeric units and myriad of configurations to make an amazing number of natural products, including fat-soluble vitamins (Thulasiram *et al.*, 2007).

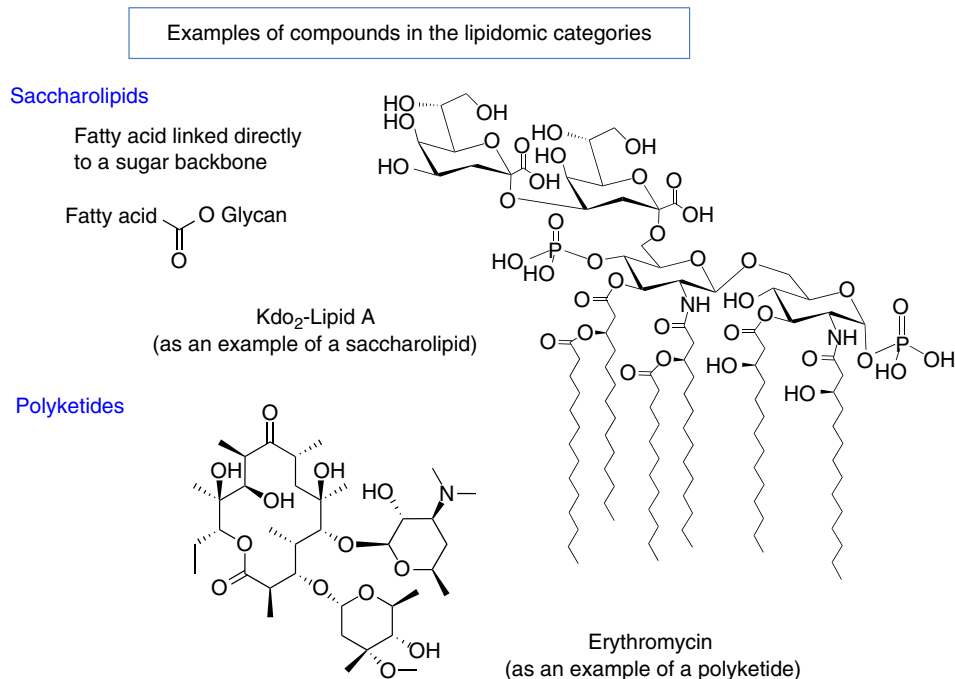


Figure 7 Representative features of two major lipid categories: saccharolipids and polyketides. A general prototype for saccharolipids (acylated glycans) is shown at the top, with an example of the Kdo₂-Lipid A portion of the lipopolysaccharide of gram-negative bacteria (Kdo is the abbreviation for 3-deoxy-D-manno-oct-2-ulosonic acid). The lower portion of the figures shows a polyketide, erythromycin, but this category of compounds is too varied in structure to be represented by a general prototype.

When there are over four isoprene units, they are usually referred to as polyprenols (Vranova et al., 2012), and subdivided into the isoprenoids and quinones/hydroquinones (such as the ubiquinones that are important in mitochondrial electron transport).

Some of the prenyls (i.e., dolichols for eukaryotes and bactoprenol for bacteria) are used in sugar transport across membranes and transfer of carbohydrates to acceptors, with attachment of the carbohydrate to the hydroxyl on the terminal isoprene via a phosphodiester for simple monohexoses and a pyrophosphate for larger oligosaccharides, such as those that are used for N-linked glycoprotein biosynthesis (Welti, 2013). As indicated in the example in Figure 6, the number of isoprene units varies from 18 to 22 for dolichols and 10 to 12 for bactoprenols, and there are also differences in the degree of unsaturation among different sources. The structural variations are becoming more evident as mass spectrometric methods are facilitating the analysis of these compounds (Garrett et al., 2007; Guan and Eichler, 2011).

Saccharolipids

These are compounds in which fatty acids are linked directly to a sugar backbone, as shown in Figure 7 by the general ester bond and example compound. One of the most widely known saccharolipids is Kdo₂-lipid A (Figure 7), which is a component of lipopolysaccharides (LPS) that are present in the outer membranes of most gram-negative bacteria (Raetz et al., 2009). Saccharolipids considerable structural variation among organisms (Banoub et al., 2010), including differences in the

carbohydrates as well as lipids, as represented by the acylated trehalose of *Mycobacteria* (Rombouts et al., 2011). Saccharolipids are often analyzed by a combination of methods to identify both the lipid moieties and glycans (Henderson et al., 2013) since both can exist in many isomeric forms.

Polyketides

These are a structurally diverse group of natural products that are biosynthetically derived from acetyl- or malonyl- units and their structural features (including polyphenols, macrolides, polyenes, enediynes, and polyethers) meet the criterion of being a 'lipid' because most are relatively soluble in organic solvents (Hertweck, 2009). No structural scheme can encompass all of the members of this subclass, so an example, erythromycin, is shown in Figure 7. Other polyketides also have antimicrobial, antiparasitic, and anticancer uses, but some are also toxic, such as the aflatoxin-family carcinogens. Methods for analysis of these compounds are usually focused structural variants and metabolites of each type (e.g., aflatoxins B₁, B₂, M₁, etc.) rather than attempting to encompass all polyketides.

Lipidomics Databases and Other Online Tools

In parallel with the growth of analytical methods for lipidomics, there has been an explosion of lipidomics data and tools in press and online. Besides searching the recent literature to keep abreast of this topic and how it fits into systems biology (e.g., Hyotylainen and Oresic, 2014), one can rapidly

access a large number of tools by examining the LIPID MAPS web site, where there are LIPID MAPS bioinformatics tools and data (as have also been explained in several published reports (Fahy *et al.*, 2007b, 2011)), links to tutorials on how to use the tools and databases, and discussion of their utilization to develop a systems biology of lipids (Subramaniam *et al.*, 2011).

The tools on this website meet the need for high-quality bioinformatics to manage and integrate experimental data at multiple levels (Fahy *et al.*, 2007a): (1) definition of lipid classification and ontologies; (2) relational database design; (3) capture and automated pipelining of experimental data; (4) efficient management of metadata; (5) development of lipid-centric search tools; (6) analysis and visual display of results; and (7) integration of the lipid knowledge base into biochemical pathways and interactive maps. The databases can be queried by text, mass, formula, classification, and other criteria, or most simply by 'Quick search' to obtain information not only about lipids but also MS standards, the lipid proteome database, and pathways. In addition to developing these databases from known compounds, LIPID MAPS has developed a software package called LipidMapsTools for the template-based combinatorial enumeration of virtual compound libraries for lipids (Sud *et al.*, 2012).

Examples of Findings Using Lipidomics Approaches

There have now been such a large number of studies using these types of lipidomics approaches that one should be able to locate publications in their particular area of interest quite easily using online search engines. Therefore, only a few examples will be given here.

CLASS-Type Lipidomic Analysis

The lipidomics studies of the LIPID MAPS Consortium are examples of CLASS-type analyses. Studies using this approach:

- Elucidated changes in the macrophage lipidome upon activation of the cells by Kdo₂-Lipid A (Dennis *et al.*, 2010). These included immediate responses in fatty acid metabolism (increases in eicosanoid synthesis), delayed responses in sphingolipid and sterol biosynthesis, and lipid remodeling for glycerolipids, glycerophospholipids, and prenols. These studies also resulted in numerous 'spin-off' discoveries, including the role of *de novo* sphingolipid biosynthesis in induction of autophagy in these cells (Sims *et al.*, 2010), and the production of 25-hydroxycholesterol and its role in immunomodulation (Shibata *et al.*, 2013).
- Quantified over 500 different molecular species of lipids in six main categories (fatty acyls, glycerolipids, glycerophospholipids, sphingolipids, sterols, and prenols) in human plasma, including detection of 100 subspecies of sphingomyelins (Quehenberger *et al.*, 2010).
- Identified a panel of 20 plasma metabolites that included glycerophospholipids, sphingolipids, and sterols as well as aqueous small molecular weight components that could differentiate between NASH and steatosis in a double-blinded study (Gorden *et al.*, 2015) of samples from individuals with different stages of nonalcoholic fatty liver

disease (NAFLD), which includes steatosis, nonalcoholic steatohepatitis (NASH), and cirrhosis). The lipids included compounds with the novel 1-deoxy-sphingoid base backbone, which have been associated with diabetes.

Shotgun Lipidomics

Use of shotgun lipidomics is illustrated by these studies:

- A targeted shotgun lipidomics analysis of lipid extracts from postmortem brains of subjects with preclinical Alzheimer's disease (i.e., persons who were cognitively normal at death, but displayed neuropathology), found sulfatide levels were significantly lower and ethanolamine glycerophospholipids were marginally lower in subjects with preclinical Alzheimer's disease compared to those without neuropathology (Cheng *et al.*, 2013).
- A study of lipids extracted from 685 plasma samples found that of the levels of individual species of cholesterol esters, lysophosphatidylcholines, phosphatidylcholines, phosphatidylethanolamines, sphingomyelins, and triacylglycerols were associated with cardiovascular disease over a 10-year observation period (Stegemann *et al.*, 2014).

Tissue-Imaging Mass Spectrometry

Tissue-imaging mass spectrometry has been used to:

- Determine the ability of DESI mass spectrometry to identify and differentiate breast tumors from normal breast tissue (Calligaris *et al.*, 2014), which found that several fatty acids (including oleic acid) were more abundant in the cancerous tissue than in normal tissues.
- Profile lipids and proteins in clear cell renal cell carcinoma using MALDI imaging mass spectrometry (Jones *et al.*, 2014), which found a refined panel of 26 proteins and 39 lipid species that could either distinguish tumor from nontumor tissues, or tissues from recurrent disease progressors from nonrecurrent disease individuals.
- Examine coronal sections in a mouse model of chronic alcohol exposure using matrix-implanted laser-desorption/ionization mass spectrometry imaging (MILDI-MSI) which involves implanting silver nanoparticulate matrices followed by focused laser desorption (Roux *et al.*, 2014). This found qualitative regional changes in ion intensities for ceramide and sphingomyelin pairs as a function of group and brain region, in cortex, striatum, accumbens, habenula, and hippocampus, and highly correlated changes in certain white matter regions.
- Establish the spatial organization of lipids in the human retina and optic nerve by MALDI imaging mass spectrometry (Zemski Berry *et al.*, 2014), which found that triacylglycerols were highly localized in accessory tissue, whereas sulfatide and plasmalogen glycerophosphoethanolamines with a monounsaturated fatty acid were enriched in the optic nerve. Several lipids were associated solely with the inner retina, photoreceptors, or retinal pigment epithelium.

Summary

Great strides have been made in lipidomics analysis over the past decade. While only a fraction of the entire lipidome can currently be quantified (somewhat) easily, development of more sophisticated mass spectrometers continues at a break-neck pace and investigators are as quickly applying them to lipids. It will still be challenging to interpret the large amounts of information this technology provides, but that is the inevitable price – pleasant prospect – of having lipids join the ‘omics’ fields.

See also: Horizontal Integration: Omics: Metabolomics in Cell Biology. Molecular Principles, Components, Technology, and Concepts: Lipids: Cholesterol and Other Steroids; Glycolipids; Lipid Signaling; Synthesis and Structure of Glycerolipids

References

- Baker, P.R., Armando, A.M., Campbell, J.L., Quehenberger, O., Dennis, E.A., 2014. Three-dimensional enhanced lipidomics analysis combining UPLC, differential ion mobility spectrometry, and mass spectrometric separation strategies. *Journal of Lipid Research* 55, 2432–2442.
- Banoub, J.H., El Anead, A., Cohen, A.M., Joly, N., 2010. Structural investigation of bacterial lipopolysaccharides by mass spectrometry and tandem mass spectrometry. *Mass Spectrometry Reviews* 29, 606–650.
- Brown, H.A. (Ed.), 2007. Lipidomics and bioactive lipids: Mass-spectrometry-based lipid analysis. *Methods in Enzymology* 432, 1–387.
- Callender, H.L., Forrester, J.S., Ivanova, P., *et al.*, 2007. Quantification of diacylglycerol species from cellular extracts by electrospray ionization mass spectrometry using a linear regression algorithm. *Analytical Chemistry* 79, 263–272.
- Calligaris, D., Caragacianu, D., Liu, X., *et al.*, 2014. Application of desorption electrospray ionization mass spectrometry imaging in breast cancer margin analysis. *Proceedings of the National Academy of Sciences of the United States of America* 111, 15184–15189.
- Castro-Perez, J., Roddy, T.P., Nibbering, N.M., *et al.*, 2011. Localization of fatty acyl and double bond positions in phosphatidylcholines using a dual stage CID fragmentation coupled with ion mobility mass spectrometry. *Journal of the American Society for Mass Spectrometry* 22, 1552–1567.
- Cheng, H., Wang, M., Li, J.L., Cairns, N.J., Han, X., 2013. Specific changes of sulfatide levels in individuals with pre-clinical Alzheimer's disease: An early event in disease pathogenesis. *Journal of Neurochemistry* 127, 733–738.
- Dennis, E.A., 2009. Lipidomics joins the omics evolution. *Proceedings of the National Academy of Sciences of the United States of America* 106, 2089–2090.
- Dennis, E.A., Deems, R.A., Harkewicz, R., *et al.*, 2010. A mouse macrophage lipidome. *Journal of Biological Chemistry* 285, 39976–39985.
- Fahy, E., Cotter, D., Byrnes, R., *et al.*, 2007a. Bioinformatics for lipidomics. *Methods in Enzymology* 432, 247–273.
- Fahy, E., Cotter, D., Sud, M., Subramaniam, S., 2011. Lipid classification, structures and tools. *Biochimica et Biophysica Acta* 1811, 637–647.
- Fahy, E., Subramaniam, S., Brown, H.A., *et al.*, 2005. A comprehensive classification system for lipids. *Journal of Lipid Research* 46, 839–861.
- Fahy, E., Subramaniam, S., Murphy, R.C., *et al.*, 2009. Update of the LIPID MAPS comprehensive classification system for lipids. *Journal of Lipid Research* 50 (Suppl.), S9–14.
- Fahy, E., Sud, M., Cotter, D., Subramaniam, S., 2007b. LIPID MAPS online tools for lipid research. *Nucleic Acids Research* 35, W606–W612.
- Farwanah, H., Kolter, T., 2012. Lipidomics of glycosphingolipids. *Metabolites* 2, 134–164.
- Garrett, T.A., Guan, Z., Raetz, C.R., 2007. Analysis of ubiquinones, dolichols, and dolichol diphosphate-oligosaccharides by liquid chromatography-electrospray ionization-mass spectrometry. *Methods in Enzymology* 432, 117–143.
- Gorden, D.L., Myers, D.S., Ivanova, P.T., *et al.*, 2015. Biomarkers of NAFLD progression: A lipidomics approach to an epidemic. *Journal of Lipid Research* 56, 722–736.
- Griffiths, W.J., Wang, Y., Alvelius, G., *et al.*, 2006. Analysis of oxysterols by electrospray tandem mass spectrometry. *Journal of the American Society for Mass Spectrometry* 17, 341–362.
- Guan, Z., Eichler, J., 2011. Liquid chromatography/tandem mass spectrometry of dolichols and polyprenols, lipid sugar carriers across evolution. *Biochimica et Biophysica Acta* 1811, 800–806.
- Han, X., Gross, R.W., 2003. Global analyses of cellular lipidomes directly from crude extracts of biological samples by ESI mass spectrometry: A bridge to lipidomics. *Journal of Lipid Research* 44, 1071–1079.
- Harkewicz, R., Dennis, E.A., 2011. Applications of mass spectrometry to lipids and membranes. *Annual Review of Biochemistry* 80, 301–325.
- Hellmuth, C., Weber, M., Koletzko, B., Peissner, W., 2012. Nonesterified fatty acid determination for functional lipidomics: Comprehensive ultrahigh performance liquid chromatography-tandem mass spectrometry quantitation, qualification, and parameter prediction. *Analytical Chemistry* 84, 1483–1490.
- Henderson, J.C., O'Brien, J.P., Brodbelt, J.S., Trent, M.S., 2013. Isolation and chemical characterization of lipid A from gram-negative bacteria. *Journal of Visualized Experiments*, e50623.
- Hertweck, C., 2009. The biosynthetic logic of polyketide diversity. *Angewandte Chemie, International Edition in English* 48, 4688–4716.
- Holzl, G., Dormann, P., 2007. Structure and function of glycosylcerolipids in plants and bacteria. *Progress in Lipid Research* 46, 225–243.
- Hyötylainen, T., Oresic, M., 2014. Systems biology strategies to study lipidomes in health and disease. *Progress in Lipid Research* 55, 43–60.
- Ishibashi, Y., Kohyama-Koganeya, A., Hirabayashi, Y., 2013. New insights on glucosylated lipids: Metabolism and functions. *Biochimica et Biophysica Acta* 1831, 1475–1485.
- Ito, E., Tominaga, A., Waki, H., *et al.*, 2012. Structural characterization of monosialo-, disialo- and trisialo-gangliosides by negative ion AP-MALDI-QIT-TOF mass spectrometry with MS(n) switching. *Neurochemical Research* 37, 1315–1324.
- John, C., Werner, P., Worthmann, A., *et al.*, 2014. A liquid chromatography-tandem mass spectrometry-based method for the simultaneous determination of hydroxy sterols and bile acids. *Journal of Chromatography A* 1371, 184–195.
- Jones, E.E., Powers, T.W., Neely, B.A., *et al.*, 2014. MALDI imaging mass spectrometry profiling of proteins and lipids in clear cell renal cell carcinoma. *Proteomics* 14, 924–935.
- Kind, T., Wohlgemuth, G., Lee Do, Y., *et al.*, 2009. FiehnLib: Mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Analytical Chemistry* 81, 10038–10048.
- Kishimoto, K., Urade, R., Ogawa, T., Moriyama, T., 2001. Nondestructive quantification of neutral lipids by thin-layer chromatography and laser-fluorescent scanning: Suitable methods for "lipidome" analysis. *Biochemical and Biophysical Research Communications* 281, 657–662.
- Kiiman, M., May, J.C., McLean, J.A., 2011. Lipid analysis and lipidomics by structurally selective ion mobility-mass spectrometry. *Biochimica et Biophysica Acta* 1811, 935–945.
- Kouzel, I.U., Pirkil, A., Pohlentz, G., *et al.*, 2014. Progress in detection and structural characterization of glycosphingolipids in crude lipid extracts by enzymatic phospholipid disintegration combined with thin-layer chromatography immunodetection and IR-MALDI mass spectrometry. *Analytical Chemistry* 86, 1215–1222.
- Lydic, T.A., Busik, J.V., Reid, G.E., 2014. A monophasic extraction strategy for the simultaneous lipidome analysis of polar and nonpolar retina lipids. *Journal of Lipid Research* 55, 1797–1809.
- Maccarone, A.T., Duldig, J., Mitchell, T.W., *et al.*, 2014. Characterization of acyl chain position in unsaturated phosphatidylcholines using differential mobility-mass spectrometry. *Journal of Lipid Research* 55, 1668–1677.
- Matysik, S., Klunemann, H.H., Schmitz, G., 2012. Gas chromatography-tandem mass spectrometry method for the simultaneous determination of oxysterols, plant sterols, and cholesterol precursors. *Clinical Chemistry* 58, 1557–1564.
- McDonald, J.G., Smith, D.D., Stiles, A.R., Russell, D.W., 2012. A comprehensive method for extraction and quantitative analysis of sterols and secosteroids from human plasma. *Journal of Lipid Research* 53, 1399–1409.
- Merrill Jr., A.H., 2011. Sphingolipid and glycosphingolipid metabolic pathways in the era of sphingolipidomics. *Chemical Reviews* 111, 6387–6422.
- Morii, H., Ogawa, M., Fukuda, K., Taniguchi, H., 2014. Ubiquitous distribution of phosphatidylinositol phosphate synthase and archaetidylinositol phosphate synthase in Bacteria and Archaea, which contain inositol phospholipid. *Biochemical and Biophysical Research Communications* 443, 86–90.

- Murphy, R., Merrill, A.H. (Eds.), 2011. Lipidomics and imaging mass spectrometry. *Biochimica et Biophysica Acta* 1811, 635–1000.
- Murphy, R.C., 2015. Tandem Mass Spectrometry of Lipids: Molecular Analysis of Complex Lipids. London: Royal Society of Chemistry.
- Murphy, R.C., Axelsen, P.H., 2011. Mass spectrometric analysis of long-chain lipids. *Mass Spectrometry Reviews* 30, 579–599.
- Murphy, R.C., Hankin, J.A., Barkley, R.M., 2009. Imaging of lipid species by MALDI mass spectrometry. *Journal of Lipid Research* 50 (Suppl.), S317–S322.
- Murphy, R.C., Leiker, T.J., Barkley, R.M., 2011. Glycerolipid and cholesterol ester analyses in biological samples by mass spectrometry. *Biochimica et Biophysica Acta* 1811, 776–783.
- Parry, S., Ledger, V., Tissot, B., *et al.*, 2007. Integrated mass spectrometric strategy for characterizing the glycans from glycosphingolipids and glycoproteins: Direct identification of sialyl Le(x) in mice. *Glycobiology* 17, 646–654.
- Paulick, M.G., Bertozzi, C.R., 2008. The glycosylphosphatidylinositol anchor: A complex membrane-anchoring structure for proteins. *Biochemistry* 47, 6991–7000.
- Pruett, S.T., Bushnev, A., Hagedorn, K., *et al.*, 2008. Biodiversity of sphingoid bases ("sphingosines") and related amino alcohols. *Journal of Lipid Research* 49, 1621–1639.
- Quehenberger, O., Armando, A.M., Brown, A.H., *et al.*, 2010. Lipidomics reveals a remarkable diversity of lipids in human plasma. *Journal of Lipid Research* 51, 3299–3305.
- Raetz, C.R., Guan, Z., Ingram, B.O., *et al.*, 2009. Discovery of new biosynthetic pathways: The lipid A story. *Journal of Lipid Research* 50 (Suppl.), S103–S108.
- Reis, A., Rudnitskaya, A., Blackburn, G.J., *et al.*, 2013. A comparison of five lipid extraction solvent systems for lipidomic studies of human LDL. *Journal of Lipid Research* 54, 1812–1824.
- Rombouts, Y., Alibaud, L., Carrere-Kremer, S., *et al.*, 2011. Fatty acyl chains of *Mycobacterium marinum* lipooligosaccharides: Structure, localization and acylation by PapA4 (MMAR_2343) protein. *Journal of Biological Chemistry* 286, 33678–33688.
- Roux, A., Muller, L., Jackson, S.N., *et al.*, 2014. Chronic ethanol consumption profoundly alters regional brain ceramide and sphingomyelin content in rodents. *ACS Chemical Neuroscience* 6, 247–259.
- Schiesel, S., Lammerhofer, M., Lindner, W., 2010. Quantitative LC-ESI-MS/MS metabolic profiling method for fatty acids and lipophilic metabolites in fermentation broths from beta-lactam antibiotics production. *Analytical and Bioanalytical Chemistry* 397, 147–160.
- Shaner, R.L., Allegood, J.C., Park, H., *et al.*, 2009. Quantitative analysis of sphingolipids for lipidomics using triple quadrupole and quadrupole linear ion trap mass spectrometers. *Journal of Lipid Research* 50, 1692–1707.
- Shibata, N., Carlin, A.F., Spann, N.J., *et al.*, 2013. 25-Hydroxycholesterol activates the integrated stress response to reprogram transcription and translation in macrophages. *Journal of Biological Chemistry* 288, 35812–35823.
- Sims, K., Haynes, C.A., Kelly, S., *et al.*, 2010. Kdo2-lipid A, a TLR4-specific agonist, induces de novo sphingolipid biosynthesis in RAW264.7 macrophages, which is essential for induction of autophagy. *Journal of Biological Chemistry* 285, 38568–38579.
- Sims, K.H., Tytler, E.M., Tipton, J., *et al.*, 2014. Avanti lipid tools: Connecting lipids, technology, and cell biology. *Biochimica et Biophysica Acta* 1841, 1038–1048.
- Spener, F., Legarde, M., Geloën, A., Record, M., 2003. What is lipidomics? *European Journal of Lipid Science and Technology* 105, 481–482.
- Stegemann, C., Pechlaner, R., Willeit, P., *et al.*, 2014. Lipidomics profiling and risk of cardiovascular disease in the prospective population-based Bruneck study. *Circulation* 129, 1821–1831.
- Subramaniam, S., Fahy, E., Gupta, S., *et al.*, 2011. Bioinformatics and systems biology of the lipidome. *Chemical Reviews* 111, 6452–6490.
- Sud, M., Fahy, E., Subramaniam, S., 2012. Template-based combinatorial enumeration of virtual compound libraries for lipids. *Journal of Cheminformatics* 4, 23.
- Sullards, M.C., Liu, Y., Chen, Y., Merrill Jr., A.H., 2011. Analysis of mammalian sphingolipids by liquid chromatography tandem mass spectrometry (LC-MS/MS) and tissue imaging mass spectrometry (TIMS). *Biochimica et Biophysica Acta* 1811, 838–853.
- Takats, Z., Wiseman, J.M., Cooks, R.G., 2005. Ambient mass spectrometry using desorption electrospray ionization (DESI): Instrumentation, mechanisms and applications in forensics, chemistry, and biology. *Journal of Mass Spectrometry* 40, 1261–1275.
- Thulasiram, H.V., Erickson, H.K., Poulter, C.D., 2007. Chimeras of two isoprenoid synthases catalyze all four coupling reactions in isoprenoid biosynthesis. *Science* 316, 73–76.
- Vranova, E., Coman, D., Griessem, W., 2012. Structure and dynamics of the isoprenoid pathway network. *Molecular Plant* 5, 318–333.
- Wang, M., Han, X., 2014. Multidimensional mass spectrometry-based shotgun lipidomics. *Methods in Molecular Biology* 1198, 203–220.
- Wang, M., Huang, Y., Han, X., 2014. Accurate mass searching of individual lipid species candidates from high-resolution mass spectra for shotgun lipidomics. *Rapid Communications in Mass Spectrometry* 28, 2201–2210.
- Welti, M., 2013. Regulation of dolichol-linked glycosylation. *Glycoconjugate Journal* 30, 51–56.
- Winograd, N., Garrison, B.J., 2010. Biological cluster mass spectrometry. *Annual Review of Physical Chemistry* 61, 305–322.
- Yang, K., Zhao, Z., Gross, R.W., Han, X., 2011. Identification and quantitation of unsaturated fatty acid isomers by electrospray ionization tandem mass spectrometry: A shotgun lipidomics approach. *Analytical Chemistry* 83, 4243–4250.
- Yu, R.K., Yanagisawa, M., Ariga, T., 2007. Glycosphingolipid structures. In: Kamerling, J.P. (Ed.), *Comprehensive Glycoscience. From Chemistry to Systems Biology*. Oxford: Elsevier.
- Zemski Berry, K.A., Gordon, W.C., Murphy, R.C., Bazan, N.G., 2014. Spatial organization of lipids in the human retina and optic nerve by MALDI imaging mass spectrometry. *Journal of Lipid Research* 55, 504–515.
- Zhang, J., Li, C., Yu, G., Guan, H., 2014. Total synthesis and structure-activity relationship of glycosylcerolipids from marine organisms. *Marine Drugs* 12, 3634–3659.

Relevant Websites

<http://lipidlibrary.aocs.org/index.html>
American Oil Chemistry Society (AOCS) Lipid Library.

<http://www.cyberlipid.org/index.htm>
Cyberlipid.

<http://lipidbank.jp/wiki/Category:LB>
Lipid Bank Wiki.

<http://www.lipidmaps.org>
Lipid Maps.

www.sphingomap.org
SphinGOMAP.

Synthesis and Structure of Glycerolipids

W Dowhan, University of Texas Medical School at Houston, Houston, TX, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Archaea The kingdom of life composed of single cell microbes lacking a nucleus and internal membrane-enclosed organelles. Although classified as prokaryotes, they are distinct from bacteria in that they contain unique membrane lipids and metabolic pathways of both bacteria and eukaryotes. Many are extremophiles living in harsh environments.

Chirality A system displays chirality if it cannot be superimposed on the mirror image of itself such as the human hand.

Eubacteria The kingdom of life composed of true bacteria as opposed to Archaea. They are single cell microbes lacking a nucleus and internal membrane-enclosed organelles, which distinguishes them from eukaryotes.

Eukaryote The kingdom of life composed of single or multiple cells containing membrane-enclosed organelles including a membrane-enclosed nucleus with defined chromosomes.

Gram-negative bacteria This subgroup of bacteria does not stain with a reagent developed by Hans Gram specific for the cell wall of most bacteria. In Gram-negative bacteria the cell wall, which is exterior to the inner cytoplasmic membrane, is protected by a second outer membrane structure.

Gram-positive bacteria This subgroup of bacteria does stain with a reagent developed by Hans Gram specific for the cell wall of most bacteria. Gram-positive bacteria lack the second outer membrane structure of Gram-negative bacteria thus have an exposed cell wall.

Introduction

Glycerolipid formally refers to molecules with glycerol esterified with long-chain fatty acids at anyone or all of the hydroxyl groups (Figure 1). Triacylglycerol is the storage form of lipids in adipose tissue. Diacylglycerol in trace amounts acts as a metabolic signaling molecule at several locations within cells (Almena and Mérida, 2011). Diacylglycerols with a mono-, di-, or trisaccharide in glycosidic linkage to the free hydroxyl form a group of glycolipids found in eubacteria and plants.

Glycerophospholipids (hereafter referred to as phospholipids) contain fatty acids at the central and one of the distal hydroxyls of glycerol with the remaining hydroxyl in ester linkage to either phosphate or a phosphate in phosphoester linkage to a variety of alcohols as denoted by the X. The

chirality of the glycerophosphate backbone of phospholipids in all forms of life except Archaea is the same with the simplest phospholipid (phosphatidic acid) formally named 1,2 diacyl-*sn*-glycerol-3-phosphate. All glycerolipids in Archaea contain the mirror image configuration of *sn*-glycerol-1-phosphate with long-chain hydrocarbons at the *sn*-2 and *sn*-3 positions. In Archaea the hydrocarbons are made up of 20–40 branched-chain isoprenoid (phytanyl) units, which are in ether rather than ester linkage to the glycerol backbone. This article will cover only phospholipids found in eubacteria, the yeast *Saccharomyces cerevisiae* and mammalian cells. In these organisms the hydrocarbon components of lipids are generally long-chain fatty acids in ester linkage except for the ether lipid subclass of phospholipids (McIntyre *et al.*, 2008) synthesized in mammalian peroxisomes. The hydrocarbons of fatty acids

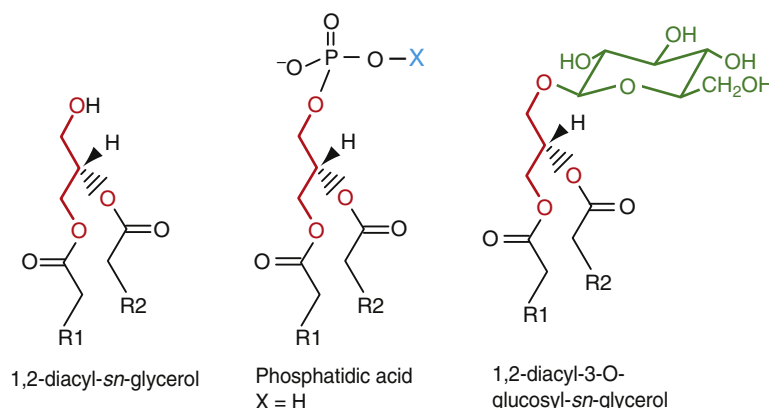


Figure 1 Basic structure of various glycerolipids. The glycerol carbon backbone (red) is depicted as the conformational isomer found in eubacteria and eukaryotes. R1 and R2 indicate long-chain hydrocarbons, which vary in length, number of double bonds, and degree of branching. The various classes of phospholipid vary in the alcohol in phosphoester linkage at X in the middle diagram. The saccharide (green) is generally glucose (as shown) or galactose and can be up to three units in length.

in extensively studied Gram-negative bacteria, yeast, and mammalian cells are not branched and contain single or multiple unsaturated bonds in over half the chains. In extensively studied Gram-positive bacterial the hydrocarbons chains are usually fully saturated but are branched in nature. Plants and mammalian cells use similar pathways for the synthesis of phospholipids common to both species. Plants contain additional glycerolipids not found in other forms of life.

Synthesis of Phosphatidic Acid

Synthesis of all glycerolipids except the plasmalogens begins with the *sn*-glycerol-3-phosphate derived from direct phosphorylation of glycerol with ATP or the reduction of dihydroxyacetone phosphate. In plasmalogen biosynthesis the free hydroxyl (*sn*-1 position) of dihydroxyacetone phosphate is first acylated with a long-chain fatty acid followed by an exchange reaction where the fatty acid is replaced by a long-chain fatty alcohol resulting in an ether linkage (McIntyre *et al.*, 2008). The ketone at the *sn*-2 position is then reduced to an alcohol and acylated. Although there are differences in the fatty acid donors among organisms, the basic pathway is conserved (Figure 2). Fatty acid synthesis in all organisms proceeds via sequential addition of acetyl Coenzyme A (CoA) units to an acyl chain covalently attached to acyl carrier protein (ACP). In eukaryotic cells ACP is a tightly associated component of the fatty acid synthase complex so that transfer to CoA releases the completed fatty acyl chain. In eukaryotes acyl-CoA derivatives are the source of fatty acid for phosphatidic acid biosynthesis. In eubacteria *de novo* fatty acid synthesis ends with the carboxyl moiety in thioester bond to free ACP, which is not part of a larger synthase complex.

The most common initial acylation step in eubacteria proceeds by first conversion of the fatty acyl-ACP to an acyl phosphate followed by transfer of the acyl group to the *sn*-1 position of glycerol to form lysophosphatidic acid (Yao and Rock, 2013); the *plsX* and *plsY* gene products, respectively, catalyze these two steps. Virtually all of the genes referred to in this article have been mapped, cloned, and sequenced. *Escherichia coli* and closely related Gram-negative bacteria are an exception in that in addition to the above gene products they contain the *plsB* gene product (Larson *et al.*, 1980), which directly utilizes either CoA or ACP derivatives of fatty acids as donors in the first step. The *plsC* gene product catalyzes the second acylation step using ACP linked fatty acids in all bacteria with only *E. coli* and related bacteria also utilizing CoA fatty acid derivatives (Coleman, 1992). All of the enzymes of these pathways are integral to the cytoplasmic membrane of bacteria.

The same basic pathway is utilized by yeast, such as *S. cerevisiae* (Henry *et al.*, 2012), and mammalian cells (Vance and Vance, 2008) to synthesis phosphatidic acid in the endoplasmic reticulum and outer mitochondrial membranes (Figure 2). All acylation steps in eukaryotic cells utilize acyl-CoA derivations since such cells do not contain free ACP. As is generally the case in eukaryotic cells, there is redundancy in the numbers of genes products in each cell, the significance of which has not been fully clarified. In mammalian cells the *GPAT1* and 2 gene products localized to the outer mitochondrial membrane and the *GPAT3* and 4 gene products in the endoplasmic reticulum membrane acylate the *sn*-1 position. The second acylation step is carried out by any of five gene products (*AGPAT1-5*) in the endoplasmic reticulum. In yeast gene products localized to the endoplasmic reticulum catalyze the first (*SCT1* or *GPT2*) and second (*SLC1* or *ALE2*) acylation step.

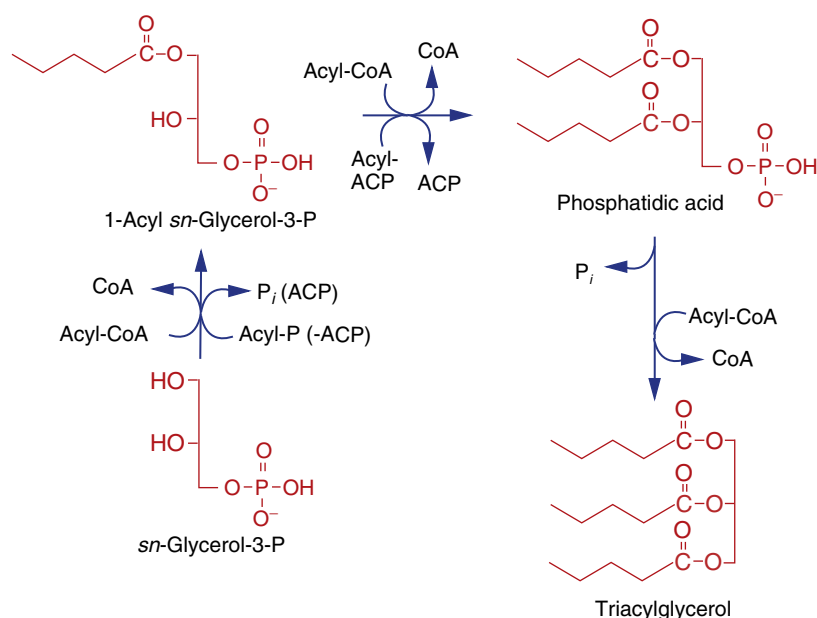


Figure 2 Biosynthetic pathway for phosphatidic acid and triacylglycerol. Synthesis begins in the lower left corner and utilizes either coenzyme A (CoA), phosphate (P_i) or ACP derivatives of long-chain fatty acids as discussed in the text. Formation of triacylglycerol in eukaryotes involves first remove of phosphate followed by acylation at the *sn*-3 position.

In all organisms phosphatidic acid can be synthesized from diacylglycerol by phosphorylation utilizing ATP. There are multiple processes that generate diacylglycerol within all bacteria and specific organelles within eukaryotic cells. Particularly in the eukaryotic cells where diacylglycerol is a metabolic signaling molecule, mechanisms are required for its rapid generation and removal. Phosphatidic acid is also an important metabolic signaling molecule in eukaryotic cells (Cai *et al.*, 2009). The interconversion of diacylglycerol and phosphatidic acid has important regulatory consequences.

Synthesis of Phospholipids in Bacteria

There are many similarities between eubacterial and eukaryotic pathways of phospholipid biosynthesis and many of the enzymes show high degrees of sequence homology. However, as expected there are many differences. Although Archaea will not be covered in any detail, the pathways resulting in different alcohols at the 'X' position of phosphatidic acid are similar between eubacteria and Archaea (Boucher, 2007).

Synthesis of Amine-Containing Phospholipids

The phospholipids containing an amino group are phosphatidylethanolamine, phosphatidylcholine, and phosphatidylserine where the X position of phosphatidic acid shown in Figure 1 is replaced by ethanolamine, choline, or serine, respectively. The greatest diversity in pathways between bacteria and eukaryotic cells is seen in the synthesis of these phospholipids.

Although the pathways are essentially the same in all bacteria, the best studied are those in the Gram-negative bacterium *E. coli* initially by Eugene P. Kennedy's laboratory in the 1960s. In the following years all of the genes responsible for this pathway were identified, cloned, and sequenced and most of the gene products have been purified and extensively studied (Figure 3). The pathway begins with the synthesis of the common precursor, CDP-diacylglycerol, to all of the major phospholipids found in bacteria (Carter, 1968; Sparrow and Raetz, 1985). The formation of the high-energy pyrophosphate bond in the reaction of phosphatidic acid with CTP provides the energy and driving force to assure that the equilibrium lies in the direction of synthesis of the final products. Displacement of CMP by serine results in conversion of a pyrophosphate bond to a lower energy phosphoester bond with the equilibrium lying far to the formation of phosphatidylserine (Kanfer and Kennedy, 1962; Raetz, 1975, 1976; Larson and Dowhan, 1976). Final decarboxylation of phosphatidylserine to form phosphatidylethanolamine with the release of CO₂ again assures completion of the reaction (Dowhan *et al.*, 1974; Hawrot and Kennedy, 1976; Li and Dowhan, 1988, 1990). Phosphatidylethanolamine ranges from 70 to 80% of the total phospholipid of *E. coli*; virtually all of the phospholipids are found as components of the membrane lipid bilayer with no significant amounts found in the cytoplasm. The precursors to phosphatidylethanolamine are found in amounts of 1% or less.

All the enzymes of phospholipid synthesis in *E. coli*, except the phosphatidylserine synthase, are integral membrane

proteins of the inner cytoplasmic membrane. They are only released from the membrane by disruption of the lipid bilayer by detergents. On the other hand, phosphatidylserine synthase is loosely associated as a peripheral membrane protein easily displaced by high-ionic strength buffers (Louie and Dowhan, 1980). In Gram-positive bacteria, the phosphatidylserine synthase is an integral membrane protein and displays no sequence homology to the *E. coli* enzyme (Matsumoto, 1997).

E. coli and closely related bacteria do not contain phosphatidylcholine, which is present in high amounts in all eukaryotic cells. However, more distant Gram-negative bacteria (Sohlenkamp *et al.*, 2003) contain phosphatidylcholine. One mode of synthesis is by trimethylation of phosphatidylethanolamine by processes similar to that observed in eukaryotic cells (see Figure 6). A second mode is displacement of CMP from CDP-diacylglycerol by choline catalyzed by the *pcs* gene product via a unique bacterial pathway resulting in the formation of phosphatidylcholine (see Figure 4). Both pathways are energetically favorable resulting in high levels of phosphatidylcholine in these bacteria.

Synthesis of Non-Amine-Containing Phospholipids

Beginning with CDP-diacylglycerol, phosphatidylglycerol is formed in two steps by first displacement of CMP by *sn*-glycerol-3-phosphate to form phosphatidylglycerophosphate followed by dephosphorylation of this product to form phosphatidylglycerol (Chang and Kennedy, 1967a,b). Again the equilibrium lies far from completion for these two reactions. It should also be noted that the unacylated glycerol-3-phosphate is the *sn*-1 isomer in contrast to the lipid backbone *sn*-glycerol-3-phosphate. There are three genes, *pgpABC* (Icho and Raetz, 1983; Funk *et al.*, 1992; Lu *et al.*, 2011), that encode phosphatases but only one gene, *pgsA* (Hirabayashi *et al.*, 1976; Ohta *et al.*, 1981) that encodes the synthase. The phosphatases reside in different cellular compartments such as the inner and outer membranes of *E. coli* and their active sites may be facing different sides of their respective membranes. These phosphatases also act on other lipid phosphates. The *pgpC* gene product appears to be the major biosynthetic enzyme. These phosphatases have numerous homologs in other bacteria. The *pgsA* gene product is highly homologous across the bacterial kingdom, but shows no homology to its eukaryotic counterpart.

The final step in this pathway is the formation of the unique phospholipid cardiolipin by the condensation of two phosphatidylglycerol molecules with the release of glycerol (Hirschberg and Kennedy, 1972; Pluschke *et al.*, 1978; Ohta *et al.*, 1985; Hiraoka *et al.*, 1991; Guo and Tropp, 2000; Tan *et al.*, 2012). Since the reaction involves an interchange among phosphodiester bonds of near equal energy, this reaction is reversible. The result is that phosphatidylglycerol comprises about 20% and cardiolipin about 5–10% of the total phospholipid. In fact in stationary phase where little new net synthesis of phospholipids occurs, cardiolipin becomes the predominate lipid of this branch of the pathway. Again there are multiple genes encoding cardiolipin synthases in all bacteria. Of the three *clsABC* gene products, *clsA* appears to be the primary biosynthetic enzyme and catalyzes the condensation

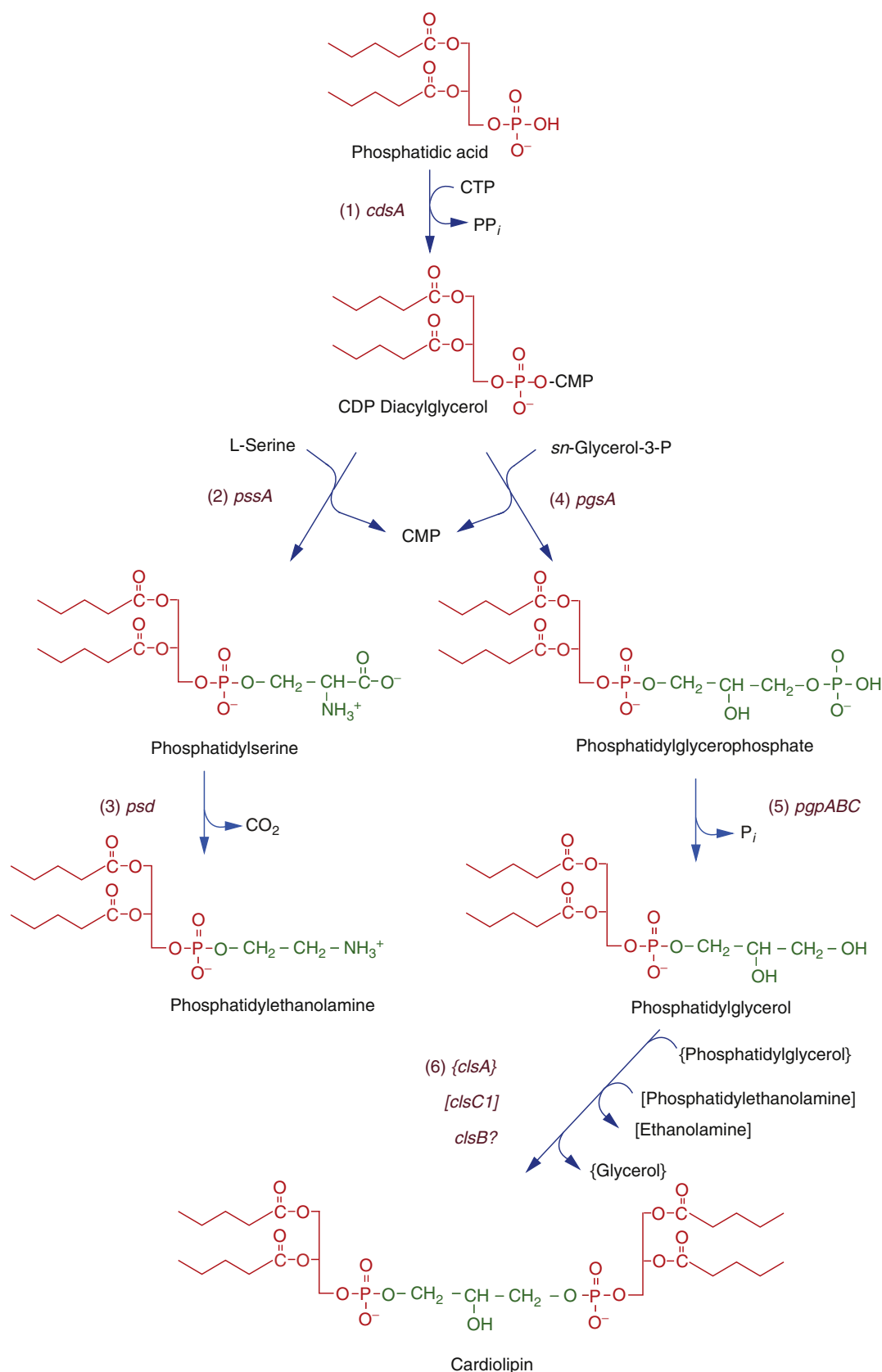


Figure 3 Biosynthetic pathway for phospholipids in *E. coli*. Each step of the pathway is catalyzed by the following enzyme encoded by the genes indicated next to each number: (1) CDP-diacylglycerol synthase; (2) phosphatidylserine synthase; (3) phosphatidylserine decarboxylase; (4) phosphatidylglycerophosphate synthase; (5) phosphatidylglycerophosphate phosphatases; and (6) cardiolipin synthases. The phosphatidyl donor for the *clsB* gene product has not been fully characterized but may also be phosphatidylethanolamine. Reprinted from Dowhan, W., 2013. A retrospective: Use of *Escherichia coli* as a vehicle to study phospholipid synthesis and function. *Biochimica et Biophysica Acta* 1831, 471–494, with permission from Elsevier B.V.

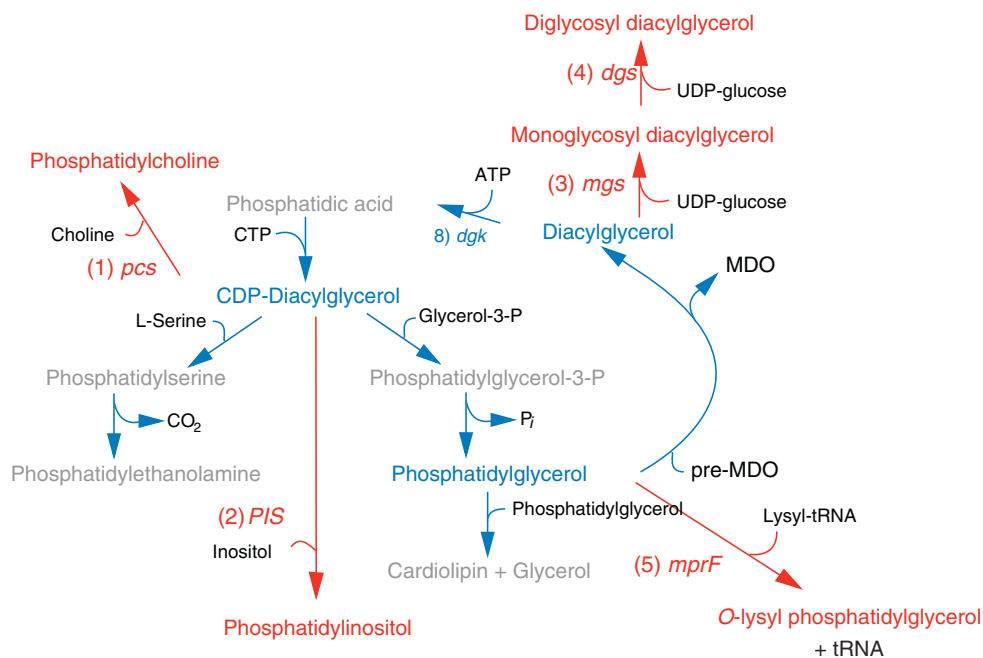


Figure 4 Synthesis of foreign lipids in *E. coli*. The pathways found in *E. coli* are noted by blue arrows. The following enzymes with their respective genes named and source indicated carry out (red arrows) the following steps: (1) phosphatidylcholine synthase (*Legionella pneumophila*) (Conover *et al.*, 2008); (2) phosphatidylinositol synthase (*S. cerevisiae*) (Xia and Dowhan, 1995); (3) monoglucosyl diacylglycerol synthase (*A. laidlawii*) (Xie *et al.*, 2006); (4) diglucosyl diacylglycerol synthase (*A. laidlawii*) (Wikstrom *et al.*, 2009); and (5) lysyl t-RNA: phosphatidylglycerol lysine transferase (*S. aureus*) (Oku *et al.*, 2004). Adapted from Dowhan, W., 2013. A retrospective: Use of *Escherichia coli* as a vehicle to study phospholipid synthesis and function. *Biochimica et Biophysica Acta* 1831, 471–494 and reprinted with permission from Elsevier B.V.

of two molecules of phosphatidylglycerol (Hirschberg and Kennedy, 1972). The other two enzymes form cardiolipin by condensation of phosphatidylglycerol with phosphatidylethanolamine with the release of ethanolamine (Tan *et al.*, 2012). While *clsA* is expressed under all growth conditions, the other two genes are not. The significance of this multiplicity of enzymatic activities is not yet understood.

Gram-positive bacteria and bacteria lacking a cell wall such as *Acholeplasma laidlawii* contain the completely neutral glycolipids with mono-, di-, or trisaccharides at the *sn*-3 position of glycerol (see Figure 1). These are made by successive transglycosylation reactions first to diacylglycerol followed by extension of the number of sugars with UDP-sugars as the donors (Dahlqvist *et al.*, 1992). These glycerolipids comprise 20–40% of the membrane lipids of these organisms. Not positively charged phospholipids are also found in many Gram-positive bacteria such as *Staphylococcus aureus* (Oku *et al.*, 2004). These are formed by the addition of an amino acid to the free hydroxyls of phosphatidylglycerol using an amino acid charged t-RNA as a donor (see Figure 4).

Viable strains of *E. coli* have been made completely devoid of gene products at each step of phospholipid synthesis after the synthesis of CDP-diacylglycerol (Dowhan, 2009). These mutants still form efficient membrane bilayers that maintain cell integrity but lack specific phospholipids. Each displays a set of phenotypes that have been correlated with a requirement for the missing lipids. These mutants have been used to implicate specific lipids in electron transport and energy production, protein translocation across and insertion into

membranes, the topological organization of membrane protein transmembrane domains, initiation of DNA replication, cell division, chemotaxis, and formation of lipid domains in cell membranes. Domains enriched in specific phospholipids appear to be sites for organization of molecular machines involved in DNA replication and cell division (Mileyskovskaya and Dowhan, 2005). In addition foreign lipids have been synthesized in *E. coli* by introduction of genes responsible for synthesis of these lipids in other organisms (Dowhan and Bogdanov, 2012). Figure 4 summarizes the lipids that have been introduced into *E. coli*. Such strains have assisted in understanding the properties of lipids that support various cell processes defective in lipid mutants of *E. coli*.

Synthesis of Phospholipids in Eukaryotes

Phospholipid synthesis in mammalian cells (Vance and Vance, 2008) and the yeast *S. cerevisiae* (Henry *et al.*, 2012) have been extensively detailed both biochemically and genetically.

Synthesis of Amine-Containing Phospholipids

The synthesis of amine-containing phospholipids in mammalian cells is vastly different from that observed in bacteria. However, yeast possess pathways of both mammalian cells and bacteria. Yeast synthesis of phosphatidylserine and phosphatidylethanolamine occurs via the bacterial pathway

(Figure 3) by a membrane-associated enzyme (*CHO1/PSS1* gene product) in the endoplasmic reticulum followed by transfer to the inner mitochondrial membrane (*PSD1* product) and to Golgi/vacuole membranes (*PSD2* gene product) for decarboxylation to phosphatidylethanolamine (Clancey *et al.*, 1993; Trotter *et al.*, 1993; Voelker, 1997). The two decarboxylases show limited homology, but the mitochondrial-localized *PSD1* gene product shows significant homology with the *E. coli* enzyme. How phosphatidylserine and phosphatidylethanolamine traffic between the mitochondria and the

endoplasmic reticulum is an area of active investigation (Osman *et al.*, 2011).

The pathway for the synthesis of these phospholipids found in all eukaryotic cells is referred to as the Kennedy Pathway (Kennedy, 1956; Kennedy *et al.*, 1956, 1959; Kennedy and Weiss, 1956; Borkenhagen and Kennedy, 1957; Smith *et al.*, 1957; Weiss *et al.*, 1958) in recognition of its founder (Figure 5). The pathway is localized initially to the cytoplasm for the formation of water-soluble precursors and then to the endoplasmic reticulum membrane for the synthesis of the final

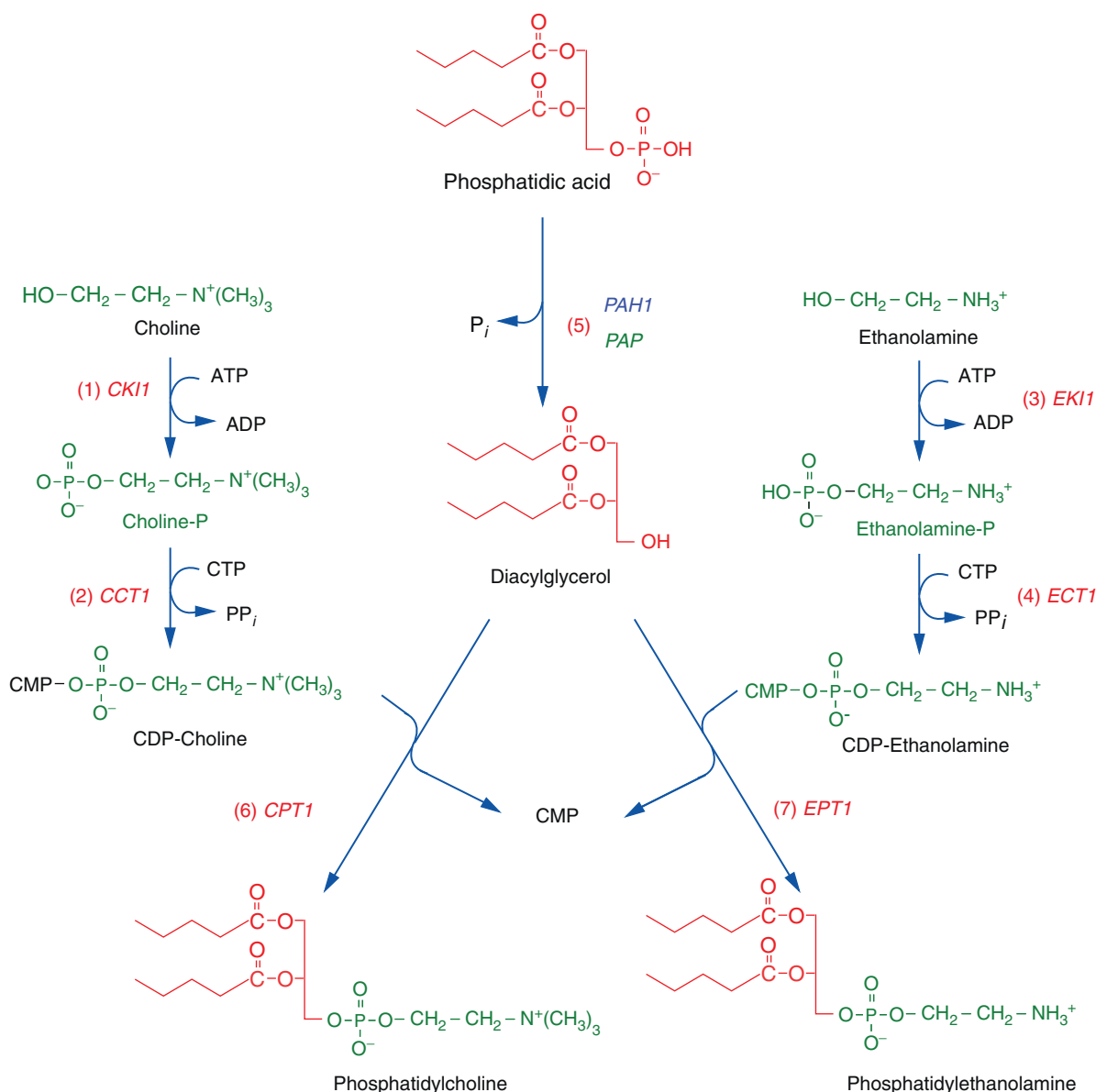


Figure 5 Biosynthetic pathway for amine-containing phospholipids in eukaryotes. The genes encoding the enzymes responsible for each step are noted next to the numbers. Gene names in red are the same for yeast and mammalian cells. Genes in blue or green denote a yeast or mammalian cell gene, respectively. (1) choline kinase; (2) CTP: phosphocholine cytidyltransferase; (3) ethanolamine kinase; (4) CTP: phosphoethanolamine cytidyltransferase; (5) phosphatidic acid phosphatase; (6) CDP-choline:1,2-*sn*-diacylglycerol cholinephosphotransferase; and (7) CDP-ethanolamine:1,2-*sn*-diacylglycerol ethanolaminephosphotransferase. Adapted from Dowhan, W., 2013. A retrospective: Use of *Escherichia coli* as a vehicle to study phospholipid synthesis and function. *Biochimica et Biophysica Acta* 1831, 471–494 and reprinted with permission from Elsevier B.V.

lipid products. How phospholipids are transported from their site of synthesis in the endoplasmic reticulum to the membranes of other organelles is not completely understood. The choline and ethanolamine precursors are first phosphorylated by ATP and further activated by formation of a pyrophosphate bond with CMP to form CDP-derivatives. Specific phosphatases remove the phosphate of phosphatidic acid to form diacylglycerol (Smith *et al.*, 1957), which displaces CMP from the activated intermediates to form the final products phosphatidylcholine and phosphatidylethanolamine. In liver and adipose tissue the diacylglycerol is also further acylated to form triacylglycerol for storage of fat. Studies of the interconversion of diacylglycerol and phosphatidic acid by kinases and phosphatases are an important area of investigation because of the branch point in metabolism between phospholipid synthesis and triacylglycerol storage. Purification of the yeast phosphatidic acid phosphatase followed by cloning of the *PAH1* gene encoding this enzyme (Han *et al.*, 2006) resulted in identifying a class of lipins as mammalian cell phosphatidic acid phosphatases (*PAP1-3*). This was a critical discovery because mice mutants in *lipin-1* displayed a low amount of liver and adipose triglyceride storage (Phan and Reue, 2005). The yeast discovery provided the means to identify the enzymatic activity of lipins.

All eukaryotes can convert phosphatidylethanolamine into phosphatidylcholine by trimethylation using *S*-adenosyl methionine as a donor (Figure 6). Mammalian cells express two enzymes (*PEMT1* and 2 gene products) that catalyze all three methylations while yeast express one enzyme (*PEM1/CHO2* gene product) catalyzing the first methylation (Ridgway and Vance, 1987; Vance and Ridgway, 1988; Cui *et al.*, 1993) and a second enzyme (*PEM2/OPI3* gene product) catalyzing the last two methylations (Summers *et al.*, 1988; Preitschopf *et al.*, 1993).

In mammalian cells phosphatidylserine is made in the endoplasmic reticulum membrane by two exchange enzymes

(Vance and Steenbergen, 2005) that catalyze the displacement by serine of the head group of either phosphatidylethanolamine (*PSS2* gene product) or phosphatidylcholine (*PSS1* gene product) (Figure 6). Phosphatidylserine is an important lipid in mammalian cells and is found in all of the multiple membranes of the cell. A primary role of phosphatidylserine is to organize protein molecular machines on membrane surfaces. In yeast, phosphatidylserine appears not to be important except as a means of making phosphatidylethanolamine or phosphatidylcholine in the absence of ethanolamine or choline in the growth medium.

The formation of phosphatidylcholine by methylation of phosphatidylcholine occurs primarily in the liver of mammals where high levels of phosphatidylcholine are needed to synthesize the serum lipoproteins necessary for transport of lipids from the liver to peripheral tissue. The pathway shown in Figure 6 is essential for new synthesis of phosphatidylcholine and phosphatidylethanolamine during starvation or fasting where the only source of ethanolamine and choline is through continuous displacement from phospholipids by serine derived from protein breakdown. This became evident in mice mutants lacking the methylation pathway that died within days on a diet lacking these two phospholipid precursors (Walkey *et al.*, 1998).

Synthesis of Non-Amine-Containing Phospholipids

The non-amine-containing phospholipids are phosphatidylinositol, phosphatidylglycerol, and cardiolipin (Figure 7). Phosphatidylinositol is synthesized in the endoplasmic reticulum membrane and is found in all membranes of eukaryotic cells. Phosphatidylglycerol and cardiolipin are synthesized exclusively in the inner mitochondrial membrane by gene products encoded by nuclear genes (as is the case with the *PSD1* gene product). Cardiolipin appears to be exclusively localized to the mitochondrial inner and outer membranes

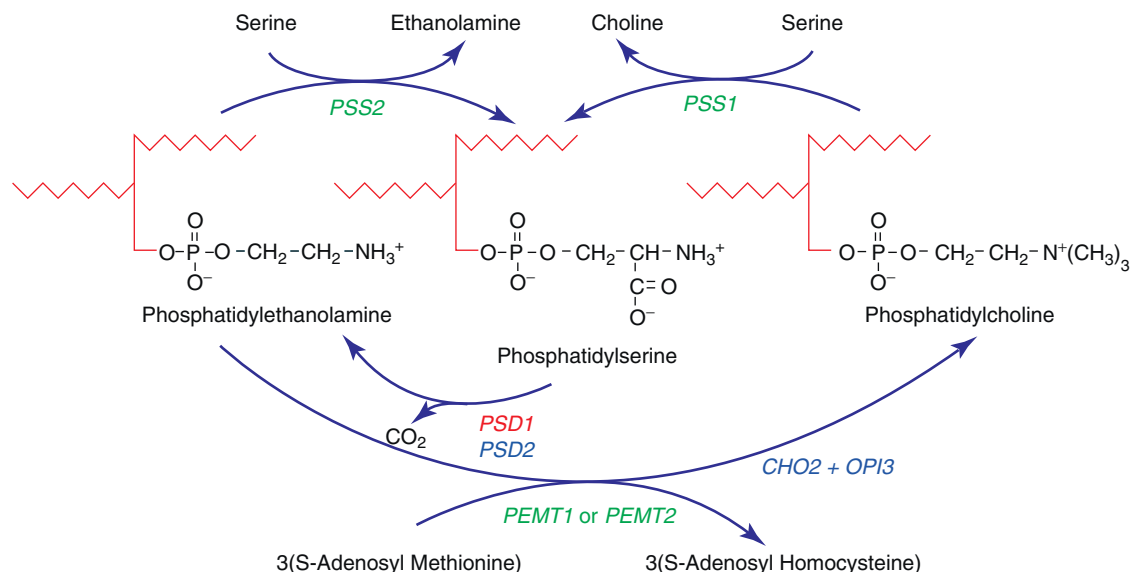


Figure 6 Interconversion of phosphatidylserine, phosphatidylethanolamine, and phosphatidylcholine in yeast and mammalian cells. Genes encoding enzymes for each step are indicated next to the respective arrow. Gene names in red are the same for yeast and mammalian cells. Genes in blue or green denote a yeast or mammalian cell gene, respectively.

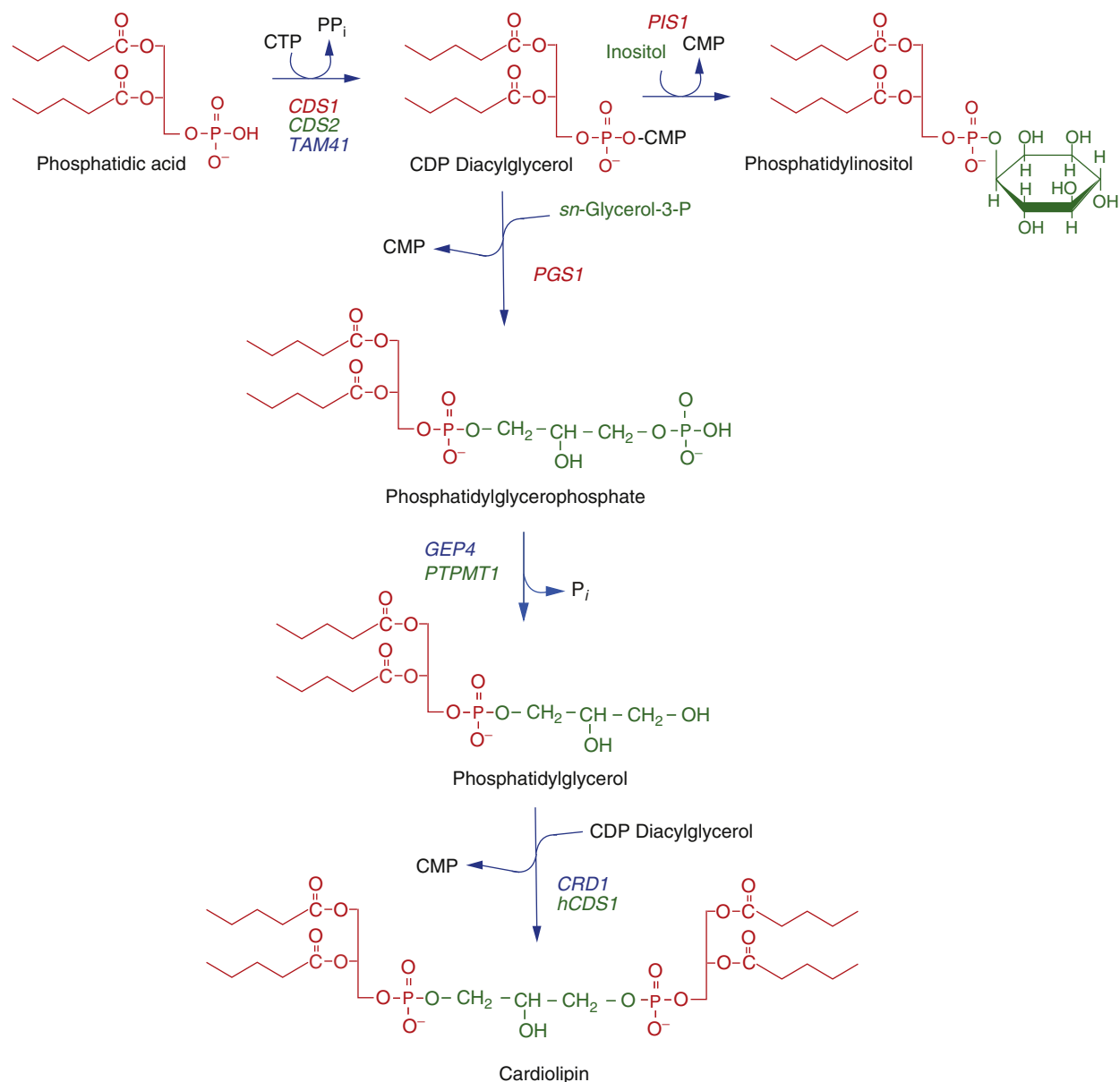


Figure 7 Synthesis of non-amine-containing phospholipids in eukaryotic cells. Genes encoding enzymes for each step are indicated next to the respective arrow. Gene names in red are the same for yeast and mammalian cells. Genes in blue or green denote a yeast or mammalian cell gene, respectively.

where it accounts for about 20% of mitochondrial phospholipids. Phosphatidylglycerol being a precursor to cardiolipin is only about 1% of mitochondrial phospholipids but is also found as an important component of lung surfactant where it is secreted by type II alveoli cells of the lung (Mason and Voelker, 1988). Phosphatidylglycerol suppresses viral and bacterial infection by blocking adhesion of these agents to the lung cell surface (Numata *et al.*, 2013).

CDP-diacylglycerol is synthesized from phosphatidic acid and CTP in yeast endoplasmic reticulum (Kelley and Carman, 1987) and mitochondria by the CDS1 (Shen *et al.*, 1996) and TAM41 (Tamura *et al.*, 2013) gene products, respectively. Mammalian cells appear to only contain CDP-diacylglycerol synthases in the endoplasmic reticulum (Paulus and Kennedy,

1960) encoded by the CDS1 (Weeks *et al.*, 1997) and CDS2 (Halford *et al.*, 1998) gene products. The mammalian cell enzymes are homologous to the *E. coli cdsA* gene (encodes CDP-diacylglycerol synthase) and yeast CDS1 gene products but not the yeast TAM41 gene product.

Phosphatidylinositol synthesis occurs by the displacement of CMP from CDP-diacylglycerol by inositol (Paulus and Kennedy, 1960; Fischl and Carman, 1983). Phosphatidylinositol is an essential lipid in eukaryotic cells. The polyphosphorylated phosphatidylinositols and inositols are important metabolic signaling and regulatory molecules (York *et al.*, 2001). They are also involved in the formation of the glycosylphosphatidylinositol (GPI) linkage that covalently tethers many proteins to the membrane surface (Tsai *et al.*, 2012).

Phosphatidylglycerophosphate synthesis takes place in the inner mitochondrial membrane of yeast (Chang *et al.*, 1998a) and mammalian cells (Kiyasu *et al.*, 1963; Kawasaki *et al.*, 1999) by the *PGS1* gene product following the bacterial pathway. However, there is little homology with the bacterial enzyme. As in bacteria, phosphatidylglycerophosphate is acted on by phosphatases to form phosphatidylglycerol primarily by the yeast *GEP4* (Osman *et al.*, 2010) and mammalian cell *PTPMT1* (Zhang *et al.*, 2011) gene products. There is no homology among the bacterial, yeast, and mammalian cell phosphatases.

Rather than a condensation of two phosphatidylglycerol molecules, cardiolipin synthesis in yeast (Schlame and Greenberg, 1997) and mammalian cells (Hostetler *et al.*, 1972) proceeds by a displacement of CMP from CDP-diacylglycerol by a molecule of phosphatidylglycerol. Due to the conversion of a pyrophosphate bond to a phosphoester bond, the equilibrium lies far in the direction of cardiolipin synthesis and is not reversible thus accounting for the low phosphatidylglycerol and high cardiolipin content of mitochondria. The *hCLS1* gene (Chen *et al.*, 2006; Houtkooper *et al.*, 2006; Lu *et al.*, 2006) and *CRD1* gene (Jiang *et al.*, 1997; Chang *et al.*, 1998b; Tuller *et al.*, 1998) encoding cardiolipin synthase in humans and yeast, respectively, have been identified and cloned.

Yeast null mutants in the *CRD1* gene accumulate the cardiolipin precursor phosphatidylglycerol (Chang *et al.*, 1998b) and are compromised but not completely lacking in oxidative

phosphorylation (Zhang *et al.*, 2002, 2005), the primary source of cellular metabolic energy. Mutants in the *PGS1* gene of yeast (Ostrander *et al.*, 2001b) and mammalian cells (Ohtsuka *et al.*, 1993a,b) display defects in mitochondrial morphology and complete loss of mitochondrial energy production. Therefore, the accumulated phosphatidylglycerol in the *CRD1* mutants appears to partially compensate for the lack of cardiolipin. Lack of or decreased levels of cardiolipin in eukaryotic cells results in failure of individual mitochondrial respiratory complexes to organize into higher order supermolecular complexes with an apparent reduction in their catalytic efficiency (Zhang *et al.*, 2002, 2005; Mileykovskaya *et al.*, 2012; Bazán *et al.*, 2013; Mileykovskaya and Dowhan, 2014).

In eukaryotic cells, composition of the acyl chains of cardiolipin does not reflect that of the general phospholipid pool. Given the four acyl chains of cardiolipin, there are potentially 81 compositional isomers of cardiolipin containing three different fatty acid species (Schlame, 2008). However, the combination of a cardiolipin monodeacylase (removes one fatty acid from cardiolipin) and the *TAZ1* gene product (a phospholipid transacylase that adds a fatty acid from another phospholipid) in all eukaryotic cells rapidly remodels newly synthesized cardiolipin to only a few of the possible isomers (Schlame, 2013). For instances, 90% of heart cardiolipin contains only linoleic acid (18 carbons with two unsaturated bonds). Respiratory function is compromised in pathological conditions or in the inherited disease Barth syndrome

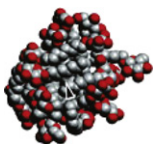
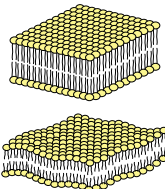
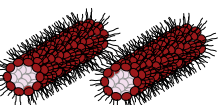
Lipids	Phase	Molecular shape
Lysophospholipids Detergents	 Micellar	 Inverted cone
Phosphatidylcholine Sphingomyelin Phosphatidylserine Phosphatidylinositol Phosphatidylglycerol Phosphatidic acid Cardiolipin Diglucoyldiglyceride	 L β L α Bilayer	 Cylindrical
Phosphatidylethanolamine Cardiolipin – Ca ²⁺ Phosphatidic acid – Ca ²⁺ Phosphatidic acid (pH<3.0) Phosphatidylserine (pH<4.0) Monoglucoyldiglyceride	 Hexagonal (H _{II})	 Cone

Figure 8 Phases and molecular shapes of glycerolipids in solution. For micelles the red depicts hydrophilic groups and the gray depicts hydrophobic groups. For the bilayer structure the yellow depicts the hydrophilic lipid head groups that sandwich the hydrophobic fatty acid chains. In the H_{II} structure the red depicts the hydrophilic head groups of lipids attached to the fatty acid chains. Adapted from Dowhan, W., 2013. A retrospective: Use of *Escherichia coli* as a vehicle to study phospholipid synthesis and function. *Biochimica et Biophysica Acta* 1831, 471–494 and reprinted with permission from Elsevier B.V.

(Schlame *et al.*, 2003) (TAZ1 mutant) where the levels of cardiolipin are abnormally low or where remodeling of the cardiolipin fatty acid composition is defective, respectively.

Roles of Phospholipids in Cell Function

The primary role of phospholipids is to form the lipid bilayer and permeability barrier of all cell membranes (Dowhan, 1997). Other lipid components such as cholesterol and sphingolipids are also important components of the membrane bilayer in eukaryotic cells. If one considers all these lipid classes and the diversity of individual species within each class resulting from the multiple fatty acids they contain, the liposome of cells ranges in the tens to hundreds of thousand individual species. This clearly outstrips the composition of the proteome. Therefore, if only one or two specific phospholipids are required to form a membrane bilayer, why are there so many lipid species (Dowhan, 1997)? The answer lies in the collective physical and chemical properties of a lipid bilayer made up of a mixture of lipid species and specific roles

for individual lipid species in defined cellular processes as partially alluded to earlier.

The ability of a glycerolipid to form a simple lipid bilayer depends mainly on the shape of the space occupied by the hydrophilic head group relative to the hydrophobic fatty acid domain (Cullis and de Kruijff, 1979; Sjolund *et al.*, 1989; Thurmond *et al.*, 1993). Lysophospholipids, phospholipids with short-chain fatty acids and detergents organize into micelles due to their inverted conical shape resulting from small hydrophilic domain relative to a large hydrophobic head group (Figure 8). The predominant fatty acids found in most membrane glycerolipids contain 16–18 carbons. The length and the physical state of the chains determine the width of the bilayer. Some variation in length is tolerated but these are usually minor fatty acid components. Some specialized membranes contain much longer but uniform lengths. The bonds within fatty acids at the *sn*-1 position are generally fully saturated while single and multiple unsaturated bonds are found in fatty acids that populate the *sn*-2 position. Hence, phospholipids and glycolipids with long-chain fatty acids and a cylindrical shape organize into extended sheets that become

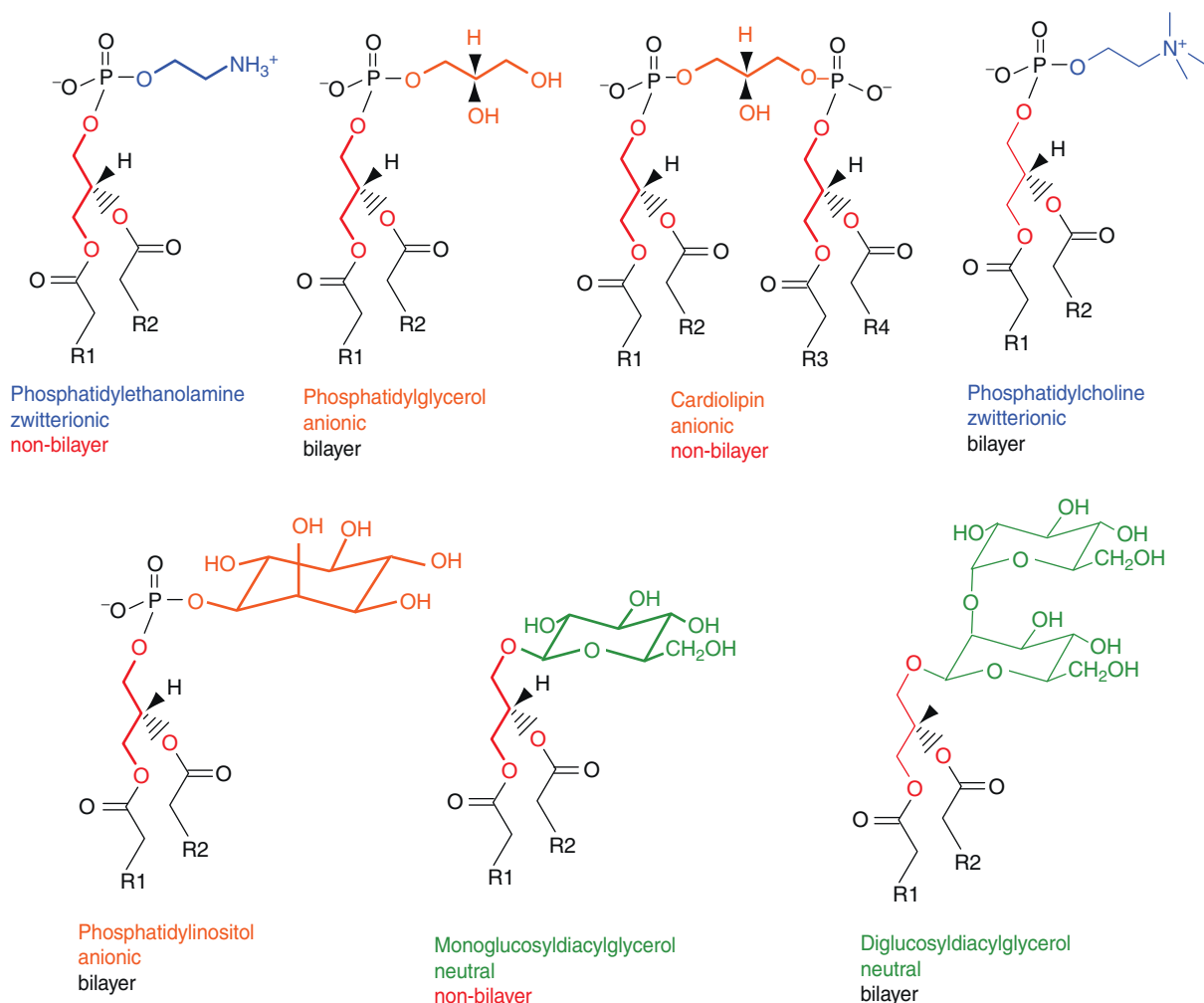


Figure 9 Physical and chemical properties of glycerolipids. The charge character of the lipid head group and the phase properties are noted below each name. See Figure 1 for more details. Adapted from Dowhan, W., 2013. A retrospective: Use of *Escherichia coli* as a vehicle to study phospholipid synthesis and function. *Biochimica et Biophysica Acta* 1831, 471–494 and reprinted with permission from Elsevier B.V.

the two leaflets of the lipid bilayer. Each of the structures shown in [Figure 8](#) minimizes the exposure of hydrophobic surfaces to water.

The fluidity or the ability of lipid molecules or proteins to move laterally within the lipid bilayer increases with the degree of unsaturation of the fatty acid chains and the temperature. Bilayers of individual phospholipid species go through distinct physical transition from an ordered gel phase (L β) to a liquid crystalline phase (L α) as the temperature is raised ([Figure 8](#)). Each of these temperature-dependent transitions is lowered by increased unsaturated fatty acid content. Lipids in which the head group diameter is smaller than that of the hydrophobic domain assume a cone-like structure, which for pure lipid mixtures results in the hexagonal II or similar inverted structures, normally referred to as the non-bilayer phase. Unsaturated bonds in naturally occurring fatty acids are in the *cis* configuration, which increases the space they occupy and their mobility at any given temperature. For non-bilayer prone lipids the degree of unsaturation of their fatty acids and the temperature is a factor determining the transition temperature from the bilayer to non-bilayer phase. As an example all species of phosphatidylethanolamine tend to organize into bilayers at low temperature, but will transition to the non-bilayer phase as the temperature is raised. This transition temperature is inversely related to the degree of unsaturation of the fatty acids.

The mixture of lipids in natural membranes is maintained in the L α phase and the content of bilayer to non-bilayer prone lipids is adjusted either by changing the lipid head group or fatty acid composition in those organisms normally exposed to variations in temperatures. A secondary consequence of the presence of non-bilayer prone lipids such as phosphatidylethanolamine, cardiolipin, or monoglucosyl diacylglycerol in biological membranes is that their conical shape introduces lateral stress within the bilayer, which appears to be important for the function of many membrane-associated proteins. The CDP-choline: diacylglycerol phosphotransferase is a cytoplasmic protein, but in response to an increase in membrane non-bilayer prone lipids, it partially inserts into one leaflet of the lipid bilayer where it is activated to synthesize the bilayer prone phosphatidylcholine. Such association appears to be energetically driven by relief of lateral stress upon membrane association ([Cornell and Northwood, 2000](#); [Kitos et al., 2006](#); [Chong et al., 2014](#)). A summary of the physical and chemical properties of several glycerolipids can be found in [Figure 9](#).

Charge density of the membrane surface is a function of the ratio of net anionic lipids such as phosphatidylserine, phosphatidylinositol, phosphatidylglycerol, and cardiolipin to net neutral lipids such as phosphatidylcholine, phosphatidylethanolamine, and the glycolipids. As an example of the importance of surface charge, the orientation of membrane protein transmembrane domains with respect to the plan of the lipid bilayer is determined by the charge nature of the protein extramembrane domains and the net negative charge density of the lipid head groups exposed at the membrane surface ([Bogdanov et al., 2013](#)). Some protein functions are dependent on a specific lipid head group as well as the extent of unsaturation of the fatty acid chains. An example is the mammalian cardiolipin synthase, which both *in vivo* and *in vitro* requires CDP-diacylglycerols with at least one

unsaturated fatty acid in order to be a substrate while the phosphatidylglycerophosphate synthase is insensitive to the fatty acid composition ([Ostrander et al., 2001a](#)).

See also: Molecular Principles, Components, Technology, and Concepts: Lipids: Cholesterol and Other Steroids

References

- Almena, M., Mérida, I., 2011. Shaping up the membrane: Diacylglycerol coordinates spatial orientation of signaling. *Trends in Biochemical Science* 36, 593–603.
- Bazán, S., Mileyskova, E., Mallampalli, V.K., et al., 2013. Cardiolipin-dependent reconstitution of respiratory supercomplexes from purified *Saccharomyces cerevisiae* complexes III and IV. *Journal of Biological Chemistry* 288, 401–411.
- Bogdanov, M., Dowhan, W., Vitrac, H., 2014. Lipids and topological rules governing membrane protein assembly. *Biochimica et Biophysica Acta* 1843 (8), 1427–1784.
- Borkenhagen, L.F., Kennedy, E.P., 1957. The enzymatic synthesis of cytidine diphosphate choline. *Journal of Biological Chemistry* 227, 951–962.
- Boucher, Y., 2007. Lipids: Biosynthesis, function, and evolution. In: Cavicchioli, R. (Ed.), *Archaea: Molecular and Cellular Biology*. Washington, DC: ASM Press, pp. 341–353.
- Cai, J., Abramovici, H., Gee, S.H., Topham, M.K., 2009. Diacylglycerol kinases as sources of phosphatidic acid. *Biochimica et Biophysica Acta* 1791, 942–948.
- Carter Jr., J.R., 1968. Cytidine triphosphate: Phosphatidic acid cytidyltransferase in *Escherichia coli*. *Journal of Lipid Research* 9, 748–754.
- Chang, S.C., Heacock, P.N., Clancey, C.J., Dowhan, W., 1998a. The *PEL1* gene (renamed *PGS1*) encodes the phosphatidylglycero-phosphate synthase of *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 273, 9829–9836.
- Chang, S.C., Heacock, P.N., Mileyskova, E., Voelker, D.R., Dowhan, W., 1998b. Isolation and characterization of the gene (*CLS1*) encoding cardiolipin synthase in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 273, 14933–14941.
- Chang, Y.Y., Kennedy, E.P., 1967a. Biosynthesis of phosphatidyl glycerophosphate in *Escherichia coli*. *Journal of Lipid Research* 8, 447–455.
- Chang, Y.Y., Kennedy, E.P., 1967b. Phosphatidyl glycerophosphate phosphatase. *Journal of Lipid Research* 8, 456–462.
- Chen, D., Zhang, X.Y., Shi, Y., 2006. Identification and functional characterization of hCLS1, a human cardiolipin synthase localized in mitochondria. *Biochemical Journal* 398, 169–176.
- Chong, S.S., Taneva, S.G., Lee, J.M., Cornell, R.B., 2014. The curvature sensitivity of a membrane-binding amphipathic helix can be modulated by the charge on a flanking region. *Biochemistry* 53, 450–461.
- Clancey, C.J., Chang, S.C., Dowhan, W., 1993. Cloning of a gene (*PSD1*) encoding phosphatidylserine decarboxylase from *Saccharomyces cerevisiae* by complementation of an *Escherichia coli* mutant. *Journal of Biological Chemistry* 268, 24580–24590.
- Coleman, J., 1992. Characterization of the *Escherichia coli* gene for 1-acyl-sn-glycerol-3-phosphate acyltransferase (*plsC*). *Molecular and General Genetics* 232, 295–303.
- Conover, G.M., Martinez-Morales, F., Heidtman, M.I., et al., 2008. Phosphatidylcholine synthesis is required for optimal function of *Legionella pneumophila* virulence determinants. *Cellular Microbiology* 10, 514–528.
- Cornell, R.B., Northwood, I.C., 2000. Regulation of CTP:Phosphocholine cytidyltransferase by amphitropism and relocalization. *Trends in Biochemical Sciences* 25, 441–447.
- Cui, Z., Vance, J.E., Chen, M.H., Voelker, D.R., Vance, D.E., 1993. Cloning and expression of a novel phosphatidylethanolamine N-methyltransferase. A specific biochemical and cytological marker for a unique membrane fraction in rat liver. *Journal of Biological Chemistry* 268, 16655–16663.
- Cullis, P.R., de Kruijff, B., 1979. Lipid polymorphism and the functional roles of lipids in biological membranes. *Biochimica et Biophysica Acta* 559, 399–420.
- Dahlqvist, A., Andersson, S., Wieslander, A., 1992. The enzymatic synthesis of membrane glucolipids in *Acholeplasma laidlawii*. *Biochimica et Biophysica Acta* 1105, 131–140.

- Dowhan, W., 1997. Molecular basis for membrane phospholipid diversity: Why are there so many lipids? Annual Review of Biochemistry 66, 199–232.
- Dowhan, W., 2009. Molecular genetic approaches to defining lipid function. Journal of Lipid Research 50 (Suppl.), S305–S310.
- Dowhan, W., 2013. A retrospective: Use of *Escherichia coli* as a vehicle to study phospholipid synthesis and function. Biochimica et Biophysica Acta 1831, 471–494.
- Dowhan, W., Bogdanov, M., 2012. Molecular genetic and biochemical approaches for defining lipid-dependent membrane protein folding. Biochimica et Biophysica Acta 1818, 1097–1107.
- Dowhan, W., Wickner, W.T., Kennedy, E.P., 1974. Purification and properties of phosphatidylserine decarboxylase from *Escherichia coli*. Journal of Biological Chemistry 249, 3079–3084.
- Fischl, A.S., Carman, G.M., 1983. Phosphatidylinositol biosynthesis in *Saccharomyces cerevisiae*: Purification and properties of microsome-associated phosphatidylinositol synthase. Journal of Bacteriology 154, 304–311.
- Funk, C.R., Zimniak, L., Dowhan, W., 1992. The *pgpA* and *pgpB* genes of *Escherichia coli* are not essential: Evidence for a third phosphatidylglycerophosphate phosphatase. Journal of Bacteriology 174, 205–213.
- Guo, D., Tropp, B.E., 2000. A second *Escherichia coli* protein with CL synthase activity. Biochimica et Biophysica Acta 1483, 263–274.
- Halford, S., Dulai, K.S., Daw, S.C., Fitzgibbon, J., Hunt, D.M., 1998. Isolation and chromosomal localization of two human CDP-diacylglycerol synthase (CDS) genes. Genomics 54, 140–144.
- Han, G.S., Wu, W.I., Carman, G.M., 2006. The *Saccharomyces cerevisiae* Lipin homolog is a Mg²⁺-dependent phosphatidate phosphatase enzyme. Journal of Biological Chemistry 281, 9210–9218.
- Hawrot, E., Kennedy, E.P., 1976. Conditional lethal phosphatidylserine decarboxylase mutants of *Escherichia coli*. Mapping of the structural gene for phosphatidylserine decarboxylase. Molecular and General Genetics 148, 271–279.
- Henry, S.A., Kohlwein, S.D., Carman, G.M., 2012. Metabolism and regulation of glycerolipids in the yeast *Saccharomyces cerevisiae*. Genetics 190, 317–349.
- Hirabayashi, T., Larson, T.J., Dowhan, W., 1976. Membrane-associated phosphatidylglycerophosphate synthetase from *Escherichia coli*. Purification by substrate affinity chromatography on cytidine 5'-diphospho-1,2-diacyl-sn-glycerol sepharose. Biochemistry 15, 5205–5211.
- Hiraoka, S., Nukui, K., Uetake, N., Ohta, A., Shibuya, I., 1991. Amplification and substantial purification of cardiolipin synthase of *Escherichia coli*. Journal of Biochemistry 110, 443–449.
- Hirschberg, C.B., Kennedy, E.P., 1972. Mechanism of the enzymatic synthesis of cardiolipin in *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America 69, 648–651.
- Hostetler, K.Y., van den Bosch, H., van Deenen, L.L., 1972. The mechanism of cardiolipin biosynthesis in liver mitochondria. Biochimica et Biophysica Acta 260, 507–513.
- Houtkooper, R.H., Akbari, H., van Lenthe, H., et al., 2006. Identification and characterization of human cardiolipin synthase. FEBS Letters 580, 3059–3064.
- Icho, T., Raetz, C.R., 1983. Multiple genes for membrane-bound phosphatases in *Escherichia coli* and their action on phospholipid precursors. Journal of Bacteriology 153, 722–730.
- Jiang, F., Rizavi, H.S., Greenberg, M.L., 1997. Cardiolipin is not essential for the growth of *Saccharomyces cerevisiae* on fermentable or non-fermentable carbon sources. Molecular Microbiology 26, 481–491.
- Kanfer, J.N., Kennedy, E.P., 1962. Synthesis of phosphatidylserine by *Escherichia coli*. Journal of Biological Chemistry 237, PC270–PC271.
- Kawasaki, K., Kuge, O., Chang, S.C., et al., 1999. Isolation of a chinese hamster ovary (CHO) cDNA encoding phosphatidylglycerophosphate (PGP) synthase, expression of which corrects the mitochondrial abnormalities of a PGP synthase-defective mutant of CHO-K1 cells. Journal of Biological Chemistry 274, 1828–1834.
- Kelley, M.J., Carman, G.M., 1987. Purification and characterization of CDP-diacylglycerol synthase from *Saccharomyces cerevisiae*. Journal of Biological Chemistry 262, 14563–14570.
- Kennedy, E.P., 1956. The synthesis of cytidine diphosphate choline, cytidine diphosphate ethanolamine, and related compounds. Journal of Biological Chemistry 222, 185–191.
- Kennedy, E.P., Borkenhagen, L.F., Smith, S.W., 1959. Possible metabolic functions of deoxycytidine diphosphate choline and deoxycytidine diphosphate ethanolamine. Journal of Biological Chemistry 234, 1998–2000.
- Kennedy, E.P., Smith, S.W., Weiss, S.B., 1956. New synthesis of lecithin in an isolated enzyme system. Nature 178, 594–595.
- Kennedy, E.P., Weiss, S.B., 1956. The function of cytidine coenzymes in the biosynthesis of phospholipides. Journal of Biological Chemistry 222, 193–214.
- Kitos, T.E., Drobnie, A., Ng, M.N., Wen, Y., Cornell, R.B., 2006. Contribution of lipid mediators to the regulation of phosphatidylcholine synthesis by angiotensin. Biochimica et Biophysica Acta 1761, 261–271.
- Kiyasu, J.Y., Pieringer, R.A., Paulus, H., Kennedy, E.P., 1963. The biosynthesis of phosphatidylglycerol. Journal of Biological Chemistry 238, 2293–2298.
- Larson, T.J., Dowhan, W., 1976. Ribosomal-associated phosphatidylserine synthetase from *Escherichia coli*: Purification by substrate-specific elution from phosphocellulose using cytidine 5'-diphospho-1,2-diacyl-sn-glycerol. Biochemistry 15, 5212–5218.
- Larson, T.J., Lightner, V.A., Green, P.R., Modrich, P., Bell, R.M., 1980. Membrane phospholipid synthesis in *Escherichia coli*. Identification of the sn-glycerol-3-phosphate acyltransferase polypeptide as the *plsB* gene product. Journal of Biological Chemistry 255, 9421–9426.
- Li, Q.X., Dowhan, W., 1988. Structural characterization of *Escherichia coli* phosphatidylserine decarboxylase. Journal of Biological Chemistry 263, 11516–11522.
- Li, Q.X., Dowhan, W., 1990. Studies on the mechanism of formation of the pyruvate prosthetic group of phosphatidylserine decarboxylase from *Escherichia coli*. Journal of Biological Chemistry 265, 4111–4115.
- Louie, K., Dowhan, W., 1980. Investigations on the association of phosphatidylserine synthase with the ribosomal component from *Escherichia coli*. Journal of Biological Chemistry 255, 1124–1127.
- Lu, B., Xu, F.Y., Jiang, Y.J., et al., 2006. Cloning and characterization of a cDNA encoding human cardiolipin synthase (hCLS1). Journal of Lipid Research 47, 1140–1145.
- Lu, Y.H., Guan, Z., Zhao, J., Raetz, C.R., 2011. Three phosphatidylglycerol-phosphate phosphatases in the inner membrane of *Escherichia coli*. Journal of Biological Chemistry 286, 5506–5518.
- Mason, R.J., Voelker, D.R., 1988. Role of intracellular glycerol-3-phosphate in the synthesis of phosphatidylglycerol by freshly isolated adult rat alveolar type II cells. American Review of Respiratory Disease 137, 519–524.
- Matsumoto, K., 1997. Phosphatidylserine synthase from bacteria. Biochimica et Biophysica Acta 1348, 214–227.
- McIntyre, T.M., Snyder, F., G.K., Marathe, 2008. Ether-linked lipids and their bioactive species. In: Vance, D.E., Vance, J.E. (Eds.), Biochemistry of Lipids, Lipoproteins and Membranes, fifth ed. Amsterdam: Elsevier, pp. 245–276.
- Mileykovskaya, E., Dowhan, W., 2005. Role of membrane lipids in bacterial division-site selection. Current Opinion in Microbiology 8, 135–142.
- Mileykovskaya, E., Dowhan, W., 2014. Cardiolipin-dependent formation of mitochondrial respiratory supercomplexes. Chemistry and Physics of Lipids 179, 42–48.
- Mileykovskaya, E., Penczek, P.A., Fang, J., et al., 2012. Arrangement of the respiratory chain complexes in *Saccharomyces cerevisiae* supercomplex III₂IV₂ revealed by single particle cryo-electron microscopy. Journal of Biological Chemistry 287, 23095–23103.
- Numata, M., Nagashima, Y., Moore, M.L., et al., 2013. Phosphatidylglycerol provides short-term prophylaxis against respiratory syncytial virus infection. Journal of Lipid Research 54, 2133–2143.
- Ohta, A., Obara, T., Asami, Y., Shibuya, I., 1985. Molecular cloning of the *cls* gene responsible for cardiolipin synthesis in *Escherichia coli* and phenotypic consequences of its amplification. Journal of Bacteriology 163, 506–514.
- Ohta, A., Waggoner, K., Radomska-Pyrek, A., Dowhan, W., 1981. Cloning of genes involved in membrane lipid synthesis: Effects of amplification of phosphatidylglycerophosphate synthase in *Escherichia coli*. Journal of Bacteriology 147, 552–562.
- Ohtsuka, T., Nishijima, M., Akamatsu, Y., 1993a. A somatic cell mutant defective in phosphatidylglycerophosphate synthase, with impaired phosphatidylglycerol and cardiolipin biosynthesis. Journal of Biological Chemistry 268, 22908–22913.
- Ohtsuka, T., Nishijima, M., Suzuki, K., Akamatsu, Y., 1993b. Mitochondrial dysfunction of a cultured Chinese hamster ovary cell mutant deficient in cardiolipin. Journal of Biological Chemistry 268, 22914–22919.
- Oku, Y., Kurokawa, K., Ichihashi, N., Sekimizu, K., 2004. Characterization of the *Staphylococcus aureus* *mprF* gene, involved in lysinylation of phosphatidylglycerol. Microbiology 150, 45–51.
- Osman, C., Haag, M., Wieland, F.T., Brugger, B., Langer, T., 2010. A mitochondrial phosphatase required for cardiolipin biosynthesis: The PGP phosphatase Gep4. EMBO Journal 29, 1976–1987.
- Osman, C., Voelker, D.R., Langer, T., 2011. Making heads or tails of phospholipids in mitochondria. Journal of Cell Biology 192, 7–16.
- Ostrander, D.B., Sparagna, G.C., Amoscato, A.A., McMillin, J.B., Dowhan, W., 2001a. Decreased cardiolipin synthesis corresponds with cytochrome c release in

- palmitate-induced cardiomyocyte apoptosis. *Journal of Biological Chemistry* 276, 38061–38067.
- Ostrand, D.B., Zhang, M., Mileykovskaya, E., Rho, M., Dowhan, W., 2001b. Lack of mitochondrial anionic phospholipids causes an inhibition of translation of protein components of the electron transport chain. A yeast genetic model system for the study of anionic phospholipid function in mitochondria. *Journal of Biological Chemistry* 276, 25262–25272.
- Paulus, H., Kennedy, E.P., 1960. The enzymatic synthesis of inositol monophosphatide. *Journal of Biological Chemistry* 235, 1303–1311.
- Phan, J., Reue, K., 2005. Lipin, a lipodystrophy and obesity gene. *Cell Metabolism* 1, 73–83.
- Pluschke, G., Hirota, Y., Overath, P., 1978. Function of phospholipids in *Escherichia coli*. Characterization of a mutant deficient in cardiolipin synthesis. *Journal of Biological Chemistry* 253, 5048–5055.
- Preitschopf, W., Luckl, H., Summers, E., *et al.*, 1993. Molecular cloning of the yeast OPI3 gene as a high copy number suppressor of the *cho2* mutation. *Current Genetics* 23, 95–101.
- Raetz, C.R., 1975. Isolation of *Escherichia coli* mutants defective in enzymes of membrane lipid synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 72, 2274–2278.
- Raetz, C.R., 1976. Phosphatidylserine synthetase mutants of *Escherichia coli*. Genetic mapping and membrane phospholipid composition. *Journal of Biological Chemistry* 251, 3242–3249.
- Ridgway, N.D., Vance, D.E., 1987. Purification of phosphatidylethanolamine N-methyltransferase from rat liver. *Journal of Biological Chemistry* 262, 17231–17239.
- Schlame, M., 2008. Cardiolipin synthesis for the assembly of bacterial and mitochondrial membranes. *Journal of Lipid Research* 49, 1607–1620.
- Schlame, M., 2013. Cardiolipin remodeling and the function of tafazzin. *Biochimica et Biophysica Acta* 1831, 582–588.
- Schlame, M., Greenberg, M.L., 1997. Cardiolipin synthase from yeast. *Biochimica et Biophysica Acta* 1348, 201–206.
- Schlame, M., Kelley, R.I., Feigenbaum, A., *et al.*, 2003. Phospholipid abnormalities in children with Barth syndrome. *Journal of the American College of Cardiology* 42, 1994–1999.
- Shen, H., Heacock, P.N., Clancey, C.J., Dowhan, W., 1996. The *CDS1* gene encoding CDP-diacylglycerol synthase in *Saccharomyces cerevisiae* is essential for cell growth. *Journal of Biological Chemistry* 271, 789–795.
- Sjölund, M., Rilfors, L., Lindblom, G., 1989. Reversed hexagonal phase formation in lecithin-alkane-water systems with different acyl chain unsaturation and alkane length. *Biochemistry* 28, 1323–1329.
- Smith, S.W., Weiss, S.B., Kennedy, E.P., 1957. The enzymatic dephosphorylation of phosphatidic acids. *Journal of Biological Chemistry* 228, 915–922.
- Sohlenkamp, C., Lopez-Lara, I.M., Geiger, O., 2003. Biosynthesis of phosphatidylcholine in bacteria. *Progress in Lipid Research* 42, 115–162.
- Sparrow, C.P., Raetz, C.R., 1985. Purification and properties of the membrane-bound CDP-diglyceride synthetase from *Escherichia coli*. *Journal of Biological Chemistry* 260, 12084–12091.
- Summers, E.F., Letts, V.A., McGraw, P., Henry, S.A., 1988. *Saccharomyces cerevisiae* *cho2* mutants are deficient in phospholipid methylation and cross-pathway regulation of inositol synthesis. *Genetics* 120, 909–922.
- Tamura, Y., Harada, Y., Nishikawa, S., *et al.*, 2013. Tam41 is a CDP-diacylglycerol synthase required for cardiolipin biosynthesis in mitochondria. *Cell Metabolism* 17, 709–718.
- Tan, B.K., Bogdanov, M., Zhao, J., *et al.*, 2012. Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. *Proceedings of the National Academy of Sciences of the United States of America* 109, 16402–16403.
- Thurmond, R.L., Lindblom, G., Brown, M.F., 1993. Curvature, order, and dynamics of lipid hexagonal phases studied by deuterium NMR spectroscopy. *Biochemistry* 32, 5394–5410.
- Trotter, P.J., Pedretti, J., Voelker, D.R., 1993. Phosphatidylserine decarboxylase from *Saccharomyces cerevisiae*. Isolation of mutants, cloning of the gene, and creation of a null allele. *Journal of Biological Chemistry* 268, 21416–21424.
- Tsai, Y.H., Liu, X., Seeberger, P.H., 2012. Chemical biology of glycosylphosphatidylinositol anchors. *Angewandte Chemie International Edition in English* 51, 11438–11456.
- Tuller, G., Hrasnik, C., Achleitner, G., *et al.*, 1998. YDL142c encodes cardiolipin synthase (Cis1p) and is non-essential for aerobic growth of *Saccharomyces cerevisiae*. *FEBS Letters* 421, 15–18.
- Vance, D.E., Ridgway, N.D., 1988. The methylation of phosphatidylethanolamine. *Progress in Lipid Research* 27, 61–79.
- Vance, D.E., Vance, J.E., 2008. Phospholipid biosynthesis in eukaryotes. In: Vance, D.E., Vance, J.E. (Eds.), *Biochemistry of Lipids, Lipoproteins and Membranes*, fifth ed. Amsterdam: Elsevier, pp. 213–244.
- Vance, J.E., Steenbergen, R., 2005. Metabolism and functions of phosphatidylserine. *Progress in Lipid Research* 44, 207–234.
- Voelker, D.R., 1997. Phosphatidylserine decarboxylase. *Biochimica et Biophysica Acta* 1348, 236–244.
- Walkey, C.J., Yu, L., Agellon, L.B., Vance, D.E., 1998. Biochemical and evolutionary significance of phospholipid methylation. *Journal of Biological Chemistry* 273, 27043–27046.
- Weeks, R., Dowhan, W., Shen, H., *et al.*, 1997. Isolation and expression of an isoform of human CDP-diacylglycerol synthase cDNA. *DNA and Cell Biology* 16, 281–289.
- Weiss, S.B., Smith, S.W., Kennedy, E.P., 1958. The enzymatic formation of lecithin from cytidine diphosphate choline and *o*-1,2-diglyceride. *Journal of Biological Chemistry* 231, 53–64.
- Wikstrom, M., Kelly, A.A., Georgiev, A., *et al.*, 2009. Lipid-engineered *Escherichia coli* membranes reveal critical lipid headgroup size for protein function. *Journal of Biological Chemistry* 284, 954–965.
- Xia, W., Dowhan, W., 1995. Phosphatidylinositol cannot substitute for phosphatidylglycerol in supporting cell growth of *Escherichia coli*. *Journal of Bacteriology* 177, 2926–2928.
- Xie, J., Bogdanov, M., Heacock, P., Dowhan, W., 2006. Phosphatidylethanolamine and monoglucosyldiacylglycerol are interchangeable in supporting topogenesis and function of the polytopic membrane protein lactose permease. *Journal of Biological Chemistry* 281, 19172–19178.
- Yao, J., Rock, C.O., 2013. Phosphatidic acid synthesis in bacteria. *Biochimica et Biophysica Acta* 1831, 495–502.
- York, J.D., Guo, S., Odom, A.R., Spiegelberg, B.D., Stolz, L.E., 2001. An expanded view of inositol signaling. *Advances in Enzyme Regulation* 41, 57–71.
- Zhang, J., Guan, Z., Murphy, A.N., *et al.*, 2011. Mitochondrial phosphatase PTPMT1 is essential for cardiolipin biosynthesis. *Cell Metabolism* 13, 690–700.
- Zhang, M., Mileykovskaya, E., Dowhan, W., 2002. Gluing the respiratory chain together. Cardiolipin is required for supercomplex formation in the inner mitochondrial membrane. *Journal of Biological Chemistry* 277, 43553–43556.
- Zhang, M., Mileykovskaya, E., Dowhan, W., 2005. Cardiolipin is essential for organization of complexes III and IV into a supercomplex in intact yeast mitochondria. *Journal of Biological Chemistry* 280, 29403–29408.

Further Reading

- Dowhan, W., 2008. Functional roles of lipids in membranes. In: Vance, D.E., Vance, J.E. (Eds.), *Biochemistry of Lipids, Lipoproteins and Membranes*, fifth ed. Amsterdam: Elsevier, pp. 1–38.
- Dowhan, W., 2013. A retrospective: Use of *Escherichia coli* as a vehicle to study phospholipid synthesis and function. *Biochimica et Biophysica Acta* 1831, 471–494.
- Schmid, K.M., Ohlrogge, J.B., 2008. Lipid synthesis in plants. In: Vance, D.E., Vance, J.E. (Eds.), *Biochemistry of Lipids, Lipoproteins and Membranes*, fifth ed. Amsterdam: Elsevier, pp. 98–1320.
- Sprott, G.D., 1992. Structures of archaeobacterial membrane lipids. *Journal of Bioenergetics and Biomembranes* 24, 555–566.

Relevant Websites

- <http://www.lipidbank.jp>
Lipid Bank in Japan.
- <http://www.lipidmaps.org>
LIPID MAPS consortium in the United States.
- <http://www.lipidomicnet.org>
LipidomicNet in Europe.

Cholesterol and Other Steroids

S Sitaula, The Scripps Research Institute, Jupiter, FL, USA

TP Burris, Saint Louis University School of Medicine, St. Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Cholesterol is a major component of cell membranes and a precursor to many important biological molecules. Since the discovery of cholesterol from gallstones in the eighteenth century, it has been a molecule of great interest and has been studied extensively. As a result, 13 Nobel Prizes have been awarded to researchers who contributed to the study of cholesterol, its chemistry, biochemistry, physiological role, and its clinical significance. Cholesterol metabolism has complex regulatory mechanisms and new functions for this molecule in cell biology and disease processes are being uncovered even hundreds of years after its initial discovery.

There is considerable evidence demonstrating that cholesterol is indispensable for normal functioning of a cell. Cholesterol is important for many biological processes and its accumulation or depletion can be detrimental to the body as evidenced by multiple disorders identified in animal models and in humans with defective cholesterol homeostasis. Due to the fine balance that exists between its essential role in normal biological function and its role in disease, complex regulatory mechanisms that strictly monitor and regulate the synthesis, uptake, transport, and elimination of cholesterol have evolved. This regulation is achieved by feedback regulatory mechanisms involving a family of transcriptional factors, cholesterol synthesis enzymes, sterol-sensing proteins, and lipoprotein receptors.

Cholesterol has important functions in membrane dynamics, bile acid synthesis, signal transduction, myelin sheath composition, and steroid hormone formation (Ikonen, 2008; Dowhan *et al.*, 2008). In humans, cholesterol is obtained from both the diet and endogenous synthesis. Cholesterol is synthesized endogenously by many tissues but the liver is the predominant site of cholesterol synthesis. Cholesterol synthesis progresses through a complex pathway that contains many enzymes and is tightly regulated in many steps. The rate of endogenous synthesis of cholesterol depends on a number of factors including the amount of cholesterol uptake from diet. A feedback mechanism exists that allows the liver to sense plasma cholesterol concentrations and regulate cholesterol biosynthesis in cells. Defects in lipid and cholesterol metabolism have been associated with conditions such as cholelithiasis, Smith-Lemli-Opitz syndrome (SLOS), Tangier disease, cancer, neurological conditions, and, most commonly, several cardiovascular disorders including atherosclerosis (Maxfield and Tabas, 2005; Orth and Bellosta, 2012). A detailed understanding of important players in the cholesterol synthesis pathway and cholesterol homeostasis processes is fundamental to understanding and treating these cholesterol-associated disorders. The structure, function, cellular cholesterol homeostasis, and the role of cholesterol in several diseases are discussed in detail below.

Structure

The structure of cholesterol was elucidated in the twentieth century by the work of many scientists including Nobel Laureates Otto Diels, Robert Robinson, Heinrich Wieland, Adolf Windaus, and Leopold Ruzicka (Brock, 2000). Cholesterol belongs to the family of polycyclic compounds known as sterols. This 27-carbon structure has rigid planar rings and a flexible tail (Figure 1). Cholesterol is comprised of four hydrocarbon rings, a hydroxyl group attached to C3 and an eight-carbon tail. The four planar rings contribute to the hydrophobicity of the molecule while the hydroxyl group is hydrophilic, making cholesterol amphipathic (Bloch, 1983). Although cholesterol is slightly polar, it is highly insoluble in water. In the cell membrane cholesterol orients with the hydrophobic rings facing the fatty acid chain of the phospholipid and the polar hydroxyl group facing the aqueous layer. While the hydroxyl group confers polarity to the molecule allowing it to form a part of the cell membrane, it is also important as a site for esterification of cholesterol. These chemical properties allow cholesterol to form an integral part of the cell membrane and perform its functions effectively, but interestingly, the same properties make it toxic to cells in free form.

Cholesteryl Esters

Most of the cholesterol in humans exists in ester form. Cholesterol converted into cholesteryl ester can be stored and transported more efficiently. Since cholesteryl esters are more hydrophobic than cholesterol, they can be packaged inside the hydrophobic core of lipoproteins and transported throughout the body (Ginsburg *et al.*, 1984). Esterification of cholesterol is required for both delivery of cholesterol to the tissues and for removal of cholesterol from peripheral tissues. Cholesterol esterification is one of the several mechanisms cells use to prevent accumulation of free cholesterol that is toxic to cells (Tabas, 2002). Free cholesterol is esterified in liver and plasma by specialized mechanisms. The enzymes lecithin:cholesterol acyltransferase (LCAT) and acyl-coenzyme A:cholesterol acyltransferase (ACAT) catalyze the esterification of cholesterol (Peelman *et al.*, 2000; Chang *et al.*, 1997). Another enzyme central to cholesterol homeostasis is cholesteryl ester transfer protein (CETP). CETP plays an important role in transferring cholesteryl esters and triglycerides between high-density lipoprotein (HDL) and low-density lipoprotein (LDL) particles (Barter, 2003). Lysosomal acid lipase is an enzyme that is required for breaking down cholesteryl esters into free cholesterol. Deficiency of this enzyme is associated with disorders such as Wolman disease and cholesteryl ester storage disease characterized by abnormal accumulation of cholesteryl ester in liver eventually leading to chronic liver disease (Fouchier and

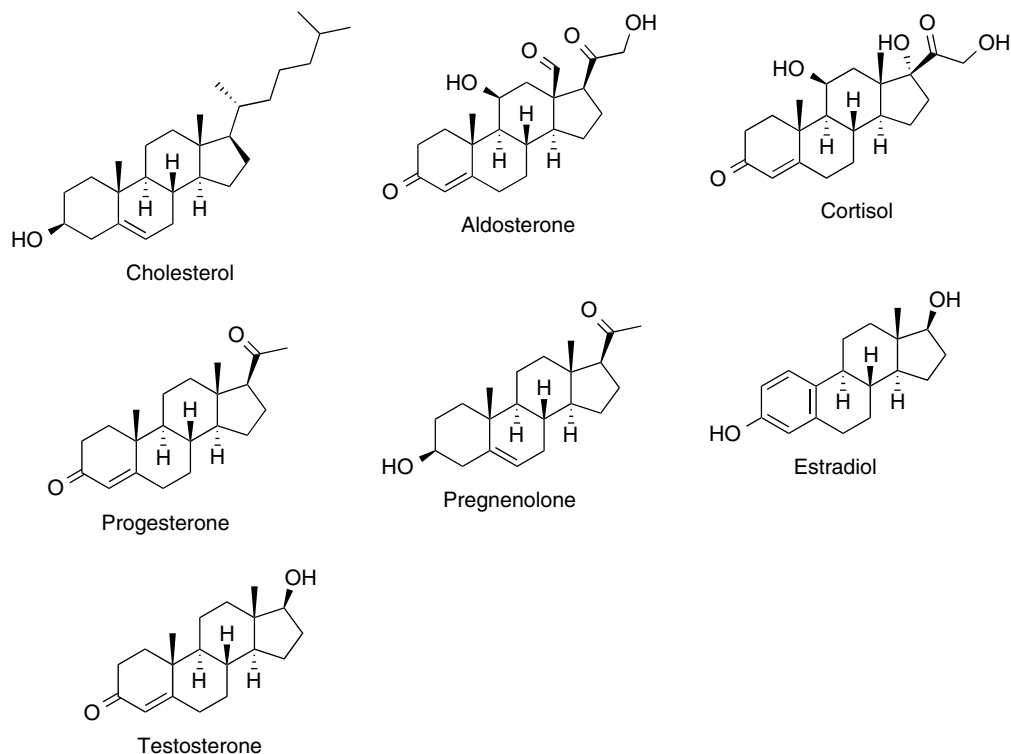


Figure 1 Structure of cholesterol and other steroids.

Defesche, 2013). Proper formation, transport, and hydrolysis of cholesteryl esters are essential in overall cholesterol homeostasis in the body.

Cholesterol in Cell Membranes

Cholesterol levels within the cell membrane influence the biophysical properties of the membrane and play a critical role in cell membrane ordering, permeability, phase transitions, and in maintaining cell membrane integrity. The plasma membrane is made up of phospholipid bilayer, proteins, cholesterol, glycoproteins, and glycolipids (Figure 2). Cholesterol is not arranged uniformly throughout the plasma membrane, but is found in varying concentrations. The hydroxyl group of cholesterol in phospholipid bilayer is oriented toward the hydrophilic head of the phospholipids in the membrane. At high temperatures, cholesterol reduces phospholipid movement and prevents the membrane from becoming too fluid or permeable. However, at low temperature, cholesterol prevents crystallization of membranes that can occur due to close packing of fatty acid tails. Cholesterol has also been shown to be associated with sphingolipids to form lipid rafts (George and Wu, 2012; Brown and London, 2000). These rafts have an important role in protein interaction during cell signaling (Lemaire-Ewing *et al.*, 2012; Simons and Toomre, 2000; George and Wu, 2012). Moreover, cholesterol has been recognized as an essential modulator of sonic hedgehog signaling. It is also required for the formation of clathrin-coated pits in cell membranes during endocytosis. There is direct evidence that depletion of cholesterol in the

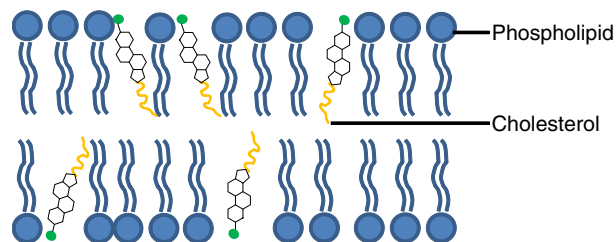


Figure 2 Cholesterol in plasma membrane.

membrane inhibits the internalization of clathrin-coated pits (Subtil *et al.*, 1999). Thus, cholesterol not only alters the properties of the cell membrane but plays an active role in the functional aspect of the cell membrane.

Cholesterol Biosynthesis

Cholesterol synthesis, also called cholesterologenesis, is a multistep enzymatic biosynthetic process that begins with acetyl-coenzyme A. A simplified schematic of the pathway that displaces the most important steps is shown in Figure 3. Cholesterol synthesis takes place in the cytoplasm and in the endoplasmic reticulum (ER). The first step in the pathway catalyzed by 3-hydroxy-3-methylglutaryl (HMG)-CoA synthase (HMGCS) occurs in the cytosol while the subsequent steps occur in the ER. Thus, the ER is the main site of cholesterol synthesis (Simons and Ikonen, 2000). Cholesterol synthesis begins with acetyl-coenzyme A derived from mitochondria and transported to the cytosol. One molecule of acetyl-coenzyme A and one molecule of acetoacetyl-CoA are

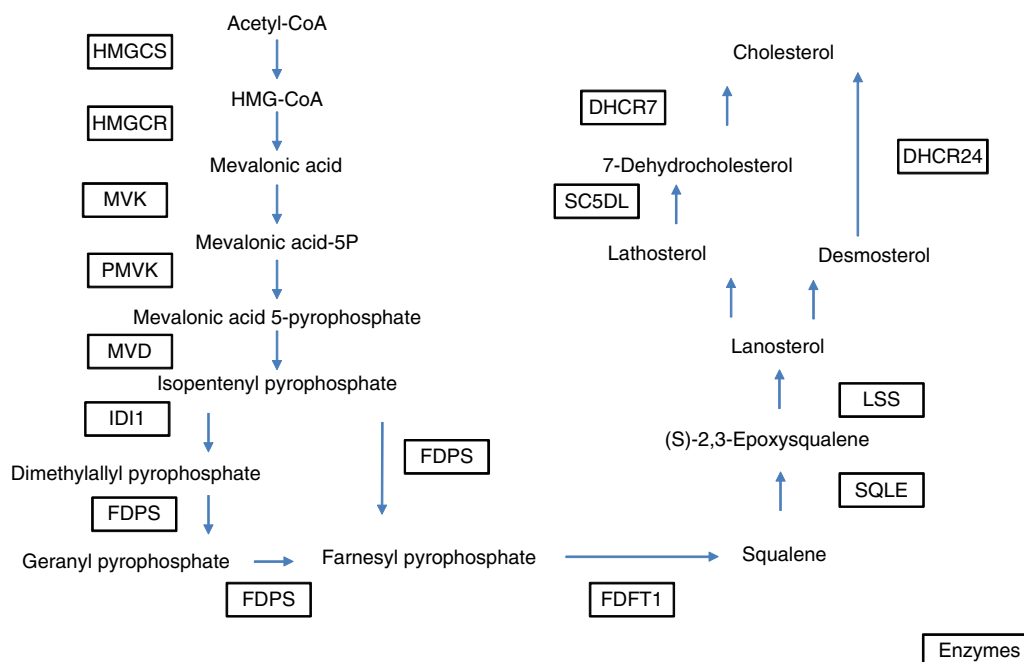


Figure 3 Cholesterol biosynthesis pathway.

converted to HMG-CoA. HMG-CoA is then reduced to mevalonate by HMG-CoA reductase (HMGCR). Mevalonate is further phosphorylated to isopentenyl pyrophosphate, which is converted to geranyl pyrophosphate. Condensation with another isopentenyl pyrophosphate yields farnesyl pyrophosphate. Squalene synthase catalyzes the condensation of two molecules of farnesyl pyrophosphate to yield squalene. Squalene is then cyclized to form lanosterol. Finally lanosterol is converted to cholesterol in 19 more reactions.

HMGCR is the rate-limiting enzyme of the pathway. It is tightly regulated at transcriptional and posttranscriptional levels. Mechanisms such as phosphorylation–dephosphorylation, feedback from sterol and non-sterol metabolites of the pathway, and ubiquitination control the activity and levels of HMGCR. This enzyme is a pharmacological target of a class of drugs called statins, which are commonly used to lower cholesterol by reducing cholesterol biosynthesis. Statins are competitive inhibitors of HMGCR enzyme.

The cholesterol biosynthesis pathway yields several molecules as intermediates that are essential for other biological pathways. Farnesyl and geranylgeranyl membrane anchors are important for signaling proteins that regulate progression through cell cycle (Edwards and Ericsson, 1999). Ubiquinone and dolichols are other intermediates that have roles in electron transport and synthesis of glycoproteins, respectively. Cholesterol can be further modified into steroid hormones as well as vitamin D.

Regulation of Cholesterol Biosynthesis

The complex regulation of cholesterol biosynthesis takes place at several levels. The rate of synthesis is highly responsive to cellular level of cholesterol. There exists a feedback regulation

mediated by changes in HMGCR. High cholesterol levels cause reduction in transcription of *HMGCR* gene, leading to the reduction in the amount of HMGCR (messenger RNA) mRNA. HMGCR activity is also regulated through phosphorylation. AMP (adenosine monophosphate)-dependent protein kinase (AMPK) catalyzes the phosphorylation and inhibition of the enzyme (Motoshima *et al.*, 2006). Various cholesterol metabolites such as oxidized derivatives of cholesterol as well as cholesterol biosynthesis intermediates such as mevalonate and farnesol are additional components of the negative feedback loop that regulates the levels of HMGCR (Correll *et al.*, 1994). Cholesterol can also induce ubiquitination and degradation of HMGCR. Increases in cellular cholesterol concentrations triggers binding of HMGCR to the ER membrane proteins Insig-1 and Insig-2. This binding leads to the recruitment of gp78, the ubiquitin ligase that leads to ubiquitination of the enzyme. In this manner, HMGCR is marked for degradation and transported to proteasomes for degradation (DeBose-Boyd, 2008).

A complex of proteins that sense cellular cholesterol levels also regulates the rate of cholesterol synthesis. Sterol response element binding protein (SREBP), the SREBP cleavage activating protein (SCAP), and two SREBP-specific proteases (S1P and S2P) are the key components of this pathway (Horton *et al.*, 2002; Bengoechea-Alonso and Ericsson, 2007). SREBP precursor proteins are embedded in the membrane of ER. The N-terminal domain of SREBP acts as a transcription factor, whereas the C-terminal domain interacts with C-terminal domain of another ER protein SCAP, which contains a sterol-sensing domain. As shown in Figure 4, when sterol levels are high, SCAP interacts with ER membrane protein, insulin regulated protein (Insig) (Yang *et al.*, 2002). Association with Insig leads to the retention of SREBP–SCAP complex in the ER (Yang *et al.*, 2002). When sterol levels are low, SCAP and Insig do not interact, which leads to a conformational change in

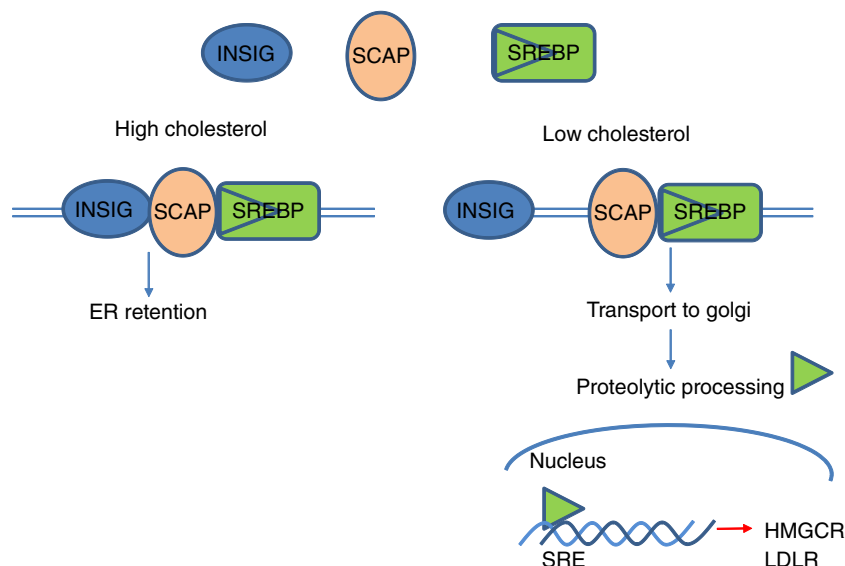


Figure 4 SREBP and SCAP in transcriptional control of cholesterol synthesis.

SCAP (Brown *et al.*, 2002). The SREBP–SCAP complex then translocates from ER to Golgi apparatus (Figure 4). Two proteases, S1P and S2P, act on SREBP in the Golgi apparatus (Nohturfft *et al.*, 2000). SREBP is first cleaved by protease S1P, yielding a product that then acts as a substrate for protease S2P. This cleaved SREBP is then released to the cytosol and travels to the nucleus where it binds to sterol response element (SRE) DNA sites resulting in increased transcription of several genes involved in cholesterol synthesis pathway (Bengoechea-Alonso and Ericsson, 2007).

Cellular Cholesterol Homeostasis

In addition to intracellular synthesis, cholesterol can also enter the cells through uptake of cholesterol-containing lipoproteins from plasma. A simplified schematic of cholesterol homeostasis is presented in Figure 5. Cholesterol from diet is absorbed in the intestine, solubilized in micelles, and transported to the periphery packaged in the form of chylomicrons along with triglycerides (Wang, 2007). The enzyme lipoprotein lipase interacts with chylomicrons and hydrolyzes triglycerides to release fatty acids. The resulting chylomicron remnants enter hepatocytes by binding to apolipoprotein E receptors. Cholesterol is subjected to endocytosis and is either converted to bile acids, secreted into circulation, esterified, or packaged into very low-density lipoprotein (VLDL) particles. These VLDL particles containing cholesterol and triglycerides are secreted into circulation. Interaction of VLDL particles with lipases results in change in the composition of these particles. These particles contain less triglycerides and all apolipoproteins are lost except apolipoprotein B (ApoB). These modified particles are called LDLs. LDL particles are responsible for transporting cholesterol to the periphery.

LDL particles are mostly composed of cholesterol esters. Cells that require cholesterol express receptors that recognize and bind LDL particles resulting in receptor-mediated

endocytosis. LDL particles and their receptors have an important role in cholesterol homeostasis (Go and Mani, 2012). Genetic mutation in the LDL receptor is associated with a condition called familial hypercholesterolemia, which is associated with increased atherosclerosis and heart disease. LDL-derived cholesterol is shown to inhibit *HMGCR* gene transcription by inhibiting SREBP pathway (Brown and Goldstein, 1980). Increased LDL cholesterol leads to the reduction of synthesis of LDLR through negative feedback mechanism. When the cholesterol level in cells is low, more LDLRs are produced that help take up LDL from plasma. Along with inhibiting *HMGCR*, statins upregulate LDLR resulting in clearing of LDL from the circulation. These LDLRs undergo endocytosis taking LDL inside the cell through endosome and are later released from endosome and are either degraded or released back to the surface (Ikonen, 2008). Due to the cellular toxicity of free cholesterol, excess free cholesterol must be esterified (by the enzyme ACAT) and packaged in lipid droplets (Chang *et al.*, 1997).

LDL particles undergo a range of modifications. Oxidized LDL particles are specifically targeted for uptake by scavenger receptors found in macrophages. These help reduce the amount of plasma LDL cholesterol and return them to the liver in the form of HDL, another lipoprotein particle, via a process called reverse cholesterol transport. HDL particles play a critical role in reverse cholesterol transport where cholesterol is returned to the liver from the periphery. Nascent HDL particles interact with cholesterol transporter proteins on the cell surface, ATP-binding cassette, sub-family A, member 1 (ABCA1), and ATP-binding cassette, sub-family G, member 1 (ABCG1), and accept cholesterol. The enzyme LCAT in HDL converts free cholesterol to cholesteryl esters and stores them in the HDL core. A defect in *ABCA1* gene is associated with Tangier disease, which is associated with very low levels of HDL and high risk of atherosclerosis. Cholesterol transported back to the liver is either recycled, stored, or converted to bile acids.

Nuclear receptors such as liver X-receptors (LXR), farnesoid X-receptor (FXR), and peroxisome proliferator-activated

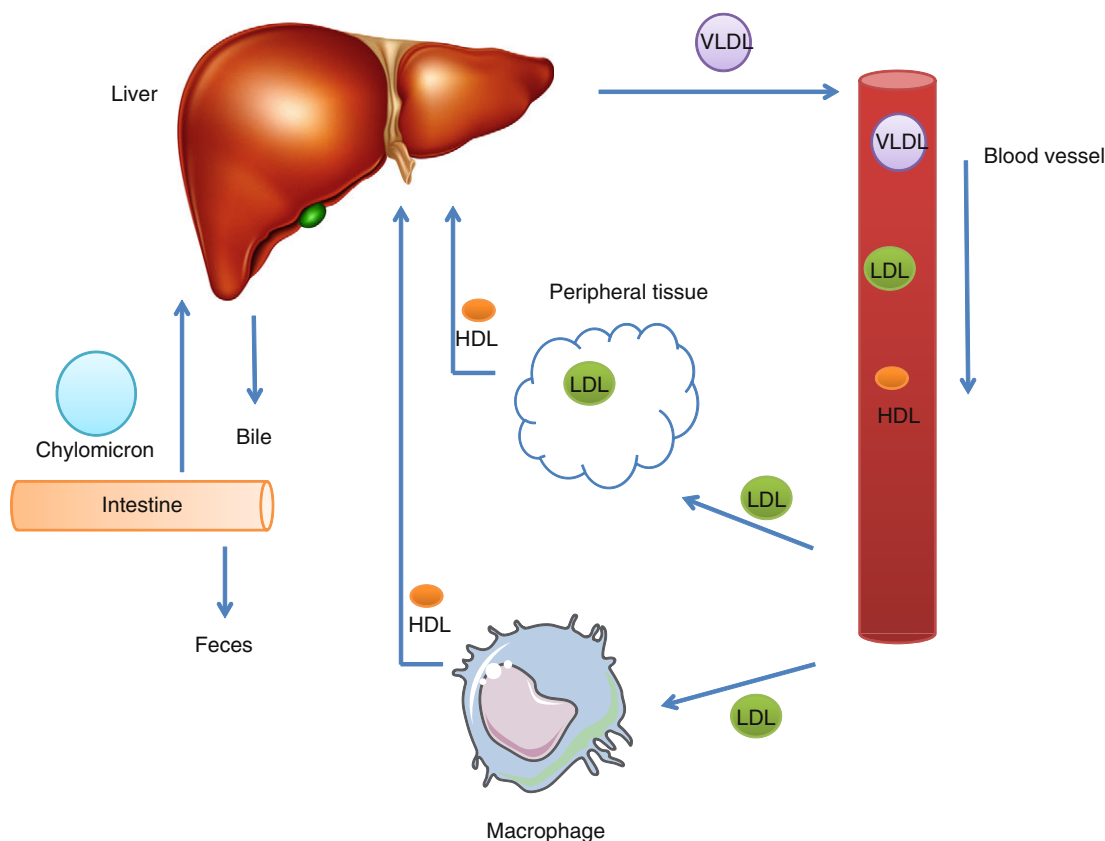


Figure 5 Overview of cholesterol homeostasis.

receptor γ (PPAR γ) are crucial in cholesterol homeostasis. LXRs are activated by cholesterol derivatives, oxysterols. LXRs are transcription factors and regulate the expression of important genes such as *ABCA1*, *ABCG1*, and *ApoE* important in the process of reverse cholesterol transport (Laffitte *et al.*, 2001; Beyea *et al.*, 2007). FXRs play a role in bile acid metabolism and in regulation of cholesterol removal from body. The ability of these nuclear receptors to affect cholesterol metabolism makes them an attractive therapeutic target for treatment of metabolic diseases associated with dyslipidemia.

Cholesterol in Brain

The brain contains approximately one-fourth of the total amount of cholesterol in the body, making it the most cholesterol-rich organ. Cholesterol is a major component of the myelin sheath and is crucial for proper brain development (Dietschy and Turley, 2001). A distinct pool of cholesterol is maintained in the brain by local *de novo* synthesis. The blood-brain barrier prevents the exchange of lipoproteins in the circulation, thus the brain must maintain its own cholesterol transport system. In adult brain, astrocytes synthesize most of the cholesterol and send them to neurons in ApoE-containing lipoproteins particles. Cholesterol released from astrocytes through ABCA1 and ABCG1 transporters are shuttled to neurons and taken up by receptor-mediated endocytosis via LDLR family receptors.

Unlike other peripheral organs, cholesterol is recycled at a very high rate in the brain. As a result, the half-life of cholesterol in adult brain is about five years (Björkhem *et al.*, 1998). While most of the cholesterol is recycled, the brain also excretes some cholesterol in the form of 24(S)-hydroxycholesterol. Cholesterol is converted to 24(S)-hydroxycholesterol by the CYP46 (cholesterol 24-hydroxylase) enzyme. As evident from numerous studies, cholesterol metabolism in the brain is independent of that in rest of the body (Björkhem *et al.*, 1998; Dietschy and Turley, 2001). The importance of separate cholesterol regulatory mechanisms in brain is still not completely understood. However, tight regulation of cholesterol synthesis in brain appears to be critical and defects in cholesterol metabolism have been implicated in many neurological disorders.

Cholesterol in Diseases

Many pathological conditions are associated with abnormal cholesterol levels or impaired cholesterol metabolism in the body. Some of these are briefly discussed below.

Cardiovascular diseases: Elevated cholesterol levels are associated with cardiovascular conditions such as atherosclerotic plaque formation, stroke, and myocardial infarction. Formation of atherosclerotic plaque is an intricate process involving inflammation and lipid deposition that eventually narrows the arteries which results in luminal obstruction.

Hypercholesterolemia is a well-established risk factor for the incidence of atherosclerosis and its pathologic complications. Modified LDL is taken up by mature macrophages in the artery. These macrophages accumulate in the subendothelial space and become cholesterol-laden foam cells that eventually harden and become plaques. Statins (inhibitors of HMGCR) that effectively lower plasma cholesterol are the most commonly prescribed drug.

Cholesterol gallstone disease (Cholelithiasis): Gallstones are masses of cholesterol crystals, calcium bilirubinate, and proteins. Gallstones are formed as a result of excess cholesterol in bile or due to insufficient bile acids. This condition is highly prevalent in developed countries. Dyslipidemia is one of the major risk factors of cholesterol gallstones.

Familial hypercholesterolemia: Mutations within the low-density lipoprotein receptor gene are one of the major causes of familial hypercholesterolemia. Due to the defect in this receptor, LDL accumulates in the plasma leading to enhanced atherosclerosis.

Tangier disease: Mutations in the gene encoding the cholesterol transporter ABCA1 result in very low levels of HDL particles in blood (Bodzioch *et al.*, 1999). Tangier disease is inherited as an autosomal recessive trait. There is excessive accumulation of cholesterol in several parts of the body due to impaired reverse cholesterol transport and the disease is associated with increased atherosclerosis.

SLOS: SLOS is a developmental disorder characterized by several physical, mental, and behavioral abnormalities related to deficiency of cholesterol synthesis. DHCR7 codes for the enzyme involved in the last step in cholesterol synthesis. Mutation in this gene causes the buildup of the 7-dehydrocholesterol metabolite, resulting in insufficient cholesterol synthesis; both combined are responsible for the multiple developmental disorders.

Neurological disorders: Defects in cholesterol metabolism in the brain have been associated with many neurological disorders such as Alzheimer's disease, Parkinson's disease, Huntington's disease, Niemen-Pick disease, and SLOS (Vance, 2012). The complexity of these diseases together with very limited knowledge of cholesterol metabolism in brain has proved to be a major challenge for scientists trying to study the pathology of the disease.

Steroids

Cholesterol is the precursor of numerous sterol molecules such as oxysterols, bile acids, and steroid hormones. These cholesterol metabolites function as activators of nuclear receptors, regulate expression of genes in lipid and fat metabolism, have important function in bile acid synthesis and steroid hormone production, and are required for the transport of sterols from peripheral tissues such as brain to the liver.

Bile acids are oxidized derivatives of cholesterol. These are amphipathic molecules required for processing of dietary fat and for maintaining cholesterol homeostasis. Conversion of cholesterol into bile is the only mechanism of cholesterol excretion from the body. Bile is formed in the liver by the action of many enzymes and cholesterol 7- α hydroxylase (CYP7A1) is the rate-limiting enzyme for bile acid synthesis (Jelinek *et al.*, 1990).

Cholesterol also serves as a precursor for all steroid hormones. These hormones regulate many important physiological processes such as stress response, cognition, ion balance, secondary sexual characteristics, electrolyte balance, and carbohydrate metabolism. Steroid hormones are produced primarily in response to trophic hormones in several steroidogenic tissues including adrenal cortex, testes, ovary, placenta, and brain.

During steroid synthesis, cholesterol in the cell is transported to the mitochondria where the enzyme Cyp11 converts it to pregnenolone, which is the common precursor of all steroid hormones (Figure 6; Miller and Auchus, 2011). Since cholesterol cannot travel from outer to inner mitochondrial membrane due to its hydrophobicity, the delivery of cholesterol to the inner membrane is the rate-limiting step in steroidogenesis. The enzyme steroidogenic acute regulatory protein (STAR) regulates this rate-limiting step by facilitating the transfer of cholesterol into the inner membrane of mitochondria (Stocco, 2001; Miller and Auchus, 2011).

Testosterone, progesterone, estradiol, cortisol, and aldosterone are five major classes of steroids. These hormones are all derived from pregnenolone and control a wide variety of physiological processes. Progesterone is required for maintenance of pregnancy, follicular growth, and ovulation and is produced by ovarian cells and corpus luteum (Graham and

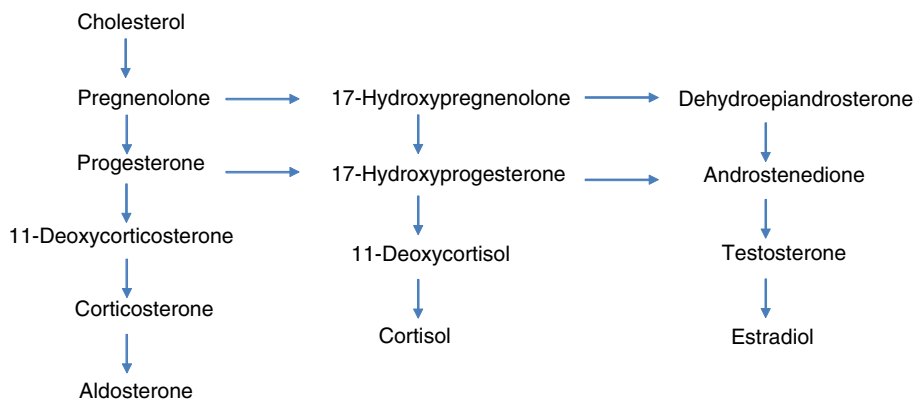


Figure 6 The biosynthetic pathway of steroid hormones.

Clarke, 1997). Cortisol is produced by adrenal cortex and has a number of diverse functions including modulating blood pressure and Na⁺ uptake, inflammation, glucose metabolism, and stress response. Testicular leydig cells produce testosterone, which is critical for the development and maintenance of male sex characteristics and functions. Similarly, estrogen, produced by ovary, is required for the maintenance and development of female sex characteristics and is also found to be important for maintaining bone integrity. Finally, aldosterone, produced in adrenal cortex, regulates blood pressure and enhances sodium reabsorption in the kidney and sweat glands (Müller, 1995). Apart from these classical steroids, neurosteroids are another class of steroids synthesized in the brain. These steroids are found to play a critical role in neuronal survival, plasticity, and functions (Sarkar *et al.*, 2011; Di Michele *et al.*, 2013; Tsutsui, 2012).

See also: Molecular Principles, Components, Technology, and Concepts: Lipids: Lipidomics

References

- Barter, P.J., 2003. Cholesteryl ester transfer protein: A novel target for raising HDL and inhibiting atherosclerosis. *Arteriosclerosis, Thrombosis, and Vascular Biology* 23, 160–167.
- Bengoechea-Alonso, M.T., Ericsson, J., 2007. SREBP in signal transduction: Cholesterol metabolism and beyond. *Current Opinion in Cell Biology* 19, 215–222.
- Beyea, M.M., Heslop, C.L., Sawyez, C.G., *et al.*, 2007. Selective up-regulation of LXR-regulated genes ABCA1, ABCG1, and APOE in macrophages through increased endogenous synthesis of 24(S),25-epoxycholesterol. *Journal of Biological Chemistry* 282, 5207–5216.
- Björkhem, I., Lütjohann, D., Diczfalusy, U., *et al.*, 1998. Cholesterol homeostasis in human brain: Turnover of 24S-hydroxycholesterol and evidence for a cerebral origin of most of this oxysterol in the circulation. *Journal of Lipid Research* 39, 1594–1600.
- Bloch, K.E., 1983. Sterol structure and membrane function. *CRC Critical Reviews in Biochemistry* 14, 47–92.
- Bodzioch, M., Orsó, E., Klucken, J., *et al.*, 1999. The gene encoding ATP-binding cassette transporter 1 is mutated in Tangier disease. *Nature Genetics* 22, 347–351.
- Brock, W.H., 2000. Proteins, enzymes, genes: The interplay of chemistry and biology. *Medical History* 44, 409–410.
- Brown, A.J., Sun, L., Feramisco, J.D., Brown, M.S., Goldstein, J.L., 2002. Cholesterol addition to ER membranes alters conformation of SCAP, the SREBP escort protein that regulates cholesterol metabolism. *Molecular Cell* 10, 237–245.
- Brown, D.A., London, E., 2000. Structure and function of sphingolipid- and cholesterol-rich membrane rafts. *Journal of Biological Chemistry* 275, 17221–17224.
- Brown, M.S., Goldstein, J.L., 1980. Multivalent feedback regulation of HMG CoA reductase, a control mechanism coordinating isoprenoid synthesis and cell growth. *Journal of Lipid Research* 21, 505–517.
- Chang, T.Y., Chang, C.C., Cheng, D., 1997. Acyl-coenzyme A: Cholesterol acyltransferase. *Annual Review of Biochemistry* 66, 613–638.
- Correll, C.C., Ng, L., Edwards, P.A., 1994. Identification of farnesol as the non-sterol derivative of mevalonic acid required for the accelerated degradation of 3-hydroxy-3-methylglutaryl-coenzyme A reductase. *Journal of Biological Chemistry* 269, 17390–17393.
- DeBose-Boyd, R.A., 2008. Feedback regulation of cholesterol synthesis: Sterol-accelerated ubiquitination and degradation of HMG CoA reductase. *Cell Research* 18, 609–621.
- Dietschy, J.M., Turley, S.D., 2001. Cholesterol metabolism in the brain. *Current Opinion in Lipidology* 12, 105–112.
- Di Michele, F., Luchetti, S., Bernardi, G., Romeo, E., Longone, P., 2013. Neurosteroid and neurotransmitter alterations in Parkinson's disease. *Frontiers in Neuroendocrinology* 34, 132–142.
- Dowhan, W., Bogdanov, M., Mileyskaya, E., 2008. Functional roles of lipids in membranes. *Biochemistry of Lipids, Lipoproteins and Membranes*. Amsterdam: Elsevier Press, pp. 1–37.
- Edwards, P.A., Ericsson, J., 1999. Sterols and isoprenoids: Signaling molecules derived from the cholesterol biosynthetic pathway. *Annual Review of Biochemistry* 68, 157–185.
- Fouchier, S.W., Defesche, J.C., 2013. Lysosomal acid lipase A and the hypercholesterolaemic phenotype. *Current Opinion in Lipidology* 24, 332–338. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23652569> (accessed 24.02.15).
- George, K.S., Wu, S., 2012. Lipid raft: A floating island of death or survival. *Toxicology and Applied Pharmacology* 259, 311–319.
- Graham, J.D., Clarke, C.L., 1997. Physiological action of progesterone in target tissues. *Endocrine Reviews* 18, 502–519.
- Ginsburg, G.S., Atkinson, D., Small, D.M., 1984. Physical properties of cholesteryl esters. *Progress in Lipid Research* 23, 135–167.
- Go, G.W., Mani, A., 2012. Low-density lipoprotein receptor (LDLR) family orchestrates cholesterol homeostasis. *Yale Journal of Biology and Medicine* 85, 19–28.
- Horton, J.D., Goldstein, J.L., Brown, M.S., 2002. SREBPs: Activators of the complete program of cholesterol and fatty acid synthesis in the liver. *Journal of Clinical Investigation* 109, 1125–1131.
- Ikonen, E., 2008. Cellular cholesterol trafficking and compartmentalization. *Nature Reviews Molecular Cell Biology* 9, 125–138.
- Jelinek, D.F., Andersson, S., Slaughter, C.A., Russell, D.W., 1990. Cloning and regulation of cholesterol 7 α -hydroxylase, the rate-limiting enzyme in bile acid biosynthesis. *Journal of Biological Chemistry* 265, 8190–8197.
- Laffitte, B.A., Repa, J.J., Joseph, S.B., *et al.*, 2001. LXRs control lipid-inducible expression of the apolipoprotein E gene in macrophages and adipocytes. *Proceedings of the National Academy of Sciences of the United States of America* 98, 507–512.
- Lemaire-Ewing, S., Lagrost, L., Néel, D., 2012. Lipid rafts: A signalling platform linking lipoprotein metabolism to atherogenesis. *Atherosclerosis* 221, 303–310.
- Maxfield, F.R., Tabas, I., 2005. Role of cholesterol and lipid organization in disease. *Nature* 438, 612–621.
- Miller, W.L., Auchus, R.J., 2011. The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. *Endocrine Reviews* 32, 81–151.
- Motoshima, H., Goldstein, B.J., Igata, M., Araki, E., 2006. AMPK and cell proliferation – AMPK as a therapeutic target for atherosclerosis and cancer. *Journal of Physiology* 574, 63–71.
- Müller, J., 1995. Aldosterone: The minority hormone of the adrenal cortex. *Steroids* 60, 2–9.
- Nohturfft, A., Yabe, D., Goldstein, J.L., Brown, M.S., Espenshade, P.J., 2000. Regulated step in cholesterol feedback localized to budding of SCAP from ER membranes. *Cell* 102, 315–323.
- Orth, M., Bellosta, S., 2012. Cholesterol: Its regulation and role in central nervous system disorders. *Cholesterol* 2012, 1–19.
- Peelman, F., Vandekerckhove, J., Rosseneu, M., 2000. Structure and function of lecithin cholesterol acyl transferase: New insights from structural predictions and animal models. *Current Opinion in Lipidology* 11, 155–160.
- Sarkar, J., Wakefield, S., MacKenzie, G., Moss, S.J., Maguire, J., 2011. Neurosteroidogenesis is required for the physiological response to stress: Role of neurosteroid-sensitive GABA_A receptors. *Journal of Neuroscience* 31, 18198–18210.
- Simons, K., Ikonen, E., 2000. How cells handle cholesterol. *Science* 290, 1721–1726.
- Simons, K., Toomre, D., 2000. Lipid rafts and signal transduction. *Nature Reviews Molecular Cell Biology* 1, 31–39.
- Stocco, D.M., 2001. StAR protein and the regulation of steroid hormone biosynthesis. *Annual Review of Physiology* 63, 193–213.
- Subtil, A., Gaidarov, I., Kobylarz, K., *et al.*, 1999. Acute cholesterol depletion inhibits clathrin-coated pit budding. *Proceedings of the National Academy of Sciences of the United States of America* 96, 6775–6780.
- Tabas, I., 2002. Consequences of cellular cholesterol accumulation: Basic concepts and physiological implications. *Journal of Clinical Investigation* 110 (7), 905–911.
- Tsutsui, K., 2012. Neurosteroid biosynthesis and action during cerebellar development. *Cerebellum* 11 (2), 414–415.
- Vance, J.E., 2012. Dysregulation of cholesterol balance in the brain: Contribution to neurodegenerative diseases. *Disease Models & Mechanisms* 5, 746–755.
- Wang, D.Q.-H., 2007. Regulation of intestinal cholesterol absorption. *Annual Review of Physiology* 69, 221–248.
- Yang, T., Espenshade, P.J., Wright, M.E., *et al.*, 2002. Crucial step in cholesterol homeostasis: Sterols promote binding of SCAP to INSIG-1, a membrane protein that facilitates retention of SREBPs in ER. *Cell* 110, 489–500.

Glycolipids

AH Merrill Jr. and MN Vu, Georgia Institute of Technology, Atlanta, GA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Depending on how you look at them, glycolipids are either enchanting biomolecules, being comprised of the most structurally diverse and complex families of compounds (carbohydrates and lipids), or they seem so mind-bogglingly complicated that you can hardly bear to think about them. To help minimize the reader's inclination toward the latter attitude, this article will summarize the types of compounds that fall in this category of biomolecules and clarify the fairly straightforward structural features that define them. Then, the review will illustrate with one category, the glycosphingolipids, how such structural complexity is generated biochemically and how thinking about the compounds in pathway maps simplifies the complexity somewhat. Although the enormity of this field precludes discussion of every type of glycolipid, the terms and concepts covered in this review should help the newcomer feel less intimidated about approaching them, and perhaps even those with knowledge about glycolipidology will be surprised by how many different ways nature has married these two families of compounds.

Definition of 'Glycolipid' and Subcategories of Glycolipids

The term 'glyco-' refers to the presence of a 'glycan' as part of the compound, which is defined by the glossary from Essentials of Glycobiology (Varki *et al.*, 2009) (available online) as 'a generic term for any sugar or assembly of sugars, in free form or attached to another molecule' and is often used interchangeably with carbohydrate or saccharide.

The complexity of glycobiology arises from the existence of many different carbohydrates (glucose, galactose, etc.) multiplied by the many ways they can be combined in polysaccharides. If all of the theoretical combinations of these compounds were produced, the numbers would be staggering because it has been estimated that six different hexoses can be combined to form $>10^{12}$ different hexasaccharides, $\sim 10^{15}$ heptasaccharides, $>10^{18}$ octasaccharides, and nearly Avogadro's number for nonasaccharides (Laine, 1994).

The term 'lipid' refers to the portion of these compounds that (as has been summarized by the LIPID MAPS Consortium) (Fahy *et al.*, 2005) '... have been loosely defined as biological substances that are generally hydrophobic in nature and in many cases soluble in organic solvents...' Lipids are, thus, also structurally diverse and have been subdivided into eight categories: fatty acyls, glycerolipids, glycerophospholipids, sphingolipids, sterol lipids, prenol lipids, saccharolipids, and polyketides. At least a few of the lipids from each of these categories have been found as glycoconjugates in nature, and for some categories (such as sphingolipids) the discrete molecular subspecies for the lipid backbone probably numbers in the thousands.

Thus, this review considers glycolipids to encompass all of the compounds that are comprised of a carbohydrate and a lipid. The term lipid-linked oligosaccharide (LLO) is sometimes also used for glycolipids, but is most often applied to the particular case of the dolichol pyrophosphate-linked oligosaccharide that is transferred to asparagine residues of nascent polypeptides in the endoplasmic reticulum during N-linked glycoprotein biosynthesis (this will be discussed under prenol lipids below).

Overview of the Different Types of Glycolipids Found in Nature

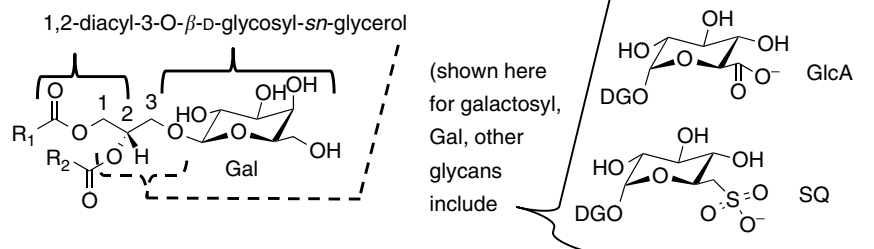
There are several useful web sites for glycolipids that provide information about their structures and nomenclature, tools to facilitate their analysis, metabolic pathways and genes, and sometimes updates and notices of upcoming conferences. Of note are the American Oil Chemistry Society (AOCS) Lipid Library, Glycoforum, LIPID MAPS, and the Consortium for Functional Glycomics. In an attempt to obtain uniformity in nomenclature and the way to display the structures of these compounds, the LIPID MAPS lipid classification system (Fahy *et al.*, 2005) has been adopted by some journals, and the website provides downloadable structure diagrams for many glycolipids (when used in this review, the LIPID MAPS number, 'LM____,' has been provided). Other 'toolboxes' have recently been summarized (Campbell *et al.*, 2014), and the list continues to grow as this field does. It is also helpful sometimes to consult the glycolipid nomenclature recommendations of the International Union of Pure and Applied Chemistry and the International Union of Biochemistry and Molecular Biology (IUPAC-IUB) Joint Commission on Biochemical Nomenclature. This review does not attempt to cover all the structural variations for all organisms, which would require a much longer treatise, but to illustrate some of the major features and the fascinating differences among these compounds.

Glycoglycerolipids

The compounds termed glycoglycerolipids typically have a carbohydrate connected to a 1,2-diacyl-*sn*-glycerol moiety (as shown in the prototype in Figure 1) where R1 and R2 are typically long-chain fatty acids in ester linkage to the glycerol backbone. In some cases, one or more of the glycerol hydroxyls is linked to a fatty alcohol in ether linkage, as shown for seminolipid in the bottom example in this figure. Since glycerol *per se* does not have stereochemistry, the '*sn*' designation of glycerolipids (standing for 'stereospecific numbering') takes into account that stereoisomers are generated once asymmetric substituents have been added to the hydroxyls. The structures shown in Figure 1 and elsewhere in this review are numbered following this convention. Likewise, commonly accepted names and abbreviations have been used for the fatty acids and fatty alcohols (see online LIPID MAPS

Glycoglycerolipids

Prototype:



Examples:

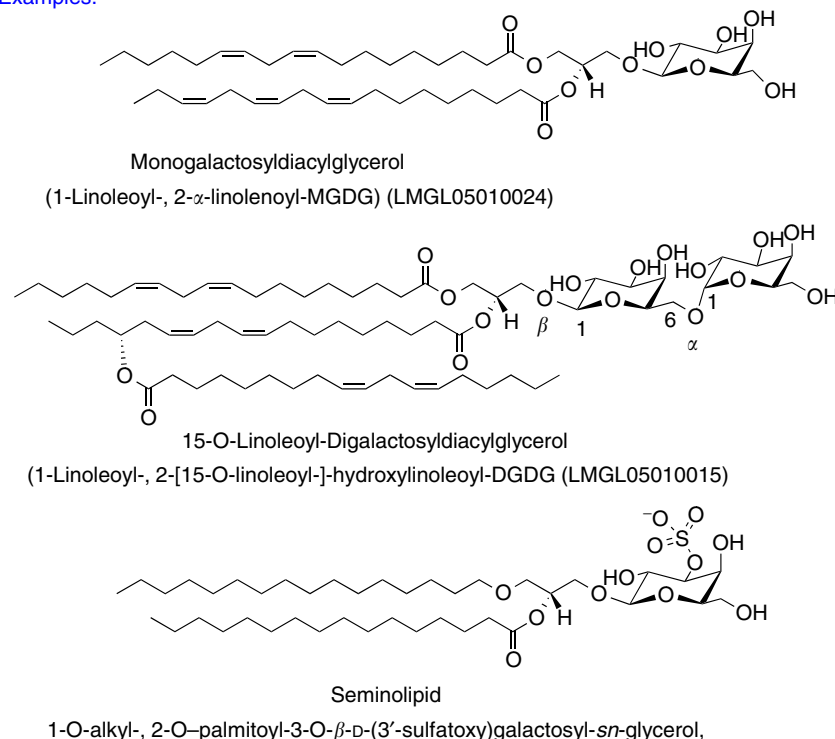


Figure 1 Representative structures of glycoglycerolipids. A general prototype for these compounds is shown in the upper diagram with R1 and R2 representing the alkyl chains of fatty acids attached to the glycerol backbone at these two positions and at position 3 is attached the glycan. The abbreviations for the carbohydrates shown are: Gal, galactose; Glc, glucose; GlcA, glucuronic acid; SQ, sulfoquinovose. Examples of specific compounds are shown below with common and full names, and where the structures were downloaded from the LIPID MAPS Website, with the 'LM' number.

Lipid Classification system); i.e., for the first number to designate how many carbon atoms are in the alkyl chain followed by a slash, colon or semicolon and a second number for how many double bonds are present. Unless otherwise stated, it is assumed that the double bonds are *cis*- (Z, in the E/Z designation of *trans/cis* double bond configurations). Fatty acids can have other features, such as hydroxyl-, methyl-, cyclopropyl- and other groups, and a somewhat unusual, but interesting, backbone is seen in 15-O-linoleoyl-digalactosyldiacylglycerol (Figure 1, middle), which was isolated from oat seeds. Its lipid diacylglycerol backbone has a hydroxyl-fatty acid with a third fatty acid in ester linkage to this hydroxyl – a feature found in so-called 'estolides.'

The carbohydrate that is attached to the 3-hydroxyl of glycerol can be in α - or β -glycosidic linkage (β -linked galactose,

abbreviated Gal, is shown in all of the examples in Figure 1) and this type of monosaccharide is usually abbreviated 'MGDC' for monogalactosyldiglyceride. Other carbohydrates include those to the right of the prototype (glucose, Glc; glucuronic acid, GlcA; and sulfoquinovose, SQ) as well as mannose, rhamnose and aminosugars (not shown). Di- and polysaccharides are designated by the type of linkage (α - or β -) and location, with the anomeric carbon numbered 1 followed by the site of attachment to the neighboring carbohydrate (e.g., 1 $\rightarrow$ 2, 1 $\rightarrow$ 3, 1 $\rightarrow$ 4, or 1 $\rightarrow$ 6). This nomenclature is shown for the digalactosyldiglyceride (DGDG) in Figure 1. In some cases, hydroxyls of the carbohydrates are esterified with fatty acids.

Glycoglycerolipids are commonly found in plants, algae and bacteria, and are the predominant lipids in chloroplasts of

many organisms (Holzl and Dormann, 2007; Zhang *et al.*, 2014). They have attracted interest for their potential use as antiviral, antitumor, and anti-inflammatory agents. Glycoglycerolipids are less prevalent in mammals, with the exception of testis where seminolipid (the bottom structure of Figure 1) has been reported to comprise more than 90% of total glycolipid (Honke, 2013).

Glycophosphoglycerolipids

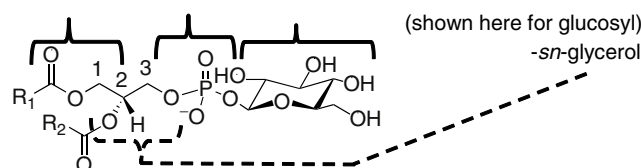
For this article, the term glycophosphoglycerolipids is used to define the lipids that are comprised of a phosphatidyl-backbone (i.e., 1,2-diacylglycerol 3-phospho-) (or in some cases, O-alkyl rather than acyl) for the lipid moiety and a carbohydrate headgroup attached in a phosphodiester linkage, as

shown in Figure 2. The AOCS Lipid Library describes a similar category and notes that the term conveniently describes these compounds even if it does not appear to be widely used. The Library also notes that there are compounds that can be referred to as 'phosphoglycerolipids' because they are glycolipids in which the sugar moiety is phosphorylated, but that is distinct from the compounds discussed here.

The most widely known glycophosphoglycerolipids are undoubtedly the inositol phospholipids (Figure 2), which are ubiquitously distributed in Eukarya, most of Archaea, and some Bacteria (Michell, 2013). The lipid backbones differ considerably between eukaryotes/bacteria (which are primarily 1,2-diacylglycerols) versus Archaea, which have two isoprenoid chains attached to positions 2,3 of glycerol via

Glycophosphoglycerolipids

Prototype: 1,2-diacyl-3-phospho- β -D-glycan (or phosphoglycan)



Examples:

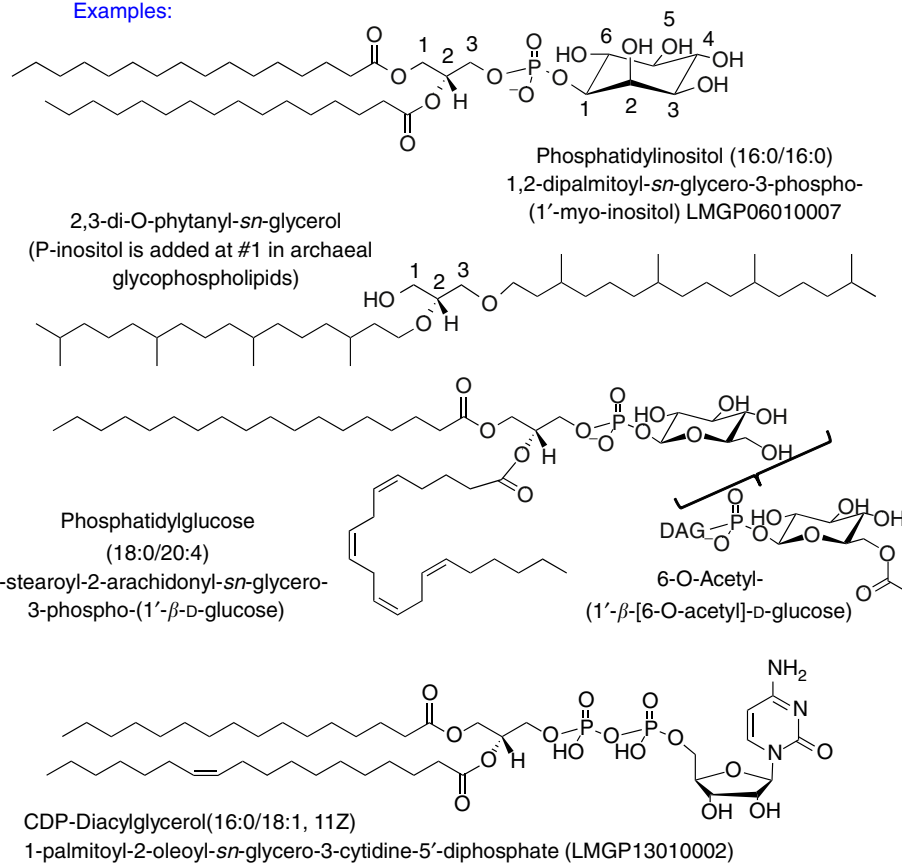


Figure 2 Representative structures of glycophosphoglycerolipids. A general prototype for these compounds is shown in the upper diagram with R1 and R2 representing the alkyl chains of fatty acids attached to the glycerol backbone at these two positions and at position 3 is attached the glycan. Also shown is the phytanyl backbone that is found in archaeal glycophosphoglycerolipids and a lipid used as an intermediate in phosphoglycerolipid biosynthesis (CDP-diacylglycerol) that falls under this category because the CDP-moiety contains a sugar. Where a structure was downloaded from the LIPID MAPS Website, the 'LM' number has been given.

ether bonds, as shown by 2,3-diphytanyl-*sn*-glycerol (Morii *et al.*, 2014) in Figure 2. Much of the interest in inositol phospholipids revolves around two further structural modifications: (1) phosphorylation of the inositol at positions 3, 4, and 5 in different combinations and the attendant biological functions of these compounds in normal cell behavior and disease (Michell, 2013; Balla, 2013); and (2) extension of the glycan headgroup with glucosamine, mannose, and phosphoethanolamine groups to form a 'Glycosylphosphatidylinositol Anchor' (GPI) for some categories of membrane bound proteins (Paulick and Bertozzi, 2008).

Phosphatidylglucose (Figure 2) is another glycosphingolipid that has been isolated from several mammalian sources, with the interesting feature that the lipid backbone from neutrophils has saturated fatty acyl chains (C18:0/C18:0 and C18:0/C20:0), whereas fetal rodent brain has 1-stearyl-2-arachidoyl- as shown in Figure 2. A portion has an acetylated glucose (6-O-acetylglucose), as is also shown

(Ishibashi *et al.*, 2013). The functions of these lipids are still being elucidated.

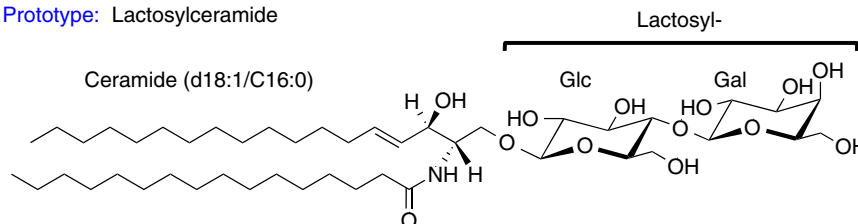
The glycosphingolipid category also includes compounds with a pyrophosphate linkage to a carbohydrate-containing headgroup, as seen in CDP-diacylglycerol, a biosynthetic intermediate for phospholipid biosynthesis by many organisms (Figure 2).

Glycosphingolipids

Glycosphingolipids are defined as glycolipids with lipid backbones called 'ceramides' that are comprised of sphingoid bases (hence, the origin of the 'sphingo' of sphingolipids) that are N-acylated with a long- or very-long-chain fatty acid as shown in Figure 3 (Merrill, 2011). Although most depictions of sphingolipids show sphingosine as the backbone (as in this prototype), a large number of structural variations have been reported (Pruett *et al.*, 2008), particularly with respect to alkyl chain length, number and position of double bonds and

Glycosphingolipids

Prototype: Lactosylceramide



Examples:

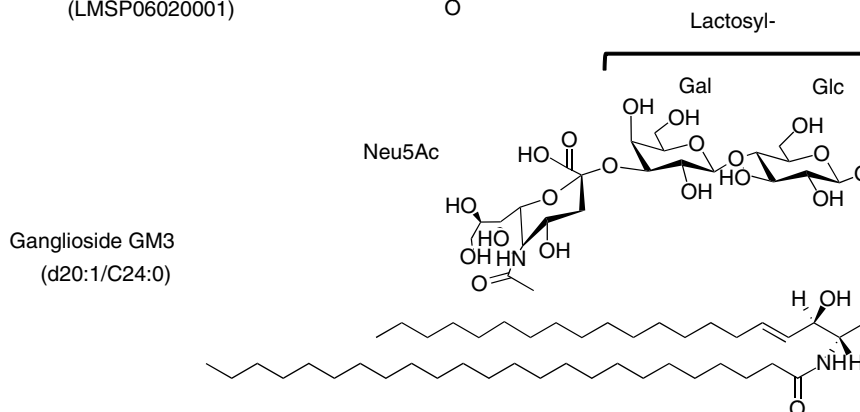
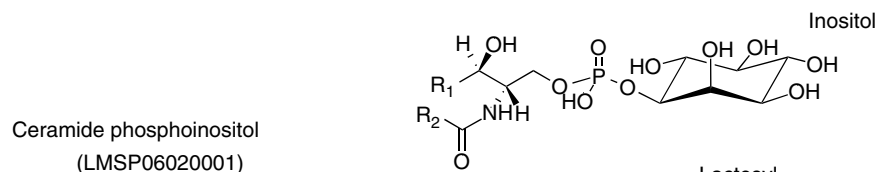
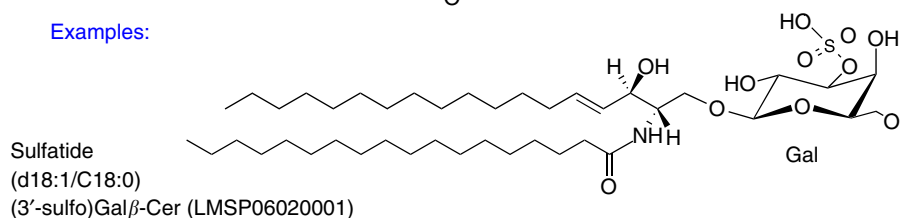


Figure 3 Representative structures of glycosphingolipids. A general prototype for these compounds is shown in the upper diagram using lactosylceramide. The other examples also illustrate structural variations, and for further examples see Figures 4, 10 and 11. Where a structure was downloaded from the LIPID MAPS Website, the 'LM' number has been given.

number of hydroxyls. The shorthand abbreviations for sphingoid base are: (1) to depict the number of hydroxyls by 'd' (dihydroxy, as in sphingosine), 't' (trihydroxy, which is most commonly found in 4-hydroxysphinganine, or 'phyto-sphingosine' – a major species in epithelial cells and skin, plants and fungi) and 'm' (monohydroxy) (1-deoxy-sphingoid bases are interesting because they are made when serine palmitoyltransferase (SPT) utilizes alanine instead of serine (Zitomer *et al.*, 2009; Penno *et al.*, 2010), but lacking the usual site of glycosylation, they are probably not backbones for mammalian glycolipids); (b) to depict the number of carbon atoms and double bonds in a manner similar to fatty acids (i.e., 18:1 for the 18-carbon atoms and single double bond of sphingosine) – sphingoid bases with no double bonds are also called 'sphinganines,' those with one double bond are 'sphingenines' (although 'sphingosine' is often used for d18:1 by tradition), and those with two double bonds are sphingediene; (c) if the double bond position is not specified, it is assumed to be between carbons 4 and 5 and *trans* (E), as shown. The stereochemical relationship between the 2-amino and 3-hydroxy-groups are as shown (termed 'D-erythro-' or '2 S,3 R' for sphingosine), although some organisms have sphingoid bases that deviate from this stereochemistry (Pruett *et al.*, 2008).

The fatty acids of ceramides are usually 14–26 carbons in length (and longer for skin), with one or no double bonds and sometimes an α -hydroxyl (and, again, as a special case for skin, an ω -hydroxyl). Although there is often less structural variation in the fatty acids of sphingolipids than for glycerolipids, the number of backbone subspecies can still be considerable when variation in the sphingoid base and fatty acids are both considered. For example, a LIPID MAPS analysis (Quehenberger *et al.*, 2010) of human plasma sphingolipids using liquid chromatography-tandem mass spectrometry (Shaner *et al.*, 2009) found ~100 sphingomyelin subspecies varying in the ceramide backbone due to the presence of several sphingoid bases and amide-linked fatty acids in numerous combinations.

Multiple categories of complex glycosphingolipids are made by addition of carbohydrate headgroups at carbon 1. The major subcategories (highlighted by the first three structures in Figure 3) are: (1) glucosylceramides (GlcCer) and downstream metabolites; (2) galactosylceramides (GalCer) and downstream metabolites; and (c) ceramide phosphoinositols (which are common in plants, fungi, and many other organisms, but not in mammals) that can be further elaborated by additional sugars (such as mannose). Ceramide phosphoinositols are also utilized for formation of the GPI anchor for membrane proteins in some organisms, such as yeast (note that these are often referred to as GPI anchors although the lipid backbone is not a 'phosphatidyl-' group). Small amounts of 'lyso-' glycosphingolipids (i.e., sphingoid bases plus a headgroup but lacking the N-acyl-substituent), such as 'psychosine' (1- β -D-galactosylsphingosine) are occasionally encountered, especially in disease.

After these ceramide monohexoses have been made, additional carbohydrates are added to produce an astonishing variety of complex glycosphingolipids. Indeed, the total number of glycan variants is not known, but in 2007, there were approximately 400 structurally characterized headgroups (174 neutral glycosphingolipids, 190 gangliosides and 24

sulfated glycosphingolipids) (Yu *et al.*, 2007). If one adds compounds that have been subsequently discovered and theoretical compounds (intermediates that are likely to exist between two known compounds in a plausible biosynthetic scheme for the larger compound), the total is probably closer to 600 (Merrill, 2011) (the subspecies can be seen online at SPHINGOMAP). This is a large number, but considerably fewer than Avogadro's number.

The majority of the more complex glycosphingolipids are made from GlcCer, and for mammals have been assigned to five 'root structure' families: globo-, isoglobo-, lacto-, neolacto-, and ganglio-) based on first four carbohydrates (designated I to IV), which are shown in Figure 4. This figure also illustrates a useful shorthand for depiction of the carbohydrates from the Consortium for Functional Glycomics (available online) in which the sugars are designated by color (e.g., Glc vs. Gal as blue vs. yellow circles, respectively) and shape (e.g., Glc as a blue circle but N-acetyl-glucosamine, NAcGlc, as a blue square). The linkage between the sugars is also indicated; for example, β 4 inside the Gal of lactosylceramide implies a β 1 $\rightarrow$ 4 bond between Gal II and Glc I (this is written between the symbols rather than inside them) and the orientation of the symbols also indicates the linkage (note for the symbolic representation of ganglioside GM1 that the substituents on Gal II are linked to positions 3 and 4 in the Lewis diagram (i.e., at ca 7 o'clock and 10 o'clock) are displayed in analogous positions in the symbolic representation).

The functions of glycosphingolipids are too diverse to include in this overview, but the underlying theme, whether a simple monohexose containing compound (Ishibashi *et al.*, 2013) or a complex ganglioside (Schnaar *et al.*, 2014), appears to be that these compounds behave differently in biological membrane (therefore and help form regions with special purposes) and the glycan moieties offer a rich library of structures for interactions with cellular proteins (receptors, transporters, membrane trafficking machinery, etc.), extracellular matrix proteins and proteins on neighboring cells, components of immunoregulation, etc. There is also evidence that some of the glycosphingolipids interact sufficiently strongly with each other that they might fulfill the function of both the ligand and target. Furthermore, some sphingolipids undergo substantial metabolism as part of fulfilling their functions, such as removal of sialic acid or release of the bioactive lipid backbone.

Glycosylated sterols

The glycosylated sterols (also called steryl glycosides) refer to a range of sterols (reflective of the type of organism – cholesterol for mammals and sitosterol, ergosterol, campesterol, stigmasterol, brassicasterol and many others for vascular plants, fungi, algae, and other marine organisms – such as starfish and sponges, and some bacteria) that are glycosylated at the 3 β -hydroxyl group with one or more sugars (glucose, galactose, xylose, arabinose, glucuronic acid, and others – again, depending on the organism) (Grille *et al.*, 2010; Ivanchina *et al.*, 2011). The structural prototype is shown with cholesteryl β -D-glucoside in Figure 5 with an example of a more structurally complex steryl glycoside (Mycaloside G from the Caribbean marine sponge *Mycale laxissima*) (Ebada *et al.*, 2010). Some are also acylated on the glycan.

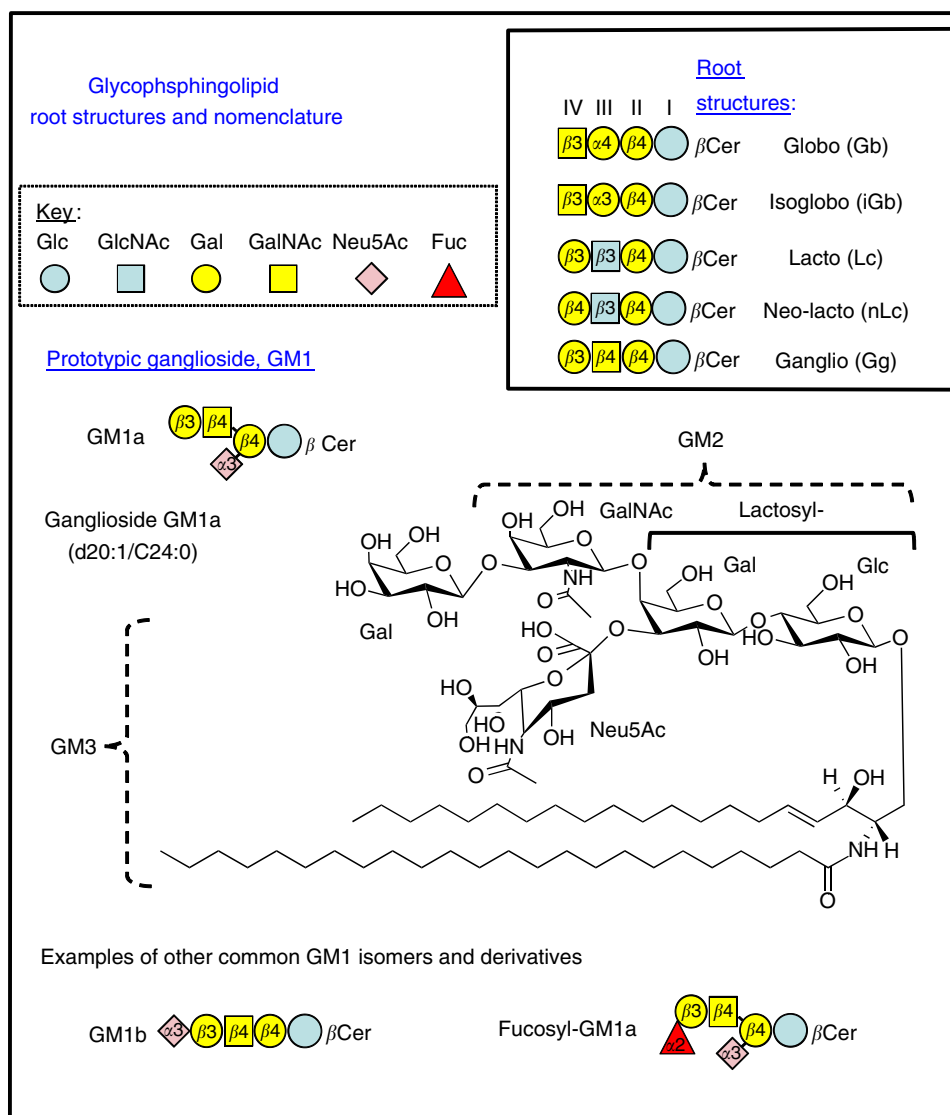


Figure 4 Representative structures of glycosphingolipids – root structures and nomenclature. The structures for these glycosphingolipids are shown using the shape/color symbols that are conventionally used in the glycobiology field (see text). The symbols are defined in the upper left, and an example is given for ganglioside GM1a which is shown both by an explicit structural diagram and symbols. The upper right shows the root structures for glycosphingolipids using the symbol representation. The bottom of the figure shows other types of GM1, an isomeric form (GM1b) and fucosyl-GM1a.

The functions of sterol glycosides are still being elucidated, but they appear to affect the biophysical properties of cell membranes, confer resistance against stresses such as freezing in plants, play a role in seed development and modulation of host immune functions (Grille *et al.*, 2010; Shimamura, 2012). The presence of cholesterol glucoside in mammalian cells has only been recently appreciated, and is interestingly formed from glucosylceramide as the sugar donor and synthesis is induced by heat shock (Ishibashi *et al.*, 2013).

Glycosylated prenols

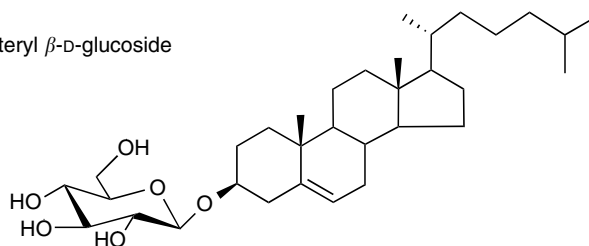
Glycosylated prenols are comprised of a polyisoprene (polyterpenoid) alcohol that is attached to a glycan via glycosidic, phosphate or pyrophosphate bonds, as seen in the prototype and examples in Figure 6. Dolichol (eukaryotes and archaea)

and bactoprenol (bacteria) are used in sugar transport across membranes and transfer of carbohydrates to acceptors, with attachment of the carbohydrate to the hydroxyl on the terminal isoprene via a phosphodiester for simple monohexoses (Dol-P-Man and Dol-P-Glc) and a pyrophosphate for the larger oligosaccharide (GlcNAc₂Man₉Glc₃), which is used for N-linked glycoprotein biosynthesis (Welti, 2013). The number of isoprene units varies from 18 to 22 for dolichols and 10 to 12 for bactoprenols, and there are also differences in the degree of unsaturation among different sources. The structural variations are becoming more evident as mass spectrometric methods are facilitating the analysis of these compounds (Garrett *et al.*, 2007; Guan and Eichler, 2011).

The structure shown by glycan symbols in Figure 6, Glc₃Man₉GlcNAc₂-PP-Dol, is used for protein N-glycosylation

Glycosylated sterols

Prototype: cholesteryl β -D-glucoside



Example:

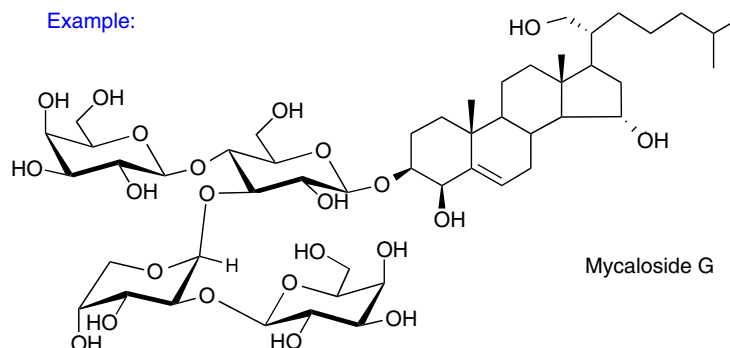


Figure 5 Representative structures of glycosylated sterols. A general prototype for these compounds is shown by cholesteryl β -D-glucoside at the top. An example of an oligosaccharide linked to a sterol is shown by mycaloside G.

in the lumen of the endoplasmic reticulum. This and related compounds in the pathway are often referred to as lipid-linked oligosaccharides (LLO) and have been characterized in considerable detail because defects in the pathway are known to cause disease (Welti, 2013).

An example of a polyisoprenol attached to a carbohydrate via glycosidic linkage is also shown in Figure 6 for a plakopolyprenoside, a compound that has been isolated from a marine sponge, *Plakortis simplex* (Costantino *et al.*, 2000).

Saccharolipids

These are compounds in which fatty acids are linked directly to a sugar backbone, as shown in Figure 7. One of the most widely known saccharolipids is the glucosamine-based saccharolipid 'lipid A' component of lipopolysaccharide (LPS) (Raetz *et al.*, 2009) (Figure 7). They show considerable structural variation among organisms (Banoub *et al.*, 2010) and have complex functions and effects on the GI tract (Rhee, 2014). Another example of a saccharolipid is the acylated trehalose of *Mycobacteria* (Rombouts *et al.*, 2011).

Glycosylated polyketides

These are a structurally diverse group of natural products that are comprised of a 'lipidic' backbone (in the sense that it is relatively soluble in organic solvents and biosynthetically derived from acetyl- or malonyl- units) and a glycan (Hertweck, 2009). The structures of glycosylated polyketides are so diverse that no single figure can encompass all of the members of this subclass, so two examples, avermectin and erythromycin, are shown in Figure 8.

The lipidic portion of polyketides encompasses a wide range of structural features, including polyphenols, macrolides, polyenes, enediynes, and polyethers. The glycosyl groups are

also structurally diverse, but have been suggested to offer an easier handle to 'mine' for new compounds using glygenomics and mass spectrometry (Kersten *et al.*, 2013).

These compounds are produced by a wide range of organisms and include many common antimicrobial, antiparasitic, and anticancer agents, such as the avermectin and erythromycin shown in Figure 8.

Glycosphingolipids – The Power of Combinatorial Biochemistry

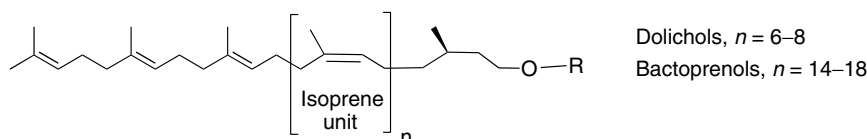
The term 'combinatorial' was applied to chemical synthesis when methods were developed to make large numbers of compounds by relatively simple operations. Its aptness for sphingolipid biosynthesis was soon noted (Hannun *et al.*, 2001; Kolter *et al.*, 2002), and continues to be pertinent as more genes and enzymes for glycolipid biosynthesis are found. In essence, then, the idea is that a very large number of glycolipids can be made because there is a large number of enzymes that make the lipid backbones and a large number of glycosyltransferases that can elaborate the headgroups. However, this also limits the number of products that can be made due to the specificities of these enzymes, and because different combinations are expressed in different tissues. Other factors that can affect the types of compounds that are made include subcellular localization of the enzymes and substrates, the presence of binding and transport proteins, etc.

Sphingolipid Backbone Biosynthesis

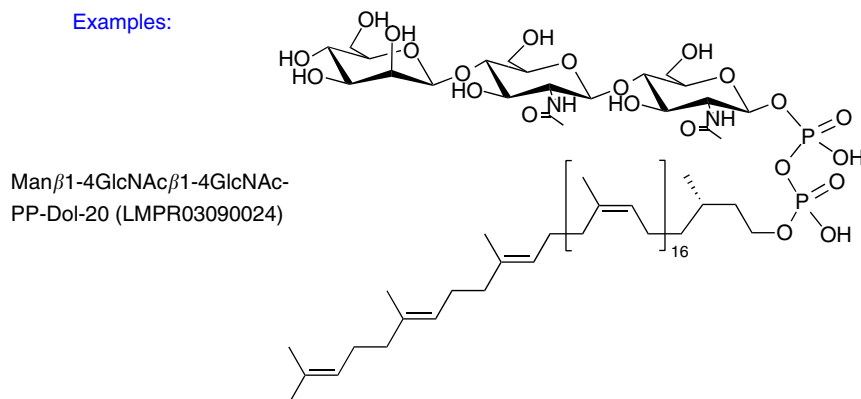
Sphingolipid biosynthesis *de novo* (Merrill, 2011; Yamaji and Hanada, 2014) begins with the condensation of serine and a

Glycosylated prenols

Prototype: Dolichol-{glycan, phospho-glycan or pyrophospho-glycan}



Examples:



Glc₃Man₉GlcNAc₂-PP-Dolichol

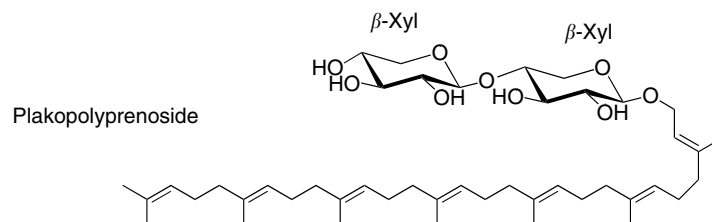
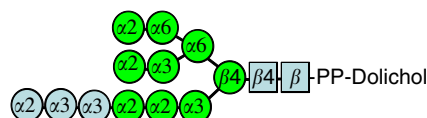


Figure 6 Representative structures of glycosylated prenols. A general prototype for these compounds is shown at the top, with the types of chain lengths for bacterial ($n=6-8$) versus plant and animal ($n=14-18$) shown. The examples include the types of carbohydrates found in the dolichol pyrophosphates used in glycoprotein biosynthesis, with the full oligosaccharide shown in shape/color symbols. The bottom structure demonstrated that other glycans are also found in some organisms.

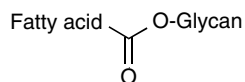
fatty acyl-CoA (myristoyl-CoA for d16-, palmitoyl-CoA for d18-, and stearoyl-CoA for d20-sphingoid bases) by SPT to produce 3-ketosphinganine with these chain lengths (Figure 9). Thus, this is the first step where the structure of the backbone is defined and is controlled by expression of SPT isoforms or accessory proteins that modify its fatty acyl-CoA selectivity.

The 3-keto-group is rapidly reduced and the product, sphinganine, is also at an important branchpoint (Figure 9) – determination of the nature of the fatty acyl group of the backbone. Sphinganine is N-acylated to dihydroceramides (DHCer) by ceramide synthases (CerS) that control the fatty acyl chain length as shown (e.g., CerS1 adds mainly stearic acid, C18:0, as shown in the highlighted portion of the diagram).

DHCers are also at a key branchpoint because they can be desaturated or hydroxylated to 4-hydroxyDHCer (also referred to as phytoceramides) as shown in this diagram as the lower octagon. Most mammalian tissues have mainly ceramides, but 4-hydroxy-ceramides are common in skin and epithelial tissues. The other option for DHCer (and for ceramides, as shown in this diagram) is for a headgroup to be added to make sphingomyelins (SM), ceramide phosphate (CerP), glucosylceramides (GlcCer), or galactosylceramides (GalCer) (Figures 9 and 10). Small amounts of ceramide phosphoethanolamines are also made by mammals, and these are the major phosphosphingolipid for some organisms (such as fruit flies). Some organisms, such as yeast, also produce ceramide phosphoinositols as their major phosphosphingolipid.

Saccharolipids

Prototype: Fatty acid linked directly to a sugar backbone



Example:

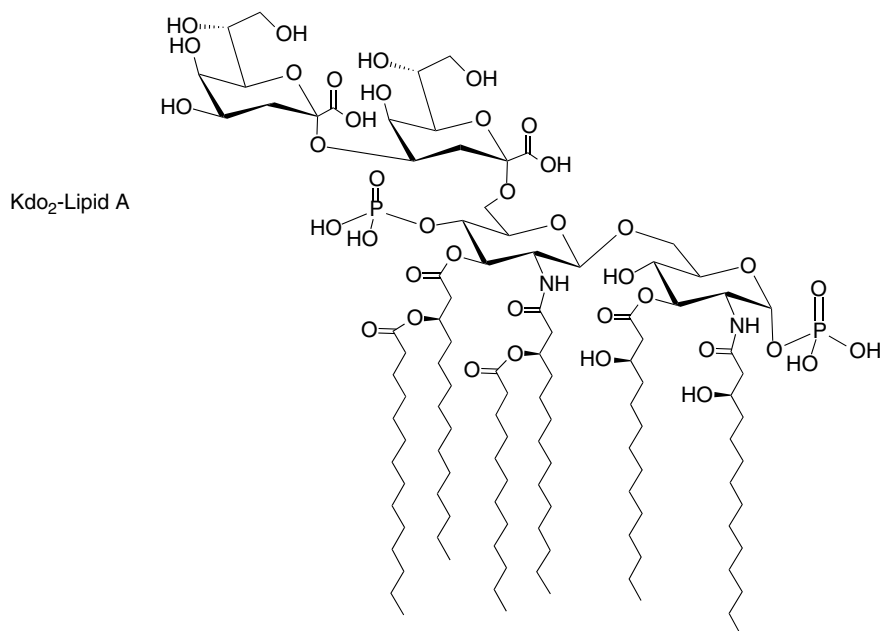


Figure 7 Representative structures of saccharolipids. A general prototype for these compounds is shown at the top, with an example of the Ldo2-Lipid A portion of the lipopolysaccharide of gram-negative bacteria.

Glycosylated polyketides

Examples:

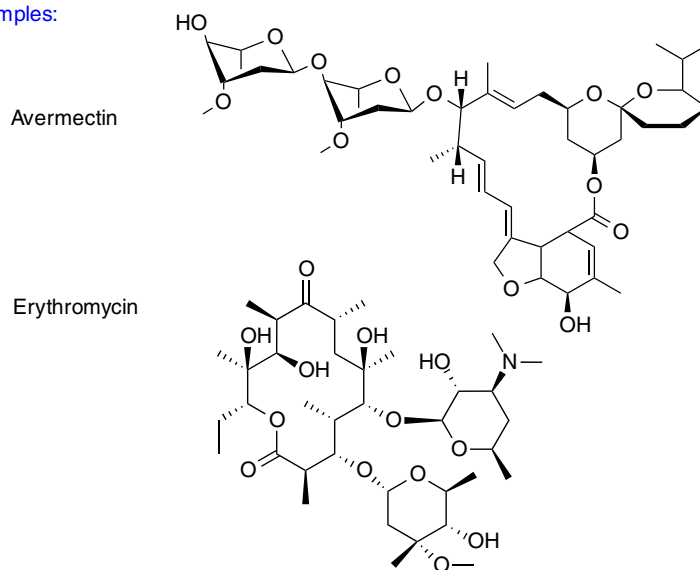


Figure 8 Representative structures of glycosylated polyketides. These compounds are too varied in structure to represent by a general prototype, but two examples are shown in avermectin and erythromycin.

Glycosphingolipid root structure biosynthetic pathways

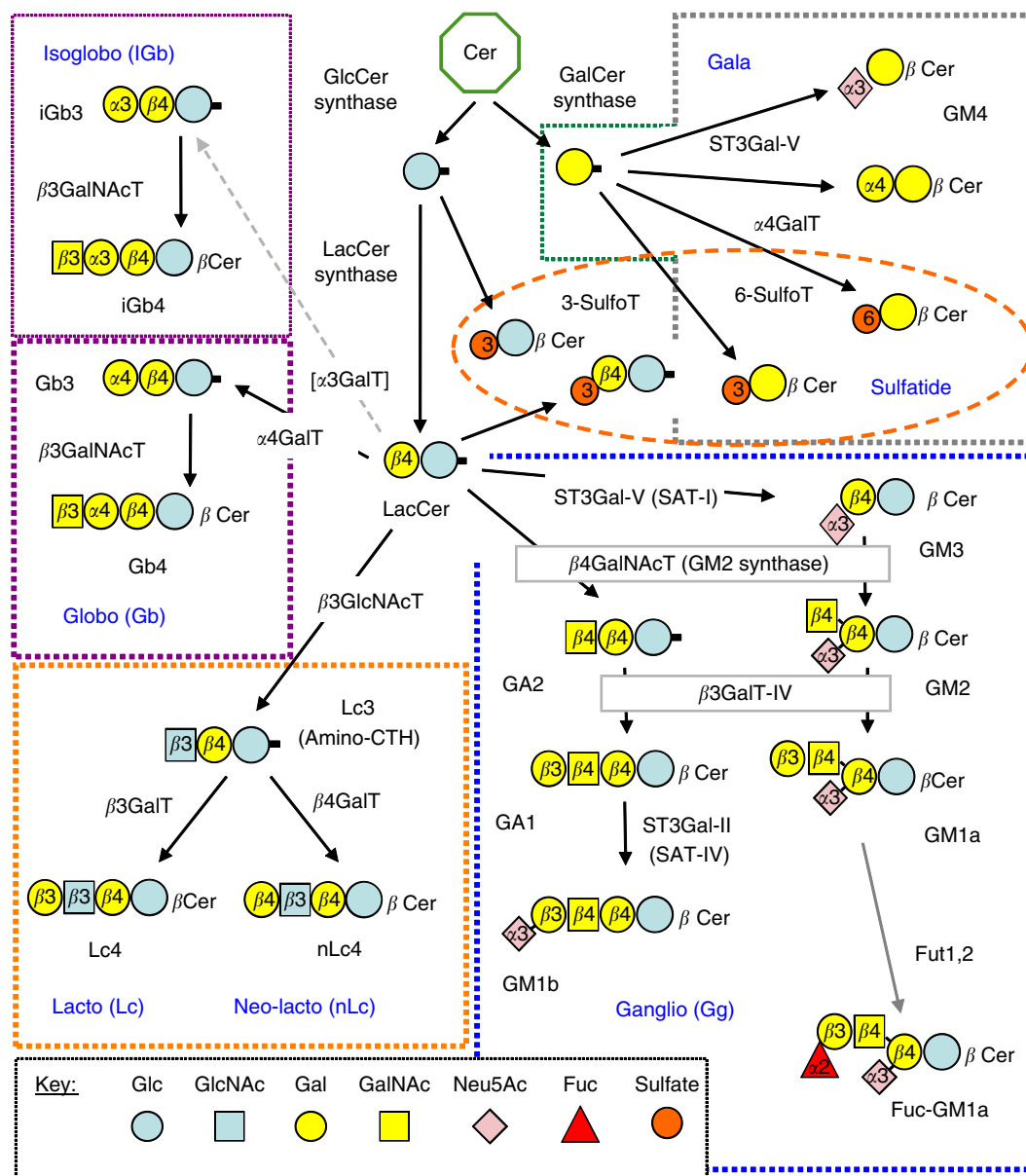


Figure 10 Glycosphingolipid root structure biosynthetic pathways. This diagram continues the metabolic pathway begun in [Figure 9](#) with the subsequent metabolic steps for the galactosylCer family and the root structure families downstream from lactosylCer. For more information about this pathway, see Merrill, A.H., Jr. 2011. Sphingolipid and glycosphingolipid metabolic pathways in the era of sphingolipidomics. *Chemical Reviews* 111, 6387–6422, from which it was modified.

[Figure 11](#) displays another important concept in thinking about glycosphingolipid biosynthesis: that a single glycosyltransferase might recognize multiple substrates and thereby produce many products ([Kolter et al., 2002](#)). This helps explain how cells can produce many more glycosphingolipids than there are genes that code for glycosyltransferases. Thus, as one examines the gene expression profile for a given system, the possible appearance of downstream metabolites might be manifested in side pathways rather than the one of original focus.

A useful source of information on glycosyltransferases is the series edited by N. Taniguchi *et al.* (eds.), *Handbook of*

Glycosyltransferases and Related Genes, DOI 10.1007/978-4-431-54240-7_33.

Analysis of Glycolipids by ‘Omic’ Technologies

Since no method is available to truly conduct a comprehensive glycolipidomic analysis, studies will actually encompass a subgroup of the glycolipidome that is defined by the investigator’s technical capabilities and/or goal for the analysis. This said, if an investigator is mainly interested in particular glycan

Ganglioside biosynthesis as a 'combinatorial' pathway

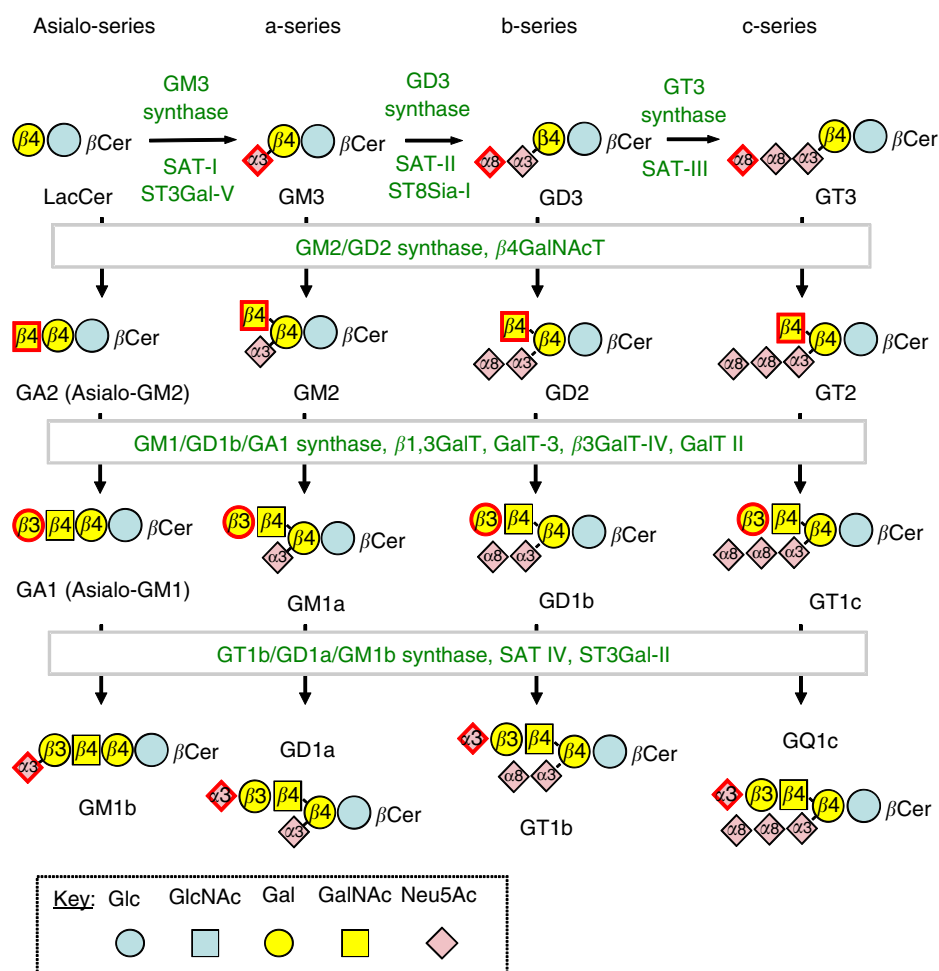


Figure 11 Ganglioside biosynthesis as a 'combinatorial' pathway. This diagram continues the metabolic pathway begun in [Figure 10](#), illustrating how glycosyltransferases can act on different substrates that share critical features but differ otherwise; therefore, multiple downstream metabolites can be made from a limited number of glycosyltransferases. For each glycosyltransferase, the added carbohydrate is indicated by a red border. For more information about this pathway, see Merrill, A.H., Jr. 2011. Sphingolipid and glycosphingolipid metabolic pathways in the era of sphingolipidomics. *Chemical Reviews* 111, 6387–6422, from which it was modified.

subspecies, there are methods to profile them using lectins or antibodies ([Cummins and Etzler, 2009](#)). Likewise, mass spectrometry has been used based to profile the glycans released by chemical or enzymatic cleavage followed by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry ([Parry et al., 2007](#)).

Intact glycolipids can be more difficult to analyze. Analysis of glycosphingolipids, as an example, usually requires some form of chromatography for the separation of isomeric species (as reviewed in [Haynes et al., 2009](#)). Nonetheless, in some respects, glycomics mass spectrometry is easier than other approaches for identification of new aglycones for glycolipids ([Wuhrer, 2013](#)).

Perspective on the Future of Glycolipid Research

Glycomics has been revolutionized over the past decade by the development of powerful tools, but glycolipid analysis

remains an especially daunting challenge because it must establish the structures, and quantity, of compounds that are sometimes present in trace amounts and comprised of two difficult-to-analyze components – glycans and lipids. One might argue that many of these compounds are present in relatively small amounts (for example, GlcCer is often only 5 to 10% of the abundance of sphingomyelin); however, cells still contain a lot of glycolipid molecules! For example, RAW264.7 macrophages have an average 53 million molecules of GlcCer/cell, and ca 1 million molecules of the less prevalent GalCer. Furthermore, upon activation of the cells, these numbers increase 2–4 fold ([Sims et al., 2010](#)). Such large increases in amount provocatively suggest that something important is happening. Changes in membrane dynamics? Interactions between glycolipids and membrane proteins?

The large numbers of exogenous glycolipids might also impact us in ways that are only beginning to be appreciated. The microflora that live on and in us produce many bioactive

glycolipids, including a potent immunomodulator, α -galactosylceramide, that was initially found in sponges but has recently been found to be produced by *Bacteroides fragilis* in the human gut microbiota (Wieland Brown *et al.*, 2013). There is undoubtedly much left to learn about glycolipids and how these compounds affect our lives.

See also: Molecular Principles, Components, Technology, and Concepts: Lipids: Lipidomics

References

- Allende, M.L., Proia, R.L., 2014. Simplifying complexity: Genetically rescuing glycosphingolipid synthesis pathways in mice to reveal function. *Glycoconjugate Journal* 31, 613–622.
- Balla, T., 2013. Phosphoinositides: Tiny lipids with giant impact on cell regulation. *Physiological Reviews* 93, 1019–1137.
- Banoub, J.H., El Anead, A., Cohen, A.M., Joly, N., 2010. Structural investigation of bacterial lipopolysaccharides by mass spectrometry and tandem mass spectrometry. *Mass Spectrometry Reviews* 29, 606–650.
- Campbell, M.P., Ranzinger, R., Luthe, T., *et al.*, 2014. Toolboxes for a standardised and systematic study of glycans. *BMC Bioinformatics* 15 (Suppl. 1), S9.
- Chigorno, V., Giannotta, C., Ottico, E., *et al.*, 2005. Sphingolipid uptake by cultured cells: Complex aggregates of cell sphingolipids with serum proteins and lipoproteins are rapidly catabolized. *Journal of Biological Chemistry* 280, 2668–2675.
- Costantino, V., Fattorusso, E., Mangoni, A., Di Rosa, M., Ianaro, A., 2000. Glycolipids from sponges. Part 8: 1 plakoprenol from the marine sponge plakortis simplex. An improved procedure for isolation of glycolipids as peracetyl derivatives. *Tetrahedron* 56, 1393–1395.
- Cummings, R.D., Etzler, M.E., 2009. Antibodies and lectins in glycan analysis. In: Varki, A., Cummings, R.D., Esko, J.D. *et al.* (Eds.), *Essentials of Glycobiology*, second ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Ebada, S.S., Lin, W., Proksch, P., 2010. Bioactive sesterterpenes and triterpenes from marine sponges: Occurrence and pharmacological significance. *Marine Drugs* 8, 313–346.
- Fahy, E., Subramaniam, S., Brown, H.A., *et al.*, 2005. A comprehensive classification system for lipids. *Journal of Lipid Research* 46, 839–861.
- Furukawa, K., Tsuchida, A., Furukawa, K., 2007. Biosynthesis of glycolipids. In: Kamerling, J.P. (Ed.), *Comprehensive Glycoscience. From Chemistry to Systems Biology*. Oxford: Elsevier, pp. 105–114.
- Garrett, T.A., Guan, Z., Rietz, C.R., 2007. Analysis of ubiquinones, dolichols, and dolichol diphosphate-oligosaccharides by liquid chromatography-electrospray ionization-mass spectrometry. *Methods in Enzymology* 432, 117–143.
- Grille, S., Zaslawski, A., Thiele, S., Plat, J., Warnecke, D., 2010. The functions of sterol glycosides come to those who wait: Recent advances in plants, fungi, bacteria and animals. *Progress in Lipid Research* 49, 262–288.
- Guan, Z., Eichler, J., 2011. Liquid chromatography/tandem mass spectrometry of dolichols and polyprenols, lipid sugar carriers across evolution. *Biochimica et Biophysica Acta* 1811, 800–806.
- Hannun, Y.A., Luberto, C., Argraves, K.M., 2001. Enzymes of sphingolipid metabolism: From modular to integrative signaling. *Biochemistry* 40, 4893–4903.
- Hansen, L., Lind-Thomsen, A., Joshi, H.J., *et al.*, 2014. A glycogene mutation map for discovery of diseases of glycosylation. *Glycobiology* 25, 211–224.
- Haynes, C.A., Allegood, J.C., Park, H., Sullards, M.C., 2009. Sphingolipidomics: Methods for the comprehensive analysis of sphingolipids. *Journal of Chromatography. B, Analytical Technologies in the Biomedical and Life Sciences* 877, 2696–2708.
- Hertweck, C., 2009. The biosynthetic logic of polyketide diversity. *Angewandte Chemie, International Edition in English* 48, 4688–4716.
- Holz, G., Dormann, P., 2007. Structure and function of glycosylglycerolipids in plants and bacteria. *Progress in Lipid Research* 46, 225–243.
- Honke, K., 2013. Biosynthesis and biological function of sulfoglycolipids. *Proceedings of the Japan Academy. Series B, Physical and Biological Sciences* 89, 129–138.
- Ishibashi, Y., Kohyama-Koganeya, A., Hirabayashi, Y., 2013. New insights on glucosylated lipids: Metabolism and functions. *Biochimica et Biophysica Acta* 1831, 1475–1485.
- Ivanchina, N.V., Kicha, A.A., Stonik, V.A., 2011. Steroid glycosides from marine organisms. *Steroids* 76, 425–454.
- Kersten, R.D., Ziemert, N., Gonzalez, D.J., *et al.*, 2013. Glycogenomics as a mass spectrometry-guided genome-mining method for microbial glycosylated molecules. *Proceedings of the National Academy of Sciences of the United States of America* 110, E4407–E4416.
- Kolter, T., Proia, R.L., Sandhoff, K., 2002. Combinatorial ganglioside biosynthesis. *Journal of Biological Chemistry* 277, 25859–25862.
- Laine, R.A., 1994. A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: The isomer barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* 4, 759–767.
- Merrill Jr., A.H., 2011. Sphingolipid and glycosphingolipid metabolic pathways in the era of sphingolipidomics. *Chemical Reviews* 111, 6387–6422.
- Michell, R.H., 2013. Inositol lipids: from an archaeal origin to phosphatidylinositol 3,5-bisphosphate faults in human disease. *FEBS Journal* 280, 6281–6294.
- Momin, A.A., Park, H., Portz, B.J., *et al.*, 2011. A method for visualization of “omic” datasets for sphingolipid metabolism to predict potentially interesting differences. *Journal of Lipid Research* 52, 1073–1083.
- Morii, H., Ogawa, M., Fukuda, K., Taniguchi, H., 2014. Ubiquitous distribution of phosphatidylinositol phosphate synthase and archaetidylinositol phosphate synthase in Bacteria and Archaea, which contain inositol phospholipid. *Biochemical and Biophysical Research Communications* 443, 86–90.
- Nairn, A.V., York, W.S., Harris, K., *et al.*, 2008. Regulation of glycan structures in animal tissues: Transcript profiling of glycan-related genes. *Journal of Biological Chemistry* 283, 17298–17313.
- Parry, S., Ledger, V., Tissot, B., *et al.*, 2007. Integrated mass spectrometric strategy for characterizing the glycans from glycosphingolipids and glycoproteins: Direct identification of sialyl Le(x) in mice. *Glycobiology* 17, 646–654.
- Paulick, M.G., Bertozzi, C.R., 2008. The glycosylphosphatidylinositol anchor: A complex membrane-anchoring structure for proteins. *Biochemistry* 47, 6991–7000.
- Penno, A., Reilly, M.M., Houlden, H., *et al.*, 2010. Hereditary sensory neuropathy type 1 is caused by the accumulation of two neurotoxic sphingolipids. *Journal of Biological Chemistry* 285, 11178–11187.
- Pruett, S.T., Bushnev, A., Hagedorn, K., *et al.*, 2008. Biodiversity of sphingoid bases (“sphingosines”) and related amino alcohols. *Journal of Lipid Research* 49, 1621–1639.
- Quehenberger, O., Armando, A.M., Brown, A.H., *et al.*, 2010. Lipidomics reveals a remarkable diversity of lipids in human plasma. *Journal of Lipid Research* 51, 3299–3305.
- Raetz, C.R., Guan, Z., Ingram, B.O., *et al.*, 2009. Discovery of new biosynthetic pathways: The lipid A story. *Journal of Lipid Research* 50 (Suppl.), S103–S108.
- Rhee, S.H., 2014. Lipopolysaccharide: Basic biochemistry, intracellular signaling, and physiological impacts in the gut. *Intestinal Research* 12, 90–95.
- Rombouts, Y., Alibaud, L., Carrere-Kremer, S., *et al.*, 2011. Fatty acyl chains of *Mycobacterium marinum* lipooligosaccharides: Structure, localization and acylation by PapA4 (MMAR_2343) protein. *Journal of Biological Chemistry* 286, 33678–33688.
- Schnaar, R.L., Gerardy-Schahn, R., Hildebrandt, H., 2014. Sialic acids in the brain: Gangliosides and polysialic acid in nervous system development, stability, disease, and regeneration. *Physiological Reviews* 94, 461–518.
- Shaner, R.L., Allegood, J.C., Park, H., *et al.*, 2009. Quantitative analysis of sphingolipids for lipidomics using triple quadrupole and quadrupole linear ion trap mass spectrometers. *Journal of Lipid Research* 50, 1692–1707.
- Shimamura, M., 2012. Immunological functions of sterol glycosides. *Archivum Immunologiae et Therapiae Experimentalis (Warsz)* 60, 351–359.
- Sims, K., Haynes, C.A., Kelly, S., *et al.*, 2010. Kdo2-lipid A, a TLR4-specific agonist, induces de novo sphingolipid biosynthesis in RAW264.7 macrophages, which is essential for induction of autophagy. *Journal of Biological Chemistry* 285, 38568–38579.
- Suzuki, A., 2002. *Map 3. Biosynthetic Pathways of Glycosphingolipids*. Tokyo: Springer-Verlag.
- Tettamanti, G., Bassi, R., Viani, P., Riboni, L., 2003. Salvage pathways in glycosphingolipid metabolism. *Biochimie* 85, 423–437.
- Varki, A., Cummings, R., Esko, J., *et al.* (Eds.), 2009. *Essentials of Glycobiology*, second ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Welfi, M., 2013. Regulation of dolichol-linked glycosylation. *Glycoconjugate Journal* 30, 51–56.
- Wieland Brown, L.C., Penaranda, C., Kashyap, P.C., *et al.*, 2013. Production of alpha-galactosylceramide by a prominent member of the human gut microbiota. *PLoS Biology* 11, e1001610.
- Wuhrer, M., 2013. Glycomics using mass spectrometry. *Glycoconjugate Journal* 30, 11–22.

- Yamaji, T., Hanada, K., 2014. Sphingolipid metabolism and interorganelle transport: Localization of sphingolipid enzymes and lipid transfer proteins. *Traffic* 16, 101–122.
- Yu, R.K., Yanagisawa, M., Ariga, T., 2007. Glycosphingolipid structures. In: Kamerling, J.P. (Ed.), *Comprehensive Glycoscience. From Chemistry to Systems Biology*. Oxford: Elsevier.
- Zhang, J., Li, C., Yu, G., Guan, H., 2014. Total synthesis and structure-activity relationship of glycosylcerolipids from marine organisms. *Marine Drugs* 12, 3634–3659.
- Zitomer, N.C., Mitchell, T., Voss, K.A., *et al.*, 2009. Ceramide synthase inhibition by fumonisin B1 causes accumulation of 1-deoxysphinganine: A novel category of bioactive 1-deoxysphingoid bases and 1-deoxydihydroceramides biosynthesized by mammalian cell lines and animals. *Journal of Biological Chemistry* 284, 4786–4795.
- <http://www.functionalglycomics.org/static/consortium/Nomenclature.shtml>
Consortium for Functional Glycomics.
- <http://www.ncbi.nlm.nih.gov/books/NBK1908/>
Essentials of Glycobiology.
- <http://www.glycoforum.gr.jp/>
GlycoForum.
- <http://www.chem.qmul.ac.uk/iupac/misc/glylp.html>
(IUPAC-IUB) Joint Commission on Biochemical Nomenclature.
- <http://www.lipidmaps.org>
LIPID MAPS.
- www.sphingomap.org
SPHINGOMAP.

Relevant Websites

<http://lipidlibrary.aocs.org/index.html>
American Oil Chemistry Society (AOCS) Lipid Library.

Lipid Signaling

B Tu-Sekine and DM Raben, Johns Hopkins School of Medicine, Baltimore, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Anionic lipid A lipid with an overall negative charge at physiological pH; also called an acidic lipid. PtdOH, PIPs, and PtdSer are anionic lipids.

Bioactive lipid A cellular lipid that stimulates or propagates biological signaling pathway(s).

Eicosanoid A signaling molecule produced through oxidation of a 20 carbon fatty acid such as arachidonic acid or eicosapentanoic acid. Leukotrienes, prostaglandins, prostacyclins, and thromboxanes are all eicosanoids.

Endocannabinoid The arachidonic acid-derived neuromodulatory lipids anandamide (AEA) and 2-arachidonoylglycerol (2-AG) present in neurons.

G-protein coupled receptor Seven pass transmembrane receptors that constitute a large superfamily of receptors in eukaryotes. Activation of a GPCR by an agonist stimulates intracellular signaling cascades.

Lipidome The complete set of lipids in a cell or organism.

Lipid raft Membrane microdomains comprised primarily of cholesterol and sphingomyelin that are less fluid than surrounding membrane and serve as organizing centers for protein signaling complexes.

Metabolome The complete set of metabolites in a cell or organism.

Signalsomes A complex of signaling proteins, typically present at the plasma membrane, but known to occur at various membranes throughout the cell to a lesser extent.

Introduction

Lipid signaling refers to the lipid-dependent activation or propagation of a signal within a cell. For example, priming of synaptic vesicles for release from a synapse, splicing of messenger RNAs, and T-cell receptor responses are all dependent on specific lipids to activate, recruit, or regulate proteins involved in the process. Lipids also form the membrane barriers that separate organelles within a cell, and one cell from another. In the form of triglycerides, lipids store energy analogous to the storage of glucose as glycogen, but in a more efficient manner, yielding 2385 kcal mol⁻¹ (palmitate) versus 670 kcal mol⁻¹ (glucose). And, in a method unique to lipids, there is significant interconversion of one lipid type to another by headgroup exchange or modification and by fatty acid exchange or remodeling, leading to a highly dynamic system.

The majority of research identifying and analyzing lipid second messengers have been conducted in cultured cells. This can occasionally result in confusion due to variations in lipid and protein compositions in different cell types. Where possible, information is provided in a general context and signaling pathways that are restricted to a particular cell type are noted.

This article provides an overview of the primary classes of mammalian signaling lipids. Particular emphasis is placed on the regulation of signaling lipids, and the mechanisms by which they participate in signal transduction. Because signaling lipids participate in most signaling pathways and processes, we have not attempted to review their physiological functions. Instead, we recommend several resources at the end of this article.

Lipid Classification and Nomenclature

The Lipid Maps Consortium has developed a classification system for lipids that has become internationally accepted.

Under this system lipids are divided into eight categories, six of which are found in mammals and listed in Table 1. All six categories contain members that play dominant roles in signaling pathways, although the prenol lipids and sterol lipids are not discussed in this review. Prenol lipids are covalently attached to proteins, and exert their effects on signaling by altering the localization of the protein, which is outside the scope of this article. Sterols include a wide range of lipid molecules that have many functions, and contribute to signaling through both structural effects and intracellular receptor binding (see Maxfield and van Meer, 2010; Hulce et al., 2013).

Each lipid category is divided into classes that are distinguished by headgroup and lipid backbone structure. These classes are further separated into subclasses based on additional chemical criteria such as headgroup type, lipid tail organization, and tail linkage (e.g., ester vs. ether linkages). The final organization is similar to the taxonomic rankings of phylum, species, and class hierarchy of the biological Tree of Life, and provides a familiar framework for non-lipid

Table 1 Lipid classification system from the Lipid MAPS Consortium

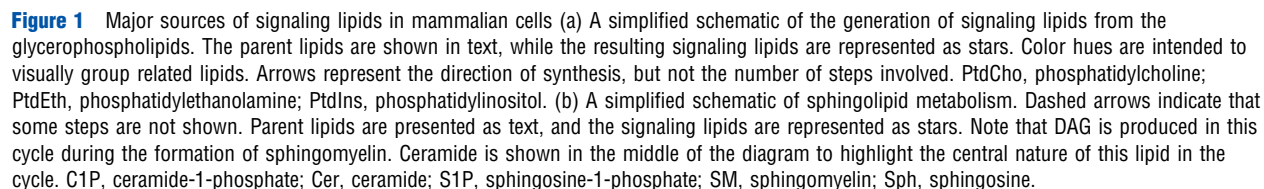
Category	Number of classes	Members discussed in this article
Fatty acyls	14	Arachidonic acid and Eicosinoids
Glycerolipids	6	DAG
Glycerophospholipids	21	PtdOH and phosphoinositides
Sphingolipids	10	S1P
Sterol lipids	6	n/a
Prenol lipids	5	n/a

Note: The above table presents the six categories of lipids present in mammalian cells, as well as the number of classes within each category. For clarity, the subclasses are not listed. The members of each category that are discussed in the text are noted.

Exceptions to this include lysolipids such as lysophosphatidic acid (LPA) and sphingosine-1-phosphate (S1P), and fatty acids such as arachidonic acid and the eicosanoids. Due to their charged state and reduced hydrophobic nature fatty acids are soluble in water to millimolar concentrations and can readily leave membranes.

While there are thousands of species of signaling lipids, the vast majority are the products of two main groups: the glycerophospholipids and the sphingolipids. Additionally, there are two primary metabolic networks that produce the majority of the signaling lipids of the cell: the PI cycle and the sphingolipid cycle. A simplified representation of the metabolic interconversions of the glycerophospholipids to form signaling lipids is presented in [Figure 1\(a\)](#). Because the PI cycle is a central source of signaling lipids, PtdIns is separated from the other phospholipids in the diagram. However, PtdCho is also an important source of signaling PtdOH, and PtdEth has been implicated as a major source of arachidonic acid (the precursor to the major signaling eicosanoids) in some tissues. The arrows represent the typical direction of synthesis, and double arrows are placed where signaling lipids are known to frequently interconvert. The signaling lipids themselves are depicted as stars. It is important to note that while PtdOH is the metabolic precursor for *de novo* synthesis of lipids at the endoplasmic reticulum (ER), this pool of PtdOH contains a distinctly more saturated fatty acid profile and is not utilized in signaling pathways.

The canonical signaling lipids include the phosphatidylinositols (PIs, primarily phosphatidylinositol-4-phosphate, phosphatidylinositol-(4,5)-bisphosphate and phosphatidylinositol (3,4,5)-trisphosphate), diacylglycerol, phosphatidic acid, sphingosine-1-phosphate, ceramide-1-phosphate, and eicosanoids. With the exception of diacylglycerol and the sterols, the major signaling lipids carry one or more negative charges that provide binding specificity and stabilize protein-lipid interaction. These lipids are confined to membranes due to the strongly hydrophobic nature of the fatty acids attached to the lipid backbone, which are below the soluble plane of the membrane. Lipid-binding proteins (LBPs) typically interact with the lipid headgroup and some of the lipid backbone, but rarely have direct interactions with the fatty acid chains. The bound lipid does not leave the membrane, but rather acts as a temporary tether to retain an effector protein at the interface.



a sphingoid base as the backbone rather than a glycerol moiety, and are metabolically and chemically distinct from the glycerophospholipids. An abbreviated representation of the sphingolipid pathway is shown in [Figure 1\(b\)](#). Note that ceramide is the central lipid in the sphingolipid cycle and the source from which the other sphingolipids are derived. There is significant interconversion between ceramide and the other pathway constituents (represented by double arrows), however C1P and S1P are the most well-studied signaling molecules in this group. While there is a wealth of experimental data showing that ceramide and sphingomyelin contribute to the formation of lipid domains (aka lipid rafts) which promote the clustering of signaling proteins into signalsomes, signaling through rafts is not discussed in this article because the lipid effect is general and does not involve specific partner interactions. Note that diacylglycerol (DAG) is produced in the sphingolipid cycle during the formation of sphingomyelin. It is important to remember that these pools of DAG are often spatially distinct from DAG derived from the PI cycle, and contain fatty acid species that are more saturated than those derived from PI(4,5)P₂, which may influence their signaling properties.

Signaling Lipids are Spatially Restricted and Scarce

Tight regulation of signaling lipids is reflected in their low abundance, even in tissues that specialize in processes that depend on bioactive lipids. For example, phospholipids represent approximately 65% of total cellular lipids, but the combined pool of phospholipid-derived signaling lipids may represent less than 1% of that value. However, quantitative methods such as mass spectroscopy have revealed general patterns in the relative abundance of many bioactive lipids. For example, DAG and PtdOH are approximately an order of magnitude more abundant than the phosphatidylinositols, with PI(4,5)P₂ and PI(4)P contributing approximately 61% and 37% or the PIPs, respectively ([Guillou et al., 2007](#)). Overall, it has become clear that the location and abundance of signaling lipids is carefully regulated to maintain coherent signaling networks ([Figure 2](#)).

While the temporal and spatial resolution of signaling lipids is still poor due to the high mobility, rapid metabolism, and low abundance of bioactive lipids within a cell, the combined data from fluorescent lipid-binding probes and lipidomic studies of organelles indicate that the distribution of signaling lipids varies throughout the cell. These observed distributions typically correspond with the functional requirements of the cell. For example, trafficking pathways are highly dependent on PtdOH for fusion and budding of vesicles from the Golgi, and PtdOH can be seen concentrated at the Golgi apparatus. Interestingly, individual phosphatidylinositol subclasses are also enriched in various organelles, and control the recruitment of stage-specific proteins to different compartments. For example, PI(4,5)P₂ is highly concentrated in the cytosolic leaflet of the plasma membrane where it participates in the production of second messengers and is instrumental in endocytosis, PI(3)P is associated with early endosomes where it recruits adapter proteins, and PI(4)P is abundant at the Golgi apparatus where it is important for

trafficking ([Di Paolo and De Camilli, 2006](#); [Mayinger, 2012](#)). Cellular partitioning is also evident for other lipids that serve structural roles, which is beyond the scope of this article but is outlined in several excellent reviews ([Bader et al., 2011](#); [Nicolson, 2014](#); [Coskun and Simons, 2011](#)).

Signaling lipids have recently been examined at the mitochondria, leading to several unexpected findings. Approximately 5% of the cellular PI(4,5)P₂ is found at the mitochondrial outer membrane. While this seems like a minor amount, when normalized to the membrane surface area, the resulting density is approximately 25% of that found at the plasma membrane ([Nemoto and De Camilli, 1999](#)). This pool of PIP₂ has been shown to recruit a form of synaptojanin 2, a phosphatidylinositol phosphatase important for uncoating vesicles, to mitochondria, and appears to be important for the intracellular localization of mitochondria. The phosphatidylinositol-4-phosphate 5-kinases which generate PI(4,5)P₂ and a form of PLD that produces the substrate lipid for PIP₂ production have also been identified at the mitochondrial membrane, indicating a separate PIP₂ signaling network at this organelle. Subsequent studies have shown this network regulates fusion, fission, and localization of mitochondria in cells ([Huang and Frohman, 2009](#)).

Another important finding resulting from combined research efforts is the identification of independent lipid signaling pathways within the nuclear matrix. This finding was controversial for many years due to lack of evidence for any membranous structures in the nuclear matrix. However, independent laboratories identified all the necessary components for a nuclear PI cycle, as well as nuclear lipid effectors such as the Star-PAP polyA polymerase ([Li et al., 2013](#)), PIP₂-activated splicing factors, and nuclear DAG-binding proteins such as PKC α . PI(4,5)P₂ has been detected *in situ* using antibodies ([Tabellini et al., 2003](#)), and lipid microdomains ([Cascianelli et al., 2008](#)) and lipid droplets ([Layerenza et al., 2013](#)) have been detected in the nucleus. The combined evidence has resulted in the general acceptance of a separate lipid signaling web in the nucleus of cells.

The Complexity of Lipid Signaling

Bioactive lipids participate in a vast number of signaling pathways from proliferative to apoptotic, throughout development and aging, in metabolic, catabolic, trafficking, and regulatory processes. This diversity arises from the wide array of interacting protein partners, the integration of multiple, simultaneously active signaling pathways, and the substantial diversity inherent within the lipid molecules themselves. The Lipid MAPS Consortium currently lists over 28 000 individual lipids in the categories covered in this article, encompassing variations in lipid class, headgroup, fatty acid linkage, and fatty acid composition such as elements of unsaturation and a variety of lipid chain modifications. The purpose(s) of the extraordinary diversity within the lipid complement remains enigmatic. Not all will be relevant to signaling, but there are likely to be additional as yet unidentified patterns responsible for variations in signaling responses.

Phosphatidylinositol lipids provide a significant amount of diversity on their own. Variations in phosphorylation patterns

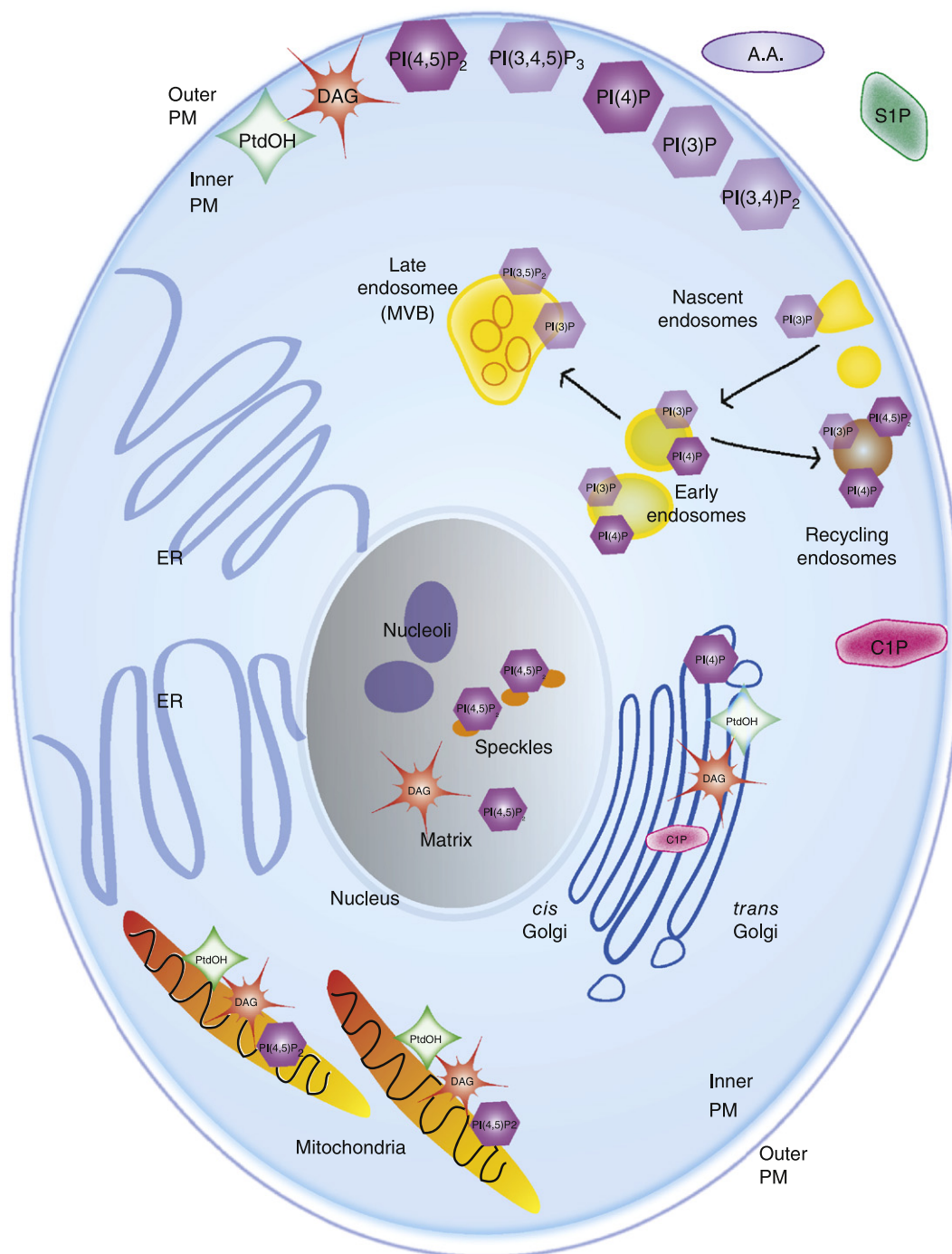


Figure 2 Spatial organization of signaling lipids. The subcellular locations of the major signaling lipids are represented in a stylized schematic of a mammalian cell. Note that all signaling lipids are present at the inner plasma membrane, but are not present at the outer plasma membrane. The most abundant phosphatidylinositols (by mass) have darker shading than the less abundant members. Remodeling of PIPs results in enriched populations that are reliable markers for the various vesicle subtypes as they proceed through the endocytic cycle or split off into recycling endosomes (indicated by arrows). Although DAG and PtdOH are abundant at the ER (the main site of phospholipid synthesis), these pools do not appear to be involved in signaling reactions and therefore are not represented in the diagram. DAG, diacylglycerol; ER, endoplasmic reticulum; MVB, multivesicular body; PtdOH, phosphatidic acid.

on the inositol ring produce seven species of phosphatidylinositols, which are recognized by a variety of lipid-binding domains in proteins. PIP₂ represents a special class of signaling lipid since it is also a precursor for diacylglycerol and phosphatidic acid for

second messenger signaling, and for the synthesis and release of endocannabinoids from arachidonic acid.

At a general level, lipid signaling can be separated into two basic categories: first messenger signaling, in which a bioactive

lipid binds to and activates a receptor at the cell surface, and second messenger signaling, in which a bioactive lipid produced in response to receptor stimulation triggers downstream signaling cascade(s).

Extracellular Signaling Lipids Interact with Receptors

A number of signaling lipids are released from cells and initiate signaling pathways by stimulating extracellular receptors. These lipid ligands include sphingosine-1-phosphate, lysophospholipids (primarily platelet activating factor and lysophosphatidic acid), and members of the arachidonate-derived eicosanoid family (prostaglandin, leukotriene, thromboxane, and the neuronal cannabinoids). These relatively water-soluble lipids are produced by enzyme-mediated hydrolysis of membrane-associated parent lipids and released into the extracellular space where they associate with cognate receptors on the same or nearby cells. With few exceptions, extracellular lipid receptors belong to the G-protein coupled receptor (GPCR) superfamily, and downstream events are defined by the individual receptor dynamics. One notable exception is the apparent stimulation of inflammatory toll-like receptors (TLRs) by long chain fatty acids such as palmitate, suggesting a potential contribution of free fatty acid levels to symptoms of type 2 diabetes (Mamedova *et al.*, 2013).

Intracellular Signaling Lipids Interact with Proteins

Second messenger lipids are produced as a direct or indirect consequence of the stimulation of a vast array of extracellular receptors. The archetypal example of lipid signaling is the PI-PLC-mediated cleavage of PI(4,5)P₂ to produce IP₃ (which increases intracellular calcium) and the bioactive lipid DAG – a strong activator of PKC (and other) enzymes. The most

common mechanism of lipid-induced protein activation is through interaction of the lipid headgroup with a structural motif on the effector protein often referred to as the lipid-binding domain (LBD). The lipid-binding protein is also referred to as an effector protein because it propagates the signal initiated at the membrane receptor.

LBD – Lipid interactions typically involve the lipid headgroup, backbone, and one or more of the fatty acid moieties, but can also include some of the fatty acid chain itself. Fatty acid binding may indeed be a driving factor in ceramide and C1P associations with their protein partners. However, it is more common that the fatty acid physiochemistry facilitates binding indirectly by altering the lipid headgroup positioning within the membrane, or by promoting the formation of lipid domains that regionally concentrate protein partners and/or alter membrane properties (Lipowsky, 2002; Bigay and Antonny, 2012).

Proteomics studies have identified thousands of LBD-containing proteins in organisms from bacteria through primates. By rough calculation, approximately 6–8% (Based on the assumption that there are 20 000 protein-coding genes in the human genome, and using reported numbers of lipid-binding proteins published by the Lipid MAPs Consortium or by manual summation of LBDs extracted from the EBI InterPro database.) of the protein-coding genes in the human genome contain one or more LBDs, and the majority of these have reported signaling functions (Hunter *et al.*, 2012). This is likely to be an underestimate based on the observation that some LBDs, such as those that bind PtdOH, ceramide, and C1P, do not contain a readily identifiable lipid-binding motif, but may rely on more general biophysical characteristic(s) of the folded protein for lipid recognition. LBDs identified in proteins using InterPro are shown in Table 2, along with the number of human proteins predicted to contain each one. Note that some proteins contain more than one type of LBD and may be

Table 2 Lipid-binding domains (LBDs) in the human genome

Bioactive lipid	LBD	Number of human proteins containing LBD	InterPro designation
Arachadonic acid-based leukotriene	Leukotriene B4 receptor	8	IPR003981
Arachadonic acid-based thromboxane, prostaglandin	Prostanoid receptors	33	IPR008365
DAG	C1	279	IPR002219 (PKC-like)
PI(3)P	FYVE	78	IPR000306
PI(3,4)P ₂ , PI(4,5)P ₂ , PI(3,4,5)P ₃ (possibly PI(3)P and PI(4)P)	PH	1020	IPR001849
PI(4,5)P ₂ , PI(3,4,5)P ₃	ENTH	55	IPR013809
PI(4,5)P ₂	ANTH	26	IPR011417
PIPs (various)	PX	53	IPR000270
Anionic lipids (PtdSer, PI(4,5)P ₂ PtdOH) in membranes of high curvature	BAR	71	IPR004148
Anionic lipids (PtdSer, PI(4,5)P ₂ , calcium-dependent binding	C2	48	IPR002420 (PI3KC-like)
PtdOH	Unidentified proposed polybasic motifs in proteins	47 Unknown	IPR014020 (PKC- and tensin-like) n/a
Steroids	Nuclear hormone receptor	261	IPR001723
S1P	EDG receptors	8	IPR004061

Note: The designations of the common LBDs are listed, along with the number of proteins that contain the domain as noted in the EBI InterPro database. Proteins may contain more than one lipid-binding domain and therefore may be represented in multiple categories.

represented multiple times in this table. The abundance of phosphatidylinositol binding proteins is conspicuous, underscoring the dominance of the phosphatidylinositols as second messenger lipid signals. Indeed, the PH domain category can be subdivided based on empirically determined specificities for various phosphoinositide headgroups. For example, the PH domain in Akt binds PI(3,4)P₂ and PI(3,4,5)P₃ (see Rong *et al.*, 2001), while the PH domain of phospholipase C δ binds PI(4,5)P₂ (Garcia *et al.*, 1995).

However, this specificity is not always evident in the primary sequence of a protein, which limits *in silico* detection of lipid-binding domains. This has resulted in a stumbling block in the characterization of lipid-signaling networks since empirical methods for detecting lipid-binding domains are often laborious and expensive. In fact, several lipids appear to interact with proteins primarily through nonspecific electrostatic effects. This is the suggested mechanism for signaling through PtdOH and C1P, which are presumed to bind to proteins via a positively charged patch on the protein's three-dimensional structure.

Physiologic Roles of Signaling Lipids

Bioactive lipids are involved in signaling in every cell of every organism – it is difficult to imagine a process or pathway that does not use them. They are particularly prominent signaling molecules in the immune system where various bioactive lipids are required for proper functioning of both the innate and adaptive immune system, in the vesicle traffic required to deliver many proteins to the appropriate organelles and to maintain structure and order within the cell. In addition, metabolic pathways are highly dependent on lipid signals, particularly PIP₃, as are pathways that are dependent on endocytosis and exocytosis, a basic function of all cells, and a specialized function of endocrine cells and neurons.

Signaling Lipids and Disease

Despite their ubiquitous presence, it is rare that a signaling lipid is diagnosed as a causative agent in disease. Rather the protein(s) involved in the regulation of or response to the lipid are regarded as the causative factor. This is due in part to the fact that signaling lipids are much less plentiful and more quickly metabolized than proteins (and therefore much more difficult to detect), and in part to the technological difficulties in analyzing the very hydrophobic lipid molecules. Fortunately, breakthroughs in mass spectroscopic techniques now permit the reliable detection of many lipid species, resulting in the establishment of two relatively new fields: metabolomics and lipidomics. While lipidomic studies directly examine the lipid complements of cells, metabolomics provides complementary information on small molecule metabolites, acting as a reporter of the state of lipid metabolism.

Studies of lipid populations in disease states are providing new insights into lipid function based on lipid–protein mapping efforts. Some of the more prominent findings include roles of lipids in cancer, diabetes, and lipid storage disease (Shevchenko and Simons, 2010), the identification of

endocannabinoids as feedback regulators of synaptic transmission (D'Addario *et al.*, 2013), and the role of saturated free fatty acids as instigators of systemic inflammation in metabolic syndrome (Wymann and Schneider, 2008). In addition, comparative studies have revealed lipid dysregulation in neuronal pathologies including Alzheimer's disease, Parkinson's disease, and bipolar disorder (Touboul and Gaudin, 2014; Rittiner *et al.*, 2014).

Recent Advances in Signaling Lipid Technologies

The study of lipids as signaling entities began in the 1950s with the discovery of the 'PI effect' (Hokin, 1985). Recent advances in mass spectroscopy have made this powerful tool accessible for lipid analysis, ushering in the lipidomics area. Lipidomics is the study of the lipid population, or lipidome, of a tissue, cell or organelle, including the identity of the fatty acid species associated with each lipid. This type of analysis provides detailed information about the changes in cellular lipids resulting from an event of interest, such as the stimulation of a cell by an agonist, or the invasion of a cell by a virus. By following changes in lipid profiles, particularly those of signaling lipids, important information can be gained about the signaling networks that are engaged. This is exemplified in the recent construction of a protein–lipid interactome in the yeast *Saccharomyces cerevisiae* (Gallego *et al.*, 2010).

Additional tools that have become commonplace are the use of fluorescently labeled lipids for evaluating macroscopic lipid organization such as lipid domains and lipid rafts, as well as lipid biosensors, which are modified lipid-binding domains containing an optogenetic tag (often GFP) to allow *in vivo* imaging of specific signaling lipids, particularly the phosphoinositides and diacylglycerol, at the cellular and subcellular level (Gao and Zheng, 2013; Smith *et al.*, 2009; Takatori *et al.*, 2014). These tools provide information on the relative strength, location and duration of lipid signals in response to various stimuli.

As advances in the detection and interpretation of the signaling roles in bioactive lipids progress, so, too, will our understanding of their roles in signaling and disease.

See also: Molecular Principles, Components, Technology, and Concepts: Lipids: Lipidomics; Synthesis and Structure of Glycerolipids. Molecular Principles, Components, Technology, and Concepts: Membranes: Composition, Physical Properties, and Curvature; The Outer Mitochondrial Membrane, a Smooth 'Coat' with Many Holes and Many Roles: Preparation, Protein Components, Interactions with Other Membranes, Involvement in Health, Disease, and as a Drug Target

References

- Bader, A.N., Hoetzel, S., Hofman, E.G., *et al.*, 2011. Homo-FRET imaging as a tool to quantify protein and lipid clustering. *Chemphyschem: A European Journal of Chemical Physics and Physical Chemistry* 12, 475–483.

- Bigay, J., Antonny, B., 2012. Curvature, lipid packing, and electrostatics of membrane organelles: Defining cellular territories in determining specificity. *Developmental Cell* 23, 886–895.
- Cascianelli, G., Villani, M., Tosti, M., *et al.*, 2008. Lipid microdomains in cell nucleus. *Molecular Biology of the Cell* 19, 5289–5295.
- Coskun, U., Simons, K., 2011. Cell membranes: The lipid perspective. *Structure* 19, 1543–1548.
- D'Addario, C., Di Francesco, A., Pucci, M., Finazzi Agro, A., Maccarrone, M., 2013. Epigenetic mechanisms and endocannabinoid signalling. *FEBS Journal* 280, 1905–1917.
- Di Paolo, G., De Camilli, P., 2006. Phosphoinositides in cell regulation and membrane dynamics. *Nature* 443, 651–657.
- Gallego, O., Betts, M.J., Gvozdenovic-Jeremic, J., *et al.*, 2010. A systematic screen for protein-lipid interactions in *Saccharomyces cerevisiae*. *Molecular Systems Biology* 6, 430.
- Gao, J., Zheng, H., 2013. Illuminating the lipidome to advance biomedical research: Peptide-based probes of membrane lipids. *Future Medicinal Chemistry* 5, 947–959.
- Garcia, P., Gupta, R., Shah, S., *et al.*, 1995. The pleckstrin homology domain of phospholipase C-delta 1 binds with high affinity to phosphatidylinositol 4,5-bisphosphate in bilayer membranes. *Biochemistry* 34, 16228–16234.
- Guillou, H., Stephens, L.R., Hawkins, P.T., 2007. Quantitative measurement of phosphatidylinositol 3,4,5-trisphosphate. *Methods in Enzymology* 434, 117–130.
- Hokin, L.E., 1985. Receptors and phosphoinositide-generated second messengers. *Annual Review of Biochemistry* 54, 205–235.
- Huang, H., Frohman, M.A., 2009. Lipid signaling on the mitochondrial surface. *Biochimica et Biophysica Acta* 1791, 839–844.
- Hulce, J.J., Cognetta, A.B., Niphakis, M.J., Tully, S.E., Cravatt, B.F., 2013. Proteome-wide mapping of cholesterol-interacting proteins in mammalian cells. *Nature Methods* 10, 259–264.
- Hunter, S., Jones, P., Mitchell, A., *et al.*, 2012. InterPro in 2011: New developments in the family and domain prediction database. *Nucleic Acids Research* 40, D306–D312.
- Layerenza, J.P., González, P., García de Bravo, M.M., *et al.*, 2013. Nuclear lipid droplets: A novel nuclear domain. *Biochimica et Biophysica Acta (BBA) – Molecular and Cell Biology of Lipids* 1831, 327–340.
- Li, W., Laishram, R.S., Anderson, R.A., 2013. The novel poly(A) polymerase Star-PAP is a signal-regulated switch at the 3'-end of mRNAs. *Advances in Biological Regulation* 53, 64–76.
- Lipowsky, R., 2002. Domains and rafts in membranes – Hidden dimensions of selforganization. *Journal of Biological Physics* 28, 195–210.
- Mamedova, L.K., Yuan, K., Laudick, A.N., *et al.*, 2013. Toll-like receptor 4 signaling is required for induction of gluconeogenic gene expression by palmitate in human hepatic carcinoma cells. *Journal of Nutritional Biochemistry* 24, 1499–1507.
- Maxfield, F.R., van Meer, G., 2010. Cholesterol, the central lipid of mammalian cells. *Current Opinion in Cell Biology* 22, 422–429.
- Mayinger, P., 2012. Phosphoinositides and vesicular membrane traffic. *Biochimica et Biophysica Acta* 1821, 1104–1113.
- Nemoto, Y., De Camilli, P., 1999. Recruitment of an alternatively spliced form of synaptotagmin 2 to mitochondria by the interaction with the PDZ domain of a mitochondrial outer membrane protein. *EMBO Journal* 18, 2991–3006.
- Nicolson, G.L., 2014. The fluid-mosaic model of membrane structure: Still relevant to understanding the structure, function and dynamics of biological membranes after more than 40 years. *Biochimica et Biophysica Acta* 1838, 1451–1466.
- Rittiner, J.E., Brings, V.E., Zylka, M.J., 2014. Overexpression of diacylglycerol kinase ϵ (DGK ϵ) prolongs Galphagq-coupled GPCR signaling. *Molecular Pharmacology* 85 (5), 800–810.
- Rong, S.B., Hu, Y., Eryedy, I., *et al.*, 2001. Molecular modeling studies of the Akt PH domain and its interaction with phosphoinositides. *Journal of Medicinal Chemistry* 44, 898–908.
- Shevchenko, A., Simons, K., 2010. Lipidomics: Coming to grips with lipid diversity. *Nature Reviews Molecular Cell Biology* 11, 593–598.
- Smith, M.D., Sudhakar, C.G., Gong, D., Stahelin, R.V., Best, M.D., 2009. Modular synthesis of biologically active phosphatidic acid probes using click chemistry. *Molecular Biosystems* 5, 962–972.
- Tabellini, G., Bortol, R., Santi, S., *et al.*, 2003. Diacylglycerol kinase- θ is localized in the speckle domains of the nucleus. *Experimental Cell Research* 287, 143–154.
- Takatori, S., Mesman, R., Fujimoto, T., 2014. Microscopic methods to observe the distribution of lipids in the cellular membrane. *Biochemistry* 53, 639–653.
- Touboul, D., Gaudin, M., 2014. Lipidomics of Alzheimer's disease. *Bioanalysis* 6, 541–561.
- Wymann, M.P., Schneider, R., 2008. Lipid signalling in disease. *Nature Reviews Molecular Cell Biology* 9, 162–176.

Further Reading

- Boije, A.F., Gennas, G., Talman, V., *et al.*, 2011. Current status and future prospects of C1 domain ligands as drug candidates. *Current Topics in Medicinal Chemistry* 11, 1370–1392.
- Clayton, E.L., Minogue, S., Waugh, M.G., 2013. Mammalian phosphatidylinositol 4-kinases as modulators of membrane trafficking and lipid signaling networks. *Progress in Lipid Research* 52, 294–304.
- De Matteis, M.A., Wilson, C., D'Angelo, G., 2013. Phosphatidylinositol-4-phosphate: The Golgi and beyond. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 35, 612–622.
- Donati, C., Cencetti, F., Bruni, P., 2013. New insights into the role of sphingosine 1-phosphate and lysophosphatidic acid in the regulation of skeletal muscle cell biology. *Biochimica et Biophysica Acta* 1831, 176–184.
- Gericke, A., Leslie, N.R., Losche, M., Ross, A.H., 2013. PtdIns(4,5)P₂-mediated cell signaling: Emerging principles and PTEN as a paradigm for regulatory mechanism. *Advances in Experimental Medicine and Biology* 991, 85–104.
- Gomez-Munoz, A., 2006. Ceramide 1-phosphate/ceramide, a switch between life and death. *Biochimica et Biophysica Acta* 1758, 2049–2056.
- Gomez-Munoz, A., Gangoiti, P., Granado, M.H., Arana, L., Ouro, A., 2010. Ceramide-1-phosphate in cell survival and inflammatory signaling. *Advances in Experimental Medicine and Biology* 688, 118–130.
- Ishisaka, M., Hara, H., 2014. The roles of diacylglycerol kinases in the central nervous system: Review of genetic studies in mice. *Journal of Pharmacological Sciences* 124, 336–343.
- Joshi, R.P., Koretzky, G.A., 2013. Diacylglycerol kinases: Regulated controllers of T cell activation, function, and development. *International Journal of Molecular Sciences* 14, 6649–6673.
- Kalish, B.T., Kieran, M.W., Puder, M., Panigrahy, D., 2013. The growing role of eicosanoids in tissue regeneration, repair, and wound healing. *Prostaglandins & Other Lipid Mediators* 104–105, 130–138.
- Lee, H.J., Rao, J.S., Rapoport, S.I., Bazinet, R.P., 2007. Antimanic therapies target brain arachidonic acid signaling: Lessons learned about the regulation of brain fatty acid metabolism. *Prostaglandins, Leukotrienes, and Essential Fatty Acids* 77, 239–246.
- Mencarelli, C., Martinez-Martinez, P., 2013. Ceramide function in the brain: When a slight tilt is enough. *Cellular and Molecular Life Sciences* 70, 181–203.
- Michell, R.H., 2013. Inositol lipids: From an archaean origin to phosphatidylinositol 3,5-bisphosphate faults in human disease. *FEBS Journal* 280, 6281–6294.
- Oliveira, T.G., Di Paolo, G., 2010. Phospholipase D in brain function and Alzheimer's disease. *Biochimica et Biophysica Acta* 1801, 799–805.
- Olivera, A., Rivera, J., 2011. An emerging role for the lipid mediator sphingosine-1-phosphate in mast cell effector function and allergic disease. *Advances in Experimental Medicine and Biology* 716, 123–142.
- Serrano-Mollar, A., Closa, D., 2005. Arachidonic acid signaling in pathogenesis of allergy: Therapeutic implications. *Current Drug Targets. Inflammation and Allergy* 4, 151–155.
- Soberman, R.J., Christmas, P., 2003. The organization and consequences of eicosanoid signaling. *Journal of Clinical Investigation* 111, 1107–1113.
- Takuwa, Y., Okamoto, Y., Yoshioka, K., Takuwa, N., 2012. Sphingosine-1-phosphate signaling in physiology and diseases. *BioFactors (Oxford, England)* 38, 329–337.
- Wang, X., Devaiah, S., Zhang, W., Welti, R., 2006. Signaling functions of phosphatidic acid. *Progress in Lipid Research* 45, 250–278.
- Yuyama, K., Mitsutake, S., Igarashi, Y., 2013. Pathological roles of ceramide and its metabolites in metabolic syndrome and Alzheimer's disease. *Biochimica et Biophysica Acta* 1841, 793–798.
- Zhang, Y., Du, G., 2009. Phosphatidic acid signaling regulation of Ras superfamily of small guanine triphosphatases. *Biochimica et Biophysica Acta* 1791, 850–855.

Relevant Websites

- <http://amigo2.geneontology.org/cgi-bin/amigo2/amigo2>
Amigo2.
- www.ebi.ac.uk/interpro/
InterPro.
- <http://www.lipidmaps.org/>
Lipid MAPS Consortium.

MOLECULAR PRINCIPLES, COMPONENTS, TECHNOLOGY, AND CONCEPTS: MEMBRANES

Contents

Composition, Physical Properties, and Curvature

Lipid Rafts/Membrane Rafts

Membrane Potential: Concepts

The Outer Mitochondrial Membrane, a Smooth 'Coat' with Many Holes and Many Roles: Preparation, Protein Components, Interactions with Other Membranes, Involvement in Health, Disease, and as a Drug Target

Neuronal Action Potentials and Ion Channel Allostery

Cystic Fibrosis

Composition, Physical Properties, and Curvature

B Antonyy, Université de Nice Sophia Antipolis et CNRS, Valbonne, France

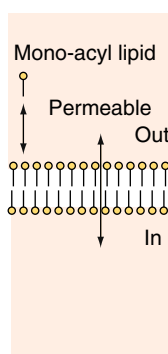
© 2016 Elsevier Inc. All rights reserved.

Origins

Making a membrane was an essential step in life appearance allowing chemical reactions to occur in a confined environment. Experiments aimed at defining minimal membrane systems capable of self-generation by division suggest that very simple lipid species are endowed with such properties ([Budin and Szostak, 2011](#)). According to these studies, primitive membranes would have self-assembled from simple single-chain lipids like fatty acids ([Figure 1](#)). Such membranes are loosely

packed, highly permeable to small metabolites, constantly exchange their components with the environment, and can spontaneously divide. Thereafter, the ability to synthesize double-chain lipids would have gradually reduced the probability of lipid loss, making membranes less dependent on the supply of molecules from the environment. However, this increase in membrane stability correlates with a decrease in permeability and a reduced ability to undergo spontaneous morphological changes. Therefore, the transition from single-chain lipids to double-chain lipids might have coevolved with the acquisition

Primitive membrane



Modern cell membrane

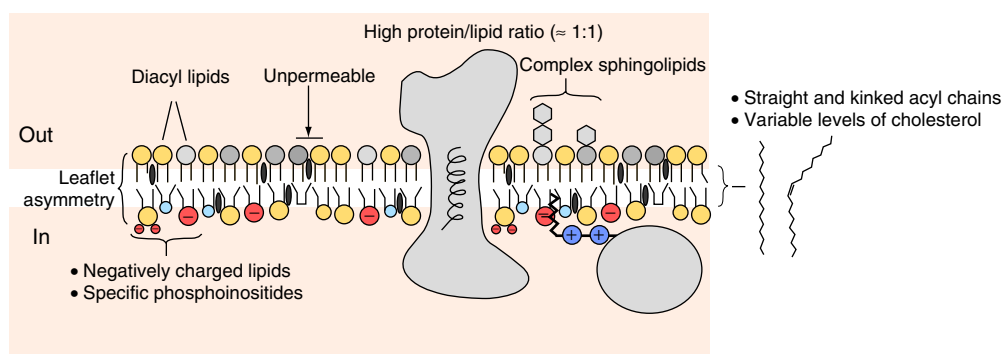


Figure 1 From primitive to modern cell membranes. Primitive cellular membranes might have formed from the self-assembly of simple molecules containing one hydrocarbon acyl chain. These membranes are highly permeable and their stability requires continuous exchange with the environment. The ability to synthesize diacyl lipid was probably a turning point to make membranes more stable and therefore less dependent on the environment. In turn, this transition would have required the emergence of intracellular mechanisms to control exchanges across the bilayer and to promote membrane division. Modern cell membranes are extremely complex. They are very rich in transmembrane and peripheral proteins, they contain hundreds of different lipid species, and their two leaflets can be very different in composition. The composition of both bulk and rare lipids is also different between cellular organelles.

of primitive mechanisms to shape membranes and transport molecules across them (Budin and Szostak, 2011).

General Traits of Biological Membranes

In currently living organisms, cellular membranes are extremely complex in composition, using a repertoire of hundreds if not thousands of lipid species (Figure 1). One of the most detailed pictures of a real biological membrane comes from studies of synaptic vesicles, which encapsulate neurotransmitters and are at the basis of communication between neurons (Takamori *et al.*, 2006). Comprehensive lipid and protein analysis (i.e., proteomic and lipidomic studies) reveal a very crowded membrane with numerous proteins of different structure and function and about 50 different lipid species. This complexity is intimidating and raises a central question: Do all details matter or can we extract some general traits that distinguish or unify membrane-bound organelles?

The abundance of transmembrane proteins in the synaptic vesicle is not unique. Although the bilayer structure is governed by the amphiphilic properties of lipids, biological membranes generally contain high amounts of proteins (Engelman, 2005); some of them cross the bilayers through transmembrane helical segments; others are peripherally associated through various means. This crowding probably influences the ability of proteins to move laterally (diffusion), modifies the mechanical properties of membranes, and can lead to membrane heterogeneities.

The building blocks of cellular membranes are glycerophospholipids, sphingolipids, and sterols (Holthuis and Levine, 2005). Glycerophospholipids encompass both bulk lipids, i.e., abundant lipids that allow the membrane to be a bilayer, and rare lipids, which act as second messengers or as a flag of subcellular compartments (Di Paolo and De Camilli, 2006). $PI_{(3,4,5)}P_3$, a phosphoinositide with three phosphate groups on its polar head, is an example of a lipid messenger, acting as a compass for cell polarity (Servant *et al.*, 2000). $PI_{(3)}P$, with one phosphate group on its polar head group is an example of a flag lipid, being present exclusively on some organelles (endosomes) (Burd and Emr, 1998). The most abundant phospholipids are phosphatidylcholine (PC), phosphatidylethanolamine (PE), and phosphatidylserine (PS). PC and PE are zwitterionic since their polar heads contain one negative and one positive charge, whereas PS is anionic since its polar head contains two negative charges and one positive charge (Holthuis and Levine, 2005).

Sphingolipids are deceptively similar to phospholipids as they also contain two acyl chains and one polar head group. However, subtle difference in the chemistry of sphingolipids compared to glycerophospholipids (longer acyl chains and more possibilities of hydrogen bonds) and the fact that complex sugar groups decorate the polar head make their physicochemical properties different. Sphingolipids are abundant on the external face of the cell plasma membrane and play essential roles in cell protection and cell–cell communication (Holthuis and Levine, 2005; Simons and Sampaio, 2011).

Sterols are very different from both sphingolipids and phospholipids being made by a rigid carbon ring structure decorated, on the one hand, by a minimal polar group

(a hydroxyl group) and, on the other hand, by a short carbon chain. A high concentration of cholesterol helps membranes to be impermeable to small molecules including water. However, the effects of cholesterol are complex and depend on the nature of the neighboring lipids.

Overall, the identity of each membrane-bound organelle results from the presence of unique proteins and of rare lipids, from the relative amounts of bulk lipids, and from some physical features such as shape or interaction with a protein matrix (e.g., the actin cytoskeleton underneath the plasma membrane). In the case of synaptic vesicles that we have used before as an example, their noticeable features are the presence of key proteins involved in neurotransmitter uptake and release, the abundance of polyunsaturated phospholipids and cholesterol (40 mol% of all lipids), their extreme curvature (radius 20 nm), and their embedding in a dense protein matrix (Takamori *et al.*, 2006). Endocytic vesicles are also very dynamic organelles whose formation and consumption correlate with abrupt changes in specific phosphoinositides species (Posor *et al.*, 2013).

Most organelles exchange vesicles or other membrane-bound intermediates, a process referred as to membrane traffic. Consequently, the physicochemical features of many organelles vary according to their position along major trafficking routes. Thus, the endoplasmic reticulum (ER) and the plasma membrane define two extreme cases in term of lipid composition, whereas organelles such as the Golgi apparatus or endosomes display intermediate traits. In this respect, two parameters define a first layer of organelle identity: membrane electrostatics, which depends on the abundance of negatively charged lipids (PS and phosphoinositides); and lipid packing, which depends on the presence of ‘kinks’ in fatty acid chains and on cholesterol levels (Bigay and Antonny, 2012; Holthuis and Levine, 2005).

Membrane Electrostatics

The gene product Src has a central history in cancer biology, as it was the first proto-oncogene to be discovered. In addition, illuminating experiments on the membrane targeting of this tyrosine kinase had been instrumental for our understanding of the organelle distribution of many cytosolic proteins (McLaughlin and Aderem, 1995). Src is attached to the cytosolic leaflet of the plasma membrane through an N-terminal sequence, which led to the hypothesis that a membrane protein recognizes this sequence. Because experiments aimed at identifying this putative Src receptor failed, another hypothesis emerged: the N-terminal sequence of Src, which is very rich in positively charged residues, interacts directly with negatively charged lipids. This second hypothesis turned out to be true: the membrane targeting of Src, both *in vitro* and *in vivo*, as well as its transforming activity correlate with the density of positively charged residues in the N-terminus, regardless of its exact amino acid sequence (McLaughlin and Aderem, 1995).

As rudimentary as it may seem, membrane electrostatics is a powerful mean to target cytosolic proteins to the plasma membrane. The cytosolic leaflet of this membrane is indeed very rich in negatively charged lipids, notably PS (net charge = -1) and $PI_{(4,5)}P_2$ (net charge = -4), as compared to

other organelles. Thus, artificial constructs bearing 2, 4, 6, or 8 accessible positively charged residues show a gradual enrichment at the plasma membrane at the expense of internal organelles (Yeung *et al.*, 2006). In many cases, close variants of the same protein (e.g., small G proteins of the Ras or Rho subfamilies) are targeted to different organelles owing to differences in the number of positively charged residues of their membrane targeting sequence.

Nonspecific electrostatic interactions between peripheral proteins and membranes are quite weak. Consequently, they generally combine with other elementary mechanisms of membrane anchorage. In the case of Src, a myristate (a 14-carbon acyl chain) is covalently attached to a glycine upstream of the N-terminal-basic sequence. Strong binding of Src to the membrane results both from myristate insertion in the hydrophobic core of the membrane and from electrostatics interactions of the N-terminal sequence with the membrane surface (McLaughlin and Aderem, 1995). The combination of weak interactions is a recurrent theme in membrane targeting mechanisms and has two advantages. First, it allows cycles of membrane association-dissociation because disrupting one elementary membrane interaction (e.g., phosphorylation of a basic sequence) is generally sufficient to promote protein detachment. Second, it improves the specificity of membrane targeting: binding occurs only when the membrane displays features that are compatible with the association of each elementary membrane-anchoring motif (e.g., negatively charged lipids and curvature).

Overall, membrane electrostatics governs a membrane territory that starts at the *trans* Golgi, includes various endosomal membranes and culminates at the plasma membrane. Not surprisingly, protein machineries that work in or on these membranes take advantage of this feature. A good example is cortical actin, whose nucleators at the plasma membrane interact with the phosphoinositide PI(4,5)P₂ and PS (Papayannopoulos *et al.*, 2005).

Lipid Packing

The majority of lipids that form biological membranes have a cylindrical shape: their polar heads have approximately the same cross-sectional area as their acyl chains. The best example of a cylindrical lipid is C16:0-C18:1 PC, which contains one saturated (C16:0) and one monounsaturated acyl chain (C18:1). This balance in molecular size allows the formation of stable bilayers (Holthuis and Levine, 2005). Yet, biological membranes also contain lipids whose shape would result in assemblies different from bilayers if those components were the majority. Consequently, most biological bilayers are under stress: the packing of their lipid components is not perfect and probably results in the formation of voids or 'defects' (Bigay and Antonny, 2012).

Although somehow counterintuitive, the presence of lipid packing defects in biological membranes is functionally important. First, lipids with a non-bilayer shape facilitate complex reactions such as membrane fusion, in which the membrane transiently and locally loses its bilayer structure (Zick *et al.*, 2014). Second, the propensity of cellular membranes to be loosely or tightly packed seems correlated with

their functions. The ER contains low levels of cholesterol and its glycerolphospholipids are generally monounsaturated: their acyl chains contain a *cis* double bond (e.g., C18:1), which creates a kink right in the middle of the hydrocarbon chain. In contrast, the plasma membrane is rich in cholesterol and in saturated lipids (Schneider *et al.*, 1999). Lipid homeostasis pathways help to maintain these parameters within narrow ranges. For example, transmembrane proteins of the ER sense the level of cholesterol and trigger a feedback loop response that ultimately controls sterol synthesis (Radhakrishnan *et al.*, 2008).

The loosely packed structure of the ER membrane probably facilitates protein and lipid synthesis reactions that occur in the hydrophobic matrix (Bretscher and Munro, 1993; Deguil *et al.*, 2011; Nilsson *et al.*, 2001). Conversely, tight packing between cholesterol and saturated (straight) acyl chains at the plasma membrane is important for making this membrane an efficient cell barrier. In specialized cells, other adaptations of the hydrophobic matrix have been reported. The discs of photoreceptor discs are made almost exclusively of polyunsaturated PLs. These highly flexible lipids facilitate the conformational change of rhodopsin, thereby contributing to the exceptional signal to noise ratio of phototransduction (Mitchell *et al.*, 2001). More recently, a gradient of polyunsaturated phospholipids has been reported along the axon (Yang *et al.*, 2012). The function of this gradient is not known but it is probably related to the extraordinary enrichment of polyunsaturated PLs in synaptic vesicles (Takamori *et al.*, 2006). *In vitro*, polyunsaturated PLs facilitate membrane deformation and fission (Pinot *et al.*, 2014).

In conclusion, the membrane territory governed by lipid packing roughly mirrors that defined by electrostatics: membranes of the early secretory pathway (ER to *cis* Golgi) seem to combine loose lipid packing and low electrostatic, whereas membranes of the late secretory pathway (*trans* Golgi, endosomes, and plasma membrane) seem to combine tight lipid packing and high electrostatics (Figure 2).

Membrane Subdomains

In a closed membrane, the general tendency of proteins and lipids is to diffuse within the full membrane area to maximize entropy. Typically, a lipid explores the entire surface of the plasma membrane in a few minutes. Yet, biological membranes are not uniform. The plasma membrane of epithelial cell is divided into a basolateral region and a highly protective apical membrane barrier (Simons and van Meer, 1988). When the yeast *Saccharomyces cerevisiae* divides, the continuity of the ER is interrupted to prevent the daughter cell from acquiring 'aging factors' that have accumulated on the ER membrane of the mother cell (Clay *et al.*, 2014). At the *trans* Golgi network or in endosomes, membrane domains form to prepare the emergence of different types of transport intermediates (Sonnichsen *et al.*, 2000).

A complex issue in membrane biology is to determine what factors govern the emergence of membrane subdomains. A popular but controversial hypothesis is that these domains arise from the tendency of some lipid species to form different phases; classically a liquid ordered phase rich in saturated

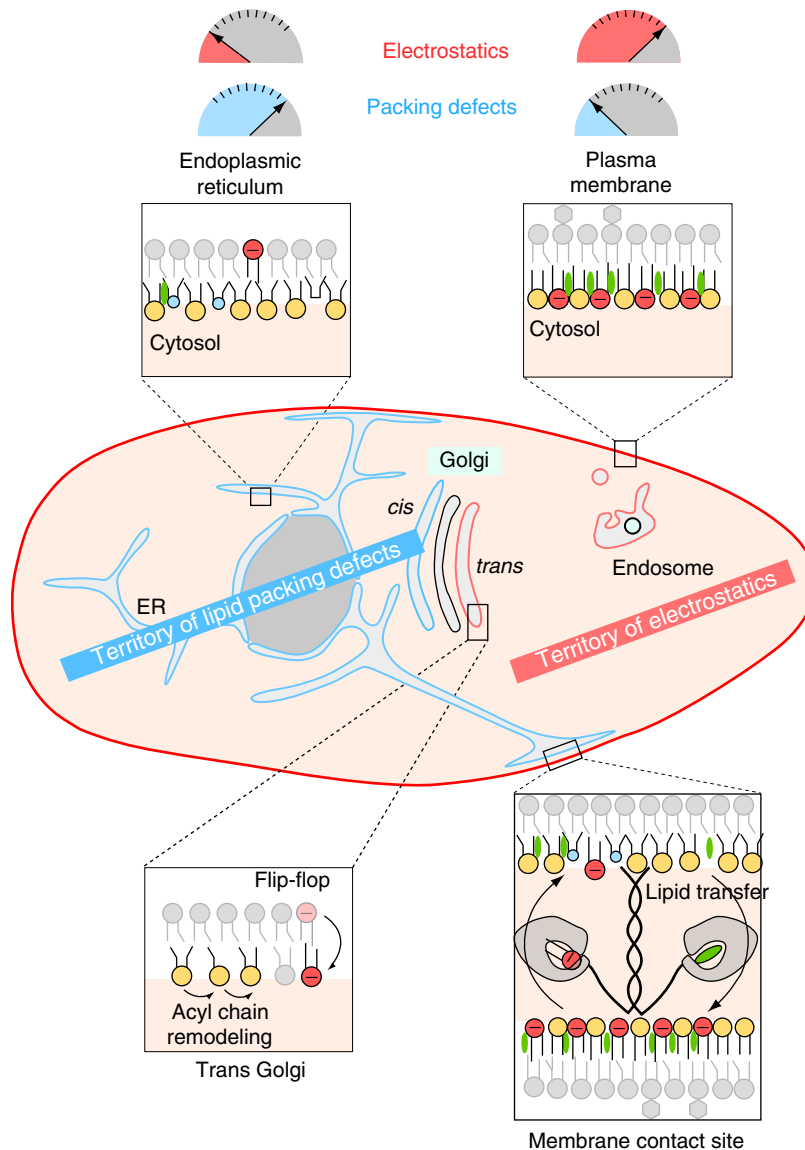


Figure 2 Two membrane territories. The ER and *cis* Golgi membranes are poor in negatively charged lipids and display lipid packing defects owing to the presence of lipids with monounsaturated chains. Membranes from the *trans* Golgi to the plasma membrane harbor negatively charged lipids on their cytosolic face and are tightly packed as they are rich in cholesterol and saturated lipids. These opposite changes in packing and electrostatics are controlled by enzymatic activities, notably at the Golgi apparatus: flippases move phosphatidylserine from the luminal leaflet to the cytoplasmic leaflet (lipid flip-flop). Phospholipases and acyltransferases promote the replacement of the second acyl chain of many phospholipids (lipid remodeling). In addition, membrane contact sites are privileged zones where lipids are exchanged between the ER and compartments such as the *trans* Golgi, endosomes, the plasma membrane, or mitochondria.

sphingolipids and cholesterol and a liquid disordered phase with less saturated lipids (Simons and Sampaio, 2011). However, if such domains can readily form in artificial lipid membranes, their existence in biological organelles remains elusive owing to the difficulty in isolating them (Munro, 2003). Moreover, the bulk lipid composition of most organelles does not seem favorable for the formation of large liquid ordered domains, a notable exception being the outer leaflet of the plasma membrane. Notwithstanding, many other factors might also contribute to the formation of lateral membrane inhomogeneities, including interactions with the underneath cytoskeleton, assemblies of large protein oligomers such as

septins, and numerous other lipid-protein interactions. Overall, if membrane subdomains clearly exist, our understanding of the factors that drive their formation remains fragmental.

Mechanisms that Modifies the Bulk Lipid Composition of Organelles

Several mechanisms contribute to the nonrandom distribution of negatively charged lipids in the cell. First PI-kinases and PI-phosphatases are localized to well-defined organelles

(Di Paolo and De Camilli, 2006). Second, PS is synthesized at the ER but is mostly present on the luminal side of this organelle and as such is inaccessible to cytosolic proteins (Fairn *et al.*, 2011). At the Golgi, transmembrane ATPases or 'flippases' promote the flip-flop of PS from the luminal to the cytosolic leaflet (Zhou and Graham, 2009), thereby making it available for the targeting of cytosolic proteins containing positively charged sequences.

How the cell gradually remodels the acyl chains of its glycerophospholipids, a process referred as to the Land's cycle, has remained mysterious for decades. Lipid remodeling involves the sequential action of phospholipases and acyltransferases to replace the esterified acyl chains on the phospholipid glycerol backbone; for example, to convert a C18:1-C18:1 into a C16:0-C18:1 lipid. The recent cloning of many acyltransferases and the realization that some members reside at the Golgi suggest that this organelle is a privileged region for changing the hydrophobic matrix properties (Shindou *et al.*, 2009).

The activities of lipid flippases, acyltransferases, and other lipid-modifying enzymes likely contribute to the large changes in membrane electrostatics, lipid packing, membrane asymmetry, and membrane thickness that occur at the Golgi, which defines the frontier of two main membrane territories (pre- and post-Golgi membranes). Of note, this frontier also acts as a filter to sort membrane proteins according to the physicochemical properties of their transmembrane domains (Sharpe *et al.*, 2010).

Other mechanisms are being disclosed that maintain different lipid compositions between organelles. Cytosolic lipid transfer proteins extract one lipid from a donor membrane, convey it in the cytosol, and then deliver it to an acceptor membrane. Thus, ceramide transport protein extracts ceramide from the ER and delivers it at the Golgi where ceramide is then transformed into sphingolipids (Hanada *et al.*, 2003). Whether such lipid transfer activities are fast and specific enough to explain the distribution of some lipid species is a long-standing question. Recent studies suggest that two mechanisms might improve the efficiency of lipid transfer reactions (Kim *et al.*, 2013; Mesmin *et al.*, 2013). First, some lipid transfer proteins are equipped with domains that bridge two membranes, leading to the formation of contact sites. Membrane contact sites allow the connection of almost any organelle to the ER, where most lipid building blocks are synthesized, and have a typical width of few tens of nanometers. Within this short distance, many lipid transfer reactions might occur in a coordinated manner. Second, some lipid transfer reactions are coupled: the same protein can transfer one lipid in one direction and another lipid in the opposite direction. As such, phosphoinositide gradients might provide the energy for the fast transfer of lipids such as cholesterol of vitamin E (de Saint-Jean *et al.*, 2011; Kim *et al.*, 2013; Kono *et al.*, 2013; Mesmin *et al.*, 2013).

Membrane Curvature

Pure lipid bilayers are poorly extensible as their area can extend by about 5% before rupture. Because the surface of the plasma membrane is much larger than what would be

required to encapsulate the cell, the cell responds to stretching through other mechanisms such as rupture of plasma membrane-cytoskeleton interactions or takes advantage small invaginations, the caveolae, which act as surface reservoirs (Raucher *et al.*, 2000; Sinha *et al.*, 2011). In contrast, lipid bilayers can be highly bent (Rawicz *et al.*, 2000). In the cell, membrane bending results from the actions of numerous protein machineries (McMahon and Gallop, 2005). The synaptic vesicle (radius 20 nm) is an example of a highly curved membrane that is formed by specialized membrane bending proteins.

The most complex organelle in term of morphology is the ER, which forms a continuous membrane network that emerges from the nuclear envelope and spreads throughout the entire cytoplasm. Many types of membrane geometries are met in this compartment (Shibata *et al.*, 2009). This includes not only long tubules and membrane sheets, but also small spherical vesicles, which transport cargoes toward the Golgi, nuclear pores with a torus-like shape, through which exchanges between the cytosol and the nucleus occur, and flat membrane contact sites with the plasma membrane.

The reticulons and the DP1/Yop1p proteins are transmembrane proteins that contribute to the formation of ER tubules and sheets (Figure 3(a)). These proteins display hairpin-like transmembrane segments that are not long enough to completely cross the bilayer. These hairpins are believed to occupy more volume in the cytosolic leaflet of the ER than its luminal leaflet, thereby leading to membrane deformation. Membrane curvature induced by asymmetric protein insertions in the bilayer matrix is referred as to the wedge mechanism.

The protein machineries involved in vesicle budding from the ER and in nuclear pore formation, the COPII coat and the nuclear pore complex (NPC), respectively, are based on a lattice-like arrangement of cytosolic proteins (Figure 3(b); McMahon and Gallop, 2005). NPC and coat complexes might have derived from the same ancestor (Schwartz, 2013). In both cases, the lattice is intrinsically curved and contacts the membrane indirectly *via* specific adaptors. This mode of membrane deformation where a cytosolic protein or a protein complex imposes its intrinsic shape to the membrane without directly penetrating into the hydrophobic membrane matrix is referred as to the scaffolding mechanism.

The division between scaffolding and wedge mechanisms is instrumental to understand how simple proteins deform membranes (McMahon and Gallop, 2005). However, in many circumstances the wedge and scaffold mechanisms cooperate to shape biological membranes. In the case of reticulon-like proteins, membrane deformation arises not only from the unbalanced chemistry of the hairpin segments but also from the ability of these proteins to oligomerize into arc-like structures. In the case of the COPII coat, the membrane deforming activity of the lattice is assisted by the membrane insertion of amphipathic helices from other proteins (Lee *et al.*, 2005). Amphipathic helices do not cross membrane bilayers but lie at the interface between the polar and nonpolar membrane region. Such shallow mode of membrane insertion is reversible and can be controlled by conformational changes (e.g., the GDP/GTP switch of small G Proteins).

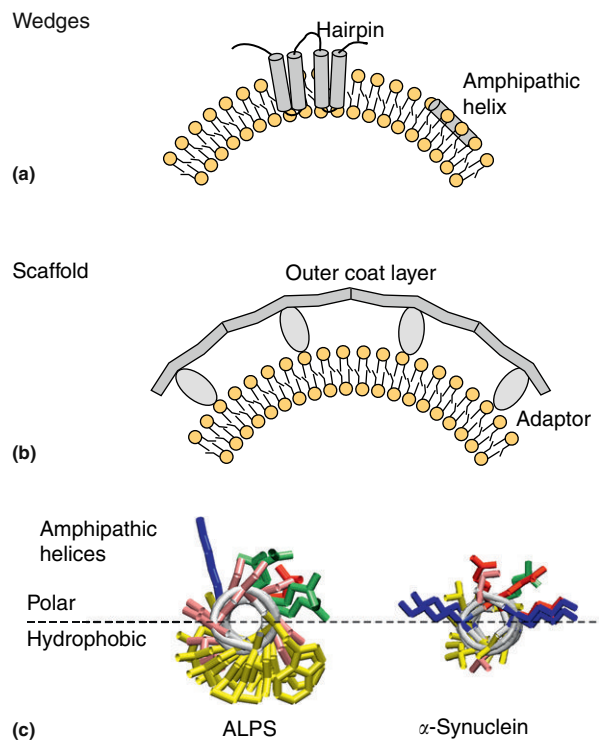


Figure 3 Shaping biological membranes. (a) Membrane deformation by wedges. Reticulon hairpins cross the membrane of the ER in an asymmetric manner, occupying more volume in the cytosolic leaflet than in the luminal one. Amphipathic helices are embedded in the interfacial membrane region, hence causing local deformations. (b) Membrane deformation by scaffolds. Protein coats are typically made of two layers: one contacting directly the membrane and one polymerizing into a spherical scaffold. (c) Atomic models of two amphipathic helices (as viewed from the front). These helices are adapted to different organelles. The helix on the left is an Amphipathic Lipid Packing Sensor: its polar face is made almost exclusively of small and uncharged residues (GLY, SER, THR; in pink), and its hydrophobic face contains bulky hydrophobic residues (LEU, PHE; in yellow). This helix inserts into large lipid packing defects that form in the outer leaflet of small vesicles from the ER or the *cis* Golgi. The helix on the right is the N-terminal region of α -synuclein: its polar face contains two wings of positively charged residues (LYS in blue) and one crest of negatively charged residues (ASP, GLU; in red). Its nonpolar face contains very small hydrophobic residues (VAL, ALA; in yellow). This helix is adapted to vesicles containing negatively charged lipids and small lipid packing defects (e.g., endocytic vesicles arising from the plasma membrane).

As we leave the ER, cross the Golgi apparatus and reach the plasma membrane, we meet other membrane deforming machineries. The most well known are the COPI coat and the clathrin-AP adaptor coats, which share with COPII the same basic architecture: a lattice-like spherical outer layer and an inner layer made of cargo adaptors and amphipathic wedges. However, each coat is adapted to the membrane on which it acts. The COPII coat is attached to the ER *via* the insertion of bulky hydrophobic amino acids of the small G protein Sar1 in the bilayer (Lee *et al.*, 2005), whereas the clathrin coat contacts the plasma membrane *via* adaptors, which recognize the polar head of the phosphoinositide PIP₂ and other anionic

lipids (Jackson *et al.*, 2010). The predominate use of either hydrophobic or electrostatic interactions is also illustrated in Figure 3, which compares the chemistry of two amphipathic helices that adsorb on small vesicles budding from either early membranes (thus characterized by large lipid packing defects) or late membranes (thus containing anionic lipids). This comparison is suggestive of a molecular ‘camouflage’ where each helix is chemically adapted to the membrane interface in which it inserts (Pranke *et al.*, 2011).

Many mechanisms of membrane deformation remain to be analyzed. The entry of some pathogens into the cell is suggestive of primitive modes of membrane deformations that bypass the requirement for specialized cytosolic machineries (Romer *et al.*, 2007). Other interesting examples are the formation of cup-shape membrane structures during autophagy or the deformations induced by endosomal sorting complexes, which occur in the opposite direction as budding by protein coats. Even in the case of protein coats for which spectacular advances have been made, key issues remain; notably how these well-defined protein assemblies cope with extremely crowded membrane surfaces or huge protein cargoes (Copic *et al.*, 2012).

Conclusions

Despite their complexities in composition, the membranes of most cellular organelles can be broadly divided in two groups according to their lipid packing properties and to the presence of negatively charged lipids (Bigay and Antonny, 2012). The ER membrane is poorly packed, a feature probably decisive for its function in protein and in lipid synthesis. The plasma membrane is tightly packed, thus forming an impermeable barrier, and its cytosolic leaflet is very rich in negatively charged lipids allowing the cell to control its shape by using plasma membrane–cytoskeleton interactions. The Golgi apparatus forms the frontier between these two territories and is a privileged place for changes in lipid acyl chain unsaturation, cholesterol content, membrane electrostatics, membrane thickness, and membrane asymmetry.

See also: Dynamic Integration: Dynamics: Theory of Cargo and Membrane Trafficking. Interorganellar Communication: Components: Adaptor Proteins: Inter-Organelle Traffic Controllers. Molecular Principles, Components, Technology, and Concepts: Lipids: Cholesterol and Other Steroids. Molecular Principles, Components, Technology, and Concepts: Membranes: Lipid Rafts/Membrane Rafts. Organelles: Structure and Function: At the Center of Autophagy: Autophagosomes

References

- Bigay, J., Antonny, B., 2012. Curvature, lipid packing, and electrostatics of membrane organelles: Defining cellular territories in determining specificity. *Developmental Cell* 23, 886–895.
- Bretscher, M.S., Munro, S., 1993. Cholesterol and the Golgi apparatus. *Science* 261, 1280–1281.

- Budin, I., Szostak, J.W., 2011. Physical effects underlying the transition from primitive to modern cell membranes. *Proceedings of the National Academy of Sciences of the United States of America* 108, 5249–5254.
- Burd, C.G., Emr, S.D., 1998. Phosphatidylinositol(3)-phosphate signaling mediated by specific binding to RING FYVE domains. *Molecular Cell* 2, 157–162.
- Clay, L., Caudron, F., Denoth-Lippuner, A., *et al.*, 2014. A sphingolipid-dependent diffusion barrier confines ER stress to the yeast mother cell. *eLife* 3, e01883.
- Copic, A., Latham, C.F., Horlbeck, M.A., D'Arcangelo, J.G., Miller, E.A., 2012. ER cargo properties specify a requirement for COPII coat rigidity mediated by Sec13p. *Science* 335, 1359–1362.
- Deguil, J., Pineau, L., Rowland Snyder, E.C., *et al.*, 2011. Modulation of lipid-induced ER stress by fatty acid shape. *Traffic* 12, 349–362.
- Di Paolo, G., De Camilli, P., 2006. Phosphoinositides in cell regulation and membrane dynamics. *Nature* 443, 651–657.
- Engelman, D.M., 2005. Membranes are more mosaic than fluid. *Nature* 438, 578–580.
- Fairn, G.D., Schieber, N.L., Ariotti, N., *et al.*, 2011. High-resolution mapping reveals topologically distinct cellular pools of phosphatidylserine. *Journal of Cell Biology* 194, 257–275.
- Hanada, K., Kumagai, K., Yasuda, S., *et al.*, 2003. Molecular machinery for non-vesicular trafficking of ceramide. *Nature* 426, 803–809.
- Holthuis, J.C., Levine, T.P., 2005. Lipid traffic: Floppy drives and a superhighway. *Nature Reviews Molecular Cell Biology* 6, 209–220.
- Jackson, L.P., Kelly, B.T., McCoy, A.J., *et al.*, 2010. A large-scale conformational change couples membrane recruitment to cargo binding in the AP2 clathrin adaptor complex. *Cell* 141, 1220–1229.
- Kim, Y.J., Hernandez, M.L., Balla, T., 2013. Inositol lipid regulation of lipid transfer in specialized membrane domains. *Trends in Cell Biology* 23, 270–278.
- Kono, N., Ohto, U., Hiramatsu, T., *et al.*, 2013. Impaired alpha-TTP-PIPs interaction underlies familial vitamin E deficiency. *Science* 340, 1106–1110.
- Lee, M.C., Orci, L., Hamamoto, S., *et al.*, 2005. Sar1p N-terminal helix initiates membrane curvature and completes the fission of a COPII vesicle. *Cell* 122, 605–617.
- McLaughlin, S., Aderem, A., 1995. The myristoyl-electrostatic switch: A modulator of reversible protein-membrane interactions. *Trends in Biochemical Sciences* 20, 272–276.
- McMahon, H.T., Gallop, J.L., 2005. Membrane curvature and mechanisms of dynamic cell membrane remodelling. *Nature* 438, 590–596.
- Mesmin, B., Bigay, J., Moser von Filseck, J., *et al.*, 2013. A four-step cycle driven by PI(4)P hydrolysis directs sterol/PI(4)P exchange by the ER-Golgi tether OSBP. *Cell* 155, 830–843.
- Mitchell, D.C., Niu, S.L., Litman, B.J., 2001. Optimization of receptor-G protein coupling by bilayer lipid composition I: Kinetics of rhodopsin-transducin binding. *Journal of Biological Chemistry* 276, 42801–42806.
- Munro, S., 2003. Lipid rafts: Elusive or illusive? *Cell* 115, 377–388.
- Nilsson, I., Ohvo-Rekila, H., Slotte, J.P., Johnson, A.E., von Heijne, G., 2001. Inhibition of protein translocation across the endoplasmic reticulum membrane by sterols. *Journal of Biological Chemistry* 276, 41748–41754.
- Papayannopoulos, V., Co, C., Prehoda, K.E., *et al.*, 2005. A polybasic motif allows N-WASP to act as a sensor of PIP(2) density. *Molecular Cell* 17, 181–191.
- Pinot, M., Vanni, S., Pagnotta, S., *et al.*, 2014. Polyunsaturated phospholipids facilitate membrane deformation and fission by endocytic proteins. *Science* 345, 693–697.
- Posor, Y., Eichhorn-Gruenig, M., Puchkov, D., *et al.*, 2013. Spatiotemporal control of endocytosis by phosphatidylinositol-3,4-bisphosphate. *Nature* 499, 233–237.
- Pranke, I.M., Morello, V., Bigay, J., *et al.*, 2011. alpha-Synuclein and ALPS motifs are membrane curvature sensors whose contrasting chemistry mediates selective vesicle binding. *Journal of Cell Biology* 194, 89–103.
- Radhakrishnan, A., Goldstein, J.L., McDonald, J.G., Brown, M.S., 2008. Switch-like control of SREBP-2 transport triggered by small changes in ER cholesterol: A delicate balance. *Cell Metabolism* 8, 512–521.
- Raucher, D., Stauffer, T., Chen, W., *et al.*, 2000. Phosphatidylinositol 4,5-bisphosphate functions as a second messenger that regulates cytoskeleton-plasma membrane adhesion. *Cell* 100, 221–228.
- Rawicz, W., Olbrich, K.C., McIntosh, T., Needham, D., Evans, E., 2000. Effect of chain length and unsaturation on elasticity of lipid bilayers. *Biophysical Journal* 79, 328–339.
- Romer, W., Berland, L., Chambon, V., *et al.*, 2007. Shiga toxin induces tubular membrane invaginations for its uptake into cells. *Nature* 450, 670–675.
- de Saint-Jean, M., Delfosse, V., Douguet, D., *et al.*, 2011. Osh4p exchanges sterols for phosphatidylinositol 4-phosphate between lipid bilayers. *Journal of Cell Biology* 195, 965–978.
- Schneider, R., Brugger, B., Sandhoff, R., *et al.*, 1999. Electrospray ionization tandem mass spectrometry (ESI-MS/MS) analysis of the lipid molecular species composition of yeast subcellular membranes reveals acyl chain-based sorting/remodeling of distinct molecular species en route to the plasma membrane. *Journal of Cell Biology* 146, 741–754.
- Schwartz, T., 2013. Functional insights from studies on the structure of the nuclear pore and coat protein complexes. *Cold Spring Harbor Perspectives* 5, a013375.
- Servant, G., Weiner, O.D., Herzmark, P., Balla, T., Sedat, J.W., Bourne, H.R., 2000. Polarization of chemoattractant receptor signaling during neutrophil chemotaxis. *Science* 287, 1037–1040.
- Sharpe, H.J., Stevens, T.J., Munro, S., 2010. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell* 142, 158–169.
- Shibata, Y., Hu, J., Kozlov, M.M., Rapoport, T.A., 2009. Mechanisms shaping the membranes of cellular organelles. *Annual Review of Cell and Developmental Biology* 25, 329–354.
- Shindou, H., Hishikawa, D., Harayama, T., Yuki, K., Shimizu, T., 2009. Recent progress on acyl CoA: Lysophospholipid acyltransferase research. *Journal of Lipid Research* 50 (Suppl.), S46–51.
- Simons, K., Sampaio, J.L., 2011. Membrane organization and lipid rafts. *Cold Spring Harbor Perspectives in Biology* 3, a004697.
- Simons, K., van Meer, G., 1988. Lipid sorting in epithelial cells. *Biochemistry* 27, 6197–6202.
- Sinha, B., Koster, D., Ruez, R., *et al.*, 2011. Cells respond to mechanical stress by rapid disassembly of caveolae. *Cell* 144, 402–413.
- Sonnichsen, B., De Renzis, S., Nielsen, E., Rietdorf, J., Zerial, M., 2000. Distinct membrane domains on endosomes in the recycling pathway visualized by multicolor imaging of Rab4, Rab5, and Rab11. *Journal of Cell Biology* 149, 901–914.
- Takamori, S., Holt, M., Stenius, K., *et al.*, 2006. Molecular anatomy of a trafficking organelle. *Cell* 127, 831–846.
- Yang, H.J., Sugiura, Y., Ikegami, K., Konishi, Y., Setou, M., 2012. Axonal gradient of arachidonic acid-containing phosphatidylcholine and its dependence on actin dynamics. *Journal of Biological Chemistry* 287, 5290–5300.
- Yeung, T., Terebiznik, M., Yu, L., *et al.*, 2006. Receptor activation alters inner surface potential during phagocytosis. *Science* 313, 347–351.
- Zhou, X., Graham, T.R., 2009. Reconstitution of phospholipid translocase activity with purified Drs2p, a type-IV P-type ATPase from budding yeast. *Proceedings of the National Academy of Sciences of the United States of America* 106, 16586–16591.
- Zick, M., Stroupe, C., Orr, A., Douville, D., Wickner, W.T., 2014. Membranes linked by trans-SNARE complexes require lipids prone to non-bilayer structure for progression to fusion. *eLife* 3, e01879.

Relevant Website

<http://www.lipidmaps.org/>
LIPID MAPS Lipidomics Gateway.

Lipid Rafts/Membrane Rafts

E London, Stony Brook University, Stony Brook, NY, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The fluid mosaic model of membrane structure is often thought of in terms of a homogeneous mixture of lipids in a lipid bilayer studded with proteins. An alternative is that there are domains consisting of spatially segregated regions with specialized lipid and protein compositions. Lipid rafts or membrane rafts are the most widely studied of such domains. The name raft was adopted to describe the concept of a specialized small (protein-containing) lipid cluster coexisting with a surrounding ocean of less specialized lipids and proteins. Even though this name has been applied to diverse concepts, for simplicity the name raft will be used here in its most commonly defined sense, physically segregated membrane domains of lipids, generally rich in sphingolipids and sterols, plus associated proteins, existing in the tightly packed liquid ordered (Lo) state. The subject of lipid or membrane rafts has been very controversial. However, recent studies have unequivocally identified raft-type domains in living cells. This article concentrates upon the properties of rafts and how they may be studied.

The Physical States of Lipids in Lipid Bilayers

Phospholipids and sphingolipids can form lipid bilayers in either a tightly packed solid-like (gel) state or loosely packed fluid (liquid disordered, Ld) state. For membrane vesicles composed of a specific lipid, the gel state melts to form the Ld state at a characteristic melting temperature (T_m). As simple physical chemistry predicts, gel and Ld phases often coexist in mixtures of high- T_m value lipids (e.g., natural sphingolipids, which easily form a tightly packed ordered state due to having saturated acyl chains lacking *cis* double bonds) and low- T_m lipids (ordinary phospholipids, which generally have one or more *cis* double bonds in one of their two acyl chains) (Korlach *et al.*, 1999). Since the T_m for sphingolipids is generally at or above physiological temperature, the possibility that natural membranes containing both sphingolipids and phospholipids could contain coexisting solid-like and liquid-like membrane domains was long considered (Thompson and Tillack, 1985). However, large phase-separated membrane domains were not observed in cells. Furthermore, it was thought cholesterol would prevent formation of coexisting solid and liquid domains. When cholesterol is mixed with other membrane lipids, especially lipids containing saturated acyl chains, i.e., lacking double bonds, it was found that a state with intermediate properties forms. This was first named the gamma state, but later called the Lo state (Ipsen *et al.*, 1987; Lentz *et al.*, 1980; Recktenwald and McConnell, 1981; Vist and Davis, 1990). The Lo state is characterized by fast lipid lateral diffusion, as in liquids, but tightly packed and ordered as are solids (London, 2005). It was widely thought that a

homogeneous Lo state might replace separate Ld and gel domains in cholesterol-rich membranes (London, 2002).

Formulation of the Lipid Raft Hypothesis

Interest in membrane domains was revived by studies in both cell and model membranes. Work on cells from van Meer and Simons in 1988 proposed formation of sphingolipid microdomains functioning in sorting of glycosphingolipids and glycosylphosphatidylinositol (GPI)-anchored proteins. It was envisioned that H-bonding induced microdomain formation and GPI-anchored protein association with glycosphingolipids (Simons and van Meer, 1988). It was proposed in another study that detergent resistance might be related to this sorting based on influenza hemagglutinin behavior (Skibbens *et al.*, 1989). The work of Brown and Rose in 1992 provided a breakthrough involving the isolation and characterization of detergent-resistant membranes (DRM) from cells (Brown and Rose, 1992). DRM membrane fractions were rich in sphingolipids, cholesterol, and GPI-anchored proteins, and their detergent resistance provided a potential method to readily isolate membrane domains. (In much earlier studies, a detergent insoluble membrane fraction was isolated from red blood cells by Steck and colleagues, and it was speculated that this might arise from lipid domains, but this study was not widely pursued (Yu *et al.*, 1973).) Brown, London, and colleagues found in 1994 that liposomes composed both of lipids found in DRM and of GPI-anchored proteins exhibited a high degree of sphingolipid, cholesterol, and GPI-anchored protein insolubility in TX-100, as did Lo state liposomes composed of just saturated lipids and cholesterol (Schroeder *et al.*, 1994). It was proposed that DRM arose from Lo-like domains in cells, and that the packing between saturated acyl chains and cholesterol drove both Lo-like domain formation in cells and the association of GPI-anchored proteins (which are anchored in membranes by saturated acyl chains) with such domains (Schroeder *et al.*, 1994). An important report by Ikonen and Simons in 1997 proposed a synthesis of many of the early studies, and popularized the name lipid rafts for cholesterol and sphingolipid-rich membrane domains (Simons and Ikonen, 1997). In addition, studies in 1996 and 1997 detected segregation of lipids into coexisting Ld and Lo domains in ternary mixtures of high- T_m lipid (including sphingomyelin (SM)), low- T_m lipid and cholesterol (Ahmed *et al.*, 1997; Silvius *et al.*, 1996) including that at physiologically relevant compositions of sphingolipid and cholesterol (Ahmed *et al.*, 1997). It was also found that detergent resistance only appeared in mixtures of SM, low- T_m phosphatidylcholine (PC), and cholesterol when SM levels were high enough that Lo domains had formed as defined spectroscopically, confirming Lo domain formation was linked to detergent insolubility (Ahmed *et al.*, 1997). Early reviews in 1997 and 1998 raised many of the issues surrounding domain properties that remain

of interest and concern today (Brown and London, 1997, 1998; Rietveld and Simons, 1998).

What Hold Rafts Together?

The tight packing of lipid molecules in an ordered state bilayer arises from both van der Waals interactions between acyl chains and polar headgroup interactions. Acyl chains with kinks due to *cis* double bonds, especially in the center of the chain, tend to resist tight packing, while long saturated acyl chains favor tight packing. However, a lipid having long saturated acyl chains does not guarantee that it has a strong tendency to form a tightly packed state with neighboring lipids because very large polar headgroups can inhibit tight acyl chain packing due to steric clashes with neighboring lipid headgroups. H-bonding between lipids may also contribute significantly to ordered domain formation (Li *et al.*, 2001; Ramstedt and Slotte, 1999). In this regard, it should be noted that H-bonding groups in the polar headgroup and upper acyl chain region (e.g., alpha OH groups on fatty acids) are in an environment significantly less polar than aqueous solution, and thus may H-bond to each other more strongly than they would in a fully aqueous environment.

Close packing between sterols and other membrane lipids has been detected via the so-called condensing effect of sterols, in which the lateral area taken up by lipids at a given lateral pressure in monolayers is less than additive when sterols are mixed with phospholipid or sphingolipids. Favorable van der Waals interactions are likely to help drive this close packing. In addition, close packing is enhanced by the hydrophobic effect. The small sterol OH group is insufficient to prevent contact of sterol hydrocarbon with water, and close packing allows the sterol to be hidden from water by the larger polar headgroups of neighboring membrane lipids. This is known as the 'umbrella effect' (Huang and Feigenson, 1999). Free ceramide also has a small polar group, and its packing with other sphingolipids is also thought to be partly driven by an umbrella effect (Bjorkqvist *et al.*, 2005; Megha and London, 2004; Zitzer *et al.*, 2001). In fact, when ceramide and sterol are both present in a mixture of high-T_m and low-T_m lipids, competition for sites under polar headgroups of other lipids can lead to displacement of sterols from ordered domains by ceramide (Bjorkqvist *et al.*, 2005; Megha and London, 2004), or a competition that affects binding to proteins (Zitzer *et al.*, 2001). Similar competition reflecting umbrella effects can be observed for several lipid-like molecules with small polar headgroups (Alanko *et al.*, 2005). It should be noted that ceramide-rich ordered membrane domains potentially exist in the gel state. Studies in yeast suggest gel-state domains can sometimes exist in cells (Aresta-Branco *et al.*, 2011).

The basis of raft formation in the cytosolic ('inner') lipid leaflet/monolayer of membranes is unclear. Membrane lipid asymmetry is such that there is little sphingolipid in the cytosolic leaflet of eukaryotic cell membranes. Furthermore, model membrane studies on vesicles roughly mimicking either the exofacial (outer) or cytosolic leaflets of plasma membranes suggest that the lipid composition of the outer leaflet supports spontaneous Lo domain formation while that of the inner leaflet does not (Silvius, 2003; Wang and Silvius, 2001).

Nevertheless, numerous lipid-anchored cytosolic peripheral proteins appear to be raft-associated (Melkonian *et al.*, 1999). How do rafts form in cytosolic leaflets? Our knowledge of lipid asymmetry is very incomplete. Perhaps some fraction of inner leaflet lipids are more saturated than generally thought, and thus can help form rafts. A different, long-held notion is that there is interleaflet coupling of lipid physical properties such that lipids organized into Lo domains in the exofacial leaflet might induce Lo state domains in the cytosolic leaflet. Studies using planar lipid bilayers (unsupported or cushioned supported) or lipid vesicles have shown that coupling can be detected in asymmetric model membranes when one leaflet contains SM or a saturated phospholipid and the other does not, and that acyl chain structure can modulate coupling (Chiantia and London, 2012; Collins and Keller, 2008; Wan *et al.*, 2008). However, the details of the relationship between lipid structure and coupling remain largely uncharacterized.

It should also be noted that membrane proteins could influence raft formation and interleaflet coupling. A protein that has a favorable affinity for rafts by definition contributes to the free energy of raft formation, and can even trigger raft formation in situations in which rafts formed by lipids alone would have a borderline stability (see below).

The Basis of Membrane Protein Interaction with Rafts

How and when proteins interact with lipid rafts is central to raft function. Localization of different proteins to rafts could increase protein-protein interaction, while segregation into different domains would prevent their interaction. In addition, the different physical properties of lipid domains (in terms of bilayer width and molecular packing) could support different conformational states for proteins in Lo and Ld domains.

Association of lipid-anchored peripheral proteins with Lo domains is likely to be driven by tight packing between the saturated acyl chains attached to proteins and the lipids in Lo domains (i.e., saturated acyl chain sphingolipids and cholesterol) (Schroeder *et al.*, 1994). Many palmitate and myristate anchored proteins that insert into the cytosolic leaflet of membranes have a significant affinity for rafts (Levental *et al.*, 2010; Melkonian *et al.*, 1999; Wang *et al.*, 2001). In contrast, proteins covalently anchored by hydrocarbon chains that pack poorly into ordered domains, such as unsaturated acyl chains and polyisoprene chains have low affinity for Lo domains (Lindwasser and Resh, 2002; Melkonian *et al.*, 1999; Moffett *et al.*, 2000; Wang *et al.*, 2001). Furthermore, GPI-anchored proteins from mutant cells producing GPI anchors with unsaturated acyl chains lose association with DRM, suggesting lower raft affinity (Maeda *et al.*, 2007).

Similar considerations apply to peripheral membrane proteins that associate with lipids non-covalently. For example, the strong affinity of cholera toxin for the sphingolipid ganglioside, ganglioside M1 (GM1), is clearly the source of the strong interaction of cholera toxin with ordered domains (Dietrich *et al.*, 2001), and cholera toxin trafficking in cells is significantly altered when GM1 is anchored by acyl chains that do interact well with rafts (Chinnappen *et al.*, 2012). Raft affinity arising from lipid binding appears to be true for other glycosphingolipid binding proteins (Kovbasnjuk *et al.*, 2001).

Nevertheless, it should be noted that the absolute affinity of saturated lipid anchors for ordered domains is not necessarily very strong (Wang *et al.*, 2001), and is likely to be dependent on the actual lipid composition of the ordered and disordered domains. In cells, oligomerization of lipid-anchored proteins that only weakly associate with rafts as monomers may greatly increase raft affinity (Brown and London, 1998; Harder *et al.*, 1998), but this has not been seen in every example studied (Keller *et al.*, 2013).

The origin of the affinity of integral transmembrane (TM) proteins for rafts has been more difficult to define. The rough surface of a TM helix should not favor very strong van der Waals interactions with lipid acyl chains, and should favor association of TM proteins with Ld domains. Indeed, several studies with single TM helices have reported a stronger affinity for Ld relative to Lo domains (Fastenberg *et al.*, 2003; Shogomori *et al.*, 2005; van Duyl *et al.*, 2002; Vidal and McIntosh, 2005). However, many TM proteins have been reported to associate with rafts. One explanation involves hydrophobic mismatch between TM segment length and the width of the hydrophobic portion of the bilayer. Lo domains generally form a wider lipid bilayer than do Ld domains (Figure 1). If hydrophobic TM segment length exceeds bilayer width, i.e., if there is positive mismatch, then unfavorable exposure of hydrophobic residues to a polar environment could occur. This unfavorable exposure would be less in Lo domains, thus favoring association with Lo domains. Early studies of mismatch may have been defeated by the ability of single TM helices to tilt in thin bilayers, thus circumventing positive mismatch. However, when perfringolysin O (PFO), a protein with multiple TM segments and which cannot avoid mismatch by tilting, was engineered to have longer and shorter lengths, it was found that mismatch can indeed control association with Lo domains (Lin and London, 2013a). It is possible that a reduced ability to tilt upon helix oligomerization may dynamically regulate TM protein raft affinity in some cases. In this regard it is noteworthy that oligomerized aquaporins show a stronger affinity for Lo domains than aquaporin monomers (Tong *et al.*, 2009). The effect of oligomerization upon raft affinity could also explain the common observation that raft affinity seems to increase upon activation of biological processes that assemble TM protein signal transduction complexes.

Another factor controlling the raft affinity of TM proteins may be attachment or binding to specific raft-associating lipids. In the case of the TM state of PFO, a protein which non-covalently binds to sterols, association with rafts is dependent upon the raft affinity of the sterol to which it is non-covalently bound (Lin and London, 2013b). Also, a functionally important interaction of the epidermal growth factor receptor (EGFR) protein with GM3 has been reported, and may impart affinity for ordered domains (Coskun *et al.*, 2011). Many TM proteins are covalently palmitoylated, and the packing of the palmitoyl group with Lo domains could contribute to raft affinity (Levental *et al.*, 2010). Indeed, blocked palmitoylation is associated with reduced raft affinity (Levental *et al.*, 2010). However, it has been difficult to confirm whether this involves a direct effect upon raft affinity, because in model membranes the effect of palmitoylation upon TM helix raft affinity has been small at most (Shogomori *et al.*, 2005; van Duyl *et al.*, 2002). This might reflect a loss of TM helix oligomerization in model membranes, or inexact mimicking of natural membranes (in these cases, plasma membrane) by the model membranes used. It is also possible that palmitoylation influences raft affinity indirectly, for example, by altering the depth of membrane immersion by a TM segment, and thus by factors such as hydrophobic mismatch or altering the subcellular membrane in which a protein is localized.

For extrapolating TM raft affinity from model to real membranes, lipid asymmetry may be an important issue. Suggestively, it has been reported that plasma membrane proteins exhibit an amino acid asymmetry in the inner and outer leaflet segments of TM sequences (Sharpe *et al.*, 2010). Thus, it is possible that differences in packing in the exofacial and cytofacial leaflets influences protein behavior, and thus raft affinity. The recent development of model membrane planar bilayers and vesicles with lipid asymmetry should make it possible to address this issue (Cheng *et al.*, 2009; Cheng and London, 2011; Chiantia and London, 2012; Collins and Keller, 2008; Hussain *et al.*, 2013; Lin and London, 2014; Wan *et al.*, 2008).

It should also be kept in mind that some membrane proteins may prefer to locate at the edges of lipid domains (Lin and London, 2013a). This would allow them to interact with proteins that are both inside and outside of domains.

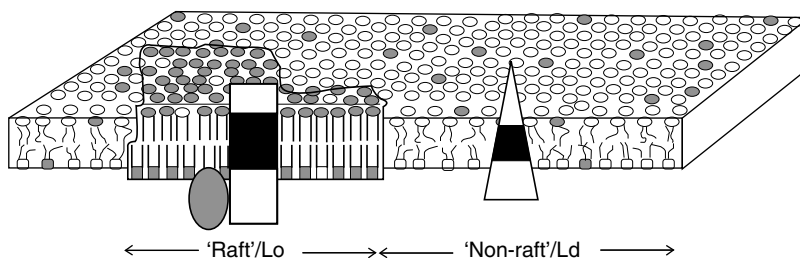


Figure 1 Schematic illustration of coexisting raft and non-raft domains. Raft-favoring lipids are shown with gray polar headgroups, and are enriched in raft domain. Note raft lipids are more tightly packed and ordered and tend to form a bilayer with a thicker width than the more disordered, loosely packed non-raft domain lipids. Raft-associating proteins are either saturated acyl chain lipid-anchored (oval) or are transmembrane (rectangle) with hydrophobic amino acid surfaces favoring raft association, for example, having a wide hydrophobic surface (black band) that matches the width of the hydrophobic portion of the raft domain lipids. Non-raft protein (triangle) has a hydrophobic surface that matches thinner width of non-raft domain lipids.

Lipid Domain Size

Under most conditions, lipid domains are submicroscopic. Attempts to measure domain size in cells using methods such as single-particle tracking, various lateral diffusion measurements, and super-resolution microscopy have not provided a single accepted value for domain size but have led to the thinking that rafts in cells are generally very small, perhaps sometimes as small as a few nanometers (Anderson and Jacobson, 2002; Brown and London, 1998). However, cell membrane domains rich in a particular lipid can sometimes be large enough (micron sized) to visualize by ordinary microscopy. These may represent large domains, or just regions richer in submicroscopic raft domains than the remainder of a membrane. In addition, because membranes are dynamic, a particular region in a membrane may simply be transiently enriched in a particular lipid because of local metabolic processes such as localized synthesis or delivery from internal organelles, rather than possess stable domains. Regions of spectroscopically identified increased membrane order that are large enough to detect by ordinary microscopy are more likely to be indicative of the presence of Lo domains, especially when raft-associated molecules co-localize with such regions (Gaus *et al.*, 2006a). This may again represent a region rich in very small raft domains, rather than a single raft.

Studies of domain size in model membranes are helpful for studying control of domain size in cells. In vesicles composed of mixtures of high-T_m saturated lipids, low-T_m unsaturated phospholipids, and cholesterol, multimicron size coexisting Lo and Ld domains can often be observed by microscopy (Hammond *et al.*, 2005; Veatch and Keller, 2003). However, lipid domain size in model membranes is highly lipid structure and temperature dependent, and can be as small as a few nanometer scale (Brown *et al.*, 2007; Pathak and London, 2011; Petruzielo *et al.*, 2013). Detecting separation of lipids into different domains in binary mixtures of one lipid and cholesterol has been more difficult to detect than in ternary mixtures of high-T_m lipid, low-T_m lipid, and cholesterol, although nuclear magnetic resonance (NMR), electron spin resonance (ESR) imply that domains form in binary mixtures (Sankaram and Thompson, 1990; Vist and Davis, 1990). Domains in binary mixtures may be extremely small or highly transient.

Why nanodomains form is not well-understood. If the energy of lipid demixing is favorable, small domains should coalesce into a single large domain because this would minimize the number of lipids at the energetically unfavorable boundaries between domains, at which lipids in different domains are in contact. The 'critical fluctuation' model proposes that tiny domains only exist at compositions and temperatures just above that at which phase separation is stable, i.e., just above a critical point, and then decrease in size as temperature is further increased (Veatch *et al.*, 2007, 2008). Consistent with this, nanodomain size decreases as temperature increases (Pathak and London, 2011). However, in some ternary lipid mixtures nanodomains predominate under a wide range of lipid compositions (Feigenson and Buboltz, 2001; Petruzielo *et al.*, 2013). Also, they can behave like miniature true phases, as judged by their different affinities for various fluorescent probes that are similar to those in large

scale phases (Pathak and London, 2011). It is possible that there is some interaction counteracting the tendency of domains to coalesce. A long standing idea has been that some lipids may have favorable energetic interactions at domain edges, which would favor small domains, which have a large circumference to area ratio. However, it is not clear to what extent this is important (Hassan-Zadeh *et al.*, 2014; Heberle *et al.*, 2013). Another factor likely to influence size is the extent of mismatch between the widths of the bilayer in different domains. The smaller the mismatch, the smaller the energetic penalty (line tension) at domain boundaries, which would lead to a smaller domain size (Heberle *et al.*, 2013). In contrast, clustering of raft preferring molecules by crosslinking tends to increase domain size or even trigger phase separation (Hammond *et al.*, 2005; Harder *et al.*, 1998; Lingwood *et al.*, 2008). Other possible biophysical influences on domain size have been proposed (Amazon *et al.*, 2013; Shlomovitz and Schick, 2013).

Raft size is an important issue when considering how they may influence protein-protein interactions. Unlike the situation in large raft domains, very small raft domains with only a single protein molecule could not enhance interactions between two different raft-associated proteins. However, they might enhance interaction between proteins in different domains, if proteins in different domains can approach close enough to establish contact. Raft size could also affect protein-protein interactions involving proteins that prefer to locate at the boundary between raft and non-raft domains (edges may have an intermediate bilayer width relative to Lo and Ld domains, and can attract TM proteins with a matched hydrophobic width (Lin and London, 2013a)). In this regard it is noteworthy, that some TM proteins surrounded by a layer or two of Ld lipids could form tiny Ld nanodomain inclusions within a raft domain (Almeida *et al.*, 2005; London, 2005; Nelson *et al.*, 2010; Polozova and Litman, 2000). The ability of Ld state lipids to stretch could allow formation of Ld inclusion thicker than bulk Ld domains, and even matching the surrounding Lo domains in thickness (Nelson *et al.*, 2010).

In cells, dynamic changes in lipid composition must also be considered. It is unlikely that this could prevent the formation of small domains, which do not need large-scale movements of lipids, but could limit the establishment of equilibrium conditions that drive the coalescence of small domains into larger ones.

Little is known about temporal persistence of domains. Large domains are clearly long-lived. The situation for nanodomains is less clear. Lo nanodomains can clearly persist long enough to have partition properties (i.e., affinities for other molecules) that are similar to large Lo phases. In addition, proteins that strongly associate with Lo or Ld domains may not only nucleate domain formation, but lengthen domain persistence. This is an important issue because domains persisting for only milliseconds might not be able to regulate protein function.

Detecting Rafts in Model Membranes

Formation of coexisting ordered and disordered domains in model membrane vesicles (liposomes) is readily detected, and

large domains can be readily visualized by microscopy for a wide variety of lipid mixtures (Korlach *et al.*, 1999). Submicroscopic domains can be detected by a calorimetry and a wide variety of spectroscopic methods (Heberle *et al.*, 2010; Pathak and London, 2011; Petruzielo *et al.*, 2013). However, an important consideration is that different methods have different special resolutions (Petruzielo *et al.*, 2013). When domains are not detected by one spectroscopic method, it does not prove they do not exist; they may just be below the resolution limit of that method. The fact that a liquid–liquid phase separation has formed is readily ascertained from domain shapes (rounded) (Samsonov *et al.*, 2001) and properties such as lateral diffusion, which is very slow in the solid-like gel states relative to Ld or Lo states (Almeida *et al.*, 1992; Rubenstein *et al.*, 1979). Based on such measurements, phase diagrams, which give the lipid composition and abundance of each type of domain as a function of overall lipid composition and temperature, have been formulated for several raft-forming ternary lipid mixtures containing cholesterol (Konyakhina *et al.*, 2013; Petruzielo *et al.*, 2013). Sometimes, different phase diagrams have been proposed for identical lipid mixtures by different groups, probably because methods with different special resolutions have been used.

Detecting Rafts in Natural Membranes: Giant Plasma Membrane Vesicles

In contrast to the situation in model membranes, rafts have been difficult to visualize in cells. An important advance in the field has been the studies of rafts in natural giant plasma membrane vesicles, giant plasma membrane vesicles (GPMV) (Baumgart *et al.*, 2007). These vesicles and closely related plasma membrane spheres (PMS), can be prepared from cell plasma membranes by methods that involve mild-to-moderate treatments (Baumgart *et al.*, 2007; Lingwood *et al.*, 2008). The vesicles are large enough for microscopy, and maintain natural membrane complexity in terms of lipid and protein composition, although cytoskeletal connections are lost, lipid composition mildly perturbed, and lipid asymmetry at least somewhat compromised (Keller *et al.*, 2009). At lower temperatures classical large Lo domains can be easily visualized in these vesicles by microscopy. This shows that the heterogeneity induced by the mere presence of hundreds of different lipids and proteins is insufficient to block lipid domain formation. In addition, in many cases the association of putative raft-associated proteins with the Lo domains in GPMV parallels that in simple giant unilamellar vesicle (GUV) and that defined by other methods (e.g., detergent insolubility) (Keller *et al.*, 2009; Sengupta *et al.*, 2008). However, the absolute affinity for Lo state domains is not identical in GPMV and simple GUV (Keller *et al.*, 2009). It is not clear why large domains are seen much more readily in GPMV than cells, but some factor connected to a loss of cytoskeletal connections is one possibility. Although domain size in GPMV shrinks as temperature is increased, it is likely that submicroscopic domains persist to physiological temperatures. In addition, formation of large Lo domains can be readily triggered in PMS at physiological temperatures by clustering events, for example, induced by addition of cholera toxin (Lingwood *et al.*, 2008).

Using Detergents to Detect Rafts in Cells: Rationale and Limitations

The first method to detect rafts in cells was detergent insolubility (Brown and Rose, 1992). Detergents (especially Triton X-100) have been widely used to isolate DRM fractions that might arise from lipid rafts. The rationale is that loosely packed Ld state membranes dissolves in detergent while tightly packed ordered membranes (gel and Lo) tend to be insoluble in detergent (London and Brown, 2000; Schroeder *et al.*, 1994). This is easily explained in terms of the stronger lipid–lipid interactions in tightly packed domains. Spectroscopic studies studying the association of protein and lipid with ordered domains in model membranes show a marked correlation with their affinity for DRM (Dietrich *et al.*, 2001; Wang *et al.*, 2001). However, binding of detergent to membranes changes membrane structure and thus can perturb the lipid composition and properties of the lipid domains. For this method to work properly the detergent has to dissolve the Ld domains before membrane rearrangements can occur (London and Brown, 2000; Nelson *et al.*, 2010; Schroeder *et al.*, 1994).

Many studies in model membranes have shown that upon addition of detergent to membranes with coexisting ordered and Ld domains there is selective solubilization of Ld domains, which suggests little perturbation of domain formation by detergent. However, interactions with detergent can be lipid composition specific, and when lipids form small membrane domains (e.g., in SM, 1-palmitoyl 2-oleoyl phosphatidylcholine (POPC), cholesterol mixtures) they might be particularly susceptible to perturbation by detergent. It has been reported that Triton X-100 can increase immiscibility of POPC and SM in cholesterol-containing membranes, and thus induce domain segregation under conditions in which it would not form otherwise (Heerklotz, 2002; Heerklotz *et al.*, 2003). However, a subsequent study reported that in SM, POPC, cholesterol mixtures TX-100 does NOT increase immiscibility or induce domain formation, but rather induces coalescence of pre-existing ordered nanodomains (Pathak and London, 2011). Coalescence into large domains is of less concern than artifactual domain formation, and is a necessary property for formation of DRM which can be readily isolated.

Nevertheless, concerns about the use of detergents remain. In some cases, but not all, DRM association can reflect rearrangements after solubilization (London and Brown, 2000; Nelson *et al.*, 2010; Schroeder *et al.*, 1994). Appropriate controls are necessary (Nelson *et al.*, 2010; Schroeder *et al.*, 1994). Another complication is under some experimental conditions (e.g., too low a detergent to lipid ratio), detergents like Triton X-100 can incompletely dissolve Ld domains, and can sometimes dissolve ordered domains, especially those with low or no cholesterol (London and Brown, 2000). As a result, interpreting differences obtained when diverse detergents or conditions are used is problematic. Finally, solubilization is usually carried out at 4 °C, and low temperature can certainly induce formation of ordered domains that would not exist at physiological temperature, for example, 37 °C. In this regard, methods to carry out solubilization at 37 °C are promising (Drevot *et al.*, 2002; Morris *et al.*, 2011). Methods to fragment membrane into raft and non-raft domains without using detergent are also sometimes used (Macdonald and Pike, 2005),

perhaps aided by domain coalescence under conditions of membrane tension (Ayuyan and Cohen, 2008) although the efficiency of these methods in terms of separating Lo and Ld domains is often uncertain. It has also been found that physical fragmentation of membranes after adding beads binding a protein with a strong raft affinity isolate raft-type lipids but control beads binding a non-raft protein do not (Zech *et al.*, 2009).

To summarize, DRM studies represent a useful starting point in studies of lipid raft in cells, and can provide an idea about the raft affinity of different molecules, and changes in raft affinity in different physiological states. However, they cannot be relied upon on their own to define when rafts are present in cells or the exact extent to which a specific molecule is or is not raft-associated under physiological conditions. Other methods must be used in conjunction with detergent solubilization.

Detecting Rafts in Cells: Microscopy and Spectroscopy

As noted above, in intact cells, apparent membrane domains, which may be single domains or regions rich in a large number of nanodomains, can sometimes be visualized by microscopy, especially when raft-associating molecules are clustered. It is important to note that a portion of a cell having more membrane lipid per unit volume than other regions (e.g., due to submicroscopic membrane irregularities) will show an increased local concentration of all membrane molecules that can be mistaken for co-localization of proteins in rafts (Glebov and Nichols, 2004). It is important to show depletion of non-raft molecules in a region enriched in raft-forming molecules before concluding that there is co-localization of a protein with a raft.

A variety of advanced microscopy methods have been used to detect rafts in cells. Membrane-bound fluorescent probes (e.g., Laurdan and di-4-ANEPPDHQ) that have distinct spectroscopic signal when associated with regions having elevated membrane order have been widely used to detect domains (Dietrich *et al.*, 2001; Gaus *et al.*, 2003, 2005, 2006b; Harder *et al.*, 2007; Owen *et al.*, 2006, 2007; Zech *et al.*, 2009) and new probes are being steadily introduced (Klymchenko and Kreder, 2014). Because of the small size of rafts, super-resolution microscopy has the promise of more facile raft detection (Honigsmann *et al.*, 2013, 2014; Sezgin *et al.*, 2012). However, raft size may often be below the resolution level achievable using this method. Microscopy methods based on mass spectrometric analysis of lipid composition have recently been used to detect lipid heterogeneity in cell membranes (Frisz *et al.*, 2013), but again resolution may be an issue. Single-particle tracking and lateral diffusion methods can be used to detect domains via their effect on reducing or confining diffusion (Lenne *et al.*, 2006; Pinaud *et al.*, 2009; Suzuki *et al.*, 2012). The fact that other membrane properties may affect diffusion can complicate interpretation of such experiments. Finally, fluorescence resonance energy transfer (FRET) methods suitable for use in cells have also been able to detect domain formation (Sengupta *et al.*, 2007; Sohn *et al.*, 2008). Depending on donor and acceptor used, FRET should be able

to detect all but the smallest of nanodomains (Pathak and London, 2011). However, FRET requires that the donor and acceptors have suitable affinities for the domain type in question.

Despite all these issues, some recent microscopy studies provide relatively compelling evidence for lipid rafts in living cells. Large-scale domain separation in yeast vacuoles which exists in the absence of any clustering of raft-markers has been observed (Toulmay and Prinz, 2013). The domains appear almost identical in scale to those seen in GUV with Lo domains, with the rounded shapes characteristic of liquid-liquid phase separations as seen in membranes with coexisting large Ld and Lo phases. In the bacterium *Borrelia burgdorferi*, raft-like domains have been observed that are large enough to easily visualize using transmission electron microscopy (LaRocca *et al.*, 2010, 2013). These domains can also be detected in living *Borrelia* using FRET, and detected biochemically by isolation of DRM (LaRocca *et al.*, 2010, 2013). They appear to be equivalent to eukaryotic lipid rafts based upon their selective affinity for saturated vs. unsaturated lipids and because the presence of raft-forming sterols are necessary and sufficient for their formation (see below).

Detecting Rafts in Cells Using Sterol Modification

One common method to identify likely raft-dependent functions is to define biological processes inhibited upon (partial) extraction of cholesterol from cells using cyclodextrins (Zidovetzki and Levitan, 2007). Restoration of function by replenishment of cellular cholesterol using cholesterol-cyclodextrin complexes is then used to demonstrate that it is extraction of cholesterol that altered function. These studies have the limitation that sterol extraction, which changes overall mass balance in the membrane, changes more than just raft levels. It is also claimed that the loss of DRM upon sterol extraction demonstrates that rafts have been destroyed by cyclodextrin treatment, but this is not correct. Ordered domains with low levels of sterol or no sterol can sometimes dissolve in Triton X-100 (London and Brown, 2000).

A more reliable, albeit more complex, approach involves sterol substitution. The abilities of various sterols and steroids to stabilize or destabilize raft formation have been defined (Megha *et al.*, 2006; Wang *et al.*, 2004; Wenz and Barrantes, 2003; Xu and London, 2000; Xu *et al.*, 2001). In sterol substitution partial extraction of cellular sterols is followed by replenishment with a diverse series of exogenous sterols with different raft-forming abilities. Overall lipid mass balance can be (roughly) maintained in these experiments. When there is a strong correlation between a biological process and the raft-forming ability of sterols with diverse chemical structures, then it becomes likely that the biological process is raft-dependent. To be more precise, if every raft-stabilizing sterol supports a function while every raft-destabilizing sterol does not, and there are no other properties that distinguish function-supporting from the function-inhibiting sterols, then it can be inferred that a sterol having raft-stabilizing properties is both necessary (i.e., function is not supported by raft-destabilizing sterols) and sufficient (i.e., function is supported by all raft-stabilizing sterols) for the function. This pattern establishes

raft-dependence as the favored working hypothesis for the process under study. Perhaps just as important, by identifying biological processes that do not show this pattern, sterol substitution is an efficient approach to define processes unlikely to be raft-dependent (Singh *et al.*, 2009). A similar strategy is to introduce other molecules that should stabilize or destabilize rafts into membranes, such as polyunsaturated fatty acids or ceramides (Gidwani *et al.*, 2003; Zech *et al.*, 2009). In all such studies additional issues to consider include rapid metabolic transformations of the added molecules and signaling events that such molecules may trigger.

Several sterol substitution studies have been carried out. Sterol substitution studies in the bacterium *B. burgdorferi* show raft properties were necessary and sufficient for domain formation. These studies were especially informative because ordered lipid domains can readily be detected in this organism by electron microscopy and FRET (LaRocca *et al.*, 2010, 2013). In addition, sterols having raft-stabilizing properties were necessary and sufficient to maintain membrane integrity in *B. burgdorferi* (LaRocca *et al.*, 2013). In T cells, 7-ketocholesterol has been substituted for cholesterol. The raft-forming abilities of 7-ketocholesterol may only be slightly weaker than those of cholesterol (Wang *et al.*, 2004), but the difference appears to be sufficient to disrupt raft formation and T cell signaling (Rentero *et al.*, 2008). In the case of HIV, it has been shown that substitution of raft-destabilizing sterols into virions inhibits infectivity while substitution of raft-stabilizing sterols does not (Campbell *et al.*, 2004).

Biological Functions of Raft Domains in Cells

It has been reported that lipid rafts are associated with almost every process involving cell membranes. Given this, the diverse criteria that have been used to assign whether a particular function is raft dependent and the disagreements about what methods prove that rafts exist under a particular set of conditions, an extensive discussion of claimed raft functions is beyond the scope of this article.

Although there are few, if any, membrane functions, in which every study supports a role for lipid rafts, there are several topics in which a preponderance of studies indicate that ordered lipid domain formation is likely to play an important functional role. One is in sorting of molecules to different membranes. Fluorescent probes and lipid-anchored proteins can be sorted to a degree or rate that is dependent upon whether they are attached to Ld or Lo associating lipids (Mukherjee *et al.*, 1999; Refaei *et al.*, 2011). In this regard, the observation that sorting vesicles targeted to plasma membrane are richer in Lo lipids than the membrane from they arise may be relevant (Surma *et al.*, 2011). The difference in Lo and Ld domain bendability is likely to be involved in membrane shape-dependent domain formation. In this regard, studies in model membrane vesicles have shown that formation of tubules from large vesicles, a common event in sorting, can induce domain formation/segregation (Roux *et al.*, 2005).

The second important function is signaling, especially in the immune system. Studies have implicated rafts having a role in organizing immune system proteins in response to stimulation in mast, T and B cells. In T cells, a wide variety of

studies have been carried out. Assembly of activated T cell receptor and associated proteins into rafts upon activation has been detected and assessed by isolation techniques without use of detergent, association with fluorescently detected ordered regions of the lipid bilayer, sterol alterations, and temperature perturbation methods (Harder and Sangani, 2009; Harder, 2012; Magee *et al.*, 2005; Mahammad *et al.*, 2010; Rentero *et al.*, 2008; Zech *et al.*, 2009), although it has also been found that clustering of both raft and non-raft lipids can induce T cell signaling (Wang *et al.*, 2005). In B cells both detergent insolubility and FRET methods have detected raft association, and FRET shows that raft association is an intermediate event in activation (Liu *et al.*, 2010). In mast cells, clustering in rafts may trigger signaling, and be involved in segregation of kinases and phosphatases (Holowka *et al.*, 2005). Rafts may also be involved in many other signaling processes (Simons and Toomre, 2000).

Rafts seem to have a role in certain types of infection. HIV and influenza infection have been most extensively studied in this regard (Rossman *et al.*, 2010; Veit *et al.*, 2013; Waheed and Freed, 2009). Raft interactions during infection or viral assembly can be important. However, differences in viral protein behavior in cells and model membranes complicate our understanding of the molecular role of rafts in these cases.

Two other areas in which the participation of rafts has been extensively investigated are the formation of amyloid from APP protein in studies of Alzheimer's disease and in sperm cell membrane structure and maturation (Nixon and Aitken, 2009; Vetrivel and Thinakaran, 2010). It is also possible that altered rafts are involved in diseases of cholesterol biosynthesis, such as Smith-Lemli-Opitz syndrome (Megha *et al.*, 2006; Nowaczyk and Irons, 2012).

A problem in defining the role of rafts in cells is that studies of protein association with rafts in model membranes and cells do not always agree (Keller *et al.*, 2013; Nikolaus *et al.*, 2010; Veit *et al.*, 2013). This may reflect the fact that model membrane vesicles are not always good mimics of cell membranes. Model membranes lack the high protein concentration and cytoskeletal proteins found in natural membranes. Also, they tend to have simple lipid compositions and generally use cholesterol concentrations lower than that in plasma membranes. However, the most serious problem may be that most model membranes lack the lipid asymmetry characteristic of natural membranes. Recent advances in forming highly asymmetric planar membranes and vesicles that closely mimic cell membranes (Cheng *et al.*, 2009; Cheng and London, 2011; Hussain *et al.*, 2013; Lin and London, 2014; Wan *et al.*, 2008) suggest future studies in model membranes may reveal domain behavior even closer to that in cells.

Acknowledgment

This work was supported by NIH grant GM099892.

See also: Molecular Principles, Components, Technology, and Concepts: Lipids: Lipid Signaling. Molecular Principles,

Components, Technology, and Concepts: Membranes: Composition, Physical Properties, and Curvature

References

- Ahmed, S.N., Brown, D.A., London, E., 1997. On the origin of sphingolipid/cholesterol-rich detergent-insoluble cell membranes: Physiological concentrations of cholesterol and sphingolipid induce formation of a detergent-insoluble, liquid-ordered lipid phase in model membranes. *Biochemistry* 36, 10944–10953.
- Alanko, S.M., Halling, K.K., Maunula, S., Slotte, J.P., Ramstedt, B., 2005. Displacement of sterols from sterol/sphingomyelin domains in fluid bilayer membranes by competing molecules. *Biochimica et Biophysica Acta* 1715, 111–121.
- Almeida, P.F., Pokorny, A., Hinderliter, A., 2005. Thermodynamics of membrane domains. *Biochimica et Biophysica Acta* 1720, 1–13.
- Almeida, P.F., Vaz, W.L., Thompson, T.E., 1992. Lateral diffusion in the liquid phases of dimyristoylphosphatidylcholine/cholesterol lipid bilayers: A free volume analysis. *Biochemistry* 31, 6739–6747.
- Amazon, J.J., Goh, S.L., Feigenson, G.W., 2013. Competition between line tension and curvature stabilizes modulated phase patterns on the surface of giant unilamellar vesicles: A simulation study. *Physical Review E* 87, 022708.
- Anderson, R.G., Jacobson, K., 2002. A role for lipid shells in targeting proteins to caveolae, rafts, and other lipid domains. *Science* 296, 1821–1825.
- Aresta-Branco, F., Cordeiro, A.M., Marinho, H.S., *et al.*, 2011. Gel domains in the plasma membrane of *Saccharomyces cerevisiae*: Highly ordered, ergosterol-free, and sphingolipid-enriched lipid rafts. *Journal of Biological Chemistry* 286, 5043–5054.
- Ayuyan, A.G., Cohen, F.S., 2008. Raft composition at physiological temperature and pH in the absence of detergents. *Biophysical Journal* 94, 2654–2666.
- Baumgart, T., Hammond, A.T., Sengupta, P., *et al.*, 2007. Large-scale fluid/fluid phase separation of proteins and lipids in giant plasma membrane vesicles. *Proceedings of the National Academy of Sciences of the United States of America* 104, 3165–3170.
- Bjorkqvist, Y.J., Nyholm, T.K., Slotte, J.P., Ramstedt, B., 2005. Domain formation and stability in complex lipid bilayers as reported by cholesterol. *Biophysical Journal* 88, 4054–4063.
- Brown, A.C., Towles, K.B., Wrenn, S.P., 2007. Measuring raft size as a function of membrane composition in PC-based systems: Part II—ternary systems. *Langmuir* 23, 11188–11196.
- Brown, D.A., London, E., 1997. Structure of detergent-resistant membrane domains: Does phase separation occur in biological membranes? *Biochemical and Biophysical Research Communications* 240, 1–7.
- Brown, D.A., London, E., 1998. Structure and origin of ordered lipid domains in biological membranes. *Journal of Membrane Biology* 164, 103–114.
- Brown, D.A., Rose, J.K., 1992. Sorting of GPI-anchored proteins to glycolipid-enriched membrane subdomains during transport to the apical cell surface. *Cell* 68, 533–544.
- Campbell, S., Gaus, K., Bittman, R., *et al.*, 2004. The raft-promoting property of virion-associated cholesterol, but not the presence of virion-associated Brij 98 rafts, is a determinant of human immunodeficiency virus type 1 infectivity. *Journal of Virology* 78, 10556–10565.
- Cheng, H.T., London, E., 2011. Preparation and properties of asymmetric large unilamellar vesicles: Interleaflet coupling in asymmetric vesicles is dependent on temperature but not curvature. *Biophysical Journal* 100, 2671–2678.
- Cheng, H.T., Megha, London, E., 2009. Preparation and properties of asymmetric vesicles that mimic cell membranes: Effect upon lipid raft formation and transmembrane helix orientation. *Journal of Biological Chemistry* 284, 6079–6092.
- Chiantia, S., London, E., 2012. Acyl chain length and saturation modulate interleaflet coupling in asymmetric bilayers: Effects on dynamics and structural order. *Biophysical Journal* 103, 2311–2319.
- Chinnappen, D.J., Hsieh, W.T., te Welscher, Y.M., *et al.*, 2012. Lipid sorting by ceramide structure from plasma membrane to ER for the cholera toxin receptor ganglioside GM1. *Developmental Cell* 23, 573–586.
- Collins, M.D., Keller, S.L., 2008. Tuning lipid mixtures to induce or suppress domain formation across leaflets of unsupported asymmetric bilayers. *Proceedings of the National Academy of Sciences of the United States of America* 105, 124–128.
- Coskun, U., Grzybek, M., Drechsel, D., Simons, K., 2011. Regulation of human EGF receptor by lipids. *Proceedings of the National Academy of Sciences of the United States of America* 108, 9044–9048.
- Dietrich, C., Bagatolli, L.A., Volovyk, Z.N., *et al.*, 2001. Lipid rafts reconstituted in model membranes. *Biophysical Journal* 80, 1417–1428.
- Dietrich, C., Volovyk, Z.N., Levi, M., Thompson, N.L., Jacobson, K., 2001. Partitioning of Thy-1, GM1, and cross-linked phospholipid analogs into lipid rafts reconstituted in supported model membrane monolayers. *Proceedings of the National Academy of Sciences of the United States of America* 98, 10642–10647.
- Drevot, P., Langlet, C., Guo, X.J., *et al.*, 2002. TCR signal initiation machinery is pre-assembled and activated in a subset of membrane rafts. *EMBO Journal* 21, 1899–1908.
- van Duyl, B.Y., Rijkers, D.T., de Kruijff, B., Killian, J.A., 2002. Influence of hydrophobic mismatch and palmitoylation on the association of transmembrane alpha-helical peptides with detergent-resistant membranes. *FEBS Letters* 523, 79–84.
- Fastenberg, M.E., Shogomori, H., Xu, X., Brown, D.A., London, E., 2003. Exclusion of a transmembrane-type peptide from ordered-lipid domains (rafts) detected by fluorescence quenching: Extension of quenching analysis to account for the effects of domain size and domain boundaries. *Biochemistry* 42, 12376–12390.
- Feigenson, G.W., Buboltz, J.T., 2001. Ternary phase diagram of dipalmitoyl-PC/dilauroyl-PC/cholesterol: Nanoscopic domain formation driven by cholesterol. *Biophysical Journal* 80, 2775–2788.
- Frisz, J.F., Klitzing, H.A., Lou, K., *et al.*, 2013. Sphingolipid domains in the plasma membranes of fibroblasts are not enriched with cholesterol. *Journal of Biological Chemistry* 288, 16855–16861.
- Gaus, K., Chklovskaya, E., Fazekas de St Groth, B., Jessup, W., Harder, T., 2005. Condensation of the plasma membrane at the site of T lymphocyte activation. *Journal of Cell Biology* 171, 121–131.
- Gaus, K., Gratton, E., Kable, E.P., *et al.*, 2003. Visualizing lipid structure and raft domains in living cells with two-photon microscopy. *Proceedings of the National Academy of Sciences of the United States of America* 100, 15554–15559.
- Gaus, K., Le Lay, S., Balasubramanian, N., Schwartz, M.A., 2006a. Integrin-mediated adhesion regulates membrane order. *Journal of Cell Biology* 174, 725–734.
- Gaus, K., Zech, T., Harder, T., 2006b. Visualizing membrane microdomains by Laurdan 2-photon microscopy. *Molecular Membrane Biology* 23, 41–48.
- Gidwani, A., Brown, H.A., Holowka, D., Baird, B., 2003. Disruption of lipid order by short-chain ceramides correlates with inhibition of phospholipase D and downstream signaling by FcεRI. *Journal of Cell Science* 116, 3177–3187.
- Glebov, O.O., Nichols, B.J., 2004. Lipid raft proteins have a random distribution during localized activation of the T-cell receptor. *Nature Cell Biology* 6, 238–243.
- Hammond, A.T., Heberle, F.A., Baumgart, T., *et al.*, 2005. Crosslinking a lipid raft component triggers liquid ordered-liquid disordered phase separation in model plasma membranes. *Proceedings of the National Academy of Sciences of the United States of America* 102, 6320–6325.
- Harder, T., 2012. The T cell plasma membrane lipid bilayer stages TCR-proximal signaling events. *Frontiers in Immunology* 3, 50.
- Harder, T., Rentero, C., Zech, T., Gaus, K., 2007. Plasma membrane segregation during T cell activation: Probing the order of domains. *Current Opinion in Immunology* 19, 470–475.
- Harder, T., Sangani, D., 2009. Plasma membrane rafts engaged in T cell signalling: New developments in an old concept. *Cell Communication and Signaling* 7, 21.
- Harder, T., Scheiffele, P., Verkade, P., Simons, K., 1998. Lipid domain structure of the plasma membrane revealed by patching of membrane components. *Journal of Cell Biology* 141, 929–942.
- Hassan-Zadeh, E., Baykal-Caglar, E., Alwarawrah, M., Huang, J., 2014. Complex roles of hybrid lipids in the composition, order, and size of lipid membrane domains. *Langmuir* 30, 1361–1369.
- Heberle, F.A., Doktorova, M., Goh, S.L., *et al.*, 2013. Hybrid and nonhybrid lipids exert common effects on membrane raft size and morphology. *Journal of the American Chemical Society* 135, 14932–14935.
- Heberle, F.A., Petruziello, R.S., Pan, J., *et al.*, 2013. Bilayer thickness mismatch controls domain size in model membranes. *Journal of the American Chemical Society* 135, 6853–6859.
- Heberle, F.A., Wu, J., Goh, S.L., Petruziello, R.S., Feigenson, G.W., 2010. Comparison of three ternary lipid bilayer mixtures: FRET and ESR reveal nanodomains. *Biophysical Journal* 99, 3309–3318.
- Heerklotz, H., 2002. Triton promotes domain formation in lipid raft mixtures. *Biophysical Journal* 83, 2693–2701.
- Heerklotz, H., Szadkowska, H., Anderson, T., Seelig, J., 2003. The sensitivity of lipid domains to small perturbations demonstrated by the effect of Triton. *Journal of Molecular Biology* 329, 793–799.
- Holowka, D., Gosse, J.A., Hammond, A.T., *et al.*, 2005. Lipid segregation and IgE receptor signaling: A decade of progress. *Biochimica et Biophysica Acta* 1746, 252–259.

- Honigsmann, A., Mueller, V., Hell, S.W., Eggeling, C., 2013. STED microscopy detects and quantifies lipid phase separation in lipid membranes using a new far-red emitting fluorescent phosphoglycerolipid analogue. *Faraday Discussions* 161, 77–89. discussion 113–150.
- Honigsmann, A., Sadeghi, S., Keller, J., *et al.*, 2014. A lipid bound actin meshwork organizes liquid phase separation in model membranes. *Elife* 3, e01671.
- Huang, J., Feigenson, G.W., 1999. A microscopic interaction model of maximum solubility of cholesterol in lipid bilayers. *Biophysical Journal* 76, 2142–2157.
- Hussain, N.F., Siegel, A.P., Ge, Y., Jordan, R., Naumann, C.A., 2013. Bilayer asymmetry influences integrin sequestering in raft-mimicking lipid mixtures. *Biophysical Journal* 104, 2212–2221.
- Ipsen, J.H., Karlstrom, G., Mouritsen, O.G., Wennerstrom, H., Zuckermann, M.J., 1987. Phase equilibria in the phosphatidylcholine-cholesterol system. *Biochimica et Biophysica Acta* 905, 162–172.
- Keller, H., Krausslich, H.G., Schwille, P., 2013. Multimerizable HIV Gag derivative binds to the liquid-disordered phase in model membranes. *Cellular Microbiology* 15, 237–247.
- Keller, H., Lorzate, M., Schwille, P., 2009. PI(4,5)P₂ degradation promotes the formation of cytoskeleton-free model membrane systems. *ChemPhysChem* 10, 2805–2812.
- Klymchenko, A.S., Kreder, R., 2014. Fluorescent probes for lipid rafts: From model membranes to living cells. *Chemistry & Biology* 21, 97–113.
- Konyakhina, T.M., Wu, J., Mastroianni, J.D., Heberle, F.A., Feigenson, G.W., 2013. Phase diagram of a 4-component lipid mixture: DSPC/DOPC/POPC/cholesterol. *Biochimica et Biophysica Acta* 1828, 2204–2214.
- Korlach, J., Schwille, P., Webb, W.W., Feigenson, G.W., 1999. Characterization of lipid bilayer phases by confocal microscopy and fluorescence correlation spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* 96, 8461–8466.
- Kovbasnjuk, O., Edidin, M., Donowitz, M., 2001. Role of lipid rafts in Shiga toxin 1 interaction with the apical surface of Caco-2 cells. *Journal of Cell Science* 114, 4025–4031.
- LaRocca, T.J., Crowley, J.T., Cusack, B.J., *et al.*, 2010. Cholesterol lipids of *Borrelia burgdorferi* form lipid rafts and are required for the bactericidal activity of a complement-independent antibody. *Cell Host & Microbe* 8, 331–342.
- LaRocca, T.J., Pathak, P., Chiantia, S., *et al.*, 2013. Proving lipid rafts exist: Membrane domains in the prokaryote *Borrelia burgdorferi* have the same properties as eukaryotic lipid rafts. *PLoS Pathogens* 9, e1003353.
- Lenne, P.F., Wawrezynieck, L., Conchonaud, F., *et al.*, 2006. Dynamic molecular confinement in the plasma membrane by microdomains and the cytoskeleton meshwork. *EMBO Journal* 25, 3245–3256.
- Lentz, B.R., Barrow, D.A., Hoeghli, M., 1980. Cholesterol-phosphatidylcholine interactions in multilamellar vesicles. *Biochemistry* 19, 1943–1954.
- Levental, I., Grzybek, M., Simons, K., 2010. Greasing their way: Lipid modifications determine protein association with membrane rafts. *Biochemistry* 49, 6305–6316.
- Levental, I., Lingwood, D., Grzybek, M., Coskun, U., Simons, K., 2010. Palmitoylation regulates raft affinity for the majority of integral raft proteins. *Proceedings of the National Academy of Sciences of the United States of America* 107, 22050–22054.
- Li, X.M., Momsen, M.M., Smaby, J.M., Brockman, H.L., Brown, R.E., 2001. Cholesterol decreases the interfacial elasticity and detergent solubility of sphingomyelins. *Biochemistry* 40, 5954–5963.
- Lin, Q., London, E., 2013a. Altering hydrophobic sequence lengths shows that hydrophobic mismatch controls affinity for ordered lipid domains (rafts) in the multitransmembrane strand protein perfringolysin O. *Journal of Biological Chemistry* 288, 1340–1352.
- Lin, Q., London, E., 2013b. Transmembrane protein (perfringolysin o) association with ordered membrane domains (rafts) depends upon the raft-associating properties of protein-bound sterol. *Biophysical Journal* 105, 2733–2742.
- Lin, Q., London, E., 2014. Preparation of artificial plasma membrane mimicking vesicles with lipid asymmetry. *PLoS One* 9, e87903.
- Lindwasser, O.W., Resh, M.D., 2002. Myristoylation as a target for inhibiting HIV assembly: Unsaturated fatty acids block viral budding. *Proceedings of the National Academy of Sciences of the United States of America* 99, 13037–13042.
- Lingwood, D., Ries, J., Schwille, P., Simons, K., 2008. Plasma membranes are poised for activation of raft phase coalescence at physiological temperature. *Proceedings of the National Academy of Sciences of the United States of America* 105, 10005–10010.
- Liu, W., Sohn, H.W., Tolar, P., Pierce, S.K., 2010. It's all about change: The antigen-driven initiation of B-cell receptor signaling. *Cold Spring Harbor Perspectives in Biology* 2, a002295.
- London, E., 2002. Insights into lipid raft structure and formation from experiments in model membranes. *Current Opinion in Structural Biology* 12, 480–486.
- London, E., 2005. How principles of domain formation in model membranes may explain ambiguities concerning lipid raft formation in cells. *Biochimica et Biophysica Acta* 1746, 203–220.
- London, E., Brown, D.A., 2000. Insolubility of lipids in triton X-100: Physical origin and relationship to sphingolipid/cholesterol membrane domains (rafts). *Biochimica et Biophysica Acta* 1508, 182–195.
- Macdonald, J.L., Pike, L.J., 2005. A simplified method for the preparation of detergent-free lipid rafts. *Journal of Lipid Research* 46, 1061–1067.
- Maeda, Y., Tashima, Y., Houjou, T., *et al.*, 2007. Fatty acid remodeling of GPI-anchored proteins is required for their raft association. *Molecular Biology of the Cell* 18, 1497–1506.
- Magee, A.I., Adler, J., Parmryd, I., 2005. Cold-induced coalescence of T-cell plasma membrane microdomains activates signalling pathways. *Journal of Cell Science* 118, 3141–3151.
- Mahammad, S., Dinic, J., Adler, J., Parmryd, I., 2010. Limited cholesterol depletion causes aggregation of plasma membrane lipid rafts inducing T cell activation. *Biochimica et Biophysica Acta* 1801, 625–634.
- Megha, Bakht, O., London, E., 2006. Cholesterol precursors stabilize ordinary and ceramide-rich ordered lipid domains (lipid rafts) to different degrees. Implications for the Bloch hypothesis and sterol biosynthesis disorders. *Journal of Biological Chemistry* 281, 21903–21913.
- Megha, London, E., 2004. Ceramide selectively displaces cholesterol from ordered lipid domains (rafts): Implications for lipid raft structure and function. *Journal of Biological Chemistry* 279, 9997–10004.
- Melkonian, K.A., Ostermeyer, A.G., Chen, J.Z., Roth, M.G., Brown, D.A., 1999. Role of lipid modifications in targeting proteins to detergent-resistant membrane rafts. Many raft proteins are acylated, while few are prenylated. *Journal of Biological Chemistry* 274, 3910–3917.
- Moffett, S., Brown, D.A., Linder, M.E., 2000. Lipid-dependent targeting of G proteins into rafts. *Journal of Biological Chemistry* 275, 2191–2198.
- Morris, R.J., Jen, A., Warley, A., 2011. Isolation of nano-meso scale detergent resistant membrane that has properties expected of lipid 'rafts'. *Journal of Neurochemistry* 116, 671–677.
- Mukherjee, S., Soe, T.T., Maxfield, F.R., 1999. Endocytic sorting of lipid analogues differing solely in the chemistry of their hydrophobic tails. *Journal of Cell Biology* 144, 1271–1284.
- Nelson, L.D., Chiantia, S., London, E., 2010. Perfringolysin O association with ordered lipid domains: Implications for transmembrane protein raft affinity. *Biophysical Journal* 99, 3255–3263.
- Nikolaus, J., Scolari, S., Bayraktarov, E., *et al.*, 2010. Hemagglutinin of influenza virus partitions into the nonraft domain of model membranes. *Biophysical Journal* 99, 489–498.
- Nixon, B., Aitken, R.J., 2009. The biological significance of detergent-resistant membranes in spermatozoa. *Journal of Reproductive Immunology* 83, 8–13.
- Nowaczyk, M.J., Irons, M.B., 2012. Smith-Lemli-Opitz syndrome: Phenotype, natural history, and epidemiology. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*. 160C, 250–262.
- Owen, D.M., Lanigan, P.M., Dunsby, C., *et al.*, 2006. Fluorescence lifetime imaging provides enhanced contrast when imaging the phase-sensitive dye di-4-ANEPPDHQ in model membranes and live cells. *Biophysical Journal* 90, L80–L82.
- Owen, D.M., Neil, M.A., French, P.M., Magee, A.I., 2007. Optical techniques for imaging membrane lipid microdomains in living cells. *Seminars in Cell & Developmental Biology* 18, 591–598.
- Pathak, P., London, E., 2011. Measurement of lipid nanodomain (raft) formation and size in sphingomyelin/POPC/cholesterol vesicles shows TX-100 and transmembrane helices increase domain size by coalescing preexisting nanodomains but do not induce domain formation. *Biophysical Journal* 101, 2417–2425.
- Petruziolo, R.S., Heberle, F.A., Drazba, P., Katsaras, J., Feigenson, G.W., 2013. Phase behavior and domain size in sphingomyelin-containing lipid bilayers. *Biochimica et Biophysica Acta* 1828, 1302–1313.
- Pinaud, F., Michalet, X., Iyer, G., *et al.*, 2009. Dynamic partitioning of a glycosyl-phosphatidylinositol-anchored protein in glycosphingolipid-rich microdomains imaged by single-quantum dot tracking. *Traffic* 10, 691–712.
- Polozova, A., Litman, B.J., 2000. Cholesterol dependent recruitment of di22:6-PC by a G protein-coupled receptor into lateral domains. *Biophysical Journal* 79, 2632–2643.
- Ramstedt, B., Slotte, J.P., 1999. Interaction of cholesterol with sphingomyelins and acyl-chain-matched phosphatidylcholines: A comparative study of the effect of the chain length. *Biophysical Journal* 76, 908–915.

- Recktenwald, D.J., McConnell, H.M., 1981. Phase equilibria in binary mixtures of phosphatidylcholine and cholesterol. *Biochemistry* 20, 4505–4510.
- Refaei, M., Leventis, R., Silvius, J.R., 2011. Assessment of the roles of ordered lipid microdomains in post-endocytic trafficking of glycosyl-phosphatidylinositol-anchored proteins in mammalian fibroblasts. *Traffic* 12, 1012–1024.
- Rentero, C., Zech, T., Quinn, C.M., *et al.*, 2008. Functional implications of plasma membrane condensation for T cell activation. *PLoS One* 3, e2262.
- Rietveld, A., Simons, K., 1998. The differential miscibility of lipids as the basis for the formation of functional membrane rafts. *Biochimica et Biophysica Acta* 1376, 467–479.
- Rossman, J.S., Jing, X., Leser, G.P., Lamb, R.A., 2010. Influenza virus M2 protein mediates ESCRT-independent membrane scission. *Cell* 142, 902–913.
- Roux, A., Cuvelier, D., Nassoy, P., *et al.*, 2005. Role of curvature and phase transition in lipid sorting and fission of membrane tubules. *EMBO Journal* 24, 1537–1545.
- Rubenstein, J.L., Smith, B.A., McConnell, H.M., 1979. Lateral diffusion in binary mixtures of cholesterol and phosphatidylcholines. *Proceedings of the National Academy of Sciences of the United States of America* 76, 15–18.
- Samsonov, A.V., Mihalov, I., Cohen, F.S., 2001. Characterization of cholesterol-sphingomyelin domains and their dynamics in bilayer membranes. *Biophysical Journal* 81, 1486–1500.
- Sankaram, M.B., Thompson, T.E., 1990. Interaction of cholesterol with various glycerophospholipids and sphingomyelin. *Biochemistry* 29, 10670–10675.
- Schroeder, R., London, E., Brown, D., 1994. Interactions between saturated acyl chains confer detergent resistance on lipids and glycosylphosphatidylinositol (GPI)-anchored proteins: GPI-anchored proteins in liposomes and cells show similar behavior. *Proceedings of the National Academy of Sciences of the United States of America* 91, 12130–12134.
- Sengupta, P., Hammond, A., Holowka, D., Baird, B., 2008. Structural determinants for partitioning of lipids and proteins between coexisting fluid phases in giant plasma membrane vesicles. *Biochimica et Biophysica Acta* 1778, 20–32.
- Sengupta, P., Holowka, D., Baird, B., 2007. Fluorescence resonance energy transfer between lipid probes detects nanoscopic heterogeneity in the plasma membrane of live cells. *Biophysical Journal* 92, 3564–3574.
- Sezgin, E., Levental, I., Grzybek, M., *et al.*, 2012. Partitioning, diffusion, and ligand binding of raft lipid analogs in model and cellular plasma membranes. *Biochimica et Biophysica Acta* 1818, 1777–1784.
- Sharpe, H.J., Stevens, T.J., Munro, S., 2010. A comprehensive comparison of transmembrane domains reveals organelle-specific properties. *Cell* 142, 158–169.
- Shlomovitz, R., Schick, M., 2013. Model of a raft in both leaves of an asymmetric lipid bilayer. *Biophysical Journal* 105, 1406–1413.
- Shogomori, H., Hammond, A.T., Ostermeyer-Fay, A.G., *et al.*, 2005. Palmitoylation and intracellular domain interactions both contribute to raft targeting of linker for activation of T cells. *Journal of Biological Chemistry* 280, 18931–18942.
- Silvius, J.R., 2003. Fluorescence energy transfer reveals microdomain formation at physiological temperatures in lipid mixtures modeling the outer leaflet of the plasma membrane. *Biophysical Journal* 85, 1034–1045.
- Silvius, J.R., del Giudice, D., Laffleur, M., 1996. Cholesterol at different bilayer concentrations can promote or antagonize lateral segregation of phospholipids of differing acyl chain length. *Biochemistry* 35, 15198–15208.
- Simons, K., Ikonen, E., 1997. Functional rafts in cell membranes. *Nature* 387, 569–572.
- Simons, K., Toomre, D., 2000. Lipid rafts and signal transduction. *Nature Reviews Molecular Cell Biology* 1, 31–39.
- Simons, K., van Meer, G., 1988. Lipid sorting in epithelial cells. *Biochemistry* 27, 6197–6202.
- Singh, D.K., Rosenhouse-Dantsker, A., Nichols, C.G., Enkvetchakul, D., Levitan, I., 2009. Direct regulation of prokaryotic Kir channel by cholesterol. *Journal of Biological Chemistry* 284, 30727–30736.
- Skibbens, J.E., Roth, M.G., Matlin, K.S., 1989. Differential extractability of influenza virus hemagglutinin during intracellular transport in polarized epithelial cells and nonpolar fibroblasts. *Journal of Cell Biology* 108, 821–832.
- Sohn, H.W., Pierce, S.K., Tzeng, S.J., 2008. Live cell imaging reveals that the inhibitory FcγRIIB destabilizes B cell receptor membrane-lipid interactions and blocks immune synapse formation. *Journal of Immunology* 180, 793–799.
- Surma, M.A., Klose, C., Klemm, R.W., Ejsing, C.S., Simons, K., 2011. Generic sorting of raft lipids into secretory vesicles in yeast. *Traffic* 12, 1139–1147.
- Suzuki, K.G., Kasai, R.S., Hirosawa, K.M., *et al.*, 2012. Transient GPI-anchored protein homodimers are units for raft organization and function. *Nature Chemical Biology* 8, 774–783.
- Thompson, T.E., Tillack, T.W., 1985. Organization of glycosphingolipids in bilayers and plasma membranes of mammalian cells. *Annual Review of Biophysics and Biophysical Chemistry* 14, 361–386.
- Tong, X., Briggs, M.M., Mlaver, D., Vidal, A., McIntosh, T.J., 2009. Sorting of lens aquaporins and connexins into raft and nonraft bilayers: Role of protein homo-oligomerization. *Biophysical Journal* 97, 2493–2502.
- Toulmay, A., Prinz, W.A., 2013. Direct imaging reveals stable, micrometer-scale lipid domains that segregate proteins in live cells. *Journal of Cell Biology* 202, 35–44.
- Veatch, S.L., Cicuta, P., Sengupta, P., *et al.*, 2008. Critical fluctuations in plasma membrane vesicles. *ACS Chemical Biology* 3, 287–293.
- Veatch, S.L., Keller, S.L., 2003. Separation of liquid phases in giant vesicles of ternary mixtures of phospholipids and cholesterol. *Biophysical Journal* 85, 3074–3083.
- Veatch, S.L., Soubias, O., Keller, S.L., Gawrisch, K., 2007. Critical fluctuations in domain-forming lipid mixtures. *Proceedings of the National Academy of Sciences of the United States of America* 104, 17650–17655.
- Veit, M., Engel, S., Thaa, B., Scolari, S., Herrmann, A., 2013. Lipid domain association of influenza virus proteins detected by dynamic fluorescence microscopy techniques. *Cellular Microbiology* 15, 179–189.
- Vetrivel, K.S., Thinakaran, G., 2010. Membrane rafts in Alzheimer's disease beta-amyloid production. *Biochimica et Biophysica Acta* 1801, 860–867.
- Vidal, A., McIntosh, T.J., 2005. Transbilayer peptide sorting between raft and nonraft bilayers: Comparisons of detergent extraction and confocal microscopy. *Biophysical Journal* 89, 1102–1108.
- Vist, M.R., Davis, J.H., 1990. Phase equilibria of cholesterol/dipalmitoylphosphatidylcholine mixtures: ²H nuclear magnetic resonance and differential scanning calorimetry. *Biochemistry* 29, 451–464.
- Waheed, A.A., Freed, E.O., 2009. Lipids and membrane microdomains in HIV-1 replication. *Virus Research* 143, 162–176.
- Wan, C., Kiessling, V., Tamm, L.K., 2008. Coupling of cholesterol-rich lipid phases in asymmetric bilayers. *Biochemistry* 47, 2190–2198.
- Wang, J., Megha, London, E., 2004. Relationship between sterol/steroid structure and participation in ordered lipid domains (lipid rafts): Implications for lipid raft structure and function. *Biochemistry* 43, 1010–1018.
- Wang, T.Y., Leventis, R., Silvius, J.R., 2001. Partitioning of lipidated peptide sequences into liquid-ordered lipid domains in model and biological membranes. *Biochemistry* 40, 13031–13040.
- Wang, T.Y., Leventis, R., Silvius, J.R., 2005. Artificially lipid-anchored proteins can elicit clustering-induced intracellular signaling events in Jurkat T-lymphocytes independent of lipid Raft association. *Journal of Biological Chemistry* 280, 22839–22846.
- Wang, T.Y., Silvius, J.R., 2001. Cholesterol does not induce segregation of liquid-ordered domains in bilayers modeling the inner leaflet of the plasma membrane. *Biophysical Journal* 81, 2762–2773.
- Wenz, J.J., Barrantes, F.J., 2003. Steroid structural requirements for stabilizing or disrupting lipid domains. *Biochemistry* 42, 14267–14276.
- Xu, X., Bittman, R., Duportail, G., *et al.*, 2001. Effect of the structure of natural sterols and sphingolipids on the formation of ordered sphingolipid/sterol domains (rafts). Comparison of cholesterol to plant, fungal, and disease-associated sterols and comparison of sphingomyelin, cerebroside, and ceramide. *Journal of Biological Chemistry* 276, 33540–33546.
- Xu, X., London, E., 2000. The effect of sterol structure on membrane lipid domains reveals how cholesterol can induce lipid domain formation. *Biochemistry* 39, 843–849.
- Yu, J., Fischman, D.A., Steck, T.L., 1973. Selective solubilization of proteins and phospholipids from red blood cell membranes by nonionic detergents. *Journal of Supramolecular Structure* 1, 233–248.
- Zech, T., Ejsing, C.S., Gaus, K., *et al.*, 2009. Accumulation of raft lipids in T-cell plasma membrane domains engaged in TCR signalling. *EMBO Journal* 28, 466–476.
- Zidovetzki, R., Levitan, I., 2007. Use of cyclodextrins to manipulate plasma membrane cholesterol content: Evidence, misconceptions and control strategies. *Biochimica et Biophysica Acta* 1768, 1311–1324.
- Zitzer, A., Bittman, R., Verbikey, C.A., *et al.*, 2001. Coupling of cholesterol and cone-shaped lipids in bilayers augments membrane permeabilization by the cholesterol-specific toxins streptolysin O and *Vibrio cholerae* cytolysin. *Journal of Biological Chemistry* 276, 14628–14633.

Membrane Potential: Concepts

AJ Moorhouse, UNSW Australia, Sydney, NSW, Australia

© 2016 Elsevier Inc. All rights reserved.

Introduction

Everything we do, from playing a challenging piano concerto or sprinting in a 100 m race, or even simply walking and breathing and admiring the birds and trees, involves a complex array of muscle activity and sensory inputs all coordinated by the brain and nerves. As we eat and metabolize our food, sleep and grow from fertilization to maturity, our digestive and endocrine systems regulate our bodies and maintain homeostasis under very different environments. Our respiratory and cardiovascular systems provide the energy to maintain our bodily functions. All of these wonderful activities rely on cellular functions that depend on changes in the small voltage difference across the thin plasma membrane that encapsulates all animal cells, the 'membrane potential.' This 20–100 mV gradient across our cell membrane is a key component to cell and body function, and a large fraction of our cellular energy is spent on establishing and regulating this transmembrane voltage. In this article, the author reviews the principles of how membrane potentials are generated by the combination of selective membrane permeability and ion electrochemical gradients. Section 'Introduction' will provide some basic principles regarding electrical and chemical properties of ions and cell membranes, section 'Determinants of the Membrane Potential' will build on this to describe more specifically how membrane potentials arise as a consequence of electrochemical gradients and selective membrane permeability, and section 'Measuring Membrane Potential and Relative Membrane Permeability' will review approaches to measure membrane selectivity and membrane potentials, providing some exemplar values from the literature.

Definitions of Different Types of Membrane Potentials and Related Terminology

The following list defines some of the different types of membrane potentials, and some related terminology. The article assumes familiarity with these terms (Box 1).

Physical–Chemical Properties of the Cell Membrane and Physiological Solutions

The membrane potential, like any voltage difference, reflects a different electrical charge from one point to another point. In this case, the uneven charge distribution is across the cell membrane, from the intracellular surface of the membrane to the extracellular surface of the membrane. This arises due to a different concentration of charged ions on each side of the membrane. So an important introductory concept is that in biology, positively charged cations and negatively charged anions contribute to the charge on either side of the cell membrane, and the flow of charge (or current) across the

membrane is carried by the movement of ions across the membrane. To understand this, we first need to review some simple aspects of electricity and of physical chemistry.

A crystal of common table salt, or NaCl, contains Na and Cl atoms bound tightly together in a lattice bound by ionic bonds. The Na atom loses its electron to the Cl atom to give rise to Na^+ and Cl^- ions. When a pinch of salt is placed into water, the ionic bonds are disrupted by electrical interactions between water and the ions, and the salt crystals dissociate into hydrated Na^+ and Cl^- ions, that is the ions surrounded by a shell of water molecules (Figure 1(a)). The water molecules are 'polar,' meaning that the distribution of electrons is not evenly spread across the H_2O molecules – the electrons spend more time around the oxygen atom as compared to the two hydrogen atoms. Hence the oxygen in water has a partial negative charge, while the hydrogen atoms have a partial positive charge. Ions are energetically unstable in isolation but are stabilized in solution by being surrounded by water orientated accordingly with Na^+ ions surrounded or shielded by negative oxygen water molecules, while the Cl^- anion is surrounded by water with the positive hydrogen facing the

Box 1

Membrane potential (V_m): A difference in voltage across the cell membrane arising due to small difference in the distribution of ions in the intracellular and extracellular solutions. Defined as intracellular voltage with respect to extracellular voltage.

Resting membrane potential: The membrane potential during a period when the cell is not active or being stimulated. Typically, a stable, steady-state condition of about -80 to -60 mV. Note however that many cells have an unstable or fluctuating resting V_m that if in a regular pattern can be called pacemaker activity.

Action potential: A basic unit of cell signaling that involves a transient and rapid change in membrane potential. Although the time course and amplitude of this V_m change varies widely in cells, in many nerve and muscle cells V_m changes from about -70 mV up to about $+20$ mV before returning to -70 mV.

Depolarization: A change in the V_m in a positive direction. Typically increases the probability of a cell firing an action potential, hence typically excitatory.

Hyperpolarization: A change in the V_m in a negative direction. Typically decreases the probability of a cell firing an action potential, hence typically inhibitory.

Repolarisation: A return of the membrane potential to resting value following a large depolarization, such the hyperpolarisation phase of the action potential

Synaptic potential: A change in membrane potential in a postsynaptic cell arising from action of a neurotransmitter released by a presynaptic nerve. It can be in the hyperpolarising direction (an inhibitory postsynaptic potential) or in a depolarizing direction (an excitatory postsynaptic potential).

Sensory potential: A change in the membrane potential of a sensory receptor neuron arising due to a change in a sensory modality input (touch, smell, taste, vision, and hearing).



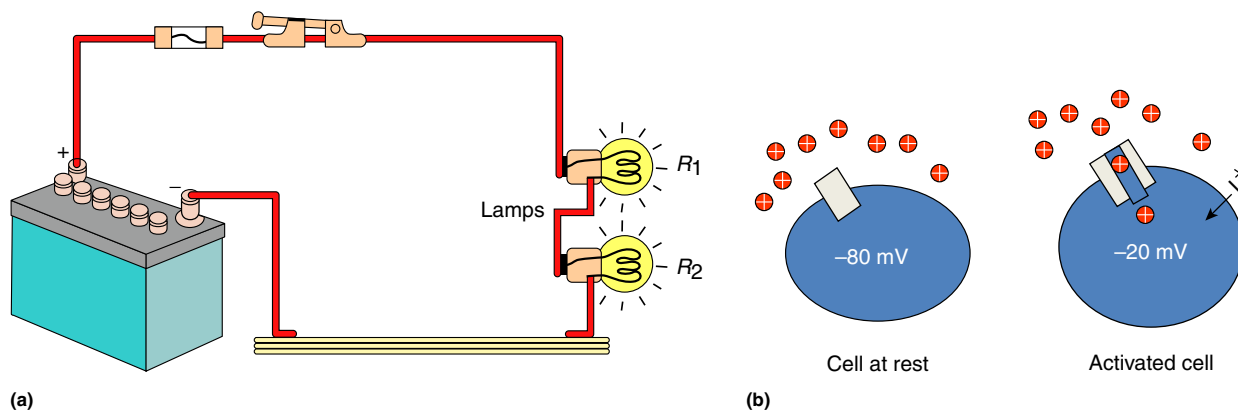


Figure 2 Electrical circuits: hardwired vs. biology. (a) Schematic diagram of an electrical circuit showing a car battery connected via a wire and switch to two car headlights/lamps. The battery has a different distribution of charges providing a negative and positive pole (cathode and anode, respectively) that drives current in the form of electrons through the wire conductor. Current only flows when the switch is closed to complete the circuit. Current flows through the resistance provided by the lamp filaments, heating the filaments to cause the response – light. Available at: <http://electricalengproject.blogspot.com.au/2012/06/electrical-circuits.html> (accessed 17.02.15). (b) In biology, the electrical battery is equivalent to the membrane potential with a negative charge (cathode) on the inside and a positive charge (anode) outside. When a switch (ion channel) is open, current in the form of charged ions can flow into the cell to cause a response – a depolarization.

anion. Molecules, like ions, that are readily dissolved in water (or become hydrated by water) are said to be ‘hydrophilic.’

In contrast to water, the interior of the cell membrane is ‘nonpolar’ and ‘hydrophobic.’ The long hydrocarbon chains of the lipid molecules that comprise the membrane have no partial charges and cannot readily dissolve in water. This is readily seen when oil is spilled into water by the way oil forms a film floating on the surface of the water. Nonpolar molecules will however readily interact with other nonpolar molecules (just as polar molecules readily interact). Hence the hydrated ions cannot freely pass through the cell membrane due to this nonpolar lipid core. The energetics of moving an ion from an aqueous solution (with a high dielectric constant of ≈ 80) to the interior of the lipid bilayer (with a dielectric constant of ≈ 2) is so high that less than 1 in 10^{50} ions would partition into the membrane (Coster, 2003). The lipid bilayer can therefore readily sustain a charge separation corresponding to a physiological V_m of $\pm 100 \text{ mV}$. As the membrane is only 5–10 nm thick, this corresponds to a very large electric field strength ($\approx 10^7 \text{ V m}^{-1}$), demonstrating the strong insulating capacity of the lipid bilayer. If the V_m is artificially increased to a few hundred mV, say to 400 mV, the membrane may transiently break down. This ‘punchthrough’ or dielectric breakdown may become irreversible if even larger voltages are applied across the membrane (1 V or more). The voltage-induced transient membrane rupture has actually been utilized

to enable large molecules to be inserted into the cell by electroporation (Coster, 2009).

If the hydrocarbon tails of lipids are nonpolar, how can they exist bathed by the polar salt solutions that comprise intracellular and extracellular solutions? The answer is that the lipid molecules that make up mammalian cell membranes are phospholipids and ‘amphiphilic,’ containing polar headgroups attached to the nonpolar hydrocarbon tails (Figure 1 (b)). The polar headgroups can readily interact with water. These phospholipids will spontaneously adopt a bilayer configuration when placed in a polar solvent such as water – the nonpolar tails will join up and the polar headgroups will form the border with the polar solvent. Hence the plasma membrane of a mammalian cell contains a phospholipid bilayer which physically and electrically separates the intracellular and extracellular solutions. Within this bilayer membrane are inserted numerous proteins which function as membrane transport proteins, enzymes, and receptors to communicate with other cells.

Electrical Properties of the Cell and Ohm’s Law

Consider a simple hardwired electrical circuit such as that found in a hand-held torch or car battery (Figure 2(a)). The driving force for providing electrical current through this

Figure 1 Schematic structures of hydrated ions and the cell membrane lipids. (a) Schematic diagram of a crystal of common table salt (NaCl; upper panel) with the Na^+ and Cl^- bound in a tight ionic lattice. Upon addition of the polar water (H_2O) molecule, the Na^+ and Cl^- separate and become hydrated ions (lower panels). Reproduced from Allman, R. A helpful introduction to polar and nonpolar molecules. The chemistry webpages of Robert Allman. Available at: <http://www.chemstone.net/Principles/7.Solutions/Soln.html> (accessed 17.02.15). (b) Schematic diagram of a single typical membrane phospholipid molecule (this example is ‘phosphatidylcholine’) showing its chemical structure and a schematic depiction of its hydrophobic fatty acid tail, and its polar head (phosphate and choline), joined by the glycerol backbone. The lower right panel depicts a lipid bilayer such as that which forms the cell membrane, with the nonpolar tails joined together, and the polar head groups facing the extracellular and intracellular solutions. Figure accessed from the CK-12 Foundation under a creative commons CC BY-NC license (<http://www.ck12.org/biology/Phospholipid-Bilayers/lesson/Phospholipid-Bilayers/>).

circuit illustrated in [Figure 2\(a\)](#) is a 12 V battery. A copper wire that connects the two poles of the battery – the cathode and anode – acts as a conductor of the electrical current. The current is movement of charge in the form of electrons traveling along this conductor. The switch is a break in the wire that only allows current to flow when the circuit between the cathode and anode is complete. The resistance of the circuit is determined by factors such as the thickness, length, and conductivity of the wire, and in this case by the properties of the high-resistance filament placed in series in this circuit. When the switch is on, current flows through the circuit driven by the potential difference (voltage) of the battery, and the filament heats up and provides light to mediate the car headlamp or torch's function.

In biology, the principles are just the same but the elements are different ([Figure 2\(b\)](#)). The circuit is the flow of current across the membrane. The current is carried by ions flowing through the circuit. The battery providing the driving force for ionic current flow is the 'electrochemical driving force' for the ion that carries the current. This concept is described in more detail below – let us assume for now that the current carrying ion, say Na^+ , is equally distributed across the membrane and so only the potential difference across the cell membrane, the V_m , equates to the battery. At a typical resting V_m of around -80 mV, the cathode (or negative pole) is the intracellular solution and the anode the extracellular solution. The conductors are the membrane transport proteins (ion channels) that traverse the cell membrane. These channels have a central water filled pore that allows the ion to flow across the membrane, and generally also contain small molecular switches or gates that open or close this channel pore. The resistance of this circuit depends on the properties of these ion channels (how many ions they allow to pass) and how many of these ion channels are open in the cell membrane. When the channel gate is open, Na^+ ions flow into the cell causing a depolarization and activating some cell function (such as triggering an action potential or the release of a neurotransmitter).

In both the biological circuit and the hardwired circuit, the relationship between the potential difference (the battery, voltage, V), the amount of current (I) flowing through the circuit, and the resistance (R) or conductance (G) of the circuit is given by 'Ohm's law' (eqn [1]):

$$V = IR = I/G \quad [1]$$

From Ohm's law we can see that when current is injected into a typical cell, for example, by cations entering the cell from the external solution, the extent of the resultant depolarization produced depends on the resistance of the cell. The intracellular salt solution has a very low resistivity, so the cell's resistance depends on the amount of cell membrane, and the resistance of each unit of cell membrane. A cell membrane with a lot of open ion channels has a low resistance (or a high conductance, or is 'leaky') and will generate a smaller voltage change than that of a higher resistance cell in response to the same input of ionic current. Indeed, opening of ion channels to reduce a cell's input resistance and thereby reduce the voltage response to current injection is an important mechanism of inhibition in the brain ('shunting' inhibition; [Farrant and Kaila, 2007](#)). Given the same distribution of open ion

channels in a cell membrane, a larger cell will have a lower input resistance, and will require a greater input of ionic current to reach a specific V_m level. A physiological illustration of this is the 'size principle' for motoneuron recruitment during muscle contraction ([Henneman et al., 1965](#)). At modest levels of ionic current from afferent nerve drive, only the higher resistance smaller motoneurons are depolarized to the voltage threshold and activated. These small motoneurons innervate the 'slow-twitch' type muscle fibers that produce only modest muscle force. When stronger muscle contractions are needed, a greater afferent drive is needed to depolarize and activate the larger motoneurons that supply the fast twitch, high-force fibers. Hence by recruiting motoneurons in order of size and force, the brain can control the extent of motor force needed for a task (increasing force by increasing the frequency of activation is also important).

Ohm's law illustrates a very simple circuit with current flowing through a resistor driven by a voltage gradient. However, the cell membrane behaves like a resistor and a capacitor in parallel ([Figure 3](#)). A capacitor is an electrical device designed to store charge, and consists of two conducting plates separated by insulating material. The conducting plates are the salt solutions, and the insulating material is the lipid core. The lipid can support a different distribution of charges, or a different potential on each side. In a perfect capacitor, this charge separation would be maintained without decrement. However, the cell membrane contains ion channels that act as a resistor to allow the flow of charge (current) across the membrane. Therefore, the membrane is like a leaky capacitor. When charge is moved from one side to the other, via ionic current flow through open ion channels, the voltage on the two plates of the capacitor (the intracellular and extracellular surface) starts to change, but the change in V_m is not instantaneous. Current flows through the capacitive element of the circuit as the voltage across the capacitor changes. A new voltage difference is eventually established across the membrane, with an amplitude dependent on the membrane current, I_m , and the membrane resistance, R_m , as given by Ohm's law. Although the current is applied virtually instantaneously as ion channels rapidly open and close, the change in V_m changes with a membrane time constant (τ), given by the product of the resistance and capacitance of the circuit. [Figure 3](#) demonstrates the electrical circuit equivalent for a cell membrane, and the time course of a change in V_m when current flows into a cell.

Determinants of the Membrane Potential

The membrane potential (V_m) arises due to a different distribution of ions on each side of the membrane. One needs to understand two broad concepts to appreciate how this different distribution of ions can occur. Firstly, one needs to appreciate the forces that determine whether an ion moves in or out of the membrane. Secondly, there needs to be a means by which ions can move across the membrane – the membrane needs to be permeable to that ion. In this section we describe these two factors that determine the distribution of ions across the membrane, and thereby determine the V_m .

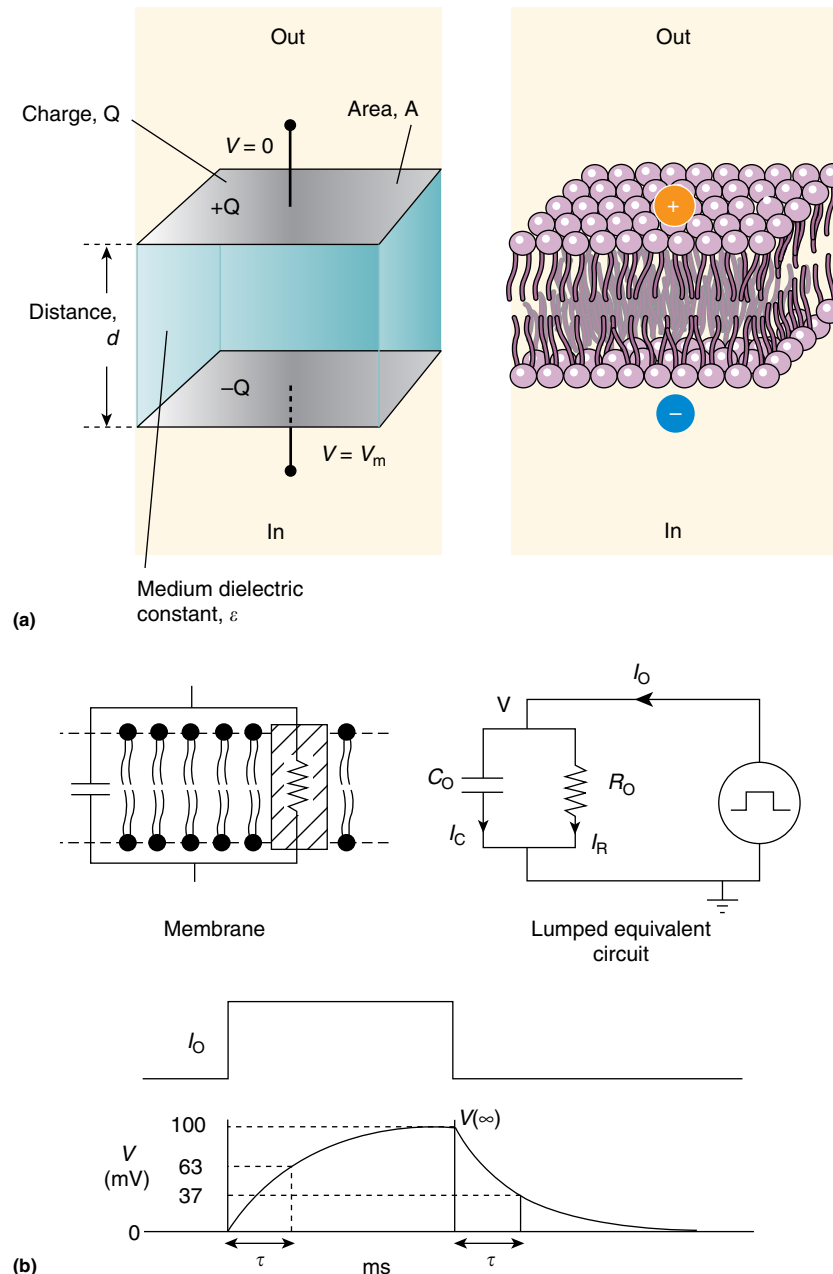


Figure 3 Electrical equivalent and voltage response of the cell membrane. (a) The lipid bilayer (right) acts like an ideal capacitor (left) that separates a charge, Q , on the two conducting plates, separated by a distance, d . The charge separation gives rise to a Voltage (V) across the capacitor (C), related by the equation $Q = CV$. The capacitance is directly proportional to the area of the plates, A , the dielectric constant of the medium, ϵ , the permittivity constant ϵ_0 , and inversely proportional to the distance between the plates ($C = Ae_0\epsilon/d$). (b) The membrane behaves electrically as an 'RC' circuit, an electrical circuit with a capacitor (from the lipid bilayer) and a resistance (from the ion channels) in parallel. Injection of a square pulse of current into this circuit (as may happen when an ion channel opens) gives rise to a slowly rising and decaying change in voltage as shown in the lower panel. The current rises to a value of $V = IR$ with a time course dependent on τ , the membrane time constant ($= RC$, or the time to decay to 37% of maximum, = the time to reach 63% of maximum). The voltage at time, t , during the rising phase is given by $V_t = IR \times (1 - e^{-t/\tau})$ while the voltage at any time, t , during the decay phase is given by $V_t = IR \times e^{-t/\tau}$. Panel (a) is reproduced from panel C of Figure 6.9 in Boron, W.F., Boulpaep, E.L., 2009. Medical Physiology. A Cellular and Molecular Approach, second ed. Philadelphia, PA: Saunders Elsevier, © 2002 Elsevier Science, USA; Panel (b) with permission from Prof. Peter Barry, © PH Barry, 2002.

Electrochemical Driving Force for Ion Movement

The major ionic species inside and surrounding cells are K^+ , Na^+ , and Cl^- . Larger molecules that carry negative charges,

such as proteins and organic anions, and that find it difficult to cross the membrane form a pool of impermeant anions inside cells. Smaller concentrations of other ions such as bicarbonate, Ca^{2+} , inorganic phosphates ($H_2PO_4^-$, HPO_4^{2-}), and Mg^{2+}

Table 1 Physiological concentrations and equilibrium potentials of common physiological electrolytes

<i>Ion</i>	<i>Extracellular (mM)</i>	<i>Intracellular (mM)</i>	<i>E_{ion} (mV; 37 °C)</i>	<i>V_m–E_{ion} (mV)</i>	<i>Gradient at rest</i>
Na ⁺	145	18	+56	–132	Influx
K ⁺	3.0	135	–102	+26	Efflux
Ca ²⁺	1.2	10 ^{–7}	+125	–201	Influx
Cl [–]	120	7	–76	0	
HCO ₃ [–]	23	15	–11	–65	Efflux
pH [H ⁺]	7.4	7.2	–12	–64	Influx
Osmolality	290	290	Not applicable	Not applicable	Nil H ₂ O gradient

Note: Concentrations of Na⁺, K⁺, Ca²⁺, and Cl[–] are based on McCormick (2008); [HCO₃[–]] which was calculated using the Henderson–Hasselbalch equation (pH=6.1 + log [HCO₃[–]] – log (0.031 × PCO₂)) using the given pH values, and with CO₂=5% (PCO₂=38 mmHg). *E_{ion}* calculated for 37 °C using concentrations, assuming equal activities in both solutions. Resting *V_m* was taken as –76 mV, as measured for hippocampal neurons *in vivo* by Tyzio *et al.* (2008).

also exist in the extracellular and intracellular solutions. These ions do not exist in equal concentrations in these two solution compartments. Table 1 lists the brain concentrations of some physiological ions important in the context of membrane potentials. Cells use a significant amount of energy to actively transport or ‘pump’ ions across the membrane to establish these different concentrations. The Na⁺ pump, for example, hydrolyzes adenosine triphosphate (ATP) to pump Na⁺ ions out of the cell and to pump K⁺ ions into the cell (reviewed in Kaplan, 2002). The Na⁺ pump is one of the most important proteins in our bodies and found in virtually all polarized mammalian cells, where it creates a concentration gradient for K⁺ and Na⁺ across the cell membrane. The Law of Diffusion dictates that substances move from high to low concentrations, as the thermal motions of molecules give rise to collisions between closely spaced molecules that drives them further apart. Hence a chemical force exists for ions to move across the membrane from a high concentration to a low concentration. The chemical force acts to move K⁺ from the intracellular solution to the extracellular solution, and to move Na⁺ from extracellular to intracellular solution. However, as ions are charged substances, they are also subject to electrical forces. When placed in an electric field, a cation will be attracted to the negative pole and repelled from the positive pole, another Law of Physics – opposite charges attract and like charges repel. The cell membrane potential is an electric field across the membrane, and hence will also attract and repel ions from its positive and negative poles. The next paragraph examines how these two forces combine to produce a specific *V_m*.

Consider a hypothetical cell (Figure 4) in which the Na⁺ pump has established concentration gradients for Na⁺ and K⁺ across the cell membrane. This hypothetical cell is simplified by balancing the positive charge of these cations by including equimolar concentrations of Cl[–] on both sides. At hypothetical time=0 the membrane does not allow any ions to cross – it is impermeant to all the ions. The ionic charges are balanced in both the inside solution and outside solution (Figure 4(a(i))). Let us now make the membrane permeable to K⁺ by including open K⁺ – selective ion channels in the membrane. These channels we assume to be perfectly selective for K⁺, so the membrane remains impermeant to Cl[–] and Na⁺. The K⁺ ions will diffuse down their concentration gradient, from inside to out, carrying their positive charge as they move across the membrane. Their diffusion results in an excess

of K⁺ on the extracellular side of the membrane, and a deficit of K⁺ on the intracellular side. There exists an uneven distribution of charges, and hence the membrane is now polarized; negatively charged on the inside with respect to the extracellular side (Figure 4(a(ii))). This provides an electrical force on the ions. The negative *V_m* acts to attract the K⁺ toward the intracellular solution, countering the effects of the chemical force. An electrochemical equilibrium will be reached, where the chemical and electrical forces are equal, and the efflux of K⁺ due to the concentration gradient will be balanced by equal influx of K⁺ attracted by the negative *V_m*. The *V_m* at which this equilibrium is reached is called the equilibrium potential (*E_{ion}*) or Nernst potential. The Nernst equation (Box 2 below) equates the electrical and chemical forces that act on a specific ion, and is used to calculate the equilibrium potential for each specific ion (Table 1).

Relative Membrane Permeability

As shown in Table 1, the equilibrium potential for K⁺ under physiological ion concentrations in the brain and at body temperature is about –100 mV. Hence, if only highly selective K⁺ channels were open in a cell the *V_m* would rapidly equilibrate at –100 mV (Figure 4(b(i))). Imagine that (Figure 4(b(ii))) highly selective Na⁺ channels were opened instead of the K⁺ channels. Na⁺ would flow into the cell down its concentration gradient, causing an excess of positive Na⁺ inside the cell and a deficit outside the cell. The membrane would again become polarized with a positive *V_m*, that provides an electrical force to oppose the further influx of Na⁺. An electrochemical equilibrium would again be established, but this time at a positive *V_m*. The *V_m* would reach the equilibrium potential of Na⁺ which under typical physiological conditions would be about +60 mV (Table 1). Similarly, if a cell only allows Ca²⁺ or Cl[–] to move across the membrane the *V_m* would equal the equilibrium potentials of these ions, being +123 mV or –76 mV at the physiological concentrations of these ions (Table 1).

Imagine now (Figure 4(b(iii))) that a hypothetical cell was permeable to both K⁺ and Na⁺ (and no other ions) and contained physiological ion concentrations as in Table 1. K⁺ would leave the cell according to the chemical force, causing a negative *V_m*, and would continue to leave until the *V_m* reached the K⁺ equilibrium potential (*E_K*, –102 mV). At the same time, Na⁺ would enter, driving the *V_m* positive, and would

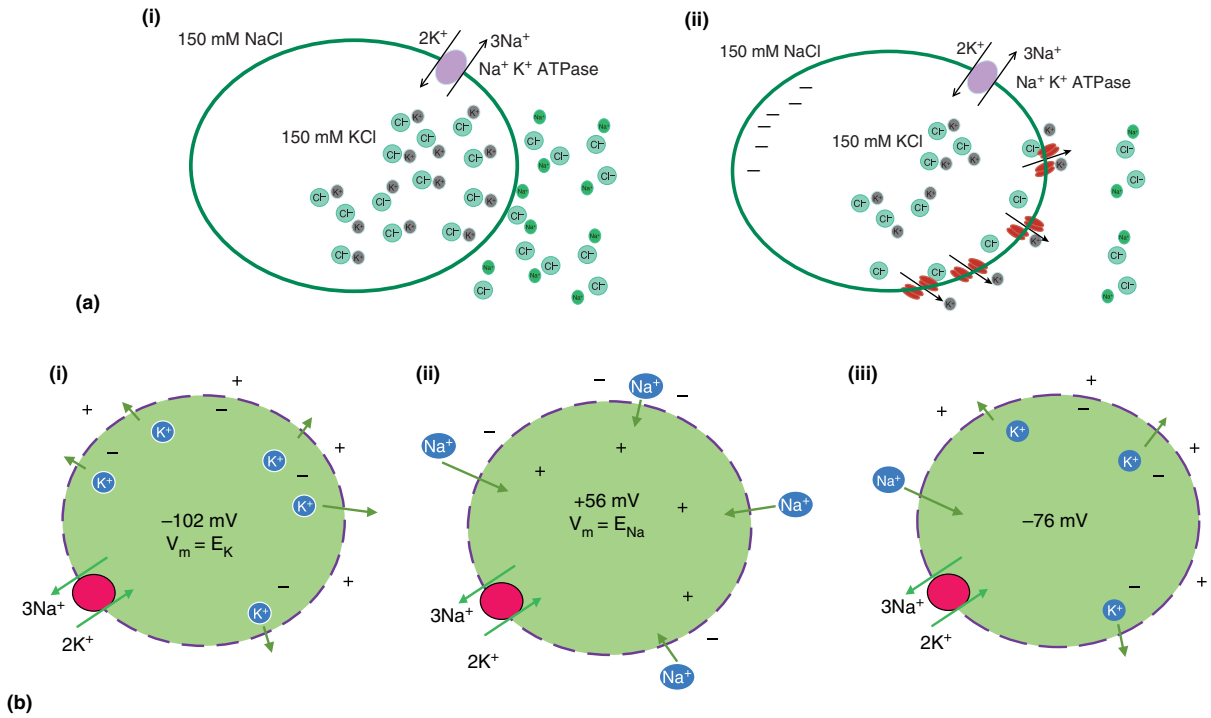


Figure 4 Electrochemical equilibrium and the resting membrane potential. (a) Schematic diagram illustrating how chemical and electrical forces combine to generate a membrane potential. In panel (i) (left) a hypothetical cell has a Na^+ pump that establishes concentration gradients for Na^+ and K^+ . In (ii) the cell becomes permeable to K^+ , which flows down its concentration gradient to establish the different ion charge distribution that is the membrane potential (V_m). (b) Hypothetical cells as in (a). In (i) only K^+ is permeable, K^+ flows out of the cell until the V_m reaches -102 mV , the equilibrium potential for K^+ (E_K). In (ii) only Na^+ is permeable, Na^+ flows in to the cell until the V_m reaches $+56 \text{ mV}$, E_{Na} . In (iii) both K^+ and Na^+ are permeable, with $P_K \gg P_{\text{Na}}$. K^+ flows out of the cell to drive V_m toward E_K , while Na^+ flows in to drive V_m toward E_{Na} . The V_m reaches a steady-state value somewhere in between, depending on the relative permeability toward K^+ and Na^+ . Neither ion is in equilibrium, so continue to flow down their electrochemical energy gradients, but the charge movement of Na^+ influx is balanced by that from K^+ efflux. This is similar to the resting V_m condition in many cells.

continue to enter until V_m reached the Na^+ equilibrium potential (E_{Na} , $+56 \text{ mV}$). The membrane potential would reach some steady-state value that balances these two opposing electrochemical driving forces. The exact value of the V_m would depend on the relative permeability of the membrane to K^+ and Na^+ . If the membrane was mostly permeable to K^+ ($P_K \gg P_{\text{Na}}$), the V_m would be closer to -100 mV ; if mostly permeable to Na^+ ($P_{\text{Na}} \gg P_K$) the V_m would be closer to $+60$. If the membrane was equally permeable to both K^+ and Na^+ ($P_K = P_{\text{Na}}$), the V_m would reach a steady-state value half way between E_K and E_{Na} (about -20 mV). Hence the V_m is a balance of the electrochemical driving forces for physiologically relevant ions, and the extent to which the membrane is permeable to each of these ions. This can be expressed mathematically using the Goldman-Hodgkin-Katz (GHK) equation (eqn [2]; Hodgkin and Katz, 1949). For further discussion of GHK and electrodiffusion equations applied to membrane potentials, see Hodgkin and Katz (1949), Keramidas *et al.* (2004), and/or Barry (2006). If the membrane is only permeable to a single ion, the GHK equation reduces to the Nernst equation.

$$V_m = RT/F \ln \left\{ \frac{P_{\text{Na}}[\text{Na}^+]_{\text{out}} + P_K[\text{K}^+]_{\text{out}} + P_{\text{Cl}}[\text{Cl}^-]_{\text{in}}}{P_{\text{Na}}[\text{Na}^+]_{\text{in}} + P_K[\text{K}^+]_{\text{in}} + P_{\text{Cl}}[\text{Cl}^-]_{\text{out}}} \right\} \quad [2]$$

$$V_m = RT/F \ln \left\{ \frac{[\text{Na}^+]_{\text{out}} + P_K/P_{\text{Na}}[\text{K}^+]_{\text{out}}}{[\text{Na}^+]_{\text{in}} + P_K/P_{\text{Na}}[\text{K}^+]_{\text{in}}} \right\} \quad [3]$$

$$V_m = RT/F \ln \left\{ \frac{[\text{Na}^+]_{\text{out}} + r.P_K/P_{\text{Na}}[\text{K}^+]_{\text{out}}}{[\text{Na}^+]_{\text{in}} + r.P_K/P_{\text{Na}}[\text{K}^+]_{\text{in}}} \right\} \quad [4]$$

Equations [2]–[4]. The GHK equation expressing the membrane potential (V_m) in terms of the permeabilities (P_x) for the most prevalent ions in physiological solutions (Na^+ , K^+ , and Cl^-). R , T , and F have their usual meaning and values as in Box 2. The subscripts in and out refer to intracellular and extracellular concentrations (or more correctly activities). The permeability constants give the velocity of ion movement through the membrane (cm per sec) and depend on the mobility and solubility of the ions in the membrane. P_{Na} , for example, $= RT/Fa \times v_{\text{Na}} \times K_{\text{Na}}$, where a = membrane thickness, v_{Na} = the mobility of the ion within the membrane, and K_{Na} is the partition coefficient between the membrane and aqueous solution. Assuming the membrane is only permeable to Na^+ and K^+ (a close approximation for a nerve axon) or that Cl^- is passively distributed (so $E_{\text{Cl}} = V_m$, example in some skeletal muscle fibers) then eqn [2] can be reduced to eqn [3] where P_{Na}/P_K is the relative Na^+ to K^+ permeability, a parameter that can be more readily

Box 2 Nernst potential and its derivation

The Nernst potential is derived considering the energy of ions inside and outside the cell at equilibrium – a special situation where there is no energy gradient between inside and outside. The free energy of a mole of ions includes components related to the chemical energy of the ion, the electrical energy acting on the ion and the standard energy on the ion in the absence of applied electrical or chemical energies, a property inherent to each ion species. This can be expressed as:

$$\mu_x = RT \ln[A_x] + zFV + \mu_o^X + PV_n$$

where μ_x = the free energy of the ion X, R = the gas constant (8.314 J K⁻¹ mol⁻¹, Joules per Kelvin per mole (as 1 V = 1 J C⁻¹, R can also be expressed in terms of VC per K⁻¹mol⁻¹)), T = temperature in Kelvins (K = °C + 273.15), A is the activity of ion X (which is related to molar concentration, C , by an activity coefficient, γ , specific for a particular salt solution: $A = \gamma \cdot C$), $\ln$ is the natural logarithm, z is the valency of the ion, F is Faraday's constant (96 485 C mol⁻¹, Coulombs per mole), V is the electrical potential (volts, or J C⁻¹), μ_o^X is the standard state potential for ion x, P is the hydrostatic pressure, and V_n is the volume of solution occupied by 1 mol of ions, the partial molar volume.

At equilibrium, the free energy of an ion on the outside will equal that of the ion on the inside of the cell so that:

$$\begin{aligned} \mu_{x \text{ outside}} &= \mu_{x \text{ inside}}; \text{ or} \\ (RT \ln[A_x] + zFV + \mu_o^X + PV_n)_{\text{outside}} \\ &= (RT \ln[A_x] + zFV + \mu_o^X + PV_n)_{\text{inside}} \end{aligned}$$

The standard state potential is the same inside and outside so can be deleted from both sides. The hydrostatic component is very small in animal cells (but can be large in plant cells) compared to the electrical and chemical forces on the ions (and is also roughly the same on both sides) so can also be eliminated from the equation. The activity coefficient for NaCl (0.75) is very similar to KCl (0.74) when the ionic strength of the intracellular and extracellular solutions are both at physiological values (~0.15 M; see, e.g., Barry *et al.*, 2013) and so can be eliminated from the equation and activity can be replaced by concentration.

So, after simplification and rearranging, our equation for the transmembrane ion energy difference at equilibrium becomes:

$$\begin{aligned} RT \ln[C_x]_{\text{outside}} - RT \ln[C_x]_{\text{inside}} &= zFV_{\text{inside}} - zFV_{\text{outside}} \text{ or} \\ zF(V_{\text{inside}} - V_{\text{outside}}) &= RT \ln[C_x]_{\text{outside}}/[C_x]_{\text{inside}} \text{ or} \\ zFV_m &= RT \ln[C_x]_o/[C_x]_i \end{aligned}$$

where V_m is the membrane potential defined as the inside potential relative to the outside potential, and the subscripts o and i refer to outside and inside, respectively.

This is the Nernst equation, where V_m is the membrane potential where ion X is at electrochemical equilibrium, often referred to as the Nernst potential or Equilibrium potential (E_x). In other words, V_m is the potential at which the chemical and electrical forces on an ion are exactly balanced.

experimentally measured. Equation [4] includes a coupling coefficient, r , to account for the contribution of the electrogenic Na⁺ pump (see text).

Steady State versus Electrochemical Equilibrium – the ‘Resting’ Membrane Potential

In most excitable or polarized cells (e.g., Figure 4(b(iii))), the resting V_m is about –70 mV, as the membrane permeability is dominated by K⁺ relative to other physiological electrolytes.

This high K⁺ permeability is due to most cells containing more open K⁺ channels (or ‘leak’ channels) at rest as compared to other types of ion channels. At a resting V_m of –70, neither K⁺ nor Na⁺ is at electrochemical equilibrium. For K⁺, the negative V_m that acts to attract K⁺ to the intracellular solution is insufficient to balance the chemical force driving K⁺ out. A relatively large number of open K⁺ channels combined with a small electrochemical driving force result in a continued small efflux of K⁺. For Na⁺, both the negative V_m and the concentration gradient provide a strong electrochemical driving force for Na⁺ influx. However, there are a relative small number of open Na⁺ channels and so only a small Na⁺ influx occurs. At rest when the V_m is not changing, we have a steady-state situation, in which the efflux of K⁺ is balanced by the influx of Na⁺. Changing the relative membrane permeability will disrupt this delicate balance. Opening of more K⁺ channels will allow K⁺ efflux to exceed Na⁺ influx and V_m becomes hyperpolarized. Conversely, closing some K⁺ channels, or opening of more Na⁺ channels, will result in less K⁺ efflux or more Na⁺ influx, respectively, and a depolarization of the V_m . Therefore by opening and closing K⁺ and/or Na⁺ channels, a cell can readily change its membrane potential. Cells also contain ion channels that allow other physiological ions such as Cl⁻, HCO₃⁻, and Ca²⁺ to diffuse across the membrane according to their electrochemical gradients. Hence by combining concentration gradients established by active transport mechanisms, and changes in relative membrane permeability by opening and closing an array of different types of ion channels, cells can change the V_m to a range of different levels and thereby induce and transduce physiologically relevant electrical signals. Such electrical signals include resting membrane potentials, action potentials, synaptic potentials, sensory, and receptor potentials.

Do Ion Concentrations Change When the V_m Changes?

A logical question to ask regards whether the efflux of K⁺ and influx of Na⁺ that occurs to establish the resting membrane potential changes the concentration gradients. Similarly with changes in V_m , such as action potentials and synaptic potentials, will these ion fluxes change the intracellular and extracellular ion concentrations? The answer is that only a relatively small number of ions flow across the membrane to generate the resting V_m or changes in V_m associated with action potentials or other membrane potential changes. The intracellular and extracellular concentrations do change, but for Na⁺ and K⁺ (and most other ions) this does not significantly change the concentration gradients. Box 3 below estimates that a single action potential in a typical nerve cell soma causes an increase in intracellular Na⁺ concentration of less than 1 μM, or <0.01% change from a resting concentration of about 18 mM. An even smaller relative change would occur with the efflux of K⁺ ions that generates the typical negative resting V_m . In a small diameter (1 μm) axon, a similar calculation gives a change of 0.2 mM Na⁺ during a single action potential. Prolonged high-frequency action potentials may elevate Na⁺ significantly in such small spaces (and some nerve axons or processes can be even thinner), although the Na⁺ pump is strongly stimulated by intracellular Na⁺ to counteract such buildup. During some pathological conditions such as seizures, the high frequency of action

Box 3 How many ions flow across the membrane to cause a voltage change of 100 mV?

An estimate of this can be done using some biophysical equations and approximated values. We consider the case of a neonatal rat motor neuron, for which we have reliable morphological and electrical data, with a somatic surface area of approximately $2 \times 10^{-5} \text{ cm}^2$ ($2000 \mu\text{m}^2$) and a total membrane capacitance of about 40 pF (Thurbon *et al.*, 1998).

$$\begin{aligned}\text{Charge (Q)} &= \text{Capacitance (Cm)} \times \text{Voltage (Vm)} \\ &= 40 \text{ p Farads} \times 100 \text{ mVolts} \\ &= 40 \times 10^{-12} \text{ F} \times 100 \times 10^{-3} \\ &= 4 \times 10^{-12} \text{ Coulombs (C; } 1 \text{ C} = 1 \text{ F.V.)}\end{aligned}$$

The charge on a single ion is the elementary charge, $1.6 \times 10^{-19} \text{ C}$. Hence $4 \times 10^{-12} \text{ C} = 4 \times 10^{-12} / 1.6 \times 10^{-19} = 2.5 \times 10^7$ ions.

1 mol of ions = 6.02×10^{23} ions (Avogadro's number) so 2.5×10^7 ions = 4.2×10^{-17} mol.

We will consider now that the charge carrier for the 100 mV V_m change is Na^+ , coming in to the neuron during an action potential. An efflux of K^+ to generate a -100 mV resting V_m would be the same charge movement, but would result in an even smaller relative change in the intracellular $[\text{K}^+]$ due to the higher resting $[\text{K}^+]$.

The surface area (SA) of a sphere is $4 \times \pi \times r^2$, the radius (r) is $\frac{1}{2} \times \sqrt{\text{SA}/\pi}$ while the volume is $\frac{4}{3} \times \pi \times r^3$. Assuming the soma equates to a sphere (an approximation), and neglecting the dendritic tree (for simplicity) the radius of the neuronal soma is:

$$\begin{aligned}\text{Radius} &= \frac{1}{2} \times \sqrt{\left[(2 \times 10^{-5} \text{ cm}^2) / \pi \right]} \\ &= 2.52 \times 10^{-3} \text{ cm (or } 25 \mu\text{m})\end{aligned}$$

$$\begin{aligned}\text{Volume} &= \frac{4}{3} \times \pi \times (2.52 \times 10^{-3} \text{ cm})^3 \\ &= 6.73 \times 10^{-8} \text{ cm}^3 \\ &= 6.73 \times 10^{-8} \text{ ml} \\ &= 6.73 \times 10^{-11} \text{ l}\end{aligned}$$

Hence a 100 mV action potential is an influx of 4.2×10^{-17} mol of Na^+ into a volume of 6.73×10^{-11} litres. Molar concentration is expressed as moles per litre, so:

$$\begin{aligned}&= 4.2 \times 10^{-17} \text{ mol} / 6.73 \times 10^{-11} \text{ l} \\ &= 0.62 \times 10^{-6} \text{ mol l}^{-1} \\ &= 0.62 \times 10^{-3} \text{ mmol l}^{-1} (\text{mM}) (\approx 0.0006 \text{ mM})\end{aligned}$$

Assuming an intracellular Na^+ concentration of 12 mM, a 0.0006 mM increase in Na^+ concentration is an increase of 0.005%, an insignificant change in total concentration.

potentials occurring synchronously in many neurons can also result in significant buildup of extracellular K^+ (Somjen, 2002). Usually extracellular K^+ increases will be buffered by surrounding glial cells or diffuse through the extracellular spaces, and saturation of this capacity can have detrimental effects such as helping to further propagate seizures and/or spreading depression (Somjen, 2002). Pathological changes in blood ion concentrations, often a result of kidney dysfunction or as a side effect of diuretics, can also result in changes in electrochemical driving forces and in the membrane potential. Chronic kidney disease, for example, is associated with hyperkalemia that results in chronic peripheral nerve depolarization and alterations in excitability that likely contributes to a range of uremic neuropathy symptoms (Krishnan and Kiernan, 2007).

Active Ion Transport via Primary and Secondary Active Transporters

For a specific ion to carry its charge across the membrane and change the membrane potential the membrane must be permeable to that ion. The membrane transporters that mediate movement of ions across the membrane fall into two broad categories: active and facilitated diffusion transporters (Table 2). Primary active transporters, such as the sodium pump ($\text{Na}^+/\text{K}^+/\text{ATPase}$) or the Ca^{2+} pump ($\text{Ca}^{2+}\text{-ATPase}$) directly utilize energy to establish ion concentration gradients. These gradients are used to fuel secondary active transport, in which the movement of one ion species down its electrochemical gradient is coupled to uphill movement of other ion species or solutes against their electrochemical gradient. These secondary transporters, such as the K^+/Cl^- cotransporter or the Na^+/H^+ exchanger, can also significantly influence the concentration gradients of ions that contribute to shaping the membrane potential. The K^+/Cl^- cotransporter and the $\text{Na}^+/\text{K}^+/\text{Cl}^-$ cotransporter are particularly important in determining Cl^- concentration gradients, as a primary active transport pump for Cl^- does not seem to exist in most cells. Some of these active transporters are 'electrogenic,' that is, they result in a separation of charge across the membrane in each transport cycle. The Na^+ pump, for example, exports 3 Na^+ ions and imports two K^+ ions in each transport cycle, utilizing one molecule of ATP in the process. Hence a net movement of 1 positive charge out of the cell occurs in each transport cycle resulting in an increase in negative voltage inside the cell. Blocking the Na^+ pump using ouabain has been reported to depolarize the V_m by 2–3 mV, as well as resulting in a gradual build up of intracellular Na^+ and depletion of intracellular K^+ (e.g., Thomas, 1972). The modest effects on V_m of the Na^+ pump as compared to ion channels relate to the 1000-fold or more lower rates of ion transport through the Na^+ pump (and other active transporters) as compared to ion channels. Hence their contribution to relative membrane permeability is also minimal. The Na^+ pump undergoes numerous conformational changes during the transport cycle, binding and trapping Na^+ and K^+ and undergoing phosphorylation and dephosphorylation in a series of discrete conformation states. Even when the Na^+ pump is strongly stimulated, the rate at which it moves Na^+ or K^+ across the membrane is of the order of 10^3 – 10^4 ions per second, whereas ion channels can mediate ionic fluxes of at rates of 10^6 – 10^8 ions per second (Hille, 1992). The small contribution that the electrogenic Na^+ pump contributes to the resting membrane potential can be accounted for by modifying the relative Na^+ and K^+ permeability using a coupling coefficient, the parameter r in eqn [4] (Mullins and Noda, 1963). This coefficient reflects the pumps transport stoichiometry, being 3/2 for the Na^+ pump. Using physiological values of Na^+ and K^+ concentrations and relative permeabilities, eqn [4] calculates that the Na^+ pump contributes a 2.0 mV hyperpolarization to the resting membrane potential.

Passive Ion Transport via Selective, Facilitated Diffusion Ion Channels

Although primary and secondary active transporters establish and maintain ion gradients, ion channels, being facilitated

Table 2 Classification and distinguishing characteristics of membrane transport processes

	Passive processes			Active processes		
	Facilitated diffusion			Primary active transport		
	Carrier	Channel		Co-transport	Counter transport	
Alternate name	Lipid diffusion	Diffusion carrier	Channels or pores	Pump	Uniporter	Exchanger
Energy used	Nil	Nil	Nil	Yes, direct	Yes, indirect	Yes, indirect
Chemical nature of transported substance	Nonpolar/lipid soluble molecules	Ions or other polar substances	Ions or other polar substances	Ions or other polar substances	Ions or other polar substances	Ions or other polar substances
Permeation pathway/mechanism	Via the lipid	Via membrane protein	Via membrane protein	Via membrane protein	Via membrane protein	Via membrane protein
Driving force for solute movement	Chemical (concentration gradient)	Chemical (concentration gradient)	Electrochemical (concentration gradient and electrical force)	ATP hydrolysis fuels uphill solute flux	Solute movement down electrochemical gradient drives another solute uphill in same direction	Solute movement down electrochemical gradient drives another solute uphill in opposite direction
Example of substances transported	O ₂ , CO ₂ , steroids	Glucose, amino acids	Na ⁺ , K ⁺ , Cl ⁻ , Ca ²⁺ , H ₂ O	Na ⁺ , K ⁺ , Ca ²⁺ , H ⁺	Glucose, amino acids, Cl ⁻	Ca ²⁺ , H ⁺
Example of transporter	Not applicable	GLUT4 glucose transporter	Voltage-dependent Na ⁺ channel	Na ⁺ /K ⁺ /ATPase (Na ⁺ pump)	GLUT1 glucose transporter	Na ⁺ /Ca ²⁺ exchanger
Example of physiological function	Oxygen flow from blood to cells	Glucose movement into cells following a meal	Generation and propagation of the action potential	Establish and maintain physiological [Na ⁺] and [K ⁺]	Store glucose into liver cells	Intracellular [Ca ²⁺] homeostasis

diffusion transporters, dissipate these ionic gradients. The study of ion channels has been remarkably advanced over the past two decades by the elucidation of the crystal structure of ion channels from bacteria and other simple organisms which can be grown to yield large quantities of membrane proteins amenable to crystallization. These ion channels have gene sequences and protein structures homologous to our own ion channels, so provide insights into mammalian ion channels. The importance of this work was recognized by the award of the Nobel Prize in chemistry in 2002, to Professors Rod Mackinnon and Peter Agre for their discoveries concerning K^+ channels and water channels, respectively (see Relevant Websites). The structure of the yeast KcsA K^+ channel discovered by Prof Mackinnon's laboratory in 1998 is shown in Figure 5 (a) (Doyle *et al.*, 1998). It illustrates some of the key structural features of ion channels – a 'selectivity filter' that allows certain ions to pass into the aqueous channel pore but not others, and a 'channel gate' that opens and/or closes in response to certain stimuli and thereby allows different changes in membrane permeability under different physiological conditions.

The selectivity properties of K^+ channels such as the KcsA channel are quite remarkable – they can allow K^+ ions to pass through at high rates of up to $1 \text{ million sec}^{-1}$ while the very similar and slightly smaller (when dehydrated) Na^+ ion is strongly excluded (the $K^+ : Na^+$ selectivity of some K^+ channels is 100:1 or more; Hille, 1992; Doyle *et al.*, 1998). Such high selectivity can be achieved as the dehydrated K^+ ion fits perfectly into the somewhat rigid selectivity filter, where it can be momentarily stabilized by interactions with negative dipoles from carbonyl oxygen groups arising from some of the amino acids that form the selectivity filter (Figure 5(b)). The energy lost by liberating the hydrating H_2O molecules can be regained by interactions with the carbonyl oxygens. Na^+ ions, in contrast, cannot interact as closely with the negative dipoles in the selectivity filter (the diameter of a bare Na^+ is slightly smaller than K^+ – 0.09 nm compared to 0.10 nm) and hence prefers to stay hydrated by H_2O . An important principle of relative ion selectivity is this balance between energy of dehydration and energy of interaction with residues within the selectivity filter. Voltage-dependent Na^+ , K^+ , and Ca^{2+} channels are, in general, very highly selective, but other channels have lower selectivity. The synaptic glutamate and acetylcholine receptors can pass both Ca^{2+} , Na^+ , and K^+ to various degrees amongst the different subtypes. Nicotinic acetylcholine channels in frog skeletal muscles do not discriminate much amongst monovalent cations and have a relative $Ca^{2+} : Na^+$ permeability of about 0.2 (Adams *et al.*, 1980) while the neuronal $\alpha 7$ subtype is more permeable to Ca^{2+} than Na^+ by about 10:1 (Bertrand *et al.*, 1993). Anion channels, such as gamma-aminobutyric acid (GABA) and glycine-receptor channels, are generally also weakly selective amongst different anions, and can even let a small degree of cations to pass through (Keramidas *et al.*, 2004; Franciolini and Nonner, 1987; Sugiharto *et al.*, 2008). Note in the KcsA channel the negatively charged residues surrounding the entrance to the channel pore. This illustrates another common theme in ion channel permeation – charged residues surrounding the entrances to the channel pore can concentrate permeant ions and hence increase the rate at which they can then flow through the channel pore (Moorhouse *et al.*, 2002; Carland *et al.*, 2009). These charged residues may also

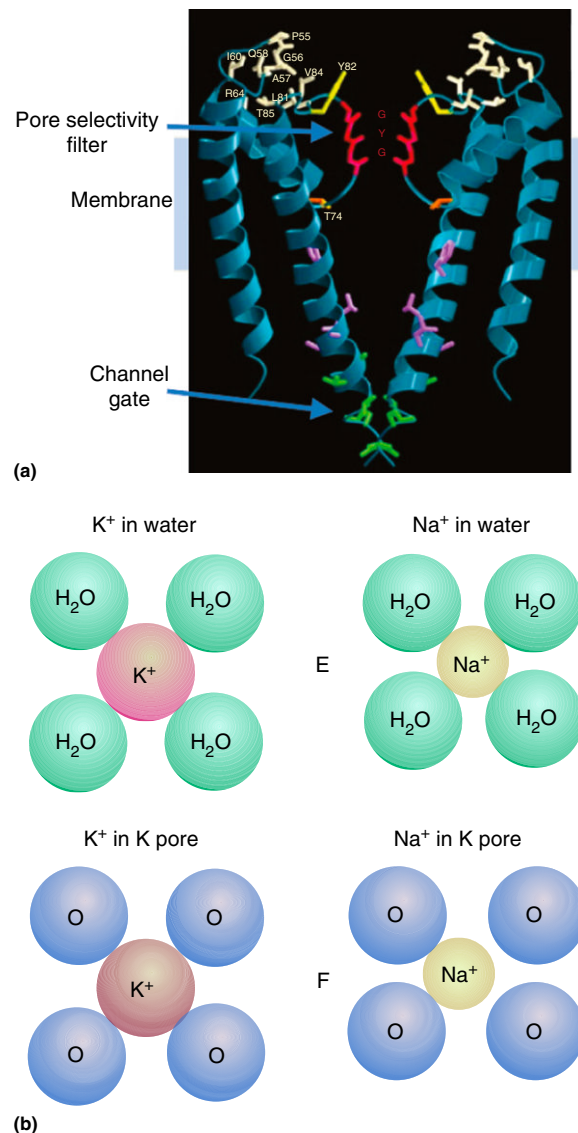


Figure 5 Structure of a K^+ -selective ion channel. (a) X-ray structure of the bacteria *Streptomyces lividans* KcsA K^+ channel (Doyle *et al.*, 1998). Two adjacent polypeptide subunits (of the four in total) are shown. The two transmembrane alpha-helical segments, and the angled pore-helix, are indicated in blue, some other amino acids are numbered and labeled with their single letter code. Of note is the 'GYG pore selectivity filter' that lines the top of the channel pore, and the 'channel gate' at the intracellular end. Adapted with permission from part of Figure 4 of Doyle, D.A., Morais, C.J., Pfuetzner, R.A., *et al.*, 1998. The structure of the potassium channel: Molecular basis of K^+ conduction and selectivity. *Science* 280, 69–77. (b) Schematic diagram illustrating energetics of why the larger diameter K^+ ion is selected over the smaller diameter Na^+ ion in the pore of the KcsA K^+ channel (panels C–F of Armstrong, C., 1998. *Science* 280 (5360), 56–57, with permission). Both ions are hydrated and stable when in the extracellular and intracellular solutions (and also in the central cavity of the channel). However, in the restricted region of the selectivity filter, the ions need to be dehydrated to pass through, K^+ 'fits' better in the selectivity filter of Na^+ , as it can be more effectively stabilized by the negative dipole of carbonyl oxygens from the pore structure.

contribute more directly to determining which type of ion – an anion or cation – is preferred through the pore (Keramidas *et al.*, 2002; Cymes and Grosman, 2012).

Ion Channel Gating

Figure 5(a) also illustrates the occlusion of the aqueous pore that forms the channel gate. The KcSA structure in Figure 5(a) represents a nonconducting state through which no ions can pass. A decrease in the pH of the bacteria cell induces conformational changes that include large opening at the inner end of the membrane domains and that ultimately results in a conducting channel (Liu *et al.*, 2001). For voltage-dependent channels, membrane potential changes are detected by specialized voltage-sensor domains that respond with conformational changes that initiate opening of the channel pore (Bezaniilla, 2000). In the cys-loop family of ligand-gated ion channels, the channel gate lies within the transmembrane pore that is more than 50 Å away from the specific ligand-binding domain – yet when the neurotransmitters bind to this site, a conformational wave is initiated that opens the channel in less than 0.1 ms (Sine and Engel, 2006). Membrane stretch is another way of activating channels, as well as being able to change the sensitivity of channels to other forms of gating. Other channels are ‘leaky,’ or open in absence of any particular stimuli. These channels include the K^+ selective channels referred to above in the context of setting the resting membrane potential. Ion channels, however, do not spend all their time in just one conformational state and may spontaneously switch between closed and open conformations. Some resting ‘leak’ channels, for example, the inward rectifying K^+ channels can be gated to either increase or decrease the amount of time spent in the open conformation: binding of G-protein subunits can open channels of the Kir gene family 3, while K^+ channels of the Kir family 6 decrease their open time when ATP levels in cells increase (Reimann and Ashcroft, 1999). Furthermore, channels have multiple gates – voltage-dependent channels, for example, activate in response to a depolarization and then inactivate during sustained depolarization via a second distinct ‘inactivation’ gate that occludes the open pore. This fast inactivation process is important for rapid electrical signaling; toxins or genetic mutations that disrupt the inactivation gate of voltage-dependent Na^+ or Cl^- channels can result in muscle diseases and epilepsy (Ashcroft, 2000). In an analogous fashion, some ligand activated channels can also transit to a closed conformation in the sustained presence of ligand through a process known as ‘desensitization.’ The precise nature of this desensitization gate remains unknown. Clearly the modes and mechanisms of channel gating are numerous and complex, yet enables changes in membrane permeability at different times in response to different physiological stimuli that result in a myriad of membrane potential changes such as action potentials, synaptic potentials, sensory potentials, and more.

Passive Membrane Potential Changes

The membrane potential changes underlying an action potential, generated in response to sensory stimulation, or arising in response to release of neurotransmitters at a synapse, all involve ionic currents flowing across the membrane through different

ion channels. Membrane potential changes can also occur in the absence of a specific activating stimulus, but rather arise ‘passively’ (although ions also flow passively through channels) or ‘electrotonically’ by conduction between two regions of different potential within the cell. A geometrically simple cell, such as a small and round neuron without much processes, or a red blood cell, is isopotential, meaning that the membrane potential at each point inside the cell is the same. For cells of more complex geometry, such as an elongated muscle cell, a long nerve axon, or a nerve cell with extensive dendritic elongations, the V_m at one point of the cell can differ from that at other points of the cell. Consider a motoneuron innervating a muscle in the foot: an action potential depolarizes the soma membrane within the ventral horn of the spinal cord, but the neurotransmitter that initiates the muscle response is packaged in a nerve terminal up to 1 m away. When the soma is depolarized by the action potential, the V_m at the nerve terminal is still at rest, and the depolarization needs to spread along the nerve cell to the terminal to initiate neurotransmission. Similarly, a layer V pyramidal neuron gets depolarized at a distal synapse in a different layer of the cortex, and this V_m change needs to spread to the soma where an action potential response could potentially result. In both these cases, the V_m is not the same at these two points in the same cell, creating a potential difference between these two points that drives the movement of charge, or current. Cations will flow along the core of the dendrite or axon from the depolarized region to the more hyperpolarized region (with anions flowing in the opposite direction). As the positive current moves along the cell membrane, it causes a spread of depolarization. This local current flow is ‘passive’ or ‘electrotonic’ propagation, and works just as effectively for the spread of either depolarizations or hyperpolarizations. The distance and speed at which this passive depolarization (or hyperpolarization) is propagated along a cell membrane depends on (1) the amplitude of the difference in potential between the two points, (2) how many ions will ‘leak’ out across the plasma membrane, relative to how many will flow through the core of the axon, dendrite, or other excitable cell, and (3) how quickly will this movement of charge causes a change in the V_m . Figure 6 demonstrates the decay of these passive responses as one gets further away from the point at which the membrane is most polarized. Figure 3 demonstrates how the passive response recorded at a single point in response to a step change in current changes over time. The two key intrinsic or passive properties of the cell membrane that control these temporal and spatial aspects of the membrane potential are the ‘space constant’ and ‘time constant’ – and these can vary markedly across cells of different shapes and sizes. A series of cable equations relate the resistances, capacitances, and geometries of the membrane and interior of a nerve ‘cable’ to the distance and speed a passive propagation can travel (Figure 6; see also Aidley, 1989; Boron and Boulpaep, 2009; Byrne, 2015). Over short distances of less than 1 mm, from the membrane of a synaptic cleft to the extrasynaptic membrane in a skeletal muscle, or from one Node of Ranvier to another in a myelinated nerve axon, passive depolarizations can be effective. But they will not suffice to propagate a depolarization in a long motoneuron nor even in a long dendritic tree. For long distance spread of membrane potentials, active depolarizations are required. In nerve axons, local (passive) currents spread the depolarization to an adjacent region

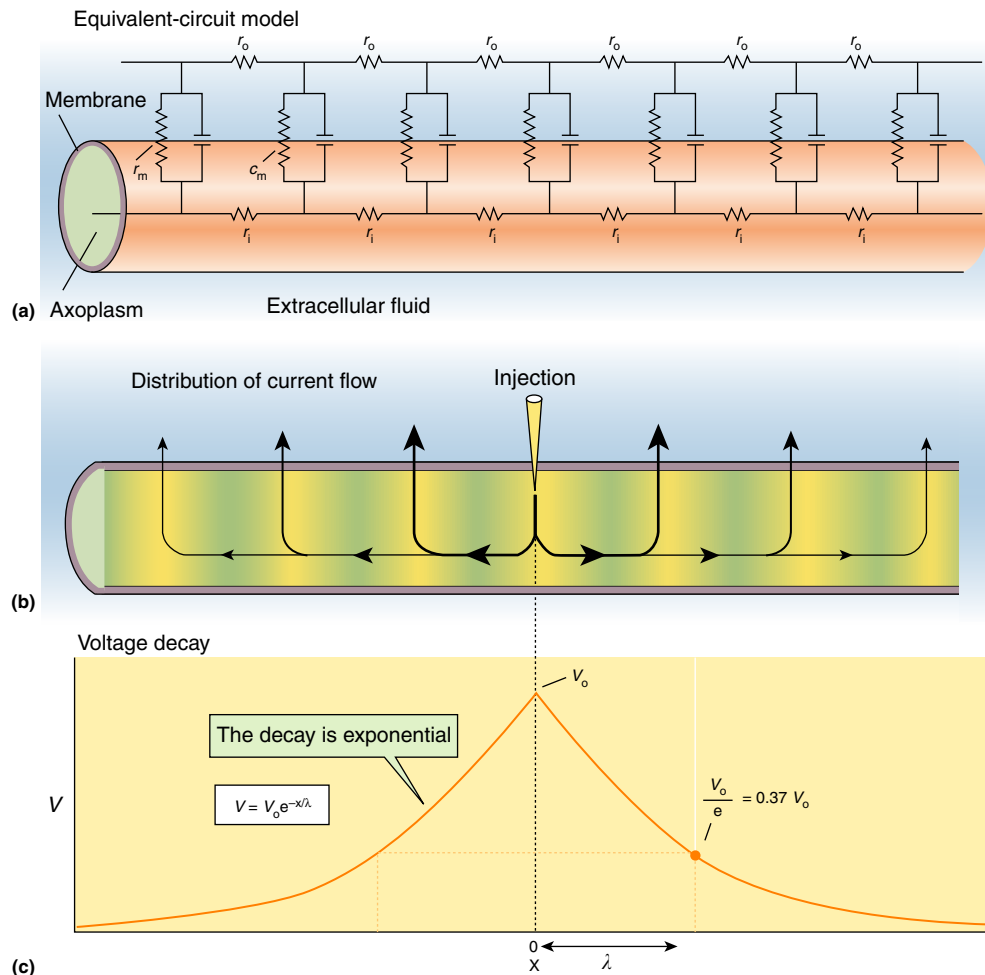


Figure 6 Passive membrane properties of a cylindrical cell. (a) Schematic diagram of a nerve axon, indicating the equivalent electrical properties of axoplasmic or internal resistance, r_i , an external resistance (r_o), a membrane resistance (r_m), and membrane capacitance (c_m) (see also Figure 3). These values defined for a unit length of a cylinder or cable can be converted to those of a unit area by incorporating the axon radius, a , so that $r_i = R_i/\pi a^2$ ($\Omega \text{ cm}^{-2}$), $r_m = R_m/2\pi a$ ($\Omega \text{ cm}^{-2}$), and $c_m = C_m \times 2\pi a$ ($\mu\text{F cm}^{-2}$). (b) As ionic current is injected into a particular point along a cylindrical cell via a microelectrode (or via the opening of an ion channel) the current passively spreads laterally. As the cell membrane is an imperfect insulator, some of the current that was injected 'leaks' across the membrane, meaning that less and less current is available to continue to spread along the axon. (c) A consequence of this leak of current is that the voltage response, or V_m , exponentially decreases in magnitude as distance from the source of the current increases. The decay depends on the amount of current that stays inside the fibre (or does not leak out), which depends directly on r_m and inversely on the sum of r_i and r_o . One can define a parameter, the space constant, λ as $\lambda = \sqrt{(r_m/r_o + r_i)}$, and show (e.g., Aidley, 1989, p. 37) that this equals the distance over which the response decays to 37% of its original or peak value (V_o). The voltage at distance x , $V = V_o e^{-x/\lambda}$. As $r_o \ll r_i$, one can simplify as $\lambda = \sqrt{(r_m/r_i)}$. For a unit area of axon, $\lambda = \sqrt{(aR_m/2r_i)}$; and hence λ increases as axon diameter increases (see Aidley, 1989, p. 38). This is particularly important for the propagation of the action potential along an axon, as the conduction velocity is directly proportional to the λ (and inversely proportional to the time constant τ , the product of $R_m \times C_m$). Reproduced from Figure 7.22 in Boron, W.F., Boulpaep, E.L., 2009. Medical Physiology. A Cellular and Molecular Approach, second ed. Philadelphia, PA: Saunders Elsevier. © 2002 Elsevier Science, USA.

where the depolarization crosses a threshold and activates the voltage-dependent Na^+ and K^+ channels to regenerate the action potential (that again spreads passively to the next region to be regenerated and so on).

Measuring Membrane Potential and Relative Membrane Permeability

Direct Measurements of Membrane Potential and Membrane Selectivity

The membrane potential is directly measured by inserting an electrode into the inside of a cell and recording the potential

difference between this electrode and one in contact with the solution outside the cell. This was first done for nerve cells by Hodgkin and Huxley (1939) who lowered a wire inside the giant axon of the squid, *Loligo forbesi*, to 1st observe the nerve membrane potential both at rest and in response to stimulation. The experimental approach is shown in a lovely video produced at the Marine Biology Laboratory in the 1970s and posted on youtube (see Section Relevant Websites). A more complete description of this experimental approach and the results was subsequently published in the *Journal of Physiology* (Hodgkin and Huxley, 1945) and is illustrated in Figure 7. The resting V_m of about -50 mV became positive (about $+40 \text{ mV}$) during

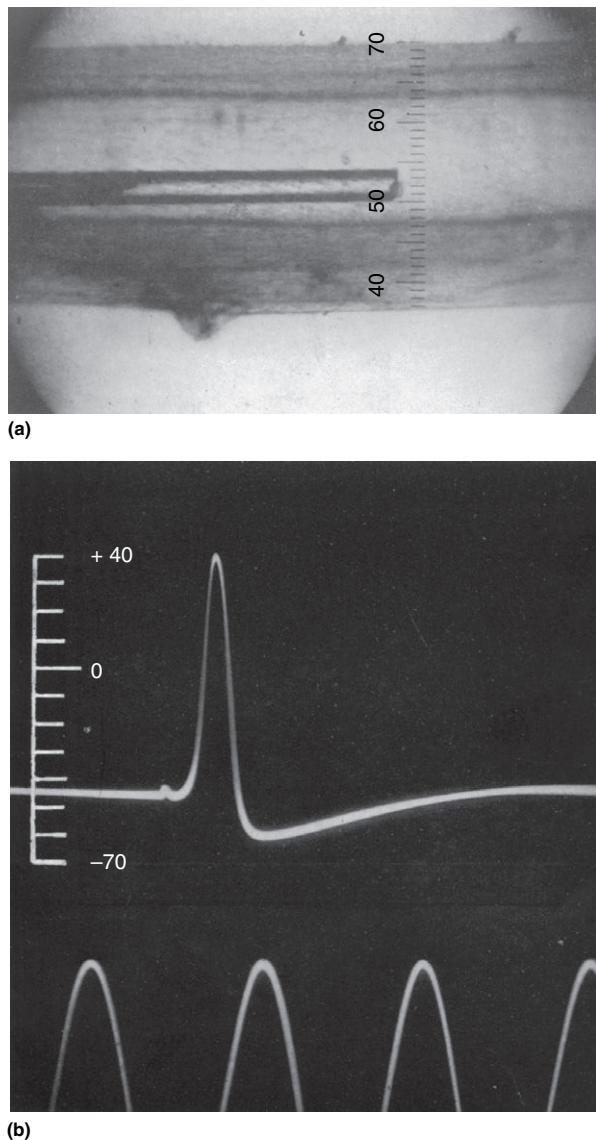


Figure 7 Direct recording of a nerve membrane potential using an intracellular microelectrode. (a) Photomicrograph of portion of the isolated giant axon of the squid, *Loligo forbesi*, with a 20 μm silver wire enclosed within a ≈100 μm glass capillary that is visible inside the ≈0.5 mm fibre (each small division represents 33 μm). (b) Direct recording of the V_m using the configuration shown in (a). The potential difference is between the silver wire microelectrode and a ground electrode connected to the seawater outside the nerve. The y axis is in mV, the time marker is cycles at 500 Hz (or 1 peak per 2 ms). Reproduced with permission from panel a, Plate 1 and Plate 2 of Hodgkin, A.L., Huxley, A.F., 1945. Resting and action potentials in single nerve fibres. *Journal of Physiology* 104, 176–195.

the action potential – this observation that the V_m overshoot zero revealed important insights into the mechanisms of the action potential. Their subsequent application of the voltage-clamp technique, coupled with rigorous and thoughtful analysis, elucidated the basic biophysical mechanisms of the resting and action potential in nerve axons. They provided a highly influential quantitative description of the membrane potential

changes in terms of changes in the membrane conductances to Na^+ and K^+ , and the relative membrane permeabilities of the membrane to these ions. On the 50th anniversary of the publication of a series of five seminal papers in the *Journal of Physiology* (1952), a special celebratory issue was published that describes the impact and legacy of their contribution (see Section Relevant Websites).

Using this intracellular recording wire electrode and squid axons, Hodgkin and Katz (1949; also analyzing earlier data of Curtis and Cole) investigated how the resting V_m and the peak amplitude of the action potential varied as a function of external K^+ and Na^+ . They demonstrated that the resting V_m was sensitive to K^+ changes but not Na^+ changes, and vice versa for the V_m at the peak of the action potential. Fitting the data to the GHK equation gave relative permeabilities, $P_{\text{K}}/P_{\text{Na}}$, of about 25:1 for the resting V_m , and about 1:20 for the peak of the action potential. However, mammalian cells are much smaller than these giant squid axons and not amenable to insertion of a fine wire. Using instead a sharp glass microelectrode (with a tip of <0.1 μm), Hodgkin and Horowitz (1959) conducted similar experiment on isolated skeletal muscle fibers of the frog, measuring resting V_m and changing the external K^+ , Na^+ , and Cl^- concentration in various combinations. The resting V_m was similarly highly dependent on external K^+ , with a linear relationship over most of the concentration ranges tested, but with a clear deviation from this relationship at lower extracellular K^+ (see their data in Figure 8). This deviation is consistent with the membrane being not only permeable to just K^+ , and fitting the data to the GHK equation gave a $P_{\text{Na}}/P_{\text{K}}$ of 0.01 for the resting V_m (Figure 8). When the external $[\text{K}^+]$ is low, the V_m is very hyperpolarized so that the electrochemical gradient for Na^+ influx is very strong (while that for K^+ is very weak), and the deviation occurs because we now see a significant effect of the Na^+ influx on V_m , despite the low relative Na^+ permeability. This work also appreciated (and quantified) that muscle had a high resting Cl^- conductance and permeability, but that the steady-state V_m was independent of changes in $[\text{Cl}^-]_o$ (Hodgkin and Horowitz, 1959). This seeming contradiction arises because Cl^- is not actively transported across the membrane of skeletal muscle and therefore distributes its concentration gradient passively with V_m so that $E_{\text{Cl}} = V_m$. When muscle V_m changes, the high Cl^- conductance of muscle plays an important role in reducing the extent of V_m changes and stabilizing muscle excitability. During an action potential, for example, Cl^- influx into the depolarized muscle is important for repolarization. Indeed, recessive or dominant mutations in the major human skeletal muscle Cl^- channel (CLC1) can cause severe muscle diseases known as generalized myotonia (Becker's disease) and myotonia congenita (Thomsen's disease), respectively (Ashcroft, 2000).

Quantifying Membrane and Ion Channel Selectivity Using Voltage-Clamp

In neurons, where active transporters for Cl^- typically exist (e.g., KCC2 and NKCC1), K^+ , Na^+ , and Cl^- (and to some extent other ions such as Ca^{2+} and HCO_3^-) all contribute to the V_m . The extent to which they do so varies considerably – as the membrane permeability to these ions is constantly

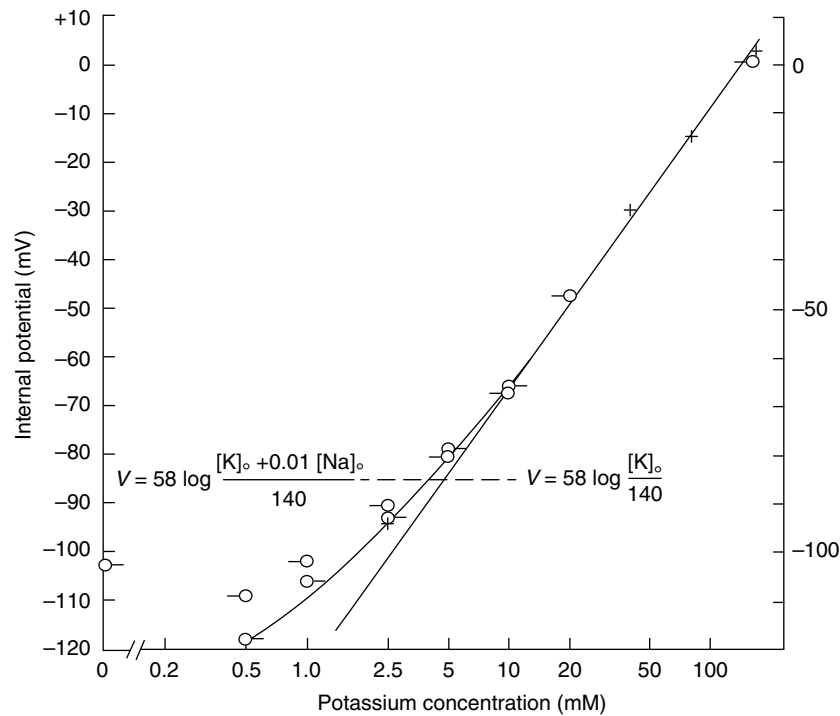


Figure 8 Quantifying the relative permeability of the cell membrane at rest. Direct recording of the resting V_m of frog skeletal muscle using an intracellular microelectrode, and the dependence of this resting V_m on the concentration of K^+ in the external solution. The experimental data points at different times after solution change (and/or after decreases and increases in K^+) are given by the open circles and crosses. This data is fit by the Nernst equation (right straight line) and by the GHK equation (left curved line). The better fit is obtained with the GHK equation assuming a relative $P_{Na}:P_K$ value of 0.01. Reproduced with permission from Figure 5 of Hodgkin, A.L., Horowicz, P., 1959. The influence of potassium and chloride ions on the membrane potential of single muscle fibres. *Journal of Physiology* 148, 127–160.

changing. Knowing what ion channels are opened by different stimuli, and the ion selectivity properties of these channels – that together determines the membrane permeability – is an important aspect of understanding and predicting the different types of membrane potential changes seen *in vivo*. Measuring the ionic selectivity of a specific ion channel, or of the membrane as a whole, involves the same basic experimental approach as illustrated by the early recordings of Hodgkin, Huxley, and their contemporaries – changing the external solutions and measuring the V_m . However, this is more precisely done using the voltage-clamp technique. Direct V_m measurements can be complicated by the fact that changes in the membrane potential can themselves change the channels that are open or the driving forces for ion movement. Greater experimental control is achieved using voltage-clamp recordings – an experimental approach where one controls the voltage by using a feedback amplifier to inject positive or negative current into the cell. The experimenter therefore controls the voltage and measures the current flowing through any channels open at that voltage. To apply to selectivity measurements, one first records the voltage at which no ionic currents flow across the membrane, the current reversal potential (V_{rev}), and then changes the solutions on one side of the membrane, and measures the new V_{rev} . The V_{rev} is the voltage where the ionic currents are at equilibrium. With knowledge of the ionic composition of the solutions, one can calculate the relative permeability of the membrane using the GHK equation as described above. If a single ion channel is responsible for the

current being measured (either because it has been over-expressed in a cultured cell, or because it has been selectively activated), then these relative permeability measurements correspond to the selectivity properties of the channel itself. Good solution and voltage control, and accounting for various voltage offsets inherent in these recordings (such as liquid junction potentials) is needed to accurately quantify membrane or channel selectivity. Using ionic activities (rather than concentrations) is particularly important when the control and test solutions have different ionic strengths, such as in dilution potential experiments often used to quantify selectivity. The procedures to accurately measure and quantify membrane selectivity are described in more detail elsewhere (Barry, 2006; Sugiharto *et al.*, 2008; 2010; Moorhouse *et al.*, *in press*), and the results from such an experiment are illustrated in Figure 9.

Caveats in Measuring Physiological V_m

Use of a microelectrode or patch clamp pipette to measure the potential difference between electrodes inside and outside of the cell is a direct and simple way to measure the membrane potential, but is associated with a number of limitations. Inserting a sharp microelectrode through the cell membrane causes some damage or holes through the membrane. These holes are nonselective ‘leak’ conductances that will change the composition of the cell and thereby potentially affect V_m . In hippocampal neurons, this damage-induced leak

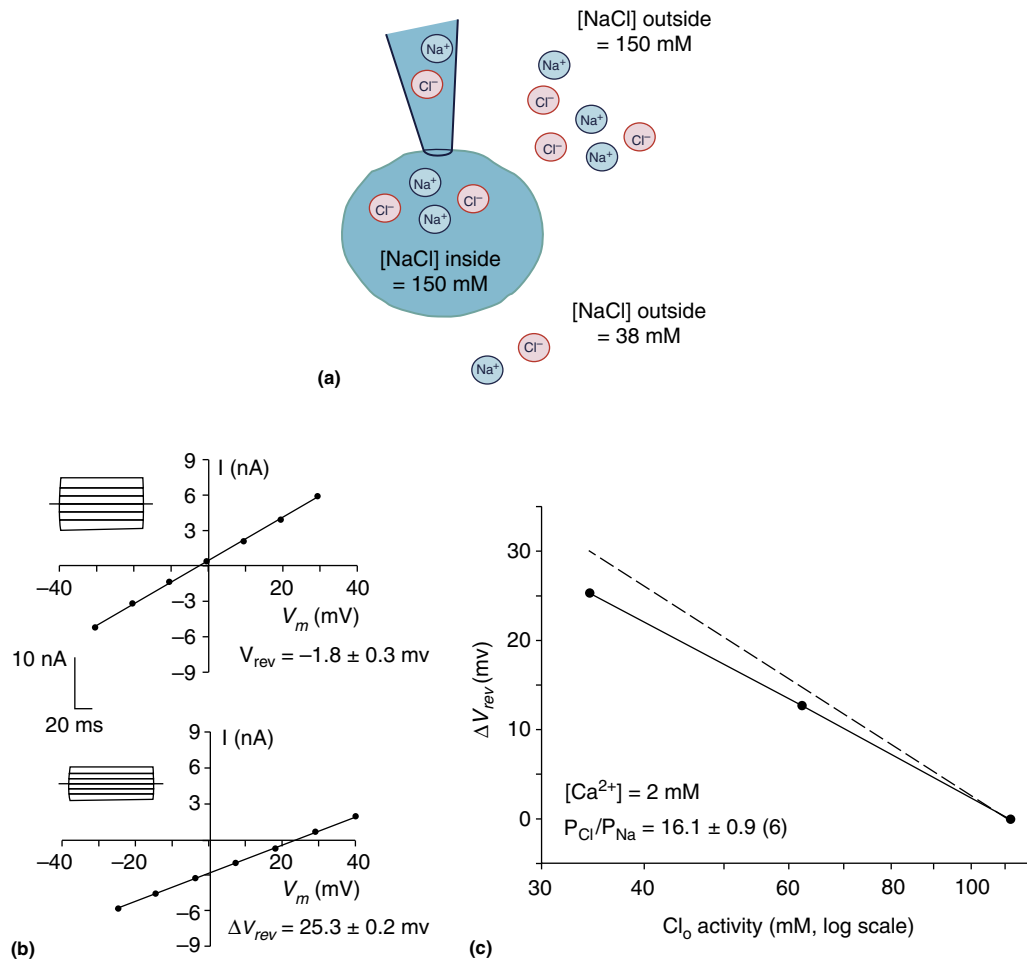


Figure 9 Quantifying ion channel selectivity using whole-cell patch clamp recordings. (a) Schematic diagram of the whole-cell patch clamp recording that allows the cell to be perfused with an intracellular solution of known ionic concentration, in this case largely 150 mM NaCl. The external solution can also be controlled, and changed from a symmetrical 150 mM NaCl (upper) to a 25% diluted solution (containing 38 mM NaCl; lower). The cell can also be voltage-clamped by the patch clamp configuration, allowing the experimenter to control the V_m and record the currents flowing across the membrane. (b) Example of glycine-receptor activated ionic currents recorded at different membrane potentials using whole-cell configuration as in (a), in symmetrical (upper) and in 25% dilution NaCl solutions (lower). The inserts show ionic currents in the presence of glycine, at different V_m s. The graphs plot these V_m s vs. current, I - V , relationships with the averaged V_m at which the current was zero (V_{rev} s) shown in the two graphs. (c) The averaged shift in V_{rev} as external NaCl is diluted to 50% and then 25% across six different cells is plotted against the external Cl activity, and fit with the GHK equation to obtain a P_{Cl}/P_{Na} value of about 12 for the recombinant human $\alpha 1$ glycine receptor. For further detail see Sugiharto *et al.* (2010). As the external NaCl is diluted, a chemical gradient is created for efflux of both Na^+ and Cl^- . As the GlyR is more permeant to Cl^- , a positive V_m is needed to counterbalance this chemical force, and hence V_{rev} shifts to positive values. A negative shift is seen if $P_{Na} > P_{Cl}$. Panels (b) and (c) have been published in different formats as parts of Figures 1 and 2 in Sugiharto *et al.* (2010).

causes the subsequent activation of ion channel conductances that reverse near the resting V_m , hence the cell's resting input resistance is changed, but there is not so much change in the resting V_m (Staley *et al.*, 1992; Spruston and Johnston, 1992). This damage is minimized with patch clamp recordings in which the membrane is sealed against a glass pipette, rather than the membrane pierced by the pipette. The resting cell input resistance (a measure of such 'leak') is 3–10-fold higher when the same neuron is recorded using patch clamp electrodes as compared to sharp microelectrodes (Staley *et al.*, 1992; Spruston and Johnston, 1992). So in this respect, a more physiological V_m value could be obtained with patch clamp recordings. A caveat with both patch clamp and sharp

microelectrodes is that ions leaking from the recording pipette into the cell will change the intracellular ionic composition and hence the physiological V_m . While this is much more significant for large patch clamp electrodes *cf* sharp microelectrodes, it can be better defined with patch electrodes. Indeed in standard whole-cell patch clamp recordings, the cell interior is almost completely (with time) dialysed by the pipette solution, so one can try to replicate the ionic concentrations and include other constituents to match the physiological milieu as close as possible. Perforated patch clamp recordings can maintain more closely the physiological ion concentrations. This variant of patch clamp recording typically uses different antibiotics inside the pipette solution that forms membrane pores with

permeability limited to ions within a restricted size or polarity, thereby enabling electrical continuity with the intracellular solution while maintaining the physiologically relevant signaling molecules and other larger constituents (Ishibashi *et al.*, 2012). Gramicidin perforated patch, for example, are impermeable to anions and hence have been useful for recording physiological responses to Cl-selective ion channels such as GABA and glycine-activated receptor-channels. While perforated patch can reduce inaccuracies in measuring physiological V_m related to disrupted intracellular cytoplasm, the technique is associated with a poorly defined voltage offset, relating to the fact that the larger charged molecules left inside the cell may not be properly balanced by charges inside the pipette solution (Horn and Marty, 1988). The liquid junction potential arising from the interface between pipette and bath solutions during the patch clamp procedure, that is between 3–8 mV with typical experimental solutions, must also be taken into account (Barry and Lynch, 1991). Furthermore, one must also appreciate that the voltage-clamp technique can only accurately control the membrane voltage of the membrane isopotential with the pipette – such as the soma in a neuron, while distal regions of cells with complex geometry – such as neuronal dendrites – are less-well controlled, particularly when the voltage or currents across the membrane are rapidly changing (Spruston *et al.*, 1993). Finally, much of our data on membrane potentials comes from *in vitro* conditions where tissues are more readily accessible and solutions and voltage more easily controlled, one must also appreciate that this situation may be quite different from *in vivo* recordings. Indeed, the V_m fluctuates much more widely in *in vivo* conditions, with much more influences of synaptic and neural circuit activity.

Physiological Values of V_m

Despite the caveats above, one should not be discouraged from measuring directly V_m using electrophysiological approaches – it is after all a necessary step in quantifying and understanding brain and body function and mechanisms. Using electrophysiological approaches, a wide range of resting V_m values have been reported for neurons (Table 3). Some of this variability may be technical: for example, Tyzio *et al.* (2003) have demonstrated that in small-sized neurons, a lower resistance pipette-neuron seal, and the procedure of going whole-cell can both depolarize the V_m somewhat. Avoiding this using a cell-attached variant gave a resting V_m of about -77 mV, both *in vitro* and *in vivo* (Tyzio *et al.*, 2003, 2008). However, even more hyperpolarized resting V_m have been reported using standard patch clamp recordings, suggesting minimal artifactual depolarizations in these studies (Table 3). Table 3 is populated with studies where V_m has been recorded in the same conditions but for different experimental techniques. A number of noteworthy points are: (1) the V_m can vary markedly in different cell types, both *in vitro* and *in vivo* (e.g., 15 mV difference in averaged V_m of mitral cells and granule cells within the olfactory bulb), (2) there seems little effect of age and preparation type when V_m is recorded using the same conditions (e.g., Tyzio *et al.* papers report approximately the same V_m value in slices, in culture, *in vivo*, and across development), and (3) the results reported by different groups using the same preparation can vary. The variability seems particularly large *in vivo*, perhaps not surprising considering the large spontaneous V_m fluctuations reported. In the somatosensory

Table 3 Some reported resting neuronal V_m values, recorded under different conditions to illustrate potential sources of variability

Neuron type	Preparation	Technique	V_m	Reference
<i>In vitro</i>				
Hippocampus DG ^a	Brain slice	WC ^c patch clamp	-84 ± 1.0 (31)	Staley <i>et al.</i> (1992)
Hippocampus DG	Brain slice	Sharp Microelectrode	-74 ± 2.1 (11)	Staley <i>et al.</i> (1992)
Hippocampus CA1	Brain slice	Perforated patch	-64 ± 2 (12)	Spruston and Johnston (1992)
Hippocampus CA3	Brain slice	Perforated patch	-66 ± 1 (12)	Spruston and Johnston (1992)
Hippocampus DG	Brain slice	Perforated patch	-74 ± 2 (12)	Spruston and Johnston (1992)
Hippocampus CA3	Brain slice P2	Perforated patch	-48 ± 3 (6)	Tyzio <i>et al.</i> (2003)
Hippocampus CA3	Brain slice P13–14	Perforated patch	-67 ± 2 (12)	Tyzio <i>et al.</i> (2003)
Hippocampus CA3	Brain slice P2	Cell-attached patch	-77 ± 2 (9)	Tyzio <i>et al.</i> (2003)
Hippocampus CA3	Brain slice P13–14	Cell-attached patch	-77 ± 2 (6)	Tyzio <i>et al.</i> (2003)
Hippocampus CA3	Cultured neurons, 16–21 days	Cell-attached patch	-76 ± 2 (8)	Tyzio <i>et al.</i> (2008)
Hippocampus CA3	Brain slice P2–3, 5 G Ω seal	WC patch clamp	-50 ± 3 (10)	Tyzio <i>et al.</i> (2003)
Hippocampus CA3	Brain slice P2–3, 14 G Ω seal	WC patch clamp	-65 ± 3 (10)	Tyzio <i>et al.</i> (2003)
Hippocampus CA3	Cultured neurons, 16–21 days	WC patch clamp	-56 ± 3 (72)	Tyzio <i>et al.</i> (2008)
<i>In vivo</i> ^b				
Thalamic neurons (VPML)	<i>In vivo</i> , anaesthetized	WC patch clamp	-65.1 ± 4.3 (31)	Margrie <i>et al.</i> (2002)
Layer 4 cortex	<i>In vivo</i> , anaesthetized	WC patch clamp	-81.8 ± 6.3 (23)	Margrie <i>et al.</i> (2002)
Layer 2/3 cortex (excitatory)	<i>In vivo</i> , anaesthetized	WC patch clamp	-83.8 ± 5.2 (31)	Margrie <i>et al.</i> (2002)
Layer 2/3 cortex (excitatory)	<i>In vivo</i> , awake (quiet)	WC patch clamp	-58 ± 1 (32)	Gentet <i>et al.</i> (2010)
Layer 2/3 cortex (inhibitory)	<i>In vivo</i> , awake (quiet)	WC patch clamp	-52 ± 1 (62) ^d	Gentet <i>et al.</i> (2010)
Olfactory bulb granule cells	<i>In vivo</i> , anaesthetized	WC patch clamp	-72 ± 2 (42)	Margrie <i>et al.</i> (2002)
Olfactory bulb mitral cells	<i>In vivo</i> , anaesthetized	WC patch clamp	-57 ± 2 (48)	Margrie <i>et al.</i> (2002)

^aHippocampal subfields as dentate gyrus (DG) and Cornu Ammonis (CA) 1 and 3.

^b*In vivo* during quiet wakefulness, average of both the up and down states but during relative nonspiking periods;

^cWC = whole-cell.

^dThis value combines the reported V_m for fast spiking and nonfast spiking groups.

cortex of the awake, behaving mouse, the 'resting' V_m oscillates at a frequency of 1–5 Hz and with an amplitude of up to 20 mV (Poulet and Petersen, 2008; Gentet *et al.*, 2010). During active behaviors, such as whisking, the V_m oscillations in corresponding cortical neurons are reduced and the baseline V_m becomes a few mV more depolarized (Poulet and Petersen, 2008). Averaging the V_m fluctuations over a few seconds gives values of -60 mV to -50 mV, but the exact value depends very much on the oscillation frequency and the brain state when recordings are measured. Furthermore, the frequency and amplitude of the currents responsible for these oscillations are markedly affected by different anesthetics (Doi *et al.*, 2007). Indeed, the concept of a 'resting V_m ' is perhaps meaningless for some types of neurons *in vivo*, highlighting the dynamic nature of neuronal activity that can be obscured *in vitro*.

In summary, one cannot give a uniform and single answer to 'What is the resting V_m ?' In neurons, it may be as depolarized as -50 mV, and as hyperpolarized as -85 mV. In less dynamic and more homogeneous cell types, it may be more consistent. *In vivo*, V_m may never be 'resting.' Clearly it depends on the region examined, but can also be affected by the recording approach. With careful consideration of voltage offsets, solutions and other potential caveats described above, measurements can provide good estimates of physiological V_m . While caution is needed in comparing across different conditions, changes in V_m induced by some physiological or experimental manipulation can be quite valid and informative.

Conclusion

This article provides an overview of the membrane potential – the polarization of cell membranes arising from a different distribution of ionic charges between the intracellular and extracellular surface of the cell membrane. Some physical and chemical properties of the ions and the membrane were introduced. The lipid nature of the membrane provides an energetically high barrier to passage of hydrated ions, making it a good separator of charge. Active transporters use energy, directly or indirectly, to establish ion concentration gradients and to move ions and other solutes against their energy gradients. Ions can rapidly cross the membrane via ion channels that contain integral membrane pores allowing ions to move passively down their electrochemical gradients. Ion channels have selectivity filters that control the flow of ions across the membrane, and gates that can be opened and closed in response to different physiological stimuli. The nature and number of ion channels open gives rise to the selectivity of the membrane for particular ions, and it is the combination of the relative membrane selectivity to different ions and the electrochemical driving force for each ion that sets the membrane potential at any particular time. Quantifying the relative selectivity of the membrane, or of individual channels, is a key component of understanding the molecular and cellular mechanisms of membrane potentials – this can be readily undertaken but one must be aware of various caveats and offsets involved in such measurements. Similarly, careful consideration is needed to accurately measure membrane potentials. However, perseverance to quantify and understand ion selectivity and membrane potentials will be rewarded, as changes in V_m are the primary means by which

cells mediate their function and communicate and coordinate with other cells in our body.

Acknowledgments

Research in the author's laboratory referred to above has been gratefully funded by UNSW Faculty Research Grants, the Australian Research Council, and the National Health and Medical Research Council of Australia. The author thanks Professor Peter Barry for helpful discussions and for the kind permission to use his schematic in Figure 3(b).

References

- Aidley, D.J., 1989. *Physiology of Excitable Cells*, third ed. Cambridge: Cambridge University Press.
- Adams, D.J., Dwyer, T.M., Hille, B., 1980. The permeability of endplate channels to monovalent and divalent metal cations. *Journal of General Physiology* 75, 493–510.
- Ashcroft, F.M., 2000. *Ion Channels and Diseases*. San Diego, CA: Academic Press.
- Barry, P.H., 2006. The reliability of relative anion-cation permeabilities deduced from reversal (dilution) potential measurements in ion channel studies. *Cell Biochemistry and Biophysics* 46, 143–154.
- Barry, P.H., Lewis, T.M., Moorhouse, A.J., 2013. An optimised 3 M KCl salt-bridge technique used to measure and validate theoretical liquid junction potential values in patch-clamping and electrophysiology. *European Biophysics Journal* 42, 631–646.
- Barry, P.H., Lynch, J.W., 1991. Liquid junction potentials and small cell effects in patch-clamp analysis. *Journal of Membrane Biology* 121, 101–117.
- Bertrand, D., Galzi, J.-L., Devillers-Thiery, A., Bertrand, S., Changeux, J.-P., 1993. Mutations at two distinct sites within the channel domain M2 alter calcium permeability of neuronal $\alpha 7$ nicotinic receptor. *Proceedings of the National Academy of Sciences of the United States of America* 90, 6971–6975.
- Bezanilla, F., 2000. The voltage sensor in voltage-dependent ion channels. *Physiological Reviews* 80, 555–592.
- Boron, W.F., Boulpaep, E.L., 2009. *Medical Physiology: A Cellular and Molecular Approach*, second ed. Philadelphia, PA: Saunders Elsevier.
- Byrne, J.H., 2015. Changes in the spatial distribution of charge. *Neuroscience* (Chapter 3) Online. Available at: <http://neuroscience.uth.tmc.edu/s1/chapter03.html> (accessed 10.02.15).
- Carland, J.E., Cooper, M.A., Sugiharto, S., *et al.*, 2009. Characterization of the effects of charged residues in the intracellular loop on ion permeation in $\alpha 1$ glycine receptor channels. *Journal of Biological Chemistry* 284, 2023–2030.
- Coster, H.G.L., 2003. The physics of cell membranes. *Journal of Biological Physics* 29, 363–399.
- Coster, H.G.L., 2009. Discovery of "punch-through" or membrane electrical breakdown and electroporation. *European Biophysics Journal* 39, 185–189.
- Cymes, G.D., Grosman, C., 2012. Tunable pKa values and the basis of opposite charge selectivities in nicotinic-type receptors. *Nature* 474, 526–530.
- Doi, A., Mizuno, M., Kataguchi, T., *et al.*, 2007. Slow oscillation of membrane currents mediated by glutamatergic inputs of rat somatosensory cortical neurons: *In vivo* patch-clamp analysis. *European Journal of Neuroscience* 26, 2565–2575.
- Doyle, D.A., Morais Cabral, J., Pfuetzner, R.A., *et al.*, 1998. The structure of the potassium channel: Molecular basis of K^+ conduction and selectivity. *Science* 280, 69–77.
- Farrant, M., Kaila, K., 2007. The cellular, molecular and ionic basis of GABA(A) receptor signalling. *Progress in Brain Research* 160, 59–87.
- Franciolini, F., Nonner, W., 1987. Anion and cation permeability of a chloride channel in rat hippocampal neurons. *Journal of General Physiology* 90, 453–478.
- Gentet, L.J., Avermann, M., Matyas, F., Staiger, J.F., Petersen, C.C., 2010. Membrane potential dynamics of GABAergic neurons in the barrel cortex of behaving mice. *Neuron* 65, 422–435.
- Henneman, E., Somjen, G., Carpenter, D.O., 1965. Functional significance of cell size in spinal motoneurons. *Journal of Neurophysiology* 28, 560–580.
- Hille, B., 1992. *Ion Channels of Excitable Membranes*, second ed. Sunderland, MA: Sinauer.
- Hodgkin, A.L., Horowitz, P., 1959. The influence of potassium and chloride ions on the membrane potential of single muscle fibres. *Journal of Physiology* 148, 127–160.

- Hodgkin, A.L., Huxley, A.F., 1939. Action potentials recorded from inside a nerve fibre. *Nature* 144, 710–711.
- Hodgkin, A.L., Huxley, A.F., 1945. Resting and action potentials in single nerve fibres. *Journal of Physiology* 104, 176–195.
- Hodgkin, A.L., Katz, B., 1949. The effect of sodium ions on the electrical activity of giant axon of the squid. *Journal of Physiology* 108, 37–77.
- Horn, R., Marty, A., 1988. Muscarinic activation of ionic currents measured by a new whole-cell recording method. *Journal of General Physiology* 92, 145–159.
- Ishibashi, H., Moorhouse, A.J., Nabekura, J., 2012. Perforated patch clamp technique. In: Okada, Y. (Ed.), *Patch Clamp Techniques: From Beginning to Advanced Protocols*. Springer Protocols Handbooks Series. Berlin Heidelberg: Springer-Verlag, pp. 71–83. (Chapter 4).
- Kaplan, J.H., 2002. Biochemistry of Na, K-ATPase. *Annual Review of Biochemistry* 71, 511–535.
- Keramidas, A., Moorhouse, A.J., Pierce, K.D., Schofield, P.R., Barry, P.H., 2002. Cation-selective mutations in the M2 domain of the inhibitory glycine receptor channel reveal determinants of ion-charge selectivity. *Journal of General Physiology* 119, 393–410.
- Keramidas, A., Moorhouse, A.J., Schofield, P.R., Barry, P.H., 2004. Ligand-gated ion channels: Mechanisms underlying ion selectivity. *Progress in Biophysics and Molecular Biology* 86, 161–204.
- Krishnan, A.V., Kiernan, M.C., 2007. Uremic neuropathy: Clinical features and new pathophysiological insights. *Muscle Nerve* 35, 273–290.
- Liu, Y.S., Somporpisut, P., Perozo, E., 2001. Structure of the KcsA channel intracellular gate in the open state. *Nature Structural & Molecular Biology* 8, 883–887.
- Margrie, T.W., Brecht, M., Sakmann, B., 2002. In vivo, low-resistance, whole-cell recordings from neurons in the anaesthetized and awake mammalian brain. *Pflügers Archives* 444, 491–498.
- McCormick, D.A., 2008. Membrane potential and action potential. In: Squire, L.A. *et al.* (Eds.), *Fundamental Neuroscience*, third ed. San Diego, CA: Elsevier Press (Chapter 6).
- Moorhouse, A.J., Keramidas, A., Zaykin, A., Schofield, P.R., Barry, P.H., 2002. Single channel analysis of conductance and rectification in cation-selective, mutant glycine receptor channels. *Journal of General Physiology* 119, 411–425.
- Moorhouse, A.J., Lewis, T.M., Barry, P.H., in press. Analysing ion permeation in channels, pumps and transporters using patch clamp recording. In: Clark, R.J., Khalid, M.A.A. (Eds.), *Pumps, Channels and Transporters: Methods of Functional Analysis*. NJ: John Wiley & Sons.
- Mullins, L.J., Noda, K., 1963. The influence of sodium-free solutions on the membrane potential of frog muscle fibres. *Journal of General Physiology* 47, 117–132.
- Poulet, J.F., Petersen, C.C., 2008. Internal brain state regulates membrane potential synchrony in barrel cortex of behaving mice. *Nature* 454, 881–885.
- Reimann, F., Ashcroft, F.M., 1999. Inwardly rectifying potassium channels. *Current Opinion in Cell Biology* 11, 503–508.
- Sine, S.M., Engel, A.G., 2006. Recent advances in Cys-loop receptor structure and function. *Nature* 440, 448–455.
- Somjen, G.G., 2002. Ion regulation in the brain: Implications for pathophysiology. *Neuroscientist* 8, 254–267.
- Spruston, N., Jaffe, D.B., Williams, S.H., Johnston, D., 1993. Voltage- and space-clamp errors associated with the measurement of electrotonically remote synaptic events. *Journal of Neurophysiology* 70, 781–802.
- Spruston, N., Johnston, D., 1992. Perforated patch-clamp analysis of the passive membrane properties of three classes of hippocampal neurons. *Journal of Neurophysiology* 67, 508–529.
- Staley, K.J., Otis, T.S., Mody, I., 1992. Membrane properties of dentate granule cells: Comparison of sharp microelectrode and whole-cell recordings. *Journal of Neurophysiology* 67, 1346–1358.
- Sugiharto, S., Carland, J.E., Lewis, T.M., Moorhouse, A.J., Barry, P.H., 2010. External divalent cations increase anion–cation permeability ratio in glycine receptor channels. *Pflügers Archiv: European Journal of Physiology* 460, 131–152.
- Sugiharto, S., Lewis, T.M., Moorhouse, A.J., Schofield, P.R., Barry, P.H., 2008. Anion–cation permeability correlates with hydrated counterion size in glycine receptor channels. *Biophysical Journal* 95, 4698–4715.
- Thomas, R.C., 1972. Intracellular sodium activity and the sodium pump in snail neurones. *Journal of Physiology* 220, 55–71.
- Thurbon, D., Lüscher, H.-R., Hofstetter, T., Redman, S.J., 1998. Passive electrical properties of ventral horn neurons in rat spinal cord slices. *Journal of Neurophysiology* 79, 2485–2502.
- Tyzio, R., Ivanov, A., Bernard, C., *et al.*, 2003. Membrane potential of CA3 hippocampal pyramidal cells during postnatal development. *Journal of Neurophysiology* 90, 2964–2972.
- Tyzio, R., Minlebaev, M., Rheims, S., *et al.*, 2008. Postnatal changes in somatic gamma-aminobutyric acid signalling in the rat hippocampus. *European Journal of Neuroscience* 27, 2515–2528.

Relevant Websites

www.nobel.org
 Nobel Organizations.
[http://onlinelibrary.wiley.com/journal/10.1111/\(ISSN\)1469-7793/homepage/celebrating_the_work_of_alan_hodgkin_and_andrew_huxley.htm](http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1469-7793/homepage/celebrating_the_work_of_alan_hodgkin_and_andrew_huxley.htm)
 John Wiley & Sons, Inc.
<http://www.youtube.com/watch?v=k48jXzFGMc8>

The Outer Mitochondrial Membrane, a Smooth ‘Coat’ with Many Holes and Many Roles: Preparation, Protein Components, Interactions with Other Membranes, Involvement in Health, Disease, and as a Drug Target

PL Pedersen, Johns Hopkins University School of Medicine, Baltimore, MD, USA
YH Ko, University of Maryland BioPark, Baltimore, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

In 1948, Lehninger and Kennedy showed mitochondria to be the intracellular site of ATP (adenosine triphosphate) synthesis by oxidative phosphorylation (Lehninger and Kennedy, 1948), and in 1953 they were visualized by electron microscopy by Palade (1953). These observations followed from the earlier work in the 1930s of Sir Hans Krebs that animal cells have two important metabolic cycles, i.e., the tricarboxylic acid cycle (TCA cycle), reviewed in Krebs and Johnson (1937), and the urea cycle, reviewed in (Krebs, 1973). Later, the TCA cycle and part of the urea cycle were localized to the mitochondria. Subsequent work revealed that mitochondria have their own DNA (Nass and Nass, 1963), RNA (Wang *et al.*, 2010), and the capacity to synthesize some of their own proteins (Pel and Grivell, 1994). Finally, studies using electron microscopy and other technologies revealed that mitochondria, in terms of numbers, are exceptionally high in nucleated cells. Thus, heart cells may contain about 5000 mitochondria per cell, muscle cells about 2500, and liver cells, 1000–2000. Red blood cells (erythrocytes) that have no nucleus also have no mitochondria. Another view is that a cell has a single mitochondrial reticulum.

By 1967, the use of electron microscopy had broadened, making it possible for several different laboratories to visualize mitochondria clearly and in some detail. Significantly, such studies with either liver mitochondria (Hackenbrock, 1968) or cancer (hepatoma) mitochondria (Pedersen *et al.*, 1970) confirmed that these ATP-producing organelles consist of four major compartments illustrated in Figure 1 and listed as follows:

1. An outer membrane, the subject of this brief article.
2. An inner membrane that consists of invaginations named ‘cristae.’
3. A matrix, a soluble compartment enclosed by the inner membrane.
4. An intermembrane space, the soluble compartment between the inner and outer membranes.

This new information posed many questions, the most obvious being ‘In which of the above four compartments (or major compartments) do known mitochondrial processes take place?’

Separation and Isolation of the Four Mitochondrial Compartments

To address the above question Carl Schnaitman working in the laboratory of Dr. Jack Greenawalt at Johns Hopkins University in Baltimore in the late 1960s took on the challenging project of ‘nailing down’ the submitochondrial locations of known mitochondrial processes and the enzymes/proteins involved.

Several other laboratories at other institutions had also begun similar studies, all with different approaches.

Schnaitman's approach included refrigerated centrifuges with the capacity to sediment cellular homogenates at a relatively low speed, others that would sediment them at a very high speed, and still others that operated somewhere in between the two. To enter the laboratory in which Schnaitman worked in the middle of the night was like entering an aircraft company testing plane engines with the exception that rather than the noise coming from motors inducing propellers to spin, it came from centrifuge motors inducing rotors to spin. Schnaitman also used electron microscopy to visualize the mitochondria before and after treatment with various detergents, mainly digitonin and lubrol WX. In collaboration with V.G. Erwin, he deduced that the enzyme referred to as monoamine oxidase would likely be a good marker for the outer mitochondrial membrane (Schnaitman *et al.*, 1967) and, via titration of a known amount of mitochondria, he found a concentration of the detergent digitonin that ‘teased off’ both (1) the outer membrane and (2) the intermembrane space. The remaining parts of the mitochondria, i.e., the inner membrane and matrix, were left intact together and named ‘mitoplast’ for one and ‘mitoplasts’ for more than one. By treating the mitoplasts with the nonionic detergent Lubrol WX, he was able to separate them into (3) the inner membrane fraction and (4) the soluble ‘matrix’ fraction. Thus, in a single experiment, with appropriate centrifugation steps to separate 1 from 2 and 3 from 4, he could isolate the four components of the mitochondria (Schnaitman and Greenawalt, 1968). Schnaitman also learned how to assay the activity of almost every enzyme known at that time to be associated with the mitochondria. These included enzymes associated with the TCA cycle, urea cycle, and the oxidative phosphorylation system involved in ATP synthesis.

Results: Protein/Enzyme Assignments of the Four Mitochondrial Compartments with an Emphasis on the Outer Mitochondrial Membrane

Subsequent studies by Schnaitman (Schnaitman *et al.*, 1967; Schnaitman and Greenawalt, 1968), in collaboration with the laboratory of Pedersen (Schnaitman and Pedersen, 1968; Pedersen and Schnaitman, 1969; Chan *et al.*, 1970), and work conducted independently in Stockholm (Sottocasa *et al.*, 1967) eventually led to a general consensus that:

1. Enzymes/proteins involved in the ‘electron transport chain’ are located in or associated with the mitochondrial inner membrane.

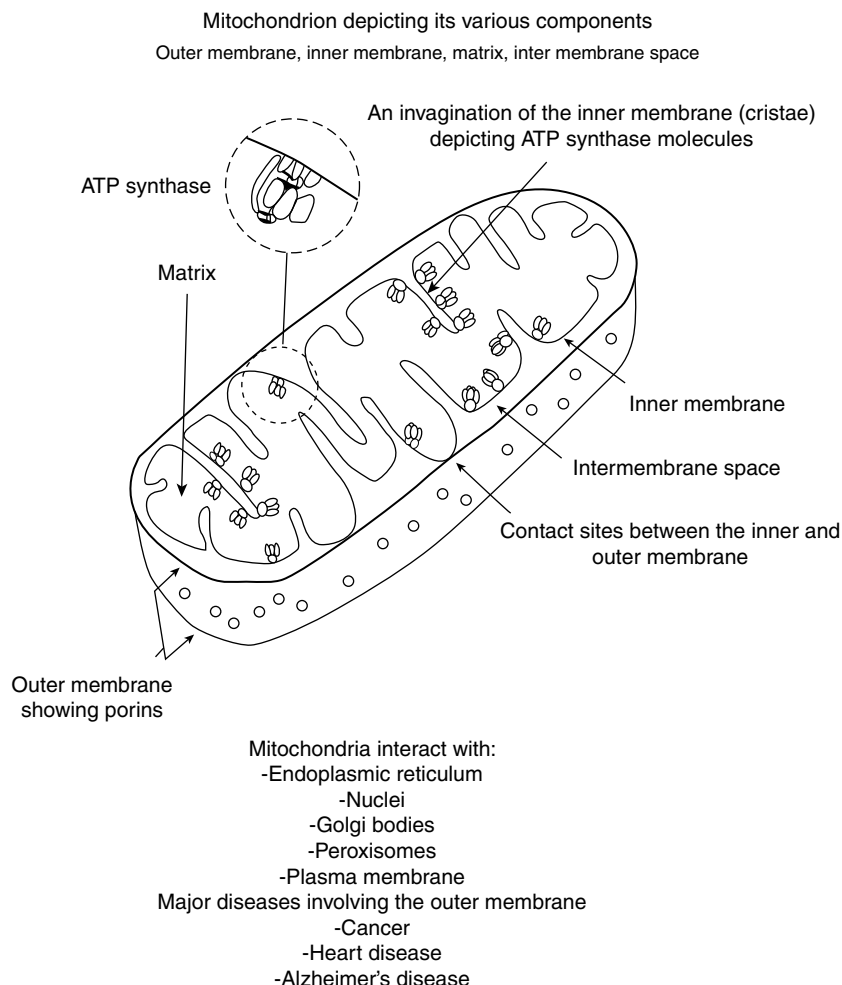


Figure 1 Schematic diagram of the bottom half of a mitochondrion that allows its various components to be visualized (artwork by Albert Ko, Gerstell Academy, Finksburg, MD, USA).

2. The enzyme complex that makes ATP (ATP synthase) and catalyzes an ATP/ADP exchange reaction inhibited by the ATP synthesis inhibitor 'oligomycin' is associated also with the mitochondrial inner membrane.
3. The key transport systems (adenine nucleotide carrier and phosphate carrier) that transport, respectively, ADP and phosphate into the mitochondria to provide substrates for the catalytic headpiece (F_1) of ATP synthase are localized within the inner membrane. (Once made, the adenine nucleotide carrier also transports the newly synthesized ATP out of the mitochondrial matrix in exchange for another ADP molecule entering the inner compartment.)
4. Most enzymes of the TCA cycle are located within the mitochondrial matrix, but transporters that may move some of the intermediates in or out of the mitochondrial matrix are localized in the inner mitochondrial membrane.
5. Two enzymes involved in the urea cycle are located in the mitochondrial matrix, and three are located in the cytosol.

Based on all the above, early work (late 1960s early 1970s) was targeted much more toward what was happening in the

mitochondrial inner compartments (inner membrane and/or matrix) and less attention was focused on the outer compartments (outer membrane and intermembrane space). However, work conducted in many different laboratories ultimately succeeded in elucidating the protein/enzyme composition of both membranes. Many of these findings and their significance (or possible significance) are summarized in [Table 1](#) as they specifically relate to the outer mitochondrial membrane.

'Moonlighting' Roles of the Mitochondrial Outer Membrane with Its Intracellular Neighbors

An area of mitochondrial research that has not received the attention that has been given to biochemical processes relates to what might be regarded as the 'moonlighting' roles of this organelle, i.e., its interactions with its intracellular neighbors. Such interactions are expected to, and in fact must, involve the outer mitochondrial membrane because it is readily accessible.

Table 1 Enzymes localized in the outer mitochondrial membrane

Enzyme(s)/protein(s)	Tissue/cell type(s)	Significance/possible significance
Acetyl CoA synthase	Liver	Activates breakdown of long-chain fatty acids (Norum <i>et al.</i> , 1966)
Arginosuccinate synthetase	Liver	Converts citrulline into arginosuccinate in the urea cycle (Demarquoy <i>et al.</i> , 1994)
Carnitine palmitoyl transferase	Heart, muscle	Carnitine is an essential cofactor in the beta oxidation of long-chain fatty acids that must be transported into mitochondria (Murthy and Pandey, 1987)
Choline dehydrogenase	Several	Role in betaine aldehyde formation. Also, a role in mitophagy, process by which cells eliminate mitochondria (Park <i>et al.</i> , 2014)
Creatine kinase	Heart, muscle	Bound to VDAC (Bridiczka <i>et al.</i> , 1994); helps provide a temporal and spatial energy buffer to maintain cellular homeostasis (Schlattner <i>et al.</i> , 2001)
Cytochrome <i>b5</i> -like hemoprotein	Rat liver	Cyt <i>b5</i> -like hemoprotein of the outer mitochondrial membrane (Ito, 1980); believed to be involved in NADH-semidehydroascorbic acid reductase activity (Ito <i>et al.</i> , 1981)
Deubiquitinating (Ub-16)	Yeast	May help regulate the turnover or activity of proteins at the outer membrane (Kinner and Kolling, 2003)
Deubiquitinating (USP-30)	Hela and Cos 7 cells	Maintenance of mitochondrial morphology (Nakamura and Hirose, 2008)
DMT1 (divalent metal transporter 1)	Most	Helps mitochondria acquire iron (Wolff <i>et al.</i> , 2014)
Glycolytic (All)	<i>Arabidopsis</i>	Allows pyruvate to be provided directly to mitochondria where it is used as a respiratory substrate (Giege <i>et al.</i> , 2003)
Hexokinase 2 (Hk-2)	Heart	Required for optimal cardiac function. Reduction in Hk-2 levels results in decreased cardiac function and altered remodeling after ischemia/reperfusion injury (Wu <i>et al.</i> , 2012) Also, Hk-2 promotes glycolysis and the production of ATP when oxygen is limiting. Finally, Hk-2 bound to VDAC may prevent cell death as in cancer (Gottlob <i>et al.</i> , 2001; Pastorino <i>et al.</i> , 2002) (see below)
Hexokinase 2 (Hk-2)	Adipocytes	Hk-2 is the main hexokinase in fat pads (Katzen and Schimke, 1965); thus it is likely involved in initiating the conversion of glucose to fat in adipose tissue (Flatt, 1970)
Hexokinase 2 (Hk-2)	Most cancers	Hk-2 bound to mitochondrial VDAC in cancers (Nakashima <i>et al.</i> , 1986) plays a pivotal role (Bustamante and Pedersen, 1977) in the 'Warburg effect,' i.e., high rates of conversion of glucose to lactic acid even in the presence of oxygen. Hk-2 provides glucose-6-phosphate for biosynthesis, and immortalizes cancer cells (Gottlob <i>et al.</i> , 2001; Pastorino <i>et al.</i> , 2002). Also, Hk-2 bound to VDAC is the enzymatic basis of PET imaging used worldwide to detect and monitor cancers
Kynurenine hydroxylase	Widespread	Involved in degradation pathway for tryptophan (Guillemin <i>et al.</i> , 2007). Associated with disorders of the brain (Shilov and Bezrukov, 2013)
Mitochondrial Ubiquitin ligase (MITOL)	Widespread	Involved in mitochondrial dynamics Eliminates misfolded proteins that localize to mitochondria, for example, superoxide dismutase (mSOD1) (Yonshiro, 2006). Mitol also regulates ER contacts via mitofusin 2 (Sugiura <i>et al.</i> , 2013)
Monoamine oxidase	Widespread	Maintains neuron firing rates throughout the body by metabolizing in the liver bioactive amines absorbed into the blood from food (Richardson, 1993)
NADH cyt <i>b5</i> reductase	Rat liver	Not clear but similar or identical with the microsomal NADH-cyt <i>b5</i> reductase (Sottocasa <i>et al.</i> , 1967)
OM 14	Yeast	OM 14 is a receptor for NAC, ribosome-associated chaperone nascent-chain associated complex (Lesnik <i>et al.</i> , 2014)
NADH cyt <i>c</i> reductase TOM	Rat liver Most cells	May participate in NADH semi-hydroascorbic acid reductase activity (Ito <i>et al.</i> , 1981) Translocase of the outer membrane that moves proteins into the intermembrane space (Neupert and Herrmann, 2007). There are a number of different TOMs, for example, TOM 5, TOM 6, TOM 7, TOM 20, TOM 22, TOM 40, and TOM 70. The TOM proteins work with TIM proteins (translocases of the inner membrane), for example, TIM 17, TIM 18, TIM 22, TIM 23, TIM 44, TIM 50, TIM and TIM 54, to move proteins into the mitochondrial matrix
VDAC	Most cells	A voltage-dependent anion channel in the outer mitochondrial membrane that is implicated as a governor of mitochondrial function (Colombini, 2004); its roles in health and disease are numerous. VDAC is essential for mitochondrial function in almost all cell types in humans and animals. However, when cells become cancerous in animals and humans VDAC, perhaps wanting to keep its job, participates in their immortalization by hijacking Hk2 (Pastorino <i>et al.</i> , 2002)

Mitochondrial Endoplasmic Reticulum Interactions

Several studies have shown that parts of the endoplasmic reticulum (ER) are intimately associated with the mitochondrial

outer membrane with one of the earliest reports occurring in Copeland and Dalton (1959). Shore and Tata (1977) reported rapidly sedimenting ER in association with mitochondria. In a follow-up study, Pickett *et al.* (1980), via sedimentation and

microscopic analysis, provided evidence for rough ER mitochondrial complexes. Filippin *et al.* (2003) concluded that stable interactions do occur between mitochondria and the ER and that this allows for a rapid accumulation of calcium in a subpopulation of mitochondria. Malhotra and Kaufman (2011) commented on what they referred to as 'ER stress,' its functional link to mitochondria and role in cell survival and death. Finally, Flis and Daum (2013) in a detailed report emphasize the importance of lipid transport between the ER and mitochondria that would necessitate involvement of the outer mitochondrial membrane.

Mitochondrial–Nuclear Interactions

Although a limited amount of work has been conducted on mitochondrial–nuclear interactions this is likely to prove to be quite significant as more work is completed. Five studies are briefly summarized here. First, in a report published, Ornstein (1956) described 'a remarkable relationship which occurs during a limited phase of this early development between the nuclear processes and the mitochondria of the oocyte.' The author goes on to state, "Here we see that the mitochondria are closely applied to the denser material which seems to be continuous with the nuclear membrane processes-as if we have caught the nucleus and the mitochondria in the process of cooperatively producing the denser masses". Mota (1963) described mitochondria that he had observed in the plant *Chlorophytum capense* (L.) Kuntz. He stated, "Electron microscope studies of tips of aerial roots of *Chlorophytum capense* (L.) Kuntz showed the formation in the nucleus of a large number of invaginations in which mitochondria accumulate. These mitochondria are seen in many cases attached to the nuclear membrane, which appears destroyed at the point of attachment, as if the mitochondria were exchanging some substances with the nucleus or were being absorbed by the nucleus. This phenomenon can conceivably be a mechanism to transfer ATP and possibly other substances from the mitochondria to the nucleus. Intimate contact between mitochondria and the nuclear membrane may be only a 'touch and go' and only involve a small portion of the mitochondrial surface". (This would of course be the outer membrane.)

Jensen *et al.* (1976) described myocardial cells derived from a 60-year-old patient with rheumatic heart disease undergoing cardiopulmonary bypass surgery. Surprisingly, such cells were found to contain mitochondria with well-preserved membranes within their nuclei. For such mitochondria to have entered the nuclei, the mitochondrial outer membrane would have had to interact first with the nuclear membrane. Two other later studies provided further evidence for mitochondrial–nuclear interactions and/or communication. Dzeja *et al.* (2002) monitored the import of histone H1, a constitutive nuclear protein into the nucleus and found mitochondria to cluster around the nucleus reducing the distance for energy transfer and in so doing undergoing interactions of their outer membrane with the nuclear membrane. One year later Prachar (2003), in briefly summarizing the study, states, "Mitochondria in close proximity to the nuclear envelop can be found virtually in all metabolically active cells. We used transmission electron microscopy to demonstrate this entity in

different leukemia cells of human origin (patient's blood) and in a mouse leukemia cell line. At high resolution, not only close proximity but even fusion of mitochondrial and nuclear membranes can be seen. Based on available data about mRNA transport through the nuclear pore complex and observed contacts of mitochondria with nuclei, we hypothesize that such contacts can provide a gateway for energy delivery to power mRNA export from the nucleus to the cytoplasm."

Mitochondrial Interactions with the Golgi, Peroxisomes, and Plasma Membrane

From the observations noted, it is clear that mitochondria have developed quite a social network within cells. This network is considered to be highly dependent on the outer mitochondrial membrane as it is the only compartment of a mitochondrion that is in a position to enter into direct contact with intracellular neighbors. In addition to having been reported to interact with the ER and nucleus of cells, mitochondria have been reported to interact also with the Golgi, peroxisomes, and the plasma/cell membrane. Dolman *et al.* (2005) reported stable Golgi–mitochondrial complexes and the formation of Golgi Ca^{2+} gradients in pancreatic acinar cells. They state, "The Golgi is ideally placed to be preferentially supplied by ATP from adjacent mitochondria." As it relates to interaction of mitochondria with peroxisomes, both Neuspiel *et al.* (2008) and Borgue and Demarquoy (2012) have reported such interactions, the latter article related to a possible role in fatty acid metabolism, and referred to as the 'mitochondria–peroxisome connection,' by Smallridge (2008). Finally, and likely equally or more significant is the finding of mitochondrial plasma membrane contacts (Klecker *et al.*, 2013). The work involving yeast characterized mitochondrial–plasma membrane interactions identifying Num1 (nuclear migration one) as a protein involved in the process. Num 1 is a 313 kDa protein that binds to the plasma membrane by a plekstrin homology (PH) domain and assembles into cortical patches. Num 1 is sometimes referred to as a cell cortex protein, where the cell cortex is defined as a specialized layer of cytoplasm on the inner face of the plasma membrane that functions as a mechanical support for it.

Involvement of the Outer Mitochondrial Membrane in Major Diseases and Potential Drug Target

Cancer

Cancer, according to the World Health Organizations Fact Sheet N 297 updated in 2014 (Stewart and Wild, 2015), figures among the leading causes of morbidity and mortality worldwide with approximately 14 million new cases and 8.2 million cancer-related deaths in 2012. Moreover, the number of new cases is expected to rise about 70% over the next 2 decades. It is known that the outer mitochondrial membrane is one of cancers' best friends. This was first recognized by Bustamante and Pedersen (1977). Specifically, they demonstrated that a form of hexokinase bound to a protein in mitochondria, later shown to be the outer membrane VDAC protein (Nakashima *et al.*, 1986), is intimately involved in the 'Warburg effect,' i.e., high glycolysis even in the presence of

oxygen. Significantly, this produces large amounts of the glucose-6-phosphate utilized in biosynthetic reactions essential for the rapid growth and division of cancer cells. However, as shown later by others, Hk-2 bound to VDAC is involved also in the immortalization of cancer cells (Gottlob *et al.*, 2001; Pastorino *et al.*, 2002) and a detection system (positron emission tomography, PET) was developed a number of years ago based on hexokinase (Som *et al.*, 1980) and is now used worldwide to detect cancers and monitor their treatment.

In more recent research, it was shown that one of the targets of the rapidly effective anticancer agent 3-bromopyruvate (3BP) is outer mitochondrial membrane bound Hk-2 (Ko *et al.*, 2001). In studies with 19 animals bearing advanced cancers carried out by Ko *et al.* (2004), 3BP was found to eradicate them all, i.e., 100%. Moreover, the cancers never returned throughout the lifetime of the animals. 3BP, when properly formulated, was found to be applicable also to human cancer. Thus, the life of one teenage boy suffering from liver cancer was significantly extended (Ko *et al.*, 2012).

In summary, the outer mitochondrial membrane is intimately involved in the rapid growth and immortality of cancer via bound Hk-2, and this enzyme has provided via PET imaging the basis for cancer detection and monitoring as well as a target for the potent anticancer agent 3BP.

Heart Disease

According to a report by Murphy *et al.* (2013), about 600 000 people die of heart disease in the United States every year. It accounts for one in four deaths. As for cancer cells, heart cells (cardiomyocytes) have Hk-2 bound to their outer mitochondrial membrane, and at this location it is believed to be essential not only for (1) every beat, but also for (2) the lifetime of the heart and, therefore, (3) human/animal lifetime. As Smeele *et al.* (2011) emphasized from results of their studies, "Disruption of Hexokinase II (Hk-2) mitochondrial binding blocks ischemic preconditioning and causes rapid cardiac necrosis." They go on to state that "The present data demonstrate that HK II (Hk-2) detachment from mitochondria results in acute mitochondrial depolarization, which is likely the mechanism of cardiac dysfunction." Also, Wu *et al.* (2012) reported that a reduction in Hk-2 levels results in decreased cardiac function and altered remodeling after ischemia/reperfusion injury. Finally, Anflous *et al.* (2007) have shown recently that VDAC1 serves as a mitochondrial binding site for hexokinase in oxidative muscles. The authors refer to "Hk-2 detachment from mitochondria," which most likely means detachment from its binding to VDAC1 in the outer mitochondrial membrane.

Alzheimer's Disease (AD)

First, in the United States, it was reported (Alzheimer's Association, 2014) that:

1. An estimated 5.2 million Americans of all ages have AD (year 2014).
2. One in nine aged 65 and older (11%) has AD.
3. About 1/3 of people 85 and older (32%) have AD.

4. Of those with AD, the vast majority (82%) are aged 75 or older.

In a second report that appeared also this past year (Prince *et al.*, 2014), worldwide:

1. Dementia, including AD, remains one of the biggest global health challenges facing our generation.
2. The number of people living with dementia worldwide today is estimated at 44 million, set to double by 2030 and more than triple by 2050.
3. The global cost was estimated in 2010 to be US\$604 billion. Therefore, by the year 2050 the total cost will exceed US\$1812 billion.

So what is the relationship of the mitochondrial outer membrane to AD and dementia? It has been known for some time that a component of the brains of deceased Alzheimer's patients contain plaques referred to as 'amyloid plaques' that are comprised of peptides 36–43 amino acids in length. Some of these are imported into mitochondria via the outer membrane TOM import machinery and localize to mitochondrial cristae (Hansson Petersen *et al.*, 2008). In particular, a channel forming subunit of TOM referred to as TOMM40 that is essential for protein transport into mitochondria has received recent attention. In humans certain alleles of the gene encoding TOMM40 have been statistically associated with an increased risk of late on-set AD. Thus, in a genome-wide association study on 381 participants in the ADNI (Alzheimer's Disease Neuroimaging Initiative) study, TOMM40 risk alleles were approximately twice as frequent in AD subjects as controls (Potkin *et al.*, 2009).

Concluding Remarks

The mitochondrial outer membrane plays highly important roles in both health and disease. It is not simply a protective cover for the mitochondrial inner membrane plus matrix compartments where the TCA cycle and oxidative phosphorylation take place, but rather is very active, in fact much more active than most investigators realize. As described, the outer membrane on a day-to-day basis interacts with the ER, the nucleus, Golgi, peroxisomes, and plasma membranes. Interestingly, when cells become diseased, the mitochondria, perhaps to survive, appear to take on properties to facilitate the disease. It seems that whether we are healthy or contract a disease, a better understanding of the role of mitochondria in these processes will be important for many aspects of human and animal health.

Acknowledgments

PLP was supported for many years by the National Cancer Institute, for which he is grateful.

See also: Molecular Principles, Components, Technology, and Concepts: Membranes: Composition, Physical Properties, and Curvature

References

- Alzheimer's Association, 2014. Alzheimer's disease facts and figures. *Alzheimer's & Dementia* 10 (2), 1–80.
- Anflous, K., Cai, Z.J., Craigen, W.J., 2007. VDAC1 serves as a mitochondrial binding site for hexokinase in oxidative muscles. *Biochimica et Biophysica Acta* 1767, 136–142.
- Borgue, F.L., Demarquoy, J., 2012. Interaction between peroxisomes and mitochondria in fatty acid metabolism. *Open Journal of Molecular and Integrative Physiology* 2, 27–33.
- Bridiczka, D., Kalds, P., Wallimann, T., 1994. In vitro complex formation between the octamer of mitochondrial creatine kinase and porin. *Journal of Biological Chemistry* 269, 27640–27644.
- Bustamante, E., Pedersen, P.L., 1977. High aerobic glycolysis of rat hepatoma cells in culture: Role of mitochondrial hexokinase. *Proceedings of the National Academy of Sciences of the United States of America* 74, 3735–3738.
- Chan, T.L., Greenawalt, J.W., Pedersen, P.L., 1970. Biochemical and ultrastructural properties of a mitochondrial inner membrane fraction deficient in outer membrane and matrix activities. *Journal of Cell Biology* 45, 291–305.
- Colombini, M., 2004. VDAC: The channel at the interface between mitochondria and the cytosol. *Molecular and Cellular Biochemistry* 256–257, 107–115.
- Copeland, D.E., Dalton, A.J., 1959. An association between mitochondria and the endoplasmic reticulum in cells of the pseudobranch gland of a teleost. *Journal of Biophysical and Biochemical Cytology* 5, 393–396.
- Demarquoy, J., Fairand, A., Gautier, C., Vaillant, R., 1994. Demonstration of argininosuccinate synthetase activity associated with mitochondrial membrane: Characterization and hormonal regulation. *Molecular and Cellular Biochemistry* 136, 145–155.
- Dolman, N.J., Gerasimenko, J.V., Gerasimenko, O.V., *et al.*, 2005. Stable golgi-mitochondrial complexes and formation of golgi Ca^{2+} gradients in pancreatic acinar cells. *Journal of Biological Chemistry* 280, 15794–15799.
- Dzeja, P.P., Bortolon, R., Perez-Terzic, C., Holmuhamedov, E.L., Terzic, A., 2002. Energetic communication between mitochondria and nucleus directed by catalyzed phosphotransfer. *Proceedings of the National Academy of Sciences of the United States of America* 99, 10156–10161.
- Filippin, L., Magalhaes, P.J., DiBenedetto, G., Colella, M., Pozzan, T., 2003. Stable interaction between mitochondria and endoplasmic reticulum allow rapid accumulation of calcium in a subpopulation of mitochondria. *Journal of Biological Chemistry* 278, 39224–39234.
- Flatt, J.P., 1970. Conversion of carbohydrate to fat in adipose tissue: An energy-yielding and, therefore, self limiting process. *Journal of Lipid Research* 11, 131–143.
- Flis, V.V., Daum, G., 2013. Lipid transport between endoplasmic reticulum and mitochondria. *Cold Spring Harbor Perspectives in Biology* 5, 1–22.
- Giege, P., Heazlewood, J.L., Roessner-Tunali, U., *et al.*, 2003. Enzymes of glycolysis are functionally associated with the mitochondria in *Arabidopsis* cells. *Plant Cell* 15, 2140–2151.
- Gottlob, K., Majewski, N., Kennedy, S., *et al.*, 2001. Inhibition of early apoptotic events by AKT/PKB is dependent on the first committed step of glycolysis and mitochondrial hexokinase. *Genes & Development* 15, 1406–1418.
- Guillemin, G.J., Cullen, K.M., Lim, C.K., *et al.*, 2007. Characterization of the kynurenine pathway in human neurons. *Journal of Neuroscience* 27, 12884–12892.
- Hackenbrock, C.R., 1968. Chemical and physical fixation of isolated mitochondria in low-energy and high-energy states. *Proceedings of the National Academy of Sciences of the United States of America* 61, 598–605.
- Hansson Petersen, C.A., Behbahani, H., Wiehager, B., *et al.*, 2008. The amyloid beta-peptide is imported into mitochondria via the TOM import machinery and localized to mitochondrial cristae. *Proceedings of the National Academy of Sciences of the United States of America* 105, 13145–13150.
- Ito, A., 1980. Cytochrome b5-like hemoprotein of outer mitochondrial membrane; OM cytochrome b5. I. Purification of OM cytochrome b from rat liver mitochondria and comparison of its molecular properties with those of cytochrome b5. *Journal of Biochemistry* 87, 63–71.
- Ito, A., Hayashi, S., Yoshida, T., 1981. Participation of a cytochrome b5-like hemoprotein of the outer mitochondrial membrane (OM cytochrome b) in NADH-semidehydroascorbic acid reductase activity of rat liver. *Biochemical and Biophysical Research Communications* 101, 591–598.
- Jensen, H., Engedal, H., Sætersdal, T.S., 1976. Ultrastructure of mitochondria-containing nuclei in human myocardial cells. *Virchows Archiv – B: Cell Pathology* 21, 1–12.
- Katzen, H.M., Schimke, R.T., 1965. Multiple forms of hexokinase in the rat: Tissue distribution, age dependency, and properties. *Proceedings of the National Academy of Sciences of the United States of America* 54, 1218–1225.
- Kinner, A., Kolling, R., 2003. The yeast deubiquitinating enzyme Ubp16 is anchored to the outer mitochondrial membrane. *FEBS Letters* 549, 135–140.
- Klecker, T., Scholz, D., Fortsch, J., Westermann, B., 2013. The yeast cell cortical protein Num 1 integrates mitochondrial dynamics into cellular architecture. *Journal of Cell Science* 126, 2924–2930.
- Ko, Y.H., Pedersen, P.L., Geschwind, J.F., 2001. Glucose catabolism in the rabbit VX2 model for liver cancer: Characterization and targeting hexokinase. *Cancer Letters* 173, 83–91.
- Ko, Y.H., Smith, B.L., Wang, Y., *et al.*, 2004. Advanced cancers: eradication in all cases using 3-bromopyruvate therapy to deplete ATP. *Biochemical and Biophysical Research Communications* 324, 269–275.
- Ko, Y.H., Verhoven, H.A., Lee, M.J., *et al.*, 2012. A translational study “case report” on the small molecule “energy blocker” 3-bromopyruvate (3BP) as a potent anti-cancer agent. *Journal of Bioenergetics and Biomembranes* 44, 163–170.
- Krebs, H., 1973. The discovery of the ornithine cycle of urea synthesis. *Biochemical Education* 1, 19–23.
- Krebs, H.A., Johnson, W.A., 1937. Metabolism of ketonic acids in animal tissues. *Biochemical Journal* 31, 645–660.
- Lehninger, A.L., Kennedy, E.P., 1948. The requirements of the fatty acid oxidase complex of rat liver mitochondria. *Journal of Biological Chemistry* 173, 753–771.
- Lesnik, C., Cohen, Y., Atir-Lande, A., Schuldiner, M., Arava, Y., 2014. OM14 is a mitochondrial receptor for cytosolic ribosomes that supports co-translational import into mitochondria. *Nature Communications* 6711, 1–10.
- Malhotra, J.D., Kaufman, R.J., 2011. ER stress and its functional link to mitochondria: Role in cell survival and death. *Cold Spring Harbor Perspectives in Biology* 3, 1–13.
- Mota, M., 1963. Electron microscope study of the relationship between the nucleus and the mitochondria in *Chlorophytum capense* (L.) Kuntz. *Cytologia* 28, 409–416.
- Murphy, S.L., Xu, J.Q., Kochanek, K.D., 2013. Deaths: Final data for 2010. *National Vital Statistics Reports* 61, 1–117.
- Murthy, M.S.R., Pandey, S.V., 1987. Malonyl-CoA binding site and overt carnitine palmitoyltransferase activity reside on the opposite sides of the outer mitochondrial membrane. *Proceedings of the National Academy of Sciences of the United States of America* 84, 378–382.
- Nakamura, N., Hirose, S., 2008. Regulation of mitochondrial morphology by USP 30, a deubiquitinating enzyme present in the mitochondrial outer membrane. *Molecular Biology of the Cell* 19, 1903–1911.
- Nakashima, R.A., Mangan, P.S., Colombini, M., Pedersen, P.L., 1986. Hexokinase receptor complex in hepatoma mitochondria: evidence from N, N'-dicyclohexylcarbodiimide-labeling studies for the involvement of the pore-forming protein VDAC. *Biochemistry* 25, 1015–1021.
- Nass, S., Nass, M.M.K., 1963. Intramitochondrial fibers with DNA characteristics II. Enzymatic and other hydrolytic treatments. *Journal of Cell Biology* 19, 613–629.
- Neupert, W., Herrmann, J.M., 2007. Translocation of protein into mitochondria. *Annual Review of Biochemistry* 76, 723–749.
- Neuspiel, M., Schauss, A.C., Braschi, E., *et al.*, 2008. Cargo-selected from mitochondria to peroxisomes is mediated by vesicular carriers. *Current Biology* 18, 102–108.
- Norun, K.R., Farstad, M., Bremer, J., 1966. The submitochondrial distribution of acid:coA ligase (AMP) and palmitoyl-CoA: Carnitine palmitoyltransferase in rat liver mitochondria. *Biochemical and Biophysical Research Communications* 24, 797–804.
- Ornstein, L., 1956. Mitochondrial and nuclear interaction. *Journal of Biophysical and Biochemical Cytology* 2, 351–352.
- Palade, G.E., 1953. An electron microscope study of the mitochondrial structure. *Journal of Histochemistry and Cytochemistry* 1, 188–211.
- Park, S., Choi, S.-G., Yoo, S.-M., Son, J.H., Jung, Y.-K., 2014. Choline dehydrogenase interacts with SQSTM1 p62 to recruit LC3 and stimulate mitophagy. *Autophagy* 10, 1906–1920.
- Pastorino, J.G., Nataly, S., Hoek, J.B., 2002. Mitochondrial binding of hexokinase II inhibits bax-induced cytochrome c release and apoptosis. *Journal of Biological Chemistry* 277, 7610–7618.
- Pedersen, P.L., Greenawalt, J.W., Chan, T.L., 1970. A comparison of some ultrastructural and biochemical properties of mitochondria from Morris Hepatomas 9618A, 7800, and 3924A. *Cancer Research* 30, 2620–2626.
- Pedersen, P.L., Schnaitman, C.A., 1969. The oligomycin-sensitive adenosine diphosphate-adenosine triphosphate exchange in an inner membrane fraction of rat liver mitochondria. *Journal of Biological Chemistry* 244, 5065–5074.
- Pel, H.J., Grivell, L.A., 1994. Protein synthesis in mitochondria. *Molecular Biology Reports* 19, 183–194.

- Pickett, C.B., Montisano, D., Eisner, D., Cascarano, J., 1980. The physical association between rat liver mitochondria and rough endoplasmic reticulum I. Isolation electron microscopic examination and sedimentation equilibrium centrifugation analyses of rough endoplasmic reticulum-mitochondrial complexes. *Experimental Cell Research* 128, 343–352.
- Potkin, S.G., Guffanti, G., Lakatos, A., *et al.*, 2009. Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. *PLoS One* 4, e6501.
- Prachar, J., 2003. Intimate contacts of mitochondria with nuclear envelope as a potential energy gateway for nucleo-cytoplasmic mRNA transport. *General Physiology and Biophysics* 22, 525–534.
- Prince M., Albanese E., Guerchet M., Prina M., 2014. Dementia and risk reduction: An analysis of protective and modifiable factors. *Alzheimer's Disease International*, London, pp. 1–104.
- Richardson, J.S., 1993. On the functions of monoamine oxidase, the emotions and adaptations to stress. *International Journal of Neuroscience* 70, 75–84.
- Schlattner, U., Dolder, M., Wallimann, T., Malgorzata, T.-S., 2001. Mitochondrial creatine kinase and mitochondrial outer membrane porin show a direct interaction that is modulated by calcium. *Journal of Biological Chemistry* 276, 48027–48030.
- Schnaitman, C., Erwin, V.G., Greenawalt, J.W., 1967. The submitochondrial localization of monoamine oxidase, an enzymatic marker for the outer mitochondrial membrane of rat liver mitochondria. *Journal of Cell Biology* 32, 719–735.
- Schnaitman, C., Greenawalt, J.W., 1968. Enzymatic properties of the inner and outer membranes of rat liver mitochondria. *Journal of Cell Biology* 38, 158–175.
- Schnaitman, C.A., Pedersen, P.L., 1968. Localization of oligomycin-sensitive ADP-ATP exchange activity in rat liver mitochondria. *Biochemical and Biophysical Research Communications* 30, 428–433.
- Shilov, I.E., Bezrukov, M.V., 2013. Kynurenines in pathogenesis of endogenous psychiatric disorders. *Vestnik Ross Akad. Med. Nauk.* 2, 35–41.
- Shore, G.C., Tata, J.R., 1977. Two fractions of endoplasmic reticulum from rat liver I. Recovery of rapidly sedimenting endoplasmic reticulum in association with the mitochondria. *Journal of Cell Biology* 72, 714–725.
- Smallridge, R., 2008. Membrane trafficking: Mitochondria-peroxisome connection. *Nature Reviews Molecular Cell Biology* 9, 186–187.
- Smeele, K.M., Southworth, R., Wu, R., *et al.*, 2011. Disruption of hexokinase II-mitochondrial binding blocks ischemic preconditioning and causes rapid cardiac necrosis. *Circulation* 108, 1165–1169.
- Som, P., Atkins, H.L., Bandyopadhyay, D., *et al.*, 1980. A fluorinated glucose analog, 2-fluoro-2-deoxy-D-glucose (F-18): non-toxic tracer for rapid tumor detection. *Journal of Nuclear Medicine* 21, 670–675.
- Sottocasa, G.L., Kuylenskierna, B., Ernster, L., Bergström, A., 1967. An electron transport system associated with the outer mitochondrial membrane of liver mitochondria. A biochemical and morphological study. *Journal of Cell Biology* 32, 415–437.
- Stewart, B.N., Wild, C.P. (Eds.), 2015. *World Cancer Report 2014*. France: World Health Organization.
- Sugiura, A., Nagashima, S., Tokuyama, T., *et al.*, 2013. Mitol regulates endoplasmic reticulum-mitochondria contacts via mitofusin 2. *Molecular Cell* 51, 20–34.
- Wang, G., Chen, H.W., Oktay, Y., *et al.*, 2010. PNPase regulates RNA import into mitochondria. *Cell* 142, 456–467.
- Wolff, N.A., Garrick, L.M., Zhao, L., Garrick, M.D., Thevenod, F., 2014. Mitochondria represent another locale for the divalent metal transporter 1 (DMT1). *Channels* 8, 458–466.
- Wu, R., Wyatt, E., Chawia, K., *et al.*, 2012. Hexokinase II knockdown results in exaggerated cardiac hypertrophy via increased ROS production. *EMBO Molecular Medicine* 4, 633–646.
- Yonshiro, R., 2006. A novel ubiquitin ligase plays a critical role in mitochondrial dynamics. *EMBO Journal* 25, 3618–3626.

Neuronal Action Potentials and Ion Channel Allostery

SE Kotermanski and M Cascio, Duquesne University, Pittsburgh, PA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Action potential A highly coordinated sequence of events that requires the activation of voltage-gated Na^+ and K^+ channels which generates an all-or-nothing wave of electrical activity. Initiated at the axon hillock an action potential travels along the length of the axon until reaching the axon terminal where it results in the activation of voltage-gated Ca^{2+} channels and release of neurotransmitter.

Electrochemical gradient A result of ionic concentration differences between two areas (such as the intracellular and extracellular space) and the electrical potential associated with relevant ions. These two properties determine which direction a permeant ion will move between the two areas. Electrochemical gradients provide a source of energy in which cells can use to transmit information as well as various other processes (such as an energy source for transport proteins).

Excitatory postsynaptic potentials (EPSP) A transient depolarization of the neuronal cell membrane that results from the opening of ligand-gated ion channels that are typically permeable to Na^+ , K^+ , and/or Ca^{2+} . EPSPs increase the probability of an action potential being generated by driving the neuronal membrane potential toward the threshold potential.

Inhibitory postsynaptic potentials (IPSP) A transient hyperpolarization of the neuronal cell membrane that results from the opening of ligand-gated ion channels that are typically permeable to Cl^- . IPSPs decrease the probability of an action potential being generated by driving the neuronal membrane potential away from the threshold potential.

Ionotropic receptors Receptors that consist of a pore region that opens in response to a variety of stimuli (such as ligand binding, voltage, or mechanical stimuli). Once activated, permeant ions (typically Na^+ , K^+ , Ca^{2+} , or Cl^-) rapidly pass through the pore of the channel in a direction determined by their electrochemical gradient.

Membrane potential The electrical potential difference (usually stated in mV) between the intracellular and extracellular surface of the cell membrane at any given

moment. The membrane potential varies over time as a result of the activity of ionotropic receptors.

Metabotropic receptors Unlike ionotropic receptors, metabotropic receptors do not contain a permeant ion pore which allows for the passage of ions upon activation. Metabotropic receptors contain an extracellular ligand-binding domain and an intracellular G-protein binding domain. The binding of ligand to metabotropic receptors results in activation of G-proteins that dissociate from the receptor and transduce the signal by activating intracellular pathways, resulting in a less rapid signaling cascade.

Resting membrane potential The electrical potential of a cell in the absence of synaptic activity. Results from the unequal distribution of charged particles between the intracellular and extracellular space. The unequal distribution of charges is the result of the activity of electrogenic pumps. The extracellular space is conventionally defined as neutral and the inside of the cell typically lies between -60 and -70 mV.

Saltatory conduction Used to describe action potential conduction along myelinated axons. In myelinated axons action potentials jump between nodes of Ranvier which contain a high concentration of voltage-gated Na^+ channels that serve to regenerate the signal strength which then passively propagates along the internode regions. This mechanism dramatically increases the rate of action potential propagation.

Synapse The region between the axon terminal of a neuron and its target. Synapses are areas where the cell membrane of a neuron and its target come into close proximity of each other. Information is then passed from the presynaptic cell to the postsynaptic cell either by an electrical connection via gap junctions or through the release from the presynaptic cell of chemical neurotransmitters that bind to receptors on the postsynaptic cell.

Threshold potential The critical membrane potential needed to be reached in order to open voltage-gated Na^+ channels and initiate an action potential. The threshold potential typically lies between -40 and -55 mV.

Introduction

An organism's ability to sense, process, and respond to cues in their environment as well as regulate essential physiological processes (such as heart rate and respiration) necessary for life is dependent upon rapid and reliable communication between neurons of the nervous system as well as between neurons and their target tissues. Neurons use a combination of electrical and chemical signals to accomplish this task. Electrical transmission of signals within (and across) neurons as well as

the conversion of chemical signals at the synapse into an electrical signal is possible due the presence of highly selective channels that allow for the passage of ions across the cell membrane. The activity of these ion channels is regulated by either changes in the neurons membrane potential (voltage-gated), the binding of chemical transmitter (ligand-gated), or mechanical stimuli (stretch receptors). The direction of ion flux through activated channels is dependent upon their electrochemical gradient and can shift the neuronal membrane potential more negative (hyperpolarization) or more positive

(depolarization) in relation to the neurons resting membrane potential. A large enough depolarizing stimulus (~ 10 mV from rest) can result in the firing of an action potential, which leads to transmitter release into the synapse and subsequent passage of information to the postsynaptic neuron (Kandel *et al.*, 2000; Hille, 2001; Purves *et al.*, 2012). (These sources were invaluable and were a source for much of the general information throughout this article.)

Electrical Properties of Neurons

The structure of the neuronal cell membrane creates an effective barrier between the intracellular and extracellular environments. The organization of the lipid bilayer (~ 6 – 8 nm thick) consists of two leaflets of non-covalently associated amphiphilic lipid molecules with their exterior facing water accessible polar head groups separated by a thin layer of interior hydrophobic lipid tails (Pfenninger, 1978). The close-packing of the hydrophobic interior of the leaflets, comprised of apposed lipid tails, excludes water and provides a very high energy barrier for the passage of charged particles across the cell membrane, while the polar head groups allow for the accumulation of charged particles along the surface of the cell membrane. This separation of charged particles allows for a potential difference between the inside and outside of the neuron. The only effective way for charged particles to traverse the cell membrane is through specialized proteins incorporated into the lipid bilayer. These proteins may be transporters or ion channels, and amongst the many stimuli that they are responsive to, may be gated by ligands, mechanical stress, temperature, or voltage.

The resting membrane potential is the voltage difference that exists across the neuronal membrane and is typically between -60 mV and -70 mV. By convention the outside of the cell is generally considered to be neutral, resulting in the inside of the cell resting at a negative potential. The value of the resting membrane potential is determined by the concentration of ions in the intracellular and extracellular space as well as the presence of non-gated (leak) channels and electrogenic pumps that move ions across the cell membrane at the expense of energy. The concentration difference of two highly important ions (sodium (Na^+) and potassium (K^+)) involved in neuronal signaling is also maintained by the activity of the Na^+/K^+ pump (Na^+/K^+ ATPase) (Yu, 2003; Wright, 2004). A large percentage of cellular energy is utilized by Na^+/K^+ ATPase, an integral membrane protein (Shinoda *et al.*, 2009) that pumps out 3 Na^+ for every 2 K^+ in, at the cost of a single ATP (with this stoichiometry contributing to the relatively electronegative interior of resting cells).

The direction of ion flow through leak channels is determined by the concentration difference on either side of the cell membrane (as maintained by pumps such as Na^+/K^+ ATPase) as well as the electrical charge associated with that ion. Simple diffusion states that a particle will move from an area of high concentration to an area of low concentration, however, since ions are charged their direction of movement is also influenced by the voltage difference between the intracellular and extracellular space. Since the resting membrane potential is negative, cations would be pulled inside the

cell with the reverse being true for anions, based on electro-motive force alone. The direction of the passive flow of ions through open channels is determined by the electrochemical gradient, the sum of the chemical and electrical gradients, and the equilibrium potential, E_{ion} , where these two forces are equivalent, resulting in no net flux of the ion through an open channel, was defined by Nernst, such that:

$$E_{\text{ion}} = \frac{RT}{zF} \ln \frac{[\text{ion}]_{\text{ext}}}{[\text{ion}]_{\text{int}}}$$

where R is the gas constant, T is the temperature in degrees Kelvin, z is the valence of the ion, F is the Faraday constant, and $[\text{ion}]_{\text{ext}}$ and $[\text{ion}]_{\text{int}}$ is the extracellular and intracellular ionic concentration, respectively (Wright, 2004).

Typically the electrochemical gradient for the major ions involved in neuronal signaling would result in Na^+ , calcium (Ca^{2+}) and chloride (Cl^-) influx and K^+ efflux when channels permeant for each respective ion are activated. An influx of Na^+ or Ca^{2+} would result in a depolarizing stimulus making the membrane potential more positive and K^+ efflux would hyperpolarize the membrane potential making it more negative. Since Cl^- is associated with a negative charge its influx would also serve to hyperpolarize the membrane potential. Of note, under typical conditions the equilibrium potential for Cl^- , E_{Cl} , is very close to the resting potential of a neuron. Thus, small changes in the low $[\text{Cl}^-]_{\text{int}}$ can dramatically affect the direction of flow through Cl^- selective channels. During early development the $[\text{Cl}^-]_{\text{int}}$ of immature neurons is higher than in mature neurons of the adult resulting in an efflux of Cl^- through activated channels, thereby leading to a depolarizing stimulus. During development, there is increased expression of a K^+/Cl^- cotransporter (KCC2), which shuttles one K^+ and one Cl^- ion out of the cell thereby reducing the $[\text{Cl}^-]_{\text{int}}$, resulting in a change in activated Cl^- channels from being depolarizing (inward flux) to hyperpolarizing (outward flux) (Laube *et al.*, 2002; Kirsch, 2006; Avila *et al.*, 2013).

Anatomy of a Neuron and Flow of Information

The neuron doctrine states that the basic signaling unit of the nervous system are separate discrete cells with processes arising out of the cell body. Neurons can be divided into many different classifications based on the number of processes arising from the cell body. Typical neurons share four regions: the cell body (soma), dendrites, axon, and axon terminal. Each region is distinctly organized for its role in the transmission of information. The cell body contains the nucleus and is responsible for gene expression and cellular metabolism. The dendrites and axon arise from the cell body. Dendrites typically have many branches and can form elaborate dendritic trees that receive incoming information to the neuron. The expansive dendritic tree of Purkinje cells of the cerebellum can receive information from approximately 150 000 synapses. Chemical transmission at the synapses of the dendrites gets converted into an electrical signal that gets passively conducted toward the cell body. The cell body integrates this information along with incoming signals arising from synapses onto the

cell body itself. Also arising from the cell body is a specialized region called the axon hillock which gives rise to the axon. The axon hillock is a region of the neuron with a high density of voltage-gated Na^+ channels (for review of voltage-sensing, see Catterall, 2010). If the summation of signals in this region of the neuron depolarizes the cell beyond a threshold potential, an action potential is initiated which is actively propagated along the axon to the axon terminal (Kress and Mennerick, 2009). The action potential activates voltage-gated Ca^{2+} channels in the axon terminal. The resulting transient increase in $[\text{Ca}^{2+}]_{\text{int}}$ results in fusion of neurotransmitter-containing vesicles, and the release of the neurotransmitter from these synaptic vesicles into the synaptic cleft (Sudhof, 2004, 2013) leading to chemical activation of corresponding ligand-gated ion channels on the postsynaptic neuron and subsequent passage of information.

For a neuron to fire an action potential a depolarizing stimulus must be strong enough to reach threshold at the axon hillock. The activation of excitatory ion channels due to release of glutamate (Glu), acetylcholine (ACh), or serotonin (5-hydroxytryptamine (5-HT)) at a single synapse on the dendrite of a neuron results in a subsequent depolarizing electrical signal, however that lone signal will not be large enough to generate an action potential. The signal may also be inhibitory, as the Cl^- channels activated by γ -aminobutyrate (GABA) or glycine (Gly) typically lead to hyperpolarization in adult neurons (as described earlier, increased $[\text{Cl}^-]_{\text{int}}$ can result in depolarization upon activation of these channels) (Laube *et al.*, 2002; Kirsch, 2006; Avila *et al.*, 2013). Neurons receive thousands of incoming depolarizing (excitatory postsynaptic potentials (EPSP)) and hyperpolarizing (inhibitory postsynaptic potentials (IPSP)) stimuli. IPSP and EPSP stimuli are integrated together as the signal travels toward the axon hillock, and it is the summation of these incoming signals that determines whether or not an action potential will be generated (Zheng and Raman, 2010).

Depolarizing and hyperpolarizing stimuli are influenced by the neuron's axial (r_a) and membrane resistances (r_m). The magnitude of incoming stimuli to a neuron decreases as it travels away from its point of origin. This occurs due to the effects of r_a , the resistance of the cytoplasm that the signal encounters as it travels along the length of the dendrite. The membrane resistance of a pure lipid bilayer is very high, however, the presence of leak channels within the membrane allows for current to pass out of the neuron instead of coursing through it, also decreasing the magnitude of the stimuli as a function of the distance the signal propagates. The nervous system has developed a way of increasing the resistance of regions of its cell membrane by adding insulating layers of membrane, called myelination (Pfenninger, 1978). Myelination decreases the loss of current through the cell membrane, thereby allowing for a stimulus to travel further with limited degradation. The length constant (λ) of a neuron is the distance in which a signal decays to $1/e$, or approximately 37%, of its original magnitude and can be calculated by the following equation:

$$\lambda = \sqrt{\frac{r_m}{r_a}}$$

A measure of how fast a stimulus decays at the point of initiation is the time constant (τ). The longer the time constant the longer the stimulus will remain at the point of initiation

leading to a greater chance of summation with subsequent signals to occur.

Taken together, these parameters describe how far from the point of initiation and how long after initiation a given stimulus will alter the membrane potential from rest. Spatial summation results when the effects of two stimuli arising from different areas of the neuron interact, and temporal summation occurs when a second stimulus arrives at a synapse before a previous stimulus has had time to fully decay. The final effect of integrating the multitude of incoming signals to a given neuron determines whether that neuron will fire an action potential (the threshold potential, described in the next section), resulting in the postsynaptic release of chemical neurotransmitter and the propagation of information to other neurons to which it is synaptically connected. If the summation of all synaptic stimuli arriving along the dendrites and cell body of a neuron result in a depolarization strong enough to reach threshold at the axon hillock an action potential will be initiated.

Action Potentials

The action potential is an all-or-nothing electrical wave that is initiated at the axon hillock and propagates toward the axon terminal via highly coordinated sequential activation of various ion channels that have differential selectivities with respect to ion permeation (Figure 1). Action potentials are not graded signals like EPSPs or IPSPs, and are either initiated fully or not at all. Although the shape of an action potential is relatively consistent across the many types of neurons within the nervous system, variations in ion channel properties do introduce some differences (Kress and Mennerick, 2009).

The axon hillock contains a very high concentration of voltage-gated Na^+ channels that become activated once a critical membrane potential is reached, the threshold potential. The threshold potential is a membrane depolarization of approximately 10 mV from rest. This change in voltage results in a conformational change of the voltage-gated Na^+ channel that opens up a central pore that is highly selective for Na^+ , allowing its influx as directed by its electrochemical gradient (Bezaniila, 2006). This influx further depolarizes nearby areas of the cell. Voltage-gated Na^+ channels are a single polypeptide chain of amino acids that contain four repeating homologous domains each containing six transmembrane segments (S1–S6) and a re-entrant pore-lining region (P region) that lies between the S5 and S6 segments (Payandeh *et al.*, 2011). The four P regions of the protein contain chemical groups that effectively replace the hydration shell of Na^+ and stabilize the charge on the dehydrated Na^+ as it passes through the selectivity filter of the pore (which is specific for Na^+ , and impermeant to K^+). The S4 region, although mostly hydrophobic, contains positively charged amino acids that sense the voltage difference across the cell membrane. It is hypothesized that this portion of the protein moves in response to membrane depolarization, thereby causing a conformation change in the channel allowing for the opening of an activation gate and flux of Na^+ through the channel. Voltage-gated Na^+ channels inactivate in a manner that is distinct from just the simple closing of the

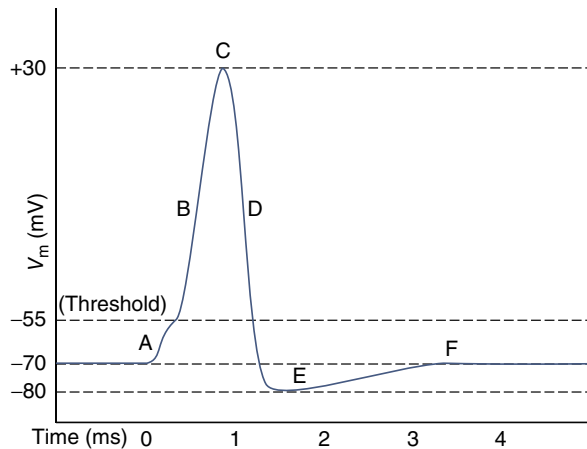


Figure 1 Representative trace of an action potential. EPSPs resulting from the activation of ligand-gated ion channels (such as members of the ligand-gated family of glutamate receptors) depolarizes the neuronal membrane potential, V_m , from its resting potential (~ -70 mV) (A) toward the threshold potential (~ -55 mV). Once the threshold potential is reached voltage-gated Na^+ channels begin to open allowing for Na^+ influx and further depolarization of the neuronal membrane and (B) the rise of the membrane potential toward E_{Na} ($\sim +55$ mV) resulting in the upswing (or rising phase) of the action potential. The membrane depolarization also opens voltage-gated K^+ channels, whose opening kinetics are slower than that of voltage-gated Na^+ channels, (C) resulting in the peak of the action potential and preventing it from rising all the way to E_{Na} . The efflux of K^+ as well as the slow inactivation of voltage-gated Na^+ channels results in (D) the downswing (or falling phase) of the action potential. V_m falls toward E_{K} (~ -80 mV) (the after hyperpolarization) and (E) gradually returns toward the resting membrane potential (F) as voltage-gated K^+ channels close and resting ionic gradients are returned by the actions of leak channels and the Na^+/K^+ ATPase.

activation gate. The inactivation gate of voltage-gated Na^+ channels also closes as a result of depolarization, however, more slowly than the opening of the activation gate. This results in a brief window (milliseconds) during which both gates are open and (Na^+) can pass and allow for the neuronal membrane potential to approach E_{Na} (typically $\sim +55$ mV). This sequence of events results in the upswing or rising phase of the action potential. It is not until the membrane potential returns to near rest that both gates reset and voltage-gated Na^+ channels resume their resting conformation of a closed activation gate and open inactivation gate. Until voltage-gated Na^+ channels return to their resting conformation, another action potential cannot be initiated since the channels are inactivated and unable to allow Na^+ influx. The time it takes for voltage-gated Na^+ channels to return to their resting confirmations and therefore allow for the initiation of another action potential is called the absolute refractory period.

The downswing, or falling phase, of action potentials results from the subsequent opening of voltage-gated K^+ channels, as well as voltage-gated Na^+ channel inactivation. Voltage-gated K^+ channels also contain six transmembrane regions (S1–S6) and a re-entrant P region between S5 and S6 (Jiang *et al.*, 2003). However, each gene product of voltage-

gated K^+ channels contains only one copy of this sequence. Four subunits assemble to produce the functional channel that is highly selective for K^+ . Similar to voltage-gated Na^+ channels, it has been shown that the P region lines the central pore and the S4 region is involved in channel activation of voltage-gated K^+ channels. Voltage-gated K^+ channels are activated by membrane depolarization as well, however their activation occurs more slowly than voltage-gated Na^+ channels. When activated, K^+ flows out of the cell along its electrochemical gradient, thereby bringing the membrane potential back near E_{K} (typically ~ -80 mV), which is more negative than the resting membrane potential (after hyperpolarization). The return of the membrane potential toward rest results in the closing of voltage-gated K^+ channels (unlike voltage-gated Na^+ channels, they do not contain an inactivation gate). The slow closing of voltage-gated K^+ channels also makes it more difficult to initiate another action potential since the latent permeability to K^+ will keep the neuronal membrane further from the threshold potential necessary to activate voltage-gated Na^+ channels. This relative refractory period lasts until voltage-gated K^+ channels return to their closed resting confirmations. During the relative refractory period, it is possible to generate another action potential, however, a stronger stimulus is required.

Action Potential Propagation

Action potentials are rapid changes in the membrane potential of neurons that transmit information throughout the nervous system. Once initiated at the axon hillock, the action potential will propagate down the length of the axon (which may be as long as a meter in length in humans) until it reaches the axon terminal. Membrane depolarization occurring at one area of the axon passively spreads to and depolarizes adjacent areas. This activates voltage-gated Na^+ channels in the adjacent areas resulting in Na^+ influx and propagation of the signal. Although the influx of Na^+ would diffuse in both directions along the length of the axon, the refractory periods prevent back propagation of the signal. While still rapid, action potential propagation rate is limited by the relatively small diameter of the axon which creates high r_a . Increasing the diameter of the axon would increase the rate of signal propagation, however, this is not evolutionarily feasible due to the large number of axons that exist in the vertebrate nervous system. Myelination of axons by glial cells increases the rate of action potential propagation while still allowing for the small diameter of axons. In general the conduction velocity of an action potential in unmyelinated axons ranges from 0.5 to 10 m s^{-1} and up to approximately 150 m s^{-1} in myelinated axons.

Oligodendrocytes and Schwann cells are specialized glial cells of the nervous system that wrap the axons of neurons with multiple layers of their own cell membrane. This increases r_m preventing the loss of current across the axonal membrane resulting in the signal propagating further along its length. However, the multiple overlapping layers of membrane prevents the incorporation of voltage-gated Na^+ channels that are needed to regenerate the action potential, therefore the signal strength decreases along the length of the axon. This degradation in signal strength is compensated by

leaving spaces ($\sim 2 \mu\text{m}$ in length) of the axon without myelination, called the nodes of Ranvier that contain a high concentration of voltage-gated Na^+ channels. The action potential signal passively propagates along the myelinated internode areas ($\sim 1\text{--}2 \text{ mm}$ in length) until reaching these nodes and activating their Na^+ channels, regenerating the strength of the action potential. This method of saltatory conduction where the action potential signal jumps from node to node drastically increases the speed of action potential propagation.

Action Potential Invasion of the Nerve Terminal

Action potential propagation terminates at the axon terminal, which contains synaptic vesicles that contain neurotransmitter, as well as the machinery needed to release that neurotransmitter into the synaptic cleft (described below). Of critical importance to the release of neurotransmitter upon action potential invasion are voltage-gated Ca^{2+} channels located in the axonal terminus. The structure of voltage-gated Ca^{2+} channels are thought to be homologous to that of voltage-gated Na^+ channels, consisting of a single polypeptide chain that contains four repeating segments of six transmembrane domains (S1–S6) and a re-entrant pore-lining region (P region) between S5 and S6 (Simms and Zamponi, 2014). Like voltage-gated Na^+ channels, voltage-gated Ca^{2+} channels activate in response to membrane depolarization, although much more slowly, allowing for Ca^{2+} influx along its electrochemical gradient. Voltage-gated Ca^{2+} channels begin to allow Ca^{2+} influx near the end of the action potential and the duration of their open time depends on the time course of the action potential. Voltage-gated Ca^{2+} channels inactivate in response to the repolarization of the membrane potential back to resting levels as well as a Ca^{2+} dependent feedback inactivation via cytosolic regulation. $[\text{Ca}^{2+}]_{\text{int}}$ is highly regulated and maintained in the cytoplasm at very low levels, as higher concentrations are toxic. Ca^{2+} is a common second messenger in cells, as this divalent cation is an allosteric regulator of a wide range of Ca^{2+} binding proteins (Villereal and Palfrey, 1989).

Only a small percentage ($\sim 1\%$) of synaptic vesicles present at the axon terminal are immediately available to participate in neurotransmitter release. These vesicles are docked near the neuronal plasma membrane and constitute the readily releasable pool of vesicles. The other pools of vesicles are the reserve pool and the recycling pool (Rizzoli and Betz, 2005). Synaptic vesicles contain transport proteins that are responsible for filling the vesicle with neurotransmitter and soluble N-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) proteins which facilitate vesicular docking and fusion. According to the SNARE hypothesis, the vesicular SNARE protein synaptobrevin forms a complex of four α helices with the neuronal membrane SNARE proteins syntaxin and SNAP-25, which primes the vesicle for fusion with the neuronal membrane (Rizo and Rosenmund, 2008). Other cytosolic proteins have also been proposed to assist in the SNARE complex formation (Weimer and Richmond, 2005; Shin *et al.*, 2010; Trimbuch *et al.*, 2014). The synaptic vesicle associated protein, synaptotagmin, contains multiple Ca^{2+} coordination sites and acts as a Ca^{2+} sensor for triggering vesicle fusion

with the neuronal membrane. The exact mechanism of synaptic vesicle fusion and release of neurotransmitter is still unknown. Evidence exists for a full fusion model where the synaptic vesicle fully fuses and collapses into the neuronal membrane releasing its transmitter into the synaptic cleft. However, evidence also exists for a kiss-and-run model of transmitter release where only a small fusion pore connecting the vesicular and neuronal membranes opens briefly allowing for neurotransmitter to diffuse into the synaptic cleft (Rizzoli and Betz, 2005; Rizo and Rosenmund, 2008).

Fate of Neurotransmitter in the Synaptic Cleft

Once released into the synaptic cleft, neurotransmitter can bind to receptors located on the postsynaptic membrane resulting in the transfer of information from one neuron to the next. The extent of the postsynaptic response to transmitter release is dependent on neurotransmitter local concentration, the total number of receptors present on the postsynaptic membrane, and their functional characteristics (e.g., ligand affinity and kinetics). The postsynaptic response is dependent on the duration of time that the transmitter is present in the synaptic cleft. Removal of neurotransmitter from the synaptic cleft occurs through simple diffusion out of the cleft, enzymatic degradation of neurotransmitter, binding of neurotransmitters by binding proteins, or reuptake of neurotransmitter back into the presynaptic nerve terminal by protein transporters.

Many neurological agents act by affecting the nervous systems mechanism of removing neurotransmitter from the synaptic cleft. For example, nerve gases such as sarin inhibit the activity of acetylcholinesterase (AChE), the enzyme that acts to degrade Ach into acetate and choline, thereby terminating its signal (Abu-Qare and Abou-Donia, 2002). The mechanism of action of cocaine is to inhibit the reuptake activity of the dopamine, norepinephrine and 5-HT transporters that function to transport their respective ligands back into the presynaptic neuron. An area of the brain that plays an important part in processing rewarding and pleasurable stimuli, the nucleus accumbens, receives signals from dopaminergic neurons of the ventral tegmental area. The rewarding effects of cocaine is believed to lie in its ability to increase activity of neurons in the nucleus accumbens by inhibiting the reuptake of dopamine (DA) into the presynaptic nerve, thereby prolonging the postsynaptic response to DA release (Nestler, 2005). The neurotransmitter 5-HT regulates mood and selective serotonin reuptake inhibitors (SSRIs) such as fluoxetine and citalopram are highly specific inhibitors of the 5-HT transporter, thereby increasing the duration of 5-HT's presence in the synaptic cleft and are used in the treatment of depression (Vaswani *et al.*, 2003).

Postsynaptic Receptors/Ligand-Gated Ion Channels

The postsynaptic effect of neurotransmitter release is mediated by two classes of ligand-binding receptors, ionotropic and metabotropic receptors. The binding of neurotransmitter to ionotropic receptors lowers the transitional energy barrier

allowing for gating of the channel to occur. Once activated, permeable ions are allowed to pass across the neuronal membrane along their electrochemical gradient. The major classes of ionotropic neurotransmitter-gated channels are the superfamilies of tetrameric Glu receptors (GluR) (Kumar and Mayer, 2013) and the pentameric ligand-gated ion channel superfamily (pLGIC, also referred to as Cys-loop receptors), such as the nicotinic AChR, GlyR, GABA_AR, and 5-HT₃R (Corringer *et al.*, 2012). Metabotropic receptors mediate postsynaptic responses to neurotransmitter release by activation of G-protein signaling pathways (Katritch *et al.*, 2013). These responses are slower, but may profoundly influence cellular metabolism via transcriptional and translational regulation. Members of the muscarinic AChR (mAChR), GABA_BR, 5-HT₂R, norepinephrine, epinephrine, histamine, dopamine, neuropeptide, endocannabinoid, as well as metabotropic GluR are examples of G-protein coupled receptors.

As described previously, ligand-gated ionotropic receptors can be either excitatory or inhibitory depending on their ion permeability and the electrochemical gradient of the permeate ion. Typically, excitatory receptors are permeable to Na⁺ and K⁺ (and Ca²⁺ as is the case for the N-methyl-D-aspartate (NMDA) subtype of GluR and 5-HT₃ subtype of serotonin receptor upon activation. Activation of these receptors tends to depolarize the membrane potential, increasing the possibility of action potential generation. In mammals, the majority of excitatory neurotransmitter consists of ionotropic receptors for ACh, Glu, and 5-HT. Inhibitory receptors are permeable to Cl⁻ under cellular conditions that result in Cl⁻ influx, leading to further hyperpolarization of the membrane potential, thereby making it more difficult in reaching the threshold for action potential generation. Inhibitory neurotransmitter receptors contain members of the GlyRs and GABA_ARs (Bowerly and Smart, 2006). It should be noted that these receptors may also be excitatory under certain cellular conditions. For example, as described previously, during development activated GlyRs and GABA_ARs are excitatory due to an increased [Cl⁻]_{int} in immature neurons provides an electrochemical gradient that drives Cl⁻ efflux rather than influx (Laube *et al.*, 2002; Kirsch, 2006; Avila *et al.*, 2013).

Allosteric Effectors

The ion channels, receptors and transporters involved in action potential generation and propagation that are described in this chapter are highly allosteric large macromolecular complexes embedded in the membrane. As such, their activities may be modulated by local lipid composition, posttranslational modifications, hormones, modulatory ligands, metals (e.g., Zn²⁺), redox conditions, and other effectors such as drugs. Much of our knowledge of these molecules was predicated on the discovery of inhibitors, both specific and general, found in venoms of snails, spiders, scorpions, and snakes, amongst other organisms (Terlau and Olivera, 2004; Kularatne and Senanayake, 2014). Given the wide array of compounds that modulate activities and the wide diversity of proteins involved in neuronal signaling, we have chosen to focus the discussion on a single subclass of receptors involved in the generation of action potentials. The remainder of this section will briefly

review allosteric effectors involved in modulating the activity of inhibitory pLGICs (Corringer *et al.*, 2012), given the importance of these channels as therapeutic targets (Lynch and Callister, 2006; D'hulst *et al.*, 2009).

Inhibitory ligand-gated receptors in the adult are gated by the neurotransmitters GABA and Gly and are permeable to Cl⁻. It should be mentioned that Gly also has excitatory actions as a co-agonist at the NMDA subtype of GluRs (Johnson and Ascher, 1987). Under typical conditions in adults, the reversal potential for Cl⁻ (E_{Cl}) is approximately -70 mV, which is usually more negative than the neuronal resting potential (~ -65 mV). As a result, activation of GABA_AR or GlyR results in Cl⁻ influx, resulting in hyperpolarization of the membrane potential, making it difficult for action potentials to be generated. In cases where the neuronal resting potential is equal to E_{Cl} activation of Cl⁻ permeable channels will not hyperpolarize the membrane potential but acts to keep the potential near E_{Cl} thereby limiting the amplitude of EPSPs and decreasing the probability of action potential generation. This process of inhibition through increasing the membrane conductance is referred to as shunting.

GABA_ARs mediate a majority of inhibitory neurotransmission in the brain and spinal cord while GlyRs are mostly localized to the spinal cord, brain stem, caudal brain, and retina (Laube *et al.*, 2002). As briefly described previously, this subset of pLGICs, which also contains the excitatory nAChR, and 5-HT₃R, are pentameric assemblies, often heteromers comprised of multiple gene products, arranged quasi-symmetrically around a central pore. Each subunit contains an N-terminal ligand-binding extracellular domain (ECD) with a conserved Cys-disulfide loop (the source of their alternate nomenclature of Cys-loop receptor superfamily), four transmembrane segments (M1–M4) and a C-terminal extracellular tail (Hibbs and Gouaux, 2011). The M2 segments line the pore, and in GABA_AR and GlyR these pore-lining regions contain positively charged amino acids (at physiological pH) at critical sites that impart anion selectivity (the pore is wide enough for hydrated ions, and while broadly anion-specific, under physiological conditions primarily conducts Cl⁻ current). Three GABA_AR subunit isoforms (α , β , and γ) combine in various stoichiometries to form most of the GABA_AR in the mammalian brain, however, other isoforms exist (σ , ϵ , π , θ , and ρ). The most common GABA_AR type in the brain is comprised of 2 α , 2 β , and 1 γ subunit. An important modulator of GABA_ARs, benzodiazepines, is believed to bind between the α and γ subunits. GlyRs are formed by only two subunit isoforms, α and β . The subunit composition of each type of receptor modulates functional characteristics such as ligand affinity, gating kinetics, and channel conductance. In all pLGICs, the neurotransmitter binding site is located between adjacent subunits in their ECD (Lynagh and Pless, 2014).

Barbiturates, benzodiazepines, neurosteroids, ethanol, and inhaled anesthetics are all positive allosteric modulators of GlyRs and GABA_ARs. Given their inhibitory role, increasing the activity of these receptors can result in sedative hypnotic, anxiolytic, and muscle relaxant effects. The variety and extent of effects experienced by these inhibitory receptor targeted drugs is likely mediated by the differential subunit composition of pLGICs. These modulatory effectors bind to sites that are distinct from the neurotransmitter binding site

and their activities, in part, are due to their preferential affinity to and stabilization of different allosteric states of their receptor targets (Taly *et al.*, 2009; Corringer *et al.*, 2012). For example, the binding of phenobarbital to GABA_ARs does not directly activate the receptor, but potentiates activity of GABA bound receptors by altering channel properties such as the channel's mean open time. Inhibitors of GABA_ARs exist, however, they are therapeutically less useful due to anxiety-inducing and convulsant effects. The convulsive alkaloid strychnine, is a highly specific and potent inhibitor of adult GlyRs (it does not bind to neonatally expressed GlyRs) that is fatal in high doses, but is a performance enhancer at very low concentrations.

The inhibitory role of pLGICs has been exploited by Lester and colleagues to express transgene insect homologs in sections of the brain to study the effect of targeted neuronal silencing (Slimko *et al.*, 2002). These studies utilized a homologous insect GluCl, a Glu-gated Cl⁻ channel that is gated by ivermectin. Recently the structure of this protein has been resolved (Hibbs and Gouaux, 2011), and its structure is highly conserved to that of other bacterial pLGICs whose structure had been determined (Hilf and Dutzler, 2008; Bocquet *et al.*, 2009; Hilf and Dutzler, 2009). Given that GlyRs mediate inhibition of signals in the spinal cord and brain stem, they are attractive candidates for treating pain. Transgene expression of human GlyR in peripheral nerves using neurotrophic Herpes simplex virus engineered to express the $\alpha 1$ subunit of GlyR was shown to be effective in reducing pain in animal studies (Goss *et al.*, 2011). More recently, mutant forms of human GlyR that are activated by low levels of ivermectin, an approved FDA drug with limited off-site targets at low concentrations in mammals, has been characterized (Lynagh and Lynch, 2010), and these studies suggest that one may potentially utilize transgene expression of inhibitory pLGICs to selectively inhibit targeted neurons that may be safely activated via systemic delivery of low levels of ivermectin.

See also: Molecular Principles, Components, Technology, and Concepts: Membranes: Membrane Potential: Concepts; Composition, Physical Properties, and Curvature. **Organelles:** Structure and Function: Synaptosomes and Synaptic Vesicles

References

- Abu-Qare, A.W., Abou-Donia, M.B., 2002. Sarin: Health effects, metabolism, and methods of analysis. *Food and Chemical Toxicology* 40, 1327–1333.
- Avila, A., Nguyen, L., Rigo, J.M., 2013. Glycine receptors and brain development. *Frontiers in Cellular Neuroscience* 7, 184.
- Bezanilla, F., 2006. The action potential: From voltage-gated conductances to molecular structures. *Biological Research* 39, 425–435.
- Bocquet, N., Nury, H., Baaden, M., *et al.*, 2009. X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. *Nature* 457, 111–114.
- Bowery, N.G., Smart, T.G., 2006. GABA and glycine as neurotransmitters: A brief history. *British Journal of Pharmacology* 147 (Suppl 1), S109–S119.
- Catterall, W.A., 2010. Ion channel voltage sensors: Structure, function, and pathophysiology. *Neuron* 67, 915–928.
- Corringer, P.J., Poitevin, F., Prevost, M.S., *et al.*, 2012. Structure and pharmacology of pentameric receptor channels: From bacteria to brain. *Structure* 20, 941–956.
- D'hulst, C., Atack, J.R., Kooy, R.F., 2009. The complexity of the GABAA receptor shapes unique pharmacological profiles. *Drug Discovery Today* 14, 866–875.
- Goss, J.R., Cascio, M., Goins, W.F., *et al.*, 2011. HSV delivery of a ligand-regulated endogenous ion channel gene to sensory neurons results in pain control following channel activation. *Molecular Therapy* 19, 500–506.
- Hibbs, R.E., Gouaux, E., 2011. Principles of activation and permeation in an anion-selective Cys-loop receptor. *Nature* 474, 54–60.
- Hilf, R.J., Dutzler, R., 2008. X-ray structure of a prokaryotic pentameric ligand-gated ion channel. *Nature* 452, 375–379.
- Hilf, R.J., Dutzler, R., 2009. Structure of a potentially open state of a proton-activated pentameric ligand-gated ion channel. *Nature* 457, 115–118.
- Hille, B., 2001. *Ion Channels of Excitable Membranes*. Sunderland, MA: Sinauer Assoc., Inc.
- Jiang, Y., Lee, A., Chen, J., *et al.*, 2003. X-ray structure of a voltage-dependent K⁺ channel. *Nature* 423, 33–41.
- Johnson, J.W., Ascher, P., 1987. Glycine potentiates the NMDA response in cultured mouse brain neurons. *Nature* 325, 529–531.
- Kandel, E., Schwartz, J., Jessell, T., 2000. *Principles of Neural Science*. New York, NY: McGraw-Hill.
- Katritch, V., Cherezov, V., Stevens, R.C., 2013. Structure-function of the G protein-coupled receptor superfamily. *Annual Review of Pharmacology and Toxicology* 53, 531–556.
- Kirsch, J., 2006. Glycinergic transmission. *Cell and Tissue Research* 326, 535–540.
- Kress, G.J., Mennerick, S., 2009. Action potential initiation and propagation: Upstream influences on neurotransmission. *Neuroscience* 158, 211–222.
- Kularatne, S.A., Senanayake, N., 2014. Venomous snake bites, scorpions, and spiders. *Handbook of Clinical Neurology* 120, 987–1001.
- Kumar, J., Mayer, M.L., 2013. Functional insights from glutamate receptor ion channel structures. *Annual Review of Physiology* 75, 313–337.
- Laube, B., Maksay, G., Schemm, R., Betz, H., 2002. Modulation of glycine receptor function: A novel approach for therapeutic intervention at inhibitory synapses? *Trends in Pharmacological Sciences* 23, 519–527.
- Lynagh, T., Lynch, J.W., 2010. An improved ivermectin-activated chloride channel receptor for inhibiting electrical activity in defined neuronal populations. *Journal of Biological Chemistry* 285, 14890–14897.
- Lynagh, T., Pless, S.A., 2014. Principles of agonist recognition in Cys-loop receptors. *Frontiers in Physiology* 5, 160.
- Lynch, J.W., Callister, R.J., 2006. Glycine receptors: A new therapeutic target in pain pathways. *Current Opinion in Investigational Drugs* 7, 48–53.
- Nestler, E.J., 2005. The neurobiology of cocaine addiction. *Science & Practice Perspectives* 3, 4–10.
- Payandeh, J., Scheuer, T., Zheng, N., Catterall, W.A., 2011. The crystal structure of a voltage-gated sodium channel. *Nature* 475, 353–358.
- Pfenninger, K.H., 1978. Organization of neuronal membranes. *Annual Review of Neuroscience* 1, 445–471.
- Purves, D., Augustine, G.J., Fitzpatrick, D., *et al.*, 2012. *Neuroscience*. Sunderland, MA: Sinauer Associates.
- Rizo, J., Rosenmund, C., 2008. Synaptic vesicle fusion. *Nature Structural & Molecular Biology* 15, 665–674.
- Rizzoli, S.O., Betz, W.J., 2005. Synaptic vesicle pools. *Nature Reviews Neuroscience* 6, 57–69.
- Shin, O.H., Lu, J., Rhee, J.S., *et al.*, 2010. Munc13 C2B domain is an activity-dependent Ca²⁺ regulator of synaptic exocytosis. *Nature Structural & Molecular Biology* 17, 280–288.
- Shinoda, T., Ogawa, H., Cornelius, F., Toyoshima, C., 2009. Crystal structure of the sodium-potassium pump at 2.4 Å resolution. *Nature* 459, 446–450.
- Simms, B.A., Zamponi, G.W., 2014. Neuronal voltage-gated calcium channels: Structure, function, and dysfunction. *Neuron* 82, 24–45.
- Slimko, E.M., McKinney, S., Anderson, D.J., Davidson, N., Lester, H.A., 2002. Selective electrical silencing of mammalian neurons *in vitro* by the use of invertebrate ligand-gated chloride channels. *Journal of Neuroscience* 22, 7373–7379.
- Sudhof, T.C., 2004. The synaptic vesicle cycle. *Annual Review of Neuroscience* 27, 509–547.
- Sudhof, T.C., 2013. Neurotransmitter release: The last millisecond in the life of a synaptic vesicle. *Neuron* 80, 675–690.
- Taly, A., Corringer, P.J., Guedin, D., Lestage, P., Changeux, J.P., 2009. Nicotinic receptors: Allosteric transitions and therapeutic targets in the nervous system. *Nature Reviews Drug Discovery* 8, 733–750.
- Terlau, H., Olivera, B.M., 2004. Conus venoms: A rich source of novel ion channel-targeted peptides. *Physiological Reviews* 84, 41–68.
- Trimbuch, T., Xu, J., Flaherty, D., *et al.*, 2014. Re-examining how complexin inhibits neurotransmitter release: SNARE complex insertion or electrostatic hindrance? *Elife*. e02391.

- Vaswani, M., Linda, F.K., Ramesh, S., 2003. Role of selective serotonin reuptake inhibitors in psychiatric disorders: A comprehensive review. *Progress in Neuro-Psychopharmacology & Biological Psychiatry* 27, 85–102.
- Villereal, M.L., Palfrey, H.C., 1989. Intracellular calcium and cell function. *Annual Review of Nutrition* 9, 347–376.
- Weimer, R.M., Richmond, J.E., 2005. Synaptic vesicle docking: A putative role for the Munc18/Sec1 protein family. *Current Topics in Developmental Biology* 65, 83–113.
- Wright, S.H., 2004. Generation of resting membrane potential. *Advances in Physiology Education* 28, 139–142.
- Yu, S.P., 2003. Na(+), K(+)-ATPase: The new face of an old player in pathogenesis and apoptotic/hybrid cell death. *Biochemical Pharmacology* 66, 1601–1609.
- Zheng, N., Raman, I.M., 2010. Synaptic inhibition, excitation, and plasticity in neurons of the cerebellar nuclei. *Cerebellum* 9, 56–66.

Cystic Fibrosis

NA Bradbury, Chicago Medical School, North Chicago, IL, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

ABC transporter A superfamily of proteins comprised of 2 membrane domains and 2 highly conserved adenosine triphosphate (ATP)-binding cassettes. Energy from ATP hydrolysis is used to move peptides, polysaccharides, signaling molecules, drugs, and ions across the cell membrane.

Airway surface liquid (ASL) A thin layer of aqueous fluid and mucus covering the airway surface between the airway cells and the air space. Important in mucociliary clearance and bacterial killing. The height of the ASL is considerably reduced in cystic fibrosis.

Apical membrane A subdomain of the plasma membrane on the apical side (the region of the membrane facing the lumen or space) of epithelial cells in a body tube or cavity, separated from the basolateral domain by tight junctions. Typically bearing micro-villi and forming the membrane across which secretion occurs.

Basolateral membrane A subdomain of the plasma membrane adjacent to the basement membrane in

epithelial cells. Separated from the apical membrane by tight junctions. The sodium pump is always found on the basolateral membrane.

Corrector A class of CFTR drugs that act to facilitate exit of mutant CFTR from the endoplasmic reticulum for traffic to the cell surface.

Ivacaftor Ivacaftor, trade name Kalydecoc, developed by Vertex Pharmaceuticals as VX-770, is a CFTR potentiator drug approved for the treatment of cystic fibrosis patients with the specific G551D mutation. Ivacaftor is N-[2,4-Bis (1,1-dimethylethyl)-5-hydroxyphenyl]-4-oxo-1,4, dihydroquinoline-3-carboxamide.

Mucociliary clearance The clearance of mucus, bacteria, and other debris from the airways by unidirectional beating of airway cell cilia.

Potentiator A class of CFTR drugs that act to increase the activity of mutant CFTR proteins that reach the cell surface.

Introduction

The finding of the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene in 1987 was a landmark in human genetics, and represented the first gene ever identified or sequenced without any knowledge of its protein product (Kerem *et al.*, 1989; Riordan *et al.*, 1989; Rommens *et al.*, 1989). The discovery of the gene caused a major paradigm shift in cystic fibrosis (CF) research, culminating in the discovery and marketing of the first drug to treat the basic defect in CF (McPhail and Clancy, 2013). Research into CF continues to have a marked impact upon many areas of biochemistry, cell biology, and physiology, providing new insights into normal and disease processes. Moreover, in the years since the first description of CF as a unique disease in 1938 (Anderson, 1938), when it was unusual to see a child live past 5 years of age, there has been a dramatic increase in survival such that many patients are now living into their 3rd and 4th decades of life (Moskowitz *et al.*, 2008; Proesmans *et al.*, 2008). Much of the research in CF would not have been possible without the steadfast support of the Cystic Fibrosis Foundation (CFF, see Relevant Websites section). Indeed, the clinical availability of drugs like 'Ivacaftor' would have likely taken considerably longer without research funds from the CFF. As we gain a greater understanding of how mutations in *CFTR* lead to CF disease, we will not only learn more about basic biology, but will also provide clinicians with more therapeutic options as we move into the exciting arena of pharmacogenetics and patient care based upon the individual patient's genotype.

Key Highlights

- CF is caused by mutations in the *CFTR* gene.
- The protein coded by *CFTR* gene functions as a regulated anion channel in many epithelial cells, including the airways, pancreas, intestine, and reproductive tracts, where it regulates salt, fluid, and pH homeostasis.
- Absence or dysfunction of CFTR protein leads to marked lung disease, as well as disruption in the function of the pancreas and other organ systems.
- Patients with CF have characteristically high levels of salt (sodium chloride) in their sweat.

Protein

CF is caused by mutations in the *CFTR* gene. The *CFTR* gene encodes a protein that functions as a regulated anion channel in epithelial cells, transporting negatively charged ions such as chloride (Cl^-), and bicarbonate (HCO_3^-). *CFTR* is a member of a large family of transport proteins, called the adenosine triphosphate binding cassette (ABC) proteins that are expressed in a wide-range of phyla from yeast to human (Higgins, 1992; ter Beek *et al.*, 2014). *CFTR*, however, is only expressed in vertebrates, and is the only known ABC transporter to function specifically as an ion channel. Mutations in ABC transporter family members are linked to many human diseases in addition to CF, including Stargardt macular dystrophy (Allikmets *et al.*, 1997), Tangier disease

(Rust *et al.*, 1999), and Adrenoleukodystrophy (Morita and Imanaka, 2012).

The CFTR channel is a large multidomain membrane glycoprotein with two similar halves (Figure 1(a)), comprised of 1480 amino acids and a calculated molecular mass of 168 kDa. It consists of two, six member, transmembrane domains (TMD1 and TMD2), 2 nucleotide binding domains (NBD1 and NBD2), and, unique among ABC transporters, a regulatory (R) domain (Riordan *et al.*, 1989). Present in the extracellular portion of the protein, in TMD2, are 2 sites

for that attachment of N-linked (asparagine) carbohydrate units (Figure 1(a)). The role of the carbohydrate on CFTR is still not certain. CFTR lacking sugar residues goes to the cell surface, and functions as an ion channel in the same way as sugar containing CFTR (Chang *et al.*, 2008), however the sugar residues may impart protein stability at the cell surface, preventing premature protein degradation. Although the crystal structure of CFTR has not yet been determined, structures for domains of CFTR are available, which, along with homology mapping against known structures of similar proteins, gives

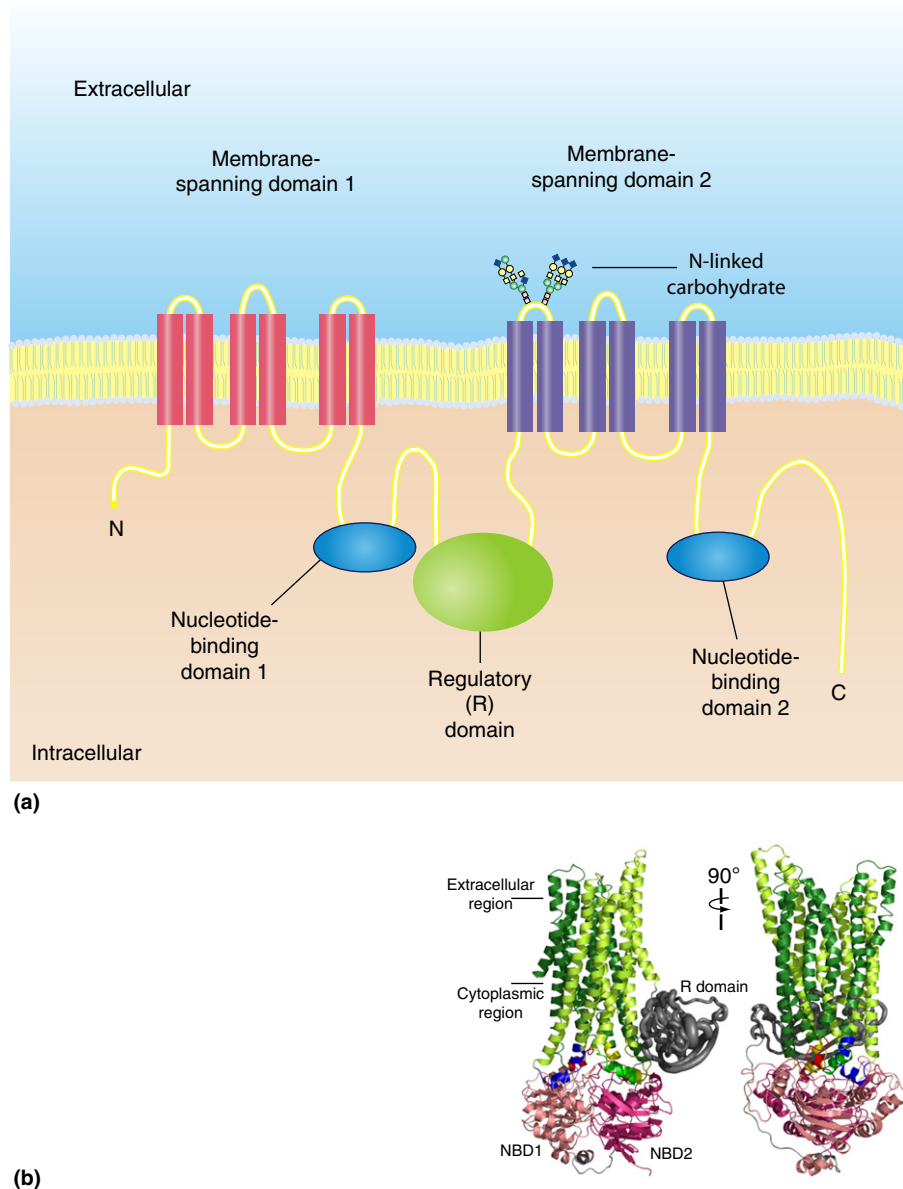


Figure 1 (a) Model showing the domain structure of CFTR. The protein has 1480 amino acids, with both the amino and carboxy termini within the cell cytoplasm. The two membrane spanning domains each contain six transmembrane segments, forming the ion channel. There are six extracellular loops. And glycosylation occurs on asparagine residues at two sites on extracellular loop 4. Also shown are two nucleotide binding domains (the $\Delta F508$ mutation is located in NBD1). Phosphorylation of the regulatory (R) domain is necessary for channel function. (b) Homology model of CFTR based on the ABC transporter Sav1866. The R domain is largely unstructured. Reprinted with permission from Serohijos, A.W., Hegedus, T., Aleksandrov, A.A., *et al.*, 2008. Phenylalanine-508 mediates a cytoplasmic-membrane domain contact in the CFTR 3D structure crucial to assembly and channel function. *Proceedings of the National Academy of Sciences of the USA* 105 (9), 3256–3261.

rise to the likely structure seen in [Figure 1\(b\)](#). The 2 nucleotide binding domains come together to form a pocket for adenosine triphosphate (ATP) binding and hydrolysis, which serves to regulate the gating (opening and closing) of the CFTR channel. The R, or regulatory, domain is mostly unstructured, but possesses multiple consensus phosphorylation sites for protein kinases. The R domain must be phosphorylated by cyclic adenosine monophosphate (cAMP)-dependent protein kinase (PKA) in order to conduct chloride ions ([Riordan, 2008](#)).

CFTR and Protein Trafficking

During its life, CFTR undergoes several trafficking pathways, taking it through various cellular compartments ([Figure 2](#);

[Ameen et al., 2007](#)). Following gene transcription, the mRNA coding for CFTR leaves the nucleus and interacts with ribosomes associated with the endoplasmic reticulum (ER). As the various domains of CFTR are synthesized, the protein starts to fold and achieve its final three-dimensional conformation. CFTR is helped in its folding by a set of proteins, called chaperones, that bind and release CFTR, as CFTR tries to achieve its correct structure. Once CFTR has achieved its final conformation, all the chaperones disengage from CFTR. It takes 7–8 mins for a single CFTR molecule to be made at the rate at which most proteins are made. CFTR folds in a surprisingly inefficient manner, with over half of wild-type CFTR molecules unable to achieve their final appropriate conformation. As well as helping proteins fold, chaperones also form part of the cell's quality control machinery. CFTR molecules that are unable to fold properly remain bound by

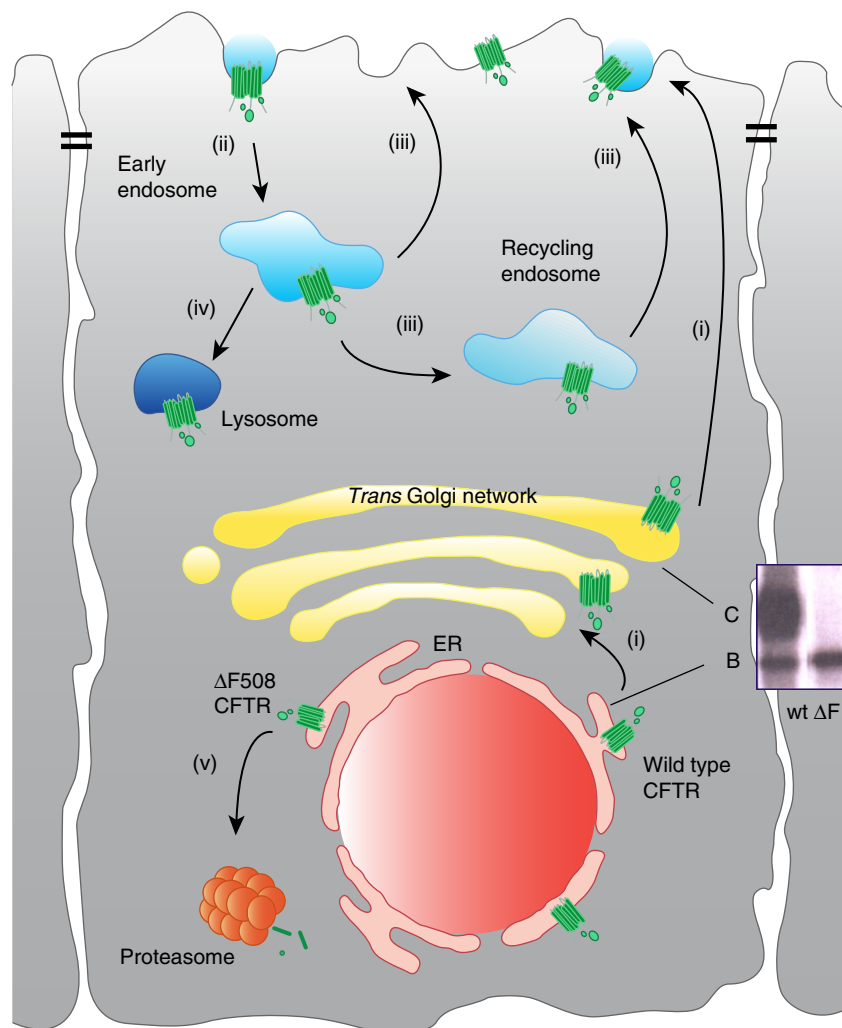


Figure 2 Trafficking pathways for wild-type and $\Delta F508$ CFTR (i) CFTR is translated in the endoplasmic reticulum (ER) where core sugars are added to the protein. Most $\Delta F508$ is recognized as misfolded and subject to proteasomal degradation (v). Wild-type CFTR traffics to the TGN where the core sugars are modified into complex carbohydrates, and then trafficked to the apical plasma membrane (i). (ii) CFTR is efficiently removed from the cell surface by clathrin-mediated endocytosis using trafficking signals embedded in the amino acid sequence of CFTR. (iii) From endosomes, CFTR can recycle back to the cell surface in a direct manner, or via recycling endosomes. (iv) Internalized CFTR can also be directed to lysosomes for degradation. Modified from Ameen, N., Silvis, M., Bradbury, N.A., et al., 2007. Endocytic trafficking of CFTR in health and disease. *Journal of Cystic Fibrosis* 6 (1), 1–14.

chaperones, which then recruit enzymes to tag misfolded CFTR with the small protein ubiquitin. Ubiquitin acts as a signal for degradation, leading to digestion of aberrant proteins by the proteasome. CFTR also undergoes co-translational modifications in the ER, with two small sugar units added initially to asparagine residues in the fourth extracellular loop. This is referred to as 'core-glycosylated' or 'Band B' CFTR.

Once folded and core-glycosylated in the ER, acidic sequences within CFTR interact with coat protein complex II (COPII). This highly regulated process generates COPII coated vesicles that facilitate the transfer of CFTR from the ER to the Golgi complex. In the Golgi CFTR undergoes its final processing, with the conversion of the small mannose-rich core sugar units on the fourth extracellular loop to a complex highly branched carbohydrate moiety. The addition of a large number of sugar residues to CFTR in the Golgi causes a marked increase in the molecular weight of CFTR to around 180 kDa, yielding 'mature' or 'Band C' CFTR. Bands B and C can be readily visualized by Western blot analysis of cells expressing CFTR.

In the final layers of the Golgi (*Trans* Golgi Network (TGN)), CFTR associates with another coat protein, clathrin, to enter clathrin coated vesicles (CCV) for the journey from the Golgi to the cell surface. Since CFTR is a protein expressed in epithelial cells, not only must CFTR be targeted to the plasma membrane, but it must be targeted specifically to the apical or luminal plasma membrane. The one exception is the sweat duct cell (see below) that expresses CFTR in both the apical and basolateral membranes. Once in the membrane, CFTR can be activated by kinases, allowing conduction of anions. CFTR does not stay in the plasma membrane for a long time, but turns over at a rate of 10% per minute. The carboxyl tail of CFTR contains a tyrosine-based endocytosis signal that drives CFTR into endocytic CCV (Weixel and Bradbury, 2000, 2001). Once internalized, most CFTR is rapidly recycled back to the cell surface (Picciano *et al.*, 2003), where it can again function as an anion channel. CFTR molecules will undergo several rounds of internalization and recycling before being degraded; with a half life of between 12 and 24 h. Why CFTR undergoes constant internalization and recycling back to the cell surface is still not entirely clear. However, it may be a means of fine-tuning the exact amount of functional CFTR at the cell surface. For example, the rates of CFTR endocytosis and CFTR recycling can be modified by agonists. This is particularly evident in the intestine, where CFTR can move from a predominantly intracellular location to a marked cell surface location (Golin-Bisello *et al.*, 2005).

Cystic Fibrosis Is an Ion Transport Disease

Before CFTR was cloned, and before the function of CFTR was established, work by Paul Quinton in the 1980s showed for the first time that there was an underlying defect in ion transport across the epithelia of patients with CF. Quinton went on to show that, in fact, the defect was in an inability of CF patients to move chloride ions (Quinton, 1983). Interestingly, the initial cloning of the *CFTR* gene suggested that it might be a regulator of a chloride channel rather than a chloride channel itself; hence the name 'conductance

regulator.' However, within 2–3 years of *CFTR*'s cloning, definitive proof that CFTR itself could conduct chloride ions was provided (Anderson *et al.*, 1991; Bear *et al.*, 1992). More recently it has been shown that CFTR can also conduct bicarbonate ions (Bridges, 2012; Quinton, 2010), and this process is particularly important in the secretion of a bicarbonate-rich fluid from the exocrine pancreas (Steward *et al.*, 2005). Thus, CFTR is more properly described as an anion channel rather than a chloride channel.

CFTR and Chloride Transport

Since CFTR is an ion channel, the direction of ion movement through the channel depends on the direction of the anion chemical gradient across the cell. As anions also carry an electric charge, their movement is also influenced by the magnitude of electrical gradients across the membrane. The net direction of ion flow is therefore a net effect of chemical and electrical gradients. Using chloride as an example, this means that in some cell types CFTR mediates the movement of chloride out of the cell, whereas in other cells types, CFTR mediates the movement of chloride ions into the cell. In secretory epithelial cells, such as the airways, intestine, and pancreas, CFTR causes chloride ions to leave the cell. Shown in Figure 3 is a model for chloride secretion, and relies on the separation of different transport proteins between apical and basolateral membrane domains. Chloride is brought in from the blood supply across the basolateral membrane through the concerted actions of three separate transport proteins: (1) an Na-K pump, (2) an Na/K/2Cl cotransporter, and (3) a potassium channel. The basolateral Na-K pump generates an inward Na gradient, providing the driving force for uptake of chloride by the Na/K/2Cl cotransporter. The net result is a raised intracellular chloride concentration, sufficiently high that the chloride electrochemical gradient favors passive chloride exit through activated CFTR. In order to maintain efficient apical chloride efflux, basolateral potassium channels must remain active allowing potassium exit and maintaining the electrical driving force for chloride to leave through CFTR. A consequence of luminal chloride is the generation of a more lumen negative electrical potential, which promotes voltage driven Na^+ transport from the blood, through the tight junctions, into the lumen. Sodium chloride in the lumen generates an osmotic gradient allowing water flow from the blood into the lumen. In this manner, CFTR underlies the net movement of salt and water into the lumen. This kind of transport can be seen in sweat glands, salivary glands, intestinal crypts, exocrine pancreatic acini, and airways.

Normally the stimulation of chloride and fluid secretion helps hydrate the airways, and flushes out mucus from specialized glands in the airways so that a thin film of mucin lines the airways where it can trap particles of dust, bacteria, and viruses, allowing the conducting airways to remain sterile. The thin layer of mucin lies on top of a fluid layer that bathes the cilia of cells lining the airways (Figure 4). Under normal conditions, there is a balance between fluid secretion and fluid absorption (under the control of CFTR) that permits just the right height of airway surface liquid (ASL), facilitating the beating of cilia and the movement of mucin-trapped debris out of the airways. In patients with CF, the absence of CFTR

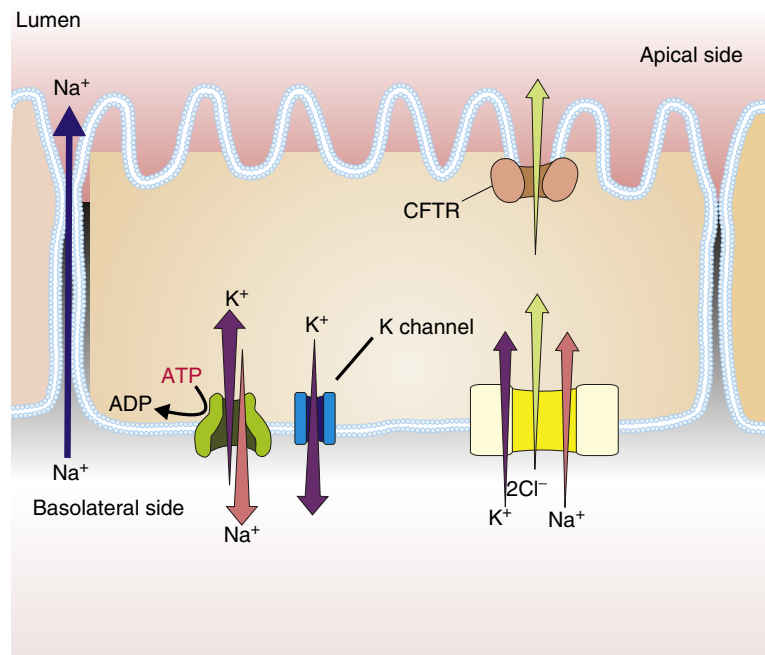


Figure 3 Model for CFTR-dependent chloride secretion.

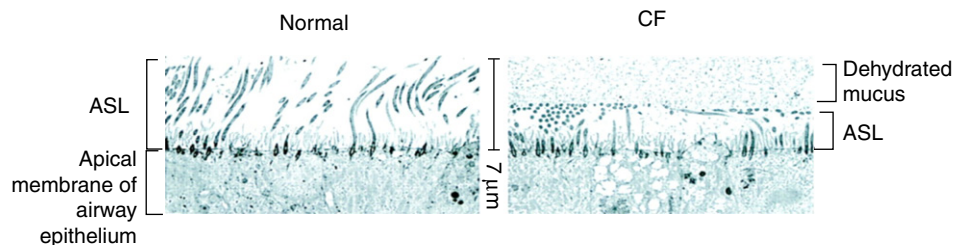


Figure 4 Airway cilia from normal and CF patients. Electron micrographs of airway cells grown in culture from normal and CF patients. The ASL height from normal cells is $\sim 7 \mu$, allowing the cilia to beat and remove debris. In the CF cells, the ASL height is dramatically reduced, and thick dehydrated mucus helps compress the cilia so that they can no longer beat. Modified from Boucher, R.C., 2007. Evidence for airway dehydration as the initiating event in CF airway disease. *Journal of International Medical Research* 261 (1), 5–16. © Wiley.

leads to an imbalance between fluid secretion and fluid absorption, resulting in insufficient amounts of ASL. This causes the collapse of the cilia and an inability of the cilia to expel the mucin-trapped debris from the airways. In addition, the lack of fluid secretion when mucin is released means that the mucins are not properly hydrated and form a sticky viscous aggregate that is hard to dislodge. Since the mucins are not moved out of the airways, trapped bacteria rapidly grow and multiply, leading to chronic lung infections. Characteristic of CF lung disease is the presence of *Pseudomonas aeruginosa* and *Staphylococcus aureus* infections. These infections are typically gained early on in life, and once infected, CF patients are unable to eradicate the infections despite aggressive antibiotic therapy. Chronic lung infections and heavy mucus production cause airway obstruction and eventual lung destruction.

In addition to helping flush out mucins from glands just below the airway surface, fluid secretion from the salivary

glands and exocrine pancreas helps flush out digestive enzymes into the mouth and small intestine respectively. Fluid is also secreted into the intestine to help move and solubilize food through the intestine, so that it can be properly digested. In the case of intestinal crypts, little chloride is secreted under resting (i.e., unstimulated conditions) because apical CFTR chloride channels are closed. Chloride secretion requires activation of CFTR through phosphorylation of CFTR's R domain. CFTR activity can be increased by gut hormones such as vasoactive intestinal peptide, which binds to membrane receptors and activates adenylyl cyclase to raise intracellular cAMP and stimulate cAMP-dependent protein kinase (PKA). Interestingly, this pathway can be activated by bacterial toxins. Cholera toxin permanently activates adenylyl cyclase, leading to chronic activation of CFTR-mediated chloride and fluid secretion. Untreated, the amount of salt and fluid lost by a patient suffering from cholera will result in the death of the patient within a few days.

Folklore, Heat Waves, and a Diagnostic Test

Early European folklore stated, "Woe to that child which when kissed on the forehead tastes salty. He is bewitched and soon must die." Thus from early times there has been a tight association between salty sweat and high morbidity and mortality. Although a salty brow is one of the more benign manifestations of CF, the discovery of the underlying cause of salty sweat by Paul Quinton marked the start of modern CF research. Sweating on a hot day is an important mechanism that the body uses to cool down. Evaporation of water from the surface of the skin requires energy in the form of body heat. So as the sweat evaporates, the body is cooled down. In order for this cooling process to work efficiently, the sweat arriving on the surface of the skin should not be very salty. In the summer of 1949, an unusually intense heat wave hit New York City. At this time, Dorothy Andersen, who was the first to recognize CF as a genetic medical disease, was working as a physician at Columbia Presbyterian Medical School. She noted that a number of her patients (with the newly diagnosed condition of CF) were stricken with heat prostration (overheating of the body and dehydration). Paul di Sant'Agnese, a young physician working with Andersen ascribed the patients' dehydration to excess loss of salt, and identified the sweat glands as the source of that loss.

Understanding how sweat is generated provided the first clues to understanding what goes wrong in patients with CF. The sweat gland is divided into two components: the secretory coil embedded in the dermis of the skin that generates the 'primary' fluid, and a sweat duct that modifies the primary fluid, removing most of the salt from the fluid before it reaches the skin surface. The generation of an isotonic 'salty' fluid by the secretory coil essentially follows the mechanisms described above for salt and fluid secretion; requiring a sodium pump, $\text{Na}^+/\text{K}^+/\text{2Cl}^-$ co-transporter and a K^+ channel on the basolateral membrane, and CFTR in the apical membrane (Figure 5). Sodium and water follow chloride into the lumen of the secretory coil to create a 'primary fluid.' As fluid is secreted into the lumen of the coil, hydrostatic pressure forces the primary fluid up a duct or tubule onto the surface of the skin. As sweat moves up the duct, salt is selectively removed from the fluid. Since the ducts are highly impermeable to water, water cannot follow the salt, and so a hypotonic 'salt-reduced' fluid reaches the surface of the skin where it can evaporate and cool the body. The arrangement of transport proteins in sweat duct cells is different from that of the secretory coil, consistent with the duct cells being an absorptive cell type and the coil cells having a secretory phenotype. Chloride enters the cell through apical CFTR channels and, unique to sweat duct cells, also exits the cell through basolateral CFTR channels. The absorption of chloride ions raises the driving force for the absorption of sodium ions through apical membrane sodium channels, and the exit of sodium from the cell through the sodium pump. The problem in CF is an inability of the sweat duct to absorb chloride ions, due to the absence of CFTR. Since chloride is not absorbed, there is no driving force for sodium absorption, and so both sodium and chloride ions stay in the sweat and appear at the skin surface. The high levels of salt in the sweat lead both to poor sweat evaporation and to excess salt loss from the body.

The high levels of salt in the sweat of CF patients is diagnostic for the disease (Figure 6). Early methods to collect sweat involved 'thermal stress,' either by placing the patient in a heated room (90 °F/32.2 °C) for 1–2 h, or by having the patient wear a plastic body bag and wrapped in blankets for 1–2 h. Although this latter method was widely adopted, it was not without risk, and one fatality was reported using this approach.

The question arises then, that if CF duct cells cannot absorb chloride due to the absence of CFTR, how does the secretory coil generate any primary fluid in the absence of CFTR? The answer lies in the observation that some secretory coil cells also express a calcium-activated chloride channel that can be activated by cholinergic agonists. The cholinergic agonist pilocarpine is used to stimulate sweat production in the 'CF Sweat Test'. Other secretory coil cells express adrenergic agonists that raise intracellular cAMP, leading to the activation of CFTR and the generation of adrenergic-induced sweat. In patients with CF, the adrenergic pathway is ineffective due to the absence of CFTR, but the cholinergic pathway is intact. It is the stimulation of this cholinergic pathway to increase sweat, that forms the basis for the modern 'sweat test.' Despite the ability to perform genetic screens for CFTR mutations, elevated sweat chloride remains the gold standard for the diagnosis of CF disease. A sweat chloride level of ≥ 60 mM is diagnostic of the disease (Figure 6). The level of sweat chloride generally follows the severity of disease, such that patients with milder CF disease generally have lower sweat chloride levels than those who have severe disease.

CFTR and Bicarbonate Transport

The exocrine pancreas produces and releases a large volume of alkaline fluid into the intestine to neutralize acid released from the stomach, and to raise the pH of the intestine lumen to the alkaline pH optimum required for pancreatic enzymes. The pH of the exocrine pancreatic fluid from patients with CF is significantly less alkaline than that of exocrine pancreatic fluid from non-CF individuals. The reduced pH of pancreatic secretions in patients with CF causes the acidity of the stomach contents to be incompletely neutralized once they enter the small intestine. Pancreatic enzymes, particularly lipases, are extremely sensitive to pH, and only breakdown substrates when the pH in the small intestine is sufficiently alkaline. In most CF patients, the pancreatic juice is not alkaline enough and lipases, for example, are inactive in the small intestine. This results in malabsorption of fats, poor weight gain and greasy foul-smelling stool (steatorrhea). In addition to reduced production of alkaline pancreatic juice, lack of CFTR also impedes the production of fluid required to flush out enzymes from the exocrine pancreas. This means that digestive enzymes are not flushed out of the pancreas into the intestine, leading to blockage of the ducts exiting the pancreas and causing autodigestion of the exocrine pancreatic tissue. Although generally not initially affected, the endocrine pancreas can eventually be destroyed, leading to CF related diabetes (CFRD).

CFTR plays both an indirect and a direct role in the generation of an alkaline pancreatic fluid. Similar to the sweat

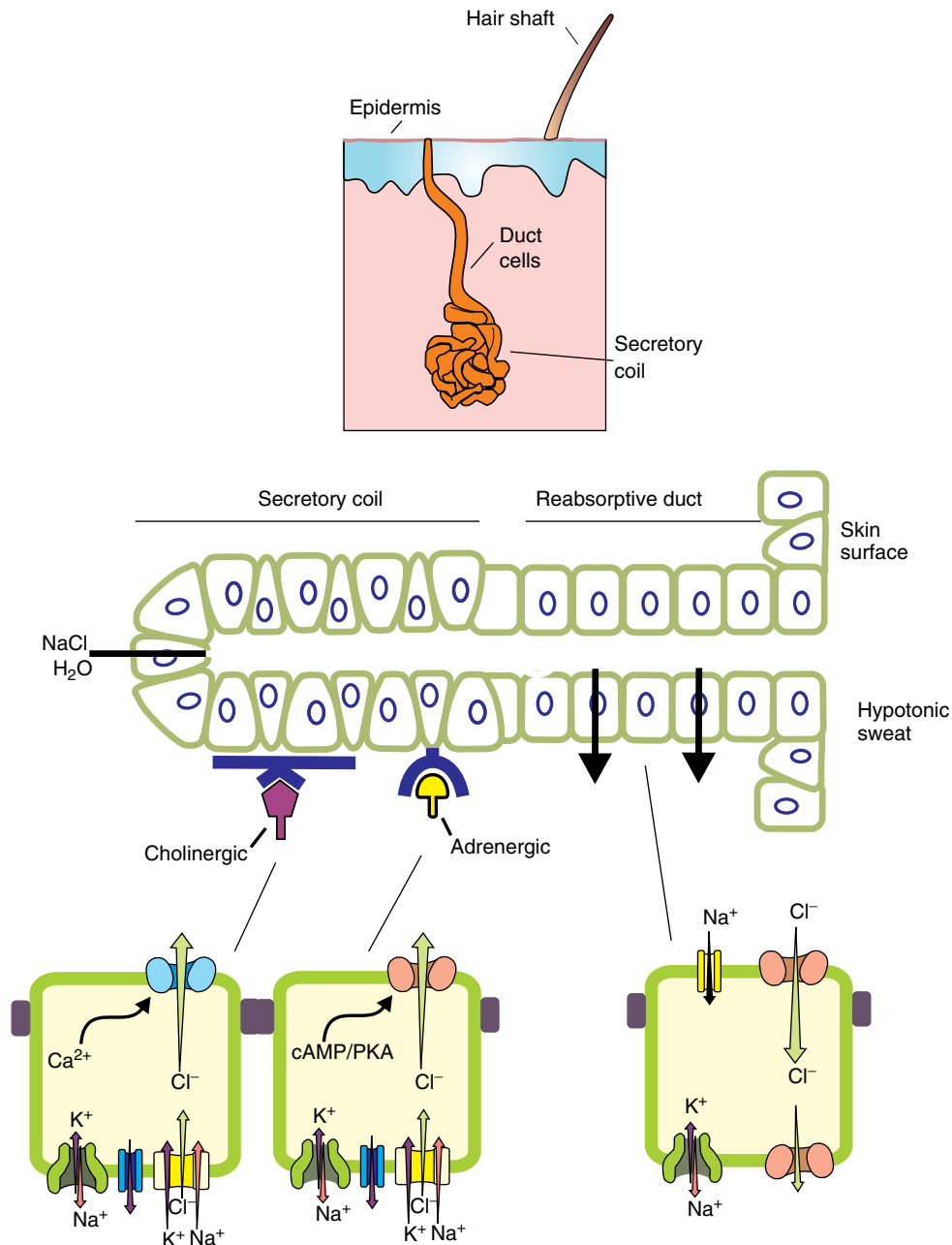


Figure 5 The eccrine sweat gland is composed of a secretory coil that generates an isotonic primary fluid. The secretory coil is permeable to water. The reabsorptive duct absorbs sodium and chloride from the sweat in excess of water, due to the high water impermeability of the duct cells, leading to a hypotonic sweat on the surface of the skin. Different arrangements of transport proteins in the coil and duct cells allow coil cells to perform a secretory function and duct cells to perform an absorptive function.

gland, the distal portions of the exocrine pancreas (the acini, which also secrete the digestive enzymes) secrete a 'primary fluid' with a similar ionic composition to the plasma. The mechanism by which the primary fluid is generated in the pancreas is identical to that in the intestine and the sweat gland secretory coil. As the fluid moves down the pancreatic ducts, it is modified to produce a final bicarbonate (HCO_3^-)-rich ($\sim 140 \text{ mM}$) alkaline fluid. In contrast to the sweat ducts where sodium and chloride ions are both absorbed to yield a hypotonic fluid, the pancreatic ducts gradually replace

the chloride ions with bicarbonate ions. How this is achieved depends on where it occurs within the pancreatic duct (Figure 7). In the proximal region of the duct, CFTR functions as a chloride channel allowing exit of chloride into the lumen of the duct. This chloride re-enters the cell using a chloride-bicarbonate co-transporter, which brings chloride into the cell in exchange for the exit of bicarbonate into the lumen. The more distal regions of the pancreatic duct lack a chloride-bicarbonate exchanger, and in this region bicarbonate exits the cell directly through CFTR. Therefore in the proximal duct

region, CFTR is a chloride channel, but in the distal duct region, CFTR is a bicarbonate channel. It can easily be appreciated that loss of CFTR channel activity will not only lead to a reduced production of primary fluid, but also an inability to properly alkalize the pancreatic fluid. The lack of an alkaline fluid or digestive enzymes was once a major cause of morbidity in patients with CF; many patients dying of malnourishment. The advent of pancreatic enzyme supplements that can be taken with food has alleviated many of the

problems of malnutrition, such that lung disease is now the major cause of morbidity and mortality in CF patients. However, the digestive process is still inefficient and many CF patients require high calorie diets and still have problems associated with impaired lipid digestion.

Genetics

CF is an autosomal recessive disease, so affected individuals must inherit one *CFTR* mutation from each parent in accordance with classical Mendelian genetics. Parents who are heterozygous generally do not display any clinical phenotype. The gene for *CFTR* is located on the middle of the long arm of chromosome 7 (at 7q31.2). The *CFTR* gene contains at least 27 exons (coding regions) and encompasses ~250 kb of DNA. The gene is transcribed into a mature mRNA of 6.5 kb. *CFTR* gene expression is highly regulated, being confined to certain tissues and displaying cell-type specificity. The incidence of CF is 1 in 2000 to 3000 live births among Caucasian populations, giving a carrier frequency of around 1:25. Although *CFTR* gene mutations associated with disease are most prevalent in populations of European descent, mutations have been detected in other ethnic groups. Despite the discovery of approximately 2000 different disease causing mutations in the *CFTR* gene, fewer than 10 account for nearly all the mutations observed in US patients with CF. The most common mutation

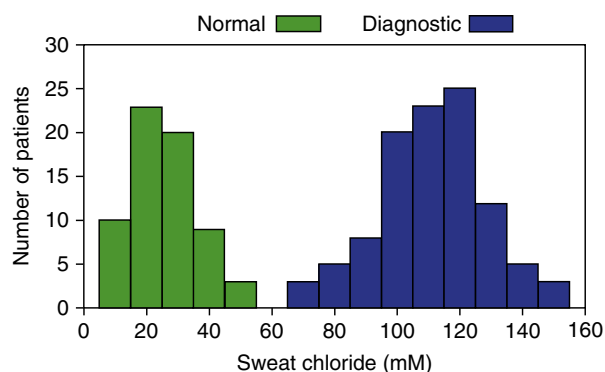


Figure 6 Stimulated chloride concentration (mM) in sweat from non-CF and CF patients. Sweat chloride concentrations above 60 mM are considered diagnostic for CF.

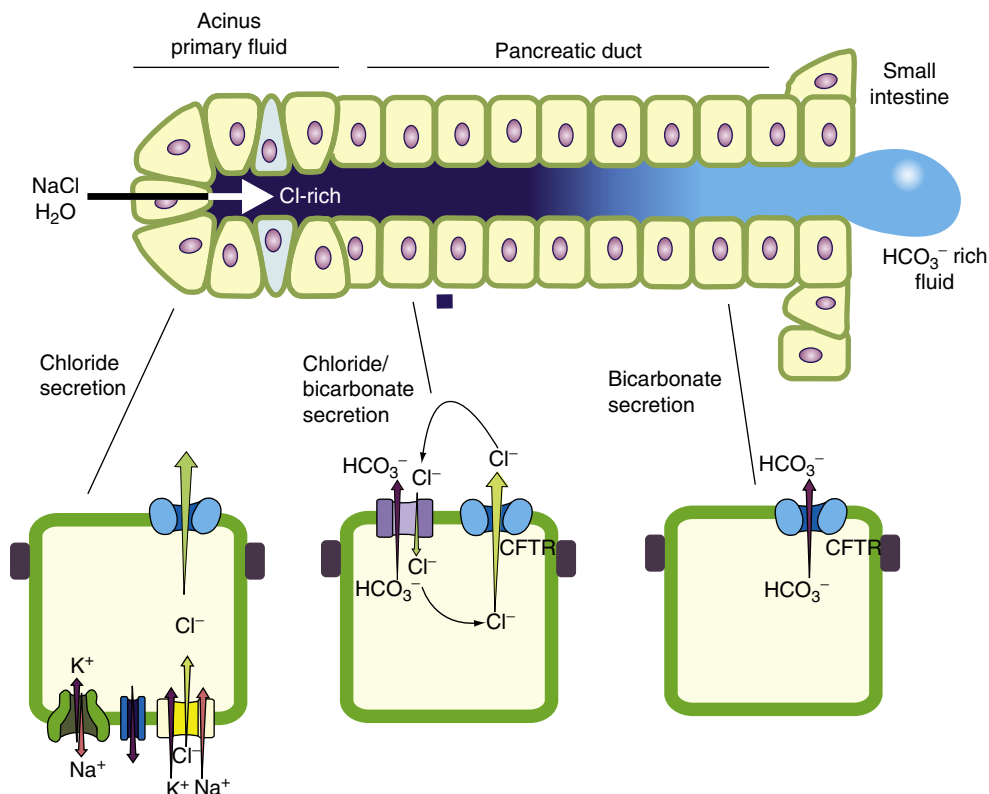


Figure 7 Production of alkaline pancreatic fluid. The pancreas secretes 140 mM NaHCO_3 solution. The mechanism for secretion of HCO_3^- depends on the section of the pancreatic duct. Close to the production of the acinar primary fluid, CFTR works in concert with a $\text{Cl}^-/\text{HCO}_3^-$ exchanger with chloride exiting the cell through CFTR. In distal portions of the duct, HCO_3^- exits directly through CFTR. R117H CFTR shows impaired Cl^- transport through CFTR, with a much smaller effect on HCO_3^- transport through CFTR.

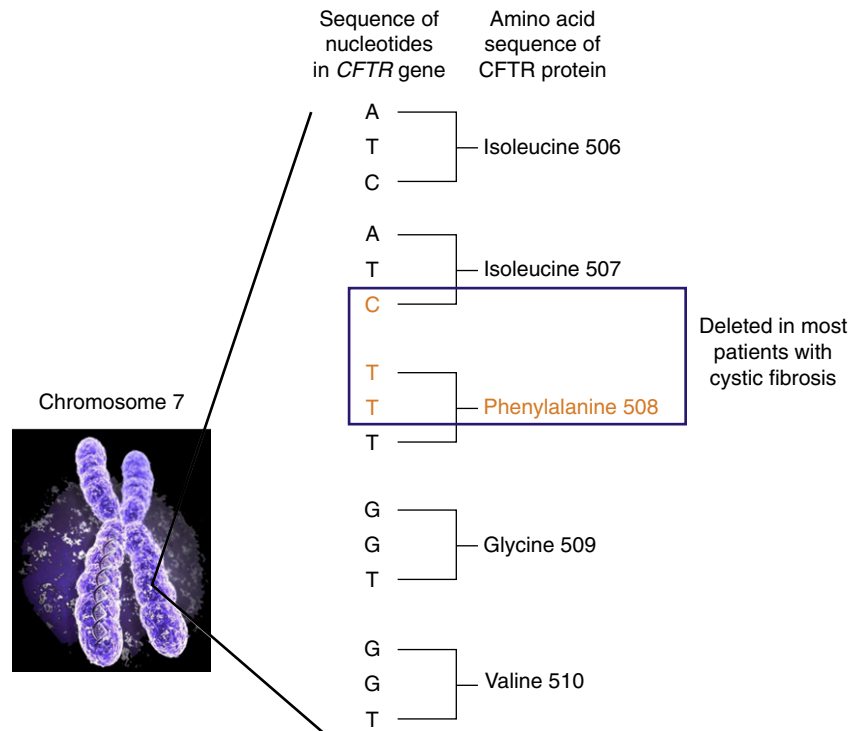


Figure 8 CFTR resides on chromosome 7. The defect that most often leads to cystic fibrosis is the deletion of three nucleotides from the gene. This mutation ($\Delta F508$) results from the loss of one amino acid – phenylalanine at position 508– in the CFTR protein. Although there is a loss of two base pairs from one codon and one base pair from another codon, the remaining triplet ATT is an alternative codon for isoleucine. The mutant protein still has isoleucine at position 507, followed by the GGT sequence for the glycine that would normally follow phenylalanine.

is the loss of a phenylalanine residue at position 508 ($\Delta F508$, also called *F508del*) and accounts for over two-thirds of all disease causing alleles in patients with CF; indeed 90% of CF patients have at least one $\Delta F508$ allele (Figure 8). Other common mutations include G551D, G542X, W1282X, and N1303K. With around 2000 mutations, many of which only appear in one family, it can be appreciated why the sweat test remains the gold standard for diagnosis. Despite the numerous mutations giving rise to the clinical symptoms associated with CF, all mutations can fall into two basic categories: mutations that affect the quantity of CFTR, and mutations that affect the quality of CFTR at the cell surface.

Mutations in CFTR

Given the number of possible mutations giving rise to CF, it is not surprising that different mutations can affect CFTR function in different ways and in different degrees (Figure 8). Even though many mutations are present only in a few patients, they nonetheless have provided insight into the workings of CFTR at the protein and cellular level. It has proven useful to classify mutations according to their functional consequences. CF-causing mutations can reduce CFTR ion transport by reducing the quantity of functional channels at the surface of epithelial cells, or by reducing the quality/function of surface CFTR as an ion channel. The degree to which either the quantity or the quality of CFTR channels is affected is reflected in the amount of anions that can be transported by CFTR, and ultimately in the severity of the disease experienced by patients

with CF. For example, mutations that severely affect protein synthesis and/or channel gating (e.g., $\Delta F508$, W1282X, and G551D) often eliminate or significantly reduce the ability of CFTR to transport anions, and are often associated with severe disease. In contrast, CF-causing mutations that only reduce protein production to a modest degree, or modestly reduce ion conductance (e.g., 3849 + 10 kb C $\rightarrow$ T, R347P) usually retain some ion transport capability. While not sufficient to maintain full/normal ion channel function, these latter mutations nonetheless are associated with much milder CF.

Mutations affecting CFTR quantity

Mutations affecting channel quantity run the spectrum from no CFTR being expressed, to a little CFTR reaching the cell surface (Figure 9). Many mutations that result in premature stop codons (e.g., G542X and R553X) yield essentially no CFTR at the cell surface. Other mutations reduce surface CFTR by producing a protein that is defective in processing or trafficking. For example, the most common mutation, $\Delta F508$, is a severe processing mutation. $\Delta F508$ is unable to achieve an appropriate stable three-dimensional structure in the ER and is targeted by the ER quality control system for degradation; resulting in little or no $\Delta F508$ CFTR at the cell surface. The lack of CFTR at the cell surface results in little or no CFTR-mediated ion transport. Certain mutations result in some CFTR reaching the cell surface. Many of these mutations occur at splice sites, resulting in inefficient splicing out of introns. The CFTR mRNA that is correctly spliced produces fully normal CFTR; however because of the splicing difficulties very

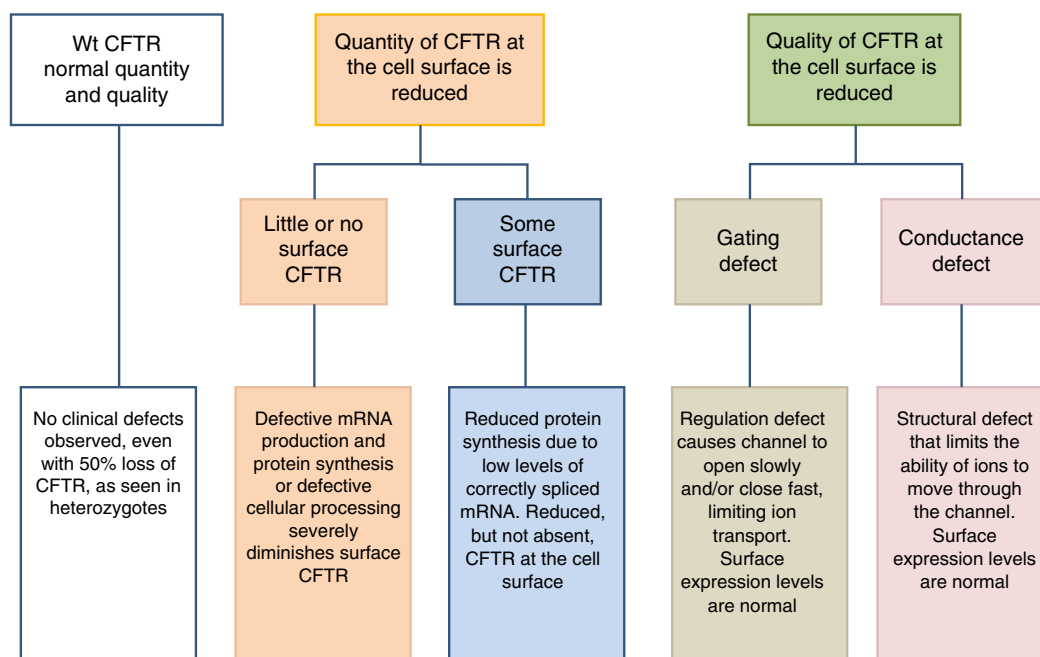


Figure 9 Categories of CF-causing mutations that result in CFTR dysfunction.

little mRNA is made. The most common example is the 3849 + 10 kb C→T mutation. Even though little CFTR reaches the surface in splice mutations, the CFTR that does reach the plasma membrane is fully normal. Although these patients may have lung disease characteristic of CF, many have sufficient levels of functional CFTR in the pancreas that allows them to retain some pancreatic function (pancreatic sufficient).

Mutations affecting CFTR quality

As an ion channel, CFTR needs appropriate gating (the ability of the channel to open and close), and appropriate conductance (the ability of ions to move through the channel). Many CFTR gating mutations result in the proper amount of CFTR being inserted into the plasma membrane, however the channel has difficulty opening or once open closes too quickly (Figure 9). These mutations are said to have a low open probability (P_o). The most common example of this type of mutation is the G551D mutation, a mutation that is present in ~5% of the CF population. Despite the presence of G551D CFTR at the cell surface, its gating defect limits the number of ions that can move through the channel. It should be noted that the $\Delta F508$ mutation also displays gating abnormalities. Thus, the $\Delta F508$ mutation affects both the quantity and quality of CFTR. Far from an academic curiosity, it means that 2 things must be corrected in $\Delta F508$ CFTR in order to provide an effective therapy for patients bearing this mutation. Mutations that alter CFTR conductance also generally result in good levels of CFTR at the cell surface, but the ability of ions to flow through the channel is severely compromised. These mutations usually occur within the pore region of the channel and restrict the number of ions able to pass through the pore. R117H and R347P are common examples of this type of mutation. As discussed, CFTR has the

ability to transport both chloride and bicarbonate ions; important especially in pancreatic function. The R117H mutation affects the ability of CFTR to conduct bicarbonate ions to a lesser degree than chloride ions. R117H CFTR allows little chloride to pass, but bicarbonate is not impeded as much. As a consequence, many CF patients with the R117H mutation retain some pancreatic function even though other characteristic features of CF, such as elevated sweat chloride, are present.

Pharmacologic Approaches to Treat Patients with CF

In 2012, the first drug was approved for the treatment of the basic defect in CF. The drug, Ivacaftor (kalydeco; Vertex Pharmaceuticals) was designed to treat gating mutations, specifically, the G551D mutation. This class of drugs, called 'potentiators' works with G551D to increase the amount of time the channel spends open and conducting ions. Since its introduction, Ivacaftor has proven to be effective for the treatment of patients with the G551D mutation, leading to a marked improvement in lung function (the most important cause of morbidity and mortality in CF patients). Clinical trials with other gating mutations are underway. Despite the success of Ivacaftor, it is currently only approved to treat those patients with the G551D mutation (~5% of the CF population). Another approach is required to help the common $\Delta F508$ mutation. Although Ivacaftor has some benefit in correcting the gating defect associated with $\Delta F508$ CFTR, the lack of $\Delta F508$ at the cell surface renders Ivacaftor ineffective on its own. Other compounds are required to allow $\Delta F508$ to bypass the ER quality control allowing $\Delta F508$ to exit the ER and traffic to the cell surface. These compounds are known as 'correctors.' Although several correctors are currently being investigated, none is yet approved for clinical use.

See also: Molecular Principles, Components, Technology, and Concepts: Proteins: Folding, Misfolding, Disordered Proteins, and Related Diseases; Posttranslational Modifications: Key Players in Health and Disease

References

- Allikmets, R., Singh, N., Sun, H., *et al.*, 1997. A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is mutated in recessive Stargardt macular dystrophy. *Nature Genetics* 15 (3), 236–246.
- Ameen, N., Silvius, M., Bradbury, N.A., 2007. Endocytic trafficking of CFTR in health and disease. *Journal of Cystic Fibrosis* 6 (1), 1–14.
- Anderson, D., 1938. Cystic fibrosis of the pancreas: Clinical and pathological study. *American Journal of Diseases of Children* 56, 344–399.
- Anderson, M.P., Gregory, R.J., Thompson, S., *et al.*, 1991. Demonstration that CFTR is a chloride channel by alteration of its anion selectivity. *Science* 253 (5016), 202–205.
- Bear, C.E., Li, C.H., Kartner, N., *et al.*, 1992. Purification and functional reconstitution of the cystic fibrosis transmembrane conductance regulator (CFTR). *Cell* 68 (4), 809–818.
- ter Beek, J., Guskov, A., Slotboom, D.J., 2014. Structural diversity of ABC transporters. *Journal of General Physiology* 143 (4), 419–435.
- Bridges, R.J., 2012. Mechanisms of bicarbonate secretion: Lessons from the airways. *Cold Spring Harbor Perspectives in Medicine* 2 (8), pii: a0150160.
- Chang, X.B., Mengos, A., Hou, Y.X., *et al.*, 2008. Role of N-linked oligosaccharides in the biosynthetic processing of the cystic fibrosis membrane conductance regulator. *Journal of Cell Science* 121 (Pt 17), 2814–2823.
- Golin-Bisello, F., Bradbury, N., Ameen, N., 2005. StA and cGMP stimulate CFTR translocation to the surface of villus enterocytes in rat jejunum and is regulated by protein kinase G. *American Journal of Physiology – Cell Physiology* 289 (3), C708–C716.
- Higgins, C.F., 1992. ABC transporters: From microorganisms to man. *Annual Review of Cell Biology* 8, 67–113.
- Kerem, B., Rommens, J.M., Buchanan, J.A., *et al.*, 1989. Identification of the cystic fibrosis gene: Genetic analysis. *Science* 245 (4922), 1073–1080.
- McPhail, G.L., Clancy, J.P., 2013. Ivacaftor: The first therapy acting on the primary cause of cystic fibrosis. *Drugs Today (Barcelona, Spain)* 49 (4), 253–260.
- Morita, M., Imanaka, T., 2012. Peroxisomal ABC transporters: Structure, function and role in disease. *Biochimica et Biophysica Acta* 1822 (9), 1387–1396.
- Moskowitz, S.M., Chmiel, J.F., Stern, D.L., *et al.*, 2008. Clinical practice and genetic counseling for cystic fibrosis and CFTR-related disorders. *Genetics in Medicine* 10 (12), 851–868.
- Picciano, J.A., Ameen, N., Grant, B.D., Bradbury, N.A., 2003. Rme-1 regulates the recycling of the cystic fibrosis transmembrane conductance regulator. *American Journal of Physiology – Cell Physiology* 285 (5), C1009–C1018.
- Proesmans, M., Vermeulen, F., De Boeck, K., 2008. What's new in cystic fibrosis? From treating symptoms to correction of the basic defect. *European Journal of Pediatrics* 167 (8), 839–849.
- Quinton, P.M., 1983. Chloride impermeability in cystic fibrosis. *Nature* 301 (5899), 421–422.
- Quinton, P.M., 2010. Role of epithelial HCO₃⁻ transport in mucin secretion: Lessons from cystic fibrosis. *American Journal of Physiology – Cell Physiology* 299 (6), C1222–C1233.
- Riordan, J.R., 2008. CFTR function and prospects for therapy. *Annual Review of Biochemistry* 77, 701–726.
- Riordan, J.R., Rommens, J.M., Kerem, B., *et al.*, 1989. Identification of the cystic fibrosis gene: Cloning and characterization of complementary DNA. *Science* 245 (4922), 1066–1073.
- Rommens, J.M., Iannuzzi, M.C., Kerem, B., *et al.*, 1989. Identification of the cystic fibrosis gene: Chromosome walking and jumping. *Science* 245 (4922), 1059–1065.
- Rust, S., Rosier, M., Funke, H., *et al.*, 1999. Tangier disease is caused by mutations in the gene encoding ATP-binding cassette transporter 1. *Nature Genetics* 22 (4), 352–355.
- Steward, M.C., Ishiguro, H., Case, R.M., 2005. Mechanisms of bicarbonate secretion in the pancreatic duct. *Annual Review of Physiology* 67, 377–409.
- Weixel, K.M., Bradbury, N.A., 2000. The carboxyl terminus of the cystic fibrosis transmembrane conductance regulator binds to AP-2 clathrin adaptors. *Journal of Biological Chemistry* 275 (5), 3655–3660.
- Weixel, K.M., Bradbury, N.A., 2001. Mu 2 binding directs the cystic fibrosis transmembrane conductance regulator to the clathrin-mediated endocytic pathway. *Journal of Biological Chemistry* 276 (49), 46251–46259.

Further Reading

- Boyle, M.P., De Boeck, K., 2013. A new era in the treatment of cystic fibrosis: Correction of the underlying CFTR defect. *Lancet Respiratory Medicine* 1 (2), 158–163.
- Cant, N., Pollock, N., Ford, R.C., 2014. CFTR structure and cystic fibrosis. *International Journal of Biochemistry & Cell Biology* 52C, 15–25.
- Welsh, M.J., Ramsey, B.W., Accurso, F., Cutting, G.R., 2001. Cystic fibrosis. In: Scriver, C.R., Beaudet, A.L., Sly, W.S., Valle, D. (Eds.), *The Metabolic and Molecular Bases of Inherited Disease*, eighth ed. New York: McGraw-Hill, pp. 5121–5188.

Relevant Websites

- www.cff.org
Cystic Fibrosis Foundation.
- www.genet.sickkids.on.ca
Cystic Fibrosis Mutation Database.
- www.cftrscience.com
Vertex Pharmaceuticals Cystic Fibrosis Website.
- <https://www.youtube.com/watchv=AHJ5ZroGhKk>

MOLECULAR PRINCIPLES, COMPONENTS, TECHNOLOGY, AND CONCEPTS: CARBOHYDRATES

Contents

Glycogen and Starch

Proteoglycans

Hyaluronan

Glycogen and Starch

PJ Roach, Indiana University School of Medicine, Indianapolis, IN, USA

SC Zeeman, ETH Zürich, Zürich, Switzerland

© 2016 Elsevier Inc. All rights reserved.

Introduction

Polysaccharides represent a major class of macromolecules in nature, with a wide range of structures and functions. Glucose, in contrast, is the most common of nutrients, whether as an exogenous source of energy and carbon or as a nutritional intermediate within cells or organisms. Polymers of glucose constitute only a small subset of naturally occurring non-structural polysaccharides but are almost universally present within cells, as glycogen, starch, or chemical variants thereof, where they serve as osmotically neutral repositories of metabolically accessible glucose. For millennia, starch has been produced from grains and tubers and used in a variety of human applications, but understanding its nature has taken centuries (Seetharaman and Bertoft, 2012). The first insight into its structure came in 1716 when Leeuwenhoek identified discrete starch granules. By the early 1800s, it was recognized that starch could be converted into sugars, of which glucose was identified as a component by mid-century. In the 1850s, Claude Bernard described a compound in the liver of animals, glycogen, that could be converted to glucose (Young, 1957). However, knowledge of the chemical structures of glycogen and starch in modern terms, as well as their chemical relatedness, was not complete until well into the twentieth century. The reader is directed to recent reviews of the metabolism of starch in plants (Zeeman *et al.*, 2010) and of glycogen in mammals (Roach *et al.*, 2012) and microorganisms (Wilson *et al.*, 2010).

Structures of Glycogen and Starch

Common Chemistry

Glycogen and starch are both polymers of glucose in which the main polymerizing linkages are α -1,4-glycosidic bonds

between glucose residues (Figure 1; for reviews, see Roach *et al.*, 2012; Zeeman *et al.*, 2010). Starch granules contain both amylose and amylopectin, the former an essentially linear polymer and the latter branched by the presence of α -1,6-glycosidic linkages. Glycogen resembles amylopectin in that it is also branched by α -1,6-glycosidic linkages. Glycogen and starch are both polydisperse, meaning that individual molecules are not identical due to the stochastic nature of their syntheses and so a given sample can be characterized only by population averages of properties like molecular weight, chain length, and branching frequency (ratio of α -1,6-linkages to α -1,4-linkages). Both glycogen and starch tend to be large molecules, with masses that can reach beyond 10^8 , the equivalent of $\sim 0.5 \times 10^6$ glucose residues, depending on the source. Because of the polydispersity, unique three-dimensional structures for these polysaccharides are impossible to obtain by the usual methods of structural biology but significant insights into structure have emerged. Early applications of X-ray diffraction established that starch granules produced coherent diffraction patterns consistent with a significant degree of structural organization and indeed crystallinity. The idea later evolved that polyglucose chains in starch and glycogen can form helices that are stabilized by hydrogen bonding. Although glycogen and starch are composed predominantly of glucose, other trace constituents have been reported, the most important of which is covalently attached phosphate.

Glycogen

Glycogen is cytosolic and relatively soluble, whether in single-celled organisms like yeasts and bacteria or in the organs of mammals. However, there is the potential for localization and interaction with other cellular components that can lead to its detection in insoluble fractions. The two most prominent mammalian stores of glycogen are in liver and skeletal muscle

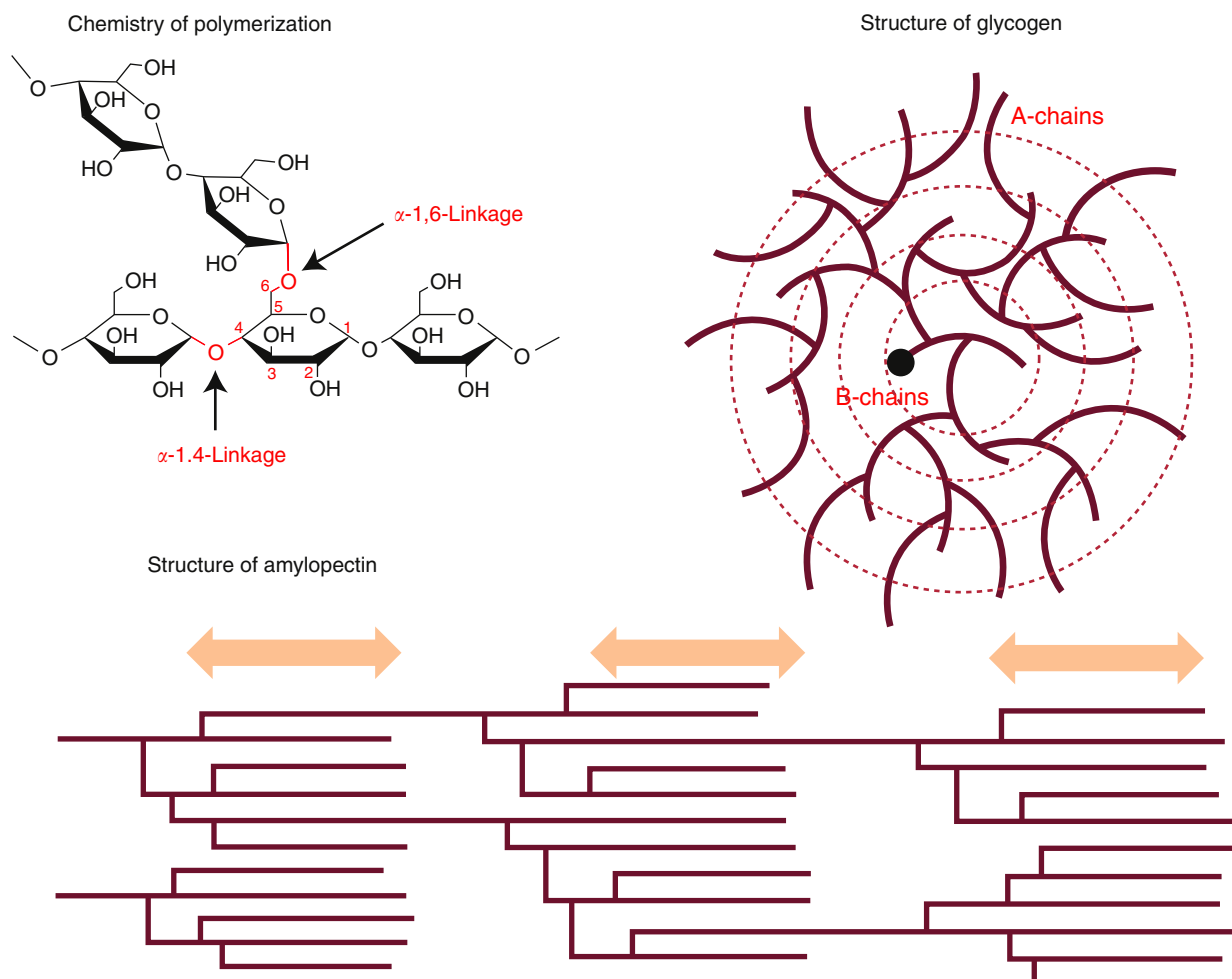


Figure 1 Structures of glycogen and amylopectin. Both polysaccharides are polymers of glucose generated by the same chemical linkages between glucose residues. The differences are in their topologies, which are determined largely by the locations of the branch points, and the average chain lengths, which are longer in amylopectin. In glycogen, the branch points are evenly distributed, average two per inner B-chain, in a tiered structure. The outer A-chains are unbranched. In amylopectin, the branches are clustered, allowing for ordered semicrystalline regions (double headed arrows), where adjacent chains form double helices, alternating with disordered regions containing the branch points. In addition, longer chains span and connect the cluster-containing semicrystalline regions.

although other tissues can synthesize glycogen, notably heart, brain, kidney, and fat.

A current model for animal glycogen structure (**Figure 1**) envisions a fundamentally spherical structure formed of tiers of chains emanating from a single initial chain. The average chain length is typically 13 residues. Each of the inner or B-chains would carry two branch points, so that the number of chains, and hence glucoses, roughly doubles with each additional tier. As a corollary, the number of glucose residues in the outer tier will equal the sum of the glucose residues in all inner tiers. The outermost, A-chains, would be unbranched on the model. Thus, the overall branching frequency is approximately one branch per 12 glucoses. After 12 tiers, such a molecule reaches a theoretical maximum size of $\sim 55\,000$ glucose residues, $M_r \sim 10^7$ and diameter ~ 44 nm. These molecules would correspond to the so-called β -particles observed by electron microscopy in muscle, though in practice few full-size particles are present, rather a distribution with average diameter ~ 25 nm. In liver, larger α -particles are reported, possibly composed of β -particles,

although the chemical basis for the linkage between β -particles is unclear. Glycogen forms assemblages with proteins to form particles that can be isolated biochemically. The particles result from the ability of the associated proteins to bind directly to glycogen and/or to each other. Some of the bound proteins can interact with membranes to explain the observed association of glycogen particles with endoplasmic or sarcoplasmic reticulum in liver and muscle cells. Most of the glycogen-associated proteins are involved in its metabolism, either directly or by regulating other enzymes. The topology of glycogen, whereby many outer chains per molecule are available to bind enzymes for elongation or breakdown, contributes to the efficiency of its metabolism.

Starch

Starch is unique to plants and represents an evolution from a simpler, ancestral ability to synthesize glycogen. In higher

plants, starch is present as insoluble granules in leaves as well as non-photosynthetic organs such as stems, roots, tubers, and seeds. These are the parts of plants that are harvested for sustenance or as animal feed: seeds from cereals (e.g., rice, maize, wheat), tubers (e.g., potato), storage roots (e.g., cassava), and seeds of peas and beans.

Starch granules are much larger and are more complex than glycogen particles. They are formed in plastids (amyloplasts or, in photosynthetic tissue, in chloroplasts) and are composed mainly of amylopectin (>75%), the architecture of which allows for a highly organized, semicrystalline structure (Figure 1; for a comprehensive review of starch structure see Pérez and Bertoft, 2010). Amylopectin has molecular weight 10^7 – 10^9 Da with polyglucose chains ranging from six to more than a hundred residues long. An essential difference between glycogen and amylopectin is the frequency and distribution of the α -1,6-branch points. In glycogen, the chains are relatively short and branching is relatively uniform. In amylopectin, the chains are longer and the branch points arranged so that clusters of adjacent, unbranched chain segments interact to form double helices. These align and pack to form stable lamellae, which underlie the semicrystalline nature of starch (Figure 1). Semicrystalline lamellae alternate with amorphous regions containing branch points. This lamellar organization has a periodicity of 9–10 nm, which can be viewed by transmission electron microscopy and detected by X-ray diffraction. Within the clusters of amylopectin, the chain lengths are between 12 and 15 residues. Chains that span two or three clusters average 35–40 or 70–80 residues, respectively, resulting in a characteristic polymodality of chain lengths. The lesser component of the starch granule, amylose, is smaller (molecular weight 10^5 – 10^6 Da) and minimally branched. Amylose is thought to be present in amorphous regions of the granule in an unorganized form and plants lacking amylose have starch granules that appear normal. Although elements of the structure of the starch granule are quite conserved, note that there is nonetheless significant variation in the detailed structures and physicochemical properties of starches derived from different plant species and organs. As in the case of glycogen, proteins associated with starch granules and several synthetic enzymes have been proposed to act together in protein complexes during granule formation.

Function

From unicellular organisms to plants and mammals, the primary role of glycogen and starch is the storage of glucose during times of nutritional or energetic plenty for retrieval during times of deprivation. The evolution of polysaccharides as vehicles for sequestering glucose is usually attributed to the osmotic advantage of utilizing polymers, despite the energetic cost of their biosynthesis. In mammals, after feeding, ingested glucose is converted into glycogen in liver and muscle, quantitatively the two major glycogen deposits in the body. Besides maintaining stores of the polysaccharide, this conversion also serves to reduce blood glucose levels after a meal, part of the mechanism of blood glucose homeostasis. Impaired glucose handling, and elevated blood glucose (hyperglycemia), is the hallmark of diabetes, the most prevalent and threatening metabolic

disease of the twenty-first century. Liver glycogen serves primarily for the hepatic production of glucose in response to short-term fasting, to prevent blood glucose levels from dropping excessively (hypoglycemia). Some tissues, notably the brain and red blood cells, need glucose for normal function. Muscle glycogen serves as a short-term supply of energy for muscle contraction in some muscle fibers and does not directly contribute to increasing blood glucose levels.

In the leaves of plants, some of the carbon captured by photosynthesis in the day time is retained in the chloroplasts as starch instead of being converted into sucrose for export to other tissues. This so-called transitory starch is degraded at night to provide energy for the leaf and to be converted to sucrose for the rest of the plant. This partitioning of assimilates between starch and sucrose during the day is finely controlled. For example, a larger fraction of the assimilated carbon is stored as starch when the days are short than when the days are long, presumably so that sufficient carbon is available for the long nights that follow short days. Furthermore, starch is degraded gradually so that it lasts the entire length of the night. Impaired starch metabolism in leaves (e.g., in mutants lacking enzymes for precursor production) is associated with reduced growth rates of the plant (Stitt and Zeeman, 2012). In non-photosynthetic tissues (stems, roots, tubers, and seeds), sucrose translocated from the leaves is the precursor for starch biosynthesis. In these tissues, starch serves for longer-term storage, often to support specific functions such as seasonal regrowth, flowering and fruit set, or seedling development. This process could be seen as analogous to the utilization of glycogen stores during sporulation by some bacteria and yeasts.

Glycogen Metabolism

Glycogen Synthesis

As in the biosynthesis of almost all polysaccharides, the immediate sugar donor for glycogen formation is a nucleoside diphosphate sugar, usually UDP (uridine diphosphate)-glucose although prokaryotic species generally utilize ADP (adenosine diphosphate)-glucose (Figures 2 and 3). The origin of the glucose is normally from the environment of the cell although under certain special circumstances the glucose may be produced in the cell from other metabolites, for example, from lactate, pyruvate, alanine, or other amino acids in mammalian liver cells when transitioning from a fasted to a fed state. In eukaryotes, much of glycogen metabolism involves the growth and degradation of existing glycogen molecules. However, *de novo* synthesis of glycogen molecules can occur and sometimes glycogen molecules are completely consumed.

The current model for the *de novo* synthesis of glycogen invokes a specialized initiator protein, a unique type of glucosyltransferase enzyme, called glycogenin, that is able to transfer glucose from UDP-glucose first to a tyrosine residue on itself and then to elongate the attached glucose via α -1,4-glycosidic linkages to generate a short polyglucose chain of 8–12 residues in length. Glycogenin protein interacts directly

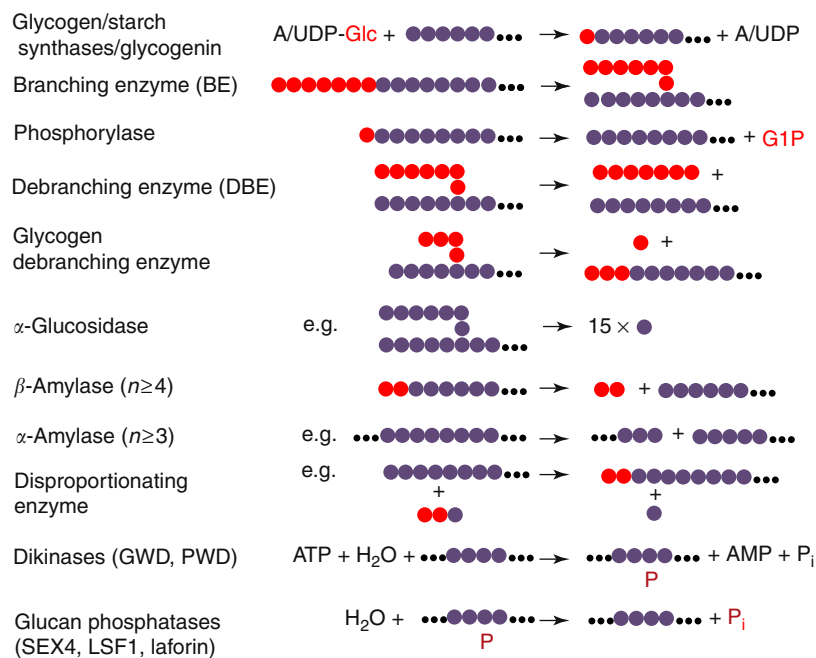


Figure 2 Enzyme activities involved in glycogen and starch metabolism. Synthesis of the α -1,4-glycosidic linkages of glycogen, amylopectin, and amylose is catalyzed by glucosyltransferases that add glucose from a high energy donor, ADP-glucose (starch) or UDP-glucose (glycogen) to the nonreducing ends of growing polyglucose chains. BEs are glucanotransferases that form branch points by breaking α -1,4-glycosidic linkages and reforming α -1,6-glycosidic linkages. Phosphorylases are exoenzymes that catalyze phosphorolytic cleavage of an α -1,4-glycosidic linkage at the nonreducing end of a chain to release glucose-1-P. Debranching enzymes (e.g., isoamylases) directly hydrolyze the α -1,6-glycosidic linkages of amylopectin. Glycogen is also debranched in this way in prokaryotes but in yeast and animals debranching is brought about by a bifunctional enzyme, combining glucanotransferase and amyloglucosidase activities. α -Glucosidases, including the mammalian lysosomal enzyme, can hydrolyze branched glucose polymers to glucose. BAM are exoglycosidases that release maltose units from the nonreducing ends of polyglucose chains of four residues or longer. α -Amylases are endoglycosidases that cleave α -1,4-glycosidic linkages typically in chains longer than three residues. Disproportionating enzymes are glucanotransferases that transfer maltose from the nonreducing end of a polyglucose chain to an acceptor chain. In the degradation of starch, they enable the metabolism of maltotriose to glucose and a longer glucan that is susceptible to further degradation by other enzymes. The glucan, water dikinases GWD and PWD phosphorylate glucose residues in amylopectin using ATP as the phosphoryl donor. Glucan phosphatases (SEX4, LSF1, and laforin) hydrolyze phosphomonoesters attached to glucose residues in poly- or oligosaccharides. In the figure, the colored circles represent glucose residues, red being used to identify residues transferred or cleaved. P, covalent phosphate monoester of a glucose residue.

with glycogen synthase which is responsible for the bulk polymerization required for glycogen formation. Yeasts, but not bacteria, have a similar enzyme activity. Experiments using genetically manipulated mice and yeast, however, suggest that the mechanism of glycogen initiation may be more complex than this simple model (Torija *et al.*, 2005).

Glycogen synthase adds glucose residues to the nonreducing ends of the outer chains of a glycogen molecule. Accounting for the formation of the UDP-glucose precursor from uridine triphosphate (UTP) and the subsequent hydrolysis of the PP_i produced, the cost of elongating glycogen by one glucose residue is two ATP equivalents, making synthesis energetically expensive. Branches are introduced by a branching enzyme (BE) that transfers a chain of ~ 7 glucose residues, breaking an α -1,4-glycosidic linkage and reforming an α -1,6-glycosidic linkage. Although isoforms of the biosynthetic enzymes exist in mammals, these appear to be tissue specific so that a given cell type can be totally competent for the *de novo* biosynthesis of glycogen with the presence of just three enzymes: glycogenin, glycogen synthase, and BE.

Glycogen Degradation

The breakdown of glycogen can occur via two distinct pathways (Figures 2 and 3). The cytosolic pathway requires two enzymes, glycogen phosphorylase and debranching enzyme, and is present from microorganisms to mammals. Phosphorylase catalyzes the phosphorolytic cleavage of the outermost glucose residues of the nonreducing ends to generate glucose-1-P which is either converted into glucose (liver) or fed into the glycolytic pathway for energy production (muscle). Phosphorylase cannot operate when the glucose chain is four residues or fewer from a branch point. Degradation through a branch point requires the debranching enzyme whose action eliminates the α -1,6-glycosidic linkage and generates a free glucose molecule. In a given cell, glycogen degradation requires just two enzymes. The second degradative pathway, which is best characterized in mammals, involves the transfer of glycogen molecules to lysosomes by a vesicular trafficking mechanism that likely resembles autophagy. Within the lysosome, glycogen is hydrolyzed completely by an α -glucosidase

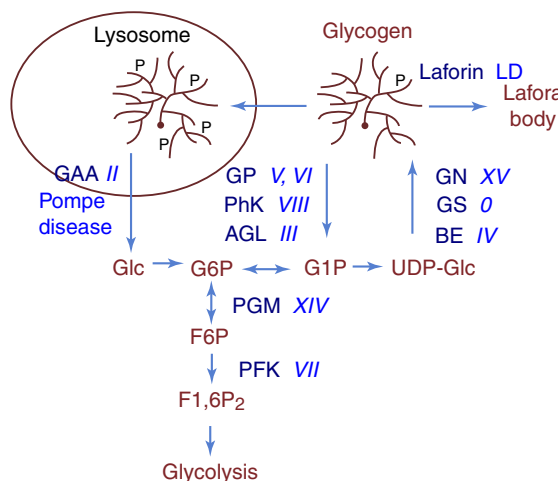


Figure 3 Glycogen metabolism and glycogen storage diseases. The figure depicts glycogen metabolism and associated intermediary metabolism. BE, branching enzyme; DBE, debranching enzyme; GN, glycogenin; GP, glycogen phosphorylase; GS, glycogen synthase; G6Pase, glucose-6-phosphatase; LD, Lafora disease; PFK, phosphofructokinase; PGM, phosphoglucosmutase; PhK, phosphorylase kinase. Roman numerals denote the glycogen storage disease associated with mutations in the gene encoding the corresponding enzyme.

enzyme (acid maltase, GAA) to glucose that can be utilized by the cell as a source of carbon and/or energy. Yeast can similarly transport glycogen to its vacuole, the approximate equivalent of the lysosome, and degrade it as needed.

Regulation

Study of the control of the cytosolic glycogen-metabolizing enzymes in mammals has played an important historical role in the development of several key biochemical concepts of cellular regulation, namely control by allosteric ligand binding, reversible covalent phosphorylation of proteins, and hormone-mediated second messenger signaling systems. Glycogen synthase and phosphorylase are the key loci for regulation. Phosphorylation of these enzymes is promoted by the cyclic AMP (adenosine monophosphate) pathway, in the liver as a response to elevated glucagon indicative of the fasted state and in muscle as a response to elevated epinephrine indicative of increased muscular activity as in the 'fight or flight' response. Under these conditions, glycogen synthase is inactivated and phosphorylase is activated, promoting glycogen breakdown to supply either glucose (in liver) or ATP (in muscle), whilst suppressing wasteful 'futile cycling' by the coordinated control of both activities. In the fed state, insulin promotes dephosphorylation of these enzymes, thereby promoting glycogen synthesis in a similarly coordinated fashion. Glucose-6-P allosterically activates glycogen synthase and inhibits phosphorylase allowing for controls that reflect the intracellular energetic status. AMP, by activating glycogen phosphorylase, provides another intracellular regulatory input.

Glycogen Phosphorylation

Glycogen contains a small proportion of covalently linked phosphate which is present as monoesters of C2, C3, and C6

carbons of glucose residues. In contrast to starch phosphorylation, the function of the phosphate in glycogen is unclear, as is its origin. No glycogen phosphorylating enzymes, comparable to the glucan dikinases discovered in plants (see below), have been identified in mammals. One proposal that could account for C2 and possibly C3 phosphomonoesters is that phosphate is derived from a rare side reaction catalyzed by glycogen synthase in which the β -phosphate of the substrate UDP-glucose is transferred as a glucose phosphate unit rather than the normal glucose residue. However, this mechanism cannot explain the presence of C6 phosphate. Nevertheless, it is well established that glycogen phosphate is removed by a glycogen phosphatase called laforin which contains both a phosphatase domain and a glycogen-binding domain. Genetic defects in laforin lead to a serious neurological disorder called Lafora disease, strongly suggesting that controlling the level of glycogen phosphorylation is of biological importance.

Starch Metabolism

Starch Synthesis

The immediate precursor for starch synthesis in higher plants is ADP-glucose (Figure 4), although in red algae and glaucophytes some starch synthases utilize UDP-glucose. Although some aspects of amylopectin synthesis resemble those of glycogen, namely the formation of the α -1,4-glycosidic and α -1,6-glycosidic linkages, the relative topological complexity of starch is coupled to a more complicated enzymatic machinery. In higher plants, there are five gene classes of starch synthase (GBSS, SSI-IV) with distinguishable biochemical properties and two types of BE. The relative levels of these isoenzymes vary among starch-producing tissues resulting in subtle, yet significant differences in the topology of the synthesized polysaccharide.

A major difference with glycogen formation is the requirement during amylopectin synthesis of debranching enzymes that help to ensure the correct distribution and positioning of the branch points (Figures 2 and 4). Two families of debranching enzyme exist in plants, isoamylase (composed of three isoforms ISA1-3) and limit-dextrinase (LDA). Genetic impairment of one debranching enzyme activity (which contains either ISA1 subunits or ISA1 and ISA2 subunits) can lead to plants that produce a more soluble, glycogen-like polymer (termed phytoglycogen) consistent with proper debranching activity being linked to the characteristic branched structure of amylopectin. Evidence is also accumulating for the existence of complexes of amylopectin-synthesizing enzymes, with starch synthases and BEs being detected in different combinations which might orchestrate their actions to generate specific glucan structures. Although plants contain glycogenin-like proteins and other reversibly glycosylated proteins, there is so far no compelling evidence for their role in the initiation of starch synthesis comparable to glycogen initiation. Rather, some studies have suggested that specific starch synthase isoforms may themselves contribute to the initiation process. Synthesis of the amylose component of starch is mediated by granule-bound starch synthase (GBSS) which, unlike the other isoforms, is

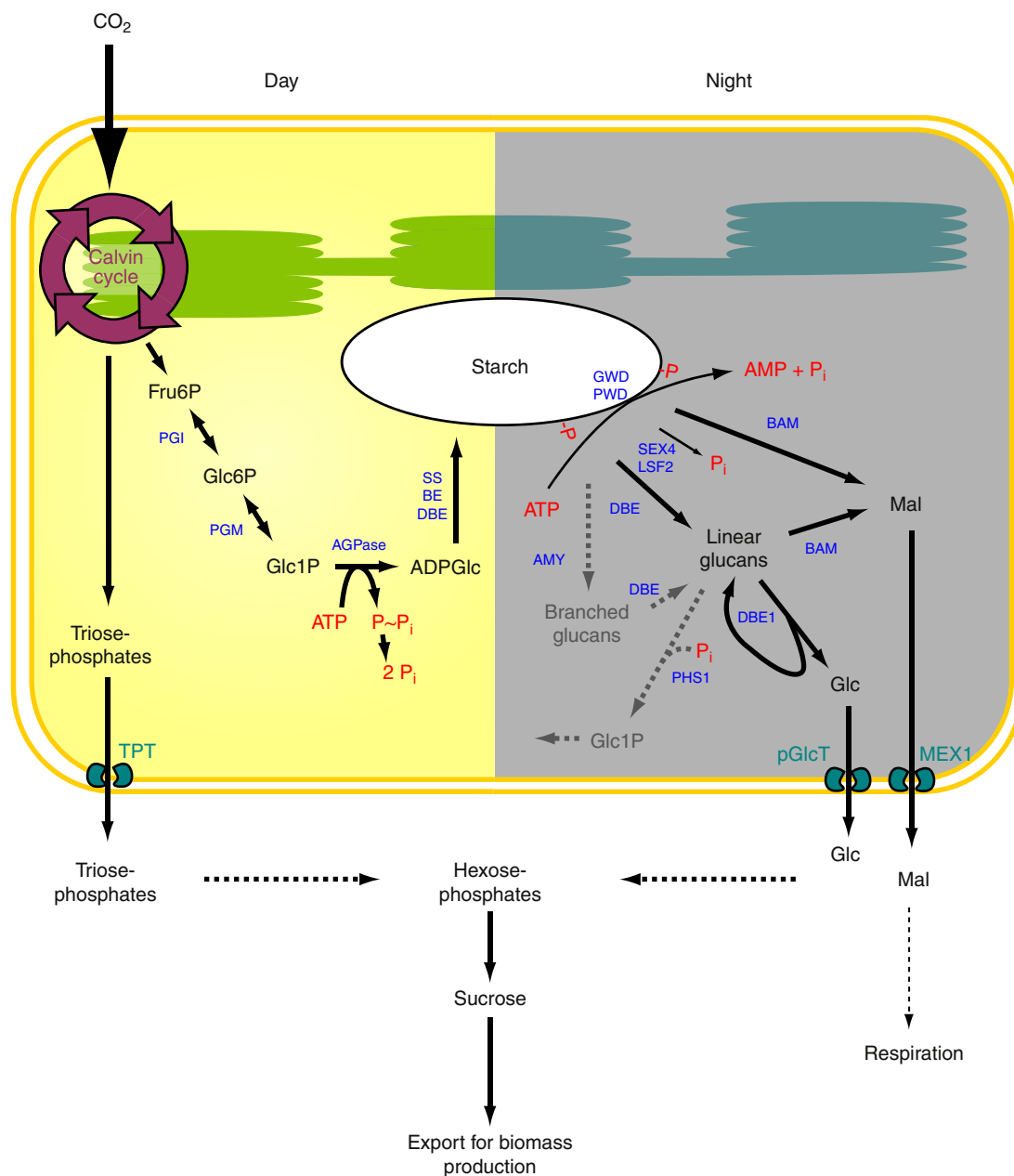


Figure 4 Starch synthesis and degradation in *Arabidopsis* leaves. The substrate for starch synthesis, ADP-glucose, is derived from photosynthesis and the Calvin cycle during the day. The insoluble starch is produced by the coordinated activities of multiple starch synthases, BEs, and debranching enzymes. At night, the starch granule surface is disrupted/solubilized by glucan phosphorylation, followed by simultaneous dephosphorylation and glucan hydrolysis. The main exported products are the neutral sugars, maltose and glucose. Some Glc1P may be generated through phosphorolysis and used for plastidial metabolism or exported. In the cytosol, glucose and maltose are metabolized to support respiration and continued sucrose biosynthesis. AGPase, ADP-glucose pyrophosphorylase; AMY, α -amylase; BAM, β -amylase; BE, branching enzyme; DBE, debranching enzyme; DPE, disproportionating enzyme; GWD, glucan water dikinase; MEX1, maltose transporter; PGI, phosphoglucose isomerase; pGlcT, glucose transporter; PGM, phosphoglucomutase; PHS, α -glucan phosphorylase; PWD, phosphoglucan water dikinase; SEX4 and LSF2, phosphoglucan phosphatases; SS, starch synthase; TPT, triose-phosphate/phosphate translocator. Modified from Stitt, M., Zeeman, S.C., 2012. Starch turnover: Pathways, regulation and role in growth. *Current Opinion in Plant Biology* 15, 282–292.

exclusively localized to the starch granule and catalyzes a processive addition of glucose residues from the ADP-glucose donor. It was recently discovered that a mechanism exists to 'deliver' GBSS to the starch granule surface, via the action of a nonenzymatic, starch-binding protein.

Starch Degradation

Multiple pathways for starch degradation exist and these can differ markedly between plant organs. The processes in leaves, where starch is broken down at night, and in cereal endosperms,

where it is degraded after germination, are the best understood. In the chloroplasts of leaves, starch is converted primarily to maltose and glucose, which can be exported to the cytosol for further metabolism, including oxidative respiration and the synthesis of sucrose for export to other tissues (Figure 4). It is now believed that the first stage in starch breakdown is the covalent phosphorylation of amylopectin by a pair of dikinase enzymes which transfer the β -phosphate of ATP to form a C3 or C6 phosphomonoester of a glucose residue concomitant with the production of AMP and inorganic phosphate. These enzymes, glucan water dikinase (GWD) and phosphoglucan water dikinase (PWD), act in sequence to modify the C6 and C3 positions, respectively of different glucosyl residues of amylopectin. The idea is that the introduction of phosphate groups disrupts the hydrogen bonding and packing of the polyglucose helices in the semicrystalline regions at the granule surface, helping to solubilize the glucans. The exposed chains are then susceptible to β -amylases (BAM) which hydrolyze α -1,4-glycosidic linkages. However, BAM cannot degrade α -1,6-linkages for which debranching enzymes are required (ISA3 and LDA in *Arabidopsis*) to generate short, linear oligosaccharides. Nor can BAM degrade past phosphate residues, and these are removed by a pair of phosphoglucan phosphatases, encoded by the SEX4 and LSF2 (for Starch EXcess 4 and Like Starch excess Four 2, respectively). Study of *sex4* and *lsf2* mutants indicates that dephosphorylation of glucans is necessary for complete degradation of starch since mutation of both results in a severe starch excess (SEX) phenotype. Interestingly, these phosphatases have similarity to the laforin phosphatase that acts on glycogen in animals. Other enzymes, including the disproportionating enzymes, plastidial α -amylase and plastidial α -glucan phosphorylase are also involved in the complete degradation of starch.

The breakdown of starch in cereal endosperm occurs in a quite different environment. At seed maturity, the starchy endosperm – the dried-out tissue consisting of starch granules, storage proteins, and cell walls – is dead. After germination, the starch is degraded by enzymes either preformed in the endosperm or secreted from living parts of the seed to generate primarily glucose. This released sugar fuels the growth of the embryo. The four enzymes, α -amylase, β -amylase, debranching enzymes (specifically LDA), and α -glucosidase, are all active in the process, although the relative importance of each is not known precisely. Nevertheless, it is generally accepted that α -amylase, which is present as several isoforms, plays a central role.

Genetic Modification of Starch and Glycogen Metabolism

Genetic defects in glycogen metabolism have generally presented as rare familial diseases in humans, occasionally in domesticated animals including horses, pigs, and dogs. Many of these glycogen storage diseases (GSDs) or glycogenoses have been known for many years and have been well characterized biochemically and clinically (Chen and Burchell, 1995); however, new examples continue to be discovered (Oldfors and DiMauro, 2013). GSDs can basically be divided into two categories: mutations that directly affect the enzymes of glycogen metabolism (glycogenin, glycogen synthase, BE,

debranching enzyme, phosphorylase, lysosomal α -glucosidase, laforin) and mutations in enzymes that affect glycogen accumulation indirectly, such as via excessive accumulation of glycolytic intermediates (phosphofructokinase, phosphoglucomutase, aldolase, glucose-6-phosphatase, etc.). The number of genes implicated so far is in the twenties. Besides their clinical relevance, GSDs have proven especially important in validating the postulated functions of the glycogen-metabolizing enzymes: for example, defective BE (Andersen disease or adult polyglucosan disease) results in glycogen with a less branched structure, resembling amylopectin in some of its physicochemical properties. Defective lysosomal α -glucosidase (Pompe disease) can be fatal due to cardiomyopathy, underscoring the importance of the lysosomal pathway of glycogen degradation. Lafora disease, which is characterized by accumulation of abnormal glycogen with excessive phosphate content in deposits called Lafora bodies, is linked to loss of function of the laforin phosphatase or of malin, an E3 ubiquitin ligase of unclear function. The abnormal glycogen has devastating effects on neurons leading to neuronal cell death, a progressive myoclonic epilepsy, neurological decline, and inevitable death usually in the twenties. Though much remains to be understood about glycogen phosphorylation, its role in Lafora disease emphasizes the importance of avoiding excessive glycogen phosphorylation. While deliberate genetic manipulation of glycogen metabolism has been widely exploited as an experimental approach in mice, only now are gene therapy approaches beginning to be applied to patients though these therapies could become important in the future. Enzyme replacement therapy, in which active enzyme is administered systemically with the goal that some proportion can replace the defective endogenous protein, has already been used with some success, especially in Pompe disease.

Besides its application in scientific research, genetic manipulation of starch metabolism in plants has a whole different scope compared with glycogen, given that starch is of such major global importance, in the human diet, as animal feed, and as an industrial material, including its use as a precursor of bioethanol (see Zeeman *et al.*, 2010 and references therein). Starch from different natural sources has a wide range of polymer compositions and structures leading to significant differences in physicochemical properties, all of which could be modified genetically. However, the structural and enzymological complexity of starch poses many challenges. As a means of simply increasing starch accumulation, transgenic expression of an unregulated ADP-glucose pyrophosphorylase has had mixed success. Promoting plastid ATP production, via transgenic plastid adenylate translocase or down-regulation of adenylate kinase, does promote starch accumulation. Suppressing starch breakdown, in particular by down-regulating GWD, has successfully elevated starch yield in several plants and in potatoes has the benefit of inhibiting the generation of reducing sugars, glucose and fructose upon storage (cold sweetening) which is detrimental during frying. Enhancing the capacity for starch degradation has been applied to crops intended for ethanol production. By expressing a thermostable α -amylase in corn, and targeting it to cellular compartments not involved in starch synthesis, the enzyme does not interfere with starch production. However, upon heating corn kernels in water, the thermostable α -amylase initiates the conversion of starch to fermentable

sugars. Another broad approach is to attempt manipulation of the relative proportions of amylose and amylopectin. For example, transgenic potato and cassava plants lacking amylose have been produced. While such research – in both the academic and private sectors – has revealed the potential benefits of genetic modification in starch crops, the commercialization of such ‘GM-plants’ has stalled in many parts of the world, including Europe. This is because of the very stringent regulatory measures that need to be met before transgenic plants are allowed to enter the agricultural environment and the food chain. In the future, alternative approaches should bypass such regulatory hurdles providing legislatively acceptable routes toward the same goal. Advances in breeding techniques should allow naturally occurring alleles to be bred into commercial crop cultivars more easily. Cisgenic approaches, where endogenous gene sequences are used for genetic improvement may be seen as an advance over transgenics. Furthermore, the advent of targeted genome editing may prove to be of major importance for the commercialization of genetically designed plants (Voytas, 2013).

Elucidating and understanding the structures of normal, defective, or genetically engineered starch and glycogen clearly have everything to do with the biological function and commercial utility of the polysaccharides. Whilst significant progress has been made, better understanding and improved analytical methods to describe their structures would greatly facilitate progress, especially in the design of new starches for industrial purposes. In the genetic manipulation of starch structure, not only must we confront the imposing complexity of the

metabolic processes but also societal issues regarding the perceived value and acceptability of genetically altered natural products.

References

- Chen, Y.-T., Burchell, A., 1995. Glycogen storage diseases. In: Scriver, C.R., Beaudet, A.L., Sly, W.S., Valle, D. (Eds.), *The Metabolic and Molecular Bases of Inherited Disease*, seventh ed. New York: McGraw-Hill.
- Oldfors, A., DiMauro, S., 2013. New insights in the field of muscle glycogenoses. *Current Opinion in Neurology* 26, 544–553.
- Pérez, S., Bertoft, E., 2010. The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review. *Starch-Stärke* 62, 389–420.
- Roach, P.J., Depaoli-Roach, A.A., Hurley, T.D., Tagliabracci, V.S., 2012. Glycogen and its metabolism: Some new developments and old themes. *Biochemical Journal* 441, 763–787.
- Seetharaman, K., Bertoft, E., 2012. Perspectives on the history of research on starch. Part 1: On the linkages in starch. *Starch-Stärke* 64, 677–682.
- Stitt, M., Zeeman, S.C., 2012. Starch turnover: Pathways, regulation and role in growth. *Current Opinion in Plant Biology* 15, 282–292.
- Torija, M.J., Novo, M., Lemassu, A., *et al.*, 2005. Glycogen synthesis in the absence of glycogenin in the yeast *Saccharomyces cerevisiae*. *FEBS Letters* 579, 3999–4004.
- Voytas, D.F., 2013. Plant genome engineering with sequence-specific nucleases. *Annual Review of Plant Biology* 64, 327–350.
- Wilson, W.A., Roach, P.J., Montero, M., *et al.*, 2010. Regulation of glycogen metabolism in yeast and bacteria. *FEMS Microbiology Reviews* 34, 952–985.
- Young, F.G., 1957. Claude Bernard and the discovery of glycogen; a century of retrospect. *British Medical Journal* 1, 1431–1437.
- Zeeman, S.C., Kossmann, J., Smith, A.M., 2010. Starch: Its metabolism, evolution, and biotechnological modification in plants. *Annual Review of Plant Biology* 61, 209–234.

Proteoglycans

PW Park, K Hayashida, RS Aquino, and A Jinno, Harvard Medical School, Boston, MA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Sulfated glycosaminoglycans (GAGs) are linear polysaccharides expressed ubiquitously in intracellular compartments, on the cell surface, and in the extracellular *milieu*. GAGs are ancient molecules that have withstood limited structural variation over evolution. For example, heparan sulfate (HS) is present in *Cnidaria* and all metazoans analyzed to date, with the exception of *Porifera*. *In vivo*, the sulfated GAGs, HS, heparin (HP), chondroitin sulfate (CS), dermatan sulfate (DS), and keratan sulfate (KS), are found covalently conjugated to specific core proteins as proteoglycans. Most cells express multiple proteoglycans. Studies of proteoglycans date back to the late nineteenth century when aggrecan (chondromucoid) was identified in cartilage. This was followed by the unexpected identification of HP (cephalin), a highly sulfated version of HS, as a potent anticoagulant in liver extracts (thus, the name heparin) in 1916 by a medical student who was trying to isolate a procoagulant molecule. HP was later determined to function as an anticoagulant by binding to antithrombin III and inducing a conformational change that facilitates its inhibitory action against procoagulant enzymes, such as thrombin and factor Xa. HS was initially recognized as a contaminant in the HP preparation, but was later distinguished from HP in the late 1940s by differences in the extent of sulfation and greater structural variability. For a long time, the biological functions of proteoglycans were largely speculative and, in fact, most proteoglycans were thought to be mainly specific to cartilage, functioning as lubricating, load-bearing cushions for variable, compressive forces in joints. We now know that proteoglycans also function as key modulators of many molecular and cellular interactions, which strongly influence diverse pathophysiological processes, such as development, immunity, infection, and cancer.

The type of sulfated GAGs attached to core proteins defines proteoglycans as HS proteoglycans (HSPGs), CS proteoglycans (CSPGs), or KS proteoglycans (KSPGs). In cases where the core protein carries several types of GAGs, the functionally dominant GAG chain defines the proteoglycan. For example, syndecan-1 contains both HS and CS chains, but is considered an HSPG based on this definition. GAGs are polysaccharides of repeating disaccharide units, which consist of uronic acid (UA) or galactose (Gal) and hexosamines. The basic disaccharide units of HS and HP are glucuronic acid (GlcA) or iduronic acid (IdoA) and *N*-acetylglucosamine (GlcNAc), whereas those of CS are GlcA and *N*-acetylgalactosamine (GalNAc), DS are GlcA or IdoA and GalNAc, and those of KS are Gal and GlcNAc. HS is differentially sulfated throughout the chain, whereas CS usually contains 1 or 2 sulfate groups per disaccharide. Sulfated GAGs are highly hydrophilic due to their negative charge, and are capable of holding large amounts of water molecules and have a large hydrodynamic volume. These properties allow several proteoglycans (e.g., aggrecan) to function as cushions for variable, compressive loads in joints. In contrast,

owing to the highly heterogeneous structure of HS chains, HSPGs can interact specifically with large number of proteins and regulate their biological activities in various cellular compartments.

Structure and Synthesis of Proteoglycans

Core Protein Structure

So far, more than 40 genetically distinct proteoglycan core proteins that bear one or several sulfated GAGs have been identified (Bernfield *et al.*, 1999; Filmus and Capurro, 2014; Scully *et al.*, 2012; Iozzo, 2005; Wight *et al.*, 2014; Iozzo and Schaefer, 2010). A partial list of proteoglycans and their sulfated GAGs is shown in Table 1. The core proteins have distinct structural designs with specific GAG attachment sites. The number of GAG chains on a core protein may vary from 1 to 100, and their length may vary from a few disaccharide units to hundreds. Furthermore, proteoglycans are expressed on different cells and cellular compartments, at different times and levels. In general, the core proteins direct the specific expression pattern of proteoglycans, although there are few exceptions (e.g., glypicans). In intracellular compartments, serglycin is mainly expressed in secretory vesicles of hematopoietic cells, but also found in vesicles of endothelial, endocrine, and tumor cells (Scully *et al.*, 2012). Serglycin contains a large number of Ser-Gly repeats (≥ 16), to which HP, CS, or DS chains are attached. Serglycin of connective tissue mast cells mostly contains HP, whereas that of mucosal tissue mast cell bears CS.

Cell surface proteoglycans include the syndecan and glypican families with four and six members in mammals, NG2, neuropilin, betaglycan, CD44, invariant chain, and thrombomodulin (Bernfield *et al.*, 1999; Couchman, 2010; Filmus and Capurro, 2014; Park *et al.*, 2000b). The latter four can exist with or without GAG chains and are therefore considered as part-time proteoglycans. Syndecans, NG2, betaglycan, CD44, invariant chain, and thrombomodulin are type I transmembrane core proteins whereas glypicans are linked to the cell surface by a glycosylphosphatidylinositol (GPI) anchor. Syndecan core proteins contain an extracellular domain extended in conformation due to a high Pro content. The sequence identity in glypicans could be as low as 25%, but their 3D structure is highly similar because the 14 highly conserved Cys residues form intramolecular disulfide bonds, resulting in a conserved globular structure. Two to four HS chains are attached distal to the plasma membrane on all syndecans and CS chains are also attached proximal to the plasma membrane on syndecan-1 and -3 (Kokenyesi and Bernfield, 1994), suggesting that HS interacts with extracellular proteins, whereas CS may interact with membrane proteins. In contrast, 2–5 HS chains are attached proximal to the plasma membrane in glypicans, suggesting that GAGs on glypicans primarily interact

Table 1 Partial list of proteoglycans and their sulfated GAG chains

<i>Proteoglycan</i>	<i>GAGs</i>	<i>Core protein (kDa)</i>	<i>Cellular compartment</i>	<i>Adult cell/tissue expression</i>
HSPGs				
Agrin	HS	200	Extracellular	Neuromuscular junction
Collagen XVIII	HS	147	Extracellular	Basement membrane
Glypican-1	HS	56	Cell surface	Broadly expressed
Glypican-2	HS	59	Cell surface	Oncofetal, neurons during development
Glypican-3	HS	59	Cell surface	Epithelial cells, hepatocellular carcinoma marker
Glypican-4	HS	58	Cell surface	Broadly expressed
Glypican-5	HS, CS	59	Cell surface	Brain
Glypican-6	HS	58	Cell surface	Broadly expressed
Perlecan	HS	400	Extracellular	Basement membrane, cartilage
Serglycin	HP, CS, DS	10–19	Intracellular	Hematopoietic, endocrine cells
Syndecan-1	HS, CS	33	Cell surface	Epithelial, plasma cells
Syndecan-2	HS	23	Cell surface	Endothelial, mesenchymal cells
Syndecan-3	HS, CS	43	Cell surface	Neural crest-derived cells
Syndecan-4	HS	22	Cell surface	Broadly expressed
Testican-1	HS, CS	48	Extracellular	Brain
CSPGs				
Aggrecan	CS, KS II	210–250	Extracellular	Cartilage
Bamacan	CS	138	Extracellular	Basement membrane
Bikunin	CS	25	Extracellular	Circulation
Brevican	CS	96	Extracellular	Brain
NG2	CS	251	Cell surface	Neural cells
Neurocan	CS	136	Extracellular	Brain
Neuropilin-1	CS	130	Cell surface	Endothelial, tumor cells
Phosphacan	CS	175	Cell surface	Brain
Versican	CS	265	Extracellular	Connective tissue
DSPGs				
Biglycan	DS, CS	38	Extracellular	Connective tissue
Decorin	DS, CS	36	Extracellular	Connective tissue, tumor cells
Epiphygan	DS	34	Extracellular	Cartilage
KSPGs				
Fibromodulin	KS I	42	Extracellular	Broadly expressed
Keratocan	KS I	37	Extracellular	Broadly expressed
Lumican	KS I	38	Extracellular	Broadly expressed
Mimecan	KS I	25	Extracellular	Broadly expressed
Part-time PGs				
Betaglycan	HS, CS	110	Cell surface	Broadly expressed
CD44	HS, CS, DS	37–81	Cell surface	Broadly expressed
Invariant chain	CS	31	Cell surface	Antigen processing cells
Thrombomodulin	CS	58	Cell surface	Endothelial cells

with other cell surface molecules. Glypican-5 can also harbor CS chains. GAG attachment sites in NG2 (CS) and CD44 (HS/CS/DS) are located in the middle portion of the core protein, whereas those of thrombomodulin (CS) and betaglycan (HS/CS) are proximal to the cell surface. The transmembrane and short cytoplasmic domains of syndecans are highly conserved across species, and along with the specific GAG attachment sites, are the signature motifs of this HSPG family. The transmembrane domain contains a GxxxG dimerization motif that mediates both homotypic and heterotypic oligomerization of syndecans (Dews and Mackenzie, 2007). The syndecan cytoplasmic domain contains several signaling and scaffolding motifs, such as three highly conserved Tyr, one highly conserved Ser, and a C-terminal Glu-Phe-Tyr-Ala (EFYA) PDZ binding domain. The invariant Tyr residues have been shown to be important in the regulation of syndecan-1 shedding

(Hayashida *et al.*, 2008b), whereas the EFYA motif binds to the PDZ protein syntenin (Grootjans *et al.*, 1997) and regulates syndecan sorting to the basolateral surface (Maday *et al.*, 2008) and exosome biogenesis (Baietti *et al.*, 2012), among other functions. The NG2 cytoplasmic domain also contains a PDZ binding domain and several Thr residues that can be phosphorylated. The CD44 cytoplasmic domain contains an ezrin, radixin, moesin (ERM) motif that links cell surface CD44 to the actin cytoskeleton.

Expression of cell surface proteoglycans is polarized and developmentally regulated. For example, syndecan-1 is first detected at the 4-cell stage in mouse embryos (Sutherland *et al.*, 1991), indicating that its expression is zygotically activated. In adult tissues, syndecan-1 is mainly expressed on the surface of epithelial and plasma cells. In culture, syndecan-1 is expressed broadly on the cell surface in subconfluent epithelial

cells, but its expression becomes highly polarized to basolateral surfaces in confluent cells, consistent with its role in mediating cell–cell and cell–matrix adhesions (Figure 1). This expression pattern of syndecan-1 in polarized, confluent cells resembles that of syndecan-1 in simple epithelial sheets *in vivo*. Glypicans are mainly targeted to the apical surfaces, though they can sometimes be expressed on basolateral surfaces. Apical sorting of glypicans is inversely related to its HS content (Mertens *et al.*, 1996).

Perlecan, agrin, and collagen XVIII are secreted and are the major HSPGs of basement membranes (Iozzo, 2005). Perlecan is so named from its appearance as a string of pearls in rotary shadowing images. Perlecan is widely expressed in basement membranes of various tissues. Agrin is localized to the basement membrane of neuromuscular junctions. Perlecan harbors 3–4 HS chains, whereas agrin is predicted to have 4 HS chains. Perlecan also contains *N*- and *O*-linked branched sugars, which are thought to function in its secretion. The aggrecan family of secreted proteoglycans comprised of aggrecan,

versican, brevican, and neurocan contains an amino-terminal hyaluronan binding domain, a central region that harbors CS chains, and a C-terminal C-type lectin domain. Aggrecan is a large, aggregating CS and KS containing proteoglycan found mainly in cartilaginous tissues. About 100 CS and 30 KS chains typically decorate aggrecan. Versican is expressed in the pericellular environment of most tissues. Versican exists in at least 4 isoforms generated by alternative splicing of exons that code for 2 CS attachment domains in the core protein. The CS chains attached to versican vary in size and composition, depending on cell type and culture conditions (Wight *et al.*, 2014).

Biglycan, decorin, fibromodulin, lumican, epiphykan, keratan, and mimecan comprise the family of small leucine-rich proteoglycans (SLRPs) (Iozzo and Schaefer, 2010). SLRPs are secreted and primarily found in ECMs. The core proteins of SLRPs contain leucine-rich repeats, which occupy more than 70% of the core proteins. Similar to other proteoglycans, the core proteins of SLRPs have distinct structural

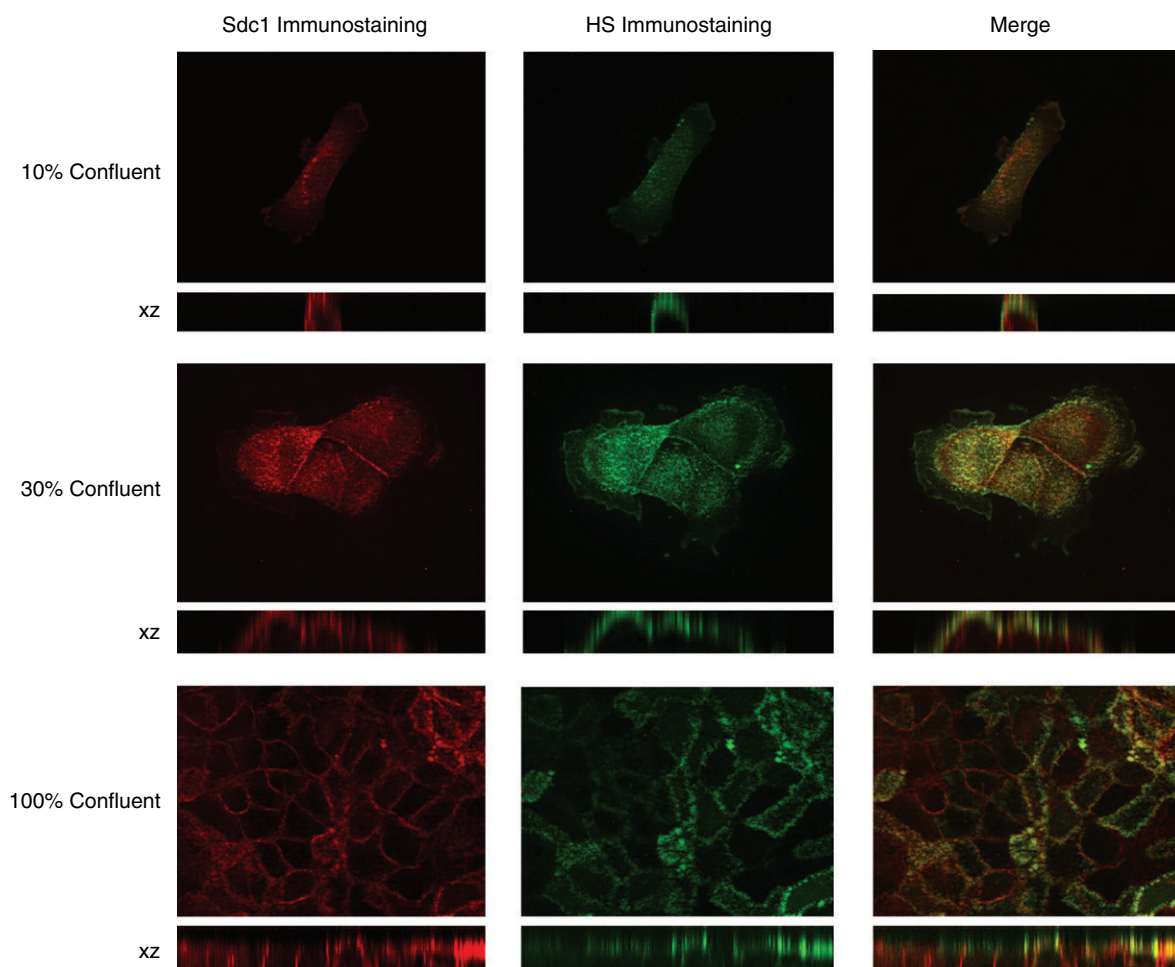


Figure 1 Expression of Syndecan-1 in Epithelial Cells. Normal murine mammary gland (NMG) epithelial cells at 10%, 30%, and 100% confluency were fixed in 4% paraformaldehyde, permeabilized, and immunostained with monoclonal rat anti-mouse syndecan-1 (281-2) and Alexa 594 (red) or mouse anti-HS (10E4) and Alexa 488 (green) antibodies and visualized. XZ images were generated by Image J from the Z stack of images taken through Zeiss Imager Z1 with apotome. Syndecan-1 and the HS epitope recognized by the 10E4 antibody are expressed mostly on the apical cell surface and in intracellular compartments in subconfluent cells, whereas syndecan-1 expression is polarized and predominantly on the basolateral cell surface in confluent cells.

designs with specific GAG attachment sites. For example, biglycan contains 2 CS or DS chains on a 38 kDa core protein. Decorin contains 1 CS or DS chain on a 36 kDa core protein. The CS containing decorin is found in developing bones, whereas the DS containing molecule is found in articular cartilage or tendon.

GAG Structure and Synthesis

GAGs are linear polysaccharides comprised of repeating disaccharide units that are defined by the composition and chemical linkage of the amino sugar and uronic acid monosaccharides in the disaccharide unit (Funderburgh, 2000; Esko and Selleck, 2002; Trowbridge and Gallo, 2002). The signature disaccharide repeat of a HS/HP is (GlcA β 1-4GlcNAc α 1-4) $_n$, CS is (GlcA β 1-3GalNAc β 1-4) $_n$, and DS is (GlcA/IdoA β 1-3GalNAc β 1-4) $_n$. CS and DS are distinguished from each other by the presence of GlcA and IdoA epimers. Except for KS, GAGs are attached to and polymerized on certain Ser residues of a Ser-Gly dipeptide sequence often repeated two or more times. KS chains are sulfated poly-*N*-acetyl lactosamines and classified based upon their core protein linkages. KS I is *N*-linked to specific Asn via high mannose type precursor oligosaccharide, KS II is *O*-linked to Ser or Thr via GalNAc, and KS III is *O*-linked to Ser or Thr via 2-*O*-mannose. Elongation of KS chains occurs through the enzymatic addition of Gal and GlcNAc, and sulfation of the polymer occurs at the 6-*O*-position of both sugar residues. The interaction of GAGs with proteins depends on their 3D structure, their sulfate substitution pattern, carboxyl groups, rotations around the glycosidic linkages, and the flexibility of the pyranose rings in the IdoA acid residues. The pyranose ring of the IdoA residue can assume both the 4C_1 chair and 4S_0 skew boat conformations and is flexible, which explains in part the extraordinary ability of IdoA containing GAGs, HP, HS, and DS, to bind to and regulate a wide range of proteins by facilitating the interaction of their anionic groups with cationic binding sites on proteins. Furthermore, HS chains are typically 50–200 disaccharide units in length, which equates to 40–160 nm in an extended, right-handed helical structure, suggesting that HS chains are a dominant feature of cell surfaces.

GAGs are attached to core proteins in a very specific and sequential manner by membrane-bound enzymes in the ER and multiple Golgi compartments (Funderburgh, 2000; Esko and Selleck, 2002; Trowbridge and Gallo, 2002). The biosynthesis of proteoglycans is a complex, nontemplate-driven process that requires many glycosyltransferases, an epimerase and sulfotransferases, which polymerize disaccharide units, epimerize GlcA to IdoA and sulfate nascent GAG chains. For example, HSPG synthesis is initiated by the formation of a tetrasaccharide linkage region (GlcA β 1-3Gal β 1-3Gal β 1-4Xyl β 1-*O*-Ser) on select Ser residues in the core protein. HS/HP synthesis is specified by the addition of an α 1,4-linked GlcNAc residue to the nonreducing end of the tetrasaccharide link attached to a Ser surrounded by a motif containing a hydrophobic amino acid adjacent to the Ser-Gly sequence and a cluster of acidic amino acids on the C-terminal side of Ser-Gly (Zhang *et al.*, 1995; Zhang and Esko, 1994). Synthesis of CS and DS uses an identical tetrasaccharide linkage region on

Ser residues in the Ser-Gly repeat, but the core protein motifs that specify CS/DS synthesis are not known.

HS is the most structurally complex GAG. Polymerization of the HS backbone proceeds by the addition of β 1,4-linked GlcA and α 1,4-linked GlcNAc units by glycosyltransferases of the exostosin family in alternating sequence to the nonreducing end of the growing polymer. Concomitant with HS chain elongation, several modifications occur through an epimerase and various sulfotransferases to generate a complex polysaccharide containing *N*-acetylated and *N*-sulfated glucosamine residues, GlcA and IdoA units, as well as *O*-sulfate groups in various positions. Various mutant CHO cell lines lacking specific HS biosynthetic enzymes have confirmed the specificity and sequential manner of HS biosynthesis, and the biological importance of HS and its modifications (Zhang *et al.*, 2006). Because the polymerization and modification reactions do not go to completion, the biosynthetic process generates an exceptionally diverse array of HS structure, both in length and extent of modification. Adding to the complexity is the fact that the composition of *N*- and *O*-sulfate groups in HS varies with cell type and tissue source. Thus, a mere HS decasaccharide can potentially assume over 10^6 distinct sequences, which is already in vast excess of the estimated gene products that the whole human genome can generate. This enormous structural diversity largely explains why and how HS binds to and regulates so many proteins.

Cellular Functions of Proteoglycans in Health and Disease

Proteoglycans have important developmental and post-developmental functions (Bernfield *et al.*, 1999; Iozzo and Sander-son, 2011; Wight *et al.*, 2014; Filmus and Capurro, 2014; Teng *et al.*, 2012; Perrimon and Bernfield, 2000). GAGs mediate the majority of ligand binding activities and endow proteoglycans the ability to interact with and regulate a multitude of protein ligands. The list includes growth factors, ECM components, cytokines, chemokines, antimicrobial peptides, coagulation factors, virulence factors, and many more (Xu and Esko, 2014; Schaefer and Iozzo, 2008; Filmus and Capurro, 2014; Bartlett and Park, 2010). Ligand binding by proteoglycans is primarily electrostatic, involving the anionic sulfate and carboxyl groups of GAGs and cationic ammonium, guanidinium, and imidazolium side chain functional groups of proteins.

Serglycin is critical in the maturation of mast cells, and regulates the storage and secretion of granule proteases, serotonin, and histamine in mast cells, platelets, neutrophils, and several other cells of hematopoietic origin (Scully *et al.*, 2012). These activities are mainly mediated by the HP chains since mice lacking *N*-deacetylase *N*-sulfotransferase-2, and hence HP chains, have abnormal mast cells with altered morphology and significantly reduced amounts of histamine and mast cell proteases (Forsberg *et al.*, 1999). HP is one of the oldest injectable drugs still in widespread clinical use (Wardrop and Keeling, 2008). HP has the highest negative charge density among the GAGs and also among any known biological molecule. HP functions as a potent anticoagulant by binding to and inducing a conformational change in antithrombin III, that facilitates the formation of tight, equimolar complexes of

antithrombin III with the serine proteases thrombin and factor Xa, resulting in efficient inhibition of these procoagulant enzymes. The minimum antithrombin III binding region in HP has been identified as a pentasaccharide sequence composed of 2 uronic acids and 3 glucosamine residues, with a 3-O-sulfated central glucosamine residue.

Cell surface proteoglycans, especially those harboring the structurally diverse HS chains, bind to a wide range of protein ligands and regulate cell–cell, cell–matrix, cell–growth factor, cell–morphogen, and cell–cytokine/chemokine interactions (Xu and Esko, 2014; Bernfield *et al.*, 1999; Couchman, 2010; Perrimon and Bernfield, 2000). Surface proteoglycans function as a receptor that translates the composition, organization, and stability of the extracellular environment into cellular activities. Surface proteoglycans can tether ligands and increase their local concentration at the cell surface, act as scaffolds and catalyze the encounter between ligands and their signaling receptors, increase the stability of ligands or receptors by protecting them from proteases, facilitate the oligomerization of ligands or receptors, or induce a conformational change of ligands or receptors (Xu and Esko, 2014; Bernfield *et al.*, 1999; Park *et al.*, 2000b). Cell surface proteoglycans function primarily as coreceptors for both soluble (e.g., growth factors) and insoluble ligands (e.g., ECM components) at the cell surface, though they can also serve as primary receptors for certain ligands, such as triglyceride-rich lipoproteins (Stanford *et al.*, 2009). Furthermore, cell surface HSPGs serve as endocytic receptors for several molecular cargo, such as cell penetrating peptides, polycation-nucleic acid complexes and morphogens, and modulate endocytic signaling (Christianson and Belting, 2014), whereas syndecan-1 can be transported to the nucleus where they have been shown to regulate transcription (Stewart and Sanderson, 2014).

The nature of the interaction of cell surface proteoglycans with growth factors has been the subject of intense study since the initial observation in 1991 that cell surface HSPGs are required for FGF-2 signaling (Yayon *et al.*, 1991; Rapraeger *et al.*, 1991). In the absence of cell surface HS, FGF-2 interacts poorly with its high affinity receptor (FGFR-1) and intracellular signaling is not activated. Cell surface HSPGs, such as syndecans, form a ternary complex with FGF-2 and FGFR-1, and this facilitates growth factor signaling. Subsequent studies have shown that HSPGs may regulate the oligomerization of both ligand and receptor (Schlessinger *et al.*, 1995), although some have also suggested that oligomerization is not essential (Pye and Gallagher, 1999). While the regulatory mechanism is not fully elucidated, it is clear that HSPGs bind specifically and tightly to FGF-2. In fact, crystal structures of complexes of FGF-2 with a HP-derived tetrasaccharide and hexasaccharide were the first reported for a protein-GAG oligosaccharide complex (Faham *et al.*, 1996).

In addition to FGF-2, cell surface HSPGs enhance the signaling of other growth factors and morphogens, such as EGF, VEGF, TGF β , HGF, sonic hedgehog (Hh), BMP, ephrin, and Wnt family proteins, among others (Bernfield *et al.*, 1999). Glypicans in particular are important regulators of growth factors and morphogens during development. Studies with *Drosophila* mutants have shown that glypicans control a diverse set of patterning events (Perrimon and Bernfield, 2000). Glypicans also function in axon guidance and formation of

excitatory synapses (Filmus and Capurro, 2014; Filmus *et al.*, 2008). Glypicans regulate the formation of Hh gradients, which is thought to be important in Simpson–Golabi–Behmel syndrome, a rare X-linked condition caused by a loss-of-function mutation of glypican-3 and characterized by pre- and postnatal overgrowth (Filmus and Capurro, 2014). Glypican-3 acts as an inhibitor of Hh signaling by competing with Patched for Hh binding. Interestingly, glypican-5 promotes Hh–Patched interaction by binding to both Hh and Patched through HS. How this is accomplished is not fully understood, but GPC5 is thought to function in this manner because HS chains of glypican-5 are more highly sulfated compared to those of glypican-3. In fact, glypican-5 is upregulated in rhabdomyosarcoma, and increased glypican-5 is thought to increase Hh signaling and promote cancer progression. Furthermore, a loss of function mutation in glypican-6 has been identified in patients with autosomal recessive omodyplasia, characterized by short-limbed short stature and craniofacial dysmorphism, consistent with reduced Hh activity in bones, indicating that glypicans are critical regulators of Hh signaling in many tissues. However, the activity of glypicans in the Wnt, Hh, and BMP signaling pathways is only partially dependent on the HS chains (Filmus *et al.*, 2008). Cell surface HSPGs also play important developmental roles in left-right development in the migrating mesoderm (Kramer *et al.*, 2002) and midline axon guidance (Inatani *et al.*, 2003). While differential phosphorylation of syndecan-2 by PKC γ in right ectodermal cells but not left cells is known to regulate left-right development, the HSPG that regulates midline axon guidance remains to be identified.

Cell surface HSPGs also control the onset and progression of various post-developmental diseases, including protein losing enteropathy (Bode *et al.*, 2008), Alzheimer's disease (Scholefield *et al.*, 2003), inflammatory disorders (Bartlett *et al.*, 2007), infection (Aquino *et al.*, 2010), and cancer (Iozzo and Sanderson, 2011). In protein losing enteropathy, syndecan-1 is required to maintain the integrity of intestinal epithelial barriers. While the underlying mechanism is not understood, HS chains of syndecan-1 are thought to bind to IFN γ and TNF α in an *N*-sulfate-dependent manner, and inhibit their detrimental effects on epithelial barrier integrity. In Alzheimer's disease, cell surface HSPGs bind to and inhibit Alzheimer's β -secretase (BACE-1) and the generation of amyloid β -peptide. Here, HS binds to and regulates BACE-1 in a 2-O- and 6-O-sulfate-dependent manner, but the identity of the cell surface HSPG responsible for this activity is not known. Cell surface HSPGs also regulate several key steps of leukocyte recruitment to sites of injury and infection. HSPGs bind to most chemokines in an HS-dependent manner, increase the local concentration of chemokines, and facilitate the formation of a chemokine gradient that guides the migration of inflammatory cells (Handel *et al.*, 2005). Solution NMR spectroscopy has clearly shown that GAGs play a major role in chemokine oligomerization and function (Pomin, 2014). Consistent with these observations, syndecan-1 has been shown to modulate leukocyte recruitment in various mouse models of inflammatory diseases (Teng *et al.*, 2012), including acute lung injury (Li *et al.*, 2002), allergic lung inflammation (Xu *et al.*, 2005) and systemic shock (Hayashida *et al.*, 2008a, 2009). In these inflammatory disorders, syndecan-1 promotes, or attenuates,

inflammation by regulating the compartmentalization, activity, or removal of various chemokines. In infectious disease, cell surface proteoglycans serve as attachment and invasion receptors for a wide variety of microbial pathogens, including viruses, bacteria, parasites, and fungi (Aquino *et al.*, 2010). For instance, *Neisseria gonorrhoeae*, one of the most common causes of bacterial sexually transmitted disease, binds to syndecan-1 and -4 via Opa proteins and stimulates a signaling pathway that involves phosphatidylcholine-specific phospholipase C, diacylglycerol, acidic sphingomyelinase, and ceramide, to invade host cells (Freissler *et al.*, 2000; Grassmé *et al.*, 1997). Herpes simplex virus has been shown to bind to an unusual *N*-unsubstituted and 3-*O*-sulfated glucosamine unit of cell surface HSPGs via gD protein to promote its infection (Shukla *et al.*, 1999). In cancer, the activity of cell surface HSPGs impinge on several key processes of tumorigenesis, such as cancer cell survival and apoptosis, angiogenesis, and metastasis (Iozzo and Sanderson, 2011; Teng *et al.*, 2012). Expression of syndecans and glypicans are highly altered in many cancers, suggesting that tumor cells regulate the expression of these cell surface HSPGs to promote their pathogenesis.

Interestingly many cell surface proteoglycans, including all members of the syndecan and glypican family, are released from the cell surface by highly regulated enzymatic mechanisms. The released proteoglycans are replete with their GAG chains and function as either paracrine or autocrine regulators of molecular interactions at the cell surface and in the extracellular milieu (Hayashida *et al.*, 2010; Teng *et al.*, 2012). Syndecans are shed from the cell surface by several metalloproteinases (Hayashida *et al.*, 2010; Teng *et al.*, 2012), whereas glypicans are mainly released by phospholipases, such as Notum (Traister *et al.*, 2008). Once released from the cell surface, proteoglycans show functions similar to or distinct from their immobilized counterparts. For example, several bacterial pathogens induce the release of syndecan-1 from the cell surface by activating an endogenous PTK- and MMP-dependent shedding mechanism (Park *et al.*, 2000a, 2004). Shed syndecan-1 has been shown to promote infection by inhibiting antimicrobial peptides and neutrophil-mediated bacterial killing mechanisms (Hayashida *et al.*, 2011; Park *et al.*, 2001).

Several secreted proteoglycans surround the cell and are ideally situated to regulate signaling by growth factors and other biological effectors. For instance, perlecan controls the pericellular concentration of various mitogens and morphogens, particularly growth factors of the FGF family (FGF-2, -7, and -10) (Aviezer *et al.*, 1994; Iozzo, 2005). Mice that lack perlecan have a complex series of developmental phenotypes, which is not confined to one tissue or organ system (Arikawa-Hirasawa *et al.*, 1999; Costell *et al.*, 1999). Mice that lack collagen XVIII have abnormalities in eye development and some effects on angiogenesis (Iozzo, 2005), whereas mice lacking agrin have defective neuromuscular junctions due to the inability to correctly cluster acetylcholine receptors on muscle fiber membranes (Gautam *et al.*, 1996). The negative charge of basement membrane proteoglycans also regulates the filtration properties of macromolecules in the vasculature, which is especially important in the kidney.

The aggrecan family of pericellular proteoglycans has important roles in cell-cell and cell-matrix adhesion, assembly

of ECM, and in providing the hyperosmotic properties necessary to counter compressive loads on tissues. Degradation of aggrecan by aggrecanases such as ADAMTS5 is a major cause of osteoarthritis (Stanton *et al.*, 2005), and mutations in genes for the aggrecan family core proteins and CS and DS biosynthetic enzymes can lead to chondrodysplasia and various disorders of connective tissues (Schwartz and Domowicz, 2002; Mizumoto *et al.*, 2013). Versican affects the adhesion and activation of chondrocytes and several leukocytes, including T cells and monocytes, and regulates the assembly of elastic fibers (Wight *et al.*, 2014).

The SLRP family of secreted proteoglycans regulates cellular responses to growth factors, immune responses, and assembly of collagen fibers in several tissues (Schaefer and Iozzo, 2008; Iozzo and Schaefer, 2010). Decorin has anti-proliferative effects on cancer cells via EGFR and c-Met (receptor for HGF) suppression. Lumican also inhibits tumor cell growth in soft agar by increasing the expression of the CDK inhibitor p21WAF1. However, in normal cells, decorin signaling through insulin-like growth factor receptor type I has anti-apoptotic and proliferative effects, favoring cellular growth. The SLRPs regulate collagen fibrillogenesis in the tendon and corneal stroma (Chakravarti *et al.*, 1998; Ezura *et al.*, 2000). Ablation of lumican in mice leads to disorganized collagen fibers and produces fragile skin and opaque corneas (Chakravarti *et al.*, 1998). Decorin, fibromodulin, and keratocan also bind to interstitial collagens and regulate collagen fiber assembly. Interestingly, lumican has been shown to bind to and present LPS to CD14, thereby activating TLR4 signaling in macrophages, whereas biglycan directly binds to and activates TLR2 and TLR4 and functions as a DAMP (danger-associated molecular pattern) (Iozzo and Schaefer, 2010).

Post-synthesis processing of GAGs also has a strong impact on cellular functions in health and disease. For example, membrane-bound HS sulfatases (SULF1 and SULF2), which remove 6-*O*-sulfate groups from HS, regulate FGF signaling and angiogenesis (Dhoot *et al.*, 2001; Wang *et al.*, 2004). Heparanase is an HS-specific endo- β -D-glucuronidase that plays multiple roles in growth factor signaling, cancer metastasis, diabetes, and atherosclerosis (Vlodavsky *et al.*, 2013). Heparanase is highly active in acidic conditions (pH 4–6.5), suggesting that it is primarily a lysosomal enzyme, but its expression and activity have been detected in cell membranes and in secreted forms. Proteoglycans are naturally degraded by an endocytic mechanism that involves proteolysis and endoglycosidic cleavage in intracellular vesicles, and degradation into monosaccharides and free sulfates by exoglycosidases and sulfatases in lysosomes. The inability to breakdown proteoglycans is characteristic of a group of genetic disorders called mucopolysaccharidoses, where deficiencies occur in one or more enzymes involved in the degradation of sulfated GAGs and hyaluronan (Lawrence *et al.*, 2014).

Research on proteoglycans and sulfated GAGs has come a long way since the discovery of aggrecan in cartilage. We now know that proteoglycans not only function as hyperosmotic cushions for variable, compressive forces in joints, but also as critical regulators of both developmental and post-developmental processes, involving cell-cell, cell-matrix, and receptor-ligand interactions. GAGs on proteoglycans bind to and regulate a wide variety of biological effectors in

intracellular, cell surface, and extracellular compartments, where they facilitate molecular interactions in various ways. However, much of their structure–function relationship remains to be elucidated as methodologies to analyze GAGs still lag behind those of protein biochemistry. Our understanding of how the synthesis and activities of proteoglycans are precisely regulated under normal and disease conditions also needs to be expanded to better understand their cell biology and to design and develop novel proteoglycan- and GAG-based therapies.

References

- Aquino, R.S., Lee, E.S., Park, P.W., 2010. Diverse functions of glycosaminoglycans in infectious diseases. *Progress in molecular biology and translational science* 93, 373–394.
- Arikawa-Hirasawa, E., Watanabe, H., Takami, H., Hassell, J.R., Yamada, Y., 1999. Perlecan is essential for cartilage and cephalic development. *Nature Genetics* 23, 354–358.
- Aviezer, D., Hecht, D., Safran, M., *et al.*, 1994. Perlecan, a basal lamina proteoglycan, promotes basic fibroblast growth factor-receptor binding, mitogenesis, and angiogenesis. *Cell* 79, 1005–1013.
- Baietti, M.F., Zhang, Z., Mortier, E., *et al.*, 2012. Syndecan-syntenin-ALIX regulates the biogenesis of exosomes. *Nature Cell Biology* 14, 677–685.
- Bartlett, A.H., Hayashida, K., Park, P.W., 2007. Molecular and cellular mechanisms of syndecans in tissue injury and inflammation. *Molecules and Cells* 24, 153–166.
- Bartlett, A.H., Park, P.W., 2010. Proteoglycans in host-pathogen interactions: Molecular mechanisms and therapeutic implications. *Expert Reviews in Molecular Medicine* 12, e5.
- Bernfield, M., Götte, M., Park, P.W., *et al.*, 1999. Functions of cell surface heparan sulfate proteoglycans. *Annual Review of Biochemistry* 68, 729–777.
- Bode, L., Salvestrini, C., Park, P.W., *et al.*, 2008. Heparan sulfate and syndecan-1 are essential in maintaining murine and human intestinal epithelial barrier function. *Journal of Clinical Investigation* 118, 229–238.
- Chakravarti, S., Magnusson, T., Lass, J.H., *et al.*, 1998. Lumican regulates collagen fibril assembly: Skin fragility and corneal opacity in the absence of lumican. *Journal of Cell Biology* 141, 1277–1286.
- Christianson, H.C., Belting, M., 2014. Heparan sulfate proteoglycan as a cell-surface endocytosis receptor. *Matrix Biology* 35, 51–55.
- Costell, M., Gustafsson, E., Aszodi, A., *et al.*, 1999. Perlecan maintains the integrity of cartilage and some basement membranes. *Journal of Cell Biology* 147, 1109–1122.
- Couchman, J.R., 2010. Transmembrane signaling proteoglycans. *Annual Review of Cell and Developmental Biology* 26, 89–114.
- Dews, I.C., Mackenzie, K.R., 2007. Transmembrane domains of the syndecan family of growth factor coreceptors display a hierarchy of homotypic and heterotypic interactions. *Proceedings of the National Academy of Sciences of the United States of America* 104, 20782–20787.
- Dhoot, G.K., Gustafsson, M.K., Ai, X., *et al.*, 2001. Regulation of Wnt signaling and embryo patterning by an extracellular sulfatase. *Science* 293, 1663–1666.
- Esko, J.D., Selleck, S.B., 2002. Order out of chaos: Assembly of ligand binding sites in heparan sulfate. *Annual Review of Biochemistry* 71, 435–471.
- Ezura, Y., Chakravarti, S., Oldberg, A., Chervoneva, I., Birk, D.E., 2000. Differential expression of lumican and fibromodulin regulate collagen fibrillogenesis in developing mouse tendons. *Journal of Cell Biology* 151, 779–788.
- Faham, S., Hileman, R.E., Fromm, J.R., Linhardt, R.J., Rees, D.C., 1996. Heparin structure and interactions with basic fibroblast growth factor. *Science* 271, 1116–1120.
- Filmus, J., Capurro, M., 2014. The role of glypicans in Hedgehog signaling. *Matrix Biology* 35, 248–252.
- Filmus, J., Capurro, M., Rast, J., 2008. Glypicans. *Genome Biology* 9, 224.
- Forsberg, E., Pejler, G., Ringvall, M., *et al.*, 1999. Abnormal mast cells in mice deficient in a heparin-synthesizing enzyme. *Nature* 400, 773–776.
- Freissler, E., Meyer auf der Heyde, A., David, G., Meyer, T.F., Dehio, C., 2000. Syndecan-1 and syndecan-4 can mediate the invasion of *Op_{HSPG}*-expressing *Neisseria gonorrhoeae* into epithelial cells. *Cellular Microbiology* 2, 69–82.
- Funderburgh, J.L., 2000. Keratan sulfate: Structure, biosynthesis, and function. *Glycobiology* 10, 951–958.
- Gautam, M., Noakes, P.G., Moscoso, L., *et al.*, 1996. Defective neuromuscular synaptogenesis in agrin-deficient mice. *Cell* 85, 525–535.
- Grassmé, H., Gulbins, E., Brenner, B., *et al.*, 1997. Acidic sphingomyelinase mediates entry of *N. gonorrhoeae* into nonphagocytic cells. *Cell* 91, 605–615.
- Grootjans, J.J., Zimmermann, P., Reekmans, G., *et al.*, 1997. Syntenin, a PDZ protein that binds syndecan cytoplasmic domains. *Proceedings of the National Academy of Sciences of the United States of America* 94, 13683–13688.
- Handel, T.M., Johnson, Z., Crown, S.E., Lau, E.K., Proudfoot, A.E., 2005. Regulation of protein function by glycosaminoglycans – As exemplified by chemokines. *Annual Review of Biochemistry* 74, 385–410.
- Hayashida, A., Amano, S., Park, P.W., 2011. Syndecan-1 promotes *Staphylococcus aureus* corneal infection by counteracting neutrophil-mediated host defense. *Journal of Biological Chemistry* 285, 3288–3297.
- Hayashida, K., Bartlett, A.H., Chen, Y., Park, P.W., 2010. Molecular and cellular mechanisms of ectodomain shedding. *Anatomical Record* 293, 925–937.
- Hayashida, K., Chen, Y., Bartlett, A.H., Park, P.W., 2008a. Syndecan-1 is an *in vivo* suppressor of Gram-positive toxic shock. *Journal of Biological Chemistry* 283, 19895–19903.
- Hayashida, K., Parks, W.C., Park, P.W., 2009. Syndecan-1 shedding facilitates the resolution of neutrophilic inflammation by removing sequestered CXC chemokines. *Blood* 114, 3033–3043.
- Hayashida, K., Stahl, P.D., Park, P.W., 2008b. Syndecan-1 ectodomain shedding is regulated by the small GTPase Rab5. *Journal of Biological Chemistry* 283, 35435–35444.
- Inatani, M., Irie, F., Plump, A.S., Tessier-Lavigne, M., Yamaguchi, Y., 2003. Mammalian brain morphogenesis and midline axon guidance require heparan sulfate. *Science* 302, 1044–1046.
- Iozzo, R.V., 2005. Basement membrane proteoglycans: From cellar to ceiling. *Nature Reviews Molecular Cell Biology* 6, 646–656.
- Iozzo, R.V., Sanderson, R.D., 2011. Proteoglycans in cancer biology, tumour microenvironment and angiogenesis. *Journal of Cellular and Molecular Medicine* 15, 1013–1031.
- Iozzo, R.V., Schaefer, L., 2010. Proteoglycans in health and disease: Novel regulatory signaling mechanisms evoked by the small leucine-rich proteoglycans. *FEBS Journal* 277, 3864–3875.
- Kokenyesi, R., Bernfield, M., 1994. Core protein structure and sequence determine the site and presence of heparan sulfate and chondroitin sulfate on syndecan-1. *Journal of Biological Chemistry* 269, 12304–12309.
- Kramer, K.L., Barnette, J.E., Yost, H.J., 2002. PKCγ regulates syndecan-2 inside-out signaling during xenopus left-right development. *Cell* 111, 981–990.
- Lawrence, R., Brown, J.R., Lorey, F., *et al.*, 2014. Glycan-based biomarkers for mucopolysaccharidoses. *Molecular Genetics and Metabolism* 111, 73–83.
- Li, Q., Park, P.W., Wilson, C.L., Parks, W.C., 2002. Matrilysin shedding of syndecan-1 regulates chemokine mobilization and transepithelial efflux of neutrophils in acute lung injury. *Cell* 111, 635–646.
- Maday, S., Anderson, E., Chang, H.C., *et al.*, 2008. A PDZ-binding motif controls basolateral targeting of syndecan-1 along the biosynthetic pathway in polarized epithelial cells. *Traffic* 9, 1915–1924.
- Mertens, G., Van der Schueren, B., van den Berghe, H., David, G., 1996. Heparan sulfate expression in polarized epithelial cells: The apical sorting of glypican (GPI-anchored proteoglycan) is inversely related to its heparan sulfate content. *Journal of Cell Biology* 132, 487–497.
- Mizumoto, S., Ikegawa, S., Sugahara, K., 2013. Human genetic disorders caused by mutations in genes encoding biosynthetic enzymes for sulfated glycosaminoglycans. *Journal of Biological Chemistry* 288, 10953–10961.
- Park, P.W., Foster, T.J., Nishi, E., *et al.*, 2004. Activation of syndecan-1 ectodomain shedding by *Staphylococcus aureus* alpha-toxin and beta-toxin. *Journal of Biological Chemistry* 279, 251–258.
- Park, P.W., Pier, G.B., Hinkes, M.T., Bernfield, M., 2001. Exploitation of syndecan-1 shedding by *Pseudomonas aeruginosa* enhances virulence. *Nature* 411, 98–102.
- Park, P.W., Pier, G.B., Preston, M.J., *et al.*, 2000a. Syndecan-1 shedding is enhanced by LasA, a secreted virulence factor of *Pseudomonas aeruginosa*. *Journal of Biological Chemistry* 275, 3057–3064.
- Park, P.W., Reizes, O., Bernfield, M., 2000b. Cell surface heparan sulfate proteoglycans: Selective regulators of ligand-receptor encounters. *Journal of Biological Chemistry* 275, 29923–29926.
- Perrimon, N., Bernfield, M., 2000. Specificities of heparan sulphate proteoglycans in developmental processes. *Nature* 404, 725–728.
- Pomin, V.H., 2014. Biological findings from the recent NMR-based studies of glycosaminoglycan-protein interactions. *Glycobiology* 24, 991–1003.
- Pye, D.A., Gallagher, J.T., 1999. Monomer complexes of basic fibroblast growth factor and heparan sulfate oligosaccharides are the minimal functional unit for cell activation. *Journal of Biological Chemistry* 274, 13456–13461.

- Rapraeger, A.C., Krufka, A., Olwin, B.B., 1991. Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. *Science* 252, 1705–1708.
- Schaefer, L., Iozzo, R.V., 2008. Biological functions of the small leucine-rich proteoglycans: From genetics to signal transduction. *Journal of Biological Chemistry* 283, 21305–21309.
- Schlessinger, J., Lax, I., Lemmon, M., 1995. Regulation of growth factor activation by proteoglycans: What is the role of the low affinity receptors? *Cell* 83, 357–360.
- Scholefield, Z., Yates, E.A., Wayne, G., *et al.*, 2003. Heparan sulfate regulates amyloid precursor protein processing by BACE1, the Alzheimer's beta-secretase. *Journal of Cell Biology* 163, 97–107.
- Schwartz, N.B., Domowicz, M., 2002. Chondrodysplasias due to proteoglycan defects. *Glycobiology* 12, 57R–68R.
- Scully, O.J., Chua, P.J., Harve, K.S., Bay, B.H., Yip, G.W., 2012. Serglycin in health and diseases. *Anatomical Record (Hoboken)* 295, 1415–1420.
- Shukla, D., Liu, J., Blaiklock, P., *et al.*, 1999. A novel role for 3-O-sulfated heparan sulfate in herpes simplex virus 1 entry. *Cell* 99, 13–22.
- Stanford, K.I., Bishop, J.R., Foley, E.M., *et al.*, 2009. Syndecan-1 is the primary heparan sulfate proteoglycan mediating hepatic clearance of triglyceride-rich lipoproteins in mice. *Journal of Clinical Investigation* 119, 3236–3245.
- Stanton, H., Rogerson, F.M., East, C.J., *et al.*, 2005. ADAMTS5 is the major aggrecanase in mouse cartilage *in vivo* and *in vitro*. *Nature* 434, 648–652.
- Stewart, M.D., Sanderson, R.D., 2014. Heparan sulfate in the nucleus and its control of cellular functions. *Matrix Biology* 35, 56–59.
- Sutherland, A.E., Sanderson, R.D., Mayes, M., *et al.*, 1991. Expression of syndecan, a putative low affinity fibroblast growth factor receptor, in the early mouse embryo. *Development* 113, 339–351.
- Teng, Y.H., Aquino, R.S., Park, P.W., 2012. Molecular functions of syndecan-1 in disease. *Matrix Biology* 31, 3–16.
- Traister, A., Shi, W., Filmus, J., 2008. Mammalian Notum induces the release of glypicans and other GPI-anchored proteins from the cell surface. *Biochemical Journal* 410, 503–511.
- Trowbridge, J.M., Gallo, R.L., 2002. Dermatan sulfate: New functions from an old glycosaminoglycan. *Glycobiology* 12, 117R–125R.
- Vlodavsky, I., Iozzo, R.V., Sanderson, R.D., 2013. Heparanase: Multiple functions in inflammation, diabetes and atherosclerosis. *Matrix Biology* 32, 220–222.
- Wang, S., Ai, X., Freeman, S.D., *et al.*, 2004. QSulf1, a heparan sulfate 6-O-endosulfatase, inhibits fibroblast growth factor signaling in mesoderm induction and angiogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 101, 4833–4838.
- Wardrop, D., Keeling, D., 2008. The story of the discovery of heparin and warfarin. *British Journal of Haematology* 141, 757–763.
- Wight, T.N., Kinsella, M.G., Evanko, S.P., Potter-Perigo, S., Merrilees, M.J., 2014. Versican and the regulation of cell phenotype in disease. *Biochimica et Biophysica Acta* 1840, 2441–2451.
- Xu, D., Esko, J.D., 2014. Demystifying heparan sulfate-protein interactions. *Annual Review of Biochemistry* 83, 129–157.
- Xu, J., Park, P.W., Kheradmand, F., Corry, D.B., 2005. Endogenous attenuation of allergic lung inflammation by syndecan-1. *Journal of Immunology* 174, 5758–5765.
- Yayon, A., Klagsbrun, M., Esko, J.D., Leder, P., Ornitz, D.M., 1991. Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell* 64, 841–848.
- Zhang, L., David, G., Esko, J.D., 1995. Repetitive Ser-Gly sequences enhance heparan sulfate assembly in proteoglycans. *Journal of Biological Chemistry* 270, 27127–27135.
- Zhang, L., Esko, J.D., 1994. Amino acid determinants that drive heparan sulfate assembly in a proteoglycan. *Journal of Biological Chemistry* 269, 19295–19299.
- Zhang, L., Lawrence, R., Frazier, B.A., Esko, J.D., 2006. CHO glycosylation mutants: Proteoglycans. *Methods in Enzymology* 416, 205–221.

Further Reading

- Bernfield, M., Kokenyesi, R., Kato, M., *et al.*, 1992. Biology of the syndecans: A family of transmembrane heparan sulfate proteoglycans. *Annual Review of Cell and Developmental Biology* 8, 365–393.
- Bishop, J.R., Schuksz, M., Esko, J.D., 2007. Heparan sulphate proteoglycans fine-tune mammalian physiology. *Nature* 446, 1030–1037.
- Busse-Wicher, M., Wicher, K.B., Kusche-Gullberg, M., 2014. The extostoin family: Proteins with many functions. *Matrix Biology* 35, 25–33.
- Conrad, H.E., 1998. Heparin-Binding Proteins. San Diego, CA: Academic Press, Inc..
- Fuerer, C., Habib, S.J., Nusse, R., 2010. A study on the interactions between heparan sulfate proteoglycans and Wnt proteins. *Developmental Dynamics* 239, 184–190.
- Kraemer, P.M., 1971. Heparan sulfates of cultured cells. II. Acid-soluble and – precipitable species of different cell lines. *Biochemistry* 10, 1445–1451.
- Rabenstein, D.L., 2002. Heparin and heparan sulfate: Structure and function. *Natural Product Reports* 19, 312–331.
- Rosenberg, R.D., Shworak, N.W., Liu, J., Schwartz, J.J., Zhang, L., 1997. Heparan sulfate proteoglycans of the cardiovascular system. Specific structures emerge but how is synthesis regulated? *Journal of Clinical Investigation* 99, 2062–2070.
- Sajdera, S.W., Hascall, V.C., 1969. Proteinopolysaccharide complex from bovine nasal cartilage. A comparison of low and high shear extraction procedures. *Journal of Biological Chemistry* 244, 77–87.
- Sanderson, R.D., 2004. Heparan sulfate proteoglycans and heparanase: Partners in osteolytic tumor growth and metastasis. *Matrix Biology* 23, 341–352.
- Sasisekharan, R., Raman, R., Prabhakar, V., 2006. Glycomics approach to structure-function relationships of glycosaminoglycans. *Annual Review of Biomedical Engineering* 8, 181–231.
- Teder, P., Vandivier, R.W., Jiang, D., *et al.*, 2002. Resolution of lung inflammation by CD44. *Science* 296, 155–158.
- Xu, Y., Masuko, S., Takiyeddin, M., *et al.*, 2011. Chemoenzymatic synthesis of homogeneous ultralow molecular weight heparins. *Science* 334, 498–501.

Hyaluronan

VC Hascall, Cleveland Clinic, Cleveland, OH, USA

PH Weigel, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

BP Toole, Medical University of South Carolina, Charleston, SC, USA

© 2016 Elsevier Inc. All rights reserved.

Primary Sources

'Hyaluronan Today,' accessible at the Glycoforum website, was initiated and sponsored by Seikagaku in 1997. Its purpose is to provide articles about the history and status of hyaluronan research in its many biological roles and functions. It presently contains 37 articles through to 2011 that provide overviews and references to the many hyaluronan 'branches' that are so nicely described in article 23, 'The Tree': Hyaluronan Research in the twentieth century, by Torvard Laurent (Laurent, 2002), one of the early pioneers in hyaluronan research. Hyaluronan is unique in many ways as described in this article (see Relevant Websites section).

Hyaluronan also has its own Society (International Society for Hyaluronan Sciences (ISHAS)) that was founded in 2004 by Endre Balazs, another early pioneer in hyaluronan research who focused on its roles and uses in medical applications. ISHAS now supports international hyaluronan conferences, and the program booklets from many previous conferences are accessible on the ISHAS website along with other current information about hyaluronan research. These sites are valuable sources of information and for contact information of many leaders in current hyaluronan research (see Relevant Websites section).

Another valuable source of information is the series of 5 hyaluronan books edited by Endre Balazs and published by the Matrix Biology Institute. This series includes interviews of many hyaluronan researchers conducted during 2000–2010 and chapters on the early history of research relevant to hyaluronan before the seminal work of Karl Meyer (see Relevant Websites section).

Introduction

In 1934, Karl Meyer and his assistant, John Palmer, described a procedure for isolating a novel glycosaminoglycan from the vitreous of bovine eyes (Meyer and Palmer, 1934; Hascall and Laurent, 1997). They showed that this substance contained a uronic acid and an aminosugar, but no sulfoesters. In their words: "we propose for convenience, the name 'hyaluronic acid,' from hyaloid (vitreous) + uronic acid." This marked the birth announcement for one of nature's most versatile and fascinating macromolecules. Today, this macromolecule is most frequently referred to as 'Hyaluronan,' reflecting the fact that it exists *in vivo* as a polyanion and not in a protonated form.

Structure and Cellular Metabolism

It took Karl Meyer's laboratory 20 years to determine the structure of the disaccharide (Weissman and Meyer, 1954; Figure 1(b)) that is the basic building block of the hyaluronan

glycosaminoglycan, typically synthesized with more than 10 000 repeat units for molecular masses $>5 \times 10^6$ Da (5 MDa; Figure 1(a)). Unlike the majority of other biomolecules, hyaluronan has no true molecular weight because in nature it does not have a unique size (number of sugars). All natural preparations of purified hyaluronan are very polydisperse, usually with sizes that span a 100-fold range. How do cells manage to do that? It requires only one enzyme, a hyaluronan synthase (HAS), and there are 3 such isozymes (HAS1-3) in vertebrate genomes (Weigel and DeAngelis, 2007; Spicer and MacDonald, 1998). In contrast to the other glycosaminoglycans (chondroitin/dermatan sulfates, heparin/heparan sulfate, and keratan sulfate), which are synthesized as proteoglycans on core proteins in the Golgi, hyaluronan is normally synthesized (not attached to a protein) at the plasma membrane using the cytosolic substrates, UDP-glucuronate and UDP-N-acetylglucosamine. The UDP-sugars are alternately

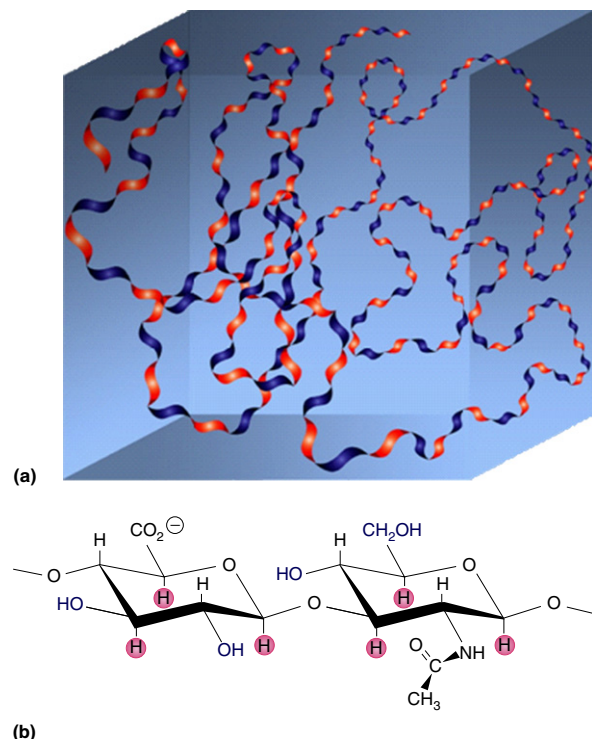


Figure 1 Disaccharide and domain structures of hyaluronan. (a) The ribbon structure of hyaluronan. The model shows how the hydrophilic structure of the large number of repeat disaccharides (GlcUA- β 1,3-GlcNAc- β 1,4) (b) of a hyaluronan macromolecule can occupy a large solvent domain that would hinder transport of large molecules while being permeable to small molecules. Modified from Hascall, V., Laurent, T., 1997. Hyaluronan: Structure and physical properties. Hyaluronan Today, article 1.

used as acceptors by being added onto the reducing end of a growing hyaluronan-UDP chain as the activating UDP is released, and the elongating chain is ratcheted through the enzyme itself within the plasma membrane into the extracellular space (Figure 2(a)). An animated cartoon (Weigel, 1998, Update 2005) illustrates how HAS might couple the addition of new sugars with the translocation of hyaluronan to the cell exterior, using only the energy provided by the activated sugar nucleotides. This novel intraprotein translocation mechanism, first proposed in 1999 (Tlapak-Simmons *et al.*, 1999), was later confirmed (Hubbard *et al.*, 2012). This product extrusion mechanism allows the final hyaluronan macromolecules to become very large as they are not constrained by synthesis in intracellular compartments. It is important to transport newly synthesized HAS enzymes to the plasma membrane before

they are activated. They are integrated into specific regions on the membrane that can form spikes when the enzyme is overexpressed and active, forming hedgehog-like cell structures (Kultti *et al.*, 2006; Figure 3). These regions are also involved in formation of exosomes that can package and transport proteins between cells and tissues (Rilla *et al.*, 2010).

Most cells are able to dynamically synthesize hyaluronan and form cell-associated glycocalyxes that include chondroitin sulfate proteoglycans (e.g., versican and aggrecan) bound to the hyaluronan (Figure 4). Some cells also have surface organelles that dynamically catabolize these glycocalyxes through cooperative action of a hyaluronan receptor, often CD44, a protease (e.g., an aggrecanase) and a hyaluronidase (Hyal) such as glycosylphosphatidylinositol (GPI)-anchored Hyal2 (Figure 4). The fragments are internalized into an

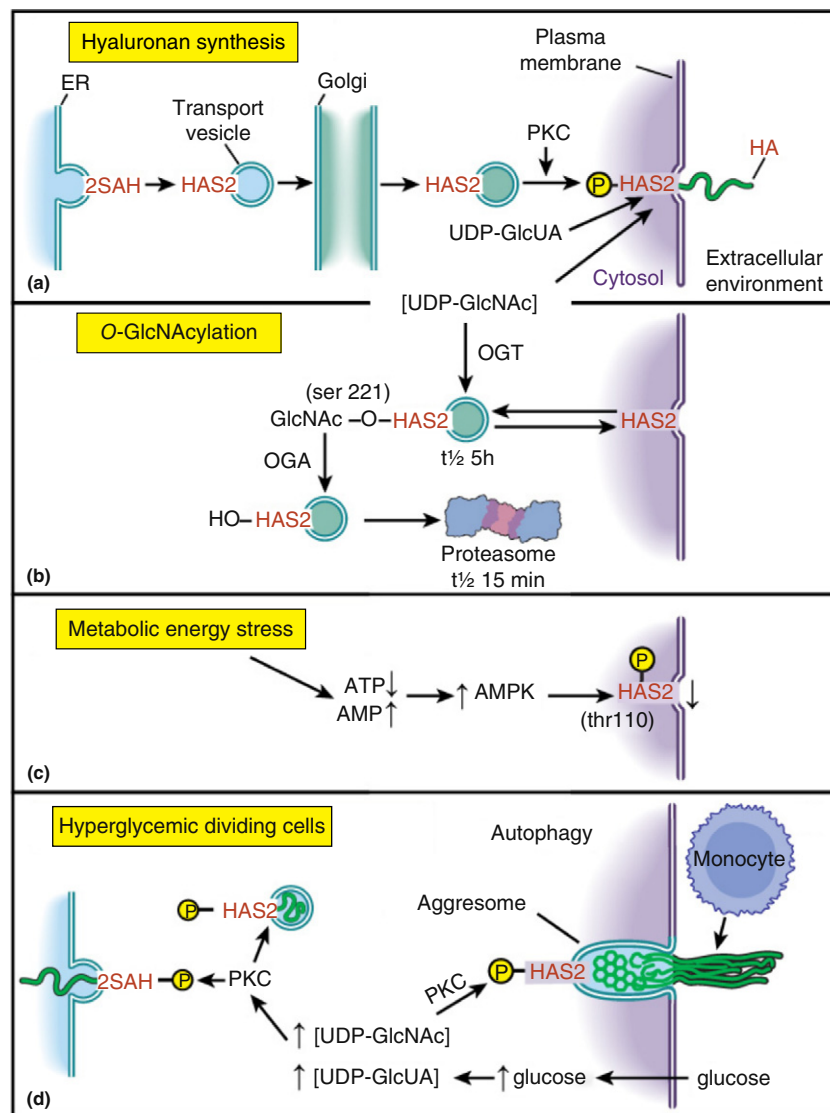


Figure 2 Models for hyaluronan synthesis. (a) Mechanism for transport of hyaluronan synthase 2 (HAS2) from the endoplasmic reticulum to the plasma membrane where it is normally activated. (b) Mechanism for regulating the half-life and maintaining inactive HAS2 in cytosolic vesicles by O-GlcNAcylation. (c) Mechanism for inactivating HAS2 by AMP kinase (AMPK) during metabolic energy stress. (d) Mechanism for activating HAS2 in intracellular compartments in hyperglycemic dividing cells. Modified from Hascall, V., Wang, A., Tammi, M., *et al.*, 2014. The dynamic metabolism of hyaluronan regulates the cytosolic concentration of UDP-GlcNAc. *Matrix Biology* 35, 14–17.

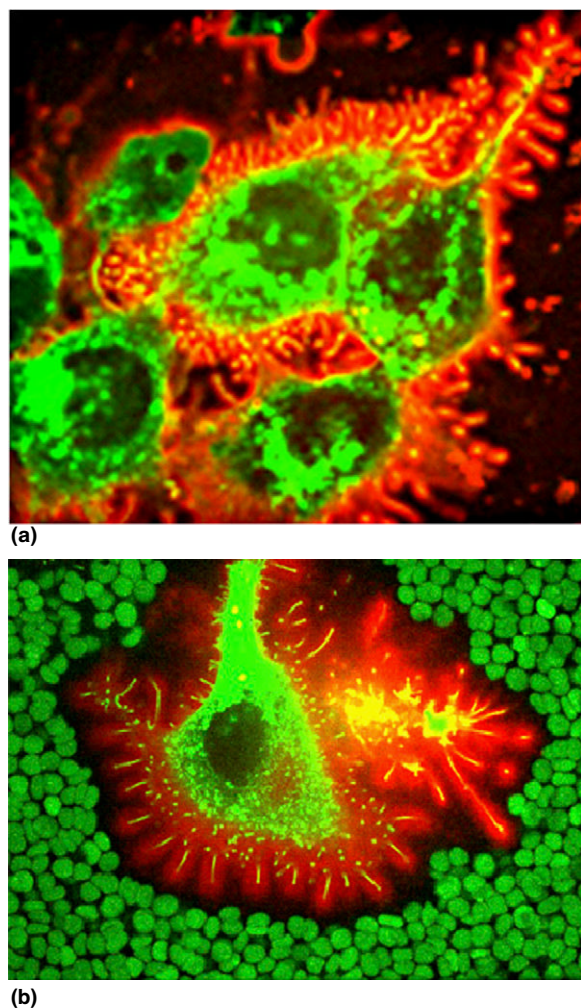


Figure 3 Hyaluronan-dependent glycoalyces. (a) GFP-hyaluronan synthase 3 was overexpressed in MCF7 breast cancer cells with formation of the hedgehog-like protrusions. Hyaluronan is stained red in the micrograph of a live cell and shows the extracellular hyaluronan matrix on the cellular protrusions. The intracellular enzyme localization is indicated in green and its active plasma membrane location in yellow. (Micrograph courtesy of the Tammi research group). (b) Exclusion assay. Visualization of the hyaluronan-dependent glycoalyx (red) and exclusion of erythrocytes (green) by confocal microscopy. Reproduced from Rilla, K., Tiihonen, R., Kultti, A., Tammi, M., Tammi, R., 2008. Pericellular hyaluronan coat visualized in live cells with a fluorescent probe is scaffolded by plasma membrane protrusions. *Journal of Histochemistry and Cytochemistry* 10, 901–910.

endosomal compartment ($t_{1/2} \sim 15$ min) that feeds them into lysosomes ($t_{1/2} \sim 3$ h) (Tammi *et al.*, 2001), and the receptor, protease, and hyaluronidase are cycled back to the cell surface.

This dynamic balance between hyaluronan synthesis and catabolism at the cellular level can regulate the cytosolic concentration of one of its substrates, UDP-GlcNAc, which is also a substrate for the intracellular O-GlcNAc transferase (OGT) (Figure 2(b)). This enzyme is critical for regulating many intracellular signaling pathways and cellular functions by adding GlcNAc to specific serine or threonine residues of target proteins, including kinases and key enzymes in metabolic pathways. If the cytosolic UDP-GlcNAc concentration is

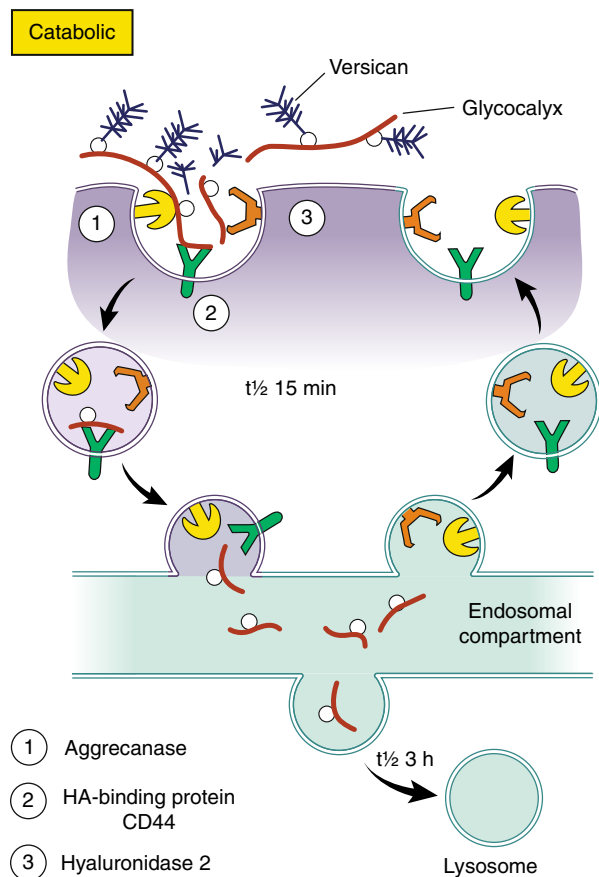


Figure 4 Model for catabolism of hyaluronan glycoalyces. Cell surface proteins that bind hyaluronan (CD44), and cleave versican (aggrecanase) and hyaluronan (GPI-anchored hyaluronidase 2), internalize fragments of hyaluronan into an endosomal compartment. The proteins recycle to the cell surface, and the fragments are shuttled to lysosomes for complete degradation. Reproduced from Hascall, V., Wang, A., Tammi, M., *et al.*, 2014. The dynamic metabolism of hyaluronan regulates the cytosolic concentration of UDP-GlcNAc. *Matrix Biology* 35, 14–17.

elevated, as occurs in diabetic hyperglycemia, OGT adds GlcNAc to a large number of additional intracellular proteins, which alters and compromises cellular functions. A recent perspective (Hascall *et al.*, 2014) describes how OGT can regulate the activity of HAS2, the only synthase required for embryonic survival, by adding GlcNAc to a key serine, and how adenine monophosphate (AMP) kinase can stop HAS2 activity by phosphorylating a key threonine during metabolic stress (Figure 2(c)). This indicates that hyaluronan metabolism is likely a rheostat for maintaining stability of the cytosolic UDP-GlcNAc concentrations within an acceptable range.

Whole-Body Turnover of Hyaluronan

The pioneering work of Laurent and Fraser in the 1980s discovered that the synthesis and degradation of hyaluronan in mammals is extremely extensive and dynamic (Laurent and Fraser, 1992). Hyaluronan is continuously synthesized by most tissues, and a small fraction, as noted above, is deposited as a pericellular surface coat, while the majority is

incorporated into tissue biomatrices. This tissue hyaluronan is predominantly turned over systemically, rather than locally, via the lymphatics and vasculature through the action of receptor-mediated endocytosis in the lymph nodes and liver, and about one-third (~5 g) of total body hyaluronan is degraded and replaced every day (Figure 5). For example, the hyaluronan in the eye vitreous fluid is degraded in lymph nodes and liver. Very large (MDa) hyaluronan fragments (containing 'piggy-backed' bound proteins and chondroitin sulfate proteoglycans) are released from tissues undergoing continuous normal partial degradation of their biomatrices (i.e., turnover; Figure 5(a)). After entering the lymphatic circulation, approximately 85% of the hyaluronan is rapidly cleared by high-affinity endocytic isoreceptors in sinusoidal endothelial cells (SECs) of lymph nodes. The remaining hyaluronan that enters blood is then removed by the same receptors in the SECs of liver. Endocytosed hyaluronan is degraded intracellularly and the component sugars are reutilized for energy production or biosynthesis (Figure 5(b)).

Based on this process, first identified by Laurent, Fraser, and coworkers, the hyaluronan scavenger receptor was purified, cloned, and named the HA receptor for endocytosis (HARE) (Weigel and Yik, 2002). It was also named Stabilin-2 after the gene was found to be related to Stabilin1 (Politz *et al.*, 2002). Native tissues express two protein isoforms of different mass (e.g., 190 kDa and 315 kDa in human) that are not splice variants; the smaller variant is the C-terminal approximately 55% of the full-length protein. The smaller receptor is designated as HARE and the full-length receptor as Stab2. HARE/Stab2 is constitutively recycling receptors that are internalized via clathrin-coated pit-mediated endocytosis. Internalized ligands are delivered to lysosomes and receptors are recycled back to the cell surface (Figure 5(b)). HARE/Stab2 are highly expressed in SECs of lymph nodes, liver, spleen, and bone marrow, and also found in macrophages, corneal and lens epithelium, mesenchymal heart valve cells, ependymal brain ventricle cells, prismatic epithelial cells covering renal papillae, and oviduct. HARE/Stab2 are also the systemic clearance receptors for apoptotic cells and other ligands including heparin, chondroitin/dermatan sulfates, acetylated low-density lipoprotein (LDL), collagen N-terminal propeptides and advanced glycation end products. HARE mediates intracellular ERK and NF- κ B signaling in response to hyaluronan uptake (Pandey and Weigel, 2014).

Hyaluronan in Structural Matrices

Hyaluronan provides the backbone to anchor proteoglycans that have hyaluronan-binding domains as indicated for versican in cell glycocalyxes (Figure 4). Cartilages were the major tissues for early proteoglycan research that provided the data for the current aggregate model for aggrecan (Figure 6), the proteoglycan that provides the resistance of cartilage to compressive loading (Hardingham, 1998). Many matrix and cell receptors that interact with hyaluronan do so by having a highly conserved approximately 93 amino acid protein domain known as a link module. The G1 domain of aggrecan contains two link modules (Day, 2001) that bind to a decasaccharide length on the hyaluronan. This binding is

reversible until it is 'locked in' by the link protein that also has two link modules and a domain that interacts with the G1 domain. HASs evolved late in evolution, likely from a chitin synthase precursor, with subsequent evolution of the 23 genes encoding link modules in the human genome. The presence of this structural domain indicates that the protein has evolved to interact with hyaluronan as part of its biological functions. For example, the cell surface receptor CD44 has one link module that normally interacts with a hexasaccharide length on hyaluronan and is involved in several functions including roles in inflammations and cancer discussed below.

Hyaluronan in Fertilization and Embryogenesis

Cumulus-oocyte complex (COC) expansion and successful ovulation requires formation of an extensive hyaluronan matrix that must be penetrated by a sperm for fertilization via a membrane hyaluronidase in the sperm head (Salustri and Fulop, 1998). The structure of this matrix requires modification of the hyaluronan by a mechanism that is also prevalent in inflammation as discussed below. Prior to ovulation, the cumulus cells upregulate the levels of HAS2, TNF α -stimulated gene 6 (TSG6), and pentraxin 3. The follicle becomes permeable to serum, which allows entry of inter- α -inhibitor (I α I), an unusual proteoglycan with one chondroitin sulfate chain. The core protein of I α I is bikunin, a trypsin inhibitor, and the chondroitin sulfate chain contains 2 proteins (heavy chains (HCs)) bound by aspartate ester linkages to the 6-OH of GalNAc residues. TSG6 is an enzyme that migrates the HCs by trans-esterification from GalNAc on I α I to the 6-OH of GlcNAc on hyaluronan forming HC-hyaluronan chains (Figure 7). The presence of pentraxin 3 stabilizes the COC hyaluronan matrix (Ievoli *et al.*, 2011). Mice null in either TSG6, or bikunin, or pentraxin 3 do not form the matrix, have aberrant ovulation, and are female infertile.

Hyaluronan is required for several other embryonic processes. Has2-null mice are embryonic lethal due to failure to form heart valves, most likely arising from deficient ErbB2-ErbB3-Ras signaling and blockade of endothelial-mesenchymal transitions that occur during early heart development (Camenisch *et al.*, 2000, 2002). Hyaluronan is also critical for development of musculoskeletal and neural tissues, as well as formation of their tissue stem cell niches, due to its binding to signal-inducing cell surface receptors such as CD44 as well as its biophysical properties and interactions with other extracellular macromolecules (Vigetti *et al.*, 2010; Sherman and Feistel, 2008; Matsumoto *et al.*, 2009).

Hyaluronan in Inflammatory Pathologies

The HC-hyaluronan structure was first identified in the 1960s in analyses of hyaluronan isolated from synovial fluids of patients with rheumatoid arthritis (Sandson *et al.*, 1965). Immunodiffusion assays showed that hyaluronidase treatment released a protein that gave precipitates with antiserum against I α I, which was thought to be a single protein at that time. The HC protein was not present on hyaluronan from synovial fluids from individuals who did not have rheumatoid arthritis. This mechanism was rediscovered in the 1990s as a

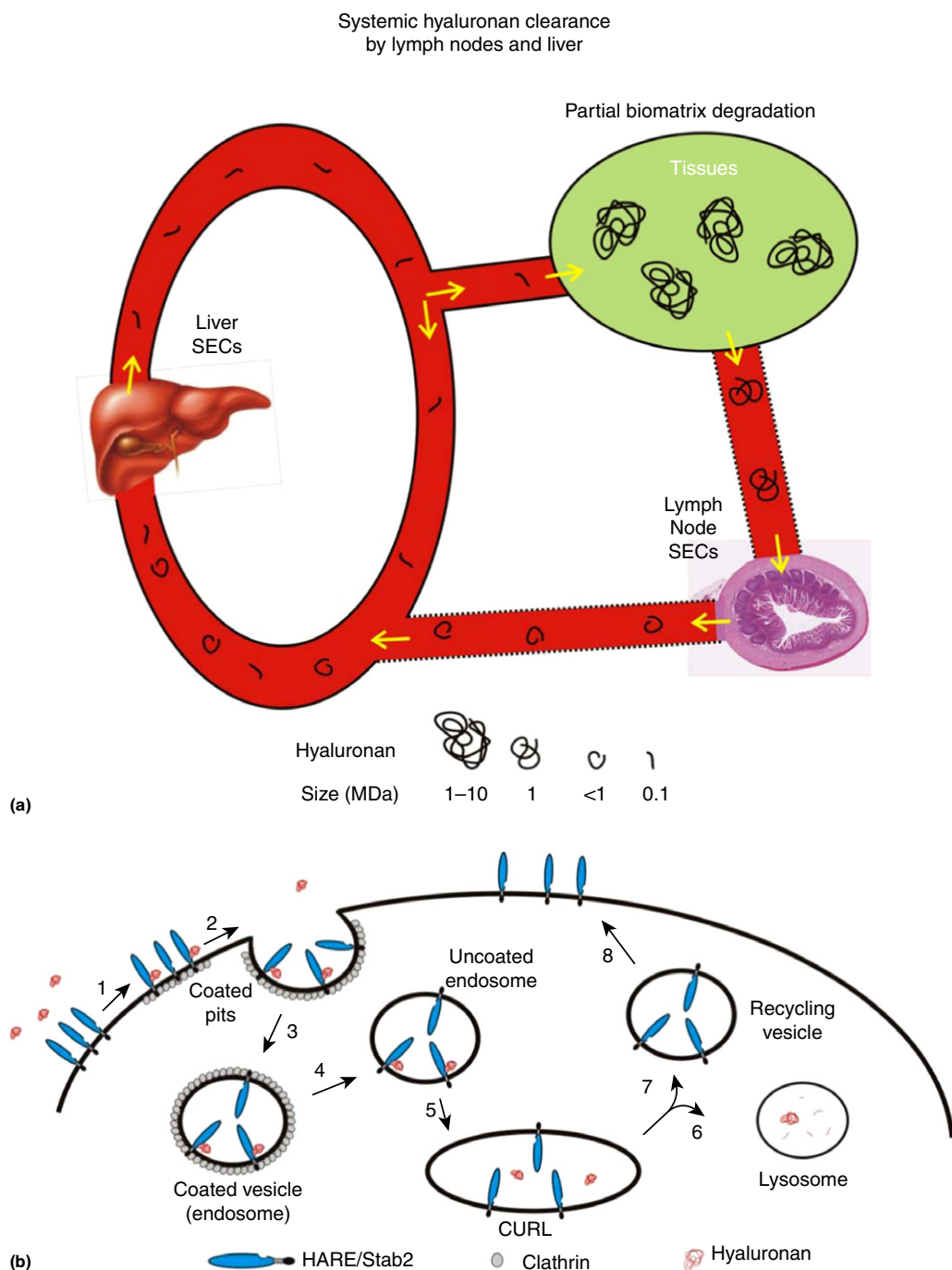


Figure 5 Systemic clearance of hyaluronan by HARE/Stab2 in lymph nodes and liver. (a) The schematic overview, based on the discoveries of Laurent and Fraser (1992), depicts the dynamic process by which mammals synthesize and degrade (i.e., turnover) about one-third of total body hyaluronan every day (~ 5 g in humans). The mass range of hyaluronan during this process is indicated (bottom) by red curved lines. All tissue biomatrices (upper right) normally undergo slow and continuous partial degradation that releases large fragments of hyaluronan (≥ 1 mDa) from native tissue hyaluronan (up to 10 MDa); the hyaluronan fragments also contain bound proteoglycans and other proteins (e.g., link protein and hyaladherins). As they flow into lymph nodes (lower right), most of these fragments are internalized and degraded by SECs. The hyaluronan that escapes clearance ($\sim 15\%$) and enters the blood (bottom) is smaller (<1 MDa) and is then internalized by the same HARE/Stab2 receptors in SECs of the liver (left). At normal steady-state in healthy people, the hyaluronan size in blood is approximately 100–150 kDa and the concentration is $<40 \mu\text{g l}^{-1}$. (b) HARE/Stab2 clearance receptors (blue) on the surface of SECs are continuously targeted to clathrin (gray) coated pits (#1), whether they are bound to hyaluronan (red coils) or not; they are constitutively endocytosed and recycled. After coated pit-mediated endocytosis (#3) and clathrin uncoating (#4), endosomes containing receptor-ligand complexes are delivered (#5) to a compartment for uncoupling receptor-ligand (CURL), in which dissociation occurs facilitated by low pH, to produce free hyaluronan that is delivered to lysosomes for degradation (#6) and free receptors, which are directed to recycling endosomes (#7). These vesicles return HARE/Stab2 to the plasma membrane (#8) where they begin another cycle of internalization. Reproduced from Weigel, P., Yik, J., 2002. Glycans as endocytosis signals: The cases of the asialoglycoprotein and hyaluronan/chondroitin sulfate receptors. *Biochimica et Biophysica Acta* 1572, 341–363.

'serum-derived hyaluronan-associated protein' (SHAP) that modified hyaluronan in fibroblast cultures, and was shown to be the heavy chains from IαI (Kimata and Zhuo, 2001).

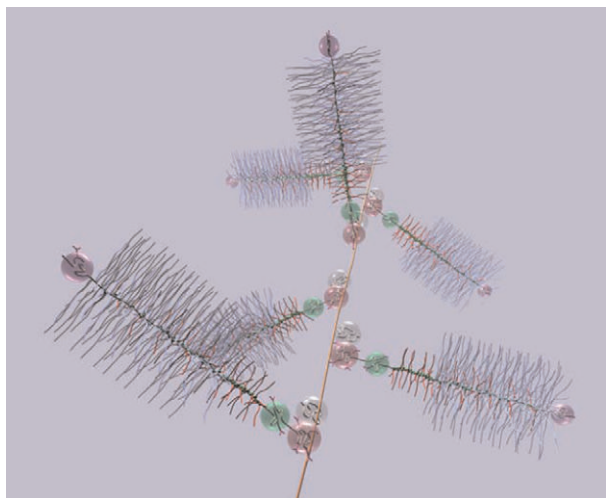


Figure 6 Model for aggregates of the proteoglycan aggrecan. The G1 domain of aggrecan (pink globe) and the associated link protein (white globe) are anchored on the central hyaluronan strand. Aggrecan contains two additional globular domains (G2, green globe and G3, purple globe) and has approximately 100 chondroitin sulfate chains (blue lines) and approximately 50 keratan sulfate chains (red lines).

In binding assays, lymphocytes bound to HC-hyaluronan coated surfaces >50-fold more than to equivalent concentrations of hyaluronan coated surfaces, indicating that this covalent protein modification of hyaluronan matrices has a significant role in inflammatory pathologies.

Experiments during this time also showed that colon smooth muscle cells treated with a virus or poly I:C, a viral mimetic, synthesized extensive structures referred to as hyaluronan cables. Monocytes bound to these structures at 4 °C and rapidly degraded them at 37 °C with capping of CD44 over phagocytosed hyaluronan fragments (Strong and de la Motte, 1999; Wang *et al.*, 2011) (Figure 8). The importance of CD44 for this process was shown in CD44-null mice treated with bleomycin, a noxious agent that induces inflammation in the lung (Teder *et al.*, 2002). Hyaluronan matrices increased significantly, and monocytes/macrophages extravasated into the inflamed areas. In the wild type mice, the hyaluronan matrix was removed and the inflammatory cells disappeared with restoration of normal lung function. In contrast, the monocytes continued to enter the lung connective tissue in the CD44-null mice but were unable to remove the hyaluronan matrix, and the mice died.

Almost all inflammatory pathologies, from wound healing to atherosclerosis to inflammatory bowel disease, involve early synthesis of the monocyte-adhesive matrix by the stressed cells (Strong and de la Motte, 1999; Wight, 1999). This matrix defines the location of the pathology and initiates the

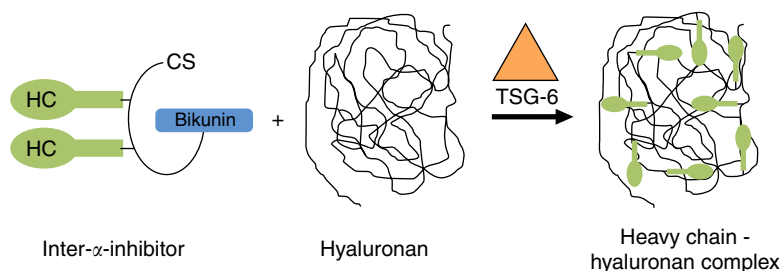


Figure 7 Model for heavy chain (HC) transfer from inter- α -trypsin inhibitor (I α I) to hyaluronan by TNF α -stimulated gene 6 (TSG6). TSG6 (orange triangle) transfers HCs (green) from the chondroitin sulfate chain of I α I to hyaluronan in many inflammations and during hyaluronan matrix formation in cumulus cell-oocyte complex expansion and ovulation.

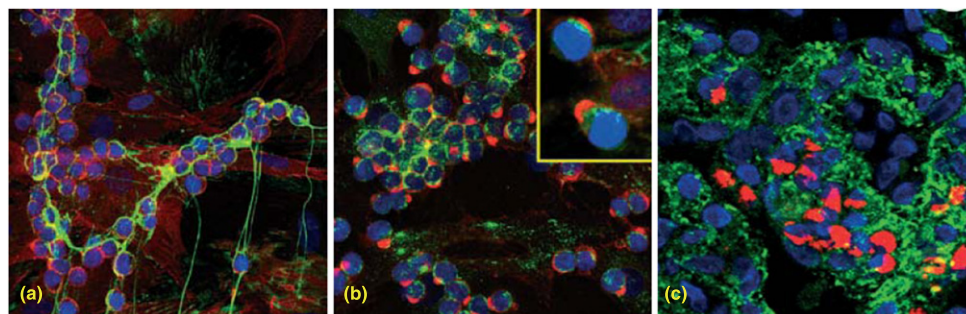


Figure 8 Hyaluronan cable structure and phagocytosis. (a) Monocytes (U937 cells) bind to a hyaluronan cable structure at 4 °C on colon smooth muscle cells treated with a viral mimetic poly I:C. Hyaluronan is stained green and CD44 is red. The monocytes do not interact with the hyaluronan coat on the smooth muscle cells in the background. (b) Monocytes phagocytose the hyaluronan matrix when cells as in (a) are warmed to 37 °C for 15 min. The inset shows the capped CD44 receptors with internalized hyaluronan. (c) A micrograph of an asthmatic lung biopsy with CD44-capped monocytes embedded in the pathological hyaluronan matrix. Modified from Wang, A., de la Motte, C., Lauer, M., Hascall, V., 2011. Hyaluronan matrices in pathobiological processes. *FEBS Journal* 278, 1412–1418.

sequence of events that follows – the influx of neutrophils (wound healing) or eosinophils (asthma) followed by the monocytes/macrophages, the removal of the hyaluronan matrix, and subsequent reconstruction. This process can restore the matrix to give normal tissue or fibrotic tissue, as in scar formation after wound healing.

Hyaluronan in Diabetic Pathologies

An experiment with a surprising result showed that vascular cells dividing in the medium commonly used for their culture induced hyaluronan accumulation intracellularly (Evanko and Wight, 2001). This presented a conundrum described in a subsequent review (Hascall *et al.*, 2004). However, the medium used was hyperglycemic, approximately 5x normal. Subsequent experiments with mesangial cells (Ren *et al.*, 2009) and bone marrow stromal (stem) cells (Wang *et al.*, 2014) showed that cells dividing in medium with more than 3x normal glucose levels activate HASs in intracellular compartments, including endoplasmic reticulum and transport vesicles (Figure 2(d)). This induced an autophagy response and a cyclin D3-mediated extrusion of hyaluronan to form a monocyte-adhesive hyaluronan matrix at the end of cell division. This mechanism compromised mesangial glomerular function in streptozotocin diabetic rats with development of proteinuria and nephropathy. The hyperglycemic bone marrow stromal cells differentiated into pathological adipocytes *in vitro* and *in vivo* with extensive formation of the monocyte-adhesive hyaluronan matrix that recruited large numbers of inflammatory cells *in vivo* with subsequent loss of mineral in trabecular bone. The diversion of stem cells that divide in hyperglycemia to this pathological adipogenesis pathway is highly likely to be a major mechanism underlying the ongoing diabetic epidemic with obesity and elevated risks of inflammatory pathologies.

Hyaluronan in Cancer

Hyaluronan deposition around tumor cells and in the tumor stroma correlates with progression of many types of human cancer (Tammi *et al.*, 2008). Correspondingly, a large amount of experimental data supports a pro-tumorigenic role for hyaluronan (Toole, 2009). However, the role of hyaluronan is complex, since experimental over-expression of hyaluronan in tumor cells can be either pro- or anti-tumorigenic, depending on the context. One contributing factor is hyaluronan turnover (Simpson and Lokeshwar, 2008), but other factors, such as molecular size of hyaluronan, are also likely to be important (Tian *et al.*, 2013). Irrespective of these complexities, it is clear that hyaluronan has a critical role in cancer. In addition to expected structural contributions to the tumor stroma, hyaluronan interacts with cell surface receptors that collaborate in driving intracellular signaling pathways and transporter activities that promote tumor progression.

Pericellular hyaluronan interacts with several cell surface receptors, including CD44, RHAMM, LYVE-1, HARE/stabilin-2, and Toll-like receptors-2 and 4 (Day, 2001). The major hyaluronan receptor implicated in cancer is CD44. However, RHAMM and CD44 can exhibit both cooperative and

interchangeable signaling functions. For example, interactions at the plasma membrane between RHAMM and CD44 have been shown to activate CD44 signaling through ERK1/2 and promote cancer cell motility (Maxwell *et al.*, 2008). In some cases, such as animal models of autoimmune diseases, RHAMM can compensate for CD44, a very important consideration when interpreting experiments in CD44-null mice (Naor *et al.*, 2007). LYVE-1 has been shown to have a similar role to CD44 in virally infected lymphoma cells (Qin *et al.*, 2011). HARE/Stab2 can also serve as homing receptors for hyaluronan-mediated tumor cell metastasis, since they are most prominently expressed in liver, lymph nodes, and bone marrow, all of which tissues are frequently colonized by metastatic tumor cells of epithelial origin bearing surface hyaluronan coats. Blocking hyaluronan binding to HARE in mice implanted with human prostate tumor cells almost completely prevents metastasis to lymph nodes (Simpson *et al.*, 2012).

CD44 is a single chain, single-pass, transmembrane glycoprotein that is very widely expressed in physiological and pathological systems. CD44 was first characterized through the confluence of several areas of investigation, including hyaluronan–cell interactions, lymphocyte homing, and cell adhesion (Toole, 1990; Knudson and Knudson, 2005). Although CD44 arises from a single gene, different transcripts are formed by alternative splicing. In several types of cancer cells, binding of hyaluronan to CD44 results in direct or indirect interaction of CD44 with signaling receptors, such as ErbB2, EGFR, and TGF- β receptor type I, and influences the activity of these receptors (Figure 9). It can also lead to interaction with and altered activity of non-receptor kinases of the Src family or Ras family GTPases (Bourguignon, 2008). Thus, hyaluronan–CD44 binding influences the activity of a variety of downstream signaling pathways, especially the MAP kinase and PI3 kinase/Akt pathways. In addition, binding of hyaluronan to CD44 stimulates multidrug and metabolic transporters that are important in malignancy and therapy resistance, and the cytosolic domain of CD44 interacts with the actin cytoskeleton via ankyrin or members of the ezrin-radixin-moesin family, consequently promoting tumor cell proliferation, survival, motility, invasiveness, and chemoresistance (Toole, 2009; Bourguignon, 2008; Ponta *et al.*, 2003). CD44 is often localized at least in part to lipid microdomains with the properties of lipid rafts, and associates indirectly or directly therein with signaling proteins and transporters (Toole, 2009). Figure 9 emphasizes cell autonomous activities influenced by hyaluronan produced by tumor cells. Hyaluronan produced by stromal cells may have overlapping or different activities, and interactions between the two compartments very likely contribute further to the malignant properties of tumors (Knudson *et al.*, 1984; Itano and Kimata, 2008). Given the large number of pathways affected by CD44, it is likely that indirect and direct interactions with a wide variety of effectors within such domains can be induced and/or stabilized by multivalent interactions of hyaluronan with CD44, which provides a likely mechanism whereby hyaluronan–CD44 interactions have such broad effects (Figure 9). Importantly, antagonists of these interactions that result in attenuation of pro-tumorigenic or therapy-resistant activities provide promising new therapeutic approaches in cancer patients (Toole, 2009; Misra *et al.*, 2009).

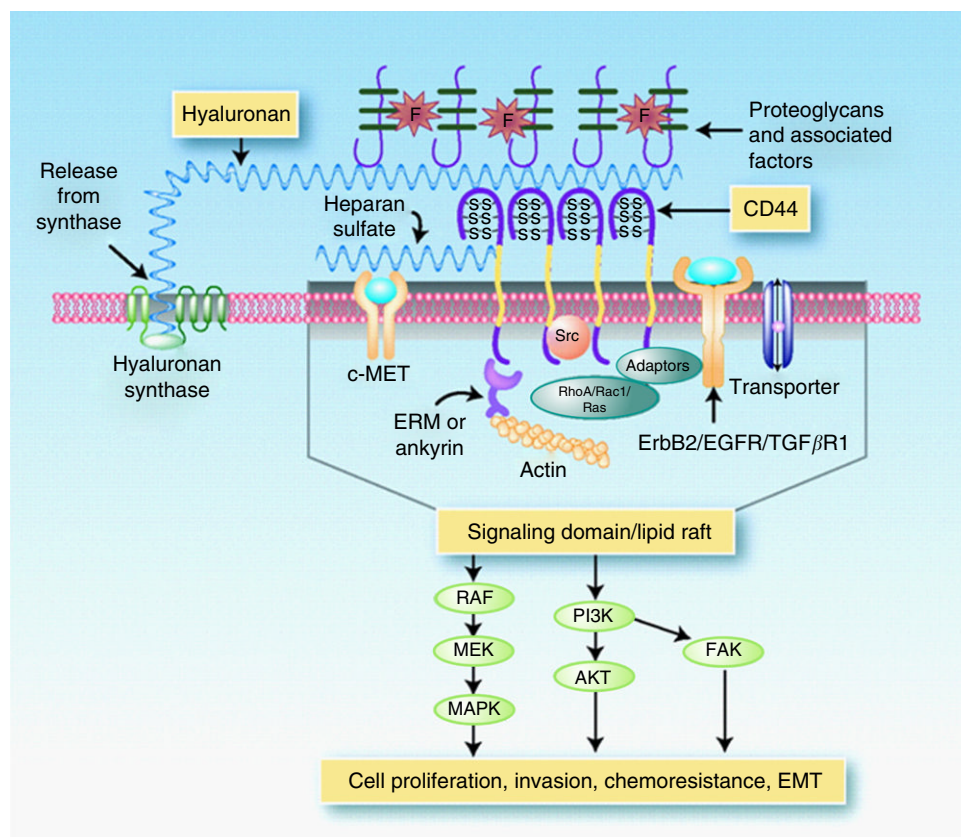


Figure 9 Regulation of signaling cascades in cancer cells by hyaluronan-CD44 interactions. Hyaluronan synthases produce and extrude hyaluronan, which interacts multivalently with CD44 to induce and/or stabilize signaling domains within the plasma membrane. These signaling domains contain various signaling receptors and non-receptor kinases that drive cell proliferation and survival pathways, as well as various transporters that participate in drug resistance and altered metabolic properties in cancer cells. Moreover, hyaluronan-CD44 interactions induce cytoskeletal changes that promote cell motility and invasion. Proteoglycans and associated factors attached to pericellular hyaluronan may also influence these activities. Reproduced from Toole, B.P., 2009. Hyaluronan-CD44 interactions in cancer: Paradoxes and possibilities. *Clinical Cancer Research* 15, 7462–7468.

Conclusions

This article provides an overview of the dynamic biochemistry and biology of hyaluronan. Research on this intriguing macromolecule continues at a rapid pace and is likely to discover yet more of its intriguing functions. The article has not addressed its expanding use in clinical and therapeutic treatments, regenerative medicine or as a material to form matrices that can incorporate other molecules or cells. These important areas are expanding steadily and would be appropriate for a future hyaluronan chapter.

References

- Bourguignon, L.Y., 2008. Hyaluronan-mediated CD44 activation of RhoGTPase signaling and cytoskeleton function promotes tumor progression. *Seminars in Cancer Biology* 18, 251–259.
- Camenisch, T.D., Schroeder, J.A., Bradley, J., Klewer, S.E., McDonald, J.A., 2002. Heart-valve mesenchyme formation is dependent on hyaluronan-augmented activation of ErbB2-ErbB3 receptors. *Nature Medicine* 8, 850–855.
- Camenisch, T.D., Spicer, A.P., Brehm-Gibson, T., *et al.*, 2000. Disruption of hyaluronan synthase-2 abrogates normal cardiac morphogenesis and hyaluronan-mediated transformation of epithelium to mesenchyme. *Journal of Clinical Investigation* 106, 349–360.
- Day, A.J., 2001. Understanding hyaluronan-protein interactions. *Hyaluronan Today*, article 16.
- Evanko, S., Wight, T., 2001. Intracellular hyaluronan. *Hyaluronan Today*, article 20.
- Hardingham, T., 1998. Cartilage: Aggrecan – link protein – hyaluronan aggregates. *Hyaluronan Today*, article 5.
- Hascall, V., Laurent, T., 1997. Hyaluronan: Structure and physical properties. *Hyaluronan Today*, article 1.
- Hascall, V., Majors, A., de la Motte, C., *et al.*, 2004. Intracellular hyaluronan: A new frontier for inflammation? *Biochimica et Biophysica Acta* 1672, 3–12.
- Hascall, V., Wang, A., Tammi, M., *et al.*, 2014. The dynamic metabolism of hyaluronan regulates the cytosolic concentration of UDP-GlcNAc. *Matrix Biology* 35, 14–17.
- Hubbard, C., McNamara, J.T., Azumaya, C., Patel, M.S., Zimmer, J., 2012. The hyaluronan synthase catalyzes the synthesis and membrane translocation of hyaluronan. *Journal of Molecular Biology* 418, 21–31.
- Ievoli, E., Lindstedt, R., Inforzato, A., *et al.*, 2011. Implication of the oligomeric state of the N-terminal PTX3 domain in cumulus matrix assembly. *Matrix Biology* 30, 330–337.
- Itano, N., Kimata, K., 2008. Altered hyaluronan biosynthesis in cancer progression. *Seminars in Cancer Biology* 18, 268–274.
- Kimata, K., Zhuo, L., 2001. SHAP, a protein covalently bound to hyaluronan. *Hyaluronan Today*, article 22.
- Knudson, W., Biswas, C., Toole, B.P., 1984. Interactions between human tumor cells and fibroblasts stimulate hyaluronate synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 81, 6767–6771.

- Knudson, W., Knudson, C.B., 2005. The hyaluronan receptor, CD44. *Hyaluronan Today*, article 10.
- Kultti, A., Rilla, K., Tiihonen, R., *et al.*, 2006. Hyaluronan synthesis induces microvillus-like cell surface protrusions. *Journal of Biological Chemistry* 281, 15821–15828.
- Laurent, T.C., 2002. 'The Tree': Hyaluronan research in the twentieth century. *Hyaluronan Today*, article 23.
- Laurent, T.C., Fraser, J.R.E., 1992. Hyaluronan. *FASEB Journal* 6, 2397–2404.
- Matsumoto, K., Li, Y., Yakuba, C., *et al.*, 2009. Conditional inactivation of Has2 reveals a crucial role for hyaluronan in skeletal growth, patterning, chondrocyte maturation and joint formation in the developing limb. *Development* 136, 2825–2835.
- Maxwell, C.A., McCarthy, J., Turley, E., 2008. Cell-surface and mitotic-spindle RHAMM: Moonlighting or dual oncogenic functions? *Journal of Cell Science* 121, 925–932.
- Meyer, K., Palmer, J.W., 1934. The polysaccharide of the vitreous humor. *Journal of Biological Chemistry* 107, 629–634.
- Misra, S., Hascall, V.C., Giovanni, C.D., Markwald, R.R., Ghatak, S., 2009. Delivery of CD44 shRNA/nanoparticles within cancer cells: Perturbation of hyaluronan/CD44v6 interactions and reduction in adenoma growth in Apc Min/+ mice. *Journal of Biological Chemistry* 284, 12432–12446.
- Naor, D., Nedvetzki, S., Walmsley, M., *et al.*, 2007. CD44 involvement in autoimmune inflammations: The lesson to be learned from CD44-targeting by antibody or from knockout mice. *Annals of the New York Academy of Sciences* 1110, 233–247.
- Pandey, M.S., Weigel, P.H., 2014. Hyaluronic acid receptor for endocytosis (HARE)-mediated endocytosis of hyaluronan, heparin, dermatan sulfate, and acetylated low density lipoprotein (AcLDL), but not chondroitin sulfate types A, C, D, or E, activates NF-kappaB-regulated gene expression. *Journal of Biological Chemistry* 289, 1756–1767.
- Politz, O., Gratchev, A., McCourt, P.A.G., *et al.*, 2002. Stabilin-1 and-2 constitute a novel family of fasciclin-like hyaluronan receptor homologues. *Biochemical Journal* 362, 155–164.
- Ponta, H., Sherman, L., Herrlich, P., 2003. CD44: From adhesion molecules to signaling regulators. *Nature Reviews Molecular Cell Biology* 4, 33–45.
- Qin, Z., Dai, L., Bratoveva, M., *et al.*, 2011. Cooperative roles for emmprin and LYVE-1 in the regulation of chemoresistance for primary effusion lymphoma. *Leukemia* 25, 1598–1609.
- Ren, J., Hascall, V., Wang, A., 2009. Cyclin D3 mediates synthesis of a hyaluronan matrix that is adhesive for monocytes in mesangial cells stimulated to divide in hyperglycemic medium. *Journal of Biological Chemistry* 284, 16621–16632.
- Rilla, K., Tiihonen, R., Kultti, A., Tammi, M., Tammi, R., 2008. Pericellular hyaluronan coat visualized in live cells with a fluorescent probe is scaffolded by plasma membrane protrusions. *Journal of Histochemistry and Cytochemistry* 10, 901–910.
- Rilla, R., Tammi, M., Tammi, R., 2010. The enigma of hyaluronan-synthesizing membrane protrusions and hyaluronan coats. *Hyaluronan Today*, article 36.
- Salustri, A., Fulop, C., 1998. Role of hyaluronan during ovulation and fertilization. *Hyaluronan Today*, article 3.
- Sandson, J., Hamerman, D., Schwick, G., 1965. Altered properties of pathological hyaluronate due to a bound inter-alpha trypsin inhibitor. *Transactions of the Association of American Physicians* 78, 304–313.
- Sherman, L.S., Feistel, K., 2008. Hyaluronan as a regulatory component of neural stem cell and progenitor cell niches. *Hyaluronan Today*, article 31.
- Simpson, M.A., Lokeshwar, V.B., 2008. Hyaluronan and hyaluronidase in genitourinary tumors. *Frontiers in Bioscience* 13, 5664–5680.
- Simpson, M.A., Weigel, J.A., Weigel, P.H., 2012. Systemic blockade of the hyaluronan receptor for endocytosis prevents lymph node metastasis of prostate cancer. *International Journal of Cancer* 131, E836–E840.
- Spicer, A.P., McDonald, J.A., 1998. Eukaryotic hyaluronan synthases. *Hyaluronan Today*, article 7.
- Strong, S., de la Motte, C., 1999. Hyaluronan and colon smooth muscle cells. *Hyaluronan Today*, article 12.
- Tammi, R., Rilla, K., Pienimäki, J., *et al.*, 2001. Hyaluronan enters keratinocytes by a novel endocytic route for catabolism. *Journal of Biological Chemistry* 276, 35111–35122.
- Tammi, R.H., Kultti, A., Kosma, V.M., *et al.*, 2008. Hyaluronan in human tumors: Pathobiological and prognostic messages from cell-associated and stromal hyaluronan. *Seminars in Cancer Biology* 18, 288–295.
- Teder, P., Vandivier, R.W., Jiang, D., *et al.*, 2002. Resolution of lung inflammation by CD44. *Science* 296, 155–158.
- Tian, X., Azpurua, J., Hine, C., *et al.*, 2013. High-molecular-mass hyaluronan mediates the cancer resistance of the naked mole rat. *Nature* 499, 346–349.
- Tlapak-Simmons, V.L., Baggenstoss, B.A., Clyne, T., Weigel, P.H., 1999. Purification and lipid dependence of the recombinant hyaluronan synthases from *Streptococcus pyogenes* and *Streptococcus equisimilis*. *Journal of Biological Chemistry* 274, 4239–4245.
- Toole, B.P., 1990. Hyaluronan and its binding proteins, the hyaladherins. *Current Opinion in Cell Biology* 2, 839–844.
- Toole, B.P., 2009. Hyaluronan-CD44 interactions in cancer: Paradoxes and possibilities. *Clinical Cancer Research* 15, 7462–7468.
- Vigetti, D., Ori, M., Passi, A., 2010. The *Xenopus* model for evaluating hyaluronan during development. *Hyaluronan Today*, article 35.
- Wang, A., de la Motte, C., Lauer, M., Hascall, V., 2011. Hyaluronan matrices in pathobiological processes. *FEBS Journal* 278, 1412–1418.
- Wang, A., Midura, R., Vasanji, A., Wang, A.J., Hascall, V., 2014. Hyperglycemia diverts dividing osteoblastic precursor cells to an adipogenic pathway and induces synthesis of a hyaluronan matrix that is adhesive for monocytes. *Journal of Biological Chemistry* 289, 11410–11420.
- Weigel, P.H., 2005. Bacterial Hyaluronan Synthases. 1998, Update *Hyaluronan Today*, article 6.
- Weigel, P.H., DeAngelis, P.L., 2007. Minireview — Hyaluronan synthases: A decade-plus of novel glycosyltransferases. *Journal of Biological Chemistry* 282, 36777–36781.
- Weigel, P.H., Yik, J.H., 2002. Glycans as endocytosis signals: The cases of the asialoglycoprotein and hyaluronan/chondroitin sulfate receptors. *Biochimica et Biophysica Acta — General Subjects* 1572, 341–363.
- Weissman, B., Meyer, K., 1954. The structure of hyalobiuronic acid and of hyaluronic acid from umbilical cord. *Journal of the American Chemical Society* 76, 1753–1757.
- Wight, T., 1999. Hyaluronan in atherosclerosis and restenosis. *Hyaluronan Today*, article 9.

Relevant Websites

<http://www.glycoforum.gr.jp/science/hyaluronan/hyaluronanE.html>
Glycoforum.

<http://ishas.org/>
International Society for Hyaluronan Sciences.

www.matrixbiologyinstitute.org
Matrix Biology Institute.

MOLECULAR PRINCIPLES, COMPONENTS, TECHNOLOGY, AND CONCEPTS: METABOLISM

Contents

Metabolic Regulation

A Structure Perspective on Organelle Bioenergetics

Metabolic Regulation

MS Patel, University at Buffalo, The State University of New York, Buffalo, NY, USA

RA Harris, Indiana University School of Medicine, Indianapolis, IN, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Acetyl-CoA/CoA ratio An index of the acetylation potential of the cell. Set at different ratios in the cytosol and mitochondrial matrix space and varies with nutritional state. It is usually determined in the mitochondrial matrix space by the relative rates of fatty acid oxidation and citric acid cycle.

Allosteric control A control mechanism by which a small molecule affects one or more of the kinetic parameters of an enzyme by binding to a site on the protein other than the active site. Positive allosteric effectors increase enzyme activity. Negative allosteric effectors decrease enzyme activity.

Cellular compartments Spaces enclosed by biological membranes create cellular compartments. These include the cytosol which is enclosed by the plasma membrane and the mitochondrial matrix space which is enclosed by the mitochondrial inner membrane. These compartments serve as a way to separate metabolic pathways that involve common intermediates, for example, fatty acid oxidation in the mitochondrial matrix space and fatty acid synthesis in the cytosol.

Covalent modification A control mechanism by which covalent modification of the side chain of an amino acid residue of an enzyme affects the kinetic parameters of that enzyme. Depending upon the enzyme and the modification, the effect can have a positive or negative effect on enzyme activity. Common covalent modifications include phosphorylation, acetylation, acylation, succinylation, and methylation.

Energy charge An index of the energy status of cells. It is defined as $([ATP] + \frac{1}{2} [ADP]) / ([ATP] + [ADP] + [AMP])$. Energy charge values range from 0 (all AMP) to 1.0 (all ATP). Cells maintain the energy charge within the narrow range of 0.8 to 0.95 by modifying rates of ATP producing and ATP utilizing reactions.

Isozymes or isoenzymes Enzymes that are different gene products but catalyze the same reaction, for example, HK 1, 2, 3, and 4 all catalyze the reaction of glucose + ATP yields G6P + ADP. Isozymes invariably differ with respect to their kinetic characteristics (K_m , V_{max}) and regulatory mechanisms (sensitivity to allosteric effectors, whether subject to covalent modification, whether regulated by gene transcription).

Malate/aspartate shuttle Cyclic pathway that transports reducing equivalents in the form of NADH from the cytosol into the mitochondrial matrix space (mitosol). The sum reaction for the cycle is $NADH_{cytosol} + NAD^+_{mitosol}$ yields $NAD^+_{cytosol} + NADH_{mitosol}$.

NAD^+ /NADH ratio An index of the redox state of the cell. Set at different ratios in the cytosol and mitochondrial matrix space and varies with the nutritional state. It is usually determined in the mitochondrial matrix space by the relative rates of fatty acid oxidation and the electron transport chain.

Oxidative phosphorylation The process by which the energy requiring reaction of ATP synthesis from ADP and Pi is coupled to the energy producing pathway of electron transport.

Phosphorylation potential An index of the energy status of cells. It is calculated from the equation $[ATP]/[ADP][Pi]$ which derives from the free energy change for ATP hydrolysis. Accurate determination requires estimation of the free concentrations of the nucleotides and inorganic phosphate. Cells maintain very high phosphorylation potentials to drive high energy requiring anabolic processes.

Tricarboxylic acid (TCA) cycle, Krebs cycle, citric acid cycle Cycle that oxidizes acetyl-CoA to CO_2 with the production of NADH and FADH₂ which in turn are oxidized by the electron transport chain with the production of ATP by oxidative phosphorylation. It is named after Hans Krebs who discovered the citric acid cycle.

Introduction

Metabolic regulation is a term used to describe the process by which metabolic pathways (both the anabolic/biosynthetic and catabolic/degradative pathways) are regulated in mammals. Living organisms need to generate energy continuously to maintain cellular processes and functions. The ability to oxidize available substrates (termed as fuels) to maintain energy needs (energy homeostasis) is central to survival of an organism. In mammals a near constant level of blood glucose (glucose homeostasis) is maintained to supply this fuel for energy production by the brain and other tissues. Hence, the maintenance of energy homeostasis and glucose homeostasis is critical for the function and survival of mammals during the fed as well as the fasting states (due to intermittent consumption of dietary fuels). Insulin and glucagon produce opposite effects upon metabolic processes. Caloric homeostasis in the fed and the fasted states depends upon a continuous monitoring and adjustment of the blood concentrations of insulin and glucagon. A change in the flux in a metabolic pathway is achieved by modulating the activity of one or more key enzymes (regulatory enzymes) which are subject to a variety of mechanisms to control their activities. Thus, by modulating the flux through various metabolic pathways to meet the metabolic needs of different organs, the body is able to maintain both its energy homeostasis and glucose homeostasis by utilizing the available fuels either from the dietary sources or from the internal tissue deposits.

The purpose of this article is to provide a summary of well-established principles of metabolic regulation in cells and tissues of animals. Rather than an exhaustive description of what is known about the regulation of a particular enzyme or a metabolic pathway, we discuss metabolic regulation in the context of physiological conditions that will be of interest to most readers. In our experience as investigators in this field as well as teachers of medical biochemistry, we have found that the remarkable ways in which the liver is able to change from an organ that synthesizes fuels for storage in the fed state to an organ that provides fuels for the rest of the body is a good way to gain an appreciation of the different mechanisms by which metabolic pathways are regulated (Figure 1).

Mechanisms for Regulation of Key Enzymes

Before we discuss the alterations in the pathways involved in glucose metabolism in the liver during the fed and fasted states, the mechanisms involved in controlling the rates of key regulatory enzymes need to be discussed. Although a change in the overall flux in a given pathway determines the contribution of that pathway in the metabolism in tissues, the control of this change in the flux is achieved by regulating key enzymes in the given pathway (usually more than one). This control of key enzymes can be achieved by five different mechanisms.

Substrates (Fuel) Availability

In the fed state, markedly increased availability of the circulating fuels (glucose, amino acids, and triacylglycerols) along

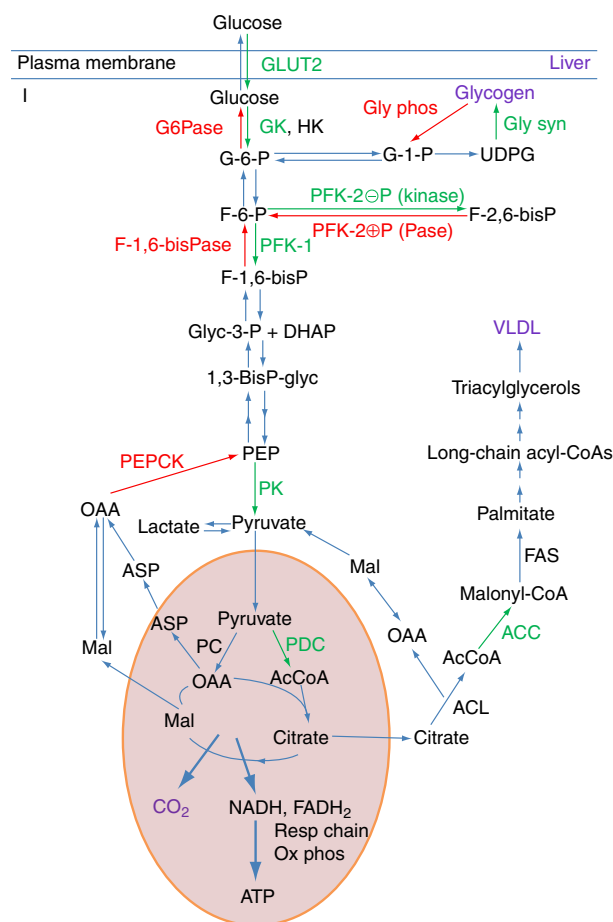


Figure 1 The pathways involved in glucose metabolism in the liver during the fed and fasted states. The major pathways depicted here are: glycolysis, glycogen synthesis and degradation, the tricarboxylic acid cycle, lipogenesis, and gluconeogenesis. These pathways are abbreviated to highlight the key regulatory enzymes. The key enzymes and the transporter in the biosynthetic pathways in the liver during the fed state are shown in green color and the key enzymes in the degradative pathways operating in the fasted state are in red color. Enzyme abbreviation used are: HK, hexokinase; GK, glucokinase; GLUT2, glucose transporter 2; Gly syn, glycogen synthase; Gly phos, glycogen phosphorylase; PFK-1, phosphofructokinase-1; PFK-2, phosphofructokinase-2; F-1,6-bisPase, fructose-1,6-bisphosphatase; PK, pyruvate kinase; PDC, pyruvate dehydrogenase complex; PC, pyruvate carboxylase; ACL, ATP-citrate lyase; ACC, acetyl-CoA carboxylase; FAS, fatty acid synthase. Abbreviations used for intermediates are commonly used in the metabolic pathways.

with the increased levels of circulating insulin set the stage for anabolic metabolism in all tissues. In the fasted state, a change in the levels of circulating hormones favoring the action of glucagon signals for mobilization (as glucose, free fatty acid, glycerol, lactate, pyruvate, and amino acids) of the stored fuels in the forms of glycogen in the liver, triacylglycerols (TAGs) in adipose tissues and proteins in skeletal muscle. Several enzymes and transporters have high K_m s for their substrates and hence they operate mainly when the concentration of the required substrates is increased (e.g., hepatic glucokinase, an isoenzyme of hexokinase, when the portal level of glucose is markedly increased during absorption after a mixed meal).

Allosteric Modulation

A change in enzyme activity by an allosteric mechanism brought about by a conformational change in the enzyme by binding of a small molecule in the site other than the active site. This type of enzyme is generally a dimer or higher order of organization. The change in activity (activation or inhibition) occurs as rapidly as the level of an allosteric compound changes in the cell and hence is referred to as short-term control.

Covalent Modification (or Posttranslational Modification)

A covalent modification refers to attachment/removal of a group (such as phosphoryl, acetyl, etc.) to/from a specific amino acid residue on an enzyme/protein. This change results in either activation or inactivation of enzyme activity. A change in enzyme activity by a covalent modification (e.g., phosphorylation of a serine residue on an enzyme by a post-translational modification) is rapid (short-term control) with often all-or-none activity of a modified enzyme molecule. This is probably the most effective way to regulate enzyme activity.

Interconversion of a key regulatory enzyme between its two forms (active and inactive) by covalent modification imparting different kinetic and allosteric properties often provides not only a unique control for its regulation but also by its interactions with a group of key enzymes involved in the biosynthetic and degradative pathways involved in a given metabolic process (e.g., protein-protein interactions among phosphorylase kinase, glycogen phosphorylase, and glycogen synthase in glycogen metabolism).

Inhibitory Protein (Regulatory Protein) Interaction

Binding of an inhibitory protein (aka regulatory protein) to an enzyme interferes with enzyme action. For example, binding of glucokinase inhibitory protein to glucokinase inhibits its activity. Similarly, binding of a specific inhibitory protein (aka regulatory protein) to protein kinase A (catalytic protein) inactivates protein kinase A. When cAMP binds to the inhibitory protein, it dissociates from protein kinase A making the latter an active enzyme.

Transcriptional and Degradational Control

A change (either increase or decrease) in the transcriptional rate of a gene by its transcriptional induction or repression eventually results in an alteration in the level of enzyme content (actual number of enzyme molecules) per cell. This is a relatively slow response requiring hours if not days for a significant response (and hence referred to as a long-term control). This control mechanism modulates the metabolic capacities of tissues over a longer time period in days and weeks under a given dietary condition. With some enzymes, levels are also regulated by degradation of the steady state amount of the protein.

Metabolism in the Fed State

An average American diet for adults provides approximately 2500 kcal day⁻¹ as percentage of calories from three major components: 40–50% from carbohydrates, 35–40% from fats, and 12–15% from proteins. This level of calories per day is commonly ingested in the form of three meals supplemented with a couple of small snacks. Hence each meal (initiating the fed state) provides a larger excess of calories than what can be utilized immediately for the cellular needs. The fed state refers to the time period following the consumption of a meal in which the glucose, amino acids, and fatty acids derived from the carbohydrate, protein, and fat in the diet are being absorbed from the intestinal tract and distributed in the body by the blood. The excess of the consumed calorie as fuels need to be converted into storage forms such as glycogen in liver (**Figure 1**) and skeletal muscle and TAGs in adipose tissues. Absorptive state initiated after ingestion of a mixed meal causes transient increases in plasma glucose, amino acids, and TAGs (as chylomicrons and very low density lipoproteins called VLDL). Increased levels of glucose and amino acids in the plasma stimulate insulin secretion and diminish glucagon secretion by the endocrine pancreas, resulting in an increased molar ratio of insulin:glucagon in the plasma in the absorptive state. This change in the hormonal levels sets the direction of anabolic metabolism in all tissues and lasts for about 2–4 h depending upon the size and composition of a meal.

Action of Insulin

The binding of insulin to its receptor (a protein tyrosine kinase) on the plasma membrane results in the activation of a series of protein kinases that increase the enzymatic capacity for glucose uptake, glycogen synthesis, and fat synthesis (**Figure 2**). Glucose uptake in skeletal muscle and adipose tissue is increased by stimulating the translocation of glucose transporter 4 (GLUT4) loaded intracellular vesicles to the plasma membrane (**Figure 2**). The synthesis of glycogen is increased in liver and skeletal muscle by activation of glycogen-bound phosphatases (such as protein phosphatase-1, PP-1) that activate glycogen synthase by dephosphorylation (**Figure 2**). The synthesis of fat is increased in liver and adipose tissue by promoting gene transcription and subsequent synthesis of lipogenic enzymes that include acetyl-CoA carboxylase (ACC) (**Figure 1**), fatty acid synthase (FAS), malic enzyme (ME), and fatty acid desaturases (**Figure 1**). Furthermore, there is increased VLDL synthesis and secretion by the liver (**Figure 1**), and increased synthesis of TAGs in adipose tissues as well as deposition of TAG derived from chylomicrons and VLDL in adipose tissue to store excess calories derived from dietary carbohydrates and fats. Also, protein synthesis is enhanced in all tissues especially in the skeletal muscle. The major tissues of interest in this metabolic interplay are liver, skeletal muscle, adipose tissues, and the brain. To illustrate the major regulatory mechanisms, we will discuss the metabolism of glucose by the liver during the fed state.

Glucose Metabolism in the Liver (in the Fed State)

Glucose transport and its phosphorylation

It is no wonder that the liver is considered a processing plant for handling of excess fuels for storage in the fed state and

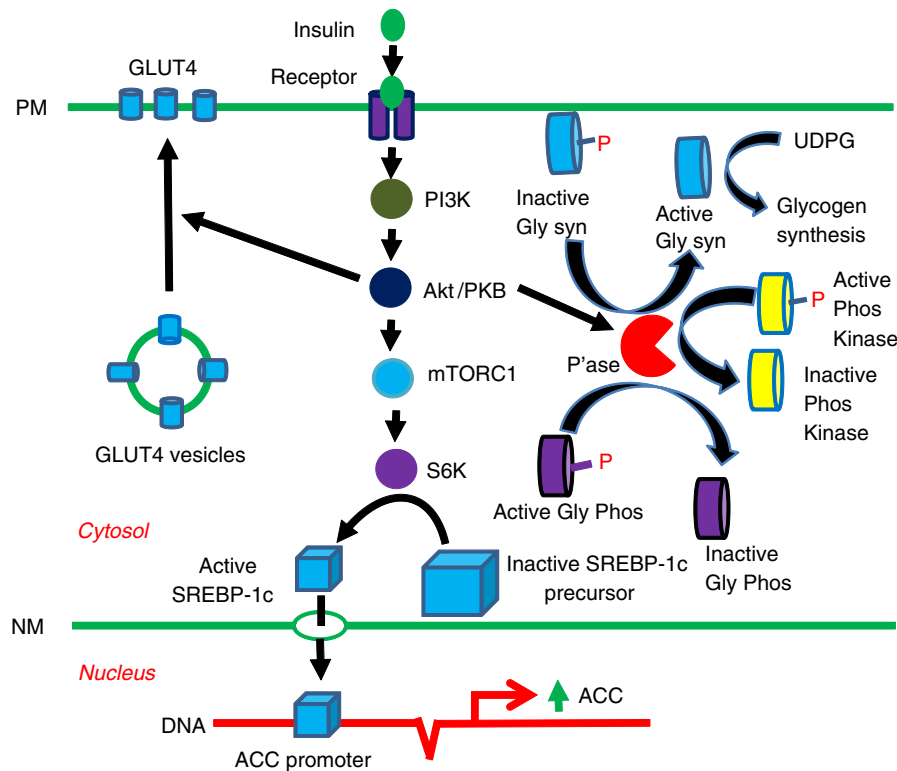


Figure 2 Insulin promotes glucose uptake, glycogen synthesis, and fatty acid synthesis in tissues. The insulin signaling pathway involves activation of PI3K (phosphoinositide 3-kinase), Akt which is also known as PKB (protein kinase B), mTORC1 (mammalian target of rapamycin complex 1), and S6K (S6 ribosomal protein kinase). Akt/PKB promotes activation of glycogen-associated protein phosphatases (probably both PP1 and PP2A) which activate Gly syn (glycogen synthase) and inactivate Phos kinase (phosphorylase kinase) and Gly Phos (glycogen phosphorylase) by dephosphorylation in the liver and skeletal muscle. Akt/PKB also promotes the translocation of vesicles with membrane-bound GLUT4 (glucose transporter 4) from cytosol to the plasma membrane in skeletal muscle and adipose tissues where they fuse and thereby increase the capacity for glucose uptake from the blood. By a mechanism that is a subject of current research, S6K activates proteolytic processing of SREBP-1c (sterol response element binding protein 1c) to produce an active SREBP-1c peptide fragment that binds to the promoter for ACC (acetyl-CoA carboxylase) and promotes ACC transcription. Increased ACC expression along with increased expression of several other lipogenic enzymes increases the rate of fatty acid synthesis in the liver and adipose tissue.

manufacturing/supplying the fuels to the rest of the body during the fasting state. Liver possesses two unique glucose sensors, namely glucose transporter 2 (GLUT2) and glucokinase (both with high K_m for glucose) (Agius, 2008). Hence the transport of glucose in hepatocytes is increased via GLUT2 and its phosphorylation to glucose-6-phosphate is enhanced by glucokinase (in addition to hexokinase with a low K_m , <0.1 mM, for glucose present in hepatocytes). Hence glucokinase activity increases rapidly as glucose concentration in the liver rises (via blood supply by the portal vein) during the absorptive state. Since glucokinase is not inhibited by its reaction product (in contrast to hexokinase), it allows rapid phosphorylation of glucose to glucose-6-phosphate for its metabolism via different pathways. Interestingly, activity of glucokinase is inhibited by its interaction with glucokinase regulatory protein in the presence of fructose-6-phosphate (which is in equilibrium with glucose-6-phosphate by the action of phosphoglucose isomerase) and this inhibition is overcome by fructose-1-phosphate formed from fructose metabolism (from dietary source as sucrose) in the liver (not shown in figure).

Regulation of glycogen synthesis

Glucose-6-phosphate is readily utilized for the synthesis and storage of glycogen and its metabolism is enhanced to pyruvate via the glycolytic pathway due to the action of several regulatory enzymes under the control of insulin-mediated actions. Acetyl-CoA generated from pyruvate via the action of the pyruvate dehydrogenase complex (PDC) is largely diverted for the synthesis of long-chain fatty acids and cholesterol for the formation of VLDL, and is also oxidized via the tricarboxylic acid (TCA) cycle to generate ATP to support the biosynthetic processes, as indicated. These biosynthetic pathways are enhanced by activation of several key regulatory enzymes by their dephosphorylation (covalent modification) (Figure 2) as well as the changes in the levels of allosteric modulators (Table 1). Dephosphorylation of these regulatory enzymes is achieved by the activation of phosphoprotein phosphatase-1 (PP-1) mediated by the action of insulin (Figure 2) (Roach *et al.*, 2012).

Glycogen synthase is the regulatory enzyme in the pathway of glycogen synthesis and is subject to regulation by three mechanisms listed above. The phosphorylated form of glycogen synthase (also referred to as 'D' form for Dependent form;

Table 1 Regulation of several key enzymes in glucose metabolism by their phosphorylation status and allosteric modifiers in the liver

Enzyme	Compounds (+ activator/– inhibitor)	Active enzyme (phosphorylation status)
Glycogen synthase	+ Glucose-6-P + Glucose – ATP, – ADP, – Pi	Non-phospho form
Phosphorylase kinase		Phospho form
Glycogen phosphorylase	– Glucose – ATP – Glucose-6-P	Phospho form
Phosphofructokinase-1	+ Fructose-2,6-bisP + AMP – ATP, – Citrate	
Phosphofructokinase-2 (kinase)		Non-phospho form
Phosphofructokinase-2 (phosphatase)		Phospho form
Pyruvate kinase	+ Fructose-1,6-bisP	Non-phospho form
Pyruvate dehydrogenase complex		Non-phospho form
Pyruvate dehydrogenase kinases	+ Acetyl-CoA + NADH – Pyruvate, – CoA – ADP, – NAD ⁺	
Acetyl-CoA carboxylase	+ Citrate – Acyl-CoA	Non-phospho form
Pyruvate carboxylase	+ Acetyl-CoA	
Fructose-1,6-bisphosphatase	– Fructose-2,6-bisP – AMP	

'inactive' form or glycogen synthase 'b') is strongly inhibited by physiological concentrations of ATP, ADP, and Pi, but this inhibition can be overcome by glucose-6-phosphate. A non-phosphorylated 'I' form (Independent form or 'active form' or glycogen synthase 'a' form) does not require glucose-6-phosphate for its activity. PP-1, activated by the action of an insulin-dependent protein kinase, dephosphorylates the phosphorylated form of glycogen synthase ('D' form; inactive form) into an 'active' form ('I' form) (Figure 2), resulting in an increased flux through this enzyme. Binding of glucose (above 7 mM) to glycogen phosphorylase 'a' promotes its conversion from 'a' form to inactive 'b' form, resulting in its inactivation (and hence inhibiting glycogen degradation). Dephosphorylation of phosphorylase kinase and glycogen phosphorylase (Figure 2) results in the conversion of their 'active' forms into 'inactive' forms. Inactivation of glycogen phosphorylase results in release of PP-1 which then dephosphorylates phosphorylated glycogen synthase 'D' (inactive form) to its non-phosphorylated 'I' form (active form) stimulating glycogen synthesis (Figure 2).

PP-1 is active when it is associated with glycogen through its glycogen-binding G protein. G protein is subject to phosphorylation at two distinct sites: (1) phosphorylation of site 1 by an insulin-stimulated protein kinase activates PP-1 and (2) phosphorylation of site 2 (and also of site 1) by cAMP-dependent protein kinase A (PKA) results in its dissociation from glycogen-associated enzyme complex. In the cytosol, PP-1 inhibitor binds to PP-1, rendering it inactive. Interestingly, PP-1 inhibitor is also subject to phosphorylation (activation) by PKA and dephosphorylation (inactivation) by PP-1 (Roach *et al.*, 2012).

Furthermore, for a long-term regulation of glycogen synthase an intracellular signaling pathway mediated by insulin enhances transcription of the glycogen synthase gene, resulting in increased synthesis of this enzyme protein.

Regulation of the glycolytic pathway

The liver has a limited capacity to store glucose as glycogen (80–100 g) and hence the excess glucose-6-phosphate is processed via the glycolytic pathway and the hexose-monophosphate (pentose) pathway to generate eventually acetyl-CoA and NADPH to support lipid biosynthesis in the liver. The glycolytic pathway has two regulatory enzymes, namely phosphofructokinase-1 (PFK-1) and pyruvate kinase (Figure 1). PFK-1 (the first committed enzyme in the glycolytic pathway) is subjected to allosteric regulation by a number of molecules (namely fructose-2,6-bisphosphate, AMP, ATP, citrate, etc.). Among these allosteric modulators of PFK-1, fructose-2,6-bisphosphate, a product of phosphofructokinase-2 (PFK-2) derived from fructose-6-phosphate, is the most potent activator of PFK-1 (Figure 3) (Kurland and Pilkis, 1995). Interestingly, PFK-2 (which is a bifunctional enzyme with alternate functional kinase or phosphatase activity) is regulated by its covalent modification. Dephosphorylated PFK-2 protein possesses kinase activity, resulting in increased synthesis of fructose-2,6-bisphosphate which then acts as an activator of PFK-1 (Figure 3). In contrast, PFK-2 in its phosphorylated form expresses phosphatase activity and causes a breakdown of fructose-2,6-bisphosphate to fructose-6-phosphate (hence lowering the intracellular concentration of fructose-2,6-bisphosphate) (Figure 3). Hepatic PFK-1 is also subject to long-term regulation at the transcriptional level. Activity of

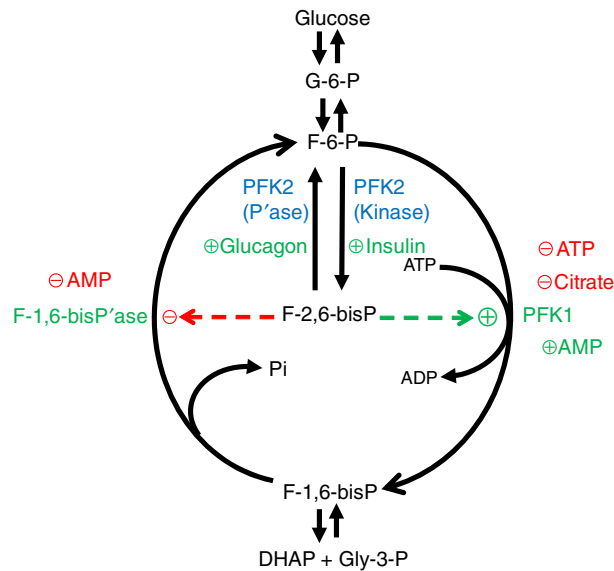


Figure 3 Regulation of glycolysis and gluconeogenesis at the level of PFK1 and F-1,6-bisP'ase. PFK1 uses ATP to convert F-6-P to F-1,6-bisP in the glycolytic pathway. F-1,6-bisP'ase catalyzes the hydrolysis of F-1,6-bisP to F-6-P in the gluconeogenic pathway. Flux in the direction of glycolysis is greatly increased by allosteric activation of PFK1 by F-2,6-bisP coupled with allosteric inhibition of F-1,6-bisP'ase by this same compound. PFK2 is a bi-functional enzyme that functions as a kinase for the conversion of F-6-P to F-2,6-bisP in its dephosphorylated state and as a phosphatase for the conversion of F-2,6-bisP back to F-6-P in its phosphorylated state. By signaling the dephosphorylation and therefore the activation of the kinase moiety of PFK2, insulin increases the level of F-2,6-bisP which stimulates glycolysis and inhibits gluconeogenesis. By signaling the phosphorylation (through PKA) and therefore the activation of the phosphatase moiety of PFK2, glucagon decreases the level of F-2,6-bisP which inhibits glycolysis and stimulates gluconeogenesis. AMP acts as an allosteric effector in manner similar to F-2,6-bisP, that is, it activates PFK1 and inhibits F-1,6-P'ase and ATP opposes stimulatory effect of AMP on PFK1. On the other hand, citrate allosterically increases the effectiveness of ATP as an allosteric inhibitor of PFK1.

pyruvate kinase is subject to regulation by all three mechanisms. Fructose-1,6-bisphosphate is an allosteric activator of (a feed-forward stimulation in the pathway) of liver pyruvate kinase (Table 1). This enzyme is also subject to a phosphorylation/dephosphorylation mechanism. Dephosphorylation of pyruvate kinase by PP-1 results in the conversion to its active form. Insulin-mediated transcription of the pyruvate kinase gene also plays a role in maintaining a steady state level of this enzyme in the liver. Collectively, the increased activities of three enzymes, namely glucokinase, PFK-1 and pyruvate kinase contribute to an increased flux in the glycolytic pathway (glycolysis) in the liver during the fed state.

Regulation of the pyruvate dehydrogenase complex

Pyruvate generated from glucose via the glycolytic pathway is transported by a carrier into the mitochondria where it is metabolized to acetyl-CoA and oxaloacetate by the PDC and pyruvate carboxylase, respectively. Dephosphorylation of

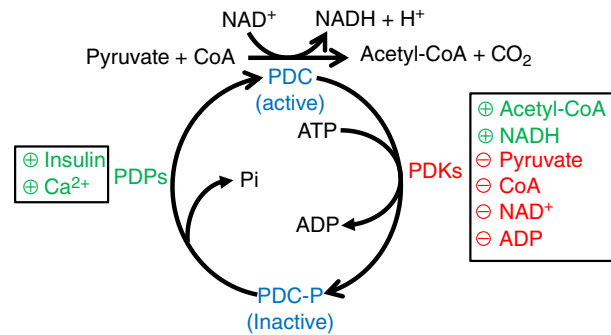


Figure 4 Regulation of the pyruvate dehydrogenase complex. PDC, a multi-enzyme complex found in the mitochondrial matrix space, is responsible for the oxidation of pyruvate to acetyl-CoA and CO_2 with concurrent reduction of NAD^+ to NADH. Subject to regulation by covalent modification, PDC is inactive in phosphorylated state catalyzed by PDKs (pyruvate dehydrogenase kinases) and active in the dephosphorylated state catalyzed by PDPs (pyruvate dehydrogenase phosphatases). Positive allosteric effectors of PDKs are given in green and negative allosteric effectors in red. One of the PDPs is allosterically activated by Ca^{2+} . Although the mechanism remains to be defined, insulin stimulates PDC activity by promoting PDP activity. PDKs and PDPs are integral proteins of the PDC complex.

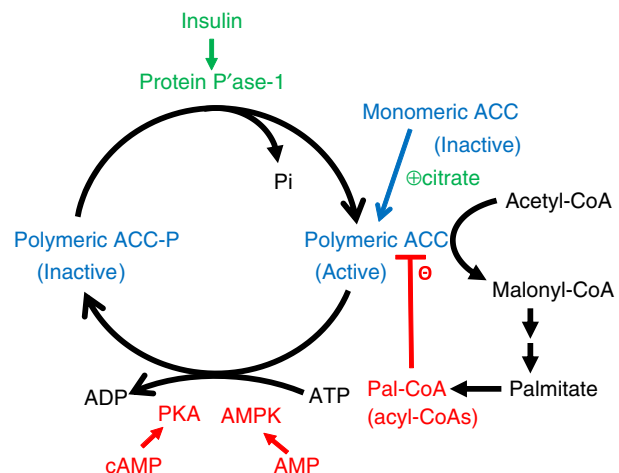


Figure 5 Regulation of fatty acid synthesis at the level of acetyl-CoA carboxylase. ACC is responsible for the conversion of acetyl-CoA to malonyl-CoA for fatty acid synthesis. It exists in the liver cytosol as inactive monomers and large active polymers that are subject to inactivation by phosphorylation by PKA and AMP kinase (AMPK). By increasing cAMP, glucagon promotes inactivation of ACC and therefore inhibition of fatty acid synthesis. Likewise, an energy shortage that causes a fall in ATP and consequently an increase in AMP will also inactivate ACC and inhibit fatty acid synthesis. By signaling the activation of protein phosphatase-1 (PP-1), insulin promotes dephosphorylation and therefore activation of ACC and stimulates fatty acid synthesis. Allosteric inhibition of ACC by long-chain acyl-CoA esters is a good example of regulation of a pathway (lipogenesis) by feedback inhibition.

phospho-PDC by its specific pyruvate dehydrogenase phosphatases results in its activation resulting in increased formation of acetyl-CoA for the formation of citrate (Figure 4) (Harris et al., 2002; Patel and Korotchikina, 2003). Acetyl-CoA

also allosterically activates pyruvate carboxylase to increase oxaloacetate synthesis for citrate formation. A small portion of citrate is metabolized to CO_2 in the TCA cycle to form NADH and FADH₂ which are further processed via the respiratory chain and the oxidative phosphorylation pathway to generate ATP. The terms, the energy charge and phosphorylation potential, are generally used to calculate energy status of cells.

Regulation of fatty acid synthesis

The bulk of citrate is transported out to the cytosol to regenerate acetyl-CoA by the action of ATP-citrate lyase (dephosphorylated 'active' form). Acetyl-CoA in the cytosol is utilized for the synthesis of long-chain fatty acids (palmitate) (Figure 1) and cholesterol. ACC is the first committed step in the pathway of fatty acid synthesis, and is also subject to regulation by the three mechanisms (Figures 2 and 5). ACC in its monomeric form is inactive. Citrate enhances the formation of a polymeric, active form (a chain-like conformation) (Figure 5). Dephosphorylation of a polymeric form is the most active form and its dephosphorylation is carried out by PP-1 (Figure 5). For its long-term regulation, its gene transcription is up-regulated by the action of insulin-mediated signaling pathway (Figure 2). Activation of ACC increases the formation of malonyl-CoA which is further utilized by fatty acid synthase for the biosynthesis of long-chain fatty acids, palmitate (Figure 1). For the reductive synthesis of fatty acids, NADPH is also required. This requirement is met by the action of three enzymes, namely glucose-6-phosphate dehydrogenase, 6-phosphogluconate dehydrogenase, and NADP-malate dehydrogenase (aka malic enzyme) in the cytosol. The protein levels of these enzymes are regulated transcriptionally by the action of insulin (Owen *et al.*, 2012; Zoncu *et al.*, 2011). Increased levels of malonyl-CoA also act as an allosteric inhibitor of carnitine acyl-CoA transferase-I (McGarry, 1998) so that newly synthesized fatty acids (after conversion into their CoA derivatives) are not transported into mitochondria for oxidation but rather utilized for the synthesis of TAGs and incorporated into VLDL for transport. The separation of the pathway for fatty acid synthesis in the cytosol from the pathway of fatty acid oxidation in the mitochondria is a good example of cellular compartmentation for achieving metabolic regulation. The cytosolic pool of acetyl-CoA is also used for the synthesis of cholesterol in the liver. The key regulatory enzyme in this pathway is 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) reductase which is subject to regulation by sterol-dependent regulation of its gene expression, sterol-accelerated enzyme degradation, covalent modification via a phosphorylation/dephosphorylation mechanism, and insulin-dependent up-regulation of its gene transcription. Hence, HMG-CoA reductase activity is increased in the fed state by its dephosphorylation and hence enhancing cholesterol synthesis in the liver. In the fed state, the newly synthesized TAGs and cholesterol are used for the synthesis of VLDL and transported into the blood. Liver also processes chylomicron remnants and reutilizes their TAGs for the formation of VLDL.

Protein synthesis

In the fed state, amino acids derived from dietary proteins are preferentially utilized for protein synthesis in the liver. Excess

amino acids are catabolized and their nitrogen is directed for urea synthesis. Acetyl-CoA derived from the degradation of carbon-skeletons of amino acids can be utilized for ATP production and/or lipid synthesis in the fed state. This, however, may represent a minor contribution to the total hepatic lipid synthesis.

Metabolism in the Fasted (Starvation) State

Fasted (starvation) state can arbitrarily be subdivided in three stages: (1) early fasted state, (2) intermediate fasted (starvation) state, and (3) prolonged fasted (starvation) state. Early fasted state is initiated after the post-absorptive period and continues for about 24–48 h after the last meal. There are metabolic changes which take place in tissues to maintain energy homeostasis and glucose homeostasis during this period (Cahill, 2006). Towards the end of the post-absorptive state as plasma glucose gradually returns to its basal level (~5 mM) so is the level of plasma insulin. During the early fasting state as the plasma levels of glucose and insulin continue to decrease, the plasma level of glucagon rises resulting in a reduction in the molar ratio of circulating insulin: glucagon ratio which favors the action of glucagon at the cellular level.

Action of Glucagon

Glucagon via interaction with its receptors (a G-protein coupled receptor, GPCR) increases synthesis of cAMP by adenylate cyclase, resulting in increased activity of cAMP-dependent protein kinase (PKA) (Figure 6). Increased PKA activity causes phosphorylation of the key regulatory enzymes in the metabolic pathways in the tissues (Berglund *et al.*, 2009; Harris and Crabb, 2010). This action increases the liver's capacity for glucose production by glycogenolysis and gluconeogenesis (Figures 1 and 6). The overall effect of this covalent modification is glycogenolysis since phosphorylation activates the key enzymes localized in the catabolic pathways (e.g., phosphorylase kinase and glycogen phosphorylase 'a') in the glycogen degradation pathway in the liver (Figure 6), and hormone-sensitive lipase in the TAG degradation pathway in adipose tissues. Glucagon increases glucose synthesis by promoting gene transcription and subsequent synthesis of gluconeogenic enzymes that include phosphoenolpyruvate carboxykinase (PEPCK) (Figure 6), F-1,6-bisphosphatase (F-1,6-bisPase), and glucose-6-phosphatase (G-6-Pase) (Figure 1).

Additionally, phosphorylation of the key enzymes (e.g., glycogen synthase, pyruvate kinase, ACC, etc.) in the biosynthetic pathways (Figure 6) results in inactivation of these enzymes to avoid futile cycling. The carbon-fluxes via the biosynthetic pathways (such as glycogen synthesis, fatty acid biosynthesis, and cholesterol synthesis) are markedly inhibited due to phosphorylation and hence inactivation of the key enzymes in these pathways (Figure 1). Protein synthesis is also markedly down regulated by reduction in insulin-mediated actions. Hence, this change in the plasma hormonal levels sets the direction from anabolic (biosynthetic mode) to catabolic state (degradative mode) in all tissues.

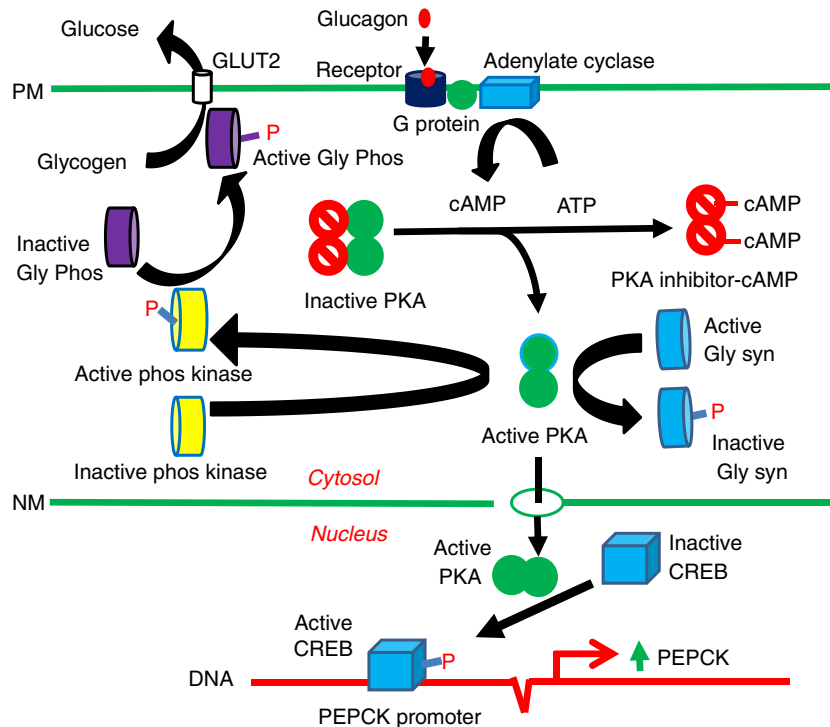


Figure 6 Glucagon promotes glucose production by glycogenolysis and gluconeogenesis. The glucagon signaling pathway involves activation of adenylate cyclase to produce cAMP which activates PKA (protein kinase A) by releasing an inhibitory protein. Active PKA promotes glycogen degradation by inactivating Gly syn (glycogen synthase) and activating Phos kinase (phosphorylase kinase) by phosphorylation. Active phosphorylase kinase activates Gly Phos (glycogen phosphorylase) by phosphorylation. Active PKA moves into the nucleus where it activates CREB (cAMP-responsive element-binding protein) by phosphorylation. Active CREB binds to the promoter of the PEPCK gene to promote transcription of the gluconeogenic enzyme PEPCK. Increased expression of PEPCK along with increased expression of other gluconeogenic enzymes increases the rate of gluconeogenesis.

Metabolism during Prolonged Starvation

After this period (~48 h) fasting extends into an intermediate starvation state lasting for about 2 weeks with major metabolic shifts in tissue fuel choices (Cahill, 2006). During this state the levels of plasma ketone bodies gradually rise to very high levels (in the range of 5 mM), resulting in increased transport and oxidation by the brain (due to high K_{ms} for ketone bodies by the monocarboxylic acid transporter). This shift allows the brain to meet its energy needs (to maintain energy homeostasis) but reducing its total dependency on glucose oxidation. This reduction in glucose demand by the brain coincides with reduction in the rate of hepatic gluconeogenesis from its precursors such as amino acids, glycerol, and lactate and pyruvate. Reduction in hepatic gluconeogenesis reflects in reduction in proteolysis in the skeletal muscle, a sparing effect on protein degradation. Selective fuel utilization in the brain oxidative metabolism from primarily glucose-dependent to a state of mixed fuel dependency (glucose plus ketone bodies) allows humans to survive for a much longer period of total starvation with a limited level of available protein mass (to support gluconeogenesis) and TAG mass in adipose tissues (to support energy homeostasis). During the prolonged starvation state (beyond about 2 weeks of starvation), this adaptive metabolic profile of the brain continues to operate with ketone bodies serving as major fuel

(approximately two-third of the energy needs) and glucose oxidation providing the rest. If starvation continues in this period, a continued demand on body fat and protein reserves cannot be met, leading to a catastrophic condition as death if feeding is not gradually initiated.

The normal American meal-eating pattern results in a re-occurring cycle of absorptive, post-absorptive and early fasted states. In a modern life-style with easily available foods, fasting beyond 24–48 h can only occur by predetermined reasons such as medical, religious, and occasionally political causes. Hence, in this section we will limit discussion on liver metabolism to the early fasting (starvation) state during the first 24–48 h.

Glucose Metabolism in the Liver (During Fasting)

Regulation of glycogen degradation

During the early fasted state the rates of glucose utilization by most tissues decrease (except the brain and tissues with anaerobic metabolism), and the rate of glucose release by the liver increases steadily due to breakdown of stored glycogen. The liver is the sole source of glucose provider for the rest of the body. The contribution to glucose release from hepatic glycogen gradually decreases after about 12 h of fasting and it is compensated by a gradual increase in the rate of hepatic gluconeogenesis during this period. After about 24 h of

fasting, liver glycogen is nearly depleted and the rate of glucose output by the liver is primarily supported by hepatic gluconeogenesis to maintain whole body glucose homeostasis. The rate of hepatic glycogenolysis is increased by phosphorylation (and hence activation) of its two regulatory enzymes, namely phosphorylase kinase and glycogen phosphorylase 'a' (Figure 6) (Roach *et al.*, 2012). These two enzymes are covalently modified by PKA which phosphorylates (and hence activates) phosphorylase kinase 'b' form to its 'a' form. Phosphorylase kinase 'a' then phosphorylates glycogen phosphorylase converting its 'b' (inactive) form into 'a' (active) form (Figure 6). This two-step activation cascade enhances rapid release of glucose from stored glycogen. Activities of these enzymes are also influenced by small molecules such as glucose, glucose-6-phosphate, and other compounds (allosteric effectors) (Table 1). Glucose-1-phosphate generated by this degradative pathway is first converted to glucose-6-phosphate and then to free glucose by glucose-6-phosphatase for its release from the liver.

Regulation of gluconeogenesis

Sources of gluconeogenic precursors are derived from different metabolic pathways: (1) lactate, pyruvate, and alanine are largely derived from inhibition of glucose-derived pyruvate oxidation by the PDC due to the metabolites (NADH and acetyl-CoA) generated from β -oxidation of fatty acids in the skeletal muscle, (2) glycerol generated from TAG breakdown by adipose tissues, and (3) release of gluconeogenic amino acids from proteolysis in the skeletal muscle. The carbon-skeletons of these amino acids are utilized for glucose synthesis and their amino-nitrogen is used for the synthesis of urea for nitrogen disposal. During the early state of fasting (about 24 h), all three sources contribute to glucose synthesis and its release by the liver.

Increases in hepatic cAMP levels also increase transcription of the PEPCK gene (Figure 6) (Yang *et al.*, 2009; Lin and Accili, 2011), resulting in a rapid accumulation of its enzyme protein within a few hours. Increased PEPCK activity together with increased availability of gluconeogenic precursors (such as lactate, pyruvate, and amino acid carbon-skeletons) results in an increased carbon-flux through this initial reaction in the gluconeogenic pathway (mitochondrial oxaloacetate to malate or aspartate and their transport to the cytosol via the malate/aspartate shuttles) for conversion to oxaloacetate and then to phosphoenolpyruvate (Figure 1). Phosphorylation and hence inactivation of hepatic pyruvate kinase in the cytosol and the PDC by its specific pyruvate dehydrogenase kinases (PDKs) in the mitochondria (Figure 4) direct the carbon-flux in the gluconeogenic pathway by inhibiting the former enzyme and minimizing the oxidation of the 3-carbon compound, pyruvate, by the latter enzyme complex (Harris *et al.*, 2002; Patel and Korotchikina, 2003). Additionally, phosphorylation of PFK-2 by PKA converts this bifunctional enzyme from its kinase activity to its phosphatase activity (Figure 3). PFK-2 (phosphatase) activity dephosphorylates fructose-2,6-phosphate to fructose-6-phosphate and hence lowers its intracellular concentrations and reduces its inhibition on fructose-1,6-phosphatase, a gluconeogenic enzyme, converting fructose-1,6-phosphate to fructose-6-phosphate. Glycerol derived from lipolysis in adipose tissues is

phosphorylated to glycerol-3-phosphate and then converted to dihydroxyacetone phosphate, a glycolytic intermediate. The final step in the gluconeogenic pathway is glucose-6-phosphatase which converts glucose-6-phosphate to glucose for its release. A long-term regulation of this enzyme is exerted at its transcriptional level.

Fatty acid oxidation and ketogenesis

This early fasting state is also characterized by increased mobilization of non-esterified fatty acids from TAGs breakdown (lipolysis) in adipose tissues. This results in increased levels of plasma non-esterified fatty acids, transported on serum albumin, which contribute to the energy demands (energy homeostasis) of liver, skeletal muscles, kidneys, heart, and other tissues by increased oxidation of fatty acids. This process also suppresses glucose oxidation and hence contributes to the maintenance of glucose homeostasis (Randle, 1998). Fatty acid oxidation via the β -oxidation pathway generates acetyl-CoA and reduced nucleotides (NADH and FADH₂), resulting in altered intramitochondrial acetyl-CoA/CoA ratio and also NAD⁺/NADH ratio, supporting gluconeogenesis and ketogenesis. Re-oxidation of these nucleotides via the respiratory chain and the oxidative phosphorylation pathway generates ATP to support gluconeogenesis. Acetyl-CoA acts as an allosteric activator of pyruvate carboxylase, an anaplerotic enzyme, converting pyruvate to oxaloacetate which is then converted to phosphoenolpyruvate by PEPCK. Both Acetyl-CoA and NADH act as activators of PDK (Table 1) (Harris *et al.*, 2002; Patel and Korotchikina, 2003) and cause phosphorylation/inactivation of PDC to conserve pyruvate for gluconeogenesis. During this period the liver also initiates the synthesis of ketone bodies (ketogenesis) from acetyl-CoA generated from fatty acid oxidation via the β -oxidation pathway. Ketone bodies are released by the liver, and are readily oxidized by several tissues (such as skeletal muscle, kidneys, heart, etc.) but not by the brain (due to its transporters with high K_ms for ketone bodies). The plasma levels of ketone bodies are maintained at a relatively low level (about 1–2 mM) during this early fasting state.

Dedication

This article is dedicated to Dr. Richard W. Hanson of Case Western Reserve University School of Medicine for his devotion and enormous contributions in the area of regulation of gluconeogenesis, for passionate teaching of metabolism, and for mentoring the next generation of investigators in the area of metabolic regulation. One of us (MSP) deeply acknowledges Dr. Hanson's generous mentorship, collegiality, and friendship over a period of 45 years.

See also: Molecular Principles, Components, Technology, and Concepts: Basic Principles: Biocatalysis. Molecular Principles, Components, Technology, and Concepts: Carbohydrates: Glycogen and Starch. Molecular Principles, Components, Technology, and Concepts: Metabolism: A Structure Perspective on Organelle Bioenergetics

References

- Agius, L., 2008. Glucokinase and molecular aspects of liver glycogen metabolism. Review article. *Biochemical Journal* 414, 1–18.
- Berglund, E.D., Lee-Young, R.S., Lustig, D.G., *et al.*, 2009. Hepatic energy state is regulated by glucagon receptor signaling in mice. *Journal of Clinical Investigation* 119, 2412–2422.
- Cahill, G.F., 2006. Fuel metabolism in starvation. *Annual Review of Nutrition* 26, 1–22.
- Harris, R.A., Bowker-Kinley, M.M., Huang, B., Wu, P., 2002. Regulation of the activity of the pyruvate dehydrogenase complex. *Advances in Enzyme Regulation* 42, 249–259.
- Harris, R.A., Crabb, D.W., 2010. Metabolic interrelationships. In: Devlin, T.M. (Ed.), *Textbook of Biochemistry with Clinical Correlations*, seventh ed. New York, NY: John Wiley & Sons, pp. 839–882.
- Kurland, I.J., Pilkis, S.J., 1995. Covalent control of 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase: Insights into autoregulation of a bifunctional enzyme. *Protein Science* 4, 1023–1037.
- Lin, H.V., Accili, D., 2011. Hormonal regulation of hepatic glucose production in health and disease. *Cell Metabolism* 14, 9–19.
- McGarry, J.D., 1998. Glucose-fatty acid interactions in health and disease. *American Journal of Clinical Nutrition* 67, 500S–504S.
- Owen, J.L., Zhang, Y., Bae, S.-H., *et al.*, 2012. Insulin stimulation of SREBP-1c processing in transgenic rat hepatocytes requires p70 S6-kinase. *Proceedings of the National Academy of Sciences of the United States of America* 109, 16184–16189.
- Patel, M.S., Korotchkina, L.G., 2003. The biochemistry of the pyruvate dehydrogenase complex. *Biochemistry and Molecular Biology Education* 31, 5–15.
- Randle, P.J., 1998. Regulatory Interactions between lipids and carbohydrates: The glucose fatty acid cycle after 35 years. *Diabetes/Metabolism Reviews* 14, 263–283.
- Roach, P.J., DePaoli-Roach, A.A., Hurley, T.D., Taqiliabracchi, V.S., 2012. Glycogen and its metabolism: Some new developments and old themes. *Biochemical Journal* 441, 763–787.
- Yang, J., Reshef, L., Cassuto, H., Aleman, G., Hanson, R.W., 2009. Aspects of the control of phosphoenolpyruvate carboxykinase gene transcription. *Journal of Biological Chemistry* 284, 27031–27035.
- Zoncu, R., Efeyan, A., Sabatini, D.M., 2011. mTOR: From growth signal integration to cancer, diabetes and ageing. *Nature Reviews. Molecular Cell Biology* 12, 21–35.

Further Reading

- Eron, D.M., Ignatova, I.D., Yonemitsu, S., *et al.*, 2009. Prevention of hepatic steatosis and hepatic insulin resistance by knockdown of cAMP Response Element Binding Protein (CREB). *Cell Metabolism* 10, 499–506.
- Gibson, D.M., Harris, R.A., 2002. *Metabolic Regulation in Mammals*. New York, NY: Taylor and Francis.
- Harris, R.A., 2013. Glycolysis overview. In: Lennarz, W.J., Lane, M.D. (Eds.), *The Encyclopedia of Biological Chemistry*, vol. 2. Walham, MA: Academic Press, pp. 443–447.
- Harris, R.A., Jeoung, N.H., Joshi, M., Hwang, B., 2008. The α -keto acid dehydrogenase complexes and the glycine cleavage system: Their involvement in pathways of carbohydrate, protein, and fat metabolism. In: Patel, M.S., Packer, L. (Eds.), *Alpha-Lipoic Acid: Energy Production, Antioxidant Activity, and Health Effects*. Boca Raton, FL: CRC, Taylor and Francis Group, pp. 99–148.
- Jeoung, N.H., Harris, C.R., Harris, R.A., 2014. Regulation of pyruvate metabolism in metabolic-related diseases. *Reviews in Endocrine and Metabolic Disorders* 15, 99–110. doi:10.1007/s11154-013-9284-2.
- Locasale, J.W., Cantley, L.C., 2011. Metabolic flux and the regulation of mammalian cell growth. *Cell Metabolism* 14, 443–451.
- Newsholme, E.A., Leech, A.R., 1983. *Biochemistry for the Medical Sciences*. New York, NY: Wiley.
- Samuel, V.T., Peterson, K.F., Shulman, G.I., 2010. Lipid-induced insulin resistance: Unravelling the mechanism. *Lancet* 375, 2267–2277.

Relevant Website

<http://themedicalbiochemistrypage.org/>
Indiana University School of Medicine and Indiana State University.

A Structure Perspective on Organelle Bioenergetics

WA Cramer and SK Singh, Purdue University, West Lafayette, IN, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Free energy Description of energy that includes contributions from both physical forces and statistics.

Membrane electron transfer chain Organization of proteins containing electron transferring groups such as

hemes, in which the energetically downhill transfer of electrons is coupled to the transfer of protons across the membrane.

Energetics of Membrane-Based Adenosine Triphosphate Synthesis

The free energy available for the active processes of adenosine triphosphate (ATP) synthesis and active transport is derived from the transmembrane proton, or sodium, electrochemical potential gradient (Mitchell, 1979), written as $\Delta\tilde{\mu}_{H^+}$ for protons. The gradient consists of two components, (1) the difference in the concentration of protons, H^+ , between the two sides of the membrane, written as the pH gradient, ΔpH ; (2) the electrical potential, $\Delta\Psi$, across the membrane, expressed in volts or millivolts (mV), which results from the movement of uncompensated electrical charge across the apolar, hydrophobic, and low dielectric membrane. The proton electrochemical gradient, $\Delta\tilde{\mu}_{H^+}$, written in units of kcal mol^{-1} , is:

$$\Delta\tilde{\mu}_{H^+} = F\Delta\Psi - 2.3 RT \times \Delta pH \quad [1]$$

where F is the Faraday constant, $23 \text{ kcal mol}^{-1}\text{-volt}^{-1}$, and the coefficient $2.3RT = 1.36 \text{ kcal mol}^{-1}$. The free energy, ΔG , stored in the transfer on n protons across the organelle membrane is:

$$\Delta G = n \cdot \Delta\tilde{\mu}_{H^+} \quad [2]$$

Thus, in the case of the mitochondrial inner membrane, the thermodynamic contribution of the membrane potential, *ca.* -150 mV , inside (mitochondrial matrix) negative, tends to exceed that of the pH gradient, *ca.* 0.5 , inside more alkaline. Then, the free energy stored for each mole of protons translocated across the membrane from the matrix to extra-membrane domain, via coupling to the respiratory electron transport chain, is described by the following formula (ambient temperature, 25°C):

$$\Delta\tilde{\mu}_{H^+} = 23(0.15) - 1.36(-0.5) = +4.1 \text{ kcal mol}^{-1}$$

Based on structure studies of the extrinsic domain of the ATP synthase that defined three α - β pairs which contain the catalytic site in the F1 rotational nanomotor (Walker, 1998), and an intrinsic Fo sector with eight subunits in the c-ring of the bovine mitochondrial ATP synthase (Watt *et al.*, 2010), on the average $2.7H^+$ are utilized for the synthesis of one ATP molecule (Watt *et al.*, 2010), then the free energy utilized $= 4.1 \times 2.7 = 11.1 \text{ kcal mol}^{-1}$. This is greater than the standard free energy of ATP synthesis, which is approximately $7.5 \text{ kcal mol}^{-1}$ at pH 7 for the synthesis from ADP and

orthophosphate, and large enough to accommodate non-standard concentrations of these reactants.

For oxygenic photosynthesis in the chloroplast thylakoid membrane, the contribution of the ΔpH to the $\Delta\tilde{\mu}_{H^+}$ generated across the thylakoid membrane is believed to dominate the $\Delta\tilde{\mu}_{H^+}$. In a classical experiment that demonstrated the ability of an artificially generated $\Delta\tilde{\mu}_{H^+}$ in thylakoid membranes to drive the synthesis of ATP (Jagendorf and Uribe, 1966; Cramer and Knaff, 1991), a pH gradient of four units, inside negative, was shown to generate more than 1 ATP:10 Chl, a ratio more than one order of magnitude greater than the stoichiometry of the electron transport proteins in the membrane. The transmembrane potential across the thylakoid membrane can also contribute significantly to ATP synthesis (Sacksteder *et al.*, 2000). However, a major difference in the net free energy or $\Delta\tilde{\mu}_{H^+}$ requirement for ATP synthesis in chloroplasts expressed through either mode of energy storage was revealed by the structure analysis of the Fo sector of the chloroplast ATPase by atomic force microscopy. AFM analysis of thylakoid membranes data shows 14 and 15 c-subunits, respectively, in the intra-membrane Fo sector c-ring of the chloroplast (Seelert *et al.*, 2003) and cyanobacterial (Pogoryelov *et al.*, 2005) ATPase. Recalling that one rotation of the c-ring results in the synthesis of three ATP molecules, $4.7 (14/3)$ or $5H^+ (15/3)$ are utilized per ATP synthesized by the ATP synthetic mechanism associated with oxygenic photosynthesis. Referring to the calculation above for mitochondria, this implies that the $\Delta\tilde{\mu}_{H^+}$ required for ATP synthesis in chloroplasts is significantly smaller, $8/14$ or $8/15$, of that utilized in mitochondria. For chloroplasts, the requirement would be $(8/14) (4.1 \text{ kcal mol}^{-1}) = 2.3 \text{ kcal mol}^{-1}$, a value satisfied by a pH gradient, $\Delta pH = 1.7$, acidic inside the thylakoid lumen.

Atomic Structures of Membrane Protein Complexes Responsible for Energization, That Is, Generation of the $\Delta\tilde{\mu}_{H^+}$, of Mitochondria and Chloroplasts

There are three general processes or mechanisms for H^+ translocation known in energy-transducing membranes: (1) light-induced pK changes and resultant H^+ translocation, as in bacteriorhodopsin (Lanyi, 2004); (2) H^+ translocation coupled to electron transfer as in cytochrome oxidase (Yoshikawa *et al.*, 2012); or (3) oxidation/reduction of lipophilic quinol/quinone, as in (a) the cytochrome b_6f complex, and (b)

cytochrome bc_1 complex that function, respectively, in (a) oxygenic photosynthesis and (b) respiration or anoxygenic photosynthesis.

A Caveat

The pathways defined by crystal structures for proton transfer in the whole set of bioenergetic membrane proteins, and thereby the mechanisms for transmembrane proton transfer are in most cases still incomplete. The reasons for this gap are partly the complexity of the structures and partly lack of resolution in the structures. One aspect of the latter problem is that intra-membrane water is usually part of the transmembrane H^+ transfer chain, as described in the Grootthus mechanism (Cukierman, 2006). For the latter problem, the resolution of intrinsic waters, i.e., of the O atom of the water requires at the least a resolution in the quality of the diffraction pattern $\leq 2.5 \text{ \AA}$, a level that has not been achieved in the diffraction analysis of many crystallized membrane proteins. The system for which the pathway of H^+ transfer and generation of the membrane potential and proton gradient has been most clearly elucidated is that of bacteriorhodopsin, for which an array of high-resolution crystal structures has been obtained (Lanyi, 2004), in which an array of H_2O molecules have been shown to connect key carboxylate residues in the

pathway of transmembrane H^+ transfer driven by the light-induced retinal-based proton pump. Even here, the fundamental mechanisms by which the proton-carrying residues undergo a sequential increase and decrease in effective pK value, which allows these residues to act alternately as H^+ donors and acceptors, has not been described.

Primary Charge Separation and Formation of a Membrane Potential; Photosynthetic Reaction Centers

A primitive and fundamental mechanism for generation of the transmembrane potential, $\Delta\Psi$, is associated with the primary charge separation in photosynthetic reaction centers.

Reaction Center of Anoxic Photosynthetic Bacteria

The reaction center from the purple photosynthetic bacteria, *Blastochloris viridis*, documented in the protein data base (PDB) with PDB code 3T6D, is shown (Figure 1). The reaction centers are the long wavelength 'traps' to which light energy absorbed by the light-harvesting pigments is transferred, and which also serve as the interface between the processes of photosynthetic light harvesting and subsequent transmembrane electron transfer. This dimeric reaction center complex consists of four

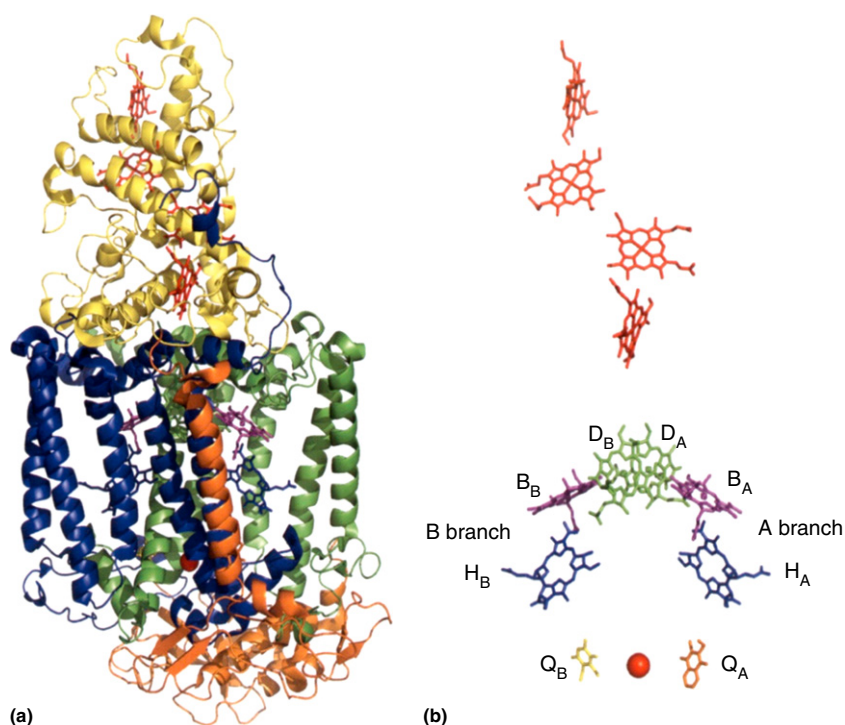


Figure 1 Structure of the photosynthetic reaction center from the purple photosynthetic bacterium, *Blastochloris viridis* (Roszak et al., 2012). Left: The diagram shows the backbone polypeptide chain of a four heme cytochrome (yellow) that provides the electron donor to the reaction center, and the M ('medium,' purple), L ('light,' green), and H ('heavy,' orange) subunits, which contain 5, 5, and 1 transmembrane helices, respectively. The terms 'M', 'L', and 'H' refer to the relative positions of these subunits when they are separated by electrophoresis under denaturing conditions. Right: Prosthetic electron and proton transfer groups in reaction center: top, extra-membrane set of four hemes that function as the p-side electron donor to the complex; bottom, transmembrane electron transfer pathway in the dimeric structure from the 'special pair' electron donor (D_A , D_B) to the transmembrane pathway, via bacteriochlorophylls B_A and B_B , bacteriopheophytins (bacteriochlorophyll without the central Mg^{2+}), n-side electron and proton acceptors, quinones Q_A and Q_B . Reprinted with permission from Blankenship, R.E., 2014. Molecular Mechanisms of Photosynthesis, second ed. Wiley-Blackwell.

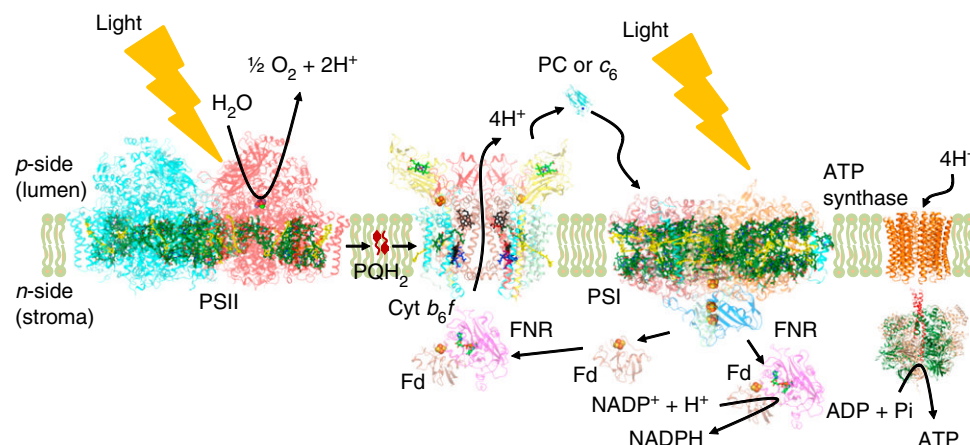


Figure 2 Photosynthetic electron transport chain in membranes that support oxygenic photosynthesis (Hasan *et al.*, 2013). Figure emphasizes the structures of the three hetero-oligomeric complexes in the chain: (a) the hetero-dimeric photosystem II complex (Figure 3); molecular weight of cyanobacterial PSII complex ≈ 350 kDa, containing 20 subunits, 35 chlorophyll molecules, 11 carotenoids, 14 lipids, 2 hemes, 1 non-heme iron, and the water-splitting Mn_4CaO_5 -cluster; (b) homo-trimeric photosystem I reaction center (trimer MW of cyanobacterial complex ≈ 1 MDa (Jordan *et al.*, 2001); 12 protein subunits and 127 cofactors comprising 96 chlorophylls, 2 phyloquinones, 3 Fe_4S_4 clusters, 22 carotenoids, 4 lipids, a putative Ca^{2+} ion, and 201 water molecules); monomer MW of plant complex ≈ 600 kDa (Amunts *et al.*, 2007), (c) the homo-dimeric cytochrome b_6f complex whose structure (Hasan *et al.*, 2013; Baniulis *et al.*, 2009; Cramer and Zhang, 2006; Hasan and Cramer, 2014; Kurisu *et al.*, 2003; Stroebel *et al.*, 2003; Yamashita *et al.*, 2007) is described below (Figure 4).

polypeptide subunits in the reaction center from the bacteria *Rhodospseudomonas viridis* (Deisenhofer *et al.*, 1984, 1995; Deisenhofer and Michel, 1989), and *B. viridis* (Roszak *et al.*, 2012), the L (light), M (medium), and H (heavy) subunits, which span the membrane, respectively, and a fourth subunit, a c-type cytochrome consisting four hemes, which serves as an electron donor to the complex (Figure 1). A structure of a three subunit complex, consisting of only the L, M, and H subunits was obtained from *Rhodobacter sphaeroides* (Allen *et al.*, 1987). Electrons are donated by the four heme cytochrome subunit to a bacteriochlorophyll dimer (the 'special pair,' 'P870,' referring to the wavelength of the absorbance peak, in *Rps. viridis*). Each monomer contains a p-side bacteriochlorophyll that is coordinated by a histidine residue (not shown) from transmembrane helix IV of the L- and M-subunits on the p- or cytochrome side of the complex. Light absorbed by the 'special pair' bacteriochlorophyll dimer drives electron transfer across the membrane to a quinone acceptor, with the initial transfer on a picosecond (10^{-12} s) time scale via a neighboring (several Å separation) bound bacteriochlorophyll and bacteriopheophytin to the quinone seen on the n-side boundary of the complex. The transmembrane electron transfer across the 30–35 Å hydrophobic domain generates a membrane potential, with a membrane specific capacitance of approximately $1 \mu\text{farad cm}^{-2}$, the transfer of one electron per 10^4 Å^2 generates a transmembrane potential, $\Delta\Psi$, on the order of 100 mV, negative on the side of the H subunit, where the quinone electron acceptor is seen. Electronic reduction of the quinone is concomitant with proton uptake to form the protonated reduced quinol, QH_2 , that functions in the transfer of protons across the membrane (see below, Pathways of Proton Transfer). The bacterial reaction center has also been solved to high resolution from the bacterium, *R. sphaeroides*, (Allen *et al.*, 1987), with similar and complementary structure information on electron charge transfer.

Reaction Centers of Photosynthetic Plants, Algae, and Cyanobacteria

The two photosystems responsible for charge transfer, oxygen evolution, and NADP^+ reduction/ CO_2 fixation in oxygenic photosynthesis are arranged in a linear electron transfer chain (Figure 2) in which electron transfer is initiated by water serving as the electron donor to the PSII reaction center. The PSII reaction center has been solved at a resolution of 1.95 Å (Suga *et al.*, 2015), which allowed a description of the Mn_4CaO_5 -cluster that catalyzes the oxidation of water to O_2 via redox intermediates of this cluster that support and define the four-step oxidation of H_2O to O_2 . The transmembrane electron transfer chain that carries the light-driven transport of electrons from water to a reaction center dimer ('P680') on the lumen or p-side of the complex and across the membrane to reduce plastoquinone on the stromal (n-side) of the membrane is structurally similar to the primary electron transfer chain in the photosynthetic bacteria (Figure 1) except that a chlorophyll and pheophytin in the system operating in plants, algae, and cyanobacteria replace the bacteriochlorophyll and bacteriopheophytin in the bacterial system. The arrangement of most of the 20 polypeptide subunits of the PSII reaction center along with the light-harvesting complexes, CP43 and CP47, is shown (Figure 3). The structure contains two polypeptide subunits in its core, D1 and D2, whose arrangement in the photosystem II reaction center complex is shown (Figure 3). The hetero-dimeric heme cross-linked cytochrome-b-559 (Widger *et al.*, 1985; Shinopoulos and Brudvig, 2012), drawn in purple in Figure 3, whose heme potential is dependent upon the hydrophobic environment in the membrane (Krishtalik *et al.*, 1993), is a PSII reaction center component whose function is not understood in the context of a linear electron transport function of the reaction center.

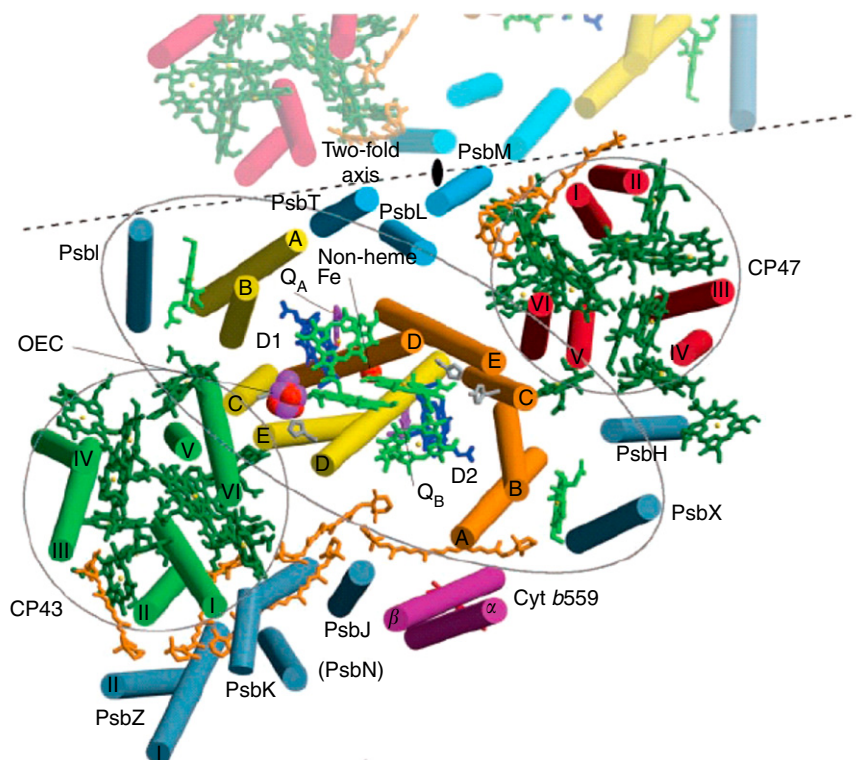


Figure 3 Monomer unit of the photosystem II reaction center complex. View of the PSII monomer from the luminal side normal to the membrane plane. Circles define the reaction core of the complex containing subunits D1, D2, PsbI, and PsbX, and separately the light-harvesting complexes CP43 and CP47. The dashed line defines the boundary of the monomeric unit in the dimer (Ferreira *et al.*, 2004). Reprinted with permission from Shinopoulos, K.E., Brudvig, G.W., 2012. Cytochrome b(5)(5)(9) and cyclic electron transfer within photosystem II. *Biochimica et Biophysica Acta* 1817, 66–75.

Reactions on the electrochemically positive p-side of the membrane support (1) the splitting of water (H_2O) to O_2 , and plastoquinol diffusion to the b_6f complex leading to reduction of the p-side [2Fe-2S] iron-sulfur protein (ISP), which supplies protons to the lumen (p-) side of the membrane and thereby contributes perhaps two-third of the protons derived from the linear electron transport chain, to the positive proton electrochemical potential. (2) Reduction of the ISP is followed by electron transfer to the heme of cytochrome *f* (not seen), then to the soluble copper protein plastocyanin or, in the presence of low copper concentrations, to a c-type cytochrome, either of which serves as a donor to the trimeric PSI, photosynthetic reaction center complex. (3) Photochemically driven electron transfer in the PSI reaction center across the membrane through an Fe and quinone electron transfer network reduces ferredoxin.

Reactions on the electrochemical n-side of the membrane: (1) Plastoquinone, PQ, in PSII, accepts electrons donated to the P680 special pair and transferred across the membrane on a picosecond–nanosecond time scale, and accepts protons (H^+) from the n-side aqueous phase, so that PQ is reduced to the hydrogen donor, PQH_2 . PQH_2 diffuses through the membrane bilayer to the b_6f complex. (2) Ferredoxin reduced by the PSI reaction center reduces NADP^+ to NADPH that is required for fixation of CO_2 . Ferredoxin can also be used in the ‘PSI cyclic electron transport pathway’ to feed electrons back to plastoquinone in the linear or ‘noncyclic’ electron transport

chain, with transfer via the b_6f complex (Zhang *et al.*, 2001), one possible pathway for closing the cyclic pathway (Figure 2).

Cytochrome b_6f Complex in Oxygenic Photosynthesis Provides Electronic Connection between the Two Reaction Center Complexes

The cytochrome b_6f complex (‘*f*’ for leaf; folium (latin)) (Figure 4), 270 kDa including lipids and prosthetic groups, connects the two photosynthetic reaction centers (Figure 3), and through its function in oxidizing hydrogenated plastoquinol, is coupled to proton translocation into the lumen-side aqueous phase and thereby mediates formation of most of the proton transmembrane electrochemical gradient, $\Delta\tilde{\mu}_{\text{H}^+}$, generated by the electron transport chain. The complex is a symmetric dimer that contains 13 TMH/monomer, 7 prosthetic groups ((5 redox, 4 hemes + 1 [2Fe-2S] cluster); 1 β -carotene, 1 chlorophyll *a*), and 23 lipid-binding sites/monomer. The b_6f complex is linked in evolution (Dibrova *et al.*, 2013; Nitschke *et al.*, 2010; Widger *et al.*, 1984; Schutz *et al.*, 2000) to the cytochrome bc_1 complex (Solmaz and Hunte, 2008; Esser *et al.*, 2004) that has similar functions in the mitochondrial respiratory chain and photosynthetic bacteria. There are specific differences in prosthetic group composition and structure that relate to some differences in detailed function. Four significant

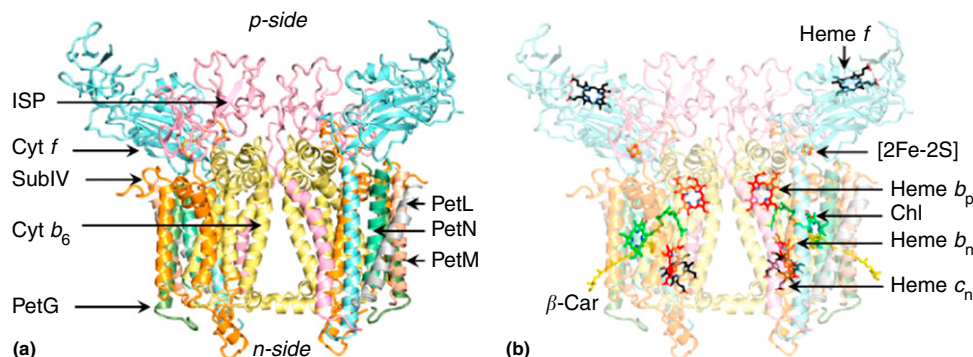


Figure 4 Dimeric cytochrome b_6f complex in the cyanobacterium *Nostoc pcc 7120*. (a) Ribbon diagram shows the four major subunits, cytochrome f (cyan), cytochrome b_6 (yellow), ISP (pink), and SubIV (orange), and four smaller 'Pet' (Photosynthetic electron transfer) subunits, PetG (dark green), PetL (gray), PetM (light brown), and PetN (light green). (b) Semi-transparent diagram shows the positions of prosthetic groups, with redox (1) or other (2) functions: (1) hemes f and c_n (covalently bound c -type hemes), hemes b_p and b_n (non-covalently bound hemes on the p- and n-sides of the complex), and a [2Fe-2S] ISP. 'p' and 'n' denote the electrochemically positive (lumen) and negative (stroma) sides of the membrane. (2) the 20 carbon chain of a bound chlorophyll a molecule appears to control ('gate') access/exit of plastoquinol to the [2Fe-2S] center, the electron/proton acceptor of the quinol (Hasan *et al.*, 2014). The β -carotene has been proposed to function as a 'latch' (Hasan and Cramer, 2012) for a 'super-complex' formed with the photosystem I reaction center (Iwai *et al.*, 2010). Figure modified from Hasan, S.S., Yamashita, E., Cramer, W.A., 2013. Transmembrane signaling and assembly of the cytochrome b_6 -lipidic charge transfer complex *Biochimica et Biophysica Acta* 1827, 1295–1308.

differences in function involve the ability to (1) participate in ferredoxin-mediated cyclic electron transport and coupled phosphorylation (Zhang *et al.*, 2001; Arnon *et al.*, 1967), (2) support a transmembrane signaling mechanism that involves activation via p-side quinol oxidation of an n-side kinase activity that phosphorylates light-harvesting chlorophyll proteins (Vener *et al.*, 1997; Rochaix, 2014), (3) generate superoxide generated by plastosemiquinone reduction of molecular oxygen at a rate more than an order of magnitude greater than measured for the bc_1 complex (Baniulis *et al.*, 2013), and (4) from the affinity of quinone analogue inhibitors for the n- and p-side quinone-binding sites (Yamashita *et al.*, 2007; Hasan *et al.*, 2014).

The cytochrome b_6f complex as a lipoprotein

A relatively new perspective in structure–function of membrane proteins is that these proteins are not only surrounded by the lipid of the bilayer membrane, but lipid is also intercalated in hetero-oligomeric membrane proteins such as the mitochondrial cytochrome bc_1 complex (Hunte and Richers, 2008) and cytochrome oxidase (Shinzawa-Itoh *et al.*, 2007), the photosynthetic photosystem II reaction center (Guskov *et al.*, 2009), and the cytochrome b_6f complex (Hasan and Cramer, 2014; Figure 5). The functions of these intercalated lipids are presently a major area of investigation. The other basic set of questions that result from the new structure information on lipids in membrane proteins involves the mechanisms of the assembly of membrane proteins. The data and concepts about translocon function in membrane protein folding and assembly (Park and Rapoport, 2012; White and von Heijne, 2008) constitute a major advance. However, based on the now seemingly universal presence of lipids in the core of oligomeric membrane proteins, concepts of membrane protein assembly and folding must now take into account

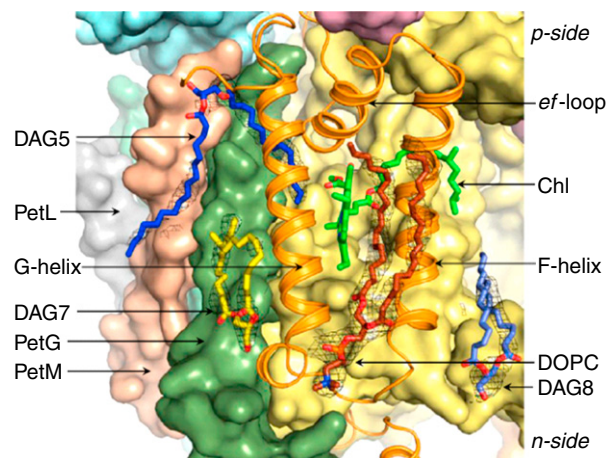


Figure 5 Lipid-binding sites in the Cytochrome b_6f complex. From the 2.5 Å crystal structure of the cyanobacterial b_6f complex (Hasan and Cramer, 2014), 23 lipid-binding sites that mediate internal interactions in the complex were identified per monomer. The sites consist of a mixture of physiological lipids, synthetic lipid (DOPC) uniquely needed for crystallization (Zhang *et al.*, 2003), and detergent used for purification and crystallization: (1) the chlorophyll-binding site is stabilized by lipid; (2) the lipid, diacyl-glycerols, DAG7 (yellow/red sticks) and DAG8 (blue/red), interact with the subunit IV 'G' and 'F'-helices, (3) DAG7 also interacts with the single transmembrane helix of the PetG subunit in the outside 'picket fence.' Figure reproduced with permission from Hasan, S.S., Cramer, W.A., 2014. Internal lipid architecture of the hetero-oligomeric cytochrome b_6f complex. *Structure* 22, 1008–1015.

mechanisms for lipid insertion into membrane protein complexes.

Other functions associated with internal lipids are the following: (1) boundary lipids provide structure and perhaps

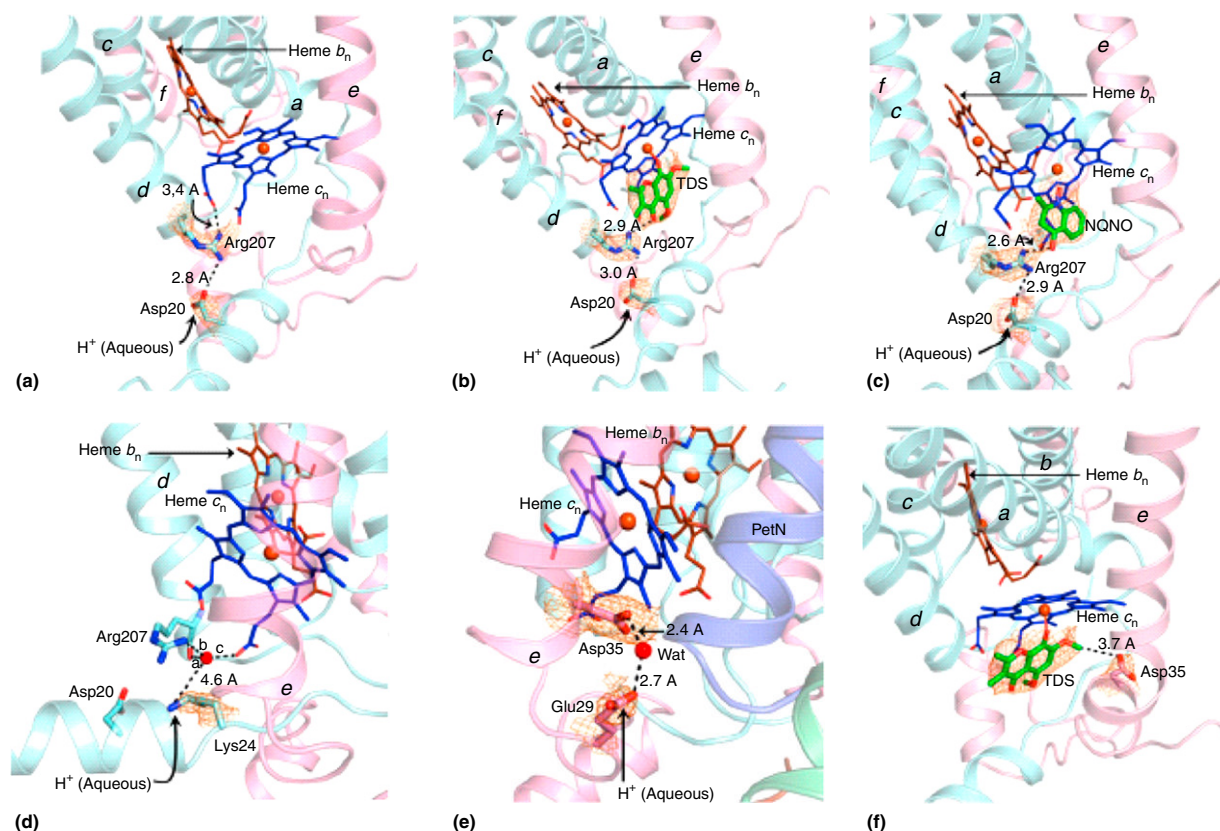


Figure 6 n-Side proton (H^+) uptake pathways of the b_6f complex. (a) The D (aspartate, Asp)-R (arginine, Arg) H^+ uptake pathway, described by a 2.70 Å structure (Baniulis *et al.*, 2009). An Asp20 → Arg207 route of H^+ transfer is described. (b,c) Arg207 as ligand to the quinone analogue inhibitors TDS (b) and NQNO (c), which mimic the physiological quinone, bound at the n-side quinone (Qn) binding site. Arg207 side chain interacts with the natural quinone bound at the Qn-site. Color code: heme b_p , brown lines; heme c_n , blue lines; TDS and NQNO, green-red-blue sticks; Peripheral Pet G, L, M, and N subunits and hydrocarbon tails of TDS and NQNO, not shown. (d) Potential H^+ uptake pathway mediated by water. Water 416 (red sphere) is coordinated by Arg207 backbone carbonyl oxygen and side chain, along with the propionate-A carboxylic acid group of heme c_n (blue). Wat416 is separated from the basic side chain of Lys24 in the cytochrome b subunit (cyan) by 4.6 Å. (e,f) E/D pathway of proton entry on the n-side of the complex. (e) In a 2.70 Å structure, Glu29 (subunit IV) located on the n-side surface of the complex interacts with a water molecule ('Wat,' red sphere), which forms a hydrogen bond with the acidic carboxylate side chain of Asp35 (subunit IV). (f) The Asp35 residue interacts with the quinone analogue TDS that is bound at the Qn-site. Transmembrane helices of cytochrome b (a–d) and subunit IV (e, f) are labeled. Reprinted with permission from Hasan, S.S., Yamashita, E., Baniulis, D., Cramer, W.A., 2013. Quinone-dependent proton transfer pathways in the photosynthetic cytochrome b_6f complex. *Proceedings of the National Academy of Sciences of the United States of America* 110, 4297–4302.

negative electrical charge adjacent to the β -carotene, whose 11 Å extension has been proposed to serve as a 'latch' to mediate super-complex formation (Iwai *et al.*, 2010) with the photosystem I reaction center complex and (2) structural interaction with the putative transmembrane kinase phosphorylates the light-harvesting chlorophyll protein for transmembrane signaling (Hasan and Cramer, 2012), and contributes to the dielectric heterogeneity that influences electron transfer between the transmembrane b -hemes in each monomer of the complex (Hasan *et al.*, 2014).

Structures, mechanisms, and mechanisms of H^+ transfer in the cytochrome b_6f complex

Understanding of the mechanisms of formation of the transmembrane proton electrochemical potential gradient requires an understanding of the intra-membrane and intra-protein pathways of proton transfer. As noted above, the energy-

transducing membrane protein for which these pathways have been mapped most completely is bacteriorhodopsin (Lanyi, 2004). Here, we discuss the pathways of H^+ transfer in the cytochrome b_6f complex. The fundamental mechanisms of transmembrane H^+ transfer discussed here for the b_6f complex, involvement of quinone/quinol, chains of protonatable amino acids, and water chains, also apply to the bc_1 complex.

Pathways of proton uptake into the membrane

As shown in Figure 6, the mechanism of transmembrane proton transfer utilizes quinone molecules, ubiquinone in the mitochondrial respiratory and photosynthetic bacteria bc_1 complex, and plastoquinone in oxygenic photosynthesis, as H^+ donor and acceptor, respectively, on the p- and n-side of the membrane. A unique aspect of these quinone-binding sites in mitochondria is that they can serve as binding sites for

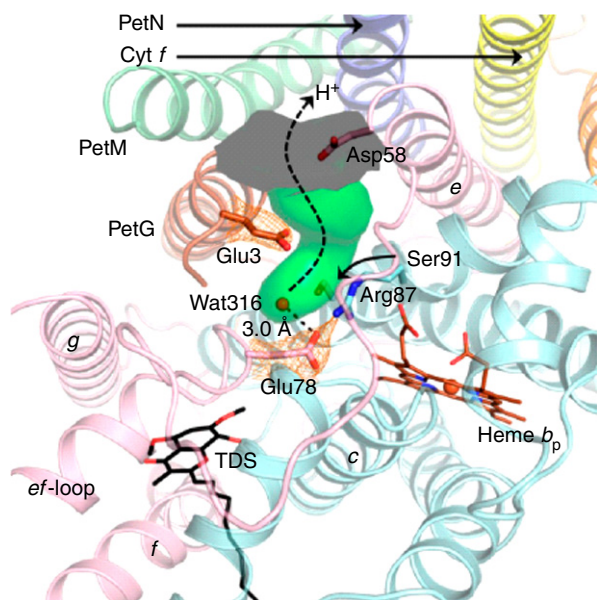


Figure 7 p-Side H^+ exit pathway. Proposed proton exit pathway from the p-side binding niche of plastoquinol of the b_6f complex comprised of a hydrophilic portal (green), which is lined by the amino acids Glu78, Arg87 (cytochrome b_6), Ser91 (cytochrome b_6), Glu3 (PetG), and Asp58 (subunit IV). A broken arrow (black) marks the exit pathway of protons. Glu78 is connected to water 316 (wat316 in the figure) through a hydrogen bond that lies at the end of this portal. Tridecyl-stigmatellin (TDS), a quinone analogue inhibitor, has been inserted from the structure, PDB 2E76 (Yamashita *et al.*, 2007) to define the position of the Qp-site used by the physiological plastoquinone. Reprinted with permission from Hasan, S.S., Yamashita, E., Baniulis, D., Cramer, W.A., 2013. Quinone-dependent proton transfer pathways in the photosynthetic cytochrome b_6f complex. *Proceedings of the National Academy of Sciences of the United States of America* 110, 4297–4302.

quinone analogues that are antimalarials, for example, atovaquone at the p-side quinone-binding niche (Birth *et al.*, 2014), and a specialized pyridone class of inhibitors at the n-side quinone-binding site (Capper *et al.*, 2015).

Proton donation to the p-side aqueous phase; water intercalation into the cytochrome complex and the membrane

The bound TDS quinone analogue inhibitor marks the position that the plastoquinone occupies transiently while delivering two protons to the aqueous phase, one via a histidine residue of the p-side [2Fe-2S] protein and the other through the glutamic acid residue, Glu78 (Figure 7), to a nearby water marked by water oxygen 316. Lipophilic quinone has a central role in proton translocation in the cytochrome b_6f and bc_1 complexes. The mechanism, the 'Q cycle' is particularly well documented for the mitochondrial and photosynthetic bc_1 complex (Mitchell, 1975, 1979; Berry *et al.*, 2000; Crofts *et al.*, 2008; Cooley *et al.*, 2009).

The intrusion of water, shown in Figure 7, into membrane proteins has been noted in other membrane transport proteins, for example, the bc_1 complex (Solmaz and Hunte, 2008) and the lactose transporter (Kaback, 2004), and may have a universal function in these proteins in facilitating the transfer of protons across part of the membrane.

The Mitochondrial Respiratory Chain

Central properties of the respiratory chain shown in Figure 8 are:

1. The maximum number of H^+ translocated across the membrane from the n (matrix)-side to the p-(cytoplasmic) side of the membrane for each two electrons transferred through the electron transport chain, from the maximally reducing end (NADH dehydrogenase) of the electron transport chain, to the oxidizing end at which O_2 is reduced

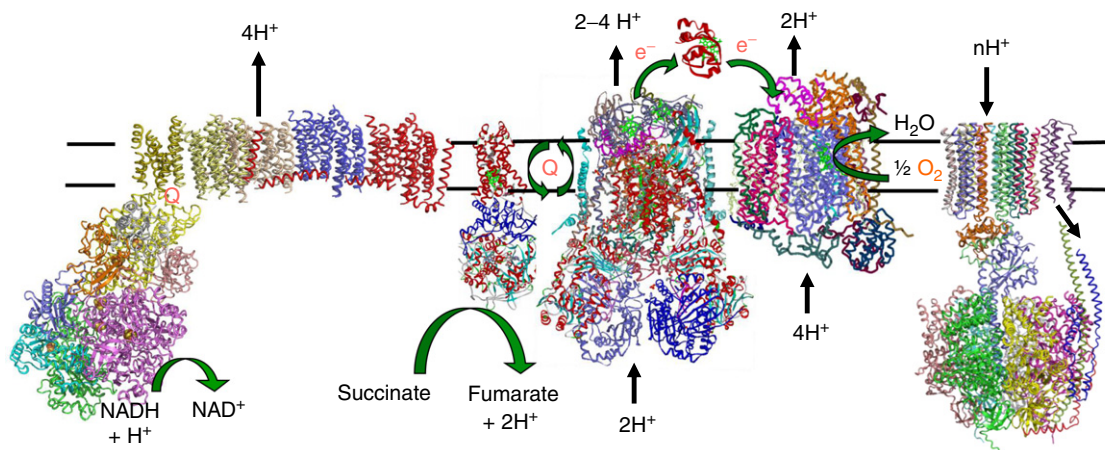


Figure 8 Electron and proton transfer complexes of the respiratory chain in bovine mitochondria summarized in Hosler *et al.* (2006). These complexes include the NADH dehydrogenase, succinate dehydrogenase (pdb, 1NEN), cytochrome bc_1 complex (pdb, 1PP9), cytochrome c oxidase (pdb, 1V54); and Q, lipophilic ubiquinone/ol. Reprinted with permission from Hosler, J.P., Ferguson-Miller, S., Mills, D.A., 2006. Energy transduction: Proton transfer through the respiratory complexes. *Annual Review of Biochemistry* 75, 165–187, updated by S. Ferguson-Miller.

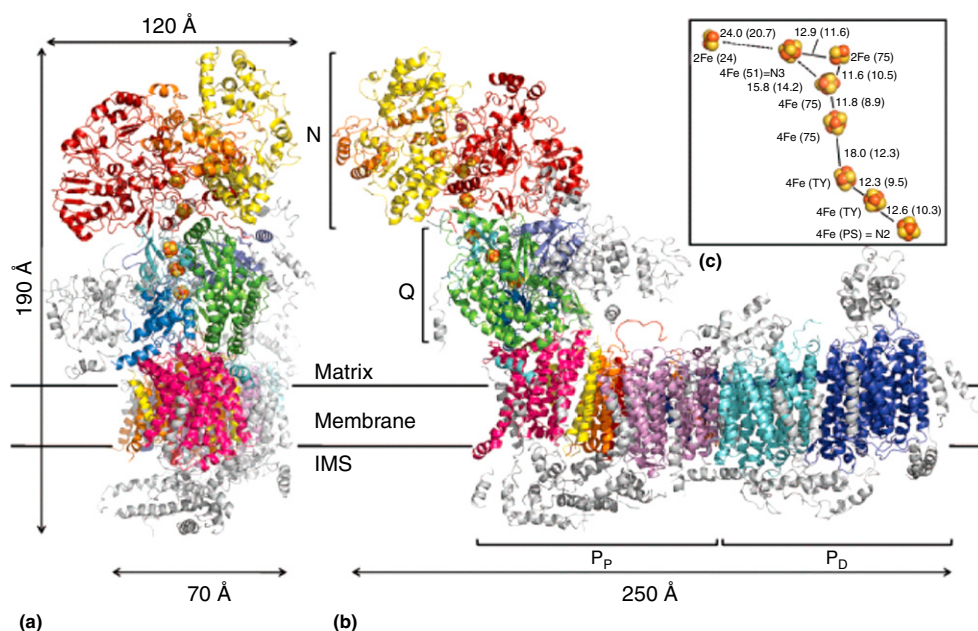


Figure 9 Structure of the mitochondrial complex I from *Y. lipolytica*. Ribbon diagram of L-shaped structure of NADH:ubiquinone oxidoreductase (complex I) from peripheral arm (a) and (b) rotated 90°. Accessory subunits are shown in gray. (c) Arrangement of cofactors in the peripheral arm. Center-to-center and edge-to-edge (in brackets) distances are in angstroms. Reprinted with permission from Zickermann, V., Wirth, C., Nasiri, H., *et al.*, 2015. Structural biology. Mechanistic insight from the crystal structure of mitochondrial complex I. Science (New York, NY) 347, 44–49.

to H_2O in the cytochrome oxidase is 10 (i.e., $\text{H}^+ / 2\text{e}^- = 10$ for the respiratory chain).

- The other parameters associated with mitochondrial energy transduction and ATP synthesis are the specific efficiency of ATP production in the total electron transport chain initiated by the oxidation of NADPH, which is approximately 2.8 ATP/2e⁻ (Hinkle, 2005), if the average stoichiometry of H^+ utilization associated with ATP synthesis is 2/3 (8/3) H^+ utilized per ATP synthesized (Watt *et al.*, 2010) by the 'F1-Fo' rotational nanomotor (Walker, 1998; Junge *et al.*, 2009). A different perspective on the organization of the mitochondrial respiratory chain, obtained from analysis of the entire chain and sub-complexes is that it is organized as a 'super-complex' (Althoff *et al.*, 2011). Although this structure perspective needs verification, particularly with respect to gaps between the individual complexes, it has the conceptually important perspective that it solves the problem of inter-complex connections which, in the 'classical textbook models' (e.g., Figure 8) involves random diffusion of quinone within the membrane, and of cytochrome *c* at the external p-side interface, to electronically connect the complexes.
- Complex I, or NADH dehydrogenase, is at the reducing end of the respiratory chain. The eukaryotic (e.g., human) (Zickermann *et al.*, 2015) and bacterial (NDH-1) (Baradaran *et al.*, 2013) NADH dehydrogenase ('complex I') contain 45 subunits and 16 subunits, respectively. The overall proton/electron stoichiometry for proton translocation in complex I is $\text{H}^+ / 2\text{e}^- = 4$. Of all the mitochondrial proton-translocating complexes, the intra-protein mechanism of proton translocation is least well understood in complex I.

Complex I, NADH dehydrogenase

Atomic structures of the entire complex I (536 kDa, 16 subunits, 9 Fe-S clusters, 64 TM helices) have been obtained from the bacterium, *Thermus thermophilus* (Baradaran *et al.*, 2013) (536 kDa, 16 subunits, 9 Fe-S clusters, 64 TM helices), and from the yeast model, *Yarrowia lipolytica* (Figure 9); for the eukaryotic complex, defects in complex I are responsible for the majority of mitochondrial-based diseases (myopathies), presumably because this complex is operating at the most negative oxidation–reduction potential (midpoint oxidation–reduction potential of $\text{NAD}^+ / \text{NADH}$ at pH 7 = −0.32 V) in the respiratory chain, with which it can reduce O_2 to superoxide (oxidation–reduction potential = −0.15 V). The NADH oxidation site is at the exterior of the long (>100 Å) peripheral domain (panel C) extending from the n- or matrix side of the membrane. This extension contains eight iron–sulfur clusters that transfer electrons from NADH to the ubiquinone that is approximately 20 Å from the membrane interface and, somewhat surprisingly not positioned well into the hydrophobic membrane. The structure suggests a unique coupling mechanism, with redox energy of electron transfer driving proton translocation via long range (up to ~200 Å) conformational changes. It resembles a steam engine, with coupling elements (akin to coupling rods) linking parts of this molecular machine. A continuous axis of basic and acidic residues running centrally through the membrane arm connects the ubiquinone reduction site in the hydrophilic arm to four putative proton-pumping units.

Cytochrome *bc₁* Complex

Many of the structure–function aspects of the cytochrome *bc₁* complex discussed above in the context of Figures 4–7

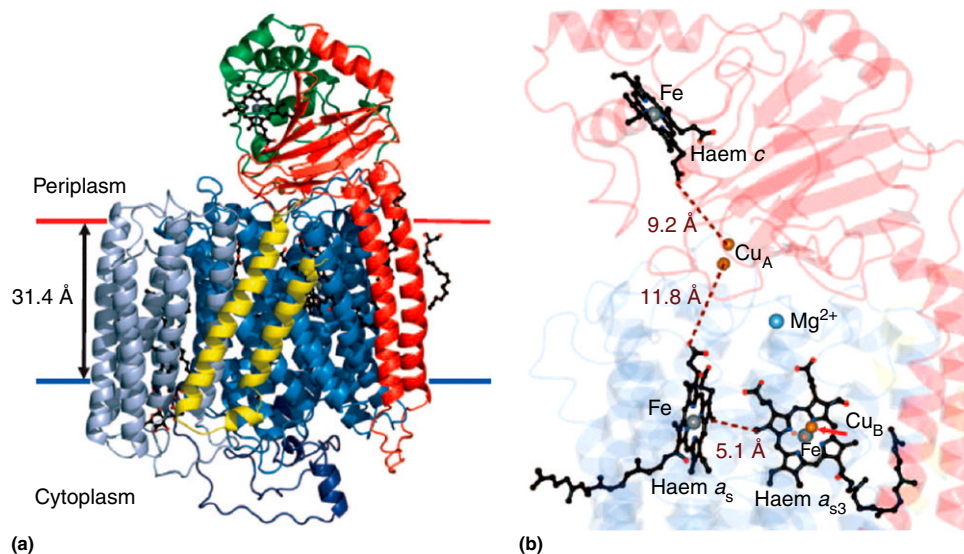


Figure 10 Structure and cofactor arrangement in cytochrome caa₃-oxidase (Lyons *et al.*, 2012). (a) Structure (ribbon model). SU I/III colored as: SU I, blue; SU III, blue-gray and the fusion linker, dark blue. SU IIc colored to highlight the classical SU II (red) and the fused cytochrome *c* domain (green). SU IV in yellow. Hemes in ball and stick with the iron and copper metal centers as gray and copper spheres, respectively. Membrane boundaries based on hydrophobic thickness calculations from the OPM server31. (b) Cofactor arrangement. Hemes *c*, *as*, and *as*₃, and the iron and copper ions shown as in (a). Mg²⁺ represented as light blue sphere. Distances, in brown. SU I/III, IIc, and IV are color-coded as faded blue, red, and yellow ribbons, respectively. Reprinted with permission from Lyons, J.A., Aragao, D., Slattery, O., *et al.*, 2012. Structural insights into electron transfer in caa₃-type cytochrome oxidase. *Nature* 487, 514–518.

apply to the similarly but, by far, not identically connected *bc*₁ complexes, which have been discussed separately for the mitochondrial (Berry *et al.*, 2013) and bacterial (Crofts *et al.*, 2008; Esser *et al.*, 2008) electron transport chains, and also compared with the *b*₆*f* complex (Cramer *et al.*, 2011).

Cytochrome Oxidase

Cytochrome *c* oxidase functions as the terminal enzymes in the respiratory chain of mitochondria and aerobic prokaryotes, coupling molecular oxygen reduction to transmembrane proton pumping. The enzyme's function is the transfer of electrons from cytochrome *c* to the oxidase via a transient association of the two proteins. The crystal structure shown in Figure 10, is the caa₃-type cytochrome oxidase from *Thermus thermophilus*, which has a covalently tethered cytochrome *c* domain (Lyons *et al.*, 2012).

Acknowledgments

The authors' studies associated with the composition of this manuscript were supported by U.S. NIH grant GM-038323, and significantly aided by discussions with Dr. S. Saif Hasan.

See also: Molecular Principles, Components, Technology, and Concepts: Membranes: The Outer Mitochondrial Membrane, a Smooth 'Coat' with Many Holes and Many Roles: Preparation,

Protein Components, Interactions with Other Membranes, Involvement in Health, Disease, and as a Drug Target. Organelles: Structure and Function: Mechanisms and Functions of Mitochondrial Dynamics

References

- Allen, J.P., Feher, G., Yeates, T.O., Komiya, H., Rees, D.C., 1987. Structure of the reaction center from *Rhodobacter sphaeroides* R-26: The cofactors. *Proceedings of the National Academy of Sciences of the United States of America* 84, 5730–5734.
- Althoff, T., Mills, D.J., Popot, J.L., Kuhlbrandt, W., 2011. Arrangement of electron transport chain components in bovine mitochondrial supercomplex I1III2IV1. *EMBO Journal* 30, 4652–4664.
- Amunts, A., Drory, O., Nelson, N., 2007. The structure of a plant photosystem I supercomplex at 3.4 Å resolution. *Nature* 447, 58–63.
- Arnon, D.I., Tsujimoto, H.Y., McSwain, B.D., 1967. Ferredoxin and photosynthetic phosphorylation. *Nature* 214, 562–566.
- Baniulis, D., Hasan, S.S., Stofleth, J.T., A., C.W., 2013. Mechanism of enhanced superoxide production in the cytochrome *b*₆*f* complex of oxygenic photosynthesis. *Biochemistry* 52, 8975–8983.
- Baniulis, D., Yamashita, E., Whitelegge, J.P., *et al.*, 2009. Structure-function, stability, and chemical modification of the cyanobacterial cytochrome *b*₆*f* complex from *Nostoc* sp. PCC 7120. *Journal of Biological Chemistry* 284, 9861–9869.
- Baradaran, R., Berrisford, J.M., Minhas, G.S., Sazanov, L.A., 2013. Crystal structure of the entire respiratory complex I. *Nature* 494, 443–448.
- Berry, E.A., De Bari, H., Huang, L.S., 2013. Unanswered questions about the structure of cytochrome *bc*₁ complexes. *Biochimica et Biophysica Acta* 1827, 1258–1277.
- Berry, E.A., Guergova-Kuras, M., Huang, L.-S., Crofts, A.R., 2000. Structure and function of cytochrome *bc* complexes. *Annual Review of Biochemistry* 69, 1005–1075.
- Birth, D., Kao, W.C., Hunte, C., 2014. Structural analysis of atovaquone-inhibited cytochrome *bc*₁ complex reveals the molecular basis of antimalarial drug action. *Nature Communications* 5, 4029.

- Capper, M.J., O'Neill, P.M., Fisher, N., *et al.*, 2015. Antimalarial 4(1H)-pyridones bind to the Q_i site of cytochrome bc₁. *Proceedings of the National Academy of Sciences of the United States of America* 112, 755–760.
- Cooley, J.W., Lee, D.W., Daldal, F., 2009. Across membrane communication between the Q_o and Q_i active sites of cytochrome bc₁. *Biochemistry* 48, 1888–1899.
- Cramer, W.A., Hasan, S.S., Yamashita, E., 2011. The Q cycle of cytochrome bc complexes: A structure perspective. *Biochimica et Biophysica Acta* 1807, 788–802.
- Cramer, W.A., Knaff, D.B., 1991. *Energy Transduction in Biological Membranes; A Textbook of Bioenergetics*. New York, NY: Springer-Verlag (Chapter 3).
- Cramer, W.A., Zhang, H., 2006. Consequences of the structure of the cytochrome b₆f complex for its charge transfer pathways. *Biochimica et Biophysica Acta* 1757, 339–345.
- Crofts, A.R., Holland, J.T., Victoria, D., *et al.*, 2008. The Q-cycle reviewed: How well does a monomeric mechanism of the bc₁ complex account for the function of a dimeric complex? *Biochimica et Biophysica Acta* 1777, 1001–1019.
- Cukierman, S., 2006. Et tu, Grothuss! and other unfinished stories. *Biochimica et Biophysica Acta* 1757, 876–885.
- Deisenhofer, J., Epp, O., Miki, K., Huber, R., Michel, H., 1984. X-ray structure analysis of a membrane protein complex. Electron density map at 3 Å resolution and a model of the chromophores of the photosynthetic reaction center from *Rhodospseudomonas viridis*. *Journal of Molecular Biology* 180, 385–398.
- Deisenhofer, J., Epp, O., Sinning, I., Michel, H., 1995. Crystallographic refinement at 2.3 Å resolution and refined model of the photosynthetic reaction centre from *Rhodospseudomonas viridis*. *Journal of Molecular Biology* 246, 429–457.
- Deisenhofer, J., Michel, H., 1989. The Nobel lecture: The photosynthetic reaction center from the purple bacterium *Rhodospseudomonas viridis*. *EMBO Journal* 8, 2149–2170.
- Dibrova, D.V., Cherepanov, D.A., Galperin, M.Y., Skulachev, V.P., Mulikidjanian, A. Y., 2013. Evolution of cytochrome bc complexes: From membrane-anchored dehydrogenases of ancient bacteria to triggers of apoptosis in vertebrates. *Biochimica et Biophysica Acta* 1827, 1407–1427.
- Esser, L., Elberry, M., Zhou, F., *et al.*, 2008. Inhibitor-complexed structures of the cytochrome bc₁ from the photosynthetic bacterium *Rhodobacter sphaeroides*. *Journal of Biological Chemistry* 283, 2846–2857.
- Esser, L., Quinn, B., Li, Y.F., *et al.*, 2004. Crystallographic studies of quinol oxidation site inhibitors: A modified classification of inhibitors for the cytochrome bc₁ complex. *Journal of Molecular Biology* 341, 281–302.
- Ferreira, K.N., Iverson, T.M., Maghlaoui, K., Barber, J., Iwata, S., 2004. Architecture of the photosynthetic oxygen-evolving center. *Science (New York, NY)* 303, 1831–1838.
- Guskov, A., Kern, J., Gabdulkhakov, A., *et al.*, 2009. Cyanobacterial photosystem II at 2.9-Å resolution and the role of quinones, lipids, channels and chloride. *Nature Structural & Molecular Biology* 16, 334–342.
- Hasan, S.S., Cramer, W.A., 2012. Lipid functions in cytochrome bc complexes: An odd evolutionary transition in a membrane protein. *Philosophical Transactions of the Royal Society of London. Series B* 367, 3406–3411.
- Hasan, S.S., Cramer, W.A., 2014. Internal lipid architecture of the hetero-oligomeric cytochrome b₆f complex. *Structure* 22, 1008–1015.
- Hasan, S.S., Proctor, E.A., Yamashita, E., Dokholyan, N.V., Cramer, W.A., 2014. Traffic within the cytochrome b₆f lipoprotein complex: Gating of the quinone portal. *Biophysical Journal* 107, 1620–1628.
- Hasan, S.S., Yamashita, E., Cramer, W.A., 2013. Trans-membrane signaling and assembly of the cytochrome b₆-lipidic charge transfer complex. *Biochimica et Biophysica Acta* 1827, 1295–1308.
- Hasan, S.S., Zakharov, S.D., Chauvet, A., *et al.*, 2014. A map of dielectric heterogeneity in a membrane protein: The hetero-oligomeric cytochrome b₆f complex. *Journal of Physical Chemistry B* 118, 6614–6625.
- Hinkle, P.C., 2005. P/O ratios of mitochondrial oxidative phosphorylation. *Biochimica et Biophysica Acta* 1706, 1–11.
- Hosler, J.P., Ferguson-Miller, S., Mills, D.A., 2006. Energy transduction: Proton transfer through the respiratory complexes. *Annual Review of Biochemistry* 75, 165–187.
- Hunte, C., Richers, S., 2008. Lipids and membrane protein structures. *Current Opinion in Structural Biology* 18, 406–411.
- Iwai, M., Takizawa, K., Tokutsu, R., *et al.*, 2010. Isolation of the supercomplex that drives cyclic electron flow in photosynthesis. *Nature* 464, 1210–1213.
- Jagendorf, A.T., Uribe, E., 1966. ATP formation caused by acid-base transition of spinach chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America* 55, 170–177.
- Jordan, P., Fromme, P., Witt, H.T., *et al.*, 2001. Three-dimensional structure of cyanobacterial photosystem I at 2.5 Å resolution. *Nature* 411, 909–917.
- Junge, W., Sielaff, H., Engelbrecht, S., 2009. Torque generation and elastic power transmission in the rotary F₀F₁-ATPase. *Nature* 459, 364–370.
- Kaback, R., 2004. The structure of the lactose permease symporter. *Biophysical Journal* 86, 4.
- Krishtalik, L.I., Tae, G.-S., Cherepanov, D.A., Cramer, W.A., 1993. The redox properties of cytochromes b imposed by the membrane electrostatic environment. *Biophysical Journal* 65, 184–195.
- Kurisu, G., Zhang, H., Smith, J.L., Cramer, W.A., 2003. Structure of the cytochrome b₆f complex of oxygenic photosynthesis: Tuning the cavity. *Science (New York, NY)* 302, 1009–1014.
- Lanyi, J.K., 2004. Bacteriorhodopsin. *Annual Review of Physiology* 66, 665–688.
- Lyons, J.A., Aragao, D., Slattery, O., *et al.*, 2012. Structural insights into electron transfer in caa3-type cytochrome oxidase. *Nature* 487, 514–518.
- Mitchell, P., 1975. The protonmotive Q cycle: A general formulation. *FEBS Letters* 59, 137–139.
- Mitchell, P., 1979. Keilin's respiratory chain concept and its chemiosmotic consequences. *Science (New York, NY)* 206, 1148–1159.
- Nitschke, W., van Lis, R., Schoep-Cothenet, B., Baymann, F., 2010. The "green" phylogenetic clade of Rieske/cyt b complexes. *Photosynthesis Research* 104, 347–355.
- Park, E., Rapoport, T.A., 2012. Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annual Review of Biophysics* 41, 21–40.
- Pogoryelov, D., Yu, J., Meier, T., *et al.*, 2005. The c15 ring of the *Spirulina platensis* F₁F₀ ATP synthase: F₁F₀ symmetry mismatch is not obligatory. *EMBO Reports* 6, 1040–1044.
- Rochaix, J.D., 2014. Regulation and dynamics of the light-harvesting system. *Annual Review of Plant Biology* 65, 287–309.
- Roszak, A.W., Moulisova, V., Reksodipuro, A.D., *et al.*, 2012. New insights into the structure of the reaction centre from *Blastochloris viridis*: Evolution in the laboratory. *Biochemical Journal* 442, 27–37.
- Sacksteder, C.A., Kanazawa, A., Jacoby, M.E., Kramer, D.M., 2000. The proton to electron stoichiometry of steady-state photosynthesis in living plants: A proton-pumping Q cycle is continuously engaged. *Proceedings of the National Academy of Sciences of the United States of America* 97, 14283–14288.
- Schutz, M., Brugna, M., Lebrun, E., *et al.*, 2000. Early evolution of cytochrome bc complexes. *Journal of Molecular Biology* 300, 663–675.
- Seelert, H., Dencher, N.A., Muller, D.J., 2003. Fourteen protomers compose the oligomer III of the proton rotor in spinach chloroplast ATP synthase. *Journal of Molecular Biology* 333, 337–344.
- Shinopoulos, K.E., Brudvig, G.W., 2012. Cytochrome b₅(5)(9) and cyclic electron transfer within photosystem II. *Biochimica et Biophysica Acta* 1817, 66–75.
- Shinzawa-Itoh, K., Aoyama, H., Muramoto, K., *et al.*, 2007. Structures and physiological roles of 13 integral lipids of bovine heart cytochrome c oxidase. *EMBO Journal* 26, 1713–1725.
- Solmaz, S.R., Hunte, C., 2008. Structure of complex III with bound cytochrome c in reduced state and definition of a minimal core interface for electron transfer. *Journal of Biological Chemistry* 283, 17542–17549.
- Stroebel, D., Choquet, Y., Popot, J.-L., Picot, D., 2003. An atypical haem in the cytochrome b₆f complex. *Nature* 426, 413–418.
- Suga, M., Akita, F., Hirata, K., *et al.*, 2015. Native structure of photosystem II at 1.95 Å resolution viewed by femtosecond X-ray pulses. *Nature* 517, 99–103.
- Vener, A.V., van Kan, P.J., Rich, P.R., Ohad, I., Andersson, B., 1997. Plastoquinol at the quinol oxidation site of reduced cytochrome b₆f mediates signal transduction between light and protein phosphorylation: Thylakoid protein kinase deactivation by a single-turnover flash. *Proceedings of the National Academy of Sciences of the United States of America* 94, 1585–1590.
- Walker, J., 1998. ATP synthesis by rotary catalysis [Nobel Lecture]. *Angewandte Chemie* 37, 5000–5011.
- Watt, I.N., Montgomery, M.G., Runswick, M.J., Leslie, A.G., Walker, J.E., 2010. Bioenergetic cost of making an adenosine triphosphate molecule in animal mitochondria. *Proceedings of the National Academy of Sciences of the United States of America* 107, 16823–16827.
- White, S.H., von Heijne, G., 2008. How translocons select transmembrane helices. *Annual Review of Biophysics* 37, 23–42.
- Widger, W.R., Cramer, W.A., Hermodson, M., G., H.R., 1985. Evidence for a hetero-oligomeric structure of the chloroplast cytochrome b-559. *FEBS Letters* 191, 186–190.
- Widger, W.R., Cramer, W.A., Herrmann, R.G., Trebst, A., 1984. Sequence homology and structural similarity between the b cytochrome of mitochondrial complex III and the chloroplast b₆f complex: Position of the cytochrome b hemes in the membrane. *Proceedings of the National Academy of Sciences of the United States of America* 81, 674–678.

- Yamashita, E., Zhang, H., Cramer, W.A., 2007. Structure of the cytochrome *b₆f* complex: Quinone analogue inhibitors as ligands of heme *c_H*. *Journal of Molecular Biology* 370, 39–52.
- Yoshikawa, S., Muramoto, K., Shinzawa-Itoh, K., Mochizuki, M., 2012. Structural studies on bovine heart cytochrome c oxidase. *Biochimica et Biophysica Acta* 1817, 579–589.
- Zhang, H., Kurisu, G., Smith, J.L., Cramer, W.A., 2003. A defined protein-detergent-lipid complex for crystallization of integral membrane proteins: The cytochrome *b₆f* complex of oxygenic photosynthesis. *Proceedings of the National Academy of Sciences of the United States of America* 100, 5160–5163.
- Zhang, H., Whitelegge, J.P., Cramer, W.A., 2001. Ferredoxin: NADP⁺ oxidoreductase is a subunit of the chloroplast cytochrome *b₆f* complex. *Journal of Biological Chemistry* 276, 38159–38165.
- Zickermann, V., Wirth, C., Nasiri, H., *et al.*, 2015. Structural biology. Mechanistic insight from the crystal structure of mitochondrial complex I. *Science (New York, NY)* 347, 44–49.

Further Reading

- Cramer, W.A., Knaff, D.B., 1991. *Energy Transduction in Biological Membranes*. New York, NY: Springer-Verlag (Chapters 1–5).
- Eisenberg, D.S., Crothers, D., 1979. *Physical Chemistry with Application to the Life Sciences*. California: Benjamin Cummings (Chapters 2–5).
- Nicholls, D.G., Ferguson, S.J., 2013. *Bioenergetics*. New York, NY: Academic Press.

NUCLEIC ACID SYNTHESIS/BREAKDOWN: RNA SYNTHESIS/FUNCTION

Contents

Transfer RNA

Messenger RNA (mRNA): The Link between DNA and Protein

The Interplay between Eukaryotic mRNA Degradation and Translation

miRNAs/Small Noncoding RNAs

Small RNAs/Cancer

Riboswitches and Ribozymes

Ribosomal RNAs and Protein Synthesis

Comparison of Bacterial and Eukaryotic Replisome Components

Transfer RNA

J Doherty and M Guo, The Scripps Research Institute, Jupiter, FL, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

aaRS Aminoacyl-tRNA synthetase include 22 members one for each amino acid in general.

AAS Amino acid accepting stem of tRNA.

AA-tRNA Aminoacyl-tRNA generated by aminoacyl-tRNA Synthetase through conjugation of amino acid with the 3'-OH of tRNA.

Amber codon UGA, a normal termination codon, but also be used to encode pyrrolysine (Pyl) or selenocysteine (Sec) in many species.

Aminoacyl-AMP Amino acyl adenosine monophosphate, a high energy reaction intermediate generated by aaRS by activating amino acid using ATP.

ASL anticodon stem loop of tRNA.

ATF4 Activating transcription factor 4 (tax-responsive enhancer element B67).

Aza-IP 5-Azacytidine-mediated RNA immunoprecipitation.

CCA-adding enzyme ATP(CTP):tRNA nucleotidyltransferases, add CCA onto the 3' end of tRNA precursors without using a nucleic acid template.

CHOP DNA damage-inducible transcript 3, also known as C/EBP homologous protein (CHOP), is a transcription factor downstream of GCN2.

Class I aaRS Class I aaRS whose aminoacylation domain has a typical Rossmann fold, including ArgRS, CysRS, GluRS, GlnRS, IleRS, LeuRS, LysRS-I, MetRS, ValRS, TrpRS, and TyrRS.

Class II aaRS Class II aminoacylation domain has 7 β -strand sheet fold, including AlaRS, AsnRS, AspRS, PheRS, GlyRS, HisRS, LysRS, ProRS, SerRS, ThrRS, and PylRS.

Class I tRNA Majority of tRNAs, contains regular variable loop region.

Class II tRNA tRNAs containing long variable loop region, include tRNA^{Leu}, tRNA^{Ser}, tRNA^{Sec}, and bacterial tRNA^{Tyr}.

Editing/proofreading domain Independent structure containing a hydrolytic site for removing the incorrect aminoacyl-tRNA, existing either in the same aaRS or in a freestanding editing enzyme.

eEF1 Eukaryotic elongation factor 1, formed by eEF1-alpha-beta-gamma. eEF1-alpha is the eukaryotic equivalent of EF-Tu; eEF1-beta-gamma are eukaryotic equivalents of EF-Ts.

eEF2 Eukaryotic elongation factor 2, is the equivalent of prokaryotic EF-G and regulates ribosome translocation.

EF-G Elongation factor G, catalyzes the recycling of ribosomes after one round of protein synthesis.

EF-Ts Elongation factor thermo stable in bacteria.

EF-Tu Elongation factor thermo unstable in bacteria.

eIF2 eukaryotic initiation factor 2, contains eIF2 $\alpha/\beta/\gamma$.

eIF2 $\alpha/\beta/\gamma$ bound to GTP delivers the initiator methionyl-tRNA (Met-tRNA^{iMet}) to the small ribosomal subunit in the first step of translation.

eIF2B Guanine exchange factor (GEF) of eIF2.

GCN2 General control nonderepressing 2 protein kinase (GCN2) as the primary sensor of amino acid starvation.

GCN4 General control nonderepressing 4 protein in yeast. GCN4 is the primary regulator in response to amino acid starvation, termed general amino acid control (GAAC). It acts as a transcription factor and activates several genes required for amino acid synthesis.

Identity element Nucleotide acid on tRNA that is responsible for the correct recognition of tRNA by the corresponding aaRS, or features that is rejected by the inappropriate aaRSs.

Initiation codon Protein synthesis in general starts with an AUG codon near the beginning of the mRNA, following the 5'-UTR, which often contains regulator elements for translation control.

Mischarged tRNA An incorrect amino acid linked to the tRNA.

mt-tRNA Mitochondrial tRNA – tRNAs that are encoded in mitochondrial genome.

N-end rule A rule states that the N-terminal amino acid of a protein determines its half-life. The rule applies to both eukaryotic and prokaryotic organisms, but with different strength.

Peptidyl-tRNA An intermediate product during peptide synthesis. It transfers the growing polypeptide to the aminoacyl-tRNA bound in the A/A site, is bound in the P/P site.

Post-editing Proofreading hydrolysis of incorrect aminoacyl-tRNA in a separate editing domain, either in the same aaRS, or in a free-standing editing enzyme.

Pre-editing Proofreading hydrolysis of incorrect aminoacyl-AMP in the same catalytic site prior to its transfer to 3'-OH of tRNA.

RNase P A type of ribonuclease which cleaves off an extra sequence of RNA on the 5' leader sequence of precursor tRNA.

RRFs Release factors (RFs) or ribosomal recycling factors (RRFs) recognize the termination codons and terminate translation.

tDNA Gene encoding tRNA.

tmRNA Transfer-messenger RNA (abbreviated tmRNA, also known as 10Sa RNA and by its genetic name SsrA) is a bacterial RNA molecule with dual tRNA-like and messenger RNA-like properties.

Termination codons Three stop codons (UAA, UAG, and UGA) signaling a termination of translation in mRNA.

TLS Certain positive-strand RNA plant viral genomes possess 3'-tRNA-like structures (TLSs) that are built quite differently from authentic tRNAs.

Wobble base pair is a pairing between two nucleotides in RNA molecules, other than the Watson–Crick base pairs. The four main wobble base pairs are guanine–uracil (G–U), hypoxanthine–uracil (I–U), hypoxanthine–adenine (I–A), and hypoxanthine–cytosine (I–C). The thermodynamic stability of a wobble base pair is comparable to that of a Watson–Crick base pair.

Wobble position A position in the 5' anticodon of tRNA. In the wobble position a G can base-pair with the two corresponding codons that have either pyrimidine (C or U) in the third position.

Basis of Transfer RNA

Transfer RNA (tRNA) is a short nucleotide RNA chain, about 80 nucleotides, which carries specific amino acids to the ribosome for addition to growing polypeptide chains during translation. tRNA functions as an 'adaptor' molecule that mediates the recognition of three nucleotide codon sequences in the messenger RNA (mRNA) and allows suitable amino acid translation of that codon. tRNAs contain a single stranded loop that base pairs with mRNA codon sequences and a CCA nucleotide sequence at the 3' end for connecting to specific amino acids. Francis Crick first proposed the hypothesis of an adaptor molecule that recognizes amino acid-related mRNA sequence and provides the appropriate amino acid for protein synthesis (Crick, 1958).

tRNA and genetic code

tRNA is the key to deciphering the genetic code in mRNA. Each amino acid has its own subset of tRNAs, which are attached to the amino acid and carry it to the growing end of a polypeptide chain. The tRNA contains a three-nucleotide sequence, an anticodon, that base pairs with its complementary codon in the mRNA (Crick, 1968; Chapeville *et al.*, 1962). tRNA anticodons are read in the 3'→5' direction to match the mRNA codons that are read in the 5'→3' direction. For example, if the anticodon base sequence is 3'-AAG-5', the corresponding codon in mRNA is 5'-UUC-3'. There are 64 potential codons (4³), 61 of these (in general) are recognized by tRNAs and encode for amino acids while 3 function as termination codons (Figure 1).

In 1966 Francis Crick proposed the wobble hypothesis to explain the observed degeneration in the third position of mRNA codons (Crick, 1966). He proposed that the 5' base of tRNA anticodon ('Wobble base') is spatially less constricted than the other two bases and can make non-standard base pairing with the 3' end of the mRNA codon. The first two bases of an mRNA codon always makes powerful Watson–Crick base pairing with tRNA anticodon while there is flexibility in the third position relieving the obligation to obey base-pairing rules. In this regard the G•U base pair is particular important since G•U structurally fits almost as well as the G•C pair. A given anticodon in tRNA with G in the 5' (wobble) position can base-pair with the two corresponding codons that have either pyrimidine (C or U) in the third position. For example, the phenylalanine codons UUU and UUC (5'→3') are both recognized by the tRNA that has AAG (3'→5') as the anticodon. Because of this, codons of NNU and NNC always encode the same amino acid and are decoded by a single tRNA with NNG anticodon (with G in the 5' wobble position).

tRNAs containing inosine in the wobble position have the greatest variability in codon recognition. Inosine is a deaminated product of adenine and therefore does not contain an amino group at the 2' carbon atom. In the 5' wobble position of tRNA anticodons, inosine can match with cytosine, uracil, or adenine. For example, a tRNA whose anticodon is CCI, will fit with GGU, GGC or GGA mRNA codons and add glycine to the growing protein chain. According to this hypothesis, 61 different anticodons (or 61 different tRNAs) are not needed to decode the 61 mRNA codons (Crick, 1966; Varani and McClain, 2000) and in fact, many cells contain fewer than 61 tRNAs.

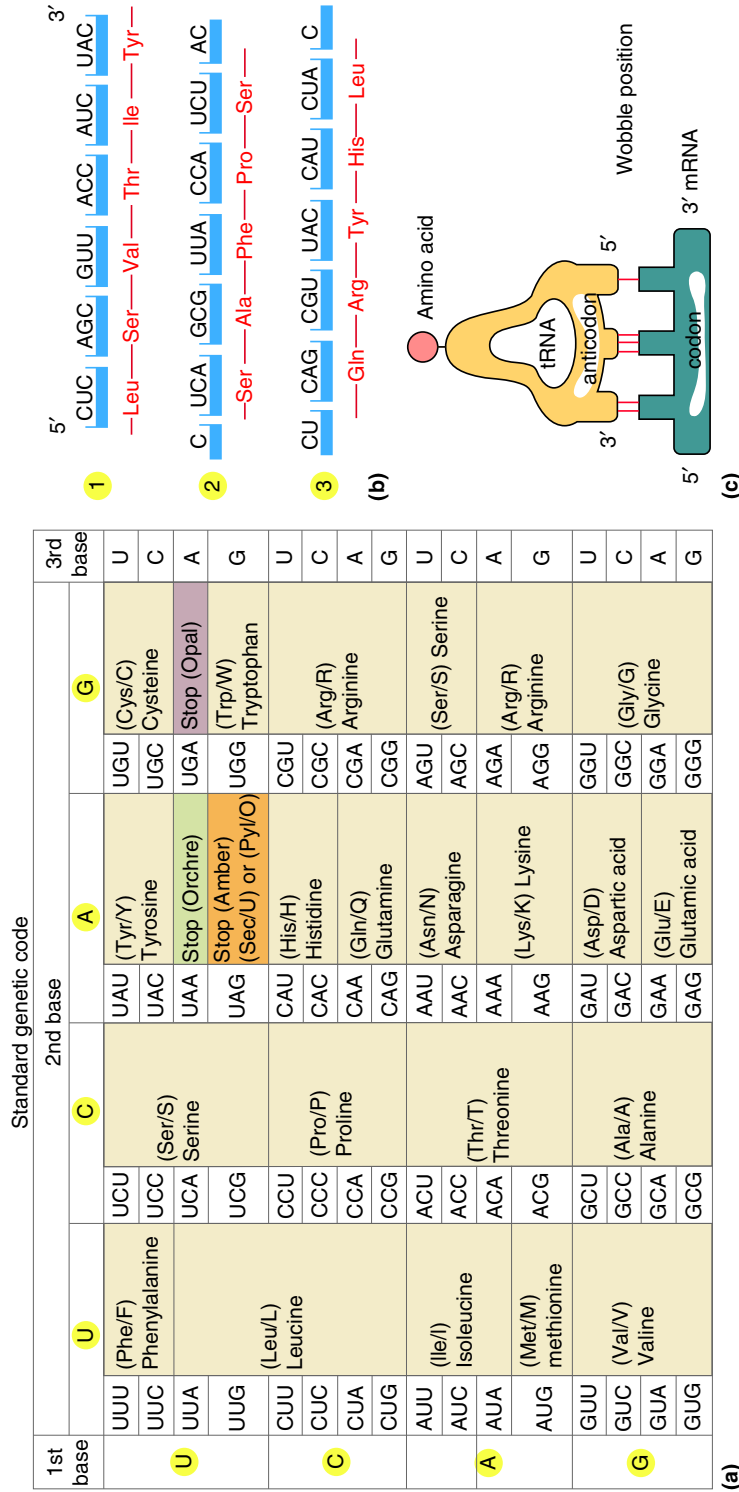


Figure 1 The genetic code. The correspondence between each of the 4³ (64) possible codon-triplets and the 20 amino acids of cellular proteome (22 in some organisms) is called the genetic code. Of the 64 codons, 61 are usually sense codons. Most of these are organized in so-called degenerated codon family boxes where synonymous triplets code for the same amino acid. Only a few codons are unassigned, usually UAA, UAG and UGA that are used as termination codons. UAG also codes for the 21st amino acid Sec (selenocysteine) or 22nd amino acid Pyl (Pyrrolysine) in some species (a). (b) The three possible reading frames in protein synthesis. In translation, the sequence of nucleotides in an mRNA molecule is read from the 5' end to the 3' end in consecutive sets of three nucleotides (blue), where in principle, the same RNA sequence can specify three completely different amino acid sequences (red). However, in most cases, only one of these reading frames is decoded by tRNA and contains the actual message. (c) The genetic code is read through a matching between anticodon of tRNA and the codon of mRNA, where 1st and 2nd position are conventional base pair. The 3rd, or wobble position is less stringent.

Figure 1 shows that most amino acids are encoded by more than one codon. Only two, methionine and tryptophan, have a single codon, while leucine, serine, and arginine each have six different synonymous codons.

The amino acid specified by each codon is conserved in most known organisms with a few exceptions. In many mitochondria, ciliated protozoans and in a single-celled plant called *Acetabularia* there is some divergence, however, in most other cases this involves reading of a normal stop codon (amber codon) as an amino acid codon and not an exchange of one amino acid for another. For example, in all three kingdoms, amber stop codon in a few mRNAs, followed by a specific sequence in the 3' untranslated region (3'-UTR, in archaea and eukarya) or in the following reading frame (bacteria) are

recognized by a specific tRNA for selenocysteine (Sec, the 21st amino acid). There are 25 proteins containing Sec in human.

A special case of codon use is found with the cytoplasmic tRNA^{Met}. Nuclear genomes of all eukaryotes contain two genes coding for two different types of tRNA^{Met} (eMet and iMet), which both harbor the UAC anticodon. One gene (often in multicopy) encodes the elongator tRNA (tRNA^{eMet}) and is used for incorporating internal methionines, while the other gene encodes the initiator tRNA (tRNA^{iMet}) and is used for initiating the synthesis of the polypeptidic chain. Although both cytoplasmic tRNA^{Met}s are charged by the same cognate Met-tRNA synthetase (MetRS), their sequences are markedly different. Indeed, the specific role of tRNA^{iMet} is highlighted by the conservation of a number of sequence features that make it

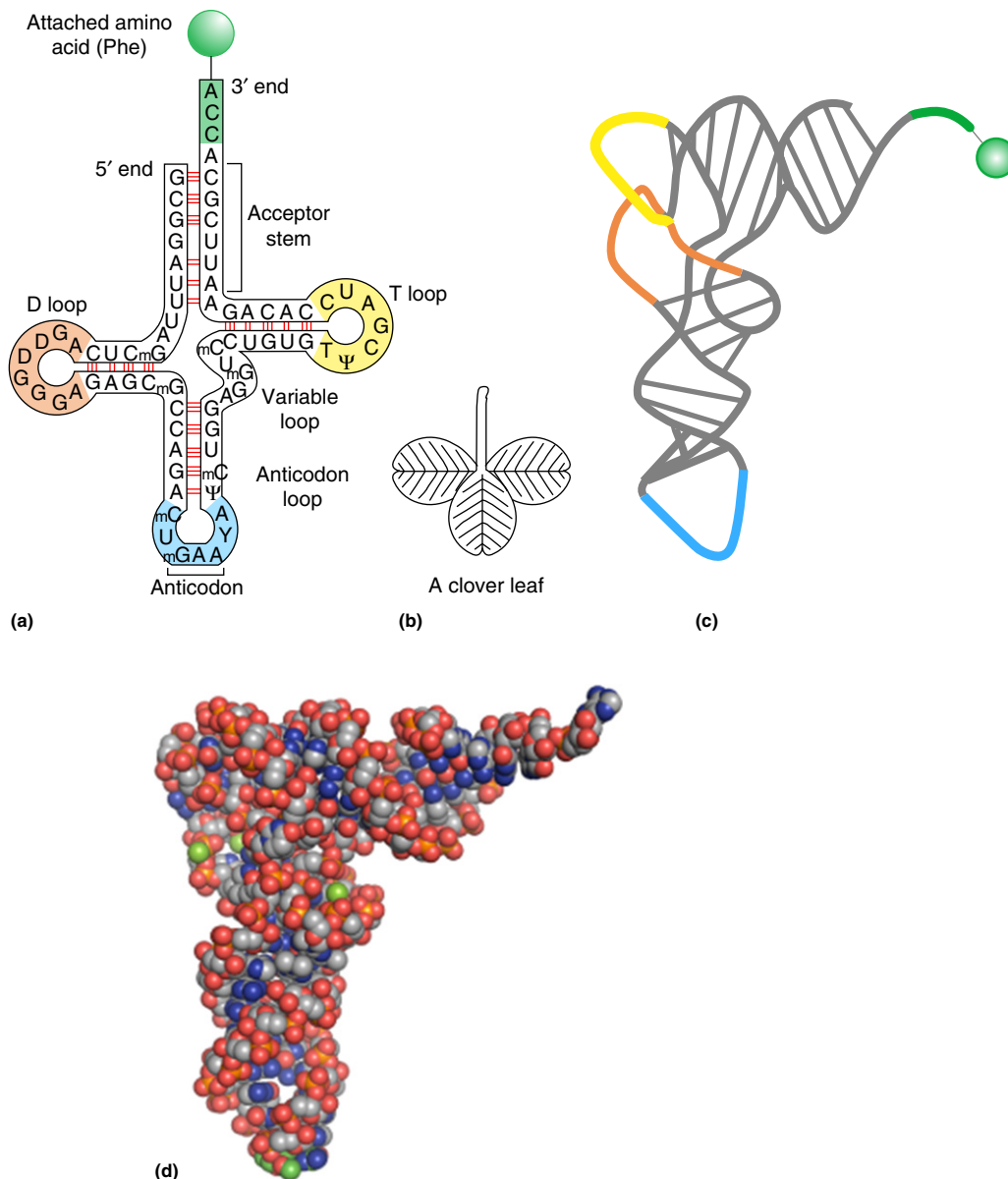


Figure 2 A tRNA molecule. The structure of yeast phenylalanyl tRNA is illustrated in open 'cloverleaf' form (a, b) to show complementary base pairing. Modified bases are indicated as mG, methylguanosine; mC, methylcytosine; DHU, dihydrouridine; T, ribothymidine; Y, a modified purine (usually adenosine); and ψ , pseudouridine. The folded form of the molecule is shown in (c) and a space-filling model in (d).

the most widely conserved tRNA throughout the three kingdoms of life (Marck and Grosjean, 2002; Grosjean *et al.*, 2010). In yeast *Saccharomyces cerevisiae*, the 2'-O-ribosyl phosphate modification at position 64 of tRNA^{Met} also serves for discrimination between the tRNA^{iMet} and the elongator tRNA^{eMet} (Astrom and Bystrom, 1994). Thus the same codon, AUG, depending on its location in the coding sequence, is used for two different purposes yet for the same amino acid. Therefore, the actual number of different codons for the 20 amino acids in any eukaryotic cell is not 61, but 61 + 1, as the two iMet and eMet codons are read by two different tRNAs.

Two Classes of tRNAs

The function of tRNA molecules is dependent on their unique three-dimensional structures. Since the first *S. cerevisiae* tRNA^{Phe} was determined by X-ray analysis (Kim *et al.*, 1974; Robertus *et al.*, 1974), nearly 200 structures of tRNAs and their fragments have been determined (Giegé *et al.*, 2012; Abe *et al.*, 2011). All tRNA molecules fold into a similar cloverleaf-like stem-loop arrangement (Figure 2). The four stems are short double helices stabilized by Watson–Crick base pairing; three of the four stems have loops containing seven or eight bases at their ends, while the un-looped stem contains the free 3' and 5' ends of the tRNA (amino acid accepting stem, AAS). The three nucleotides composing the anticodon are located at the center of the middle loop (anticodon stem loop, ASL) in an accessible position that facilitates codon-anticodon base pairing. All tRNAs are first transcribed into a precursor tRNA and several processing steps are required to convert the precursor transcript into a properly folded, functional tRNA molecule. A ribozyme (RNase P) is responsible for the generation of the mature 5'-end of tRNAs. The removal of tRNA 3' sequences is done by an endonuclease cleaving immediately 3' to the last nucleotide of the tRNA gene. All tRNAs contain a CCA sequence at the 3' end and this is usually added after synthesis and processing of the tRNA by CCA-adding enzyme (Jackman and Alfonzo, 2013; Heinemann *et al.*, 2010; Schurer *et al.*, 2001; Wolin and Matera, 1999). Several bases in tRNAs are also commonly modified after synthesis and this is important for proper folding into the tertiary structure (Giegé *et al.*, 2012; Wolin and Matera, 1999), which has a characteristic L shape with the anticodon loop and acceptor stem at each end (Figure 2).

Therefore, the canonical cloverleaf tRNA can be divided in six subdomains, 1) AAS of 7 bp (base pair) terminated at the 3' strand by $-N_{73}CCA_{OH}$ (N_{73} being the discriminator, N for any nucleotide) and at the 5' strand pN_1 (additional N_{-1} in tRNA^{His}); 2) A 2 nt (nucleotide) connector (U_8, U_9) between AAS and DSL; 3) DSL (D-stem loop of tRNA; N_{10} to N_{25}), a 4 bp stem closed by 7–11 nt D-loop (With A_{14}, R_{15} , and a variable content of Dihydrouridines located 3' or 5' of conserved $G_{18}G_{19}$); 4) ASL (N_{27} to N_{43}), a 5 bp stem and a 7-nt loop (with U_{33}, R_{37} and the anti-codon triplet N_{34}, N_{35}, N_{36} ; R for purine); 5) V (variable region; N_{44} to Y_{48}) of 4 to 24-nucleotides; 6) TSL (T-Stem Loop, T_{54} to N_{60}) with a 5 bp stem (53–61 with a conserved G-C base pair) and a 7-nt loop (with $T_{54}, \psi_{55}, C_{56}, A_{58}$).

Canonical tRNAs are divided into two classes depending on the size of the variable region (N_{44} to Y_{48} ; Figure 3). Class I tRNAs contains small variable region (4–5 nucleotides) and

includes most of the tRNAs (Giegé *et al.*, 2012; Hall *et al.*, 1989). Class II tRNAs are characterized by a large variable regions and include tRNA^{Leu}, tRNA^{Ser}, tRNA^{Sec}, and bacterial tRNA^{Tyr}. The variable region in class II tRNAs forms a stem-loop structure and ranges in size from 15 to 24 nucleotides. Further, the D-stem is 13–22 bp in class II tRNAs versus 4 bp in class I tRNAs. These large variable regions are often used for recognition by the corresponding aminoacyl-tRNA synthetases.

The L-shaped architecture of both classes of canonical tRNAs is formed by two helical domains of 12 base pair each. These two helical domains are held together by similar tertiary interaction networks at the elbow position and involve 12 conserved or semiconserved bases ($U_8, Y_{12}, A_{14}, R_{15}, G_{18}, G_{19}, R_{20}, R_{23}, Y_{48}, \psi_{55}, C_{56}$, and R_{57}). These bases form a high packed core with seven layers that corner the L-shaped structure at the D- and T-loops. The seven layers are stacked in the order of $Y_{48} \bullet R_{15}$ Levitt trans W-C pair; $A_{21} (A_8 \bullet U_{11})$ triple base; $R_{46} \bullet (N_{22}-N_{13})$; $R_9 \bullet (R_{23}-Y_{12})$ base pair $N_{24}-N_{11}$ base pair; $N_{45} \bullet (N_{10}-N_{25})$ base pair; $R_{33} \bullet R_{26}$ base pair. The D-loop and T-loop is further connected by $G_{19} \bullet G_{56}$ Watson–Crick base pair, atypical H-bonding of ψ_{55} with G_{18} and stacking of R_{57} with G_{18} and G_{19} . Note that $N_{45} \bullet N_{10}$ and $R_{46} \bullet N_{22}$ are missing in class II tRNAs.

One exception to canonical tRNAs is the tRNA for pyrrolysine (tRNA^{Pyl}), the 22nd amino acid. Pyrrolysine (Pyl) is a natural, genetically coded amino acid used by some methanogenic archaea and bacteria, with a UAG 'amber' stop codon (Krzycki, 2005). While tRNA^{Pyl} exhibits a rather canonical L-shape, tRNA^{Pyl} lacks many of the invariant structural elements conserved in the canonical tRNAs (Nozawa *et al.*, 2009). The most prominent differences include an elongated anticodon stem of six base pairs instead of five; one nucleotide at the junction connecting the acceptor and D-stem, instead of two; a short variable region of only three bases; a small D-loop of only five bases; absence of the universally conserved $G_{18}G_{19}D_{20}$ (D-loop) and $T_{54}\psi_{55}C_{56}$ (T ψ C-loop) sequences and absence of the conserved modifications, such as m^2G_{26} , m^7G_{46} and D_{47} . The resulting tertiary interactions between the non-standard D-arm, the T ψ C-loop and variable region are so different that tRNA^{Pyl} has a quite compact core structure with re-organized tertiary base pairs (Figure 3). Therefore, the structure of tRNA^{Pyl} represents an interaction pattern that is rather idiosyncratic to the tertiary structure of all other tRNAs.

Paralogous tRNA Genes/tRNA Isoacceptors

The wobble base in tRNAs and codons enables a single tRNA to recognize several codons so that although there are a maximum of 61 amino acid codons, not every codon has a unique tRNA and the total number of tRNAs in one species can be degenerated to less than 61 (this is usually seen in organelles). Vice versa, several different tRNA genes may contain the same anticodon sequence but be slightly different in the rest of the sequence. An example of the variability of tRNA gene number is demonstrated by comparing *Escherichia coli*, which has 88 tRNA genes (tDNA), with humans, which have 647 tRNA genes (625 cytoplasmic tRNA genes and 22 mitochondrial tRNA genes). Ranges from 1 to 600 tDNAs per amino acid and 0.3–200 per codon are seen across species (Table 1).

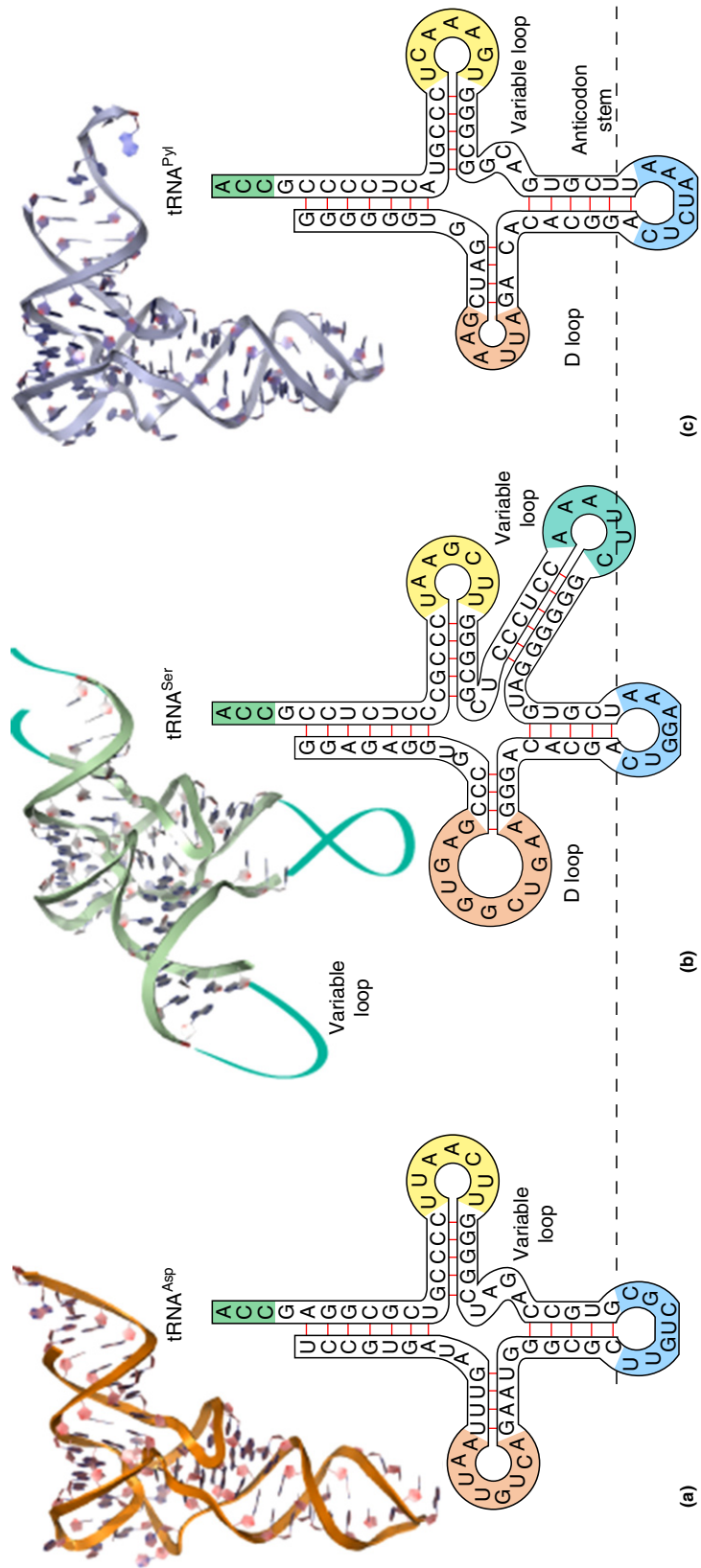


Figure 3 Two classes of tRNA. The standard L-shaped 3D-structure of class I tRNAs represents a majority of tRNAs as shown in the *Sce* tRNA^{Asp} (pdb2tra, a) versus the long various loop containing class II tRNA as shown in the *Tth* tRNA^{Ser} (pdb1ser, b). The atypical tRNA^{Pyl} represents a shape idiosyncratic to other tRNAs (pdb2zni, c).

Table 1 Statistics of number of tRNA genes in different organisms

Organism	Total number of tRNA genes	Total number of decoding amino acid	Average number of tRNA gene/codon	Average number of tRNA gene/amino acid	Maximum number of tRNA gene/amino acid	Kingdom
<i>Pyrococcus abyssi</i>	46	46	0.72	2.3	5	Archaea
<i>Escherichia coli</i>	88	87	1.36	4.14	8	Bacteria
Human mitochondria	22	22	0.34	1.1	2	Organelle
<i>Arabidopsis thaliana</i> chloroplast	38	38	0.59	1.9	4	Organelle
<i>Saccharomyces cerevisiae</i>	295	287	4.61	13.67	22	Eukarya
<i>Arabidopsis thaliana</i>	639	630	9.84	31.5	76	Eukarya
<i>Caenorhabditis elegans</i>	820	606	8.47	28.9	55	Eukarya
<i>Drosophila melanogaster</i>	304	299	4.67	14.95	26	Eukarya
Danio rerio	12 844	12 808	200.13	609.9	1583	Eukarya
<i>Xenopus tropicalis</i>	2 586	2 582	40.41	122.95	266	Eukarya
<i>Mus musculus</i>	435	434	6.80	20.67	57	Eukarya
<i>Homo sapiens</i>	625	509	9.77	24.23	43	Eukarya

Note: Total numbers of predicted genes from different genomes are retracted from the Genomic tRNA Database. It is further divided as tRNAs decoding the 20 standard amino acid and selenocysteine, suppressor tRNAs (CTA, TTA), tRNAs with undertermined or unknown isotypes, and pseduogenes. The human mitochondrial tRNA genes were cross-checked with mitoRNAdb and Mamit-tRNA. The plant *Arabidopsis thaliana* chloroplast tRNA genes were retracted from tRNADB-CE.

Genomic analyses have greatly advanced our knowledge on the number of tRNA genes and there are several good resources available to explore this information. The tRNomics project has analyzed more than 4000 sequences of cytosolic tRNAs from 50 genomes (Marck and Grosjean, 2002). The Bayreuth database of tRNAdb system (see Relevant Websites) contains more than 12 000 tRNA genes from 577 species and 623 curated tRNA sequences from 104 species, as well as more than 30 000 mitochondrial tRNA genes from over 1500 metazoan species. The GtRNAdb Genomic tRNA database (see Relevant Websites) contains annotation of 797 genomes from Eukarya (82 taxa), Bacteria (629 taxa), and Archaea (86 taxa) (Chan and Lowe, 2009). Another comprehensive database is tRNADB-CE (see Relevant Websites) that provides annotations of 595,115 tRNA genes from more than 7000 genomes, except Metazoa (Chan and Lowe, 2009). Other information on mitochondrial tRNA (mt-tRNAs) from mammals (see Relevant Websites) and tetrapods is also available from specialized databases (Putz et al., 2007; Popadin et al., 2007).

Analysis by tRNAscan-SE indicates that the reference human genome contains 625 cytoplasmic tRNA genes (Abe et al., 2011). Among these, 446 genes (280 different sequences) are predicted to fold into the canonical cloverleaf structure with zero or at most one mismatch in the stem; 167 genes (160 different sequences) are predicted to have more than one mismatch in the stems (Parisien et al., 2013). Since a mature tRNA contains 74–93 nucleotides, the entire repertoire of human tRNA sequence is made of ~60 000 base pairs representing nearly 0.02% of the genome. Considering that nearly all tRNA genes contain extra sequences that are required to be trimmed off during processing and 32 of them have introns, the total size of tRNA genes in the genome is even more than this number.

The recent 1000-genomes project, which was designed to examine human genetic variation by sequencing a large number of people, demonstrates the great sequence diversity of tRNA genes in the human genome (Parisien et al., 2013). This project also identified new tRNA sequences that are only encoded in a few percent or less of the human population; 76 new tRNA genes exist in >0.2% of all individuals. Interestingly, some tRNA genes encode for tRNAs that could not be aminoacylated, while other genes unexpectedly contained base-pair mismatches in the tRNA structure, suggesting that these genes produced inactive tRNA molecules or tRNAs with new, non-canonical, extra-translation functions. Together this highlights the evolutionary changes occurring in the human population and suggests that these changes in tRNAs may be linked to human complexity (Parisien et al., 2013).

Degenerative tRNAs in Mitochondria

Despite the cloverleaf structure that exists in most tRNAs, tRNAs with atypical cloverleaf structures have been identified. For example, the variation of base pair number in the stem regions of tRNA^{Sec}, tRNA^{Pyl}, and the degenerative sequence in mt-tRNAs.

A major group of tRNAs that deviate from the canonical structure of tRNAs are mitochondrial tRNAs (mt-tRNAs). Mitochondrial genomes encode most of the tRNAs to be used in mitochondrial translation (Kurland, 1992). Due to the evolutionary pressure for minimizing the mitochondrial genome size, many eukaryotes contain only a minimal set of mt-tRNA genes, while in others the set is partial or totally absent from the mitochondrial genome and requires importing missing tRNAs from the cytosol. Large-scale sequencing of mt-genomes has compiled 22 mt-tRNA genes in mammals with

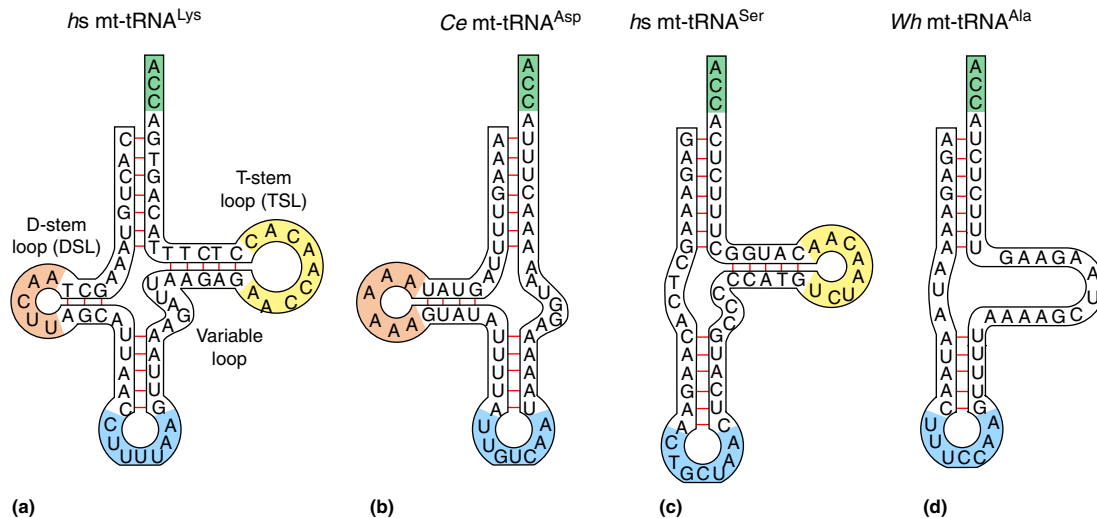


Figure 4 Structure of mitochondrial tRNAs. The mitochondrial tRNAs show divergent structures with atypical (a) or resected (b–d) cloverleaf folds. (a) human (*Hs*) mt-tRNA^{ASP} with atypical D- and T-loops. (b) *Caenorhabditis elegans* (*Ce*) mt-tRNA^{ASP} missing TSL. (c) human mt-tRNA^{Ser} missing DSL. (d) *Walchia hayashii* (*Wh*) mt-tRNA^{Ala} missing both TSL and DSL. Detailed illustration of diverse mt-tRNA structures can be found at Mamit-tRNA (see Relevant Websites section) and was reviewed elsewhere (Giegé *et al.*, 2012).

one per amino acid, two for leucine and serine isoacceptors, respectively. A characteristic feature of mt-tRNAs is the preferential use of light nucleotides (A, U, and C-rich tRNAs) in most Metazoa, especially in Nematoda and Arthropoda (Helm *et al.*, 2000; Klimov and Oconnor, 2009; Yuan *et al.*, 2010). The low GC content is due to the absence of signature motifs (e.g., G₁₈G₁₉ in D-loop) and other variations in D-loop and T-loop that are standard in canonical tRNAs (Figure 4). Particularly, the conserved residues in the D-loop and T-loop in cytosolic tRNAs are often missing in the mt-tRNAs. Some mt-tRNAs even have resected cloverleaf structures, missing the entire DSL or TSL or both. In fact, all mammalian mt-genomes encode a D-armless tRNA^{Ser}, tRNA^{Ser(GCU)}, and a second serine isoacceptor, tRNA^{Ser(UCA)} of atypical size (Putz *et al.*, 2007). Larger deletions of structural domains are popular in nematode mt-tRNAs such as *Caenorhabditis elegans*, including a systematic absence of cloverleaf arm TSL (T-stem loop) or DSL for nearly all tRNAs (Wolstenholme *et al.*, 1987). tRNAs with extremely short length (44–65 nucleotides) lacking either the TSL or DSL or both are also found in Arthropoda mt-genomes. Despite the extensive degeneration in sequence and progressively simplified core domain, all mt-tRNAs nonetheless are expected to fold into a similar canonical tRNA L-shape for functioning in protein synthesis.

This degenerative feature of mt-tRNAs suggests that alternative scaffolds of tRNAs could exist as long as they maintain the basic function of tRNAs. On the other hand, even the degenerative tRNAs still share the same global structural features including the similar overall size, L-shaped 3D structure, anticodon stem loop, acceptor stem, and ~70 Å distance between the anticodon to terminal Adenosine, (for interacting with aminoacyl-tRNA and ribosome). Other molecules that mimic the function of tRNAs also share this global structure, as seen in the case of viral tRNA-like structure (tmRNA), translation termination recycling factor, and other tRNA mimics.

Aminoacyl-tRNA Synthetases

Decoding the genetic code from mRNA to protein includes two steps, attachment of the amino acid to tRNA and the pairing of tRNA with mRNA codon. This decoding is established at the first step of translation by enzymes called aminoacyl-tRNA synthetases, which were discovered by Paul Zamecnik and Mahlon Hoagland in 1957 (Hoagland *et al.*, 1957). Recognition of the codon on mRNA by a particular aminoacyl-tRNA happens at the ribosome. Therefore, the tRNA synthetases occupy a special position in molecular biology and the fidelity of the genetic code.

tRNA Aminoacylation

Aminoacyl-tRNA Synthetase (aaRSs) enzymes are highly conserved proteins that exist in every single cell. There are in general 20 aaRSs that each corresponds to one amino acid (Schimmel and Soll, 1979). The 20 aaRSs are thought to have arisen early in evolution, amongst the first proteins to appear in transition from the primordial RNA world to the theatre of proteins (Woese *et al.*, 2000). They can be divided into two classes, each class containing about 10 members (Figure 5). The Class I aaRSs access the minor groove of tRNA and transfer the amino acid onto the 2' -OH group on the ribose of the terminal adenosine (A76); the Class II aaRSs access the major groove of tRNA and transfer the amino acid onto the 3'-OH on the ribose of A76. The exception is PheRS, which is a Class II enzyme but charges the amino acid to 3'-OH of A76 (Ruff *et al.*, 1991; Eriani *et al.*, 1990). All aaRSs catalyze the reaction in two steps; in the first step, ATP and amino acid bind to the aaRS and the carboxyl group of the amino acid is activated by forming a high energy aminoacyl-AMP (adenosine monophosphate) bond and pyrophosphate (PPi); in the second step, the amino acid carboxyl group is further transferred (charged) to the 2' or 3'-OH group of the ribose at the

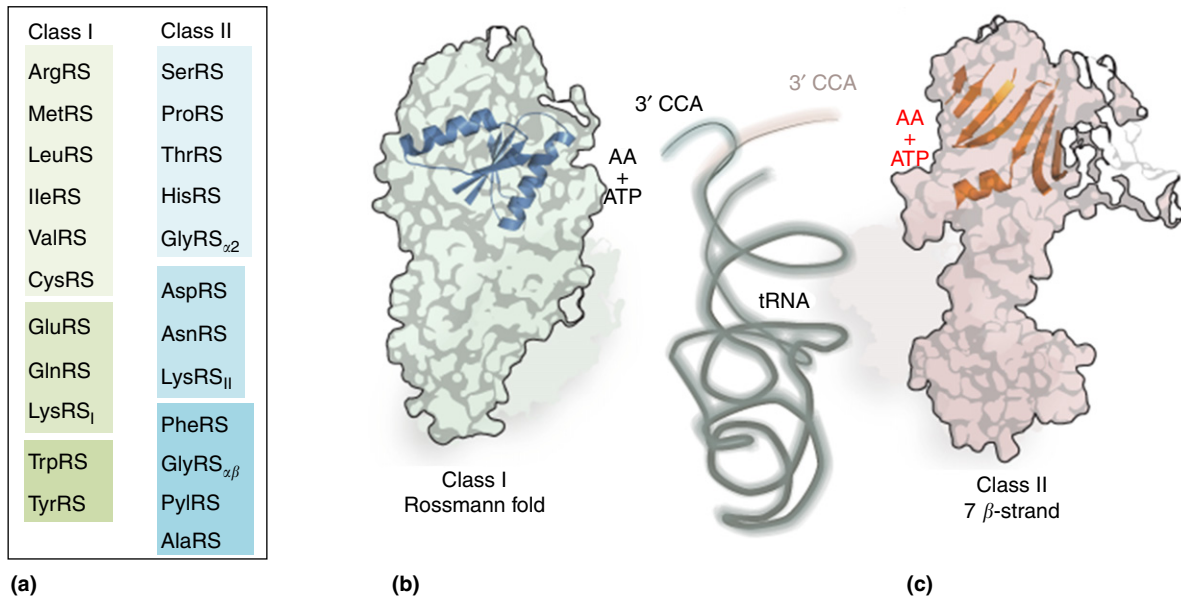


Figure 5 Two classes of aminoacyl-tRNA synthetases. (a) The specific ligation of 22 amino acids is carried out by the 21 aminoacyl-tRNA synthetases (including PylRS that is not universally present. tRNA^{Sec} is initially charged by SerRS with Ser, then converted to Sec-tRNA^{Sec}), that equally divided into two classes based on their distinct evolutionary origins. The LysRS contains both class I (α-proteobacteria) and class II (most organisms). (b) Class I aminoacylation domain has a typical Rossmann fold. (c) Class II aminoacylation domain has 7 β-strand sheet fold, which in general bind to the opposite sides of the acceptor stem helix of tRNA.

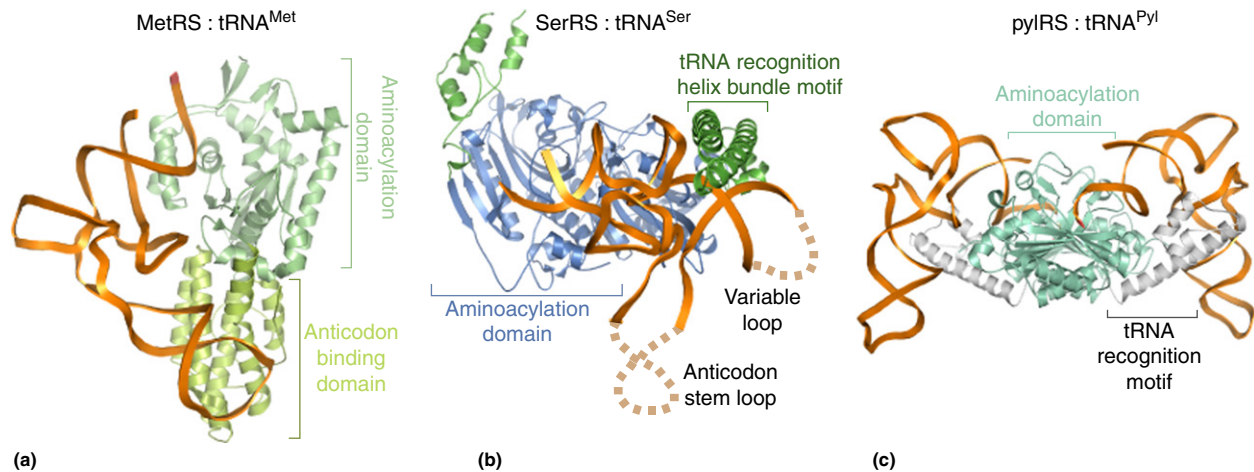
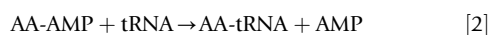
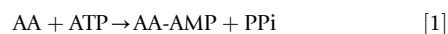


Figure 6 Recognition of tRNAs by aaRSs. (a) The standard aaRS:tRNA complex recognizes the anticodon loop and acceptor stem end of the cognate tRNA as shown in the *Aquifex aeolicus* MetRS:tRNA^{Met} (pdb2csx) versus (b) the non-anticodon recognition mode of tRNA as shown in the *Thermus thermophilus* SerRS:tRNA^{Ser} (pdb1ser) and (c) in the *Desulfitobacterium hafniense* PylRS:tRNA^{Pyl} (pdb2zni).

terminal 3' end of tRNA (A76) and AMP is released (First, 2005).



Combined: $\text{AA} + \text{ATP} + \text{tRNA} \rightarrow \text{AA-tRNA} + \text{PPi}$ (Harms *et al.*, 2001).

While all aaRSs follow this two-step reaction in a consecutive independent manner, some aaRSs, namely GlnRS,

GluRS, and ArgRS, catalyzes the first step reaction only when the tRNA is bound to the enzyme.

Identify Element of tRNAs

The fidelity of protein synthesis depends on specific tRNA aminoacylation by aminoacyl-tRNA synthetase enzymes. This in turn depends on the recognition of particular nucleotides and structural features in the cognate substrate tRNA. For this reason, each aaRS enzyme is specific for one amino acid and

one or more isoaccepting tRNAs. The general scheme of interaction between aaRS and tRNA is seen in Figure 5. In the complex of *E. coli* tRNA^{Gln} with GlnRS, tRNA^{Gln} is bound to the major groove side of the acceptor stem helix, with its anticodon loop being directly recognized by the anticodon binding domain of GlnRS (Rould *et al.*, 1991). The 3' C₇₄C₇₅A₇₆ end of tRNA^{Gln} turns toward the inside of the catalytic cavity in GlnRS to access the bound L-Gln and ATP. The discriminator base (N₇₃) provides the sharp turn and also makes contact with the catalytic domain of GlnRS. Other interactions also occur at the acceptor stem and the anticodon stem. A similar binding mode was also found in other class I aaRSs, such as MetRS (Figure 6). The complex of yeast tRNA^{Asp} with Class II AspRS occurs in a mirrored pattern, with similar interactions to the anticodon loop, the CCA end and the stems, however, with AspRS bound to the minor groove side of the tRNA^{Asp} acceptor helix (Ruff *et al.*, 1991). This mirrored interaction between Class I and Class II aaRSs with tRNAs prompted the hypothesis that the two classes evolved from a single primordial aaRS gene, or a tRNA chaperon protein, that was transcribed in both direction to produce two 'mirror image' aaRS proteins (Ribas de Pouplana and Schimmel, 2001).

Surprisingly, the identity elements of tRNAs (i.e., the features recognized by the appropriate aaRS, but rejected by inappropriate aaRSs) are not restricted to the anticodon. In some cases the anticodon is completely unrecognized by the aaRSs. These features generally reside within the acceptor helix, the long variable loop, and the anticodon stem-loop of the tRNA.

As an example, a typical recognition of tRNA by aaRSs can be found with aspartic acid tRNA (tRNA^{Asp}) and yeast AspRS. *S. cerevisiae* AspRS was the first synthetase that yielded crystals when complexed with tRNA (Ruff *et al.*, 1991; Giegé *et al.*, 1980). The fidelity of correct aminoacylation on tRNA^{Asp} is insured by a series of interactions that span from the anticodon to the acceptor end of the tRNA. Specifically these are governed by direct base interactions in the anticodon, as well as backbone-mediated interactions within the entire acceptor helix of tRNA^{Asp}. All three bases from the tRNA anticodon are universal aspartate identity determinants of the tRNA^{Asp} and are recognized by AspRS. Anticodon Q34 (in *E. coli*, G34 in yeast) and U35 are the strongest identity determinants (Giegé *et al.*, 1998; Putz *et al.*, 1991), making direct interactions with the N-terminal anticodon binding domain of AspRS and a single mutation affecting the interaction with G34 inactivates the enzyme (Ador *et al.*, 1999). The G73 discriminator base in tRNA^{Asp} serves as another major aspartate identity element, particularly as mutation of this base results in mischarging of the tRNA (Ruff *et al.*, 1991; Putz *et al.*, 1991). Direct proof that G73 is a discriminator base was later demonstrated with the complex structure of yeast AspRS: tRNA^{Asp}. In this structure, both bases with the strongest identity determinant, G73 and C74, are specifically recognized by the active site of AspRS. The 3'-end nucleotide A76 of the CCA-acceptor extremity is also secured in the active site by hydrogen bonding with the 2'-OH of the ribose, placing the ribose in the right position for binding the aspartyl group on O3' (Moulinier *et al.*, 2001). Similar anticodon recognition and discriminator recognition of major identity elements is found in most aaRSs, such as MetRS, GluRS, GlnRS, CysRS,

TrpRS (Class I) as well as LysRS, GlyRS, AsnRS, ProRS, etc. (Class II).

Conversely, other aaRSs recognize the cognate tRNAs based on different identity features in addition to, and sometimes completely independent of, the anticodon and discriminator sequence. For example, HisRS also recognizes the unique G-1 nucleotide at the 5'-end of tRNA, which is only present in histidine tRNAs (Rosen and Musier-Forsyth, 2004; Himeno *et al.*, 1989; Yuan *et al.*, 2011).

In the alanine system, fidelity is ensured by a G•U base pair located at the third position (G3•U70) within the acceptor helix of alanine tRNA (tRNA^{Ala}). Genetic, biochemical and biophysical data suggest that the G•U pair provides a distinctive structure that is directly recognized by the AlaRS to position the tRNA acceptor end into the active site of AlaRS (Gabriel *et al.*, 1996). In fact, AlaRS could aminoacylate an acceptor stem-loop moiety of tRNA^{Ala} (minihelix) that is completely deprived of the anticodon stem-loop (Beuning *et al.*, 2002). Further, introduction of the G₃:U₇₀ base pair into other tRNAs enables efficient aminoacylation by AlaRS (Lovato *et al.*, 2004). Structural analysis of AlaRS:tRNA^{Ala} indicates that AlaRS only interacts with the upper helix of the L-shaped tRNA^{Ala}, the elbow region of D- and T-loops, with no contact to the bottom helix and the anticodon (Sokabe *et al.*, 2009; Naganuma *et al.*, 2009; Guo *et al.*, 2009). AlaRS recognizes the G3•U70 by directly recognizing the base pair from both major and minor grooves, by which it controls the 3'-CCA end to reach the catalytic site in a productive form (Naganuma *et al.*, 2014). These findings spawned the hypothesis that before the formation of the modern genetic code, the primordial tRNA evolved from duplication of a simple helical stem-loop (minihelix) that was promiscuous for amino acid charging (Schimmel *et al.*, 1993).

The evolution of tRNA synthetases reveals that they all contain an ancient catalytic domain positioned at the root of all life followed by the addition of a second domain later in evolution, which in many cases interacts with the anticodon of its cognate tRNA. This suggests that early tRNA recognition was based in large part on nucleotide determinants in the tRNA acceptor stem and this was followed by the development of a system of recognition based on the anticodon loop. This development is considered as a second genetic code (de Duve, 1988). In support of this, the major recognition elements of *Methanococcus jannaschii* tRNA^{Tyr} are the discriminator base A73 and the first C₁•G₇₂ base pair in the acceptor stem, whereas the anticodon triplet participates only weakly in identity determination (Fechter *et al.*, 2001; Steer and Schimmel, 1999). For that reason, the anticodon triplet can be changed without changing the aminoacylation activity of *M. jannaschii* TyrRS (Wang and Schimmel, 1999). For example the anticodon of *M. jannaschii* tRNA^{Tyr} can be converted to the amber suppressor anticodon sequence CUA (or any other anticodon of interest) so that the mutant *M. jannaschii* TyrRS will now insert Tyr during translation of the amber stop codon UAG in mRNA. This finding has led to the beginning of an exciting new field of protein engineering by site-specific incorporation of unnatural amino acid (reviewed elsewhere Xie and Schultz, 2005).

There are other examples of tRNAs that are charged without direct recognition of the anticodon and that include tRNA^{Leu},

tRNA^{Pyl}, tRNA^{Ser} and tRNA^{Sec}. The tRNA^{Ser} adopts a tripod-like structure with a unique long variable hairpin-loop protruding out ($\sim 45^\circ$) from the plane of the L-shaped tRNA (Figure 6). SerRS uses a special two-helical bundle to recognize this variable region as a major determinant, with no interaction between SerRS and the anticodon of tRNA^{Ser} (Biou *et al.*, 1994). This pattern of recognition may be a requisite for SerRS to recognizing multiple tRNA^{Ser} isoacceptor for equal reading of the six serine codons. SerRS is also required for translation of the 7th codon, UGA termination codon on tRNA^{Sec}. The synthesis of tRNA^{Sec} requires it first to be charged with Ser by SerRS and then converted into selenocysteine, the 21st amino acid (Bock *et al.*, 1991). Similarly, tRNA^{Sec} also contains a tripod-like structure with a long variable hairpin recognized by SerRS (Kryukov *et al.*, 2003). This requirement for multiple isoaccepting tRNAs, is in sharp contrast with the previous case of AspRS, where aspartate specificity involves a single tRNA^{Asp}.

The idiosyncratic structure of tRNA^{Pyl} also allows a unique recognition pattern by PylRS. Distinct from all other canonical tRNAs, tRNA^{Pyl} has a compact core structure with re-organized tertiary base pairs (Figure 6). PylRS mainly recognizes tRNA^{Pyl} through this unusual core (Nozawa *et al.*, 2009). From the inner side of the elbow of tRNA^{Pyl}, the C-terminal tRNA binding domain of PylRS makes extensive contacts with the sugar backbones of the D-stem base pairs (G₁₀•C₂₅ to C₁₃:C₂₂), the T ψ C-loop from the minor groove side. It also forms stacking interaction with a 'flipped out' guanine base of G9 in the tRNA^{Pyl}-specific unusual minimal core (Nozawa *et al.*, 2009). PylRS also recognizes the special U-shaped acceptor helix structure of tRNA^{Pyl}. In contrast, the anticodon arm of tRNA^{Pyl} does not make any specific contacts with PylRS. Structural analysis suggests that PylRS is probably one of the earliest evolved aaRSs among the total 21 aaRSs (Kavran *et al.*, 2007) and in this and other cases the anticodon does not participate in amino acid identity, such as AlaRS. The exceptional recognitions for these aaRSs are found throughout all three kingdoms of life, further suggesting the essential role

of aaRSs in establishing the genetic code in the initial step of evolution.

Proofreading of Aminoacyl-tRNAs

The similarity of some amino acids can result in the mischarging of tRNAs and these errors are corrected by either reversal of the charging or by modifying the amino acid to the correct form. The earliest hint of mischarging of tRNA by non-cognate amino acid came from the fact that, although there are 20 standard amino acids, many organisms do not possess all 20 aaRSs. This was first observed for *Bacillus subtilis* when it was shown that Gln-tRNA^{Gln} is formed via the transfer of an amino group onto Glu-tRNA^{Gln} (Kunst *et al.*, 1997; Curnow *et al.*, 1997b). Later, it was demonstrated that both tRNA^{Glu} and tRNA^{Gln} are charged with glutamate by a single GluRS. The amidotransferase that subsequently converts Glu-tRNA^{Gln} to Gln-tRNA^{Gln} has been characterized and shown to co-bind with tRNA^{Gln} together with the non-discriminating GluRS thus preventing leaking of the intermediate Glu-tRNA^{Gln} to translation (Curnow *et al.*, 1998; Curnow *et al.*, 1997a). A similar situation was later found in gram-positive eubacteria, archaea, chloroplasts, as well as mitochondria. A similar misacylation-transamidation pathway to produce Asn-tRNA^{Asn} was also found in the halophilic archaea (Min *et al.*, 2002; Kim *et al.*, 1996). In these examples a system has evolved to make use of the mischarging by GluRS and AspRS eliminating the need for the GlnRS and AsnRS genes.

The correct recognition of amino acids by modern aaRSs is essential for the accuracy of translation and uncorrected mischarging events will result in mistranslation leading to corruption of the proteome. However, the differences between similar amino acids are not large enough for the enzymes to rigorously discriminate (Pauling, 1958). To achieve higher discrimination, amino acid recognition is split into two-steps that involve a partially accurate synthetic step followed by a proofreading or editing step that eliminates misactivated

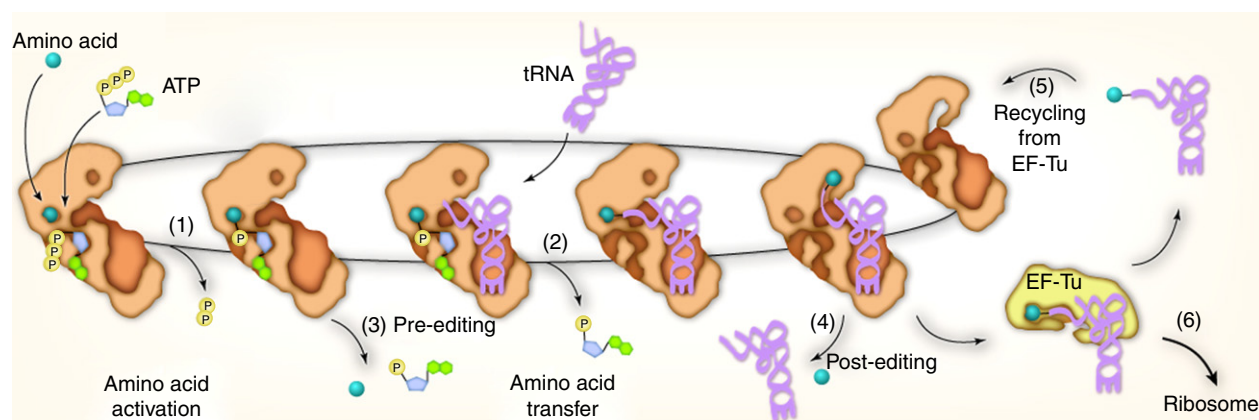


Figure 7 Pre-transfer and post-transfer editing of mischarged aa-tRNAs. A cognate amino acid (blue) is activated at the aminoacyl-tRNA synthetase (aaRS) active site with ATP to form an aminoacyl adenylate (1), releasing pyrophosphate (PPi). Subsequently, a cognate tRNA binds to the aaRS, and the amino acid is attached to the 3' end of the tRNA and AMP is released, forming an aminoacyl-tRNA (2). The non-cognate amino acid could be either removed before transferring to tRNA (3) or after (4). The aminoacyl-tRNA is then released and binds to elongation factor Tu (EF-Tu). In some cases, the EF-Tu bound aminoacyl-tRNAs could also be further checked by recycling back to the empty aaRS (5). The proofreading by free-standing editing enzymes may also happen at this recycling circle. Final aminoacyl-tRNA that passes the quality control steps will be delivered to ribosome for protein synthesis.

noncognate amino acids. There are two checkpoints to monitor the fidelity of amino acid/tRNA charging (Figure 7); (1) After the AMP activation of amino acids (AA-AMP, pre-transfer, reaction (3)); (2) After the charging of tRNA (AA-tRNA, post-transfer, reaction (4)). The mis-activated aminoacyl-AMPs are cleared at the synthetic site – pre-editing; or the mischarged tRNAs are cleared by hydrolytic editing (Fersht, 1977; Jakubowski and Goldman, 1992; Ibba and Soll, 1999; Ling *et al.*, 2009), at a second, discrete active site within an editing domain – post-editing (Sokabe *et al.*, 2009; Cusack *et al.*, 2000; Fukai *et al.*, 2000; Ahel *et al.*, 2003; Wong *et al.*, 2003; An and Musier-Forsyth, 2004; Fukunaga and Yokoyama, 2005; Sokabe *et al.*, 2005). For example, IleRS is not able to discriminate between isoleucine and valine; however, valine-tRNA^{Ile} can be recognized by the editing domains and hydrolyzed (reaction (3)).



These editing domains have been identified for the class I IleRS, ValRS, and LeuRS and for the class II AlaRS, ThrRS, PheRS, and ProRS (Sokabe *et al.*, 2009; Naganuma *et al.*, 2009; Cusack *et al.*, 2000; Fukai *et al.*, 2000; Ahel *et al.*, 2003; Wong *et al.*, 2003; An and Musier-Forsyth, 2004; Fukunaga and Yokoyama, 2005a,b,c; Sokabe *et al.*, 2005; Tukalo *et al.*, 2005; Crepin *et al.*, 2006; Hussain *et al.*, 2006). Early structures of ValRS/IleRS/LeuRS re-capitulated the double-sieve (two active sites) model proposed by Fersht for the three close-related class I aaRSs (Fukai *et al.*, 2000). If the wrong amino acid becomes attached to a tRNA, the bound synthetase catalyzes removal of the amino acid from the tRNA, through a separated editing domain that specifically recognizes the mis-linked amino acid (CP1 domain). This process requires the flip-over of the 3' CCA end of the tRNA to present its linked amino acyl-group to the editing domain, and is called post-editing reaction. After dissociating from aaRSs, the aminoacyl-tRNA will be bound by elongation factor and delivered to ribosome for translation.

In addition to the immediate editing steps at the aaRS, the released AA-tRNAs that have already been bound by elongation factor will also sometimes recycle back to the tRNA

synthetase for a third check and hydrolysis if it is inappropriate. One example shown for the *E. coli* PheRS was discovered in *in vitro* translation system, where mischarged Tyr-tRNA^{Phe} could be dissociated from the elongation factor EF-Tu and rebound to the PheRS. PheRS will then double check the mischarged Tyr-tRNA^{Phe} and hydrolyzes it through its editing domain (Ling *et al.*, 2009). Although elongation factor EF-Tu, shows some preference for correctly charged AA-tRNAs, it binds efficiently to mischarged AA-tRNAs as well. This can be exemplified by the extreme cases where hundreds of different non-natural amino acids have been successfully incorporated into protein by engineering the specificity of tRNA synthetases (Xie and Schultz, 2005; Malyshev *et al.*, 2014; Wang *et al.*, 2007). Thus the elongation factors are not major contributors to editing of mischarged tRNAs.

Finally in addition to editing by the synthetases, there is a separate set of free-standing editing enzymes that bind and hydrolyze mischarged tRNAs that leak from the aaRS. These enzymes show homology to the editing domains of tRNA synthetases and as examples include AlaXp (AlaXp-I, -II, -III, for tRNA^{Ala}), ThrX (for tRNA^{Thr}), ProX and Ybak (for tRNA^{Pro}) (Fukunaga and Yokoyama, 2007; Chong *et al.*, 2008; Ruan and Soll, 2005; Murayama *et al.*, 2005). They are widely present in multiple species throughout the three kingdoms of life (Guo *et al.*, 2009; Ahel *et al.*, 2003; Ruan and Soll, 2005; Korencic *et al.*, 2004) and provide yet another layer for proofreading and removal of mischarged tRNAs.

In all, these proofreading mechanisms highlight the importance of high fidelity in translating the genetic code and the essential role that the tRNA synthetases, by supplying the correct AA-tRNA for peptidyl formation, play in this fundamental process.

tRNA in Translation

Translation of mRNA occurs at ribosomes where two different ribosomal subunits assemble on mRNA, a small subunit (30S in prokaryotes and 40S in eukaryotes) and a large subunit (50S in prokaryote or 60S in eukaryote). A typical eukaryotic cell contains millions of ribosomes in its cytoplasm (Farrell, 2010). The complete ribosome contains a binding site for

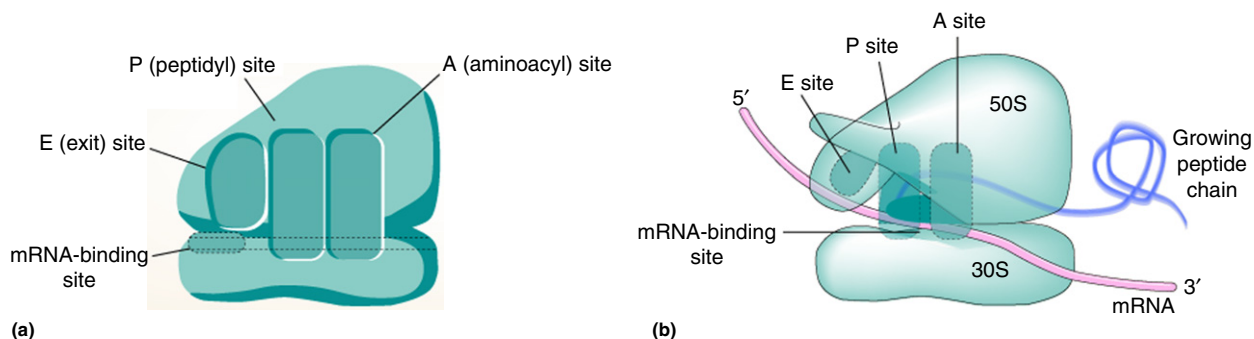


Figure 8 Three tRNAs bound in ribosome for protein translation. (a) Each ribosome has one binding site for mRNA and three binding sites for tRNA: the A-, P-, and E-sites (short for aminoacyl-tRNA, peptidyl-tRNA, and exit, respectively). (b) Each amino acid added to the growing end of a polypeptide chain is selected by complementary base-pairing between the anticodon on its attached tRNA molecule and the next codon on the mRNA chain. Because only one of the many types of tRNA molecules in a cell can base-pair with each codon, the codon determines the specific amino acid to be added to the growing polypeptide chain.

mRNA and three binding sites for tRNA called the A-site, P-site, and the E-site (**Figure 8**).

A-Site (Aminoacyl Site): It is the site where aminoacyl-tRNA enters ribosome. The anticodon of the specific AA charged tRNA base pairs with the codon on mRNA providing the selectivity for decoding the mRNA.

P-Site (Peptidyl Site): It is the site where the peptidyl tRNA is formed in the ribosome. The amino acid is transferred from the tRNA in A-site.

E-Site (Exit Site): It is the site where tRNA moves to after transfer of the amino acid in the P-site and before leaving the ribosome. The designs and functions of eukaryotic, archaeal and bacterial ribosomes are similar ([Ramakrishnan, 2002](#); [Harms et al., 2001](#)).

The translation can be divided into three stages: initiation, elongation and termination, where tRNAs are re-located from A site to P site and E site accordingly. High resolution structures

of ribosome obtained from X-ray crystallography and Cryo-EM confirmed that tRNA keeps the L-shaped conformation in all three sites, with considerable conformational changes to promote translation at each step ([Fernandez et al., 2014](#); [Amunts et al., 2014](#); [Tourigny et al., 2013](#); [Schmeing and Ramakrishnan, 2009](#); [Gao et al., 2009](#); [Weixlbaumer et al., 2008](#); [Selmer et al., 2006](#); [Valle et al., 2003](#); [Brodersen et al., 2000](#); [Wimberly et al., 2000](#); [Lomakin and Steitz, 2013](#); [Polikanov et al., 2012](#); [Blaha et al., 2009](#); [Ban et al., 2000](#)).

tRNA in Translation Initiation

Translation is a complex enzymatic process that requires ribosome, mRNAs, tRNAs, and multiple protein translation factors. In prokaryotes, three initiation factors (IF1, IF2 and IF3) are responsible for the initiation of translation ([Dever, 2002](#)). In eukaryotes, initiation involves more than 12 protein

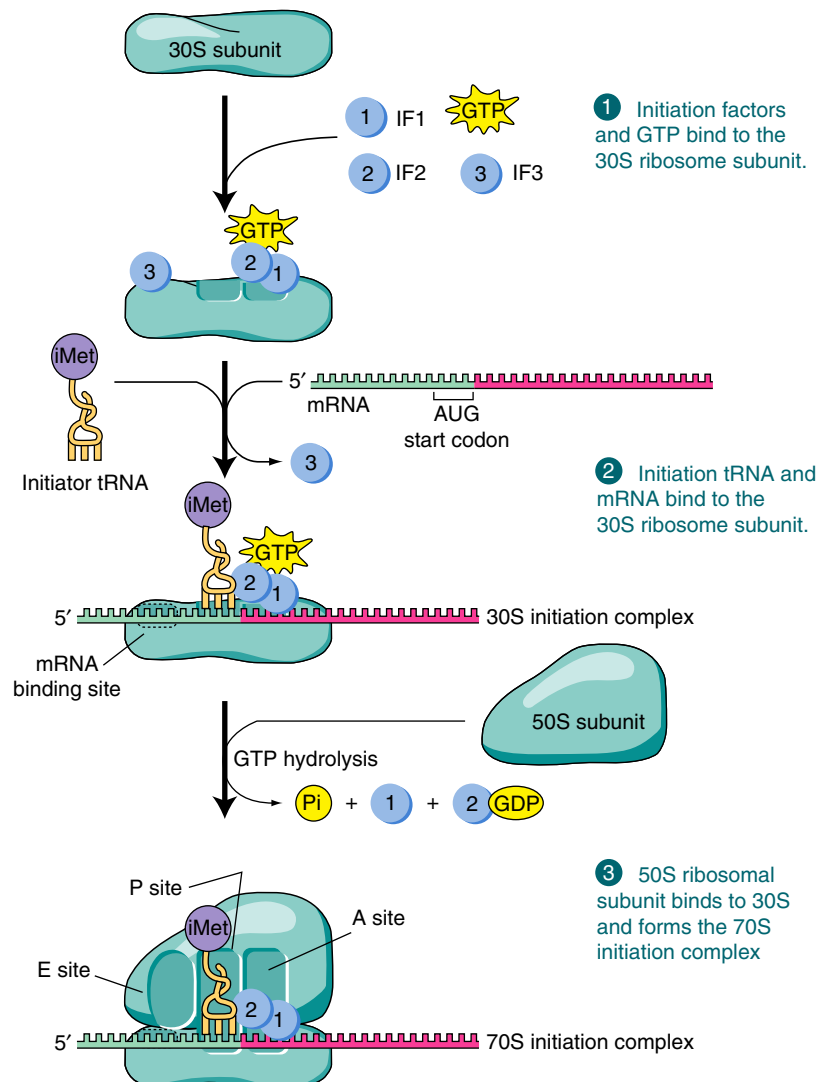


Figure 9 tRNA in translation initiation complex. Three initiation factors (IF1, IF2, and IF3) first bind to the 30S ribosomal subunit in bacteria. This step is followed by binding of the mRNA and the initiator *N*-formyl-methionyl (fMet) tRNA^{fMet}, which is recognized by IF2 bound to GTP. IF3 is then released, and a 50S subunit binds to the complex, triggering the hydrolysis of bound GTP, followed by the release of IF1 and IF2 bound to GDP.

factors including EIF1, eIF1A, eIF2, eIF2B, eIF3, eIF4A, eIF4E, eIF4G, eIF5, and eIF5B, where many of them play important regulatory roles during this step (Kozak, 2005; Allen and Frank, 2007; Hinnebusch and Lorsch, 2012).

Translation Initiation consists of three steps. First, mRNA is bound to the ribosome small subunit. Second, the initiator tRNA will recognize and bind to the mRNA initiation AUG codon, through a codon–anticodon match. Third, the large subunit is recruited to this initiation complex and proceeds to protein synthesis (Ramakrishnan, 2002; Wilson and Hunt, 2008).

Protein synthesis in general starts with an AUG codon near the beginning of the mRNA, following the 5'-UTR, which often contains regulator elements for translation control. To distinguish the initiation codon from internal methionine codons, prokaryotes contain a specific sequence, Shine–Dalgarno sequence (UAAGGAGG), which is located about 5–10

nucleotides before the initiation codon. Close to the 3' end of 16S RNA of 30S ribosomal subunit, there is a nucleotide sequence complementary to Shine–Dalgarno sequence. The interaction of these two sequences accelerates the binding of mRNA to the 30S ribosomal subunit (Figure 9). In eukaryotes, the first AUG codon is recognized by a specific initiation tRNA. A distinction between bacteria and eukaryotes also exists in this initiation tRNA. In prokaryotes, proteins synthesis is initiated with *N*-formyl methionyl tRNA^{fMet}. In eukaryotes, protein synthesis is initiated with non-formylated, but a special methionyl initiator tRNA (tRNA^{iMet}) (Varshney *et al.*, 1993; Myasnikov *et al.*, 2009). During initiation, IF3 binds to the 30S ribosome small subunit and prevents the binding of large subunit. Bacterial initiation complex includes initiation factor (IFs), GTP, ribosome subunits, mRNA and fMet-tRNA^{fMet} (Ramakrishnan, 2002; Varshney *et al.*, 1993; Myasnikov *et al.*, 2009).

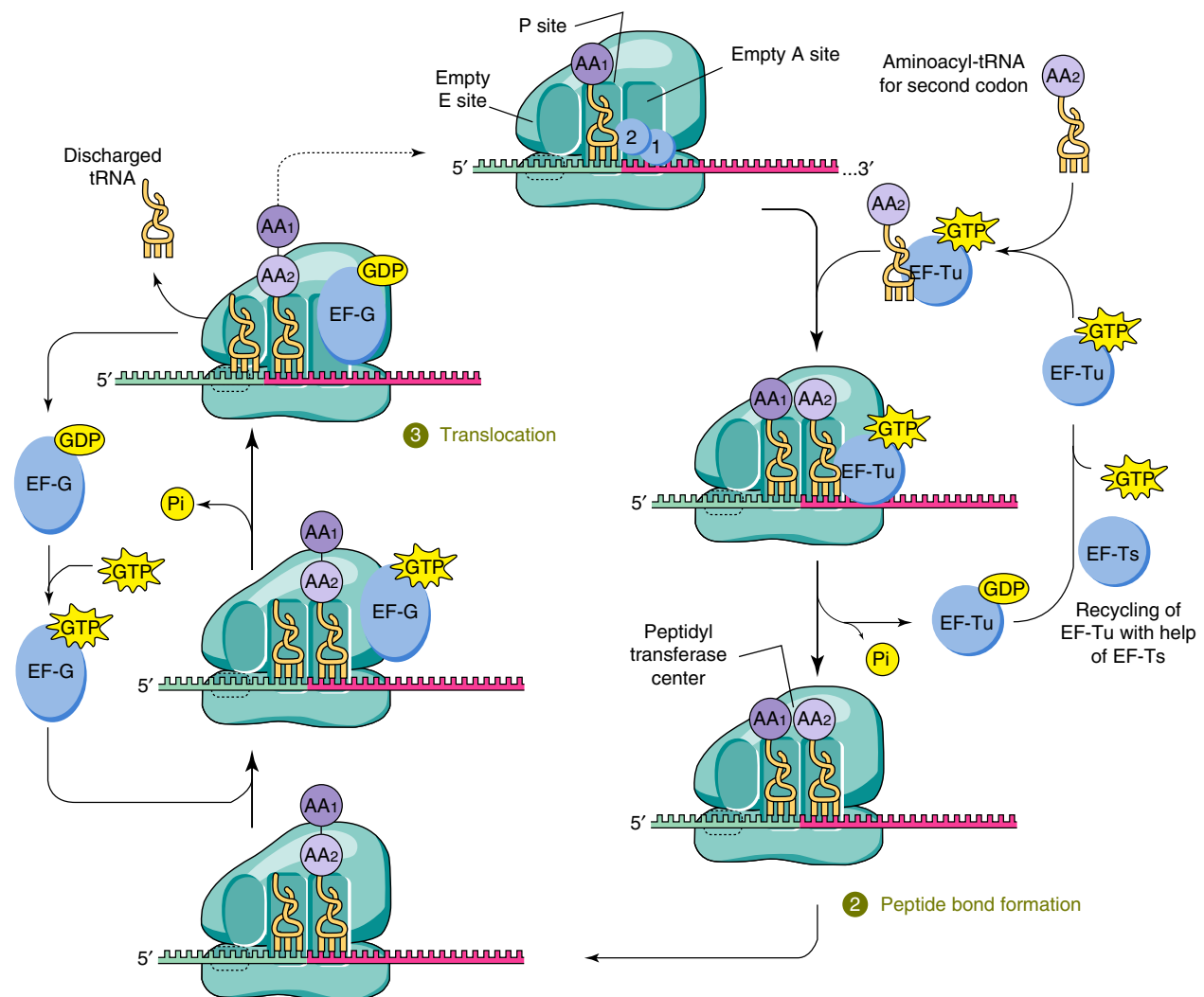


Figure 10 tRNA in translation elongation complex. A cognate aminoacyl-tRNA bound to EF-Tu binds to its cognate codon of the mRNA in the ribosome at the acceptor site (A site), adjacent to the tRNA in the peptidyl site (P site), which is bound to the nascent polypeptide (1). Decoding occurs on the 30S subunit of the ribosome, GTP is hydrolysed by EF-Tu, which leads to release of aminoacyl-tRNA into the A site if there is correct codon–anticodon pairing. Subsequent peptide bond formation between the nascent polypeptide in the P site and the amino acid attached to the tRNA in the A site results in a peptidyl-tRNA in the A site that is one amino acid longer (2). After translocation of the ribosome, the tRNA attached to the nascent polypeptide is moved to the P site and a new aminoacyl-tRNA is delivered to the empty A site (3).

The second step of initiation involves the binding of initiator fMet-tRNA^{fMet} to the initiation codon. In this step, IF2 plays a key role by controlling entry of tRNA into the P site of ribosome. IF2 in GTP-bound form is directly bound to the P site in 30S subunit. Then fMet-tRNA^{fMet} binds to IF2 and IF2 transfers fMet-tRNA to the P site. IF1 remains associated with the 30S ribosomal subunit in A site and maintains the separation of small and large subunits, further preventing entry of other aminoacyl-tRNAs to the A site.

At the final step of initiation, the 50S large subunit is joined to the 30S small subunit. Energy required for the initiation step is provided from the hydrolysis of guanosine triphosphate (GTP) by IF1 and IF2. IF2 is released and the initiation complex of protein synthesis is formed. The resulting initiation complex is a 70S ribosome with an empty A site, carrying initiator fMet-tRNA^{fMet} in P site bound to mRNA (Schmeing *et al.*, 2009). Eukaryotic translation initiation is similar to that of bacteria but more complex. In addition to the different initiator tRNA^{iMet}, eukaryotic initiation complex is also associated with the 3' poly(A) tail of mRNA, the 5'-cap structure (5' end of eukaryotic mRNAs), the Kozak sequence in mRNA that increases the efficiency of initiator tRNA entering the ribosome, and more eukaryotic specific inhibition factors (eIFs) (Polikanov *et al.*, 2012; Ban *et al.*, 2000).

tRNA in Elongation

In the elongation step, amino acids are combined to form a polypeptide chain in the ribosome. It includes all reactions from the first peptide bond formation to the last peptide bond formation in the synthesis of a single protein (Ramakrishnan, 2002; Alberts, 2008).

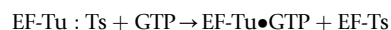
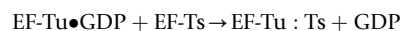
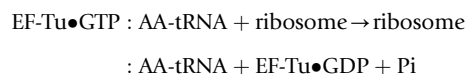
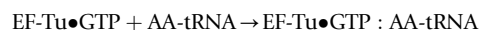
After initiation the ribosome is bound to mRNA with the AUG initiation codon positioned by fMet-tRNA^{fMet} at P site while the A site is empty (Schmeing *et al.*, 2009). Prokaryotic translation elongation involves this initiation complex as well as tRNAs charged with their cognate amino acids (aminoacyl-tRNAs), elongation factors EF-Ts (elongation factor thermo stable), EF-Tu (elongation factor thermo unstable) and EF-G (historically known as translocase) and GTP (Ramakrishnan, 2002; Tourigny *et al.*, 2013; Schmeing and Ramakrishnan, 2009; Selmer *et al.*, 2006).

Elongation occurs in three steps (Ramakrishnan, 2002; Frank and Gonzalez, 2010; Figure 10). The first step is the binding of aminoacyl-tRNAs matching the codon in the mRNA at the ribosome A site. Aminoacyl-tRNAs (AA-tRNAs) generated by the aminoacyl-tRNA synthetases are prone to hydrolysis in water. To deliver aminoacyl-tRNA to ribosomes efficiently, a GTP-bound EF-Tu will bind specifically to the charged tRNAs and protect amino acid (AA) from hydrolysis as well as masks it from peptide bond formation. After formation of the initiation complex, thanks to hydrolysis of GTP and elongation factor (EF-Tu) at the A site of this complex, aminoacylation-tRNA uses the anticodon to recognize the codon sequence on mRNA. The EF-Tu•GTP:AA-tRNA ternary complex enters the ribosome A site by matching the anticodon of tRNA to the codon on mRNA. EF-Tu then interacts with a factor binding site in the ribosome large subunit that triggers its GTPase activity. GTP hydrolysis results in the separation of the complex into EF-Tu•GDP and phosphate (Pi) and

aminoacyl-tRNA is released (Ramakrishnan, 2002; Schmeing and Ramakrishnan, 2009; Schmeing *et al.*, 2009; Frank and Gonzalez, 2010).

The crystal structure of ribosome associated with the ternary complex EF-Tu•GTP:AA-tRNA revealed a strongly distorted structure of tRNA bound at the A site. With the anticodon binding to the codon on mRNA, the interaction leads to a ~30° bend of the anticodon (AC)-stem and a 5 Å swing of the D-stem moving away from the T-stem. This distortion allows the AA-tRNA to interact simultaneously with elongation factor and cognate codon in the decoding center of 30S subunit. The relaxation of the bent conformation may accompany the hydrolysis of GTP to promote tRNA translation from the A to P site (Ramakrishnan, 2002; Tourigny *et al.*, 2013; Schmeing and Ramakrishnan, 2009; Selmer *et al.*, 2006).

EF-Tu•GDP is inactive and must be activated before the next elongation cycle. For this, EF-Ts is required, because affinity of EF-Tu for GDP is 40 times greater than its affinity for GTP. EF-Ts activates EF-Tu exchanging GDP with GTP. EF-Tu does not interact with fMet-tRNA^{fMet} because initiator fMet-tRNA^{fMet} never enters the A site (Ramakrishnan, 2002; Alberts, 2008).



The second step in elongation involves the transfer of the initial methionine or peptide from the tRNA in the P site to the amino acid attached to tRNA in the A site. This occurs through the hydrolysis of the ester bond between the amino acid and the 3' -OH group of tRNA in the P site and the formation of a peptide bond with the amino acid -NH₂ group in the A site tRNA. The reaction breaks the bond between tRNA in the P site and the initial methionine or peptide, and then transfers the methionine or peptide onto the amino acid bound to the tRNA in the A site. This reaction is called peptidyl transferase reaction, with the responsible region in the ribosome is called peptidyl transferase center (Figure 10).

Translocation is the third step of the elongation phase and involves positioning of the ribosome to read the next codon and requires GTP hydrolysis of elongation factor G (EF-G). The hydrolysis of GTP to GDP alters the three dimensional structure of EF-G. This alteration of EF-G moves the peptide-tRNA in the A site to the P site. Because this peptide-tRNA is base paired with the mRNA, the mRNA also shifts, positioning the next codon in the empty A site. At the same time, the original tRNA in the P site, now relieved of methionine or peptide, moves to the E (exit) site breaking the base pairing with mRNA. The uncharged tRNA is released from the E site and EF-G•GDP is released from the ribosome (Ramakrishnan, 2002; Schmeing and Ramakrishnan, 2009; Alberts, 2008).

Elongation and translocation reactions in eukaryotes are very similar to those in prokaryotes with minor differences in

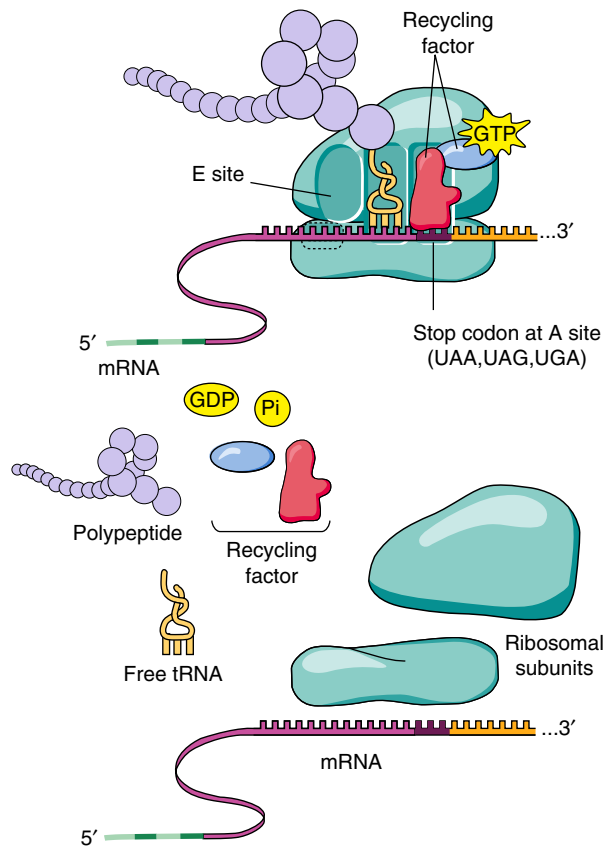


Figure 11 tRNA in translation termination complex. When a stop codon (UAG, UAA, or UGA) arrives at the A site, it is recognized and bound by a ribosomal recycling factor. This causes the polypeptide to be transferred to a molecule of water and the release of the polypeptide from the tRNA, followed by the dissociation of the other components of the elongation complex.

proteins. Instead of elongation factors EF-Tu and EF-Ts there is a stable ternary complex made up of eEF1- α - β and γ . eEF1- α is the eukaryotic equivalent of EF-Tu; eEF1- β - γ are eukaryotic equivalents of EF-Ts. eEF2 (eukaryotic elongation factor 2) is the equivalent of prokaryotic EF-G and regulates ribosome translocation. Phosphorylation of eEF2 by EF2 kinase regulates translation elongation by completely inactivating the eEF2-dependent ribosomal translocation (Hershey *et al.*, 2012; Dever and Green, 2012).

tRNA in Termination

Termination phase is the release of the completed polypeptide from the ribosome and occurs when the ribosome arrives at one of the three termination codons (UAA, UAG, and UGA) to the A site. Because there are no complementary tRNAs for the termination codons, no tRNA binds to the A site. Instead, proteins called release factors (RFs) or ribosomal recycling factors (RRFs) recognize the termination codons and terminate translation (Schmeing and Ramakrishnan, 2009; Weixlbaumer *et al.*, 2008; Alberts, 2008).

Ribosomal releasing factors are divided into two classes, Class 1 RRFs and Class 2 RRFs. In prokaryotes, there are two

types of Class 1 RRFs: RF1 and RF2. In eukaryotes, Class 1 RRF has only one; eRF1. In prokaryotes and eukaryotes there is only one Class II factor: RF3 and eRF3, respectively (Nakamura and Ito, 2003; Carraro *et al.*, 2010).

In prokaryotes RF1 is responsible for the recognition of termination codons UAA and UAG and RF2 triggers polypeptide hydrolysis from tRNA in the P site. RF1 and RF2 are similar in size and shape to tRNA and interact with the A site of the ribosome as tRNA does (Nakamura and Ito, 2003). RF3 is structurally similar to EF-G complex and hydrolysis of RF3•GTP to RF3. GDP promotes the separation of the polypeptide chain from the linked tRNA 3'-OH in P site and the release of the polypeptide chain from the ribosome (Figure 11).

The termination of protein synthesis in eukaryotic cells is similar to that in prokaryotes. However, two different releasing factors have been identified. The first one, eRF1, recognizes all three termination codon and is equivalent to RF1 and RF2 in prokaryotes. eRF3 binds with GTP and stimulates the separation of the polypeptide chain from tRNA, similar to RF3 in prokaryotes. Termination by the release factors enable the protein synthesis cycle to end and the mRNA to be translated again or degraded (Dever and Green, 2012; Nirenberg and Tampe, 2013).

Ribosome Recycling

After the release of the polypeptide chain and RRFs, the ribosome (in its P and E site) is still connected to mRNA and uncharged tRNA. For the ribosome to enter a new cycle of polypeptide synthesis, it needs to release the mRNA and tRNA and separate the large and small subunits of the ribosome. EF-G and IF3 help to activate this process and are necessary for ribosome recycling after the release of polypeptide (Franckenberg *et al.*, 2012).

Surprising Number of Roles of tRNAs beyond Translation

Molecular biologists first encountered tRNAs as a key component of the translation machinery and because of this history it is all too easy to think of translation as the primary or proper function of tRNA. But in fact, tRNA and tRNA processing enzymes impact much more than protein production. Studies have uncovered roles for tRNA in the regulation of transcription, translation and protein turnover. Induced by stress or as part of a developmental program, non-random tRNA fragments can guide mRNA cleavage, inhibit translation and promote morphological changes. Similarly, tRNA processing enzymes, such as RNaseP and aminoacyl-tRNA synthetases participate in tasks affecting more than tRNA function (i.e., mRNA function and cellular signaling). Unraveling the complexities of all these functions will increase our understanding of the downstream consequences of mutations in these genes and how they impact health and disease. As we will discuss in detail, tRNA and tRNA-like molecules also play key roles in a wide variety of cellular processes including cell wall and membrane biosynthesis (Shepherd and Ibba, 2013;

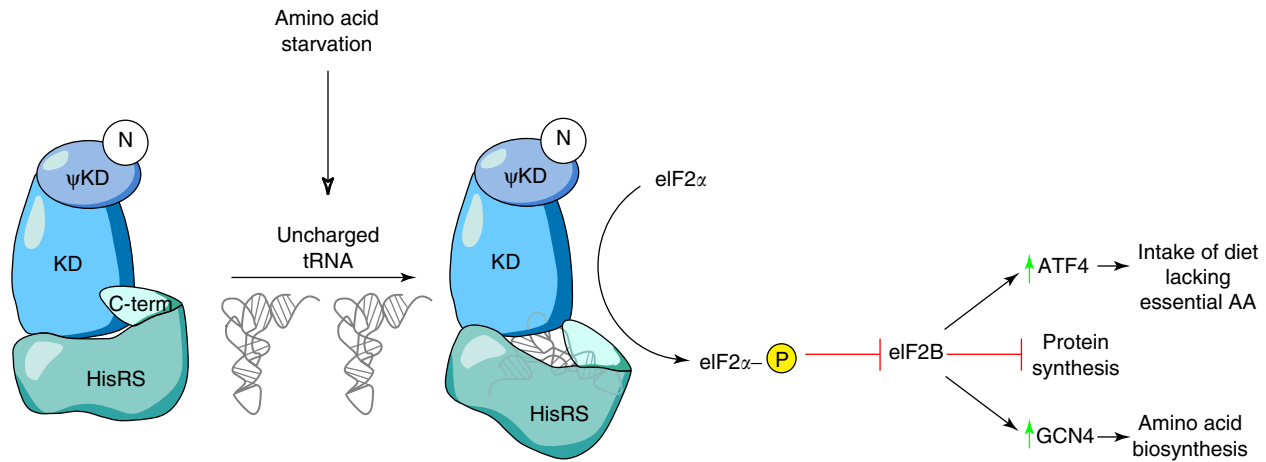


Figure 12 GCN2 and tRNA. Amino acid starvation causes the accumulation of uncharged tRNAs that bind to the histidyl-tRNA synthetase (HisRS)-related domain of GCN2 and activate the adjacent protein kinase domain (KD). GCN2 also contains a pseudo-protein kinase domain (ΨKD) of unknown function and an *N*-terminal binding site (N) for the GCN1/GCN20 complex that is essential for recognition of the starvation signal. GCN2 phosphorylates the translation factor eIF2 α , preventing it to be recycled by its guanine nucleotide exchange factor eIF2B and leading to inhibition of protein synthesis. At the same time, phosphorylation of eIF2 α also leads to the translational induction of GCN4 synthesis in yeast and, ATF4 in animals. GCN4 activates the transcription of genes necessary for amino acid biosynthesis. In mice, activation of GCN2 in the brain mediates an aversive response toward intake of diets lacking essential amino acids.

Francklyn and Minajigi, 2010; Dare and Ibba, 2012), antibiotics synthesis (Francklyn and Minajigi, 2010; Banerjee *et al.*, 2010), metabolic regulation (Castilho *et al.*, 2014; Murguia and Serrano, 2012), protein degradation (Mogk *et al.*, 2007), and replication of RNA or DNA from viruses (Saadatmand and Kleiman, 2012; Klasse, 2012; Dreher, 2009; Gomez *et al.*, 2004; Hurto, 2011), etc.

tRNA and GCN2 Pathway

Eukaryotes from yeast to animals select foods to achieve a balanced diet that includes adequate levels of the essential amino acids. When starved for any of several amino acids, yeast cells induce the transcription of over 70 genes encoding enzymes that function in the biosynthesis of all 20 amino acids (Hinnebusch and Natarajan, 2002). Mammals are not capable of synthesizing all 20 amino acids and must obtain essential amino acids from the diet. The presence of essential amino acids is somehow detected and a diet missing any of these amino acids will be rejected by rats. The molecular mechanisms governing this universal 'feeding' behavior is through the accumulation of uncharged tRNAs.

Genetic studies in yeast identified the 'general CONTROL Nonderepressing 2 protein kinase (GCN2)' as the primary sensor of amino acid starvation and the domain structure of GCN2 helps to explain this assignment. In addition to a typical eukaryotic protein kinase domain, GCN2 contains a domain related to histidyl-tRNA synthetase (HisRS), which binds uncharged tRNAs with higher affinity than the corresponding charged tRNAs (Dong *et al.*, 2000). Amino acid starvation results in the accumulation of uncharged tRNAs that bind to the synthetase-like domain of GCN2 and activates the kinase domain (Wek *et al.*, 1995; Figure 12). Active GCN2 phosphorylates the alpha subunit of eukaryotic initiation factor 2 (p-eIF2 α) on serine 51, which inactivates eIF2 α /b/g

and inhibits translation initiation. eIF2 α /b/g bound to GTP delivers the initiator methionyl-tRNA (Met-tRNA^{Met}) to the small ribosomal subunit in the first step of translation. eIF2 α /b/g-GDP is released from the ribosome and is recycled to eIF2 α /b/g-GTP by its guanine exchange factor (GEF) eIF2B, however p-eIF2 α blocks this recycling reaction thereby preventing translation initiation.

Paradoxically, phosphorylation of eIF2 α leads to the translational induction of GCN4 synthesis in yeast, and GCN4 activates the transcription of genes necessary for amino acid biosynthesis. In mice, activation of GCN2 in the anterior piriform cortex of the brain mediates an aversive response toward intake of diets lacking an essential amino acid. Subsequent cellular adaptation to amino acid deprivation is marked by decreases in global protein synthesis complemented by increased transcription of genes related to amino acid biosynthesis (Castilho *et al.*, 2014; Murguia and Serrano, 2012). The components of the pathway linking eIF2 α phosphorylation and downstream behaviors in animals involve the transcription factor ATF4 (Carraro *et al.*, 2010; Harding *et al.*, 2000; Sikalidis *et al.*, 2014), although the complete mechanism remains unknown (Peng *et al.*, 2012).

GCN2 binds and is activated by many different uncharged tRNAs (Dong *et al.*, 2000). A genome-wide analysis of tRNA activation of GCN2 in yeast has shown that GCN2 was activated by multiple uncharged tRNAs during amino acid starvation and, interestingly, these uncharged tRNAs are not necessarily the tRNAs for the missing amino acids. Other environmental stresses such as high salt could also lead to selective deacylation of tRNAs and activation of GCN2. Nonetheless, inhibition of aminoacyl-tRNA synthetases is absolutely responsible for the accumulation of uncharged tRNA. Mutation of an aminoacyl-tRNA synthetase triggers GCN2 activation even in medium replete with amino acids (Hinnebusch and Natarajan, 2002; Dong *et al.*, 2000).

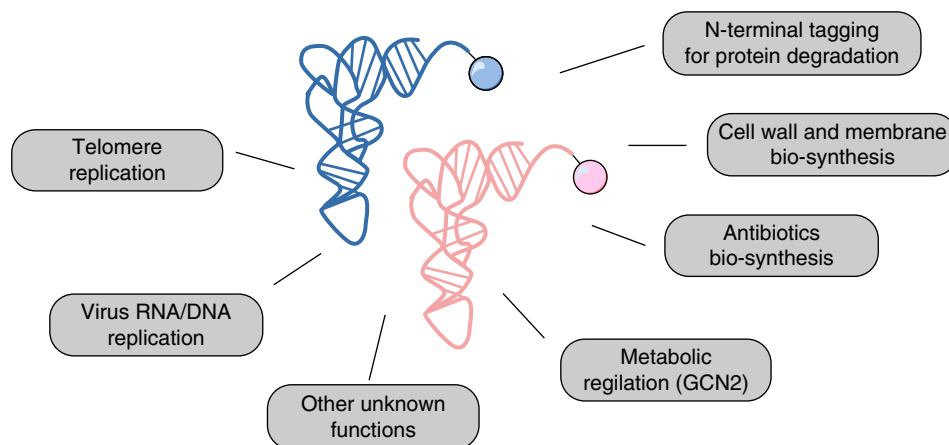


Figure 13 Nontranslational function of tRNAs. In addition to the canonical function of tRNA as substrate for translation, tRNAs have been exploited for various roles in cellular processes, including metabolic regulation, proteolysis, bio-synthesis, replication of telomere and even viral RNA and DNA.

tRNAs for Regulated Proteolysis

The *N*-end rule states that the half-life of a protein is determined by the nature of its *N*-terminal residue. Although prokaryotes and eukaryotes employ distinct proteolytic machineries for degradation of *N*-end rule substrates, recent findings indicate that they share common principles of substrate modification (Mogk *et al.*, 2007).

There are three ways of generating specific *N*-terminal residue that target proteins for degradation; *N*-ter or endopeptidase cleavage, and transfer of amino acids onto *N*-termini (Tasaki *et al.*, 2012). In addition to the methionine aminopeptidase (in *E. coli*) and endopeptidase (in eukaryotes), both *E. coli* and eukaryotes can transfer amino acids from aminoacyl-tRNAs (AA-tRNAs) to certain *N*-termini using enzymatic transferases (Figure 13). This fundamental principle of regulated proteolysis is conserved from bacteria to mammals. For example in eukaryotes, Arg is conjugated to specific proteins by the ATE1-encoded arginyl-tRNA protein transferase (R-transferase), which targets the protein for Ub-dependent degradation (Balzi *et al.*, 1990; Rai and Kashina, 2005). The human pathogen *Vibrio vulnificus* encodes a second L-transferase (Bpt) that is homologous to eukaryotic ATE1. In *E. coli*, Arg and Lys function as secondary destabilizing residues that recruit primary destabilizing residues (Phe, Leu), which are transferred using a leucyl/phenylalanyl-tRNA protein transferase (L/F-transferase or Aat) (Ninnis *et al.*, 2009). The crystal structure of *E. coli* L/F-transferase suggests that the enzyme has preference for different aminoacyl-tRNAs with specific amino acids binding pockets (Watanabe *et al.*, 2007).

tRNAs for Cell Wall Synthesis

In the bacterial cell wall, amino acid modification of peptidoglycan facilitates cross-linkage with the plasma membrane resulting in decreased membrane permeability that is often a prerequisite for high-level antibiotic resistance. Many clinically relevant bacteria, including *Streptococcus pneumoniae*, *Staphylococcus aureus*, *Enterococcus faecalis*, and *Pseudomonas aeruginosa* recruit specific aminoacylated-tRNAs into peptidoglycan

biosynthesis and membrane phospholipid modification (Francklyn and Minajigi, 2010; Dare and Ibba, 2012; Banerjee *et al.*, 2010). The cross-linking of short intra-strand peptides provides the cell with additional rigidity allowing for resistance of *S. aureus* to β -lactam antibiotics, including methicillin (Biarrotte-Sorin *et al.*, 2004). These amino acids, usually Gly and Ala, are transferred from the aminoacyl-tRNA (AA-tRNA) to a hexapeptide lipid intermediate by a series of enzymes, such as FemA/B/X. Each of the Fem proteins attaches several amino acids to build the inter-peptide bridge in a sequential manner (Biarrotte-Sorin *et al.*, 2004; Benson *et al.*, 2002).

In order to participate in peptidoglycan biosynthesis, the AA-tRNAs must escape the cytoplasmic protein translation circuit. Some tRNAs are specially optimized for escaping translation by lacking certain GT ψ C and GG sequence in the acceptor stem loop and are thus unable bind with EF-Tu•GTP and ribosome. Further, the FemX proteins specifically recognize the amino acid moiety (Gly and Ala), the discriminator base of tRNA^{Gly}, and the acceptor stem of tRNA^{Ala}, to recruit these tRNAs for peptidoglycan biosynthesis. Other amino acids, including Ser, Thr, Lys and Arg and their tRNAs are also used for peptidoglycan cross-linking or cell membrane lipid modification (Dare and Ibba, 2012).

In Streptomycetes, aminoacylated tRNAs are also used for antibiotic synthesis as well as antibiotic resistance. The antibiotic valanimycin produced by *Streptomyces viridifaciens* is derived from Val (L-valine) and Ser (L-serine). Val is first transformed into isobutylhydroxylamine that must be reacted with Ser during valanimycin biosynthesis (Garg *et al.*, 2008). A seryl-tRNA synthetase gene unexpectedly identified in the valanimycin biosynthetic gene cluster (vlm) indicates that the seryl residue could be transferred from seryl-tRNA to the hydroxyl group of isobutylhydroxylamine (Garg *et al.*, 2006). Another example of tRNA-dependent antibiotic synthesis is in the biosynthesis of albonoursin (alb) in *Streptomyces noursei* (Gondry *et al.*, 2009). Albonoursin is a cyclodeptide antibiotic made from Phe and Leu by the AlbC enzyme and it was found that charged *E. coli* tRNA^{Phe} and tRNA^{Leu} were the required substrates. The following studies detected numerous other cyclodeptides containing Ala, Val, and Met from the cell extract,

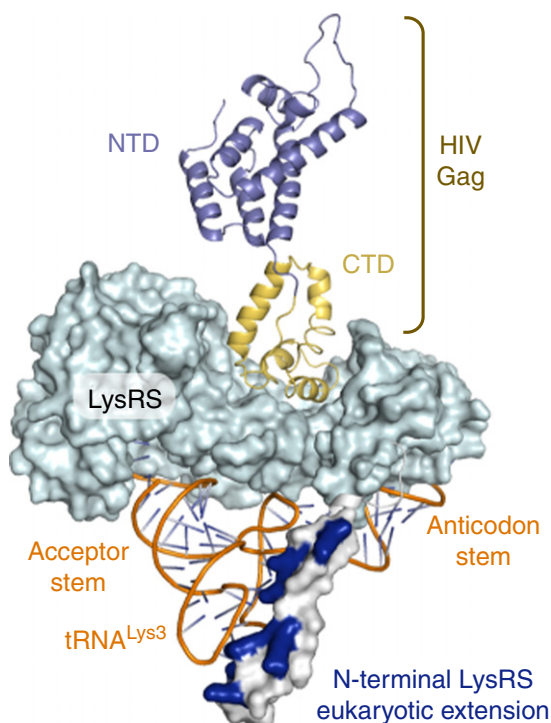


Figure 14 Mechanism of the LysRS/tRNA^{Lys}/Gag/Pol packaging complex in HIV-1. Reverse transcription of the HIV genome is primed by a human lysine-specific tRNA (tRNA^{Lys3}) that is packaged (into the virion) by the HIV Gag protein with simultaneous packaging of tRNA^{Lys3}, LysRS, and Gag. A model here showing Gag, tRNA^{Lys}, and LysRS form a ternary complex. The model suggests that tRNA^{Lys} is mainly anchored through the interaction of the anticodon stem-loop with the anticodon-binding domain, with minimal interaction being seen with the catalytic domain of the synthetase. Other interactions between tRNA^{Lys3} with Gag-Pol may further stabilize the ternary complex for tRNA^{Lys3} incorporation into the HIV virion.

indicating that AlbC and its homologs could use other tRNAs as substrates besides tRNA^{Phe} and tRNA^{Leu}. These examples suggest that tRNAs have many other roles that remain to be uncovered.

tRNAs for Virus Reverse Transcription Initiation

During the replication of RNA viruses, the viral RNA genome is converted into double-stranded proviral DNA by reverse transcription. Initiation of reverse transcription is primed by a cellular tRNA, which is selectively incorporated into the virus during assembly in the cell of origin. The primer tRNA is annealed to an 18-base sequence near the 5' end of the viral RNA genome termed the primer binding site (PBS), and is used to prime the reverse transcriptase-catalyzed synthesis of minus strand cDNA in the newly infected cell.

Different viruses use different cellular tRNAs as primer (Mak and Kleiman, 1997). In lentiviruses, such as human immunodeficiency virus 1 (HIV-1) and mouse mammary tumor virus, tRNA^{Lys3} serves as the primer tRNA (Marquet *et al.*, 1995). However, in avian retroviruses, tRNA^{Trp} is the primer for all members of the avian sarcoma and leukosis virus group examined to date (Swanstrom and Wills, 1997; Morris and Leis, 1999), whereas tRNA^{Pro} is the common primer for murine

leukemia virus (MuLV) (Peters and Hu, 1980). tRNA^{Lys1,2}, two other tRNA^{Lys} isoacceptors differing by one base pair in the anticodon stem (differ from tRNA^{Lys3} by 14 or 16 bases), is also the primer tRNA for several retroviruses, including Mason-Pfizer monkey virus (MPMV) and human foamy virus (HFV).

Primer tRNA is present in the mature virus capsid core prior to initiation of reverse transcription, which may happen before capsid disassembly (Mak and Kleiman, 1997; Kleiman, 2002). Thus, during virus assembly, viral genomic RNA and cellular tRNA isoacceptors (such as tRNA^{Lys3} and tRNA^{Lys1,2}) are selectively concentrated at the site of assembly, and packaged by virus proteins. The packaging complex in HIV-1 includes the capsid Gag protein, reverse transcriptase Pol, and the host human lysyl-tRNA synthetase (LysRS; Figure 14). Both LysRS and Pol interact with tRNA^{Lys} and additional contacts between the viral RNA and LysRS may also occur (Saadatmand and Kleiman, 2012; Guo *et al.*, 2005; Kovaleski *et al.*, 2006; Kovaleski *et al.*, 2007; Kleiman *et al.*, 2010). In HIV-1, estimates of approximately 20 molecules of tRNA^{Lys}/virion have been reported, with about 8 molecules tRNA^{Lys3} and 12 molecules tRNA^{Lys1,2} per virion, reflecting their cytoplasmic ratio (Guo *et al.*, 2003; Javanbakht *et al.*, 2003; Halwani *et al.*, 2004). A portion of the incorporated tRNA^{Lys} anneals to the 18-nucleotide PBS sequence near the 5'-end of the viral genome, which is perfectly complementary to the 3'-18 nucleotides of tRNA^{Lys3}. It is not known whether tRNA annealing occurs prior to, or after, viral budding. The co-packaged tRNA^{Lys1,2} not involved in priming the virus RNA may play a role in the import of the pre-integration complex into the nucleus of the HIV-1 infected cell (Kleiman and Cen, 2004). Efficient packaging tRNA^{Lys} into viral particles is required for both annealing and optimizing viral infection of the host cells (Saadatmand and Kleiman, 2012; Guo *et al.*, 2003, 2005; Kleiman *et al.*, 2010; Gabor *et al.*, 2002).

tRNA-Like Molecules

Transfer RNA-like structures (tRNA-like structures, TLSs) are RNA or even protein sequences that have a similar tertiary structure to tRNA. tRNA-like structures have been demonstrated in bacteria (tmRNA), many plant RNA virus genomes that are linked to regulation of virus replication, as well as ribosomal recycling factors (RRF).

tmRNAs

Transfer messenger RNA (tmRNA) is a small bacterial RNA that has dual structural and functional similarities to both tRNA and messenger RNA (Zwieb *et al.*, 1999). tmRNA contains structures similar to the upper half of tRNA, including a complete acceptor stem with 3' CCA end, a T stem-loop and a degenerative D-loop (Figure 15). The 5' end of tmRNA is also processed by RNase P with typical tRNA-specific base modifications (Ray and Apirion, 1979; Keiler, 2008). tmRNA acceptor stem contains the same conserved G₃•U₇₀ wobble pair as tRNA^{Ala} and is aminoacylated with alanine by AlaRS (Komine *et al.*, 1994). In place of the anticodon stem loop is a coding sequence for a short peptide called a tag peptide. Translation of the tmRNA into a single tag-peptide requires the

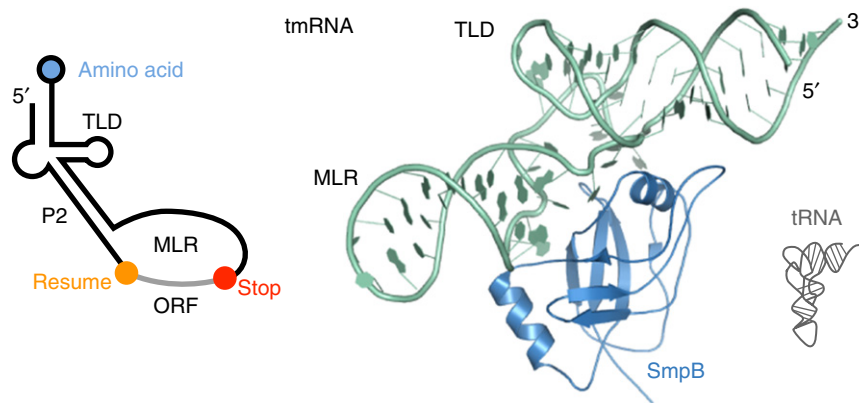


Figure 15 Structure of tmRNA. The schematic diagram shows the secondary structure of tmRNA. The mRNA-like region (MLR) in standard tmRNA is a large loop containing pseudoknots and a coding sequence (ORF) for the tag peptide, marked by the resume codon and the stop codon. Overall structure of the entire tRNA domain of tmRNA complexed with SmpB is shown on the right as determined by X-ray crystallography. The SmpB mimics the missing anticodon stem-loop of a canonical tRNA.

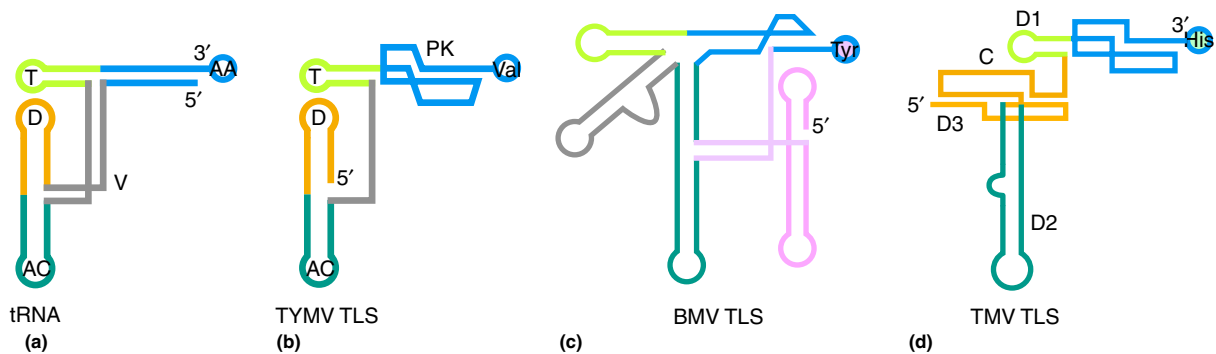


Figure 16 Viral tRNA-like structures. Schematic diagrams show the secondary structure of tRNA with the T loop (T), D loop (D), anticodon loop (AC), variable arm (V), and acceptor stem colored (a), the secondary structure of the TYMV tRNA-like structure (TLS) (b), BMV TLS (c), TMV TLS (d). Structures analogous to those in tRNA are colored, together with the pseudoknot (PK) in the acceptor stem.

elaborate interplay of tmRNA and a tmRNA-binding protein, SmpB, and is termed trans-translation (Keiler, 2008). In the majority of bacteria these functions are carried out by a single chain tmRNA; however, in other bacterial species and mitochondria, a permuted *ssrA* gene produces a two-piece tmRNA in which two separate RNA chains are joined by base-pairing (Gur and Sauer, 2008; Hafez et al., 2013).

The bacterial quality control system monitors protein synthesis and recycles stalled translation complexes in a rescue process that involves tmRNA and trans-translation (Janssen and Hayes, 2012). For example, ribosomes that have stalled upon reaching the end of a messenger RNA that has lost its stop codon are rescued by this pathway. During rescue, tmRNA recycles the stalled ribosome, adds a proteolysis-inducing tag to the unfinished polypeptide, and facilitates the degradation of the aberrant messenger RNA (Keiler, 2008). First, alanine charged tmRNA (Ala-tmRNA) binds to EF-Tu · GTP and SmpB and this complex enters the vacant A-site of the stalled ribosome, similar to aminoacyl-tRNA but without the codon–anticodon interaction. The bound mRNA is subsequently replaced by the tag-encoding region of tmRNA allowing translation of the tag onto the nascent peptide. These tagged proteins are preferentially degraded by AAA + proteases so that they do not accumulate in the

cell. tmRNA and SmpB structurally and functionally mimics both tRNA and mRNA during these processes. Although several structural elements are known to be essential, the molecular mechanism for correct trans-translation is still not fully understood. This pathway is unique in that it employs a small RNA to prevent the accumulation of non-functional proteins produced from truncated mRNA. Although alternative rescue systems have recently been revealed, trans-translation is the only system that universally exists in bacteria and therefore is thought to play a major role in rescuing stalled translation in bacteria (Janssen and Hayes, 2012; Shpanchenko et al., 2010).

Viral tRNA-Like Molecules

Viruses commonly exploit or modify some aspect of tRNA biology. Large DNA viruses, especially bacteriophages, phycodnaviruses, and mimiviruses, produce their own tRNAs, which are used to adjust translational capacity during infection (Dreher, 2009). Retroviruses recruit specific host tRNAs for use in priming the reverse transcription of their genome. Certain positive-strand RNA plant viral genomes possess 3'-tRNA-like structures (TLSs) that are built quite differently from authentic tRNAs, yet efficiently recapitulate several properties of tRNAs

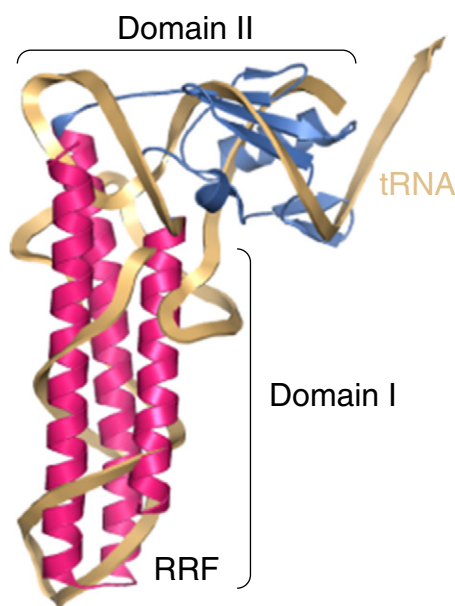


Figure 17 tRNA-like ribosomal recycling factors. Structure representation of the superposed *Thermotoga maritima* RRF (red and blue, pdb1dd5) and yeast tRNA^{Phe} (yellow, pdb2tra).

(Dreher, 2010). For example, in tobacco mosaic virus (TMV), replication takes place in two stages: synthesis of a minus RNA strand using the virus plus-strand RNA as a template and synthesis of progeny plus strand RNA using the minus strand as a template (Osman *et al.*, 2000). The 3' UTR of tobacco mosaic virus plus-strand RNA can be folded into a tRNA-like structure (TLS) containing a 3' pseudoknotted domain (D1) that mimics a tRNA acceptor branch terminating in an unpaired CCA sequence and a domain (D2) that resembles a tRNA anticodon branch (Felden *et al.*, 1996; van Belkum *et al.*, 1985). The central core, C, connects domains D1, D2, and the upstream domain D3 and determines their relative orientations. All of these TLS regions are important for minus-strand synthesis of TMV RNA (Figure 16).

The TMV 3' TLS, like those of Brome mosaic virus (BMV) and Turnip yellow mosaic virus (TYMV), can be aminoacylated. TMV constructs containing the 3'-most 182 nucleotides (D1, D2, and D3), 108 nucleotides (D1 and D2), or 38 nucleotides (D1) are all substrates for yeast histidyl-tRNA synthetase (Felden *et al.*, 1996). Experiments with BMV suggested that aminoacylation may be required for *in vivo* replication of two of the three viral genome RNAs (van Belkum *et al.*, 1985; Dreher *et al.*, 1984, 1989; Rao and Hall, 1991), but not an absolute requirement for *in vivo* replication of all tymoviruses, since TYMV chimeras containing the 3' TLS of Erysimum latent virus, which cannot be aminoacylated, are infectious in plants and protoplasts (Gultyaev *et al.*, 1994). Furthermore, aminoacylation is not required for TYMV minus-strand synthesis *in vitro* (Deiman *et al.*, 2000; Singh and Dreher, 1997; Singh and Dreher, 1998).

tRNA-Like Ribosome Recycling Factors

In addition to tRNA-like RNA structures, several proteins also mimic the structure of tRNA and play major roles in modern

biology. Most highly characterized proteins demonstrating 'tRNA mimicry' are the ribosome release/recycling factors (RRF) that function to terminate protein synthesis. Ribosome releasing/recycling factor (RRF), together with elongation factor G (EF-G), catalyzes the recycling of ribosomes after one round of protein synthesis (Franckenberg *et al.*, 2012). Bacteria have two RRFs, RRF1, and RRF2, with high specificity to decipher the three stop codons (Petry *et al.*, 2008). In the past decade, the crystal structures of the translation termination complex between the ribosome and RRFs have been determined at atomic resolution (Weixlbaumer *et al.*, 2008; Hershey *et al.*, 2012; Dever and Green, 2012; Agrawal *et al.*, 2004; Wilson *et al.*, 2005). Atomic structures of RRF from five different species, including *E. coli*, show that it is comprised of two structural domains: domain I, consisting of three long α -helix bundles, and the smaller domain II, which is an α/β motif (Selmer *et al.*, 1999; Toyoda *et al.*, 2000; Kim *et al.*, 2000; Yoshida *et al.*, 2001; Saikrishnan *et al.*, 2005). The molecule superimposes almost perfectly onto the structure of tRNA except that the amino acid-binding 3' end is missing (Figure 17). This mimicry is further supported by the identification of two classes of crucial RRF peptide motifs, P(A/V)T/SPF and GGQ, in bacteria, which were functionally equivalent to the anticodon and aminoacyl-CCA terminus of tRNA, respectively. RRF is a near-perfect structural mimic of tRNA in both size and dimensions.

Despite the tRNA-mimicry, RRFs bind to ribosomes quite differently from the way tRNA does. A three-dimensional cryo-electron microscopy map of the *E. coli* 70S ribosome-RRF complex (Borovinskaya *et al.*, 2007), together with the crystal structures of RF2 bound to ribosome (Weixlbaumer *et al.*, 2008; Jin *et al.*, 2010), indicates RRF binds further inside the intersubunit space of the ribosome, with domain II of RRF being oriented more toward the 30S ribosomal subunit and the tip of domain I of RRF is shifted toward the 50S subunit. These important conformational differences of RRF bound ribosome suggest that the mechanism of RRF in directing the dissociation of ribosome is beyond simple structural mimicry of the binding positions of tRNA. In addition to the RRFs, tRNA mimicry has also been demonstrated in proteins involved in translation initiation, elongation, as well as mRNA surveillance pathways for protein synthesis (Nakamura and Ito, 2011).

tRNA and Human Diseases

Because of the essential and multiplexed roles for tRNA in and outside of translation, alterations of tRNAs have a major impact on human health and diseases. Comprehensive inspection of alterations in tRNA and tRNA biosynthetic genes indicate that: (1) Disease-related mutations in mitochondrial tRNA are prevalent; however, (2) mutations in genes encoding cytoplasmic tRNAs are less often associated with disease, and (3) disease-related defects in modification and processing of cytoplasmic tRNA are common. This bipartite pattern may result from several differences including the different numbers of tRNA genes (22 in mitochondria, while over 600 for cytoplasmic), different number of genome copies per cell (tens to hundreds for mitochondrial, while only two nuclear genomic copies for cytoplasmic tRNAs per cell), as well as their different rates of mutation (higher in mitochondria) and functional

Table 2 Mutations of tRNA and tRNA modification enzymes in human diseases

<i>Disease</i>	<i>mt-tRNA involved</i>
ADPD/hearing loss and migraine	tRNA Gln
Adult Leigh syndrome	tRNA Val
AMDF	tRNA Val
Ataxia, PEO, deafness	tRNA Phe
Ataxia + RP + deafness	tRNA Pro
Axial myopathy with encephalopathy	tRNA Phe
Cardiomyopathy/SNHL/possible hypertension factor	tRNA Lys
CIPO	tRNA Ser (AGY)
CIPO/encephalopathy	tRNA Gly
Combined OXPHOS defects	tRNA Trp
Combined OXPHOS defects and severe multisystem disorder	tRNA Arg
CPEO	tRNA Ile, tRNA Leu (CUN)
CPEO/DEAF enhancer	tRNA Ala
CPEO/KSS	tRNA Leu (CUN)
CPEO/MM	tRNA Asn
CPEO/Motor neuron disease	tRNA Ile
CPEO/MS	tRNA Ile
CPEO/possible hypertension factor	tRNA Leu (UUR)
CPEO/Stroke/CM/Breast and renal and prostate cancer risk/altered brain pH	tRNA Leu (CUN)
CPEO + Myopathy	tRNA Glu
CPEO + ptosis	tRNA Ser (UCN) precursor
CPEO + ptosis + myopathy + exercise intolerance + diabetes	tRNA Leu (CUN)
CPEO + ptosis + proximal myopathy	tRNA Asn, tRNA Phe, tRNA Ser (AGY), tRNA Ser (UCN), tRNA Trp, tRNA Ala, tRNA Trp
DEAF	tRNA Phe
DEAF enhancer	tRNA Ala
DEAF helper mut	tRNA Arg, tRNA Cys, tRNA Ser (AGY), tRNA Thr, tRNA Cys
DEMCHO	tRNA Trp
Developmental delay, optic atrophy, cataract, hearing loss, myopathy	tRNA His
Dilated cardiomyopathy (15 bp dup)	tRNA Pro, tRNA Leu
DM	tRNA Leu (UUR)
DMD/MERRF/HCM/epilepsy	tRNA Lys
DMD/RP + SNHL	tRNA Ser (AGY)
Dopaminergic nerve cell death (PD)	tRNA Pro, tRNA Thr
Dystonia and stroke-like episodes	tRNA Lys
Encephalomyopathy	tRNA Asn, tRNA Leu (UUR), tRNA Thr
Encephalomyopathy/DEAF	tRNA Trp
Encephalomyopathy + retinopathy	tRNA Glu
Encephalopathy	tRNA Lys
Encephalopathy/MELAS	tRNA Gln
Epilepsy + ataxia + visual disturbance + deafness	tRNA Lys
EXIT	tRNA Glu, tRNA Leu (UUR), tRNA Tyr, tRNA Met
EXIT and Deafness	tRNA Phe
Extrapryamidal disorder with akinesia-rigidity, psychosis, and SNHL	tRNA Phe
Fatal neonatal lactic acidosis	tRNA Ser (UCN)
FICP	tRNA Ile
FSGS/mitochondrial cytopathy	tRNA Tyr
FSHD	tRNA Leu (CUN)
Gastrointestinal syndrome	tRNA Trp
Gestational diabetes (GDM)	tRNA Leu (UUR)
HCM severe multisystem disorder	tRNA Trp
HCM with hearing loss/possible hypertension factor	tRNA Ile
HCM + MELAS	tRNA Val
Hypomagnesemic metabolic syndrome	tRNA Ile
Hypotonia, seizure, muscle weakness, lactic acidosis, hearing loss	tRNA Met
Isolated complex I deficiency	tRNA Glu
Kearns-Sayre syndrome	tRNA Lys
KSS	tRNA Leu(UUR)
Late infantile onset fatal mito disease	tRNA Val
Leigh syndrome	tRNA Ile, tRNA Trp, tRNA Val
LHON	tRNA Gln, tRNA Leu (UUR), tRNA Glu, and tRNA Met

(Continued)

Table 2 Continued

<i>Disease</i>	<i>mt-tRNA involved</i>
LIMM	tRNA Thr
Maternally inherited epilepsy	tRNA Phe
Maternally inherited essential hypertension	tRNA Ile
Maternally inherited hearing loss	tRNA Ser (UCN)
Maternally inherited non-syndromic deafness	tRNA His
MELAS	tRNA Leu (CUN), tRNA Leu (UUR), tRNA (Lys), tRNA (Val)
MELAS/DM	tRNA Leu (UUR)
MELAS/LHON/DEAF/hypertension helper	tRNA Glu
MELAS/LS/DMD/MD/ID/SNHL/CPEO/MM/FSGS/Cardiac + multi-organ dysfunction	tRNA Leu (UUR)
MELAS/MM	tRNA Phe
MELAS/myopathy	tRNA Leu (UUR)
MELAS/myopathy/deafness + cognitive impairment	tRNA Leu (UUR)
MELAS-like encephalopathy + bilateral optic atrophy	tRNA His
MELAS + stroke-like episodes and cortical blindness + MRI shows occipital lobe infarct	tRNA Trp
MEPR	tRNA Asp
MERRF	tRNA Lys, tRNA Phe, tRNA Leu (UUR)
MERRF Other – LD/depressive mood disorder/leukoencephalopathy/HiCM	tRNA Lys
MERRF-like disease	tRNA Pro
MERRF-MELAS/enkephalopathy	tRNA His
MHCM	tRNA Gly, tRNA Ile
MICM	tRNA His, tRNA Ile
MICM + DEAF/MERRF/autism/LS/ataxia + lipomas	tRNA Lys
Migraine + pigmentary retinopathy + deafness + leukariosis	tRNA Pro
Mito encephalomyopathy	tRNA Trp
Mito encephalopathy/EXIT with myopathy and ptosis	tRNA Lys
Mito leukoencephalopathy	tRNA Glu
Mito myopathy with respiratory failure	tRNA Glu
Mitochondrial cytopathy	tRNA Lys, tRNA Pro
Mitochondrial encephalocardiomyopathy	tRNA Ile
Mitochondrial encephalomyopathy	tRNA Glu
Mitochondrial encephalopathy	tRNA Cys
Mitochondriahypertension factoral myopathy	tRNA Arg, tRNA Asp, tRNA Phe
MM	tRNA Leu, tRNA Met, tRNA Phe, tRNA Pro, tRNA Ser (UCN), tRNA Thr, tRNA Trp
MM/CPEO	tRNA Ile, tRNA Leu (UUR)
MM/DMD/MD/ID modulator	tRNA Ser (UCN)
MM/EXIT	tRNA Ser (UCN)
MM/HCM + renal tubular dysfunction	tRNA Leu (UUR)
MM/MELAS/SNHL/CPEO	tRNA Leu (UUR)
MM + DMD/MD/ID/encephalomyopathy/dementia + diabetes + ophthalmoplegia	tRNA Glu
MMC	tRNA Leu (UUR)
MMC/MELAS	tRNA Leu (UUR)
MNGIE/progressive mito cytopathy	tRNA Lys
MNGIE-like disease/MELAS	tRNA Val
Movement disorder	tRNA Val
Multiorgan failure	tRNA Asn
Multiple sclerosis/DEAF1555 increased penetrance	tRNA Thr
Multiple sclerosis/idiopathic repeat miscarriage/AD protection	tRNA Thr
Multisystem disease with cataracts/myopathy + epilepsy + DEAF + atypical autism	tRNA Ser (AGY)
Myoglobinuria	tRNA Phe
Myopathy	tRNA Ala, tRNA Gln, tRNA Gly, tRNA Leu (UUR), tRNA Lys, tRNA Met, tRNA Trp
Myopathy/encephalopathy	tRNA Ser (AGY)
Myopathy/exercise intolerance	tRNA Lys
Myopathy deafness	tRNA Cys
Myopathy + ataxia + nystagmus + migraines + lactic acidosis	tRNA Asp
Myopathy + epilepsy + retinal degeneration + DEAF	tRNA Ser (AGY)
Neonatal onset mito disease	tRNA Trp
Neuropsychiatric syndrome + cataract	tRNA Leu (UUR)
Ocular myopathy	tRNA Leu (UUR)

(Continued)

Table 2 Continued

<i>Disease</i>	<i>mt-tRNA involved</i>
PEM	tRNA Gly, tRNA Leu (UUR)
PEM/AMDF/motor neuron disease-like	tRNA Ser (UCN)
PEM/MERME	tRNA Ser (UCN)
PEO	tRNA Ala, tRNA Ser (UCN)
PEO and myoclonus	tRNA Lys
PEO with hearing loss	tRNA Ser (UCN)
Possible hypertension factor	tRNA Gln, tRNA Ile, tRNA Leu (UUR), tRNA Lys, tRNA Met
Possible contributor to mito dysfunction/hypertension	tRNA Met
Possible DEAF modifier	tRNA Ala
Possible PD risk factor	tRNA Lys
Possibly associated w DEAF + RP + dev delay/hypertension	tRNA Gln
Possibly LVNC-associated	tRNA Gln, tRNA Thr
Progressive dystonia	tRNA Cys
Progressive encephalopathy	tRNA Arg, tRNA Glu, tRNA Ile
Progressive MM + deafness + seizures	tRNA Ser (AGY)
Ptosis CPEO MM	tRNA Phe
Recurrent myoglobinuria	tRNA Ile
Retinopathy + diabetes + dysphagia + cerebral atrophy	tRNA Ile
Reversible COX deficiency myopathy	tRNA Glu
RP + DEAF	tRNA His
SIDS	tRNA Gly
SNHL	tRNA Cys, tRNA Ser (UCN)
SNHL and epilepsy	tRNA Phe
Sporadic bilateral optic neuropathy	tRNA Asp, tRNA Leu (UUR)
Suspected mito disease	tRNA Gln, tRNA Val
Tubulo-interstitial nephritis	tRNA Phe
Varied familial presentation/spastic paraparesis	tRNA Ile
tRNA Modification Related Human Disease	Modification, Gene involved
Neurological intellectual disability	2'-O-methylribose, FTSJ1b m ² G, TRM1 m ⁵ C, NSUN2 m ⁷ G, WDR4c A-to-I editing, ADAT3 mcm ⁵ s ² U, IKBKAP mcm ⁵ s ² U, ELP3 mcm ⁵ s ² U, ELP4 m ⁵ C, NSUN2 m ⁵ C, NSUN2 mcm ⁵ s ² U, IKBKAP m ⁵ C, NSUN2 wybutosine, TRMT12 m ¹ G, HRG9MTD2e mcm ⁵ U, HABH8 (HALKBH8) mcm ⁵ U, HTRM9L m ⁵ C, DNMT2 mS ² t ⁶ A, CDKAL1 rm ⁵ U, mt-tRNA Leu (UAA) tm ⁵ S ² U, mt-tRNA Lys (UUU) S ² U, MTU1 (TRMU)
Familial dysautonomia	
Amyotrophic lateral sclerosis	
Rolandic epilepsy	
Dubowitz-like syndrome	
Cardiac Noonan-like syndromed	
Respiratory bronchial asthma	
Cancer Skin, breast, and colorectal	
Breast cancer	
Colorectal cancer	
Urothelial cancer	
Breast, bladder, colorectal, cervix, testicular cancer	
Epigenetic cancer	
Metabolic Type 2 diabetes	
MELAS	
MERRF	
Infantile liver failure	
Aminoacyl-tRNA Synthetase Related Human Disease	Gene involved
Heart hypertrophic cardiomyopathy	AARS2 (Mitochondrial)
Brain leukoencephalopathy with brain stem and spinal cord involvement and lactate elevation (LBSL)	DARS2 (Mitochondrial)
Brain leukoencephalopathy with thalamus and brainstem involvement and high lactate (LTBL)	EARS2 (Mitochondrial)
Brain Alpers encephalopathy	FARS2 (Mitochondrial)
Cochlea ovary progressive sensorineural hearing loss and ovarian dysgenesis (Perrault syndrome)	HARS2 (Mitochondrial)
Brain autosomal recessive spastic ataxia with leukoencephalopathy (ARSAL)	MARS2 (Mitochondrial)
Brain pontocerebellar hypoplasia type 6 (PCH6)	RARS2 (Mitochondrial)

(Continued)

Table 2 Continued

Disease	mt-tRNA involved
Kidney tubulopathy (hyperuricemia, metabolic alkalosis), pulmonary hypertension, and progressive renal failure (HUPRA)	SARS2 (Mitochondrial)
Muscle myopathy, lactic acidosis, and sideroblastic anemia (MLSA)	YARS2 (Mitochondrial)
Nerve Charcot-Marie-Tooth disease (CMT)	GARS, KARS, MARS, AARS, YARS

Abbreviations: ADPD, Alzheimer's disease and Parkinson disease; AMDF, ataxia, myopathy, and deafness; CPEO, chronic progressive ophthalmoplegia; DEAF/SNHL, deafness/sensorineural hearing loss; DEMCHO, dementia and chorea; DM, diabetes mellitus; DMDF, diabetes mellitus and deafness; FSGS, focal segmental glomerulosclerosis; GER, gastrointestinal reflux; MERRF, myoclonic epilepsy and ragged red fiber disease; MELAS, mitochondrial encephalomyopathy, lactic acidosis and stroke-like episodes; MICM, maternally inherited cardiomyopathy; MIDD, maternally inherited diabetes and deafness; MILS, maternally inherited leigh syndrome; MMC, mitochondrial myopathy and cardiomyopathy; MNGIE, mitochondrial neurogastrointestinal encephalopathy; PEM, progressive encephalomyopathy; RP, retinitis pigmentosa; SIDS, sudden infant death syndrome; SNHL, sensorineural hearing loss. mt, mitochondrial; m²G, N²,N²-dimethyl guanosine; m⁵C, 5-methylcytosine; m⁷G, 7-methylguanosine; mcm⁵s²U, 5-methoxycarbonylmethyl-2-thiouridine; m⁵U, 5-methyl uridine; m¹G, 1-methylguanosine; mcm⁵U, 5-methoxycarbonylmethyluridine; ms²t⁶A, 2-methylthio-N⁶-threonyl carbamoyl adenosine; tm⁵U, 5-taurinomethyluridine; tm⁵s²U, 5-taurinomethyl-2-thiouridine; s²U, 2-thiouridine. GARS and KARS are genes for glycyl-tRNA synthetase and lysyl-tRNA synthetase shared between cytoplasm and mitochondria. The table is extracted from the mitomap database and the references (Suzuki *et al.*, 2011; Torres and Batlle, 2014; Konovalova and Tyynismaa, 2013).

tolerance to structure/sequence alteration (mit-tRNAs are degenerative and more fragile to alteration). In addition, the multiple copies of mitochondrial genome in human cell may also allow the slow accumulation of mit-tRNA mutations through generations. As the final section of this article, we will summarize the current knowledge of defects in tRNA modification, mutations in mitochondrial tRNA in human disease, and the related therapeutic interventions.

Mitochondrial tRNA and Human Diseases

The human mitochondrial genome encodes 13 proteins, all subunits of the respiratory chain complexes involved in energy metabolism (Mercer *et al.*, 2011). These genes are translated by 22 tRNAs encoded by the mitochondrial genome (Putz *et al.*, 2007). It is the minimal set required for reading all codons. Human mitochondrial tRNAs gained interest with the rapid discovery of correlations between point mutations in their genes and various multisystemic disorders including neuromuscular and neurodegenerative disorders etc.

Recent full-spectrum surveys concluded that mutations in mt-tRNA genes contribute to the etiology of more than half of human disorders caused by mutations in the mitochondrial genome, while tRNA genes comprise only 10% of the mt-genome (Bannwarth *et al.*, 2013). A total of 249 mutations were detected in all 22 mt-tRNA genes and were found spread all over the structural domains of the corresponding tRNAs (Table 2). Nearly half of the mutations affect highly conserved nucleotides (Brandon *et al.*, 2005).

Considerable interest in mitochondrial tRNAs centers on the occurrence of diseases arising from mutations in their genes that lead to maternally inherited genetic disorders (Nakada *et al.*, 2001; Wittenhagen and Kelley, 2003; Sternberg *et al.*, 2001; Enriquez *et al.*, 1995). The diseases associated with mitochondrial tRNA mutations may arise from effects at different steps of tRNA biogenesis, such as failure in the processing of the tRNA (Levinger *et al.*, 2001), reduced stability of the tRNA (Hao and Moraes, 1997; Kelley *et al.*, 2000), or a reduction in aminoacylation (Enriquez *et al.*, 1995; Ling *et al.*, 2007; Vachon *et al.*, 1990), a reduced ability of the mutated aminoacyl-tRNA to interact with mitochondrial elongation factor Tu (mt-EF-Tu) (Ling *et al.*, 2007), and from the failure of

the tRNA to be correctly modified leading to translational defects (Kirino *et al.*, 2004). For example, MELAS (mitochondrial encephalomyopathy, lactic acidosis, and stroke-like episodes) is a maternally inherited disease that is caused by mutations in mitochondrial genes resulting in respiratory defects arising from complex I (and complex IV) deficiencies. Approximately 90% of cases are caused by point mutations in the mitochondrial mt-tRNA^{Leu(UUR)} gene (Goto *et al.*, 1990; Kobayashi *et al.*, 1990; Gotz *et al.*, 2011; Shoffner *et al.*, 1990). The most common mutation, occurring in roughly 80% of patients, have an A to G mutation at position 3243 (A3243G) in the mt tRNA^{Leu(UUR)} gene (Goto *et al.*, 1990; Kobayashi *et al.*, 1990) and another 10% of patients have a T to C mutation at position 3271 (T3271C) in the same tRNA gene (Goto *et al.*, 1991). These mutations are believed to interfere with the efficient translation of mitochondrial complex I and IV proteins.

Emerging fundamental knowledge of the structure/function relationships of these particular tRNAs has expedited our understanding of etiology of mitochondrial disorders (Suzuki *et al.*, 2011). A point mutation (T4409C) in the gene for human mitochondrial tRNA^{Met} (mt-tRNA^{Met}) has been found to cause mitochondrial myopathy resulting in dystrophic muscles and exercise intolerance (Vissing *et al.*, 1998). This mutation results in the replacement of U8 in mt-tRNA^{Met} with a C8, at the corner of the acceptor stem and D-stem of mt-tRNA^{Met}. Examination of the U8C mutation on the structure and function of mt-tRNA^{Met} shows the mutations disrupts a critical Mg²⁺-binding site on the tRNA such that the mutant can no longer form the standard cloverleaf structure in the presence of Mg²⁺. Mt-tRNA^{Met} is the only tRNA^{Met} in human mitochondria, serving as both the initiator and elongator tRNA^{Met}, and this mutant consequently results in defective translation initiation and elongation. Combination of this mutation with other mutations in the mitochondrial DNA further increases the severity of disease and this is seen in Leber's Hereditary Optic Neuropathy (Jones *et al.*, 2008). Similarly another mutation (G4450A) leads to loss of the final base pair in the T-stem of mt-tRNA^{Met}. This mutation presents as splenic lymphoma, is largely confined to lymphocyte cells, and results in severely abnormal mitochondria leading to serious defects in energy production (Lombes *et al.*, 1998).

Another disease associated mutation (A4435G) was found to change A37 to G37 in the anticodon loop of tRNA^{Met} (Qu *et al.*, 2006). An overview of the large variety of mechanisms within translation, affected by mutations can also be found elsewhere (Suzuki *et al.*, 2011).

tRNA Modification and Human Diseases

tRNAs are heavily modified post-transcriptionally during their maturation process and many of these modifications are kingdom specific (Novoa *et al.*, 2012). In eukaryotes there are more than 50 different chemical modifications described affecting different positions on the tRNA (The tRNA Modification Database, see Relevant Websites). Most of these modifications and the enzymes responsible for catalyzing them are well described in the yeast (Phizicky and Hopper, 2010) and human homologs (Phizicky and Hopper, 2010). Particularly, the biological roles of the modifications they catalyze and the link to human diseases have been documented (Torres and Batlle, 2014).

In contrast to mitochondrial tRNA genes, diseases associated with cytosolic tRNA mutations are less frequent, however, diseases associated with aberrant modification of cytosolic tRNAs are more common. There are nearly 20 different tRNA modification enzymes that have been associated with 14 different human diseases, including neurological disorders (Intellectual disability, Familial dysautonomia, Amyotrophic lateral sclerosis, etc.), cardiac, respiratory diseases, diabetes, and cancer. For example, mutations in the tRNA modification enzymes FtsJ RNA methyltransferase homolog 1 (FTSJ1) and tRNA methyltransferase 1 (TRM1) have been found with strong association with intellectual disability. *FTSJ1* and *TRM1* genes are responsible for methylation of tRNA^{Leu} and tRNA^{Phe} at positions 32 and 34, and dimethylating guanosines (m²G) at position 26 of tRNA^{Tyr} (Townsend and Begley, 2012; Feder *et al.*, 2003). In addition, NSUN2, a tRNA methyltransferase, was first described as a downstream target of the proto-oncogene Myc and shown to be responsible for Myc-induced keratinocyte proliferation and cell cycle progression (Ador *et al.*, 1999). NSUN2 is expressed at low levels in normal tissues, but it is abundant in a range of human and mice tumor types, including squamous cell carcinoma, colorectal cancer, and breast cancer (Frye and Watt, 2006). Indeed, knockdown of NSUN2 was shown to reduce the growth of human squamous cell carcinoma xenografts in nude mice (Frye and Watt, 2006). tRNA methyltransferase homolog 12 (TRMT12) is involved in formation of wybutosine at position 37 on tRNA^{Phe} (Townsend and Begley, 2012). This gene is amplified in breast cancer cell lines and overexpressed in 26 out of 30 analyzed breast cancer tumors (Rodriguez *et al.*, 2007). TRMT2A (a paralog of TRMT12) gene is actually used as one of the five biomarkers in Mammostrat, a diagnosis tool for predicting breast cancer recurrence after tamoxifen treatment (Bartlett *et al.*, 2010). Other cancer-related tRNA modification genes include *HRG9MTD2* (Human RNA (guanine-9-) methyltransferase domain containing 2), *DNMT2* (DNA methyltransferase 2) (Schaefer *et al.*, 2009; Goll *et al.*, 2006; Himeno *et al.*, 1989; Sokabe *et al.*, 2009), *HABH8* (*HALKBH8*, human

AlkB homolog 8), and *HTRM9L* (tRNA methyltransferase 9-like) (Torres and Batlle, 2014).

Disease-associated tRNA modification could also affect mt-tRNAs. For example, both MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis, and stroke-like episodes) and MERRF (myoclonus epilepsy associated with ragged-red fibers) are characterized by lack of modifications on mt-tRNAs. The mutations in the mt-tRNA genes that cause both diseases lie outside of the tRNA anticodon, but result in hypomodification of uridine 34. In particular, MELAS patients lack 5-taurinomethyluridine (tm⁵U) on mt-tRNA^{Leu(UAA)}, whereas MERRF patients lack 5-taurinomethyl-2-thiouridine (tm⁵s²U) on mt-tRNA^{Lys(UUU)} (Suzuki *et al.*, 2011).

Therapeutics of tRNA-Related Diseases

Various potential approaches have been suggested to treat tRNA-related human diseases. For example, a decrease of aminoacylation levels of affected mitochondrial tRNA have been demonstrated in patients' tissues and in cultured cells, as well as for pathogenic mutations in nuclear genes encoding mitochondrial aminoacyl-tRNA-synthetases (Belostotsky *et al.*, 2012). Consequently, over-expression of mitochondrial aminoacyl-tRNA synthetases or elongation factor EF-Tu rescued mutated tRNAs from degradation (Gutman, 1976; Rinaldi *et al.*, 1997; Li and Guan, 2010). Similarly, because tRNA modifications can affect translation accuracy and efficiency as well as general tRNA stability, overexpression of tRNAs could potentially be beneficial for diseases where the target tRNA cannot be properly folded or modified (Karicheva *et al.*, 2011). A very interesting approach can be learned from the MELAS system. It was found that whereas the U34 could not be modified in the mutant mt-tRNA^{Leu(UAA)}, the other mt-tRNA^{Leu} isoacceptor (UAG) carrying a mutated anticodon to (UAA) could be modified to tm⁵U34. This modified mutated mt-tRNA^{Leu} isoacceptor restored the mitochondrial deficiencies of MELAS in a lung carcinoma cell line bearing 99% MELAS mutant mt-DNA (Kirino *et al.*, 2006). Finally, gene therapy could be used to express tRNA modification enzymes in those cases with defective mutations in these enzymes (Shimizu *et al.*, 2014).

On the other hand, modulation of tRNA modification enzymes could be a promising therapeutic strategy to treat cancer. For example, the down-regulation of NSUN2 (Frye and Watt, 2006) or HABH8 (Shimada *et al.*, 2009) has been shown to inhibit squamous cell carcinoma and bladder cancer. Overexpression of HTRM9L (Begley *et al.*, 2013) could inhibit colorectal cancer. Similar strategies have also been suggested to use for targeting TRMT12 against breast cancer (Torres and Batlle, 2014; Rodriguez *et al.*, 2007). There are other suggested approaches to develop diagnosis tool kit of cancer, based on levels of the tRNA modification enzymes, or hypomodified tRNAs themselves. Practical tools for tRNA modification detection such as 'miCLIP' (methylation individual-nucleotide resolution crosslinking and immunoprecipitation) and 5-azacytidine-mediated RNA immunoprecipitation (Aza-IP) have been developed to detect m⁵C modifications in transcriptomes (Hussain *et al.*, 2013; Khoddami and Cairns, 2013). Mass spectrometry-based technique could now monitor the levels of a whole collection of tRNA modifications (Chan *et al.*, 2010).

Other targets such as *HRG9MTD2* and the levels of m¹G9 modification (in early-onset and late-onset colorectal cancer) (Berg *et al.*, 2010), the levels of m⁵U at positions 42 and 54 of tRNAs (for breast cancer) (Bartlett *et al.*, 2010), combined with direct RNA sequencing have been suggested to create a personal 'epi-tRNAomes' for cancer diagnosis (Torres and Batlle, 2014).

Other work using mitochondrial gene transfer in mice (trans-mitochondrial mice), showed that mutation of mtDNA is randomly segregated during maternal inheritance, where subsequent generations with high proportions of mutant mtDNA exclusively expressed disease-related phenotypes. They also showed that the proportion of mutated mtDNA varied markedly among the pups born to each dam, suggesting that selecting oocytes with a high proportion of normal mtDNA from mothers affected with mutant-tRNA-based mitochondrial diseases may be an effective prevention (Shimizu *et al.*, 2014).

Concluding Remarks

It has been 56 years since Francis Crick first predicted the existence of tRNA in 1958 (Crick, 1958) and since that time a tremendous amount of knowledge has been learned about tRNA biology. This progress is attributed to the collaborative work in biochemistry, molecular biology, cell biology, structure and genomic studies by scientists worldwide, which unfortunately could not be comprehensively covered in this book section. tRNAs or a precursor most likely existed in the very beginning of life on planet earth, and these nucleic acid molecules have overseen, and in fact, carried out, the transition from the primordial RNA world to the current world. As the link between amino acids and nucleic acids, tRNAs determine the genetic code, but their functions have expanded beyond protein translation and include a role in cell stress response, bacterial cell wall biosynthesis, and viral replication. Not surprisingly defects in these critical molecules are also involved in various human diseases. With the exponential rise of information about tRNA sequences in the post-genomic era, the increasing recognition of the multiplexed roles of RNAs in the modern life, we are watching a new surge of tRNA research in multiple fields from protein translation to the extreme complexity of biology.

Acknowledgments

M.G. is supported by The Scripps Research Institute and the State of Florida. This work was supported by the Sidney Kimmel Foundation for Cancer Research, and grants from the National Institutes of Health GM100136 and GM106134 to M.G.

References

Abe, T., Ikemura, T., Sugahara, J., *et al.*, 2011. tRNADB-CE 2011: tRNA gene database curated manually by experts. *Nucleic Acids Research* 39 (database issue), D210–D213.

- Ador, L., Camasses, A., Erbs, P., *et al.*, 1999. Active site mapping of yeast aspartyl-tRNA synthetase by in vivo selection of enzyme mutations lethal for cell growth. *Journal of Molecular Biology* 288 (2), 231–242.
- Agrawal, R.K., Sharma, M.R., Kiel, M.C., *et al.*, 2004. Visualization of ribosome-recycling factor on the Escherichia coli 70S ribosome: Functional implications. *Proceedings of National Academy of Sciences of the United States of America* 101 (24), 8900–8905.
- Ahel, I., Korencic, D., Ibba, M., Söll, D., 2003. Trans-editing of mischarged tRNAs. *Proceedings of National Academy of Sciences of the United States of America* 100 (26), 15422–15427.
- Alberts, B., 2008. *Molecular Biology of the Cell*, fifth ed. New York, NY: Garland Science. xxxiii, 1268, G 40, 1–49 p.
- Allen, G.S., Frank, J., 2007. Structural insights on the translation initiation complex: Ghosts of a universal initiation complex. *Molecular Microbiology* 63 (4), 941–950.
- Amunts, A., Brown, A., Bai, X.C., *et al.*, 2014. Structure of the yeast mitochondrial large ribosomal subunit. *Science* 343 (6178), 1485–1489.
- An, S., Musier-Forsyth, K., 2004. Trans-editing of Cys-tRNA^{Pro} by Haemophilus influenzae YbaK protein. *Journal of Biological Chemistry* 279 (41), 42359–42362.
- Astrom, S.U., Bystrom, A.S., 1994. Rit1, a tRNA backbone-modifying enzyme that mediates initiator and elongator tRNA discrimination. *Cell* 79 (3), 535–546.
- Balzi, E., Choder, M., Chen, W.N., Varshavsky, A., Goffeau, A., 1990. Cloning and functional analysis of the arginyl-tRNA-protein transferase gene ATE1 of *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 265 (13), 7464–7471.
- Ban, N., Nissen, P., Hansen, J., Moore, P.B., Steitz, T.A., 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289 (5481), 905–920.
- Banerjee, R., Chen, S., Dare, K., *et al.*, 2010. tRNAs: Cellular barcodes for amino acids. *FEBS Letters* 584 (2), 387–395.
- Bannwarth, S., Procaccio, V., Lebre, A.S., *et al.*, 2013. Prevalence of rare mitochondrial DNA mutations in mitochondrial disorders. *Journal of Medical Genetics* 50 (10), 704–714.
- Bartlett, J.M., Thomas, J., Ross, D.T., *et al.*, 2010. Mammostrat as a tool to stratify breast cancer patients at risk of recurrence during endocrine therapy. *Breast Cancer Research* 12 (4), R47.
- Begley, U., Sosa, M.S., Avivar-Valderas, A., *et al.*, 2013. A human tRNA methyltransferase 9-like protein prevents tumour growth by regulating LIN9 and HIF1-α. *EMBO Molecular Medicine* 5 (3), 366–383.
- van Belkum, A., Abrahams, J.P., Pleij, C.W., Bosch, L., 1985. Five pseudoknots are present at the 204 nucleotides long 3' noncoding region of tobacco mosaic virus RNA. *Nucleic Acids Research* 13 (21), 7673–7686.
- Belostotsky, R., Frishberg, Y., Entelis, N., 2012. Human mitochondrial tRNA quality control in health and disease: A channelling mechanism? *RNA Biology* 9 (1), 33–39.
- Benson, T.E., Prince, D.B., Mutchler, V.T., *et al.*, 2002. X-ray crystal structure of *Staphylococcus aureus* FemA. *Structure* 10 (8), 1107–1115.
- Berg, M., Agesen, T.H., Thiis-Evensen, E., *et al.*, 2010. Distinct high resolution genome profiles of early onset and late onset colorectal cancer integrated with gene expression data identify candidate susceptibility loci. *Molecular Cancer* 9, 100.
- Beuning, P.J., Nagan, M.C., Cramer, C.J., *et al.*, 2002. Efficient aminoacylation of the tRNA(Ala) acceptor stem: Dependence on the 2:71 base pair. *RNA* 8 (5), 659–670.
- Biarrotte-Sorin, S., Maillard, A.P., Deletré, J., *et al.*, 2004. Crystal structures of *Weissella viridescens* FemX and its complex with UDP-MurNAc-pentapeptide: Insights into FemABX family substrates recognition. *Structure* 12 (2), 257–267.
- Biou, V., Yaremchuk, A., Tukalo, M., Cusack, S., 1994. The 2.9 Å crystal structure of *T. thermophilus* seryl-tRNA synthetase complexed with tRNA(Ser). *Science* 263 (5152), 1404–1410.
- Blaha, G., Stanley, R.E., Steitz, T.A., 2009. Formation of the first peptide bond: The structure of EF-P bound to the 70S ribosome. *Science* 325 (5943), 966–970.
- Bock, A., Forchhammer, K., Heider, J., Baron, C., 1991. Selenoprotein synthesis: An expansion of the genetic code. *Trends in Biochemical Sciences* 16 (12), 463–467.
- Borovinskaya, M.A., Pai, R.D., Zhang, W., *et al.*, 2007. Structural basis for aminoglycoside inhibition of bacterial ribosome recycling. *Nature Structural & Molecular Biology* 14 (8), 727–732.
- Brandon, M.C., Lott, M.T., Nguyen, K.C., *et al.*, 2005. MITOMAP: A human mitochondrial genome database – 2004 update. *Nucleic Acids Research* 33 (database issue), D611–D613.
- Brodersen, D.E., Clemons Jr., W.M., Carter, A.P., *et al.*, 2000. The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell* 103 (7), 1143–1154.

- Carraro, V., Maurin, A.C., Lambert-Langlais, S., *et al.*, 2010. Amino acid availability controls TRB3 transcription in liver through the GCN2/eIF2alpha/ATF4 pathway. *PLoS One* 5 (12), e15716.
- Castilho, B.A., Shanmugam, R., Silva, R.C., *et al.*, 2014. Keeping the eIF2 alpha kinase Gcn2 in check. *Biochimica et Biophysica Acta* 1843 (9), 1948–1968.
- Chan, C.T., Dyavaiah, M., DeMott, M.S., *et al.*, 2010. A quantitative systems approach reveals dynamic control of tRNA modifications during cellular stress. *PLoS Genetics* 6 (12), e1001247.
- Chan, P.P., Lowe, T.M., 2009. GtRNAdb: A database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Research* 37 (database issue), D93–D97.
- Chapeville, F., Lipmann, F., Von Ehrenstein, G., *et al.*, 1962. On the role of soluble ribonucleic acid in coding for amino acids. *Proceedings of National Academy of Sciences of the United States of America* 48, 1086–1092.
- Chong, Y.E., Yang, X.L., Schimmel, P., 2008. Natural homolog of tRNA synthetase editing domain rescues conditional lethality caused by mistranslation. *Journal of Biological Chemistry* 283 (44), 30073–30078.
- Crepin, T., Yaremchuk, A., Tukalo, M., Cusack, S., 2006. Structures of two bacterial prolyl-tRNA synthetases with and without a cis-editing domain. *Structure* 14 (10), 1511–1525.
- Crick, F.H., 1958. On protein synthesis. *Symposia of the Society for Experimental Biology* 12, 138–163.
- Crick, F.H., 1966. Codon–anticodon pairing: The wobble hypothesis. *Journal of Molecular Biology* 19 (2), 548–555.
- Crick, F.H., 1968. The origin of the genetic code. *Journal of Molecular Biology* 38 (3), 367–379.
- Curnow, A.W., Hong, K.W., Yuan, R., *et al.*, 1997a. Glu-tRNA^{Gln} amidotransferase: A novel heterotrimeric enzyme required for correct decoding of glutamine codons during translation. *Proceedings of National Academy of Sciences of the United States of America* 94 (22), 11819–11826.
- Curnow, A.W., Hong, K.W., Yuan, R., Söll, D., 1997b. tRNA-dependent amino acid transformations. *Nucleic Acids Symposium Series*, 36, 2–4.
- Curnow, A.W., Tumbula, D.L., Pelaschier, J.T., Min, B., Söll, D., 1998. Glutamyl-tRNA^{Gln} amidotransferase in *Deinococcus radiodurans* may be confined to asparagine biosynthesis. *Proceedings of National Academy of Sciences of the United States of America* 95 (22), 12838–12843.
- Cusack, S., Yaremchuk, A., Tukalo, M., 2000. The 2A crystal structure of leucyl-tRNA synthetase and its complex with a leucyl-adenylate analogue. *EMBO Journal* 19 (10), 2351–2361.
- Dare, K., Ibba, M., 2012. Roles of tRNA in cell wall biosynthesis. *Wiley Interdisciplinary Reviews: RNA* 3 (2), 247–264.
- Deiman, B.A., Verlaan, P.W., Pleij, C.W., 2000. In vitro transcription by the turnip yellow mosaic virus RNA polymerase: A comparison with the alfalfa mosaic virus and brome mosaic virus replicases. *Journal of Virology* 74 (1), 264–271.
- Dever, T.E., 2002. Gene-specific regulation by general translation factors. *Cell* 108 (4), 545–556.
- Dever, T.E., Green, R., 2012. The elongation, termination, and recycling phases of translation in eukaryotes. *Cold Spring Harbor Perspectives in Biology* 4 (7), a013706.
- Dong, J., Qiu, H., Garcia-Barrio, M., Anderson, J., Hinnebusch, A.G., 2000. Uncharged tRNA activates GCN2 by displacing the protein kinase moiety from a bipartite tRNA-binding domain. *Molecular Cell* 6 (2), 269–279.
- Dreher, T.W., 2009. Role of tRNA-like structures in controlling plant virus replication. *Virus Research* 139 (2), 217–229.
- Dreher, T.W., 2010. Viral tRNAs and tRNA-like structures. *Wiley Interdisciplinary Reviews: RNA* 1 (3), 402–414.
- Dreher, T.W., Bujarski, J.J., Hall, T.C., 1984. Mutant viral RNAs synthesized in vitro show altered aminoacylation and replicase template activities. *Nature* 311 (5982), 171–175.
- Dreher, T.W., Rao, A.L., Hall, T.C., 1989. Replication in vivo of mutant brome mosaic virus RNAs defective in aminoacylation. *Journal of Molecular Biology* 206 (3), 425–438.
- de Duve, C., 1988. Transfer RNAs: The second genetic code. *Nature* 333 (6169), 117–118.
- Enriquez, J.A., Chomyn, A., Attardi, G., 1995. Mitochondrial mutation in MERRF syndrome causes defective aminoacylation of tRNA(Lys) and premature translation termination. *Nature Genetics* 10 (1), 47–55.
- Eriani, G., Delarue, M., Poch, O., Gangloff, J., Moras, D., 1990. Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature* 347 (6289), 203–206.
- Farrell, R.E., 2010. RNA methodologies: Laboratory guide for isolation and characterization, fourth ed. Boston, MA: Elsevier/Academic Press. xxiii, 717 p.
- Fechter, P., Rudinger-Thirion, J., Tukalo, M., Giegé, R., 2001. Major tyrosine identity determinants in *Methanococcus jannaschii* and *Saccharomyces cerevisiae* tRNA (Tyr) are conserved but expressed differently. *European Journal of Biochemistry* 268 (3), 761–767.
- Feder, M., Pas, J., Wyrwicz, L.S., Bujnicki, J.M., 2003. Molecular phylogenetics of the RrmJ/fibrillarlin superfamily of ribose 2'-O-methyltransferases. *Gene* 302 (1–2), 129–138.
- Felden, B., Florentz, C., Giegé, R., Westhof, E., 1996. A central pseudoknotted three-way junction imposes tRNA-like mimicry and the orientation of three 5' upstream pseudoknots in the 3' terminus of tobacco mosaic virus RNA. *RNA* 2 (3), 201–212.
- Fernandez, I.S., Bai, X.C., Murshudov, G., Scheres, S.H., Ramakrishnan, V., 2014. Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell* 157 (4), 823–831.
- Fersht, A.R., 1977. Editing mechanisms in protein synthesis. Rejection of valine by the isoleucyl-tRNA synthetase. *Biochemistry* 16 (5), 1025–1030.
- First, E.A., 2005. Catalysis of the tRNA aminoacylation reaction. In: Ibba, M., Francklyn, C., Cusack, S. (Eds.), *The Aminoacyl-tRNA Synthetases*. Georgetown: Eurekah, pp. 328–352.
- Franckenberg, S., Becker, T., Beckmann, R., 2012. Structural view on recycling of archaeal and eukaryotic ribosomes after canonical termination and ribosome rescue. *Current Opinion in Structural Biology* 22 (6), 786–796.
- Francklyn, C.S., Minajigi, A., 2010. tRNA as an active chemical scaffold for diverse chemical transformations. *FEBS Letters* 584 (2), 366–375.
- Frank, J., Gonzalez Jr., R.L., 2010. Structure and dynamics of a processive Brownian motor: The translating ribosome. *Annual Review of Biochemistry* 79, 381–412.
- Frye, M., Watt, F.M., 2006. The RNA methyltransferase Misu (NSun2) mediates Myc-induced proliferation and is upregulated in tumors. *Current Biology* 16 (10), 971–981.
- Fukai, S., Nureki, O., Sekine, S., *et al.*, 2000. Structural basis for double-sieve discrimination of L-valine from L-isoleucine and L-threonine by the complex of tRNA(Val) and valyl-tRNA synthetase. *Cell* 103 (5), 793–803.
- Fukunaga, R., Yokoyama, S., 2005a. Structural basis for non-cognate amino acid discrimination by the valyl-tRNA synthetase editing domain. *Journal of Biological Chemistry* 280 (33), 29937–29945.
- Fukunaga, R., Yokoyama, S., 2005b. Aminoacylation complex structures of leucyl-tRNA synthetase and tRNA^{Leu} reveal two modes of discriminator-base recognition. *Nature Structural & Molecular Biology* 12 (10), 915–922.
- Fukunaga, R., Yokoyama, S., 2005c. Crystal structure of leucyl-tRNA synthetase from the archaeon *Pyrococcus horikoshii* reveals a novel editing domain orientation. *Journal of Molecular Biology* 346 (1), 57–71.
- Fukunaga, R., Yokoyama, S., 2007. Structure of the AlaX-M trans-editing enzyme from *Pyrococcus horikoshii*. *Acta Crystallographica Section D: Biological Crystallography* 63 (Pt 3), 390–400.
- Gabor, J., Cen, S., Javanbakht, H., Niu, M., Kleiman, L., 2002. Effect of altering the tRNA(Lys)(3) concentration in human immunodeficiency virus type 1 upon its annealing to viral RNA, GagPol incorporation, and viral infectivity. *Journal of Virology* 76 (18), 9096–9102.
- Gabriel, K., Schneider, J., McClain, W.H., 1996. Functional evidence for indirect recognition of G.U in tRNA(Ala) by alanyl-tRNA synthetase. *Science* 271 (5246), 195–197.
- Gao, Y.G., Selmer, M., Dunham, C.M., *et al.*, 2009. The structure of the ribosome with elongation factor G trapped in the posttranslocational state. *Science* 326 (5953), 694–699.
- Garg, R.P., Gonzalez, J.M., Parry, R.J., 2006. Biochemical characterization of VlmL, a Seryl-tRNA synthetase encoded by the valanimycin biosynthetic gene cluster. *Journal of Biological Chemistry* 281 (37), 26785–26791.
- Garg, R.P., Qian, X.L., Alemany, L.B., Moran, S., Parry, R.J., 2008. Investigations of valanimycin biosynthesis: Elucidation of the role of seryl-tRNA. *Proceedings of National Academy of Sciences of the United States of America* 105 (18), 6543–6547.
- Giegé, R., Jühling, F., Pütz, J., *et al.*, 2012. Structure of transfer RNAs: Similarity and variability. *Wiley Interdisciplinary Reviews: RNA* 3 (1), 37–61.
- Giegé, R., Lorber, B., Ebel, J.P., Moras, D., Thierry, J.C., 1980. Crystallization of the complex formed between yeast aspartyl tRNA and its specific aminoacyl tRNA synthetase. *Comptes Rendus de l'Académie des Sciences* 291 (4), 393–396.
- Giegé, R., Sissler, M., Florentz, C., 1998. Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Research* 26 (22), 5017–5035.
- Goll, M.G., Kirkekar, F., Maggert, K.A., *et al.*, 2006. Methylation of tRNA^{Asp} by the DNA methyltransferase homolog Dnmt2. *Science* 311 (5759), 395–398.

- Gomez, J., Nadal, A., Sabariego, R., *et al.*, 2004. Three properties of the hepatitis C virus RNA genome related to antiviral strategies based on RNA-therapeutics: Variability, structural conformation and tRNA mimicry. *Current Pharmaceutical Design* 10 (30), 3741–3756.
- Gondry, M., Sauguet, L., Belin, P., *et al.*, 2009. Cyclodipeptide synthases are a family of tRNA-dependent peptide bond-forming enzymes. *Nature Chemical Biology* 5 (6), 414–420.
- Goto, Y., Nonaka, I., Horai, S., 1990. A mutation in the tRNA(Leu)(UUR) gene associated with the MELAS subgroup of mitochondrial encephalomyopathies. *Nature* 348 (6302), 651–653.
- Goto, Y., Nonaka, I., Horai, S., 1991. A new mtDNA mutation associated with mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes (MELAS). *Biochimica et Biophysica Acta* 1097 (3), 238–240.
- Gotz, A., Tyynismaa, H., Euro, L., *et al.*, 2011. Exome sequencing identifies mitochondrial alanyl-tRNA synthetase mutations in infantile mitochondrial cardiomyopathy. *American Journal of Human Genetics* 88 (5), 635–642.
- Grosjean, H., de Crecy-Lagard, V., Marck, C., 2010. Deciphering synonymous codons in the three domains of life: Co-evolution with specific tRNA modification enzymes. *FEBS Letters* 584 (2), 252–264.
- Gulyaev, A.P., van Batenburg, E., Pleij, C.W., 1994. Similarities between the secondary structure of satellite tobacco mosaic virus and tobamovirus RNAs. *Journal of General Virology* 75 (Pt 10), 2851–2856.
- Guo, F., Cen, S., Niu, M., Javanbakht, H., Kleiman, L., 2003. Specific inhibition of the synthesis of human lysyl-tRNA synthetase results in decreases in tRNA(Lys) incorporation, tRNA(3)(Lys) annealing to viral RNA, and viral infectivity in human immunodeficiency virus type 1. *Journal of Virology* 77 (18), 9817–9822.
- Guo, F., Gabor, J., Cen, S., *et al.*, 2005. Inhibition of cellular HIV-1 protease activity by lysyl-tRNA synthetase. *Journal of Biological Chemistry* 280 (28), 26018–26023.
- Guo, M., Chong, Y.E., Beebe, K., *et al.*, 2009. The C-Ala domain brings together editing and aminoacylation functions on one tRNA. *Science* 325 (5941), 744–747.
- Gur, E., Sauer, R.T., 2008. Evolution of the ssrA degradation tag in *Mycoplasma*: Specificity switch to a different protease. *Proceedings of National Academy of Sciences of the United States of America* 105 (42), 16113–16118.
- Gutman, M., 1976. The effect of opposing effectors on activation level of succinate dehydrogenase: Equilibrium and kinetic studies. *Biochemistry* 15 (6), 1342–1348.
- Hafez, M., Burger, G., Steinberg, S.V., Lang, B.F., 2013. A second eukaryotic group with mitochondrion-encoded tmRNA: In silico identification and experimental confirmation. *RNA Biology* 10 (7), 1117–1124.
- Hall, K.B., Sampson, J.R., Uhlenbeck, O.C., Redfield, A.G., 1989. Structure of an unmodified tRNA molecule. *Biochemistry* 28 (14), 5794–5801.
- Halwani, R., Cen, S., Javanbakht, H., *et al.*, 2004. Cellular distribution of Lysyl-tRNA synthetase and its interaction with Gag during human immunodeficiency virus type 1 assembly. *Journal of Virology* 78 (14), 7553–7564.
- Hao, H., Moraes, C.T., 1997. A disease-associated G5703A mutation in human mitochondrial DNA causes a conformational change and a marked decrease in steady-state levels of mitochondrial tRNA(Asn). *Molecular and Cellular Biology* 17 (12), 6831–6837.
- Harding, H.P., Novoa, I., Zhang, Y., *et al.*, 2000. Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Molecular Cell* 6 (5), 1099–1108.
- Harms, J., Schlutzen, F., Zarivach, R., *et al.*, 2001. High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. *Cell* 107 (5), 679–688.
- Heinemann, I.U., Soll, D., Randau, L., 2010. Transfer RNA processing in archaea: Unusual pathways and enzymes. *FEBS Letters* 584 (2), 303–309.
- Helm, M., Brulé, H., Friede, D., *et al.*, 2000. Search for characteristic structural features of mammalian mitochondrial tRNAs. *RNA* 6 (10), 1356–1379.
- Hershey, J.W., Sonenberg, N., Mathews, M.B., 2012. Principles of translational control: An overview. *Cold Spring Harbor Perspectives in Biology* 4, 12.
- Himeno, H., Hasegawa, T., Ueda, T., *et al.*, 1989. Role of the extra G-C pair at the end of the acceptor stem of tRNA(His) in aminoacylation. *Nucleic Acids Research* 17 (19), 7855–7863.
- Hinnebusch, A.G., Lorsch, J.R., 2012. The mechanism of eukaryotic translation initiation: New insights and challenges. *Cold Spring Harbor Perspectives in Biology* 4, 10.
- Hinnebusch, A.G., Natarajan, K., 2002. Gcn4p, a master regulator of gene expression, is controlled at multiple levels by diverse signals of starvation and stress. *Eukaryotic Cell* 1 (1), 22–32.
- Hoagland, M.B., Zamecnik, P.C., Stephenson, M.L., 1957. Intermediate reactions in protein biosynthesis. *Biochimica et Biophysica Acta* 24 (1), 215–216.
- Hurto, R.L., 2011. Unexpected functions of tRNA and tRNA processing enzymes. *Advances in Experimental Medicine and Biology* 722, 137–155.
- Hussain, S., Sajini, A.A., Blanco, S., *et al.*, 2013. NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. *Cell Reports* 4 (2), 255–261.
- Hussain, T., Kruparani, S.P., Pal, B., *et al.*, 2006. Post-transfer editing mechanism of a D-aminoacyl-tRNA deacylase-like domain in threonyl-tRNA synthetase from archaea. *EMBO Journal* 25 (17), 4152–4162.
- Ibba, M., Soll, D., 1999. Quality control mechanisms during translation. *Science* 286 (5446), 1893–1897.
- Jackman, J.E., Alfonzo, J.D., 2013. Transfer RNA modifications: Nature's combinatorial chemistry playground. *Wiley Interdisciplinary Reviews: RNA* 4 (1), 35–48.
- Jakubowski, H., Goldman, E., 1992. Editing of errors in selection of amino acids for protein synthesis. *Microbiological Reviews* 56 (3), 412–429.
- Janssen, B.D., Hayes, C.S., 2012. The tmRNA ribosome-rescue system. *Advances in Protein Chemistry and Structural Biology* 86, 151–191.
- Javanbakht, H., Halwani, R., Cen, S., *et al.*, 2003. The interaction between HIV-1 Gag and human lysyl-tRNA synthetase during viral assembly. *Journal of Biological Chemistry* 278 (30), 27644–27651.
- Jin, H., Kelley, A.C., Loakes, D., Ramakrishnan, V., 2010. Structure of the 70S ribosome bound to release factor 2 and a substrate analog provides insights into catalysis of peptide release. *Proceedings of National Academy of Sciences of the United States of America* 107 (19), 8593–8598.
- Jones, C.N., Jones, C.I., Graham, W.D., Agris, P.F., Spremulli, L.L., 2008. A disease-causing point mutation in human mitochondrial tRNA^{Met} results in tRNA misfolding leading to defects in translational initiation and elongation. *Journal of Biological Chemistry* 283 (49), 34445–34456.
- Karicheva, O.Z., Kolesnikova, O.A., Schirtz, T., *et al.*, 2011. Correction of the consequences of mitochondrial 3243A>G mutation in the MT-TL1 gene causing the MELAS syndrome by tRNA import into mitochondria. *Nucleic Acids Research* 39 (18), 8173–8186.
- Kavran, J.M., Gundllapalli, S., O'Donoghue, P., *et al.*, 2007. Structure of pyrrolysyl-tRNA synthetase, an archaeal enzyme for genetic code innovation. *Proceedings of National Academy of Sciences of the United States of America* 104 (27), 11268–11273.
- Keiler, K.C., 2008. Biology of trans-translation. *Annual Review of Microbiology* 62, 133–151.
- Kelley, S.O., Steinberg, S.V., Schimmel, P., 2000. Functional defects of pathogenic human mitochondrial tRNAs related to structural fragility. *Nature Structural Biology* 7 (10), 862–865.
- Khoddami, V., Cairns, B.R., 2013. Identification of direct targets and modified bases of RNA cytosine methyltransferases. *Nature Biotechnology* 31 (5), 458–464.
- Kim, K.K., Min, K., Suh, S.W., 2000. Crystal structure of the ribosome recycling factor from *Escherichia coli*. *EMBO Journal* 19 (10), 2362–2370.
- Kim, S.H., Suddath, F.L., Quigley, G.J., *et al.*, 1974. Three-dimensional tertiary structure of yeast phenylalanine transfer RNA. *Science* 185 (4149), 435–440.
- Kim, S.I., Nalaskowska, M., Germond, J.E., Pridmore, D., Söll, D., 1996. Asn-tRNA in *Lactobacillus bulgaricus* is formed by asparaginylation of tRNA and not by transamidation of Asp-tRNA. *Nucleic Acids Research* 24 (14), 2648–2651.
- Kirino, Y., Yasukawa, T., Ohta, S., *et al.*, 2004. Codon-specific translational defect caused by a wobble modification deficiency in mutant tRNA from a human mitochondrial disease. *Proceedings of National Academy of Sciences of the United States of America* 101 (42), 15070–15075.
- Kirino, Y., Yasukawa, T., Marjawaara, S.K., *et al.*, 2006. Acquisition of the wobble modification in mitochondrial tRNA^{Leu}(CUN) bearing the G12300A mutation suppresses the MELAS molecular defect. *Human Molecular Genetics* 15 (6), 897–904.
- Klasse, P.J., 2012. The molecular basis of HIV entry. *Cell Microbiology* 14 (8), 1183–1192.
- Kleiman, L., 2002. tRNA(Lys3): The primer tRNA for reverse transcription in HIV-1. *IUBMB Life* 53 (2), 107–114.
- Kleiman, L., Cen, S., 2004. The tRNA^{Lys} packaging complex in HIV-1. *International Journal of Biochemistry & Cell Biology* 36 (9), 1776–1786.
- Kleiman, L., Jones, C.P., Musier-Forsyth, K., 2010. Formation of the tRNA^{Lys} packaging complex in HIV-1. *FEBS Letters* 584 (2), 359–365.
- Klimov, P.B., O'Connor, B.M., 2009. Improved tRNA prediction in the American house dust mite reveals widespread occurrence of extremely short minimal tRNAs in acariform mites. *BMC Genomics* 10, 598.
- Kobayashi, Y., Momoi, M.Y., Tominaga, K., *et al.*, 1990. A point mutation in the mitochondrial tRNA(Leu)(UUR) gene in MELAS (mitochondrial myopathy, encephalopathy, lactic acidosis and stroke-like episodes). *Biochemical and Biophysical Research Communications* 173 (3), 816–822.
- Komine, Y., Kitabatake, M., Yokogawa, T., Nishikawa, K., Inokuchi, H., 1994. A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia*

- coli. Proceedings of National Academy of Sciences of the United States of America 91 (20), 9223–9227.
- Konovalova, S., Tyynismaa, H., 2013. Mitochondrial aminoacyl-tRNA synthetases in human disease. *Molecular Genetics and Metabolism* 108 (4), 206–211.
- Korencic, D., Ahel, I., Schelert, J., *et al.*, 2004. A freestanding proofreading domain is required for protein synthesis quality control in Archaea. *Proceedings of National Academy of Sciences of the United States of America* 101 (28), 10260–10265.
- Kovaleski, B.J., Kennedy, R., Hong, M.K., *et al.*, 2006. In vitro characterization of the interaction between HIV-1 Gag and human lysyl-tRNA synthetase. *Journal of Biological Chemistry* 281 (28), 19449–19456.
- Kovaleski, B.J., Kennedy, R., Khorchid, A., *et al.*, 2007. Critical role of helix 4 of HIV-1 capsid C-terminal domain in interactions with human lysyl-tRNA synthetase. *Journal of Biological Chemistry* 282 (44), 32274–32279.
- Kozak, M., 2005. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* 361, 13–37.
- Kryukov, G.V., Castellano, S., Novoselov, S.V., *et al.*, 2003. Characterization of mammalian selenoproteomes. *Science* 300 (5624), 1439–1443.
- Krzycki, J.A., 2005. The direct genetic encoding of pyrrolysine. *Current Opinion in Microbiology* 8 (6), 706–712.
- Kunst, F., Ogasawara, N., Moszer, I., *et al.*, 1997. The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature* 390 (6657), 249–256.
- Kurland, C.G., 1992. Evolution of mitochondrial genomes and the genetic code. *Bioessays* 14 (10), 709–714.
- Levinger, L., Jacobs, O., James, M., 2001. In vitro 3'-end endonucleolytic processing defect in a human mitochondrial tRNA(Ser(UCN)) precursor with the U7445C substitution, which causes non-syndromic deafness. *Nucleic Acids Research* 29 (21), 4334–4340.
- Li, R., Guan, M.X., 2010. Human mitochondrial leucyl-tRNA synthetase corrects mitochondrial dysfunctions due to the tRNA^{Leu}(UUR) A3243G mutation, associated with mitochondrial encephalomyopathy, lactic acidosis, and stroke-like symptoms and diabetes. *Molecular and Cellular Biology* 30 (9), 2147–2154.
- Ling, J., Reynolds, N., Ibba, M., 2009. Aminoacyl-tRNA synthesis and translational quality control. *Annual Review of Microbiology* 63, 61–78.
- Ling, J., Roy, H., Qin, D., *et al.*, 2007. Pathogenic mechanism of a human mitochondrial tRNA^{Phe} mutation associated with myoclonic epilepsy with ragged red fibers syndrome. *Proceedings of National Academy of Sciences of the United States of America* 104 (39), 15299–15304.
- Ling, J., So, B.R., Yadavalli, S.S., *et al.*, 2009. Resampling and editing of mischarged tRNA prior to translation elongation. *Molecular Cell* 33 (5), 654–660.
- Lomakin, I.B., Steitz, T.A., 2013. The initiation of mammalian protein synthesis and mRNA scanning mechanism. *Nature* 500 (7462), 307–311.
- Lombes, A., Bories, D., Girodon, E., *et al.*, 1998. The first pathogenic mitochondrial methionine tRNA point mutation is discovered in splenic lymphoma. *Human Mutation* Suppl 1, S175–S183.
- Lovato, M.A., Swairjo, M.A., Schimmel, P., 2004. Positional recognition of a tRNA determinant dependent on a peptide insertion. *Molecular Cell* 13 (6), 843–851.
- Mak, J., Kleiman, L., 1997. Primer tRNAs for reverse transcription. *Journal of Virology* 71 (11), 8087–8095.
- Malyshev, D.A., Dhami, K., Laverne, T., *et al.*, 2014. A semi-synthetic organism with an expanded genetic alphabet. *Nature* 509 (7500), 385–388.
- Marck, C., Grosjean, H., 2002. tRNomics: Analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA* 8 (10), 1189–1232.
- Marquet, R., Isel, C., Ehresmann, C., Ehresmann, B., 1995. tRNAs as primer of reverse transcriptases. *Biochimie* 77 (1–2), 113–124.
- Mercer, T.R., Neph, S., Dinger, M.E., *et al.*, 2011. The human mitochondrial transcriptome. *Cell* 146 (4), 645–658.
- Min, B., Pelaschier, J.T., Graham, D.E., Tumbula-Hansen, D., Söhl, D., 2002. Transfer RNA-dependent amino acid biosynthesis: An essential route to asparagine formation. *Proceedings of National Academy of Sciences of the United States of America* 99 (5), 2678–2683.
- Mogk, A., Schmidt, R., Bukau, B., 2007. The N-end rule pathway for regulated proteolysis: Prokaryotic and eukaryotic strategies. *Trends in Cell Biology* 17 (4), 165–172.
- Morris, S., Leis, J., 1999. Changes in Rous sarcoma virus RNA secondary structure near the primer binding site upon tRNA^{Trp} primer annealing. *Journal of Virology* 73 (8), 6307–6318.
- Moulinier, L., Eiler, S., Eriani, G., *et al.*, 2001. The structure of an AspRS-tRNA(Asp) complex reveals a tRNA-dependent control mechanism. *EMBO Journal* 20 (18), 5290–5301.
- Murayama, K., Kato-Murayama, M., Katsura, K., *et al.*, 2005. Structure of a putative trans-editing enzyme for prolyl-tRNA synthetase from *Aeropyrum pernix* K1 at 1.7 Å resolution. *Acta Crystallographica. Section F, Structural Biology Communications* 61 (Pt 1), 26–29.
- Murguía, J.R., Serrano, R., 2012. New functions of protein kinase Gcn2 in yeast and mammals. *IUBMB Life* 64 (12), 971–974.
- Myasnikov, A.G., Simonetti, A., Marzi, S., Klaholz, B.P., 2009. Structure-function insights into prokaryotic and eukaryotic translation initiation. *Current Opinion in Structural Biology* 19 (3), 300–309.
- Naganuma, M., Sekine, S., Chong, Y.E., *et al.*, 2014. The selective tRNA aminoacylation mechanism based on a single G*U pair. *Nature* 510 (7506), 507–511.
- Naganuma, M., Sekine, S., Fukunaga, R., Yokoyama, S., 2009. Unique protein architecture of alanyl-tRNA synthetase for aminoacylation, editing, and dimerization. *Proceedings of National Academy of Sciences of the United States of America* 106 (21), 8489–8494.
- Nakada, K., Inoue, K., Hayashi, J.I., 2001. Mito-mice: Animal models for mitochondrial DNA-based diseases. *Seminars in Cell & Developmental Biology* 12 (6), 459–465.
- Nakamura, Y., Ito, K., 2003. Making sense of mimic in translation termination. *Trends in Biochemical Sciences* 28 (2), 99–105.
- Nakamura, Y., Ito, K., 2011. tRNA mimicry in translation termination and beyond. *Wiley Interdisciplinary Reviews: RNA* 2 (5), 647–668.
- Ninnis, R.L., Spall, S.K., Talbot, G.H., Truscott, K.N., Dougan, D.A., 2009. Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in *Escherichia coli*. *EMBO Journal* 28 (12), 1732–1744.
- Novoa, E.M., Pavon-Eternod, M., Pan, T., Ribas de Pouplana, L., 2012. A role for tRNA modifications in genome structure and codon usage. *Cell* 149 (1), 202–213.
- Nozawa, K., O'Donoghue, P., Gundllapalli, S., *et al.*, 2009. Pyrrolysyl-tRNA synthetase-tRNA(Pyl) structure reveals the molecular basis of orthogonality. *Nature* 457 (7233), 1163–1167.
- Nurenberg, E., Tampe, R., 2013. Tying up loose ends: Ribosome recycling in eukaryotes and archaea. *Trends in Biochemical Sciences* 38 (2), 64–74.
- Osman, T.A., Hemenway, C.L., Buck, K.W., 2000. Role of the 3' tRNA-like structure in tobacco mosaic virus minus-strand RNA synthesis by the viral RNA-dependent RNA polymerase *in vitro*. *Journal of Virology* 74 (24), 11671–11680.
- Parisien, M., Wang, X., Pan, T., 2013. Diversity of human tRNA genes from the 1000-genomes project. *RNA Biology* 10 (12), 1853–1867.
- Pauling, L., 1958. *Festschrift für Prof. Dr. Arthur Stoll*. Basel: Birkhauser-Verlag. pp. 597–602.
- Peng, W., Robertson, L., Gallinetti, J., *et al.*, 2012. Surgical stress resistance induced by single amino acid deprivation requires Gcn2 in mice. *Science Translational Medicine* 4 (118), 118ra11.
- Peters, G.G., Hu, J., 1980. Reverse transcriptase as the major determinant for selective packaging of tRNA's into Avian sarcoma virus particles. *Journal of Virology* 36 (3), 692–700.
- Petry, S., Weixlbaumer, A., Ramakrishnan, V., 2008. The termination of translation. *Current Opinion in Structural Biology* 18 (1), 70–77.
- Phizicky, E.M., Hopper, A.K., 2010. tRNA biology charges to the front. *Genes & Development* 24 (17), 1832–1860.
- Polikanov, Y.S., Blaha, G.M., Steitz, T.A., 2012. How hibernation factors RMF, HPF, and YfiA turn off protein synthesis. *Science* 336 (6083), 915–918.
- Popadin, K.Y., Mamirova, L.A., Kondrashov, F.A., 2007. A manually curated database of tetrapod mitochondrially encoded tRNA sequences and secondary structures. *BMC Bioinformatics* 8, 441.
- Putz, J., Dupuis, B., Sissler, M., Florentz, C., 2007. Mamit-tRNA, a database of mammalian mitochondrial tRNA primary and secondary structures. *RNA* 13 (8), 1184–1190.
- Putz, J., Puglisi, J.D., Florentz, C., Giegé, R., 1991. Identity elements for specific aminoacylation of yeast tRNA(Asp) by cognate aspartyl-tRNA synthetase. *Science* 252 (5013), 1696–1699.
- Qu, J., Li, R., Zhou, X., *et al.*, 2006. The novel A4435G mutation in the mitochondrial tRNA^{Met} may modulate the phenotypic expression of the LHON-associated ND4 G11778A mutation. *Investigative Ophthalmology & Visual Science* 47 (2), 475–483.
- Rai, R., Kashina, A., 2005. Identification of mammalian arginyltransferases that modify a specific subset of protein substrates. *Proceedings of National Academy of Sciences of the United States of America* 102 (29), 10123–10128.
- Ramakrishnan, V., 2002. Ribosome structure and the mechanism of translation. *Cell* 108 (4), 557–572.
- Rao, A.L., Hall, T.C., 1991. Interference in trans with brome mosaic virus replication by RNA-2 bearing aminoacylation-deficient mutants. *Virology* 180 (1), 16–22.

- Ray, B.K., Apirion, D., 1979. Characterization of 10S RNA: A new stable rna molecule from *Escherichia coli*. *Molecular and General Genetics* 174 (1), 25–32.
- Ribas de Pouplana, L., Schimmel, P., 2001. Two classes of tRNA synthetases suggested by sterically compatible dockings on tRNA acceptor stem. *Cell* 104 (2), 191–193.
- Rinaldi, T., Lande, R., Bolotin-Fukuhara, M., Frontali, L., 1997. Additional copies of the mitochondrial Ef-Tu and aspartyl-tRNA synthetase genes can compensate for a mutation affecting the maturation of the mitochondrial tRNA^{Asp}. *Current Genetics* 31 (6), 494–496.
- Robertus, J.D., Ladner, J.E., Finch, J.T., *et al.*, 1974. Structure of yeast phenylalanine tRNA at 3 Å resolution. *Nature* 250 (467), 546–551.
- Rodriguez, V., Chen, Y., Elkahoul, A., *et al.*, 2007. Chromosome 8 BAC array comparative genomic hybridization and expression analysis identify amplification and overexpression of TRMT12 in breast cancer. *Genes, Chromosomes & Cancer* 46 (7), 694–707.
- Rosen, A.E., Musier-Forsyth, K., 2004. Recognition of G-1:C73 atomic groups by *Escherichia coli* histidyl-tRNA synthetase. *Journal of the American Chemical Society* 126 (1), 64–65.
- Rould, M.A., Perona, J.J., Steitz, T.A., 1991. Structural basis of anticodon loop recognition by glutamyl-tRNA synthetase. *Nature* 352 (6332), 213–218.
- Ruan, B., Soll, D., 2005. The bacterial YbaK protein is a Cys-tRNA^{Pro} and Cys-tRNA^{Cys} deacylase. *Journal of Biological Chemistry* 280 (27), 25887–25891.
- Ruff, M., Krishnaswamy, S., Boeglin, M., *et al.*, 1991. Class II aminoacyl transfer RNA synthetases: Crystal structure of yeast aspartyl-tRNA synthetase complexed with tRNA^{Asp}. *Science* 252 (5013), 1682–1689.
- Saadatmand, J., Kleiman, L., 2012. Aspects of HIV-1 assembly that promote primer tRNA(Lys3) annealing to viral RNA. *Virus Research* 169 (2), 340–348.
- Saikrishnan, K., Kalapala, S.K., Varshney, U., Vijayan, M., 2005. X-ray structural studies of Mycobacterium tuberculosis RRF and a comparative study of RRFs of known structure. *Molecular plasticity and biological implications*. *Journal of Molecular Biology* 345 (1), 29–38.
- Schaefer, M., Hagemann, S., Hanna, K., Lyko, F., 2009. Azacytidine inhibits RNA methylation at DNMT2 target sites in human cancer cell lines. *Cancer Research* 69 (20), 8127–8132.
- Schimmel, P., Giegé, R., Moras, D., Yokoyama, S., 1993. An operational RNA code for amino acids and possible relationship to genetic code. *Proceedings of National Academy of Sciences of the United States of America* 90 (19), 8763–8768.
- Schimmel, P.R., Soll, D., 1979. Aminoacyl-tRNA synthetases: General features and recognition of transfer RNAs. *Annual Review of Biochemistry* 48, 601–648.
- Schmeing, T.M., Ramakrishnan, V., 2009. What recent ribosome structures have revealed about the mechanism of translation. *Nature* 461 (7268), 1234–1242.
- Schmeing, T.M., Voorhees, R.M., Kelley, A.C., *et al.*, 2009. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science* 326 (5953), 688–694.
- Schurer, H., Schiffer, S., Marchfelder, A., Mörl, M., 2001. This is the end: Processing, editing and repair at the tRNA 3'-terminus. *Biological Chemistry* 382 (8), 1147–1156.
- Selmer, M., Al-Karadaghi, S., Hirokawa, G., Kaji, A., Liljas, A., 1999. Crystal structure of Thermotoga maritima ribosome recycling factor: A tRNA mimic. *Science* 286 (5448), 2349–2352.
- Selmer, M., Dunham, C.M., Murphy 4th, F.V., *et al.*, 2006. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science* 313 (5795), 1935–1942.
- Shepherd, J., Ibba, M., 2013. Direction of aminoacylated transfer RNAs into antibiotic synthesis and peptidoglycan-mediated antibiotic resistance. *FEBS Letters* 587 (18), 2895–2904.
- Shimada, K., Nakamura, M., Anai, S., *et al.*, 2009. A novel human AlkB homologue, ALKBH8, contributes to human bladder cancer progression. *Cancer Research* 69 (7), 3157–3164.
- Shimizu, A., Mito, T., Hayashi, C., *et al.*, 2014. Transmitochondrial mice as models for primary prevention of diseases caused by mutation in the tRNA(Lys) gene. *Proceedings of National Academy of Sciences of the United States of America* 111 (8), 3104–3109.
- Shoffner, J.M., Lott, M.T., Lezza, A.M., *et al.*, 1990. Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA(Lys) mutation. *Cell* 61 (6), 931–937.
- Shpanchenko, O.V., Golovin, A.V., Bugaeva, E.Y., Isaksson, L.A., Dontsova, O.A., 2010. Structural aspects of trans-translation. *IUBMB Life* 62 (2), 120–124.
- Sikalidis, A.K., Mazor, K.M., Lee, J.I., *et al.*, 2014. Upregulation of capacity for glutathione synthesis in response to amino acid deprivation: Regulation of glutamate-cysteine ligase subunits. *Amino Acids* 46 (5), 1285–1296.
- Singh, R.N., Dreher, T.W., 1997. Turnip yellow mosaic virus RNA-dependent RNA polymerase: Initiation of minus strand synthesis in vitro. *Virology* 233 (2), 430–439.
- Singh, R.N., Dreher, T.W., 1998. Specific site selection in RNA resulting from a combination of nonspecific secondary structure and -CCR- boxes: Initiation of minus strand synthesis by turnip yellow mosaic virus RNA-dependent RNA polymerase. *RNA* 4 (9), 1083–1095.
- Sokabe, M., Okada, A., Yao, M., Nakashima, T., Tanaka, I., 2005. Molecular basis of alanine discrimination in editing site. *Proceedings of National Academy of Sciences of the United States of America* 102 (33), 11669–11674.
- Sokabe, M., Ose, T., Nakamura, A., *et al.*, 2009. The structure of alanyl-tRNA synthetase with editing domain. *Proceedings of National Academy of Sciences of the United States of America* 106 (27), 11028–11033.
- Steer, B.A., Schimmel, P., 1999. Major anticodon-binding region missing from an archaeobacterial tRNA synthetase. *Journal of Biological Chemistry* 274 (50), 35601–35606.
- Sternberg, D., Chatzoglou, E., Laforêt, P., *et al.*, 2001. Mitochondrial DNA transfer RNA gene sequence variations in patients with mitochondrial disorders. *Brain* 124 (Pt 5), 984–994.
- Suzuki, T., Nagao, A., Suzuki, T., 2011. Human mitochondrial tRNAs: Biogenesis, function, structural aspects, and diseases. *Annual Review of Genetics* 45, 299–329.
- Swanstrom, R., Wills, J.W., 1997. Synthesis, assembly, and processing of viral proteins. In: Coffin, J.M., Hughes, S.H., Varmus, H.E. (Eds.), *Retroviruses*. New York, NY: Cold Spring Harbor.
- Tasaki, T., Sriram, S.M., Park, K.S., Kwon, Y.T., 2012. The N-end rule pathway. *Annual Review of Biochemistry* 81, 261–289.
- Torres, A.G., Batlle, E., Ribas de Pouplana, L., 2014. Role of tRNA modifications in human diseases. *Trends in Molecular Medicine* 20 (6), 306–314.
- Tourigny, D.S., Fernández, I.S., Kelley, A.C., Ramakrishnan, V., 2013. Elongation factor G bound to the ribosome in an intermediate state of translocation. *Science* 340 (6140), 1235490.
- Towns, W.L., Begley, T.J., 2012. Transfer RNA methyltransferases and their corresponding modifications in budding yeast and humans: Activities, predilections, and potential roles in human health. *DNA and Cell Biology* 31 (4), 434–454.
- Toyoda, T., Tin, O.F., Ito, K., *et al.*, 2000. Crystal structure combined with genetic analysis of the *Thermus thermophilus* ribosome recycling factor shows that a flexible hinge may act as a functional switch. *RNA* 6 (10), 1432–1444.
- Tukalo, M., Yaremchuk, A., Fukunaga, R., Yokoyama, S., Cusack, S., 2005. The crystal structure of leucyl-tRNA synthetase complexed with tRNA^{Leu} in the post-transfer-editing conformation. *Nature Structural & Molecular Biology* 12 (10), 923–930.
- Vachon, G., Laalami, S., Grunberg-Manago, M., Julien, R., Ceniatiempo, Y., 1990. Purified internal G-domain of translational initiation factor IF-2 displays guanine nucleotide binding properties. *Biochemistry* 29 (41), 9728–9733.
- Valle, M., Gillet, R., Kaur, S., *et al.*, 2003. Visualizing tmRNA entry into a stalled ribosome. *Science* 300 (5616), 127–130.
- Varani, G., McClain, W.H., 2000. The G x U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Reports* 1 (1), 18–23.
- Varshney, U., Lee, C.P., RajBhandary, U.L., 1993. From elongator tRNA to initiator tRNA. *Proceedings of National Academy of Sciences of the United States of America* 90 (6), 2305–2309.
- Vissing, J., Salamon, M.B., Arlien-Søborg, P., *et al.*, 1998. A new mitochondrial tRNA(Met) gene mutation in a patient with dystrophic muscle and exercise intolerance. *Neurology* 50 (6), 1875–1878.
- Wang, C.C., Schimmel, P., 1999. Species barrier to RNA recognition overcome with nonspecific RNA binding domains. *Journal of Biological Chemistry* 274 (23), 16508–16512.
- Wang, W., Takimoto, J.K., Louie, G.V., *et al.*, 2007. Genetically encoding unnatural amino acids for cellular and neuronal studies. *Nature Neuroscience* 10 (8), 1063–1072.
- Watanabe, K., Toh, Y., Suto, K., *et al.*, 2007. Protein-based peptide-bond formation by aminoacyl-tRNA protein transferase. *Nature* 449 (7164), 867–871.
- Weixlbaumer, A., Jin, H., Neubauer, C., *et al.*, 2008. Insights into translational termination from the structure of RF2 bound to the ribosome. *Science* 322 (5903), 953–956.
- Wek, S.A., Zhu, S., Wek, R.C., 1995. The histidyl-tRNA synthetase-related sequence in the eIF-2 alpha protein kinase GCN2 interacts with tRNA and is required for activation in response to starvation for different amino acids. *Molecular and Cellular Biology* 15 (8), 4497–4506.

- Wilson, D.N., Schlutzenzen, F., Harms, J.M., *et al.*, 2005. X-ray crystallography study on ribosome recycling: The mechanism of binding and action of RRF on the 50S ribosomal subunit. *EMBO Journal* 24 (2), 251–260.
- Wilson, J.H., Hunt, T., 2008. *Molecular Biology of the Cell: The Problems Book*, fifth ed. New York, NY: Garland Science. xviii, 587 p.
- Wimberly, B.T., Brodersen, D.E., Clemons Jr., W.M., *et al.*, 2000. Structure of the 30S ribosomal subunit. *Nature* 407 (6802), 327–339.
- Wittenhagen, L.M., Kelley, S.O., 2003. Impact of disease-related mitochondrial mutations on tRNA structure and function. *Trends in Biochemical Sciences* 28 (11), 605–611.
- Woese, C.R., Olsen, G.J., Ibba, M., Söll, D., 2000. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiology and Molecular Biology Reviews* 64 (1), 202–236.
- Wolin, S.L., Matera, A.G., 1999. The trials and travels of tRNA. *Genes & Development* 13 (1), 1–10.
- Wolstenholme, D.R., Macfarlane, J.L., Okimoto, R., Clary, D.O., Wahleithner, J.A., 1987. Bizarre tRNAs inferred from DNA sequences of mitochondrial genomes of nematode worms. *Proceedings of National Academy of Sciences of the United States of America* 84 (5), 1324–1328.
- Wong, F.C., Beuning, P.J., Silvers, C., Musier-Forsyth, K., 2003. An isolated class II aminoacyl-tRNA synthetase insertion domain is functional in amino acid editing. *Journal of Biological Chemistry* 278 (52), 52857–52864.
- Xie, J., Schultz, P.G., 2005. Adding amino acids to the genetic repertoire. *Current Opinion in Chemical Biology* 9 (6), 548–554.
- Yoshida, T., Uchiyama, S., Nakano, H., *et al.*, 2001. Solution structure of the ribosome recycling factor from *Aquifex aeolicus*. *Biochemistry* 40 (8), 2387–2396.
- Yuan, J., Gogakos, T., Babina, A.M., Söll, D., Randau, L., 2011. Change of tRNA identity leads to a divergent orthogonal histidyl-tRNA synthetase/tRNA^{His} pair. *Nucleic Acids Research* 39 (6), 2286–2293.
- Yuan, M.L., Wei, D.D., Wang, B.J., Dou, W., Wang, J.J., 2010. The complete mitochondrial genome of the citrus red mite *Panonychus citri* (Acari: Tetranychidae): High genome rearrangement and extremely truncated tRNAs. *BMC Genomics* 11, 597.
- Zwieb, C., Wower, I., Wower, J., 1999. Comparative sequence analysis of tmRNA. *Nucleic Acids Research* 27 (10), 2063–2071.

Relevant Websites

<http://gtrnadb.ucsc.edu>
Genomic tRNA Database.

<http://mamit-trna.u-strasbg.fr/>
Mamit-tRNA.

<http://www.mitomap.org/>
Mitomap.

<http://mttrna.bioinf.uni-leipzig.de>
mitotRNAdb.

<http://www.rcsb.org/pdb/>
Protein Database.

<http://mods.rna.albany.edu/>
The tRNA Modification Database.

<http://trna.ie.niigata-u.ac.jp>
tRNADB-CE.

<http://tRNAdb.bioinf.uni-leipzig.de/>
tRNAdb.

Messenger RNA (mRNA): The Link between DNA and Protein

DJ Goss, Hunter College CUNY, New York, NY, USA

AV Domashevskiy, John Jay College of Criminal Justice, New York, NY, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

ARE AU-rich elements containing multiple copies of the sequence AUUUA, located in the 3' UTR.

Cap A methylated Guanosine at the 5' end of eukaryotic mRNA joined to the first base by a 5' to 5' triphosphate linkage.

Polycistronic mRNA The RNA transcript of a bacteria operon. It encodes several proteins.

UTR A sequence of nucleotides at the 5' or 3' end of mRNA that does not code for a protein product.

Messenger RNA (mRNA)

Messenger RNA is a large family of RNA molecules that are complementary to DNA molecules and convey genetic information from the DNA to be translated by ribosomes into proteins (Brenner *et al.*, 1961). mRNAs, like DNA, are nucleic acids that contain a specific sequence of nucleotides. These nucleotides are 'read' (translated) by ribosomes to assemble a polymer of amino acids, a protein. The mRNA plays a key role in the 'central dogma' of molecular biology, which deals with the transfer of sequence information from DNA to RNA to protein (Brenner *et al.*, 1961). There are a series of transfers possible as framed by this dogma. Transfers that occur normally in most cells are the transfer of information from DNA to RNA (transcription), DNA can be copied to DNA (DNA replication) or proteins can be synthesized from the information in mRNA (translation or protein synthesis). Rarely, information can be transferred from RNA to DNA, usually in the case of retroviruses (Baltimore, 1970; Temin and Mizutani, 1970). Particles of these viruses (e.g., HIV) contain two copies of RNA as their genetic information. A rare enzyme, reverse transcriptase, copies the RNA to a single strand of DNA which is then integrated into the host's DNA. Information is not transferred from proteins to either RNA or DNA. These transfers describe the general flow of biological information: DNA to RNA to protein. Messenger RNA is sometimes also called the Rosetta Stone of biological information because it is the link in transferring genetic information from DNA (nucleic acid) into protein.

While there are differences in both structure and mechanism of function of prokaryotic and eukaryotic mRNAs (Berg, 2002), there are also similarities (Figure 1). In mRNA, genetic information is encoded in a four-base alphabet of nucleotides, which form codons consisting of three bases each. Each codon codes for a specific amino acid, with the exception of stop codons which determine where protein synthesis terminates. The mRNA are translated by the ribosome which reads the codons. The start or initiator codon for both prokaryotes and eukaryotes is an AUG sequence and the sequence is read from the 5' to 3' direction. Eukaryotic mRNA generally code for a single protein (monocistronic) (Gerlinger *et al.*, 1977) while prokaryotic mRNA often code for a series of related proteins (polycistronic) on the same mRNA. Polycistronic (Gerlinger *et al.*, 1977) mRNA direct the more or less simultaneous synthesis of each of the

encoded polypeptides. For example, the *trp* operon (Hiraga and Yanofsky, 1972) is a DNA that is transcribed into mRNA that codes for six polypeptides that catalyze tryptophan synthesis. Messenger RNA are short-lived compared to DNA and the lifetime of the RNA serves as a regulatory site (Guhaniyogi and Brewer, 2001) of gene expression. After transcription, an mRNA molecule may be processed, edited, and transported prior to translation. Eukaryotic mRNA require transport and often extensive processing, while prokaryotic mRNA molecules do not (Clark and Pazdernik, 2013).

Prokaryotic mRNA Structure

Prokaryotic mRNA are synthesized in the cytoplasm and do not require transport from the nucleus (Clark and Pazdernik, 2013). They also do not require processing and can begin translation immediately after the transcription is complete. Most mRNA contain a sequence at the 5' end of the mRNA prior to the AUG start codon, termed 5' untranslated region (5' UTR) (Meijer and Thomas, 2002) and a region following the stop codon, the 3' UTR. Most prokaryotic mRNA contain a sequence in the 5' UTR to position ribosomes for translation. This sequence is named after its discoverers as the Shine–Dalgarno sequence.

Shine–Dalgarno Sequence

Most extensively studied in *Escherichia coli*, the Shine–Dalgarno (Shine and Dalgarno, 1974) sequence is a purine-rich tract of nucleotides located in the 5' UTR. This tract of 3–10 nucleotides is complementary to the pyrimidine-rich (Shine and Dalgarno, 1975) sequence of the 16S rRNA (ribosomal RNA) and is centered approximately 10 nucleotides upstream of the AUG start codon. Base pairing between the Shine–Dalgarno sequence of the mRNA and the 16S rRNA (Kozak, 1987) permit the ribosome to bind at the correct position on the mRNA in order to select the proper initiation codon. The start codon, AUG, codes for internal methionine residues as well as the initiating Met residue of the polypeptide; here the translation initiation site must be specified by more than the start codon. The Shine–Dalgarno sequence is present in nearly all prokaryotic mRNA; however, recent evidence has shown that there are prokaryotic mRNAs that lack a

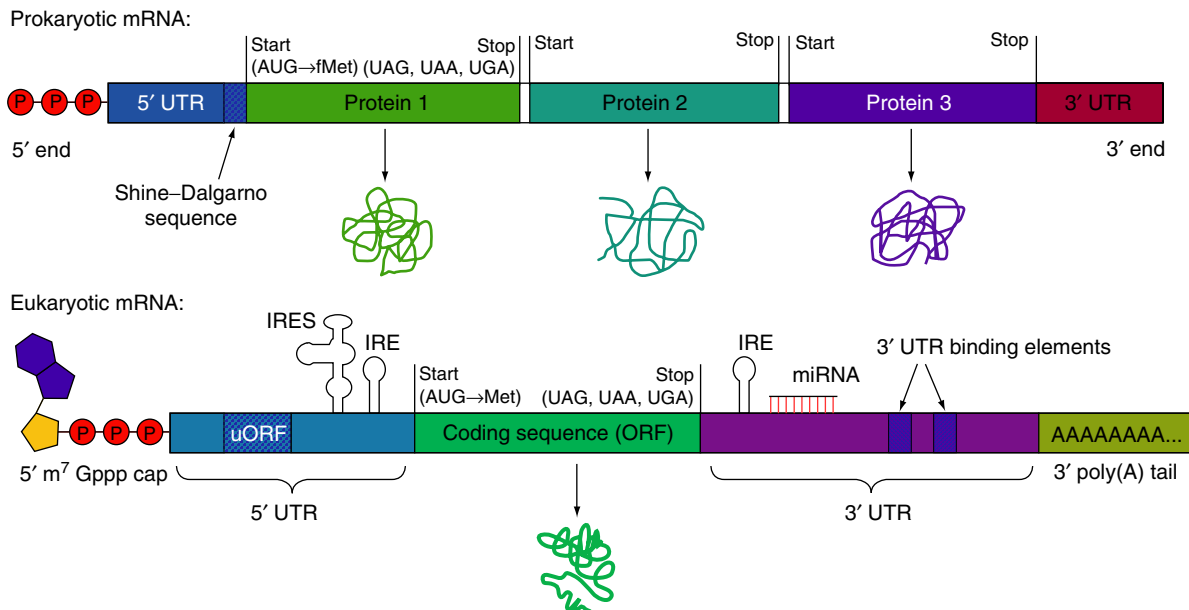


Figure 1 Schematic diagram of prokaryotic (top) and eukaryotic (bottom) mRNA. Bars indicate the relative length of the regions.

Shine–Dalgarno sequence in the 5' UTR (Londei, 2005) or lack a 5' UTR completely. These mRNAs appear to be more common in primitive prokaryotes, such as archaea, in which initiation of translation on leaderless transcripts is thought to be the evolutionary oldest mechanism. The mechanism of how these prokaryotes distinguish the start codon is not known.

Polycistronic mRNA

In prokaryotic cells, a single mRNA may code for several proteins. Each message on the mRNA is contained in a single 'open reading frame,' a sequence of codons bound by start and stop codons. There are no start or stop codons within the reading frame itself. The arrangement of messages in tandem along a single strand of mRNA allows the proteins (often called gene products) to be translated simultaneously; these gene products are often related in function. Because mRNAs are single stranded, some mRNA molecules are able to base-pair within themselves and can form secondary and tertiary three-dimensional structures. These structures can regulate the synthesis of polypeptides in the polycistronic mRNA. One example of this mechanism is MS2 bacteriophage (Kozak, 1983). The A protein is coded at the 5' end of the polycistronic message, but is needed in only small quantities. The 5' end of the mRNA is often blocked by tertiary folding of the mRNA allowing only limited translation of the A protein while allowing translation to occur at more accessible sites downstream from the A gene.

Eukaryotic mRNA

Processing

Unlike prokaryotic mRNAs which are produced and translated in the cytoplasm, eukaryotic mRNAs are produced in the

nucleus and translated in the cytosol, requiring transport. In addition, eukaryotic mRNAs undergo processing (Alifano *et al.*, 1994; Lodish, 2000), described in detail elsewhere. We will mainly focus on the structural features of the mRNA resulting from this processing. These 'pre-mRNA' undergo a number of modifications, the most striking of which is splicing, involving the excision of nonexpressed intervening sequences, following which flanking sequences are spliced or joined together. The mRNA itself also undergoes capping and addition of a poly(A) tail, the structural features of which are described below.

Cap

The cap is an altered nucleotide attached to the 5' end of mRNA (Figure 2). 5' capping is essential for creating mature mRNA for efficient translation. Capping occurs in the cell nucleus and is required for export of the mRNA to the cytosol for translation. The 5' cap itself consists of a guanosine residue methylated in the 7 position and attached to the 5' end of the mRNA through a 5' to 5' triphosphate linkage (Figure 2). Additional methylation may occur at the first two nucleosides of the transcript. In most multicellular eukaryotes, the first nucleoside is methylated at the O2' position, the 2' hydroxyl group of the first ribose sugar. Additional methylations can occur at the 2' hydroxyl group of the ribose sugar of the second nucleoside, and if the first nucleoside is an adenosine, it may also be methylated at the N6 position. The 5' cap functions to regulate nuclear export of the mRNA, prevent degradation by exonucleases, promote 5' exon excision, and promote efficient translation of the mRNA.

The cap is the binding site for eukaryotic initiation factor 4E (eIF4E) and this interaction is believed to be the rate-limiting step for protein synthesis for most eukaryotic mRNA.

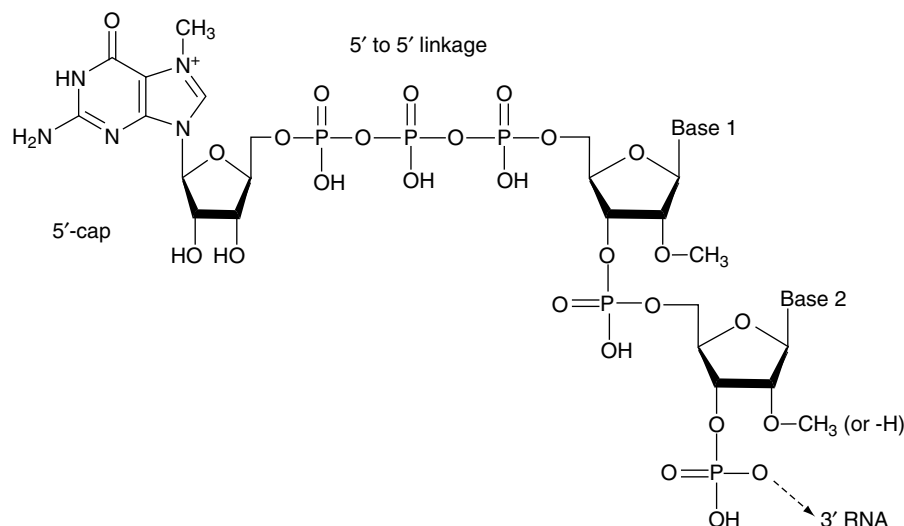


Figure 2 Chemical structure of m7G cap and first nucleotide. The linkage between the m7G and first nucleotide is a 5' to 5' triphosphate. Possible additional phosphorylation sites are indicated.

5' UTR

The 5' UTR, or leader sequence, begins at the 5' terminal end and ends one nucleotide before the AUG start site. In eukaryotes, this tends to be long, anywhere from 100 to several thousand nucleotides. In warm-blooded eukaryotes, this is a G C rich region with a G + C percentage of about 60%. This long stretch of nucleotides contains features often used to regulate translation of the mRNA. Unlike prokaryotic mRNA, eukaryotic mRNA do not contain a Shine–Dalgarno sequence, but usually contain the Kozak consensus sequence (ACCAUGG), which contains the initiator AUG and is the optimum sequence for ribosome initiation. While termed the untranslated region, a portion of the 5' UTR is sometimes translated to make a protein product which can then regulate translation of the main coding region of the mRNA. These coding sequences, termed upstream open reading frames (uORF) contain their own initiation codon and are fairly common as a means to regulate translation of the mRNA, occurring in 35–50% of all human mRNA (Chatterjee and Pal, 2009).

In other mRNA, the 5' UTR region is untranslated but can form secondary structures to regulate translation. One of the most studied examples of this is the iron response element (IRE). Iron levels are maintained by translation of proteins involved in iron storage and metabolism. The 5' UTR contains a stem loop structure, the IRE, which binds iron regulatory proteins 1 and 2 (This element also occurs in the 3' UTR of some mRNA coding for proteins involved in iron regulation.). When cellular iron levels are low, iron response protein (IRP) binds to the IRE element and blocks translation by steric hindrance of the ribosome binding. At higher iron levels, the IRE RNA element binds iron (Ma *et al.*, 2012) causing release of the IRP protein and enhanced binding of protein synthesis initiation factor 4F (eIF4F) to increase synthesis of proteins such as ferritin, required for storage of iron. In general, secondary structures in the 5' UTR are thought to inhibit translation by making scanning of the ribosome from the 5' cap to the ORF AUG more difficult.

A further means of regulating translation through mRNA structure is the internal ribosome entry site (IRES). These nucleotide sequences allow the ribosome to bind internally on the 5' UTR, rather than at the cap. First discovered in viruses, these sequences have since been found in cellular mRNA, often those involved in cell growth regulation and mitosis. The secondary structures of these regions can be quite complex and to date there is no consensus motif. The mechanism of IRES-dependent translation is much better understood in viruses than in cellular mRNA and continues to be an area of active research (Hellen and Sarnow, 2001).

3' UTR

The 3' UTR is comprised of a series of bases following the stop codon that are not translated into protein. Like the 5' UTR, this region contains numerous regulatory elements which can influence translational efficiency, mRNA stability, mRNA localization, and polyadenylation. The poly(A) tail is a sequence of -A- residues added to the 3' end of the mRNA after transcription. In humans, the average length of the 3' UTR is approximately 800 nucleotides, but can range from 60 to about 4000 nucleotides in eukaryotes. This long sequence in the 3' UTR reflects the complexity of regulation. The 3' UTRs contain binding sites for both microRNA (miRNA) and proteins. In general, longer 3' UTRs are associated with reduced gene expression since they contain more protein and miRNA binding sites that can inhibit translation. Overall these sites can lead to a highly complex regulation of gene expression and mRNA stability. Further, the 3' UTR contains sequences that promote the binding of proteins that associate the mRNA with the cytoskeleton, transport from the nucleus and other types of mRNA localization (Matoulikova *et al.*, 2012).

AU-Rich Elements and Stem Loops in mRNA

AU-rich elements (AREs) are 50–150 bases in length and usually contain multiple copies of the sequence AUUUA.

These elements are binding sites for proteins and these proteins, in response to different intracellular and extracellular signals, can promote mRNA decay, affect mRNA stability (Winzen *et al.*, 2004), or promote translation. It provides rapid means of cell growth regulation, differentiation, and adaptation. Transcripts encoding cytokines, proto-oncogenes, growth factors, and tumor suppressors are among the mRNA utilizing this type of mechanism (Winzen *et al.*, 2004).

Stem loops occur in the 3'UTR as well as the 5' UTR and can provide a scaffold for protein binding. IREs described earlier also occur in the 3' UTR of mRNAs that encode proteins involved in cellular iron metabolism. The mRNA transcript containing the IRE is either stabilized or degraded depending on the intracellular iron concentration which affects protein binding (Svoboda and Di Cara, 2006).

MicroRNA Response Elements

miRNA response elements are sequences often contained in the 3' UTR to which miRNAs bind. miRNAs are naturally occurring short, noncoding RNA molecules, usually between 18 and 25 nucleotides in length that bind to mRNA and regulate gene expression. This interaction is mediated by proteins in the miRNA-RNA Induced Silencing Complex (RISC). Several hundred miRNAs have been identified in mammals, although many of their mRNA targets have not been characterized. The miRNA interaction with miRNA response elements usually involves partial base pairing between the miRNA and the mRNA within the 3' UTR and results in translational repression (Macfarlane and Murphy, 2010).

The miRNA-RISC complex can also disrupt the closed loop between the cap and Poly(A) tail that forms during translation of the mRNA which exposes the mRNA to exonucleases and decapping enzymes resulting in mRNA decay. These interactions allow for differential gene expression at developmental stages and in various tissues.

Poly(A) Tail

Eukaryotic mRNAs have a sequence of 100–200 adenylic acid residues attached to the 3' UTR designated the poly(A) tail. This sequence is added posttranscriptionally and is essential for efficient translation and mRNA stability (Goss and Kleiman, 2013). The poly(A) tail is a binding site for poly(A) binding proteins (PABPs), which interact with other factors to affect translation, decay, and export of mRNA. PABPs interact with translation initiation factors bound to the 5' UTR cap to promote circularization of the mRNA which allows for efficient recycling of ribosomes during translation. Polyadenylation is regulated by sequences within the 3' UTR, including cytoplasmic polyadenylation elements (CPEs). These CPEs are uridine-rich sequences which bind other proteins to contribute to polyadenylation activation and repression.

mRNA and Disease

Mutations in eukaryotic mRNA have been identified with a number of diseases. These mutations can affect the open reading frame or coding sequence of the mRNA and result in

mutated proteins which can be defective in function (Liu *et al.*, 1999; Wen *et al.*, 2009; Wiestner *et al.*, 1998). In addition, both the 5' UTR and the 3' UTR provide regulatory elements for expression of the mRNA and mutations in these regions can lead to disease states by altering translation of the mRNA (Scheper *et al.*, 2007), either overexpression or repression of key proteins (Calvo *et al.*, 2009), changing the stability of the mRNA (Hollams *et al.*, 2002), or possibly altering the localization of the mRNA (Holt *et al.*, 2007). These mutations alter key features of the mRNA required for protein interactions.

There are a number of cases where mutations in the 5' UTR cause altered translation and lead to human disease. The most studied and best understood case is the dysregulation of ferritin translation (Goforth *et al.*, 2010). A mutation in the IRE element (described above) decreased binding of the IRP protein and increased ferritin translation leading to hereditary hyperferritinemia-cataract syndrome (Roetto *et al.*, 2002). The IRE element has been identified in the 5' UTRs of other mRNAs with disease implications. Amyloid- β precursor protein (APP) is implicated in Alzheimer disease and Down syndrome (Folin *et al.*, 2003). APP is cleaved into peptides which are the main components of neurotoxic amyloid plaques and overexpression of APP is due to increased translation.

Mutations in the 5' UTR IRES have also been associated with human disease (Chatterjee and Pal, 2009). Cellular IRES containing mRNA include a number of oncogenes, including members of the myc gene family, growth factors such as VEGF, and the anti-apoptotic protein Bag1. It is not surprising that deregulation of these IRES may affect the progression of cancer and other diseases. Many of the IRES contain mRNA function under stress conditions and this is particularly important for tumor progression where IRES translation of growth factors such as VEGF and FGF2 functions under hypoxic conditions. Tumor growth leads to the center of the tumor becoming hypoxic and increased vascularization is necessary for continued tumor growth. IRES-driven expression of growth factors is necessary for increased vascularization and resulting growth. Other point mutations in the 5' UTR have been associated with human disease: a mutation in the connexin-32 IRES causes Chalcot-Marie-Tooth disease, a neurodegenerative disease (Hudder and Werner, 2000), and a point mutation in the 5' UTR of BRCA1 has been identified in aggressive breast cancer (Signori *et al.*, 2001).

The 3' UTR contains a number of elements involved in gene expression and mutation of these elements can lead to disease states, usually through affecting protein binding. Dysregulation of ARE-binding proteins through mutations in the AU-rich regions can lead to tumorigenesis and hematopoietic malignancies. An expanded number of trinucleotide CTG repeats in the mRNA of dystrophin myotonia protein kinase causes myotonic dystrophy and other 3' UTR elements have been linked to alpha thalassemia, neuroblastoma, and congenital heart defects among other diseases (Halvorsen *et al.*, 2010).

Because UTR sequences often contain regulatory motifs that are common to members of the same metabolic family, they could provide an opportunity for development of highly specific RNA-based therapies to target a single or related genes (Wilkie *et al.*, 2003).

See also: Cell Division/Death: Regulation of Cell Growth:
Targeted mRNA Degradation

References

- Alifano, P., Bruni, C.B., Carlomagno, M.S., 1994. Control of mRNA processing and decay in prokaryotes. *Genetica* 94, 157–172.
- Baltimore, D., 1970. RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226, 1209–1211.
- Berg, J.M., Tymoczko, J.L., Stryer, L., 2002. *Biochemistry*. New York, NY: W. H. Freeman.
- Brenner, S., Jacob, F., Meselson, M., 1961. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* 190, 576–581.
- Calvo, S.E., Pagliarini, D.J., Mootha, V.K., 2009. Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proceedings of the National Academy of Sciences of the United States of America* 106, 7507–7512.
- Chatterjee, S., Pal, J.K., 2009. Role of 5′- and 3′-untranslated regions of mRNAs in human diseases. *Biology of the Cell* 101, 251–262.
- Clark, D.P., Pazdernik, N.J., 2013. *Academic Cell: Molecular Biology*. Waltham, MA: Academic Press, Elsevier.
- Folin, M., Baiguera, S., Conconi, M.T., et al., 2003. The impact of risk factors of Alzheimer's disease in the Down syndrome. *International Journal of Molecular Medicine* 11, 267–270.
- Gerlinger, P., Mohier, E., Le Meur, M.A., Hirth, L., 1977. Monocistronic translation of alfalfa mosaic virus RNAs. *Nucleic Acids Research* 4, 813–826.
- Goforth, J.B., Anderson, S.A., Nizzi, C.P., Eisenstein, R.S., 2010. Multiple determinants within iron-responsive elements dictate iron regulatory protein binding and regulatory hierarchy. *RNA* 16, 154–169.
- Goss, D.J., Kleiman, F.E., 2013. Poly(A) binding proteins: Are they all created equal? *Wiley Interdisciplinary Reviews: RNA* 4, 167–179.
- Guhaniyogi, J., Brewer, G., 2001. Regulation of mRNA stability in mammalian cells. *Gene* 265, 11–23.
- Halvorsen, M., Martin, J.S., Broadaway, S., Laederach, A., 2010. Disease-associated mutations that alter the RNA structural ensemble. *PLoS Genetics* 6, e1001074.
- Hellen, C.U., Sarnow, P., 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes and Development* 15, 1593–1612.
- Hiraga, S., Yanofsky, C., 1972. Normal repression in a deletion mutant lacking almost the entire operator-proximal gene of the tryptophan operon of *E. coli*. *Nature: New Biology* 237, 47–49.
- Hollams, E.M., Giles, K.M., Thomson, A.M., Leedman, P.J., 2002. mRNA stability and the control of gene expression: Implications for human disease. *Neurochemical Research* 27, 957–980.
- Holt, I., Mittal, S., Furling, D., et al., 2007. Defective mRNA in myotonic dystrophy accumulates at the periphery of nuclear splicing speckles. *Genes to Cells* 12, 1035–1048.
- Hudder, A., Werner, R., 2000. Analysis of a Charcot-Marie-Tooth disease mutation reveals an essential internal ribosome entry site element in the connexin-32 gene. *Journal of Biological Chemistry* 275, 34586–34591.
- Kozak, M., 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiological Reviews* 47, 1–45.
- Kozak, M., 1987. An analysis of 5′-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Research* 15, 8125–8148.
- Liu, L., Dilworth, D., Gao, L., et al., 1999. Mutation of the CDKN2A 5′ UTR creates an aberrant codon and predisposes to melanoma. *Nature Genetics* 21, 128–132.
- Lodish, H., Berk, A., Zipursky, S.L., et al., 2000. *Molecular Cell Biology*. New York, NY: W. H. Freeman.
- Londei, P., 2005. Evolution of translational initiation: New insights from the archaea. *FEMS Microbiology Reviews* 29, 185–200.
- Ma, J., Haldar, S., Khan, M.A., et al., 2012. Fe²⁺ binds iron responsive element-RNA, selectively changing protein-binding affinities and regulating mRNA repression and activation. *Proceedings of the National Academy of Sciences of the United States of America* 109, 8417–8422.
- Macfarlane, L.A., Murphy, P.R., 2010. MicroRNA: Biogenesis, function and role in cancer. *Current Genomics* 11, 537–561.
- Matoukova, E., Michalova, E., Vojtesek, B., Hrstka, R., 2012. The role of the 3′ untranslated region in post-transcriptional regulation of protein expression in mammalian cells. *RNA Biology* 9, 563–576.
- Meijer, H.A., Thomas, A.A., 2002. Control of eukaryotic protein synthesis by upstream open reading frames in the 5′-untranslated region of an mRNA. *Biochemical Journal* 367, 1–11.
- Roetto, A., Bosio, S., Gramaglia, E., et al., 2002. Pathogenesis of hyperferritinemia cataract syndrome. *Blood Cells, Molecules and Diseases* 29, 532–535.
- Scheper, G.C., van der Knaap, M.S., Proud, C.G., 2007. Translation matters: Protein synthesis defects in inherited diseases. *Nature Reviews Genetics* 8, 711–723.
- Shine, J., Dalgarno, L., 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America* 71, 1342–1346.
- Shine, J., Dalgarno, L., 1975. Determinant of cistron specificity in bacterial ribosomes. *Nature* 254, 34–38.
- Signori, E., Bagni, C., Papa, S., et al., 2001. A somatic mutation in the 5′UTR of BRCA1 gene in sporadic breast cancer causes down-modulation of translation efficiency. *Oncogene* 20, 4596–4600.
- Svoboda, P., Di Cara, A., 2006. Hairpin RNA: A secondary structure of primary importance. *Cellular and Molecular Life Sciences* 63, 901–908.
- Temin, H.M., Mizutani, S., 1970. RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 226, 1211–1213.
- Wen, Y., et al., 2009. Loss-of-function mutations of an inhibitory upstream ORF in the human hairless transcript cause Marie Unna hereditary hypotrichosis. *Nature Genetics* 41, 228–233.
- Wiestner, A., Schlemper, R.J., van der Maas, A.P., Skoda, R.C., 1998. An activating splice donor mutation in the thrombopoietin gene causes hereditary thrombocythaemia. *Nature Genetics* 18, 49–52.
- Wilkie, G.S., Diskson, K.S., Gray, N.K., 2003. Regulation of mRNA translation by 5′- and 3′-UTR-binding factors. *Trends in Biochemical Sciences* 28, 182–188.
- Winzen, R., Gowrishankar, G., Bollig, F., et al., 2004. Distinct domains of AU-rich elements exert different functions in mRNA destabilization and stabilization by p38 mitogen-activated protein kinase or HuR. *Molecular and Cellular Biology* 24, 4835–4847.

Further Reading

- Marchfelder, A., Hess, W., 2012. *Regulatory RNA in Prokaryotes*, 272 pp. (Google eBook).
- McManus, M.T., Sharp, P.A., 2003. Gene silencing in mammals by small interfering RNAs. *Nature Reviews Genetics* 3, 737–747.
- Piao, X., Zhang, X., Wu, L., Belasco, J.G., 2010. CCR4-NOT deadenylates mRNA associated with RNA-induced silencing complexes in human. *Molecular and Cellular Biology* 30 (6), 1486–1489.
- Qi, M.Y., Wang, Z.Z., Zhang, Z., et al., 2012. AU-rich-element-dependent translation repression requires the cooperation of tristetraprolin and RCK/P54. *Molecular and Cellular Biology* 32 (5), 913–928.
- Valencia-Sanchez, M.A., Liu, J., Hannon, G.J., Parker, R., 2006. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes & Development* 20, 515–524.

The Interplay between Eukaryotic mRNA Degradation and Translation

W Hu, Whitehead Institute for Biomedical Research, Cambridge, MA, USA

© 2016 Elsevier Inc. All rights reserved.

The destruction of mRNA is essential for proper regulation of gene expression. All mRNA must be destroyed to ensure that nuclear regulatory decisions are manifested as cytoplasmic events. In other words, ceasing mRNA transcription in response to environmental or physiological cues is irrelevant as long as cytoplasmic mRNA exists and is translated. From this, it can be said that mRNA decay contributes significantly to the proteomes' overall architecture by allowing cells to quickly adapt to environmental and physiological changes. Although much is known about mRNA decay in eukaryotes, understanding how the process is modulated is far from complete.

In this section, we first discuss our current understanding of eukaryotic mRNA decay. Then we put mRNA degradation in the context of mRNA translation, which is another important cytoplasmic event of gene expression, as these two events are known to be tightly integrated with each other for decades. In the last few years, many studies have concentrated on this relationship as a means to elucidate regulation of mRNA degradation. From this, two views have arisen. The first suggests that mRNA must be removed from ribosomes to be destroyed in ribosome-free areas, such as P-bodies. The other view proposes mRNA decay occurs in concert with protein synthesis. Here we summarize our current understanding of how eukaryotic mRNA decay interconnects with mRNA translation, and then analyze the data that have led to these two diametrically opposed models of where mRNA decay occurs within the cell. Finally, we attempt to reconcile these two views and suggest important areas for future investigation.

Overview of Eukaryotic mRNA turnover

Cytoplasmic mRNAs are destroyed by two major pathways. Both are initiated by the exonucleolytic digestion of the 3' poly (A) tail (a process termed deadenylation). Deadenylation is a key regulatory step in mRNA half-life determination because its rate is highly variable among transcripts. In some cases, deadenylation leads to mRNA destruction 3'–5' by the cytoplasmic exosome. In other cases, deadenylation causes the rapid removal of the 5' 7mGpppN cap and digestion of the transcript body 5'–3' (reviewed in [Coller and Parker, 2004](#)). Which pathway predominates is not clearly understood. Most likely transcript, tissue, and species specific differences dictate the predominate pathway employed. In this article, we focus exclusively on the deadenylation-dependent decapping pathway since it has been extensively characterized and its relationship to translation is more established than that of the cytoplasmic exosome-dependent pathway.

As mentioned, deadenylation is the first step of mRNA decay and is often thought to be rate limiting. This reaction is carried out by several enzymes (deadenylases) that digest the poly(A) tail in a 3'–5' direction. Generally, mRNA deadenylation is biphasic, the first step is initiated by the deadenylase dependent

poly(A) nuclease 2 (PAN2). Following this initial trimming event, the bulk of the tail is digested by the enzyme complex CCR4/POP2/NOT ([Yamashita et al., 2005](#)). Other deadenylases, however, exist including poly(A)-specific ribonuclease (PARN), Nocturin, and Angel proteins (reviewed in [Goldstrohm and Wickens, 2008](#)). It would appear, then, that deadenylation is facilitated by a seemingly redundant class of enzymes. Diversity in transcript stability might occur in part by the manner in which these enzymes achieve transcript specificity. For example, 3'UTR-binding proteins and miRNAs can enhance deadenylation by tethering a specific enzyme ([Giraldez et al., 2006](#)).

In yeast as well as in mammalian cells, deadenylation usually leads to destruction of the mRNA. In some systems, deadenylation can lead to translational storage of the mRNA. Although still unclear, it is thought that the poly(A) tail mediates its effect on mRNA translation and decay by way of an interaction between the poly(A)-binding protein (PAB) and the translation initiation factor eIF4G ([Coller et al., 1998](#); [Jacobson and Peltz, 1996](#)). Deadenylation serves to displace PAB from the mRNA thereby liberating its protective and/or translational stimulatory function.

In the cases where transcript destruction follows deadenylation, once the mRNA tail has been trimmed to a length no longer supporting PAB binding, the 7-methyl-guanosine cap structure is rapidly cleaved from the mRNA's 5' end. The decapping reaction requires the enzymatic activity of the DCP2 protein. Enzymatically, the pyrophosphatase activity of DCP2 cleaves the alpha-beta bond of the 5' cap liberating m⁷GDP and leaving a 5' monophosphate on the transcript. *In vitro*, DCP2 is sufficient for this activity ([Lykke-Andersen, 2002](#)). Although simple in chemistry *in vitro*, the *in vivo* orchestration of mRNA decapping requires a suite of proteins including DCP1, DHH1(RCK/p54), Hedls (only in metazoans), Pat1p (P100), Lsm1–7, Lsm12, Scd6 (Trailerhitch...), and Edc1–3 (reviewed in [Franks and Lykke-Andersen, 2008](#)). The precise function(s) of these factors is murky. Some decapping activators like LSM1–7 and Hedls may serve structural roles in promoting binding of DCP2 to the mRNA. Others like DHH1 and PAT1 are hypothesized to promote DCP2's accessibility to the cap by dissociating factors bound at the mRNA 5' end, such as eIF4F ([Coller and Parker, 2005](#); [Chu and Rana, 2006](#)). Irrespective of their true function, the complexity of this process indicates that mRNA decapping is tightly regulated. Control of decapping makes sense since this step of decay commits the mRNA to destruction and unlike deadenylation is largely irreversible although exceptions have been documented (reviewed in [Schoenberg and Maquat, 2009](#)).

Subsequent to decapping, the transcript is quickly and efficiently destroyed by the 5'–3' exonuclease, XRN1. This enzyme is highly processive and decay intermediates are hard to detect ([Stevens, 2001](#)). To date, no regulation of XRN1 activity has been detected, thus, it is generally thought that the 5'–3' exonucleolytic decay is not a rate-limiting step in turnover.

mRNA Degradation is Intimately Linked to Translation

All mRNA must and will succumb to decay. Transcript decay is, therefore, a default state. The spectrum of observed half-lives most likely represents acceleration or deceleration of this default rate. The translatability of an mRNA is a critical contributor to its overall stability (reviewed in [Franks and Lykke-Andersen, 2008](#); [Jacobson and Peltz, 1996](#)). Specifically, an inverse correlation has been established showing efficiently translated mRNAs have longer half-lives; while poorly translated mRNAs are unstable. These findings suggest that a major determinant of mRNA stability is the translation machinery. Curiously, however, modulation of each specific sub-step of translation, i.e., initiation, elongation, or termination, is perceived to have disparate effects on mRNA's half-life. It is important to note that this perception is based largely on the manner of analysis (see below).

The when and where of mRNA decapping and exonucleolytic decay has been the subject of some debate. Models for where mRNA decay takes place are in part based on the distinct manner in which translational inhibition affects mRNA decay. Experimental evidence has suggested two possibilities. On one side data seems to indicate that before decapping can occur, the translational machinery must be liberated and then the message is destroyed in specialized places within the cell. On the other side, there are data to suggest that mRNA never leave polysomes and are degraded co-translationally. In the following sections we highlight the work of many individuals whose observations have generated these two distinct views.

The Ying-Yang of Translational Initiation and Transcript Stability

Translational initiation is posited to be in direct competition with mRNA decapping (reviewed in [Franks and Lykke-Andersen, 2008](#)). From a priori view, this hypothesis makes sense because both translational initiation and mRNA decapping are dependent on factors binding on or near the 5' cap ([Cougot et al., 2004](#)). This theory is also supported by experimental data. In a simplified view, translation initiation consists of three major events: first recognition of the cap by eIF4F, which is composed of eIF4E, eIF4G, and eIF4A; second, recruitment of the 40 S ribosomal subunit; third, scanning of the 40 S ribosomal subunit to the first AUG (reviewed in [Kapp and Lorsch, 2004](#)). Altering the rate of any of these three events has a profound effect on mRNA half-life. For instance, mutations in eIF4F reducing its affinity for the cap result in rapid and efficient deadenylation and decapping ([Schwartz and Parker, 1999](#)). Similarly, inhibiting 40 S ribosomal subunit recruitment and scanning accelerates mRNA deadenylation and decapping. For example, inactivation of eIF3, thereby blocking 40 S recruitment ([Kapp and Lorsch, 2004](#)), results in rapid mRNA decay ([Heikkinen et al., 2003](#)). *Cis*-acting RNA structures that inhibit 40 S ribosomal subunit scanning also accelerate mRNA turnover ([Muhlrads et al., 1995](#); [Coller and Parker, 2005](#)). Lastly, the context of the AUG start codon influences mRNA stability. Specifically, nucleotide changes around the translation start codon predicted to decrease translational efficiency also dramatically destabilize reporter transcripts ([LaGrande and](#)

[Parker, 1999](#)). Taken together these correlations between translational initiation and mRNA decapping predict that decreasing ribosome occupancy on the 5'UTR allows the decay machinery to associate more efficiently with the transcript. Moreover, these data suggest that the translation initiation factors and decapping factors vie for the mRNA *in vivo*.

A direct competition between decapping and eIF4E has been documented *in vitro*. Purification of DCP1 from yeast extracts allows decapping to be monitored *in vitro* (presumable because purification of DCP1 brings along the decapping enzyme DCP2). Addition of recombinant eIF4E protein efficiently inhibits *in vitro* decapping mediated by the purified decapping enzyme ([Schwartz and Parker, 2000](#)). In conjunction with the aforementioned *in vivo* work, these data have led to the model that dissociation of eIF4E is required before mRNA decapping can occur. Importantly, this notion is supported by genetic analysis. Specifically, loss of eIF4E will restore the decapping activity when it is partially impaired genetically (i.e., DCP1^{Ts}; [Schwartz and Parker, 2000](#)). Collectively, these results indicate that the mRNA decapping enzyme (Dcp2/Dcp1) and eIF4E are probably in competition for the 5'-end cap structure on an mRNA. Clearly, mRNA decapping is stimulated by loss of eIF4E function, but it is also important to note that *in vivo* inhibition of translational initiation in any manner also greatly enhances deadenylation rate. The effects of translation rate on deadenylation have remained unexplored.

Undoubtedly there is a positive correlation between the efficiency of translational initiation and mRNA stability. The displacement of eIF4E from the cap, therefore, is proposed to be a necessary first step in mRNA decapping (reviewed in [Coller and Parker, 2004](#)). At this point, it is important to note that the inverse correlation between translation rate and mRNA decapping does not hold true under every physiological condition. For example, cellular stress can have profound effects on translational initiation and mRNA decay. Under these extreme conditions the predictable correlation between translation and decay does not occur. Specifically, translational initiation is inhibited when cells are deprived of glucose ([Ashe et al., 2000](#)). Despite this, mRNAs are dramatically stabilized rather than destabilized ([Hilgers et al., 2006](#)). On the other hand, amino acid deprivation also powerfully inhibits translation initiation (ref) but mRNAs are greatly destabilized under this condition. Lastly, the unfolded protein response also triggers repression of translation ([Scheuner et al., 2001](#)), but this stress does not affect mRNA stability at all ([Hilgers et al., 2006](#)). One possible explanation for these discrepancies is that each distinct stress alters one or more of the mRNA decay factors. Alternatively, each stress might change mRNP dynamics and therefore the manner by which it is perceived by the decay machinery. It seems likely, therefore, that a more detailed understanding of how cellular stress impacts mRNA stability will shed more light on the complex interplay between decay and translation.

Translation Elongation and mRNA Turnover

It is postulated that once an mRNA is engaged in elongation, ribosomes provide a protective quality that insulates the

transcript from decay (Jacobson and Peltz, 1996). This theory, however, is controversial. Unlike translational initiation where slowing the process enhances decay, the manner in which elongation is perturbed dramatically influences the result obtained. For example, treating cells with the translation elongation inhibitor, cycloheximide (a drug that freezes ribosomes on mRNA), results in a significant inhibition of decay in budding yeast and high eukaryotes (Beelman and Parker, 1994; Ross, 1995). Similar results are obtained when elongation is stopped using genetic mutation. Specifically, depleting charged tRNAs by inhibiting maturation of their 3' end will dramatically stabilize mRNA (Peltz *et al.*, 1992). In both cases, blocking elongation is suggested to inhibit decapping (Beelman and Parker, 1994), deadenylation still occurs, albeit at a slightly reduced rate (Hilgers *et al.*, 2006). It is important to point out, that both of these procedures do not kinetically slow elongation, rather, they freeze ribosomes on mRNA and block elongation altogether. Paradoxically, slowing the rate of translation elongation is similar to slowing translational initiation in that it destabilizes mRNAs. For example, rare codons are recognized by tRNAs with low concentrations within the cell, thus, mRNAs with rare codons have a relatively slow translation rate compared to the mRNAs with cognate normal codons. The presence of rare codons within an mRNA saturates the transcript with ribosomes and dramatically shortens half-life (Caponigro *et al.*, 1993; Hu *et al.*, 2009). Slowing ribosome translocation in this manner, therefore, destabilizes mRNAs. One possible explanation for these differences is that freezing mRNA on ribosomes may 'lock' the transcript in an RNP conformation that is resistant to mRNA decay factors. Conversely slowing elongation enhances decay by a yet unknown mechanism.

Translation Termination and mRNA Stability

Eukaryotic mRNA translation termination requires two release factors, eRF1 and eRF3. Translation termination process can influence mRNA half-life. Specifically, it was observed that the N-terminal domain of eRF3, which is not required for translation termination, can interact with Pab1, and this interaction is involved in modulating mRNA stability (Hosoda *et al.*, 2003). Disruption of this interaction results in translation-dependent stabilization of mRNA caused by decreased deadenylation rate (Hosoda *et al.*, 2003). Interestingly, it was further found that certain deadenylase complexes can also bind to the same site on Pab1 that is involved in the interaction with eRF3 (Funakoshi *et al.*, 2007). Thus, it has been postulated that eRF3 can regulate mRNA deadenylation by competitively binding to the Pab1, which then modulates the recruitment and activation of deadenylase complexes (Funakoshi *et al.*, 2007). In addition to the release factors, other proteins that can modulate translation termination can also influence mRNA stability. For example, a recent characterized protein named Tpa1 can interact with the two release factors and regulate the readthrough of stop codons (Keeling *et al.*, 2006). Interestingly, although the detailed mechanisms still remain elusive, knocking out this protein can have decreased deadenylation rate and increased mRNA stability (Keeling *et al.*, 2006). Collectively, these results suggest that

mRNA translation termination can result in mRNP conformational changes that can influence mRNA stability, likely via modulating mRNA deadenylation.

Where Does Eukaryotic mRNA Degradation Occur Within the Cell?

In recent years, understanding where mRNA degradation occurs within the cell has been given a great deal of scientific investigation. The thought is that insight into how the large repertoire of observed mRNA half-lives is achieved may come from knowing when and where degradation occurs within the cytoplasm. At present, there are two different views on where mRNA degradation takes place. The first comes from the discovery that mRNA decapping and decay factors aggregate into punctuate, microscopically visible structures. These structures have been given the epitaph of processing bodies or P-bodies for short. Importantly, P-bodies do not contain ribosomes nor do they contain most translational initiation factors. What they do contain is the full complement of decapping proteins, certain mRNA decay intermediates, and in higher eukaryotes, the miRNA machinery (reviewed in Franks and Lykke-Andersen, 2008; Parker and Sheth, 2007). In conjunction with the aforementioned findings that mRNA translation is generally positively correlated with half-life, these data have led to a popular 'two-step' model for regulating transcript stability in which ribosome dissociation first occurs, then the mRNA is trafficked to P-bodies where mRNA decapping ensues (reviewed in Franks and Lykke-Andersen, 2008; Parker and Sheth, 2007). We refer to this model as the 'mRNA cycle hypothesis,' because it is postulated that mRNA can cycle in and out of polyribosomes and into a quiescent state that can be either stored/reutilized or degraded (Brenques *et al.*, 2005). A second model supported in the literature, is that mRNA decay does not involve a fundamental transition between a polyribosome-bound state and a translationally repressed state, but rather occurs co-translationally (Hu *et al.*, 2009). This hypothesis dates back almost 20 years (Beelman and Parker, 1994; Mangus and Jacobson, 1999), but has lacked strong experimental evidence until recently. Nonetheless, both models are supported but have their limitations. Here, we will examine these two models in detail and at the end attempt to provide a unifying theory.

The mRNA Cycle Hypothesis – From Polysome to p-Bodies (and Perhaps Back)

The theory that translation and decay are partitioned events is based in part on the aforementioned distinct effects translation initiation versus elongation have on mRNA half-life. The mRNA cycle hypothesis suggests that mRNA is translated and then at some point becomes translationally silenced and moved into a state of quiescence. Once this inactive state is achieved, the mRNA can either be destroyed or perhaps reutilized in response to certain cues. In theory, translational repression of mRNA makes sense from a logistical stand point. The 5'-cap structure is occupied by a suite of proteins that promote 40 S ribosome joining. At some

point, DCP2 must have access to the cap and 5'UTR, and thus it seems logical that the 5' end is remodeled and initiation factors displaced (Steiger *et al.*, 2003). It is predicted that the downstream consequence of this remodeling must be translational repression, as manifest by ribosome dissociation.

The RNA cycle hypothesis is rooted in studies of translational initiation and its effect on decay. As discussed in the previous section, Schwartz and Parker (1999) found that mutations in eIF4F result in mRNA destabilization. Cap binding by eIF4F, therefore was proposed to be in competition with mRNA decapping. Similar results were seen when other aspects of initiation were compromised, including 40 S joining and AUG recognition. Moreover, *in vitro* experiments demonstrated that addition of purified eIF-4E could reduce decapping in cell extracts (Schwartz and Parker, 2000). The combined findings were interpreted as a direct competition between eIF4E and DCP2 exists. A corollary of this hypothesis is that translational repression, defined as ribosome dissociation, is required before mRNA decapping.

The mRNA cycle model is also supported by the discovery that several decapping regulators influence mRNA translation. Specifically, DHH1 and PAT1 were shown to be bona fide translational repressors in yeast and humans (Coller and Parker, 2005; Chu and Rana, 2006). Consistent with the notion, DHH1 is homologous to factors implicated in maternal mRNA storage (reviewed in Rajyaguru and Parker, 2009) thus decapping was hypothesized to be similar to posttranscriptional events that occur during early development (Coller *et al.*, 2001). The nature of how mRNA decapping regulators influence mRNA translation is still unknown, but has been proposed to inhibit translation initiation (Coller and Parker, 2005).

Other aspects of mRNA decay are also thought to result in translational silencing, and possible ribosome dissociation. Specifically, mRNA deadenylation can lead to translational repression (Huarte *et al.*, 1992). This is especially well documented during development. The poly(A) tail and its binding protein, PAB, are hypothesized to interact with the 5' cap via a protein interaction with eIF4G. Poly(A) is popularly believed to stimulate translation by way of this 'close-loop' (Jacobson and Peltz, 1996). Collectively, the first step of mRNA decay, deadenylation, would be predicted to result in a loss of translational efficiency. It is important to note, however, that controversy exists whether deadenylation is a cause or effect of translational repression in some context.

P-bodies, above all things, bolstered the notion that ribosome dissociation is required for mRNA decapping and that these events are similar to storage events that occur in other contexts (reviewed in Parker and Sheth, 2007; Rajyaguru and Parker, 2009). The discovery of P-bodies dates back to 1997 when Bashkirov *et al.* (1997) cloned the mouse homolog of the known 5'-3' exonuclease Xrn1. Using mouse E10 fibroblast cells, it was observed that mXrn1 localized to punctuate cytoplasmic structures. At this time, the authors proposed that these structures may represent either sites for RNA turnover or sites in which the enzyme is stored until used. P-bodies were rediscovered in 2002 when Eystathioy *et al.* (2002) found that the human auto-antigen protein, GW182, was also found in cytoplasmic granules. Although it was not clear at this time,

GW182 is a critical factor in mediating miRNA translational control (Tritschler *et al.*, 2010). One year later, a flurry of studies simultaneously demonstrated that, in the addition to XRN1, the major decapping factors also are found in cytoplasmic foci (many refs). Sheth and Parker (2003) extended this analysis by showing that mRNA decay intermediates co-localized with these granules in a manner that is dependent on the enzymatic activity of yeast XRN1. Similar results were seen in humans (Cougot *et al.*, 2004). Thus the aggregation of decay factors in granules was suggested to be of functional significance and that these foci represented the place where mRNAs were degraded.

P-bodies are devoid of ribosomes and other vital translational initiation factors (Sheth and Parker, 2003; Cougot *et al.*, 2004). Coupled with the aforementioned results indicating decapping is in competition with translational initiation, it was further proposed that mRNA must be removed from ribosomes and trafficked to a P-body before decay ensues (reviewed in Franks and Lykke-Andersen, 2008). Indeed the very term P-bodies, or processing bodies was coined, in part, to draw similarity to P-granules or polar bodies, known sites of mRNA storage in germ-lines, oocytes, and embryos (reviewed in Parker and Sheth, 2007; Rajyaguru and Parker, 2009). Consistent with this, P-bodies have also been suggested to contain translational quiescent mRNA (Bregues *et al.*, 2005; Parker and Sheth, 2007; Franks and Lykke-Andersen, 2008). In mammalian systems, the miRNA machinery co-localizes to P-bodies further supporting a model in which P-bodies are sites of mRNA storage. In yeast, stress conditions like glucose deprivation result in global shut-down of mRNA translation and mRNA decay. This correlates with an increase in P-body size and abundance. An increase in P-body size has been interpreted as an influx of mRNA into these structures (Bregues *et al.*, 2005; Sheth and Parker, 2003). Taken together, mRNA decapping has hypothesized to require removal of mRNA from the translational apparatus followed by packaging into P-bodies where either decay or storage ensues.

The RNA cycle model is very provocative, but there are still many questions left to be answered. For instance, one important feature of this hypothesis is that DCP2 and eIF-4E are in direct competition with each other for cap access. While this competition is clear *in vitro*, *in vivo* the relationship between DCP2 and eIF4E is more complex. Specifically, if a simple competitive relationship existed, then loss of eIF4E function would result in promiscuous mRNA decapping independent of deadenylation. Although perturbations in initiation result in accelerated decapping, in all cases when initiation is impaired (i.e., initiation factor mutants, AUG context changes, and *cis*-acting initiation blocks) decapping is not uncoupled from deadenylation. Indeed, loss of deadenylation by *ccr4* mutation stabilizes non-translating mRNAs. Rather, loss of eIF-4F function results in rapid deadenylation followed by rapid decapping (Schwartz and Parker, 1999). These data indicate that loss of eIF-4E is not sufficient to elicit decapping directly, rather dissolution of the initiation complex sensitizes the mRNA toward decay in a fashion that is still unknown. In this light, the notion that eIF-4E displacement is a critical event in decapping stimulated by deadenylation requires reinvestigation.

The RNA cycle hypothesis is also based on the notion that removing mRNAs from ribosomes accelerates their decay. Based on this, polysomes would provide a protective quality to the transcript, for example, sequestering the message from P-bodies. While generally, impeding translational initiation promotes rapid mRNA decay, important caveats exist. For instance, there is no correlation between changes of translation status and mRNA stability under conditions of stress known to impede protein synthesis (see above). Moreover, little evidence exists that ribosomes protect the mRNA from destruction. In fact, slowing translation elongation by use of rare codons saturates the message with ribosomes, but, paradoxically, greatly enhances its decay by the normal deadenylation-dependent decapping machinery. It appears then that an mRNA that never has ribosomes is degraded quickly, and an mRNA that is saturated with slowly moving ribosomes is also degraded quickly. Thus, there is no obvious need for the mRNA to dissociate from the translational machinery in order to be degraded efficiently (see below).

Lastly, an important aspect of the RNA cycle hypothesis is that once mRNA leaves ribosomes they enter P-bodies. From a quantitative point of view, it is still unclear to what extent mRNA decapping and 5'–3' exonucleolytic decay occur in ribosome-free areas in the cell. Moreover, it is unclear how much of the decapping complex assembles into P-bodies versus the soluble cytosol. In some cases, it has been suggested to be only a minor fraction. For instance quantification of the P-body protein Ago2 in mammalian cells revealed that less than 5% of total Ago2 is localized in P-bodies (Leung *et al.*, 2006). Second, P-bodies, whose abundance correlates well with events that enhance translation repression and mRNA decay, are not required for either of these events. Specifically, in both yeast and higher eukaryotes, several groups have shown that P-body formation can be uncoupled from mRNA turnover and translation repression (Sweet *et al.*, 2007; Decker *et al.*, 2007; Eulalio *et al.*, 2007). Together, these data demonstrate that the function of P-bodies remains enigmatic. Clearly, however, the aggregation of decapping factors into P-bodies is not required for normal rates of mRNA turnover.

In summary, the RNA cycle hypothesis is provocative. It provides a simple explanation for how mRNA turnover and translation are interconnected. Importantly, however, significant caveats exist suggesting that mRNA turnover is not as simplistic as the model suggest.

Parting on Polysomes

A second model that has emerged in the literature is that message decay takes place co-translationally. In other words, mRNA decapping and 5'–3' exonucleolytic digestion occur primarily on polyribosomes and not in a ribosome-free state (e.g., a P-body). The hypothesis that mRNAs are destroyed on polysomes is not new. In most cases, however, this scenario has only been observed in special cases; specifically albumin and β -tubulin mRNA (Pastori *et al.*, 1991; Theodorakis and Cleveland, 1992). Nonetheless, small clues are seen throughout the literature. Early evidence was first seen yeast. As mentioned in previous sections, treating yeast cells with the translational elongation inhibitor cycloheximide leads to the

dramatic stabilization of mRNAs. Recent interpretations of this effect are that the mRNA is sequestered in polyribosomes under these conditions and away from P-bodies (reviewed in Parker and Sheth, 2007; Franks and Lykke-Andersen, 2008). Interestingly, however, Beelman and Parker (1994) investigated mRNA decay in cells treated with cycloheximide and found the mRNA accumulated as a slightly shorter species over time. Although this finding wasn't further characterized, they proposed the truncated mRNA was the result of decapping and digestion up to a ribosome stalled at the AUG. Consistent with the hypothesis of co-translational decapping/decay, it was also found that in humans decapping activity co-sediments with polysomes (Wang *et al.*, 2002a,b). Mangus and Jacobson (1999) observed that decay intermediates similar to those observed in P-bodies were also associated with polyribosomes. Lastly, we provided additional supportive by showing that the majority of decapping is observed while the transcript is saturated with ribosomes (Hu *et al.*, 2009). Collectively, these data argue that under normal physiological conditions, mRNA decapping and 5'–3' exonucleolytic digestion occur on polyribosomes. A corollary of this hypothesis is that a transition into a translational repressed state, a predicted by the RNA cycle hypothesis, is not required for efficient mRNA decay.

Like the RNA cycle hypothesis, the model of co-translational decay also has important caveats to mention. First, numerous reports have shown that events like deadenylation inhibit protein synthesis (Huarte *et al.*, 1992; Thompson *et al.*, 2000; Beilharz *et al.*, 2009), yet no apparent difference in ribosome association can be seen for mRNA before and after deadenylation in our analysis. It is important to note here that polysome analysis provides direct physical evidence for the binding of ribosomes to an mRNA, it does not provide any information about translation rate. For example, it has been documented that miRNA targets are saturated with ribosomes, yet no protein output is detectable (Maroney *et al.*, 2006; Nottrott *et al.*, 2006). It is unclear how this is possible, but most likely indicates that there are enigmatic mechanisms of translational control that have yet to be revealed. Indeed in our own analysis we cannot conclude if nonadenylated mRNAs are translating at the same rate as fully adenylated mRNA. A clear and important area of future investigation is to correlate ribosome occupancy with adenylation status and protein output. Perhaps this analysis will shed light on this unexpected finding.

Second, it is clear that the decapping regulators DHH1 and PAT1 inhibit protein synthesis most likely by altering translational initiation (Coller and Parker, 2005) and that mutations in initiation factors can accelerate decapping. In this regard, it is important to realize that polysome analysis is a gross measurement of ribosome-associated mRNPs. Subtle changes such as rearrangement of the mRNA's 5' UTR are beyond the detection limit of this assay, unless they result in dramatic changes in polysome association.

Lastly, the contribution of initiation factors to polyribosome association/maintenance is still unclear. In recent years it has been observed that loss of initiation factor function *in vivo* has only a mild effect on polyribosome formation. Thus, it is formally possible that 5'UTR remodeling can occur with little or no appreciable change in ribosome occupancy. The explanation for this paradox remains unclear, but might indicate that initiation factors are only rate limiting for

translation under certain circumstances or during early events in protein synthesis before the establishment of polyribosomes.

What Does It All Mean?

The interplay between translation and message decay is of great importance for understanding the overall regulation of gene expression. As discussed in the previous two sections, there are two points of view. One is that most mRNA decay occurs after ribosome dissociation and perhaps in P-bodies. The other suggests that mRNA never leave ribosomes and decay is co-translational. Both hypothesizes are based on clear experimental observations. Neither model is perfect and both have important caveats for consideration. In the end, the truth most likely lies somewhere in between.

Biochemical data indicate that polyribosomes represent a major site of mRNA decapping within the cell under normal growth conditions. Importantly, however, evidence also exists to suggest that decay occurs, at some level, in P-bodies. We propose that all mRNA degradation is initiated on polyribosomes at the step of deadenylation. Where subsequent transcript degradation occurs is a function of the messages' translation rate versus relative to its decapping rate. We hypothesize that deadenylation does indeed reduce translational efficiency (Huarte *et al.*, 1992) perhaps through loss of the PAB, Pab1p (Jacobson and Peltz, 1996), or association of decapping regulators such as DHH1 and PAT1 (Coller and Parker, 2005). These events allow for the remodeling of the 5'UTR and deposition of a decapping complex. Under normal conditions, where mRNA translation is robust, mRNA decapping will occur on polyribosomes because it is kinetically faster than ribosomal run-off. We envision that decay in P-bodies is restricted to certain circumstances, such as under stress (when global translation is altered) or for mRNA populations in which ribosomal run-off is kinetically faster than mRNA decapping. Under conditions in which mRNA translational initiation or decapping are rate limiting (such as during stress), ribosomal run-off would predominate even in the absence of mRNA deadenylation and decay would be predicted to occur on ribosome-free mRNAs which may assemble into cytoplasmic P-bodies. Under normal conditions, however, we propose that remodeling of the 5' UTR and deposition of the decapping complex is faster than the rate at which the mRNA can be cleared of ribosomes. In this regard, the mRNA is degraded co-translationally.

Perspectives

Despite all that we have learned about mRNA decay in the past 20 years, there are still many mysteries. For instance, we still do not really understand how mRNA half-lives are determined for individual mRNAs. We are aware of elements that promote stability or instability, but by and large, it is unclear how these mRNA features really work. Perhaps central to this line of study is determining exactly how mRNA deadenylation rate is regulated. Deadenylation initiates mRNA decay and it clearly occurs at a message-specific rate, thereby setting the transcript's

overall half-life. Despite its importance, the control of mRNA deadenylation rate is still largely unexplored. We do know that mRNA deadenylation occurs on polyribosomes (Hilgers *et al.*, 2006). As mentioned previously, altering the translation rate of a message greatly enhances deadenylation. It would be of great interest to explore how mRNA translation modulates mRNA deadenylation rate.

Regarding decapping, this step of decay commits the message to be destroyed. Although a large collection of information on the decapping reaction has been gathered, such as characterization of the decapping enzyme, identification of global and message-specific decapping activators, and the potential interplay between the decapping enzyme and translation initiation factors, several outstanding and fundamental questions remain. For instance, the decapping complex is composed of more than 10 proteins in addition to the Dcp2/Dcp1 but we do not even have the simplest understanding of the temporal and special relationship of these factors during to each other and during the process of decay. Moreover, it has been documented that the same mRNA can have different stabilities under different environmental conditions (e.g., normal physiological conditions vs. stress conditions), thus determining how the decapping activity is regulated under different conditions will further provide insights into the regulation of mRNA decapping.

Third, what is the role of P-bodies? P-bodies are evolutionarily conserved ribosome-free cellular foci with mRNA decay factors and certain mRNA decay intermediates. Interestingly, there is a reverse correlation between mRNA decay status and the formation and sizes of P-bodies under certain conditions (Franks and Lykke-Andersen, 2008; Parker and Sheth, 2007). It has been demonstrated by many different labs that P-bodies are not required for mRNA decay nor for translational repression (Chu and Rana, 2006; Decker *et al.*, 2007; Eulalio *et al.*, 2007; Sweet *et al.*, 2007). The contribution, therefore, of P-bodies to mRNA metabolism is yet to be understood but of clear importance.

Lastly, how is mRNA half-life determined in the context of on-going translation? Clearly subtle changes in messages translatability can lead to dramatic changes in the messages' stability. For instance, dramatic disruption of translation events can elicit quality control pathways like nonsense-mediated mRNA decay (NMD), non-stop decay, or No-Go decay. It is important also to think about how subtle changes alter the way the normal pathway responds to the message. For instance, slowing initiation accelerates decay. Slowing elongation accelerates decay. Indeed, it appears that if translation is aberrant at any step, mRNA decay accelerates. Perhaps is it most appropriate to think of the normal degradation machinery as a monitor for the quality of protein synthesis events; its function is to respond when translation goes awry. mRNAs are constantly associated with proteins, from its birth to death (reviewed in Moore, 2005). As we learn more, it seems that this mRNP has to be just right in order to facilitate mRNA metabolism events, like splicing, transport, and translation. It is tempting to speculate that changes to the mRNP occur during mRNA translation and this signals that the message is no longer translating efficiently and it is time for it to be cleared from the polyribosome pool. Perhaps the key event is deadenylation itself, which is a default reaction that

dramatically impacts the mRNP structure and translatability of the mRNA. As the tail is shortened, translation rate changes and this is 'sensed' as a translational aberrancy. Once this occurs the messages is cleared from the cytoplasm by other decay factors. In this light, perhaps the normal degradation machinery becomes a critical aspect of monitoring the quality of gene expression as well as a regulator of its overall levels. Time will tell.

References

- Ashe, M.P., De Long, S.K., Sachs, A.B., 2000. Glucose depletion rapidly inhibits translation initiation in yeast. *Molecular Biology of the Cell* 11, 833–848.
- Bashkurov, V.I., Scherthan, H., Solinger, J.A., Buerstedde, J.M., Heyer, W.D., 1997. A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. *Journal of Cell Biology* 136, 761–773.
- Beelman, C.A., Parker, R., 1994. Differential effects of translational inhibition in cis and in trans on the decay of the unstable yeast MFA2 mRNA. *Journal of Biological Chemistry* 269, 9687–9692.
- Beilharz, T.H., Humphreys, D.T., Clancy, J.L., et al., 2009. microRNA-mediated messenger RNA deadenylation contributes to translational repression in mammalian cells. *PLoS One* 4, e6783.
- Bregues, M., Teixeira, D., Parker, R., 2005. Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. *Science* 310, 486–489.
- Caponigro, G., Muhrad, D., Parker, R., 1993. A small segment of the MAT alpha 1 transcript promotes mRNA decay in *Saccharomyces cerevisiae*: A stimulatory role for rare codons. *Molecular and Cellular Biology* 13, 5141–5148.
- Chu, C.Y., Rana, T.M., 2006. Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biology* 4, e210.
- Collier, J., Parker, R., 2004. Eukaryotic mRNA decapping. *Annual Review of Biochemistry* 73, 861–890.
- Collier, J., Parker, R., 2005. General translational repression by activators of mRNA decapping. *Cell* 122, 875–886.
- Collier, J.M., Gray, N.K., Wickens, M.P., 1998. mRNA stabilization by poly(A) binding protein is independent of poly(A) and requires translation. *Genes & Development* 12, 3226–3235.
- Collier, J.M., Tucker, M., Sheth, U., Valencia-Sanchez, M.A., Parker, R., 2001. The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. *RNA* 7, 1717–1727.
- Cougot, N., Babajko, S., Seraphin, B., 2004. Cytoplasmic foci are sites of mRNA decay in human cells. *Journal of Cell Biology* 165, 31–40.
- Decker, C.J., Teixeira, D., Parker, R., 2007. Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in *Saccharomyces cerevisiae*. *Journal of Cell Biology* 179, 437–449.
- Eulalio, A., Behm-Ansmant, I., Schweizer, D., Izaurralde, E., 2007. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Molecular and Cellular Biology* 27, 3970–3981.
- Eystathiou, T., Chan, E.K., Tenenbaum, S.A., et al., 2002. A phosphorylated cytoplasmic autoantigen, GW182, associates with a unique population of human mRNAs within novel cytoplasmic speckles. *Molecular Biology of the Cell* 13, 1338–1351.
- Franks, T.M., Lykke-Andersen, J., 2008. The control of mRNA decapping and P-body formation. *Molecular Cell* 32, 605–615.
- Funakoshi, Y., Doi, Y., Hosoda, N., et al., 2007. Mechanism of mRNA deadenylation: Evidence for a molecular interplay between translation termination factor eRF3 and mRNA deadenylases. *Genes & Development* 21, 3135–3148.
- Giraldez, A.J., Mishima, Y., Rihel, J., et al., 2006. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* 312, 75–79.
- Goldstrohm, A.C., Wickens, M., 2008. Multifunctional deadenylase complexes diversify mRNA control. *Nature Reviews* 9, 337–344.
- Heikkinen, H.L., Llewellyn, S.A., Barnes, C.A., 2003. Initiation-mediated mRNA decay in yeast affects heat-shock mRNAs, and works through decapping and 5'-to-3' hydrolysis. *Nucleic Acids Research* 31, 4006–4016.
- Hilgers, V., Teixeira, D., Parker, R., 2006. Translation-independent inhibition of mRNA deadenylation during stress in *Saccharomyces cerevisiae*. *RNA* 12, 1835–1845.
- Hosoda, N., Kobayashi, T., Uchida, N., et al., 2003. Translation termination factor eRF3 mediates mRNA decay through the regulation of deadenylation. *Journal of Biological Chemistry* 278, 38287–38291.
- Hu, W., Sweet, T.J., Chamnongpol, S., Baker, K.E., Collier, J., 2009. Co-translational mRNA decay in *Saccharomyces cerevisiae*. *Nature* 461, 225–229.
- Huarte, J., Stutz, A., O'Connell, M.L., et al., 1992. Transient translational silencing by reversible mRNA deadenylation. *Cell* 69, 1021–1030.
- Jacobson, A., Peltz, S.W., 1996. Interrelationships of the pathways of mRNA decay and translation in eukaryotic cells. *Annual Review of Biochemistry* 65, 693–739.
- Kapp, L.D., Lorsch, J.R., 2004. The molecular mechanics of eukaryotic translation. *Annual Review of Biochemistry* 73, 657–704.
- Keeling, K.M., Salas-Marco, J., Oshovich, L.Z., Bedwell, D.M., 2006. Tpa1p is part of an mRNP complex that influences translation termination, mRNA deadenylation, and mRNA turnover in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 26, 5237–5248.
- LaGrandeur, T., Parker, R., 1999. The cis acting sequences responsible for the differential decay of the unstable MFA2 and stable PGK1 transcripts in yeast include the context of the translational start codon. *RNA* 5, 420–433.
- Leung, A.K., Calabrese, J.M., Sharp, P.A., 2006. Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. *Proceedings of the National Academy of Sciences of the United States of America* 103, 18125–18130.
- Lykke-Andersen, J., 2002. Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. *Molecular and Cellular Biology* 22, 8114–8121.
- Mangus, D.A., Jacobson, A., 1999. Linking mRNA turnover and translation: Assessing the polyribosomal association of mRNA decay factors and degradative intermediates. *Methods* 17, 28–37.
- Maroney, P.A., Yu, Y., Fisher, J., Nilsen, T.W., 2006. Evidence that microRNAs are associated with translating messenger RNAs in human cells. *Nature Structural & Molecular Biology* 13, 1102–1107.
- Moore, M.J., 2005. From birth to death: The complex lives of eukaryotic mRNAs. *Science* 309, 1514–1518.
- Muhrad, D., Decker, C.J., Parker, R., 1995. Turnover mechanisms of the stable yeast PGK1 mRNA. *Molecular and Cellular Biology* 15, 2145–2156.
- Nottrott, S., Simard, M.J., Richter, J.D., 2006. Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nature Structural & Molecular Biology* 13, 1108–1114.
- Parker, R., Sheth, U., 2007. P bodies and the control of mRNA translation and degradation. *Molecular Cell* 25, 635–646.
- Pastori, R.L., Moskaitis, J.E., Buzek, S.W., Schoenberg, D.R., 1991. Coordinate estrogen-regulated instability of serum protein-coding messenger RNAs in *Xenopus laevis*. *Molecular Endocrinology* 5, 461–468.
- Peltz, S.W., Donahue, J.L., Jacobson, A., 1992. A mutation in the tRNA nucleotidyltransferase gene promotes stabilization of mRNAs in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 12, 5778–5784.
- Rajyaguru, P., Parker, R., 2009. CGH-1 and the control of maternal mRNAs. *Trends in Cell Biology* 19, 24–28.
- Ross, J., 1995. mRNA stability in mammalian cells. *Microbiological Reviews* 59, 423–450.
- Scheuner, D., Song, B., McEwen, E., et al., 2001. Translational control is required for the unfolded protein response and *in vivo* glucose homeostasis. *Molecular Cell* 7, 1165–1176.
- Schoenberg, D.R., Maquat, L.E., 2009. Re-capping the message. *Trends in Biochemical Sciences* 34, 435–442.
- Schwartz, D.C., Parker, R., 1999. Mutations in translation initiation factors lead to increased rates of deadenylation and decapping of mRNAs in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 19, 5247–5256.
- Schwartz, D.C., Parker, R., 2000. mRNA decapping in yeast requires dissociation of the cap binding protein, eukaryotic translation initiation factor 4E. *Molecular and Cellular Biology* 20, 7933–7942.
- Sheth, U., Parker, R., 2003. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* 300, 805–808.
- Steiger, M., Carr-Schmid, A., Schwartz, D.C., Kiledjian, M., Parker, R., 2003. Analysis of recombinant yeast decapping enzyme. *RNA* 9, 231–238.
- Stevens, A., 2001. 5'-Exoribonuclease 1: Xrn1. *Methods in Enzymology* 342, 251–259.
- Sweet, T.J., Boyer, B., Hu, W., Baker, K.E., Collier, J., 2007. Microtubule disruption stimulates P-body formation. *RNA* 13, 493–502.
- Theodorakis, N.G., Cleveland, D.W., 1992. Physical evidence for cotranslational regulation of beta-tubulin mRNA degradation. *Molecular and Cellular Biology* 12, 791–799.
- Thompson, S.R., Goodwin, E.B., Wickens, M., 2000. Rapid deadenylation and Poly (A)-dependent translational repression mediated by the *Caenorhabditis elegans* tra-2 3' untranslated region in *Xenopus* embryos. *Molecular and Cellular Biology* 20, 2129–2137.

- Tritschler, F., Huntzinger, E., Izaurralde, E., 2010. Role of GW182 proteins and PABPC1 in the miRNA pathway: A sense of déjà vu. *Nature Reviews* 11, 379–384.
- Wang, Y., Liu, C.L., Storey, J.D., *et al.*, 2002a. Precision and functional specificity in mRNA decay. *Proceedings of the National Academy of Sciences of the United States of America* 99, 5860–5865.
- Wang, Z., Jiao, X., Carr-Schmid, A., Kiledjian, M., 2002b. The hDcp2 protein is a mammalian mRNA decapping enzyme. *Proceedings of the National Academy of Sciences of the United States of America* 99, 12663–12668.
- Yamashita, A., Chang, T.C., Yamashita, Y., *et al.*, 2005. Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nature Structural & Molecular Biology* 12, 1054–1063.

miRNAs/Small Noncoding RNAs

R Fukunaga, Johns Hopkins University School of Medicine, Baltimore, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Argonaute (AGO) Proteins that are guided to mRNA targets by small silencing RNAs including microRNAs.

Dicer Ribonuclease III endonuclease that liberates microRNA–microRNA* duplexes from precursor microRNAs (pre-miRNAs). Dicer associates with partner proteins containing multiple double-stranded RNA-binding domains (dsRBDs), such as transactivation-response RNA-binding protein (TRBP), protein kinase R-activating protein (PACT), and Adenosine deaminase acting on RNA 1 (ADAR1) in mammals. The plant Dicer (DCL1, Dicer-like 1) cleaves primary microRNAs (pri-miRNAs) into pre-miRNAs, as well as pre-miRNAs into microRNA–microRNA* duplexes.

Drosha The nuclear RNase III endonuclease in animals that cleaves the base of a stem–loop structure contained in pri-miRNAs to produce a pre-miRNA. Drosha associates with partner proteins DGCR8 or Pasha, which contain dsRBDs. The Drosha–DGCR8/Pasha complex is called Microprocessor. Plants lack Drosha and its partner proteins.

dsRNA-binding domain (dsRBD) Approximately 70 amino acid domain that binds to double-stranded RNA (dsRNA). Proteins containing dsRBDs are called double-stranded RNA-binding proteins (dsRBPs). Many of the protein factors in the miRNA pathway have dsRBDs.

microRNA (miRNA) Approximately 22 nucleotide long, small noncoding RNAs that mediate posttranscriptional gene silencing. miRNAs are produced in a form of microRNA–microRNA* duplex by Dicer cleavage of pre-miRNA. microRNA* is the partially complementary strand to the miRNA strand. The miRNA strand becomes single stranded within the Argonaute protein after the miRNA* strand is released from the Argonaute protein, and forms mature miRNA-induced silencing complex (miRISC) capable of binding and silencing target mRNAs.

Precursor miRNA (pre-miRNA) Approximately 60–70-nt long stem–loop RNAs produced from pri-miRNAs by Drosha cleavage. Pre-miRNAs have a single-stranded loop that connects two partially complementary sequences, which will become the miRNA strand and the miRNA* strand after Dicer cleavage. Pre-miRNAs have a 5' monophosphate and a two nucleotide 3' overhang.

Primary miRNA (pri-miRNA) RNA polymerase II transcripts containing 7 methylguanosine cap at the 5' end, polyadenylated tail at the 3' end, and a stem–loop structure that serves as a substrate for Drosha in animals and DCL1 in plants. Pri-miRNAs are processed to liberate a pre-miRNA and two unstable, single-stranded by-products.

Ribonuclease III (RNase III) Double-stranded RNA-specific endoribonuclease that generates products with two nucleotide 3' overhangs, a 5' monophosphate and a 3' hydroxyl group. Drosha and Dicer are Ribonuclease III enzymes.

RNA-induced silencing complex

(RISC) Ribonucleoprotein complex that consists of a small RNA guide strand bound to an Argonaute protein and mediates silencing of target RNAs. RISC containing miRNA is called miRISC.

Seed sequence Seven nucleotide motif in the nucleotide positions 2–8 of microRNA, counting from the 5' end. Seed sequence is organized by Argonaute to determine target-RNA recognition and binds target mRNAs.

Small interfering RNA (siRNA) Approximately 21 nucleotide long, small noncoding RNAs that mediate posttranscriptional gene silencing. siRNAs are produced by Dicer cleavage of precursor long dsRNA or short hairpin RNA (shRNA). The siRNA duplex becomes single stranded (the guide strand) within the Argonaute protein after one of the strands (the passenger strand) is released from the Argonaute protein. The siRNA guide strand bound by Argonaute forms mature siRNA-induced silencing complex (siRISC) capable of cleaving target RNAs.

Introduction

microRNAs (miRNAs) are small noncoding RNAs (~22 nucleotide long) found in animals, plants, and some viruses, which function in posttranscriptional gene silencing (Figure 1(a)). The miRNA functions are important in every aspect of biology, including cell physiology, cell cycle control, metabolism, immune response, apoptosis, development, differentiation, stress response, and aging. miRNAs bind target mRNAs via sequence complementary and silence the target mRNAs by translational suppression or mRNA destabilization (Figure 1(b)). miRNAs are generated in a highly regulated

pathway, where conserved protein factors play essential roles. Human genome contains over 1800 miRNAs, which target as many as 60% of human mRNAs. Mutations in the miRNA genes and the miRNA pathway genes are associated with human diseases. Small interfering RNAs (siRNAs) are also small noncoding RNAs (~21 nucleotide long), but they are distinct from miRNAs. siRNAs are typically produced from exogenous long double-stranded RNA (dsRNAs) such as virus and transposon or from artificially introduced short hairpin RNAs (shRNAs). siRNAs silence the target RNAs by cleaving them. In this article, biogenesis and function of miRNAs are reviewed.

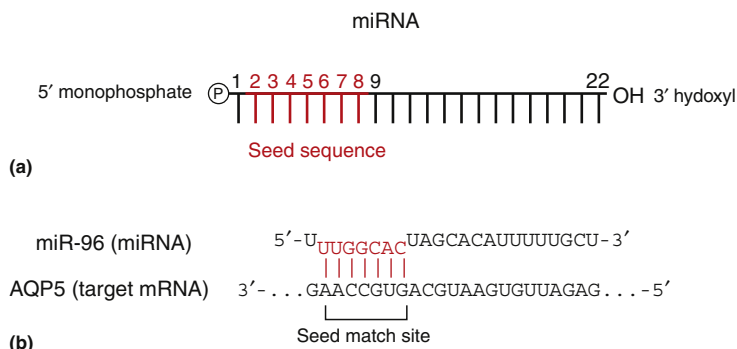


Figure 1 miRNA structure and the mode of miRNA binding to its target (a) miRNA structure. miRNA is approximately 22 nucleotide long, small RNA with 5' monophosphate and 3' hydroxyl at its ends. The sequence of the nucleotide positions 2–8 counting from the 5' end is called 'seed' sequence and used for binding to the target mRNAs. (b) miRNA binds its target mRNAs that contain sequence complementarity to its seed sequence. The complementary site in the target mRNA is called 'seed match' site. The binding of human miR-96 and one of its target mRNAs, AQP5, is shown.

Biogenesis of miRNAs in Animals

miRNAs are Transcribed by RNA Polymerase II

miRNAs are encoded in genomic DNA. miRNAs are first transcribed as long primary miRNA (pri-miRNA) transcripts containing hairpin loop structure by RNA polymerase II (Figure 2). Pri-miRNAs are transcribed either from independent genomic transcription units or from the introns of protein-coding genes. Pri-miRNAs have a 7-methylguanosine cap and a 3' polyA tail (Cai *et al.*, 2004; Lee *et al.*, 2004a; Rodriguez *et al.*, 2004). About half of all mammalian miRNAs originate from polycistronic units containing multiple discrete hairpin loops from which mature miRNAs are produced (Altuvia *et al.*, 2005).

Drosha Cleaves Pri-miRNAs into Pre-miRNAs

Pri-miRNAs are cleaved in the nucleus by the Ribonuclease III (RNase II)I enzyme Drosha, aided by its partner protein DGCR8 (DiGeorge syndrome critical region gene 8) in mammals or Pasha in flies and worms, into an approximately 60–70-nt long hairpin RNA, called precursor miRNAs (pre-miRNAs) (Figure 2; Lee *et al.*, 2003, 2004a; Denli *et al.*, 2004; Gregory *et al.*, 2004; Han *et al.*, 2004; Landthaler *et al.*, 2004).

DGCR8 and Pasha contain two dsRNA-binding domains (dsRBDs) (Figure 3).

The complex containing Drosha and DGCR8/Pasha is called Microprocessor complex. The Microprocessor complex also contains a variety of cofactors such as RNA helicases and heterogeneous nuclear ribonucleoproteins (hnRNPs). These auxiliary factors may function to control the activity, specificity, and localization, of Drosha cleavage.

In the Microprocessor complex, DGCR8/Pasha orients the catalytic RNase III domains of Drosha to liberate pre-miRNA hairpins from pri-miRNA by cleaving RNA about 11 nts from the hairpin base, a junction between single-stranded RNA (ssRNA) and dsRNA, which corresponds to one helical turn into the stem (Han *et al.*, 2006). pri-miRNAs can produce several pre-miRNAs for polycistronic miRNAs. pre-miRNAs

bear the hallmarks common to all RNase III products: 2-nt 3'-overhang end with 5'-monophosphate and 3'-hydroxyl.

Exportin-5/Ran-GTP Complex Transports Pre-miRNAs from Nucleus to Cytoplasm

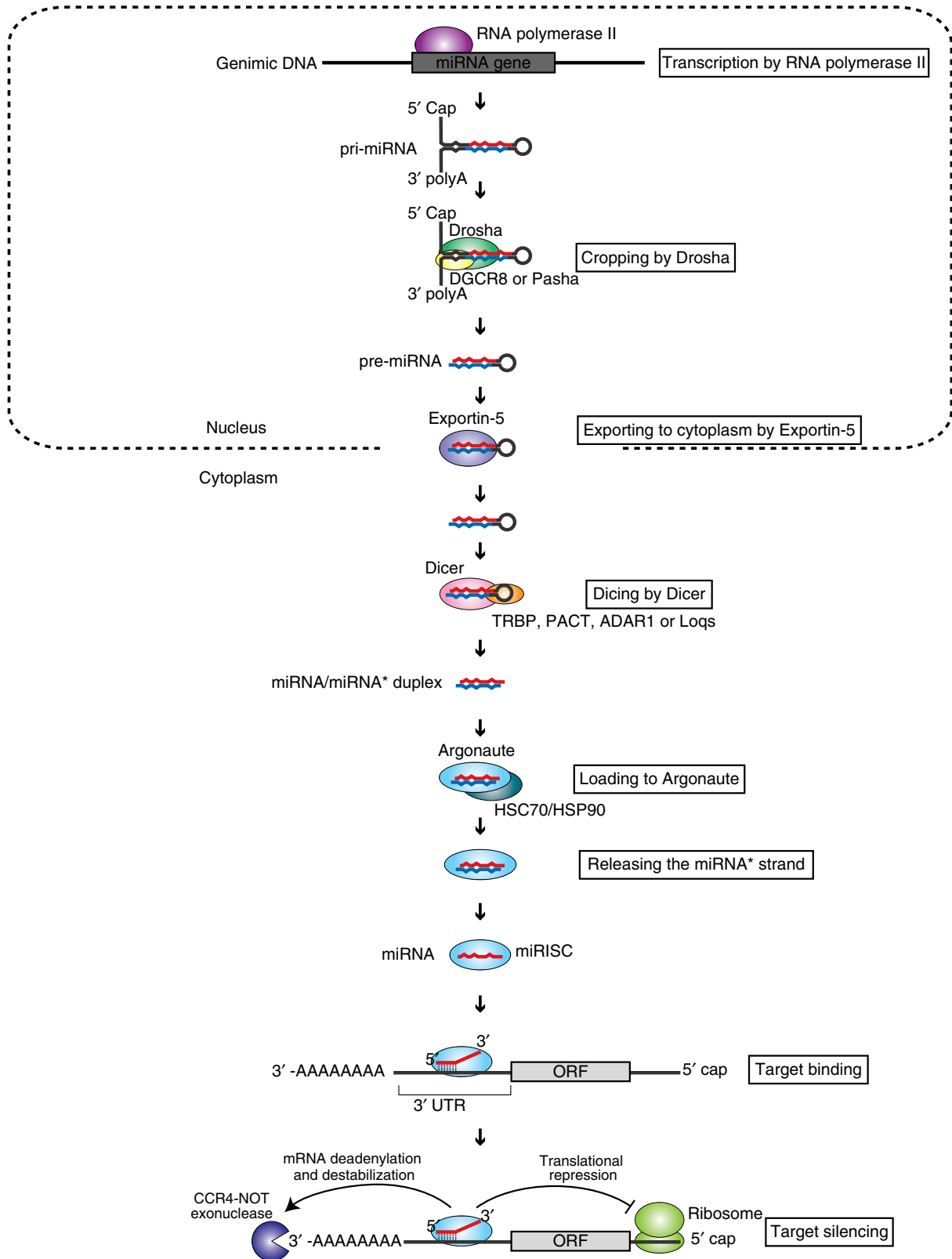
The Exportin-5/Ran-GTP complex forms a baseball mitt-like structure, recognizes the ends and the stem of pre-miRNAs, and transports them from the nucleus to the cytoplasm via the nuclear pore (Figure 2; Yi *et al.*, 2003; Bohnsack *et al.*, 2004; Lund *et al.*, 2004; Okada *et al.*, 2009).

Dicer Cleaves Pre-miRNAs into miRNA-miRNA* Duplexes

In cytoplasm, the second RNase III enzyme, Dicer, cleaves pre-miRNAs into approximately 22 nucleotide long miRNA-miRNA* duplexes comprising approximately 20 bp with 2-nt 3'-overhang ends (Figure 2; Bernstein *et al.*, 2001; Grishok *et al.*, 2001; Hutvagner *et al.*, 2001; Ketting *et al.*, 2001). The two strands of the duplex correspond to the mature miRNA and its miRNA*, a partially complementary small RNA derived from the opposite arm of the pre-miRNA stem. Both miRNA and miRNA* strands have 5'-monophosphate and 3'-hydroxyl (Figure 1(a)).

In mammals and worms, the single Dicer enzyme functions in converting pre-miRNAs into miRNA-miRNA* duplexes. In flies, which has two Dicer enzymes, Dicer-1 cleaves pre-miRNAs into miRNA-miRNA* duplexes, while Dicer-2 functions in biogenesis of siRNA duplexes (Lee *et al.*, 2004b). Like Drosha, Dicer recognizes defined RNA structures of pre-miRNAs and then cleaves at a fixed distance away from the base of the pre-miRNA stem, cutting off the loop joining the 5' and 3' arms to produce approximately 22 nucleotide mature miRNA-miRNA* duplexes.

As Drosha binds with partner proteins DGCR8/Pasha encompassing tandem dsRBDs, Dicer also associates with partner proteins containing multiple dsRBDs (Figures 2 and 3). Mammalian Dicer has three partner proteins, transactivation-response RNA-binding protein (TRBP), protein kinase R-activating protein (PACT), and Adenosine deaminase acting



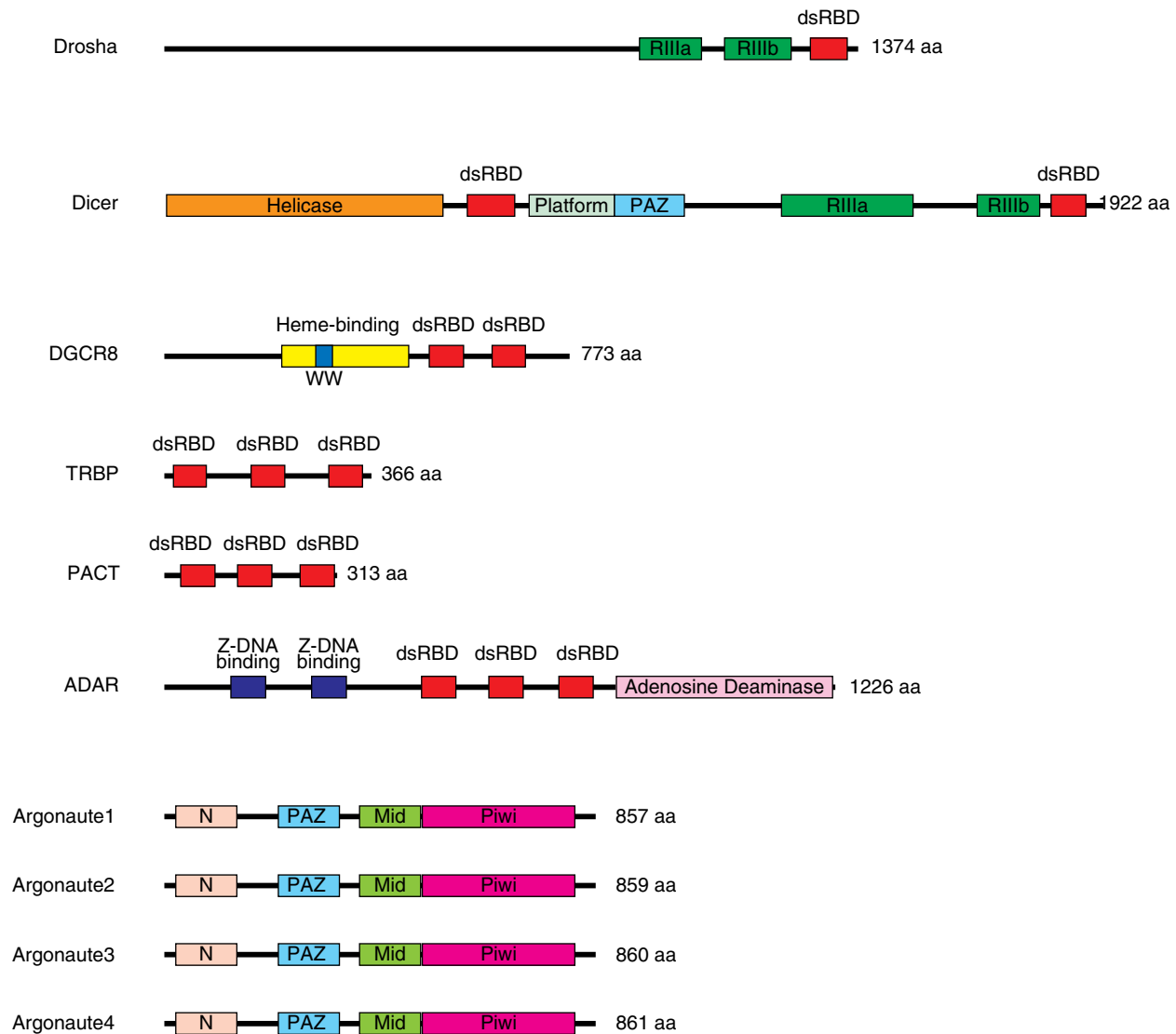


Figure 3 Domain structures of the miRNA pathway proteins. Domain structures of human proteins are shown. Drosha and Dicer are Ribonuclease III enzymes, DGCR8, TRBP, PACT, and ADAR are their partner proteins. Argonaute 1–4 are Argonaute proteins; RIIIa and RIIIb, Ribonuclease III domain; dsRBD, dsRNA-binding domain; Helicase, Helicase domain; Platform, Platform domain; PAZ, PAZ domain; Heme-binding, Heme-binding domain; WW, tryptophan–tryptophan (WW) motif; Z-DNA binding, Z-DNA binding domain; Adenosine Deaminase, Adenosine deaminase domain; N, N domain; Mid, Mid domain; Piwi, Piwi domain.

Figure 2 miRNA pathway. miRNA pathway in animals. miRNAs are transcribed from genomic DNA by RNA polymerase II (RNA pol II) as pri-miRNAs bearing a 7-methylguanosine cap and polyA tail. Drosha, assisted by its partner protein DiGeorge syndrome critical region gene 8 (DGCR8) in mammals or Pasha in flies and worms, cleaves pri-miRNAs into pre-miRNAs in the nucleus. Exportin-5 (Exp-5), assisted by Ran-GTP, transfers pre-miRNAs from the nucleus to the cytoplasm. Dicer, assisted by its partner proteins TRBP, PACT, and ADAR1 in mammals or Loquacious in flies, cleaves pre-miRNAs into miRNA–miRNA* duplexes. The miRNA–miRNA* duplexes are loaded into Argonaute with the aid of the chaperone machinery HSP90 and HSC70. Then miRNA* is released from the Argonaute complex and miRNA becomes single stranded within the Argonaute complex. The mature complex composed of an Argonaute protein and a single-stranded small silencing RNA is called RISC, and when the small silencing RNA is miRNA, the complex is called miRISC. miRISC binds target mRNAs using the sequence complementarity between the seed region of miRNA and the target mRNAs. The seed region is located at the nucleotide positions 2–8 of miRNA counting from the 5' arm of miRNA. The binding sites (seed match sites) are usually located within the 3' untranslated region (3' UTR), rather than 5' UTR or open reading frame (ORF), of mRNAs. miRISC silences the target mRNAs by repressing its translation or destabilizing the mRNAs. For the translational repression, miRISC inhibits the initiation step of translation. For mRNA destabilization, miRISC causes deadenylation of poly A tails of the mRNA targets by the CCR4–NOT exonuclease, leading to decay of the mRNAs. Other molecular mechanisms for target mRNA suppressions are also proposed.

on RNA 1 (ADAR1) (Chendrimada *et al.*, 2005; Haase *et al.*, 2005; Lee *et al.*, 2006; Ota *et al.*, 2013). Fly Dicer-1 associates with Loquacious-PA and Loquacious-PB (Forstemann *et al.*, 2005; Jiang *et al.*, 2005; Saito *et al.*, 2005). The Dicer partner proteins can enhance miRNA processing by Dicer by increasing the affinity or enzymatic turn over of Dicer for pre-miRNAs and refine the substrate specificity of Dicer (Chakravarthy *et al.*, 2010; Fukunaga *et al.*, 2012; Ota *et al.*, 2013). For a subset of pre-miRNAs, the Dicer partner proteins can also alter the position of Dicer-mediated cleavage within the pre-miRNAs, producing different miRNA isoforms (isomiRs) with different length compared with the miRNA isoforms produced by Dicer alone or Dicer bound by alternative partner proteins (Fukunaga *et al.*, 2012; Lee and Doudna, 2012).

miRNA-miRNA* Duplexes are Loaded to Argonaute

miRNA-miRNA* duplexes are loaded into Argonaute proteins (Figure 2). The loading step is supported by the chaperone machinery HSC70-HSP90 consuming ATP, which presumably accelerates the conversion of the Argonaute protein conformation from a closed to a more open structure receptive to binding miRNA-miRNA* duplexes (Iwasaki *et al.*, 2010; Miyoshi *et al.*, 2010).

After the miRNA* strand is released, the Argonaute complex becomes functional miRNA-induced silencing complex (miRISC) encompassing the single-stranded miRNA strand capable of binding and silencing target mRNAs. Mismatches between the miRNA and miRNA*, particularly in the miRNA seed sequence (nucleotide positions 2–8) region, promote release of the miRNA* strand (Tomari *et al.*, 2007; Kawamata *et al.*, 2009).

Argonaute loading of the miRNA-miRNA* duplexes occur in such orientation that the miRNA strand, instead of the miRNA* strand, is maintained at the later step. The inherent features of the miRNA-miRNA* duplex determine the loading orientation. The thermodynamic asymmetry of the miRNA duplex is one important parameter (Khvorova *et al.*, 2003; Schwarz *et al.*, 2003). The strand whose 5' end is less stably base-paired will be more frequently chosen as the miRNA strand or the guide strand. In contrast, the strand whose 5' end is more stably base-paired serves as the miRNA* strand or the passenger strand, and is excluded from the Argonaute and is eventually degraded. The sequence of the 5' end of the strand is another important parameter. A strand with 5' uridine is favored as a miRNA strand.

Flies have two Argonaute (Ago1 and Ago2) and miRNA-miRNA* duplexes are selectively loaded to Ago1, while siRNA duplexes are loaded to Ago2. Mammals have four Argonaute proteins (AGO1, AGO2, AGO3, and AGO4), which exhibit similar preferences for the structures of small RNA duplexes, suggesting that there is not clear selective loading mechanism in mammals.

Domain Structures of the miRNA Pathway Proteins

Both Drosha and Dicer have two RNase III domains (RNase IIIa and RNase IIIb) followed by the C-terminal dsRBD (Figure 3). The two RNase III domains form an intramolecular heterodimer and make staggered cleavages in the two arms of a

pri-miRNA or pre-miRNA (Zhang *et al.*, 2004; Macrae *et al.*, 2006). Cleavage by RNase III domains results in 2-nt 3'-overhang end with a 5'-terminal monophosphate, and a 3'-hydroxyl in the product RNA.

In addition, Dicer has an amino-terminal 'helicase' domain, a central atypical dsRBD (previously known as DUF283), a platform domain, and a PAZ domain. PAZ domains are found in both Dicer and Argonaute proteins and recognize the characteristic 2-nt 3'-overhang of pre-miRNA and miRNA-miRNA* duplex left by Drosha and Dicer cleavage, respectively (Lingel *et al.*, 2003, 2004). In addition, the platform domain and the PAZ domain of Dicer recognizes the 5' monophosphate on pre-miRNAs (Park *et al.*, 2011; Fukunaga *et al.*, 2014; Tian *et al.*, 2014).

The entire Dicer protein forms an L-shaped structure (Wang *et al.*, 2009; Lau *et al.*, 2009, 2012; Taylor *et al.*, 2013). The two RNase III domains and the C-terminal dsRBD are positioned in the body, while the PAZ domain is located at the head of the longer axis of the L. The helicase domain forms a clamp-shaped structure at base.

DGCR8 has a heme-binding domain and two dsRBDs. A tryptophan-tryptophan (WW) motif is located within the heme-binding domain. DGCR8 forms homodimer via a heme-binding domain interaction, where the WW motif functions as a platform for extensive dimerization interaction (Senturia *et al.*, 2012). Heme binding to DGCR8 is required for efficient pri-miRNA processing by the Drosha-DGCR8 complex (Faller *et al.*, 2007; Weitz *et al.*, 2014).

TRBP and PACT have three dsRBDs and are paralogous proteins. In addition to their role in the miRNA pathway by association with Dicer, these two proteins are involved in interferon response. Currently it is not known if there is any crosstalk between the miRNA pathway and the interferon pathway.

Like TRBP and PACT, ADAR1 has three tandem dsRBDs. ADAR1 also has N-terminal two tandem Z-DNA binding domains and a C-terminal adenosine deaminase domain. The Z-DNA binding domains bind to Z-DNA, the left-handed conformer of DNA. The adenosine deaminase domain converts adenosine residues into inosine (A-to-I editing) in RNA substrate. The deaminase catalytic activity is not required for ADAR1 to bind Dicer and promote its miRNA biogenesis activity (Ota *et al.*, 2013).

Human Argonaute1, 2, 3, and 4 are highly paralogous proteins encompassing N domain, PAZ domain, Mid domain, and Piwi domain. The N domain is required for releasing of miRNA* strand after Argonaute is loaded with a miRNA-miRNA* duplex (Kwak and Tomari, 2012). The PAZ and Mid domains anchors the 3'- and 5'-ends of the guide miRNA strand, respectively. The Piwi domain adopts RNase-H like fold (Elkayam *et al.*, 2012; Schirle and Macrae, 2012). The piwi domain of Ago2, but not those of the other three Argonaute proteins, can cleave target mRNAs when the guide miRNA strand or the guide siRNA strand has extensive complementarity to the target RNAs.

miRNAs Suppress Target mRNAs

The Argonaute protein loaded with a single-stranded miRNA (miRISC) binds target mRNAs via sequence complementarity.

Here, only a limited region of the miRNA sequence, called the 'seed' sequence, a 7-nt long sequence located at the nucleotide positions 2–8 counting from the 5' end of the miRNA, is used for the complementary binding (Figures 1 and 2; Lewis *et al.*, 2005). Sequences that are complementary to the seed sequence, called 'seed matches,' are bound by miRISC. The complementary binding between a seed sequence and a seed match sequence contributes most of the energy for target binding (Haley and Zamore, 2004; Ameres *et al.*, 2007; Wee *et al.*, 2012). In the Argonaute–miRNA complex, the nucleotide bases of the seed sequence of miRNA are prearranged for binding to mRNA targets (Elkayam *et al.*, 2012; Schirle and Macrae, 2012). Seed matches can occur in any region of an mRNA but are more likely to silence target mRNAs effectively when they are in the 3' UTR (Grimson *et al.*, 2007; Gu *et al.*, 2009; Forman and Collier, 2010). The small size (7 nt) of the seed sequence means that a single miRNA can regulate many of different mRNAs. Conversely, a single mRNA encompassing multiple seed matches for different miRNAs can be bound and regulated by multiple different miRNAs. In fact, approximately 60% of all protein-coding genes in mammals are predicted to be regulated by miRNAs (Friedman *et al.*, 2009). Thousands of other mRNAs seem to have experienced negative selection to avoid seed matches with miRNAs.

miRISC binding to mRNAs causes suppression of translation or degradation of the mRNAs. Despite extensive studies, the precise molecular mechanism by which miRISC causes translational repression and mRNA degradation are not well known. Many distinct mechanistic models for miRNA suppression of target genes have been proposed (Fabian *et al.*, 2010), including (1) inhibition of translation initiation by inhibiting interaction between an mRNA 5' cap and the 40S Ribosome subunit. (2) Inhibition of translation initiation by inhibiting 60S Ribosomal subunit joining. (3) Inhibition of translation elongation by inhibiting 80S Ribosome movement. (4) Causing premature termination of translation by inducing Ribosome drop-off. (5) Degradation of produced protein by co-translational nascent peptide degradation. (6) Destabilizing mRNA by enhancing deadenylation of polyA tail by CCR4-NOT exonuclease. (7) Degrading mRNA by cleaving them. (8) Decreasing mRNA abundance available for translation by sequestration mRNA in discrete cellular compartment called Processing-bodies (P-bodies). (9) Inhibition of transcription by reorganizing chromatin structure. Among these, the models (1) and (6) seem to have more supporting data than the others (Figure 2). It is likely that multiple of these proposed mechanisms operate and choice of the mechanisms might dependent on species, cellular context, Argonaute proteins, and miRNA–target context.

Some Argonaute proteins, such as mammalian AGO2, have catalytic activity to cleave RNA targets, while other Argonaute proteins, such as mammalian AGO1, AGO3, and Ago4, do not. When the miRNA loaded to the cleavage competent Argonaute proteins has extensive sequence complementarity to target mRNAs, then the Argonaute complex can cleave the target mRNAs. However, animal miRNAs generally do not have enough sequence complementarity to their endogenous mRNA targets and thus target mRNA regulation via cleavage is rare in animals.

miRNA Turnover

For rapid changes in miRNA expression profiles during development and when responding to cellular condition changes, turnover of mature miRNA is needed. During miRNA maturation in the cytoplasm, uptake by the Argonaute protein is thought to stabilize the guide miRNA strand. The released miRNA* strand is preferentially destroyed and miRNA–miRNA* duplexes that are not efficiently loaded to Argonaute for some reasons are believed to be unstable and degraded.

When Argonaute-loaded miRNA has unusually extensive complementarity to target RNAs, that miRNA is degraded by unknown mechanisms (Ameres *et al.*, 2010). Then the Argonaute protein becomes empty and can accommodate a next miRNA–miRNA* duplex (De *et al.*, 2013). In *Caenorhabditis elegans*, 5'→3' exoribonuclease XRN-2, degrades miRNA (Chatterjee and Grosshans, 2009).

Alternative Biogenesis Pathways of miRNAs

As described above, in the standard pathway of miRNA biogenesis, miRNAs are produced by sequential cleavage by the two RNase III enzymes, Drosha and Dicer. However, some miRNAs are produced by alternative pathways, which bypass one of the two cleavage steps.

Some pre-miRNAs, called mirtrons, are produced by splicing and debranching instead of Drosha cleavage (Okamura *et al.*, 2007; Ruby *et al.*, 2007). A few pre-miRNAs are produced directly by transcription by RNA polymerase II, independent of Drosha, and has a 7-methylguanosine cap at the 5' end and the 3' end is defined by transcription termination and lacks polyA tail (Xie *et al.*, 2013).

Like cropping of pri-miRNAs by Drosha can be bypassed for some pri-miRNAs, dicing of pre-miRNAs can also be bypassed. Pri-miR-451 is cropped into pre-miR-451 by Drosha in the standard pathway. However, the hairpin of pre-miR-451 in zebrafish and mice is too short to be recognized by Dicer; instead, it is directly loaded into AGO2 for further processing by AGO2 and Poly(A)-specific ribonuclease into a mature miR-451 within AGO2 (Cifuentes *et al.*, 2010; Yang *et al.*, 2010; Yoda *et al.*, 2013).

In the standard pathway, the miRNA strand is retained in the Argonaute protein while the miRNA* strand is released and degraded. There are cases in which both miRNA and miRNA* strands are functional (Okamura *et al.*, 2008, 2009; Seitz *et al.*, 2008; Czech *et al.*, 2009; Ghildiyal *et al.*, 2010). In such cases, a miRNA–miRNA* duplex can be loaded to Argonaute proteins in two different orientations: one orientation that favors the miRNA strand retention and the miRNA* strand release, the other orientation that favors the miRNA* strand retention and the miRNA strand release. Relative expression levels between the miRNA strand and the miRNA* strand reflect relative retention ratio in the Argonaute proteins and vary widely among tissues, indicating that strand selection factors other than sequence features might exist. Furthermore, for select pre-miRNAs, the loop strand, which is cleaved off by Dicer as a ssRNA and is degraded, can be loaded to Argonaute by unknown mechanism and silence target mRNAs (Okamura *et al.*, 2013). Thus, pre-miRNAs can be multifunctional, with individual strands adopting different fates within the miRNA biogenesis pathway.

miRNAs in Plants

Fungi do not have miRNAs. Plants have distinct miRNA sequences, miRNA precursor structures and mechanisms of biogenesis, compared with those in animals. These suggest that the miRNA pathway evolved independently between in animals and plants (Jones-Rhoades *et al.*, 2006). miRNAs are present in the earliest diverging extant lineage of animal life, sponge *Amphimedon*, suggesting possible correlation between the emergence of miRNAs and multicellularity of animals (Grimson *et al.*, 2008). In contrast, the unicellular alga *Chlamydomonas* has miRNAs and the biogenesis factors that are similar to those in higher plants, showing that the miRNA pathway evolved before multicellularity in plants (Molnar *et al.*, 2007).

Like in animals, plant miRNAs are transcribed by RNA polymerase II and pri-miRNAs have 5' 7-methylguanosine cap and 3' polyA tail (Jones-Rhoades and Bartel, 2004; Xie *et al.*, 2004, 2005; Zhang *et al.*, 2005). Unlike animals, plants lack Drosha enzyme. Instead of the sequential processing of pri-miRNAs by Drosha and then Dicer in animals, plant pri-miRNAs are cleaved twice by Dicer (DCL1, DICER-LIKE 1), first to produce pre-miRNA and then to liberate miRNA-miRNA* duplex. DCL1 is associated with two partner proteins: HYPONASTIC LEAVES 1 (HYL1) containing dsRBDs and SERRATE (SE) containing a zinc-finger domain (Hiraguri *et al.*, 2005; Yang *et al.*, 2006). Unlike in animals, the 3' end nucleotides of plant miRNA/miRNA* duplexes are 2'-O-methylated by the methyltransferase HUA-ENHANCER 1 (HEN1) before loading to Argonaute (Yu *et al.*, 2005).

Plant *Arabidopsis* has 10 Argonaute proteins. Canonical miRNAs are loaded to AGO1. Unlike in animals, plant miRNAs are highly complementary to target RNAs through their entire length and the high degree of complementarity is required for efficient target slicing. Plant miRISCs suppress target RNAs mainly by cleaving them, although they can induce translational repression as well.

Mutations in the miRNA and the Pathway Genes Cause Human Diseases

Since miRNAs are important for normal functioning of cells, dysregulation of miRNAs and the miRNA pathway are associated with various human diseases.

Mutations in the miRNA Pathway Genes Cause Diseases

Mutations in the Dicer gene are associated with tumorigenesis such as pleuropulmonary blastoma, Sertoli-Leydig cell tumor, multinodular goiter, and embryonal rhabdomyosarcoma (Hill *et al.*, 2009; Sabbaghian *et al.*, 2014). Mouse models suggest that Dicer may function as a haplo-insufficient tumor suppressor; a gene product whose normal function is to inhibit or control cell division in which one normal gene copy alone cannot produce enough product to prevent tumor formation. Furthermore, Dicer mutation is associated with age-related macular degeneration, in which toxic Alu RNA is no longer degraded by Dicer and thus accumulates (Kaneko *et al.*, 2011).

Some particular SNP types in the Exportin-5 and Drosha genes are suggested to be linked with cancer progression (Ye *et al.*, 2008; Wang *et al.*, 2010; Zhang *et al.*, 2010). Somatic mutations were found in the TRBP gene in colorectal and endometrial tumors and in gastric cancer cell lines and in the AGO2 gene in colorectal and gastric cancers (Melo *et al.*, 2009; Kim *et al.*, 2010). Mutations in the PACT gene are associated with Dyt16, an autosomal recessive, young-onset dystonia-parkinsonism disorder (Camargos *et al.*, 2008; Seibler *et al.*, 2008).

DiGeorge syndrome is caused by the chromosomal deletion 22q11.2, a region that includes the DGCR8 gene locus (Shiohama *et al.*, 2003). Haploinsufficiency of DGCR8 in the mouse model for this deletion causes abnormal miRNA biogenesis in the brain and the behavioral and neurological defects associated with DiGeorge syndrome (Stark *et al.*, 2008; Fenelon *et al.*, 2011).

Mutations in the miRNA Genes Cause Diseases

Germline mutations in the miR-96, miR-184, and miR-17 approximately 92 genes cause hearing loss, keratoconus, and skeletal and growth defects, respectively (Mencia *et al.*, 2009; De Pontual *et al.*, 2011; Hughes *et al.*, 2011).

About half of the miRNA genes map to cancer associated loci or in fragile sites (Calin *et al.*, 2004). Overexpression, down regulation, and deletion of many kinds of miRNAs have been reported to impact oncogenesis. For example, mutations in miR-15a and miR-16-1 are linked with tumorigenesis (Zhang *et al.*, 2006; Mavrakis *et al.*, 2010).

There are extensive studies being conducted that attempt to use miRNA expression profile as a biomarker in diagnosis and prognosis of various diseases including cancers. In addition, inhibiting particular miRNA functions using oligonucleotide with complementary sequence to the miRNA can be used as therapeutic approach.

Nomenclature of miRNA

The prefix 'miR' is followed by a dash and a number (e.g., miR-8), the latter often indicating order of naming. Exceptions are the miRNAs discovered early in the research history, such as lin-4 and let-7, which are the first (in 1993 by the Victor Ambros group) and the second (in 2000 by the Gary Ruvkun group) miRNAs to be discovered, respectively (Lee *et al.*, 1993; Pasquinelli *et al.*, 2000; Reinhart *et al.*, 2000). miRNAs with nearly identical sequences except for one or two nucleotides are annotated with an additional lower case letter. For example, miR-20a would be closely related to miR-20b. let-7a, let-7b, let-7c, let-7d, let-7e, let-7f, let-7g, and let-7i are closely related each other. pre-miRNAs that lead to 100% identical mature miRNAs but that are located at different places in the genome are indicated with an additional dash-number suffix. For example, the pre-miRNAs pre-miR-7-1, pre-miR-7-2, and pre-miR-7-3 lead to an identical mature miRNA (miR-7) but are located in different regions of the genome. When two mature miRNAs originate from opposite arms of the same pre-miRNA, they are denoted with a -3p or -5p suffix. For example, miR-22-5p and miR-22-3p derive from the 5' arm and 3' arm,

respectively, of pre-miR-22. Alternatively, when relative expression levels of two miRNAs deriving from the opposite arms of a pre-miRNA are known, an asterisk (star) following the name indicates a miRNA expressed at low levels relative to the miRNA in the opposite arm of a hairpin. For example, miR-22 and miR-22* (miR-22 star) would share a pre-miRNA hairpin (pre-miR-22), but miR-22 is more abundant than miR-22* in the cell. Species of origin is designated with a three-letter prefix, for example, hsa-miR-22 is a human (*Homo sapiens*) miRNA and mmu-miR-22 is a mouse (*Mus musculus*) miRNA. hsa-pri-miR-22 and has-pre-miR-22 are human pri-miR-22 and pre-miR-22, respectively.

See also: Molecular Principles, Components, Technology, and Concepts: Nucleic Acids: DNA, RNA Chemical Properties (Including Sequencing and Next-Generation Sequencing). Nucleic Acid Synthesis/Breakdown: RNA Synthesis/Function: Small RNAs/Cancer; The Interplay between Eukaryotic mRNA Degradation and Translation. Nucleic Acid Synthesis/Breakdown: Transcription: Eukaryotic Transcriptional Regulation

References

- Altuvia, Y., Landgraf, P., Lithwick, G., *et al.*, 2005. Clustering and conservation patterns of human microRNAs. *Nucleic Acids Research* 33, 2697–2706.
- Ameres, S.L., Horwich, M.D., Hung, J.H., *et al.*, 2010. Target RNA-directed trimming and tailing of small silencing RNAs. *Science* 328, 1534–1539.
- Ameres, S.L., Martinez, J., Schroeder, R., 2007. Molecular basis for target RNA recognition and cleavage by human RISC. *Cell* 130, 101–112.
- Bernstein, E., Caudy, A.A., Hammond, S.M., Hannon, G.J., 2001. Role for a bidirectional ribonuclease in the initiation step of RNA interference. *Nature* 409, 363–366.
- Bohnsack, M.T., Czapinski, K., Gorlich, D., 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* 10, 185–191.
- Cai, X., Hagedorn, C.H., Cullen, B.R., 2004. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* 10, 1957–1966.
- Calin, G.A., Sevignani, C., Dumitru, C.D., *et al.*, 2004. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences of the United States of America* 101, 2999–3004.
- Camargos, S., Scholz, S., Simon-Sanchez, J., *et al.*, 2008. DYT16, a novel young-onset dystonia-parkinsonism disorder: Identification of a segregating mutation in the stress-response protein PRKRA. *Lancet Neurology* 7, 207–215.
- Chakravarthy, S., Sternberg, S.H., Kellenberger, C.A., Doudna, J.A., 2010. Substrate-specific kinetics of Dicer-catalyzed RNA processing. *Journal of Molecular Biology* 404, 392–402.
- Chatterjee, S., Grosshans, H., 2009. Active turnover modulates mature microRNA activity in *Caenorhabditis elegans*. *Nature* 461, 546–549.
- Chendrimada, T.P., Gregory, R.I., Kumaraswamy, E., *et al.*, 2005. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* 436, 740–744.
- Cifuentes, D., Xue, H., Taylor, D.W., *et al.*, 2010. A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* 328, 1694–1698.
- Czech, B., Zhou, R., Erlich, Y., *et al.*, 2009. Hierarchical rules for Argonaute loading in *Drosophila*. *Molecular Cell* 36, 445–456.
- De, N., Young, L., Lau, P.W., *et al.*, 2013. Highly complementary target RNAs promote release of guide RNAs from human Argonaute2. *Molecular Cell* 50, 344–355.
- Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., Hannon, G.J., 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature* 432, 231–235.
- De Pontual, L., Yao, E., Callier, P., *et al.*, 2011. Germline deletion of the miR-17 approximately 92 cluster causes skeletal and growth defects in humans. *Nature Genetics* 43, 1026–1030.
- Elkayam, E., Kuhn, C.D., Tocilj, A., *et al.*, 2012. The structure of human argonaute-2 in complex with miR-20a. *Cell* 150, 100–110.
- Fabian, M.R., Sonenberg, N., Filipowicz, W., 2010. Regulation of mRNA translation and stability by microRNAs. *Annual Review of Biochemistry* 79, 351–379.
- Faller, M., Matsunaga, M., Yin, S., Loo, J.A., Guo, F., 2007. Heme is involved in microRNA processing. *Nature Structural & Molecular Biology* 14, 23–29.
- Fenelon, K., Mukai, J., Xu, B., *et al.*, 2011. Deficiency of Dgcr8, a gene disrupted by the 22q11.2 microdeletion, results in altered short-term plasticity in the prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America* 108, 4447–4452.
- Forman, J.J., Collier, H.A., 2010. The code within the code: MicroRNAs target coding regions. *Cell Cycle* 9, 1533–1541.
- Forstemann, K., Tomari, Y., Du, T., *et al.*, 2005. Normal microRNA maturation and germ-line stem cell maintenance requires Loquacious, a double-stranded RNA-binding domain protein. *PLoS Biology* 3, e236.
- Friedman, R.C., Farh, K.K., Burge, C.B., Bartel, D.P., 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research* 19, 92–105.
- Fukunaga, R., Colpan, C., Han, B.W., Zamore, P.D., 2014. Inorganic phosphate blocks binding of pre-miRNA to Dicer-2 via its PAZ domain. *EMBO Journal* 33, 371–384.
- Fukunaga, R., Han, B.W., Hung, J.H., *et al.*, 2012. Dicer partner proteins tune the length of mature miRNAs in flies and mammals. *Cell* 151, 533–546.
- Ghildiyal, M., Xu, J., Seitz, H., Weng, Z., Zamore, P.D., 2010. Sorting of *Drosophila* small silencing RNAs partitions microRNA* strands into the RNA interference pathway. *RNA* 16, 43–56.
- Gregory, R.I., Yan, K.P., Amuthan, G., *et al.*, 2004. The microprocessor complex mediates the genesis of microRNAs. *Nature* 432, 235–240.
- Grimson, A., Farh, K.K., Johnston, W.K., *et al.*, 2007. MicroRNA targeting specificity in mammals: Determinants beyond seed pairing. *Molecular Cell* 27, 91–105.
- Grimson, A., Srivastava, M., Fahey, B., *et al.*, 2008. Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. *Nature* 455, 1193–1197.
- Grishok, A., Pasquinelli, A.E., Conte, D., *et al.*, 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 106, 23–34.
- Gu, S., Jin, L., Zhang, F., Sarnow, P., Kay, M.A., 2009. Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs. *Nature Structural & Molecular Biology* 16, 144–150.
- Haase, A.D., Jaskiewicz, L., Zhang, H., *et al.*, 2005. TRBP, a regulator of cellular PKR and HIV-1 virus expression, interacts with Dicer and functions in RNA silencing. *EMBO Reports* 6, 961–967.
- Haley, B., Zamore, P.D., 2004. Kinetic analysis of the RNAi enzyme complex. *Nature Structural & Molecular Biology* 11, 599–606.
- Han, J., Lee, Y., Yeom, K.H., *et al.*, 2004. The Drosha-DGCR8 complex in primary microRNA processing. *Genes & Development* 18, 3016–3027.
- Han, J., Lee, Y., Yeom, K.H., *et al.*, 2006. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell* 125, 887–901.
- Hill, D.A., Ivanovich, J., Priest, J.R., *et al.*, 2009. DICER1 mutations in familial pleuropulmonary blastoma. *Science* 325, 965.
- Hiraguri, A., Itoh, R., Kondo, N., *et al.*, 2005. Specific interactions between Dicer-like proteins and HYL1/DRB-family dsRNA-binding proteins in Arabidopsis thaliana. *Plant Molecular Biology* 57, 173–188.
- Hughes, A.E., Bradley, D.T., Campbell, M., *et al.*, 2011. Mutation altering the miR-184 seed region causes familial keratoconus with cataract. *American Journal of Human Genetics* 89, 628–633.
- Hutvagner, G., McLachlan, J., Pasquinelli, A.E., *et al.*, 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293, 834–838.
- Iwasaki, S., Kobayashi, M., Yoda, M., *et al.*, 2010. Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes. *Molecular Cell* 39, 292–299.
- Jiang, F., Ye, X., Liu, X., *et al.*, 2005. Dicer-1 and R3D1-L catalyze microRNA maturation in *Drosophila*. *Genes & Development* 19, 1674–1679.
- Jones-Rhoades, M.W., Bartel, D.P., 2004. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Molecular Cell* 14, 787–799.
- Jones-Rhoades, M.W., Bartel, D.P., Bartel, B., 2006. MicroRNAs and their regulatory roles in plants. *Annual Review of Plant Biology* 57, 19–53.
- Kaneko, H., Dridi, S., Tarallo, V., *et al.*, 2011. DICER1 deficit induces Alu RNA toxicity in age-related macular degeneration. *Nature* 471, 325–330.
- Kawamata, T., Seitz, H., Tomari, Y., 2009. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nature Structural & Molecular Biology* 16, 953–960.

- Ketting, R.F., Fischer, S.E., Bernstein, E., *et al.*, 2001. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes & Development* 15, 2654–2659.
- Khvorov, A., Reynolds, A., Jayasena, S.D., 2003. Functional siRNAs and miRNAs exhibit strand bias. *Cell* 115, 209–216.
- Kim, M.S., Oh, J.E., Kim, Y.R., *et al.*, 2010. Somatic mutations and losses of expression of microRNA regulation-related genes AGO2 and TNRC6A in gastric and colorectal cancers. *Journal of Pathology* 221, 139–146.
- Kwak, P.B., Tomari, Y., 2012. The N domain of Argonaute drives duplex unwinding during RISC assembly. *Nature Structural & Molecular Biology* 19, 145–151.
- Landthaler, M., Yalcin, A., Tuschl, T., 2004. The human DiGeorge syndrome critical region gene 8 and its *D. melanogaster* homolog are required for miRNA biogenesis. *Current Biology* 14, 2162–2167.
- Lau, P.W., Guiley, K.Z., De, N., *et al.*, 2012. The molecular architecture of human Dicer. *Nature Structural & Molecular Biology* 19, 436–440.
- Lau, P.W., Potter, C.S., Carragher, B., Macrae, I.J., 2009. Structure of the human Dicer-TRBP complex by electron microscopy. *Structure* 17, 1326–1332.
- Lee, H.Y., Doudna, J.A., 2012. TRBP alters human precursor microRNA processing *in vitro*. *RNA* 18, 2012–2019.
- Lee, R.C., Feinbaum, R.L., Ambros, V., 1993. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843–854.
- Lee, Y., Ahn, C., Han, J., *et al.*, 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415–419.
- Lee, Y., Hur, I., Park, S.Y., *et al.*, 2006. The role of PACT in the RNA silencing pathway. *EMBO Journal* 25, 522–532.
- Lee, Y., Kim, M., Han, J., *et al.*, 2004a. MicroRNA genes are transcribed by RNA polymerase II. *EMBO Journal* 23, 4051–4060.
- Lee, Y.S., Nakahara, K., Pham, J.W., *et al.*, 2004b. Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* 117, 69–81.
- Lewis, B.P., Burge, C.B., Bartel, D.P., 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 120, 15–20.
- Lingel, A., Simon, B., Izaurralde, E., Sattler, M., 2003. Structure and nucleic-acid binding of the Drosophila Argonaute 2 PAZ domain. *Nature* 426, 465–469.
- Lingel, A., Simon, B., Izaurralde, E., Sattler, M., 2004. Nucleic acid 3'-end recognition by the Argonaute2 PAZ domain. *Nature Structural & Molecular Biology* 11, 576–577.
- Lund, E., Guttinger, S., Calado, A., Dahlberg, J.E., Kutay, U., 2004. Nuclear export of microRNA precursors. *Science* 303, 95–98.
- Macrae, I.J., Zhou, K., Li, F., *et al.*, 2006. Structural basis for double-stranded RNA processing by Dicer. *Science* 311, 195–198.
- Mavrikis, K.J., Wolffe, A.L., Oricchio, E., *et al.*, 2010. Genome-wide RNA-mediated interference screen identifies miR-19 targets in Notch-induced T-cell acute lymphoblastic leukaemia. *Nature Cell Biology* 12, 372–379.
- Melo, S.A., Ropero, S., Moutinho, C., *et al.*, 2009. A TARBP2 mutation in human cancer impairs microRNA processing and DICER1 function. *Nature Genetics* 41, 365–370.
- Mencia, A., Modamio-Hoybjør, S., Redshaw, N., *et al.*, 2009. Mutations in the seed region of human miR-96 are responsible for nonsyndromic progressive hearing loss. *Nature Genetics* 41, 609–613.
- Miyoshi, T., Takeuchi, A., Siomi, H., Siomi, M.C., 2010. A direct role for Hsp90 in pre-RISC formation in Drosophila. *Nature Structural & Molecular Biology* 17, 1024–1026.
- Molnar, A., Schwach, F., Studholme, D.J., Thuenemann, E.C., Baulcombe, D.C., 2007. miRNAs control gene expression in the single-cell alga *Chlamydomonas reinhardtii*. *Nature* 447, 1126–1129.
- Okada, C., Yamashita, E., Lee, S.J., *et al.*, 2009. A high-resolution structure of the pre-microRNA nuclear export machinery. *Science* 326, 1275–1279.
- Okamura, K., Hagen, J.W., Duan, H., Tyler, D.M., Lai, E.C., 2007. The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. *Cell* 130, 89–100.
- Okamura, K., Ladewig, E., Zhou, L., Lai, E.C., 2013. Functional small RNAs are generated from select miRNA hairpin loops in flies and mammals. *Genes & Development* 27, 778–792.
- Okamura, K., Liu, N., Lai, E.C., 2009. Distinct mechanisms for microRNA strand selection by *Drosophila* Argonautes. *Molecular Cell* 36, 431–444.
- Okamura, K., Phillips, M.D., Tyler, D.M., *et al.*, 2008. The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution. *Nature Structural & Molecular Biology* 15, 354–363.
- Ota, H., Sakurai, M., Gupta, R., *et al.*, 2013. ADAR1 forms a complex with Dicer to promote microRNA processing and RNA-induced gene silencing. *Cell* 153, 575–589.
- Park, J.E., Heo, I., Tian, Y., *et al.*, 2011. Dicer recognizes the 5' end of RNA for efficient and accurate processing. *Nature* 475, 201–205.
- Pasquinelli, A.E., Reinhart, B.J., Slack, F., *et al.*, 2000. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* 408, 86–89.
- Reinhart, B.J., Slack, F.J., Basson, M., *et al.*, 2000. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901–906.
- Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., Bradley, A., 2004. Identification of mammalian microRNA host genes and transcription units. *Genome Research* 14, 1902–1910.
- Ruby, J.G., Jan, C.H., Bartel, D.P., 2007. Intronic microRNA precursors that bypass Drosha processing. *Nature* 448, 83–86.
- Sabbaghian, N., Srivastava, A., Hamel, N., *et al.*, 2014. Germ-line deletion in DICER1 revealed by a novel MLPA assay using synthetic oligonucleotides. *European Journal of Human Genetics* 22, 564–567.
- Saito, K., Ishizuka, A., Siomi, H., Siomi, M.C., 2005. Processing of pre-microRNAs by the Dicer-1-Loquacious complex in *Drosophila* cells. *PLoS Biology* 3, e235.
- Schirle, N.T., Macrae, I.J., 2012. The crystal structure of human Argonaute2. *Science* 336, 1037–1040.
- Schwarz, D.S., Hutvagner, G., Du, T., *et al.*, 2003. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* 115, 199–208.
- Seibler, P., Djarmati, A., Langpap, B., *et al.*, 2008. A heterozygous frameshift mutation in PRKRA (DYT16) associated with generalised dystonia in a German patient. *Lancet Neurology* 7, 380–381.
- Seitz, H., Gildiyal, M., Zamore, P.D., 2008. Argonaute loading improves the 5' precision of both MicroRNAs and their miRNA* strands in flies. *Current Biology* 18, 147–151.
- Senturia, R., Laganowsky, A., Barr, I., Scheidemantle, B.D., Guo, F., 2012. Dimerization and heme binding are conserved in amphibian and starfish homologues of the microRNA processing protein DGCR8. *PLoS One* 7, e39688.
- Shiohama, A., Sasaki, T., Noda, S., Minoshima, S., Shimizu, N., 2003. Molecular cloning and expression analysis of a novel gene DGCR8 located in the DiGeorge syndrome chromosomal region. *Biochemical and Biophysical Research Communications* 304, 184–190.
- Stark, K.L., Xu, B., Bagchi, A., *et al.*, 2008. Altered brain microRNA biogenesis contributes to phenotypic deficits in a 22q11-deletion mouse model. *Nature Genetics* 40, 751–760.
- Taylor, D.W., Ma, E., Shigematsu, H., *et al.*, 2013. Substrate-specific structural rearrangements of human Dicer. *Nature Structural & Molecular Biology* 20, 662–670.
- Tian, Y., Simanshu, D.K., Ma, J.B., *et al.*, 2014. A phosphate-binding pocket within the platform-PAZ-connector helix cassette of human Dicer. *Molecular Cell*.
- Tomari, Y., Du, T., Zamore, P.D., 2007. Sorting of *Drosophila* small silencing RNAs. *Cell* 130, 299–308.
- Wang, H.W., Noland, C., Siridechadilok, B., *et al.*, 2009. Structural insights into RNA processing by the human RISC-loading complex. *Nature Structural & Molecular Biology* 16, 1148–1153.
- Wang, J., Lippman, S.M., Lee, J.J., *et al.*, 2010. Genetic variations in regulator of G-protein signaling genes as susceptibility loci for second primary tumor/recurrence in head and neck squamous cell carcinoma. *Carcinogenesis* 31, 1755–1761.
- Wee, L.M., Flores-Jasso, C.F., Salomon, W.E., Zamore, P.D., 2012. Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. *Cell* 151, 1055–1067.
- Weitz, S.H., Gong, M., Barr, I., Weiss, S., Guo, F., 2014. Processing of microRNA primary transcripts requires heme in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* 111, 1861–1866.
- Xie, M., Li, M., Vilborg, A., *et al.*, 2013. Mammalian 5'-capped microRNA precursors that generate a single microRNA. *Cell* 155, 1568–1580.
- Xie, Z., Allen, E., Fahlgren, N., *et al.*, 2005. Expression of Arabidopsis MIRNA genes. *Plant Physiology* 138, 2145–2154.
- Xie, Z., Johansen, L.K., Gustafson, A.M., *et al.*, 2004. Genetic and functional diversification of small RNA pathways in plants. *PLoS Biology* 2, E104.
- Yang, J.S., Maurin, T., Robine, N., *et al.*, 2010. Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 107, 15163–15168.
- Yang, L., Liu, Z., Lu, F., Dong, A., Huang, H., 2006. SERRATE is a novel nuclear regulator in primary microRNA processing in Arabidopsis. *Plant Journal* 47, 841–850.
- Ye, Y., Wang, K.K., Gu, J., *et al.*, 2008. Genetic variations in microRNA-related genes are novel susceptibility loci for esophageal cancer risk. *Cancer Prevention Research (Phila)* 1, 460–469.
- Yi, R., Qin, Y., Macara, I.G., Cullen, B.R., 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development* 17, 3011–3016.

- Yoda, M., Cifuentes, D., Izumi, N., *et al.*, 2013. Poly(A)-specific ribonuclease mediates 3'-end trimming of Argonaute2-cleaved precursor microRNAs. *Cell Reports* 5, 715–726.
- Yu, B., Yang, Z., Li, J., *et al.*, 2005. Methylation as a crucial step in plant microRNA biogenesis. *Science* 307, 932–935.
- Zhang, B.H., Pan, X.P., Wang, Q.L., Cobb, G.P., Anderson, T.A., 2005. Identification and characterization of new plant microRNAs using EST analysis. *Cell Research* 15, 336–360.
- Zhang, H., Kolb, F.A., Jaskiewicz, L., Westhof, E., Filipowicz, W., 2004. Single processing center models for human Dicer and bacterial RNase III. *Cell* 118, 57–68.
- Zhang, L., Huang, J., Yang, N., *et al.*, 2006. microRNAs exhibit high frequency genomic alterations in human cancer. *Proceedings of the National Academy of Sciences of the United States of America* 103, 9136–9141.
- Zhang, X., Yang, H., Lee, J.J., *et al.*, 2010. MicroRNA-related genetic variations as predictors for risk of second primary tumor and/or recurrence in patients with early-stage head and neck cancer. *Carcinogenesis* 31, 2118–2123.
- Jinek, M., Doudna, J.A., 2009. A three-dimensional view of the molecular machinery of RNA interference. *Nature* 457 (7228), 405–412. PMID:19158786. doi:10.1038/nature07755.
- Kim, V.N., Han, J., Siomi, M.C., 2009. Biogenesis of small RNAs in animals. *Nature Reviews Molecular Cell Biology* 10 (2), 126–139. PMID:19165215. doi:10.1038/nrm2632.
- Kwak, P.B., Iwasaki, S., Tomari, Y., 2010. The microRNA pathway and cancer. *Cancer Science* 101 (11), 2309–2315. PMID:20726859. doi:10.1111/j.1349-7006.2010.01683.x.
- Mendell, J.T., Olson, E.N., 2012. MicroRNAs in stress signaling and human disease. *Cell* 148 (6), 1172–1187. PMID:22424228. doi:10.1016/j.cell.2012.02.005.
- Rogers, K., Chen, X., 2013. Biogenesis, turnover, and mode of action of plant microRNAs. *Plant Cell* 25 (7), 2383–2399. PMID:23881412. doi:10.1105/tpc.113.113159.
- Sayed, D., Abdellatif, M., 2011. MicroRNAs in development and disease. *Physiological Reviews* 91 (3), 827–887. PMID:21742789. doi:10.1152/physrev.00006.2010.
- Siomi, H., Siomi, M.C., 2010. Posttranscriptional regulation of microRNA biogenesis in animals. *Molecular Cell* 38 (3), 323–332. PMID:20471939. doi:10.1016/j.molcel.2010.03.013.
- Wilson, R.C., Doudna, J.A., 2013. Molecular mechanisms of RNA interference. *Annual Review of Biophysics* 42, 217–239. PMID:23654304. doi:10.1146/annurev-biophys-083012-130404.
- Yang, J.S., Lai, E.C., 2011. Alternative miRNA biogenesis pathways and the interpretation of core miRNA pathway mutants. *Molecular Cell* 43 (6), 892–903. PMID:21925378. doi:10.1016/j.molcel.2011.07.024.

Further Reading

- Ameres, S.L., Zamore, P.D., 2013. Diversifying microRNA sequence and function. *Nature Reviews Molecular Cell Biology* 14 (8), 475–488. PMID:23800994. doi:10.1038/nrm3611.
- Carthew, R.W., Sontheimer, E.J., 2009. Origins and mechanisms of miRNAs and siRNAs. *Cell* 136 (4), 642–655. PMID:19239886. doi:10.1016/j.cell.2009.01.035.
- Czech, B., Hannon, G.J., 2011. Small RNA sorting: Matchmaking for Argonautes. *Nature Reviews Genetics* 12 (1), 19–31. PMID:21116305. doi:10.1038/nrg2916.
- Di Leva, G., Garofalo, M., Croce, C.M., 2014. MicroRNAs in cancer. *Annual Review of Pathology* 9, 287–314. PMID:24079833. doi:10.1146/annurev-pathol-012513-104715.
- Ghildiyal, M., Zamore, P.D., 2009. Small silencing RNAs: An expanding universe. *Nature Reviews Genetics* 10 (2), 94–108. PMID:19148191. doi:10.1038/nrg2504.
- Gurtan, A.M., Sharp, P.A., 2013. The role of miRNAs in regulating gene expression networks. *Journal of Molecular Biology* 425 (19), 3582–3600. PMID:23500488. doi:10.1016/j.jmb.2013.03.007.

Relevant Website

<http://www.mirbase.org/>
miRBase.

Small RNAs/Cancer

BM Ryan, Laboratory of Human Carcinogenesis Center for Cancer Research, Bethesda, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

3' compensatory binding Occurs when insufficient 5' seed pairing in a miRNA is compensated by extensive 3' end pairing.

Kindred A group of related individuals.

Lac repressor A DNA-binding protein that inhibits the expression of genes coding for proteins involved in the metabolism of lactose in bacteria.

Single nucleotide polymorphism A single nucleotide (A, T, C, G) variation in the DNA sequence.

Introduction to Noncoding RNAs

Less than 2% of our genome codes for protein-coding genes. However, it has become apparent in recent years that up to 90% of our genome is transcribed in one form or another. In addition to approximately 21 000 protein-coding genes, current estimates suggest that the human genome includes 9000 small RNAs, 10–32 000 long noncoding RNAs (lncRNAs) and 11 000 pseudogenes (Encode Project Consortium, 2012; Volders *et al.*, 2013). The list of small RNAs is long and includes; miRNAs (involved in posttranscriptional regulating of gene expression), transfer RNAs (involved in mRNA translation), small-interfering RNAs (implicated in posttranscriptional regulation of gene silencing), small nucleolar RNAs (involved in ribosomal RNA modifications), small nuclear RNAs (implicated in splicing), PIWI-interacting RNAs (involved in transposon repression), transcription initiation RNAs, promoter upstream transcripts, and promoter-associated small RNAs (all of which seem to be involved in regulation transcription) (Tay *et al.*, 2014). It is likely that this nonprotein coding portion of the genome contributed to our evolution. In yeast, the protein-coding portion of the genome accounts for most of the genome, while as noted above, in mammals the protein-coding portion accounts for only 2% (Encode Project Consortium, 2012).

Introduction to miRNAs: Discovery, Biogenesis, and Nomenclature

Discovery

The most studied class of small noncoding RNAs is microRNAs or miRNAs, single-stranded mRNA sequences of 19–25 nucleotides in length (Bartel, 2005). Their primary function is to negatively mediate posttranscriptional regulation of mRNA (Wakiyama *et al.*, 2007; Bartel, 2004), although there are some reports showing upregulation of a target gene (Vasudevan *et al.*, 2007; Orom *et al.*, 2008). The first miRNA was discovered in 1993 by Victor Ambros' laboratory (Lee *et al.*, 1993). While studying *Caenorhabditis elegans* development, they discovered that a small RNA encoded by the *lin-4* gene regulated *lin-14* protein abundance. At the same time, Gary Ruvkun's laboratory identified the first miRNA-target gene, also in *C. elegans* (Wightman *et al.*, 1993). Combined, these two reports uncovered a novel mechanism of posttranscriptional

gene regulation, and it was not until 7 years later that the second miRNA, *let-7*, was discovered (Reinhart *et al.*, 2000). In 2001, the Bartel, Tuschl, and Ambros laboratories independently identified additional miRNAs and miRNA conservation across species started to emerge (Lagos-Quintana *et al.*, 2001; Lau *et al.*, 2001; Lee and Ambros, 2001). Indeed, >50% of miRNAs found in *C. elegans* have homologs in humans, and although there are some examples of primate specific miRNAs (Robles and Harris, 2013; Bentwich *et al.*, 2005; Perdomo *et al.*, 2013), the vast majority of a miRNA's family tree is traced back as far as plants (Bartel, 2005). Although the discovery of the actual genes was novel, their mechanism of action, i.e., of acting in a posttranscriptional manner on a mRNA strand, had been postulated by Jacob and Monod in 'model II' as far back as 1961 when they were studying the mechanisms of the lac repressor (Jacob and Monod, 1961).

miRNA genes are smaller than protein-coding genes. They can be found within genes or as independent genes in intergenic regions. Intergenic miRNAs often lie in an antisense orientation to neighboring genes, which suggests that they have independent transcriptional control (for example, miR-375). For miRNAs that lie within genes, the host gene can be either protein-coding (e.g., miR-101-2) or nonprotein-coding (e.g., *let-7c*). Intragenic miRNAs are often co-transcribed with the host gene and tend to be located within introns (e.g., miR-608); miRNAs located within introns are also called miRtrons. However, some miRNAs can overlap with exons (e.g., miR-711) (He *et al.*, 2012; Rodriguez *et al.*, 2004) and miRBase currently counts 89 exonic miRNAs (Marsico *et al.*, 2013). More than 50% of miRNAs exist within clusters, the so-called *cistronic* miRNAs (they can be bi- or poly-cistronic) (Tanzer and Stadler, 2004). One of the main examples of a *cistronic* miRNA is the miR-17–92 cluster. It is transcribed as a single sequence and subsequently processed into 6 individual miRNAs. A useful website has been developed that lists miRNAs according to which area of the genome/types of genes they reside in (see Integratome TIME in Relevant Websites) (Godnic *et al.*, 2013).

Biogenesis

Both intergenic and within gene (intragenic) miRNAs are transcribed by RNA polymerase II (Lee *et al.*, 2004), have a 5' cap and 3' poly (A) tail (Cai *et al.*, 2004). A small number of human miRNAs, approximately 50, lie within Alu repeat

regions (e.g., the chromosome 19 miRNA cluster [C19MC]) (Bentwich *et al.*, 2005). Alu transcription is mediated by RNA polymerase III and miRNA genes within repetitive elements are also transcribed by RNA polymerase III (Borchert *et al.*, 2006). Intergenic miRNAs can be several kilobases in length, something that makes the annotation of their promoters difficult. However, several online bioinformatics tools have been developed to overcome this (Megraw *et al.*, 2009), including ARTS (Sonnenburg *et al.*, 2006) and CoreBoost_HM (Wang *et al.*, 2009). While intronic miRNAs in exons tend to be transcribed with the host gene, emerging evidence suggests that some intronic miRs have their own promoters (Monteys *et al.*, 2010; Oszolak *et al.*, 2008; Ladewig *et al.*, 2012) and interestingly, these promoters tend to lie within a different intron to the one in which the miRNA is located (Oszolak *et al.*, 2008). The role of these promoters is unclear, but a recently developed program that bioinformatically identifies promoters in intronic miRs could be used to further experimentally clarify the importance of such promoters (Marsico *et al.*, 2013). Interestingly, intragenic miRNAs that lie in exons are generally not predicted to have an independent promoter, suggesting that they may arise through splicing of the host gene.

The processing of intergenic and intragenic miRNA differs in some ways (Winter *et al.*, 2009) and not all components are fully understood. Following transcription of intergenic miRNAs into pri-miRNA sequences, the hairpin loops are recognized by DGCR8, which then associates with Drosha, an RNase III family member that cleaves the hairpin approximately 11 nucleotides from the base of the hairpin (Kim, 2005; Kim *et al.*, 2009). This new pre-miRNA has a two-nucleotide overhang at its 3' end; it also has 3' hydroxyl and 5' phosphate groups (it is usually 60–100 nucleotides in length). Intragenic miRNAs are often spliced directly from their host gene's mRNA transcript and bypass DGCR8 and Drosha processing, though for those with their own promoter a mechanism similar to intergenic miRNAs is likely to occur.

Pre-miRNA hairpins derived from either intergenic or intragenic regions are exported from the nucleus by Exportin-5/RAN-GTP – Exportin-5 recognizes the two-nucleotide overhang. In the cytoplasm, Dicer, TRBP, and Ago proteins bind to the hairpin and load it onto the RNA-induced silencing complex (RISC) where a mature miRNA duplex is formed by the RNase III action of Dicer. Each pre-miR produces two mature miRNA strands. In general, one of the strands is degraded and the other remains loaded in the RISC complex, ready for mRNA targeting (Lin *et al.*, 2005). The decision regarding which strand remains in the RISC complex seems to largely depend on thermodynamics.

When a miRNA finds its complimentary mRNA strand, Ago will either cleave the mRNA transcript (Lim *et al.*, 2005) or the complex will repress translation. The higher the degree of overlap between the miRNA sequence and the mRNA sequence to which it binds, the higher the likelihood that the sequence will be degraded (Bartel, 2004; Bartel, 2005). In plants, each miRNA normally targets a single mRNA and has very high sequence complementarity to its mRNA. Therefore miRNA-target binding in plants usually leads to cleavage of the mRNA strand via endonucleolytic activity. Recent work predicts that >60% of protein-coding genes are regulated by

miRNAs (Friedman *et al.*, 2009) and mRNA transcripts can have multiple miRNA binding sites within their 3'UTR (Hobert, 2004; Krek *et al.*, 2005; Hon and Zhang, 2007). Binding to an mRNA sequence by a miRNA is governed by a series of biochemical rules, though not all criteria are fully known. Current bioinformatics tools for predicting miRNA targets consider the degree of conservation, the 'seed' sequence (nucleotides 2–7 of the mature miRNA sequence), the number of sites within an mRNA target, and features of the surrounding mRNA sequence, such as secondary structure (Lewis *et al.*, 2003; Krek *et al.*, 2005). Common examples of bioinformatics tools that help to predict mRNA targets include TargetScan (Lewis *et al.*, 2003), miRanda (John *et al.*, 2004), rna22 (Loher and Rigoutsos, 2012), DIANA-microT (Margakakis *et al.*, 2011) and PicTar (Krek *et al.*, 2005). However, the gold standard approach for confirming miRNA targets remains empirical experimental testing.

The seed region is the primary sequence within the miRNA that confers specificity to a particular mRNA target; it can be considered analogous to that of consensus sequences in DNA that recognize transcription factors (Hobert, 2004). Although the binding between the seed region is (mostly) in perfect Watson–Crick complementarity, flanking regions do not have to bind with equal precision. However, studies show that key features of the transcript, including the exact configuration of mismatches in the mRNA sequences outside of the seed region, can influence the efficacy of translation repression (Aleman *et al.*, 2007; Grimson *et al.*, 2007). It should be noted that not all miRNA-mRNA binding takes place in the 3'UTR. Indeed, miRNAs have also been found to bind to coding regions (Xu *et al.*, 2009) as well as 5'UTRs (Orom *et al.*, 2008; Place *et al.*, 2008); these binding patterns however appear to be less frequent than canonical 3'UTR binding, but they are biologically pertinent. For example, binding of miR-10a to the 5'UTR of several ribosomal proteins actually enhances their translation (Orom *et al.*, 2008), while miR-122 binds to the 5' UTR of the hepatitis C RNA and promotes viral accumulation (Jopling *et al.*, 2005). A recent analysis of miRNA binding sites located within coding sequences suggests that they function to effectively inhibit translation, rather than mediate mRNA degradation (Xu *et al.*, 2009; Hausser *et al.*, 2013).

Nomenclature

The nomenclature of miRNAs has evolved since their discovery in 1993 (Griffiths-Jones *et al.*, 2006). They are initially annotated by species and start with a three or four letter prefix (for example, hsa for *Homo sapiens*) that is followed by a dash (for example, hsa-miR-145). An un-capitalized r refers to the pre-miRNA (for example mir-9), while the capitalized R refers to the mature miRNA (for example miR-9). Some miRNAs are annotated with a lower case letter, such as a or b (for example, miR-34a, miR-34b and miR-34c). This designation is used for miRNAs whose sequences are almost identical except for one or two nucleotides. Human miRNAs also have paralogs, likely to have emerged through gene duplication events. These miRNAs are annotated by a dash followed by a number (e.g., miR-29b-1 and miR-29b-2). As mentioned, a single pre-miRNA duplex will lead to 2 mature miRNA strands. These are annotated with a 5p or 3p, depending on which arm of the

hairpin the miRNA arose from. In the past, these sequences were designated with *, with the * assigned to the least thermodynamically stable, and therefore least abundant, transcript. In the past, the 5p strand was oft referred to as the guide strand, while the 3p strand was called the passenger strand. When referring to a miRNA gene, the consensus is similar to that of protein-coding genes, where italics are used (i.e., *hsa-mir-21*).

The latest release of the miRBase database in 2013 (V20) cataloged 1872 precursor miRNA sequences and 2578 mature sequences in humans. In recent times, miRBase has started annotating 'high confidence' miRNAs based on the rise in deposition of next generation sequencing reads (Kozomara and Griffiths-Jones, 2014). In general, their criteria for calling a miRNA sequence as a *bone fide* transcript is based on having reads that map to both mature sequences of a mature miRNA duplex, presence of 2 nucleotide 3' overhangs (a remnant of Drosha activity) and relatively consistent 5' ends. Currently, only 20% of miRNAs are classified as high confidence in *Homo Sapiens* (Kozomara and Griffiths-Jones, 2014).

miRNAs and Cancer

Given their key role in acting as a rheostat for protein output, it is not surprising that miRNAs have been linked to the etiology, progression and prognosis of cancer (Di Leva *et al.*, 2014). The first evidence that miRNAs were connected to cancer came from the laboratory of Carlo Croce in 2002 where he, George Calin and their colleagues were studying Chr13q14, a region that is frequently deleted in chronic lymphocytic leukemia (CLL). They long suspected the presence of a tumor suppressor in the region but detailed genetic analyses failed to identify any causal (classical) genes. And so it was that two miRNAs, *mir-15* and *mir-16-1*, were identified and demonstrated to function as tumor suppressors (Calin *et al.*, 2002). Since then, studies have shown that expression of miRNAs are deregulated in almost every human cancer studied (Calin and Croce, 2006; Lu *et al.*, 2005; Volinia *et al.*, 2006); this deregulation could be an early event in the carcinogenesis process and mediate the effect of carcinogenic environmental exposures (Schembri *et al.*, 2009). miRNAs can also define specific histological subtypes (Mathe *et al.*, 2009; Yanaihara *et al.*, 2006; Landi *et al.*, 2010) and are expressed in tissue-specific patterns (Lagos-Quintana *et al.*, 2002). miR-124, for example, is predominantly expressed in the brain where it mediates neuronal differentiation (Makeyev *et al.*, 2007; Lagos-Quintana *et al.*, 2002), while miR-122 is mainly found in the liver where it can promote viral accumulation (Lagos-Quintana *et al.*, 2002; Jopling *et al.*, 2005). In addition, approximately 50% of all annotated human miRNA genes are located in fragile sites or areas of the genome that are associated with cancer (Calin *et al.*, 2002, 2004; Sevignani *et al.*, 2007).

Etiology

When Croce and colleagues discovered miR-15 and miR-16-1, they also discovered the first inherited mutation in a miRNA. A single nucleotide polymorphism (SNP) was found within *pri-mir-16-1* in a kindred with familial CLL (Calin *et al.*, 2005). This SNP was functional and led to decreased levels of miR-16-

1, most likely due to disruption of pri-pre processing. Interestingly, the same germ line mutation is found in New Zealand black mice that naturally develop CLL-like disease (Raveche *et al.*, 2007). Since then, SNPs in *mir-196a-2* (rs11614913), *mir-146a* (rs2910164) and others have been associated with risk of cancer in multiple studies (Wang *et al.*, 2014; Chen *et al.*, 2013; Zhang *et al.*, 2012; Jazdzewski *et al.*, 2008, 2009). Sequencing has shown that SNPs in miRNA genes, and specifically in miRNA seed regions, are rare (Saunders *et al.*, 2007; Chen and Rajewsky, 2006) and, owing to their conservation, any changes in a miRNA sequence are likely to be functional. Most changes will affect biogenesis and processing (as in the case of miR-16-1), or alter targeting (as in the case of miR-146a). rs2910164 in miR-146a was originally found to be associated with risk of papillary thyroid carcinoma – it was an usual case though, as the heterozygote genotype (GC) was associated with increased risk of cancer, rather than either the wild-type (CC) or homozygous variant (CC) genotypes. This phenomenon, called over-dominance (Parsons and Bodmer, 1961), is rare in genetic association studies but there was a novel biological reasoning behind it that is somewhat specific to the biology of miRNAs. As each pre-miR strand generates two mature miRNAs, when the genotype is homozygous for either the G (ancestral) or C allele (derived allele), there are two possible mature stands that are produced from the mother and father's copy of chromosome 5. However, if someone carries a G allele from their mother and the C allele from their father, then there are now three potential mature miRNA produced. Jazdzewski and colleagues found that these strands had different mRNA targets and that this effectively contributed to an increased risk of cancer (Jazdzewski *et al.*, 2008, 2009).

SNPs in the 3' UTR of a gene may create, as well as destroy, a miRNA binding site (Ryan *et al.*, 2010). Abelson *et al.* (2005) elegantly depicted the first demonstration of a 3' UTR miRNA-modulating SNP in 2005, which found that a mutation in the miR-189 binding site of *SLITRK1* was associated with Tourette's syndrome. We now know that disruption of miRNA-dependent regulation by SNPs in the miRNA binding site of target mRNAs is a *bona fide* mechanism for altered gene expression in cancer. In 2005 Frank Slack's laboratory was the first to show that the miRNA let-7 binds to the 3' UTR of KRAS and regulates its expression (Johnson *et al.*, 2005). Subsequently, a novel SNP within a let-7 binding site was discovered and found to be associated with a 2.3 fold increased risk of lung cancer among moderate smokers (Chin *et al.*, 2008). This SNP has also been associated with ovarian (Pharoah *et al.*, 2011), oral (Christensen *et al.*, 2009) and colorectal cancer outcome (Ryan *et al.*, 2012b; Ruzzo *et al.*, 2011; Smits *et al.*, 2011). There are many other examples of 3'UTR SNPs associated with risk of cancer, many of which have been shown to be functional (Zanetti *et al.*, 2012). Some data also suggests that SNPs in the miRNA processing machinery are linked to the etiology of cancer (Ye *et al.*, 2008; Clague *et al.*, 2010; Yang *et al.*, 2008).

Predicting Prognosis and Response to Therapy

miRNAs are also associated with prognosis and response to therapy. miR-21 was initially shown to be associated with both

lung and colon cancer survival by the Harris laboratory (Schetter *et al.*, 2008; Yanaihara *et al.*, 2006). This work has since been replicated by many groups and indeed high miR-21 levels are associated with the prognosis of over 12 cancer types (Nair *et al.*, 2012; Fu *et al.*, 2011; Oue *et al.*, 2014; Mathe *et al.*, 2009; Landi *et al.*, 2010). Let-7 is downregulated in lung cancer, and this decreased expression is associated with poor prognosis (Takamizawa *et al.*, 2004; Yanaihara *et al.*, 2006). Some miRNA-related SNPs have been associated with prognosis (Zheng *et al.*, 2013; Ryan *et al.*, 2012a,b).

In tandem with predicting prognosis, an important part of the clinical management of cancer patients is knowing which therapy a patient is most likely to respond to. There has been considerable success in breast and lung cancer in using the molecular features of a tumor to guide and predict response to therapy (estrogen receptor positivity for TamoxifenTM response, HER-2 positivity for Herceptin[®] response, EGFR-mutation positivity for Erlotinib response and the *EML4-ALK* fusion gene positivity for guiding response to Crizotinib). Given the documented role of miRNAs in cancer (discussed below), there is an increasing recognition that miRNA patterns could also predict response to treatment. High levels of miR-21, for example, have been associated with resistance to chemotherapy in several types of cancer (Zhang *et al.*, 2011; Kurashige *et al.*, 2012; Wang *et al.*, 2013; Valeri *et al.*, 2010; Tomimaru *et al.*, 2010; Oue *et al.*, 2014). This correlation is most likely derived from (1) its upregulation in the cancer versus normal tissue, and (2) its role as a negative regulator of several cell death genes. Studies have also documented a decrease in circulating miRNAs in the weeks following surgery or treatment for cancer (Le *et al.*, 2012; Jung *et al.*, 2012; Zheng *et al.*, 2011; Sun *et al.*, 2012; Schwarzenbach *et al.*, 2014), suggesting that miRNAs could also be used as noninvasive predictors of tumor recurrence (Ohyashiki *et al.*, 2011; Schwarzenbach *et al.*, 2014).

The Role of miRNAs in Carcinogenesis

While dysregulation of miRNA expression in cancer could reflect a consequence rather than a cause of cancer, multiple lines of evidence suggest that miRNAs are indeed directly involved in the process of carcinogenesis. Examples of miRNAs acting as oncogenes (the so-called oncomiRs), tumor repressors, or both, have been described. For example, transgenic expression of miR-155 (Costinean *et al.*, 2006) or miR-21 (Medina *et al.*, 2010) leads to malignancy. Also, the aforementioned miR-17–92 cluster accelerates cancer induced by c-myc in mouse models of B-cell lymphoma (He *et al.*, 2005); transgenic mice expressing the miR-17–92 cluster developed tumors twice as early as tumors without the cistronic cluster of miRNAs. In addition to being involved in CLL, the miR-15 and miR-16-1 cluster is also lost in prostate cancer and multiple myeloma (Ambs *et al.*, 2008; Rocco *et al.*, 2009), while attempts at restoring the cluster actually sensitized cisplatin resistant cells (Pouliot *et al.*, 2012). Conversely, systemic delivery of tumor suppressive miRNAs, including let-7 and miR-34, inhibits tumor progression *in vivo* (Kota *et al.*, 2009; Pramanik *et al.*, 2011; Trang *et al.*, 2010; Wiggins *et al.*, 2010). Indeed, in 2013, the first clinical trial involving a human miRNA was initiated by Mirna Therapeutics, Inc. The Phase I

trial is being conducted in patients with un-resectable primary liver cancer or metastatic cancer with liver involvement.

As mentioned, the same miRNA can function as both an oncogene and tumor suppressor – depending on the tissue in which is expressed. Thus one cannot automatically infer a function from one tissue type to another. For example, miR-221 and miR-222 inhibit the growth of erythroblastic leukemia by targeting the oncogene KIT (Felli *et al.*, 2005). However, in models of liver cancer and hepatocellular carcinoma, miR-221 and miR-222 target PTEN and TIMP3 and induce TRAIL resistance, thereby acting as an oncomiR (Garofalo *et al.*, 2009). Thus, experimental evidence will always be needed to verify the cancer promoting or inhibiting effects of a miRNA.

miRNAs may also be related to cancer stem cell biology. The cancer stem hypothesis posits that a sub-fraction of cancer stem cells is responsible for the initiation, maintenance and progression of cancer. Central to this hypothesis is the stem cell characteristics that these tumor-initiating cells possess (Bonnet and Dick, 1997). Dysregulation of miRNAs has been found in cancer stem cells compared with healthy tissue. The miR-200 family, for example, is decreased in breast cancer stem cells (Shimono *et al.*, 2009; Polytaichou *et al.*, 2012). These miRNAs play a key role in suppressing epithelial-to-mesenchymal transition, migration, invasion and metastasis (Mongroo and Rustgi, 2010) through targets such as ZEB1 and ZEB2. An *in vitro* screen of miRNAs that could inhibit the growth of breast cancer stem cells also identified miR-15 and miR-16-1, miR-103 and miR-107, miR-145, miR-335, and miR-128b. In hepatocellular carcinoma, cancer stem cells defined by EpCAM expression express high levels of miR-181 family members (Ji *et al.*, 2009). Ji and colleagues found that inhibiting miR-181 led to a reduction in the cancer stem cell population and tumor-initiating ability, suggesting that targeting miRNAs could be an effective clinical strategy.

miRNAs as a Target for Therapy

As mentioned, cancers of nearly all kinds display global changes in miRNA expression. It is likely that some of these altered miRNAs are consequences, rather than causes, of carcinogenesis and it has been difficult to identify the true miRNA drivers. However, a few key oncogenic and tumor suppressor miRNAs have emerged and as stated above, a clinical trial involving a miRNA has already been initiated. The miR-34 family of genes, which includes miR-34a, miR-34b and miR-34c, are directly regulated by p53. They are classified as tumor suppressors, can be hyper-methylated in cancer (Gallardo *et al.*, 2009) and are located in fragile sites (Calin *et al.*, 2004). miR-34 antagonizes pathways that are essential for cancer cell viability as well as cancer stemness, metastasis, and chemoresistance. The results of this clinical trial will be eagerly awaited.

The goal of miRNA replacement therapy, for example, with miRNAs like miR-34, is to restore a loss of normal function (Bader *et al.*, 2010), but issues with RISC saturation have already emerged (Grimm *et al.*, 2006), as has the potential of inducing unwanted side-effects when new miRNAs are introduced into a cell. Encouragingly, however, preliminary animal models suggest that these side effects could be minimal (Kota *et al.*, 2009; Wiggins *et al.*, 2010).

Inhibiting oncogenic miRNAs can be initiated by several strategies, including anti-miRs, locked nucleic acids (LNAs), and antagomiRs, developed by Tuschl and colleagues. miRNA antagonists are single-stranded RNA molecules, approximately 21–23 nt long, that act through complementary base-pairing with miRNAs (Krutzfeldt *et al.*, 2007). To increase the pharmacological inhibition of target miRNAs, various chemical modifications are used to enhance binding affinity, confer nuclease resistance, and facilitate cellular uptake. These include adding phosphorothioate backbones, 2'sugar modifications such as 2'-O-methyl (2'-Ome) (the so-called antagomiRs) and LNAs (Krutzfeldt *et al.*, 2005; Levin, 1999). *In vivo* experiments with an antagomiR targeting miR-122 have demonstrated effective inhibition of miR-122 and decreases in plasma cholesterol levels (Esau *et al.*, 2006), while an LNA-modified form of a miR-122 anti-miR led to long-lasting suppression of hepatitis C virus in primates (Lanford *et al.*, 2010).

miRNAs are found in several biological fluids – not just tissue. Indeed, the discovery of miRNAs in the circulation led to the identification of a new function for miRNAs, i.e., as activators or ligands for Toll-like receptors (TLR) (Fabbri *et al.*, 2012; He *et al.*, 2014). Fabbri and colleagues found that miR-21 excreted from lung cancer cells in exosomes was taken up by macrophages wherein the miRNA bound to, and activated, TLR8, leading to the production of pro-inflammatory cytokines. Given that miRNAs are over-expressed and decreased in many forms of human cancer, this raises the possibility that miRNAs could be used as a form of 'liquid biopsy' and be useful in cancer diagnosis and predicting outcome (Cortez *et al.*, 2011). They are remarkably stable in these biospecimens, in large part due to their small size and protection from endogenous RNase activity (Mitchell *et al.*, 2008). In the circulation, cell-free miRNAs exist in 2 main forms – either bound to Ago2 or encapsulated within microvesicles, such as exosomes (Mitchell *et al.*, 2008). When trying to detect miRNAs in fluids such as serum or plasma, considerable attention needs to be paid to both sample preparation and RNA extraction – both of which have been shown to affect the endpoint of miRNA detection. Differential centrifugation will isolate different populations of cell-free miRNAs – centrifuging at >110 000 g, for example, will isolate exosomal miRNAs. Other factors to consider are the choice of serum versus plasma – hemolysis floods the circulating miRNA pool with myeloid-derived miRNAs, such as miR-223, miR-197, miR-594-3p, and let-7a, and conditions that lead to a fluctuation in blood cell levels can influence miRNA abundance levels by up to 50-fold (Pritchard *et al.*, 2012). RT-PCR and deep sequencing are the main methods used to identify circulating miRNAs at present. For a good review of this area see (Schwarzenbach *et al.*, 2014). Despite the challenges in identifying and isolating miRNAs in bodily fluids, some studies have shown significant associations with diagnosis and prognosis of cancer (Schwarzenbach *et al.*, 2014). However, robust and extensive validation of methods and results will be needed to verify these associations and move them into clinical use.

IsomiRs

Before their official annotation as 'isomiRs' by Morin *et al.* (2008), miRNA variants had been identified in cloning studies

but were ambiguously misclassified as putative experimental error (Ruby *et al.*, 2006; Landgraf *et al.*, 2007; Ebhardt *et al.*, 2009). Their discovery owes much to the acceleration in next generation sequencing technology and researchers are now confident that they are not due to RNA degradation during sample preparation for RNA sequencing (Bizuayehu *et al.*, 2012). miRbase records the most abundant miRNA sequence as the reference sequence. In recent years, it has started to include miRNA sequence heterogeneity as part of its database.

IsomiRs are formed via posttranscriptional editing. Four main forms of editing are known; 3' trimming, 5' trimming, 3' nucleotide addition (usually adenylation or uridylation) and internal nucleotide substitution. The latter two are less common than trimming events. The location of trimming variations suggests that they generally arise from variable cleavage sites for DROSHA and DICER1 in the hairpin (Lu *et al.*, 2009). Wyman and colleagues implicated the nucleotidyl transferases MTPAP, PAPD4, PAPD5, ZCCHC6, ZCCHC11, and TUT1 in 3' miRNA modification in human cells (Wyman *et al.*, 2011).

IsomiRs were initially described in plants and lower organisms (Morin *et al.*, 2008; Aravin and Tuschl, 2005). However, isomiRs are conserved from plants to humans (Wyman *et al.*, 2011). Recently, isomiRs have been detected in animals (Bizuayehu *et al.*, 2012) and *Drosophila melanogaster* (Fernandez-Valverde *et al.*, 2010). Furthermore, isomiRs have been described in human stem cells (Pantano *et al.*, 2010) and human brain tissue (Marti *et al.*, 2010). Although they are detected, a key question is whether they are biologically relevant. Recent work suggests that they are. For example, miR-282, and miR-312 have 3'adenosine additions during early development of *D. melanogaster*. These newly edited miRNAs show subsequent enrichment in target sites for developmental genes expressed during late embryogenesis, suggesting that non-DNA templated additions to a miRNA can strengthen miRNA-target gene interactions (Fernandez-Valverde *et al.*, 2010). The functionality of adenylation or uridylylation at the 3' end has been related to stability (Burroughs *et al.*, 2010). A recent study found that the addition of an adenine to miR-122 enhanced the stability of the miRNA, while plant miRNAs that end with an adenine have slower decay rates (Kai and Pasquinelli, 2010). The addition of uridines to the 3' end of miRNAs has been reported (Li *et al.*, 2005; Heo *et al.*, 2009; Heo *et al.*, 2008), but its significance is less understood. Some suggest that uridylation at the 3' end could mediate miRNA turnover or facilitate mRNA-miRNA binding in cases in which 3' compensatory binding is predominant. This addition may also function as a degradation signal, perhaps in a manner that is analogous to protein ubiquitylation (Li *et al.*, 2005).

miRNAs with 5' modifications are especially interesting as they will have a seed sequence that differs from the reference sequence, opening up the possibility that the isomiRs will have different pools of mRNA targets. Though rarer than 3' end modification events, modification at the 5' end of miRNAs has been observed in T cells (Wu *et al.*, 2007) and isomiRs of mir-142 are differentially expressed during antigen-induced T cell differentiation (Wu *et al.*, 2007). 5' end modifications are a conserved phenomenon (Wu *et al.*, 2007; Waidner *et al.*, 2009; Ruby *et al.*, 2006; Seitz *et al.*, 2008). An analysis of brain-derived miR-101 detected a 5' isomiR that was highly abundant, incorporated into RISC, and bound to miR-101 targets,

albeit with less efficiency than canonical miR-101 (Llorens *et al.*, 2013a). The identity of the 5' nucleotide directs miRNAs to Ago, though 3' end nucleotides could also have an effect. There are several members of the Ago family (Liu *et al.*, 2004), thus any changes in nucleotide identity could alter isoform-specific loading of miRNAs (Burroughs *et al.*, 2011; Juvvuna *et al.*, 2012).

Internal editing of mature miRNAs also occurs (Reid *et al.*, 2008), although there is a selection against nucleotide alterations in nucleotides 2–7 (seed) and 10–15 (cleavage and anchor sites). Internal editing could also affect stability and the pool of mRNA targets. Work suggests that this RNA-editing can alter the specificity by which miRNAs associate with Ago proteins (Landgraf *et al.*, 2007).

In a recent analysis of miRNAs from healthy individuals and patients with Huntington disease, 80–90% of small RNAs had some form of editing. Most of this editing occurred at the 3' end, and was in the form of trimming or nucleotide additions (Marti *et al.*, 2010). Differential expression of isomiRs has also been identified during differentiation (Wyman *et al.*, 2011). Strikingly, expression of isomiRs varied between healthy tissue and samples from patients with Huntington disease (Marti *et al.*, 2010). IsomiRs have also been implicated in cancer. In a mouse model of leukemia, several isomiRs were differentially expressed between healthy and diseased cells (Kuchenbauer *et al.*, 2008). For example, let-7a-5p had 78 sequences derived from various combinations of 5' and 3' modification: some of these sequences had counts greater than 4000 and were therefore highly expressed. The miR-181 family of putative tumor suppressors and miR-21, an oncomiR, also has a remarkable level of sequence variation. Several studies have shown differential expression of isomiRs in normal compared with cancer tissue (Li *et al.*, 2012; Llorens *et al.*, 2013b; Chan *et al.*, 2013). Although these variations have not been functionally interrogated in either normal or cancer biology, it is plausible that they are relevant to tumorigenesis and it is likely that studying and understanding miRNA editing will add greater understanding to both normal and cancer cell biology.

Other Forms of Noncoding RNAs and their Relevance to Cancer

ceRNAs

While it was first thought that miRNAs only bound to message RNA transcripts, it has become increasingly clear the much of the RNAome, including lncRNAs and pseudogenes, contains binding sites for miRNAs. This realization, and the possibility of cross talk regulation, led to the development of the term competing endogenous RNAs, or ceRNAs (Salmena *et al.*, 2011). ceRNAs may function as a decoy for miRNA binding. In T cells transformed with Herpesvirus saimiri, seven noncoding RNAs are produced that have binding sites for host miRNAs, including miR-27a. Binding between miR-27a with the viral noncoding RNAs leads to degradation of the miR-27a and thus downstream consequences for the targets of miR-27a. Such a cunning viral strategy illustrates a key example of how a noncoding RNA exerts a cross talk regulation of host-cell gene

expression via miRNAs (Cazalla *et al.*, 2010). In another example, binding of miRNAs to a PTEN pseudogene, PTENP1, regulates the effect of miRNAs by competing for their binding (Poliseno *et al.*, 2010).

eRNAs

Enhancers, genetic elements that regulate gene transcription upstream from the promoter region, were identified in 1981 by Schaffner and colleagues while studying the beta-globin gene (Banerji *et al.*, 1981). At the time, they were considered to be DNA elements and since their discovery, it has become clear that an enhancer can exist far from the gene it regulates. For example, in the case of sonic hedgehog its enhancer lies almost a megabase away (the enhancer lies in an 800-base pair region within intron 5 of the *Lmbr1* gene) (Lettice *et al.*, 2003). In 2010, Natoli and colleagues discovered that 70% of extragenic RNA polymerase II peaks were associated with enhancers (De Santa *et al.*, 2010). Michael Greenberg and colleagues subsequently found that enhancers activated by neuron depolarization were transcribed in multiple places throughout the genome and that expression of the enhancer RNAs (eRNA) correlated with mRNA of the nearest gene (Kim *et al.*, 2010). eRNAs can be transcribed both a single and bi-directional manner. Unidirectional transcription of enhancer regions usually generates long (4 kb) polyadenylated transcripts (Heintzman *et al.*, 2007), while those transcribed in a bi-directional manner are usually shorter (0.5–2 kb) and not polyadenylated (Koch *et al.*, 2011). It is not entirely clear whether expression of eRNAs in tandem with their mRNA target is a mark of active transcription, or whether it has a specific function, with some suggesting that eRNAs are indeed just a by-product of transcription. Whether or not it is a widespread phenomenon of all eRNAs, physiological functions have been noted for eRNAs. For example, Rosenfeld and colleagues found that when 17 β -estradiol is bound to its receptor in breast cancer cells (ER- α), an increase in eRNAs beside estrogen-regulated genes occurs. Removing the eRNAs abrogated their ability to increase gene expression (Li *et al.*, 2013). While more work is needed to understand whether eRNAs are broadly functional, their importance for regulating gene expression and potential contributions to disease (Lettice *et al.*, 2003) are established.

LncRNAs

LncRNAs can be between 200 nucleotides and 100 kilobases in length. One of the best known lncRNAs, XIST, was known and functionally characterized many years before lncRNAs were an official class of noncoding RNAs (Brockdorff *et al.*, 1992; Brockdorff *et al.*, 1991) (mouse XIST was discovered in 1991). XIST is a 15 kb length transcript and causes X-inactivation, a compensation mechanism developed in mammals to offset the unequal number of sex chromosomes in males and females. In female cells, XIST initially covers gene-rich islands on the X-chromosome before spreading to less gene dense regions (Simon *et al.*, 2013).

LncRNAs are primarily transcribed by RNA polymerase II though some cases of transcription by RNA polymerase III are described. (Bierhoff *et al.*, 2010). They are classified in terms of

their orientation and location relative to neighboring genes (Rinn and Chang, 2012). In addition to the scaffolding function as described for XIST, lncRNAs can also function as enhancers and decoys. Since their discovery as a class of noncoding RNAs, lncRNAs have undergone a few changes in nomenclature, but an official naming system has been established (Wright and Bruford, 2011). An in-depth discussion of lncRNAs is beyond the scope of this article, but it is important to mention that, as observed with miRNAs, deregulated lncRNA expression has been identified in a multitude of cancers (Tang *et al.*, 2013; Zhang *et al.*, 2009; Gutschner and Diederichs, 2012; Luo *et al.*, 2006; Kim *et al.*, 2013). The lncRNA MALAT-1 is associated with prognosis in lung cancer (Schmidt *et al.*, 2011). Recently, it was also identified as a target for miR-9 in an Ago2-dependent manner (Leucci *et al.*, 2013).

Transposons and piRNAs

A transposon, or transposable element, is a small piece of DNA that can insert itself into another part of the genome. They were discovered by Barbara McClintock in 1940s while she was studying corn and differing coloration patterns in kernels (McClintock, 1950). Transposons can travel or 'jump' to alternate regions of the genome, while in other cases they make copies of themselves and then 'paste' into new regions of the genome. In bacteria, for example, transposons often carry genes that allow these microorganisms to develop antibiotic resistance. Recently, The Cancer Genome Atlas (TCGA) identified transposons as a probable source of novel mutations across many tumor types. Specifically, almost 200 transposable element insertions were identified, 62 of which had 'jumped' to known tumor suppressors and often altered the expression of their new host gene (Lee *et al.*, 2012).

Small RNA-dependent silencing mechanisms have evolved to regulate the activity of transposable elements (Malone and Hannon, 2009). In *Drosophila*, for example, transposons are silenced by repeat associated siRNAs (rasiRNA) (Klattenhoff *et al.*, 2007). rasiRNAs are now called piRNAs (for piwi-interacting RNA) (Aravin and Tuschl, 2005). piRNAs are the biggest of the small noncoding RNAs at around 26–31 nucleotides in length – they are specifically expressed in the testis and thus are necessary for spermatogenesis (Malone and Hannon, 2009; Girard *et al.*, 2006). In mammals, the main activity of piRNAs is in transposon silencing, specifically during development of the embryo (Aravin and Tuschl, 2005). They are evolutionarily conserved (Wang and Reinke, 2008) and although their biogenesis is not very well understood, there is emerging evidence to suggest that piRNAs are involved in the epigenetic regulation of cancer (Yan *et al.*, 2014; Kwon *et al.*, 2014; Ferreira *et al.*, 2014; Mei *et al.*, 2013).

Summary

It is clear that the multitude of RNA species within our cells, and indeed, outside of our cells, have an intricate, elegant and complicated mode of communication (Tay *et al.*, 2014). Understanding the function and implications for disease of these small RNAs significantly lags behind their rate of

discovery. Furthermore, the emerging cross talk in the RNA world demonstrates an almost overwhelming level of complexity. Even among protein-coding genes, the mRNA transcripts that code for protein undergo significant shortening at the 3'UTR end in proliferating cells and in cancer cells, thereby creating a mechanism to escape miRNA-mediated repression and increases in the bioavailability of miRNAs (Sandberg *et al.*, 2008; Mayr and Bartel, 2009). Understanding the matrix of RNA communication will not only further our understanding of gene regulation, but given the knowledge developed over the last two decades on miRNAs and their contribution to health and disease, it is likely that such knowledge will further our understanding of diseases and inform new ways to both diagnose and treat disease more effectively.

See also: Cell Division/Death: Regulation of Cell Growth: Targeted mRNA Degradation. Nucleic acid Synthesis/Breakdown: Transcription: Distant Activation of Transcription by Enhancers; Pre-mRNA Splicing: Function and Dysfunction; The Spliceosome and Pre-mRNA Splicing. Protein Synthesis/Degradation: Translation: Components, Initiation, Elongation, Termination, and Regulation

References

- Abelson, J.F., Kwan, K.Y., O'Roak, B.J., *et al.*, 2005. Sequence variants in SLITRK1 are associated with Tourette's syndrome. *Science* 310, 317–320.
- Aleman, L.M., Doench, J., Sharp, P.A., 2007. Comparison of siRNA-induced off-target RNA and protein effects. *RNA* 13, 385–395.
- Ambs, S., Prueitt, R.L., Yi, M., *et al.*, 2008. Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. *Cancer Research* 68, 6162–6170.
- Aravin, A., Tuschl, T., 2005. Identification and characterization of small RNAs involved in RNA silencing. *FEBS Letters* 579, 5830–5840.
- Bader, A.G., Brown, D., Winkler, M., 2010. The promise of microRNA replacement therapy. *Cancer Research* 70, 7027–7030.
- Banerji, J., Rusconi, S., Schaffner, W., 1981. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27, 299–308.
- Bartel, B., 2005. MicroRNAs directing siRNA biogenesis. *Nature Structural & Molecular Biology* 12, 569–571.
- Bartel, D.P., 2004. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297.
- Bentwich, I., Avniel, A., Karov, Y., *et al.*, 2005. Identification of hundreds of conserved and nonconserved human microRNAs. *Nature Genetics* 37, 766–770.
- Bierhoff, H., Schmitz, K., Maass, F., Ye, J., Grummt, I., 2010. Noncoding transcripts in sense and antisense orientation regulate the epigenetic state of ribosomal RNA genes. *Cold Spring Harbor Symposia on Quantitative Biology* 75, 357–364.
- Bizuyehu, T.T., Lanes, C.F., Furmanek, T., *et al.*, 2012. Differential expression patterns of conserved miRNAs and isomiRs during Atlantic halibut development. *BMC Genomics* 13, 11.
- Bonnet, D., Dick, J.E., 1997. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature Medicine* 3, 730–737.
- Borchert, G.M., Lanier, W., Davidson, B.L., 2006. RNA polymerase III transcribes human microRNAs. *Nature Structural & Molecular Biology* 13, 1097–1101.
- Brockdorff, N., Ashworth, A., Kay, G.F., *et al.*, 1991. Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* 351, 329–331.
- Brockdorff, N., Ashworth, A., Kay, G.F., *et al.*, 1992. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* 71, 515–526.

- Burroughs, A.M., Ando, Y., De Hoon, M.J., *et al.*, 2010. A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. *Genome Research* 20, 1398–1410.
- Burroughs, A.M., Ando, Y., Hoon, D.E., *et al.*, 2011. Deep-sequencing of human Argonaute-associated small RNAs provides insight into miRNA sorting and reveals Argonaute association with RNA fragments of diverse origin. *RNA Biology* 8, 158–177.
- Cai, X., Hagedorn, C.H., Cullen, B.R., 2004. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA* 10, 1957–1966.
- Calin, G.A., Croce, C.M., 2006. MicroRNA signatures in human cancers. *Nature Reviews Cancer* 6, 857–866.
- Calin, G.A., Dumitru, C.D., Shimizu, M., *et al.*, 2002. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proceedings of the National Academy of Sciences of the United States of America* 99, 15524–15529.
- Calin, G.A., Ferracin, M., Cimmino, A., *et al.*, 2005. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *New England Journal of Medicine* 353, 1793–1801.
- Calin, G.A., Sevignani, C., Dumitru, C.D., *et al.*, 2004. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proceedings of the National Academy of Sciences of the United States of America* 101, 2999–3004.
- Cazalla, D., Yario, T., Steitz, J.A., 2010. Down-regulation of a host microRNA by a Herpesvirus saimiri noncoding RNA. *Science* 328, 1563–1566.
- Chan, Y.T., Lin, Y.C., Lin, R.J., *et al.*, 2013. Concordant and discordant regulation of target genes by miR-31 and its isoforms. *PLoS One* 8, e58169.
- Chen, K., Rajewsky, N., 2006. Natural selection on human microRNA binding sites inferred from SNP data. *Nature Genetics* 38, 1452–1456.
- Chen, Z., Xu, L., Ye, X., *et al.*, 2013. Polymorphisms of microRNA sequences or binding sites and lung cancer: A meta-analysis and systematic review. *PLoS One* 8, e61008.
- Chin, L.J., Ratner, E., Leng, S., *et al.*, 2008. A SNP in a let-7 microRNA complementary site in the KRAS 3' untranslated region increases non-small cell lung cancer risk. *Cancer Research* 68, 8535–8540.
- Christensen, B.C., Moyer, B.J., Avissar, M., *et al.*, 2009. A let-7 microRNA-binding site polymorphism in the KRAS 3' UTR is associated with reduced survival in oral cancers. *Carcinogenesis* 30, 1003–1007.
- Clague, J., Lippman, S.M., Yang, H., *et al.*, 2010. Genetic variation in MicroRNA genes and risk of oral premalignant lesions. *Molecular Carcinogenesis* 49, 183–189.
- Cortez, M.A., Bueso-Ramos, C., Ferdin, J., *et al.*, 2011. MicroRNAs in body fluids – The mix of hormones and biomarkers. *Nature Reviews Clinical Oncology* 8, 467–477.
- Costinean, S., Zanesi, N., Pekarsky, Y., *et al.*, 2006. Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* 103, 7024–7029.
- De Santa, F., Barozzi, I., Miettinen, F., *et al.*, 2010. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biology* 8, e1000384.
- Di Leva, G., Garofalo, M., Croce, C.M., 2014. MicroRNAs in cancer. *Annual Review of Pathology* 9, 287–314.
- Ehhardt, H.A., Tsang, H.H., Dai, D.C., *et al.*, 2009. Meta-analysis of small RNA-sequencing errors reveals ubiquitous post-transcriptional RNA modifications. *Nucleic Acids Research* 37, 2461–2470.
- Encode Project Consortium, 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 58–74.
- Esau, C., Davis, S., Murray, S.F., *et al.*, 2006. miR-122 regulation of lipid metabolism revealed by *in vivo* antisense targeting. *Cell Metabolism* 3, 87–98.
- Fabbri, M., Paone, A., Calore, F., *et al.*, 2012. MicroRNAs bind to Toll-like receptors to induce prometastatic inflammatory response. *Proceedings of the National Academy of Sciences of the United States of America* 109, E2110–E2116.
- Felli, N., Fontana, L., Pelosi, E., *et al.*, 2005. MicroRNAs 221 and 222 inhibit normal erythropoiesis and erythroleukemic cell growth via kit receptor down-modulation. *Proceedings of the National Academy of Sciences of the United States of America* 102, 18081–18086.
- Fernandez-Valverde, S.L., Taft, R.J., Mattick, J.S., 2010. Dynamic isomiR regulation in *Drosophila* development. *RNA* 16, 1881–1888.
- Ferreira, H.J., Heyn, H., Garcia Del Muro, X., *et al.*, 2014. Epigenetic loss of the PIWI/piRNA machinery in human testicular tumorigenesis. *Epigenetics* 9, 113–118.
- Friedman, R.C., Farh, K.K., Burge, C.B., Bartel, D.P., 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research* 19, 92–105.
- Fu, X., Han, Y., Wu, Y., *et al.*, 2011. Prognostic role of microRNA-21 in various carcinomas: A systematic review and meta-analysis. *European Journal of Clinical Investigation* 41, 1245–1253.
- Gallardo, E., Navarro, A., Vinolas, N., *et al.*, 2009. miR-34a as a prognostic marker of relapse in surgically resected non-small-cell lung cancer. *Carcinogenesis* 30, 1903–1909.
- Garofalo, M., Di Leva, G., Romano, G., *et al.*, 2009. miR-221&222 regulate TRAIL resistance and enhance tumorigenicity through PTEN and TIMP3 downregulation. *Cancer Cell* 16, 498–509.
- Girard, A., Sachidanandam, R., Hannon, G.J., Carmell, M.A., 2006. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* 442, 199–202.
- Godnic, I., Zorc, M., Jevsinek Skok, D., *et al.*, 2013. Genome-wide and species-wide *in silico* screening for intragenic MicroRNAs in human, mouse and chicken. *PLoS One* 8, e65165.
- Griffiths-Jones, S., Grocock, R.J., Dongen, V.A.N., Bateman, S., Enright, A. J., A., 2006. miRBase: MicroRNA sequences, targets and gene nomenclature. *Nucleic Acids Research* 34, D140–D144.
- Grimm, D., Streetz, K.L., Jopling, C.L., *et al.*, 2006. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature* 441, 537–541.
- Grimson, A., Farh, K.K., Johnston, W.K., *et al.*, 2007. MicroRNA targeting specificity in mammals: Determinants beyond seed pairing. *Molecular Cell* 27, 91–105.
- Gutschner, T., Diederichs, S., 2012. The hallmarks of cancer: A long non-coding RNA point of view. *RNA Biology* 9, 703–719.
- Hausser, J., Syed, A.P., Bilen, B., Zavolan, M., 2013. Analysis of CDS-located miRNA target sites suggests that they can effectively inhibit translation. *Genome Research* 23, 604–615.
- He, C., Li, Z., Chen, P., *et al.*, 2012. Young intragenic miRNAs are less coexpressed with host genes than old ones: Implications of miRNA-host gene coevolution. *Nucleic Acids Research* 40, 4002–4012.
- He, L., Thomson, J.M., Hemann, M.T., *et al.*, 2005. A microRNA polycistron as a potential human oncogene. *Nature* 435, 828–833.
- He, W.A., Calore, F., Londhe, P., *et al.*, 2014. Microvesicles containing miRNAs promote muscle cell death in cancer cachexia via TLR7. *Proceedings of the National Academy of Sciences of the United States of America* 111, 4525–4529.
- Heintzman, N.D., Stuart, R.K., Hon, G., *et al.*, 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature Genetics* 39, 311–318.
- Heo, I., Joo, C., Cho, J., *et al.*, 2008. Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. *Molecular Cell* 32, 276–284.
- Heo, I., Joo, C., Kim, Y.K., *et al.*, 2009. TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell* 138, 696–708.
- Hobert, O., 2004. Common logic of transcription factor and microRNA action. *Trends in Biochemical Sciences* 29, 462–468.
- Hon, L.S., Zhang, Z., 2007. The roles of binding site arrangement and combinatorial targeting in microRNA repression of gene expression. *Genome Biology* 8, R166.
- Jacob, F., Monod, J., 1961. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* 3, 318–356.
- Jazdzewski, K., Liyanarachchi, S., Swierniak, M., *et al.*, 2009. Polymorphic mature microRNAs from passenger strand of pre-miR-146a contribute to thyroid cancer. *Proceedings of the National Academy of Sciences of the United States of America* 106, 1502–1505.
- Jazdzewski, K., Murray, E.L., Franssila, K., *et al.*, 2008. Common SNP in pre-miR-146a decreases mature miR expression and predisposes to papillary thyroid carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* 105, 7269–7274.
- Ji, J., Yamashita, T., Budhu, A., *et al.*, 2009. Identification of microRNA-181 by genome-wide screening as a critical player in EpCAM-positive hepatic cancer stem cells. *Hepatology* 50, 472–480.
- John, B., Enright, A.J., Aravin, A., *et al.*, 2004. Human MicroRNA targets. *PLoS Biology* 2, e363.
- Johnson, S.M., Grosshans, H., Shingara, J., *et al.*, 2005. RAS is regulated by the let-7 microRNA family. *Cell* 120, 635–647.
- Jopling, C.L., Yi, M., Lancaster, A.M., Lemon, S.M., Sarnow, P., 2005. Modulation of hepatitis C virus RNA abundance by a liver-specific MicroRNA. *Science* 309, 1577–1581.
- Jung, E.J., Santarpia, L., Kim, J., *et al.*, 2012. Plasma microRNA 210 levels correlate with sensitivity to trastuzumab and tumor presence in breast cancer patients. *Cancer* 118, 2603–2614.
- Juvvuna, P.K., Khandelja, P., Lee, L.M., Makeyev, E.V., 2012. Argonaute identity defines the length of mature mammalian microRNAs. *Nucleic Acids Research* 40, 6808–6820.

- Kai, Z.S., Pasquinelli, A.E., 2010. MicroRNA assassins: Factors that regulate the disappearance of miRNAs. *Nature Structural & Molecular Biology* 17, 5–10.
- Kim, K., Jutooru, I., Chadalapaka, G., *et al.*, 2013. HOTAIR is a negative prognostic factor and exhibits pro-oncogenic activity in pancreatic cancer. *Oncogene* 32, 1616–1625.
- Kim, T.K., Hemberg, M., Gray, J.M., *et al.*, 2010. Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182–187.
- Kim, V.N., 2005. MicroRNA biogenesis: Coordinated cropping and dicing. *Nature Reviews Molecular Cell Biology* 6, 376–385.
- Kim, V.N., Han, J., Siomi, M.C., 2009. Biogenesis of small RNAs in animals. *Nature Reviews Molecular Cell Biology* 10, 126–139.
- Klattenhoff, C., Bratu, D.P., McGinnis-Schultz, N., *et al.*, 2007. Drosophila rasiRNA pathway mutations disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response. *Developmental Cell* 12, 45–55.
- Koch, F., Fenouil, R., Gut, M., *et al.*, 2011. Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. *Nature Structural & Molecular Biology* 18, 956–963.
- Kota, J., Chivukula, R.R., O'donnell, K.A., *et al.*, 2009. Therapeutic microRNA delivery suppresses tumorigenesis in a murine liver cancer model. *Cell* 137, 1005–1017.
- Kozomara, A., Griffiths-Jones, S., 2014. miRBase: Annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research* 42, D68–D73.
- Krek, A., Grun, D., Poy, M.N., *et al.*, 2005. Combinatorial microRNA target predictions. *Nature Genetics* 37, 495–500.
- Krutzfeldt, J., Kuwajima, S., Braich, R., *et al.*, 2007. Specificity, duplex degradation and subcellular localization of antagomirs. *Nucleic Acids Research* 35, 2885–2892.
- Krutzfeldt, J., Rajewsky, N., Braich, R., *et al.*, 2005. Silencing of microRNAs *in vivo* with 'antagomirs'. *Nature* 438, 685–689.
- Kuchenbauer, F., Morin, R.D., Argiropoulos, B., *et al.*, 2008. In-depth characterization of the microRNA transcriptome in a leukemia progression model. *Genome Research* 18, 1787–1797.
- Kurashige, J., Kamohara, H., Watanabe, M., *et al.*, 2012. Serum microRNA-21 is a novel biomarker in patients with esophageal squamous cell carcinoma. *Journal of Surgical Oncology* 106, 188–192.
- Kwon, C., Tak, H., Rho, M., *et al.*, 2014. Detection of PIWI and piRNAs in the mitochondria of mammalian cancer cells. *Biochemical and Biophysical Research Communications* 446, 218–223.
- Ladewig, E., Okamura, K., Flynt, A.S., Westholm, J.O., Lai, E.C., 2012. Discovery of hundreds of mirtrons in mouse and human small RNA data. *Genome Research* 22, 1634–1645.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., Tuschl, T., 2001. Identification of novel genes coding for small expressed RNAs. *Science* 294, 853–858.
- Lagos-Quintana, M., Rauhut, R., Yalcin, A., *et al.*, 2002. Identification of tissue-specific microRNAs from mouse. *Current Biology* 12, 735–739.
- Landgraf, P., Rusu, M., Sheridan, R., *et al.*, 2007. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* 129, 1401–1414.
- Landi, M.T., Zhao, Y., Rotunno, M., *et al.*, 2010. MicroRNA expression differentiates histology and predicts survival of lung cancer. *Clinical Cancer Research* 16, 430–441.
- Lanford, R.E., Hildebrandt-Eriksen, E.S., Petri, A., *et al.*, 2010. Therapeutic silencing of microRNA-122 in primates with chronic hepatitis C virus infection. *Science* 327, 198–201.
- Lau, N.C., Lim, L.P., Weinstein, E.G., Bartel, D.P., 2001. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858–862.
- Le, H.B., Zhu, W.Y., Chen, D.D., *et al.*, 2012. Evaluation of dynamic change of serum miR-21 and miR-24 in pre- and post-operative lung carcinoma patients. *Medical Oncology* 29, 3190–3197.
- Lee, E., Iskrow, R., Yang, L., *et al.*, 2012. Landscape of somatic retrotransposition in human cancers. *Science* 337, 967–971.
- Lee, R.C., Ambros, V., 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294, 862–864.
- Lee, R.C., Feinbaum, R.L., Ambros, V., 1993. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843–854.
- Lee, Y., Kim, M., Han, J., *et al.*, 2004. MicroRNA genes are transcribed by RNA polymerase II. *EMBO Journal* 23, 4051–4060.
- Lettice, L.A., Heaney, S.J., Purdie, L.A., *et al.*, 2003. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Human Molecular Genetics* 12, 1725–1735.
- Leucci, E., Patella, F., Waage, J., *et al.*, 2013. microRNA-9 targets the long non-coding RNA MALAT1 for degradation in the nucleus. *Scientific Reports* 3, 2535.
- Levin, A.A., 1999. A review of the issues in the pharmacokinetics and toxicology of phosphorothioate antisense oligonucleotides. *Biochimica et Biophysica Acta* 1489, 69–84.
- Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., Burge, C.B., 2003. Prediction of mammalian microRNA targets. *Cell* 115, 787–798.
- Li, J., Yang, Z., Yu, B., Liu, J., Chen, X., 2005. Methylation protects miRNAs and siRNAs from a 3'-end uridylation activity in Arabidopsis. *Current Biology* 15, 1501–1507.
- Li, S.C., Liao, Y.L., Ho, M.R., *et al.*, 2012. miRNA arm selection and isomiR distribution in gastric cancer. *BMC Genomics* 13 (Suppl 1), S13.
- Li, W., Notani, D., Ma, Q., *et al.*, 2013. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498, 516–520.
- Lim, L.P., Lau, N.C., Garrett-Engle, P., *et al.*, 2005. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature* 433, 769–773.
- Lin, S.L., Chang, D., Ying, S.Y., 2005. Asymmetry of intronic pre-miRNA structures in functional RISC assembly. *Gene* 356, 32–38.
- Liu, J., Carmell, M.A., Rivas, F.V., *et al.*, 2004. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* 305, 1437–1441.
- Llorens, F., Banez-Coronel, M., Pantano, L., *et al.*, 2013a. A highly expressed miR-101 isomiR is a functional silencing small RNA. *BMC Genomics* 14, 104.
- Llorens, F., Hummel, M., Pantano, L., *et al.*, 2013b. Microarray and deep sequencing cross-platform analysis of the miRNome and isomiR variation in response to epidermal growth factor. *BMC Genomics* 14, 371.
- Loher, P., Rigoutsos, I., 2012. Interactive exploration of RNA22 microRNA target predictions. *Bioinformatics* 28, 3322–3323.
- Lu, J., Getz, G., Miska, E.A., *et al.*, 2005. MicroRNA expression profiles classify human cancers. *Nature* 435, 834–838.
- Lu, S., Sun, Y.H., Chiang, V.L., 2009. Adenylation of plant miRNAs. *Nucleic Acids Research* 37, 1878–1885.
- Luo, J.H., Ren, B., Keryanov, S., *et al.*, 2006. Transcriptomic and genomic analysis of human hepatocellular carcinomas and hepatoblastomas. *Hepatology* 44, 1012–1024.
- Makeyev, E.V., Zhang, J., Carrasco, M.A., Maniatis, T., 2007. The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. *Molecular Cell* 27, 435–448.
- Malone, C.D., Hannon, G.J., 2009. Small RNAs as guardians of the genome. *Cell* 136, 656–668.
- Maragkakis, M., Vergoulis, T., Alexiou, P., *et al.*, 2011. DIANA-microT web server upgrade supports fly and worm miRNA target prediction and bibliographic miRNA to disease association. *Nucleic Acids Research* 39, W145–W148.
- Marsico, A., Huska, M.R., Lasserre, J., *et al.*, 2013. PROMiRNA: A new miRNA promoter recognition method uncovers the complex regulation of intronic miRNAs. *Genome Biology* 14, R84.
- Marti, E., Pantano, L., Banez-Coronel, M., *et al.*, 2010. A myriad of miRNA variants in control and Huntington's disease brain regions detected by massively parallel sequencing. *Nucleic Acids Research* 38, 7219–7235.
- Mathe, E.A., Nguyen, G.H., Bowman, E.D., *et al.*, 2009. MicroRNA expression in squamous cell carcinoma and adenocarcinoma of the esophagus: Associations with survival. *Clinical Cancer Research* 15, 6192–6200.
- Mayr, C., Bartel, D.P., 2009. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell* 138, 673–684.
- McClintock, B., 1950. The origin and behavior of mutable loci in maize. *Proceedings of the National Academy of Sciences of the United States of America* 36, 344–355.
- Medina, P.P., Nolde, M., Slack, F.J., 2010. OncomiR addiction in an *in vivo* model of microRNA-21-induced pre-B-cell lymphoma. *Nature* 467, 86–90.
- Megraw, M., Pereira, F., Jensen, S.T., Ohler, U., Hatzigeorgiou, A.G., 2009. A transcription factor affinity-based code for mammalian transcription initiation. *Genome Research* 19, 644–656.
- Mei, Y., Clark, D., Mao, L., 2013. Novel dimensions of piRNAs in cancer. *Cancer Letters* 336, 46–52.
- Mitchell, P.S., Parkin, R.K., Kroh, E.M., *et al.*, 2008. Circulating microRNAs as stable blood-based markers for cancer detection. *Proceedings of the National Academy of Sciences of the United States of America* 105, 10513–10518.
- Mongroo, P.S., Rustgi, A.K., 2010. The role of the miR-200 family in epithelial-mesenchymal transition. *Cancer Biology & Therapy* 10, 219–222.
- Monteys, A.M., Spengler, R.M., Wan, J., *et al.*, 2010. Structure and activity of putative intronic miRNA promoters. *RNA — A Publication of the RNA Society* 16, 495–505.
- Morin, R.D., O'Connor, M.D., Griffith, M., *et al.*, 2008. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Research* 18, 610–621.

- Nair, V.S., Maeda, L.S., Ioannidis, J.P., 2012. Clinical outcome prediction by microRNAs in human cancer: A systematic review. *Journal of the National Cancer Institute* 104, 528–540.
- Ohyashiki, K., Umez, T., Yoshizawa, S., *et al.*, 2011. Clinical impact of down-regulated plasma miR-92a levels in non-Hodgkin's lymphoma. *PLoS One* 6, e16408.
- Orom, U.A., Nielsen, F.C., Lund, A.H., 2008. MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Molecular Cell* 30, 460–471.
- Oue, N., Anami, K., Schetter, A.J., *et al.*, 2014. High miR-21 expression from FFPE tissues is associated with poor survival and response to adjuvant chemotherapy in colon cancer. *International Journal of Cancer* 134, 1926–1934.
- Ozsolak, F., Poling, L.L., Wang, Z.X., *et al.*, 2008. Chromatin structure analyses identify miRNA promoters. *Genes & Development* 22, 3172–3183.
- Pantano, L., Estivill, X., Marti, E., 2010. SeqBuster, a bioinformatic tool for the processing and analysis of small RNAs datasets, reveals ubiquitous miRNA modifications in human embryonic cells. *Nucleic Acids Research* 38, e34.
- Parsons, P.A., Bodmer, W.F., 1961. The evolution of overdominance: Natural selection and heterozygote advantage. *Nature* 190, 7–12.
- Perdomo, C., Campbell, J.D., Gerrein, J., *et al.*, 2013. MicroRNA 4423 is a primate-specific regulator of airway epithelial cell differentiation and lung carcinogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 110, 18946–18951.
- Pharoah, P.D., Palmieri, R.T., Ramus, S.J., *et al.*, 2011. The role of KRAS rs61764370 in invasive epithelial ovarian cancer: Implications for clinical testing. *Clinical Cancer Research* 17, 3742–3750.
- Place, R.F., Li, L.C., Pookot, D., Noonan, E.J., Dahiya, R., 2008. MicroRNA-373 induces expression of genes with complementary promoter sequences. *Proceedings of the National Academy of Sciences of the United States of America* 105, 1608–1613.
- Poliseno, L., Salmena, L., Zhang, J., *et al.*, 2010. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 465, 1033–1038.
- Polytarchou, C., Iliopoulos, D., Struhl, K., 2012. An integrated transcriptional regulatory circuit that reinforces the breast cancer stem cell state. *Proceedings of the National Academy of Sciences of the United States of America* 109, 14470–14475.
- Pouliot, L.M., Chen, Y.C., Bai, J., *et al.*, 2012. Cisplatin sensitivity mediated by WEE1 and CHK1 is mediated by miR-155 and the miR-15 family. *Cancer Research* 72, 5945–5955.
- Pramanik, D., Campbell, N.R., Karikari, C., *et al.*, 2011. Restitution of tumor suppressor microRNAs using a systemic nanovector inhibits pancreatic cancer growth in mice. *Molecular Cancer Therapeutics* 10, 1470–1480.
- Pritchard, C.C., Kroh, E., Wood, B., *et al.*, 2012. Blood cell origin of circulating microRNAs: A cautionary note for cancer biomarker studies. *Cancer Prevention Research (Philadelphia, PA)* 5, 492–497.
- Raveche, E.S., Salerno, E., Scaglione, B.J., *et al.*, 2007. Abnormal microRNA-16 locus with synteny to human 13q14 linked to CLL in NZB mice. *Blood* 109, 5079–5086.
- Reid, J.G., Nagaraja, A.K., Lynn, F.C., *et al.*, 2008. Mouse let-7 miRNA populations exhibit RNA editing that is constrained in the 5'-seed/ cleavage/anchor regions and stabilize predicted mmu-let-7a:miRNA duplexes. *Genome Research* 18, 1571–1581.
- Reinhart, B.J., Slack, F.J., Basson, M., *et al.*, 2000. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901–906.
- Rinn, J.L., Chang, H.Y., 2012. Genome regulation by long noncoding RNAs. *Annual Review of Biochemistry* 81, 145–166.
- Robles, A.I., Harris, C.C., 2013. A primate-specific microRNA enters the lung cancer landscape. *Proceedings of the National Academy of Sciences of the United States of America* 110, 18748–18749.
- Roccaro, A.M., Sacco, A., Thompson, B., *et al.*, 2009. MicroRNAs 15a and 16 regulate tumor proliferation in multiple myeloma. *Blood* 113, 6669–6680.
- Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., Bradley, A., 2004. Identification of mammalian microRNA host genes and transcription units. *Genome Research* 14, 1902–1910.
- Ruby, J.G., Jan, C., Player, C., *et al.*, 2006. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* 127, 1193–1207.
- Ruzzo, A., Canestrari, E., Galluccio, N., *et al.*, 2011. Role of KRAS let-7 LCS6 SNP in metastatic colorectal cancer patients. *Annals of Oncology* 22, 234–235.
- Ryan, B.M., McClary, A.C., Valeri, N., *et al.*, 2012a. rs4919510 in hsa-mir-608 is associated with outcome but not risk of colorectal cancer. *PLoS One* 7, e36306.
- Ryan, B.M., Robles, A.I., Harris, C.C., 2010. Genetic variation in microRNA networks: The implications for cancer research. *Nature Reviews Cancer* 10, 389–402.
- Ryan, B.M., Robles, A.I., Harris, C.C., 2012b. KRAS-LCS6 genotype as a prognostic marker in early-stage CRC — letter. *Clinical Cancer Research* 18, 3487–3488. [author reply 3489.](#)
- Salmena, L., Poliseno, L., Tay, Y., Kats, L., Pandolfi, P.P., 2011. A ceRNA hypothesis: The Rosetta Stone of a hidden RNA language? *Cell* 146, 353–358.
- Sandberg, R., Neilson, J.R., Sarma, A., Sharp, P.A., Burge, C.B., 2008. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science* 320, 1643–1647.
- Saunders, M.A., Liang, H., Li, W.H., 2007. Human polymorphism at microRNAs and microRNA target sites. *Proceedings of the National Academy of Sciences of the United States of America* 104, 3300–3305.
- Schembri, F., Sridhar, S., Perdomo, C., *et al.*, 2009. MicroRNAs as modulators of smoking-induced gene expression changes in human airway epithelium. *Proceedings of the National Academy of Sciences of the United States of America* 106, 2319–2324.
- Schetter, A.J., Leung, S.Y., Sohn, J.J., *et al.*, 2008. MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *JAMA* 299, 425–436.
- Schmidt, L.H., Spieker, T., Koschmieder, S., *et al.*, 2011. The long noncoding MALAT-1 RNA indicates a poor prognosis in non-small cell lung cancer and induces migration and tumor growth. *Journal of Thoracic Oncology* 6, 1984–1992.
- Schwarzenbach, H., Nishida, N., Calin, G.A., Pantel, K., 2014. Clinical relevance of circulating cell-free microRNAs in cancer. *Nature Reviews Clinical Oncology* 11, 145–156.
- Seitz, H., Ghildiyal, M., Zamore, P.D., 2008. Argonaute loading improves the 5' precision of both MicroRNAs and their miRNA* strands in flies. *Current Biology* 18, 147–151.
- Sevignani, C., Calin, G.A., Nnadi, S.C., *et al.*, 2007. MicroRNA genes are frequently located near mouse cancer susceptibility loci. *Proceedings of the National Academy of Sciences of the United States of America* 104, 8017–8022.
- Shimono, Y., Zabala, M., Cho, R.W., *et al.*, 2009. Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. *Cell* 138, 592–603.
- Simon, M.D., Pinter, S.F., Fang, R., *et al.*, 2013. High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* 504, 465–469.
- Smits, K.M., Paranjape, T., Nallur, S., *et al.*, 2011. A let-7 microRNA SNP in the KRAS 3'UTR is prognostic in early-stage colorectal cancer. *Clinical Cancer Research* 17, 7723–7731.
- Sonnenburg, S., Zien, A., Ratsch, G., 2006. ARTS: Accurate recognition of transcription starts in human. *Bioinformatics* 22, E472–E480.
- Sun, Y., Wang, M., Lin, G., *et al.*, 2012. Serum microRNA-155 as a potential biomarker to track disease in breast cancer. *PLoS One* 7, e47003.
- Takamizawa, J., Konishi, H., Yanagisawa, K., *et al.*, 2004. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Research* 64, 3753–3756.
- Tang, J.Y., Lee, J.C., Chang, Y.T., *et al.*, 2013. Long noncoding RNAs-related diseases, cancers, and drugs. *Scientific World Journal* 2013, 943539.
- Tanzer, A., Stadler, P.F., 2004. Molecular evolution of a microRNA cluster. *Journal of Molecular Biology* 339, 327–335.
- Tay, Y., Rinn, J., Pandolfi, P.P., 2014. The multilayered complexity of ceRNA crosstalk and competition. *Nature* 505, 344–352.
- Tomimaru, Y., Eguchi, H., Nagano, H., *et al.*, 2010. MicroRNA-21 induces resistance to the anti-tumour effect of interferon-alpha/5-fluorouracil in hepatocellular carcinoma cells. *British Journal of Cancer* 103, 1617–1626.
- Trang, P., Medina, P.P., Wiggins, J.F., *et al.*, 2010. Regression of murine lung tumors by the let-7 microRNA. *Oncogene* 29, 1580–1587.
- Valeri, N., Gasparini, P., Braconi, C., *et al.*, 2010. MicroRNA-21 induces resistance to 5-fluorouracil by down-regulating human DNA MutS homolog 2 (hMSH2). *Proceedings of the National Academy of Sciences of the United States of America* 107, 21098–21103.
- Vasudevan, S., Tong, Y., Steitz, J.A., 2007. Switching from repression to activation: microRNAs can up-regulate translation. *Science* 318, 1931–1934.
- Volders, P.J., Helsens, K., Wang, X., *et al.*, 2013. LNCipedia: A database for annotated human lncRNA transcript sequences and structures. *Nucleic Acids Research* 41, D246–D251.
- Volinia, S., Calin, G.A., Liu, C.G., *et al.*, 2006. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proceedings of the National Academy of Sciences of the United States of America* 103, 2257–2261.
- Waidner, L.A., Morgan, R.W., Anderson, A.S., *et al.*, 2009. MicroRNAs of Gallid and Meleagrid herpesviruses show generally conserved genomic locations and are virus-specific. *Virology* 388, 128–136.

- Wakiyama, M., Takimoto, K., Ohara, O., Yokoyama, S., 2007. Let-7 microRNA-mediated mRNA deadenylation and translational repression in a mammalian cell-free system. *Genes & Development* 21, 1857–1862.
- Wang, G., Reinke, V., 2008. A *C. elegans* Piwi, PRG-1, regulates 21U-RNAs during spermatogenesis. *Current Biology* 18, 861–867.
- Wang, G., Wang, W., Gao, W., Lv, J., Fang, J., 2014. Two functional polymorphisms in microRNAs and lung cancer risk: A meta-analysis. *Tumour Biology* 35, 2693–2699.
- Wang, P., Zhuang, L., Zhang, J., *et al.*, 2013. The serum miR-21 level serves as a predictor for the chemosensitivity of advanced pancreatic cancer, and miR-21 expression confers chemoresistance by targeting FasL. *Molecular Oncology* 7, 334–345.
- Wang, X.W., Xuan, Z.Y., Zhao, X.Y., Li, Y.D., Zhang, M.Q., 2009. High-resolution human core-promoter prediction with CoreBoost_HM. *Genome Research* 19, 266–275.
- Wiggins, J.F., Ruffino, L., Kelnar, K., *et al.*, 2010. Development of a lung cancer therapeutic based on the tumor suppressor microRNA-34. *Cancer Research* 70, 5923–5930.
- Wightman, B., Ha, I., Ruvkun, G., 1993. Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855–862.
- Winter, J., Jung, S., Keller, S., Gregory, R.I., Diederichs, S., 2009. Many roads to maturity: MicroRNA biogenesis pathways and their regulation. *Nature Cell Biology* 11, 228–234.
- Wright, M.W., Bruford, E.A., 2011. Naming ‘junk’: Human non-protein coding RNA (ncRNA) gene nomenclature. *Human Genomics* 5, 90–98.
- Wu, H., Neilson, J.R., Kumar, P., *et al.*, 2007. miRNA profiling of naive, effector and memory CD8 T cells. *PLoS One* 2, e1020.
- Wyman, S.K., Knouf, E.C., Parkin, R.K., *et al.*, 2011. Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. *Genome Research* 21, 1450–1461.
- Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J.A., Kosik, K.S., 2009. MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell* 137, 647–658.
- Yan, H., Wu, Q.L., Sun, C.Y., *et al.*, 2014. piRNA-823 contributes to tumorigenesis by regulating de novo DNA methylation and angiogenesis in multiple myeloma. *Leukemia*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24732595> (accessed 27.11.14).
- Yanaihara, N., Caplen, N., Bowman, E., *et al.*, 2006. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* 9, 189–198.
- Yang, H., Dinney, C.P., Ye, Y., *et al.*, 2008. Evaluation of genetic variants in microRNA-related genes and risk of bladder cancer. *Cancer Research* 68, 2530–2537.
- Ye, Y., Wang, K.K., Gu, J., *et al.*, 2008. Genetic variations in microRNA-related genes are novel susceptibility loci for esophageal cancer risk. *Cancer Prevention Research (Philadelphia, PA)* 1, 460–469.
- Zanetti, K.A., Haznadar, M., Welsh, J.A., *et al.*, 2012. 3'-UTR and functional secretor haplotypes in mannose-binding lectin 2 are associated with increased colon cancer risk in African Americans. *Cancer Research* 72, 1467–1477.
- Zhang, H., Gao, S., De Geyter, C., 2009. A natural antisense transcript, BOKAS, regulates the pro-apoptotic activity of human Bok. *International Journal of Oncology* 34, 1135–1138.
- Zhang, H.L., Yang, L.F., Zhu, Y., *et al.*, 2011. Serum miRNA-21: Elevated levels in patients with metastatic hormone-refractory prostate cancer and potential predictive factor for the efficacy of docetaxel-based chemotherapy. *Prostate* 71, 326–331.
- Zhang, M.W., Jin, M.J., Yu, Y.X., *et al.*, 2012. Associations of lifestyle-related factors, hsa-miR-149 and hsa-miR-605 gene polymorphisms with gastrointestinal cancer risk. *Molecular Carcinogenesis* 51 (Suppl 1), E21–E31.
- Zheng, D., Haddadin, S., Wang, Y., *et al.*, 2011. Plasma microRNAs as novel biomarkers for early detection of lung cancer. *International Journal of Clinical and Experimental Pathology* 4, 575–586.
- Zheng, J., Deng, J., Xiao, M., *et al.*, 2013. A sequence polymorphism in miR-608 predicts recurrence after radiotherapy for nasopharyngeal carcinoma. *Cancer Research* 73, 5151–5162.

Relevant Websites

<http://genome.ucsc.edu/ENCODE/>
 Encyclopedia of DNA Elements at UCSC.
<http://www.integratomics-time.com/miR-host/catalog>
 Integratomics TIME.
<http://www.mirbase.org/>
 miRBase.
<http://www.targetscan.org/>
 TargetScanHuman.
<http://cancergenome.nih.gov/>
 The Cancer Genome Atlas.

Riboswitches and Ribozymes

MWL Lau and AR Ferré-D'Amaré, National Heart, Lung and Blood Institute, Bethesda, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Aptamer An RNA or DNA molecule that has been evolved to bind specifically to a small molecule, another nucleic acid, or a protein.

Ribonucleoprotein (RNP) Supramolecular complex formed by association of RNA and protein molecules.

Riboswitch Domain in a messenger RNA (mRNA) that regulates the expression of the mRNA in response to specific binding of a regulatory small molecule.

Ribozyme An RNA molecule that functions as an enzyme and catalyzes specific biochemical reactions.

RNA world hypothesis A proposition that at an early stage in the evolution of life, RNA molecules served as both,

repositories of genetic information, and biochemical catalysts.

Systematic evolution of ligands by exponential enrichment (SELEX) A technique for isolating DNA or RNA sequences with functional properties from a population of nucleic acid molecules.

Viroid Plant pathogens that consist solely of a short, circular, single-stranded RNA.

Virusoid Infectious agents composed of circular, single-stranded RNAs that require assistance from other viruses for replication and propagation.

Evolutionary Importance of RNA

Evolution of all organisms, from unicellular archaea and bacteria to multicellular eukaryotes, requires the inheritance of genetic information. In modern cell-based biology, this information is encoded in DNA and directs the synthesis of cellular proteins through RNA (many viruses use RNA as the repository for genetic information). Although RNA was thought, early on, to act primarily as a conduit of information between DNA and proteins, already in 1961 Jacob and Monod proposed that some RNAs might function directly in genetic regulation by sensing the intracellular concentration of small molecule metabolites (Jacob and Monod, 1961). The description of the secondary structure of tRNA by Holley and colleagues in 1965 suggested that RNAs can adopt complex three-dimensional structures resembling those of proteins (Holley *et al.*, 1965). Such proposals, and central roles played by RNA in the synthesis of proteins in contemporary cells, led Crick and Orgel to hypothesize in 1968 that at an early stage in the evolution of life, RNAs could have served both as carriers of genetic information and as genetically encoded catalysts (Orgel, 1968; Crick, 1968).

Catalytic RNAs: Ribozymes

Important support for this 'RNA world hypothesis' came from the discovery in the early 1980's of enzymes made entirely of RNA, or ribozymes (Table 1). The first two classes of ribozymes described were the self-splicing 'group I' introns (Kruger *et al.*, 1982), and the tRNA precursor processing enzyme, ribonuclease (RNase) P (Guerrier-Takada *et al.*, 1983). The ability of RNA to catalyze chemical reactions provided strong evidence that some RNAs, just like protein enzymes, can adopt complex three-dimensional structures that can precisely position reactants and catalytic functional groups, and preferentially stabilize transition states. Since the initial discovery of ribozymes, much work has been devoted to examining the

ability of RNAs to catalyze metabolic reactions, both in contemporary cells and in putative pre-biotic metabolisms.

In Vitro Selection of Aptamers and Ribozymes

A powerful technique for the discovery and characterization RNAs that have intrinsic biochemical functions is *in vitro* selection, or SELEX. In the late 1980s, the groups of Gold, Joyce and Szostak invented this methodology, which consists of subjecting a pool of RNAs of diverse sequence (pools comprising as many as $\sim 10^{15}$ different sequences are employed) to a selective pressure, and amplifying the most active species (Joyce, 1989; Ellington and Szostak, 1990; Tuerk and Gold, 1990). The process is iterated until a handful of RNAs that are very efficient in the activity for which they are being selected is obtained. This process, which mimics Darwinian evolution, was initially used to obtain RNA molecules capable of high affinity and specificity binding to small molecules. Such RNAs were named 'aptamers.' Since then, SELEX has been used to isolate many different aptamers, which bind selectively to a variety of ligands ranging from simple ions to large proteins. In addition, SELEX has been used to isolate ribozymes with new activities. Unlike catalytic antibodies, which are typically selected for tight binding to transition-state mimics, SELEX for ribozymes has been most successful by searching for RNA molecules capable of catalyzing the desired chemical transformation on themselves, or a covalently bound substrate. *In vitro* selection has led to the isolation of RNAs capable of catalyzing a broad range of chemical transformations, greatly expanding the known chemical versatility of this nucleic acid (Table 2).

From Aptamers to Riboswitches

Even though the sequence pools employed for SELEX are very complex, they typically only cover a minuscule fraction of the

Table 1 Natural ribozymes

Ribozyme	Bond formation	Reaction	Functional groups in reaction	k_{obs} (min^{-1})	References
Hammer-head	P–O	Transesterification (self-cleavage)	2'-OH of RNA, internal RNA phosphate	~ 10 (~ 200 at 200 mM Mg^{2+})	Canny <i>et al.</i> (2004)
HDV/CPEB3	P–O	Transesterification (self-cleavage)	2'-OH of RNA, internal RNA phosphate	$> 60/0.01$	Thill <i>et al.</i> (1993)
Hairpin	P–O	Transesterification (self-cleavage)	2'-OH of RNA, internal RNA phosphate	~ 0.5	Hegg and Fedor (1995)
VS	P–O	Transesterification (self-cleavage)	2'-OH of RNA, internal RNA phosphate	~ 1 (~ 600 for mutant)	Zamel <i>et al.</i> (2004)
<i>glmS</i> ribozyme	P–O	Transesterification (self-cleavage)	2'-OH of RNA, internal RNA phosphate	~ 70	Lau and Ferre-D'Amare (2013)
Twister	P–O	Transesterification (self-cleavage)	Internal 2'-OH of RNA, internal RNA phosphate	Est. ~ 1000	Roth <i>et al.</i> (2014)
CoTC	P–O	Transesterification (self-cleavage)	Internal 2'-OH of RNA, internal RNA phosphate	~ 0.015	Teixeira <i>et al.</i> (2004)
Ribosome	C–N	Peptidyl transfer	α -amine of amino acid, carbonyl of peptidyl-RNA	up to 3000	Katunin <i>et al.</i> (2002)
Group I intron	P–O	Splicing (2 steps)	2'-OH of internal RNA adenosine (or OH ⁻), internal RNA phosphate, 3'-OH of RNA	~ 1	Bass and Cech (1984)
Group II intron	P–O	Splicing (2 steps)	2'-OH of guanosine (or OH ⁻), internal RNA phosphate, 3'-OH of RNA	~ 0.1	Xiang <i>et al.</i> (1998)
GIR 1	P–O	Transesterification (branching)	Internal 2'-OH of RNA, internal RNA phosphate	~ 0.3	Tang <i>et al.</i> (2011)
RNAse P	P–O	Transesterification (cleavage)	OH ⁻ ion (proposed), internal RNA phosphate	~ 300	Beebe and Fierke (1994)

Table 2 Examples of artificially selected ribozymes

Ribozyme	Bond formation	Reaction	Functional groups in reaction	k_{obs} (min^{-1})	References
Class I Ligase	P–O	Ligation	3'-OH of RNA, 5' α -phosphate of RNA	60 (~ 800 at pH 9)	Glasner <i>et al.</i> (2002)
Class II Kin.25	P–O	Phosphorylation	5'-OH of RNA, 5' γ -thio-phosphate of γ -thio-ATP	~ 0.3	Lorsch and Szostak (1994)
Iso6	P–O	Capping	5' β -phosphate of GDP, 5' α -phosphate of RNA	0.08	Huang and Yarus (1997)
Leadzyme	P–O	Transesterification, Hydrolysis (2 steps)	2'-OH of RNA, internal RNA phosphate	~ 0.5	Pan and Uhlenbeck (1992)
CoES7	P–O	Adenylation	Phosphate of NMN/FMN / 4'-phosphopantetheine, 5' α -phosphate of RNA	0.060	Huang <i>et al.</i> (2000)
DA-22	C–C	Diels–Alder	C–C double bond of maleimide, C–C double bond of anthracene	~ 0.7	Seelig and Jäschke (1999)
Clone 10	C–C	Claisen condensation	Acyl group of acyl-coenzyme A, carboxyl group of malonyl	N/A	Ryu <i>et al.</i> (2006)
Pre-24	C–O	Aminoacylation	3'-OH of tRNA, carbonyl of amino acids	0.15	Saito <i>et al.</i> (2001)
RA, MA	C–N	Glycosidic bond formation	N7 of 6-thioguanine, C1 of 5-phosphoribosyl 1-pyrophosphate	0.02	Lau <i>et al.</i> (2004)
pR1	C–N	Schiff base	N7 of 6-thioguanine, aldehyde of ribose-5-phosphate	0.001	Lau and Unrau (2009)
UV5	C–S	Michael addition	Sulfhydryl group, C–C double bond of fumaramide	0.01	Sengle <i>et al.</i> (2001)

total possible number of sequences for a given RNA length. For instance, for a 50-nucleotide long RNA, there are $\sim 1.3 \times 10^{30}$ possible unique sequences (4^{50}). Nonetheless, aptamers

targeting a variety of ligands can be found quite readily. This implies that biochemically active RNA species are relatively common in RNA 'sequence space.' In turn, such abundance

suggests that aptamer-like RNAs may exist in contemporary cells. Indeed, the T-box, a first example of such an RNA, was discovered in 1993 (Grundy and Henkin, 1993). This is a structured domain that lies in the 5' untranslated region (UTR) of mRNAs encoding aminoacyl-tRNA synthetases (ARSs) and is conserved across Gram-positive bacteria (Miranda-Ríos *et al.*, 2001). T-boxes bind sequence-specifically to a tRNA (engaging its anticodon), and control the expression of the mRNA they are part of depending on whether the tRNA is charged with its cognate amino acid (i.e., aminoacylated) or not. In this way, this cellular aptamer can respond with a gene expression decision to starvation for a particular amino acid. Many naturally occurring aptamer-like domains of mRNA that control gene expression *in-cis* by directly binding to their cognate intracellular ligands have been discovered in the present century (Mironov *et al.*, 2002; Nahvi *et al.*, 2002; Serganov and Nudler, 2013). Such RNA genetic switches, or 'riboswitches' are now known to respond selectively to simple ions (fluoride and magnesium), amino acids (glycine and lysine), vitamins and other enzymatic cofactors (adenosine cobalamine, flavin mononucleotide, thiamine pyrophosphate (TPP) and S-adenosyl methionine (SAM)), second messengers (cyclic-di-GMP), etc. For some ligands such as SAM, riboswitches appear to have evolved multiple times, independently; sequence and crystallographic analysis reveals the existence of at least four structurally distinct classes of SAM riboswitches. At the time of writing, there are 28 known riboswitch classes that collectively respond to 22 different ligands (Table 3). Although examples of the TPP riboswitch have been found to function in plants, fungi and archaea, the vast majority of currently characterized riboswitches appear to be exclusively bacterial. Whether this distribution is a true indication of the under-representation of gene-regulatory aptamers in eukarya, or reflects biases in the bioinformatic, genetic and biochemical methodology employed thus far to discover riboswitches, remains to be established.

Control of Gene Expression by Riboswitches

By definition, all riboswitches are capable of high specificity binding to their cognate ligand. However, the mechanisms by which they transduce the initial binding event into a gene-regulatory decision vary, even among different examples of riboswitches from the same structural class. At present, riboswitches are known to control expression of genes in five different ways (Figure 1): (1) Many bacterial riboswitches function by adopting alternate conformations that result in the stabilization or destabilization of an intrinsic (p -independent) transcriptional terminator (Serganov *et al.*, 2008). Such terminators consist of a stable stem-loop followed by several uridine residues, and riboswitches can, for instance, form a structure that sequesters one of the strands of the terminator stem-loop into their ligand binding-stabilized aptamer domain structure. In such an example, the riboswitch would function as a transcriptional 'on' switch. Transcriptional 'off' switches are also legion. (2) Riboswitch-mediated transcriptional termination, as just described, gives rise to a truncated transcript encompassed by the transcriptional initiation and termination sites. In some instances, these short RNAs have

been found to function in *trans*-, that is, by binding to other mRNAs and controlling their expression, analogous to the way that bacterial sRNA or eukaryal miRNA function (Loh *et al.*, 2010). (3) Riboswitches can adopt alternate conformations as a function of ligand binding that expose or occlude the Shine-Dalgarno (ribosome-binding) site of the mRNA of which they are part (Serganov *et al.*, 2009). Depending on which conformation of the 5'-UTR is stabilized by ligand binding, such riboswitches can be translational 'on' or 'off' switches. (4) The *glmS* ribozyme-riboswitch is a catalytic RNA that cleaves the 5'-UTR of the mRNA it is part of in response to ligand binding (Klein and Ferre-D'Amare, 2006). This cleavage leads to rapid degradation of the mRNA. Thus, this is one example where gene expression is regulated at the level of mRNA decay by a riboswitch. (5) Several examples of plant and fungal TPP riboswitches are known to control alternative pre-mRNA splicing, by exposing splice enhancer sequences as a function of ligand binding (Li and Breaker, 2013). Some bacterial Class-II cyclic-di-GMP riboswitches also function by modulating alternative splicing, in this case, of a self-splicing group I intron with which they are tightly coupled structurally. This mechanistic diversity (likely to be expanded by future discoveries), coupled with the versatility of RNA in specifically recognizing diverse ligands, underpins the widespread use of riboswitches as genetic regulators.

Ribozymes are Ubiquitous in Nature

Ribozymes are found in all living organisms, where they catalyze peptide bond formation or phosphoryl transfer reactions. Crick speculated that the primordial ribosome would have been a catalytic RNA (Crick, 1968), and biochemical and structural analyses in the past twenty years have demonstrated that the enzymatic active site of modern ribosomes is indeed constructed exclusively of RNA. That is, the ribosome is a ribozyme (Ban *et al.*, 2000). Many eukaryotic genes are interrupted by noncoding elements called introns. Some introns harbor catalytic activity, and are capable of self-splicing. Two major structural classes of such introns (group I and group II) have been extensively characterized. Some of these have been demonstrated to be selfish mobile genetic elements. The biochemical mechanism and structure of group II introns resembles those of the RNA components of the spliceosome, the *trans*-acting ribonucleoprotein (RNP) that excises most introns in eukaryotes. Thus, it is thought that the active site of this large RNP, like that of the ribosome, consists exclusively of RNA. Further structural studies will be needed to confirm this hypothesis. RNase P is responsible for excising the universal 5' leader sequence that must be removed for tRNA maturation. In virtually all organisms, RNase P is an RNP with an all-RNA active site. At least six structurally distinct classes of RNA-cleaving ribozymes that share an internal transesterification mechanism have been described, and these play key roles in gene regulation in bacteria, in eukaryotes, as well as in subgenomic elements such as viroids and virusoids. Despite their wide distribution and abundance (the ribosome is the most abundant particle, by mass, in an actively growing yeast), naturally occurring ribozymes are known to catalyze only a limited subset of the reactions that RNA has been shown to be

Table 3 Natural riboswitches

Riboswitch (organism)	Ligand bound	K _d (nM)	Tertiary structure	Accession PDB code	References
M-box riboswitch (<i>Bacillus subtilis</i>)	Magnesium ion	$\sim 2 \times 10^6$	3 parallel helices, 2 coaxial stacks	2QBZ	Dann <i>et al.</i> (2007)
Fluoride riboswitch (<i>Thermotoga petrophila</i>)	Fluoride ion	$\sim 1.5 \times 10^5$	2 helical stems, 1 pseudoknot	4ENC	Ren <i>et al.</i> (2012)
Moco riboswitch (<i>Escherichia coli</i>)	Molybdenum	N/A	N/A	N/A	Regulski <i>et al.</i> (2008)
Tuco riboswitch (<i>E. coli</i>)	Tungsten	N/A	N/A	N/A	Regulski <i>et al.</i> (2008)
THF riboswitch (<i>Streptococcus mutans</i>)	Tetrahydrofolate/7-deazaguanine	1.8×10^4 / 2.5×10^4	5 helices, 2 coaxial stacks, a 3-way junction, 1 pseudoknot	3SD1	Trausch <i>et al.</i> (2011)
add A-riboswitch (<i>Vibrio vulnificus</i>)	Adenine/2,6-diaminopurine	300/10	3 helices, a 3-way junction	1Y26	Serganov <i>et al.</i> (2004)
xpt G-riboswitch (<i>B. subtilis</i>)	Guanine/hypoxanthine	5/50	3 helices, a 3-way junction	1Y27	Serganov <i>et al.</i> (2004)
Type IA dG riboswitch (<i>Mesoplasma florum</i>)	2' deoxyguanosine	80	3 helices, a 3-way junction	3SKI	Pikovskaya <i>et al.</i> (2011)
Pre-Q ₁ -I (<i>B. subtilis</i>)	7-aminomethyl-7-deazaguanine	20	3 helices, 1 coaxial stack, H-type pseudoknot	3FU2	Klein <i>et al.</i> (2009)
Pre-Q ₁ -II (<i>Lactobacillus rhamnosus</i>)	7-aminomethyl-7-deazaguanine	18	4 helices, 1 coaxial stack, a 3-way junction, H-type pseudoknot	4JF2	Liberman <i>et al.</i> (2013)
C-di-AMP riboswitch (<i>B. subtilis</i>)	Cyclic di-adenosine monophosphate	≤ 0.1	N/A	N/A	Nelson <i>et al.</i> (2013)
C-di-GMP riboswitch-I (<i>Vibrio cholerae</i>)	Cyclic di-guanosine monophosphate	~ 1	H-shaped, 3 helices, 1 coaxial stack, a 3-way junction	3IWN	Kulshina <i>et al.</i> (2009)
C-di-GMP riboswitch-II (<i>Clostridium acetobutylicum</i>)	Cyclic di-guanosine monophosphate	~ 1	4 helices, 1 coaxial stack, a 3-way junction, 1 pseudoknot	3IRW	Smith <i>et al.</i> (2011)
Glycine riboswitch (<i>Vibrio cholerae</i>)	Glycine	1300	3 helices, 1 coaxial stack formed from central loop	3OWI	Huang <i>et al.</i> (2010)
Lysine riboswitch (<i>Thermotoga maritima</i>)	Lysine	1800	5 helices, 2 coaxial stacks, a 5-way junction	3DIL	Serganov <i>et al.</i> (2008)
Glutamine riboswitch (<i>Synechococcus elongates</i>)	L-Glutamine	6×10^5	N/A	N/A	Ames and Breaker, (2011)
RFN element (<i>Fusobacterium nucleatum</i>)	Flavin mononucleotide	12	6 helices, a 6-way junction	3F2Q	Serganov <i>et al.</i> (2009)
SAH riboswitch (<i>Ralstonia solanacearum</i>)	S-adenosyl-homocysteine	1350	3 helices, 1 coaxial stack, H-type pseudoknot	3NPQ	Edwards <i>et al.</i> (2010)
S-box (SAM I) riboswitch (<i>Thermoanaerobacter tengcongensis</i>)	S-adenosyl-methionine	1350	4 helices, 2 coaxial stacks, a 4-way junction	2GIS	Montange and Batey (2006)
SAM II riboswitch (from <i>Sargasso Sea metagenome</i>)	S-adenosyl-methionine	670	3 helical regions, 2 loop regions, 1 coaxial stack, H-type pseudoknot	2QWY	Gilbert <i>et al.</i> (2008)
S _{MK} box (SAM III) riboswitch (<i>Enterococcus faecalis</i>)	S-adenosyl-methionine	~ 850	4 helices, 1 coaxial stack, Y-shaped	3E5C	Lu <i>et al.</i> (2008)
SAM IV riboswitches (<i>Mycobacterium tuberculosis</i>)	S-adenosyl-methionine	~ 150	N/A	N/A	Weinberg <i>et al.</i> (2008)
SAM V riboswitch (<i>Cand. P. ubique</i>)	S-adenosyl-methionine	~ 150	N/A	N/A	Poiata <i>et al.</i> (2009)
AdoCbl riboswitch (<i>T. tengcongensis</i>)	Adenosyl-cobalamin	~ 250	2 coaxial stacks joined by T-loop-T-loop motifs	4GMA	Johnson <i>et al.</i> (2012)
AqCbl riboswitch (<i>env. metagenomes</i>)	Aquocobalamin	~ 7.5	2 coaxial stacks joined by T-loop-T-loop motifs	4FRN	Johnson <i>et al.</i> (2012)
thi-box riboswitch (<i>E. coli</i>)	Thiamine pyrophosphate	495	3 irregular helices, Y-shape	2HOJ	Edwards and Ferré-D'Amaré (2006)
GlyQ T-box Stem I (<i>B. subtilis</i>)	tRNA ^{Gly}	~ 150	Irregular helix, C-shape formed from K-turn	4LCK	Zhang and Ferré-D'Amaré (2013)
glmS ribozyme (<i>T. tengcongensis</i>)	Glucosamine-6-phosphate	$\sim 2 \times 10^5$	6 helices, 3 coaxial stacks, 3 pseudoknots	2Z75	Klein <i>et al.</i> (2007)

capable of catalyzing through *in vitro* evolution experiments (Tables 1 and 2). This apparently limited use of ribozymes by contemporary organisms could reflect the intrinsic chemical

limitations of RNA as a catalyst, compared to proteins. On the other hand, key cellular processes (such as protein synthesis and tRNA maturation) are RNA-catalyzed. Thus, some

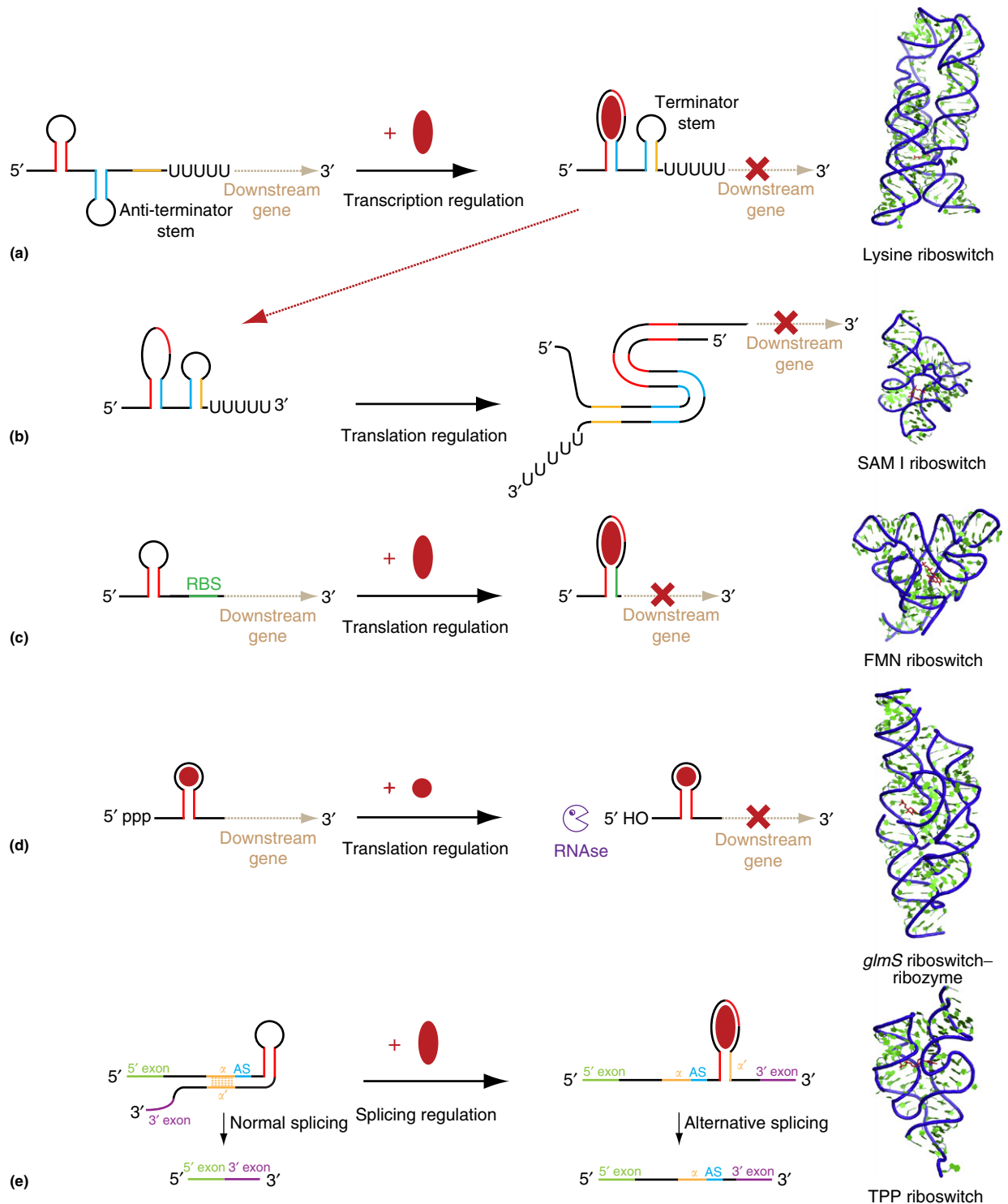


Figure 1 Cartoon representation of gene-regulation strategies employed by riboswitches. (a) Regulation at the transcription level. In the absence of ligand, the riboswitch forms an anti-terminator helix, which permits read-through by RNA polymerase. Upon binding the ligand, the riboswitch changes its secondary structure and forms instead a transcription terminator stem-loop. (b) The prematurely terminated RNA transcript from panel (a), in some cases, can also function as a noncoding RNA in *trans* and regulate translation initiation of a different mRNA. (c) The ribosomal binding site (RBS) is sequestered by the riboswitch in the presence of ligand. Sequestration of the RBS blocks translation initiation. (d) The *glmS* riboswitch-ribozyme is catalytically activated by ligand binding, cleaving the mRNA it is part of, resulting in a newly exposed 5'-OH. This is recognized by a ribonuclease, leading to degradation of the mRNA. (e) Regulation by splicing. In the absence of ligand, base-pairing between the α and α' elements (orange) brings the 5' and 3' splice sites close to each other favoring the principal splicing product. This base-pairing is disrupted when the riboswitch binds its ligand and sequesters the α' element, favoring splicing at alternative sites (AS) instead. The crystal structure of a riboswitch that exemplifies each of the five regulatory pathways is shown on the right.

properties of RNA must endow it with selective advantage over protein catalysts. Two possible such properties are that RNA is produced in one step from DNA (and thus ribozymes and riboswitches can be generated without the energetic cost or delay of protein synthesis), and that RNA, compared to protein, is much more mechanically stiff (thus allowing long-range communication within large macromolecular machines, such as the ribosome and the spliceosome).

Ribozymes in the Replication of Viroids and Virusoids

Ribozymes function in the replication of some viroids and virusoids. These are plant pathogens that consist solely of a single-stranded circular RNA, and are the smallest known replicating nucleic acids (only ~300 nucleotides). Unlike viruses, these pathogens lack protein coding capability, and are replicated by host RNA polymerases. Replication proceeds through a 'rolling-circle' mechanism (Figure 2(a)), in which the infectious RNA (the 'genomic RNA') is copied into concatenated RNAs of the opposite polarity (the 'antigenomic RNA'). This concatamer is cut into unit-length pieces, which then circularize to form antigenomic circles. Copying of these and processing into circles regenerates the infectious RNA. In many viroids and virusoids, ribozymes (such as the hairpin and hammerhead ribozymes) catalyze the cleavage and circularization reactions. A fungal parasitic DNA (the Varkud satellite of *Neurospora* mitochondria) employs a ribozyme in a related reaction. A human satellite virus of the Hepatitis B virus called the hepatitis δ virus (HDV) is strikingly similar to viroids (albeit somewhat larger), is copied by the human RNA polymerase II, and relies on self-cleaving ribozyme domains in both the genomic and antigenomic RNAs for processing. Remarkably, recent biochemical and bioinformatic experiments have shown that RNAs resembling the HDV ribozyme are scattered through phylogeny, from bacteriophage and bacteria through all animal phyla examined. The strong association of these HDV ribozyme-like RNAs with non-LTR retrotransposons suggests that the catalytic RNA serves an essential function in the physiology of this selfish genetic elements. The hammerhead ribozyme has also been found to be distributed across all of phylogeny, but its biological function is not well understood.

Ribozymes as Mobile Genetic Elements

Group II introns are self-splicing ribozymes present mostly in organellar pre-mRNAs. They catalyze two sequential steps (Figure 2(b)), and the transesterification reactions are reversible. The spliced intron sequence, interestingly, often includes an open reading frame encoding a large protein with reverse-transcriptase activity. This protein has been shown to associate with the ribozyme. The RNP thus formed is a mobile genetic element that can hop from its location in an intron-containing allele to the equivalent location in an intron-free allele, in a process of gene conversion. This is catalyzed both by the RNA and the protein component of the RNP. The spliced intron is responsible for locating the site of insertion into the intron-free allele, and cleaving one of the strands of the recipient DNA. The protein is responsible for the rest of the reaction in which the RNA is reverse-transcribed into the recipient. Although no

mobile group II introns have been detected in eukaryotic nuclear genomes, the mechanistic and structural similarity between these introns and the spliceosome suggests that pre-mRNA introns, together with the splicing machinery, might be remnants of ancestral group II intron-like mobile genetic elements.

The Ribosome as a Ribozyme and Riboswitch

In addition to being an RNA that catalyzes peptide bond formation, the ribosome is responsible for the process of decoding, that is, binding to tRNAs that are specified by the bound mRNA. Structural studies have demonstrated that the small subunit of the RNA carries out proof-reading of the match formed between the mRNA codons and the tRNA anticodons employing exclusively RNA. In this sense, the decoding site of the ribosome can be thought of as a programmable riboswitch. In order to function in protein synthesis, the ribosome must translocate to successive codons, and repeatedly form peptide bonds between the nascent polypeptide and the next amino acid, which is brought into the active site esterified to its cognate tRNA. Although, biologically, the process of translocation is catalyzed by proteins (elongation factors), studies using antibiotics have shown that the ribosome can translocate on its own (Konevega *et al.*, 2007). Such studies demonstrate that the ribosome can function solely on the basis of the free energy of peptide bond formation, and by co-opting the thermal energy of its surrounding solvent molecules.

The *glmS* Ribozyme–Riboswitch

The *glmS* riboswitch–ribozyme is another example of a riboswitch that is also a catalytic RNA. This ribozyme resides in the 5' UTR of the gene encoding glucosamine-6-phosphate (GlcN6P) synthetase (GlmS) in Gram-positive bacteria. Unlike other ribozymes that cleave RNA sequence-specifically, this RNA is inactive until it binds to GlcN6P. GlcN6P functions as a true coenzyme of the ribozyme, providing a key catalytic group to the otherwise preformed (but inert) ribozyme active site. In Gram-positive bacteria, intact mRNAs have a 5' triphosphate group. Self-cleavage by the *glmS* ribozyme–riboswitch domain exposes a 5'-OH group, and this is the signal that triggers degradation of the rest of the mRNA by the ribonuclease J1 protein (Collins *et al.*, 2007). Because the GlmS protein is unstable, destruction of its mRNA results in shutdown of GlcN6P production, thus completing negative feedback regulation mediated by this riboswitch–ribozyme. This RNA is of particular interest in evolutionary considerations because it is thus far the only known example of a natural RNA that employs a catalytic coenzyme. Analogously to the use of diverse coenzymes and prosthetic groups by protein enzymes, if RNAs can employ exogenous small molecules in their active sites, they could readily overcome limitations arising from the relatively simple chemical composition of nucleic acids. Indeed, it has been argued, based on the chemical similarity between many coenzymes and nucleic acids, that coenzyme utilization by modern protein enzymes is a remnant of an earlier metabolism where RNAs were responsible for forming the active sites of primordial proteins or peptide aggregates.

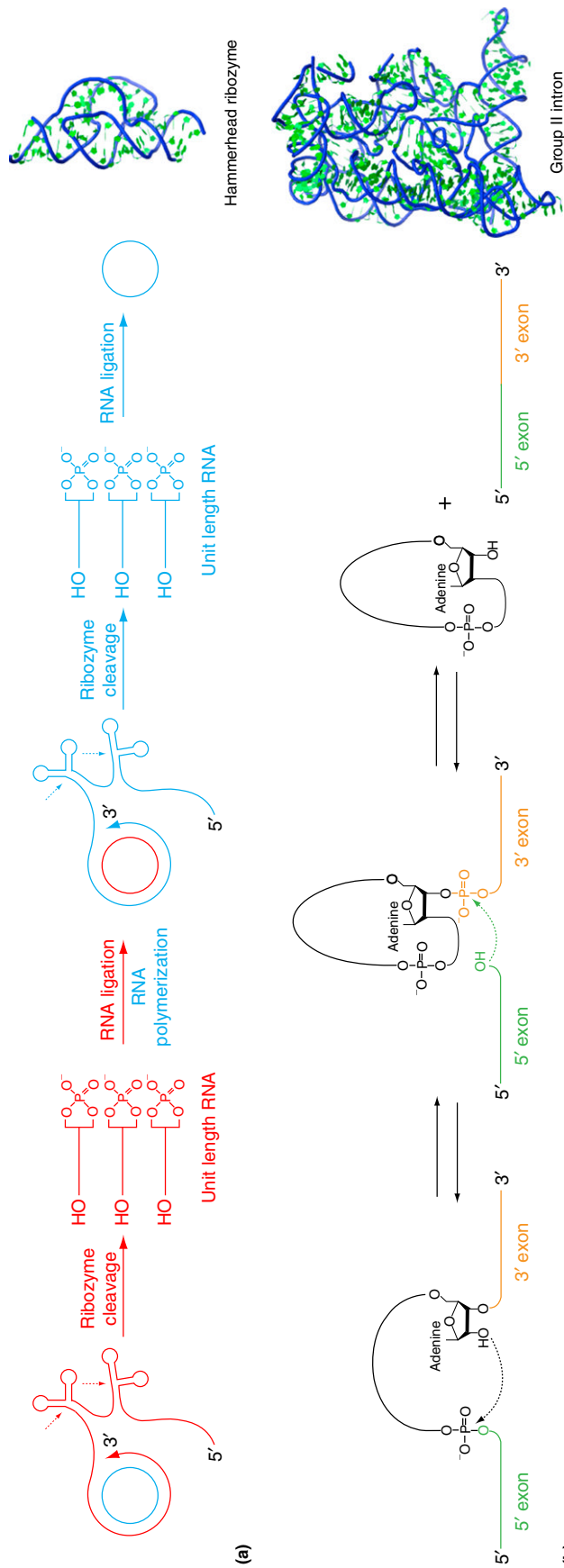


Figure 2 Examples of RNA processing mediated by ribozymes. (a) Rolling-circle replication employed by viroids and virusoids. Upon infection, the monomeric, circular viroid RNA ('genomic' sense, blue) is transcribed by a host RNA polymerase to generate a multimeric 'antigenomic' RNA (red). The antigenomic RNA is processed to unit-length copies by self-cleaving ribozyme domains (dotted arrow indicates cleavage site) and circularized by ligation. This process is repeated in the opposite direction to regenerate the genomic strand. (b) Splicing by Group II introns. This process involves two sequential steps: (1) nucleophilic attack of the 5' exon-intron junction, (2) nucleophilic attack of the 3' exon-intron junction. This results in the ligation of the two exons, and liberation of the intron as a lariat. Group II introns can alternatively employ water as a nucleophile in the first splicing reaction, yielding a linear intron and ligated exons (hydrolytic splicing). The crystal structure of a ribozyme involved in each of the two RNA processing pathways is shown on the right.

Ribozymes and Riboswitches as Cell Biological Tools

Ribozymes and riboswitches perform numerous gene regulatory tasks in modern cells. In addition, they have the potential to be employed as cell biological tools. Unlike tools such as anti-sense RNAs, siRNAs, and CRISPR, which rely on Watson–Crick base complementarity, riboswitches and ribozymes can function by directly recognizing the three-dimensional structures of their ligands, cofactors and substrates. Therefore, ribozymes and riboswitches have the potential to access a larger fraction of cellular molecules. For instance, ‘Spinach’ an aptamer RNA that binds to a latent fluorophore (a compound that does not become fluorescent until bound to the aptamer RNA) has been converted into a sensor for small molecule metabolites by fusing it to the ligand-binding domains of natural riboswitches (Paige *et al.*, 2012). Several artificial riboswitches that control gene expression in living cells in response to exogenous small molecules have been described (Sinha *et al.*, 2010). These overcome the difficulty in using natural riboswitches in recombinant contexts that arises from the fact that natural riboswitches are controlled by ubiquitous and essential cellular metabolites. Ribozymes that function as mobile genetic elements have the potential to be utilized for manipulating genomes *in vivo*. Molecular engineering efforts to harness the structural versatility of ribozymes and riboswitches represent a promising avenue of research.

Acknowledgment

This work was supported in part by the intramural program of the National Heart, Lung and Blood Institute, NIH.

See also: Molecular Principles, Components, Technology, and Concepts: Nucleic Acids: DNA, RNA Chemical Properties (Including Sequencing and Next-Generation Sequencing). Nucleic Acid Synthesis/Breakdown: Nucleic Acid Technology: Transgenesis and Gene Replacement; Viral Nucleic Acids. Nucleic Acid Synthesis/Breakdown: RNA Synthesis/Function: miRNAs/Small Noncoding RNAs; The Interplay between Eukaryotic mRNA Degradation and Translation. Nucleic Acid Synthesis/Breakdown: Transcription: Pre-mRNA Splicing: Function and Dysfunction; The Spliceosome and Pre-mRNA Splicing

References

- Ames, T.D., Breaker, R.R., 2011. Bacterial aptamers that selectively bind glutamine. *RNA Biology* 8, 82–89.
- Ban, N., Nissen, P., Hansen, J., *et al.*, 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289, 905–920.
- Bass, B.L., Cech, T.R., 1984. Specific interaction between the self-splicing RNA of *Tetrahymena* and its guanosine substrate: Implications for biological catalysis by RNA. *Nature* 308, 820–826.
- Beebe, J.A., Fierke, C.A., 1994. A kinetic mechanism for cleavage of precursor tRNA^{Asp} catalyzed by the RNA component of *Bacillus subtilis* ribonuclease P. *Biochemistry* 33, 10294–10304.
- Canny, M.D., Jucker, F.M., Kellogg, E., *et al.*, 2004. Fast cleavage kinetics of a natural hammerhead ribozyme. *Journal of the American Chemical Society* 126, 10848–10849.
- Collins, J.A., Irnov, I., Baker, S., Winkler, W.C., 2007. Mechanism of mRNA destabilization by the *glmS* ribozyme. *Genes & Development* 21, 3356–3368.
- Crick, F.H., 1968. The origin of the genetic code. *Journal of Molecular Biology* 38, 367–379.
- Dann, C.E., Wakeman, C.A., Sieling, C.L., *et al.*, 2007. Structure and mechanism of a metal-sensing regulatory RNA. *Cell* 130, 878–892.
- Edwards, A.L., Reyes, F.E., Heroux, A., Batey, R.T., 2010. Structural basis for recognition of S-adenosylhomocysteine by riboswitches. *RNA* 16, 2144–2155.
- Edwards, T.E., Ferré-D’Amaré, A.R., 2006. Crystal structures of the *thi*-box riboswitch bound to thiamine pyrophosphate analogs reveal adaptive RNA-small molecule recognition. *Structure* 14, 1459–1468.
- Ellington, A.D., Szostak, J.W., 1990. *In vitro* selection of RNA molecules that bind specific ligands. *Nature* 346, 818–822.
- Gilbert, S.D., Rambo, R.P., Van Tyne, D., Batey, R.T., 2008. Structure of the SAM-II riboswitch bound to S-adenosylmethionine. *Nature Structural & Molecular Biology* 15, 177–182.
- Glasner, M.E., Bergman, N.H., Bartel, D.P., 2002. Metal ion requirements for structure and catalysis of an RNA ligase ribozyme. *Biochemistry* 41, 8103–8112.
- Grundy, F.J., Henkin, T.M., 1993. tRNA as a positive regulator of transcription antitermination in *B. subtilis*. *Cell* 74, 475–482.
- Guerrier-Takada, C., Gardiner, K., Marsh, T., *et al.*, 1983. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell* 35, 849–857.
- Hegg, L.A., Fedor, M.J., 1995. Kinetics and thermodynamics of intermolecular catalysis by hairpin ribozymes. *Biochemistry* 34, 15813–15828.
- Holley, R.W., Apgar, J., Everett, G.A., *et al.*, 1965. Structure of a ribonucleic acid. *Science* 147, 1462–1465.
- Huang, F., Bugg, C.W., Yarus, M., 2000. RNA-Catalyzed CoA, NAD, and FAD synthesis from phosphopantetheine, NMN, and FMN. *Biochemistry* 39, 15548–15555.
- Huang, F., Yarus, M., 1997. 5'-RNA self-capping from guanosine diphosphate. *Biochemistry* 36, 6557–6563.
- Huang, L., Serganov, A., Patel, D.J., 2010. Structural insights into ligand recognition by a sensing domain of the cooperative glycine riboswitch. *Molecular Cell* 40, 774–786.
- Jacob, F., Monod, J., 1961. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* 3, 318–356.
- Johnson Jr., J.E., Reyes, F.E., Polaski, J.T., Batey, R.T., 2012. B12 cofactors directly stabilize an mRNA regulatory switch. *Nature* 492, 133–137.
- Joyce, G.F., 1989. Amplification, mutation and selection of catalytic RNA. *Gene* 82, 83–87.
- Katunin, V.I., Muth, G.W., Strobel, S.A., *et al.*, 2002. Important contribution to catalysis of peptide bond formation by a single ionizing group within the ribosome. *Molecular Cell* 10, 339–346.
- Klein, D., Edwards, T., Ferré-D’Amaré, A., 2009. Cocrystal structure of a class I preQ₁ riboswitch reveals a pseudoknot recognizing an essential hypermodified nucleobase. *Nature Structural & Molecular Biology* 16, 343–344.
- Klein, D.J., Ferré-D’Amaré, A.R., 2006. Structural basis of *glmS* ribozyme activation by glucosamine-6-phosphate. *Science* 313, 1752–1756.
- Klein, D.J., Wilkinson, S.R., Been, M.D., Ferré-D’Amaré, A.R., 2007. Requirement of helix P2.2 and nucleotide G1 for positioning the cleavage site and cofactor of the *glmS* ribozyme. *Journal of Molecular Biology* 373, 178–189.
- Konevega, A.L., Fischer, N., Semenov, Y.P., *et al.*, 2007. Spontaneous reverse movement of mRNA-bound tRNA through the ribosome. *Nature Structural & Molecular Biology* 14, 318–324.
- Kruger, K., Grabowski, P.J., Zaug, A.J., *et al.*, 1982. Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* 31, 147–157.
- Kulshina, N., Baird, N.J., Ferré-D’Amaré, A.R., 2009. Recognition of the bacterial second messenger cyclic diguanylate by its cognate riboswitch. *Nature Structural & Molecular Biology* 16, 1212–1217.
- Lau, M.W., Cadieux, K.E., Unrau, P.J., 2004. Isolation of fast purine nucleotide synthase ribozymes. *Journal of the American Chemical Society* 126, 15686–15693.
- Lau, M.W., Ferré-D’Amaré, A.R., 2013. An *in vitro* evolved *glmS* ribozyme has the wild-type fold but loses coenzyme dependence. *Nature Chemical Biology* 9, 805–810.
- Lau, M.W., Unrau, P.J., 2009. A promiscuous ribozyme promotes nucleotide synthesis in addition to ribose chemistry. *Chemistry & Biology* 16, 815–825.
- Li, S., Breaker, R.R., 2013. Eukaryotic TPP riboswitch regulation of alternative splicing involving long-distance base pairing. *Nucleic Acids Research* 41, 3022–3031.
- Lieberman, J.A., Salim, M., Krucinska, J., Wedekind, J.E., 2013. Structure of a class II preQ₁ riboswitch reveals ligand recognition by a new fold. *Nature Chemical Biology* 9, 353–355.

- Loh, E., Dussurget, O., Gripenland, J., *et al.*, 2010. A *trans*-acting riboswitch controls expression of the virulence regulator PrfA in *Listeria monocytogenes*. *Cell* 139, 770–779.
- Lorsch, J.R., Szostak, J.W., 1994. *In vitro* evolution of new ribozymes with polynucleotide kinase activity. *Nature* 371, 31–36.
- Lu, C., Smith, A.M., Fuchs, R.T., *et al.*, 2008. Crystal structures of the SAM-III/SMK riboswitch reveal the SAM-dependent translation inhibition mechanism. *Nature Structural & Molecular Biology* 15, 1076–1083.
- Miranda-Ríos, J., Navarro, M., Soberón, M., 2001. A conserved RNA structure (*thi* box) is involved in regulation of thiamin biosynthetic gene expression in bacteria. *Proceedings of the National Academy of Sciences of the USA* 98, 9736–9741.
- Mironov, A.S., Gusarov, I., Rafikov, R., *et al.*, 2002. Sensing small molecules by nascent RNA: A mechanism to control transcription in bacteria. *Cell* 111, 747–756.
- Montange, R.K., Batey, R.T., 2006. Structure of the S-adenosylmethionine riboswitch regulatory mRNA element. *Nature* 441, 1172–1175.
- Nahvi, A., Sudarsan, N., Ebert, M.S., *et al.*, 2002. Genetic control by a metabolite binding mRNA. *Chemistry & Biology* 9, 1043.
- Nelson, J.W., Sudarsan, N., Furukawa, K., *et al.*, 2013. Riboswitches in eubacteria sense the second messenger c-di-AMP. *Nature Chemical Biology* 9, 834–839.
- Orgel, L.E., 1968. Evolution of the genetic apparatus. *Journal of Molecular Biology* 38, 381–393.
- Paige, J.S., Nguyen-Duc, T., Song, W., Jaffrey, S.R., 2012. Fluorescence imaging of cellular metabolites with RNA. *Science* 335, 1194.
- Pan, T., Uhlenbeck, O.C., 1992. A small metalloribozyme with a two-step mechanism. *Nature* 358, 560–563.
- Pikovskaya, O., Polonskaia, A., Patel, D.J., Serganov, A., 2011. Structural principles of nucleoside selectivity in a 2'-deoxyguanosine riboswitch. *Nature Chemical Biology* 7, 748–755.
- Poïata, E., Meyer, M.M., Ames, T.D., Breaker, R.R., 2009. A variant riboswitch aptamer class for S-adenosylmethionine common in marine bacteria. *RNA* 15, 2046–2056.
- Regulski, E.E., Moy, R.H., Weinberg, Z., *et al.*, 2008. A widespread riboswitch candidate that controls bacterial genes involved in molybdenum cofactor and tungsten cofactor metabolism. *Molecular Microbiology* 68, 918–932.
- Ren, A., Rajashankar, K.R., Patel, D.J., 2012. Fluoride ion encapsulation by Mg^{2+} ions and phosphates in a fluoride riboswitch. *Nature* 486, 85–89.
- Roth, A., Weinberg, Z., Chen, A.G., *et al.*, 2014. A widespread self-cleaving ribozyme class is revealed by bioinformatics. *Nature Chemical Biology* 10, 56–60.
- Ryu, Y., Kim, K.J., Roessner, C.A., Scott, A.I., 2006. Decarboxylative Claisen condensation catalyzed by *in vitro* selected ribozymes. *Chemical Communications (Cambridge, England)*. 1439–1441.
- Saito, H., Kourouklis, D., Suga, H., 2001. An *in vitro* evolved precursor tRNA with aminoacylation activity. *EMBO Journal* 20, 1797–1806.
- Seelig, B., Jäschke, A., 1999. A small catalytic RNA motif with Diels-Alderase activity. *Chemistry & Biology* 6, 167–176.
- Sengle, G., Eisenfuhr, A., Arora, P.S., *et al.*, 2001. Novel RNA catalysts for the Michael reaction. *Chemistry & Biology* 8, 459–473.
- Serganov, A., Huang, L., Patel, D., 2008. Structural insights into amino acid binding and gene control by a lysine riboswitch. *Nature* 455, 1263–1267.
- Serganov, A., Huang, L., Patel, D., 2009. Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature* 457, 233–237.
- Serganov, A., Nudler, E., 2013. A decade of riboswitches. *Cell* 152, 17–24.
- Serganov, A., Yuan, Y.R., Pikovskaya, O., *et al.*, 2004. Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chemistry & Biology* 11, 1729–1741.
- Sinha, J., Reyes, S.J., Gallivan, J.P., 2010. Reprogramming bacteria to seek and destroy an herbicide. *Nature Chemical Biology* 6, 464–470.
- Smith, K.D., Shanahan, C.A., Moore, E.L., *et al.*, 2011. Structural basis of differential ligand recognition by two classes of bis-(3'-5')-cyclic dimeric guanosine monophosphate-binding riboswitches. *Proceedings of the National Academy of Sciences of the USA* 108, 7757–7762.
- Tang, Y., Nielsen, H., Birgisdóttir, A.B., Johansen, S., 2011. A natural fast-cleaving branching ribozyme from the amoebal flagellate *Naegleria gruberi*. *RNA Biology* 8, 997–1004.
- Teixeira, A., Tahiri-Alaoui, A., West, S., *et al.*, 2004. Autocatalytic RNA cleavage in the human β -globin pre-mRNA promotes transcription termination. *Nature* 432, 526–530.
- Thill, G., Vasseur, M., Tanner, N.K., 1993. Structural and sequence elements required for the self-cleaving activity of the hepatitis delta virus ribozyme. *Biochemistry* 32, 4254–4262.
- Trausch, J.J., Ceres, P., Reyes, F.E., Batey, R.T., 2011. The structure of a tetrahydrofolate-sensing riboswitch reveals two ligand binding sites in a single aptamer. *Structure* 19, 1413–1423.
- Tuerk, C., Gold, L., 1990. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* 249, 505–510.
- Weinberg, Z., Regulski, E.E., Hammond, M.C., *et al.*, 2008. The aptamer core of SAM-IV riboswitches mimics the ligand-binding site of SAM-I riboswitches. *RNA* 14, 822–828.
- Xiang, Q., Qin, P.Z., Michels, W.J., *et al.*, 1998. Sequence specificity of a group II intron ribozyme: Multiple mechanisms for promoting unusually high discrimination against mismatched targets. *Biochemistry* 37, 3839–3849.
- Zamel, R., Poon, A., Jaikaran, D., *et al.*, 2004. Exceptionally fast self-cleavage by a *Neurospora* Varkud satellite ribozyme. *Proceedings of the National Academy of Sciences of the USA* 101, 1467–1472.
- Zhang, J., Ferre-D'Amare, A.R., 2013. Co-crystal structure of a T-box riboswitch stem I domain in complex with its cognate tRNA. *Nature* 500, 363–366.

Ribosomal RNAs and Protein Synthesis

S Joseph, University of California, San Diego, La Jolla, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Elongation factors Proteins that promote the extension of the protein chain by the ribosome.

Initiation factors Proteins that help the ribosome to start the process of protein synthesis on an mRNA.

Messenger RNA RNA molecules that carry information to the ribosome for protein synthesis.

Ribosome Cellular structures made of RNA and protein molecules that synthesize proteins that the cells need for growth and division.

Termination factors Proteins that are required for the release of the newly synthesized protein from the ribosome and for the recycling of the ribosome.

Transfer RNA RNA molecules that bring amino acids to the ribosome for protein synthesis.

Introduction

Ribosomes synthesize proteins by translating the genetic information in messenger RNAs and are one of the few macromolecular complexes in the cell that are composed of both proteins and RNAs. The active role played by the ribosomal RNAs (rRNAs) in protein synthesis is evidence that the ribosomes are a relic of the primitive RNA world. This article will focus on the functional role of the rRNAs in bacterial protein synthesis. In bacteria, the 70S ribosome (2.3 MDa) consists of the 50S large ribosomal subunit and the 30S small ribosomal subunit. The 50S ribosomal subunit is composed of 23S rRNA (2900 nts), 5S rRNA (120 nts) and more than 30 ribosomal proteins and is responsible for catalyzing peptide bond formation. The 30S ribosomal subunit is composed of 16S rRNA (1500 nts) and about 20 ribosomal proteins and is responsible for decoding the mRNA sequence. Aminoacyl-tRNAs are the substrates for protein synthesis and each ribosomal subunit has three tRNA binding sites designated aminoacyl site (A site), peptidyl site (P site), and exit site (E site). During protein synthesis, tRNAs bind first to the A site, then to the P site and finally to the E site in a step-wise fashion. This iterative tRNA binding process is the basis for the elongation cycle of protein synthesis.

Overview of Protein Synthesis

Protein synthesis is an energy intensive process requiring about 10 translation factors in addition to aminoacyl-tRNAs and the ribosome. The process of protein synthesis can be divided into four stages: (1) initiation, (2) elongation, (3) termination, and (4) recycling (Figure 1). During the initiation stage of protein synthesis, the 30S ribosomal subunit binds the mRNA and the initiator tRNA (formyl-methionine-tRNA^{Met}) with the help of initiation factors 1, 2, and 3 (IF1, IF2, and IF3) (reviewed in Boelens and Gualerzi, 2002). The initiator tRNA binds to the P site in the 30S subunit and interacts with the start codon in the mRNA. Correct base pair formation between the anticodon of the initiator tRNA and the start codon promotes the association of the 50S subunit to the 30S subunit and the release of IF1,

IF2, and IF3 from the ribosome. The ribosome then begins the elongation cycle of protein synthesis by binding elongation factor Tu•GTP•aminoacyl tRNA (EF-Tu ternary complex) to the ribosomal A site. Proper Watson–Crick base pair formation between the anticodon of the tRNA and the mRNA codon in the A site triggers the accommodation of the tRNA into the 50S subunit and the release of EF-Tu•GDP from the ribosome (reviewed in Rodnina *et al.*, 2005). The peptidyl transferase center in the 50S subunit then catalyzes peptide bond formation resulting in the extension of the nascent peptide by one amino acid. Next, the deacylated tRNA in the P site and the peptidyl-tRNA in the A site are translocated to the E and P sites, respectively, by the activity of elongation factor G (EF-G) (reviewed in Shoji *et al.*, 2009). This process also brings the next mRNA codon into the A site for interaction with EF-Tu ternary complexes. This cycle continues until a stop codon is placed in the A site. The stop codon in the A site is recognized not by an EF-Tu ternary complex but by release factors 1 or 2 (RF1 recognizes UAA and UAG where as RF2 recognizes UAA and UGA). RF1 and RF2 promote the release of the newly synthesized protein from the peptidyl-tRNA in the P site (reviewed in Zhou *et al.*, 2012a). RF1 and RF2 are released from the ribosome by the activity of RF3•GTP, which in turn hydrolyze GTP to dissociate from the ribosome. The ribosome is now left with a deacylated tRNA in the P site and the mRNA. This ribosomal complex binds RRF and EF-G, which separates the ribosome into the 50S and 30S subunits and the mRNA and tRNA are released from the 30S subunit by the proof-reading activity of IF3 (reviewed in Petty *et al.*, 2008). The vacant 50S and 30S subunits are free to start the whole process of protein synthesis again on another mRNA.

Functional Role of 16S and 23S rRNAs During Initiation

Interactions with mRNA

As described above, the process of protein synthesis begins with the 30S subunit binding an mRNA. The mRNA binds in a

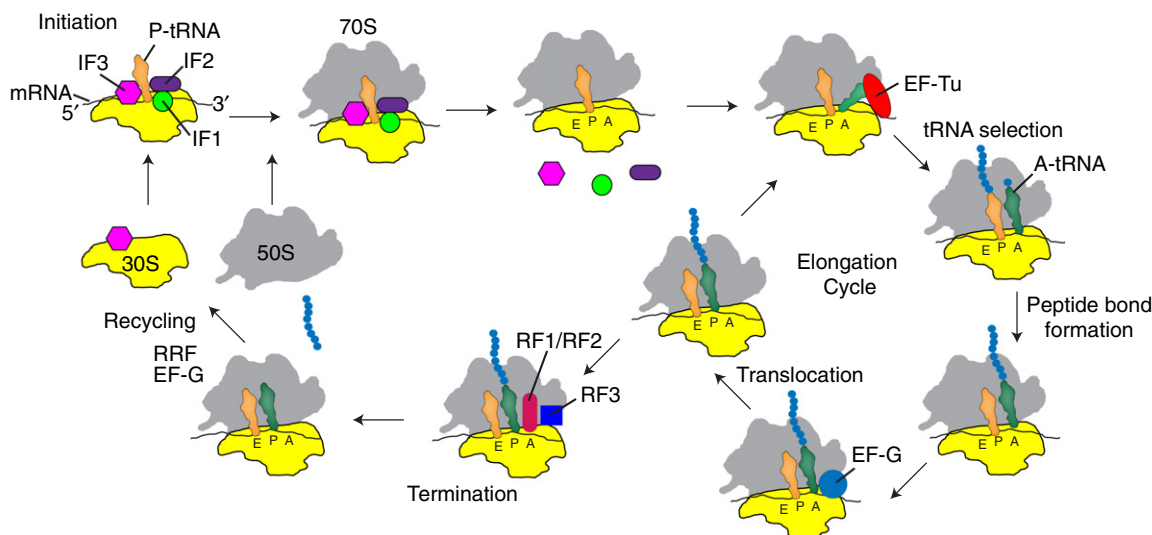


Figure 1 Overview of bacterial protein synthesis. During translation initiation the 30S subunit binds an mRNA and the initiator tRNA with the help of IF1, IF2, and IF3. The 30S initiation complex then binds the 50S subunit to form 70S ribosome. The 70S ribosome enters the elongation cycle by binding EF-Tu-GTP-aminoacyl tRNA complex to the A site and catalyzing peptide bond formation. This is followed by mRNA-tRNA translocation catalyzed by EF-G-GTP. In the termination stage, RF1/RF2 binds to the ribosome with a stop codon in the A site and catalyzes the release of the newly synthesized protein. RF3-GTP then triggers the release of RF/RF2 from the ribosome. The 70S ribosome is recycled into the 30S and 50S subunits by the combined activities of RRF and EF-G.

channel between the 'head' and the 'body' of the 30S subunit and about 30 nucleotides of the mRNA from position -18 to $+12$ are accommodated by the ribosome (Figure 2; Yusupova *et al.*, 2001). Most bacterial mRNAs have a purine-rich sequence 4–8 nucleotides upstream of the start codon called the Shine–Dalgarno (SD) sequence (Shine and Dalgarno, 1974). The SD sequence in the mRNA forms Watson–Crick base pairs with a complementary sequence, the anti-SD sequence, present at the 3' end of the 16S rRNA. The SD/anti-SD duplex is located in a cleft at the back of the 'head' and 'platform' domains of the 30S subunit and interacts with helices 23, 28, and 37 of 16S rRNA and ribosomal proteins S11 and S18 (Figure 2 (d); Yusupova *et al.*, 2001; Korostelev *et al.*, 2007). The mRNA (positions -4 to -1) then passes through a short tunnel between the head and the platform of the 30S subunit to the subunit interface side, where it interacts with the 690 loop, the 790 loop, the 925 region of helix 28, and helix 45 of 16S rRNA. The mRNA then undergoes a turn around position -1 and the P and A site codons (positions $+1$ to $+6$) are exposed on the interface surface for base pairing interactions with the anticodons of P site and A site tRNAs. Interestingly, the mRNA has a sharp kink between the P and A site codon to allow for the simultaneous binding of two tRNAs. This kink in the mRNA is caused by the phosphate of nucleotide 1401 of 16S rRNA blocking the path of the mRNA. Some of the 16S rRNA nucleotides that contact the P site codon are 790, 791, 926, and 1498. The A site codon interacts with the universally conserved bases G530, A1492, and A1493 of 16S rRNA and protein S12. The mRNA downstream to the A site codon (positions $+7$ to $+10$) passes through another tunnel formed by helix 34 at the top, helix 28 on the right, the 530 loop on the left, and the 5' terminus of 16S rRNA, which forms a hairpin loop at the bottom. Finally, the mRNA (positions $+11$ to $+15$) passes through a ring of proteins (S3, S4, and

S5) to emerge on the backside of the 30S subunit. Proteins S3, S4, and S5 facilitate the disruption of mRNA secondary structure during translation (Takyar *et al.*, 2005). Thus, several highly conserved nucleotides of 16S rRNA interact intimately with the mRNA as it traverses the 30S subunit. Chief among these interactions for translation initiation is the formation of the SD/anti-SD duplex. Masking of the SD sequence by secondary structure in the mRNA or by the binding of proteins reduces the efficiency of translation initiation (de Smit and Van Duin, 1990; Brunel *et al.*, 1995; Studer and Joseph, 2006). Indeed, cells regulate the expression of some genes at the post-transcriptional level by changing the availability of the SD sequence. Finally, it is not clear exactly when the SD/anti-SD duplex is disrupted during protein synthesis.

Interactions with Initiation Factors

IF1, IF2, and IF3 are essential translational factors that promote the recruitment of mRNA and initiator tRNA to the 30S subunit to form the 30S initiation complex (30SIC) (reviewed in Boelens and Gualerzi, 2002; Milón *et al.*, 2012). These factors also ensure that the initiator tRNA interacts with the correct start codon so that the appropriate mRNA reading frame is translated. The role of IF1 is not clear, it has been proposed to stimulate the activities of IF2 and IF3. IF1 binds to the 30S subunit near the A site such that it would preclude the binding of a tRNA to the A site (Figure 3(a); Carter *et al.*, 2001). IF1 interacts with the 530 loop (helix 18) and helix 44 of 16S rRNA and with ribosomal protein S12 (Figure 3(b)). IF1 interacts specifically with A1492 and A1493 of 16S rRNA, inducing these bases to flip out of helix 44. IF1 also induces conformational changes within the 30S subunit that may affect the ability the 30S subunit to bind to the 50S subunit.

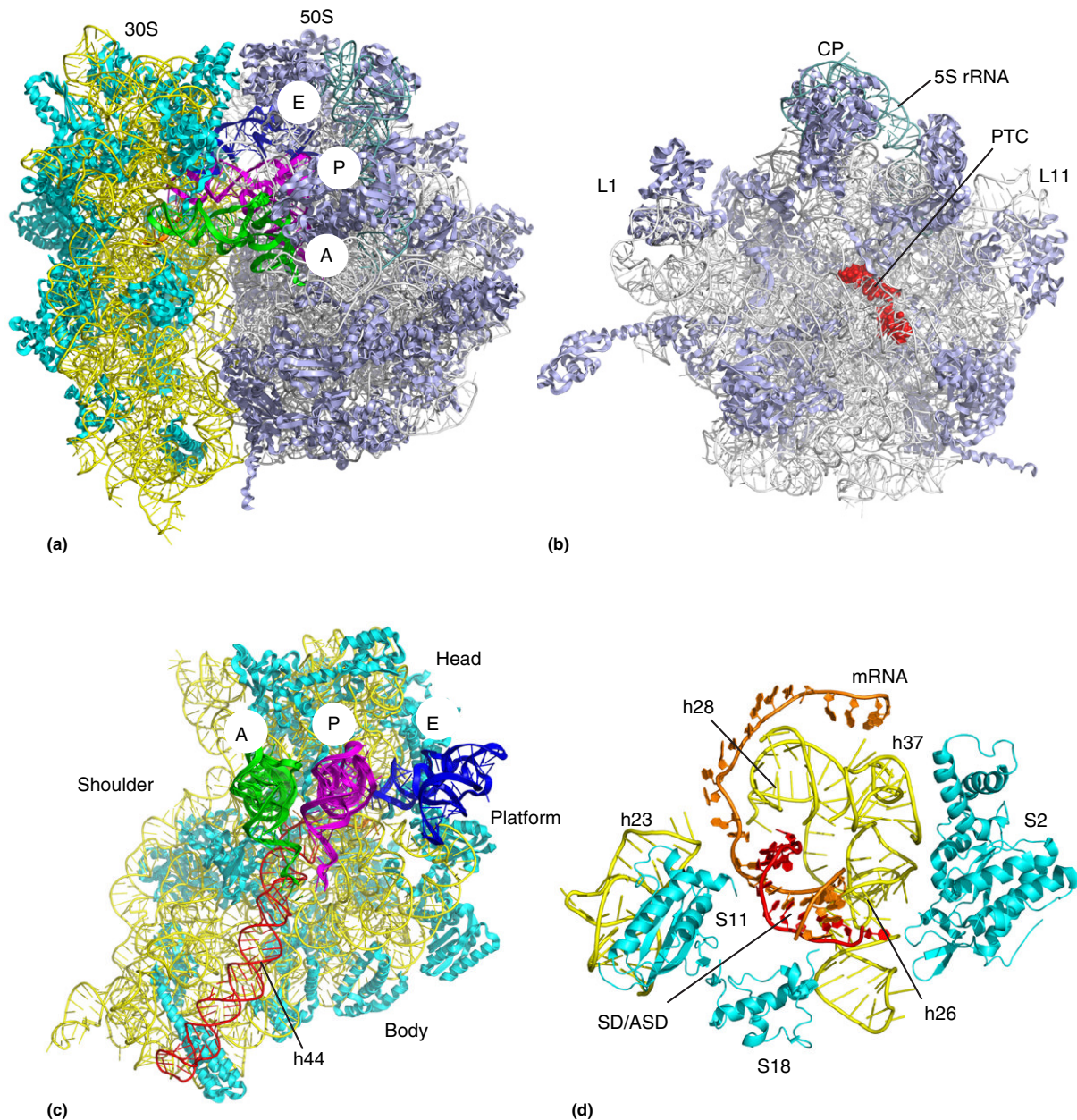


Figure 2 Structure of the ribosome. (a) Ribosome with three tRNAs in the A, P, and E sites. The 30S subunit consists of 16S rRNA (yellow) and 20 proteins (cyan) and the 50S subunit consists of 5S rRNA (teal), 23S rRNA (white), and 30 proteins (light blue). A site tRNA (green), P site tRNA (purple), and E site tRNA (blue) bind between the 30S and 50S subunits (PDB ID: 2WDK and 2WDL). (b) The 50S subunit viewed from the intersubunit side. The peptidyl transferase center (PTC in red), L1 stalk, Central Protuberance (CP), L11 region and 5S rRNA are indicated (PDB ID: 2WDL). (c) The 30S subunit viewed from the intersubunit side. The A, P, and E site tRNAs are shown. The 'head,' 'platform,' 'body,' and 'shoulder' domains of the 30S subunit are indicated. Helix 44 is in red (PDB ID: 2WDK). (d) Interactions of mRNA with the 30S subunit. The Shine-Dalgarno sequence of the mRNA (orange) base pairs with the 3' end of 16S rRNA forming a helix (SD/ASD in red) and interacts with 16S rRNA (yellow) and proteins S2, S11, and S18 (all in cyan) (PDB ID: 2HGR).

IF2 selectively stabilizes the initiator tRNA in the P site of the 30S subunit (Milon *et al.*, 2010). IF2 is a GTP/GDP binding protein that has a 'chalice' shaped structure with a globular N-terminal domain that interacts with the 30S subunit, G domain that binds GTP and C-terminal region that is used to recognize the formylated α -amino group of the initiator tRNA (Figure 3(c); Roll-Mecak *et al.*, 2000). IF2 binds on the interface side of the 30S subunit with the N-terminal region

close to helices 5 and 14 of 16S rRNA, the G domain contacting the GTPase-associated center of the 50S subunit and the C-terminal region interacting with the 3'-end of the initiator tRNA (Allen *et al.*, 2005; Myasnikov *et al.*, 2005; Simonetti *et al.*, 2008). IF3 prevents the premature association of the 30SIC with the 50S subunit (Antoun *et al.*, 2006a,b). It also carries out a similar function during ribosome recycling by preventing the re-association of the 30S and 50S subunits. IF3

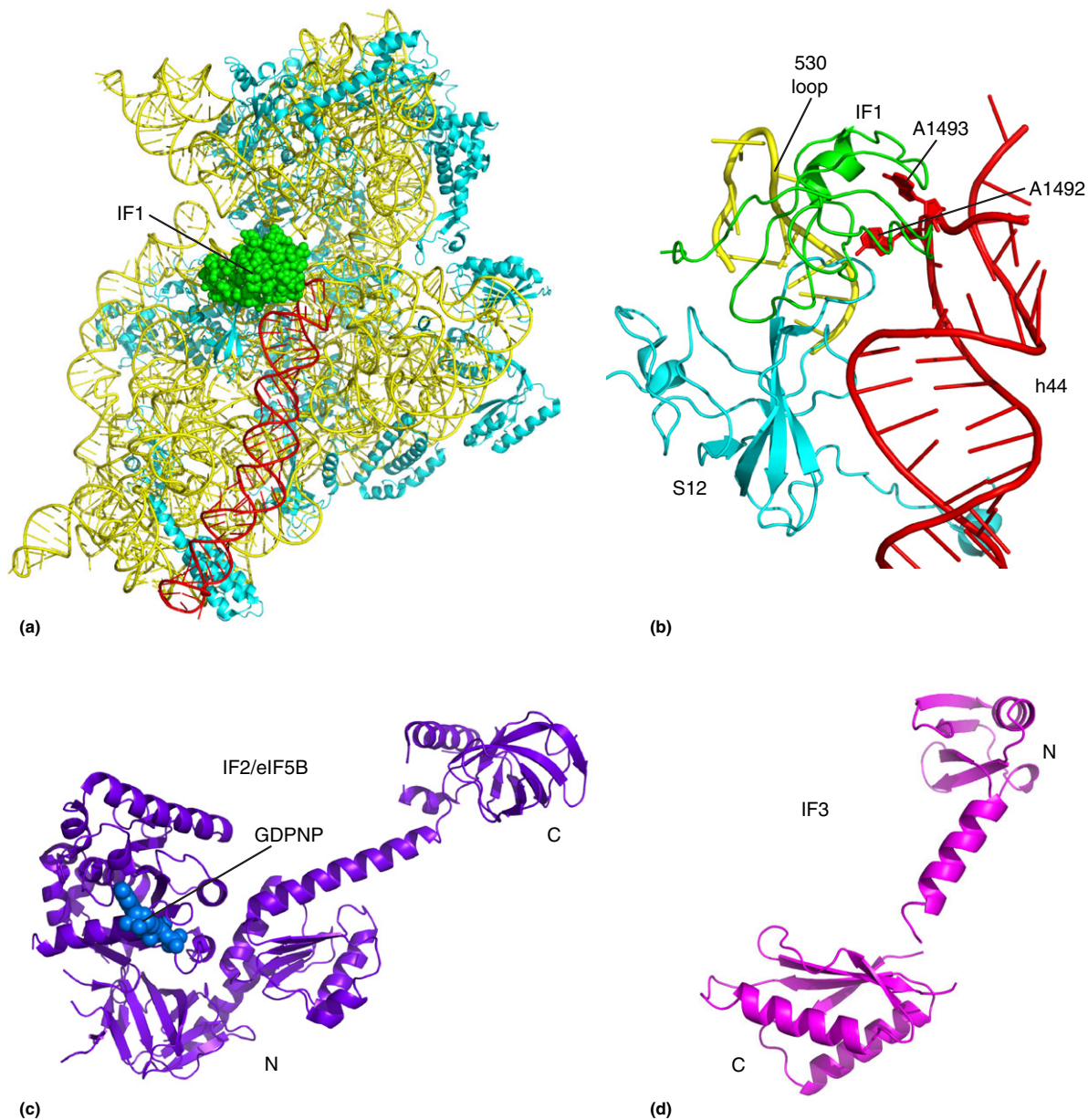


Figure 3 Interactions of initiation factors with the ribosome. (a) Structure of IF1 bound to the 30S subunit. IF1 is shown as green spheres and helix 44 of 16S rRNA is shown in red (PDB ID: 1HR0). (b) Close-up of IF1 interacting with 16S rRNA and protein S12 (cyan). (c) Structure of IF2/eIF5B. The N- and C-terminal domains are indicated and the GTP analog, GDPNP, is indicated by the blue spheres (PDB ID: 1G7T). (d) Structure of IF3 N and C-terminal domains (PDB ID: 1TIF and 1TIG).

enhances the fidelity of translation initiation by destabilizing the interaction of an incorrectly bound near- or non-cognate tRNA in the P site of the 30SIC (Antoun *et al.*, 2006a; Elvekrog and Gonzalez, 2013). It has N- and C-terminal globular domains connected by a flexible linker (Figure 3(d)). IF3 binds on the subunit interface side of the platform domain of the 30S subunit. The C-terminal domain interacts with helices 23, 24, and 45 of 16S rRNA, whereas the N-terminal domain interacts with the E site (Dallas and Noller, 2001; McCutcheon *et al.*, 1999). The binding of IF3 to 30S subunit is proposed to disrupt key intersubunit bridges formed by the platform domain with helix 69 of 23S rRNA (Dallas and Noller, 2001;

McCutcheon *et al.*, 1999), which is consistent with its function as an anti-association factor.

Interactions with Initiator tRNA

The initiator tRNA binds directly to the 30S subunit P site, which is mainly composed of 16S rRNA (Noller *et al.*, 2005). The interaction of 16S rRNA bases G1338 and A1339 in the P site with the initiator tRNA is important for IF3-dependent discrimination of initiator tRNA from elongator tRNAs (Lancaster and Noller, 2005). Other 16S rRNA nucleotides that

contact the P-tRNA are A790, G926, G966, A1229, C1230, C1400, and U1498 (Yusupov *et al.*, 2001; Korostelev *et al.*, 2006; Selmer *et al.*, 2006). Mutations of these bases have only a modest effect on translation. Similarly, the interactions of ribosomal proteins S9 and S13 with the anticodon arm of the P-tRNA are also not essential for translation (Hoang *et al.*, 2004). These results suggest that the interactions in the 30S subunit P site are redundant and initiation will occur as long as the initiator tRNA can bind with reasonable affinity to the P site (Hoang *et al.*, 2004).

Subunit Association and rRNA

Once the initiator tRNA has been properly placed in the P site, the 30SIC undergoes a conformational change that results in the anticodon of the initiator tRNA forming base pairs with the start codon (La Teana *et al.*, 1996). This also triggers the release of IF1 and IF3 from the 30SIC, followed by the binding of the 50S subunit to the 30SIC to form the 70S initiation complex (Tsai *et al.*, 2012). The G domain of IF2 interacts with the GTPase-associated center of the 50S subunit resulting in GTP hydrolysis on IF2 and the release of IF2•GDP from the 70S ribosome (Tomsic *et al.*, 2000). At the end of the initiation stage the ribosome contains a correctly positioned mRNA with the start codon interacting with the anticodon of the initiator tRNA in the P site and a vacant A site.

There are about a dozen contact points between the 30S and 50S subunits called intersubunit bridges that hold the subunits together in the 70S ribosome (Gabashvili *et al.*, 2000; Yusupov *et al.*, 2001; Korostelev *et al.*, 2006; Selmer *et al.*, 2006). Most of these bridges are composed of 16S rRNA interacting directly with the 23S rRNA; however, some of the bridges involve contacts between the rRNAs and ribosomal proteins. In particular, helix 44 of 16S rRNA, which lies at the interface of the 30S subunit, makes a number of contacts with the 50S subunit and IF1-dependent changes to the structure of helix 44 may control the rates of subunit association and dissociation (Carter *et al.*, 2001). Finally, some of bridges that connect the 30S subunit head with the 50S subunit are dynamically altered as the ribosome undergoes conformational changes during the translational cycle (Gao *et al.*, 2003; Frank *et al.*, 2007; Frank and Gonzalez, 2010). These bridges may play a functional role by coordinating the large-scale conformational changes in the ribosome and by biasing the movement of tRNAs.

Functional Role of 16S and 23S RNAs During the Elongation Cycle

Interactions of EF-Tu-GTP-Aminoacyl tRNA with the Ribosomal RNA

During the elongation stage, aminoacyl-tRNAs bind as an EF-Tu•GTP ternary complex to the A site (Figure 4(a)). The ribosome selects the cognate EF-Tu ternary complex while rejecting non-cognate EF-Tu ternary complexes with an error frequency of about 10^{-3} . Universally conserved 16S rRNA bases G530, A1492, and A1493 located in the decoding center of the 30S subunit play a crucial role in achieving this high

level of accuracy during tRNA selection (Ogle *et al.*, 2001). Upon cognate tRNA binding to the A site, G530, A1492, and A1493 undergo conformational changes to interact with the minor groove of the codon-anticodon helix (Ogle *et al.*, 2001). The first codon-anticodon base pair interacts with A1493. The second codon-anticodon base pair interacts with G530 and A1492. Finally, the third codon-anticodon base pair interacts less stringently with G530, which allows for wobble base pairs at the third position of the codon. G530, A1492, and A1493 interact stably with the cognate codon-anticodon helix because they have the proper A-form helical geometry. In contrast, the binding of a non-cognate tRNA with mismatches in the codon-anticodon helix deviates from A-form helical geometry, which destabilizes the interactions of G530, A1492, and A1493. In addition, non-cognate tRNAs fail to trigger a conformational change of the 30S subunit called 'domain closure' resulting ultimately in the rejection of the non-cognate tRNA (Ogle *et al.*, 2002). Thus, differences in the structure of the cognate tRNA•ribosome complex versus non-cognate tRNA•ribosome complex are proposed to be the molecular basis for fidelity. However, an alternate proposal is that the decoding center clamps tightly around the codon-anticodon helix forcing the mismatched base pair to adopt the geometry of a normal Watson-Crick pair, which is energetically unfavorable resulting in the rejection of the non-cognate and near cognate tRNAs (Demeshkina *et al.*, 2012).

GTP Hydrolysis

Binding of a cognate EF-Tu ternary complex to the ribosome accelerates GTP hydrolysis on EF-Tu, whereas non-cognate EF-Tu ternary complexes fail to stimulate rapid GTP hydrolysis and are released from the ribosome (reviewed in Rodnina and Wintermeyer, 2001). Interaction of the G domain of EF-Tu with the GTPase-associated center of the 50S subunit is critical for activating GTP hydrolysis. The GTPase-associated center consists of the universally conserved sarcin-ricin loop (helix 95) of 23S rRNA, the L11 protein and proximal rRNA (helix 43), and the L7/L12 ribosomal proteins (Figure 4(b); Schmeing *et al.*, 2009). The phosphate of A2662 in the sarcin-ricin loop is proposed to properly position the catalytic histidine residue in EF-Tu to allow a water molecule to carry out GTP hydrolysis (Voorhees *et al.*, 2010). A similar mechanism may be used by other translational GTPases to catalyze GTP hydrolysis on the ribosome. However, more biochemical and structural studies are necessary to precisely understand the steps that lead to GTPase activation and GTP hydrolysis.

Peptide Bond Formation

After GTP hydrolysis, EF-Tu•GDP dissociates from the ribosome and the aminoacyl-tRNA is accommodated into the 50S subunit A site. Peptidyl transferase reaction occurs in a deep cleft in the interface side of the 50S subunit (Ban *et al.*, 2000; Nissen *et al.*, 2000). The 3' aminoacyl end of the A site tRNA is in close proximity to the 3'-end of the peptidyl-tRNA in the P site (Korostelev *et al.*, 2006; Selmer *et al.*, 2006). The universally conserved CCA nucleotides at the 3' end of both tRNAs interact with conserved nucleotides in 23S rRNA. These

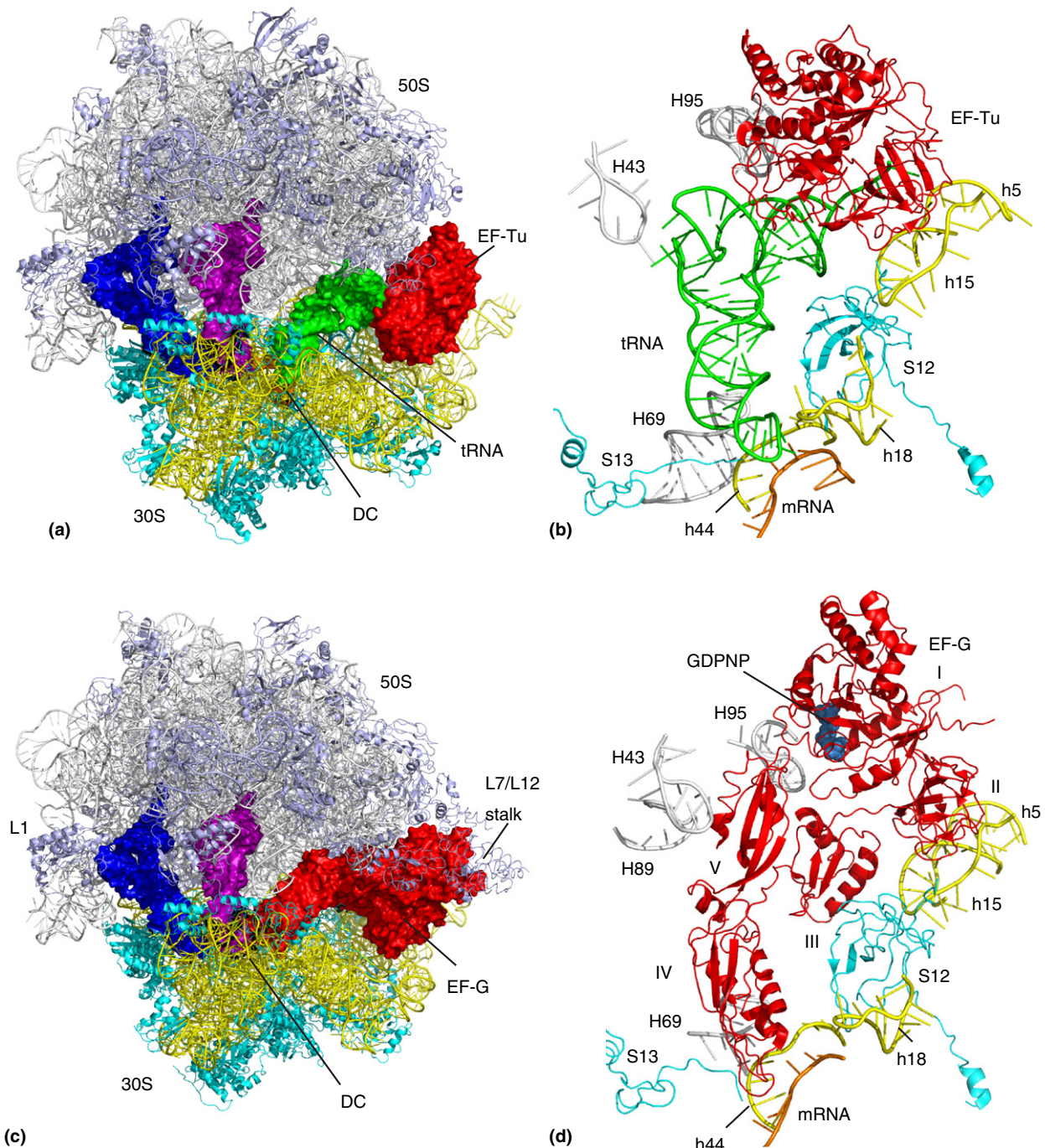


Figure 4 Interactions of elongation factors with the ribosome. (a) Structure of EF-Tu ternary complex bound to the 70S ribosome. EF-Tu ternary complex consist of EF-Tu (red), aminoacyl tRNA (green), and GTP (not shown). Also indicated are the decoding center (DC), P site tRNA (purple) and E site tRNA (blue) (PDB ID: 2XQD and 2XQE). (b) Close-up of EF-Tu ternary complex interacting with mRNA (orange), 23S rRNA (H43, H69, and H95 in white color), 16S rRNA (h5, h15, h18, and h44 in yellow color) and proteins (S12 and S13 in cyan color). (c) Structure of EF-G bound to the 70S ribosome. EF-G (red) interacts with the L7/L12 stalk and with the decoding center (DC) in the post-translocation state (PDB ID: 2WRI and 2WRJ). Close-up of EF-G interacting with 23S rRNA (H43, H69, H89, and H95 in white color), 16S rRNA (h5, h15, h18, and h44 in yellow color) and proteins (S12 and S13 in cyan color). (d) EF-G has five domains (labeled I to V) and domain I binds GTP (the GTP analog GDPNP is indicated by the blue spheres).

include Watson–Crick base pairs by C74 and C75 of the P site tRNA with G2252 and G2251, respectively, of the 23S rRNA (Samaha et al., 1995). Similarly, C75 of the A site tRNA forms a Watson–Crick base pair with G2553 of the 23S rRNA (Kim

and Green, 1999). Interestingly, the peptidyl transferase active site of the ribosome is composed only of 23S rRNA indicating that peptide bond formation is catalyzed by RNA and not by protein (Ban et al., 2000). Consistent with the structural data,

earlier biochemical studies showed that 23S rRNA extracted from the ribosome and free of most ribosomal proteins is capable of catalyzing peptide bond formation (Noller *et al.*, 1992). Most of the catalytic enhancement comes from the ability of the ribosome to precisely position the peptidyl-tRNA and the aminoacyl tRNA for the reaction (Beringer and Rodnina, 2007; Sievers *et al.*, 2004). Peptide bond formation results in a deacylated tRNA in the P site and the peptidyl-tRNA with a new amino acid added in the A site. Now the ribosome can spontaneously undergo a ratchet-like motion caused by the 30S subunit moving in a counter-clockwise manner relative to the 50S subunit (Frank and Agrawal, 2000; Cornish *et al.*, 2008). The ratchet-like movement partly translocates the 3'-end of the P and A site tRNAs into E and P sites, respectively, in the 50S subunit (Valle *et al.*, 2003). This creates hybrid tRNA binding states designated P/E and A/P states (Moazed and Noller, 1989; Spiegel *et al.*, 2007). The deacylated tRNA in the P/E state interacts with the P site in the 30S subunit and with the E site in the 50S subunit. The peptidyl-tRNA in the A/P state interacts with the A site in the 30S subunit and with the P site in the 50S subunit. The ratchet-like motion of the ribosomal subunits and the hybrid state formation by the two tRNAs are important for the translocation step of protein synthesis.

Interactions of EF-G with the Ribosomal RNAs

Translocation of the mRNA and the two tRNAs in the ribosome is catalyzed by EF-G. EF-G is a GTP binding protein and it has five domains (I–V). At present, there is no structure of ribosome•EF-G complex in the pre-translocation state. However, structures of the ribosome•EF-G complex in intermediate (Ratje *et al.*, 2010; Tourigny *et al.*, 2013; Zhou *et al.*, 2013; Pulk and Cate, 2013; Ramrath *et al.*, 2013) and in post-translocation states are available (Gao *et al.*, 2009). EF-G binds at the intersubunit space with domain I of EF-G contacting L12 ribosomal protein (Figure 4(c)). Domain II interacts with nucleotides in helices 5 and 15 of 16S rRNA in the shoulder domain of the 30S subunit (Figure 4(d)). Domain III, which is the GTP binding domain, interacts with helix 5 of 16S rRNA and the sarcin-ricin loop (helix 95) of 23S rRNA. Domain IV interacts with helices 44 and 69 of the 16S and 23S rRNAs, respectively. Finally, domain V interacts with helices 43, 44, 89 and the sarcin-ricin loop of 23S rRNA. How the interactions of EF-G with the ribosome promote GTP hydrolysis and translocation is not fully understood. Studies indicate that EF-G•GTP binds to the ribosome and stabilizes both the ratcheted state of the ribosome and tRNAs in the P/E and A/P hybrid states (Frank and Agrawal, 2000; Cornish *et al.*, 2008; Valle *et al.*, 2003; Spiegel *et al.*, 2007; Agirrezabala *et al.*, 2008). Interaction of EF-G•GTP with the ratcheted ribosome triggers GTP hydrolysis, which accelerates translocation of the two tRNAs from the P/E and A/P hybrid states to E/E and P/P states (Rodnina *et al.*, 1997; Cunha *et al.*, 2013). Thus, EF-G mainly catalyzes the movement of the anticodon arms of the two tRNAs relative to the 30S subunit. The mRNA also moves by one codon towards the E site by virtue of forming base pairs with the anticodons of the P and A site tRNAs. During translocation, both EF-G and the ribosomes undergo dynamic

changes that may bias the directional movement of the tRNAs. For example, domain IV of EF-G moves deeper into the decoding center, the head domain of the 30S subunit undergoes a rotation relative to the body of the 30S subunit, L1 stalk moves closer to the P/E tRNA, and the L7/L12 stalk contacts EF-G are some of the conformational changes that occur during translocation (reviewed in Frank, 2012; Noeske and Cate, 2012; Korostelev *et al.*, 2008). Thus, translocation is one of most complex process orchestrated by the ribosome.

Interaction of Ribosomal RNAs with E-tRNA

Translocation is completed when the deacylated tRNA is in the E/E state and the peptidyl-tRNA is the P/P state. The 50S subunit E site has a higher affinity for deacylated tRNA than for aminoacylated tRNA partly explaining the spontaneous movement of the 3'-end of the P site tRNA into the E site after peptide bond formation. The 3' end of E site tRNA interacts with a deep pocket that is separate from the peptidyl transferase cleft and is composed of helices 11, 74, and 75 of 23S rRNA (Yusupov *et al.*, 2001; Korostelev *et al.*, 2006; Selmer *et al.*, 2006). The anticodon arm of E site tRNA is located between the head and platform domains of the 30S subunit and interacts with the 690 loop, 790 loop, helices 28, 29, and 42 of 16S rRNA. Additionally, ribosomal proteins S7, L1 and L33 contact the E site tRNA. The deacylated tRNA spontaneously dissociates from the E site of the post-translocation state ribosome. This resets the ribosome to the start of the elongation cycle with the ribosome having the peptidyl-tRNA in the P site and vacant E and A sites. It has been debated in the literature whether the anticodon of the E site tRNA remains base paired to the mRNA codon after translocation (Wilson and Nierhaus, 2006). This issue has not been settled with the current structural data because they contain incorrect codon–anticodon combinations that were formed presumably by deacylated tRNAs binding directly to the E site.

Functional Role of 16S and 23S RNAs During Termination

Interactions with Release Factors 1 and 2

The elongation cycle continues to add amino acids to the growing nascent polypeptide chain as dictated by the mRNA codons until a stop codon enters the A site. The nearly universal stop codons UAA, UAG, and UGA are recognized by class I release factors. In bacteria, two class I release factors (RF1 and RF2) recognize the stop codons with overlapping specificity. RF1 recognizes the stop codons UAA and UAG, while RF2 recognizes the stop codons UAA and UGA (Scolnick *et al.*, 1968). RF1/RF2 binds to the ribosome with a stop codon in the A site and catalyze hydrolysis of the ester bond linking the newly synthesized protein to the tRNA in the P site. RF1 and RF2 bind to the ribosomal A site and occupy the same space as the aminoacyl tRNA (Figure 5(a); Klaholz *et al.*, 2003; Rawat *et al.*, 2003; Petry *et al.*, 2005; Weixlbaumer *et al.*, 2008; Laurberg *et al.*, 2008). Residues important for stop codon recognition, including a tripeptide anticodon motif in domain

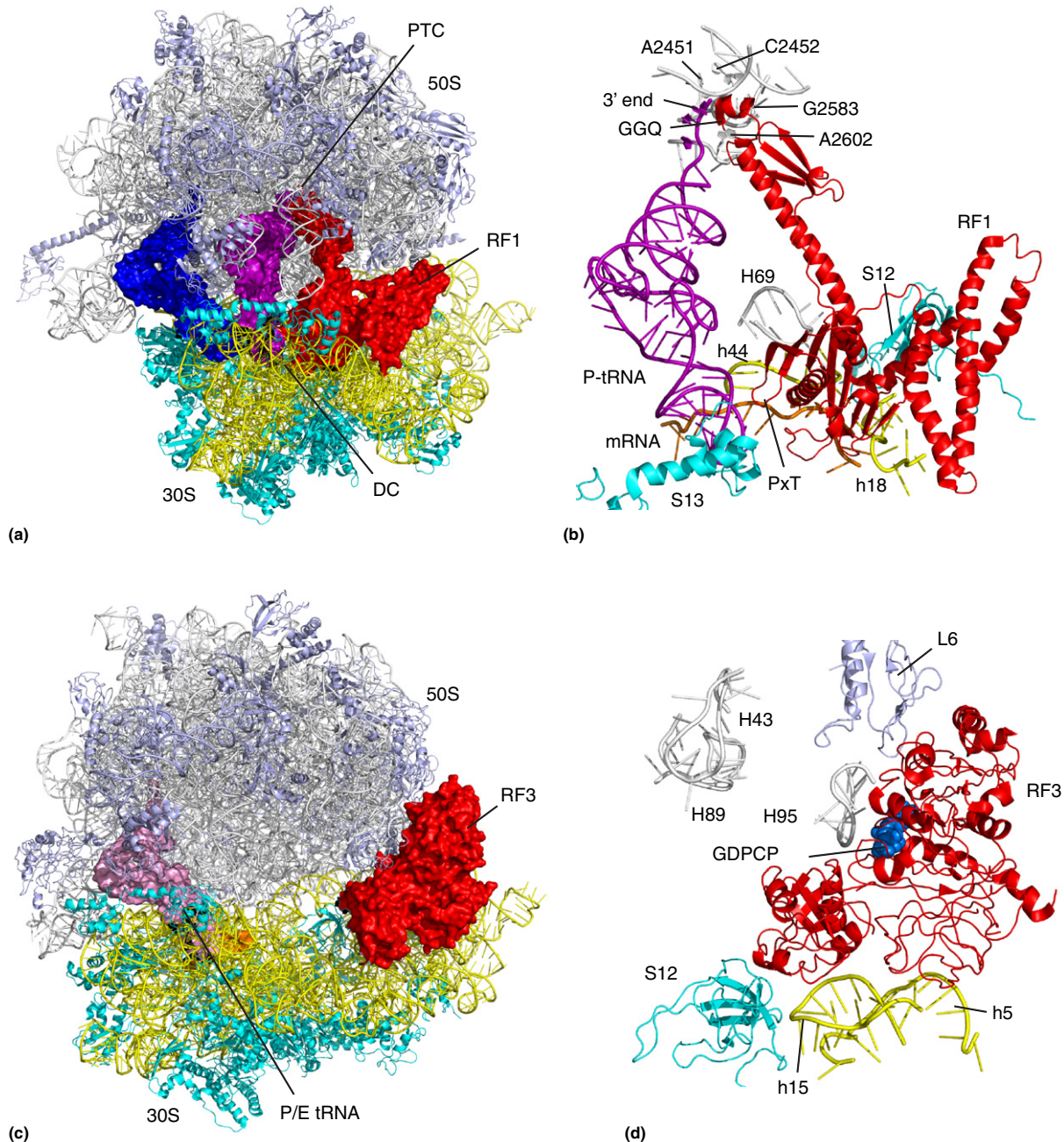


Figure 5 Interactions of release factors 1 and 3 with the ribosome. (a) Structure of RF1 bound to the 70S ribosome. Indicated are RF1 (red), P site tRNA (purple), E site tRNA (blue), the peptidyl transferase center (PTC) and the decoding center (DC) (PDB ID: 3D5A and 3D5B). (b) Close-up of RF1 interacting with 16S rRNA (h18 and h44 in yellow color), 23S rRNA (H69 and bases A2451, C2452, G2583, and A2602 of the PTC; all in white color) and proteins (S12 and S13 in cyan color). The GGQ and PxT motifs of RF1 are indicated. (c) Structure of RF3 bound to the 70S ribosome. RF3 (red) and tRNA in the P/E hybrid state (pink) are shown (PDB ID: 3ZVO and 3ZVP). (d) Close-up of RF3 interacting with 16S rRNA (h5 and h15 in yellow color), 23S rRNA (H43, H89, and H95 in white color), and proteins (S12 in cyan and L6 in light blue). The GTP analog, GDPCP bound to RF3 is indicated by the blue spheres.

II of RF1 and RF2 (PxT in RF1 and SPF in RF2) (Ito *et al.*, 2000) are located in the decoding center of the 30S subunit. In addition, all class I release factors have a GGQ motif in domain III, which is important for peptidyl-tRNA hydrolysis, is located in the peptidyl transferase center next to the 3' end of P site tRNA (Figure 5(b)). The two glycines in the GGQ motif are conserved to correctly position the glutamine in the active

site for peptide hydrolysis. The backbone amide of glutamine contributes directly to catalysis by transition state or product stabilization (Laurberg *et al.*, 2008; Youngman *et al.*, 2008). Interestingly, several of the conserved bases in 23S rRNA that form the peptidyl transferase center are essential for peptide release rather than for peptide bond formation (Youngman *et al.*, 2004). Thus, the peptide release reaction is very sensitive

to disruptions in the active center. This may be exploited by the ribosome to increase the fidelity of termination by coupling stop codon recognition by RF1/RF2 to the proper placement of the GGQ motif in the peptidyl transferase center (He and Green, 2010; Field *et al.*, 2010).

Interactions with Release Factor 3

After peptide release, RF1 and RF2 remain bound to the ribosome. The next step in termination is the removal of RF1 or RF2 from the ribosome by the action of RF3. RF3 is a member of the GTPase superfamily and consists of three domains. Domain I of RF3 binds GTP and is known as the G domain. The current model is that RF3•GDP binds to the release complex consisting of a deacylated tRNA in the P site and either RF1 or RF2 in the A site (Zavialov *et al.*, 2001). RF3 then exchanges GDP for GTP on the ribosome. The exchange reaction results in an increase in the affinity of RF3•GTP for the ribosome and promotes the dissociation of RF1/RF2. RF3 hydrolyzes GTP, switches to its low affinity GDP form, and leaves the ribosome. However, more recent studies indicate that RF3•GTP binds to the ribosome to trigger the release of RF1/RF2 from the ribosome (Peske *et al.*, 2013).

Structural studies showed that RF3 binds at the entry to the intersubunit cavity and its binding site overlaps with the binding sites for EF-Tu ternary complex and EF-G (Figure 5(c); Klaholz *et al.*, 2004; Jin *et al.*, 2011; Zhou *et al.*, 2012b). As expected, domain I of RF3 interacts with the sarcin-ricin loop (helix 95) of 23S rRNA and protein L6 (Figure 5(d)). A conserved histidine in domain I of RF3 that is important for GTP hydrolysis is located close to the sarcin-ricin

loop. Domains II and III interact with helices 5 and 15 of 16S rRNA and protein S12. The binding of RF3 to the ribosome induces the ratchet-like rotation of the subunits and the movement of the deacylated tRNA from P/P state to P/E state (Jin *et al.*, 2011; Sternberg *et al.*, 2009). Structural changes also occur in the GTPase-associated center (helices 42–44 of 23S rRNA) and in the decoding center (Jin *et al.*, 2011; Zhou *et al.*, 2012b). The ratcheted state of the ribosome induced by RF3 creates steric clash between the head domain of the 30S subunit and domain IV of RF1/RF2, and between the L11 stalk of the 50S subunit and domain I of RF1/RF2. These unfavorable interactions between RF1/RF2 and the ribosome are proposed trigger their release from the ribosome.

Functional Role of 16S and 23S RNAs During Recycling

Interactions with Ribosome Recycling Factor

In the final stage of protein synthesis the ribosomes with an mRNA and a tRNA in the P site are separated into 30S and 50S subunits by the activities of RRF and EF-G (Zavialov *et al.*, 2005; Peske *et al.*, 2005). RRF has two domains and forms an L-shaped structure, which mimics the overall shape of a tRNA (Selmer, 1999). However, RRF does not bind to the same sites as tRNAs on the ribosome. Instead, RRF binds in the subunit interface with domain I spanning the A and P sites of the 50S subunit and domain II is close to protein S12 of the 30S subunit (Figure 6(a); Lancaster *et al.*, 2002; Wilson *et al.*, 2004; Agrawal *et al.*, 2004; Weixlbaumer *et al.*, 2007;

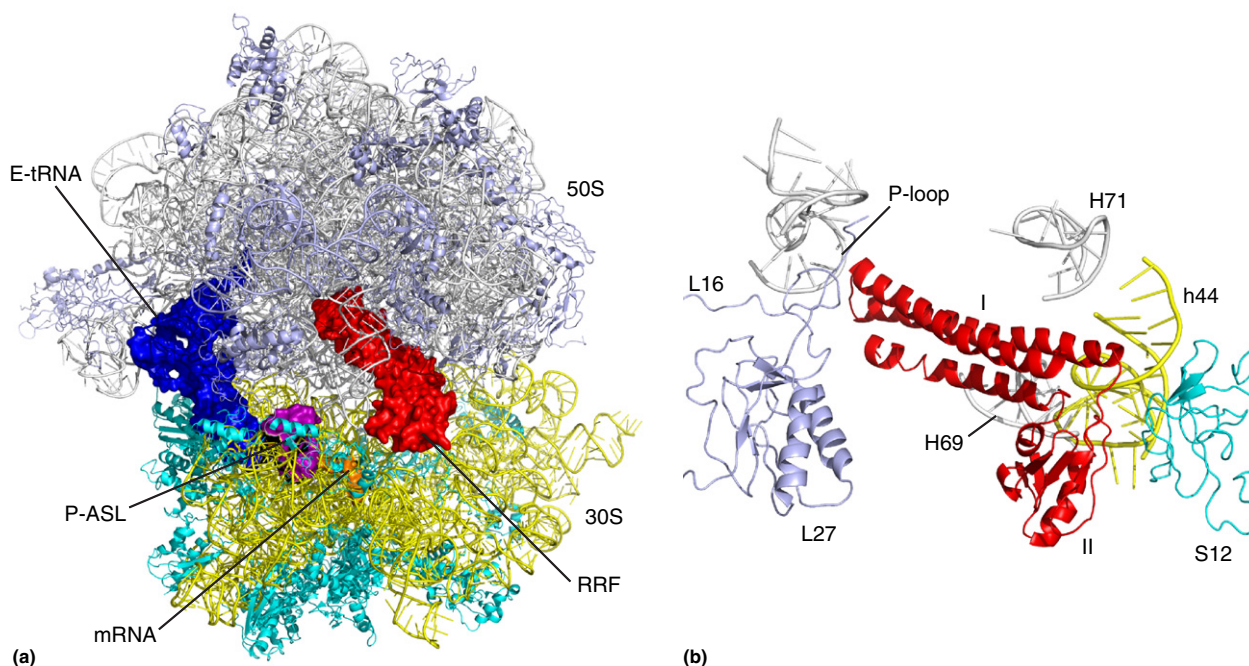


Figure 6 Interactions of ribosome recycling factor with the ribosome. (a) Structure of RRF bound to the 70S ribosome. Indicated are RRF (red), mRNA (orange), P site anticodon stem-loop (P-ASL in purple), and E site tRNA (blue) (PDB ID: 2V46 and 2V48). (b) Close-up of RRF interacting with 16S rRNA (h44 in yellow color), 23S rRNA (H69, H71 and the P-loop; all in white color) and proteins (S12 in cyan, L16 and L27 in light blue).

Borovinskaya *et al.*, 2007). The tip of domain I contacts the P-loop of 23S rRNA (residues 2246–2258) and will clash with the acceptor arm of tRNA in the P/P state (Figure 6(b)). This suggests that the binding of RRF to the ribosome will favor the movement of the P-tRNA to the P/E hybrid state. Additionally, RRF may change the conformation of helix 69 of 23S rRNA, which forms important intersubunit bridges (Lancaster *et al.*, 2002; Wilson *et al.*, 2004; Agrawal *et al.*, 2004). There is no structure of RRF and EF-G bound simultaneously to the ribosome. Therefore, the exact mechanism for how RRF and EF-G trigger subunit dissociation is not known. It is possible that the binding of both RRF and EF-G to the ribosome may drive the P site tRNA to the P/E hybrid state and the ribosome to undergo the ratchet-like rotation, which ultimately disrupts the intersubunit bridges and cause the 30S and 50S subunits to separate (Gao *et al.*, 2005). IF3 then binds to the 30S subunit, preventing it from reassociating with the 50S subunit and also promoting the dissociation of the mRNA and tRNA from the 30S subunit (Zavialov *et al.*, 2005; Peske *et al.*, 2005). The 30S subunit•IF3 complex is free to initiate protein synthesis on a new mRNA.

Functional Role of 5S rRNA

The 5S rRNA is the smallest of the three rRNAs and is universally present in all organisms. Although 5S rRNA is essential for the activity of the ribosome, its functional role is still a mystery. The 5S rRNA is located in the central protuberance of the large ribosomal subunit and it interacts with 23S rRNA and ribosomal proteins L5, L16, L18, L25, L27, L30, L33, and L35 (Figure 2; Ban *et al.*, 2000; Selmer *et al.*, 2006; Korostelev *et al.*, 2006; Schuwirth *et al.*, 2005). The 5S rRNA and associated proteins contacts domain II (helices H38, H39) and domain V (H83–H85) of 23S rRNA, thus connecting these two structural domains of the 23S rRNA. This unique location for the 5S rRNA has led to the proposal that it may regulate the activities of these two functional centers of the ribosome (reviewed in Dinman, 2005; Gongadze, 2012). Additionally, since the central protuberance of the 50S subunit interacts with the head domain of the 30S subunit to form two intersubunit bridges (B1a and B1b) (Selmer *et al.*, 2006; Korostelev *et al.*, 2006; Schuwirth *et al.*, 2005), it is possible that the 5S rRNA-protein complex plays a role in subunit association. Finally, the head domain of the 30S subunit undergoes a rotation during mRNA–tRNA translocation (Zhang *et al.*, 2009), and the 5S rRNA-protein complex may coordinate this motion of the 30S subunit. More studies are needed to understand the precise function of this evolutionarily conserved RNA found in all three branches of life.

Acknowledgments

This work was supported by an NSF grant (MCB 1158127).

See also: Protein Synthesis/Degradation: Translation: Components, Initiation, Elongation, Termination, and Regulation

References

- Agirrezabala, X., *et al.*, 2008. Visualization of the hybrid state of tRNA binding promoted by spontaneous ratcheting of the ribosome. *Molecular Cell* 32 (2), 190–197.
- Agrawal, R.K., *et al.*, 2004. Visualization of ribosome-recycling factor on the *Escherichia coli* 70S ribosome: Functional implications. *Proceedings of the National Academy of Sciences of the United States of America* 101 (24), 8900–8905.
- Allen, G.S., Zavialov, A., Gursky, R., Ehrenberg, M., Frank, J., 2005. The cryo-EM structure of a translation initiation complex from *Escherichia coli*. *Cell* 121 (5), 703–712.
- Antoun, A., Pavlov, M.Y., Lovmar, M., Ehrenberg, M., 2006a. How initiation factors maximize the accuracy of tRNA selection in initiation of bacterial protein synthesis. *Molecular Cell* 23 (2), 183–193.
- Antoun, A., Pavlov, M.Y., Lovmar, M., Ehrenberg, M., 2006b. How initiation factors tune the rate of initiation of protein synthesis in bacteria. *The EMBO Journal* 25 (11), 2539–2550.
- Ban, N., Nissen, P., Hansen, J., Moore, P.B., Steitz, T.A., 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* (New York, NY) 289 (5481), 905–920.
- Beringer, M., Rodnina, M.V., 2007. The ribosomal peptidyl transferase. *Molecular Cell* 26 (3), 311–321.
- Boelens, R., Gualerzi, C.O., 2002. Structure and function of bacterial initiation factors. *Current Protein & Peptide Science* 3 (1), 107–119.
- Borovinskaya, M.A., *et al.*, 2007. Structural basis for aminoglycoside inhibition of bacterial ribosome recycling. *Nature Structural & Molecular Biology* 14 (8), 727–732.
- Brunel, C., *et al.*, 1995. Stabilised secondary structure at a ribosomal binding site enhances translational repression in *E. coli*. *Journal of Molecular Biology* 253 (2), 277–290.
- Carter, A.P., *et al.*, 2001. Crystal structure of an initiation factor bound to the 30S ribosomal subunit. *Science* (New York, NY) 291 (5503), 498–501.
- Cornish, P.V., Ermolenko, D.N., Noller, H.F., Ha, T., 2008. Spontaneous intersubunit rotation in single ribosomes. *Molecular Cell* 30 (5), 578–588.
- Cunha, C., *et al.*, 2013. Dual use of GTP hydrolysis by elongation factor G on the ribosome. *Translation* 1 (1), e24315.
- Dallas, A., Noller, H.F., 2001. Interaction of translation initiation factor 3 with the 30S ribosomal subunit. *Molecular Cell* 8 (4), 855–864.
- Demeshekina, N., Jenner, L., Westhof, E., Yusupov, M., Yusupova, G., 2012. A new understanding of the decoding principle on the ribosome. *Nature* 484, 256–259.
- Dinman, J.D., 2005. 5S rRNA: Structure and Function from Head to Toe. *International Journal of Biomedical Science: IJBS* 1 (1), 2–7.
- Elvekrog, M.M., Gonzalez, R.L., 2013. Conformational selection of translation initiation factor 3 signals proper substrate selection. *Nature Structural & Molecular Biology* 20 (5), 628–633.
- Field, A., Hetrick, B., Mathew, M., Joseph, S., 2010. Histidine 197 in release factor 1 is essential for a site binding and peptide release. *Biochemistry* 49 (43), 9385–9390.
- Frank, J., 2012. Intermediate states during mRNA–tRNA translocation. *Current Opinion in Structural Biology* 22 (6), 778–785.
- Frank, J., Agrawal, R.K., 2000. A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature* 406 (6793), 318–322.
- Frank, J., Gao, H., Sengupta, J., Gao, N., Taylor, D.J., 2007. The process of mRNA–tRNA translocation. *Proceedings of the National Academy of Sciences of the United States of America* 104 (50), 19671–19678.
- Frank, J., Gonzalez, R.L., 2010. Structure and dynamics of a processive Brownian motor: The translating ribosome. *Annual Review of Biochemistry* 79, 381–412.
- Gabashvili, I.S., *et al.*, 2000. Solution structure of the *E. coli* 70S ribosome at 11.5 Å resolution. *Cell* 100 (5), 537–549.
- Gao, H., *et al.*, 2003. Study of the structural dynamics of the *E. coli* 70S ribosome using real-space refinement. *Cell* 113 (6), 789–801.
- Gao, N., *et al.*, 2005. Mechanism for the disassembly of the posttermination complex inferred from cryo-EM studies. *Molecular Cell* 18 (6), 663–674.
- Gao, Y.-G., *et al.*, 2009. The structure of the ribosome with elongation factor G trapped in the posttranslocational state. *Science* (New York, NY) 326 (5953), 694–699.
- Gongadze, G.M., 2012. 5S rRNA and ribosome. *Biochemistry (Moscow)* 76 (13), 1450–1464.
- He, S.L., Green, R., 2010. Visualization of codon-dependent conformational rearrangements during translation termination. *Nature Structural & Molecular Biology* 17 (4), 465–470.

- Hoang, L., Fredrick, K., Noller, H.F., 2004. Creating ribosomes with an all-RNA 30S subunit P site. *Proceedings of the National Academy of Sciences of the United States of America* 101 (34), 12439–12443.
- Ito, K., Uno, M., Nakamura, Y., 2000. A tripeptide “anticodon” deciphers stop codons in messenger RNA. *Nature* 403 (6770), 680–684.
- Jin, H., Kelley, A.C., Ramakrishnan, V., 2011. Crystal structure of the hybrid state of ribosome in complex with the guanosine triphosphatase release factor 3. *Proceedings of the National Academy of Sciences of the United States of America* 108 (38), 15798–15803.
- Kim, D.F., Green, R., 1999. Base-pairing between 23S rRNA and tRNA in the ribosomal A site. *Molecular Cell* 4 (5), 859–864.
- Klaholz, B.P., *et al.*, 2003. Structure of the *Escherichia coli* ribosomal termination complex with release factor 2. *Nature* 421 (6918), 90–94.
- Klaholz, B.P., Myasnikov, A.G., van Heel, M., 2004. Visualization of release factor 3 on the ribosome during termination of protein synthesis. *Nature* 427 (6977), 862–865.
- Korostelev, A., *et al.*, 2007. Interactions and dynamics of the Shine Dalgarno helix in the 70S ribosome. *Proceedings of the National Academy of Sciences of the United States of America* 104 (43), 16840–16843.
- Korostelev, A., Ermolenko, D.N., Noller, H.F., 2008. Structural dynamics of the ribosome. *Current Opinion in Chemical Biology* 12 (6), 674–683.
- Korostelev, A., Trakhanov, S., Laurberg, M., Noller, H.F., 2006. Crystal structure of a 70S ribosome-tRNA complex reveals functional interactions and rearrangements. *Cell* 126 (6), 1065–1077.
- La Teana, A., Pon, C.L., Gualerzi, C.O., 1996. Late events in translation initiation. Adjustment of fMet-tRNA in the ribosomal P-site. *Journal of Molecular Biology* 256 (4), 667–675.
- Lancaster, L., Kiel, M.C., Kaji, A., Noller, H.F., 2002. Orientation of ribosome recycling factor in the ribosome from directed hydroxyl radical probing. *Cell* 111 (1), 129–140.
- Lancaster, L., Noller, H.F., 2005. Involvement of 16S rRNA nucleotides G1338 and A1339 in discrimination of initiator tRNA. *Molecular Cell* 20 (4), 623–632.
- Laurberg, M., *et al.*, 2008. Structural basis for translation termination on the 70S ribosome. *Nature* 454 (7206), 852–857.
- McCutcheon, J.P., *et al.*, 1999. Location of translational initiation factor IF3 on the small ribosomal subunit. *Proceedings of the National Academy of Sciences of the United States of America* 96 (8), 4301–4306.
- Milon, P., *et al.*, 2010. The ribosome-bound initiation factor 2 recruits initiator tRNA to the 30S initiation complex. *Nature Publishing Group* 11 (4), 312–316.
- Milón, P., Maracci, C., Filonava, L., Gualerzi, C.O., Rodnina, M.V., 2012. Real-time assembly landscape of bacterial 30S translation initiation complex. *Nature Structural & Molecular Biology* 19 (6), 609–615.
- Moazed, D., Noller, H.F., 1989. Intermediate states in the movement of transfer RNA in the ribosome. *Nature* 342 (6246), 142–148.
- Myasnikov, A.G., *et al.*, 2005. Conformational transition of initiation factor 2 from the GTP- to GDP-bound state visualized on the ribosome. *Nature Structural & Molecular Biology* 12 (12), 1145–1149.
- Nissen, P., Hansen, J., Ban, N., Moore, P.B., Steitz, T.A., 2000. The structural basis of ribosome activity in peptide bond synthesis. *Science (New York, NY)* 289 (5481), 920–930.
- Noeske, J., Cate, J.H., 2012. Structural basis for protein synthesis: Snapshots of the ribosome in motion. *Current Opinion in Structural Biology* 22 (6), 743–749.
- Noller, H.F., Hoang, L., Fredrick, K., 2005. The 30S ribosomal P site: A function of 16S rRNA. *FEBS Letters* 579 (4), 855–858.
- Noller, H.F., Hoffarth, V., Zimniak, L., 1992. Unusual resistance of peptidyl transferase to protein extraction procedures. *Science (New York, NY)* 256 (5062), 1416–1419.
- Ogle, J.M., *et al.*, 2001. Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science (New York, NY)* 292 (5518), 897–902.
- Ogle, J.M., Murphy, F.V., Tarry, M.J., Ramakrishnan, V., 2002. Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* 111 (5), 721–732.
- Peske, F., Kuhlencoetter, S., Rodnina, M.V., Wintermeyer, W., 2013. Timing of GTP binding and hydrolysis by translation termination factor RF3. *Nucleic Acids Research* 42 (3), 1812–1820.
- Peske, F., Rodnina, M.V., Wintermeyer, W., 2005. Sequence of steps in ribosome recycling as defined by kinetic analysis. *Molecular Cell* 18 (4), 403–412.
- Petry, S., *et al.*, 2005. Crystal structures of the ribosome in complex with release factors RF1 and RF2 bound to a cognate stop codon. *Cell* 123 (7), 1255–1266.
- Petry, S., Weixlbaumer, A., Ramakrishnan, V., 2008. The termination of translation. *Current Opinion in Structural Biology* 18 (1), 70–77.
- Pulk, A., Cate, J.H.D., 2013. Control of ribosomal subunit rotation by elongation factor G. *Science (New York, NY)* 340 (6140), 1235970.
- Ramrath, D.J.F., *et al.*, 2013. Visualization of two transfer RNAs trapped in transit during elongation factor G-mediated translocation. *Proceedings of the National Academy of Sciences of the United States of America* 110 (52), 20964–20969.
- Ratje, A.H., *et al.*, 2010. Head swivel on the ribosome facilitates translocation by means of intra-subunit tRNA hybrid sites. *Nature* 468 (7324), 713–716.
- Rawat, U.B.S., *et al.*, 2003. A cryo-electron microscopic study of ribosome-bound termination factor RF2. *Nature* 421 (6918), 87–90.
- Rodnina, M.V., Gromadski, K.B., Kothe, U., Wieden, H.J., 2005. Recognition and selection of tRNA in translation. *FEBS Letters* 579 (4), 938–942.
- Rodnina, M.V., Savelsbergh, A., Katunin, V.I., Wintermeyer, W., 1997. Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome. *Nature* 385 (6611), 37–41.
- Rodnina, M.V., Wintermeyer, W., 2001. Fidelity of aminoacyl-tRNA selection on the ribosome: Kinetic and structural mechanisms. *Annual Review of Biochemistry* 70, 415–435.
- Roll-Mecak, A., Cao, C., Dever, T.E., Burley, S.K., 2000. X-Ray structures of the universal translation initiation factor IF2/eIF5B: Conformational changes on GDP and GTP binding. *Cell* 103 (5), 781–792.
- Samaha, R.R., Green, R., Noller, H.F., 1995. A base pair between tRNA and 23S rRNA in the peptidyl transferase centre of the ribosome. *Nature* 377 (6547), 309–314.
- Schmeing, T.M., *et al.*, 2009. The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science (New York, NY)* 326 (5953), 688–694.
- Schuwirth, B.S., *et al.*, 2005. Structures of the bacterial ribosome at 3.5 Å resolution. *Science (New York, NY)* 310 (5749), 827–834.
- Scolnick, E., Tompkins, R., Caskey, T., Nirenberg, M., 1968. Release factors differing in specificity for terminator codons. *Proceedings of the National Academy of Sciences of the United States of America* 61 (2), 768–774.
- Selmer, M., 1999. Crystal structure of thermotoga maritima ribosome recycling factor: A tRNA mimic. *Science (New York, NY)* 286 (5448), 2349–2352.
- Selmer, M., *et al.*, 2006. Structure of the 70S ribosome complexed with mRNA and tRNA. *Science (New York, NY)* 313 (5795), 1935–1942.
- Shine, J., Dalgarno, L., 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America* 71 (4), 1342–1346.
- Shoji, S., Walker, S.E., Fredrick, K., 2009. Ribosomal translocation: One step closer to the molecular mechanism. *ACS Chemical Biology* 4 (2), 93–107.
- Sievers, A., Beringer, M., Rodnina, M.V., Wolfenden, R., 2004. The ribosome as an entropy trap. *Proceedings of the National Academy of Sciences of the United States of America* 101 (21), 7897–7901.
- Simonetti, A., *et al.*, 2008. Structure of the 30S translation initiation complex. *Nature* 455 (7211), 416–420.
- de Smit, M.H., Van Duin, J., 1990. Secondary structure of the ribosome binding site determines translational efficiency: A quantitative analysis. *Proceedings of the National Academy of Sciences of the United States of America* 87 (19), 7668–7672.
- Spiegel, P.C., Ermolenko, D.N., Noller, H.F., 2007. Elongation factor G stabilizes the hybrid-state conformation of the 70S ribosome. *RNA (New York, NY)* 13 (9), 1473–1482.
- Sternberg, S.H., *et al.*, 2009. Translation factors direct intrinsic ribosome dynamics during translation termination and ribosome recycling. *Nature Structural & Molecular Biology* 16 (8), 861–868.
- Studer, S.M., Joseph, S., 2006. Unfolding of mRNA secondary structure by the bacterial translation initiation complex. *Molecular Cell* 22 (1), 105–115.
- Takyar, S., Hickerson, R.P., Noller, H.F., 2005. mRNA helicase activity of the ribosome. *Cell* 120 (1), 49–58.
- Tomsic, J., *et al.*, 2000. Late events of translation initiation in bacteria: A kinetic analysis. *The EMBO Journal* 19 (9), 2127–2136.
- Tourigny, D.S., Fernández, I.S., Kelley, A.C., Ramakrishnan, V., 2013. Elongation factor G bound to the ribosome in an intermediate state of translocation. *Science (New York, NY)* 340 (6140), 1235490.
- Tsai, A., *et al.*, 2012. Heterogeneous pathways and timing of factor departure during translation initiation. *Nature* 484, 256–259.
- Valle, M., *et al.*, 2003. Locking and unlocking of ribosomal motions. *Cell* 114 (1), 123–134.
- Voorhees, R.M., Schmeing, T.M., Kelley, A.C., Ramakrishnan, V., 2010. The mechanism for activation of GTP hydrolysis on the ribosome. *Science (New York, NY)* 330 (6005), 835–838.
- Weixlbaumer, A., *et al.*, 2007. Crystal structure of the ribosome recycling factor bound to the ribosome. *Nature Structural & Molecular Biology* 14 (8), 733–737.
- Weixlbaumer, A., *et al.*, 2008. Insights into translational termination from the structure of RF2 bound to the ribosome. *Science (New York, NY)* 322 (5903), 953–956.

- Wilson, D., *et al.*, 2004. X-ray crystallography study on ribosome recycling: The mechanism of binding and action of RRF on the 50S ribosomal subunit. *The EMBO Journal* 24 (2), 251–260.
- Wilson, D.N., Nierhaus, K.H., 2006. The E-site story: The importance of maintaining two tRNAs on the ribosome during protein synthesis. *Cellular and Molecular Life Sciences: CMLS* 63 (23), 2725–2737.
- Youngman, E.M., Brunelle, J.L., Kochaniak, A.B., Green, R., 2004. The active site of the ribosome is composed of two layers of conserved nucleotides with distinct roles in peptide bond formation and peptide release. *Cell* 117 (5), 589–599.
- Youngman, E.M., McDonald, M.E., Green, R., 2008. Peptide release on the ribosome: Mechanism and implications for translational control. *Annual review of microbiology* 62, 353–373.
- Yusupov, M.M., *et al.*, 2001. Crystal structure of the ribosome at 5.5Å resolution. *Science (New York, NY)* 292 (5518), 883–896.
- Yusupova, G.Z., Yusupov, M.M., Cate, J.H., Noller, H.F., 2001. The path of messenger RNA through the ribosome. *Cell* 106 (2), 233–241.
- Zavialov, A.V., Buckingham, R.H., Ehrenberg, M., 2001. A posttermination ribosomal complex is the guanine nucleotide exchange factor for peptide release factor RF3. *Cell* 107 (1), 115–124.
- Zavialov, A.V., Hauryliuk, V.V., Ehrenberg, M., 2005. Splitting of the posttermination ribosome into subunits by the concerted action of RRF and EF-G. *Molecular Cell* 18 (6), 675–686.
- Zhang, W., Dunkle, J.A., Cate, J.H.D., 2009. Structures of the Ribosome in Intermediate States of Ratcheting. *Science (New York, NY)* 325 (5943), 1014–1017.
- Zhou, J., Korostelev, A., Lancaster, L., Noller, H.F., 2012a. Crystal structures of 70S ribosomes bound to release factors RF1, RF2 and RF3. *Current Opinion in Structural Biology* 22 (6), 733–742.
- Zhou, J., Lancaster, L., Donohue, J.P., Noller, H.F., 2013. Crystal structures of EF-G-ribosome complexes trapped in intermediate states of translocation. *Science (New York, NY)* 340 (6140), 1236086.
- Zhou, J., Lancaster, L., Trakhanov, S., Noller, H.F., 2012b. Crystal structure of release factor RF3 trapped in the GTP state on a rotated conformation of the ribosome. *RNA (New York, NY)* 18 (2), 230–240.

Comparison of Bacterial and Eukaryotic Replisome Components

NY Yao and ME O'Donnell, The Rockefeller University, New York, NY, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

AAA + protein AAA + stands for ATPases Associated with diverse cellular Activities. They contain a similar 'fold' of about 220 residues.

Clamp loader The clamp loader is a multiprotein assembly that uses ATP to assemble circular sliding clamps onto DNA. The clamp loader binds three substrates in this reaction, DNA, the sliding clamp, and ATP.

Distributive An enzyme is said to be distributive when it only performs one catalytic turnover for each substrate binding event.

DNA helicase DNA helicases are enzymes that couple ATP hydrolysis to separation of the strands of duplex DNA.

DNA polymerase DNA polymerases extend DNA by using dNTP substrates which they add to the 3' OH terminus of a preexisting RNA or DNA chain that is annealed to a template ssDNA strand. The DNA polymerase 'reads' the sequence of nucleotide bases in the template strand and matches complementary dNTPs to it.

LUCA This term is an abbreviation for 'last universal common ancestor' and is often thought of as a complete cell, from which all modern cells evolved. The nature of LUCA is highly disputed.

OB fold This is a common polypeptide fold within oligonucleotide and oligosaccharide binding proteins. These proteins typically bind single-strand DNA, RNA, or protein.

Okazaki fragment The term 'Okazaki fragment,' sometimes called 'lagging strand fragment' refers to the short segments of DNA that are synthesized discontinuously along the lagging strand. They are typically tipped at the 5' terminus with RNA, made by primase, and the RNA requires removal and replacement by DNA before Okazaki fragments can be joined together by the action of ligase.

Processive This term refers to the number of catalytic turnover events in one substrate binding event. For example, in polymerase action, the processivity of a polymerase is characterized by the number of dNTPs it incorporates into a DNA molecule before it dissociates from that molecule. In helicase action, this term refers to the number of base pairs that are melted (broken apart) in one helicase binding event on the DNA that it unwinds.

Proofreading exonuclease This is a 3'-5' exonuclease that is associated with the DNA polymerase, and 'proofreads' the product of nucleotide incorporation. If the polymerase makes a mistake, the new 3' terminus will not be correctly paired, and the 3'-5' exonuclease will remove the

mismatched 3' terminal nucleotide, enabling the DNA polymerase to try again.

Protein family These are proteins that are grouped together, most often due to having homologous sequences and therefore similar 3D structures. The similar sequence and structure is interpreted as having a common evolutionary ancestor with other proteins in the same family.

Protein fold This refers to the topological way that the polypeptide chain of a protein folds into a three-dimensional shape. A region within a protein that has the same 3D folding pattern as within another protein is referred to as having the same 'fold.'

RecA fold These are proteins that have a similar protein folding pattern as the RecA recombinase.

Replication Replication refers to the process of converting one DNA duplex into two new DNA duplexes.

Replication fork This is a commonly used term to indicate the point at which the two strands of parental DNA have been separated for duplication of the two daughter strands.

Replisome The term 'replisome' refers to the group of proteins that function together at a replication fork to convert one duplex into two new daughter duplexes.

Ribonucleotide reductase This is an enzyme that converts rNMPs into dNTPs, by removing the 2' hydroxyl group from the ribose, and replacing it with a proton. The fact that dNTPs are made from rNTPs, and are not synthesized de novo, is consistent with the RNA world hypothesis that invokes the presence of RNA before DNA.

RNA primase Primase is an enzyme that starts nucleic acid polymers with complementary sequence to a template single-strand, which it uses as a substrate. The primer oligonucleotides it creates are typically made of RNA and are short, 4–12 nucleotides.

Sliding clamp The sliding clamp is a term used to refer to the ring-shaped proteins that encircle DNA and slide freely along the duplex. Sliding clamps bind DNA polymerase, acting as a mobile tether to hold the polymerase to DNA for many rounds of dNTP incorporation.

Single-strand binding protein (SSB) SSB refers to a 'single-strand DNA-binding protein.' There are many types of SSBs. But when referred to in the context of DNA replication, SSB is reserved for the bacterial SSB homotetramer, required for replication. The eukaryotic and archaeal functional equivalent is referred to as replication protein A (RPA), a heterotrimer with distinct structure from bacterial SSB.

Introduction

The elegant double helical structure of deoxyribonucleic acid (DNA) belied the secret of just how difficult it would be to

replicate this simple structure. Watson and Crick recognized the need for considerable untwisting of the DNA strands for its duplication ([Watson and Crick, 1953a,b](#)), which we now know are resolved by topoisomerases, but they could not have

envisioned just how many proteins are actually required to perform the job of DNA replication. Here we provide an overview of the current state of knowledge about the proteins that function at replication forks in bacteria and eukaryotes.

The fact that the DNA strands are complementary led Watson and Crick to suggest that DNA may self-replicate without needing a protein, although they were open to the idea of an enzyme that ‘reads’ each strand to replicate DNA (Watson and Crick, 1953a). This issue was put to rest with the discovery of DNA polymerase (Pol) I by Arthur Kornberg (Lehman *et al.*, 1958). We now know that all cells contain multiple types of DNA Pols for replication and for DNA repair (Hubscher and Maga, 2011; Kornberg and Baker, 1992; Yang and Woodgate, 2007). However, there are many other, completely unanticipated proteins needed for replication in all cell types. Two of these are the circular sliding clamp, which tethers DNA polymerases to DNA, and the clamp loader machine that assembles these clamps onto DNA (Kelch *et al.*, 2012; O'Donnell and Kuriyan, 2006). Another is ribonucleic acid (RNA) primase, required to initiate DNA synthesis (Frick and Richardson, 2001; Kornberg and Baker, 1992). Replication also requires a DNA helicase to melt the parental duplex (Kornberg and Baker, 1992). Another essential protein, ubiquitous in all life forms, is a single-strand (ss) DNA binding protein that protects and organizes single-stranded DNA (ssDNA) produced as the helicase unwinds DNA (Kornberg and Baker, 1992). We came to understand that these proteins function in a coordinated manner, as gears within a ‘replisome’ machine (Alberts, 1984). And with time, learned that the proteins of the replisome are not fixed, but instead are highly dynamic and can interchange parts during replication (Langston *et al.*, 2009).

Comparisons of bacterial, eukaryotic, and archaeal genomes reveal that the sequences of the core translation and transcription machinery are conserved in all three domains of life. Given the conservation in these information processing systems it is surprising that the core replication enzymes of bacteria and eukaryotes/archaea are not conserved (Leipe *et al.*, 1999). Replication forks require the direct action of DNA polymerases, RNA primase, DNA helicase, and single-strand binding protein (SSB), yet these central proteins in bacteria and in eukaryotes share no common ancestor, while these same components in archaea are related to those in eukaryotes (Forterre, 2013; Leipe *et al.*, 1999). The implication is that the last universal common ancestor (LUCA) cell did not replicate DNA the same way as any modern cell, and may have had an RNA genome (Forterre *et al.*, 2004; Leipe *et al.*, 1999). Thus the DNA replication machinery developed later, after the branch point of bacteria and archaea/eukaryotes (Figure 1). Bacteria evolved one solution to replicate DNA while eukaryotes/archaea developed a different one.

The discovery of catalytic RNA (Kruger *et al.*, 1982) led to the RNA world hypothesis (Woese, 1968), which posits that the first enzymatic functions were performed by ribozymes, most of which were later replaced by proteins. The ribosome is a prime example of a ribozyme in all modern day cells. Sequence homologies reveal that the ribosomal RNAs, transfer RNAs (tRNAs), and aminoacyl tRNA synthetases in all domains of life share a common ancestor. All cells also use the same genetic code. These facts imply that protein synthesis was

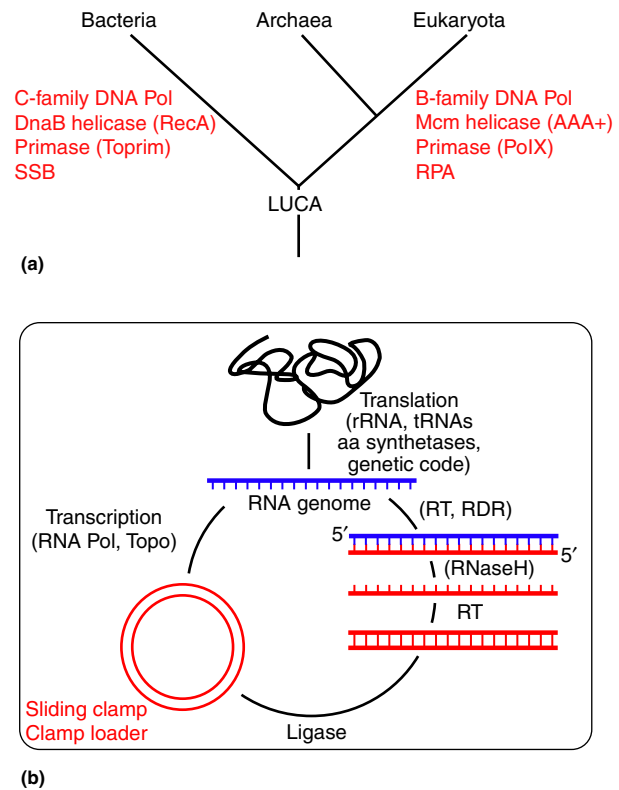


Figure 1 Independent evolution of two DNA replication systems in bacteria and eukaryotes/archaea. (a) Common genes for transcription (RNA polymerase) and translation (rRNA, tRNAs, and amino acyl (aa) synthetases), and the same genetic code among all three domains of life indicate that LUCA had a fully formed transcription and translation apparatus. The universal homology of ribonucleotide reductase (RDR), ligase, and DNA dependent RNA polymerase indicate LUCA had DNA. However, the central enzymes of replication are nonhomologous between bacteria and the eukaryotic/archaeal domains of life, indicating that bacterial and eukaryotic replication enzymes evolved independently. (b) One hypothesis of information transfer in LUCA that included DNA (illustrated) is that LUCA had an RNA genome which was reverse transcribed into duplex DNA, as occurs in the life cycle of many modern RNA viruses. The DNA is then transcribed to form RNA for the genome and for translation.

highly developed in LUCA. In fact, the RNA polymerases of bacteria and eukaryotes are also homologous in all cell types, and they use DNA as a substrate. Ribonucleotide reductase, which converts NMPs to dNMPs is also conserved in all domains of life. These facts strongly suggest that LUCA had DNA, even if it did use RNA as the genome repository of information.

The finding that key replication enzymes of bacteria and eukaryotes are not homologous, yet LUCA had DNA, suggests that LUCA used another means of DNA synthesis (Leipe *et al.*, 1999). If LUCA had an RNA genome, why did it make DNA and how did it replicate? One proposal is that LUCA replicated in similar fashion to certain modern retroviruses that have RNA genomes, but use duplex DNA as a substrate for translation and replication. The RNA is reverse transcribed into DNA, then the RNA strand is digested by a nuclease followed

Table 1 Evolutionary relationship of replication fork proteins in bacteria and eukaryotes

Protein	Common ancestor	Fold pattern	Bacterial	Eukaryotic
Polymerase	No	C- vs B family	Pol C/Pol III	Pol α, δ, ϵ
Clamp	Yes	Same	β	PCNA
Clamp loader	Yes	Same	τ complex	RFC
Helicase	No	RecA vs AAA +	DnaB	Mcm2–7 (CMG)
Primase	No	Toprim vs Pol X	DnaG	Pri1
SSB	No	OB	SSB	RPA
Additional factors	No			Mcm10, Ctf4, GINS, and Cdc45

by conversion to duplex DNA (Figure 1). The DNA then serves as substrate for transcription by RNA polymerase to form transcripts and new copies of the RNA genome. It is interesting to note that this pathway does not require helicases, primase, clamps, and clamp loaders.

RNA is intrinsically less stable than DNA, and the enhanced stability likely led, over time, to the use of DNA as the genetic repository, and thus evolution of the replication machinery. These and other hypothesis have been proposed in the comparative genomics field to explain the non-homology of central replication enzymes in bacteria compared to eukaryotes/archaea (Forterre, 2013; Leipe *et al.*, 1999). Other possibilities exist, such as LUCA used DNA for its genome and had two replication systems, one being retained in bacteria and the other in eukaryotes/archaea. But this implies that replication in LUCA was much more complicated than in modern cells. One can consider whether the large size and nucleosome packaging of eukaryotic genomes selected for changes in replication strategy. But archaeal and eukaryotic replication proteins are homologous yet archaea, like bacteria, do not have nucleosomes or large genomes; they also generally have circular chromosomes without telomeres, and replicate rapidly like bacteria.

Regardless of how the different replication enzymes evolved, it is remarkable that the process is performed by proteins with the same biochemical function, even though the enzymes have distinct mechanisms and structures to perform these tasks (Table 1). Given the distinct evolutionary paths to replication, it is possible, even probable that the mechanics by which the proteins function together to accomplish replication will have significant differences. This review gives an overview of the structure and function of the central proteins at replication forks in bacteria and eukaryotes. We conclude with discussion of how these several enzymes coordinate their activities within replisome machines. Protein nomenclature used in this review utilizes that of *Escherichia coli* for bacteria, and of budding yeast (*Saccharomyces cerevisiae*) for eukaryotes. Where relevant, the names for human and archaeal proteins are given in parenthesis.

Bacterial Replicative DNA Polymerases Are Distinct from Those of Eukaryotes

Functional Overview of Bacterial (C Family) and Eukaryotic (B Family) Replicative DNA Polymerases

DNA polymerases do not start their own DNA chain; they can only extend from a preexisting primed template junction

(Kornberg and Baker, 1992). All DNA polymerases extend DNA unidirectionally, 3'-5', a natural consequence of dNTP substrates that are activated at the 5' position of the ribose and not the 3' position. There are a variety of different DNA polymerases with diverse sequences, and they have been as-sorted into classes of homologous sequences, referred to as families. The major DNA polymerase families are A, B, C, D, X, and Y (Steitz, 1999; Yang, 2014). Crystal structure analysis reveals that all DNA polymerases are shaped like a right hand, with palm, fingers, and thumb domains (see Figure 2(a)). The folding patterns of these domains are distinct among the different families, consistent with their nonhomologous sequences.

The replicative polymerase of all bacterial cells (i.e., Pol III/C) is in the C family. *Escherichia coli* Pol III 'core' contains three subunits, the α subunit polymerase, the ϵ subunit proof-reading 3'-5' exonuclease, and the θ subunit of unknown function. Most bacteria lack θ , and many bacterial C-family polymerases contain a 3'-5' exonuclease within the same polypeptide chain instead of a separate protein (Barros *et al.*, 2013; McHenry, 2011). Some bacteria also contain a second C-family polymerase, thought to be involved in repair or an ancillary role in replication (McHenry, 2011).

Eukaryotes require three different DNA polymerases for chromosome replication, each in the B family (Garg and Burgers, 2005a; Johansson and Macneill, 2010; Stillman, 2008). The leading strand is performed by Pol ϵ (Kunkel and Burgers, 2008), the lagging strand is performed by Pol δ (Pursell *et al.*, 2007) and also by Pol α , which makes RNA/DNA primers (Conaway and Lehman, 1982). Each of these polymerases consists of multiple subunits. The second largest subunit is referred to as the B-subunit; the B-subunits of all three replicative polymerases share homology (Johansson and Macneill, 2010). The functions of the B-subunits are largely unknown. Each replicative polymerase has additional accessory subunits as well. The polymerase subunit of each replicative enzyme has recently been found to contain a zinc finger and an iron-sulfur (FeS) cluster (Netz *et al.*, 2011). The zinc finger is necessary to bind the B-subunit and the function of the FeS cluster is still under investigation (Sanchez Garcia *et al.*, 2004). Archaeal cells also utilize two distinct polymerases for leading (Pol B) and lagging (Pol D) synthesis and have a primase related to the primase of eukaryotic Pol α (discussed later) (Beattie and Bell, 2011; Ishino *et al.*, 2013).

The leading strand Pol ϵ consists of Pol2 (polymerase/exonuclease, p261 in human), the B-subunit/Dpb2 (p59 in human), and the small accessory subunits Dpb3 (p17 in

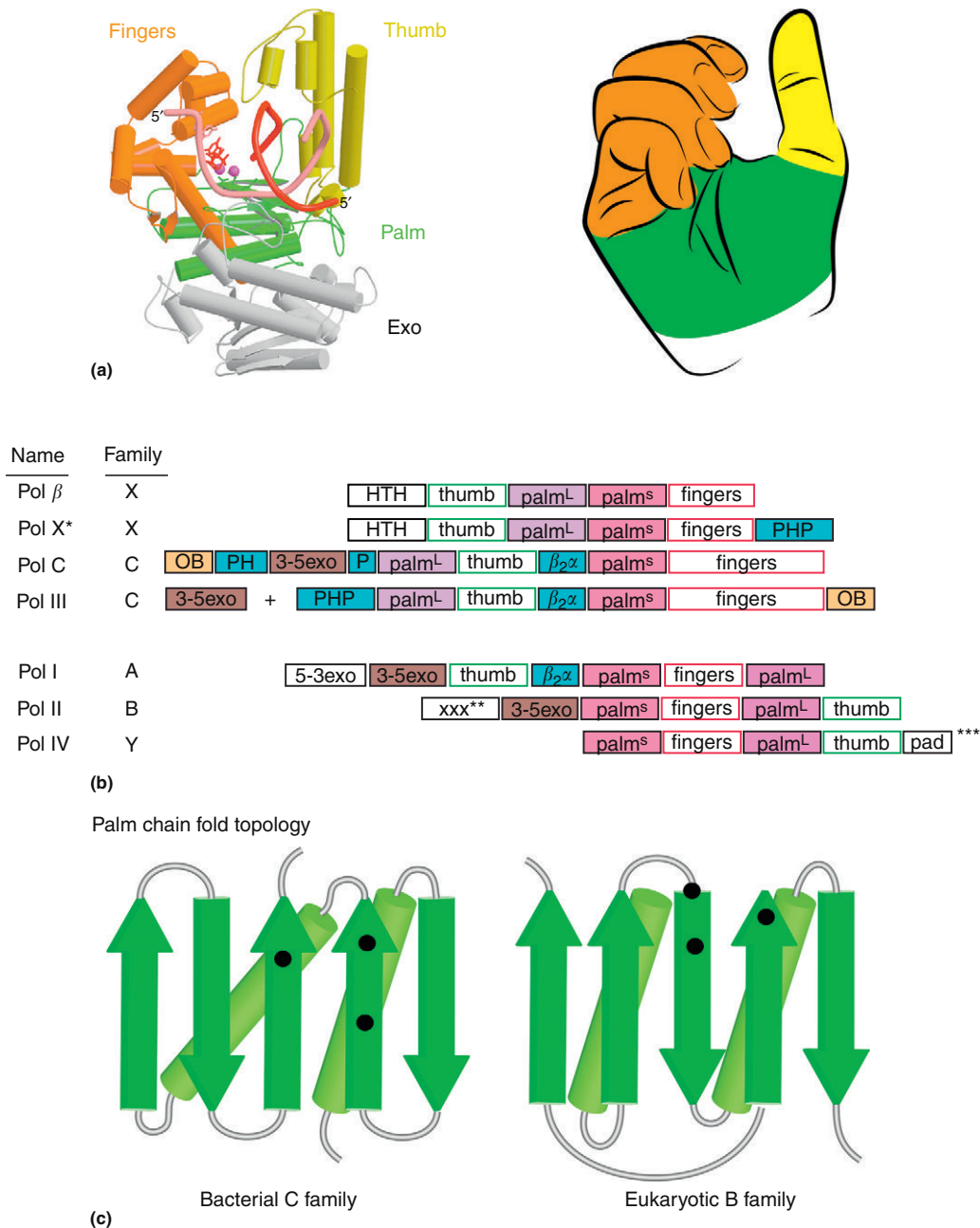


Figure 2 Distinct structure and organization of bacterial and eukaryotic DNA polymerases. (a) DNA polymerases have a right hand shape. Left: Cartoon illustration of T7 DNA polymerase with fingers (orange), palm (green), thumb (yellow) domains; the 3'-5' exonuclease domain is in gray. Reprinted from Beard, W.A., Wilson, S.H., 2003. Structural insights into the origins of DNA polymerase fidelity. Structure 11, 489–496, Copyright (2003) with permission from Elsevier. Right: Schematic of a right hand with anatomical features that are colored similar to the polymerase domains that carry the same name. (b) Organization of domains of polymerases in different families. Domains with similar structure and function are in full colors (e.g., palm of C and X families, palm of A, B, and Y families). Domains with similar function but different structure are outline colored (i.e., thumb and fingers). Domains of unrelated or unclear function are black and white. (c) Chain folding topology diagrams of palm domains in the eukaryotic B family and the bacterial C family. The conserved acidic residues that coordinate metal ions for catalysis are shown as three black dots. OB, OB fold; HTH, helix-turn-helix; pad, polymerase associated domain, also called 'wrist' or 'little finger'; PHP, polymerase and histidinol phosphatase. Reprinted from Lamers, M.H., Georgescu, R.E., Lee, S.-G., O'Donnell, M., Kuriyan, J., 2006. Crystal structure of the catalytic α subunit of *E. coli* replicative DNA polymerase III. Cell 126(5), 881–892, Copyright (2006) with permission from Elsevier.

human) and Dpb4 (p12 in human) (Pursell and Kunkel, 2008). Dpb3 and Dpb4 are dispensable for viability in budding yeast, but Dpb2 is essential. Pol δ contains the Pol3 subunit (polymerase/exonuclease; p125 in human), the

B-subunit Pol31 (p50 in human), and accessory subunit Pol32 (p66 in human) (Garg and Burgers, 2005a; Stillman, 2008). Pol32 is not essential and Pol δ from higher eukaryotes contains a nonessential fourth subunit (p12 in human). Pol α is a

four-subunit complex consisting of Pol I (polymerase, p180 in human), the B-subunit Pol II (unknown function, p70 in human), Pri I (primase, p48 in human, Pri S in archaea), and Pri II (p58 in human, Pri L in archaea), which modulates primase activity (Garg and Burgers, 2005a; Johansson and Macneill, 2010; Stillman, 2008). All four subunits of Pol α are essential to viability in all cells examined thus far.

Crystal Structure of B- and C-Family DNA Polymerases

The first structure of a DNA polymerase was *E. coli* Pol I, a member of the A family that is mostly involved in repair (Pursell and Kunkel, 2008). Since that time, representative polymerases in all families have been crystallized, and all have been found to contain the same 'right hand' shape, with three major domains: a palm, fingers, and thumb (Figure 2(a); Johansson and Macneill, 2010; Steitz, 1999). The different polymerase families result in many variations in the structural architecture of these domains, but their folding into the right hand shape is universal. The three acidic residues that bind the two metal ions for catalysis are always found in the palm domain, and the dNTP binding site is located in the fingers domain (Beese and Steitz, 1991; Steitz, 2006). The thumb domain facilitates binding duplex DNA. Many DNA polymerases incorporate multiple dNTPs in one DNA binding event by sliding along DNA after successive dNTP additions. Binding of nucleotide induces a conformational change in the fingers domain that pairs the dNTP to the template strand, and checks for base pair accuracy by a conformational-induced fit mechanism in which only base pairs of the correct geometry trigger the catalytic step (Doublié *et al.*, 1999; Johnson *et al.*, 2003; Joyce *et al.*, 2008; Li *et al.*, 1998; Luo *et al.*, 2007; Santoso *et al.*, 2010). There are often one or more additional domains in DNA polymerases. An example of this is illustrated in Figure 2(a), for the 3'-5' proofreading exonuclease domain in T7 DNA polymerase. The alignment of the polymerase sequences, and the locations of the domains from their respective structures are shown in Figure 2(b).

The C-family polymerase of both *E. coli* and *Thermus aquaticus* contain the palm, thumb, fingers domains, and also a polymerase and histidinol phosphatase domain (PHP domain), a C-terminal domain, and a domain that binds a sliding clamp (see Figure 2(b); Bailey *et al.*, 2006; Lamers *et al.*, 2006; Lamers and O'Donnell, 2008; Wing *et al.*, 2008). Some PHP domains of C-family polymerases are metal-dependent nucleases and may play a role in proofreading (Barros *et al.*, 2013; Stano *et al.*, 2006). But in many C-family polymerases the PHP domain has lost ability to bind metal and appears inactive (Barros *et al.*, 2013; Lamers *et al.*, 2006). In these cases the replicative C-family polymerase either contains a separate 3'-5' exonuclease domain, or recruits a separate 3'-5' exonuclease subunit for proofreading, as with *E. coli* Pol III which recruits ϵ for proofreading (compare Pol C and Pol III in Figure 2(b)).

The first crystal structure of a B-family polymerase was RB69 bacteriophage gp43, homologous to eukaryotic replicative polymerases α , δ , and ϵ (Stano *et al.*, 2006). Like eukaryotic Pols δ and ϵ , RB69 Pol contains a proofreading 3'-5' exonuclease. The structures of the fingers and thumb domains

of different families are distinct (Figure 2(b)), but the palm domains that contain the 'active site' acidic residues in A-, B-, and Y-family polymerases are similar (the Y family are error-prone polymerases discussed later). The C-family palm fold, and location of the active site acidic residues, is distinct from the A, B, C families, but has similarity to the palm of X-family polymerases (Bailey *et al.*, 2006; Lamers *et al.*, 2006; Lamers and O'Donnell, 2008; Wing *et al.*, 2008). The chain folding topology of the palms of B and C families are shown illustrated in Figure 2(c). The significant sequence and structural differences between bacterial and eukaryotic polymerases underscore their distinctive evolutionary lineage.

Sliding Clamps Are Conserved in All Cell Types

Unlike the DNA polymerases (and primase, helicase, and SSB), bacterial and eukaryotic sliding clamps evolved from a common ancestor (Leipe *et al.*, 1999). Sliding clamps are ring-shaped proteins that encircle duplex DNA and freely slide along it (Gulbis *et al.*, 1996; Kong *et al.*, 1992; Krishna *et al.*, 1994; Moarefi *et al.*, 2000). One function of sliding clamps is to bind directly to replicative polymerases and tether them to DNA, endowing them with high processivity (Stukenberg *et al.*, 1991). Originally discovered in *E. coli*, bacterial clamps are a ring-shaped homodimer called β (Figure 3(a); Kong *et al.*, 1992; Stukenberg *et al.*, 1991). The β dimer has a 6-fold appearance that derives from three globular domains in each subunit; the three domains share the same chain fold. The outside of the ring is a continuous layer of antiparallel sheet that also forms the dimer interface. The inside of the ring is lined by helices that form a cavity with a diameter of about 35 angstroms; the width of β is equal to the length of one turn of B-form duplex DNA. The eukaryotic and archaeal clamps, proliferating cell nuclear antigen (PCNA), are also ring shaped and have the same chain folding pattern as bacterial β (Gulbis *et al.*, 1996; Krishna *et al.*, 1994). PCNA is a trimer and each monomer consists of two globular domains, giving the ring the same 6-fold appearance as *E. coli* β . The outer surfaces of both clamps carry a net negative charge, while the insides of the rings are positively charged.

The similar structures of β and PCNA reveal that they evolved from a common ancestor, but their amino acid sequences have diverged significantly. Rapid sequence divergence is common for proteins that serve structural functions, as they are not constrained by the precise geometries of active site residues needed at a catalytic center. Indeed, advanced sequence comparison algorithms are required to identify the common origin of the domains within clamp subunits, and the homology between β and PCNA (Leipe *et al.*, 1999).

The bacterial and eukaryotic clamps interact with numerous proteins, not just their respective replicative polymerases (Maga and Hubscher, 2003). The interaction of clamps with partner proteins is mediated by a peptide motif that binds a hydrophobic pocket on the surface of the clamp (Dalrymple *et al.*, 2001; Gulbis *et al.*, 1996). PCNA and β interact with proteins of repair, cell cycle regulatory proteins, DNA ligase, and specialized DNA polymerases involved in traversing DNA

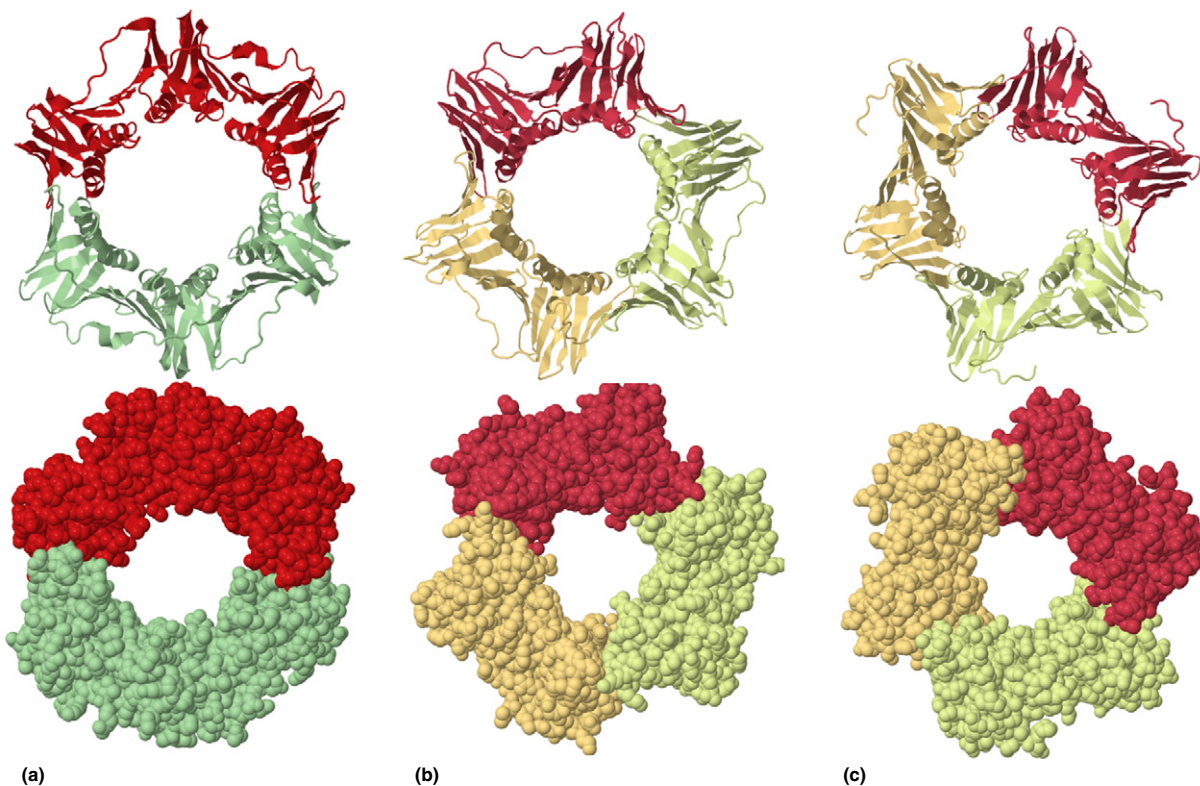


Figure 3 Sliding clamps are conserved in all three domains of life. Ribbon (top) and space filling (bottom) representations of: (a) *E. coli* b (2POL), (b) yeast PCNA (1PLQ), and (c) *Pyrococcus furiosus* (1GE8).

lesions (De Biasio and Blanco, 2013; Maga and Hubscher, 2003).

The Clamp Loader Was Also Present in LUCA

Clamps do not get onto DNA by themselves, they require a multiprotein clamp loader that uses adenosine triphosphate (ATP) to open and close the clamps around DNA (Ellison and Stillman, 1998; Fien and Stillman, 1992; Jeruzalmi *et al.*, 2001; Stukenberg *et al.*, 1991). Clamp loaders, both bacterial and eukaryotic are composed of five subunits essential to clamp loading activity (Bowman *et al.*, 2004; Cullmann *et al.*, 1995; Jeruzalmi *et al.*, 2001). Clamp loader subunits also share sequence homology to one another and are members of one gene family (Cullmann *et al.*, 1995). Moreover, the sequences of clamp loading subunits of bacteria are clearly homologous to the clamp loading subunits of eukaryotes (Bunz *et al.*, 1993; Cullmann *et al.*, 1995; O'Donnell *et al.*, 1993), indicating that LUCA contained the clamp loader.

Overview of Clamp Loader Mechanism

The mechanism of the clamp loader has emerged from several key clamp loading crystal structures and biochemical studies (Ason *et al.*, 2003; Bowman *et al.*, 2004; Jeruzalmi *et al.*, 2001; Kelch *et al.*, 2011; O'Donnell and Kuriyan, 2006; Turner *et al.*, 1999). In overview, the clamp loader associates with the clamp

upon binding to ATP and opens the clamp at one interface (Dionne *et al.*, 2008; Kelch *et al.*, 2011; Paschall *et al.*, 2011; Zhuang *et al.*, 2006). The clamp loader has a notch in one side through which DNA can enter a central chamber, positioning the DNA through the clamp as illustrated in Figure 4(a) (Kelch *et al.*, 2011). The central DNA binding chamber is capped by a 'collar' composed of the C-terminal domains that have no opening for DNA to go through. Hence, only DNA with a flexible hinge can fit into the central chamber, because the DNA must make a sharp bend to exit the notch in the side of the chamber. This requisite flexibility is provided by the ssDNA portion of a primed template (Bowman *et al.*, 2005; Kelch *et al.*, 2011; Simonetta *et al.*, 2009). Once on primed DNA, the ATP sites are brought into register to hydrolyze ATP, thus ejecting the clamp loader and allowing the clamp to close around DNA.

The function of the clamp loader was initially identified in the *E. coli* system, referred to as γ complex (Stukenberg *et al.*, 1991). Each subunit contains a sequence motif of the AAA+ family (ATPases Associated with a variety of cellular Activities) (Cullmann *et al.*, 1995; Jeruzalmi *et al.*, 2002; Kelch *et al.*, 2012). The AAA+ family is huge and as its name implies, its members function in a wide range of biological actions that include vesicle fusion, protease, helicase, and origin initiation (Erzberger and Berger, 2006; Neuwald *et al.*, 1999). Like the clamp loader subunits, most AAA+ proteins are multiprotein complexes. The first AAA+ protein to be solved structurally was the δ subunit of the *E. coli* clamp loader (Guenther *et al.*, 1997). This C-shaped subunit revealed features of AAA+ proteins that

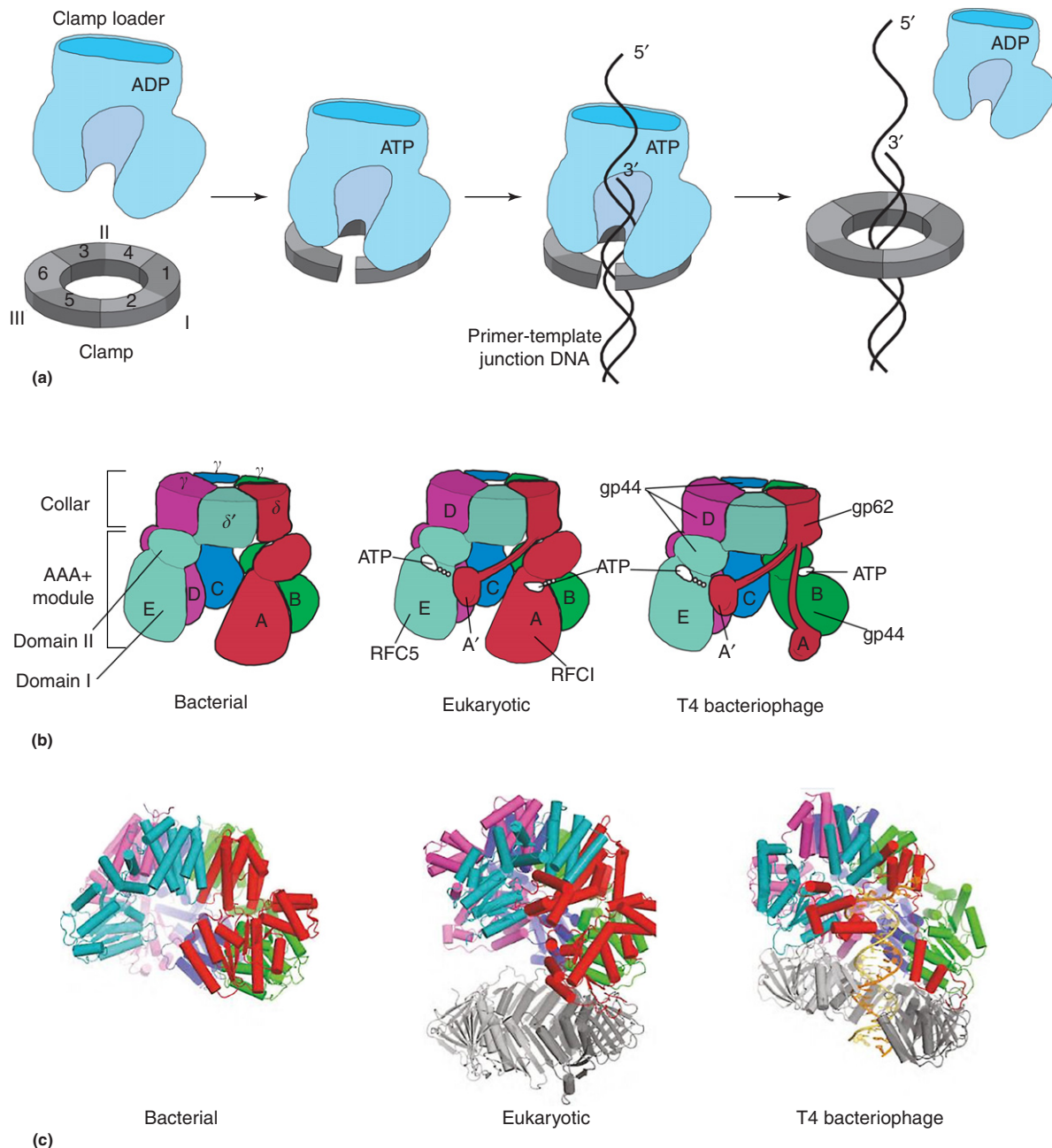


Figure 4 Clamp loaders spanning bacteria to eukaryotes have similar structure and function. (a) Overview of clamp loader mechanism. First diagram: The clamp loader has a central chamber accessible to DNA by a notch in the side. The clamp loader cannot bind the clamp when in the ADP bound form. Second diagram: ATP enables the clamp loader to bind and open the clamp. Third diagram: Primed DNA binds the clamp-clamp loader complex by passing through the opened ring and the notch in the side of the clamp loader. Fourth diagram: ATP hydrolysis is triggered and the clamp loader ejects, leaving the clamp to close around the primed site. Fifth diagram: DNA polymerase binds the clamp for processive DNA synthesis. (b) Illustration of the domain architecture of clamp loaders. Clamp loaders are circular pentamers, and each subunit has at least three domains. The N-terminal domains I and II are the AAA + ATPase region and domain III forms a tight circular collar. The five subunits are denoted, counterclockwise, A–E. A gap exists between AAA + domains of subunits A and E. The common names of each subunit are indicated for each clamp loader. The A subunit of the eukaryotic and T4 clamp loaders contain an extra C-terminal region that reaches across the notch to the E subunit (this domain is labeled A'). From Kelch, B.A., Makino, D. L., O'Donnell, M., Kuriyan, J., 2011. How a DNA polymerase clamp loader opens a sliding clamp. *Science* 334, 1675–1680. Reprinted with permission from AAAS. (c) Ribbon representations of clamp loaders. Each subunit is a different color. The clamp, when present, is gray. The DNA in the T4 bacteriophage clamp-clamp loader structure is yellow. Reproduced from Kelch, B.A., Makino, D.L., O'Donnell, M., Kuriyan, J., 2012. Clamp loader ATPases and the evolution of DNA replication machinery. *BMC Biology* 10 (34), with permission from BioMed Central.

have been proven to generalize to all crystal structures of AAA+ proteins. The AAA+ region folds into two domains, one domain contains the walker A and B sites that bind the nucleotide and the other domain forms a helical 'lid' that contains residues that interact with the nucleotide and may be used to regulate hydrolysis. Two key residues of the lid domain are basic residues referred to as sensor 1 and sensor 2 that sense ATP binding and modulate intersubunit communication (Guenther *et al.*, 1997). Different AAA+ proteins that serve various cellular functions contain additional domains at the N- and/or C-terminus of the AAA+ region. In the case of clamp loaders, all the subunits contain a C-terminal domain that forms a tightly associated 'collar' from which the AAA+ domains are suspended (Bowman *et al.*, 2004; Jeruzalmi *et al.*, 2001; Kelch *et al.*, 2011; O'Donnell and Kuriyan, 2006).

Structure Analysis of Clamp Loaders

Crystal structure analysis of the *E. coli* clamp loader reveals a circular arrangement of the five subunits, one δ , three γ , and one δ' , referred to as subunits A–E as illustrated in the leftmost diagrams of Figures 4(b) and 4(c). The notch in the side of the clamp loader is formed by a gap between the AAA+ domains of δ and δ' (Figure 4(b); Jeruzalmi *et al.*, 2001). The AAA+ domains form a shallow spiral. Only the γ subunits contain ATP sites, as δ and δ' lack a functional Walker A motif (i.e., P loop). The three ATP sites are located at the γ – δ' and the two γ – γ interfaces (Jeruzalmi *et al.*, 2001). The δ' subunit contributes a catalytic arginine finger residue to ATP bound in γ , and this bi-partite ATP site construction holds true for the other two ATP sites as well. The location of ATP sites at subunit interfaces is a general feature of AAA+ oligomers and enables communication and coordination among subunits during ATP hydrolysis. The structure of the eukaryotic clamp loader, replication factor C (RFC) in complex with the PCNA clamp (middle diagrams in Figures 4(b) and 4(c)) shows a similar circular arrangement of the five AAA+ subunits (each are distinct polypeptides), with the PCNA clamp located under surface generated by the AAA+ domains (Bowman *et al.*, 2004). RFC contains four ATP sites, but only three, at the positions analogous to those of γ complex, are required for clamp loading. One subunit of RFC, RFC1 (a subunit), contains additional N- and C-terminal extensions that are not present in bacterial clamp loaders. The additional RFC1 N-terminal region is not essential to clamp loading, or to cell viability (Gomes *et al.*, 2000; Majka and Burgers, 2004). But the additional C-terminal region of RFC1 (not present in the structure of Figures 4(b) and 4(c)) reaches across the notch between the A and E subunits and EM studies reveal that it binds to PCNA (Miyata *et al.*, 2005).

The bacterial γ complex and eukaryotic RFC contain a central chamber with dimensions consistent with binding duplex DNA, confirmed in a cocrystal of *E. coli* γ complex with a primed DNA substrate (Simonetta *et al.*, 2009). Interestingly, DNA acts as a scaffold to bring the subunits into a spiral with a pitch similar to A-form duplex DNA. The main protein connections to the duplex are formed by a loop in the AAA+ domains of each subunit that interacts only with the phosphodiester backbone of the template strand. Template ssDNA bends out the notch

between δ and δ' (Simonetta *et al.*, 2009). Significant advances in our knowledge of the clamp loader mechanism derive from structural studies of the T4 bacteriophage clamp loader, composed of four gp44 subunits and one gp61 protein, in complex with the gp45 clamp, primed DNA, and an ATP analog (Kelch *et al.*, 2011). Archaeal clamp loaders are also often constructed from four identical subunits and one additional subunit (Dionne *et al.*, 2008; Seybert *et al.*, 2002). These studies confirmed that the clamp loader binds the entire surface of the clamp in the active complex, and open the clamp by bending it into a right hand open spiral with DNA positioned through the open clamp (Figure 4(c)). Structures of different nucleotide-bound forms of the T4 clamp-clamp loader-primed DNA complex indicate that ATP hydrolysis proceeds stepwise, from one ATP site to the next, counterclockwise around the ring (i.e., from ATP bound to subunit B to C to D).

Why were clamps and the clamp loaders in LUCA, while other replicative machinery was not? If LUCA used an RNA genome and reverse transcriptase, as suggested in Figure 1, a simple presumption is that clamps were used to help reverse transcriptase stay on the nucleic acid, much as clamps are used by DNA polymerases in modern day cells.

Bacterial and Eukaryotic Replicative Helicases

Proposed Strand Exclusion Mechanism of Bacterial and Eukaryotic Replicative Helicases

DNA replication requires a helicase that harnesses the energy of ATP hydrolysis to separate the strands of DNA. Replicative helicases are circular hexamers that encircle one single strand and use ATP to motor along it, excluding the other strand (Ahnert and Patel, 1997; Hacker and Johnson, 1997; Lee *et al.*, 2014). The excluded single strand is thus prevented from annealing to the strand inside the helicase ring. Helicases generally operate with other proteins that keep the strands separated during unwinding, otherwise the two strands could simply reanneal in back of the helicase (von Hippel and Delagoutte, 2001). Single-strand DNA binding protein, discussed later, is one such protein. DNA polymerase also prevents strand reannealing by converting one unwound strand into a duplex. Helicases may also 'back slide,' allowing strands to reanneal, and a polymerase and/or single-strand DNA-binding protein may prevent backsliding and increase helicase efficiency (Donmez and Patel, 2008).

Numerous helicases exist for various functions, and based on sequence analysis are assorted into six superfamilies (SF1–SF6) (Singleton *et al.*, 2007). All but the SF1 and SF2 families are hexamers (Donmez and Patel, 2006; Enemark and Joshua-Tor, 2008). The ATP motor domains of the bacterial replicative helicase, exemplified by *E. coli* DnaB, are constructed from the RecA fold and are members of the SF4 superfamily. The eukaryotic/archaeal Mcm helicase is in the SF6 family based on the AAA+ fold. Bacterial DnaB and eukaryotic Mcm helicases travel in opposite directions (Bochman *et al.*, 2008; Bochman and Schwacha, 2008; LeBowitz and McMacken, 1986). Thus at a replication fork the prokaryotic helicase encircles the lagging strand while the eukaryotic/archaeal helicase surrounds the leading strand. Given the nonhomologous sequences, distinct

RecA versus AAA+ folds, and opposite directions of unwinding, the bacterial and eukaryotic/archaeal helicases are thought to have evolved independently (Leipe *et al.*, 1999).

Helicases use the energy of ATP hydrolysis to translocate along ssDNA, but it is not yet certain that the energy of ATP hydrolysis is coupled to DNA unwinding. Thermal breathing of DNA may occur with sufficient speed for unwinding (Johnson *et al.*, 2007). Specifically, ATP-driven helicase translocation along ssDNA and consequent occupation of thermally unwound DNA may be sufficient to explain helicase unwinding (Jeong *et al.*, 2004; Johnson *et al.*, 2007; Jose *et al.*, 2012). Helicases use one ATP every 1–2 bp unwound, and given the –12 kcal per ATP hydrolyzed and –3.6 kcal for 2 bp of duplex DNA, the energetics is consistent with either scenario (Delagoutte and von Hippel, 2001; Donmez and Patel, 2008; von Hippel and Delagoutte, 2001).

Bacterial Helicase Structure

E. coli DnaB helicase and the helicases of bacteriophage T4 and T7 are based on the RecA fold and the hexameric rings have the appearance of two tiers of rings, because the protein has N- and C-terminal globular regions (Bailey *et al.*, 2007; Wang *et al.*, 2008). The RecA-based domains form the C-terminal tier and act as the motor, while the N-terminal domains form a ring-shaped tier that acts as a collar. Both ring-shaped tiers have holes in their center that can accommodate DNA. Crystal structures of DnaB-ssDNA and the T7 gp4 helicase have been solved (Itsathitphaisarn *et al.*, 2012; Sawaya *et al.*, 1999; Singleton *et al.*, 2000; Toth *et al.*, 2003). The RecA motor domains form an asymmetric out-of-plane spiral arrangement of the domains due to different nucleotide-bound states within distinct subunits. The DnaB-ssDNA structure bound to ssDNA suggests the following mechanism of action, illustrated in Figure 5. Firstly, the binding of ssDNA to the DnaB hexamer molds the RecA tier of DnaB into a right hand spiral having the pitch of A-form DNA. Each subunit binds the phosphodiester bonds of two nucleotides, for a total of 12 nucleotides bound per hexamer. ATP is hydrolyzed in an ordered fashion around the hexamer, one ATP at a time. The subunit at the bottom of the 'spiral staircase' hydrolyzes ATP and the DNA interaction is broken. This subunit then traverses 12 nucleotides to the top of the staircase and rebinds ssDNA and ATP. As this process repeats itself, the staircase moves hand-over-hand along the ssDNA, two nucleotides for each ATP hydrolyzed.

Eukaryotic Helicase Structure

The eukaryotic Mcm2–7 helicase is composed of six unique subunits that form a ring, but Mcm2–7 is nearly inactive and it associates with five other proteins, Cdc45 and the four subunit GINS complex (Ilves *et al.*, 2010; Moyer *et al.*, 2006). Little is known about the function of Cdc45 and GINS (none contain ATP sites). The four GINS subunits are related by sequence and structure, but are distinct and are known to bind other replication proteins (Pol ϵ and Ctf4), hinting that the GINS may act as a scaffold. Thus the eukaryotic helicase is referred to as CMG (Cdc45, Mcm2–7, and GINS). CryoEM studies of CMG reveal that the Cdc45 and GINS subunits form an additional 'hole'

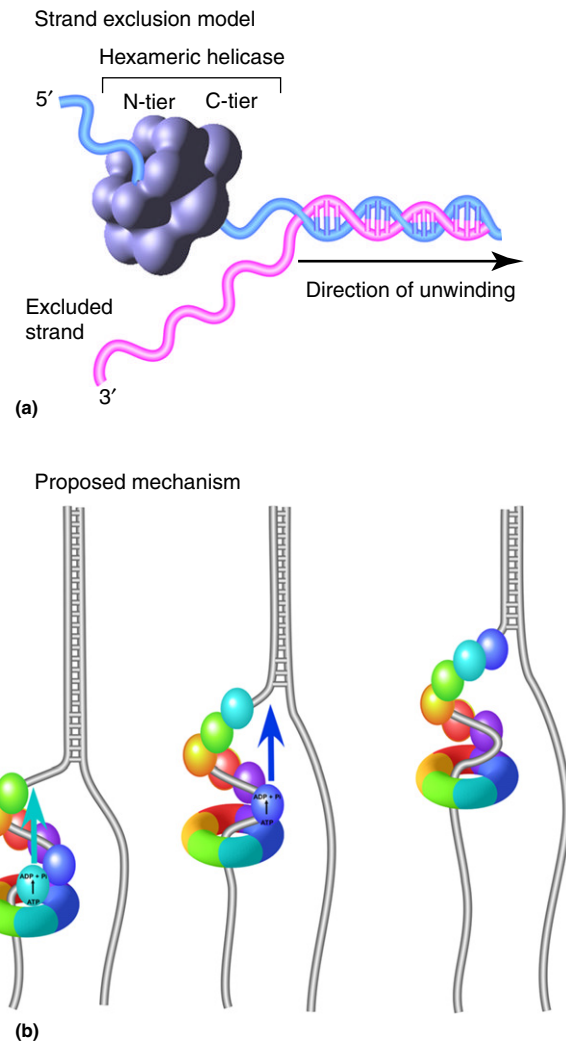


Figure 5 Hexameric helicase mechanism. (a) Hexameric helicases encircle one strand of DNA and exclude the other. ATP hydrolysis is used to propel the helicase along the strand it encircles, driving it into the forked junction. Although bacterial and eukaryotic helicases are constructed from nonhomologous chain folds and sequence they are both circular hexamers and operate by this strand exclusion mechanism. Both types of helicases have a double-ring appearance, the C-terminal domains form a 'tier' that faces the duplex and contains the ATP sites; the unwound DNA product extrudes from the domains that comprise the N-terminal 'tier.' Bacterial and eukaryotic helicases travel opposite directions, and thus encircle either the lagging or leading strands, respectively. (b) Schematic of proposed mechanism of hexameric helicases. The C-terminal domains form an A-form spiral 'staircase' upon binding ssDNA (left). Hydrolysis of ATP in the subunit at the bottom leads to dissociation of that domain from DNA, followed by relocation to the top of the staircase and rebinding to ATP and ssDNA (middle), moving the helicase either one (eukaryotic) or two (bacterial) nucleotides. The right diagram indicates a second round of ATP hydrolysis.

on one side of the Mcm2–7 ring, suggesting CMG may encircle both single strands at a forked junction (Costa *et al.*, 2011). Further studies will be needed to determine if this is the case. The Mcm2–7 hexamer contains the motor subunits, but unlike DnaB, the Mcm2–7 ring contains a gap, between subunits

Mcm2 and Mcm5 (Costa *et al.*, 2011; Lyubimov *et al.*, 2012). The archaeal Mcm is a homohexamer (and sometimes double hexamer) that does not appear to contain a gap between subunits; and it is an active helicase without accessory factors (Beattie and Bell, 2011; Bell and Botchan, 2013; Chia *et al.*, 2010; Chong *et al.*, 2000; Graham *et al.*, 2011; Pan *et al.*, 2011; Sakakibara *et al.*, 2009). However, archaeal cells contain sequence homologs of the GINS and Cdc45, and thus may form a CMG complex like eukaryotes (Makarova *et al.*, 2012).

Most of the knowledge about how AAA+ hexameric helicases function is derived from studies of SV40 T-antigen and papilloma virus E1 helicases. These helicases are based on the AAA+ domain and form homohexameric rings, suggesting they share a common ancestor with the Mcm proteins of eukaryotes and archaea (Chia *et al.*, 2010; Leipe *et al.*, 1999; Makarova and Koonin, 2013). A major breakthrough on the mechanism of these helicases came with the crystal structure of the papilloma virus E1 helicase bound to ssDNA in the presence of ADP (Enemark and Joshua-Tor, 2006). This study revealed a right hand spiral arrangement of ssDNA within the E1 homohexamer. Nucleotide states of subunits within the ring included ATP-like bound subunits, ADP-like bound subunits, and subunits having no bound nucleotide. As with DnaB, the E1 hexamer contains two tiers of rings, an N-terminal tier serving as a collar with a hole for ssDNA to exit, and a C-terminal tier containing the AAA+ domains. The AAA+ domains contain a long loop, each of which binds only one nucleotide per monomer, for a total of six nucleotides per helicase. E1 translocates along ssDNA in the opposite direction to DnaB, but the helical structure implied an 'escort' mechanism of action not unlike the proposed mechanism of DnaB. With each ATP hydrolytic event, the DNA binding loop within the hydrolyzing subunit detaches from ssDNA, then upon rebinding ATP, the subunit reattaches to ssDNA at the top of the helix, after sliding along the six nucleotide long binding site and moving the helicase one base pair in the process. This proposed mechanism is referred to as the 'escort model,' because the ssDNA is escorted through the hexamer by long loops one base at a time.

Primases Are Very Different in Bacteria Compared to Eukaryotes

To initiate DNA chains, cells from all domains of life contain a primase enzyme that synthesizes a short RNA primer on a template ssDNA (Kornberg and Baker, 1992). The difficult step in this process is the formation of the first dinucleotide. This requires the enzyme to have two NTP binding sites, positioning them adjacent to one another to form the first phosphodiester bond. In the cell, rNTPs are at 10- to 100-fold higher concentration than dNTPs and thus use of RNA to initiate synthesis of DNA may reflect the need for a high NTP concentration to bind two NTPs at the same time.

The Bacterial Primase

The sequence and structure of bacterial and eukaryotic primases are completely different (Frick and Richardson, 2001). The structure of bacterial primase, referred to as DnaG, reveals

that it is related to topoisomerases (Figure 6(a); Keck *et al.*, 2000; Podobnik *et al.*, 2000). This folding pattern is referred to as the 'toprim fold,' to reflect this relationship (Aravind *et al.*, 1998). DnaG primase starts chains with a purine and generates RNA primers of 10–12 nucleotides (Rowen and Kornberg, 1978). After primer synthesis, the β clamp is assembled on the primer by the clamp loader and extended into a 1–2 kb Okazaki fragment by the replicative C-family polymerase. RNA primers are removed and filled in by the concerted action of the 5'-3' exonuclease and polymerase activities of Pol I, followed by ligase which seals the fragments together (Kornberg and Baker, 1992).

The Eukaryotic Primase

The eukaryotic primase requires two different subunits, Pri1 and Pri2, which are associated within the four subunit Pol α (Kaguni *et al.*, 1983a). Archaeal primase is composed of two subunits without the additional polymerase and B-subunit of Pol α (Liu *et al.*, 2001; Hubscher *et al.*, 2002). The small subunit (Pri1, p48 kDa in human, PriS in archaea) is the catalytic subunit (Copeland and Wang, 1993; Hubscher *et al.*, 2002; Kaguni *et al.*, 1983a; Kuchta and Stengel, 2010; Liu *et al.*, 2001). The large subunits of archaeal (PriL) and eukaryotic (PriL, p58 in human) primase share limited homology (Kuchta and Stengel, 2010). The large subunit is required, or greatly stimulates priming and it contains an FeS cluster needed for primase function (Klinge *et al.*, 2007; Weiner *et al.*, 2007; Zerbe and Kuchta, 2002). In eukaryotic Pol α , the RNA primer is transferred to the DNA polymerase subunit and extended another 20–25 nucleotides with DNA (Conaway and Lehman, 1982; Kilkenny *et al.*, 2013; Perera *et al.*, 2013). The reason that eukaryotes synthesize an RNA-DNA primer is not fully understood. The RNA-DNA primer is then recruited by PCNA/RFC and extended by Pol δ in a 'Pol α /Pol δ switch' (Tsurimoto and Stillman, 1991). Since DNA polymerase α lacks a proofreading exonuclease, the DNA portion of the primer is efficiently proofread by mismatch repair and Exo I (Hombauer *et al.*, 2011; Liberti *et al.*, 2013). The RNA portion, and some excision of the DNA portion involves limited strand displacement by Pol δ , and the Fen1 5'-3' nuclease or a combination of the Dna2 and Fen1 nucleases depending on the length of strand displacement (Balakrishnan and Bambara, 2013; Stith *et al.*, 2008).

The eukaryotic Pri1 catalytic subunit has sequence similarity to the Pol X family of DNA polymerases (Kirk and Kuchta, 1999), and comparison of their structures show that the 'active site' acidic residues superimpose (Figure 6(b)), indicating that eukaryotic primase probably functions by a two-metal mechanism like DNA polymerases (Augustin *et al.*, 2001). In a departure from eukaryotes, archaeal primases can initiate primers using dNTPs and can even extend them with DNA for one to several kilobases (Bocquier *et al.*, 2001; Liu *et al.*, 2001). However, archaeal primases bind rNTPs much tighter than they bind dNTPs (Lao-Sirieix and Bell, 2004). Importantly, in archaeal cells, Okazaki fragments contain 5' terminal RNA (Matsunaga *et al.*, 2003). Thus in the cell, archaeal primase makes RNA primers. Interestingly, some bacteria contain a LigB protein involved in repair that is

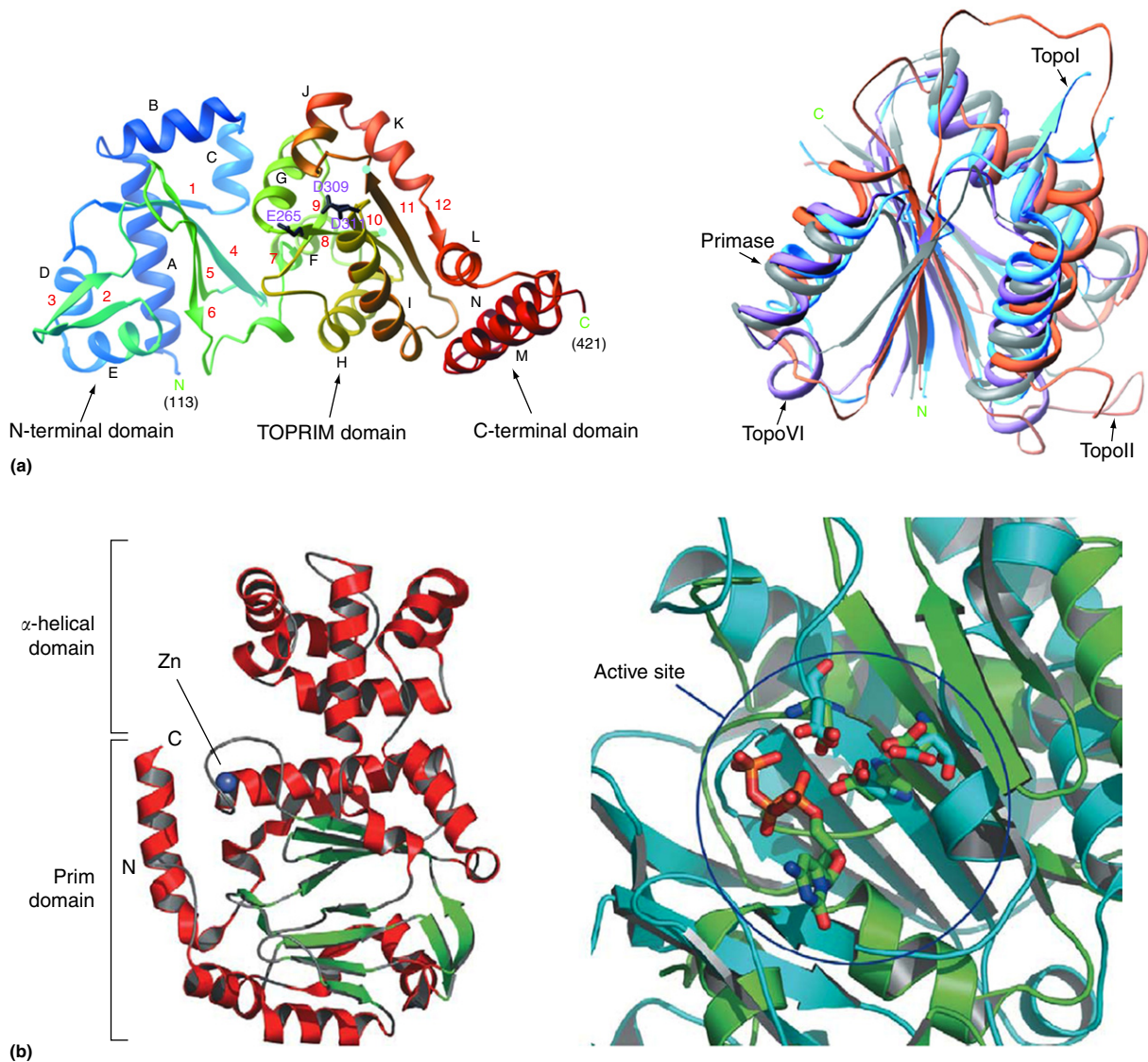


Figure 6 Structures of bacterial and eukaryotic/archaeal primases are unrelated. (a) Structure of *E. coli* DnaG primase and relationship to topoisomerases. Left: Cartoon representation of bacterial primase. The central catalytic domain has a 'toprim fold.' Right: Superimposition of DnaG primase toprim fold (green) with the homologous regions of three different topoisomerases (blue, purple, orange). Reprinted from Podobnik, M., McInerney, P., O'Donnell, M., Kuriyan, J., 2000. A TOPRIM domain in the crystal structure of the catalytic core of *Escherichia coli* primase confirms a structural link to DNA topoisomerases. *Journal of Molecular Biology* 300 (2), 353–362, Copyright (2000) with permission from Elsevier. (b) Structure of archaeal primase and relationship to Pol X. Left: Cartoon representation of the catalytic primase subunit of *Pyrococcus furiosus*. Right: Superposition of the catalytic sites of *Pyrococcus horikoshii* primase (Asp residues 95, 97, and 280) and human DNA polymerase β , an X-family DNA polymerase. Reprinted from Lao-Sirieix, S., Pellegrini, L., Bell, S.D., 2005. The promiscuous primase. *TRENDS in Genetics* 21 (10), 568–572, Copyright (2005) with permission from Elsevier.

homologous to archaeal primase and has similar catalytic activities (Zhu and Shuman, 2005).

The PrimPol Primase

Recent studies have identified a new type of eukaryotic primase-polymerase (primpol) in eukaryotic cells that shares homology to archaeal/eukaryotic primases, yet is a distinct protein (Garcia-Gomez et al., 2013). PrimPol can initiate synthesis using dNTPs or rNTPs, with an apparent preference

for dNTPs, and can extend DNA over common types of oxidative damage, abasic sites, and even UV photoproducts (Bianchi et al., 2013; Garcia-Gomez et al., 2013; Keen et al., 2014; Rudd et al., 2013; Wan et al., 2013). These properties suggest primPol may help nuclear forks to advance over a damaged nucleotide, and in fact some evidence for this has been demonstrated in cellular studies (Bianchi et al., 2013; Mouron et al., 2013). PrimPol is also located inside mitochondria, and gene silencing in cell culture affects mitochondrial replication (Garcia-Gomez et al., 2013). The mitochondrial location suggests that primPol may act in

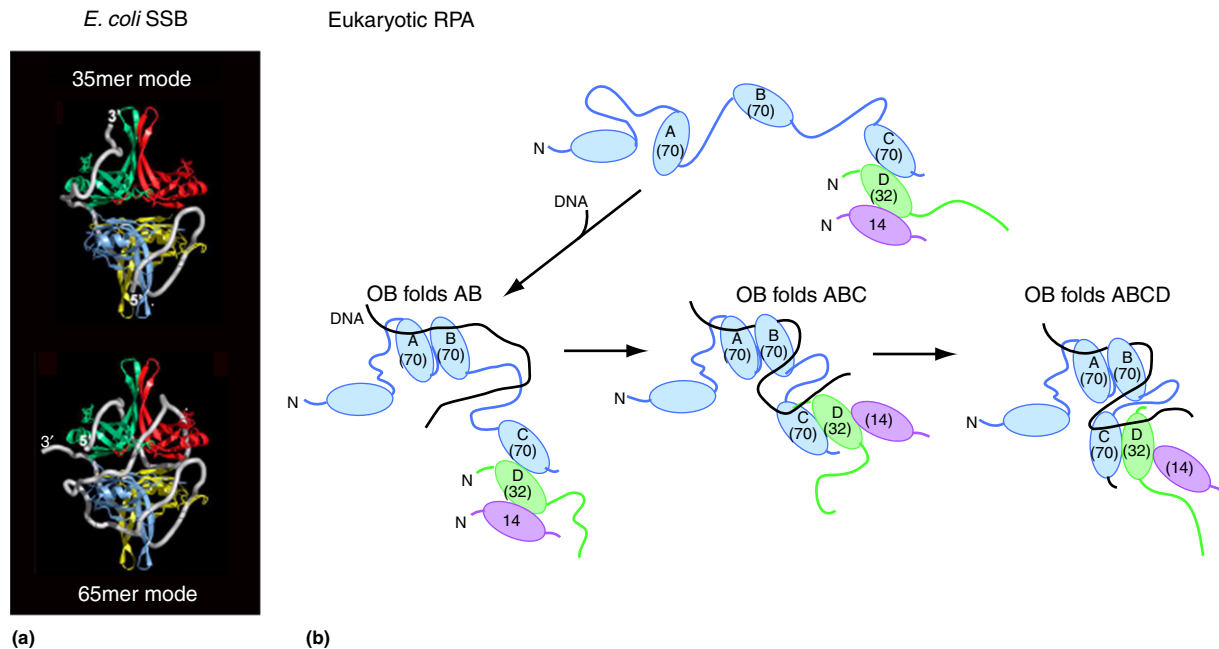


Figure 7 Single-strand DNA-binding proteins of bacteria and eukaryotes. (a) Model of *E. coli* SSB tetramer-ssDNA complexes based on crystal structure information. SSB is a homotetramer of subunits (colored differently); each subunit has one OB fold. The SSB tetramer has two binding modes, 35 nucleotides bound to two subunits and 65 nucleotides bound to all four subunits. The models of these two modes are reproduced with permission from Figure 1a,b in Roy *et al.* (2007). (b) Illustration eukaryotic RPA composed of three different subunits that have six OB folds. Four OB folds bind DNA, labeled A–D. Three binding modes have been detected: Domains AB, then domains ABC, and finally domains ABCD. See text for details. Reprinted from Roy, R., Kozlov, A.G., Lohman, T.M., Ha, T., 2007. Dynamic structural rearrangements between DNA binding modes of *E. coli* SSB Protein. *Journal of Molecular Biology* 369 (5), 1244–1257, Copyright (2007) with permission from Elsevier.

mitochondrial replication (Garcia-Gomez *et al.*, 2013), but knock out mice survive and thus primpol is not essential. Studies of primpol have only just begun, and caution may need to be exercised, considering that many biochemical studies of primpol utilize Mn^{++} instead of Mg^{++} , and Mn^{++} is known to alter the specificity of DNA polymerases for nucleic acid substrates (Kornberg and Baker, 1992).

Single-Strand DNA-Binding Protein

Single-strand DNA-binding proteins are required in all domains of life and fulfill diverse roles in genome metabolism (Flynn and Zou, 2010). During replication, single-strand DNA-binding protein protects ssDNA from nucleases, removes hairpin blocks, and prevents strand reannealing. Bacterial single-strand DNA-binding protein is a homotetramer referred to as SSB, and in eukaryotes it is a heterotrimer referred to as replication protein A (RPA) (Kornberg and Baker, 1992). The sequence and structure of bacterial SSB and eukaryotic RPA indicate that they do not share a common ancestor (Leipe *et al.*, 1999). However, both SSB and RPA bind ssDNA using OB folds, a widespread ‘oligonucleotide and oligosaccharide binding’ motif (Murzin, 1993). The sequence of OB folds can vary considerably, and range from 70 to 150 amino acids. OB folds are comprised of a closed β barrel consisting of 5 tightly coiled antiparallel β sheets with a helix at one end and the

binding site at the other end (Murzin, 1993). OB folds often bind ssDNA but they can also bind RNA and protein.

The Bacterial SSB

The structure of the *E. coli* SSB homotetramer reveals four OB folds, one per monomer, illustrated in Figure 7(a) (Raghuathan *et al.*, 2000). SSB can bind ssDNA in two modes that occupy either 35 or 65 nucleotides (Roy *et al.*, 2007). RPA contains six OB folds, although only four OB folds bind to ssDNA (Figure 7(b); Bochkarev and Bochkareva, 2004; Fanning *et al.*, 2006; Wold, 1997). The RPA1 subunit (p70 in human) contains four OB folds; the OB fold near the N-terminus binds a checkpoint response signaling factor, and the other three bind ssDNA, designated A, B, and C in order of decreasing affinity. RPA2 (p32 in human) also contains a ssDNA binding OB fold, designated D. RPA3 (p14 in human) contains an OB fold that interacts with RPA2.

The Eukaryotic RPA

Unlike bacterial SSB, the OB folds of RPA are connected by flexible linkers that are substantially disordered in the RPA trimer, making structure determination of intact RPA elusive. Regardless of the inherent flexibility in RPA, it binds ssDNA very tightly with an affinity of 10^{-9} – 10^{-10} M (Kim *et al.*, 1994). Notably, the asymmetric structure of RPA binds ssDNA

with a definite 5'-3' polarity, and can impart this directional information to other proteins (de Laat *et al.*, 1998; Iftode and Borowiec, 2000). Biochemical study indicates that RPA can bind ssDNA in three states (illustrated in Figure 7(b)): (1) binding of OB folds A and B occludes about 10 nucleotides, (2) binding of OB folds A, B and C occludes 12–23 nucleotides, and (3) binding of the four OB folds occludes 28–30 nucleotides (Arunkumar *et al.*, 2003; Bastin-Shanower and Brill, 2001; Bochkareva *et al.*, 2002; Kolpashchikov *et al.*, 2001; Kumaran *et al.*, 2006). The functional significance of different conformers is not certain. There also exist different varieties of RPA subunits that form alternative RPA complexes; their functions are insecure, and lie outside of replication (Flynn and Zou, 2010). For example, telomere maintenance requires an alternative form of RPA (Flynn and Zou, 2010).

Comparison of Bacterial and Eukaryotic Replisomes

The proteins described in this review must interdigitate their actions for efficient replication of the DNA duplex. Information of how bacterial replication proteins work together largely come from studies of reconstituted replisomes in the *E. coli* system and its bacteriophages T4 and T7 (Benkovic *et al.*, 2001; Duderstadt *et al.*, 2014; Geertsema and van Oijen, 2013; Johnson and O'Donnell, 2005). Our current knowledge of eukaryotic replisome function largely stems from numerous genetic and cellular-based studies (Bell and Kaguni, 2013; Burgers, 2009; Heller *et al.*, 2011; Kunkel, 2011; Siddiqui *et al.*, 2013). *In vitro* studies of the eukaryotic replisome have centered around the SV40 virus, which uses host proteins except for the viral T-antigen helicase (Fanning and Zhao, 2009; Sowd and Fanning, 2012). Recently a reconstituted system of cellular replication using CMG helicase, RFC, PCNA, RPA and the DNA polymerases has been reported (Georgescu *et al.*, 2014).

In both bacteria and eukaryotes, the helicase is tightly associated with DNA at the replication fork. In contrast, DNA polymerases within replisomes are less tightly associated to replication fork DNA. One reason for this may lie in the fact that replisomes encounter DNA lesions that require other DNA polymerases to traverse them (Indiani *et al.*, 2009; Johnson *et al.*, 2007; Yang *et al.*, 2004; Zhuang *et al.*, 2008). DNA lesions occur frequently in cells due to endogenous attack of DNA by water, reactive oxygen species generated by oxidative metabolism, as well as exogenous genotoxic insults. DNA lesions are difficult for the high fidelity replicative DNA polymerases to pass, stopping a replication fork. One process that enables replication forks to bypass DNA lesions is exchange of the replicative polymerase with one of several translesion synthesis DNA polymerases (TLS Pols), low fidelity enzymes that can extend DNA across a template lesion (Goodman and Woodgate, 2013; Yang, 2014). Insight into polymerases exchange during replication came from studies in the T4 system which showed that two T4 DNA polymerases can trade places on the same sliding clamp while the replisome is moving (Yang *et al.*, 2004). Studies in bacteria demonstrated that TLS polymerases bind the β clamp and can trade places with Pol III stalled on template DNA (Indiani *et al.*, 2009). In this processes, the clamp can accommodate two DNA polymerases at

once, at least transiently. In eukaryotes, the PCNA clamp is ubiquitinated in response to DNA damage (Hoegge *et al.*, 2002), which is associated with TLS DNA polymerase exchange on PCNA (Garg and Burgers, 2005b; Zhuang *et al.*, 2008). The emerging view is that replisomes are not a single entity, but are highly dynamic and more malleable than ever before envisioned (Langston *et al.*, 2009).

The Bacterial Replisome

The central replication proteins of *E. coli* are associated in a particle organized by the clamp loader as illustrated in Figure 8(a) (O'Donnell *et al.*, 2001). The central organizer of this 'replisome' is a trimer of three identical τ subunits in the clamp loader, encoded by the *dnaX* gene (Johnson and O'Donnell, 2005). The C-terminal sequence of τ is not required for the clamp loading reaction, but is essential to cell viability and encodes two domains that extrude from the collar of the clamp loader. These three C-terminal regions bind one DnaB helicase and three molecules of Pol III core. For the last 30 years the paradigm of replisome structure has included only two DNA polymerases (Kornberg and Baker, 1992), and viral systems may, in fact, utilize only two polymerases (Chen *et al.*, 2013). However, the triple polymerase structure of the *E. coli* replisome has been confirmed biochemically and by *in vivo* single-molecule studies of living cells (Georgescu *et al.*, 2011; McInerney *et al.*, 2007; Reyes-Lamothé *et al.*, 2010). Single-molecule studies of the *E. coli* replisome, both *in vitro* and *in vivo*, reveal that all three polymerases are functional at a fork, one for leading strand synthesis, and the other two co-operate to extend the numerous Okazaki fragments (Georgescu *et al.*, 2011; Lia *et al.*, 2012).

The connection between Pol III and DnaB, via τ , results in a 10- to 20-fold stimulation of helicase unwinding (Kim *et al.*, 1996). This stimulation requires an active polymerase, not just binding to τ , making it unlikely that stimulation is derived from a simple allosteric change induced by binding τ (Dallmann *et al.*, 2000). Presumably, the polymerase–helicase connection enables transfer of the energy of polymerization to the helicase. Indeed, functional coupling of the leading polymerase to the helicase generalizes to both phage T4 and phage T7 systems (Dong *et al.*, 1996; Kulczyk *et al.*, 2012; Manosas *et al.*, 2012).

The single clamp loader within the replisome assembles clamps onto both the leading and lagging strands (Turner *et al.*, 1999). Once associated with a β clamp, Pol III is highly processive in DNA synthesis and can extend DNA thousands of nucleotides at a rate exceeding 500 nucleotides, consistent with *in vivo* studies that demonstrate an average rate of the *E. coli* replication fork to be 650 nucleotides (Breier *et al.*, 2005; Yao *et al.*, 2009). In some bacteria, including *E. coli*, the *dnaX* gene produces two proteins, the full length τ , and a smaller γ subunit that is truncated by a translational frameshift. These cells contain a second clamp loader containing three γ subunits instead of three τ subunits. The τ and γ clamp loaders function with equal efficiency in clamp loading (McInerney *et al.*, 2007), but only τ complex binds polymerase and helicase (Gao and McHenry, 2001). The γ complex is proposed to load β clamps onto DNA for the many different DNA

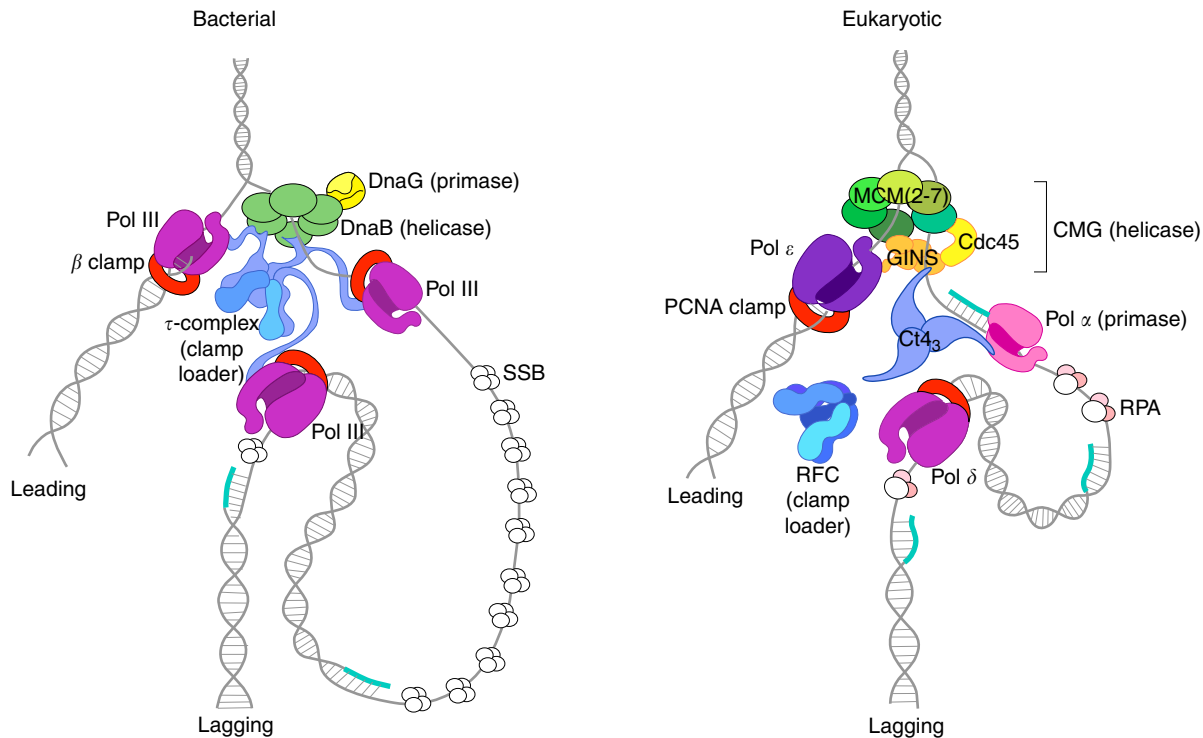


Figure 8 Proposed replisome architectures in bacteria and eukaryotes. Left: The *E. coli* replisome. The DnaB hexamer (purple) encircles the lagging strand. Primase (blue) transiently binds DnaB to produce RNA primers. The clamp loader (green) binds DnaB and three molecules of Pol III (yellow). Sliding β clamps are shown in red. SSB (light blue) coats ssDNA on the lagging strand. The leading strand Pol ϵ binds CMG through contact to Psf1 of the GINS complex within CMG. Ctf4 is a trimer that binds the Sld5 subunit of GINS and the Pol1 subunit of Pol α polymerase/primase. Ctf4 has a weak connection to Pol δ . Both Pol ϵ and Pol δ can function with PCNA (red). The clamp loader, RFC, is not known to be connected to the replisome. The ssDNA on the lagging strand is bound by the RPA heterotrimer.

metabolic processes that utilize β outside the context of a DNA replication fork (McInerney *et al.*, 2007; Yao and O'Donnell, 2009).

Another general feature of *E. coli* and phages T4 and T7 is the association of primase with the helicase in order for primase to become active (Frick and Richardson, 2001; Kornberg and Baker, 1992). The requirement that primase associate with helicase localizes RNA primers to replication forks where they are needed. *E. coli* DnaG primase transiently interacts with DnaB to form RNA primers of 10–12 nucleotides. The T4 primase does likewise, forming 5-nt RNA primers. T7 primase is a separate domain(s) within the gp4 helicase and generates 4mer RNA primers at unique sequences.

Due to the antiparallel structure of duplex DNA, replisomes synthesize one strand (lagging) in the opposite direction of the other (leading strand). The leading strand is continuously extended in the direction of replication fork movement, while the lagging strand is discontinuous and made as a series of numerous Okazaki fragments, necessitated by the unidirectional nature of DNA polymerases. Each Okazaki fragment is initiated by an RNA primer. Extension of RNA primers into Okazaki fragments results in DNA loops, due to the fact that the lagging polymerase remains in contact with the replisome. This was first proposed by Bruce Alberts in the T4 system, and Okazaki loops were referred to as 'trombone loops' because the DNA loops mimic the motion of a

trombone slide while the instrument is played (Sinha *et al.*, 1980). DNA looping has been observed in the electron microscope in the phage T4 and T7 systems (Nossal *et al.*, 2007; Park *et al.*, 1998). DNA looping has also been observed in real-time by single-molecule studies in the T7 system (Hamdan *et al.*, 2009). Single-molecule studies of the *E. coli* replisome, performed in a flow cell in the absence of added polymerase, clamp loader, and helicase, also confirm that DNA loops must form because the replisome synthesizes both strands at an average of 86 kb before dissociating (Yao *et al.*, 2009). Hence, the helicase, polymerase, and clamp loader components form a tightly associated replisome that remains intact during movement, and therefore numerous Okazaki fragment loops must be formed during fork progression (Breier *et al.*, 2005).

The fact that the β clamp endows Pol III with exceptional processivity, well above the length of an Okazaki fragment, raises the question of how the lagging strand polymerase dissociates from DNA upon completing an Okazaki fragment. The release step is triggered by multiple processes, all of which require polymerase to detach from the clamp, leaving the clamp on DNA (O'Donnell, 1987; Yang *et al.*, 2006). The clamp loader repeatedly loads clamps onto RNA primers as they are produced by primase, enabling the lagging strand polymerase to reassociate with a new clamp for each Okazaki fragment. This mechanism of polymerase hopping among clamps was initially

demonstrated in *E. coli* (O'Donnell, 1987; Studwell, *et al.* 1989; Stukenberg *et al.*, 1994), and has also been shown to occur in the T4 system (Hacker and Alberts, 1994).

The mechanism of polymerase release from an Okazaki fragment has been examined in the *E. coli*, T4 and T7 systems. In the context of a working T4 replisome, the lagging strand polymerase is often 'signaled' to release from an Okazaki fragment before the fragment has been completed, leaving a ssDNA gap to be filled in by soluble polymerases (Yang *et al.*, 2006). This 'signal release' of polymerase before completing an Okazaki fragment generalizes to *E. coli* and T7, and is associated with formation of a new primed site. Signal release may be triggered by primase, the new primed site, a new clamp assembled on the primed site, or build up of torsional stress caused by coupled leading/lagging strand synthesis (Georgescu *et al.*, 2011; Hamdan *et al.*, 2009; Kurth *et al.*, 2013; Yang *et al.*, 2006). There also exists a second type of polymerase release, called 'collision release,' because the Okazaki fragment is extended to completion and the polymerase collides into the 5' terminus of the previous Okazaki fragment (Hacker and Alberts, 1994; O'Donnell, 1987; Stukenberg *et al.*, 1994). Dissociation of polymerase from DNA during collision release is a process that is intrinsic to the DNA polymerase and does not require other factors (Georgescu *et al.*, 2009). Both signal release and collision release occur during replisome progression (Duderstadt *et al.*, 2014).

The Eukaryotic Replisome

Pol α was the first eukaryotic replicative DNA polymerase to be purified and characterized (Kaguni *et al.*, 1983b). Identification of a primase activity within the four subunit holoenzyme in Bob Lehman's group (Conaway and Lehman, 1982) was a remarkable surprise, as never before had an enzyme contained both primase and polymerase in one tight complex. Another huge advance was development of the *in vitro* SV40 DNA replication system (Li and Kelly, 1984). The SV40 virus encodes only one replication protein, the T-antigen helicase; all the rest of its replication proteins are encoded by the cell. Intensive studies of the SV40 system in the Kelly, Stillman, and Hurwitz laboratories identified Pol δ , RFC, PCNA and RPA (Fanning and Zhao, 2009; Sowd and Fanning, 2012). Upon discovery that Pol δ was a replicative enzyme, Pol α was thought to replicate the lagging strand (since it makes primers) and that Pol δ replicated the leading strand. But further work showed that Pol δ replicated both strands while Pol α served as a primase (Stillman, 2008). Genetic studies soon thereafter identified Pol ϵ as a third replicative cellular polymerase, while SV40 did not utilize Pol ϵ , indicating that SV40 did not replicate quite the same as cellular replisomes (Zlotkin *et al.*, 1996). The question of which strands Pol δ and Pol ϵ operate on during chromosome replication was unresolved for many years. Cleaver genetic studies by Tom Kunkel's group and his collaborators have now conclusively shown that Pol ϵ replicates the bulk of the leading strand, while Pol δ replicates the lagging strand (Kunkel and Burgers, 2008; Pursell *et al.*, 2007).

At the head of the replisome is the 11 subunit CMG helicase (Costa *et al.*, 2011; Ilves *et al.*, 2010; Kang *et al.*, 2012;

Moyer *et al.*, 2006). As discussed earlier in this review, CMG tracks along the leading strand, and CMG is known to bind numerous proteins. Proteins that bind CMG have been identified by epitope tagging CMG subunit genes, followed by antibody pull-outs from cells to identify CMG interactive proteins by mass spectrometry (Gambus *et al.*, 2006). The protein assemblage thus identified is referred to as the replisome progression complex (RPC), and includes CMG, Ctf4, Mcm10, Fact (a nucleosome remodeling complex), Topo I, and three proteins involved in the DNA damage checkpoint (Csm3, Mrc1, and Tof1). Interestingly, RPC does not contain DNA polymerases, implying that DNA polymerases are not tightly associated with CMG. This may not be surprising given the inability to isolate helicase-polymerase complexes in bacterial systems, even though the helicase and polymerase are known to interact. However, when CMG is isolated under mild conditions Pol α co-purifies with RPC (Gambus *et al.*, 2009) and this association requires the presence of Ctf4, an oligomer that has long been known to bind Pol α (Miles and Formosa, 1992). The Ctf4 oligomer also binds to CMG through the Sld5 subunit of GINS (Simon *et al.*, 2014). Interestingly, the GINS complex binds to Pol ϵ through the Psf1 subunit of GINS (Bermudez *et al.*, 2011; Sengupta *et al.*, 2013). This connection implies that Pol ϵ binds CMG directly, which is somewhat analogous to bacterial replisomes in which the leading polymerase binds the helicase during coupled polymerase-helicase fork progression. The rate of eukaryotic replication forks is much slower than in bacteria and archaeal cells. In budding yeast, fork progression is estimated at 30–60 bp/s (Raghuraman *et al.*, 2001; Sekedat *et al.*, 2010; Yabuki *et al.*, 2002), 10- to 20-fold slower than in *E. coli*.

Interestingly, the crystal structure of Ctf4 shows it is a homotrimer (Simon *et al.*, 2014). Hence, it is tempting to speculate that Ctf4 may be the scaffolding equivalent of the bacterial τ subunit, a trimer that links the helicase with DNA polymerases (compare Figures 8(a) and 8(b)). Indeed, a weak interaction of Pol δ with Ctf4 has been reported (Bermudez *et al.*, 2010). Although Ctf4 is not essential in budding yeast, Ctf4 is essential in fission yeast and other eukaryotes. Mcm10 is essential to cell viability and it binds to Pol α , although the exact function of Mcm10 in DNA replication remains unknown (Warren *et al.*, 2009). Some studies suggest Mcm10 is required during origin activation, while other studies indicate Mcm10 acts at a late stage, when replisomes are assembled (Heller *et al.*, 2011; Kanke *et al.*, 2012; Siddiqui *et al.*, 2013; van Deursen *et al.*, 2012). The Csm3, Tof1, Mrc1 proteins of the RPC are not essential, but their involvement in the checkpoint response suggests they may regulate replisome function in response to DNA damage. It is likely that many other proteins will be shown to move and function with replication forks in future studies.

It seems likely that the eukaryotic replisome is more dynamic than the bacterial replisome. The affinity of Ctf4 for CMG helicase appears stable to glycerol gradient analysis, implying a tight connection between these two proteins (Kang *et al.*, 2013). But the affinities of the polymerases ϵ , α , and δ to the replisome appear much weaker (Gambus *et al.*, 2009; Bermudez *et al.*, 2010). Therefore the eukaryotic polymerase may loosely interact with the replisome during fork progression, rather than forming a stable connection between the

leading and lagging strand polymerases as observed in the bacterial replisome in which the τ subunit tightly binds three molecules of Pol III in a stable fashion. Looser connections of the eukaryotic polymerases with the replisome implies that the DNA polymerases are more dynamic and form transient connections to Ctf4 and/or CMG subunits, producing a dynamic replisome. It is possible that a more dynamic replisome, with more transient polymerase connections enables translesion synthesis DNA polymerases to enter the replisome to help the replisome move forward if it encounters a template lesion. Dynamic polymerase action in the replisome may also facilitate the many protein exchanges required in the confines of short Okazaki fragments during their maturation prior to ligation. For example, eukaryotic Okazaki fragments only average 160 bp, and many operations occur within this short section of DNA. First Pol α makes a RNA-DNA primer, then RFC loads PCNA onto the primed site. Pol δ then extends the short Okazaki fragment and a nucleosome is thought to assemble on the new Okazaki fragment even before fragments are sealed. The RNA-DNA primers must also be displaced and removed by nuclease action before ligase can seal two Okazaki fragments together. Thus a dynamic replisome with DNA polymerases capable of coming on and off the replisome may be important to make way for these many actions within a short region of DNA.

DNA looping has served as a paradigm in the bacterial replication field. However, the more dynamic a replisome, the less likely that connections among replisome components will create DNA loops during lagging strand synthesis. It is interesting to note that DNA loops have no known function; they are simply documented to exist in bacterial systems. Perhaps DNA loops are simply an artifact of the use of identical polymerases on leading and lagging strands, in which the leading polymerase binds the helicase, which itself is an assembly of identical subunits. Hence, the lagging strand polymerase may contact the helicase simply as a consequence of the interactions that evolved to enhance function on the leading strand, resulting in DNA loops during lagging strand synthesis. DNA looping is a very low energetic process. Assuming that DNA loops carry no energetic penalty, there would be no selective pressure to prevent their formation on the lagging strand.

The mechanistic basis for the asymmetric assembly of different DNA polymerases onto their respective leading and lagging strands of the eukaryotic fork has recently been determined (Georgescu *et al.*, 2014). CMG helicase stabilizes Pol ϵ on the leading strand, enabling processive DNA synthesis by Pol ϵ even in the absence of PCNA. But Pol δ is not stabilized by CMG, and acts distributively on the leading strand. The distributive behavior of Pol δ with CMG on the leading strand contrasts sharply with the high processivity demonstrated by yeast Pol δ -PCNA on primed ssDNA (Langston and O'Donnell, 2008). The stable function of Pol ϵ with CMG, and unstable function of Pol δ with CMG, likely underlies the basis by which Pol ϵ is directed to the leading strand. Protein binding experiments show that Pol δ binds PCNA much tighter than Pol ϵ (Chilkova *et al.*, 2007), and competition experiments show that Pol δ effectively out-competes Pol ϵ for use of PCNA clamps on primed ssDNA (Georgescu *et al.*, 2014). The greater affinity of Pol δ for PCNA clamps compared

to Pol ϵ may explain how Pol δ is selected for function on the lagging strand over Pol ϵ . However, it is worth noting that the human Pol δ is not highly processive with PCNA (Hu *et al.*, 2012), and therefore further studies will be needed to determine if these principles generalize to human replisomes.

Fork progression in eukaryotes is regulated by the DNA damage checkpoint response. This pathway recognizes the presence of DNA lesions and activates a kinase that phosphorylates a variety of targets. Among the targets of these kinases are Mcm and GINS subunits of CMG (Cortez *et al.*, 2004; Ilves *et al.*, 2012; Segurado and Diffley, 2008). Activity assays of phosphorylated CMG demonstrate that helicase action is down-regulated in response to phosphorylation (De Piccoli *et al.*, 2012; Ilves *et al.*, 2012). Cell-based studies indicate that replication forks stop, or move slowly upon activation of the DNA damage checkpoint (Labib and De Piccoli, 2011). However, phosphorylation may not be the only explanation for how the replisome is halted upon DNA damage, and recent biochemical studies provide another possibility (Bruck and Kaplan, 2013). The Cdc45 subunit of CMG contains a ssDNA binding activity that requires about 40 nucleotides of ssDNA. Mutation of the DNA interactive residues in Cdc45 yields a Cdc45 mutant with reduced affinity for ssDNA (Bruck and Kaplan, 2013). Interestingly, study of cells with this Cdc45 ssDNA binding mutant indicate that replication forks no longer stop upon activation of the DNA damage checkpoint (Bruck and Kaplan, 2013). These findings have been interpreted as follows. When the replication fork encounters a leading strand lesion, Pol ϵ stops at the lesion, but CMG continues unwinding. When 40 nucleotides of ssDNA are exposed, Cdc45 binds the ssDNA and stops the fork. The Cdc45 DNA binding mutant is unable to bind ssDNA and stop the fork upon encounter with DNA damage. Further study will be needed to develop a more complete understanding of how replisomes are controlled by the DNA damage checkpoint. In addition to phosphorylation in response to DNA damage, numerous replication fork proteins are phosphorylated in a cell cycle specific fashion, including Pols α , δ , and ϵ , RPA, and particular Mcm subunits (Errico and Costanzo, 2012; Gangavarapu *et al.*, 2011; Olson *et al.*, 2006; Sheu *et al.*, 2014; Sheu and Stillman, 2010). In most cases, the exact kinases and the roles these modifications are not yet understood.

What Is Needed for the Future?

The process of replication has been studied for many decades, and the basic outline of the process is understood. The facts that replication requires 5'-3' DNA polymerases, dNTPs, a proofreading 3'-5' exonuclease, hexameric helicase, RNA primase, and SSB to replicate DNA in a semi-discontinuous manner required the effort of numerous labs for about 30 years after the discovery of the DNA structure. Identification of accessory protein function as clamps and clamp loaders, also a universal feature of cells, took another 20 years. The organization of these proteins, and how they deal with DNA lesions is also reasonably well understood, although significant details about the mechanism and how the

replisome interfaces with repair and recombination is still in progress.

While it may seem that the major questions about bacterial replication are answered, this is certainly not the case for eukaryotic cells. New proteins that act at eukaryotic replication forks are being identified at a rapid pace. The functional consequences of cell cycle specific phosphorylation is not understood, nor are the targets and function of other modifications such as acetylation, ubiquitination, and sumoylation, most of which are not covered in this review. It remains unknown how eukaryotic replisomes progress through nucleosomes, especially highly condensed chromatin. Furthermore, episomal modifications of nucleosomes that regulate the transcriptome of a cell must be preserved during development of a multicellular organism, and the mechanism by which 'marked nucleosomes' are transferred from the parental duplex to sister chromatids is essentially unexplored. The process of sister chromatid cohesion is also intimately involved with the replication process. It is presumed that the replisome progresses through cohesion rings, although whether this occurs, and if so, what factors are required, is not at all clear. The process of assembly of cohesion rings is also still unknown. There exist at least three alternative clamp loaders in which the RFC1 subunit is exchanged for another protein and one of these alternative clamp loaders is genetically linked to the process of cohesion, but no data exists to suggest what special function this alternative clamp loader performs for this action.

Replisomes are assembled at origins in a highly regulated multistep process, not reviewed here. It is during this origin activation process that the CMG is assembled. First the Mcm rings are placed onto duplex DNA and at some point, GINS and Cdc45 are brought into the Mcm complex and the complex must transit from encircling duplex DNA to encircling leading strand ssDNA. Reconstitution of this process and an understanding of how the many assembly factors function is shrouded in mystery. Reconstitution of a functioning eukaryotic replisome from pure proteins also remains an important goal. Study *in vitro* of these multiprotein reactions will be required to gain detailed insight into the structure and function of the many proteins that regulate origin firing and perform the mechanics of replication fork progression. Many repair processes are also known to require replication fork proteins, sometimes occurring at the fork itself. For example, break-induced repair involves the initiation of a functional fork at a DNA break that continues for many kilobases until the end of a linear chromosome. This process appears to require many of the same proteins that move normal replication forks, yet it is mutagenic and its exact mechanism has only begun to be explored. Eukaryotes also have telomeres that require the telomerase polymerase, and several other proteins, including some of those at the fork. Many human diseases are also the result of defects in one of these many processes. We can hope that some of the insights gained from a greater understanding of their mechanistic basis will prevent, or at least treat some of these diseases. These, and many other important questions will require intensive genetic, cellular, biochemical, and structural studies. It is only reasonable to expect that the answers to these important questions will be very exciting, but also will take many years of research from numerous bright and energetic scientists.

Acknowledgment

The authors are grateful for funding from the National Institutes of Health, US (GM38839) and the Howard Hughes Medical Institute.

See also: Nucleic acid Synthesis/Breakdown: DNA Synthesis/Repair: Telomeres and Telomerase; The Base Excision Repair Pathway. Nucleic acid Synthesis/Breakdown: RNA Synthesis/Function: Riboswitches and Ribozymes

References

- Ahnert, P., Patel, S.S., 1997. Asymmetric interactions of hexameric bacteriophage T7 DNA helicase with the 5'- and 3'-tails of the forked DNA substrate. *Journal of Biological Chemistry* 272, 32267–32273.
- Alberts, B.M., 1984. The DNA enzymology of protein machines. *Cold Spring Harbor Symposia on Quantitative Biology* 49, 1–12.
- Aravind, L., Leippe, D.D., Koonin, E.V., 1998. Toprim – A conserved catalytic domain in type IA and II topoisomerases, DnaG-type primases, OLD family nucleases and RecR proteins. *Nucleic Acids Research* 26, 4205–4213.
- Arunkumar, A.I., Stauffer, M.E., Bochkareva, E., Bochkarev, A., Chazin, W.J., 2003. Independent and coordinated functions of replication protein A tandem high affinity single-stranded DNA binding domains. *Journal of Biological Chemistry* 278, 41077–41082.
- Ason, B., Handayani, R., Williams, C.R., *et al.*, 2003. Mechanism of loading the *Escherichia coli* DNA polymerase III beta sliding clamp on DNA. Bona fide primer/templates preferentially trigger the gamma complex to hydrolyze ATP and load the clamp. *Journal of Biological Chemistry* 278, 10033–10040.
- Augustin, M.A., Huber, R., Kaiser, J.T., 2001. Crystal structure of a DNA-dependent RNA polymerase (DNA primase). *Nature Structural & Molecular Biology* 8, 57–61.
- Bailey, S., Eliason, W.K., Steitz, T.A., 2007. Structure of hexameric DnaB helicase and its complex with a domain of DnaG primase. *Science* 318, 459–463.
- Bailey, S., Wing, R.A., Steitz, T.A., 2006. The structure of *T. aquaticus* DNA polymerase III is distinct from eukaryotic replicative DNA polymerases. *Cell* 126, 893–904.
- Balakrishnan, L., Bambara, R.A., 2013. Flap endonuclease 1. *Annual Review of Biochemistry* 82, 119–138.
- Barros, T., Guenther, J., Kelch, B., *et al.*, 2013. A structural role for the PHP domain in *E. coli* DNA polymerase III. *BMC Structural Biology* 13, 8.
- Bastin-Shanower, S.A., Brill, S.J., 2001. Functional analysis of the four DNA binding domains of replication protein A. The role of RPA2 in ssDNA binding. *Journal of Biological Chemistry* 276, 36446–36453.
- Beattie, T.R., Bell, S.D., 2011. Molecular machines in archaeal DNA replication. *Current Opinion in Chemical Biology* 15, 614–619.
- Beese, L.S., Steitz, T.A., 1991. Structural basis for the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I: A two metal ion mechanism. *EMBO Journal* 10, 25–33.
- Bell, S.D., Botchan, M.R., 2013. The minichromosome maintenance replicative helicase. *Cold Spring Harbor Perspectives in Biology* 5, a012807.
- Bell, S.P., Kaguni, J.M., 2013. Helicase loading at chromosomal origins of replication. *Cold Spring Harbor Perspectives in Biology* 5.
- Benkovic, S.J., Valentine, A.M., Salinas, F., 2001. Replisome-mediated DNA replication. *Annual Review of Biochemistry* 70, 181–208, pii: a010124.
- Bermudez, V.P., Farina, A., Raghavan, V., Tappin, I., Hurwitz, J., 2011. Studies on human DNA polymerase epsilon and GINS complex and their role in DNA replication. *Journal of Biological Chemistry* 286, 28963–28977.
- Bermudez, V.P., Farina, A., Tappin, I., Hurwitz, J., 2010. Influence of the human cohesion establishment factor Ctf4/AND-1 on DNA replication. *Journal of Biological Chemistry* 285, 9493–9505.
- Bianchi, J., Rudd, S.G., Jozwiakowski, S.K., *et al.*, 2013. PrimPol bypasses UV photoproducts during eukaryotic chromosomal DNA replication. *Molecular Cell* 52, 566–573.
- Bochkarev, A., Bochkareva, E., 2004. From RPA to BRCA2: Lessons from single-stranded DNA binding by the OB-fold. *Current Opinion in Structural Biology* 14, 36–42.

- Bochkareva, E., Korolev, S., Lees-Miller, S.P., Bochkarev, A., 2002. Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO Journal* 21, 1855–1863.
- Bochman, M.L., Bell, S.P., Schwacha, A., 2008. Subunit organization of Mcm2-7 and the unequal role of active sites in ATP hydrolysis and viability. *Molecular and Cellular Biology* 28, 5865–5873.
- Bochman, M.L., Schwacha, A., 2008. The Mcm2-7 complex has *in vitro* helicase activity. *Molecular Cell* 31, 287–293.
- Bocquier, A.A., Liu, L., Cann, I.K., *et al.*, 2001. Archaeal primase: Bridging the gap between RNA and DNA polymerases. *Current Biology* 11, 452–456.
- Bowman, G.D., Goedken, E.R., Kazmirski, S.L., O'Donnell, M., Kuriyan, J., 2005. DNA polymerase clamp loaders and DNA recognition. *FEBS Letters* 579, 863–867.
- Bowman, G.D., O'Donnell, M., Kuriyan, J., 2004. Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature* 429, 724–730.
- Breier, A.M., Weier, H.U., Cozzarelli, N.R., 2005. Independence of replisomes in *Escherichia coli* chromosomal replication. *Proceedings of the National Academy of Sciences of the United States of America* 102, 3942–3947.
- Bruck, I., Kaplan, D.L., 2013. Cdc45 protein-single-stranded DNA interaction is important for stalling the helicase during replication stress. *Journal of Biological Chemistry* 288, 7550–7563.
- Bunz, F., Kobayashi, R., Stillman, B., 1993. cDNAs encoding the large subunit of human replication factor C. *Proceedings of the National Academy of Sciences of the United States of America* 90, 11014–11018.
- Burgers, P.M., 2009. Polymerase dynamics at the eukaryotic DNA replication fork. *Journal of Biological Chemistry* 284, 4041–4045.
- Chen, D., Yue, H., Spiering, M.M., Benkovic, S.J., 2013. Insights into Okazaki fragment synthesis by the T4 replisome: The fate of lagging-strand holoenzyme components and their influence on Okazaki fragment size. *Journal of Biological Chemistry* 288, 20807–20816.
- Chia, N., Cann, I., Olsen, G.J., 2010. Evolution of DNA replication protein complexes in eukaryotes and Archaea. *PLoS One* 5, e10866.
- Chilkova, O., Stenlund, P., Isoz, I., *et al.*, 2007. The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA. *Nucleic Acids Research* 35, 6588–6597.
- Chong, J.P., Hayashi, M.K., Simon, M.N., Xu, R.M., Stillman, B., 2000. A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. *Proceedings of the National Academy of Sciences of the United States of America* 97, 1530–1535.
- Conaway, R.C., Lehman, I.R., 1982. A DNA primase activity associated with DNA polymerase alpha from *Drosophila melanogaster* embryos. *Proceedings of the National Academy of Sciences of the United States of America* 79, 2523–2527.
- Copeland, W.C., Wang, T.S., 1993. Enzymatic characterization of the individual mammalian primase subunits reveals a biphasic mechanism for initiation of DNA replication. *Journal of Biological Chemistry* 268, 26179–26189.
- Cortez, D., Glick, G., Elledge, S.J., 2004. Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proceedings of the National Academy of Sciences of the United States of America* 101, 10078–10083.
- Costa, A., Ilves, I., Tamberg, N., *et al.*, 2011. The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nature Structural & Molecular Biology* 18, 471–477.
- Cullmann, G., Fien, K., Kobayashi, R., Stillman, B., 1995. Characterization of the five replication factor C genes of *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 15, 4661–4671.
- Dallmann, H.G., Kim, S., Pritchard, A.E., Mariani, K.J., McHenry, C.S., 2000. Characterization of the unique C terminus of the *Escherichia coli* tau DnaX protein. Monomeric C-tau binds alpha AND DnaB and can partially replace tau in reconstituted replication forks. *Journal of Biological Chemistry* 275, 15512–15519.
- Dalrymple, B.P., Kongsuwan, K., Wijffels, G., Dixon, N.E., Jennings, P.A., 2001. A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proceedings of the National Academy of Sciences of the United States of America* 98, 11627–11632.
- De Biasio, A., Blanco, F.J., 2013. Proliferating cell nuclear antigen structure and interactions: Too many partners for one dancer? *Advances in Protein Chemistry and Structural Biology* 91, 1–36.
- Delagoutte, E., von Hippel, P.H., 2001. Molecular mechanisms of the functional coupling of the helicase (gp41) and polymerase (gp43) of bacteriophage T4 within the DNA replication fork. *Biochemistry* 40, 4459–4477.
- De Piccoli, G., Katou, Y., Itoh, T., *et al.*, 2012. Replisome stability at defective DNA replication forks is independent of S phase checkpoint kinases. *Molecular Cell* 45, 696–704.
- van Deursen, F., Sengupta, S., De Piccoli, G., Sanchez-Diaz, A., Labib, K., 2012. Mcm10 associates with the loaded DNA helicase at replication origins and defines a novel step in its activation. *EMBO Journal* 31, 2195–2206.
- Dionne, I., Brown, N.J., Woodgate, R., Bell, S.D., 2008. On the mechanism of loading the PCNA sliding clamp by RFC. *Molecular Microbiology* 68, 216–222.
- Dong, F., Weitzel, S.E., von Hippel, P.H., 1996. A coupled complex of T4 DNA replication helicase (gp41) and polymerase (gp43) can perform rapid and processive DNA strand-displacement synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 93, 14456–14461.
- Donmez, I., Patel, S.S., 2006. Mechanisms of a ring shaped helicase. *Nucleic Acids Research* 34, 4216–4224.
- Donmez, I., Patel, S.S., 2008. Coupling of DNA unwinding to nucleotide hydrolysis in a ring-shaped helicase. *EMBO Journal* 27, 1718–1726.
- Doublie, S., Sawaya, M.R., Ellenberger, T., 1999. An open and closed case for all polymerases. *Structure* 7, R31–R35.
- Duderstadt, K.E., Reyes-Lamothe, R., van Oijen, A.M., Sherratt, D.J., 2014. Replication-fork dynamics. *Cold Spring Harbor Perspectives in Biology* 6, pii: a010157.
- Ellison, V., Stillman, B., 1998. Reconstitution of recombinant human replication factor C (RFC) and identification of an RFC subcomplex possessing DNA-dependent ATPase activity. *Journal of Biological Chemistry* 273, 5979–5987.
- Enemark, E.J., Joshua-Tor, L., 2006. Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* 442, 270–275.
- Enemark, E.J., Joshua-Tor, L., 2008. On helicases and other motor proteins. *Current Opinion in Structural Biology* 18, 243–257.
- Errico, A., Costanzo, V., 2012. Mechanisms of replication fork protection: A safeguard for genome stability. *Critical Reviews in Biochemistry and Molecular Biology* 47, 222–235.
- Erzberger, J.P., Berger, J.M., 2006. Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annual Review of Biophysics and Biomolecular Structure* 35, 93–114.
- Fanning, E., Klimovich, V., Nager, A.R., 2006. A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Research* 34, 4126–4137.
- Fanning, E., Zhao, K., 2009. SV40 DNA replication: From the A gene to a nanomachine. *Virology* 384, 352–359.
- Fien, K., Stillman, B., 1992. Identification of replication factor C from *Saccharomyces cerevisiae*: A component of the leading-strand DNA replication complex. *Molecular and Cellular Biology* 12, 155–163.
- Flynn, R.L., Zou, L., 2010. Oligonucleotide/oligosaccharide-binding fold proteins: A growing family of genome guardians. *Critical Reviews in Biochemistry and Molecular Biology* 45, 266–275.
- Forterre, P., 2013. The common ancestor of archaea and eukarya was not an archaeon. *Archaea* 2013, 372396.
- Forterre, P., Filee, J., Myllykallio, H., 2004. Origin and Evolution of DNA and DNA Replication Machinery, October 4, 2004 edit. Georgetown, TX: Landes Bioscience, pp. 24.
- Frick, D.N., Richardson, C.C., 2001. DNA primases. *Annual Review of Biochemistry* 70, 39–80.
- Gambus, A., van Deursen, F., Polychronopoulos, D., *et al.*, 2009. A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. *EMBO Journal* 28, 2992–3004.
- Gambus, A., Jones, R.C., Sanchez-Diaz, A., *et al.*, 2006. GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. *Nature Cell Biology* 8, 358–366.
- Gangavarapu, V., Santa Maria, S.R., Prakash, S., Prakash, L., 2011. Requirement of replication checkpoint protein kinases Mec1/Rad53 for postreplication repair in yeast. *MBio* 2, e00079–00011.
- Gao, D., McHenry, C.S., 2001. Tau binds and organizes *Escherichia coli* replication proteins through distinct domains. Domain IV, located within the unique C terminus of tau, binds the replication fork, helicase, DnaB. *Journal of Biological Chemistry* 276, 4441–4446.
- Garcia-Gomez, S., Reyes, A., Martinez-Jimenez, M.I., *et al.*, 2013. PrimPol, an archaic primase/polymerase operating in human cells. *Molecular Cell* 52, 541–553.
- Garg, P., Burgers, P.M., 2005a. DNA polymerases that propagate the eukaryotic DNA replication fork. *Critical Reviews in Biochemistry and Molecular Biology* 40, 115–128.
- Garg, P., Burgers, P.M., 2005b. Ubiquitinated proliferating cell nuclear antigen activates translesion DNA polymerases eta and REV1. *Proceedings of the National Academy of Sciences of the United States of America* 102, 18361–18366.

- Geertsema, H.J., van Oijen, A.M., 2013. A single-molecule view of DNA replication: The dynamic nature of multi-protein complexes revealed. *Current Opinion in Structural Biology* 23, 788–793.
- Georgescu, R.E., Kurth, I., O'Donnell, M.E., 2011. Single-molecule studies reveal the function of a third polymerase in the replisome. *Nature Structural & Molecular Biology* 19, 113–116.
- Georgescu, R.E., Kurth, I., Yao, N.Y., *et al.*, 2009. Mechanism of polymerase collision release from sliding clamps on the lagging strand. *EMBO Journal* 28, 2981–2991.
- Georgescu, R.E., Langston, L.D., Yao, N.Y., *et al.*, 2014. Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. *Nature Structural & Molecular Biology* 21, 664–670.
- Gomes, X.V., Gary, S.L., Burgers, P.M., 2000. Overproduction in *Escherichia coli* and characterization of yeast replication factor C lacking the ligase homology domain. *Journal of Biological Chemistry* 275, 14541–14549.
- Goodman, M.F., Woodgate, R., 2013. Translesion DNA polymerases. *Cold Spring Harbor Perspectives in Biology* 5, a010363.
- Graham, B.W., Schauer, G.D., Leuba, S.H., Trakselis, M.A., 2011. Steric exclusion and wrapping of the excluded DNA strand occurs along discrete external binding paths during MCM helicase unwinding. *Nucleic Acids Research* 39, 6585–6595.
- Guenther, B., Onrust, R., Sali, A., O'Donnell, M., Kuriyan, J., 1997. Crystal structure of the delta' subunit of the clamp-loader complex of *E. coli* DNA polymerase III. *Cell* 91, 335–345.
- Gulbis, J.M., Kelman, Z., Hurwitz, J., O'Donnell, M., Kuriyan, J., 1996. Structure of the C-terminal region of p21(WAF1/CIP1) complexed with human PCNA. *Cell* 87, 297–306.
- Hacker, K.J., Alberts, B.M., 1994. The rapid dissociation of the T4 DNA polymerase holoenzyme when stopped by a DNA hairpin helix. A model for polymerase release following the termination of each Okazaki fragment. *Journal of Biological Chemistry* 269, 24221–24228.
- Hacker, K.J., Johnson, K.A., 1997. A hexameric helicase encircles one DNA strand and excludes the other during DNA unwinding. *Biochemistry* 36, 14080–14087.
- Hamdan, S.M., Loparo, J.J., Takahashi, M., Richardson, C.C., van Oijen, A.M., 2009. Dynamics of DNA replication loops reveal temporal control of lagging-strand synthesis. *Nature* 457, 336–339.
- Heller, R.C., Kang, S., Lam, W.M., *et al.*, 2011. Eukaryotic origin-dependent DNA replication *in vitro* reveals sequential action of DDK and S-CDK kinases. *Cell* 146, 80–91.
- von Hippel, P.H., Delagoutte, E., 2001. A general model for nucleic acid helicases and their "coupling" within macromolecular machines. *Cell* 104, 177–190.
- Hoege, C., Pfander, B., Moldovan, G.L., Pyrowolakis, G., Jentsch, S., 2002. RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. *Nature* 419, 135–141.
- Hombauer, H., Campbell, C.S., Smith, C.E., Desai, A., Kolodner, R.D., 2011. Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell* 147, 1040–1053.
- Hu, Z., Perumal, S.K., Yue, H., Benkovic, S.J., 2012. The human lagging strand DNA polymerase delta holoenzyme is distributive. *Journal of Biological Chemistry* 287, 38442–38448.
- Hubscher, U., Maga, G., 2011. DNA replication and repair bypass machines. *Current Opinion in Chemical Biology* 15, 627–635.
- Hubscher, U., Maga, G., Spadari, S., 2002. Eukaryotic DNA polymerases. *Annual Review of Biochemistry* 71, 133–163.
- Iftode, C., Borowiec, J.A., 2000. 5' → 3' molecular polarity of human replication protein A (hRPA) binding to pseudo-origin DNA substrates. *Biochemistry* 39, 11970–11981.
- Ivles, I., Petojevic, T., Pesavento, J.J., Botchan, M.R., 2010. Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Molecular Cell* 37, 247–258.
- Ivles, I., Tamberg, N., Botchan, M.R., 2012. Checkpoint kinase 2 (Chk2) inhibits the activity of the Cdc45/MCM2-7/GINS (CMG) replicative helicase complex. *Proceedings of the National Academy of Sciences of the United States of America* 109, 13163–13170.
- Indiani, C., Langston, L.D., Yurieva, O., Goodman, M.F., O'Donnell, M., 2009. Translesion DNA polymerases remodel the replisome and alter the speed of the replicative helicase. *Proceedings of the National Academy of Sciences of the United States of America* 106, 6031–6038.
- Ishino, S., Kelman, L.M., Kelman, Z., Ishino, Y., 2013. The archaeal DNA replication machinery: Past, present and future. *Genes & Genetic Systems* 88, 315–319.
- Itsathiphaisarn, O., Wing, R.A., Eliason, W.K., Wang, J., Steitz, T.A., 2012. The hexameric helicase DnaB adopts a nonplanar conformation during translocation. *Cell* 151, 267–277.
- Jeong, Y.J., Levin, M.K., Patel, S.S., 2004. The DNA-unwinding mechanism of the ring helicase of bacteriophage T7. *Proceedings of the National Academy of Sciences of the United States of America* 101, 7264–7269.
- Jeruzalmi, D., O'Donnell, M., Kuriyan, J., 2001. Crystal structure of the processivity clamp loader gamma (gamma) complex of *E. coli* DNA polymerase III. *Cell* 106, 429–441.
- Jeruzalmi, D., O'Donnell, M., Kuriyan, J., 2002. Clamp loaders and sliding clamps. *Current Opinion in Structural Biology* 12, 217–224.
- Johansson, E., Macneill, S.A., 2010. The eukaryotic replicative DNA polymerases take shape. *Trends in Biochemical Sciences* 35, 339–347.
- Johnson, A., O'Donnell, M., 2005. Cellular DNA replicases: Components and dynamics at the replication fork. *Annual Review of Biochemistry* 74, 283–315.
- Johnson, D.E., Takahashi, M., Hamdan, S.M., Lee, S.J., Richardson, C.C., 2007. Exchange of DNA polymerases at the replication fork of bacteriophage T7. *Proceedings of the National Academy of Sciences of the United States of America* 104, 5312–5317.
- Johnson, D.S., Bai, L., Smith, B.Y., Patel, S.S., Wang, M.D., 2007. Single-molecule studies reveal dynamics of DNA unwinding by the ring-shaped T7 helicase. *Cell* 129, 1299–1309.
- Johnson, S.J., Taylor, J.S., Beese, L.S., 2003. Processive DNA synthesis observed in a polymerase crystal suggests a mechanism for the prevention of frameshift mutations. *Proceedings of the National Academy of Sciences of the United States of America* 100, 3895–3900.
- Jose, D., Weitzel, S.E., von Hippel, P.H., 2012. Breathing fluctuations in position-specific DNA base pairs are involved in regulating helicase movement into the replication fork. *Proceedings of the National Academy of Sciences of the United States of America* 109, 14428–14433.
- Joyce, C.M., Potapova, O., Delucia, A.M., *et al.*, 2008. Fingers-closing and other rapid conformational changes in DNA polymerase I (Klenow fragment) and their role in nucleotide selectivity. *Biochemistry* 47, 6103–6116.
- Kaguni, L.S., Rossignol, J.M., Conaway, R.C., Banks, G.R., Lehman, I.R., 1983a. Association of DNA primase with the beta/gamma subunits of DNA polymerase alpha from *Drosophila melanogaster* embryos. *Journal of Biological Chemistry* 258, 9037–9039.
- Kaguni, L.S., Rossignol, J.M., Conaway, R.C., Lehman, I.R., 1983b. Isolation of an intact DNA polymerase-primase from embryos of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America* 80, 2221–2225.
- Kang, Y.H., Farina, A., Bermudez, V.P., *et al.*, 2013. Interaction between human Ctf4 and the Cdc45/Mcm2-7/GINS (CMG) replicative helicase. *Proceedings of the National Academy of Sciences of the United States of America* 110, 19760–19765.
- Kang, Y.H., Galal, W.C., Farina, A., Tappin, I., Hurwitz, J., 2012. Properties of the human Cdc45/Mcm2-7/GINS helicase complex and its action with DNA polymerase epsilon in rolling circle DNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 109, 6042–6047.
- Kanke, M., Kodama, Y., Takahashi, T.S., Nakagawa, T., Masukata, H., 2012. Mcm10 plays an essential role in origin DNA unwinding after loading of the CMG components. *EMBO Journal* 31, 2182–2194.
- Keck, J.L., Roche, D.D., Lynch, A.S., Berger, J.M., 2000. Structure of the RNA polymerase domain of *E. coli* primase. *Science* 287, 2482–2486.
- Keen, B.A., Jozwiakowski, S.K., Bailey, L.J., Bianchi, J., Doherty, A.J., 2014. Molecular dissection of the domain architecture and catalytic activities of human PrimPol. *Nucleic Acids Research* 42, 5830–5845.
- Kelch, B.A., Makino, D.L., O'Donnell, M., Kuriyan, J., 2011. How a DNA polymerase clamp loader opens a sliding clamp. *Science* 334, 1675–1680.
- Kelch, B.A., Makino, D.L., O'Donnell, M., Kuriyan, J., 2012. Clamp loader ATPases and the evolution of DNA replication machinery. *BMC Biology* 10 (34), doi:10.1186/1741-7007-10-34.
- Kilkenny, M.L., Longo, M.A., Perera, R.L., Pellegrini, L., 2013. Structures of human primase reveal design of nucleotide elongation site and mode of Pol alpha tethering. *Proceedings of the National Academy of Sciences of the United States of America* 110, 15961–15966.
- Kim, C., Paulus, B.F., Wold, M.S., 1994. Interactions of human replication protein A with oligonucleotides. *Biochemistry* 33, 14197–14206.
- Kim, S., Dallmann, H.G., McHenry, C.S., Marians, K.J., 1996. Coupling of a replicative polymerase and helicase: A tau-DnaB interaction mediates rapid replication fork movement. *Cell* 84, 643–650.
- Kirk, B.W., Kuchta, R.D., 1999. Arg304 of human DNA primase is a key contributor to catalysis and NTP binding: Primase and the family X polymerases share significant sequence homology. *Biochemistry* 38, 7727–7736.

- Klinge, S., Hirst, J., Maman, J.D., Krude, T., Pellegrini, L., 2007. An iron-sulfur domain of the eukaryotic primase is essential for RNA primer synthesis. *Nature Structural & Molecular Biology* 14, 875–877.
- Kolpashchikov, D.M., Khodyreva, S.N., Khlimankov, D.Y., *et al.*, 2001. Polarity of human replication protein A binding to DNA. *Nucleic Acids Research* 29, 373–379.
- Kong, X.P., Onrust, R., O'Donnell, M., Kuriyan, J., 1992. Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: A sliding DNA clamp. *Cell* 69, 425–437.
- Kornberg, A., Baker, T.A., 1992. *DNA Replication*, second ed. New York, NY: W. H. Freeman.
- Krishna, T.S., Kong, X.P., Gary, S., Burgers, P.M., Kuriyan, J., 1994. Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* 79, 1233–1243.
- Kruger, K., Grabowski, P.J., Zaug, A.J., *et al.*, 1982. Self-splicing RNA: Autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* 31, 147–157.
- Kuchta, R.D., Stengel, G., 2010. Mechanism and evolution of DNA primases. *Biochimica et Biophysica Acta* 1804, 1180–1189.
- Kulczyk, A.W., Akabayov, B., Lee, S.J., *et al.*, 2012. An interaction between DNA polymerase and helicase is essential for the high processivity of the bacteriophage T7 replisome. *Journal of Biological Chemistry* 287, 39050–39060.
- Kumaran, S., Kozlov, A.G., Lohman, T.M., 2006. *Saccharomyces cerevisiae* replication protein A binds to single-stranded DNA in multiple salt-dependent modes. *Biochemistry* 45, 11958–11973.
- Kunkel, T.A., 2011. Balancing eukaryotic replication asymmetry with replication fidelity. *Current Opinion in Chemical Biology* 15, 620–626.
- Kunkel, T.A., Burgers, P.M., 2008. Dividing the workload at a eukaryotic replication fork. *Trends in Cell Biology* 18, 521–527.
- Kurth, I., Georgescu, R.E., O'Donnell, M.E., 2013. A solution to release twisted DNA during chromosome replication by coupled DNA polymerases. *Nature* 496, 119–122.
- de Laat, W.L., Appeldoorn, E., Sugawara, K., *et al.*, 1998. DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes & Development* 12, 2598–2609.
- Labib, K., De Piccoli, G., 2011. Surviving chromosome replication: The many roles of the S-phase checkpoint pathway. *Philosophical Transactions of the Royal Society B: Biological Sciences* 366, 3554–3561.
- Lamers, M.H., Georgescu, R.E., Lee, S.G., O'Donnell, M., Kuriyan, J., 2006. Crystal structure of the catalytic alpha subunit of *E. coli* replicative DNA polymerase III. *Cell* 126, 881–892.
- Lamers, M.H., O'Donnell, M., 2008. A consensus view of DNA binding by the C family of replicative DNA polymerases. *Proceedings of the National Academy of Sciences of the United States of America* 105, 20565–20566.
- Langston, L.D., Indiani, C., O'Donnell, M., 2009. Whither the replisome: Emerging perspectives on the dynamic nature of the DNA replication machinery. *Cell Cycle* 8, 2686–2691.
- Langston, L.D., O'Donnell, M., 2008. DNA polymerase delta is highly processive with proliferating cell nuclear antigen and undergoes collision release upon completing DNA. *Journal of Biological Chemistry* 283, 29522–29531.
- Lao-Sirieix, S.H., Bell, S.D., 2004. The heterodimeric primase of the hyperthermophilic archaeon *Sulfolobus solfataricus* possesses DNA and RNA primase, polymerase and 3'-terminal nucleotidyl transferase activities. *Journal of Molecular Biology* 344, 1251–1263.
- LeBowitz, J.H., McMacken, R., 1986. The *Escherichia coli* dnaB replication protein is a DNA helicase. *Journal of Biological Chemistry* 261, 4738–4748.
- Lee, S.J., Syed, S., Enemark, E.J., *et al.*, 2014. Dynamic look at DNA unwinding by a replicative helicase. *Proceedings of the National Academy of Sciences of the United States of America* 111, E827–E835.
- Lehman, I.R., Bessman, M.J., Simms, E.S., Kornberg, A., 1958. Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from *Escherichia coli*. *Journal of Biological Chemistry* 233, 163–170.
- Leipe, D.D., Aravind, L., Koonin, E.V., 1999. Did DNA replication evolve twice independently? *Nucleic Acids Research* 27, 3389–3401.
- Li, J.J., Kelly, T.J., 1984. Simian virus 40 DNA replication *in vitro*. *Proceedings of the National Academy of Sciences of the United States of America* 81, 6973–6977.
- Li, Y., Korolev, S., Waksman, G., 1998. Crystal structures of open and closed forms of binary and ternary complexes of the large fragment of *Thermus aquaticus* DNA polymerase I: Structural basis for nucleotide incorporation. *EMBO Journal* 17, 7514–7525.
- Lia, G., Michel, B., Allemand, J.F., 2012. Polymerase exchange during Okazaki fragment synthesis observed in living cells. *Science* 335, 328–331.
- Liberti, S.E., Larrea, A.A., Kunkel, T.A., 2013. Exonuclease 1 preferentially repairs mismatches generated by DNA polymerase alpha. *DNA Repair (Amst)* 12, 92–96.
- Liu, L., Komori, K., Ishino, S., *et al.*, 2001. The archaeal DNA primase: Biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*. *Journal of Biological Chemistry* 276, 45484–45490.
- Luo, G., Wang, M., Konigsberg, W.H., Xie, X.S., 2007. Single-molecule and ensemble fluorescence assays for a functionally important conformational change in T7 DNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America* 104, 12610–12615.
- Lyubimov, A.Y., Costa, A., Bleichert, F., Botchan, M.R., Berger, J.M., 2012. ATP-dependent conformational dynamics underlie the functional asymmetry of the replicative helicase from a minimalist eukaryote. *Proceedings of the National Academy of Sciences of the United States of America* 109, 11999–12004.
- Maga, G., Hubscher, U., 2003. Proliferating cell nuclear antigen (PCNA): A dancer with many partners. *Journal of Cell Science* 116, 3051–3060.
- Majka, J., Burgers, P.M., 2004. The PCNA-RFC families of DNA clamps and clamp loaders. *Progress in Nucleic Acid Research and Molecular Biology* 78, 227–260.
- Makarova, K.S., Koonin, E.V., 2013. Archaeology of eukaryotic DNA replication. *Cold Spring Harbor Perspectives in Medicine* 3, a012963.
- Makarova, K.S., Koonin, E.V., Kelman, Z., 2012. The CMG (CDC45/RecJ, MCM, GINS) complex is a conserved component of the DNA replication system in all archaea and eukaryotes. *Biology Direct* 7, 7.
- Manosas, M., Spiering, M.M., Ding, F., Croquette, V., Benkovic, S.J., 2012. Collaborative coupling between polymerase and helicase for leading-strand synthesis. *Nucleic Acids Research* 40, 6187–6198.
- Matsunaga, F., Norais, C., Forterre, P., Myllykallio, H., 2003. Identification of short 'eukaryotic' Okazaki fragments synthesized from a prokaryotic replication origin. *EMBO Reports* 4, 154–158.
- McHenry, C.S., 2011. Bacterial replicases and related polymerases. *Current Opinion in Chemical Biology* 15, 587–594.
- McInerney, P., Johnson, A., Katz, F., O'Donnell, M., 2007. Characterization of a triple DNA polymerase replisome. *Molecular Cell* 27, 527–538.
- Miles, J., Formosa, T., 1992. Protein affinity chromatography with purified yeast DNA polymerase alpha detects proteins that bind to DNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America* 89, 1276–1280.
- Miyata, T., Suzuki, H., Oyama, T., *et al.*, 2005. Open clamp structure in the clamp-loading complex visualized by electron microscopic image analysis. *Proceedings of the National Academy of Sciences of the United States of America* 102, 13795–13800.
- Moarefi, I., Jeruzalmi, D., Turner, J., O'Donnell, M., Kuriyan, J., 2000. Crystal structure of the DNA polymerase processivity factor of T4 bacteriophage. *Journal of Molecular Biology* 296, 1215–1223.
- Mouron, S., Rodríguez-Acebes, S., Martínez-Jiménez, M.I., *et al.*, 2013. Repriming of DNA synthesis at stalled replication forks by human PrimPol. *Nature Structural & Molecular Biology* 20, 1383–1389.
- Moyer, S.E., Lewis, P.W., Botchan, M.R., 2006. Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proceedings of the National Academy of Sciences of the United States of America* 103, 10236–10241.
- Murzin, A.G., 1993. OB(oligonucleotide/oligosaccharide binding)-fold: Common structural and functional solution for non-homologous sequences. *EMBO Journal* 12, 861–867.
- Netz, D.J., Stith, C.M., Stumpf, M., *et al.*, 2011. Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. *Nature Chemical Biology* 8, 125–132.
- Neuwald, A.F., Aravind, L., Spouge, J.L., Koonin, E.V., 1999. AAA+ : A class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Research* 9, 27–43.
- Nossal, N.G., Makhov, A.M., Chastain 2nd, P.D., Jones, C.E., Griffith, J.D., 2007. Architecture of the bacteriophage T4 replication complex revealed with nanoscale biointerfaces. *Journal of Biological Chemistry* 282, 1098–1108.
- O'Donnell, M., Jeruzalmi, D., Kuriyan, J., 2001. Clamp loader structure predicts the architecture of DNA polymerase III holoenzyme and RFC. *Current Biology* 11, R935–R946.
- O'Donnell, M., Kuriyan, J., 2006. Clamp loaders and replication initiation. *Current Opinion in Structural Biology* 16, 35–41.
- O'Donnell, M., Onrust, R., Dean, F.B., Chen, M., Hurwitz, J., 1993. Homology in accessory proteins of replicative polymerases — *E. coli* to humans. *Nucleic Acids Research* 21, 1–3.
- O'Donnell, M.E., 1987. Accessory proteins bind a primed template and mediate rapid cycling of DNA polymerase III holoenzyme from *Escherichia coli*. *Journal of Biological Chemistry* 262, 16558–16565.

- Olson, E., Nievera, C.J., Klimovich, V., Fanning, E., Wu, X., 2006. RPA2 is a direct downstream target for ATR to regulate the S-phase checkpoint. *Journal of Biological Chemistry* 281, 39517–39533.
- Pan, M., Santangelo, T.J., Li, Z., Reeve, J.N., Kelman, Z., 2011. *Thermococcus kodakarensis* encodes three MCM homologs but only one is essential. *Nucleic Acids Research* 39, 9671–9680.
- Park, K., Debyser, Z., Tabor, S., Richardson, C.C., Griffith, J.D., 1998. Formation of a DNA loop at the replication fork generated by bacteriophage T7 replication proteins. *Journal of Biological Chemistry* 273, 5260–5270.
- Paschall, C.O., Thompson, J.A., Marzahn, M.R., *et al.*, 2011. The *Escherichia coli* clamp loader can actively pry open the beta-sliding clamp. *Journal of Biological Chemistry* 286, 42704–42714.
- Perera, R.L., Torella, R., Klinge, S., *et al.*, 2013. Mechanism for Priming DNA Synthesis by Yeast DNA Polymerase alpha. *Elife* 2, e00482.
- Podobnik, M., McNerney, P., O'Donnell, M., Kuriyan, J., 2000. A TOPRIM domain in the crystal structure of the catalytic core of *Escherichia coli* primase confirms a structural link to DNA topoisomerases. *Journal of Molecular Biology* 300, 353–362.
- Pursell, Z.F., Isoz, I., Lundstrom, E.B., Johansson, E., Kunkel, T.A., 2007. Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* 317, 127–130.
- Pursell, Z.F., Kunkel, T.A., 2008. DNA polymerase epsilon: A polymerase of unusual size (and complexity). *Progress in Nucleic Acid Research and Molecular Biology* 82, 101–145.
- Raghuathan, S., Kozlov, A.G., Lohman, T.M., Waksman, G., 2000. Structure of the DNA binding domain of *E. coli* SSB bound to ssDNA. *Nature Structural & Molecular Biology* 7, 648–652.
- Raghuraman, M.K., Winzler, E.A., Collingwood, D., *et al.*, 2001. Replication dynamics of the yeast genome. *Science* 294, 115–121.
- Reyes-Lamothe, R., Sherratt, D.J., Leake, M.C., 2010. Stoichiometry and architecture of active DNA replication machinery in *Escherichia coli*. *Science* 328, 498–501.
- Rowen, L., Kornberg, A., 1978. Primase, the dnaG protein of *Escherichia coli*. An enzyme which starts DNA chains. *Journal of Biological Chemistry* 253, 758–764.
- Roy, R., Kozlov, A.G., Lohman, T.M., Ha, T., 2007. Dynamic structural rearrangements between DNA binding modes of *E. coli* SSB protein. *Journal of Molecular Biology* 369, 1244–1257.
- Rudd, S.G., Glover, L., Jozwiakowski, S.K., Horn, D., Doherty, A.J., 2013. PPL2 translesion polymerase is essential for the completion of chromosomal DNA replication in the African trypanosome. *Molecular Cell* 52, 554–565.
- Sakakibara, N., Kelman, L.M., Kelman, Z., 2009. Unwinding the structure and function of the archaeal MCM helicase. *Molecular Microbiology* 72, 286–296.
- Sanchez Garcia, J., Ciufo, L.F., Yang, X., Kearsley, S.E., MacNeill, S.A., 2004. The C-terminal zinc finger of the catalytic subunit of DNA polymerase delta is responsible for direct interaction with the B-subunit. *Nucleic Acids Research* 32, 3005–3016.
- Santoso, Y., Joyce, C.M., Potapova, O., *et al.*, 2010. Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proceedings of the National Academy of Sciences of the United States of America* 107, 715–720.
- Sawaya, M.R., Guo, S., Tabor, S., Richardson, C.C., Ellenberger, T., 1999. Crystal structure of the helicase domain from the replicative helicase-primase of bacteriophage T7. *Cell* 99, 167–177.
- Segurado, M., Diffley, J.F., 2008. Separate roles for the DNA damage checkpoint protein kinases in stabilizing DNA replication forks. *Genes & Development* 22, 1816–1827.
- Sekedat, M.D., Fenyo, D., Rogers, R.S., *et al.*, 2010. GINS motion reveals replication fork progression is remarkably uniform throughout the yeast genome. *Molecular Systems Biology* 6, 353.
- Sengupta, S., van Deursen, F., de Piccoli, G., Labib, K., 2013. Dpb2 integrates the leading-strand DNA polymerase into the eukaryotic replisome. *Current Biology* 23, 543–552.
- Seybert, A., Scott, D.J., Scaife, S., Singleton, M.R., Wigley, D.B., 2002. Biochemical characterisation of the clamp/clamp loader proteins from the euryarchaeon *Archaeoglobus fulgidus*. *Nucleic Acids Research* 30, 4329–4338.
- Sheu, Y.J., Kinney, J.B., Lengronne, A., Pasero, P., Stillman, B., 2014. Domain within the helicase subunit Mcm4 integrates multiple kinase signals to control DNA replication initiation and fork progression. *Proceedings of the National Academy of Sciences of the United States of America* 111, E1899–E1908.
- Sheu, Y.J., Stillman, B., 2010. The Dbf4-Cdc7 kinase promotes S phase by alleviating an inhibitory activity in Mcm4. *Nature* 463, 113–117.
- Siddiqui, K., On, K.F., Diffley, J.F., 2013. Regulating DNA replication in eukarya. *Cold Spring Harbor Perspectives in Biology* 5.pii: a012930.
- Simon, A.C., Zhou, J.C., Perera, R.L., *et al.*, 2014. A Ctf4 trimer couples the CMG helicase to DNA polymerase alpha in the eukaryotic replisome. *Nature* 510, 293–297.
- Simonetta, K.R., Kazmirski, S.L., Goedken, E.R., *et al.*, 2009. The mechanism of ATP-dependent primer-template recognition by a clamp loader complex. *Cell* 137, 659–671.
- Singleton, M.R., Dillingham, M.S., Wigley, D.B., 2007. Structure and mechanism of helicases and nucleic acid translocases. *Annual Review of Biochemistry* 76, 23–50.
- Singleton, M.R., Sawaya, M.R., Ellenberger, T., Wigley, D.B., 2000. Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* 101, 589–600.
- Sinha, N.K., Morris, C.F., Alberts, B.M., 1980. Efficient *in vitro* replication of double-stranded DNA templates by a purified T4 bacteriophage replication system. *Journal of Biological Chemistry* 255, 4290–4293.
- Sowd, G.A., Fanning, E., 2012. A wolf in sheep's clothing: SV40 co-opts host genome maintenance proteins to replicate viral DNA. *PLoS Pathogens* 8, e1002994.
- Stano, N.M., Chen, J., McHenry, C.S., 2006. A coproofreading Zn(2+) -dependent exonuclease within a bacterial replicase. *Nature Structural & Molecular Biology* 13, 458–459.
- Steitz, T.A., 1999. DNA polymerases: Structural diversity and common mechanisms. *Journal of Biological Chemistry* 274, 17395–17398.
- Steitz, T.A., 2006. Visualizing polynucleotide polymerase machines at work. *EMBO Journal* 25, 3458–3468.
- Stillman, B., 2008. DNA polymerases at the replication fork in eukaryotes. *Molecular Cell* 30, 259–260.
- Stith, C.M., Sterling, J., Resnick, M.A., Gordenin, D.A., Burgers, P.M., 2008. Flexibility of eukaryotic Okazaki fragment maturation through regulated strand displacement synthesis. *Journal of Biological Chemistry* 283, 34129–34140.
- Studwell, P.S., Stukenberg, P.T., Onrust, R., Skangalis, M., O'Donnell, M., 1989. Replication of the lagging strand by DNA polymerase III holoenzyme. In: Richardson, C.C., Lehman, I.R. (Eds.), *Molecular Mechanisms in DNA Replication and Recombination*, vol. 127. Hoboken, NJ: John Wiley and Sons, Inc, pp. 153–164.
- Stukenberg, P.T., Studwell-Vaughan, P.S., O'Donnell, M., 1991. Mechanism of the sliding beta-clamp of DNA polymerase III holoenzyme. *Journal of Biological Chemistry* 266, 11328–11334.
- Stukenberg, P.T., Turner, J., O'Donnell, M., 1994. An explanation for lagging strand replication: Polymerase hopping among DNA sliding clamps. *Cell* 78, 877–887.
- Toth, E.A., Li, Y., Sawaya, M.R., Cheng, Y., Ellenberger, T., 2003. The crystal structure of the bifunctional primase-helicase of bacteriophage T7. *Molecular Cell* 12, 1113–1123.
- Tsurimoto, T., Stillman, B., 1991. Replication factors required for SV40 DNA replication *in vitro*. II. Switching of DNA polymerase alpha and delta during initiation of leading and lagging strand synthesis. *Journal of Biological Chemistry* 266, 1961–1968.
- Turner, J., Hingorani, M.M., Kelman, Z., O'Donnell, M., 1999. The internal workings of a DNA polymerase clamp-loading machine. *EMBO Journal* 18, 771–783.
- Wan, L., Lou, J., Xia, Y., *et al.*, 2013. hPrip1/CCDC111 is a human DNA primase-polymerase required for the maintenance of genome integrity. *EMBO Reports* 14, 1104–1112.
- Wang, G., Klein, M.G., Tokonzaba, E., *et al.*, 2008. The structure of a DnaB-family replicative helicase and its interactions with primase. *Nature Structural & Molecular Biology* 15, 94–100.
- Warren, E.M., Huang, H., Fanning, E., Chazin, W.J., Eichman, B.F., 2009. Physical interactions between Mcm10, DNA, and DNA polymerase alpha. *Journal of Biological Chemistry* 284, 24662–24672.
- Watson, J.D., Crick, F.H., 1953a. Genetical implications of the structure of deoxyribonucleic acid. *Nature* 171, 964–967.
- Watson, J.D., Crick, F.H., 1953b. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* 171, 737–738.
- Weiner, B.E., Huang, H., Dattilo, B.M., *et al.*, 2007. An iron-sulfur cluster in the C-terminal domain of the p58 subunit of human DNA primase. *Journal of Biological Chemistry* 282, 33444–33451.
- Wing, R.A., Bailey, S., Steitz, T.A., 2008. Insights into the replisome from the structure of a ternary complex of the DNA polymerase III alpha-subunit. *Journal of Molecular Biology* 382, 859–869.
- Woese, C.R., 1968. *Genetic Code*, New Edition Edit. New York, NY: Joanna Cotler Books.

- Wold, M.S., 1997. Replication protein A: A heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annual Review of Biochemistry* 66, 61–92.
- Yabuki, N., Terashima, H., Kitada, K., 2002. Mapping of early firing origins on a replication profile of budding yeast. *Genes to Cells* 7, 781–789.
- Yang, J., Nelson, S.W., Benkovic, S.J., 2006. The control mechanism for lagging strand polymerase recycling during bacteriophage T4 DNA replication. *Molecular Cell* 21, 153–164.
- Yang, J., Zhuang, Z., Roccasecca, R.M., Trakselis, M.A., Benkovic, S.J., 2004. The dynamic processivity of the T4 DNA polymerase during replication. *Proceedings of the National Academy of Sciences of the United States of America* 101, 8289–8294.
- Yang, W., 2014. An overview of Y-family DNA polymerases and a case study of human DNA polymerase η . *Biochemistry* 53, 2793–2803.
- Yang, W., Woodgate, R., 2007. What a difference a decade makes: Insights into translesion DNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 104, 15591–15598.
- Yao, N.Y., Georgescu, R.E., Finkelstein, J., O'Donnell, M.E., 2009. Single-molecule analysis reveals that the lagging strand increases replisome processivity but slows replication fork progression. *Proceedings of the National Academy of Sciences of the United States of America* 106, 13236–13241.
- Yao, N.Y., O'Donnell, M., 2009. Replisome structure and conformational dynamics underlie fork progression past obstacles. *Current Opinion in Cell Biology* 21, 336–343.
- Zerbe, L.K., Kuchta, R.D., 2002. The p58 subunit of human DNA primase is important for primer initiation, elongation, and counting. *Biochemistry* 41, 4891–4900.
- Zhu, H., Shuman, S., 2005. A primer-dependent polymerase function of *Pseudomonas aeruginosa* ATP-dependent DNA ligase (LigD). *Journal of Biological Chemistry* 280, 418–427.
- Zhuang, Z., Johnson, R.E., Haracska, L., *et al.*, 2008. Regulation of polymerase exchange between Pol ϵ and Pol δ by monoubiquitination of PCNA and the movement of DNA polymerase holoenzyme. *Proceedings of the National Academy of Sciences of the United States of America* 105, 5361–5366.
- Zhuang, Z., Yoder, B.L., Burgers, P.M., Benkovic, S.J., 2006. The structure of a ring-opened proliferating cell nuclear antigen-replication factor C complex revealed by fluorescence energy transfer. *Proceedings of the National Academy of Sciences of the United States of America* 103, 2546–2551.
- Zlotkin, T., Kaufmann, G., Jiang, Y., *et al.*, 1996. DNA polymerase epsilon may be dispensable for SV40- but not cellular-DNA replication. *EMBO Journal* 15, 2298–2305.

NUCLEIC ACID SYNTHESIS/BREAKDOWN: DNA SYNTHESIS/REPAIR

Contents

Telomeres and Telomerase

Telomere Biology

Eukaryotic Nucleotide Excision Repair

The Base Excision Repair Pathway

Nonhomologous DNA End Joining

DNA Repair by Homologous Recombination

Telomeres and Telomerase

JJ-L Chen and JD Podlevsky, Arizona State University, Tempe, AZ, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Exonuclease An enzyme that catalyzes the hydrolysis of nucleic acids from the ends of the molecule.

Heterodimer A molecule composed of two different, simpler molecules.

Homodimer A molecule composed of two identical, simpler molecules.

Nucleoprotein A substance composed of DNA and protein.

Polymerase An enzyme that catalyzes the formation of a long-chain DNA molecule by linking DNA nucleotides together.

Retrotransposon A segment of DNA which goes through an RNA intermediate prior to inserting DNA copies of itself into the genome of the same cell.

Reverse transcriptase (RT) A polymerase that catalyzes the formation of DNA from an RNA template in the process of reverse transcription.

Ribonucleoprotein A substance composed of RNA and protein.

Senescence When a cell is no longer capable of dividing but still alive and metabolically active.

Somatic Of the body and distinguished from the germ line or stem cells.

Telomeres

In the late 1930s, Hermann Muller discovered that the native ends of linear eukaryotic chromosomes are functionally distinguished from chromosome breaks generated by X-ray irradiation. Muller coined the term 'telomere,' a combination of the Greek words 'telo-' and '-mere' meaning end part, to designate the distinctive and protective terminal portion of the chromosome. Shortly thereafter, Barbara McClintock described similar findings in maize with dicentric chromosomes – chromosomes with two centromeres – which are extremely unstable due to the high frequency of chromosome breakage and fusion events (McClintock, 1941). During cell division, dicentric chromosomes often result in chromosomal bridges between daughter cells and subsequent chromosome breaks leading to additional events of fusion and dicentric chromosome formation. This breakage–fusion–bridge cycle can be stopped in embryonic cells where the broken chromosome ends are healed by *de novo* addition of protective telomeres.

The findings of Muller and McClintock set the foundation for telomere biology years before DNA was known as the genetic material of inheritance.

The End-Replication Problem

The inability of conventional DNA polymerases to fully replicate the ends of linear DNA was recognized as 'the end-replication problem' in the early 1970s (Watson, 1972; Olovnikov, 1973). DNA polymerases synthesize DNA in a 5' to 3' direction and require a free 3'-hydroxyl group for the catalysis of nucleotide addition. The 3'-hydroxyl group is initially provided by an RNA primer which is later degraded and in-filled with DNA. It was postulated that the end-most RNA primer cannot be in-filled with DNA, which in theory results in DNA products shorter than the parental DNA each time the cell divides (Figure 1). However, it was later understood that eukaryotic chromosome ends have 3'-overhangs of

The length and shortening rate of telomeres is viewed as a molecular clock that determines the replicative capacity of somatic cells. Human somatic cells continuously cultured outside of the body eventually senesce and lose viability. Early work by Leonard Hayflick showed that terminally differentiated somatic cells had a limited number of cell divisions possible, termed the 'Hayflick limit' (Hayflick, 1965). The progressive loss of telomeric DNA from chromosome ends eventually induces cell senescence and is the limiting factor for the number of cell divisions possible (d'Adda di Fagagna *et al.*, 2003).

Telomeres are nucleoprotein complexes composed of telomeric DNA that consists of a vast array of highly repetitive short DNA sequences. In the late 1970s, Elizabeth Blackburn sequenced telomeric DNA from the free-living unicellular eukaryote *Tetrahymena thermophila* that uniquely contains millions of copies of linear minichromosomes amplified during its vegetative growth phase. These amplified linear minichromosomes are appended with telomeric DNA at the ends, providing a rich sample source for sequencing telomeric DNA. Blackburn discovered that *Tetrahymena* telomeric DNA contains hundreds of copies of the repeated sequence 'TTGGGG,' later termed the 'G-strand' in reference to the high propensity of guanosine nucleotides (Blackburn and Gall, 1978). The complementary 'C-strand' was named for the high incidence of cytosine nucleotides. Human telomeric DNA contains the repeat sequence 'TTAGGG,' which is slightly different from *Tetrahymena* (Moyzis *et al.*, 1988; Figure 2). The telomeric DNA sequence TTAGGG is not unique to humans. Rather, it is found throughout all known vertebrate species as well as select species ranging from marine invertebrates to fungi to plants to protozoans (Meyne *et al.*, 1989; Podlevsky *et al.*, 2008). While the length of telomeric DNA in *Tetrahymena* is only hundreds of base-pairs, in humans it is 10 000–15 000 base-pairs with a 3' G-strand overhang of approximately 200 nucleotides in length (Makarov *et al.*, 1997). In mammals, the 3' G-strand overhang is protected by invading the double-stranded telomeric DNA, and forms a D-shaped loop structure, termed a T-loop (Griffith *et al.*, 1999). Telomeric DNA is bound by unique groups of proteins which distinguish telomere ends from damage-induced chromosome breaks.

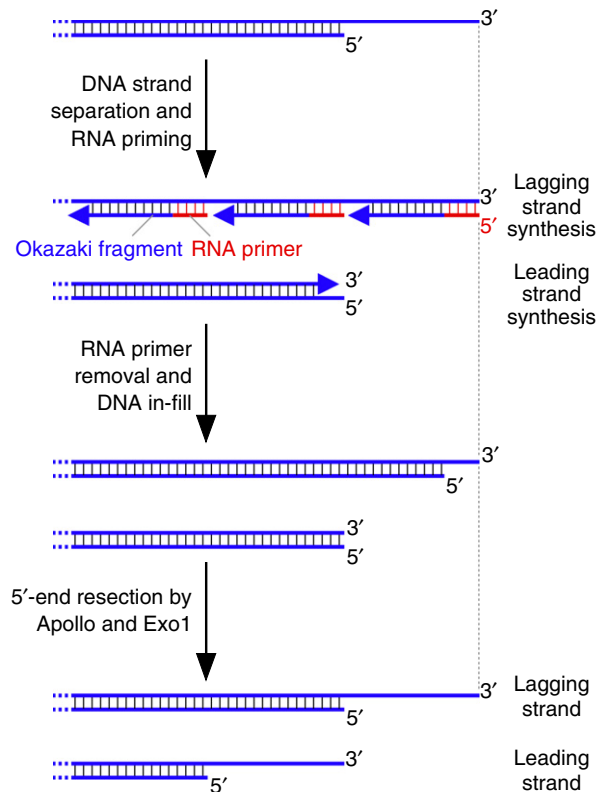


Figure 1 The end-replication problem results in DNA loss. Conventional DNA polymerases are unable to fully replicate the ends of linear eukaryotic chromosomes leading to shorter DNA products than the parental DNA strands. The lagging strand terminal-most RNA primer (red) cannot be in-filled with DNA resulting in a slightly shorter DNA strand (blue). In contrast, leading strand synthesis generates a blunt-end which is processed by exonucleases Apollo and Exo1 and results in a much shorter DNA product (blue).

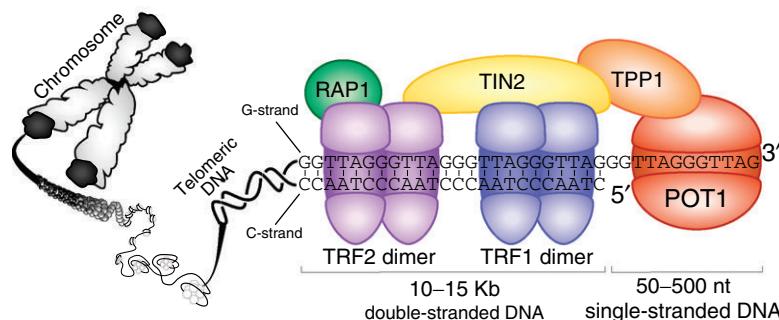


Figure 2 Mammalian telomeric DNA is coated by the Shelterin protein complex. The Shelterin complex is composed of three DNA binding proteins(TRF1 (blue), TF2 (violet), and POT1 (red)) and three associated proteins. RAP1 (green) binds selectively to TRF2 and not TRF1, while TIN2 (yellow) and TPP1 (orange) bridge the single-stranded and double-stranded regions of the telomere.

Telomeric Proteins

In mammals, telomeric DNA is coated by the Shelterin protein complex. The Shelterin complex is composed of six proteins: three directly bind to telomeric DNA and are in turn bound by the remaining three proteins (de Lange, 2005; Figure 2). Shelterin proteins are continuously and specifically associated with telomeric DNA throughout the cell cycle as opposed to numerous additional proteins which transiently bind to mammalian telomeric DNA. The first Shelterin proteins identified were aptly termed Telomeric Repeat-binding Factors 1 and 2 (TRF1 and TRF2, respectively). TRF1 was identified by purification of human cell extracts and later characterization revealed TRF1 bound telomeric DNA through a unique DNA binding 'Myb' sequence (Zhong *et al.*, 1992; Chong *et al.*, 1995). A search for this unique Myb sequence in the human proteome identified TRF2, a protein highly similar in sequence and structure to TRF1 (Bilaud *et al.*, 1997; Broccoli *et al.*, 1997). TRF1 and TRF2 bind double-stranded telomeric DNA as homodimers – protein complexes composed of two copies of the same protein (Figure 2). Despite the immense similarity between these two proteins, TRF1 and TRF2 cannot form heterodimers. TRF1 and TRF2 homodimers specifically recognize double-stranded telomeric DNA through the nine base-pair sequence 'TTAGGGTTA' (Choi *et al.*, 2011). The last Shelterin direct telomeric DNA binding protein, termed Protection Of Telomeres 1 (POT1) was identified as binding single-stranded telomeric DNA located in the 3' G-strand overhang and recognizes the 10 nucleotide sequence 'TTAGGGTTAG' (Baumann and Cech, 2001; Choi *et al.*, 2011; Figure 2).

Telomeric DNA binding by TRF1, TRF2, and POT1 is necessary for the remaining three Shelterin proteins to associate with telomeric DNA (Figure 2). Human Repressor/Activator Protein 1 (RAP1) binds specifically to TRF2 and not to its highly similar counterpart TRF1 (Li *et al.*, 2000). While RAP1 forms a complex with TRF2, this complex is not essential for TRF2 to bind telomeric DNA. Instead, RAP1 functions to prevent inappropriate interactions between neighboring telomeres (Martinez *et al.*, 2010). In the absence of RAP1, these inappropriate telomere interactions induce telomere fragility and breaks which increase the rate of telomere shortening. The last two Shelterin proteins function as molecular bridges, connecting the disparate and physically separate Shelterin protein components (Figure 2). TRF1- and TRF2-interacting nuclear protein 2 (TIN2) function as the central link between the numerous copies of TRF1 and TRF2 proteins which coat double-stranded telomeric DNA (Kim *et al.*, 1999). TIN2 and POT1-interacting protein 1 (TPP1), as the name implies, interacts mutually with POT1 and TIN2 (Houghtaling *et al.*, 2004; Liu *et al.*, 2004; Ye and de Lange, 2004). TPP1 connects single-stranded and double-stranded telomeric DNA together allowing for interaction and communication between these distinct regions of the telomere (Figure 2). The interactions between these unique Shelterin protein components are essential for the overall structure of the telomere and chromosome stability.

Yeast telomeres have two distinct telomeric protein complexes, one specific for double-stranded telomeric DNA and the other for the single-stranded region. The yeast double-stranded telomeric DNA binding complex is formed by the yeast variant of RAP1 and its binding partner proteins RAP1-

Interacting Factors 1 and 2 (RIF1 and RIF2, respectively) (Conrad *et al.*, 1990; Hardy *et al.*, 1992; Wotton and Shore, 1997). Unlike mammalian RAP1, the yeast RAP1 protein directly binds to double-stranded telomeric DNA. Yeast single-stranded telomeric DNA is bound by the three protein complex composed of Cdc13, Stn1, and Ten1, commonly referred to as the CST complex (Lin and Zakian, 1996; Nugent *et al.*, 1996; Grandin *et al.*, 1997, 2001). Stn1 and Ten1 bind to Cdc13 which directly binds single-stranded telomeric DNA. The CST complex is important for regulating the exonucleases responsible for 3' G-strand overhang length. The loss of the CST complex results in an exacerbated C-strand degradation, producing longer 3' G-strand overhangs. A human variant of the CST complex was found to associate with human telomeric DNA and function similarly for regulating the length of the 3' G-strand overhang (Miyake *et al.*, 2009). While the yeast CST complex represses C-strand degradation, the human variant promotes C-strand in-fill (Wang *et al.*, 2012). The human CST complex promotes the association of DNA polymerase alpha (pol α) to the telomere for the in-fill of the C-strand following G-strand extension by the telomerase enzyme (Figure 2).

Telomeric Repeat-Containing RNA

Telomeres were initially viewed as transcriptionally silent, packed into the higher-order heterochromatin structure. It was later discovered that the C-strand of the telomere is used as template for transcription of the long non-coding RNA, termed telomeric repeat-containing RNA (TERRA), which as the name suggests contains telomeric DNA repeat sequences. TERRA has been identified within distantly related species, ranging from vertebrates to yeasts to plants, and is actively transcribed from approximately 25% of the telomeres within a cell (Azzalin *et al.*, 2007). Human TERRA is highly heterogeneous, varying in length from 100 nucleotides to more than 9000 nucleotides, while yeast TERRA has a more consistent length of approximately 400 nucleotides. The levels of TERRA transcribed are seemingly dependent on telomere length, with shorter telomeres promoting TERRA transcription and longer telomeres repressing expression (Iglesias *et al.*, 2011; Arnoult *et al.*, 2012). Shorter telomeres have fewer telomere-binding proteins that are responsible for transcriptional silencing of the telomere. TERRA appears to aid in telomerase recruitment and the increased levels of TERRA at short telomeres increase the rate of telomere extension at these eroded telomere ends (Cusanelli *et al.*, 2013). Additional functions for TERRA have been proposed, which include preventing separate telomeres from interacting and supporting chromosome end stability.

Telomerase

The telomerase enzyme solves the end-replication problem by extending the 3'-end of linear chromosomes to offset telomere shortening. In the mid-1980s, Carol Greider and Elizabeth Blackburn identified telomerase from *Tetrahymena thermophila*, an enzyme with the unique ability to add telomeric DNA repeats onto the 3'-end of DNA (Greider and Blackburn, 1985). Telomerase functions as a ribonucleoprotein enzyme that

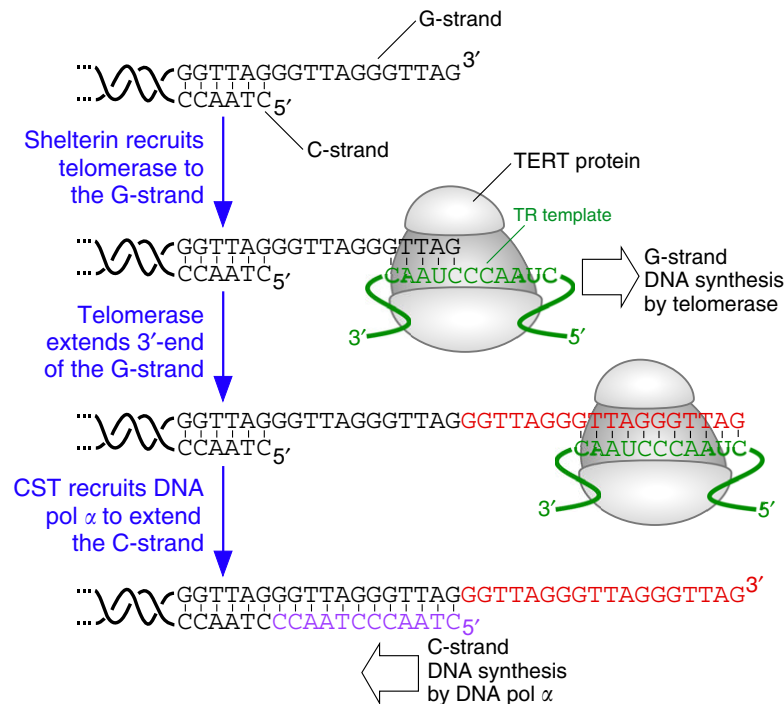


Figure 3 The telomerase enzyme extends the G-strand of telomeric DNA. The Shelterin complex recruits telomerase selectively to the ends of shortened telomeres. Telomerase contains the template for telomeric DNA synthesis located within the integral TR (green) component for reiteratively catalyzing six nucleotide telomeric DNA repeats (red) to extend the G-strand. The mammalian CST complex recruits DNA pol α for C-strand in-fill (violet).

contains the catalytic telomerase reverse transcriptase (TERT) protein component and the integral telomerase RNA (TR) component that provides the template for DNA synthesis (Greider and Blackburn, 1987, 1989). With an internal template, telomerase adds single-stranded telomeric DNA repeats to the 3'-end of the G-strand, bypassing the requirement of a parental DNA template (Figure 3). The *de novo* addition of DNA repeats onto the chromosome 3'-end offsets DNA loss (Shippen-Lentz and Blackburn, 1990).

Telomerase synthesizes telomeric DNA during the late S-phase of the cell cycle. The mammalian Shelterin complex proteins POT1 and TPP1 actively recruit telomerase to the ends of critically short telomeres and stimulate telomerase DNA repeat addition (Xin *et al.*, 2007; Wang *et al.*, 2007). Telomerase binds to the 3'-end of the telomere G-strand overhang and catalyzes the addition of six nucleotides, GGTTAG in humans, which form the basic unit of a telomeric DNA repeat (Figure 3). Telomerase is processive, capable of adding hundreds of telomeric repeats in reiterative cycles of DNA synthesis. While telomerase specifically extends the G-strand overhang, C-strand in-fill requires a separate DNA synthesis mechanism. In vertebrates, following G-strand extension by telomerase, the CST complex binds to the single-stranded telomeric DNA and recruits DNA pol α to synthesize the complementary C-strand DNA using the G-strand as template (Wang *et al.*, 2012). The cell orchestrates a precise balance between recruitment of telomerase for G-strand extension, DNA pol α for C-strand in-fill, and exonucleases Apollo and Exo1 for generating the G-strand overhang (Wu *et al.*, 2012).

Telomerase Reverse Transcriptase

The TERT protein is the catalytic component of the telomerase enzyme. This protein is responsible for the synthesis of telomeric DNA repeats from the RNA template located within TR. The TERT protein component is present across all known taxa – with the exception of a subgroup of select insects – and is composed of four independently folded protein segments or domains. The four domains of TERT include the telomerase essential N-terminal (TEN) domain, the telomerase RNA binding domain (TRBD), the reverse transcriptase (RT) domain, and the C-terminal extension (CTE) domain (Figure 4). The central RT domain and the CTE domain contain motifs highly conserved in other RTs and DNA polymerases, while the TEN and TRBD domains are unique to the TERT protein (Lingner *et al.*, 1997).

The overall structure of the RT and CTE domains, like all known DNA polymerases, resembles a right hand with finger, palm, and thumb subdomains (Nakamura *et al.*, 1997; Gillis *et al.*, 2008). The 'fingers' in right handed DNA polymerases bind incoming nucleotides, while the 'palm' constitutes the catalytic site for nucleotide polymerization (Bosoy and Lue, 2001). Like that of all other DNA polymerases, within the palm of the TERT protein lie three invariant aspartic acids (Lingner *et al.*, 1997). This aspartic acid triad coordinates the positioning of two magnesium atoms. This acid-metal chemistry for nucleotide addition is common to all DNA polymerases. The TERT CTE domain has a similar overall structure to viral RTs 'thumb' domains, yet functions distinctly within the TERT protein (Gillis

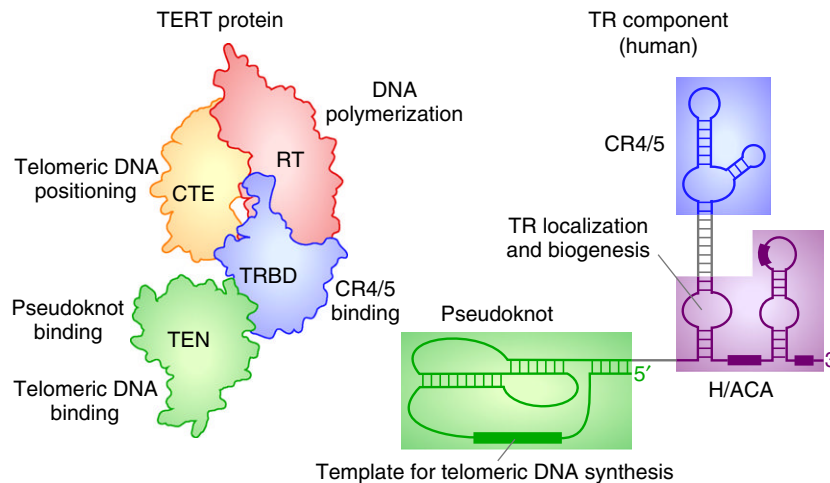


Figure 4 The core telomerase enzyme is minimally composed of the TERT protein and TR component. The TERT protein is composed of four distinct functional domains: TEN (green), TRBD (blue), RT (red), and CTE (orange). The human TR is composed of three distinct functional domains: Pseudoknot (green), CR4/5 (blue), and H/ACA (violet). TEN and TRBD bind the similarly colored Pseudoknot and CR4/5 domains for formation of the telomerase RNP.

et al., 2008). The RT thumb domain positions the RNA template base-paired with a DNA primer, while the TERT CTE domain binds the 3'-end of the telomeric DNA to enhance DNA polymerization (Hossain *et al.*, 2002).

The TEN and TRBD domain contains TR binding sites critical for ribonucleoprotein assembly (Moriarty *et al.*, 2002; Bley *et al.*, 2011; Huang *et al.*, 2014). In addition, the TEN domain contains 'anchor' sites for binding single-stranded telomeric DNA (Sealey *et al.*, 2010). This DNA anchor site is important for delaying the complete dissociation of the telomerase enzyme from the telomeric DNA, enhancing the number of telomeric DNA repeats synthesized onto the end of the chromosome (Jacobs *et al.*, 2006).

Telomerase RNA

TR is the integral RNA component of the telomerase enzyme. In addition to providing the template for telomeric DNA synthesis, TR is essential for the formation of a functional telomerase ribonucleoprotein enzyme. The structure of TR is a remarkably divergent and complex, containing essential elements for defining the template region from the remainder of the RNA as well as specific regions necessary for TR biogenesis and telomerase assembly. TR from vertebrates and most fungi contains two universal structural domains: the pseudoknot domain and the Conserved Region 4 and 5 (CR4/5) domain (Chen *et al.*, 2002, 2000; Qi *et al.*, 2013; Figure 4). These two domains bind the TERT protein independently and are essential for telomerase activity (Tesmer *et al.*, 1999).

The TR template region is properly positioned within the TERT catalytic site by the pseudoknot domain (Zhang *et al.*, 2011). The pseudoknot domain contains the forenamed pseudoknot structure wherein the loop from a stem-loop base-pairs with a nearby single-stranded region (Figure 4). The TR pseudoknot structure includes a triple helix, a third RNA strand base-pairing with a canonically Watson-Crick base-paired RNA helix (Theimer *et al.*, 2005; Shefer *et al.*, 2007). Immediately

upstream of the 5'-end of the TR template lies the template boundary element which defines and safeguards the end of the template region (Chen and Greider, 2003). The template boundary element restricts access of the template adjacent RNA to the TERT catalytic site, preventing the use of this region as template by telomerase and synthesis of non-telomeric DNA.

The CR4/5 domain, termed three-way-junction in budding yeasts, is located at a distance downstream from the pseudoknot domain and is essential for telomerase assembly and activity (Chen *et al.*, 2002; Brown *et al.*, 2007; Figure 4). The high-affinity binding between the CR4/5 helical structure and the TRBD is specific and essential for telomerase to function (Bley *et al.*, 2011; Huang *et al.*, 2014). While the CR4/5 domain is conserved in vertebrates and the vast majority of fungal TRs, ciliate TRs contain a stem-loop IV that may be functionally analogous to the vertebrate CR4/5 domain (Robart *et al.*, 2010).

In vertebrate TR, a third domain called H/ACA domain located at the 3' distal region is necessary for TR biogenesis and localization to the correct nuclear compartment. The H/ACA domain, named for the similarity to H/ACA small nucleolar (sno) and small Cajal body (sca) RNAs, is composed of a tandem array of stem-loops with the forenamed boxes H and ACA intervening (Mitchell *et al.*, 1999; Jády *et al.*, 2004). In common with H/ACA snoRNAs and scaRNAs, the vertebrate TR H/ACA domain is bound by two copies of the dyskerin complex (Egan and Collins, 2010). The dyskerin complex is a tetrad of proteins comprising dyskerin, NOP10, NHP2, and GAR1 (Girard *et al.*, 1993; Maiorano *et al.*, 1999; Pogacíc *et al.*, 2000; Cheng and Roberts, 2001; Hamma *et al.*, 2005). Additionally, the 3'-most stem-loop in the human TR H/ACA domain is bound by telomerase Cajal body protein 1 (TCAB1) for localization to Cajal bodies, a nuclear compartment rich in RNA splicing and post-transcriptional modification machinery, prior to TR assembly with the TERT protein (Venteicher *et al.*, 2009).

The TR demonstrates astonishing divergence in length, sequence, and the folded RNA structure (Podlevsky *et al.*, 2008). Ciliates encompass the smallest and most compact TRs discovered to date, ranging from 147 to 205 nucleotides in length

(McCormick-Graham and Romero, 1995). In contrast, yeast and filamentous fungi have vastly larger TRs which range from 928 to 2425 nucleotides (Dandjinou *et al.*, 2004; Qi *et al.*, 2013). Vertebrate TRs are more modest in length, ranging from 312 to 559 nucleotides (Chen *et al.*, 2000; Xie *et al.*, 2008). The tremendous difference in TR length accommodates specific and distinct groups of TR binding proteins. While vertebrate TRs are bound by the dyskerin complex, yeast TRs are bound by Ever Shorter Telomeres 1 (EST1), the Ku heterodimer, and the Sm protein complex (Lendvay *et al.*, 1996; Peterson *et al.*, 2001), while ciliate TRs are bound by p50 and p65 proteins (Witkin and Collins, 2004; Min and Collins, 2009). Additionally, ciliate TRs are transcribed by RNA pol III and contain a terminal poly-U tract, while vertebrate and yeast TRs are transcribed by RNA pol II (McCormick-Graham and Romero, 1995; Mitchell *et al.*, 1999; Chapon *et al.*, 1997; Chen *et al.*, 2000). The different TR binding protein partners and even the RNA polymerase employed for transcription demonstrate the innate flexibility of TR biogenesis amongst distinct groups of species.

Syndromes of Short Telomeres

Telomere dysfunction leads to a myriad of human diseases and cellular ageing. Sufficient telomere length is essential for chromosome stability and permitting continuous cellular proliferation. Telomeres of insufficient length are unable to safeguard the ends of the chromosome from recognition by DNA damage repair, permitting chromosome fusions. Additionally, telomere length functions as a checkpoint for cell growth and division (d'Adda di Fagagna *et al.*, 2003). While the average length of telomeric DNA within humans is 10 000–15 000 base-pairs, as little as 1000 base-pairs is sufficient for chromosome end protection (Damm *et al.*, 2001). Since telomeres shorten with each cell division, the length of the telomere is a predictor of the number of cell divisions possible. This has earned telomeres the moniker 'mitotic clock,' since the telomeres countdown to senescence.

Numerous mutations within six of the genes that encode for the Shelterin complex and the telomerase holoenzyme are linked with telomere shortening and result in several human diseases. These syndromes of short telomeres include dyskeratosis congenita, aplastic anemia, and idiopathic pulmonary fibrosis (Armanios, 2009). Dyskeratosis congenita presents as a combination of skin and nail disorders, aplastic anemia as low peripheral blood cell counts, and idiopathic pulmonary fibrosis as the accumulation of fibrous material in the lungs (Fogarty *et al.*, 2003; Vulliamy *et al.*, 2006; Armanios *et al.*, 2007). Despite the different clinical presentations, all these diseases arise from the inability to maintain cell populations necessary for organ and tissue renewal. Telomeres shrink with human development and ageing; thus ageing itself can be considered a telomere-mediated disease (Harley *et al.*, 1990). Reciprocally, the vast majority of cancers appropriate telomerase for maintaining their continuous, deleterious growth (Kim *et al.*, 1994).

Alternative Lengthening of Telomeres

While telomerase is the prominent solution to the end-replication problem within eukaryotes and the vast majority of

cancers, there are telomerase-independent means of telomere length maintenance employed by specific cancers (Bryan *et al.*, 1995). These telomerase-independent pathways have been termed Alternative Lengthening of Telomeres (ALT). ALT employs telomere homologous recombination in which one telomere strand invades the other, using the other telomere as the template for elongation (Londoño-Vallejo *et al.*, 2004). Cells which employ ALT for telomere elongation have increased heterogeneity in telomere length and the presence of circular extra-chromosomal telomeric DNA repeats (Bryan *et al.*, 1997; Cesare and Griffith, 2004).

Flies from the dipteran insect lineage lack telomerase for telomere length maintenance (Biessmann and Mason, 1997; Pardue *et al.*, 1997). Specific *Drosophila* species telomeres are composed of tandem arrays of HeT-A and TART retrotransposons – parasitic genetic elements capable of self-replication through an RNA intermediate (Biessmann *et al.*, 1990). The appropriation of these retrotransposons is limited to only a select group of species even within the *Drosophila* genus. In most dipteran fly species, telomeres are instead composed of satellite sequences – repetitive DNA sequences of 50–800 base-pairs and employ a form of recombination similar to ALT for satellite sequence amplification for the extension and maintenance of these telomere lengths (Nielsen and Edström, 1993; Roth *et al.*, 1997; Biessmann *et al.*, 2000). The lack of telomerase within such a small and closely related group of species implies that telomerase was lost recently within this insect lineage.

The evolution of linear eukaryotic chromosomes set-in-motion the need for ancillary mechanisms to overcome the end-replication problem imposed by the limitations of conventional DNA polymerases. Additionally, a means of protection to delineate chromosome ends from damaged DNA was necessary. The solution was to cap the ends of chromosomes with a vast array of short telomeric DNA repeats bound by unique telomeric protein complexes, the Shelterin complex in humans. Telomere shortening following each cell division was virtually universally offset by the unique telomerase enzyme.

See also: Nucleic Acid Synthesis/Breakdown: DNA Synthesis/Repair: DNA Repair by Homologous Recombination. Nucleic Acid Synthesis/Breakdown: RNA Synthesis/Function: Comparison of Bacterial and Eukaryotic Replisome Components. Nucleic Acid Synthesis/Breakdown: Transcription: Eukaryotic Transcriptional Regulation

References

- d'Adda di Fagagna, F., Reaper, P.M., Clay-Farrace, L., *et al.*, 2003. A DNA damage checkpoint response in telomere-initiated senescence. *Nature* 426, 194–198.
- Armanios, M.Y., 2009. Syndromes of telomere shortening. *Annual Review of Genomics and Human Genetics* 10, 45–61.
- Armanios, M.Y., Chen, J.J.-L., Cogan, J.D., *et al.*, 2007. Telomerase mutations in families with idiopathic pulmonary fibrosis. *New England Journal of Medicine* 356, 1317–1326.
- Arnoult, N., van Beneden, A., Decottignies, A., 2012. Telomere length regulates TERRA levels through increased trimethylation of telomeric H3K9 and HP1 α . *Nature Structural & Molecular Biology* 19, 948–956.
- Azzalin, C.M., Reichenbach, P., Khoraiuli, L., Giulotto, E., Lingner, J., 2007. Telomeric repeat containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* 318, 798–801.

- Baumann, P., Cech, T.R., 2001. Pot1, the putative telomere end-binding protein in fission yeast and humans. *Science* 292, 1171–1175.
- Biessmann, H., Carter, S.B., Mason, J.M., 1990. Chromosome ends in *Drosophila* without telomeric DNA sequences. *Proceedings of the National Academy of Sciences of the United States of America* 87, 1758–1761.
- Biessmann, H., Mason, J.M., 1997. Telomere maintenance without telomerase. *Chromosoma* 106, 63–69.
- Biessmann, H., Zurovcova, M., Yao, J.G., Lozovskaya, E., Walter, M.F., 2000. A telomeric satellite in *Drosophila virilis* and its sibling species. *Chromosoma* 109, 372–380.
- Bilaud, T., Brun, C., Ancelin, K., *et al.*, 1997. Telomeric localization of TRF2, a novel human telobox protein. *Nature Genetics* 17, 236–239.
- Blackburn, E.H., Gall, J.G., 1978. A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena. *Journal of Molecular Biology* 120, 33–53.
- Bley, C.J., Qi, X., Rand, D.P., *et al.*, 2011. RNA-protein binding interface in the telomerase ribonucleoprotein. *Proceedings of the National Academy of Sciences of the United States of America* 108, 20333–20338.
- Bosoy, D., Lue, N.F., 2001. Functional analysis of conserved residues in the putative “finger” domain of telomerase reverse transcriptase. *Journal of Biological Chemistry* 276, 46305–46312.
- Broccoli, D., Smogorzewska, A., Chong, L., de Lange, T., 1997. Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nature Genetics* 17, 231–235.
- Brown, Y., Abraham, M., Pearl, S., *et al.*, 2007. A critical three-way junction is conserved in budding yeast and vertebrate telomerase RNAs. *Nucleic Acids Research* 35, 6280–6289.
- Bryan, T.M., Englezou, A., Dalla-Pozza, L., Dunham, M.A., Reddel, R.R., 1997. Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nature Medicine* 3, 1271–1274.
- Bryan, T.M., Englezou, A., Gupta, J., Bacchetti, S., Reddel, R.R., 1995. Telomere elongation in immortal human cells without detectable telomerase activity. *EMBO Journal* 14, 4240–4248.
- Cesare, A.J., Griffith, J.D., 2004. Telomeric DNA in ALT cells is characterized by free telomeric circles and heterogeneous t-loops. *Molecular and Cellular Biology* 24, 9948–9957.
- Chapon, C., Cech, T.R., Zaug, A.J., 1997. Polyadenylation of telomerase RNA in budding yeast. *RNA* 3, 1337–1351.
- Chen, J. J.-L., Blasco, M.A., Greider, C.W., 2000. Secondary structure of vertebrate telomerase RNA. *Cell* 100, 503–514.
- Chen, J. J.-L., Greider, C.W., 2003. Template boundary definition in mammalian telomerase. *Genes & Development* 17, 2747–2752.
- Chen, J. J.-L., Opperman, K.K., Greider, C.W., 2002. A critical stem-loop structure in the CR4-CR5 domain of mammalian telomerase RNA. *Nucleic Acids Research* 30, 592–597.
- Cheng, X., Roberts, R.J., 2001. AdoMet-dependent methylation, DNA methyltransferases and base flipping. *Nucleic Acids Research* 29, 3784–3795.
- Choi, K.H., Farrell, A.S., Lakamp, A.S., Ouellette, M.M., 2011. Characterization of the DNA binding specificity of Shelterin complexes. *Nucleic Acids Research* 39, 9206–9223.
- Chong, L., Van Steensel, B., Broccoli, D., *et al.*, 1995. A human telomeric protein. *Science* 270, 1663–1667.
- Conrad, M.N., Wright, J.H., Wolf, A.J., Zakian, V.A., 1990. RAP1 protein interacts with yeast telomeres *in vivo*: Overproduction alters telomere structure and decreases chromosome stability. *Cell* 63, 739–750.
- Cusanelli, E., Romero, C.A.P., Chartrand, P., 2013. Telomeric noncoding RNA TERRA is induced by telomere shortening to nucleate telomerase molecules at short telomeres. *Molecular Cell* 51, 780–791.
- Damm, K., Hemmann, U., Garin-Chesa, P., *et al.*, 2001. A highly selective telomerase inhibitor limiting human cancer cell proliferation. *EMBO Journal* 20, 6958–6968.
- Dandjinou, A.T., Lévesque, N., Larose, S., *et al.*, 2004. A phylogenetically based secondary structure for the yeast telomerase RNA. *Current Biology* 14, 1148–1158.
- Egan, E.D., Collins, K.L., 2010. Specificity and stoichiometry of subunit interactions in the human telomerase holoenzyme assembled *in vivo*. *Molecular and Cellular Biology* 30, 2775–2786.
- Fogarty, P.F., Yamaguchi, H., Wiestner, A., *et al.*, 2003. Late presentation of dyskeratosis congenita as apparently acquired aplastic anaemia due to mutations in telomerase RNA. *Lancet* 362, 1628–1630.
- Gillis, A.J., Schuller, A.P., Skordalakes, E., 2008. Structure of the *Tribolium castaneum* telomerase catalytic subunit TERT. *Nature* 455, 633–637.
- Girard, J.P., Caizergues-Ferrer, M., Lapeyre, B., 1993. The SpGAR1 gene of *Schizosaccharomyces pombe* encodes the functional homologue of the snoRNP protein GAR1 of *Saccharomyces cerevisiae*. *Nucleic Acids Research* 21, 2149–2155.
- Grandin, N., Damon, C., Charbonneau, M., 2001. Ten1 functions in telomere end protection and length regulation in association with Stn1 and Cdc13. *EMBO Journal* 20, 1173–1183.
- Grandin, N., Reed, S.I., Charbonneau, M., 1997. Stn1, a new *Saccharomyces cerevisiae* protein, is implicated in telomere size regulation in association with Cdc13. *Genes & Development* 11, 512–527.
- Greider, C.W., Blackburn, E.H., 1985. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell* 43, 405–413.
- Greider, C.W., Blackburn, E.H., 1987. The telomere terminal transferase of Tetrahymena is a ribonucleoprotein enzyme with two kinds of primer specificity. *Cell* 51, 887–898.
- Greider, C.W., Blackburn, E.H., 1989. A telomeric sequence in the RNA of Tetrahymena telomerase required for telomere repeat synthesis. *Nature* 337, 331–337.
- Griffith, J.D., Comeau, L., Rosenfield, S., *et al.*, 1999. Mammalian telomeres end in a large duplex loop. *Cell* 97, 503–514.
- Hamma, T., Reichow, S.L., Varani, G., Ferré-D’amaré, A.R., 2005. The Cbf5-Nop10 complex is a molecular bracket that organizes box H/ACA RNPs. *Nature Structural & Molecular Biology* 12, 1101–1107.
- Hardy, C.F., Sussel, L., Shore, D., 1992. A RAP1-interacting protein involved in transcriptional silencing and telomere length regulation. *Genes & Development* 6, 801–814.
- Harley, C.B., Futcher, A.B., Greider, C.W., 1990. Telomeres shorten during ageing of human fibroblasts. *Nature* 345, 458–460.
- Hayflick, L., 1965. The limited *in vitro* lifetime of human diploid cell strains. *Experimental Cell Research* 37, 614–636.
- Hossain, S., Singh, S., Lue, N.F., 2002. Functional analysis of the C-terminal extension of telomerase reverse transcriptase. A putative “thumb” domain. *Journal of Biological Chemistry* 277, 36174–36180.
- Houghtaling, B.R., Cuttonaro, L., Chang, W., Smith, S., 2004. A dynamic molecular link between the telomere length regulator TRF1 and the chromosome end protector TRF2. *Current Biology: CB* 14, 1621–1631.
- Huang, J., Brown, A.F., Wu, J., *et al.*, 2014. Structural basis for protein-RNA recognition in telomerase. *Nature Structural & Molecular Biology* 21, 507–512.
- Huffman, K.E., Levene, S.D., Tesmer, V.M., Shay, J.W., Wright, W.E., 2000. Telomere shortening is proportional to the size of the G-rich telomeric 3'-overhang. *Journal of Biological Chemistry* 275, 19719–19722.
- Iglesias, N., Redon, S., Pfeiffer, V., *et al.*, 2011. Subtelomeric repetitive elements determine TERRA regulation by Rap1/Rif and Rap1/Sir complexes in yeast. *EMBO Reports* 12, 587–593.
- Jacobs, S.A., Podell, E.R., Cech, T.R., 2006. Crystal structure of the essential N-terminal domain of telomerase reverse transcriptase. *Nature Structural & Molecular Biology* 13, 218–225.
- Jády, B.E., Bertrand, E., Kiss, T., 2004. Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *Journal of Cell Biology* 164, 647–652.
- Kim, N.W., Piatyszek, M.A., Prowse, K.R., *et al.*, 1994. Specific association of human telomerase activity with immortal cells and cancer. *Science* 266, 2011–2015.
- Kim, S.H., Kaminker, P., Campisi, J., 1999. TIN2, a new regulator of telomere length in human cells. *Nature Genetics* 23, 405–412.
- de Lange, T., 2005. Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes & Development* 19, 2100–2110.
- Lendvay, T.S., Morris, D.K., Sah, J., Balasubramanian, B., Lundblad, V., 1996. Senescence mutants of *Saccharomyces cerevisiae* with a defect in telomere replication identify three additional EST genes. *Genetics* 144, 1399–1412.
- Li, B., Oestreich, S., de Lange, T., 2000. Identification of human Rap1: Implications for telomere evolution. *Cell* 101, 471–483.
- Lin, J.J., Zakian, V.A., 1996. The *Saccharomyces cerevisiae* CDC13 protein is a single-strand TG1-3 telomeric DNA-binding protein *in vitro* that affects telomere behavior *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America* 93, 13760–13765.
- Lingner, J., Hughes, T.R., Shevchenko, A., *et al.*, 1997. Reverse transcriptase motifs in the catalytic subunit of telomerase. *Science* 276, 561–567.
- Liu, D., Safari, A., O'Connor, M.S., *et al.*, 2004. PTOP interacts with POT1 and regulates its localization to telomeres. *Nature Cell Biology* 6, 673–680.
- Londoño-Vallejo, J.A., Der-Sarkissian, H., Cazes, L., Bacchetti, S., Reddel, R.R., 2004. Alternative lengthening of telomeres is characterized by high rates of telomeric exchange. *Cancer Research* 64, 2324–2327.
- Maiorano, D., Brimage, L.J., Leroy, D., Kearsey, S.E., 1999. Functional conservation and cell cycle localization of the Nhp2 core component of H + ACA

- snoRNPs in fission and budding yeasts. *Experimental Cell Research* 252, 165–174.
- Makarov, V.L., Hirose, Y., Langmore, J.P., 1997. Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell* 88, 657–666.
- Martinez, P., Thanasoula, M., Carlos, A.R., *et al.*, 2010. Mammalian Rap1 controls telomere function and gene expression through binding to telomeric and extratelomeric sites. *Nature Cell Biology* 12, 768–780.
- McClintock, B., 1941. The stability of broken ends of chromosomes in *Zea Mays*. *Genetics* 26, 234–282.
- McCormick-Graham, M., Romero, D.P., 1995. Ciliate telomerase RNA structural features. *Nucleic Acids Research* 23, 1091–1097.
- Meyne, J., Ratliff, R.L., Moyzis, R.K., 1989. Conservation of the human telomere sequence (TTAGGG)_n among vertebrates. *Proceedings of the National Academy of Sciences of the United States of America* 86, 7049–7053.
- Min, B., Collins, K.L., 2009. An RPA-related sequence-specific DNA-binding subunit of telomerase holoenzyme is required for elongation processivity and telomere maintenance. *Molecular Cell* 36, 609–619.
- Mitchell, J.R., Cheng, J., Collins, K.L., 1999. A box H/ACA small nucleolar RNA-like domain at the human telomerase RNA 3' end. *Molecular and Cellular Biology* 19, 567–576.
- Miyake, Y., Nakamura, M., Nabetani, A., *et al.*, 2009. RPA-like mammalian Ctc1-Stn1-Ten1 complex binds to single-stranded DNA and protects telomeres independently of the Pot1 pathway. *Molecular Cell* 36, 193–206.
- Moriarty, T.J., Huard, S., Dupuis, S., Autexier, C., 2002. Functional multimerization of human telomerase requires an RNA interaction domain in the N terminus of the catalytic subunit. *Molecular and Cellular Biology* 22, 1253–1265.
- Moyzis, R.K., Buckingham, J.M., Cram, L.S., *et al.*, 1988. A highly conserved repetitive DNA sequence, (TTAGGG)_n, present at the telomeres of human chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* 85, 6622–6626.
- Nakamura, T.M., Morin, G.B., Chapman, K.B., *et al.*, 1997. Telomerase catalytic subunit homologs from fission yeast and human. *Science* 277, 955–959.
- Nielsen, L., Edström, J.E., 1993. Complex telomere-associated repeat units in members of the genus *Chironomus* evolve from sequences similar to simple telomeric repeats. *Molecular and Cellular Biology* 13, 1583–1589.
- Nugent, C.I., Hughes, T.R., Lue, N.F., Lundblad, V., 1996. Cdc13p: A single-strand telomeric DNA-binding protein with a dual role in yeast telomere maintenance. *Science* 274, 249–252.
- Olovnikov, A.M., 1973. A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *Journal of Theoretical Biology* 41, 181–190.
- Pardue, M.L., Danilevskaya, O.N., Traverse, K.L., Lowenhaupt, K., 1997. Evolutionary links between telomeres and transposable elements. *Genetica* 100, 73–84.
- Peterson, S.E., Stellwagen, A.E., Diede, S.J., *et al.*, 2001. The function of a stem-loop in telomerase RNA is linked to the DNA repair protein Ku. *Nature Genetics* 27, 64–67.
- Podlevsky, J.D., Bley, C.J., Omana, R.V., Qi, X., Chen, J. J.-L., 2008. The telomerase database. *Nucleic Acids Research* 36, D339–D343.
- Pogacic, V., Dragon, F., Filipowicz, W., 2000. Human H/ACA small nucleolar RNPs and telomerase share evolutionarily conserved proteins NHP2 and NOP10. *Molecular and Cellular Biology* 20, 9028–9040.
- Qi, X., Li, Y., Honda, S., *et al.*, 2013. The common ancestral core of vertebrate and fungal telomerase RNAs. *Nucleic Acids Research* 41, 450–462.
- Robart, A.R., O'Connor, C.M., Collins, K.L., 2010. Ciliate telomerase RNA loop IV nucleotides promote hierarchical RNP assembly and holoenzyme stability. *RNA* 16, 563–571.
- Roth, C.W., Kobeski, F., Walter, M.F., Biessmann, H., 1997. Chromosome end elongation by recombination in the mosquito *Anopheles gambiae*. *Molecular and Cellular Biology* 17, 5176–5183.
- Sealey, D.C.F., Zheng, L., Taboski, M.A.S., *et al.*, 2010. The N-terminus of hTERT contains a DNA-binding domain and is required for telomerase activity and cellular immortalization. *Nucleic Acids Research* 38, 2019–2035.
- Steir, A.J., Chai, W., Shay, J.W., Wright, W.E., 2005. Telomere-end processing the terminal nucleotides of human chromosomes. *Molecular Cell* 18, 131–138.
- Shefer, K., Brown, Y., Gorkovoy, V., *et al.*, 2007. A triple helix within a pseudoknot is a conserved and essential element of telomerase RNA. *Molecular and Cellular Biology* 27, 2130–2143.
- Shippen-Lentz, D., Blackburn, E.H., 1990. Functional evidence for an RNA template in telomerase. *Science* 247, 546–552.
- Tesmer, V.M., Ford, L.P., Holt, S.E., *et al.*, 1999. Two inactive fragments of the integral RNA cooperate to assemble active telomerase with the human protein catalytic subunit (hTERT) *in vitro*. *Molecular and Cellular Biology* 19, 6207–6216.
- Theimer, C.A., Blois, C.A., Feigon, J., 2005. Structure of the human telomerase RNA pseudoknot reveals conserved tertiary interactions essential for function. *Molecular Cell* 17, 671–682.
- Venteicher, A.S., Abreu, E.B., Meng, Z., *et al.*, 2009. A human telomerase holoenzyme protein required for Cajal body localization and telomere synthesis. *Science* 323, 644–648.
- Vuliamy, T.J., Marrone, A., Knight, S.W., *et al.*, 2006. Mutations in dyskeratosis congenita: Their impact on telomere length and the diversity of clinical presentation. *Blood* 107, 2680–2685.
- Wang, F., Podell, E.R., Zaug, A.J., *et al.*, 2007. The POT1-TPP1 telomere complex is a telomerase processivity factor. *Nature* 445, 506–510.
- Wang, F., Stewart, J.A., Kasbek, C., *et al.*, 2012. Human CST has independent functions during telomere duplex replication and C-strand fill-in. *Cell Reports* 2, 1096–1103.
- Watson, J.D., 1972. Origin of concatemeric T7 DNA. *Nature: New Biology* 239, 197–201.
- Witkin, K.L., Collins, K., 2004. Holoenzyme proteins required for the physiological assembly and activity of telomerase. *Genes & Development* 18, 1107–1118.
- Wotton, D., Shore, D., 1997. A novel Rap1p-interacting factor, Rif2p, cooperates with Rif1p to regulate telomere length in *Saccharomyces cerevisiae*. *Genes & Development* 11, 748–760.
- Wu, P., Takai, H., de Lange, T., 2012. Telomeric 3' overhangs derive from resection by Exo1 and Apollo and fill-in by POT1b-associated CST. *Cell* 150, 39–52.
- Xie, M., Mosig, A., Qi, X., *et al.*, 2008. Structure and function of the smallest vertebrate telomerase RNA from teleost fish. *Journal of Biological Chemistry* 283, 2049–2059.
- Xin, H., Liu, D., Wan, M., *et al.*, 2007. TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase. *Nature* 445, 559–562.
- Ye, J.Z.-S., de Lange, T., 2004. TIN2 is a tankyrase 1 PARP modulator in the TRF1 telomere length control complex. *Nature Genetics* 36, 618–623.
- Zhang, Q., Kim, N.-K., Feigon, J., 2011. Architecture of human telomerase RNA. *Proceedings of the National Academy of Sciences of the United States of America* 108, 20325–20332.
- Zhong, Z., Shiue, L., Kaplan, S., de Lange, T., 1992. A mammalian factor that binds telomeric TTAGGG repeats *in vitro*. *Molecular and Cellular Biology* 12, 4834–4843.

Further Reading

- Armanios, M., 2012. Telomerase and idiopathic pulmonary fibrosis. *Mutation Research* 730, 52–58.
- Blackburn, E.H., Collins, K.L., 2011. Telomerase: An RNP enzyme synthesizes DNA. *Cold Spring Harbor Perspectives in Biology* 3, 1–9.
- Cristofari, G., Sikora, K., Lingner, J., 2007. Telomerase unplugged. *ACS Chemical Biology* 2, 155–158.
- Harley, C.B., 1991. Telomere loss: Mitotic clock or genetic time bomb? *Mutation Research* 256, 271–282.
- Hodes, R.J., Hathcock, K.S., Weng, N.-P., 2002. Telomeres in T and B cells. *Nature Reviews. Immunology* 2, 699–706.
- Lue, N.F., 2004. Adding to the ends: What makes telomerase processive and how important is it? *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 26, 955–962.
- Nandakumar, J., Cech, T.R., 2013. Finding the end: Recruitment of telomerase to telomeres. *Nature Reviews Molecular Cell Biology* 14, 69–82.
- Palm, W., de Lange, T., 2008. How shelterin protects mammalian telomeres. *Annual Review of Genetics* 42, 301–334.
- Podlevsky, J.D., Chen, J. J.-L., 2012. It all comes together at the ends: Telomerase structure, function, and biogenesis. *Mutation Research* 730, 3–11.
- Sandin, S., Rhodes, D., 2014. Telomerase structure. *Current Opinion in Structural Biology* 25C, 104–110.
- Wyatt, H.D.M., West, S.C., Beattie, T.L., 2010. InTERTpreting telomerase structure and function. *Nucleic Acids Research* 38, 5609–5622.

Relevant Website

<http://telomerase.asu.edu>
The telomerase database.

Telomere Biology

DC Teasley and SA Stewart, Washington University School of Medicine, Saint Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

ALT Alternative lengthening of telomeres, a recombination-based mechanism for lengthening telomeres independent of telomerase.

Breakage–fusion–bridge cycle A mechanism of chromosome instability induced by telomere loss; following replication, telomere-free sister chromatids fuse and are subsequently broken apart during anaphase, not necessarily at the site of fusion, creating new telomere-free ends which may continue the cycle.

End replication problem The progressive shortening of telomeres that occurs with each cell division as a consequence of the mechanism of lagging strand DNA synthesis and telomere resection.

G-quadruplex A secondary structure possible in certain G-rich sequences, whereby four single-stranded portions of DNA or RNA form a helical structure stabilized by cyclic hydrogen bonding between guanine bases.

Senescence A state of irreversible cell cycle arrest induced by either telomere shortening (replicative senescence) or stress (stress-induced premature senescence).

Shelterin A six-member telomere-specific protein complex consisting of TRF1, TRF2, POT1, Rap1, TPP1, and TIN2; shelterin is involved in regulating numerous processes occurring at the telomere.

Telomerase A ribonucleoprotein holoenzyme that adds telomeric DNA repeats to the terminal end of an existing telomere, minimally composed of the telomerase reverse transcriptase protein (TERT) and the telomerase RNA component (TERC or TR).

Telomere crisis A state of excessive telomere shortening occurring in p53-deficient cells; critically short telomeres fuse to one another, initiating the breakage–fusion–bridge cycle and causing substantial genomic instability.

TERRA Telomeric repeat-containing RNA, the functional RNA product resulting from RNA polymerase II transcription of telomeres.

t-Loop A secondary structure that forms at telomeres; strand invasion of the 3' ssDNA telomere overhang into telomeric dsDNA (of the same telomere) producing a dsDNA loop (the t-loop) and a small ssDNA displacement loop (D-loop) at the site of invasion.

Introduction

All eukaryotic genomes are organized into linear chromosomes. As a result, each chromosome possesses termini that must be protected from two distinct problems not present in organisms with circular genomes. The first of these is the end replication problem, a consequence of the mechanism of lagging strand DNA synthesis. During lagging strand synthesis, when the last Okazaki fragment on the lagging strand template is processed to remove the ribonucleotide primer, an irreparable 5' gap remains. This gap on the lagging strand, as well as nucleolytic processing of the chromosome ends, results in a progressive shortening of the linear chromosome with each cell division (Zakian, 2012). Telomeres are the means by which cells avoid the loss of genetic information that would occur as a result of the end replication problem – instead of a progressive loss of protein-coding sequences, a progressive loss of noncoding telomere sequence occurs instead. In serving as a stopgap to the end replication problem, telomeres also act as a molecular clock that regulates cellular lifespan. In the absence of elongation mechanisms, telomeres eventually erode to lengths that are no longer sufficient to protect against loss of genetic information. Once telomeres become critically short, they trigger cellular senescence to prevent continued division that might result in a loss of genetic information and genome instability (Frias *et al.*, 2012). The second potential problem resulting from linear chromosome ends arises from the fact that all cells possess exquisitely sensitive mechanisms to recognize free DNA ends as DNA damage. By

coordinating the activities of DNA repair proteins, telomeres protect natural chromosome ends from DNA repair events that would result in chromosome fusions (de Lange, 2004). Since inappropriate chromosome fusions lead to genome instability, telomere integrity is essential for cell proliferation (Frias *et al.*, 2012).

The molecular events that telomeres are associated with have consequences reaching far beyond the proliferation of a given cell. Telomere dysfunction has been linked to a number of diseases including dyskeratosis congenita (DC), aplastic anemia, and emphysema (Armanios and Blackburn, 2012). Despite the wide range of organismal effects that can occur as a result of telomere dysfunction, perhaps no disease is more strongly linked to telomere biology than cancer. Telomere attrition is classically considered to be a tumor-suppressive mechanism due to the fact that shortened telomeres cause checkpoint activation and cellular senescence (Xu *et al.*, 2013). Because senescence is a potent obstacle to transformation, cells that become neoplastic must stabilize their telomeres. Along similar lines, telomere shortening or loss has the potential to cause substantial genome instability resulting in apoptosis or mitotic catastrophe (Xu *et al.*, 2013). Despite these strong tumor-suppressive effects, telomere dysfunction can also act as a tumor promoting mechanism. Since genome instability can lead to oncogenic translocations, gene amplification, and loss of heterozygosity of tumor suppressor genes, telomere dysfunction can also enhance the transformation process (Xu *et al.*, 2013). These opposing roles – preventing and promoting cancer – underscore the complexity of the molecular activities

that maintain telomeres, and the activities that take place in response to telomere shortening.

Telomere Structure and Proteins

Telomere DNA

The telomere sequence in eukaryotic organisms is canonically composed of short, G-rich repeats. Indeed, all identified vertebrate telomeres are composed of the sequence 5'-(TTAGGG)_n-3' oriented toward the chromosome terminus (McEachern *et al.*, 2000). Similar highly repetitive, G-rich sequences exist in other organisms; for instance, *Tetrahymena thermophila* telomeres consist of 5'-(TTGGGG)_n-3' repeats (Fulcher *et al.*, 2014). Telomeres are composed largely of double-stranded DNA (dsDNA), with a relatively short 3' single strand DNA (ssDNA) overhang of the G-rich strand (Sfeir, 2012). The length of the overhang can vary – in humans it is typically between 30 and 500 nucleotides – but its presence is essential, as evidenced by the fact that it is actively produced by resection following DNA replication (Novo and Londoño-Vallejo, 2013; Sfeir, 2012). Telomeres span a wide range of total lengths, from as short as 300 base pairs in yeasts, to between 2 and 15 kb in humans, and as long as 150 kb in tobacco (Fulcher *et al.*, 2014).

Not all organisms follow the theme of compact, regular repeats; for instance, *Saccharomyces cerevisiae* telomeres are usually designated 5'-(C₁₋₃A/TG₁₋₃)-3' to indicate a consensus core telomere sequence (Wellinger and Zakian, 2012). Budding yeast telomere repeats in fact range from 8 to 26 bp in length, and are on average less G-rich than typical telomere repeats (McEachern *et al.*, 2000). Even human telomeres show slight variability, generally having perfect repeats in the centromere-distal majority of the telomere with variant repeats occurring in the centromere-proximal end of the telomere (McEachern *et al.*, 2000). Small variations like those found in human telomeres are likely the consequence of errors by the DNA replication machinery. Since the centromere-proximal end of the telomere is more likely exclusively produced by the DNA replication machinery rather than telomerase, sequence polymorphisms can manifest and remain (McEachern *et al.*, 2000).

Telomere Secondary Structure

Because of the repetitive sequence and ssDNA overhang, telomeres can form secondary structures more complex than a linear dsDNA stretch with an ssDNA terminus. The most documented of these is the telomeric loop, or t-loop. T-loops form by a strand invasion event in which the 3' ssDNA overhang invades the dsDNA portion of the telomere. This event produces a small (approximately 150 nucleotide) ssDNA displacement loop (D-loop) of G-rich sequence at the site of invasion, as well as a large dsDNA loop (t-loop) (de Lange, 2004). T-loops were first identified in electron micrographs of telomeric DNA, and have since been observed using super-resolution fluorescence microscopy (de Lange, 2004; Doksan *et al.*, 2013). At face value, t-loops would seem to be an effective means to prevent end-to-end chromosome fusions;

by sequestering the 3' overhang, t-loops inhibit ATM signaling and nonhomologous end joining. However, t-loops themselves resemble strand invasion intermediate structures produced during homologous recombination; it is unclear if such a structure elicits a DNA damage response (DDR) if persistent, and how telomeres might avoid this response. It is likely that telomeric proteins play a role in both the formation and stabilization of t-loops; the telomere protein TRF2 in particular is sufficient for t-loop formation *in vitro*, and is necessary for t-loop formation and maintenance in mammalian cells (de Lange, 2004; Doksan *et al.*, 2013).

In addition to t-loops, telomeres are also capable of forming G-quadruplexes. G-quadruplexes form by an association of four single strands of DNA or RNA (monomeric, dimeric, or tetrameric in origin) in a helical structure, where the strands assemble such that four guanines align in a cyclic Hoogsteen hydrogen-bonded tetrad (Paeschke *et al.*, 2010). G-quadruplexes are characterized by the stacking of multiple tetrad cores that are connected by linker 'loops,' and can form in multiple orientations (Phan, 2010). Both the ssDNA overhang and ssDNA portions of the G-strand that form during replication or repair can presumably form into G-quadruplexes. The ability of the telomere to form G-quadruplexes is a significant phenomenon, as G-quadruplexes inhibit both semiconservative DNA replication and telomerase-mediated telomere lengthening (Paeschke *et al.*, 2011). Data from ciliates demonstrating that G-quadruplexes form at telomeres in a cell cycle-specific manner, and the observation that treatment of mammalian cells with G-quadruplex-stabilizing small molecules triggers telomere dysfunction, suggest that G-quadruplexes at the telomere must be actively regulated by the cell (Lipps and Rhodes, 2009). Among the strongest candidates for G-quadruplex regulation in the cell are helicases; in particular, the Fanconi anemia group helicase FANCD1 and RecQ helicase BLM are known to unwind G-quadruplexes, and are known to contribute to telomere stability (Lipps and Rhodes, 2009).

Proteins Associated with the Telomere

In addition to the DNA itself, telomeres are host to a number of proteins important for telomere maintenance and function. In mammals, the network of proteins present at the telomere are coordinated by six telomere-specific proteins: TRF1 (telomeric repeat-binding factor 1, also TERF2), TRF2 (telomeric repeat-binding factor 2, also TERF2), POT1 (protection of telomeres protein 1), Rap1 (telomeric repeat-binding factor 2-interacting protein 1, also TERF2IP), TPP1 (adrenocortical dysplasia protein homolog, also ACD), and TIN2 (TRF1-interacting factor 2, also TIN2) (de Lange, 2005). These six proteins together form the shelterin complex, which exclusively binds telomeres due to the DNA-binding specificities of TRF1 and TRF2 for telomeric dsDNA, and of POT1 for telomeric ssDNA.

TRF1 and TRF2 share homology in the form of a central TRFH dimerization domain and C-terminal Myb DNA-binding domain, though TRF1 possesses an acidic N-terminus, while TRF2's N-terminus is basic. TRF1 is also substantially more divergent (65% identity between human and mouse) in

mammalian evolution than TRF2 (82% identity) (Broccoli *et al.*, 1997). While both proteins bind telomeric dsDNA and are abundant at telomeres, their functions in telomere maintenance and stability are very different. TRF1 is required for semiconservative DNA replication through the telomere, and prevents a phenotype known as telomere fragility by recruiting helicases to the telomere to facilitate replication fork progression (Sfeir, 2012). TRF1 is also capable of looping, bending, and pairing arrays of telomere repeats, which may be involved in positioning or folding of telomeres, though TRF1 is dispensable for t-loop formation (de Lange, 2005; Doksan *et al.*, 2013). TRF2 suppresses ATM kinase activation at the telomere, thus preventing a DDR (H2AX phosphorylation and 53BP1 accumulation) and p53 activation (Sfeir, 2012). In the absence of TRF2, ligase IV- and Ku-mediated nonhomologous end joining occurs, resulting in aberrant end-to-end chromosome fusions and early embryonic lethality in mice. The protective activities of TRF2 likely originate from its ability to facilitate t-loop formation, as well as its role (along with TRF1) in recruiting TPP1 to the telomere (de Lange, 2005, 2004; Doksan *et al.*, 2013). Putative TRF1/2 homologs have been identified in *Schizosaccharomyces pombe*, trypanosomes, and plants (Sfeir, 2012).

The other mammalian shelterin protein with telomere sequence binding specificity, POT1, binds to the ssDNA overhang. POT1 is responsible for suppression of ATR kinase activation at the telomere, which is achieved by exclusion of the ssDNA-binding protein RPA from the ssDNA overhang (Baumann and Price, 2010). POT1 also restricts the length of the ssDNA overhang and regulates the activity of telomerase by competing for binding at the overhang, which is the substrate for telomerase elongation (Longhese *et al.*, 2012; Sfeir, 2012). POT1 is largely conserved among eukaryotes, with homologs in *S. pombe*, *Tetrahymena*, nematodes, and plants (Baumann and Price, 2010). In mouse and *Tetrahymena*, there are two POT1 gene homologs, *Pot1a* and *Pot1b*, with each playing a subset of the roles attributed to the single POT1 gene observed in human and yeast.

The three other shelterin proteins (Rap1, TPP1, and TIN2) lack the ability to bind telomeric sequence directly in vertebrates, but are telomere-specific by virtue of their direct or indirect binding to TRF1 and TRF2. Indeed, simultaneous deletion of TRF1 and TRF2 from cells results in telomeres completely devoid of shelterin (Sfeir and de Lange, 2012). Rap1 is the most conserved of all the shelterin proteins and the only shelterin protein with well-defined extra-telomeric functions – it is a transcriptional regulator and can modulate nuclear factor kappa B (NF- κ B) signaling. At the mammalian telomere, Rap1 is a negative regulator of telomere length. It also acts with POT1 to repress aberrant homology-directed repair through an unclear mechanism (Sfeir, 2012). In *S. cerevisiae*, which lacks a TRF1/2 homolog, Rap1 is highly diverged from its orthologs and directly binds telomeric dsDNA; it effectively serves as the shelterin core in this yeast (de Lange, 2005). TIN2 is a 'bridging' protein that connects TRF1 and TRF2 to one another, and recruits TPP1 to the telomere. TPP1 in turn is required for efficient recruitment of POT1 to the telomere, as POT1's DNA-binding affinity is insufficient to keep it tethered to the ssDNA overhang (Sfeir, 2012). Both TIN2 and TPP1 are exclusive to vertebrate telomeres, and their

emergence may coincide with the appearance of two TRF genes (de Lange, 2005; Figure 1).

In addition to the telomere-specific shelterin proteins, a growing number of other proteins localize to telomeres and are important for telomere function. The majority of these proteins are associated with DNA metabolism. The protein complex CST, consisting of Cdc13, Stn1, and Ten1 in *S. cerevisiae* and CTC1, STN1, and TEN1 in mammals, associates with the ssDNA telomere overhang and due to structural similarity to the heterotrimeric RPA complex has been proposed to function as a 'telomere-specific RPA' (Longhese *et al.*, 2012). The RecQ helicases WRN and BLM interact with shelterin and play roles in lagging strand telomere synthesis and replication fork progression (de Lange, 2005). These are but a few of the numerous proteins that localize to the telomere and contribute to telomere synthesis and maintenance.

Telomere Replication and Length Maintenance

Telomeres, like the rest of the genome, must be replicated during S phase to ensure genome continuity during cell division. Telomeres can be replicated primarily in two ways: first, the semiconservative DNA replication machinery replicates telomeres along with the rest of the genome, and second, the ribonucleoprotein (RNP) enzyme telomerase adds telomere repeats to extend telomeres. In addition to these two main mechanisms for telomere replication and maintenance, there are two additional identified means of telomere length maintenance: recombination-based telomere maintenance and the alternative lengthening of telomeres (ALT) mechanism. Lastly, some insects maintain their telomeres by a unique mechanism involving retrotransposition, but this will not be discussed here (de Lange, 2004).

Semiconservative Telomere Replication

Telomere replication poses unique challenges for the semiconservative DNA replication machinery. The first of these challenges is the consequence of the telomere's location at chromosome termini. Unlike the rest of the genome, which can be replicated by one or more DNA replication origins with replication forks approaching from either direction, the most centromere-distal origin is thought to be the sole origin responsible for the replication of a given telomere. Additionally, because there are no known replication origins within the telomere, the 'last' origin to fire must replicate the entire length of the telomere; in the event of fork collapse, telomere replication may remain incomplete (Gilson and Géli, 2007). This potential problem is exacerbated by the second main challenge in telomere replication: the repetitive, G-rich sequence, which leads to increased replication fork stalling compared to the rest of the genome (Cesare and Karlseder, 2012). The secondary structures that telomeric DNA forms, in particular G-quadruplexes and t-loops, must be resolved into linear dsDNA for replication to occur.

The combination of structures that induce replication fork stalling and the lack of a replication fork approaching from the opposite direction to rescue a stalled fork necessitates robust

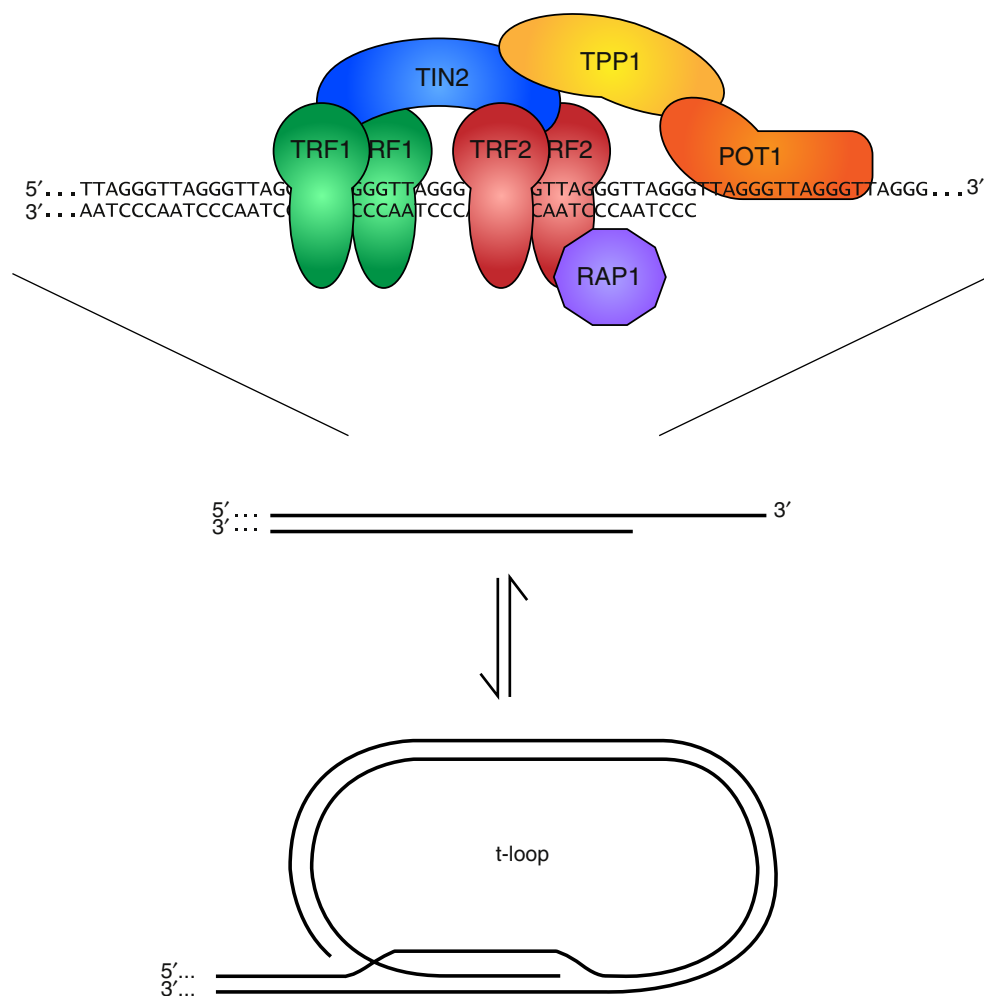


Figure 1 A diagrammatic representation of telomere structure and telomere-binding proteins. TRF1 (green) and TRF2 (red) form homodimers and directly bind telomere dsDNA. TIN2 (blue) acts as a bridge between the homodimers and interacts with TPP1 (yellow). TPP1 also interacts with POT1 (orange), which binds telomere ssDNA at the 3' overhang. Below, the t-loop structure is shown, where the G-rich 3' end loops and hybridizes with the C-rich strand, forming a displacement loop.

mechanisms to ensure that telomere replication is completed. These mechanisms are coordinated by the shelterin complex, which recruits DNA replication and repair proteins to the telomere. In *S. pombe*, the TRF1/2 homolog Taz1 prevents replication fork arrest and telomere loss in a telomerase-negative background (Gilson and Géli, 2007). Subsequent work in murine cells has shown that TRF1 is required to facilitate replication fork progression through the telomere and prevent telomere fragility, and that this activity depends upon the helicases BLM and RTEL1 (Sfeir *et al.*, 2009). TRF2 also plays critical roles in telomere replication – both by recruiting the RecQ helicases BLM and WRN to the telomere, and by acting in a pathway with Apollo and topoisomerase 2 α to relieve topological stress during replication (Ye *et al.*, 2010). The recruitment of two RecQ helicases, BLM and WRN, by shelterin is not coincidental; both proteins are able to unwind G-quadruplexes and are theorized to have overlapping functions in resolving G-quadruplexes formed on the lagging strand during replication.

The cooperation between shelterin and its binding partners described above ensures that replication forks can progress through the telomere with as little stalling as possible. Nevertheless, replication forks do stall in telomeric sequence, and additional mechanisms are in place to ensure successful fork restart. The DNA replication and repair protein flap endonuclease 1 (FEN1) localizes to telomeres during and after replication, and ensures that replication forks are reinitiated on the lagging strand template following stalling (Saharia *et al.*, 2010). The mammalian CST complex rescues stalled replication forks that occur as a result of replication stress by facilitating dormant origin firing (Stewart *et al.*, 2012). Stalled replication forks at the telomere appear to require ATR, which is recruited to telomeres during S phase, for restart (Verdun and Karlseder, 2006). These are unlikely to be the only contributing mechanisms for replication fork restart at the telomere. The preponderance of DNA replication and repair proteins that are recruited to the telomere to facilitate replication fork progression and fork restart following stalling

illustrates the difficulty of telomere replication by the semi-conservative DNA replication machinery, as well as the robust series of mechanisms in place to ensure that replication is completed in spite of the challenges.

The fact that both leading and lagging strand-replicated telomeres possess similar 3' ssDNA overhangs implies that at least the leading strand, which would be expected to produce a blunt end, has to be processed to produce an overhang. In fact, leading and lagging strand-replicated telomere ends are significantly resected and processed following the completion of semiconservative DNA replication (Gilson and Géli, 2007). This process must be highly regulated by the cell, as too little resection would generate a short overhang which precludes POT1 binding and t-loop formation, while too much resection would accelerate the end replication problem. In *S. cerevisiae*, resection of the leading strand-replicated telomere to produce the G-rich overhang is facilitated by the same DNA repair proteins involved in 5' resection at DNA double strand breaks prior to homologous recombination, notably Cdk1 and the Mre11–Rad50–Xrs2 (MRX) complex. Indeed, yeast strains with deficient Mre11 exhibit shorter 3' ssDNA overhangs (Gilson and Géli, 2007). Following recognition of the end by MRX, the nucleases Exo1 and/or Dna2 are recruited to the 5' end, and in concert with the helicase Sgs1, cleave the 5' end to produce the 3' ssDNA overhang (Stewart *et al.*, 2012). In mammalian cells, the Apollo nuclease is recruited to telomere ends by its interaction with TRF2. At the leading strand-replicated telomere, Apollo initiates resection of the 5' end. At the lagging strand-replicated telomere, where an overhang already exists following replication, POT1b binds the ssDNA and inhibits the resection activity of Apollo; similarly, after a sufficient ssDNA overhang is generated at the leading strand-replicated telomere, POT1b binds and inhibits further resection by Apollo (Wu *et al.*, 2012). Exo1 then further and transiently resects the 5' ends at both strands to produce long overhangs. Lastly, POT1b recruits CST and polymerase α to shorten the extended overhangs, presumably by fill-in C-strand synthesis (Wu *et al.*, 2012).

Telomerase

Semiconservative replication and telomere end processing cause the end replication problem, which manifests as progressive telomere shortening with each cell division cycle. The solution to the end replication problem is telomerase, a holoenzyme minimally composed of the telomerase reverse transcriptase protein (TERT) and the telomerase RNA component (TERC or TR). TERC binds accessory proteins in addition to TERT, which contribute to telomerase localization and processing (Martínez and Blasco, 2011). Telomerase adds telomeric DNA repeats to the terminal end of an existing telomere, thus lengthening telomeres and offsetting the end replication problem.

Telomerase is unique among known reverse transcriptase enzymes in being an RNP. The TERT protein is the catalytic portion of the holoenzyme, and consists of four conserved structural domains. Two of these, the reverse transcriptase and C-terminal extension domains, are conserved between TERT and other reverse transcriptases; the telomerase essential N-terminal domain and telomerase RNA-binding domain are

unique to TERT (Podlevsky and Chen, 2012). The RNA portion of the telomerase RNP, TERC, is required for telomerase activity not only because it provides the telomere repeat sequence template, but also because two conserved regions, the template/pseudoknot domain and the CR4/5 domain, contribute to template positioning in the active site and provide important protein–nucleic acid contacts (Podlevsky and Chen, 2012). In addition to these two regions, which are conserved among all known species, vertebrate TERCs contain a conserved H/ACA domain. Each of two stems in the H/ACA domain binds a protein complex consisting of dyskerin, NOP10, NHP2, and GAR1. These proteins are required for TERC maturation and processing, RNP biogenesis, and Cajal body localization of TERC (Podlevsky and Chen, 2012).

TERT protein production occurs via the canonical processes of mRNA transcription and cytoplasmic translation, after which TERT is first recruited to nucleoli, and then Cajal bodies. TERC is transcribed by RNA polymerase II, after which the ends are processed; TERC binds the dyskerin-anchored complex of accessory proteins, which facilitate maturation and localization to Cajal bodies. The chaperone proteins HSP90 and p23 facilitate the assembly of the TERC–accessory protein complex with TERT into the active telomerase RNP holoenzyme, after which it localizes to telomeres (Podlevsky and Chen, 2012).

At the telomere, the catalytic process of telomere elongation occurs in two steps. In the first step, the 3' end of the telomere base pairs with the 5' region of TERC; the TERT active site then uses the 3' end of the telomere as a primer to reverse transcribe a telomere repeat (in human, the six nucleotides 5'-GGTTAG-3') using TERC as a template. In the second step, TERC dissociates from the telomeric DNA, translocates 5'-to-3' along the DNA, and re-anneals for a new round of nucleotide addition (Podlevsky and Chen, 2012). By repeating the synthesis and translocation steps, telomerase can processively add repeats to a single telomere end without ever completely dissociating from the DNA. Here, shelterin plays a role in telomere synthesis, as the POT1–TPP1 complex has been shown to hold the telomeric DNA primer close to telomerase, inhibiting primer release and enhancing processivity (Podlevsky and Chen, 2012). Telomerase activity is also restricted in a shelterin-dependent manner; POT1 bound to the ssDNA overhang is thought to limit the initial binding of telomerase to the telomere, and the t-loop structure prevents telomerase activity (de Lange, 2005). Once telomerase dissociates from its template, polymerase α synthesizes the complementary C-rich strand. It is thought that the CST complex, which is known to interact with the polymerase α complex, facilitates this event (Gilson and Géli, 2007). In yeast, telomere repeat-containing RNA (TERRA), the noncoding RNA product of telomere transcription, may act as a seed to nucleate clusters of telomerase prior to telomere recruitment; TERRA transcription is induced by telomere shortening (Cusanelli *et al.*, 2013).

Recombination at Telomeres and ALT

In addition to lengthening by telomerase activity, cells have evolved other mechanisms to elongate or maintain telomeres. In *S. cerevisiae*, deletion of the telomerase reverse transcriptase

EST1 produces colonies of survivor cells. Genetic analysis revealed the requirement for recombination in virtually all *est1Δ* survivor cells, as *est1Δ rad52Δ* strains produce virtually no survivors (Wellinger and Zakian, 2012). In addition to *RAD52*, all survivor cells require the Pol32p replication protein. Survivors are categorized into one of two types (I and II), which appear to have different mechanisms but are not mutually exclusive. Type I survivors have telomeres with multiple, repeated Y' subtelomere elements and short, terminal telomere repeats; they also possess extrachromosomal circular Y' elements thought to serve as recombination substrates. These cells grow relatively slowly, easily convert to type II survivors, and require *RAD51*, *RAD54*, and *RAD57* in addition to *RAD52* and *POL32* (Wellinger and Zakian, 2012). Type II survivors have telomeres with few subtelomere repeats but extensive amplification in telomere repeats. The telomeres in these cells may depend on rolling circle amplification as an initiating event to lengthen their telomeres. Type II survivors require *MRX*, *RAD59*, and *SGS1* (Wellinger and Zakian, 2012). Interestingly, the recombination-mediated mechanisms of telomere maintenance in yeast appear to be promoted by RNA:DNA hybrid formation between telomeric DNA and TERRA (Balk *et al.*, 2013). This observation suggests that TERRA may be required for both telomerase-mediated and recombination-mediated telomere elongation pathways.

In mammalian cells lacking telomerase activity, the ALT mechanism provides a means to maintain telomere length. Found in 10–15% of human cancers, ALT appears to be recombination-based and is characterized by several striking phenotypes. ALT cells possess an abundance of extrachromosomal telomeric DNA, much of which is in the form of double-stranded telomeric circles (t-circles) and C-rich single-stranded telomeric circles (C-circles) (Cesare and Reddel, 2010). ALT cells also exhibit telomere localization to promyelocytic leukemia (PML) nuclear bodies, elevated levels of telomere sister chromatid exchange (T-SCE) events, and heterogeneous chromosomal telomere lengths. The molecular mechanism(s) responsible for ALT have remained elusive, and two models have emerged to explain it. First, the unequal T-SCE model proposes that T-SCE events produce long and short telomere lengths in the chromosomes experiencing T-SCE; if the chromosomes with longer telomeres were able to segregate into one of the two daughter cells, the enhanced proliferative capacity of one daughter over the other could produce selection at the population level for theoretically unlimited proliferation (Cesare and Reddel, 2010). The second model, which is not mutually exclusive from the unequal T-SCE model, proposes that ALT is the result of homologous recombination-dependent DNA replication. In this model, shorter telomeres extend by recombination-based DNA synthesis using an existing telomere sequence substrate in the cell. The telomere substrate for the elongating telomere could be a telomere on another chromosome, the same telomere via t-loop formation, linear extrachromosomal telomeric DNA, t-circles, and/or C-circles (Cesare and Reddel, 2010). Both telomere elongation and length maintenance in ALT cells appear to depend on a striking number of DNA replication and repair proteins, which suggests that ALT results from deregulation of normal DNA metabolic processes. Indeed, the shelterin proteins TRF2 and POT1 have been suggested as

suppressors of ALT due to both proteins' ability to inhibit recombination at the telomere; ALT cells' vast expansion of telomeric DNA content may dilute shelterin saturation at telomeres and contribute to the phenotype (Cesare and Reddel, 2010). Like in recombination-mediated telomere elongation in yeast, TERRA may play a role in ALT. Recent data suggests that RNA:DNA hybrid levels between TERRA and the telomere are 'fine-tuned' by RNase H1 expression in ALT cells (Arora *et al.*, 2014).

Telomere Physiology

Senescence and Telomere Crisis

One of the most significant consequences of telomere erosion at the cell biological level is the induction of cellular senescence. Indeed, the fact that telomere length is the primary determinant of cellular lifespan is well established (Deng *et al.*, 2008). Telomere shortening beyond a critical length causes deprotection of the telomere via a loss of TRF2. Deprotection induced by replicative shortening or other telomere dysfunction induces a DDR. The DDR induced by telomere dysfunction causes phosphorylation of H2AX and recruitment of 53BP1 and other DNA repair proteins to the telomere. Like the classical DDR, the DDR induced by telomere attrition activates both the ATM and ATR checkpoints, which in turn activate p53 via the CHK1 and CHK2 kinases (Deng *et al.*, 2008). p53 activation by telomere shortening induces one of two physiological consequences: either entry into cellular senescence by p21 and RB signaling, or apoptosis (Frias *et al.*, 2012). Both senescence and apoptosis serve as potent antiproliferative mechanisms in cells with eroded telomeres (Figure 2).

Absence of a functional p53 and Rb checkpoint, as often occurs in cancer cells, abrogates both senescence and apoptosis as responses to telomere dysfunction and attrition. Because this causes checkpoint bypass and continued proliferation, telomeres in p53-deficient cells will continue to shorten, ultimately reaching a condition called crisis. Crisis inherently acts as a second checkpoint to continued proliferation; critically short telomeres induce aberrant end-to-end fusions between chromosomes followed by chromosome breakage and additional fusion events in what is referred to as the breakage–fusion–bridge cycle (Deng *et al.*, 2008). The breakage–fusion–bridge cycle causes massive genomic instability and aneuploidy, both of which lead to rapid cell death in a majority of cells. However, a small fraction of cells are able to escape crisis by activating either telomerase or the ALT pathway; in these cells, the chromosome fusions resulting from crisis can produce oncogene amplification, tumor suppressor loss of function, and gene fusions that contribute to rapid proliferation and tumorigenesis (Deng *et al.*, 2008).

Telomeres in Cancer Cells

Telomere biology contributes to cancer phenotypes via two distinct but related mechanisms. First, telomere shortening induces senescence and/or apoptosis, acting as a significant antiproliferative barrier to the incipient tumor cell; these cells must

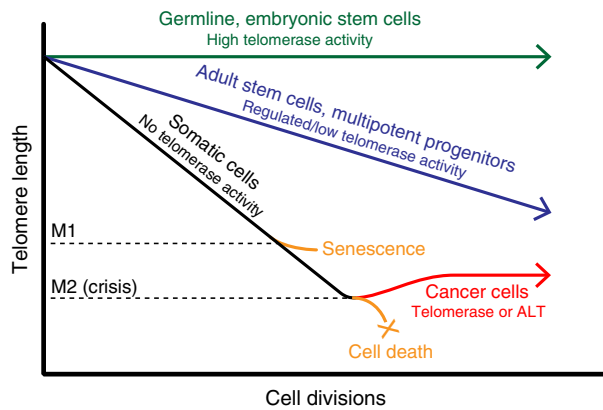


Figure 2 A depiction of telomere length over the course of cell divisions. Germline cells and embryonic stem cells (green) possess high telomerase activity and maintain long telomeres. Adult stem cells and multipotent progenitor cells (purple) possess detectable telomerase activity, but do not express telomerase sufficiently to entirely prevent telomere shortening; the telomeres in these cells shorten gradually over time. Somatic cells have no detectable telomerase activity, and the telomeres in these cells shorten with each cell division. When telomeres shorten to a particular length (M1), most cells enter senescence and cease dividing. In cells without functional p53/Rb checkpoints, senescence is bypassed; cell division and telomere shortening continue until telomere length becomes short enough to trigger crisis (M2). The majority of cells that reach crisis die; however, a very small subset of incipient cancer cells that reach crisis activate either telomerase or the ALT pathway to stabilize their short telomeres, thus becoming immortal.

activate either telomerase or ALT to continue proliferating. Second, telomere dysfunction, whether induced by DNA damage, genetic mutation, or avoidance of replicative senescence causes genomic instability that can drive tumorigenesis.

The observation that telomerase is aberrantly activated in as many as 85–90% human cancers, with ALT mechanisms active in the remainder, demonstrates the absolute barrier that telomere shortening imposes on incipient tumor cells' ability to survive (Gomez *et al.*, 2012). Despite the universal requirement for a telomere maintenance mechanism for unlimited proliferation, the significance of a tumor cell's use of telomerase versus ALT remains unclear. Several studies have identified an ability of cells to switch between the two mechanisms; of particular note is the ability of telomerase-positive cancer cells to become ALT-positive upon genetic or pharmacologic inactivation of telomerase (Hu *et al.*, 2012; Queisser *et al.*, 2013). In tumors that are telomerase-positive, studies conflict on the relevance of telomerase expression levels or activity as a prognostic marker. In some late-stage non-small cell lung cancers, colorectal cancers, and soft tissue sarcomas, telomerase expression correlates with poor prognosis, but many of these studies remain controversial (Chen and Chen, 2011). Tumors that are ALT-positive (commonly, glioblastoma multiforme and sarcomas) frequently have poor prognosis (Cesare and Reddel, 2010).

While aberrant telomere elongation by telomerase or ALT contributes directly to the proliferative capacity of cancer cells, telomere-related genomic instability is perhaps just as significant a contributor to tumorigenesis. As described in the Senescence and Telomere Crisis section, the breakage–fusion–

bridge cycle can arise from critically short or lost telomeres. In cancer cells, critically short telomeres are likely to induce the breakage–fusion–bridge cycle, resulting in genomic instability and generating tumor-promoting conditions. Indeed, mice with a *TERC* deletion and p53 deletion experience a wide range of carcinomas consistent with human cancers; when examined for cytogenetic abnormalities, the tumor cells in these mice show chromosome rearrangements with inverted repeats that are consistent with the breakage–fusion–bridge model (Murnane, 2010). Human cancer cells frequently contain amplified regions in the genome with inverted repeats (Murnane, 2012). Analysis of tumor samples by chromosome painting indicates that the breakage–fusion–bridge cycle frequently involves a single chromosome involved in multiple sister chromatid fusions, continuing over multiple generations until telomere stability can be achieved; this allows the cell to survive and facilitates substantial amplification of loci on the involved chromosome(s) (Murnane, 2012). In addition to amplification events generated by the breakage–fusion–bridge cycle, telomere loss can trigger end-to-end chromosome fusions that ultimately produce gross chromosomal rearrangements. These rearrangements are capable of producing fusion genes that encode oncogenic proteins in specific human cancers (Jones *et al.*, 2012).

Noncancer Telomere-Related Diseases

Diseases with telomeric origin or contribution can largely be grouped into two categories: those that are caused by mutation to specific genes with telomeric functions and those that are associated with changes in telomere length. The first category, often referred to as telomeropathies or telomere syndromes, constitutes a diverse group of diseases that are monogenic in origin, with mutations typically occurring in one of several telomere-associated genes. The most well-known telomeropathy is DC, a disease initially presenting with nail dystrophy, leukoplakia, and abnormal skin pigmentation; DC patients frequently progress to experience aplastic anemia and increased cancer incidence (Savage and Bertuch, 2010). Among the most frequently mutated genes in DC are *TERC* and *TERT*; DC is also associated with mutations in *DKC1* (common), *TINF2* (common), *NHP2* (rare), and *NOP10* (rare) (Holohan *et al.*, 2014). In addition, three other disorders presenting with overlapping symptoms – Hoyerall-Hreidarsson, Revesz, and Coats plus syndromes – are associated with mutations in the same group of genes, and may constitute a group of diseases including DC with common origins that exhibit distinct phenotypes due to different mutations or penetrance (Armanios and Blackburn, 2012). Most of the mutations identified in DC and the DC-related syndromes cause decreased telomerase function; in affected cells, telomere length is frequently below the first percentile in the population for the age of the patient (Savage and Bertuch, 2010). Mutations in *TERC* and *TERT* have also been identified in studies as contributors to aplastic anemia, idiopathic pulmonary fibrosis, and unexplained severe liver disease (Savage and Bertuch, 2010). The symptoms of these conditions are all observed in DC patients as either complications or comorbidities, suggesting that abrogation of telomerase function by mutation of *TERC*, *TERT*, or another protein can produce a wide spectrum of disease presentations that are all caused by similar molecular mechanisms.

Two mutated genes identified in diseases of the DC spectrum, *TINF2* (common) and *CTC1* (rare), produce proteins not known to directly affect telomerase function. *TINF2*, a shelterin complex member, could indirectly affect telomerase recruitment by modulating the binding of other shelterin proteins to the telomere, but the mechanism remains unclear. *CTC1*, associated only with Coats plus syndrome, is a member of the mammalian CST complex and thus facilitates telomere replication, but there is little evidence about how it might contribute to the disease (Armanios and Blackburn, 2012).

In addition to the monogenic telomeropathies described above, several age-related human diseases show association with telomere length. A wide number of risk factors for atherosclerosis (smoking, hypertension, and obesity), as well as atherosclerosis itself, have been studied with regard to telomere length. Reports indicate that tobacco use, hypertension with sclerotic plaques, and weight gain correlate with shorter telomere lengths; multiple studies have demonstrated that shorter telomeres or increased telomere attrition are present in atherosclerotic patients (Khan *et al.*, 2012). Some of these studies, particularly those examining atherosclerosis risk factors as opposed to the disease itself, have weaknesses: several studies were unable to control all the confounding factors, while others showed statistically significant effects only in certain subgroups (i.e., males or females). Nonetheless, the overall picture of data certainly indicates that shorter telomere length is associated with atherosclerosis (Khan *et al.*, 2012). Age-related musculoskeletal diseases have also been associated with telomere length. Chondrocytes from patients with osteoarthritis exhibit shorter telomere lengths than control patients; given that chondrocyte senescence causes changes known to accelerate osteoarthritis, it is possible that accelerated telomere loss might contribute to osteoarthritis (Li *et al.*, 2012). Some studies have indicated that shorter telomeres are associated with lower bone mineral density, a characteristic of osteoporosis (Li *et al.*, 2012). Other studies, however, have found no such association, and so associations between telomere length and age-related musculoskeletal diseases are challenging to identify and interpret. In mouse models, a number of other diseases appear to be telomere-dependent. Mice with short telomeres display a wide variety of phenotypes – these include: immune problems (opportunistic infections, impaired vaccine responses, and altered T-cell ratios), gastrointestinal problems (enterocolitis, villous blunting), defects in fibrosis following injury, and altered β -cell function leading to insulin resistance (Armanios and Blackburn, 2012). Many of the phenotypes observed in mice appear to be related to telomere-induced senescence. While most of the noncancer diseases with known telomere dependence in humans fall into the DC spectrum of disorders and phenotypes, the evidence in mice suggests that telomere-dependent senescence contributes to age-related diseases.

See also: Nucleic acid Synthesis/Breakdown: DNA Synthesis/Repair: Telomeres and Telomerase

References

- Armanios, M., Blackburn, E.H., 2012. The telomere syndromes. *Nature Reviews: Genetics* 13, 693–704.
- Arora, R., Lee, Y., Wischniewski, H., *et al.*, 2014. RNaseH1 regulates TERRA-telomeric DNA hybrids and telomere maintenance in ALT tumour cells. *Nature Communications* 5, 5220.
- Balk, B., Maicher, A., Dees, M., *et al.*, 2013. Telomeric RNA-DNA hybrids affect telomere-length dynamics and senescence. *Nature Structural & Molecular Biology* 20, 1199–1205.
- Baumann, P., Price, C., 2010. Pot1 and telomere maintenance. *FEBS Letters* 584, 3779–3784.
- Broccoli, D., Smogorzewska, A., Chong, L., de Lange, T., 1997. Human telomeres contain two distinct Myb-related proteins, TRF1 and TRF2. *Nature Genetics* 17, 231–235.
- Cesare, A.J., Karlseder, J., 2012. A three-state model of telomere control over human proliferative boundaries. *Current Opinion in Cell Biology* 24, 731–738.
- Cesare, A.J., Reddel, R.R., 2010. Alternative lengthening of telomeres: Models, mechanisms and implications. *Nature Reviews: Genetics* 11, 319–330.
- Chen, C.-H., Chen, R.-J., 2011. Prevalence of telomerase activity in human cancer. *Journal of the Formosan Medical Association* 110, 275–289.
- Cusanelli, E., Romero, C.A.P., Chartrand, P., 2013. Telomeric noncoding RNA TERRA is induced by telomere shortening to nucleate telomerase molecules at short telomeres. *Molecular Cell* 51, 780–791.
- Deng, Y., Chan, S.S., Chang, S., 2008. Telomere dysfunction and tumour suppression: The senescence connection. *Nature Reviews: Cancer* 8, 450–458.
- Doksani, Y., Wu, J.Y., de Lange, T., Zhuang, X., 2013. Super-resolution fluorescence imaging of telomeres reveals TRF2-dependent t-loop formation. *Cell* 155, 345–356.
- Frias, C., Pampalona, J., Genesca, A., Tusell, L., 2012. Telomere dysfunction and genome instability. *Frontiers in Bioscience* 17, 2181–2196.
- Fulcher, N., Derboven, E., Valuchova, S., Riha, K., 2014. If the cap fits, wear it: An overview of telomeric structures over evolution. *Cellular and Molecular Life Sciences* 71, 847–865.
- Gilson, E., Géli, V., 2007. How telomeres are replicated. *Nature Reviews* 8, 825–838.
- Gomez, D.E., Armando, R.G., Farina, H.G., *et al.*, 2012. Telomere structure and telomerase in health and disease (review). *International Journal of Oncology* 41, 1561–1569.
- Holohan, B., Wright, W.E., Shay, J.W., 2014. Telomeropathies: An emerging spectrum disorder. *Journal of Cell Biology* 205, 289–299.
- Hu, J., Hwang, S.S., Liesa, M., *et al.*, 2012. Antitelomerase therapy provokes ALT and mitochondrial adaptive mechanisms in cancer. *Cell* 148, 651–663.
- Jones, C.H., Pepper, C., Baird, D.M., 2012. Telomere dysfunction and its role in haematological cancer. *British Journal of Haematology* 156, 573–587.
- Khan, S., Chuturgoon, A.A., Naidoo, D.P., 2012. Telomeres and atherosclerosis. *Cardiovascular Journal of Africa* 23, 563–571.
- de Lange, T., 2004. T-loops and the origin of telomeres. *Nature Reviews: Molecular Cell Biology* 5, 323–329.
- de Lange, T., 2005. Shelterin: The protein complex that shapes and safeguards human telomeres. *Genes & Development* 19, 2100–2110.
- Li, D., Yuan, Q., Wang, W., 2012. The role of telomeres in musculoskeletal diseases. *Journal of International Medical Research* 40, 1242–1250.
- Lipps, H.J., Rhodes, D., 2009. G-quadruplex structures: In vivo evidence and function. *Trends in Cell Biology* 19, 414–422.
- Longhese, M.P., Anbalagan, S., Martina, M., Bonetti, D., 2012. The role of shelterin in maintaining telomere integrity. *Frontiers in Bioscience* 17, 1715–1728.
- Martínez, P., Blasco, M.A., 2011. Telomeric and extra-telomeric roles for telomerase and the telomere-binding proteins. *Nature Reviews: Cancer* 11, 161–176.
- McEachern, M.J., Krauskopf, A., Blackburn, E.H., 2000. Telomeres and their control. *Annual Review of Genetics* 34, 331–358.
- Murnane, J.P., 2010. Telomere loss as a mechanism for chromosome instability in human cancer. *Cancer Research* 70, 4255–4259.
- Murnane, J.P., 2012. Telomere dysfunction and chromosome instability. *Mutation Research* 730, 28–36.
- Novo, C.L., Londoño-Vallejo, J.A., 2013. Telomeres and the nucleus. *Seminars in Cancer Biology* 23, 116–124.
- Paeschke, K., Capra, J.A., Zakian, V.A., 2011. DNA replication through G-quadruplex motifs is promoted by the *Saccharomyces cerevisiae* Pif1 DNA helicase. *Cell* 145, 678–691.
- Paeschke, K., McDonald, K.R., Zakian, V.A., 2010. Telomeres: Structures in need of unwinding. *FEBS Letters* 584, 3760–3772.
- Phan, A.T., 2010. Human telomeric G-quadruplex: Structures of DNA and RNA sequences. *FEBS Journal* 277, 1107–1117.
- Podlevsky, J.D., Chen, J.J.-L., 2012. It all comes together at the ends: Telomerase structure, function, and biogenesis. *Mutation Research* 730, 3–11.

- Queisser, A., Heeg, S., Thaler, M., *et al.*, 2013. Inhibition of telomerase induces alternative lengthening of telomeres during human esophageal carcinogenesis. *Cancer Genetics* 206, 374–386.
- Saharia, A., Teasley, D.C., Duxin, J.P., *et al.*, 2010. FEN1 ensures telomere stability by facilitating replication fork re-initiation. *Journal of Biological Chemistry* 285, 27057–27066.
- Savage, S.A., Bertuch, A.A., 2010. The genetics and clinical manifestations of telomere biology disorders. *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 12, 753–764.
- Steir, A., 2012. Telomeres at a glance. *Journal of Cell Science* 125, 4173–4178.
- Steir, A.J., de Lange, T., 2012. Removal of shelterin reveals the telomere end-protection problem. *Science* 336, 593–597.
- Steir, A.J., Kosiyatrakul, S.T., Hockemeyer, D., *et al.*, 2009. Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. *Cell* 138, 90–103.
- Stewart, J.A., Chaiken, M.F., Wang, F., Price, C.M., 2012. Maintaining the end: Roles of telomere proteins in end-protection, telomere replication and length regulation. *Mutation Research* 730, 12–19.
- Stewart, J.A., Wang, F., Chaiken, M.F., *et al.*, 2012. Human CST promotes telomere duplex replication and general replication restart after fork stalling. *EMBO Journal* 31, 3537–3549.
- Verdun, R.E., Karlseder, J., 2006. The DNA damage machinery and homologous recombination pathway act consecutively to protect human telomeres. *Cell* 127, 709–720.
- Wellinger, R.J., Zakian, V.A., 2012. Everything you ever wanted to know about *Saccharomyces cerevisiae* telomeres: Beginning to end. *Genetics* 191, 1073–1105.
- Wu, P., Takai, H., de Lange, T., 2012. Telomeric 3' overhangs derive from resection by Exo1 and Apollo and fill-in by POT1b-associated CST. *Cell* 150, 39–52.
- Xu, L., Li, S., Stohr, B.A., 2013. The role of telomere biology in cancer. *Annual Review of Pathology* 8, 49–78.
- Ye, J., Lenain, C., Bauwens, S., *et al.*, 2010. TRF2 and Apollo cooperate with topoisomerase 2alpha to protect human telomeres from replicative damage. *Cell* 142, 230–242.
- Zakian, V.A., 2012. Telomeres: The beginnings and ends of eukaryotic chromosomes. *Experimental Cell Research* 318, 1456–1460.

Eukaryotic Nucleotide Excision Repair

B Van Houten and M Kong, University of Pittsburgh, Pittsburgh, PA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

DNA polymerase A protein or protein complex which synthesizes a daughter strand of DNA using a parental strand as a template.

Helicase A protein which actively separates the two strands of DNA using the energy of adenosine triphosphate (ATP) hydrolysis.

Nucleotide excision repair (NER) A highly conserved multistep process in which several protein machines identify and remove bulky damage from DNA using a dual incision mechanism.

Perspective

Deoxyribonucleic acid (DNA) repair mechanisms have co-evolved with life, and maintaining genome integrity is essential for the health and fitness of any organism. One of the most highly conserved genome maintenance processes is nucleotide excision repair (NER). This repair pathway is found in all kingdoms of life and is critical for the removal of a wide variety of DNA lesions produced by environmental insult, including: ultraviolet (UV)-induced photoproducts, polycyclic aromatic hydrocarbons found in air pollution, toxins produced by fungi and plants, and several anticancer agents such as cisplatin.

NER was discovered in 1964 simultaneously in four laboratories under the direction of Richard Setlow, Paul Howard-Flanders, Robert Painter, and Philip Hanawalt (Boyce and Howard-Flanders, 1964; Kusumoto *et al.*, 2001; Pettijohn and Hanawalt, 1964; Setlow and Carrier, 1964). The first two studies showed the removal of UV-induced cyclobutane pyrimidine dimers (CPD) from bacterial and last two studies showed the incorporation of new nucleotides into the repairing sites in mammalian and bacterial cells, respectively. NER can be initiated in two general ways: (1) global genome repair (GGR), in which damage recognition proteins scan the entire genome for damaged sites and (2) transcription-coupled repair (TCR), in which RNA polymerase (RNAP) that has been stalled by a lesion in the transcribed strand recruits a DNA translocase, which functions to simultaneously backtrack RNAP from the damaged site and recruit DNA repair enzymes. This article explores the process of NER in human cells.

Mechanistically, NER occurs in four highly integrated steps (Figure 1):

1. Damage recognition, allows the initial identification of specific lesions in the context of chromatin.
2. Damage verification, in which the repair machinery assembles at sites of damage and uses ATP-driven processes to verify the damaged nucleotide.
3. Assembly of the incision complex, involving large conformational changes in both the DNA and protein complexes.
4. Excision, repair synthesis, and ligation, in which nucleases cleave two phosphodiester bonds both 5' and 3' to the

damaged nucleotide, releasing an oligonucleotide containing the damage during repair synthesis by DNA polymerases. DNA ligase then seals the resulting nick made at the repair patch.

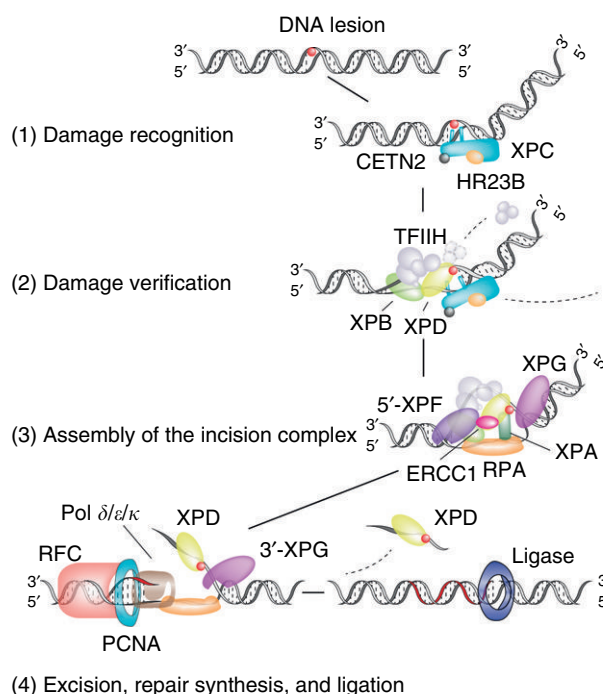


Figure 1 Mechanism of global genome nucleotide excision repair. This molecular process can be viewed in four steps: (1) 'Damage recognition' occurs when XPC (blue)-HR23B (orange)-Centrin-2 (dark gray) bind to a lesion through XPC's beta-hairpins causing bending of the DNA; (2) 'Damage verification' is mediated by the recruitment of XPD (light green) and XPB (dark green) as part of TFIIH (gray), which loses the three protein CAK subunit (gray); (3) Once the DNA surrounding the lesion is opened the incision proteins, XPF-ERCC1 (dark purple-magenta) makes the 5' incision; and (4) DNA polymerase with the help of RFC (light red cylinder) and PCNA (blue ring) begin 'repair synthesis,' which signals to XPG to make the 3' incision and allow the 'excision' of the damage containing oligonucleotide and XPD. DNA 'ligase' (dark purple ring) seals the newly created repair patch.

As discussed more at the end of this article, deficiencies in several of these key proteins cause several human diseases characterized by increased sensitivity to sun light, premature aging, neurodegeneration, and in some cases highly elevated skin cancers. The entire process of human NER was reconstituted first from cell extracts and then from purified proteins in three different laboratories under the direction of Richard Wood, Aziz Sancar, and Jean-Marc Egly (Aboussekhra *et al.*, 1995; Araujo *et al.*, 2000; Mu *et al.*, 1995). Recent work by a number of laboratories has provided biochemical and structural insights into the nature of the proteins that mediate this complex process, which is reviewed in more detail (Scharer, 2013).

Damage Recognition

UV-DNA Damage Binding (DDB)

Photoproducts produced by sunlight are such ubiquitous DNA lesions, human cells have developed a specialized damage recognition complex, UV-DDB factor, which is a heterodimer consisting of DDB1 and DDB2 (encoded by xeroderma pigmentosum complementation group E (XPE) gene). DDB1 forms a complex with Culin4A and Rbx and is an ubiquitin-E3 ligase. UV-DDB has high affinity for UV-induced CPD, 6-4 photoproducts, and abasic sites; however, the physiological relevance of the latter lesion is not clear. UV-DDB initiates GGR and helps identify photoproducts in chromatin by adding ubiquitin moieties to histones, UV-DDB, and xeroderma pigmentosum complementation group C protein (XPC). It is believed that histone ubiquitylation facilitates nucleosome eviction providing access by the second GGR damage recognition protein, XPC-HR23B (Figure 2(a)). UV-DDB has been crystallized with DNA damage (Scrima *et al.*, 2008) and recent co-crystals with full-length human protein suggest that UV-DDB forms a dimer of dimers (DDB1-DDB2)₂ at damaged sites on two DNA molecules, in which the FQH domain of one of the DDB2 molecule probes the DNA for the damaged site, while the N-terminal domain of the other DDB2 contacts the DNA away from the lesion to help buttress the extremely stable interactions (Yeh *et al.*, 2012), Figure 3(a). Electron microscopy and atomic force microscopy have confirmed this mode of binding to DNA damage (Yeh *et al.*, 2012). This dimer is apparently remarkably stable on DNA, and it is believed that ubiquitylation is essential to allow UV-DDB to dissociate from damaged sites. However, multi-ubiquitylation steps leads to UV-DDB degradation and would thus prevent subsequent damage recognition. Recent data suggest that poly-ADP-ribose polymerase (PARP1) can PAR-ylate UV-DDB and allow it to cycle through multiple rounds of damage recognition (Pines *et al.*, 2012).

XPC-HR23B-CETN2

Damage recognition in genomic DNA is primarily achieved by a trimeric protein complex consisting of XPC, human Rad23B (HR23B), and centrin-2 (CETN2). It has been suggested that HR23B dissociates from XPC after damage recognition occurs (Bergink *et al.*, 2012). XPC-HR23B binds to a wide variety of

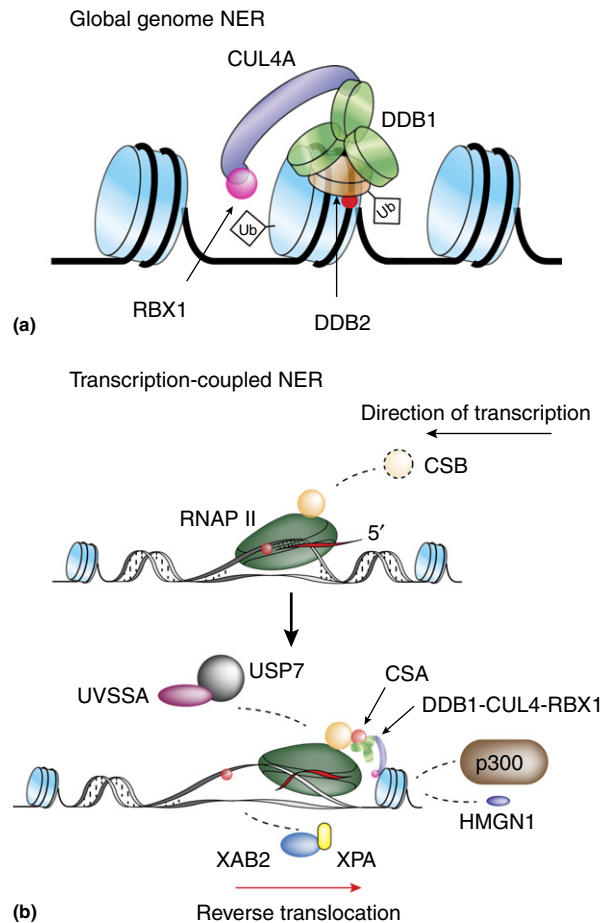


Figure 2 Initiation of NER at UV-induced photoproducts by UV-DDB or TCR. (a) Initiation of GG-NER on chromatinized DNA is achieved through lesion recognition by the UV-DDB heterodimer (DDB1 (green) and DDB2 (orange)). The E3 ubiquitin ligase CRL4DDB2, consisting of the UV-DDB heterodimer, Cul4A (purple), and Rbx1 (magenta), then ubiquitylates core histone as well as DDB2, leading to eviction of nucleosomes. (b) Transcription-coupled NER is initiated when transcribing RNAP (dark green) is stalled at the site of lesion and recruits CSB (orange). Binding of CSB recruits E3 ubiquitin ligase CRL4CSA, consisting of DDB1, CUL4, RBX1, and CSA (red). Stalled RNAP is then able to backtrack with the help of remodeling and sliding of upstream nucleosomes, facilitated by p300 (brown) and HMGN1 (purple). XAB2 (navy) in complex with XPA (yellow) are then recruited to site of lesion. Simultaneously, UVSSA (maroon) and USP7 (black) form complex at stalled RNAP, protecting it from degradation.

DNA lesions that differ in structure and conformation, and can even bind to helical distortions created by several mismatched bases (Kusumoto *et al.*, 2001; Sugasawa *et al.*, 2009; Sugasawa *et al.*, 2001). Further analysis showed that the ability of XPC-HR23B to open the DNA around the lesion correlated well with overall repair efficiency (Mocquet *et al.*, 2007). Structural analysis of the yeast XPC homolog, Rad4–Rad23 (see Table 1) showed that damage recognition is achieved by insertion of a beta-hairpin at the site of the damage (Figure 3(c)). However, the damage is flipped out of the helix and is disordered in the crystal structure (Min and Pavletich, 2007), indicating that XPC-HR23B makes large contacts with the non-damaged

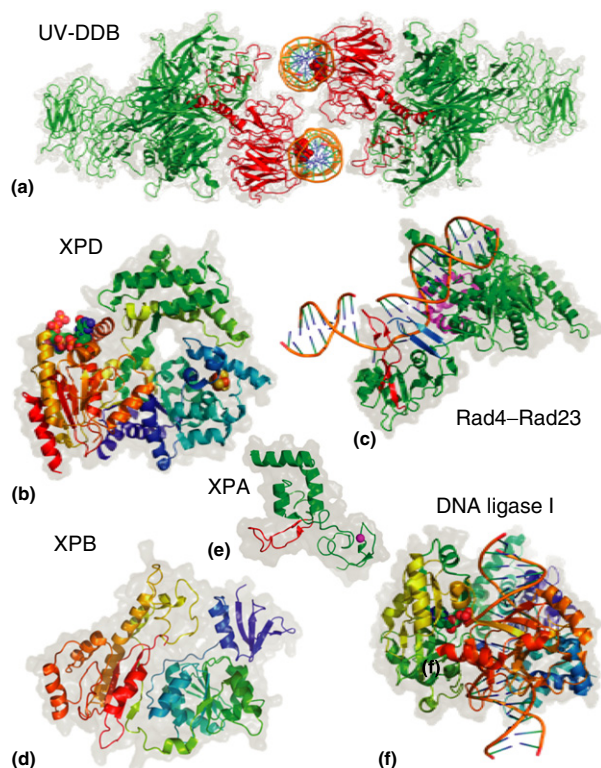


Figure 3 Structural motifs involved in damage recognition and verification. Carbon backbone cartoon in various colors and space filling outline in gray are given for each structure. (a) Human UV-DDB (PDB ID: 4E5Z) consists of DDB1 (green) and DDB2 (red), the damage probing phenylalanine, glutamine, and histone wedge are shown in contact with the DNA as space-filled red spheres. (b) XPD (PDB ID: 4A15) from *Thermoplasma acidophilum* bound to a four-base oligonucleotide (shown in space filling); the iron sulfur center (space filled, brown and yellow) is shown in the blue domain; the arch domain is shown in green. (c) Rad4-Rad23 (PDB ID: 2QSH), the *Saccharomyces cerevisiae* homolog of XPC-HR23B is shown in green and magenta, respectively. The two beta-hairpins that make contact at and near the damaged site are shown in red and blue, respectively. DNA is shown in gold and blue. (d) XPB (PDB ID: 2FWR), from *Archaeoglobus fulgidus*. (e) Human XPA (PDB ID: 1XPA), structure of the core domain solved by NMR showing the zinc finger (zinc in magenta) and the presumed DNA binding loop XPA (red). (f) Human DNA ligase I (PDB ID: 1X9N) shown bound to DNA (gold and blue).

strand and does not make direct contact with the damaged bases. The fact that a beta-hairpin is used to interrogate the DNA by probing into the DNA nicely explains why several mismatched bases are sufficient for recognition by XPC-HR23B. Centrin-2, a EF-hand calcium binding protein, forms a complex with XPC-HR23B and has been shown to stimulate damage binding and repair (Krasikova *et al.*, 2012; Nishi *et al.*, 2005). Further analysis of Centrin-2 suggests that it may mediate the interaction of XPC-HR23B with xeroderma pigmentosum group A protein (XPA) (Nishi *et al.*, 2013).

Transcription-Coupled Repair

Many lesions which are acted on by NER also stall RNAP (Hanawalt and Spivak, 2008). Stalled RNAP represents a

critical problem to the DNA transaction machinery, as it can cause loss of critical gene expression, trigger apoptosis, and create replication blocks. Nature has dealt with this problem in a resourceful way by directing key motor proteins to push RNAP away from the lesion and simultaneously recruit the DNA repair machinery. Several key factors in human cells are involved in moving RNAP upstream from the damaged site. Cockayne syndrome A and B (CSA and CSB) proteins are recruited to the stalled RNAP. CSB is a chromatin remodeling protein with a helicase-like fold, and it interacts with CSA which forms a complex with DDB1-culin4A-RBX1 (CRL4), which can ubiquitylate both itself and CSB. A recently discovered protein UV-stimulated scaffold protein A (UVSSA), which has been found to be mutated in certain UV-sensitive humans (discussed below), forms a complex with ubiquitin-specific protease 7 (USP7), a deubiquitylating enzyme (Figure 2(b)). Together the CSB-CRL4^{CSA} complex helps recruit the XPA-binding protein XAB2, which is an essential protein for pre-mRNA (messenger RNA) splicing, to the stalled RNAP. XAB2 participates in repair by attracting XPA to the damaged site and may be necessary for RNAP to restart transcription after repair. Once XPA and TFIIF are recruited to the damaged site, repair proceeds as shown in Figure 1.

Damage Verification

Prior to making the commitment to incise and remove the DNA lesion and surrounding nucleotides, the cell works to process the damaged site into a new conformation in an energy-dependent step, involving ATP hydrolysis, through a kinetic proof-reading mechanism (Reardon and Sancar, 2004). This processing is mediated by the dual action of transcription-factor IIH (TFIIH) and XPA-RPA complex.

TFIIH

TFIIH is a multi-subunit protein complex consisting of two DNA helicases, xeroderma pigmentosum complementation group B protein (XPB) (Figure 3(d)), and xeroderma pigmentosum complementation group D protein (XPD) (Figure 3(b)); and a kinase domain (CAK), consisting of cyclin H (CCNH), cyclin-dependent kinase 7 (CDK7) and ménage a trois 1 (MNAT1) proteins, and five other scaffold proteins (Table 1). TFIIH is an interesting protein machine that has several roles in the nucleus (Svejstrup *et al.*, 1996). First, it functions in transcription by opening the promoter to allow RNAP initiation. Second, TFIIH functions in NER, by helping to open up the DNA around the lesion to allow damage verification (Oksenyich *et al.*, 2009). Finally, TFIIH, through its kinase domain, CAK, is directly involved in cell cycle regulation. The two helicases within TFIIH play different roles in each process. The 3' → 5' helicase activity of XPB is essential for transcriptional initiation, whereas XPB's ATPase and XPD's 5' → 3' helicase activity are essential for DNA damage verification and subsequent repair. TFIIH unwinds the DNA about 20 nucleotides around the DNA lesions. It has been shown that XPD makes direct contact at the damaged site and its helicase activity is inhibited by damage (Buechner *et al.*, 2014;

Table 1 Human and yeast nucleotide excision repair (NER) proteins

Human			Yeast			Function
Gene	AA	MW (kD)	Gene	AA	MW (kD)	
CCNH (cyclin H)	323	38	CCL1	393	45	Kinase subunit of TFIIH
CDK7 (cyclin-dependent kinase 7)	346	39	KIN28	306	35	Kinase subunit of TFIIH
CSA (ERCC8)	396	44	RAD28	515	57	Interaction with Cockayne syndrome type B (CSB) protein
CSB (ERCC6)	1493	168	RAD26	1085	125	A member of the SWI2/SNF2 family of ATP-dependent chromatin remodeling factors TCR
CETN2 (centrin-2)	172	20	CDC31	161	18	Interaction with XPC causing conformational changes
DDB1	1140	127				DDB subunit, with CUL4-RBX1 forms a E3 platform and interacts with various WD repeats proteins
CUL4A	667	78				Cullin4A, E3 ligase
RBX1	108	12				Ring-box 1, E3 ubiquitin protein ligase
XPE (DDB2)	427	48	PRP4	465	52	DDB subunit, defective in XPE
ERCC1	323	36	RAD10	210	24	Forms complex with XPF
XPF (ERCC4)	916	107	RAD1	1100	126	5' incision nuclease
FBL3 (FBXL2)	423	47	RAD7	565	64	With RAD16 forms E3 ubiquitin ligase and damage binding
GTF2H1	548	62	TFB1	642	73	TFIIH subunit
GTF2H2	395	44	SSL1			TFIIH subunit
GTF2H3	308	34	TFB4	338	37	TFIIH subunit
GTF2H4	462	52	TFB2	513	59	TFIIH subunit
GTF2H5 (TTDA)	71	8	TFB5	72	8	TFIIH subunit
LIG1	919	102	CDC9	755	84	DNA ligase
MMS19L (MMS19)	1030	113	MET18	1032	118	Required for transcription and NER
MNAT1	309	36	TFB3			TFIIH subunit
			RAD16	790	91	With RAD7 forms E3 ubiquitin ligase and damage binding
HR23A	363	40	RAD23	398	42	RAD23B paralog
HR23B	409	43	RAD23	398	42	Forms complex with XPC and binds distorted DNA
RPA1	616	68	RFA1	621	73	RPA subunit, binds ssDNA intermediates
RPA2	270	30	RFA2	273	30	Interaction with ssDNA intermediates
RPA3	121	14				Interaction with ssDNA intermediates
XAB2	1140	127	SYF1	859	100	Interaction with XPA
XPA	273	31	RAD14	371	43	Interaction with DNA and proteins of the pre-incision complex
XPB (ERCC3)	782	89	SSL2	843	95	3' to 5' helicase DNA helicase TFIIH subunit
XPC	940	106	RAD4	754	87	Interaction with distorted DNA as complex with RAD23B
XPB (ERCC2)	760	87	RAD3	778	90	5'-to-3' DNA helicase TFIIH subunit
XPG (ERCC5)	1186	133	RAD2	273	30	3' incision nuclease

Abbreviations: CSA, Cockayne syndrome A; CSB, Cockayne syndrome B; DNA, deoxyribonucleic acid; ERCC, excision repair cross-complementing; RPA, replication factor protein; ssDNA, single-stranded DNA; TCR, transcription-coupled repair; XPA, xeroderma pigmentosum group A protein; XPB, xeroderma pigmentosum complementation group B; XPC, xeroderma pigmentosum C protein; XPE, xeroderma pigmentosum E protein; XPF, xeroderma pigmentosum complementation group F protein; XPG, xeroderma pigmentosum complementation group G protein.

Mathieu *et al.*, 2013; Reardon and Sancar, 2002). It is thus believed that XPD provides damage verification by being directly inhibited by the DNA damage as it translocates in a 5'→3' direction on the damaged strand (Kuper *et al.*, 2012; Mathieu *et al.*, 2013). Recent data suggest that MNAT1 interacts directly with the arch domain of XPD turning off the helicase activity of XPD (Abdulrahman *et al.*, 2013). Thus, the CAK domain is a molecular switch that provides important

phosphorylation of RNAP and its release allows damage verification to proceed (Coin *et al.*, 2008).

XPA-RPA

XPA is a key protein in both GGR and TCR. XPA, while being a relatively small protein of 273 amino acids (Table 1), interacts

with many NER proteins, including: TFIIH, XPC-HR23B, DDB2, excision repair cross-complementing 1 (ERCC1), proliferation cell nuclear antigen (PCNA), and replication factor protein (RPA) (Scharer, 2013). RPA is a three-subunit protein which binds to single-strand DNA and apparently helps stabilize XPA binding to several DNA damaged substrates. The XPA-RPA complex has affinity for helically distorted DNA (Camenisch *et al.*, 2006; Missura *et al.*, 2001), and RPA is believed to make significant contact with the non-damaged strand (Mocquet *et al.*, 2008). XPA contains several interesting structural motifs including a zinc and a beta-hairpin; the latter is believed to be necessary for damage recognition, reviewed in Kuper and Kisker (2012); see Figure 3(e).

Assembly of the Incision Complex

TFIIH and XPA/RPA having verified the presence of a damaged nucleotide, serve as a scaffold for the recruitment of the two endonucleases, xeroderma pigmentosum complementation group F protein (XPF), and xeroderma pigmentosum complementation group G protein (XPG). Working in sequential fashion, XPF and XPG incise the damage strand facilitating the excision of the damage containing oligonucleotide.

XPF-ERCC1

XPF forms a heterodimeric complex with the ERCC1 protein. XPF-ERCC1 is a structure-specific endonuclease and is apparently recruited to the TFIIH verification complex through its interaction with XPA and RPA (Croteau *et al.*, 2008; Tripsianes *et al.*, 2007; Tsodikov *et al.*, 2007). It has been shown that mutation of two highly conserved residues in XPA (Asn-110 and Tyr-145) located in the XPA binding site of ERCC1 decreases overall NER activity, but does not decrease the nuclease activity of XPF-ERCC1 on model nuclease substrates consisting of single-strand/double-strand junctions (Orelli *et al.*, 2010). XPF-ERCC1's nuclease hydrolyzes the phosphodiester bond approximately 20 nucleotide 5' to the damaged site. XPF-ERCC1's nuclease activity is also necessary for the repair of interstrand DNA cross-links in a mechanism that is different from its role in NER (Clauson *et al.*, 2013). Since many anticancer drugs such as cisplatin are repaired by NER, it is interesting to note that high levels of ERCC1 mRNA, but not protein are correlated with worse prognosis in women with advanced-stage ovarian cancer (Deloia *et al.*, 2012).

XPG

XPG is the 3' nuclease and shares homology with the flap-endonuclease, FEN1 (Scharer, 2008). XPG is recruited to TFIIH, where its nuclease activity remains inactive. XPG is actually required for the recruitment and the stimulation of XPF-ERCC1's nuclease activity (Mocquet *et al.*, 2008). XPG is also a structure-specific nuclease and incises the approximately 5th phosphodiester bond 3' to the damaged site. However, XPG's incision activity is not stimulated until repair synthesis proceeds from the 5' nick generated by XPF-ERCC1 (Fagbemi *et al.*, 2011; Staresincic *et al.*, 2009).

Excision, Repair Synthesis, and Ligation

5' incision appears to be the signal for the recruitment of repair synthesis machinery which fills in the repair patch triggering XPG incision and oligonucleotide release and subsequent DNA ligation.

DNA Polymerases

Reconstitution experiments with purified proteins indicated that both DNA polymerases δ and ϵ in the presence of the clamp loader, replication factor C (RFC), and the polymerase accessory sliding clamp, PCNA were necessary to fill in the repair patch (Araujo *et al.*, 2000; Shivji *et al.*, 1995). Surprisingly, the error-prone translesion polymerase DNA polymerase κ has more recently been shown to participate in the gap-filling step of NER (Ogi and Lehmann, 2006). It has been suggested that 50% of the repair sites use DNA polymerases κ and δ to fill in the patch, while the other 50% require DNA polymerase ϵ (Ogi *et al.*, 2010). Cell cycle analysis suggests that DNA polymerase δ and κ are absolutely required in non-cycling cells, whereas DNA polymerase ϵ is used in cells that are rapidly proliferating. Excision of the oligonucleotide containing the damage is accompanied by the loss of TFIIH, and it appears that XPD stays bound to the excision product (Kemp *et al.*, 2012).

Ligation

DNA ligase I was first identified as the main ligase that is essential for sealing the newly synthesized repair patch (Figures 1 and 3(f)). However, recent studies suggest that in non-cycling cells DNA ligase III α and XRCC1 can help carry out the ligation step of NER (Moser *et al.*, 2007). These cell cycle effects have important implications in considering NER in the context of an entire organism in which many of the cells are terminally differentiated and nondividing.

Human Diseases

Because of the prevalence of endogenously and exogenously induced DNA damage, loss of several key NER proteins are the cause of a variety of rare human diseases (Hoeijmakers, 2001).

Cockayne Syndrome

Cockayne syndrome was first recognized in 1936 as patients with dwarfism, retinal atrophy, and deafness, and presents with a wide spectrum of symptoms that greatly differ in severity (Cockayne, 1936). These symptoms include cutaneous photosensitivity, cataracts, dental abnormalities, progressive neurodegeneration, mental retardation, developmental failure, deep sunken eyes, and progeroid appearance (Vermeulen and Foustieri, 2013). Mutations in several different genes can cause this syndrome; these include genes encoding CSA, CSB, XPB, XPD, and XPG proteins (Cleaver *et al.*, 2009).

Trichothiodystrophy

Trichothiodystrophy (TTD) was first described by Davies in 1968 as a rare autosomal syndrome in which patients exhibited sulfur-deficient brittle hair, scaly skin and mental, physical retardation, and in about 50% of the cases photosensitivity, without an elevation in skin cancer (Cleaver *et al.*, 2009; Pollitt *et al.*, 1968). Mutations in several genes encoding subunits of TFIIH, including the small 8 kDa subunit, XPB, and XPD can cause TTD. It has been suggested that this is a disease caused by dysfunction transcription (Bergmann and Egly, 2001; Coin *et al.*, 1998).

UV-Sensitive Syndrome

UV-Sensitive Syndrome (UVSS) was first observed in two Japanese siblings by Itoh *et al.* (1994). Patients with this syndrome display a three- to fourfold higher sensitivity to sunlight (Itoh *et al.*, 1994). Patients with this disorder display increased photosensitivity, telangiectasia, freckling, but no increase in skin tumors (Spivak, 2005). The genetics of UVSS are interesting in that mutations in CSA and CSB genes are responsible for some, but not all of the patients with this syndrome. Recently, the gene that causes UVSS-A was identified by exome sequencing and proteomics and was originally identified as KIAA1530, and is now called UVSSA gene (Fei and Chen, 2012; Nakazawa *et al.*, 2012; Schwertman *et al.*, 2012; Zhang *et al.*, 2012). The gene product interacts with CSA and CSB to help recruit USP7 to sites of TCR (Schwertman *et al.*, 2012). After damage UVSS-A binds to stalled RNAP, where it recruits USP7, which de-ubiquitinates CSB to help stabilize it at damage sites.

Xeroderma Pigmentosum

Xeroderma pigmentosum (XP) was first described by Kaposi and Ritter in 1870 and was characterized by flaky and hyper- and hypopigmentation of sun-exposed skins. XP patients show a remarkable sensitivity to sunlight with a 2000-fold increase in basal cell and squamous cell carcinomas (Cleaver *et al.*, 2009; Hoeijmakers, 2001). XP can be caused by mutations in one of seven genes encoding nucleotide excision repair proteins (XPA-G), or a mutation in a translesion DNA polymerase η that can bypass UV-induced pyrimidine dimers. Many of these XP patients (complementation groups A, B, D, and G) also show neurodegeneration (Cleaver *et al.*, 2009). It has been suggested that this neurodegeneration may be caused by the accumulation of rare forms of oxidative damage, cyclo-dA and cyclo-dG which are repaired by NER (Brooks *et al.*, 2000).

XFE (XPF-ERCC1) Progeria

Since ERCC1 and XPF help stabilize each other, as loss of one protein causes a concomitant loss of the other. Loss of either ERCC1 or XPF is associated with a premature aging phenotype (Niedernhofer *et al.*, 2006). Mice carrying a complete loss of ERCC1 show an inability to thrive, dwarfness, kyphosis, and neurological defects and died within 4 weeks of birth (Niedernhofer *et al.*, 2006). Hypomorphic mice expressing less than 20% of the normal levels of ERCC1 show a constellation

of phenotypes all of which are consistent with premature aging (Gregg *et al.*, 2011). It is believed that the aging defect is not due to loss of NER, but is due to the direct role of XPF and ERCC1 in the repair of interstrand cross-links (Clauson *et al.*, 2013). Taken together, these data suggest that endogenous metabolites such as lipid peroxidation can mediate interstrand cross-links which if not repaired can cause premature aging.

Acknowledgments

We would like to apologize to the many friends and colleagues whose work, in the interest of space, could not be included here. This work was supported by an NIH grant, R01ES019566.

See also: Nucleic Acid Synthesis/Breakdown: RNA Synthesis/Function: Comparison of Bacterial and Eukaryotic Replisome Components

References

- Abdulrahman, W., Iltis, I., Radu, L., *et al.*, 2013. ARCH domain of XPD, an anchoring platform for CAK that conditions TFIIH DNA repair and transcription activities. *Proceedings of the National Academy of Sciences of the United States of America* 110, E633–E642.
- Aboussekhra, A., Biggerstaff, M., Shivji, M.K., *et al.*, 1995. Mammalian DNA nucleotide excision repair reconstituted with purified protein components. *Cell* 80, 859–868.
- Araujo, S.J., Tirode, F., Coin, F., *et al.*, 2000. Nucleotide excision repair of DNA with recombinant human proteins: Definition of the minimal set of factors, active forms of TFIIH, and modulation by CAK. *Genes & Development* 14, 349–359.
- Bergink, S., Toussaint, W., Luijsterburg, M.S., *et al.*, 2012. Recognition of DNA damage by XPC coincides with disruption of the XPC-RAD23 complex. *Journal of Cell Biology* 196, 681–688.
- Bergmann, E., Egly, J.M., 2001. Trichothiodystrophy, a transcription syndrome. *Trends in Genetics* 17, 279–286.
- Boyce, R.P., Howard-Flanders, P., 1964. Release of ultraviolet light-induced thymine dimers from DNA in *E. coli* K-12. *Proceedings of the National Academy of Sciences of the United States of America* 51, 293–300.
- Brooks, P.J., Wise, D.S., Berry, D.A., *et al.*, 2000. The oxidative DNA lesion 8,5'-(S)-cyclo-2'-deoxyadenosine is repaired by the nucleotide excision repair pathway and blocks gene expression in mammalian cells. *Journal of Biological Chemistry* 275, 22355–22362.
- Buechner, C.N., Heil, K., Michels, G., *et al.*, 2014. Strand-specific recognition of DNA damages by XPD provides insights into nucleotide excision repair substrate versatility. *Journal of Biological Chemistry* 289, 3613–3624.
- Camenisch, U., Dip, R., Schumacher, S.B., Schuler, B., Naegeli, H., 2006. Recognition of helical kinks by xeroderma pigmentosum group A protein triggers DNA excision repair. *Nature Structural & Molecular Biology* 13, 278–284.
- Clauson, C., Scharer, O.D., Niedernhofer, L., 2013. Advances in understanding the complex mechanisms of DNA interstrand cross-link repair. *Cold Spring Harbor Perspectives in Biology* 5, a012732.
- Cleaver, J.E., Lam, E.T., Revet, I., 2009. Disorders of nucleotide excision repair: The genetic and molecular basis of heterogeneity. *Nature Reviews Genetics* 10, 756–768.
- Cockayne, A.E., 1936. Dwarfism with retinal atrophy and deafness. *Archives of Disease in Childhood* 11, 1–8.
- Coin, F., Marinoni, J.C., Rodolfo, C., *et al.*, 1998. Mutations in the XPD helicase gene result in XP and TTD phenotypes, preventing interaction between XPD and the p44 subunit of TFIIH. *Nature Genetics* 20, 184–188.
- Coin, F., Oksenyich, V., Mocquet, V., *et al.*, 2008. Nucleotide excision repair driven by the dissociation of CAK from TFIIH. *Molecular Cell* 31, 9–20.
- Croteau, D.L., Peng, Y., Van Houten, B., 2008. DNA repair gets physical: Mapping an XPA-binding site on ERCC1. *DNA Repair (Amst.)* 7, 819–826.

- Deloia, J.A., Bhagwat, N.R., Darcy, K.M., *et al.*, 2012. Comparison of ERCC1/XPF genetic variation, mRNA and protein levels in women with advanced stage ovarian cancer treated with intraperitoneal platinum. *Gynecologic Oncology* 126, 448–454.
- Fagbemi, A.F., Orelli, B., Schärer, O.D., 2011. Regulation of endonuclease activity in human nucleotide excision repair. *DNA Repair* 10, 722–729.
- Fei, J., Chen, J., 2012. KIAA1530 protein is recruited by Cockayne syndrome complementation group protein A (CSA) to participate in transcription-coupled repair (TCR). *Journal of Biological Chemistry* 287, 35118–35126.
- Gregg, S.Q., Robinson, A.R., Niedernhofer, L.J., 2011. Physiological consequences of defects in ERCC1-XPF DNA repair endonuclease. *DNA Repair* 10, 781–791.
- Hanawalt, P.C., Spivak, G., 2008. Transcription-coupled DNA repair: Two decades of progress and surprises. *Nature Reviews Molecular Cell Biology* 9, 958–970.
- Hoeijmakers, J.H., 2001. Genome maintenance mechanisms for preventing cancer. *Nature* 411, 366–374.
- Itoh, T., Ono, T., Yamaizumi, M., 1994. A new UV-sensitive syndrome not belonging to any complementation groups of xeroderma pigmentosum or Cockayne syndrome: Siblings showing biochemical characteristics of Cockayne syndrome without typical clinical manifestations. *Mutation Research* 314, 233–248.
- Kemp, M.G., Reardon, J.T., Lindsey-Boltz, L.A., Sancar, A., 2012. Mechanism of release and fate of excised oligonucleotides during nucleotide excision repair. *Journal of Biological Chemistry* 287, 22889–22899.
- Krasikova, Y.S., Rechkunova, N.I., Maltseva, E.A., *et al.*, 2012. Influence of centrin 2 on the interaction of nucleotide excision repair factors with damaged DNA. *Biochemistry Biokhimiia* 77, 346–353.
- Kuper, J., Kisker, C., 2012. Damage recognition in nucleotide excision DNA repair. *Current Opinion in Structural Biology* 22, 88–93.
- Kuper, J., Wolski, S.C., Michels, G., Kisker, C., 2012. Functional and structural studies of the nucleotide excision repair helicase XPD suggest a polarity for DNA translocation. *EMBO Journal* 31, 494–502.
- Kusumoto, R., Masutani, C., Sugawara, K., *et al.*, 2001. Diversity of the damage recognition step in the global genomic nucleotide excision repair *in vitro*. *Mutation Research* 485, 219–227.
- Mathieu, N., Kaczmarek, N., Ruthemann, P., Luch, A., Naegeli, H., 2013. DNA quality control by a lesion sensor pocket of the xeroderma pigmentosum group D helicase subunit of TFIIH. *Current Biology* 23, 204–212.
- Min, J.H., Pavletich, N.P., 2007. Recognition of DNA damage by the Rad4 nucleotide excision repair protein. *Nature* 449, 570–575.
- Missura, M., Buterin, T., Hindges, R., *et al.*, 2001. Double-check probing of DNA bending and unwinding by XPA-RPA: An architectural function in DNA repair. *EMBO Journal* 20, 3554–3564.
- Mocquet, V., Kropachev, K., Kolbanovskiy, M., *et al.*, 2007. The human DNA repair factor XPC-HR23B distinguishes stereoisomeric benzo[a]pyrenyl-DNA lesions. *EMBO Journal* 26, 2923–2932.
- Mocquet, V., Laine, J.P., Riedl, T., *et al.*, 2008. Sequential recruitment of the repair factors during NER: The role of XPG in initiating the resynthesis step. *EMBO Journal* 27, 155–167.
- Moser, J., Kool, H., Giakzidis, I., *et al.*, 2007. Sealing of chromosomal DNA nicks during nucleotide excision repair requires XRCC1 and DNA ligase III alpha in a cell-cycle-specific manner. *Molecular Cell* 27, 311–323.
- Mu, D., Park, C.H., Matsunaga, T., *et al.*, 1995. Reconstitution of human DNA repair excision nuclease in a highly defined system. *Journal of Biological Chemistry* 270, 2415–2418.
- Nakazawa, Y., Sasaki, K., Mitsutake, N., *et al.*, 2012. Mutations in UVSSA cause UV-sensitive syndrome and impair RNA polymerase I processing in transcription-coupled nucleotide-excision repair. *Nature Genetics* 44, 586–592.
- Niedernhofer, L.J., Garinis, G.A., Raams, A., *et al.*, 2006. A new progeroid syndrome reveals that genotoxic stress suppresses the somatotrophic axis. *Nature* 444, 1038–1043.
- Nishii, R., Okuda, Y., Watanabe, E., *et al.*, 2005. Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein. *Molecular and Cellular Biology* 25, 5664–5674.
- Nishi, R., Sakai, W., Tone, D., Hanaoka, F., Sugawara, K., 2013. Structure-function analysis of the EF-hand protein centrin-2 for its intracellular localization and nucleotide excision repair. *Nucleic Acids Research* 41, 6917–6929.
- Ogi, T., Lehmann, A.R., 2006. The Y-family DNA polymerase kappa (pol kappa) functions in mammalian nucleotide-excision repair. *Nature Cell Biology* 8, 640–642.
- Ogi, T., Limsirichaikul, S., Overmeer, R.M., *et al.*, 2010. Three DNA polymerases, recruited by different mechanisms, carry out NER repair synthesis in human cells. *Molecular Cell* 37, 714–727.
- Oksenyshyn, V., de Jesus, B.B., Zhovmer, A., Egly, J.M., Coin, F., 2009. Molecular insights into the recruitment of TFIIH to sites of DNA damage. *EMBO Journal* 28, 2971–2980.
- Orelli, B., McClendon, T.B., Tsodikov, O.V., *et al.*, 2010. The XPA-binding domain of ERCC1 is required for nucleotide excision repair but not other DNA repair pathways. *Journal of Biological Chemistry* 285, 3705–3712.
- Pettijohn, D., Hanawalt, P., 1964. Evidence for repair — Replication of ultraviolet damaged DNA in bacteria. *Journal of Molecular Biology* 93, 395–410.
- Pines, A., Vrouwe, M.G., Marteijn, J.A., *et al.*, 2012. PARP1 promotes nucleotide excision repair through DDB2 stabilization and recruitment of ALC1. *Journal of Cell Biology* 199, 235–249.
- Pollitt, R.J., Jenner, F.A., Davies, M., 1968. Sibs with mental and physical retardation and trichorrhexis nodosa with abnormal amino acid composition of the hair. *Archives of Disease in Childhood* 43, 211–216.
- Reardon, J.T., Sancar, A., 2002. Molecular anatomy of the human excision nuclease assembled at sites of DNA damage. *Molecular and Cellular Biology* 22, 5938–5945.
- Reardon, J.T., Sancar, A., 2004. Thermodynamic cooperativity and kinetic proofreading in DNA damage recognition and repair. *Cell Cycle* 3, 141–144.
- Scharer, O.D., 2008. XPG: Its products and biological roles. *Advances in Experimental Medicine and Biology* 637, 83–92.
- Scharer, O.D., 2013. Nucleotide excision repair in eukaryotes. *Cold Spring Harbor Perspectives in Biology* 5, a012609.
- Schwertman, P., Lagarou, A., Dekkers, D.H., *et al.*, 2012. UV-sensitive syndrome protein UVSSA recruits USP7 to regulate transcription-coupled repair. *Nature Genetics* 44, 598–602.
- Scriba, A., Konickova, R., Czyzewski, B.K., *et al.*, 2008. Structural basis of UV DNA-damage recognition by the DDB1-DDB2 complex. *Cell* 135, 1213–1223.
- Setlow, R.B., Carrier, W.L., 1964. The disappearance of thymine dimers from DNA: An error-correcting mechanism. *Proceedings of the National Academy of Sciences of the United States of America* 51, 226–231.
- Shivji, M.K., Podust, V.N., Hubscher, U., Wood, R.D., 1995. Nucleotide excision repair DNA synthesis by DNA polymerase epsilon in the presence of PCNA, RFC, and RPA. *Biochemistry* 34, 5011–5017.
- Spivak, G., 2005. UV-sensitive syndrome. *Mutation Research* 577, 162–169.
- Staresinic, L., Fagbemi, A.F., Enzlin, J.H., *et al.*, 2009. Coordination of dual incision and repair synthesis in human nucleotide excision repair. *EMBO Journal* 28, 1111–1120.
- Sugawara, K., Akagi, J., Nishi, R., Iwai, S., Hanaoka, F., 2009. Two-step recognition of DNA damage for mammalian nucleotide excision repair: Directional binding of the XPC complex and DNA strand scanning. *Molecular Cell* 36, 642–653.
- Sugawara, K., Okamoto, T., Shimizu, Y., *et al.*, 2001. A multistep damage recognition mechanism for global genomic nucleotide excision repair. *Genes & Development* 15, 507–521.
- Svejstrup, J.Q., Vichi, P., Egly, J.M., 1996. The multiple roles of transcription/repair factor TFIIH. *Trends in Biochemical Sciences* 21, 346–350.
- Tripsianes, K., Folkers, G.E., Zheng, C., *et al.*, 2007. Analysis of the XPA and ssDNA-binding surfaces on the central domain of human ERCC1 reveals evidence for subfunctionalization. *Nucleic Acids Research* 35, 5789–5798.
- Tsodikov, O.V., Ivanov, D., Orelli, B., *et al.*, 2007. Structural basis for the recruitment of ERCC1-XPF to nucleotide excision repair complexes by XPA. *EMBO Journal* 26, 4768–4776.
- Vermeulen, W., Foustier, M., 2013. Mammalian transcription-coupled excision repair. *Cold Spring Harbor Perspectives in Biology* 5, a012625.
- Yeh, J.I., Levine, A.S., Du, S., *et al.*, 2012. Damaged DNA induced UV-damaged DNA-binding protein (UV-DDB) dimerization and its roles in chromatinized DNA repair. *Proceedings of the National Academy of Sciences of the United States of America* 109, E2737–E2746.
- Zhang, X., Horibata, K., Saijo, M., *et al.*, 2012. Mutations in UVSSA cause UV-sensitive syndrome and destabilize ERCC6 in transcription-coupled DNA repair. *Nature Genetics* 44, 593–597.

Relevant Website

<https://dnapittcrew.upmc.com/>
DNA Repair Database.

The Base Excision Repair Pathway

JL Parsons and MJ Edmonds, University of Liverpool, Liverpool, UK

© 2016 Elsevier Inc. All rights reserved.

Glossary

Haploinsufficiency The presence of only one functional copy of a gene in a diploid organism which is not sufficient for performing its complete functions, leading to an impaired or diseased state.

HeLa cells An immortal cervical cancer cell line originally isolated from a single patient, Henrietta Lacks, and commonly used in cell biology research.

Mouse embryonic fibroblasts Immortalized cell lines derived from mouse embryos. In this context, they allow researchers to study cells from knockout mice with an embryonic lethal phenotype.

Processivity factor Accessory protein which enhances the activity of an enzyme. For DNA polymerases, they prevent the dissociation of the enzyme from the DNA strand it is extending.

Maintenance of Genome Stability and the Role of Base Excision Repair

The chemical structure of DNA in all cells is subject to attack from both exogenous sources, such as ionizing radiation and environmental toxins, as well as endogenous sources, including cellular reactive oxygen species produced as a by-product of normal metabolism. Indeed, human cells have been estimated to suffer greater than 10 000 DNA base damage events per cell per day (Lindahl, 1993). Without a mechanism for repairing such alterations to DNA, these DNA base lesions have the potential to cause mutations, leading to the disruption of essential proteins and subsequently promote the development of diseases such as premature aging, metabolic diseases, neurodegenerative diseases, and cancer. The base excision repair (BER) pathway has evolved to correct a plethora of different small DNA base damage lesions through a coordinated process of several steps involving a specific enzyme, or class of enzymes, whose goal is to restore genome integrity and prevent disease development (Table 1). It is present in evolutionarily diverse organisms, from bacteria and yeast through to humans. Indeed, some of the BER enzymes were first discovered in *Escherichia coli* (Verly and Paquette, 1972; Lindahl, 1974). The individual steps and the enzymes involved are now well defined, and these can be broadly divided into: (1) DNA base damage recognition and removal,

(2) abasic site incision and processing of DNA strand break ends, (3) insertion of the correct undamaged nucleotide, and (4) DNA nick ligation. These will be described sequentially in more detail below.

Basic Overview of the BER Pathway

BER repairs small lesions to DNA, which fall into several categories (Figure 1). Broadly, these are: (1) oxidative DNA base damage caused by reactive oxygen species produced by cellular metabolism and ionizing radiation, (2) spontaneous hydrolysis of the glycosidic bond connecting the base with the phosphodiester backbone, a process called depurination, (3) deamination of bases (i.e., removal of NH_2 group), such as deamination of cytosine to uracil, although this cannot occur on thymine residues as they do not have an amine group, and (4) DNA base alkylation. Some of these DNA base lesions are known to be mutagenic, demonstrating the importance of their efficient removal from genomic DNA (Bjelland and Seeberg, 2003; Cooke *et al.*, 2003). For example, 8-oxoguanine can form base pairs equally well with cytosine, adenine, or thymine; depurination leaves a gap in the base sequence termed an abasic or apurinic/apyrimidinic (AP) site which can stall transcription; and deamination of cytosine to uracil can result in a change on the complementary strand from a guanine to an adenine base during DNA replication.

The initial step of BER is damage recognition by a base damage-specific DNA glycosylase enzyme. Once a damaged base is identified, the DNA glycosylase cleaves the *N*-glycosidic bond connecting the base to the phosphodiester backbone. This predominantly creates an AP site, where the base is missing, although a break in the DNA strand can also be generated depending on the DNA glycosylase involved. This AP site is usually recognized by AP endonuclease 1 (APE1) which incises the phosphodiester backbone to create a DNA single-strand break. Alternatively, depending on the type of DNA ends generated through specific DNA glycosylases, these ends require processing by APE1 itself, or by polynucleotide kinase phosphatase (PNKP). DNA polymerase β then fills the gap with the correct, undamaged nucleotide, using the base on the opposing strand as a template, and further processes the

Table 1 Accession numbers for the major human proteins in base excision repair in the NCBI and UniProt databases

Protein	Accession no.
AP endonuclease 1	P27695
DNA glycosylases	See Table 2
DNA ligase I	P18858
DNA ligase III α	P49916
DNA polymerase β	P06746
DNA polymerase δ	P28340
DNA polymerase ϵ	Q07864
Flap endonuclease 1	P39748
Polynucleotide kinase phosphatase	Q96T60
Proliferating cell nuclear antigen	P12004
X-ray cross complementing 1	P18887

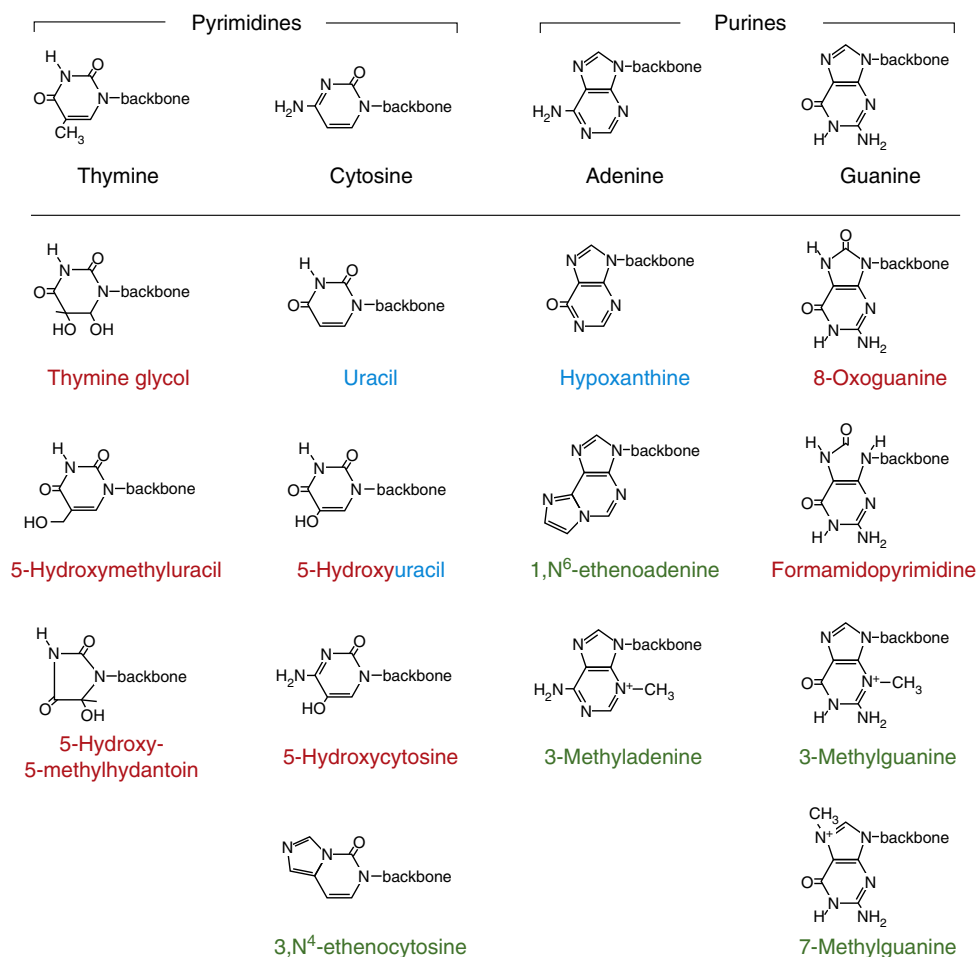


Figure 1 Examples of small DNA base lesions. The chemical structures of the four undamaged DNA bases are presented, and their damaged derivatives which are considered in this article are listed below. The damaged bases are color-coded according to the type of damage suffered: oxidative damage (red), deamination (blue), or alkylation (green).

phosphodiester backbone. The remaining nick in the phosphodiester backbone is sealed by DNA ligase III α in complex with the scaffold protein X-ray cross-complementing protein 1 (XRCC1). This completes the pathway known as short-patch BER, in which a single nucleotide is removed and replaced (Figure 2). This is the main mechanism through which the majority (>80%) of DNA base damage is repaired (Dianov *et al.*, 1992; Dianov and Parsons, 2007). A small proportion of DNA base damage is repaired through the long-patch BER pathway, which will be discussed in more detail later in this article.

DNA Base Damage Recognition and Removal by DNA Glycosylases

The first step of BER is initiated by the excision of damaged bases from DNA by DNA glycosylases. Humans have 11 known DNA glycosylases, each specific to a particular class of base damage or a specific lesion (Table 2). These enzymes can be divided into four distinct families: (1) uracil DNA glycosylases (UNGs), (2) helix-hairpin-helix (HhH) glycosylases,

(3) endonuclease VIII-like (NEIL) glycosylases, and (4) methylpurine DNA glycosylases (Jacobs and Schar, 2012; Brooks *et al.*, 2013; Wallace, 2013). The DNA glycosylases can be further subdivided based upon their catalytic mechanism, either monofunctional (base excision only, leaving an intact phosphodiester bond) or bifunctional (base excision and phosphodiester bond cleavage). Nevertheless, these enzymes all share a common mechanism to assess whether a base is correct and undamaged, or alternatively damaged, whereby the DNA glycosylase runs along the DNA backbone assessing each base pair in turn. If a damaged DNA base is detected, this is excised by the DNA glycosylase using a so-called base-flipping mechanism. This involves rotating the base around the phosphodiester backbone axis to 180° relative to its normal orientation. This moves the damaged base into the enzyme's active site, where this interaction is stabilized by specific amino acid interactions between the DNA glycosylase and the DNA. Subsequently the *N*-glycosidic bond linking the damaged base to the phosphodiester backbone is cleaved, creating an AP site (Figure 3). Bifunctional DNA glycosylases are able to further cleave the phosphodiester backbone to create a DNA single-strand break.

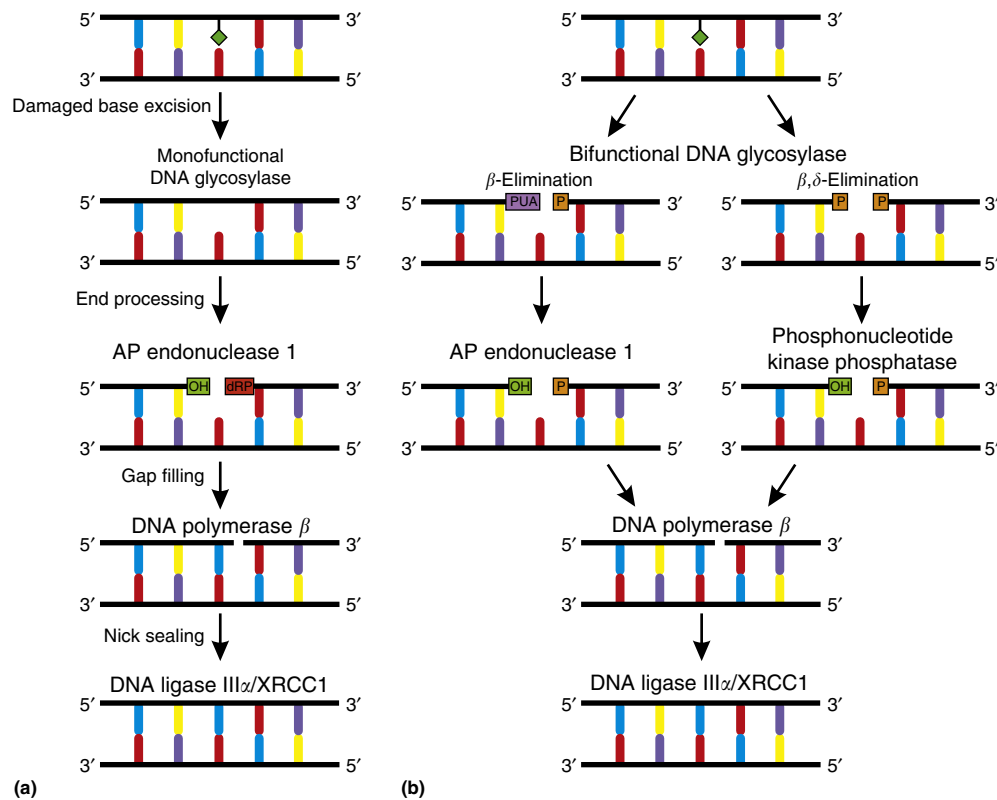


Figure 2 Overview of short-patch BER. DNA glycosylases excise the damaged base, creating different products depending on their mechanism of action. (a) Monofunctional glycosylases create an AP site which is recognized and cleaved by APE1 generating 3'-hydroxyl and 5'-deoxyribose phosphate ends. The 5'-deoxyribose phosphate moiety is subsequently removed by DNA polymerase β . (b) Bifunctional glycosylases perform either β -elimination (left pathway) or β,δ -elimination (right pathway) where a 3'-phospho- α,β -unsaturated aldehyde or 3'-phosphate, respectively are generated in addition to a 5'-phosphate. The 3'-phospho- α,β -unsaturated aldehyde is excised by APE1 whereas the 3'-phosphate is excised by PNKP. Following DNA strand break processing, in all cases the correct base is inserted by DNA polymerase β and the nick is sealed by the DNA ligase III α -XRCC1 complex. dRP, deoxyribose phosphate; PUA, 3' phospho- α,β -unsaturated aldehyde.

UNGs

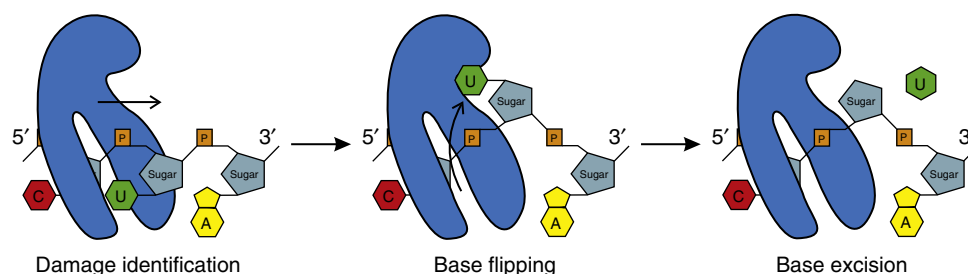
The major enzyme involved in the excision of uracil base damage from both single- and double-stranded DNA is UNG, which demonstrates a very high specificity for this type of base damage (Kavli *et al.*, 2002). Whilst the single-strand-selective monofunctional uracil-DNA glycosylase (SMUG1) suggests in its name that it has a preference for uracil in single-stranded DNA, in fact it can also excise the base damage from double-stranded DNA. Consequently, SMUG1 is thought to be a backup for UNG, although it also has distinct activity against derivatives of uracil base damage, such as 5-hydroxymethyluracil. Thymine DNA glycosylase (TDG) was originally discovered as an enzyme that can excise thymine from mismatches with guanine. However, its enzymatic activity is more efficient against uracil residues opposite guanine. In fact, TDG has a relatively broad substrate specificity, including 5-hydroxymethyluracil and 3,N⁴-ethenocytosine (Cortazar *et al.*, 2007). Methyl-CpG-binding domain protein 4, similar to TDG, is a mismatch-specific DNA glycosylase that is able to excise thymine, uracil, and 3,N⁴-ethenocytosine opposite guanine residues (Petronzelli *et al.*, 2000). All of the UNGs are monofunctional glycosylases, in that they only perform excision of the damaged base.

HhH DNA Glycosylases

This class of enzymes is characterized by the presence of a HhH structural motif contained within the protein. The 8-oxoguanine DNA glycosylase (OGG1), as the name suggests, is the major enzyme involved in the removal of 8-oxoguanine lesions from DNA when paired opposite cytosine (Boiteux and Radicella, 2000). This enzyme is important since adenine will preferentially pair opposite 8-oxoguanine, so in the next round of DNA replication the original guanine will be mutated to thymine (termed as G to T transversion). OGG1 can also excise formamidopyrimidine lesions from DNA. The importance of removing 8-oxoguanine residues from DNA is further highlighted by the existence of the MutY homologue (MYH) DNA glycosylase, which removes the adenine residues which have mispaired with 8-oxoguanine. This enzyme therefore allows for the subsequent replacement of adenine with the correct cytosine residue opposite 8-oxoguanine, in order that the latter can be excised by OGG1. Interestingly, MYH can also excise adenine residues when aberrantly base paired opposite guanine or cytosine. The endonuclease III homologue (NTH1) predominantly acts upon oxidized pyrimidine base damage, such as thymine glycol, 5-hydroxyuracil, and 5-hydroxycytosine (Aspinwall *et al.*, 1997). Both OGG1 and

Table 2 The human DNA glycosylases and their specificities

Name	Abbreviation	Accession no.	Escherichia coli orthologue	Classes		Major type of damage excised
8-Oxoguanine DNA glycosylase 1	OGG1	O15527	Fpg	HhH DNA glycosylases	Bifunctional (β -elimination)	8-Oxoguanine formamidopyrimidines
Endonuclease III-like 1	NTH1	P78549	Nth	HhH DNA glycosylases	Bifunctional (β -elimination)	Oxidative damage to pyrimidines, including thymine glycol, 5-hydroxyuracil, 5-hydroxycytosine
Endonuclease VIII-like 1	NEIL1	Q96F14	Nei	NEIL DNA glycosylases	Bifunctional (β , δ -elimination)	Oxidative damage, with preference for pyrimidines
Endonuclease VIII-like 2	NEIL2	Q969S2	Nei	NEIL DNA glycosylases	Bifunctional (β , δ -elimination)	Oxidized cytosine derivatives
Endonuclease VIII-like 3	NEIL3	Q8TAT5	Nei	NEIL DNA glycosylases	Bifunctional (β , δ -elimination)	Hydantoins; prefers single-stranded DNA structures
Methyl-CpG binding domain protein 4	MBD4	O95243	—	UNGs	Monofunctional	Mismatches opposite guanine
MutY homologue	MYH	Q9UIF7	MutY	HhH DNA glycosylases	Monofunctional	Mispaired adenine
N-methylpurine DNA glycosylase	MPG	P29372	AlkA	Methylpurine DNA glycosylase	Monofunctional	Alkylated bases
Single-strand-selective monofunctional uracil-DNA glycosylase	SMUG1	Q53HV7	—	UNGs	Monofunctional	Uracil and derivatives, preferably from single-stranded DNA
Thymine DNA glycosylase	TDG	Q13569	Mug	UNGs	Monofunctional	Thymine or uracil from mispairs with guanine
Uracil DNA glycosylase	UNG	P13051	Ung	UNGs	Monofunctional	Uracil residues

**Figure 3** Mechanism of action of damaged DNA base excision by DNA glycosylases. DNA glycosylases scan one strand of the double helix, probing each DNA base in turn. If they encounter a damaged base they can excise (in this case a uracil residue), this is flipped outside of the double helix to 180° relative to the phosphodiester backbone, and the *N*-glycosidic bond is cleaved. This releases the damaged base and creates an AP site.

NTH1 are bifunctional DNA glycosylases which carry out β -elimination, cleaving the phosphodiester backbone once (explained in Section NEIL DNA Glycosylases).

NEIL DNA Glycosylases

The NEIL DNA glycosylases consists of three family members, namely NEIL1, NEIL2, and NEIL3 (Wallace, 2013). When originally discovered, NEIL1 and NEIL2 were thought to act as backup repair enzymes to OGG1 and NTH1, since their substrates include oxidized pyrimidines (including thymine

glycol, 5-hydroxycytosine, and 5-hydroxyuracil), 8-oxoguanine, and formamidopyrimidine (Hazra *et al.*, 2002a, 2002b; Rosenquist *et al.*, 2003). However unlike OGG1 and NTH1, NEIL1 and NEIL2 can remove this type of base damage from single-stranded DNA and from bubble structures mimicking those that are formed during transcription or replication (Dou *et al.*, 2003), and NEIL1 can also excise DNA base damage in close proximity to another DNA lesion, which can be typically induced by ionizing radiation (Parsons *et al.*, 2005, 2007). NEIL3 also appears to remove oxidized pyrimidines and purines preferentially from single-stranded DNA and has been reported, in addition to NEIL1, to remove base damage from

quadruplex and telomeric DNA (Zhou *et al.*, 2013). This suggests a more specific and defined role for NEIL DNA glycosylases during the cell cycle. Importantly, these enzymes are bifunctional enzymes which carry out β,δ -elimination, cleaving the phosphodiester backbone twice (explained in Section Abasic Site Incision and Processing of DNA Strand Break Ends).

Methylpurine DNA Glycosylase

Methylpurine DNA glycosylase is specific for the removal of alkylated base damage from DNA. This includes the excision of 3-methyladenine, 7-methylguanine, and 3-methylguanine from both single- and double-stranded DNA (O'Connor, 1993). It is also capable of excising hypoxanthine and 1,N⁶-ethenoadenine.

Abasic Site Incision and Processing of DNA Strand Break Ends

Monofunctional DNA glycosylases only excise the damaged DNA base creating an AP site with an intact phosphodiester

backbone. The abasic site is subsequently recognized and processed by APE1, which hydrolyzes the phosphodiester bond between the phosphate moiety of the AP site and the deoxyribose moiety of the nucleotide 5' to the AP site (Figure 4(a); Demple *et al.*, 1991; Robson and Hickson, 1991). This creates a gap in the DNA strand which is flanked by a 3'-hydroxyl group and a 5'-deoxyribose phosphate, which is what remains of the original damaged nucleotide. This is important as the next step in the BER pathway, nucleotide incorporation by a DNA polymerase, requires a 3'-hydroxyl end so that the new, undamaged and correct nucleotide can be added.

Bifunctional DNA glycosylases possess an additional activity that cleaves the phosphodiester backbone creating a DNA strand break that is flanked by different DNA ends depending on the DNA glycosylase employed. Following bifunctional DNA glycosylase action, the 3'-end of the strand break requires further processing in order to generate the 3'-hydroxyl end necessary for downstream DNA polymerase action. OGG1 and NTH1 perform β -elimination, leaving the DNA strand break flanked by a 3'-phospho- α,β -unsaturated aldehyde and a 5'-phosphate. Therefore the 3'- α,β -unsaturated

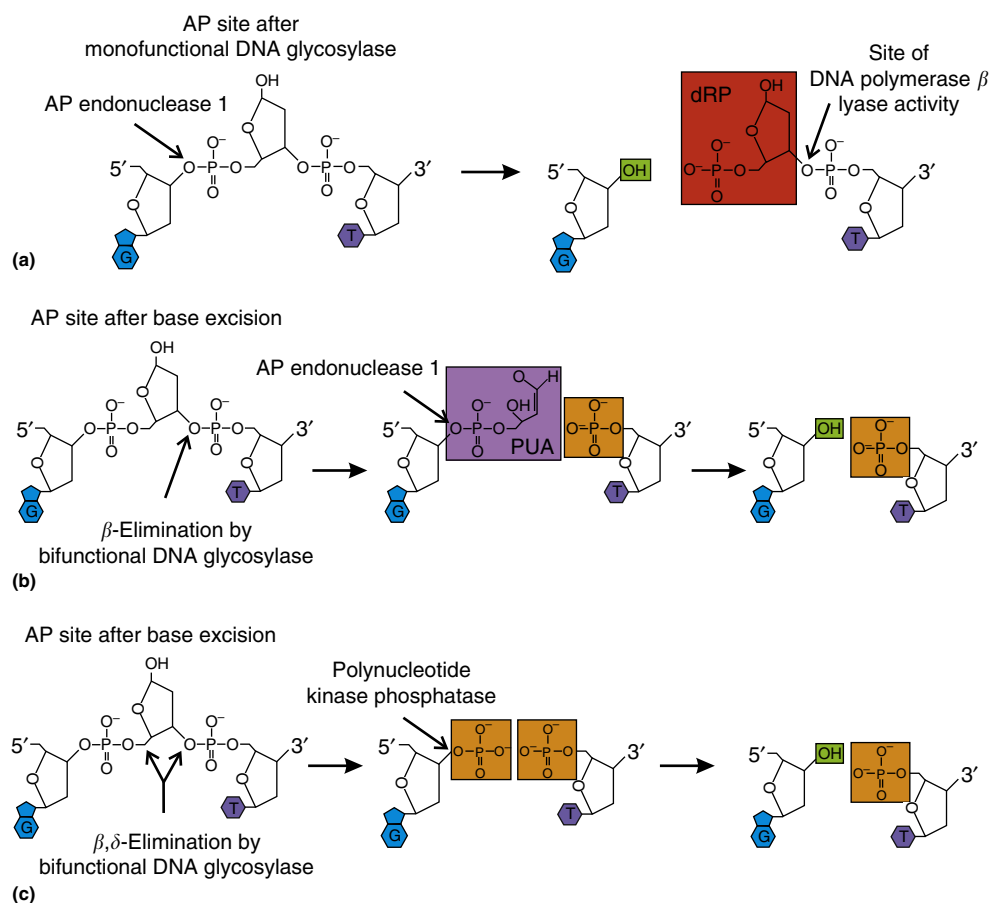


Figure 4 Monofunctional and bifunctional glycosylases generate different DNA ends. (a) Monofunctional DNA glycosylases only excise the damaged base, leaving an AP site with an intact phosphodiester backbone. This is then processed by APE1, which generates 3'-hydroxyl and 5'-deoxyribose phosphate ends, the latter of which is excised by the lyase activity of DNA polymerase β generating a 5'-phosphate end. (b) Bifunctional DNA glycosylases which perform β -elimination also cleave the phosphodiester backbone, in addition to damaged DNA base removal, leaving a 3'-phospho- α,β -unsaturated aldehyde and a 5' phosphate. APE1 processes the 3'-end to leave a 3'-hydroxyl group. (c) Bifunctional DNA glycosylases which perform β,δ -elimination generate 3'-phosphate and 5'-phosphate ends. PNKP subsequently processes the 3'-phosphate to generate a 3'-hydroxyl end.

aldehyde requires removal to generate a 3'-hydroxyl end, and in fact this is performed by the action of APE1 (Figure 4(b)). In contrast, the NEIL family of bifunctional DNA glycosylases performs β,δ -elimination which through two cleavages creates a DNA strand break flanked by a 3'-phosphate and a 5'-phosphate. The 3'-phosphate end requires processing to generate a 3'-hydroxyl end by a different enzyme, called PNKP (Figure 4(c); Wiederhold *et al.*, 2004).

Insertion of the Correct Undamaged Nucleotide

The end product of AP site incision and DNA strand break processing is the generation of a single nucleotide gap that contains a 3'-hydroxyl end which is a substrate for DNA polymerase. The major DNA polymerase employed during BER is DNA polymerase β (Sobol *et al.*, 1996). Although it is not a replicative DNA polymerase, its incorporation of a new nucleotide into the repair gap follows the same basic mechanism. Energy is released as pyrophosphate from the reaction between the incoming deoxyribonucleoside phosphate and the 3' hydroxyl group of the nucleotide immediately upstream. As with DNA replication, the fact that the genomic sequence is encoded on two strands is essential, as the undamaged strand opposite the repair site is used as a template to determine which nucleotide to insert.

Despite the successful replacement of the damaged nucleotide with the correct undamaged one, if BER was initiated by a monofunctional DNA glycosylase, the repair site still contains a 5'-deoxyribose phosphate moiety which cannot be ligated to the proximal 3'-hydroxyl end. In fact, in addition to nucleotide incorporation, DNA polymerase β also contains lyase activity which simultaneously removes the 5'-deoxyribose phosphate to generate the correct 5'-phosphate end (Figure 4(a); Matsumoto and Kim, 1995; Allinson *et al.*, 2001). BER which has proceeded through a pathway initiated by a bifunctional DNA glycosylase will already have this 5'-phosphate end, and so no further processing is required in this case (Figures 4(b) and 4(c)).

DNA Nick Ligation

Although the genetic sequence has now been corrected following DNA polymerase β action, there still remains a DNA single-strand break 3' to the inserted nucleotide. Therefore the phosphodiester backbone must be sealed in order to restore the genetic integrity of the DNA. This step is performed by a DNA ligase and is dependent on adenosine triphosphate (ATP), which activates the 5'-phosphate on the inserted nucleotide. This makes the reaction with the 3'-hydroxyl on the other side of the single-strand break energetically favorable, and so the nick is sealed with the release of adenosine monophosphate. The major DNA ligase employed during BER is DNA ligase III α , which is actually in a stable complex with the scaffold protein XRCC1 (Cappelli *et al.*, 1997; Nash *et al.*, 1997). The XRCC1-DNA ligase III α complex therefore binds the DNA strand break and ligates the nick in an ATP-dependent process.

As described previously, the BER process described above is commonly referred to as the short-patch BER pathway, since

the DNA base damage is replaced by a single, correct, undamaged nucleotide, and is the predominant mode of BER.

The Long-Patch BER Pathway

In some instances, particularly following the formation of reduced or oxidized abasic sites, the 5'-deoxyribose phosphate group produced by the action of APE1 is resistant to processing by the lyase activity of DNA polymerase β . Consequently the XRCC1-DNA ligase III α complex is then unable to ligate the nick containing this 5'-end blocking group. In this case an alternative pathway called long-patch BER is employed. Following incorporation of the correct, undamaged nucleotide by DNA polymerase β , there is then a polymerase switch to the replicative DNA polymerases δ or ϵ (Frosina *et al.*, 1996; Podlutzky *et al.*, 2001). These replicative polymerases then synthesize two to eight more nucleotides into the repair gap from the 3'-end, therefore displacing those already in position to form a 5'-flap structure (Figure 5). The processivity factor,

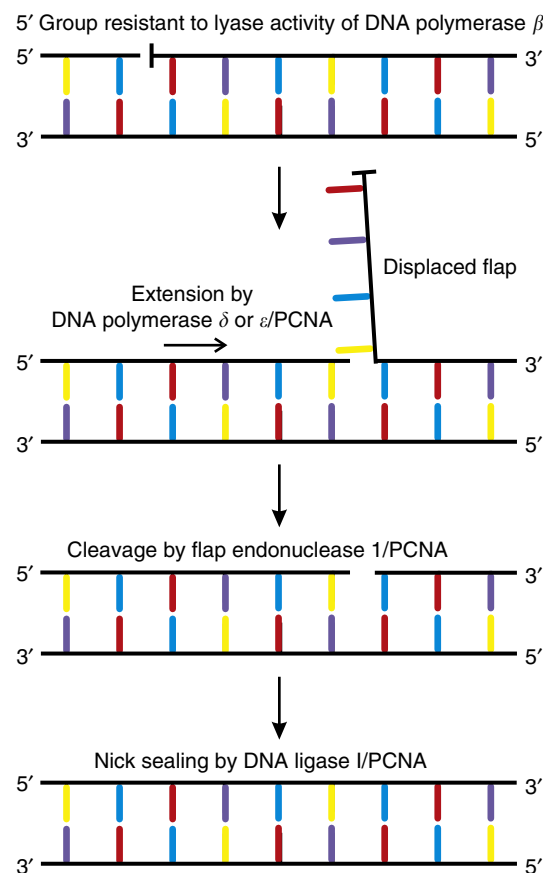


Figure 5 Overview of long-patch BER. If the lyase activity of DNA polymerase β is unable to process the 5' end once the correct nucleotide is inserted, long-patch BER is performed. This involves the addition of 2–8 more nucleotides into the repair gap by DNA polymerase δ or ϵ causing strand displacement and the formation of a 5'-flap structure. This structure is cleaved by flap endonuclease 1 (FEN1) before the nick is sealed by DNA ligase I. PCNA acts as a stabilizing factor throughout these steps.

proliferating cell nuclear antigen (PCNA), is additionally required for the correct function of DNA polymerase δ and ϵ . The 5'-flap structure produced during long-patch BER is consequently recognized and removed by the enzyme flap endonuclease (FEN1) to leave a single-strand break flanked by 3'-hydroxyl and 5'-phosphate ends (Klungland and Lindahl, 1997). This single-strand break is consequently sealed by DNA ligase I, in concert with PCNA, in an ATP-dependent process. This is in contrast to the XRCC1–DNA ligase III α complex that is specific to the short-patch BER pathway.

The Importance of Key BER Proteins is Highlighted by Knockout Mouse Models

Several mouse knockout models of BER proteins have been attempted in order to elucidate in detail their functions and roles in genome stability and the etiology of human diseases such as premature aging, metabolic diseases, neurodegenerative diseases, and cancer. A trend has emerged where deletion of the proteins which are essential for a particular step of BER causes embryonic lethality, demonstrating they are essential for organisms to develop and function at a cellular level. For example, mouse knockouts of APE1 (Xanthoudakis *et al.*, 1996), DNA polymerase β (Sobol *et al.*, 1996), DNA ligase III (Puebla-Orsorio *et al.*, 2006), XRCC1 (Tebbs *et al.*, 1999), and FEN1 (Kucherlapati *et al.*, 2002) are all embryonic lethal. Even in mouse haploinsufficiency models, containing only a single functional copy of the gene of interest, higher levels of spontaneous tumors and a high sensitivity to DNA damaging agents are seen in relation to APE1, DNA polymerase β , and XRCC1 (Cabelof, 2012). Mouse embryonic fibroblasts have been generated from BER-deficient embryos with a lethal phenotype and used for further studies to finely dissect the cellular effect of lacking these key BER proteins. These cells likewise usually display hallmarks of genomic instability. For example, both XRCC1 and DNA polymerase β -deficient mouse embryonic fibroblasts are hypersensitive to DNA damaging agents (Horton *et al.*, 2008). Cumulatively, this further underlines the crucial role that key BER proteins play in genome stability and in the prevention of human diseases, such as cancer.

Interestingly, knockout models of individual DNA glycosylases are not embryonic lethal (Parsons and Elder, 2003). This is not surprising given as there are multiple DNA glycosylases which can initiate BER, and there is a large degree of overlap in their substrate specificities (Table 2). However, some DNA glycosylase-knockout mice do show an increased propensity to DNA damage and indeed some develop specific phenotypes, indicating it is advantageous to possess each member of the family of DNA glycosylases. For example, OGG1-knockout mice demonstrate an increased level of 8-oxoguanine and formamidopyrimidine lesions but do not appear to be prone to a significant elevation in spontaneous tumor formation (Klungland *et al.*, 1999). Likewise, NTH1-deficient mice reveal no phenotypic abnormalities (Ocampo *et al.*, 2002; Takao *et al.*, 2002). In contrast, NEIL1-knockout mice develop an obese phenotype associated with a metabolic syndrome (Vartanian *et al.*, 2006), which was also recently reported in the OGG1-deficient mice (Sampath *et al.*, 2012), but do not develop spontaneous tumors. The theory of

overlapping DNA base damage substrate specificity being a cause for the lack of phenotype in DNA glycosylase-knockout mice is further strengthened by the isolation of NEIL1/NTH1 double knockout mice. These animals develop spontaneous pulmonary adenomas and carcinomas and hepatocellular carcinomas at a higher rate than in the corresponding single gene knockout mice (Chan *et al.*, 2009).

Regulation of BER Protein Levels

Consistent with the essential role of key BER proteins, such as XRCC1 and DNA polymerase β , in genome stability and normal development as displayed by embryonic lethality in knockout mouse models, BER proteins are usually very tightly regulated. One of the major mechanisms through which this can be achieved is by protein post-translational modifications. Indeed, phosphorylation and acetylation of BER proteins has been observed (Almeida and Sobol, 2007). However, more recently it has emerged that one of the major mechanisms of regulating cellular levels of BER proteins is through ubiquitylation (Parsons and Dianov, 2013; Edmonds and Parsons, 2014). Ubiquitylation involves the addition of the small protein ubiquitin (8 kDa), either singularly (monoubiquitylation) or through linked ubiquitin chains (polyubiquitylation), to specific lysine residues on target proteins. Ubiquitylation is, however, a reversible process with the addition and removal of ubiquitin catalyzed by different families of enzymes, E3 ubiquitin ligases and deubiquitylation enzymes, respectively. The major function of polyubiquitylation is in targeting proteins for degradation by the proteasome. Ubiquitylation of many BER proteins employed during the various stages of the pathway is increasingly being reported, although particularly significant progress has been made in our understanding of the role of ubiquitylation in regulating DNA polymerase β . Specifically, the E3 ubiquitin ligase Mule has been shown to monoubiquitylate DNA polymerase β which retains the protein in the cytoplasmic compartment of the cell (Parsons *et al.*, 2009), enabling another E3 ubiquitin ligase enzyme called CHIP to polyubiquitylate it and target the protein for degradation (Parsons *et al.*, 2008). In contrast, the deubiquitylating enzyme ubiquitin specific protease 47 (USP47) can remove these ubiquitin molecules and reverse their effects, therefore allowing DNA polymerase β protein levels to accumulate and enter the nucleus to assist in DNA repair during genotoxic stress conditions (Parsons *et al.*, 2011). This ensures that cells are able to immediately respond to the induction of DNA damage, and avoid genomic instability. Further post-translational modifications of BER proteins are increasingly being identified as a major mechanism for controlling BER protein cellular localization, activity, and cellular levels. These mechanisms are important in the maintenance of genome stability in the cellular response to DNA damage and, ultimately, in the survival of the organism.

Conclusions

- Maintenance of genome stability is an essential function of cell biology.

- Small base lesions to DNA are repaired by the BER pathway.
- BER is a tightly coordinated process involving the sequential action of a subset of enzymes from specific families.
- DNA glycosylases have distinct but overlapping specificities for different types of DNA base damage.
- Short-patch BER is the predominant mode, where only the single damaged nucleotide is replaced.
- Long-patch BER, involving the replacement of several nucleotides, may be employed when the DNA ends are resistant to DNA polymerase β processing.
- Disruption of the key enzymes involved in BER leads to embryonic lethality in mouse models and increased sensitivity to DNA damaging agents in deficient cells, correlating with an increased risk of development of diseases such as cancer.
- The cellular levels of BER proteins are regulated by post-translational modifications in response to DNA damage levels.

See also: Nucleic Acid Synthesis/Breakdown: DNA Synthesis/Repair: Eukaryotic Nucleotide Excision Repair

References

- Allinson, S.L., Dianova, I.I., Dianov, G.L., 2001. DNA polymerase beta is the major dRP lyase involved in repair of oxidative base lesions in DNA by mammalian cell extracts. *EMBO Journal* 20, 6919–6926.
- Almeida, K.H., Sobol, R.W., 2007. A unified view of base excision repair: Lesion-dependent protein complexes regulated by post-translational modification. *DNA Repair (Amst)* 6, 695–711.
- Aspinwall, R., Rothwell, D.G., Roldan-Arjona, T., *et al.*, 1997. Cloning and characterization of a functional human homolog of *Escherichia coli* endonuclease III. *Proceedings of the National Academy of Sciences of the United States of America* 94, 109–114.
- Bjelland, S., Seeberg, E., 2003. Mutagenicity, toxicity and repair of DNA base damage induced by oxidation. *Mutation Research* 531, 37–80.
- Boiteux, S., Radicella, J.P., 2000. The human OGG1 gene: Structure, functions, and its implication in the process of carcinogenesis. *Archives of Biochemistry and Biophysics* 377, 1–8.
- Brooks, S.C., Adhikary, S., Robinson, E.H., Eichman, B.F., 2013. Recent advances in the structural mechanisms of DNA glycosylases. *Biochimica et Biophysica Acta* 1834, 247–271.
- Cabelof, D.C., 2012. Haploinsufficiency in mouse models of DNA repair deficiency: Modifiers of penetrance. *Cellular and Molecular Life Sciences* 69, 727–740.
- Cappelli, E., Taylor, R., Cevasco, M., *et al.*, 1997. Involvement of XRCC1 and DNA ligase III gene products in DNA base excision repair. *Journal of Biological Chemistry* 272, 23970–23975.
- Chan, M.K., Ocampo-Hafalla, M.T., Vartanian, V., *et al.*, 2009. Targeted deletion of the genes encoding NTH1 and NEIL1 DNA N-glycosylases reveals the existence of novel carcinogenic oxidative damage to DNA. *DNA Repair (Amst)* 8, 786–794.
- Cooke, M.S., Evans, M.D., Dizdaroglu, M., Lunec, J., 2003. Oxidative DNA damage: Mechanisms, mutation, and disease. *FASEB Journal* 17, 1195–1214.
- Cortazar, D., Kunz, C., Saito, Y., Steinacher, R., Schar, P., 2007. The enigmatic thymine DNA glycosylase. *DNA Repair (Amst)* 6, 489–504.
- Demple, B., Herman, T., Chen, D.S., 1991. Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease: Definition of a family of DNA repair enzymes. *Proceedings of the National Academy of Sciences of the United States of America* 88, 11450–11454.
- Dianov, G., Price, A., Lindahl, T., 1992. Generation of single-nucleotide repair patches following excision of uracil residues from DNA. *Molecular and Cellular Biology*, 12, 1605–1612.
- Dianov, G.L., Parsons, J.L., 2007. Co-ordination of DNA single strand break repair. *DNA Repair (Amst)* 6, 454–460.
- Dou, H., Mitra, S., Hazra, T.K., 2003. Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. *Journal of Biological Chemistry* 278, 49679–49684.
- Edmonds, M.J., Parsons, J.L., 2014. Regulation of base excision repair proteins by ubiquitylation. *Experimental Cell Research* 329 (1), 132–138.
- Frosina, G., Fortini, P., Rossi, O., *et al.*, 1996. Two pathways for base excision repair in mammalian cells. *Journal of Biological Chemistry* 271, 9573–9578.
- Hazra, T.K., Izumi, T., Boldogh, I., *et al.*, 2002a. Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proceedings of the National Academy of Sciences of the United States of America* 99, 3523–3528.
- Hazra, T.K., Kow, Y.W., Hatahet, Z., *et al.*, 2002b. Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. *Journal of Biological Chemistry* 277, 30417–30420.
- Horton, J.K., Watson, M., Stefanick, D.F., *et al.*, 2008. XRCC1 and DNA polymerase beta in cellular protection against cytotoxic DNA single-strand breaks. *Cell Research* 18, 48–63.
- Jacobs, A.L., Schar, P., 2012. DNA glycosylases: In DNA repair and beyond. *Chromosoma* 121, 1–20.
- Kavli, B., Sundheim, O., Akbari, M., *et al.*, 2002. hUNG2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hSMUG1 as a broad specificity backup. *Journal of Biological Chemistry* 277, 39926–39936.
- Klungland, A., Lindahl, T., 1997. Second pathway for completion of human DNA base excision-repair: Reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO Journal* 16, 3341–3348.
- Klungland, A., Rosewell, I., Hollenbach, S., *et al.*, 1999. Accumulation of premutagenic DNA lesions in mice defective in removal of oxidative base damage. *Proceedings of the National Academy of Sciences of the United States of America* 96, 13300–13305.
- Kucherlapati, M., Yang, K., Kuraguchi, M., *et al.*, 2002. Haploinsufficiency of Flap endonuclease (Fen1) leads to rapid tumor progression. *Proceedings of the National Academy of Sciences of the United States of America* 99, 9924–9929.
- Lindahl, T., 1974. An N-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues. *Proceedings of the National Academy of Sciences of the United States of America* 71, 3649–3653.
- Lindahl, T., 1993. Instability and decay of the primary structure of DNA. *Nature* 362, 709–715.
- Matsumoto, Y., Kim, K., 1995. Excision of deoxyribose phosphate residues by DNA polymerase β during DNA repair. *Science* 269, 699–702.
- Nash, R.A., Caldecott, K.W., Barnes, D.E., Lindahl, T., 1997. XRCC1 protein interacts with one of two distinct forms of DNA ligase III. *Biochemistry* 36, 5207–5211.
- O'Connor, T.R., 1993. Purification and characterization of human 3-methyladenine-DNA glycosylase. *Nucleic Acids Research* 21, 5561–5569.
- Ocampo, M.T., Chaung, W., Marenstein, D.R., *et al.*, 2002. Targeted deletion of mNth1 reveals a novel DNA repair enzyme activity. *Molecular and Cellular Biology* 22, 6111–6121.
- Parsons, J.L., Dianov, G.L., 2013. Co-ordination of base excision repair and genome stability. *DNA Repair (Amst)* 12, 326–333.
- Parsons, J.L., Dianova, I.I., Khoronenkova, S.V., *et al.*, 2011. USP47 is a deubiquitylating enzyme that regulates base excision repair by controlling steady-state levels of DNA polymerase beta. *Molecular Cell* 41, 609–615.
- Parsons, J.L., Elder, R.H., 2003. DNA N-glycosylase deficient mice: A tale of redundancy. *Mutation Research* 531, 165–175.
- Parsons, J.L., Kavli, B., Slupphaug, G., Dianov, G.L., 2007. NEIL1 is the major DNA glycosylase that processes 5-hydroxyuracil in the proximity of a DNA single-strand break. *Biochemistry* 46, 4158–4163.
- Parsons, J.L., Tait, P.S., Finch, D., *et al.*, 2008. CHIP-mediated degradation and DNA damage-dependent stabilization regulate base excision repair proteins. *Molecular Cell* 29, 477–487.
- Parsons, J.L., Tait, P.S., Finch, D., *et al.*, 2009. Ubiquitin ligase ARF-BP1/Mule modulates base excision repair. *EMBO Journal* 28, 3207–3215.
- Parsons, J.L., Zharkov, D.O., Dianov, G.L., 2005. NEIL1 excises 3' end proximal oxidative DNA lesions resistant to cleavage by NTH1 and OGG1. *Nucleic Acids Research* 33, 4849–4856.
- Petronzelli, F., Riccio, A., Markham, G.D., *et al.*, 2000. Investigation of the substrate spectrum of the human mismatch-specific DNA N-glycosylase MED1 (MBD4): fundamental role of the catalytic domain. *Journal of Cellular Physiology* 185, 473–480.

- Podlutzky, A.J., Dianova, I.I., Podust, V.N., Bohr, V.A., Dianov, G.L., 2001. Human DNA polymerase beta initiates DNA synthesis during long-patch repair of reduced AP sites in DNA. *EMBO Journal* 20, 1477–1482.
- Puebla-Osorio, N., Lacey, D.B., Alt, F.W., Zhu, C., 2006. Early embryonic lethality due to targeted inactivation of DNA ligase III. *Molecular and Cellular Biology* 26, 3935–3941.
- Robson, C.N., Hickson, I.D., 1991. Isolation of cDNA clones encoding a human apurinic/aprimidinic endonuclease that corrects DNA repair and mutagenesis defects in *E. coli xth* (exonuclease III) mutants. *Nucleic Acids Research* 19, 5519–5523.
- Rosenquist, T.A., Zaika, E., Fernandes, A.S., *et al.*, 2003. The novel DNA glycosylase, NEIL1, protects mammalian cells from radiation-mediated cell death. *DNA Repair (Amst)* 2, 581–591.
- Sampath, H., Vartanian, V., Rollins, M.R., *et al.*, 2012. 8-Oxoguanine DNA glycosylase (OGG1) deficiency increases susceptibility to obesity and metabolic dysfunction. *PLoS One* 7, e51697.
- Sobol, R.W., Horton, J.K., Kuhn, R., *et al.*, 1996. Requirement of mammalian DNA polymerase-beta in base-excision repair. *Nature* 379, 183–186.
- Takao, M., Kanno, S., Kobayashi, K., *et al.*, 2002. A back-up glycosylase in Nth1 knock-out mice is a functional Nei (endonuclease VIII) homologue. *Journal of Biological Chemistry* 277, 42205–42213.
- Tebbs, R.S., Flannery, M.L., Meneses, J.J., *et al.*, 1999. Requirement for the Xrcc1 DNA base excision repair gene during early mouse development. *Developmental Biology* 208, 513–529.
- Vartanian, V., Lowell, B., Minko, I.G., *et al.*, 2006. The metabolic syndrome resulting from a knockout of the NEIL1 DNA glycosylase. *Proceedings of the National Academy of Sciences of the United States of America* 103, 1864–1869.
- Verly, W.G., Paquette, Y., 1972. An endonuclease for depurinated DNA in *Escherichia coli* B. *Canadian Journal of Biochemistry* 50, 217–224.
- Wallace, S.S., 2013. DNA glycosylases search for and remove oxidized DNA bases. *Environmental and Molecular Mutagenesis* 54, 691–704.
- Wiederhold, L., Leppard, J.B., Kedar, P., *et al.*, 2004. AP endonuclease-independent DNA base excision repair in human cells. *Molecular Cell* 15, 209–220.
- Xanthoudakis, S., Smeyne, R.J., Wallace, J.D., Curran, T., 1996. The redox/DNA repair protein, Ref-1, is essential for early embryonic development in mice. *Proceedings of the National Academy of Sciences of the United States of America* 93, 8919–8923.
- Zhou, J., Liu, M., Fleming, A.M., Burrows, C.J., Wallace, S.S., 2013. Neil3 and NEIL1 DNA glycosylases remove oxidative damages from quadruplex DNA and exhibit preferences for lesions in the telomeric sequence context. *Journal of Biological Chemistry* 288, 27263–27272.

Further Reading

Friedberg, E.C., Elledge, S.J., Lehmann, A.R., Lindahl, T., Muzi-Falconi, M. (Eds.), 2013. *DNA Repair, Mutagenesis, and Other Responses to DNA Damage*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Relevant Website

http://sciencepark.mdanderson.org/labs/wood/dna_repair_genes.html
University of Texas MD Anderson Cancer Center.

Nonhomologous DNA End Joining

S Sharma, Indian Institute of Science, Bangalore, India and FIRC Institute of Molecular Oncology, Milan, Italy
SC Raghavan, Indian Institute of Science, Bangalore, India

© 2016 Elsevier Inc. All rights reserved.

Introduction

DNA in each mammalian cell accumulates various types of lesions every day. DNA damage may be due to external agents like exposure to ultraviolet, ionization radiation, or internal agents like reactive oxygen species generated during metabolic or hydrolytic processes. These lesions can lead to break or breaks in strand or strands of DNA, an alteration in its chemical structure, loss, or a chemical change in a base of DNA. DNA damage if left unrepaired or incorrectly repaired may lead to genomic instability and chromosomal aberrations, eventually leading to cell death or cancer. Therefore, multiple repair pathways and complex DNA damage responses that recognize and repair these DNA damages have evolved to maintain the continuity and integrity of the genome.

Among various damages, DNA double-strand break repair (DSBR) is considered as the most lethal, as unrepaired double-strand breaks (DSBs) can lead to chromosomal translocations and consequently, cancer. Homologous recombination (HR) and nonhomologous DNA end joining (NHEJ) are two distinct pathways for the repair of DSB. HR, in general, requires regions of extensive homology and hence is conservative in nature and is the predominant pathway for DSBR in bacteria and lower eukaryotes. NHEJ, on the other hand, involves end-to-end ligation of the DNA termini accompanied by end editing with little or no requirement of homology.

NHEJ is considered as the major DSB repair pathway in higher eukaryotes and is believed to be active throughout the cell cycle. During lymphocytic development, both B-cells and T-cells utilize NHEJ to generate diversity in their repertoire of B-cell receptors and T-cell receptors, respectively. Even during maturation phase of B-cells for the generation of antibodies during class switch recombination, NHEJ plays a vital role. The mechanism and proteins involved in NHEJ is still an active area of research. Extensive genetic (using cell lines that were radiosensitive due to defective DSB repair) and biochemical studies led to the discovery of various NHEJ proteins, which have provided novel insights into the mechanism of NHEJ.

Mechanism of NHEJ

Recognition of DSBs

KU heterodimer that consists of two subunits, KU70 and KU80, is important for binding to DNA ends, once a DSB occurs during G0, G1, or early S phase of cell cycle (Figure 1; Chu, 1997; Dynan and Yoo, 1998). Interestingly, there appears to be a competition between HR and NHEJ machinery during late S and G2 phase (Bunting *et al.*, 2010; Adachi *et al.*, 2001; Sonoda *et al.*, 2006). KU complex along with DNA-PKcs gets activated upon binding to DNA, leading to a change in the conformation of KU (Lieber *et al.*, 2006). The crystal structure of KU heterodimer complexed with DNA showed that it has

toroidal shape with a hole for DNA (Walker *et al.*, 2001). Positively charged amino acids line the asymmetrical ring which interact with negatively charged phosphate backbone of DNA which allows DNA end-dependent movement of KU to interior of DNA (Walker *et al.*, 2001). Hence, KU heterodimer in association with DNA-PKcs forms a synaptic complex which helps in bringing the broken ends together (Bliss and Lane, 1997; Cary *et al.*, 1997; Ramsden and Gellert, 1998). However, some studies also suggested that recruitment of KU and DNA-PKcs could follow any order (Nick McElhinny *et al.*, 2000). Interestingly, mathematical modeling of this pathway has shown that DSB repair is similar, irrespective of the order as long as any protein is not rate-limiting (Li and Cucinotta, 2011). Very recently, another protein PAXX (paralog of XRCC4 and XLF) has been discovered which has been shown to interact with KU and may help in assembling core NHEJ factors on damaged chromatin (Ochi *et al.*, 2015).

Nucleolytic Processing by ARTEMIS–DNA-PKcs Complex

As explained before, KU slides internally after bringing the two DNA ends together, which leads to the contact of DNA-PKcs with DNA (Figure 1). This results in the activation of its serine/threonine kinase activity (Chu, 1997; Yoo and Dynan, 1999; Hammarsten *et al.*, 2000; Ma *et al.*, 2002). Once activated it phosphorylates itself as well as ARTEMIS. It has been shown that phosphorylated ARTEMIS along with DNA-PKcs can act as an endonuclease that can cleave 5'-overhangs, 3'-overhangs, and other DNA structures. It can also cleave hairpin intermediates of V(D)J recombination (Ma *et al.*, 2002, 2004, 2005; Yannone *et al.*, 2008). Additionally, ARTEMIS possesses an exonuclease activity by itself (Ma *et al.*, 2005). Hence, both the exonuclease and endonuclease activity of ARTEMIS is utilized in processing the ends (Ma *et al.*, 2002; Niewolik *et al.*, 2006).

Template-Dependent and Template-Independent DNA Synthesis

Once the ends are processed, they are end-filled using polymerases POL λ , POL μ , and terminal deoxynucleotidyl transferase (TdT), in order to make them ligatable (Figure 1). These polymerases belong to POLX family and they contain C-terminal portion of the BRCA-1 gene (BRCT) domain. The role of this family in NHEJ has been shown in both yeast and mammalian cells by genetic and biochemical studies. Efficacy and efficiency of fill-in synthesis varies between POL λ (fill-in of 5' overhangs) and POL μ (fill-in of 3' overhangs) (Lieber *et al.*, 2008). TdT has a major role during V(D)J recombination (Lieber, 2010).

Ligation

Once the ends are processed, they are ligated using DNA Ligase IV. The stability of Ligase IV is enhanced by XRCC4 (X-ray

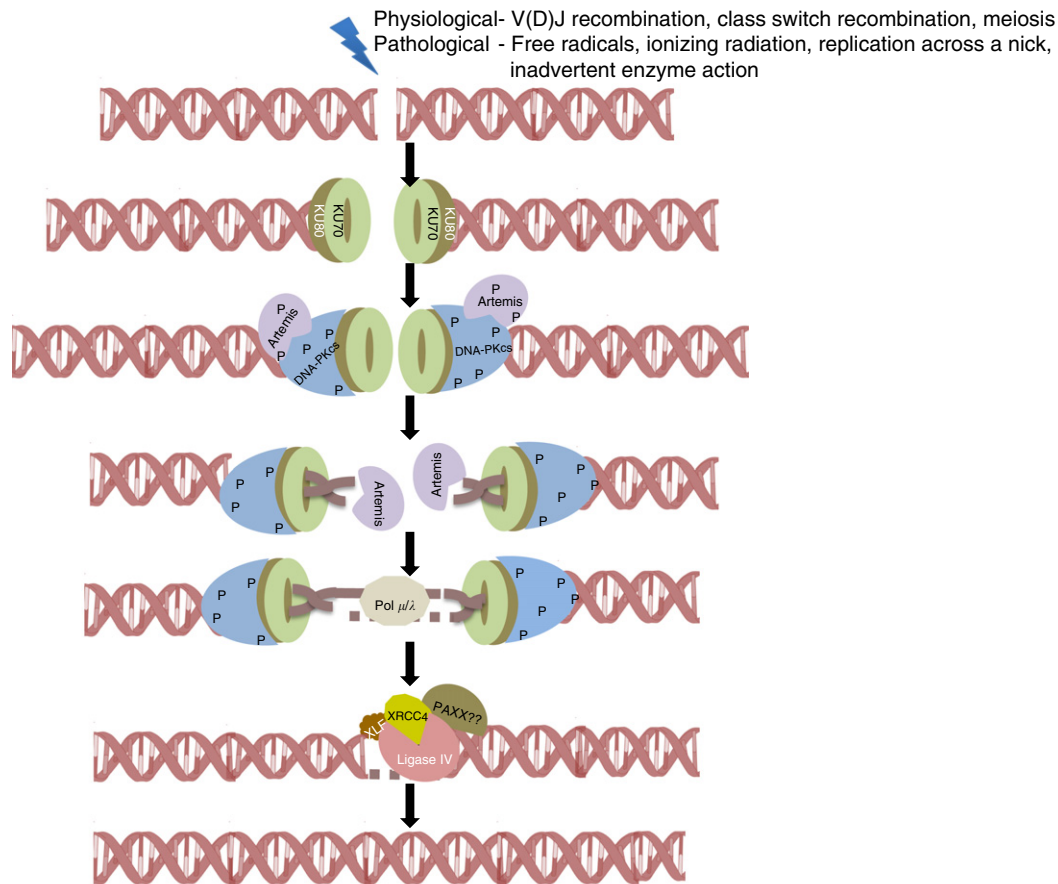


Figure 1 Mechanism of NHEJ. DSBs are generated either physiologically during V(D)J recombination etc. or pathologically by ionizing radiation, free radicals etc. Once a DSB is generated and is committed to repair by NHEJ, KU70/KU80 heterodimer is recruited to the break. This not only inhibits nucleolytic process but also helps in recruitment of DNA-PKcs, which autophosphorylates and then phosphorylates Artemis. Endonucleolytic activity of Artemis processes the DNA ends followed by strand synthesis by Pol μ or Pol λ , the Pol X family members. XLF–XRCC4–Ligase IV is recruited to the DNA ends to ligate DNA ends. PAXX may interact with XRCC4 and XLF to mediate ligation.

repair cross-complementing protein 4) that stimulates adenylation and helps DNA Ligase IV to interact with DNA (Grawunder *et al.*, 1997; Schar *et al.*, 1997; Teo and Jackson, 1997; Wilson *et al.*, 1997; Modesti *et al.*, 1999). The activity of XRCC4–Ligase IV is further stimulated by XLF (XRCC4-like factor) (Ahnesorg *et al.*, 2006; Buck *et al.*, 2006; Callebaut *et al.*, 2006). Interestingly, XRCC4–Ligase IV can ligate one strand of DNA irrespective of the other strand (Gu *et al.*, 2007a,b) and is also dependent on sequence of overhangs (Gu *et al.*, 2007a,b). It has also been shown that PAXX works along with XRCC4 and XLF (Ochi *et al.*, 2015). Therefore, XLF–XRCC4–PAXX–Ligase IV complex plays a very important role in ligating incompatible DNA ends efficiently (Figure 1).

A-NHEJ

Another distinct and highly error-prone but less-characterized pathway that has been shown to join DSBs in the absence of key NHEJ factor, has been named as A-NHEJ or backup NHEJ or microhomology-mediated end joining (MMEJ) (Figure 2).

A-NHEJ has been believed to have a role in chromosomal translocations and genomic instability.

A-NHEJ pathway was first identified when mice deficient for *p53* and *Xrcc4* showed co-amplification of *c-Myc* and *IgH* locus from pro-B lymphomas (Zhu *et al.*, 2002). Earlier studies have shown that this pathway was dominant only when key NHEJ proteins like XRCC4 or Ligase IV were absent (Yan *et al.*, 2007; Corneo *et al.*, 2007; Guirouilh-Barbat *et al.*, 2007). Although little is known about the detailed mechanism and protein components, potential role of MRE11–RAD50–NBS1, PARP1, PARP2, DNA Ligase III α , and WRN in A-NHEJ has been shown (Figure 2; Dinkelman *et al.*, 2009; Deriano *et al.*, 2009; Robert *et al.*, 2009; Sallmyr *et al.*, 2008). Although the existence of this distinct pathway in normal mammalian cells is being explored (Truong *et al.*, 2013) its existence as a separate DSB repair pathway is still not clear (Pannunzio *et al.*, 2014).

Choice of DSB Repair Pathway

Since both NHEJ and HR are an absolute requirement for repairing DSBs, the choice of one pathway over the other

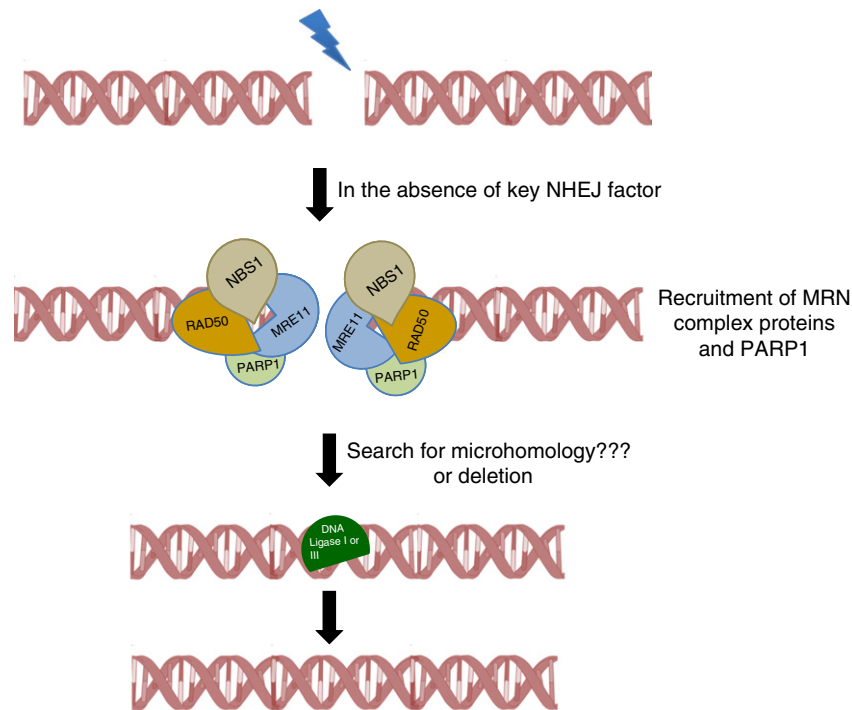


Figure 2 Probable mechanism of alternative or backup NHEJ. Once a DSB is generated in the absence of a key NHEJ factor, proteins like MRE11, NBS1, and PARP1 are recruited. This may further help in homology search and deletions at the ends. Phosphodiester bonds are recreated to maintain the continuity of genome by DNA Ligase I or III.

depends on factors like cell cycle, end resection, or DSB end structure. The availability of a homologous template during post-replicative stages of S phase makes HR the predominant pathway in S phase while condensed chromatin makes homology search difficult in G2-M phase restricting its dominance (Figure 3). It was recently shown that phosphorylation of 53BP1 along with Rif1 inhibit the recruitment of BRCA1 during G1 and hence limiting end resection thereby promoting NHEJ (Figure 3; Chapman *et al.*, 2013; Zimmermann and de Lange, 2014). On the other hand, upon phosphorylation of CtIP during S and G2 phases, end resection is activated promoting HR (Yun and Hiom, 2009).

NHEJ and Disease

Since NHEJ is an imprecise mode of joining, the ligated DNA molecules may have some loss of nucleotides 'information scar' which may accumulate in the genome and contribute to biological aging (Lieber *et al.*, 2008). A subset of such mutations may lead to excessive cellular proliferation or cancer. On the other hand, it is also known that *XRCC4* or *Ligase IV* knock out mice are embryonically lethal (Barnes *et al.*, 1998). It has been observed that cells possessing either mutations or altered levels of key NHEJ proteins render them radiosensitive and predispose them to cancer (Marangoni *et al.*, 2000; Riballo *et al.*, 1999; van der Burg *et al.*, 2009). Role of Ligase IV, XLF, and Artemis has been implicated in genetic disorders, namely, Ligase IV syndrome, immunodeficiency, and radiosensitive SCID (Rass *et al.*, 2007). Hence, deregulation of repair proteins

may lead to genomic rearrangements, higher mutations, and survival advantages to cancer cells (Lieberman, 2008; Powell and Bindra, 2009). In fact, KU and DNA-PKcs were found to be over-expressed in oral squamous, lung carcinoma, and breast, gastric and esophageal cancers. In essence, NHEJ proteins can be targeted to sensitize tumor cells (Srivastava and Raghavan, 2015).

NHEJ and Inhibitors

Small-molecule inhibitors are currently being developed which can lead to the death of cancer cells. One of the earliest inhibitors, wortmannin, an inhibitor of DNA-PKcs was not selective and acts as a pan PI3K inhibitor (Oliveira *et al.*, 2002). LY2940002 is another PI3K inhibitor known to inhibit DNA-PKcs (Rosenzweig *et al.*, 1997). NU7026 was then identified as a selective and potent inhibitor of DNA-PK (Willmore *et al.*, 2004). Since inhibition at an early stage of NHEJ may divert DSB repair toward alternative pathways, blocking DNA Ligase IV would lead to accumulation of DSBs, leading to apoptosis. Although L189 was developed as a potential Ligase IV inhibitor, it was active against all three ligases, Ligase I, III, and IV (Chen *et al.*, 2008). Recently, a selective inhibitor of Ligase IV, SCR7, was discovered which inhibited NHEJ both *in vitro* and *in vivo* (Srivastava *et al.*, 2012). It showed significant tumor regression in mouse tumor models and improved the efficacy of radio and chemotherapy, when coadministered (Srivastava *et al.*, 2012). It has been shown that encapsulated SCR7 (ESCR7) was fivefold better than its

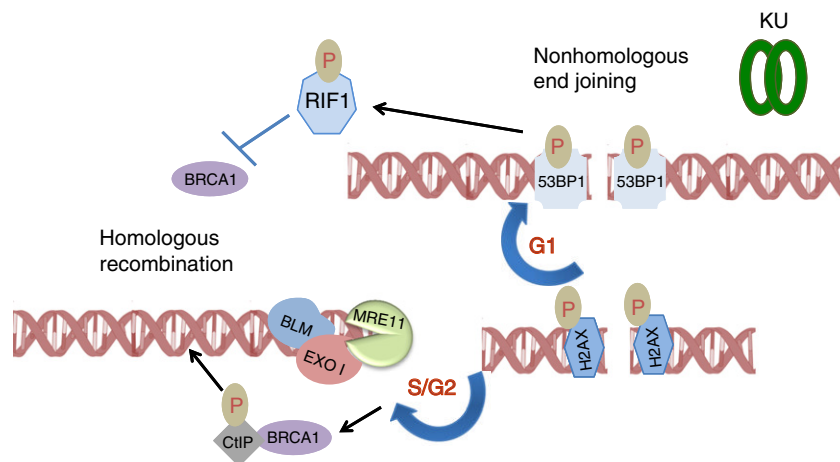


Figure 3 Factors determining choice of DSB repair pathways. Numerous regulatory mechanisms are involved in choosing NHEJ as the double strand repair pathway for the repair of break. During S or G2 phase of the cell cycle, HR is the preferred pathway as the sister chromatid is readily available. Cell-cycle-dependent kinases (CDKs) phosphorylate CtIP that along with BRCA1 helps in end resection to initiate HR. Although NHEJ is active throughout the cell cycle, it plays a vital role during M or G1 phase. Phosphorylation of Rif1 by 53BP1 inhibits BRCA1 that contains end resection by Mre11/EXO1/BLM. This leads to recruitment of KU to the DNA ends and initiating NHEJ.

parental form (John *et al.*, 2015a,b). Thus, different therapeutic strategies need to be developed using DSB repair as a target selectively against cancer cells.

Acknowledgments

This work was supported by grant from DBT, India BT/HRD/NBA/34/01/2012 (ix) for SCR. SS acknowledges Postdoctoral fellowship from IISc, India.

See also: Nucleic Acid Synthesis/Breakdown: DNA Synthesis/Repair: DNA Repair by Homologous Recombination

References

- Adachi, N., Ishino, T., Ishii, Y., Takeda, S., Koyama, H., 2001. DNA ligase IV-deficient cells are more resistant to ionizing radiation in the absence of Ku70: Implications for DNA double-strand break repair. *Proceedings of the National Academy of Sciences of the United States of America* 98 (21), 12109–12113.
- Ahnesorg, P., Smith, P., Jackson, S.P., 2006. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell* 124 (2), 301–313.
- Barnes, D.E., Stamp, G., Rosewell, I., Denzel, A., Lindahl, T., 1998. Targeted disruption of the gene encoding DNA ligase IV leads to lethality in embryonic mice. *Current Biology* 8 (25), 1395–1398.
- Bliss, T.M., Lane, D.P., 1997. Ku selectively transfers between DNA molecules with homologous ends. *Journal of Biological Chemistry* 272 (9), 5765–5773.
- Buck, D., Malivert, L., de Chasseval, R., *et al.*, 2006. Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell* 124 (2), 287–299.
- Bunting, S.F., Callen, E., Wong, N., *et al.*, 2010. 53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks. *Cell* 141 (2), 243–254.
- van der Burg, M., Ijspeert, H., Verkaik, N.S., *et al.*, 2009. A DNA-PKcs mutation in a radiosensitive T-B- SCID patient inhibits Artemis activation and nonhomologous end-joining. *Journal of Clinical Investigation* 119 (1), 91–98.
- Callebaut, I., Malivert, L., Fischer, A., *et al.*, 2006. Cernunnos interacts with the XRCC4 x DNA-ligase IV complex and is homologous to the yeast nonhomologous end-joining factor Nej1. *Journal of Biological Chemistry* 281 (20), 13857–13860.
- Cary, R.B., Peterson, S.R., Wang, J., *et al.*, 1997. DNA looping by Ku and the DNA-dependent protein kinase. *Proceedings of the National Academy of Sciences of the United States of America* 94 (9), 4267–4272.
- Chapman, J.R., Barral, P., Vannier, J.B., *et al.*, 2013. RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection. *Molecular Cell* 49 (5), 858–871.
- Chu, G., 1997. Double strand break repair. *Journal of Biological Chemistry* 272 (39), 24097–24100.
- Corneo, B., Wendland, R.L., Deriano, L., *et al.*, 2007. Rag mutations reveal robust alternative end joining. *Nature* 449 (7161), 483–486.
- Deriano, L., Stracker, T.H., Baker, A., Petrini, J.H., Roth, D.B., 2009. Roles for NBS1 in alternative nonhomologous end-joining of V(D)J recombination intermediates. *Molecular Cell* 34 (1), 13–25.
- Dinkelmann, M., Spehalski, E., Stoneham, T., *et al.*, 2009. Multiple functions of MRN in end-joining pathways during isotype class switching. *Nature Structural & Molecular Biology* 16 (8), 808–813.
- Dynan, W.S., Yoo, S., 1998. Interaction of Ku protein and DNA-dependent protein kinase catalytic subunit with nucleic acids. *Nucleic Acids Research* 26 (7), 1551–1559.
- Grawunder, U., Wilm, M., Wu, X., *et al.*, 1997. Activity of DNA ligase IV stimulated by complex formation with XRCC4 protein in mammalian cells. *Nature* 388 (6641), 492–495.
- Gu, J., Lu, H., Tiffin, B., *et al.*, 2007a. XRCC4:DNA ligase IV can ligate incompatible DNA ends and can ligate across gaps. *EMBO Journal* 26 (4), 1010–1023.
- Gu, J., Lu, H., Tsai, A.G., Schwarz, K., Lieber, M.R., 2007b. Single-stranded DNA ligation and XLF-stimulated incompatible DNA end ligation by the XRCC4-DNA ligase IV complex: Influence of terminal DNA sequence. *Nucleic Acids Research* 35 (17), 5755–5762.
- Guirouilh-Barbat, J., Rass, E., Plo, I., Bertrand, P., Lopez, B.S., 2007. Defects in XRCC4 and KU80 differentially affect the joining of distal nonhomologous ends. *Proceedings of the National Academy of Sciences of the United States of America* 104 (52), 20902–20907.
- Hammarsten, O., DeFazio, L.G., Chu, G., 2000. Activation of DNA-dependent protein kinase by single-stranded DNA ends. *Journal of Biological Chemistry* 275 (3), 1541–1550.
- John, F., George, J., Srivastava, M., *et al.*, 2015a. Pluronic copolymer encapsulated SCR7 as a potential anticancer agent. *Faraday Discussions* 177, 155–161.
- John, F., George, J., Vartak, S.V., *et al.*, 2015b. Enhanced efficacy of pluronic copolymer micelle encapsulated SCR7 against cancer cell proliferation. *Macromolecular Bioscience* 15 (4), 521–534.

- Li, Y., Cucinotta, F.A., 2011. Modeling non-homologous end joining. *Journal of Theoretical Biology* 283 (1), 122–135.
- Lieber, M.R., 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual Review of Biochemistry* 79, 181–211.
- Lieber, M.R., Lu, H., Gu, J., Schwarz, K., 2008. Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligase of vertebrate non-homologous DNA end joining: Relevance to cancer, aging, and the immune system. *Cell Research* 18 (1), 125–133.
- Lieber, M.R., Yu, K., Raghavan, S.C., 2006. Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations. *DNA Repair (Amst)* 5 (9–10), 1234–1245.
- Lieberman, H.B., 2008. DNA damage repair and response proteins as targets for cancer therapy. *Current Medicinal Chemistry* 15 (4), 360–367.
- Ma, Y., Lu, H., Tippin, B., *et al.*, 2004. A biochemically defined system for mammalian nonhomologous DNA end joining. *Molecular Cell* 16 (5), 701–713.
- Ma, Y., Pannicke, U., Schwarz, K., Lieber, M.R., 2002. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* 108 (6), 781–794.
- Ma, Y., Schwarz, K., Lieber, M.R., 2005. The Artemis: DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps. *DNA Repair (Amsterdam)* 4 (7), 845–851.
- Marangoni, E., Foray, N., O'Driscoll, M., *et al.*, 2000. A Ku80 fragment with dominant negative activity imparts a radiosensitive phenotype to CHO-K1 cells. *Nucleic Acids Research* 28 (23), 4778–4782.
- Modesti, M., Hesse, J.E., Gellert, M., 1999. DNA binding of Xrcc4 protein is associated with V(D)J recombination but not with stimulation of DNA ligase IV activity. *EMBO Journal* 18 (7), 2008–2018.
- Nick McElhinny, S.A., Snowden, C.M., McCarville, J., Ramsden, D.A., 2000. Ku recruits the XRCC4-ligase IV complex to DNA ends. *Molecular and Cellular Biology* 20 (9), 2996–3003.
- Niewolik, D., Pannicke, U., Lu, H., *et al.*, 2006. DNA-PKcs dependence of Artemis endonucleolytic activity, differences between hairpins and 5' or 3' overhangs. *Journal of Biological Chemistry* 281 (45), 33900–33909.
- Ochi, T., Blackford, A.N., Coates, J., *et al.*, 2015. DNA repair. PAXX, a paralog of XRCC4 and XLF, interacts with Ku to promote DNA double-strand break repair. *Science* 347 (6218), 185–188.
- Oliveira, N.G., Castro, M., Rodrigues, A.S., *et al.*, 2002. DNA-PK inhibitor wortmannin enhances DNA damage induced by bleomycin in V79 Chinese hamster cells. *Teratogenesis, Carcinogenesis, and Mutagenesis* 22 (5), 343–351.
- Pannunzio, N.R., Li, S., Watanabe, G., Lieber, M.R., 2014. Non-homologous end joining often uses microhomology: Implications for alternative end joining. *DNA Repair* 17, 74–80.
- Powell, S.N., Bindra, R.S., 2009. Targeting the DNA damage response for cancer therapy. *DNA Repair* 8 (9), 1153–1165.
- Ramsden, D.A., Gellert, M., 1998. Ku protein stimulates DNA end joining by mammalian DNA ligases: A direct role for Ku in repair of DNA double-strand breaks. *EMBO Journal* 17 (2), 609–614.
- Rass, U., Ahel, I., West, S.C., 2007. Defective DNA repair and neurodegenerative disease. *Cell* 130 (6), 991–1004.
- Riballo, E., Critchlow, S.E., Teo, S.H., *et al.*, 1999. Identification of a defect in DNA ligase IV in a radiosensitive leukaemia patient. *Current Biology* 9 (13), 699–702.
- Robert, I., Dantzer, F., Reina-San-Martin, B., 2009. Parp1 facilitates alternative NHEJ, whereas Parp2 suppresses IgH/c-myc translocations during immunoglobulin class switch recombination. *Journal of Experimental Medicine* 206 (5), 1047–1056.
- Rosenzweig, K.E., Youmell, M.B., Palayoor, S.T., Price, B.D., 1997. Radiosensitization of human tumor cells by the phosphatidylinositol(3)-kinase inhibitors wortmannin and LY294002 correlates with inhibition of DNA-dependent protein kinase and prolonged G2-M delay. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 3 (7), 1149–1156.
- Sallmyr, A., Tomkinson, A.E., Rassool, F.V., 2008. Up-regulation of WRN and DNA ligase IIIalpha in chronic myeloid leukemia: Consequences for the repair of DNA double-strand breaks. *Blood* 112 (4), 1413–1423.
- Schar, P., Herrmann, G., Daly, G., Lindahl, T., 1997. A newly identified DNA ligase of *Saccharomyces cerevisiae* involved in RAD52-independent repair of DNA double-strand breaks. *Genes & Development* 11 (15), 1912–1924.
- Sonoda, E., Hochegger, H., Saberi, A., Taniguchi, Y., Takeda, S., 2006. Differential usage of non-homologous end-joining and homologous recombination in double strand break repair. *DNA Repair (Amsterdam)* 5 (9–10), 1021–1029.
- Srivastava, M., Nambiar, M., Sharma, S., *et al.*, 2012. An inhibitor of nonhomologous end-joining abrogates double-strand break repair and impedes cancer progression. *Cell* 151 (7), 1474–1487.
- Srivastava, M., Raghavan, S.C., 2015. DNA double-strand break repair inhibitors as cancer therapeutics. *Chemistry & Biology* 22 (1), 17–29.
- Teo, S.H., Jackson, S.P., 1997. Identification of *Saccharomyces cerevisiae* DNA ligase IV: Involvement in DNA double-strand break repair. *EMBO Journal* 16 (15), 4788–4795.
- Truong, L.N., Li, Y., Shi, L.Z., *et al.*, 2013. Microhomology-mediated end joining and homologous recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* 110 (19), 7720–7725.
- Walker, J.R., Corpina, R.A., Goldberg, J., 2001. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* 412 (6847), 607–614.
- Willmore, E., de Caux, S., Sunter, N.J., *et al.*, 2004. A novel DNA-dependent protein kinase inhibitor, NU7026, potentiates the cytotoxicity of topoisomerase II poisons used in the treatment of leukemia. *Blood* 103 (12), 4659–4665.
- Wilson, T.E., Grawunder, U., Lieber, M.R., 1997. Yeast DNA ligase IV mediates non-homologous DNA end joining. *Nature* 388 (6641), 495–498.
- Yan, C.T., Boboila, C., Souza, E.K., *et al.*, 2007. IgH class switching and translocations use a robust non-classical end-joining pathway. *Nature* 449 (7161), 478–482.
- Yannone, S.M., Khan, I.S., Zhou, R.Z., *et al.*, 2008. Coordinate 5' and 3' endonucleolytic trimming of terminally blocked blunt DNA double-strand break ends by Artemis nuclease and DNA-dependent protein kinase. *Nucleic Acids Research* 36 (10), 3354–3365.
- Yoo, S., Dynan, W.S., 1999. Geometry of a complex formed by double strand break repair proteins at a single DNA end: Recruitment of DNA-PKcs induces inward translocation of Ku protein. *Nucleic Acids Research* 27 (24), 4679–4686.
- Yun, M.H., Hiom, K., 2009. CtIP-BRCA1 modulates the choice of DNA double-strand-break repair pathway throughout the cell cycle. *Nature* 459 (7245), 460–463.
- Zhu, C., Mills, K.D., Ferguson, D.O., *et al.*, 2002. Unrepaired DNA breaks in p53-deficient cells lead to oncogenic gene amplification subsequent to translocations. *Cell* 109 (7), 811–821.
- Zimmermann, M., de Lange, T., 2014. 53BP1: Pro choice in DNA repair. *Trends in Cell Biology* 24 (2), 108–117.

DNA Repair by Homologous Recombination

SS Jenkins, S Mukherjee, and W-D Heyer, University of California, Davis, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Crossover/noncrossover Homologous recombination can lead to reciprocal exchange of DNA sequences flanking a recombination event (crossover) or not (noncrossover).

D-loop Primary DNA strand invasion product by the Rad51-ssDNA filament, which involves the formation of hDNA.

hDNA Hybrid DNA or heteroduplex (hDNA) refers to duplex DNA generated by DNA strand invasion during recombination. In case this involves a DNA sequence difference, hDNA may be subjected to mismatch repair, which may lead to gene conversion. Absence of mismatch repair maintains different genetic information on the Watson and Crick strands of the duplex DNA.

Hemi-catenane Specific topological entanglement of duplex DNA involving two single-strands of each duplex (see step 7c in [Figure 1](#)).

Holliday junction, double Holliday junction Four-stranded cross-strand structures mediated by homologous recombination. While single HJs can only be resolved to crossover or noncrossover outcome, dHJs can be resolved to crossover or noncrossover outcomes as well as dissolved to noncrossover outcome (see step 6a and ensuing steps in [Figure 1](#)).

Resolution/dissolution Resolution refers to the processing of recombination-mediated joint molecules by nucleases, whereas dissolution refers to alternative processes that do not involve nucleases.

Introduction

Exogenous DNA damage caused by radiation and environmental sources as well as endogenous DNA damage as a result of cellular metabolism (e.g., reactive oxygen species) or replication-associated genotoxic stress require constant attention by DNA damage repair and tolerance mechanisms ([Friedberg et al., 2006](#)). Homologous recombination (HR) comprises a number of distinct but related sub-pathways ([Figure 1](#)) that provide high-fidelity repair or tolerance of DNA damage by taking advantage of Watson–Crick base-pairing principles in a template-dependent fashion. The fundamental reaction involves DNA strand exchange that enables a 3'-DNA end engaging a suitable template for DNA synthesis to repair or bypass DNA damage and accurately restore contiguous chromosomes ([Figures 1–3](#)). Besides its direct role in repairing complex DNA damage such as DNA double-stranded breaks (DSB; [Figure 1](#)), DNA gaps ([Figure 2](#)), interstrand cross-links (ICL; [Figure 3](#)), and protein–DNA cross-links, HR also interfaces with DNA replication ([Figure 2](#)) and telomere maintenance ([Figure 1](#)) to ensure faithful duplication of the genome. This article focuses on the role of HR in somatic DNA damage repair and tolerance, drawing mostly from studies performed with budding yeast and humans. HR also plays a key role in meiotic chromosome segregation. In the section 'further reading,' the reader is referred to outstanding reviews that highlight specific aspects of HR in somatic DNA repair as well as its essential role in meiosis.

The Core Mechanism of Homologous Recombination

The core mechanism of HR involves generating single-stranded DNA (ssDNA) that allows assembly of the Rad51-ssDNA filament, which catalyzes the signature reactions of homology search and DNA strand invasion. DSB end-processing is well

understood and involves two successive resection steps: short-range resection by the Mre11–Rad50–Xrs2 complex (MRX; human MRE11–RAD50–NBS1; see [Table 1](#)) and long-range resection carried out by two partially overlapping pathways ([Figure 1](#), step 1). It is currently unclear, whether replication-associated gaps require processing ([Figure 2](#), step 3c).

DSB Resection

The MRX complex (human MRN; [Table 1](#)) initially recognizes and binds to DSBs ([Lisby et al., 2004](#)). The MRX complex integrates end-processing, DNA tethering, and signaling functions ([Stracker and Petrini, 2011](#)). Mre11 displays endonuclease and 3'–5' exonuclease activities that are essential for end-processing, when breaks contain chemically modified ends or bound proteins. RAD50 is an SMC-family protein with a long coiled-coil domain that can tether two DSB ends or different DNA molecules. Xrs2 (NBS1 in human) recruits the Tel1 kinase (human ATM) to the MRX complex to trigger the DNA damage response signaling cascade in response to DSBs. Sae2 (human CtIP) associates with the MRX (human MRN) complex, and it is still unclear whether it acts as a nuclease-cofactor or contains autonomous nuclease activity.

The initial short-range resection by MRX/Sae2 (human MRN/CtIP) appears to initiate with an endonucleolytic nick 50–100 nucleotides distant from the DSB ([Symington, 2014](#)). The Mre11 3'–5' exonuclease activity removes the terminal 5'-ending oligonucleotides resulting from the endonucleolytic incision. Following the initial short-range resection mediated by MRX/Sae2 (human MRN/CtIP), the broken DNA ends undergo additional long-range resection, generating long stretches of ssDNA. Two different pathways are involved in the long-range 5'–3' resection of the DSB ([Mimitou and Symington, 2008](#); [Zhu et al., 2008](#)). The processive 5'–3' exonuclease EXO1 defines the first long-range DSB resection pathway. The second pathway is significantly more complex and involves the

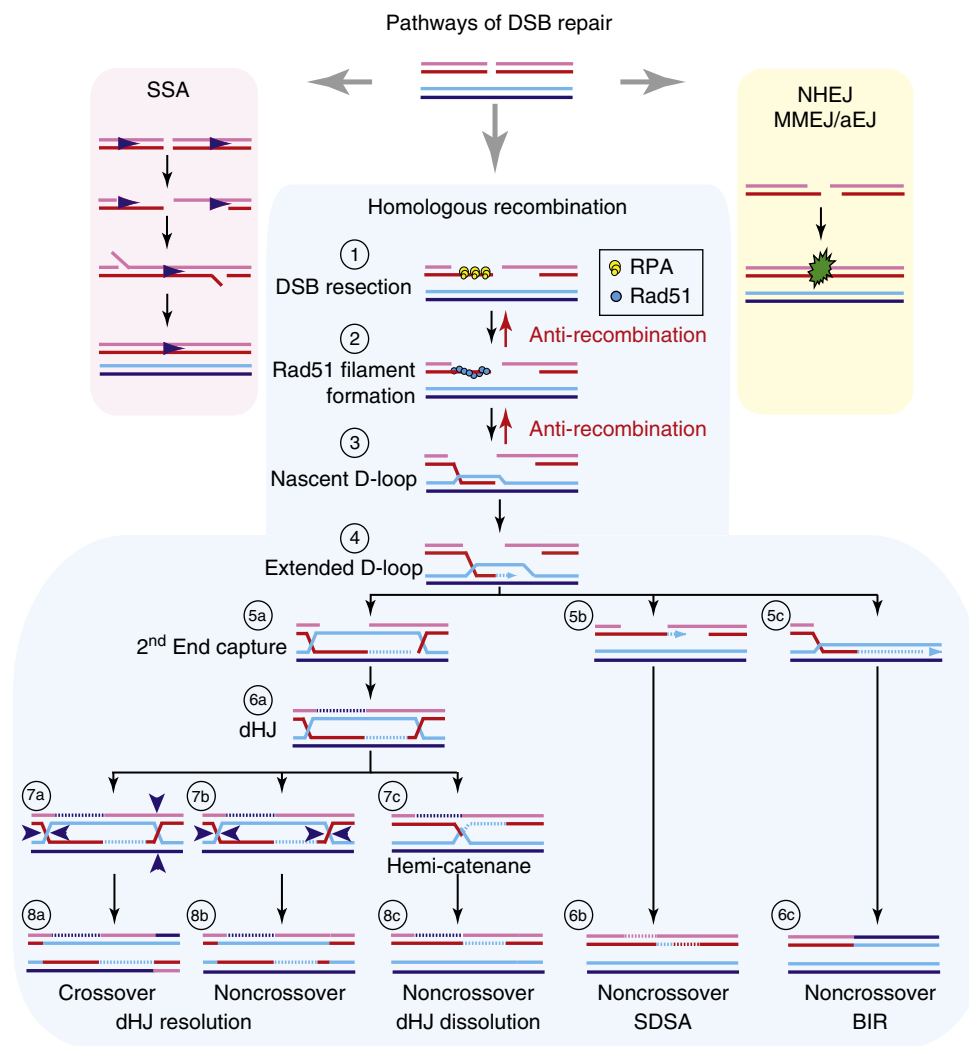


Figure 1 Pathways of DNA double-strand break (DSB) repair. DSBs are repaired using three different pathways. Single-strand annealing (SSA, red box) requires direct DNA sequence repeats flanking the DSB. After DNA end resection, the homologous sequences anneal leading to deletion of one repeat and of the intervening sequences. Nonhomologous end-joining (NHEJ, yellow box) rejoins broken ends by ligation in a template-independent manner with variable fidelity that depends on the exact chemical structure of the DSB. Microhomology-mediated end-joining (MMEJ), also called alternative end-joining (aEJ), involves limited end-processing to uncover short stretches (5–25 bp) of homology prior to ligation. Homologous recombination (HR, blue box) comprises a number of template-dependent pathways that exhibit high fidelity compared to SSA and NHEJ/MMEJ/aEJ. The HR pathways share the DNA strand invasion product, the 'D-loop,' but vary in the ensuing steps. The double Holliday junction (dHJ) pathway involves formation of the signature double Holliday junction intermediate, which can be resolved nucleolytically to generate crossover or noncrossover products. Alternatively, dHJs can be dissolved by the combined action of the BLM-TOPOIII α -RMI1-RMI2 complex (*Saccharomyces cerevisiae* Sgs1–Top3–Rmi1) to noncrossover products exclusively. The synthesis-dependent strand annealing (SDSA) pathway requires 'dissolution' of the 'D-loop' after extension by DNA polymerase and annealing of the extended strand with the second DSB end, leading exclusively to noncrossover products. During break-induced replication (BIR), 'D-loop' extension leads to DNA synthesis that copies the entire chromosome arm leading.

Sgs1–Top3–Rmi1 complex (human BLM-TOPOIII α -RMI1-RMI2; see [Table 1](#)) and the nuclease activity of the Dna2 helicase/nuclease ([Nimonkar et al., 2011](#)). The helicase activity of Sgs1 but not the catalytic activity of Top3 is required in this process ([Cejka et al., 2010a](#); [Niu et al., 2010](#)).

Rad51 Filament Assembly and Stability

The product of DSB resection is a 3'-OH ending ssDNA tail that allows assembly of the Rad51 nucleoprotein filament for

homology search and DNA strand invasion ([Figure 1](#), step 2; [Sung, 1994](#)). Binding of ATP endows Rad51 with a high-affinity DNA-binding state that allows its cooperative binding to ssDNA or double-stranded DNA (dsDNA). One Rad51 protomer binds 3–4 nucleotides, resulting in a right-handed Rad51 filament with a helical pitch of 130 Å and a 1.5-fold extension compared to B-form DNA ([Figure 1](#), step 2; [Conway et al., 2004](#)). Elegant X-ray crystallographic studies with RecA protein, the bacterial Rad51 homolog, showed nonuniform stretching of the DNA leaving triplets of normal B-form DNA

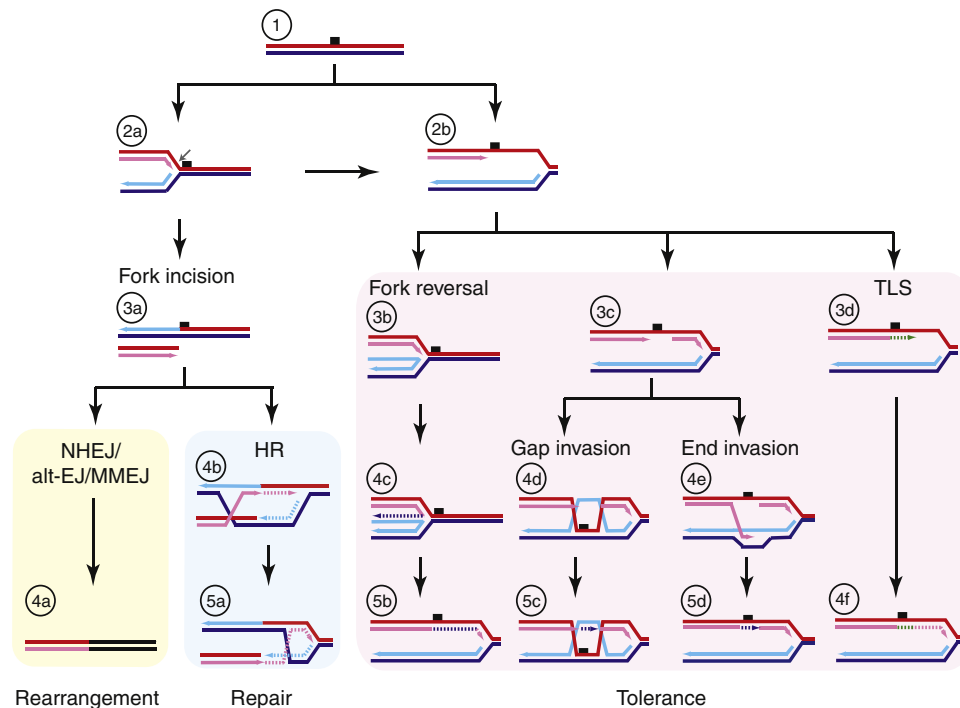


Figure 2 Replication fork support by homologous recombination (HR). A replication fork encountering a stalling lesion (black box) (step 1) may block the replicative helicase (step 2a) or the replicative DNA polymerase involving decoupling from the replicative helicase and generation of ssDNA gaps (step 2b). Multiple pathways, including homologous recombination, can engage under fork stalling conditions. The fork can be incised nucleolytically (step 3a) leading to a one-sided double strand break (DSB). Repair of a one-sided DSB by any end-joining pathway (NHEJ, MMEJ/aEJ) will result in a chromosomal rearrangement (step 4a), alternatively a telomere could be added leading to loss of the distal DNA (not shown). HR is the only pathway for accurate repair engaging the intact sister chromatid (step 4b) to restart replication (step 5a) after lesion removal (not shown). Uncoupled forks (step 2b) can engage at least three different DNA damage tolerance pathways to complete and/or restart replication. Replication fork reversal (step 3b) generates the so-called 'chicken-foot' intermediate, which provides an undamaged template for the blocked 3'-OH end (step 4c) leading the reestablishment of the replication fork (step 5b). Restart of DNA synthesis leaves a persisting gap (step 3c) that can be repaired by HR using two different pathways, depending on the specific mechanism (gap invasion, steps 4d, 5c; end invasion steps 4e, 5d). Alternatively, translesion synthesis (TLS) polymerases can directly synthesize across the blocking lesion (DNA synthesis marked in green) (step 3d) followed by switchback to the replicative polymerase for replication restart (step 4f). aEJ, alternative end-joining; NHEJ, non-homologous end-joining; MMEJ, microhomology-mediated end-joining.

flanked by large 8.4 Å gaps (Chen *et al.*, 2008). Presumably, this facilitates triplets to flip out during homology search suggesting that the fundamental search unit is a base triplet. Two challenges need to be addressed to target Rad51 to an ssDNA tail. First, ssDNA is avidly bound by RPA (replication protein A), the eukaryotic hetero-trimeric ssDNA-binding protein, which inhibits Rad51 filament formation. As detailed below, the recombination mediator proteins facilitate Rad51 filament nucleation on RPA-coated ssDNA (Liu *et al.*, 2011a). Second, Rad51 binds ssDNA only with slight preference over dsDNA. Given the vast excess of dsDNA over processed DSBs under typical DNA repair conditions, it is unclear how Rad51 is kept available for HR. Two potential solutions emanate: The dsDNA motor proteins Rad54 and Rdh4/Tid1 target dsDNA complexes of Rad51 or its meiotic homolog Dmc1 and dissociate Rad51 (or its meiotic homolog Dmc1) from dsDNA (Solinger *et al.*, 2002; Holzen *et al.*, 2006). Moreover, it is possible that Rad51 is kept available for repair by binding to specific protein factors to prevent duplex DNA binding, a scenario that has been proposed for BRCA2 and human RAD51 (Reuter *et al.*, 2014).

The role of RPA in HR is complex and multifaceted. RPA eliminates secondary structures in ssDNA to allow formation of productive Rad51-ssDNA filaments (Figure 1, step 1). In the absence of RPA, Rad51 also binds to secondary structures (dsDNA) resulting in Rad51-dsDNA filaments, which are incapable of homology search and DNA strand invasion. Hence the transition from RPA-coated ssDNA to Rad51-ssDNA filaments represents a critical quality control step. RPA also inhibits 3'-5' end resection by the Dna2 nuclease, an activity that would interfere with generating the 3' ssDNA tails needed for initiating HR (Cejka *et al.*, 2010a). During DNA strand invasion RPA binds the displaced strand in the 'D-loop' favoring heteroduplex DNA (hDNA) formation (Figure 1, step 3). Finally, there is evidence that RPA also binds the template strand within the 'D-loop' in front of the DNA polymerase extending the invading 3'-OH end to present a more optimal template for DNA synthesis, and therefore prevent polymerase stalling (Figure 1, step 4; Sneed *et al.*, 2013).

Recombination mediators permit stable Rad51 filament formation on RPA-coated ssDNA by facilitating nucleation of the Rad51 filament and/or stabilizing the Rad51-ssDNA

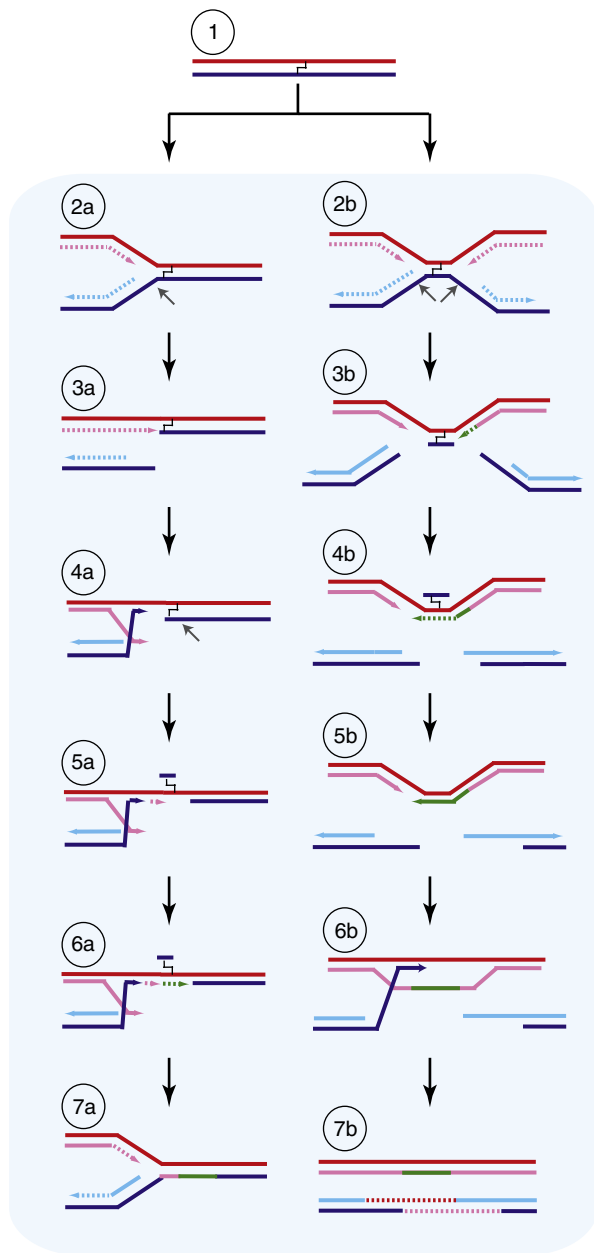


Figure 3 Interstrand cross-link (ICL) repair and homologous recombination. ICLs (black line connecting the Watson and Crick strand of the double helix) stall replication by blocking the replicative helicase (step 1). Fork incision of a unidirectional fork (step 2a) creates a one-sided double strand break (DSB) (step 3a) that liberates a 3'-end that can be resected for invasion of the homologous template (step 4a). After some synthesis, the replicative polymerase is blocked by the ICL and a nick is required to unhook the ICL (step 5a). This is followed by recruitment of a translesion synthesis (TLS) polymerase for lesion bypass (DNA synthesis shown in green) (step 6a) and the fork is reestablished with intact templates (step 7a) after removal of the ICL by nucleotide excision repair (NER) (not shown). Alternatively, dual incisions can unhook the damaged template (steps 1–2b) and generate a DSB including a short dsDNA gap (step 3b). First, DNA synthesis across lesion by TLS polymerases (DNA synthesis shown in green) establishes an intact template for DSB repair after removal of the ICL by excision repair (step 4b). DSB repair (steps 6b, 7b) is illustrated as an SDSA mechanism (see Figure 1).

filament from dissociation (Zelensky *et al.*, 2014). In yeast, Rad52 protein is the dominant mediator protein forming a multi-meric ring structure that binds ssDNA, Rad51, and RPA and facilitates the displacement of RPA by Rad51 on the ssDNA (Sung, 1997a). In mammals, the major mediator function is provided by the BRCA2 tumor suppressor protein (Jensen *et al.*, 2010), which explains why mutants or knock-downs of RAD52 in mammals elicit only modest phenotypes (Rijkers *et al.*, 1998; Feng *et al.*, 2010). The mechanism employed by both proteins is possibly different, as BRCA2, unlike Rad52, does not directly bind RPA (Jensen *et al.*, 2010). BRCA2 engages multiple strategies to favor formation of stable RAD51-ssDNA filaments by directly nucleating Rad51 filaments on RPA-coated ssDNA, impeding RAD51 binding to dsDNA and stabilizing the resulting filaments by reducing ATP turnover within the Rad51 filament (Liu *et al.*, 2010; Jensen *et al.*, 2010; Pellegrini *et al.*, 2002). Moreover, BRCA2 physically interacts with PALB2 and DSS1, both of which are required for HR (Table 1; Xia *et al.*, 2006; Gudmundsdottir *et al.*, 2004; Liu *et al.*, 2010). PALB2 localizes BRCA2 to sites of DNA damage (Xia *et al.*, 2006). The precise mechanism by which PALB2 and DSS1 promote HR through their interaction with BRCA2 is not well understood.

In addition to its role in promoting Rad51 filament formation on RPA-coated ssDNA, Rad52 plays a second role during HR in yeast and possibly also in humans. Yeast and human RAD52 are unique in their ability to mediate the annealing of complementary ssDNA that is fully coated by RPA, displaying specificity for the cognate RPA (Mortensen *et al.*, 1996; Sugiyama *et al.*, 1998; Grimme *et al.*, 2010). This reaction is critical for SSA (Figure 1) as well as for second end capture in the double Holliday junctions (dHJ) pathway (Figure 1, step 5a) and second end annealing in the synthesis-dependent strand annealing (SDSA) pathway (Figure 1, step 5b), as discussed in more detail below.

Another group of proteins implicated in the nucleation and stability of Rad51-ssDNA filaments are the Rad51 paralogs and their binding partners (Table 1; Liu *et al.*, 2011a). Budding yeast contains two protein complexes with Rad51 paralogs: the hetero-dimeric Rad55–Rad57 complex and the Shu complex consisting of two Rad51 paralogs, Psy3–Csm2, and their binding partners Shu1 and Shu2. Rad55 and Rad57 share the Walker A and B boxes with Rad51 enabling the hetero-dimer to bind and hydrolyze ATP. Rad55–Rad57 allows assembly of Rad51 filaments on RPA-coated ssDNA (Sung, 1997b). The homology of Psy3–Csm2 with Rad51 was only appreciated through the high-resolution crystal structure (Sasanuma *et al.*, 2013). These proteins lack ATP-binding motifs and the ability to bind or hydrolyze ATP. In humans, five RAD51 paralogs have been identified so far, but the sequence homology is insufficient to assign clear correspondence between XRCC2, XRCC3, RAD51B, RAD51C, or RAD51D with the yeast Rad51 paralogs (Liu *et al.*, 2011a). SWS1, a possible homolog of yeast Shu2, and its binding partner SWSAP1 associate with RAD51D in human cells. The Rad51 paralogs display distinct non-overlapping functions in Rad51 filament formation and/or stability, but their precise function is incompletely understood. The Rad55–Rad57 complex forms a co-filament with Rad51 possibly capping Rad51-ssDNA filaments and insulating them from disassembly by the Srs2 anti-recombinase

Table 1 Homologous recombination proteins in *Saccharomyces cerevisiae* and *Homo sapiens*

HR phase	Saccharomyces cerevisiae	Homo sapiens	Function
DSB processing	Dna2	DNA2	Nuclease/helicase
	Mre11–Rad50–Xrs2	MRE11–RAD50–NBS1	Nuclease, end-binding
	Sae2	CtIP	Nuclease?, cofactor?
	Exo1	EXO1	Nuclease
	Sgs1–Top3–Rmi1	BLM –TOPBP1a–RMI1–RMI2	Helicase/motor, type IA topoisomerase
Rad51 filament formation, homology search, and DNA strand exchange	Rad51	RAD51	Homology search
	RPA	RPA	ssDNA-binding protein
	Rad55–Rad57	XRCC2/3	Rad51 paralogs
	Psy3–Csm2	RAD51B/C/D	Rad51 paralogs
	Shu1–Shu2	SWS1–SWSAP1	Bind Rad51 paralogs
	Rad52	RAD52	Strand annealing
		BRCA2 ^a /DSS1	Mediator
		PALB2 ^a	Mediator
		RAD51AP1	Binds Rad51
	Hop2–Mnd1 ^b	HOP2–MND1	Binds Rad51 filament
	Rad54	RAD54	dsDNA translocase
	Rdh54	RAD54B	dsDNA translocase
Branch migration	Rad54	RAD54	dsDNA translocase
	Rdh54 (Tid1)	RAD54B	dsDNA translocase
	Sgs1	BLM	Helicase
Junction processing	Mus81–Mms4	MUS81–EME1/2	Nuclease
	Slx1–Slx4	SLX1– SLX4 ^a	Nuclease
	Yen1	GEN1	Nuclease
	Sgs1–Top3–Rmi1	BLM –TOPBP1a–RMI1–RMI2	Helicase/motor, type IA topoisomerase

^aFanconi anemia syndrome gene, for details see text.

^bHop2–Mnd1 are meiosis-specific in *S. cerevisiae*.

Note: The factors emphasized by the brackets represent a group of Rad51 paralogs, which due to their shared homologies are difficult to assign to specific homolog pairs in yeast and humans. Confirmed tumor suppressor proteins are highlighted in bold face.

Source: Modified from San Filippo, J., Sung, P., Klein, H., 2008. Mechanism of eukaryotic homologous recombination. Annual Review of Biochemistry 77, 229–257 and Symington, L.S., Rothstein, R., Lisby, M., 2014. Mechanisms and regulation of mitotic recombination in *Saccharomyces cerevisiae*. Genetics 198, 795–835.

(Liu *et al.*, 2011b). These data suggest that the structure of the Rad51-ssDNA filament conducting homology search and DNA strand invasion is considerably more complex than a simple homotypic filament of Rad51 protomers.

HOP2-MND1 is another mediator, forming a stable heterodimer required for HR in meiotic and mitotic cells in humans, whereas these proteins are meiosis-specific in yeast (Zierhut *et al.*, 2004). *hop2* knockout mice exhibit normal RAD51 and DMC1 (the meiosis-specific RAD51 paralog) foci formation; however, these foci persist much longer, suggesting that HOP2-MND1 act after filament assembly (Petukhova *et al.*, 2003).

RAD51-Associated Protein 1 (RAD51AP1, Table 1) physically interacts with RAD51 and DMC1, and is required for recombination in vertebrates but is not present in yeast. Depletion of RAD51AP1 does not affect the formation of RAD51 foci in human cells suggesting a function after filament assembly (Modesti *et al.*, 2007; Wiese *et al.*, 2007).

DNA Strand Invasion and Associated DNA Synthesis

Rad54 is a dsDNA motor protein functioning in conjunction with Rad51 during homology search and DNA strand invasion (Table 1; Figure 1, step 3; Petukhova *et al.*, 1998). Rad54

associates with and stabilizes the Rad51-ssDNA filament, which delivers Rad54 to the pairing site (Mazin *et al.*, 2000). In yeast, Rad54 is indispensable for 'D-loop' formation *in vitro* and the *in vivo* DNA repair phenotypes of *rad54* mutants are equal to those of *rad51* mutants (Petukhova *et al.*, 1998). In mammals, however, RAD51 is an essential gene, whereas RAD54 is not (Lim and Hasty, 1996; Essers *et al.*, 1997). This might be related to a role of mammalian RAD51 in stabilizing replication forks from degradation by MRE11, a function that is not shared with RAD54 (Schlachter *et al.*, 2011). Rad54 is a potent ATP-driven dsDNA translocase that disrupts dsDNA-Rad51 filaments (Solinger *et al.*, 2002). Importantly, during DNA strand invasion, Rad51 switches from binding the initial ssDNA to *hDNA*, i.e., duplex DNA (Figure 1, steps 2–3). Rad54 pumps the Rad51-bound duplex DNA to extend the nascent *hDNA* while simultaneously dissociating Rad51, thus allowing access of DNA polymerases to the invading 3'-end (Figure 1, steps 3–4; Wright and Heyer, 2014). Rad54 may play additional roles in remodeling nucleosomes during HR (Alexeev *et al.*, 2003). Rad54 paralogs in yeast (Rdh54) and humans (RAD54B) exert highly similar dsDNA-dependent ATPase activities exerting likely a similar, somewhat overlapping function in conjunction with Rad51 and Dmc1, the meiotic Rad51 paralog (Tanaka *et al.*, 2000; Holzen *et al.*, 2006).

Extension of the invading 3'-end in the nascent 'D-loop' by DNA polymerase (Figure 1, step 4) initiates recombination-associated DNA repair synthesis. The reaction is conceptually similar to leading strand synthesis during DNA replication but involves displacement of a contiguous DNA strand. Yeast and human DNA polymerase δ holoenzyme in conjunction with the PCNA (proliferating cell nuclear antigen) processivity clamp and the RFC (replication factor C) clamploader are highly capable of performing this specific type of DNA synthesis in reconstituted *in vitro* reactions (Li *et al.*, 2009; Sneed *et al.*, 2013; Wang *et al.*, 2004). Genetic evidence in yeast is consistent with the lagging strand DNA polymerase δ being the primary polymerase for first end DNA synthesis during HR (Li *et al.*, 2009; Sneed *et al.*, 2013; Wang *et al.*, 2004). It will be of interest to explore the potential contributions of other DNA polymerases, including translesion synthesis (TLS) polymerases, in D-loop extension.

The common core reactions of HR include DNA processing to produce ssDNA and the subsequent formation of the Rad51 nucleoprotein filament, which conducts homology search and DNA strand invasion. The resulting 'D-loop' contains the invading 3'-OH end, which primes HR-associated DNA repair synthesis, restoring sequence information lost at the break site. Processing of the extended 'D-loop' occurs in at least three distinct sub-pathways of HR, which determine the repair outcome of the event (Figure 1, after step 4), as discussed below.

Recombination and Double-Strand Break Repair

DSB repair is critical to maintain genome stability and can be achieved by three distinct pathways: End-joining, single-strand annealing (SSA), and HR (Figure 1). DNA end-joining, either the classic nonhomologous end-joining pathway (NHEJ) or the microhomology-mediated end-joining pathway (MMEJ), also called alternative end-joining (aEJ) in humans, requires no or minimal resection to restore contiguous chromosomes (Lieber, 2010). Depending on the exact chemical structure of the original DSB, end-joining is of variable fidelity but often involves deletions. SSA can occur when direct repeated DNA sequences flank the DSB. The process involves DSB resection, annealing of the repeated DNA sequences, flap-trimming, and gap repair by DNA synthesis and ligation to restore contiguous chromosomes. The process is inherently error-prone, deleting one repeat and the intervening sequence. The distinction between SSA, MMEJ, and NHEJ relates to different protein factors involved and the extent of homology ranging from > 25 bp for SSA to as little as 5 bp in MMEJ, whereas NHEJ processes ends with 0–4 bp homology.

The focus of this article is HR. The central HR reactions have been illustrated in the context of DSB repair in the preceding section up to the point, where the various HR sub-pathways diverge in the processing of the extended 'D-loop' (Figure 1, step 4). Below we discuss the ensuing steps.

Double Holliday Junction Formation, Resolution, and Dissolution

dHJ, a key intermediate in meiotic recombination, also occur during somatic DNA repair involving primarily sister

chromatids (Bzymek *et al.*, 2010). The displaced strand in the extended 'D-loop' captures the second end by DNA strand annealing involving Rad52 (Figure 1, step 5a) to generate a dHJ by DNA synthesis (Step 6a). It is unclear whether all strands are fully ligated before further processing. dHJs can be resolved by structure-selective DNA endonucleases into 'crossover' and 'noncrossover' products (steps 7a–8a, 7b–8b). A candidate protein is Yen1 (human GEN1, Table 1), a nuclease that is capable of resolving Holliday junctions (HJs) with fully ligated strands (Matos *et al.*, 2011). Yen1 is activated in the cell cycle just around the time of chromosome segregation in anaphase, and likely represents a failsafe mechanism of unresolved joint molecules. Yen1 mutants do not display an overt HR phenotype in otherwise wild type cells. In contrast to Yen1/GEN1, the structure-selective endonuclease Mus81–Mms4 is very inefficient in cleaving HJs or dHJs when all strands are ligated, suggesting that it cleaves earlier intermediates before ligation (Ehmsen and Heyer, 2008). Mutants in the Mus81–Mms4 complex show defects in 'crossover' formation in somatic DSB repair, confirming the involvement of this endonuclease in HR (Ho *et al.*, 2010). Slx1–Slx4, another structure-selective nuclease, appears to play no obvious role in HR in budding yeast, based on the mutant phenotype, but has been implicated in HJ/dHJ processing in mammalian cells (Fricke and Brill, 2003; Castor *et al.*, 2013). Seminal work with the human BLM–TOPOIII α –RMI1 uncovered a novel mechanism to process dHJs by a mechanism termed dissolution (Wu and Hickson, 2003). The significance is that dHJ dissolution leads only to 'noncrossover' outcome. This mechanism provides a satisfying interpretation for the hyper-sister chromatid exchange phenotype of BLM-deficient human cells as an inability to dissolve dHJs, which instead give rise to 'crossovers' by resolution. dHJ dissolution is mechanistically complex and requires the coordinate movement of both HJs in the dHJ toward each other by the BLM helicase (step 7c). The topological stress accumulating between the two individual HJs is addressed by TOPOIII α and its RMI1 cofactor. The final hemi-catenane intermediate (step 8c) is an excellent substrate for decatenation by Top3–Rmi1 (Cejka *et al.*, 2010b). The role of the RMI2 cofactor in mammals remains to be determined. The structure and mechanism of the BLM–TOPOIII α –RMI1 complex is highly conserved in eukaryotes (Table 1), and it is possible that this complex also targets other HR-dependent joint molecules for dissolution.

Synthesis-Dependent Strand Annealing

SDSA involves alternative processing of the extended 'D-loop' by dissolution and annealing of the newly synthesized strand with the ssDNA tail of the second end of the frank (two-sided) DSB (Figure 1, step 5b). The candidate proteins for dissolution of the extended D-loop are listed in Table 2 and discussed later. The annealing step is likely catalyzed by Rad52, the protein capable of annealing complementary ssDNA fully coated by RPA. DNA synthesis and ligation are required to restore contiguous chromosomes. The DNA synthesis (step 6b), as well as DNA synthesis from the second end in the dHJ sub-pathway (step 6a) is simpler than extension of the first end in the 'D-loop' as no displacement synthesis is involved. It is presently unclear which DNA polymerases are involved in

Table 2 Homologous recombination proteins in *Saccharomyces cerevisiae* and *Homo sapiens* with anti-recombination and anti-crossover functions

HR role	<i>Saccharomyces cerevisiae</i>	<i>Homo sapiens</i>	Function
Anti-rec	Srs2	FBH1, PARI BLM, RECQ5	Dissociation Rad51-ssDNA filament
	Sgs1	FANCI	Dissociation of nascent D-loop
	Top3-Rmi1	BLM TOPOIII α -RMI1-RMI2	
Anti-crossover	Srs2	BLM, RECQ1 RTEL1	Dissociation of extended D-loop
	Mph1	FANCM	dHJ dissolution
	Sgs1-Top3-Rmi1	BLM-TOPOIII α -RMI1-RMI2	

Note: Antirecombination (Anti-rec) refers to activities that lead to an overall decrease in recombination repair products. Anti-crossover refers to activities that specifically block formation of crossover products.

Source: Modified from Heyer, W.D., Ehmsen, K.T., Liu, J., 2010. Regulation of homologous recombination in eukaryotes. Annual Review of Genetics 44, 113–139.

this step. Importantly, SDSA always leads to a ‘noncrossover’ outcome of DSB repair.

Break-Induced Replication

Break-induced replication (BIR) engages the core mechanism of HR to establish a noncanonical replication structure that copies the entire chromosome arm in the absence of a second DSB end in a conservative manner (Figure 1, steps 5c–6c; Malkova *et al.*, 1996). BIR is formally analogous to one-sided DSB repair, but BIR may occur outside the S phase context emblematic for one-sided DSB repair (Figure 2). BIR has also been invoked as the conceptual mechanism for the alternative lengthening of telomeres (ALT) in cancer cells to maintain telomeres in the absence of telomerase. It is presently unclear what distinguishes BIR from the other HR sub-pathways, in particular how the absence of the second end can be sensed. While BIR can copy an entire chromosome arm involving the components of a regular replication fork, the associated DNA synthesis is significantly more error-prone than DNA replication (Malkova and Haber, 2012). BIR leads to loss-of-heterozygosity distal to the DSB site when using the homologous chromosomes as a template, or to nonreciprocal translocations when copying from a nonallelic locus.

HR achieves high-fidelity repair of DSBs in comparison to alternative mechanisms such as SSA, NHEJ, or MMEJ. However, the DSB repair by HR is not error-free and a 1000-fold elevation in mutation frequencies close to DSBs repaired by HR have been recorded with even larger mutagenic increases in BIR-mediated events (Malkova and Haber, 2012). Foremost, dispersed repeats present throughout genomes provide opportunity for nonallelic recombination, with the potential to generate large-scale chromosomal rearrangements.

Recombination and DNA Replication

Replication of DNA is central to life and tightly regulated to ensure genomic stability. Changes to the nucleotide pools or DNA lesions may lead to stalling of the replicative DNA polymerases and/or replicative helicase (Figure 2). Even in the

absence of DNA damage, repetitive DNA elements, DNA secondary structures, transcription complexes, RNA–DNA hybrids, and fragile sites can lead to slowing or stalling of DNA replication (Zeman and Cimprich, 2014). Figure 2 illustrates how HR serves as a repair or tolerance mechanism in the recovery of stalled or broken replication forks.

Multiple HR, fork remodeling, and TLS factors are involved in the recovery of stalled replication forks. Stalled replication forks can be cleaved, resulting in a one-sided DSB to initiate HR using the sister chromatid as a template (Figure 2, steps 2a, 3a, 4b, 5a). Repair of one-sided DSBs by HR is crucial, as events mediated by any end-joining pathway would result in associated chromosomal rearrangements (Figure 2, step 4a; Pace *et al.*, 2010; Adamo *et al.*, 2010). Upon prolonged replication stress and fork stalling by extended hydroxyurea (HU) or camptothecin (CPT) treatment (Hanada *et al.*, 2007; Regairaz *et al.*, 2011) cells accumulate DSBs in a Mus81–Mms4/EME1-dependent fashion. These DSBs likely reflect direct cleavage of the stalled fork by Mus81–Mms4/EME1, as DNA substrates mimicking stalled replication forks are among the best substrates for its endonuclease activity (Ehmsen and Heyer, 2008; Ciccio *et al.*, 2003). In all organisms tested, cells lacking *MUS81* or *EME1* orthologs are sensitive to replication blocking agents, but not IR which causes frank (two-sided) DSBs (Dendouga *et al.*, 2005; Ciccio *et al.*, 2008). ‘Resolution’ of a single HJ intermediate that arises in the process is necessary before chromosome segregation (Figure 2, steps 4a–5a).

Alternatively, DNA damage may be tolerated instead of repaired to ensure completion of DNA replication and hence survival (Figure 2). When the 3′ end of the leading strand is blocked, fork regression can cause a template switch to the nascent sister chromatid, forming the chicken-foot intermediate which is analogous to an HJ (Figure 2, steps 3b). Direct visualization provided evidence for reversed forks in checkpoint-proficient mammalian cells after exposure to sublethal doses of topoisomerase 1 inhibitors, camptothecin, and topotecan (Chaudhuri *et al.*, 2012). DNA synthesis on the sister chromatid (step 4c) allows bypass of the lesion after reversing the chicken-foot structure (step 5b). Several potential factors have been identified, including the DNA motor proteins Rad5/HLTF, FANCM, SMARCA1, and ZRANB3 that may

be involved in fork and/or chicken-foot reversal (Zeman and Cimprich, 2014). This mechanism is likely independent of RAD51 as no DNA strand invasion is required, but RAD52-mediated strand annealing may be involved.

DNA lesions that block leading or lagging strand synthesis can be tolerated through S phase by gap repair (Figure 2, steps 3c; Zeman and Cimprich, 2014) (Note that Figure 2 only represents leading strand blockage but the following mechanisms described can also occur if there is a lagging strand block.). Gap repair can ensue in two ways. During gap invasion, the uninterrupted strand forms a Rad51 filament to invade the intact sister chromatid (steps 4d, 5c). Following joint migration, the displaced strand of the joint molecule serves as the template for DNA synthesis. Since the invading DNA strand has no free ends, this type of DNA strand invasion poses specific topological problems. Alternatively, during end invasion the leading strand can invade the sister chromatid template for DNA synthesis past the lesion (step 4e). End invasion requires a helicase to generate a ssDNA template for Rad51 filament assembly. The gap is closed after withdrawal of the newly synthesized strand (step 5d), analogous to SDSA (Figure 1). Inherent in the depiction of gap repair in Figure 2 is restart of leading strand and lagging strand (not shown) replication. While restart of lagging strand replication is considered intrinsic to regular DNA synthesis, it has only been recently realized that also restart of leading strand replication is common (Lopes *et al.*, 2006).

Finally, lesions blocking the replicative DNA polymerases can be bypassed by a polymerase switch to TLS polymerases (Figure 2, steps 3d, 4f; Goodman and Woodgate, 2013). These specialized low fidelity DNA polymerases lack proofreading activity and contain enlarged active sites to accommodate damaged bases. Bypass is likely achieved by a combination of specialized TLS polymerases that have the ability to add a base opposite the damaged template and other TLS polymerases than can extend the unusual primer template structure before switching back to the replicative polymerase (Prakash and Prakash, 2002).

Pathway choice at stalled replication forks is an important issue and posttranslational modification of the processivity clamp PCNA serves as the nexus to regulate fork regression, HR-mediated template-switch, and translesion DNA synthesis (Moldovan *et al.*, 2007). Mono-ubiquitylation of K164 on PCNA promotes DNA polymerase switching from replicative DNA polymerases to TLS polymerases, which have, in addition to a PCNA-interacting motif, a distinct ubiquitin-interaction motif that enhances their affinity to ubiquitylated PCNA. Switchback to replicative DNA polymerases may involve deubiquitylation or unloading of modified PCNA. Poly-ubiquitylation of K164 on PCNA involving K63-linked ubiquitin chains favors fork regression, potentially by providing a specific binding site for fork regression motors such as ZRANB3. Instead, sumoylation of the same K164 residue on PCNA recruits the anti-recombinogenic helicase Srs2 that dissociates Rad51 from ssDNA and effectively inhibits HR (see below, Table 2).

Processing of replication forks and collisions of replication forks with active transcription complexes (see Aguilera and Garcia-Muse, 2013) are a major source of genomic instability. HR competes with TLS and fork regression to process stalled

forks or complete DNA replication by gap repair to suppress genomic rearrangements and enable complete and faithful replication of the genome.

Recombination and Interstrand Cross-Link Repair

DNA ICLs are cytotoxic lesions that covalently link the Watson and Crick strands of duplex DNA effectively blocking replication and transcription. Exogenous ICL-inducing agents include nitrogen mustards, psoralens, mitomycin C, and platinum-like cisplatin that are utilized as anticancer therapeutics (Deans and West, 2011). Nitrous acids that form as by-products of nitrite metabolism and aldehydes from alcohol or fat metabolism are endogenous sources of ICLs (Deans and West, 2011; Langevin *et al.*, 2011). Figure 3 illustrates two models for how HR participates in the repair of ICLs initiated by stalling of one or two converging replication forks.

Sensitivity to ICL agents is a hallmark of cells lacking the Fanconi anemia (FA) pathway. FA is a rare genetic disorder that is characterized by childhood onset of aplastic anemia, bone marrow failure, cancer, and leukemia susceptibility, and cellular hypersensitivity to ICL agents (Kim and D'Andrea, 2012). Cells derived from FA patients show increased chromosome breakage and radial chromosomes after exposure to ICL-inducing agents, likely because DSBs left unrepaired can fuse with other chromosome arms through NHEJ (Figure 2, step 4a). The FA pathway may function to actively stabilize the replication fork to allow for HR-mediated repair and suppress NHEJ to decrease genomic instability (Pace *et al.*, 2010; Adamo *et al.*, 2010).

To date, identification of 17 FA genes has increased the molecular understanding of this disease (Wang and Smogorzewska, 2015). Eight proteins (FANCA, B, C, E, F, G, L, and FANCM/FAAP24) form the FA core complex that is responsible for recognizing a cross-link in DNA. FANCM is an ATPase and translocase that is essential for recognizing the ICL at the stalled replication fork, recruiting the FA core complex to the covalently linked duplex DNA. FANCD2 and FANCI are mono-ubiquitylated by FANCL and this serves as a reversible signal that causes these proteins to localize to DNA damage. The other FANCA genes encode proteins factors that appear to be more involved in the ensuing DNA transactions (Deans and West, 2011). FANCI/BRIP1 is a helicase and ATPase that has been shown to bind to BRCA1 (FANCS), another important factor involved in genomic stability. FANCD1 is identical to BRCA2, the central HR mediator protein in mammals described above (Howlett *et al.*, 2002). The BRCA2-interacting partner PALB2 was identified as FANCN, whereas the RAD51C protein as FANCO. These findings suggest that possibly additional HR proteins (see Table 2) might be involved in FA, as the recent identification of the HR nuclease scaffold protein SLX4 as FANCP has already shown (Kim *et al.*, 2011).

ICLs are most directly recognized genome-wide during S phase because they stall replication forks, as illustrated in Figure 3. The FA core complex recognizes ICLs that stall replication forks and signals downstream repair factors including the aforementioned FANCA proteins that act on DNA. Stalling of transcription complexes offers an alternative pathway of ICL recognition, independent of DNA replication. In S phase, ICL

repair is initiated by ubiquitylated FANCI–FANCD2 and involves factors from Nucleotide Excision Repair (NER), HR, and TLS DNA polymerases. DSBs have been identified as a major reaction intermediate and substrate for HR during ICL repair (Figure 3; Long *et al.*, 2011; Knipscheer *et al.*, 2009). A replication fork approaching an ICL from one side, for example near the telomere or in mid/early S phase, leads to a one-sided DSB after cleavage of the stalled fork, which can be repaired by HR. Resection of the DSB permits RAD51 filament formation for invasion of the intact sister to create a ‘D-loop’ (Figure 3, steps 3a, 4a). After the initial extension at the ‘D-loop,’ an additional nick on the other side of the ICL unhooks the damaged base (step 5a). A polymerase switch to a TLS polymerase to synthesize across the lesions and subsequent excision of the lesion by NER (Figure 3, step 6a) restores the chromosome potentially leading to reestablishment of the replication fork (Figure 3, step 7a; Deans and West, 2011; Zhang and Walter, 2014).

The mechanism of repair of an ICL flanked by two converging replication forks, as for example in late S phase, has been elucidated in an elegant system using cell-free *Xenopus laevis* egg extracts and a plasmid substrate with a single site-specific ICL. When the replicative polymerase encounters the ICL, leading strand synthesis stalls 20–40 nt away from the ICL (Figure 3, step 2b; Raschle *et al.*, 2008). After replisome remodeling, successive nicks on either side of the ICL unhook the damaged template (step 3b) and a polymerase switch recruits TLS polymerases (step 4b). Although the specificity and molecular mechanism of recruitment of the low fidelity DNA polymerase is not well understood, REV1–REV3–REV7 (REV1–Pol ζ) is a candidate complex (Kim and D’Andrea, 2012). The initial synthesis up to one nucleotide from the ICL (step 3b) is followed by synthesis across the lesion (steps 4b). NER removes the ICL (step 5b), restoring a proper template for DSB repair by HR (steps 6b, 7b; Raschle *et al.*, 2008).

Six nucleases with varying substrate specificity for joint molecule DNA substrates have been implicated in processing ICLs because mutations in corresponding genes increase sensitivity to ICL agents (Zhang and Walter, 2014; Kim and D’Andrea, 2012; Deans and West, 2011). SNM1A and SNM1B are exonucleases that may function downstream of the incisions instead of unhooking. Another nuclease, FAN1, was identified by its interaction with FANCD2 but the physiological relevance is unclear, as *FAN1* deficiency does not lead to an FA-like phenotype. MUS81–EME1 and SLX1 cleave nicked DNA joint molecules that could arise when a replication fork encounters an ICL; however, even though human MUS81–EME1 or SLX1 patients have not been described, deficiency of animal models in either endonuclease does not phenocopy FA (Dendouga *et al.*, 2005; Castor *et al.*, 2013). These endonucleases may have partially overlapping roles in ICL repair. Lastly, mutations in the endonuclease XPF (FANCD2)–ERCC1 lead to FA, and XPF–ERCC1 have been shown to unhook ICLs *in vivo* and *in vitro* (Zhang and Walter, 2014). Intriguingly, the SLX4 (FANCD2) scaffold protein interacts with XPF–ERCC1, MUS81–EME1, as well as SLX1, and SLX4 deficiency causes FA (Fekairi *et al.*, 2009). This raises the possibility that all three nucleases could be recruited to the ICL for unhooking to allow for repair of the damaged template. It is unclear what role FANCD2 plays in this context and whether

its ability to dissociate RAD51 from ssDNA is important for ICL repair.

Taken together, the FA pathway shows extraordinary complexity as it detects ICLs in replicating cells and coordinates the interplay between HR, certain NER factors, and TLS polymerases to repair ICLs and suppress chromosomal rearrangements induced by NHEJ.

Pathway Regulation and Crossover Control

Several layers of regulation operate throughout the HR pathway targeting multiple proteins and DNA intermediates. The sister chromatid is the preferred template, and HR is largely limited to the S and G2 phases of the cell cycle. The major control step to limit HR during DSB repair in the G1 phase of the cell cycle is to regulate DSB end resection involving the cyclin-dependent kinases targeting Sae2/CtIP, elaborate histone modifications, and the mammalian 53BP1 protein (Daley and Sung, 2014; Symington and Gautier, 2011). Inhibition of DSB resection favors end-joining pathways (Figure 1), which involve no or little end resection.

An additional target of regulation is the Rad51–ssDNA filament (Figure 1, steps 2 to 1). Seminal work on the budding yeast anti-recombinase Srs2 established a mechanism, by which Srs2 translocates on ssDNA to dissociate Rad51–ssDNA filaments (Table 2; Veaute *et al.*, 2003; Krejci *et al.*, 2003). Four motor proteins, FBH1, BLM, RECQ5, and FANCD2, as well as a non-motor protein, PARI, have been postulated to exert this function in mammalian cells (Table 2). The diversity of anti-recombinases may suggest functional specialization as well as cooperation between motor proteins and PARI, which lacks a motor domain and may target motor proteins to RAD51. Rad51 filament dissociation by Srs2 is regulated by Rad51 paralogs in *Saccharomyces cerevisiae* (Liu *et al.*, 2011b), and the mammalian RAD51 paralogs may play similar roles.

A final tier of HR regulation dissociates the nascent ‘D-loops’ (Figure 1, steps 3 to 2), a process that has been called heteroduplex rejection (Hombauer *et al.*, 2011). This quality control mechanism employs components of the mismatch repair pathway and the Sgs1–Top3–Rmi1 complex to abort DNA strand invasions, where the template is not fully homologous with the invading DNA to prevent recombination between nonallelic sites (Table 2). The exact mechanism still needs to be established, but in order to be biologically effective as an anti-recombination mechanism, heteroduplex rejection should engage before DNA polymerase extends the ‘D-loop’ (Figure 1). In sum, HR is regulated during at least three distinct phases, DSB end resection, Rad51 filament, and the nascent ‘D-loop’ to ensure proper DSB repair pathway choice and enforce quality control.

HR can result in ‘crossover,’ the reciprocal exchange of chromosome arms flanking the event, which manifest themselves as sister chromatid exchanges during somatic DNA repair. In this context, ‘crossovers’ are actively avoided by at least two distinct mechanisms (Table 2). As discussed above, the BLM–TOPO3 α –RMI1 (yeast Sgs1–Top3–Rmi1) complex dissolves dHJ (Figure 1, steps 7c, 8c; Table 2) as a mechanism to finalize HR events as ‘noncrossover’ events (Wu and Hickson, 2003; Cejka *et al.*, 2010b). A second ‘anti-crossover’ mechanism

is represented by the SDSA sub-pathway of HR (Figure 1, steps 5b, 6b). Dissociation of the 'D-loop' after extension by DNA polymerases leads to reannealing of the newly synthesized strand with the second DSB end, which was not involved in the DNA strand invasion. A number of motor proteins, Srs2 and Mph1 in yeast, as well as BLM, RECQ1, FANCM, RTEL1 in mammals have been shown or suggested to dissolve 'D-loops' *in vitro* and postulated to do so in cells (see Table 2 and Further Reading). Genetic data are consistent with an anti-crossover role for these proteins, although in the case of BLM, dHJ dissolution may be the primary mechanism involved. The removal of Rad51 on the second end to allow annealing may be an additional requirement for SDSA. The specificity of the various motor proteins for a specific 'D-loop' structure (nascent or extended), in particular their relationship to 'D-loops' being extended by DNA polymerase, still needs to be established.

In conclusion, while meiotic recombination is geared to generate 'crossovers' that physically connect the two homologous chromosomes during chromosome segregation, HR during somatic DNA repair engages at least two distinct mechanisms to actively avoid 'crossover' formation between participating sister chromatids.

HR and Human Disease

Maintaining genomic stability plays a pivotal role in preventing human disease (Moynahan and Jasin, 2010). The human tumor suppressor protein BRCA2 illustrates most strikingly the connection between HR, human disease, and cancer (Moynahan *et al.*, 2001). The molecular function of BRCA2 is best understood as a mediator protein to nucleate Rad51 filaments, as discussed above. Rad51 filaments protect stalled forks from degradation by MRE11 and also perform homology search and DNA strand invasion during HR (Schlachter *et al.*, 2011; Jensen *et al.*, 2010). Both aspects may be distinct functions to maintain genomic stability. Bi-allelic mutations in BRCA2 cause FA (Howlett *et al.*, 2002), a complex human syndrome discussed more in detail above. Heterozygous BRCA2 mutants carry a greatly enhanced risk for cancer, especially for breast and ovarian cancer. Consistently, the BRCA2-interacting protein PALB2 has also been identified as a FANCD1 gene (FANCD1) and tumor suppressor (Reid *et al.*, 2007). Considering our knowledge of the HR pathway (Figure 1), more candidate genes may be involved at the nexus of HR, human disease and cancer, as already recognized for the RAD51 paralogs RAD51B/C/D, SLX4, and BLM (see Table 1).

HR has dual significance for cancer. As discussed above, the involvement of HR in the etiology of cancer is well documented (Moynahan and Jasin, 2010). In addition, HR addresses many types of DNA damage induced by current modalities of anticancer therapy. These include DSBs induced directly by ionizing radiation or drugs as well as indirectly by targeting topoisomerase II, topoisomerase I-targeted drugs, ICL-inducing drugs, as well as drugs interfering directly or indirectly with DNA replication. This offers the prospect of exploiting genetic or pharmacological approaches to disable HR to sensitize tumor cells to such DNA damage-based therapies (Powell and Bindra, 2009; Carvalho and Kanaar, 2014). Pioneering work with BRCA-deficient cells and tumors

exploits the synthetic lethality between an HR-defect and inhibition of the poly-ADP ribose polymerase (PARP) as a strategy to treat HR-deficient cancers (Bryant *et al.*, 2005; Farmer *et al.*, 2005). These exciting developments have given the HR field an additional purpose.

Conclusions

- HR comprises a web of interrelated pathways supporting genomic stability through their function to repair or tolerate DNA damage.
- HR is highly regulated to avoid recombination in the G1 phase of the cell cycle and revert inappropriate formation of Rad51-ssDNA filaments or DNA strand invasions.
- The full extent of the involvement of HR genes in cancer predisposition and possibly other human diseases still needs to be established.
- The role of HR in DNA damage-based modalities of cancer treatment is beginning to be appreciated and therapeutically exploited.

Acknowledgments

Work in the author's laboratory was supported by the NIH (GM58015, CA92276, CA154920), and the DOD (BC133980). SM was partially supported by a predoctoral fellowship from the Tobacco-Related Disease Research Program (20DT-0036) and the NIGMS T32 Pharmacology: Bench to Bedside T32 training grant (GM099608). SSJ was partially supported by the T32 Training Grant Molecular and Cellular Biology (GM007377).

See also: Nucleic Acid Synthesis/Breakdown: DNA Synthesis/Repair: Nonhomologous DNA End Joining

References

- Adamo, A., Collis, S.J., Adelman, C.A., *et al.*, 2010. Preventing nonhomologous end joining suppresses DNA repair defects of fanconi anemia. *Molecular Cell* 39, 25–35.
- Aguilera, A., Garcia-Muse, T., 2013. Causes of genome instability. *Annual Review of Genetics* 47, 1–32.
- Alexeev, A., Mazin, A., Kowalczykowski, S.C., 2003. Rad54 protein possesses chromatin-remodeling activity stimulated by a Rad51-ssDNA nucleoprotein filament. *Nature Structural Biology* 10, 182–186.
- Bryant, H.E., Schultz, N., Thomas, H.D., *et al.*, 2005. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* 434, 913–917.
- Bzymek, M., Thayer, N.H., Oh, S.D., Kleckner, N., Hunter, N., 2010. Double Holliday junctions are intermediates of DNA break repair. *Nature* 464, 937–941.
- Carvalho, J.F.S., Kanaar, R., 2014. Targeting homologous recombination-mediated DNA repair in cancer. *Expert Opinion on Therapeutic Targets* 18, 427–458.
- Castor, D., Nair, N., Declais, A.C., *et al.*, 2013. Cooperative control of holliday junction resolution and DNA repair by the SLX1 and MUS81-EME1 nucleases. *Molecular Cell* 52, 221–233.
- Cejka, P., Cannavo, E., Polaczek, P., *et al.*, 2010a. DNA end resection by Dna2-Sgs1-RPA and its stimulation by Top3-Rmi1 and Mre11-Rad50-Xrs2. *Nature* 467, 112–116.

- Cejka, P., Plank, J.L., Bachrati, C.Z., Hickson, I.D., Kowalczykowski, S.C., 2010b. Rrm1 stimulates decatenation of double Holliday junctions during dissolution by Sgs1-Top3. *Nature Structural & Molecular Biology* 17, 1377–1382.
- Chen, Z., Yang, H., Pavletich, N.P., 2008. Mechanism of homologous recombination from the RecA-ssDNA/dsDNA structures. *Nature* 453, 489–494.
- Ciccia, A., Constantinou, A., West, S.C., 2003. Identification and characterization of the human Mus81/Eme1 endonuclease. *Journal of Biological Chemistry* 278, 25172–25178.
- Ciccia, A., McDonald, N., West, S.C., 2008. Structural and functional relationships of the XPF/MUS81 family of proteins. *Annual Review of Biochemistry* 77, 259–287.
- Conway, A.B., Lynch, T.W., Zhang, Y., *et al.*, 2004. Crystal structure of a Rad51 filament. *Nature Structural & Molecular Biology* 11, 791–796.
- Daley, J.M., Sung, P., 2014. 53BP1, BRCA1, and the choice between recombination and end joining at DNA double-strand breaks. *Molecular and Cellular Biology* 34, 1380–1388.
- Deans, A.J., West, S.C., 2011. DNA interstrand crosslink repair and cancer. *Nature Reviews Cancer* 11, 467–480.
- Dendouga, N., Gao, H., Moechars, D., *et al.*, 2005. Disruption of murine Mus81 increases genomic instability and DNA damage sensitivity but does not promote tumorigenesis. *Molecular and Cellular Biology* 25, 7569–7579.
- Ehmsen, K.T., Heyer, W.D., 2008. *Saccharomyces cerevisiae* Mus81-Mms4 is a catalytic structure-selective endonuclease. *Nucleic Acids Research* 36, 2182–2195.
- Essers, J., Hendriks, R.W., Swagemakers, S.M.A., *et al.*, 1997. Disruption of mouse *RAD54* reduces ionizing radiation resistance. *Cell* 89, 195–204.
- Farmer, H., McCabe, N., Lord, C.J., *et al.*, 2005. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. *Nature* 434, 917–921.
- Fekairi, S., Scaglione, S., Chahwan, C., *et al.*, 2009. Human SLX4 is a holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell* 138, 78–89.
- Feng, Z.H., Scott, S.P., Bussen, W., *et al.*, 2010. Rad52 inactivation is synthetically lethal with BRCA2 deficiency. *Proceedings of the National Academy of Sciences of the United States of America* 108, 686–691.
- Fricke, W.M., Brill, S.J., 2003. Slx1-Slx4 is a second structure-specific endonuclease functionally redundant with Sgs1-Top3. *Genes & Development* 17, 1768–1778.
- Friedberg, E.C., Walker, G.C., Siede, W., *et al.*, 2006. *DNA Repair and Mutagenesis*, second ed. Washington, D.C.: ASM Press.
- Goodman, M.F., Woodgate, R., 2013. Translesion DNA polymerases. *Cold Spring Harbor Perspectives in Biology* 5, a010363.
- Grimme, J.M., Honda, M., Wright, R., *et al.*, 2010. Human Rad52 binds and wraps single-stranded DNA and mediates annealing via two hRad52-ssDNA complexes. *Nucleic Acids Research* 38, 2917–2930.
- Gudmundsdottir, K., Lord, C.J., Witt, E., Tutt, A.N.J., Ashworth, A., 2004. DSS1 is required for RAD51 focus formation and genomic stability in mammalian cells. *EMBO Reports* 5, 989–993.
- Hanada, K., Budzowska, M., Davies, S.L., *et al.*, 2007. The structure-specific endonuclease Mus81 contributes to replication restart by generating double-strand DNA breaks. *Nature Structural & Molecular Biology* 14, 1096–1104.
- Ho, C.K., Mazón, G., Lam, A.F., Symington, L.S., 2010. Mus81 and Yen1 promote reciprocal exchange during mitotic recombination to maintain genome integrity in budding yeast. *Molecular Cell* 40, 988–1000.
- Holzen, T.M., Shah, P.P., Olivares, H.A., Bishop, D.K., 2006. Tid1/Rdh54 promotes dissociation of Dmc1 from nonrecombinogenic sites on meiotic chromatin. *Genes & Development* 20, 2593–2604.
- Hombauer, H., Srivatsan, A., Putnam, C.D., Kolodner, R.D., 2011. Mismatch repair, but not heteroduplex rejection, is temporally coupled to DNA replication. *Science* 334, 1713–1716.
- Howlett, N.G., Taniguchi, T., Olson, S., *et al.*, 2002. Biallelic inactivation of BRCA2 in Fanconi anemia. *Science* 297, 606–609.
- Jensen, R.B., Carreira, A., Kowalczykowski, S.C., 2010. Purified human BRCA2 stimulates RAD51-mediated recombination. *Nature* 467, 678–683.
- Kim, H., D'Andrea, A.D., 2012. Regulation of DNA cross-link repair by the Fanconi anemia/BRCA pathway. *Genes & Development* 26, 1393–1408.
- Kim, Y., Lach, F.P., Desetty, R., *et al.*, 2011. Mutations of the SLX4 gene in Fanconi anemia. *Nature Genetics* 43, 142–146.
- Knipscheer, P., Raschle, M., Smogorzewska, A., *et al.*, 2009. The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science* 326, 1698–1701.
- Krejci, L., van Komen, S., Li, Y., *et al.*, 2003. DNA helicase Srs2 disrupts the Rad51 presynaptic filament. *Nature* 423, 305–309.
- Langevin, F., Crossan, G.P., Rosado, I.V., Arends, M.J., Patel, K.J., 2011. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature* 475, 53–58.
- Li, X., Stith, C.M., Burgers, P.M., Heyer, W.-D., 2009. PCNA is required for initiating recombination-associated DNA synthesis by DNA polymerase δ . *Molecular Cell* 36, 704–713.
- Lieber, M.R., 2010. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annual Review of Biochemistry* 79, 181–211.
- Lim, D.S., Hasty, P., 1996. A mutation in mouse *rad51* results in an early embryonic lethal that is suppressed by a mutation in p53. *Molecular and Cellular Biology* 16, 7133–7143.
- Lisby, M., Barlow, J.H., Burgess, R.C., Rothstein, R., 2004. Choreography of the DNA damage response: Spatiotemporal relationships among checkpoint and repair proteins. *Cell* 118, 699–713.
- Liu, J., Doty, T., Gibson, B., Heyer, W.D., 2010. Human BRCA2 protein promotes RAD51 filament formation on RPA-covered single-stranded DNA. *Nature Structural & Molecular Biology* 17, 1260–1262.
- Liu, J., Ehmsen, K.T., Heyer, W.D., Morrical, S.W., 2011a. Presynaptic filament dynamics in homologous recombination and DNA repair. *Critical Reviews in Biochemistry and Molecular Biology* 46, 240–270.
- Liu, J., Renault, L., Veaute, X., *et al.*, 2011b. Rad51 paralogs Rad55-Rad57 balance the anti-recombinase function of Srs2 in Rad51 pre-synaptic filament formation. *Nature* 479, 245–248.
- Long, D.T., Raschle, M., Joukov, V., Walter, J.C., 2011. Mechanism of RAD51-dependent DNA interstrand cross-link repair. *Science* 333, 84–87.
- Lopes, M., Foiani, M., Sogo, J.M., 2006. Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Molecular Cell* 21, 15–27.
- Malkova, A., Haber, J.E., 2012. Mutations Arising During Repair of Chromosome Breaks. *Annual Review of Genetics* 46, 455–473.
- Malkova, A., Ivanov, E.L., Haber, J.E., 1996. Double-strand break repair in the absence of RAD51 in yeast: A possible role for break-induced DNA replication. *Proceedings of the National Academy of Sciences of the United States of America* 93, 7131–7136.
- Matos, J., Blanco, M.G., Maslen, S., Skehel, J.M., West, S.C., 2011. Regulatory control of the resolution of DNA recombination intermediates during meiosis and mitosis. *Cell* 147, 158–172.
- Mazin, A.V., Bornarth, C.J., Solinger, J.A., Heyer, W.-D., Kowalczykowski, S.C., 2000. Rad54 protein is targeted to pairing loci by the Rad51 nucleoprotein filament. *Molecular Cell* 6, 583–592.
- Mimitou, E.P., Symington, L.S., 2008. Sae2, Exo1 and Sgs1 collaborate in DNA double-strand break processing. *Nature* 455, 770–774.
- Modesti, M., Budzowska, M., Baldeyron, C., *et al.*, 2007. RAD51AP1 is a structure-specific DNA binding protein that stimulates joint molecule formation during RAD51-mediated homologous recombination. *Molecular Cell* 28, 468–481.
- Moldovan, G.L., Pfander, B., Jentsch, S., 2007. PCNA, the maestro of the replication fork. *Cell* 129, 665–679.
- Mortensen, U.H., Bendixen, C., Sunjevaric, I., Rothstein, R., 1996. DNA strand annealing is promoted by the yeast Rad52 protein. *Proceedings of the National Academy of Sciences of the United States of America* 93, 10729–10734.
- Moynahan, M.E., Jasin, M., 2010. Mitotic homologous recombination maintains genomic stability and suppresses tumorigenesis. *Nature Reviews Molecular Cell Biology* 11, 196–207.
- Moynahan, M.E., Pierce, A.J., Jasin, M., 2001. BRCA2 is required for homology-directed repair of chromosomal breaks. *Molecular Cell* 7, 263–272.
- Nimonkar, A.V., Genschel, J., Kinoshita, E., *et al.*, 2011. BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. *Genes & Development* 25, 350–362.
- Niu, H.Y., Chung, W.H., Zhu, Z., *et al.*, 2010. Mechanism of the ATP-dependent DNA end-resection machinery from *Saccharomyces cerevisiae*. *Nature* 467, 108–111.
- Pace, P., Mosedale, G., Hodskinson, M.R., *et al.*, 2010. Ku70 corrupts DNA repair in the absence of the Fanconi anemia pathway. *Science* 329, 219–223.
- Pellegrini, L., Yu, D.S., Lo, T., *et al.*, 2002. Insights into DNA recombination from the structure of a RAD51-BRCA2 complex. *Nature* 420, 287–293.
- Petukhova, G., Stratton, S., Sung, P., 1998. Catalysis of homologous DNA pairing by yeast Rad51 and Rad54 proteins. *Nature* 393, 91–94.
- Petukhova, G.V., Romanienko, P.J., Camerini-Otero, R.D., 2003. The Hop2 protein has a direct role in promoting interhomolog interactions during mouse meiosis. *Developmental Cell* 5, 927–936.
- Powell, S.N., Bindra, R.S., 2009. Targeting the DNA damage response for cancer therapy. *DNA Repair* 8, 1153–1165.
- Prakash, S., Prakash, L., 2002. Translesion DNA synthesis in eukaryotes: A one- or two-polymerase affair. *Genes & Development* 16, 1872–1883.

- Raschle, M., Knipsheer, P., Enoiu, M., *et al.*, 2008. Mechanism of replication-coupled DNA interstrand crosslink repair. *Cell* 134, 969–980.
- Ray Chaudhuri, A., Hashimoto, Y., Herrador, R., *et al.*, 2012. Topoisomerase I poisoning results in PARP-mediated replication fork reversal. *Nature Structural & Molecular Biology* 19, 417–423.
- Regairaz, M., Zhang, Y.V., Fu, H.Q., *et al.*, 2011. Mus81-mediated DNA cleavage resolves replication forks stalled by topoisomerase I-DNA complexes. *Journal of Cell Biology* 195, 739–749.
- Reid, S., Schindler, D., Hanenberg, H., *et al.*, 2007. Biallelic mutations in PALB2 cause Fanconi anemia subtype FA-N and predispose to childhood cancer. *Nature Genetics* 39, 162–164.
- Reuter, M., Zelensky, A., Smal, I., *et al.*, 2014. BRCA2 diffuses as oligomeric clusters with RAD51 and changes mobility after DNA damage in live cells. *Journal of Cell Biology* 207, 599–613.
- Rijkers, T., van de Nieuweland, J., Morolli, B., *et al.*, 1998. Targeted inactivation of mouse *RAD52* reduces homologous recombination but not resistance to ionizing radiation. *Molecular and Cellular Biology* 18, 6423–6429.
- Sasanuma, H., Tawaramoto, M.S., Lao, J.P., *et al.*, 2013. A new protein complex promoting the assembly of Rad51 filaments. *Nature Communications* 4, 1676.
- Schlacher, K., Christ, N., Siaud, N., *et al.*, 2011. Double-strand break repair-independent role for BRCA2 in blocking stalled replication fork degradation by MRE11. *Cell* 145, 529–542.
- Sneeden, J.L., Grossi, S.M., Tappin, I., Hurwitz, J., Heyer, W.D., 2013. Reconstitution of recombination-associated DNA synthesis with human proteins. *Nucleic Acids Research* 41, 4913–4925.
- Solinger, J.A., Kiianitsa, K., Heyer, W.-D., 2002. Rad54, a Swi2/Snf2-like recombinational repair protein, disassembles Rad51:dsDNA filaments. *Molecular Cell* 10, 1175–1188.
- Stracker, T.H., Petrini, J.H.J., 2011. The MRE11 complex: Starting from the ends. *Nature Reviews Molecular Cell Biology* 12, 90–103.
- Sugiyama, T., New, J.H., Kowalczykowski, S.C., 1998. DNA annealing by Rad52 protein is stimulated by specific interaction with the complex of replication protein A and single-stranded DNA. *Proceedings of the National Academy of Sciences of the United States of America* 95, 6049–6054.
- Sung, P., 1994. Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein. *Science* 265, 1241–1243.
- Sung, P., 1997a. Function of yeast Rad52 protein as a mediator between replication protein A and the Rad51 recombinase. *Journal of Biological Chemistry* 272, 28194–28197.
- Sung, P., 1997b. Yeast Rad55 and Rad57 proteins form a heterodimer that functions with replication protein A to promote DNA strand exchange by Rad51 recombinase. *Genes & Development* 11, 1111–1121.
- Symington, L.S., 2014. End resection at double-strand breaks: Mechanism and regulation. *Cold Spring Harbor Perspectives in Biology* 6, a016436.
- Symington, L.S., Gautier, J., 2011. Double-strand break end resection and repair pathway choice. *Annual Review of Genetics* 45, 247–271.
- Tanaka, K., Hiramoto, T., Fukuda, T., Miyagawa, K., 2000. A novel human rad54 homologue, Rad54B, associates with Rad51. *Journal of Biological Chemistry* 275, 26316–26321.
- Veaute, X., Jeusset, J., Soustelle, C., *et al.*, 2003. The Srs2 helicase prevents recombination by disrupting Rad51 nucleoprotein filaments. *Nature* 423, 309–312.
- Wang, A.T., Smogorzewska, A., 2015. Snapshot: Fanconi anemia and associated proteins. *Cell* 160, 354.e1.
- Wang, X.A., Ira, G., Tercero, J.A., *et al.*, 2004. Role of DNA replication proteins in double-strand break-induced recombination in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 24, 6891–6899.
- Wiese, C., Dray, E., Groesser, T., *et al.*, 2007. Promotion of homologous recombination and genomic stability by RAD51AP1 via RAD51 recombinase enhancement. *Molecular Cell* 28, 482–490.
- Wright, W.D., Heyer, W.D., 2014. Rad54 functions as a heteroduplex DNA pump modulated by its DNA substrates and Rad51 during d loop formation. *Molecular Cell* 53, 420–432.
- Wu, L.J., Hickson, I.D., 2003. The Bloom's syndrome helicase suppresses crossing-over during homologous recombination. *Nature* 426, 870–874.
- Xia, B., Sheng, Q., Nakanishi, K., *et al.*, 2006. Control of BRCA2 cellular and clinical functions by a nuclear partner, PALB2. *Molecular Cell* 22, 719–729.
- Zelensky, A., Kanaar, R., Wyman, C., 2014. Mediators of homologous DNA pairing. *Cold Spring Harbor Perspectives in Biology* 6, a016451.
- Zeman, M.K., Cimprich, K.A., 2014. Causes and consequences of replication stress. *Nature Cell Biology* 16, 2–9.
- Zhang, J.Q., Walter, J.C., 2014. Mechanism and regulation of incisions during DNA interstrand cross-link repair. *DNA Repair* 19, 135–142.
- Zhu, Z., Chung, W.H., Shim, E.Y., Lee, S.E., Ira, G., 2008. Sgs1 helicase and two nucleases Dna2 and Exo1 resect DNA double-strand break ends. *Cell* 134, 981–994.
- Zierhut, C., Berlinger, M., Rupp, C., Shinohara, A., Klein, F., 2004. Mnd1 is required for meiotic interhomolog repair. *Current Biology* 14, 752–762.

Further Reading

- Aguilera, A., Garcia-Muse, T., 2013. Causes of genome instability. *Annual Review of Genetics* 47, 1–32.
- Ceballos, S.J., Heyer, W.D., 2011. Functions of the Snf2/Swi2 family Rad54 motor protein in homologous recombination. *Biochimica et Biophysica Acta* 1809, 509–523.
- Heyer, W.D., Ehmsen, K.T., Liu, J., 2010. Regulation of homologous recombination in eukaryotes. *Annual Review of Genetics* 44, 113–139.
- Hunter, N., 2007. Meiotic recombination. In: Aguilera, A., Rothstein, R. (Eds.), *Homologous Recombination*. Berlin-Heidelberg: Springer-Verlag.
- Mehta, A., Haber, J.E., 2014. Sources of DNA double-strand breaks and models of recombinational DNA repair. *Cold Spring Harbor Perspectives in Biology* 6, a016428.
- San Filippo, J., Sung, P., Klein, H., 2008. Mechanism of eukaryotic homologous recombination. *Annual Review of Biochemistry* 77, 229–257.
- Schwartz, E.K., Heyer, W.D., 2011. Processing of joint molecule intermediates by structure-selective endonucleases during homologous recombination in eukaryotes. *Chromosoma* 120, 109–127.
- Symington, L.S., Rothstein, R., Lisby, M., 2014. Mechanisms and regulation of mitotic recombination in *Saccharomyces cerevisiae*. *Genetics* 198, 795–835.

NUCLEIC ACID SYNTHESIS/BREAKDOWN: TRANSCRIPTION

Contents

Prokaryotic Transcription

Eukaryotic Transcriptional Regulation

Distant Activation of Transcription by Enhancers

The Spliceosome and Pre-mRNA Splicing

Pre-mRNA Splicing: Function and Dysfunction

Prokaryotic Transcription

DM Hinton, National Institutes of Health, Bethesda, MD, USA

Published by Elsevier Inc.

Glossary

Nontemplate strand DNA strand that is same sequence as transcribed RNA, the 'top' strand.

Operator A DNA sequence bound by a repressor or activator.

Promoter DNA sequence recognized by RNA polymerase that sets the start site for transcription.

Ribonucleoside triphosphates (rNTPs) rNTPs used by RNA polymerase to synthesize RNA.

Template strand DNA strand that is complementary to transcribed RNA, the 'bottom' strand.

Introduction

Transcription, the synthesis of a 5'→3' RNA chain complementary to a DNA template, is the first step in the expression of a gene. Consequently, the control of transcription is fundamentally important to all organisms. In all three kingdoms of life, RNA synthesis is catalyzed by multi-subunit DNA-dependent RNA polymerases, which generate RNA through the hydrolysis of pyrophosphate from ribonucleoside triphosphates (rNTPs). In this process, RNA polymerase (RNAP) first binds to a specific DNA sequence (a promoter) that sets the start site for transcription. RNAP then starts RNA synthesis (initiation), proceeds to transcribe through the gene(s) (elongation), and finally stops transcription at a specific site (termination). Additional factors can control the process by interacting with the polymerase, the DNA, or the RNA, which in turn affects the various steps. Thus, a myriad of outcomes arising from these various interactions results in the sophisticated regulation needed for the proper growth and development of a cell and for the cell to adapt to various environmental stresses.

Despite their seeming simplicity, prokaryotes (bacteria and archaea) have elegant mechanisms to perform and regulate the process of transcription. Because prokaryotes not only survive but flourish in a wide range of environments, one might surmise that the process of transcription varies widely among prokaryotes and between the prokaryotic and eukaryotic

worlds. However, fundamental mechanisms underlie how transcription is done throughout biology. In this article we will review this process in prokaryotes and also indicate some of the basic similarities among all organisms (reviewed in [Decker and Hinton, 2013](#)).

Core RNA Polymerase

RNA polymerase 'core' is the RNA-synthesizing machine. In bacteria it is composed of 5 subunits: β , β' (the two largest proteins), two α 's, and ω ([Figure 1\(a\)](#), left; [Zhang *et al.*, 1999](#); reviewed in [Murakami and Darst, 2003](#)). Although archaeal core ([Figure 1\(a\)](#), right) and eukaryotic cores (for Pol I, Pol II, and Pol III) are composed of more subunits, they too contain proteins that are homologous to these bacterial proteins ([Hirata *et al.*, 2008](#); reviewed in [Vannini and Cramer, 2012](#); [Ebright, 2000](#); [Jun *et al.*, 2011](#)). All cores are arranged like a 'claw' with the active site and a catalytic Mg^{++} located at the center. The largest subunits, such as β and β' in bacteria, generate the top and bottom 'pincers' of the claw with a long alpha helix within β' (or its equivalent in archaeal or eukaryotic core) extending from the top to the bottom. This is the β' bridge helix (shown as red helix in [Figure 1\(a\)](#)).

Bacterial core is assembled as: $\alpha_2 + \beta + \beta' / \omega$. Each α subunit is composed of an N-terminal domain (NTD) tethered to its C-

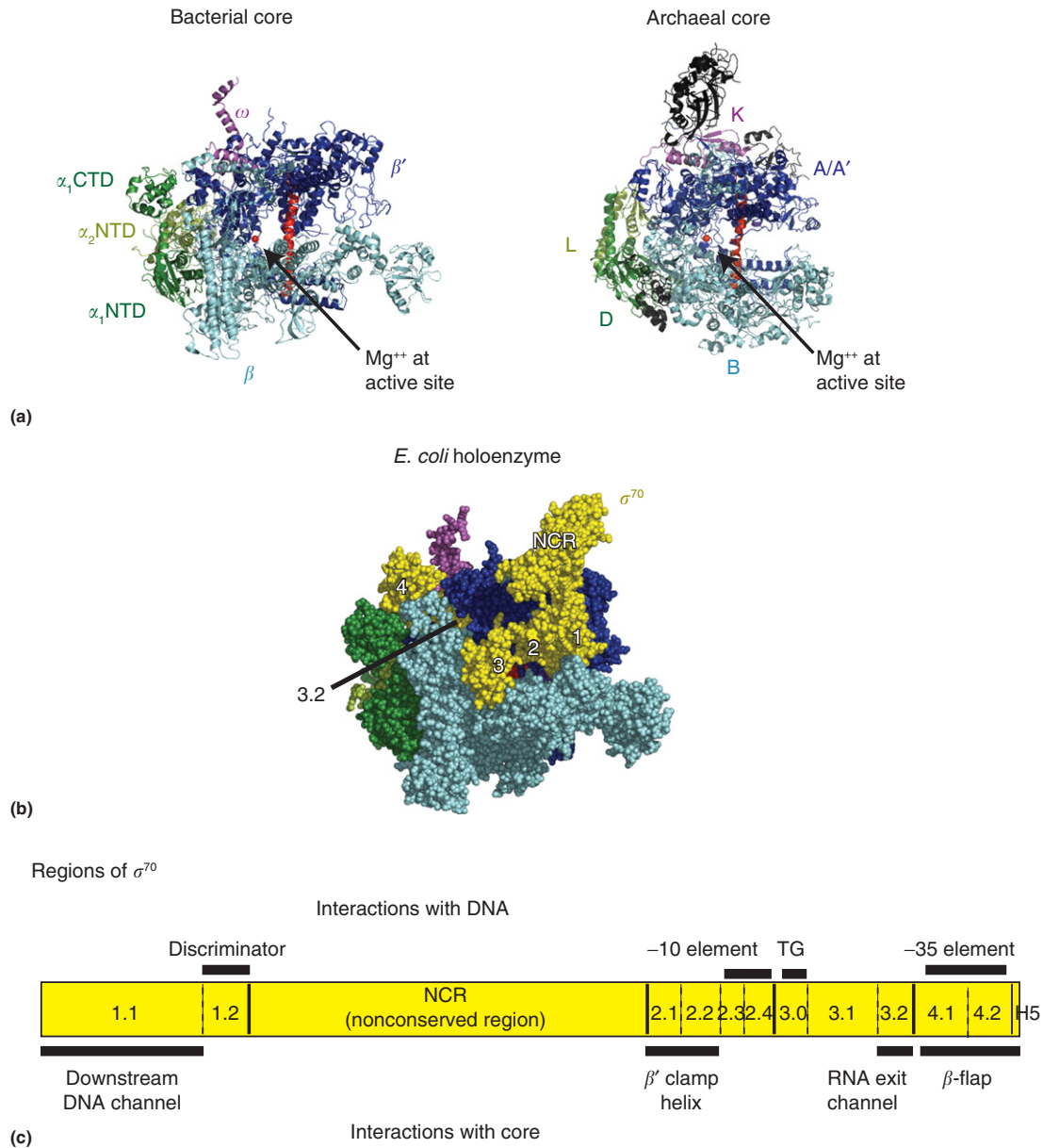


Figure 1 Structures of bacterial and archaeal RNA polymerases and Regions of σ^{70} . (a) RNA polymerase core of *E. coli* (Murakami, 2013; PDB: 4IGC without the σ subunit; left) and *Sulfolobus solfataricus* (Hirata *et al.*, 2008; PDB: 2PMZ; right). Homologous subunits in the bacterial and archaeal cores are colored similarly as follows: β' , A/A', dark blue (except for the bridge helix); β , B, light blue; α_1 , D, dark green; α_2 , L, light green; ω , K, magenta. The β' bridge helix is shown in red; the Mg^{++} at the active site is the red sphere. Only one of the α CTDs (α_1 CTD) can be seen in the *E. coli* structure. Archaeal subunits E', F, H, N, and P that are missing in bacteria are shown in black. (b) *E. coli* RNA polymerase holoenzyme (Murakami, 2013; PDB: 4IGC). The σ^{70} subunit is shown in yellow and the positions of the 4 major Regions (1, 2, 3, and 4), the nonconserved region (NCR), and subregion 3.2 are indicated. (c) Regions and subregions of σ^{70} . Important interactions between σ^{70} and promoter DNA or core are shown on the top or bottom, respectively.

terminal domain (CTD) by a flexible peptide linker. (Only the CTD of α_1 is observed in the structure in Figure 1(a).) Dimerization of the two α 's occurs through their NTDs. As each NTD also interacts with either β or β' , the two α subunits are distinguished within core by their interaction with the particular large subunit. ω , the smallest subunit, serves as a chaperone for β' (Mathew and Chatterji, 2006; and references therein) and when present in core, generates part of the binding site for the small molecule effector ppGpp, which is

involved in the ability of some bacteria, such as *Escherichia coli*, to respond to limiting nutrient conditions (Mechold *et al.*, 2013; Zuo *et al.*, 2013).

Bacterial RNA Polymerase Holoenzyme

Besides the RNA-synthesizing core, RNA polymerases also require a specificity factor to recognize DNA promoter sequences

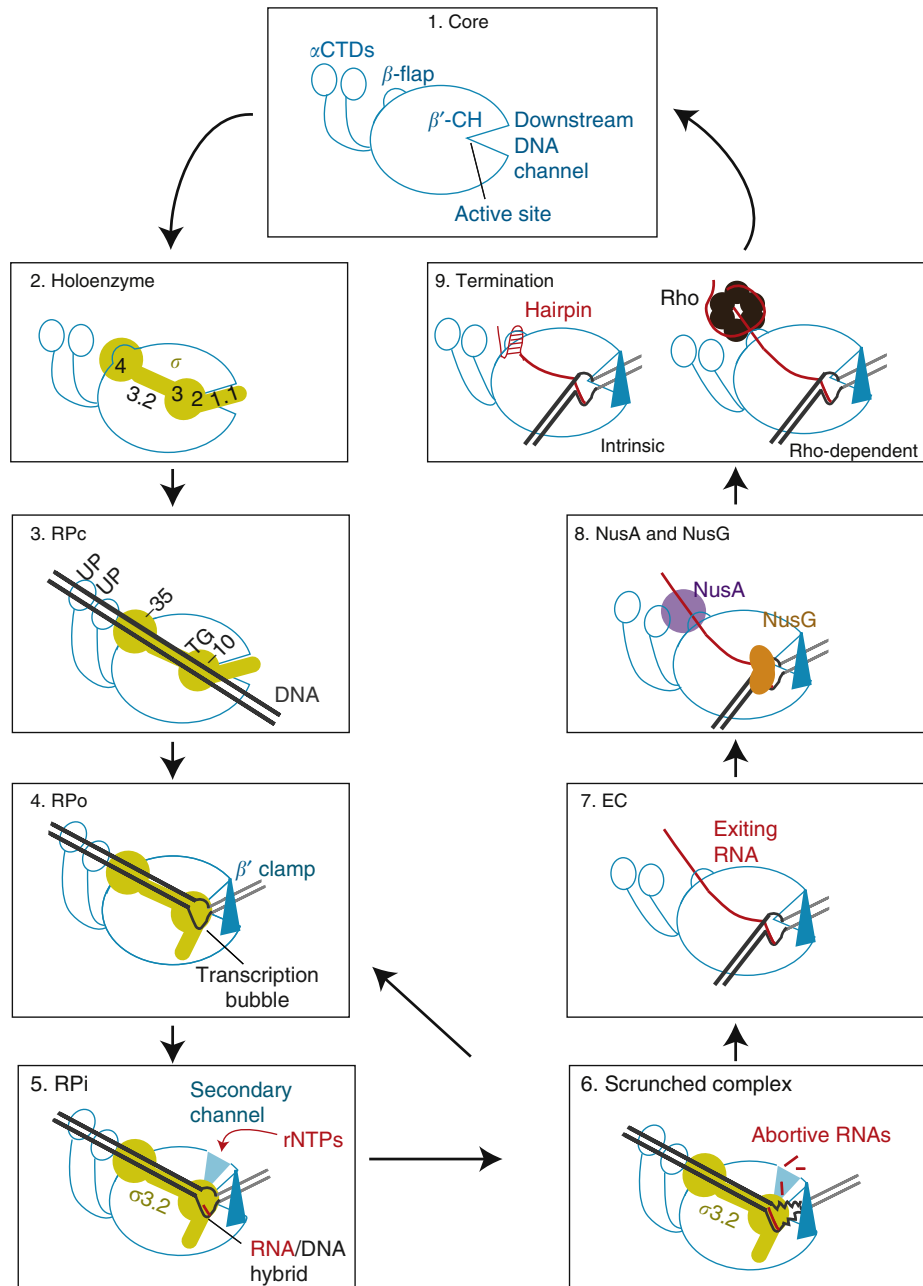


Figure 2 The process of transcription in bacteria. (1) Core polymerase. Schematic of core based on structural information (Figure 1(a)). Positions of the α CTDs, β -flap, β' -CH (clamp helix), downstream DNA channel, and active site are shown. (2) Holoenzyme. σ is shown in yellow; core is shown in cyan. Positions of σ Regions 1.1, 2, 3, 3.2, and 4 are indicated. (3) Closed complex (RPc). Promoter DNA lies across the face of RNAP. Positions of promoter elements (UP, -35, -15 TG $^{-14}$, and -10) are indicated. (4) Open complex (RPo). Isomerization of RNAP and bending, unwinding of the DNA generates the transcription bubble surrounding the start site and the β' clamp, which secures the DNA within the downstream DNA channel. (5) Initiating complex (RPi). The presence of rNTPs, which enter through the secondary channel, starts RNA synthesis. The initially transcribed RNA remains annealed to the DNA template, creating a RNA/DNA hybrid of 8–9 bp. (6) Scrunched complex. As RNA is made, RNAP moves forward pulling the DNA into the active site but without releasing the promoter. The release of small abortive transcripts collapses the scrunch returning the complex back to RPo. (7) Elongating complex (EC). RNA exits through the RNA exit channel, which displaces σ Regions 3.2 and 4, eventually resulting in the loss of σ . (8) Elongation factors NusA and NusG. NusA interacts with the β -flap, α CTD, and RNA. NusG interacts with the β' -CH, extending over the DNA upstream of the transcription bubble to hold the DNA even more securely. (9) Termination. The RNAP/DNA/RNA complex is disrupted by the presence of a terminating hairpin (intrinsic) or a hexamer of Rho protein (Rho-dependent).

that set the start site for transcription. In bacteria, these factors are the σ subunits (reviewed in Gruber and Gross, 2003). The combination of core with σ constitutes the holoenzyme. All

bacteria have a primary ('housekeeping') σ , such as σ^{70} of *E. coli*, that is needed to transcribe genes during exponential growth (Figure 1(b) shows the structure of *E. coli* σ^{70}).

containing holoenzyme with σ in yellow (Murakami, 2013).). In addition, most bacteria have a few to dozens of σ factors that are used during other growth conditions or times of stress (reviewed in Osterberg *et al.*, 2011). Primary σ 's contain 4 broad regions of sequence similarities (Regions 1, 2, 3, and 4) with subregions (1.1, 1.2, 2.1, 2.2, etc.) identified by sequence, structure, and function (Figures 1(b) and 1(c)). Primary σ factors also can contain a nonconserved region (NCR) between Regions 1 and 2. Alternate σ 's fall into two categories: those related to σ^{70} and those of the disparate σ^{54} family. Within the σ^{70} family, one category of alternate σ 's is quite similar to primary σ factors, retaining similarities within Regions 2, 3, and 4. The other class, extracytoplasmic (ECF) σ 's, are less similar, only having recognizable Regions 2 and 4. ECF σ 's are involved in processes concerning cell surface or transport as well as other cellular needs. The σ^{54} family represents a separate class of factors that are not related to the σ^{70} family members.

Extensive interactions between core and σ secure σ within the holoenzyme (Figure 1(b); Vassilyev *et al.*, 2002; Murakami *et al.*, 2002b; Murakami, 2013; reviewed in Decker and Hinton, 2013). Within the σ^{70} family there are three primary contacts: σ Regions 2.1 and 2.2 with the β' clamp helix (CH), a coiled-coil motif; the C-terminal portion of σ (Regions 4.1/4.2/H5) with a portion of β called the β -flap; and σ Region 3.2 with the RNA exit channel, a portion of β and β' that will provide the path for the emerging RNA during transcription (Figures 1(b) and 1(c); shown schematically in Figure 2, steps 2 and 6). Although the ECF σ 's lack a homologous Region 3, presumably they too have residues that make interactions that are similar to those between Region 3.2 and core. In addition, crosslinking analyses suggest that like σ^{70} , the disparate σ^{54} family members also have interactions with the β' CH and the β -flap (reviewed in Wigneshweraraj *et al.*, 2008). Finally, primary σ factors have an additional interaction made by Region 1.1 and the β , β' – cleft the portion of core that will hold the DNA downstream of the transcription start site once the promoter is stably bound (Decker and Hinton, 2013; Murakami, 2013; Bae *et al.*, 2013). Thus, Region 1.1, with its abundance of negatively charged residues, is thought to provide a 'placeholder' in the absence of the DNA.

The Process of Transcription in Bacteria

Initiation

Transcription is a highly ordered process. A detailed understanding of bacterial transcription at a molecular level has been obtained from decades of extensive genetic and biochemical analyses together with more recent structures of bacterial and eukaryotic polymerases with and without DNA and RNA (reviewed in Saecker *et al.*, 2011; Hook-Barnard and Hinton, 2007; Decker and Hinton, 2013). This article focuses on *E. coli* σ^{70} -containing RNAP since most biochemical analyses have used this polymerase as the model. However, it should be noted that there are differences depending on the particular σ factor and the particular bacterium.

Transcription initiation begins with the recognition of double-stranded promoter elements: σ Region 4 and the – 35

element ($^{-35}\text{TTGACA}^{-30}$), σ Region 3.0 and the extended – 10 ($^{-15}\text{TG}^{-14}$), σ Region 2.4 and the first bp of the – 10 element (^{-12}T), and the α CTDs and A/T-rich sequences called UP elements (– 40 to – 60) (Figure 1(c) and Figure 2, step 3). Because σ by itself binds extremely poorly to DNA, this recognition and binding occurs only within the context of the holoenzyme. The positions indicated for the elements represent the ideal, but there is limited flexibility. For example, the spacing between the 3' end of the – 35 element (typically position – 30) and the start of the – 10 element (typically position – 12) can vary from 16 to 18 bp. In addition, at the very strong ribosomal promoters (detailed below), the – 10 element actually occurs from – 9 to – 14.

As promoter elements are modular, the UP elements may or may not be present, and it is not necessary for all of the σ elements to be recognized. However, normally at least two good σ elements must be present for recognition by RNAP in the absence of other factors. Thus, promoters can be classified as one of the three types: – 35/– 10 (the most prevalent in *E. coli*), $^{-15}\text{TG}^{-14}/-10$ (referred to as an extended – 10 promoter), and – 35/ $^{-15}\text{TG}^{-14}$. The major groove of the DNA contains the specific determinants for the σ interactions with the double-stranded DNA. For example, a classic helix-turn-helix within σ Region 4 interacts with specific base determinants on the template strand at positions – 33 and – 31 (Campbell *et al.*, 2002 and references therein). In contrast, UP elements are recognized through contacts with the minor groove of the A/T-rich sequence (Ross *et al.*, 1993; Naryshkin *et al.*, 2000). The complex that is formed by this initial binding of RNAP with the promoter is called the closed complex (RPC), since the DNA remains double-stranded or 'closed'. RPC's are typically unstable; they are usually observed only at low temperatures (4 °C) and can be challenged by the addition of a substrate that competes with the DNA, such as the polyanion heparin. Biochemical analyses suggest that in RPC, the DNA simply lies across the face of the polymerase.

Incubation of the RNA polymerase with the promoter DNA at higher temperatures (typically above 20 °C) results in the formation of the open complex (RPO), in which the DNA is unwound, generating a bubble surrounding the start site for transcription (– 12/– 11 to ~ + 5) (Figure 2, step 4). Formation of RPO involves several intermediate steps and requires major conformational changes (isomerization) within polymerase as well as the unwinding and bending of the DNA as it enters the active site of polymerase (Murakami *et al.*, 2002a; Schickor *et al.*, 1990). These changes include the movement of σ Region 1.1 out of the downstream DNA channel so that the DNA can enter (Mekler *et al.*, 2002) and the rearrangement of a portion of β' , called the clamp, into a structure that secures the downstream DNA (Saecker *et al.*, 2011; Chakraborty *et al.*, 2012). In addition, the DNA itself must bend by approximately 90° and the template strand surrounding the transcription start site must descend into the active site.

The generation of single-stranded DNA within the transcription bubble provides additional specific sites for interaction with polymerase (Lee *et al.*, 2004; Feklistov and Darst, 2011; Zhang *et al.*, 2012; Schneider, 2001; Schroeder *et al.*, 2009). σ Region 2.3 interacts with the nontemplate single-stranded DNA from – 11 to – 7 ($^{-11}\text{ATAAT}^{-7}$) with specific recognition of base determinants at the – 11A and the – 7T.

At these positions, the nontemplate base itself is 'flipped out', stacking against the aromatic ring of a residue or held in a protein pocket. Through these interactions as well as other nonspecific interactions with the -10 element, polymerase captures the nontemplate strand so that the bubble is kept open.

Other residues within σ and β also contribute important interactions with the single-stranded DNA (Zhang *et al.*, 2012; Severinov *et al.*, 1994). A structure within σ Region 3.2, the σ 'finger,' interacts with template positions -3 and -4 to place the $+1$ start site in the correct position for the incoming ribonucleoside triphosphates (rNTPs). When a GC-rich sequence (discriminator) is present in the promoter at positions -4 to -6 , residues within σ Region 1.2 can bury a nontemplate ^{-6}G within a protein pocket (Figure 1(c)) (Haugen *et al.*, 2008; Feklistov *et al.*, 2006; Travers, 1984; Zhang *et al.*, 2012). Finally, β subunit residues can interact with Core Recognition Element (CRE) present at positions -4 to $+2$ (Zhang *et al.*, 2012).

Once RPo has formed, addition of rNTPs results in the initiating complex (RPI) (Figure 2, step 5). The rNTPs enter through the secondary channel, a small channel within RNAP that leads to the active site. The RNA remains annealed to the DNA template strand for 8–9 bp. Because the polymerase is still attached to the promoter, continued synthesis 'scrunches' the DNA into the active site as the front end of the polymerase moves forward but the back end RNAP remains attached to the promoter (Figure 2, step 6; Kapanidis *et al.*, 2006). RNA longer than 9 nucleotides in length enters the RNA exit channel. However, at the start of initiation this channel is still occupied by σ Region 3.2. Consequently, promoter clearance typically requires multiple attempts, in which short RNAs are released as abortive products, causing synthesis to restart at position $+1$ (reviewed in Hsu, 2002). Finally, the emerging RNA successfully enters the channel, pushing σ Region 3.2 and helping to release σ Region 4/H5 from the grip of the β -flap. Complete release of σ , which involves disrupting the strong interaction of σ Region 2.2 with the β' CH, often accompanies formation of the elongating complex (EC) containing DNA/RNA/core polymerase (Figure 2, step 7). However, as detailed below, under certain conditions σ can remain associated with core, interacting with sequences within the DNA.

RPo is typically a very stable complex. However, for certain promoters in enteric bacteria, including the strong ribosomal promoters that account for about 70% of transcription during fast exponential growth, RPo is unstable (reviewed in Paul *et al.*, 2004; Srivatsan and Wang, 2008). Under good growth conditions, this is not a problem. The presence of a high concentration of rNTPs results in the rapid transition of the kinetically unstable RPo to a stable, initiating RPI that proceeds to elongation. When cells are slow growing or starved, the concentration of rNTPs remains high, but there is also an increase in the intracellular concentration of the small molecule ppGpp. The presence of ppGpp together with the DksA protein, which binds within the secondary channel, destabilizes the RNAP ribosomal promoter complex even further, essentially eliminating initiation (Rutherford *et al.*, 2009). In contrast, promoters within the amino acid biosynthetic pathways are activated by DksA (Gummesson *et al.*, 2013 and references therein). In this way, starved cells shift their focus

from ribosomal synthesis to pathways for amino acid synthesis and stress adaptation. Although gram positive bacteria, such as *Bacillus subtilis*, lack DksA and generate stable RPo complexes at ribosomal promoters, they still regulate expression of ribosomal RNA using ppGpp. In this case, the presence of ppGpp limits formation of GTP, which in turn restricts initiation of rRNA chains that start with GTP.

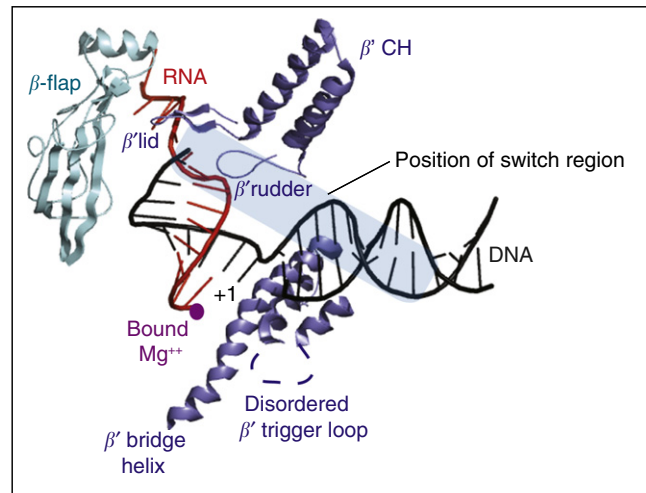
Elongation

The elongating complex (EC) is a processive transcription machine that synthesizes RNA at a rate of approximately 70 nt/s. Structures indicate that the main DNA binding channel of EC is smaller than that observed with RNAP alone. Furthermore, the 90° bend in the downstream DNA is stabilized by specific interactions with β and β' residues. These properties contribute to stability and processivity (Vassilyev *et al.*, 2007).

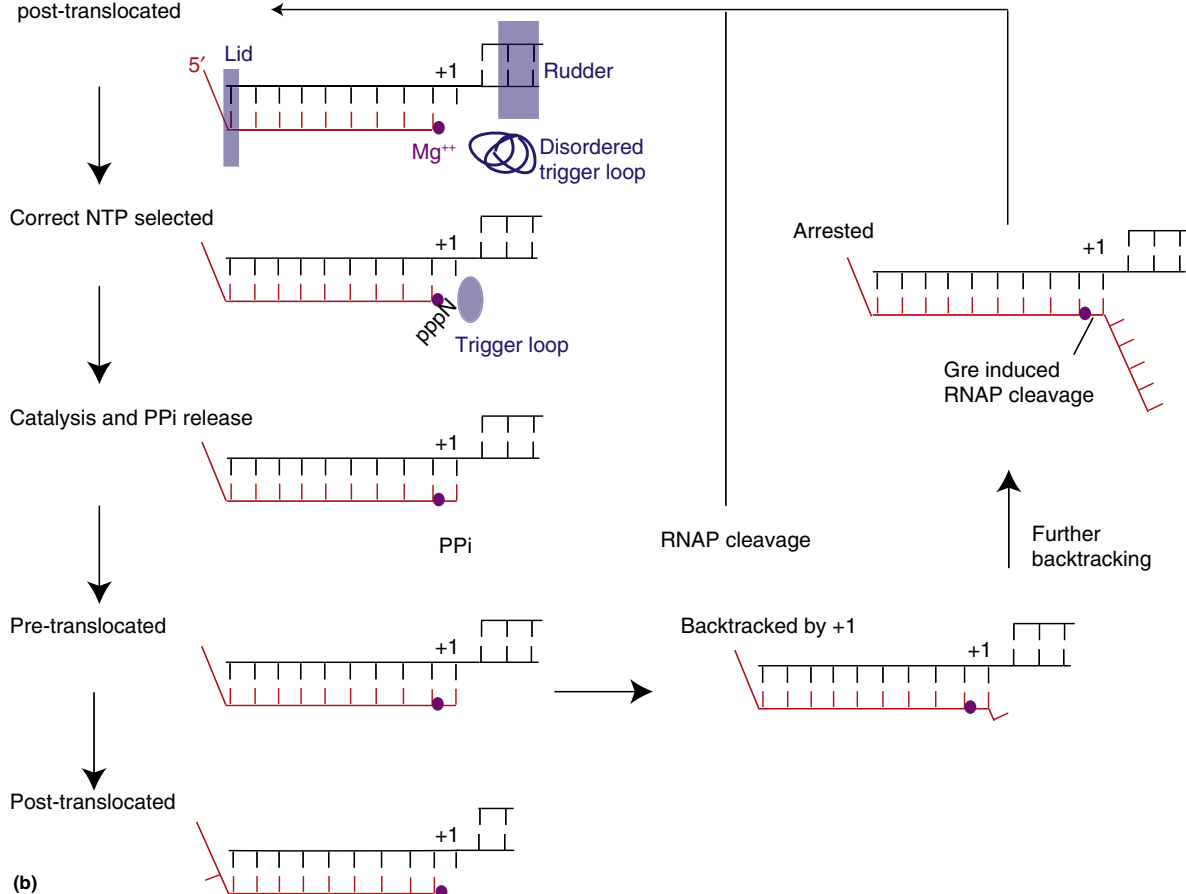
The heart of the EC is the active site channel where 8–9 bases of the newly synthesized RNA (red in Figure 3) are annealed to the DNA template strand at positions -1 to -9 , and the DNA template strand is located such that the $+1$ site is next to a catalytic Mg^{++} bound by core (Vassilyev *et al.*, 2007; Wang *et al.*, 2006; reviewed in Martinez-Rucobo and Cramer, 2013). For transcription to occur smoothly, the complementary NTP must be positioned correctly for catalysis, the generated pyrophosphate (PPi) must be released, and the DNA and RNA must translocate 1 base after catalysis. Several structures within the major subunits of RNAP (β and β') are thought to drive this process in an orderly manner. The β' lid loop and β' rudder lie on either side of the β' CH (Figures 3(a) and 3(b)). The β' lid loop sits at the upstream edge of the RNA: DNA hybrid, preventing an extension of more than 9 bp in the hybrid, while the β' rudder sits toward the downstream edge of the hybrid, helping to stabilize the transcription bubble. A 'switch region', composed of β and β' residues, sits at the base of the β' clamp and functions like a hinge, opening and closing the RNAP claw (reviewed in Srivastava *et al.*, 2011). This regulates how tightly the RNAP grips the DNA. One of these switches, β switch loop 3, sits just where the 5' end of the RNA extends toward the RNA exit channel and is thought to help guide the RNA toward the channel. Finally, the β' trigger loop and the β' bridge helix, which are located at the active site itself, are intimately involved in catalysis and translocation.

Structures of the EC with different substrates and inhibitors have revealed various conformations of the β' trigger loop and the β' bridge helix, leading to the following mechanism for ordered catalysis (Figure 3(b); reviewed in Kaplan, 2010; Martinez-Rucobo and Cramer, 2013). At the beginning of the process (the post-translocated state) the trigger loop is in a disordered state. If the base of the incoming rNTP (pppN) is complementary to the template base, the refolding of the trigger loop into an alpha helical structure will secure the rNTP in the correct position for insertion. Thus, the trigger loop helps to ensure accurate synthesis of the RNA. Catalysis is thought to occur through a nucleophilic substitution $\text{S}_{\text{N}}2$ mechanism, in which the 3'OH at the end of the RNA attacks the α -phosphate of the substrate rNTP. Two Mg^{++} s promote the nucleophilic attack needed to generate the phosphodiester bond. One Mg^{++} is held by core at the 3' end of the RNA; the

Active site of EC



(a)

Elongation process
post-translocated

(b)

Figure 3 Structural arrangements in EC and the process of elongation. (a) Portion of the EC of *Thermus thermophilus* (Vassilyev *et al.*, 2007; PDB: 2051), showing the DNA (black), the RNA (red), the β -flap (cyan), the β' CH, the β' rudder, and the β' bridge helix. Positions of the bound Mg^{++} (magenta circle), the disordered β' trigger loop, and the switch region, composed of both β and β' residues, are drawn. (b) Cartoon depicting elongation. Positions of the β' lid, β' rudder, β' trigger loop (all shown in purple), the Mg^{++} (shown as magenta circle), and the 5' end of the RNA (red) are shown. PPi, pyrophosphate; pppN, ribonucleoside triphosphate.

other is transiently bound with the rNTP. After catalysis, the DNA is left in the pre-translocated state and must be translocated in order to repeat catalysis. Because different structures have revealed both straight and bent bridge helix conformations, it has been suggested that movement of the bridge helix toward the hybrid may facilitate translocation of the DNA. In addition, it is thought that PPi release may cause the trigger loop to become disordered, which returns the machine back to a state that can incorporate the next NTP.

Pausing

Although the EC is a highly processive machine, RNAP does pause for a number of reasons, such as the incorporation of a nontemplated nucleotide, the ability of RNA to adopt a secondary structure within the RNA exit channel, the presence of a lesion within the DNA, or the influence of a particular sequence. Initially, the EC enters what is called an elemental pause (reviewed in [Martinez-Rucobo and Cramer, 2013](#)). This paused state has a number of properties that disfavor synthesis: the bridge helix at the active site is in a kinked rather than straight position, the RNA exit channel is widened, the position of the switch region favors a looser binding of the RNA/DNA hybrid, and the β' clamp is in a more open state, meaning that the DNA is not as tightly held ([Weixlbaumer et al., 2013](#)). Consequently, the 3' end of the RNA misaligns relative to the template. This is referred to as 'fraying'. Elemental pauses are important because they are on the pathway for longer pauses that can lead to termination.

In the case of fraying because of a misincorporated base, translocation of the RNAP (backtracking) by 1 bp allows cleavage of the misincorporated nucleotide by the intrinsic endonuclease activity present within RNAP and the resumption of transcription ([Figure 3\(b\)](#)). However, RNAP can also backtrack by several nucleotides, causing the nucleotides at the 3' end of the RNA to become completely out of register with the template and extruded through the secondary channel ([Figure 3\(b\)](#)). In this scenario, the EC is arrested since there is no way to continue RNA synthesis. This problem can be solved by proteins called Gre factors (GreA and GreB), which insert into the secondary channel, reach the active site, and stimulate the RNAP endonuclease so that the extruded RNA is released and transcription can resume again at the +1 site ([Orlova et al., 1995](#); reviewed in [Borukhov et al., 2005](#)).

Pausing and termination are also modulated by NusA, a protein that associates with EC ([Yang et al., 2009](#); [Yang and Lewis, 2010](#)). NusA contains domains that bind to the β -flap region of RNAP and interact with RNA emerging from the exit channel ([Figure 2](#), step 8). Through these interactions, NusA promotes pausing and termination although it can also be hijacked by other proteins that enhance processivity and anti-termination. In addition, a domain of *E. coli* NusA also interacts with the α CTDs.

Another factor that associates with the EC is NusG, which binds to the β' CH, the same interaction site used by σ Region 2.1/2.2 during transcription initiation ([Figure 2](#), step 8). In *E. coli*, this factor reduces pausing and increases the rate of RNA synthesis ([Herbert et al., 2010](#)). Structures obtained using Spt5, the archaeal orthologue of NusG, together with archaeal

polymerase or a portion of the polymerase, indicate that NusG binding extends over the active site cleft and suggest a direct interaction between NusG and the nontemplate strand in the transcription bubble ([Klein et al., 2011](#); [Martinez-Rucobo et al., 2011](#)). The additional clamping provided by NusG increases the elongation rate of RNAP and discourages pausing, thereby increasing processivity. In addition, NusG interacts with the ribosomal protein S10. This is thought to connect the transcriptional and translational machineries, coordinating the rates of transcription and translation. Interestingly, NusG also interacts with the termination protein Rho (detailed below), and in this case stimulates termination ([Burmann et al., 2010](#) and references therein).

A special category of pauses is σ -dependent (reviewed in [Perdue and Roberts, 2011](#)). These pauses occur close to the promoter before σ has been released from the initiating RNAP and involve the presence of two specific sequences. First, there is a sequence that resembles one of the σ binding elements, such as AnnnT (similar to the -10 element TATAAT) together with a discriminator-like sequence or TGGATT (similar to the -35 element TTGACA). Secondly, there is a sequence that can induce RNAP backtracking (G + C rich sequence followed by an A + T-rich sequence) just upstream of the pause. It is thought that the presence of the σ binding element creates a scrunched complex much like when RNAP is trying to clear a promoter ([Figure 2](#), step 6); that is, as RNAP synthesizes RNA, the DNA is scrunched since RNAP is held by the σ /DNA interactions. (Whether σ binds the element first or RNAP pauses first is unclear.) At this point, synthesis must stop. RNAP can backtrack, relaxing the scrunch but leaving the RNA out of register. Gre-induced RNA cleavage is then needed to rescue the complex. Even though promoter-proximal pausing is important for eukaryotic gene expression and approximately 15% of *E. coli* promoters appear to contain these pauses, the purpose of the bacterial pauses is not yet clear. However, lambdoid phages use σ -dependent pausing at specific phage promoters to increase transcription of their own DNA. In this case, a phage anti-termination protein binds to the paused, scrunched complex to generate an RNAP that clears the pause and is then insensitive to downstream termination signals within the phage genome.

Termination

Correct termination of transcription at specific sites is crucial as it serves to prevent the unwanted expression of genes that are downstream of a transcribed gene. Termination occurs through two processes: (1) intrinsic termination, in which a sequence within the message alone causes RNAP to release both the DNA and RNA, and (2) factor-dependent termination, in which additional factors together with a sequence result in termination ([Figure 2](#), step 9; reviewed in [Santangelo and Artsimovitch, 2011](#); [Peters et al., 2011](#)). The typical intrinsic sequence is a GC-rich hairpin followed by a stretch of uridines. It is thought that the U-stretch results in a pause, which allows the formation of the hairpin within the RNA exit channel because of the widening of the RNA exit channel that occurs with an elemental pause. This then 'pulls' the RNA from the RNA/DNA hybrid, leading to destabilization of the EC.

Most factor-dependent termination is dependent on Rho protein, an RNA binding, ATP-dependent translocase/helicase. A hexamer of Rho proteins binds to a typically pyrimidine-rich, unstructured part of the message, called a *rut* (Rho utilization) site, and then moves along the RNA to reach the transcription machinery. Rho associates with the transcribing RNAP through an interaction with NusG, which stimulates some Rho-dependent termination. Rho causes termination at specific sequences; however, these sites are not readily predicted, although they typically correspond to a RNAP pause site. Rho is thought to function by using its translocase to push RNAP, by using its helicase to unwind the RNA/DNA hybrid, and/or by using its RNA-binding activity to pull the RNA out of the channel.

Rho serves several global needs of the cell. It clears stalled elongation complexes from the DNA, which could obstruct DNA polymerase and lead to the generation of double strand DNA breaks within the chromosome. Rho helps remove extended regions of RNA/DNA hybridization (R-loops). Finally, Rho down-regulates the expression of horizontally acquired DNA, such as prophages. In fact, deletion of such DNA results in a strain that is much less sensitive to bicyclomycin, an antibiotic that specifically inhibits Rho (Cardinale *et al.*, 2008).

Both intrinsic and Rho-dependent terminations require a conformation or a sequence within the message. Thus, factors or processes that disrupt the formation of an intrinsic hairpin or disrupt the binding of Rho to a *rut* site can result in anti-termination. Since the presence of ribosomes on RNA can prevent hairpin formation or Rho binding, active translation is a common way for the cell to prevent termination in both bacteria and archaea, where transcription and translation occur concurrently. In some cases, this provides specific regulation. For example, at the beginning of the *trp* operon, it is possible for the message to form two different hairpins (reviewed in Yanofsky, 2004; Santangelo and Artsimovitch, 2011). One of these hairpins generates an intrinsic terminator while the other hairpin (the anti-terminator) prevents the terminator hairpin from forming. Two tandem Trp codons are present in a small open reading frame just upstream of this region. When the tryptophan concentration is low, the ribosome stalls at these codons, which leads to the formation of the anti-terminator hairpin. Thus, when tryptophan is needed, the RNAP does not terminate, the *trp* operon is expressed, and tryptophan is synthesized. Conversely, transcription of the *trp* operon is prevented when more tryptophan is not needed. This type of regulation is called attenuation.

In other cases, anti-termination by stalled ribosomes is not so specific. For example, since bacterial genes are typically arranged as operons, the inability to translate a message typically allows Rho to bind, which then results in termination, preventing the expression of downstream genes. This process is called polarity (Adhya and Gottesman, 1978). Rho termination is stimulated by its interaction with NusG. However, NusG also binds to the ribosomal protein S10, and it cannot interact with Rho and the ribosome simultaneously (Burmam *et al.*, 2010). Thus, polarity may also be enhanced by the availability of NusG to interact with Rho when the ribosome is absent.

Another example occurs through a riboswitch, a specific structure in RNA whose creation is modulated by the binding

of a small molecule. If present at the appropriate location in the RNA, the folding of the riboswitch can promote or prevent the formation of an intrinsic terminator or may affect the interaction of Rho with the RNA (Hollands *et al.*, 2012 reviewed in Henkin, 2008). In this way, expression of the gene is regulated by termination in response to the controlling ligand.

Mechanisms that Regulate Bacterial Transcription

Although the presence of optimal promoter elements leads to efficient transcription, this is not ideal if expression of a gene must be controlled. Regulation requires the ability to express genes at specific times and to turn off (repress) or turn on (activate) gene expression in response to the correct conditions. Since transcription is the first step in gene expression, the process of transcription provides multiple steps for this regulation.

Repression

The easiest way to repress transcription is simply to prevent the binding of RNAP to the DNA, thus stopping R_Pc formation. Transcription of the *lac* operon is regulated by repression (Jacob and Monod, 1961). In this operon, the main operator binding site for lac repressor overlaps the promoter. Consequently, binding of a dimer of lac repressors to the operator occludes RNAP directly and prevents expression of genes needed to import and use lactose, an alternate sugar source for the cell. However, there are also weaker operator sites, upstream and downstream of the lac promoter, which can also bind lac repressor dimers. These bound dimers then form tetramers through protein-protein interactions to generate a more tightly repressed complex, a 'repressosome' (reviewed in Bell and Lewis, 2001; Wilson *et al.*, 2007).

In the case of the *gal* operon, which is needed to utilize galactose, the creation of the repressosome is required, since the operator sites only occur upstream and downstream of the promoter and binding by the repressor does not prevent RNAP binding to one of the *gal* promoters. In this case, a dimer of gal repressor binds to each operator and with the aid of the histone-like protein HU, 'loops out' the RNAP/promoter complex, preventing R_Po formation. Interactions between the repressor and the α CTDs of RNAP are thought to stabilize the RNAP/DNA complex in the inactive state (Semsey *et al.*, 2006; Choy *et al.*, 1995).

Repression of either *gal* or *lac* is relieved by the presence of a specific inducer (allolactose, a lactose metabolite for *lac* or D-galactose for *gal*). Binding of the inducer to the repressor changes its conformation and prevents its binding to the operator sites, thus allowing transcription.

Transcription can also be repressed at the step of promoter clearance. At the ribosomal promoter *rrb P1*, RNAP generates an active initiating complex R_Pi in the presence of rNTPs. However, the histone-like protein H-NS can bind simultaneously with RNAP. When H-NS is present, RNAP generates abortive transcripts, but does not proceed to elongation (Schroder and Wagner, 2000).

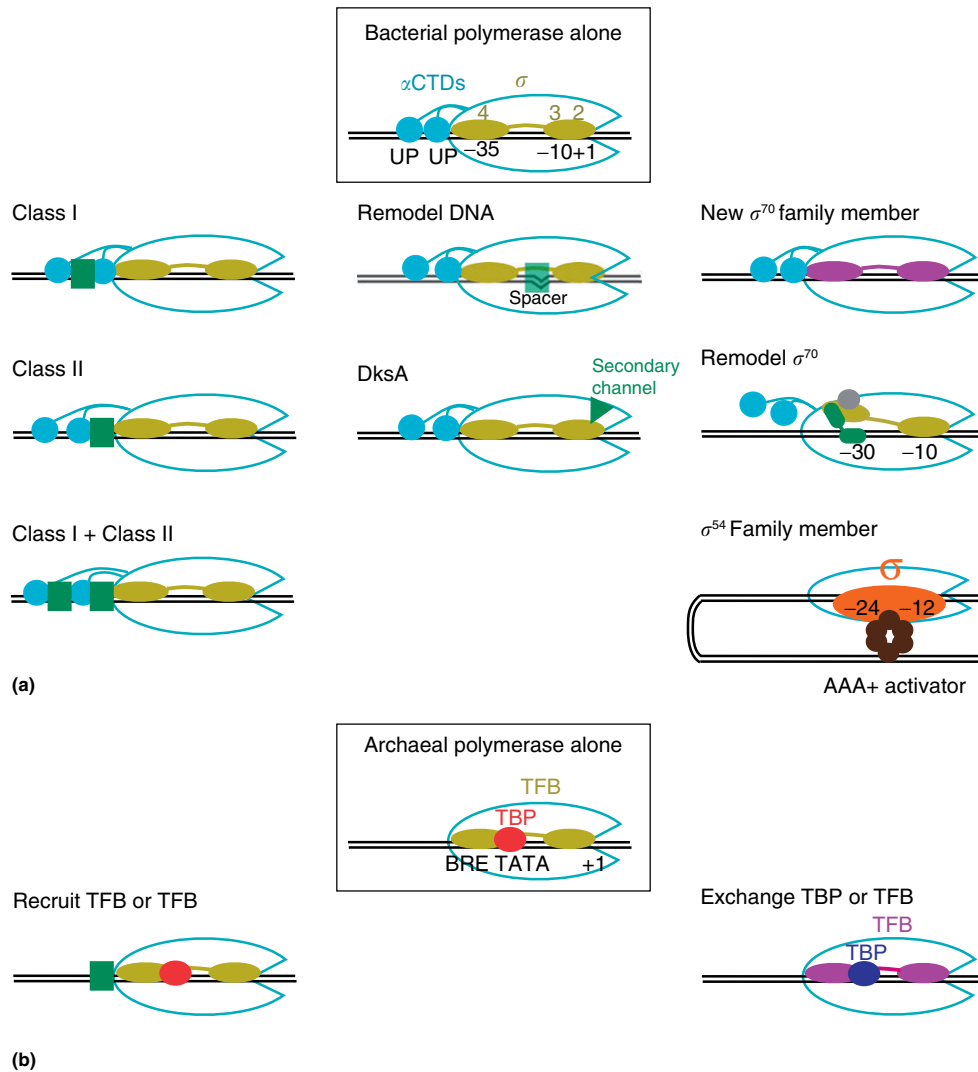


Figure 4 Processes that activate transcription or change promoter specificity. (a) Bacteria. Top shows RNAP/promoter complex made with σ^{70} -containing RNAP. Interactions of α CTDs and σ^{70} with promoter elements are indicated. Underneath shows examples of activation, σ^{70} remodeling, or σ switching. The position of the activator is shown in dark green while the co-activator needed for remodeling is in gray. New σ^{70} family member is in magenta. New σ^{54} member is in orange with the AAA+ activator in black. (b) Archaea. Top shows RNAP/promoter complex made at an archaeal promoter. TFB (in yellow) and TBP (in red) make up the functional equivalent of σ with TFB interacting with the BRE element and TBP interacting with the TATA element. Underneath shows example of an activator in green that is needed to recruit either TBP or TFB or the exchange of the typical TFB or TBP by TFB and TBP-related factors.

The interaction of DksA with the secondary channel of RNAP (Figure 2, step 5) also represses transcription from ribosomal promoters as well as promoters for other components of the translational machinery. DksA binds within the secondary channel, decreasing the stability of transcription initiation complexes at these promoters through the interaction of a portion of DksA (the 'tip') with the β' trigger loop (Lennon *et al.*, 2012). For repression at physiological DksA levels, this process also needs the small molecule ppGpp, which binds to a pocket formed by the ω subunit and a portion of β' , located far from the active site (Mechold *et al.*, 2013; Zuo *et al.*, 2013). Thus, allosteric changes within polymerase must be transmitted by the binding of ppGpp, which then

affect RNAP/DNA interactions in the presence of DksA. As the concentration of ppGpp rises under limiting nutrient conditions, this process redirects transcription to favor transcription of amino acid biosynthesis and alternate σ -dependent redistribution of core polymerase activity.

Activation

As promoters are composed of modular units (UP elements, -35 element, $^{-15}\text{TG}^{-14}$, -10 element), a classic activated promoter consists of a suboptimal number of elements, elements whose sequences deviate substantially from the

consensus, or elements that are not spaced correctly. Thus, these promoters will not be used efficiently by RNAP alone. However, the presence of a factor that interacts with RNAP or the DNA can facilitate the interaction of RNAP with the promoter, resulting in transcription (Figure 4(a); reviewed in Decker and Hinton, 2013; Lee *et al.*, 2012)).

Common mechanisms of activation are called Class I and Class II (reviewed in Busby and Ebright, 1999). In Class I activation, an activator binds upstream of the promoter, commonly centered at -61.5 , and interacts with α CTD's that contact DNA on either side of the activator binding site. The activator/ α CTD interactions help to 'recruit' polymerase to the promoter, thus facilitating R_Pc formation. In Class II activation, the activator binding site is just beside (commonly centered at -41.5) or even overlapping the -35 element. At this position, the activator interacts with both an α CTD that binds to the DNA upstream or to σ Region 4, which interacts with the -35 DNA. Class II activation typically facilitates both R_Pc and R_Po formation. Interestingly, some factors, such as the global activator c-AMP Receptor Protein (CRP), can function as either a Class I or Class II activator depending on the promoter. Class I and Class II activators can also function together, either using multiple activators or one activator that has more than one binding site. Such is the case for CRP, when it binds at multiple sites upstream of -70 at the *malT* promoter, and in this case, facilitates promoter clearance, the last step in transcription initiation (Eichenberger *et al.*, 1997).

Another class of activators works by 'remodeling' the DNA. These factors, exemplified by MerR, are needed at promoters whose spacer region (between the -35 and -10 elements) is longer than ideal (reviewed in Brown *et al.*, 2003). MerR binds to the spacer of the *mer* promoter and together with RNAP, generates an inactive, closed state. Upon binding Hg(II), MerR undergoes a conformation change that distorts the DNA, positioning the -10 and -35 elements correctly and allowing open complex formation. Expression of the *mer* operon then generates the proteins needed for resistance to mercury.

In contrast to its inhibition at ribosomal promoters, the interaction of DksA with the secondary channel of RNAP activates transcription from promoters that direct transcription of genes needed to respond to nutritional stress, such as promoters that drive amino acid biosynthesis. These promoters share A+T-rich discriminator sequences (as opposed to the G+C-rich discriminators at the ribosomal promoters). It is thought that the presence of DksA helps RNAP to clear the promoter, but how DksA modulates activity in response to this sequence is not yet understood (Gummesson *et al.*, 2013).

Switching the Sequence Specificity of Polymerase

The presence of alternate σ factors allows a bacterium to 'switch out' the polymerase specificity subunit, simply exchanging recognition from one set of promoters to another (Figure 4(a); reviewed in Gruber and Gross, 2003; Osterberg *et al.*, 2011). Besides σ^{70} , *E. coli* has six additional σ factors, used for flagella biosynthesis (σ^F), iron transport (σ^{FecI}), nitrogen metabolism and other cellular processes (σ^{54}) and to respond to starvation and general environmental stress (σ^S), heat shock (σ^{H1}), and unfolded proteins in the cell envelope

(σ^E). The number of σ factors encoded by a bacterial genome generally correlates with the complexity of its lifestyle. Organisms that have a varied lifestyle tend to have more σ factors. For example, *Streptomyces coelicolor*, a soil bacterium that engages in mycelial growth and spore formation, has dozens of σ factors. Because the presence of multiple σ factors would mean that there would be competition for core, there are various mechanisms, including targeted proteolysis, regulation by small RNAs, and sequestration of a σ factor by anti-sigma factors (and their regulation by anti-anti-sigma factors!) that assure that the correct σ is present at the correct time.

Since the domains of σ interact with modular promoter elements, it is possible to create a σ with different promoter specificity by remodeling one of its DNA binding domains (Figure 4(a)). Bacteriophage T4 and some T4-like phages do this by using a phage co-activator that binds and structurally alters σ Region 4 (Lambert *et al.*, 2004; reviewed in Hinton, 2010). This restructuring prevents the interaction of σ Region 4 with the -35 DNA and the interaction of σ Region 4/H5 with the β -flap. A phage activator then interacts with a specific sequence within the -30 region of the promoter and with σ H5. Thus, a new ' σ ' Region 4/promoter element is generated. Because the phage promoters retain the typical σ Region 2/ -10 element interaction, the specificity of RNAP switches from recognition of $-35/-10$ promoters to the phage $-30/-10$ promoters.

Activation of RNAP containing σ^{54}

σ^{54} , which represents the other family of σ factors, is distinct in that it requires an activating factor to generate the active transcription complex (reviewed in Wigneshweraraj *et al.*, 2005). Unlike σ^{70} , σ^{54} can bind to its promoter DNA with or without core, interacting with elements positioned at -12 and -24 (Figure 4(a)). However, the σ^{54} /RNAP/DNA complex is a stable R_Pc, which does not by itself proceed to R_Po. Isomerization to the open complex requires an ATPase-activator of the AAA+ class (ATPases Associated with various cellular Activities). An oligomer of activator binds to a region of DNA from 80 to 150 bp upstream of the promoter (referred to as a bacterial enhancer site), and then reaches the σ^{54} /RNAP/promoter complex by looping of the intervening DNA. The conformational change needed to generate the active open complex is driven by the hydrolysis of ATP. Interestingly, while most of these activators need the enhancer binding site, there are ones that work in the absence of DNA binding, suggesting that the enhancer site simply increases the local concentration of the activator (reviewed in Beck *et al.*, 2007).

Transcription and Regulation in Archaea

As seen in Figure 1(a), archaeal core contains subunits that are homologous to β' , β , the 2 α 's, and ω . More importantly, the structural features employed for RNA catalysis (β' bridge helix, β' trigger loop, switch regions, etc.) are conserved throughout biology (reviewed in Decker and Hinton, 2013; Grohmann and Werner, 2011). In particular, the structures of eukaryotic EC's are quite similar to that observed for

the bacterial EC of *T. thermophilus*. Consequently, although the structure of an archaeal EC has not yet been solved, it seems likely the fundamental process of RNA catalysis detailed for bacteria will extend to archaea. However, transcription initiation requires a specificity factor(s), and there are no archaeal or eukaryotic σ factors. In archaea, specificity is provided by TBP (TATA-binding protein) that, like eukaryotic TBP, recognizes a TATA-box [(t/a)(t/a)TATATA] centered at approximately -27 , together with TFB (Transcription Factor B), a homolog of the basal eukaryotic transcription factor TFIIB, that recognizes a specific element (called BRE) located just upstream of the TATA-box sequence (Figure 4(b)). Because TBP and TFB do not share sequence or structural homology with σ factors, the recognition of promoter elements as well as the process of open complex formation was initially thought to be substantially different between bacteria and higher organisms. However, structures (eukaryotic RNA polymerase Pol II with TFIIB (Bushnell *et al.*, 2004) and eukaryotic TBP/TFIIB/DNA (Nikolov *et al.*, 1995; Tsai and Sigler, 2000)) demonstrate that TBP/TFIIB is functionally equivalent to a σ factor. In addition, structure (Littlefield *et al.*, 1999) and photocrosslinking and EM analyses ((Renfrow *et al.*, 2004; Bartlett *et al.*, 2004; De Carlo *et al.*, 2010) reviewed in Jun *et al.*, 2011) indicate that the arrangement of archaeal TBP/TFB within polymerase is similar to that seen in eukaryotic Pol II. In particular, the important interactions of σ with core (Region 2.1/2.2 with the β' CH, Region 3.2 with the RNA exit channel, and Region 4 with the β -flap) are recapitulated by interactions between domains of TFIIB and core polymerase. Furthermore, the protein–DNA interactions of an initiating Pol II/promoter complex resemble those found in bacterial RPo, suggesting that fundamental mechanisms are retained in all organisms (Cheung *et al.*, 2011). Finally, Spt5, the homolog of bacterial NusG in archaea and eukaryotes, is also functionally homologous to NusG, increasing processivity of the elongating polymerase by interacting with the DNA upstream of the RNA/DNA hybrid (reviewed in Blombach *et al.*, 2013).

Given these similarities, it is not surprising that transcriptional regulation in archaea shares features with that observed in bacteria (Figure 4(b)). First, archaea engage in the functional equivalent of ‘sigma switching’ except that in the archaeal version, both TBP and TFB can be exchanged. For example, *Halobacterium* NRC-1 uses six TBPs and seven TFBs in various combinations (Facciotti *et al.*, 2007). As in bacteria with alternate σ factors, alternate TBP/TFB combinations can be helpful to cope with different growth conditions or times of stress. Given that TBP only interacts with one of the promoter elements (the TATA-box), the ability to exchange alternate TBPs generates a situation that bears some resemblance to bacterial σ remodeling by T4 phage.

As detailed earlier, bacterial Class I and Class II activators bind to DNA sites, interacting with subunits of RNAP (σ Region 4 and/or α CTDs) and increasing the binding of polymerase to a nonideal promoter sequence. Similarly, the *Methanococcus jannaschii* activator Ptr2 (Ouhammouch *et al.*, 2003) and the *Pyrococcus furiosus* activator PF1088 (Ochs *et al.*, 2012) bind just upstream of the BRE to recruit TBP or TFB, respectively. Thus, these archaeal activators appear to recapitulate some of the mechanisms that have been characterized in bacterial systems.

Summary

Efficient and well-controlled transcription of genomic DNA into RNA is vital to all organisms. Using a catalytic core and a σ specificity factor, the bacterial RNA polymerase proceeds through an ordered process to ensure that the genome is transcribed accurately. Transcription is regulated by factors that interact with polymerase, DNA, and RNA. Although archaeal core shares homology with the bacterial core, the archaeal specificity factor (TFB plus TBP) is not homologous to σ . Nonetheless, the interactions of TBP and TFP with core demonstrate that they are functionally similar to σ . In addition, mechanisms that activate or switch promoter specificity resemble processes seen in bacteria.

Acknowledgments

This research was supported by the Intramural Research Program of the NIH, National Institute of Diabetes and Digestive and Kidney Diseases.

See also: Nucleic Acid Synthesis/Breakdown: Transcription: Eukaryotic Transcriptional Regulation

References

- Adhya, S., Gottesman, M., 1978. Control of transcription termination. *Annual Review of Biochemistry* 47, 967–996.
- Bae, B., Davis, E., Brown, D., *et al.*, 2013. Phage T7 Gp2 inhibition of *Escherichia coli* RNA polymerase involves misappropriation of sigma70 domain 1.1. *Proceedings of the National Academy of Sciences of the United States of America* 110, 19772–19777.
- Bartlett, M.S., Thomm, M., Geiduschek, E.P., 2004. Topography of the euryarchaeal transcription initiation complex. *Journal of Biological Chemistry* 279, 5894–5903.
- Beck, L.L., Smith, T.G., Hoover, T.R., 2007. Look, no hands! Unconventional transcriptional activators in bacteria. *Trends in Microbiology* 15, 530–537.
- Bell, C.E., Lewis, M., 2001. The Lac repressor: A second generation of structural and functional studies. *Current Opinion in Structural Biology* 11, 19–25.
- Blombach, F., Daviter, T., Fielden, D., *et al.*, 2013. Archaeology of RNA polymerase: Factor swapping during the transcription cycle. *Biochemical Society Transactions* 41, 362–367.
- Borukhov, S., Lee, J., Laptenko, O., 2005. Bacterial transcription elongation factors: New insights into molecular mechanism of action. *Molecular Microbiology* 55, 1315–1324.
- Brown, N.L., Stoyanov, J.V., Kidd, S.P., Hobman, J.L., 2003. The MerR family of transcriptional regulators. *FEMS Microbiology Reviews* 27, 145–163.
- Burmann, B.M., Schweimer, K., Luo, X., *et al.*, 2010. A NusE:NusG complex links transcription and translation. *Science* 328, 501–504.
- Busby, S., Ebricht, R.H., 1999. Transcription activation by catabolite activator protein (CAP). *Journal of Molecular Biology* 293, 199–213.
- Bushnell, D.A., Westover, K.D., Davis, R.E., Kornberg, R.D., 2004. Structural basis of transcription: An RNA polymerase II–TFIIB cocrystal at 4.5 Ångströms. *Science* 303 (5660), 983–988.
- Campbell, E.A., Muzzin, O., Chlenov, M., *et al.*, 2002. Structure of the bacterial RNA polymerase promoter specificity sigma subunit. *Molecular Cell* 9, 527–539.
- Cardinale, C.J., Washburn, R.S., Tadigotla, V.R., *et al.*, 2008. Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science* 320, 935–938.
- Chakraborty, A., Wang, D., Ebricht, Y.W., *et al.*, 2012. Opening and closing of the bacterial RNA polymerase clamp. *Science* 337, 591–595.
- Cheung, A.C., Sainsbury, S., Cramer, P., 2011. Structural basis of initial RNA polymerase II transcription. *EMBO Journal* 30, 4755–4763.

- Choy, H.E., Park, S.W., Aki, T., *et al.*, 1995. Repression and activation of transcription by Gal and Lac repressors: Involvement of alpha subunit of RNA polymerase. *EMBO Journal* 14, 4523–4529.
- De Carlo, S., Lin, S.C., Taatjes, D.J., Hoenger, A., 2010. Molecular basis of transcription initiation in Archaea. *Transcription* 1, 103–111.
- Decker, K.B., Hinton, D.M., 2013. Transcription regulation at the core: Similarities among bacterial, archaeal, and eukaryotic RNA polymerases. *Annual Review of Microbiology* 67:113–139.
- Ebright, R.H., 2000. RNA polymerase: Structural similarities between bacterial RNA polymerase and eukaryotic RNA polymerase II. *Journal of Molecular Biology* 304, 687–698.
- Eichenberger, P., Dethiollaz, S., Buc, H., Geiselmann, J., 1997. Structural kinetics of transcription activation at the *malT* promoter of *Escherichia coli* by UV laser footprinting. *Proceedings of the National Academy of Sciences of the United States of America* 94, 9022–9027.
- Facciotti, M.T., Reiss, D.J., Pan, M., 2007. General transcription factor specified global gene regulation in archaea. *Proceedings of the National Academy of Sciences of the United States of America* 104, 4630–4635.
- Feklistov, A., Barinova, N., Sevostyanova, A., *et al.*, 2006. A basal promoter element recognized by free RNA polymerase sigma subunit determines promoter recognition by RNA polymerase holoenzyme. *Molecular Cell* 23, 97–107.
- Feklistov, A., Darst, S.A., 2011. Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell* 147, 1257–1269.
- Grohmann, D., Werner, F., 2011. Recent advances in the understanding of archaeal transcription. *Current Opinion in Microbiology* 14, 328–334.
- Gruber, T.M., Gross, C.A., 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual Review of Microbiology* 57, 441–466.
- Gummesson, B., Lovmar, M., Nystrom, T., 2013. A proximal promoter element required for positive transcriptional control by guanosine tetraphosphate and DksA protein during the stringent response. *Journal of Biological Chemistry* 288, 21055–21064.
- Haugen, S.P., Ross, W., Manrique, M., Gourse, R.L., 2008. Fine structure of the promoter-sigma region 1.2 interaction. *Proceedings of the National Academy of Sciences of the United States of America* 105, 3292–3297.
- Henkin, T.M., 2008. Riboswitch RNAs: Using RNA to sense cellular metabolism. *Genes & Development* 22, 3383–3390.
- Herbert, K.M., Zhou, J., Mooney, R.A., *et al.*, 2010. *E. coli* NusG inhibits backtracking and accelerates pause-free transcription by promoting forward translocation of RNA polymerase. *Journal of Molecular Biology* 399, 17–30.
- Hinton, D.M., 2010. Transcriptional control in the prereplicative phase of T4 development. *Virology Journal* 7, 289.
- Hirata, A., Klein, B.J., Murakami, K.S., 2008. The X-ray crystal structure of RNA polymerase from Archaea. *Nature* 451, 851–854.
- Hollands, K., Proshkin, S., Sklyarova, S., *et al.*, 2012. Riboswitch control of Rho-dependent transcription termination. *Proceedings of the National Academy of Sciences of the United States of America* 109, 5376–5381.
- Hook-Barnard, I.G., Hinton, D.M., 2007. Transcription initiation by mix and match elements: Flexibility for polymerase binding to bacterial promoters. *Gene Regulation and Systems Biology*. 275–293. Available at: <http://www.lapress.com/transcription-initiation-by-mix-and-match-elements-flexibility-for-pol-article-a481>
- Hsu, L.M., 2002. Promoter clearance and escape in prokaryotes. *Biochimica et Biophysica Acta* 1577, 191–207.
- Jacob, F., Monod, J., 1961. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* 3, 318–356.
- Jun, S.H., Reichlen, M.J., Tajiri, M., Murakami, K.S., 2011. Archaeal RNA polymerase and transcription regulation. *Critical Reviews in Biochemistry and Molecular Biology* 46, 27–40.
- Kapanidis, A.N., Margeat, E., Ho, S.O., *et al.*, 2006. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science* 314, 1144–1147.
- Kaplan, C.D., 2010. The architecture of RNA polymerase fidelity. *BMC Biology* 8, 85–88.
- Klein, B.J., Bose, D., Baker, K.J., *et al.*, 2011. RNA polymerase and transcription elongation factor Spt4/5 complex structure. *Proceedings of the National Academy of Sciences of the United States of America* 108, 546–550.
- Lambert, L.J., Wei, Y., Schirf, V., Demeler, B., Werner, M.H., 2004. T4 AsiA blocks DNA recognition by remodeling sigma(70) region 4. *EMBO Journal* 23, 2952–2962.
- Lee, D.J., Minchin, S.D., Busby, S.J., 2012. Activating transcription in bacteria. *Annual Review of Microbiology* 66, 125–152.
- Lee, H.J., Lim, H.M., Adhya, S., 2004. An unsubstituted C2 hydrogen of adenine is critical and sufficient at the -11 position of a promoter to signal base pair deformation. *Journal of Biological Chemistry* 279, 16899–16902.
- Lennon, C.W., Ross, W., Martin-Tomasz, S., *et al.*, 2012. Direct interactions between the coiled-coil tip of DksA and the trigger loop of RNA polymerase mediate transcriptional regulation. *Genes & Development* 26, 2634–2646.
- Littlefield, O., Korkhin, Y., Sigler, P.B., 1999. The structural basis for the oriented assembly of a TBP/TFB/promoter complex. *Proceedings of the National Academy of Sciences of the United States of America* 96, 13668–13673.
- Martinez-Rucobo, F.W., Cramer, P., 2013. Structural basis of transcription elongation. *Biochimica et Biophysica Acta* 1829, 9–19.
- Martinez-Rucobo, F.W., Sainsbury, S., Cheung, A.C., Cramer, P., 2011. Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO Journal* 30, 1302–1310.
- Mathew, R., Chatterji, D., 2006. The evolving story of the omega subunit of bacterial RNA polymerase. *Trends in Microbiology* 14, 450–455.
- Mechold, U., Potrykus, K., Murphy, H., Murakami, K.S., Cashel, M., 2013. Differential regulation by ppGpp versus pppGpp in *Escherichia coli*. *Nucleic Acids Research* 41, 6175–6189.
- Mekler, V., Kortkhonja, E., Mukhopadhyay, J., *et al.*, 2002. Structural organization of bacterial RNA polymerase holoenzyme and the RNA polymerase-promoter open complex. *Cell* 108, 599–614.
- Murakami, K.S., 2013. X-ray Crystal Structure of *Escherichia coli* RNA Polymerase sigma70 Holoenzyme. *Journal of Biological Chemistry* 288, 9126–9134.
- Murakami, K.S., Darst, S.A., 2003. Bacterial RNA polymerases: The whole story. *Current Opinion in Structural Biology* 13, 31–39.
- Murakami, K.S., Masuda, S., Campbell, E.A., Muzzin, O., Darst, S.A., 2002a. Structural basis of transcription initiation: An RNA polymerase holoenzyme-DNA complex. *Science* 296, 1285–1290.
- Murakami, K.S., Masuda, S., Darst, S.A., 2002b. Structural basis of transcription initiation: RNA polymerase holoenzyme at 4 Å resolution. *Science* 296, 1280–1284.
- Naryshkin, N., Revyakin, A., Kim, Y., Mekler, V., Ebright, R.H., 2000. Structural organization of the RNA polymerase-promoter open complex. *Cell* 101, 601–611.
- Nikolov, D.B., Chen, H., Halay, E.D., *et al.*, 1995. Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* 377, 119–128.
- Ochs, S.M., Thumann, S., Richau, R., *et al.*, 2012. Activation of archaeal transcription mediated by recruitment of transcription factor B. *Journal of Biological Chemistry* 287, 18863–18871.
- Orlova, M., Newlands, J., Das, A., Goldfarb, A., Borukhov, S., 1995. Intrinsic transcript cleavage activity of RNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America* 92, 4596–4600.
- Osterberg, S., Del Peso-Santos, T., Shingler, V., 2011. Regulation of alternative sigma factor use. *Annual Review of Microbiology* 65, 37–55.
- Ouhammouch, M., Dewhurst, R.E., Hausner, W., Thomm, M., Geiduschek, E.P., 2003. Activation of archaeal transcription by recruitment of the TATA-binding protein. *Proceedings of the National Academy of Sciences of the United States of America* 100, 5097–5102.
- Paul, B.J., Ross, W., Gaal, T., Gourse, R.L., 2004. rRNA transcription in *Escherichia coli*. *Annual Review of Genetics* 38, 749–770.
- Perdue, S.A., Roberts, J.W., 2011. Sigma(70)-dependent transcription pausing in *Escherichia coli*. *Journal of Molecular Biology* 412, 782–792.
- Peters, J.M., Vangeloff, A.D., Landick, R., 2011. Bacterial transcription terminators: The RNA 3'-end chronicles. *Journal of Molecular Biology* 412, 793–813.
- Renfrow, M.B., Naryshkin, N., Lewis, L.M., *et al.*, 2004. Transcription factor B contacts promoter DNA near the transcription start site of the archaeal transcription initiation complex. *Journal of Biological Chemistry* 279, 2825–2831.
- Ross, W., Gosink, K.K., Salomon, J., *et al.*, 1993. A third recognition element in bacterial promoters: DNA binding by the alpha subunit of RNA polymerase. *Science* 262, 1407–1413.
- Rutherford, S.T., Villers, C.L., Lee, J.H., Ross, W., Gourse, R.L., 2009. Allosteric control of *Escherichia coli* rRNA promoter complexes by DksA. *Genes & Development* 23, 236–248.
- Saecker, R.M., Record, M.T., Dehaseth, J.R., 2011. Mechanism of bacterial transcription initiation: RNA polymerase-promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *Journal of Molecular Biology* 412, 754–771.
- Santangelo, T.J., Artsimovitch, I., 2011. Termination and antitermination: RNA polymerase runs a stop sign. *Nature Reviews Microbiology* 9, 319–329.
- Schickor, P., Metzger, W., Werel, W., Lederer, H., Heumann, H., 1990. Topography of intermediates in transcription initiation of *E. coli*. *EMBO Journal* 9, 2215–2220.

- Schneider, T.D., 2001. Strong minor groove base conservation in sequence logos implies DNA distortion or base flipping during replication and transcription initiation. *Nucleic Acids Research* 29, 4881–4891.
- Schroder, O., Wagner, R., 2000. The bacterial DNA-binding protein H-NS represses ribosomal RNA transcription by trapping RNA polymerase in the initiation complex. *Journal of Molecular Biology* 298, 737–748.
- Schroeder, L.A., Gries, T.J., Saecker, R.M., *et al.*, 2009. Evidence for a tyrosine-adenine stacking interaction and for a short-lived open intermediate subsequent to initial binding of *Escherichia coli* RNA polymerase to promoter DNA. *Journal of Molecular Biology* 385, 339–349.
- Semsey, S., Virnik, K., Adhya, S., 2006. Three-stage regulation of the amphibolic gal operon: From repressosome to GalR-free DNA. *Journal of Molecular Biology* 358, 355–363.
- Severinov, K., Fenyó, D., Severinova, E., *et al.*, 1994. The sigma subunit conserved region 3 is part of “5'-face” of active center of *Escherichia coli* RNA polymerase. *Journal of Biological Chemistry* 269, 20826–20828.
- Srivastava, A., Talaue, M., Liu, S., *et al.*, 2011. New target for inhibition of bacterial RNA polymerase: ‘Switch region’. *Current Opinion in Microbiology* 14, 532–543.
- Srivatsan, A., Wang, J.D., 2008. Control of bacterial transcription, translation and replication by (p)ppGpp. *Current Opinion in Microbiology* 11, 100–105.
- Travers, A.A., 1984. Conserved features of coordinately regulated *E. coli* promoters. *Nucleic Acids Research* 12, 2605–2618.
- Tsai, F.T., Sigler, P.B., 2000. Structural basis of preinitiation complex assembly on human pol II promoters. *EMBO Journal* 19, 25–36.
- Vannini, A., Cramer, P., 2012. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Molecular Cell* 45, 439–446.
- Vassilyev, D.G., Sekine, S., Laptchenko, O., *et al.*, 2002. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417, 712–719.
- Vassilyev, D.G., Vassilyeva, M.N., Perederina, A., Tahirov, T.H., Artsimovitch, I., 2007. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* 448, 157–162.
- Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D., Kornberg, R.D., 2006. Structural basis of transcription: Role of the trigger loop in substrate specificity and catalysis. *Cell* 127, 941–954.
- Weixlbaumer, A., Leon, K., Landick, R., Darst, S.A., 2013. Structural basis of transcriptional pausing in bacteria. *Cell* 152, 431–441.
- Wigneshweraraj, S., Bose, D., Burrows, P.C., *et al.*, 2008. Modus operandi of the bacterial RNA polymerase containing the sigma54 promoter-specificity factor. *Molecular Microbiology* 68, 538–546.
- Wigneshweraraj, S.R., Burrows, P.C., Bordes, P., *et al.*, 2005. The second paradigm for activation of transcription. Progress in Nucleic Acid Research and Molecular Biology 79, 339–369.
- Wilson, C.J., Zhan, H., Swint-Kruse, L., Matthews, K.S., 2007. The lactose repressor system: Paradigms for regulation, allosteric behavior and protein folding. *Cellular and Molecular Life Sciences* 64, 3–16.
- Yang, X., Lewis, P.J., 2010. The interaction between bacterial transcription factors and RNA polymerase during the transition from initiation to elongation. *Transcription* 1, 66–69.
- Yang, X., Molimau, S., Doherty, G.P., *et al.*, 2009. The structure of bacterial RNA polymerase in complex with the essential transcription elongation factor NusA. *EMBO Reports* 10, 997–1002.
- Yanofsky, C., 2004. The different roles of tryptophan transfer RNA in regulating trp operon expression in *E. coli* versus *B. subtilis*. *Trends in Genetics* 20, 367–374.
- Zhang, G., Campbell, E.A., Minakhin, L., *et al.*, 1999. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* 98, 811–824.
- Zhang, Y., Feng, Y., Chatterjee, S., *et al.*, 2012. Structural basis of transcription initiation. *Science* 338, 1076–1080.
- Zuo, Y., Wang, Y., Steitz, T.A., 2013. The mechanism of *E. coli* RNA polymerase regulation by ppGpp is suggested by the structure of their complex. *Molecular Cell* 50, 430–436.

Further Reading

- Decker, K.B., Hinton, D.M., 2013. Transcription regulation at the core: Similarities among bacterial, archaeal, and eukaryotic RNA polymerases. *Annual Review of Microbiology* 67:113–139.
- Grohmann, D., Werner, F., 2011. Recent advances in the understanding of archaeal transcription. *Current Opinion in Microbiology* 14, 328–334.
- Gruber, T.M., Gross, C.A., 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annual Review of Microbiology* 57, 441–466.
- Hook-Barnard, I.G., Hinton, D.M., 2007. Transcription initiation by mix and match elements: Flexibility for polymerase binding to bacterial promoters. *Gene Regulation and Systems Biology*. 275–293. Available at: <http://www.la-press.com/transcription-initiation-by-mix-and-match-elements-flexibility-for-pol-article-a481>.
- Lee, D.J., Minchin, S.D., Busby, S.J., 2012. Activating transcription in bacteria. *Annual Review of Microbiology* 66, 125–152.
- Murakami, K.S., 2013. X-ray Crystal Structure of *Escherichia coli* RNA Polymerase sigma70 Holoenzyme. *Journal of Biological Chemistry* 288, 9126–9134.
- Saecker, R.M., Record Jr, M.T., Dehaseth, P.L., 2011. Mechanism of bacterial transcription initiation: RNA polymerase - promoter binding, isomerization to initiation-competent open complexes, and initiation of RNA synthesis. *Journal of Molecular Biology* 412, 754–771.
- Santangelo, T.J., Artsimovitch, I., 2011. Termination and antitermination: RNA polymerase runs a stop sign. *Nature Reviews Microbiology* 9, 319–329.
- Tomar, S.K., Artsimovitch, I., 2013. NusG-Spt5 proteins – Universal tools for transcription modification and communication. *Chemical Reviews* 113, 8604–8619.
- Vassilyev, D.G., Vassilyeva, M.N., Perederina, A., Tahirov, T.H., Artsimovitch, I., 2007. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* 448, 157–162.
- Weixlbaumer, A., Leon, K., Landick, R., Darst, S.A., 2013. Structural basis of transcriptional pausing in bacteria. *Cell* 152, 431–441.

Eukaryotic Transcriptional Regulation

BA Lewis, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Eukaryotic transcriptional regulation is an enormous topic that has been researched extensively over the past 30 years or so. Its relevance to virtually every biological process and human disease cannot be understated. Cancer, for example, is clearly due to the dysregulation of gene expression, wherein growth promoting genes (such as the c-myc oncogene) are on for too long, making too much RNA and thus too much protein. Thus, an understanding of how genes are turned on and off is of paramount importance, both as a fundamental biological problem and as a problem in human disease.

Historically, it may be useful to begin with a brief mention of some of the major thematic ideas in eukaryotic transcriptional regulation. The first major discovery was the identification of the three eukaryotic polymerases in 1969 (Roeder and Rutter, 1969), surprising because it deviated from the at-the-time paradigmatic one bacterial RNA polymerase. The next significant step was the discovery of discrete well-defined DNA sequences called promoters that established starting points for the polymerase to begin copying the gene into RNA (Dierks *et al.*, 1983, 1981). This was closely followed by the identification and purification of DNA binding proteins that could activate transcription. Concomitantly, the general factors that established transcriptionally active polymerase were purified and cloned (Thomas and Chiang, 2006). These discoveries were followed by an understanding of how packaging DNA into chromosomes affected transcriptional regulation. These thematic ideas were further extended with the discovery of chromatin modification enzymes and epigenetics. These discoveries often overlapped with each other and the demarcation of the time line of discoveries above should only be considered very coarsely. Additionally, these advances were often extensively influenced by previous work on transcription in bacteria, where several of these major themes were initially discovered. Nevertheless, these are roughly how the field has evolved, and in this article one should get a sense of the basics of transcriptional regulation in eukaryotes.

RNA Polymerases I, II, and III

There are three RNA polymerases in eukaryotes, in contrast to one in bacteria. These three polymerases, I, II, and III, transcribe ribosomal genes (pol I), protein-coding genes (pol II; and some non-coding genes/RNAs), and the 5S RNA and tRNA genes (pol III). All three are large multi-subunit enzyme complexes which share five subunits but also contain polymerase-specific subunits (Vannini and Cramer, 2012). Each polymerase has its own catalytic domain, which in pol II consists of the two largest subunits, Rpb1 and Rpb2. These two subunits comprise the cleft and trigger loop of pol II through which one strand of the DNA template threads itself. The addition of a ribonucleotide begins with the binding of an

NTP to the active site where presumably proofreading for the correct NTP occurs, followed by a metal-based catalysis that extends the RNA one base while releasing pyrophosphate. This is followed by a translocation step to move the RNA out of the active site, and allowing the cycle to begin again (Martinez-Rucobo and Cramer, 2013).

None of the three polymerases is capable of initiating transcription on its own. Instead, each polymerase requires a unique but partially overlapping set of accessory factors. The partial overlap is due to the requirement for the TATA binding protein (TBP) for all three polymerases (Vannini and Cramer, 2012). Originally discovered as a pol II-specific protein that bound to TATA box sequences in promoters (see below), it was later found to be part of the pol I and III machinery. In addition to TBP, the general transcription factor (GTF) proteins and protein complexes are necessary for the proper recruitment of the polymerase to a promoter. Among the three polymerases, the protein complement and promoter structure of the pol II genes is by far the most complex and diverse.

Promoters and Transcription Factors

The structure of a gene in general consists of an upstream promoter region that encompasses the transcriptional start site (TSS) and perhaps 50 bp of DNA downstream of the TSS. The upstream boundary of the promoter region varies but can be several kilobases in length. The core promoter comprises the region in the immediate vicinity of the TSS and in some cases defines the position of a discrete TSS. However, many genes have multiple TSSs whose positioning is not understood. Core promoter elements (CPEs) are discrete functionally defined DNA sequences that when mutated result in reduced levels of transcription. The most well-known is the TATA box (consensus sequence TATAAA) and it is found approximately 30 bp upstream of the TSS (Dierks *et al.*, 1983; McKnight *et al.*, 1981; McKnight and Kingsbury, 1982). Some promoters have an initiator element (Inr) at the TSS (consensus sequence YYA₊₁NA/TYY) (Smale and Baltimore, 1989). The Inr element can establish a TSS without need of a TATA box, although they can be found together. Bioinformatic analysis suggests that only about 10% of genes in the human genome contain TATA boxes and possibly half of all genes can be assigned known core promoter sequences. This suggests that we do not really have an understanding of core promoter architecture and TSS selection and regulation for at least half of the genome.

For many years it was assumed that no CPEs existed downstream of the TSS, but such elements have now been described in detail and display the most heterogeneity of all CPEs. The first exact sequence was the downstream promoter element (DPE), preluded by additional evidence that such elements existed (Burke and Kadonaga, 1996; Nakatani *et al.*, 1990; Smale and Baltimore, 1989). Afterwards, the downstream core element (DCE), motif ten element (MTE),

and X Core Promoter Element 1 and 2 (XCPE1 and 2) were characterized (Lim *et al.*, 2004; Lee *et al.*, 2005; Lewis *et al.*, 2000; Anish *et al.*, 2009; Tokusumi *et al.*, 2007). The DPE and DCE are the most extensively characterized, and illustrate some of the heterogeneity in downstream elements. Spacing relative to the TSS is critical as moving either the DPE or any of the three DCE subelements relative to the TSS decreases transcription (Burke and Kadonaga, 1996; Lee *et al.*, 2005; Lewis *et al.*, 2000). Similarly, spacing between the TATA box and Inr is important (Smale and Kadonaga, 2003). In the case of the DCE these spacing disruptions likely change the position of the DNA elements away from an optimal helical surface necessary for binding of the TFIID complex necessary for core promoter recognition (Lee *et al.*, 2005; Lewis *et al.*, 2000). Additionally, the sequences of all of these downstream elements are different from each other and likely bind different proteins. For example, the DPE binds the TAF6/9 component of the TFIID complex while the DCE subelements bind the large TAF1 component of TFIID (Burke and Kadonaga, 1997; Lee *et al.*, 2005). The additional distinction between the DCE and DPE in human systems is the requirement for only TFIID for DPE function while DPE function additionally requires other cofactors such as a kinase activity (Lewis *et al.*, 2005).

The core promoter functions to establish a TSS and recruit RNA polymerase II to the gene. It does so by serving as a nucleating point, recruiting the GTFs to the promoter via the various CPEs. The GTFs are the family of factors thought to be necessary for transcription from all RNA pol II-dependent genes. There are similar sets of GTFs necessary for RNA pol I and III transcription, albeit of a somewhat reduced complexity and number. The RNA pol II GTFs are TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH (Thomas and Chiang, 2006; Buratowski *et al.*, 1989; Horikoshi *et al.*, 1988; Lu *et al.*, 1991). Biochemically, the order of recruitment has been worked out. The TFIID complex binds to the core promoter via TATA box/Inr/DPE/DCE recognition along with TFIIA and TFIIB. The large TFIID complex consists of the TBP and 12 TAFs (TATA binding protein associated factors) (Thomas and Chiang, 2006). The TAFs contain the Inr, DPE, and DCE DNA binding activities as well as functioning to interact with activators (Burke and Kadonaga, 1997; Chalkley and Verrijzer, 1999; Lee *et al.*, 2005). In some cases, an activator recruits TFIID to the promoter where it then binds to various CPEs in the promoter. TFIIA/B/D recruitment to the promoter is followed by the tandem recruitment of RNA pol II and TFIIF, which in turn recruit TFIIE and TFIIH. This complex in general is referred to as the pre-initiation complex or PIC. The subsequent addition of ribonucleotides results in the melting of the DNA by the TFIIH complex helicase ERCC2 allowing RNA pol II access to a single-stranded DNA template necessary for RNA synthesis (Thomas and Chiang, 2006). It is not clear to what extent this ordered recruitment occurs *in vivo*, to what extent it occurs on different promoters with different activators and enhancers, or whether or not a paused pol II is being established. It is also not clear to what extent this model is true in the context of the discovery of many more factors involved in initiation of transcription. Indeed, *in vivo* experiments that assessed GTF recruitment during the activation of a promoter indicated that the ordered recruitment model was not observed *in vivo* (Soutoglou and Talianidis, 2002). Secondly, GTF requirements

for genes with multiple TSSs and the factor requirements for several of the DEs are not known.

The region around the TSS is typically not sufficient for higher levels of transcription; only low basal levels of transcription are obtained. Upstream of the core promoter region are often additional discrete DNA sequences that bind transcriptional activators and repressors (see below) that are necessary for high levels of transcription. Typically, these proteins are modular in their structure, consisting of a discrete DNA binding domain (DBD) such as a zinc finger motif (Ptashne, 2004). Activation of a promoter is mediated by a second domain of the protein, the transcriptional activation domain (TAD). TADs physically interact with the GTFs and the Mediator coactivator complex, a large two megadalton complex, containing 20 proteins, which interact with a variety of transcriptional activators to stimulate transcription (Poss *et al.*, 2013). Elongation factors such as P-TEFb and coactivators such as the acetyltransferases can be recruited as well (Peterlin and Price, 2006; Nagy and Tora, 2007). Both DBDs and TADs are truly modular and can function independently of one another, often being used to make hybrid DNA binding proteins and activators and which greatly facilitated their study. Furthermore, different DBDs and TADs bind different DNA sequences and coactivator proteins, respectively (Cox and Phillips, 2007). This likely allows for considerable combinatorial regulation to occur on any one gene. But it is still not clear how or whether multiple activators can stimulate a promoter concomitantly.

Enhancers are a second type of promoter activation DNA region, but can act many kilobases away from the TSS and can be positioned up- or downstream of the target promoter (Bulger and Groudine, 2010). Enhancers are a collection of activator binding sites but their histone modification signatures are distinct from promoters (Smallwood and Ren, 2013). The exact mechanism of enhancer function is not completely understood but significant advances have come from the discovery that RNA pol II-dependent transcription can begin within an enhancer and that the resulting RNA is necessary for enhancer-promoter communication. There is increasing evidence that this communication entails the physical looping of an enhancer to the target promoter (Orom and Shiekhattar, 2011).

Lastly, one final class of DNA sequences is the boundary or insulator elements. These elements function to delineate domains between genes and 'insulate' promoters from improper activation from enhancers. These boundaries are established by the CTCF protein that interacts with a boundary element, possibly by determining discrete topological domains within a chromosome that then prevent access of enhancers outside of that domain (Ong and Corces, 2014).

Transcription Initiation and Elongation

Once the PIC (pre-initiation complex) has formed at the promoter, RNA pol II is ready to begin transcribing the gene. Biochemically, the addition of the four ribonucleotides is sufficient to begin transcription. Additionally, the TFIIH complex DNA helicases, ERCC2 and ERCC3 utilize ATP to melt the double-stranded DNA, creating a transcription

bubble. RNA pol II requires bubble formation and access to single-stranded DNA to begin RNA synthesis. Concomitantly, the serine 5 residues of the RNA pol II carboxy-terminal domain (CTD) are phosphorylated by the CDK7/cyclin H kinase activity within the TFIH complex (Sims *et al.*, 2004).

Within the past 5 years or so the importance of transcriptional elongation and its regulation has become increasingly clear. In the late 1980s it was found that a subset of genes in *Drosophila* involved in the heat shock response contained an RNA pol II that was not at the promoter but was transcriptionally engaged and paused roughly 50 bp downstream of the TSS (Rougvie and Lis, 1988). It is an engaged pol II, meaning that it has a short capped mRNA and has an open DNA bubble bound by pol II. These paused polymerases were thought to encompass a minority of genes such as heat shock promoters and growth-factor-induced genes such as c-myc. The advent of genome-wide techniques changed this view and now paused pol II is thought to exist at the lower estimate of 40% of genes in the genome and thus has become quite significant (Adelman and Lis, 2012).

The predominant model of pausing requires recognition and binding of RNA pol II by the DSIF–NELF complex, which stops the elongating polymerase. The release into productive elongation occurs with the phosphorylation of DSIF (and the RNA pol II CTD serine 2 residue) by a second CTD kinase, P-TEFb. P-TEFb is thought to be regulated by various mechanisms such as recruitment by an activating transcription factor. Additionally, there are several proteins that directly affect elongation rates of RNA pol II, such as elongins, eleven nineteen lysine-rich leukemia (ELL), and TFIIS. ELL also participates in elongation as a member of the super elongation complex (SEC) (Zhou *et al.*, 2012). Additionally, the CDK7 CTD kinase activity of TFIH is also required for pausing (Glover-Cutter *et al.*, 2009; Schwartz *et al.*, 2003). Although pausing was originally thought to be necessary for rapid induction of transcription in response to stress such as heat shock, it is not a prerequisite of rapid induction. Whether pausing has a particular regulatory function or whether it is part of the evolution of additional transcriptional initiation mechanisms in higher eukaryotes is not clear.

Pol II C-Terminal Domain

The CTD of RNA pol II consists of a heptad repeat of YSPTSPS over the first 26 repeats and then diverges somewhat from the consensus sequence, predominantly at serine 7, replaced by lysine residues. Y1, S2, T4, S5, and S7 can all be phosphorylated by multiple kinases. Serine 5 phosphorylation is a prerequisite for the recruitment of capping enzyme (required for the addition of the m7G cap to the transcribed RNA) which itself binds to the phosphoserine 5 residue (Wen and Shatkin, 1999; Fabrega *et al.*, 2003). Serine 2 and serine 7 phosphorylations also recruit factors to the transcribing pol II (Egloff *et al.*, 2010, 2012). The two proline residues are targeted by the Pin1 prolyl isomerase, which converts the proline from a *trans* to *cis* conformation (Hanes, 2014). This conformational change is necessary for the recognition of phosphoserine 2 by the FCP1 phosphatase. T4 and S5 can also be O-GlcNAcylated by the O-GlcNAc transferase (Kelly *et al.*, 1993; Ranuncolo

et al., 2012). The lysines in the C-terminal half of the CTD can be acetylated by the p300 acetyltransferase, and the one R residue can be methylated by CARM1 (Sims *et al.*, 2011; Schroder *et al.*, 2013). Additionally, FCP1, RPAP2 are phosphatases acting on phosphoserines 2 and 5 respectively (Eick and Geyer, 2013; Phatnani and Greenleaf, 2006).

The phosphorylation state of the CTD is mostly important for various posttranscriptional events (Corden, 2013). These include the recruitment of proteins involved in mRNA capping, splicing, addition of the poly A tail to the mRNA, and termination of transcription. In many cases, these events are cotranscriptional, where the mRNA is processed while the RNA pol II continues the synthesis of the RNA. Histone methyltransferases (HMTs) are likewise recruited by the phosphorylated CTD cotranscriptionally and then add various combinations of methyl groups to the histone H3 tail (see below).

Termination of Transcription

The final stage of transcription is the actual termination of RNA synthesis by RNA pol II and its release from the DNA template (Proudfoot, 2011; Richard and Manley, 2009). The process begins with the cleavage factor Pcf11 binding to the serine 2 phosphorylated CTD of pol II. RNA cleavage occurs 3' to the poly A site sequence, releasing the mature mRNA from the pol II. Then a 5'–3' exonuclease (Xrn2 in humans and Rat1 in yeast) is recruited to degrade the remaining RNA presumably until contact with pol II is made and termination of RNA synthesis occurs. This model is commonly referred to as the torpedo model. Some genes also need an RNA–DNA helicase called Senataxin in humans and Sen1 in yeast, which is also recruited to the phosphorylated CTD. Among these genes are the snoRNA genes which require a complex containing Nrd1 and Nab3, along with the Sen1 helicase. More recent work suggests that these two models should be combined to create a hybrid model that incorporates the mechanisms from each, as Nrd1 has been found to colocalize with Pcf11. The Nrd1 pathway at this time only exists in yeast and no homologs of Nrd1 and Nab3 have been found in humans. Although the RNA cleavage and exonuclease degradation of the RNA are understood, it is not clear what the mechanism of termination of transcription is.

Chromatin

In the nucleus the linear DNA strand is compacted by wrapping itself around the nucleosome. A nucleosome consists of four pairs of histone proteins, H2A, H2B, H3, and H4. DNA wraps around the nucleosome 1.8 times (including the H1 linker), equivalent to 180 bp of DNA. The presence of a nucleosome creates a problem – how does a polymerase move through nucleosomal DNA (Kulaeva *et al.*, 2013)? This physical barrier requires some mechanism by which the polymerase can transcribe a gene efficiently. Facilitates chromatin transcription (FACT) was isolated using an assay designed to detect activities that could aid RNA pol II transcription on a chromatin template *in vitro*. It is thought that FACT functions as a histone chaperone, disassembling and reassembling the

histone octamer as the polymerase transcribes the DNA (Belotserkovskaya *et al.*, 2003; Hsieh *et al.*, 2013). A second problem with chromatin is that DNA binding proteins often are unable to recognize and bind their sites in nucleosomal DNA. Various chromatin remodeling complexes can move a nucleosome in an ATP-dependent manner to expose the DNA binding site (Clapier and Cairns, 2009).

One additional complexity that chromatin has is an extensive array of posttranslational modifications on the histones (Rothbart and Strahl, 2014). The N-terminus 'tail' of each histone is approximately 30 residues in length which can be acetylated, methylated, ubiquitylated, O-GlcNAcylated, phosphorylated, and citrullinated. These modifications are thought to comprise a 'histone code' and many of these serve as binding sites to recruit additional proteins involved in a variety of functions (Gardner *et al.*, 2011). For example, Brd4 has a bromodomain that binds acetylated lysines, and several proteins have chromodomains that bind methylated residues. These modifications are removed by phosphatases, demethylases, deacetylases, etc.

Each modification is associated with particular functions or states of transcription. For example, trimethyl lysine 4 (H3K4me3) modification of the H3 tail is found concentrated at the 5' ends of actively transcribed genes while H3 K36 methylation is found within the gene itself. Repressed states are associated with H3 K9 and K27 methylation, and respectively recruit the repressive HP1 and Polycomb repressive complex (PRC) to repress transcription. Similarly, active genes are heavily acetylated while histone deacetylases are parts of repressive protein complexes. The interplay and regulation of the histone code modifications is known as epigenetics and is currently under intense study.

***In Vivo* Dynamics and Behavior**

Biochemical approaches indicated that DNA binding proteins interacted with their respective specific DNA sequences with very high affinity. Stable binding was essential for their detection and characterization. However, *in vivo* studies using a variety of techniques suggest a far different picture. Residence times of transcription factors bound to DNA sites indicate rapid on/off times measured in seconds (Voss and Hager, 2014; Hager *et al.*, 2009). This has been observed with a variety of DNA binding factors and so presents a significant departure from the biochemical view of protein–DNA interactions.

Expression analysis of promoters both *in vitro* and *in vivo* has relied on a population-based approach. A promoter is linked to a gene whose expression is easily detected. For example, luciferase expression is transfected into a population of cells and changes in luciferase levels are measured upon stimulation of the promoter or by making mutations in the promoter. Similarly, *in vitro* transcription assays measure the expression of many copies of a cloned promoter added to a nuclear extract. These population assays show a continuum of expression levels.

Single-cell analysis with fluorescent microscopy shows a far different set of expression dynamics (Darzacq *et al.*, 2009). Expression levels from individual genes are stochastic, consisting of bursts of transcription followed by periods of

inactivity. This bursting can be modeled as a probability of the promoter being in either an 'on' or 'off' state. Additionally, these behaviors indicate that the equilibrium models of transcription are not correct and instead nonequilibrium conditions exist, wherein energy requirements and the thermodynamics of the kinetic processes become important. These considerations led to the random telegraph model of transcription where the frequency of initiation and the length of on–off state are stochastic processes resembling random messages sent along a telegraph wire (Larson, 2011; Palangat and Larson, 2012).

Dynamics can be seen at other levels as well. Networks of genes contain specific motifs or 'rules' that dictate time-dependent promoter behavior (Alon, 2007b; Tyson and Novak, 2010). An example of this is a feed-forward loop that involves transcription factor A turning on the expression of transcription factor B and in turn both A and B act in concert to turn on gene C. The internal behaviors of A and B then dictate the temporal expression pattern of C. In turn, these motifs are often part of a larger gene regulatory network whose cumulative expression output is governed by the collective behavior of these individual circuits.

Supercoiling

The consequence of any polymerase moving through a gene and unwinding the DNA in front of it results in the accumulation of positive supercoiling in front and negative supercoiling behind the polymerase. The positive supercoiling creates a barrier that resists any further progress of the polymerase. Continued polymerase movement requires a topoisomerase to remove this torsional stress by nicking one strand of DNA to allow its unwinding. Supercoiling thus occurs as a consequence of transcription and has important consequences for gene expression (Baranello *et al.*, 2012; Kouzine *et al.*, 2014).

Experimental Approaches

Protein Biochemistry and Transcription *In Vitro*

Gene regulation, and indeed life in general, is the manifestation of biochemistry at many different scales. The cell itself is a biochemical machine of enormous complexity. Therefore, a major effort at understanding gene regulation by necessity must comprise a thorough understanding of the underlying protein biochemistry. Protein purifications were essential to identifying and isolating the RNA polymerases, GTFs, and elongation factors, and the order they assemble on a promoter (Maldonado *et al.*, 1996). Many coactivators, DNA binding proteins, repressors, and activators were purified based on binding to their DNA sites (Ryu and Tjian, 1999; Briggs *et al.*, 1986). Similarly, histone modification enzymes were purified using assays that detected a specific posttranslational modification (Nishioka *et al.*, 2002a,b). What makes protein biochemistry so incredibly powerful is the ability to design functional assays to detect new activities and ultimately new proteins or known proteins having previously unknown functions. These assays typically rely on nuclear extracts, which are crude mixtures of proteins extracted from the nuclei of

cells. These extracts can specifically activate promoters with only the addition of ribonucleotides and the promoter DNA. The power of the biochemistry is also in its unbiased approach to identifying the proteins involved. A functionally based biochemical assay simply asks for the reconstruction of a particular activity or effect. The number of proteins or what proteins they are is irrelevant. The assay tells the researcher what protein(s) are important. In the end, it is the functional biochemistry that allows the researcher to determine protein functions directly. Many of these and other assays are discussed in detail in [Carey et al. \(2009\)](#). *The Methods in Enzymology* series and *Current Protocols* are also excellent resources.

Genetics

Biochemical approaches are complemented by the genetics of some systems, such as *Saccharomyces cerevisiae* (yeast) and *Drosophila melanogaster* (fruit flies). The success of such genetic screen approaches using yeast is apparent in the genetic identification of many of the general factors and elongation machinery, and it became apparent that many thematic ideas in transcription as well as the proteins necessary for gene expression were conserved from yeast to humans. Even before the biochemistry was worked out, *Drosophila* genetics had identified transcription factors involved in many developmental processes. Similarly, these developmental processes were highly conserved from flies to mice and humans.

The power of a genetic approach was further increased with the discovery and application of short hairpin RNAs (shRNA) to target endogenous RNAs and thereby decrease expression of the gene the shRNA is targeting. This technique has greatly expanded the tools of the biochemist in that biochemical results now can be readily tested *in vivo* without resorting to the more laborious mouse gene ablation approaches. The development of targeted RNA ablation technologies (RNAi) such as shRNA were perhaps even more important in that they basically gave researchers a tool to easily target and ablate specific genes where no genetic approaches were feasible, such as with cell lines that are growing in incubators (there is no germ-line to target). RNA ablation assays are essential tools in gathering further evidence of protein functions determined biochemically. The caveat to all genetic approaches to keep in mind is that effects may often be indirect and pleiotropic, but when combined with the more direct evidence using biochemistry, allow the researcher to understand gene regulation from several perspectives.

Other techniques greatly expanded our early understanding of transcription. Electrophoretic mobility shift DNA binding assays, and DNase I and chemical footprinting assays gave clear pictures of the DNA bound by transcription factors. Nuclear run-on experiments were essential in determining polymerase distributions across a gene and were used to first identify paused RNA pol II on *Drosophila* heat shock genes. Many of these and other assays are discussed in detail in [Carey et al. \(2009\)](#).

Molecular Biological Techniques

A significant amount of work was achieved through the use of various reporter assays that could be conducted in tissue culture. These reporter assays consisted of a DNA plasmid

containing the promoter of interest fused to a gene whose expression could be easily detected such as chloramphenicol acetyltransferase or luciferase genes. The assays for these genetic reporters are straightforward and could be done using a variety of permutations also aided by the various mutagenesis techniques developed over the years.

Genome-Wide Analysis

The past 10 years or so have seen the emergence of powerful genome-wide approaches, facilitated by the development of high-throughput DNA sequencing technologies ([van Dijk et al., 2014](#); [Furey, 2012](#)). The invention of ChIP and ChIP-seq (chromatin immunoprecipitation-high-throughput sequencing or ChIP-seq) assays were a powerful addition to the biochemist's arsenal. These approaches have given a window into the genome-wide distribution of transcription factors and unprecedented information on RNA expression levels ([Wilhelm and Landry, 2009](#); [Core et al., 2008](#); [Kwak et al., 2013](#); [Churchman and Weissman, 2011](#)). These datasets are analyzed by very sophisticated bioinformatics and computational software in order to gain insight into regulatory behaviors genome-wide.

Computational Analysis and Systems Biology

One of the newest experimental approaches to emerge is a computational one. These studies are mathematically based, utilizing data from experiments to model and understand more completely the complex behavior regulatory systems employ ([Alon, 2007a,b](#); [Milo et al., 2002](#); [Rue and Garcia-Ojalvo, 2013](#)). The data sources can be based on the ones we have already mentioned. It is these computational approaches that have added significantly to our understanding of the stochastic *in vivo* behavior of transcription dynamics ([Munsky et al., 2012](#); [Maheshri and O'Shea, 2007](#)).

Some Outstanding Issues and Questions

The following is a short list of some of the unresolved issues and questions in transcriptional regulation. There are surely more than those listed here but the list should give the reader some idea of current and future research in transcriptional regulation.

1. What is the entire content of the pre-initiation complex at promoters? What are the more physiologically relevant factors involved in transcription initiation?
2. What is the mechanism of initiation? That is, what is the signal that the PIC receives to begin transcription? What causes initiation to stop?
3. How are some genes paused and some genes not? Why are they regulated differently?
4. What is the mechanism of tissue-specific gene expression, especially when considered *in vivo* where multiple transcriptional activators can bind to the same promoter? What is the mechanism of transcription when multiple activators are employed?
5. How do transcription factors find their correct binding sites at genes where activation (or repression) is intended?

6. What is the mechanism of enhancer-mediated transcriptional activation?
7. How does the CTD of RNA polymerase II manage the variety of posttranslational modifications throughout the transcription cycle?
8. Do larger-scale structures and domains exist in the nucleus and do they have regulatory functions?
9. What is the mechanism of transcription termination?

See also: Dynamic Integration: Dynamics: Transcription Factor Networks. Nucleic Acid Synthesis/Breakdown: Transcription: Distant Activation of Transcription by Enhancers

References

- Adelman, K., Lis, J.T., 2012. Promoter-proximal pausing of RNA polymerase II: Emerging roles in metazoans. *Nature Reviews Genetics* 13, 720–731.
- Alon, U., 2007a. An Introduction to Systems Biology: Design Principles of Biological Circuits. Boca Raton, FL: Chapman & Hall/CRC.
- Alon, U., 2007b. Network motifs: Theory and experimental approaches. *Nature Reviews Genetics* 8, 450–461.
- Anish, R., Hossain, M.B., Jacobson, R.H., Takada, S., 2009. Characterization of transcription from TATA-less promoters: Identification of a new core promoter element XCPE2 and analysis of factor requirements. *PLoS One* 4, e5103.
- Baranello, L., Levens, D., Gupta, A., Kouzine, F., 2012. The importance of being supercoiled: How DNA mechanics regulate dynamic processes. *Biochimica et Biophysica Acta* 1819, 632–638.
- Belotserkovskaya, R., Oh, S., Bondarenko, V.A., et al., 2003. FACT facilitates transcription-dependent nucleosome alteration. *Science* 301, 1090–1093.
- Briggs, M.R., Kadonaga, J.T., Bell, S.P., Tjian, R., 1986. Purification and biochemical characterization of the promoter-specific transcription factor, Sp1. *Science* 234, 47–52.
- Bulger, M., Groudine, M., 2010. Enhancers: The abundance and function of regulatory sequences beyond promoters. *Developmental Biology* 339, 250–257.
- Buratowski, S., Hahn, S., Guarente, L., Sharp, P.A., 1989. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* 56, 549–561.
- Burke, T.W., Kadonaga, J.T., 1996. *Drosophila* TFIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes & Development* 10, 711–724.
- Burke, T.W., Kadonaga, J.T., 1997. The downstream core promoter element, DPE, is conserved from *Drosophila* to humans and is recognized by TAFII 60 of *Drosophila*. *Genes & Development* 11, 3020–3031.
- Carey, M., Peterson, C.L., Smale, S.T., 2009. Transcriptional Regulation in Eukaryotes: Concepts, Strategies, and Techniques. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Chalkley, G.E., Verrijzer, C.P., 1999. DNA binding site selection by RNA polymerase II TAFs: A TAF(II)250-TAF(II)150 complex recognizes the initiator. *EMBO Journal* 18, 4835–4845.
- Churchman, L.S., Weissman, J.S., 2011. Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* 469, 368–373.
- Clapier, C.R., Cairns, B.R., 2009. The biology of chromatin remodeling complexes. *Annual Review of Biochemistry* 78, 273–304.
- Corden, J.L., 2013. RNA polymerase II c-terminal domain: Tethering transcription to transcript and template. *Chemical Reviews* 113, 8423–8455.
- Core, L.J., Waterfall, J.J., Lis, J.T., 2008. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 322, 1845–1848.
- Cox, M.M., Phillips, G.N., 2007. Handbook of Proteins: Structure, Function, and Methods. Chichester, West Sussex, England; Hoboken, NJ: John Wiley & Sons.
- Darzacq, X., Yao, J., Larson, D.R., et al., 2009. Imaging transcription in living cells. *Annual Review of Biophysics* 38, 173–196.
- Dierks, P., van Ooyen, A., Cochran, M.D., et al., 1983. Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit beta-globin gene in mouse 3T6 cells. *Cell* 32, 695–706.
- Dierks, P., van Ooyen, A., Mantei, N., Weissmann, C., 1981. DNA sequences preceding the rabbit beta-globin gene are required for formation in mouse L cells of beta-globin RNA with the correct 5' terminus. *Proceedings of the National Academy of Sciences of the United States of America* 78, 1411–1415.
- van Dijk, E.L., Auger, H., Jaszczyzyn, Y., Thermes, C., 2014. Ten years of next-generation sequencing technology. *Trends in Genetics* 30, 418–426.
- Egloff, S., Szczepaniak, S.A., Dienstbier, M., et al., 2010. The integrator complex recognizes a new double mark on the RNA polymerase II carboxyl-terminal domain. *Journal of Biological Chemistry* 285, 20564–20569.
- Egloff, S., Zaborowska, J., Laitem, C., Kiss, T., Murphy, S., 2012. Ser7 phosphorylation of the CTD recruits the RPA2 Ser5 phosphatase to snRNA genes. *Molecular Cell* 45, 111–122.
- Eick, D., Geyer, M., 2013. The RNA polymerase II carboxy-terminal domain (CTD) code. *Chemical Reviews* 113, 8456–8490.
- Fabrega, C., Shen, V., Shuman, S., Lima, C.D., 2003. Structure of an mRNA capping enzyme bound to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Molecular Cell* 11, 1549–1561.
- Furey, T.S., 2012. ChIP-seq and beyond: New and improved methodologies to detect and characterize protein-DNA interactions. *Nature Reviews Genetics* 13, 840–852.
- Gardner, K.E., Allis, C.D., Strahl, B.D., 2011. Operating on chromatin, a colorful language where context matters. *Journal of Molecular Biology* 409, 36–46.
- Glover-Cutter, K., Laroche, S., Erickson, B., et al., 2009. TFIH-associated Cdk7 kinase functions in phosphorylation of c-terminal domain ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. *Molecular and Cellular Biology* 29, 5455–5464.
- Hager, G.L., McNally, J.G., Misteli, T., 2009. Transcription dynamics. *Molecular Cell* 35, 741–753.
- Hanes, S.D., 2014. The Ess1 prolyl isomerase: Traffic cop of the RNA polymerase II transcription cycle. *Biochimica et Biophysica Acta* 1839, 316–333.
- Horikoshi, M., Hai, T., Lin, Y.S., Green, M.R., Roeder, R.G., 1988. Transcription factor ATF interacts with the TATA factor to facilitate establishment of a preinitiation complex. *Cell* 54, 1033–1042.
- Hsieh, F.K., Kulaeva, O.I., Patel, S.S., et al., 2013. Histone chaperone FACT action during transcription through chromatin by RNA polymerase II. *Proceedings of the National Academy of Sciences of the United States of America* 110, 7654–7659.
- Kelly, W., Dahmus, M., Hart, G., 1993. RNA polymerase II is a glycoprotein. modification of the COOH-terminal domain by O-GlcNAc. *Journal of Biological Chemistry* 268, 10416–10424.
- Kouzine, F., Levens, D., Baranello, L., 2014. DNA topology and transcription. *Nucleus* 5, 195–202.
- Kulaeva, O.I., Hsieh, F.K., Chang, H.W., Luse, D.S., Studitsky, V.M., 2013. Mechanism of transcription through a nucleosome by RNA polymerase II. *Biochimica et Biophysica Acta* 1829, 76–83.
- Kwak, H., Fuda, N.J., Core, L.J., Lis, J.T., 2013. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* 339, 950–953.
- Larson, D.R., 2011. What do expression dynamics tell us about the mechanism of transcription? *Current Opinion in Genetics & Development* 21, 591–599.
- Lee, D.H., Gershenson, N., Gupta, M., et al., 2005. Functional characterization of core promoter elements: The downstream core element is recognized by TAF1. *Molecular and Cellular Biology* 25, 9674–9686.
- Lewis, B.A., Kim, T.K., Orkin, S.H., 2000. A downstream element in the human beta-globin promoter: Evidence of extended sequence-specific transcription factor IID contacts. *Proceedings of the National Academy of Sciences of the United States of America* 97, 7172–7177.
- Lewis, B.A., Sims 3RD, R.J., Lane, W.S., Reinberg, D., 2005. Functional characterization of core promoter elements: DPE-specific transcription requires the protein kinase CK2 and the PC4 coactivator. *Molecular Cell* 18, 471–481.
- Lim, C.Y., Santoso, B., Boulay, T., et al., 2004. The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes & Development* 18, 1606–1617.
- Lu, H., Flores, O., Weinmann, R., Reinberg, D., 1991. The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proceedings of the National Academy of Sciences of the United States of America* 88, 10004–10008.
- Maheshri, N., O'Shea, E.K., 2007. Living with noisy genes: How cells function reliably with inherent variability in gene expression. *Annual Review of Biophysics and Biomolecular Structure* 36, 413–434.
- Maldonado, E., Drapkin, R., Reinberg, D., 1996. Purification of human RNA polymerase II and general transcription factors. *Methods in Enzymology* 274, 72–100.

- Martinez-Rucobo, F.W., Cramer, P., 2013. Structural basis of transcription elongation. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1829, 9–19.
- McKnight, S.L., Gavis, E.R., Kingsbury, R., Axel, R., 1981. Analysis of transcriptional regulatory signals of the HSV thymidine kinase gene: Identification of an upstream control region. *Cell* 25, 385–398.
- McKnight, S.L., Kingsbury, R., 1982. Transcriptional control signals of a eukaryotic protein-coding gene. *Science* 217, 316–324.
- Milo, R., Shen-Orr, S., Itzkovitz, S., *et al.*, 2002. Network motifs: Simple building blocks of complex networks. *Science* 298, 824–827.
- Munsky, B., Neuert, G., van Oudenaarden, A., 2012. Using gene expression noise to understand gene regulation. *Science* 336, 183–187.
- Nagy, Z., Tora, L., 2007. Distinct GCN5/PCAF-containing complexes function as co-activators and are involved in transcription factor and global histone acetylation. *Oncogene* 26, 5341–5357.
- Nakatani, Y., Horikoshi, M., Brenner, M., *et al.*, 1990. A downstream initiation element required for efficient TATA box binding and *in vitro* function of TFIID. *Nature* 348, 86–88.
- Nishioka, K., Chuikov, S., Sarma, K., *et al.*, 2002a. Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. *Genes & Development* 16, 479–489.
- Nishioka, K., Rice, J.C., Sarma, K., *et al.*, 2002b. PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. *Molecular Cell* 9, 1201–1213.
- Ong, C.T., Corces, V.G., 2014. CTCF: An architectural protein bridging genome topology and function. *Nature Reviews Genetics* 15, 234–246.
- Orom, U.A., Shiekhattar, R., 2011. Long non-coding RNAs and enhancers. *Current Opinion in Genetics & Development* 21, 194–198.
- Palangat, M., Larson, D.R., 2012. Complexity of RNA polymerase II elongation dynamics. *Biochimica et Biophysica Acta* 1819, 667–672.
- Peterlin, B.M., Price, D.H., 2006. Controlling the elongation phase of transcription with P-TEFb. *Molecular Cell* 23, 297–305.
- Phatnani, H.P., Greenleaf, A.L., 2006. Phosphorylation and functions of the RNA polymerase II CTD. *Genes & Development* 20, 2922–2936.
- Poss, Z.C., Ebmeier, C.C., Taatjes, D.J., 2013. The Mediator complex and transcription regulation. *Critical Reviews in Biochemistry and Molecular Biology* 48, 575–608.
- Proudfoot, N.J., 2011. Ending the message: Poly(A) signals then and now. *Genes & Development* 25, 1770–1782.
- Ptashne, M., 2004. *A Genetic Switch: Phage Lambda Revisited*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Ranuncolo, S.M., Ghosh, S., Hanover, J.A., Hart, G.W., Lewis, B.A., 2012. Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription *in vitro* and *in vivo*. *Journal of Biological Chemistry* 287, 23549–23561.
- Richard, P., Manley, J.L., 2009. Transcription termination by nuclear RNA polymerases. *Genes & Development* 23, 1247–1269.
- Roeder, R.G., Rutter, W.J., 1969. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* 224, 234–237.
- Rothbart, S.B., Strahl, B.D., 2014. Interpreting the language of histone and DNA modifications. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 1839, 627–643.
- Rougvie, A.E., Lis, J.T., 1988. The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged. *Cell* 54, 795–804.
- Rue, P., Garcia-Ojalvo, J., 2013. Modeling gene expression in time and space. *Annual Review of Biophysics* 42, 605–627.
- Ryu, S., Tjian, R., 1999. Purification of transcription cofactor complex CRSP. *Proceedings of the National Academy of Sciences of the United States of America* 96, 7137–7142.
- Schroder, S., Herker, E., Itzen, F., *et al.*, 2013. Acetylation of RNA polymerase II regulates growth-factor-induced gene transcription in mammalian cells. *Molecular Cell* 52, 314–324.
- Schwartz, B.E., Larochelle, S., Suter, B., Lis, J.T., 2003. Cdk7 is required for full activation of *Drosophila* heat shock genes and RNA polymerase II phosphorylation *in vivo*. *Molecular and Cellular Biology* 23, 6876–6886.
- Sims 3rd, R.J., Belotserkovskaya, R., Reinberg, D., 2004. Elongation by RNA polymerase II: The short and long of it. *Genes & Development* 18, 2437–2468.
- Sims 3rd, R.J., Rojas, L.A., Beck, D., *et al.*, 2011. The c-terminal domain of RNA polymerase II is modified by site-specific methylation. *Science* 332, 99–103.
- Smale, S.T., Baltimore, D., 1989. The “initiator” as a transcription control element. *Cell* 57, 103–113.
- Smale, S.T., Kadonaga, J.T., 2003. The RNA polymerase II core promoter. *Annual Review of Biochemistry* 72, 449–479.
- Smallwood, A., Ren, B., 2013. Genome organization and long-range regulation of gene expression by enhancers. *Current Opinion in Cell Biology* 25, 387–394.
- Soutoglou, E., Talianidis, I., 2002. Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation. *Science* 295, 1901–1904.
- Thomas, M.C., Chiang, C.M., 2006. The general transcription machinery and general cofactors. *Critical Reviews in Biochemistry and Molecular Biology* 41, 105–178.
- Tokusumi, Y., Ma, Y., Song, X., Jacobson, R.H., Takada, S., 2007. The new core promoter element XCPE1 (X core promoter element 1) directs activator-, mediator-, and TATA-binding protein-dependent but TFIID-independent RNA polymerase II transcription from TATA-less promoters. *Molecular and Cellular Biology* 27, 1844–1858.
- Tyson, J.J., Novak, B., 2010. Functional motifs in biochemical reaction networks. *Annual Review of Physical Chemistry* 61, 219–240.
- Vannini, A., Cramer, P., 2012. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Molecular Cell* 45, 439–446.
- Voss, T.C., Hager, G.L., 2014. Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nature Reviews Genetics* 15, 69–81.
- Wen, Y., Shatkin, A.J., 1999. Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes & Development* 13, 1774–1779.
- Wilhelm, B.T., Landry, J.-R., 2009. RNA-Seq — quantitative measurement of expression through massively parallel RNA-sequencing. *Methods* 48, 249–257.
- Zhou, Q., Li, T., Price, D.H., 2012. RNA polymerase II elongation control. *Annual Review of Biochemistry* 81, 119–143.

Distant Activation of Transcription by Enhancers

O Kulaeva, Fox Chase Cancer Center, Philadelphia, PA, USA

VM Studitsky, Fox Chase Cancer Center, Philadelphia, PA, USA and Russian Academy of Sciences, Moscow, Russia

© 2016 Elsevier Inc. All rights reserved.

Glossary

Activator A protein that increases gene transcription of a gene or set of genes. Most activators are DNA-binding proteins.

ATP Adenosine triphosphate, coenzyme used as an energy carrier in the cells of all known organisms.

ATP-dependent chromatin remodeler Protein complex driving the dynamic modification of chromatin architecture to allow access of condensed genomic DNA to transcription activators. Its action is driven by ATP hydrolysis.

Coactivator A protein that increases gene expression by binding to an activator (or transcription factor), which contains a DNA-binding domain. The coactivator is unable to bind DNA by itself.

Enhancer (E) A short DNA sequence capable of binding transcription regulation factors that can activate

transcription over variable distances, and independently of its orientation on DNA and position upstream or downstream of the target promoter.

Insulator A genetic boundary element that blocks the interaction between enhancers and promoters. Insulators therefore determine the set of genes an enhancer can influence.

Promoter (P) DNA region where RNA polymerase binds and initiates transcription.

Silencer A DNA sequence capable of binding transcription regulation factors that can work over a large distance and causes repression of transcription.

Super-enhancer An unusually long (~10 kb) enhancer that often activate genes defining cell identity.

Introduction

Transcriptional enhancers (Enhancers) are short (typically averaging about 700 bp in human cells) DNA sequences that can activate transcription from target promoters (Promoters, DNA regions where RNA polymerase binds and initiates transcription) over variable distances (up to more than 1 Mb), independently of enhancer orientation on DNA and its position upstream or downstream of the target promoter (Wasylyk *et al.*, 1984). Enhancers operate both in bacteria and in eukaryotes (Bondarenko *et al.*, 2003); they are predominant regulatory elements in higher eukaryotes while distant gene regulation is much less common in bacteria. However, yeast enhancer-like elements (UASs) work only over a short distance (Guarente, 1988). Over a million of enhancers regulate cell- and tissue-specific expression of human genes (Heinz *et al.*, 2010).

Action of enhancers involves communication between activated enhancer with otherwise transcriptionally silent promoter target, both already containing specifically bound proteins (Figure 1). Enhancer-promoter communication (EPC) could occur by 'hopping' or 'scanning' mechanisms proposed for translocation of *lac* repressor on DNA (Winter *et al.*, 1981). However, one class of bacterial enhancers uses a considerably different (tracking) mechanism for EPC and interaction (see below). Regardless of the precise mechanism of EPC, once the enhancer-bound protein(s) reach the target promoter, they establish direct interaction with the promoter-bound protein(s). This interaction results in local melting of promoter DNA and transcription activation. Once RNA polymerase leaves the promoter, inactive transcription complexes have to be re-established and re-activated there.

Misregulation of enhancers can lead to numerous human diseases. Thus, 93–95% of sites of single-nucleotide

polymorphism (SNPs) associated with various human diseases are localized within the non-coding regions of the genome, including multiple enhancers (Hnisz *et al.*, 2013; Thurman *et al.*, 2012; Maurano *et al.*, 2012), suggesting that enhancer malfunction is involved in development of many human diseases caused by genetic abnormalities.

Recent studies have revealed that a relatively small subgroup of enhancers called super-enhancers plays particularly important roles in cell differentiation and disease (see Lee and Young (2013) for review). Super-enhancers were discovered as regulatory elements guiding expression of genes controlling the pluripotent cell state (Whyte *et al.*, 2013). They consist of extended (~10 kb) clusters of regulatory DNA sequences that are densely occupied by the master regulators, transcription factors and mediator (Whyte *et al.*, 2013). In differentiated cells super-enhancers are present at many genes defining cell identity. The importance of super-enhancers for human health was emphasized by recent discovery of their association with many potent oncogenes expressed in cancer cells (Loven *et al.*, 2013; Hnisz *et al.*, 2013).

There are several excellent recent reviews on various aspects of enhancer action (Kulaeva *et al.*, 2012a; Krivega and Dean, 2012; de Laat and Duboule, 2013; Calo and Wysocka, 2013; Xie and Ren, 2013; Gibcus and Dekker, 2013).

Bacterial Enhancer Action

Bacteria have one four-subunit RNA polymerase core responsible for transcription of all genes and associated with one of the two types (σ^{70} - or σ^{54} -type) of sigma factor, which recognizes the promoter DNA sequence. There are two types of bacterial transcriptional enhancers using tracking and looping mechanisms for EPC, respectively.

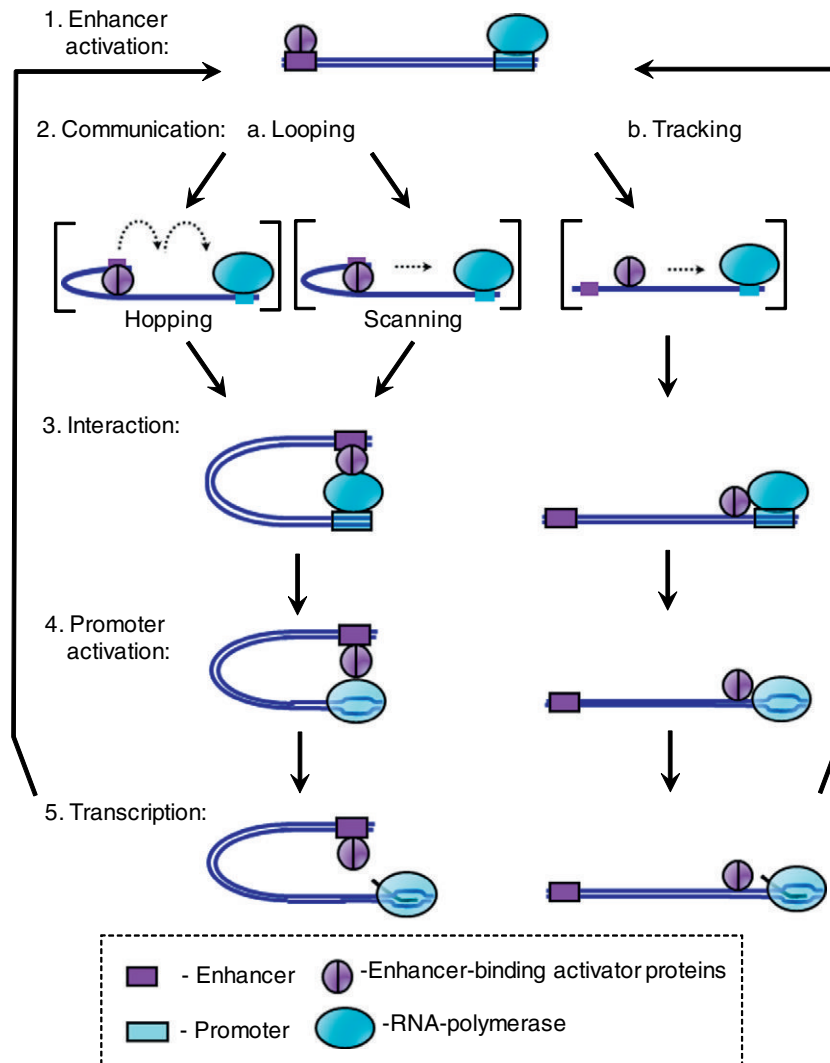


Figure 1 Mechanisms of enhancer action over a distance. (a) Enhancer activation involves either covalent modification of the protein(s) bound to the enhancer or *de novo* protein binding; target promoters are occupied by target proteins before activation. (b) Enhancer-promoter communication occurs by one of the following ways: (a) Looping: Enhancer-bound proteins either directly interact with the target promoter, or search the spacer DNA using a hopping or scanning mechanisms. (b) Alternatively, enhancer-promoter communication occurs by tracking of enhancer-bound protein along DNA. (3) Enhancer-bound or tracking protein(s) reach the target promoter, establishing direct interaction with the promoter-bound protein(s). This interaction results in local DNA melting and promoter activation (4), followed by transcription (5). After the initial round of transcription, inactive transcription complexes have to be re-established at the target promoters to begin the activation cycle.

Distant Promoter Activation Using Tracking Mechanism

Transcription of the late genes of bacteriophage T4 is activated by the 'mobile enhancer' – a sliding clamp of T4 DNA polymerase (gp45) that is loaded on DNA at an internal nick T4-encoded protein complex gp44/62 (Latham *et al.*, 1999; Kolesky *et al.*, 2002). Once the gp45 clamp is loaded, it slides along DNA in 5'→3' direction relative to the DNA strand containing the nick. As a result, an unobstructed DNA region between the enhancer and promoter is required for enhancer action (Tinker *et al.*, 1994). When the clamp reaches the promoter, the closed initiation complex is converted into the open, functionally active complex (Kolesky *et al.*, 2002).

Distant Promoter Activation Accompanied by DNA Looping

In bacteria, the looping mechanism is employed by transcriptional enhancers activating only σ^{54} -dependent promoters (Ray *et al.*, 1990). Bacterial enhancers are involved in regulation of different metabolic pathways, ranging from nitrogen metabolism to aromatic amino acid biosynthesis (Bush and Dixon, 2012). Bacterial σ^{54} -dependent and eukaryotic enhancers share several key properties, including high stability of the closed, inactive initiation complexes, entire dependence of transcription on the presence of an activator, distant transcriptional activation (independent on exact location and orientation of the enhancer) and tight coupling of DNA melting with ATP hydrolysis (Ghosh *et al.*, 2010).

The mechanism of bacterial enhancer action is shown in **Figure 2**. Before activation, the σ^{54} -containing RNAP recognizes the promoter and forms an unusually stable closed initiation complex there (Ninfa *et al.*, 1987). Transcription initiation fully requires activators – bacterial enhancer-binding proteins (bEBPs). In contrast, regulation of σ^{70} -dependent promoters largely depends on RNA polymerase recruitment. bEBPs are members of the AAA+ protein subfamily and function by coupling the energy of ATP hydrolysis to the conversion of the closed complex to active, open initiation complex (Schumacher *et al.*, 2004). Before activation bEBPs are bound to a short (20–40 bp) enhancer DNA sequence, typically forming inactive protein dimers (**Figure 2**). Activation of bEBPs by phosphorylation or ligand binding results in a conformational change and formation of functionally

active protein heptamer (Tucker *et al.*, 2010). Then enhancer-bound bEBPs interact with the promoter-bound initiation complex and stimulate formation of the open, transcriptionally active complex at the promoter in an ATP hydrolysis-dependent way (Wedel and Kustu, 1995). Enhancer-promoter interaction is accompanied by looping of intervening DNA (Rippe *et al.*, 1997; Bose *et al.*, 2008). During short-distance EPC DNA looping is often facilitated by DNA bending induced by IHF protein (Huo *et al.*, 2009), while long-distance (over more than 1-kb distances) enhancer action requires DNA supercoiling (Liu *et al.*, 2001). DNA supercoiling induces formation of DNA ‘branches’ and thus increases the probability of juxtaposition between linearly separated enhancer and promoter (Huang *et al.*, 2001; Polikanov *et al.*, 2007). Once the bEBP is engaged with the closed initiation

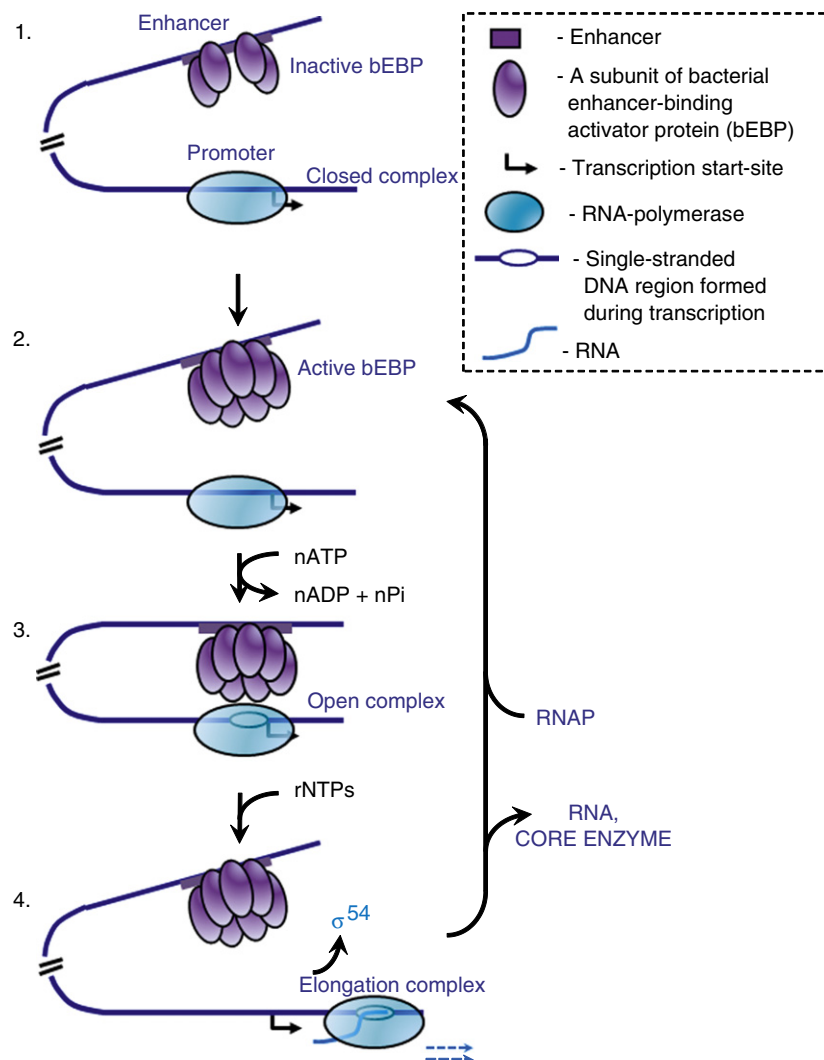


Figure 2 Mechanism of action of bacterial σ^{54} -dependent enhancer. (1) Before transcription is induced, the RNA polymerase forms a closed complex at the promoter but cannot initiate transcription. Bacterial enhancer-binding protein (bEBP) is bound to the enhancer in the form of two functionally inactive dimers and does not communicate with the promoter. (2) After induction, phosphorylated bEBP assumes the heptameric oligomerization state, and (3) interacts with the closed complex causing looping of the intervening DNA and ATP hydrolysis-dependent formation of the open, functionally active initiation complex at the promoter. After formation of the active initiation complex the DNA loop is opened. (4) As the RNA polymerase leaves the promoter, the σ^{54} subunit dissociates into solution, and the closed initiation complex has to be re-assembled again.

complex through interaction with σ^{54} subunit within the ignition complex (Bose *et al.*, 2008), ATPase activity of the activator is required to induce a change in conformation of the closed complex, facilitating formation of the productive open initiation complex (Burrows *et al.*, 2009; Sharma *et al.*, 2014). As RNAP escapes from the promoter, the enhancer-promoter looping is disrupted, the σ^{54} subunit dissociates into solution and the next round of transcription requires assembly of new initiation complex at the promoter (Bondarenko *et al.*, 2002).

Enhancer Action in Eukaryotes

Gene expression in higher eukaryotes is regulated through combined action of many distant regulatory elements (enhancers, silencers (Vokes *et al.*, 2008) and insulators (Ghirlando *et al.*, 2012)) acting on core promoters (Lenhard *et al.*, 2012). Enhancers play a dominant role in regulation of gene expression and typically carry clusters of different transcription factor (TF)-binding sites. Indeed, genome-wide studies revealed that during gene activation the majority (typically >90%) of TF-binding events occur at enhancers, typically localized hundreds kb away from the regulated promoters (Young, 2011; Chen *et al.*, 2008). Histone modification patterns at enhancers are more cell type specific than promoter-associated patterns (Visel *et al.*, 2009; Ernst *et al.*, 2011), suggesting that enhancers dictate patterns of tissue-specific gene expression. There are about a million of human enhancers (Consortium *et al.*, 2012; Thurman *et al.*, 2012), with thousands of them active in a particular cell type (Visel *et al.*, 2009; Ernst *et al.*, 2011). Most promoters associate with a single enhancer, but about 25% associate with two or more enhancers (Chepelev *et al.*, 2012). During development or after induction of gene expression, the occupancy of enhancers by corresponding transcription factors is increased, inducing a particular program of gene expression (Spitz and Furlong, 2012).

The events that typically occur during enhancer-dependent promoter activation during development are illustrated in Figure 3. In early development, a subset of promoters, to-be-transcribed later in development, can be 'pre-marked' with histone modifications H3K4me3 (histone H3 tri-methylated at lysine 4 (Lindeman *et al.*, 2011; Wu *et al.*, 2011)). Later during development, pioneer factors can bind to chromatin-covered DNA that is inaccessible to other factors and recruit chromatin remodeling complexes and histone variants that destabilize nucleosomes on the enhancer and make the enhancer available for binding by other factors (Zaret and Carroll, 2011).

Later during development both enhancers and promoters destined for activation become more extensively pre-marked (Figure 3). Thus, in human and mouse embryonic stem cells, the enhancers of genes expressed early during differentiation share many properties with active enhancers (open chromatin structure, the presence of H3K4me1, TFs and coactivators (Rada-Iglesias *et al.*, 2011; Zentner *et al.*, 2011; Creyghton *et al.*, 2010)). However, in contrast to active enhancers, these 'poised' enhancers are functionally inactive (Rada-Iglesias *et al.*, 2011), contain H3K27me3 and do not contain H3K27ac (histone H3 acetylated at lysine 27 (Rada-Iglesias *et al.*, 2011; Zentner *et al.*, 2011; Creyghton *et al.*, 2010)). Corresponding

poised core promoters are inactive and contain 'bivalent' marks (H3K4me3 and H3K27me3), histone variants H3.3 and H2A.Z, and elongating (paused) RNA polymerase II (Pol II) complexes making short aborted transcripts (Marks *et al.*, 2012; Kellner *et al.*, 2012; Jin *et al.*, 2009; Golob *et al.*, 2011).

As cell differentiation unravels, different subsets of poised enhancers in different cell types are activated by cell-specific TFs and become associated with Pol II, short enhancer RNAs (eRNAs) and H3K27ac (Rada-Iglesias *et al.*, 2011; Bonn *et al.*, 2012; Figure 3). Activation of enhancers by TFs involves recruitment of coactivators, which lack sequence-specific DNA-binding competency and function as histone modifiers (e.g., acetyltransferases p300, CBP and ATAC complex), ATP-dependent chromatin remodelers (e.g., CHD7 and BAF complexes), and mediators of interaction with transcriptional machinery at promoters (e.g., Mediator complex) (D'Alessio *et al.*, 2009; Weake and Workman, 2010; Borggreffe and Yue, 2011). Active enhancers are present in nucleosome-free regions with a high rate of histone replacement and enriched for the histone variants H2A.Z and H3.3 (Mito *et al.*, 2007; Jin *et al.*, 2009).

To be activated, a gene has to be relocated from the chromosome territory (Fraser and Bickmore, 2007) into a different nuclear compartment, and chromatin structure decondensed to the level of the 30-nm chromatin fiber (Morey *et al.*, 2008). Both 'poised' (potentially active in transcription) and active chromatin domains exist in form of the 30-nm fibers (Gilbert *et al.*, 2004; Naughton *et al.*, 2010).

Similar to the NtrC-dependent enhancers, all well-studied eukaryotic enhancers work by looping: in the active state the enhancer and target promoter are in physical proximity (see Kulaeva *et al.* (2012a), Krivega and Dean (2012), Ong and Corces (2011) for review). Enhancer-promoter interaction can occur *in cis* and *in trans*; for a given E, the majority of looping interactions occur over a distance of hundreds kb (Buecker and Wysocka, 2012) and only rarely with the nearest gene (Sanyal *et al.*, 2012). In metazoan genomes, there are extensive networks of long-range interactions (see Gibcus and Dekker (2013) for review).

The 30-nm chromatin fiber is a dynamic structure that maintains efficient, distance-independent EPC; bending of the chromatin fiber required for communication is facilitated by internucleosomal interactions involving the histone tails (Kulaeva *et al.*, 2012b). Additionally, transcription of enhancers and the spacer DNA is required for enhancer action (Zhu *et al.*, 2007; Orom *et al.*, 2010; Kim *et al.*, 2010; Wang *et al.*, 2011). Two models have been proposed to explain this requirement. According to the first model, the process of transcription itself is essential, since a terminator inserted between the enhancer and promoter prevents enhancer action (Kim *et al.*, 2007; Zhu *et al.*, 2007; Ling *et al.*, 2004). Alternatively, RNA could stabilize the enhancer-promoter chromatin loop, either during or after communication (Orom *et al.*, 2010; Kim *et al.*, 2010; Wang *et al.*, 2011; Orom and Shiekhata, 2013). Numerous additional proteins and TFs are involved in the formation of chromatin loops that bring enhancers and promoters together (Marsman and Horsfield, 2012). Various enhancer-binding transcription factors (EKLF, GATA-1, and OCA-B) and cofactors (Mediator, BRG1, NLI/Ldb1, and cohesin) have been implicated in loop formation

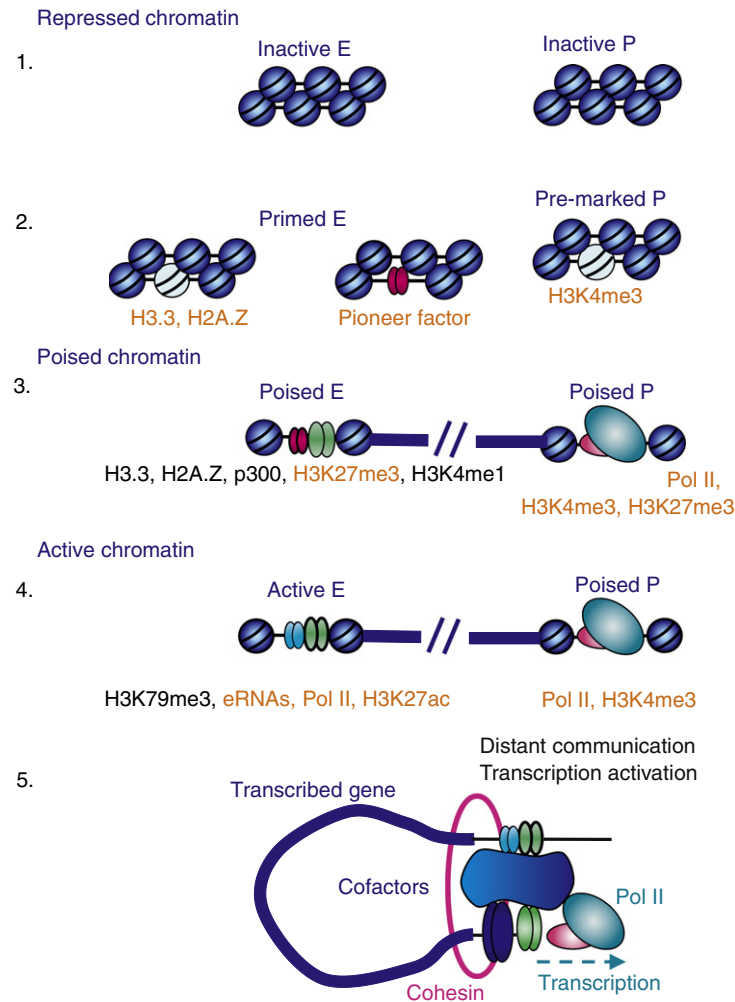


Figure 3 Proposed mechanisms of activation of eukaryotic enhancers in development. (1) Early during development many promoters that become active later in development are pre-marked by the histone modification. (2) Later, during cell differentiation inactive enhancers are primed with pioneering transcription factors and marked by histone variants H2A.Z and H3.3. (3) Multiple inactive promoters and enhancers are poised and specifically marked in human and mouse stem cells. (4) During enhancer activation, a set of specific transcription factors binds to the enhancer, inducing enhancer-promoter communication, chromatin looping and promoter activation (5). Histone modifications and other marks that are specific for a particular development stage are shown in orange.

on different genes (Marsman and Horsfield, 2012). It should be noted that almost every factor binding at the interacting enhancer or promoter (direct or indirect, through DNA-interacting proteins) could contribute to the looping efficiency. Possible underlying mechanisms include cooperative binding of protein factors at these elements (with multiple direct and indirect mechanisms involved (Spitz and Furlong, 2012)) and stabilization of the looping *per se*, either through DNA-protein interactions (CTCF), protein-protein interactions (Mediator) or topologically enclosing interacting chromatin fibers (cohesin) (Figure 3; Marsman and Horsfield, 2012; Felsenfeld and Dekker, 2012).

After formation of the enhancer-promoter loop, activation of the promoter can occur by one of the following two mechanisms. The first mechanism involves recruiting of general transcription factor TFIID to the promoter and its activation through interaction with TFIIA, resulting in a change of TFIID conformation leading to formation of functionally

active initiation complex (Burley and Roeder, 1996; Papai *et al.*, 2010). Alternatively, enhancers can activate Pol II paused on poised promoters and protect these from inactivation by competing chromatin assembly (Min *et al.*, 2011; Golob *et al.*, 2011; Gilchrist *et al.*, 2010; Cheng *et al.*, 2012). In this case, Gdown1 protein binds to transcribing complex immediately after transcription initiation and stabilizes poised polymerase complexes stalled after transcribing 20–40 bp (Cheng *et al.*, 2012; Jishage *et al.*, 2012; Davis *et al.*, 2014). After activation by enhancer, the Pol II elongation complexes are phosphorylated by protein kinase P-TEFb and released (Czudnochowski *et al.*, 2012).

Conclusions

In summary, recent studies have identified fundamental principles of enhancer action that are valid both for bacteria

and eukaryotes. In both cases, the enhancer and promoter to-be-activated are pre-marked by inactive, stable protein-DNA complexes. Activation is induced either by modification of enhancer-bound proteins or by binding of new protein complexes at the enhancer. Then activated enhancer searches for the target using a facilitated looping mechanism (e.g., tracking) and establishes a direct interaction through protein-protein interactions with the promoter, accompanied by looping of intervening chromatin. Promoter activation typically involves a conformational change in promoter-bound protein complexes or their covalent modification.

Since enhancers are primary drivers of gene regulation in eukaryotes, they are involved in development of numerous human diseases, ranging from cardiovascular abnormalities to cancers (see Lee and Young (2013) for review).

Acknowledgment

This work was supported by NSF MCB-1050470 grant to V.M.S.

See also: Dynamic Integration: Dynamics: Transcription Factor Networks. Nucleic Acid Synthesis/Breakdown: RNA Synthesis/Function: Comparison of Bacterial and Eukaryotic Replisome Components; miRNAs/Small Noncoding RNAs; Small RNAs/Cancer. Nucleic Acid Synthesis/Breakdown: Transcription: Eukaryotic Transcriptional Regulation; Prokaryotic Transcription. Organelles: Structure and Function: Nuclear Organization (Nuclear Structure and Dynamics)

References

- Bondarenko, V., Liu, Y., Ninfa, A., Studitsky, V.M., 2002. Action of prokaryotic enhancer over a distance does not require continued presence of promoter-bound sigma54 subunit. *Nucleic Acids Research* 30, 636–642.
- Bondarenko, V.A., Liu, Y.V., Jiang, Y.I., Studitsky, V.M., 2003. Communication over a large distance: Enhancers and insulators. *Biochemistry and Cell Biology* 81, 241–251.
- Bonn, S., Zinnen, R.P., Girardot, C., *et al.*, 2012. Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nature Genetics* 44, 148–156.
- Borggreffe, T., Yue, X., 2011. Interactions between subunits of the Mediator complex with gene-specific transcription factors. *Seminars in Cell & Developmental Biology* 22, 759–768.
- Bose, D., Pape, T., Burrows, P.C., *et al.*, 2008. Organization of an activator-bound RNA polymerase holoenzyme. *Molecular Cell* 32, 337–346.
- Buecker, C., Wysocka, J., 2012. Enhancers as information integration hubs in development: Lessons from genomics. *Trends in Genetics* 28, 276–284.
- Burley, S.K., Roeder, R.G., 1996. Biochemistry and structural biology of transcription factor IID (TFIID). *Annual Review of Biochemistry* 65, 769–799.
- Burrows, P.C., Joly, N., Cannon, W.V., *et al.*, 2009. Coupling sigma factor conformation to RNA polymerase reorganization for DNA melting. *Journal of Molecular Biology* 387, 306–319.
- Bush, M., Dixon, R., 2012. The role of bacterial enhancer binding proteins as specialized activators of sigma54-dependent transcription. *Microbiology and Molecular Biology Reviews* 76, 497–529.
- Calo, E., Wysocka, J., 2013. Modification of enhancer chromatin: What, how, and why? *Molecular Cell* 49, 825–837.
- Chen, X., Xu, H., Yuan, P., *et al.*, 2008. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133, 1106–1117.
- Cheng, B., Li, T., Rahl, P.B., *et al.*, 2012. Functional association of Gdown1 with RNA polymerase II poised on human genes. *Molecular Cell* 45, 38–50.
- Chepelev, I., Wei, G., Wangsa, D., *et al.*, 2012. Characterization of genome-wide enhancer-promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization. *Cell Research* 22, 490–503.
- Consortium, E.P., Bernstein, B.E., Birney, E., *et al.*, 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.
- Creyghton, M.P., Cheng, A.W., Welstead, G.G., *et al.*, 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences of the United States of America* 107, 21931–21936.
- Czudnochowski, N., Bosken, C.A., Geyer, M., 2012. Serine-7 but not serine-5 phosphorylation primes RNA polymerase II CTD for P-TEFb recognition. *Nature Communications* 3, 842.
- D'Alessio, J.A., Wright, K.J., Tjian, R., 2009. Shifting players and paradigms in cell-specific transcription. *Molecular Cell* 36, 924–931.
- Davis, M.A., Guo, J., Price, D.H., Luse, D.S., 2014. Functional interactions of the RNA polymerase II-interacting proteins Gdown1 and TFIIF. *Journal of Biological Chemistry* 289, 11143–11152.
- Ernst, J., Kheradpour, P., Mikkelsen, T.S., *et al.*, 2011. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473, 43–49.
- Felsenfeld, G., Dekker, J., 2012. Genome architecture and expression. *Current Opinion in Genetics & Development* 22, 59–61.
- Fraser, P., Bickmore, W., 2007. Nuclear organization of the genome and the potential for gene regulation. *Nature* 447, 413–417.
- Ghirlando, R., Giles, K., Gowher, H., *et al.*, 2012. Chromatin domains, insulators, and the regulation of gene expression. *Biochimica et Biophysica Acta* 1819, 644–651.
- Ghosh, T., Bose, D., Zhang, X., 2010. Mechanisms for activating bacterial RNA polymerase. *FEMS Microbiology Reviews* 34, 611–627.
- Gibcus, J.H., Dekker, J., 2013. The hierarchy of the 3D genome. *Molecular Cell* 49, 773–782.
- Gilbert, N., Boyle, S., Fiegler, H., *et al.*, 2004. Chromatin architecture of the human genome: Gene-rich domains are enriched in open chromatin fibers. *Cell* 118, 555–566.
- Gilchrist, D.A., Dos Santos, G., Fargo, D.C., *et al.*, 2010. Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* 143, 540–551.
- Golob, J.L., Kumar, R.M., Guenther, M.G., *et al.*, 2011. Evidence that gene activation and silencing during stem cell differentiation requires a transcriptionally paused intermediate state. *PLoS One* 6, e22416.
- Guarente, L., 1988. UASs and enhancers: Common mechanism of transcriptional activation in yeast and mammals. *Cell* 52, 303–305.
- Heinz, S., Benner, C., Spann, N., *et al.*, 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular Cell* 38, 576–589.
- Hnisz, D., Abraham, B.J., Lee, T.I., *et al.*, 2013. Super-enhancers in the control of cell identity and disease. *Cell* 155, 934–947.
- Huang, J., Schlick, T., Vologodskii, A., 2001. Dynamics of site juxtaposition in supercoiled DNA. *Proceedings of the National Academy of Sciences of the United States of America* 98, 968–973.
- Huo, Y.X., Zhang, Y.T., Xiao, Y., *et al.*, 2009. IHF-binding sites inhibit DNA loop formation and transcription initiation. *Nucleic Acids Research* 37, 3878–3886.
- Jin, C., Zang, C., Wei, G., *et al.*, 2009. H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nature Genetics* 41, 941–945.
- Jishage, M., Malik, S., Wagner, U., *et al.*, 2012. Transcriptional regulation by Pol II (G) involving mediator and competitive interactions of Gdown1 and TFIIF with Pol II. *Molecular Cell* 45, 51–63.
- Kellner, W.A., Ramos, E., Van Bortle, K., *et al.*, 2012. Genome-wide phosphoacetylation of histone H3 at Drosophila enhancers and promoters. *Genome Research* 22, 1081–1088.
- Kim, A., Zhao, H., Ifrim, I., Dean, A., 2007. Beta-globin intergenic transcription and histone acetylation dependent on an enhancer. *Molecular and Cellular Biology* 27, 2980–2986.
- Kim, T.K., Hemberg, M., Gray, J.M., *et al.*, 2010. Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182–187.
- Kolesky, S.E., Ouhammouch, M., Heiduschek, E.P., 2002. The mechanism of transcriptional activation by the topologically DNA-linked sliding clamp of bacteriophage T4. *Journal of Molecular Biology* 321, 767–784.
- Krivega, I., Dean, A., 2012. Enhancer and promoter interactions-long distance calls. *Current Opinion in Genetics & Development* 22, 79–85.

- Kulaeva, O.I., Nizovtseva, E.V., Polikanov, Y.S., *et al.*, 2012a. Distant activation of transcription: Mechanisms of enhancer action. *Molecular and Cellular Biology* 32, 4892–4897.
- Kulaeva, O.I., Zheng, G., Polikanov, Y.S., *et al.*, 2012b. Internucleosomal interactions mediated by histone tails allow distant communication in chromatin. *Journal of Biological Chemistry* 287, 20248–20257.
- de Laat, W., Duboule, D., 2013. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* 502, 499–506.
- Latham, G.J., Dong, F., Pietroni, P., *et al.*, 1999. Opening of a monomer-monomer interface of the trimeric bacteriophage T4-coded GP45 sliding clamp is required for clamp loading onto DNA. *Proceedings of the National Academy of Sciences of the United States of America* 96, 12448–12453.
- Lee, T.I., Young, R.A., 2013. Transcriptional regulation and its misregulation in disease. *Cell* 152, 1237–1251.
- Lenhard, B., Sandelin, A., Carninci, P., 2012. Metazoan promoters: Emerging characteristics and insights into transcriptional regulation. *Nature Reviews Genetics* 13, 233–245.
- Lindeman, L.C., Andersen, I.S., Reiner, A.H., *et al.*, 2011. Prepatterning of developmental gene expression by modified histones before zygotic genome activation. *Developmental Cell* 21, 993–1004.
- Ling, J., Ainol, L., Zhang, L., *et al.*, 2004. HS2 enhancer function is blocked by a transcriptional terminator inserted between the enhancer and the promoter. *Journal of Biological Chemistry* 279, 51704–51713.
- Liu, Y., Bondarenko, V., Ninfa, A., Studitsky, V.M., 2001. DNA supercoiling allows enhancer action over a large distance. *Proceedings of the National Academy of Sciences of the United States of America* 98, 14883–14888.
- Loven, J., Hoke, H.A., Lin, C.Y., *et al.*, 2013. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* 153, 320–334.
- Marks, H., Kalkan, T., Menafra, R., *et al.*, 2012. The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* 149, 590–604.
- Marsman, J., Horsfield, J.A., 2012. Long distance relationships: Enhancer-promoter communication and dynamic gene transcription. *Biochimica et Biophysica Acta* 1819, 1217–1227.
- Maurano, M.T., Humbert, R., Rynes, E., *et al.*, 2012. Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337, 1190–1195.
- Min, I.M., Waterfall, J.J., Core, L.J., *et al.*, 2011. Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes & Development* 25, 742–754.
- Mito, Y., Henikoff, J.G., Henikoff, S., 2007. Histone replacement marks the boundaries of cis-regulatory domains. *Science* 315, 1408–1411.
- Morey, C., Da Silva, N.R., Kmita, M., *et al.*, 2008. Ectopic nuclear reorganisation driven by a *Hoxb1* transgene transposed into *Hoxd*. *Journal of Cell Science* 121, 571–577.
- Naughton, C., Sproul, D., Hamilton, C., Gilbert, N., 2010. Analysis of active and inactive X chromosome architecture reveals the independent organization of 30 nm and large-scale chromatin structures. *Molecular Cell* 40, 397–409.
- Ninfa, A.J., Reitzer, L.J., Magasanik, B., 1987. Initiation of transcription at the bacterial *glnAp2* promoter by purified *E. coli* components is facilitated by enhancers. *Cell* 50, 1039–1046.
- Ong, C.T., Corces, V.G., 2011. Enhancer function: New insights into the regulation of tissue-specific gene expression. *Nature Reviews Genetics* 12, 283–293.
- Orom, U.A., Derrien, T., Beringer, M., *et al.*, 2010. Long noncoding RNAs with enhancer-like function in human cells. *Cell* 143, 46–58.
- Orom, U.A., Shiekhattar, R., 2013. Long noncoding RNAs usher in a new era in the biology of enhancers. *Cell* 154, 1190–1193.
- Papai, G., Tripathi, M.K., Ruhlmann, C., *et al.*, 2010. TFIIA and the transactivator Rap1 cooperate to commit TFIID for transcription initiation. *Nature* 465, 956–960.
- Polikanov, Y.S., Bondarenko, V.A., Tchernenko, V., *et al.*, 2007. Probability of the site juxtaposition determines the rate of protein-mediated DNA looping. *Biophysical Journal* 93, 2726–2731.
- Rada-Iglesias, A., Bajpai, R., Swigut, T., *et al.*, 2011. A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470, 279–283.
- Ray, L., Claverie-Martin, F., Weglenski, P., Magasanik, B., 1990. Role of the promoter in activation of transcription by nitrogen regulator I phosphate in *Escherichia coli*. *Journal of Bacteriology* 172, 818–823.
- Rippe, K., Guthold, M., von Hippel, P.H., Bustamante, C., 1997. Transcriptional activation via DNA-looping: Visualization of intermediates in the activation pathway of *E. coli* RNA polymerase $\times$ sigma 54 holoenzyme by scanning force microscopy. *Journal of Molecular Biology* 270, 125–138.
- Sanyal, A., Lajoie, B.R., Jain, G., Dekker, J., 2012. The long-range interaction landscape of gene promoters. *Nature* 489, 109–113.
- Schumacher, J., Zhang, X., Jones, S., *et al.*, 2004. ATP-dependent transcriptional activation by bacterial PspF AAA + protein. *Journal of Molecular Biology* 338, 863–875.
- Sharma, A., Leach, R.N., Gell, C., *et al.*, 2014. Domain movements of the enhancer-dependent sigma factor drive DNA delivery into the RNA polymerase active site: Insights from single molecule studies. *Nucleic Acids Research* 42, 5177–5190.
- Spitz, F., Furlong, E.E., 2012. Transcription factors: From enhancer binding to developmental control. *Nature Reviews Genetics* 13, 613–626.
- Thurman, R.E., Rynes, E., Humbert, R., *et al.*, 2012. The accessible chromatin landscape of the human genome. *Nature* 489, 75–82.
- Tinker, R.L., Williams, K.P., Kassavetis, G.A., Geiduschek, E.P., 1994. Transcriptional activation by a DNA-tracking protein: Structural consequences of enhancement at the T4 late promoter. *Cell* 77, 225–237.
- Tucker, N.P., Ghosh, T., Bush, M., *et al.*, 2010. Essential roles of three enhancer sites in sigma54-dependent transcription by the nitric oxide sensing regulatory protein NorR. *Nucleic Acids Research* 38, 1182–1194.
- Visel, A., Blow, M.J., Li, Z., *et al.*, 2009. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457, 854–858.
- Vokes, S.A., Ji, H., Wong, W.H., McMahon, A.P., 2008. A genome-scale analysis of the cis-regulatory circuitry underlying sonic hedgehog-mediated patterning of the mammalian limb. *Genes & Development* 22, 2651–2663.
- Wang, K.C., Yang, Y.W., Liu, B., *et al.*, 2011. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature* 472, 120–124.
- Wasylyk, B., Wasylyk, C., Chambon, P., 1984. Short and long range activation by the SV40 enhancer. *Nucleic Acids Research* 12, 5589–5608.
- Weake, V.M., Workman, J.L., 2010. Inducible gene expression: Diverse regulatory mechanisms. *Nature Reviews Genetics* 11, 426–437.
- Wedel, A., Kustu, S., 1995. The bacterial enhancer-binding protein NTRC is a molecular machine: ATP hydrolysis is coupled to transcriptional activation. *Genes & Development* 9, 2042–2052.
- Whyte, W.A., Orlando, D.A., Hnisz, D., *et al.*, 2013. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153, 307–319.
- Winter, R.B., Berg, O.G., von Hippel, P.H., 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The *Escherichia coli lac* repressor–operator interaction: Kinetic measurements and conclusions. *Biochemistry* 20, 6961–6977.
- Wu, S.F., Zhang, H., Cairns, B.R., 2011. Genes for embryo development are packaged in blocks of multivalent chromatin in zebrafish sperm. *Genome Research* 21, 578–589.
- Xie, W., Ren, B., 2013. Developmental biology. Enhancing pluripotency and lineage specification. *Science* 341, 245–247.
- Young, R.A., 2011. Control of the embryonic stem cell state. *Cell* 144, 940–954.
- Zaret, K.S., Carroll, J.S., 2011. Pioneer transcription factors: Establishing competence for gene expression. *Genes & Development* 25, 2227–2241.
- Zentner, G.E., Tesar, P.J., Scacheri, P.C., 2011. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Research* 21, 1273–1283.
- Zhu, X., Ling, J., Zhang, L., *et al.*, 2007. A facilitated tracking and transcription mechanism of long-range enhancer function. *Nucleic Acids Research* 35, 5532–5544.

Relevant Websites

<http://www.ndsu.edu/pubweb/~mcclean/plsc431/geneexpress/eukaryex4.htm>
Control of Gene expression in Eukaryotes.

<http://www.gene-regulation.com/pub/databases.html>
Gene Regulation.

<http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/P/Promoter.html>
Gene Regulation in Eukaryotes.

<http://ghr.nlm.nih.gov/>
Genetics Home Reference.

<http://www.nature.com/scitable/topic/gene-expression-and-regulation-15>
Nature.

<http://zlab.umassmed.edu/zlab/gene.shtml>
UMass Medical School.

The Spliceosome and Pre-mRNA Splicing

TJ Carrocci and AA Hoskins, University of Wisconsin–Madison, Madison, WI, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Alternative Splicing Different ways of splicing a given pre-mRNA to produce multiple distinct mRNAs.

DExD/H box A family of RNA-dependent ATPases that contain a signature DExD or DExH amino acid motif.

Exon Regions of pre-mRNA that are ligated together by the spliceosome.

Intron The regions of a pre-mRNA removed by the spliceosome.

Lariat An RNA containing a 2', 5' phosphodiester linkage that results from 5' exon cleavage during the first chemical step of splicing.

Metazoa Multicellular organisms that are members of the Animalia (Animal) kingdom.

Pre-mRNA Precursor mRNA molecules that have not been spliced and contain introns.

snRNA Small nuclear RNAs. snRNAs assemble with proteins to form the spliceosomal snRNPs.

snRNP Small nuclear ribonucleoproteins. Assemblies of snRNAs and proteins that can form spliceosomes.

Introduction

In the mid-twentieth century, eukaryotic biologists were presented with a dilemma. It had been well-established in prokaryotes that messenger RNAs (mRNAs) were direct copies of genes found within DNA. It was widely predicted that the same would hold true for eukaryotes; however, it was clear that RNA metabolism was occurring differently in organisms with a nucleus. Eukaryotic cells were found to transcribe a huge amount of nuclear RNA (called heterogeneous nuclear RNA or hnRNA) that never appeared to be exported to the cytoplasm (Sharp 1993; Soeiro *et al.*, 1966). Furthermore, both hnRNAs and mRNAs were found to contain 5' cap structures and 3' polyadenylation tracts. If hnRNAs were the precursor molecules (pre-mRNAs) to cytosolic mRNAs, how could the larger hnRNA be shortened in the middle but keep both the 5' cap and 3' polyadenylation signals at the ends intact in the mRNA? To answer this question, the laboratories of Richard Roberts and Phillip Sharp independently carried out an elegant experiment: adenovirus RNA was hybridized to its coding DNA and directly visualized by electron microscopy (Berget *et al.*, 1977; Chow *et al.*, 1977). The RNA/DNA hybrid structures revealed loops – regions of DNA not present in the mRNA. The answer was then apparent: eukaryotic mRNAs are made by cutting and ligating together different regions of a pre-mRNA. This process is called RNA splicing, and Roberts and Sharp were awarded the Nobel Prize in Medicine in 1993 for their discovery.

In the years since the discovery of RNA splicing, much has been learned about how splicing is catalyzed and how cells integrate splicing with gene regulation. The intragenic regions removed from pre-mRNAs are called introns and the retained segments are called exons. Introns are ubiquitous in eukaryotic genomes, with most human transcripts having an average of 8 introns, and ultimately several hundred thousand introns are present in the human genome (Scherer, 2008). Introns must be removed precisely: a splicing error of just one nucleotide will disrupt the correct reading frame of the mRNA. The

splicing reaction is carried out by large cellular machine made up of both RNA and proteins, dubbed the spliceosome (Brody and Abelson, 1985). The spliceosome minimally performs two functions during splicing: it defines the boundaries of the intron and then catalyzes its removal (Figure 1).

The Chemistry of Splicing and the Architecture of the Intron

The chemistry of splicing is relatively simple, comprising only two consecutive S_N2 -type transesterification reactions (Moore and Sharp, 1993). During the first catalytic step, the phosphodiester bond between the 5' exon and the intron is cleaved by nucleophilic attack from the 2' hydroxyl of the branchpoint adenosine (Figure 1(b)). The first step of catalysis results in the formation of a 2', 5' branched RNA (known as a lariat intermediate) and a free 5' exon. Rearrangements within the active site of the spliceosome then occur that permit the free 3' hydroxyl of the 5' exon to attack the intron-3' exon junction and result in the simultaneous ligation of the exons and excision of the lariat intron.

Human introns vary greatly in size from less than a hundred nucleotides to several hundred thousand (Scherer, 2008). Despite the range of intron sizes, most exons are much shorter, ranging just a few hundred nucleotides. Because of this disparity in length, selecting the proper position of the pre-mRNA for excision represents a formidable task. This problem is further complicated by the relative lack of sequence conservation between introns. Nonetheless, the spliceosome manages to recognize introns through direct interaction with consensus sequences known as the 5' splice site, the branchsite sequence, and the 3' splice site. The 5' splice site is found at the junction between the 5' exon and the intron. Downstream of the 5' splice site is the branchsite sequence containing the adenosine that serves as the branching point in the lariat intron product. The final components of the intron necessary for splicing are the polypyrimidine tract and the 3' splice site,

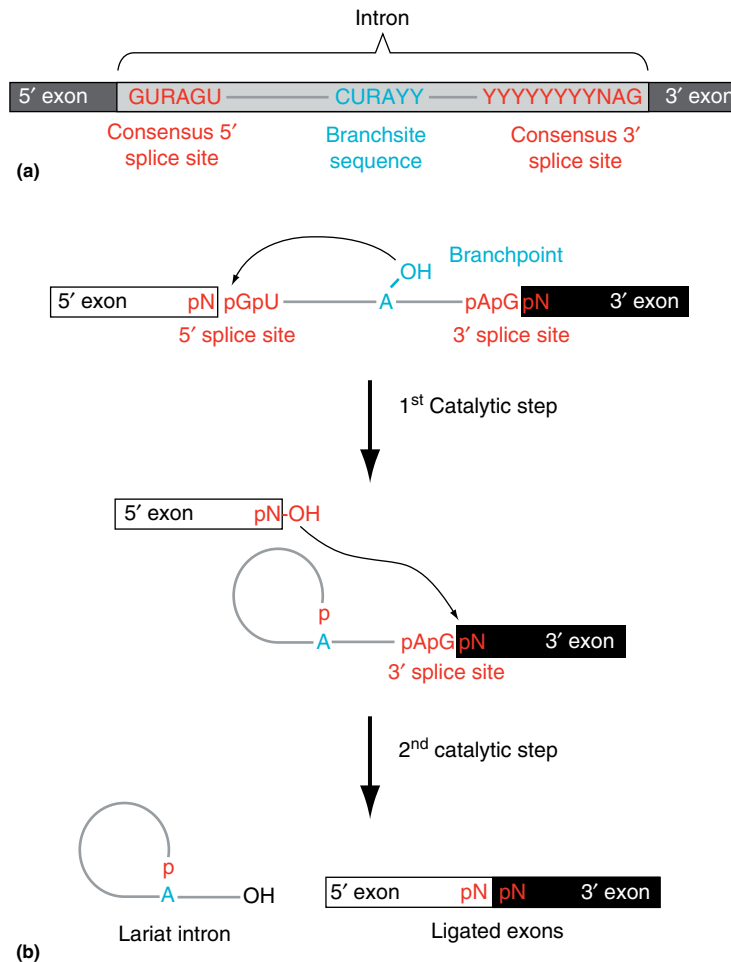


Figure 1 (a) The architecture of the intron. The boundaries of the intron are known as the 5' and 3' splice sites (shown in red). In addition to the splice sites, the pre-mRNA contains a branchsite region (blue) that contains the reactive adenosine (bold) necessary for catalysis. Polypyrimidine tracts are often found between the branchsite sequence and the 3' splice site. The consensus sequences as shown with *R* being any purine, *Y* being any pyrimidine, and *N* being any base. (b) The chemical steps of splicing. During the first step, the 2' hydroxyl of the branchsite adenosine (blue), attacks the 5' splice site to liberate the 5' exon and generate a lariat intermediate. In the second step, the 3' hydroxyl of the free 5' exon attacks the 3' splice site to generate the spliced mRNA and lariat intron products.

with the latter marking the junction between the intron and 3' exon. The consensus splice site and branchsite sequences in eukaryotes such as the yeast *Saccharomyces cerevisiae* are well conserved whereas these sequences are more variable in metazoans. The human consensus sequences are shown in Figure 1(a).

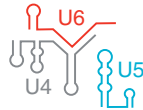
Despite the existence of consensus sequences, splice site recognition is complicated by the presence of both non-consensus splice sites that must be used by the spliceosome and pseudo-splice sites that must be avoided. The spliceosome can accurately discriminate *bona fide* splice sites from false signals and does so with input from other factors in nuclear RNA processing. Amazingly, the human spliceosome is proposed to properly recognize >2500 different 5' splice site sequences (Roca and Krainer, 2009). Identification of the correct splice sites is accomplished through multiple types of noncovalent interactions occurring during spliceosome assembly and activation. The spliceosome is an incredibly

sensitive and flexible enzyme that can discriminate between subtle differences in a large number of RNA sequences.

The Spliceosome is Composed of snRNAs and Proteins

Along with the discovery of introns and splicing by Roberts and Sharp, a second major breakthrough in splicing occurred when Joan Steitz proposed that small nuclear ribonucleoproteins (snRNPs, pronounced 'snurps') were the complexes that carried out the splicing reaction (Lerner *et al.*, 1980). Spliceosomal snRNPs are large complexes containing both small nuclear, uridine-rich RNAs (the 'U' snRNAs) and proteins. Sequencing of the U1 snRNA revealed that its 5' end was complementary to many 5' splice site sequences, suggesting that the U1 snRNP played a role in the splicing reaction. It was later determined that U1 along with other snRNPs assembled on pre-mRNAs into spliceosomal complexes (Wahl *et al.*, 2009).

Table 1 Comparison of the snRNP components for *Saccharomyces cerevisiae* and *Homo sapiens* spliceosomes. Cartoon structures of the human snRNAs are shown

Complex	S. cerevisiae		H. sapiens		snRNA structure
	snRNA length	No. of proteins	snRNA length	No. of proteins	
U1 snRNP	568nt	17	164nt	10	
U2 snRNP	1175nt	19	187nt	19	
U4/U6.U5 tri-snRNP	112nt 160nt 179–214nt	34	107nt 145nt 116nt	36	
Nineteen complex (NTC)	–	8	–	7	–

The core of the spliceosome is composed of four major, highly conserved subcomplexes: the U1 and U2 snRNPs, the U4/U6.U5 tri-snRNP (composed of three individual snRNPs), and the protein-only nineteen complex (NTC) (Will and Lührmann, 2011). These subcomplexes work together with a host of additional protein factors to process the RNA transcript. Each snRNP contains one or more snRNA and a variable number of snRNP-specific proteins (Table 1). The U1, U2, U4, and U5 snRNPs all contain a heptameric Sm protein ring bound to a U-rich sequence near their 3' end, while the U6 snRNA is bound by the LSM2-8 heptamer (Matera and Wang, 2014). Each snRNP serves a distinct purpose in splicing. The U1 and U2 snRNPs are responsible for recognizing the 5' splice site and branchsite sequences, respectively, through snRNA:pre-mRNA interactions. Basepairing between the U2 snRNA and the pre-mRNA branchsite sequence is responsible for positioning the reactive branchpoint adenosine by bulging it from a U2 snRNA/pre-mRNA duplex to allow for catalysis (Smith et al., 2009).

The U4/U6.U5 tri-snRNP is preassembled prior to spliceosome formation from individual U4/U6 di-snRNPs and U5 snRNPs. U4/U6 contains extensive basepairing between the U4 and U6 snRNAs. The U6 snRNA is the mostly highly conserved of the spliceosomal snRNAs, and it plays a critical role in catalysis by the spliceosome. The U6 snRNA contains at least three distinct features that are absolutely required for catalysis: the ACAGA-box (pronounced 'ah-cah-ga'), the U6 internal stem loop (ISL), and the AGC catalytic triad (Brow, 2002). The U6 ISL and AGC triad are thought to play important roles in catalysis through the coordination of catalytic metal ions, while the ACAGA-box positions the 5' splice site for exon ligation. Proper formation and alignment of these features is aided, in part, by duplex formation between the U2 and U6 snRNAs in the spliceosome. Possibly to prevent premature or aberrant splicing, the U6 ISL is disrupted in the tri-snRNP by annealing of the U6 snRNA to the U4 snRNA, which acts as a chaperone for U6. Therefore, the U4/U6 di-snRNA must be unwound prior to catalysis (Brow, 2002; Wahl et al., 2009).

The U5 snRNA is responsible for contacting the exons of the pre-mRNA to ensure proper splice site alignment

(Newman, 1997). Furthermore, the U5 snRNP contains three important protein components: Prp8, Snu114, and Brr2. Prp8 is a highly conserved 220 kDa protein that is thought to help structure and modulate the active site of the spliceosome as evidenced by a large number of chemical crosslinks between the protein and critical RNA components (Teigelkamp et al., 1995). Prp8 has been shown to interact with the U5 and U6 snRNAs, the 5' and 3' splice sites, and the branchpoint adenosine of the pre-mRNA substrate. Prp8 contains multiple protein domains including those resembling endonucleases, reverse transcriptases, and RNase H (Galej et al., 2013). The endonuclease and reverse transcriptase domains are particularly noteworthy since these resemble proteins encoded by group II self-splicing introns that facilitate DNA homing and intron insertion. This similarity suggests that at least some protein components of the spliceosome, like the snRNAs as described below, evolved from a common ancestor shared by the group II intron machinery. In addition to scaffolding the active site, Prp8 also regulates the activity of Brr2, which is responsible for unwinding the U4/U6 snRNA duplex to allow catalytic structures within U6 to form. The last unique protein component of U5, Snu114, is homologous to the translation elongation factor EF-G and uses GTP hydrolysis to further coordinate Brr2 activity (Small et al., 2006).

The final major subcomplex of the spliceosome is the NTC (Hogg et al., 2010). The NTC was named for a core spliceosomal protein, Prp19, and contains approximately seven other proteins. The NTC is required for stabilizing interactions between the U5 and U6 snRNAs with the pre-mRNA. Additionally, the NTC protein Cwc2 is essential for splicing and contacts the U6 ISL and the duplex formed between the ACAGA-box and 5' splice site (Rasche et al., 2012). This suggests that one function of the NTC may be to bring distal regions of the spliceosome into close proximity for catalysis.

The spliceosome requires many other accessory proteins not stably associated with the U snRNPs or the NTC complex to properly excise introns. For example, two important components for spliceosome assembly are splicing factor 1 (SF1), which recognizes the pre-mRNA branchsite prior to U2

association, and the U2 auxiliary factor heterodimer (U2AF35/65, also known as U2AF1/2) that binds the polypyrimidine tract and 3' splice site AG dinucleotide (Wahl *et al.*, 2009). SF1 is ultimately replaced at the branchsite by U2. Another major class of proteins that associate with the spliceosome, are the DExH/D box proteins. These proteins contain a signature sequence motif of aspartate (D), glutamate (E), usually a hydrophobic amino acid (x), and either a second aspartate or histidine (H). DExH/D box proteins are RNA-dependent ATPases that are extensively involved in RNA metabolism often through destabilization or unwinding of RNA duplexes (Jankowsky, 2011). Finally, metazoan spliceosomes are frequently associated with splicing regulatory (SR) proteins. SR proteins usually contain arginine-serine (RS) repeat motifs and help to guide spliceosome formation at the proper locations in a pre-mRNA (Long and Cáceres, 2009).

In sum, a spliceosome contains 5 snRNAs and a core set of 75 proteins that have been conserved between yeast and humans. The number of accessory proteins that have been isolated with spliceosomes in humans numbers over 300 (Jurica and Moore, 2003). The snRNPs and spliceosomes are consequently some of the largest macromolecular machines inside the cell. Due to their size and compositional complexity, only limited high-resolution crystallographic structural information is available for the individual snRNPs or spliceosomal proteins. However, lower resolution cryo-electron microscopy structures are available for several snRNPs and spliceosomal complexes. Obtaining high-resolution structural information of spliceosomes and spliceosomal components is currently being pursued by many laboratories.

Spliceosomes are Assembled from snRNPs on Introns

The spliceosome is not a static complex, but rather a dynamic machine that assembles stepwise from preformed catalytically inactive subcomplexes for each round of splicing (Hoskins *et al.*, 2011). In the canonical stepwise assembly pathway, the U1 snRNP first recognizes the 5' splice site through base-pairing of the 5' end of the U1 snRNA to the pre-mRNA (Figure 2). Elsewhere on the intron, the branchsite, polypyrimidine tract, and 3' splice sites are recognized by SF1, U2AF65, and U2AF35, respectively. Together, this forms the first intermediate in spliceosome assembly, the E complex. E complex is then converted into A complex upon ATP-dependent exchange of SF1 for U2 at the branchsite. On long introns, U1 and U2 may first interact across exons to form an exon definition complex prior to interactions that occur across the intron to be spliced (Wahl *et al.*, 2009). Regardless, A complex formation is often followed by the recruitment of the U4/U6.U5 tri-snRNP to the pre-mRNA substrate to form the catalytically inactive B complex spliceosome that contains all five of the U snRNPs. Assembly of B complex is followed by Br2-dependent activation and the formation of B^{ACT}, wherein the NTC has joined the complex, U1 and U4 snRNPs have been lost, and the active site of the spliceosome has begun to form. Activation involves multiple rearrangements that result in U6 replacement of U1 at the 5' splice site, formation of the U6 ISL, and duplex formation between the U2 and U6 snRNAs. The action of the DExD/H box protein Prp2 helps to promote the first step of catalysis (Kim and Lin, 1993), and spliceosomes capable of carrying out the chemistry of splicing are called C complexes. The first step of catalysis occurs,

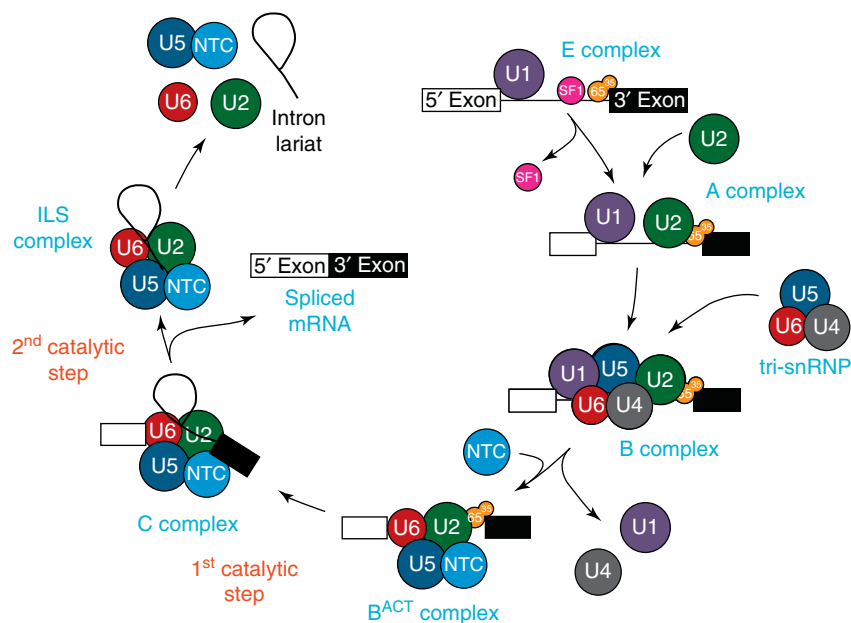


Figure 2 Stepwise assembly mechanism of the spliceosome. Spliceosome assembly begins at E complex, with U1 bound at the 5' splice site, SF1 at the branchsite sequence, and U2AF65/35 at the 3' splice site. Exchange of SF1 for U2 generates A complex, which is converted into B complex upon addition of the U4/U6.U5 tri-snRNP. Activation of the spliceosome releases the U1 and U4 snRNPs from the spliceosome and generates the catalytically ready B^{ACT} complex. It is unknown if U2AF65/35 are present in the B^{ACT} spliceosome and are thus shown in lighter colors. The first step of catalysis occurs and C complex forms. After the second step of catalysis, the mRNA is released and the ILS is disassembled for further rounds of splicing.

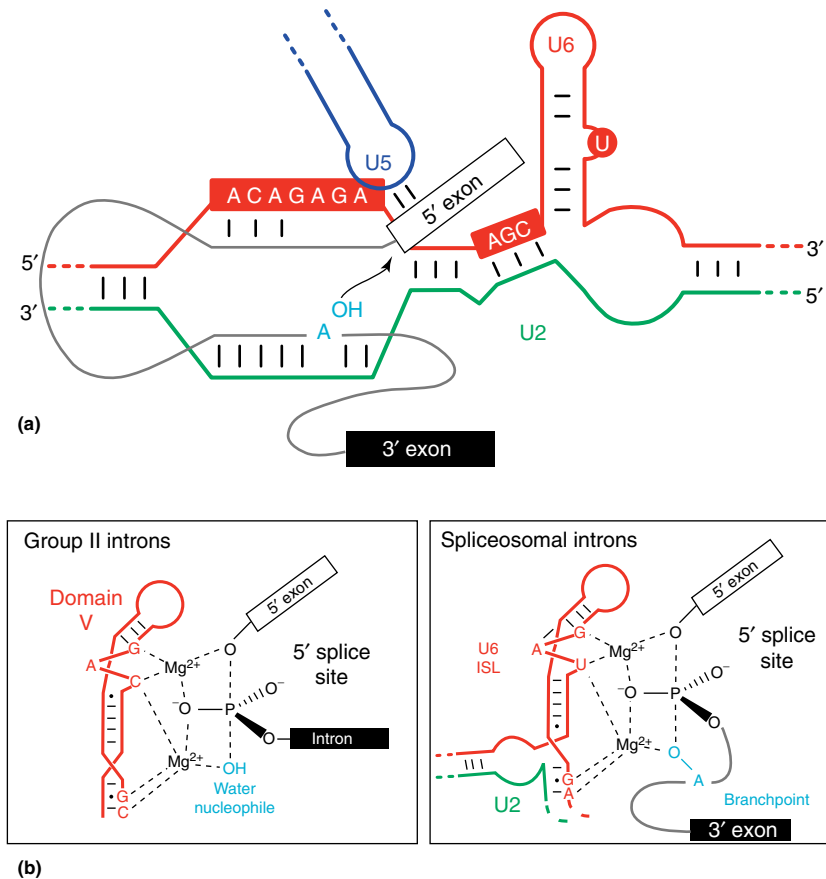


Figure 3 The active site of the spliceosome. (a) The catalytically active spliceosome contains a complex containing the U2, U6, and U5 snRNPs. U2 and U6 snRNAs are extensively base paired with each other and the substrate. U2 makes contact with the branchsite region and results in the formation of the bulged adenosine residue that provides the nucleophile for the first step. U6 contacts the 5' splice site with the conserved ACAGAGA sequence and helps to coordinate catalytic metal ions with the AGC triad, the U6 ISL, and a bulged uridine. The U5 snRNA contacts the 5' exon to help position the substrate for catalysis. (b) Comparison of the group II intron and spliceosomal active sites. The U6 ISL is homologous the domain V of the group II intron. Both coordinate catalytic metal ions (probably Mg^{2+}) and contain similar geometries to promote catalysis. The first step of catalysis is shown for both enzymes and some group II introns can use water nucleophiles to promote 5' exon cleavage. Figures in (a) and (b) were adapted from Konarska, M.M., Vilardell, J., Query, C.C., 2006. Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Molecular Cell* 21 (4), 543–553; and Fica, S.M., Tuttle, N., Novak, T., *et al.* 2013. RNA catalyses nuclear pre-mRNA splicing. *Nature* 503 (7475), 229–234, respectively and have been used with permission of the publishers.

followed by several rearrangements within the active site facilitated by the DEXD/H box protein Prp16 that reposition the components for the second step of splicing (Schwer and Guthrie, 1991). After the second step, the spliced mRNA is released from the intron lariat spliceosome (ILS), and the post-catalytic complex is disassembled and the components are recycled for another round of splicing (Fourmann *et al.*, 2013). The spliceosome is a single turnover enzyme, meaning that the assembly and activation must occur a new on each intron to be spliced.

The Active Site of the Spliceosome

Spliceosomes must bring together distant regions of the pre-mRNA along with spliceosomal snRNAs and proteins that enable catalysis. Alignment of the reactive groups occurs in part through scaffolding of the snRNA and pre-mRNA through the basepairing interactions between the intron and the U2

and U6 snRNAs (Figure 3(a)). These reactive groups must also be aligned with the U6 ISL, and it is thought that spliceosomal proteins such as Prp8 and the NTC play key roles in juxtaposing all of these functional groups. A number of experiments have found evidence that the U6 ISL plays a key role in coordinating essential magnesium ions for catalysis (Fica *et al.*, 2013). Despite the necessity for proper alignment, the spliceosome active site must also be flexible since it is remodeled between the first and second steps of splicing to permit juxtaposition of the 5' exon and 3' splice site for exon ligation.

One striking feature of the spliceosomal active site is its similarity to that found in group II introns. Group II introns are catalytic pieces of 'self-splicing' RNA that carry out identical chemical steps (5' splice site cleavage/lariat formation and exon ligation) to those of the spliceosome. For both spliceosomes and group II introns, the chemical steps have been shown to be reversible and metal-ion dependent (Chin and Pyle, 1995; Tseng and Cheng, 2008). Remarkably, domain V of group II introns possesses many of the same catalytic

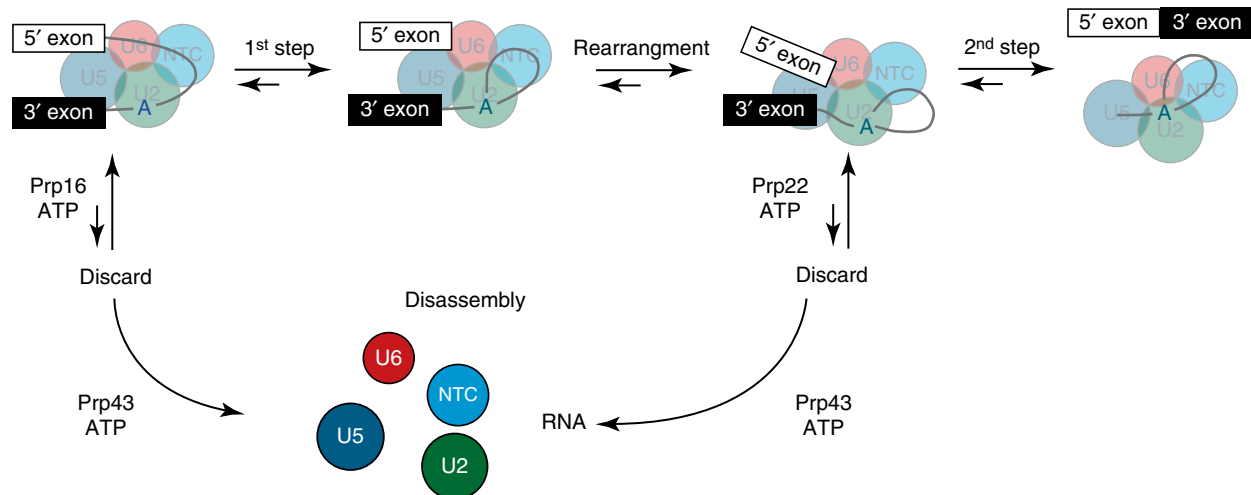


Figure 4 Kinetic proofreading during splicing. The forward reactions (toward spliced mRNA) are favored but suboptimal substrates can enter into the pre-mRNA discard pathway. Entry to these pathways is controlled by the action of the DEAH-box ATPases Prp16 and Prp22. While entry into the discard pathway is reversible, discard becomes irreversible if the spliceosomes are subsequently disassembled by the DEAH-box ATPase Prp43.

features as the U6 snRNA (Figure 3(b)). Crystal structures of group II introns have been used to model the spliceosomal active site (Fica *et al.*, 2014) and enabled experiments that found direct evidence for the presence of two catalytic magnesium ions (Toor *et al.*, 2008; Fica *et al.*, 2013). It is likely that the spliceosome and group II intron use a ‘two-metal mechanism’ (Steitz and Steitz, 1993) for catalyzing phosphodiester bond cleavage and formation with each magnesium ion playing a role in stabilizing the nucleophile or leaving group. In addition to the similarity between Prp8 domains and group II intron encoded proteins, this conservation of active site structure and mechanism between the spliceosome and group II introns provide strong evidence for evolution of the spliceosome from a group II intron-like ancestor.

DExH Box Proteins are Essential Cofactors of the Spliceosome

DExH box proteins are ubiquitously associated with processes that involve RNA, ranging from pre-mRNA splicing to mRNA translation and degradation. These enzymes couple ATP hydrolysis to RNA binding or duplex unwinding (helicase) activity in order to facilitate structural or compositional rearrangements in different RNA–protein complexes. Eight different DExH box ATPases are required for splicing: the DEAD-box proteins UAP56, Prp5, and Prp28; the DEIH-box U4/U6 unwindase Brr2; and the DEAH-box proteins Prp2, Prp16, Prp22, and Prp43 (Chang *et al.*, 2013). Each plays a distinct role in splicing. UAP56 and Prp5 are involved in spliceosome assembly; Prp28, Brr2, and Prp2 are involved in activating the spliceosome for catalysis; Prp16 and Prp22 are responsible for remodeling the catalytic spliceosome; and Prp43 is necessary for spliceosome disassembly. Consequently, while the chemical steps of splicing do not require ATP, ATP hydrolysis by DExH box proteins is necessary for splicing.

In addition to their roles in regulating spliceosomal conformational changes, several of the ATPases also have roles in splicing fidelity. Examples of this can be found in studies of the yeast homologs of Prp16 and Prp22 (Mayas *et al.*, 2006; Koodathingal *et al.*, 2010). The ATPase activity of Prp16 is involved in transitioning the spliceosome from the first to second step of catalysis. Prior to exon cleavage, Prp16 is also involved in ‘kinetic proofreading’ of suboptimal substrates (Figure 4). Spliceosomes that have assembled on pre-mRNAs with mutations in the branchsite region proceed through the catalytic steps more slowly than RNAs that contain the consensus sequences. In these suboptimal spliceosomes, hydrolysis of ATP diverts the complex from the splicing pathway to a spliceosome discard pathway. In this way, Prp16 is able to increase the fidelity of pre-mRNA splicing by rejecting suboptimal substrates or poorly assembled spliceosomes prior to the first step. This kinetic proofreading step is also seen with Prp22, which monitors the spliceosome after exon cleavage but prior to exon ligation. The DEAH-box protein Prp43 provides a critical, irreversible step during discard by disassembling these spliceosomes.

Alternative Splicing Creates Multiple Products from a Single Gene

While many introns present in eukaryotic genes are spliced from the pre-mRNA transcript constitutively, the removal of other introns can be subject to additional layers of regulation. By altering the splice site decisions during intron removal, multi-exonic transcripts can be spliced together to produce many isoforms of the same gene (Figure 5). This process is known as alternative splicing. In mammalian cells, 95% of gene transcripts are thought to undergo alternative splicing (Nilsen and Graveley, 2010). The generation of multiple mRNA isoforms from a single pre-mRNA is a powerful method to increase the protein-encoding capacity of the genome

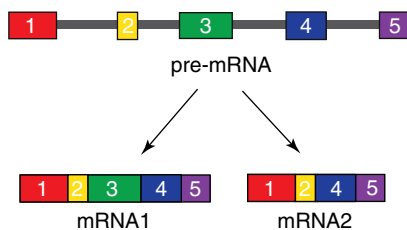


Figure 5 Alternative splicing of a pre-mRNA. Single pre-mRNAs can generate multiple splicing products through the use of alternative splice sites. mRNA 1 contains exon 3 (green) which is not found in mRNA 2.

without increasing the size of the genome. Alternative splicing represents an important source of proteomic diversity. The most common form of alternative splicing is exon skipping in which specific exons are expressed in only a fraction of the total transcripts. Exon skipping can commonly be found in different tissue isoforms of the same gene. The first example of alternative splicing in humans was for gene encoding the peptide hormone calcitonin (Leff and Rosenfeld, 1986). Two distinct isoforms of the calcitonin mRNA exist, one that contains only exons 1–4 and a second that contains exons 1–3, 5, and 6. The first mRNA transcript is found predominately in the brain and encodes for calcitonin and the second encodes another peptide called calcitonin-gene-related peptide, which predominates in the thyroid.

Splicing is Coupled to Other Cellular Processes

Splicing does not occur in isolation but rather is connected to other cellular processes such as transcription, RNA capping, polyadenylation, and chromatin remodeling. Work in several organisms suggests that splicing can occur during or after the synthesis of pre-mRNA. Indeed, interactions between RNA polymerase II and the spliceosomal snRNPs are thought to recruit them to the nascent transcript for processing (Bentley, 2014). These interactions may ultimately influence the alternative splicing fate of the pre-mRNA. In humans and other metazoans, nuclear splicing and cytosolic translation and RNA decay can even be linked (Nott *et al.*, 2004). This occurs through deposition of a set of proteins called the exon junction complex (EJC) by the spliceosome just upstream of ligated exons. The presence of EJCs on an mRNA can have profound consequences for the fate of that mRNA once it enters the cytosol. EJCs can influence the number of proteins that can be translated from a mRNA, target the mRNA for rapid degradation, or impact other aspects of mRNA biology (Moore and Proudfoot, 2009).

Defects in Splicing Cause Disease

More than 15% of human genetic diseases arise from single point mutations that cause defects related to pre-mRNA splicing (Faustino and Cooper, 2003). These mutations can be classified as either *cis*- or *trans*-acting. *Cis*-acting mutations affect the pre-mRNA substrate, usually at constitutive or alternative splice sites but other regions of the RNA can be

affected as well. These mutations change how the pre-mRNA is processed by the spliceosome and can result in, for example, failure to remove introns, choice of incorrect splice sites, or deletions of entire exons in the mRNA. Ultimately this may change the protein that is encoded by the mRNA; sometimes by just a few amino acids as is the case in Frasier syndrome. In this disease a mutant 5' splice site causes deletion of just three amino acids from the Wilms Tumor suppressor protein, WT1, and results in severe defects in urogenital development.

Trans-acting mutations affecting the basal splicing machinery can also cause severe human diseases such as retinitis pigmentosa (RP), cancer, or spinal muscular atrophy (SMA). RP is the leading cause of inherited blindness in the world and is characterized by progressive retinal degeneration. This eventually results in total blindness from loss of rod photoreceptor cells. Several genetic causes of RP exist, with some resulting from defects in components of the U4/U6.U5 tri-snRNP, such as Prp8 or Brr2. Some cancers, such as the blood cell cancers caused by myelodysplastic syndromes, have also been found to be specifically associated with single point mutations in components of the U2 snRNP, SF1, or U2AF35/65. Degenerative neuromuscular disorders such as SMA can arise from defects in production and assembly of the snRNPs themselves. SMA is attributed to mutations in the survival of motor neuron complex, which plays a major role in the assembly of newly produced U1, U2, U4, and U5 snRNPs. Patients with SMA exhibit a progressive loss of spinal cord motor neurons leading to total paralysis of the voluntary muscles and early death.

Conclusion

The discovery of RNA splicing represented one of the great surprises in biology during the last century. Equally surprising was the complexity of the spliceosomal machinery that carries out this reaction: a multimegadalton complex of snRNAs and proteins that briefly associate together on a transcript to remove an intron before being dismantled. In addition to this complexity and transience, splicing is absolutely essential for processing of the vast majority of human mRNAs. It is likely that the complexity of the spliceosome is precisely what allows it to function on a vast array of substrates and integrate with other cellular processes to produce correct mRNAs. In the coming years advances in structural biology, biophysical methods, and genetic engineering will lead to new insights into the structures of the spliceosome and the underlying biochemical mechanisms of RNA splicing.

See also: Nucleic Acid Synthesis/Breakdown: Transcription: Pre-mRNA Splicing: Function and Dysfunction

References

- Bentley, D.L., 2014. Coupling mRNA processing with transcription in time and space. *Nature Reviews Genetics* 15, 163–175.
- Berget, S.M., Moore, C., Sharp, P.A., 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proceedings of the National Academy of Sciences of the United States of America* 74 (8), 3171–3175.

- Brody, E., Abelson, J., 1985. The "spliceosome": Yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science* 228 (4702), 963–967.
- Brow, D.A., 2002. Allosteric cascade of spliceosome activation. *Genetics* 36, 333–360.
- Chang, T.-H., Tung, L., Yeh, F.-L., Chen, J.-H., Chang, S.-L., 2013. Functions of the DExD/H-box proteins in nuclear pre-mRNA splicing. *Biochimica et Biophysica Acta — Gene Regulatory Mechanisms* 1829 (8), 764–774.
- Chin, K., Pyle, A.M., 1995. Branch-point attack in group II introns is a highly reversible transesterification, providing a potential proofreading mechanism for 5'-splice site selection. *RNA* 1 (4), 391–406.
- Chow, L.T., Gelinas, R.E., Borker, T.R., Roberts, R.J., 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* 12 (1), 1–8.
- Faustino, N.A., Cooper, T.A., 2003. Pre-mRNA splicing and human disease. *Genes & Development* 17 (4), 419–437.
- Fica, S.M., Mefford, M.A., Piccirilli, J.A., Staley, J.P., 2014. Evidence for a group II intron-like catalytic triplex in the spliceosome. *Nature Structural & Molecular Biology* 21 (5), 464–471.
- Fica, S.M., Tuttle, N., Novak, T., *et al.*, 2013. RNA catalyses nuclear pre-mRNA splicing. *Nature* 503 (7475), 229–234.
- Fourmann, J.B., Schmitzová, J., Christian, H., *et al.*, 2013. Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. *Genes & Development* 27 (4), 413–428.
- Galej, W.P., Oubridge, C., Newman, A.J., Nagai, K., 2013. Crystal structure of Prp8 reveals active site cavity of the spliceosome. *Nature* 493 (7437), 1–7.
- Hogg, R., McGrail, J., O'Keefe, R., 2010. The function of the NineTeen Complex (NTC) in regulating spliceosome conformations and fidelity during pre-mRNA splicing. *Biochemical Society Transactions* 38 (4), 1110–1115.
- Hoskins, A.A., Friedman, L.J., Gallagher, S.S., *et al.*, 2011. Ordered and dynamic assembly of single spliceosomes. *Science* 331 (6022), 1289–1295.
- Jankowsky, E., 2011. RNA helicases at work: Binding and rearranging. *Trends in Biochemical Sciences* 36 (1), 19–29.
- Jurica, M.S., Moore, M.J., 2003. Pre-mRNA splicing: Awash in a sea of proteins. *Molecular Cell* 12 (1), 5–14.
- Kim, S.-H., Lin, R.-J., 1993. Pre-mRNA splicing within an assembled yeast spliceosome requires an RNA-dependent ATPase and ATP hydrolysis. *Proceedings of the National Academy of Sciences of the United States of America* 90 (3), 888–892.
- Konarska, M.M., Vilardell, J., Query, C.C., 2006. Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Molecular Cell* 21 (4), 543–553.
- Koodathingal, P., Novak, T., Piccirilli, J.A., Staley, J.P., 2010. The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing. *Molecular Cell* 39 (3), 385–395.
- Leff, S.E., Rosenfeld, M.G., 1986. Complex transcriptional units: Diversity in gene expression by alternative RNA processing. *Annual Review of Biochemistry* 55 (1), 1091–1117.
- Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L., Steitz, J.A., 1980. Are snRNPs involved in splicing? *Nature* 283 (5743), 220–224.
- Long, J.C., Caceres, J.F., 2009. The SR protein family of splicing factors: Master regulators of gene expression. *Biochemical Journal* 417 (1), 15.
- Matera, A.G., Wang, Z., 2014. A day in the life of the spliceosome. *Nature Reviews Molecular Cell Biology* 15 (2), 108–121.
- Mayas, R.M., Maita, H., Staley, J.P., 2006. Exon ligation is proofread by the DExD/H-box ATPase Prp22p. *Nature Structural & Molecular Biology* 13 (6), 482–490.
- Moore, M.J., Proudfoot, N.J., 2009. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* 136 (4), 688–700.
- Moore, M.J., Sharp, P.A., 1993. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature* 365 (6444), 364–368.
- Newman, A.J., 1997. The role of U5 snRNP in pre-mRNA splicing. *The EMBO Journal* 16 (19), 5797–5800.
- Nilsen, T.W., Graveley, B.R., 2010. Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463 (7280), 457–463.
- Nott, A., Le Hir, H., Moore, M.J., 2004. Splicing enhances translation in mammalian cells: An additional function of the exon junction complex. *Genes & Development* 18 (2), 210–222.
- Rasche, N., Dybkov, O., Schmitzová, J., *et al.*, 2012. Cwc2 and its human homologue RBM22 promote an active conformation of the spliceosome catalytic centre. *The EMBO Journal* 31 (6), 1591–1604.
- Roca, X., Krainer, A.R., 2009. Recognition of atypical 5' splice sites by shifted base-pairing to U1 snRNA. *Nature Structural & Molecular Biology* 16 (2), 176–182.
- Scherer, S., 2008. *A Short Guide to the Human Genome*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Schwer, B.B., Guthrie, C.C., 1991. PRP16 is an RNA-dependent ATPase that interacts transiently with the spliceosome. *Nature* 349 (6309), 494–499.
- Sharp, P.A., 1993. Split genes and RNA splicing. *Nobel Lectures, Physiology or Medicine* 1995, 145–174.
- Small, E.C., Leggett, S.R., Winans, A.A., Staley, J.P., 2006. The EF-G-like GTPase Snu114p regulates spliceosome dynamics mediated by Brr2p, a DExD/H Box ATPase. *Molecular Cell* 23 (3), 389–399.
- Smith, D.J., Konarska, M.M., Query, C.C., 2009. Insights into branch nucleophile positioning and activation from an orthogonal pre-mRNA splicing system in yeast. *Molecular Cell* 34 (3), 333–343.
- Soeiro, E., Birnboim, H.C., Darnell, J.E., 1966. Rapidly labeled HeLa cell nuclear RNA: II. Base composition and cellular localization of a heterogeneous RNA fraction. *Journal of Molecular Biology* 19 (2), 362–372.
- Steitz, T.A., Steitz, J.A., 1993. A general two-metal-ion mechanism for catalytic RNA. *Proceedings of the National Academy of Sciences of the United States of America* 90 (14), 6498–6502.
- Teigelkamp, S., Newman, A.J., Beggs, J.D., 1995. Extensive interactions of PRP8 protein with the 5' and 3' splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. *The EMBO Journal* 14 (11), 2602–2612.
- Toor, N., Keating, K.S., Taylor, S.D., Pyle, A.M., 2008. Crystal structure of a self-spliced group II intron. *Science* 320 (5872), 77–82.
- Tseng, C.K., Cheng, S.C., 2008. Both catalytic steps of nuclear pre-mRNA splicing are reversible. *Science* 320 (5884), 1782–1784.
- Wahl, M.C., Will, C.L., Lührmann, R., 2009. The spliceosome: Design principles of a dynamic RNP machine. *Cell* 136 (4), 701–718.
- Will, C.L., Lührmann, R., 2011. Spliceosome structure and function. *Cold Spring Harbor Perspectives in Biology* 3 (7), a003707.

Further Reading

- Atkins, J.F., Gesteland, R.F., Cech, T.R. (Eds.), 2011. *RNA Worlds: From Life's Origins to Diversity in Gene Regulation*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Darnell, J., 2011. *RNA: Life's Indispensable Molecule*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Pre-mRNA Splicing: Function and Dysfunction

AM Case, Abbott Molecular Inc., Des Plaines, IL, USA

IA Sawyer, M Dundr, and ML Hastings, Chicago Medical School, Rosalind Franklin University of Medicine and Science, North Chicago, IL, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Most genes in higher eukaryotes are transcribed into a pre-mRNA that is comprised of expressed sequences, called exons, and intervening sequences, termed introns. Discovered in 1977, introns are removed from pre-messenger RNA (pre-mRNA), in a process called splicing, to generate mature mRNA (Berget *et al.*, 1977; Chow *et al.*, 1977). Most pre-mRNA transcripts can undergo different patterns of splicing, defined as alternative splicing (Pan *et al.*, 2008; Wang *et al.*, 2008). The diversity in transcripts that arises through alternate splicing, and the regulation of their production, is thought to be one of the benefits of a multi-exon gene system. At the same time, a drawback of requiring the removal of introns for appropriate protein production is apparent from the prevalence of human disease-causing mutations that disrupt splicing. By some estimates, up to 60% of pathological mutations may cause aberrant splicing (Lopez-Bigas *et al.*, 2005). Mutations that affect splicing can, for example, change splicing patterns, induce the use of aberrant splice sites or alter splicing efficiency to affect gene expression by gain or loss of function coding changes that result in pathogenesis. In recent years, screens for causative mutations in human disease have identified likely pathological mutations in novel components of the splicing machinery. Interestingly, there are phenotypic themes shared between groups of splicing factors, which provide insight into the function and dysfunction of splicing in humans.

Splicing

Splicing must occur with high fidelity in order to ensure the integrity of the resulting mRNA (Semlow and Staley, 2012). Any errors in the process can result in changes in the reading frame, which would disrupt protein coding or other functional aspects of the mRNA (Hsu and Hertel, 2009). Consensus sequences, called splice sites, are found at the junctions of introns and exons and identify where splicing should occur. The 5' splice site (5'ss) defines the 5' end of an intron and the 3' splice site (3'ss) defines the 3' end (Figure 1(a)). The most common dinucleotides at the 5' and 3' ends of an intron are GU and AG, respectively, though there are noncanonical splice sites that do not conform to this rule. Many introns have a polypyrimidine tract upstream of the 3'ss which is preceded by a branchpoint sequence (BPS). The consensus sequences for these sequence elements are relatively degenerate in mammals and occur at a high frequency within a genome. Thus, they do not, themselves, provide enough sequence information to distinguish bona fide splice sites from those that match the consensus sequence of a splice site but are not used for splicing (Stephens and Schneider, 1992; Roca *et al.*, 2013). It has been a major goal of splicing research to understand how splicing

occurs with high fidelity and specificity at authentic splice sites. Much progress has been made toward understanding a so-called 'splicing code,' which predicts how splice sites, additional RNA features and proteins work together to coordinate and regulate splicing (Barash *et al.*, 2010).

Splicing Catalysis

Splicing involves two transesterification reactions that result in cleavage at the splice sites at the ends of the introns, removal of the intron, and joining the flanking exons together (Horowitz, 2012; Horowitz and Abelson, 1993). The first step of splicing involves rearrangement of snRNPs to allow a nucleophilic attack by the 2'-hydroxyl group of an adenosine in the BPS on the guanine nucleotide at the 5'ss, which results in cleavage of the 5'ss and production of an intron lariat (Reed and Maniatis, 1988). This reaction results in the formation of a 'free' upstream exon and a lariat intron contiguous with the downstream RNA. In the next transesterification reaction, the hydroxyl group at the 3' end of the upstream exon attacks the phosphodiester bond at the 3'ss. The adjoining exons are thereby covalently bound and the intron is released in the form of a lariat, which is degraded by an exonuclease, or, less commonly, processed to yield various noncoding RNAs (Hesselberth, 2013). Single-molecule studies of the splicing reaction have begun to reveal the complexity of the RNA and protein rearrangements and structural changes in the spliceosome that are required for splicing catalysis (Abelson *et al.*, 2010; Hoskins *et al.*, 2011; Figure 1(b)).

The Spliceosome

The splicing reaction is carried out by the spliceosome, a multimegadalton ribonucleoprotein complex that is comprised of small RNAs and hundreds of auxiliary proteins (Jurica and Moore, 2003). Splicing is coordinated, in large part, by five distinct uridine-rich small nuclear ribonucleoproteins (U snRNPs). Each U snRNP is comprised of a small nuclear RNA (snRNA) and auxiliary protein factors (Will and Luhrmann, 2011). The snRNA component of a snRNP confers specificity to the splicing reaction in that it base-pairs directly with the pre-mRNA to identify splice sites and direct splicing. Base-pairing between the snRNAs is essential for positioning of the pre-mRNA for the catalytic steps of splicing. The auxiliary protein components of snRNPs stabilize these relatively weak bonds and also mediate interactions with other splicing factors within the spliceosome to facilitate intron excision and exon ligation during splicing.

Spliceosomal snRNPs coordinate splicing through a process involving the specific recognition of splice sites and repositioning of RNA:RNA and RNA:protein interactions, that ultimately allows the catalytic steps of splicing. For most

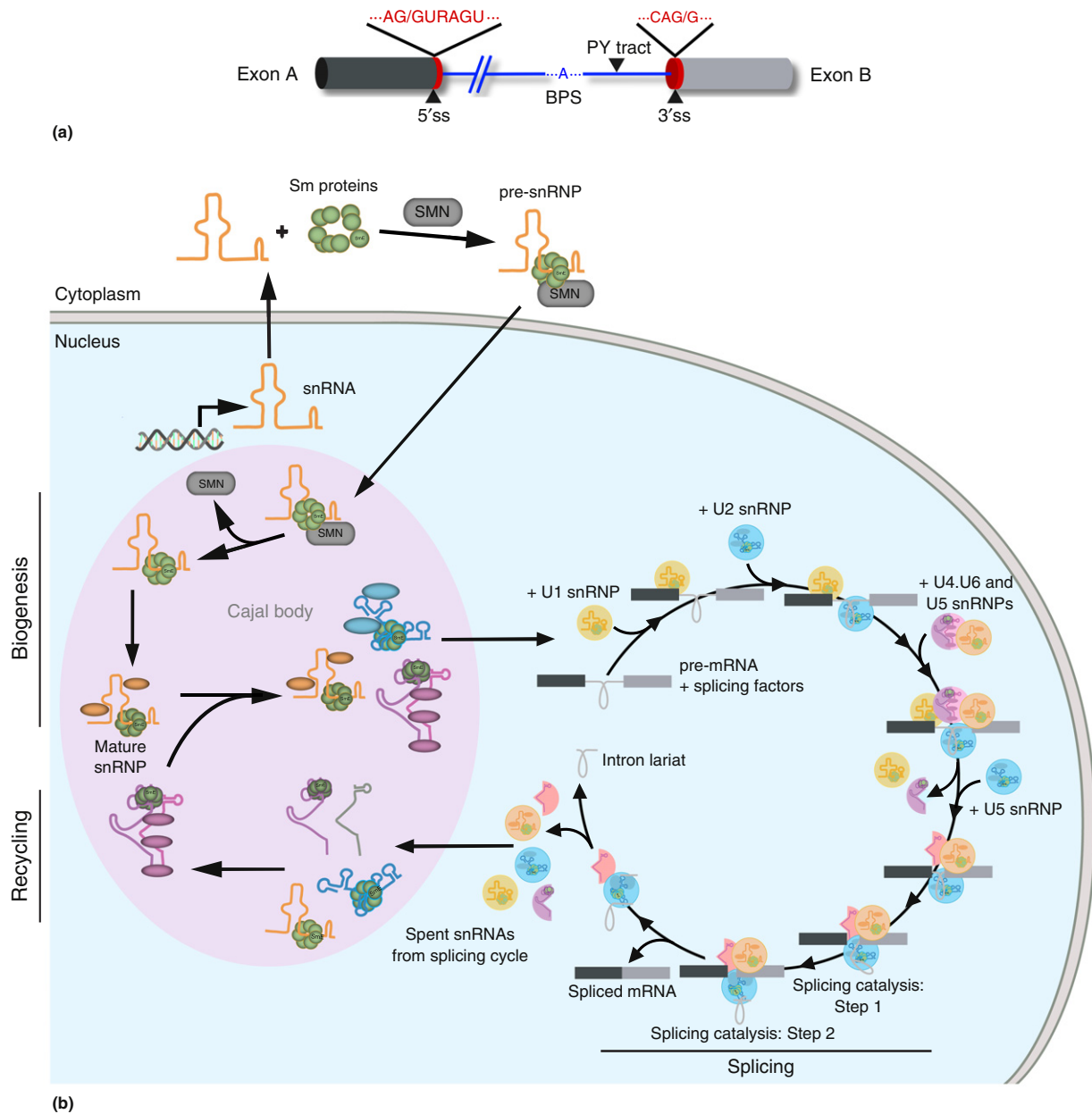


Figure 1 snRNP biogenesis, recycling and pre-mRNA splicing. (a) Sequence features of human exons and introns, relevant to pre-mRNA splicing. The 5' splice site (5'ss) and 3' splice site (3'ss) define exon–intron boundaries and are bound by snRNPs, whereas the polypyrimidine tract (PY tract) and branch point sequence (BPS) facilitate spliceosome assembly and splicing catalysis respectively. (b) snRNP biogenesis begins with transcription, which occurs near the Cajal body, and shuttling to the cytoplasm for Sm protein loading via a SMN multiprotein complex. The SMN complex is shuttled with the nascent pre-snRNP back to the Cajal body within the nuclear compartment, which serves as a site of snRNP maturation and recycling. snRNPs then enter the splicing cycle, where introns are removed from pre-mRNA by a complex rearrangement mediated primarily by snRNPs and splicing factors.

introns, U1 snRNP binds to the pre-mRNA through base-pairing between the U1 snRNA and specific nucleotides of the 5'ss consensus sequence (Mount *et al.*, 1983; Roca *et al.*, 2013). U2 snRNP associates with the intronic branch site region (Query *et al.*, 1996; Will and Luhrmann, 2011). The snRNPs bind to pre-mRNAs in an ATP-dependent fashion that is assisted by auxiliary proteins such as SF1, U2AF1 (U2AF35), U2AF2 (U2AF65), and precursor RNA processing proteins (Prps) to promote and stabilize the relatively weak

associations of snRNPs to the pre-mRNA (Matera and Wang, 2014). Additional intronic features, summarized in Figure 1(a), such as the polypyrimidine tract (PY tract), which resides between the BPS and the 3'ss, also aid in the splicing process by facilitating the binding of auxiliary splicing factors such as SR proteins, hnRNP proteins, U2AF65, PUF60, and polypyrimidine tract binding proteins (PTBs) to enable subsequent spliceosomal assembly (De Conti *et al.*, 2013; Hastings *et al.*, 2007; Hastings and Krainer, 2001b).

Following U2 snRNP binding, a tri-snRNP complex, consisting of U4/U6 and U5 snRNPs, stably associates with the pre-mRNA in an ATP-dependent manner (Mroczek and Dziembowski, 2013; Roscigno and Garcia-Blanco, 1995; Turner *et al.*, 2004). U5 and U6 snRNPs interact with the 5' splice region (Kandels-Lewis and Seraphin, 1993; Newman and Norman, 1991). The U4/U6 di-snRNPs exhibit extensive base-pairing between their snRNAs, and additional shared associations between protein auxiliary factors aid binding within U4/U6 and between U4/U6 and U5 snRNP (Konarska and Sharp, 1987). Major conformational changes occur between snRNPs, pre-mRNA, and associated proteins to make the spliceosome catalytically competent for the splicing reactions.

There are two types of spliceosomes in eukaryotic cells (Patel and Steitz, 2003). The major or U2-dependent spliceosome recognizes consensus splice site sequences that define most introns. The minor or U12-dependent spliceosome, recognizes a distinct, rare type of splice site that has a unique consensus sequence (Hall and Padgett, 1996; Tarn and Steitz, 1996; Turunen *et al.*, 2013). The 5' and 3' ends of U12-type introns can be either an AU and AC, or a GU and AG. These introns were originally termed AT-AC introns for their unique AU/AC splice site dinucleotides. Both spliceosomes work in parallel, each with their own subset of U snRNPs, except for U5 snRNP, which is a component of both spliceosomes (Will and Luhrmann, 2005). Components of the minor spliceosome – U11, U12, and U4atac/U6atac snRNPs – are functional homologs of the major spliceosomal U1, U2, and U4/U6 snRNPs in the major spliceosome, respectively (Tarn and Steitz, 1996). The minor spliceosome removes ~0.3% of all introns (Sheth *et al.*, 2006; Levine and Durbin, 2001). Splicing of U12-dependent introns is similar to splicing of U2-type introns (Hastings and Krainer, 2001a; Luo *et al.*, 1999; Will *et al.*, 1999). There are some exceptions to this rule, for example, the U11/U12 di-snRNP is preformed and binds the intron as a unit rather than independently as do U1 and U2

snRNPs in the major spliceosomal reaction (Frilander and Steitz, 1999). Additionally, consensus sequence base pairing in U12-type introns is mediated via the U11-48K protein (Turunen *et al.*, 2008).

Alternative Splicing

Most splice sites can be linked together in different combinations to create more than one type of mRNA transcript from a single gene. These alternatively spliced mRNA isoforms may have different stabilities or encode protein isoforms with distinct functions (Kelemen *et al.*, 2013). Alternative splicing is a major mechanism of protein diversity in metazoans. In mammals, more than 90% of multi-exon gene transcripts are alternatively spliced (Pan *et al.*, 2008; Wang *et al.*, 2008). The main patterns of alternative splicing events are alternative 5' and 3' splice sites, mutually exclusive exons, exon inclusion/exclusion (cassette exons), and intron retention (Figure 2). Exon inclusion/exon skipping appears to be the most abundant alternative splicing event in mammals, followed by the use of alternative splice sites (Sugnet *et al.*, 2004). Although the majority of alternative splicing falls into one of these categories, there are other, less common types of alternative splicing that have been described, including the recognition of the same sequence alternatively as either a 5' splice site or a 3' splice site (Zhang *et al.*, 2007).

Alternative splicing is regulated by many of the same proteins that are involved in constitutive splicing. However, the expression of a particular splice variant is the culmination of many inputs and can be regulated at multiple levels. As may be hypothesized from a dynamic process populated by a diverse range of players, not all splice variants are generated in the same manner (Braunschweig *et al.*, 2013). Furthermore, disease-related changes to splicing dynamics can further increase the heterogeneity of splice site determination (Zhang and Manley, 2013).

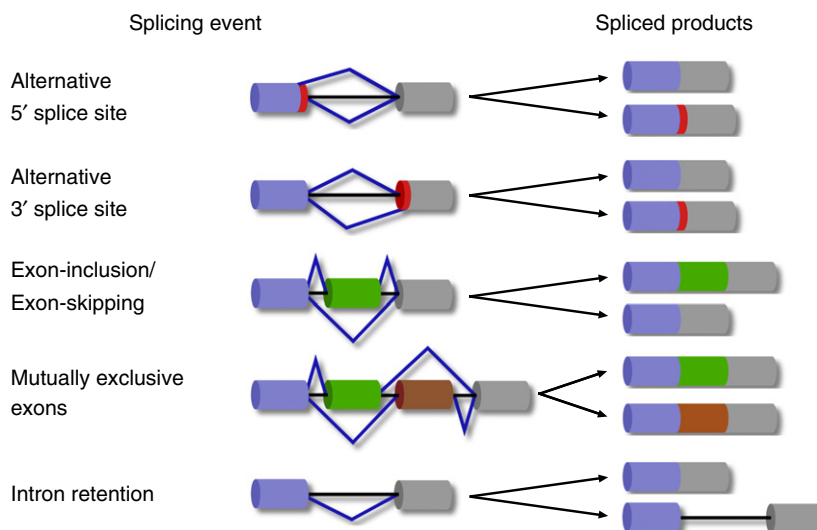


Figure 2 Alternative splicing. Dominant patterns of alternative splicing include alternative use of 5' or 3' splice sites, exon inclusion or skipping (including mutually exclusive exons) and intron retention. Exons are represented by colored cylinders and introns by black lines. Alternate splicing events are shown in blue lines, which span the region removed from the pre-mRNA.

Cis-Acting Sequences and Trans-Acting Factors

A continuing issue in splicing is the question of how exons and introns are differentially identified in constitutive splicing and how recognition of splice sites is regulated in alternative splicing. There are a multitude of sequence features, both positive and negative, that contribute to the regulation of alternative splicing, including intron and exon length, protein binding sites, and secondary structure (McManus and Graveley, 2011). Splicing pathways are the result of complex combinatorial control resulting from the influence of any and all of these features within a transcript.

One example for how exon–intron sequence and organization influence RNA splicing was first proposed by Susan Berget in 1990 as the exon definition model of splicing (Robberson *et al.*, 1990). This pioneering work demonstrated that the exon was the initiating unit of spliceosome assembly and that interactions between the 3′ss and 5′ss across an exon were critical for distinguishing the correct splice sites and avoiding exon skipping. The exon definition model was supported by evidence that increasing exon sequences to more than 300 nucleotides ablated alternative splicing in most cases (Robberson *et al.*, 1990). Later, using a genome-wide approach, Bolisetty and Beemon (2012) showed that approximately 40% of large internal exons (>1000 nts) differentially spliced. The influence of exon length on exon skipping was further demonstrated by experiments showing that artificial elongation (Black, 1991) or shortening (Dominski and Kole, 1991) of exon sequences increases or decreases skipping, respectively. Furthermore, exon inclusion, even in longer exons, is promoted when the exons are flanked by shorter introns, thereby promoting interactions between the 5′ss and 3′ss across the intron (Sterner *et al.*, 1996). Overall, this and other work over that last 20 years support models that invoke splice site pairing across introns and/or exons, depending on size and other sequence features.

Splicing regulation is not limited to structural features of the mRNA transcript, as the spliceosome is associated with 100 or more protein components, many of which contribute to regulation of splicing. A large body of work exists describing the individual role of many of these proteins in alternative splicing. Two of the more well-characterized families of pre-mRNA binding splicing factors are heterogeneous ribonucleoproteins (hnRNPs), which typically bind exon-exclusion motifs (termed repressor or silencer sequences) and generally repress splicing, and Serine-Arginine rich (SR) proteins, which are best-characterized for their recognition and binding to exon inclusion sites (or enhancer sequences) (Busch and Hertel, 2012; Dreyfuss *et al.*, 2002; Manley and Krainer, 2010; Zhou and Fu, 2013). These proteins influence splicing at early stages of spliceosomal assembly. For example, SR proteins bind exon enhancer sequences and mediate the interaction between both U1 snRNA and U2AF (Anko, 2014). These splicing enhancers may also stabilize base pairing between U2 snRNA and the BPS. However, the interaction between SR proteins and pre-mRNA can be blocked by repressor sequences bound by hnRNPs, which often decrease exon inclusion (Zhu *et al.*, 2001; Martinez-Contreras *et al.*, 2007). The inhibitory effect of hnRNP binding to mRNA can also act as a steric deterrent to the core spliceosomal machinery, preventing

binding of core splicing components. In addition to their role in the early stages of spliceosome assembly, SR proteins and hnRNPs may also influence splicing later, such as the association of U6 snRNP with the 5′ splice site (Roca *et al.*, 2013). Although much work has been done to understand the role of these proteins in splicing the complexity of their interactions with each other and with RNA is still being revealed (Huelga *et al.*, 2012).

It is becoming increasingly clear that the activity of most RNA binding proteins that function in splicing is dependent on the location of binding in relationship to splice sites and binding sites for other transacting splicing factors, potentially leading to opposing functions for the same splicing factor at different splice sites (Erkelenz *et al.*, 2013; Zhou and Fu, 2013). The development of methods to identify splicing protein binding sites on RNA *in vivo* has allowed the building of binding or splicing maps for individual splicing proteins in different cell or tissue types and under different cellular conditions (Ule *et al.*, 2003). This information, along with advances in genomic and high-throughput sequencing capabilities has been invaluable as the field challenges itself with the task of understanding the splicing code that governs how splicing and alternative splicing is regulated (Hartmann and Valcarcel, 2009; Nilsen and Graveley, 2010; Wang and Burge, 2008; Barash *et al.*, 2010).

Transcription, Chromatin Structure, and Splicing

The earliest suggested point of splicing regulation is during transcription. The number of introns that are co-transcriptionally spliced is not entirely clear and can vary between species (Bhatt *et al.*, 2012; Iannone and Valcarcel, 2013; Khodor *et al.*, 2012). However, pausing of the RNA polymerase within terminal exons appears to enforce co-transcriptional splicing of the majority of introns (Carrillo Oesterreich *et al.*, 2010). Co-transcriptional splicing can regulate alternative splicing by a number of different mechanisms. First, the rate of transcription can influence competition between alternative splice sites. Slower transcription may make weaker, upstream splice sites more competitive for splicing selection as they will have more time to recruit the spliceosome before a competing splice site is transcribed (Kornblihtt, 2007). Indeed, RNA polymerase II processivity has been shown to correlate strongly with exon inclusion, with weaker splice sites included as a result of decreased transcription rate, although exceptions to this rule highlight the complexity of the interactions between splicing and transcription (Dujardin *et al.*, 2014). Transcription can also affect splice site choice through interactions between splicing and transcription factors. For example, the C-terminal domain (CTD) of RNA polymerase II interacts with a number of splicing proteins and can regulate splicing through differential phosphorylation (Du and Warren, 1997; Hirose *et al.*, 1999; McCracken *et al.*, 1997). Chromatin structure has also been implicated in alternative splicing outcomes as denser, nucleosome-rich chromatin slows transcription speed. In particular, chromatin markers, such as histone methylation, may influence splicing factor recruitment and mRNA transcript diversity, although the extent to which this affects alternative splicing is still unclear (Darnell, 2013; Luco and Misteli, 2011; Schor *et al.*, 2013).

Signaling Pathways

The effect of signal transduction pathways on alternative splicing are not well-understood, although it is likely that such signals are involved in the initiation of many, if not all, condition-specific changes in alternative splicing (Lynch, 2007). Cellular stress, immune responses, cellular differentiation, (Biamonti and Caceres, 2009; Gabut *et al.*, 2011; Martinez and Lynch, 2013; Xie and Black, 2001), and depolarization of ion channels are a few examples of potential signaling pathways that induce specific alternative splicing changes. Although work in this area is in its infancy, posttranslational modifications (PTMs) are known to influence activity, localization, expression, and protein–protein interactions. In particular, SR proteins are especially responsive to kinase activity, and many phosphorylation sites have been identified which cycle during splicing (Zhou and Fu, 2013). HnRNP and SR proteins also undergo a variety of modifications, including phosphorylation, methylation, acetylation, and SUMOylation (Sinha *et al.*, 2010). These various points of regulation suggest the spliceosome is a malleable system, which is responsive to changes in cellular dynamics.

snRNP Biogenesis

Splicing is a demanding process and the production, maturation, and recycling of the snRNPs is likely an essential part of what keeps the reaction occurring at the efficiency required to fulfill the gene expression demands of the cell and to help regulate the changes in splicing and gene expression that occur under different physiological conditions. Because of their central importance to the efficiency of the splicing reaction, the relative efficiency of these processes in different cell-types and tissues may influence splicing and play a role in the regulation of alternative splicing. The genes encoding spliceosomal snRNAs are frequently transcribed adjacent to Cajal bodies followed by export of newly synthesized snRNAs to the cytoplasm (Matera and Wang, 2014; Figure 1(b)). In the cytoplasm the SMN complex chaperones the loading of seven Sm proteins, SmB, D1, D2, D3, E, F, and G, onto an Sm consensus sequence within U1, U2, U4, and U5 snRNAs to form the heptameric Sm ring in the major spliceosome, and their corresponding snRNAs in the minor spliceosome (Fischer *et al.*, 1997; Pellizzoni *et al.*, 2002). Subsequently, snRNAs are hypermethylated at the 5'-end and trimmed at the 3'-end (Mouaikel *et al.*, 2003). Then, immature snRNPs are re-imported into the nucleus where they are initially concentrated in Cajal bodies for the secondary maturation steps. These steps involve the addition of auxiliary protein subunits specific to each particular snRNP and posttranscriptional RNA modifications, methylation, and pseudouridylation (Fischer *et al.*, 2011; Jady *et al.*, 2003). Only U6 snRNP escapes this process, and is loaded into the assembly factor SART3, which brings together U4 and U6 snRNPs primarily in Cajal bodies, prior to the assembly into the U4/U6.U5 tri-snRNP complex (Mroczek and Dziembowski, 2013; Novotny *et al.*, 2011). Mature snRNPs are released from Cajal bodies after dissociation from the SMN complex into the nucleoplasm where they participate in splicing. Thus, the biogenesis of snRNPs requires multistep maturation at cellular locations that are distinct from sites of

function. Cajal bodies are also a site of snRNP recycling after each round of splicing (Fischer *et al.*, 2011; Schaffert *et al.*, 2004; Stanek *et al.*, 2008). It should be noted that Cajal bodies are not present in all cell types (Cioce and Lamond, 2005; Dundr, 2012), but rather these nuclear bodies are present in highly proliferative aneuploid transformed cells such as cancer cells, and absent in primary diploid cells. Exceptions are fetal cells and nondividing, terminally differentiated neurons where Cajal bodies accelerate the processing of spliceosomal components more efficiently than in cells lacking them (Nizami *et al.*, 2010). Taken together, these findings strongly indicate that Cajal bodies actively participate in RNP biogenesis in both early transcriptional and late posttranscriptional nuclear steps of snRNP formation and may play a role in cell-type specific alternative splicing.

Diseases of the Spliceosome: Dysfunction Sheds Light on Function

Aberrant pre-mRNA splicing is the cause of many human diseases (Singh and Cooper, 2012). Mutations in splicing proteins or *cis*-acting splicing regulatory elements can disrupt gene expression by altering the specificity or fidelity of splicing. Mutations in *cis*-acting sequence elements can alter splicing of a single gene transcript, whereas mutations in splicing proteins may affect splicing globally. Indeed, most disease-causing mutations in splicing proteins are autosomal dominant, with the disease arising from a mutation in only one copy of the gene. These heterozygous gene mutations may result in haploinsufficiency or in the creation of a dominant negative form of the protein. The dominant nature of splicing factor mutations is not surprising considering the central role of many of the proteins in the splicing process. Different models have been proposed to explain the disease-specific phenotypes associated with spliceosomal mutations. One model is that loss of gene function reduces splicing efficiency in general, perhaps through differential affinities for various mRNA target sites. A second possibility is that poor splicing of developmental- and tissue-specific transcript(s) makes them more susceptible to deregulation following the mutation of a key regulatory splicing factor. Both models may be correct with the two outcomes operating in concert to produce each condition.

The disease conditions that are associated with mutations of particular splicing proteins offer insight into the requirements of these proteins for normal development and biological function at the organismal level. Interestingly, groups of splicing factors with similar functions have been linked to similar phenotypes and diseases, offering insights into the functional relationship between splicing processes in the cell. For example, germ-line mutations in some splicing factors cause retinitis pigmentosa (Liu and Zack, 2013; Mordes *et al.*, 2006), whereas mutations in another group of splicing factors are associated with developmental deficits including neurological, skeletal, and craniofacial abnormalities. Furthermore, somatic mutations in a number of splicing factors have been associated with hematopoietic malignancies and solid tumors (Padgett, 2012; Scott and Rebel, 2013). The reason for the manifestation of specific pathologies associated with these groups of splicing factors is not clear, though it is clear that the

human condition provides important information about the consequences of aberrant splicing at an organismal level.

Splicing Factor Mutations and Cancer

Somatic mutations in spliceosomal components have been identified in six different types of cancer to date and approximately 85% of different myeloid neoplasms with myelodysplastic features (Papaemmanuil *et al.*, 2011; Yoshida *et al.*, 2011; Makishima *et al.*, 2012), though they may not be the only causative driver of these conditions. These mutations primarily affect either auxiliary components of U2 snRNA or splicing factors involved in early steps in the splicing process during splice site recognition. Specifically, SF3B1, a component of U2 and U12 snRNPs, has been mutated in five different cancerous tissues (Biankin *et al.*, 2012; Cancer Genome Atlas, 2012; Furney *et al.*, 2013; Quesada *et al.*, 2012; Yoshida *et al.*, 2011). Hemizygous somatic mutations in another core component of the spliceosome, PRPF8, have also been identified in myeloid neoplasms and were associated with poor prognosis (Kurtovic-Kozarik *et al.*, 2014). In addition, mutations in U2AF1 and U2AF2, a non-snRNP auxiliary factor homodimer that together are required for U2 to bind pre-mRNA introns, have been linked to myelodysplasia and leukemia, respectively (Graubert *et al.*, 2012; Quesada *et al.*, 2012; Yoshida *et al.*, 2011). Similarly, ZRSR2, which interacts with the U2AF1 and the intronic sequence of the 3'ss, has also been found mutated in myelodysplasia (Shen *et al.*, 2010; Yoshida *et al.*, 2011). Also linked to cancer and involved in early events in splice site recognition and alternative splicing, SRSF1 and SRSF2 are both members of the Serine-Arginine (SR) family of splicing factors, and have been found mutated in leukemia and myeloma, respectively (Quesada *et al.*, 2012; Yoshida *et al.*, 2011). Much work is now being done to identify key splicing events that may be disrupted by mutations in these splicing factors (Brooks *et al.*, 2014; Przychodzen *et al.*, 2013).

Splicing Mutations in Retinitis Pigmentosa and Neurodegeneration

Autosomal dominant retinitis pigmentosa (RP) is a form of blindness caused by degeneration of retinal tissue that commences with either retinal pigmented epithelial cells or photoreceptor neurons. Germ-line mutations in one of 60 different genes can cause RP. Six of these RP-associated genes are linked to structure or function of the U4/U6.U5 tri-snRNP. Three of these, PRPF3, PRPF6, and PRPF31, are all auxiliary factors of the U5 snRNP or U4/U6 di-snRNP (Chakarova *et al.*, 2002; Saini *et al.*, 2012; Utz *et al.*, 2013; Vithana *et al.*, 2001; Xia *et al.*, 2004; McKie *et al.*, 2001; Tanackovic *et al.*, 2011; Grainger and Beggs, 2005; Makarova *et al.*, 2002). Interestingly, somatic mutations in one of the RP-causative genes, PRPF8 leads to myelodysplastic syndromes, as discussed above. It is not clear whether the PRPF8 germ-line mutations that cause RP also predispose individuals to myeloid neoplasms. PRPF8 has been shown to prevent the helicase hBrr2 from unwinding U4/U6 snRNPs prematurely (Mozaffari-Jovin *et al.*, 2013). Mutations in SNRNP200, the gene that encodes hBrr2 have also been shown to cause retinitis pigmentosa

(Liu *et al.*, 2012; Zhao *et al.*, 2009). Finally, RP9 (PAP-1) has been suggested to be a splicing factor, but may act as a bridge between pre-mRNA and U4/U6.U5 tri-snRNP as it has weak associations between both the tri-snRNP and U2AF1 (Keen *et al.*, 2002; Maita *et al.*, 2004).

The vulnerability of certain cell-types to mutations in components of the splicing machinery may be due to the resulting dearth of functional spliceosomes experienced by tissue that has a high amount of mRNA transcripts and therefore a high splicing demand. For example, a human tissue survey showed that retinal tissue has a disproportionately large expression of major spliceosome snRNAs relative to other tissues (Tanackovic *et al.*, 2011). Furthermore, there tends to be a higher number of Cajal bodies in cells with higher metabolic requirement such as neurons (Pena *et al.*, 2001; Spector *et al.*, 1992), which are also affected by mutations in genes encoding splicing-associated proteins. For example, mutations in the gene encoding SMN, involved in snRNP biogenesis, cause the pediatric neurodegenerative disease spinal muscular atrophy (Li *et al.*, 2014).

Splicing Mutations in Growth and Development

Germ-line mutations in a number of splicing factors have been identified as the cause of abnormal developmental and growth phenotypes in humans. These splicing factors function in relatively diverse aspects of the splicing reaction. For example, splicing factors specifically involved in the splicing of U12-type introns have been found to be mutated in conditions leading to growth hormone deficiency. Recessive mutations in the U4atac snRNA have been shown to cause the clinically most severe form of microcephalic osteodysplastic primordial dwarfism type I (MOPD I), also known as Taybi-Linder syndrome (TALS) (Edery *et al.*, 2011; He *et al.*, 2011). This condition is characterized by growth retardation, abnormal bone growth, and short stature, severe brain abnormalities in the cerebral cortex, and a life expectancy, typically, of less than 3 years. A less severe symptomatology is seen in a family with mutations in the RNPC3 gene that encodes the U11/U12-65K protein, an auxiliary protein for U12 snRNP (Argente *et al.*, 2014). The affected family members were each diagnosed with anterior pituitary hypoplasia and isolated growth hormone deficiency, type I, which leads to mild microcephaly and postnatal growth retardation. The 65K protein binds to U12 snRNA and interacts with U11 snRNP, and the mutation disrupts these associations. In these patients, an RNPC3 mutation was observed to result in splicing defects in U12-specific genes, including those related to pituitary function, possibly explaining the growth phenotype.

In addition to the mutations in U4atac and RNPC3 U12 spliceosome-associated proteins, a number of other splicing proteins have been found in individuals with short stature and digit malformations. These include PUF60 (Dauber *et al.*, 2013), and SF3B4 (Bernier *et al.*, 2012), which are associated with U2 snRNP and 3'ss recognition, and EFTUD2 which is a U5 snRNP protein (Bernier *et al.*, 2012; Luquetti *et al.*, 2013; Need *et al.*, 2012).

This partial list of splicing components encoded by genes that, when mutated, cause specific pathologies, can offer valuable information for decoding the splicing code and

understanding splicing regulation. It should be noted, that, here, we have classified the diseases by the most-dominant pathological feature arising from the mutation. However, the diseases associated with mutations in splicing factor genes often have additional phenotypes that may be less severe. This multisystem involvement highlights the central role of splicing and splicing regulation in normal cellular function.

Summary

The spliceosome is comprised of a tightly regulated, yet highly dynamic cohort of proteins and RNA transcripts. These spliceosomal players act in an organized manner, efficiently catalyzing the generation of mature mRNA transcripts. Despite this efficiency, the splicing process can also be selective as demonstrated by the prevalence of alternative splicing. Given the importance of splicing fidelity, efficiency, and regulation for proper gene expression, it is not unexpected that perturbations in core spliceosomal units and their regulatory partners have been found to be associated with various pathological dysfunctions. Intriguingly, mutations in particular splicing factors that lead to common disease phenotypes may suggest mechanistic links between the factors, the investigation of which will serve to inform our understanding of the splicing process and expose underappreciated avenues for biomedical research and subsequent innovations in clinical intervention.

Acknowledgments

This work was funded by NIH grants R01 DC012596 (M.L.H.), R01 NS069759 (M.L.H.), and R01 GM090156 (M.D.).

See also: Nucleic Acid Synthesis/Breakdown: Transcription: The Spliceosome and Pre-mRNA Splicing

References

- Abelson, J., Blanco, M., Ditzler, M.A., *et al.*, 2010. Conformational dynamics of single pre-mRNA molecules during *in vitro* splicing. *Nature Structural & Molecular Biology* 17, 504–512.
- Anko, M.L., 2014. Regulation of gene expression programmes by serine-arginine rich splicing factors. *Seminars in Cell & Developmental Biology* 32C, 11–21.
- Argente, J., Flores, R., Gutierrez-Arumi, A., *et al.*, 2014. Defective minor spliceosome mRNA processing results in isolated familial growth hormone deficiency. *EMBO Molecular Medicine* 6, 299–306.
- Barash, Y., Calarco, J.A., Gao, W., *et al.*, 2010. Deciphering the splicing code. *Nature* 465, 53–59.
- Bergert, S.M., Moore, C., Sharp, P.A., 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proceedings of the National Academy of Sciences of the United States of America* 74, 3171–3175.
- Bernier, F.P., Caluseriu, O., Ng, S., *et al.*, 2012. Haploinsufficiency of SF3B4, a component of the pre-mRNA spliceosomal complex, causes Nager syndrome. *American Journal of Human Genetics* 90, 925–933.
- Bhatt, D.M., Pandya-Jones, A., Tong, A.J., *et al.*, 2012. Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. *Cell* 150, 279–290.
- Biamonti, G., Caceres, J.F., 2009. Cellular stress and RNA splicing. *Trends in Biochemical Sciences* 34, 146–153.
- Biankin, A.V., Waddell, N., Kassahn, K.S., *et al.*, 2012. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature* 491, 399–405.
- Black, D.L., 1991. Does steric interference between splice sites block the splicing of a short c-src neuron-specific exon in non-neuronal cells? *Genes & Development* 5, 389–402.
- Bolisetty, M.T., Beemon, K.L., 2012. Splicing of internal large exons is defined by novel cis-acting sequence elements. *Nucleic Acids Research* 40, 9244–9254.
- Braunschweig, U., Guerousov, S., Plocik, A.M., Graveley, B.R., Blencowe, B.J., 2013. Dynamic integration of splicing within gene regulatory pathways. *Cell* 152, 1252–1269.
- Brooks, A.N., Choi, P.S., de Waal, L., *et al.*, 2014. A pan-cancer analysis of transcriptome changes associated with somatic mutations in U2AF1 reveals commonly altered splicing events. *PLoS One* 9, e87361.
- Busch, A., Hertel, K.J., 2012. Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdisciplinary Reviews: RNA* 3, 1–12.
- Cancer Genome Atlas, N., 2012. Comprehensive molecular portraits of human breast tumours. *Nature* 490, 61–70.
- Carrillo Oesterreich, F., Preibisch, S., Neugebauer, K.M., 2010. Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Molecular Cell* 40, 571–581.
- Chakarova, C.F., Hims, M.M., Bolz, H., *et al.*, 2002. Mutations in HPRP3, a third member of pre-mRNA splicing factor genes, implicated in autosomal dominant retinitis pigmentosa. *Human Molecular Genetics* 11, 87–92.
- Chow, L.T., Gelinas, R.E., Broker, T.R., Roberts, R.J., 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* 12, 1–8.
- Cioce, M., Lamond, A.I., 2005. Cajal bodies: A long history of discovery. *Annual Review of Cell and Developmental Biology* 21, 105–131.
- Darnell Jr., J.E., 2013. Reflections on the history of pre-mRNA processing and highlights of current knowledge: A unified picture. *RNA* 19, 443–460.
- Dauber, A., Golzio, C., Guenot, C., *et al.*, 2013. SCRIB and PUF60 are primary drivers of the multisystemic phenotypes of the 8q24.3 copy-number variant. *American Journal of Human Genetics* 93, 798–811.
- De Conti, L., Baralle, M., Buratti, E., 2013. Exon and intron definition in pre-mRNA splicing. *Wiley Interdisciplinary Reviews: RNA* 4, 49–60.
- Dominski, Z., Kole, R., 1991. Selection of splice sites in pre-mRNAs with short internal exons. *Molecular and Cellular Biology* 11 (12), 6075–6083.
- Dreyfuss, G., Kim, V.N., Kataoka, N., 2002. Messenger-RNA-binding proteins and the messages they carry. *Nature Reviews Molecular Cell Biology* 3, 195–205.
- Du, L., Warren, S.L., 1997. A functional interaction between the carboxy-terminal domain of RNA polymerase II and pre-mRNA splicing. *Journal of Cell Biology* 136, 5–18.
- Dujardin, G., Lafaille, C., de la Mata, M., *et al.*, 2014. How slow RNA polymerase II elongation favors alternative exon skipping. *Molecular Cell* 54, 683–690.
- Dundr, M., 2012. Nuclear bodies: Multifunctional companions of the genome. *Current Opinion in Cell Biology* 24, 415–422.
- Ederly, P., Marcaillou, C., Sahbatou, M., *et al.*, 2011. Association of TALS developmental disorder with defect in minor splicing component U4atac snRNA. *Science* 332, 240–243.
- Erkelenz, S., Mueller, W.F., Evans, M.S., *et al.*, 2013. Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms. *RNA* 19, 96–102.
- Fischer, U., Englbrecht, C., Chari, A., 2011. Biogenesis of spliceosomal small nuclear ribonucleoproteins. *Wiley Interdisciplinary Reviews: RNA* 2, 718–731.
- Fischer, U., Liu, Q., Dreyfuss, G., 1997. The SMN-SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* 90, 1023–1029.
- Frilander, M.J., Steitz, J.A., 1999. Initial recognition of U12-dependent introns requires both U11/5' splice-site and U12/branchpoint interactions. *Genes & Development* 13, 851–863.
- Furney, S.J., Pedersen, M., Gentien, D., *et al.*, 2013. SF3B1 mutations are associated with alternative splicing in uveal melanoma. *Cancer Discovery* 3, 1122–1129.
- Gabut, M., Samavarchi-Tehrani, P., Wang, X., *et al.*, 2011. An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. *Cell* 147, 132–146.
- Grainger, R.J., Beggs, J.D., 2005. Prp8 protein: At the heart of the spliceosome. *RNA* 11, 533–557.
- Grauert, T.A., Shen, D., Ding, L., *et al.*, 2012. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nature Genetics* 44, 53–57.
- Hall, S.L., Padgett, R.A., 1996. Requirement of U12 snRNA for *in vivo* splicing of a minor class of eukaryotic nuclear pre-mRNA introns. *Science* 271, 1716–1718.
- Hartmann, B., Valcarcel, J., 2009. Decrypting the genome's alternative messages. *Current Opinion in Cell Biology* 21, 377–386.

- Hastings, M.L., Allemand, E., Duelli, D.M., Myers, M.P., Krainer, A.R., 2007. Control of pre-mRNA splicing by the general splicing factors PUF60 and U2AF(65). *PLoS One* 2, e538.
- Hastings, M.L., Krainer, A.R., 2001a. Functions of SR proteins in the U12-dependent AT-AC pre-mRNA splicing pathway. *RNA* 7, 471–482.
- Hastings, M.L., Krainer, A.R., 2001b. Pre-mRNA splicing in the new millennium. *Current Opinion in Cell Biology* 13, 302–309.
- He, H., Liyanarachchi, S., Akagi, K., *et al.*, 2011. Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental disorder MOPD I. *Science* 332, 238–240.
- Hesselberth, J.R., 2013. Lives that introns lead after splicing. *Wiley Interdisciplinary Reviews: RNA* 4, 677–691.
- Hirose, Y., Tacke, R., Manley, J.L., 1999. Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes & Development* 13, 1234–1239.
- Horowitz, D.S., 2012. The mechanism of the second step of pre-mRNA splicing. *Wiley Interdisciplinary Reviews: RNA* 3, 331–350.
- Horowitz, D.S., Abelson, J., 1993. Stages in the second reaction of pre-mRNA splicing: The final step is ATP independent. *Genes & Development* 7, 320–329.
- Hoskins, A.A., Friedman, L.J., Gallagher, S.S., *et al.*, 2011. Ordered and dynamic assembly of single spliceosomes. *Science* 331, 1289–1295.
- Hsu, S.N., Hertel, K.J., 2009. Spliceosomes walk the line: Splicing errors and their impact on cellular function. *RNA Biology* 6, 526–530.
- Huelga, S.C., Vu, A.Q., Arnold, J.D., *et al.*, 2012. Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Reports* 1, 167–178.
- Iannone, C., Valcarcel, J., 2013. Chromatin's thread to alternative splicing regulation. *Chromosoma* 122, 465–474.
- Jady, B.E., Darzacq, X., Tucker, K.E., *et al.*, 2003. Modification of Sm small nuclear RNAs occurs in the nucleoplasmic Cajal body following import from the cytoplasm. *EMBO Journal* 22, 1878–1888.
- Jurica, M.S., Moore, M.J., 2003. Pre-mRNA splicing: Awash in a sea of proteins. *Molecular Cell* 12, 5–14.
- Kandels-Lewis, S., Seraphin, B., 1993. Involvement of U6 snRNA in 5' splice site selection. *Science* 262, 2035–2039.
- Keen, T.J., Hims, M.M., McKie, A.B., *et al.*, 2002. Mutations in a protein target of the Pim-1 kinase associated with the RP9 form of autosomal dominant retinitis pigmentosa. *European Journal of Human Genetics: EJHG* 10, 245–249.
- Kelemen, O., Convertini, P., Zhang, Z., *et al.*, 2013. Function of alternative splicing. *Gene* 514, 1–30.
- Khodor, Y.L., Menet, J.S., Tolan, M., Rosbash, M., 2012. Cotranscriptional splicing efficiency differs dramatically between *Drosophila* and mouse. *RNA* 18, 2174–2186.
- Konarska, M.M., Sharp, P.A., 1987. Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell* 49, 763–774.
- Kornblihtt, A.R., 2007. Coupling transcription and alternative splicing. *Advances in Experimental Medicine and Biology* 623, 175–189.
- Kurtovic-Kozaric, A., Przychodzen, B., Singh, J., *et al.*, 2014. PRPF8 defects cause missplicing in myeloid malignancies. *Leukemia*. doi:10.1038/leu.2014.144.
- Levine, A., Durbin, R., 2001. A computational scan for U12-dependent introns in the human genome sequence. *Nucleic Acids Research* 29, 4006–4013.
- Li, D.K., Tisdale, S., Lotti, F., Pellizzoni, L., 2014. SMN control of RNP assembly: From post-transcriptional gene regulation to motor neuron disease. *Seminars in Cell & Developmental Biology* 32C, 22–29.
- Liu, M.M., Zack, D.J., 2013. Alternative splicing and retinal degeneration. *Clinical Genetics* 84, 142–149.
- Liu, T., Jin, X., Zhang, X., *et al.*, 2012. A novel missense SNRNP200 mutation associated with autosomal dominant retinitis pigmentosa in a Chinese family. *PLoS One* 7, e45464.
- Lopez-Bigas, N., Audit, B., Ouzounis, C., Parra, G., Guigo, R., 2005. Are splicing mutations the most frequent cause of hereditary disease? *FEBS Letters* 579, 1900–1903.
- Luco, R.F., Misteli, T., 2011. More than a splicing code: Integrating the role of RNA, chromatin and non-coding RNA in alternative splicing regulation. *Current Opinion in Genetics & Development* 21, 366–372.
- Luo, H.R., Moreau, G.A., Levin, N., Moore, M.J., 1999. The human Prp8 protein is a component of both U2- and U12-dependent spliceosomes. *RNA* 5, 893–908.
- Luquetti, D.V., Hing, A.V., Rieder, M.J., *et al.*, 2013. "Mandibulofacial dysostosis with microcephaly" caused by EFTUD2 mutations: Expanding the phenotype. *American Journal of Medical Genetics Part A* 161A, 108–113.
- Lynch, K.W., 2007. Regulation of alternative splicing by signal transduction pathways. *Advances in Experimental Medicine and Biology* 623, 161–174.
- Maita, H., Kitaura, H., Keen, T.J., *et al.*, 2004. PAP-1, the mutated gene underlying the RP9 form of dominant retinitis pigmentosa, is a splicing factor. *Experimental Cell Research* 300, 283–296.
- Makarova, O.V., Makarov, E.M., Liu, S., Vornlocher, H.P., Luhrmann, R., 2002. Protein 61K, encoded by a gene (PRPF31) linked to autosomal dominant retinitis pigmentosa, is required for U4/U6*U5 tri-snRNP formation and pre-mRNA splicing. *EMBO Journal* 21, 1148–1157.
- Makishima, H., Visconte, V., Sakaguchi, H., *et al.*, 2012. Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. *Blood* 119, 3203–3210.
- Manley, J.L., Krainer, A.R., 2010. A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). *Genes & Development* 24, 1073–1074.
- Martinez, N.M., Lynch, K.W., 2013. Control of alternative splicing in immune responses: Many regulators, many predictions, much still to learn. *Immunological Reviews* 253, 216–236.
- Martinez-Contreras, R., Cloutier, P., Shkreta, L., *et al.*, 2007. hnRNP proteins and splicing control. *Advances in Experimental Medicine and Biology* 623, 123–147.
- Matera, A.G., Wang, Z., 2014. A day in the life of the spliceosome. *Nature Reviews Molecular Cell Biology* 15, 108–121.
- McCracken, S., Fong, N., Yankulov, K., *et al.*, 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature* 385, 357–361.
- McKie, A.B., McHale, J.C., Keen, T.J., *et al.*, 2001. Mutations in the pre-mRNA splicing factor gene PRPC8 in autosomal dominant retinitis pigmentosa (RP13). *Human Molecular Genetics* 10, 1555–1562.
- McManus, C.J., Graveley, B.R., 2011. RNA structure and the mechanisms of alternative splicing. *Current Opinion in Genetics & Development* 21, 373–379.
- Mordes, D., Luo, X., Kar, A., *et al.*, 2006. Pre-mRNA splicing and retinitis pigmentosa. *Molecular Vision* 12, 1259–1271.
- Mouaikel, J., Narayanan, U., Verheggen, C., *et al.*, 2003. Interaction between the small-nuclear-RNA cap hypermethylase and the spinal muscular atrophy protein, survival of motor neuron. *EMBO Reports* 4, 616–622.
- Mount, S.M., Pettersson, I., Hinterberger, M., Karmas, A., Steitz, J.A., 1983. The U1 small nuclear RNA-protein complex selectively binds a 5' splice site *in vitro*. *Cell* 33, 509–518.
- Mozaffari-Jovin, S., Wandersleben, T., Santos, K.F., *et al.*, 2013. Inhibition of RNA helicase Brr2 by the C-terminal tail of the spliceosomal protein Prp8. *Science* 341, 80–84.
- Mroczek, S., Dziembowski, A., 2013. U6 RNA biogenesis and disease association. *Wiley Interdisciplinary Reviews: RNA* 4, 581–592.
- Need, A.C., Shashi, V., Hitomi, Y., *et al.*, 2012. Clinical application of exome sequencing in undiagnosed genetic conditions. *Journal of Medical Genetics* 49, 353–361.
- Newman, A., Norman, C., 1991. Mutations in yeast U5 snRNA alter the specificity of 5' splice-site cleavage. *Cell* 65, 115–123.
- Nielsen, T.W., Graveley, B.R., 2010. Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463, 457–463.
- Nizami, Z.F., Deryusheva, S., Gall, J.G., 2010. Cajal bodies and histone locus bodies in *Drosophila* and *Xenopus*. *Cold Spring Harbor Symposia on Quantitative Biology* 75, 313–320.
- Novotny, I., Blazikova, M., Stanek, D., Herman, P., Malinsky, J., 2011. *In vivo* kinetics of U4/U6.U5 tri-snRNP formation in Cajal bodies. *Molecular Biology of the Cell* 22, 513–523.
- Padgett, R.A., 2012. New connections between splicing and human disease. *Trends in Genetics* 28, 147–154.
- Pan, Q., Shai, O., Lee, L.J., Frey, B.J., Blencowe, B.J., 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics* 40, 1413–1415.
- Papaemmanuil, E., Cazzola, M., Boulwood, J., *et al.*, 2011. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *New England Journal of Medicine* 365, 1384–1395.
- Patel, A.A., Steitz, J.A., 2003. Splicing double: Insights from the second spliceosome. *Nature Reviews Molecular Cell Biology* 4, 960–970.
- Pellizzoni, L., Yong, J., Dreyfuss, G., 2002. Essential role for the SMN complex in the specificity of snRNP assembly. *Science* 298, 1775–1779.
- Pena, E., Berciano, M.T., Fernandez, R., Ojeda, J.L., Lafarga, M., 2001. Neuronal body size correlates with the number of nucleoli and Cajal bodies, and with the organization of the splicing machinery in rat trigeminal ganglion neurons. *Journal of Comparative Neurology* 430, 250–263.
- Przychodzen, B., Jerez, A., Guinta, K., *et al.*, 2013. Patterns of missplicing due to somatic mutations in myeloid neoplasms. *Blood* 122, 999–1006.
- Query, C.C., Strobel, S.A., Sharp, P.A., 1996. Three recognition events at the branch-site adenine. *EMBO Journal* 15, 1392–1402.

- Quesada, V., Conde, L., Villamor, N., *et al.*, 2012. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nature Genetics* 44, 47–52.
- Reed, R., Maniatis, T., 1988. The role of the mammalian branchpoint sequence in pre-mRNA splicing. *Genes & Development* 2, 1268–1276.
- Robberson, B.L., Cote, G.J., Berget, S.M., 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. *Molecular and Cellular Biology* 10, 84–94.
- Roca, X., Krainer, A.R., Eperon, I.C., 2013. Pick one, but be quick: 5' Splice sites and the problems of too many choices. *Genes & Development* 27, 129–144.
- Roscigno, R.F., Garcia-Blanco, M.A., 1995. SR proteins escort the U4/U6.U5 tri-snRNP to the spliceosome. *RNA* 1, 692–706.
- Saini, S., Robinson, P.N., Singh, J.R., Vanita, V., 2012. A novel 7 bp deletion in PRPF31 associated with autosomal dominant retinitis pigmentosa with incomplete penetrance in an Indian family. *Experimental Eye Research* 104, 82–88.
- Schaffert, N., Hossbach, M., Heintzmann, R., Achsel, T., Luhrmann, R., 2004. RNAi knockdown of hPrp31 leads to an accumulation of U4/U6 di-snRNPs in Cajal bodies. *EMBO Journal* 23, 3000–3009.
- Schor, I.E., Gomez Acuna, L.I., Kornblihtt, A.R., 2013. Coupling between transcription and alternative splicing. *Cancer Treatment and Research* 158, 1–24.
- Scott, L.M., Rebel, V.I., 2013. Acquired mutations that affect pre-mRNA splicing in hematologic malignancies and solid tumors. *Journal of the National Cancer Institute* 105, 1540–1549.
- Semlow, D.R., Staley, J.P., 2012. Staying on message: Ensuring fidelity in pre-mRNA splicing. *Trends in Biochemical Sciences* 37, 263–273.
- Shen, H., Zheng, X., Luecke, S., Green, M.R., 2010. The U2AF35-related protein Urp contacts the 3' splice site to promote U12-type intron splicing and the second step of U2-type intron splicing. *Genes & Development* 24, 2389–2394.
- Sheth, N., Roca, X., Hastings, M.L., *et al.*, 2006. Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Research* 34, 3955–3967.
- Singh, R.K., Cooper, T.A., 2012. Pre-mRNA splicing in disease and therapeutics. *Trends in Molecular Medicine* 18, 472–482.
- Sinha, R., Allemand, E., Zhang, Z., *et al.*, 2010. Arginine methylation controls the subcellular localization and functions of the oncoprotein splicing factor SF2/ASF. *Molecular and Cellular Biology* 30, 2762–2774.
- Spector, D.L., Lark, G., Huang, S., 1992. Differences in snRNP localization between transformed and nontransformed cells. *Molecular Biology of the Cell* 3, 555–569.
- Stanek, D., Pridalova-Hnilicova, J., Novotny, I., *et al.*, 2008. Spliceosomal small nuclear ribonucleoprotein particles repeatedly cycle through Cajal bodies. *Molecular Biology of the Cell* 19, 2534–2543.
- Stephens, R.M., Schneider, T.D., 1992. Features of spliceosome evolution and function inferred from an analysis of the information at human splice sites. *Journal of Molecular Biology* 228, 1124–1136.
- Stern, D.A., Carlo, T., Berget, S.M., 1996. Architectural limits on split genes. *Proceedings of the National Academy of Sciences of the United States of America* 93 (26), 15081–15085.
- Sugnet, C.W., Kent, W.J., Ares Jr., M., Haussler, D., 2004. Transcriptome and genome conservation of alternative splicing events in humans and mice. *Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing* 9, 66–77.
- Tarn, W.Y., Steitz, J.A., 1996. A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron *in vitro*. *Cell* 84, 801–811.
- Tanackovic, G., Ransijn, A., Ayuso, C., *et al.*, 2011. A missense mutation in PRPF6 causes impairment of pre-mRNA splicing and autosomal-dominant retinitis pigmentosa. *American Journal of Human Genetics* 88, 643–649.
- Turner, I.A., Norman, C.M., Churcher, M.J., Newman, A.J., 2004. Roles of the U5 snRNP in spliceosome dynamics and catalysis. *Biochemical Society Transactions* 32, 928–931.
- Turunen, J.J., Niemela, E.H., Verma, B., Frilander, M.J., 2013. The significant other: Splicing by the minor spliceosome. *Wiley Interdisciplinary Reviews: RNA* 4, 61–76.
- Turunen, J.J., Will, C.L., Grote, M., Luhrmann, R., Frilander, M.J., 2008. The U11-48K protein contacts the 5' splice site of U12-type introns and the U11-59K protein. *Molecular and Cellular Biology* 28, 3548–3560.
- Ule, J., Jensen, K.B., Ruggiu, M., *et al.*, 2003. CLIP identifies Nova-regulated RNA networks in the brain. *Science* 302, 1212–1215.
- Utz, V.M., Beight, C.D., Marino, M.J., Hagstrom, S.A., Traboulsi, E.I., 2013. Autosomal dominant retinitis pigmentosa secondary to pre-mRNA splicing-factor gene PRPF31 (RP11): Review of disease mechanism and report of a family with a novel 3-base pair insertion. *Ophthalmic Genetics* 34, 183–188.
- Vithana, E.N., Abu-Safieh, L., Allen, M.J., *et al.*, 2001. A human homolog of yeast pre-mRNA splicing gene, PRP31, underlies autosomal dominant retinitis pigmentosa on chromosome 19q13.4 (RP11). *Molecular Cell* 8, 375–381.
- Wang, E.T., Sandberg, R., Luo, S., *et al.*, 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476.
- Wang, Z., Burge, C.B., 2008. Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *RNA* 14, 802–813.
- Will, C.L., Luhrmann, R., 2005. Splicing of a rare class of introns by the U12-dependent spliceosome. *Biological Chemistry* 386, 713–724.
- Will, C.L., Luhrmann, R., 2011. Spliceosome structure and function. *Cold Spring Harbor Perspectives in Biology* 3, a003707.
- Will, C.L., Schneider, C., Reed, R., Luhrmann, R., 1999. Identification of both shared and distinct proteins in the major and minor spliceosomes. *Science* 284, 2003–2005.
- Xia, K., Zheng, D., Pan, Q., *et al.*, 2004. A novel PRPF31 splice-site mutation in a Chinese family with autosomal dominant retinitis pigmentosa. *Molecular Vision* 10, 361–365.
- Xie, J., Black, D.L., 2001. A CaMK IV responsive RNA element mediates depolarization-induced alternative splicing of ion channels. *Nature* 410, 936–939.
- Yoshida, K., Sanada, M., Shiraishi, Y., *et al.*, 2011. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 478, 64–69.
- Zhang, C., Hastings, M.L., Krainer, A.R., Zhang, M.Q., 2007. Dual-specificity splice sites function alternatively as 5' and 3' splice sites. *Proceedings of the National Academy of Sciences of the United States of America* 104, 15028–15033.
- Zhang, J., Manley, J.L., 2013. Misregulation of pre-mRNA alternative splicing in cancer. *Cancer Discovery* 3, 1228–1237.
- Zhao, C., Bellur, D.L., Lu, S., *et al.*, 2009. Autosomal-dominant retinitis pigmentosa caused by a mutation in SNRNP200, a gene required for unwinding of U4/U6 snRNAs. *American Journal of Human Genetics* 85, 617–627.
- Zhou, Z., Fu, X.D., 2013. Regulation of splicing by SR proteins and SR protein-specific kinases. *Chromosoma* 122, 191–207.
- Zhu, J., Mayeda, A., Krainer, A.R., 2001. Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Molecular Cell* 8, 1351–1361.

NUCLEIC ACID SYNTHESIS/BREAKDOWN: NUCLEIC ACID TECHNOLOGY

Contents

Transgenesis and Gene Replacement
Viral Nucleic Acids

Transgenesis and Gene Replacement

MJ Wolfgang, Johns Hopkins University School of Medicine, Baltimore, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Transgenic and gene replacement strategies function to link basic biological mechanisms to organismic biology. They enable the transfer of discoveries made *in vitro* to intact organisms to provide insight into the relevance of genes and pathways in a proper context. Cell culture systems or simple unicellular organisms often do not recapitulate gene function of multicellular organisms given the interactions between cells and tissues. The role of genes within an organism is often more complex and nuanced than can be modeled reliably *in vitro*. Due to their small size and relative ease of genetic manipulation, the mouse has become the preferred model to dissect gene function and regulation related to human biology. Although much work has been devoted to mouse models, many of these technologies can and have been adapted to other organisms with unique features. For example, pigs have been genetically modified as a first step in enabling organ transplantation (Prather *et al.*, 2013). Other biological questions related specifically to humans can only be modeled in non-human primates, which have also been successfully engineered using techniques pioneered in lower organisms (Wolfgang and Golos, 2002; Sasaki *et al.*, 2009).

Genetic engineering in animals requires reliable reproductive technologies. Introducing genetic material or manipulating genes that can be transmitted through progeny requires the culture of germ cells or stem cells in order to facilitate genome engineering. Genetic material can be introduced into fertilized zygotes. Shortly after fertilization, both maternal and paternal pronuclei (encoding the haploid genomes) are prominent and easily visualized in most species. They can be directly injected before fusion (Figure 1(a)). This occurs well before germline differentiation therefore germline transmission is usually robust. Simply implanting these manipulated embryos back into surrogate females can result in genetically engineered animals even in the F0 generation. The other critical advance in reproductive technologies is the culture of embryonic stem cells.

These cells, derived from the inner cell mass of preimplantation blastocysts, can give rise to most embryonic structures including the germplasm. The specialized culture of these cells enables the manipulation and selection of genetically altered cells *in vitro*.

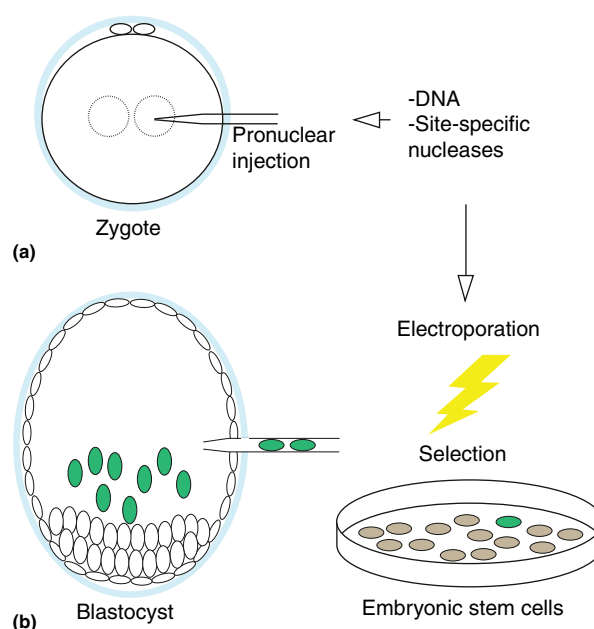


Figure 1 Genetic manipulation of mammals. (a) Foreign DNA can be introduced by directly injecting genetic material into the pronuclei of fertilized zygotes. These can then be transplanted back into surrogate females to generate transgenic offspring. (b) DNA can be introduced into ES cells by electroporation. Selected ES cell clones (green colonies) can be injected directly into the blastocoel of blastocyst stage embryos. Transplantation of the resultant embryos to surrogate females can result in the production of chimeric offsprings. Breeding of chimeric animals can yield ES cell-derived offsprings if the ES cells contributed to the germplasm.

The addition of these cells back to recipient blastocytes generates chimeras. Breeding of the chimeras can result in an ES cell-derived offspring given the incorporation of the ES cells into the germline (Figure 1(b)). These reproductive technologies have enabled the facile manipulation of genomes to determine gene function and regulation. The landscape of genetic manipulation is vast and seems to be inhibited only by the imagination of the investigator. There are seemingly endless variations to these basic themes that are easily adapted to answer specific questions. Here, an overview of the most widely used and informative techniques will be discussed mainly in the context of mouse models.

Transgenesis

Foreign DNA can be integrated into the germline to facilitate the understanding of gene regulation or function. Transgenics can be made in a rather straightforward manner by injecting DNA (usually linearized plasmid DNA) into the pronuclei of mammalian zygotes. This facilitates random integration of DNA into the genome that can be stably propagated to successive generations through the germline. Gene regulation can be deciphered in this manner by linking putative regulatory sequences of a gene of interest to a reporter gene (Zimmerman *et al.*, 1994). There are many convenient reporter genes available to monitor gene regulation in a cell-specific manner in both fixed tissue and live mice. β -Galactosidase activity, encoded by LacZ, can be used to understand tissue-specific gene expression even in harshly fixed tissue. This strategy has been used widely to understand gene regulation in transgenics due to its stability and convenience. To understand gene regulation in living cells, fluorescent proteins such as green fluorescent protein (GFP) can be used to directly visualize fluorescent cells *ex vivo* or via intravital imaging *in vivo* (Weigert *et al.*, 2013; Ellenbroek and van Rheen, 2014). This can enable rather sophisticated analysis of gene expression in living cells. Most fluorescent proteins have a long half-life, however, short-lived or photoconvertible fluorescent proteins can be used to obtain a more precise timing of expression. Additionally, noninvasive means of monitoring gene expression in living mice can be accommodated by using luciferase-based reporter genes (Wilsbacher *et al.*, 2002). Luciferase bioluminescence can be imaged directly in live mice noninvasively over a long time-scale. Even secreted forms of luciferase (e.g., Gaussia princeps luciferase) can be used to understand gene regulation by minimally invasive techniques such as repeated blood sampling (Wurdinger *et al.*, 2008).

The transcriptional regulation of genes sometimes goes well beyond what can be accommodated within a simple plasmid. For genes that have critical regulatory elements encoded far away from their coding sequences, one can turn to bacterial artificial chromosomes (BACs) to produce transgenics. BACs often recapitulate gene expression more precisely as they contain hundreds of kilobases of genomic DNA surrounding the coding sequence. BACs, although more laborious to work with than simple plasmids, can be recombineered by standard techniques in bacteria to encode reporter genes within the coding sequence (Gong *et al.*, 2010). This approach has been successfully adapted to large-scale gene mapping projects to

mark gene expression patterns *in vivo* (Gong *et al.*, 2003). One drawback of BACs besides their unwieldy size is that they often contain other genes linked to the gene of interest, which has the potential to unintentionally modify phenotypes.

Going beyond marking cells of interest or unraveling the regulation of genes *in vivo*, understanding gene function using transgenics is also of great interest. Over-expression of genes or expressing constitutively active or dominant negative genes aids in unraveling gene functions *in vivo* (Moitra *et al.*, 1998). Additionally, early work defining cell- and tissue-specific regulatory elements have made tissue-specific expression of transgenes routine. Therefore, one can define the role of genes in a tissue-specific manner by expressing transgenes *in vivo*. Inclusion of a synthetic intron or an insulator sequence is an important consideration that can greatly enhance the expression of transgenes.

Some of the earliest studies facilitated transgenesis using simple retroviruses like Moloney murine leukemia virus (MMLV) to integrate new genetic material into mouse embryos and subsequently transmitted to progeny through the germline (Jaenisch, 1976). Unfortunately, simple retroviruses are transcriptionally suppressed under most circumstances, limiting their usefulness in promoting transgenesis. Circumventing this transcriptional suppression is the use of lentiviruses that can support the expression of genes in mammals (Pfeifer *et al.*, 2002; Lois *et al.*, 2002). Injecting DNA into pronuclei often generates concatemeric random insertion that can effect transgene expression and function depending on the local genetic landscape. One advantage of lentivirus-mediated transgenesis is that single transgene copies are inserted randomly throughout the genome therefore permitting single transgene integration. These random integration methods can sometimes disrupt endogenous gene function and can be heavily influenced by their genomic environment. These issues can be greatly mitigated by introducing transgenes into predetermined transcriptionally permissive integration sites such as the R26 or H11 loci. Several mouse lines have been produced to facilitate entry of exogenously provided DNA into these loci (Tasic *et al.*, 2011).

Temporal Regulation of Transgenes

It is often not possible to regulate the temporal expression of transgenes using tissue-specific promoters. Some promoters can be regulated by exogenous factors (e.g., interferon or metal ions), however, these are often hindered by significant off-target effects. Steroid-responsive promoters that are engineered to respond to unique steroids such as the synthetic progesterone receptor antagonist mifepristone or the insect steroid ecdysone can be used to facilitate temporal expression (Wang *et al.*, 1997). These methods are robust and specific temporal regulators but are somewhat cost prohibitive for *in vivo* experimentation. A more popular and versatile means of regulating mammalian transgene expression is the use of the bacterial tetracycline-responsive operon. Tetracycline or its analogue doxycycline binds to and regulates the tetracycline repressor (TetR). The TetR has been modified to allow both positive (Tet-On) and negative (Tet-Off) transgene expression in response to doxycycline. Tet-Off works by activating

transgene expression in the absence of doxycycline, whereas Tet-On activates transgene expression in the presence of doxycycline. The Tet system has been widely used to regulate gene expression in transgenic animals. Doxycycline crosses the blood–brain barrier and can be given simply by adding it to drinking water (Chen *et al.*, 1998). This allows versatile regulation of transgenes in many settings.

In addition to regulating the transcriptional control of transgenes, transgenes can be regulated via posttranslational means as well. Introduction of fusion proteins that can be regulated by small molecules can control the localization of transgenes. Addition of a steroid receptor to a transcription factor or to Cre recombinase can sequester the protein within the cytoplasm. Providing the ligand can enable the translocation to the target and access to the substrate (Feil *et al.*, 1996; Kohn *et al.*, 1998; Metzger *et al.*, 1995). Additionally, fusion of proteins with small molecule-regulated destabilizing domains or inducible dimerization domains can mediate temporal regulation of transgene function (Rodriguez *et al.*, 2014; Rodriguez and Wolfgang, 2012; Banaszynski *et al.*, 2006; Rakhit *et al.*, 2014). Some transgenes can even be activated by light. These techniques can provide not only tissue-specific expression but also temporal-specific expression or function.

Loss of Function

A basic reductionist approach to determining gene function *in vivo* is the generation of knockout animals. Loss-of-function strategies can greatly inform the role of a gene in the context of an organism by deciphering the gene's requirements for a phenotype of interest. Many investigations into lower organisms make use of forward genetic screens using randomly induced mutagenesis to assign gene function to an assayed phenotype. Although this strategy has been used in mammals, the cost and scale of genetic screens is largely prohibitive. Understanding mammalian gene function is usually done in one or a few genes at a time via gene-targeting strategies of well-annotated genes. Although laborious, the results of these reverse genetic strategies are usually enormously enlightening, focus on further experimentation, and form the foundation for much of modern biomedical sciences.

Gene Trap

One means of determining gene requirements is the use of gene traps to both mark and inactivate genes (Skarnes *et al.*, 2004). Gene traps utilize simple retroviruses containing a splice acceptor linked to a reporter gene (e.g., LacZ or GFP) and a selectable cassette with a polyadenylation sequence to short circuit gene expression. Self-inactivating MMLV vectors can be used to limit the transcriptional inhibition inherent in these vectors, therefore mitigating some of their limitations. MMLV retroviruses pseudo-randomly integrate into the genome by preferentially integrating into large permissive introns making them ideal for gene trapping in ES cells. In this way, clonal heterozygous ES cell traps can be isolated and selected for the production of chimeras and subsequent breeding for mice with homozygous gene disruptions. Additionally, if a

reporter gene is placed into the vector downstream of the splice acceptor then one can determine the endogenous gene expression pattern of the trapped gene by assaying for the reporter gene in the heterozygous or homozygous offspring (Zambrowicz *et al.*, 1997). One of the major limitations to this method is that alternative splicing, which occurs widely throughout the genome, can often mask true null phenotypes. More sophisticated trapping methods can be used to understand, for example, the regulation of transcriptional targets of critical signaling pathways and subsequently the requirements for these targets *in vivo* (Chen *et al.*, 2004). Gene trapping methods have been used extensively to understand gene requirements *in vivo* but have been somewhat outpaced by more stringent gene-replacement technologies.

Homologous Recombination

To generate a more reliable means of reverse engineering mammalian animal models, investigators have employed homologous recombination in ES cells to ensure correctly targeted alleles. This is done routinely in several inbred strains of mice. In species where germline competent ES cells have not been routinely used, homologous recombination can be performed in somatic cells followed by nuclear transfer (i.e., cloning) (Dai *et al.*, 2002). Homologous recombination in ES cells is by far the most popular method owing to the stringency and user-defined nature of reverse genetics and is now routinely used throughout biomedical sciences. Homologous recombination is mediated by the electroporation of a selectable cassette (usually antibiotic resistance) that is flanked by 3–10 kb of homologous arms of genomic DNA. The classic method for disrupting a gene is to target a critical exon by omitting the exon from the targeting vector. Therefore, the removal of the exon and replacement with an antibiotic resistant minigene disrupts the gene throughout the organism. Knockin alleles can also be produced in this manner by replacing an exon with exon-containing amino acid substitutions to determine the role of specific mutations on gene function.

Simple gene replacement has been largely supplanted by the generation of more versatile conditional alleles via Cre/Lox technology. Cre recombinase recognizes and recombines a 34 bp lox site (Araki *et al.*, 2002). Two such sites placed in tandem in the same orientation can be recombined generating a minicircle and thus the genomic sequence can be removed between the sites. This allows for tissue-specific or temporal deletion of genes within organisms expressing Cre in a tissue-specific or inducible manner. Many mice have been generated and are widely available that express a tissue-specific or temporally regulated Cre recombinase for most applications. Lox sites are generally placed surrounding a critical exon and the antibiotic resistance cassette is placed within an intron. The resistance cassette is often flanked by Frt sites that can be recombined within germplasm by Flpe recombinase. The goal is to preserve wild-type function of the gene by having minimal foreign-derived sequences that are embedded in putatively inert introns. Cre recombinase can be used to remove sequences if placed in the same orientation or they can be used to invert sequences if placed in the opposite orientation. Cre recombinase is used routinely for the deletion of small

sequences but has also been used to rearrange large sections of chromosomes (Zheng *et al.*, 1999). Additionally, transgenes can be controlled by Cre-mediated recombination by placing genes after a lox-stop-lox cassette (Reamy and Wolfgang, 2011; Rodriguez and Wolfgang, 2012). In this way the large array of Cre recombinase-expressing mice can be brought to bear to regulate transgene expression in a cell-type or temporal manner *in vivo*. Homologous recombination and gene replacement are used in almost every field of biomedical science to inform gene function and regulation in a stringent manner *in vivo*.

Precision Genome Editing

The loss or gain of gene function in transgenic animals has been invaluable for understanding the role of these genes in human development, physiology, or pathophysiology. Removing gene function is not always the most valuable means of studying a gene particularly in the context of understanding the role of human mutations in disease progression. Human mutations are rarely true null alleles, but are usually point mutations that disrupt the proteins' function. Understanding how specific mutations or amino acid modifications affect protein function also has enormous value. The classic means of studying specific mutations is to introduce the mutation by homologous recombination to generate a knockin allele essentially as described above. This is often informative but is laborious to develop these animal models. Genome editing via DNA targeted endonucleases has been extremely efficient at targeting specific DNA sequences via a streamlined production that enables multiple gene editing simultaneously (Hauschild *et al.*, 2011). Originally, zinc finger nucleases (ZFNs) directed toward specific DNA sequences were used to target an endonuclease to generate DNA breaks at specified sequences (Miller *et al.*, 2007). Generating double-stranded DNA breaks induces an error-prone non-homologous end joining that often results in site-specific frame-shift mutations. Additionally, the induction of a DNA break greatly increases the efficiency of homologous recombination at the site of the break. Therefore, providing homologous DNA encoding a site-specific mutation along with the nuclease enables efficient gene editing that can be accomplished within the zygote. Injection of the target-specific nucleases into pronuclear embryos or ES cells can rapidly induce frame-shift mutations or enable gene editing.

Generating target-specific ZFNs requires a great deal of selection and optimization that has limited their broader use. In a similar fashion, transcription activator-like effector nucleases (TALENs) have been used to target endonucleases to specific sequences. The DNA-binding properties of the Tal effectors from *Xanthomonas* are straightforward such that the generation of site-specific binding TALENs can be done in a rather programmable manner. TALENs have proven to be more user friendly, target specific, and effective and have been used to genetically engineer a wide range of species (Sung *et al.*, 2013). Another leap in this technology has been the development of the clustered regularly interspaced short palindromic repeat (CRISPR) system for site-specific DNA engineering of mammalian cells (Mali *et al.*, 2013). CRISPR utilizes a Cas9 endonuclease that is targeted to DNA by a short guide RNA (gRNA). Because the

DNA-binding specificity is encoded by the gRNA, this system is incredibly user friendly and is easily adapted to the species of interest with little optimization (Li *et al.*, 2013; Niu *et al.*, 2014; Wang *et al.*, 2013; Yang *et al.*, 2013). Additionally, since the endonuclease is targeted by the gRNA, multiple sites can be engineered simultaneously by simply providing multiple gRNA sequences. Remarkably, this has been used to target and mutate multiple alleles to homozygosity in a single generation. These facile and exciting genome-editing technologies will likely enable even greater dissection of gene function *in vivo*.

Conclusion

Generating transgenic and knockout animals has proven to be enormously effective at deciphering the molecular mechanisms by which genes operate within complex organismal settings. Genome engineering that was once only feasible in simple organisms such as fungi or invertebrates is not routinely done in mammalian species, even toward the use in nonhuman primate models. Transgenic and gene-replacement technologies that are now available to investigators are vast and can be either straightforward such as the removal of a single gene of interest or can be incredibly nuanced to answer ever more specific questions. Enhancing, removing, or manipulating genes in mammals is central to many research programs aimed at understanding human biomedical sciences and these technologies are continuously growing and improving to provide ever more accessible and informative *in vivo* models.

References

- Araki, K., Araki, M., Yamamura, K., 2002. Site-directed integration of the cre gene mediated by Cre recombinase using a combination of mutant lox sites. *Nucleic Acids Research* 30 (19), e103.
- Banaszynski, L.A., Chen, L.C., Maynard-Smith, L.A., Ooi, A.G., Wandless, T.J., 2006. A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules. *Cell* 126, 995–1004.
- Chen, J., Kelz, M.B., Zeng, G., *et al.*, 1998. Transgenic animals with inducible, targeted gene expression in brain. *Molecular Pharmacology* 54, 495–503.
- Chen, W.V., Delrow, J., Corrin, P.D., Frazier, J.P., Soriano, P., 2004. Identification and validation of PDGF transcriptional targets by microarray-coupled gene-trap mutagenesis. *Nature Genetics* 36, 304–312.
- Dai, Y., Vaught, T.D., Boone, J., *et al.*, 2002. Targeted disruption of the alpha1,3-galactosyltransferase gene in cloned pigs. *Nature Biotechnology* 20, 251–255.
- Ellenbroek, S.I., van Rhee, J., 2014. Imaging hallmarks of cancer in living mice. *Nature Reviews Cancer* 14, 406–418.
- Feil, R., Brocard, J., Mascres, B., *et al.*, 1996. Ligand-activated site-specific recombination in mice. *Proceedings of the National Academy of Sciences of the United States of America* 93, 10887–10890.
- Gong, S., Kus, L., Heintz, N., 2010. Rapid bacterial artificial chromosome modification for large-scale mouse transgenesis. *Nature Protocols* 5, 1678–1696.
- Gong, S., Zheng, C., Doughty, M.L., *et al.*, 2003. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. *Nature* 425, 917–925.
- Hauschild, J., Petersen, B., Santiago, Y., *et al.*, 2011. Efficient generation of a biallelic knockout in pigs using zinc-finger nucleases. *Proceedings of the National Academy of Sciences of the United States of America* 108, 12013–12017.
- Jaenisch, R., 1976. Germ line integration and Mendelian transmission of the exogenous Moloney leukemia virus. *Proceedings of the National Academy of Sciences of the United States of America* 73, 1260–1264.

- Kohn, A.D., Barthel, A., Kovacina, K.S., *et al.*, 1998. Construction and characterization of a conditionally active version of the serine/threonine kinase Akt. *Journal of Biological Chemistry* 273, 11937–11943.
- Li, W., Teng, F., Li, T., Zhou, Q., 2013. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nature Biotechnology* 31, 684–686.
- Lois, C., Hong, E.J., Pease, S., Brown, E.J., Baltimore, D., 2002. Germline transmission and tissue-specific expression of transgenes delivered by lentiviral vectors. *Science* 295, 868–872.
- Mali, P., Yang, L., Esvelt, K.M., *et al.*, 2013. RNA-guided human genome engineering via Cas9. *Science* 339, 823–826.
- Metzger, D., Clifford, J., Chiba, H., Chambon, P., 1995. Conditional site-specific recombination in mammalian cells using a ligand-dependent chimeric Cre recombinase. *Proceedings of the National Academy of Sciences of the United States of America* 92, 6991–6995.
- Miller, J.C., Holmes, M.C., Wang, J., *et al.*, 2007. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nature Biotechnology* 25, 778–785.
- Moitra, J., Mason, M.M., Olive, M., *et al.*, 1998. Life without white fat: A transgenic mouse. *Genes & Development* 12, 3168–3181.
- Niu, Y., Shen, B., Cui, Y., *et al.*, 2014. Generation of gene-modified cynomolgus monkey via Cas9/RNA-mediated gene targeting in one-cell embryos. *Cell* 156, 836–843.
- Pfeifer, A., Ikawa, M., Dayn, Y., Verma, I.M., 2002. Transgenesis by lentiviral vectors: Lack of gene silencing in mammalian embryonic stem cells and preimplantation embryos. *Proceedings of the National Academy of Sciences of the United States of America* 99, 2140–2145.
- Prather, R.S., Lorson, M., Ross, J.W., Whyte, J.J., Walters, E., 2013. Genetically engineered pig models for human diseases. *Annual Review of Animal Biosciences* 1, 203–219.
- Rakhit, R., Navarro, R., Wandless, T.J., 2014. Chemical biology strategies for posttranslational control of protein function. *Chemistry & Biology* 21, 1238–1252.
- Reamy, A.A., Wolfgang, M.J., 2011. Carnitine palmitoyltransferase-1c gain-of-function in the brain results in postnatal microencephaly. *Journal of Neurochemistry* 118, 388–398.
- Rodriguez, S., Ellis, J.M., Wolfgang, M.J., 2014. Chemical-genetic induction of malonyl-CoA decarboxylase in skeletal muscle. *BMC Biochemistry* 15, 20.
- Rodriguez, S., Wolfgang, M.J., 2012. Targeted chemical-genetic regulation of protein stability *in vivo*. *Chemistry & Biology* 19, 391–398.
- Sasaki, E., Suemizu, H., Shimada, A., *et al.*, 2009. Generation of transgenic non-human primates with germline transmission. *Nature* 459, 523–527.
- Skarnes, W.C., Von Melchner, H., Wurst, W., *et al.*, 2004. A public gene trap resource for mouse functional genomics. *Nature Genetics* 36, 543–544.
- Sung, Y.H., Baek, I.J., Kim, D.H., *et al.*, 2013. Knockout mice created by TALEN-mediated gene targeting. *Nature Biotechnology* 31, 23–24.
- Tasic, B., Hippenmeyer, S., Wang, C., *et al.*, 2011. Site-specific integrase-mediated transgenesis in mice via pronuclear injection. *Proceedings of the National Academy of Sciences of the United States of America* 108, 7902–7907.
- Wang, H., Yang, H., Shivalila, C.S., *et al.*, 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell* 153, 910–918.
- Wang, Y., Demayo, F.J., Tsai, S.Y., O'Malley, B.W., 1997. Ligand-inducible and liver-specific target gene expression in transgenic mice. *Nature Biotechnology* 15, 239–243.
- Weigert, R., Porat-Shliom, N., Amornphimoltham, P., 2013. Imaging cell biology in live animals: Ready for prime time. *Journal of Cell Biology* 201, 969–979.
- Wilsbacher, L.D., Yamazaki, S., Herzog, E.D., *et al.*, 2002. Photic and circadian expression of luciferase in mPeriod1-luc transgenic mice *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America* 99, 489–494.
- Wolfgang, M.J., Golos, T.G., 2002. Nonhuman primate transgenesis: Progress and prospects. *Trends in Biotechnology* 20, 479–484.
- Wurdinger, T., Badr, C., Pike, L., *et al.*, 2008. A secreted luciferase for ex vivo monitoring of *in vivo* processes. *Nature Methods* 5, 171–173.
- Yang, H., Wang, H., Shivalila, C.S., *et al.*, 2013. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154, 1370–1379.
- Zambrowicz, B.P., Imamoto, A., Fiering, S., *et al.*, 1997. Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells. *Proceedings of the National Academy of Sciences of the United States of America* 94, 3789–3794.
- Zheng, B., Sage, M., Cai, W.W., *et al.*, 1999. Engineering a mouse balancer chromosome. *Nature Genetics* 22, 375–378.
- Zimmerman, L., Parr, B., Lendahl, U., *et al.*, 1994. Independent regulatory elements in the nestin gene direct transgene expression to neural stem cells or muscle precursors. *Neuron* 12, 11–24.

Viral Nucleic Acids

IP O'Carroll and A Rein, National Cancer Institute at Frederick, Frederick, MD, USA

Published by Elsevier Inc.

Glossary

Cis-acting signal A sequence or structure in a nucleic acid molecule that confers some functional property on the molecule, but this property is confined to the molecule containing the signal and cannot be transferred to other molecules.

Encapsidation Incorporation of nucleic acid into an assembling virus particle.

Icosahedron A solid with 20 faces. Many 'spherical' viruses are regular icosahedra, with 20 equilateral triangular faces and 12 axes of 5-fold symmetry.

Trans-acting factor A factor, typically a protein, produced within a cell and capable of conferring a functional property on other molecules or complexes. For example, a viral genome might contain a *cis*-acting signal enabling it to be packaged into assembling virus particles, while a protein supplied by an expression vector might be incorporated into the virus particles and affect their host range.

Virion Virus particle.

Introduction

Viruses are the most abundant organisms on earth (Breitbart and Rohwer, 2005; Suttle, 2005). An Avogadro's number of infections occurs every second in the world's oceans (Suttle, 2013). Their role in the ecology of the oceans is so significant that they affect the global carbon cycle and contribute to climate change (Danovaro *et al.*, 2011). If placed end to end, marine viruses would span a distance 70 times the diameter of our galaxy.

At the most basic level, a virus particle is a package containing nucleic acid. The packaging material is protein encoded by the nucleic acid. The particle can enter cells and, taking advantage of the cellular biosynthetic machinery, direct the production of progeny virus particles identical to the infecting parent.

This scheme in turn places two requirements on the nucleic acid: it must be replicated in the virus-producing cell, to provide the genetic material encased in the progeny virus particles; and it must encode the proteins needed for the production of the progeny particles, including at a minimum the structural proteins from which the particles will be assembled. The particles are always composed of multiple copies of a limited number of proteins; this fact has fundamental implications regarding particle structure (Crick and Watson, 1956).

The diversity of nucleic acids found in viruses is startling. Viruses may contain: linear double-stranded (ds) DNAs; circular ds DNAs; linear or circular single-stranded (ss) DNAs; coding ('+' strand) RNAs; RNAs ('-' strand) whose complements encode viral proteins; and dsRNAs. In some RNA-containing viruses ('ambisense'), part of the RNA is + strand and part is - strand; in some, individual virus particles contain multiple RNA molecules, which together constitute the viral genome and all of which are required for successful replication of the virus; in others, the RNAs required for replication are distributed among different particles. Some virus particles contain RNA, but this RNA is copied into DNA when the virus infects a new host cell. The size and complexity of viruses also varies over an enormous range: the nucleic acids may contain

as few as 2000 bases or more than 2×10^6 bases, corresponding to coding capacities ranging from two proteins to over 2000 proteins.

Obviously, all viruses extant today have survived by virtue of successful strategies for replicating themselves (Koonin and Dolja, 2013). Therefore, each virus has a way of both replicating and expressing its nucleic acid; these strategies exploit the tools available within the susceptible host cell. In turn, viruses are incomparable sources of information about cellular processes; indeed, much of our most fundamental knowledge about biology, including the fact that DNA is the repository of genetic information in cells and that DNA is copied into mRNA for its expression, has been obtained using viruses (Brenner *et al.*, 1961; Hershey and Chase, 1952; Judson, 1996).

Table 1 presents a partial listing of the nucleic acids in different virus families. One of the most fundamental distinctions between viruses is whether they contain DNA or RNA. As cellular genetic information is carried in DNA, many DNA viruses rely on cellular machinery for both replication of their genomes and production of mRNA transcripts for gene expression. In contrast, cells do not normally make copies of RNA molecules; thus, RNA viruses must encode the proteins needed for RNA replication, including the RNA-dependent RNA polymerase.

As the cellular machinery for replication and transcription of DNA is in the cell nucleus, those viruses using this machinery must perform these tasks in the nucleus. This includes all but very large DNA viruses: the latter, including the poxviruses and the newly discovered *Megaviridae*, encode their own synthetic machinery and replicate in the cytoplasm. Another striking difference between DNA and RNA viruses is in the range of their genome sizes: while the largest RNA virus genomes, those of the coronaviruses, are approximately 32 kb, DNA viral genomes can be nearly 100 times larger, as Pandoravirus DNA is approximately 2.5×10^6 bp.

Despite the amazing variety in the nucleic acids present in virus particles, these nucleic acids must all be replicated by standard Watson-Crick mechanisms: every coding strand must be copied into a noncoding strand of opposite polarity, and

Table 1 Diversity of viral genomes^a

	<i>ds/ss</i>	<i>Polarity</i>	<i>Genome structure</i>	<i>Segments</i>	<i>Factors bound to genome</i>	<i>Genome length</i>	<i>Example of virus family/genus^b</i>		
DNA	ds		Linear	1	Viral terminal protein at 5' end	15–2500 kb	<i>Poxviridae</i>		
			Linear	10		150–250 kb	<i>Polydnnaviridae</i>		
			Linear	1		36–48 kb	<i>Adenoviridae</i>		
			Circular	1		4.5–300 kb	<i>Polyomaviridae</i>		
	ss	+ or –	Linear	1		4–6 kb	<i>Parvoviridae</i>		
			Linear	2		12.5 kb	<i>Bidnaviridae</i>		
			Circular	1		1.8–2.9 kb	<i>Circoviridae</i>		
			Circular	2–8		3–9 kb	<i>Nanoviridae</i>		
	Partially ds		Circular and gapped	1		3–8 kb	<i>Hepadnaviridae</i>		
	RNA	ds		Linear	1	Viral protein (VPg)	4.6–17.6 kb	<i>Totiviridae</i>	
Linear				2	5.6–6.5 kb		<i>Birnaviridae</i>		
Linear				2–4	3.7–16 kb		<i>Cystoviridae</i>		
Linear				10–12	5' cap		18–32 kb	<i>Reoviridae</i>	
ss		+	Linear	1		4.2	<i>Umbravirus</i> (satellite virus)		
			Linear	2–3		6.3–12 kb	<i>Virgaviridae</i>		
			Linear	1		5' cap	3.4–32 kb	<i>Coronaviridae</i>	
			Linear	2–5		5' cap	4.5–15.8	<i>Bromoviridae</i>	
			Linear	1		Viral protein (VPg)	4–9.7 kb	<i>Picornaviridae</i>	
			Linear	2		Viral protein (VPg)	8.5–15.7 kb	<i>Potyviridae</i>	
			Linear	1 (2 copies)		5' cap, cellular tRNA primer	7–10 kb	<i>Retroviridae</i>	
			–	Linear		1		11–19 kb	<i>Paramyxoviridae</i>
				Linear		2–8		11–25 kb	<i>Orthomyxoviridae</i>
				Circular		1		1.7 kb	<i>Deltavirus</i> (satellite virus)
			Ambisense	Linear		2–3		11–19 kb	<i>Arenaviridae</i>

^aViral genomes are grouped based on type of nucleic acid (DNA/RNA), whether it is double-stranded (ds) or single-stranded (ss), polarity (+/–), genome structure (Linear/Circular), number of segments, and the presence or absence of factors bound to the genome. For each group, the range of genome size is given.

^bFor simplicity, only one virus family is shown as an example for each group. It is important to remember that this table indicates the size-range of genomes for all members of a given group, and not merely for the example shown in the last column.

vice versa. The other absolute requirement is that somehow an RNA (or RNAs) of the coding strand must be produced for translation into the virus-coded proteins. The various pathways of information transfer to mRNA are summarized in [Figure 1](#), and were well summarized over 40 years ago ([Baltimore, 1971](#)). It is important to realize that viral nucleic acids possess, in addition to protein-coding sequences, *cis*-acting signals that often perform essential functions in virus replication. Examples are sequences recognized by viral proteins for packaging the genomic nucleic acid into progeny virus particles and sequences required for initiation of complementary nucleic acid strands. In general, when viral genomes are engineered for use as vectors, the sequences encoding viral proteins can be replaced and the needed proteins provided in *trans*, but the *cis*-acting signals must be preserved to permit replication of the vector.

Group I: dsDNA Viruses

We have known for many years that dsDNA viruses vary over a wide range in size and complexity, extending from the

polyomaviruses (~5 kbp) to the poxviruses (~130–365 kbp). Very recently, even far larger dsDNA viruses have also been discovered; these ‘giant viruses’ that have been isolated to date only infect amoebae. Although the genomes of these newly isolated viruses have been sequenced, little is known about their biology or the structure of their DNA as yet ([Hendrix, 2009](#); [Yutin and Koonin, 2013](#)).

As far as is known, all DNA synthesis in living systems is primer-dependent: in other words, DNA chains are not synthesized *de novo*, but only by addition of deoxynucleotides to a preexisting macromolecule (the primer). In cellular DNA synthesis, the primers for one of the two strands are short RNA molecules. dsDNA viruses use a variety of primers in DNA synthesis. The DNA of SV40, a member of the Polyomavirus family, is replicated by the same basic mechanisms as cellular DNA ([Figure 2\(a\)](#); [Stenlund, 2003](#)). In contrast, in both adenoviruses and hepadnaviruses, the primer is a viral protein: the new DNA chain is initiated by addition of a deoxynucleotide to an –OH group on the side-chain of a specific amino acid in a virus-coded protein ([Figure 2\(b\)](#); [Van der Vliet, 1995](#)). In other dsDNA viruses, special structures in the DNA are involved in its replication. For example, poxvirus genomic DNAs

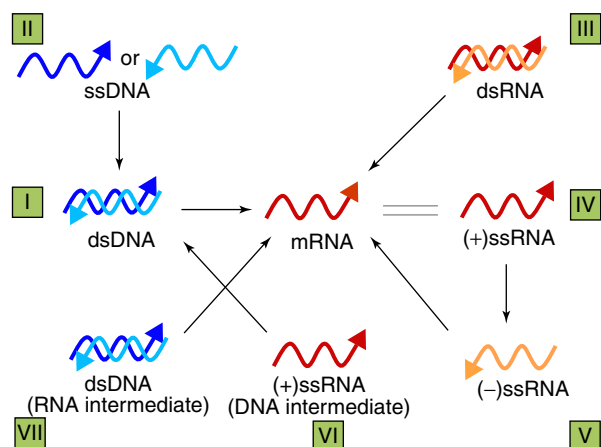


Figure 1 Information pathways in viral mRNA synthesis. The nature of the nucleic acid in the virus particle is indicated in groups I–VII (green boxes). Dark blue and light blue wavy arrows represent positive and negative sense DNA strands, respectively, whereas red and orange wavy arrows represent positive and negative sense RNA strands, respectively. The direction of the wavy arrows indicates polarity (5' to 3'). Straight, black arrows represent a copying step. The two parallel gray lines represent equivalency.

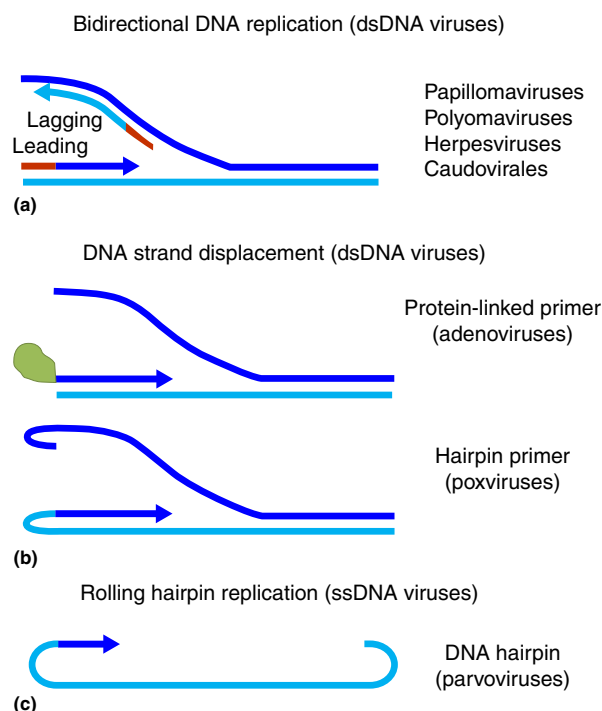


Figure 2 Replication strategies of DNA viruses. (a) Bidirectional replication in dsDNA viruses, (b) DNA strand displacement in dsDNA viruses, (c) rolling hairpin replication in ssDNA viruses. Dark blue lines represent DNA template in the 5' to 3' direction and light blue lines represent the complementary DNA template strand. Dark blue arrows indicate synthesis of the new DNA strand in the 5' to 3' direction. Red lines represent RNA primers. Green-filled shape represents a protein that serves as a primer.

are covalently closed: while replication of their DNAs is still not fully understood, these structures are critical for this process (Figure 2(b)). In many dsDNA viruses, the protein shell is preformed, and the DNA is then pumped into the shell, with the help of energy from ATP, by elaborate 'motor' machinery. This apparatus is best characterized in the case of dsDNA bacteriophages, but analogous mechanisms are apparently used by the herpesviruses as well. The DNA is packed at extremely high density in the fully formed virus particle.

Group II: ssDNA Viruses

The ssDNA viruses include some of the smallest and simplest viruses, with genomes only approximately 2–6 kb in length. One of these viruses is the familiar dog pathogen, canine parvovirus. These small viruses rely on the host cell (and, for some ssDNA viruses, coinfecting larger viruses) for much of their replication machinery. These 'minimalist' viruses may encode only a single structural protein and a single protein involved in their DNA replication. It seems reasonable to speculate that the genome size of ssDNA viruses is limited because a single nick in viral DNA will be a fatal disruption of their genome, unlike in organisms whose information is carried in dsDNA. However, some ssDNA viruses are more complex: these include the Nanoviridae, in which the entire genome is composed of 6 to 8 ~1-kb ssDNA segments, each packaged in a separate particle; the Pleolipoviridae, which infect Archaea and contain a single circular ssDNA molecule ranging up to approximately 10 kb in size; and the Bidnaviridae, which infect silkworms and whose genomes consist of a 6 kb and a 6.5 kb DNA molecule, encapsidated separately.

The DNA in these viruses can be either circular or linear; in the latter case, 'hairpin' structures at the ends of the DNA participate in priming of DNA replication (Figure 2(c)). In some ssDNA viruses, either of the two complementary strands can be packaged, so that a virus preparation is a mixture of particles with either strand; in others, only one of the strands is packaged into virions. The DNAs of ssDNA viruses are replicated by a mechanism similar to 'rolling circle' replication, involving synthesis of dsDNA intermediates containing multiple tandem copies of the viral genome. It should be noted that the template for the mRNAs of ssDNA viruses is not the ssDNA that the infecting virion brings into the cell, but rather the dsDNA produced intracellularly in infected cells (Figure 1).

Group III: dsRNA Viruses

The dsRNA viruses form a large and diverse group of viruses. They include the *Reoviridae*, which in turn include rotaviruses, a major cause of morbidity and mortality from childhood diarrhea; rice dwarf virus, and other important plant pathogens; bluetongue virus, an important disease of sheep and other livestock; and several bacteriophages. The L-A 'virus' of *Saccharomyces cerevisiae* is also classified with these viruses, although, unlike authentic viruses, this agent is not released from cells, but spreads from cell to cell during mating of the host.

The genomes of these viruses can be segmented, with anywhere from 1 to 12 dsRNA molecules packaged together in a virion. Each of these RNAs carries the information for only one (or sometimes two) proteins. RNA-dependent RNA polymerase molecules are also present in the particle. The dsRNA genomes cannot, of course, be translated directly, but they serve as templates for production of (+) strand mRNA molecules. In fact, after infection the dsRNAs in the cytoplasm are retained in a subviral particle; the 'progeny' mRNA molecules are extruded from these particles through pores at 5-fold axes of icosahedral symmetry. (As an icosahedron has 12 5-fold axes, this appears to place an absolute limit on the number of genomic RNA segments that can be accommodated in these viruses.) This sequestration of the dsRNA in an intracellular subviral particle may serve to shield this RNA from detection by elements of the innate immune system. Upon release from the subviral particle, the (+) strand mRNAs are translated in the cytoplasm; they are also incorporated into newly assembling subviral particles, where they are copied by the viral polymerase into (−) strand RNAs, thus reproducing the dsRNA genomic RNA of the virus.

Interestingly, in some dsRNA viruses, RNA replication is semiconservative, while in others it is conservative: in other words, in some, both the (+) and (−) strands of the incoming, parental dsRNA are copied, forming progeny molecules that will give rise to new dsRNA (semiconservative replication). In others, the parental (−) strand is copied multiple times, forming multiple (+) strand progeny RNAs; these are the templates, in these viruses, for synthesis of new (−) strand molecules. The former class includes infectious bursal disease virus, an avian dsRNA virus, while the latter includes the reoviruses.

Group IV: (+) Strand RNA Viruses

The (+) strand RNA viruses include a number of familiar pathogens, such as poliovirus; hepatitis C virus; and the common cold virus (rhinovirus). Many plant viruses, including tobacco mosaic virus (the first virus to be crystallized and the first to be reconstituted *in vitro* from its protein and RNA components), are also (+) strand RNA viruses. These viruses are unique in that their genomic RNA is translated immediately upon infection; that is, the virus particle is simply a package that introduces an mRNA molecule into the cell. It is the translation of this RNA, and the resulting synthesis of the virus-specific proteins, that initiates the virus replication process. Obviously, these proteins include the RNA-dependent RNA polymerase that will replicate the genome; as this protein can be produced in the cell from the viral RNA, it does not need to be imported into the cell in the virus particle. These are the only RNA-containing viruses that do not require a polymerase in the particle; not coincidentally, they are also the only RNA viruses in which infection can be initiated if naked RNA is artificially introduced into the cell.

Most (+) strand RNA viruses have relatively small genomes (4–9 kb). However, the largest known RNAs in nature are the genomes of the coronaviruses, which are 27–32 kb in length (Gorbalenya *et al.*, 2006; Sawicki *et al.*, 2007). Some

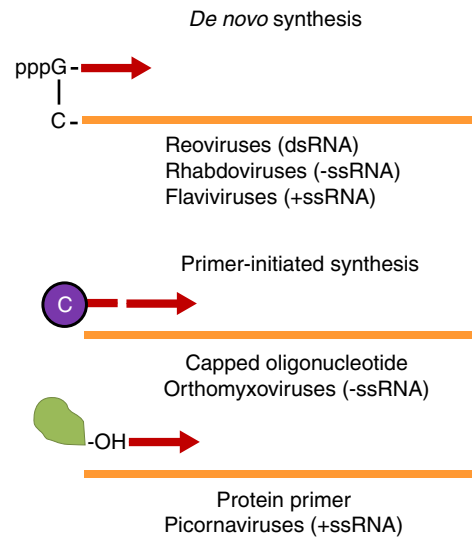


Figure 3 RNA-dependent RNA synthesis can be primer-independent (*de novo* synthesis) or primer-initiated. Orange lines represent RNA template (3' to 5' direction). Red arrows represent RNA synthesis (5' to 3' direction). Purple circle linked to a red line represents a 5' cap linked to a RNA oligonucleotide. Green-filled shape represents a protein that serves as a primer.

(+) strand genomic RNAs are replicated *de novo*, while others are primed by attachment of the first nucleotide to a viral protein (see Figure 3). Many (+) strand RNA viruses infecting plants have structures resembling tRNAs at the 3' end of their genomes; in fact, in many cases these structures can be acylated *in vitro* by tRNA synthetases. These structures appear to play a variety of roles in virus replication, contributing both to translation initiation and to genome replication.

Group V: (−) Strand RNA Viruses

Many viruses, including influenza, Ebola, and rabies viruses, contain RNAs that are complementary to the mRNAs for the virus-coded proteins. Obviously, these (−) strand viruses must contain RNA-dependent RNA polymerase in order to produce the viral mRNAs upon infection. The total length of their genomes ranges from ~11 kb to ~25 kb, but the genomes are frequently segmented, with individual virions containing anywhere from 2 to 8 RNA molecules which collectively comprise the viral genome. Interestingly, no (−) strand RNA viruses infecting prokaryotes have yet been described.

The virus-coded polymerase in a (−) strand RNA virus must be able to copy the genomic RNA(s) to produce the viral mRNAs; copy the genomic RNA(s) into (+) strand RNAs that will serve as templates for production of new genomes; and copy these (+) strand molecules into new, (−) strand genomic RNAs. In many cases, initiation of synthesis does not require a primer: a single nucleoside triphosphate provides the 3'OH for elongation (*de novo* synthesis, Figure 3). In some (−) strand RNA viruses, such as vesicular stomatitis virus, a single genomic RNA molecule contains as many as 6–7 genes. The polymerase produces the individual mRNAs for these

proteins by copying the genomic RNA, beginning by copying the 3' end and proceeding to the 5' end. However, the transcription is interrupted, stopping at the end of each gene and reinitiating RNA synthesis at the beginning of the next gene. In many cases, the switch in polymerase function from mRNA production to genome replication is governed by the accumulation of one or more virus-coded proteins.

Influenza virus is an orthomyxovirus with a segmented (–) strand RNA genome: the virus particle contains 8 distinct RNAs. The 8 RNAs have common sequences of ~12 bases at both of their ends, flanking the individual coding sequences. The sequences at the two ends of each segment are complementary to each other, and it is thought that they pair with each other, producing a circular topology in the RNA. This paired region appears to be the promoter for synthesis of the templated mRNA. Unlike many RNA viruses, influenza replicates its RNA entirely in the nucleus. The 5' ends of the mature viral mRNAs are obtained by 'cap-snatching,' i.e., physically removing the 5' ends from cellular pre-mRNA molecules (including the cap structure) and extending them with viral coding sequences (Figure 3). The 8 RNAs can produce as many as 12 proteins, as some of the mRNAs copied from the genomic RNA are translated from alternative initiation codons ('leaky scanning,' Figure 4(f)) and others are inefficiently spliced, using the splicing machinery in the nucleus. How the virus ensures that all 8 segments will be packaged into progeny virions is a long-standing, unsolved problem in virology.

One more variation on the theme of (–) strand RNA viruses is provided by 'ambisense' RNA viruses. These include members of the bunyavirus family, such as Rift Valley fever virus, and of the Arenavirus family, such as lymphocytic choriomeningitis virus and Lassa virus. In these viruses, the RNA molecules produced by copying the incoming genomic RNAs are not entirely mRNA, as described above; rather, portions of these RNAs are recopied in the infected cell, forming subgenomic mRNAs of the same polarity as the incoming viral RNA.

Group VI: (+) Strand RNA Viruses with DNA Intermediates

The viruses we now call 'retroviruses,' or more specifically 'orthoretroviruses,' have long excited great scientific interest. They were originally detected very early in the twentieth century by their ability to induce tumors in animals. Most recently, they have been the focus of an extraordinary level of attention, because human immunodeficiency virus (HIV-1), the causative agent of AIDS, is a retrovirus. Human T-cell leukemia virus (HTLV-1), which causes leukemias as well as other diseases, is also a retrovirus.

A retrovirus particle contains two RNA copies of its genome. They are ~8–10 kb in length and are both of the same, (+) strand polarity, and they are joined together into a dimeric structure by a limited number of base pairs (D'Souza and Summers, 2005). When the particle infects a new host cell, the RNA-dependent DNA polymerase or 'reverse transcriptase' in the virus copies this RNA into dsDNA. Although two copies of the RNA are present, the virus is best described as 'pseudodiploid,' as only a single genomic DNA copy is synthesized

during the infection. Reverse transcriptase frequently jumps between the two RNAs while it is making the DNA copy, resulting in recombination: this is an important source of genetic variation in these viruses.

After the DNA is synthesized, it is imported into the nucleus along with some of the proteins from the virus particle, including a second enzyme, 'integrase.' This enzyme catalyzes the insertion of the viral DNA into the chromosomal DNA of the cell. The viral sequences are then replicated as part of the chromosome.

The mechanisms by which the viral genes, now resident in the chromosome, are expressed are similar, although not quite identical, to those of cellular genes. The viral RNA carries, near its 3' end, sequences which are copied into the 5' end of the DNA version of the genome; these sequences include promoters recognized by Pol II. (In other words, the viral genomic RNA includes its own promoter, despite the fact that Pol II promoters are normally outside the transcriptional unit.) Like cellular mRNAs, the transcripts are capped at their 5' ends and polyadenylated at their 3' ends.

Some of these RNA copies are destined to be encapsidated into progeny virus particles, while some are mRNAs for Gag (the building block of the virus particle) and Gag-Pol (a fusion protein containing both Gag and the three virus-coded enzymes, i.e., protease, reverse transcriptase, and integrase). All of these RNAs are full-length, intact copies of the viral RNA that entered the cell in the infecting virus particle; thus, they must be exported to the cytoplasm without being spliced, unlike nearly all cellular Pol II transcripts. However, other virus-coded proteins are made from mRNAs that are formed by splicing the full-length viral transcript (Figure 4(c)). These proteins include the Env protein, which functions in attachment and penetration of progeny virus particles into new host cells. Thus, successful virus replication requires that both spliced and unspliced transcripts, in the proper ratio, be exported to the cytoplasm. Retroviral RNAs contain structures which facilitate their escape from the splicing machinery, ensuring that some unspliced transcripts are sent to the cytoplasm. Some retroviruses, including HIV-1, also encode proteins that participate in this process.

One retrovirus, HTLV-1, produces a protein from mRNAs of the polarity complementary to the viral RNA. These mRNA molecules are transcribed from a promoter at the 3' end of the viral DNA. Thus HTLV-1, unlike other retroviruses, is really an 'ambisense' RNA virus. Remarkably, the same stretch of viral sequence that encodes this protein in the 'antisense' direction is also translated from other mRNA molecules in the 'sense' direction: in other words, both complementary versions of this stretch of the viral genome are translated, producing two completely different proteins.

It is interesting to note that the primer for the first strand of retroviral DNA synthesis from the viral RNA is a tRNA molecule. All retroviral genomes contain an 18-base stretch, called the 'primer-binding site,' that is complementary to the 3' 18 bases of a cellular tRNA. Different retroviruses use different tRNAs as primers. The tRNA is annealed to the primer-binding site before or during the assembly of the virus particle; the Gag protein is a highly active nucleic acid chaperone and unwinds the tRNA, making the annealing possible (Rein, 2010).

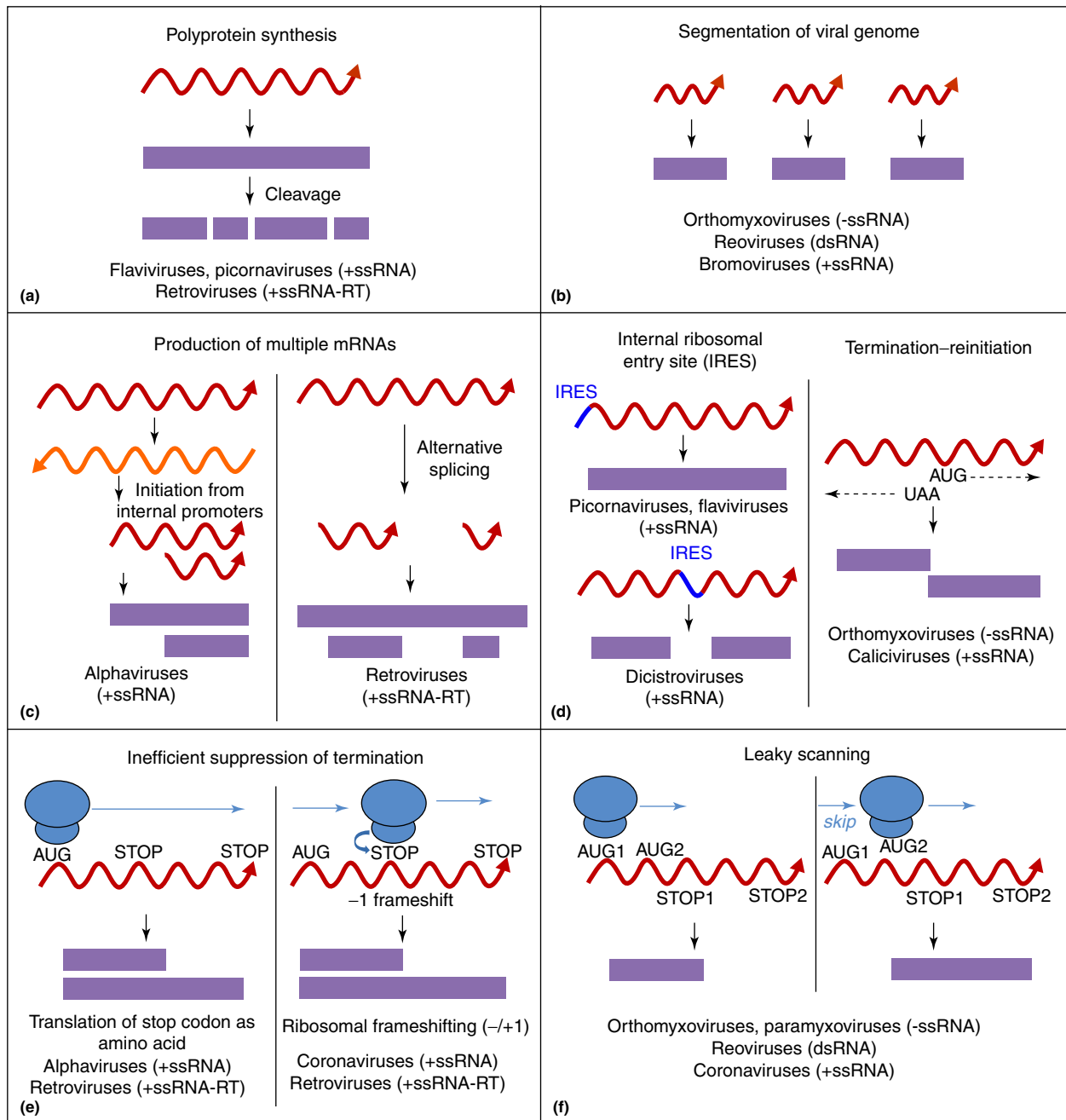


Figure 4 Viral translation strategies. (a) polyprotein synthesis, (b) segmentation of viral genome, (c) production of multiple mRNAs, (d) IRES and termination–reinitiation, (e) inefficient suppression of termination, and (f) leaky scanning. Red and orange wavy lines represent RNAs in the 5' to 3' and 3' to 5' direction, respectively. Solid purple rectangles represent protein products. Blue-filled shapes represent ribosomes. '+ ssRNA-RT' refers to RNA-containing viruses that produce a DNA intermediate in the cell.

The genomic RNA is chemically identical to the mRNA for Gag and Gag-Pol. (Although it has often been suggested that the incoming viral RNA is translated upon infection, there is no definitive evidence that this occurs.) While these RNAs have identical sequences, recent research indicates that they have crucial differences in secondary and tertiary structure. Specifically, dimerization entails changes in the secondary structure of each monomer, as well as the formation of new, inter-molecular base pairs. These changes evidently produce

structures in the RNA that are recognized by the Gag protein, leading to the highly selective packaging of genomic RNA. In fact, there may be two distinct pools of full-length viral RNA in the virus-producing cell: monomers being translated into Gag and Gag-Pol, and dimers destined for encapsidation.

There are six genera of orthoretroviruses, and it is important to recognize the diversity within this family of viruses. Their replication mechanisms and interactions with their host cells are quite different from one genus to another. While a cell

infected with HIV-1 usually dies within a few days after infection, cells productively infected with murine leukemia viruses or avian leukosis viruses almost never die.

It should also be noted that an astonishingly large fraction (~40%) of the sequences in our own DNA comes from RNA, as a result of infections of germline cells with retroviruses or the action of intracellular elements called retrotransposons. Some of these latter elements are remarkably similar to retroviruses, except that there is no extracellular phase in their life-cycle, while others are much more divergent.

Group VII: dsDNA Viruses with RNA Intermediates

There are three families of viruses in which the virion contains DNA, but this DNA is produced in the virus-producing cell by reverse transcription of an RNA intermediate. In essence, this replication scheme is a permutation of the retroviral replication cycle, such that reverse transcription precedes, rather than follows, the release of the virus from one cell and its entry into another. These three families are the Spumaretroviruses or 'foamy viruses,' the hepadnaviruses, and the caulimoviruses. The spumaretroviruses are endemic in many species of primates, but are not known to be associated with any diseases. The prototypical hepadnavirus is hepatitis B virus, an extremely important public health problem associated with hepatocellular carcinogenesis as well as liver cirrhosis. The caulimoviruses are plant viruses exemplified by cauliflower mosaic virus. In these viruses, as in orthoretroviruses, genetic information is continually cycled between DNA and RNA during virus replication; however, there are striking differences between them.

Spumaretroviral genomes are roughly the same size as orthoretroviral genomes (~9–11 kb), and there is close analogy between the genes and many aspects of the replication of these two virus families. One important similarity between orthoretroviruses and spumaretroviruses is that in both, the DNA is inserted by the viral integrase into the chromosomal DNA of the host cell.

In contrast, hepadnaviruses are far smaller and simpler, with genomes of only ~3 kb. (They are not quite as simple as this suggests, as more than half of the genome is translated in more than one reading frame.) Virions contain dsDNA in which the (–) strand is covalently closed while there are gaps in the (+) strand. Upon infection, the DNA enters the nucleus and the gaps in the (+) strand are repaired. Integration of this dsDNA is not necessary for virus replication; rather it is maintained and replicated as an unintegrated circular DNA in the nucleus. (On the other hand, hepatocellular carcinomas in infected humans and animals frequently contain integrated DNA copies of the viral genome; their role in tumorigenesis is not clear at this time.) Transcripts are exported to the cytoplasm, where the virus-coded reverse transcriptase copies them into DNA that will be encapsidated. The first DNA strand to be synthesized (the (–) strand) is primed from a tyrosine residue within the reverse transcriptase protein.

The caulimoviruses infect plants. Virus particles contain DNA with gaps in both strands. Remarkably, one of the virus-coded proteins induces the ribosomes to initiate translation on internal AUG codons in the viral RNA.

Gene Expression Strategies of RNA Viruses

In general, mRNAs in eukaryotic cells are translated beginning at the first AUG initiation codon; as virtually all viruses encode more than one protein, RNA viruses must possess a mechanism for production of multiple proteins from the RNA genome. Different viruses have a wide variety of solutions to this problem. These include: (1) translation of the genome into a single, large 'polyprotein' that is cleaved post-translationally into the mature proteins that function in virus replication (Figure 4(a)). Both hepatitis C virus (a flavivirus) and poliovirus (a picornavirus) use this mechanism. (2) segmentation of the viral genome: it consists of several discrete RNAs, each encoding one (or two) of the required proteins (Figure 4(b)). A number of (+) strand RNA viruses infecting plants (e.g., brome mosaic virus) and insects (e.g., flock house virus) use this strategy, as do orthomyxoviruses such as influenza and dsRNA viruses such as reoviruses. (3) production of multiple mRNAs (Figure 4(c)). For example, the alphaviruses (e.g., Sindbis virus) are (+) strand RNA viruses whose incoming viral RNA is translated at the outset of the infection. However, the (–) strand produced by copying the viral RNA contains one or more internal promoters for the viral RNA-dependent RNA polymerase. Thus, the RNAs produced during replication include not only full-length (+) and (–) strand copies of the genome, but smaller RNAs of (+) strand polarity, representing only a portion of the genomic sequence, that will also serve as mRNAs. In a somewhat analogous strategy, coronaviruses produce a nested series of mRNAs encoding different viral proteins. (4) the presence in the genome of a special RNA structure, the 'internal ribosomal entry site' or IRES (Figure 4(d), left panel; Fraser and Doudna, 2007; Martinez-Salas *et al.*, 2008). An IRES obviates the requirement for the 5' cap structure in initiation of translation. In many viruses, it is located near the 5' end of the viral RNA, and functions principally in protecting the viral mRNA from the inhibitory effects that the virus exerts upon normal, 5' cap-dependent translation of host mRNAs. However, in some viruses including cricket paralysis virus, an IRES in the interior of the viral RNA allows translation of an internally placed open reading frame (Figure 4(d), bottom left). IRES elements derived from viral genomes have been extremely useful in the design of vectors for simultaneous expression of two genes from a single mRNA. A somewhat analogous mechanism is called termination–reinitiation: here the AUG of a second open reading frame in the RNA overlaps the termination codon of a first reading frame within the 5-base sequence UAAUG (Figure 4(d), right panel). Initiation at this AUG is dependent upon a special sequence placed 5' to this junction. Termination–reinitiation is used in (+) strand RNA viruses such as caliciviruses, dsRNA viruses, and (–) strand viruses. (5) inefficient suppression of the termination codon at the end of an open reading frame, so that some ribosomes continue translating past the termination codon (Figure 4(e)). These ribosomes produce a fusion protein, containing some or all of the sequence encoded 5' of the termination codon and, in addition, the sequence 3' of this codon. The suppression can either be by inefficient translation of the termination codon as an amino acid (as in alphaviruses) or by ribosomal frame-shifting at a site 5' of the termination codon (as in

coronaviruses). Both of these mechanisms are also used by retroviruses (Hatfield *et al.*, 1992). Successful virus replication in these cases requires the optimal efficiency of suppression, yielding the proper ratio of extended to terminated translation products. This ratio is often in the range of 1:10. In some retroviruses, there are two successive frameshift sites, resulting in the production of three proteins from the same mRNA molecule. (6) 'leaky scanning,' in which some ribosomes initiate translation at the 5'-most AUG on an RNA, while others bypass this AUG and initiate at a downstream AUG (Figure 4 (f)). The efficiency of utilization of the AUGs is largely governed by their surrounding sequence contexts. In some retroviruses, the viral RNA contains a CUG in an excellent context for initiation, with an AUG further downstream in the same reading frame; both of these codons are used for initiation, resulting in the production of two protein species, identical except for an N-terminal extension on one of them.

It is important to note that these diverse mechanisms for gene expression are not mutually exclusive: many viruses use more than one of these strategies.

Concluding Remarks

As far as we know, all living organisms are hosts to viruses. We have attempted here to briefly indicate the amazing diversity of information-transmission mechanisms found among viruses. Indeed, it is striking that viruses exhibit far more diversity in this regard than do cellular organisms, whose genomes are exclusively dsDNA. However, all viral genomes larger than those of coronaviruses are also composed of dsDNA. Perhaps this is because dsDNA is more resistant than other nucleic acids to the vicissitudes of mutation and chemical and physical damage. In any case, the study of viruses can only inspire wonder at the unimaginable variety found among living things.

Note on Sources

The primary literature on viral nucleic acids is massive. We have cited a few reviews here, but two textbooks (Flint *et al.*, 2009; Knipe and Howley, 2013) are excellent sources of information. In addition, the website (See Relevant Website – ViralZone) is an invaluable resource. We apologize to the many researchers whose results we have summarized without attribution here.

Work of I.P.O. and A.R. is supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute Center for Cancer Research.

See also: Cell Division/Death: Regulation of Cell Growth: Internal Ribosome Entry Site-Mediated Translation. Intracellular Infectiology: Infectious Agents: Virus Factories and Mini-Organelles Generated for Virus Replication. Molecular Principles, Components, Technology, and Concepts: Proteins: Expression

Systems. Nucleic acid Synthesis/Breakdown: RNA Synthesis/Function: Messenger RNA (mRNA): The Link between DNA and Protein

References

- Baltimore, D., 1971. Expression of animal virus genomes. *Bacteriological Reviews* 35 (3), 235–241.
- Breitbart, M., Rohwer, F., 2005. Here a virus, there a virus, everywhere the same virus? *Trends in Microbiology* 13 (6), 278–284.
- Brenner, S., Jacob, F., Meselson, M., 1961. An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature* 190, 576–581.
- Crick, F., Watson, J.D., 1956. The structure of small viruses. *Nature* 177, 473–475.
- Danovaro, R., Corinaldesi, C., Dell'anno, A., *et al.*, 2011. Marine viruses and global climate change. *FEMS Microbiology Reviews* 35 (6), 993–1034.
- D'Souza, V., Summers, M.F., 2005. How retroviruses select their genomes. *Nature Reviews. Microbiology* 3 (8), 643–655.
- Flint, S.J., Racaniello, V.R., Enquist, L.W., Skalka, A.M., 2009. *Principles of Virology, Third Edition, Volume I: Molecular Biology*. Washington, DC: American Society of Microbiology.
- Fraser, C.S., Doudna, J.A., 2007. Structural and mechanistic insights into hepatitis C viral translation initiation. *Nature Reviews. Microbiology* 5 (1), 29–38.
- Gorbalenya, A.E., Enjuanes, L., Ziebuhr, J., Snijder, E.J., 2006. *Nidovirales: Evolving the largest RNA virus genome*. *Virus Research* 117 (1), 17–37.
- Hatfield, D.L., Levin, J.G., Rein, A., Oroszlan, S., 1992. Translational suppression in retroviral gene expression. *Advances in Virus Research* 41, 193–239.
- Hendrix, R.W., 2009. Jumbo bacteriophages. *Current Topics in Microbiology and Immunology* 328, 229–240.
- Hershey, A.D., Chase, M., 1952. Independent functions of viral protein and nucleic acid in growth of bacteriophage. *Journal of General Physiology* 36 (1), 39–56.
- Judson, H.F., 1996. *The Eighth Day of Creation: Makers of the Revolution in Biology*, expanded ed. Plainview, NY: CSHL Press.
- Knipe, D.M., Howley, P.M. (Eds.), 2013. *Fields Virology*, sixth ed. Philadelphia, PA: Lippincott Williams & Wilkins.
- Koonin, E.V., Dolja, V.V., 2013. A virocentric perspective on the evolution of life. *Current Opinion in Virology* 3 (5), 546–557.
- Martinez-Salas, E., Pacheco, A., Serrano, P., Fernandez, N., 2008. New insights into internal ribosome entry site elements relevant for viral gene expression. *Journal of General Virology* 89 (Pt 3), 611–626.
- Rein, A., 2010. Nucleic acid chaperone activity of retroviral Gag proteins. *RNA Biology* 7 (6), 700–705.
- Sawicki, S.G., Sawicki, D.L., Siddell, S.G., 2007. A contemporary view of coronavirus transcription. *Journal of Virology* 81 (1), 20–29.
- Stenlund, A., 2003. Initiation of DNA replication: Lessons from viral initiator proteins. *Nature Reviews. Molecular Cell Biology* 4 (10), 777–785.
- Suttle, C.A., 2005. Viruses in the sea. *Nature* 437 (7057), 356–361.
- Suttle, C.A., 2013. Viruses: Unlocking the greatest biodiversity on Earth. *Genome* 56, 542–544.
- Van der Vliet, P.C., 1995. Adenovirus DNA replication. *Current Topics in Microbiology and Immunology* 199 (Pt 2), 1–30.
- Yutin, N., Koonin, E.V., 2013. Pandoraviruses are highly derived phycodnaviruses. *Biology Direct* 8 (1), 25.

Relevant Website

<http://viralzone.expasy.org/>
ViralZone.

PROTEIN SYNTHESIS/DEGRADATION: TRANSLATION

Contents

Components, Initiation, Elongation, Termination, and Regulation

Biogenesis of Secretory Proteins

The Protein Biosynthetic Machinery of Mitochondria

Regulated Proteolysis of Signaling Molecules: The Proprotein Convertases

Components, Initiation, Elongation, Termination, and Regulation

AA Komar, Cleveland State University, Cleveland, OH, USA

WC Merrick, Case Western Reserve University, Cleveland, OH, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

A, P, and E sites Physical locations on the surface of the ribosome that are occupied by aminoacyl-, peptidyl-tRNA, and deacylated tRNA.

Elongation (of protein synthesis) The sequential steps that lead to the addition of one amino acid at a time to the growing polypeptide chain.

Elongation Factor (EF) A non-ribosomal protein that facilitates the process of elongation only. (Note: eukaryotic elongation factors are designated eEF where the lower case 'e' signifies 'eukaryotic'.)

Initiation (of protein synthesis) The required steps that lead to the placement of the initiator tRNA in the P site of the ribosome, correctly base paired with the initiating AUG codon.

Initiation Factors(IF) A non-ribosomal protein that facilitates the process of initiation. (Note: eukaryotic

initiation factors are designated eIF where the lower case 'e' signifies 'eukaryotic'.)

Protein synthesis The process of joining amino acids in a specific sequence through the α carbonyl and α amino groups via a peptide bond that is templated by an mRNA molecule.

Ribosome recycling The process of splitting of post-termination ribosomes into subunits and release of mRNA and deacylated tRNA, thereby making available ribosomal subunits for the subsequent rounds of initiation/translation.

Termination (of protein synthesis) The codon-directed (UAA, UAG, or UGA) process of cleavage (and therefore, release) of the polypeptide chain from the tRNA in the P site of the ribosome.

Components of Protein Synthesis

The general dogma established some 40–50 years ago said that 'DNA makes RNA makes protein.' Although this general view has become considerably more complex, the basic components associated with the process of converting RNA into protein have remained the same, Transfer RNA (tRNA), messenger RNA (mRNA), ribosomes, and translation factors. These essential components are discussed below.

tRNA

In the process of converting the linear nucleotide sequence information in an mRNA into amino acid sequence information, an adapter is used to convert the three nucleotide genetic code into a single amino acid, tRNA. The genetic code is nearly

universal, meaning that the majority of organisms from all kingdoms use the same codons to specify the incorporation of the same amino acid into a synthesized polypeptide. As there are 61 different code words (or codons; 61 out of the 64 possible codon combinations) that specify one of the 20 'standard' amino acids incorporated into polypeptides during protein synthesis, a large number of tRNAs exist in cells to allow for the optimal decoding of an mRNA. The requirement of tRNA molecules reflects two basic needs, the carrier of an activated amino acid and a matching of an amino acid with the correct codon in the mRNA. In general, the activation of an amino acid can be described by the following series of reactions (see [Figure 1](#); reviewed in [Yadavalli and Ibba, 2012](#)).

In the first step of the reaction, the amino acid is activated by the formation of an acid anhydride bond between its carboxylic acid and the phosphate of AMP, thus preserving the

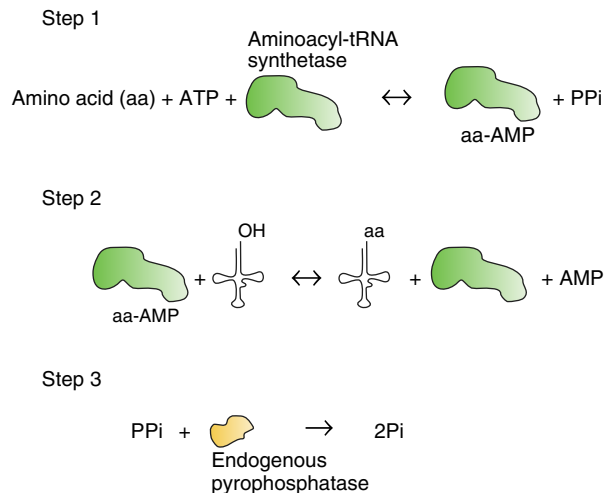


Figure 1 Aminoacylation of tRNA – this figure shows the three steps of aminoacylation with the activation of the amino acid (step 1), the transfer of the activated amino acid to the cognate tRNA (step 2) and the normal physiologic splitting of the pyrophosphate into 2 inorganic phosphate molecules (step 3). This final step pulls the reaction in the direction of aminoacyl-tRNA formation and is aided by the subsequent formation of the elongation ternary complex of eEF1A•GTP•aminoacyl-tRNA.

high energy bond that originally existed in ATP. In the second step, the transfer of the activated amino acid to the tRNA, the high energy is preserved through an inductive affect that results from the neighboring 2' (or 3') hydroxyl group.

The ability of the tRNA to serve as an adapter (between the amino acid and the nucleic acid sequence in the mRNA) is the critical component in translation, the conversion of nucleic acid sequence information into amino acid sequence information. This is facilitated based upon the observation that the recognition of the nucleic acid sequence in an mRNA is affected by base pairing with the anticodon loop of the tRNA. However, the geometry of this pairing is not identical to Watson–Crick base pairing as the 3' nucleotide in the codon does not form a perfect Watson–Crick base pair and thus there is 'wobble' in the recognition (Crick, 1966). The 'Wobble Rules' are:

Nucleotide in the 5' position of the anticodon	Nucleotide in the 3' position of the codon
U	A, G
C	G
A	U
G	U, C
I	U, C, A

mRNA

The mRNA is a direct transcript of the genetic information or DNA in that the sequence information is maintained in nucleic acid sequence (although in eukaryotes, some of this information is spliced out). In general, an mRNA is described as having a 5' UTR (untranslated region), a coding region (made of codons of three nucleotides) and a 3' UTR. In both prokaryotes and

eukaryotes, the initiating code word is AUG (with few exceptions from the universal or standard rule) and the terminating code words are UAA, UAG, and UGA (which are recognized by specific protein factors and for which there is no corresponding aminoacyl-tRNA). It should be noted, however, that the so-called 21st (Selenocysteine (Sec)) and 22nd (pyrrolysine (Pyl)) amino acids that have been recently found to be incorporated into proteins during protein synthesis are encoded by codons that normally function as stop signals. Sec is incorporated via a UGA codon and Pyl is inserted by a UAG codon. So far, Sec and Pyl are the only known additions to the pool of 20 universal/standard amino acids. The mechanism of Sec incorporation is quite well understood and requires the presence of the so-called selenocysteine insertion sequence (SECIS) element, a specific structural RNA element present in mRNA (Bellinger *et al.*, 2009; Driscoll and Copeland, 2003). The mechanism of Pyl incorporation is known in less detail (Dinman, 2012). Both mechanisms require the specific Sec-tRNA^{Sec} and Pyl-tRNA^{Pyl}, responsible for decoding of the respective stop codons. Both prokaryotic and eukaryotic systems initiate translation of the majority of their mRNAs with a methionyl-tRNA species (where in bacteria, the amino nitrogen in methionine is formylated) and perform elongation with methionine using a separate methionyl-tRNA. These two species of methionyl-tRNA are differentially recognized by either initiation factors or elongation factors. This ensures that there will be independent supplies of the methionyl-tRNAs, both for initiation and elongation.

The initiation step is important as it establishes the reading frame. But which AUG? In both bacterial and eukaryotic systems, a second element comes into play that distinguishes the initiating code word. In the bacterial system, there is a conserved nucleic acid sequence 5' of the initiating AUG that is complimentary to the 3' end of the rRNA in the small ribosomal subunit (16S) and is 4 to 7 nucleotides upstream (referred to as the 'Shine–Delgarno sequence'; Shine and Delgarno, 1975). In eukaryotic mRNAs, it is (as a rule) the AUG that is closest to the 5' end of the mRNA that is recognized.

Ribosomes

Ribosomes are the catalytic ribonucleoprotein particles on which protein synthesis occurs. Their general composition is as described below (Melnikov *et al.*, 2012):

Bacteria	Eukaryotes
30S subunit	40S subunit
16S rRNA	18S rRNA
~ 20 proteins	~ 30 proteins
50S subunit	60S subunit
5S rRNA	5S rRNA
23S rRNA	5.8S rRNA
	26–28S rRNA
~ 30 proteins	~ 50 proteins
MW = 2.3×10^6	MW = $3.3 - 4.3 \times 10^6$

Of the total mass of the ribosome from either bacteria or eukaryotes, the ribosome is roughly 50% RNA and 50%

protein. Prokaryotic ribosomes on average usually contain slightly less protein (35–45%) than RNA. In spite of this distribution, much of the actual surface of the ribosome is RNA. Thus, the recognition of particular positions on the ribosome by translation factors is often guided by contacts between both protein and rRNA.

If one follows the development of the rRNA from bacteria to eukaryotes, one finds all the ancestral parts in the eukaryotic ribosome with insertions at various points within the linear rRNA sequence. Thus the overall folding of the rRNA and the shape of the bacterial and eukaryotic ribosomes are quite similar. In spite of this similarity, bacterial translation factors do not function on eukaryotic ribosomes and eukaryotic factors do not function on bacterial ribosomes.

Eukaryotic Translation Factors

Listed below are all the translation factors that participate in eukaryotic protein synthesis at the level of initiation (eIF, eukaryotic initiation factor), elongation (eEF, eukaryotic elongation factor) or termination (eRF, eukaryotic release factor). Historically, these proteins were named ‘factors’ as they facilitated binding reactions but may or may not have been associated with an enzymatic process, the making or breaking of a covalent bond.

Name	Molecular weight ^a	Function
eIF1	15000	Accuracy in start site selection, Pi release from eIF2
eIF1A	15000	Accuracy in start site selection
eIF2	125000	Forms ternary complex with GTP and Met-tRNA ^{iMet}
eIF2A	67000	Suppresses IRES-mediated translation
eIF2B	260000	Facilitates GTP exchange for eIF2•GDP
eIF2D	65000	Mediates ribosomal recruitment of Met-tRNA ^{iMet} in a GTP-independent manner within minor pathways
eIF3	650000	Provides the free pool of 40S subunits for initiation
eIF4A	45000	ATP-dependent RNA helicase
eIF4B	160000	Stimulates eIF4A and eIF4F activities
eIF4F	250000	Recognizes m ⁷ G cap, unwinds RNA
eIF4H	25000	Stimulates eIF4A and eIF4F activities
eIF5	65000	Stimulates GTP hydrolysis in the ternary complex
eIF5A	15000	Activates peptide bond formation from the E site
eIF5B	125000	GTP-dependent subunit joining
eEF1A	50000	Forms a ternary complex with GTP and aa-tRNA
eEF1B	110000	Stimulates GDP exchange of eEF1A•GDP
eEF2	100000	Drives translocation using GTP
eEF3	116000	ATPase, required for elongation (in fungi only)
eRF1	47000	Recognizes stop codons UAA, UAG, UGA
eRF3	75000	Binds eRF1, binds and hydrolyzes GTP
ABCE1	65000	Uses ATP to recycle ribosomes
DHX29	155000	Promotes scanning on mRNAs with highly structured 5' UTRs
Ded1	66000	Promotes scanning in yeast <i>Saccharomyces cerevisiae</i>
PABP	71000	Facilitates recruitment of recycled 40S subunits to the 5' end of mRNA via mRNA circularization

^aAggregate molecular weight: eIF2, eIF2B, eIF3, eIF4B, eIF4F, and eEF1B are multisubunit proteins.

When one considers the relative amounts of most of the components in the translation system, the rough numbers come out as follows:

1 ribosome: 0.3–1.0 initiation factors and eEF2; 8–10 aminoacyl-tRNAs: 20–30 eEF1A: 0.1–0.2 mRNAs (encoding a 50 000 molecular weight protein).

The relatively large excess of eEF1A ensures that all of the aminoacyl-tRNA is present in the cell as a ternary complex (eEF1A•GTP•aminoacyl-tRNA) ready to participate in protein synthesis. This is also true for the initiating methionyl-tRNA as there is a molar excess of eIF2. Additionally, complexes of the aminoacyl-tRNA with either eEF1A or eIF2 stabilize the aminoacyl linkage to extend the half life of the chemical bond approximately 5-fold over free aminoacyl-tRNA.

Cap-Dependent Protein Synthesis, the Predominant Pathway

Approximately 90% of the mRNAs within eukaryotic cells are translated via the ‘cap-dependent pathway’ (Aitken and Lorsch, 2012; Hinnebusch, 2011; Jackson *et al.*, 2010; Merrick, 2003). Simply put, the mRNA’s m⁷G cap and neighboring nucleotides are bound to a 40S subunit containing bound ternary complex. Subsequent movement of the mRNA (scanning) along the 40S subunits leads to the recognition of the AUG start codon via the anticodon of the Met-tRNA^{iMet} in the ternary complex. This is followed by subunit joining.

This process has been broken down into considerably greater detail, both mechanically (when proteins are bound or released from the 40S subunit) and kinetically (where are the slow/rate limiting steps). Figure 2 presents a useful pathway that reflects the more detailed mechanics of eukaryotic initiation (Merrick, 2010). As with bacteria, the eukaryotic ribosomal subunits tend to associate forming inactive 80S ribosomes that are in equilibrium with 40S and 60S subunits. The binding of eIF3 (and eIF1 and eIF1A) to the 40S subunit shifts the equilibrium to 40S subunit complexes and this serves as the original pool of 40S subunits on which to build an initiation complex. The next step is the addition of the ternary complex (eIF2•GTP•Met-tRNA^{iMet}) to the 40S complex, a reaction facilitated by the interaction of the ternary complex with both the 40S subunit and eIF3.

To go further in the pathway requires the addition of mRNA. In eukaryotic cells with the synthesis and processing of the mRNA in the nucleus, the molecule that enters the cytoplasm is an mRNP which is about 50% protein and 50% RNA. In order to bind the 5' end of the mRNA to the 43S complex, the 5' end needs to be single stranded and absent of protein so that the RNA strand can be bound into the 43S mRNA channel. The conversion of the mRNP into an mRNA suitable for binding to the 43S complex is a process referred to as ‘activation’ and requires eIF4A, eIF4B, and eIF4F and ATP. It is the helicase activity of eIF4A, a DEAD-box helicase, that accounts for this conversion although the relative contribution of eIF4A versus eIF4F (a three subunit protein composed of eIF4A, eIF4E, and eIF4G) is unknown. When examined *in vitro*, eIF4F is a more rigorous helicase and therefore has been assumed to be the predominant protein responsible for the activation

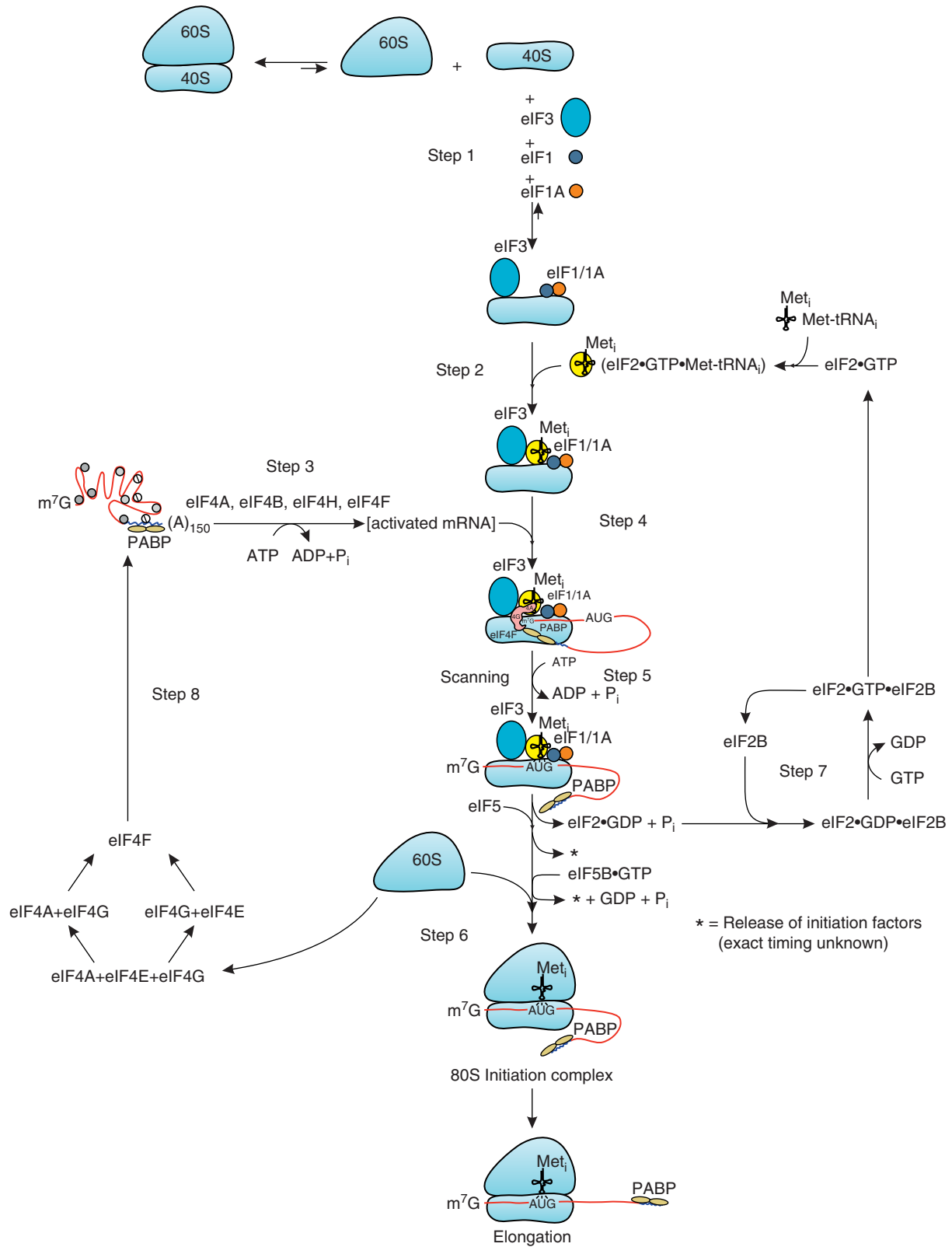


Figure 2 80S initiation complex formation – formation of an initiation complex as seen for the attachment of the first ribosome to an mRNA. Steps 1 through 6 are involved in the synthesis of the 80S initiation complex, ready to enter the elongation phase of translation. Steps 7 and 8 represent the recycling steps and these are the steps subject to global regulation via protein phosphorylation. Adapted from Merrick, W.C., 2010. Eukaryotic protein synthesis: Still a mystery. *Journal of Biological Chemistry* 285, 21197–21201 and used with permission of the Journal of Biological Chemistry.

process. In addition, the association of eIF4F with the mRNA helps in the binding of the mRNA to the 43S subunit via a protein–protein interaction with eIF3.

Once the mRNA is associated with the 43S complex, making it a 48S complex, the process of scanning begins where the mRNA is inspected in a 5′ to 3′ direction in search of the initiating AUG codon. This process requires ATP and for most mRNAs, eIF4A is sufficient to locate the start codon. More recent experiments have implicated several other DEAD-box helicases in the process of scanning as a requirement for the optimal expression of proteins encoded by mRNAs that have either long 5′ UTRs or 5′ UTRs with extensive secondary structure (Pisareva *et al.*, 2008).

The sensing of the initiating AUG codon is performed by the anticodon of the Met-tRNA^{Met} and this accounts for the highly specific recognition of the AUG codon (Cigan *et al.*, 1988). However, this does not account for the finding that the optimal start codon context is A/GXXAUG (Kozak, 1986). The preference for a purine at the minus 3 position (the A in AUG is nucleotide +1) reflects an interaction with the α subunit of eIF2 (Pisarev *et al.*, 2006). In mammalian systems, it was also noted that often the presence of a G following the AUG provided a preferred context. However, this interpretation is not entirely straight forward. In the synthesis of proteins, the N-terminal methionine is often a subject to co- and posttranslational modifications (Jha and Komar, 2011), i.e., deformylation in bacteria and removal in both prokaryotes and eukaryotes, which is expected to occur in about half of all synthesized proteins and depending on the final amino terminus (which amino acid, is it N-acetylated), the protein may be stable or quite labile (sensitive to proteasome degradation). Thus, in the absence of data relating to the half life of the protein, the actual preference for nucleotides 3′ of the initiating AUG must be more carefully investigated.

With the correct start codon identified, there is hydrolysis of the GTP in the ternary complex (if this has not already happened), release of eIF1 and release of the phosphate bound to eIF2 (Nanda *et al.*, 2009, 2013). The release of the phosphate drives the release of the eIF2•GDP complex. Subsequent binding of eIF5B•GTP allows for subunit joining with the 60S subunit and the correct placement of the Met-tRNA^{Met} in the P site. At present, the release of the factors associated with the early steps of initiation (eIF1, eIF1A, eIF3, eIF4A, B or F, and eIF5) has been accomplished with the release of eIF1 or at the latest, with the release of the eIF2•GDP complex. Retention of the other initiation factors on the 40S subunit blocks subunit joining and the completion of the initiation process.

While the above steps have led to an initiation complex prepared to form the first peptide bond, there are two steps that are not included for the full catalytic use of the initiation factors. The first is the recycling of eIF2. As a classic G protein, eIF2 exits the initiation pathway as an eIF2•GDP complex. To re-enter the pathway the bound GDP needs to be exchanged for GTP and this requires the protein complex termed eIF2B. As seen in step 7 of Figure 2, eIF2B accomplished the exchange for GTP made necessary by the very tight binding of GDP by eIF2, $K_d = 10^{-8}$ M, compared to the looser binding of GTP, $K_d = 10^{-6}$ M. Once the GDP is exchanged, the eIF2•GTP complex can then rebind Met-tRNAi.

In a similar manner, during the process of initiation the subunits of eIF4F come apart and then need to be reassembled although the precise time and place for the disassembly is not known. Given the ability of eIF4G to bind either eIF4E or eIF4A, it is likely that the reassembly process would allow for the initial binding of either eIF4A or eIF4E with the other subunit binding subsequently.

Both of the recycling steps, for eIF2•GDP and the reassembly of eIF4F, are the major points for global regulation via various protein kinases and this regulation will be discussed in greater detail below. However, the mechanistic outcome of this regulation depends on the process (Merrick, 2003). When down regulation is affected by eIF2 α phosphorylation (on Ser51), the reduction in the level of ternary complexes (eIF2•GTP•met-tRNA^{Met}) tends to reduce the translation of the majority of mRNAs equally, with the exception of a few mRNAs that use alternative ways or codons to initiate translation. In contrast, down regulation of the level of eIF4F activity through 4E-BP drives mRNA competition such that poorly translated mRNAs may be completely silenced while only a modest reduction in expression is seen for the efficiently translated mRNAs (perhaps only a reduction of 10 to 20%).

Other Initiation Events

Although cap-dependent translation is the most common form, a reasonable number of mRNAs contain internal ribosome entry sites (IRESs) which allow for the direct binding of the 40S subunit at or near the internal initiating AUG (Komar *et al.*, 2012; Thompson, 2012). This 40S subunit may or may not already have a bound ternary complex and/or eIF3. The differential requirement for initiation factors depends on the mRNA and is best characterized by viral IRES elements which may require all of the initiation factors except eIF4E (poliovirus), all of the initiation factors except the mRNA specific ones (eIF4A, eIF4B, and eIF4F; hepatitis C virus) or none of the initiation factors (cricket paralysis virus). Although not proven, it would appear that most cellular IRES-containing mRNAs likely require all of the initiation factors except eIF4E.

IRES elements are generally characterized by being GC rich, having more extensive secondary structure and containing several hundred nucleotides in the 5′ UTR relative to the mRNAs that use the cap-dependent pathway. For what few structures are known by cryo-EM, the IRES element is commonly found bound to the equivalent location of the E site of the ribosome and it is this association that facilitates the binding of the mRNA in the absence of binding of the m⁷G cap and scanning as described above. However, the binding of the mRNA in this fashion does not preclude that subsequent scanning of the IRES-containing mRNA might also occur.

Cap-dependent and IRES-mediated translation appear to account for 95 to 99% of cellular translation. Other rare initiation events include shunting (Yeuh and Schneider, 2000), re-initiation (Hinnebusch, 2011), cap-dependent internal initiation (Shatsky *et al.*, 2010), initiation with short 5′ UTRs (TISU or translation initiation of short 5′ UTRs) (Dikstein, 2012), and RAN translation (repeat associated, non-AUG) (Zu *et al.*, 2011) are also known to occur, but the mechanism by

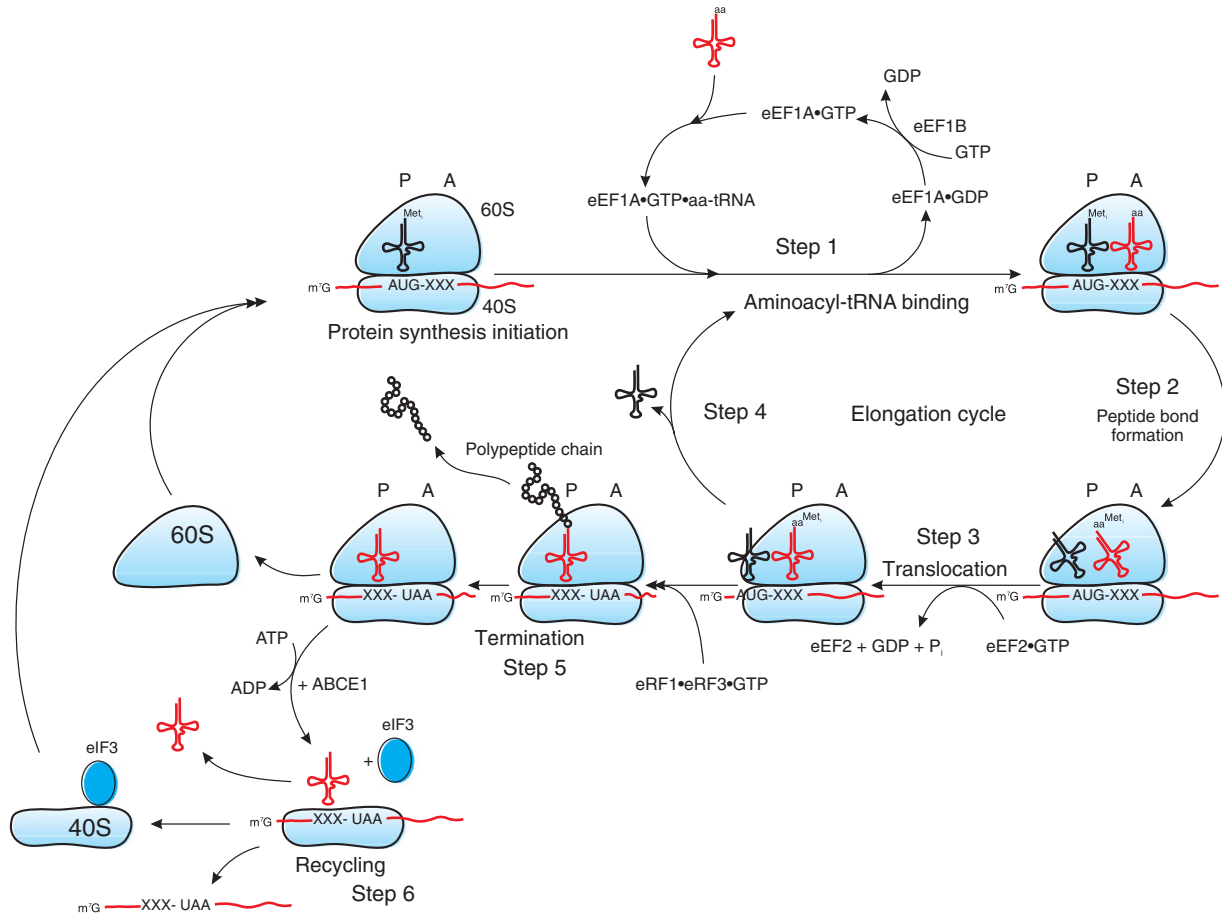


Figure 3 Elongation, termination, and ribosome recycling – shown above is the movement of an 80S initiation complex into the elongation cycle which uses sequentially eEF1A•GTP•aminoacyl-tRNA, the peptidyl transferase center of the ribosome and eEF2•GTP. These three steps are then repeated until a termination codon is reached (UAA, UAG, or UGA). When a stop codon is reached, the complex of eRF1•eRF3•GTP recognizes the stop codon and triggers the release of the polypeptide chain from the tRNA in the P site. The resulting ribosome complex is then recycled in an ATP-dependent manner by ABCE1 which drives the release of the mRNA and tRNA. This process is stimulated by the presence of eIF3 which helps to begin the next initiation event. As in initiation, eEF1A•GDP requires a recycling protein to regenerate eEF1A•GTP which can then rebound an aminoacyl-tRNA.

which these events happen remains understudied and poorly described. As such, they will not be described in further detail within this review.

The Elongation Cycle

The completed initiation complex contains the initiating Met-tRNA^{Met} in the P site and has an exposed codon in the A site. The next step in translation is the forming of a ternary complex of eEF1A•GTP•aa-tRNA^{aa} which can then bind to the A site codon in response to the standard use of the genetic code (where a specific triplet in the mRNA corresponds to only one aa-tRNA; see Figure 3). The correct selection of the aa-tRNA resides in two steps, one where the GTP might be hydrolyzed to stabilize the complex and allow the release of either the eEF1A•GDP•aa-tRNA^{aa} from the ribosome or the release of eEF1A•GDP with an imperfectly paired aa-tRNA^{aa} in the A site that is then subsequently lost. The key to the correct selection is the conformational change in the ribosome that takes place

with the correctly paired aa-tRNA and adds several additional hydrogen bonds for stabilization (thereby explaining the difference in stability of the correctly vs. the incorrectly paired aa-tRNA in the A site).

As perhaps the fastest step in the elongation cycle, correctly paired aa-tRNA in the A site immediately undergoes condensation with the Met-tRNA^{Met} to form a peptide bond on the tRNA in the A site. The chemistry of this reaction appears to be the nucleophilic attack by the amino group of the A site aa-tRNA aided by the 2' hydroxyl of the 3' terminal adenosine of the tRNA. With peptidyl-tRNA bound in the A site and a free tRNA in the P site, the next step is to move the mRNA by one codon's worth to be able to extend the growing polypeptide chain by another amino acid. Here eEF2•GTP binds to the A site of the ribosome and with hydrolysis of the GTP triggers a conformational change that moves the peptidyl-tRNA into the P site and the free tRNA into the E site. This exposes a new codon in the A site of the ribosome and thus allows the elongation cycle to continue. It has been noted that experimentally, the accuracy of matching the incoming aa-tRNA with

the A site codon is enhanced by having the free tRNA in the E site and that the so-called translocation (movement of the ribosome from one codon to another) is occurring through a number of substeps, which are not discussed here for simplicity. With each subsequent binding of an aminoacyl-tRNA to the A site, there is the release of the free tRNA from the E site. Thus, there are at least two tRNAs associated with the ribosome at all times post initiation.

Not pictured in the elongation cycle is the protein eIF5A originally identified as stimulating methionyl-puromycin synthesis, a model initiation assay. Current structural information indicates that eIF5A binds in the vicinity of the E site on the ribosome and biochemically, this enhances the rate of peptide bond formation between an aminoacyl-tRNA or peptidyl-tRNA in the P site with an aminoacyl-tRNA in the A site. While the extent of increase in the elongation rate may only be about 2-fold, for stretches of polyprolines (at least three), the requirement for eIF5A becomes nearly absolute (Buskirk and Green, 2013; Gutierrez *et al.*, 2013). From the databases, it would appear that most eukaryotes (i.e., humans) contain about 7000 different proteins with at least three prolines in a row which may account for the essential nature of eIF5A (Pell *et al.*, 2013).

As was observed in the initiation pathway, eEF1A is also a G protein with an approximate 100-fold preference for GDP over GTP. As a consequence, a GTP exchange factor is required, eEF1B. In this manner, eEF1A•GDP is converted to eEF1A•GTP which can then bind a new aminoacyl-tRNA to form the ternary complex eEF1A•GTP•aa-tRNA^{aa} and continue to participate in protein synthesis. Perhaps curiously, although eEF2 is also a protein that uses GTP for the translocation step, it has the 'more usual' higher affinity for the substrate GTP than the product GDP and thus no recycling protein is required for eEF2.

In bacterial systems, it has been estimated that in log phase growth, approximately 60% of the cellular energy is devoted to protein synthesis. Equivalent amounts of energy may also be devoted to protein synthesis in log phase growth of single cellular eukaryotes or cells in tissue culture. This large expenditure of energy is the result of the requirement for 4 high energy phosphates for each peptide bond formed (aminoacylation, 1 ATP → AMP + PPi → 2 Pi; elongation, 2 GTP → 2 GDP + 2 Pi) as the amount of energy associated with initiation (2 GTP → 2 GDP + 2 Pi, some small number of ATPs to ADPs) or termination (1 GTP → 1 GDP + 1 Pi, 1 ATP → 1 ADP + 1 Pi) is quite small by comparison as each of these events only happens once.

In yeast, however, an additional fungi-specific elongation factor eEF3 is required to facilitate release of deacylated tRNA from the E site in an ATP-dependent manner, thereby allowing binding of the next aminoacyl-tRNA to the A site (Chakraburty, 2001; Andersen *et al.*, 2006). Therefore, in yeast the translation process is even more energy consuming. eEF3 is an essential protein and is required for each cycle of translation. It is unclear why translation in yeast requires eEF3, but translation in bacteria or higher eukaryotes does not. The requirement for eEF3 resides in the yeast ribosome as mammalian ribosomes do not require eEF3 when using yeast eEF1A and eEF2 as the elongation factors. No eEF3 homologs can be found in genomes of other organisms. Comparison of the

yeast and bacterial ribosomes does not reveal any particular structural features that would indicate the need for eEF3 in case of yeast elongation cycle. Thus, at present the evolutionary origin of eEF3 and its requirement remain rather enigmatic.

Termination of Translation and Ribosome Recycling

At the end of the coding region is one of three termination codons (UAA, UAG, or UGA). In all of translation, this is the one time that the nucleotide sequence is recognized by a protein, eRF1 (Dever and Green, 2012; Shoemaker and Green, 2011). The complex of eRF1•eRF3•GTP binds to the A site of the ribosome in response to a stop codon in the A site (see Figure 3). The binding of just eRF1 is sufficient to trigger the hydrolysis of the aminoacyl linkage to the tRNA in the P site thus releasing the growing polypeptide chain from its tethered position. However, *in vivo*, it is thought that this event is coupled with the hydrolysis of GTP in eRF3 so that release of the growing polypeptide chain and separation of eRF1 and eRF3 occur at nearly the same time.

Unfortunately, the product, a ribosome with a bound mRNA and unacylated tRNAs in the P and E sites, is relatively stable. To promote the recycling and reuse of the ribosome, following peptide release the protein ABCE1 binds and in the presence of ATP (and eIF3), drives the dissociation of this complex into its individual parts (Pisarev *et al.*, 2010). This is enhanced by the presence of eIF3 as its binding to the 40S subunits prevents possible reassociation with the 60S subunit (as was observed in initiation; see Figure 2). ABCE1 is essential and highly conserved in eukaryotes and archaea.

Protein Synthesis Beyond the First Ribosome

The pathways shown in Figures 2 and 3 provide the mechanisms for the synthesis of a protein from the first binding ribosome. In actuality, mRNAs generally are bound by 3 to 5 ribosomes and this complex is referred to as a polysome. This reflects the general observation that mRNAs can be initiated while there is still a ribosome attempting to complete the synthesis of the encoded protein. In general, it appears that ribosomes that are released at termination begin the initiation process more efficiently. In part, this is because that in log phase growing cells there is a shortage of free 40S subunits as these appear to be participating in elongation. However, there is also an element of intramolecular interaction that is facilitated by the circularization of the mRNA bringing the 3' end of the mRNA in proximity of the 5' end of the mRNA. This is achieved through physical interactions of the poly(A) binding protein (PABP) at the 3' end with eIF4F at the 5' end of the mRNA. This 'intramolecular effect' yields 40S subunits at termination that are 3 to 10-fold more likely to begin a new series of steps in the initiation pathway than is a free 40S subunit. In model systems, this accounts for much of the loss of translational efficiency when the poly(A) tail is omitted and hence there is no PABP at the 3' end of the mRNA.

Molecular Mimicry

The primary element bound in the P site of the ribosome is either aminoacyl-tRNA or peptidyl-tRNA. In contrast, binding to the A site can involve eEF1A•GTP•aminoacyl-tRNA, eEF2•GTP, eRF1•eRF2•GTP or ABCE1•ATP. Yet three of these complexes lack nucleic acid. From three dimensional structures, both eRF1 and a domain of eEF2 have a protein projection that takes on the appearance of the anticodon stem and loop (negatively charged) and for eRF1 an amino acid sequence that recognizes the stop codons through hydrogen bonding in a manner similar to nucleotides. This ability of a protein domain to be a mimic of the anticodon stem and loop has been referred to as molecular mimicry (Nissen *et al.*, 1995).

Mistakes in Translation

There are two different sites where mistakes (the insertion of an incorrect amino acid) can occur. The first is in the attachment of the amino acid to its cognate tRNA. This step has been shown to have errors in that similar amino acids can be attached to a tRNA (i.e., valine, leucine, and isoleucine). However, the slight degree of incorrect amino acid attachment is compensated by the editing domain of the aminoacyl-tRNA synthetase that quickly removes non-cognate amino acids from the tRNA. Thus, this is not a source of error under most circumstances. The other source of error is in the matching of the aminoacyl-tRNA to the correct codon in the A site on the ribosome. This error rate is about 3 errors in 10^4 polymerized amino acids. Thus, for a protein that is 50 000 in molecular weight, out of 20 copies made, 17 would be perfect and 3 would be mutant. These mistakes in many instances may not be biologically significant as the errors in reading the genetic code tend to replace one amino acid with a similar one (i.e., leucine with valine). In addition, while there are numerous proteins that have molecular weights above 500 000, in most instances these proteins are not a single polypeptide chain, but an aggregate of multiple subunits. In this way, high molecular weight complexes can be assembled with perfect subunits.

Global Regulation of Protein Synthesis

At a first glance, control of translation is quite complex and incompletely understood as there are a number of initiation factors that undergo co- and post-translational modification (for a total of 40 or more co- and post-translational modifications). Some of these may be small contributors (perhaps 5–10% increase or decrease in activity) or these modifications may influence the half life of the proteins and thus their total abundance. However, there are three well-established mechanisms for regulation that are affected by phosphorylation, two of which are inhibitory (eIF2 and eEF2) and one that is stimulatory (4E-BP). The regulation of eIF2 and 4E-BP activities stems directly from the 80S initiation pathway shown in Figure 2. The site of regulation is the recycling steps, steps 7, and 8. Generally, in eukaryotes, initiation is a predominant step at which regulation of protein synthesis is achieved.

There are 4 specific eIF2 kinases that phosphorylate serine 51 of the α subunit of eIF2, therefore inhibiting its activity and substantially downregulating cap-dependent initiation: GCN2 (activated by amino acid starvation); HRI (activated by low heme concentrations, mostly in erythrocytes), PKR (activated by double-stranded RNA) and PERK, an endoplasmic reticulum kinase located in the ER membrane (activated by oxidative stress or unfolded proteins) (Baird and Wek, 2012; Kimball and Jefferson, 2010). Although not well characterized, one or more of these kinases seems to also be activated by other stress signals as well. This phosphorylation yields an eIF2 that when it exits the initiation pathway as eIF2PO₄•GDP binds to eIF2B tightly, but does not exchange the bound GDP for GTP. As there is roughly 5 to 10 times as much eIF2 as eIF2B, when 20% of the eIF2 is phosphorylated, protein synthesis is almost 100% inhibited as there is no free eIF2B to catalyze the nucleotide exchange. This leads to a marked reduction in ternary complexes (eIF2•GTP•Met-tRNAi^{Met}) for initiation. As might be predicted from the initiation flow scheme in Figure 2, this reduction in protein expression tends to be non-mRNA specific (i.e., the translation of most mRNAs will be affected to the same degree), with the exception of those that might rely on the activity of initiation factor eIF2D (Dmitriev *et al.*, 2010), capable of mediating ribosomal recruitment of Met-tRNAi^{Met} in a GTP-independent manner, or which utilize initiation factor and Met-tRNAi^{Met}-independent pathways to initiate translation of their mRNAs, like cricket paralysis virus (CrPV) IRES. However, these cases are believed to be rather rare.

In a similar manner, when eIF4F exits the initiation pathway, it appears to do so as individual subunits that then need to be recombined to form active eIF4F (Merrick, 2010). During the time that eIF4E exists as a free subunit, it is susceptible to 4E-BP binding protein which sequesters the eIF4E and prevents the reformation of an active eIF4F molecule. The binding by 4E-BP is controlled by phosphorylation and is in the mTOR pathway where upregulated phosphorylation correlates with increased protein synthesis (Livingstone *et al.*, 2010; Huo *et al.*, 2011; Hinnebusch and Lorsch, 2012). In this case, phosphorylation of 4E-BP blocks its ability to compete with eIF4G for binding eIF4E and thus, active eIF4F is formed. This generally occurs under conditions favorable for growth. In contrast, when a stress occurs and there is down regulation of the mTOR pathway, 4E-BP becomes dephosphorylated and binds tightly to eIF4E reducing the pool of active eIF4F. As eIF4F is associated with mRNA activation (Figure 2, step 3), a reduction in eIF4F activity leads to a reduction in total protein expression. However, unlike eIF2, it tends to be mRNA specific in that inefficient mRNAs may have their level of expression reduced by 90% while efficiently translated mRNAs may only see a 10–20% reduction in expression; that is, the reduction of eIF4F activity drives competition between mRNAs for the limiting eIF4F available.

A secondary consequence of either regulation type is that some mRNAs are upregulated under these circumstances. mRNAs that contain an upstream ORF (uORF) are often upregulated under conditions of eIF2 phosphorylation while mRNAs that contain an internal ribosome entry site (IRES) are often upregulated under conditions where eIF4F activity is reduced (as there is reduced competition for initiation factors used in the normal cap-dependent translation). Both of these

types of responses are to trigger adaptation to the particular stress encountered.

The regulation of eEF2 activity is through the eEF2 kinase which is activated by calcium. This regulation makes sense given the release of calcium in the muscle during exercise. Under these circumstances, slowing the elongation rates reduces the energy drain of protein synthesis at a time when energy is needed for work (remembering that the incorporation of one amino acid into the growing polypeptide chain requires 4 high energy phosphates). Release of calcium under other circumstances is also consistent with energy saving (apoptosis and autophagy) under conditions of stress. This reduction in eEF2 activity slows the elongation rate of all proteins and thus reduces the expression from mRNAs in a roughly equivalent manner (as was seen with reduction of eIF2 activity and the level of ternary complexes). eEF2 is also a target of diphtheria toxin produced by *Corynebacterium diphtheriae*, that causes diphtheria. Diphtheria toxin catalyzes the transfer of NAD⁺ to a diphthamide (a modified histidine) residue in eEF2, thereby inactivating the protein. *Pseudomonas aeruginosa* exotoxin A and *Vibrio cholerae* cholera toxin also utilize a similar ADP-ribosylation mechanism to inactivate eEF2. These toxins are extremely potent and irreversibly inhibit the elongation step of protein synthesis.

mRNA Specific Regulation

The examples cited below are examples of mRNAs whose expression is regulated via a protein that binds to this specific mRNA (or a family of mRNAs) and affects its translational efficiency (through influencing initiation, elongation, or termination steps), half life or cellular localization. Due to space restrictions, we have mostly referenced these types of regulation to a series of books from Cold Spring Harbor Press that provide chapters to explore this regulation in further detail. The advantage of mRNA specific regulation, in contrast to just control of mRNA levels by regulating transcription, is the rapid change possible, often triggered by covalent modification.

Regulation of Protein Expression

5' UTR

RNA binding proteins, affects translational efficiency.

Cap-dependent translation (i.e., iron response element-binding protein) (Rouault and Harford, 2000).

IRES-mediated translation (i.e., IRES transacting acting factors or ITAFs) (Doudna and Sarnow, 2007; Elroy-Stein and Merrick, 2007; Komar and Hatzoglou, 2011).

3' UTR

RNA binding proteins affect mRNA half life (i.e., ARE or adenosine-uracil rich elements and ARE-BPs) (Darnell and Richter, 2012).

RNA binding proteins affect eIF4F/PABP interactions and translational efficiency (i.e., cytoplasmic polyadenylation element-binding protein or CPEB) (Thompson *et al.*, 2007; Gebauer *et al.*, 2012; Darnell and Richter, 2012) and block eIF4F(eIF4G)-eIF3 interaction and initiation (Jia *et al.*, 2013).

RNA binding proteins affect cellular localization of mRNA (i.e., Smaug and nanos mRNA) (Gavis *et al.*, 2007; Gebauer *et al.*, 2012; Lasko, 2012).

See also: Nucleic Acid Synthesis/Breakdown: RNA Synthesis/Function: Messenger RNA (mRNA): The Link between DNA and Protein; Transfer RNA

References

- Aitken, C.E., Lorsch, J.R., 2012. A mechanistic view of translation initiation in eukaryotes. *Nature Structural and Molecular Biology* 19, 568–576.
- Andersen, C.B., Becker, T., Blau, M., *et al.*, 2006. Structure of eEF3 and the mechanism of transfer RNA release from the E-site. *Nature* 443, 663–668.
- Baird, T.D., Wek, R.C., 2012. Eukaryotic initiation factor 2 phosphorylation and translational control in metabolism. *Advances in Nutrition* 3, 307–321.
- Bellinger, F.P., Raman, A.V., Reeves, M.A., Berry, M.J., 2009. Regulation and function of selenoproteins in human disease. *Biochemical Journal* 422, 11–22.
- Buskirk, A.R., Green, R., 2013. *Biochemistry: Getting past polyproline pauses.* *Science* 339, 38–39.
- Chakraborty, K., 2001. Translational regulation by ABC systems. *Research in Microbiology* 152, 391–399.
- Cigan, A.M., Feng, L., Donahue, T.F., 1988. tRNAⁱ(met) functions in directing the scanning ribosome to the start site of translation. *Science* 242, 93–97.
- Crick, F.H., 1966. Codon-anticodon pairing: The wobble hypothesis. *Journal of Molecular Biology* 19, 548–555.
- Darnell, J.C., Richter, J.D., 2012. Cytoplasmic RNA-binding proteins and the control of complex brain function. In: Hershey, J.W.B., Mathews, M.B., Sonenberg, N. (Eds.), *Protein Synthesis and Translational Control*. Cold Spring Harbor, NY: Cold Spring Harbor Press, pp. 281–297.
- Dever, T.E., Green, R., 2012. The elongation, termination and recycling phases of translation in eukaryotes. *Cold Spring Harbor Perspectives in Biology* 4, a013706.
- Dikstein, R., 2012. Transcription and translation in a package deal: The TISU element. *Gene* 491, 1–4.
- Dinman, J.D., 2012. Control of gene expression by translational recoding. *Advances in Protein Chemistry and Structural Biology* 86, 129–149.
- Dmitriev, S.E., Terenin, I.M., Andreev, D.E., *et al.*, 2010. GTP-independent tRNA delivery to the ribosomal P-site by a novel eukaryotic translation factor. *Journal of Biological Chemistry* 285, 26779–26787.
- Doudna, J.A., Sarnow, P., 2007. Translation initiation by viral internal ribosome entry sites. In: Mathews, M.B., Sonenberg, N., Hershey, J.W.B. (Eds.), *Translational Control in Biology and Medicine*. Cold Spring Harbor, NY: Cold Spring Harbor Press, pp. 129–153.
- Driscoll, D.M., Copeland, P.R., 2003. Mechanism and regulation of selenoprotein synthesis. *Annual Review of Nutrition* 23, 17–40.
- Elroy-Stein, O., Merrick, W.C., 2007. Translation initiation via cellular internal ribosome entry sites. In: Mathews, M.B., Sonenberg, N., Hershey, J.W.B. (Eds.), *Translational Control in Biology and Medicine*. Cold Spring Harbor, NY: Cold Spring Harbor Press, pp. 155–172.
- Gavis, E.R., Singer, R.H., Huttelmaier, S., 2007. Localized translation through messenger RNA localization. In: Mathews, M.B., Sonenberg, N., Hershey, J.W.B. (Eds.), *Translational Control in Biology and Medicine*. Cold Spring Harbor, NY: Cold Spring Harbor Press, pp. 689–717.
- Gebauer, F., Preiss, T., Hentze, M.W., 2012. From cis-regulatory elements to complex RNPs and back. In: Hershey, J.W.B., Mathews, M.B., Sonenberg, N. (Eds.), *Protein Synthesis and Translational Control*. Cold Spring Harbor, NY: Cold Spring Harbor Press, pp. 109–122.
- Gutierrez, E., Shin, B.S., Woolstenhulme, C.J., *et al.*, 2013. eIF5A promotes translation of polyproline motifs. *Molecular Cell* 51, 35–45.
- Hinnebusch, A.G., 2011. Molecular mechanism of scanning and start codon selection in eukaryotes. *Microbiology and Molecular Biology Reviews* 75, 434–467.
- Hinnebusch, A.G., Lorsch, J.R., 2012. The mechanism of eukaryotic translation initiation: New insights and challenges. *Cold Spring Harbor Perspectives in Biology* 4, a011544.

- Huo, Y., Iadevaia, V., Proud, C.G., 2011. Differing effects of rapamycin and mTOR kinase inhibitors on protein synthesis. *Biochemical Society Transactions* 39, 446–450.
- Jackson, R.J., Hellen, C.U., Pestova, T.V., 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Reviews of Molecular and Cellular Biology* 11, 113–127.
- Jha, S., Komar, A.A., 2011. Birth, life and death of nascent polypeptide chains. *Biotechnology Journal* 6, 623–640.
- Jia, J., Yao, P., Arif, A., Fox, P.L., 2013. Regulation and dysregulation of 3'UTR-mediated translational control. *Current Opinion in Genetics and Development* 23, 29–34.
- Kimball, S.R., Jefferson, L.S., 2010. Control of translation initiation through integration of signals generated by hormones, nutrients and exercise. *Journal of Biological Chemistry* 285, 29027–29032.
- Komar, A.A., Hatzoglou, M., 2011. Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states. *Cell Cycle* 10, 229–240.
- Komar, A.A., Mazumder, B., Merrick, W.C., 2012. A new framework for understanding IRES-mediated translation. *Gene* 502, 75–86.
- Kozak, M., 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44, 283–292.
- Lasko, P., 2012. mRNA localization and translational control in *Drosophila* oogenesis. In: Hershey, J.W.B., Mathews, M.B., Sonenberg, N. (Eds.), *Protein Synthesis and Translational Control*. Cold Spring Harbor, NY: Cold Spring Harbor Press, pp. 193–207.
- Livingstone, M., Atas, E., Meller, A., Sonenberg, N., 2010. Mechanisms governing the control of mRNA translation. *Physical Biology* 7, 1–8.
- Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., *et al.*, 2012. One core, two shells: Bacterial and eukaryotic ribosomes. *Nature Structural & Molecular Biology* 19, 560–567.
- Merrick, W.C., 2003. Initiation of protein biosynthesis in eukaryotes. *Biochemistry and Molecular Biology Education* 31, 378–385.
- Merrick, W.C., 2010. Eukaryotic protein synthesis: Still a mystery. *Journal of Biological Chemistry* 285, 21197–21201.
- Nanda, J.S., Cheung, Y.N., Takacs, J.E., *et al.*, 2009. eIF1 controls multiple steps in start codon recognition during eukaryotic translation initiation. *Journal of Molecular Biology* 394, 268–285.
- Nanda, J.S., Saini, A.K., Munoz, A.M., Hinnebusch, A.G., Lorsch, J.R., 2013. Coordinated movements of eukaryotic translation initiation factors eIF1, eIF1A, and eIF5 trigger phosphate release from eIF2 in response to start codon recognition by the ribosomal preinitiation complex. *Journal of Biological Chemistry* 288, 5316–5329.
- Nissen, P., Kjeldgaard, M., Thirup, S., *et al.*, 1995. Crystal structure of the ternary complex of Phe-tRNA^{Phe}, EF-Tu and a GTP analog. *Science* 270, 1464–1472.
- Pell, L., Starosta, A.L., Lassak, J., *et al.*, 2013. Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. *Proceedings of the National Academy of Sciences of the USA* 110, 15265–15270.
- Pisarev, A.V., Kolupaeva, V.G., Pisareva, V.P., *et al.*, 2006. Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. *Genes & Development* 20, 624–636.
- Pisarev, A.V., Skabkin, M.A., Pisareva, V.P., *et al.*, 2010. The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Molecular Cell* 37, 196–210.
- Pisareva, V.P., Pisarev, A.V., Komar, A.A., Hellen, C.U., Pestova, T.V., 2008. Translation initiation on mammalian mRNAs with structured 5' UTRs requires DExH-box protein DHX29. *Cell* 135, 1237–1250.
- Rouault, T.A., Harford, J.B., 2000. Translational control of ferritin synthesis. In: Sonenberg, N., Hershey, J.W.B., Mathews, M.B. (Eds.), *Translational Control of Gene Expression*. Cold Spring Harbor, NY: Cold Spring Harbor Press, pp. 655–670.
- Shatsky, I.N., Dmitriev, S.E., Terenin, I.M., Andreev, D.E., 2010. Cap- and IRES-independent scanning mechanisms of translation initiation as an alternative to the concept of cellular IRESs. *Molecules and Cells* 30, 285–293.
- Shine, J., Dalgarno, L., 1975. Terminal-sequence analysis of bacterial ribosomal RNA. Correlation between the 3'-terminal polypyrimidine sequence of 16S RNA and translational specificity of the ribosome. *European Journal of Biochemistry* 57, 221–230.
- Shoemaker, C.J., Green, R., 2011. Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proceedings of the National Academy of Sciences of the USA* 108, 1392–1398.
- Thompson, B., Wickens, M., Kimble, J., 2007. Translational control in development. In: Mathews, M.B., Sonenberg, N., Hershey, J.W.B. (Eds.), *Translational Control in Biology and Medicine*. Cold Spring Harbor, NY: Cold Spring Harbor Press, pp. 507–544.
- Thompson, S.R., 2012. Tricks an IRES uses to enslave ribosomes. *Trends in Microbiology* 20, 558–566.
- Yadavalli, S.S., Ibba, M., 2012. Quality control in aminoacyl-tRNA synthesis: Its role in translational fidelity. *Advances in Protein Chemistry and Structural Biology* 86, 1–43.
- Yeuh, A., Schneider, R.J., 2000. Translation by ribosome shunting on adenovirus and hsp70 mRNAs facilitated by complementarity to 18S rRNA. *Genes & Development* 14, 414–421.
- Zu, T., Gibbens, B., Doty, N.S., *et al.*, 2011. Non-ATG initiated translation directed by microsatellite expansions. *Proceedings of the National Academy of Sciences of the USA* 108, 260–265.

Biogenesis of Secretory Proteins

Z Chang, Peking University, Beijing, China

© 2016 Elsevier Inc. All rights reserved.

Glossary

Calnexin/calreticulin binding and release cycle It describes a cycle, in which nascent polypeptide of glycoproteins carrying the core oligosaccharide with one glucose residue ($\text{GlcNAc}_2\text{Man}_9\text{Glc}_1$) will bind to calnexin or calreticulin, but to be released when the glucose residue is removed. A reglucosylation of the oligosaccharide, occurring when the nascent polypeptide remains incompletely folded, will lead to the glycoprotein into another binding to and release from calnexin or calreticulin.

ER-associated degradation (ERAD) The process through which improperly folded substrate proteins in the endoplasmic reticulum (ER) lumen are selectively delivered back to the cytosol for eventual cleavage.

ER-related folding diseases Diseases caused by the defects of the folding and assembly of one particular protein or all secretory proteins in general in the ER, which will result in either the loss of function of a particular protein or formation of deposits of aggregated proteins in the ER lumen and thus cause ER malfunctioning, all to be harmful to the cells.

Molecular chaperone It is also called heat shock protein or stress protein. It initially interacts with its clients, being fully or partially unfolded nascent polypeptides or mature proteins, to prevent their aggregation but later release the clients, thus to facilitate their translocation, folding, assembly, or degradation.

Nascent polypeptides Polypeptides that are being synthesized but not yet folded/assembled into their native and functional structures in cells.

Oligosaccharyltransferase (OST) A multi-subunit ER membrane-associated protein complex that catalyzes the *en bloc* transferring of the core oligosaccharide of 14 monosaccharide residues ($\text{GlcNAc}_2\text{Man}_9\text{Glc}_3$), that is preformed and carried by the ER membrane-associated lipid dolichol, to the Asn residue of an Asn-X-Ser motif present in nascent polypeptides during its translation and translocation into the ER lumen.

Protein biogenesis The whole process to form functional proteins in living cells.

Protein disulfide isomerase (PDI) A protein that facilitates the formation of correct disulfide bonds in proteins by shuffling disulfide bonds, stabilizing proper intermediates, and resolving aberrant disulfide bonds via a thiol-disulfide exchange reaction, with the transient formation of mixed disulfide bond between PDI and its substrate proteins.

Retrotranslocon The ER membrane-embedded protein complex that facilitates the translocation of improperly folded substrate proteins in the ER lumen back into the cytosol for elimination. This complex has not yet been clearly characterized.

Sec61 translocon The multi-subunit protein complex that forms the conducting channel on the ER membrane, facilitating the translocation of nascent polypeptides of secretory proteins into the ER lumen.

Signal peptidase complex (SPC) The multi-subunit protease associated with the ER membrane, that specifically cleaves at C-terminal end of the signal peptides of nascent polypeptide chains.

Signal peptide A short N-terminal extension of sequences present in the nascent polypeptides of secretory proteins, that target them to the ER lumen (or certain other organelles) but be subsequently removed from the mature protein.

Signal recognition particle (SRP) A cytosolic ribonucleoprotein (containing both protein and RNA) complex that selectively delivers the ribosomes that are synthesizing secretory proteins to the ER membrane of eukaryotic cells (or the plasma membrane in prokaryotic cells).

SRP receptor The protein embedded on the ER membrane (or plasma membrane in prokaryotic cells) that interacts with the SRP that is bound to a ribosome which is synthesizing a secretory protein.

Unfolded protein response (UPR) The multiple ER to nucleus signaling transduction pathways that are switched on in responding to the accumulation of unfolded proteins in the ER lumen, resulting in the induction of a large number of proteins, whose function will strengthen the function of the ER and recover the homeostasis in ER.

For a protein to exhibit biological function, it has to be delivered to a proper cellular location after being synthesized on the cytosolic ribosomes. Proteins to be secreted to the extracellular milieu, like hormones, certain enzymes, and defense proteins (like IgG), are first translocated into the membrane-enclosed space, or lumen of the endoplasmic reticulum (ER) in all eukaryotic cells. There they are folded, assembled, and modified before being further transported to their final destination via the Golgi apparatus. Additionally, many

membrane proteins are also folded and assembled in the ER membrane for their production in cells (i.e., biogenesis), using many similar protein factors and mechanisms to be discussed here. Such a targeting or sorting process was found to be specified by intrinsic structural elements, commonly linked to the N-terminal of the precursor protein but will be removed from the mature protein. Cells that are specified to produce secretory proteins, like those of the pancreas (producing peptide hormones insulin and glucagon) and the mature B cells

(producing antibodies such as IgG) commonly possess a more extended ER system. It was the work of such scientists as George Palade and Gunter Blobel that unveiled the key role of ER in such protein biogenesis (to be explained in the Glossary) processes. For their remarkable contributions, Palade and Blobel were respectively awarded the Nobel prizes in Physiology or Medicine in 1974 and 1999 (Palade, 1975; Günter Blobel, 2000).

The ER membrane typically constitutes more than a half of the total membrane of an average animal cell, extending as a continuous enclosing of tubules and flattened sacs throughout the cytosol. The ER lumen is topologically equivalent to the outside of the cell with a composition significantly different from the cytosol, providing a unique versatile and robust compartment for the posttranslational maturation of secretory proteins. It meanwhile provides a stringent quality control checkpoint for the proteins matured there, such that only proteins, which successfully form their structures will be further transported, those otherwise will be retained and removed. As a matter of fact, most of the resident ER proteins are involved in one way or another for the biogenesis or quality control of the secretory or membrane proteins that are translocated across the ER.

Nascent Polypeptides Are Specifically Recognized and Delivered to the ER Membrane in a Co- or Posttranslational Manner

Delivery to the ER is an early and decisive step for nascent polypeptides to be exported to the extracellular space. Such an export process might be roughly divided into such key steps as recognition and delivery to the ER membrane, translocation across the membrane (or into it for the membrane proteins), formation of native structure (or folding and assembly), and departure from the ER for further transport to the final destination.

Although early biochemical and genetic studies demonstrated that nascent polypeptides are mostly transported across the ER membrane in a co-translational ('during translation') fashion, some, especially those of small size (as well as 'tail-anchored' membrane proteins, i.e. those anchored via a C-terminal tail), are believed to depend on a post-translational ('after translation') pathway. Both the co- and posttranslational translocations are considered to use the same protein translocation channel, or translocon, on the ER membrane, with the initial targeting stage differing from each other.

Nascent Polypeptides Possessing a Unique Signal Peptide Are Selectively Delivered to the ER Membrane

It has been a major scientific discovery of Blobel's laboratory that proteins rely on intrinsic structural signals, either as an extension at their N-terminal (for the secretory proteins) or as an internal sequence (for the membrane proteins), for their selective delivery to the ER. The N-terminal signal peptides of different secretory proteins commonly consist of 16–30 amino acid residues that share hardly any sequence similarity but only a tripartite organization as follows: a core of 6–12 hydrophobic residues flanked by one or more positively charged residues at its N-terminal side and polar uncharged residues at its C-terminal side (Milstein *et al.*, 1972; Blobel and Dobberstein, 1975; Martoglio and Dobberstein, 1998). It is this signal peptide that directs a nascent polypeptide to the ER membrane, with the help of a group of cytosolic factors that will be described in detail below (see Figure 1 and Table 1). Other than this rough feature of amino acid composition, much awaits to be learned about the signal peptide on how it exactly functions during the biogenesis of secretory proteins in the ER, for example, to retard the folding of the nascent polypeptides, to select and go through a particular co- or posttranslational translocation pathway.

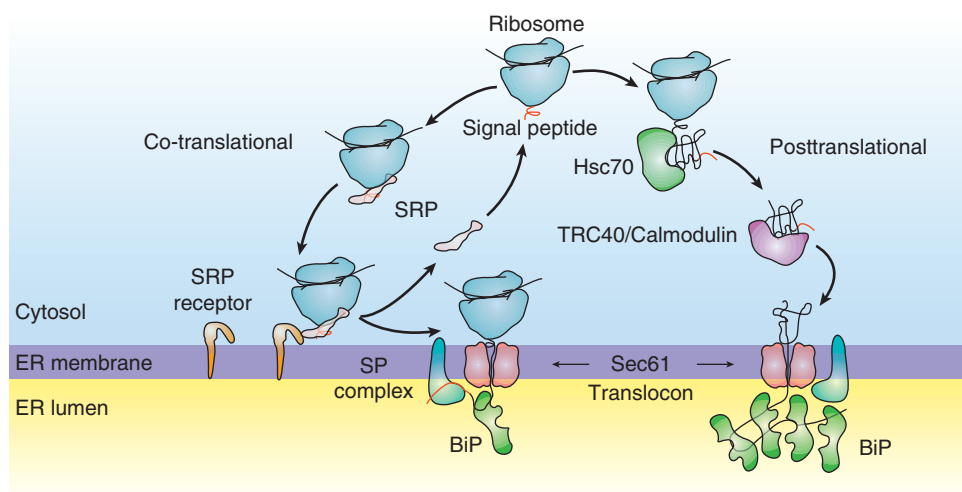


Figure 1 Major events in delivering to and translocating across the ER membrane of nascent polypeptides for the secretory proteins. Both the co- or posttranslational translocation pathways are illustrated here. BiP, an isoform of the Hsp70 chaperone family present in the ER lumen; ER, endoplasmic reticulum; Hsc70, an isoform of the Hsp70 molecular chaperone family present in the cytoplasm; SP, signal peptidase; SRP, signal recognition particle. For further details, please refer to Table 1 and the text.

Table 1 Major complexes/factors involved in the translocation of secretory protein precursors into the endoplasmic reticulum of mammalian cells

<i>Name of complexes/factors</i>	<i>Known/proposed functions</i>
<i>Signal Recognition Particle (SRP)</i>	
SRP72	Forms heterodimers with SRP68; binds to the 7S RNA
SRP68	Forms heterodimers with SRP72; binds to the 7S RNA
SRP54	Binds to the signal peptide sequence of substrate proteins; interacts with the SRP receptor located on the ER membrane
SRP19	Binds to the S domain of the 7S RNA
SRP14	Forms heterodimers with SRP9; binds to the Alu domain of the 7S RNA
SRP9	Forms heterodimers with SRP14; binds to the Alu domain of the 7S RNA
7S RNA	A central regulator of SRP; mediating global reorganization of the SRP; coupling the GTPase cycle of SRP and SRP receptor to the cargo loading and unloading events of SRP
<i>SRP Receptor</i>	
SR α	A peripheral GTP-binding membrane protein associated with the integrated membrane protein SR β interacts with SRP54 to mediate SRP docking to the SRP receptor
SR β	A GTP-binding membrane-integrated protein that interacts with SR α and mediates docking to the Sec61 translocon
<i>Cytosolic Chaperones for Posttranslational Translocation</i>	
Hsc70	A cytosolic Hsp70 molecular chaperone that keeps the nascent polypeptide competent for posttranslational translocation
Hsp40	A co-chaperone of Hsc70 that functions in posttranslational translocation
TRC40	Delivers small secretory protein precursors to the Sec61 translocon for posttranslational translocation
Calmodulin	Delivers small secretory protein precursors to the Sec61 translocon for posttranslational translocation
<i>Sec61 Translocon</i>	
Sec61 α	Forming the major part of the protein-conducting channel on ER membrane
Sec61 β	A nonessential component of the Sec61 translocon
Sec61 γ	An essential component of the Sec61 translocon
<i>Auxiliary Translocation Factors</i>	
TRAM	The 'translocating-chain associated membrane' protein; interacts with weak signal peptide sequences, regulating membrane protein biogenesis
TRAP	The 'translocon-associated protein complex'; facilitate substrate specific translocation; assists membrane protein topogenesis
RAMP4	The 'ribosome associated membrane protein 4'; regulating protein glycosylation; stabilizing newly synthesized membrane proteins under ER stress; maintaining ER homeostasis and translocation efficiency
Sec62	Mediate posttranslational translocation of small secretory proteins; binding to the ribosome exit tunnel and Sec63; stabilizing the translocon
Sec63 (ERdj2)	A co-chaperone of BiP; mediates posttranslational translocation of small secretory proteins; regulate Sec62 recruitment to the Sec61 translocon

The signal recognition particle (SRP) helps deliver a nascent polypeptide to the ER for co-translational translocation

Co-translational translocation is viewed as the most common and effective mode for the delivery of nascent polypeptides to the ER. Its advantage is derived from the fact that the nascent polypeptide chains are not released into the cytosol; thus their potential aggregation and misfolding is intrinsically avoided. In animals, this pathway of translocation is known to depend on the cytosol-located SRP, a ribonucleoprotein complex containing six different protein subunits and a 300-nucleotide RNA molecule (Walter and Blobel, 1982; Gilmore *et al.*, 1982; Tajima *et al.*, 1986; Akopian *et al.*, 2013; for details, see Table 1). For the SRP, the SRP54 subunit is the most conserved across all species and is primarily responsible for binding to the signal peptide sequence of the nascent polypeptide. Interestingly, the SRP composition in prokaryotic cells was found to be much simpler (as a common phenomenon, many molecular processes are found to be much simpler in prokaryotes than in eukaryotes), only containing one single protein that is homologous to SRP54, thus termed Ffh

(abbreviated from 'fifty four homology') and a much shorter (only 110 nucleotides) RNA molecule.

The SRP is believed to recognize and bind to the ribosome-nascent polypeptide complex once its signal peptide sequence emerges from the exit of the ribosome (Figure 1). This SRP binding was found to temporarily halt the translation process until the nascent polypeptide is delivered to the translocation machinery on the ER membrane. SRP is thought to bring the bound ribosome-nascent polypeptide complex (its 'cargo') to the ER mainly by binding to the SRP receptor that is associated with the ER membrane and composed of two different subunits (for details, see Table 1). Uniquely, both the SRP and SRP receptor (a pair of molecular matchmakers) possess GTPase activities, which are only reciprocally activated upon their binding to each other, with the subsequent GTP hydrolysis apparently leading to certain conformation modulation that releases SRP from its receptor, allowing the SRP for a new round of cargo recruitment (Figure 1). This unique 'GTPase cycle' of the SRP and SRP receptor provides an exquisite spatial and temporal coordination in targeting the ribosome-nascent polypeptide complex to the ER membrane.

This dissociation of SRP and SRP receptor should occur to result in the transferring the ribosome–nascent polypeptide complex to the translocation machinery of the ER membrane, the mechanism of which is hardly known at this time. The RNA component in SRP plays a central regulatory role for the multiple and dynamic interactions involved in this delivery process.

Cytosolic molecular chaperone proteins may help deliver completely synthesized nascent polypeptides to the ER for posttranslational translocation

Despite the aforementioned advantages of the co-translational translocation pathway, it is becoming apparent that a significant portion of secretory proteins are targeted to the ER membrane using a SRP-independent posttranslational translocation pathway (Figure 1). This was initially observed to occur for the precursor of mating factor α in yeast cells (Hansen *et al.*, 1966). It was later found to occur for secretory protein whose size is too small (e.g., peptide hormones) for the ribosome to stably retain the nascent polypeptide, whose signal peptide is not of sufficient hydrophobicity for SRP recognition, or whose signal peptide is located at the C-terminal end (for the ‘tail-anchored membrane proteins’) and will emerge only after it is completely synthesized (Johnson *et al.*, 2013). For this, cytosolic ATP-dependent molecular chaperones such as Hsc70 are needed to transiently maintain the nascent polypeptide chain in a translocation-competent form (i.e., loosely folded or unfolded) and meanwhile to prevent their aggregation. In mammals, calmodulin was found to also play a role for bringing the nascent polypeptides to the ER for their posttranslational translocation. Meanwhile, the ER lumen-located BiP (also a molecular chaperone of the Hsp70 family) is believed to act as an ATP-dependent ‘molecular ratchet’ that minimizes the passive backward movement of the nascent polypeptide chain through the translocation channel, thus to ‘pull’ it into the ER lumen (Figure 1).

Conducting Channels Made of Integrated Membrane Proteins Translocate Nascent Polypeptides across the ER Membrane

The nascent polypeptides, recruited to the ER membrane either via the co- or posttranslational pathway, are subsequently translocated into the ER lumen by using basically the same protein-conducting channel, termed the Sec61 translocon in eukaryotic cells (or SecY complex in prokaryotic cells), that is facilitated by different auxiliary factors for different nascent polypeptides (see Table 1 and Figure 1).

The Sec61 Translocon Functions as the Protein-Conducting Channel on the ER Membrane

The Sec61 translocon on the ER membrane is a highly conserved multi-subunit protein complex that consists of three subunits, Sec61 α , Sec61 β , and Sec61 γ (of 476, 96, and 68 amino acid residues, respectively) in eukaryotic cells (Görlich and Rapoport, 1993; Osborne *et al.*, 2005; Table 1). The pore

of the channel is largely formed by the large Sec61 α subunit (or the homologous Sec Y subunit in prokaryotic cells). The Sec61 translocon is associated with many other protein factors, for example, the Sec62/Sec63 proteins that are needed for the posttranslational translocation (Table 1). Several additional proteins, such as the translocating-chain associated protein (TRAM), the translocon-associating protein (TRAP), and the ribosome associated membrane protein RAMP4, also apparently assist the translocon, but their detail functions are barely known (Table 1).

The driving force for co-translational translocation is apparently provided by the cytosol-located GTP-consuming ribosome that may push the nascent polypeptide through the Sec61 translocon, while that for posttranslational translocation is provided the ER lumen-located BiP. In addition, the translocon is presumed to open only when a signal peptide sequence of a transporting substrate becomes available, otherwise remain closed to maintain the barrier of the ER membrane lipid bilayer.

The Signal Peptide Is Cleaved at a Certain Time Point after the Nascent Polypeptide Chain Reaches the ER Lumen

The N-terminal signal peptide sequence of a nascent polypeptide chain is cleaved by a specific multi-subunit signal peptidase complex (SPC), that is associated with the ER membrane. The mammalian enzyme consists of five non-identical subunits (again only one single subunit in the bacteria), all being integrated in the ER membrane (Evans *et al.*, 1986; Dalbey *et al.*, 1997). The C-terminal region of the signal peptide contains crucial structural features required for this cleavage. Statistical analysis of the amino acid residues in the cleavage site led to the formulation of the (–1, –3) rule, by which the residues at the –1 and –3 positions (–1 is immediately before the cleavage site) are typically made up of small, neutral residues, such as Ala, Gly, Cys, and Ser, more often being the Ala-X-Ala motif.

The Nascent Polypeptides Form Their Native Structures in the ER Lumen with the Help of Multiple Factors

The secretory proteins fold and assemble in the ER, where for many proteins disulfide bonds are also formed and N-glycans (i.e., oligosaccharides linked to particular Asn residues) added. The ER is highly enriched in protein factors that promote efficient protein folding and ensure only correctly folded and assembled proteins be transported to their final destination. In this regard, improperly folded/assembled proteins are retained for further rounds of folding/assembly or eventual removal (Braakman and Bulleid, 2011; Fewell *et al.*, 2001; see Table 2, and Figure 2). Accordingly, proteins that have evolved to fold in the ER fail to fold correctly in the cytosol and vice versa. Undoubtedly, folding and assembly of secretory proteins is one of the main functions evolved for ER (in addition to lipid synthesis and calcium storage). Numerous chaperones and enzymes function, often redundantly, in the ER to ensure the formation of the correct native structures for a great diversity of secretory proteins.

Table 2 Major complexes/factors involved in the processing, folding, and quality control of secretory proteins in the lumen of the endoplasmic reticulum of mammalian cells

<i>Name of factors/complexes</i>	<i>Known/proposed functions</i>
<i>Modifying Enzymes</i>	
SPC	The 'signal peptidase complex' that specifically removes the signal peptide sequence in the ER lumen
OST	The 'oligosaccharyltransferase' that adds the core N-linked oligosaccharide to the substrate glycoproteins
Glucosidase I	Removes the outmost glucose residue from the core N-linked glycan
Glucosidase II	Removes the two inner glucose residues from the core N-linked glycan; releases glycoproteins from the lectin-like chaperones CNX/CRT
UGGT	The 'UDP-glucose:glycoprotein glucosyl transferase,' a protein folding sensor that re-glucosylates incompletely folded glycoproteins
α -mannosidase	Cleave a single mannose residue from the glucose-lacking N-linked glycan to extract glycoproteins from the CNX/CRT binding and release cycle
<i>Chaperones in ER Lumen</i>	
Calnexin (CNX)	An ER membrane associated lectin-like chaperone that binds to substrate glycoproteins via their monoglucosylated N-linked oligosaccharides, leading to their retention in ER
Calreticulin (CRT)	A soluble lectin-like chaperone of identical function as calnexin
BiP (GRP78)	A key molecular chaperone of the Hsp70 family in the ER lumen, function in translocation of nascent polypeptide into ER lumen, in unfolded protein response, in protein folding, in ERAD, in holding the IgG heavy chains before associating with IgG light chains during IgG formation in mature B cells
GRP94	A molecular chaperone of the Hsp90 family essential for secretion of specific clients and also for the removal of substrate proteins via the ERAD pathway
GRP170	Nucleotide exchange factor for BiP
Si1 (BAP)	Nucleotide exchange factor for BiP
ERdj1	BiP co-chaperone functioning in protein biosynthesis
ERdj3	BiP co-chaperone functioning in protein folding and ERAD
ERdj4	BiP co-chaperone functions in ERAD
ERdj5	BiP co-chaperone; a potential PDI involved in ERAD
ERdj6	BiP co-chaperone defending cells against ER stress
<i>Peptidyl Prolyl cis-trans Isomerases (PPlases)</i>	
Cyclophilin B	PPlase
FKBP7	PPlase; associates with BiP
FKBP9, 10	PPlase
FKBP11	Pancreas-specific PPlase
<i>Protein Disulfide Isomerases (PDIs)</i>	
PDI	Formation, isomerization, and reduction of disulfide bonds; potential chaperone; also function during ERAD pathway
ERp46	Protection against ER stress
ERp57	Calnexin/calreticulin-associated oxidoreductase
ERp72	Potential oxidoreductase
P5	Potential oxidoreductase; associates with BiP and forms mixed disulfide bonds with client proteins of BiP
ERp18	Potential oxidase
ERp29	Molecular chaperones preventing premature oligomerization, stimulating secretion
ERp44	Thiol-dependent ER retention; inositol 1,4,5-triphosphate receptor regulation
Ero1 α	Oxidase, regulation of redox conditions in ER; often work with PDIs
Ero β	Oxidase, induced by UPR, high expression in pancreas
<i>ERAD-Related Proteins</i>	
P97 complex	Associates with the ER membrane, function in ERAD to facilitate substrate retrotranslocation from ER lumen to cytosol
EDEM1-3	Lectin-like chaperones that are key ER factors for the recognition and targeting of glycosylated ERAD substrates
Os9	A lectin-like chaperone that delivers glycosylated ERAD substrates to the retrotranslocon
XTP3-B	A lectin-like chaperone that delivers glycosylated ERAD substrates to the retrotranslocon
HERP	Delivers non-glycosylated BiP substrates to the ERAD pathway
Derlin1-3	Candidates of retrotranslocation channel on the ER membrane
HRD1-SEL1L	Recruits misfolded ERAD substrates or other adapter proteins (e.g., Os9 and XTP3-B) to the E3 ligase complex on the ER membrane
P97-UFD-NPL4	A cytosolic complex for retrotranslocation of ERAD substrate proteins
<i>Unfolded Protein Response (UPR) Factors</i>	
Ire1 α	The 'inositol-requiring enzyme 1'; an ER transmembrane stress transducer, activated by forming homo-oligomers when unfolded protein accumulates in the ER lumen

(Continued)

Table 2 Continued

Name of factors/ complexes	Known/proposed functions
PERK	The 'protein kinase R-like endoplasmic reticulum kinase'; an ER transmembrane kinase serving as stress transducer, activated by forming homo-oligomers when unfolded protein accumulates in ER lumen
ATF6	The 'activating transcription factor 6'; an ER transmembrane transcription factor precursor that will exist as homo-oligomers under nonstress conditions, but dissociate into monomers that will be transported to the Golgi apparatus for cleavage and activation in response to unfolded protein accumulation in the ER lumen
XBP-1	A transcription activator produced upon Ire1 α activation
ATF4	A transcription factor activated in response to PERK activation
S1P and S2P	Golgi-resident proteases that will cleave ATF6 monomer precursor to release the activated ATF6 transcription factor in the cytosol

Protein Folding and Assembly in ER Occur as an Interplay between a Polypeptide's Primary Structural Information and a Variety of Folding Factors

The ER lumen provides an optimal and specialized milieu for the maturation of the itinerant nascent polypeptides. This is made possible by the presence of a large number of molecular chaperones and enzymes. This network of folding factors work collaboratively to minimize protein aggregation, facilitate native structure formation, ensure oligomeric assembly and monitor the quality of the products to be exported. These folding factors in the ER lumen might be divided into chaperones/co-chaperones, peptidyl prolyl *cis/trans* isomerases (PPIases), protein disulfide isomerases (PDIs) and glycan-binding proteins (such as calnexin and calreticulin discussed above). Many of these factors seem to play multiple or redundant functions. For example, the molecular chaperone BiP functions in the translocation, folding, assembly, and degradation of the substrate proteins (to be discussed below; Figure 2) and many proteins function as PDIs or PPIases (Table 2).

Similar to the molecular chaperones functioning in the cytosol, the ER-resident molecular chaperones such as BiP, GRP94 (an Hsp90), GRP170 (a nucleotide exchange factor of BiP), ERdj1-6 (Hsp40 family co-chaperones of BiP) usually function to shield the hydrophobic sequences of the nascent polypeptides and thus prevent their aggregation. In addition, most other folding factors present in the ER, for example, the PPIases and PDIs seem to also exhibit chaperone activities.

Secretory Proteins Form Their Disulfide Bonds in the ER Lumen

The disulfide bonds, often present in secretory proteins and virtually absent in cytosolic proteins, are formed in the ER lumen where a relatively high oxidative redox potential is commonly maintained and PDIs are abundant (Braakman and Bulleid, 2011). The PDI was initially discovered by Christian Anfinsen as an ER enzyme that facilitates the formation of correct disulfide bonds in proteins. This occurs as a protein-based relay of oxidation/reduction reactions, involving, for example, a PDI and the oxidoreductases Ero1 (associated with the ER membrane). The PDIs function to shuffle disulfide bonds, stabilize proper intermediates, and resolve aberrant

disulfide bonds via a thiol-disulfide exchange reaction, with the transient formation of mixed disulfide bond between PDI and its substrate protein (Figure 2). A remarkable number of PDI isoforms exists in the ER lumen (Table 2), implicating either a redundancy or specification in their functions.

Glycan-Recognizing Molecular Chaperones Facilitate the Folding of Glycoproteins in the ER Lumen

A majority of the secreted proteins are glycoproteins. As an early and common modification, the core oligosaccharide of 14 monosaccharide residues (GlcNAc₂Man₉Glc₃, in which GlcNAc is N-acetylglucosamine, Man is mannose and Glc is glucose) is preformed and carried by the ER membrane-associated lipid dolichol. It is then co-translationally transferred *en bloc* to an Asn-sequon (i.e., the Asn-X-Ser/Thr sequence) present in nascent polypeptides by the multi-subunit oligosaccharyltransferase (OST) in ER lumen near the translocon (Figure 2). Such a common glycan molecule seems to act as a key marker to reflect the folding status of the proteins it linked and their binding to the lectin-like chaperones, calnexin, or calreticulin.

Subsequent glucose trimming and reglucosylation will lead the glycoproteins into a calnexin/calreticulin binding and release cycle to promote their folding (Hammond *et al.*, 1994; Helenius and Aebi, 2004; Figure 2). For this, the consecutive trimming of two glucose residues by glucosidases I and II generates GlcNAc₂Man₉Glc₁, which becomes a high affinity ligand for the two lectin-like chaperones, the membrane-associated calnexin and the soluble calreticulin. This binding allows calnexin/calreticulin to bind/retain the glycoprotein substrates in the ER lumen and facilitates a new round of folding, which meanwhile prevents their aggregation and degradation. It has been demonstrated that calnexin/calreticulin also recruits the PDI ERp57 for promoting disulfide bond formation in the glycoproteins (Figure 2).

According to current understanding, the oligosaccharide becomes no longer recognizable by calnexin/calreticulin if the last glucose residue is removed from the oligosaccharide. It is presumed that glycoproteins released from these chaperones will exit the ER if correctly folded, but will be recognized and re-glucosylated by the folding sensor UDP-glucose: glycoprotein glucosyltransferase (UGGT) if remain unfolded, thus entering another round of calnexin/calreticulin binding

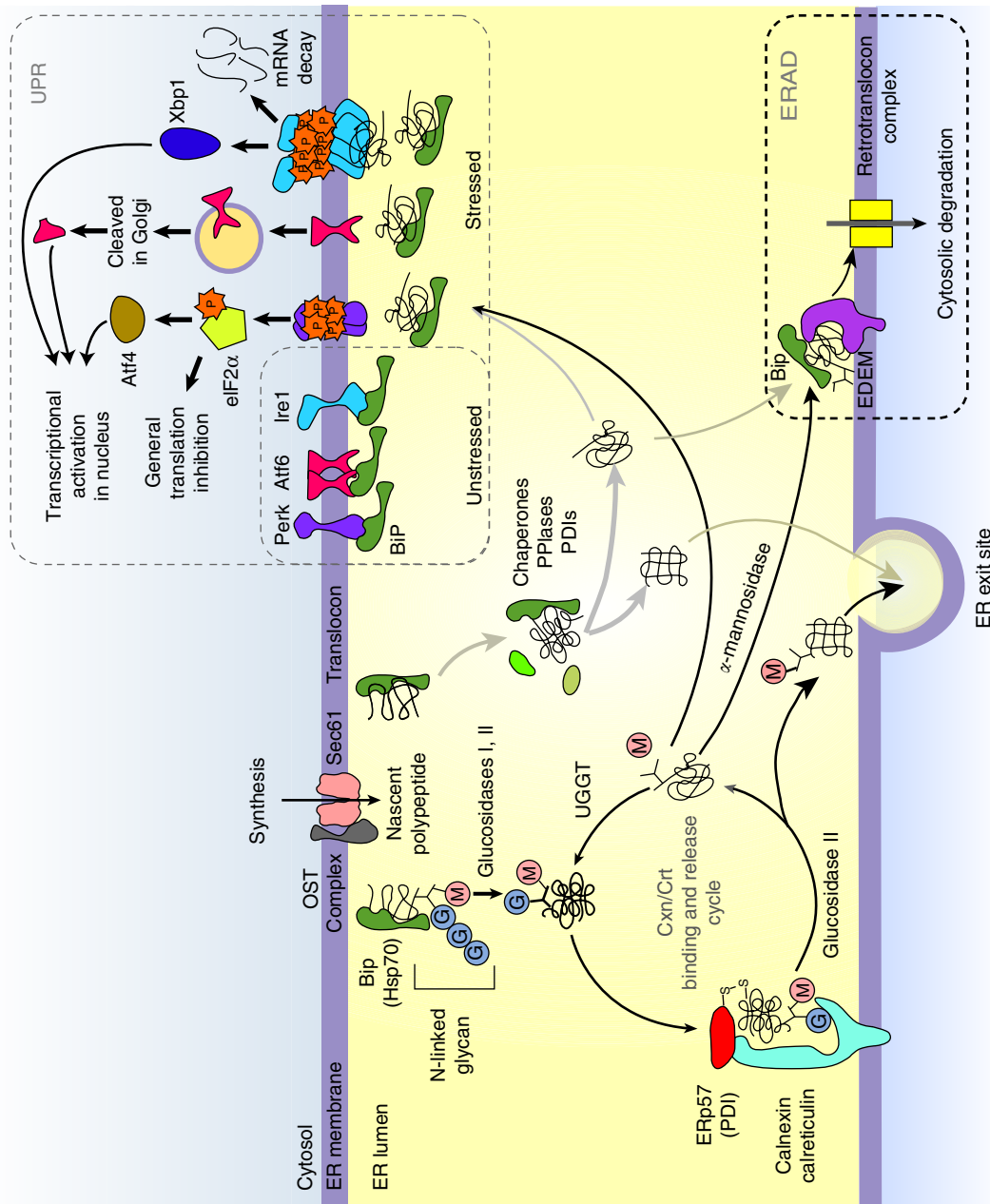


Figure 2 Major events for the folding and quality control of secretory proteins in the ER lumen. Illustrated here are the following: the calnexin/calreticulin (Cxn/Crt) binding and release cycle for the glycoproteins, the master regulatory roles of the Hsp70 member molecular chaperone Bip, the ER-associated degradation (ERAD) pathway, and the three signaling pathways for the unfolded protein response (UPR). Atf, activating transcription factor; Bip, an isoform of the Hsp70 molecular chaperone family; Crt, Calreticulin; Cxn, calnexin; EDEM, ER degradation-enhancing-mannosidase like protein; eIF, eukaryotic initiation factor; ERp57, a protein disulfide isomerase; Ire1, inositol-requiring enzyme 1; OST, oligosaccharyl transferase; PDI, protein disulfide isomerase; Perk, protein kinase R-like endoplasmic reticulum kinase; UGGT, UDP-glucose:glycoprotein glucosyl transferase; Xbp-1, X-box binding protein 1, a transcription factor. For further details, please refer to [Table 2](#) and the text.

(Figure 2). Certain unfolded forms of the glycoproteins, either due to an overtime existence or an extensive structure distortion, may eventually get its mannose residues trimmed and thus become recognizable by other ER lectin-like chaperones that deliver them for degradation (see Figure 2 and below). How exactly these different fates of the glycoproteins (and nonglycoproteins) are determined in the ER lumen is still unclear. For glycoproteins, folding, trafficking, and degradation apparently rely on a common machinery and are thus highly integrated.

The Folding Factors in the ER Lumen Work as a Network or to Form Complexes in Folding Each Unique Client Protein

Studies indicate that the different folding factors, including the chaperones, PDIs and PPIases, do not work alone, but as a dynamic network for the maturation of each unique substrate proteins. For example, the BiP works sequentially with Hsp40 family co-chaperones, nucleotide exchange factors, as well as the Hsp90 protein (GRP94) in fulfilling its multiple functions for the translocation, folding and assembly of nascent polypeptides, removal of misfolded proteins, and transduction of ER stress signals to the nucleus. On the other hand, chaperones such as BiP, GRP94, calreticulin, ERp72, and cyclophilin have been found to form complexes with such large secretory proteins as immunoglobulins, apolipoprotein B-100, and thyroglobulin. Apparently, they come together to meet the 'needs' of individual substrate proteins, as well as on the stage of folding that is being promoted.

The Folding and Assembly of IgG in the ER Lumen of the B Cell Involves Many Different Types of Folding Factors

The IgG molecule is a large and complex secretory protein that is assembled from two light and two heavy chains of multiple Ig (Immunoglobulin) domains, forming multiple intra- and interchain disulfide bonds. The biogenesis of IgG in effector B cells would conceivably depend on the functioning of many of the folding factors we discuss here (Haas and Wabl, 1984; Feige *et al.*, 2009, 2010). As a matter of fact, the key ER lumen molecular chaperone BiP was originally discovered as an 'Immunoglobulin heavy chain binding protein' (thus the name BiP) and was identified by virtue of its association with the unassembled IgG heavy chains, preventing the premature escape or aggregation of the incompletely assembled IgG molecules (Haas and Wabl, 1984).

Although a detail pathway remains to be defined, the folding and assembly of IgG apparently are considered to involve the following major events (Feige *et al.*, 2009, 2010). Folding of the individual Ig domains and formation of intra-chain disulfide bonds seem to begin during their co-translational translocation. BiP transiently interact with most of the domains during their folding. All constant domains except C_H1 ('constant domain 1 in the heavy chain') and most of the variable domains seem to fold autonomously but need PPIases to promote the rate-limiting correct formation of certain peptidyl prolyl *cis-trans* isoforms on the way to their final native state. The dimerization of two heavy chains is induced by their folded C_H3 structures, which will be further fixed by

the formation of interchain disulfide bonds between the hinge regions of the heavy and light chains. Interestingly, the C_H1 domain seems to remain unfolded, with its disulfide bonds unformed and thus stably bound to BiP until it interacts with the C_L domain ('constant domain of the light chain') that simultaneously displaces BiP. In other words, the folding of the C_H1 domain will be initiated upon its binding to the C_L. This includes the formation of the correct isoforms of a prolyl peptide bond in the C_H1 domain as promoted by a PPIase. A disulfide bond will then be formed between the interacting C_H1 and C_L domains, rendering the IgG molecules ready for secretion into the extracellular space. Chaperones, PDIs, and PPIases contribute to these multiple individual folding/assembly steps in IgG biogenesis in a cooperative manner.

A Comprehensive Quality Control System in the ER Lumen Ensures that Only Correctly Folded/Assembled Proteins Are Transported with the Misfolded/Misassembled Ones to be Degraded

Protein folding and assembly occur through non-covalent weak interactions and are thus inherently error prone. Folding factors have evolved not only to facilitate protein folding and assembly, but also to ensure a high standard of quality control, such that only correctly formed products be delivered to their final destination and those failed to reach the standard to be retained for another round of folding/assembly attempts or eventual removal (Braakman and Bulleid, 2011; Fewell *et al.*, 2001; Figure 2). As molecular sensors, the protein folding factors often play dual roles of both assisting the folding process and dispatching improperly folded proteins for destruction. As a result of such 'proof-reading,' the number of accumulated errors in the secreted proteins that the cells ultimately deploy is extremely low. Prolonged retention of misfolded and incompletely folded proteins in the ER actuates a dual stress response, which will help to remove the improperly folded proteins through degradation (see below) on the one hand, and to induce a comprehensive adaptive unfolded protein response (UPR) to increase the level of the quality control proteins (Figure 2).

Improperly Folded Proteins Are Disposed Through the ER-Associated Degradation (ERAD) Pathway

In spite of efforts to identify active proteases in the ER, none have been found, except the rather specific signal peptidase. Strikingly, the improperly folded proteins in the ER lumen were recently found to be degraded in the cytosol through a process termed ER-associated degradation (ERAD in short). For ERAD to occur, the unfolded and unassembled substrate proteins in the ER lumen are actually retrograded back to the cytosol, where they are ubiquitinated and degraded via the proteasome system (Werner *et al.*, 1996; Hiller *et al.*, 1996; Vembar and Brodsky, 2008). The ER membrane protein responsible for the retrotranslocation of substrate proteins from ER to cytosol has been intensively searched, but still with no clear consensus, with candidates include the Sec61 translocon itself and the Derlin proteins (Table 2).

According to the current understanding, the ERAD pathway occurs as follows (Figure 2). The misfolded substrate proteins in the ER lumen are first recognized by BiP for non-glycosylated proteins, by calnexin or calreticulin for glycosylated proteins, or by PDI for disulfide bond-containing proteins. The substrates are then targeted to the retrotranslocation machinery termed the 'retrotranslocon', through which the unfolded substrates are delivered back into the cytosol, the driving energy may derive from the ATP hydrolysis by the P97 protein complex (an AAA + ATPase). Upon exiting the retrotranslocon, the polypeptide chain is polyubiquitinated by an E₃ ubiquitin ligase, after which, further retrotranslocation is aided by the cytosolic ubiquitin-binding proteins. Once a polyubiquitinated substrate is displaced into the cytosol, it is then degraded by the 26S proteasomes.

Extensive trimming of the N-glycans in a misfolded glycoprotein leads to an increased exposure of hydrophobicity, marking it for ERAD. For this, when the three glucose and two mannose residues are trimmed from the N-glycan, the Man₇-GlcNAc₂ oligosaccharide-carrying glycoprotein becomes more likely to be recognized by such lectins as EDEM1 ('ER degradation-enhancing α -mannosidase like protein 1') or OS9 (Table 2), which then guides the misfolded substrate proteins to an E₃ ligase complex and thus the ERAD pathway (Figure 2). Molecular cues that direct different folding sensors to various nonnative substrate proteins remain largely unknown. One model suggests that the efficiency of productive folding and trafficking is not defined by a single feature, but rather by the combination of multiple ones, including, for example, protein stability and folding rate.

UPR Prepares the Cells to Recover from an ER Stress

When the ER is burdened with unfolded proteins, a comprehensive so-called UPR will be initiated to reestablish homeostasis through a sophisticated transcriptionally and translationally regulated signaling network (Kozutsumi *et al.*, 1988; Cox *et al.*, 1993; Haze *et al.*, 1999; Harding *et al.*, 1999; Moore and Hollien, 2012; Figure 2). This will result in a general remodeling of the whole secretory pathway, such that the capacity of the ER increases and the ER may even become enlarged to accommodate the increased requirements for its function.

The UPR is known to occur via three different ER to nucleus signaling pathways in mammalian cells (Figure 2). In each pathway, a different ER membrane-integrated protein functions as the stress sensor and transducer, which are respectively found to be the Ire1, the Perk, and the Atf6 (Table 2). The BiP protein again seems to play a key role by binding to and inhibiting them in the ER lumen under nonstress conditions, but dissociating from them under stress conditions. The dissociation of BiP will result in the homo-oligomerization of Ire1, which will lead to the autophosphorylation and activation of the ribonuclease domain of Ire1, promoting the synthesis of the transcription factor XBP-1 (Cox *et al.*, 1993; Moore and Hollien, 2012). Similarly, the dissociation of BiP will result in homo-oligomerization of Perk, which will attenuate cell protein translation in general, but conversely enhances translation of Atf4 (analogous to the heat shock response), another transcription factor (Haze *et al.*, 1999; Moore and Hollien, 2012). In contrast to Ire1 and Perk, the dissociation of BiP results in the

disassembly of the Atf6 homo-oligomer into monomers, which will then be translocated to the Golgi apparatus where it is cleaved by the protease S1P ('site-1 protease') or S2P (Harding *et al.*, 1999; Moore and Hollien, 2012). The cytosolic domain cleaved from Atf6 also functions as a transcription factor. The three transcription factors that are activated through the three signaling pathways, XBP-1, Atf4, and Atf6, will subsequently induce the expression of a large number of proteins that are involved in such processes as protein biogenesis, protein quality control, and ERAD in ER. Together, these transcriptional regulations, in combination with the additional Ire1-mediated decay of mRNAs associated with the ER membrane and Perk-mediated general translational suppression, are believed to restore and strengthen ER function under stress conditions. The UPR signaling pathways will be attenuated and turned off upon the alleviation of the ER stress condition, apparently through the action of a few negative feedback loops. Although the general output of the UPR is to restore ER homeostasis, sustained and unresolved ER stress may induce apoptosis through the intrinsic mitochondrial apoptosis pathway.

The BiP Chaperone Is a Central Player for Protein Biogenesis and Quality Control in the ER

BiP was initially identified as a glucose-regulated protein named as Grp78 and independently as an immunoglobulin heavy chain binding protein (Haas and Wabl, 1984; Munro and Pelham, 1986). It was then demonstrated to be an ER-located member of the ATP-dependent Hsp70 family of molecular chaperones and play a central role for the protein biogenesis and quality control in the ER (Dudek *et al.*, 2009).

The major roles of BiP in ER may be summarized as follows (Figure 2). First, it facilitates the translocation of nascent polypeptides into the ER lumen by acting as a molecular ratchet (see above). For this, BiP is apparently recruited to the translocon by binding to its co-chaperone Sec63, a member of the Hsp40 family. This transient binding of BiP meanwhile will inhibit the intra- or intermolecular interaction of the nascent polypeptides to facilitate their subsequent maturation in the ER lumen. Second, BiP is also known to facilitate the disposal of those more permanently misfolded and unassembled proteins. For this, BiP directs the substrate proteins to the ER-associated degradation pathway. How BiP distinguishes the incompletely folded and the permanently misfolded polypeptides is an issue remains unresolved. One proposal is that the ER is divided into different sub-domains in which folding and degradation occur in a compartmentalized manner, thus each could be worked upon by a distinct set of chaperones and co-chaperones. For example, in yeast, the BiP co-chaperones ERdj3 or ERdj6 apparently mediates folding, while ERdj4 or ERdj5 mediates degradation (Table 2). Thirdly, BiP is also known to play key modulatory roles in the three signal transduction pathways between the ER lumen and the nucleus for the UPR (Figure 2).

The Malfunctioning of Protein Folding in ER Are Known to Cause Folding Diseases

The ER is a key cellular site for the biogenesis of many secretory (and membrane proteins). Failed folding, transport, and

degradation in ER have been found to be the basis of many human diseases termed ER-related folding diseases (Aridora and Hannan, 2000; Hebert and Molinari, 2007). These diseases can be roughly categorized into three groups. The first group is due to the retention and subsequent degradation of mutated cargo protein molecules that fail to be exported by the ER machinery. One disease for this is the chronic lung defect called emphysema, which is caused by a mutation leading to the replacement of Glu342 by a Lys residue in the α 1-antitrypsin, a glycoprotein inhibitor of the elastase (a serine protease) in the lung. This mutant protein is retained in the ER of hepatocytes, a major site of α 1-antitrypsin production and secretion, and is subsequently degraded. This lack of inhibition of elastase leads to the degradation of the elastin protein in the lungs, causing emphysema. The second group is due to the accumulation of substrate proteins that are retained but inefficiently degraded in the ER. This will initiate the ER UPR, the signaling that may culminate in apoptosis of the host cell. The Glu324 to Lys mutant protein of the α 1-antitrypsin described above was also found to accumulate in the liver cells, forming intracellular deposits that damage the liver cells and cause such liver disease as cirrhosis. The third group is due to defects in the ER transport machinery or the UPR signaling.

A great number of human diseases have been related to the biogenesis defects of different secretory (and membrane proteins) in the ER, the detail understanding of the molecular mechanism of these diseases will certainly allow us to find ways to prevent or treat these diseases.

See also: Interorganellar Communication: Interplay and Processes: Endoplasmic Reticulum Stress in Disease; ER–Golgi Transport; N-Linked Glycans (N-Glycans). Molecular Principles, Components, Technology, and Concepts: Proteins: Posttranslational Modifications: Key Players in Health and Disease. Protein Synthesis/Degradation: Protein Degradation – Intracellular: Endoplasmic Reticulum-Associated Degradation and Protein Quality Control. Protein Synthesis/Degradation: Translation: Components, Initiation, Elongation, Termination, and Regulation; Regulated Proteolysis of Signaling Molecules: The Proprotein Convertases

References

- Akopian, D., Shen, K., Zhang, X., Shan, S.O., 2013. Signal recognition particle: An essential protein-targeting machine. *Annual Review of Biochemistry* 82, 693–721 (Review).
- Aridora, M., Hannan, L.A., 2000. Traffic jam: A compendium of human diseases that affect intracellular transport processes. *Traffic* 1, 836–851 (Review).
- Blobel, G., Dobberstein, B., 1975. Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *Journal of Cell Biology* 67, 835–851.
- Braakman, I., Bulleid, N.J., 2011. Protein folding and modification in the mammalian endoplasmic reticulum. *Annual Review of Biochemistry* 80, 71–99 (Review).
- Cox, J.S., Shamu, C.E., Walter, P., 1993. Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase. *Cell* 73, 1197–1206 (Discovery of IRE1).
- Dalbey, R.E., Lively, M.O., Bron, S., van Dijk, J.M., 1997. The chemistry and enzymology of the type I signal peptidases. *Protein Science* 6, 1129–1138 (Review).
- Dudek, J., Benedix, J., Cappel, S., *et al.*, 2009. Functions and pathologies of BiP and its interaction partners. *Cellular and Molecular Life Sciences* 66, 1556–1569 (Review).
- Evans, E.A., Gilmore, R., Blobel, G., 1986. Purification of microsomal signal peptidase as a complex. *Proceedings of the National Academy of Sciences of the United States of America* 83, 581–585.
- Feige, M.J., Groscurth, S., Marciniowski, M., *et al.*, 2009. An unfolded C_{H1} domain controls the assembly and secretion of IgG antibodies. *Molecular Cell* 34, 569–579.
- Feige, M.J., Hendershot, L.M., Buchner, J., 2010. How antibodies fold. *Trends in Biochemical Sciences* 35, 189–198 (Review).
- Fewell, S.W., Travers, K.J., Weissman, J.S., Brodsky, J.L., 2001. The action of molecular chaperones in the early secretory pathway. *Annual Review of Genetics* 35, 149–191 (Review).
- Gilmore, R., Walter, P., Blobel, G., 1982. Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor. *Journal of Cell Biology* 95, 470–477.
- Görlach, D., Rapoport, T.A., 1993. Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell* 75, 615–630.
- Günter Blobel, G., 2000. Protein targeting (Nobel Lecture). *ChemBioChem* 1, 86–102 (Review).
- Haas, G., Wabl, M., 1984. Immunoglobulin heavy chain binding protein. *Nature* 306, 387–389.
- Hammond, C., Braakman, I., Helenius, A., 1994. Role of N-linked oligosaccharide recognition, glucose trimming, and calnexin in glycoprotein folding and quality control. *Proceedings of the National Academy of Sciences of the United States of America* 91, 913–917.
- Hansen, W., Garcia, P.D., Walter, P., 1966. *In vitro* protein translocation across the yeast endoplasmic reticulum: ATP-dependent posttranslational translocation of the prepro- α -factor. *Cell* 45, 397–406.
- Harding, H., Zhang, Y., Ron, D., 1999. Translation and protein folding are coupled by an endoplasmic reticulum resident kinase. *Nature* 397, 271–274 (Discovery of PERK).
- Haze, K., Yoshida, H., Yanagi, H., Yura, T., Mori, K., 1999. Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. *Molecular Biology of the Cell* 10, 3787–3799.
- Hebert, D.N., Molinari, M., 2007. In and out of the ER: Protein folding, quality control, degradation, and related human diseases. *Physiology Review* 87, 1377–1408 (Review).
- Helenius, A., Aebi, M., 2004. Roles of N-linked glycans in the endoplasmic reticulum. *Annual Review of Biochemistry* 73, 1019–1049 (Review).
- Hiller, M.M., Finger, A., Schweiger, M., Wolf, D.H., 1996. ER degradation of a misfolded luminal protein by the cytosolic ubiquitin-proteasome pathway. *Science* 273, 1725–1728.
- Johnson, N., Powis, K., High, S., 2013. Post-translational translocation into the endoplasmic reticulum. *Biochimica et Biophysica Acta* 1833, 2403–2409 (Review).
- Kozutsumi, Y., Segal, M., Normington, K., Gething, M.J., Sambrook, J., 1988. The presence of misfolded proteins in the endoplasmic reticulum signals the induction of glucose-regulated proteins. *Nature* 332, 462–464.
- Martoglio, B., Dobberstein, B., 1998. Signal sequences: More than just greasy peptides. *Trends in Cell Biology* 8, 410–415 (Review).
- Milstein, C., Brownlee, G.G., Harrison, T.M., Matthews, M.B., 1972. A possible precursor of immunoglobulin light chains. *Nature New Biology* 239, 117–120.
- Moore, K.A., Hollien, J., 2012. The unfolded protein response in secretory cell function. *Annual Review of Genetics* 46, 165–183 (Review).
- Munro, S., Pelham, H.R.B., 1986. An Hsp70-like protein in the ER: Identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. *Cell* 46, 291–300.
- Osborne, A.R., Rapoport, T.A., van den Berg, B., 2005. Protein translocation by the Sec61/SecY channel. *Annual Review of Cell and Developmental Biology* 21, 529–550 (Review).
- Palade, G., 1975. Intracellular aspects of the process of protein synthesis. *Science* 189, 347–358 (Review).
- Tajima, S., Lauffer, L., Rath, V.L., Walter, P., 1986. The signal recognition particle receptor is a complex that contains two distinct polypeptide chains. *Journal of Cell Biology* 103, 1167–1178.
- Vembar, S.S., Brodsky, J.L., 2008. One step at a time: Endoplasmic reticulum-associated degradation. *Nature Reviews Molecular Cell Biology* 9, 944–957 (Review).
- Walter, P., Blobel, G., 1982. Signal recognition particle contains a 7S RNA essential for protein translocation across the endoplasmic reticulum. *Nature* 299, 691–698.
- Werner, E.D., Brodsky, J.L., McCracken, A.A., 1996. Proteasome-dependent endoplasmic reticulum associated protein degradation: An unconventional route to a familiar fate. *Proceedings of the National Academy of Sciences of the United States of America* 93, 13797–13801.

The Protein Biosynthetic Machinery of Mitochondria

LL Spemulli, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

© 2016 Elsevier Inc. All rights reserved.

General Background

Mitochondria are subcellular organelles found in eukaryotic organisms that function as the powerhouses of the cell. They produce about 90% of the energy used by higher cells as well as carrying out additional metabolic and cellular functions. Mitochondria are thought to have developed from endosymbiotic bacteria. These bacteria are thought to have been engulfed by other cells during the evolution of eukaryotes. They transferred much of their original genetic information to the nucleus leaving a small amount in the residual structures, mitochondria. These organelles retain a small amount of DNA and carry out transcription and translation for the expression of the limited number of genes remaining within the organelle.

Information Content of Mitochondrial DNA

The mitochondrial DNA from about 3000 organisms has been sequenced. In most mammals, mitochondrial DNA is a very compact double-stranded DNA about 16 000 base pairs in length. This DNA generally codes for 2 ribosomal RNAs (rRNAs), 22 transfer RNAs (tRNAs) and 13 polypeptides. The proteins encoded are all membrane embedded subunits of the respiratory chain complexes including 7 subunits of Complex I (NADH-ubiquinone oxidoreductase), one subunit of Complex III (ubiquinone-cytochrome c oxidoreductase); three subunits of complex IV (cytochrome oxidase) and two subunits of complex V (the ATP synthase).

Plant mitochondrial genomes are much larger than those found in metazoans ranging from 200 to 2400 kilobase pairs (kbp). In some cases there is a large master circle and other various sized DNAs that appear to be derived from it (Scheffler, 2008). The arrangement of the genes is quite variable but the information content is limited. For example, the *Arabidopsis* mitochondrial genome is just under 370 000 base pairs and consists of three different configurations of circular molecules. It encodes 3 rRNAs, 22 tRNAs, the standard proteins of the respiratory chain complexes with an extra subunit of complex V included. Plant mitochondrial genomes often encode a few ribosomal proteins. There are introns in certain genes and RNA editing must occur.

Fungal mitochondrial genomes also show significant variability although most are between 40 and 60 kbp. The best studied of these is from the yeast *Saccharomyces cerevisiae* which is 85 kbp and encodes 2 rRNAs, 24 tRNAs, 7 proteins of the respiratory chain and ATP synthase as well as a ribosomal protein. Yeast mitochondrial genes often contain introns and certain yeast strains encode additional proteins involved in the removal of these introns.

Components of the Translational System

Mitochondrial Messenger RNAs

Mitochondrial DNA is transcribed to create messenger RNAs (mRNAs) used by the organellar translational apparatus (Figure 1). In humans and other mammals, both strands of the DNA are transcribed into long polycistronic RNAs, transcripts carrying the coding information for multiple products. These transcripts are processed into 9 monocistronic and 2 dicistronic mRNAs before translation. The polycistronic transcripts also include the coding information for the rRNAs and tRNAs. In many cases, the tRNAs appear to serve as 'punctuation marks' specifying where endonucleases should cut the primary transcript to release the appropriate RNAs (Ojala *et al.*, 1981). In mammals, the mitochondrial mRNAs are quite unusual. They have an almost complete lack of 5' and 3' untranslated nucleotides. In humans the translational start codon is directly at the 5' end in 8 of the eleven 5' reading frames. It is located within 3 nucleotides of the 5' end of the other three mRNAs (Anderson *et al.*, 1982; Montoya *et al.*, 1981). The unusual structure of these mRNAs argues that the mechanism used by mitochondrial ribosomes to recognize the translational start site differs significantly from the mechanisms found in bacteria and the eukaryotic cytoplasm. In prokaryotes, the 16S rRNA contains a polypyrimidine tract near the 3' end which plays an important role in initiation complex formation by hydrogen bonding to a polypurine sequence (called the Shine/Dalgarno sequence) in the mRNA. In contrast, eukaryotic cytoplasmic mRNAs have a 5' cap structure which facilitates the binding of the 40S subunit to the mRNA. The 40S subunit then migrates in a 5' to 3' direction and usually initiates at the first AUG codon (Merrick, 1992; Kozak, 1992). Since mitochondrial mRNAs are not capped (Wolstenholme, 1992) and lack a Shine/Dalgarno sequence, neither the prokaryotic nor cytoplasmic mechanism for translational initiation can occur in this system. Mammalian mitochondrial mRNAs also generally lack significant 3' untranslated regions (UTRs) and, in many cases, the termination codon is completed by the polyadenylation of the 3' end of the mRNAs (Hirsch and Penman, 1974). The poly(A) tail at the 3' end of the mRNA is generally 40–60 residues in length and is believed to increase the stability of the mRNA.

In contrast to the mRNAs in mammals, plant mitochondrial mRNAs often have extensive 5' and 3' UTRs. They undergo extensive processing including the generation of multiple 5' ends. RNA editing (C to U) is observed and both cis- and trans-splicing is used for the removal of introns in the protein coding genes (Scheffler, 2008). The 3' ends of these mRNAs may be polyadenylated; however, unlike the situation in mammals, polyadenylation appears to serve as a signal for the degradation of the mRNA rather than serving to stabilize them (Chang and Tong, 2012).

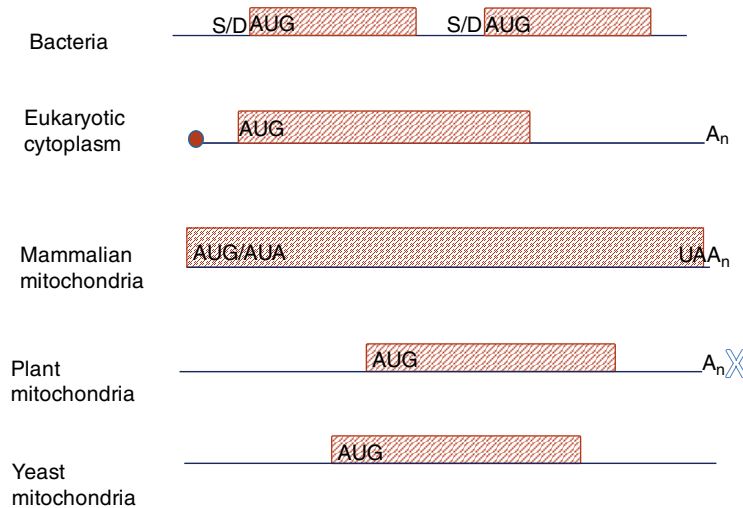


Figure 1 Comparison of mRNAs from mitochondria with those from bacteria and the eukaryotic cell cytoplasm: the black lines indicate untranslated regions. The lengths of these lines correspond roughly to the sizes of the 5' and 3' untranslated regions. The red boxes are open reading frames (ORFs) encoding proteins. The sizes on the ORFs vary considerably. S/D indicates the presence of a Shine/Dalgarno sequence. Codons used for initiation (AUG or AUA) are shown. The small circle on the eukaryotic mRNA represents the 5' cap structure. The X indicates that polyadenylation signals destabilization of the mRNA.

In *S. cerevisiae*, the processed mRNAs are monocistronic and have 5' UTR sequences that play a role in translational initiation as described below. In *S. cerevisiae* the mitochondrial mRNAs are not polyadenylated (Chang and Tong, 2012).

A particularly interesting situation is observed with the kinetoplast mRNAs observed in the trypanosomatid protozoa including *Leishmania major* and *Trypanosoma brucei*. The kinetoplast is a network of circular DNA molecules contained within a large mitochondrion located near the base of the flagellum. In these organisms, initial polycistronic mitochondrial precursor RNAs are endonucleolytically processed. Many of the pre-mRNAs produced are extensively edited through a process that inserts and deletes uridines at the direction of small guide RNAs. Mature mRNAs in the kinetoplast have long poly(A/U) tails of 200–300 residues that stabilize them and make them competent for translation (Aphasizhev and Aphasizheva, 2013).

Mitochondrial tRNAs

Genetic code in mitochondria

Mitochondrial genomes of many organisms use a variation of the universal genetic code. Further, there is clearly not a universal mitochondrial genetic code and codon assignments vary in different lineages (Watanabe, 2010). One of the most interesting changes is the use of the Ile codon AUA for Met in most metazoans, making AUA an important initiation codon. Further, all animals have reassigned the 'universal' termination codon UGA as a Trp codon. In some cases, novel post-transcriptional modifications of mitochondrial tRNAs have developed to facilitate reading the altered code. Unlike the mitochondria of many organisms, the mitochondria of green plants use the universal code.

Import of tRNAs

Mammalian mitochondria, with the exception of certain marsupials, encode all 22 tRNAs that are required for the translation of the modified genetic code used in this organelle. Of particular note is that animal mitochondrial DNA encodes a single gene for tRNA^{Met} that functions in both initiation and elongation. Normally, protein biosynthetic systems have two tRNA^{Met} species, one used for initiation and the other for chain elongation (Schneider, 2011).

In contrast to animal systems, plant mitochondria often import a number of their tRNA species. However, the number and the amino acid specificity of the imported tRNAs varies widely from one plant system to another even among closely related species. The underlying biological basis for the differences observed is not clear, but suggests that the evolution of the import systems undergoes rapid evolutionary change (Schneider, 2011).

Although *S. cerevisiae* mitochondrial DNA has tRNA genes for all 20 amino acids, this organism imports a tRNA^{Lys} and possibly two tRNA^{Gln} species from the cytosol (Duchene et al., 2009). The mitochondrial genomes of trypanosomatids do not contain tRNA genes and these organisms import all of their tRNAs from the cytosol (Tan et al., 2002).

Structural features of mitochondrial tRNAs

Animal mitochondrial tRNAs have a number of unusual features distinguishing them from the tRNAs found in other systems. They are shorter and structurally weaker than their prokaryotic or eukaryotic cytoplasmic counterparts. Many of these tRNAs can be folded into an L-shaped tertiary structure, but the D-loop/T-loop interactions in the elbow region do not exist or are considerably different from those of canonical tRNAs. Mitochondrial tRNAs have been classified into five

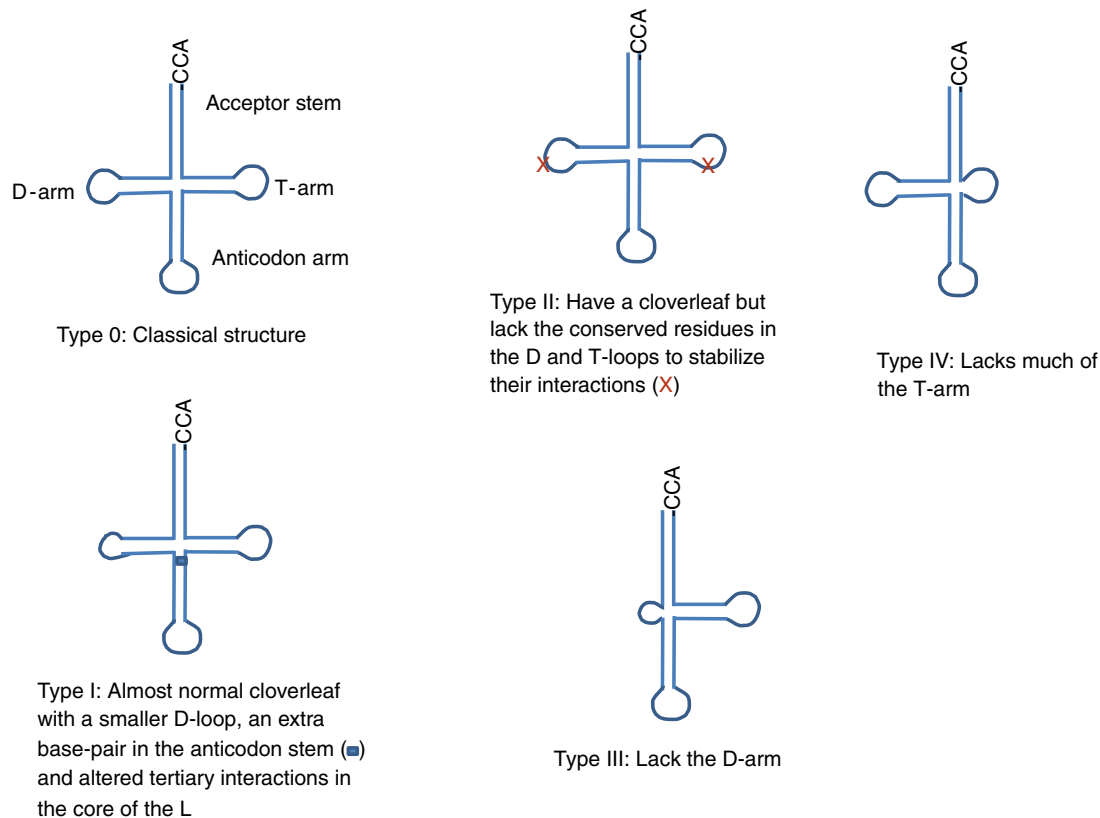


Figure 2 Classes of tRNAs found in animal mitochondria. Type 0 is considered the classical cloverleaf structure with acceptor stem, D-arm, T-arm and anticodon stem/loop found in the tRNAs for prokaryotes and the eukaryotic cell cytoplasm. Many mitochondrial systems have tRNAs that deviate from this classical structure as indicated in types I-IV.

groups on the basis of their structural characteristics as summarized in [Figure 2](#) ([Watanabe, 2010](#)).

In contrast to observations in animals, mitochondrial tRNAs in higher plants have retained the canonical tRNA structure observed in prokaryotes and the eukaryotic cytosol. Since many plant species import a number of cytosolic tRNAs for use in the mitochondrion, this conservation of structure is expected. Fungal and trypanosomatid mitochondrial tRNAs also generally fold into the classical structure.

Aminoacylation of mitochondrial tRNAs

The genes for 38 aminoacyl-tRNA synthetases (aaRS) have been annotated in the human genome. Of these, 17 encode mitochondrial aaRS distinct from the corresponding cytosolic enzymes. Two genes encode both mitochondrial and cytosolic forms. There is no apparent gene for GlnRS and Gln-tRNA^{Gln} is thought to be generated through transamidation of Glu-tRNA^{Gln}.

In the *S. cerevisiae* genome, the genes for 35 aaRS have been convincingly annotated. In this organism, two genes are found for 15 aaRS; one of these encodes the cytosolic form of the enzyme and the other codes for the mitochondrial form. Five aaRS are shared between the two compartments. In *T. brucei* and *L. major*, the same aaRS is used in both the cytosol and mitochondria. This idea is compatible with the import of cytosolic tRNA for use in the kinetoplast ([Duchene et al., 2009](#)).

The situation in plants can be more complex since there is an additional translational system in the chloroplast. No aaRS are encoded in the mitochondrial or chloroplast DNA. Specifics vary in different plant species but, in general, the aaRS can be divided into several classes: (1) enzymes used only in the cytosol, (2) aaRS imported into both mitochondria and chloroplasts, (3) enzymes that function both in the cytosol and mitochondria, (4) enzymes that are localized only in chloroplasts. In general, there do not appear to be enzymes that are targeted solely to mitochondria.

Mitochondrial tRNA and aaRS in human disease

Genetic diseases leading to mitochondrial malfunction are classified into two groups: (1) mutations in nuclear genes encoding mitochondrial proteins and (2) mutations in mitochondrial DNA. This latter class is characterized by maternal inheritance. Diseases arising from mitochondrial dysfunction display a wide variety of clinical symptoms generally involving tissues that have high energy requirements including skeletal and cardiac muscle, neurological tissues and the renal and endocrine systems ([Kononova and Tynismaa, 2013](#); [Rotig, 2011](#)). When a mutation arises in mitochondrial DNA, it coexists with wild-type DNA since there are many copies of the mitochondrial genome in each organelle and cell. When the mutated DNA exceeds a threshold level, clinical symptoms appear.

While the genes for tRNAs constitute only about 10% of the human mitochondrial genome, mutations in these genes

account for about 50% of the diseases arising from mutations in mitochondrial DNA (Suzuki *et al.*, 2011). One of the most common defects in mitochondrial protein synthesis and respiratory chain activity is caused by a tRNA^{Leu} mutation associated with mitochondrial myopathy, encephalopathy, lactic acid acidosis, and stroke like episodes (MELAS). Approximately 80% of MELAS patients have an A to G mutation at position 3243 (A3243G) in the mitochondrial tRNA^{Leu} gene. This mutation leads in part to defective processing, reduced aminoacylation and an abnormal tertiary structure. Perhaps the most serious defect is that the cell cannot create the modified nucleotide, 5-taurinomethyluridine (tm5U), in the 5'-position of the anticodon of the affected tRNA. As a result, the translation of the Leu codons is deficient leading to a reduction in the synthesis of the mitochondrially encoded subunits of the respiratory chain complexes.

All of the mitochondrial aaRS are encoded in the nuclear genome. Mutations in nine of these have been associated with human genetic diseases. The clinical symptoms differ from one synthetase to another but the most common clinical symptom is encephalopathy (Kononova and Tynismaa, 2013). Only mutations that allow a certain level of residual aminoacylation have been observed since a full loss of function would be lethal.

Mitochondrial Ribosomes

As a group mitochondrial ribosomes present a diverse set of particles that differ from both prokaryotic and eukaryotic cytosolic ribosomes in their specific properties (Table 1). The most detailed structural studies to date have been carried out on bovine mitochondrial ribosomes and on kinetoplast ribosomes.

Mammalian mitochondrial ribosomes have sedimentation coefficients of about 55S and consist of a 28S small subunit (SSU) and a 39S large subunit (LSU) (O'Brien, 1971). Only two rRNA species, 12S in the SSU and 16S in the LSU, are observed. These ribosomes are only 25–30% RNA compared to bacterial ribosomes which are 60–70% RNA and eukaryotic cytoplasmic ribosomes which are 50–60% RNA (van Holde and Hill, 1974). The rRNAs show little primary sequence identity to other rRNAs. They fold into a secondary structure that has overall similarity to that of cytoplasmic rRNAs but show a complete erosion of certain features of the secondary structure.

The bulk of the mass of the mammalian mitochondrial ribosome consists of proteins (Koc *et al.*, 2010, 2013). These proteins fall into two classes – those with homology to bacterial

ribosomal proteins and those that are new and that can be unique to different organisms (Smits *et al.*, 2007; Rackham and Filipovska, 2013). There is a large variation in the proteins present in the mitochondrial ribosomes of different organisms. A number of these mitochondrial-specific ribosomal proteins are thought to play additional roles in the cell or to coordinate mitochondrial translation with other cellular processes. These include two SSU proteins that have been reported to play a role in apoptosis. A number of the ribosomal proteins in mammals are posttranslationally modified and these modifications appear to play an important regulatory role in the synthesis of the respiratory chain complexes (Koc and Koc, 2012).

Cryo-EM studies on bovine and porcine mitochondrial ribosomes (Agrawal and Sharma, 2012; Greber *et al.*, 2014; Kaushal *et al.*, 2014) indicate the presence of several interesting features (Figure 3). These ribosomes have a porous structure with several large masses protruding from a main compact globular core. No E-Site for deacylated tRNA as it leaves the ribosome is observed, a feature that distinguishes these ribosomes from bacterial and eukaryotic cytoplasmic ribosomes. There is a gate-like structure at the entrance to the mRNA-binding channel which may be required for the stable binding of the leaderless mRNAs during initiation. The polypeptide exit tunnel is quite unusual and readily allows premature access to the solvent before the exit site, suggesting an unusual nascent-polypeptide exit mechanism. All of the products of mammalian mitochondrial protein synthesis are inserted into the inner membrane of the organelle. The unusual structure near the exit tunnel may be involved in the interaction of the ribosome with the membrane and with the insertion of the polypeptide products into the hydrophobic core of the membrane (Agrawal and Sharma, 2012).

Perhaps the most unusual ribosomes investigated to date are those present in the kinetoplast of the Trypanosomes (Sharma *et al.*, 2009). These ribosomes have very small rRNAs, less than half the size of the bacterial rRNAs (Table 1) and a large number of proteins, most of which are unique to kinetoplast ribosomes. Their functions remain unknown. The cryo-EM map of these ribosomes (Figure 3) indicates that they have many of the structures expected of ribosomes, but this map has also revealed several unique features. Perhaps the most striking observation is the porosity of the particle. The space between the large and small subunits is quite large and shows many tunnels connecting it to the solvent. The mRNA-binding channel in the SSU is rich with unidentified proteins and the entrance to the channel is quite wide. The E-site appears to be

Table 1 Properties of mitochondrial ribosomes from various sources

Source of ribosomes	Sedimentation coefficient	Subunits	rRNAs	Proteins
Bacterial (Wittmann, 1986)	70S	30S and 50S	16S, 23S, and 5S	~54
Eukaryotic cytosol (Wittmann, 1986)	80S	40S and 60S	18S, 28S, and 5.8S	~82
Mammalian mitochondria (O'Brien and Kalf, 1967; Koc <i>et al.</i> , 2010)	55S	28S and 39S	12S and 16S ^a	~85
<i>S. cerevisiae</i> mitochondria (Grivell, 1995; Smits <i>et al.</i> , 2007)	74S	37S and 54S	15S and 21S	~78
Plant mitochondria (Bonen and Calixte, 2006)	~78S	~33S and 50S	18S, 25S, and 5S	~80
<i>Trypanosoma</i> kinetoplast (Agrawal and Sharma, 2012)	50S	28–30S and 40S	9S and 12S	~133

^aA second small unidentified RNA has been observed in the cryoEM structure of the porcine large subunit (Greber *et al.*, 2014). It is unlikely that this RNA is the 5S species or is derived from it (Lightowlers *et al.*, 2014).

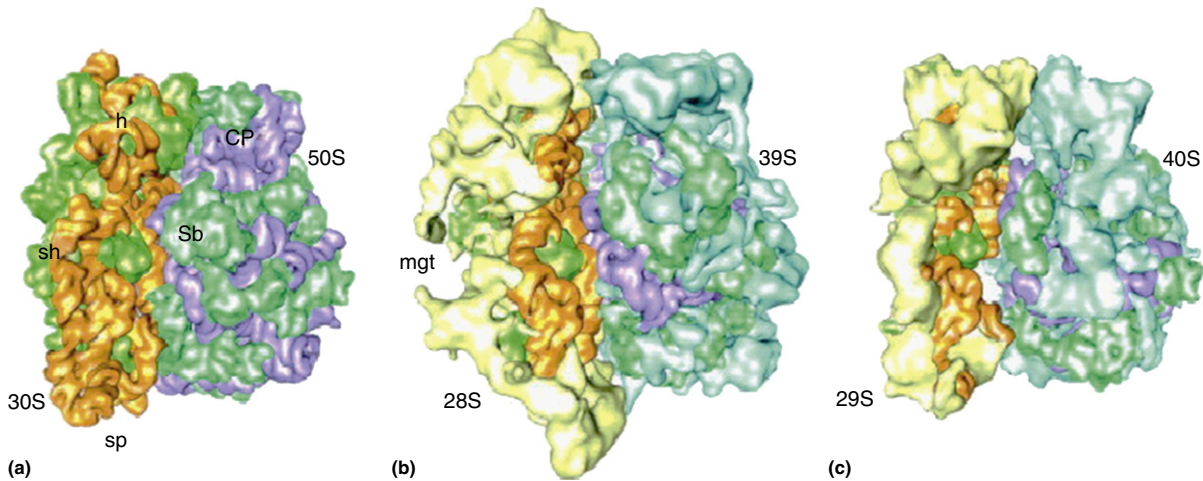


Figure 3 Cryo-EM structures of bacterial and mitochondrial ribosomes. (a) The 70S *E. coli* ribosome, (b) the 55S bovine mitochondrial ribosome, and (c) the 50S ribosome from the protistan, *Leishmania tarentolae*, mitochondria. The rRNAs of the small subunit (SSU) and the large subunit (LSU) are colored orange and purple, respectively; the bacterial ribosomal proteins and proteins homologous to bacterial ribosomal proteins are colored green and aquamarine, respectively; the mitochondrial specific ribosomal proteins are colored yellow and blue. Landmarks of the SSU: h, head; mgt, mRNA gate; sh, shoulder; sp, spur. Landmarks of the LSU: CP, central protuberance; Sb, Stalk base. Reprinted from Agrawal, R.K., Sharma, M.R., 2012. Structural aspects of mitochondrial translational apparatus. Current Opinion Structural Biology 22, 797–803, with permission from Elsevier.

absent as observed in the mammalian mitochondrial ribosome. The polypeptide exit tunnel in the LSU has two openings as observed in mammalian mitochondrial ribosomes. The most highly conserved region of the ribosome is the peptidyl transferase center of the LSU.

Mechanism of Mitochondrial Protein Biosynthesis

Introductory Comments

Protein synthesis in mitochondria follows the same basic steps seen in bacterial and eukaryotic cytoplasmic translational systems. The process is divided into four major stages – initiation, elongation, termination and ribosome recycling. Given the presumed prokaryotic origin of mitochondria, it is expected that the process of protein synthesis in this organelle will be more closely related to that of bacteria than to that of the eukaryotic cell cytoplasm. This idea is borne out by studies of the translational machinery; however, there are a number of interesting and fundamental differences between mitochondrial and bacterial translation. Further, there are clearly distinct differences in how this process takes place in the mitochondria of different organisms. The most detailed studies have been carried out with the mammalian and yeast mitochondrial systems as summarized below.

Initiation

Mammalian mitochondrial system

During initiation, the start site on the mRNA is selected; the initiator tRNA is placed at the P-site of the ribosome and the reading frame on the mRNA is established. In standard translational systems, there are two tRNAs for methionine, one

for initiation and the second for elongation. Mammalian mitochondria are unusual in having a single tRNA^{Met}. A portion of this tRNA becomes formylated and the fMet-tRNA formed is used in chain initiation. The importance of formylation for initiation is illustrated by the serious medical effects arising from mutations in the methionyl-tRNA transformylase (Tucker *et al.*, 2011). The remainder of the tRNA remains as Met-tRNA and participates in elongation. The regulation of this partitioning is not currently understood.

Investigations into the mechanism of initiation indicate that two initiation factors (IFs) are present in mammalian mitochondria, initiation factor 2 (IF2_{mt}) and initiation factor 3 (IF3_{mt}). Based on the translational mechanism of bacteria, one would expect a third factor, initiation factor 1, IF1. However, in mammalian mitochondria, IF2_{mt} serves the function of both IF1 and IF2 (Gaur *et al.*, 2008). In the current model (Figure 4), initiation takes place through a number of discrete steps: (1) IF3_{mt} interacts with 55S ribosomes, forming a transient [IF3_{mt}:55S] complex. (2) The 39S subunit dissociates leading to the formation of an IF3_{mt}:28S complex. (3) IF2_{mt}:GTP binds to the small subunit followed by (4) the binding of the fMet-tRNA and mRNA. When the first 17 nucleotides of the mRNA have slid through the mRNA entrance gate, the movement of the mRNA pauses. During this pause, IF2_{mt}:GTP promotes the binding of fMet-tRNA to the ribosome. If there is an AUG or AUA codon exposed in the P-site, codon/anticodon interactions between the fMet-tRNA and the start codon lead to a stable initiation complex (Christian and Spremulli, 2010) locking the mRNA into place. If no codon/anticodon interactions can form due to a lack of fMet-tRNA or due to the absence of a 5' start codon, the inspection step fails and the mRNA resumes sliding through the small subunit and eventually dissociates. (5) With both mRNA and fMet-tRNA bound to the 28S subunit, the 39S subunit joins the initiation

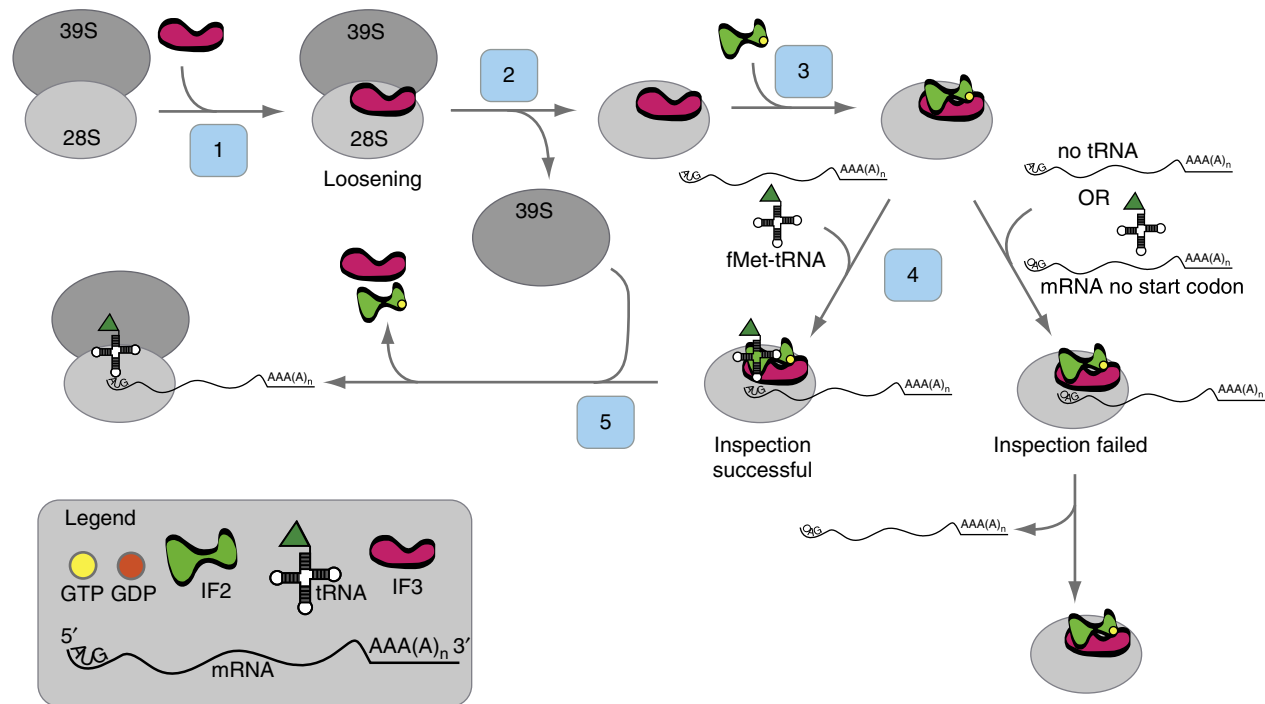


Figure 4 Model for the initiation of protein synthesis in mammalian mitochondria. The steps are described in detail in the text. Reprinted from Christian, B.E., Spemulli, L.L., 2012. Mechanism of protein biosynthesis in mammalian mitochondria. *Biochimica et Biophysica Acta, Gene Regulatory Mechanisms* 1819, 1035–1054, with permission from Elsevier.

complex, GTP is hydrolyzed to GDP by IF2_{mt} and the initiation factors are released. The resulting 55S initiation complex can then enter the elongation process.

Initiation in yeast mitochondria

The mechanism of translational initiation in yeast mitochondria is quite different from that observed in mammals (Herrmann *et al.*, 2013). Initiation factors corresponding to IF2_{mt} and IF3_{mt} have been identified. While the gene for *S. cerevisiae* IF2_{mt} can be readily identified by homology searches, yeast IF3_{mt} (Aim23p) is very different in sequence from mammalian IF3_{mt}. It was not identified until highly sensitive alignment methods were used (Atkinson *et al.*, 2012). No factor equivalent to IF1 has been identified biochemically nor is a gene corresponding to this factor identifiable in the yeast genome. Unlike animal mitochondrial mRNAs, yeast mitochondrial mRNAs have long 5' UTR's (Figure 1). They do not have a Shine/Dalgarno sequence and the mechanism by which the start codon is recognized is unknown. A Met-tRNA transformylase has been identified, but unlike human mitochondria, yeast mitochondria can initiate protein synthesis with an unformylated Met-tRNA under some circumstances.

One of the most interesting features of translational initiation in yeast mitochondria is the requirement for nuclear-encoded mRNA-specific translational activator proteins. Genetic studies in *S. cerevisiae* have demonstrated that there are specific components that interact with individual mitochondrial mRNAs or ribosomes to control the level of synthesis of proteins encoded within the mitochondrial genome (Herrmann *et al.*, 2013). These factors have the ability to

recognize and bind to the 5' UTR of their specified mRNA (Figure 5). They also bind mitochondrial ribosomes and promote the interaction of the translational machinery with the inner membrane facilitating the insertion of the polypeptide products into the assembling respiratory chain complexes. It is also possible that some of these factors interact directly with the nascent chain and promote its incorporation into the appropriate complex. Mitochondrial ribosomes are strongly associated with the inner membrane. This interaction is facilitated in part by proteins such as Oxa1p which interact with the large subunit near the exit tunnel and may also interact with the nascent chain. An understanding of the precise mechanism by which these factors are involved in translational activation and membrane insertion must await the development of an *in vitro* translational system from this organelle.

Polypeptide Chain Elongation

The mechanism of chain elongation in mitochondria has significant similarities to the process in bacteria. In general, this phase of translation has been more highly conserved during evolution than either initiation or termination. The steps involved have been extensively studied only in the mammalian mitochondrial system (Spemulli *et al.*, 2004; Christian and Spemulli, 2012) and are outlined in Figure 6. In the first step, a ternary complex EF-Tu_{mt}-GTP:aa-tRNA enters the decoding site (A-site) of the ribosome and, if cognate codon:anticodon interactions can take place, it is stably bound (Step 1). Following selection, the GTP is hydrolyzed by EF-Tu_{mt} with the

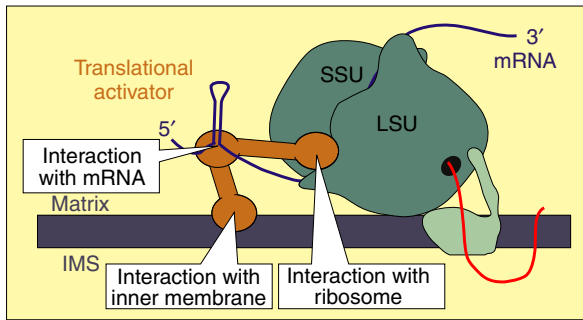


Figure 5 Role of trans-acting factors in yeast mitochondria during translational initiation. Initiation of translation on yeast mitochondrial mRNAs requires the cooperative action of mRNA-specific translational activators. A number of these factors bind to the 5' end of the mRNA. They also appear to bind to the small subunit of the ribosome. They are often found associated with the inner membrane of the mitochondrion and may be directly involved in tethering the ribosome to the membrane. This interaction ensures that mitochondrial translation occurs at or near the membrane to allow the protein products to be inserted into the respiratory chain complexes. IMS, intermembrane space. Reprinted from Herrmann, J.M., Woellhaf, M. W., Bonnefoy, N., 2013. Control of protein synthesis in yeast mitochondria: The concept of translational activators. *Biochimica et Biophysica Acta* 1833, 286–294, with permission from Elsevier.

ribosome acting as the GTPase activating protein. $\text{EF-Tu}_{\text{mt}}\text{:GDP}$ is released (Step 2). In most mitochondrial systems including those from mammals, plants, and many fungi EF-Ts_{mt} acts as a guanine nucleotide exchange factor and displaces the GDP forming an $\text{EF-Tu}_{\text{mt}}\text{:EF-Ts}_{\text{mt}}$ complex (Step 3). Binding of GTP releases EF-Ts_{mt} leading to the formation of $\text{EF-Tu}_{\text{mt}}\text{:GTP}$ (Step 4) allowing another ternary complex to form. One interesting exception to this process occurs in *S. cerevisiae* mitochondria which do not have a factor corresponding to EF-Ts . In the yeast system, EF-Tu_{mt} binds guanine nucleotides very weakly obviating the need for the nucleotide exchange factor. In Step 5, the ribosome itself catalyzes peptide bond formation leaving a deacylated tRNA in the P-site and the peptidyl-tRNA in the A-site. Mitochondrial Elongation Factor G1 (EF-G1_{mt}) catalyzes the translocation step removing the deacylated tRNA from the P-site and moving the peptidyl-tRNA from the A-site to the P-site (Steps 6 and 7). In the mammalian mitochondrial system, the ribosomes do not appear to have an exit site (E-site) and the deacylated tRNA is thought to leave the ribosome during the translocation step.

Termination of Protein Synthesis and Ribosome Recycling in Mitochondria

From a mechanistic perspective, the best studied mitochondrial termination system is that from mammals ([Richter et al.](#),

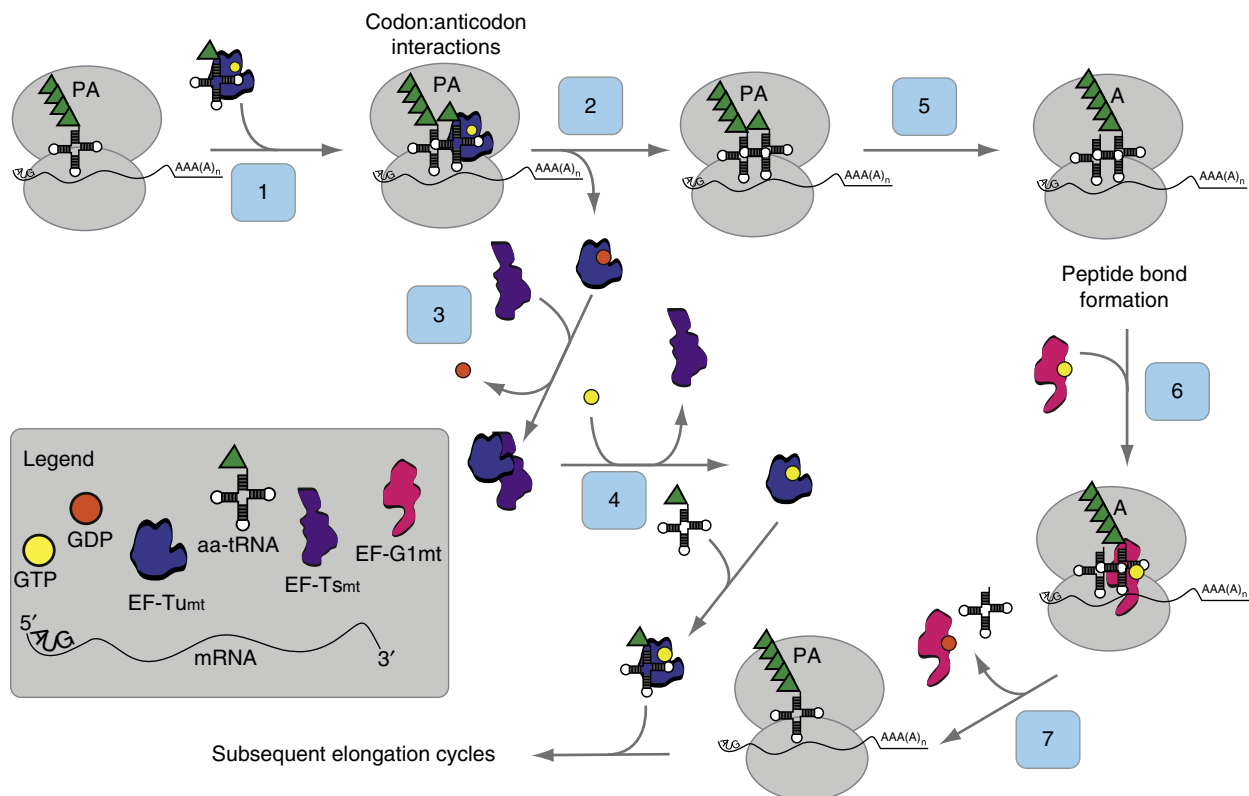


Figure 6 Elongation of protein synthesis in mitochondria. The steps are described in detail in the text. Reprinted from Christian, B.E., Spremulli, L.L., 2012. Mechanism of protein biosynthesis in mammalian mitochondria. *Biochimica et Biophysica Acta, Gene Regulatory Mechanisms* 1819, 1035–1054, with permission from Elsevier.

2010; Lightowlers and Chrzanowska-Lightowlers, 2010). When mitochondrial DNA sequences were analyzed, it became clear that both UAA and UAG serve as stop codons while the standard UGA stop codon codes for tryptophan. Two reading frames end in AGA and AGG and these codons were initially believed to function as additional stop codons rather than as arginine codons. However, it is now thought that a -1 frame-shift during termination leads to the use of a UAG codon for termination in these mRNAs. Thus, only UAA and UAG actually represent stop codons in mammals.

Recent models for chain termination and ribosome recycling (Figure 7) suggest that when a stop codon arrives at the A-site of the translating ribosome, it is recognized by a release factor designated mtRF1a bound to GTP (Step 1). This binding event triggers the peptidyl transferase center on the large subunit to hydrolyze the peptide from the tRNA in the P-site releasing it as a completed product (Step 2). This step leaves a 55S ribosome:mRNA:tRNA complex which must be disassembled allowing the reuse of the components. This recycling step is carried out by the combined action of the mitochondrial ribosome recycling factor (mtRRF also designated RRF1_{mt}) and a translocase (EFG2_{mt} also referred to as RRF2_{mt}) related to EF-G1_{mt} used in chain elongation (Steps 3 and 4).

In yeast as in humans, two codons (UAA and UAG) function as stop signals. Both are recognized by mRF1

corresponding to the human mtRF1a. Like other mitochondrial systems, yeast mitochondria have factors corresponding to RRF1_{mt} and EFG2_{mt}. The overall mechanism of chain termination is expected to be similar to that observed in mammalian mitochondria.

Future Directions

In animals all of the products of mitochondrial protein synthesis are inserted into the respiratory chain complexes in the inner membrane. A similar observation holds true for most of the products in the mitochondrial translational systems of plants and the lower eukaryotes. Other subunits of these complexes are encoded in the nuclear genome, synthesized in the cell cytoplasm and imported into the mitochondrion for assembly into the respiratory chain complexes. Mitochondrial protein synthesis must certainly be regulated and coordinated with the expression of nuclear genes and with the assembly of the multi-subunit complexes. How this is accomplished remains to be elucidated.

Mitochondrial ribosomes are associated with the inner membrane allowing the incorporation of the nascent chain into the membrane. In yeast several proteins, including Oxa1p, have been identified that play a role in this process (Bonnefoy *et al.*, 2009). However, little is known about the details of how

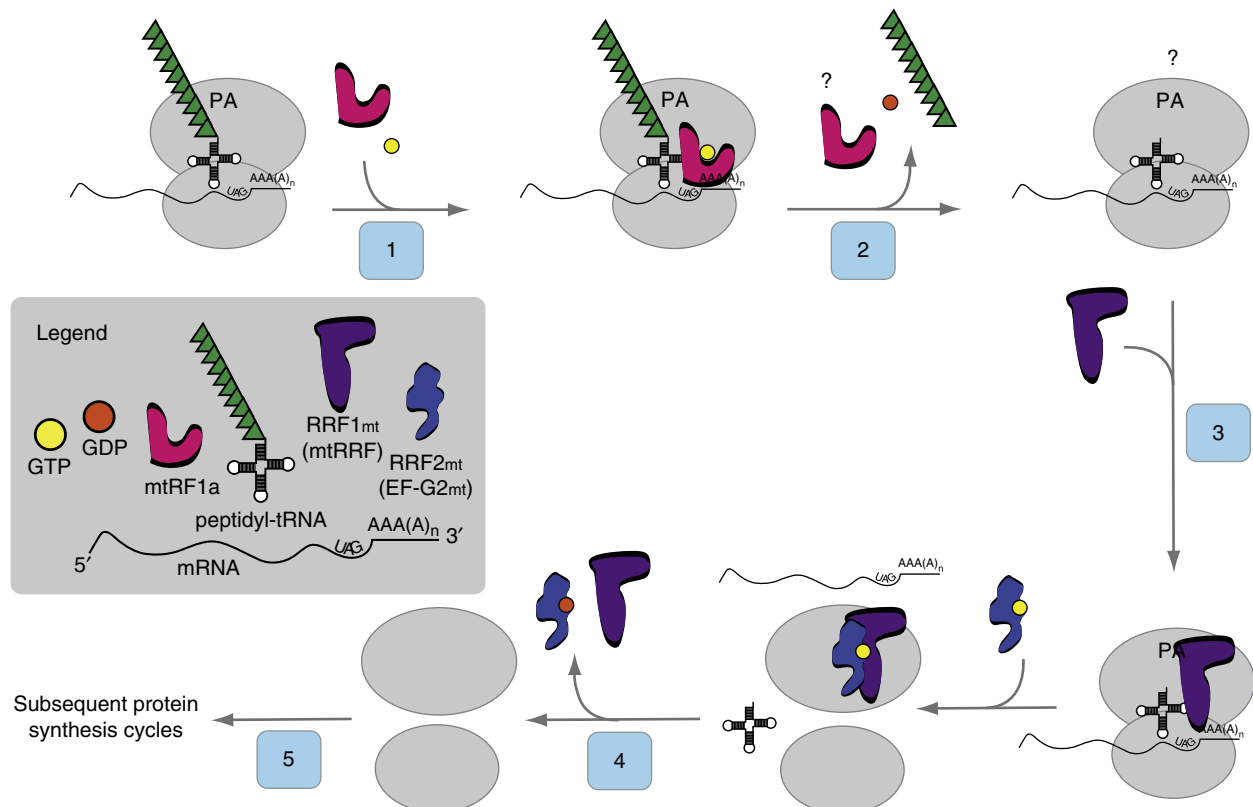


Figure 7 Termination of mitochondrial translation and ribosome recycling. The steps are described in the text although many of the details remain to be elucidated. The ? indicates that it is unclear at the present time which release factor carries out this step. Reprinted from Christian, B. E., Spremulli, L.L., 2012. Mechanism of protein biosynthesis in mammalian mitochondria. *Biochimica et Biophysica Acta, Gene Regulatory Mechanisms* 1819, 1035–1054, with permission from Elsevier.

this interaction occurs and how it participates in the assembly of the respiratory complexes. Even less information is available in mammalian systems, and an understanding of the integration of mitochondrial translation with the physiology of the cell will require the identification and characterization of additional factors.

See also: Nucleic acid Synthesis/Breakdown: RNA Synthesis/Function: Messenger RNA (mRNA): The Link between DNA and Protein. **Organelles: Structure and Function: Mechanisms and Functions of Mitochondrial Dynamics. Protein Synthesis/Degradation: Translation: Components, Initiation, Elongation, Termination, and Regulation**

References

- Agrawal, R.K., Sharma, M.R., 2012. Structural aspects of mitochondrial translational apparatus. *Current Opinion in Structural Biology* 22, 797–803.
- Anderson, S., de Bruijn, M., Coulson, A., *et al.*, 1982. Complete sequence of bovine mitochondrial DNA: Conserved features of the mammalian mitochondrial genome. *Journal of Molecular Biology* 156, 683–717.
- Aphasizhev, R., Aphasizheva, I., 2013. Emerging roles of PPR proteins in trypanosomes: Switches, blocks, and triggers. *RNA Biology* 10, 1495–1500.
- Atkinson, G.C., Kuzmenko, A., Kaminski, P., *et al.*, 2012. Evolutionary and genetic analyses of mitochondrial translation initiation factors identify the missing mitochondrial IF3 in *S. cerevisiae*. *Nucleic Acids Research* 40, 6122–6134.
- Bonen, L., Calixte, S., 2006. Comparative analysis of bacterial-origin genes for plant mitochondrial ribosomal proteins. *Molecular Biology and Evolution* 23, 701–712.
- Bonnefoy, N., Fiumera, H.L., Dujardin, G., Fox, T.D., 2009. Roles of Oxa1-related inner-membrane translocases in assembly of respiratory chain complexes. *Biochimica et Biophysica Acta* 1793, 60–70.
- Chang, J.H., Tong, L., 2012. Mitochondrial poly(A) polymerase and polyadenylation. *Biochimica et Biophysica Acta* 1819, 992–997.
- Christian, B.E., Spremulli, L.L., 2010. Preferential selection of the 5'-terminal start codon on leaderless mRNAs by mammalian mitochondrial ribosomes. *Journal of Biological Chemistry* 285, 28379–28386.
- Christian, B.E., Spremulli, L.L., 2012. Mechanism of protein biosynthesis in mammalian mitochondria. *Biochimica et Biophysica Acta* 1819, 1035–1054.
- Duchene, A.M., Pujol, C., Marechal-Drouard, L., 2009. Import of tRNAs and aminoacyl-tRNA synthetases into mitochondria. *Current Genetics* 55, 1–18.
- Gaur, R., Grasso, D., Datta, P.P., *et al.*, 2008. A single mammalian mitochondrial translation initiation factor functionally replaces two bacterial factors. *Molecular Cell* 29, 180–190.
- Greber, B.J., Boehringer, D., Leitner, A., *et al.*, 2014. Architecture of the large subunit of the mammalian mitochondrial ribosome. *Nature* 505, 515–519.
- Grivell, L., 1995. Nucleo-mitochondrial interactions in mitochondrial gene expression. *Critical Reviews in Biochemistry and Molecular Biology* 30, 121–164.
- Herrmann, J.M., Woellhaf, M.W., Bonnefoy, N., 2013. Control of protein synthesis in yeast mitochondria: The concept of translational activators. *Biochimica et Biophysica Acta* 1833, 286–294.
- Hirsch, M., Penman, S., 1974. Post-transcriptional addition of polyadenylic acid to mitochondrial RNA by a cordycepin-insensitive process. *Journal of Molecular Biology* 83, 131–142.
- van Holde, K., Hill, W., 1974. General physical properties of ribosomes. In: Nomura, M., Tissieres, A., Lengyel, P. (Eds.), *Ribosomes*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, pp. 53–91.
- Kaushal, P.S., Sharma, M.R., Booth, T.M., *et al.*, 2014. Cryo-EM structure of the small subunit of the mammalian mitochondrial ribosome. *Proceedings of the National Academy of Sciences of the USA* 111, 7284–7289.
- Koc, E.C., Cimen, H., Kumcuoglu, B., *et al.*, 2013. Identification and characterization of CHCHD1, AURKAIP1, and CRIF1 as new members of the mammalian mitochondrial ribosome. *Frontiers in Physiology* 4, 183.
- Koc, E.C., Haque, M.d.E., Spremulli, L.L., 2010. Current views of the structure of the mammalian mitochondrial ribosome. *Israel Journal of Chemistry* 50, 45–59.
- Koc, E.C., Koc, H., 2012. Regulation of mammalian mitochondrial translation by post-translational modifications. *Biochimica et Biophysica Acta* 1819, 1055–1066.
- Konovalova, S., Tyynismaa, H., 2013. Mitochondrial aminoacyl-tRNA synthetases in human disease. *Molecular Genetics and Metabolism* 108, 206–211.
- Kozak, M., 1992. A consideration of alternative models for the initiation of translation in eukaryotes. *Critical Reviews in Biochemistry and Molecular Biology* 27, 385–402.
- Lightowlers, R.N., Chrzanowska-Lightowlers, Z.M., 2010. Terminating human mitochondrial protein synthesis: A shift in our thinking. *RNA Biology* 7, 282–286.
- Lightowlers, R.N., Rozanska, A., Chrzanowska-Lightowlers, Z.M., 2014. Mitochondrial protein synthesis: figuring the fundamentals, complexities and complications, of mammalian mitochondrial translation. *FEBS Letters* 588, 2496–2503.
- Merrick, W., 1992. Mechanism and regulation of eukaryotic protein synthesis. *Microbiological Reviews* 56, 291–315.
- Montoya, J., Ojala, D., Attardi, G., 1981. Distinctive features of the 5'-terminal sequences of the human mitochondrial mRNAs. *Nature* 290, 465–470.
- O'Brien, T.W., 1971. The general occurrence of 55S ribosomes in mammalian liver mitochondria. *Journal of Biological Chemistry* 246, 3409–3417.
- O'Brien, T.W., Kalf, G.F., 1967. Ribosomes from rat liver mitochondria. I. Isolation procedure and contamination studies. *Journal of Biological Chemistry* 242, 2172–2179.
- Ojala, D., Montoya, J., Attardi, G., 1981. tRNA punctuation model of RNA processing in human mitochondria. *Nature* 290, 470–474.
- Rackham, O., Filipovska, A., 2013. Supernumerary proteins of mitochondrial ribosomes. *Biochimica et Biophysica Acta* 1840, 1227–1232.
- Richter, R., Pajak, A., Dennerlein, S., *et al.*, 2010. Translation termination in human mitochondrial ribosomes. *Biochemical Society Transactions* 38, 1523–1526.
- Rotig, A., 2011. Human diseases with impaired mitochondrial protein synthesis. *Biochimica et Biophysica Acta* 1807, 1198–1205.
- Scheffler, I., 2008. *Mitochondria*, second ed. Hoboken, NJ: Wiley.
- Schneider, A., 2011. Mitochondrial tRNA import and its consequences for mitochondrial translation. *Annual Review of Biochemistry* 80, 1033–1053.
- Sharma, M.R., Booth, T.M., Simpson, L., Maslov, D.A., Agrawal, R.K., 2009. Structure of a mitochondrial ribosome with minimal RNA. *Proceedings of the National Academy of Sciences of the USA* 106, 9637–9642.
- Smits, P., Smeitink, J.A., van den Heuvel, L.P., Huynen, M.A., Ettema, T.J., 2007. Reconstructing the evolution of the mitochondrial ribosomal proteome. *Nucleic Acids Research* 35, 4686–4703.
- Spremulli, L.L., Coursey, A., Navratil, T., Hunter, S.E., 2004. Initiation and elongation factors in mammalian mitochondrial protein biosynthesis. *Progress in Nucleic Acid Research and Molecular Biology* 77, 211–261.
- Suzuki, T., Nagao, A., Suzuki, T., 2011. Human mitochondrial tRNAs: Biogenesis, function, structural aspects, and diseases. *Annual Review of Genetics* 45, 299–329.
- Tan, T.H.P., Pach, R., Crausaz, A., Ivens, A., Schneider, A., 2002. tRNAs in *Trypanosoma brucei*: Genomic organization, expression, and mitochondrial import. *Molecular and Cellular Biology* 22, 3707–3717.
- Tucker, E.J., Hershman, S.G., Kohrer, C., *et al.*, 2011. Mutations in MTFMT underlie a human disorder of formylation causing impaired mitochondrial translation. *Cell Metabolism* 14, 428–434.
- Watanabe, K., 2010. Unique features of animal mitochondrial translation systems. The non-universal genetic code, unusual features of the translational apparatus and their relevance to human mitochondrial diseases. *Proceedings of the Japan Academy. Series B, Physical and Biological Sciences* 86, 11–39.
- Wittmann, H., 1986. Structure of ribosomes. In: Hardesty, B., Kramer, G. (Eds.), *Structure, Function and Genetics of Ribosomes*. New York: Springer-Verlag, pp. 1–27.
- Wolstenholme, D., 1992. Animal mitochondrial DNA: Structure and evolution. In: Wolstenholme, D., Jeon, K. (Eds.), *Mitochondrial Genomes*, vol. 141. New York, NY: Academic Press, pp. 173–216.

Further Reading

- Akama, K., Christian, B.E., Jones, C.N., *et al.*, 2010. Analysis of the functional consequences of lethal mutations in mitochondrial translational elongation factors. *Biochimica et Biophysica Acta* 1802, 692–698.
- Bonnefoy, N., Fiumera, H.L., Dujardin, G., Fox, T.D., 2009. Roles of Oxa1-related inner-membrane translocases in assembly of respiratory chain complexes. *Biochimica et Biophysica Acta* 1793, 60–70.

- Chrzanowska-Lightowlers, Z.M., Pajak, A., Lightowlers, R.N., 2011. Termination of protein synthesis in mammalian mitochondria. *Journal of Biological Chemistry* 286, 34479–34485.
- Davies, S.M., Rackham, O., Shearwood, A.M., *et al.*, 2009. Pentatricopeptide repeat domain protein 3 associates with the mitochondrial small ribosomal subunit and regulates translation. *FEBS Letters* 583, 1853–1858.
- Gruschke, S., Groene, K., Heublein, M., *et al.*, 2010. Proteins at the polypeptide tunnel exit of the yeast mitochondrial ribosome. *Journal of Biological Chemistry* 285, 19022–19028.
- Hallberg, B.M., Larsson, N.G., 2014. Making proteins in the powerhouse. *Cell Metabolism* 20, 226–240.
- Haque, M.E., Elmore, K.B., Tripathy, A., *et al.*, 2010. Properties of the C-terminal tail of human mitochondrial inner membrane protein Oxa1L and its interactions with mammalian mitochondrial ribosomes. *Journal of Biological Chemistry* 285, 28353–28362.
- Jia, L., Dienhart, M., Schramm, M., *et al.*, 2003. Yeast Oxa1 interacts with mitochondrial ribosomes: the importance of the C-terminal region of Oxa1. *EMBO Journal* 22, 6438–6447.
- Jones, C.N., Jones, C.I., Graham, W.D., Agris, P.F., Spremulli, L.L., 2008. A disease-causing point mutation in human mitochondrial tRNA^{Met} results in tRNA misfolding leading to defects in translational initiation and elongation. *Journal of Biological Chemistry* 283, 34445–34456.
- Jones, C.N., Miller, C., Tenenbaum, A., Spremulli, L.L., Saada, A., 2009. Antibiotic effects on mitochondrial translation and in patients with mitochondrial translational defects. *Mitochondrion* 9, 429–437.
- Jones, C.N., Wilkinson, K.A., Hung, K.T., Weeks, K.M., Spremulli, L.L., 2008. Lack of secondary structure characterizes the 5' ends of mammalian mitochondrial mRNAs. *RNA* 14, 862–871.
- Liu, M., Spremulli, L.L., 2000. Interaction of mammalian mitochondrial ribosomes with the inner membrane. *Journal of Biological Chemistry* 275, 29400–29406.
- Ott, M., Herrmann, J.M., 2010. Co-translational membrane insertion of mitochondrially encoded proteins. *Biochimica et Biophysica Acta* 1803, 767–775.
- Richter, R., Rorbach, J., Pajak, A., *et al.*, 2010. A functional peptidyl tRNA hydrolase, ICT1, has been recruited into the human mitochondrial ribosome. *EMBO Journal* 29, 1116–1125.
- Sissler, M., Lorber, B., Messmer, M., *et al.*, 2008. Handling mammalian mitochondrial tRNAs and aminoacyl-tRNA synthetases for functional and structural characterization. *Methods* 44, 176–189.
- Tsuboi, M., Morita, H., Nozaki, Y., *et al.*, 2009. EF-G2mt is an exclusive recycling factor in mammalian mitochondrial protein synthesis. *Molecular Cell* 35, 502–510.
- Weraarpachai, W., Antonicka, H., Sasarman, F., *et al.*, 2009. Mutation in TACO1, encoding a translational activator of COX I, results in cytochrome c oxidase deficiency and late-onset Leigh syndrome. *Nature Genetics* 41, 833–837.

Relevant Websites

- <http://www.rna.ccbb.utexas.edu/>
Comparative RNA WebSite and Project.
- <http://mamit-trna.u-strasbg.fr/>
Compilation of Mammalian Mitochondrial tRNA Genes.
- <http://www.bch.umontreal.ca/People/lang/FMGP/seqprojects.html>
Fungal Mitochondrial Genome Project: Mitochondrial DNAs of Fungi.
- <http://www.genenames.org/genefamilies/AARS>
HUGO Gene Nomenclature Committee.
- <http://www.genenames.org/genefamilies/MRP>
HUGO: Gene Nomenclature Committee.
- <http://mips.gsf.de/genre/proj/yeast>
MIPS *Saccharomyces cerevisiae* Genome Database.
- <http://mitomap.org/MITOMAP>
MitoMap: A Human Mitochondrial Genome Database.
- <http://mi.caspur.it/mitozoa/>
MitoZoa DB: Mitochondrial DNA of Metazoans.
- <http://www.ndsu.edu/pubweb/~mcclean/plsc731/genome/genome8.htm>
Plant Genome Organization and Structure: Mitochondrial Genome Organization.
- <http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi>
The Genetic Codes: National Center for Biotechnology Information.

Regulated Proteolysis of Signaling Molecules: The Proprotein Convertases

B Ramos-Molina and I Lindberg, University of Maryland-Baltimore, MD, USA
JR Peinado, University of Castilla-La Mancha, Ciudad Real, Spain

© 2016 Elsevier Inc. All rights reserved.

Glossary

Chaperone A protein which assists another protein in some way, usually to block inappropriate protein–protein interactions.

Consensus sequence An amino acid sequence required for a given posttranslational modification.

Cys-palmitoylation A reversible lipid modification of proteins that occurs in the secretory pathway (endoplasmic reticulum and Golgi apparatus), used for membrane targeting.

Genome-wide association studies Approach that involves genome screening to find genetic variations associated with a particular disease.

Heparan sulfate proteoglycan A core protein containing heparan sulfate chains (a linear sugar) which is present on the cell surface or in the extracellular matrix

N-glycosylation Attachment of a sugar chain to asparagines present in the glycosylation consensus sequence Asn-X-Ser.

Sulfation A posttranslational modification of secretory proteins consisting of the addition of a sulfate group to a tyrosine residue.

Transfection To introduce DNA into a cell to be expressed, usually performed together with a specialized lipid chemical.

Transmembrane domain A region of a protein capable of traversing the plasma membrane; normally an α -helix.

Zymogen An inactive enzyme precursor; also termed 'proenzyme.'

Introduction to Regulated Proteolysis

Activation of proteins via proteolytic maturation is a common mechanism used by cells to deploy specific biological activities in a defined temporal and spatial manner. Examples of proteolytic activation occur in all cellular compartments and in all tissues. Probably the most well-recognized example is the activation of pancreatic trypsinogen by the intestinal enzyme enterokinase, which prevents premature- and possibly destructive-production of active trypsin until it reaches the portion of the digestive tract in which it carries out degradative cleavage of nutritive proteins: the upper intestine. Another example of controlled protease action is the activation of cytoplasmic enzymes known as caspases, which exist normally as precursors activated by various messenger systems; these enzymes carry out proteolytic actions on mitochondrial and other proteins that ultimately result in cell death ('apoptosis').

However, proteolytic activation is not restricted to proteases themselves; many structural proteins are activated by proteolytic maturation. One of the most well-known examples is the cleavage of fibrinogen into fibrin, the protein which forms blood clots. While blood clot formation and dissolution involve protease cascades in which one protease activates another in exponential manner, the final substrate is not a protease but a structural protein. Another example of structural protein activation by proteolysis is the cleavage of procollagen to collagen, the scaffold of the extracellular matrix; this cleavage takes place only outside the cell, as intracellular formation of collagen fibrils would be damaging to the structural integrity of the cell. On the other hand, an example of a nonstructural proteolytic activation event is the maturation of protease-activated receptors by circulating proteases; this cleavage results in the formation of a new N-terminus which

acts as ligand for the receptor, thus activating it and triggering an intracellular signaling mechanism.

These few examples highlight the many ways in which proteases are used to control biological processes by proteolytic activation of specific substrates. In this article, we will describe in depth the regulatory proteases that are used to generate peptide signaling molecules, with a special emphasis on the prohormone convertases PC1/3 and PC2. We will present information on their biochemical properties, as well as the substrates they cleave, and how they are regulated. Lastly, we will discuss their relationship to disease states, and how human genetic studies have helped us understand their contribution to diseases states.

The Proprotein Convertases

The family of eukaryotic proprotein convertases, serine proteases with a subtilisin-like arrangement of the catalytic triad, originated with the discovery of the yeast proteinase Kex2 in 1988 (Fuller *et al.*, 1989a). A search for vertebrate homologs of Kex2 led the Thorner group to a human homolog (Fuller *et al.*, 1989b) at about the same time as the discovery by the group of van de Ven that the *fes* upstream region ('fur'), interrupted in lymphoma, encodes a sequence with strong homology to the Kex2 protease (van de Ven *et al.*, 1990); the protein that encodes this activity was therefore named furin. During the following decade, six other family members were discovered: PC2, PC1/3 (PC1; PC3; sPC3), PACE4, PC4, PC5/6, and PC7 (PC8; LPC). Alternative names used in early articles are shown in parentheses; the protein names used here were agreed upon at a Gordon Conference in 2004 (Fugere and Day, 2005). All of these enzymes cleave precursor proteins at basic residue

motifs. This article will not concern itself with another eukaryotic subtilase, PC9, which does not exhibit specificity for paired basic residues, and which has been recently amply reviewed (Seidah *et al.*, 2013, 2014; Seidah, 2013).

All convertases contain four domains in addition to the signal peptide (Figure 1). These include the propeptide, involved in intramolecular chaperoning; the catalytic domain, which carries out catalysis; an interesting novel structure called the 'P' or 'homo B' domain; and a C-terminal section, which, for certain family members, can contain a transmembrane domain. The order of the residues in the catalytic triad which accomplishes substrate cleavage, Asp, His, and Ser, is similar to that of the bacterial enzyme subtilisin; for this reason, this family of enzymes is known as the eukaryotic subtilases. The propeptide is thought to be required for efficient folding of the enzyme during synthesis, serving as an intramolecular chaperone. The P domain has only been studied for a few enzymes, but seems to contribute to enzymatic properties and is clearly important to enzyme stability since truncation of this domain by even a few residues results in instability and intracellular degradation. Lastly, the C-terminal domains, including the cytoplasmic segment which is present in enzymes containing a transmembrane domain (furin, PC5/6, and PC7), are thought to be involved in intracellular trafficking of the PCs.

In this article, we will discuss the expression, regulation, and functional characteristics of each of the convertases in turn, focusing in depth on the two convertases involved in the production of peptide hormones from precursors, PC1/3 and PC2.

Furin (*PCSK3*)

General Properties

Furin, the original founder of the mammalian subtilase family (van de Ven *et al.*, 1990), is now known to be involved in the cleavage of a wide variety of proteins within the constitutive secretory pathway and is ubiquitously expressed (Molloy *et al.*, 1999). At this time, over 2000 articles have been written on furin, making it the best-studied member of the family of proprotein convertases. The gene encoding furin is known as *PCSK3* and it is located on chromosome 15 in humans. Five different transcripts (spliced forms) of furin have been reported, with transcript A (resulting in a protein of 794 amino acids) most commonly expressed (Hatsuzawa *et al.*, 1990; Seidah *et al.*, 1991). The role of so many different transcripts is as yet unclear.

Furin is one of only three proprotein convertases possessing a transmembrane domain/cytoplasmic tail in the C-terminal region (the others are PC5/6B and PC7). This feature enables furin to bind cytoplasmic routing proteins and to cleave its substrates in three distinct subcellular compartments, the *trans*-Golgi network (TGN); the plasma membrane, and the endosomal compartment following its retrieval from the plasma membrane. Molloy *et al.* (1999) have written an excellent article on furin action in various compartments. In addition to these domains, the C-terminal region of furin is particularly rich in cysteine residues (Figure 1); the function of these cysteines is currently unknown.

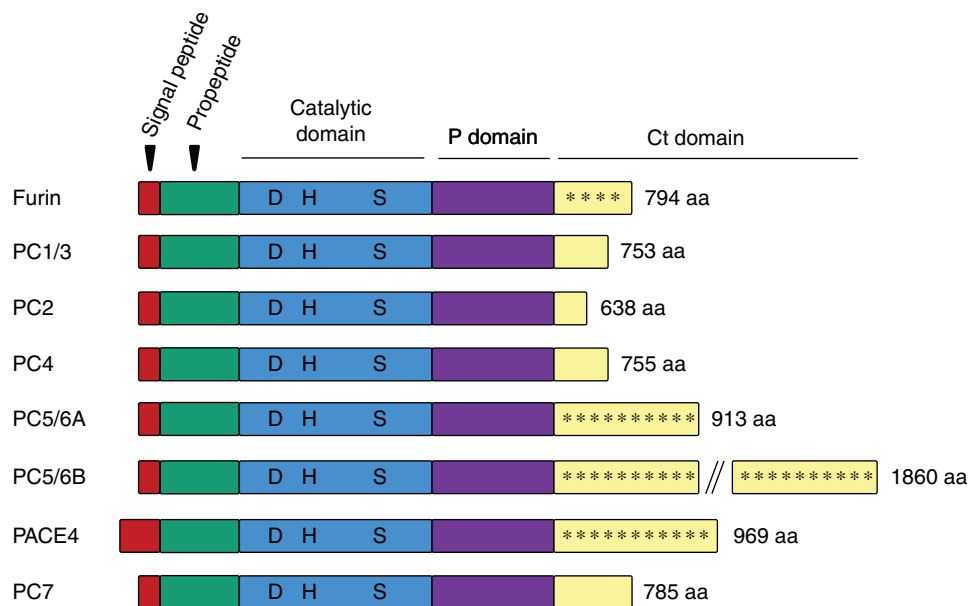


Figure 1 Schematic representation of human proprotein convertases. The subtilisin/kexin-like proprotein convertases family is composed of seven members: PC1/3 (*PCSK1*), PC2 (*PCSK2*), furin (*PCSK3*), PC4 (*PCSK4*), PC5/6 (*PCSK5*), PACE4 (*PCSK6*), and PC7 (*PCSK7*). All convertases contain four main domains (a prodomain, a catalytic domain, a P domain, and a C-terminal (Ct) domain of variable length) and a signal peptide (SP). The catalytic domain is the most conserved region among proprotein convertases and includes the Asp-His-Ser (D-H-S) catalytic triad typical of serine proteases. Note that PACE4 exhibits an unusually long signal peptide in comparison to the other convertases. In the case of furin, PC5/6 and PACE4, the C-terminal domain is rich in Cys residues (highlighted as asterisks). This Cys-rich domain has been shown to be involved in the binding of PC5/6 and PACE4 to proteoglycans at the cell surface. PC5/6 is the only family member that undergoes major alternative splicing, resulting in the two isoforms PC5/6A and PC5/6B that differ considerably in the length of their C-terminal tails.

Enzymatic Properties and Substrate Cleavage

Furin exhibits specificity for a basic residue motif consisting of Arg-X-Arg, where the middle residues are also often basic residues. The enzyme has a neutral pH optimum, requires calcium, and its propeptide is rapidly cleaved from its precursor in the endoplasmic reticulum, though the propeptide remains attached until late in the secretory pathway (Anderson *et al.*, 1997). Some internal cleavage ('shedding') of the membrane anchor occurs when the enzyme reaches the cell surface; the enzyme that performs this reaction has not yet been identified. Like other convertases, furin requires autocatalytic cleavage of its prodomain for activation (Anderson *et al.*, 1997). This process occurs in two different steps and is compartment-specific. First, the prodomain is rapidly cleaved in the endoplasmic reticulum at neutral pH; the propeptide remains attached. During trafficking within the secretory pathway, propeptide removal occurs under mildly acidic pH conditions, thus ensuring that furin acts in a compartment-specific manner.

Because of its complicated cellular itinerary, furin is capable of cleaving a wide range of precursor proteins. Some of the more notable substrates of furin include blood proteins, such as albumin, secreted from the liver; matrix metalloproteinases (MMPs) such as MMP2 and 9; and bacterial toxins such as those produced by the anthrax and *Pseudomonas* bacteria (reviewed in Molloy *et al.* (1999)). Thus far the crystal structure of furin is the only known structure of a member of the proprotein convertase family in higher animals (Henrich *et al.*, 2003).

Physiological Significance of Furin

The furin knockout mouse dies early during embryonic development; death is thought to derive from an inability to make growth factors (Roebroek *et al.*, 1998). Interesting tissue-specific knockouts have been made (e.g., Louagie *et al.*, 2008) which show unexpected cellular roles for this enzyme in identifying new substrates. Furin is not, however, required for tissue viability after development, as tissue-specific knockouts exist (Roebroek *et al.*, 2004; Creemers and Khatib, 2008), as do cell lines which do not express furin (e.g., LoVo; Takahashi *et al.*, 1993).

Overexpression of human furin is correlated with increased carcinogenic potential (Mbikay *et al.*, 1997; Bassi *et al.*, 2000; Thomas, 2002) and furin upregulation in tumors has been associated with increased tumor aggressiveness (Bassi *et al.*, 2001). High furin activity has been related to increased processing of the substrates membrane type 1-MMP (MT1-MMP) (Bassi *et al.*, 2001) and insulin-like growth factor-1 (IGF1) as well as IGF1 receptor present on the surface of tumor cells (Khatib *et al.*, 2001). Administration of furin inhibitors blocks tumor metastasis in mice (Khatib *et al.*, 2001; Bassi *et al.*, 2001). A number of synthetic inhibitors against furin have been recently generated using the crystal structure of furin (Jiao *et al.*, 2006; Coppola *et al.*, 2008; Komiyama *et al.*, 2009; Becker *et al.*, 2010). Interestingly, these furin inhibitors have been also proposed for the treatment of viral and pathogenic infections (Komiyama *et al.*, 2005; Ozden *et al.*, 2008).

Prohormone Convertase 1/3 (PCSK1)

General Properties

Prohormone convertase 1/3 (PC1/3), formerly known as PC1, was one of the earliest PCs discovered (Smeekens *et al.*, 1991; Seidah *et al.*, 1991). This enzyme is predominantly located in neuroendocrine tissues (such as brain, pituitary, adrenal, and pancreas) (Hoshino and Lindberg, 2012). PC1/3 specifically cleaves peptide hormone precursors at the carboxyl side of the amino acids lysine or arginine. This enzyme recognizes a dibasic pair of amino acids, generally lysine-arginine, although rarely a single arginine residue can be used. PC1/3-mediated cleavage of precursor proteins occurs almost exclusively in the TGN and the secretory granules of neuronal and endocrine cells, where it cleaves a wide variety of neuropeptides and peptide hormones, including pro-opiomelanocortin (POMC), proinsulin, proglucagon, parathyroid hormone, proenkephalin, and prodynorphin (reviewed in Hoshino and Lindberg (2012)).

In the brain, PC1/3 is highly expressed in the hypothalamus, for example, in POMC-expressing neurons. In peripheral tissues, its expression is especially strong in pancreas (mainly in the β cells of the islets); adrenal medulla; adenohypophysis; thyroid; and small intestine (where it resides in the enteroendocrine cells). Figure 2 shows the PC1/3-mediated processing of POMC, proinsulin, and proglucagon. In the case of POMC, PC1/3 cleaves the protein at two different dibasic residue sites, producing three different peptides: the N-terminal region of POMC, adrenocorticotrophic hormone (ACTH), and β -lipotropin (β -LPH) (Zhou and Mains, 1994). In pancreatic β cells, PC1/3 initiates the processing of proinsulin by cleaving the bond between the β -chain and the connecting peptide ('C peptide') (Smeekens *et al.*, 1992). In the enteroendocrine cells of the small intestine, PC1/3 enzymatic activity results in the processing of proglucagon to glicentin, GLP-1, and 2 (GLP-2) (Rouille *et al.*, 1995). Note that in this case PC1/3 is able to cleave the precursor at a single arginine residue rather than the double basic residues which are much more common consensus sequences.

PC1/3 Maturation in the Regulated Secretory Pathway

The *PCSK1* gene consists of 14 exons that encode the 753-amino acid precursor preproPC1/3. The signal peptide is co-translationally removed in the endoplasmic reticulum, resulting in an inactive proPC1/3 zymogen (97 kDa) formed by the four different domains described earlier (Figure 3): (1) a prodomain involved in intramolecular chaperoning and early enzyme inhibition (Baker *et al.*, 1993; Rabah *et al.*, 2006), (2) a catalytic domain, which contains the catalytic triad Asp-His-Ser which accomplishes catalysis, (3) a P domain critical to convertase maturation and which is known to contribute to enzymatic properties (calcium binding, low pH dependence, and substrate specificity) (Zhou *et al.*, 1998), and (4) a C-terminal domain, involved in targeting PC1/3 to the secretory granules and in enzyme stability and activity (Rovere *et al.*, 1999; Bernard *et al.*, 2003).

After translation, the PC1/3 precursor proPC1/3 rapidly cleaves its own prodomain within the endoplasmic reticulum

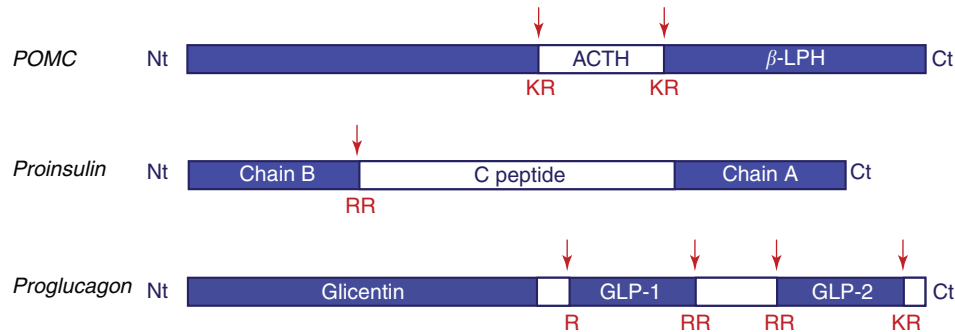


Figure 2 Initial prohormone processing by PC1/3. In POMC-expressing cells the POMC precursor is cleaved by PC1/3 at two KR (lysine-arginine) dibasic sites. ACTH can be secreted (anterior lobe of the pituitary gland) or be further processed by PC2 (intermediate lobe of the pituitary gland) (see Section on PC2). In pancreatic β cells, PC1/3 initiates the cleavage of proinsulin by catalyzing the breakage of the bond between the B chain and the C peptide at an arginine-arginine site. In the endocrine L cells of the small intestine (which lack PC2), PC1/3 catalyzes the cleavage of proglucagon at three dibasic sites (RR, KR) and one arginine (R), producing glicentin and the glucagon-like peptides (GLP) 1 and 2.

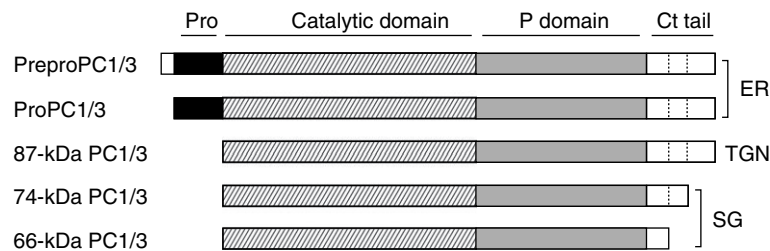


Figure 3 The structure of PC1/3. The preproPC1/3 precursor is co-translationally converted into proPC1/3 during synthesis. In the endoplasmic reticulum (ER), the prodomain (Pro) rapidly undergoes autocatalytic cleavage. In the TGN the prodomain is finally removed, initiating activation of this enzyme. The 87-kDa form is catalytically active, but undergoes additional processing within the C-terminal (Ct) tail in the secretory granules (SG), generating more active truncated forms (74/66 kDa) with somewhat different biochemical properties.

(Lindberg, 1994; Goodman and Gorman, 1994; Zhou and Mains, 1994). The cleaved prodomain remains bound to the rest of the protein, participating in folding and later activation of this protein in the Golgi apparatus and inhibiting the PC1/3 enzymatic activity during the transition between the endoplasmic reticulum and the Golgi and the protein maturation (Rabah *et al.*, 2006). In the TGN, the prodomain finally dissociates, and PC1/3 assumes its active 87-kDa form (Figure 3). In addition to prodomain cleavage and removal, PC1/3 must also undergo certain posttranslational modifications, such as N-glycosylation (Benjannet *et al.*, 1993) for correct folding and maturation within the secretory pathway; the enzyme is also sulfated (Benjannet *et al.*, 1993), though the role of this posttranslational modification is not known.

In comparison to other convertases, such as PC2 and furin, the 87-kDa form of PC1/3 is considerably less active (Zhou and Lindberg, 1993); its autocatalytic processing at the C-terminal domain in the secretory granules results in much more active, but unstable, forms of 74 and 66 kDa (Figure 3) (Zhou and Lindberg, 1994). After cleavage of the C-terminal tail, 74- and 66-kDa forms are able to act at a lower and narrower pH range (5–5.5) (Zhou and Lindberg, 1993), in contrast to the broader pH optimum (5–6.5) of the 87-kDa protein (Zhou and Lindberg, 1994). Like other convertases, PC1/3 requires calcium for enzymatic activity (mM range), and the truncated forms exhibit somewhat greater calcium

dependence (Zhou and Lindberg, 1993, 1994). A recent study has found that in mammalian cells and tissues a high percentage of 87-kDa PC1/3 is present as inactive oligomers and aggregates, whereas 74/66-kDa forms exist primarily in an active monomeric state (Hoshino *et al.*, 2011).

ProSAAS is a neuroendocrine PC1/3 binding protein involved in endogenous inhibition of this enzyme within the regulated secretory pathway (Fricker *et al.*, 2000; Qian *et al.*, 2000; Fortenberry *et al.*, 2002). In endocrine cells, the majority of PC1/3-expressing cells—excepting pancreatic β cells—co-express high amounts of proSAAS protein, which suggests that proSAAS participates in the regulation of PC1/3 activity in neuronal and endocrine tissues (Fricker *et al.*, 2000). However, proSAAS is much more widely expressed in brain than is PC1/3, suggesting other actions (Lanoue and Day, 2001). ProSAAS has been recently shown to serve as a potent antiaggregant chaperone protein in blocking amyloid polypeptide fibrillation (Peinado *et al.*, 2013; Hoshino *et al.*, 2014). These data hint at a possible role for this protein in the development of neurodegenerative diseases involving aggregative processes.

Physiological Significance of PC1/3

In humans as well as in mice, the loss of PC1/3 activity is associated with an increase in circulating prohormone levels.

In humans, a (rare) deficiency in PC1/3 activity caused by inactivating mutations on both alleles has been related to early-onset obesity, diarrhea, diabetes, and hypogonadism (Jackson *et al.*, 1997, 2003; Farooqi and Volders, 2007; Yourshaw *et al.*, 2013; Martin *et al.*, 2013). Thus far, 17 patients exhibiting total or partial loss of PC1/3 activity (via the presence of inactivating mutations) have been reported (Jackson *et al.*, 1997, 2003; Farooqi and Volders, 2007; Yourshaw *et al.*, 2013; Martin *et al.*, 2013). The *PCSK1* mutations found in these individuals are summarized in Table 1. It is important to mention that 14 of the 17 subjects in this table are homozygous for the mutant PC1/3 allele, and in only three cases do individuals exhibit different inactivating mutations in each *PCSK1* allele (see Table 1 legend for further information). Interestingly, the G593R variant, which affects the removal of the prodomain and exit from the endoplasmic reticulum, is present in two of these 17 cases and was present in the original 1997 report of a *PCSK1* mutation; this is thus the most common of these rare mutations (Jackson *et al.*, 1997, 2003; Farooqi and Volders, 2007; Yourshaw *et al.*, 2013; Martin *et al.*, 2013).

Recently, individuals heterozygous for 8 non-synonymous mutations in PC1/3 that partially or totally reduce enzymatic activity have been identified (Creemers *et al.*, 2012). This screening, performed in lean and obese European cohorts (both adults and children), indicates that individuals in possession of just one mutant PC1/3 allele have a much higher risk of developing obesity (by 8.7 times). Recent genome-wide association studies showed that there are three common single nucleotide polymorphisms (SNPs) in *PCSK1* (rs6232, rs6234, and rs6235) strongly associated with obesity in European (Benzinou *et al.*, 2008; Kilpelainen *et al.*, 2009; Heni *et al.*, 2010; Rouskas *et al.*, 2012), Asian (Hsiao *et al.*,

2014), and Mexican populations (Villalobos-Comparan *et al.*, 2012). The SNPs rs6232, rs6234, and rs6235 encode for the N221D, Q665E, and S690T variants, respectively. Interestingly, the non-deleterious Q665E and S690T variants are very common in the population (~25%) and are also associated positively with the incidence of diabetes (Benzinou *et al.*, 2008; Heni *et al.*, 2010; Strawbridge *et al.*, 2011). Unlike the double substitution Q665E/S690T, which does not cause a significant loss of enzyme activity, the N221D mutation reduces PC1/3 activity (Benzinou *et al.*, 2008; Mbikay *et al.*, 2011; Pickett *et al.*, 2013). The recently reported *PCSK1* SNP rs1799904, encoding an R80Q mutation in the propeptide, also has a significant presence in the general population (0.87% minor allele frequency). Similarly to the N221D variant, this mutant exhibits partial loss of enzymatic activity (Pickett *et al.*, 2013).

In mice, total loss of the *PCSK1* gene does not result in increased body weight. In fact, homozygous animals are even smaller than control littermates, mainly because PC1/3 deficiency causes a decrease in growth hormone synthesis due to differences in the sequences of growth hormone between humans and mice (Zhu *et al.*, 2002). However, heterozygous PC1/3 null adult mice do show increased susceptibility to obesity (Zhu *et al.*, 2002; Mbikay *et al.*, 2007). In addition, mice homozygous for the N222D mutation, which results in a partial decrease in PC1/3 activity, are obese and exhibit impaired plasma insulin levels (Lloyd *et al.*, 2006). Interestingly, transgenic mice overexpressing the PC1/3 inhibitor proSAAS develop obesity and hyperglycemia, possibly due to inhibition of endogenous PC1/3 activity (Wei *et al.*, 2004). Conversely, proSAAS knockout mice are lean (Morgan *et al.*, 2010). All of these data point to an important role for PC1/3 and its binding protein, proSAAS, in body weight regulation.

Table 1 Human mutations reported in 17 individuals lacking PC1/3 activity

<i>PCS KI mutation</i>	<i>Mutation Type</i>	<i>PC1/3 domain</i>	<i>PC1/3 properties</i>
G209R	Missense	Catalytic domain	Retained in ER
P258T	Missense	Catalytic domain	Secreted but less active
S307L	Missense	Catalytic domain	Secreted but inactive
N423K	Missense	Pdomain	Secreted but less active
F548S	Missense	Pdomain	Secreted but inactive
G593R	Missense	P domain	Retained in ER
M1X	Nonsense	Signal peptide	Truncated
Y231X	Nonsense	Catalytic domain	Truncated
G250X	Nonsense	Catalytic domain	Truncated
Q337X	Nonsense	Catalytic domain	Truncated
Y343X	Nonsense	Catalytic domain	Truncated
R405X	Nonsense	Catalytic domain	Truncated
A213del	Deletion	Catalytic domain	Partially processed/low secretion
V450fsxl	Deletion	Pdomain	Truncated
A → C + 4(intron 5)	Splice site	Catalytic domain	Truncated
IVS8 + 1G > T	Splice site	Catalytic domain	—
IVS8 + 1G > A	Splice site	Catalytic domain	—

Note: Two of the missense mutants, G209R and G593R, are retained in the endoplasmic reticulum. The remainder of the missense variants, P258T, S307L, N423K, and F548S, are secreted but either display lower enzymatic activity, or are completely inactive. All nonsense mutations result in highly unstable truncated variants (due to a premature stop codon) lacking catalytic activity. The V450fsxl mutation in the frame shift site results in a truncated protein that lacks the P and CT domains. The A → C + 4 transversion in the donor splice site of exon 5 results in a premature stop codon in the catalytic domain. The splice site mutations IVS8 + 1G > T and IVS8 + 1G > A are located at the same acceptor nucleotide in intron 8, and are predicted to critically affect the correct splicing of the *PCSK1* gene. Whereas 14 of the 17 patients are homozygous for their mutations, three individuals have an inactivating mutation in each allele (G593R/A → C + 4; G250X/A213del; and G209R/P258T).

Prohormone Convertase 2 (PCSK2)

General Properties

The prohormone convertase 2 (PC2) is encoded by the gene known as PCSK2, and was the second convertase discovered (Smeekens and Steiner, 1990). Like PC1/3, PC2 is highly expressed in brain, pituitary gland, adrenal medulla, and pancreas and is a key activator of prohormones and neuropeptide precursors within the regulated secretory pathway of neuroendocrine cells. Within the brain, PC2 expression is rich in the hypothalamus, where it acts to cleave a variety of peptide precursors such as proCART (see table in reference Hoshino and Lindberg, 2012). In the pituitary gland, PC2 is mostly located in the intermediate lobe, where it also carries out the cleavage of ACTH to α -melanocyte-stimulating hormone (α -MSH) (Pritchard *et al.*, 2003; Tanaka, 2003). In the pancreas, PC2 is present exclusively in islet tissue, both in α and β cells (but mostly in α cells), where it participates in the synthesis of insulin and glucagon (Furuta *et al.*, 1998; Dey *et al.*, 2004). Note that depending on the tissue where a given precursor is produced, different peptides are generated; this is due to the tissue-specific expression of each convertase. For example, in the anterior lobe of the human pituitary gland, POMC is processed to ACTH but not to α -MSH; whereas, in the intermediate lobe PC2 is able to process ACTH to α -MSH (Figure 4; Pritchard *et al.*, 2003; Tanaka, 2003; Seidah *et al.*, 1999). In the β -cells of pancreatic islets, PC2 and PC1/3 are required for the complete processing of proinsulin. In this case PC2 cleaves the bond between the α -chain and the C peptide, generating insulin (Steiner, 2011). However, PC1/3 null mice can still produce some insulin (Zhu *et al.*, 2002), so it is clear

that functional redundancy must exist between the two prohormone convertases.

Similarly to PC1/3, PC2 is a calcium-dependent endoprotease with a low pH optimum. However, these properties differ considerably from that of PC1/3; the calcium requirement is much lower (micromolar range) and PC2 operates at a lower pH for catalysis, pH 5.0. Indeed, PC2 has the lowest pH optimum of the identified convertases. The substrate specificity of PC2 is also much broader than PC1/3; it is capable of cleaving precursors at a pair of basic residues (usually Lys-Arg or Arg-Arg), and even at a single basic residue if an upstream proline is present (Seidah and Prat, 2012). A number of synthetic non-peptide compounds with potent inhibitory action against PC2 have recently been described (Kowalska *et al.*, 2009; Vivoli *et al.*, 2012; Yongye *et al.*, 2013).

ProPC2 Maturation in the Regulated Secretory Pathway

Like PC1/3, proPC2 contains four different domains: the propeptide, the catalytic domain, the P domain, which is critical for both expression as well as zymogen cleavage (Zhou *et al.*, 1995), and a short C-terminal tail domain that seems to be responsible for granule targeting. PC2 is itself synthesized as a precursor protein, preproPC2, and requires proteolytic maturation for activation, a step where the neuroendocrine chaperone 7B2 plays an important role (Benjannet *et al.*, 1995a,b). In contrast to many of the convertases, PC2 does not require cleavage of its prodomain to exit the endoplasmic reticulum. In fact, the prodomain is involved in the binding of proPC2 to 7B2, a process which occurs after proPC2 folding in the endoplasmic reticulum. Once binding

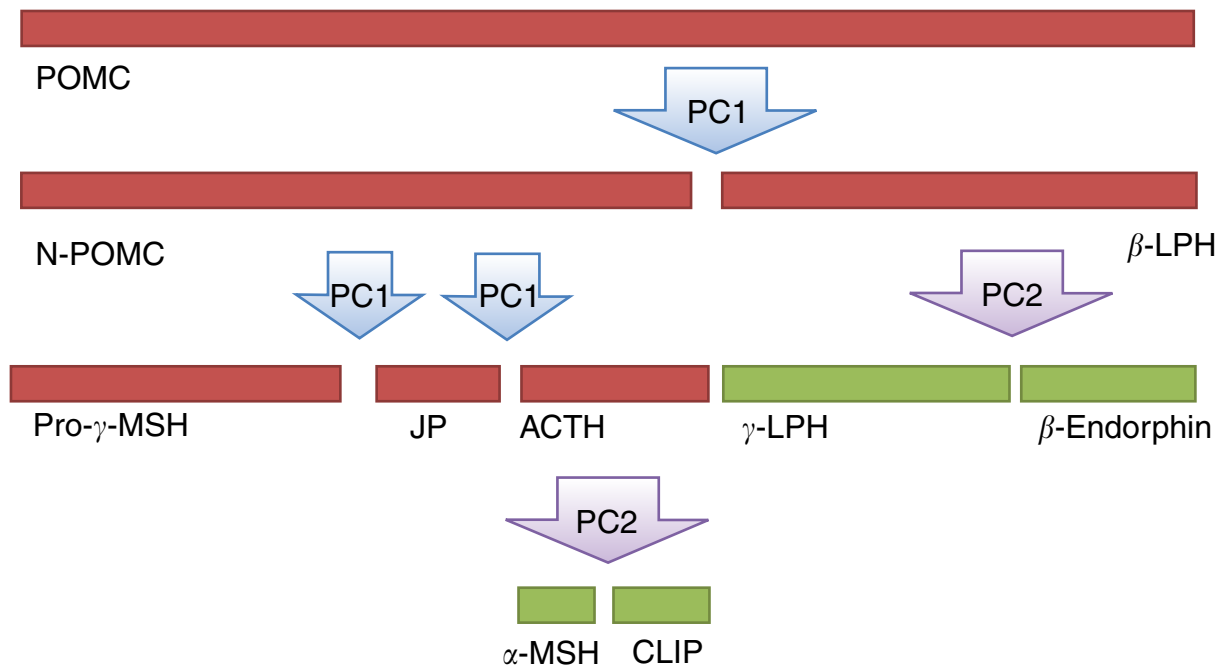


Figure 4 POMC processing by the prohormone convertases 1/3 and 2. Prohormone convertase cleavage of POMC defines the final hormone that will be secreted from neuroendocrine tissues. PC1/3 cleaves POMC to produce β -LPH and ACTH and other peptides, whereas PC2, and only PC2, can further cleave ACTH to produce α -MSH and corticotrophin intermediate-like peptide (CLIP).

occurs, the 7B2/proPC2 complex exits this compartment and travels to the TGN. PC2 activity is thought to be endogenously inhibited by the C-terminal domain of 7B2 during trafficking through the secretory pathway, although overexpression of this domain alone does not decrease peptide processing (Fortenberry *et al.*, 1999). In the acidic environment of the TGN and secretory granules, 7B2 dissociates from proPC2 either before or during its autocatalytic activation via autocatalytic prodomain removal (Tanaka, 2003; Seidah and Chretien, 1999; Figure 5). Unlike PC1/3, PC2 is fully active without further C-terminal cleavage. Recent studies show that 7B2 facilitates proPC2 maturation by preventing its aggregation and inactivating oligomerization (Lee and Lindberg, 2008). Interestingly, 7B2 is much more widely distributed in neuronal and endocrine tissues when compared to the more limited distribution of PC2 (Seidel *et al.*, 1998). These studies raise the idea that 7B2 may possess additional functions other than assisting proPC2 trafficking. In this regard, it has been recently reported that 7B2 is capable of blocking the aggregation of amyloid-containing polypeptides (Peinado *et al.*, 2013; Helwig *et al.*, 2013).

Physiological Significance of PC2

Although PC2 null mice are viable and appear normal at birth, they display retarded growth (Scamuffa *et al.*, 2006). Analysis of these mice reveals chronic fasting hypoglycemia and a deficiency in circulating glucagon. Glucagon is produced in islet α -cells through cleavage of proglucagon by PC2, whereas in intestinal L cells GLPs with different profiles of bioactivity are produced through cleavage by PC1/3 (Hayashi, 2011). In humans, variants in the *PCSK2* gene sequence are associated with reduced insulin secretion in nondiabetic individuals (Jonsson *et al.*, 2012). Variants in the *PCSK2* gene have previously been also linked to type 2 diabetes in African (Leak *et al.*, 2007), Japanese (Yoshida *et al.*, 1995) and European (Jonsson *et al.*, 2012) individuals. However, linkage of *PCSK1* variants to obesity and diabetes is far more common than *PCSK2* variants, suggesting that the PC2 sequence may be more evolutionarily constrained.

Tumors of neuroendocrine origin display increased PC2 expression (Scopsi *et al.*, 1995). PC2 is also clearly expressed in pituitary adenomas, where it is associated with increased

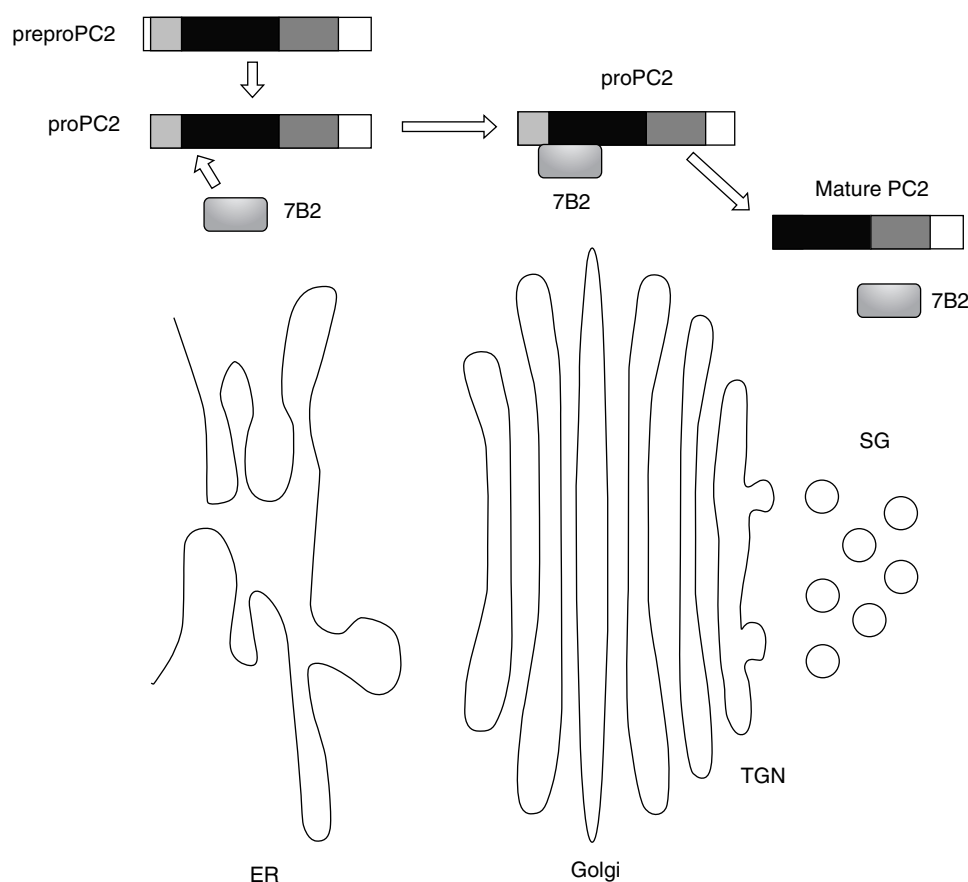


Figure 5 Biosynthesis and maturation of PC2 during the secretory pathway. The signal peptide of preproPC2 is co-translationally removed in the endoplasmic reticulum (ER). The resulting product, proPC2, requires a long time to fold; when folded, it then binds to the secretory chaperone 7B2 (the intact 27-kDa form). 7B2 binding blocks proPC2 aggregation and is essential to maintain the competence of proPC2 for later activation in the secretory granules. Within the acidic environment of secretory granules, proPC2 is activated by autoproteolytic cleavage of the prodomain, and 7B2 (now present as the furin-cleaved 21-kDa form) dissociates from the complex. The presence of 7B2 is not required for proPC2 activation, an intramolecular process initiated by exposure to acid pH.

circulating levels of α -MSH. Also, PC2 expression is present in bronchial carcinoid tumors expressing ACTH, lung carcinomas, adrenomedullary tumors, and pheochromocytomas; PC2 has also been shown to be expressed in two breast cancer lines (Hoshino and Lindberg, 2012). Many neuroendocrine tumors also express 7B2 (reviewed in Mbikay and Seidah, 2001). Thus the expression of PC2 and 7B2 seems to be associated with a differentiated neuroendocrine phenotype.

Proprotein Convertase 4 (PCSK4)

The proprotein convertase 4 (PC4), discovered in 1992 (Nakayama *et al.*, 1992), has the most restricted expression pattern among all members of the family of convertases (Gyamera-Acheampong and Mbikay, 2009). In males, PC4 is expressed exclusively in the testis, where it is mainly located in the haploid spermatids and the sperm cells (Nakayama *et al.*, 1992; Seidah *et al.*, 1992; Torii *et al.*, 1993; Gyamera-Acheampong *et al.*, 2006). PC4 is also detected in spermatocytes and in the residual bodies engulfed by Sertoli cells (Gyamera-Acheampong *et al.*, 2006). PC4 is specifically located in the acrosomal granules and ridges of round and elongated spermatids, respectively. In sperm cells PC4 accumulates in the plasma membrane overlying the acrosome, where it is thought to participate in the fertilization process (Gyamera-Acheampong *et al.*, 2006; Iamsaard *et al.*, 2011). In females, the distribution of PC4 is limited to the ovary and placenta (Tadros *et al.*, 2001; Qiu *et al.*, 2005).

Thus far no data exist on the biosynthesis and trafficking of PC4 within the secretory pathway. Data on the biochemical properties of PC4 have been obtained *in vitro* by using purified recombinant protein after transfection into somatic cell lines (Basak *et al.*, 2008; Remacle *et al.*, 2008). Like all kexin-like convertases, PC4 is capable of cleaving peptide substrates after an arginine residue at a dibasic site (Lys-Arg and Arg-Arg) (Remacle *et al.*, 2008), but preferentially recognizes and cleaves the sequence motif Lys-X-X-Arg following the Arg residue, where X is any amino acid (Basak *et al.*, 2008, 1999, 2004). PC4 enzymatic activity requires neutral pH and mM calcium concentrations (Basak *et al.*, 1999).

Mutant male mice lacking PC4 display severely impaired fertility, and in the *PCSK4* null, spermatozoan capacity to fertilize eggs *in vitro* is significantly reduced (Mbikay *et al.*, 1997; Tardif *et al.*, 2012). The loss of PC4 activity in spermatozoa has been associated with altered processing of acrosin-binding protein (ACRBP), a sperm molecule involved in fertilization (Tardif *et al.*, 2012), which may represent the only known substrate of this enzyme. In addition, several groups have proposed that PC4 might be involved in fertilization via the activation of sperm surface proteins such as the metalloproteases ADAM-1, -2, -3, and -5 (Gyamera-Acheampong *et al.*, 2006; Basak *et al.*, 2004).

Paired Basic Amino Acid Cleaving Enzyme 4 (PACE4; PCSK6)

Like many other convertases, the paired basic amino acid cleaving enzyme 4 (PACE4), encoded by the gene *PCSK6*, has

a widespread tissue distribution, including cerebellum, heart, intestine, kidney, liver, and muscle (Kiefer *et al.*, 1991; Dong *et al.*, 1995). PACE4, discovered in 1991 (Kiefer *et al.*, 1991), is very similar to PC5/6 in many aspects. For example, both convertases can bind tissue inhibitors of metalloproteases and heparan sulfate proteoglycans at the cell surface or in the extracellular matrix via their C-terminal Cys-rich domains (Tsuji *et al.*, 2003; Nour *et al.*, 2005; Mayer *et al.*, 2008; Seidah *et al.*, 2008). Unlike other convertases, PACE4 and PC5/6 can be activated at the cell surface when they are in contact with heparan sulfate proteoglycans (Mayer *et al.*, 2008; Seidah *et al.*, 2008). In this case, the second cleavage of the prodomain, and the subsequent activation, occurs at the cell surface, restricting its activity to the cell surface and extracellular matrix: an example of regulated proteolysis.

In vitro, PACE4 processes the same substrates as other convertases such as furin and PACE5/6, exhibiting strong redundancy with these convertases (Seidah, 2011). *In vivo*, PACE4 (and also furin) is able to cleave Nodal, a secreted protein in the transforming growth factor β family, at the membrane (Mesnard *et al.*, 2011). Other proposed *in vivo* substrates are the metalloprotease ADAM-TS4 (Tortorella *et al.*, 2005), angiopoietin-like protein 3 (Liu *et al.*, 2010), a secreted protein that functions in angiogenesis, and the viral glycoprotein Vpr (Xiao *et al.*, 2008).

In mice, knockout of the PACE4 gene causes about 25% of homozygous mutant embryos to die prenatally with severe cardiac malformations, and the viable knockout mice display bone morphogenetic defects (Constam and Robertson, 2000). Interestingly, PACE4 and furin are overexpressed in the cartilage of patients with arthritis (Seidah and Prat, 2012), and inhibitors of these convertases have been added in cartilage explants for arthritis treatment (Byun *et al.*, 2010). PACE4/furin overexpression has recently been associated with the development of prostate and skin cancers (Bassi *et al.*, 2010; D'Anjou *et al.*, 2011). The inhibition of furin or PACE4 has been also proposed for the treatment of some viral and pathogenic infections (Seidah and Prat, 2012) as well as in cancer (Seidah and Prat, 2012; D'Anjou *et al.*, 2011; Levesque *et al.*, 2012).

Proprotein Convertase 5/6 (PCSK5)

The proprotein convertase 5/6 (PC5 or PC6; now known exclusively as PC5/6), was discovered in 1993 simultaneously by the Seidah and Nakayama (Nakagawa *et al.*, 1993; Lusson *et al.*, 1993) groups. This convertase can be alternatively synthesized as one of two transcripts, one bearing a transmembrane domain (PC5B/6B); and one lacking such a domain (PC5A/6A). Like furin and PACE4, this enzyme also has a widespread tissue distribution. While PC5/6A is a soluble 915-aa protein with a C-terminal domain rich in Cys residues, the PC5/6B form has an extended C-terminal tail of almost 1000 residues that includes a transmembrane domain (Nakagawa *et al.*, 1993). In brain PC5/6 mRNA has been found in regions rich in neuropeptides, including cortex, hippocampus, and hypothalamus. PC5/6 is also expressed in adrenal cortex, ovary, pituitary (anterior lobe), lung, and thyroid (Lusson *et al.*, 1993). PC5/6B transcript expression is especially high in

small intestine and kidney (Seidah, 2011). Like PACE4, PC5/6 is also activated at the cell surface where it contacts heparan sulfate proteoglycans (Mayer *et al.*, 2008). However, because of its transmembrane domain, PC5/6, but not PACE4, is able to cycle from the cell surface back to the TGN through endosomes (Xiang *et al.*, 2000), similarly to furin and PC7 (see corresponding sections).

As discussed in the previous section, certain redundancies between the substrate specificity of the convertases furin, PACE4, PC5/6, and PC7 have been demonstrated in a number of *in vitro* experiments, often leading to difficulty in identification of the protease responsible for a particular event. *In vivo*, one known PC5/6 substrate is growth differentiation factor 11, GDF11 (also known as bone morphogenetic protein 11, BMP-11) (Essalmani *et al.*, 2008). Interestingly, PC5/6 deficiency causes a similar phenotypic pattern as GDF11 knockout, with death shortly after birth, the absence of kidneys and tails, and spinal abnormalities (Essalmani *et al.*, 2008). PC5/6 knockout mice also exhibit GDF11-independent phenotypes, suggesting that the lack of this convertase also affects the processing of other substrates. In addition, PC5/6 mediates the proteolytic activation of α -integrins (Paule *et al.*, 2012) and the phosphoprotein EBP50 (Heng *et al.*, 2011) in endometrial cells, suggesting a key role of this convertase in successful embryo implantation.

In human colon carcinoma tissue, PC5/6 expression is significantly lower than in the adjacent normal tissues (Sun *et al.*, 2009), and PC5/6 expression is not detected in breast tumors (Cheng *et al.*, 1997). These analyses suggest this convertase could have a protective role in tumor growth. In agreement with this idea, crossing mutant mice lacking PC5/6 only in enterocytes together with $Apc^{min/+}$ mice (animals with a point mutation in the *Apc* gene that are used as a model for the human disease familial adenomatous polyposis) has revealed that the loss of PC5/6 both increases tumor number and causes premature mortality (Sun *et al.*, 2009). The interesting paradox of furin and PACE4 acting in a pro-carcinogenic fashion while PC5/6 expression is anticarcinogenic has been previously discussed (Seidah, 2011).

Proprotein Convertase 7 (PCSK7)

Human proprotein convertase 7 (PC7; also known as PC8 and LPC) was first identified as a sequence within a chromosome breakpoint region in lymphoma tissue (Meerabux *et al.*, 1996). While PC7 is thought to be the most ancestral of the basic amino acid-specific convertases, only recently several of its physiological functions have been clarified. Several studies have shown a widespread expression of PC7 mRNA in nearly all tissues and several cell lines (Seidah *et al.*, 1996). Specifically, abundant expression of PC7 has been found in colon, kidney, duodenum, and heart in adult mouse, which indicates that PC7 may exert multiple physiological functions. Similarly to furin, PACE4 and PC5/6, PC7 cleaves its substrates following specific arginine residues both *in vitro* and within cells (Rousselet *et al.*, 2011).

Like other members of the convertase family, proPC7 undergoes propeptide cleavage within the endoplasmic reticulum (Rousselet *et al.*, 2011). The exact mechanism of

activation is still unknown. PC7 has been shown to undergo many posttranslational modifications such as N-glycosylation and sulfation, modifications which likely confer stability (Seidah *et al.*, 1996). PC7 is the only convertase that is palmitoylated at two cysteines present in the cytosolic domain, a modification which normally affects cellular routing (van de Loo *et al.*, 1997, 2000). Unlike other convertases, the propeptide of PC7 is secreted into the medium (Zhong *et al.*, 1999). Recent studies indicate that the active convertase reaches the cell surface by a conventional, but also by an as yet poorly understood unconventional secretory pathway, while the propeptide traffics through the usual Golgi-dependent route (Rousselet *et al.*, 2011). The transmembrane region of PC7 contains critical and unusual elements which control its intracellular trafficking (Rousselet *et al.*, 2011).

PC7 null mice reveal no apparent abnormal phenotype except loss of anxiety (Besnard *et al.*, 2012); this has been explained by functional redundancy with furin (Seidah, 2011). Interestingly, PC7 colocalizes with brain-derived neurotrophic factor in mouse hippocampus and amygdala; in PC7 null mice, the levels of this neuropeptide are reduced in these brain areas, which affects learning and memory (Wetsel *et al.*, 2013). Among specific functions that differ from furin, PC7 action constitutes a quality control checkpoint in rescuing unstable MHC class I proteins, whereas the related convertase furin is completely dispensable for this process (Leonhardt *et al.*, 2010). PC7 has been implicated in the proteolytic activation of bone morphogenetic protein 4 (BMP4) in *Xenopus* oocytes (Nelsen and Christian, 2009). Lastly, a specific function of PC7, not exhibited by other convertases, is its ability to activate the epidermal growth factor receptor (Rousselet *et al.*, 2011). Much still remains to be learned about this interesting convertase.

Summary and Conclusions

During the past 20 years, we have made considerable progress in elucidating the family members responsible for the regulated proteolysis of signaling molecules. With the exception of PC9, members of the proprotein convertase family of enzymes show relatively similar biochemistry and substrate specificity, which likely accounts for their redundant action in many tissues. Substrate redundancy likely underlies the surprisingly mild phenotype of many knockout mice, for example, those of the endocrine convertases PC1/3 and PC2. While the literature to date indicates that furin is likely involved in hundreds of intracellular basic residue cleavage events, there are many cases where a specific convertase plays the major role in cleaving a certain substrate. It is also possible that redundant convertases are brought into play during specific physiological events. Future studies will undoubtedly identify many more specific convertase substrates. The unique features of each convertase are summarized in Table 2.

The biochemistry of the convertases provides examples of many different kinds of regulated proteolysis. Firstly, the convertases themselves are initially synthesized as inactive forms, and must be activated in the correct subcellular compartment (e.g., the late activation of proPC2 in the mature secretory granules). If PC2 cleaved its substrates prematurely,

Table 2 Expression and special features of proprotein convertases

Enzyme	Tissue distribution	Special features
PC1/3	Neural/endocrine	Slow activity; acidic pH optimum (5.5); operates in TGN and secretory granules; aggregation; and binds proSAAS
PC2	Neural/endocrine	Requires 7B2 to form an activatable precursor; acidic pH optimum (5.0); and operates within secretory granules
Furin	Ubiquitous	Compartment-specific substrate cleavage and transmembrane domain in the C-terminal tail
PC4	Reproductive system	Acrosomal granules (sperm cells) and fertilization
PC5/6	Widespread	Alternative splicing; activation at the cell surface (HSPGs) and transmembrane domain in the C-terminal tail of B forms
PACE4	Widespread	Activation at the cell surface (HSPGs)
PC7	Ubiquitous	Not secreted; Cys-palmitoylation and transmembrane domain in the C-terminal tail

Abbreviations: HSPGs, heparan sulfate proteoglycans; PC, proprotein convertase; TM, transmembrane.

the products would likely not become correctly targeted to the compartment from which they are efficiently released, the secretory granules. Another example is that of the convertase PC5/6, which is maintained as a latent enzymatic activity until the later stages of the secretory pathway, for example, for substrates such as proGdf11 (Sun *et al.*, 2011). A last example of regulated proteolysis is the cell membrane-localized activation of the MMP2 precursor by furin; active MMP2 then acts to cleave the extracellular matrix, a process integral to matrix remodeling as well as to the migration of cancer cells to new locations ('metastasis').

As yet, none of the pharmacological potential of the convertases (other than PC9) has been reached. Challenges for the future include establishing the crystal structures of other convertases aside from furin, and identifying potential therapeutic targets in various diseases. PACE4 is an attractive target for arthritis; furin, and PACE4 can be targeted in cancer; and PC2 can be targeted to inhibit the production of glucagon in diabetes. In the coming years, additional substrates for each enzyme will be identified, and these will likely result in additional pharmacologic targets. As with regulated proteolysis, it may be necessary to restrict therapeutic drug delivery to a specific tissue or time frame in order to achieve a beneficial effect.

See also: Interorganellar Communication: Interplay and Processes: Regulation of the Secretory Pathway. Protein Synthesis/Degradation: Protein Degradation – Intracellular: Endoplasmic Reticulum-Associated Degradation and Protein Quality Control

References

- Anderson, E.D., VanSlyke, J.K., Thulin, C.D., Jean, F., Thomas, G., 1997. Activation of the furin endoprotease is a multiple-step process: Requirements for acidification and internal propeptide cleavage. *EMBO Journal* 16, 1508–1518.
- Baker, D., Shiau, A.K., Agard, D.A., 1993. The role of pro regions in protein folding. *Current Opinion in Cell Biology* 5, 966–970.
- Basak, A., Shervani, N.J., Mbikay, M., Kolajova, M., 2008. Recombinant proprotein convertase 4 (PC4) from *Leishmania tarentolae* expression system: Purification, biochemical study and inhibitor design. *Protein Expression and Purification* 60, 117–126.
- Basak, A., Toure, B.B., Lazure, C., *et al.*, 1999. Enzymic characterization *in vitro* of recombinant proprotein convertase PC4. *Biochemical Journal* 343 (Pt 1), 29–37.
- Basak, S., Chrétien, M., Mbikay, M., Basak, A., 2004. *In vitro* elucidation of substrate specificity and bioassay of proprotein convertase 4 using intramolecularly quenched fluorogenic peptides. *Biochemical Journal* 380, 505–514.
- Bassi, D., Lopez de Cicco, R., Mahloogi, H., *et al.*, 2001. Furin inhibition results in absent or decreased invasiveness and tumorigenicity of human cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* 98, 10326–10331.
- Bassi, D.E., Mahloogi, H., Al-Saleem, L., *et al.*, 2001. Elevated furin expression in aggressive human head and neck tumors and tumor cell lines. *Molecular Carcinogenesis* 31, 224–232.
- Bassi, D.E., Mahloogi, H., Klein-Szanto, A.J., 2000. The proprotein convertases furin and PACE4 play a significant role in tumor progression. *Molecular Carcinogenesis* 28, 63–69.
- Bassi, D.E., Zhang, J., Cenna, J., *et al.*, 2010. Proprotein convertase inhibition results in decreased skin cell proliferation, tumorigenesis, and metastasis. *Neoplasia* 12, 516–526.
- Becker, G.L., Sielaff, F., Than, M.E., *et al.*, 2010. Potent inhibitors of furin and furin-like proprotein convertases containing decarboxylated P1 arginine mimetics. *Journal of Medicinal Chemistry* 53, 1067–1075.
- Benjannet, S., Lussan, J., Savaria, D., Chrétien, M., Seidah, N.G., 1995a. Structure-function studies on the biosynthesis and bioactivity of the precursor convertase PC2 and the formation of the PC2/7B2 complex. *FEBS Letters* 362, 151–155.
- Benjannet, S., Rondeau, N., Paquet, L., *et al.*, 1993. Comparative biosynthesis, covalent post-translational modifications and efficiency of prosegment cleavage of the prohormone convertases PC1 and PC2: Glycosylation, sulphation and identification of the intracellular site of prosegment cleavage of PC1 and PC2. *Biochemical Journal* 294 (Pt 3), 735–743.
- Benjannet, S., Savaria, D., Chrétien, M., Seidah, N.G., 1995b. 7B2 is a specific intracellular binding protein of the prohormone convertase PC2. *Journal of Neurochemistry* 64, 2303–2311.
- Benzinou, M., Creemers, J.W., Choquet, H., *et al.*, 2008. Common nonsynonymous variants in PCSK1 confer risk of obesity. *Nature Genetics* 40, 943–945.
- Bernard, N., Kitabgi, P., Rovere-Jovene, C., 2003. The Arg617-Arg618 cleavage site in the C-terminal domain of PC1 plays a major role in the processing and targeting of the enzyme within the regulated secretory pathway. *Journal of Neurochemistry* 85, 1592–1603.
- Besnard, J., Ruda, G.F., Setola, V., *et al.*, 2012. Automated design of ligands to polypharmacological profiles. *Nature* 492, 215–220.
- Byun, S., Tortorella, M.D., Malfait, A.M., *et al.*, 2010. Transport and equilibrium uptake of a peptide inhibitor of PACE4 into articular cartilage is dominated by electrostatic interactions. *Archives of Biochemistry and Biophysics* 499, 32–39.
- Cheng, M., Watson, P.H., Paterson, J.A., *et al.*, 1997. Pro-protein convertase gene expression in human breast cancer. *International Journal of Cancer* 71, 966–971.
- Constam, D.B., Robertson, E.J., 2000. SPC4/PACE4 regulates a TGFbeta signaling network during axis formation. *Genes & Development* 14, 1146–1155.
- Coppola, J.M., Bhojani, M.S., Ross, B.D., Rehemtulla, A., 2008. A small-molecule furin inhibitor inhibits cancer cell motility and invasiveness. *Neoplasia* 10, 363–370.
- Creemers, J.W., Choquet, H., Stijnen, P., *et al.*, 2012. Heterozygous mutations causing partial prohormone convertase 1 deficiency contribute to human obesity. *Diabetes* 61, 383–390.

- Creemers, J.W., Khatib, A.M., 2008. Knock-out mouse models of proprotein convertases: Unique functions or redundancy? *Frontiers in Bioscience* 13, 4960–4971.
- D'Anjou, F., Routhier, S., Perreault, J.P., *et al.*, 2011. Molecular validation of PACE4 as a target in prostate cancer. *Translational Oncology* 4, 157–172.
- Dey, A., Lipkind, G.M., Rouille, Y., *et al.*, 2004. Significance of prohormone convertase 2, PC2, mediated initial cleavage at the proglucagon interdomain site, Lys70-Arg71, to generate glucagon. *Endocrinology* 146, 713–727.
- Dong, W., Marcinkiewicz, M., Vieau, D., *et al.*, 1995. Distinct mRNA expression of the highly homologous convertases PC5 and PACE4 in the rat brain and pituitary. *Journal of Neuroscience* 15, 1778–1796.
- Essalmani, R., Zaid, A., Marcinkiewicz, J., *et al.*, 2008. *In vivo* functions of the proprotein convertase PC5/6 during mouse development: Gdf11 is a likely substrate. *Proceedings of the National Academy of Sciences of the United States of America* 105, 5750–5755.
- Farooqi, I.S., Volders, K., Stanhope, R., *et al.*, 2007. Hyperphagia and early-onset obesity due to a novel homozygous missense mutation in prohormone convertase 1/3. *Journal of Clinical Endocrinology and Metabolism* 92, 3369–3373.
- Fortenberry, Y., Hwang, J.R., Apletalina, E.V., Lindberg, I., 2002. Functional characterization of ProSAAS: Similarities and differences with 7B2. *Journal of Biological Chemistry* 277, 5175–5186.
- Fortenberry, Y., Liu, J., Lindberg, I., 1999. The role of the 7B2 CT peptide in the inhibition of prohormone convertase 2 in endocrine cell lines. *Journal of Neurochemistry* 73, 994–1003.
- Fricker, L.D., McKinzie, A.A., Sun, J., *et al.*, 2000. Identification and characterization of proSAAS, a granin-like neuroendocrine peptide precursor that inhibits prohormone processing. *Journal of Neuroscience* 20, 639–648.
- Fugere, M., Day, R., 2005. Cutting back on pro-protein convertases: The latest approaches to pharmacological inhibition. *Trends in Pharmacological Sciences* 26, 294–301.
- Fuller, R.S., Brake, A., Thorne, J., 1989a. Yeast prohormone processing enzyme (KEX2 gene product) is a Ca²⁺-dependent serine protease. *Proceedings of the National Academy of Sciences of the United States of America* 86, 1434–1438.
- Fuller, R.S., Brake, A.J., Thorne, J., 1989b. Intracellular targeting and structural conservation of a prohormone-processing endoprotease. *Science* 246, 482–486.
- Furuta, M., Carroll, R., Martin, S., *et al.*, 1998. Incomplete processing of proinsulin to insulin accompanied by elevation of Des-31,32 proinsulin intermediates in islets of mice lacking active PC2. *Journal of Biological Chemistry* 273, 1–7.
- Goodman, L.J., Gorman, C.M., 1994. Autoproteolytic activation of the mouse prohormone convertase mPC1. *Biochemical and Biophysical Research Communications* 201, 795–804.
- Gyamra-Acheampong, C., Mbikay, M., 2009. Proprotein convertase subtilisin/kexin type 4 in mammalian fertility: A review. *Human Reproduction Update* 15, 237–247.
- Gyamra-Acheampong, C., Tantibhedhyangkul, J., Weerachayanukul, W., *et al.*, 2006. Sperm from mice genetically deficient for the PCSK4 proteinase exhibit accelerated capacitation, precocious acrosome reaction, reduced binding to egg zona pellucida, and impaired fertilizing ability. *Biology of Reproduction* 74, 666–673.
- Hatsuzawa, K., Hosaka, M., Nakagawa, T., *et al.*, 1990. Structure and expression of mouse furin, a yeast Kex2-related protease. Lack of processing of coexpressed prorenin in GH4C1 cells. *Journal of Biological Chemistry* 265, 22075–22078.
- Hayashi, Y., 2011. Metabolic impact of glucagon deficiency. *Diabetes, Obesity and Metabolism* 13, 151–157.
- Helwig, M., Hoshino, A., Berridge, C., *et al.*, 2013. The neuroendocrine protein 7B2 suppresses the aggregation of neurodegenerative disease-related proteins. *Journal of Biological Chemistry* 288, 1114–1124.
- Heng, S., Cervero, A., Simon, C., *et al.*, 2011. Proprotein convertase 5/6 is critical for embryo implantation in women: Regulating receptivity by cleaving EBP50, modulating ezrin binding, and membrane-cytoskeletal interactions. *Endocrinology* 152, 5041–5052.
- Heni, M., Haupt, A., Schafer, S.A., *et al.*, 2010. Association of obesity risk SNPs in PCSK1 with insulin sensitivity and proinsulin conversion. *BMC Medical Genetics* 11, 86.
- Henrich, S., Cameron, A., Bourenkov, G.P., *et al.*, 2003. The crystal structure of the proprotein processing proteinase furin explains its stringent specificity. *Nature Structural & Molecular Biology* 10, 520–526.
- Hoshino, A., Helwig, M., Razaee, S., *et al.*, 2014. A novel function for proSAAS as an amyloid anti-aggregant in Alzheimer's disease. *Journal of Neurochemistry* 128, 419–430.
- Hoshino, A., Kowalska, D., Jean, F., Lazure, C., Lindberg, I., 2011. Modulation of PC1/3 activity by self-interaction and substrate binding. *Endocrinology* 152, 1402–1411.
- Hoshino, A., Lindberg, I., 2012. Peptide biosynthesis: Prohormone convertases 1/3 and 2. In: Fricker, L.D., Devi, L. (Eds.), *Colloquium Series on Neuropeptides*, vol. 1, issue 1. Morgan & Claypool Life Sciences Publishers, pp. 1–112.
- Hsiao, T.J., Hwang, Y., Chang, H.M., Lin, E., 2014. Association of the rs6235 variant in the proprotein convertase subtilisin/kexin type 1 (PCSK1) gene with obesity and related traits in a Taiwanese population. *Gene* 533, 32–37.
- Iamsaard, S., Vanichviriyakit, R., Hommalai, G., *et al.*, 2011. Enzymatic activity of sperm proprotein convertase is important for mammalian fertilization. *Journal of Cellular Physiology* 226, 2817–2826.
- Jackson, R.S., Creemers, J.W., Farooqi, I.S., *et al.*, 2003. Small-intestinal dysfunction accompanies the complex endocrinopathy of human proprotein convertase 1 deficiency. *Journal of Clinical Investigation* 112, 1550–1560.
- Jackson, R.S., Creemers, J.W., Ohagi, S., *et al.*, 1997. Obesity and impaired prohormone processing associated with mutations in the human prohormone convertase 1 gene. *Nature Genetics* 16, 303–306.
- Jiao, G.S., Cregar, L., Wang, J., *et al.*, 2006. Synthetic small molecule furin inhibitors derived from 2,5-dideoxystreptamine. *Proceedings of the National Academy of Sciences of the United States of America* 103, 19707–19712.
- Jonsson, A., Isomaa, B., Tuomi, T., *et al.*, 2012. Effect of a common variant of the PCSK2 gene on reduced insulin secretion. *Diabetologia* 55, 3245–3251.
- Khatib, A.M., Siegfried, G., Prat, A., *et al.*, 2001. Inhibition of proprotein convertases is associated with loss of growth and tumorigenicity of HT-29 human colon carcinoma cells: Importance of insulin-like growth factor-1 (IGF-1) receptor processing in IGF-1-mediated functions. *Journal of Biological Chemistry* 27, 30686–30693.
- Kiefer, M.C., Tucker, J.E., Joh, R., *et al.*, 1991. Identification of a second human subtilisin-like protease gene in the fes/lps region of chromosome 15. *DNA and Cell Biology* 10, 757–769.
- Kilpelainen, T.O., Bingham, S.A., Khaw, K.T., Wareham, N.J., Loos, R.J., 2009. Association of variants in the PCSK1 gene with obesity in the EPIC-Norfolk study. *Human Molecular Genetics* 18, 3496–3501.
- Komiyama, T., Coppola, J.M., Larsen, M.J., *et al.*, 2009. Inhibition of furin/proprotein convertase-catalyzed surface and intracellular processing by small molecules. *Journal of Biological Chemistry* 284, 15729–15738.
- Komiyama, T., Swanson, J.A., Fuller, R.S., 2005. Protection from anthrax toxin-mediated killing of macrophages by the combined effects of furin inhibitors and chloroquine. *Antimicrobial Agents and Chemotherapy* 49 (9), 3875–3882.
- Kowalska, D., Liu, J., Appel, J.R., *et al.*, 2009. Synthetic small-molecule prohormone convertase 2 inhibitors. *Molecular Pharmacology* 75, 617–625.
- Lanoue, E., Day, R., 2001. Coexpression of proprotein convertase SPC3 and the neuroendocrine precursor proSAAS. *Endocrinology* 142, 4141–4149.
- Leak, T.S., Keene, K.L., Langefeld, C.D., *et al.*, 2007. Association of the proprotein convertase subtilisin/kexin-type 2 (PCSK2) gene with type 2 diabetes in an African American population. *Molecular Genetics and Metabolism* 92, 145–150.
- Lee, S.N., Lindberg, I., 2008. 7B2 prevents unfolding and aggregation of prohormone convertase 2. *Endocrinology* 149, 4116–4127.
- Leonhardt, R.M., Fiegl, D., Rufer, E., *et al.*, 2010. Post-endoplasmic reticulum rescue of unstable MHC class I requires proprotein convertase PC7. *Journal of Immunology* 184, 2985–2998.
- Levesque, C., Fugere, M., Kwiatkowska, A., *et al.*, 2012. The Multi-Leu peptide inhibitor discriminates between PACE4 and furin and exhibits antiproliferative effects on prostate cancer cells. *Journal of Medicinal Chemistry* 55, 10501–10511.
- Lindberg, I., 1994. Evidence for cleavage of the PC1/PC3 pro-segment in the endoplasmic reticulum. *Molecular and Cellular Neuroscience* 5, 263–268.
- Liu, J., Afroza, H., Rader, D.J., Jin, W., 2010. Angiopietin-like protein 3 inhibits lipoprotein lipase activity through enhancing its cleavage by proprotein convertases. *Journal of Biological Chemistry* 285, 27561–27570.
- Lloyd, D.J., Bohan, S., Gekakis, N., 2006. Obesity, hyperphagia and increased metabolic efficiency in P₁ mutant mice. *Human Molecular Genetics* 15, 1884–1893.
- van de Loo, J.H.P., Creemers, J.W.M., Bright, N.A., *et al.*, 1997. Biosynthesis, distinct posttranslational modifications, and functional characterization of lymphoma proprotein convertase. *Journal of Biological Chemistry* 272, 22713–22716.
- van de Loo, J.W., Teuchert, M., Pauli, I., *et al.*, 2000. Dynamic palmitoylation of lymphoma proprotein convertase prolongs its half-life, but is not essential for trans-Golgi network localization. *Biochemical Journal* 352, 827–833.

- Louagie, E., Taylor, N.A., Flamez, D., *et al.*, 2008. Role of furin in granular acidification in the endocrine pancreas: Identification of the V-ATPase subunit Ac45 as a candidate substrate. *Proceedings of the National Academy of Sciences of the United States of America* 105, 12319–12324.
- Lusson, J., Vieau, D., Hamelin, J., *et al.*, 1993. cDNA structure of the mouse and rat subtilisin/kexin-like PC5: A candidate proprotein convertase expressed in endocrine and nonendocrine cells. *Proceedings of the National Academy of Sciences of the United States of America* 90, 6691–6695.
- Martin, M.G., Lindberg, I., Solorzano-Vargas, R.S., *et al.*, 2013. Congenital proprotein convertase 1/3 deficiency causes malabsorptive diarrhea and other endocrinopathies in a pediatric cohort. *Gastroenterology* 145, 138–148.
- Mayer, G., Hamelin, J., Asselin, M.C., *et al.*, 2008. The regulated cell surface zymogen activation of the proprotein convertase PC5A directs the processing of its secretory substrates. *Journal of Biological Chemistry* 283, 2373–2384.
- Mbikay, M., Croissandeau, G., Sirois, F., *et al.*, 2007. A targeted deletion/insertion in the mouse Pcsk1 locus is associated with homozygous embryo preimplantation lethality, mutant allele preferential transmission and heterozygous female susceptibility to dietary fat. *Developmental Biology* 306, 584–598.
- Mbikay, M., Seidah, N.G.M.C., 2001. Neuroendocrine secretory protein 7B2: Structure, expression and functions. *Biochemical Journal* 357, 329–342.
- Mbikay, M., Sirois, F., Nkongo, K.K., Basak, A., Chretien, M., 2011. Effects of rs6234/rs6235 and rs6232/rs6234/rs6235 PCSK1 single-nucleotide polymorphism clusters on proprotein convertase 1/3 biosynthesis and activity. *Molecular Genetics and Metabolism* 104, 682–687.
- Mbikay, M., Sirois, F., Yao, J., Seidah, N.G., Chretien, M., 1997. Comparative analysis of expression of the proprotein convertases furin, PACE4, PC1 and PC2 in human lung tumours. *British Journal of Cancer* 75, 1509–1514.
- Mbikay, M., Tados, H., Ishida, N., *et al.*, 1997. Impaired fertility in mice deficient for the testicular germ-cell protease PC4. *Proceedings of the National Academy of Sciences of the United States of America* 94, 6842–6846.
- Meerabux, J., Yaspo, M.L., Roebroek, A.J., *et al.*, 1996. A new member of the proprotein convertase gene family (LPC) is located at a chromosome translocation breakpoint in lymphomas. *Cancer Research* 56, 448–451.
- Mesnard, D., Donnison, M., Fuerer, C., Pfeffer, P.L., Constam, D.B., 2011. The microenvironment patterns the pluripotent mouse epiblast through paracrine Furin and Pace4 proteolytic activities. *Genes & Development* 25, 1871–1880.
- Molloy, S.S., Anderson, E.D., Jean, F., Thomas, G., 1999. Bi-cycling the furin pathway: From TGN localization to pathogen activation and embryogenesis. *Trends in Cell Biology* 9, 28–35.
- Morgan, D.J., Wei, S., Gomes, I., *et al.*, 2010. The propeptide precursor proSAAS is involved in fetal neuropeptide processing and body weight regulation. *Journal of Neurochemistry* 113, 1275–1284.
- Nakagawa, T., Murakami, K., Nakayama, K., 1993. Identification of an isoform with an extremely large Cys-rich region of PC6, a Kex2-like processing endoprotease. *FEBS Letters* 327, 165–171.
- Nakayama, K., Kim, W.S., Torii, S., *et al.*, 1992. Identification of the fourth member of the mammalian endoprotease family homologous to the yeast Kex2 protease. *Journal of Biological Chemistry* 267, 5897–5900.
- Nelsen, S.M., Christian, J.L., 2009. Site-specific cleavage of BMP4 by furin, PC6, and PC7. *Journal of Biological Chemistry* 284, 27157–27166.
- Nour, N., Mayer, G., Mort, J.S., *et al.*, 2005. The cysteine-rich domain of the secreted proprotein convertases PC5A and PACE4 functions as a cell surface anchor and interacts with tissue inhibitors of metalloproteinases. *Molecular Biology of the Cell* 16, 5215–5226.
- Ozden, S., Lucas-Hourani, M., Ceccaldi, P.E., *et al.*, 2008. Inhibition of Chikungunya virus infection in cultured human muscle cells by furin inhibitors: Impairment of the maturation of the E2 surface glycoprotein. *Journal of Biological Chemistry* 283, 21899–21908.
- Paule, S., Aljofan, M., Simon, C., Rombauts, L.J., Nie, G., 2012. Cleavage of endometrial alpha-integrins into their functional forms is mediated by proprotein convertase 5/6. *Human Reproduction* 27, 2766–2774.
- Peinado, J.R., Sami, F., Rajpurohit, N., Lindberg, I., 2013. Blockade of islet amyloid polypeptide fibrillation and cytotoxicity by the secretory chaperones 7B2 and proSAAS. *FEBS Letters* 587 (21), 3406–3411.
- Pickett, L.A., Yourshaw, M., Albornoz, V., *et al.*, 2013. Functional consequences of a novel variant of PCSK1. *PLoS One* 8, e55065.
- Pritchard, L.E., Oliver, R.L., McLoughlin, J.D., *et al.*, 2003. Proopiomelanocortin-derived peptides in rat cerebrospinal fluid and hypothalamic extracts: Evidence that secretion is regulated with respect to energy balance. *Endocrinology* 144 (3), 760–766.
- Qian, Y., Devi, L.A., Mzhavia, N., *et al.*, 2000. The C-terminal region of proSAAS is a potent inhibitor of prohormone convertase 1. *Journal of Biological Chemistry* 275, 23596–23601.
- Qiu, Q., Basak, A., Mbikay, M., Tsang, B.K., Gruslin, A., 2005. Role of pro-IGF-II processing by proprotein convertase 4 in human placental development. *Proceedings of the National Academy of Sciences of the United States of America* 102, 11047–11052.
- Rabah, N., Gauthier, D., Wilkes, B.C., Gauthier, D.J., Lazure, C., 2006. Single amino acid substitution in the PC1/3 propeptide can induce significant modifications of its inhibitory profile toward its cognate enzyme. *Journal of Biological Chemistry* 281, 7556–7567.
- Remacle, A.G., Shiryayev, S.A., Oh, E.S., *et al.*, 2008. Substrate cleavage analysis of furin and related proprotein convertases: A comparative study. *Journal of Biological Chemistry* 283, 20897–20906.
- Roebroek, A.J., Taylor, N.A., Louagie, E., *et al.*, 2004. Limited redundancy of the proprotein convertase furin in mouse liver. *Journal of Biological Chemistry* 279, 53442–53450.
- Roebroek, A.J., Umans, L., Pauli, I.G., *et al.*, 1998. Failure of ventral closure and axial rotation in embryos lacking the proprotein convertase furin. *Development* 125, 4863–4876.
- Rouille, Y., Martin, S., Steiner, D.F., 1995. Differential processing of proglucagon by the subtilisin-like prohormone convertases PC2 and PC3 to generate either glucagon or glucagon-like peptide. *Journal of Biological Chemistry* 270, 26488–26496.
- Rouskas, K., Kouvasi, A., Paletas, K., *et al.*, 2012. Common variants in FTO, MC4R, TMEM18, PRL, AIF1, and PCSK1 show evidence of association with adult obesity in the Greek population. *Obesity (Silver Spring)* 20, 389–395.
- Rousselet, E., Benjannet, S., Hamelin, J., Canuel, M., Seidah, N.G., 2011. The proprotein convertase PC7: Unique zymogen activation and trafficking pathways. *Journal of Biological Chemistry* 286, 2728–2738.
- Rousselet, E., Benjannet, S., Marcinkiewicz, E., *et al.*, 2011. Proprotein convertase PC7 enhances the activation of the EGF receptor pathway through processing of the EGF precursor. *Journal of Biological Chemistry* 286, 9185–9195.
- Rovere, C., Luis, J., Lissitzky, J.C., *et al.*, 1999. The RGD motif and the C-terminal segment of proprotein convertase 1 are critical for its cellular trafficking but not for its intracellular binding to integrin alpha5beta1. *Journal of Biological Chemistry* 274, 12461–12467.
- Scamuffa, N., Calvo, F., Chretien, M., Seidah, N.G., Khatib, A.M., 2006. Proprotein convertases: Lessons from knockouts. *FASEB Journal* 20, 1954–1963.
- Scopsi, L., Gullo, M., Rilke, F., Martin, S., Steiner, D.F., 1995. Proprotein convertases (PC1/PC3 and PC2) in normal and neoplastic human tissues: Their use as markers of neuroendocrine differentiation. *Journal of Clinical Endocrinology and Metabolism* 80, 294–301.
- Seidah, N.G., 2011. The proprotein convertases, 20 years later. *Methods in Molecular Biology* 768, 23–57.
- Seidah, N.G., 2011. What lies ahead for the proprotein convertases? *Annals of the New York Academy of Sciences* 1220, 149–161.
- Seidah, N.G., 2013. Proprotein convertase subtilisin kexin 9 (PCSK9) inhibitors in the treatment of hypercholesterolemia and other pathologies. *Current Pharmaceutical Design* 19, 3161–3172.
- Seidah, N.G., Awan, Z., Chretien, M., Mbikay, M., 2014. PCSK9: A Key Modulator of Cardiovascular Health. *Circulation Research* 114, 1022–1036.
- Seidah, N.G., Benjannet, S., Hamelin, J., *et al.*, 1999. The subtilisin/kexin family of precursor convertases. Emphasis on PC1, PC2/7B2, POMC and the novel enzyme SKI-1. *Annals of the New York Academy of Sciences* 885, 57–74.
- Seidah, N.G., Chretien, M., 1999. Proprotein and prohormone convertases: A family of subtilases generating diverse bioactive polypeptides. *Brain Research* 848, 45–62.
- Seidah, N.G., Day, R., Hamelin, J., *et al.*, 1992. Testicular expression of PC4 in the rat: Molecular diversity of a novel germ cell-specific Kex2/subtilisin-like proprotein convertase. *Molecular Endocrinology* 6, 1559–1570.
- Seidah, N.G., Day, R., Marcinkiewicz, M., Benjannet, S., Chretien, M., 1991. Mammalian neural and endocrine pro-protein and pro-hormone convertases belonging to the subtilisin family of serine proteinases. *Enzyme* 45, 271–284.
- Seidah, N.G., Hamelin, J., Mamarbachi, M., *et al.*, 1996. cDNA structure, tissue distribution, and chromosomal localization of rat PC7, a novel mammalian proprotein convertase closest to yeast kexin-like proteinases. *Proceedings of the National Academy of Sciences of the United States of America* 93, 3388–3393.
- Seidah, N.G., Marcinkiewicz, M., Benjannet, S., *et al.*, 1991. Cloning and primary sequence of a mouse candidate prohormone convertase PC1 homologous to PC2, furin, and Kex2: Distinct chromosomal localization and messenger RNA distribution in brain and pituitary compared to PC2. *Molecular Endocrinology* 5, 111–122.
- Seidah, N.G., Mayer, G., Zaid, A., *et al.*, 2008. The activation and physiological functions of the proprotein convertases. *International Journal of Biochemistry & Cell Biology* 40, 1111–1125.

- Seidah, N.G., Prat, A., 2012. The biology and therapeutic targeting of the proprotein convertases. *Nature Review Drug Discovery* 11, 367–383.
- Seidah, N.G., Sadr, M.S., Chrétien, M., Mbikay, M., 2013. The multifaceted proprotein convertases: Their unique, redundant, complementary, and opposite functions. *Journal of Biological Chemistry* 288, 21473–21481.
- Seidel, B., Dong, W., Savaria, D., *et al.*, 1998. Neuroendocrine protein 7B2 is essential for proteolytic conversion and activation of proprotein convertase 2 *in vivo*. *DNA and Cell Biology* 17, 1017–1029.
- Smeekens, S., Montag, A.G., Thomas, G., *et al.*, 1992. Proinsulin processing by the subtilisin-related proprotein convertases furin, PC2, and PC3. *Proceedings of the National Academy of Sciences of the United States of America* 89, 8822–8826.
- Smeekens, S.P., Avruch, A.S., LaMendola, J., Chan, S.J., Steiner, D.F., 1991. Identification of a cDNA encoding a second putative prohormone convertase related to PC2 in AIT20 cells and islets of Langerhans. *Proceedings of the National Academy of Sciences of the United States of America* 88, 340–344.
- Smeekens, S.P., Steiner, D.F., 1990. Identification of a human insulinoma cDNA encoding a novel mammalian protein structurally related to the yeast dibasic processing protease Kex2. *Journal of Biological Chemistry* 265, 2997–3000.
- Steiner, D.F., 2011. On the discovery of precursor processing. *Methods in Molecular Biology* 768, 3–11.
- Strawbridge, R.J., Dupuis, J., Prokopenko, I., *et al.*, 2011. Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes. *Diabetes* 60, 2624–2634.
- Sun, X., Essalmani, R., Seidah, N.G., Prat, A., 2009. The proprotein convertase PC5/6 is protective against intestinal tumorigenesis: *In vivo* mouse model. *Molecular Cancer* 8, 73.
- Sun, X., Essalmani, R., Susan-Resiga, D., Prat, A., Seidah, N.G., 2011. Latent transforming growth factor beta-binding proteins-2 and -3 inhibit the proprotein convertase 5/6A. *Journal of Biological Chemistry* 286, 29063–29073.
- Tadros, H., Chrétien, M., Mbikay, M., 2001. The testicular germ-cell protease PC4 is also expressed in macrophage-like cells of the ovary. *Journal of Reproductive Immunology* 49, 133–152.
- Takahashi, S., Kasai, K., Hatsuzawa, K., *et al.*, 1993. A mutation of furin causes the lack of precursor-processing activity in human colon carcinoma LoVo cells. *Biochemical and Biophysical Research Communications* 195, 1019–1026.
- Tanaka, S., 2003. Comparative aspects of intracellular proteolytic processing of peptide hormone precursors: Studies of proopiomelanocortin processing. *Zoological Science* 20, 1183–1198.
- Tardif, S., Guyonnet, B., Cormier, N., 2012. Cornwall, G. A., Alteration in the processing of the ACRBP/sp32 protein and sperm head/acrosome malformations in proprotein convertase 4 (PCSK4) null mice. *Molecular Human Reproduction* 18, 298–307.
- Thomas, G., 2002. Furin at the cutting edge: From protein traffic to embryogenesis and disease. *Nature Reviews Molecular Cell Biology* 3, 753–766.
- Torii, S., Yamagishi, T., Murakami, K., Nakayama, K., 1993. Localization of Kex2-like processing endoproteases, furin and PC4, within mouse testis by *in situ* hybridization. *FEBS Letters* 316, 12–16.
- Tortorella, M.D., Arner, E.C., Hills, R., *et al.*, 2005. ADAMTS-4 (aggrecanase-1): N-terminal activation mechanisms. *Archives of Biochemistry and Biophysics* 444, 34–44.
- Tsuji, A., Sakurai, K., Kiyokage, E., *et al.*, 2003. Secretory proprotein convertases PACE4 and PC6A are heparin-binding proteins which are localized in the extracellular matrix. Potential role of PACE4 in the activation of proproteins in the extracellular matrix. *Biochimica et Biophysica Acta* 1645, 95–104.
- van de Ven, W.J., Voorberg, J., Fontijn, R., *et al.*, 1990. Furin is a subtilisin-like proprotein processing enzyme in higher eukaryotes. *Molecular Biology Reports* 14, 265–275.
- Villalobos-Compan, M., Villamil-Ramirez, H., Villarreal-Molina, T., *et al.*, 2012. PCSK1 rs6232 is associated with childhood and adult class III obesity in the Mexican population. *PLoS One* 7, e39037.
- Vivoli, M., Caulfield, T.R., Martinez-Mayorga, K., *et al.*, 2012. Inhibition of prohormone convertases PC1/3 and PC2 by 2,5-dideoxystreptamine derivatives. *Molecular Pharmacology* 81, 440–454.
- Wei, S., Feng, Y., Che, F.Y., *et al.*, 2004. Obesity and diabetes in transgenic mice expressing proSAAS. *Journal of Endocrinology* 180, 357–368.
- Wetsel, W.C., Rodriguez, R.M., Guillemot, J., *et al.*, 2013. Disruption of the expression of the proprotein convertase PC7 reduces BDNF production and affects learning and memory in mice. *Proceedings of the National Academy of Sciences of the United States of America* 110, 17362–17367.
- Xiang, Y., Molloy, S.S., Thomas, L., Thomas, G., 2000. The PC6B cytoplasmic domain contains two acidic clusters that direct sorting to distinct trans-Golgi network/endosomal compartments. *Molecular Biology of the Cell* 11, 1257–1273.
- Xiao, Y., Chen, G., Richard, J., *et al.*, 2008. Cell-surface processing of extracellular human immunodeficiency virus type 1 Vpr by proprotein convertases. *Virology* 372, 384–397.
- Yongye, A.B., Vivoli, M., Lindberg, I., *et al.*, 2013. Identification of a small molecule that selectively inhibits mouse PC2 over mouse PC1/3: A computational and experimental study. *PLoS One* 8, e56957.
- Yoshida, H., Ohagi, S., Sanke, T., *et al.*, 1995. Association of the prohormone convertase 2 gene (PCSK2) on chromosome 20 with NIDDM in Japanese subjects. *Diabetes* 44, 389–393.
- Yourshaw, M., Solorzano-Vargas, R.S., Pickett, L.A., *et al.*, 2013. Exome Sequencing Finds a Novel PCSK1 Mutation in a Child With Generalized Malabsorptive Diarrhea and Diabetes Insipidus. *Journal of Pediatric Gastroenterology and Nutrition* 57, 759–767.
- Zhong, M., Munzer, J.S., Basak, A., *et al.*, 1999. The prosegments of furin and PC7 as potent inhibitors of proprotein convertases. *In vitro* and *ex vivo* assessment of their efficacy and selectivity. *Journal of Biological Chemistry* 274, 33913–33920.
- Zhou, A., Mains, R.E., 1994. Endoproteolytic processing of proopiomelanocortin and prohormone convertases 1 and 2 in neuroendocrine cells overexpressing prohormone convertases 1 or 2. *Journal of Biological Chemistry* 269, 17440–17447.
- Zhou, A., Martin, S., Lipkind, G., LaMendola, J., Steiner, D.F., 1998. Regulatory role of the P domain of subtilisin-like prohormone convertases. *Journal of Biological Chemistry* 273, 11107–11114.
- Zhou, A., Paquet, L., Mains, R.E., 1995. Structural elements that direct specific processing of different mammalian subtilisin-like prohormone convertases. *Journal of Biological Chemistry* 270, 21509–21516.
- Zhou, Y., Lindberg, I., 1993. Purification and characterization of the prohormone convertase PC1(PC3). *Journal of Biological Chemistry* 268, 5615–5623.
- Zhou, Y., Lindberg, I., 1994. Enzymatic properties of carboxyl-terminally truncated prohormone convertase 1 (PC1/SPC3) and evidence for autocatalytic conversion. *Journal of Biological Chemistry* 269, 18408–18413.
- Zhu, X., Zhou, A., Dey, A., *et al.*, 2002. Disruption of PC1/3 expression in mice causes dwarfism and multiple neuroendocrine peptide processing defects. *Proceedings of the National Academy of Sciences of the United States of America* 99, 10293–10298.

PROTEIN SYNTHESIS/DEGRADATION: PROTEIN DEGRADATION – GENERAL

Mass Spectrometry-based Methodologies for Studying Proteolytic Networks and the Degradome

N Solis and CM Overall, University of British Columbia, Vancouver, BC, Canada

© 2016 Elsevier Inc. All rights reserved.

Glossary

Degradome The complement of all proteases, their respective substrates and their inhibitors in the proteome.

Degradomics The system-wide study of degradomes employing high-throughput technologies such as chip-arrays, genomics, and proteomics.

Endopeptidase A protease that cleaves at peptide bonds within a protein or peptide rather than at the termini of the protein or peptide.

Exopeptidase A protease that cleaves off the terminal amino acid (or penultimate dipeptide) in a peptide or protein. Aminopeptidases are exopeptidases that act at the N-terminus and carboxypeptidases act at the C-terminus.

Internal N-termini N-termini arising from proteolytic cleavage. Distinct from the original mature N-terminus. Internal N-termini are also known as neo-N termini.

Kinome The complement of all human kinases in the proteome.

LC–MS/MS Liquid chromatography tandem mass spectrometry. The process whereby peptide mixtures are resolved by liquid chromatography (typically reverse phase chromatography) with an ion-pairing agent (e.g., formic acid) that is coupled directly to a mass spectrometer. Eluting peptides are desolvated and ionized into the gas phase directly into a mass spectrometer capable of performing fragmentation. Peptide precursor ions are funneled, selected, and fragmented into fragment ions. Precursor and fragment ions in the mass spectra are used to elucidate the peptide sequence by database searching algorithms.

MS1 quantification Quantification of peptides during LC–MS/MS on the basis of their precursor ion intensities. The intensity of a precursor ion of interest is monitored over its elution profile (retention time) and compared to other peptide ion standards or peptide ion samples. Peptides are labeled with isotopic tags (e.g., dimethylation or SILAC) to enable relative quantification across samples. Fragmentation of the precursor ion yields fragment ions that enable peptide sequence elucidation, thus uncoupling identification and quantification.

MS2 quantification quantification of peptides during LC–MS/MS on the basis of reporter ions following fragmentation of a precursor ion. Peptides in each sample are labeled with different isobaric tags (i.e., the tags have

identical mass). Upon fragmentation (e.g., CID or HCD), these isotopically distinct tags yield a reporter ion unique to that tag. Fragmentation of the precursor ion (MS2) yields fragment ions enabling peptide sequence peptide elucidation and also reporter ions for each chemical tag that allows for relative quantification between samples in one mass spectrum.

MS3 quantification Quantification of peptides during LC–MS/MS on the basis of reporter ions following further fragmentation of a fragment ion. During MS2 quantification, a precursor ion (peptide ion) is selected for fragmentation to yield identification-specific and quantification-specific ions. A second round of fragmentation (MS3) of a fragment ion yields new fragment ions to validate identification of the first fragmentation event as well as reporter ions to provide quantification. MS3 is a feature of ion trap mass spectrometers and provides higher confidence in identification and quantification yet at the cost of sensitivity.

Negative selection of N-termini Enrichment of neo-N-termini by depletion of trypsin/GluC-generated internal peptides. Various strategies are employed to perform negative selection including TAILS and COFRADIC.

Neo N-termini N-termini in proteins or peptides arising from proteolysis; internal N-termini.

Positive selection of N-termini Enrichment of neo-N-termini by affinity purification. The primary amine of each protein N-terminus is tagged with biotin, either enzymatically using an engineered form of subtiligase, or chemically. Digestion with trypsin yields non-modified termini. Biotinylated neo-N-termini are enriched by streptavidin chromatography.

P residues Residues N-terminal to the scissile bond of a protease substrate. If the residue is X residues immediately N-terminal to the scissile bond, it is the PX site (e.g., the P3 site is the third residue N-terminal to the scissile bond of the substrate).

P' residues Residues C-terminal to the scissile bond of a protease substrate. If the residue is X residues immediately C-terminal to the scissile bond, it is the PX' site (e.g., the P3' site is the third residue C-terminal to the scissile bond of the substrate).

Protease web The interdependent network of interactions of all proteases and inhibitors linked by cleavages.

Proteases Enzymes that catalyze the hydrolysis of amide bonds in peptides and proteins. Proteases cleave specific amino acid residues alone or within motifs, the frequency of their occurrence determines how broadly the protease can act on proteins and on proteomes.

Proteome The complement of all proteins and their respective posttranslational modifications in a defined organelle, cell type, tissue, or organism at a particular point in time under defined conditions.

PSP Proteolytic signature peptides. Peptides that can be used to monitor the activation status of a protein. Namely, a fully tryptic unique peptide spanning across a cleavage site for a protease of interest. Detection of the entire tryptic peptide indicates lack of protease activity, whereas detection (and quantification) of the semi-tryptic peptides indicates protease activity for that cleavage site.

SRM Selected reaction monitoring. A form of label-free peptide quantification where the mass spectrometer scans combinations of precursor and then fragment ions, known as transitions, during LC-MS. As well as defined mass/charge (m/z) ratios, transitions designed for known peptides of interest have defined LC retention times and are thus specific to the (unlabeled) peptide of interest. The mass spectrometer scans several transitions during a chromatography experiment and generates an elution profile. Comparison of elution profile areas across samples provides quantification. SRM is non-discovery form of mass spectrometry that is more sensitive than discovery forms.

Subsites The surface sites on a protease into which the residues either side of the scissile bond dock. S subsites receive the P side amino acids lying N-terminal to the cleavage site whereas the S' subsites receive the P' amino acids lying carboxy-terminal to the scissile bond.

TAILS 'Terminal amine isotopic labeling of substrates' is a negative selection technique for the enrichment of protein N-termini. Before trypsin or GluC digestion the primary amine of each protein N-terminus is labeled and blocked with chemical tags that are isotopically coded to enable relative quantification between samples of identical peptides. Proteins are then digested with trypsin (or other proteases, e.g., GluC) and internal tryptic peptides are removed from the sample by binding the reactive primary amines of the newly generated tryptic peptides to a high molecular weight multi-aldehyde-containing polymer (HPG-ALD). Original and neo-N-termini remain in solution and the polymer can be easily removed by ultrafiltration.

Terminome The complement of all protein N-termini (N-terminome) and C-termini (C-terminome) of proteins in a proteome, which includes the mature protein termini and protease generated neo-termini. Terminomes add an additional layer of information of a proteome by providing information on protease dynamics.

Terminomics The study of terminomes employing high-throughput technologies.

Zymogen Inactive protease precursor. Terminal peptide chain must be cleaved or allosterically moved to expose the active site.

Introduction

Proteases were originally considered to be hydrolases with a strict degradative role and not thought to have any particular exciting roles in cellular regulation or development other than the catabolic release of free amino acids required for cellular growth (Freeman *et al.*, 1983). However, the sequencing of multiple eukaryotic genomes other than mice and humans has shown that there is a plethora of proteases throughout the genetic code. Of the 20 126 human genes there are over 560 proteases identified in *Homo sapiens* and 644 in *Mus musculus* (Overall and Blobel, 2007) representing approximately 2.8% of the genes encoded by mammals. In comparison, kinases, a well-described group of regulatory enzymes responsible for attaching phosphate groups to proteins to regulate signal transduction, cellular response to damage, growth, and a myriad of other cellular events, (Hanks *et al.*, 1988) comprise 518 proteins in humans (Manning *et al.*, 2002). The activity of kinases is strictly intracellular or membrane-bound whereas the activity of proteases has been reported intracellularly, extracellularly, on and in cell membranes (Doucet *et al.*, 2008; Lopez-Otin and Overall, 2002; Overall and Blobel, 2007). The protease repertoire in humans is distributed with 273 extracellular proteases, 277 intracellular proteases, and 16 integral membrane proteases (Overall and Blobel, 2007). Due to their

abundance and evolutionary diversity (having different families of proteases on the basis of their sequence and different clans based on structure), proteases possess regulatory roles as varied and critical as those of the human 'kinome.'

Proteases are classified by their molecular mechanism (which is an indirect reflection of their primary structure similarity): in humans, the five classes are aspartate proteases (21 total), cysteine proteases (148 total), metalloproteinases (194 total), serine proteases (175 total), and threonine proteases (28 total) (Overall and Blobel, 2007). Proteases are further distinguished on the basis of their molecular mechanism, namely, endopeptidases and exopeptidases (Barrett and McDonald, 1986). Exopeptidases act by removing the terminal amino acid or the two N-terminal amino acids from an oligopeptide or protein: aminopeptidases cleave an amino acid from the N-terminus (Taylor, 1993) whereas carboxypeptidases remove an amino acid from the C-terminus (Skidgel and Erdos, 2006). In contrast, endopeptidases are able to hydrolyze peptide bonds within a protein or oligopeptide. A protease recognizes a particular stretch of amino acids in its substrate that are accommodated at the binding surface of the protease active site. Upon substrate binding, the protease catalyzes hydrolysis of the peptide bond at the cleavage site. The residues N-terminal to the scissile bond on the substrate are termed the P residues whereas the residues

C-terminal to the scissile bond are termed the P' residues. It is the sequence across P and P' residues that determine protease sequence specificity (Schechter and Berger, 1967). In addition, the protease binding cleft has a similar nomenclature that defines the specificity of the protease: pockets that accommodate P and P' residues of the substrate are termed S subsites and S' subsites respectively (Schechter and Berger, 1967). The interaction between P/P' and S/S' sites is what defines the specificity of the protease: substrate interaction and also influences kinetics of peptide bond hydrolysis.

All proteins are processed by a protease at some point in their life cycle (Doucet et al., 2008). The non-reversible nature of proteolytic cleavage (as opposed to most other post-translational modifications in cells) demands a well-tuned, tightly controlled system for controlling the activation and deactivation of regulatory proteases. Thus proteases are not ubiquitously dispersed in a cell, or even among tissues or time during development and differentiation and are not ubiquitously active: proteases have inhibitors bound to them or exist in inactive forms (zymogens) that are activated upon a defined stimulus.

Controlled proteolytic processing modulates the functions of proteins and thus proteases play multiple crucial roles in cellular regulation. For example, directing proteins to their correct location in the cell by removal of signal sequences (von Heijne, 1990), orchestrating the recruitment of leukocytes during an inflammatory response by processing chemokines (Ajami et al., 2008; Dean et al., 2008; McQuibban et al., 2002), processing cytoplasmic proteins involved in cell cycle regulation (e.g., processing of cyclins; King et al., 1996), zymogen activation (activation of other proteases; Neurath and Walsh, 1976; Wiederanders et al., 2003), quality control (proteasome-mediated degradation; Ciechanover, 2005) as well as extracellular matrix remodeling (by matrix metalloproteinases and other proteases). Hence it is no surprise that proteases represent 5–10% of all drug targets (Overall and Blobel, 2007).

Perhaps one of the most important concepts to arise in recent years is that proteases act in concerted networks, termed the protease web, to transmit and amplify signals (Fortelny et al., 2014) – somewhat analogous to a kinase pathway. Protease cascades *in vivo* are dynamic and well regulated. For example, the proteolytic processing cascades of coagulation (Hoffman and Monroe, 2007) and caspases (Pop and Salvesen, 2009) allow for signal amplification and many points of regulation that culminate in the formation of a fibrin clot or apoptosis, respectively. In both cases, a carefully concerted action of proteases and other proteins is tightly regulated to prevent untimely clot formation or cell death. Moreover, once a signal has initiated and the appropriate response been effected, the cascade must be 'switched off.' Switching off a protease is typically achieved by biological inhibitors – molecules that can bind the active site of a protease to prevent it from binding substrate; or binding an allosteric site (remote region from the binding site) that alters the conformation of the protease, thus preventing it from binding substrate (Turk et al., 2012). For instance, in the coagulation cascade, in order to prevent further thrombin clot production, the serine protease inhibitor (serpin) antithrombin inactivates thrombin (Seegers et al., 1954) and activated protein C cleaves FVa and FVIIIa (Hoffman and Monroe, 2007).

Although in the examples of coagulation and apoptosis there is a wealth of information regarding the key proteases, of the approximately 600 proteases in humans, 50% have no known substrates (Fortelny et al., 2014) and no known regulatory functions or *in vivo* functionalities (Doucet et al., 2008). Moreover, not all substrates of a particular protease that are determined *in vitro* or bioinformatically are relevant *in vivo*. Proteases, their respective substrates and inhibitors do not always coexist in the same cellular compartment, do not exist in the concentrations typically used in *in vitro* assays, are not always in complexes required for function (such as the proteasome), or do not have substrates folded in the correct conformation. As such, despite the power of *in vitro* assays to determine substrates and protease cleavage sites, these lack the depth, throughput and *in vivo* relevance given by animal models or cell lines that display proteases in action. Consequently, in order to study the complete set of proteases, their respective substrates, spatio-temporal expression profiles and phenotypic effects (defined as the 'degradome') the field of 'degradomics' has emerged.

Degradomics

The convergence of several technologies over the past two decades including high-throughput genome sequencing (genomics), ionization methods for biological molecules in mass spectrometry (Fenn et al., 1989; Karas et al., 1985), microarray (chip) technology (Lopez-Otin and Overall, 2002), increased computational power and reliable high pressure liquid chromatography (HPLC) systems has resulted in the birth of proteomics (Washburn et al., 2001). Proteomics is the systematic characterization of all proteins and peptides and their posttranslational modifications at a particular time in a cell or tissue. Developments in this rapidly advancing field have allowed unprecedented sensitivity, throughput, robustness and relative ease in not only the systematic identification of thousands of proteins and peptides from complex biological samples but also the relative (and absolute) quantitation of such molecular species. Degradomics utilizes techniques spanning many disciplines of biochemistry such as protease substrates, cleavage specificities, substrate chips, protease-activity chips, DNA microarray chips, or protease-specific protein chips and these have been extensively described elsewhere (Lopez-Otin and Overall, 2002).

Due to the high level of success of proteomic approaches employed in degradomics we will discuss some strategies to identify and quantify the cleavage specificities of unknown proteases, the substrates of a particular protease and validation techniques that could be employed to independently prove the results obtained from high-throughput technologies.

Identification of Protease Substrates

Substrates define protease function. The identification of cleaved proteins from the neo-N-termini in a sample after treatment of that sample with a protease or in a biological model expressing a mutant (inactive) form of the protease or not expressing it at all (knockout model) compared with a

control sample (showing regular proteolytic processing patterns) can yield information on the substrates that are targeted by that protease and hence their *in vivo* roles. All biological samples contain background proteolysis products. Therefore effective unbiased methods of substrate discovery must be employed that can efficiently discriminate between basal proteolysis and the protease-specific cleavage products with high confidence. Thus methods need to be employed that are capable of 'subtracting' background products in a control sample from the induced proteolysis products. Therefore, for this aim quantitative proteomics approaches need to be employed over non-quantitative approaches where one sample can be compared to another under the same conditions in order to subtract the background cleavage products, for example, control versus protease-containing. The following section describes experimental procedures for global approaches to identify substrates and we will not dwell upon techniques that do not provide this essential feature.

Classical gel-based methods

Theoretically, a protein that has been processed by a protease will have a lower mass and potentially a different isoelectric point (pI – the pH when a protein has no net charge) indicating a shift in the ratio of charged amino acids. Such an event can be easily visualized utilizing two-dimensional gel electrophoresis (2-DE). During this process, proteins are denatured, cysteine residues are blocked to prevent disulfide bond reformation by reduction and alkylation, and are separated according to charge on a low percentage acrylamide strip using isoelectric focusing (IEF). The strip is then embedded in a gradient acrylamide gel and proteins are resolved

by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) according to their mass. Following staining of the gel, a map of all proteins (as individual protein spots) resolved on the basis of both charge and mass is obtained, allowing for visual identification of differences in spot patterns (Gorg *et al.*, 2004). Protein in each spot can be extracted, digested with a site-specific enzyme (most commonly trypsin) and the resulting peptide fragments identified utilizing mass spectrometry (MS) – commonly known as peptide mass fingerprinting (PMF) (Figure 1; Pappin *et al.*, 1993). However, 2-DE alone is not quantitative, suffers from reproducibility problems and only allows for the profiling of proteins which are highly abundant and soluble in gel-compatible buffers.

Quantitative differences between samples using 2-DE can be achieved using fluorescently labeled dyes (Cy dyes), referred to as differential in gel electrophoresis (DIGE) (Unlu *et al.*, 1997). The samples to be compared are labeled with different Cy dyes and can also be compared to an internal standard. Different samples are electrophoresed on the same gel, which eliminates gel-to-gel variation as well as providing relative quantitation not only for all intact proteins but also their modified cleaved variants by visualizing the gel under the different wavelengths used to excite the fluorescent labels (Figure 1). Despite these advantages, however, DIGE often does not resolve cleavages near the N- or C-termini, does not identify the cleavage site, is labor intensive, low throughput, and requires scanners capable of excitation at several wavelengths.

While 2-DE and DIGE have relied on protein identification mainly by PMF and also suffered from dynamic range issues an alternative gel-based strategy was developed by the Cravatt

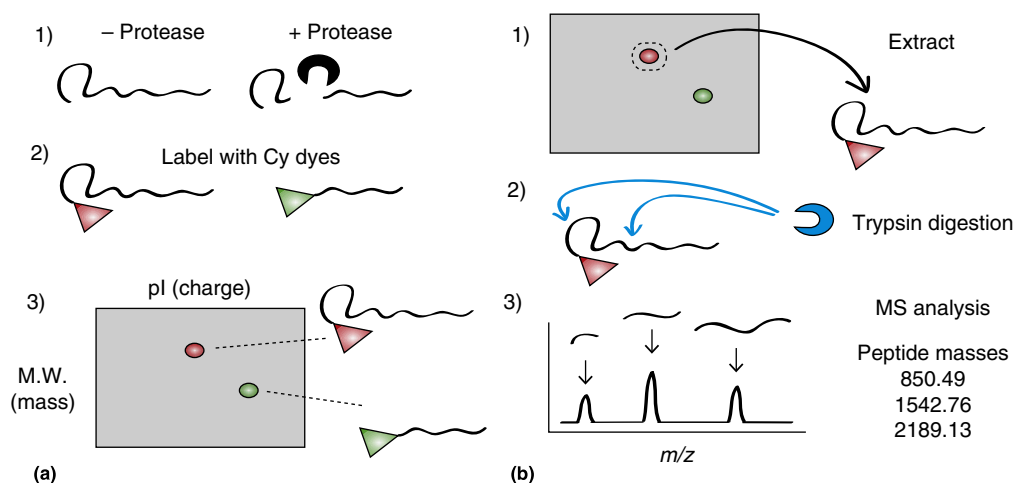


Figure 1 Differential in gel electrophoresis (DIGE). (a) (1) Two samples, one exhibiting more protease processing (shown in black) processing than the other; (2) Each is treated with a separate Cy dye (red and green in the figure); (3) Proteins are resolved by charge and mass using two-dimensional gel electrophoresis. Note that the wild-type protein (red) is of higher mass and different charge than the processed protein (green). Scanning at different wavelengths will show proteins with different Cy dyes which correspond to different samples. (b) (1) Protein of interest is extracted from the gel; (2) Protein extract is digested with a site-specific protease (e.g., trypsin); (3) MS analysis of peptide fragments derived from the protein of interest will result in peptide masses to be searched against a database yielding the identification of the protein. In addition to mass spectrometric analysis, proteins that have been excised in gel-based approaches can also have N-terminal peptides identified by Edman degradation. Edman sequencing can only target peptides with non-modified N-terminal residues as the first step is the reaction of phenylisothiocyanate with the amino group from N termini under alkaline conditions. This reaction cyclizes the N-terminus which then under heat and acidic conditions releases the N-terminal residue in the form of a phenylthiohydantoin amino acid derivative. This derivative can be identified by chromatography or electrophoresis and iteration of this many times yields the sequence of the target peptide (Edman *et al.*, 1950).

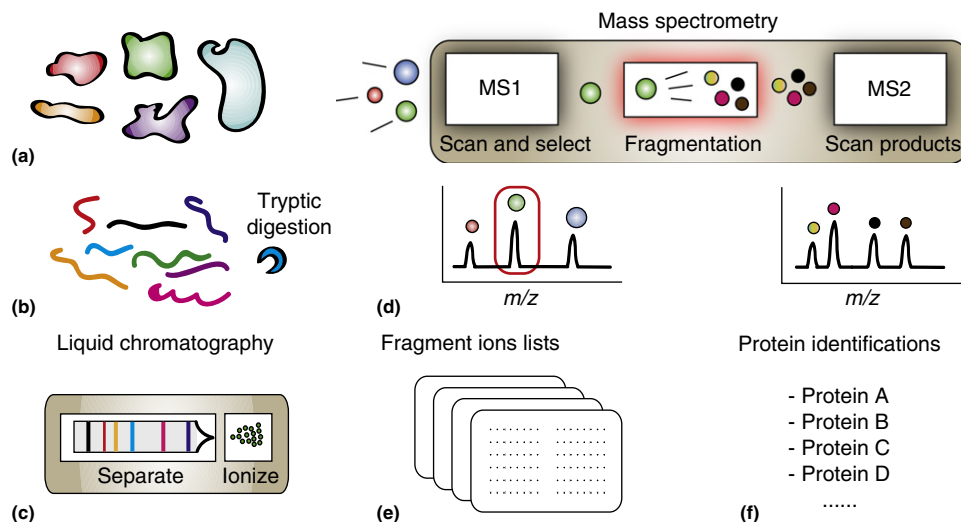


Figure 2 Liquid chromatography tandem mass spectrometry (LC-MS/MS). (a) Proteins are extracted and purified for analysis; (b) Proteins are digested with a site-specific protease (e.g., trypsin); (c) Peptide fragments are purified and resolved by reversed phase-HPLC coupled directly to ionization (e.g., ESI); (d) Ions enter the mass spectrometer where they are scanned in MS1. An ion selected for fragmentation (indicated by a red square in the MS1 spectrum) is fragmented and the fragment ions are scanned in MS2; (e) Fragmentation patterns are saved to lists which are searched with programs such as Mascot, Andromeda, and X! Tandem; (f) Protein identifications can be inferred from peptide sequences.

laboratory termed 'Protein Topography and Migration Analysis Platform – PROTOMAP' (Dix *et al.*, 2008). This is essentially identical to the 1-D gel-based substrate identification approach of Guo *et al.* (2002), but without isotopic labeling as used by Guo *et al.* (2002). Dix *et al.* utilize and distribute accessible software analysis to render the approach more amenable for widespread adoption. Guo's 1-D approach and PROTOMAP are conceptually distinct from the 2D gel-based methods. Each sample (e.g., control and test) is solubilized and electrophoresed on separate lanes of a 1-D SDS-PAGE gel. Each lane is then sliced into equal lengths and trypsinized in gel. Peptides from each lane slice can then be analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS) in order to identify the proteins in each gel band. LC-MS/MS consists of peptide separation by chromatography coupled directly to an ionization source (commonly electrospray ionization, ESI) and mass analysis in the mass spectrometer using one or various fragmentation techniques (Figure 2). In Guo's 1-D approach the proteins and cleaved fragments are manually determined. In PROTOMAP examination of protein identifications, spectral counts (the number of times a peptide ion was fragmented) and most importantly, sequence coverage, generates profiles for each protein – termed peptographs (Dix *et al.*, 2014). Guo's 1-D gel and PROTOMAP can yield semi-quantitative information on hundreds of proteins in a single analysis due to the high throughput of LC-MS/MS in comparison to PMF (Washburn *et al.*, 2001).

The effectiveness of PROTOMAP was initially demonstrated in the analysis of apoptosis-mediated protease substrates in a Jurkat T cell model (Dix *et al.*, 2008) and showed significant power over previous gel-based methodologies (PROTOMAP identified 260 proteins being cleaved during apoptosis while previous gel-based methods identified 49 (Brockstedt *et al.*, 1998; Gerner *et al.*, 2000; Thiede *et al.*, 2006; Thiede *et al.*, 2005). Furthermore, PROTOMAP circumvents

the need for Cy dyes (thus cheaper, and does not necessitate advanced equipment and imaging software), is technically easy, has a higher dynamic range than other 2D gel-based methods and provides full-length protein information and topology. However, despite the advantages and power of PROTOMAP, it was not originally intended to determine cleavage sites which can be easily missed still and it can suffer from poor resolution. Cleavage fragments near the termini are often not resolved. Nevertheless, PROTOMAP remains an inexpensive, effective and powerful method and has been extensively applied, for example, in the examination of proteolysis during malaria infection (Bowyer *et al.*, 2011), biomarker discovery in cancer (Shen *et al.*, 2012), and coagulation (Niessen *et al.*, 2011).

Gel-free methods (mass spectrometry-based)

Identification of protein fragments by gel-free methods is a non-trivial problem and still remains a technical challenge in proteomics in general, and not just degradomics, due to the high complexity and high dynamic range (i.e., highly abundant proteins obscuring low abundance ones) of biological samples being processed for MS analysis. Often a single protease-specific cleavage site must be identified amongst hundreds to thousands of proteins in the proteome in order to identify pathway involvement and develop biological knowledge. The sample is further complicated since prior to mass spectrometry the proteins in the sample need to be digested to smaller sizes to be amenable to MS. Peptides are easily handled, have well-defined chromatographic characteristics, and are amenable to MS, allowing accurate measurement on the basis of mass-to-charge ratio (m/z) (Abersold and Mann, 2003). Complex peptide mixtures are frequently fractionated by liquid column chromatography prior to MS. Furthermore, by colliding peptide ions inside the mass spectrometer (either using a variety of fragmentation methods including

collision-induced dissociation (CID) (Wells and McLuckey, 2005), higher collisional dissociation (HCD) (Olsen *et al.*, 2007), and electron transfer dissociation (ETD) (Syka *et al.*, 2004)) sequence-specific ions can be generated and used for sequence elucidation by means of database searching algorithms (e.g., Sequest, Eng *et al.*, 1994; Mascot, Perkins *et al.*, 1999; X! Tandem, Craig and Beavis, 2003 and Andromeda, Cox *et al.*, 2011). The availability of completely sequenced whole genomes has provided a rich source of information that has enabled proteomic technologies to be high-throughput. This is due to the ability to search MS/MS spectra against protein databases of the most commonly used organisms in a short period of time.

In order to identify one or a few protease cleavage sites in a whole proteome, it is imperative to simplify the sample and to enrich for the relevant neo-N or neo-C terminal peptides. Separating peptides using several different chromatographic chemistries, high speed and high resolution mass spectrometers and comprehensive computer algorithms coupled to powerful computing nodes now routinely yields peptide sequence identifications in the tens of thousands – orders of magnitude higher than gel-based methods or Edman sequencing reactions (Edman *et al.*, 1950).

In addition to ‘identification’ of peptide sequences, relative ‘quantification’ of peptides (or absolute quantification if standards with known concentration are spiked into the sample; Fahlman *et al.*, 2014) between control, protease-treated or even among different test conditions of samples can be confidently obtained using chemical labeling strategies (Ross *et al.*, 2004; Thompson *et al.*, 2003).

Quantitation can be performed in three modes: MS1 quantification (Egertson *et al.*, 2013) where quantitative information is obtained in the primary scan (MS1 scan) in the

mass spectrometer after labeling by dimethylation or SILAC and the peptide sequence can be determined from a subsequent fragmentation event of the quantified ion; MS2 quantification carried out in the tandem MS scan (MS2 scan, containing all fragment ion signals used for peptide sequence inference) (e.g., iTRAQ tags (isobaric tags for relative and absolute quantification; Ross *et al.*, 2004)) or TMT tags (tandem mass tags; Thompson *et al.*, 2003; Figure 2), and recently MS3 quantification which is performed on reporter ions generated during fragmentation of fragment ions (that is, a further iteration of quantification using MS2 methods with iTRAQ and TMT). During MS1-based quantification, quantification is provided without the need for peptide fragmentation as the tags have different masses which the mass spectrometer can resolve; the fragmentation is necessitated solely for identification. In contrast, MS2-based quantification requires fragmentation for simultaneous quantification and identification as the tags have the same mass (isobaric) and their reporter regions have different masses that are used in MS2 analysis. An additional round of fragmentation on product ions can yield more confident quantification results (due to lack of interfering peptide ions in a close mass range) in MS3 quantification. So that identical peptides with different isobaric tags fly together in MS1 mode the different tags must add up to the same mass. Therefore, balance groups within each tag are included so that the fragmented balance component of the tag has a different mass and importantly the labeled peptide displays a tag with a different mass for each of the 4, 6, 8, or recently 10 tags (Figure 3). Terminal amine isotopic labeling of substrates (TAILS), described below, can be utilized with either MS1, MS2, or MS3 quantitation techniques depending on the isotopic label, but for the purposes of illustration only isotopically labeled dimethyl tags (MS1 quantification) will be

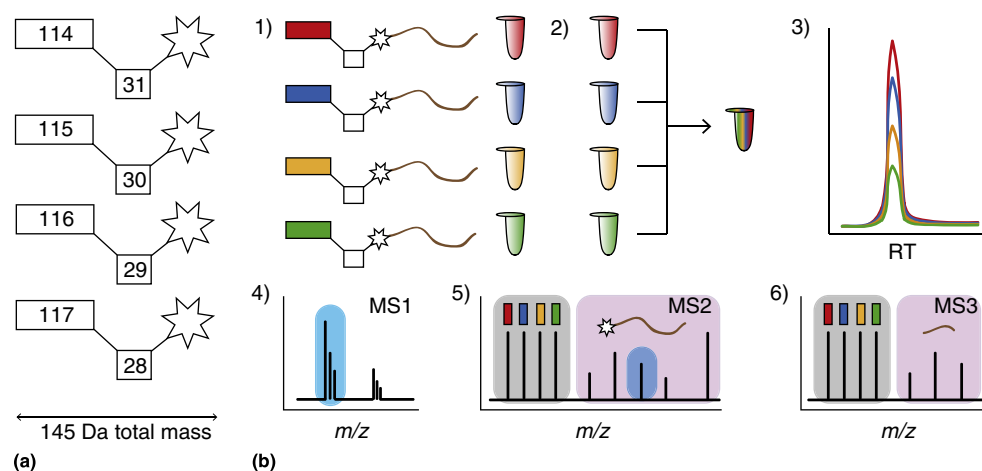


Figure 3 MS2 quantification scheme using 4-plex iTRAQ tags. (a) Schematic representation of isobaric iTRAQ tags. Each tag has a reporter ion group (114 through 117 Da), a peptide-reactive group (indicated by a star) and a balance group (31 through 28 Da). The balance group ensures that all tags add up to 145 Da. (b) (1) Each sample (here a single peptide indicated by a wavy line) is labeled with a different iTRAQ tag (reporter groups indicated by different colors and kept in separate tubes); (2) Individual labeled samples are then be combined prior to LC-MS/MS; (3) All iTRAQ tags have identical chromatographic properties allowing mixing of samples and co-elution during LC-MS/MS; (4) For a population of the same peptide labeled with different iTRAQ tags, a single peptide ion is detected during MS1 and selected for fragmentation; (5) The selected peptide ion is fragmented yielding reporter ions (highlighted in gray) and peptide sequence ions (highlighted in purple) in the MS2 scan. Relative intensities of reporter ions provides quantitative information while peptide sequence ions provides identification information; (6) Further selection and fragmentation of peptide fragment ions from the MS2 scan can yield refined reporter ions for higher accuracy quantitation (MS3 quantification).

discussed. Subtiligase biotinylation to date has not employed quantification by isotopic labeling whereas Combined Fractional Diagonal Chromatography (COFRADIC) has mainly utilized MS1 quantification.

Relative quantification using dimethylation

For relative quantification of peptides between samples, for instance a test sample with over-expressed or deleted protease of interest compared to a control sample only displaying background levels of proteolytic processing, samples are each treated with a different formaldehyde tag that reacts with primary amines of free N-termini (α -amine groups) and lysine side chains (ϵ -amine groups). These tags are chemically identical to each other – thus displaying identical chromatographic properties, equal signal in MS intensity as well as linearity in response to quantity (i.e., truly quantitative chemical labels). The only difference between the tags used for each sample are the number of neutrons – namely, the ‘light’ version of the tag is CH_2O , the ‘medium’ tag being CD_2O (D being deuterium) and the ‘heavy’ form of the tag being $^{13}\text{CD}_2\text{O}$. Hence, there is a difference of 4 Da between tags, giving each peptide in the samples a unique identifier – the intensity of each ion providing relative quantitation between samples (Boersema *et al.*, 2009) while fragmentation of the peptide ion yields product ions that provide the information for the identification of the peptide sequence.

Negative Selection of N-termini by Combined Fractional Diagonal Chromatography (COFRADIC)

The earliest methods for isolating peptides that originate from proteolytic events in a cell (native peptides) is COFRADIC (Gevaert *et al.*, 2002). COFRADIC was first used to isolate methionine-containing peptides from the model organism *Escherichia coli*. Over 800 proteins were identified from 5×10^7 cells and COFRADIC was estimated to be 100 times more sensitive than the previously described gel-based methodologies (Gevaert *et al.*, 2002). COFRADIC exploits two key physical principles that peptides display during chromatographic separation. Firstly, peptides will elute consistently at the same retention time during chromatographic separation between analytical runs if identical parameters (flow rate, buffer compositions, column dimensions, etc.) are used; and secondly, chemical modification of a peptide will result in an altered interaction with the solid-phase of the chromatographic column that manifests as a shift in retention time.

To identify peptides that have originated during protease processing in the cell, proteins are isolated and acetylated to block all free amine groups on N-termini and lysine side chains, rendering them unreactive. Proteins are then digested with trypsin that generates a variety of reactive free internal N-termini and separated by reverse phase (RP) HPLC into several fractions – with each peptide eluting after a particular retention time. Each fraction is treated with 2,4,6-trinitrobenzenesulfonic acid

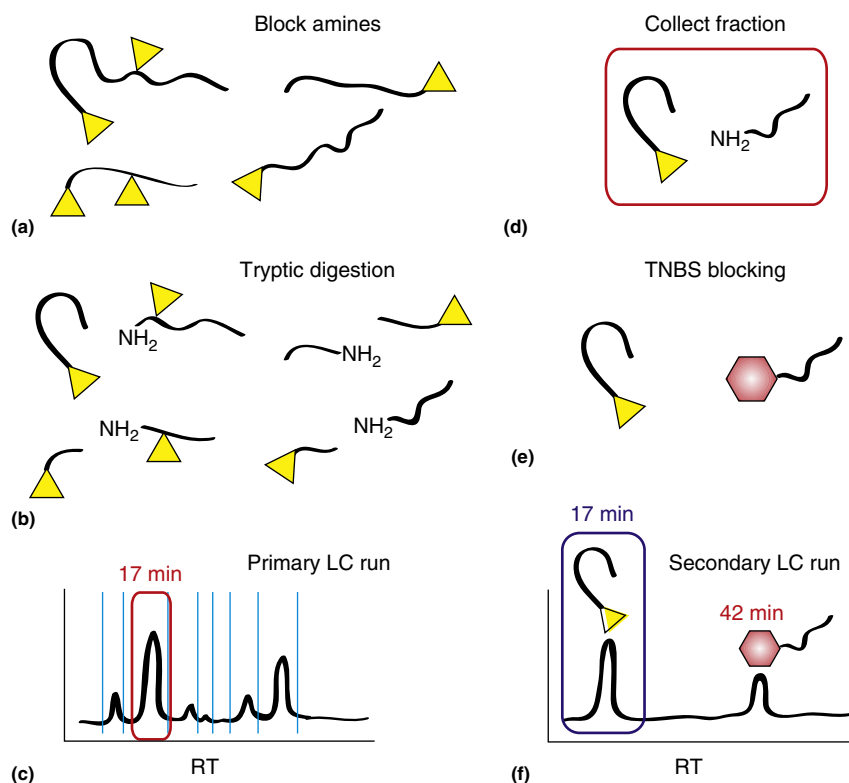


Figure 4 Combined fractional diagonal chromatography (COFRADIC). (a) Proteins have their amines blocked (e.g., by acetylation); (b) Proteins are digested with a site-specific protease (e.g., trypsin or GluC) generating a variety of internal N-termini; (c) Peptides are separated by HPLC; (d) Considered here peptides at 17 min retention time; (e) Peptide fraction is treated with TNBS; (f) Peptide fraction is again resolved using identical HPLC, however due to TNBS treatment, internal N-termini display a strong hydrophobic shift. Native peptide still elutes at 17 min but internal N-terminus elutes much later and can be discarded.

(TNBS) which reacts selectively with the free α -amino groups (namely, all internal N-termini resulting from tryptic digestion) and each fraction is then run again on RP-HPLC using identical conditions as the primary run. Neo-N-termini from proteins in the original sample were blocked (acetylated) and so will elute at the same time as in the primary run and can easily be collected. However, peptides derived from tryptic digestion are TNBS-modified and consequently display a strong hydrophobic shift (increased retention time) and can now be discarded (Figure 4; Gevaert *et al.*, 2003). Thereby, neo-N-termini have been refined from the total peptide pool and can be identified by LC-MS/MS. COFRADIC enables the separation not only of mature protein and neo-N-terminal peptides but of any desired modification provided that the reaction step between the primary and secondary runs are specific for this modification – making COFRADIC an extremely versatile and effective targeted technique for sampling the proteome with exquisite resolution that can be automated. The main caveats lie principally in the need for multiple chromatographic steps, which demand technical reproducibility, large sample amounts to account for potential losses at each step, but perhaps above all, the extensive chromatography time required to obtain such samples for one experiment followed by many MS analyses to analyze all the sample fractions, even > 100. Despite this, COFRADIC remains an attractive and versatile method to target the degradome and other post-translational modifications.

Negative selection of N-terminal peptides by Terminal Amine Isotopic Labeling of Substrates (TAILS)

A more recent simplified N-terminomics procedure termed TAILS (Kleifeld *et al.*, 2010, 2011; Doucet *et al.*, 2011) has the advantage of starting with as little as 200 μ g protein per channel followed by 1–3 MS analyses only. Amino groups comprising the α amines of N-termini of mature protein and cleaved neo-N-termini of proteins in the sample and of lysine side chains are chemically protected at the whole protein level.

This is the key as it is simple to discriminate any N-terminus of native proteins that were present in the sample before trypsinization from internal N-terminal peptides arising after tryptic digestion during proteomic processing. Chemical blocking of α - and ϵ -amine groups (i.e., those present on the N-terminus and lysines respectively) with different isotopes, renders them unreactive as well as providing each sample with a unique mass tag that allows quantitation during MS analyses.

Following amine blocking (e.g., dimethylation with CH_2O , CD_2O , or $^{13}\text{CD}_2\text{O}$, or MS2 tags like iTRAQ and TMTs) protein samples are combined and digested with trypsin to yield free internal N-termini that are reactive. For negative selection, the peptide mixture is reacted with a commercially available highly branched polymer containing 1000s of aldehyde groups (HPG-ALD, see Relevant Websites), which captures free reactive N-termini under reductive conditions, while all peptides that do not have a free amino group (i.e., protease substrate neo-N-termini and natural N-termini blocked by dimethylation in TAILS or naturally blocked by acetylation or pyroglutamate formation) will not bind and therefore remain free in solution. The polymer and bound peptides are then removed by ultrafiltration, and the blocked peptides can be collected for LC-MS/MS analysis (Kleifeld *et al.*, 2010; Figure 5). The product is a highly simplified mixture of peptides that enables easier identification and quantitation by LC-MS/MS, yet as with all methodologies there are caveats and advantages. This method offers a technically much simpler protocol than COFRADIC, as well as a statistical platform for data analysis of the identifications when using MS2 quantification (CLIPPER) (auf dem Keller *et al.*, 2010), all reagents are commercially available, and the process can be multiplexed for simultaneous analyses of several samples. Additionally, the reaction of internal tryptic peptide N-termini with HPG-ALD is efficient, rapid, and demonstrates minimal non-specific peptide binding. TAILS has been widely adopted by many companies and laboratories now by virtue of its rapid processing and ease of preparation, minimal mass spectrometry time

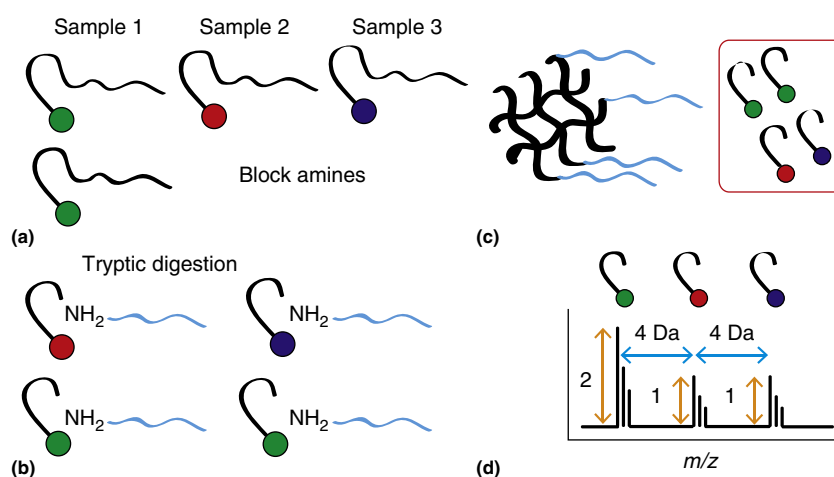


Figure 5 Terminal Amine Isotopic Labeling of Substrates (TAILS). (a) Consider three protein samples. Each sample is treated with a different dimethyl label (light, medium and heavy – each differing in mass by 4 Da). Each sample has all amine groups blocked; (b) Samples are combined and digested with a site-specific protease (e.g., trypsin); (c) Internal N-termini bind aldehyde groups in the HPG-ALD polymer, retaining all neo-N-termini in solution; (d) HPG-ALD containing all internal N-termini peptides are removed by ultrafiltration and neo-N-termini are identified and quantified by LC-MS/MS as observed by a difference of 4 Da between peptide ions corresponding to a single native N-terminus.

required for analysis with the advantage of also identifying the N-termini of mature proteins which can be used for statistical modeling and data normalization, comprehensive analysis of the entire N-terminome, as well as for identifying N-terminal peptide cleavage sites.

TAILS also has variation of the approach for the isolation of C-terminal peptides (C-TAILS) that although effective at targeting the C-terminome still cannot achieve similar number of identifications as N-TAILS (Schilling *et al.*, 2010). However, with no alternative C-terminal enrichment methods available, C-TAILS remains the method of choice with the caveat of having lower peptide identifications than for N-TAILS.

In vivo analyses are confounded by a number of different variables that makes the identification and quantification of neo-N-termini challenging regardless of the methodology employed. Not only do *in vivo* analyses suffer from biological variability between animals or tissues but the variability in quantification of protease substrates is also attributed to other factors: Tissues are responsive biological systems that upon stimulus will alter their transcriptional regulation and hence protein synthesis, vascular permeability (which alters the plasma protein composition in pathologically affected tissues), and cell migration (e.g., the infiltration of neutrophils and macrophages upon inflammation or infection (McQuibban *et al.*, 2002)). Thus, an influx of proteins from either serum entering at higher amounts due to increased blood flow and vessel permeability, or turnover products from effector cells acting at the site of stimulus can increase the amount of particular proteins that would lead to a perceived increase of protease activity on the basis of increased amount of neo-N-termini pertaining to that protein. An example of overcoming these issues was demonstrated in the examination of proteolytic events of complement activation during skin inflammation (auf dem Keller *et al.*, 2013). The authors used a murine model of skin inflammation matrix metalloproteinase 2 (MMP2)-deficient mice and utilized N-TAILS with MS2 quantification (iTRAQ-TAILS) to elucidate protease substrates *in vivo*. A very effective way to truly unravel if the protease substrates quantified are truly a biological signature in response to a stimulus is to normalize the results obtained in the N-terminome analysis to whole protein ratios. Analysis of natural N-termini ratios provides information if there is altered protein synthesis or import from plasma or infiltrating cells. TAILS offers a way to examine the changes in proteome abundance by examining the relative ratios of peptides prior to HPG-ALD depletion (preTAILS). Since proteins are labeled with iTRAQ reagent, the peptides remain labeled (on N-termini and lysine residues) following tryptic digestion which permits direct LC-MS/MS analysis that complements TAILS (preTAILS and TAILS analyses). Comparisons of preTAILS relative ratios across native N-termini and proteins to substrates identified by TAILS provided the normalization (protein context) required to infer true protease signatures during skin inflammation (auf dem Keller *et al.*, 2013).

Positive Selection of N-Terminal Peptides by Subtiligase Biotinylation

Neo-N-termini can be labeled with a biotin tag that is cleavable by means of the highly site-specific TEV (tobacco etch

virus) protease. In this case, the ϵ -amino group of lysine residues are not protected as N-terminal specific biotinylation is performed by subtiligase (Mahrus *et al.*, 2008), a genetically-engineered enzyme that exclusively attaches an engineered biotin tag (a biotin moiety linked to the TEV cleavage site) to the N-terminus of available proteins and peptides. Proteins are digested with trypsin (to allow for database identification following LC-MS/MS) and the entire peptide samples are passed through a streptavidin column washing out all trypsin-generated internal N-termini. Peptides are cleaved from the biotin tag by TEV protease, exclusively at the TEV cleavage site linking the biotin tag and native peptide. As a consequence of such processing, all TEV processed native peptides will initiate with the dipeptide sequence Ser-Tyr which can also serve as diagnostic ions during the fragmentation event in LC-MS/MS, thus providing a simple yet effective manner of validating peptide identifications by positive selection of native termini (Mahrus *et al.*, 2008; Figure 6).

Positive enrichment of subtiligase biotinylated neo-N-termini offers a number of advantages over COFRADIC: enrichment is relatively simple and the chromatographic separation is faster and requires less instrumentation than in COFRADIC, allowing for higher throughput as well as diagnostic ion confirmation in the case of the subtiligase method. However, limitations with positive enrichment include false-positive binding events in the streptavidin affinity step (an issue circumvented entirely by COFRADIC), a requirement for relatively large sample amounts (50–100 mg) due to lower efficiencies in labeling, loss of natural N-termini that are not captured if they are N-terminally blocked, for example, acetylated, so these peptides cannot be used for data normalization, N-terminomics analyses or substrate discovery (if cut within the N-terminal peptide leading to low peptide ratios due to cleavage versus the intact N-terminal peptide of the control sample), and perhaps a troublesome issue is that subtiligase is a patent-protected enzyme – making it less appealing than easily commercially available alternatives. However, COFRADIC and subtiligase-directed biotinylation are potentially compatible with isotope labeling, although only reported using MS1 quantification for COFRADIC, allowing quantification by mass spectrometry and thus providing information on protease cleavage events and protein abundance across different tissues/conditions.

Identification of Protease Cleavage Motifs

Proteomic identification of protease cleavage site specificity

The availability of complete genome sequences, largely due to high-throughput sequencing projects, now offers a rich source of information that can be exploited to profile the specificities of proteases. An entire tryptic digested proteome was conceived as an extensive peptide library offering a plethora of biologically relevant amino acid combinations that will most likely contain cleavage sequences specific to a protease. This is the basis for proteomic identification of protease 'cleavage site specificity (PICS)' (Schilling and Overall, 2008; Schilling *et al.*, 2011) and although other methods exist to profile cleavage specificities of proteases (such as phage display libraries;

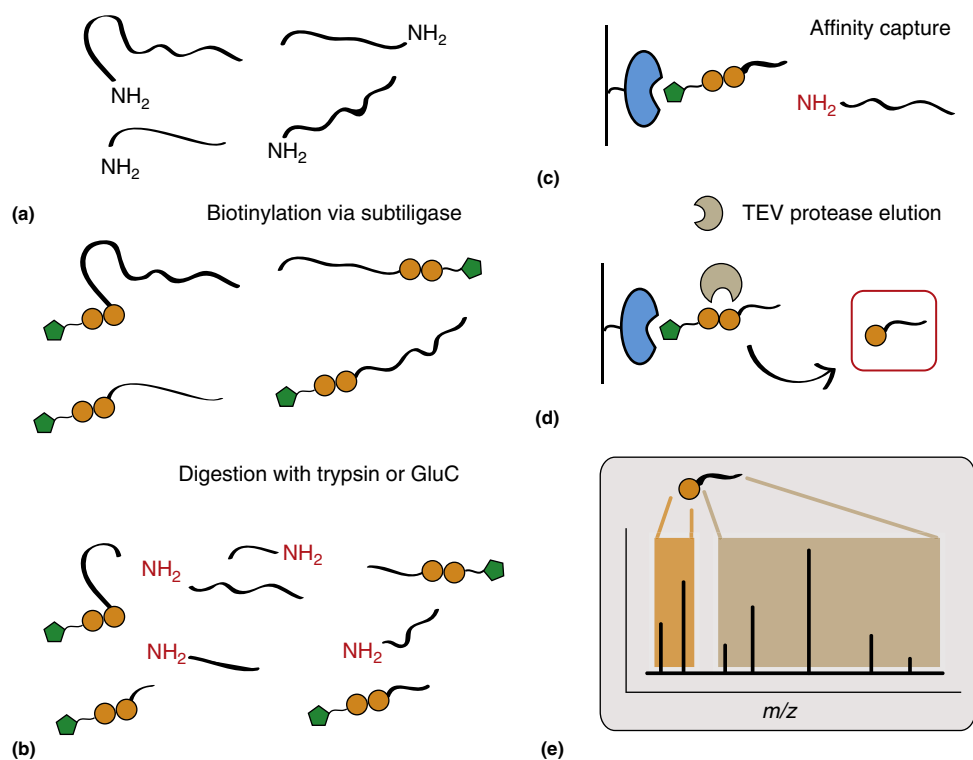


Figure 6 Enzymatic biotinylation. (a) Proteome is treated with subtiligase, which exclusively biotinylates N termini; (b) Protein sample is digested with a site-specific protease as trypsin or GluC; (c) Streptavidin affinity capture retains biotinylated neo-N-termini as native N-termini pass through the column; (d) Biotinylated peptides are released by TEV protease cleavage and collected for LC-MS/MS (the orange circle denotes the leader dipeptide Ser-Tyr characteristic after TEV protease cleavage of the biotin reagent); (e) Signature pair of ions (Ser-Tyr) can be used as diagnostic positive controls during MS analysis. Here the orange head corresponds to Ser-Tyr and the fragment ions corresponding to Ser-Tyr are highlighted orange whereas sequence-specific ions are highlighted gray.

Matthews and Wells, 1993), we will only discuss PICS as a gel-free proteomics technique as it alone has the advantage of profiling both the prime and non-prime sites simultaneously in the same analysis.

To perform a PICS experiment, an entire proteome (typically the human proteome, as it offers a larger list of proteins compared to microorganisms or less complex eukaryotes) is digested with a site-specific protease (usually trypsin and GluC are used to make separate libraries. Both trypsin and GluC are utilized to maximize coverage of the libraries in complementary fashion. Namely, the cleavage site specificity for trypsin is C-terminal to arginines and lysines whereas for GluC it is C-terminal to glutamic acid. Thus, tryptic peptide libraries contain internal Asp and Glu residues, GluC libraries contain internal Lys and Arg residues) and all amine groups are blocked (rendering them unreactive to the subsequent biotinylation steps). Such a complex peptide mixture can act as a library with the potential to contain sequences that are cleaved by the test protease. The test protease is allowed to proteolytically process the peptide library yielding a variety of unblocked neo-N-termini on the prime side of the cleavage site, which are positively selected using a chemical biotinylation strategy using NHS-biotinylation. Biotinylated peptides can then be isolated following disulfide bond reduction and identified by LC-MS/MS (Schilling and Overall, 2008).

Information obtained through LC-MS/MS yields the peptide sequence with the N-terminus of the neo-N-terminal peptide being the prime side of the cleavage site (P' site), which can then be mapped back to the genome to determine the non-prime side, P sites N-terminal to the scissile bond. Assembling hundreds of the cleavage sites enables one to infer a consensus substrate motif (Figure 7). Thus, PICS offers a rapid, inexpensive and relatively simple approach to confidently identify protease consensus cleavage sites from up to thousands of cleavage events by using readily available and established proteomic technologies and software.

Validation of Degradomic Data

Perhaps as important as having high confidence proteomic data (both in terms of identification and quantification of substrates) is to provide insight into biological processes by validation of these data. Validation experiments by biological assays which address hypotheses generated by large-scale proteomic experiments convert possible biological information into actual biological information. This is a particularly important distinction as all large-scale data is subject to large biological variability that can be statistically difficult to analyze in terms of relatively low numbers of biological

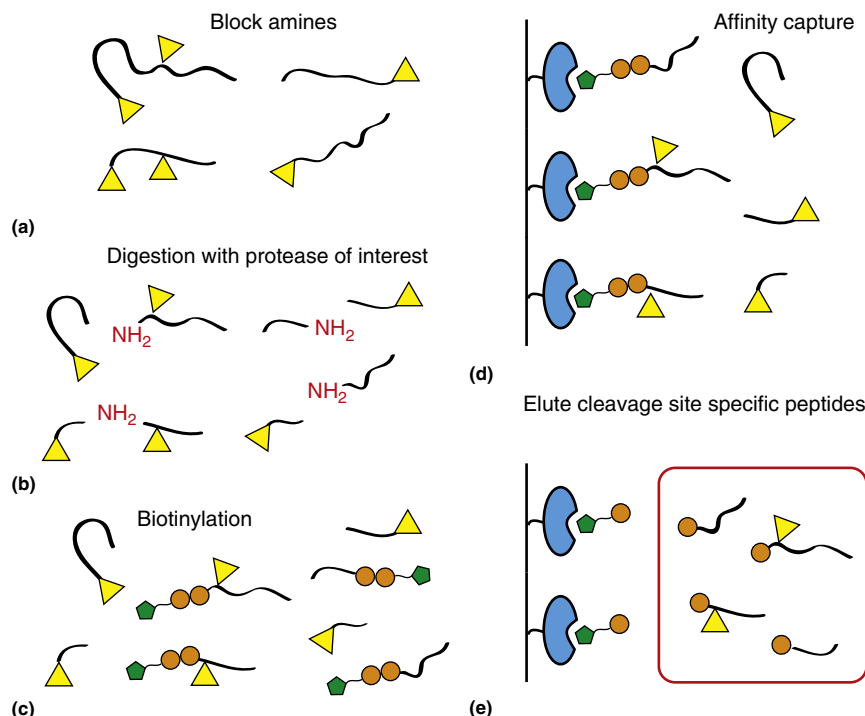


Figure 7 Proteomic identification of protease cleavage site specificity (PICS). (a) A proteome sample is digested to peptides by trypsin and then has all amines blocked; (b) the test protease is allowed to cleave the proteome derived peptide library resulting in neo-N-termini at the cleavage site; (c) neo-N-termini are chemically biotinylated; (d) utilizing a streptavidin capture column, all tryptic peptides are washed while only retaining those caused by the protease treatment; (e) biotinylated peptides are eluted and then identified by LC-MS/MS. Following bioinformatic analysis a consensus motif can be constructed on available proteomic data and genome sequences.

samples that can be processed per proteomic approach as well as the inherent technical variations in all MS-based technologies. In other words, by using complementary techniques, biological knowledge can be obtained by converting moderately quantitative information on several substrates (degradomic profiling) into highly quantitative (and phenotypic) information on few substrates.

There have been numerous publications now utilizing methods such as COFRADIC and TAILS in both *in vitro* models (e.g., Becker-Pauly *et al.*, 2011; Holmberg *et al.*, 2013; Papadioti *et al.*, 2012; Starr *et al.*, 2012; Wejda *et al.*, 2012) as well as *in vivo* analyses in human and murine models (auf dem Keller *et al.*, 2013; Jefferson *et al.*, 2011; Lange *et al.*, 2014; Lewandrowski *et al.*, 2009; Marchant *et al.*, 2014). Identification of protease substrates by mass spectrometry requires a combination of manual annotation and validation of individual mass spectra but also validation by usage of different database searching algorithms and post-processing tools (Keller *et al.*, 2002; Nesvizhskii *et al.*, 2003). These data can be confirmed by *in vitro* assays to validate cleavage. However, the consistent confirmation of cleavage sites identified by mass spectrometry with confirmatory Edman degradation has now nearly eliminated the need for *in vitro* validation of cleavage sites. This signifies that mass spectrometric data needs to be validated in terms of protease activity, and phenotypic characteristics in the model being investigated.

Why is validation of accurate mass spectrometric data one aspect for inferring protease biology? Proteases are subject to

many different stimuli that can affect their conformation, specificity, subcellular localization, activation and inhibition which all contribute to their overall activity. The use of a particular biological model, be it *in vitro* or *in vivo*, comes with advantages and disadvantages that can make the interpretation of degradomic data challenging. For instance, overexpression of a protease in tissues not normally expressing the protease will generate a situation of high protease:substrate thus resulting in an artificially high number of substrates being identified that might never occur *in vivo*. Overexpression in tissues not normally expressing the protease also would lead to substrates being cleaved that are not normally encountered by the active protease. Thus, overexpression studies have considerably less power than analysis of knockout animals.

Additionally, there is biological redundancy – that is, proteases from the same enzyme family may compensate for the induced loss of a particular protease (Fortelny *et al.*, 2014). Some of these questions may not be exclusively addressed using the proteomic data alone but can be using alternate verification methods such as selected reaction monitoring (SRM). SRMs are a label-free method for quantifying selected peptides across several samples using LC-MS/MS; and more importantly the use of a variety of cell lines and animal models that can be integrated into biological assays testing overlapping proteomic data sets to confirm hypotheses. Further to SRMs, validation of protease activity at a particular cleavage site can be performed utilizing proteolytic signature peptides (PSP). A unique tryptic peptide spanning a cleavage

site in a protein can be monitored by SRM or MS2 labeling as an indicator of activation of the protease recognizing that cleavage site motif. If the protease is active, the tryptic peptide will not be present. Rather, two semi-tryptic peptides will be present (their termini being at the cleavage site). Normalization across a protein can be achieved by also monitoring peptides on the same protein that do not cover the cleavage site motif ('Standard of expressed protein', STEP) thus accounting for protein abundance. Because synthetic standard peptide mixtures (PSP and STEP) are spiked into the sample for quantification analysis, not only relative quantification can be obtained, but also, absolute quantification over a large dynamic range and low limit of quantification (Fahlman *et al.*, 2014). This provides a relatively simple, powerful, and targeted method to monitor protease activities circumventing issues encountered by activity-based probe technologies and has been recently used to profile MMP2 activation and activity *in vivo* (Fahlman *et al.*, 2014).

There are several tools for the *in vivo* analysis of proteases that typically involve mouse models due to their relative ease of manipulation and high biological similarity that can be extrapolated for further human research. The scope of this chapter is not to examine all of these tools, but they have been reviewed extensively elsewhere (Overall and Blobel, 2007). For testing the global effects of a protease, knocking-out the protease or knocking-in an inactive protease (inserting a mutation in the active site rendering the protease inactive) may be employed to observe the phenotypic effects *in vivo* and also correlate to substrates identified by proteomic experiments as these two functionally similar methods address different questions. Namely, if the protease exists in a complex, the biological effect will likely be different if that protease is removed entirely (knockout) or simply inactivated in the complex (catalytic-site-inactivating knock-in)? In an analogous fashion, upon identification of a particular protease substrate associated with a phenotype, strategies include knocking-in the substrate with a modified protease cleavage site making it resistant to the tested protease, or knocking-out the substrate (Overall and Blobel, 2007).

Conclusions

The ability to globally profile different aspects of protease biology utilizing MS-based methods has produced a revolution in how protease function is viewed today. Proteases are compartmentalized to prevent substrate processing, they exist in complexes and trigger many signals (both proteolytic and non-proteolytic, Turk *et al.*, 2012), that result in modified phenotypes, such as the coagulation cascade (Hoffman and Monroe, 2007) or apoptosis (Elmore, 2007). Novel techniques to analyze proteolytic processing and the increasing amount of information regarding proteases being stored and made publicly available (for example, MEROPS, Rawlings *et al.*, 2012 and TopFIND databases, Lange and Overall, 2011, 2013; Lange *et al.*, 2012), has shown that proteases exist in networks, originally proposed and termed the protease web (Dean and Overall, 2007) and validated recently as a highly pervasive interdependent network of interactions (Fortelny *et al.*, 2014). Proteases are dynamic and their activities not only affect their direct substrates but in turn the entire system

to which they are connected that can dictate the balance between a signal being activated or remaining in homeostasis (Overall and Blobel, 2007; Turk *et al.*, 2012).

In addition, increasing power in technologies not only associated with MS but also cell lines, animal models, microscopy, *etc.*, have also opened a new frontier to validate these systems-wide studies and have begun to reveal the exquisite regulation these webs have and their potential as targets for drugs and novel therapeutics. Protease research is still exploring new avenues of biological regulation and the development and refinement of degradomic techniques is critical to answer questions such as interplay in the protease web, does proteolytic signaling affect kinase signaling or vice-versa, how do complexes interact once a particular proteolytic signal has targeted it, *etc.* One recent example of novel biological regulation mediated by proteases is the case of MMP12 during viral infection in a murine model (Marchant *et al.*, 2014). MMP12^{-/-} knockout animals challenged with coxsackievirus display significant morbidity and are unable to mount an effective immune response. Utilizing a combination of TAILS and chromatin immunopurification-sequencing (ChIPSEQ) MMP12 was shown to act not only as an extracellular protease, but also an intracellular protease and strikingly, as a transcription factor. ChIPSEQ analysis (which allows for the identification of transcriptionally regulated genes) revealed that MMP12 directly binds DNA and represses several genes. Furthermore, using TAILS it was found that 177 substrates were processed by MMP12 that also were transcriptionally regulated by MMP12, *i.e.*, these are dual regulated substrates by MMP12. Combination of various genomic, proteomic *in vitro* and *in vivo* models, degradomics is now flourishing into an integrated field that is now allowing for the examination of new frontiers in protease biology.

See also: Protein Synthesis/Degradation: Protein Degradation – Intracellular: Ubiquitin, Ubiquitin-Like Proteins, and Proteasome-Mediated Degradation

References

- Aebersold, R., Mann, M., 2003. Mass spectrometry-based proteomics. *Nature* 422, 198–207.
- Ajami, K., Pitman, M.R., Wilson, C.H., *et al.*, 2008. Stromal cell-derived factors 1alpha and 1beta, inflammatory protein-10 and interferon-inducible T cell chemo-attractant are novel substrates of dipeptidyl peptidase 8. *FEBS Letters* 582, 819–825.
- auf dem Keller, U., Prudova, A., Eckhard, U., Fingleton, B., Overall, C.M., 2013. Systems-level analysis of proteolytic events in increased vascular permeability and complement activation in skin inflammation. *Science Signaling* 6, rs2.
- auf dem Keller, U., Prudova, A., Gioia, M., Butler, G.S., Overall, C.M., 2010. A statistics-based platform for quantitative N-terminome analysis and identification of protease cleavage products. *Molecular & Cellular Proteomics* 9, 912–927.
- Barrett, A.J., McDonald, J.K., 1986. Nomenclature: Protease, proteinase and peptidase. *Biochemical Journal* 237, 935.
- Becker-Pauly, C., Barre, O., Schilling, O., *et al.*, 2011. Proteomic analyses reveal an acidic prime side specificity for the astacin metalloprotease family reflected by physiological substrates. *Molecular & Cellular Proteomics* 10.M111 009233.
- Boersma, P.J., Rajmakers, R., Lemeer, S., Mohammed, S., Heck, A.J., 2009. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nature Protocols* 4, 484–494.

- Bowyer, P.W., Simon, G.M., Cravatt, B.F., Bogoy, M., 2011. Global profiling of proteolysis during rupture of *Plasmodium falciparum* from the host erythrocyte. *Molecular & Cellular Proteomics* 10, M110. 001636.
- Brockstedt, E., Rickers, A., Kostka, S., *et al.*, 1998. Identification of apoptosis-associated proteins in a human Burkitt lymphoma cell line. Cleavage of heterogeneous nuclear ribonucleoprotein A1 by caspase 3. *Journal of Biological Chemistry* 273, 28057–28064.
- Ciechanover, A., 2005. Intracellular protein degradation: From a vague idea thru the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. *Cell Death and Differentiation* 12, 1178–1190.
- Cox, J., Neuhauser, N., Michalski, A., *et al.*, 2011. Andromeda: A peptide search engine integrated into the MaxQuant environment. *Journal of Proteome Research* 10, 1794–1805.
- Craig, R., Beavis, R.C., 2003. A method for reducing the time required to match protein sequences with tandem mass spectra. *Rapid Communications in Mass Spectrometry* 17, 2310–2316.
- Dean, R.A., Cox, J.H., Bellac, C.L., *et al.*, 2008. Macrophage-specific metalloelastase (MMP-12) truncates and inactivates ELR + CXC chemokines and generates CCL2, -7, -8, and -13 antagonists: Potential role of the macrophage in terminating polymorphonuclear leukocyte influx. *Blood* 112, 3455–3464.
- Dean, R.A., Overall, C.M., 2007. Proteomics discovery of metalloproteinase substrates in the cellular context by iTRAQ labeling reveals a diverse MMP-2 substrate degradome. *Molecular & Cellular Proteomics* 6, 611–623.
- Dix, M.M., Simon, G.M., Cravatt, B.F., 2008. Global mapping of the topography and magnitude of proteolytic events in apoptosis. *Cell* 134, 679–691.
- Dix, M.M., Simon, G.M., Cravatt, B.F., 2014. Global identification of caspase substrates using PROTOMAP (protein topography and migration analysis platform). *Methods in Molecular Biology* 1133, 61–70.
- Doucet, A., Butler, G.S., Rodriguez, D., Prudova, A., Overall, C.M., 2008. Metadegradomics: Toward *in vivo* quantitative degradomics of proteolytic post-translational modifications of the cancer proteome. *Molecular & Cellular Proteomics* 7, 1925–1951.
- Doucet, A., Kleifeld, O., Kizhakkedathu, J.N., Overall, C.M., 2011. Identification of proteolytic products and natural protein N-termini by Terminal Amine Isotopic Labeling of Substrates (TAILS). *Methods in Molecular Biology* 753, 273–287.
- Edman, P., Hogfeldt, E., Sillen, L.G., Kinell, P., 1950. Method for determination of the amino acid sequence in peptides. *Acta Chemica Scandinavica* 4, 283–293.
- Egertson, J.D., Kuehn, A., Merrihew, G.E., *et al.*, 2013. Multiplexed MS/MS for improved data-independent acquisition. *Nature Methods* 10, 744–746.
- Elmore, S., 2007. Apoptosis: A review of programmed cell death. *Toxicologic Pathology* 35, 495–516.
- Eng, J.K., McCormack, A.L., Yates, J.R., 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry* 5, 976–989.
- Fahlman, R.P., Chen, W., Overall, C.M., 2014. Absolute proteomic quantification of the activity state of proteases and proteolytic cleavages using proteolytic signature peptides and isobaric tags. *Journal of Proteomics* 100, 79–91.
- Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., Whitehouse, C.M., 1989. Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246, 64–71.
- Fortelny, N., Cox, J.H., Kappelhoff, R., *et al.*, 2014. Network analyses reveal pervasive functional regulation between proteases in the human protease web. *PLoS Biology* 12, e1001869.
- Freeman, H.J., Sleisenger, M.H., Kim, Y.S., 1983. Human protein digestion and absorption: Normal mechanisms and protein-energy malnutrition. *Clinical Gastroenterology* 12, 357–378.
- Gerner, C., Frohwein, U., Gotzmann, J., *et al.*, 2000. The Fas-induced apoptosis analyzed by high throughput proteome analysis. *Journal of Biological Chemistry* 275, 39018–39026.
- Gevaert, K., Goethals, M., Martens, L., *et al.*, 2003. Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. *Nature Biotechnology* 21, 566–569.
- Gevaert, K., Van Damme, J., Goethals, M., *et al.*, 2002. Chromatographic isolation of methionine-containing peptides for gel-free proteome analysis: Identification of more than 800 *Escherichia coli* proteins. *Molecular & Cellular Proteomics* 1, 896–903.
- Gorg, A., Weiss, W., Dunn, M.J., 2004. Current two-dimensional electrophoresis technology for proteomics. *Proteomics* 4, 3665–3685.
- Guo, L., Eisenman, J.R., Mahimkar, R.M., *et al.*, 2002. A proteomic approach for the identification of cell-surface proteins shed by metalloproteinases. *Molecular & Cellular Proteomics* 1, 30–36.
- Hanks, S.K., Quinn, A.M., Hunter, T., 1988. The protein kinase family: Conserved features and deduced phylogeny of the catalytic domains. *Science* 241, 42–52.
- von Heijne, G., 1990. The signal peptide. *Journal of Membrane Biology* 115, 195–201.
- Hoffman, M., Monroe, D.M., 2007. Coagulation 2006: A modern view of hemostasis. *Hematology/Oncology Clinics of North America* 21, 1–11.
- Holmberg, C., Ghesquiere, B., Impens, F., *et al.*, 2013. Mapping proteolytic processing in the secretome of gastric cancer-associated myofibroblasts reveals activation of MMP-1, MMP-2, and MMP-3. *Journal of Proteome Research* 12, 3413–3422.
- Jefferson, T., Causevic, M., auf dem Keller, U., *et al.*, 2011. Metalloprotease meprin beta generates nontoxic N-terminal amyloid precursor protein fragments *in vivo*. *Journal of Biological Chemistry* 286, 27741–27750.
- Karas, M., Bachmann, D., Hillenkamp, F., 1985. Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Analytical Chemistry* 57, 2935–2939.
- Keller, A., Nesvizhskii, A.I., Kolker, E., Aebersold, R., 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Analytical Chemistry* 74, 5383–5392.
- King, R.W., Deshaies, R.J., Peters, J.M., Kirschner, M.W., 1996. How proteolysis drives the cell cycle. *Science* 274, 1652–1659.
- Kleifeld, O., Doucet, A., auf dem Keller, U., *et al.*, 2010. Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. *Nature Biotechnology* 28, 281–288.
- Kleifeld, O., Doucet, A., Prudova, A., *et al.*, 2011. Identifying and quantifying proteolytic events and the natural N terminome by terminal amine isotopic labeling of substrates. *Nature Protocols* 6, 1578–1611.
- Lange, P.F., Huesgen, P.F., Nguyen, K., Overall, C.M., 2014. Annotating N termini for the human proteome project: N termini and N alpha-acetylation status differentiate stable cleaved protein species from degradation remnants in the human erythrocyte proteome. *Journal of Proteome Research* 13, 2028–2044.
- Lange, P.F., Huesgen, P.F., Overall, C.M., 2012. TopFIND 2.0-linking protein termini with proteolytic processing and modifications altering protein function. *Nucleic Acids Research* 40, D351–D361.
- Lange, P.F., Overall, C.M., 2011. TopFIND, a knowledgebase linking protein termini with function. *Nature Methods* 8, 703–704.
- Lange, P.F., Overall, C.M., 2013. Protein TAILS: When termini tell tales of proteolysis and function. *Current Opinion in Chemical Biology* 17, 73–82.
- Lewandrowski, U., Wortelkamp, S., Lohrig, K., *et al.*, 2009. Platelet membrane proteomics: A novel repository for functional research. *Blood* 114, e10–e19.
- Lopez-Otin, C., Overall, C.M., 2002. Protease degradomics: A new challenge for proteomics. *Nature Reviews Molecular Cell Biology* 3, 509–519.
- Mahrus, S., Trinidad, J.C., Barkan, D.T., *et al.*, 2008. Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. *Cell* 134, 866–876.
- Manning, G., Whyte, D.B., Martinez, R., Hunter, T., Sudarsanam, S., 2002. The protein kinase complement of the human genome. *Science* 298, 1912–1934.
- Marchant, D.J., Bellac, C.L., Moraes, T.J., *et al.*, 2014. A new transcriptional role for matrix metalloproteinase-12 in antiviral immunity. *Nature Medicine* 20, 493–502.
- Matthews, D.J., Wells, J.A., 1993. Substrate phase: Selection of protease substrates by monovalent phage display. *Science* 260, 1113–1117.
- McQuibban, G.A., Gong, J.H., Wong, J.P., *et al.*, 2002. Matrix metalloproteinase processing of monocyte chemoattractant proteins generates CC chemokine receptor antagonists with anti-inflammatory properties *in vivo*. *Blood* 100, 1160–1167.
- Nesvizhskii, A.I., Keller, A., Kolker, E., Aebersold, R., 2003. A statistical model for identifying proteins by tandem mass spectrometry. *Analytical Chemistry* 75, 4646–4658.
- Neurath, H., Walsh, K.A., 1976. Role of proteolytic enzymes in biological regulation (a review). *Proceedings of the National Academy of Sciences of the United States of America* 73, 3825–3832.
- Niessen, S., Hoover, H., Gale, A.J., 2011. Proteomic analysis of the coagulation reaction in plasma and whole blood using PROTOMAP. *Proteomics* 11, 2377–2388.
- Olsen, J.V., Macek, B., Lange, O., *et al.*, 2007. Higher-energy C-trap dissociation for peptide modification analysis. *Nature Methods* 4, 709–712.
- Overall, C.M., Blobel, C.P., 2007. In search of partners: Linking extracellular proteases to substrates. *Nature Reviews Molecular Cell Biology* 8, 245–257.
- Papadioti, A., de Bock, P.J., Vranakis, I., *et al.*, 2012. Study of the whole cell lysate of two *Coxiella burnetii* strains using N-terminomics. *Journal of Proteome Research* 11, 3150–3159.
- Pappin, D.J., Hojrup, P., Bleasby, A.J., 1993. Rapid identification of proteins by peptide-mass fingerprinting. *Current Biology* 3, 327–332.
- Perkins, D.N., Pappin, D.J., Creasy, D.M., Cottrell, J.S., 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551–3567.

- Pop, C., Salvesen, G.S., 2009. Human caspases: Activation, specificity, and regulation. *Journal of Biological Chemistry* 284, 21777–21781.
- Rawlings, N.D., Barrett, A.J., Bateman, A., 2012. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research* 40, D343–D350.
- Ross, P.L., Huang, Y.N., Marchese, J.N., *et al.*, 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Molecular & Cellular Proteomics* 3, 1154–1169.
- Schechter, I., Berger, A., 1967. On the size of the active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications* 27, 157–162.
- Schilling, O., Barre, O., Huesgen, P.F., Overall, C.M., 2010. Proteome-wide analysis of protein carboxy termini: C terminomics. *Nature Methods* 7, 508–511.
- Schilling, O., Huesgen, P.F., Barre, O., auf dem Keller, U., Overall, C.M., 2011. Characterization of the prime and non-prime active site specificities of proteases by proteome-derived peptide libraries and tandem mass spectrometry. *Nature Protocols* 6, 111–120.
- Schilling, O., Overall, C.M., 2008. Proteome-derived, database-searchable peptide libraries for identifying protease cleavage sites. *Nature Biotechnology* 26, 685–694.
- Seegers, W.H., Johnson, J.F., Fell, C., 1954. An antithrombin reaction to prothrombin activation. *American Journal of Physiology* 176, 97–103.
- Shen, C., Yu, Y., Li, H., *et al.*, 2012. Global profiling of proteolytically modified proteins in human metastatic hepatocellular carcinoma cell lines reveals CAPN2 centered network. *Proteomics* 12, 1917–1927.
- Skidgel, R.A., Erdos, E.G., 2006. Cellular carboxypeptidases. *Immunological Reviews* 161, 129–141.
- Starr, A.E., Bellac, C.L., Dufour, A., Goebeler, V., Overall, C.M., 2012. Biochemical characterization and N-terminomics analysis of leukolysin, the membrane-type 6 matrix metalloprotease (MMP25): Chemokine and vimentin cleavages enhance cell migration and macrophage phagocytic activities. *Journal of Biological Chemistry* 287, 13382–13395.
- Syka, J.E., Coon, J.J., Schroeder, M.J., Shabanowitz, J., Hunt, D.F., 2004. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* 101, 9528–9533.
- Taylor, A., 1993. Aminopeptidases: Structure and function. *FASEB Journal* 7, 290–298.
- Thiede, B., Kretschmer, A., Rudel, T., 2006. Quantitative proteome analysis of CD95 (Fas/Apo-1)-induced apoptosis by stable isotope labeling with amino acids in cell culture, 2-DE and MALDI-MS. *Proteomics* 6, 614–622.
- Thiede, B., Treumann, A., Kretschmer, A., Sohlke, J., Rudel, T., 2005. Shotgun proteome analysis of protein cleavage in apoptotic cells. *Proteomics* 5, 2123–2130.
- Thompson, A., Schafer, J., Kuhn, K., *et al.*, 2003. Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical Chemistry* 75, 1895–1904.
- Turk, B., Turk, D., Turk, V., 2012. Protease signalling: The cutting edge. *EMBO Journal* 31, 1630–1643.
- Unlu, M., Morgan, M.E., Minden, J.S., 1997. Difference gel electrophoresis: A single gel method for detecting changes in protein extracts. *Electrophoresis* 18, 2071–2077.
- Washburn, M.P., Wolters, D., Yates 3RD, J.R., 2001. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature Biotechnology* 19, 242–247.
- Wejda, M., Impens, F., Takahashi, N., *et al.*, 2012. Degradomics reveals that cleavage specificity profiles of caspase-2 and effector caspases are alike. *Journal of Biological Chemistry* 287, 33983–33995.
- Wells, J.M., McLuckey, S.A., 2005. Collision-induced dissociation (CID) of peptides and proteins. *Methods in Enzymology* 402, 148–185.
- Wiederanders, B., Kaulmann, G., Schilling, K., 2003. Functions of propeptide parts in cysteine proteases. *Current Protein & Peptide Science* 4, 309–326.

Relevant Website

<http://flintbox.com/public/project/1948>
Flintbox.

PROTEIN SYNTHESIS/DEGRADATION: PROTEIN DEGRADATION – INTRACELLULAR

Contents

Ubiquitin, Ubiquitin-Like Proteins, and Proteasome-Mediated Degradation
Endoplasmic Reticulum-Associated Degradation and Protein Quality Control
Role of Lysosomes in Intracellular Degradation

Ubiquitin, Ubiquitin-Like Proteins, and Proteasome-Mediated Degradation

RJ Dohmen, University of Cologne, Köln, Germany

JM Huibregtse, The University of Texas at Austin, Austin, TX, USA

M Scheffner, University of Konstanz, Konstanz, Germany

© 2016 Elsevier Inc. All rights reserved.

General Background

Intracellular protein degradation is a highly selective and regulated process as indicated by the fact that different proteins have different half-lives. Moreover, the half-life of a protein can vary in response to internal or external cues allowing, for example, the correct execution of spatio-temporally regulated processes (e.g., cell cycle, cell differentiation, and signal transduction pathways). Besides its manifold functions in controlling/regulating particular pathways or processes, protein degradation also plays a pivotal role in more general processes such as protein quality control (removal of 'abnormal', e.g., misfolded or misassembled proteins), amino acid homeostasis (e.g., under starvation conditions), and immune surveillance (generation of peptides presented at the cellular surface by major histocompatibility complexes). To ensure that proteins are not randomly destroyed – which would be deleterious for a cell – but are degraded only when appropriate, eukaryotes have evolved two complex systems for selective degradation of intracellular proteins, the autophagic-lysosomal system and the ubiquitin-proteasome system (UPS).

Ubiquitin and Ubiquitin-like Proteins

Introductory Comments

A critical issue in selective protein degradation is how a protease discriminates between proteins that should be degraded and those that should not be targeted. In principle, this can be achieved by two, non-mutually exclusive mechanisms: proteins contain a defined amino acid sequence and structure that serve as a recognition and degradation signal, or proteins are flagged for degradation by a distinct posttranslational

modification. The first mechanism is mainly followed by site-specific proteases and results in protein processing rather than complete degradation; the substrate spectrum of such proteases is limited. In contrast, the second mechanism allows the specific recognition and degradation of many different proteins by a single protease, and this mechanism has been adopted by the UPS: proteins are first modified by covalent attachment of ubiquitin ('ubiquitylation' or 'ubiquitination'), which serves as recognition signal for a multi-subunit protease complex, the 26S proteasome.

Ubiquitin

Ubiquitin was discovered in 1975 (Goldstein *et al.*, 1975) and two years later, it was shown that a certain fraction of histone 2A (H2A) is covalently modified by a single moiety of ubiquitin via isopeptide bond formation between the carboxyl group of the C-terminal glycine of ubiquitin and the ϵ -amino group of a defined lysine residue of H2A (Goldknopf and Busch, 1977; Hunt and Dayhoff, 1977). This was the first description of protein–protein modification but it took almost 30 years to uncover the function(s) of histone ubiquitylation (Weake and Workman, 2008). In the late 1970s/early 1980s, biochemical characterization of a eukaryotic non-lysosomal ATP-dependent proteolytic system revealed that its substrates are ubiquitylated prior to recognition and degradation by a protease complex, which later on was identified as the 26S proteasome (Etlinger and Goldberg, 1977; Ciechanover *et al.*, 1978, 1980; Wilkinson *et al.*, 1980; Hershko and Ciechanover, 1998). In addition, evidence was provided that proteolytic substrates are modified by a ubiquitin chain or ubiquitin polymer consisting of several ubiquitin moieties connected by isopeptide bonds rather than by a single moiety of ubiquitin (Hershko *et al.*, 1980; Chau *et al.*, 1989).

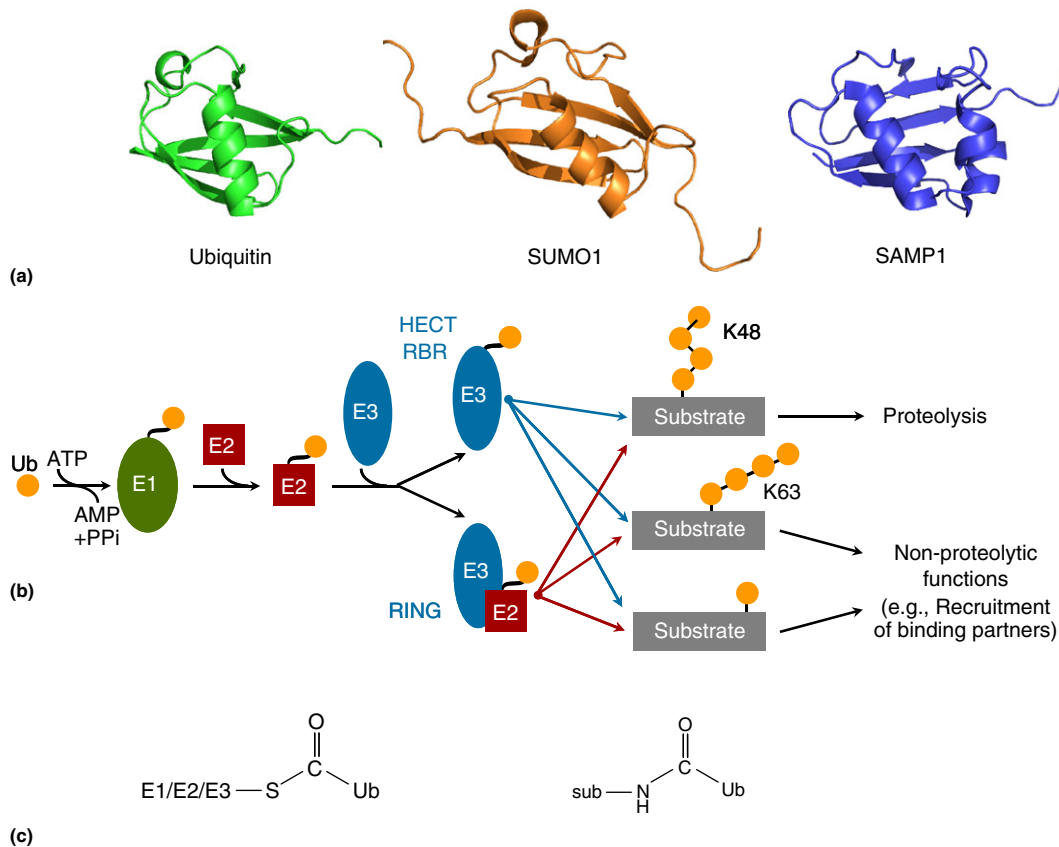


Figure 1 Conjugation of ubiquitin and ubiquitin-like proteins. (a) Structures of ubiquitin (PDB 1UBQ), SUMO (PDB 1A5R), and SAMP (PDB 3P00). (b) The ubiquitin conjugation cascade. Ubiquitin is activated by an E1 activating enzyme by forming a thioester bond between the C-terminal glycine of ubiquitin and the catalytic cysteine residue of E1 at the expense of ATP. Activated ubiquitin is then transferred to an E2 conjugating enzyme in a transesterification reaction. Finally, E2s in concert with E3 ubiquitin-protein ligases, which mediate substrate recognition, catalyze the covalent attachment of ubiquitin to a substrate protein via isopeptide bond formation. Based on their functional properties, E3s can be classified into three groups. In case of HECT E3s and RBR E3s, activated ubiquitin is transferred from the E2 to the E3 in a transesterification reaction and isopeptide bond formation is catalyzed by the E3 itself, while RING E3s act as adaptors between substrates and E2s, and at least in some cases act as allosteric activators of E2s. Depending on the type of ubiquitylation, proteins are targeted for degradation by the 26S proteasome (e.g., modification by K48-linked ubiquitin chains) or for non-proteolytic fates (e.g., monoubiquitylation, modification by K63-linked ubiquitin chains). Finally, it should be noted that ubiquitylation is a reversible process by the action of deubiquitylating enzymes (DUBs) (not shown) and that modification of proteins by ubiquitin-like proteins is catalyzed by enzymatic cascades similar to the ubiquitin-conjugation system. For further details, see main text. (c) Ubiquitylation involves two types of covalent protein-protein interactions: thioester complexes between the carboxyl group of the C-terminal glycine of ubiquitin and the catalytic cysteine residue of E1, E2s, and HECT and RBR E3s; (iso)peptide bonds between the carboxyl group of the C-terminal glycine of ubiquitin and the ϵ -amino group of lysine residues or the α -amino group of the N-terminal residue of a substrate.

Ubiquitin is a 76 amino acid polypeptide that is highly conserved among eukaryotes and adopts a compact globular fold, the β -grasp or ubiquitin-like fold (Vijay-Kumar *et al.*, 1987; Burroughs *et al.*, 2012; Figure 1). *Saccharomyces cerevisiae* and humans contain four genes encoding ‘pro-ubiquitin’ – head-to-tail-fusions of several ubiquitin moieties or ubiquitin fused to ribosomal subunits (Wiborg *et al.*, 1985; Ozkaynak *et al.*, 1987; Baker and Board, 1987). Free ubiquitin is generated from such fusion proteins by deubiquitylating enzymes (DUBs) (Komander *et al.*, 2009; Reyes-Turcu *et al.*, 2009). Covalent attachment of ubiquitin to target proteins is mediated by an elaborate enzymatic conjugation cascade (see below); in most cases, lysine residues serve as acceptors for ubiquitin but in some cases, ubiquitin can also be attached to the α -amino group of the N-terminal residue of a protein or to

serine or threonine residues via the formation of a peptide bond or ester bond, respectively. In addition to ‘mono-ubiquitylation’ (attachment of a single ubiquitin to a lysine residue) and multiple mono-ubiquitylation (several lysine residues of a substrate are modified by single ubiquitin moieties), ubiquitin can serve as its own substrate resulting in the formation of ubiquitin chains on target proteins (‘poly-ubiquitylation’). Ubiquitin contains seven lysine residues (K6, 11, 27, 29, 33, 48, 63) and each of these as well as its N-terminal residue can serve as an attachment site for ubiquitin resulting in 8 homotypic (the same lysine is used for ubiquitin attachment throughout the chain), numerous heterotypic chains (different lysines used within one chain), and branched chains (two or more lysines of one ubiquitin are ubiquitylated) (Komander and Rape, 2012).

The diversity of ubiquitylation or 'ubiquitin signals' indicates that different types of these fulfill different functions. Indeed, different ubiquitin chains adopt different topologies, and K48-linked and K11-linked ubiquitin chains mainly represent signals for proteasomal degradation, while K63-linked chains fulfill non-proteasomal functions. Similarly, mono-ubiquitylation is mainly a signal for non-proteasomal fates (e.g., endocytosis) but in some cases may also target proteins for proteasomal degradation (Komander and Rape, 2012; Kravtsova-Ivantsiv and Ciechanover, 2012; Kravtsova-Ivantsiv *et al.*, 2013). In a simplistic view, attachment of ubiquitin or ubiquitin chains alters the biochemical properties of a modified protein. This can be achieved by inducing a conformational change (e.g., allosteric activation or inhibition of enzymatic activities) or by creating or shielding a protein-protein interaction surface. In many cases, decoding the different ubiquitin signals involves proteins with ubiquitin-binding domains that recognize the ubiquitin-like fold or in some cases the specific ubiquitin chain type (e.g., K48-linked or linear chains). Different ubiquitin-binding domains can be grouped into distinct families based on structural similarities (Dikic *et al.*, 2009; Husnjak and Dikic, 2012); however, it remains to be determined if structural similarities are correlated strictly to functional similarities.

The ubiquitin-like fold consists of a four-stranded anti-parallel β -sheet with an α -helix on top of the β -sheet (Figure 1). This fold is not only found in eukaryotic proteins but also in prokaryotic ones including ThiS and MoaD that are not covalently attached to other proteins but serve as sulfur carriers in biosynthetic pathways (see below) (Hochstrasser, 2009; Burroughs *et al.*, 2012). In this context, it should be noted that the concept of earmarking proteins for degradation by protein-protein modification is also used by some eubacteria but the respective modifying protein (e.g., prokaryotic ubiquitin-like protein, Pup) does not show any structural similarity to

ubiquitin (see below, Pearce *et al.*, 2008). The small archaeal modifier proteins (SAMP) found in some archaea, in contrast, share the β -grasp fold with ubiquitin (e.g. Humbard and Maupin-Furlow, 2013). There are nearly 20 human proteins that share primary sequence and structural similarity with ubiquitin and are covalently conjugated to proteins or other molecules (note that ubiquitin-like domains are also found as part of eukaryotic proteins that cannot be covalently attached to other proteins (Hochstrasser, 2000)). Some of the more thoroughly characterized ubiquitin-like proteins (UBLs), described further below, include small ubiquitin-like modifier (SUMO) family members, neural precursor cell expressed developmentally downregulated-8 (Nedd8), IFN-stimulated gene, 15 kDa (ISG15), and two UBLs involved in autophagy, Atg8 and Atg15 (Table 1). A discussion of these UBLs highlights some commonalities and differences in the organization of their conjugation and deconjugation systems, as well as the remarkable functional divergence of these protein modifiers. The functions of other UBLs are just beginning to be understood, and in some cases are revealing the evolutionary origins of all eukaryotic UBLs (Hochstrasser, 2009; Burroughs *et al.*, 2012).

Ubiquitin-Conjugation System

Ubiquitin is conjugated to substrate proteins by the concerted action of three classes of enzymes: In the first step, ubiquitin is put into a chemically reactive state by a ubiquitin-activating enzyme (E1, UBA, UBE1) at the expense of ATP. The activated ubiquitin is then transferred from E1 to a ubiquitin-conjugating enzyme (E2, UBC, UBE2). Finally, the E2 in conjunction with a ubiquitin-protein ligase (E3, UBE3), which mediates substrate recognition, transfers ubiquitin to a substrate protein by isopeptide bond formation. This general scheme of conjugation (Figure 1) as well as the underlying biochemical mechanisms have been elucidated in the 1980s

Table 1 The enzymes (E1s, E2s, E3s, and Deconjugation Enzymes), targets, and functions for some of the most thoroughly characterized ubiquitin-like proteins (UBLs)

UBL	E1	E2	E3(s)	Deconjugating enzyme(s)	Targets	Function(s)
SUMO1-4	SAE1/UBA2 heterodimer	UBE2I (UBC9)	SIZ/PIAS proteins, RanBP2	SENP family proteases	Numerous, generally containing ψ KxD/E motif	Numerous, including localization, and recognition signal for STUBs
NEDD8	APPBP1/UBA3 heterodimer	UBE2M (UBC12)	RBX1, RBX2 DENN1	CSN complex (CSN5 catalytic subunit)	Cullin proteins	Activation of Cullin-containing ubiquitin ligases
ISG15	UBA7 (UBE1L)	UBE2L6 (UBCH8)	HERC5	USP 18 (and several virus-encoded DUBs)	Newly synthesized proteins	Innate immunity
LC3 proteins (ATG8)	ATG7	ATG3	ATG12-ATG5-ATG16 complex	ATG4	Phosphatidylethanolamine	Autophagy; phagophore membrane formation, cargo recruitment
ATG12	ATG7	ATG10	N/A	N/A	ATG5	Component of the ATG12-ATG5-ATG16 ligase for LC3 proteins

(Ciechanover *et al.*, 1982; Haas *et al.*, 1982; Hershko *et al.*, 1983; Hershko and Ciechanover, 1998) and apply also to the conjugation of UBLs (see below). The hierarchical mode of ubiquitin conjugation is reflected in the number of the respective enzymes, with one (*S. cerevisiae*) or two (human) E1 enzymes, 11 (*S. cerevisiae*) or approximately 30 (human) E2s, and several hundred E3s.

The ubiquitin E1 enzyme is a monomer with three functional entities: an adenylation domain, a domain containing the catalytically active cysteine residue, and a ubiquitin fold domain (UFD). The adenylation domain contains an ATP-binding site and catalyzes the formation of a ubiquitin (acyl)-adenylate. The acyl-adenylate is then nucleophilically attacked by the thiol group of the active site cysteine residue resulting in the formation of a thioester complex between E1 and ubiquitin. The UFD as well as part of the catalytic cysteine-containing domain provide binding sites for E2 enzymes (Hanzelmann *et al.*, 2012). The essential role of the ubiquitin-conjugation system *in vivo* and its involvement in the cell division cycle and protein quality control became clear through the identification and characterization of a mouse cell line (ts85) bearing a temperature-sensitive E1 (Finley *et al.*, 1984; Ciechanover *et al.*, 1984).

E2 enzymes are characterized by the presence of the catalytic ubiquitin-conjugating (UBC) domain of approx. 150 amino acids in length. Some E2s consist just of this domain indicating that the UBC domain unites the four main properties of E2s: association with E1 as well as with E3 enzymes, formation of ubiquitin-thioester complexes, and covalent attachment of ubiquitin to target proteins by isopeptide bond formation (Hanzelmann *et al.*, 2012). In a simplified model, upon binding to the E1 UFD, the catalytic site cysteine of the E2 becomes located in close proximity to the catalytic cysteine of E1. This results in transfer of ubiquitin from E1 to the catalytic cysteine residue of E2 in a trans-thioesterification reaction followed by release of the ubiquitin-loaded E2. This change in affinity ('high' affinity of free E2 for ubiquitin-loaded E1, 'low' affinity of ubiquitin-loaded E2 for E1) is required to allow the ubiquitin-loaded E2 to interact with its cognate E3(s), since the binding surfaces of E2 for E1 and E3s overlap (and, thus, binding of the two is mutually exclusive). From the E2, ubiquitin is then transferred either to a target protein, which in most cases occurs in conjunction with an E3, or to an E3 in a trans-thioesterification reaction (Figure 1).

Some E2s have, in addition to the UBC domain, N-terminal and C-terminal extensions. The diversity in composition/structure contributes to the functional specialization of E2s. In general, a given E2 interacts only with a distinct subset of E3s (e.g., UBC6 and UBC7 interact with E3s involved in endoplasmic reticulum-associated degradation (ERAD), UbcH10 is one of the cognate E2s of the anaphase promoting complex (APC)/cyclosome (Smith *et al.*, 2011; Meyer and Rape, 2011)) and may be found at distinct intracellular localizations (e.g., UBC6 is an integral ER membrane protein). Furthermore, besides their involvement in the control of different cellular processes, different E2s may catalyze the formation of different types of ubiquitin chains (e.g., the Mms2-Ubc13 complex forms K63-linked ubiquitin chains, UbcH10 and Ubc6 form K11-linked chains (Hanzelmann *et al.*, 2012; Meyer and Rape, 2011; Xu *et al.*, 2009)).

For the integrity of a cell, ubiquitylation has to be a highly specific and, at least in some cases, regulated process. The need for specificity is reflected by the number of E3s: the human genome encodes several hundred putative E3s, with each E3 presumably having a distinct and limited substrate spectrum. Based on functional considerations, E3s can be categorized into three groups: HECT (homologous to E6AP C terminus) E3s, RING (really interesting new gene) and RING-like (U box) E3s, and RBR (Ring-in-between-RING) E3s (Berndsen and Wolberger, 2014). HECT E3s were the first family of E3s described (Huibregtse *et al.*, 1995). They are characterized by the presence of a C-terminal HECT domain of approximately 350 amino acids in length. The HECT domain mediates the interaction with cognate E2 enzymes, contains the catalytic cysteine residue, which accepts ubiquitin from the E2 in a trans-thioesterification reaction, and catalyzes the covalent attachment of ubiquitin to substrate proteins. The human genome encodes 28 HECT proteins ranging in size from approximately 80–600 kDa. HECT proteins can be categorized into three subfamilies based on the presence of known protein–protein interaction motifs in their respective N-terminal extensions, which presumably define the substrate spectrum of HECT E3s. Furthermore, some HECT E3s preferentially build up K63-linked ubiquitin chains (e.g., yeast Rsp5), while others form K48-linked ubiquitin chains (e.g., E6AP) thereby targeting proteins for proteasomal degradation (Scheffner and Kumar, 2014).

RING E3s are the largest known E3 family with the human genome encoding approximately 500 RING domain-containing proteins (Metzger *et al.*, 2014). The RING domain consists of approximately 40–60 amino acids and comprises 6 or 7 cysteine residues and 2 or 1 histidine residues that coordinate 2 Zn^{2+} ions in a so-called cross-braced manner. The RING domain mediates the interaction with cognate E2s and at least in some cases acts as an allosteric activator of E2s (i.e., in contrast to HECT E3s, ubiquitin is transferred directly from the E2 to the substrate; Figure 1). Based on their composition, RING E3s can be roughly grouped into E3s in which the substrate recognition domain and the RING domain are located on the same polypeptide (note that such RING E3s frequently function in the form of homo-dimers or hetero-dimers; e.g., Mdm2, Mdm2-MdmX, BRCA1-BARD1), and multi-subunit E3s, in which substrate recognition is mediated by a protein other than the RING domain protein. Examples for the latter are the Cullin-RING ligases (CRLs), including SCF complexes consisting of CUL1, RBX1 (RING protein), SKP1, and an exchangeable F-box protein, which mediates substrate recognition, and CBC complexes consisting of CUL2 or CUL5, RBX1 or RBX2, Elongins B and C, and as substrate recognition components VHL-box or SOCS-box proteins (Zimmerman *et al.*, 2010; Lydeard *et al.*, 2013; Deshaies and Joazeiro, 2009). Another prominent example is the APC/cyclosome, which consists of at least 13 subunits with APC11 being a RING domain protein and CDH1 and CDC20 being substrate recognition components (Bassermann *et al.*, 2014). Since the human genome encodes approximately 90 F-box substrate adaptor proteins (the F-box represents the interaction site for SKP1, while substrate recognition is mediated by other regions that differ between different F-box proteins), the combinatorial nature of multi-subunit E3s enlarges the substrate spectrum of the ubiquitin-conjugation system.

The catalytic domain of RBR E3s consists of two RING domains (RING1, RING2) that are linked by an in-between RING domain. RING1 serves as interaction site for ubiquitin-loaded E2s, while RING2 contains a catalytic cysteine residue that accepts ubiquitin from E2 in a trans-thioesterification reaction and presumably catalyzes the covalent attachment of ubiquitin to substrate proteins (Wenzel and Klevit, 2012; Smit and Sixma, 2014). Thus, RBR E3s follow a RING/HECT hybrid-like mechanism of ubiquitylation. Prominent members of this E3 family (the human genome encodes 14 RBR E3s) are PARKIN and the HOIL–HOIP complex, which catalyzes the formation of N-terminally linked ('linear') ubiquitin chains on substrate proteins.

Although it remains to be determined if all HECT/RING/RBR domain-containing proteins indeed function as E3s, the sheer number of putative E3s indicates that the ubiquitin-conjugation system has the potential to target any protein in a cell in a highly specific manner (note that most, if not all, intracellular proteins are ubiquitylated at some stage during their lifetime; e.g., Kim *et al.*, 2011). Indeed, substrates of the ubiquitin-conjugation system are not characterized by a unifying amino acid sequence or region which serves as common recognition signal. Rather each E3 recognizes a distinct subset of sequence motifs that depends on the protein–protein interaction motifs the E3 contains (e.g., substrates of the APC/cyclosome often contain a so-called KEN box, while substrates of the NEDD4 subfamily of HECT E3s frequently display PPxY motifs (Barford, 2011; Scheffner and Kumar, 2014)). Furthermore, some proteins are targeted for degradation only at certain times during the cell cycle, development or differentiation or in response to distinct extracellular stimuli. This temporal regulation can be achieved at the level of E3 or substrate. For example, cofactors are in some cases required to activate a given E3 (e.g., Smad7-mediated activation of Smurf2 (Scheffner and Kumar, 2014)) or substrates of SCF complexes need to be phosphorylated before they are recognized by their cognate F-box proteins (Deshaies and Joazeiro, 2009). In fact, in many cases, there is a crosstalk between components of the ubiquitin-conjugation system and other posttranslational modification systems that ensures appropriate ubiquitylation and degradation of a given protein (see also below, sections 'SUMO' and 'NEDD8'). Thus, the actual mechanisms involved in regulating ubiquitylation of a given protein need to be determined for each individual substrate-E3 pair.

DUBs

Ubiquitylation is a dynamic process inasmuch as it can be reversed by the action of DUBs. The human genome encodes approximately 90 different DUBs that can be grouped into 5 families: ubiquitin C-terminal hydrolases (UCHs), ubiquitin-specific proteases (USPs), ovarian tumor domain (OTU) proteases, Josephin domain proteases (MJD), and Jab1/MPN/Mov34 (JAMM) metalloproteases (Komander *et al.*, 2009; Reyes-Turcu *et al.*, 2009; Clague *et al.*, 2012; Eletre and Wilkinson, 2014). While the first four are members of the cysteine protease superfamily, the latter are Zn-dependent enzymes. Besides their role as antagonists of E2/E3 enzymes (i.e. by removing ubiquitin from substrate proteins), DUBs are

required for processing pro-ubiquitin molecules (see section 'Ubiquitin') into mature ubiquitin monomers and for regenerating mono-ubiquitin from ubiquitin chains liberated from polyubiquitylated proteins by other DUBs, for example at the proteasome, or from free ubiquitin chains that may serve as reservoir for mono-ubiquitin in cells.

The balance between ubiquitylation and deubiquitylation eventually decides whether a protein is degraded or not, indicating that the activities of DUBs must be tightly controlled. Similar to other proteases, DUBs show little activity unless they are associated with cognate substrates or scaffolding proteins, presumably ensuring that 'non-substrates' are not coincidentally cleaved (Liz and Sousa, 2005). Furthermore, DUBs in general do not just consist of the catalytic domain but have additional domains such as ubiquitin binding domains and protein-protein interaction motifs. While the former are involved in determining the type of ubiquitin chain a DUB may recognize (Komander *et al.*, 2009; Reyes-Turcu *et al.*, 2009; Clague *et al.*, 2012; Eletre and Wilkinson, 2014), the latter may define the subcellular localization of a DUB or may affect its activity. For example, three DUBs have been associated with the 26S proteasome (POH1, UCH37, USP14) with POH1 (the human ortholog of yeast Rpn11) and UCH37 being integral components of the 19S regulatory particle (19S RP; see below). Furthermore, the yeast ortholog of USP14, Ubp6, interacts with the E3 ligase Hul5. This is thought to fine-tune the length of ubiquitin chains attached to substrates of Hul5 (Crosas *et al.*, 2006). Finally, DUB activity is subjected to regulation at the level of expression and at the posttranslational level by phosphorylation/dephosphorylation, ubiquitylation, and sumoylation (Komander *et al.*, 2009; Reyes-Turcu *et al.*, 2009; Clague *et al.*, 2012; Eletre and Wilkinson, 2014).

SUMO Family Members

SUMO is found in all eukaryotes, with four related SUMO proteins in most vertebrates (SUMO1-4). SUMO2 and 3 are nearly identical and about 50% identical to SUMO1 (Gareau and Lima, 2010). SUMO4 is most similar to SUMO2, although it remains unclear whether it is actually conjugated to other proteins. SUMO2/3 and yeast SUMO can form poly-SUMO chains, while SUMO1 does not. A single heterodimeric E1 activating enzyme (AOS1/UBA2) and a single E2 protein (Ubc9) mediate all SUMO conjugation. Unusual for UBLs, most SUMOylated proteins have a consensus modification site, consisting of ψ -K-x-D/E (where ψ is a bulky hydrophobic residue, x is any amino acid, and K is the actual modification site), and Ubc9 can generally directly recognize this modification site, at least *in vitro*. Variations on the consensus motif exist, as well, including phosphorylation-dependent recognition sites. E3s for SUMO are few in number, but include RanBP2 and the SIZ/PIAS family of proteins. The problem of how specific substrate recognition is established, and the mechanisms of SUMO E3s, remains an active area of investigation. There are several SUMO isopeptidases, including six human sentrin-specific protease (SENP) proteins, which have varying specificities for SUMO paralogs and types of SUMO linkages. As with ubiquitin, the balance between SUMOylation

and deSUMOylation is likely a key to the regulation of the biological functions of SUMO.

SUMO interacting motifs (SIMs) play a central part in mediating the biologic functions of SUMOylation. Particularly interesting examples of SIM-containing proteins are the STUBs (SUMO-targeted ubiquitin ligases), which are E3 ubiquitin ligases that ubiquitylate SUMOylated proteins (Perry *et al.*, 2008; Geoffroy and Hay, 2009; Sriramachandran and Dohmen, 2014). In this setting, SUMOylation serves as a posttranslational modification that triggers ubiquitylation through binding of the SUMOylated protein to the SIM of the STUB. SIMs also exist in some SUMO E3s (e.g., RanBP2 and Siz/PIAS proteins), aiding in the recruitment and orientation of Ubc9~SUMO complexes.

The substrates of SUMOylation are numerous and diverse (Xu and Peng, 2006). Although substrate proteins are found in many locales, proteins associated with the nucleus are particularly abundant. A wide range of transcriptional regulators, chromatin modifiers, nucleolar substrates, and centromeric and telomeric substrates have been identified (Jentsch and Psakhye, 2013; Cubenas-Potts and Matunis, 2013; Nagai *et al.*, 2011). The range of effects of SUMOylation can include changes in protein stability, protein localization, assembly of multi-protein structures, and alteration of enzyme activity. In short, it is not possible to define a single biochemical function for SUMOylation. This is certainly due, in part, to the nature of the conjugates (the specific SUMO paralog, whether a protein is poly-SUMOylated) and the interplay of SUMOylation with other types of protein modifications (e.g., ubiquitylation and protein acetylation).

NEDD8

NEDD8 is a key regulator of Cullin-Ring ubiquitin ligases (CRLs) (Sarikas *et al.*, 2011). CRLs play critical roles in a diverse array of essential cellular processes, including cell cycle regulation and signaling pathways. The E1 enzyme for NEDD8 is NAE, and the major E2 enzyme is Ubc12 (Ube2M). Like the SUMO E1, NAE is a heterodimer of proteins (APP-BP1/Uba3) that, together, structurally and functionally resemble the monomeric E1 enzyme for ubiquitin. Cullin proteins form the 'scaffold' around which the multi-subunit CRLs are built, with a RING domain protein associated at one end of the rigid cullin, and adaptor proteins and substrate binding proteins associated at the other. A RING domain protein, RBX1 or RBX2, functions to recruit ubiquitin E2 proteins; however, it binds to unmodified cullins in a conformation that restrains the RING domain from interacting with E2 proteins. The NEDDylation of the cullin at a single conserved lysine residue near the RBX binding site alters the conformation of RBX1/2, freeing the RING domain to interact productively with ubiquitin E2 proteins (Lydeard *et al.*, 2013).

Cullin proteins (seven in humans) constitute the predominant targets of NEDDylation. The de-NEDDylation of Cullins is tightly controlled by the COP9 signalosome complex, which contains a metalloproteinase subunit, COP9 signalosome subunit 5 (CSN5), with de-NEDDylase activity (Deshaies *et al.*, 2010). While there is still much to be sorted out with respect to regulation, and not all cullins may be

subject to the same regulation, it appears that de-NEDDylated cullins preferentially associate with CAND1; and CAND1 drives dissociation of what is normally a very stable association of the Cullin with its substrate adaptors. Therefore, while NEDDylation activates the ligase function of CRLs, de-NEDDylation constitutes part of a cycle that allows a limited number of Cullin proteins to dynamically associate with a wide array of substrate recruitment proteins, accounting for the tremendous variety of CRL target proteins. Consistent with the role of CRLs in regulating many cancer-related signaling pathways, a specific NAE inhibitor (MLN4924) is currently in clinical trials as a cancer therapeutic (Nawrocki *et al.*, 2012).

ISG15

ISG15 is a vertebrate-specific UBL that consists of two ubiquitin-like domains connected by a short flexible linker. The most striking feature of ISG15 is that it is strongly induced at the transcriptional level by type 1 interferon (IFN- α/β) signaling, implying a role for ISG15 conjugation in the innate immune response. The E1, E2, and predominant E3 enzymes for ISG15 are Ube11 (UBA7), UbcH8 (UBE2L6), and Herc5, respectively, and like ISG15, the genes encoding these enzymes are induced by IFN- α/β signaling (Durfee and Huibregtse, 2012). Mouse Ubp43 and its human ortholog Usp18 are interferon-induced deconjugation enzymes for ISG15 (Malakhov *et al.*, 2002). Several hundred cellular proteins have been identified as ISG15 targets in IFN-stimulated cells, although no clear pattern with respect to function or intracellular localization emerged from these analyses. A remarkable aspect of this conjugation pathway is that Herc5 mediates conjugation to the vast majority of these proteins. The basis of this broad recognition is that Herc5 is stably associated with the 60S ribosome, allowing Herc5 to conjugate ISG15 to nearly any protein that is being translated in the cell. Thus, it was proposed that ISGylation of newly synthesized proteins is essentially a stochastic process, and that, in the context of a virus-induced interferon response, newly synthesized viral proteins (rather than cellular proteins) may be the biologically relevant targets of ISG15 (Durfee *et al.*, 2010). ISG15 does not appear to target proteins for degradation or form poly-ISG15 chains. The basis of its antiviral activity may simply be that it serves to disrupt the function of viral proteins to which it is conjugated. The range of viruses that have been shown to be sensitive to the anti-viral effects of ISG15 include Influenza viruses, Herpesviruses, HIV, and Chikungunyavirus (Morales and Lenschow, 2013). Interestingly, several viruses have evolved mechanisms to either prevent ISG15 conjugation or reverse ISGylation via virus-encoded DUB-like proteins (Zhao *et al.*, 2013).

Early studies reported that ISG15 was a secreted cytokine that induced IFN- γ production (D'Cunha *et al.*, 1996). This surprising result is consistent with the recent identification of human patients harboring homozygous loss-of-function mutations in the ISG15 gene (Bogunovic *et al.*, 2012). These patients developed significant mycobacterial disease after receiving a live attenuated mycobacterial vaccine. IFN- γ signaling is critical for the response to mycobacterial infection, and further experiments confirmed that extracellular ISG15 induces IFN- γ production from peripheral blood mononuclear cells

(PBMcs). Thus, ISG15 appears to have two very different functions – one as an intracellular posttranslational modifier, and one as an extracellular cytokine-like molecule. The mode of secretion and cytokine-like signaling remain unknown.

ATG8/LC3 and ATG12: Key Players in Autophagy

Autophagy is a mechanism that eukaryotic cells use for bulk degradation of cytoplasmic organelles and other constituents (Hurley and Schulman, 2014). It is characterized by the phagophore, a double-membrane structure that engulfs cytoplasmic material (the autophagosome), which then fuses with the lysosome, delivering its contents for degradation. This process is triggered by nutrient starvation and other stresses and provides a source of building blocks for biosynthesis and energy production. Among the many proteins that participate in coordinating and regulating this cellular reorganization are two UBLs, ATG8/LC3 and ATG12 (Klionsky and Schulman, 2014). ATG8 (yeast) and LC3 (mammalian cells) are unusual in that they are not conjugated to a protein, but rather to the amine group of a lipid, phosphatidylethanolamine (PE). This complex then becomes part of the phagophore membrane and plays a role in regulating the expansion of the membrane. There are six LC3-related proteins in human cells that may have distinct functions in regulating the phagophore membrane. The β -grasp domain of the conjugate also interacts with and recruits other proteins, including cargo, to the autophagosome (Slobodkin and Elazar, 2013).

The second UBL in the autophagy pathway, ATG12, actually forms part of the ligase that is involved in conjugating ATG8/LC3 to PE (Hanada *et al.*, 2007). ATG12 is conjugated to another autophagy protein, ATG5, and the ATG12-ATG5 complex binds to ATG16. The ATG12-ATG5-ATG16 complex then acts as an E3 ligase for ATG8, covalently attaching it to PE. Other ATG proteins are involved in upstream steps in this process. ATG7 is the E1-like enzyme for both ATG8/LC3 and ATG12. ATG7 then transfers ATG8/LC3 to an E2-like protein, ATG3, and ATG12 to another E2-like protein, ATG10. From ATG10, ATG12 is passed to ATG5. ATG4 is a protease that processes ATG8 at its C terminus prior to its conjugation and can liberate the ATG8/LC3 from its PE conjugates. This amazing system illustrates the diversity of uses of UBLs. Much remains to be discovered regarding the function and regulation of ATG8 conjugation and deconjugation to phagophore lipids and the biophysical effects of this on membrane dynamics (Klionsky and Schulman, 2014).

Other UBLs

While the examples above illustrate the diversity of functions of UBLs, there are several others that promise to reveal other interesting aspects of cellular regulation. FAT10 is a vertebrate-specific UBL that, like ISG15, is a translational dimer of ubiquitin-like domains (Schmidtke *et al.*, 2014). FAT10 expression is induced by IFN- γ and TNF- α , primarily in tissues of the immune system. While its precise biologic role is unclear, like ubiquitin, it appears to target proteins for proteasomal recognition. URM1 is a UBL that is unusual in that it is involved in sulfur transfer reactions for thionucleoside

(e.g., 2-thiouridine) modification of tRNAs (Shigi, 2014). A URM1 thiocarboxylate is an intermediate in this reaction. This certainly reflects the early evolutionary origins of UBL systems, as sulfur transfer reactions in molybdopterin and thiamin biosynthesis in bacteria employ β -grasp proteins with terminal di-Glycine signatures (MoaD and ThiS) to transfer thiol groups to intermediates in their respective biosynthetic pathways (Hochstrasser, 2009; Burroughs *et al.*, 2012). Further, these bacterial UBLs are activated by E1-like enzymes, MoeB and ThiF, that are clearly structurally and mechanistically related to the E1 proteins of all UBLs.

The 26S Proteasome

Introductory Comments

The 26S proteasome is a macromolecular protease and central player in the UPS of eukaryotic cells. It mediates the degradation of proteins tagged with certain types of ubiquitin chains (Finley, 2009; Tomko and Hochstrasser, 2013). Some proteins, however, are also degraded in a ubiquitin-independent manner (Erales and Coffino, 2014). More simple evolutionary relatives of eukaryotic proteasomes are found in various archaeons and in one group of eubacteria (Actinomycetales) (Humbard and Maupin-Furrow, 2013). These prokaryotic proteasomes share the principle architecture with the catalytic core particle (CP) of eukaryotic proteasomes, which is also known as the 20S proteasome (Baumeister *et al.*, 1998). The 26S proteasome is composed of a 20S CP and one or two 19S RPs, which control the access of substrates to the CP (Finley, 2009; Tomko and Hochstrasser, 2013). 19S RPs recognize ubiquitin tags as well as unstructured domains of substrates and mediates their unfolding and translocation into the CP, with simultaneous deubiquitylation (Finley, 2009; Bhattacharyya *et al.*, 2014). The proteasome is a dynamic structure that undergoes conformational changes during substrate processing and that interacts with many other cellular proteins (Verma *et al.*, 2000; Matyskiela *et al.*, 2013; Marques *et al.*, 2009). Principle features of the proteasome and its subcomplexes are described below.

The 20S Core Particle: Structure and Catalytic Mechanism

The eukaryotic 20S CP (name refers to the sedimentation coefficient of the complex) is a barrel-shaped complex of about 700 kDa composed of 28 subunits. The complex bears identical halves made of two rings, one with seven distinct α subunits and one with seven distinct β subunits (Groll *et al.*, 1997). The α subunits form the outer platforms of the barrel and provide docking sites for regulator complexes, which control access of substrates and their translocation into the CP. In isolated CPs, the central pore is closed by N-terminal extensions of α subunits (Groll *et al.*, 2000). The two β rings form the catalytic chamber of the protease complex with two sets of three distinct active sites. These reside in the subunits β 1, β 2, and β 5 and cleave peptide bonds preferentially after acidic residues, after basic residues (tryptic activity), or after hydrophobic residues (chymotryptic activity) (Figure 2; Baumeister *et al.*, 1998). The proteasome is a threonine protease because its active sites employ the hydroxyl group of

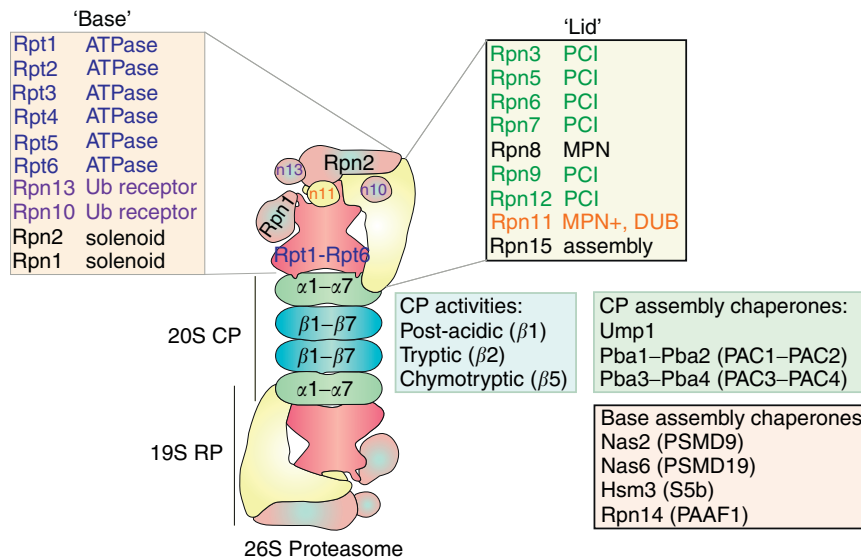


Figure 2 The 26S proteasome: assembly chaperones, subunit composition and functions. Shown is a schematic drawing of the proteasome highlighting its subcomplexes, the 20S core particle (CP) and the 19S regulatory particle (RP). Subunit compositions of the 19S subcomplexes base and lid are given with indications of structurally characteristic domains or functions of the subunits. For the assembly chaperones, both their yeast and their human names (in brackets) are given. DUB, deubiquitylating activity; Ub, ubiquitin.

N-terminal threonine residues as nucleophiles (Seemuller *et al.*, 1995). Products of the digestion of a polypeptide within the proteasome are small peptides typically with an average size of 7–8 amino acid residues (Marques *et al.*, 2009). In some cases, however, the proteasome has also been shown to incompletely degrade (process) substrate proteins, a prime example being the 105 kDa precursor that is processed in a ubiquitin-dependent manner by the proteasome to yield the p50 subunit of NF- κ B (Palombella *et al.*, 1994).

The 19S RPs: Structure and Function

The main proteasome activator complex in eukaryotic cells is the 19S RP (or PA700). It is a ~700 kDa complex composed of at least 19 subunits, which can biochemically be separated into subcomplexes known as the 'base' and the 'lid' (Figure 2; Glickman *et al.*, 1998). The central unit of the base subcomplex is a hexameric ATPase complex of the AAA+ class (ATPases associated with different cellular activities). This part of the RP is formed by 6 distinct Rpt (Regulatory Particle ATPases) subunits, which make direct contact with the α ring surfaces of the CP. This Rpt ring is structurally related to the homo-hexameric PAN complex of the archaeon *Thermophilus acidophilum*, for which a structural model was developed based upon crystal structures of parts of the complex (Zhang *et al.*, 2009). The structural information deduced from the similarity to the PAN complex, as well as structural information on other subunits, together with cryo-electron microscopic analyses of the 19S RP, have recently enabled the generation of a structural model for the entire 19S RP (Figure 3; Lander *et al.*, 2012; Unverdorben *et al.*, 2014). Three of the 19S RP subunits (Rpt2, Rpt3, and Rpt5) share a so-called HbYX motif, in which a hydrophobic residue and a tyrosine are the third last and second last residue, respectively, at the C terminus. These

C-terminal extensions insert into distinct pockets between the α subunits thereby not only contributing to the stability of the complex, but also functioning as a key in the lock to open the central gate of the CP (Smith *et al.*, 2007). As a result, the central channel of the RPT complex aligns with the open gate in the α ring (Matyskiela *et al.*, 2013). Critical residues, including a tyrosine or phenylalanine residue followed by a hydrophobic residue and a glycine (Y/F-I/L/V-G), in loops of the Rpt subunits that align the central channel are thought to mediate binding of substrate domains and their subsequent unfolding and translocation into the 20S CP (Zhang *et al.*, 2009; Martin *et al.*, 2008; Eroles *et al.*, 2012). The latter is the result of ATP-driven conformational changes in the Rpt complex. Additional noteworthy structural features of the Rpt complex are three coiled coil structures each formed by pairs of Rpt subunits. These structural elements have important scaffold function mediating interaction with other regulatory particle non-ATPase (Rpn) subunits of the RP (Beck *et al.*, 2012; Lander *et al.*, 2012). Two subunits in the base, Rpn1 and Rpn2, that share solenoid and arm domains, can be considered as important scaffold subunits and platforms for the interaction with many proteins, including ubiquitin shuttling factors such as Rad23, ubiquitin ligases, and DUBs (Xie and Varshavsky, 2000; Leggett *et al.*, 2002; Elsasser *et al.*, 2002; Husnjak *et al.*, 2008; Lander *et al.*, 2012). Attached to these Rpn subunits are the two intrinsic ubiquitin receptor subunits. Rpn10 binds to ubiquitin chains via a ubiquitin-associated (UBA) domain found in many ubiquitin binding proteins, whereas Rpn13 employs a pleckstrin-homology (Pru) domain for ubiquitin recognition (Husnjak *et al.*, 2008; Schreiner *et al.*, 2008).

The lid is composed of 6 subunits bearing PCI (proteasome, CSN, eIF3) domains, two subunits with MPN (Mpr1/Pad1 N-terminal) domains (Glickman *et al.*, 1998; Scheel and Hofmann, 2005), and the structurally flexible Rpn15 subunit

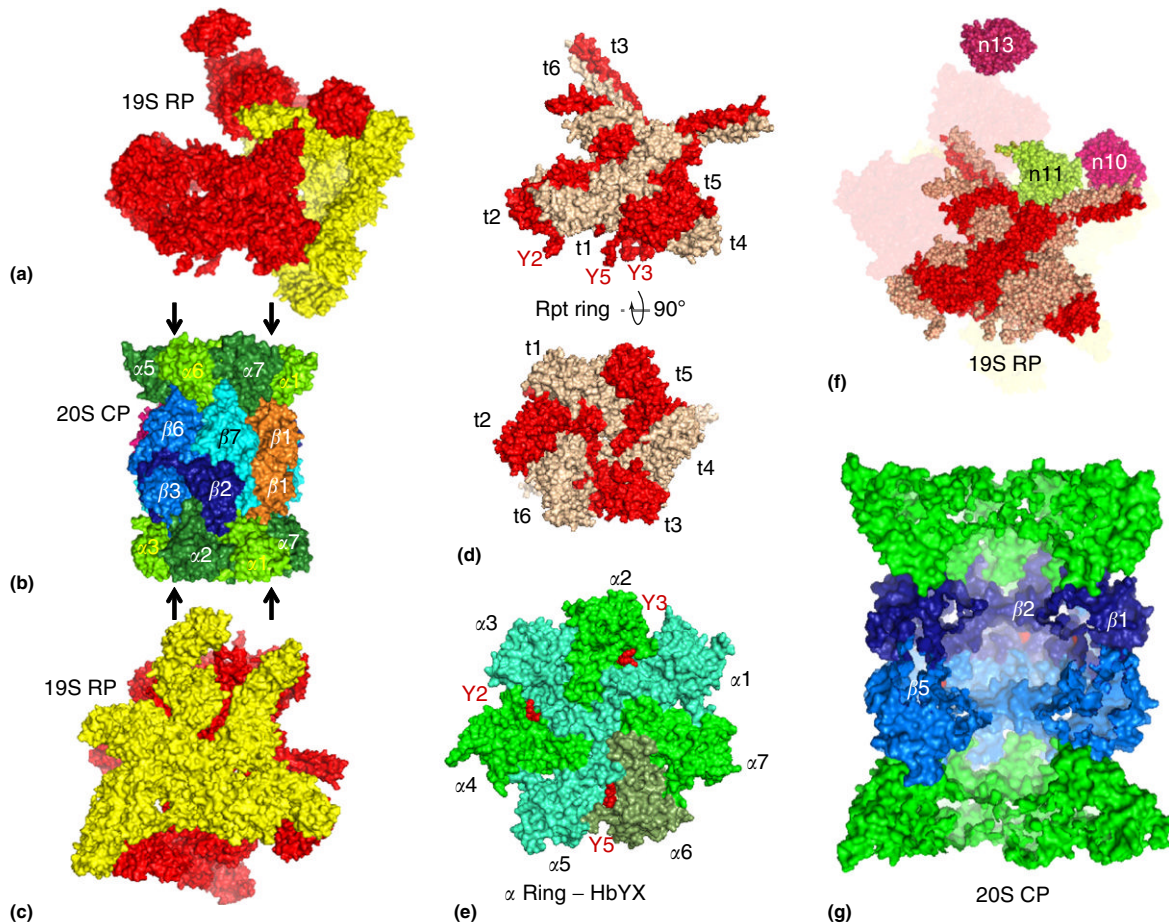


Figure 3 Structure and functional modules of the 26S proteasome. Shown are various depictions of functional modules of the proteasome, which are derived from the X-ray crystal structures of the 20S core particle (CP) of *S. cerevisiae*, and a structural model of the 19S regulatory particle (RP) complex (Groll *et al.* (1997); Unverdorben *et al.* (2014)). (a), (c) Side views of the 19S RP, with base subunits shown in red and lid subunits shown in yellow. The two particles are matching in scale and rotation to join the 20S CP (shown in b) by up or down movements indicated by the arrows. (d) Side (upper part) and bottom views (lower part) of the hexameric Rpt ring. C-terminal HbYX (Y) motifs of Rpt (t) subunits are highlighted. (e) Top view on the heptameric α ring of the 20S CP showing the insertion sites of the HbYX (Y) motifs (in red). (f) Side view of the 19S RP highlighting the Rpt ring and ubiquitin receptor subunits Rpn10 and Rpn13 in the base, as well as of the Rpn13 subunit bearing deubiquitylating activity in the lid. (g) Cut open view of the 20S CP highlighting the active sites (in red) of three of the six active subunits. Figures were prepared using Pymol and PDB files 4CR2 (19S RP) and 1RYP (20S CP) and kindly provided by Paula C. Ramos.

(known as Sem1 in yeast, and Dss1 in humans). The latter has an important function for lid assembly (Tomko and Hochstrasser, 2014). The six C-terminal PCI domains form a horseshoe-like core structure from which the amino-terminal domains of the same subunits extend laterally to form an overall hand-like structure, which is attached to one side of the base reaching down to the CP α ring (Figures 3(a) and 3(c); Beck *et al.*, 2012; Lander *et al.*, 2012; Lasker *et al.*, 2012). The only clearly defined function of the lid subcomplex is provided by its Rpn11 subunit, which bears a metalloprotease activity that mediates substrate deubiquitylation. Rpn11 is positioned right above the entry into the Rpt complex (Lander *et al.*, 2012; Unverdorben *et al.*, 2014). This structural organization fits with biochemical results showing that ATP hydrolysis is critical for Rpn11 DUB activity indicating that substrate unfolding and translocation are coupled with deubiquitylation (Verma *et al.*, 2002; Yao and Cohen, 2002).

Alternative Regulatory Particles and Activators

Aside from the 19S RP, eukaryotic cells express several alternative proteasome activator proteins or complexes. PA200 (Blm10 in yeast) is a large monomeric protein with a C-terminal HbYX motif that has been implicated in DNA damage response and the degradation of some substrates (Ustrell *et al.*, 2002; Tar *et al.*, 2014; Lopez *et al.*, 2011). PA28 is a hexameric complex formed by two distinct subunits, expression of which is induced by interferon- γ . This complex takes part in the immune response by promoting the production of antigenic peptides (Groettrup *et al.*, 2010).

Assembly of the Proteasome

The assembly of the 26S proteasome from 47 individual (33 distinct) subunits is a complex process that involves a

remarkable array of dedicated assembly chaperones, all of which are not part of the final structure (Ramos and Dohmen, 2008; Murata *et al.*, 2009; Tomko and Hochstrasser, 2013; Figure 2). A special case is that of the CP chaperone Ump1 that is enclosed by newly formed proteasomes and degraded upon active site maturation (Ramos *et al.*, 1998).

Coordinated Transcriptional Control of Proteasome Gene Expression

Expression of proteasome genes is subject to a feedback control mechanism, in which the relevant transcriptional activator is itself a proteasomal substrate. This was first discovered in *S. cerevisiae*, where the Rpn4 protein, which was known to interact with 19S RP, was found to bind distinct elements in the promoters of proteasome subunit as well as many other genes (Mannhaupt *et al.*, 1999). Rpn4 is degraded by the proteasome both by ubiquitin-independent and ubiquitin-dependent mechanisms (Xie and Varshavsky, 2001; Ju and Xie, 2004). A similar, but in some aspects also distinct, feedback mechanism operates in mammalian cells, which use the transcription factor Nrf1 to control proteasome gene expression (Steffen *et al.*, 2010; Radhakrishnan *et al.*, 2010). Inactive Nrf1 resides in the membrane of the endoplasmic reticulum (ER). Its activation involves Cdc48/p97-dependent retrotranslocation into the cytosol and processing (Radhakrishnan *et al.*, 2014). However, if sufficient proteasomal activity is present in the cells, most of Nrf1 is degraded upon its arrival at the cytosolic surface of the ER. For both Rpn4 and Nrf1, inhibition or overload of the proteasome leads to stabilization and hence up-regulation of proteasome gene expression (Xie and Varshavsky, 2001; London *et al.*, 2004; Ju *et al.*, 2004; Steffen *et al.*, 2010; Radhakrishnan *et al.*, 2010).

The Immunoproteasome and Alternative CP Subunits

The proteasome has an important function in the immune system by supplying it with antigenic peptides, which are presented by major histocompatibility complex class I molecules on the surface of cells, e.g., upon infection with a pathogen (Goldberg *et al.*, 2002; Groettrup *et al.*, 2010; Basler *et al.*, 2013). Stimulation of an immune response transmitted by interferon- γ leads to the induction of three alternative β subunits (β 1i/LMP2/PSMB9, β 2i/MECL-1/PSMB10, and β 5i/LMP7/PSMB8), which assemble into newly formed CPs instead of their standard counterparts. The resulting immunoproteasome has slightly different specificities in the selection of substrate cleavage sites, which are thought to contribute to the generation of antigenic peptides (Goldberg *et al.*, 2002; Groettrup *et al.*, 2010; Basler *et al.*, 2013).

Another alternative subunit that leads to a distinct CP subtype termed the 'thymoproteasome' is β 5t. This subunit is expressed in the thymus in cortical epithelial cells (cTECs) and is important for the positive selection of CD8⁺ T cells (Murata *et al.*, 2007).

The only alternative α subunit identified thus far is α 4s/PSMB8, which is specifically expressed and incorporated into CPs in male germ cells. The functional significance of the novel proteasome subtype remains unclear as no difference in

biochemical activities were detectable in comparison to standard proteasomes (Uechi *et al.*, 2014). It was recently suggested, however, that 'spermatoproteasomes' characterized by the presence of α 4s and the PA200 activator promote acetylation-dependent degradation of core histones during spermatogenesis (Qian *et al.*, 2013).

The UPS as Drug Target

The Ubiquitin-Conjugation System

The UPS is involved in the control of many fundamental cellular processes including cell cycle, various signal transduction pathways, differentiation, and DNA repair. Thus, not surprisingly, deregulation of substrates or components of the UPS has been associated with various human diseases including rather complex disorders such as neurodegenerative diseases and various types of cancer as well as mono-genetic disorders such as cystic fibrosis and Angelman syndrome, a neurodevelopmental disorder (Bedford *et al.*, 2011; Ciechanover, 2013; Metzger *et al.*, 2014; Scheffner and Kumar, 2014). Accordingly, the UPS has become an attractive therapeutic target (see also section NEDD8). Depending on the UPS component affected/involved and the type of deregulation (inactivation or inappropriate activation), one can envision different strategies for molecular intervention. For example, 'loss-of-function' mutations in the genes encoding the RBR E3 ligase PARKIN and the HECT E3 ligase E6AP have been associated with autosomal recessive juvenile parkinsonism and Angelman syndrome, respectively (Ardley and Robinson, 2004; Bird, 2014). Since this presumably results in hyperactivation of respective substrate proteins, these would be preferred targets for pharmacological intervention. In contrast, amplification of the gene encoding the RING E3 ligase HDM2 (human Mdm2) has been associated with the development of osteosarcomas. Since HDM2 plays a prominent role in degradation of the tumor suppressor protein p53 (Brooks and Gu, 2011; Hock and Vousden, 2014), intervention with HDM2 activity or with the interaction of HDM2 with p53 appears to be the strategy of choice. Indeed, small molecules (termed Nutlins) binding to the p53-binding pocket of HDM2 interfere with HDM2-mediated ubiquitylation of p53 within cells (Shen and Maki, 2011). In more general terms, targeting the interfaces of E3-substrate interactions seems to be an attractive and most selective option in the treatment of diseases where enhanced degradation of a substrate contributes to the disease phenotype. Alternatively, due to the notion that in contrast to most of the E3 ligases (i.e. RING ligases), DUBs have a defined catalytic center and presumably show a certain degree of substrate specificity (as suggested by the number of DUBs encoded by the human genome), DUBs represent attractive therapeutic targets (Shanmugham and Ovaa, 2008).

The Proteasome

An important advance for the investigation of proteasome function and for its targeting in human diseases was the identification of specific natural and synthetic inhibitors (Vinitsky *et al.*, 1992; Rock *et al.*, 1994; Fenteany *et al.*, 1995).

One of the first identified synthetic peptide inhibitors of the proteasome, MG132, is still widely used in addressing proteasomal functions in cell biology (Goldberg, 2012). In the meanwhile, a wide repertoire of inhibitors has been established, all of which have in common that they target the active sites of the 20S CP (Borissenko and Groll, 2007; Moore *et al.*, 2008). The proteasome inhibitor Bortezomib/Velcade was the first approved drug that specifically targets the proteasome. It is successfully used in the treatment of multiple myeloma, a cancer affecting blood cells that secrete large amounts of aberrant antibody polypeptides (Goldberg, 2012). One reason that these lymphoma cells are hypersensitive to proteasome inhibitor appears to be that they require high proteasome activity to cope with the abnormal antibody polypeptides generated by ERAD. Interfering with this process by proteasome inhibition drives these cells into apoptosis.

More recent advances are inhibitors that provide a species- or subtype-specific targeting of proteasomes. One such class of inhibitors was shown to be specific for the proteasome of the eubacterial human pathogen *Mycobacterium tuberculosis* (Lin *et al.*, 2009). Another inhibitor was shown to preferentially inhibit the immunoproteasome (Muchamuel *et al.*, 2009). Both types of inhibitors have the potential for the treatment of specific human diseases.

Conclusion

The past three to four decades have witnessed the transformation of a research field (ubiquitin-mediated protein degradation) that was studied by only a few pioneering groups into one of the most intensely studied research areas. Similarly, while in the beginning ubiquitin was assumed to function only in proteasome-mediated degradation, it has become clear that many aspects or properties of a protein can be affected by modification with ubiquitin or UBLs. In addition, ubiquitylation is not a uniform posttranslational modification (e.g., such as phosphorylation) but rather comes in many different modes/types and involves an ever-increasing number of enzymes and proteins. The complexity of the UPS, however, also poses a number of challenging questions. For example, do different types of ubiquitylation signal a modified protein for distinct fates or do they have overlapping/redundant functions? How are the different types eventually decoded? Does this solely depend on proteins that selectively recognize different types of ubiquitylation or do additional properties of the modified proteins contribute? Furthermore, the high number of putative E3s provides the UPS with the ability to operate in a highly substrate-specific manner. Nonetheless, for most E3s, the substrates have not yet been identified (and vice versa) and, at least in some cases, E3s appear to have overlapping sets of substrates. For example, more than 10 E3s have been reported to be involved in p53 ubiquitylation/degradation (Brooks and Gu, 2011; Hock and Vousden, 2014). Thus, besides determining the potential substrate spectra of E3s and how these are affected by crosstalk with other posttranslational modifications, a major challenge will be in defining when (e.g., during cell cycle or differentiation) and to what extent a given E3 contributes to the degradation of a given protein. The proteasome is a central factor in the UPS

and a valuable drug target. Recent studies have led to major advances in the understanding of its structure and function. Yet, many challenging questions remain open regarding the recognition of distinct types of substrates (e.g., ubiquitin-independent *versus* ubiquitin-dependent) and their competition in a cellular context, as well as concerning the transcriptional and posttranscriptional regulation of the proteasome.

See also: Protein Synthesis/Degradation: Protein Degradation – General: Mass Spectrometry-based Methodologies for Studying Proteolytic Networks and the Degradome

References

- Ardley, H.C., Robinson, P.A., 2004. The role of ubiquitin-protein ligases in neurodegenerative disease. *Neurodegenerative Diseases* 1, 71–87.
- Baker, R.T., Board, P.G., 1987. The human ubiquitin gene family: Structure of a gene and pseudogenes from the Ub B subfamily. *Nucleic Acids Research* 15, 443–463.
- Barford, D., 2011. Structural insights into anaphase-promoting complex function and mechanism. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 366, 3605–3624.
- Basler, M., Kirk, C.J., Groettrup, M., 2013. The immunoproteasome in antigen processing and other immunological functions. *Current Opinion in Immunology* 25, 74–80.
- Bassermann, F., Eichner, R., Pagano, M., 2014. The ubiquitin proteasome system - implications for cell cycle control and the targeted treatment of cancer. *Biochimica et Biophysica Acta* 1843, 150–162.
- Baumeister, W., Walz, J., Zuhl, F., Seemuller, E., 1998. The proteasome: Paradigm of a self-compartmentalizing protease. *Cell* 92, 367–380.
- Beck, F., Unverdorben, P., Bohn, S., *et al.*, 2012. Near-atomic resolution structural model of the yeast 26S proteasome. *Proceedings of the National Academy of Sciences of the United States of America* 109, 14870–14875.
- Bedford, L., Lowe, J., Dick, L.R., Mayer, R.J., Brownell, J.E., 2011. Ubiquitin-like protein conjugation and the ubiquitin-proteasome system as drug targets. *Nature Reviews Drug Discovery* 10, 29–46.
- Berndsen, C.E., Wolberger, C., 2014. New insights into ubiquitin E3 ligase mechanism. *Nature Structural & Molecular Biology* 21, 301–307.
- Bhattacharyya, S., Yu, H., Mim, C., Matouschek, A., 2014. Regulated protein turnover: Snapshots of the proteasome in action. *Nature Reviews Molecular Cell Biology* 15, 122–133.
- Bird, L.M., 2014. Angelman syndrome: Review of clinical and molecular aspects. *Application of Clinical Genetics* 7, 93–104.
- Bogunovic, D., Byun, M., Durfee, L.A., *et al.*, 2012. Mycobacterial disease and impaired IFN-gamma immunity in humans with inherited ISG15 deficiency. *Science* 337, 1684–1688.
- Borissenko, L., Groll, M., 2007. 20S proteasome and its inhibitors: Crystallographic knowledge for drug development. *Chemical Reviews* 107, 687–717.
- Brooks, C.L., Gu, W., 2011. p53 regulation by ubiquitin. *FEBS Letters* 585, 2803–2809.
- Burroughs, A.M., Iyer, L.M., Aravind, L., 2012. Structure and evolution of ubiquitin and ubiquitin-related domains. *Methods in Molecular Biology* 832, 15–63.
- Chau, V., Tobias, J.W., Bachmair, A., *et al.*, 1989. A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* 243, 1576–1583.
- Ciechanover, A., 2013. Intracellular protein degradation: From a vague idea through the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. *Bioorganic & Medicinal Chemistry* 21, 3400–3410.
- Ciechanover, A., Elias, S., Heller, H., Hershko, A., 1982. "Covalent affinity" purification of ubiquitin-activating enzyme. *Journal of Biological Chemistry* 257, 2537–2542.
- Ciechanover, A., Finley, D., Varshavsky, A., 1984. Ubiquitin dependence of selective protein degradation demonstrated in the mammalian cell cycle mutant ts85. *Cell* 37, 57–66.
- Ciechanover, A., Heller, H., Elias, S., Haas, A.L., Hershko, A., 1980. ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein

- degradation. *Proceedings of the National Academy of Sciences of the United States of America* 77, 1365–1368.
- Ciechanover, A., Hod, Y., Hershko, A., 1978. A heat-stable polypeptide component of an ATP-dependent proteolytic system from reticulocytes. *Biochemical and Biophysical Research Communications* 81, 1100–1105.
- Clague, M.J., Coulson, J.M., Urbe, S., 2012. Cellular functions of the DUBs. *Journal of Cell Science* 125, 277–286.
- Crosas, B., Hanna, J., Kirkpatrick, D.S., *et al.*, 2006. Ubiquitin chains are remodeled at the proteasome by opposing ubiquitin ligase and deubiquitinating activities. *Cell* 127, 1401–1413.
- Cubenas-Potts, C., Matunis, M.J., 2013. SUMO: A multifaceted modifier of chromatin structure and function. *Developmental Cell* 24, 1–12.
- D'cunha, J., Knight Jr., E., Haas, A.L., Truitt, R.L., Borden, E.C., 1996. Immunoregulatory properties of ISG15, an interferon-induced cytokine. *Proceedings of the National Academy of Sciences of the United States of America* 93, 211–215.
- Deshais, R.J., Emberley, E.D., Saha, A., 2010. Control of cullin-ring ubiquitin ligase activity by nedd8. *Subcellular Biochemistry* 54, 41–56.
- Deshais, R.J., Joazeiro, C.A., 2009. RING domain E3 ubiquitin ligases. *Annual Review of Biochemistry* 78, 399–434.
- Dikic, I., Wakatsuki, S., Walters, K.J., 2009. Ubiquitin-binding domains - from structures to functions. *Nature Reviews Molecular Cell Biology* 10, 659–671.
- Durfee, L.A., Huibregtse, J.M., 2012. The ISG15 conjugation system. *Methods in Molecular Biology* 832, 141–149.
- Durfee, L.A., Lyon, N., Seo, K., Huibregtse, J.M., 2010. The ISG15 conjugation system broadly targets newly synthesized proteins: Implications for the antiviral function of ISG15. *Molecular Cell* 38, 722–732.
- Eletr, Z.M., Wilkinson, K.D., 2014. Regulation of proteolysis by human deubiquitinating enzymes. *Biochimica et Biophysica Acta* 1843, 114–128.
- Elsasser, S., Gali, R.R., Schwickart, M., *et al.*, 2002. Proteasome subunit Rpn1 binds ubiquitin-like protein domains. *Nature Cell Biology* 4, 725–730.
- Erales, J., Coffino, P., 2014. Ubiquitin-independent proteasomal degradation. *Biochimica et Biophysica Acta* 1843, 216–221.
- Erales, J., Hoyt, M.A., Troll, F., Coffino, P., 2012. Functional asymmetries of proteasome translocase pore. *Journal of Biological Chemistry* 287, 18535–18543.
- Etlinger, J.D., Goldberg, A.L., 1977. A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. *Proceedings of the National Academy of Sciences of the United States of America* 74, 54–58.
- Fenteany, G., Standaert, R.F., Lane, W.S., *et al.*, 1995. Inhibition of proteasome activities and subunit-specific amino-terminal threonine modification by lactacystin. *Science* 268, 726–731.
- Finley, D., 2009. Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annual Review of Biochemistry* 78, 477–513.
- Finley, D., Ciechanover, A., Varshavsky, A., 1984. Thermolability of ubiquitin-activating enzyme from the mammalian cell cycle mutant ts85. *Cell* 37, 43–55.
- Gareau, J.R., Lima, C.D., 2010. The SUMO pathway: Emerging mechanisms that shape specificity, conjugation and recognition. *Nature Reviews Molecular Cell Biology* 11, 861–871.
- Geoffroy, M.C., Hay, R.T., 2009. An additional role for SUMO in ubiquitin-mediated proteolysis. *Nature Reviews Molecular Cell Biology* 10, 564–568.
- Glickman, M.H., Rubin, D.M., Coux, O., *et al.*, 1998. A subcomplex of the proteasome regulatory particle required for ubiquitin-conjugate degradation and related to the COP9-signalosome and eIF3. *Cell* 94, 615–623.
- Goldberg, A.L., 2012. Development of proteasome inhibitors as research tools and cancer drugs. *Journal of Cell Biology* 199, 583–588.
- Goldberg, A.L., Cascio, P., Saric, T., Rock, K.L., 2002. The importance of the proteasome and subsequent proteolytic steps in the generation of antigenic peptides. *Molecular Immunology* 39, 147–164.
- Goldknopf, I.L., Busch, H., 1977. Isopeptide linkage between nonhistone and histone 2A polypeptides of chromosomal conjugate-protein A24. *Proceedings of the National Academy of Sciences of the United States of America* 74, 864–868.
- Goldstein, G., Scheid, M., Hammerling, U., *et al.*, 1975. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proceedings of the National Academy of Sciences of the United States of America* 72, 11–15.
- Groettrup, M., Kirk, C.J., Basler, M., 2010. Proteasomes in immune cells: More than peptide producers? *Nature Reviews Immunology* 10, 73–78.
- Groll, M., Bajorek, M., Kohler, A., *et al.*, 2000. A gated channel into the proteasome core particle. *Nature Structural Biology* 7, 1062–1067.
- Groll, M., Ditzel, L., Lowe, J., *et al.*, 1997. Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* 386, 463–471.
- Haas, A.L., Warms, J.V., Hershko, A., Rose, I.A., 1982. Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. *Journal of Biological Chemistry* 257, 2543–2548.
- Hanada, T., Noda, N.N., Satomi, Y., *et al.*, 2007. The Atg12-Atg5 conjugate has a novel E3-like activity for protein lipidation in autophagy. *Journal of Biological Chemistry* 282, 37298–37302.
- Hanzelmann, P., Schafer, A., Voller, D., Schindelin, H., 2012. Structural insights into functional modes of proteins involved in ubiquitin family pathways. *Methods in Molecular Biology* 832, 547–576.
- Hershko, A., Ciechanover, A., 1998. The ubiquitin system. *Annual Review of Biochemistry* 67, 425–479.
- Hershko, A., Ciechanover, A., Heller, H., Haas, A.L., Rose, I.A., 1980. Proposed role of ATP in protein breakdown: Conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. *Proceedings of the National Academy of Sciences of the United States of America* 77, 1783–1786.
- Hershko, A., Heller, H., Elias, S., Ciechanover, A., 1983. Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. *Journal of Biological Chemistry* 258, 8206–8214.
- Hochstrasser, M., 2000. Evolution and function of ubiquitin-like protein-conjugation systems. *Nature Cell Biology* 2, E153–E157.
- Hochstrasser, M., 2009. Origin and function of ubiquitin-like proteins. *Nature* 458, 422–429.
- Hock, A.K., Vousden, K.H., 2014. The role of ubiquitin modification in the regulation of p53. *Biochimica et Biophysica Acta* 1843, 137–149.
- Huibregtse, J.M., Scheffner, M., Beaudenon, S., Howley, P.M., 1995. A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proceedings of the National Academy of Sciences of the United States of America* 92, 2563–2567.
- Humbard, M.A., Maupin-Furlow, J.A., 2013. Prokaryotic proteasomes: Nanocompartments of degradation. *Journal of Molecular Microbiology and Biotechnology* 23, 321–334.
- Hunt, L.T., Dayhoff, M.O., 1977. Amino-terminal sequence identity of ubiquitin and the nonhistone component of nuclear protein A24. *Biochemical and Biophysical Research Communications* 74, 650–655.
- Hurley, J.H., Schulman, B.A., 2014. Autistic autophagy: The structures of cellular self-digestion. *Cell* 157, 300–311.
- Husnjak, K., Dikic, I., 2012. Ubiquitin-binding proteins: Decoders of ubiquitin-mediated cellular functions. *Annual Review of Biochemistry* 81, 291–322.
- Husnjak, K., Elsasser, S., Zhang, N., *et al.*, 2008. Proteasome subunit Rpn13 is a novel ubiquitin receptor. *Nature* 453, 481–488.
- Jentsch, S., Psakhye, I., 2013. Control of nuclear activities by substrate-selective and protein-group SUMOylation. *Annual Review of Genetics* 47, 167–186.
- Ju, D., Wang, L., Mao, X., Xie, Y., 2004. Homeostatic regulation of the proteasome via an Rpn4-dependent feedback circuit. *Biochemical and Biophysical Research Communications* 321, 51–57.
- Ju, D., Xie, Y., 2004. Proteasomal degradation of RPN4 via two distinct mechanisms, ubiquitin-dependent and -independent. *Journal of Biological Chemistry* 279, 23851–23854.
- Kim, W., Bennett, E.J., Huttlin, E.L., *et al.*, 2011. Systematic and quantitative assessment of the ubiquitin-modified proteome. *Molecular Cell* 44, 325–340.
- Klionsky, D.J., Schulman, B.A., 2014. Dynamic regulation of macroautophagy by distinctive ubiquitin-like proteins. *Nature Structural & Molecular Biology* 21, 336–345.
- Komander, D., Clague, M.J., Urbe, S., 2009. Breaking the chains: Structure and function of the deubiquitinases. *Nature Reviews Molecular Cell Biology* 10, 550–563.
- Komander, D., Rape, M., 2012. The ubiquitin code. *Annual Review of Biochemistry* 81, 203–229.
- Kravtsova-Ivantsiv, Y., Ciechanover, A., 2012. Non-canonical ubiquitin-based signals for proteasomal degradation. *Journal of Cell Science* 125, 539–548.
- Kravtsova-Ivantsiv, Y., Sommer, T., Ciechanover, A., 2013. The lysine48-based polyubiquitin chain proteasomal signal: Not a single child anymore. *Angewandte Chemie (International ed. in English)* 52, 192–198.
- Lander, G.C., Estrin, E., Matyskiela, M.E., *et al.*, 2012. Complete subunit architecture of the proteasome regulatory particle. *Nature* 482, 186–191.
- Lasker, K., Forster, F., Bohn, S., *et al.*, 2012. Molecular architecture of the 26S proteasome holocomplex determined by an integrative approach. *Proceedings of the National Academy of Sciences of the United States of America* 109, 1380–1387.
- Leggett, D.S., Hanna, J., Borodovsky, A., *et al.*, 2002. Multiple associated proteins regulate proteasome structure and function. *Molecular Cell* 10, 495–507.

- Lin, G., Li, D., De Carvalho, L.P., *et al.*, 2009. Inhibitors selective for mycobacterial versus human proteasomes. *Nature* 461, 621–626.
- Liz, M.A., Sousa, M.M., 2005. Deciphering cryptic proteases. *Cellular and Molecular Life Sciences* 62, 989–1002.
- London, M.K., Keck, B.I., Ramos, P.C., Dohmen, R.J., 2004. Regulatory mechanisms controlling biogenesis of ubiquitin and the proteasome. *FEBS Letters* 567, 259–264.
- Lopez, A.D., Tar, K., Krugel, U., *et al.*, 2011. Proteasomal degradation of Stp1 contributes to the repression of ribosome biogenesis during starvation and is mediated by the proteasome activator Blm10. *Molecular Biology of the Cell* 22, 528–540.
- Lydeard, J.R., Schulman, B.A., Harper, J.W., 2013. Building and remodelling Cullin-RING E3 ubiquitin ligases. *EMBO Reports* 14, 1050–1061.
- Malakhov, M.P., Malakhova, O.A., Kim, K.I., Ritchie, K.J., Zhang, D.E., 2002. UBP43 (USP18) specifically removes ISG15 from conjugated proteins. *Journal of Biological Chemistry* 277, 9976–9981.
- Mannhaupt, G., Schnall, R., Karpov, V., Vetter, I., Feldmann, H., 1999. Rpn4p acts as a transcription factor by binding to PACE, a nonamer box found upstream of 26S proteasomal and other genes in yeast. *FEBS Letters* 450, 27–34.
- Marques, A.J., Palanimurugan, R., Matias, A.C., Ramos, P.C., Dohmen, R.J., 2009. Catalytic mechanism and assembly of the proteasome. *Chemical Reviews* 109, 1509–1536.
- Martin, A., Baker, T.A., Sauer, R.T., 2008. Pore loops of the AAA + ClpX machine grip substrates to drive translocation and unfolding. *Nature Structural & Molecular Biology* 15, 1147–1151.
- Matyskiela, M.E., Lander, G.C., Martin, A., 2013. Conformational switching of the 26S proteasome enables substrate degradation. *Nature Structural & Molecular Biology* 20, 781–788.
- Metzger, M.B., Pruneda, J.N., Klevit, R.E., Weissman, A.M., 2014. RING-type E3 ligases: Master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochimica et Biophysica Acta* 1843, 47–60.
- Meyer, H.J., Rape, M., 2011. Processive ubiquitin chain formation by the anaphase-promoting complex. *Seminars in Cell & Developmental Biology* 22, 544–550.
- Moore, B.S., Eustaquio, A.S., McGlinchey, R.P., 2008. Advances in and applications of proteasome inhibitors. *Current Opinion in Chemical Biology* 12, 434–440.
- Morales, D.J., Lenschow, D.J., 2013. The antiviral activities of ISG15. *Journal of Molecular Biology* 425, 4995–5008.
- Muchamuel, T., Basler, M., Aujay, M.A., *et al.*, 2009. A selective inhibitor of the immunoproteasome subunit LMP7 blocks cytokine production and attenuates progression of experimental arthritis. *Nature Medicine* 15, 781–787.
- Murata, S., Sasaki, K., Kishimoto, T., *et al.*, 2007. Regulation of CD8 + T cell development by thymus-specific proteasomes. *Science* 316, 1349–1353.
- Murata, S., Yashiroda, H., Tanaka, K., 2009. Molecular mechanisms of proteasome assembly. *Nature Reviews Molecular Cell Biology* 10, 104–115.
- Nagai, S., Davoodi, N., Gasser, S.M., 2011. Nuclear organization in genome stability: SUMO connections. *Cell Research* 21, 474–485.
- Nawrocki, S.T., Griffin, P., Kelly, K.R., Carew, J.S., 2012. MLN4924: A novel first-in-class inhibitor of NEDD8-activating enzyme for cancer therapy. *Expert Opinion on Investigational Drugs* 21, 1563–1573.
- Ozkaynak, E., Finley, D., Solomon, M.J., Varshavsky, A., 1987. The yeast ubiquitin genes: A family of natural gene fusions. *EMBO Journal* 6, 1429–1439.
- Palombella, V.J., Rando, O.J., Goldberg, A.L., Maniatis, T., 1994. The ubiquitin-proteasome pathway is required for processing the NF-kappa B1 precursor protein and the activation of NF-kappa B. *Cell* 78, 773–785.
- Pearce, M.J., Mintseris, J., Ferreyra, J., Gygi, S.P., Darwin, K.H., 2008. Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science* 322, 1104–1107.
- Perry, J.J., Tainer, J.A., Boddy, M.N., 2008. A SIM-ultaneous role for SUMO and ubiquitin. *Trends in Biochemical Sciences* 33, 201–208.
- Qian, M.X., Pang, Y., Liu, C.H., *et al.*, 2013. Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis. *Cell* 153, 1012–1024.
- Radhakrishnan, S.K., Den Besten, W., Deshaies, R.J., 2014. p97-dependent retrotranslocation and proteolytic processing govern formation of active Nrf1 upon proteasome inhibition. *Elife* 3, e01856.
- Radhakrishnan, S.K., Lee, C.S., Young, P., *et al.*, 2010. Transcription factor Nrf1 mediates the proteasome recovery pathway after proteasome inhibition in mammalian cells. *Molecular Cell* 38, 17–28.
- Ramos, P.C., Dohmen, R.J., 2008. PACemakers of proteasome core particle assembly. *Structure* 16, 1296–1304.
- Ramos, P.C., Hockendorff, J., Johnson, E.S., Varshavsky, A., Dohmen, R.J., 1998. Ump1p is required for proper maturation of the 20S proteasome and becomes its substrate upon completion of the assembly. *Cell* 92, 489–499.
- Reyes-Turcu, F.E., Ventii, K.H., Wilkinson, K.D., 2009. Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. *Annual Review of Biochemistry* 78, 363–397.
- Rock, K.L., Gramm, C., Rothstein, L., *et al.*, 1994. Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules. *Cell* 78, 761–771.
- Sarikas, A., Hartmann, T., Pan, Z.Q., 2011. The cullin protein family. *Genome Biology* 12, 220.
- Scheel, H., Hofmann, K., 2005. Prediction of a common structural scaffold for proteasome lid, COP9-signalosome and eIF3 complexes. *BMC Bioinformatics* 6, 71.
- Scheffner, M., Kumar, S., 2014. Mammalian HECT ubiquitin-protein ligases: Biological and pathophysiological aspects. *Biochimica et Biophysica Acta* 1843, 61–74.
- Schmidtke, G., Aichele, A., Groettrup, M., 2014. FAT10ylation as a signal for proteasomal degradation. *Biochimica et Biophysica Acta* 1843, 97–102.
- Schreiner, P., Chen, X., Husnjak, K., *et al.*, 2008. Ubiquitin docking at the proteasome through a novel pleckstrin-homology domain interaction. *Nature* 453, 548–552.
- Seemuller, E., Lupas, A., Stock, D., *et al.*, 1995. Proteasome from *Thermoplasma acidophilum*: A threonine protease. *Science* 268, 579–582.
- Shanmugham, A., Ovaa, H., 2008. DUBs and disease: Activity assays for inhibitor development. *Current Opinion in Drug Discovery & Development* 11, 688–696.
- Shen, H., Maki, C.G., 2011. Pharmacologic activation of p53 by small-molecule MDM2 antagonists. *Current Pharmaceutical Design* 17, 560–568.
- Shigi, N., 2014. Biosynthesis and functions of sulfur modifications in tRNA. *Frontiers in Genetics* 5, 67.
- Slobodkin, M.R., Elazar, Z., 2013. The Atg8 family: Multifunctional ubiquitin-like key regulators of autophagy. *Essays in Biochemistry* 55, 51–64.
- Smit, J.J., Sixma, T.K., 2014. RBR E3-ligases at work. *EMBO Reports* 15, 142–154.
- Smith, D.M., Chang, S.C., Park, S., *et al.*, 2007. Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's alpha ring opens the gate for substrate entry. *Molecular Cell* 27, 731–744.
- Smith, M.H., Ploegh, H.L., Weissman, J.S., 2011. Road to ruin: Targeting proteins for degradation in the endoplasmic reticulum. *Science* 334, 1086–1090.
- Sriramachandran, A.M., Dohmen, R.J., 2014. SUMO-targeted ubiquitin ligases. *Biochimica et Biophysica Acta* 1843, 75–85.
- Steffen, J., Seeger, M., Koch, A., Kruger, E., 2010. Proteasomal degradation is transcriptionally controlled by TCF11 via an ERAD-dependent feedback loop. *Molecular Cell* 40, 147–158.
- Tar, K., Dange, T., Yang, C., *et al.*, 2014. Proteasomes Associated with the Blm10 Activator Protein Antagonize Mitochondrial Fission through Degradation of the Fission Protein Dnm1. *Journal of Biological Chemistry* 289, 12145–12156.
- Tomko Jr., R.J., Hochstrasser, M., 2013. Molecular architecture and assembly of the eukaryotic proteasome. *Annual Review of Biochemistry* 82, 415–445.
- Tomko Jr., R.J., Hochstrasser, M., 2014. The intrinsically disordered Sem1 protein functions as a molecular tether during proteasome lid biogenesis. *Molecular Cell* 53, 433–443.
- Uechi, H., Hamazaki, J., Murata, S., 2014. Characterization of the Testis-specific Proteasome Subunit alpha4s in Mammals. *Journal of Biological Chemistry* 289, 12365–12374.
- Unverdorben, P., Beck, F., Sledz, P., *et al.*, 2014. Deep classification of a large cryo-EM dataset defines the conformational landscape of the 26S proteasome. *Proceedings of the National Academy of Sciences of the United States of America* 111, 5544–5549.
- Ustrell, V., Hoffman, L., Pratt, G., Rechsteiner, M., 2002. PA200, a nuclear proteasome activator involved in DNA repair. *EMBO Journal* 21, 3516–3525.
- Verma, R., Aravind, L., Oania, R., *et al.*, 2002. Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. *Science* 298, 611–615.
- Verma, R., Chen, S., Feldman, R., *et al.*, 2000. Proteasomal proteomics: Identification of nucleotide-sensitive proteasome-interacting proteins by mass spectrometric analysis of affinity-purified proteasomes. *Molecular Biology of the Cell* 11, 3425–3439.
- Vijay-Kumar, S., Bugg, C.E., Wilkinson, K.D., *et al.*, 1987. Comparison of the three-dimensional structures of human, yeast, and oat ubiquitin. *Journal of Biological Chemistry* 262, 6396–6399.
- Vinitsky, A., Michaud, C., Powers, J.C., Orłowski, M., 1992. Inhibition of the chymotrypsin-like activity of the pituitary multicatalytic proteinase complex. *Biochemistry* 31, 9421–9428.
- Weake, V.M., Workman, J.L., 2008. Histone ubiquitination: Triggering gene activity. *Molecular Cell* 29, 653–663.
- Wenzel, D.M., Klevit, R.E., 2012. Following Ariadne's thread: A new perspective on RBR ubiquitin ligases. *BMC Biology* 10, 24.

- Wiborg, O., Pedersen, M.S., Wind, A., *et al.*, 1985. The human ubiquitin multigene family: Some genes contain multiple directly repeated ubiquitin coding sequences. *EMBO Journal* 4, 755–759.
- Wilkinson, K.D., Urban, M.K., Haas, A.L., 1980. Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. *Journal of Biological Chemistry* 255, 7529–7532.
- Xie, Y., Varshavsky, A., 2000. Physical association of ubiquitin ligases and the 26S proteasome. *Proceedings of the National Academy of Sciences of the United States of America* 97, 2497–2502.
- Xie, Y., Varshavsky, A., 2001. RPN4 is a ligand, substrate, and transcriptional regulator of the 26S proteasome: A negative feedback circuit. *Proceedings of the National Academy of Sciences of the United States of America* 98, 3056–3061.
- Xu, P., Duong, D.M., Seyfried, N.T., *et al.*, 2009. Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. *Cell* 137, 133–145.
- Xu, P., Peng, J., 2006. Dissecting the ubiquitin pathway by mass spectrometry. *Biochimica et Biophysica Acta* 1764, 1940–1947.
- Yao, T., Cohen, R.E., 2002. A cryptic protease couples deubiquitination and degradation by the proteasome. *Nature* 419, 403–407.
- Zhang, F., Hu, M., Tian, G., *et al.*, 2009. Structural insights into the regulatory particle of the proteasome from *Methanocaldococcus jannaschii*. *Molecular Cell* 34, 473–484.
- Zhang, F., Wu, Z., Zhang, P., *et al.*, 2009. Mechanism of substrate unfolding and translocation by the regulatory particle of the proteasome from *Methanocaldococcus jannaschii*. *Molecular Cell* 34, 485–496.
- Zhao, C., Collins, M.N., Hsiang, T.Y., Krug, R.M., 2013. Interferon-induced ISG15 pathway: An ongoing virus-host battle. *Trends in Microbiology* 21, 181–186.
- Zimmerman, E.S., Schulman, B.A., Zheng, N., 2010. Structural assembly of cullin-RING ubiquitin ligase complexes. *Current Opinion in Structural Biology* 20, 714–721.
- Finley, D., Ulrich, H.D., Sommer, T., Kaiser, P., 2012. The ubiquitin-proteasome system of *Saccharomyces cerevisiae*. *Genetics* 192, 319–360.
- Fulda, S., Rajalingam, K., Dikic, I., 2012. Ubiquitylation in immune disorders and cancer: From molecular mechanisms to therapeutic implications. *EMBO Molecular Medicine* 4, 545–556.
- Harper, J.W., Tan, M.K., 2012. Understanding cullin-RING E3 biology through proteomics-based substrate identification. *Molecular & Cellular Proteomics* 11, 1541–1550.
- Jentsch, S., Muller, S., 2010. Regulatory Functions of Ubiquitin and SUMO in DNA Repair Pathways. *Subcellular Biochemistry* 54, 184–194.
- Lipkowitz, S., Weissman, A.M., 2011. RINGS of good and evil: RING finger ubiquitin ligases at the crossroads of tumour suppression and oncogenesis. *Nature Reviews Cancer* 11, 629–643.
- Mattiroli, F., Sixma, T.K., 2014. Lysine-targeting specificity in ubiquitin and ubiquitin-like modification pathways. *Nature Structural & Molecular Biology* 21, 308–316.
- Micel, L.N., Tentler, J.J., Smith, P.G., Eckhardt, G.S., 2013. Role of ubiquitin ligases and the proteasome in oncogenesis: Novel targets for anticancer therapies. *Journal of Clinical Oncology* 31, 1231–1238.
- Rotin, D., Kumar, S., 2009. Physiological functions of the HECT family of ubiquitin ligases. *Nature Reviews Molecular Cell Biology* 10, 398–409.
- Schulman, B.A., Harper, J.W., 2009. Ubiquitin-like protein activation by E1 enzymes: The apex for downstream signalling pathways. *Nature Reviews Molecular Cell Biology* 10, 319–331.
- Saeki, Y., Tanaka, K., 2012. Assembly and function of the proteasome. *Methods in Molecular Biology* 832, 315–337.
- Schwartz, A.L., Ciechanover, A., 2009. Targeting proteins for destruction by the ubiquitin system: Implications for human pathobiology. *Annual Review of Pharmacology and Toxicology* 49, 73–96.
- Striebel, F., Imkamp, F., Ozelik, D., Weber-Ban, E., 2014. Pupylation as a signal for proteasomal degradation in bacteria. *Biochimica et Biophysica Acta* 1843, 103–113.
- Tanaka, K., Matsuda, N., 2014. Proteostasis and neurodegeneration: The roles of proteasomal degradation and autophagy. *Biochimica et Biophysica Acta* 1843, 197–204.
- Ulrich, H.D., 2014. Two-way communications between ubiquitin-like modifiers and DNA. *Nature Structural & Molecular Biology* 21, 317–324.
- Varshavsky, A., 2012a. Three decades of studies to understand the functions of the ubiquitin family. *Methods in Molecular Biology* 832, 1–11.
- Varshavsky, A., 2012b. The ubiquitin system, an immense realm. *Annual Review of Biochemistry* 81, 167–176.
- Weathington, N.M., Mallampalli, R.K., 2014. Emerging therapies targeting the ubiquitin proteasome system in cancer. *Journal of Clinical Investigation* 124, 6–12.
- Wright, D.E., Wang, C.Y., Kao, C.F., 2012. Histone ubiquitylation and chromatin dynamics. *Frontiers in Bioscience (Landmark Ed)* 17, 1051–1078.
- Xie, Y., 2010. Feedback regulation of proteasome gene expression and its implications in cancer therapy. *Cancer Metastasis Reviews* 29, 687–693.
- Zhang, W., Sidhu, S.S., 2014. Development of inhibitors in the ubiquitination cascade. *FEBS Letters* 588, 356–367.

Further Reading

- Amm, I., Sommer, T., Wolf, D.H., 2014. Protein quality control and elimination of protein waste: The role of the ubiquitin-proteasome system. *Biochimica et Biophysica Acta* 1843, 182–196.
- Brownell, J.E., Sintchak, M.D., Gavin, J.M., *et al.*, 2010. Substrate-assisted inhibition of ubiquitin-like protein-activating enzymes: The NEDD8 E1 inhibitor MLN4924 forms a NEDD8-AMP mimetic in situ. *Molecular Cell* 37, 102–111.
- Ciechanover, A., Stanhill, A., 2014. The complexity of recognition of ubiquitinated substrates by the 26S proteasome. *Biochimica et Biophysica Acta* 1843, 86–96.
- Dohmen, R.J., Willers, I., Marques, A.J., 2007. Biting the hand that feeds: Rpn4-dependent feedback regulation of proteasome function. *Biochimica et Biophysica Acta* 1773, 1599–1604.

Endoplasmic Reticulum-Associated Degradation and Protein Quality Control

LF Zacchi and JJ Caramelo, Fundación Instituto Leloir, Buenos Aires, Argentina

AA McCracken, University of Nevada, Reno, NV, USA

JL Brodsky, University of Pittsburgh, Pittsburgh, PA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Endoplasmic reticulum-associated degradation

(ERAD) A collection of related processes that participate in the recognition and degradation of proteins localized in the endoplasmic reticulum. The ERAD of all substrates generally requires for basic steps: substrates are recognized (and brought to the ER membrane, if luminal), removed from the ER, ubiquitinated, and degraded by the cytosolic proteasome.

Native fold The properly folded and biologically active conformation of a protein, usually it is also the most thermodynamically stable.

Proteostasis (protein homeostasis) Refers to the maintenance of optimal protein biological activity and the cellular mechanisms that are responsible for this. Environmental alterations can have a profound impact on the native structure of proteins and on the folding capacity of the cells. Cells have developed a network of molecular chaperones, enzymes, and signal transducing components that participate in the synthesis, folding, quality control, and degradation of all cellular proteins, and that are able to rapidly respond to environmental stress, such that protein's biological function is maintained.

Quality control code The collection of signals displayed by a polypeptide (encoded or introduced posttranslationally) that dictate its movement through the secretory pathway and its interaction with the folding and quality control machineries. These include the ER targeting sequences, the *N*-glycosylation sequons, the exposed hydrophobic patches, the ER retention/exit signals, and the glycan and polyubiquitin chains, among others.

Retro-translocation Also known as dislocation, is a term that refers to the processes by which ER proteins are extracted from the ER into the cytosol.

Sequon A sequence of amino acids *N*-X-S/T (Asparagine-X-Serine or Threonine) found in a polypeptide, where X cannot be Proline, to which *N*-linked glycans are appended by the oligosaccharyltransferase.

Translocation Movement of a polypeptide across the ER membrane and into the ER lumen.

Translocon A proteinaceous channel in the ER membrane through which secretory polypeptides enter the ER. It is composed of a Sec61 heterotrimeric core and several associated components.

Introduction

Endoplasmic reticulum-associated degradation (ERAD) is the process that refers to the identification of proteins localized in the endoplasmic reticulum (ER) and their cytosolic destruction. One of the most critical tasks of ERAD is to aid in the selective degradation of misfolded proteins that enter the secretory pathway. If these misfolded proteins are not eliminated, they may accumulate in the ER, limiting its folding capacity by sequestering ER chaperones or by aggregating. Further, the accumulation of misfolded proteins in the ER can trigger a cellular stress response that can lead to apoptosis if unmitigated (Fribley *et al.*, 2009). Thus, the carefully controlled intracellular destruction of proteins by ERAD helps prevent the toxicity associated with aberrantly folded secretory proteins and their potential detrimental effect on cellular homeostasis.

Many human diseases are associated with the ERAD pathway. Some of these diseases occur due to mutations in secretory proteins that result in protein misfolding and that turn them into ERAD substrates. One disease in this class is Gaucher disease, a lysosomal storage disorder (Ron and Horowitz, 2005; Futerman and van Meer, 2004). In other cases the ERAD machinery can be too efficient and prematurely eliminate proteins that would eventually be able to fold. For example,

the slowly folding $\Delta F508$ variant of the cystic fibrosis transmembrane conductance regulator (CFTR) is prematurely degraded by ERAD, leading to cystic fibrosis (Jensen *et al.*, 1995; Ward *et al.*, 1995). Other human diseases are associated with mutations in the ERAD or the quality control machinery themselves. For example, mutations that affect *N*-linked glycosylation (an ER posttranslational modification linked to quality control and ERAD) can cause a variety of symptoms, including dysmorphism, encephalopathy, and organ disorders (Imbach *et al.*, 1999; Freeze *et al.*, 2014). Together, a growing number of human diseases, including neurological, respiratory, cardiovascular, and liver diseases, among many others (Guerriero and Brodsky, 2012; Jucker and Walker, 2013) are linked to ERAD and protein quality control in the secretory pathway.

The function of the secretory pathway was elucidated in the 1970s and early 1980s. This pathway can be grossly outlined as a port of entry (the ER), intermediate entities (the Golgi complex and vesicles), and a port of exit (the plasma membrane or the extracellular space). However, as we discuss below, the secretory organelles and vesicular intermediates do not simply provide a route for proteins out of the cell or to become embedded within lipid bilayers in the organelles or plasma membrane. Instead, these compartments play a critical role in preparing secretory proteins for export and in their

quality control. Research on this topic began with Palade's ground-breaking use of radioisotopes to outline the pathway (Palade, 1975), Blobel's innovative development of a cell-free system to uncover the first secretion signals and receptors (Blobel and Dobberstein, 1975), and Schekman's elegant application of yeast genetics to characterize components that constitute the secretory pathway (Novick *et al.*, 1980). By the mid 1980s, many research groups focused on determining how specific components mediate the selective versus non-selective trafficking of proteins through the ER (Pelham, 1989; Pfeffer and Rothman, 1987; McCracken and Kruse, 1989). Increasing evidence indicated that mutated proteins of medical importance accumulated in the ER (Hurtley and Helenius, 1989; McCracken *et al.*, 1989), and that mutated proteins that normally pass through the ER were apparently turned-over (Cheng *et al.*, 1990; Needham and Brodsky, 2013). It was soon clear that mutated ER proteins were degraded (McCracken and Kruse, 1993; Finger *et al.*, 1993; Hampton and Rine, 1994; Klausner and Sitia, 1990). Because lysosomal/vacuolar enzymes were dispensable for this event, it was tempting to speculate that degradation took place within the ER. However, there was no evidence of a proteolytic quality control system housed within the ER. Because of the uncertainty surrounding the nature of the protease, we named this process ER-associated degradation, or ERAD (McCracken and Brodsky, 1996).

Through the use of yeast genetics and mammalian cell systems, and by employing both *in vivo* and *in vitro* tools, compelling evidence emerged that the cytosolic proteasome provides the proteolytic activity for ERAD (Werner *et al.*, 1996; Hiller *et al.*, 1996; McCracken *et al.*, 1996; Jensen *et al.*, 1995; Ward *et al.*, 1995; Wiertz *et al.*, 1996a; Sommer and Jentsch, 1993). This discovery came as a complete surprise. Although integral membrane proteins in the ER could access a cytosolic protease, it was unexpected that soluble secreted proteins could also be degraded by the proteasome. The solution to this problem was that soluble proteins were recognized within the ER and then returned to the cytosol via an event termed 'retro-translocation' or 'dislocation.' Over the ensuing years, a significant effort has been devoted to understanding how diverse ERAD substrates are selected, retro-translocated, and targeted to the proteasome.

In the end, it was evident that ERAD represented 'an unconventional route' (retro-translocation from the ER) to a 'familiar fate' (misfolded protein degradation by the cytosolic proteasome) (Werner *et al.*, 1996; Hiller *et al.*, 1996). Because defects in the ERAD pathway exhibited synthetic, negative effects when ER stress response pathways were disabled (Travers *et al.*, 2000), it also became clear that ERAD was one of the two critical components that maintain ER homeostasis, the other being the unfolded protein response (UPR) (Walter and Ron, 2011). Together, ERAD and the UPR minimize the potentially detrimental effects of protein misfolding in the secretory pathway. Nevertheless, the ERAD pathway and ERAD-like mechanisms also regulate the stability (and thus activity) of wild type, functional proteins in the secretory pathway (Chen *et al.*, 2011a; Hampton, 2002; Lemberg, 2013) and can also be hijacked by pathogens (Noack *et al.*, 2014).

In this article, we will provide an overview of the processes that shape a secretory protein as it travels through the early secretory pathway. During this journey, secretory proteins

display signals that are scanned by numerous binding partners. These signals not only dictate if a protein will enter the ER, but also if it should be posttranslationally modified and transported to later compartments in the secretory pathway, or if it should instead be degraded. Deciphering this 'quality control code' is a critical task dutifully performed by ER protein quality control mechanisms and ERAD. It should be noted that most of the information on the function of the secretory pathway and quality control mechanisms has been obtained through the use of both the yeast *Saccharomyces cerevisiae* and mammalian cells. As expected, the mammalian system is more complex and thus there are more enzymes and chaperones involved in each process, and the ERAD-L, ERAD-C, and ERAD-M processes, which were defined in yeast, are poorly defined in mammalian cells. However, general processes are highly conserved and orthologs of the most relevant genes involved in these processes are found in both organisms. Here we discuss both systems, provide the name of the orthologs when available, and point out the differences between both models when relevant.

The Manufacturing of a Secretory Protein

Targeting and Entering the ER: The Beginning of a Perilous Journey

The first challenge that secretory proteins must overcome on their way through the secretory pathway is finding and then inserting into the ER. To solve the first problem, secretory proteins possess signals that target them to the ER. These signals influence not only if the protein will enter the ER using co- or posttranslational mechanisms, but also what chaperones will protect and guide them to the ER, and in what folding state they will reach this organelle (Cross *et al.*, 2009). The strength of the signal depends on its length, hydrophobicity, topology, or amino acid composition, and each signal is recognized by specific protein complexes and chaperones (Akopian *et al.*, 2013; Ng *et al.*, 1996; Hegde and Bernstein, 2006). The majority of proteins are targeted to the ER via an N-terminal signal sequence (SS) (Figure 1(a)). The most prevalent ER targeting mechanism is the signal recognition particle (SRP) pathway, a co-translational mRNA/protein partitioning mechanism that depends on the secretory protein's SS (Blobel *et al.*, 1979; Akopian *et al.*, 2013; Noriega *et al.*, 2014). In this pathway, soon after the SS is synthesized by the ribosome it is bound by the SRP. Binding stalls the ribosome, which prevents further translation (Siegel and Walter, 1986; Walter *et al.*, 1981) (Figure 1(a)). The SRP tethers the ribosome-nascent chain complex (RNC) to the ER through interaction with the SRP receptor at the ER surface (Gilmore *et al.*, 1982). There, the RNC is transferred to the Sec61 translocon (see below). The SRP is then recycled, and mRNA translation continues while the nascent polypeptide is translocated (i.e. transferred across the membrane) into the ER lumen, or is inserted into the ER membrane (Crowley *et al.*, 1994; Jungnickel and Rapoport, 1995) (Figures 1(a) and 1(b)). However, recent evidence demonstrated that ~40% of secretory proteins in yeast reach the ER in an SRP-independent manner (Ast *et al.*, 2013). Most of these proteins contain a tail

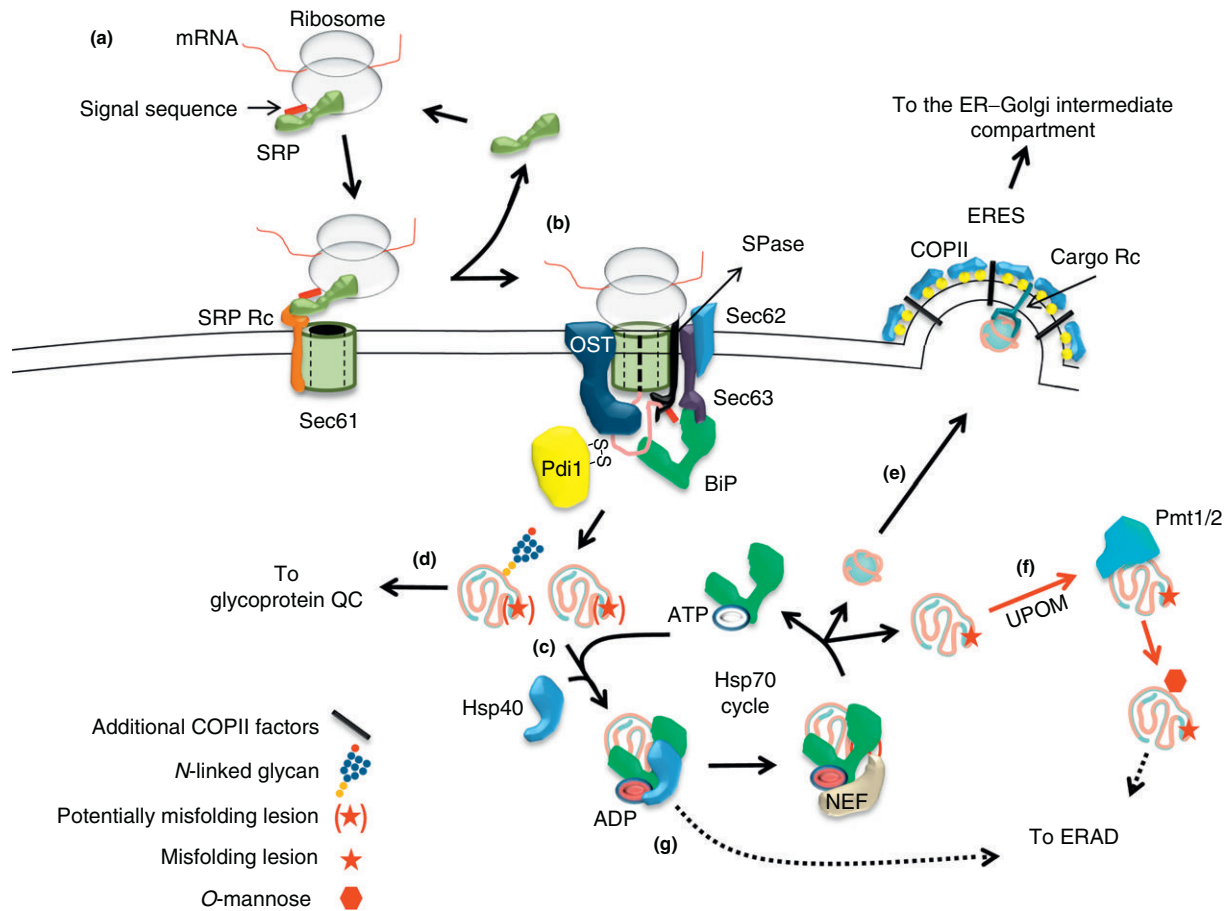


Figure 1 Translocation, folding, and ER exit of a secretory protein. (a) The signal recognition particle (SRP) recognizes the signal sequence (SS) emerging from the ribosome, and delivers the ribosome-nascent chain complex to the SRP receptor (SRP Rc) near the Sec61 translocation channel. The SRP is then recycled. (b) The polypeptide is translocated into the ER lumen and posttranslationally modified: the SS is cleaved by the signal peptidase (SPase), N-X-S/T sequons are *N*-glycosylated by the oligosaccharyltransferase (OST), and PDI facilitates disulfide bond formation. The Hsp70 chaperone BiP protects the incoming polypeptide from aggregation and is required for translocation. BiP is tethered to the translocon through its interaction with the transmembrane protein Sec63, and this interaction is required for proper protein translocation. (c) The folding of glycosylated and unglycosylated polypeptides is surveyed by BiP and its co-chaperones. ATP-bound BiP accepts substrates that may be delivered by Hsp40s co-chaperones. BiP binding to the Hsp40s and the substrate stimulates ATP-hydrolysis, forming a high affinity ADP-BiP-substrate complex. Nucleotide exchange factors (NEFs) accelerate ADP release. This reduces Hsp70's affinity for the substrate, facilitates substrate release, and recycles ATP-BiP. At this point the substrate may still not be properly folded, requiring additional cycles of Hsp70 binding and release. (d) In mammalian cells, glycoproteins are delivered to the CNX/CRT/UGGT quality control machinery (glycoprotein QC) (also see [Figure 2\(b\)](#)). Folded substrates exit the ER at ER exit sites (ERES) (known as transitional ER in yeast), where they are loaded into COPII vesicles with the help of cargo receptors (Cargo Rc). (f) In yeast, misfolded unglycosylated substrates can be *O*-mannosylated by Pmt1 and Pmt2 and targeted for ERAD (UPOM). (g) Alternatively, misfolded substrates may be targeted to ERAD due to prolonged chaperone interactions independently of the glycan.

anchor (TA) sequence (a single C-terminal transmembrane domain) or a glycosylphosphatidylinositol (GPI)-anchoring sequence. TA and GPI proteins are targeted to the ER via the guided entry of TA proteins (GET) pathway ([Ast et al., 2013](#); [Hegde and Keenan, 2011](#); [Schuldiner et al., 2008](#); [Wang et al., 2014](#)). Small soluble proteins can also be posttranslationally translocated into the ER, but their ER targeting mechanism remains poorly elucidated ([Deshaies et al., 1988](#); [Panzner et al., 1995](#)). Further, some secretory proteins reach the ER through targeting sequences in their mRNAs ([Weis et al., 2013](#); [Kraut-Cohen et al., 2013](#)).

The port of entry to the ER lumen for most secretory proteins which are translocated co- or posttranslationally is a

proteinaceous pore in the membrane called the Sec61 translocon ([Figure 1\(b\)](#)). The translocon is an oligomeric protein-conducting channel composed of a conserved Sec61-containing heterotrimeric core and associated components ([Rapoport, 2007](#)). The associated components perform a variety of functions. For example, the ribosome and the Hsp70 BiP (via the Sec63 anchor at the ER membrane) ratchets proteins into the ER (since the pore itself is passive); the signal peptidase complex removes the SS from the translocating nascent chain; and the oligosaccharyltransferase (OST) *N*-glycosylates substrates at specific motifs (see below) ([Figure 1\(b\)](#)). The yeast translocon also associates with protein mannosyltransferases which *O*-mannosylate translocating polypeptides ([Loibl et al., 2014](#)).

The yeast posttranslational translocon also contains the transmembrane proteins Sec62, Sec63, Sec71, and Sec72 (Deshaies and Schekman, 1989; Green *et al.*, 1992). In order to insert transmembrane segments into the lipid bilayer, the translocon must open laterally (Shao and Hegde, 2011). On the other hand, TA-proteins insert into the ER membrane independently of the translocon, using the yeast GET pathway or a homologous family of proteins in mammals (Hegde and Keenan, 2011). Given the variety of ER targeting signals and the structural diversity of secretory proteins, it has become clear that the ER targeting mechanisms and translocation pathways are more diverse than previously expected.

Secretory proteins must be maintained in a translocation/insertion competent state during and after translation. This is a challenging task, since these proteins contain aggregation-prone hydrophobic regions, including ER targeting signals (Hegde and Bernstein, 2006). Translocation competence is facilitated by dedicated molecular chaperones in the cytosol and the ER, which include members of the heat-shock protein (Hsp) 70 family and Hsp70 regulating co-factors, such as the Hsp40s and nucleotide exchange factors (NEFs) (Zimmermann *et al.*, 2006; Figure 1(c)). Hsp70 chaperones bind stretches of hydrophobic amino acids that are exposed in non-native proteins (Flynn *et al.*, 1991). Hsp70s bind and then release substrates in an ATP-dependent manner, thereby preventing aggregation (Hartl *et al.*, 2011; Figure 1(c), the Hsp70 cycle). Although chaperones can recognize any substrate presenting hydrophobic patches or segments, there is also client-specificity driven by cofactors. For example, the ubiquitous and abundant Hsp70s are regulated by a considerably larger number of Hsp40s and NEFs, which diversify and fine-tune the Hsp70's client choice (Kampinga and Craig, 2010). Chaperones play other critical roles during protein biogenesis and quality control, including: aiding *de novo* folding, re-folding, and oligomeric protein assembly; dissolving protein aggregates; and redirecting misfolded proteins for degradation. Chaperones can also chaperone other chaperones. For example, the ER resident Hsp70 BiP is required for the stability of the ER localized AAA+ ATPase torsinA, a protein involved in the human neurological disease Early Onset Torsion Dystonia (Zacchi *et al.*, 2014). Chaperones are also required for protein translocation. For example, BiP is required for both co- and posttranslational translocation (Brodsky *et al.*, 1995; Matlack *et al.*, 1999), possibly by regulating the activity of the translocon (Alder *et al.*, 2005; Figure 1(b)). Further, a specialized set of chaperone-like proteins localized in the ER act on the translocated polypeptide, including the lectins, which bind glycosylated (and in some cases non-glycosylated) proteins, and the protein disulfide isomerases (PDIs), which form and isomerize disulfide bonds (Buck *et al.*, 2007; Figure 1(b)). Together, the ER chaperones cooperate in the productive folding of secretory proteins and in preventing the formation of misfolded intermediates by recognizing and targeting these substrates for ERAD.

The N-Glycosylation Quality Control Pathway

Once in the ER, secretory proteins can be posttranslationally modified by lipidation, the formation of disulfide bonds, cleavage of the SS, and O- and N-linked glycosylation. Most secretory proteins are modified with high mannose-containing

glycans, a process carried out by the membrane-bound OST complex (Kelleher and Gilmore, 2006; Figure 1(b)). The biological roles of N-glycans are diverse. Bulky and hydrophilic sugar moieties can mediate protein-protein recognition, can stabilize some elements of secondary structure, and can improve protein solubility by stabilizing hydrophobic regions (Imperiali and Rickert, 1995). In most organisms the catalytic subunit of OST, STT3, transfers the oligosaccharide Glc₃Man₉GlcNAc₂ (where Glc, Man, and GlcNAc are glucose, mannose, and N-acetyl glucosamine, respectively (Figure 2(a))) from a dolichol-linked donor to the side chain of an Asn displayed within a sequon. The sequon is an Asn-X-Ser/Thr sequence, where X can be any residue except Pro. Mammals express two isoforms of STT3, STT3-A and STT3-B, which ensure a high degree of sequon occupancy (Ruiz-Canada *et al.*, 2009). While STT3A activity is mainly evident co-translocationally, STT3B is most active after a polypeptide is released from the Sec61 translocon. High mannose glycans are generally trimmed-down to the core glycan (Man₃GlcNAc₂) as they transit through the early secretory pathway. Golgi-resident glycosyltransferases complete the formation of the complex glycans usually found in mature glycoproteins. Thus, although all glycoproteins initially contain a common glycan, this structure is sharply diversified as glycoproteins traffic to their final destinations.

Why would the cell transfer a glycan that ends up being trimmed to the core, consuming energy in the process? The answer to this question emerged from the discovery that high mannose glycans are used as a platform to encode information on the folding status of glycoproteins. In other words, the glycan constitutes a signal in the 'quality control code,' and is the basis of the glycoprotein quality control system (Figures 1(d) and 2, also see below). The glycoprotein quality control process can be summarized as follows: N-glycans are modified by glycosyltransferases and glycosidases depending on the conformation and the age of the glycoprotein. Several lectins recognize the resulting N-glycans, prevent the premature exit from the ER of folding intermediates, aid in the folding process, and permit export to the Golgi or, alternatively, escort misfolded glycosylated proteins to the ERAD machinery (Figure 2(b)).

The glycoprotein quality control system provided the first demonstration that N-glycans shape a glycoprotein's fate (Caramelo and Parodi, 2008; Aebi *et al.*, 2010). Immediately after the Glc₃Man₉GlcNAc₂ N-glycan is appended onto a secretory protein, the outermost Glc residue is cleaved by glucosidase I (GI), generating Glc₂Man₉GlcNAc₂ (Figure 2). GI is localized at the ER membrane and associates with the translocon, and this ensures that the first deglycosylation step occurs rapidly (Dejgaard *et al.*, 2010). The second Glc residue is hydrolyzed by glucosidase II (GII), thus generating the monoglucosylated glycan Glc₁Man₉GlcNAc₂ (Figure 2(a)). Because it is the monoglucosylated form of the glycan the one that enters the CNX/CRT/UGGT quality control cycle (see below), the diglucosylated intermediate had not previously received much attention. However, Malectin, another ER resident lectin, specifically recognizes this diglucosylated moiety (Schallus *et al.*, 2008). Although the biological role of Malectin binding is unclear, Malectin association may regulate the flux of glycoproteins into subsequent processing steps (Chen *et al.*, 2011b). Regardless, the monoglucosylated moiety is then

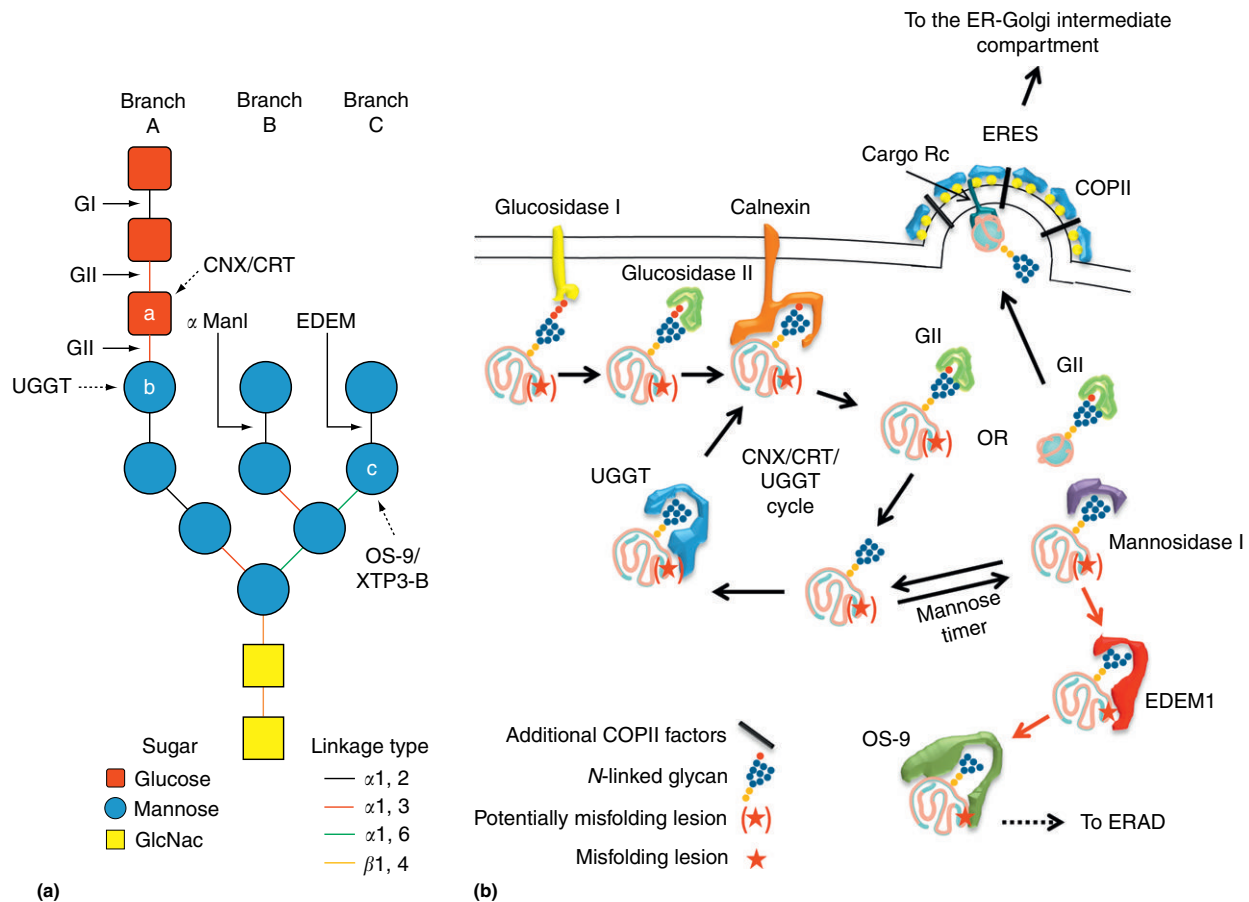


Figure 2 *N*-glycan quality control. (a) A schematic representation of the *N*-glycan transferred by OST. Residues relevant for quality control decisions are highlighted: (a) is required for CNX/CRT binding, (b) is required for UGGT binding, and (c) is required for OS-9/XTP3-B binding. The types of sugars and linkages are shown, as well as the sites where quality control enzymatic reactions take place (the linkages cleaved by glycosydases (GI and GII) and mannosydases (ER ManI and EDEMs)). (b) Glycoprotein folding is assisted by the CNX/CRT/UGGT cycle and, if unsuccessful, it is terminated by mannose trimming. *N*-glycans are processed by Glucosidase I and II (GI and GII), which produce a monoglucosylated *N*-glycan, Glc₁Man₉GlcNAc₂ that is a ligand for CNX/CRT. GII removes the last Glc residue, preventing CNX/CRT re-binding. If the glycoprotein is folded it can exit the ER. If it remains unfolded, UGGT reglucosylates it (residue 'a'). During these cycles, the misfolded Man₉GlcNAc₂ substrate may be de-mannosylated by ER α_{1,2} mannosidase I. Removal of 'b' irreversibly blocks UGGT binding, and therefore the substrate is now targeted for ERAD or ER exit. EDEM trims the substrate further, eventually exposing residue 'c', an α_{1,6} linked mannose, which is a ligand for OS-9 (Yos9) and XTP3-B and, thus, an ERAD substrate.

recognized by the ER resident lectins calreticulin (CRT) or its membrane-bound paralog calnexin (CNX) (Figure 2(a), residue 'a' is required for CNX/CRT binding) (Michalak *et al.*, 1999). CNX/CRT retain the monoglucosylated glycoprotein in the ER and improve folding efficiency by preventing aggregation and by recruiting ERp57, a PDI homolog that helps form correct disulfide bridges (Coe and Michalak, 2010). Eventually, the last Glc residue is cleaved by GII, releasing the glycoprotein from CNX and CRT (Figure 2). At this point, proteins displaying a native fold can enter COPII vesicles that bud from the ER membrane for delivery to the Golgi (Figure 1(e)). *N*-glycans also play a role during this event, since ER export via COPII vesicles in some cases is mediated by other lectins (ERGIC-53, VIP36, and VIPL) that recognize high-mannose structures (Hauri *et al.*, 2002). In contrast, unfolded or partially folded proteins are re-glucosylated by the enzyme UDP-Glc:glycoprotein glucosyltransferase (UGGT), which

adds back a Glc and regenerates the monoglucosylated substrate that interacts with CNX and CRT (Caramelo and Parodi, 2008; Figure 2). UGGT operates as a folding sensor that combines the activity of a glucosyltransferase with the specificity of a molecular chaperone. *In vitro* assays showed that UGGT preferentially recognizes exposed hydrophobic stretches displayed by advanced folding intermediates (Caramelo *et al.*, 2003; Ritter and Helenius, 2000; Taylor *et al.*, 2004; Taylor *et al.*, 2003), and cellular studies suggested that UGGT focuses its activity on advanced folding intermediates after release from BiP (Labriola *et al.*, 2011). Therefore, BiP and the CNX/CRT/UGGT cycle can cooperate during the folding of nascent proteins in the ER (Figures 1(c), 1(d), and 2). Interestingly, *S. cerevisiae* lacks a CRT ortholog as well as functional orthologs of CNX and UGGT: Cne1 (the CNX ortholog) plays a more limited role in ERAD and Kre5 (the UGGT ortholog) appears to lack enzymatic activity (Parlati *et al.*, 1995;

Fernandez *et al.*, 1994). Nevertheless, yeast can efficiently fold glycoproteins and dispose of misfolded *N*-glycosylated proteins, indicating that other chaperones or alternative mechanisms are in place in yeast that compensate for the lack of the CNX/CRT/UGGT cycle.

The cycle of deglycosylation and reglycosylation by GII and UGGT can continue until the protein folds or until it is marked for ERAD. Glycoprotein commitment to the ERAD pathway is modulated by the trimming of mannose residues (Hosokawa *et al.*, 2010a; Stigliano *et al.*, 2011). During this commitment step, ER α mannosidase (ER α ManI) (Mns1p in yeast) first cleaves the α 1,2-mannose residue located in the B branch of the *N*-glycan to generate Man₈GlcNAc₂ (Figure 2). The concentration of this enzyme in a specialized ER sub-compartment may allow further trimming of the *N*-glycans until Man₆GlcNAc₂ or Man₅GlcNAc₂ are generated (Avezov *et al.*, 2008). Mannose trimming can negatively impact GII activity, favoring the pro-folding CNX/CRT-ligand interaction (Stigliano *et al.*, 2011). Mannose trimming may also be carried out by EDEM1, EDEM2, or EDEM3, all of which are members of the EDEM (ER degradation-enhancing α -mannosidase-like protein) family (Htm1/Mnl1 in yeast) (Hirao *et al.*, 2006; Hosokawa *et al.*, 2010b; Molinari *et al.*, 2003). In fact, CNX can transfer glycoproteins directly to EDEM (Oda *et al.*, 2003). Indeed, recent evidence indicates that EDEM2 can also perform the initial de-mannosylation of the substrate from Man₉GlcNAc₂ to Man₈GlcNAc₂, which can be followed by further mannose trimming by EDEM3 (Ninagawa *et al.*, 2014). A point of no return in these reactions is the cleavage of the α 1,2-mannose residue located in the A branch of the *N*-glycan (Figure 2(a), residue 'b'), since it precludes UGGT action and definitively removes glycoproteins from the CNX/CRT/UGGT cycle. In contrast, cleavage of the α 1,2-mannose residue located in the C branch of the *N*-glycan exposes a terminal α 1,6-mannose residue (Figure 2(a), residue 'c'). The resulting *N*-glycan is recognized by the lectins OS-9 (Yos9 in yeast) and XTP3-B, which escort terminally misfolded glycoproteins to the ERAD machinery (Hosokawa *et al.*, 2010a; Bhamidipati *et al.*, 2005; Clerc *et al.*, 2009; Figure 2). Interestingly, OS-9, XTP3-B, and the EDEMs can also bind substrates outside of their lectin domain and, instead, they can use this domain to associate with glycosylated SEL1L (Hrd3 in yeast), the substrate acceptor component of the retrotranslocon (Figure 4(b); Aebi *et al.*, 2010; Christianson *et al.*, 2008; Cormier *et al.*, 2009). In this way, the lectin domain would bring substrates to the retrotranslocon instead of selecting ERAD substrates. This helps explain how these lectins are also involved in the degradation of unglycosylated substrates (see below).

A crucial, yet poorly understood event in this cycle is how this glycoprotein quality control system distinguishes terminally misfolded proteins from folding intermediates. This issue is particularly important for proteins that are difficult to fold, such as multidomain proteins with high β -sheet content or that display an intricate connectivity of disulfide bridges. The simplest model is that the time a protein resides in the ER plays the most important role in the ERAD decision. In comparison with GI and GII, the activity of ER ManI and EDEMs is low, which led to the 'mannose timer' hypothesis for glycoprotein quality control (Wu *et al.*, 2003; Figure 2(b)). According to this model, if a protein spends too much time

attempting to fold, mannose residues are eventually trimmed and the protein is sent to the ERAD machinery. However, this model fails to explain why native ER resident proteins avoid degradation. Perhaps a signal other than mannose trimming is required or, alternatively, misfolded glycoproteins are simply better substrates for mannose trimming.

ER Exit: The Continuation of a Productive Journey

Once posttranslationally modified and folded, proteins leave the ER to continue through the secretory pathway toward their next destination: the Golgi complex (Figure 1(e)). ER exit occurs in COPII vesicles that bud from specific ER regions, known as ER exit sites (ERES) in mammals, and the transitional ER in lower organisms (Orci *et al.*, 1991; Bannykh *et al.*, 1996). Specialized molecular machinery is tethered to these sites, where cargo is sorted and most misfolded proteins are selectively excluded from secretory vesicles. To exit the ER, cargo utilize spherical COPII-coated transport vesicles (Lord *et al.*, 2013), although there are examples of COPII-independent ER exit mechanisms that bypass the Golgi (Grieve and Rabouille, 2011). In the canonical pathway, liberated COPII vesicles are tethered to pre-Golgi compartments (the mammalian ER-Golgi intermediate compartment (ERGIC) or the yeast *cis*-Golgi) through interactions with receptors (Barlowe *et al.*, 1994; Beckers *et al.*, 1989; Eakle *et al.*, 1988; Sollner *et al.*, 1993; Wilson *et al.*, 1989).

Similar to ER targeting and glycan signals, the ER exit/retention signals encoded in the proteins are fundamental parts of the 'quality control code.' Proteins normally retained in the ER may exit through bulk flow in COPII vesicles but are retrieved via COPI coated vesicles (Aridor *et al.*, 1995; Letourneur *et al.*, 1994; Waters *et al.*, 1991). ER retrieval is dictated by retrieval sequences: the C-terminal K/HDEL sequence (primarily for soluble proteins), and the C-terminal di-Lys and N-terminal di-Arg motifs (primarily for membrane proteins) (Munro and Pelham, 1987; Jackson *et al.*, 1990; Michelsen *et al.*, 2005; Schutze *et al.*, 1994). There are also sequence-independent mechanisms for ER retention which depend on several factors, including the physicochemical characteristics of the transmembrane domains (e.g., length and hydrophobicity), the membrane lipid composition, and alterations in ER membrane curvature at ERES, which limit the partition of specific membrane-associated monotopic proteins (Vander Heyden *et al.*, 2011; Ronchi *et al.*, 2008; Sato *et al.*, 1996; Rayner and Pelham, 1997).

The sorting and selection of secretory proteins is carefully controlled. These substrates use a variety of signals to exit the ER. Some transmembrane proteins expose ER export signals to the cytosol that are recognized by COPII components. In turn, soluble secretory proteins can expose exit signals in the ER lumen, which are recognized by cargo receptors that associate with COPII components (Figure 1(e); Bonifacino and Click, 2004; Barlowe, 2003; Zanetti *et al.*, 2012). Although most misfolded proteins are retained in the ER and degraded via ERAD, some misfolded proteins are exported if ER exit sequences are displayed (Kincaid and Cooper, 2007). In fact, the degradation of select ERAD substrates may require ER-to-Golgi traffic (Haynes *et al.*, 2002; Taxis *et al.*, 2002), but the nature of

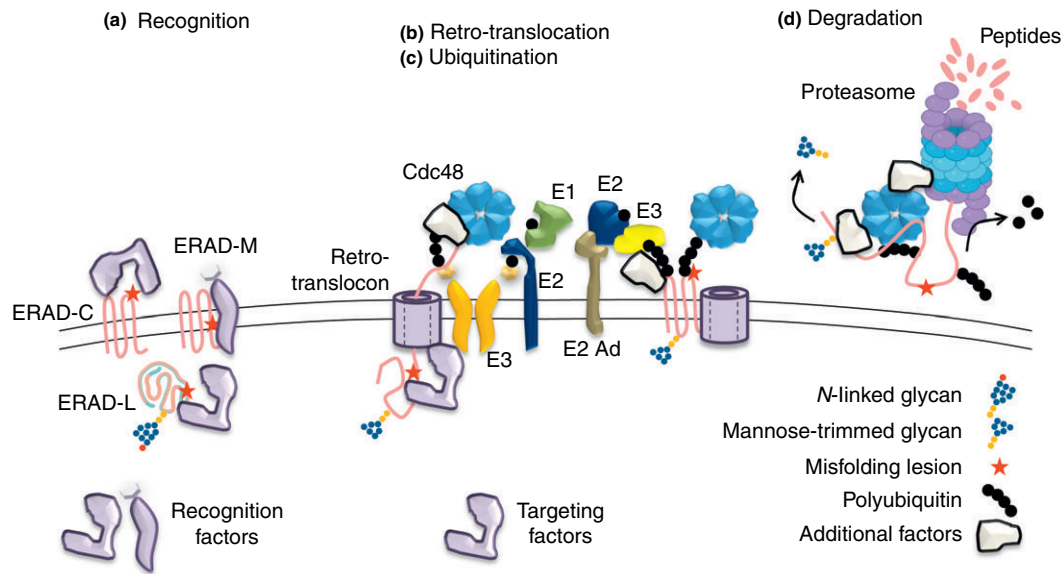


Figure 3 Basic steps during ER associated degradation in yeast. (a) ERAD substrates can expose lesions in the ER lumen (ERAD-L), ER membrane (ERAD-M), or cytosol (ERAD-C), and are recognized by distinct recognition factors (including chaperones and lectins), some of which deliver the substrate to the ER membrane for retro-translocation into the cytosol. (b) Retro-translocation is thought to occur through a channel in the ER membrane and is intimately connected with substrate ubiquitination. Some of the targeting factors may also help unfold the substrate. Ubiquitination occurs before or during substrate retro-translocation. (c) Ubiquitination requires the concerted action of E1, E2s (associated to the ER via a transmembrane domain or a membrane adapter (E2 Ad)) and ER membrane-resident E3 enzymes. Some cytosolic E2s and E3s also play a role in ERAD. The ubiquitinated substrate may tether the Cdc48 complex to the retro-translocation channel, where Cdc48 drives substrate retro-translocation. (d) A Cdc48-associated cofactor deglycosylates the substrate, and the Cdc48 complex delivers the substrate to the proteasome where it is de-ubiquitinated by proteasomal enzymes and degraded into shorter peptides.

this phenomenon is unclear. In addition, ER exit is used to regulate ERAD. For example, the lectins EDEM1 and OS-9 are exported from the ER in a COPII-independent manner and degraded via selective autophagy or in late endosomes/lysosomes, a process that down-regulates ERAD activity (Cali *et al.*, 2008; Zuber *et al.*, 2007; Park *et al.*, 2014; Bernasconi *et al.*, 2012). Unassembled subunits of oligomers can also be exported from the ER with EDEM1 in a COPII-independent manner and degraded via autophagy (Le Fourn *et al.*, 2013). Therefore, folding and export mechanisms work together in a fragile and strongly interdependent equilibrium to mediate ER protein quality control.

Unforgivable Errors: The ERADication of an Unfolded Protein

ERAD is a general term encompassing a variety of related mechanisms that lead to the proteasomal degradation of ER proteins. In all cases the ERAD pathway consists of four steps: substrate recognition, retro-translocation into the cytosol, ubiquitination, and proteasomal degradation (Figure 3). Each step is considered in the following sections.

Substrate Selection

The great majority of polypeptides enter the protein-rich environment in the ER in a non native state. Even under the best

conditions folding is inefficient, and polypeptides transition through folding intermediates before attaining their native conformations (Hartl *et al.*, 2011). During this time, hydrophobic, aggregation-prone sequences that will become buried in the native structure are exposed. Chaperones bind hydrophobic amino acid stretches, but are unable to distinguish a terminally misfolded protein from a folding intermediate. Similarly, although the CNX/CRT/UGGT cycle can recognize protein folding status, the decision to target misfolded glycoproteins for ERAD still depends on stochastic mannose trimming events. Consequently, many wild type proteins are probably targeted for degradation, even though considerable energy has been spent on the synthesis, translocation, post-translational modification, and chaperoning of a sluggish polypeptide/folding intermediate. However, because misfolded polypeptides pose a threat to cellular viability that is more costly than re-synthesizing a polypeptide, it is critical that any misfolded protein – even if it is a wild type protein – is eliminated by ERAD (Brodsky, 2012).

Location, location, location

The ERAD substrate recognition machinery varies depending on the subcellular location of the misfolding lesion in the protein. For example, integral membrane proteins can expose misfolded domains on the cytosolic side of the membrane, within the ER membrane, or in the ER lumen. In yeast, the processes that recognize these lesions are referred to as ERAD-C, ERAD-M, and ERAD-L, respectively, and rely on a specific set of chaperones and ubiquitin conjugating enzymes

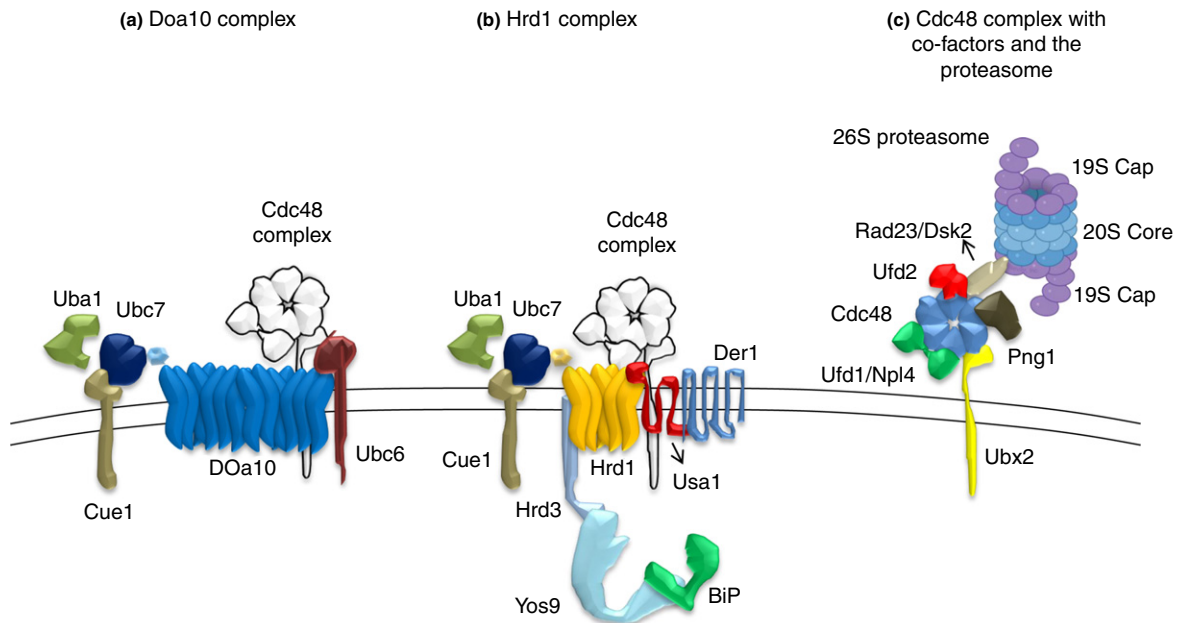


Figure 4 The yeast Doa10, Hrd1, and Cdc48 complexes, and the 26S proteasome. Schematic diagram of the (a) Doa10, (b) Hrd1, and (c) Cdc48 complexes and the 26S proteasome (please see the text for a description of the relevant components of each complex). The Cdc48 complex (with Ufd1, Npl4, and Ubx2) is a component of both the Doa10 and Hrd1 complexes and is shown in light gray in the background of (a) and (b). Shown in (c) are additional Cdc48 co-factors such as the E4 Ufd2, which lengthens the polyubiquitin chain, Dsk2 and Rad23, which target polyubiquitinated proteins to the proteasome, and the *N*-glycanase Png1. The 26S proteasome is shown, with the two 19S caps and the 20S core.

(Carvalho *et al.*, 2006; Kanehara *et al.*, 2010; Vashist and Ng, 2004; Huyer *et al.*, 2004; Denic *et al.*, 2006; Figure 3(a)). ERAD-C requires cytosolic chaperones and the Doa10 complex, whereas ERAD-L relies on luminal chaperones and requires the Hrd1 complex (Figures 4(a) and 4(b)). ERAD-M is less well understood, and only requires a subset of the Hrd1 complex components. These complexes are located in the ER membrane, and their main function is to ubiquitinate and extract ERAD substrates from the ER lumen (for soluble ERAD-L substrates) or membrane (for ERAD-M or -C substrates). Doa10 and Hrd1 are E3 ubiquitin ligases (see below), and they associate with additional factors that determine their client specificity and function and that aid in substrate extraction from the ER. Common co-factors include the Ubc7 E2 ubiquitin-conjugating enzyme and its membrane-anchoring partner and activator, Cue1 (Bagola *et al.*, 2013; Bazirgan and Hampton, 2008; Biederer *et al.*, 1997; Metzger *et al.*, 2013). The Cdc48 complex is also a common component (Figure 4(c)). The Cdc48 complex contains the cytosolic AAA ATPase Cdc48 (also known as p97 or VCP in mammals), two co-factors Ufd1 and Npl4, and a putative membrane-anchor, Ubx2 (Jarosch *et al.*, 2002; Neuber *et al.*, 2005; Schuberth and Buchberger, 2005; Ye *et al.*, 2003; Figure 4). One unique component in the Doa10 complex is the transmembrane E2 Ubc6 (Swanson *et al.*, 2001; Figure 4(a)). Unique components in the Hrd1 complex are the transmembrane proteins Der1, Usa1, and Hrd3, and the luminal lectin-like protein Yos9 (called Derlin-1, Herp, SEL1L, and OS-9 in mammals, respectively) (Figure 4(b)). These co-factors help in substrate retro-translocation (Der1) (Mehnert *et al.*, 2013; Greenblatt *et al.*, 2011), regulate Hrd1 activity (Usa1 and Hrd3) (Carroll and Hampton, 2010; Horn *et al.*, 2009; Kostova *et al.*, 2007;

Okuda-Shimizu and Hendershot, 2007), or are part of a dual recognition system for both glycosylated and unglycosylated ERAD substrates (Hrd3, Yos9, and Usa1) (Benitez *et al.*, 2011; Gauss *et al.*, 2006; Quan *et al.*, 2008; Stanley *et al.*, 2011). Hrd3 interacts with Yos9 and indirectly with BiP (through Yos9) and can therefore be considered a receptor for ERAD-L substrates. Although the core machinery is conserved from yeast to mammals, it is more difficult to distinguish ERAD-L, ERAD-C, and ERAD-M in mammalian cells (Vembar and Brodsky, 2008; Bernasconi *et al.*, 2010).

The fatty and sweet sides of destruction

What is the minimal signal that a misfolded protein needs to display for ERAD? As described above, the recognition of misfolded proteins in the lumen or even on the cytosolic side of the ER (for ERAD-C substrates) occurs through exposed hydrophobic stretches that would normally be buried in the native structure. These stretches are bound by chaperones. Prolonged chaperone interaction leads to ERAD targeting, potentially by bringing substrate and ubiquitin ligases together (Nakatsukasa *et al.*, 2008; Han *et al.*, 2007; Figure 1(g)). In the ER lumen, there is an additional signal for ERAD: the glycan (see above). However, only specific glycans within a substrate are required for ERAD selection, and they require the presence of an adjacent unfolded region, thereby functioning as a bipartite signal (Kostova and Wolf, 2005; Spear and Ng, 2005; Xie *et al.*, 2009).

Although the majority of secretory proteins contain *N*-linked glycans, some ERAD substrates are not glycosylated (Apweiler *et al.*, 1999). How are these proteins targeted for ERAD? Cell-free assays developed using yeast were critical to shed light on some of the required components (McCracken

and Brodsky, 1996; Werner *et al.*, 1996). These studies identified BiP and two BiP-associated Hsp40s as factors required for the ERAD of a non-glycosylated substrate (Brodsky *et al.*, 1999; Nishikawa *et al.*, 2001). Interestingly, Cne1 (CNX in mammals) was also required (McCracken and Brodsky, 1996). In fact, many factors that recognize *N*-linked glycosylated proteins are required for the degradation of unglycosylated substrates, including EDEM2, OS-9/Yos9, XTP3-B, and SEL1L/Hrd3 (Shenkman *et al.*, 2013; Gauss *et al.*, 2006; Cormier *et al.*, 2009; Ihara *et al.*, 1999; Tang *et al.*, 2014). Other lectins, such as EDEM1 and EDEM3, are able to interact with unglycosylated substrates but appear to play no role in their degradation (Tang *et al.*, 2014). In contrast, the yeast EDEM1 homolog, Htm1/Mnl1, does not appear to act on unglycosylated substrates, perhaps because an *N*-terminal substrate binding region is absent in Htm1 (Marin *et al.*, 2012). Similarly, Herp (Usa1 in yeast) is required for the Hrd1-dependent degradation of both glycosylated and unglycosylated substrates. Herp directly associates with Hrd1, Derlin-1, and the proteasome (Okuda-Shimizu and Hendershot, 2007; Schulze *et al.*, 2005; Huang *et al.*, 2013), and acts on CNX-independent but BiP and ERdj5 dependent unglycosylated substrates (Ushioda *et al.*, 2013) (Figure 1(g)). The ability of these lectins to bind to a variety of substrates independently of their glycan content underscores the complexity of the system and the inherent difficulties in dissecting the role these factors play in the degradation of ERAD substrates.

Similar to the mannose timer model for misfolded glycoproteins, the folding of unglycosylated proteins can also time out, at least in yeast. This timing mechanism, termed Unfolded Protein *O*-mannosylation (UPOM), requires an enzymatic reaction that adds a mannose to a Ser or Thr side chain (Xu *et al.*, 2013) (Figure 1(f)). The enzymes required for this reaction are Pmt1 and Pmt2. *O*-mannosylation terminates folding, reduces BiP binding, and helps the misfolded substrate remain soluble for the cytosolic proteasome-mediated degradation (Nakatsukasa *et al.*, 2004). It is unclear whether the folding rate or intrinsic properties of folding intermediates determine when a substrate will become *O*-mannosylated, and further studies are needed to determine the generality of this pathway. Besides playing a role during ERAD, *O*-mannosylation has also been postulated to act upstream of the ERAD decision, helping folding intermediates escape ERAD and remain soluble until they find the subcellular location or binding partner required to attain their native fold (Kleizen and Braakman, 2013).

Molecular chaperones: Critical partners during substrate selection and delivery

Chaperones are intimately involved in the decision of whether or not to degrade a substrate. One of the most important chaperone families involved in ERAD substrate recognition is the family of luminal and cytosolic Hsp70s. BiP/Grp78 (also known as Kar2 in yeast) is the only ER luminal Hsp70, while there are multiple constitutively expressed Hsc70s and stress-inducible cytosolic Hsp70s, especially in higher eukaryotes (Kabani and Martineau, 2008; Kampina and Craig, 2010). BiP binding is not only critical for ERAD-L substrate recognition, but also for maintaining ERAD substrate solubility and delivery to ubiquitin ligases (Nishikawa *et al.*, 2001; Nishikawa *et al.*, 2005). Other chaperones and chaperone-like

proteins that participate in ERAD-L substrate recognition or delivery to the retro-translocation/ubiquitination machinery include the PDIs, OS-9/Yos9, Hrd1, and Herp/Hrd3, and the mammalian CNX, UGGT, Grp94 (an ER Hsp90), and chaperone/disulfide reductase ERdj5 (Christianson *et al.*, 2008). ERAD-M substrates can be directly recognized by Hrd1, obviating the requirement for substrate recognition chaperones (Sato *et al.*, 2009). ERAD-C substrates are recognized and targeted for degradation by cytosolic factors, including Hsp70/Hsc70, Hsp40, Hsp26, and Hsp42 (Huyer *et al.*, 2004; Ahner *et al.*, 2007; Youker *et al.*, 2004). In the absence of chaperone function – and in some diseases (Guerriero and Brodsky, 2012) – ERAD substrates can aggregate and must be eliminated through other pathways. Alternative degradation pathways include autophagy and secretory pathway delivery to the vacuole/lysosome (Kruse *et al.*, 2006; Kario *et al.*, 2011; Kroeger *et al.*, 2009; Johnston *et al.*, 1998; Houck *et al.*, 2014).

Getting Back Out: The Retro-Translocation of Misfolded Substrates

Since the proteasome resides in the cytosol, luminal ERAD substrates must be moved into the cytosol for degradation and integral membrane proteins must be extracted from the lipid bilayer of the ER. This process is called retro-translocation or dislocation, and involves the following events: ER membrane targeting (of soluble substrates), enzymatic modification of the substrate with ubiquitin, and energy-dependent extraction of the substrate from the ER lumen/membrane.

Prior or during retro-translocation, some substrates are unfolded through the action of PDIs or ERdj5 (Tsai *et al.*, 2002; Ushioda *et al.*, 2008; Molinari *et al.*, 2002; Tsai *et al.*, 2001), *N*-glycoproteins are deglycosylated by the cytosolic PNGase enzyme (Png1 in yeast) (Suzuki *et al.*, 2000; Hirsch *et al.*, 2003), and most substrates in yeast are ubiquitinated by Doa10 and Hrd1 (Figures 3 and 4), although mammalian cells have a significantly larger set of ubiquitin ligases. The energy for retro-translocation is usually provided by the ATP hydrolyzing hexameric AAA ATPase, Cdc48 (see above; Wolf and Stolz, 2012; Carlson *et al.*, 2006). The Cdc48 complex binds substrates *en route* to the proteasome and participates in their dislocation (Ye *et al.*, 2003; Elkabatz *et al.*, 2004). However, Cdc48 is not always required for retro-translocation (Kothe *et al.*, 2005). In some cases, the AAA ATPases embedded in the 19S cap of the proteasome can provide the energy for retro-translocation (Lee *et al.*, 2004).

One of the most important questions in the field refers to the identity of the retro-translocation channel. Based on their topologies, soluble and integral membrane proteins present distinct challenges for dislocation. Furthermore, these proteins may contain folded sub-domains and posttranslational modifications like glycosylation or disulfide bridges. Therefore, the putative retro-translocation channel must accommodate bulky structures. In fact, there is evidence that some soluble luminal proteins can be retro-translocated while fully folded (Tirosh *et al.*, 2003; Olzmann *et al.*, 2013). Then, how are misfolded substrates exported to the cytosol? One attractive idea is that the Sec61 translocon may function as the pore. There is one example in which a soluble protein

(Apolipoprotein B) remains associated with Sec61 until either co-translational translocation finishes or retro-translocation occurs (Brodsky and Fisher, 2008). The interaction of Sec61 with ERAD components and the proteasome, and its requirement for the retro-translocation of several substrates further support this idea (Plempner *et al.*, 1999; Scott and Schekman, 2008; Wiertz *et al.*, 1996b; Tretter *et al.*, 2013). However, the small diameter of the pore may preclude the retro-translocation of glycosylated proteins and those that retain native structure (Van den Berg *et al.*, 2004). Other transmembrane proteins have been suggested to act as retro-translocation channels for soluble substrates, including Der1, Hrd1, Doa10, and the mammalian gp78 (Mehnert *et al.*, 2013; Wahlman *et al.*, 2007; Carvalho *et al.*, 2010; Bernardi *et al.*, 2010; Swanson *et al.*, 2001). Supporting a role for Hrd1 as the retro-translocator, recent evidence suggests a model in which luminal substrates traverse the ER membrane by inserting polypeptide loops in the channel formed by the Hrd1 transmembrane domains. Formation of these retro-translocation intermediates depends not only on known ERAD luminal targeting factors (Yos9, Hrd3, and Der1), but also, interestingly, on the function of cytosolic factors (including the Hrd1 ubiquitin ligase activity and the Cdc48 complex) (Carvalho *et al.*, 2010). However, mutations in any of the putative retro-translocation channels prevent the ERAD of specific and some common substrates, but none prevents the retro-translocation of all substrates (Hebert *et al.*, 2010; Swanson *et al.*, 2001). Therefore, retro-translocation channels may be functionally redundant. Alternatively, none of these candidates may reflect the sole exit strategy. Instead, misfolded proteins could exit the ER through, for example, lipid droplets (Ploegh, 2007), although lipid droplet formation appeared dispensable for ERAD in yeast (Olzmann and Kopito, 2011).

The mechanism underlying the retro-translocation of integral membrane proteins is similarly mysterious (Thibault and Ng, 2012; Ismail *et al.*, 2006). Some transmembrane substrates are fully extracted from the membrane in a Cdc48/p97-dependent manner, and this step can be uncoupled from degradation (Nakatsukasa *et al.*, 2008; Garza *et al.*, 2009; Lechner *et al.*, 2009). Other substrates remain membrane-integrated, but proteasome degradation still requires Cdc48/p97 and ubiquitination (Ikeda *et al.*, 2009). The channel – if it exists – that permits membrane protein extraction is unknown. Overall, much work remains to be done to understand the mechanism of substrate retro-translocation.

Signaling for Degradation: the Ubiquitin Mark

Protein ubiquitination occurs when ubiquitin, which is a highly conserved 76-amino acids protein, is covalently attached to a target protein. Ubiquitin is most commonly appended onto the ϵ -amino group of a Lys, forming an isopeptide bond. There are different types of ubiquitination – mono, multi-mono, and polyubiquitination – depending on how many ubiquitins are attached in tandem onto a substrate. Since each ubiquitin contains eight residues onto which additional ubiquitins can be conjugated (7 Lys and the N-terminus), polyubiquitin chains can attain a high degree of structural diversity (Winget and Mayor, 2010). This

complicated code is used to regulate a plethora of cellular activities, including ERAD.

Ubiquitination is the ultimate signal of the ‘quality control code.’ Proteasomal degradation requires a polyubiquitin chain of at least four molecules linked via the C-terminal carboxylate of one ubiquitin and onto a Lys at position 48 in the next molecule (Chau *et al.*, 1989; Flierman *et al.*, 2003). However, the proteasome is promiscuous and can degrade substrates ubiquitinated on other residues (e.g., Cys, Ser, and Thr) or with different chain types, and even those that are not ubiquitinated (Kravtsova-Ivantsiv and Ciechanover, 2012). Substrate ubiquitination requires the concerted action of three enzymes: a ubiquitin-activating enzyme (E1), a ubiquitin conjugating enzyme (E2), and a ubiquitin ligase (E3) (Figure 3(c)). There are also ubiquitin-chain-extension enzymes (E4) (e.g., Ufd2) that play a role in the ERAD of some substrates (Richly *et al.*, 2005; Nakatsukasa *et al.*, 2008; Figure 4(c)). The substrate specificity of the ubiquitination reaction is achieved through the combination of the function of the E2s, E3s, and other factors, such as chaperones (Christianson and Ye, 2014). As described above, the Doa10 and Hrd1 E3 ligases are required for the ERAD of most substrates in yeast. Both of these are RING finger E3 ligases, and bring substrates in proximity to ubiquitin-conjugated E2s (Swanson *et al.*, 2001; Bays *et al.*, 2001; Deak and Wolf, 2001). In turn, the E2 receives ubiquitin from the E1. Yeast E2s implicated in ERAD include Ubc7, Ubc6, and Ubc1 (McGrath *et al.*, 1991; Kostova *et al.*, 2007). ERAD can occasionally occur in the absence of Doa10 and Hrd1 function, and in this case other membrane-bound or cytosolic E3s (Ubr1 and Rsp5) function instead or contribute to degradation (Stolz *et al.*, 2013; Haynes *et al.*, 2002). In mammals, there is a considerably larger number of ERAD-associated E3s, including the ER residents Synoviolin/HsHrd1, TEB4, and gp78, and the cytosolic proteins Parkin and CHIP (Kostova *et al.*, 2007; Claessen *et al.*, 2012). Substrates can be ubiquitinated by multiple E3s, which act redundantly, simultaneously, or sequentially. E3s also self-ubiquitinate and ubiquitinate each other, a process that may fine tune ERAD (Olzmann *et al.*, 2013; Bernasconi *et al.*, 2013). Interestingly, the ubiquitination machinery and de-ubiquitinating enzymes also regulate the retro-translocation and degradation of non-ubiquitinated substrates, probably by acting on the ERAD machinery (Bernardi *et al.*, 2013). Thus, protein ubiquitination is a complex but universal signal for degradation that directly and indirectly impacts the fate of ERAD substrates.

At the End, the Proteasome

ERAD substrate elimination requires the 26S proteasome, which is a large (~2500 kDa), multimeric complex. The 26S proteasome is composed of a core, the 20S particle where proteins are degraded by unique ATP-dependent proteases, and the ‘caps,’ which are 19S regulatory particles that participate in substrate selection, de-ubiquitination, and unfolding (Schmidt and Finley, 2013; Figure 4(c)). Due to the presence of the 19S cap, entrance to the core is regulated, which prevents indiscriminate protein degradation in the cell. AAA+ ATPases located at the base of the caps regulate and drive substrate delivery into the 20S core (Smith *et al.*, 2007).

Nearly all ERAD substrates require the Cdc48 complex for delivery to the proteasome. In yeast, Cdc48 may hand substrates over to the proteasome via two adapter proteins, Rad23 and Dsk2 (Chen and Madura, 2002; Figure 4(c)), although Cdc48 also binds the proteasome (Barthelme and Sauer, 2012; Verma *et al.*, 2000). However, the proteasome itself binds ubiquitinated substrates by virtue of 19S-associated receptors, and can interact with the Hrd1 complex in the absence of functional Cdc48 (Nakatsukasa *et al.*, 2013). Further, there exists a proteasome subpopulation already localized at the ER membrane, which may be dedicated for ERAD (Palmer *et al.*, 1996). Therefore, coupling factors that act as intermediaries between Cdc48 and the proteasome may be dispensable for the degradation of at least some ERAD substrates. Prior to entry into the 20S core, ubiquitin is stripped from the substrate through the action of proteasome-associated deubiquitinases (Figure 3(d)). This process is required for delivery of the substrate into the 20S core chamber and is also critical to recycle ubiquitin (Smith *et al.*, 2007).

Concluding Remarks

Maintaining proteostasis is critical for cell survival. Since the discovery of the secretory pathway, a surprisingly large number of complex and interconnected cellular processes have been identified that help facilitate the quality control and the traffic of secretory proteins. Some of these processes, very recently uncovered, include quality control processes that occur in the cytosol, such as prERAD and the activity of surveying factors for mis-targeted TA proteins (Ast *et al.*, 2014; Okreglak and Walter, 2014), as well as quality control processes that occur at the ER, such as selective autophagy and ERQC autophagy (Houck *et al.*, 2014; Park *et al.*, 2014). Intense research in these areas will surely bring to light numerous novel exciting pieces to complete the protein quality control puzzle.

Even after more than two decades since ERAD was discovered many fundamental questions remain unanswered: How does the molecular machinery 'decide' on the folding status of a protein? How are the different aspects of the 'quality control code' decoded by the chaperones and lectins? Which proteins form the retro-translocation channel? How are transmembrane proteins dislocated and degraded? Furthermore, several microbes require the ERAD machinery in the host cell for pathogenesis (Loureiro and Ploegh, 2006; Cho *et al.*, 2012; Eshraghi *et al.*, 2014; Noack *et al.*, 2014). How do they co-opt the ERAD machinery and for what specific purpose? Importantly, much needs to be learnt about how ERAD and the cellular proteostasis network communicate, how these interactions change during disease or aging, and how we may modulate distinct nodes in the network to treat the expanding number of human diseases linked to the ERAD pathway.

Acknowledgments

L.F.Z. is a recipient of a post-doctoral Fellowship from the Dystonia Medical Research Foundation, and J.L.B. acknowledges support from National Institutes of Health grant GM75061 as well as grant DK79307 ('The Pittsburgh Center

for Kidney Research') and the Cystic Fibrosis Foundation grant BRODSK13XX0.

See also: Interorganellar Communication: Interplay and Processes: Endoplasmic Reticulum Stress in Disease; N-Linked Glycans (N-Glycans). Molecular Principles, Components, Technology, and Concepts: Proteins: Posttranslational Modifications: Key Players in Health and Disease. Organelles: Structure and Function: Intermediate Compartment: A Sorting Station between the Endoplasmic Reticulum and the Golgi Apparatus; The Endoplasmic Reticulum. Protein Synthesis/Degradation: Protein Degradation – Intracellular: Ubiquitin, Ubiquitin-Like Proteins, and Proteasome-Mediated Degradation. Protein Synthesis/Degradation: Translation: Biogenesis of Secretory Proteins

References

- Aebi, M., Bernasconi, R., Clerc, S., Molinari, M., 2010. N-glycan structures: Recognition and processing in the ER. *Trends in Biochemical Sciences* 35, 74–82.
- Ahner, A., Nakatsukasa, K., Zhang, H., Frizzell, R.A., Brodsky, J.L., 2007. Small heat-shock proteins select deltaF508-CFTR for endoplasmic reticulum-associated degradation. *Molecular Biology of the Cell* 18, 806–814.
- Akopian, D., Shen, K., Zhang, X., Shan, S.O., 2013. Signal recognition particle: An essential protein-targeting machine. *Annual Review of Biochemistry* 82, 693–721.
- Alder, N.N., Shen, Y., Brodsky, J.L., Hendershot, L.M., Johnson, A.E., 2005. The molecular mechanisms underlying BiP-mediated gating of the Sec61 translocon of the endoplasmic reticulum. *Journal of Cell Biology* 168, 389–399.
- Apweiler, R., Hermjakob, H., Sharon, N., 1999. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochimica et Biophysica Acta* 1473, 4–8.
- Aridor, M., Bannykh, S.I., Rowe, T., Balch, W.E., 1995. Sequential coupling between COPII and COPI vesicle coats in endoplasmic reticulum to Golgi transport. *Journal of Cell Biology* 131, 875–893.
- Ast, T., Aviram, N., Chuartzman, S.G., Schuldiner, M., 2014. A cytosolic degradation pathway, prERAD, monitors pre-inserted secretory pathway proteins. *Journal of Cell Science* 127, 3017–3023.
- Ast, T., Cohen, G., Schuldiner, M., 2013. A network of cytosolic factors targets SRP-independent proteins to the endoplasmic reticulum. *Cell* 152, 1134–1145.
- Avezov, E., Frenkel, Z., Ehrlich, M., Herscovics, A., Lederkremer, G.Z., 2008. Endoplasmic reticulum (ER) mannosidase I is compartmentalized and required for N-glycan trimming to Man5-6GlcNAc2 in glycoprotein ER-associated degradation. *Molecular Biology of the Cell* 19, 216–225.
- Bagola, K., Von Delbruck, M., Dittmar, G., *et al.*, 2013. Ubiquitin binding by a CUE domain regulates ubiquitin chain formation by ERAD E3 ligases. *Molecular Cell* 50, 528–539.
- Bannykh, S.I., Rowe, T., Balch, W.E., 1996. The organization of endoplasmic reticulum export complexes. *Journal of Cell Biology* 135, 19–35.
- Barlowe, C., 2003. Signals for COPII-dependent export from the ER: What's the ticket out? *Trends in Cell Biology* 13, 295–300.
- Barlowe, C., Orci, L., Yeung, T., *et al.*, 1994. COPII: A membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* 77, 895–907.
- Barthelme, D., Sauer, R.T., 2012. Identification of the Cdc48*20S proteasome as an ancient AAA + proteolytic machine. *Science* 337, 843–846.
- Bays, N.W., Gardner, R.G., Seelig, L.P., Joazeiro, C.A., Hampton, R.Y., 2001. Hrd1p/ Der3p is a membrane-anchored ubiquitin ligase required for ER-associated degradation. *Nature Cell Biology* 3, 24–29.
- Bazirgan, O.A., Hampton, R.Y., 2008. Cue1p is an activator of Ubc7p E2 activity *in vitro* and *in vivo*. *Journal of Biological Chemistry* 283, 12797–12810.
- Beckers, C.J., Block, M.R., Glick, B.S., Rothman, J.E., Balch, W.E., 1989. Vesicular transport between the endoplasmic reticulum and the Golgi stack requires the NEM-sensitive fusion protein. *Nature* 339, 397–398.
- Benitez, E.M., Stolz, A., Wolf, D.H., 2011. Yos9, a control protein for misfolded glycosylated and non-glycosylated proteins in ERAD. *FEBS Letters* 585, 3015–3019.

- Bernardi, K.M., Williams, J.M., Inoue, T., Schultz, A., Tsai, B., 2013. A deubiquitinase negatively regulates retro-translocation of nonubiquitinated substrates. *Molecular Biology of the Cell* 24, 3545–3556.
- Bernardi, K.M., Williams, J.M., Kikkert, M., *et al.*, 2010. The E3 ubiquitin ligases Hrd1 and gp78 bind to and promote cholera toxin retro-translocation. *Molecular Biology of the Cell* 21, 140–151.
- Bernasconi, R., Galli, C., Calanca, V., Nakajima, T., Molinari, M., 2010. Stringent requirement for HRD1, SEL1L, and OS-9/XTP3-B for disposal of ERAD-LS substrates. *Journal of Cell Biology* 188, 223–235.
- Bernasconi, R., Galli, C., Kokame, K., Molinari, M., 2013. Autoadaptive ER-associated degradation defines a preemptive unfolded protein response pathway. *Molecular Cell* 52, 783–793.
- Bernasconi, R., Galli, C., Noack, J., *et al.*, 2012. Role of the SEL1L:LC3-I complex as an ERAD tuning receptor in the mammalian ER. *Molecular Cell* 46, 809–819.
- Bhamidipati, A., Denic, V., Quan, E.M., Weissman, J.S., 2005. Exploration of the topological requirements of ERAD identifies Yos9p as a lectin sensor of misfolded glycoproteins in the ER lumen. *Molecular Cell* 19, 741–751.
- Biederer, T., Volkwein, C., Sommer, T., 1997. Role of Cue1p in ubiquitination and degradation at the ER surface. *Science* 278, 1806–1809.
- Blöbel, G., Dobberstein, B., 1975. Transfer of proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. *Journal of Cell Biology* 67, 852–862.
- Blöbel, G., Walter, P., Chang, C.N., *et al.*, 1979. Translocation of proteins across membranes: The signal hypothesis and beyond. *Symposia of the Society for Experimental Biology* 33, 9–36.
- Bonifacino, J.S., Glick, B.S., 2004. The mechanisms of vesicle budding and fusion. *Cell* 116, 153–166.
- Brodsky, J.L., 2012. Cleaning up: ER-associated degradation to the rescue. *Cell* 151, 1163–1167.
- Brodsky, J.L., Fisher, E.A., 2008. The many intersecting pathways underlying apolipoprotein B secretion and degradation. *Trends in Endocrinology & Metabolism* 19, 254–259.
- Brodsky, J.L., Goeckeler, J., Schekman, R., 1995. BiP and Sec63p are required for both co- and posttranslational protein translocation into the yeast endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America* 92, 9643–9646.
- Brodsky, J.L., Werner, E.D., Dubas, M.E., *et al.*, 1999. The requirement for molecular chaperones during endoplasmic reticulum-associated protein degradation demonstrates that protein export and import are mechanistically distinct. *Journal of Biological Chemistry* 274, 3453–3460.
- Buck, T.M., Wright, C.M., Brodsky, J.L., 2007. The activities and function of molecular chaperones in the endoplasmic reticulum. *Seminars in Cell & Developmental Biology* 18, 751–761.
- Calì, T., Galli, C., Olivari, S., Molinari, M., 2008. Segregation and rapid turnover of EDEM1 by an autophagy-like mechanism modulates standard ERAD and folding activities. *Biochemical and Biophysical Research Communications* 371, 405–410.
- Caramelo, J.J., Castro, O.A., Alonso, L.G., de Prat-Gay, G., Parodi, A.J., 2003. UDP-Glc:glycoprotein glucosyltransferase recognizes structured and solvent accessible hydrophobic patches in molten globule-like folding intermediates. *Proceedings of the National Academy of Sciences of the United States of America* 100, 86–91.
- Caramelo, J.J., Parodi, A.J., 2008. Getting in and out from calnexin/calreticulin cycles. *Journal of Biological Chemistry* 283, 10221–10225.
- Carlson, E.J., Pitzonzo, D., Skach, W.R., 2006. p97 functions as an auxiliary factor to facilitate TM domain extraction during CFTR ER-associated degradation. *EMBO Journal* 25, 4557–4566.
- Carroll, S.M., Hampton, R.Y., 2010. Usa1p is required for optimal function and regulation of the Hrd1p endoplasmic reticulum-associated degradation ubiquitin ligase. *Journal of Biological Chemistry* 285, 5146–5156.
- Carvalho, P., Goder, V., Rapoport, T.A., 2006. Distinct ubiquitin-ligase complexes define convergent pathways for the degradation of ER proteins. *Cell* 126, 361–373.
- Carvalho, P., Stanley, A.M., Rapoport, T.A., 2010. Retrotranslocation of a misfolded luminal ER protein by the ubiquitin-ligase Hrd1p. *Cell* 143, 579–591.
- Chau, V., Tobias, J.W., Bachmair, A., *et al.*, 1989. A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* 243, 1576–1583.
- Chen, L., Madura, K., 2002. Rad23 promotes the targeting of proteolytic substrates to the proteasome. *Molecular and Cellular Biology* 22, 4902–4913.
- Chen, X., Tukachinsky, H., Huang, C.H., *et al.*, 2011a. Processing and turnover of the Hedgehog protein in the endoplasmic reticulum. *Journal of Cell Biology* 192, 825–838.
- Chen, Y., Hu, D., Yabe, R., *et al.*, 2011b. Role of malexin in Glc(2)Man(9)GlcNAc(2)-dependent quality control of alpha1-antitrypsin. *Molecular Biology of the Cell* 22, 3559–3570.
- Cheng, S.H., Gregory, R.J., Marshall, J., *et al.*, 1990. Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell* 63, 827–834.
- Cho, J.A., Chinnappen, D.J., Amar, E., *et al.*, 2012. Insights on the trafficking and retro-translocation of glycosphingolipid-binding bacterial toxins. *Frontiers in Cellular and Infection Microbiology* 2, 51.
- Christianson, J.C., Shaler, T.A., Tyler, R.E., Kopito, R.R., 2008. OS-9 and GRP94 deliver mutant alpha1-antitrypsin to the Hrd1-SEL1L ubiquitin ligase complex for ERAD. *Nature Cell Biology* 10, 272–282.
- Christianson, J.C., Ye, Y., 2014. Cleaning up in the endoplasmic reticulum: Ubiquitin in charge. *Frontiers in Cellular and Infection Microbiology* 21, 325–335.
- Claessen, J.H., Kundrat, L., Ploegh, H.L., 2012. Protein quality control in the ER: Balancing the ubiquitin checkbook. *Trends in Cell Biology* 22, 22–32.
- Clerc, S., Hirsch, C., Oggier, D.M., *et al.*, 2009. Htm1 protein generates the N-glycan signal for glycoprotein degradation in the endoplasmic reticulum. *Journal of Cell Biology* 184, 159–172.
- Coe, H., Michalak, M., 2010. ERp57, a multifunctional endoplasmic reticulum resident oxidoreductase. *International Journal of Biochemistry & Cell Biology* 42, 796–799.
- Cormier, J.H., Tamura, T., Sunryd, J.C., Hebert, D.N., 2009. EDEM1 recognition and delivery of misfolded proteins to the SEL1L-containing ERAD complex. *Molecular Cell* 34, 627–633.
- Cross, B.C., Sinning, I., Lührink, J., High, S., 2009. Delivering proteins for export from the cytosol. *Nature Reviews Molecular Cellular Biology* 10, 255–264.
- Crowley, K.S., Liao, S., Worrell, V.E., Reinhart, G.D., Johnson, A.E., 1994. Secretory proteins move through the endoplasmic reticulum membrane via an aqueous, gated pore. *Cell* 78, 461–471.
- Deak, P.M., Wolf, D.H., 2001. Membrane topology and function of Der3/Hrd1p as a ubiquitin-protein ligase (E3) involved in endoplasmic reticulum degradation. *Journal of Biological Chemistry* 276, 10663–10669.
- Dejgaard, K., Theberge, J.F., Heath-Engel, H., *et al.*, 2010. Organization of the Sec61 translocon, studied by high resolution native electrophoresis. *International Journal of Biochemistry & Cell Biology* 9, 1763–1771.
- Denic, V., Quan, E.M., Weissman, J.S., 2006. A luminal surveillance complex that selects misfolded glycoproteins for ER-associated degradation. *Cell* 126, 349–359.
- Deshaies, R.J., Koch, B.D., Werner-Washburne, M., Craig, E.A., Schekman, R., 1988. A subfamily of stress proteins facilitates translocation of secretory and mitochondrial precursor polypeptides. *Nature* 332, 800–805.
- Deshaies, R.J., Schekman, R., 1989. SEC62 encodes a putative membrane protein required for protein translocation into the yeast endoplasmic reticulum. *Journal of Cell Biology* 109, 2653–2664.
- Eakle, K.A., Bernstein, M., Emr, S.D., 1988. Characterization of a component of the yeast secretion machinery: Identification of the SEC18 gene product. *Molecular and Cellular Biology* 8, 4098–4109.
- Elkabetz, Y., Shapira, I., Rabinovich, E., Bar-Nun, S., 2004. Distinct steps in dislocation of luminal endoplasmic reticulum-associated degradation substrates: Roles of endoplasmic reticulum-bound p97/Cdc48p and proteasome. *Journal of Biological Chemistry* 279, 3980–3989.
- Eshraghi, A., Dixon, S.D., Tamilselvan, B., *et al.*, 2014. Cytotoxic distending toxins require components of the ER-associated degradation pathway for host cell entry. *PLoS Pathogens* 10, e1004295.
- Fernandez, F.S., Trombetta, S.E., Hellman, U., Parodi, A.J., 1994. Purification to homogeneity of UDP-glucose:glycoprotein glucosyltransferase from *Schizosaccharomyces pombe* and apparent absence of the enzyme for *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 269, 30701–30706.
- Finger, A., Knop, M., Wolf, D.H., 1993. Analysis of two mutated vacuolar proteins reveals a degradation pathway in the endoplasmic reticulum or a related compartment of yeast. *European Journal of Biochemistry* 218, 565–574.
- Flierman, D., Ye, Y., Dai, M., Chau, V., Rapoport, T.A., 2003. Polyubiquitin serves as a recognition signal, rather than a ratcheting molecule, during retrotranslocation of proteins across the endoplasmic reticulum membrane. *Journal of Biological Chemistry* 278, 34774–34782.
- Flynn, G.C., Pohl, J., Flocco, M.T., Rothman, J.E., 1991. Peptide-binding specificity of the molecular chaperone BiP. *Nature* 353, 726–730.
- Freeze, H.H., Chong, J.X., Bamshad, M.J., Ng, B.G., 2014. Solving glycosylation disorders: Fundamental approaches reveal complicated pathways. *American Journal of Human Genetics* 94, 161–175.
- Friable, A., Zhang, K., Kaufman, R.J., 2009. Regulation of apoptosis by the unfolded protein response. *Methods in Molecular Biology* 559, 191–204.
- Futerman, A.H., van Meer, G., 2004. The cell biology of lysosomal storage disorders. *Nature Reviews Molecular Cell Biology* 5, 554–565.

- Garza, R.M., Sato, B.K., Hampton, R.Y., 2009. *In vitro* analysis of Hrd1p-mediated retrotranslocation of its multispanning membrane substrate 3-hydroxy-3-methylglutaryl (HMG)-CoA reductase. *Journal of Biological Chemistry* 284, 14710–14722.
- Gauss, R., Jarosch, E., Sommer, T., Hirsch, C., 2006. A complex of Yos9p and the HRD ligase integrates endoplasmic reticulum quality control into the degradation machinery. *Nature Cell Biology* 8, 849–854.
- Gilmore, R., Walter, P., Blobel, G., 1982. Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor. *Journal of Cell Biology* 95, 470–477.
- Green, N., Fang, H., Walter, P., 1992. Mutants in three novel complementation groups inhibit membrane protein insertion into and soluble protein translocation across the endoplasmic reticulum membrane of *Saccharomyces cerevisiae*. *Journal of Cell Biology* 116, 597–604.
- Greenblatt, E.J., Olzmann, J.A., Kopito, R.R., 2011. Derlin-1 is a rhomboid pseudoprotease required for the dislocation of mutant alpha-1 antitrypsin from the endoplasmic reticulum. *Frontiers in Cellular and Infection Microbiology* 18, 1147–1152.
- Grieve, A.G., Rabouille, C., 2011. Golgi bypass: Skirting around the heart of classical secretion. *Cold Spring Harbor Perspectives in Biology* 3, pii: a005298.
- Guerriero, C.J., Brodsky, J.L., 2012. The delicate balance between secreted protein folding and endoplasmic reticulum-associated degradation in human physiology. *Physiological Reviews* 92, 537–576.
- Hampton, R.Y., 2002. ER-associated degradation in protein quality control and cellular regulation. *Current Opinion in Cell Biology* 14, 476–482.
- Hampton, R.Y., Rine, J., 1994. Regulated degradation of HMG-CoA reductase, an integral membrane protein of the endoplasmic reticulum, in yeast. *Journal of Cell Biology* 125, 299–312.
- Han, S., Liu, Y., Chang, A., 2007. Cytoplasmic Hsp70 promotes ubiquitination for endoplasmic reticulum-associated degradation of a misfolded mutant of the yeast plasma membrane ATPase, PMA1. *Journal of Biological Chemistry* 282, 26140–26149.
- Hartl, F.U., Bracher, A., Hayer-Hartl, M., 2011. Molecular chaperones in protein folding and proteostasis. *Nature* 475, 324–332.
- Hauri, H.P., Nufer, O., Breuza, L., Tekaya, H.B., Liang, L., 2002. Lectins and protein traffic early in the secretory pathway. *Biochemical Society Symposia* 73–82.
- Haynes, C.M., Caldwell, S., Cooper, A.A., 2002. An HRD/DER-independent ER quality control mechanism involves Rsp5p-dependent ubiquitination and ER-Golgi transport. *Journal of Cell Biology* 158, 91–101.
- Hebert, D.N., Bernasconi, R., Molinari, M., 2010. ERAD substrates: Which way out? *Seminars in Cell & Developmental Biology* 21, 526–532.
- Hegde, R.S., Bernstein, H.D., 2006. The surprising complexity of signal sequences. *Trends in Biochemical Sciences* 31, 563–571.
- Hegde, R.S., Keenan, R.J., 2011. Tail-anchored membrane protein insertion into the endoplasmic reticulum. *Nature Reviews Molecular Cell Biology* 12, 787–798.
- Hiller, M.M., Finger, A., Schweiger, M., Wolf, D.H., 1996. ER degradation of a misfolded luminal protein by the cytosolic ubiquitin-proteasome pathway. *Science* 273, 1725–1728.
- Hirao, K., Natsuka, Y., Tamura, T., *et al.*, 2006. EDEM3, a soluble EDEM homolog, enhances glycoprotein endoplasmic reticulum-associated degradation and mannose trimming. *Journal of Biological Chemistry* 281, 9650–9658.
- Hirsch, C., Blom, D., Ploegh, H.L., 2003. A role for N-glycanase in the cytosolic turnover of glycoproteins. *EMBO Journal* 22, 1036–1046.
- Horn, S.C., Hanna, J., Hirsch, C., *et al.*, 2009. Usa1 functions as a scaffold of the HRD-ubiquitin ligase. *Molecular Cell* 36, 782–793.
- Hosokawa, N., Kamiya, Y., Kato, K., 2010a. The role of MRH domain-containing lectins in ERAD. *Glycobiology* 20, 651–660.
- Hosokawa, N., Tremblay, L.O., Steno, B., *et al.*, 2010b. EDEM1 accelerates the trimming of alpha1,2-linked mannose on the C branch of N-glycans. *Glycobiology* 20, 567–575.
- Houck, S.A., Ren, H.Y., Madden, V.J., *et al.*, 2014. Quality control autophagy degrades soluble ERAD-resistant conformers of the misfolded membrane protein GnRHR. *Molecular Cell* 54, 166–179.
- Huang, C.H., Chu, Y.R., YE, Y., Chen, X., 2013. Role of HERP and a HERP-related protein in HRD1-dependent protein degradation at the endoplasmic reticulum. *Journal of Biological Chemistry* 289, 4444–4454.
- Hurtley, S.M., Helenius, A., 1989. Protein oligomerization in the endoplasmic reticulum. *Annual Review of Cell and Developmental Biology* 5, 277–307.
- Huyer, G., Piluek, W.F., Fansler, Z., *et al.*, 2004. Distinct machinery is required in *Saccharomyces cerevisiae* for the endoplasmic reticulum-associated degradation of a multispanning membrane protein and a soluble luminal protein. *Journal of Biological Chemistry* 279, 38369–38378.
- Ihara, Y., Cohen-Doyle, M.F., Saito, Y., Williams, D.B., 1999. Calnexin discriminates between protein conformational states and functions as a molecular chaperone *in vitro*. *Molecular Cell* 4, 331–341.
- Ikeda, Y., Demartino, G.N., Brown, M.S., *et al.*, 2009. Regulated endoplasmic reticulum-associated degradation of a polytopic protein: p97 recruits proteasomes to Insig-1 before extraction from membranes. *Journal of Biological Chemistry* 284, 34889–34900.
- Imbach, T., Burda, P., Kuhnert, P., *et al.*, 1999. A mutation in the human ortholog of the *Saccharomyces cerevisiae* ALG6 gene causes carbohydrate-deficient glycoprotein syndrome type-Ic. *Proceedings of the National Academy of Sciences of the United States of America* 96, 6982–6987.
- Imperiali, B., Rickert, K.W., 1995. Conformational implications of asparagine-linked glycosylation. *Proceedings of the National Academy of Sciences of the United States of America* 92, 97–101.
- Ismail, N., Crawshaw, S.G., High, S., 2006. Active and passive displacement of transmembrane domains both occur during opsin biogenesis at the Sec61 translocon. *Journal of Cell Science* 119, 2826–2836.
- Jackson, M.R., Nilsson, T., Peterson, P.A., 1990. Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *EMBO Journal* 9, 3153–3162.
- Jarosch, E., Taxis, C., Volkwein, C., *et al.*, 2002. Protein dislocation from the ER requires polyubiquitination and the AAA-ATPase Cdc48. *Nature Cell Biology* 4, 134–139.
- Jensen, T.J., Loo, M.A., Pind, S., *et al.*, 1995. Multiple proteolytic systems, including the proteasome, contribute to CFTR processing. *Cell* 83, 129–135.
- Johnston, J.A., Ward, C.L., Kopito, R.R., 1998. Aggresomes: A cellular response to misfolded proteins. *Journal of Cell Biology* 143, 1883–1898.
- Jucker, M., Walker, L.C., 2013. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* 501, 45–51.
- Jungnickel, B., Rapoport, T.A., 1995. A posttargeting signal sequence recognition event in the endoplasmic reticulum membrane. *Cell* 82, 261–270.
- Kabani, M., Martineau, C.N., 2008. Multiple hsp70 isoforms in the eukaryotic cytosol: Mere redundancy or functional specificity? *Current Genomics* 9, 248–338.
- Kampinga, H.H., Craig, E.A., 2010. The HSP70 chaperone machinery: J proteins as drivers of functional specificity. *Nature Reviews Molecular Cell Biology* 11, 579–592.
- Kanehara, K., Xie, W., NG, D.T., 2010. Modularity of the Hrd1 ERAD complex underlies its diverse client range. *Journal of Cell Biology* 188, 707–716.
- Kario, E., Amar, N., Elazar, Z., Navon, A., 2011. A new autophagy-related checkpoint in the degradation of an ERAD-M target. *Journal of Biological Chemistry* 286, 11479–11491.
- Kelleher, D.J., Gilmore, R., 2006. An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* 16, 47R–62R.
- Kincaid, M.M., Cooper, A.A., 2007. Misfolded proteins traffic from the endoplasmic reticulum (ER) due to ER export signals. *Molecular Biology of the Cell* 18, 455–463.
- Klausner, R.D., Sitia, R., 1990. Protein degradation in the endoplasmic reticulum. *Cell* 62, 611–614.
- Kleizen, B., Braakman, I., 2013. Cell biology. A sweet send-off. *Science* 340, 930–931.
- Kostova, Z., Tsai, Y.C., Weissman, A.M., 2007. Ubiquitin ligases, critical mediators of endoplasmic reticulum-associated degradation. *Seminars in Cell & Developmental Biology* 18, 770–779.
- Kostova, Z., Wolf, D.H., 2005. Importance of carbohydrate positioning in the recognition of mutated CPY for ER-associated degradation. *Journal of Cell Science* 118, 1485–1492.
- Kothe, M., Ye, Y., Wagner, J.S., *et al.*, 2005. Role of p97 AAA-ATPase in the retrotranslocation of the cholera toxin A1 chain, a non-ubiquitinated substrate. *Journal of Biological Chemistry* 280, 28127–28132.
- Kraut-Cohen, J., Afanasieva, E., Haim-Vilimovsky, L., *et al.*, 2013. Translation- and SRP-independent mRNA targeting to the endoplasmic reticulum in the yeast *Saccharomyces cerevisiae*. *Molecular Biology of the Cell* 24, 3069–3084.
- Kravtsova-Ivantsiv, Y., Ciechanover, A., 2012. Non-canonical ubiquitin-based signals for proteasomal degradation. *Journal of Cell Science* 125, 539–548.
- Kroeger, H., Miranda, E., Macleod, I., *et al.*, 2009. Endoplasmic reticulum-associated degradation (ERAD) and autophagy cooperate to degrade polymerogenic mutant serpins. *Journal of Biological Chemistry* 284, 22793–22802.
- Kruse, K.B., Brodsky, J.L., McCracken, A.A., 2006. Characterization of an ERAD gene as VPS30/ATG6 reveals two alternative and functionally distinct protein quality control pathways: One for soluble Z variant of human alpha-1 proteinase inhibitor (A1PiZ) and another for aggregates of A1PiZ. *Molecular Biology of the Cell* 17, 203–212.
- Labriola, C.A., Giraldo, A.M., Parodi, A.J., Caramelo, J.J., 2011. Functional cooperation between BiP and calreticulin in the folding maturation of a

- glycoprotein in *Trypanosoma cruzi*. *Molecular and Biochemical Parasitology* 175, 112–117.
- Le Fourn, V., Park, S., Jang, I., *et al.*, 2013. Large protein complexes retained in the ER are dislocated by non-COPII vesicles and degraded by selective autophagy. *Cellular and Molecular Life Sciences* 70, 1985–2002.
- Lee, R.J., Liu, C.W., Harty, C., *et al.*, 2004. Uncoupling retro-translocation and degradation in the ER-associated degradation of a soluble protein. *EMBO Journal* 23, 2206–2215.
- Leichner, G.S., Avner, R., Harats, D., Roitelman, J., 2009. Dislocation of HMG-CoA reductase and Insig-1, two polytopic endoplasmic reticulum proteins, en route to proteasomal degradation. *Molecular Biology of the Cell* 20, 3330–3341.
- Lemberg, M.K., 2013. Sampling the membrane: Function of rhomboid-family proteins. *Trends in Cell Biology* 23, 210–217.
- Létourneur, F., Gaynor, E.C., Hennecke, S., *et al.*, 1994. Coatamer is essential for retrieval of dilysine-tagged proteins to the endoplasmic reticulum. *Cell* 79, 1199–1207.
- Loibl, M., Wunderle, L., Hutzler, J., *et al.*, 2014. Protein O-mannosyltransferases associate with the translocon to modify translocating polypeptide chains. *Journal of Biological Chemistry* 289, 8599–8611.
- Lord, C., Ferro-Novick, S., Miller, E.A., 2013. The highly conserved COPII coat complex sorts cargo from the endoplasmic reticulum and targets it to the golgi. *Cold Spring Harbor Perspectives in Biology* 5, a013367.
- Loureiro, J., Ploegh, H.L., 2006. Antigen presentation and the ubiquitin-proteasome system in host-pathogen interactions. *Advances in Immunology* 92, 225–305.
- Marin, M.B., Ghenea, S., Spiridon, L.N., *et al.*, 2012. Tyrosinase degradation is prevented when EDEM1 lacks the intrinsically disordered region. *PLoS One* 7, e42998.
- Matlack, K.E., Misselwitz, B., Plath, K., Rapoport, T.A., 1999. BiP acts as a molecular ratchet during posttranslational transport of prepro- α factor across the ER membrane. *Cell* 97, 553–564.
- McCracken, A.A., Brodsky, J.L., 1996. Assembly of ER-associated protein degradation *in vitro*: Dependence on cytosol, calnexin, and ATP. *Journal of Cell Biology* 132, 291–298.
- McCracken, A.A., Karpichev, I.V., Ernaga, J.E., *et al.*, 1996. Yeast mutants deficient in ER-associated degradation of the Z variant of α -1-protease inhibitor. *Genetics* 144, 1355–1362.
- McCracken, A.A., Kruse, K.B., 1989. Intracellular transport of rat serum albumin is altered by a genetically engineered deletion of the propeptide. *Journal of Biological Chemistry* 264, 20843–20846.
- McCracken, A.A., Kruse, K.B., 1993. Selective protein degradation in the yeast exocytic pathway. *Molecular Biology of the Cell* 4, 729–736.
- McCracken, A.A., Kruse, K.B., Brown, J.L., 1989. Molecular basis for defective secretion of the Z variant of human α -1-proteinase inhibitor: Secretion of variants having altered potential for salt bridge formation between amino acids 290 and 342. *Molecular and Cellular Biology* 9, 1406–1414.
- McGrath, J.P., Jentsch, S., Varshavsky, A., 1991. UBA 1: An essential yeast gene encoding ubiquitin-activating enzyme. *EMBO Journal* 10, 227–236.
- Mehner, M., Sommer, T., Jarosch, E., 2013. Der1 promotes movement of misfolded proteins through the endoplasmic reticulum membrane. *Nature Cell Biology* 16, 77–86.
- Metzger, M.B., Liang, Y.H., Das, R., *et al.*, 2013. A structurally unique E2-binding domain activates ubiquitination by the ERAD E2, Ubc7p, through multiple mechanisms. *Molecular Cell* 50, 516–527.
- Michalak, M., Corbett, E.F., Mesaeli, N., Nakamura, K., Opas, M., 1999. Calreticulin: One protein, one gene, many functions. *Biochemical Journal* 344 (Pt 2), 281–292.
- Michelsen, K., Yuan, H., Schwappach, B., 2005. Hide and run. Arginine-based endoplasmic-reticulum-sorting motifs in the assembly of heteromultimeric membrane proteins. *EMBO Reports* 6, 717–722.
- Molinari, M., Calanca, V., Galli, C., Lucca, P., Paganetti, P., 2003. Role of EDEM in the release of misfolded glycoproteins from the calnexin cycle. *Science* 299, 1397–1400.
- Molinari, M., Galli, C., Piccaluga, V., Pieren, M., Paganetti, P., 2002. Sequential assistance of molecular chaperones and transient formation of covalent complexes during protein degradation from the ER. *Journal of Cell Biology* 158, 247–257.
- Munro, S., Pelham, H.R., 1987. A C-terminal signal prevents secretion of luminal ER proteins. *Cell* 48, 899–907.
- Nakatsukasa, K., Brodsky, J.L., Kamura, T., 2013. A stalled retrotranslocation complex reveals physical linkage between substrate recognition and proteasomal degradation during ER-associated degradation. *Molecular Biology of the Cell* 24 (1765–1775), S1–S8.
- Nakatsukasa, K., Huyer, G., Michaelis, S., Brodsky, J.L., 2008. Dissecting the ER-associated degradation of a misfolded polytopic membrane protein. *Cell* 132, 101–112.
- Nakatsukasa, K., Okada, S., Umehayashi, K., *et al.*, 2004. Roles of O-mannosylation of aberrant proteins in reduction of the load for endoplasmic reticulum chaperones in yeast. *Journal of Biological Chemistry* 279, 49762–49772.
- Needham, P.G., Brodsky, J.L., 2013. How early studies on secreted and membrane protein quality control gave rise to the ER associated degradation (ERAD) pathway: The early history of ERAD. *Biochimica et Biophysica Acta* 1833, 2447–2457.
- Neuber, O., Jarosch, E., Volkwein, C., Walter, J., Sommer, T., 2005. Ubx2 links the Cdc48 complex to ER-associated protein degradation. *Nature Cell Biology* 7, 993–998.
- Ng, D.T., Brown, J.D., Walter, P., 1996. Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *Journal of Cell Biology* 134, 269–278.
- Ninagawa, S., Okada, T., Sumitomo, Y., *et al.*, 2014. EDEM2 initiates mammalian glycoprotein ERAD by catalyzing the first mannose trimming step. *Journal of Cell Biology* 206, 347–356.
- Nishikawa, S., Brodsky, J.L., Nakatsukasa, K., 2005. Roles of molecular chaperones in endoplasmic reticulum (ER) quality control and ER-associated degradation (ERAD). *Journal of Biochemistry* 137, 551–555.
- Nishikawa, S.I., Fewell, S.W., Kato, Y., Brodsky, J.L., Endo, T., 2001. Molecular chaperones in the yeast endoplasmic reticulum maintain the solubility of proteins for retrotranslocation and degradation. *Journal of Cell Biology* 153, 1061–1070.
- Noack, J., Bernasconi, R., Molinari, M., 2014. How viruses hijack the ERAD tuning machinery. *Journal of Virology* 88, 10272–10275.
- Noriega, T.R., Tsai, A., Elvekrog, M.M., *et al.*, 2014. Signal recognition particle-ribosome binding is sensitive to nascent chain length. *Journal of Biological Chemistry* 289, 19294–19305.
- Novick, P., Field, C., Schekman, R., 1980. Identification of 23 complementation groups required for post-translational events in the yeast secretory pathway. *Cell* 21, 205–215.
- Oda, Y., Hosokawa, N., Wada, I., Nagata, K., 2003. EDEM as an acceptor of terminally misfolded glycoproteins released from calnexin. *Science* 299, 1394–1397.
- Okreglak, V., Walter, P., 2014. The conserved AAA-ATPase Msp1 confers organelle specificity to tail-anchored proteins. *Proceedings of the National Academy of Sciences of the United States of America* 111, 8019–8024.
- Okuda-Shimizu, Y., Hendershot, L.M., 2007. Characterization of an ERAD pathway for nonglycosylated BiP substrates, which require Herp. *Molecular Cell* 28, 544–554.
- Olzmann, J.A., Kopito, R.R., 2011. Lipid droplet formation is dispensable for endoplasmic reticulum-associated degradation. *Journal of Biological Chemistry* 286, 27872–27874.
- Olzmann, J.A., Kopito, R.R., Christianson, J.C., 2013. The mammalian endoplasmic reticulum-associated degradation system. *Cold Spring Harbor Perspectives in Biology* 5, pii: a013185.
- Orci, L., Ravazzola, M., Meda, P., *et al.*, 1991. Mammalian Sec23p homologue is restricted to the endoplasmic reticulum transitional cytoplasm. *Proceedings of the National Academy of Sciences of the United States of America* 88, 8611–8615.
- Palade, G., 1975. Intracellular aspects of the process of protein synthesis. *Science* 189, 347–358.
- Palmer, A., Rivett, A.J., Thomson, S., *et al.*, 1996. Subpopulations of proteasomes in rat liver nuclei, microsomes and cytosol. *Biochemical Journal* 316 (Pt 2), 401–407.
- Panzner, S., Dreier, L., Hartmann, E., Kostka, S., Rapoport, T.A., 1995. Posttranslational protein transport in yeast reconstituted with a purified complex of Sec proteins and Kar2p. *Cell* 81, 561–570.
- Park, S., Jang, I., Zuber, C., *et al.*, 2014. ERADication of EDEM1 occurs by selective autophagy and requires deglycosylation by cytoplasmic peptide N-glycanase. *Histochemistry and Cell Biology* 142, 153–169.
- Parlati, F., Dominguez, M., Bergeron, J.J., Thomas, D.Y., 1995. *Saccharomyces cerevisiae* CNE1 encodes an endoplasmic reticulum (ER) membrane protein with sequence similarity to calnexin and calreticulin and functions as a constituent of the ER quality control apparatus. *Journal of Biological Chemistry* 270, 244–253.
- Pelham, H.R., 1989. Control of protein exit from the endoplasmic reticulum. *Annual Review of Cell and Developmental Biology* 5, 1–23.
- Pfeffer, S.R., Rothman, J.E., 1987. Biosynthetic protein transport and sorting by the endoplasmic reticulum and Golgi. *Annual Review of Biochemistry* 56, 829–852.
- Plempner, R.K., Bordallo, J., Deak, P.M., *et al.*, 1999. Genetic interactions of Hrd3p and Der3p/Hrd1p with Sec61p suggest a retro-translocation complex mediating protein transport for ER degradation. *Journal of Cell Science* 112 (Pt 22), 4123–4134.

- Ploegh, H.L., 2007. A lipid-based model for the creation of an escape hatch from the endoplasmic reticulum. *Nature* 448, 435–438.
- Quan, E.M., Kamiya, Y., Kamiya, D., *et al.*, 2008. Defining the glycan destruction signal for endoplasmic reticulum-associated degradation. *Molecular Cell* 32, 870–877.
- Rapoport, T.A., 2007. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature* 450, 663–669.
- Rayner, J.C., Pelham, H.R., 1997. Transmembrane domain-dependent sorting of proteins to the ER and plasma membrane in yeast. *EMBO Journal* 16, 1832–1841.
- Richly, H., Rape, M., Braun, S., *et al.*, 2005. A series of ubiquitin binding factors connects CDC48/p97 to substrate multiubiquitylation and proteasomal targeting. *Cell* 120, 73–84.
- Ritter, C., Helenius, A., 2000. Recognition of local glycoprotein misfolding by the ER folding sensor UDP-glucose:glycoprotein glucosyltransferase. *Nature Structural & Molecular Biology* 7, 278–280.
- Ron, I., Horowitz, M., 2005. ER retention and degradation as the molecular basis underlying Gaucher disease heterogeneity. *Human Molecular Genetics* 14, 2387–2398.
- Ronchi, P., Colombo, S., Francolini, M., Borgese, N., 2008. Transmembrane domain-dependent partitioning of membrane proteins within the endoplasmic reticulum. *Journal of Cell Biology* 181, 105–118.
- Ruiz-Canada, C., Kelleher, D.J., Gilmore, R., 2009. Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. *Cell* 136, 272–283.
- Sato, B.K., Schulz, D., Do, P.H., Hampton, R.Y., 2009. Misfolded membrane proteins are specifically recognized by the transmembrane domain of the Hrd1p ubiquitin ligase. *Molecular Cell* 34, 212–222.
- Sato, M., Sato, K., Nakano, A., 1996. Endoplasmic reticulum localization of Sec12p is achieved by two mechanisms: Rer1p-dependent retrieval that requires the transmembrane domain and Rer1p-independent retention that involves the cytoplasmic domain. *Journal of Cell Biology* 134, 279–293.
- Schallus, T., Jaech, C., Feher, K., *et al.*, 2008. Malectin: A novel carbohydrate-binding protein of the endoplasmic reticulum and a candidate player in the early steps of protein N-glycosylation. *Molecular Biology of the Cell* 19, 3404–3414.
- Schmidt, M., Finley, D., 2013. Regulation of proteasome activity in health and disease. *Biochimica et Biophysica Acta* 1843, 13–25.
- Schubert, C., Buchberger, A., 2005. Membrane-bound Ubx2 recruits Cdc48 to ubiquitin ligases and their substrates to ensure efficient ER-associated protein degradation. *Nature Cell Biology* 7, 999–1006.
- Schuldiner, M., Metz, J., Schmid, V., *et al.*, 2008. The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* 134, 634–645.
- Schulze, A., Stander, S., Buerger, E., *et al.*, 2005. The ubiquitin-domain protein HERP forms a complex with components of the endoplasmic reticulum associated degradation pathway. *Journal of Molecular Biology* 354, 1021–1027.
- Schutze, M.P., Peterson, P.A., Jackson, M.R., 1994. An N-terminal double-arginine motif maintains type II membrane proteins in the endoplasmic reticulum. *EMBO Journal* 13, 1696–1705.
- Scott, D.C., Schekman, R., 2008. Role of Sec61p in the ER-associated degradation of short-lived transmembrane proteins. *Journal of Cell Biology* 181, 1095–1105.
- Shao, S., Hegde, R.S., 2011. Membrane protein insertion at the endoplasmic reticulum. *Annual Review of Cell and Developmental Biology* 27, 25–56.
- Shenkman, M., Groisman, B., Ron, E., *et al.*, 2013. A shared endoplasmic reticulum-associated degradation pathway involving the EDEM1 protein for glycosylated and nonglycosylated proteins. *Journal of Biological Chemistry* 288, 2167–2178.
- Siegel, V., Walter, P., 1986. Removal of the Alu structural domain from signal recognition particle leaves its protein translocation activity intact. *Nature* 320, 81–84.
- Smith, D.M., Chang, S.C., Park, S., *et al.*, 2007. Docking of the proteasomal ATPases' carboxyl termini in the 20S proteasome's alpha ring opens the gate for substrate entry. *Molecular Cell* 27, 731–744.
- Sollner, T., Whiteheart, S.W., Brunner, M., *et al.*, 1993. SNAP receptors implicated in vesicle targeting and fusion. *Nature* 362, 318–324.
- Sommer, T., Jentsch, S., 1993. A protein translocation defect linked to ubiquitin conjugation at the endoplasmic reticulum. *Nature* 365, 176–179.
- Spear, E.D., Ng, D.T., 2005. Single, context-specific glycans can target misfolded glycoproteins for ER-associated degradation. *Journal of Cell Biology* 169, 73–82.
- Stanley, A.M., Carvalho, P., Rapoport, T., 2011. Recognition of an ERAD-L substrate analyzed by site-specific *in vivo* photocrosslinking. *FEBS Letters* 585, 1281–1286.
- Stigliano, I.D., Alcumbre, S.G., Labriola, C.A., Parodi, A.J., D'Alessio, C., 2011. Glucosidase II and N-glycan mannose content regulate the half-lives of monoglucosylated species *in vivo*. *Molecular Biology of the Cell* 22, 1810–1823.
- Stolz, A., Besser, S., Hottmann, H., Wolf, D.H., 2013. Previously unknown role for the ubiquitin ligase Ubr1 in endoplasmic reticulum-associated protein degradation. *Proceedings of the National Academy of Sciences of the United States of America* 110, 15271–15276.
- Suzuki, T., Park, H., Hollingsworth, N.M., Sternglanz, R., Lennarz, W.J., 2000. PNG1, a yeast gene encoding a highly conserved peptide:N-glycanase. *Journal of Cell Biology* 149, 1039–1052.
- Swanson, R., Locher, M., Hochstrasser, M., 2001. A conserved ubiquitin ligase of the nuclear envelope/endoplasmic reticulum that functions in both ER-associated and Matalpha2 repressor degradation. *Genes & Development* 15, 2660–2674.
- Tang, H.Y., Huang, C.H., Zhuang, Y.H., Christianson, J.C., Chen, X., 2014. EDEM2 and OS-9 are required for ER-associated degradation of non-glycosylated sonic hedgehog. *PLoS One* 9, e92164.
- Taxis, C., Vogel, F., Wolf, D.H., 2002. ER-golgi traffic is a prerequisite for efficient ER degradation. *Molecular Biology of the Cell* 13, 1806–1818.
- Taylor, S.C., Ferguson, A.D., Bergeron, J.J., Thomas, D.Y., 2004. The ER protein folding sensor UDP-glucose glycoprotein-glucosyltransferase modifies substrates distant to local changes in glycoprotein conformation. *Frontiers in Cellular and Infection Microbiology* 11, 128–134.
- Taylor, S.C., Thibault, P., Tessier, D.C., Bergeron, J.J., Thomas, D.Y., 2003. Glycopeptide specificity of the secretory protein folding sensor UDP-glucose glycoprotein-glucosyltransferase. *EMBO Reports* 4, 405–411.
- Thibault, G., Ng, D.T., 2012. The endoplasmic reticulum-associated degradation pathways of budding yeast. *Cold Spring Harbor Perspectives in Biology* 4, pii: a013193.
- Tirosh, B., Furman, M.H., Tortorella, D., Ploegh, H.L., 2003. Protein unfolding is not a prerequisite for endoplasmic reticulum-to-cytosol dislocation. *Journal of Biological Chemistry* 278, 6664–6672.
- Travers, K.J., Patil, C.K., Wodicka, L., *et al.*, 2000. Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. *Cell* 101, 249–258.
- Tretter, T., Pereira, F.P., Ulucan, O., *et al.*, 2013. ERAD and protein import defects in a sec61 mutant lacking ER-lumenal loop 7. *BMC Cell Biology* 14, 56.
- Tsai, B., Rodighiero, C., Lencer, W.I., Rapoport, T.A., 2001. Protein disulfide isomerase acts as a redox-dependent chaperone to unfold cholera toxin. *Cell* 104, 937–948.
- Tsai, B., Ye, Y., Rapoport, T.A., 2002. Retro-translocation of proteins from the endoplasmic reticulum into the cytosol. *Nature Reviews Molecular Cell Biology* 3, 246–255.
- Ushioda, R., Hoseki, J., Araki, K., *et al.*, 2008. ERdj5 is required as a disulfide reductase for degradation of misfolded proteins in the ER. *Science* 321, 569–572.
- Ushioda, R., Hoseki, J., Nagata, K., 2013. Glycosylation-independent ERAD pathway serves as a backup system under ER stress. *Molecular Biology of the Cell* 24, 3155–3163.
- Van den Berg, B., Clemons Jr., W.M., Collinson, I., *et al.*, 2004. X-ray structure of a protein-conducting channel. *Nature* 427, 36–44.
- Vander Heyden, A.B., Naismith, T.V., Snapp, E.L., Hanson, P.I., 2011. Static retention of the luminal monotopic membrane protein torsinA in the endoplasmic reticulum. *EMBO Journal* 30, 3217–3231.
- Vashist, S., Ng, D.T., 2004. Misfolded proteins are sorted by a sequential checkpoint mechanism of ER quality control. *Journal of Cell Biology* 165, 41–52.
- Vembar, S.S., Brodsky, J.L., 2008. One step at a time: Endoplasmic reticulum-associated degradation. *Nature Reviews Molecular Cell Biology* 9, 944–957.
- Verma, R., Chen, S., Feldman, R., *et al.*, 2000. Proteasomal proteomics: Identification of nucleotide-sensitive proteasome-interacting proteins by mass spectrometric analysis of affinity-purified proteasomes. *Molecular Biology of the Cell* 11, 3425–3439.
- Wahlman, J., Demartino, G.N., Skach, W.R., *et al.*, 2007. Real-time fluorescence detection of ERAD substrate retrotranslocation in a mammalian *in vitro* system. *Cell* 129, 943–955.
- Walter, P., Ibrahim, I., Blobel, G., 1981. Translocation of proteins across the endoplasmic reticulum. I. Signal recognition protein (SRP) binds to in-vitro-assembled polysomes synthesizing secretory protein. *Journal of Cell Biology* 91, 545–550.
- Walter, P., Ron, D., 2011. The unfolded protein response: From stress pathway to homeostatic regulation. *Science* 334, 1081–1086.
- Wang, F., Chan, C., Weir, N.R., Denic, V., 2014. The Get1/2 transmembrane complex is an endoplasmic-reticulum membrane protein insertase. *Nature* 512, 441–444.
- Ward, C.L., Omura, S., Kopito, R.R., 1995. Degradation of CFTR by the ubiquitin-proteasome pathway. *Cell* 83, 121–127.
- Waters, M.G., Serafini, T., Rothman, J.E., 1991. 'Coatomer': A cytosolic protein complex containing subunits of non-clathrin-coated Golgi transport vesicles. *Nature* 349, 248–251.

- Weis, B.L., Schleiff, E., Zerges, W., 2013. Protein targeting to subcellular organelles via MRNA localization. *Biochimica et Biophysica Acta* 1833, 260–273.
- Werner, E.D., Brodsky, J.L., McCracken, A.A., 1996. Proteasome-dependent endoplasmic reticulum-associated protein degradation: An unconventional route to a familiar fate. *Proceedings of the National Academy of Sciences of the United States of America* 93, 13797–13801.
- Wiertz, E.J., Jones, T.R., Sun, L., *et al.*, 1996a. The human cytomegalovirus US11 gene product dislocates MHC class I heavy chains from the endoplasmic reticulum to the cytosol. *Cell* 84, 769–779.
- Wiertz, E.J., Tortorella, D., Bogoy, M., *et al.*, 1996b. Sec61-mediated transfer of a membrane protein from the endoplasmic reticulum to the proteasome for destruction. *Nature* 384, 432–438.
- Wilson, D.W., Wilcox, C.A., Flynn, G.C., *et al.*, 1989. A fusion protein required for vesicle-mediated transport in both mammalian cells and yeast. *Nature* 339, 355–359.
- Winget, J.M., Mayor, T., 2010. The diversity of ubiquitin recognition: Hot spots and varied specificity. *Molecular Cell* 38, 627–635.
- Wolf, D.H., Stolz, A., 2012. The Cdc48 machine in endoplasmic reticulum associated protein degradation. *Biochimica et Biophysica Acta* 1823, 117–124.
- Wu, Y., Swulius, M.T., Moremen, K.W., Sifers, R.N., 2003. Elucidation of the molecular logic by which misfolded alpha 1-antitrypsin is preferentially selected for degradation. *Proceedings of the National Academy of Sciences of the United States of America* 100, 8229–8234.
- Xie, W., Kanehara, K., Sayeed, A., Ng, D.T., 2009. Intrinsic conformational determinants signal protein misfolding to the Hrd1/Htm1 endoplasmic reticulum-associated degradation system. *Molecular Biology of the Cell* 20, 3317–3329.
- Xu, C., Wang, S., Thibault, G., Ng, D.T., 2013. Futile protein folding cycles in the ER are terminated by the unfolded protein O-mannosylation pathway. *Science* 340, 978–981.
- Ye, Y., Meyer, H.H., Rapoport, T.A., 2003. Function of the p97-Ufd1-Npl4 complex in retrotranslocation from the ER to the cytosol: Dual recognition of nonubiquitinated polypeptide segments and polyubiquitin chains. *Journal of Cell Biology* 162, 71–84.
- Youker, R.T., Walsh, P., Beilharz, T., Lithgow, T., Brodsky, J.L., 2004. Distinct roles for the Hsp40 and Hsp90 molecular chaperones during cystic fibrosis transmembrane conductance regulator degradation in yeast. *Molecular Biology of the Cell* 15, 4787–4797.
- Zacchi, L.F., Wu, H.C., Bell, S.L., *et al.*, 2014. The BiP molecular chaperone plays multiple roles during the biogenesis of torsinA, an AAA+ ATPase associated with the neurological disease early-onset torsion dystonia. *Journal of Biological Chemistry* 289, 12727–12747.
- Zanetti, G., Pahuja, K.B., Studer, S., Shim, S., Schekman, R., 2012. COPII and the regulation of protein sorting in mammals. *Nature Cell Biology* 14, 20–28.
- Zimmermann, R., Muller, L., Wullich, B., 2006. Protein transport into the endoplasmic reticulum: Mechanisms and pathologies. *Trends in Molecular Medicine* 12, 567–573.
- Zuber, C., Cormier, J.H., Guhl, B., *et al.*, 2007. EDEM1 reveals a quality control vesicular transport pathway out of the endoplasmic reticulum not involving the COPII exit sites. *Proceedings of the National Academy of Sciences of the United States of America* 104, 4407–4412.

Role of Lysosomes in Intracellular Degradation

M Hafner Česen, Jožef Stefan Institute, Ljubljana, Slovenia

V Stoka, Jožef Stefan Institute, Ljubljana, Slovenia and International Postgraduate School Jozef Stefan, Ljubljana, Slovenia

B Turk, Jožef Stefan Institute, Ljubljana, Slovenia and Centre of Excellence for Integrated Approaches in Chemistry and Biology of Proteins, Ljubljana, Slovenia

© 2016 Elsevier Inc. All rights reserved.

Glossary

Cathepsins A group of cysteine, aspartic, and serine peptidases with broad substrate specificity that enables complete degradation of their substrates. In addition, their function can also be more specific, for example, in the immune response.

CLEAR gene network A network of genes coding for various lysosomal proteins with a specific CLEAR sequence and regulated by transcription factor EB.

Lysosomal storage disorders A group of almost 60 genetic diseases caused by deficiencies in genes coding for proteins of the lysosomal system.

Lysosome-related organelles Cell-type specific group of organelles biogenetically similar to lysosomes.

Lysosomes Highly dynamic acidic organelles responsible for degradation of various intra- and extracellular material.

Introduction

Cells are equipped with two major systems for intracellular protein degradation, the ubiquitin-proteasome system and the endolysosomal system (Turk *et al.*, 2012a). The ubiquitin-proteasome system is the major machinery for recycling of short-lived proteins, for degradation of misfolded or damaged proteins within the nucleus and the cytoplasm, thereby representing the quality control of the cell, as well as for the processing of cytosolic antigens for antigen presentation (Kleiger and Mayor, 2014; Blum *et al.*, 2013). On the other hand, the endolysosomal system is responsible for nonspecific degradation of bulk material and organelles to retain energy homeostasis, for processing of exogenous antigens for major histocompatibility complex class II (MHC-II)-mediated immune response, for bone resorption, and processing of other proteins including hormones. In addition, lysosomal proteases are involved in numerous extralysosomal processes in health and disease, including in extracellular protein processing and degradation in various pathological processes (Turk and Turk, 2009; Turk *et al.*, 2012b). In this article the focus will be on the endolysosomal system.

Lysosomes are highly dynamic compartments limited by single-phospholipid bilayer and are found in nearly all eukaryotic cells. They were first described by Christian de Duve and coworkers back in the mid-1950s, a discovery that was awarded with a Nobel Prize for medicine and physiology (de Duve, 1959). They are morphologically heterogeneous, varying from globular to tubular shape with electron-dense interior. Their main characteristics are highly acidic lumen (pH 4.5–5.0) and the absence of mannose-6-phosphate receptor (M6PR), which distinguish them from other less acidic organelles of endocytic pathway, such as early endosomes (EEs) and late endosomes (LEs) or hybrid organelles (Saftig and Klumperman, 2009).

Lysosomes represent the terminal degradative compartments that receive and degrade material from endocytic, phagocytic, and autophagic pathways, providing building

blocks for synthesis of new macromolecules. During the years a lot has been learned about their function. It is now clear that their role is not limited to degradation of bulk material, and that they are also involved in diverse cellular processes like membrane repair, pathogen defense, antigen presentation, cell death, and cell signaling (Figure 1; Luzio *et al.*, 2007; Saftig and Klumperman, 2009; van Kasteren and Overkleeft, 2014). Recently, lysosomes have also been found to have an important role in the regulation of energy metabolism through sensing nutrient availability and activating lysosome-to-nucleus signaling pathway that mediates the starvation response (Settembre *et al.*, 2013). In certain cells, particularly of hematopoietic origin and in melanocytes, the lysosomal catabolic function is complemented with lysosome-related organelles (LROs) or secretory lysosomes, which are biogenetically similar to lysosomes. The best characterized LROs are melanosomes, lytic granules, azurophilic granules, and MHC-II compartments from dendritic cells (Dell'Angelica *et al.*, 2000; Saftig and Klumperman, 2009).

Lysosomes as well as LROs are considered as important regulators of cellular well-being and homeostasis and any abnormality that affects their proper function or turnover can lead to various diseases. Classic examples of the diseases of the lysosomal system are lysosomal storage disorders (LSDs). LSDs are a group of nearly 60, mostly rare genetic diseases and are caused by mutations or deficiencies in genes coding for various lysosomal proteins, which result in accumulation of nondegraded material within the lysosome (Schroder *et al.*, 2010; Schultz *et al.*, 2011). One of the first LSDs was described by Pompe in the early 1930s. It is associated with impaired glycogen degradation, and interestingly, it was this discovery that helped to facilitate the later discovery of the lysosome (Pompe, 1932; Parkinson-Lawrence *et al.*, 2010; de Duve, 1959). The connection between lysosomes and LROs has also been highlighted by studies on multiorganellar genetic LSDs such as Hermansky-Pudlak and Chediak-Higashi syndromes (Dell'Angelica *et al.*, 2000). In addition to LSDs, deregulation of function of lysosomal proteins, such as their increased

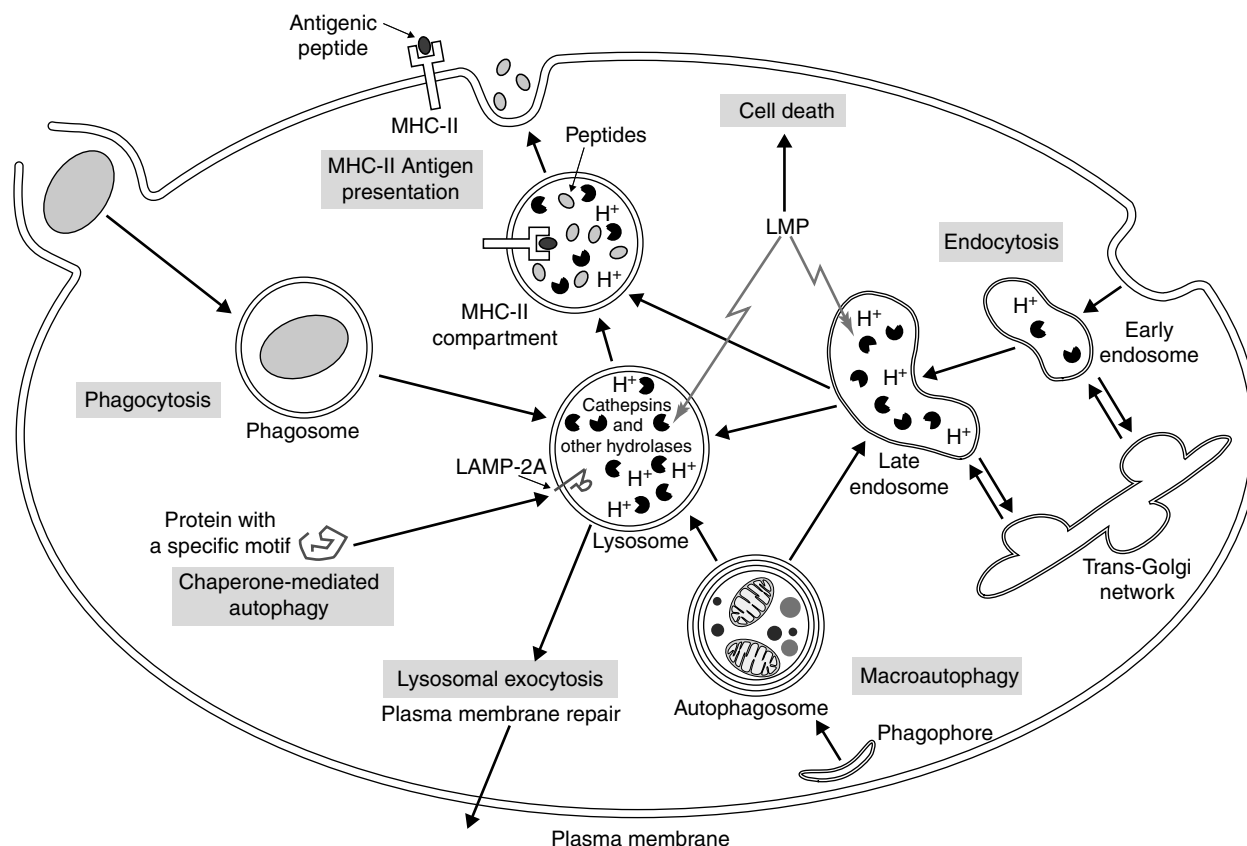


Figure 1 Major functions of lysosome and lysosomal hydrolases in intracellular degradation. Lysosome and lysosomal hydrolases, especially the cathepsins, are involved in numerous cellular processes. They are indispensable in the endocytic and phagocytic pathways, where they degrade various extracellular materials, from fluids and macromolecules to microbial organisms and apoptotic cells. Moreover, lysosomes are involved in the turnover of intracellular material, such as organelles and macromolecules in the process of macroautophagy. In another type of autophagy, CMA, lysosomal proteases degrade cytosolic proteins with a specific amino acid sequence motif, based on which they are transported directly into the lysosomes through the LAMP-2A, which serves as a receptor for proteins destined for CMA-mediated degradation. Lysosomes are important also for plasma membrane repair, which is mediated by the process of lysosomal exocytosis. Lysosomal proteases and lysosomal membrane proteins are also essential in the adaptive immune response, where they are involved in major histocompatibility complex class II (MHC-II)-dependent antigen presentation. However, when the membrane of lysosomes and LEs is permeabilized (LMP) then the released hydrolases, particularly cathepsins, can induce apoptosis, necrosis, or pyroptosis, depending on the cell type and the extent of endosome and lysosome rupture.

expression, combined with their improper localization and/or secretion, is also linked to numerous widespread diseases, including cancer, atherosclerosis, late onset neurodegenerative diseases, rheumatoid arthritis, inflammatory bowel disease, and others (Appelqvist *et al.*, 2013; Maxfield, 2014; Vasiljeva *et al.*, 2007; Reiser *et al.*, 2010).

Lysosomal Proteins

Soluble Hydrolases

As primary catabolic organelles, lysosomes exert their function through nearly 60 different soluble acid-dependent hydrolases such as proteases, phosphatases, nucleases, glycosidases, sulphatases, lipases, and others, which collectively perform degradation of various intra- and extracellular material like sphingolipids, glycogen, glycosaminoglycans, and proteins.

Among the enzymes, proteases have a major role, especially the cathepsins. In addition, several lipases have important roles, such as acid sphingomyelinase (ASMase), acid ceramidase, acid lipase, acid or lysosomal phospholipase A₂ (LPLA₂), as well as several glycosidases including α -galactosidase A, β -glucuronidase, α -iduronidase, glucocerebrosidase, and α -galactosidase (Schroder *et al.*, 2010).

Cathepsins

The name cathepsin originates from Greek word 'kathapsein' which means to digest. Cathepsins are divided into three groups based on the amino acid present in the active site: cysteine (in humans we can find cathepsin B, C, or dipeptidyl peptidase I, F, H, K, L, O, S, V, W, and X), aspartic (D and E), and serine cathepsins (A and G). The majority of them can be found in most living organisms and ubiquitously expressed

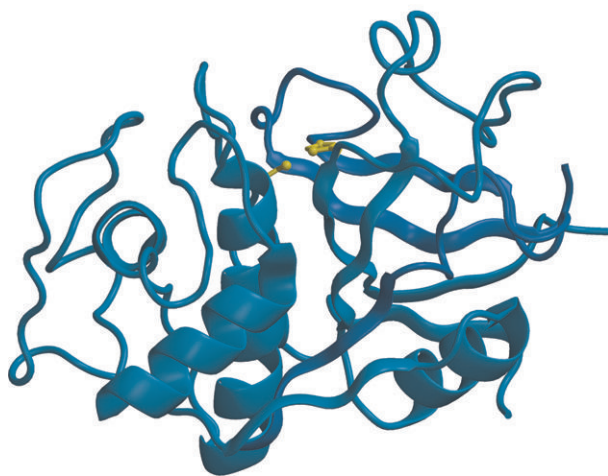


Figure 2 Fold of mature cathepsin L, a typical cysteine cathepsin endopeptidase. The fold of the two-chain form of native cathepsin L (1icf) is shown in the ribbon representation with heavy chain in light blue and light chain in dark blue. The side chains of the active-site cysteine (Cys25) and histidine (His163) residues, each residing on one side of the active site cleft, are indicated in a ball-and-stick representation.

among different tissues. However, some of them are cell-type specific, such as cathepsin K or cathepsin S and W, which are mainly found in osteoclasts or in specific immune cells, respectively. Likewise, cathepsin V is expressed predominantly in the thymus and testicles, and cathepsin G is found only in the azurophilic granules in neutrophils. The cell-type specific localization suggests a more specific role for the cathepsins in these cells. Furthermore, the members of the cysteine group and the aspartic cathepsin D are found throughout the endolysosomal system, whereas the aspartic cathepsin E is found only in the endosomes (Repnik *et al.*, 2012; Turk *et al.*, 2012b).

Cysteine cathepsins belong to the clan CA of cysteine peptidases, more specifically to the C1 family of papain-like enzymes, the largest and, in addition to the aspartic cathepsin D, the most studied group of lysosomal proteases. Cysteine cathepsins are predominantly endopeptidases (Figure 2), with the exception of cathepsins B and H, which are endo- as well as exopeptidases, and cathepsin X and C, which are strictly exopeptidases. Exopeptidase activities of these cathepsins are consequences of different structural elements, such as the occluding loop and mini loop in carboxypeptidases cathepsins B and X, respectively, and the remaining parts of their proregions such as seen in aminopeptidases cathepsins H and C, which protrude in the active site cleft, thereby sterically hindering substrate binding (Turk *et al.*, 2012a). Moreover, cysteine cathepsins exhibit broad substrate specificity, which combined with their endo- and exopeptidase activities ensures complete degradation of their substrates (Turk *et al.*, 2012b). Nevertheless, cathepsins can also have more specific functions within lysosomes or LROs, such as found for cathepsin C, which has a major role in processing of numerous other granule proteases, including granzymes A and B, neutrophil elastase, cathepsin G, and chymase (Turk *et al.*, 2001).

In order to be optimally active, lysosomal hydrolases, including cathepsins, require a reducing, slightly acidic

environment, such as found in the endolysosomal system. The acidic pH and reducing conditions in the endolysosomal system are not only optimal for cathepsin activity and stability, but also help in loosening the structure of macromolecules, which increases the number of potential cleavage sites, thereby facilitating their degradation. Because of their enormous degradative potential, their action has to be well regulated and, importantly, confined to the acidic vesicles. The first and essential barrier that separates the acidic interior-rich hydrolases from the rest of the cell is the 7–10 nm thick lysosomal membrane, which is well protected against degradation by heavily glycosylated luminal domains of membrane proteins that form a thick glycocalyx (Granger *et al.*, 1990). Secondly, all cathepsins are synthesized as inactive precursors, which can be largely activated only in the acidic environment of the endolysosomal system. Their activation can be autocatalytic or catalyzed by other proteases, including cathepsin D, and can be largely facilitated by glycosaminoglycans. Moreover, cysteine cathepsins are relatively unstable at neutral pH due to irreversible unfolding, which provides another level of protection. The exception is cathepsin S, which remains active for several hours at neutral pH. Particularly sensitive to neutral pH are the aspartic cathepsins because of the reversible deprotonation of Asp residues in the active site. Nevertheless, their limited activity at neutral pH, such as found within the cytosol or in the extracellular milieu seems to be sufficient for substrate cleavage and thus proteolytic signaling even under such unfavorable conditions, which can be very harmful and can have devastating consequences (Repnik *et al.*, 2012; Turk *et al.*, 2012b; Fonović and Turk, 2014). A major level of control of such escaped cathepsins is provided by their endogenous inhibitors present in the cytosol or extracellularly, such as stefins A and B, certain serpins, or cystatin C (Turk *et al.*, 2002, 2012b).

Cathepsins in the Immune Response

Although cathepsins were long believed to participate exclusively in nonspecific protein degradation, it is now clear that the concept of ‘regulation through limited destruction’ takes part in various physiological processes, in particular in the immune response. Especially well characterized is their role in MHC-II antigen processing and presentation, although they are implicated also in cytokine regulation, natural killer T cell development, and other processes (Colbert *et al.*, 2009).

In the adaptive immune response foreign antigens are internalized either through endocytosis or phagocytosis and degraded into peptides, which are recruited to the MHC-II molecules that present them to the CD4⁺ T-helper cells on the cell surface. In comparison to MHC-I molecules, which bind peptides generated by the proteasome system in the endoplasmic reticulum (ER), MHC-II molecules mostly bind peptides generated by endolysosomal proteolysis. Contrary to antigen processing, which is generally nonspecific process with a great amount of redundancy, processing of the MHC-II invariant chain is a highly controlled and coordinated process. MHC-II molecules are assembled within the ER, whereas the maturation occurs in endosomal compartments, which are rich with antigenic peptides. MHC-II molecules consist of

α and β subunits, and the invariant chain (Ii), responsible for assembly of the MHC-II complex and targeting it to LEs, which occurs either through trans-Golgi network (TGN) or endocytosis. Due to alternative splicing, several forms of Ii exist, and the longest, the p41 Ii variant, possesses a glycosylated domain, which was shown to inhibit cathepsin L and thereby regulate its activity in antigen presentation. Processing of MHC-II invariant chain occurs through several consecutive C-terminal cleavages that culminate in a variably extended peptide of approximately 20 residues called CLIP (class II-associated invariant chain peptide). Multiple cathepsins have been shown to be involved in the Ii processing; however, their function seems to be cell-type specific. In dendritic and B cells cathepsin S was thus found to play a crucial role in the final stages of processing, while in cortical thymic epithelial cells cathepsin L (V in humans) perform the processing. The CLIP–MHC-II complex is then, at least in humans, recognized by the human leukocyte antigen (HLA)-DM chaperone that is highly homologous to the MHC-II molecule and facilitates the substitution of the CLIP fragment with the antigenic peptides and its subsequent presentation to the CD4⁺ T-helper cells (Blum *et al.*, 2013; Cresswell, 1996; Turk *et al.*, 2012a).

In addition to immune response, cathepsins participate also in bone remodeling, keratinocyte differentiation, and prohormone activation for neuropeptide biosynthesis (Turk *et al.*, 2012b). They can also be found within the nucleus, where they regulate cell cycle (Duncan *et al.*, 2008), as well as in the extracellular space, where they participate in numerous physiological processes like bone remodeling and wound healing (Fonović and Turk, 2014).

However, excessive cathepsin activity is often associated with diseases. In most of these diseases, cathepsins are secreted into the extracellular milieu, often from the immune cells infiltrated to the disease site and therefore represent markers for numerous inflammation-associated diseases, including cancer, rheumatoid arthritis and osteoarthritis, atherosclerosis, and inflammatory bowel diseases. In addition, cathepsin K has a major role in osteoporosis. Therefore, cathepsins represent important targets for drug development and drugs targeting cathepsin K have successfully finished Phase III clinical trials for osteoporosis treatment (Vasiljeva *et al.*, 2007; Reiser *et al.*, 2010; Fonović and Turk, 2014).

Lysosomes and Lysosomal Cathepsins in Cell Death

The deadly potential of lysosomes was first noticed by Christian de Duve, who referred to lysosomes as ‘suicide bags’ (de Duve, 1959), but it was long after this discovery that the first cytosolic targets of cysteine cathepsins and the mechanism of lysosomal apoptotic pathway were revealed (Blomgran *et al.*, 2007; Cirman *et al.*, 2004; Droga-Mazovec *et al.*, 2008; Stoka *et al.*, 2001). In healthy cells, lysosomal cathepsins normally reside inside the acidic environment of endosomes and lysosomes, which prevents the undesired proteolysis. However, when their membrane is damaged and the inhibitors in the cytosol fail to inactivate the released cathepsins, then they are fatal for the cell and induce apoptosis, necrosis, or pyroptosis, depending on the cell type and the extent of endosome and lysosome rupture (Repnik *et al.*, 2014).

Cysteine cathepsins have the potential to trigger lysosomal apoptosis if the lysosomal membrane is directly damaged, or, alternatively, they can amplify the death signal if apoptosis is initiated outside the endocytic compartment as often seen in late stages of the process (Repnik *et al.*, 2012). One group of compounds that directly target lysosomal membrane are lysosomotropic detergents, which are lipophilic bases that accumulate in acidic vesicles, namely endosomes, lysosomes, and hybrid organelles, where after protonation they become trapped and develop membranolytic properties, which results in the release of the cathepsins into the cytosol (de Duve *et al.*, 1974; Miller *et al.*, 1983). Although the name suggests that lysosomotropic detergents target only lysosomes, they actually affect all acidic vesicles in the cell, i.e., the vesicles of the late endocytic pathway, including LEs, lysosomes, and hybrid organelles. The best example of a lysosomotropic detergent is the dipeptidyl ester LeuLeuOMe, which gains detergent-like properties after polymerization with cathepsin C (Thiele and Lipsky, 1990). LeuLeuOMe studies represent the basis on which the mechanism of lysosomal apoptotic pathway has been determined (Cirman *et al.*, 2004; Droga-Mazovec *et al.*, 2008; Stoka *et al.*, 2001; Uchimoto *et al.*, 1999). The crucial event in lysosomal apoptotic pathway is lysosomal membrane permeabilization (LMP) that enables the release of hydrolases into the cytosol, where they perform rather specific proapoptotic cleavages that trigger a signaling cascade ending with cell death. Cysteine cathepsins (Cirman *et al.*, 2004; Stoka *et al.*, 2001) and aspartic cathepsin D (Appelqvist *et al.*, 2012) have been shown to be the most important lysosomal proteases in the processing of the proapoptotic Bid protein, thereby generating the proapoptotic p15 tBid form (Repnik *et al.*, 2012). tBid is involved in oligomerization of proapoptotic Bak and Bax proteins, which form pores in the outer mitochondrial membrane and thereby release cytochrome *c* into the cytosol. This is also the stage where the lysosomal apoptotic pathway merges with the classical intrinsic apoptotic pathway. In the cytosol, cytochrome *c* and apoptotic protease-activating factor (Apaf-1) form the initiator caspase-9-activating platform named apoptosome. Activated caspase-9 can then proteolytically activate executioner caspase -3, -6 and -7, leading to execution of cell death (Repnik *et al.*, 2014). In addition to Bid activation, cysteine cathepsins also degrade antiapoptotic Bcl-2 molecules and X-chromosome-linked inhibitor of apoptosis (XIAP), thereby facilitating apoptotic cell death (Droga-Mazovec *et al.*, 2008). In addition, other LMP-mediated cell death pathways have been described, but the mechanisms are less clear (Kirkegaard and Jaattela, 2009; Aits and Jaattela, 2013).

Lipases

Among the lipids, best described are sphingolipids, which are involved in a variety of cellular processes, such as cell proliferation, senescence, differentiation, and apoptosis, so their cellular levels must be tightly regulated, which is achieved by a coordinated action of several specific lipases. For example, elevated activity of ASMase and acid ceramidase that convert sphingolipid sphingomyelin to ceramide and sphingosine, respectively, can influence lysosomal membrane stability and

may lead to cell death (Kågedal *et al.*, 2001). Recently, an additional function for ASMase in lysosomal function has been revealed, suggesting that ASMase deficiency impairs autophagy-lysosomal degradation pathway and increases lysosomal cholesterol that in turn increases lysosomal membrane stability (Fucho *et al.*, 2014). On the other hand, inherited mutations in ASMase cause accumulation of sphingomyelin in lysosomes, a condition called Niemann–Pick disease types A and B (Stern, 2014). Furthermore, acid lipase, essential for hydrolysis of cholesterol esters and triglycerides, has also recently been shown to be indispensable for the alternative activation of macrophages after parasite infection (Huang *et al.*, 2014), whereas the genetic deficiency in the enzyme results in autosomal recessive disorders like Wolman disease and cholesteryl ester storage disease (Porto, 2014). Another well-known lipase is LPLA₂, which has a ubiquitous role in lysosomal phospholipid degradation. Negatively charged glycerophospholipids, particularly bis (monoacylglycero)phosphate (BMP), have been found to significantly augment the LPLA₂ activity and to work as a platform for the lipid–water interfacial reaction of lysosomal lipolytic enzymes in intra-endosomal and intra-lysosomal membranes (Shayman *et al.*, 2011). Until now, no defined syndrome has been linked to LPLA₂ deficiency; however, the likelihood that the loss of the enzyme activity causes a disease is very high. Several drugs have been shown to cause a toxic condition, so-called cellular phospholipidosis, as a result of impaired degradation of phospholipids. Such drugs are called cationic amphiphilic drugs (CAD), because they contain an aromatic group and an amine that is protonated under acidic conditions. Despite the importance of the condition, the molecular mechanism that induces it and its pathological consequences at the cell and organ levels still remain unclear (Shayman *et al.*, 2011).

Glycosidases

In addition to proteases and lipases, glycosidases are indispensable for normal lysosomal function. They are responsible for the turnover of numerous endocytosed glycosphingolipids, like ceramides with variable oligosaccharide moieties, which are ubiquitously expressed on the cell surface. Two inherited prototypical glycosphingolipidoses, namely Gaucher and Fabry diseases, which were first already described in the nineteenth century (Ferraz *et al.*, 2014), further stress this importance. The diseases are caused by deficiencies in glucocerebrosidase and α -galactosidase A, respectively, and occur with a prevalence of 1:40 000–100 000. Interestingly, Gaucher disease is even more abundant in Ashkenazi Jewish population, where it occurs with a frequency of 1:900 (Cassinerio *et al.*, 2014). The diseases are classic examples of LSDs, where the non-hydrolyzed substrate, glucosylceramide (glucocerebroside) or globotriaosylceramide, accumulates within the lysosome (Ferraz *et al.*, 2014). Similarly, deficiency in the *N*-acetylglucosamine-1-phosphotransferase, responsible for synthesis of the mannose-6-phosphate (M6P) hydrolase targeting signal for lysosomes, results in mucopolisaccharidosis II and III, as a consequence of improper localization of the enzyme (Idol *et al.*, 2014).

Lysosomal Membrane Proteins

More than 120 different lysosomal membrane proteins are needed to mediate numerous functions of the lysosomal membrane. Among them the best characterized is a transmembrane multiprotein complex, vacuolar H⁺-ATPase, responsible for the acidification of the organelle's interior by using the energy derived from ATP hydrolysis to transport protons across the lysosomal membrane. In addition, there is a complex machinery of membrane proteins responsible for protein import and export from and to the cytosol, as well as proteins for lysosomal trafficking and fusion, which are mediated by a machinery of membrane-associated RAB GTPases and SNARE (SNAP (Soluble NSF Attachment Protein) REceptor) proteins. Especially important lysosomal membrane structural proteins are the so-called lysosome-associated membrane proteins (LAMPs), which regulate macroautophagy, chaperone-mediated autophagy (CMA), and direct lysosomal transport for cytosolic proteins (Ruivo *et al.*, 2009; Saftig and Klumperman, 2009). Recently, it has been shown that LAMPs regulate also the positioning of lysosomes and mitochondria, which subsequently affects the energy metabolism (Rajapakshe *et al.*, 2014). The most abundant structural protein is LAMP-1, which accounts for 50% of the total protein in the membrane (Settembre *et al.*, 2013). Deficiency in another LAMP, LAMP-2A (lysosome-associated membrane protein 2A), is known to cause Danon disease, an X-linked lysosomal disease, associated with the accumulation of autophagic vacuoles in muscle cells, resulting in cardiomyopathy, skeletal myopathy, and mental retardation. Interestingly, until now more than 60 different mutations have been found in the LAMP-2 protein only (Cheng and Fang, 2012). A mutation of another lysosomal protein Niemann–Pick C1 protein 1, involved in the export of cholesterol from the endolysosomal compartment, results in an autosomal recessive lipid storage disorder, namely Niemann–Pick disease type C, characterized by progressive neurodegeneration. Another mutation in the lysosomal membrane protein called heparin- α glucosaminide *N*-acetyltransferase impairs the stepwise degradation of heparin sulfate, resulting in mucopolysaccharidosis type IIIC (Settembre *et al.*, 2013). These and many other disease-causing mutations illustrate the importance of lysosomal membrane proteins in addition to soluble lysosomal hydrolases.

Lysosomal Biogenesis and Its Regulation

Biogenesis of lysosomal soluble and membrane proteins requires interactions between the biosynthetic and endocytic pathways. Proteins are synthesized in the ER and then transported to the TGN, from where they are targeted to the lysosome in a signal-dependent manner either directly from the TGN or indirectly through the secretory pathway to the plasma membrane and subsequent endocytosis into the cell. The majority of the lysosomal hydrolases are modified with M6P and in the TGN selected by means of M6PR for transport to endosomes (Bräulke and Bonifacino, 2009). For example, cathepsins are synthesized as inactive zymogens (pre-pro-enzymes), which are following the co-translational removal of the pre-peptide and posttranslational glycosylation in the

Golgi apparatus directed toward endosomes, where the zymogens are activated. The acidic environment of the endosome enables dissociation of the M6PR from the hydrolase. The cathepsins are then transferred to lysosomes either through endosome–lysosome fusion or through endosome maturation (Katunuma, 2010). Other soluble hydrolases and nonenzyme proteins are transferred to lysosomes independently of M6PR by alternative transport receptors, such as sortilin as a receptor for prosaposin and ASMase or lysosomal integral membrane protein LIMP-2 as a receptor for β -glucocerebrosidase (Schroder *et al.*, 2010). Compared to the soluble lysosome hydrolases, much less is known about the transport of lysosomal membrane proteins to the lysosome. They contain specific sorting sequences in the cytoplasmic part of the protein known as tyrosine- and dileucine-based motifs responsible for binding of different adaptor protein complexes, which initiate packaging in clathrin-coated vesicles (Braulke and Bonifacino, 2009; Schroder *et al.*, 2010).

A few years ago, a specific gene network, so-called coordinated lysosomal expression and regulation (CLEAR), has been discovered and many genes coding for the lysosomal proteins have been found to contain the CLEAR sequence (GTCACGTGAC). The CLEAR gene network and its master regulator transcription factor EB (TFEB) are crucial regulators of lysosomal biogenesis and function, showing that lysosomal function can be globally controlled (Sardiello *et al.*, 2009). Furthermore, it has been shown that lysosomal function can be coordinated also as a consequence of environmental stimuli. For example, upon starvation TFEB enters the nucleus and binds to CLEAR elements and triggers the transcription of genes important for autophagy as well as for lysosomal biogenesis (Settembre *et al.*, 2013). In addition, non-lysosomal proteins that regulate the functions of lysosome-resident proteins, such as cation-dependent (CD-M6PR) and cation-independent M6PR (CI-M6PR) are of major importance. These proteins circle between TGN and LEs and are responsible for targeting of lysosomal hydrolases to the lysosome (Settembre *et al.*, 2013).

Pathways to the Lysosome

Endocytosis

Extracellular material reaches the lysosomes mostly through the endocytic machinery. Cells have developed several mechanisms of endocytosis, such as phagocytosis, macropinocytosis, clathrin-dependent and clathrin-independent endocytosis, calveolae or calveolin-mediated endocytosis, depending on the physical and chemical nature of the cargo. In general, particles larger than 500 nm enter the cell via phagocytosis or macropinocytosis, whereas smaller material is internalized via other endocytic mechanisms. The best characterized are receptor-mediated clathrin-dependent endocytosis, which dominates and defines the paradigm for endocytic pathway, and phagocytosis, which is important for intake of microbial organisms, apoptotic cells, and cellular debris (Doherty and McMahon, 2009; Kumari *et al.*, 2010).

The process of endocytosis starts on the cell surface where cells internalize fluids, macromolecules, plasma membrane

components, and other particles by invagination of the plasma membrane. The cargo then passes through different endosomal vesicles and vacuoles, so-called endosomal intermediates, through membrane fission until it reaches the final destination, the lysosome (Huotari and Helenius, 2011). An important hallmark of the pathway is the progressive decrease of the luminal pH from EEs (pH ~6) and LEs (pH 5.0–6.0) to lysosomes (pH 4.5–5.0), which is optimal for the activity of hydrolytic enzymes (Saftig and Klumperman, 2009; Turk and Turk, 2009).

The recipients of the incoming cargo from the cellular surface are EEs, which represent the main sorting station of the endocytic pathway. In the EEs the majority of the internalized cargo, such as plasma membrane components and their ligands, is recycled back to the cell surface, the process important for regulation of the cell-surface expression of membrane proteins. The cargo that is not recycled to the cell surface is carried in EEs further along the endocytic pathway to the LEs and finally to the lysosomes either for degradation or for lysosomal biogenesis. Interestingly, a relatively small fraction of internalized material actually reaches the lysosomes due to a stringent selection in the EEs and later in LEs, allowing only a specific cargo to be transported and degraded in the lysosomes. In addition, EEs receive the cargo through bidirectional vesicle exchange with TGN, such as M6PR bound to newly synthesized lysosomal hydrolases, for transport to LEs and lysosomes. In the slightly acidic EEs, M6PR dissociate from the hydrolases and return back to the TGN. In addition, this transport is important for removal of the used components during endosome maturation and also occurs between TGN and maturing LEs, but seldom after the fusion with lysosome. The proteins that enable the bidirectional transport are Rab7, Rab9, and a complex of nexins. An essential component of the membrane of the EE is Rab5 protein, which follows EE through various stages of maturation and is also the major regulator of EE conversion to LE. After the formation in the peripheral cytoplasm, LEs move to the perinuclear region where they undergo transient fusions with other LEs to form larger bodies and eventually fuse with lysosomes or hybrid organelles, called endolysosomes. Finally, by fusing with lysosomes they follow a unidirectional pathway, reaching a point of no return, where some of the components get degraded, whereas the others contribute to lysosomal biogenesis (Huotari and Helenius, 2011; Saftig and Klumperman, 2009).

The endocytic pathway is an extremely dynamic and often elusive process because of the continuous maturation, transformation, fusion, and fission of the organelles. The distinction between the organelles is therefore often difficult, since the majority of proteins are either transiently associated with a certain organelle or they follow it all the way through transformation. Up to now, no specific markers for lysosomes have been identified. LAMPs and acid hydrolases are found both in lysosomes and LEs, however, some proteins, such as cation-independent M6PR, Rab7, and the regulatory (RII) domain of the cAMP-dependent protein kinase, are present only on LEs and therefore enable some discrimination (Repnik *et al.*, 2013).

However, it is worth noting that recent findings suggest that the machinery used for endocytosis is also shared in autophagy, including in regulation of membrane budding and

fusion, role of the lysosome in sensing the nutrient status of the cell, regulation of mammalian target of rapamycin complex 1 (mTORC1) activity, and initiation of autophagy (Lamb *et al.*, 2013). Moreover, there is a dynamic relationship between autophagy and phagocytosis (Oczypok *et al.*, 2013; Vernon and Tang, 2013), including newly discovered autophagic-phagocytosis 'hybrid' processes such as microtubule-associated protein 1 light chain 3-associated phagocytosis (LAP) (Vernon and Tang, 2013). Furthermore, autophagy and phagocytosis share many common features and act in a concerted manner during the clearance of dead cells or autophagy, thus acting as a backup system to fight off infection upon phagocytosis failure to eliminate a pathogen (Eskelinen and Saftig, 2009; Oczypok *et al.*, 2013).

Autophagy

Intracellular material reaches the lysosomes for their degradation through three different types of autophagy, namely, macroautophagy (Eskelinen and Saftig, 2009; Yang and Klionsky, 2010), microautophagy (Mijaljica *et al.*, 2011), and CMA (Wong and Cuervo, 2010; Park and Cuervo, 2013).

(Macro)autophagy is a bulk degradation process that mediates the clearance of macromolecules, damaged organelles, and aggregate-prone proteins (Yang and Klionsky, 2010; Eskelinen and Saftig, 2009; Ravikumar *et al.*, 2010). It is an important survival mechanism that is upregulated upon starvation, oxidative stress, or other harmful conditions. It is initiated by double-membraned structures, which engulf portions of cytoplasm. The magnitude of autophagosome formation is tightly regulated by intracellular and extracellular amino acid concentrations and ATP levels via signaling pathways that include the nutrient-sensing kinase TOR (target of rapamycin). The resulting autophagosomes ultimately fuse with lysosomes, where their contents are degraded. The formation of autophagosomes is tightly regulated by a specific set of autophagy-related (ATG) genes (Yang and Klionsky, 2010; Mizushima *et al.*, 2011). Recently, a systematic proteomic analysis of the autophagy interaction network in human cells under conditions of ongoing (basal) autophagy provided a global view of the mammalian autophagy interaction landscape (Behrends *et al.*, 2010). Interestingly, recent findings suggest that macroautophagy can also occur in the absence of some of these key autophagy proteins through the unconventional biogenesis of canonical autophagosomes (Codogno *et al.*, 2012).

Although autophagy has long been considered to be a nonselective bulk degradation pathway, in the last few years several forms of selective autophagy have been identified, leading to degradation of specific organelles, proteins, and pathogens in yeast, flies, and mammals (Kraft *et al.*, 2009). Two types of cargo selection by autophagy have been described, namely, nonselective and selective autophagy. During the nonselective process, bulk cytosol and other cytoplasmic components are randomly sequestered into autophagosomes, whereas during selective macroautophagy, a specific cargo is exclusively enwrapped into the double-membrane vesicle (Kraft *et al.*, 2009). Selective types of autophagy include the selective degradation of cytoplasmic proteins during bulk autophagy, the Cvt pathway and organellophagy such as mitophagy, pexophagy, reticulophagy, ribophagy, and

piecemeal autophagy of the nucleus (Kraft *et al.*, 2009; Mijaljica *et al.*, 2011; Okamoto, 2014). Moreover, selective types of autophagy only described in higher eukaryotes include aggrephagy, xenophagy, and MHC-II crosspresentation (Johansen and Lamark, 2011; Kraft *et al.*, 2009). Although, all selective types of autophagy use the conserved core autophagy machinery, they also possess a mechanism that assures the accuracy in the specific selection of the cargo that has to be eliminated (Kraft *et al.*, 2009). Available evidence suggests two general mechanisms for cargo recognition. First, a factor on the surface of the cargo molecule might become available to bind an adaptor or receptor, i.e., p62/SQSTM1 and NBR1 that carry out an identical function (Kraft *et al.*, 2009; Johansen and Lamark, 2011). The second principle for selective cargo recognition is the involvement of ubiquitin as a signaling molecule for targeting various types of cargo, ranging from protein aggregates to membrane-bound organelles and microbes (Kirkin *et al.*, 2009; Kraft *et al.*, 2009; Shaid *et al.*, 2013).

The identification of autophagy receptors, such as p62/SQSTM1 and NBR1, which simultaneously bind both ubiquitin and autophagy-specific ubiquitin-like modifiers, LC3 (light chain 3) or GABARAP (gamma-aminobutyric acid receptor-associated protein), has provided a molecular link between ubiquitination and autophagy (Shaid *et al.*, 2013; Kirkin *et al.*, 2009). Although these two systems bear some unique properties, there are a series of essential steps and components common to both of them and required for their functions in cellular quality control. The common steps in protein degradation are cargo selection and tagging, cargo recognition and delivery to the proteolytic machinery, degradation in the proteolytic core, and recycling of the constituent amino acids (Wong and Cuervo, 2010). Moreover, evidence supports different levels of cross talk among autophagic pathways and the ubiquitin-proteasome system where changes in one proteolytic pathway affect other proteolytic systems, components of one proteolytic pathway are degraded by another proteolytic system, and different proteolytic pathways share components and substrates (Park and Cuervo, 2013).

See also: Protein Synthesis/Degradation: Protein Degradation – Protease Classes: Cathepsin E: An Aspartic Protease with Diverse Functions and Biomedical Implications

References

- Aits, S., Jaattela, M., 2013. Lysosomal cell death at a glance. *Journal of Cell Science* 126, 1905–1912.
- Appelqvist, H., Johansson, A.C., Linderöth, E., *et al.*, 2012. Lysosome-mediated apoptosis is associated with cathepsin D-specific processing of bid at Phe24, Trp48, and Phe183. *Annals of Clinical Laboratory Science* 42, 231–242.
- Appelqvist, H., Waster, P., Kagedal, K., Ollinger, K., 2013. The lysosome: From waste bag to potential therapeutic target. *Journal of Molecular Cell Biology* 5, 214–226.
- Behrends, C., Sowa, M.E., Gygi, S.P., Harper, J.W., 2010. Network organization of the human autophagy system. *Nature* 466, 68–76.
- Blomgran, R., Zheng, L., Stendahl, O., 2007. Cathepsin-cleaved Bid promotes apoptosis in human neutrophils via oxidative stress-induced lysosomal membrane permeabilization. *Journal of Leukocyte Biology* 81, 1213–1223.

- Blum, J.S., Wearsch, P.A., Cresswell, P., 2013. Pathways of antigen processing. *Annual Review of Immunology* 31, 443–473.
- Braulke, T., Bonifacio, J.S., 2009. Sorting of lysosomal proteins. *Biochimica et Biophysica Acta* 1793, 605–614.
- Cassinero, E., Graziadei, G., Poggiali, E., 2014. Gaucher disease: A diagnostic challenge for internists. *European Journal of Internal Medicine* 25, 117–124.
- Cheng, Z., Fang, Q., 2012. Danon disease: Focusing on heart. *Journal of Human Genetics* 57, 407–410.
- Cirman, T., Orešić, K., Droga-Mazovec, G., *et al.*, 2004. Selective disruption of lysosomes in HeLa cells triggers apoptosis mediated by cleavage of Bid by multiple papain-like lysosomal cathepsins. *Journal of Biological Chemistry* 279, 3578–3587.
- Codogno, P., Mehrpour, M., Proikas-Cezanne, T., 2012. Canonical and non-canonical autophagy: Variations on a common theme of self-eating? *Nature Reviews Molecular Cell Biology* 13, 7–12.
- Colbert, J.D., Matthews, S.P., Miller, G., Watts, C., 2009. Diverse regulatory roles for lysosomal proteases in the immune response. *European Journal of Immunology* 39, 2955–2965.
- Cresswell, P., 1996. Invariant chain structure and MHC class II function. *Cell* 84, 505–507.
- de Duve, C., 1959. Lysosomes, a new group of cytoplasmic particles. In: Hayashi, T. (Ed.), *Subcellular Particles*. New York, NY: The Ronald Press Co.
- de Duve, C., de Barse, T., Poole, B., *et al.*, 1974. Commentary. lysosomotropic agents. *Biochemical Pharmacology* 23, 2495–2531.
- Dell'Angelica, E.C., Mullins, C., Caplan, S., Bonifacio, J.S., 2000. Lysosome-related organelles. *Faseb Journal* 14, 1265–1278.
- Doherty, G.J., McMahon, H.T., 2009. Mechanisms of endocytosis. *Annual Review of Biochemistry* 78, 857–902.
- Droga-Mazovec, G., Bojić, L., Petelin, A., *et al.*, 2008. Cysteine cathepsins trigger caspase-dependent cell death through cleavage of bid and antiapoptotic Bcl-2 homologues. *Journal of Biological Chemistry* 283, 19140–19150.
- Duncan, E.M., Muratore-Schroeder, T.L., Cook, R.G., *et al.*, 2008. Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation. *Cell* 135, 284–294.
- Eskelinen, E.L., Saftig, P., 2009. Autophagy: A lysosomal degradation pathway with a central role in health and disease. *Biochimica et Biophysica Acta* 1793, 664–673.
- Ferraz, M.J., Kallmeijer, W.W., Mirzaian, M., *et al.*, 2014. Gaucher disease and Fabry disease: New markers and insights in pathophysiology for two distinct glycosphingolipidoses. *Biochimica et Biophysica Acta* 1841, 811–825.
- Fonović, M., Turk, B., 2014. Cysteine cathepsins and extracellular matrix degradation. *Biochimica et Biophysica Acta* 1840, 2560–2570.
- Fucho, R., Martinez, L., Baulies, A., *et al.*, 2014. ASMase regulates autophagy and lysosomal membrane permeabilization and its inhibition prevents early stage non-alcoholic steatohepatitis. *Journal of Hepatology* 61, 1126–1134.
- Granger, B.L., Green, S.A., Gabel, C.A., *et al.*, 1990. Characterization and cloning of lgp110, a lysosomal membrane glycoprotein from mouse and rat cells. *Journal of Biological Chemistry* 265, 12036–12043.
- Huang, S.C., Everts, B., Ivanova, Y., *et al.*, 2014. Cell-intrinsic lysosomal lipolysis is essential for alternative activation of macrophages. *Nature Immunology* 15, 846–855.
- Huotari, J., Helenius, A., 2011. Endosome maturation. *EMBO Journal* 30, 3481–3500.
- Idol, R.A., Wozniak, D.F., Fujiwara, H., *et al.*, 2014. Neurologic abnormalities in mouse models of the lysosomal storage disorders mucopolipidosis II and mucopolipidosis III gamma. *PLoS One* 9, e109768.
- Johansen, T., Lamark, T., 2011. Selective autophagy mediated by autophagic adapter proteins. *Autophagy* 7, 279–296.
- Kågedal, K., Zhao, M., Svensson, I., Brunk, U.T., 2001. Sphingosine-induced apoptosis is dependent on lysosomal proteases. *Biochemical Journal* 359, 335–343.
- van Kasteren, S.I., Overkleeft, H.S., 2014. Endo-lysosomal proteases in antigen presentation. *Current Opinion in Chemical Biology* 23C, 8–15.
- Katunuma, N., 2010. Posttranslational processing and modification of cathepsins and cystatins. *Journal of Signal Transduction* 2010, 1–8.
- Kirkegaard, T., Jaattela, M., 2009. Lysosomal involvement in cell death and cancer. *Biochimica et Biophysica Acta* 1793, 746–754.
- Kirkin, V., McEwan, D.G., Novak, I., Dikic, I., 2009. A role for ubiquitin in selective autophagy. *Molecular Cell* 34, 259–269.
- Kleiger, G., Mayor, T., 2014. Perilous journey: A tour of the ubiquitin-proteasome system. *Trends in Cell Biology* 24, 352–359.
- Kraft, C., Reggiori, F., Peter, M., 2009. Selective types of autophagy in yeast. *Biochimica et Biophysica Acta* 1793, 1404–1412.
- Kumari, S., Mg, S., Mayor, S., 2010. Endocytosis unplugged: Multiple ways to enter the cell. *Cell Research* 20, 256–275.
- Lamb, C.A., Dooley, H.C., Tooze, S.A., 2013. Endocytosis and autophagy: Shared machinery for degradation. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 35, 34–45.
- Luzio, J.P., Pryor, P.R., Bright, N.A., 2007. Lysosomes: Fusion and function. *Nature Reviews Molecular Cell Biology* 8, 622–632.
- Maxfield, F.R., 2014. Role of endosomes and lysosomes in human disease. *Cold Spring Harbor Perspectives in Biology* 6, 1–15.
- Mijaljica, D., Prescott, M., Devenish, R.J., 2011. Microautophagy in mammalian cells: Revisiting a 40-year-old conundrum. *Autophagy* 7, 673–682.
- Miller, D.K., Griffiths, E., Lenard, J., Firestone, R.A., 1983. Cell killing by lysosomotropic detergents. *Journal of Cell Biology* 97, 1841–1851.
- Mizushima, N., Yoshimori, T., Ohsumi, Y., 2011. The role of Atg proteins in autophagosome formation. *Annual Review of Cell and Developmental Biology* 27, 107–132.
- Oczypok, E.A., Oury, T.D., Chu, C.T., 2013. It's a cell-eat-cell world: Autophagy and phagocytosis. *The American Journal of Pathology* 182, 612–622.
- Okamoto, K., 2014. Organellophagy: Eliminating cellular building blocks via selective autophagy. *Journal of Cell Biology* 205, 435–445.
- Park, C., Cuervo, A.M., 2013. Selective autophagy: Talking with the UPS. *Cell Biochemistry and Biophysics* 67, 3–13.
- Parkinson-Lawrence, E.J., Shandala, T., Prodoehl, M., *et al.*, 2010. Lysosomal storage disease: Revealing lysosomal function and physiology. *Physiology (Bethesda)* 25, 102–115.
- Pompe, J., 1932. Over idiopatische hypertrophie van het hart. *Nederlands Tijdschrift voor Geneeskunde* 76, 304–312.
- Porto, A.F., 2014. Lysosomal acid lipase deficiency: Diagnosis and treatment of Wolman and cholesteryl ester storage diseases. *Pediatric Endocrinology Reviews* 12 (Suppl. 1), 125–132.
- Rajapakshe, A.R., Podyma-Inoue, K.A., Terasawa, K., *et al.*, 2014. Lysosome-associated membrane proteins (LAMPs) regulate intracellular positioning of mitochondria in MC3T3-E1 cells. *Experimental Cell Research* 331, 211–222.
- Ravikumar, B., Sarkar, S., Davies, J.E., *et al.*, 2010. Regulation of mammalian autophagy in physiology and pathophysiology. *Physiological Reviews* 90, 1383–1435.
- Reiser, J., Adair, B., Reinheckel, T., 2010. Specialized roles for cysteine cathepsins in health and disease. *Journal of Clinical Investigation* 120, 3421–3431.
- Repnik, U., Hafner Česen, M., Turk, B., 2013. The endolysosomal system in cell death and survival. *Cold Spring Harbor Perspectives in Biology* 5, 1–14.
- Repnik, U., Hafner Česen, M., Turk, B., 2014. Lysosomal membrane permeabilization in cell death: Concepts and challenges. *Mitochondrion* 19, 49–57.
- Repnik, U., Stoka, V., Turk, V., Turk, B., 2012. Lysosomes and lysosomal cathepsins in cell death. *Biochimica et Biophysica Acta* 1824, 22–33.
- Ruivo, R., Anne, C., Sagne, C., Gasnier, B., 2009. Molecular and cellular basis of lysosomal transmembrane protein dysfunction. *Biochimica et Biophysica Acta* 1793, 636–649.
- Saftig, P., Klumperman, J., 2009. Lysosome biogenesis and lysosomal membrane proteins: Trafficking meets function. *Nature Reviews Molecular Cell Biology* 10, 623–635.
- Sardiello, M., Palmieri, M., Di Ronza, A., *et al.*, 2009. A gene network regulating lysosomal biogenesis and function. *Science* 325, 473–477.
- Schroder, B.A., Wrocklage, C., Hasilik, A., Saftig, P., 2010. The proteome of lysosomes. *Proteomics* 10, 4053–4076.
- Schultz, M.L., Tecedor, L., Chang, M., Davidson, B.L., 2011. Clarifying lysosomal storage diseases. *Trends in Neuroscience* 34, 401–410.
- Settembre, C., Fraldi, A., Medina, D.L., Ballabio, A., 2013. Signals from the lysosome: A control centre for cellular clearance and energy metabolism. *Nature Reviews Molecular Cell Biology* 14, 283–296.
- Shaid, S., Brandts, C.H., Serve, H., Dikic, I., 2013. Ubiquitination and selective autophagy. *Cell Death and Differentiation* 20, 21–30.
- Shayman, J.A., Kelly, R., Kollmeyer, J., He, Y., Abe, A., 2011. Group XV phospholipase A(2), a lysosomal phospholipase A(2). *Progress in Lipid Research* 50, 1–13.
- Stern, G., 2014. Niemann-Pick's and Gaucher's diseases. *Parkinsonism & Related Disorders* 20 (Suppl. 1), S143–S146.
- Stoka, V., Turk, B., Schendel, S.L., *et al.*, 2001. Lysosomal protease pathways to apoptosis. Cleavage of bid, not pro-caspases, is the most likely route. *Journal of Biological Chemistry* 276, 3149–3157.
- Thiele, D.L., Lipsky, P.E., 1990. Mechanism of L-leucyl-L-leucine methyl ester-mediated killing of cytotoxic lymphocytes: Dependence on a lysosomal thiol protease, dipeptidyl peptidase I, that is enriched in these cells.

- Proceedings of the National Academy of Sciences of the United States of America 87, 83–87.
- Turk, B., Turk, D., Salvesen, G.S., 2002. Regulating cysteine protease activity: Essential role of protease inhibitors as guardians and regulators. *Current Pharmaceutical Design* 8, 1623–1637.
- Turk, B., Turk, D., Turk, V., 2012a. Protease signaling: The cutting edge. *Embo Journal* 31, 1630–1643.
- Turk, B., Turk, V., 2009. Lysosomes as “suicide bags” in cell death: Myth or reality? *Journal of Biological Chemistry* 284, 21783–21787.
- Turk, V., Stoka, V., Vasiljeva, O., *et al.*, 2012b. Cysteine cathepsins: From structure, function and regulation to new frontiers. *Biochimica et Biophysica Acta* 1824, 68–88.
- Turk, V., Turk, B., Turk, D., 2001. Lysosomal cysteine proteases: Facts and opportunities. *EMBO Journal* 20, 4629–4633.
- Uchimoto, T., Nohara, H., Kamehara, R., *et al.*, 1999. Mechanism of apoptosis induced by a lysosomotropic agent, L-Leucyl-L-Leucine methyl ester. *Apoptosis* 4, 357–362.
- Vasiljeva, O., Reinheckel, T., Peters, C., *et al.*, 2007. Emerging roles of cysteine cathepsins in disease and their potential as drug targets. *Current Pharmaceutical Design* 13, 387–403.
- Vernon, P.J., Tang, D., 2013. Eat-me: Autophagy, phagocytosis, and reactive oxygen species signaling. *Antioxidants & Redox Signaling* 18, 677–691.
- Wong, E., Cuervo, A.M., 2010. Integration of clearance mechanisms: The proteasome and autophagy. *Cold Spring Harbor Perspectives in Biology* 2, 1–19.
- Yang, Z., Klionsky, D.J., 2010. Eaten alive: A history of macroautophagy. *Nature Cell Biology* 12, 814–822.

PROTEIN SYNTHESIS/DEGRADATION: PROTEIN DEGRADATION – PROTEASE CLASSES

Contents

Matrix Metalloproteinases

ADAMTS Proteases: Mediators of Physiological and Pathogenic Extracellular Proteolysis

ADAMs Regulate Cell–Cell Interactions by Controlling the Function of the EGF-Receptor, TNF α and Notch

Extracellular: Plasma Membrane Proteases – Serine Proteases

Aspartic Proteases of Alzheimer's Disease: β - and γ -Secretases

The Calpain Proteolytic System

Cathepsin E: An Aspartic Protease with Diverse Functions and Biomedical Implications

Metalloproteinases Meprin α and Meprin β in Health and Disease

Kallikrein

Naturally-Occurring Polypeptide Inhibitors: Cystatins/Stefins, Inhibitors of Apoptosis (IAPs), Serpins, and Tissue Inhibitors of Metalloproteinases (TIMPs)

Matrix Metalloproteinases

G Murphy, University of Cambridge, Cambridge, UK

© 2016 Elsevier Inc. All rights reserved.

Glossary

Matrix metalloproteinases (MMPs) Endopeptidases secreted from cells or bound to the cell surface and that aid degradation of extracellular matrix molecules.

Metzincins Metalloproteinases that contain the zinc-ion-binding motif HEXXHXXGXXH and a conserved methionine that forms a unique Met-turn structure, for example, matrix metalloproteinases, astacins, repolysins, ADAMs, ADAMTSs, pappalysins, serralysins, and leishmanolysin.

Proteinase Proteolytic enzymes, also known as proteases or endopeptidases that hydrolyze internal peptide bonds of proteins and polypeptide chains (cf. exopeptidase, enzymes that hydrolyze one or a few amino acids from the N- or C-terminus of polypeptides).

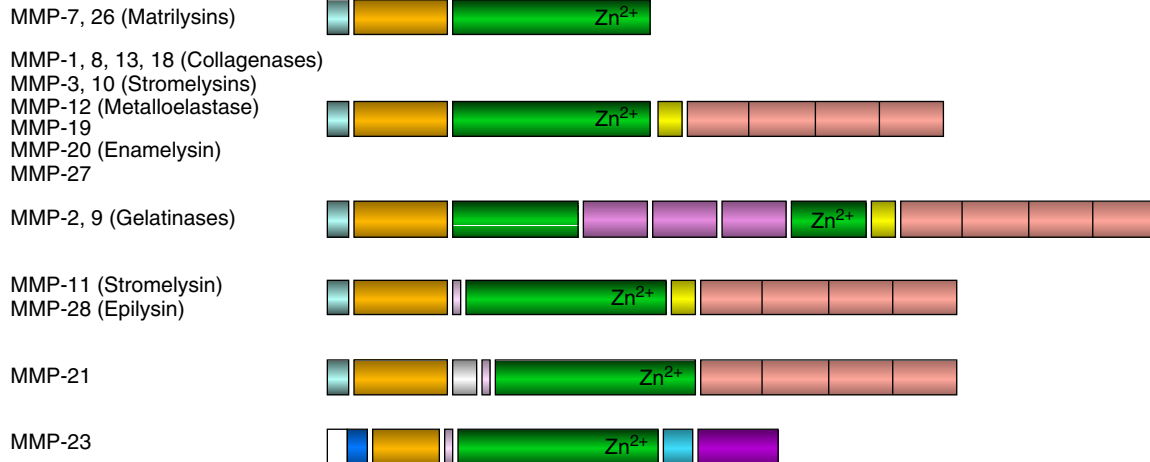
Tissue inhibitors of metalloproteinases (TIMP) Tissue inhibitors with molecular masses of 22–30 kDa that inhibit MMPs.

Structure

The matrix metalloproteinases (MMPs) are a subfamily within the M10 family of endopeptidases of the metzincin clan (M10A; Rawlings *et al.*, 2012). They are found in lower eukaryotes and in plants but diversified substantially during the evolution of the vertebrates (Fanjul-Fernandez *et al.*, 2010). In human there are 24 genes, with one gene duplication. A number of the human MMP genes (MMP-1,-3,-7,-8,-10,-12,-13, and -20) are clustered on chromosome 11 at 11q21-23. Other MMP genes are found on chromosomes 1, 8, 12, 14, 16, 20, and 22 (Ugalde *et al.*, 2010). MMPs are multidomain proteinases and their domain structure is

shown in Figure 1. All have a signal peptide, implying that they are secreted proteins. The N-terminal propeptides confer latency and need to be removed by proteolysis to generate the mature enzyme forms (see below). Most MMPs have a catalytic domain linked to a C-terminal hemopexin-like domain by a hinge region, except in the cases of MMP-7, MMP-23, and MMP-26. The 'gelatinases' MMP-2 and MMP-9 have triple fibronectin type II-like repeats inserted into the catalytic domain. The 3D structures of the catalytic domains and some ancillary domains of a number of MMPs have been resolved (Tallant *et al.*, 2010). The MMPs are largely secreted into the extracellular space, however six members are membrane anchored (membrane type, MT MMPs). Four

Soluble MMPs



Membrane-type MMPs

Transmembrane



GPI-anchored

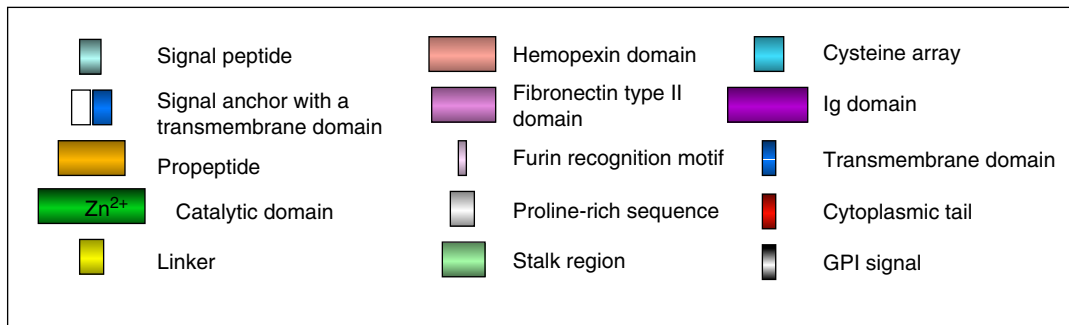


Figure 1 The domain structure of matrix metalloproteinases.

are transmembrane proteins with short cytoplasmic tails (MT1, MT2, MT3, MT5 MMPs; MMP-14,-15,-16, and -24, respectively) and two have a glycosylphosphatidylinositol (GPI) anchor (MT4, MT6 MMPs; MMP-17 and -25, respectively). MMP-23 is anomalous, containing an N-terminal type II transmembrane region, a cysteine array, and immunoglobulin domains. Although the zinc-containing catalytic domain is responsible for peptide bond hydrolysis, the hemopexin and other extra-catalytic domains, where present, exhibit a range of specific functions from interaction with substrates, TIMPs and pericellular proteins, to regulation of MMP turnover (Murphy and Nagase, 2011; Hadler-Olsen *et al.*, 2011).

Substrate Profiles: Biological and Biomedical Implications

MMPs have been implicated in many biological processes from embryonic development through wound healing, as well as in several pathologies. As secreted enzymes, functional at neutral pH, the mammalian MMPs were originally identified as significant modulators of the extracellular matrix (ECM) in tissue remodeling (Egeblad and Werb, 2002). Alongside proteolytic breakdown of ECM proteins, with a consequent loss in cell interactions, MMP activities can release growth factors bound to the ECM or generate cryptic ligands at specific extracellular protein cleavage sites, with consequent effects on cell

Table 1 Substrates of human matrix metalloproteinases

Enzymes	ECM substrates	Non-ECM substrates
<i>Secreted-type MMP</i>		
Collagenases		
Interstitial collagenase (MMP-1)	Collagens I, II, III, VII, and X, gelatins, aggrecan, link protein, entactin, tenascin, perlecan	
Neutrophil collagenase (MMP-8)	Collagens I, II, and III, gelatins, aggrecan, link protein	α 2-M, α 1-PI, α 1-antichymotrypsin, IGFBP-2, 3, 5, proIL-1 β , CTGF
Collagenase-3 (MMP-13)	Collagens I, II, III, IV, IX, X, and XIV, aggrecan, Fn, tenascin, osteonectin, Ln, perlecan	α 1-PI
Gelatinases		
Gelatinase A (MMP-2)	Gelatins, collagens IV, V, VII, X, and XI, Ln, Fn, elastin, aggrecan, link protein	CTGF, ProTGF- β , MCP-3, α 1-antichymotrypsin
Gelatinase B (MMP-9)	Gelatins, collagens III, IV, and V, aggrecan, elastin, entactin, link protein, vitronectin, N-telopeptide of collagen I	ProTGF- β , FGF receptor I, MCP-3, IGFBP-5, proIL-1 β , galectin-3, plasminogen
Stromelysins		
Stromelysin-1 (MMP-3)	Aggrecan, decorin, gelatins, Fn, Ln, collagens III, IV, IX, and X, tenascin, link protein, perlecan	ProTGF- β , IL-2 receptor α , Kit-L, IGFBP-3, proIL-1 β , ICAM-1, α 1-PI, galectin-3, plasminogen
Stromelysin-2 (MMP-10)	Aggrecan, Fn, Ln, collagens III, IV, and V, link protein	IGFBP-3, proIL-1 β , HB-EGF, CTGF, E-cadherin, α 1-antichymotrypsin, α 1-PI, α 2-M, plasminogen, uPA, proMMP-1, 7, 8, 9, 13
Matrilysins		
Matrilysin-1 (MMP-7)	Aggrecan, gelatins, Fn, Ln, elastin, entactin, collagen IV, tenascin, decorin, link protein	Pro1, 8, 10
Matrilysin-2 (MMP-26)	Gelatin, collagen IV, Fn, fibrinogen, vitronectin	Pro α -defensin, Fas-L, β 4 integrin, E-cadherin, proTNF α , CTGF, HB-EGF
Furin-activated MMP		RANKL, IGFBP-3, plasminogen
Stromelysin-3 (MMP-11)	Fn, Ln, aggrecan, gelatins	ProMMP-9, α 1-PI
Epilysin (MMP-28)	Unknown	α 1-PI, α 2-M, IGFBP-1
<i>Other secreted-type MMP</i>		
Metalloelastase (MMP-12)	Elastin, aggrecan, Fn, collagen IV, osteonectin, Ln, nidogen	Casein
RASI-1 (MMP-19)	Collagen IV, gelatin, Fn, tenascin, aggrecan, COMP, Ln, nidogen	Plasminogen, apolipoprotein(a)
Enamelysin (MMP-20)	Amelogenin, aggrecan, gelatin, COMP	IGFBP-3
MMP-21	Unknown	Unknown
MMP-27	Unknown	Unknown
<i>Membrane-anchored MMP</i>		
Type I transmembrane-type MMP		
MT1-MMP (MMP-14)	Collagens I, II, and III, gelatins, aggrecan, Fn, Ln, fibrin, Ln-5	Unknown
MT2-MMP (MMP-15)	Fn, tenascin, nidogen, aggrecan, perlecan, Ln	ProMMP-2, proMMP-13, CD44, MCP-3, tissue transglutaminase
MT3-MMP (MMP-16)	Collagen III, Fn, gelatin	ProMMP-2, tissue transglutaminase
MT5-MMP (MMP-24)	PG	ProMMP-2, tissue transglutaminase
<i>GPI-linked MMP</i>		
MT4-MMP (MMP-17)	Gelatin, fibrinogen	ProMMP-2
MT6-MMP (MMP-25)	Gelatin, collagen IV, fibrin, Fn, Ln	Unknown
<i>Type II transmembrane-type MMP</i>		
MMP-23	Gelatin	ProMMP-2
		Unknown

Major substrates of the MMPs are shown. Reproduced from Shiomi T., Lemaître V., D'Armiento J., Okada Y., 2010. Matrix metalloproteinases, a disintegrin and metalloproteinases, and a disintegrin and metalloproteinases with thrombospondin motifs in non-neoplastic diseases. *Pathology International* 60, 477–496.

Abbreviations: α 2-M, α 2-macroglobulin; α 1-PI, α 1-proteinase inhibitor; COMP, cartilage oligomeric matrix protein; CTGF, connective tissue growth factor; Fas-L, Fas ligand; FGF, fibroblast growth factor; Fn, fibronectin; HB-EGF, heparin-binding epidermal growth factor like growth factor; ICAM-1, intercellular adhesion molecule 1; IGFBP, insulin-like growth factor binding protein; Kit-L, kit ligand; Ln, laminin; MCP-3, monocyte chemoattractant protein-3; MMP, matrix metalloproteinases; MT-MMP, membrane-type MMP; PG, proteoglycan; proIL-1 β , pro interleukin-1 β ; Pro, proteinase type; ProMMP, latent MMP; proTNF- α , pro tumor necrosis factor- α ; proTGF- β , pro transforming growth factor β ; RANKL, receptor activator for nuclear factor κ B ligand; RASI-1, rheumatoid arthritis synovium inflamed-1; uPA, urokinase plasminogen activator.

interactions and signaling (Rodríguez *et al.*, 2010; Table 1). As individual MMP studies have become more focused, including proteomic studies and gene ablation models in mice, it has become clear that matrix proteins are not necessarily their only physiological or even pathological substrates. MMPs are responsible for the shedding of cell membrane proteins, including growth factors and cytokines and their receptors, cell adhesion molecules, and other proteins (Cauwe *et al.*, 2007; Lu *et al.*, 2011). MMPs also generate antimicrobial peptides, to activate or inactivate chemokines or cytokines, proteases and inhibitors, apolipoproteins, and complement components (Butler and Overall, 2009; Rodríguez *et al.*, 2010). With the discovery of the role of MMPs in the modulation of other proteases and their inhibitors in the extracellular environment, their consideration as one part of a complex interacting network of proteolytic activities, the 'protease web' has also emerged (Rodríguez *et al.*, 2010). Thus, many putative MMP substrates have been described by individual studies with different research interests. More recently rather more MMPs have been identified from the advent of the 'Degradomics' initiative. This uses proteomic techniques and the analysis of the cell and tissue proteomes in the presence versus the absence of a specific MMP. However the challenge of validation of these discoveries, in order to establish which are biologically important, still remains (Butler and Overall, 2009; Rodríguez *et al.*, 2010). Although less than ideal as a means to identify specific substrates of MMPs in humans, some interesting candidates have emerged from the analysis of transgenic mice (Gill *et al.*, 2010; Fanjul-Fernandez *et al.*, 2010). Validation, by comparing phenotypes of individual *Mmp* gene knockouts with the substrate gene ablation, has been valuable in this respect. Interpretation of the observations from *Mmp* gene ablation *in vivo* is complicated by protease web effects on other proteolytic activities and the possibility of downstream repercussions of the lack of activity in all tissues and the paucity of specific site ablations to date. It is interesting that the knockout of *Mmp* genes has few substantial developmental repercussions, with the exception of *Mmp-14* and, to a lesser extent, *Mmp-9*, *Mmp-13*, and *Mmp-20* (Page-McCaw *et al.*, 2007). This may be due the overlapping substrate profiles of many MMPs and the fact that proteases of other classes may also execute similar functions.

Currently identified MMP substrates and peptide cleavage site preferences are cataloged in the MEROPS, TOPFIND, and PMAP databases (recent aspects of this topic are reviewed in Butler and Overall, 2009; Rodríguez *et al.*, 2010; Lu *et al.*, 2011). Some potentially important substrates in both physiological and pathological processes are highlighted here. Although MMPs have been implicated in many human diseases there remain rather few definitive examples of primary causative roles.

ECM

The growth and remodeling of tissues in development or disease clearly demands the turnover of the ECM, notably the fibrillar collagens that are at the core of stromal tissues, and fibrin deposited during wound healing-related events. Although there are four MMPs defined as major fibrillar (type I,

II,III) collagen degrading activities (MMP-1, MMP-8, MMP-13, and MMP-14; note MMP-2, MMP-15, and MMP-16 also have weak activity), MMP-14 appears to be the major proteinase driving invasion through collagen when cells such as endothelial cells, fibroblasts, and macrophages invade the surrounding ECM (Willis *et al.*, 2013; Koziol *et al.*, 2012). It may also drive movement through fibrin matrices, a process key to wound healing. *In vitro* it is vital for cell growth in a 3D matrix environment (Rowe and Weiss, 2009). Of the other collagenases, there is definitive evidence that MMP-13, also known as collagenase-3, has *in vivo* collagenase activity in the mouse. *Mmp-13* null mice show an expansion of the zone of hypertrophic chondrocytes and a delay in apoptosis, indicating that MMP-13 is required for the transition from cartilage to bone at the growth plates of long bones. MMP-13 cleavage products of type II collagen and aggrecan can be recognized in a tight zone that is just distal to the zone of ossification and angiogenesis. The primary defect in *Mmp-13* deletion mutants is a failure of chondrocytes to remodel the ECM that is rich in type II collagen, as well as aggrecan. Further, *MMP-13* deletion mutants showed slower progression of an injury-induced arthritis model (Wang *et al.*, 2013). Loss of collagen II in human osteoarthritis is thought to be due to MMP-13 activity and recently developed specific inhibitors do prevent collagen loss. Evidence that the other 'collagenases,' MMP-1, MMP-8, and MMP-2; play major roles in direct collagen turnover *in vivo* remains sparse, although they have other important proteolytic roles (Page-McCaw *et al.*, 2007). Inactivating mutations in the *MMP-2* gene in humans leads to generalized osteolysis and severe osteoporosis, with arthritis in which marked progressive bone loss and joint destruction occurs. *Mmp-2*^{-/-} mice partially recapitulate the human disease due to developmentally restricted changes in osteoclasts and osteoblasts that may be due to changes in the supporting bone marrow cell system. This suggests a complex role for MMP-2 where its absence causes bone loss due to changes in the cells involved in both production and turnover. Murine MMP-1 (two forms, 'a' and 'b,' 'a' probably being the human MMP-1 homolog) is not expressed widely in the mouse and gene ablation studies suggest that protease-activated receptor1; PAR1 is a major target (Foley *et al.*, 2013). In wounded human skin, keratinocytes migrate off the basal lamina and encounter a dermal matrix rich in type I collagen; MMP-1 is thought to facilitate keratinocyte migration over the dermal matrix by lessening the affinity of collagen-integrin contacts (Pilcher *et al.*, 1997).

Many ECM proteins could be susceptible to breakdown by proteases of numerous classes and the identification of specific MMP actions in both physiological and pathological events can be gained from *in vivo* studies, for example, uncleaved galectin-3 is enriched specifically at the growth plate in *Mmp-9* null mice (Page-McCaw *et al.*, 2007). The NC1 domain of collagen XVIII is cleaved off by MMP-3,-9,-12,-13, -14 to generate the anti-angiogenic factor endostatin and plasminogen cleavage by MMP-2,-7,-9,-12 generates angiostatin that is also anti-angiogenic (Rodríguez *et al.*, 2010). MMP-2 and MMP-14 cleave the $\gamma 2$ chain of laminin-5 to release domain III, exposing a cryptic epitope that induces epithelial cell migration. Laminin-5 $\gamma 2$ fragment levels are reduced in *Mmp-14* null mice, suggesting a further mechanism by which this MMP may regulate cell movement.

Cytokines and Chemokines and Their Receptors

MMPs can proteolytically modify various cytokines and chemokines, as well as their release from ECM-binding sites, are now well established. Studies show the involvement of the MMPs in the generation of chemotactic gradients *in vivo*, for example, MMP-12 inactivates most members of the CXCL chemokine family (Rodríguez *et al.*, 2010; Hu *et al.*, 2007). MMPs are hence significant modulators of the initiation of both acute (endotoxic shock, myocardial infarction, and stroke) and chronic inflammatory conditions (autoimmune diseases and atherosclerosis), as well as ongoing disease activity and resolution. MMPs could have roles in angiogenesis by the release and activation of vascular endothelial growth factor vascular endothelial growth factor (VEGF) (MMP-1,-3,-7,-9,-14,-15,-19); MMP-9 activity has been confirmed in an *in vivo* model. bFGF is released from perlecan by MMP-1 or MMP-3 in endothelial cells and transforming growth factor β , TGF β is released from binding proteins such as the latent TGF β binding protein, LTBP complex, latency peptide, LAP, or decorin by MMP-2, -9, -13 or -14. However, MMP-released beta-glycan can bind to and act as a TGF β inhibitor (Rodríguez *et al.*, 2010; Cauwe *et al.*, 2007). Different MMP activities can promote or impair tumorigenic processes such as proliferation and apoptosis, contingent on the modulation of proteolysis of specific chemokines, growth factors, and receptors (Overall and Kleinfeld, 2006; Decock *et al.*, 2011; Hadler–Olsen *et al.*, 2013). MMP-9 expression is inversely associated with colorectal cancer metastasis which may be due to its ability to generate anti-angiogenic factors such as tumstatin and endostatin, but this is not proven. In a number of *in vivo* model studies with down-regulation or gene knockout of MMP-9, increased tumor development, progression, and metastasis was observed. On the other hand, the role of MMP-9 may be pro-tumorigenic in some stages of human cancers and the topic requires more detailed mechanistic study (Decock *et al.*, 2011; Vandooren *et al.*, 2013). Similar contradictory data exist for MMP-12 that has anti-tumorigenic activities in *in vitro* and *in vivo* models. The regulation of cytokine activities as well as protease inhibitors indicates that some MMPs have beneficial actions in immune-mediated events (Dufour and Overall, 2013). MMP-8 cleaves and activates interleukin-8 which is essential for the recruitment of neutrophils but its *in vivo* substrates are largely undefined as yet. In human cancer studies expression of MMP-8 was associated with prolonged survival in tongue squamous cell carcinoma. In breast cancer, plasma MMP-8 levels were positively associated with lymph node involvement but showed a negative correlation with the risk of distant metastasis (Decock *et al.*, 2011).

Cell–ECM and Cell–Cell Adhesion Molecules

Proteolytic modification of cell membrane proteins involved in cell adhesion could potentially be significant in the modulation of pathways associated with cell migration. These are of special relevance in tumorigenesis and angiogenesis. MMP-3 and MMP-7 disrupt E-cadherin interactions of epithelia. E-cadherin cleavage by MMPs has been linked to epithelial-to-mesenchymal transition associated with aggressive tumorigenesis. CD44 processing by MMPs facilitates cell

motility and transglutaminase that controls some ECM–integrin interactions is cleaved by MMP-2,-14,-15, and -16 (Cauwe *et al.*, 2007; Butler and Overall, 2009). MMP-14 has been shown to cleave a number of integrins or integrin precursors *in vitro* that can variously lead to their up- or down-regulation. Modulation of cell adhesion is a critical event in immune responses of leukocytes in inflammation and cancer, as well as numerous other pathologies (Cauwe *et al.*, 2007; Khokha Murthy and Weiss, 2013). Degradation of intercellular junction proteins in endothelia and in epithelia by MMPs are features of changes in permeability, for example, at the blood–brain barrier in cerebral ischemia and ischemia-related acute renal failure, respectively.

Intracellular Activities

Somewhat counter intuitive has been the recent description of MMP localization in nuclear, mitochondrial, vesicular, and cytoplasmic compartments including the cytoskeletal matrix and of ‘intracellular’ substrates for MMPs (Cauwe and Opdenakker, 2010; Butler and Overall, 2009). Proteomic approaches have demonstrated that a number of classical intracellular proteins are cleaved by MMPs, including apoptotic regulators, signal transducers, molecular chaperones, cytoskeletal proteins, systemic autoantigens, enzymes in carbohydrate metabolism and protein biosynthesis, transcriptional and translational regulators, as well as lysosomal and ubiquitination enzymes. Besides proteolysis inside cells, intracellular proteins may also be cleaved by MMPs extracellularly (Butler and Overall, 2009) since many intracellular proteins exit cells by nonclassical secretion mechanisms, or by various conditions of cell death by apoptosis, necrosis, etc. At this stage the relative importance of these activities *in vivo* is not clear.

Cellular Regulation of MMPs

Transcriptional

Pivotal modulators of cell signaling such as the MMPs necessarily have complex levels of regulation (Hadler-Olsen *et al.*, 2011). Few, or low levels, of MMPs are expressed by cells within resting tissues and they are induced in repair or remodeling processes and in disease or inflammation. At the gene level many hormones, cytokines and growth factors have been described as inducers of MMP-expression. Tissue specificity of expression of individual MMPs is mainly achieved by the combination of different transcriptional control mechanisms (Fanjul-Fernandez *et al.*, 2010). The integration of multiple signaling pathways, coupled with the cooperation between several *cis*-regulatory elements of the MMP promoters facilitates the strict spatiotemporal control of MMP- transcriptional activity. Additionally, epigenetic mechanisms, such as DNA methylation or histone acetylation, contribute to MMP regulation (Chernov and Strongin, 2012). Clinical studies have led to the identification of MMP polymorphisms and mutations causally implicated in the development of different genetic diseases (Fanjul-Fernandez *et al.*, 2010).

Posttranscriptional

This aspect of MMP biology is sparsely studied to date. mRNA stability and translational efficiency are apparent levels of regulation. Recent work has identified microRNAs that regulate MMPs, largely by translational repression or RNA target breakdown of upstream regulatory molecules (Fanjul-Fernandez *et al.*, 2010).

Posttranslational

At the posttranslational level there are also some interesting and novel regulatory mechanisms that are emerging. MMPs are synthesized in a latent 'pro' form in which a ~80 residue N-terminal propeptide harbors a cysteine residue that blocks the function of the catalytic zinc ion. Propeptide proteolysis is effected variously by autocatalytic mechanisms, probably initiated by allosteric changes in the MMP, by other MMPs, by the uPA/plasmin cascade, or by other cellular proteases. Mechanisms for the removal of the propeptide have been intensively analyzed *in vitro* (Hadler-Olsen *et al.*, 2011). Plasmin was originally identified as the major potential extracellular regulator of soluble MMPs, but it is still not clear which mechanisms are significant *in vivo*, for example, ablation of plasmin or plasminogen activators and of individual MMPs *in vivo* do not seem to prevent proMMP activation (Ra and Parks, 2007). The membrane associated MMPs and some others do appear to be activated by proprotein convertases such as furin within the secretory pathway. Potentially allosteric activation can also occur by interaction with other extracellular proteins or chemicals (Hadler-Olsen *et al.*, 2011). Such interactions may well turn out to be of some relevance *in vivo*, for example, MMP-7, MMP-2, and MMP-9 binding to proteoglycans facilitates auto-activation. MMP-9 is notable in that the extended linker sequence between the catalytic and hemopexin domains can be heterogeneously O-glycosylated. This O-glycosylated domain and its attached oligosaccharides (mucin-like) give substantial domain flexibility that modulate interactions with other proteins, for example, TIMP-1. MMP-14 also has an extensively N- and O-glycosylated linker with substantial heterogeneity that affects TIMP-2 binding and proMMP-2 activation, as well as its cellular turnover rate.

ProMMP-9 forms both homo and hetero-dimers involving the hemopexin-like domain and these can modulate its biochemical properties *in vitro* (Hadler-Olsen *et al.*, 2011; Vandooren *et al.*, 2013). In functional studies, Dufour *et al.* (2011) showed that MMP-9 multimers interact with CD44, leading to the activation of the EGF receptor and consequently the MAP kinase (extracellular signal-regulated kinases ERK1/2) pathway, thereby potentially mediating cancer cell migration. By targeting a compound to the hemopexin domain, that prevented the formation of multimers, cancer cell migration was abolished (Dufour *et al.*, 2011). Spatial regulation of MMPs may depend heavily on molecular interactions within the pericellular environment. The binding of specific secreted MMPs to cell surface receptors and other membrane proteins and ECM molecules can be invoked not only in the regulation of spatial activity, for example, CD151-MMP-7, CD44-MMP-9, and CD44-MT1-MMP; CD147 (EMMPRIN)-MMP-1, but also in regulation associated with endocytic mechanisms, for

example, Endo180-MMP-14-, low-density lipoprotein receptor-related protein-1, LRP-1-MMP-2, LRP-1-MMP-9, LRP-1-MMP-13 (Murphy and Nagase, 2011 and see below). MMPs are also down-regulated by classical endocytic mechanisms, such as clathrin coated pits or caveolae. MMP-14 may be protected from lysosomal destruction by association with tetraspanins.

Compartmentalization

The soluble MMPs are largely secreted proteins with conventional signal peptides and secretion through vesicular mechanisms and attendant regulatory processes. Polymorphonuclear leukocytes and mast cells store MMPs in vesicles and it is likely that some other cell types do similarly, although to a more limited extent. Some tumor and endothelial cells produce exosomes where vesicles persist in the extracellular space and can act as reservoirs of activity (Hadler-Olsen *et al.*, 2011). Such vesicles may be generated via diverse biological mechanisms triggered by pathways involved in oncogenic transformation, microenvironmental stimulation, cell activation, stress, or death. Exosomes are recognized as mediators of intercellular communication due to their capacity to merge with and transfer a repertoire of bioactive molecules to recipient cells. More recently MMPs have been localized in intracellular sites, for example, MMP-1,-2,-3,-9, and -13 have been found in the nucleus of different cell types but the trafficking mechanisms are not defined. Cytosolic localization has been described in various neuronal cells. Notably MMP-26 is thought to be largely located intracellularly (Hadler-Olsen *et al.*, 2011).

MMPs may be compartmentalized by association with both cell membrane and matrix proteins (Murphy and Nagase, 2011). Such interactions can orchestrate gradients or directionality of proteolytic activity and the modulation of signaling molecules. Interactions of MMPs with cell surface receptors are the basis of endocytic mechanisms to down-regulate MMP function outlined above. Although only a few MMPs have been studied, the diversity of mechanisms is enormous. These examples have led us to consider that many, if not all, of the MMPs may frequently function pericellularly, rather than at sites distal to the cells where the communication between the cells and the surrounding ECM and other bioactive molecules takes place.

Natural Inhibitors

The major natural inhibitors of the MMPs are α_2 -macroglobulin and the tissue inhibitors of metalloproteinases, TIMPs. α_2 -macroglobulin is a broad-spectrum protease inhibitor and probably largely has activities in the fluid phase of tissues. The TIMPs are the primary endogenous inhibitors; four forms (TIMPs-1-4) are found in the human and are 22–30 kDa proteins, consisting of an N-terminal metalloproteinase inhibitory domain (about 125 amino acids) and a C-terminal 'sub' domain (about 65 amino acids), each with three conserved disulfide bonds (Brew and Nagase, 2010; Murphy, 2011). The mechanism of TIMP inhibition is conserved across the MMPs, as shown by crystal structures of complexes between various

TIMPs and MMP-catalytic domains (Brew and Nagase, 2010). TIMP-3 is a key endogenous regulator of MMPs involved in ECM remodeling, inflammation, innate immunity, and cancer progression (Brew and Nagase, 2010).

Some TIMPs may interact with the hemopexin domain of specific MMPs and hence play additional biological roles, for example, the endocytic removal of MMP-2/TIMP-2 and MMP-9/TIMP-1 complexes. TIMP regulation of MMP function may itself be further regulated: receptor-mediated endocytosis occurs through low-density lipoprotein receptor-related protein 1, LRP-1 clearance of MMP-2 (complexed to thrombospondin-2 or TIMP-2) and MMP-9 (complexed to TIMP-1; Emonard *et al.*, 2005). TIMP-3 has more recently been shown to be endocytosed through LRP-1 (Scilabra *et al.*, 2013).

TIMPs regulate cell behavior independent of MMP inhibition, acting through specific cell-binding partners and signaling events. Cell signaling mechanisms mediated by TIMP-1 and TIMP-2 have been shown to promote cell proliferation, although the receptor-mediated events that are involved are not well understood. TIMP-1 has anti-apoptotic effects on some cell types, possibly acting by enhancing the activities of survival and differentiation factors in a signaling network that is mediated by the tetraspanin CD63 and by focal adhesion kinase, phosphoinositide 3-kinase and ERK. TIMP-2 binds to $\alpha 3 \beta 1$ integrin, leading to G1 phase growth arrest and enhanced *de novo* expression of the cyclin-dependent kinase inhibitor p27Kip1 (Bourboulia and Stetler-Stevenson, 2010). TIMP-3 has conflicting activities according to cell type. In some cells, it appears to promote the development of a transformed phenotype. On the other hand, TIMP-3 promotes apoptosis in several tumor cell lines and in smooth muscle cells, but this appears to involve the modulation of metalloproteinase activities. Work with *Timp-3*-null mice suggests that TIMP-3 can either promote or prevent apoptosis depending on the model system examined (Bourboulia and Stetler-Stevenson, 2010).

MMPs as Therapeutic Targets and Anti-Targets: The Use of Synthetic Inhibitors

MMPs observed functions at the ‘cutting edge’ of cell biology puts them into pivotal positions in the rapid response of cells to their environment by acting as key switches between different signaling pathways. Inevitably such enzymes should be regarded as suitable targets for therapeutic approaches to many diseases where such pathways become dysregulated. However, it is also now evident that some, as key regulators of immune responses, tissue repair, and other activities, may be classed as ‘anti-targets’ (Decock *et al.*, 2011; Dufour and Overall, 2013). A major challenge is therefore to identify which MMPs have driver roles in pathologies and could be therapeutic targets. The development of high-throughput proteomic techniques and conditional genetic modifications of mouse models is facilitating this process (Rodríguez *et al.*, 2010). A further hurdle for the development of specific inhibitors of catalysis has been the broad structural similarity of the metzincin catalytic site (Murphy, 2011; Tallant *et al.*, 2010). More detailed knowledge of active site structures has helped to some extent to resolve the problems associated with the development of more specific chemical inhibitors and selected

enzymes are now being targeted (Yiotakis and Dive, 2008). The role of exosites, both in the catalytic and hemopexin domains are being explored as determinants of substrate specificity of individual MMPs and, hence, as potential specific targeting sites for inhibitors (Hadler-Olsen *et al.*, 2011).

MMPs have well-defined roles in vascular diseases such as myocardial infarction, atherosclerosis, and aneurysms from associative studies of patient samples as well as transgenic mice and cell biological studies (Newby, 2012). Specific inhibitors of MMP-12 and MMP-13 have been developed that may be suitable for further assessment in cardiovascular diseases.

MMPs have also been implicated in blood–brain barrier dysfunction, demyelination, and the neuroinflammation and neurotoxicity underlying a range of neurological diseases, including multiple sclerosis, amyotrophic lateral sclerosis, and Alzheimer’s and Parkinson’s disease. Hence, MMP inhibition strategies have been proposed as for the treatment of multiple central nervous system (CNS) disorders. Data showing potential beneficial contributions of MMPs in key physiological and regenerative events in CNS disorders and injuries is emerging (Candelario-Jalil *et al.*, 2009; Hove *et al.*, 2012).

Imaging of MMPs

Functional analysis by genomic and proteomic techniques of MMPs *in vivo* has helped to elucidate MMP functions but studies are complicated by the many levels of posttranslational regulation and the modulation of other protease cascades. This is coupled with the challenge of designing specific active site probes, similar to the development of appropriate inhibitors. Hence *bona fide* tools for the study of this family *in vivo* have been slow to emerge. Activity-based probes have recently been designed using protein engineering to allow covalent probe–protease interactions that are useful in cell studies as proof of principle (Morell *et al.*, 2013). Fluorescence resonance energy transfer (FRET)-based probes with specific peptide sequences have been developed for MMP-12 and MMP-13 and used to image activity in an *in vivo* rodent inflammatory arthritis model (Lim *et al.*, 2013) and for MMP-13 in an osteoarthritis model (Ryu *et al.*, 2011).

MMP Activity in Drug Delivery Mechanisms

MMP activity has been used to activate cytotoxic agents *in situ*, for example, anthrax toxin in the tumor vasculature (Liu *et al.*, 2008), cytotoxic drugs coupled to nanocarriers (Samuelson *et al.*, 2013; Gao *et al.*, 2013). Vartak and Gemeinhart (2007) described the various strategies that have been adopted and have analyzed the factors that are key to successful prodrug development. As for the design of inhibitors and imaging agents, validation of the target MMP(s) and the design of cleavable peptide specificity are paramount.

Future Directions

The challenge to identify specific roles of the MMPs in development, tissue physiology, and pathology is still substantial

and no single level of investigation can suffice. The relevance of data from transgenic mice studies for the understanding of human disease remains under debate, hence efforts to study human tissue directly will be of exceptional value. At the cellular level there are many topics of outstanding interest particularly in relation to MMP trafficking and delivery at sites of activity and the relationship to cell-ECM dynamics (Lu *et al.*, 2011). The biomechanical properties of the ECM are key to cell behavior and some MMPs will have prominent roles in such processes by proteolytic modification of the tissue environment. For instance, study of ECM control of skeletal cell commitment and differentiation showed that MMP-14 in mesenchymal progenitor cells can regulate skeletal stem cell fate by changing their integrin interactions, Rho GTPase signaling and hence the nuclear transcription factors YAP and TAZ which control skeletal stem cell lineage commitment (Tang *et al.*, 2013).

The potential for MMPs as targets for therapeutic approaches is still regarded with skepticism by many and the need for further studies, particularly in human diseases has already been discussed. MMPs have not consistently been employed as biomarkers in disease as yet, although their potential as diagnostic and prognostic predictors of cancer is a burgeoning topic (Roy *et al.*, 2009). MMP-9 may be used as a marker of inflammatory pathologies and some cancers (Vandooren *et al.*, 2013), but this requires further critical analysis. Serum MMP-3 may be used to distinguish subjects with psoriatic arthritis from those with psoriasis alone and is an indicator of inflammation in rheumatoid arthritis patients. Further development of diagnostic tools capable of measuring MMP activities would advance the *in vivo* monitoring of therapeutic interventions in disease.

See also: Protein Synthesis/Degradation: Protein Degradation – Protease Classes: ADAMs Regulate Cell–Cell Interactions by Controlling the Function of the EGF-Receptor, TNF α and Notch; Extracellular: Plasma Membrane Proteases – Serine Proteases; Metalloproteases Meprin α and Meprin β in Health and Disease

References

- Bourboulia, D., Stetler-Stevenson, W.G., 2010. Matrix metalloproteinases (MMPs) and tissue inhibitors of metalloproteinases (TIMPs): Positive and negative regulators in tumor cell adhesion. *Seminars in Cancer Biology* 20, 161–168.
- Brew, K., Nagase, H., 2010. The tissue inhibitors of metalloproteinases (TIMPs): An ancient family with structural and functional diversity. *Biochimica et Biophysica Acta – Molecular Cell Research* 1803, 55–71.
- Butler, G.S., Overall, C.M., 2009. Updated biological roles for matrix metalloproteinases and new “intracellular” substrates revealed by degradomics. *Biochemistry* 48, 10830–10845.
- Candelario-Jalil, E., Yang, Y., Rosenberg, G.A., 2009. Diverse roles of matrix metalloproteinases and tissue inhibitors of metalloproteinases in neuroinflammation and cerebral ischemia. *Neuroscience* 158, 983–994.
- Cauwe, B., Van den Steen, P.E., Opdenakker, G., 2007. The biochemical, biological, and pathological kaleidoscope of cell surface substrates processed by matrix metalloproteinases. *Critical Reviews in Biochemistry and Molecular Biology* 42, 113–185.
- Cauwe, B., Opdenakker, G., 2010. Intracellular substrate cleavage: A novel dimension in the biochemistry, biology and pathology of matrix metalloproteinases. *Critical Reviews in Biochemistry and Molecular Biology* 45, 351–423.
- Chernov, A.V., Strongin, A.Y., 2012. Epigenetic regulation of matrix metalloproteinases and their collagen substrates in cancer. *Biomolecular Concepts* 2, 135–147.
- Decock, J., Thirkettle, S., Wagstaff, L., Edwards, D.R., 2011. Matrix metalloproteinases: Protective roles in cancer. *Journal of Cellular and Molecular Medicine* 15, 1254–1265.
- Dufour, A., Sampson, N.S., Li, J., *et al.*, 2011. Small-molecule anticancer compounds selectively target the hemopexin domain of matrix metalloproteinase-9. *Cancer Research* 71, 4977–4988.
- Dufour, A., Overall, C.M., 2013. Missing the target: Matrix metalloproteinase antitargets in inflammation and cancer. *Trends in Pharmaceutical Sciences* 34, 233–242.
- Egeblad, M., Werb, Z., 2002. New functions for the matrix metalloproteinases in cancer progression. *Nature Reviews Cancer* 2, 161–174.
- Emonard, H., Bellon, G., de Diesbach, P., *et al.*, 2005. Regulation of matrix metalloproteinase (MMP) activity by the low-density lipoprotein receptor-related protein (LRP). A new function for an “old friend”. *Biochimie* 87, 369–376.
- Fanjul-Fernandez, M., Folgueras, A.R., Cabrera, S., Lopez-Otin, C., 2010. Matrix metalloproteinases: Evolution, gene regulation and functional analysis in mouse models. *Biochimica et Biophysica Acta – Molecular Cell Research* 1803, 3–19.
- Foley, C.J., Fanjul-Fernandez, M., Bohm, A., *et al.*, 2013. Matrix metalloproteinase 1a deficiency suppresses tumor growth and angiogenesis. *Oncogene* 33, 2264–2272.
- Gao, W., Xiang, B., Meng, T.T., Liu, F., Qi, X.R., 2013. Chemotherapeutic drug delivery to cancer cells using a combination of folate targeting and tumor microenvironment-sensitive polypeptides. *Biomaterials* 34, 4137–4149.
- Gill, S.Y., Kassim, S.E., Birkland, T.P., Parks, W.C., 2010. Mouse models of MMP and TIMP function. *Matrix Metalloproteinase Protocols. Methods in Molecular Biology* 622, 31–52.
- Hadler-Olsen, E., Fadnes, B., Sylte, I., Uhlin-Hansen, L., Winberg, J.-O., 2011. Regulation of matrix metalloproteinase activity in health and disease. *FEBS Journal* 278, 28–45.
- Hadler-Olsen, E., Winberg, J.-O., Uhlin-Hansen, L., 2013. Matrix metalloproteinases in cancer: Their value as diagnostic and prognostic markers and therapeutic targets. *Tumor Biology* 34, 2041–2051.
- Hove, I., Lemmens, K., Van de Velde, S., Verslegers, M., Moons, L., 2012. Matrix metalloproteinase-3 in the central nervous system: A look on the bright side. *Journal of Neurochemistry* 123, 203–216.
- Hu, J., Van den Steen, P.E., Sang, Q.-X. A., Opdenakker, G., 2007. Matrix metalloproteinase inhibitors as therapy for inflammatory and vascular diseases. *Nature Reviews Drug Discovery* 6, 480–498.
- Khokha, R., Murthy, A., Weiss, A., 2013. Metalloproteinases and their natural inhibitors in inflammation and immunity. *Nature Reviews Immunology* 13, 649–665.
- Kozio, A., Martin-Alonso, M., Gonzalo, P., Arroyo, A., 2012. Site-specific cellular functions of MT1-MMP. *European Journal of Cell Biology* 91, 889–895.
- Lim, N.H., Meinjohanns, E., Bou-Gharios, G., *et al.*, 2013. *In vivo* imaging of matrix metalloproteinase 12 and matrix metalloproteinase 13 activities in the mouse model of collagen-induced arthritis. *Arthritis & Rheumatology* 66, 589–598.
- Liu, S., Wang, H., Currie, B.M., *et al.*, 2008. Matrix metalloproteinase-activated anthrax lethal toxin demonstrates high potency in targeting tumor vasculature. *Journal of Biological Chemistry* 283, 529–540.
- Lu, P., Takai, K., Weaver, V.M., Werb, Z., 2011. Extracellular matrix degradation and remodeling in development and disease. *Cold Spring Harbor Perspectives in Biology* 3 (12), a005058.
- Morell, M., Nguyen Duc, T., Willis, A.L., *et al.*, 2013. Coupling protein engineering with probe design to inhibit and image matrix metalloproteinases with controlled specificity. *Journal of the American Chemical Society* 135, 9139–9148.
- Murphy, G., 2011. Tissue inhibitors of metalloproteinases. *Genome Biology* 12, 233–240.
- Murphy, G., Nagase, H., 2011. Localizing matrix metalloproteinase activities in the pericellular environment. *FEBS Journal* 278, 2–15.
- Newby, A.C., 2012. Matrix metalloproteinase inhibition therapy for vascular diseases. *Vascular Pharmacology* 56, 232–244.
- Overall, C.M., Kleinfeld, O., 2006. Validating matrix metalloproteinases as drug targets and anti-targets for cancer therapy. *Nature Reviews Cancer* 6, 227–239.
- Page-McCaw, A., Ewald, A.J., Werb, Z., 2007. Matrix metalloproteinases and the regulation of tissue remodelling. *Nature Reviews Molecular Cell Biology* 8, 221–233.
- Pilcher, B.K., Dumin, J.A., Sudbeck, B.D., *et al.*, 1997. The activity of collagenase-1 is required for keratinocyte migration on a type I collagen matrix. *Journal of Cell Biology* 137, 1445–1457.
- Ra, H.-J., Parks, W.C., 2007. Control of matrix metalloproteinase catalytic activity. *Matrix Biology* 26, 587–596.
- Rawlings, N.D., Barrett, A., Bateman, A., 2012. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research* 40, D343–D350.

- Rodriguez, D., Morrison, C.L., Overall, C.M., 2010. Matrix metalloproteinases: What do they not do? New substrates and biological roles identified by murine models and proteomics. *Biochimica et Biophysica Acta – Molecular Cell Research* 1803, 39–54.
- Rowe, R.G., Weiss, S.J., 2009. Navigating ECM barriers at the invasive front: The cancer cell–stroma interface. *Annual Review of Cell and Developmental Biology* 25, 567–595.
- Roy, R., Yang, J., Moses, M.A., 2009. Matrix metalloproteinases as novel biomarkers and potential therapeutic targets in human cancer. *Journal of Clinical Oncology* 27, 5287–5297.
- Ryu, J.H., Lee, A., Na, J.H., *et al.*, 2011. Optimization of matrix metalloproteinase fluorogenic probes for osteoarthritis imaging. *Amino Acids* 41, 1113–1122.
- Samuelson, L.E., Scherer, R.L., Matrisian, L.M., McIntyre, J.O., Bornhop, D.J., 2013. Synthesis and *in vitro* efficacy of MMP9-activated nanodendrons. *Molecular Pharmaceutics* 10, 3164–3174.
- Scilabra, S.D., Troeberg, L., Yamamoto, K., *et al.*, 2013. Differential regulation of extracellular tissue inhibitor of metalloproteinases-3 levels by cell membrane-bound and shed low density lipoprotein receptor-related protein 1. *Journal of Biological Chemistry* 288, 332–342.
- Tallant, C., Marrero, A., Gomis-Rüth, F.X., 2010. Matrix metalloproteinases: Fold and function of their catalytic domains. *Biochimica et Biophysica Acta – Molecular Cell Research* 1803, 20–28.
- Tang, Y., Rowe, R.G., Botvinick, E.L., *et al.*, 2013. MT1-MMP-dependent control of skeletal stem cell commitment via a β 1-integrin/YAP/TAZ signaling axis. *Developmental Cell* 25, 402–416.
- Ugalde, A.P., Ordóñez, G.R., Quirós, P.M., Puente, X.S., López-Otín, C., 2010. Metalloproteinases and the degradome. *Methods in Molecular Biology* 622, 3–29.
- Vandooren, J., Van den Steen, P.E., Opdenakker, G., 2013. Biochemistry and molecular biology of gelatinase B or matrix metalloproteinase-9 (MMP-9): The next decade. *Critical Reviews in Biochemistry and Molecular Biology* 48, 222–272.
- Vartak, D.G., Gemeinhart, R.A., 2007. Matrix metalloproteinases: Underutilized targets for drug delivery. *Journal of Drug Targeting* 15, 1–20.
- Wang, M., Sampson, E.R., Jin, H., *et al.*, 2013. MMP13 is a critical target gene during the progression of osteoarthritis. *Arthritis Research & Therapy* 15, R5.
- Willis, A.L., Sabeh, F., Li, X.-Y., Weiss, S.J., 2013. Extracellular matrix determinants and the regulation of cancer cell invasion stratagems. *Journal of Microscopy* 251, 250–260.
- Yiotakis, A., Dive, V., 2008. Synthetic active site-directed inhibitors of metzincins: Achievement and perspectives. *Molecular Aspects of Medicine* 29, 329–338.

Relevant Websites

<http://merops.sanger.ac.uk/>
MEROPS.

<http://pmap.burnham.org/proteases>
PMAP.

<http://degradome.uniovi.es/met.html>
Universidad de Oviedo.

ADAMTS Proteases: Mediators of Physiological and Pathogenic Extracellular Proteolysis

SS Apte, Cleveland Clinic, Cleveland, OH, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Extracellular matrix (ECM) The highly organized and interconnected networks of proteinaceous and nonproteinaceous molecules (such as glycosaminoglycans and proteoglycans) that occupy the space around cells. ECM constitutes the scaffold that organizes each organ's hierarchical structure. It provides cells with an adhesive substrate that dictates cell polarity and regulates cell behavior. ECM endows organs with specific mechanical properties, best evident in connective tissues such as bone, tendon, and cartilage, and it has a crucial role in epithelia, where it comprises the basement membrane. The regulatory function of ECM is via direct cell-matrix interactions mediated by integrins, syndecans, and other receptors, as well as by sequestration of growth factors and cytokines, and release of bioactive ECM fragments termed matrikines.

Inherited connective tissue disorders A major category of human Mendelian disorders caused by gene mutations that affect the structure and function of connective tissue cells and their ECM. Examples include osteogenesis imperfecta (brittle bone disease), Ehlers–Danlos syndrome (causes weakness of blood vessels, skin, and internal organs), and Marfan syndrome, which affects the aorta, eyes, and skeleton.

Metzincins A class of secreted and cell-surface proteolytic enzymes characterized by a zinc-binding catalytic domain with a conserved three-dimensional structure. Protease families within the class primarily differ in regard to

additional (ancillary) domains located on the C-terminus of the catalytic domain. They constitute a major category of proteases that target ECM and cell-surface molecules, and are involved in cell migration, tissue destruction in disease, and cellular regulation by ectodomain shedding. Some metzincins (ADAM proteases, membrane-type matrix metalloproteinases (MT-MMPs), and meprins) are single-pass transmembrane molecules but the majority are secreted from cells (Astacin/tolloid/BMP-like proteases, most MMPs and all ADAMTS proteases).

Osteoarthritis (OA) A disorder of synovial joints characterized by progressive loss of articular cartilage leading to pain, swelling, and loss of joint function. In contrast to inflammatory arthritis, such as rheumatoid arthritis, which is primarily an autoimmune condition, OA is a degenerative condition resulting from altered joint mechanics, injury, or mild cartilage dysplasia, which leads to a cycle of progressive cartilage destruction and secondary inflammation.

Thrombotic thrombocytopenic purpura A hemostatic defect caused by unbridled platelet aggregation. Platelet aggregates occlude microvessels, and damage vital organs, with potentially fatal consequences. It is diagnosed by the presence of thrombocytopenia (reduced platelet count), anemia (reduced red blood cell count), and fragmented red blood cells damaged by their passage through the blocked microvessels.

Introduction, Defining Characteristics and Phylogeny of A Disintegrin-Like and Metalloprotease Domain with Thrombospondin Type 1 Motif Proteinases

At the time of its discovery, a disintegrin-like and metalloprotease domain with thrombospondin type 1 motif (ADAMTS1) was thought to be a variant disintegrin and metalloprotease domain (ADAM) because it has a similar catalytic domain active-site sequence (Kuno *et al.*, 1997). However, unlike ADAMs, ADAMTS proteases are not membrane anchored, and were found to have several distinguishing features (Hurskainen *et al.*, 1999). In particular, the ADAMTS ancillary (nonproteolytic) domains contain a canonical modular structure that always includes one or more thrombospondin type 1 repeats (TSRs, Figure 1). This characteristic domain organization, together with additional sequence and structural hallmarks (see below) clearly distinguishes the ADAMTS family from other zinc metalloproteases or metzincins, namely the matrix metalloproteinases (MMPs), astacin/BMP/tolloid/meprin proteases (astacins), and membrane-anchored ADAMs. All metzincins have structurally similar

catalytic domain folds and an active-site configuration with similar zinc-binding sequences and catalytic mechanisms. They are synthesized as proenzymes that subsequently undergo excision of the propeptide.

The zinc-binding sequence signature of the ADAMTS protease active site is of the repolysin or snake venom type (HEXXH + HD), indicating a closer evolutionary relationship to the membrane-anchored ADAMs rather than MMPs and astacins (Hurskainen *et al.*, 1999). Also, unlike many MMPs, but like all ADAMs, ADAMTS propeptide excision occurs through the proteolytic activity of proprotein convertases such as furin. However, the cysteine signature of the ADAMTS propeptide and catalytic domains is distinct from both MMPs and ADAMs; ADAMTS catalytic domain structures identified 4 disulfide bonds in ADAMTS1, 4 and 5, compared to 3 in ADAM17 and ADAM33, and none in the MMPs (Hurskainen *et al.*, 1999). Pioneering work of Kuno *et al.* on the prototypic ADAMTS protease, ADAMTS1, established its binding to the extracellular matrix (ECM, see glossary above), and the importance of the ancillary domain in this regard (Kuno and Matsushima, 1998; Kuno *et al.*, 1999). This point has assumed

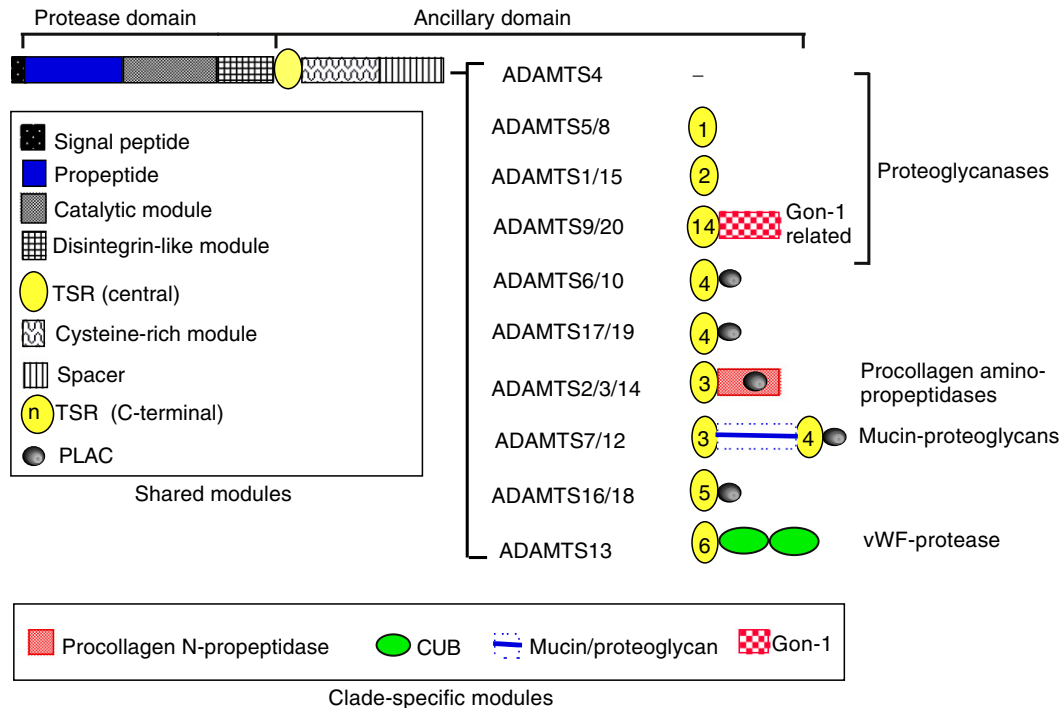


Figure 1 Overview of mammalian ADAMTS proteases. The domain backbone shared by each ADAMTS protease is shown at the top, and the shared component modules are indicated in the box at left. The modular organization of the specialized evolutionarily related subclasses of ADAMTS proteases is indicated on the right. The additional C-terminal modules shown for each subgroup are added to the backbone at its C-terminus. The number of TSRs is indicated for each variant. Modules that are unique to each subgroup are explained at the bottom of the figure. Structural or functional characteristics that define each subgroup are shown at far right. Note that the Gon-1 related proteases are also proteoglycanases but are shown as a separate subclass. ADAMTS6 and 10 constitute a separate evolutionary subclass from ADAMTS17 and 19 but are shown together as they have the same modular organization. Reproduced from Apte, S.S., 2009. A disintegrin-like and metalloprotease (reprolysin-type) with thrombospondin type 1 motif (ADAMTS) superfamily: Functions and mechanisms. *Journal of Biological Chemistry* 284, 31493–31497. CUB, complement C1r/C1s, Uegf, Bmp1 domain; PLAC, protease and lacunin domain.

importance as a general ADAMTS characteristic, since ADAMTS catalytic domains alone generally lack proteolytic activity toward substrates.

Subsequent to discovery of ADAMTS1, the explosion of genome and transcript sequencing driven by the human genome project led to identification and molecular cloning of all family members. ADAMTS11 was misnumbered (it is identical to ADAMTS5); hence although there are 20 assigned ADAMTS numbers, there are only 19 ADAMTS proteases. Together with 7 ADAMTS-like proteins, they constitute the ADAMTS superfamily (Apte, 2009). ADAMTS-like proteins resemble the ancillary domains of ADAMTS proteases and some appear to be involved in similar processes and genetic disorders as the ADAMTS proteases. They are distinct gene products and not the result of alternative splicing of ADAMTS genes (Apte, 2009).

Phylogeny of the human ADAMTS family demonstrated the existence of numerous paralogs, subdividing the family into clades of paired or triplet paralogs, i.e., proteases with highly similar primary structure and domain organization (Figure 1; Apte, 2004; Huxley-Jones et al., 2005). ADAMTS proteases appear to be a metazoan innovation, since they are absent in bacteria, algae, and fungi. Completion of the genome of metazoans such as *Caenorhabditis elegans*, *Drosophila*

melanogaster, and *Ciona intestinalis* has facilitated understanding of the evolution of human ADAMTS genes. Humans and other vertebrates have significantly more ADAMTS genes than these organisms. Each of the human clades is rooted in one of 6 predicted *Ciona* ADAMTS proteases, indicating that the ancestral gene of each clade was present in an early chordate ancestor and subsequently underwent duplication (Huxley-Jones et al., 2005). ADAMTS13, the von Willebrand factor (vWF) protease, appears to be a vertebrate innovation, possibly related to development of a closed circulation and concomitant requirement for hemostasis, since an ortholog was not found in *Ciona* or protostomes. The so-called 'proteoglycanase' clade, highlighting a shared ability to cleave proteoglycans, is the largest and appears to be the most vertebrate-specific in regard to its amplification. Overall, phylogenetic analysis suggests that ADAMTS vertebrate gene expansion reflects sub-functionalization, i.e., expansion of structurally similar and functionally overlapping genes with emergence of some distinct characteristics, such as individual spatial-temporal expression patterns, rather than neo-functionalization (i.e., a stand-apart innovation such as ADAMTS13) (Huxley-Jones et al., 2005). Sub-functionalization is supported by work using gene knockout mice that demonstrated cooperativity of ADAMTS paralogs during morphogenesis.

Domain Organization and 3-Dimensional Structure

With the exception of ADAMTS13, whose propeptide is short (Majerus *et al.*, 2003), ADAMTS proteases have a propeptide of > 200 residues that typically contains 3 Cys residues. A consensus proprotein convertase (e.g., furin) processing site (RXK/RR↓) is present at the junction of the propeptide and catalytic domain. The exception is ADAMTS10, which has a suboptimal furin processing site (GLKR) at this junction (Kutz *et al.*, 2011). However, like most other ADAMTS proteases, it has 1–2 additional furin cleavage sites within the propeptide. The ADAMTS catalytic domain and disintegrin-like domain are believed to constitute a single functional unit, and all downstream modules are regarded as constituting the ADAMTS ancillary domain. The core ancillary domain comprises TSR1, the cysteine-rich domain (CRD) and the spacer, which lacks Cys residues. TSR1 is highly conserved among ADAMTS family members. The CRD and spacer are followed in every ADAMTS except ADAMTS4 by one or more additional TSRs. Many ADAMTS proteases have a small, cysteine-rich C-terminal PLAC (protease and lacunin) module or other characteristic C-terminal domains (Figure 1). In ADAMTS7 and ADAMTS12 the TSR array is interrupted by a mucin domain, which contains glycosaminoglycan (GAG) attachment sites (Somerville *et al.*, 2004). Few ADAMTS splice variants have been reported. One, in ADAMTS4, alters the sequence of its spacer module, with no apparent significant impact on proteolytic specificity (Wainwright *et al.*, 2013). ADAMTS20 has a short splice variant that excludes 4 C-terminal TSRs and the Gon-1 domain, and appears to be the dominant ADAMTS20 transcript, whereas splicing in the ADAMTS9 propeptide and TSR array are known (Dubail *et al.*, 2014; Rao *et al.*, 2003), but have not been functionally investigated.

Because of inherent challenges in expressing and purifying these large, complex proteases, a high-resolution three-dimensional structure is presently unavailable for a full-length ADAMTS protease. A low resolution structure of ADAMTSL1/punctin-1, comprising the core ancillary domain and 3 TSRs was obtained by rotary shadowing electron microscopy and showed a hockey-stick like structure with a globular structure with an attached extended segment (Hirohata *et al.*, 2002). 3-dimensional structures obtained by X-ray crystallography are currently known for the catalytic + disintegrin-like domain of ADAMTS1, ADAMTS4, and ADAMTS5, both in apo-form and inhibitor-bound (Gerhardt *et al.*, 2007; Mosyak *et al.*, 2008; Shieh *et al.*, 2008). These studies used different preparation methods – ADAMTS1 catalytic domain + disintegrin-like domain was expressed in insect cells, ADAMTS5 catalytic domain was refolded from *Escherichia coli* inclusion bodies, and the ADAMTS4 and ADAMTS5 catalytic domain + disintegrin-like domains were purified from transfected CHO cells. Nevertheless, all structures closely resembled each other and identified several unifying features of ADAMTS proteases. The overall catalytic domain fold was similar to previously determined structures of the MMP and ADAM family members in regard to α/β structure, but differed in regard to specific loops joining β -strands and α -helices. The ADAMTS catalytic domains differed from MMPs and ADAMs in having 4 disulfide bonds and two calcium-binding sites, as well as a deeper active-site pocket containing the catalytic zinc atom. ADAMTS13

reportedly has three calcium-binding sites. An extended 9-residue strand connected the catalytic domain to the smaller disintegrin-like domain, bringing it in close proximity to the active site, suggesting it could contribute an auxiliary substrate (or inhibitor)-binding surface of the protease (Gerhardt *et al.*, 2007; Mosyak *et al.*, 2008). That the catalytic and disintegrin-like domains act as a functional unit is also supported by the domain structure of ADAMTS-like proteins, which do not have either of these domains (Hirohata *et al.*, 2002). Comparison of the ADAMTS4 and ADAMTS5 catalytic domains with or without bound inhibitor, i.e., the Apo forms, identified a dynamic active-site configuration, i.e., an autoinhibited ‘non-binding’ closed form and an open ‘binding’ form (Mosyak *et al.*, 2008). In its open state, the inhibitor-bound active sites were similar to those of MMPs and ADAMs, albeit with differences in local pockets, and showed a Zn atom coordinated by three conserved histidines. The S2’ pocket was formed by a short disulfide-containing loop with the motif CGxxxxCDTL, which was proposed as a unique feature of ADAMTS proteases. In the unliganded state, this loop moves from an open position toward the catalytic zinc by ~ 8 Å, resulting in removal of bound calcium and repositioning of the disulfide bridge. It was postulated that the open and closed forms coexisted in equilibrium, with the mobility of the active-site responsive to the dynamic nature of their proteoglycan substrates (Mosyak *et al.*, 2008).

The fold of the disintegrin-like domains resembles the CRD of the ADAM proteases, and despite sequence identity with the disintegrins (hence the name disintegrin-like), has only partial structural homology to snake venom disintegrins (Gerhardt *et al.*, 2007; Mosyak *et al.*, 2008). Determination of the crystal structure of the non-catalytic domains of ADAMTS13 (disintegrin-like, TSR1, CRD, and spacer) was a landmark in visualizing ADAMTS proteases (Akiyama *et al.*, 2009). It showed that the N-terminal portion of the ADAMTS CRD (designated C_A) resembled the disintegrin-like domain and the ADAM CRDs (although it lacked sequence homology to them), whereas the C-terminal part of this domain, designated C_B, formed a short rod, bridging C_A to the spacer. The C_A and spacer domains made direct contact. The spacer domain, which has no sequence homologies outside the ADAMTS family, and lacks Cys residues, folded into a single globular domain with 10 β -strands in a jellyroll topology, forming two obliquely facing antiparallel β -sheets. Together with structure-based mutagenesis of ADAMTS13, Akiyama *et al.* (2009) identified three VWF-binding exosites on the linearly aligned discontinuous surfaces of the disintegrin-like, cysteine-rich, and spacer domains (Wu *et al.*, 2006).

The structure of TSR1 of ADAMTS13 was analogous to the previously determined structure of TSRs 2 and 3 of the ECM protein thrombospondin-1 (TSP1) (Tan *et al.*, 2002), and comprised an antiparallel 3-stranded fold with rigidity provided by stacked Arg and Trp residues (Akiyama *et al.*, 2009). The single TSR of ADAMTS4 was modeled based on the structure of the TSP1 TSRs, and positive charges were found to cover nearly an entire face, potentially explaining the high affinity of this TSR for sulfated GAG chains of aggrecan (Tan *et al.*, 2002). In addition to GAG-binding, which could mediate cell-surface and substrate binding, a postulated function of TSRs could be to mediate proper spacing of exosites. TSR2

and TSR3 of TSP1 were arranged as an extended rigid structure with tilt and twist angles of 30° and 180° between the TSRs (Tan *et al.*, 2002). However, the precise relationship between tandem TSRs of ADAMTS proteases will depend on the sequences of the linkers between them, and is yet to be determined.

Because of strict conservation of all cysteine residues in the ancillary domain, as well as of numerous hydrophobic residues in all domains that provide additional stabilization, it is very likely that all ADAMTS proteases will have a similar structural backbone. Differences in surface residues and of peripheral loop length and sequence likely mediate different substrate binding by the various ADAMTS proteases. Three-dimensional structures for an ADAMTS propeptide and of C-terminal domains such as the Gon-1 domain and PLAC domain are presently unavailable.

Biosynthesis and Regulation

ADAMTS proteases are directed to the secretory pathway by a signal peptide, which is removed co-translationally. ADAMTS proteases (except ADAMTS4) have N-linked oligosaccharide attachment sites that are also modified co-translationally. The propeptide as well as its N-glycans were found to be crucial for secretion of the ADAMTS9 catalytic domain (Koo *et al.*, 2007), but appear not to be essential for secretion of ADAMTS13 (Majerus *et al.*, 2003). Similarly, propeptide N-glycosylation of a worm ADAMTS named MIG-17 was crucial for its localization during gonad morphogenesis (Nishiwaki *et al.*, 2004). The ADAMTS13 propeptide is unusually short (41 residues) and dispensable for maintenance of quiescence or folding of downstream domains (Majerus *et al.*, 2003). Many ADAMTS TSRs undergo two unusual and specific posttranslational modifications, named C-mannosylation and O-fucosylation. C-mannosylation occurs on W^o residues in W^oXXW⁺ motifs and has been documented in an ADAMTS-like protein, ADAMTSL1 (punctin-1), but its functional significance is presently unknown (Wang *et al.*, 2009). Because the modification occurs on unfolded peptides, it is thought to occur co-translationally. The C-mannosylation consensus sequence is most frequently present in TSR1, consistent with its high degree of conservation (Wang *et al.*, 2009).

O-fucosylation is the addition of fucose or glucose β 1-3fucose disaccharide to Ser/Thr residues within the consensus sequence CXX(S/T)CG by the tandem activity of protein O-fucosyltransferase 2 (POFUT2) and β 1,3-glucosyltransferase (B3GLCT, previously known as B3GALTL) (Luo *et al.*, 2006). This modification was shown to be a crucial requirement for efficient secretion of ADAMTS13 and ADAMTSL1 (Ricketts *et al.*, 2007; Wang *et al.*, 2007), and could constitute a form of protein quality control within the ER, since it does not occur on unfolded peptides. A human genetic disorder resulting from B3GLCT mutations, named Peters Plus syndrome (Lesnik Oberstein *et al.*, 2006), affecting the eyes and other organs, may result at least in part from reduced secretion of several family members, since half of its predicted substrates are ADAMTS proteins. ADAMTS7 and ADAMTS12 contain predicted sites for attachment of GAG chains within their mucin domains. ADAMTS7 was shown to be modified by attachment

of chondroitin sulfate, thus rendering it a proteoglycan, an unusual modification that was previously undocumented in a protease (Somerville *et al.*, 2004).

Once secreted, many ADAMTS proteases are susceptible to *trans*- or autocatalytic proteolysis within the ancillary domain. This is a poorly studied phenomenon that potentially can be activating, generate new or promiscuous activities, or inactivate the protease through loss of substrate-binding exosites. Cell-surface heparan sulfate proteoglycans such as syndecans have been proposed as an activation mechanism for aggrecanases that leads to MMP cleavage of the ancillary domain (Gao *et al.*, 2003; Echtermeyer *et al.*, 2009).

Protease inhibitors active against ADAMTS proteases include TIMP-3, one of a family of four tissue inhibitors of metalloproteases, which are well-known MMP inhibitors, and α 2-macroglobulin (Gendron *et al.*, 2003; Hashimoto *et al.*, 2001; Tortorella *et al.*, 2004). Inhibition of ADAMTS proteases by TIMP-3 reflects the close relationship of the ADAMTS catalytic domain fold to that of MMPs and ADAMs. A chondroprotective role for TIMP-3 is suggested by increased cartilage breakdown in *Timp3*^{-/-} mice (Mahmoodi *et al.*, 2005). Polyanions such as calcium pentosan polyphosphate are thought to reduce cartilage breakdown in part by enhancing TIMP-3 affinity for aggrecanases (Troeborg *et al.*, 2008). ADAMTS4 and ADAMTS5 are endocytosed by the scavenger receptor low-density lipoprotein receptor-related protein 1 (LRP1) (Yamamoto *et al.*, 2013). LRP1 binds them with different affinities raising the possibility that ADAMTS5, which binds more strongly, may be a competitor for LRP1 endocytosis of ADAMTS4. ADAMTS interacts with LRP1 via its cysteine-rich and spacer domains, whereas ADAMTS4 binds via TSR1 and the spacer (Yamamoto *et al.*, 2013).

Mechanistic Basis of Genetic and Acquired Disorders Involving ADAMTS Proteases

An important viewpoint for understanding the outcomes of ADAMTS mutations, indeed applicable to any protease, is that the biology of a protease is really the biology of its substrates. Mendelian disorders resulting from ADAMTS mutations (Table 1), albeit rare, provide insights on ADAMTS function in relation to the biology of specific substrates and the processes and pathways they participate in. Although numerous disease associations of ADAMTS proteases have been reported from transcriptome analysis and genome-wide association analysis, especially in cancer, the author have focused here on disease and functional roles that are well substantiated by altered cleavage of a specific substrate. An extensive literature has accumulated on three such ADAMTS-substrate relationships that illuminate disease mechanisms:

ADAMTS13, vWF and Thrombotic Thrombocytopenic Purpura

ADAMTS13, also known as the vWF-protease, appears to have no substrates other than vWF. vWF is synthesized by endothelial cells and megakaryocytes, and is released as ultra-large string-like structures which provide a markedly adhesive substrate for platelets (Zheng *et al.*, 2002). vWF binds glycoprotein

Table 1 Mendelian disorders resulting from ADAMTS mutations

Mendelian disorder	MIM number	Gene/chromosomal locus	Inheritance
Ehlers–Danlos syndrome (EDS), dermatosparaxis type or (VIIC)	225410	ADAMTS2, 5q35.3 (Colige <i>et al.</i> , 1999)	Autosomal recessive
Weill–Marchesani syndrome 1/Weill–Marchesani syndrome, autosomal recessive/mesodermal dysmorphodystrophy, congenital	277600	ADAMTS10/19p13.2 (Dagoneau <i>et al.</i> , 2004)	Autosomal recessive
Thrombotic thrombocytopenic purpura, congenital/Upshaw–Schulman syndrome	274150	ADAMTS13, 9q34.2 (Levy <i>et al.</i> , 2001)	Autosomal recessive
Weill–Marchesani-like syndrome	613195	ADAMTS17, 15q26.3 (Morales <i>et al.</i> , 2009)	Autosomal recessive
Microcornea, myopic chorioretinal atrophy, and telecanthus (MMCAT)	615458	ADAMTS18, 16q23.1 (Aldahmesh <i>et al.</i> , 2013)	Autosomal recessive

1b on the platelet surface and to collagen in sub-endothelium of damaged vessels. ADAMTS13, which is primarily secreted by stellate cells in the liver, and to a lesser extent by vascular endothelial cells, megakaryocytes, and platelets, cleaves the vWF strings at the Tyr¹⁶⁰⁵-Met¹⁶⁰⁶ peptide bond in the A2 domain to reduce their prothrombogenic propensity (Zheng *et al.*, 2002). ADAMTS13 is particularly efficient when vWF strings are stretched by circulatory shear force. Shear force extends the A2 domain, exposing the scissile bond, which can also be induced by denaturants such as urea, which is used in biochemical assays for plasma ADAMTS13 activity. Loss of ADAMTS13 activity to less than 5% of normal, resulting most commonly from autoantibodies (acquired thrombotic thrombocytopenic purpura (TTP)) or rarely, from ADAMTS13 mutations (congenital TTP), leads to unbridled platelet thrombogenesis (Zheng, 2013). Most autoantibodies formed in TTP react with the CRD and spacer module, most commonly the latter, and correspond to ADAMTS13 surface regions comprising exosites for vWF processing (Zheng, 2013). Additionally, the ADAMTS13 TSRs contain free thiols that may prevent higher order complex formation of vWF and reduce vWF mediated platelet aggregation. Coagulation factor VIII binds vWF with high affinity and can facilitate vWF cleavage under shear stress conditions, possibly by increasing unfolding of the vWF-A2 domain. Platelet glycoprotein 1b α , which binds vWF, also enhances cleavage of multimeric vWF by ADAMTS13. *Adamt13* deficient mice develop TTP, but only when challenged with shigatoxin or recombinant vWF (Motto *et al.*, 2005). Shigatoxin binds to A1-A2 domains of vWF and interferes with processing by ADAMTS13 (Nolasco *et al.*, 2005). Thus, the highly specific ADAMTS13-vWF interaction is clearly influenced by environmental and genetic factors. Reduced plasma ADAMTS13 activity is a risk factor for myocardial infarction, cerebral stroke, preeclampsia, and cerebral malaria (Zheng, 2013).

ADAMTS2, Procollagen Processing and Dermatosparaxis

Fibrillar collagens (e.g., collagen types I, II, III, V, XI) are synthesized as procollagens with bulky N- and C-terminal propeptides that preclude assembly into tightly packed fibrils. Hence they are excised prior to assembly, with ADAMTS proteases thought to remove the N-propeptides of procollagens I,

II, and III. ADAMTS2 mutations cause Ehlers–Danlos syndrome (dermatosparactic type) (Colige *et al.*, 1999), a connective tissue disorder first identified in cattle and named dermatosparaxis (Lapiere and Nusgens, 1993) and subsequently in humans (Colige *et al.*, 1999). In this disorder, there is failure to excise the amino-propeptide of procollagen I in the dermis of skin. The retained propeptide hinders collagen fibril assembly, and the resulting thin, ribbon-like, and branched fibrils are mechanically weak, resulting in severe skin fragility (Lapiere and Nusgens, 1993). However, arteries, bone, and tendon, which also contain collagen I as a major constituent, are not as fragile, very likely because of compensating procollagen processing activity possibly provided by ADAMTS3 (Fernandes *et al.*, 2001) and/or ADAMTS14 (Colige *et al.*, 2002). Recently, ADAMTS3, was shown to process the pro-lymphangiogenic factor VEGF-C, converting the 29/31-kDa pro-form of VEGF-C to the mature 21/23-kDa form, which is active in signaling (Jeltsch *et al.*, 2014).

ADAMTS10, Tissue Microfibrils and Weill–Marchesani Syndrome

Tissue microfibrils are macromolecular complexes in ECM comprising fibrillins and other proteins. During the embryonic period, fibrillin-2 predominates, and most adult microfibrils comprise fibrillin-1. In particular, the ocular zonule, which suspends the ocular lens in the optic path, is a cell-free structure almost entirely composed of fibrillin-1 microfibrils. In addition to providing structural support, microfibrils regulate storage and activation of transforming growth factor- β (TGF β) and related growth factors. Recessive ADAMTS10 mutations cause Weill–Marchesani syndrome (WMS1), in which the lens is dislocated (ectopia lentis), leading to severe eye problems; affected individuals are short, with short hands and feet and have thick skin, stiff joints, and heart defects (Dagoneau *et al.*, 2004). ADAMTS17 mutations cause a similar syndrome (Table 1), except that brachydactyly, joint stiffness, and heart defects are absent (Morales *et al.*, 2009). Intriguingly, fibrillin-1 mutations lead to ectopia lentis as an isolated disorder, as well as to dominantly inherited WMS, which was suggestive of a functional link between ADAMTS10 and fibrillin-1 (Faivre *et al.*, 2003). Indeed, ADAMTS10 binds to fibrillin-1 and is found in microfibrils in skin and the zonule (Kutz *et al.*, 2011).

ADAMTS10 cleaves fibrillin-1, although it does so inefficiently, and was found to enhance microfibril assembly in cultured fibroblast. The relationship of ADAM17 to fibrillins is presently unknown. Two other superfamily members, ADAMTSL2 and ADAMTSL4 also have a strong genetic relationship to fibrillin-1, which was reviewed recently (Hubmacher and Apte, 2011).

ADAMTS Proteases as Aggrecanases in Osteoarthritis

The hallmark of articular cartilage ECM is an abundance of the highly sulfated proteoglycan aggrecan, which forms large complexes with hyaluronan (HA) that are constrained by a collagenous network, primarily comprised of collagen II (Heinegard and Saxne, 2011). This combination and arrangement of ECM components is responsible for the weight-bearing properties of cartilage. Proteolysis of aggrecan and its subsequent release from cartilage was identified as a major pathogenic mechanism in osteoarthritis (OA) (Lark *et al.*, 1997). ADAMTS4 and subsequently, ADAMTS5, were isolated as the activities responsible for aggrecan loss (Abbaszade *et al.*, 1999; Tortorella *et al.*, 1999). Both proteases cleave aggrecan at multiple sites, but cleavage at the Glu³⁷³-Ala³⁷⁴ site is the most deleterious, since it releases the bulk of aggrecan from its anchorage to HA. Analysis of mice with mutant ADAMTS4 and ADAMTS5 determined that ADAMTS5 was the major aggrecanase, because *Adamts5* mutant mice were resistant to both surgically induced and inflammatory arthritis (Glasson *et al.*, 2005; Stanton *et al.*, 2005). Furthermore, mutagenesis of the Glu³⁷³-Ala³⁷⁴ scissile bond in mice prevented cartilage destruction (Little *et al.*, 2007). With this stringent demonstration of causality, ADAMTS4 and ADAMTS5 are considered as major drug targets in arthritis. Although ADAMTS1, ADAMTS9, and other ADAMTS proteases cleave aggrecan, they do so inefficiently and are not as rigorously associated with cartilage destruction in OA as ADAMTS4 and ADAMTS5 (Fosang and Little, 2008; Little *et al.*, 2005).

Functions in Mammalian Development Gleaned from Mouse Mutants

Several phenotypes are now known from mouse mutants, which are summarized in Table 2 and whose underlying mechanisms are understood to varying degrees. *Adamts1* deficient mice have reduced survival immediately after birth, for reasons that are poorly understood, and mice that do survive have stunted growth (Shindo *et al.*, 2000). *Adamts1* deficient mice have structural defects of the urinary tract, such as hydronephrosis, which can result from distal occlusion of urinary flow (Shindo *et al.*, 2000; Mittaz *et al.*, 2004). *Adamts2*-deficient mice have dermatosparaxis, as well as male infertility, the basis of which is not presently understood (Li *et al.*, 2001). *Adamts5* deficient mice are externally normal, but have pulmonic valve stenosis, bicuspid aortic valves, and myxomatous mitral valves (Dupuis *et al.*, 2011, 2013). *Adamts9*-deficient mice do not survive past gastrulation in the early embryo (Dubail *et al.*, 2014; Enomoto *et al.*, 2010), whereas mice homozygous for a spontaneous *Adamts20* mutation (*belted/bt*) survive and are apparently normal but for a belt of unpigmented skin around their torso (Rao *et al.*, 2003; Silver *et al.*, 2008). Among the ADAMTS mutant mice reported to date, only *Adamts4* and *Adamts12* mutants are developmentally normal, fertile, and have a normal lifespan (Stanton *et al.*, 2005; Boerboom *et al.*, 2011; El Hour *et al.*, 2010). However, data from combinatorial crosses of a number of ADAMTS genes suggests that the functions of ADAMTS4 and ADAMTS12 in these mice could be masked by compensation by their paralogs. Indeed, mice lacking both *Adamts1* and *Adamts4* die at birth and have severe thinning of their renal medulla (Boerboom *et al.*, 2011).

ADAMTS Proteolysis of Versican as a Crucial Requirement for Normal Embryogenesis and Neural Function

During organogenesis, the embryo ECM is of a provisional/temporary nature, rich in proteoglycans such as versican,

Table 2 Phenotypes arising from inactivation of mouse ADAMTS genes

Gene	Phenotype
<i>Adamts1</i> ^a	Lethal at birth, urogenital anomalies, myocardial non-compaction, and female infertility (Shindo <i>et al.</i> , 2000; Mittaz <i>et al.</i> , 2004; Stankunas <i>et al.</i> , 2008; Brown <i>et al.</i> , 2010)
<i>Adamts2</i>	Skin fragility and male infertility (Li <i>et al.</i> , 2001)
<i>Adamts4</i> ^a	No discernible phenotype reported (Stanton <i>et al.</i> , 2005)
<i>Adamts5</i> ^a	Pulmonic valve stenosis, mitral valve myxoma, and soft tissue syndactyly affecting hindlimbs (Dupuis <i>et al.</i> , 2011, 2013; McCulloch <i>et al.</i> , 2009)
<i>Adamts9</i> ^a	Lethal very early during embryogenesis (<i>Adamts9</i> −/−); variable cardiovascular anomalies (<i>Adamts9</i> +/−); ocular anterior segment dysgenesis (<i>Adamts9</i> +/−); soft tissue syndactyly (limb-specific conditional deletion) (Dubail <i>et al.</i> , 2014; Enomoto <i>et al.</i> , 2010; McCulloch <i>et al.</i> , 2009; Kern <i>et al.</i> , 2010; Koo <i>et al.</i> , 2010)
<i>Adamts12</i>	No developmental phenotype reported (El Hour <i>et al.</i> , 2010)
<i>Adamts13</i>	Thrombotic thrombocytopenia (Motto <i>et al.</i> , 2005)
<i>Adamts20</i> ^a	White belly spotting, soft tissue syndactyly affecting forelimbs (Rao <i>et al.</i> , 2003; Silver <i>et al.</i> , 2008; McCulloch <i>et al.</i> , 2009)

^aAdditional phenotypes were reported after combinatorial deletion with other ADAMTS alleles. *Adamts1* −/−; *Adamts4* −/− mice died within 72 h of birth and had dilated renal pelvis and thinned medulla (Boerboom *et al.*, 2011); *Adamts5* −/−; *Adamts20* −/− and *Adamts5* −/−; *Adamts9* +/− mice have greater penetrance of soft tissue syndactyly than either allele deletion alone (McCulloch *et al.*, 2009); *Adamts20* −/−; *Adamts9* +/− newborn mice have cleft palate and more severe white spotting than *Adamts20* −/− alone and die at birth (Enomoto *et al.*, 2010; Silver *et al.*, 2008).

fibronectin, and HA. As the embryo matures, and continuing into the juvenile period, provisional ECM transitions to a mature ECM containing abundant collagen, elastin, and other specialized components. Analysis of ADAMTS mutant mice has suggested that several ADAMTS proteases, such as ADAMTS1, ADAMTS4, ADAMTS5, ADAMTS9, ADAMTS15, and ADAMTS20, collectively termed proteoglycanases, participate in versican remodeling, and may work cooperatively to dismantle the provisional ECM (reviewed recently in [Nandadasa et al., 2014](#)). During the process of myocardial compaction, the versican-HA rich cardiac jelly is remodeled by ADAMTS1 ([Stankunas et al., 2008](#)). Failure of sculpting of developing pulmonic valve leaflets in *Adamts5* $-/-$ mice is a consequence of reduced cleavage of versican, since the valve phenotype can be ameliorated by reduction of versican ([Dupuis et al., 2011](#)). Reduced melanoblast colonization of skin in *Adamts20*^{bt/bt} mice is associated with reduced processing of versican in the dermis of skin ([Silver et al., 2008](#)). Mice doubly deficient in *Adamts5*, *Adamts9*, or *Adamts20* develop soft tissue syndactyly owing to failed apoptosis in interdigital web mesenchyme, accompanied by loss of versican proteolysis ([McCulloch et al., 2009](#)). Although mice singly deficient in *Adamts5* and *Adamts20* have a low penetrance of syndactyly, it is fully penetrant in mice with limb-specific deletion of *Adamts9*. The N-terminal product of versican proteolysis by ADAMTS proteases, named versikine, has been shown to induce cell death in ADAMTS-deficient interdigital webs, in which apoptosis was reduced ([McCulloch et al., 2009](#)). *Adamts9* and *Adamts20* combined deficiency leads to death at birth from a complete cleft of the secondary palate where reduced versican proteolysis is observed ([Enomoto et al., 2010](#)). These mice also have more extensively depigmented skin than *Adamts20*^{bt/bt} mice ([Silver et al., 2008](#)).

The central nervous system ECM is rich in proteoglycans, including the ADAMTS substrates aggrecan, brevican, and versican. These HA-binding molecules are particularly enriched in perineuronal nets and in glial scars. Based on upregulation of ADAMTS1, ADAMTS4, ADAMTS5, and ADAMTS9 in neural injury, it has been proposed that they have a role in mediating inflammation, repair, vascular perfusion, and neuronal plasticity. ADAMTS proteases are primarily expressed by astrocytes and expressed in several neuronal structures. ADAMTS4 cleaves brevican at the E395-S396 site and the resulting cleaved brevican can enhance glioma cell migration ([Matthews et al., 2000](#); [Viapiano et al., 2008](#)). ADAMTS4 and ADAMTS5 can cleave reelin, an ECM molecule involved in neural development and plasticity ([Hisanaga et al., 2012](#); [Koie et al., 2014](#); [Krstic et al., 2012](#)).

ADAMTS1 and Female Infertility

Adamts1 is expressed in a hormone-dependent manner during the ovulatory cycle and upregulated by the preovulatory surge of luteinizing hormone in female mice. ([Robker et al., 2000](#); [Russell et al., 2003](#)). Female *Adamts1* $-/-$ mice are subfertile due to impaired ovulation, with mature oocytes remaining trapped in unruptured follicles ([Brown et al., 2006, 2010](#)). ADAMTS1 is required for the proper assembly of the cumulus oophorus complex (COC) ECM, and is responsible for proteolysis of versican in the COC that is required for follicular

rupture and release of the oocyte. ADAMTS1 is also required for clearance of residual versican remaining after fertilization. Versican proteolysis by ADAMTS1 may be crucial for sperm penetration through the cumulus barrier prior to fertilization. ADAMTS1 and versican cleavage were associated with fertilization capacity in human oocytes. ADAMTS1 was also shown to have a role in ensuring ovarian angiogenesis and lymphangiogenesis ([Brown et al., 2006](#)) through mechanisms that are not understood.

Current Directions and Challenges

Because of the relatively recent discovery of ADAMTS proteases, their full range of biological roles remains to be unearthed. Nevertheless, it is already clear that they are extremely significant in mammalian development and human disorders. Indeed, more human genetic disorders are associated with this family than other metzincins. Their detrimental role in arthritis stimulated a great deal of interest in generating agents to block their activity, but current interest in these has waned because the normal functions of aggrecanases are not yet well understood, and there is a fear of side-effects or off-target effects. There is considerable interest in recombinant ADAMTS13 for treatment of, and engineering it to avoid targeting by autoantibodies. There have been numerous reports of altered expression of ADAMTS genes in cancer as well as their identification as tumor suppressors in various cancers, but no pivotal, direct, causal association with cancer has yet become apparent. ADAMTS1 and ADAMTS9 have been shown to be anti-angiogenic, which may be relevant to potential roles in cancer. Protein chemistry and substrate identification for this family is challenging because of the difficulty in obtaining full-length ADAMTS proteases, and the role of cofactors such as fibulins, remains to be fully explored. Relatively little is known about ADAMTS proteins in the brain and this is likely to be a fertile area for future research. Finally, very little is known about the transcriptional regulation of ADAMTS genes. This is a field with considerable potential for new discoveries.

See also: Protein Synthesis/Degradation: Protein Degradation – General: Mass Spectrometry-based Methodologies for Studying Proteolytic Networks and the Degradome. Protein Synthesis/Degradation: Protein Degradation – Protease Classes: ADAMs Regulate Cell–Cell Interactions by Controlling the Function of the EGF-Receptor, TNF α and Notch; Matrix Metalloproteinases; Metalloproteases Meprin α and Meprin β in Health and Disease

References

- Abbaszade, I., Liu, R.Q., Yang, F., et al., 1999. Cloning and characterization of ADAMTS11, an aggrecanase from the ADAMTS family. *Journal of Biological Chemistry* 274, 23443–23450.
- Akiyama, M., Takeda, S., Kokame, K., Takagi, J., Miyata, T., 2009. Crystal structures of the noncatalytic domains of ADAMTS13 reveal multiple discontinuous exosites for von Willebrand factor. *Proceedings of the National Academy of Sciences of the United States of America* 106, 19274–19279.

- Aldahmesh, M.A., Alshammari, M.J., Khan, A.O., *et al.*, 2013. The syndrome of microcornea, myopic chorioretinal atrophy, and telecanthus (MMCAT) is caused by mutations in ADAMTS18. *Human Mutation* 34, 1195–1199.
- Apte, S.S., 2004. A disintegrin-like and metalloprotease (reprolysin type) with thrombospondin type 1 motifs: The ADAMTS family. *International Journal of Biochemistry & Cell Biology* 36, 981–985.
- Apte, S.S., 2009. A disintegrin-like and metalloprotease (reprolysin-type) with thrombospondin type 1 motif (ADAMTS) superfamily: Functions and mechanisms. *Journal of Biological Chemistry* 284, 31493–31497.
- Boerboom, D., Lafond, J.F., Zheng, X., *et al.*, 2011. Partially redundant functions of Adamts1 and Adamts4 in the perinatal development of the renal medulla. *Developmental Dynamics: An Official Publication of the American Association of Anatomists* 240, 1806–1814.
- Brown, H.M., Dunning, K.R., Robker, R.L., *et al.*, 2010. ADAMTS1 cleavage of versican mediates essential structural remodeling of the ovarian follicle and cumulus-oocyte matrix during ovulation in mice. *Biology of Reproduction* 83, 549–557.
- Brown, H.M., Dunning, K.R., Robker, R.L., Pritchard, M., Russell, D.L., 2006. Requirement for ADAMTS-1 in extracellular matrix remodeling during ovarian folliculogenesis and lymphangiogenesis. *Developmental Biology* 300, 699–709.
- Colige, A., Sieron, A.L., Li, S.W., *et al.*, 1999. Human Ehlers–Danlos syndrome type VII C and bovine dermatosparaxis are caused by mutations in the procollagen I N-proteinase gene. *American Journal of Human Genetics* 65, 308–317.
- Colige, A., Vandenberghe, I., Thiry, M., *et al.*, 2002. Cloning and characterization of ADAMTS-14, a novel ADAMTS displaying high homology with ADAMTS-2 and ADAMTS-3. *Journal of Biological Chemistry* 277, 5756–5766.
- Dagoneau, N., Benoist-Lasselin, C., Huber, C., *et al.*, 2004. ADAMTS10 mutations in autosomal recessive Weill–Marchesani syndrome. *American Journal of Human Genetics* 75, 801–806.
- Dubail, J., Aramaki-Hattori, N., Bader, H.L., *et al.*, 2014. A new Adamts9 conditional mouse allele identifies its non-redundant role in interdigital web regression. *Genesis* 52 (7), 702–712.
- Dupuis, L.E., McCulloch, D.R., McGarity, J.D., *et al.*, 2011. Altered versican cleavage in ADAMTS5 deficient mice: a novel etiology of myxomatous valve disease. *Developmental Biology* 357, 152–164.
- Dupuis, L.E., Osinska, H., Weinstein, M.B., Hinton, R.B., Kern, C.B., 2013. Insufficient versican cleavage and Smad2 phosphorylation results in bicuspid aortic and pulmonary valves. *Journal of Molecular and Cellular Cardiology* 60, 50–59.
- Echtermeyer, F., Bertrand, J., Dreier, R., *et al.*, 2009. Syndecan-4 regulates ADAMTS-5 activation and cartilage breakdown in osteoarthritis. *Nature Medicine* 15, 1072–1076.
- El Hour, M., Moncada-Pazos, A., Blacher, S., *et al.*, 2010. Higher sensitivity of Adamts12-deficient mice to tumor growth and angiogenesis. *Oncogene* 29, 3025–3032.
- Enomoto, H., Nelson, C.M., Somerville, R.P., *et al.*, 2010. Cooperation of two ADAMTS metalloproteases in closure of the mouse palate identifies a requirement for versican proteolysis in regulating palatal mesenchyme proliferation. *Development* 137, 4029–4038.
- Faivre, L., Gorlin, R.J., Wirtz, M.K., *et al.*, 2003. In frame fibrillin-1 gene deletion in autosomal dominant Weill–Marchesani syndrome. *Journal of Medical Genetics* 40, 34–36.
- Fernandes, R.J., Hirohata, S., Engle, J.M., *et al.*, 2001. Procollagen II amino propeptide processing by ADAMTS-3. Insights on dermatosparaxis. *Journal of Biological Chemistry* 276, 31502–31509.
- Fosang, A.J., Little, C.B., 2008. Drug insight: Aggrecanases as therapeutic targets for osteoarthritis. *Nature Clinical Practice. Rheumatology* 4, 420–427.
- Gao, G., Plaas, A.H., Thompson, V.P., *et al.*, 2003. ADAMTS4 (aggrecanase-1) activation on the cell surface involves C-terminal cleavage by GPI-anchored MT4-MMP and binding of the activated proteinase to chondroitin sulfate and heparan sulfate on syndecan-1. *Journal of Biological Chemistry* 279 (11), 10042–10051.
- Gendron, C., Kashiwagi, M., Hughes, C., Caterson, B., Nagase, H., 2003. TIMP-3 inhibits aggrecanase-mediated glycosaminoglycan release from cartilage explants stimulated by catabolic factors. *FEBS Letters* 555, 431–436.
- Gerhardt, S., Hassall, G., Hawtin, P., *et al.*, 2007. Crystal structures of human ADAMTS-1 reveal a conserved catalytic domain and a disintegrin-like domain with a fold homologous to cysteine-rich domains. *Journal of Molecular Biology* 373, 891–902.
- Glasson, S.S., Askew, R., Sheppard, B., *et al.*, 2005. Deletion of active ADAMTS5 prevents cartilage degradation in a murine model of osteoarthritis. *Nature* 434, 644–648.
- Hashimoto, G., Aoki, T., Nakamura, H., Tanzawa, K., Okada, Y., 2001. Inhibition of ADAMTS4 (aggrecanase-1) by tissue inhibitors of metalloproteinases (TIMP-1, 2, 3 and 4). *FEBS Letters* 494, 192–195.
- Heinegard, D., Saxne, T., 2011. The role of the cartilage matrix in osteoarthritis. *Nature Reviews Rheumatology* 7, 50–56.
- Hirohata, S., Wang, L.W., Miyagi, M., *et al.*, 2002. Punctin, a novel ADAMTS-like molecule (ADAMTSL-1) in extracellular matrix. *Journal of Biological Chemistry* 277 (14), 12182–12189.
- Hisanaga, A., Morishita, S., Suzuki, K., *et al.*, 2012. A disintegrin and metalloproteinase with thrombospondin motifs 4 (ADAMTS-4) cleaves Reelin in an isoform-dependent manner. *FEBS Letters* 586, 3349–3353.
- Hubmacher, D., Apte, S.S., 2011. Genetic and functional linkage between ADAMTS superfamily proteins and fibrillin-1: A novel mechanism influencing microfibril assembly and function. *Cellular and Molecular Life Sciences* 68, 3137–3148.
- Hurskainen, T.L., Hirohata, S., Seldin, M.F., Apte, S.S., 1999. ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of zinc metalloproteases. General features and genomic distribution of the ADAM-TS family. *Journal of Biological Chemistry* 274, 25555–25563.
- Huxley-Jones, J., Apte, S.S., Robertson, D.L., Boot-Handford, R.P., 2005. The characterisation of six ADAMTS proteases in the basal chordate *Ciona intestinalis* provides new insights into the vertebrate ADAMTS family. *International Journal of Biochemistry & Cell Biology* 37, 1838–1845.
- Jeltsch, M., Jha, S.K., Tvorogov, D., *et al.*, 2014. CCBE1 enhances lymphangiogenesis via a disintegrin and metalloprotease with thrombospondin motifs-3-mediated vascular endothelial growth factor-C activation. *Circulation* 129, 1962–1971.
- Kern, C.B., Wessels, A., McGarity, J., *et al.*, 2010. Reduced versican cleavage due to Adamts9 haploinsufficiency is associated with cardiac and aortic anomalies. *Matrix biology: Journal of the International Society for Matrix Biology* 29, 304–316.
- Koie, M., Okumura, K., Hisanaga, A., *et al.*, 2014. Cleavage within Reelin repeat 3 regulates the duration and range of the signaling activity of Reelin protein. *Journal of Biological Chemistry* 289, 12922–12930.
- Koo, B.H., Coe, D.M., Dixon, L.J., *et al.*, 2010. ADAMTS9 is a cell-autonomously acting anti-angiogenic metalloprotease expressed by microvascular endothelial cells. *American Journal of Pathology* 176 (3), 1494–1504.
- Koo, B.H., Longpre, J.M., Somerville, R.P., *et al.*, 2007. Regulation of ADAMTS9 secretion and enzymatic activity by its propeptide. *Journal of Biological Chemistry* 282, 16146–16154.
- Krstic, D., Rodriguez, M., Knesel, I., 2012. Regulated proteolytic processing of Reelin through interplay of tissue plasminogen activator (tPA), ADAMTS-4, ADAMTS-5, and their modulators. *PLoS One* 7, e47793.
- Kuno, K., Kanada, N., Nakashima, E., *et al.*, 1997. Molecular cloning of a gene encoding a new type of metalloproteinase-disintegrin family protein with thrombospondin motifs as an inflammation associated gene. *Journal of Biological Chemistry* 272, 556–562.
- Kuno, K., Matsushima, K., 1998. ADAMTS-1 protein anchors at the extracellular matrix through the thrombospondin type I motifs and its spacing region. *Journal of Biological Chemistry* 273, 13912–13917.
- Kuno, K., Terashima, Y., Matsushima, K., 1999. ADAMTS-1 is an active metalloproteinase associated with the extracellular matrix. *Journal of Biological Chemistry* 274, 18821–18826.
- Kutz, W.E., Wang, L.W., Bader, H.L., *et al.*, 2011. ADAMTS10 protein interacts with fibrillin-1 and promotes its deposition in extracellular matrix of cultured fibroblasts. *Journal of Biological Chemistry* 286, 17156–17167.
- Lapiere, C.M., Nusgens, B.V., 1993. Ehlers–Danlos type VII-C, or human dermatosparaxis. The offspring of a union between basic and clinical research [editorial; comment]. *Archives of Dermatology* 129, 1316–1319.
- Lark, M.W., Bayne, E.K., Flanagan, J., *et al.*, 1997. Aggrecan degradation in human cartilage. Evidence for both matrix metalloproteinase and aggrecanase activity in normal, osteoarthritic, and rheumatoid joints. *Journal of Clinical Investigation* 100, 93–106.
- Lesnik Oberstein, S.A., Kriek, M., White, S.J., *et al.*, 2006. Peters Plus syndrome is caused by mutations in B3GALT1, a putative glycosyltransferase. *American Journal of Human Genetics* 79, 562–566.
- Levy, G.G., Nichols, W.C., Lian, E.C., *et al.*, 2001. Mutations in a member of the ADAMTS gene family cause thrombotic thrombocytopenic purpura. *Nature* 413, 488–494.
- Li, S.W., Arita, M., Fertala, A., *et al.*, 2001. Transgenic mice with inactive alleles for procollagen N-proteinase (ADAMTS-2) develop fragile skin and male sterility. *Biochemical Journal* 355, 271–278.
- Little, C.B., Meeker, C.T., Golub, S.B., *et al.*, 2007. Blocking aggrecanase cleavage in the aggrecan interglobular domain abrogates cartilage erosion and promotes cartilage repair. *Journal of Clinical Investigation* 117, 1627–1636.

- Little, C.B., Mittaz, L., Belluoccio, D., *et al.*, 2005. ADAMTS-1-knockout mice do not exhibit abnormalities in aggrecan turnover *in vitro* or *in vivo*. *Arthritis and Rheumatism* 52, 1461–1472.
- Luo, Y., Koles, K., Vorndam, W., Haltiwanger, R.S., Panin, V.M., 2006. Protein O-fucosyltransferase 2 adds O-fucose to thrombospondin type 1 repeats. *Journal of Biological Chemistry* 281, 9393–9399.
- Mahmoodi, M., Sahebjam, S., Smookler, D., Khokha, R., Mort, J.S., 2005. Lack of tissue inhibitor of metalloproteinases-3 results in an enhanced inflammatory response in antigen-induced arthritis. *American Journal of Pathology* 166, 1733–1740.
- Majerus, E.M., Zheng, X., Tuley, E.A., Sadler, J.E., 2003. Cleavage of the ADAMTS13 propeptide is not required for protease activity. *Journal of Biological Chemistry* 278, 46643–46648.
- Matthews, R.T., Gary, S.C., Zerillo, C., *et al.*, 2000. Brain-enriched hyaluronan binding (BEHAB)/brevican cleavage in a glioma cell line is mediated by a disintegrin and metalloproteinase-3 results in an enhanced inflammatory response in antigen-induced arthritis. *American Journal of Pathology* 166, 1733–1740.
- McCulloch, D.R., Nelson, C.M., Dixon, L.J., *et al.*, 2009. ADAMTS metalloproteases generate active versican fragments that regulate interdigital web regression. *Developmental Cell* 17, 687–698.
- Mittaz, L., Russell, D.L., Wilson, T., *et al.*, 2004. Adamts-1 is essential for the development and function of the urogenital system. *Biology of Reproduction* 70, 1096–1105.
- Morales, J., Al-Sharif, L., Khalil, D.S., *et al.*, 2009. Homozygous mutations in ADAMTS10 and ADAMTS17 cause lenticular myopia, ectopia lentis, glaucoma, spherophakia, and short stature. *American Journal of Human Genetics* 85, 558–568.
- Mosyak, L., Georgiadis, K., Shane, T., *et al.*, 2008. Crystal structures of the two major aggrecan degrading enzymes, ADAMTS4 and ADAMTS5. *Protein Science* 17, 16–21.
- Motto, D.G., Chauhan, A.K., Zhu, G., *et al.*, 2005. Shigatoxin triggers thrombotic thrombocytopenic purpura in genetically susceptible ADAMTS13-deficient mice. *Journal of Clinical Investigation* 115, 2752–2761.
- Nandadasa, S., Foulcer, S., Apte, S.S., 2014. The multiple, complex roles of versican and its proteolytic turnover by ADAMTS proteases during embryogenesis. *Matrix biology: Journal of the International Society for Matrix Biology* 35, 34–41.
- Nishiwaki, K., Kubota, Y., Chigira, Y., *et al.*, 2004. An NDPase links ADAM protease glycosylation with organ morphogenesis in *C. elegans*. *Nature Cell Biology* 6, 31–37.
- Nolasco, L.H., Turner, N.A., Bernardo, A., *et al.*, 2005. Hemolytic uremic syndrome-associated Shiga toxins promote endothelial-cell secretion and impair ADAMTS13 cleavage of unusually large von Willebrand factor multimers. *Blood* 106, 4199–4209.
- Rao, C., Foerzler, D., Loftus, S.K., *et al.*, 2003. A defect in a novel ADAMTS family member is the cause of the belted white-spotting mutation. *Development* 130, 4665–4672.
- Ricketts, L.M., Dlugosz, M., Luther, K.B., Haltiwanger, R.S., Majerus, E.M., 2007. O-fucosylation is required for ADAMTS13 secretion. *Journal of Biological Chemistry* 282, 17014–17023.
- Robker, R.L., Russell, D.L., Espey, L.L., *et al.*, 2000. Progesterone-regulated genes in the ovulation process: ADAMTS-1 and cathepsin L proteases. *Proceedings of the National Academy of Sciences of the United States of America* 97, 4689–4694.
- Russell, D.L., Doyle, K.M., Ochsner, S.A., Sandy, J.D., Richards, J.S., 2003. Processing and localization of ADAMTS-1 and proteolytic cleavage of versican during cumulus matrix expansion and ovulation. *Journal of Biological Chemistry* 278, 42330–42339.
- Shieh, H.S., Mathis, K.J., Williams, J.M., *et al.*, 2008. High resolution crystal structure of the catalytic domain of ADAMTS-5 (aggrecanase-2). *Journal of Biological Chemistry* 283, 1501–1507.
- Shindo, T., Kurihara, H., Kuno, K., *et al.*, 2000. ADAMTS-1: A metalloproteinase-disintegrin essential for normal growth, fertility, and organ morphology and function. *Journal of Clinical Investigation* 105, 1345–1352.
- Silver, D.L., Hou, L., Somerville, R., *et al.*, 2008. The secreted metalloprotease ADAMTS20 is required for melanoblast survival. *PLoS Genetics* 4, 1–15.
- Somerville, R.P., Longpre, J.M., Apel, E.D., *et al.*, 2004. ADAMTS7B, the full-length product of the ADAMTS7 gene, is a chondroitin sulfate proteoglycan containing a mucin domain. *Journal of Biological Chemistry* 279, 35159–35175.
- Stankunas, K., Hang, C.T., Tsun, Z.Y., *et al.*, 2008. Endocardial Brg1 represses ADAMTS1 to maintain the microenvironment for myocardial morphogenesis. *Developmental Cell* 14, 298–311.
- Stanton, H., Rogerson, F.M., East, C.J., *et al.*, 2005. ADAMTS5 is the major aggrecanase in mouse cartilage *in vivo* and *in vitro*. *Nature* 434, 648–652.
- Tan, K., Duquette, M., Liu, J.H., *et al.*, 2002. Crystal structure of the TSP-1 type 1 repeats: A novel layered fold and its biological implication. *Journal of Cell Biology* 159, 373–382.
- Tortorella, M.D., Arner, E.C., Hills, R., *et al.*, 2004. Alpha2-macroglobulin is a novel substrate for ADAMTS-4 and ADAMTS-5 and represents an endogenous inhibitor of these enzymes. *Journal of Biological Chemistry* 279, 17554–17561.
- Tortorella, M.D., Burn, T.C., Pratta, M.A., *et al.*, 1999. Purification and cloning of aggrecanase-1: A member of the ADAMTS family of proteins [see comments]. *Science* 284, 1664–1666.
- Troeberg, L., Fushimi, K., Khokha, R., *et al.*, 2008. Calcium pentosan polysulfate is a multifaceted exosite inhibitor of aggrecanases. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 22, 3515–3524.
- Viapiano, M.S., Hockfield, S., Matthews, R.T., 2008. BEHAB/brevican requires ADAMTS-mediated proteolytic cleavage to promote glioma invasion. *Journal of Neuro-oncology* 88, 261–272.
- Wainwright, S.D., Bondeson, J., Caterson, B., Hughes, C.E., 2013. ADAMTS-4_v1 is a splice variant of ADAMTS-4 that is expressed as a protein in human synovium and cleaves aggrecan at the interglobular domain. *Arthritis and Rheumatism* 65, 2866–2875.
- Wang, L.W., Dlugosz, M., Somerville, R.P., *et al.*, 2007. O-fucosylation of thrombospondin type 1 repeats in ADAMTS-like-1/punctin-1 regulates secretion: Implications for the ADAMTS superfamily. *Journal of Biological Chemistry* 282, 17024–17031.
- Wang, L.W., Leonhard-Melief, C., Haltiwanger, R.S., Apte, S.S., 2009. Post-translational modification of thrombospondin type-1 repeats in ADAMTS-like 1/punctin-1 by C-mannosylation of tryptophan. *Journal of Biological Chemistry* 284, 30004–30015.
- Wu, J.J., Fujikawa, K., McMullen, B.A., Chung, D.W., 2006. Characterization of a core binding site for ADAMTS-13 in the A2 domain of von Willebrand factor. *Proceedings of the National Academy of Sciences of the United States of America* 103, 18470–18474.
- Yamamoto, K., Troeberg, L., Scilabra, S.D., *et al.*, 2013. LRP-1-mediated endocytosis regulates extracellular activity of ADAMTS-5 in articular cartilage. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 27, 511–521.
- Zheng, X.L., 2013. Structure-function and regulation of ADAMTS-13 protease. *Journal of Thrombosis and Haemostasis: JTH* 11 (Suppl. 1), 11–23.
- Zheng, X., Majerus, E.M., Sadler, J.E., 2002. ADAMTS13 and TTP. *Current Opinion in Hematology* 9, 389–394.

Further Reading

- Frantz, C., Stewart, K.M., Weaver, V.M., 2010. The extracellular matrix at a glance. *Journal of Cell Science* 123, 4195–4200. PMID: 21123617.
- Heinegård, D., Saxne, T., 2011. The role of the cartilage matrix in osteoarthritis. *Nature Reviews Rheumatology* 7, 50–56. PMID: 21119607.
- Hubmacher, D., Apte, S.S., 2011. Genetic and functional linkage between ADAMTS superfamily proteins and fibrillin-1: A novel mechanism influencing microfibril assembly and function. *Cellular and Molecular Life Sciences* 68, 3137–3148. PMID: 21858451.
- Nandadasa, S., Foulcer, S., Apte, S.S., 2014. The multiple, complex roles of versican and its proteolytic turnover by ADAMTS proteases during embryogenesis. *Matrix Biology* 35, 34–41. PMID: 24444773.
- Turner, N., Nolasco, L., Moake, J., 2012. Generation and breakdown of soluble ultralarge von Willebrand factor multimers. *Seminars in Thrombosis and Hemostasis* 38, 38–46. PMID: 22314602.

Relevant Websites

- <http://flybase.org>
FlyBase.
- <http://informatics.jax.org>
Mouse Genome Informatics.
- <http://omim.org>
OMIM.
- www.rcsb.org
RCSB Protein Data Bank.
- www.wormbase.org
WormBase.

ADAMs Regulate Cell–Cell Interactions by Controlling the Function of the EGF-Receptor, TNF α and Notch

CP Blobel, Hospital for Special Surgery, New York, NY, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

ADAM A disintegrin and metalloproteinase; conditional knockout mice, mice engineered to carry loxP recombination sites surrounding a gene of interest, this gene can be conditionally inactivated in a specific cell type or tissue or in a temporal manner if the Cre-recombinase,

which deletes the sequence between the loxP sites, is expressed under the control of a tissue specific promoter, or a promoter that can be induced in a temporal manner.

Notch A membrane-anchored transcription factor that is activated by proteolytic processing, mutations in *Drosophila* Notch cause a notched wing tip.

Introduction

Proteolytic processing of membrane proteins is a vital component of cell–cell interactions. It is crucial for normal development and for healthy adulthood, but can also contribute to the pathogenesis of diseases such as cancer and rheumatoid arthritis (Blobel, 2005). This process, which is also referred to as protein ectodomain shedding (Figure 1), allows the release of growth factors and cytokines from their membrane anchors, and can activate or inactivate cell surface molecules or change their functional properties (Blobel, 2005). The regulated proteolytic release of a growth factor or cytokine from its membrane anchor fulfills a similar function as the regulated release of signaling molecules from a secretory vesicle: it allows storage of the growth factor and rapid and regulated release upon demand (Figure 2). A family of metalloproteinases called a disintegrin and metalloproteinase (ADAMs) because they contain a disintegrin and metalloproteinase domain (see domain organization in Figure 3) are key players in protein ectodomain shedding. Out of the approximately 30 ADAMs

found in mammalian genomes, ADAM17 and ADAM10 stand out for their essential roles in development (Blobel, 2005). ADAM17 is essential for activating the epidermal growth factor receptor (EGFR) during development (Peschon *et al.*, 1998; reviewed in Blobel, 2005) and is later required to maintain the skin and intestinal barrier in adults (Blaydon *et al.*, 2011; Franzke *et al.*, 2012; Chalaris *et al.*, 2010). In addition, ADAM17 regulates the release of the pro-inflammatory cytokine TNF α (Black *et al.*, 1997; Moss *et al.*, 1997), and thus has a role in TNF α -dependent diseases, including rheumatoid arthritis (Issuree *et al.*, 2013) and Crohn's disease (Chalaris *et al.*, 2010). The related ADAM10 is essential for the activation of the Notch signaling pathway (Rooke *et al.*, 1996; Hartmann *et al.*, 2002; Sotillos *et al.*, 1997), a pathway that is used to control a variety of cell fate decision during development and in adults (Kopan and Ilagan, 2009). Other ADAMs also have interesting functions, such as ADAM19, which is involved in heart valve development (Zhou *et al.*, 2004; Kurohara *et al.*, 2004), or ADAMs 9, 12, and 15, which are thought to have important roles in cancer (Peduto *et al.*, 2006, 2005; Najj *et al.*, 2008). However, since the functions of ADAM10 and 17 and the underlying mechanisms are currently better understood than those of other ADAMs, this article will focus on these two membrane-anchored metalloproteinases and highlight our understanding of their roles as key mediators of cell–cell interactions in development and disease.

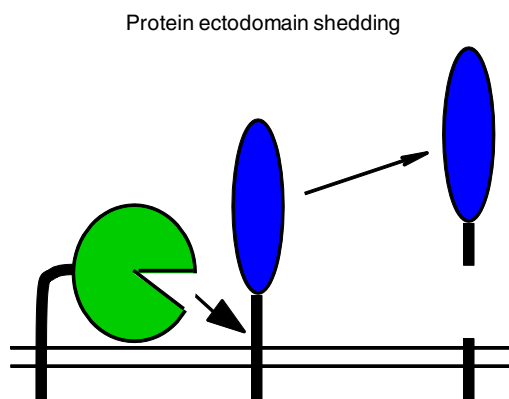


Figure 1 Protein ectodomain shedding. The term 'protein ectodomain shedding' refers to the proteolytic release of membrane-anchored molecules from cell membranes. A membrane-anchored protease (green) is shown processing a membrane-anchored growth factor close to its membrane anchor, thereby releasing the soluble extracellular domain of the growth factor. Reproduced from Blobel, C.P., Carpenter, G., Freeman, M., 2009. The role of protease activity in ErbB biology. *Experimental Cell Research* 315, 671–682, Figure 1(A).

ADAM17, the TNF α Convertase

ADAM17 was first discovered in a search for molecules that regulate the release of the potent pro-inflammatory cytokine tumor necrosis factor α (TNF α) (Black *et al.*, 1997; Moss *et al.*, 1997). TNF α is considered to be the fire alarm of the body because it can help to activate an innate immune response, but TNF α can also cause damage to the body if its production and release is inappropriately activated, leading to autoimmune disease. TNF α is synthesized as a trimeric molecule in which each subunit is tethered to the cell surface through a membrane anchor (Figure 4(a)). These membrane tethers must be cut by the TNF α convertase (TACE, an alternative name for ADAM17) to release the soluble form of TNF α (Figure 4(b)), and therefore TACE/ADAM17 is considered a potential

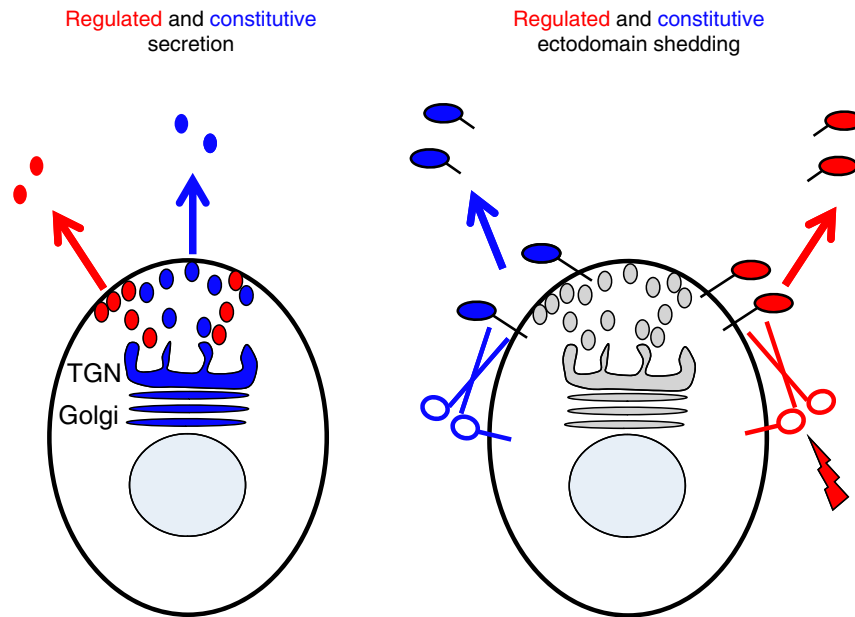


Figure 2 Conceptual similarities between secretion from vesicles and ectodomain shedding of proteins. Regulated and constitutive secretion of proteins from intracellular vesicles that fuse with the plasma membrane is a well-established means of regulating cell–cell communications. Regulated and constitutive protein ectodomain shedding uses a very different mechanism, i.e., proteolysis, to accomplish something very similar. By cleaving the membrane tether of growth factors or cytokines in a regulated or constitutive manner, the proteases involved in protein ectodomain shedding, i.e., ADAM10 and ADAM17 and in some cases also other enzymes, can control the signaling via these molecules and thus their role in cell–cell interactions.

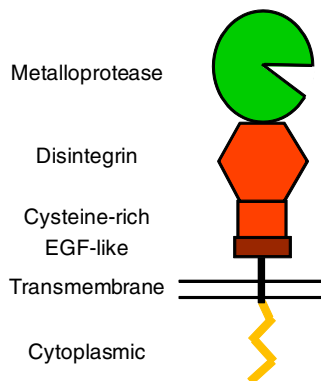


Figure 3 Domain Structure of an ADAM. ADAMs (a disintegrin and metalloprotease) are a family of membrane-anchored metalloproteinases that contain an N-terminal signal sequence and pro-domain (not shown), followed by a metalloprotease domain, a disintegrin domain that received this name because it is related to snake venom toxins called disintegrins, a cysteine-rich and EGF-like domain and a transmembrane domain followed by a cytoplasmic tail. Reproduced from Blobel, C.P., Carpenter, G., Freeman, M., 2009. The role of protease activity in ErbB biology. *Experimental Cell Research* 315, 671–682, Figure 1(B); Schlöndorff, J., Blobel, C.P., 1999. Metalloprotease-disintegrins: Modular proteins capable of promoting cell–cell interactions and triggering signals by protein–ectodomain shedding. *Journal of Cell Science* 112, 3603–3617, Figure 1, left panel, modified.

alternative target for treatment of diseases such as rheumatoid arthritis and Crohn's disease that depend on the release of soluble TNF α (Feldmann and Maini, 2008).

The typical mammalian genome contains over 500 proteases (Overall and Blobel, 2007), so once ADAM17 was identified as the TNF α convertase (Black *et al.*, 1997; Moss *et al.*, 1997), it was important to validate that it is indeed the physiologically relevant TNF α convertase and that none of the other proteases could take over its function if it is inactivated. The validation that ADAM17 is indeed the relevant TNF α convertase was accomplished using an assay that measures the function of soluble TNF α in mice. This assay, called endotoxin shock, relies on a rapid activation of the innate immune system by lipopolysaccharide (LPS, also called endotoxin), a component of the bacterial cell wall. Injection of LPS into mice can activate pattern recognition receptors (e.g., Toll-like receptor 4) to initiate the production and release of high levels of TNF α from myeloid cells, with eventually lethal consequences for the host (septic shock). Conditional knockout mice lacking ADAM17 in myeloid cells (immune cells that respond to LPS and produce TNF α) are protected from endotoxin shock and survive this challenge under conditions where wild type mice do not. Moreover, these conditional ADAM17 knockout mice have significantly lower levels of TNF α in their blood stream, corroborating that ADAM17 is the physiologically relevant TNF α convertase *in vivo*, and that no other protease in mice can effectively take over the release of TNF α in its absence (Horiuchi *et al.*, 2007).

ADAM17 is Crucial for EGFR Signaling

It is not uncommon for proteases and other biologically active molecules to have more than one major function, and this is

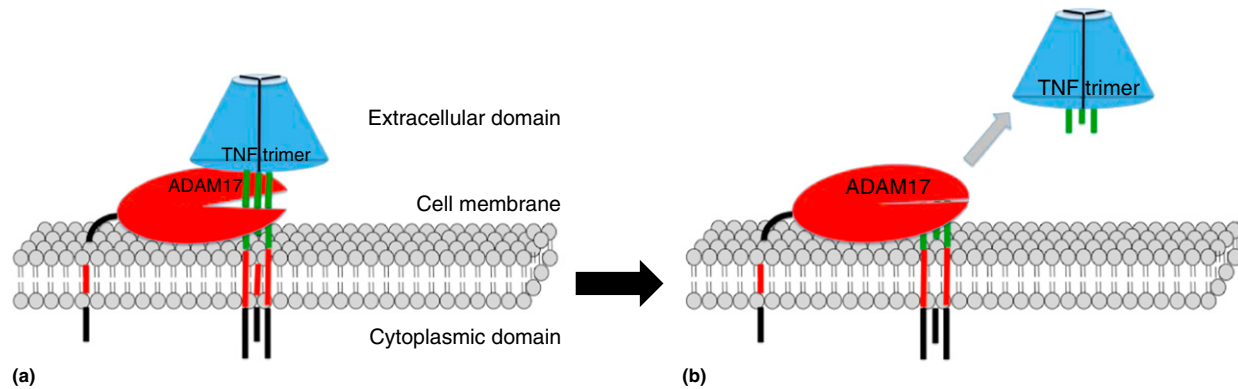


Figure 4 Protein ectodomain shedding of $\text{TNF}\alpha$ by ADAM17. The potent pro-inflammatory cytokine $\text{TNF}\alpha$ is generated as a trimeric molecule with three membrane tethers that must be cut by the $\text{TNF}\alpha$ -convertase (TACE, also referred to as ADAM17) to release the soluble form of $\text{TNF}\alpha$. Since the soluble form is thought to be responsible for autoimmune diseases such as Rheumatoid Arthritis (RA), ADAM17 is considered a potential target for treatment of RA.

also the case for ADAM17. This became evident when ADAM17 was completely inactivated in mice (knockout mice). The resulting ADAM17-deficient animals were born with open eyes and died shortly after birth (Peschon *et al.*, 1998), whereas normal wild type mice are born with closed eyes and survive. When genetically modified mice are born with open eyes, this is usually an indication of defects in the EGFR signaling pathway (Sibilia and Wagner, 1995) (although mutations in other pathways can also cause open eyes at birth). The EGFR is a cell surface tyrosine kinase receptor that is important for proper development of heart valves and for establishing a normal bone growth plate during mouse development. In adult mice and humans, a major role of the EGFR is thought to be the maintenance of the skin and intestinal barrier (see below, Jackson *et al.*, 2003; Sibilia *et al.*, 2003; Lichtenberger *et al.*, 2013). Moreover, inappropriate activation of the EGFR can lead to cancer, and therefore antibodies against the EGFR or drugs that block its kinase activity are used to treat EGFR-dependent cancers, such as colon cancer (Modjtahedi and Essapen, 2009; Wheeler *et al.*, 2010). An analysis of ADAM17 knockout mice revealed heart valve defects and abnormalities in the growth plate of the bone that closely resembled those observed in mice lacking the EGFR (Hall *et al.*, 2013; Jackson *et al.*, 2003), providing strong genetic evidence that ADAM17 has a key role in activating the EGFR signaling pathway.

Why is ADAM17 essential for the normal functioning of the EGFR tyrosine kinase receptor? The EGFR has 7 known ligands (EGFR-ligands), each of which is attached to the cell membrane via a membrane tether (Harris *et al.*, 2003). Unless the membrane tether is cut, EGFR-ligands cannot be released from cells to activate the EGFR (Blobel, 2005). ADAM17 is required for the shedding of several key EGFR-ligands, including transforming growth factor α ($\text{TGF}\alpha$), Heparin-binding epidermal growth factor (HB-EGF) and amphiregulin, which explains why inactivation of ADAM17 causes defects that closely resemble those seen in mice lacking the EGFR (Sahin *et al.*, 2004; Sunnarborg *et al.*, 2002) or HB-EGF (Jackson *et al.*, 2003), $\text{TGF}\alpha$ (Peschon *et al.*, 1998) or amphiregulin (Sternlicht *et al.*, 2005).

ADAM17-Dependent EGFR Activation Protects the Skin and Intestinal Barrier

Mouse models are commonly used to understand the function of specific molecules in the context of an intact organism *in vivo*. Because mice lacking ADAM17 die at birth with open eyes and heart valve defects (Jackson *et al.*, 2003; Peschon *et al.*, 1998), it was necessary to circumvent their postnatal lethality to gain better insights into the function of ADAM17 in adult animals. This was accomplished by generating 'hypomorphic' mice, so animals with strongly reduced activity of ADAM17 that was nevertheless sufficient for the mutant animals to survive (Chalaris *et al.*, 2010) or conditional knockout mice that allowed selective inactivation of ADAM17 in specific tissues in adult mice (Horiuchi *et al.*, 2007; Franzke *et al.*, 2012). These approaches revealed that a major function of ADAM17 is to protect the skin and intestinal barrier, most likely by activating the EGFR (Lichtenberger *et al.*, 2013). Activation of the EGFR is thought to control the production of enzymes called transglutaminases, which crosslink proteins in the skin, and presumably also in the intestine, thus ensuring proper barrier functions (Franzke *et al.*, 2012). How relevant are studies in mice for the understanding of human development and disease? The recent discovery of a patient lacking ADAM17 suggests that mouse models are highly relevant and informative in terms of understanding the function of the ADAM17/EGFR pathway in humans (Blaydon *et al.*, 2011). The ADAM17-deficient patient suffers from skin and intestinal inflammation that is in many ways very similar to the defects observed in ADAM17-deficient mice. This suggests that the skin and intestinal inflammation in the ADAM17-deficient patient can most likely be explained by inadequate activation of the EGFR, because there is no ADAM17-dependent EGFR-ligand shedding. Hopefully the information that has emerged from basic biomedical research on ADAM17 and its role in EGFR signaling using mouse models of disease will some day become useful for the treatment of ADAM17/EGFR-dependent cancers by inhibiting ADAM17 and thus the release of EGFR-ligands. On the other hand, these mechanistic studies might also help devise a custom-tailored treatment for

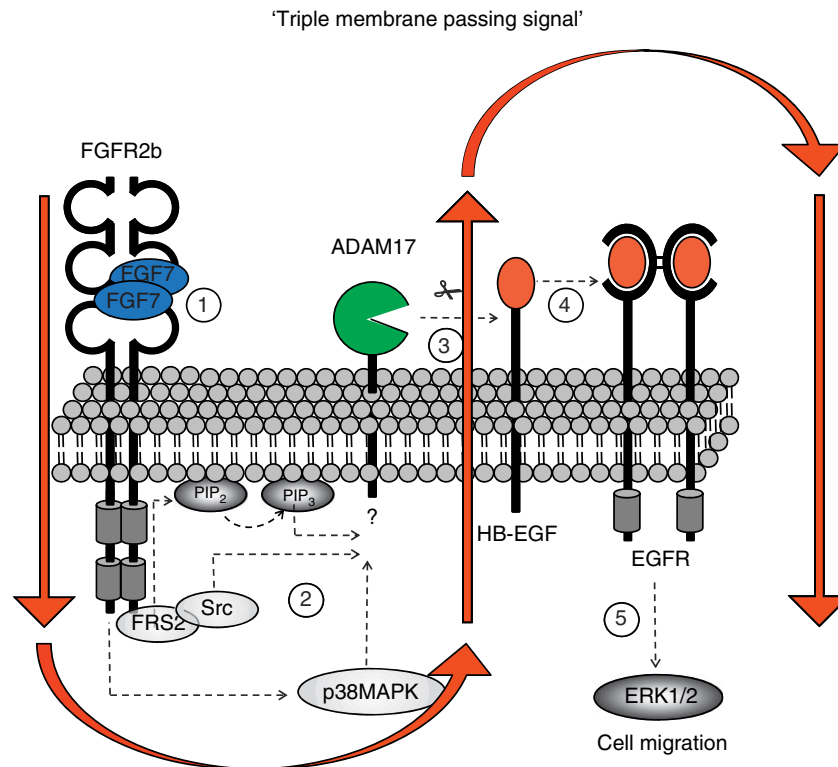


Figure 5 A model for transactivation of the EGFR/ERK1/2 signaling pathway by FGFR2b-dependent activation of ADAM17. Binding of FGF7 to the FGFR2b (1) stimulates ADAM17 in a manner that can be blocked by inhibitors of Src kinases, p38 MAP-kinase and of phosphatidylinositol 3 (PI3) kinase, which is responsible for the conversion of phosphatidylinositol 4,5-bisphosphate (PIP₂) to phosphatidylinositol 3,4,5-trisphosphate (PIP₃) (2). Stimulation of ADAM17 by FGF7 requires its transmembrane domain, but not the cytoplasmic domain. Activation of ADAM17 triggers release of the EGFR-ligand HB-EGF (3), which activates the EGFR (4) and ERK1/2 (5), thereby promoting FGF7-stimulated cell migration in keratinocytes. Both the FGFR2b and the EGFR are known to have critical roles in skin repair and inflammation of the skin. Reproduced from Maretzky, T., Evers, A., Zhou, W., *et al.*, 2011. Migration of growth factor-stimulated epithelial and endothelial cells depends on EGFR transactivation by ADAM17. *Nature Communications* 2, 229. doi:10.1038/ncomms1232, Figure 8, modified, with permission from Nature Publishing Group.

ADAM17-deficient patients, for example by applying soluble EGFR-ligands to replace the missing soluble growth factors and activate the EGFR.

ADAM17/TNF α and ADAM17/EGFR Signaling are Controlled by Upstream Regulators Called iRhoms

Protection of the skin and intestinal barrier by the ADAM17/EGFR pathway is a highly dynamic process that allows a rapid response to injury and precisely controlled activity for normal barrier maintenance. ADAM17-dependent activation of the EGFR can be very rapidly stimulated (within minutes) by a variety of different signaling pathways, including G-protein coupled receptors, tyrosine kinase receptors and even mechanical stress (Fischer *et al.*, 2003; Maretzky *et al.*, 2011). In keratinocytes, for example, the addition of keratinocyte growth factor activates ADAM17-dependent processing of EGFR-ligands, which in turn activates the EGFR signaling pathway and promotes cell migration (Maretzky *et al.*, 2011) (see Figure 5 for a diagram of this pathway, which is referred to as a triple membrane passing signal). How exactly is ADAM17 itself activated? Cell biological studies have demonstrated that the

transmembrane domain of ADAM17 is important for its rapid activation, whereas the cytoplasmic domain is not. The subsequent discovery of the seven-membrane spanning inactive Rhomboid-like protein 2 (iRhom2) as a binding partner of ADAM17 provided the first evidence for the existence of a molecule that could potentially regulate ADAM17 by interacting with its transmembrane domain (Adrain *et al.*, 2012; Maretzky *et al.*, 2013; McIlwain *et al.*, 2012) (see Figure 6 for a diagram of the putative interaction between iRhom2 and ADAM17).

iRhom2 and the related iRhom1 are similar to members of a family of intramembrane serine proteinases called Rhomboids, but lack the catalytic site and are therefore considered inactive Rhomboid-like proteins (hence the name iRhom) (Adrain and Freeman, 2012). Studies in knockout mice have shown that iRhom2 is critical for the function of ADAM17 in immune cells. Mice lacking iRhom2 have no mature ADAM17 in immune cells, and as a consequence, are protected from septic shock (no TNF α can be released) and inflammatory arthritis, a mouse model for rheumatoid arthritis, just like mice lacking ADAM17 in myeloid cells (McIlwain *et al.*, 2012; Issuree *et al.*, 2013). iRhom2-knockout mice are healthy and appear normal because the related iRhom1 can support the function and maturation of ADAM17 in all tissues except for

Model of putative interaction between ADAM17 and iRhomb2

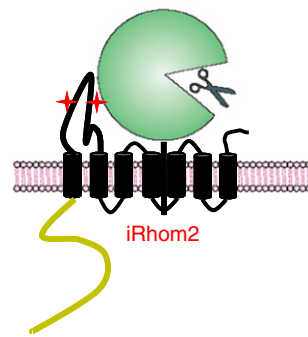


Figure 6 Model of the interaction between iRhomb2 and ADAM17. ADAM17 is thought to be regulated by an interaction with the seven-membrane spanning molecule iRhomb2. This diagram shows a putative interaction between iRhomb2 and ADAM17, although it is important to emphasize that the exact nature of the interaction between these two molecules remains to be established. Reproduced from Blobel, C.P., 2005. ADAMs: Key players in EGFR-signaling, development and disease. *Nature Reviews Molecular Cell Biology* 6, 32–43, Figure 2, modified, with permission from Nature Publishing Group.

immune cells in these animals, where iRhomb1 is not expressed at high enough levels to support the maturation of ADAM17. However, when both iRhoms 1 and 2 are inactivated, then the function of ADAM17 appears to be completely lost. Therefore iRhomb1/2 double knockout resemble ADAM17 knockout mice or EGFR knockout mice in that they have open eyes at birth, heart valve defects and a defective bone growth plate (Blobel C, manuscript submitted). This strong genetic evidence for a role of iRhoms in regulating ADAM17-dependent EGFR signaling is further backed up by cell biological studies that show no mature ADAM17 and no activated EGFR in any cell type or tissue of iRhomb1/2 double knockout mice as well as no release of EGFR-ligands from iRhomb 1/2 knockout cells.

The EGFR signaling pathway also exists in the fruit fly *Drosophila melanogaster*, which has several EGFR-ligands (Spitz, Keren, and Gurken) that are also tethered to cells with a membrane anchor (Lee *et al.*, 2001). However, the protease that releases EGFR-ligands in *Drosophila* is actually an active Rhomboid, and thus an intramembrane serine proteinase (Urban *et al.*, 2001; Urban, 2002; Figure 7). So the release of EGFR-ligands is accomplished by a very different type of protease in flies and in mammals, where the metalloprotease ADAM17 is responsible for releasing EGFR-ligands (Blobel *et al.*, 2009), aided by the inactive Rhomboid-like iRhoms 1 and 2. Perhaps the identity of the protease responsible EGFR ligand-release changed during evolution because the iRhomb/ADAM17 complex allows a more rapid response to stimuli, and thus was perhaps better suited to protect and maintain the skin and intestinal barrier than the active Rhomboids that are used to process EGFR-ligands and activate the EGFR in flies.

ADAM10, A Crucial Regulator of Notch Signaling

ADAM10 is closely related to ADAM17, yet it has very different functions in that it is crucial for Notch signaling during

The EGFR signaling pathway is regulated by different proteases in flies and mammals

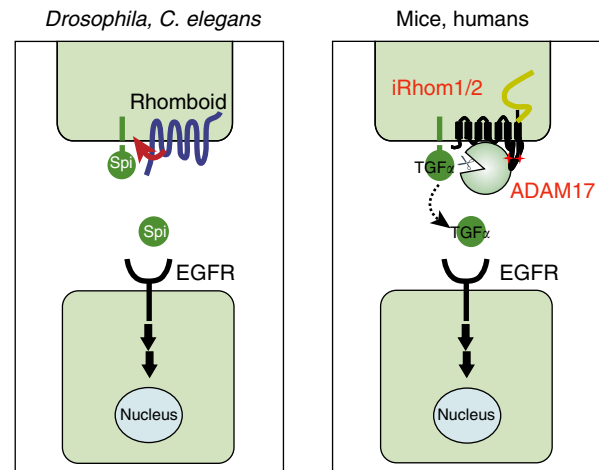


Figure 7 The EGFR pathway is regulated by different proteases in fruit flies and mammals. The main components of the EGFR pathway in *Drosophila melanogaster* are shown on the left. The intramembrane protease Rhomboid processes the EGFR-ligand Spitz (Spi, considered a functional equivalent of $\text{TGF}\alpha$ in mammals), allowing it to be released from the membrane of one cell in order to activate the EGFR on a second cell. In mammals, the EGFR-ligand $\text{TGF}\alpha$ is processed by a different protease, ADAM17, but with very similar consequences in that $\text{TGF}\alpha$ is released from its membrane anchor so that it can activate the EGFR on a different cell. The function of ADAM17 depends on the presence of iRhoms 1 or 2. The relationship between the active Rhomboid in flies and the inactive Rhomboid in mammals raises interesting questions about how and why the responsibility for cleaving EGFR-ligands was transferred from the Rhomboid in flies to the iRhomb1/2/ADAM17 complex in mammals during evolution. Since one of the main functions of the ADAM17/EGFR pathway in mammals is protection of the skin and intestinal barrier, it is tempting to speculate that the ability of the iRhomb1/2/ADAM17/EGFR pathway to rapidly respond to various stimuli, for example those released by injury, provided an evolutionary advantage that favored the mammalian proteolytic system to process EGFR-ligands. Reproduced from Blobel, C.P., Carpenter, G., Freeman, M., 2009. The role of protease activity in ErbB biology. *Experimental Cell Research* 315, 671–682, with permission from Nature Publishing Group.

development (Hartmann *et al.*, 2002; Rooke *et al.*, 1996; Sotillos *et al.*, 1997). Notch is a membrane-anchored transcriptional regulator that was initially discovered in *Drosophila*, where mutations in this gene resulted in a notched wing tip (Artavanis-Tsakonas *et al.*, 1995). Notch contains a large extracellular domain, a single transmembrane domain and an intracellular domain (Notch intracellular domain (NICD)) that can activate Notch-dependent transcription once it has entered the nucleus (Kopan and Ilagan, 2009; Figure 8). How does the membrane-anchored NICD manage to enter the nucleus? The answer is regulated intramembrane proteolysis, which consists of sequential proteolytic events that ultimately liberate the NICD from its membrane anchor (Kopan and Ilagan, 2009). First, Notch is processed by the pro-protein convertase furin during transit to the cell surface (referred to as S1 cleavage), but the actual regulatory step is the second processing event (S2 cleavage), which depends on ADAM10

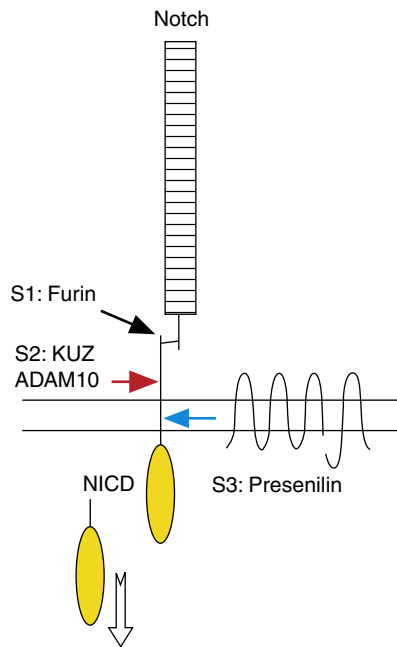


Figure 8 Domain organization of Notch and its processing by three different enzymes. Notch is a membrane-anchored transcriptional regulator with a large extracellular domain, a transmembrane domain and an intracellular domain (NICD) that can be released from the membrane and enter the nucleus to activate Notch-dependent transcription. The approximate positions of the three major cleavage sites in Notch are shown: the Furin/pro-protein convertase-dependent S1 site, the ADAM10/Kuzbanian (KUZ)-dependent S2 site, and the S3 site, which depends on processing by the γ -secretase complex (only presenilin is shown as a seven-membrane spanning molecule, please note that the γ -secretase contains additional subunits that are not shown). Reproduced from Schlöndorff, J., Blobel, C.P., 1999. Metalloprotease-disintegrins: Modular proteins capable of promoting cell-cell interactions and triggering signals by protein-ectodomain shedding. *Journal of Cell Science* 112, 3603–3617, Figure 3, right panel, with permission from the Company of Biologists Ltd.

(also referred to as Kuzbanian, KUZ, see below) (Hartmann *et al.*, 2002; Rooke *et al.*, 1996; Sotillos *et al.*, 1997). The S2 cleavage event is closely followed by intramembrane processing through the protease presenilin/ γ -secretase (S3 cleavage, Figure 8), which releases the NICD from the membrane and allows it to enter the nucleus (De Strooper *et al.*, 1999; Kopan and Ilagan, 2009; Struhl and Greenwald, 1999).

ADAM10 was first identified as a major regulator of Notch signaling in *Drosophila* through the identification of a mutant in ADAM10 called Kuzbanian (Rooke *et al.*, 1996). A major function of Notch is to regulate cell fate decisions, and in the developing eye of fruit flies, it ensures that there is only one sensory bristle per ommatidium (a component of the compound eye of fruit flies and other insects). In the Kuzbanian (ADAM10) mutant, more than one sensory bristles developed (Rooke *et al.*, 1996), giving rise to abnormal cluster of sensory bristles that resembled the spiky hair of a character on the Muppets Show called Kuzbanian, hence the name (the Muppets Show was a television puppet show in the 1970s). Later the general relevance of this discovery in fruit flies for mammalian development was confirmed when knockout mice for

ADAM10 were found to resemble knockout mice for Notch (early embryonic lethality with characteristic defects in somitogenesis) (Hartmann *et al.*, 2002). Several additional studies using conditional knockout mice lacking ADAM10 in specific tissues also predominantly uncovered defects that resembled those seen upon inactivation of Notch in these tissues (Glomski *et al.*, 2011; Jorissen *et al.*, 2010; Weber *et al.*, 2011). Taken together, genetic studies in mice, fruit flies and worms strongly suggest that the principal purpose of ADAM10 is to regulate ligand-induced canonical Notch signaling (Hartmann *et al.*, 2002; Rooke *et al.*, 1996; Sotillos *et al.*, 1997; Wen *et al.*, 1997).

The processing of Notch is tightly regulated because the membrane proximal cleavage site for ADAM10 is protected and covered in the full length Notch receptor. This cleavage site is only exposed when a Notch ligand on an adjacent cell binds to Notch and thereby initiates endocytosis of Notch and its ligand. This process essentially provides the force to 'pull' a protective cap away from the cleavage site (Kovall and Blacklow, 2010; Tiyanont *et al.*, 2011; Figure 9). Once the cleavage site is exposed, it is rapidly processed by ADAM10 (Figures 9 and 10), allowing the presenilin complex to bind to the remaining stub and cleave within the transmembrane domain of Notch. Because presenilin is thought to be unable to process the intact Notch molecule, ligand binding to Notch followed by exposure and processing of the site for ADAM10 controls Notch signaling (Kopan and Ilagan, 2009; Kovall and Blacklow, 2010). Consequently, Notch signaling can be blocked by adding a soluble ligand that cannot exert an endocytic pull, or by inactivating ADAM10 (Hartmann *et al.*, 2002; Rooke *et al.*, 1996; Sotillos *et al.*, 1997) or presenilins (De Strooper *et al.*, 1999; Kopan and Ilagan, 2009). A diagram of the canonical Notch signaling pathway is shown in Figure 11.

In this context it is important to note that ADAM17 can also function as a Notch processing enzyme, but only under non-physiological conditions such as treatment of cells with the Ca⁺⁺-chelating agent EDTA (Brou *et al.*, 2000; Bozkulak and Weinmaster, 2009; van Tetering *et al.*, 2009). The negative regulatory repeats (NRR) covering the cleavage site for ADAM10 contain Ca⁺⁺ ions, and once these are removed by EDTA, the NRR unfolds, thereby exposing the cleavage site to ADAM17 (Kovall and Blacklow, 2010; Tiyanont *et al.*, 2011). In addition, ADAM17 can also contribute to processing of mutant forms of Notch carrying disruptions in the LNR repeats (Sulis *et al.*, 2011). However, genetic studies with flies and in mice support the view that ADAM10 is the major physiologically relevant Notch processing enzyme because inactivation of ADAM10 causes defects that resemble those seen upon inactivation of Notch, whereas inactivation of ADAM17 does not cause major defects in Notch signaling, but instead leads to inactivation of EGFR signaling and of the functions of soluble TNF α .

Other Substrates of ADAM10 and ADAM17

In addition to the main substrates of ADAM10 and ADAM17 discussed above (TNF α , EGFR-ligands, and Notch), both enzymes can also cleave a variety of other membrane proteins on the cell surface. Indeed, it is likely that most of the

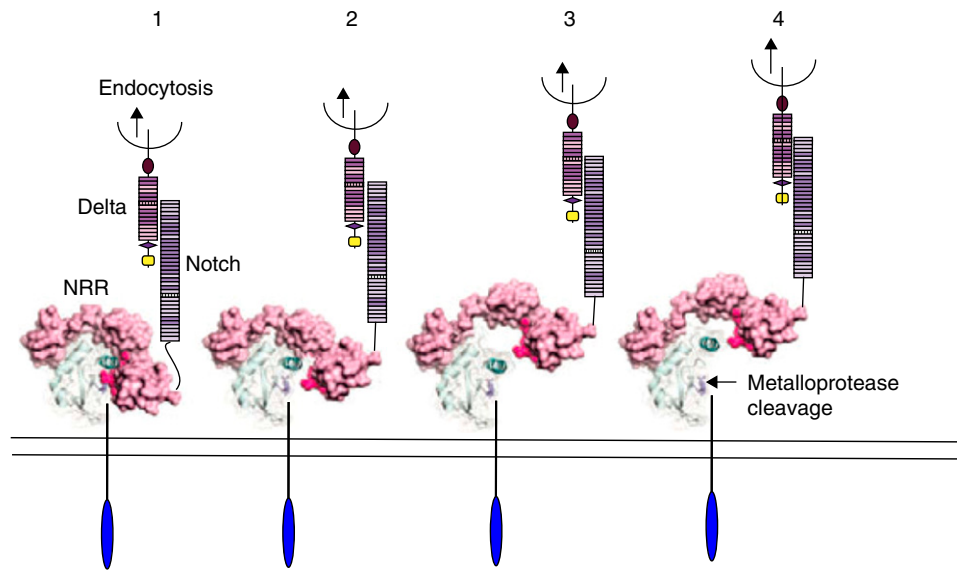


Figure 9 Diagram of the ligand-dependent activation of Notch processing by exposure of the cleavage site for ADAM10. (1) Binding of a Notch ligand such as Delta triggers endocytosis of Delta, which in turn exerts a pull on Notch. The membrane proximal negative regulatory domains (NRR, pink) are shown covering the cleavage site for ADAM10. (2–4) As endocytosis of Notch and its ligand proceed, the resulting force pulls away the NRR, exposing the cleavage site for ADAM10, and thereby allowing the S2 processing to occur. S2 processing is rapidly followed by S3 processing by the γ -secretase, which allows the NICD to enter the nucleus (see [Figure 8](#)). Modified from Blacklow, S.C., Gordon, W.R., Arnett, K.L., 2008. The molecular logic of Notch signaling—a structural and biochemical perspective. *Journal of Cell Science*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18799787>, Figure 2(C).

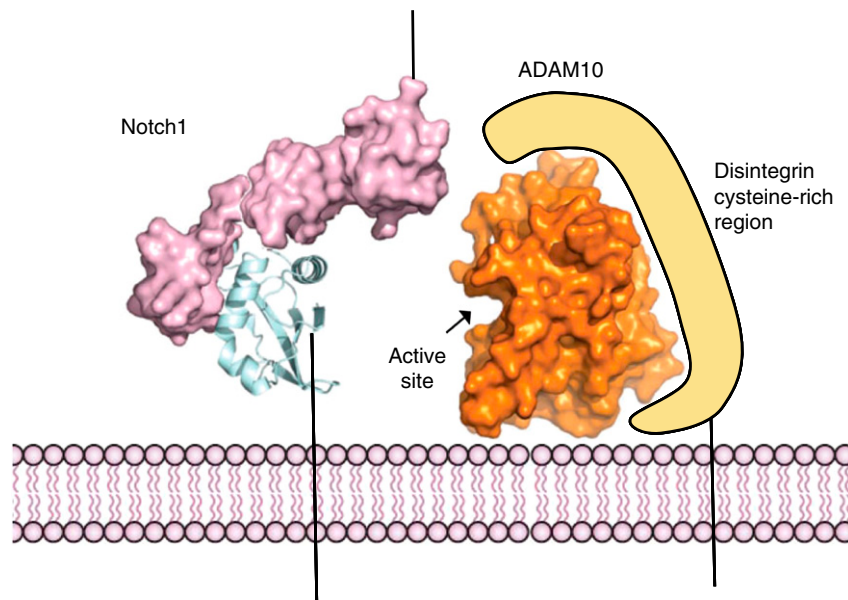


Figure 10 Processing of Notch by ADAM10 once the NRR has been pulled away. A more detailed structural view of the diagram in [Figure 9](#) panel 4 shows the S2 cleavage site in Notch exposed to the active site of ADAM10. The beige domain surrounding the metalloprotease domain of ADAM10 represents its disintegrin domain and cysteine-rich region. The approximate position of the membrane and the membrane anchors for ADAM10 and Notch1 are shown. Modified from Blacklow, S.C., Gordon, W.R., Arnett, K.L., 2008. The molecular logic of Notch signaling – A structural and biochemical perspective. *Journal of Cell Science*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18799787>, Figure 2(D).

membrane-anchored molecules on cells can be shed by ADAM17, and in some cases by ADAM10, although the functional relevance of the majority of these processing events remains to be established. There are several examples of

membrane proteins shed by ADAM10 or ADAM17 that are not essential for development, but that make important contributions to human disease, including the low affinity IgE receptor CD23 ([Weskamp *et al.*, 2006](#)), which has important

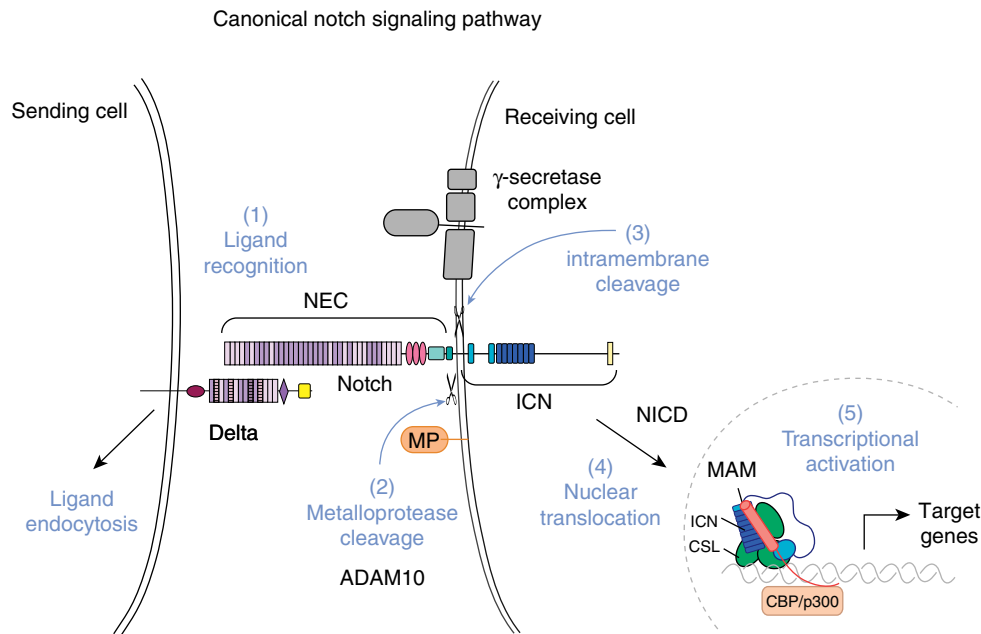


Figure 11 The Canonical Notch signaling pathway. This diagrammatic representation of the Notch signaling pathway shows the Notch-ligand Delta attached to the 'signal sending cell' on the left. (1) Delta binds to Notch through the Notch extracellular Domain (NEC), triggering ligand and receptor endocytosis (see also Figure 9). (2) This exposes the S2 cleavage site for ADAM10, which in turn triggers intramembrane cleavage of Notch by the γ -secretase complex (3). This releases the Notch intracellular domain (NICD) from the membrane, allowing it to translocate to the nucleus (4) and activate Notch-dependent transcription (5). Modified from Blacklow, S.C., Gordon, W.R., Arnett, K.L., 2008. The molecular logic of Notch signaling—a structural and biochemical perspective. *Journal of Cell Science*. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18799787>, Figure 1.

roles in asthma and allergic diseases, or the neutrophil cell adhesion molecule CD62L (L-selectin) (Li *et al.*, 2006), which is important for directing neutrophils to sites of inflammation. In addition, ADAM10 and potentially also ADAM17 are thought to have protective roles in Alzheimer's disease. A characteristic histopathological feature of brain sections from Alzheimer's disease patients is the presence of amyloid plaques (Selkoe, 1996; Figure 12). These contain large aggregates of a peptide that is called the amyloid beta peptide, which in turn is generated through proteolytic processing of a larger cell surface molecule that is called the amyloid precursor protein (Figures 13(a)–(d)). Similar to Notch, APP undergoes sequential cleavages, the first by an α -secretase (ADAM10 or ADAM17) or β -secretase (BACE), and the second by presenilin/ γ -secretase (Selkoe and Kopan, 2003). Processing of APP by the β -secretase and γ -secretase produces the A β peptide, which has a strong propensity to aggregate (Figures 13(b) and (c)), whereas processing of APP by α -secretase, i.e., ADAM10 (or ADAM17), gives rise to a shorter peptide that does not aggregate, and is therefore not amyloidogenic (Figure 13(d)). The α -secretase-dependent processing of APP preempts processing by beta secretase and is therefore considered to be protective. Attempts to block γ -secretase as a treatment for Alzheimer's disease have not been successful, in part because of toxicity of the inhibitory compounds, which also blocked Notch signaling. However, BACE remains a potential target for treatment of Alzheimer's disease (Vassar *et al.*, 1999), as does the possibility of selective activation of ADAM10 or ADAM17 in the brain (Fahrenholz and Postina, 2006), although much

remains to be learned about possible side effects associated with these potential treatments.

Summary

Taken together, ADAMs have emerged as crucial regulators of major signaling pathways with important roles in development and disease. All ligands of the EGFR must be cleaved in order to activate this receptor, which is important for maintenance of the skin and intestinal barrier, but can also promote cancer when it is inappropriately activated. ADAM17 together with iRhom1 and 2 control EGFR signaling in mammals and are thought to be able to allow rapid activation of this pathway if necessary, for example in case of an injury to the skin. In addition, the soluble form of the pro-inflammatory cytokine TNF α is generated from a membrane-anchored precursor by ADAM17 together with iRhom2. Moreover, ADAM10 is critical for signaling through the Notch receptor pathway. Notch receptors are membrane-anchored transcription factors that are processed by ADAM10 and then presenilin/ γ -secretase to release the Notch intracellular domain (NICD). This allows the NICD to enter the nucleus and promote Notch-dependent transcription, which is typically involved in cell fate decisions that control the proper development of an organism. Finally, ADAM10 and possibly also ADAM17 are thought to have an important protective role in Alzheimer's disease because they preempt the production of the A β peptide and thus help prevent the generation of amyloid plaques.

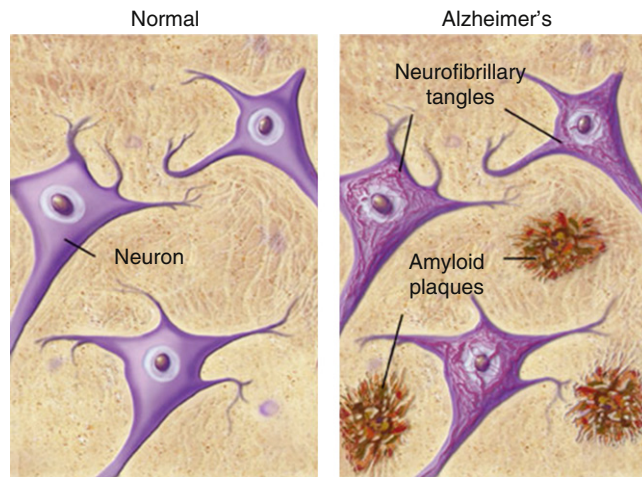


Figure 12 Diagrammatic representation of Amyloid plaques in the brain of an Alzheimer's disease patient compared to a normal brain without these plaques. <http://www.brightfocus.org/alzheimers/about/understanding/plaques-and-tangles.html>.

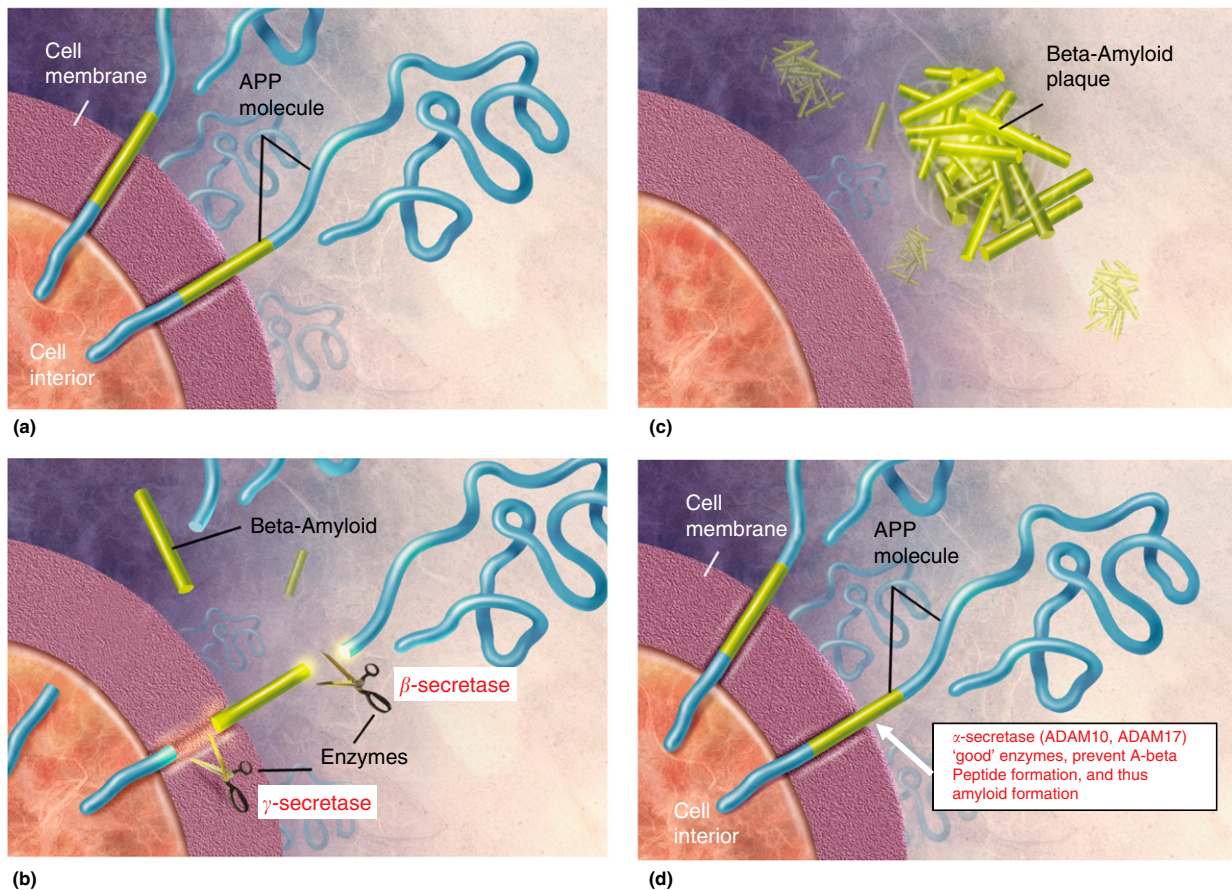


Figure 13 Diagram of the Amyloid Precursor Protein (APP) integrated into the cell membrane. (a) The APP molecule contains a large extracellular domain, a transmembrane domain and a short cytoplasmic domain (http://en.wikipedia.org/wiki/Alzheimer%27s_disease#mediaviewer/File: Amyloid_01big1.jpg). Panel (b) shows the processing of APP by β -secretase and γ -secretase, which releases the $\text{A}\beta$ peptide (highlighted in yellow) (http://en.wikipedia.org/wiki/Alzheimer%27s_disease#mediaviewer/File: Amyloid_02big1.jpg). (c) The $\text{A}\beta$ peptide is the main component of amyloid plaques (http://en.wikipedia.org/wiki/Alzheimer%27s_disease#mediaviewer/File: Amyloid_03big1.jpg). Panel (d) indicates the site of processing of APP by α -secretases (ADAM10 or ADAM17), which prevent processing by β -secretase, and are therefore considered to be protective in the context of Alzheimer's disease (http://en.wikipedia.org/wiki/Alzheimer%27s_disease#mediaviewer/File: Amyloid_01big1.jpg).

See also: Cell Communication: Growth Factor Mediated Cell Signaling: Receptor Tyrosine Kinases and Their Ligands; Signaling and Function of Death Receptors of the Tumor Necrosis Factor Receptor Superfamily; The Notch Pathway; Tumor Necrosis Factor Receptors: A Brief Digestion

References

- Adrain, C., Freeman, M., 2012. New lives for old: Evolution of pseudoenzyme function illustrated by iRhoms. *Nature Reviews Molecular Cell Biology* 13, 489–498.
- Adrain, C., Zettl, M., Christova, Y., Taylor, N., Freeman, M., 2012. Tumor necrosis factor signaling requires iRhom2 to promote trafficking and activation of TACE. *Science* 335, 225–228.
- Artavanis-Tsakonas, S., Matsuno, K., Fortini, M.E., 1995. Notch signaling. *Science* 268, 225–232.
- Black, R., Rauch, C.T., Kozlosky, C.J., *et al.*, 1997. A metalloprotease disintegrin that releases tumour-necrosis factor- α from cells. *Nature* 385, 729–733.
- Blaydon, D.C., Biancheri, P., Di, W.L., *et al.*, 2011. Inflammatory skin and bowel disease linked to ADAM17 deletion. *New England Journal of Medicine* 365, 1502–1508.
- Blobel, C.P., 2005. ADAMs: Key players in EGFR-signaling, development and disease. *Nature Reviews Molecular Cell Biology* 6, 32–43.
- Blobel, C.P., Carpenter, G., Freeman, M., 2009. The role of protease activity in ErbB biology. *Experimental Cell Research*, 315, 671–682.
- Bozkulak, E.C., Weinmaster, G., 2009. Selective use of ADAM10 and ADAM17 in activation of Notch1 signaling. *Molecular and Cellular Biology* 29, 5679–5695.
- Brou, C., Logeat, F., Gupta, N., *et al.*, 2000. A novel proteolytic cleavage involved in Notch signaling: The role of the disintegrin-metalloprotease TACE. *Molecular Cell* 5, 207–216.
- Chalaris, A., Adam, N., Sina, C., *et al.*, 2010. Critical role of the disintegrin metalloprotease ADAM17 for intestinal inflammation and regeneration in mice. *Journal of Experimental Medicine* 207, 1617–1624.
- De Strooper, B., Annaert, W., Cupers, P., *et al.*, 1999. A presenilin-1-dependent gamma-secretase-like protease mediates release of Notch intracellular domain. *Nature* 398, 518–522.
- Fahrenholz, F., Postina, R., 2006. Alpha-secretase activation – An approach to Alzheimer's disease therapy. *Neurodegenerative Diseases* 3, 255–261.
- Feldmann, M., Maini, S.R., 2008. Role of cytokines in rheumatoid arthritis: An education in pathophysiology and therapeutics. *Immunological Reviews* 223, 7–19.
- Fischer, O.M., Hart, S., Gschwind, A., Ullrich, A., 2003. EGFR signal transactivation in cancer cells. *Biochemical Society Transactions* 31, 1203–1208.
- Franzke, C.W., Cobzaru, C., Triantafyllou, A., *et al.*, 2012. Epidermal ADAM17 maintains the skin barrier by regulating EGFR ligand-dependent terminal keratinocyte differentiation. *Journal of Experimental Medicine* 209, 1105–1119.
- Glomski, K., Monette, S., Manova, K., *et al.*, 2011. Deletion of Adam10 in endothelial cells leads to defects in organ-specific vascular structures. *Blood* 118, 1163–1174.
- Hall, K.C., Hill, D., Otero, M., *et al.*, 2013. ADAM17 controls endochondral ossification by regulating terminal differentiation of chondrocytes. *Molecular and Cellular Biology* 33, 3077–3090.
- Harris, R.C., Chung, E., Coffey, R.J., 2003. EGF receptor ligands. *Experimental Cell Research* 284, 2–13.
- Hartmann, D., De Strooper, B., Serneels, L., *et al.*, 2002. The disintegrin/metalloprotease ADAM 10 is essential for Notch signaling but not for alpha-secretase activity in fibroblasts. *Human Molecular Genetics* 11, 2615–2624.
- Horiuchi, K., Kimura, T., Miyamoto, T., *et al.*, 2007. Cutting edge: TNF- α -converting enzyme (TACE/ADAM17) inactivation in mouse myeloid cells prevents lethality from endotoxin shock. *Journal of Immunology* 179, 2686–2689.
- Issuree, P.D., Maretzky, T., McIlwain, D.R., *et al.*, 2013. iRHOM2 is a critical pathogenic mediator of inflammatory arthritis. *Journal of Clinical Investigation* 123, 928–932.
- Jackson, L.F., Qiu, T.H., Sunnarborg, S.W., *et al.*, 2003. Defective valvulogenesis in HB-EGF and TACE-null mice is associated with aberrant BMP signaling. *EMBO Journal* 22, 2704–2716.
- Jorissen, E., Prox, J., Bernreuther, C., *et al.*, 2010. The disintegrin/metalloprotease ADAM10 is essential for the establishment of the brain cortex. *Journal of Neuroscience* 30, 4833–4844.
- Kopan, R., Ilagan, M.X., 2009. The canonical Notch signaling pathway: Unfolding the activation mechanism. *Cell* 137, 216–233.
- Kovall, R.A., Blacklow, S.C., 2010. Mechanistic insights into Notch receptor signaling from structural and biochemical studies. *Current Topics in Developmental Biology* 92, 31–71.
- Kurohara, K., Komatsu, K., Kurisaki, T., *et al.*, 2004. Essential roles of Meltrin beta (ADAM19) in heart development. *Developmental Biology* 267, 14–28.
- Lee, J.R., Urban, S., Garvey, C.F., Freeman, M., 2001. Regulated intracellular ligand transport and proteolysis control EGF signal activation in *Drosophila*. *Cell* 107, 161–171.
- Li, Y., Brazzell, J., Herrera, A., Walcheck, B., 2006. ADAM17 deficiency by mature neutrophils has differential effects on L-selectin shedding. *Blood* 108, 2275–2279.
- Lichtenberger, B.M., Gerber, P.A., Holcman, M., *et al.*, 2013. Epidermal EGFR controls cutaneous host defense and prevents inflammation. *Science Translational Medicine* 5, 199ra111.
- Maretzky, T., Evers, A., Zhou, W., *et al.*, 2011. Migration of growth factor-stimulated epithelial and endothelial cells depends on EGFR transactivation by ADAM17. *Nature Communications* 2, 229.
- Maretzky, T., McIlwain, D.R., Issuree, P.D., *et al.*, 2013. iRhom2 controls the substrate selectivity of stimulated ADAM17-dependent ectodomain shedding. *Proceedings of the National Academy of Sciences of the United States of America* 110, 11433–11438.
- McIlwain, D.R., Lang, P.A., Maretzky, T., *et al.*, 2012. iRhom2 regulation of TACE controls TNF-mediated protection against *Listeria* and responses to LPS. *Science* 335, 229–232.
- Modjtahedi, H., Essapen, S., 2009. Epidermal growth factor receptor inhibitors in cancer treatment: Advances, challenges and opportunities. *Anticancer Drugs* 20, 851–855.
- Moss, M.L., Jin, S.-L.C., Milla, M.E., *et al.*, 1997. Cloning of a disintegrin metalloprotease that processes precursor tumour-necrosis factor- α . *Nature* 385, 733–736.
- Najj, A.J., Day, K.C., Day, M.L., 2008. ADAM15 supports prostate cancer metastasis by modulating tumor cell-endothelial cell interaction. *Cancer Research* 68, 1092–1099.
- Overall, C.M., Blobel, C.P., 2007. In search of partners: linking extracellular proteases to substrates. *Nature Reviews Molecular Cell Biology* 8, 245–257.
- Peduto, L., Reuter, V.E., Sehara-Fujisawa, A., *et al.*, 2006. ADAM12 is highly expressed in carcinoma-associated stroma and is required for mouse prostate tumor progression. *Oncogene* 25, 5462–5466.
- Peduto, L., Reuter, V.E., Shaffer, D.R., Scher, H.I., Blobel, C.P., 2005. Critical function for ADAM9 in mouse prostate cancer. *Cancer Research* 65, 9312–9319.
- Peschon, J.J., Slack, J.L., Reddy, P., *et al.*, 1998. An essential role for ectodomain shedding in mammalian development. *Science* 282, 1281–1284.
- Rooke, J., Pan, D., Xu, T., Rubin, G.M., 1996. KUZ, a conserved metalloprotease-disintegrin protein with two roles in *Drosophila* neurogenesis. *Science* 273, 1227–1230.
- Sahin, U., Weskamp, G., Zhou, H.M., *et al.*, 2004. Distinct roles for ADAM10 and ADAM17 in ectodomain shedding of six EGFR-ligands. *Journal of Cell Biology* 164, 769–779.
- Selkoe, D.J., 1996. Amyloid beta-protein and the genetics of Alzheimer's disease. *Journal of Biological Chemistry* 271, 18295–18298.
- Selkoe, D., Kopan, R., 2003. Notch and Presenilin: regulated intramembrane proteolysis links development and degeneration. *Annual Review of Neuroscience* 26, 565–597.
- Sibilia, M., Wagner, B., Hoeberz, A., *et al.*, 2003. Mice humanised for the EGF receptor display hypomorphic phenotypes in skin, bone and heart. *Development* 130, 4515–4525.
- Sibilia, M., Wagner, E.F., 1995. Strain-dependent epithelial defects in mice lacking the EGF receptor [published erratum appears in *Science* 1995 ; 269(5226), 909]. *Science* 269, 234–238.
- Sotillos, S., Roch, F., Campuzano, S., 1997. The metalloprotease-disintegrin Kuzbanian participates in Notch activation during growth and patterning of *Drosophila* imaginal discs. *Development* 124, 4769–4779.
- Sternlicht, M.D., Sunnarborg, S.W., Kourou-Mehr, H., *et al.*, 2005. Mammary ductal morphogenesis requires paracrine activation of stromal EGFR via ADAM17-dependent shedding of epithelial amphiregulin. *Development* 132, 3923–3933.
- Struhl, G., Greenwald, I., 1999. Presenilin is required for activity and nuclear access of Notch in *Drosophila*. *Nature* 398, 522–525.
- Sulis, M.L., Saftig, P., Ferrando, A.A., 2011. Redundancy and specificity of the metalloprotease system mediating oncogenic NOTCH1 activation in T-ALL. *Leukemia* 25, 1564–1569.

- Sunnarborg, S.W., Hinkle, C.L., Stevenson, M., *et al.*, 2002. Tumor necrosis factor- α converting enzyme (TACE) regulates epidermal growth factor receptor ligand availability. *Journal of Biological Chemistry* 277, 12838–12845.
- van Tetering, G., Van Diest, P., Verlaan, I., *et al.*, 2009. Metalloprotease ADAM10 is required for Notch1 site 2 cleavage. *Journal of Biological Chemistry* 284, 31018–31027.
- Tiyanont, K., Wales, T.E., Aste-Amezaga, M., *et al.*, 2011. Evidence for increased exposure of the Notch1 metalloprotease cleavage site upon conversion to an activated conformation. *Structure* 19, 546–554.
- Urban, S., 2002. A family of Rhomboid intramembrane proteases activates all *Drosophila* membrane-tethered EGF ligands. *EMBO Journal* 21, 4277–4286.
- Urban, S., Lee, J.R., Freeman, M., 2001. *Drosophila* rhomboid-1 defines a family of putative intramembrane serine proteases. *Cell* 107, 173–182.
- Vassar, R., Bennett, B.D., Babu-Khan, S., *et al.*, 1999. Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* 286, 735–741.
- Weber, S., Niessen, M.T., Prox, J., *et al.*, 2011. The disintegrin/metalloproteinase Adam10 is essential for epidermal integrity and Notch-mediated signaling. *Development* 138, 495–505.
- Wen, C., Metzstein, M.M., Greenwald, I., 1997. SUP-17, a *Caenorhabditis elegans* ADAM protein related to *Drosophila* KUZBANIAN, and its role in LIN-12/NOTCH signaling. *Development* 124, 4759–4767.
- Weskamp, G., Ford, J., Sturgill, J., *et al.*, 2006. ADAM10 is a principal 'shedase' of the low-affinity immunoglobulin E receptor CD23. *Nature Immunology* 7, 1393–1398.
- Wheeler, D.L., Dunn, E.F., Harari, P.M., 2010. Understanding resistance to EGFR inhibitors-impact on future treatment strategies. *Nature Reviews Clinical Oncology* 7, 493–507.
- Zhou, H.M., Weskamp, G., Chesneau, V., *et al.*, 2004. Essential role for ADAM19 in cardiovascular morphogenesis. *Molecular and Cellular Biology* 24, 96–104.

Extracellular: Plasma Membrane Proteases – Serine Proteases

TM Antalis and MS Buzza, University of Maryland School of Medicine, Baltimore, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Proteases constitute approximately 2% of the human proteome and are critically important for numerous biological processes such as blood coagulation, cell death, tissue morphogenesis, inflammation, and wound healing. Proteases are ubiquitously expressed and can be found secreted into the extracellular environment, anchored to the cell surface, in the cell cytosol or compartmentalized in cellular organelles such as lysosomes. Through cleavage of specific peptide bonds in proteins that are recognized as substrates, proteases mediate many cellular functions including protein degradation, enzymatic activation, and induction of cellular signaling (reviewed in Lopez-Otin and Bond, 2008). Serine proteases are one of the largest families of proteolytic enzymes, constituting over one third of all proteolytic enzymes, and are known to play critical roles in diverse biological functions including blood coagulation, digestion and tissue homeostasis. Serine proteases are defined by the classical histidine, aspartate, and serine amino acid residues which form their catalytic triad, that mediate the process of peptide hydrolysis. Peptide hydrolysis occurs when the nucleophilic serine residue in the enzyme's active site attacks the carbonyl moiety of the substrate peptide bond, forming an acyl intermediate, and proteolysis follows which also depends upon the histidine and aspartate residues of the enzyme (Hedstrom, 2002; Rawlings and Barrett, 2004). The S1A subfamily of serine proteases are the most widely studied group of serine proteases; the prototypical members being trypsin, chymotrypsin and thrombin, which are produced as soluble proteases that are secreted into the extracellular environment. In recent years, a unique sub-group of S1A serine proteases has been identified which are found to be directly anchored to the cell surface. This review focuses on the current knowledge and *in vivo* functions of this family of membrane-anchored serine proteases. Further information on this family of proteases can be found in the following comprehensive reviews: Antalis *et al.*, 2010; Bugge *et al.*, 2009; Hooper *et al.*, 2001; Netzel-Arnett *et al.*, 2003; Szabo and Bugge, 2011.

Membrane-Anchored Serine Proteases

To date, 20 human and 22 mouse membrane-anchored serine proteases have been identified, which are classified into 3 subclasses based on the manner in which they are anchored to the cell surface (Figure 1). These proteases are anchored to the plasma membrane either via a C-terminal glycosylphosphatidylinositol (GPI) linkage (GPI-anchored), a C-terminal transmembrane domain (Type I), or an N-terminal transmembrane domain (Type II). In addition to the extracellular serine protease domain which mediates catalytic activity, the type II transmembrane serine proteases (TTSPs), possess a stem region composed of various combinations of different

accessory domains localized adjacent to the cell surface and C-terminal to the serine protease domain (Figure 1; Antalis *et al.*, 2010; Bugge *et al.*, 2009; Netzel-Arnett *et al.*, 2003).

Catalytic Activity

The catalytic serine protease domains (SPD) are highly homologous, and are 225–230 amino acids in size. Being trypsin-like serine proteases, the membrane-anchored serine proteases all prefer to cleave peptide substrates after the basic amino acids arginine or lysine, and specificity is also influenced by the residues N- and C-terminal to the cleavage site. A comprehensive list of protein and peptide substrates cleaved by the membrane-anchored serine proteases *in vitro* is found in Antalis *et al.* (2010). Like other serine proteases, membrane-anchored serine proteases are synthesized as inactive pro-enzymes called zymogens. Proteolytic cleavage of the short N-terminal pro-peptide of the SPD is required to induce a conformationally active protease, with the cleaved pro-domain remaining attached to the protease domain by a disulfide linkage. This activation-inducing cleavage may be mediated by other serine proteases (membrane-anchored or soluble), that recognize the zymogen activation sequence of the protease as a substrate. An example of such a zymogen activation cascade is the activation of the GPI-anchored protease prostasin by the TTSP matriptase in skin (Netzel-Arnett *et al.*, 2006). Once activated, several membrane-anchored-serine proteases can also induce zymogen activation of serine proteases of the digestive, coagulation, and fibrinolytic systems (reviewed in Antalis *et al.*, 2010). In addition, several of the TTSPs are thought to be capable of auto-activation (Oberst *et al.*, 2003b; Velasco *et al.*, 2002). In these cases the amino acid sequence of zymogen activation site of the protease resembles the substrate specificity of the protease, and pro-domain cleavage is thought to occur due to a low level of catalytic activity of the protease zymogen.

Stem Region Accessory Domains

While the functions of the modular domains found in the stem regions of TTSPs are mostly uncharacterized, they are thought to contribute to the cell surface localization, substrate recognition or activation of the protease. Specific mutations in these domains have been described to occur in human diseases (Antalis *et al.*, 2010). The sea urchin sperm protein (SEA)-domain appears to be important for inducing protease activation in matriptase and matriptase-2, and undergoes a spontaneous conformational non-enzymatic cleavage event required for protease activity (Oberst *et al.*, 2003b; Ramsay *et al.*, 2009). In addition, the low density lipoprotein-receptor class A-like (LDLRA) domain of matriptase-2 appears important for cell surface expression (Silvestri *et al.*, 2009), and the frizzled and LDLRA domains of corin are important for macromolecular substrate recognition (Knappe *et al.*, 2004).

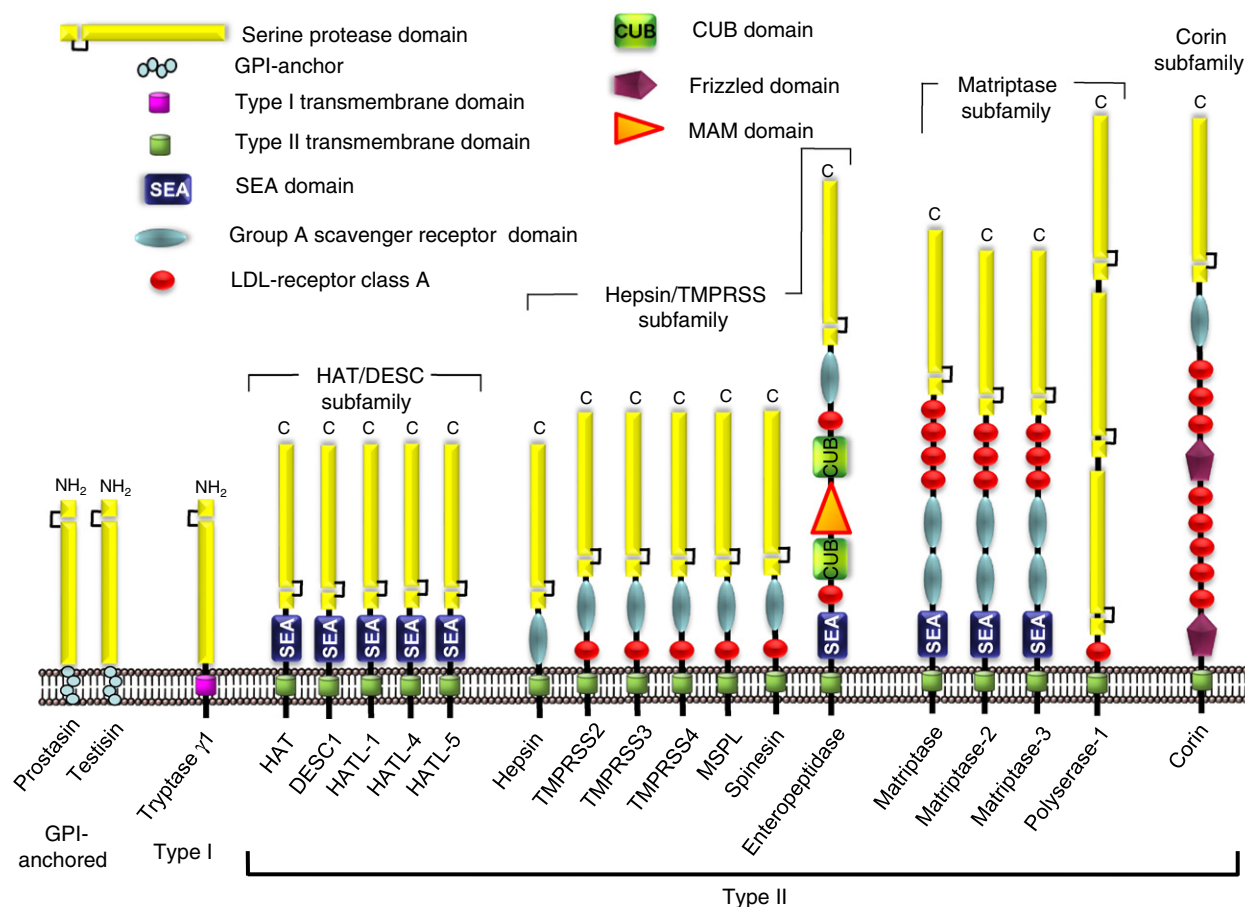


Figure 1 The membrane-anchored serine protease family. The human GPI-anchored serine proteases are prostatic and testisin. Trypsin $\gamma 1$ is the only known human type I transmembrane serine protease. The type II transmembrane serine proteases (TTSP) may be divided into four subfamilies: (1) the human airway trypsin-like protease/differentially expressed in squamous cell carcinoma (HAT/DESC) subfamily for which the stem regions are all composed of a single SEA-domain, (2) the hepsin/transmembrane protease, serine (TMPRSS) subfamily, each of which have a group A scavenger receptor domain (SRCR) in their stem region, that may be preceded by a single LDL receptor class A-like (LDLRA) domain or in enteropeptidase, an array of SEA, LDLRA, CUB, and MAM domains, (3) the matriptase subfamily, each containing a SEA-domain, two CUB domains, and 3-4 LDLRA domains in their stem region. Polyserase-1 comprises two active, and one catalytically inactive serine protease domains and a stem region containing an LDLRA domain, and (4) the corin subfamily, consisting of a single member, corin, which possesses a complex stem region composed of two frizzled domains, eight LDLRA domains, and one SRCR domain. NH₂ indicates amino terminus and C indicates carboxyl terminus. Adapted from Antalis, T.M., Buzza, M.S., Hodge, K.M., Hooper, J.D., Netzel-Arnett, S., 2010. The cutting edge membrane-anchored serine protease activities in the pericellular microenvironment. *Biochemical Journal* 428 (3), 325–346 and Bugge, T.H., Antalis, T.M., Wu, Q., 2009. Type II transmembrane serine proteases. *Journal of Biological Chemistry* 284 (35), 23177–23181.

Inhibition

After activation, the protease activity of membrane-anchored serine proteases can be regulated by interactions with endogenous protease inhibitors including membrane-anchored Kunitz-type inhibitors which form a reversible inhibitory complex with the protease. The Kunitz-type inhibitors HAI-1/SPINT1 and HAI-2/SPINT2 have been shown to regulate the activity of both matriptase and prostatic in *murine* models (Szabo *et al.*, 2007, 2009a,b, 2012). *In vitro*, the protease activities of several membrane-anchored serine proteases can also be inhibited by various members of the serpin superfamily, which form irreversible covalent complexes with serine proteases (reviewed in Antalis *et al.*, 2010). Inhibitory complexes between matriptase and anti-thrombin III (serpinC1), $\alpha 1$ -

proteinase inhibitor (serpinA1) and $\alpha 2$ -antiplasmin (serpinF2) have been detected in breast milk, suggesting an inhibitory role for these serpins *in vivo* (Tseng *et al.*, 2008).

While there is still much to learn about the physiological functions of individual members of this family, the development of murine models that are deficient in these proteases, and human diseases where expression or activity is reduced, have provided valuable evidence for the critical importance of many of these proteases in key biological functions as described below, and summarized in Tables 1 and 2.

The GPI-Anchored Serine Proteases

The two human GPI-anchored-membrane-anchored serine proteases, prostatic and testisin, are structurally the most

Table 1 Membrane-anchored serine proteases – physiological functions learnt through protease deficiency

Protease	Function	Implicated substrate	References
GPI-anchored Prostasin	Maintenance of epidermal barrier function	Unknown	(Frateschi <i>et al.</i> , 2012; Leyvraz <i>et al.</i> , 2005; Peters <i>et al.</i> , 2014)
	Regulation of ENaC-mediated fluid clearance in lung and colon epithelium	ENaC- γ subunit	(Frateschi <i>et al.</i> , 2012; Malsure <i>et al.</i> , 2014; Planes <i>et al.</i> , 2010)
	Regulation of hepatic insulin sensitivity by cleavage and inactivation of TLR-4	Toll-like receptor-4 (TLR-4)	(Uchimura <i>et al.</i> , 2014)
Testisin	Involved in sperm cell maturation and fertilizing ability	Unknown	(Aimes <i>et al.</i> , 2003; Inoue <i>et al.</i> , 1998; Netzel-Arnett <i>et al.</i> , 2009; Yamashita <i>et al.</i> , 2008)
Type II HAT/DESC subfamily			
HAT	Unknown. Not required for development, long-term health or survival in mice	Unknown	(Bertram <i>et al.</i> , 2012; Iwakiri <i>et al.</i> , 2004; Takahashi <i>et al.</i> , 2001; Yamaoka <i>et al.</i> , 1998)
HATL-1	Unknown. Not required for development, long-term health or survival in mice	Unknown	(Faller <i>et al.</i> , 2014; Kam <i>et al.</i> , 2009; Sales <i>et al.</i> , 2011)
Hepsin/TMPRSS subfamily			
Hepsin	Important for hearing in mice, where it has a role in cochlear development in inner ear	Unknown	(Faller <i>et al.</i> , 2014; Guipponi <i>et al.</i> , 2007; Kurachi <i>et al.</i> , 1994; Tsuji <i>et al.</i> , 1991)
	Maintenance of liver structural homeostasis in mice	Pro-HGF	(Hsu <i>et al.</i> , 2012)
TMPRSS2	Unknown. Not required for development, long-term health, fertility or survival in mice	Unknown	(Faller <i>et al.</i> , 2014; Jacquinet <i>et al.</i> , 2001; Kim <i>et al.</i> , 2006)
TMPRSS3	Mediates normal hearing in humans. Shown to be critical for cochlear hair cell survival in murine model of deficiency	Unknown	(Fasquelle <i>et al.</i> , 2011; Guipponi <i>et al.</i> , 2008; Lee <i>et al.</i> , 2013; Molina <i>et al.</i> , 2013)
TMPRSS5	Mediates normal hearing in humans	Unknown	(Guipponi <i>et al.</i> , 2008; Yamaguchi <i>et al.</i> , 2002)
Enteropeptidase	Critical for digestion of dietary proteins. Initiates a proteolytic cascade that results in the activation of several intestinal proteases	Trypsinogen	(Haworth <i>et al.</i> , 1971; Zheng <i>et al.</i> , 2009)
Matriptase subfamily			
Matriptase	Global role in maintenance of epidermal and epithelial barrier function and homeostasis.	Unknown	(List <i>et al.</i> , 2002, 2009; Yin <i>et al.</i> , 2014)
Matriptase-2	Essential regulator of iron homeostasis that prevents iron deficiency anemia	Hemojuvelin	(Du <i>et al.</i> , 2008; Folgueras <i>et al.</i> , 2008; Truksa <i>et al.</i> , 2009)
Corin subfamily			
Corin	Regulates systemic salt and water balance to prevent hypertension and cardiac hypertrophy	Pro-ANP, pro-BNP	(Chan <i>et al.</i> , 2005; Rame <i>et al.</i> , 2009; Wang <i>et al.</i> , 2008)

simple of this family, being composed solely of an N-terminal SPD, that is linked to the cell surface through a GPI moiety that is added to the C-terminus post-transcriptionally (Chen *et al.*, 2001; Hooper *et al.*, 1999; Figure 1). This lipid anchor is known to compartmentalize these proteases to specialized cholesterol-rich microdomains of the plasma membrane known as lipid rafts (Honda *et al.*, 2002; Verghese *et al.*, 2006). Prostasin, also known as PRSS8 and channel activating protease (CAP)-1, is found ubiquitously expressed in all epithelia (List *et al.*, 2007b). In polarized epithelia such as the gastrointestinal tract and kidney, it specifically localizes to the apical (luminal) membrane (Selzer-Plon *et al.*, 2009; Steensgaard *et al.*, 2010; Verghese *et al.*, 2006). Soluble forms of prostasin have been identified in both human seminal fluid and urine

(Koda *et al.*, 2009; Yu *et al.*, 1994), with release from the cell surface shown to be mediated by endogenous phospholipases or by proteolytic shedding (Iwashita *et al.*, 2003; Verghese *et al.*, 2006). Functionally, prostasin was the first identified membrane-anchored serine protease shown to enhance the activity of epithelial sodium channels (ENaC) by cleavage of the ENaC γ subunit to release an inhibitory peptide (Bruns *et al.*, 2007; Carattino *et al.*, 2008). ENaC activity is important for regulating sodium and water flux across polarized epithelium, and *in vivo*, increased expression or activity of prostasin, and the associated induction of ENaC activation, may be pathologically significant in increased fluid secretion in cystic fibrosis, high blood pressure and congenital sodium diarrhea (Table 2).

Table 2 Membrane-anchored serine proteases – roles in human disease

<i>Protease</i>	<i>Abnormality</i>	<i>Role in disease</i>	<i>References</i>
Prostasin	Over-expressed in lung epithelium of cystic fibrosis patients	May contribute to pathogenesis of cystic fibrosis by increasing fluid clearance	(Myerburg <i>et al.</i> , 2008; Planes <i>et al.</i> , 2010)
	Increased soluble prostasin detect in urine in hypertensive patients	May have a role in development of high blood pressure	(Maekawa <i>et al.</i> , 2009; Zhu <i>et al.</i> , 2008)
	Over-activity in colonic epithelium caused by loss of inhibitor function	Implicated role in fluid secretion in congenital sodium diarrhea	(Faller <i>et al.</i> , 2014)
Testisin	Aberrant expression in advanced stage ovarian cancer	May promote tumor growth and metastasis	(Shigemasa <i>et al.</i> , 2000; Tang <i>et al.</i> , 2005)
HAT	Lost in male germ cell tumors	Unknown	(Kempkensteffen <i>et al.</i> , 2006)
	Increased expression and shedding into airway fluids	Occurs in patients with inflammatory airway diseases such as asthma, function is unknown	(Yasuoka <i>et al.</i> , 1997)
DESC1	Lost in head and neck squamous cell carcinoma	Unknown	(Lang and Schuller, 2001; Sedghizadeh <i>et al.</i> , 2006)
HATL-5	Significantly decreased in cervical, esophageal, and head and neck carcinomas	Unknown	(Miller <i>et al.</i> , 2014)
Hepsin	Increased expression in human prostate cancers which correlates with disease severity	Increases prostate cancer progression and metastasis in mouse models	(Wu and Parry, 2007)
TMPRSS2	Frequent gene fusions between the promotor of TMPRSS2 and the ERG protooncogene and related transcription factors in prostate cancers	Androgen responsive elements in TMPRSS2 promotor drive expression of ERG transcription factor to promote prostate cancer progression	(Tomlins <i>et al.</i> , 2005; Yu <i>et al.</i> , 2010)
TMPRSS3	Point mutations that inhibit TMPRSS3 auto-activation blocking its activity	Causes non-syndromic autosomal recessive deafness	(Lee <i>et al.</i> , 2003; Scott <i>et al.</i> , 2001; Wattenhofer <i>et al.</i> , 2002)
TMPRSS4	Over-expressed in epithelial carcinomas of diverse origins	May have a role in tumor progression	(Choi <i>et al.</i> , 2008)
TMPRSS5	Point mutation that inactivates TMPRSS5	Associated with human deafness	(Guipponi <i>et al.</i> , 2008)
Enteropeptidase	Intestinal deficiency caused by point mutations	Failure to thrive due to reduced digestive function	(Holzinger <i>et al.</i> , 2002)
Matriptase	Mutations resulting in an inactive protease	ARIH, a rare human skin disease with ichthyosis and hair follicle defects	(Avrahami <i>et al.</i> , 2008; Basel-Vanagaite <i>et al.</i> , 2007; Desilets <i>et al.</i> , 2008; Lee <i>et al.</i> , 2007)
	Expression downregulated in inflammatory bowel diseases	May contribute to loss of intestinal barrier function and disease pathogenesis	(Kosa <i>et al.</i> , 2012; Netzel-Arnett <i>et al.</i> , 2012)
	Reduced expression in salivary gland epithelium	Causes loss of secretory cell function, contribute to pathogenesis of Sjogren's syndrome	(Yin <i>et al.</i> , 2014)
	Over-expressed in epithelial carcinomas of diverse origins	Possible role in tumor progression	(List, 2009)
Matriptase-2	Mutations that affect protease expression and activation	Causal factor in familial iron-refractory iron deficiency anemia	(Finberg <i>et al.</i> , 2008; Guillem <i>et al.</i> , 2008; Melis <i>et al.</i> , 2008)
Corin	Polymorphisms that cause reduced protease activity due to decreased zymogen activation	Associated with hypertension and cardiac hypertrophy, worse clinical outcome in patients with heart failure	(Dries <i>et al.</i> , 2005; Rame <i>et al.</i> , 2007, 2009; Wang <i>et al.</i> , 2008)
	Mutations and reduced expression and in pregnant uterus	May have causal role in the development of pre-eclampsia	(Cui <i>et al.</i> , 2012)

In mice, genetic deficiency of prostasin in the skin leads to complete loss of skin barrier function, which appears to be unrelated to defective ENaC activation (Frateschi *et al.*, 2012;

Leyvraz *et al.*, 2005; Peters *et al.*, 2014). The identification of prostasin's substrate in the epidermis remains uncertain, however, since genetic deficiency of the type II membrane-

anchored serine protease matriptase results in an identical epidermal defect (List *et al.*, 2002), it is thought that these proteases participate in a zymogen activation cascade that mediates epidermal barrier function. This is also supported by *in vitro* studies showing that expression of either protease is able to induce the activation of the other (Buzza *et al.*, 2013; Netzel-Arnett *et al.*, 2006). Interestingly, one study using murine models suggests that prostasin may mediate skin barrier function by a mechanism that is independent of its catalytic activity (Peters *et al.*, 2014). Using a murine model of prostasin deficiency in the liver, Uchimura *et al.* found that prostasin regulates hepatic insulin signaling by the cleavage and inactivation of the cell surface toll-like receptor, TLR-4, which suppresses inflammatory signaling (Uchimura *et al.*, 2014).

Testisin (also known as ESP-1 and PRSS21), in contrast to prostasin, exhibits an extremely specific and restricted tissue distribution, being abundantly expressed solely in male germ cells and spermatocytes, with lower expression also identified in microvascular endothelial cells, and eosinophils (Aimes *et al.*, 2003; Inoue *et al.*, 1998; Yamashita *et al.*, 2008). Functionally testisin is important for sperm cell maturation and fertilizing ability in mice (Netzel-Arnett *et al.*, 2009; Yamashita *et al.*, 2008), although the physiological substrate that mediates this activity is unknown. Pathologically, the aberrant testisin expression found in lung cancers and advanced ovarian cancers may contribute to tumour progression (Shigemasa *et al.*, 2000; Tang *et al.*, 2005), and the consequences of its loss in male germ cell tumors remains to be determined.

The Type I Transmembrane Serine Proteases

Tryptase $\gamma 1$ (also known as PRSS31, transmembrane tryptase and transmembrane protease γ -1) is found only in hematopoietic cells and is stored as the major component of mast cell secretory granules (Caughey *et al.*, 2000; Wong *et al.*, 1999, 2002b). Upon mast cell degranulation, tryptase $\gamma 1$ is retained on the cell surface by its N-terminal transmembrane domain (Wong *et al.*, 2002a). The physiological function and substrates of tryptase $\gamma 1$ remain unknown, but it is speculated that it may play a role in pathogen host defense in the human airway, where mast cells are important for protection against bacterial infections (Wong *et al.*, 2002a). In an animal model of exogenous administration to the airway, tryptase $\gamma 1$ promoted airway hyperresponsiveness and induced the expression of interleukin-13 (IL-13) in bronchiolar lavage fluid (Wong *et al.*, 2002a), which is a key cytokine implicated in the pathogenesis of allergic asthma.

The Type II Transmembrane Serine Proteases

TTSPs are by far the largest group of membrane-anchored serine proteases, with 17 members in humans and 19 members in mice (Bugge *et al.*, 2009; Szabo *et al.*, 2003; Figure 1). All are anchored to the cell surface by a C-terminal transmembrane domain, and have been phylogenetically divided into 4 subfamilies: (1) the human airway trypsin-like (HAT)/differentially expressed in squamous cell carcinoma gene

(DESC) subfamily, (2) the hepsin/transmembrane protease, serine (TMPRSS) subfamily, (3) the matriptase subfamily, and (4) the corin subfamily. In comparison to the GPI-anchored and type I serine proteases, which are composed of essentially a SPD and membrane anchor, the TTSPs possess a stem region C-terminal to the protease domain, which is comprised of a variety of modular structural accessory domains (Figure 1), most of which are currently of unknown function, but which are likely to play roles in protease activation, localization, and substrate recognition (Bugge *et al.*, 2009).

The HAT/DESC Subfamily

This subfamily is composed of 5 members in humans; HAT, DESC-1, HAT like-1 (HATL-1), HATL-4 and HATL-5, with mice having two additional members HATL-2 and HATL-3. This subfamily is the simplest of the TTSPs, having just one modular SEA-domain in their stem region. The role of this domain in HAT/DESC protease function has not been determined, but in other TTSPs such as matriptase, this domain undergoes a spontaneous conformation-induced auto-processing event at a conserved glycine residue which is important for protease activation (Macao *et al.*, 2006; Oberst *et al.*, 2003b). As its name implies, HAT protease (also known as TMPRSS11D) is found predominantly expressed in the epithelium of the airways, where it was originally identified as a soluble form found in extracellular lung fluids of asthma patients (Takahashi *et al.*, 2001; Yamaoka *et al.*, 1998; Yasuoka *et al.*, 1997). The normal physiological function of HAT is unknown, and deletion of the gene in mice caused no phenotypical defects in development or long-term health under non-challenged conditions (Sales *et al.*, 2011). Pathologically, HAT is up-regulated in chronic airway diseases (Yamaoka *et al.*, 1998), and has been shown *in vitro* to increase cell proliferation, and modulate inflammatory processes including increasing mucin production and suppression of fibrin deposition in airway epithelial cultures, suggesting it may play a role in disease suppression (Liu *et al.*, 2013; Matsushima *et al.*, 2006; Yoshinaga *et al.*, 1998). *In vitro*, HAT has been shown to cleave the urokinase plasminogen activator receptor (uPAR) (Beaufort *et al.*, 2007) and protease activated receptor (PAR)-2 (Iwakiri *et al.*, 2004), although the *in vivo* relevance of these substrates remains to be demonstrated. Interestingly, HAT may play a role in the propagation and spread of human respiratory viruses. HAT has been shown to cleave and activate the influenza virus hemagglutinin (HA) glycoprotein (Baron *et al.*, 2013; Bertram *et al.*, 2010), and the severe acute respiratory syndrome (SARS) coronavirus spike protein (Bertram *et al.*, 2011), both of which are important for mediating host cell entry, suggesting that inhibition of HAT activity may be a good target for therapeutic intervention in these infections.

Little is known regarding the *in vivo* function of the other members of this subfamily, which are expressed predominantly in epithelial cells of various organs. DESC-1 (also known as TMPRSS11E) is expressed in the epidermis, prostate, testis, placenta, thymus, and the epithelium of the head and neck (Lang and Schuller, 2001; Sales *et al.*, 2011; Sedghizadeh *et al.*, 2006), where it was originally identified based on its loss of expression in head and neck squamous cell carcinomas (Sales

et al., 2011). HATL-1 (also known as TMPRSS11A) is most highly expressed in the esophagus and trachea (Faller *et al.*, 2014; Kam *et al.*, 2009), and like HAT, the HATL-1 null mouse shows no overt phenotype (Sales *et al.*, 2011). Similarly, this protease may mediate the spread of SARS-corona virus through cleavage of the viral host cell entry spike glycoprotein (Kam *et al.*, 2009). HATL-4 (also known as TMPRSS11F) mRNA is expressed in the esophagus and testis, with lower expression in the cervix, placenta, and trachea (Sales *et al.*, 2011). HATL-5 (also known as TMPRSS11B) displays a relatively restricted tissue expression profile, being expressed in the cervix, esophagus, and oral cavity, with lower expression in the kidney and testis (Miller *et al.*, 2014; Sales *et al.*, 2011). This protease is found expressed in the more differentiated epithelial cells of cervix, esophagus, and trachea (Miller *et al.*, 2014), and is lost in squamous cell carcinomas of these tissues.

The Hepsin/TMPRSS Subfamily

This subfamily is composed of seven members in humans and mice. Hepsin possesses only one additional domain in its stem region, a group A scavenger receptor domain. TMPRSS2, TMPRSS3, TMPRSS4, mosaic serine protease large-form (MSPL), and spinesin have an additional low density lipoprotein receptor class A (LDLA) domain N-terminal to the scavenger domain, while enteropeptidase is much more complex containing multiple other domains types (Figure 1). Hepsin (also known as TMPRSS1), is abundantly expressed in the liver, but is also found at lower levels in other tissues such as the kidney, stomach, prostate, thyroid, and inner ear (Guipponi *et al.*, 2007; Kurachi *et al.*, 1994; Tsuji *et al.*, 1991). Physiological functions for hepsin have been identified in both liver and inner ear using murine models. While hepsin deficient mice do not show any strong defects in development or fertility (Wu *et al.*, 1998), mice lacking hepsin expression have abnormal cochlear structure and are deaf (Guipponi *et al.*, 2007). The physiological substrate cleaved by hepsin to mediate cochlear development has not been determined, but hepsin-deficient mice also show significantly reduced levels of the thyroid hormone thyroxine (Guipponi *et al.*, 2007) which is known to be important for cochlear development. Any relevance of hepsin expression to human hearing is yet to be reported. Recent analysis of liver specific hepsin null mice show its expression is important for maintenance of the structural integrity of the liver, with knock-out livers showing increased hepatocyte size and narrowed liver sinusoids (Hsu *et al.*, 2012). These authors demonstrated that the likely *in vivo* substrate cleaved by hepsin to regulate liver physiology is pro-hepatocyte growth factor (pro-HGF), a growth factor shown to be a substrate of hepsin *in vitro* (Kirchhofer *et al.*, 2005). Of importance, treatment of liver specific hepsin null mice with the active form of HGF was able rescue these phenotypic defects (Hsu *et al.*, 2012).

TMPRSS2 (also known as epitheliasin) is expressed in the epithelium of many tissues including the prostate, gastrointestinal tract, breast, lung, kidney, pancreas, ovary, lung, and salivary gland (Faller *et al.*, 2014; Jacquinet *et al.*, 2001). TMPRSS2 is of unknown function and is not required for

normal development or health in mice (Kim *et al.*, 2006). Similar to HAT, TMPRSS2 may augment virus entry into airway epithelium increasing virus propagation by cleaving the HA protein of influenza A virus (IAV) (Hatesuer *et al.*, 2013; Sakai *et al.*, 2014), and the SARS coronavirus spike protein (Heurich *et al.*, 2014). Of note, TMPRSS2 deficient mice are highly resistant to experimental IAV virus challenges (H1N1, H3N2, and H7N9 strains), which is associated with reduced HA cleavage *in vivo*, suggesting TMPRSS2 is a key host protease important for IAV infection and may represent a new therapeutic target in humans (Hatesuer *et al.*, 2013; Sakai *et al.*, 2014). TMPRSS3 (also known as TADG-12), is also expressed in epithelium of a variety of tissues where its function is unknown, however, its expression in the cochlea of the inner ear has been shown to be critically important for hearing in both humans and mice. At least five different point mutations in TMPRSS3 gene that affect protease activity have been linked to autosomal recessive deafness in humans (Guipponi *et al.*, 2008; Lee *et al.*, 2013; Scott *et al.*, 2001; Wattenhofer *et al.*, 2002). Furthermore, mice with TMPRSS3 deficiency are deaf, with reduced cochlear hair cell survival at the onset of hearing (Fasquelle *et al.*, 2011). The identification of the substrate cleaved by TMPRSS3 to mediate hearing is unknown, but expression of hair cell KCNMA1 potassium channels is reduced in mice with mutated TMPRSS3 (Molina *et al.*, 2013), suggesting a role in the acquisition of mature ion channels during cochlear development. Little is known about the biological function of TMPRSS4, also known as CAP-2 for its ability to activate ENaCs *in vitro* (Andreassen *et al.*, 2006; Passero *et al.*, 2012). However, it is over-expressed in a variety of epithelial carcinomas where it is correlated with disease progression and predicts poor patient survival (reviewed in Choi *et al.*, 2008). *In vitro* studies suggest TMPRSS4 mediates cancer cell invasion, epithelial-mesenchymal transition, and metastasis (Huang *et al.*, 2014; Kim *et al.*, 2010; Min *et al.*, 2014), where TMPRSS4 proteolytic activation of urokinase plasminogen activator (uPA) may be involved (Min *et al.*, 2014). TMPRSS5 (also known as spinesin), has an unusual expression pattern in comparison to other family members, being highly expressed in the brain and spinal cord (Yamaguchi *et al.*, 2002), where its function is unknown. TMPRSS5 is also expressed in the inner ear, and may be important for human hearing (Guipponi *et al.*, 2008). MSPL (also known as TMPRSS13) is expressed in the lung, placenta, prostate, and pancreas (Kim *et al.*, 2001) where its physiological function is unknown. Similar to HAT and TMPRSS2, it may play a role in influenza virus infection through cleavage of HA (Okumura *et al.*, 2010). Enteropeptidase (commonly known as enterokinase, also known as PRSS7) is expressed in upper small intestine (duodenum and jejunum) on the brush border of enterocytes of the intestinal villus (Hermon-Taylor *et al.*, 1977; Kitamoto *et al.*, 1994, 1995; Yuan *et al.*, 1998). Enteropeptidase has long been known to play a critical role in human digestion by activating a proteolytic cascade that results in the activation of numerous proteases important for human digestion (reviewed in Zheng *et al.*, 2009). Enteropeptidase cleaves and activates pancreatic trypsinogen to trypsin, which in turn activates other digestive zymogens such as chymotrypsinogen, proelastase, procarboxypeptidase, and prolipase in the lumen of the gut (Kunitz, 1939). Mutations that lead to truncated forms of the protease

that lack the active site are shown to cause malabsorption and failure to thrive in infants (Holzinger *et al.*, 2002).

The Matriptase Subfamily

This TTSP subfamily has four members – matriptase, matriptase-2, matriptase-3, and polyserase-1, which contain various combinations of accessory domains in their stem regions (Figure 1). Matriptase (also known as PRSS14, MT-SP1, CAP3, epithin, ST14) is the most well characterized of this subfamily and is found widely expressed in all epithelia (List *et al.*, 2006; Oberst *et al.*, 2003a). Matriptase has been shown to be critical for epithelial barrier function using mouse models of matriptase deficiency (List *et al.*, 2002, 2009). In humans, mutations affecting matriptase's enzymatic activity present with a rare form of skin disease (Basel-Vanagaite *et al.*, 2007; ARIH, see Table 2), which is characterized by scaly, itchy skin with increased permeability, similar to the phenotype of matriptase null mice (List *et al.*, 2002, 2007a). The substrate(s) cleaved by matriptase in the regulation of skin barrier function remain to be elucidated, although matriptase is thought to activate the pro-prostasin zymogen in this tissue (Netzel-Arnett *et al.*, 2006). In the polarized epithelium of the gastrointestinal tract where matriptase localizes to adherens junctions (Buzza *et al.*, 2010), matriptase is also critically important for regulating intestinal epithelial barrier function (Buzza *et al.*, 2010; List *et al.*, 2009), and mice deficient in intestinal matriptase develop chronic inflammation and spontaneous colitis-induced colon adenocarcinoma (Kosa *et al.*, 2012). In murine models, matriptase deficiency increases susceptibility to experimental colitis, and is found to be downregulated in human Crohn's disease and ulcerative colitis (Netzel-Arnett *et al.*, 2012). The matriptase substrate that regulates intestinal barrier function is unknown, but matriptase has been reported to induce an intracellular atypical protein kinase C signaling pathway that regulates the composition of the apical tight junctions (Buzza *et al.*, 2010). In the intestine and in contrast to the skin, prostasin appears to act upstream of matriptase and is important for inducing its activation in intestinal epithelium (Buzza *et al.*, 2013). There is also the possibility of a reciprocal zymogen activation cascade between these two proteases in different tissues, suggesting that both matriptase and prostasin zymogens may act as cofactors that can induce the activation of each other, in a mechanism that is independent of catalytic activity (Friis *et al.*, 2013). Matriptase is also lost in the salivary gland of patients with Sjogren's syndrome which exhibit salivary gland dysfunction and develop autoimmunity. Its loss may be important for disease pathogenesis, since matriptase deficiency in mice induces a primary Sjogren's syndrome-like phenotype with autoimmunity and loss of gland function (Yin *et al.*, 2014).

Matriptase also possesses strong oncogenic activities and in many epithelial cancers, its increased expression or activity (through loss of inhibition) is associated with disease progression (reviewed in List, 2009). Low level over-expression of matriptase in murine skin is sufficient to induce multistage carcinogenesis and invasive squamous cell carcinoma (List *et al.*, 2005). Based on murine models, the mechanism by which matriptase induces squamous cell carcinoma involves

cleavage and activation of the matriptase substrate pro-HGF (Lee *et al.*, 2000; Szabo *et al.*, 2011). *In vitro* studies also identified matriptase as the first membrane-anchored serine protease to activate the G-protein coupled receptor, PAR-2 (Camerer *et al.*, 2010; Takeuchi *et al.*, 2000). *In vivo* evidence suggests that cleavage of PAR-2, which induces inflammatory signaling pathways, may be an essential component for matriptase induced malignant progression of squamous cell carcinomas in animal models (Sales *et al.*, 2014). Matriptase activity in the skin also has been linked to the initiation of Netherton syndrome, which results from a loss of function mutation in the serine protease inhibitor LEKTI, which is important for regulating the activity of kallikrein proteases in the skin (Sales *et al.*, 2010). Genetic ablation of matriptase in mice deficient in the LEKTI inhibitor reduced inflammation, reduced aberrant protease activity and improved barrier function in the skin, suggesting matriptase is pathologically important in the human disease through kallikrein activation (Sales *et al.*, 2010).

Matriptase-2 (also known as TMPRSS6) is most highly expressed in the liver, but is also found in the kidney and uterine tissue (Hooper *et al.*, 2003; Velasco *et al.*, 2002). Matriptase-2 has a significant function in the regulation of iron metabolism. Humans with point mutations that affect matriptase-2 activity suffer from iron-refractory iron deficient anemia (Finberg *et al.*, 2008; Guillem *et al.*, 2008; Melis *et al.*, 2008), manifested by very low iron levels and severe microcytic anemia. Mice deficient in matriptase-2 expression also show this iron deficient phenotype (Du *et al.*, 2008; Truksa *et al.*, 2009). Matriptase-2 is thought to function by suppressing the expression of the liver hormone hepcidin which is known to be an important mediator of iron uptake by liver cells. This suppression occurs indirectly at the mRNA level, and is thought to be mediated by matriptase-2 cleavage of the bone morphogenetic protein (BMP) co-receptor hemojuvelin (Silvestri *et al.*, 2008, 2009; Truksa *et al.*, 2009), which has a role in activation of hepcidin transcription. Matriptase-3 (also known as TMPRSS7) is expressed in the brain, skin, salivary gland, and reproductive tissues (Szabo *et al.*, 2005); however, its physiological functions remain to be determined. Polyserase-1 (also known as TMPRSS9 and serase1B) is expressed in skeletal muscle, heart, kidney, liver, placenta, and brain (Cal *et al.*, 2003, 2007). This TTSP is highly unique in that it possesses 3 SPDs but its function is currently unknown.

The Corin Subfamily

Corin (also known as TMPRSS10) is the only member of this subfamily, and is unique in that it possesses two frizzled-like cysteine rich domains in its stem region, along with multiple LDLA domains and a scavenger domain (Figure 1). Corin is highly expressed in cardiomyocytes of the heart, but also in kidney, bone, brain, skin, and pregnant uterus (Hooper *et al.*, 2000; Yan *et al.*, 1999). Corin plays an important role in maintaining heart function and decreasing blood volume and blood pressure, by processing the cardiac hormone pro-atrial natriuretic peptide (pro-ANP) (Wu *et al.*, 2002; Yan *et al.*, 2000). Corin cleaves and activates pro-ANP to mature ANP which promotes natriuresis, diuresis, and vasodilation. In

humans, mutations in the corin gene which result in reduced activation and protease activity have been reported in patients with hypertension and cardiac hypertrophy, and are associated with a poorer prognosis in patients with heart failure (Dries *et al.*, 2005; Rame *et al.*, 2007, 2009; Wang *et al.*, 2008). Corin-deficient mice spontaneously develop salt-sensitive hypertension, cardiac hypertrophy and increased body weight, and possess an elevated level of pro-ANP and an absence of active ANP (Chan *et al.*, 2005; Nigrovic *et al.*, 2008). Corin and ANP have also been shown to be important for promoting placental trophoblast invasion and spiral artery remodeling in the pregnant uterus (Cui *et al.*, 2012). Pregnancy in corin- or ANP-deficient mice results in the development of high blood pressure and proteinuria, which are characteristics of pre-eclampsia. Interestingly, in human pre-eclamptic patients, uterine corin expression was found to be reduced, and several corin mutations which cause proteolytic activity were also identified in pre-eclamptic patients, suggesting defective corin may be pathologically important in this disease.

Conclusion

Following the sequencing of the human genome at the turn of the century and the discovery of the family of membrane-anchored serine proteases, these enzymes have emerged to play key roles in many diverse aspects of mammalian physiology. The misregulation of these enzymes is emerging to contribute to the pathology of a variety of diseases. The plasma membrane is a dynamic, fluid microenvironment where cell surface molecules act as cell sensors, initiating signals and relaying information. The extracellular SPDs contained in these molecules likely target specific cellular substrates, growth factors, receptors, and components of the extracellular matrix to modulate irreversible changes in the cellular environment. Clearly these unique enzymes represent potential strategic targets for diagnostic and therapeutic applications for a wide range of diseases.

See also: Protein Synthesis/Degradation: Protein Degradation – Protease Classes: ADAMs Regulate Cell–Cell Interactions by Controlling the Function of the EGF-Receptor, TNF α and Notch; ADAMTS Proteases: Mediators of Physiological and Pathogenic Extracellular Proteolysis; Matrix Metalloproteinases; Naturally-Occurring Polypeptide Inhibitors: Cystatins/Stefins, Inhibitors of Apoptosis (IAPs), Serpins, and Tissue Inhibitors of Metalloproteinases (TIMPs). Protein Synthesis/Degradation: Proteolytic Pathways: Molecular Mechanisms Underlying the Actions of the Complement System; Overview of Blood Coagulation and the Pathophysiology of Blood Coagulation Disorders

References

- Aimes, R.T., Zijlstra, A., Hooper, J.D., *et al.*, 2003. Endothelial cell serine proteases expressed during vascular morphogenesis and angiogenesis. *Thrombosis and Haemostasis* 89 (3), 561–572.
- Andreasen, D., Vuagniaux, G., Fowler-Jaeger, N., Hummler, E., Rossier, B.C., 2006. Activation of epithelial sodium channels by mouse channel activating proteases (mCAP) expressed in *Xenopus* oocytes requires catalytic activity of mCAP3 and mCAP2 but not mCAP1. *Journal of the American Society of Nephrology* 17 (4), 968–976.
- Antalis, T.M., Buzza, M.S., Hodge, K.M., Hooper, J.D., Netzel-Arnett, S., 2010. The cutting edge: Membrane-anchored serine protease activities in the pericellular microenvironment. *Biochemical Journal* 428 (3), 325–346.
- Avrahami, L., Maas, S., Pasmanik-Chor, M., *et al.*, 2008. Autosomal recessive ichthyosis with hypotrichosis syndrome: Further delineation of the phenotype. *Clinical Genetics* 74 (1), 47–53.
- Baron, J., Tarnow, C., Mayoli-Nussle, D., *et al.*, 2013. Matriptase, HAT, and TMPRSS2 activate the hemagglutinin of H9N2 influenza A viruses. *Journal of Virology* 87 (3), 1811–1820.
- Basel-Vanagaite, L., Attia, R., Ishida-Yamamoto, A., *et al.*, 2007. Autosomal recessive ichthyosis with hypotrichosis caused by a mutation in ST14, encoding type II transmembrane serine protease matriptase. *American Journal of Human Genetics* 80 (3), 467–477.
- Beaufort, N., Leduc, D., Eguchi, H., *et al.*, 2007. The human airway trypsin-like protease modulates the urokinase receptor (uPAR, CD87) structure and functions. *American Journal of Physiology – Lung Cellular and Molecular Physiology* 292 (5), L1263–L1272.
- Bertram, S., Glowacka, I., Muller, M.A., *et al.*, 2011. Cleavage and activation of the severe acute respiratory syndrome coronavirus spike protein by human airway trypsin-like protease. *Journal of Virology* 85 (24), 13363–13372.
- Bertram, S., Glowacka, I., Steffen, I., Kuhl, A., Pohlmann, S., 2010. Novel insights into proteolytic cleavage of influenza virus hemagglutinin. *Reviews in Medical Virology* 20 (5), 298–310.
- Bertram, S., Heurich, A., Lavender, H., *et al.*, 2012. Influenza and SARS-coronavirus activating proteases TMPRSS2 and HAT are expressed at multiple sites in human respiratory and gastrointestinal tracts. *PLoS One* 7 (4), e35876.
- Bruns, J.B., Carattino, M.D., Sheng, S., *et al.*, 2007. Epithelial Na⁺ channels are fully activated by furin- and prostaticin-dependent release of an inhibitory peptide from the gamma-subunit. *Journal of Biological Chemistry* 282 (9), 6153–6160.
- Bugge, T.H., Antalis, T.M., Wu, Q., 2009. Type II transmembrane serine proteases. *Journal of Biological Chemistry* 284 (35), 23177–23181.
- Buzza, M.S., Martin, E.W., Driesbaugh, K.H., *et al.*, 2013. Prostaticin is required for matriptase activation in intestinal epithelial cells to regulate closure of the paracellular pathway. *Journal of Biological Chemistry* 288 (15), 10328–10337.
- Buzza, M.S., Netzel-Arnett, S., Shea-Donohue, T., *et al.*, 2010. Membrane-anchored serine protease matriptase regulates epithelial barrier formation and permeability in the intestine. *Proceedings of the National Academy of Sciences of the United States of America* 107 (9), 4200–4205.
- Cal, S., Moncada-Pazos, A., Lopez-Otin, C., 2007. Expanding the complexity of the human degradome: Polyserases and their tandem serine protease domains. *Frontiers in Bioscience* 12, 4661–4669.
- Cal, S., Quesada, V., Garabaya, C., Lopez-Otin, C., 2003. Polyserase-I, a human polyprotease with the ability to generate independent serine protease domains from a single translation product. *Proceedings of the National Academy of Sciences of the United States of America* 100 (16), 9185–9190.
- Camerer, E., Barker, A., Duong, D.N., *et al.*, 2010. Local protease signaling contributes to neural tube closure in the mouse embryo. *Developmental Cell* 18 (1), 25–38.
- Carattino, M.D., Hughey, R.P., Kleyman, T.R., 2008. Proteolytic processing of the epithelial sodium channel gamma subunit has a dominant role in channel activation. *Journal of Biological Chemistry* 283 (37), 25290–25295.
- Caughey, G.H., Raymond, W.W., Blount, J.L., *et al.*, 2000. Characterization of human gamma-trypsinases, novel members of the chromosome 16p mast cell trypsinase and prostaticin gene families. *Journal of Immunology* 164 (12), 6566–6575.
- Chan, J.C., Knudson, O., Wu, F., *et al.*, 2005. Hypertension in mice lacking the proatrial natriuretic peptide convertase corin. *Proceedings of the National Academy of Sciences of the United States of America* 102 (3), 785–790.
- Chen, L.M., Skinner, M.L., Kauffman, S.W., *et al.*, 2001. Prostaticin is a glycosylphosphatidylinositol-anchored active serine protease. *Journal of Biological Chemistry* 276 (24), 21434–21442.
- Choi, S.Y., Shin, H.C., Kim, S.Y., Park, Y.W., 2008. Role of TMPRSS4 during cancer progression. *Drug News & Perspectives* 21 (8), 417–423.
- Cui, Y., Wang, W., Dong, N., *et al.*, 2012. Role of corin in trophoblast invasion and uterine spiral artery remodelling in pregnancy. *Nature* 484 (7393), 246–250.
- Desilets, A., Beliveau, F., Vandal, G., *et al.*, 2008. Mutation G827R in matriptase causing autosomal recessive ichthyosis with hypotrichosis yields an inactive protease. *Journal of Biological Chemistry* 283 (16), 10535–10542.

- Dries, D.L., Victor, R.G., Rame, J.E., *et al.*, 2005. Corin gene minor allele defined by 2 missense mutations is common in blacks and associated with high blood pressure and hypertension. *Circulation* 112 (16), 2403–2410.
- Du, X., She, E., Gelbart, T., *et al.*, 2008. The serine protease TMPRSS6 is required to sense iron deficiency. *Science* 320 (5879), 1088–1092.
- Faller, N., Gautschi, I., Schild, L., 2014. Functional analysis of a missense mutation in the serine protease inhibitor SPINT2 associated with congenital sodium diarrhea. *PLoS One* 9 (4), e94267.
- Fasquelle, L., Scott, H.S., Lenoir, M., *et al.*, 2011. Tmprss3, a transmembrane serine protease deficient in human DFNB8/10 deafness, is critical for cochlear hair cell survival at the onset of hearing. *Journal of Biological Chemistry* 286 (19), 17383–17397.
- Finberg, K.E., Heeney, M.M., Campagna, D.R., *et al.*, 2008. Mutations in TMPRSS6 cause iron-refractory iron deficiency anemia (IRIDA). *Nature Genetics* 40 (5), 569–571.
- Folgueras, A.R., de Lara, F.M., Pendas, A.M., *et al.*, 2008. Membrane-bound serine protease matriptase-2 (Tmprss6) is an essential regulator of iron homeostasis. *Blood* 112 (6), 2539–2545.
- Frateschi, S., Keppner, A., Malsure, S., *et al.*, 2012. Mutations of the serine protease CAP1/Prss8 lead to reduced embryonic viability, skin defects, and decreased ENaC activity. *American Journal of Pathology* 181 (2), 605–615.
- Friis, S., Uzzun, S.K., Godiksen, S., *et al.*, 2013. A matriptase-prostasin reciprocal zymogen activation complex with unique features: Prostasin as a non-enzymatic co-factor for matriptase activation. *Journal of Biological Chemistry* 288 (26), 19028–19039.
- Guillem, F., Lawson, S., Kannengiesser, C., *et al.*, 2008. Two nonsense mutations in the TMPRSS6 gene in a patient with microcytic anemia and iron deficiency. *Blood* 112 (5), 2089–2091.
- Guipponi, M., Antonarakis, S.E., Scott, H.S., 2008. TMPRSS3, a type II transmembrane serine protease mutated in non-syndromic autosomal recessive deafness. *Frontiers in Bioscience* 13, 1557–1567.
- Guipponi, M., Tan, J., Cannon, P.Z., *et al.*, 2007. Mice deficient for the type II transmembrane serine protease, TMPRSS1/hepsin, exhibit profound hearing loss. *American Journal of Pathology* 171 (2), 608–616.
- Guipponi, M., Toh, M.Y., Tan, J., *et al.*, 2008. An integrated genetic and functional analysis of the role of type II transmembrane serine proteases (TMPRSSs) in hearing loss. *Human Mutation* 29 (1), 130–141.
- Hatesuer, B., Bertram, S., Mehner, N., *et al.*, 2013. Tmprss2 is essential for influenza H1N1 virus pathogenesis in mice. *PLoS Pathogens* 9 (12), e1003774.
- Haworth, J.C., Gourley, B., Hadorn, B., Sumida, C., 1971. Malabsorption and growth failure due to intestinal enterokinase deficiency. *Journal of Pediatrics* 78 (3), 481–490.
- Hedstrom, L., 2002. Serine protease mechanism and specificity. *Chemical Reviews* 102 (12), 4501–4524.
- Hermion-Taylor, J., Perrin, J., Grant, D.A., *et al.*, 1977. Immunofluorescent localisation of enterokinase in human small intestine. *Gut* 18 (4), 259–265.
- Heurich, A., Hofmann-Winkler, H., Gierer, S., *et al.*, 2014. TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *Journal of Virology* 88 (2), 1293–1307.
- Holzinger, A., Maier, E.M., Buck, C., *et al.*, 2002. Mutations in the proenteropeptidase gene are the molecular cause of congenital enteropeptidase deficiency. *American Journal of Human Genetics* 70 (1), 20–25.
- Honda, A., Yamagata, K., Sugiura, S., Watanabe, K., Baba, T., 2002. A mouse serine protease TESP5 is selectively included into lipid rafts of sperm membrane presumably as a glycosylphosphatidylinositol-anchored protein. *Journal of Biological Chemistry* 277 (19), 16976–16984.
- Hooper, J.D., Campagnolo, L., Goodarzi, G., *et al.*, 2003. Mouse matriptase-2: Identification, characterization and comparative mRNA expression analysis with mouse hepsin in adult and embryonic tissues. *Biochemical Journal* 373 (no Pt 3), 689–702.
- Hooper, J.D., Clements, J.A., Quigley, J.P., Antalis, T.M., 2001. Type II transmembrane serine proteases. Insights into an emerging class of cell surface proteolytic enzymes. *Journal of Biological Chemistry* 276 (2), 857–860.
- Hooper, J.D., Nicol, D.L., Dickinson, J.L., *et al.*, 1999. Testisin, a new human serine proteinase expressed by premeiotic testicular germ cells and lost in testicular germ cell tumors. *Cancer Research* 59 (13), 3199–3205.
- Hooper, J.D., Scarman, A.L., Clarke, B.E., Normyle, J.F., Antalis, T.M., 2000. Localization of the mosaic transmembrane serine protease corin to heart myocytes. *European Journal of Biochemistry* 267 (23), 6931–6937.
- Hsu, Y.C., Huang, H.P., Yu, I.S., *et al.*, 2012. Serine protease hepsin regulates hepatocyte size and hemodynamic retention of tumor cells by hepatocyte growth factor signaling in mice. *Hepatology* 56 (5), 1913–1923.
- Huang, A., Zhou, H., Zhao, H., *et al.*, 2014. TMPRSS4 correlates with colorectal cancer pathological stage and regulates cell proliferation and self-renewal ability. *Cancer Biology & Therapy* 15 (3), 297–304.
- Inoue, M., Kanbe, N., Kurosawa, M., Kido, H., 1998. Cloning and tissue distribution of a novel serine protease esp-1 from human eosinophils. *Biochemical and Biophysical Research Communications* 252 (2), 307–312.
- Iwakiri, K., Ghazizadeh, M., Jin, E., *et al.*, 2004. Human airway trypsin-like protease induces PAR-2-mediated IL-8 release in psoriasis vulgaris. *Journal of Investigative Dermatology* 122 (4), 937–944.
- Iwashita, K., Kitamura, K., Narikyo, T., *et al.*, 2003. Inhibition of prostasin secretion by serine protease inhibitors in the kidney. *Journal of the American Society of Nephrology* 14 (1), 11–16.
- Jacquinet, E., Rao, N.V., Rao, G.V., *et al.*, 2001. Cloning and characterization of the cDNA and gene for human epitheliasin. *European Journal of Biochemistry* 268 (9), 2687–2699.
- Kam, Y.W., Okumura, Y., Kido, H., *et al.*, 2009. Cleavage of the SARS coronavirus spike glycoprotein by airway proteases enhances virus entry into human bronchial epithelial cells *in vitro*. *PLoS One* 4 (11), e7870.
- Kempkensteffen, C., Christoph, F., Weikert, S., *et al.*, 2006. Epigenetic silencing of the putative tumor suppressor gene testisin in testicular germ cell tumors. *Journal of Cancer Research and Clinical Oncology* 132 (12), 765–770.
- Kim, D.R., Sharmin, S., Inoue, M., Kido, H., 2001. Cloning and expression of novel mosaic serine proteases with and without a transmembrane domain from human lung. *Biochimica et Biophysica Acta* 1518 (1–2), 204–209.
- Kim, S., Kang, H.Y., Nam, E.H., *et al.*, 2010. TMPRSS4 induces invasion and epithelial-mesenchymal transition through upregulation of integrin alpha5 and its signaling pathways. *Carcinogenesis* 31 (4), 597–606.
- Kim, T.S., Heinlein, C., Hackman, R.C., Nelson, P.S., 2006. Phenotypic analysis of mice lacking the Tmprss2-encoded protease. *Molecular and Cellular Biology* 26 (3), 965–975.
- Kirchhofer, D., Peek, M., Lipari, M.T., *et al.*, 2005. Hepsin activates pro-hepatocyte growth factor and is inhibited by hepatocyte growth factor activator inhibitor-1B (HAI-1B) and HAI-2. *FEBS Letters* 579 (9), 1945–1950.
- Kitamoto, Y., Veile, R.A., Donis-Keller, H., Sadler, J.E., 1995. cDNA sequence and chromosomal localization of human enterokinase, the proteolytic activator of trypsinogen. *Biochemistry* 34 (14), 4562–4568.
- Kitamoto, Y., Yuan, X., Wu, Q., McCourt, D.W., Sadler, J.E., 1994. Enterokinase, the initiator of intestinal digestion, is a mosaic protease composed of a distinctive assortment of domains. *Proceedings of the National Academy of Sciences of the United States of America* 91 (16), 7588–7592.
- Knappe, S., Wu, F., Madlansacay, M.R., Wu, Q., 2004. Identification of domain structures in the propeptide of corin essential for the processing of proatrial natriuretic peptide. *Journal of Biological Chemistry* 279 (33), 34464–34471.
- Koda, A., Wakida, N., Toriyama, K., *et al.*, 2009. Urinary prostasin in humans: Relationships among prostasin, aldosterone and epithelial sodium channel activity. *Hypertension Research* 32 (4), 276–281.
- Kosa, P., Szabo, R., Molinolo, A.A., Bugge, T.H., 2012. Suppression of Tumorigenicity-14, encoding matriptase, is a critical suppressor of colitis and colitis-associated colon carcinogenesis. *Oncogene* 31 (32), 3679–3695.
- Kunitz, M., 1939. Formation of trypsin from crystalline trypsinogen by means of enterokinase. *Journal of General Physiology* 22 (4), 429–446.
- Kurachi, K., Torres-Rosado, A., Tsuji, A., 1994. Hepsin. *Methods in Enzymology* 244, 100–114.
- Lang, J.C., Schuller, D.E., 2001. Differential expression of a novel serine protease homologue in squamous cell carcinoma of the head and neck. *British Journal of Cancer* 84 (2), 237–243.
- Lee, J., Baek, J.I., Choi, J.Y., *et al.*, 2013. Genetic analysis of TMPRSS3 gene in the Korean population with autosomal recessive nonsyndromic hearing loss. *Gene* 532 (2), 276–280.
- Lee, M.S., Tseng, I.C., Wang, Y., *et al.*, 2007. Autoactivation of matriptase *in vitro*: Requirement for biomembrane and LDL receptor domain. *American Journal of Physiology – Cell Physiology* 293 (1), C95–105.
- Lee, S.L., Dickson, R.B., Lin, C.Y., 2000. Activation of hepatocyte growth factor and urokinase/plasminogen activator by matriptase, an epithelial membrane serine protease. *Journal of Biological Chemistry* 275 (47), 36720–36725.
- Lee, Y.J., Park, D., Kim, S.Y., Park, W.J., 2003. Pathogenic mutations but not polymorphisms in congenital and childhood onset autosomal recessive deafness disrupt the proteolytic activity of TMPRSS3. *Journal of Medical Genetics* 40 (8), 629–631.
- Leyvraz, C., Charles, R.P., Rubera, I., *et al.*, 2005. The epidermal barrier function is dependent on the serine protease CAP1/Prss8. *Journal of Cell Biology* 170 (3), 487–496.

- List, K., 2009. Matriptase: A culprit in cancer? *Future Oncology* 5 (1), 97–104.
- List, K., Currie, B., Scharschmidt, T.C., *et al.*, 2007a. Autosomal ichthyosis with hypotrichosis syndrome displays low matriptase proteolytic activity and is phenocopied in ST14 hypomorphic mice. *Journal of Biological Chemistry* 282 (50), 36714–36723.
- List, K., Haudenschild, C.C., Szabo, R., *et al.*, 2002. Matriptase/MT-SP1 is required for postnatal survival, epidermal barrier function, hair follicle development, and thymic homeostasis. *Oncogene* 21 (23), 3765–3779.
- List, K., Hobson, J.P., Molinolo, A., Bugge, T.H., 2007b. Co-localization of the channel activating protease prostaticin (CAP1/PRSS8) with its candidate activator, matriptase. *Journal of Cellular Physiology* 213 (1), 237–245.
- List, K., Kosa, P., Szabo, R., *et al.*, 2009. Epithelial integrity is maintained by a matriptase-dependent proteolytic pathway. *American Journal of Pathology* 175 (4), 1453–1463.
- List, K., Szabo, R., Molinolo, A., Nielsen, B.S., Bugge, T.H., 2006. Delineation of matriptase protein expression by enzymatic gene trapping suggests diverging roles in barrier function, hair formation, and squamous cell carcinogenesis. *American Journal of Pathology* 168 (5), 1513–1525.
- List, K., Szabo, R., Molinolo, A., *et al.*, 2005. Deregulated matriptase causes ras-independent multistage carcinogenesis and promotes ras-mediated malignant transformation. *Genes & Development* 19 (16), 1934–1950.
- Liu, C., Li, Q., Zhou, X., Kolosov, V.P., Perelman, J.M., 2013. Human airway trypsin-like protease induces mucin5AC hypersecretion via a protease-activated receptor 2-mediated pathway in human airway epithelial cells. *Archives of Biochemistry and Biophysics* 535 (2), 234–240.
- Lopez-Otin, C., Bond, J.S., 2008. Proteases: Multifunctional enzymes in life and disease. *Journal of Biological Chemistry* 283 (45), 30433–30437.
- Macao, B., Johansson, D.G., Hansson, G.C., Hard, T., 2006. Autoproteolysis coupled to protein folding in the SEA domain of the membrane-bound MUC1 mucin. *Nature Structural & Molecular Biology* 13 (1), 71–76.
- Maekawa, A., Kakizoe, Y., Miyoshi, T., *et al.*, 2009. Camostat mesilate inhibits prostaticin activity and reduces blood pressure and renal injury in salt-sensitive hypertension. *Journal of Hypertension* 27 (1), 181–189.
- Malsure, S., Wang, Q., Charles, R.P., *et al.*, 2014. Colon-specific deletion of epithelial sodium channel causes sodium loss and aldosterone resistance. *Journal of the American Society of Nephrology* 25, 1453–1464.
- Matsushima, R., Takahashi, A., Nakaya, Y., *et al.*, 2006. Human airway trypsin-like protease stimulates human bronchial fibroblast proliferation in a protease-activated receptor-2-dependent pathway. *American Journal of Physiology – Lung Cellular and Molecular Physiology* 290 (2), L385–L395.
- Melis, M.A., Cau, M., Congiu, R., *et al.*, 2008. A mutation in the TMPRSS6 gene, encoding a transmembrane serine protease that suppresses hepcidin production, in familial iron deficiency anemia refractory to oral iron. *Haematologica* 93 (10), 1473–1479.
- Miller, G.S., Zoratti, G.L., Murray, A.S., *et al.*, 2014. HATL5: A cell surface serine protease differentially expressed in epithelial cancers. *PLoS One* 9 (2), e87675.
- Min, H.J., Lee, M.K., Lee, J.W., Kim, S., 2014. TMPRSS4 induces cancer cell invasion through pro-uPA processing. *Biochemical and Biophysical Research Communications* 446 (1), 1–7.
- Molina, L., Fasquelle, L., Nouvian, R., *et al.*, 2013. Tmprss3 loss of function impairs cochlear inner hair cell Kcnma1 channel membrane expression. *Human Molecular Genetics* 22 (7), 1289–1299.
- Myerburg, M.M., McKenna, E.E., Luke, C.J., *et al.*, 2008. Prostaticin expression is regulated by airway surface liquid volume and is increased in cystic fibrosis. *American Journal of Physiology – Lung Cellular and Molecular Physiology* 294 (5), L932–L941.
- Netzel-Arnett, S., Bugge, T.H., Hess, R.A., *et al.*, 2009. The glycosylphosphatidylinositol-anchored serine protease PRSS21 (testisin) imparts murine epididymal sperm cell maturation and fertilizing ability. *Biology of Reproduction* 81 (5), 921–932.
- Netzel-Arnett, S., Buzza, M.S., Shea-Donohue, T., *et al.*, 2012. Matriptase protects against experimental colitis and promotes intestinal barrier recovery. *Inflammatory Bowel Diseases* 18 (7), 1303–1314.
- Netzel-Arnett, S., Currie, B.M., Szabo, R., *et al.*, 2006. Evidence for a matriptase-prostaticin proteolytic cascade regulating terminal epidermal differentiation. *Journal of Biological Chemistry* 281 (44), 32941–32945.
- Netzel-Arnett, S., Hooper, J.D., Szabo, R., *et al.*, 2003. Membrane anchored serine proteases: A rapidly expanding group of cell surface proteolytic enzymes with potential roles in cancer. *Cancer and Metastasis Reviews* 22 (2–3), 237–258.
- Nigrovic, P.A., Gray, D.H., Jones, T., *et al.*, 2008. Genetic inversion in mast cell-deficient (Wsh) mice interrupts corin and manifests as hematopoietic and cardiac aberrancy. *American Journal of Pathology* 173 (6), 1693–1701.
- Oberst, M.D., Singh, B., Ozdemirli, M., *et al.*, 2003a. Characterization of matriptase expression in normal human tissues. *Journal of Histochemistry & Cytochemistry* 51 (8), 1017–1025.
- Oberst, M.D., Williams, C.A., Dickson, R.B., Johnson, M.D., Lin, C.Y., 2003b. The activation of matriptase requires its noncatalytic domains, serine protease domain, and its cognate inhibitor. *Journal of Biological Chemistry* 278 (29), 26773–26779.
- Okumura, Y., Takahashi, E., Yano, M., *et al.*, 2010. Novel type II transmembrane serine proteases, MSPL and TMPRSS13, proteolytically activate membrane fusion activity of the hemagglutinin of highly pathogenic avian influenza viruses and induce their multicycle replication. *Journal of Virology* 84 (10), 5089–5096.
- Passero, C.J., Mueller, G.M., Myerburg, M.M., *et al.*, 2012. TMPRSS4-dependent activation of the epithelial sodium channel requires cleavage of the gamma-subunit distal to the furin cleavage site. *American Journal of Physiology – Renal Physiology* 302 (1), F1–F8.
- Peters, D.E., Szabo, R., Friis, S., *et al.*, 2014. The membrane-anchored serine protease prostaticin (CAP1/PRSS8) supports epidermal development and postnatal homeostasis independent of its enzymatic activity. *Journal of Biological Chemistry* 289, 14740–14749.
- Planes, C., Randrianarison, N.H., Charles, R.P., *et al.*, 2010. ENaC-mediated alveolar fluid clearance and lung fluid balance depend on the channel-activating protease 1. *EMBO Molecular Medicine* 2 (1), 26–37.
- Rame, J.E., Drazner, M.H., Post, W., *et al.*, 2007. Corin I555(P568) allele is associated with enhanced cardiac hypertrophic response to increased systemic afterload. *Hypertension* 49 (4), 857–864.
- Rame, J.E., Tam, S.W., McNamara, D., *et al.*, 2009. Dysfunctional corin I555(P568) allele is associated with impaired brain natriuretic peptide processing and adverse outcomes in Blacks with systolic heart failure: Results from the genetic risk assessment in heart failure substudy. *Circulation: Heart Failure* 2 (6), 541–548.
- Ramsay, A.J., Quesada, V., Sanchez, M., *et al.*, 2009. Matriptase-2 mutations in iron-refractory iron deficiency anemia patients provide new insights into protease activation mechanisms. *Human Molecular Genetics* 18 (19), 3673–3683.
- Rawlings, N.D., Barrett, A.J., 2004. Introduction: Serine peptidases and their clans. In: Barrett, A.J., Rawlings, N.D., Woessner, J.F. (Eds.), *Handbook of Proteolytic Enzymes*, second ed., vol. 2. London: Elsevier Ltd, pp. 1417–1439.
- Sakai, K., Ami, Y., Tahara, M., *et al.*, 2014. The host protease TMPRSS2 plays a major role in vivo replication of emerging H7N9 and seasonal influenza viruses. *Journal of Virology* 88 (10), 5608–5616.
- Sales, K.U., Friis, S., Konkil, J.E., *et al.*, 2014. Non-hematopoietic PAR-2 is essential for matriptase-driven pre-malignant progression and potentiation of ras-mediated squamous cell carcinogenesis. *Oncogene*.
- Sales, K.U., Hobson, J.P., Wagenaar-Miller, R., *et al.*, 2011. Expression and genetic loss of function analysis of the HAT/DESC cluster proteases TMPRSS11A and HAT. *PLoS One* 6 (8), e23261.
- Sales, K.U., Masedunskas, A., Bey, A.L., *et al.*, 2010. Matriptase initiates activation of epidermal pro-kallikrein and disease onset in a mouse model of Netherton syndrome. *Nature Genetics* 42 (8), 676–683.
- Scott, H.S., Kudoh, J., Wattenhofer, M., *et al.*, 2001. Insertion of beta-satellite repeats identifies a transmembrane protease causing both congenital and childhood onset autosomal recessive deafness. *Nature Genetics* 27 (1), 59–63.
- Sedghizadeh, P.P., Mallery, S.R., Thompson, S.J., *et al.*, 2006. Expression of the serine protease DESC1 correlates directly with normal keratinocyte differentiation and inversely with head and neck squamous cell carcinoma progression. *Head & Neck* 28 (5), 432–440.
- Selzer-Plon, J., Bornholdt, J., Friis, S., *et al.*, 2009. Expression of prostaticin and its inhibitors during colorectal cancer carcinogenesis. *BMC Cancer* 9, 201.
- Shigemasa, K., Underwood, L.J., Beard, J., *et al.*, 2000. Overexpression of testisin, a serine protease expressed by testicular germ cells, in epithelial ovarian tumor cells. *Journal of the Society for Gynecologic Investigation* 7 (6), 358–362.
- Silvestri, L., Guillem, F., Pagani, A., *et al.*, 2009. Molecular mechanisms of the defective hepcidin inhibition in TMPRSS6 mutations associated with iron-refractory iron deficiency anemia. *Blood* 113 (22), 5605–5608.
- Silvestri, L., Pagani, A., Nai, A., *et al.*, 2008. The serine protease matriptase-2 (TMPRSS6) inhibits hepcidin activation by cleaving membrane hemojuvelin. *Cell Metabolism* 8 (6), 502–511.
- Steensgaard, M., Svenningsen, P., Tinning, A.R., *et al.*, 2010. Apical serine protease activity is necessary for assembly of a high-resistance renal collecting duct epithelium. *Acta Physiologica (Oxford)* 200 (4), 347–359.
- Szabo, R., Bugge, T.H., 2011. Membrane-anchored serine proteases in vertebrate cell and developmental biology. *Annual Review of Cell and Developmental Biology* 27, 213–235.

- Szabo, R., Hobson, J.P., Christoph, K., *et al.*, 2009a. Regulation of cell surface protease matriptase by HAI2 is essential for placental development, neural tube closure and embryonic survival in mice. *Development* 136 (15), 2653–2663.
- Szabo, R., Kosa, P., List, K., Bugge, T.H., 2009b. Loss of matriptase suppression underlies spint1 mutation-associated ichthyosis and postnatal lethality. *American Journal of Pathology* 174 (6), 2015–2022.
- Szabo, R., Molinolo, A., List, K., Bugge, T.H., 2007. Matriptase inhibition by hepatocyte growth factor activator inhibitor-1 is essential for placental development. *Oncogene* 26 (11), 1546–1556.
- Szabo, R., Netzel-Arnett, S., Hobson, J.P., Antalis, T.M., Bugge, T.H., 2005. Matriptase-3 is a novel phylogenetically preserved membrane-anchored serine protease with broad serpin reactivity. *Biochemical Journal* 390 (Pt 1), 231–242.
- Szabo, R., Rasmussen, A.L., Moyer, A.B., *et al.*, 2011. c-Met-induced epithelial carcinogenesis is initiated by the serine protease matriptase. *Oncogene* 30 (17), 2003–2016.
- Szabo, R., Uzzun, S.K., Kosa, P., *et al.*, 2012. Reduced prostasin (CAP1/PRSS8) activity eliminates HAI-1 and HAI-2 deficiency-associated developmental defects by preventing matriptase activation. *PLOS Genetics* 8 (8), e1002937.
- Szabo, R., Wu, Q., Dickson, R.B., *et al.*, 2003. Type II transmembrane serine proteases. *Thrombosis and Haemostasis* 90 (2), 185–193.
- Takahashi, M., Sano, T., Yamaoka, K., *et al.*, 2001. Localization of human airway trypsin-like protease in the airway: An immunohistochemical study. *Histochemistry and Cell Biology* 115 (3), 181–187.
- Takeuchi, T., Harris, J.L., Huang, W., *et al.*, 2000. Cellular localization of membrane-type serine protease 1 and identification of protease-activated receptor-2 and single-chain urokinase-type plasminogen activator as substrates. *Journal of Biological Chemistry* 275 (34), 26333–26342.
- Tang, T., Kmet, M., Corral, L., *et al.*, 2005. Testisin, a glycosyl-phosphatidylinositol-linked serine protease, promotes malignant transformation in vitro and in vivo. *Cancer Research* 65 (3), 868–878.
- Tomlins, S.A., Rhodes, D.R., Perner, S., *et al.*, 2005. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 310 (5748), 644–648.
- Truksa, J., Gelbart, T., Peng, H., *et al.*, 2009. Suppression of the hepcidin-encoding gene *Hamp* permits iron overload in mice lacking both hemojuvelin and matriptase-2/TMPRSS6. *British Journal of Haematology* 147 (4), 571–581.
- Tseng, I.C., Chou, F.P., Su, S.F., *et al.*, 2008. Purification from human milk of matriptase complexes with secreted serpins: Mechanism for inhibition of matriptase other than HAI-1. *American Journal of Physiology – Cell Physiology* 295 (2), C423–C431.
- Tsuji, A., Torres-Rosado, A., Arai, T., *et al.*, 1991. Hepsin, a cell membrane-associated protease. Characterization, tissue distribution, and gene localization. *Journal of Biological Chemistry* 266 (25), 16948–16953.
- Uchimura, K., Hayata, M., Mizumoto, T., *et al.*, 2014. The serine protease prostasin regulates hepatic insulin sensitivity by modulating TLR4 signalling. *Nature Communications* 5, 3428.
- Velasco, G., Cal, S., Quesada, V., Sanchez, L.M., Lopez-Otin, C., 2002. Matriptase-2, a membrane-bound mosaic serine proteinase predominantly expressed in human liver and showing degrading activity against extracellular matrix proteins. *Journal of Biological Chemistry* 277 (40), 37637–37646.
- Verghese, G.M., Gutknecht, M.F., Caughey, G.H., 2006. Prostatin regulates epithelial monolayer function: Cell-specific Gpld1-mediated secretion and functional role for GPI anchor. *American Journal of Physiology – Cell Physiology* 291 (6), C1258–C1270.
- Wang, W., Liao, X., Fukuda, K., *et al.*, 2008. Corin variant associated with hypertension and cardiac hypertrophy exhibits impaired zymogen activation and natriuretic peptide processing activity. *Circulation Research* 103 (5), 502–508.
- Wattenhofer, M., Di, I., Rabionet, V., *et al.*, 2002. Mutations in the TMPRSS3 gene are a rare cause of childhood nonsyndromic deafness in Caucasian patients. *Journal of Molecular Medicine* 80 (2), 124–131.
- Wong, G.W., Foster, P.S., Yasuda, S., *et al.*, 2002a. Biochemical and functional characterization of human transmembrane tryptase (TMT)/tryptase gamma. TMT is an exocytosed mast cell protease that induces airway hyperresponsiveness in vivo via an interleukin-13/interleukin-4 receptor alpha/signal transducer and activator of transcription (STAT) 6-dependent pathway. *Journal of Biological Chemistry* 277 (44), 41906–41915.
- Wong, G.W., Foster, P.S., Yasuda, S., *et al.*, 2002b. Biochemical and functional characterization of human transmembrane tryptase (TMT)/tryptase gamma. TMT is an exocytosed mast cell protease that induces airway hyperresponsiveness in vivo via an IL-13/IL-4Ra/STAT6-dependent pathway. *Journal of Biological Chemistry* 277, 41906–41915.
- Wong, G.W., Tang, Y., Feyfant, E., *et al.*, 1999. Identification of a new member of the tryptase family of mouse and human mast cell proteases which possesses a novel COOH-terminal hydrophobic extension. *Journal of Biological Chemistry* 274 (43), 30784–30793.
- Wu, F., Yan, W., Pan, J., Morser, J., Wu, Q., 2002. Processing of pro-atrial natriuretic peptide by corin in cardiac myocytes. *Journal of Biological Chemistry* 277 (19), 16900–16905.
- Wu, Q., Parry, G., 2007. Hepsin and prostate cancer. *Frontiers in Bioscience* 12, 5052–5059.
- Wu, Q., Yu, D., Post, J., *et al.*, 1998. Generation and characterization of mice deficient in hepsin, a hepatic transmembrane serine protease. *Journal of Clinical Investigation* 101 (2), 321–326.
- Yamaguchi, N., Okui, A., Yamada, T., Nakazato, H., Mitsui, S., 2002. Spinesin/TMPRSS5, a novel transmembrane serine protease, cloned from human spinal cord. *Journal of Biological Chemistry* 277 (9), 6806–6812.
- Yamaoka, K., Masuda, K., Ogawa, H., *et al.*, 1998. Cloning and characterization of the cDNA for human airway trypsin-like protease. *Journal of Biological Chemistry* 273 (19), 11895–11901.
- Yamashita, M., Honda, A., Ogura, A., *et al.*, 2008. Reduced fertility of mouse epididymal sperm lacking Prss21/Tesp5 is rescued by sperm exposure to uterine microenvironment. *Genes to Cells* 13 (10), 1001–1013.
- Yan, W., Sheng, N., Seto, M., Morser, J., Wu, Q., 1999. Corin, a mosaic transmembrane serine protease encoded by a novel cDNA from human heart. *Journal of Biological Chemistry* 274 (21), 14926–14935.
- Yan, W., Wu, F., Morser, J., Wu, Q., 2000. Corin, a transmembrane cardiac serine protease, acts as a pro-atrial natriuretic peptide-converting enzyme. *Proceedings of the National Academy of Sciences of the United States of America* 97 (15), 8525–8529.
- Yasuoka, S., Ohnishi, T., Kawano, S., *et al.*, 1997. Purification, characterization, and localization of a novel trypsin-like protease found in the human airway. *American Journal of Respiratory Cell and Molecular Biology* 16 (3), 300–308.
- Yin, H., Kosa, P., Liu, X., *et al.*, 2014. Matriptase deletion initiates a Sjogren's syndrome-like disease in mice. *PLoS One* 9 (2), e82852.
- Yoshinaga, S., Nakahori, Y., Yasuoka, S., 1998. Fibrinogenolytic activity of a novel trypsin-like enzyme found in human airway. *Journal of Medical Investigation* 45 (1–4), 77–86.
- Yu, J., Yu, J., Mani, R.S., *et al.*, 2010. An integrated network of androgen receptor, polycarb, and TMPRSS2-ERG gene fusions in prostate cancer progression. *Cancer Cell* 17 (5), 443–454.
- Yu, J.X., Chao, L., Chao, J., 1994. Prostatin is a novel human serine proteinase from seminal fluid. Purification, tissue distribution, and localization in prostate gland. *Journal of Biological Chemistry* 269 (29), 18843–18848.
- Yuan, X., Zheng, X., Lu, D., *et al.*, 1998. Structure of murine enterokinase (enteropeptidase) and expression in small intestine during development. *American Journal of Physiology* 274 (2 Pt 1), G342–G349.
- Zheng, X.L., Kitamoto, Y., Sadler, J.E., 2009. Enteropeptidase, a type II transmembrane serine protease. *Frontiers in Bioscience (Elite ed)* 1, 242–249.
- Zhu, H., Guo, D., Li, K., *et al.*, 2008. Prostatin: A possible candidate gene for human hypertension. *American Journal of Hypertension* 21 (9), 1028–1033.

Aspartic Proteases of Alzheimer's Disease: β - and γ -Secretases

MS Wolfe, Harvard Medical School, Boston, MA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Amyloid β -protein Small aggregation-prone proteolytic product that pathologically deposits in the Alzheimer brain.

AmEncyclopedia of Cell Biology
amyloid β -protein precursor A membrane protein that is proteolytically processed by β - and γ -secretases to produce the amyloid β -protein.

BACE1 Also β -secretase, a membrane protein and aspartic protease that cleaves APP to produce the $A\beta$ N-terminus.

Neuregulin A cell-surface receptor that undergoes proteolytic processing by β -secretase as a part of its signaling mechanism.

Notch A cell-surface receptor that undergoes proteolytic processing by γ -secretase as part of its signaling mechanism.

Presenilin A nine-transmembrane domain protein that serves as the catalytic component of the γ -secretase complex.

γ -Secretase A membrane-embedded aspartic protease complex that cleaves APP to produce the $A\beta$ C-terminus; comprised of presenilin, nicastrin, Aph-1, and Pen-2.

Amyloid and Alzheimer's Disease

Presentation and Pathology

Alzheimer's disease (AD) is a devastating progressive neurodegenerative disorder that initially presents as gradual loss of memory and executive function. On average, the disease runs its course over 8 years, but can be as long as 20 years, leading to severe disability and ultimately death. An estimated 5 million are affected in the United States and perhaps 20–30 million worldwide. As increasing age is the greatest risk factor for AD, longer lifespans in the developed world mean more cases of AD. Projections suggest that healthcare systems could be overwhelmed with the problem of AD in the coming decades. Understanding the molecular basis of the AD and leveraging that understanding to discover and develop effective disease-modifying therapeutics is of critical importance.

Two main pathological features are seen postmortem in the AD brain: amyloid plaques and neurofibrillary tangles (Goe-dert and Spillantini, 2006). The amyloid plaques are extracellular and composed primarily of the amyloid β -peptide ($A\beta$), a 38–43 amino acid amphipathic peptide. The

neurofibrillary tangles are intraneuronal and composed of filaments of the otherwise microtubule-associated protein tau. First identified by Alois Alzheimer in 1906, these two pathological features together now define the disease. For reasons that will be detailed later, pathological $A\beta$ is thought to precede and trigger pathological tau. Tau pathology can be found in the absence of amyloid plaques in other neurodegenerative brain diseases, and dominant missense mutations cause non-AD neurodegenerative diseases such as frontotemporal lobar degeneration and progressive supranuclear palsy. These and other findings suggest that a variety of insults, including pathological $A\beta$, can lead to tau pathology and neurodegeneration.

Amyloid Production

$A\beta$ is formed via proteolytic processing of the amyloid β -protein precursor (APP), a 695–770 amino acid type I integral membrane protein (Figure 1) (Tanzi and Bertram, 2005). First, β -secretase sheds the large extracellular/luminal ectodomain, leaving a 99-amino acid C-terminal fragment (CTF- β) in

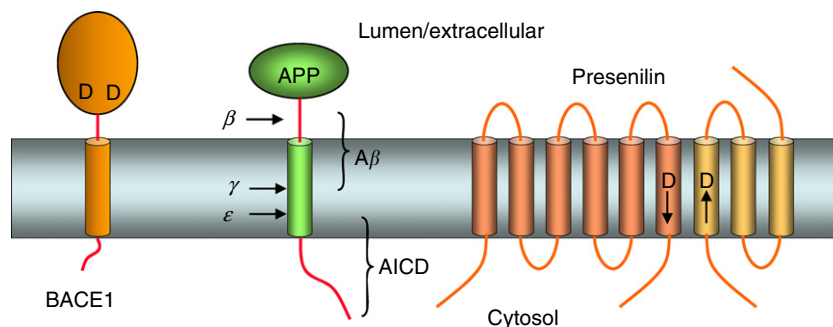


Figure 1 Schematic of BACE1, APP and presenilin and proteolytic sites within APP leading to $A\beta$ production. APP is cleaved first by β -secretase – aka BACE1 – to release the ectodomain. Subsequent cleavage by γ -secretase within the APP TMD forms $A\beta$ (through proteolysis at the γ site) and AICD (through proteolysis at the ϵ site). BACE1 is a membrane-tethered aspartyl protease of the pepsin family (two catalytic aspartates in the ectodomain each labeled as D). Presenilin is the catalytic component of the γ -secretase complex, with two catalytic TMD aspartates. Upon assembly with the other members of the complex, presenilin undergoes autoproteolysis to NTF and CTF subunits to become active γ -secretase.

the membrane. APP CTF- β is then processed within its trans-membrane domain (TMD) by γ -secretase to release the APP intracellular domain (AICD) into the cytosol and A β into the lumen/extracellular space. Alternatively, APP can cleave 17 residues into the A β region by α -secretase to produce an 83-residue CTF- α that is subsequently processed by γ -secretase to release AICD and a truncated A β called p3. These proteolytic events occur normally (i.e., they are not solely pathological occurrences), and A β is secreted constitutively from most cell types, although primarily in neurons.

The normal biological function of APP and its proteolytic processing remains largely unknown (Wolfe and Guenette, 2007). Knockout of the single APP-like gene in flies results in only a mild behavioral phenotype. Knockout of APP in mice leads to impairment in spatial learning and neuronal long-term potential; however, double and triple knockouts with other APP family members (APP-like proteins 1 and 2) are lethal. These and other studies demonstrate that the APP proteins are functionally redundant, are critical to development, and play roles in neuronal cell adhesion and migration. Regarding APP proteolytic products, efforts to identify a true signaling role for AICD has been essentially fruitless. Recent evidence suggests that shed APP ectodomain may be essential to neuronal pruning via apoptosis during development (Nikolaev *et al.*, 2009). A β may be involved in neuronal function at normal physiological levels.

APP and Presenilin Mutations

In 1991, rare, dominantly inherited missense mutation in the APP gene were discovered to be associated with early-onset familial AD (FAD) (Chartier-Harlin *et al.*, 1991; Goate *et al.*, 1991), identical in presentation, progression and pathology to the much more common sporadic AD that strikes in old age. Over 30 such mutations have been found (alzforum.org/mutations), all in and around the small A β region of the large APP protein and to affect the amount or type of secreted A β peptides in ways that increased its aggregation propensity. Mutations near the β -secretase cleavage site increased formation of APP CTF- β and therefore all A β peptides. Mutations near the γ -secretase cleavage site increased the proportion of the much more aggregation-prone 42-residue A β (A β 42), a minor secreted A β variant – A β 40 is the major one – that represents the primary A β variant found in AD amyloid plaques. Mutations within the A β region itself increase the aggregation propensity of the peptide.

In 1995, the first mutations in the presenilin genes, presenilin-1 and -2 (PS1 and PS2) were discovered to be associated with FAD (Rogaev *et al.*, 1995; Sherrington *et al.*, 1995; Levy-Lahad *et al.*, 1995). Now well over 100 dominantly inherited missense mutations have been identified (alzforum.org/mutations), with some causing AD onset even in the third or fourth decade of life. Again, these FAD cases were identical to sporadic AD except for the early age of onset. The presenilins are membrane proteins with 9 TMDs (Laudon *et al.*, 2005; Oh and Turner, 2005; Figure 1), and unlike APP, their connection to AD was entirely unclear upon discovery. The closest known homolog was an obscure gene in *Caenorhabditis elegans* (*spe-4*) associated with spermatogenesis (L'hernault and

Arduengo, 1992). Other, closer worm homologs (*sel-12* and *hop-1*), however, were quickly discovered, and these were found to be involved in processes known to be connected to the Notch receptor, a key signaling molecule in developmental biology (Leviton and Greenwald, 1995; Li and Greenwald, 1997). Still, the connect to AD was completely unclear, until proteolysis of the Notch receptor TMD was found to be essential for Notch signaling to the nucleus (Schroeter *et al.*, 1998). This will be discussed further in the section on γ -secretase.

β -Secretase

Identification, Structure, and Mechanism

β -Secretase was discovered independently by four different groups in 1999–2000 (Vassar *et al.*, 1999; Sinha *et al.*, 1999; Yan *et al.*, 1999; Hussain *et al.*, 1999; Lin *et al.*, 2000), each of which approached the search differently (e.g., expression cloning, biochemical, and bioinformatic). The first report (Vassar *et al.*, 1999) dubbed the enzyme the β -site APP-cleaving protein (BACE), a name which has endured. A close homolog – BACE2 – was quickly discovered (Saunders *et al.*, 1999), with β -secretase then called BACE1. BACE1 is a membrane-tethered member of the pepsin family of aspartic proteases (Figure 1), and a cocrystal structure of the luminal/extracellular domain (no TMD) with a bound transition-state analog inhibitor quickly confirmed the characteristic juxtaposition and coordination of the two conserved aspartic acids required for catalysis (Hong *et al.*, 2000). The structure also revealed two lobes, with the peptidomimetic inhibitor bound at the interface (i.e., the active site is located at the interface between these two lobes). The cytosolic region of BACE1 is very short but contains an acidic cluster-dileucine motif important for subcellular trafficking (Huse *et al.*, 2000; He *et al.*, 2002, 2003).

BACE1 expression is fairly ubiquitous; however, high levels are found in the brain, explaining why this organ makes relatively large amounts of A β . The enzyme cleaves just before the Asp-1 position (A β numbering) within APP, but the FAD-causing Swedish double mutation Lys-Met $\rightarrow$ Asn-Leu at positions -2/-1 is cleaved much more effectively by BACE1 than wild-type APP (Citron *et al.*, 1992), explaining the elevated A β production from this mutant. Recently, an AD-protective mutation in APP with the opposite biochemical effect was discovered in an Icelandic population (Jonsson *et al.*, 2012). This mutation – Ala-2 to Thr – correlates with a decreased risk of AD and an increased retention of cognitive abilities even into old age, along with a decrease in cleavage by BACE1. BACE2 can also cleave APP, but it is poorly expressed in the brain (Bennett *et al.*, 2000) and its APP cleavage site is different from that of BACE1, 20–21 residues within the A β sequence (Farzan *et al.*, 2000).

Knockout Mice

Homozygous deletion of the BACE1 gene in mice resulted in healthy, viable and fertile animals, initially without obvious abnormalities (Cai *et al.*, 2001; Luo *et al.*, 2001). Crossing

these knockout mice with transgenic mice overexpressing human APP in the brain revealed the absence of brain A β production, strong evidence that BACE1 is the primary β -secretase. As the APP transgene contained the Swedish double mutation near the β -secretase cleavage site, some objection has been raised to whether BACE1 is indeed β -secretase, capable of cleaving wild-type APP (Hook *et al.*, 2008). However, BACE1 knockout mice also do not produce endogenous mouse A β from mouse APP (Nishitomi *et al.*, 2006), in which the sequence around the BACE1 cleavage site is conserved. Crossing BACE1 null mice with an AD mouse model expressing a human FAD APP transgene demonstrated the absence of BACE1 prevents cerebral plaque formation and rescues memory deficits seen in these models (Ohno *et al.*, 2004).

Later studies, however, revealed that complete elimination of BACE1 is not as harmless as originally thought. Knockout mice showed reduced synaptic function and plasticity and developed behavioral abnormalities resembling schizophrenia (Savonenko *et al.*, 2008). Moreover, the mice developed substantial loss of myelination of peripheral neuronal axons, due to decreased signaling from neuregulin (Willem *et al.*, 2006). These findings do not necessarily mean that BACE1 is a poor therapeutic target for AD, as proteolytic activity from BACE1 is absent from birth in these null mice. A better model for safety and efficacy with respect to AD would be a conditional knockdown or knockout in mature mice, to avoid developmental defects. Such mice have not yet been reported. Partial reduction of BACE1 may also be safe and sufficient (Mcconlogue *et al.*, 2007).

Other Substrates

After the discovery of BACE1, the search for substrates besides APP began. Early studies identified ST6Gal I, a membrane-bound sialyltransferase (Kitazume *et al.*, 2001). Cleavage by BACE1 results in secretion of the enzyme, thereby regulating its biological role. P-selectin glycoprotein ligand 1, which mediates leukocyte adhesion in inflammatory reactions, was also shown to be a BACE1 substrate (Lichtenthaler *et al.*, 2003). The dramatic reduction in peripheral myelination in BACE1 null mice led to the finding that neuregulin-1 (NRG1) isoforms are also BACE1 substrates (Willem *et al.*, 2006), critical for NRG1 signaling and stimulation of myelination. As NRG1 signaling through its cognate receptor ErbB4 have been implicated in the etiology of schizophrenia, deficient NRG1-ErbB4 signaling may lead to the schizophrenia-like behavioral phenotype in BACE1 knockout mice (Savonenko *et al.*, 2008). These findings have raised concerns about the safety of BACE1 inhibitors as AD therapeutics. Again though, these deficiencies are seen upon complete knockout of BACE1 from conception onward and may not be an indication of safety upon partial inhibition in aging adults.

An unbiased proteomic analysis revealed dozens of putative substrates for BACE1 identified in human HEK293 embryonic kidney cells and HeLa cervical carcinoma cells, both overexpressing BACE1 (Hemming *et al.*, 2009). Many of these substrates were also validated in cell culture. Surprisingly, a clear consensus sequence was not seen, emphasizing the importance of taking an unbiased approach. Almost all of the

substrates were type I integral membrane proteins, although one was type II and three were glycosphosphatidylinositol (GPI)-anchored. The biological roles of these various BACE1-mediated proteolytic events remains largely unclear, although many of these substrates are receptors or are involved in cell-cell contact. Two more recent proteomics studies employed neurons and without overexpression (Zhou *et al.*, 2012; Kuhn *et al.*, 2012), making the physiological relevance of the findings more likely. Among the more interesting validated substrates were the neuronal cell adhesion molecules (NCAMs) L1 and CHL1, both of which localize to synapses and may affect plasticity.

Inhibitors

The first designed BACE1 inhibitors were substrate-based peptidomimetics containing a hydroxyethylene moiety that mimicked the transition-state of aspartyl protease catalysis (Ghosh *et al.*, 2000). The compounds were highly potent, inhibiting BACE1 with K_i values at or below 1 nM. Various types of transition-state analog inhibitors of BACE1 were developed, and these helped define the side chain preferences for the active site pockets on the enzyme. Consistent with loose sequence specificity of substrates (Turner *et al.*, 2002; Gruninger-Leitch *et al.*, 2002), none of the active site pockets were found to be particularly specific, although some preferences could be determined. Co-crystallization of BACE1 with one of these inhibitors, which spanned the P4-P4' positions, provided important information about the nature of the active site and substrate recognition (Hong *et al.*, 2000). The active site was located between the two lobes of the protease, each contributing one conserved aspartate, and found to be relatively long and shallow. This explained the loss of potency upon shortening the peptidomimetic inhibitors. Moreover, the inhibitor was covered by a flap, a common feature found in other aspartyl proteases of the pepsin family.

Peptidomimetic inhibitors, however, typically do not make for useful drugs. This is particularly true for BACE1 inhibitors, as compounds must traverse the blood-brain barrier, in addition to being metabolically stable in the periphery. The large size required for potency of transition-state mimicking peptide analogs did not allow good transport into the brain in general. Thus, considerable efforts have gone into identifying small, nonpeptide inhibitors with good potency, specificity and drug-like *in vivo* properties (Ghosh and Osswald, 2014). Such inhibitors have been identified through screening of chemical scaffolds followed by optimization, often using structure-based design. Another strategy has been computational screening, which has provided initial hits for modification into potent inhibitors.

Prospects for Therapeutics

The structure of BACE1 bound to an inhibitor was first reported in 2000 (Hong *et al.*, 2000), but the development of orally bioavailable BACE1 inhibitors with the ability to lower brain A β has been difficult and slow (Ghosh and Osswald, 2014). The primary reasons for this are the characteristics of the BACE1 active site (long and shallow, making the

identification of small, nonpeptidic inhibitors challenging) and the fact that many compounds that inhibit BACE1 are substrates for P-glycoprotein (which leads to expulsion of the compounds from the brain if they do manage to get in) (Meredith *et al.*, 2008). Nevertheless, compounds that lowered brain A β upon oral administration in AD mouse models (e.g., Hussain *et al.*, 2007) and nonhuman primates (e.g., Sankaranarayanan *et al.*, 2009) eventually emerged, and several compounds have advanced into clinical trials in recent years.

Scientists at Eli Lilly discovered an aminothiazine, LY2811376, as a potent, orally bioavailable nonpeptidic inhibitor that lowered A β in the central nervous system (specifically in the cerebrospinal fluid (CSF)) in healthy volunteers, and the compound appeared to be safe and well tolerated (May *et al.*, 2011). However, while the candidate progressed in clinical trials, further evaluation in animals revealed serious retinal toxicity. Reports conflict on whether retinal pathology appears in BACE1 knockout mice (e.g., Cai *et al.*, 2012). However, inhibition of the related BACE2 cannot be ruled out, as the compound is only modestly selective for BACE1 over BACE2. Another aminothiazine, LY2866721, soon followed but was recently halted due to liver toxicity. The reasons for the liver toxicity are unclear, but do not appear to be an on-target effect. As of this writing, AstraZeneca is moving into advanced clinical trials with an amino-isindole AZD3839, which displayed safety and efficacy (in terms of lowering CSF A β), as has Merck's MK-8931 (undisclosed structure).

γ -Secretase

Identification and Mechanism

Soon after their discovery, dominant FAD-associated presenilin mutations were determined to increase the ratio of aggregation-prone A β 42 to A β 40, in transfected cell lines, transgenic mice, patient-derived human fibroblasts, and patient plasma samples (Borchelt *et al.*, 1996; Citron *et al.*, 1997; Tomita *et al.*, 1997; Xia *et al.*, 1997; Duff *et al.*, 1996; Scheuner *et al.*, 1996). These findings suggested that presenilin can modulate γ -secretase activity, directly or indirectly, to change the cleavage site specificity within the APP TMD. The presenilins were also found to assemble into higher molecular weight complexes (Capell *et al.*, 1998; Yu *et al.*, 1998) and to be cleaved into stable N-terminal and C-terminal fragments (NTFs and CTFs) in a regulated manner (Thinakaran *et al.*, 1996; Ratovitski *et al.*, 1997), suggesting that functional presenilin is an NTF/CTF heterodimer and part of a larger complex. Knockout of PS1 in mice is embryonic lethal (Shen *et al.*, 1997; Wong *et al.*, 1997); however, culturing of the embryonic fibroblasts revealed a dramatic reduction in γ -secretase activity (reduced A β production and elevation of APP CTFs) (De Strooper *et al.*, 1998). PS1/PS2 double knockout cells then demonstrated that presenilins are absolutely required for γ -secretase activity (Herreman *et al.*, 2000; Zhang *et al.*, 2000).

The specific role of presenilin, however, was unclear until the recognition that these proteins contain two completely conserved transmembrane aspartates, one in TMD6 and one in TMD7, that are essential for γ -secretase activity (Wolfe *et al.*,

1999b). This finding was consistent with the fact that substrate-based peptidomimetics containing aspartyl protease transition-state mimicking moieties could inhibit γ -secretase (Wolfe *et al.*, 1999a; Shearman *et al.*, 2000). Indeed, such active site-directed inhibitors bind directly to presenilin NTF and CTF (Li *et al.*, 2000a; Esler *et al.*, 2000), suggesting that the active site is at the interface between these two presenilin subunits, each of which contribute one of the two conserved and essential TMD aspartates. Presenilin is now known to be part of a larger family of membrane-embedded aspartyl protease (Wolfe, 2009), with no sequence resemblance to the pepsin family of soluble aspartyl proteases.

The factors regulating presenilin NTF/CTF formation were then discovered to be the other protein members of the higher molecular weight complex into which presenilin assembles (Takasugi *et al.*, 2003; Kimberly *et al.*, 2003; Edbauer *et al.*, 2003). Three transmembrane proteins, nicastrin, Pen-2 and Aph-1, are all necessary for γ -secretase activity, and these proteins along with presenilin are sufficient for this proteolytic activity as well. The four components assemble together in the ER (Capell *et al.*, 2005), whereupon presenilin is cleaved into NTF and CTF subunits, a process that also requires the two TMD presenilin aspartates (i.e., presenilin is a zymogen that undergoes autoproteolysis) (Wolfe *et al.*, 1999b; Fukumori *et al.*, 2010). While there had been some controversy over the size and stoichiometry of the complex, a rigorous analysis by co-immunoprecipitation established that only one of each component is required for γ -secretase activity (Sato *et al.*, 2007), and this finding was unambiguously confirmed by electron microscopic analysis of purified, active complexes (see below).

In recent years, evidence has mounted supporting a processive proteolysis mechanism for γ -secretase cleavage of APP. The first clue for such a mechanism came from the finding of discrepancies between the C-termini of secreted A β peptides and the N-termini of the AICD (Weidemann *et al.*, 2002) (see γ and ϵ cleavage sites in the APP TMD in Figure 1). Moreover, longer, membrane-associated forms of A β , ranging from 45–49 residues, were discovered, taking the A β sequence all the way to where AICD begins (Qi-Takahara *et al.*, 2005). These findings suggested that γ -secretase first cuts the APP TMD to release AICD and form A β 48 or A β 49; the long A β peptides are then trimming to secreted A β s (Funamoto *et al.*, 2004). Furthermore, the pattern of A β peptides produced suggested that the C-terminal trimming generally occurs every three or four amino acids, an idea supported by mass spectrometric detection of the expected tri- and tetrapeptide products (Takami *et al.*, 2009). Importantly, A β 42 is a precursor to A β 38, and this step can be stimulated by certain modulators of γ -secretase activity (Okochi *et al.*, 2013) (see later).

Other Substrates

The finding that presenilin deficiency in *C. elegans* resulted in a phenotype resembling that of Notch deficiency led to the discovery that the Notch receptor, like APP, is cleaved within its TMD by a presenilin-dependent γ -secretase-like activity and that this proteolysis is required for Notch signaling (Schroeter *et al.*, 1998; De Strooper *et al.*, 1999). Release of the Notch

intracellular domain from the membrane is followed by translocation to the nucleus, with binding and activation of (CBF-1, Suppressor of Hairless, Lag-2) CSL transcription factors. Notch signaling is involved in many cell differentiation events, its physiological role often tissue- and context-dependent (Kopan and Ilgan, 2009). Complete knockout of presenilins or other γ -secretase components in other animals besides worms also leads to a Notch-deficient and embryonic lethal phenotype, suggesting that cleavage of Notch is the most important role of γ -secretase in all metazoans.

Over the years, dozens of other substrates besides APP and Notch have been described, all of which are type I integral membrane proteins that undergo ectodomain shedding (Hemming *et al.*, 2008; Haapasalo and Kovacs, 2011). Some of these γ -secretase-mediated proteolytic events are part of signaling pathways, similar to Notch. These include N-cadherin, the intracellular domain of which interacts with transcriptional activator CBP (CREB binding protein), leading to translocation to the cytosol and degradation by the proteasome (Marambaud *et al.*, 2003); and ErbB4, the intracellular domain of which binds to repressors of astrocyte gene expression to inhibit astrocyte differentiation (Sardi *et al.*, 2006). However, many γ -secretase-generated products do not appear to have a signaling function, leading to the hypothesis that γ -secretase helps clear away membrane stubs left behind after ectodomain shedding (i.e., that γ -secretase is 'the proteasome of the membrane') (Kopan and Ilgan, 2004).

Inhibitors and Modulators

As with BACE1, the first γ -secretase inhibitors were substrate-based peptidomimetics. Design of peptides based on the cleavage sites within the APP TMD that lead to A β 40 or A β 42 with aspartyl protease transition-state mimicking moieties provided compounds that inhibited A β production and elevated APP CTF γ -secretase substrates in cell culture (Wolfe *et al.*, 1998, 1999a; Shearman *et al.*, 2000). Varying the hydrophobic side chains in these peptidomimetic transition-state analogs (e.g., Ala, Val, Ile, Leu, and Phe) typically did not substantially affect inhibitor potency, suggesting that the protease has large hydrophobic pockets that can accommodate these residues (Wolfe *et al.*, 1999a; Moore *et al.*, 2000; Esler *et al.*, 2004; Bakshi and Wolfe, 2004). As mentioned above, conversion of these types of active site-directed inhibitors into affinity labeling reagents led to the tagging of presenilin NTFs and CTFs (Li *et al.*, 2000b; Esler *et al.*, 2000). Moreover, immobilization of transition-state analog inhibitors allowed purification of the γ -secretase complex (Esler *et al.*, 2002; Kimberly *et al.*, 2003; Beher *et al.*, 2003).

Other peptide analog inhibitors include APP TMD-based peptides containing the helix-inducing aminoisobutyric acid, to better mimic the substrate conformation that initially interacts with the enzyme (Das *et al.*, 2003). Such compounds could be highly potent (Bihel *et al.*, 2004), and conversion to affinity labeling reagents led to tagging of presenilin NTFs and CTFs (Kornilova *et al.*, 2005), as with transition-state analog inhibitors. However, competition experiments demonstrated that the binding sites were distinct, suggesting the presence of an initial substrate docking site on the outer surface of

presenilin, at the NTF/CTF interface. These findings suggested that substrate TMDs first dock onto presenilin before passing between NTF and CTF to access the protease active site.

As with BACE1 inhibitors, nonpeptide γ -secretase inhibitors were sought to identify leads with better drug-like properties. Many such inhibitors have been described (Kreft *et al.*, 2009), some with very high potencies. Interestingly, upon conversion of these inhibitors into affinity labeling reagents, all those that have been so converted tag presenilin, either NTF, CTF or both. No other components of the protease complex are labeled by inhibitors. Despite this, nicastrin and Aph-1 have been implicated in substrate recognition along with presenilin (Shah *et al.*, 2005; Chen *et al.*, 2010).

Concern about Notch-based toxic consequences of γ -secretase inhibitors has led to two strategies for developing safe and effective therapeutics for AD that target the protease. The first is to identify so-called 'Notch-sparing' inhibitors (Imbimbo *et al.*, 2011). These compounds show varying degrees of selectivity (with some reports of no selectivity) for inhibiting the cleavage of APP over that of Notch by γ -secretase, apparently through interaction with an allosteric site on the enzyme (again, located on presenilin) (Crump *et al.*, 2012). The second strategy is to develop modulators of γ -secretase that lower A β 42 without inhibiting total A β peptide levels (Golde *et al.*, 2013). These compounds apparently work by enhancing the conversion of A β 42 to A β 38 (Okochi *et al.*, 2013). The therapeutic potential of γ -secretase inhibitors and modulators will be described later.

Toward Structure Elucidation

A high resolution structure of the γ -secretase complex has not yet been determined, undoubtedly due to the fact that the enzyme is composed of four different membrane proteins, and with likely heterogeneity in the glycosylation of nicastrin. Another limitation is that the protease has not yet been successfully reconstituted from separately expressed and purified components: the four proteins must be expressed together in a eukaryotic cell line prior to purification for structural studies. Nevertheless, a number of structures have been reported using electron microscopy (EM) and single-particle analysis (Lazarov *et al.*, 2006; Ogura *et al.*, 2006; Osenkowski *et al.*, 2009; Renzi *et al.*, 2011; Li *et al.*, 2014), including most recently a structure solved to 4.5 Å resolution (Lu *et al.*, 2014). Critical prerequisites for structure elucidation, however, were determination of the stoichiometry of the complex (one of each component) (Sato *et al.*, 2007) and determination of its size (~230 kDa, consistent with the 1:1:1:1 stoichiometry) (Osenkowski *et al.*, 2009). Also critical were prior elucidation of the membrane topologies of PS1 (Laudon *et al.*, 2005; Oh and Turner, 2005), Aph-1 (Fortna *et al.*, 2004) and Pen-2 (Crystal *et al.*, 2003) as well as the protein-protein contacts (e.g., Pen-2 with PS1 NTF, nicastrin with Aph-1, and PS1 CTF with either of these two partial complexes) (Fraering *et al.*, 2004; Steiner *et al.*, 2008). Cysteine scanning mutagenesis and use of water accessible crosslinking reagents has also proven useful in probing the catalytic pore of PS1 (Tolia *et al.*, 2006, 2008; Sato *et al.*, 2006, 2008; Takagi *et al.*, 2010).

Earlier EM studies, with resolutions of 12 Å at best, revealed relatively featureless globular structures. However, several characteristics could be gleaned. The location of the nicastrin ectodomain was determined using lectins or antibodies, thereby orienting the complex with respect to lumenal/extracellular, cytosolic and transmembrane region. The transmembrane region appears as a high-density belt but with low-density regions that make be sites of entry for catalytic water and a surface groove that may be the substrate docking site. The most recent 4.5 Å cryo-EM structure (Lu *et al.*, 2014), solved using a combination of improved expression and purification and the more sensitive direct electron detector system. The TMDs of the complex were arranged in a horseshoe shape, with the part of the nicastrin ectodomain implicated in substrate recognition intriguingly located above the horseshoe cavity (Figure 2). Thus, nicastrin appears to be in position to guide the substrate into the horseshoe and the TMD active site of the complex. The location of presenilin within the arrangement of TMDs is unclear, despite the prior elucidation of a higher-resolution structure of an archaeal presenilin homolog.

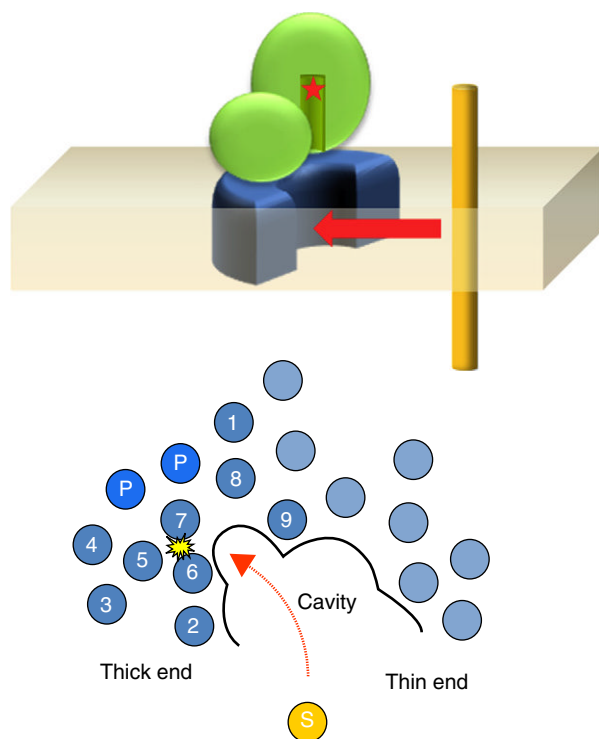


Figure 2 Architecture of the γ -secretase complex and implications for substrate access and recognition. Top: A recently reported 4.5 Å cryo-EM structure reveals a horseshoe-shaped arrangement of TMDs and a bi-lobed nicastrin ectodomain. The nicastrin sequence and structure is homologous to a peptidase, but without catalytic activity of its own. Nicastrin's catalytically dead site (red star) is part of a groove located above the opening to the horseshoe-shaped TMD arrangement and may be important for substrate recognition. Bottom: Speculated arrangement of specific TMDs of PS1 (numbered) and Pen-2 (P) in the thick end, based in part on the structure of an archaeal presenilin homolog. TMDs of Aph-1 and nicastrin are nonspecifically assigned to the thin end. Substrate (S) approaches the active site of PS1 via the cavity.

Prospects for Therapeutics

While many γ -secretase inhibitors and modulators have been reported, only a handful of these have entered into clinical trials (Crump *et al.*, 2013). The first modulator into the clinic was R-flurbiprofen (Eriksen *et al.*, 2003; Kukar *et al.*, 2007; Wilcock *et al.*, 2008), the enantiomer of the nonsteroidal anti-inflammatory drug (NSAID) S-flurbiprofen. A number of NSAIDs show A β 42-lowering effects on γ -secretase processing of APP (these were the first γ -secretase modulators discovered (Weggen *et al.*, 2001)), but this activity does not correlate with inhibition of the NSAID target cyclooxygenase. Indeed, R-flurbiprofen is not a cyclooxygenase inhibitor, and A β 42-lowering NSAIDs do so even in cyclooxygenase-deficient cells. However, R-flurbiprofen is a very weak modulator, and despite being quite safe, even very high doses were not efficacious in patients (Green *et al.*, 2009). Other, much more potent second- and third-generation modulators have since been developed (Crump *et al.*, 2013), but so far none have reached late-stage clinical trials.

A major problem with clinical trials for γ -secretase modulators, as with all candidate AD therapeutics to date, has been that patients enrolled to test these agents already had mild-to-moderate AD. As substantial neurodegeneration of the hippocampus – a brain region critical to learning and memory – has occurred by the time AD manifests itself cognitively (Barnes *et al.*, 2009), reducing A β 42 levels at this point is unlikely to be effective. Thus, there is intense interest in AD biomarkers and diagnostics, in order to identify presymptomatic subjects who are highly likely to develop AD within the span of a clinical trial (Sperling *et al.*, 2011). Presymptomatic carriers of FAD mutations are of particular interest for new drug trials.

γ -Secretase inhibitors, however, have other, much more serious problems. As mentioned earlier, γ -secretase cleavage of Notch receptors is an essential part of a signaling pathway critical for many different cell-fate determinations, both in embryogenesis and in adulthood (Kopan and Ilagan, 2009). γ -Secretase inhibitors cause a variety of toxic effects, all related to deficient Notch signaling, including immunosuppression, gastrointestinal bleeding and diarrhea, and cancerous skin lesions. For these reasons, a potent, brain-penetrant γ -secretase inhibitor (semagacestat) was halted in late-phase clinical trials (Doody *et al.*, 2013). Even more worrisome, this compound caused cognitive worsening in AD patients. While this disturbing result may be linked to inhibition of Notch signaling, it is also possible that elevation of APP CTFs is responsible, as cognitive worsening with γ -secretase inhibitors has been seen at lower doses in APP transgenic mice than in wild-type mice (Mitani *et al.*, 2012). Indeed, another clinical candidate (ava-gacestat), with selectivity for inhibiting γ -secretase cleavage of APP over that of Notch (Gillman *et al.*, 2010), also caused cognitive worsening and was halted after phase II trials (Coric *et al.*, 2012). Discerning whether elevated APP CTFs is responsible for cognitive worsening is critically important in deciding whether further development of γ -secretase inhibitors that avoid effects on Notch function is a sensible approach. If instead cognitive worsening is due to blocking Notch signaling, improving on APP/Notch selectivity would be potentially worthwhile, as would targeting specific isoforms of the γ -secretase complex.

See also: Molecular Principles, Components, Technology, and Concepts: Proteins: Diseases of Protein Folding: Huntington's Disease and Amyotrophic Lateral Sclerosis; Folding, Misfolding, Disordered Proteins, and Related Diseases. Protein Synthesis/Degradation: Protein Degradation – Protease Classes: Cathepsin E: An Aspartic Protease with Diverse Functions and Biomedical Implications

References

- Bakshi, P., Wolfe, M.S., 2004. Stereochemical analysis of (hydroxyethyl)urea peptidomimetic inhibitors of gamma-secretase. *Journal of Medicinal Chemistry* 47, 6485–6489.
- Barnes, J., Bartlett, J.W., Van De Pol, L.A., *et al.*, 2009. A meta-analysis of hippocampal atrophy rates in Alzheimer's disease. *Neurobiology of Aging* 30, 1711–1723.
- Behr, D., Fricker, M., Nadin, A., *et al.*, 2003. *In vitro* characterization of the presenilin-dependent gamma-secretase complex using a novel affinity ligand. *Biochemistry* 42, 8133–8142.
- Bennett, B.D., Babu-Khan, S., Loeloff, R., *et al.*, 2000. Expression analysis of BACE2 in brain and peripheral tissues. *Journal of Biological Chemistry* 275, 20647–20651.
- Bihel, F., Das, C., Bowman, M.J., Wolfe, M.S., 2004. Discovery of a subnanomolar helical D-tridecapeptide inhibitor of γ -secretase. *Journal of Medicinal Chemistry* 47, 3931–3933.
- Borchelt, D.R., Thinakaran, G., Eckman, C.B., *et al.*, 1996. Familial Alzheimer's disease-linked presenilin 1 variants elevate Abeta1-42/1-40 ratio *in vitro* and *in vivo*. *Neuron* 17, 1005–1013.
- Cai, H., Wang, Y., McCarthy, D., *et al.*, 2001. BACE1 is the major beta-secretase for generation of Abeta peptides by neurons. *Nature Neuroscience* 4, 233–234.
- Cai, J., Qi, X., Kociok, N., *et al.*, 2012. beta-Secretase (BACE1) inhibition causes retinal pathology by vascular dysregulation and accumulation of age pigment. *EMBO Molecular Medicine* 4, 980–991.
- Capell, A., Behr, D., Prokop, S., *et al.*, 2005. Gamma-secretase complex assembly within the early secretory pathway. *Journal of Biological Chemistry* 280, 6471–6478.
- Capell, A., Grunberg, J., Pesold, B., *et al.*, 1998. The proteolytic fragments of the Alzheimer's disease-associated presenilin-1 form heterodimers and occur as a 100–150-kDa molecular mass complex. *Journal of Biological Chemistry* 273, 3205–3211.
- Chartier-Harlin, M.C., Crawford, F., Houlden, H., *et al.*, 1991. Early-onset Alzheimer's disease caused by mutations at codon 717 of the beta-amyloid precursor protein gene. *Nature* 353, 844–846.
- Chen, A.C., Guo, L.Y., Ostaszewski, B.L., Selkoe, D.J., Lavoie, M.J., 2010. Aph-1 associates directly with full-length and C-terminal fragments of gamma-secretase substrates. *Journal of Biological Chemistry* 285, 11378–11391.
- Citron, M., Oltschdorf, T., Haass, C., *et al.*, 1992. Mutation of the beta-amyloid precursor protein in familial Alzheimer's disease increases beta-protein production. *Nature* 360, 672–674.
- Citron, M., Westaway, D., Xia, W., *et al.*, 1997. Mutant presenilins of Alzheimer's disease increase production of 42-residue amyloid beta-protein in both transfected cells and transgenic mice. *Nature Medicine* 3, 67–72.
- Coric, V., Van Dyck, C.H., Salloway, S., *et al.*, 2012. Safety and tolerability of the gamma-secretase inhibitor avagacestat in a phase 2 study of mild to moderate Alzheimer disease. *Archives of Neurology* 69, 1430–1440.
- Crump, C.J., Castro, S.V., Wang, F., *et al.*, 2012. BMS-708,163 targets presenilin and lacks notch-sparing activity. *Biochemistry* 51, 7209–7211.
- Crump, C.J., Johnson, D.S., Li, Y.M., 2013. Development and mechanism of gamma-secretase modulators for Alzheimer's disease. *Biochemistry* 52, 3197–3216.
- Crystal, A.S., Morais, V.A., Pierson, T.C., *et al.*, 2003. Membrane topology of gamma-secretase component PEN-2. *Journal of Biological Chemistry* 278, 20117–20123. doi:10.1074/jbc.M213107200.
- Das, C., Berezovska, O., Diehl, T.S., *et al.*, 2003. Designed helical peptides inhibit an intramembrane protease. *Journal of the American Chemical Society* 125, 11794–11795.
- De Strooper, B., Annaert, W., Cupers, P., *et al.*, 1999. A presenilin-1-dependent γ -secretase-like protease mediates release of Notch intracellular domain. *Nature* 398, 518–522.
- De Strooper, B., Saftig, P., Craessaerts, K., *et al.*, 1998. Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein. *Nature* 391, 387–390.
- Doody, R.S., Raman, R., Farlow, M., *et al.*, 2013. A phase 3 trial of semagacestat for treatment of Alzheimer's disease. *New England Journal of Medicine* 369, 341–350.
- Duff, K., Eckman, C., Zehr, C., *et al.*, 1996. Increased amyloid-beta42(43) in brains of mice expressing mutant presenilin 1. *Nature* 383, 710–713.
- Edbauer, D., Winkler, E., Regula, J.T., *et al.*, 2003. Reconstitution of gamma-secretase activity. *Nature Cell Biology* 5, 486–488.
- Eriksen, J.L., Sagi, S.A., Smith, T.E., *et al.*, 2003. NSAIDs and enantiomers of flurbiprofen target gamma-secretase and lower Abeta 42 *in vivo*. *Journal of Clinical Investigation* 112, 440–449.
- Esler, W.P., Das, C., Wolfe, M.S., 2004. Probing pockets S2-S4' of the gamma-secretase active site with (hydroxyethyl)urea peptidomimetics. *Bioorganic & Medicinal Chemistry Letters* 14, 1935–1938.
- Esler, W.P., Kimberly, W.T., Ostaszewski, B.L., *et al.*, 2000. Transition-state analogue inhibitors of γ -secretase bind directly to presenilin-1. *Nature Cell Biology* 2, 428–434.
- Esler, W.P., Kimberly, W.T., Ostaszewski, B.L., *et al.*, 2002. Activity-dependent isolation of the presenilin/ γ -secretase complex reveals nicastrin and a γ substrate. *Proceedings of the National Academy of Sciences of the United States of America* 99, 2720–2725.
- Farzan, M., Schnitzler, C.E., Vasilieva, N., Leung, D., Choe, H., 2000. BACE2, a beta-secretase homolog, cleaves at the beta site and within the amyloid-beta region of the amyloid-beta precursor protein. *Proceedings of the National Academy of Sciences of the United States of America* 97, 9712–9717.
- Fortna, R.R., Crystal, A.S., Morais, V.A., *et al.*, 2004. Membrane topology and nicastrin-enhanced endoproteolysis of APH-1, a component of the gamma-secretase complex. *Journal of Biological Chemistry* 279, 3685–3693.
- Fraering, P.C., Lavoie, M.J., Ye, W., *et al.*, 2004. Detergent-dependent dissociation of active gamma-secretase reveals an interaction between Pen-2 and PS1-NTF and offers a model for subunit organization within the complex. *Biochemistry* 43, 323–333.
- Fukumori, A., Fluhrer, R., Steiner, H., Haass, C., 2010. Three-amino acid spacing of presenilin endoproteolysis suggests a general stepwise cleavage of gamma-secretase-mediated intramembrane proteolysis. *Journal of Neuroscience* 30, 7853–7862.
- Funamoto, S., Morishima-Kawashima, M., Tanimura, Y., *et al.*, 2004. Truncated carboxyl-terminal fragments of beta-amyloid precursor protein are processed to amyloid beta-proteins 40 and 42. *Biochemistry* 43, 13532–13540.
- Ghosh, A.K., Osswald, H.L., 2014. BACE1 (beta-secretase) inhibitors for the treatment of Alzheimer's disease. *Chemical Society Reviews* 43, 6765–6813.
- Ghosh, A.K., Shin, D., Downs, D., *et al.*, 2000. Design of potent inhibitors for human brain memapsin 2 (beta-secretase). *Journal of the American Chemical Society* 122, 3522–3523.
- Gillman, K.W., Starrett, J.E., Parker, M.F., *et al.*, 2010. Discovery and evaluation of BMS-708163, a potent, selective and orally bioavailable gamma-secretase inhibitor. *ACS Medicinal Chemistry Letters* 1, 120–124.
- Goate, A., Chartier-Harlin, M.C., Mullan, M., *et al.*, 1991. Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature* 349, 704–706.
- Goedert, M., Spillantini, M.G., 2006. A century of Alzheimer's disease. *Science* 314, 777–781.
- Golde, T.E., Koo, E.H., Felsenstein, K.M., Osborne, B.A., Miele, L., 2013. gamma-Secretase inhibitors and modulators. *Biochimica et Biophysica Acta* 1828, 2898–2907.
- Green, R.C., Schneider, L.S., Amato, D.A., *et al.*, 2009. Effect of tarenflurbil on cognitive decline and activities of daily living in patients with mild Alzheimer disease: A randomized controlled trial. *JAMA* 302, 2557–2564.
- Gruninger-Leitch, F., Schlatter, D., Kung, E., Nelbock, P., Dobeli, H., 2002. Substrate and inhibitor profile of BACE (beta-secretase) and comparison with other mammalian aspartic proteases. *Journal of Biological Chemistry* 277, 4687–4693.
- Haapasalo, A., Kovacs, D.M., 2011. The many substrates of presenilin/gamma-secretase. *Journal of Alzheimer's Disease* 25, 3–28.
- He, X., Chang, W.P., Koelsch, G., Tang, J., 2002. Memapsin 2 (beta-secretase) cytosolic domain binds to the VHS domains of GGA1 and GGA2: Implications on the endocytosis mechanism of memapsin 2. *FEBS Letters* 524, 183–187.
- He, X., Zhu, G., Koelsch, G., *et al.*, 2003. Biochemical and structural characterization of the interaction of memapsin 2 (beta-secretase) cytosolic domain with the VHS domain of GGA proteins. *Biochemistry* 42, 12174–12180.
- Hemming, M.L., Elias, J.E., Gygi, S.P., Selkoe, D.J., 2008. Proteomic profiling of gamma-secretase substrates and mapping of substrate requirements. *PLoS Biology* 6, e257.

- Hemming, M.L., Elias, J.E., Gygi, S.P., Selkoe, D.J., 2009. Identification of beta-secretase (BACE1) substrates using quantitative proteomics. *PLoS One* 4, e8477.
- Herreman, A., Serneels, L., Annaert, W., *et al.*, 2000. Total inactivation of gamma-secretase activity in presenilin-deficient embryonic stem cells. *Nature Cell Biology* 2, 461–462.
- Hong, L., Koelsch, G., Lin, X., *et al.*, 2000. Structure of the protease domain of memapsin 2 (beta-secretase) complexed with inhibitor. *Science* 290, 150–153.
- Hook, V.Y., Kindy, M., Hook, G., 2008. Inhibitors of cathepsin B improve memory and reduce beta-amyloid in transgenic Alzheimer disease mice expressing the wild-type, but not the Swedish mutant, beta-secretase site of the amyloid precursor protein. *Journal of Biological Chemistry* 283, 7745–7753.
- Huse, J.T., Pijak, D.S., Leslie, G.J., Lee, V.M., Doms, R.W., 2000. Maturation and endosomal targeting of beta-site amyloid precursor protein-cleaving enzyme. The Alzheimer's Disease β -Secretase. *Journal of Biological Chemistry* 275, 33729–33737.
- Hussain, I., Hawkins, J., Harrison, D., *et al.*, 2007. Oral administration of a potent and selective non-peptidic BACE-1 inhibitor decreases beta-cleavage of amyloid precursor protein and amyloid-beta production *in vivo*. *Journal of Neurochemistry* 100, 802–809.
- Hussain, I., Powell, D., Howlett, D.R., *et al.*, 1999. Identification of a novel aspartic protease (Asp 2) as beta-secretase. *Molecular and Cellular Neuroscience* 14, 419–427.
- Imbimbo, B.P., Panza, F., Frisardi, V., *et al.*, 2011. Therapeutic intervention for Alzheimer's disease with gamma-secretase inhibitors: Still a viable option? *Expert Opinion on Investigational Drugs* 20, 325–341.
- Jonsson, T., Atwal, J.K., Steinberg, S., *et al.*, 2012. A mutation in APP protects against Alzheimer's disease and age-related cognitive decline. *Nature* 488, 96–99.
- Kimberly, W.T., Lavoie, M.J., Ostaszewski, B.L., *et al.*, 2003. γ -Secretase is a membrane protein complex comprised of presenilin, nicastrin, aph-1, and pen-2. *Proceedings of the National Academy of Sciences of the United States of America* 100, 6382–6387.
- Kitazume, S., Tachida, Y., Oka, R., *et al.*, 2001. Alzheimer's beta-secretase, beta-site amyloid precursor protein-cleaving enzyme, is responsible for cleavage secretion of a Golgi-resident sialyltransferase. *Proceedings of the National Academy of Sciences of the United States of America* 98, 13554–13559.
- Kopan, R., Ilagan, M.X., 2004. Gamma-secretase: Proteasome of the membrane? *Nature Reviews Molecular Cell Biology* 5, 499–504.
- Kopan, R., Ilagan, M.X., 2009. The canonical Notch signaling pathway: Unfolding the activation mechanism. *Cell* 137, 216–233.
- Kornilova, A.Y., Bihel, F., Das, C., Wolfe, M.S., 2005. The initial substrate-binding site of gamma-secretase is located on presenilin near the active site. *Proceedings of the National Academy of Sciences of the United States of America* 102, 3230–3235.
- Kreft, A.F., Martone, R., Porte, A., 2009. Recent advances in the identification of gamma-secretase inhibitors to clinically test the Abeta oligomer hypothesis of Alzheimer's disease. *Journal of Medicinal Chemistry* 52, 6169–6188.
- Kuhn, P.H., Koroniak, K., Hogg, S., *et al.*, 2012. Secretome protein enrichment identifies physiological BACE1 protease substrates in neurons. *EMBO Journal* 31, 3157–3168.
- Kukar, T., Prescott, S., Eriksen, J.L., *et al.*, 2007. Chronic administration of R-flurbiprofen attenuates learning impairments in transgenic amyloid precursor protein mice. *BMC Neuroscience* 8, 54.
- L'hernault, S.W., Arduengo, P.M., 1992. Mutation of a putative sperm membrane protein in *Caenorhabditis elegans* prevents sperm differentiation but not its associated meiotic divisions. *Journal of Cell Biology* 119, 55–68.
- Laudon, H., Hansson, E.M., Melen, K., *et al.*, 2005. A nine-transmembrane domain topology for presenilin 1. *Journal of Biological Chemistry* 280, 35352–35360.
- Lazarov, V.K., Fraering, P.C., Ye, W., *et al.*, 2006. Electron microscopic structure of purified, active gamma-secretase reveals an aqueous intramembrane chamber and two pores. *Proceedings of the National Academy of Sciences of the United States of America* 103, 6889–6894.
- Levitan, D., Greenwald, I., 1995. Facilitation of lin-12-mediated signaling by sel-12, a *Caenorhabditis elegans* S182 Alzheimer's disease gene. *Nature* 377, 351–354.
- Levy-Lahad, E., Wijsman, E.M., Nemens, E., *et al.*, 1995. A familial Alzheimer's disease locus on chromosome 1. *Science* 269, 970–973.
- Li, X., Greenwald, I., 1997. HOP-1, a *Caenorhabditis elegans* presenilin, appears to be functionally redundant with SEL-12 presenilin and to facilitate LIN-12 and GLP-1 signaling. *Proceedings of the National Academy of Sciences of the United States of America* 94, 12204–12209.
- Li, Y., Lu, S.H., Tsai, C.J., *et al.*, 2014. Structural interactions between inhibitor and substrate docking sites give insight into mechanisms of human PS1 complexes. *Structure* 22, 125–135.
- Li, Y.M., Lai, M.T., Xu, M., *et al.*, 2000a. Presenilin 1 is linked with gamma-secretase activity in the detergent solubilized state. *Proceedings of the National Academy of Sciences of the United States of America* 97, 6138–6143.
- Li, Y.M., Xu, M., Lai, M.T., *et al.*, 2000b. Photoactivated gamma-secretase inhibitors directed to the active site covalently label presenilin 1. *Nature* 405, 689–694.
- Lichtenthaler, S.F., Dominguez, D.I., Westmeyer, G.G., *et al.*, 2003. The cell adhesion protein P-selectin glycoprotein ligand-1 is a substrate for the aspartyl protease BACE1. *Journal of Biological Chemistry* 278, 24, 24.
- Lin, X., Koelsch, G., Wu, S., *et al.*, 2000. Human aspartic protease memapsin 2 cleaves the beta-secretase site of beta-amyloid precursor protein. *Proceedings of the National Academy of Sciences of the United States of America* 97, 1456–1460.
- Lu, P., Bai, X., Ma, D., *et al.*, 2014. Three-dimensional structure of human gamma-secretase. *Nature* 512, 166–170.
- Luo, Y., Bolon, B., Kahn, S., *et al.*, 2001. Mice deficient in BACE1, the Alzheimer's beta-secretase, have normal phenotype and abolished beta-amyloid generation. *Nature Neuroscience* 4, 231–232.
- Marambaud, P., Wen, P.H., Dutt, A., *et al.*, 2003. A CBP binding transcriptional repressor produced by the PS1/epsilon-cleavage of N-cadherin is inhibited by PS1 FAD mutations. *Cell* 114, 635–645.
- May, P.C., Dean, R.A., Lowe, S.L., *et al.*, 2011. Robust central reduction of amyloid-beta in humans with an orally available, non-peptidic beta-secretase inhibitor. *Journal of Neuroscience* 31, 16507–16516.
- McConlogue, L., Buttini, M., Anderson, J.P., *et al.*, 2007. Partial reduction of BACE1 has dramatic effects on Alzheimer plaque and synaptic pathology in APP transgenic mice. *Journal of Biological Chemistry* 282, 26326–26334.
- Meredith Jr., J.E., Thompson, L.A., Toyn, J.H., *et al.*, 2008. P-glycoprotein efflux and other factors limit brain amyloid beta reduction by beta-site amyloid precursor protein-cleaving enzyme 1 inhibitors in mice. *Journal of Pharmacology and Experimental Therapeutics* 326, 502–513.
- Mitani, Y., Yarimizu, J., Saita, K., *et al.*, 2012. Differential effects between gamma-secretase inhibitors and modulators on cognitive function in amyloid precursor protein-transgenic and nontransgenic mice. *Journal of Neuroscience* 32, 2037–2050.
- Moore, C.L., Leatherwood, D.D., Diehl, T.S., Selkoe, D.J., Wolfe, M.S., 2000. Difluoro ketone peptidomimetics suggest a large S1 pocket for Alzheimer's γ -secretase: Implications for inhibitor design. *Journal of Medicinal Chemistry* 43, 3434–3442.
- Nikolaev, A., McLaughlin, T., O'leary, D.D., Tessier-Lavigne, M., 2009. APP binds DR6 to trigger axon pruning and neuron death via distinct caspases. *Nature* 457, 981–989.
- Nishitomi, K., Sakaguchi, G., Horikoshi, Y., *et al.*, 2006. BACE1 inhibition reduces endogenous Abeta and alters APP processing in wild-type mice. *Journal of Neurochemistry* 99, 1555–1563.
- Ogura, T., Mio, K., Hayashi, I., *et al.*, 2006. Three-dimensional structure of the gamma-secretase complex. *Biochemical and Biophysical Research Communications* 343, 525–534.
- Oh, Y.S., Turner, R.J., 2005. Topology of the C-terminal fragment of human presenilin 1. *Biochemistry* 44, 11821–11828.
- Ohno, M., Sametsky, E.A., Younkin, L.H., *et al.*, 2004. BACE1 deficiency rescues memory deficits and cholinergic dysfunction in a mouse model of Alzheimer's disease. *Neuron* 41, 27–33.
- Okochi, M., Tagami, S., Yanagida, K., *et al.*, 2013. gamma-Secretase modulators and presenilin 1 mutants act differently on presenilin/gamma-secretase function to cleave Abeta42 and Abeta43. *Cell Report* 3, 42–51.
- Osenkowski, P., Li, H., Ye, W., *et al.*, 2009. Cryoelectron microscopy structure of purified gamma-secretase at 12 Å resolution. *Journal of Molecular Biology* 385, 642–652.
- Qi-Takahara, Y., Morishima-Kawashima, M., Tanimura, Y., *et al.*, 2005. Longer forms of amyloid beta protein: Implications for the mechanism of intramembrane cleavage by gamma-secretase. *Journal of Neuroscience* 25, 436–445.
- Ratovitski, T., Slunt, H.H., Thinakaran, G., *et al.*, 1997. Endoproteolytic processing and stabilization of wild-type and mutant presenilin. *Journal of Biological Chemistry* 272, 24536–24541.
- Renzi, F., Zhang, X., Rice, W.J., *et al.*, 2011. Structure of gamma-secretase and its trimeric pre-activation intermediate by single-particle electron microscopy. *Journal of Biological Chemistry* 286, 21440–21449.
- Rogaev, E.I., Sherrington, R., Rogaeva, E.A., *et al.*, 1995. Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature* 376, 775–778.
- Sankaranarayanan, S., Holahan, M.A., Colussi, D., *et al.*, 2009. First demonstration of cerebrospinal fluid and plasma A beta lowering with oral administration of a beta-site amyloid precursor protein-cleaving enzyme 1 inhibitor in

- nonhuman primates. *Journal of Pharmacology and Experimental Therapeutics* 328, 131–140.
- Sardi, S.P., Murtie, J., Koirala, S., Patten, B.A., Corfas, G., 2006. Presenilin-dependent ErbB4 nuclear signaling regulates the timing of astrogenesis in the developing brain. *Cell* 127, 185–197.
- Sato, C., Morohashi, Y., Tomita, T., Iwatsubo, T., 2006. Structure of the catalytic pore of gamma-secretase probed by the accessibility of substituted cysteines. *Journal of Neuroscience* 26, 12081–12088.
- Sato, C., Takagi, S., Tomita, T., Iwatsubo, T., 2008. The C-terminal PAL motif and transmembrane domain 9 of presenilin 1 are involved in the formation of the catalytic pore of the gamma-secretase. *Journal of Neuroscience* 28, 6264–6271.
- Sato, T., Diehl, T.S., Narayanan, S., *et al.*, 2007. Active gamma-secretase complexes contain only one of each component. *Journal of Biological Chemistry* 282, 33985–33993.
- Saunders, A.J., Kim, T.-W., Tanzi, R.E., *et al.*, 1999. BACE maps to chromosome 11 and a BACE homolog, BACE2, resides in the obligate Down syndrome region of chromosome 21. *Science* 286, 1255a.
- Savonenko, A.V., Melnikova, T., Laird, F.M., *et al.*, 2008. Alteration of BACE1-dependent NRG1/ErbB4 signaling and schizophrenia-like phenotypes in BACE1-null mice. *Proceedings of the National Academy of Sciences of the United States of America* 105, 5585–5590.
- Scheuner, D., Eckman, C., Jensen, M., *et al.*, 1996. Secreted amyloid beta-protein similar to that in the senile plaques of Alzheimer's disease is increased *in vivo* by the presenilin 1 and 2 and APP mutations linked to familial Alzheimer's disease. *Nature Medicine* 2, 864–870.
- Schroeter, E.H., Kisslinger, J.A., Kopan, R., 1998. Notch-1 signaling requires ligand-induced proteolytic release of intracellular domain. *Nature* 393, 382–386.
- Shah, S., Lee, S.F., Tabuchi, K., *et al.*, 2005. Nicastrin functions as a gamma-secretase-substrate receptor. *Cell* 122, 435–447.
- Shearman, M.S., Behr, D., Clarke, E.E., *et al.*, 2000. L-685,458, an aspartyl protease transition state mimic, is a potent inhibitor of amyloid beta-protein precursor gamma-secretase activity. *Biochemistry* 39, 8698–8704.
- Shen, J., Bronson, R.T., Chen, D.F., *et al.*, 1997. Skeletal and CNS defects in presenilin-1-deficient mice. *Cell* 89, 629–639.
- Sherrington, R., Rogaev, E.I., Liang, Y., *et al.*, 1995. Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 375, 754–760.
- Sinha, S., Anderson, J.P., Barbour, R., *et al.*, 1999. Purification and cloning of amyloid precursor protein beta-secretase from human brain. *Nature* 402, 537–540.
- Sperling, R.A., Jack Jr., C.R., Aisen, P.S., 2011. Testing the right target and right drug at the right stage. *Science Translational Medicine* 3, 111cm33.
- Steiner, H., Winkler, E., Haass, C., 2008. Chemical crosslinking provides a model of the gamma-secretase complex subunit architecture and evidence for close proximity of the C-terminal fragment of presenilin with APH-1. *Journal of Biological Chemistry* 283, 34677–34686.
- Takagi, S., Tominaga, A., Sato, C., Tomita, T., Iwatsubo, T., 2010. Participation of transmembrane domain 1 of presenilin 1 in the catalytic pore structure of the {gamma}-secretase. *Journal of Neuroscience* 30, 15943–15950.
- Takami, M., Nagashima, Y., Sano, Y., *et al.*, 2009. gamma-Secretase: Successive tripeptide and tetrapeptide release from the transmembrane domain of beta-carboxyl terminal fragment. *Journal of Neuroscience* 29, 13042–13052.
- Takasugi, N., Tomita, T., Hayashi, I., *et al.*, 2003. The role of presenilin cofactors in the gamma-secretase complex. *Nature* 422, 438–441.
- Tanzi, R.E., Bertram, L., 2005. Twenty years of the Alzheimer's disease amyloid hypothesis: A genetic perspective. *Cell* 120, 545–555.
- Thinakaran, G., Borchelt, D.R., Lee, M.K., *et al.*, 1996. Endoproteolysis of presenilin 1 and accumulation of processed derivatives *in vivo*. *Neuron* 17, 181–190.
- Tolia, A., Chavez-Gutierrez, L., De Strooper, B., 2006. Contribution of presenilin transmembrane domains 6 and 7 to a water-containing cavity in the gamma-secretase complex. *Journal of Biological Chemistry* 281, 27633–27642.
- Tolia, A., Horre, K., De Strooper, B., 2008. Transmembrane domain 9 of presenilin determines the dynamic conformation of the catalytic site of gamma-secretase. *Journal of Biological Chemistry* 283, 19793–19803.
- Tomita, T., Maruyama, K., Saido, T.C., *et al.*, 1997. The presenilin 2 mutation (N141I) linked to familial Alzheimer disease (Volga German families) increases the secretion of amyloid beta protein ending at the 42nd (or 43rd) residue. *Proceedings of the National Academy of Sciences of the United States of America* 94, 2025–2030.
- Turner 3rd, R.T., Loy, J.A., Nguyen, C., *et al.*, 2002. Specificity of memapsin 1 and its implications on the design of memapsin 2 (beta-secretase) inhibitor selectivity. *Biochemistry* 41, 8742–8746.
- Vassar, R., Bennett, B.D., Babu-Khan, S., *et al.*, 1999. beta-Secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science* 286, 735–741.
- Weggen, S., Eriksen, J.L., Das, P., *et al.*, 2001. A subset of NSAIDs lower amyloidogenic Abeta42 independently of cyclooxygenase activity. *Nature* 414, 212–216.
- Weidemann, A., Eggert, S., Reinhard, F.B., *et al.*, 2002. A novel var epsilon-cleavage within the transmembrane domain of the Alzheimer amyloid precursor protein demonstrates homology with Notch processing. *Biochemistry* 41, 2825–2835.
- Wilcock, G.K., Black, S.E., Hendrix, S.B., *et al.*, 2008. Efficacy and safety of tarenflurbil in mild to moderate Alzheimer's disease: A randomised phase II trial. *Lancet Neurology* 7, 483–493.
- Willem, M.S., Garratt, A.N., Novak, B., *et al.*, 2006. Control of peripheral nerve myelination by the beta-secretase BACE1. *Science* 314, 664–666.
- Wolfe, M.S., 2009. Intramembrane proteolysis. *Chemical Reviews* 109, 1599–1612.
- Wolfe, M.S., Citron, M., Diehl, T.S., *et al.*, 1998. A substrate-based difluoro ketone selectively inhibits Alzheimer's γ -secretase activity. *Journal of Medicinal Chemistry* 41, 6–9.
- Wolfe, M.S., Guenette, S.Y., 2007. APP at a glance. *Journal of Cell Science* 120, 3157–3161.
- Wolfe, M.S., Xia, W., Moore, C.L., *et al.*, 1999a. Peptidomimetic probes and molecular modeling suggest Alzheimer's γ -secretases are intramembrane-cleaving aspartyl proteases. *Biochemistry* 38, 4720–4727.
- Wolfe, M.S., Xia, W., Ostaszewski, B.L., *et al.*, 1999b. Two transmembrane aspartates in presenilin-1 required for presenilin endoproteolysis and γ -secretase activity. *Nature* 398, 513–517.
- Wong, P.C., Zheng, H., Chen, H., *et al.*, 1997. Presenilin 1 is required for Notch1 and Dll1 expression in the paraxial mesoderm. *Nature* 387, 288–292.
- Xia, W., Zhang, J., Kholodenko, D., *et al.*, 1997. Enhanced production and oligomerization of the 42-residue amyloid beta-protein by Chinese hamster ovary cells stably expressing mutant presenilins. *Journal of Biological Chemistry* 272, 7977–7982.
- Yan, R., Bienkowski, M.J., Shuck, M.E., *et al.*, 1999. Membrane-anchored aspartyl protease with Alzheimer's disease beta-secretase activity. *Nature* 402, 533–537.
- Yu, G., Chen, F., Levesque, G., *et al.*, 1998. The presenilin 1 protein is a component of a high molecular weight intracellular complex that contains beta-catenin. *Journal of Biological Chemistry* 273, 16470–16475.
- Zhang, Z., Nadeau, P., Song, W., *et al.*, 2000. Presenilins are required for gamma-secretase cleavage of beta-APP and transmembrane cleavage of Notch-1. *Nature Cell Biology* 2, 463–465.
- Zhou, L., Barao, S., Laga, M., *et al.*, 2012. The neural cell adhesion molecules L1 and CHL1 are cleaved by BACE1 protease *in vivo*. *Journal of Biological Chemistry* 287, 25927–25940.

Further Reading

- Cole, S.L., Vassar, R., 2008. The role of APP processing by BACE1, the beta-secretase, in Alzheimer's disease pathophysiology. *Journal of Biological Chemistry* 283, 29621–29625.
- De Strooper, B., Iwatsubo, T., Wolfe, M.S., 2012. Presenilins and γ -secretase: Structure, function, and role in Alzheimer disease. *Cold Spring Harbor Perspectives in Medicine* 2, a006304.

Relevant Websites

- www.alz.org
Alzheimer's Association.
- www.alzforum.org
Alzheimer's Research Forum.

The Calpain Proteolytic System

Y Luo, National Institute of Infectious Diseases, Tokyo, Japan

DF Sellitti, Uniformed Services University of the Health Sciences, Bethesda, MD, USA

K Suzuki, National Institute of Infectious Diseases, Tokyo, Japan

© 2016 Elsevier Inc. All rights reserved.

Introduction

Calpains are a distinct group of ubiquitously expressed eukaryotic cysteine proteases dependent upon Ca^{2+} and a neutral pH for their enzymatic activity. Unlike many other cysteine proteases, they are located in the cytoplasm rather than in membrane-limited compartments such as lysosomes, allowing their enzymatic products to participate in various cellular processes (Sorimachi *et al.*, 2011). Together with calpastatin, the only known endogenous inhibitor of calpain, this family of structurally similar proteases constitutes the 'calpain proteolytic system.'

The calpain proteolytic system is distinct from other intracellular proteolytic systems, such as the ubiquitin-proteasome system (Tanaka, 2009), the autophagy-lysosome system (Mizushima and Levine, 2010), and caspases (Tait and Green, 2010). The ubiquitin/proteasome and autophagy/lysosome systems recognize their substrates only after those substrates are tagged, either by ubiquitinylation, or incorporation into membrane-delimited autophagosomes, respectively. These two systems essentially operate as 'scavengers' to degrade and eliminate unneeded and damaged proteins from the cell. Caspases, on the other hand, are activated by sequence-specific substrate recognition signals, and rather than engaging directly in the elimination of worn proteins, they play important roles in cascades leading ultimately to apoptosis. Moreover, in contrast to caspases, the calpains recognize the tertiary structure of protein substrates rather than specific amino acid sequences.

The most important distinction between the calpains and other proteolytic systems is that whereas the others function in the degradation or elimination of intracellular proteins, calpain proteolysis acts to modify the structures and activities of their substrates, rather than destroying them. The altered calpain substrates are then available for participation in intracellular signaling. Thus this unique enzyme system has acquired the designation as an intracellular 'modulator' protease system (Ono *et al.*, 2004).

To date, the scope of revealed calpain action in cells includes basic processes like cell cycle progression, apoptosis, cell adhesion, and cell mobility, in addition to a number of cell-type specific functions. Furthermore, several specific pathologies have been linked to mutations and dysfunction of specific calpain subtypes, suggesting that novel treatments for these diseases may in the future employ the inactivation of the calpain proteolytic system with specific calpain inhibitors.

Structure of Conventional Calpains

The most extensively studied calpains are the two major ubiquitous mammalian calpains: calpain 1 and calpain 2, often

referred to as the 'conventional' calpains. All other calpains are referred to as 'unconventional' calpains, and their structure is compared to that of the conventional calpains as the standard reference for further classification into 'classical' and 'non-classical' forms (Figure 1).

The conventional calpains, μ - and m-calpain, are heterodimers consisting of a large catalytic subunit (CAPN1/ μ CL and CAPN2/mCL, for μ - and m-calpain, respectively) and a common small regulatory subunit (CAPNS1/30K) unique to the conventional calpains (Figure 1). Calpain 1 and calpain 2 are the heterodimers of CAPN1/ μ CL or CAPN2/mCL and CAPNS1/30K, respectively. The large subunits CAPN1/ μ CL and CAPN2/mCL possess approximately 60% amino acid identity, and therefore have almost indistinguishable substrate and inhibitor specificities. However, as reflected in their designations as μ CL ('micromolar' CL) and mCL ('millimolar' CL) respectively, these subunits differ greatly in their requirement for Ca^{2+} . Thus, μ CL activation requires Ca^{2+} in the micromolar range, whereas mCL requires Ca^{2+} at 1000-fold greater concentration. The functions of calpains 1 and 2 are fundamental and essential to the cell, and have been implicated in numerous biological processes, including the regulation of signal transduction (Pontremoli and Melloni, 1988; Noguchi *et al.*, 1997), cell motility (Leloup *et al.*, 2010; Wells *et al.*, 2005); membrane repair (Mellgren and Huang, 2007; Mellgren *et al.*, 2007), and apoptosis (Danial and Korsmeyer, 2004; Nakagawa and Yuan, 2000), although their precise roles in these events remain elusive.

The catalytic and regulatory subunits of the conventional calpains are each divided into several domains (Figure 1). The catalytic subunit is comprised of four distinct domains: an N-terminal anchor helix, a CysPc protease domain, a C2 domain-like (C2L) region and a penta-EF-hand large (PEF(L)) domain. The regulatory subunit CAPNS1 consists of two domains: an N-terminal Gly-rich (GR) domain and a penta-EF-hand small (PEF(S)) domain (Figure 1). The two PEF domains, PEF(L) and PEF(S), each have five EF-hand motifs, and the fifth motif in each contributes to the heterodimerization of the catalytic and regulatory subunits (Imajoh *et al.*, 1987).

The N-terminal anchor helix in the catalytic subunit interacts with the second EF-hand motif (EF-2) of the PEF(S) domain of the regulatory subunit CAPNS1. Upon activation of the calpain complex, the interaction of the catalytic and regulatory units is disrupted by either autolysis of N-terminal anchor helix or by binding of Ca^{2+} to the EF-2 motif. Autolysis of N-terminal anchor helix upon activation by Ca^{2+} enables the enzyme to function at a lower Ca^{2+} concentration (Cong *et al.*, 1989; Hathaway *et al.*, 1982; Inomata *et al.*, 1988; Suzuki *et al.*, 1981), and with different substrate specificity from before. Therefore, autolysis of the N-terminal anchor helix domain is considered to be an important regulatory mechanism for the activity of conventional calpains.

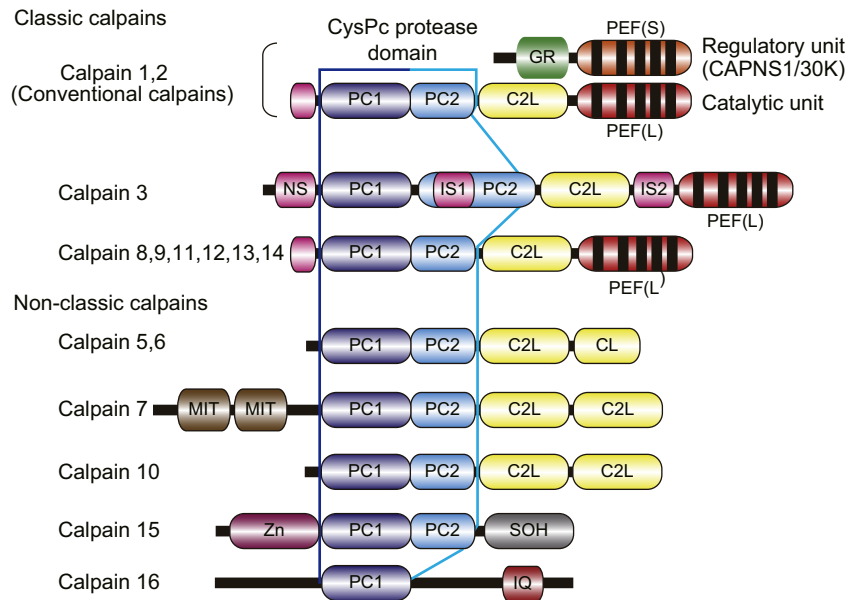


Figure 1 Schematic structures of human calpains. The two major ubiquitous human calpains: calpain 1 and calpain 2, are often referred to as the conventional calpains, while all others are referred to as the unconventional calpains. Human calpains are classified into classic and non-classic types according to their structural similarity compared to the conventional calpains, calpain 1 and calpain 2 (Sorimachi *et al.*, 2011). The structure of conventional calpains consists of two subunits: the catalytic and regulatory subunits. The catalytic subunit consists of four distinct domains: an *N*-terminal anchor helix, a CysPc protease domain, a C2 domain-like (C2L) domain and a penta-EF-hand (PEF) large domain. The regulatory subunit CAPNS1 consists of two domains: an *N*-terminal Gly-rich domain and a PEF small domain. Non-classical calpains lack C2L and PEF domains and instead often have characteristic functional domain(s) of their own. The structural domains are indicated by their abbreviation: PC1 (protease core domain1), PC2 (protease core domain2), C2L (C2 domain-like), PEF(L) (penta-EF-hand large domain), PEF(S) (penta-EF-hand small domain), GR (*N*-terminal Gly-rich domain), NS/IS1/IS2 (three unique insertions in calpain 3), C2 (C2 domain), MIT (microtubule interacting and transport motif), Zn (Zn finger containing domain), SOH (SOL homology domain), IQ (calmodulin interacting motif).

Classification of Calpains

Since the conventional calpain catalytic subunits (CAPN1/ μ CL and CAPN2/mCL) in the human serve as the reference for describing the structure of other calpains, calpains with a domain structure similar to the conventional calpains are called ‘classical’ calpains. In these calpains, the CysPc protease domain of the catalytic unit is followed by the C2L and PEF domains. In contrast, non-classical calpains lack the C2L and/or the PEF domains (Sorimachi *et al.*, 1993; Figure 1). Instead, they may possess functional domain(s) of their own. The structural domains are indicated by their abbreviation: PC1 (protease core domain1), PC2 (protease core domain2), C2L (C2 domain-like), PEF(L) (penta-EF-hand large domain), PEF(S) (penta-EF-hand small domain), GR (*N*-terminal Gly-rich domain), NS/IS1/IS2 (three unique insertions in calpain 3), C2 (C2 domain), MIT (microtubule interacting and transport motif), Zn (Zn finger containing domain), SOH (SOL homology domain), IQ (calmodulin interacting motif) (Figure 1). The nine human classical calpains are calpains 1, 2, 3, 8, 9, 11, 12, 13, and 14 (Macqueen *et al.*, 2010; Figure 1). Although they share strong similarities in amino acid sequence, their function and tissue distribution can differ greatly from each other (Table 1).

The expression of six human calpain genes, including CAPN3, 6, 8, 9, 11, and 12, are tissue-specific, while the others, including those encoding the conventional calpains, are expressed ubiquitously (Table 1). In the human, CAPN3 is expressed in skeletal muscle (Sorimachi *et al.*, 1993), CAPN6 in the placenta and embryonic muscles (Dear *et al.*, 1997), CAPN8 and CAPN9 in the gastrointestinal tract (Sorimachi *et al.*, 1993), CAPN11 in the testis (Sorimachi

et al., 1993), and CAPN12 in hair follicles (Dear *et al.*, 2000), each having a specific intracellular function or functions reflecting its tissue-specific expression. In contrast, the ubiquitous calpains are thought to play fundamental roles in the cell. The classification of calpains as having primarily ‘tissue-specific’ or ‘fundamental’ activities has been supported by the observation of several kinds of calpain gene knockout mouse models. Defects seen in *Capn2*^{-/-} and calpain small subunit *Capns1*^{-/-} mice can be embryonic lethal (Arthur *et al.*, 2000; Dutt *et al.*, 2006; Shimada *et al.*, 2008; Takano *et al.*, 2011), whereas defects in tissue-specific calpains usually cause tissue-specific phenotypes. The roles of calpains in different systems *in vivo* will be further reviewed in the following articles.

Regulation and Activation of Calpains

The Smaller Regulatory Subunit, CAPNS1/30K

CAPNS1/30 K is an important and unique regulatory subunit for conventional calpains. It has been reported that CAPNS1/30 K is indispensable for the stability and activation of conventional calpains. In *Capns1*^{-/-} mice, the expression of both subunits of CAPN1/ μ CL and CAPN2/mCL are almost completely down-regulated resulting in embryonic lethality of almost all mice before E11.5 (Arthur *et al.*, 2000; Zimmerman *et al.*, 2000). Regarding the role of the regulatory unit in

Table 1 The expression distribution and classification of human calpains

<i>Protein</i>	<i>Gene</i>	<i>Tissue distribution</i>	<i>Classification</i>	<i>Physiological function</i>	<i>Pathological involvement</i>
μ -Calpain large subunit	<i>CAPN1</i>	Ubiquitous	Classical	Normal central nervous system functions, normal cardiovascular functions ^a	Central nervous system diseases, cardiovascular diseases ^a
m-Calpain large subunit	<i>CAPN2</i>	Ubiquitous	Classical	Normal central nervous system functions, normal cardiovascular functions ^a	Central nervous system diseases, cardiovascular diseases ^a
Calpain 3/p94	<i>CAPN3</i>	Skeletal muscle	Classical	Normal skeletal muscle functions	Limb-girdle muscular dystrophy type 2A (LGMD2A)
Calpain 5/hTRA-3	<i>CAPN5</i>	Ubiquitous	Non-classical		
Calpain 6	<i>CAPN6</i>	Placenta and embryonic muscles	Non-classical		
Calpain 7/palBH	<i>CAPN7</i>	Ubiquitous	Non-classical		
Calpain 8/nCL-2	<i>CAPN8</i>	Stomach mucosa and gastrointestinal track	Classical	Gastric mucosal defense	Susceptibility to gastric ulcers
Calpain 9/nCL-4	<i>CAPN9</i>	Stomach mucosa and gastrointestinal track	Classical	Gastric mucosal defense	Susceptibility to gastric ulcers
Calpain 10	<i>CAPN10</i>	Ubiquitous	Non-classical		Susceptibility to non-insulin-dependent diabetes mellitus (NIDDM)
Calpain 11	<i>CAPN11</i>	Testis	Classical		
Calpain 12	<i>CAPN12</i>	Hair follicle	Classical		
Calpain 13	<i>CAPN13</i>	Ubiquitous	Classical		
Calpain 14	<i>CAPN14</i>	Ubiquitous	Classical		
Calpain 15/SOLH	<i>CAPN15</i>	Ubiquitous	Non-classical		
Calpain 16/demi-calpain	<i>CAPN16</i>	Ubiquitous	Non-classical		
Calpain small subunit 1	<i>CAPNS1</i>	Ubiquitous	Regulatory unit		
Calpain small subunit 2	<i>CAPNS2</i>	Ubiquitous	Regulatory unit		

^aAlso involved in variety of physiological functions and human diseases, as conventional calpains are fundamental regulators of many physiological and pathological functions.

activation of calpains, it has been shown that it is very difficult to activate CAPN2/mCL in the absence of CAPNS1/30K *in vitro* (Yoshizawa *et al.*, 1995). So far, CAPNS1/30K has only been shown to be necessary for activation of conventional calpains.

Ca²⁺: The Main Calpain Activator

Since the activation of conventional calpains is Ca²⁺-dependent, numerous studies have been aimed at identifying the location of the Ca²⁺-binding site responsible for calpain activation. 3D structure analysis has revealed that the active protease domain is formed by the fusion of core domain PC1 and PC2 upon the binding of one Ca²⁺ to each of the core domains. Thus, the activation of the conventional calpains entails the binding of Ca²⁺ to the protease domain itself (Moldoveanu *et al.*, 2002; Moldoveanu *et al.*, 2003).

As mentioned above, the activation of conventional calpains requires a high Ca²⁺ concentration (in the range of 10 mM or greater) that is seldom available *in vivo*, so the

means by which the conventional calpains are activated in the presence of low physiologic Ca²⁺ concentrations is another topic of intense study. One proposed mechanism suggests that specific intracellular molecules such as membrane phospholipids could lower the required Ca²⁺ concentration for calpain activation (Tomba *et al.*, 2001; Shao *et al.*, 2006). Another possibility is that a very small number of calpain molecules localized in a small region with a high local Ca²⁺ concentration are sufficient to excise calpain function. The neuromuscular junction (NMJ) where Ca²⁺ and Na⁺ undergo rapid, massive and localized changes in concentration is believed to be one such region favorable for calpain activation. In fact, the abnormal activation of calpains in the post-synaptic muscle fiber of the NMJ is believed to be one of the mechanisms underlying the development of slow-channel myasthenic syndrome, an inherited neuromuscular disorder (Groshong *et al.*, 2007). In this disease, Ca²⁺ overload occurs at the NMJ due to mutations in the genes encoding nicotinic acetylcholine receptor in the post-synaptic (muscle cell) membrane. The elevated Ca²⁺ at the synapse, in turn, could

result in the abnormal activation of Ca^{2+} -dependent processes including calpain activation. This hypothesis is supported by studies showing that the transgenic overexpression of calpastatin ameliorates the symptoms of slow-channel myasthenic syndrome in a mouse model of this disease.

Calpastatin: The Endogenous Calpain Inhibitor

Calpastatin is the only known endogenous inhibitor of calpains (Kiss *et al.*, 2008). Although calpastatin inhibits mainly the conventional calpains, it has been reported that calpastatin also suppresses the activity of several unconventional calpains *in vitro*, including calpain 8 (Hata *et al.*, 2007) and calpain 9 (Lee *et al.*, 1999). Furthermore, although calpastatin does not affect calpain 3 activity *per se*, calpain 3 can inactivate calpastatin through partial proteolysis, permitting a regulatory effect of calpain 3 on calpastatin and conventional calpains in skeletal muscle (Ono *et al.*, 2004).

The calpastatin molecule is divided into four inhibitor units, each of which inhibits one calpain molecule with variable efficiency (Emori *et al.*, 1987; Maki *et al.*, 1987a; Maki *et al.*, 1987b). Oligopeptides as short as 20 amino acids derived from these inhibitory units can also inhibit calpain, although with reduced efficacy compared with the whole molecule. 3D structure analysis of the co-crystallization of calpastatin and CAPN2/mCL with Ca^{2+} has revealed one of the mechanisms underlying the inhibitory effect of calpastatin on conventional calpains. When calpastatin interacts with CAPN2/mCL in the presence of Ca^{2+} , calpastatin develops a kink between its flanking residues, L173/L612 on one side of the catalytic cleft between PC1 and PC2 in CAPN2/mCL, and T179/T618 at the other side. This kink allows several adjacent amino acid residues in calpastatin to loop out from the proteolytic active site on CAPN2/mCL (Hanna *et al.*, 2008; Moldoveanu *et al.*, 2008), permitting calpastatin to bind CAPN2/mCL tightly without being proteolyzed. Calpastatin mutagenesis studies have confirmed the importance of this 'kink' in the inhibitory effect of calpastatin.

Unique Regulation of Unconventional Calpains

As described above, conventional mammalian calpains exhibit the following basic features: (1) CAPN1/30K forms heterodimers with CAPN1/ μ CL or CAPN2/mCL to maintain calpain stability and activity, (2) calpastatin exerts inhibitory effects on proteolysis by calpain, and (3) Ca^{2+} binds to the protease domain to initiate enzymatic activity. The non-conventional calpains, however, possess some unique regulatory mechanisms not seen in the conventional calpains.

Calpain 3, which is specifically expressed in skeletal muscle in the human, is approximately 50% identical to CAPN1/ μ CL and CAPN2/mCL (Figure 1). Two NS insertions, located at the N-terminus of the PC2 domain, and between C2L and the PEF domains (termed IS1 and IS2, respectively), confer a unique *in vitro* susceptibility of calpain 3 to a rapid (<10 min duration) autodegradation that distinguishes it from all other members of the calpain family (Zimmerman *et al.*, 2000; Ono *et al.*, 2007; Ono *et al.*, 2006). Calpain 3 variants missing either IS1 or IS2 degrade much slowly, indicating that the two

insertions are essential for the rapid degradation of substrates (Sorimachi *et al.*, 1993). Moreover, autodegradation has been shown to occur Na^{+} -dependently and in the absence of Ca^{2+} , thus making calpain 3 the first example of an intracellular Na^{+} -dependent enzyme (Ono *et al.*, 2010). In addition, the substrates of Na^{+} -activated calpain 3 differ from those of the Ca^{2+} -activated enzyme (Ono *et al.*, 2010).

Calpain 8 and calpain 9 are specifically expressed in the gastrointestinal tract and form the heterodimer calpain 8/calpain 9 (G-calpain) *in vivo* (Hata *et al.*, 2010). The formation of this heterodimers has been shown to be essential for the stability of these two calpains. Other calpain subtypes are undoubtedly regulated by as yet undefined mechanisms that are specific to the precise structure and function of each.

Calpains in Apoptosis

Studies showing the involvement of calpains activation in apoptosis are abundant and still growing in number. However, the role of calpains in apoptosis is very complex, and ordinarily dependent on the specific cellular context. Thus, both pro- and anti-apoptotic functions of calpains have been observed, and those functions are now considered to be associated with the pathogenesis of various human diseases. In the following sections, we will summarize various pathways linking calpain activation and apoptosis, as a prelude to describing the role of the calpain proteolytic system in specific human diseases.

Relationship between Calpain and Caspases

Caspases, best known for their essential roles in apoptosis, interact with the calpain proteolytic system in several important ways. Not only have calpains and caspases been shown to exert their enzymatic activity on many of the same substrates (Nakagawa and Yuan, 2000), but they have also been reported to serve as substrates for each other (Waterhouse *et al.*, 1998; Chua *et al.*, 2000).

Furthermore, although calpains usually activate caspases indirectly, through the cleavage and degradation of endogenous caspase inhibitors (Mikosik *et al.*, 2007; Porn-Ares *et al.*, 1998; Wang *et al.*, 1998), the direct regulation of caspases by calpains is also possible (Wood and Newcomb, 1999). As an example, it was demonstrated that activation of calpains resulted in caspase 3 activation and subsequent apoptosis in cells undergoing radiation-induced death (Waterhouse *et al.*, 1998). It was also reported that calpains can activate caspase 7 directly through cleavage at specific sites, both *in vivo* and *in vitro* (Ruiz-Vela *et al.*, 1999), and that they can also activate caspase 12 to initiate stress- and ischemic injury-induced apoptosis (Tan *et al.*, 2006). On the other hand, a study of drug-induced apoptosis in human neuroblastoma cell lines showed that calpains indirectly induce the cleavage of pro-caspase, resulting in permanent caspase inactivation (McGinnis *et al.*, 1999). Moreover, several calpains have been shown to inactivate caspases 7, 8, and 9 via proteolysis at specific cleavage sites in the caspase molecules, and conversely, calpains have been shown to possess a set of specific cleavage sites for caspase activity

(Chua *et al.*, 2000). Thus, the complex, reciprocal relationship between calpains and caspases can result in either pro- or anti-apoptotic effects of the calpains, depending on precisely which cell context, calpain subtypes, caspase subtypes, and enzymatic substrates are involved (Chua *et al.*, 2000; Waterhouse *et al.*, 1998; McCollum *et al.*, 2002).

Other Molecules Involved in Calpain-Mediated Apoptosis

Calpains also mediate apoptosis through their interactions with members of the B cell lymphoma 2 (Bcl 2) protein family, another important regulator of apoptosis whose members regulate apoptosis via changes in outer mitochondrial membrane permeability. Specifically, calpains are known to cleave and activate the Bcl-2 associated X (Bax) protein which is then able to permeabilize the mitochondrion and initiate a cytochrome *c*/caspase cascade leading to cell death (Cao *et al.*, 2003). On the other hand, the Bax fragments produced by calpain cleavage are thought to lose their ability to interact with and block the action of the anti-apoptotic proteins Bcl-2 and Bcl-xL. Thus, the cleavage of Bax by calpains could have both pro- and anti-apoptotic effects whose balance can shift with changes in the intracellular milieu. Moreover, it has been shown that calpain inhibitors can block or at least attenuate Bax-mediated apoptosis in chemotherapy drug-treated cells, suggesting that calpain dysregulation could play a role in the acquisition of chemotherapy resistance. As an example of this, an *in vitro* study has shown that pretreatment of human leukemic cell lines with calpeptin (a calpain inhibitor) blocks the ability of etoposide (a widely used chemotherapy drug) to induce Bax cleavage and subsequent apoptosis (Gao and Dou, 2000). Another substrate of calpain responsible for the initiation of an apoptotic cascade is the BH3 interacting-domain death agonist (Bid) protein, which is known to cause Bax insertion into the outer mitochondrial membrane (Mandic *et al.*, 2002). The cleavage of Bid by calpain leads to mitochondrial permeabilization and apoptosis during myocardial ischemia reperfusion (Mandic *et al.*, 2002).

Drug-Induced Apoptosis and Calpain

Chemotherapy-induced apoptosis has been reported in many studies to involve the activation of conventional calpains. For example, it was reported that activation of calpain 1 contributes to the genistein-induced apoptosis of MCF-7 breast cancer cells (Sergeev, 2004). This process could conceivably occur via the genistein-induced release of Ca^{2+} from the endoplasmic reticulum into the cytoplasm, where it would then bind to and activate calpain to modify its substrate, caspase 7, and initiate an apoptotic cascade (Shim *et al.*, 2007). Similarly, calpain 2 has been shown to be associated with genistein-induced apoptosis in hepatocellular carcinomas via activation of caspase 12 (Yeh *et al.*, 2007). Conventional calpains are also reported to mediate the induction of apoptosis by the widely used chemotherapy drugs cisplatin and oxaliplatin in human lung adenocarcinoma and cervical carcinoma cells, respectively (Liu *et al.*, 2008; Anguissola *et al.*, 2009) and by aspirin and curcumin, respectively, in cervical carcinoma and glioblastoma cells (Lee *et al.*, 2008; Karmakar *et al.*, 2007). The broad range

of anti-tumor drugs dependent on calpain activation, and the large number of tumor types affected suggest that calpains could represent a very promising target for future chemotherapies. On the other hand, a recent study has shown the deregulation of conventional calpains may be involved in the acquisition of chemotherapy resistance. In the report, the sensitivity of breast cancer cells to trastuzumab was reduced by the deregulation of calpain activation, suggesting that trastuzumab resistance in chemotherapy of human breast cancer could be overcome by targeting calpain activity (Kulkarni *et al.*, 2010). The studies described above highlight a potentially very important role of calpains in both chemotherapy of a wide range of tumors, and in overcoming the drug-resistance that often occurs during chemotherapy.

Physiological Functions of Calpains and the Role of Dysregulation in Human Disease

Calpain dysregulation is believed to enzymatically alter many additional intracellular proteins resulting in either excessive protein degradation or accumulation, or changes in function that ultimately lead to the development of a variety of human diseases. In this section, we present an overview of the molecular evidence linking calpain dysregulation to the etiology of those diseases.

Central Nervous System

Conventional calpains are ubiquitously expressed in the brain and contribute to its normal functions (Goll *et al.*, 2003) (Table 1). However, persistent activation of calpains induced by the disruption of calcium homeostasis could result in abnormal elevation in the cleavage of neuronal substrates to cause changes in neuronal structure and function. To date, the sustained activation of calpains has been reported in both acute neuronal damage resulting from trauma (Kampfl *et al.*, 1996) and in several chronic neurodegenerative diseases including Alzheimer's disease, Parkinson's disease and Huntington's disease (Vosler *et al.*, 2008) (Table 1). As a result of these findings, calpain inhibitors have been extensively studied as potential therapeutic agents in the treatment of central nervous system (CNS) diseases.

Calpain activation after experimental traumatic brain injury (TBI) was first reported in a rat model of percussion brain injury, using a combination of immunoblotting and immunohistochemical techniques to detect specific breakdown products of α II-spectrin (Saatman *et al.*, 1996a), a cytoskeletal protein that under normal circumstances provides structural support to the plasma membrane of axons and pre-synaptic terminals. Increased cleavage of α II-spectrin by calpain acts to destabilize additional proteins in the neuron cell membrane, thus making axons more vulnerable to traumatic injury. By assessing α II-spectrin proteolysis as a marker of calpain activation, TBI development has been confirmed to occur as a consequence of calpain activity in commonly used rat, mouse (Saatman *et al.*, 2003) and several other types of *in vivo* TBI models (Hall *et al.*, 2005; Kupina *et al.*, 2001), suggesting that posttraumatic calpain activation may be a key event in the subsequent development of both focal and diffuse brain

injury. Moreover, the specific inhibition of calpain activation in a rat model of brain trauma has been reported to attenuate of α II-spectrin proteolysis (Posmantur *et al.*, 1997; Saatman *et al.*, 1996b), neuronal damage, and behavioral deficits (Saatman *et al.*, 1996b), suggesting the application of calpain inhibitors as a potential therapeutic approach in the treatment of TBI.

Abnormal calpain activation has also been reported as a factor underlying the development of brain damage induced by cerebral ischemia. The mechanism is thought to involve the post-ischemic over-activation of cerebral glutamate receptors, causing a sharp increase of intracellular Ca^{2+} that could reach the threshold for calpain activation. Thus activated calpains could further increase the intracellular Ca^{2+} concentrations through the cleavage of several calcium-handling substrates, such as calcium regulatory proteins in the cell membrane (L-type calcium channels (De Jongh *et al.*, 1994; Hell *et al.*, 1996), plasma membrane calcium ATPase (PMCA) (Pottorf *et al.*, 2006), and the sodium–calcium exchanger (NCX) (Bano *et al.*, 2005)). Additionally, calpain may alter cell survival signals to increase the vulnerability of neuronal cells to apoptosis. One such pro-apoptotic mechanism is the calpain-induced uncoupling of the N-methyl-D-aspartic acid receptor (NMDAR) from cell survival-promoting signal cascades (including the phosphatidylinositol 3 kinases/protein kinase B and protein kinase A pathways), and the subsequent initiation of mitochondrial dysfunction (Gascon *et al.*, 2008). Other mechanisms resulting in neural damage include the modulation of neuronal transcription by calpains through their direct cleavage of specific transcription factors (Sp3 and Sp4) (Mao *et al.*, 2007), or through their interactions with calcineurin (Wu *et al.*, 2007), cyclin-dependent kinase (Rashidian *et al.*, 2005), and collapsin response mediator protein-3 (Hou *et al.*, 2006).

One of the most important areas of research into abnormal calpain activation in human disease is the role of calpains in Alzheimer's disease. Alzheimer's disease is characterized by pathological deposits of amyloid β -protein ($\text{A}\beta$) in senile plaques (Haass and Selkoe, 2007), intracellular neurofibrillary tangles (NFTs) comprised of hyperphosphorylated aggregates of tau protein (Grundke-Iqbal *et al.*, 1986), synaptic dysfunction and neuronal death. Although the pathological mechanisms underlying Alzheimer's disease have not yet been fully explained, convincing evidence shows that disrupted neuronal calcium homeostasis is a key factor. Elevated concentrations of soluble $\text{A}\beta$ produced by the integral membrane protein amyloid precursor protein (APP) are thought to lead to NMDAR activation and an increase of intracellular calcium (De Felice *et al.*, 2007; Demuro *et al.*, 2005). Evidence of increased calpain activation by calcium has been confirmed in the post mortem examination of brains from patients with Alzheimer's disease (Nixon *et al.*, 1994; Saito *et al.*, 1993). Moreover, a recent study suggests that the pathological activation of calpain can induce the cleavage of sodium–calcium exchanger 3 (NCX3) in the patient's brain. Reduced calcium transporter activity of the cleaved NCX3 substrate would be expected to cause the sustained increases of intracellular calcium level that are required to initiate the synaptic and neuronal dysfunction characteristic of the disease (Chen and Fernandez, 2004).

Despite the increasingly clear role of calpains in Alzheimer's disease etiology, treatments for Alzheimer's disease aimed at inhibiting calpains may actually induce some of the suspected causative conditions for this disease. For example, calpain has been proposed as a candidate α -secretase responsible for the normal regulation of APP (Li *et al.*, 1995; Chen and Fernandez, 2005). Likely as a consequence of this function, the inhibition of calpain 1 activity using specific siRNA was seen to decrease overall APP release, but actually increased $\text{A}\beta$ production (Atherton *et al.*, 2013). Effective Alzheimer's therapies would therefore need to preserve the beneficial actions of calpain inhibition, while attempting to limit the potentially damaging effects of elevated $\text{A}\beta$ production that results from it.

Hyperphosphorylation and subsequent aggregation of tau protein is yet another calpain-dependent molecular mechanism contributing to the progression Alzheimer's disease. Calpain is thought to contribute to the hyperphosphorylation of tau through the regulation of kinase substrates, one of which has been identified as mitogen-activated protein kinase Erk 1/2 (Veeranna *et al.*, 2004). An *in vitro* study showed that treatment of primary cultures of neurons with the calcium ionophore ionomycin caused tau hyperphosphorylation via the activation of calpain and the subsequent activation of Erk 1/2. Calpain may also increase kinase activity (Kishimoto *et al.*, 1989) and tau hyperphosphorylation (Lucas *et al.*, 2001) through cleavage of protein kinase C (PKC) and calcium/calmodulin kinase (CamK) II, and by removing the inhibitory domain (GSK-3 β) from glycogen synthase kinase 3 (GSK-3) (Lucas *et al.*, 2001; Goni-Oliver *et al.*, 2007). Active GSK-3 in addition to hyperphosphorylation of tau may contribute to Alzheimer's disease via inducing neuronal death through the β -catenin pathway (Lucas *et al.*, 2001).

$\text{A}\beta$, in addition to increasing calpain activity via Ca^{2+} elevation, has been shown to contribute to the calpain-induced hyperphosphorylation and cleavage of tau into a toxic 17 kDa fragment (Park and Ferreira, 2005) that promotes cell death in neurons. Neuronal damage caused by this fragment is not unique to Alzheimer's disease (Amadoro *et al.*, 2006). The mechanism underlying the toxicity of tau fragments is thought to involve fragment translocation into the nucleus, but the details of the process remain largely unknown.

Calpains have also been identified as playing important pathologic roles in the neurodegenerative diseases of external CNS sensory receptors such as the eye. One such disease is glaucoma, a disease wherein conventional calpains participate in the progressive retinal degeneration that eventually results in vision impairment and blindness. The process is thought to occur as the increased intraocular pressure due to blocked outflow of aqueous humor triggers calcium dyshomeostasis in retinal ganglion cells, which in turn activates calpain to cleave caspase and initiate an apoptotic cascade. It is the progressive loss of retinal ganglion cells in this disease that leads to blindness (Huang *et al.*, 2010). Another mechanism through which calpain dysregulation contributes to glaucoma involves the overexpression of calpain rather than its hyperactivation. The hypothesized mechanism for the elevated calpain levels is the impairment of calpain degradation due to the accumulation of lipid oxidation products, such as iso[4]levuglandin E2, in the cell. As in several other neurodegenerative diseases

in which calpain dysfunction is the suspected cause, the systemic delivery of calpain inhibitors in patients at risk for glaucoma has been proposed as a protective strategy (Govindarajan *et al.*, 2008).

Cardiovascular System

The dysregulation of conventional calpains has been linked to a wide range of cardiovascular disorders (Shao *et al.*, 2006; Iwamoto *et al.*, 1999; Li *et al.*, 2011; Mani *et al.*, 2008; Singh *et al.*, 2004; Sorimachi and Ono, 2012; Table 1), including coronary contractile dysfunction, cardiac ischemia/infarction, hypertension, atherosclerosis, myocardial stunning, and diabetes-associated cardiovascular disease. It has also been reported that many of these conditions are improved or ameliorated by inhibition of calpain activity, suggesting that over-activation of calpains might have a role in causing these diseases.

Mice receiving chronic infusions of angiotensin II or genetically lacking low-density lipoprotein receptor (*Ldlr*) are useful animal models for studying the development and treatment of such cardiovascular disorders as hypertension and atherosclerosis. Transgenic overexpression of calpastatin in chronically angiotensin II-infused mice was found to suppress hypertrophy of the left ventricle and tunica media, and the perivascular inflammation and fibrosis induced by the elevated levels of angiotensin II (Letavernier *et al.*, 2008). Similarly, the endogenous calpain inhibitor was found to ameliorate the atherosclerosis and abdominal aortic aneurysms induced by angiotensin-II infusion in *Ldlr*^{-/-} mice (Subramanian *et al.*, 2012).

Vascular endothelial cadherin (VE-cadherin) is a recognized substrate for calpain 2 expressed in coronary arteries. Miyazaki *et al.* (Miyazaki *et al.*, 2011) have reported that both calpain 2 levels and VE-cadherin proteolysis are increased in aortic endothelial cells from atherosclerotic lesions in human patients and from the *Ldlr*^{-/-} mouse model of atherosclerosis. Treatment with either specific siRNA against *CAPN2*, or with calpain inhibitors was seen to attenuate this condition, suggesting that increased calpain 2-induced VE-cadherin proteolysis represents a potential mechanism underlying the development of atherosclerosis.

Myocardial stunning (also called broken-heart syndrome or transient left ventricular apical ballooning), a form of cardiac dysfunction caused by post-ischemic reperfusion, is another cardiovascular disorder whose etiology involves the over-activation of calpain (Bolli and Marban, 1999). Specifically, myocardial stunning is suspected to be due to the increased proteolysis of the cardiac troponin I subunit (cTnI) by calpain (van der Laarse, 2002; Maekawa *et al.*, 2003). However, relevant studies are still controversial (Colantonio *et al.*, 2004; Galvez *et al.*, 2007; Thomas *et al.*, 1999).

It has also been reported that diabetes-associated cardiovascular complications in animal models are improved by the disruption of conventional calpains. In one investigation, the knockdown of either *Capn1* or both *Capn1* and *Capn2* together was seen to mitigate the myocardial hypertrophy and fibrosis in the mouse model of type 1 diabetes (Li *et al.*, 2011). In other studies, the inhibition of calpain activity mitigated

hyperglycemia in animal models of both type 1 (Smolock *et al.*, 2011) and 2 diabetes (Scalia *et al.*, 2007).

In addition, it has been proposed that calpain-induced over-activation of protein kinase C (PKC) may be a contributory factor in the development of some forms of cardiomyopathy (Kang *et al.*, 2010; Zhang *et al.*, 2011). Normally, PKC is maintained in an inactive state by the presence of a regulatory domain within the N-terminus of the molecule. However, intracellular events leading to PKC proteolysis by the conventional calpains could cut off that regulatory domain to produce a proteolytic fragment of PKC called PKM that remains constitutively active regardless of upstream regulatory signals (Kishimoto *et al.*, 1983). Constant activation of PKM, in turn, would then induce hyperphosphorylation of various substrates, including myosin-binding protein C and histone deacetylase 5 (HDAC5), resulting in morbid cellular responses, such as a constitutive export of HDAC5 from the nucleus that leads to cardiomyopathy.

Calpains in Skeletal Muscle

CAPN3 mutations were shown to be responsible for limb-girdle muscular dystrophy type 2A (LGMD2A) (also called calpainopathy) (Table 1) using a genotypic analysis of families residing in the French-owned island of Réunion and through positional cloning (Richard *et al.*, 1995; Chiannikulchai *et al.*, 1995). The involvement of calpain 3 defects in this disease received additional support from experimental studies showing that *Capn3*^{-/-} mice display a phenotype similar to human LGMD2A (Richard *et al.*, 2000). *Capn3* knock-in (*Capn3*^{CS/CS}) mice, which express a structurally intact calpain 3 mutant lacking proteolytic activity, also display an LGMD2A-like phenotype, indicating that compromised calpain 3 protease activity rather than altered calpain structure is the more important contributor to the pathology of LGMD2A (Ono *et al.*, 1998).

Subsequent analyses of many additional families have so far identified numerous pathogenic and polymorphic mutations in *CAPN3* (Richard *et al.*, 1999; Saenz *et al.*, 2005). However, no mutational hot spot has been claimed for the disease, so it is difficult to make a genetic diagnosis of LGMD2A. A total of 2530 mutations, including 456 unique ones, have been reported in the *CAPN3* gene so far. Among *CAPN3* gene mutations, approximately 62.7% are nucleotide substitutions, of which, 60.3% are missense mutations, and the rest (37.8%) are pathogenic mutations. The association of LGMD2A with *CAPN3* mutations represents the first and so far the only example of a clear cause–effect relationship between human disease and calpain gene mutations (Sorimachi *et al.*, 2011).

Calpains in the Gastrointestinal Tract

Calpains 8 and 9 in the gastrointestinal tract are mainly expressed in surface mucus cells lining the gastric pits in the stomach, and in lesser amounts in the goblet cells of the intestines (Sorimachi *et al.*, 1993; Hata *et al.*, 2006). Calpain 8 is most similar to calpain 2, having a 62% identity in amino acid sequence with the conventional calpain. However, unlike

conventional calpains, the Ca^{2+} -dependent autolytic and catalytic activities of calpain 8 do not require the regulatory subunit, CAPNS1/30K (Hata *et al.*, 2007). Calpain 9, on the other hand, shares an approximately equal amino acid sequence identity with most other calpains, suggesting that this calpain subtype is closest to the ancestral calpain form (Lee *et al.*, 1998). Recombinant human calpain 9 requires CAPNS1/30K for its Ca^{2+} -dependent activity *in vitro*, and this activity can be inhibited by calpastatin and other Cys protease inhibitors (Lee *et al.*, 1999).

The properties of calpain 8 and calpain 9 *in vivo*, however, could differ greatly from those that have been observed *in vitro*. Analysis of *Capn8*^{-/-} and *Capn9*^{-/-} mouse models has revealed that neither calpain 8 nor calpain 9 forms a stable complex with CAPNS1/30K *in vivo*. Instead, they form a hybrid calpain 8/calpain 9 heterodimer referred to as gastric calpain (G-calpain) (Hata *et al.*, 2010). Disruption of either *Capn8* or *Capn9* was observed to down-regulate the other molecule in the pair, suggesting a co-dependency of the two calpains for stability and the proper function. Although neither *Capn8*^{-/-} nor *Capn9*^{-/-} mice display obvious gastric pathologies under normal conditions, the knockout mice are significantly more susceptible to ethanol-induced gastric ulcers (Hata *et al.*, 2010). *Capn8* knock-in (*Capn8*^{CS/CS}) mice expressing a structurally intact but protease-inactive calpain 8 mutant are also more susceptible to stress-induced gastric ulcers compared with wild type mice. The localization of calpain 8 and 9 in mucus-secreting cells, and the association of gastric ulcers with their disruption suggest that these calpains, coupled together to form heterodimeric G-calpain, play important roles in gastric mucosal defense (Hata *et al.*, 2010; Table 1). Additional reported functions for calpain 9 include an involvement in neoplastic transformation and tumorigenesis (Liu *et al.*, 2000), and in adhesion molecule-induced lumen formation in the mammary gland (Chen *et al.*, 2010).

A search of the single nucleotide polymorphism (SNP) database has revealed several CAPN8 and CAPN9 SNPs that result in amino acid substitutions (Hata *et al.*, 2010). Further *in vitro* studies showed that some of these substitutions are associated with the impaired proteolytic activity of G-calpain. Additional studies are needed to further explore the possibility of a relationship between CAPN8 and CAPN9 SNPs and the susceptibility to gastrointestinal disorders.

Calpains in the Endocrine System

Human calpain 10 is a ubiquitously expressed non-classical calpain that lacks PEF(L) domains but instead has two C2L domains in tandem at the C-terminus (Figure 1). In 2000, a large-scale genetic association study using haplotype analysis performed in Mexican Americans and a Northern European population from the Botnia region of Finland reported that a single nucleotide polymorphism (SNP-43) within intron 3 of CAPN10 is a risk factor for type 2 (i.e., non-insulin-dependent) diabetes mellitus, (NIDDM) (Horikawa *et al.*, 2000; Horikawa, 2006; Table 1). Subsequent genetic association studies performed in various populations (Horikawa, 2006; e.g., Indian (Baier *et al.*, 2000; British/Irish (Evans *et al.*, 2001); African-American (Garant *et al.*, 2002); Polish (Malecki *et al.*, 2002);

Finnish/Botnia (Orho-Melander *et al.*, 2002)) confirmed CAPN10 as a susceptibility gene for NIDDM and also reported several tag SNPs (including SNP-43, -44, -19, and -63) within CAPN10 as putative risk factors for NIDDM (Horikawa, 2006). Clinical studies in Pima Indians found that individuals with G/G homozygotes in SNP-43 have a reduced CAPN10 mRNA expression level and decreased rate of oxidative utilization of glucose in their skeletal muscle, indicating that this SNP may increase susceptibility to NIDDM by decreasing the rate of insulin-mediated glucose turnover (Baier *et al.*, 2000). However, the molecular mechanisms underlying the association of the reported tag SNPs within CAPN10 with susceptibility to NIDDM remain unclear.

Conclusion

The calpain system, unlike other eukaryotic proteolytic systems such as autophagolysosomes and proteasomes, engages in only a limited proteolysis of its substrates to create new signaling molecules with specific intracellular functions downstream of calpain activation by calcium. In this way, the calpain system shares a greater similarity with other enzymatically activated signaling systems (such a kinase dependent ones) than it does to mechanisms for complete protein breakdown. With a variety of enzymatic and regulatory subunits coded by separate genes, the calpain family has been shown to play a role in a large number of both very basic cellular processes such as survival, apoptosis, necrosis, migration, adhesion, and in cell type-specific functions. Thus, the inappropriate activation of calpain has been identified or suspected in a wide range of clinical disorders, and the use of specific calpain inhibitors has been proposed to be an effective therapeutic strategy. Nonetheless, many open questions regarding the calpain proteolytic system remain, including the extent of its involvement in numerous cellular processes not described above. Given the ubiquity of the system and its strict regulation by Ca^{2+} and other factors, it is not hard to imagine that calpains will soon be recognized both as critical players in the normal maintenance of cellular function, and as key participants in the progression of diseases when calpain regulation fails.

See also: Cell Communication: Intracellular Pathways: Calcium and Calmodulin Signaling. Cell Division/Death: Apoptosis: Caspases. Protein Synthesis/Degradation: Protein Degradation – Pathological Aspects: Cancer – Proteases in the Progression and Metastasis; Inhibitors of HIV Protease

References

- Amadoro, G., Ciotti, M.T., Costanzi, M., *et al.*, 2006. NMDA receptor mediates tau-induced neurotoxicity by calpain and ERK/MAPK activation. *Proceedings of the National Academy of Sciences of the United States of America* 103, 2892–2897.
- Anguissola, S., Kohler, B., O'Byrne, R., *et al.*, 2009. Bid and calpains cooperate to trigger oxaliplatin-induced apoptosis of cervical carcinoma HeLa cells. *Molecular Pharmacology* 76, 998–1010.

- Arthur, J.S., Elce, J.S., Hegadorn, C., Williams, K., Greer, P.A., 2000. Disruption of the murine calpain small subunit gene, *Capn4*: calpain is essential for embryonic development but not for cell growth and division. *Molecular and Cellular Biology* 20, 4474–4481.
- Atherton, J., Kurbatskaya, K., Bondulich, M., *et al.*, 2013. Calpain cleavage and inactivation of the sodium calcium exchanger-3 occur downstream of Abeta in Alzheimer's disease. *Aging Cell* 13, 49–59.
- Baier, L.J., Permana, P.A., Yang, X., *et al.*, 2000. A calpain-10 gene polymorphism is associated with reduced muscle mRNA levels and insulin resistance. *Journal of Clinical Investigation* 106, R69–R73.
- Bano, D., Young, K.W., Guerin, C.J., *et al.*, 2005. Cleavage of the plasma membrane $\text{Na}^+/\text{Ca}^{2+}$ exchanger in excitotoxicity. *Cell* 120, 275–285.
- Bolli, R., Marban, E., 1999. Molecular and cellular mechanisms of myocardial stunning. *Physiological Reviews* 79, 609–634.
- Cao, X., Deng, X., May, W.S., 2003. Cleavage of Bax to p18 Bax accelerates stress-induced apoptosis, and a cathepsin-like protease may rapidly degrade p18 Bax. *Blood* 102, 2605–2614.
- Chen, C.J., Nguyen, T., Shively, J.E., 2010. Role of calpain-9 and PKC-delta in the apoptotic mechanism of lumen formation in CEACAM1 transfected breast epithelial cells. *Experimental Cell Research* 316, 638–648.
- Chen, M., Fernandez, H.L., 2004. Stimulation of beta-amyloid precursor protein alpha-processing by phorbol ester involves calcium and calpain activation. *Biochemical and Biophysical Research Communications* 316, 332–340.
- Chen, M., Fernandez, H.L., 2005. Mu-calpain is functionally required for alpha-processing of Alzheimer's beta-amyloid precursor protein. *Biochemical and Biophysical Research Communication* 330, 714–721.
- Chiannikulchai, N., Pasturaud, P., Richard, I., Auffray, C., Beckmann, J.S., 1995. A primary expression map of the chromosome 15q15 region containing the recessive form of limb-girdle muscular dystrophy (LGMD2A) gene. *Human Molecular Genetics* 4, 717–725.
- Chua, B.T., Guo, K., Li, P., 2000. Direct cleavage by the calcium-activated protease calpain can lead to inactivation of caspases. *Journal of Biological Chemistry* 275, 5131–5135.
- Colantonio, D.A., Van Eyk, J.E., Przyklen, K., 2004. Stunned peri-infarct canine myocardium is characterized by degradation of troponin T, not troponin I. *Cardiovascular Research* 63, 217–225.
- Cong, J., Goll, D.E., Peterson, A.M., Kapprell, H.P., 1989. The role of autolysis in activity of the Ca^{2+} -dependent proteinases (mu-calpain and m-calpain). *Journal of Biological Chemistry* 264, 10096–10103.
- Daniel, N.N., Korsmeyer, S.J., 2004. Cell death: Critical control points. *Cell* 116, 205–219.
- De Felice, F.G., Velasco, P.T., Lambert, M.P., *et al.*, 2007. Abeta oligomers induce neuronal oxidative stress through an N-methyl-D-aspartate receptor-dependent mechanism that is blocked by the Alzheimer drug memantine. *Journal of Biological Chemistry* 282, 11590–11601.
- De Jongh, K.S., Colvin, A.A., Wang, K.K., Catterall, W.A., 1994. Differential proteolysis of the full-length form of the L-type calcium channel alpha 1 subunit by calpain. *Journal of Neurochemistry* 63, 1558–1564.
- Dear, N., Matena, K., Vingron, M., Boehm, T., 1997. A new subfamily of vertebrate calpains lacking a calmodulin-like domain: Implications for calpain regulation and evolution. *Genomics* 45, 175–184.
- Dear, T.N., Meier, N.T., Hunn, M., Boehm, T., 2000. Gene structure, chromosomal localization, and expression pattern of *Capn12*, a new member of the calpain large subunit gene family. *Genomics* 68, 152–160.
- Demuro, A., Mina, E., Kaye, R., *et al.*, 2005. Calcium dysregulation and membrane disruption as a ubiquitous neurotoxic mechanism of soluble amyloid oligomers. *Journal of Biological Chemistry* 280, 17294–17300.
- Dutt, P., Croall, D.E., Arthur, J.S., *et al.*, 2006. m-Calpain is required for preimplantation embryonic development in mice. *BMC Developmental Biology* 6, 3.
- Emori, Y., Kawasaki, H., Imajoh, S., Imahori, K., Suzuki, K., 1987. Endogenous inhibitor for calcium-dependent cysteine protease contains four internal repeats that could be responsible for its multiple reactive sites. *Proceedings of the National Academy of Sciences of the United States of America* 84, 3590–3594.
- Evans, J.C., Frayling, T.M., Cassell, P.G., *et al.*, 2001. Studies of association between the gene for calpain-10 and type 2 diabetes mellitus in the United Kingdom. *American Journal of Human Genetics* 69, 544–552.
- Galvez, A.S., Diwan, A., Odley, A.M., *et al.*, 2007. Cardiomyocyte degeneration with calpain deficiency reveals a critical role in protein homeostasis. *Circulation Research* 100, 1071–1078.
- Gao, G., Dou, Q.P., 2000. N-terminal cleavage of bax by calpain generates a potent proapoptotic 18-kDa fragment that promotes bcl-2-independent cytochrome C release and apoptotic cell death. *Journal of Cellular Biochemistry* 80, 53–72.
- Garant, M.J., Kao, W.H., Brancati, F., *et al.*, 2002. SNP43 of CAPN10 and the risk of type 2 Diabetes in African-Americans: The Atherosclerosis Risk in Communities Study. *Diabetes* 51, 231–237.
- Gascon, S., Sobrado, M., Roda, J.M., Rodriguez-Pena, A., Diaz-Guerra, M., 2008. Excitotoxicity and focal cerebral ischemia induce truncation of the NR2A and NR2B subunits of the NMDA receptor and cleavage of the scaffolding protein PSD-95. *Molecular Psychiatry* 13, 99–114.
- Goll, D.E., Thompson, V.F., Li, H., Wei, W., Cong, J., 2003. The calpain system. *Physiological Reviews* 83, 731–801.
- Goni-Oliver, P., Lucas, J.J., Avila, J., Hernandez, F., 2007. N-terminal cleavage of GSK-3 by calpain: A new form of GSK-3 regulation. *Journal of Biological Chemistry* 282, 22406–22413.
- Govindarajan, B., Laird, J., Sherman, R., Salomon, R.G., Bhattacharya, S.K., 2008. Neuroprotection in glaucoma using calpain-1 inhibitors: Regional differences in calpain-1 activity in the trabecular meshwork, optic nerve and implications for therapeutics. *CNS & Neurological Disorders – Drug Targets* 7, 295–304.
- Groshong, J.S., Spencer, M.J., Bhattacharyya, B.J., *et al.*, 2007. Calpain activation impairs neuromuscular transmission in a mouse model of the slow-channel myasthenic syndrome. *Journal of Clinical Investigation* 117, 2903–2912.
- Grundke-Iqbal, I., Iqbal, K., Tung, Y.C., *et al.*, 1986. Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proceedings of the National Academy of Sciences of the United States of America* 83, 4913–4917.
- Haass, C., Selkoe, D.J., 2007. Soluble protein oligomers in neurodegeneration: Lessons from the Alzheimer's amyloid beta-peptide. *Nature Reviews Molecular Cell Biology* 8, 101–112.
- Hall, E.D., Sullivan, P.G., Gibson, T.R., *et al.*, 2005. Spatial and temporal characteristics of neurodegeneration after controlled cortical impact in mice: More than a focal brain injury. *Journal of Neurotrauma* 22, 252–265.
- Hanna, R.A., Campbell, R.L., Davies, P.L., 2008. Calcium-bound structure of calpain and its mechanism of inhibition by calpastatin. *Nature* 456, 409–412.
- Hata, S., Abe, M., Suzuki, H., *et al.*, 2010. Calpain 8/nCL-2 and calpain 9/nCL-4 constitute an active protease complex, G-calpain, involved in gastric mucosal defense. *PLoS Genetics* 6, e1001040.
- Hata, S., Doi, N., Kitamura, F., Sorimachi, H., 2007. Stomach-specific calpain, nCL-2/calpain 8, is active without calpain regulatory subunit and oligomerizes through C2-like domains. *Journal of Biological Chemistry* 282, 27847–27856.
- Hata, S., Koyama, S., Kawahara, H., *et al.*, 2006. Stomach-specific calpain, nCL-2, localizes in mucus cells and proteolyzes the beta-subunit of coatomer complex, beta-COP. *Journal of Biological Chemistry* 281, 11214–11224.
- Hathaway, D.R., Werth, D.K., Haeblerle, J.R., 1982. Limited autolysis reduces the Ca^{2+} requirement of a smooth muscle Ca^{2+} -activated protease. *Journal of Biological Chemistry* 257, 9072–9077.
- Hell, J.W., Westenbroek, R.E., Breeze, L.J., *et al.*, 1996. N-methyl-D-aspartate receptor-induced proteolytic conversion of postsynaptic class C L-type calcium channels in hippocampal neurons. *Proceedings of the National Academy of Sciences of the United States of America* 93, 3362–3367.
- Horikawa, Y., 2006. Calpain-10 (NIDDM1) as a susceptibility gene for common type 2 diabetes. *Endocrine Journal* 53, 567–576.
- Horikawa, Y., Oda, N., Cox, N.J., *et al.*, 2000. Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. *Nature Genetics* 26, 163–175.
- Hou, S.T., Jiang, S.X., Desbois, A., *et al.*, 2006. Calpain-cleaved collapsin response mediator protein-3 induces neuronal death after glutamate toxicity and cerebral ischemia. *Journal of Neuroscience* 26, 2241–2249.
- Huang, W., Fileta, J., Rawe, I., Qu, J., Grosskreutz, C.L., 2010. Calpain activation in experimental glaucoma. *Investigative Ophthalmology & Visual Science* 51, 3049–3054.
- Imajoh, S., Kawasaki, H., Suzuki, K., 1987. The COOH-terminal E-F hand structure of calcium-activated neutral protease (CANP) is important for the association of subunits and resulting proteolytic activity. *Journal of Biochemistry* 101, 447–452.
- Inomata, M., Kasai, Y., Nakamura, M., Kawashima, S., 1988. Activation mechanism of calcium-activated neutral protease. Evidence for the existence of intramolecular and intermolecular autolyses. *Journal of Biological Chemistry* 263, 19783–19787.
- Iwamoto, H., Miura, T., Okamura, T., *et al.*, 1999. Calpain inhibitor-1 reduces infarct size and DNA fragmentation of myocardium in ischemic/reperfused rat heart. *Journal of Cardiovascular Pharmacology* 33, 580–586.
- Kampfl, A., Posmantur, R., Nixon, R., *et al.*, 1996. mu-calpain activation and calpain-mediated cytoskeletal proteolysis following traumatic brain injury. *Journal of Neurochemistry* 67, 1575–1583.

- Kang, M.Y., Zhang, Y., Matkovich, S.J., *et al.*, 2010. Receptor-independent cardiac protein kinase Calpha activation by calpain-mediated truncation of regulatory domains. *Circulation Research* 107, 903–912.
- Karmakar, S., Banik, N.L., Ray, S.K., 2007. Curcumin suppressed anti-apoptotic signals and activated cysteine proteases for apoptosis in human malignant glioblastoma U87MG cells. *Neurochemical Research* 32, 2103–2113.
- Kishimoto, A., Kajikawa, N., Shiota, M., Nishizuka, Y., 1983. Proteolytic activation of calcium-activated, phospholipid-dependent protein kinase by calcium-dependent neutral protease. *Journal of Biological Chemistry* 258, 1156–1164.
- Kishimoto, A., Mikawa, K., Hashimoto, K., *et al.*, 1989. Limited proteolysis of protein kinase C subspecies by calcium-dependent neutral protease (calpain). *Journal of Biological Chemistry* 264, 4088–4092.
- Kiss, R., Kovacs, D., Tompa, P., Perczel, A., 2008. Local structural preferences of calpastatin, the intrinsically unstructured protein inhibitor of calpain. *Biochemistry* 47, 6936–6945.
- Kulkarni, S., Reddy, K.B., Esteva, F.J., *et al.*, 2010. Calpain regulates sensitivity to trastuzumab and survival in HER2-positive breast cancer. *Oncogene* 29, 1339–1350.
- Kupina, N.C., Nath, R., Bernath, E.E., *et al.*, 2001. The novel calpain inhibitor SJA6017 improves functional outcome after delayed administration in a mouse model of diffuse brain injury. *Journal of Neurotrauma* 18, 1229–1240.
- van der Laarse, A., 2002. Hypothesis: Troponin degradation is one of the factors responsible for deterioration of left ventricular function in heart failure. *Cardiovascular Research* 56, 8–14.
- Lee, H.J., Sorimachi, H., Jeong, S.Y., Ishiura, S., Suzuki, K., 1998. Molecular cloning and characterization of a novel tissue-specific calpain predominantly expressed in the digestive tract. *Biological Chemistry* 379, 175–183.
- Lee, H.J., Tomioka, S., Kinbara, K., *et al.*, 1999. Characterization of a human digestive tract-specific calpain, nCL-4, expressed in the baculovirus system. *Archives of Biochemistry and Biophysics* 362, 22–31.
- Lee, S.K., Park, M.S., Nam, M.J., 2008. Aspirin has antitumor effects via expression of calpain gene in cervical cancer cells. *Journal of Oncology* 2008, 285374.
- Leloup, L., Shao, H., Bae, Y.H., *et al.*, 2010. m-Calpain activation is regulated by its membrane localization and by its binding to phosphatidylinositol 4,5-bisphosphate. *Journal of Biological Chemistry* 285, 33549–33566.
- Letavernier, E., Perez, J., Belloq, A., *et al.*, 2008. Targeting the calpain/calpastatin system as a new strategy to prevent cardiovascular remodeling in angiotensin II-induced hypertension. *Circulation Research* 102, 720–728.
- Li, Q.X., Evin, G., Small, D.H., *et al.*, 1995. Proteolytic processing of Alzheimer's disease beta A4 amyloid precursor protein in human platelets. *Journal of Biological Chemistry* 270, 14140–14147.
- Li, Y., Ma, J., Zhu, H., *et al.*, 2011. Targeted inhibition of calpain reduces myocardial hypertrophy and fibrosis in mouse models of type 1 diabetes. *Diabetes* 60, 2985–2994.
- Liu, K., Li, L., Cohen, S.N., 2000. Antisense RNA-mediated deficiency of the calpain protease, nCL-4, in NIH3T3 cells is associated with neoplastic transformation and tumorigenesis. *Journal of Biological Chemistry* 275, 31093–31098.
- Liu, L., Xing, D., Chen, W.R., *et al.*, 2008. Calpain-mediated pathway dominates cisplatin-induced apoptosis in human lung adenocarcinoma cells as determined by real-time single cell analysis. *International Journal of Cancer* 122, 2210–2222.
- Lucas, J.J., Hernandez, F., Gomez-Ramos, P., *et al.*, 2001. Decreased nuclear beta-catenin, tau hyperphosphorylation and neurodegeneration in GSK3beta conditional transgenic mice. *EMBO Journal* 20, 27–39.
- Macqueen, D.J., Delbridge, M.L., Manthri, S., Johnston, I.A., 2010. A newly classified vertebrate calpain protease, directly ancestral to CAPN1 and 2, episodically evolved a restricted physiological function in placental mammals. *Molecular Biology and Evolution* 27, 1886–1902.
- Maekawa, A., Lee, J.K., Nagaya, T., *et al.*, 2003. Overexpression of calpastatin by gene transfer prevents troponin I degradation and ameliorates contractile dysfunction in rat hearts subjected to ischemia/reperfusion. *Journal of Molecular and Cellular Cardiology* 35, 1277–1284.
- Maki, M., Takano, E., Mori, H., *et al.*, 1987a. Repetitive region of calpastatin is a functional unit of the proteinase inhibitor. *Biochemical and Biophysical Research Communications* 143, 300–308.
- Maki, M., Takano, E., Mori, H., *et al.*, 1987b. All four internally repetitive domains of pig calpastatin possess inhibitory activities against calpains I and II. *FEBS Letters* 223, 174–180.
- Malecki, M.T., Moczulski, D.K., Klupa, T., *et al.*, 2002. Homozygous combination of calpain 10 gene haplotypes is associated with type 2 diabetes mellitus in a Polish population. *European Journal of Endocrinology* 146, 695–699.
- Mandic, A., Viktorsson, K., Strandberg, L., *et al.*, 2002. Calpain-mediated Bid cleavage and calpain-independent Bak modulation: Two separate pathways in cisplatin-induced apoptosis. *Molecular and Cellular Biology* 22, 3003–3013.
- Mani, S.K., Shiraishi, H., Balasubramanian, S., *et al.*, 2008. *In vivo* administration of calpeptin attenuates calpain activation and cardiomyocyte loss in pressure-overloaded feline myocardium. *American Journal of Physiology – Heart and Circulatory Physiology* 295, H314–H326.
- Mao, X., Yang, S.H., Simpkins, J.W., Barger, S.W., 2007. Glutamate receptor activation evokes calpain-mediated degradation of Sp3 and Sp4, the prominent Sp-family transcription factors in neurons. *Journal of Neurochemistry* 100, 1300–1314.
- McCollum, A.T., Nasr, P., Estus, S., 2002. Calpain activates caspase-3 during UV-induced neuronal death but only calpain is necessary for death. *Journal of Neurochemistry* 82, 1208–1220.
- McGinnis, K.M., Gnegy, M.E., Park, Y.H., Mukerjee, N., Wang, K.K., 1999. Procaspase-3 and poly(ADP)ribose polymerase (PARP) are calpain substrates. *Biochemical and Biophysical Research Communications* 263, 94–99.
- Mellgren, R.L., Huang, X., 2007. Fetuin A stabilizes m-calpain and facilitates plasma membrane repair. *Journal of Biological Chemistry* 282, 35868–35877.
- Mellgren, R.L., Zhang, W., Miyake, K., McNeil, P.L., 2007. Calpain is required for the rapid, calcium-dependent repair of wounded plasma membrane. *Journal of Biological Chemistry* 282, 2567–2575.
- Mikosik, A., Zaremba, A., Puchalska, Z., *et al.*, 2007. Ex vivo measurement of calpain activation in human peripheral blood lymphocytes by detection of immunoreactive products of calpastatin degradation. *Folia Histochemica et Cytobiologica* 45, 343–347.
- Miyazaki, T., Taketomi, Y., Takimoto, M., *et al.*, 2011. m-Calpain induction in vascular endothelial cells on human and mouse atheromas and its roles in VE-cadherin disorganization and atherosclerosis. *Circulation* 124, 2522–2532.
- Mizushima, N., Levine, B., 2010. Autophagy in mammalian development and differentiation. *Nature Cell Biology* 12, 823–830.
- Moldoveanu, T., Gehring, K., Green, D.R., 2008. Concerted multi-pronged attack by calpastatin to occlude the catalytic cleft of heterodimeric calpains. *Nature* 456, 404–408.
- Moldoveanu, T., Hosfield, C.M., Lim, D., *et al.*, 2002. A Ca(2+) switch aligns the active site of calpain. *Cell* 108, 649–660.
- Moldoveanu, T., Hosfield, C.M., Lim, D., Jia, Z., Davies, P.L., 2003. Calpain silencing by a reversible intrinsic mechanism. *Nature Structural Biology* 10, 371–378.
- Nakagawa, T., Yuan, J., 2000. Cross-talk between two cysteine protease families. Activation of caspase-12 by calpain in apoptosis. *Journal of Cell Biology* 150, 887–894.
- Nixon, R.A., Saito, K.I., Grynspan, F., *et al.*, 1994. Calcium-activated neutral proteinase (calpain) system in aging and Alzheimer's disease. *Annals of the New York Academy of Sciences* 747, 77–91.
- Noguchi, M., Sarin, A., Aman, M.J., *et al.*, 1997. Functional cleavage of the common cytokine receptor gamma chain (gamma c) by calpain. *Proceedings of the National Academy of Sciences of the United States of America* 94, 11534–11539.
- Ono, Y., Hayashi, C., Doi, N., *et al.*, 2007. Comprehensive survey of p94/calpain 3 substrates by comparative proteomics – Possible regulation of protein synthesis by p94. *Biotechnology Journal* 2, 565–576.
- Ono, Y., Kakinuma, K., Torii, F., *et al.*, 2004. Possible regulation of the conventional calpain system by skeletal muscle-specific calpain, p94/calpain 3. *Journal of Biological Chemistry* 279, 2761–2771.
- Ono, Y., Ojima, K., Torii, F., *et al.*, 2010. Skeletal muscle-specific calpain is an intracellular Na⁺-dependent protease. *Journal of Biological Chemistry* 285, 22986–22998.
- Ono, Y., Shimada, H., Sorimachi, H., *et al.*, 1998. Functional defects of a muscle-specific calpain, p94, caused by mutations associated with limb-girdle muscular dystrophy type 2A. *Journal of Biological Chemistry* 273, 17073–17078.
- Ono, Y., Torii, F., Ojima, K., *et al.*, 2006. Suppressed disassembly of autolyzing p94/CAPN3 by N2A connectin/titin in a genetic reporter system. *Journal of Biological Chemistry* 281, 18519–18531.
- Orho-Melander, M., Klannemark, M., Svensson, M.K., *et al.*, 2002. Variants in the calpain-10 gene predispose to insulin resistance and elevated free fatty acid levels. *Diabetes* 51, 2658–2664.
- Park, S.Y., Ferreira, A., 2005. The generation of a 17 kDa neurotoxic fragment: An alternative mechanism by which tau mediates beta-amyloid-induced neurodegeneration. *Journal of Neuroscience* 25, 5365–5375.
- Pontremoli, S., Melloni, E., 1988. The role of calpain and protein kinase C in activation of human neutrophils. *Progress in Clinical Biological Research* 282, 195–208.
- Porn-Ares, M.I., Samali, A., Orrenius, S., 1998. Cleavage of the calpain inhibitor, calpastatin, during apoptosis. *Cell Death and Differentiation* 5, 1028–1033.

- Posmantur, R., Kampfl, A., Siman, R., *et al.*, 1997. A calpain inhibitor attenuates cortical cytoskeletal protein loss after experimental traumatic brain injury in the rat. *Neuroscience* 77, 875–888.
- Pottorf 2nd, W.J., Johanns, T.M., Derrington, S.M., *et al.*, 2006. Glutamate-induced protease-mediated loss of plasma membrane Ca^{2+} pump activity in rat hippocampal neurons. *Journal of Neurochemistry* 98, 1646–1656.
- Rashidian, J., Iyiriharo, G., Aleyasin, H., *et al.*, 2005. Multiple cyclin-dependent kinases signals are critical mediators of ischemia/hypoxic neuronal death *in vitro* and *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America* 102, 14080–14085.
- Richard, I., Broux, O., Allamand, V., *et al.*, 1995. Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. *Cell* 81, 27–40.
- Richard, I., Roudaut, C., Marchand, S., *et al.*, 2000. Loss of calpain 3 proteolytic activity leads to muscular dystrophy and to apoptosis-associated IkappaBalpha/nuclear factor kappaB pathway perturbation in mice. *Journal of Cell Biology* 151, 1583–1590.
- Richard, I., Roudaut, C., Saenz, A., *et al.*, 1999. Calpainopathy – A survey of mutations and polymorphisms. *American Journal of Human Genetics* 64, 1524–1540.
- Ruiz-Vela, A., Gonzalez de Buitrago, G., Martinez, A.C., 1999. Implication of calpain in caspase activation during B cell clonal deletion. *EMBO Journal* 18, 4988–4998.
- Saatman, K.E., Abai, B., Grosvenor, A., *et al.*, 2003. Traumatic axonal injury results in biphasic calpain activation and retrograde transport impairment in mice. *Journal of Cerebral Blood Flow and Metabolism* 23, 34–42.
- Saatman, K.E., Bozyczko-Coyne, D., Marcy, V., Siman, R., McIntosh, T.K., 1996a. Prolonged calpain-mediated spectrin breakdown occurs regionally following experimental brain injury in the rat. *Journal of Neuropathology and Experimental Neurology* 55, 850–860.
- Saatman, K.E., Murai, H., Bartus, R.T., *et al.*, 1996b. Calpain inhibitor AK295 attenuates motor and cognitive deficits following experimental brain injury in the rat. *Proceedings of the National Academy of Sciences of the United States of America* 93, 3428–3433.
- Saenz, A., Leturcq, F., Cobo, A.M., *et al.*, 2005. LGMD2A: Genotype-phenotype correlations based on a large mutational survey on the calpain 3 gene. *Brain* 128, 732–742.
- Saito, K., Elce, J.S., Hamos, J.E., Nixon, R.A., 1993. Widespread activation of calcium-activated neutral proteinase (calpain) in the brain in Alzheimer disease: A potential molecular basis for neuronal degeneration. *Proceedings of the National Academy of Sciences of the United States of America* 90, 2628–2632.
- Scalia, R., Gong, Y., Berzins, B., Zhao, L.J., Sharma, K., 2007. Hyperglycemia is a major determinant of albumin permeability in diabetic microcirculation: The role of mu-calpain. *Diabetes* 56, 1842–1849.
- Sergeev, I.N., 2004. Genistein induces Ca^{2+} -mediated, calpain/caspase-12-dependent apoptosis in breast cancer cells. *Biochemical and Biophysical Research Communications* 321, 462–467.
- Shao, H., Chou, J., Baty, C.J., *et al.*, 2006. Spatial localization of m-calpain to the plasma membrane by phosphoinositide biphosphate binding during epidermal growth factor receptor-mediated activation. *Molecular and Cellular Biology* 26, 5481–5496.
- Shim, H.Y., Park, J.H., Paik, H.D., *et al.*, 2007. Genistein-induced apoptosis of human breast cancer MCF-7 cells involves calpain-caspase and apoptosis signaling kinase 1-p38 mitogen-activated protein kinase activation cascades. *Anticancer Drugs* 18, 649–657.
- Shimada, M., Greer, P.A., McMahon, A.P., Bouxsein, M.L., Schipani, E., 2008. *In vivo* targeted deletion of calpain small subunit, Capn4, in cells of the osteoblast lineage impairs cell proliferation, differentiation, and bone formation. *Journal of Biological Chemistry* 283, 21002–21010.
- Singh, R.B., Chohan, P.K., Dhalla, N.S., Netticadan, T., 2004. The sarcoplasmic reticulum proteins are targets for calpain action in the ischemic-reperfused heart. *Journal of Molecular and Cellular Cardiology* 37, 101–110.
- Smolock, A.R., Mishra, G., Eguchi, K., Eguchi, S., Scalia, R., 2011. Protein kinase C upregulates intercellular adhesion molecule-1 and leukocyte-endothelium interactions in hyperglycemia via activation of endothelial expressed calpain. *Arteriosclerosis, Thrombosis, and Vascular Biology* 31, 289–296.
- Sorimachi, H., Hata, S., Ono, Y., 2011. Calpain chronicle – An enzyme family under multidisciplinary characterization. *Proceedings of the Japan Academy – Series B: Physical & Biological Sciences* 87, 287–327.
- Sorimachi, H., Ishiura, S., Suzuki, K., 1993. A novel tissue-specific calpain species expressed predominantly in the stomach comprises two alternative splicing products with and without Ca^{2+} -binding domain. *Journal of Biological Chemistry* 268, 19476–19482.
- Sorimachi, H., Ono, Y., 2012. Regulation and physiological roles of the calpain system in muscular disorders. *Cardiovascular Research* 96, 11–22.
- Subramanian, V., Uchida, H.A., Ijaz, T., *et al.*, 2012. Calpain inhibition attenuates angiotensin II-induced abdominal aortic aneurysms and atherosclerosis in low-density lipoprotein receptor-deficient mice. *Journal of Cardiovascular Pharmacology* 59, 66–76.
- Suzuki, K., Tsuji, S., Kubota, S., Kimura, Y., Imahori, K., 1981. Limited autolysis of Ca^{2+} -activated neutral protease (CANP) changes its sensitivity to Ca^{2+} ions. *Journal of Biochemistry* 90, 275–278.
- Tait, S.W., Green, D.R., 2010. Mitochondria and cell death: Outer membrane permeabilization and beyond. *Nature Reviews Molecular Cell Biology* 11, 621–632.
- Takano, J., Mihira, N., Fujioka, R., *et al.*, 2011. Vital role of the calpain-calpastatin system for placental-integrity-dependent embryonic survival. *Molecular and Cellular Biology* 31, 4097–4106.
- Tan, Y., Wu, C., De Veyra, T., Greer, P.A., 2006. Ubiquitous calpains promote both apoptosis and survival signals in response to different cell death stimuli. *Journal of Biological Chemistry* 281, 17689–17698.
- Tanaka, K., 2009. The proteasome: Overview of structure and functions. *Proceedings of the Japan Academy – Series B: Physical & Biological Sciences* 85, 12–36.
- Thomas, S.A., Fallavollita, J.A., Lee, T.C., Feng, J., Canty Jr., J.M., 1999. Absence of troponin I degradation or altered sarcoplasmic reticulum uptake protein expression after reversible ischemia in swine. *Circulation Research* 85, 446–456.
- Tomba, P., Emori, Y., Sorimachi, H., Suzuki, K., Friedrich, P., 2001. Domain III of calpain is a Ca^{2+} -regulated phospholipid-binding domain. *Biochemical and Biophysical Research Communications* 280, 1333–1339.
- Veeranna, J., Kaji, T., Boland, B., *et al.*, 2004. Calpain mediates calcium-induced activation of the erk1,2 MAPK pathway and cytoskeletal phosphorylation in neurons: Relevance to Alzheimer's disease. *American Journal of Pathology* 165, 795–805.
- Vosler, P.S., Brennan, C.S., Chen, J., 2008. Calpain-mediated signaling mechanisms in neuronal injury and neurodegeneration. *Molecular Neurobiology* 38, 78–100.
- Wang, K.K., Posmantur, R., Nadimpalli, R., *et al.*, 1998. Caspase-mediated fragmentation of calpain inhibitor protein calpastatin during apoptosis. *Archives of Biochemistry and Biophysics* 356, 187–196.
- Waterhouse, N.J., Finucane, D.M., Green, D.R., *et al.*, 1998. Calpain activation is upstream of caspases in radiation-induced apoptosis. *Cell Death and Differentiation* 5, 1051–1061.
- Wells, A., Huttenlocher, A., Lauffenburger, D.A., 2005. Calpain proteases in cell adhesion and motility. *International Review of Cytology* 245, 1–16.
- Wood, D.E., Newcomb, E.W., 1999. Caspase-dependent activation of calpain during drug-induced apoptosis. *Journal of Biological Chemistry* 274, 8309–8315.
- Wu, H.Y., Tomizawa, K., Matsui, H., 2007. Calpain-calcineurin signaling in the pathogenesis of calcium-dependent disorder. *Acta Medica Okayama* 61, 123–137.
- Yeh, T.C., Chiang, P.C., Li, T.K., *et al.*, 2007. Genistein induces apoptosis in human hepatocellular carcinomas via interaction of endoplasmic reticulum stress and mitochondrial insult. *Biochemical Pharmacology* 73, 782–792.
- Yoshizawa, T., Sorimachi, H., Tomioka, S., Ishiura, S., Suzuki, K., 1995. A catalytic subunit of calpain possesses full proteolytic activity. *FEBS Letters* 358, 101–103.
- Zhang, Y., Matkovich, S.J., Duan, X., *et al.*, 2011. Receptor-independent protein kinase C alpha (PKCalpha) signaling by calpain-generated free catalytic domains induces HDAC5 nuclear export and regulates cardiac transcription. *Journal of Biological Chemistry* 286, 26943–26951.
- Zimmerman, U.J., Boring, L., Pak, J.H., Mukerjee, N., Wang, K.K., 2000. The calpain small subunit gene is essential: Its inactivation results in embryonic lethality. *IUBMB Life* 50, 63–68.

Cathepsin E: An Aspartic Protease with Diverse Functions and Biomedical Implications

K Yamamoto, Kyushu University, Fukuoka, Japan

K Okamoto and T Tsukuba, Nagasaki University, Nagasaki, Japan

© 2016 Elsevier Inc. All rights reserved.

Aspartic Peptidases

Proteolytic enzymes (peptidases, proteases) are of great relevance to biology, medicine, and biotechnology. Based on the nature of their catalytic residues and the homologous sets of their sequences, peptidases are now classified into nine families; aspartic- (A), cysteine- (C), glutamic- (G), metallo- (M), asparagine- (N), mixed- (P), serine- (S), threonine- (T) and unknown- (U) peptidases. Aspartic peptidases are widely distributed among vertebrates, yeast, plants, nematodes, fungi, protozoa, parasites, and retrovirus and have received enormous interest because their abnormal expression and uncontrolled activities are closely linked to various human diseases. Extra efforts have thus been made to develop clinically active inhibitors of each peptidase for the treatment of such diseases.

According to the MEROPS database, aspartic peptidases are further classified into 16 different subfamilies, based on their amino acid sequence homology followed by their evolutionary relationship and tertiary structures. The A1 family peptidase includes most of aspartic endopeptidases found in higher organisms and has the catalytic Asp residues within an Asp-Thr/Ser-Gly motif between both N- and C-terminal lobes of the molecules. The two lobes have similar structures, despite very little amino acid sequence similarity. Most of the A1 family peptidases display the maximum activity at acid pH, although some such as renin are active at neutral pH, and are more or less strongly inhibited by pepstatin. This family is also divided into secretory and nonsecretory groups. The former includes pepsin, gastricsin, chymosin, renin, and candida-pepsin SAP1, whereas the latter is represented by cathepsin D, cathepsin E, and β -secretase. The former members have not only well-defined physiological functions but also the possible implications in various pathological conditions, for example, cancer and hypertension. Most of the latter members are also associated with various human diseases, but mechanistic insights into how each enzyme links to pathological processes peculiar to each enzyme are not well addressed. Among them, pepsin A is well-known to contribute to proteolysis of food proteins in the vertebrate stomach, whereas cathepsins D and E serve for nonspecific and/or specific proteolysis in the endosomal/lysosomal system or the extracellular spaces. β -Secretase (BACE1, β -site amyloid precursor protein (APP) cleaving enzyme 1, or memapsin 2) is a membrane-bound aspartic peptidase and is known to cleave the type I transmembrane APP to generate a secreted form of truncated APP and a membrane-anchored C-terminal fragment of 99 amino acids (C99). C99 is further processed by another membrane peptidase γ -secretase, composed of four transmembrane proteins (presenilin, nicastrin, Pen2, and Aph1), to yield a p3 fragment and the neurotoxic amyloid β peptide (A β). Since

accumulation of A β in the brain is closely associated with the pathogenesis of Alzheimer's disease (AD), strategies to inhibit A β formation are believed to prove beneficial for AD treatment. To date, a number of inhibitors of β -secretase to prevent A β generation have been developed.

The pepsin subfamily of the A1 family (clan A1A) is significantly related to the retropepsin subfamily of the A2 family (clan A2A). Human immunodeficiency virus type-1 (HIV-1) retropepsin is a representative aspartic peptidase in the A2A subfamily, essential for processing newly synthesized *gag-pol* polyproteins at eight sites, mostly with hydrophobic residues in P1 and P'1, thereby generating the mature protein components of an infectious HIV virion. Hence this peptidase is indispensable for the life-cycle of HIV. By X-ray crystallography, this peptidase is shown to exist as a catalytically active homodimer, with each subunit made up of 99 amino acids. The active site of HIV-1 retropepsin lies between the two subunits containing one Asp-Thr-Gly motif for each. Each motif contributes to one of the pair of catalytic Asp residues (reviewed in, Navia and MaKeever, 2006; Louis *et al.*, 2007; Patel *et al.*, 2013). Therefore, the retropepsin is believed to be an important drug target for AIDS. Actually, several promising HIV-1 peptidase inhibitors have entered human clinical trials.

Cathepsin E

Historical Background

Cathepsin E was first identified in aqueous extract of rabbit bone marrow by Lapresle and Webb (1962), where it exhibited the maximal activity at pH 2.5 on human serum albumin and a higher molecular weight than the analogous cathepsin D. For 15 years since then the existence of this enzyme had long been an object of discussion. This argument seemed to reach a conclusion by the identification and characterization of two analogous aspartic peptidases, cathepsin E-like enzyme (Yamamoto *et al.*, 1978) and cathepsin D (Yamamoto *et al.*, 1979) in rat spleen and subsequently by showing clear differences in their immunochemical properties (Yamamoto *et al.*, 1980). In those days, the analogous proteinases were also isolated from a variety of sources and called by different designations, including cathepsin D-type proteases (lymphoid tissues) (Yago and Bowers, 1975), cathepsin D-like acid proteinase and slow moving proteinase (gastric mucosa) (Kageyama and Takahashi, 1980; Muto *et al.*, 1983; Samloff *et al.*, 1987), erythrocyte membrane acid proteinase (EMAP) (erythrocyte membranes) (Yamamoto and Marchesi, 1984; Takeda *et al.*, 1986), and cathepsin E (epidermis) (Hara *et al.*, 1993). Later, on the basis of biochemical and immunochemical analyses, these enzymes were proved to be identical

(Tarasova *et al.*, 1986; Yamamoto *et al.*, 1987; Muto *et al.*, 1987; Jupp *et al.*, 1988) and ultimately referred to as cathepsin E. Shortly after these studies, the cDNA clone encoding human cathepsin E was isolated and sequenced, and its chromosomal localization and sequence homology to other aspartic peptidases were defined (Azuma *et al.*, 1989). At this point, cathepsin E established its position as an independent peptidase.

Tissue Distribution and Cellular Localization

Differing from the definite localization of other related aspartic peptidases, cathepsin E has a limited distribution in certain cell types, including the immune system cells such as lymphocytes, macrophages, dendritic cells, microglia, and Langerhans cells and the rapidly regenerating cells such as gastric mucosal cells, epidermal keratinocytes (reviewed in, Yasuda *et al.*, 1999a,b; Zaidi and Kalbacher, 2008; Chlabicz *et al.*, 2011; Yamamoto *et al.*, 2012), and mammary epithelial cells (Kawakubo *et al.*, 2014). The cellular localization of cathepsin E also varies with different cell types. In antigen presenting cells (APC) such as microglia (Sastradipura *et al.*, 1998; Nishioku *et al.*, 2002), macrophages (Yanagawa *et al.*, 2007) and dendritic cells (Chain *et al.*, 2005), mast cells (Henningsson *et al.*, 2005), and mammary epithelial cells (Kawakubo *et al.*, 2014), cathepsin E is mainly confined to endosomal structures. The association of cathepsin E with plasma membranes is observed in erythrocytes (Yamamoto and Marchesi, 1984; Takeda *et al.*, 1986), intracellular canaliculi of gastric parietal cells, renal proximal tubule cells, bile canaliculi of hepatic cells (Saku *et al.*, 1991), and osteoclasts (Yoshimine *et al.*, 1995). The enzyme is also detected in the endoplasmic reticulum and/or Golgi complex (Samloff *et al.*, 1987; Fiocca *et al.*, 1990; Saku *et al.*, 1991; Finzi *et al.*, 1993; Sastradipura *et al.*, 1998), and the cytoplasmic matrix (Saku *et al.*, 1991; Sakai *et al.*, 1992; Tsukuba *et al.*, 1993; Finley and Kornfeld, 1994). In addition to its intracellular localization, cathepsin E is secreted extracellularly from activated APC (Yanagawa *et al.*, 2006), mammary epithelial cells (Kawakubo *et al.*, 2014), and cancer cells (Kawakubo *et al.*, 2007). Such strategic expression and localization of cathepsin E have been suggested its association with specific functions in the individual cell types.

Regulation of Cathepsin E Gene Expression

To date, the nucleotide sequence of the cDNA encoding procathepsin E has been defined in human (Azuma *et al.*, 1989), guinea pig (Kageyama *et al.*, 1992), rabbit (Kageyama, 1993), rat (Okamoto *et al.*, 1995), and mouse (Tatnell *et al.*, 1997). The promoter region of cathepsin E gene has some characteristic features different from those of other analogous aspartic peptidases, which seems to be linked to its limited distribution. The promoter region of cathepsin E gene had no TATA box that offers a binding site for transcription factor IID (Azuma *et al.*, 1992; Tatnell *et al.*, 1998). The tissue- and cell-specific transcription regulation of cathepsin E gene was also suggested to be associated in part with a decreased level of methylation in the gene region (Tsukada *et al.*, 1992) or the balance between the positive effect of cell-specific transcription

factors such as GATA1 and PU1 and the negative influence of the ubiquitous YY1 factor (Cook *et al.*, 2001). Cathepsin E transcription may also be regulated positively by PU1 and p300 and negatively by Class II transactivator (Yee *et al.*, 2004). Additionally, it has recently been demonstrated that the expression of murine cathepsin E gene is positively regulated by binding of the transcription factor Sp1 to its promoter region (Okamoto *et al.*, 2012). Characteristics of the transcripts of mouse cathepsin E gene are shown in Figure 1.

Biosynthesis and Trafficking

Cell-specific processing and trafficking

Cathepsin E exists as a homodimer of two catalytic active monomers with a molecular mass of 42 kDa and is N-glycosylated with high mannose- and/or complex-type oligosaccharides. Cathepsin E is present predominantly as the homodimer and partly as the catalytically active monomer. The dimer and monomer forms are easily interconverted by reducing agents such as 2-mercaptoethanol and cysteine *in vitro* (Kageyama *et al.*, 1992). Although the catalytic properties were indistinguishable between the both forms, the dimer form was more stable to pH and temperature changes than the monomer form (Kageyama *et al.*, 1992; Fowler *et al.*, 1995; Tsukuba *et al.*, 1996). Accordingly, the dimerization is thought to be necessary for structural stabilization of this molecule, but not for its proper folding, correct intracellular localization, and carbohydrate modification (Tsukuba *et al.*, 1996).

Although the potential N-glycosylation site near the N-terminal region is conserved in all of the species tested, its number near the C-terminal regions varies with different species. Human and rabbit cathepsin E has one potential N-glycosylation site at the N-terminal region only. Besides the N-terminal N-glycosylation site, rat and guinea pig cathepsin E has an additional N-glycosylation site at the C-terminal region. Mouse cathepsin E has two additional N-glycosylation sites in the C-terminal region. Cathepsin E from erythrocytes (human and rat), microglia, and thymocytes (rat) is N-glycosylated with complex-type oligosaccharides, whereas the enzyme from spleen and stomach (rat) has high-mannose-type oligosaccharides. Therefore, the nature of its oligosaccharide chains may be cell-specific or vary with its cellular localization. By pulse-chase experiments with ³⁵S-methionine, cathepsin E was shown to be initially synthesized on the membrane-bound ribosomes as a preproenzyme and then undergo different proteolytic and glycolytic processing depending on cell types before being sorted to its appropriate cellular destinations (reviewed in, Yamamoto *et al.*, 1994, 1997; Yamamoto, 1999). In rat macrophages (Yamamoto *et al.*, 1997) and microglia (Sastradipura *et al.*, 1998), procathepsin E undergoes proteolytic processing to yield the mature enzyme composing of two active monomers via the intermediate form. In these cell types, procathepsin E is initially N-glycosylated with high mannose-type oligosaccharides and then converted to the mature enzyme with complex-type oligosaccharides. In rat thymocytes and mouse Friend erythroleukemia cells, procathepsin E is initially synthesized as a 46-kDa monomer and converted to the homodimer and then processed to the 44-kDa intermediate form with complex-type oligosaccharides, which

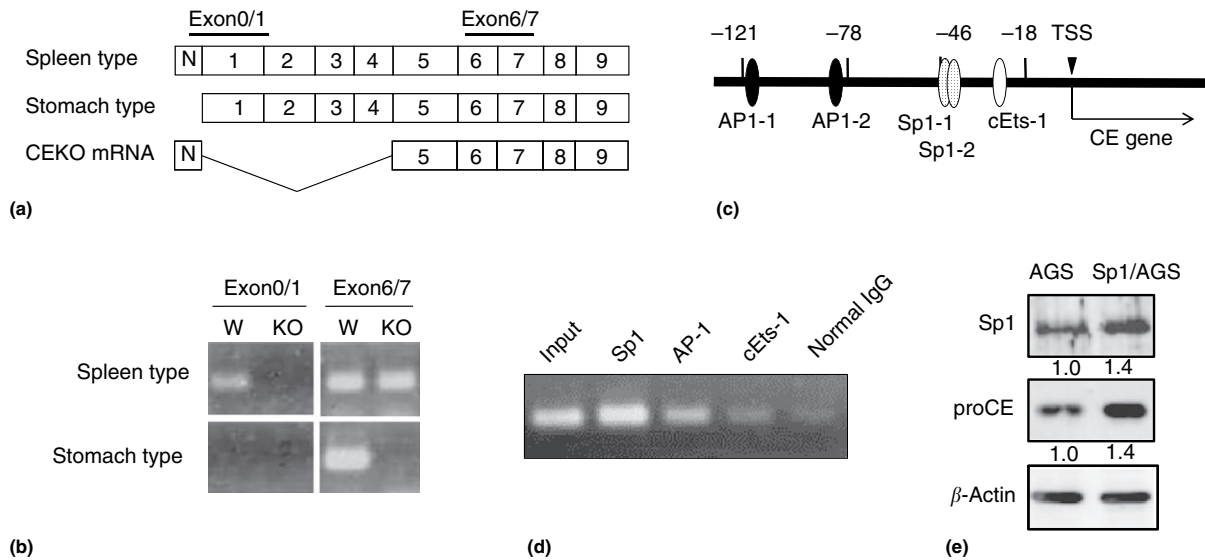


Figure 1 Characterization of the transcripts of mouse cathepsin E gene. (a) A schematic diagram of two types of cathepsin E transcripts in wild-type mice. Total RNA from various tissues of wild type (C57BL/6 J) and cathepsin E-null mice was analyzed by RT-PCR using a forward primer for noncoding exon (N) and a reverse primer in exon 1. The RNA from most of wild-type tissues such as spleen and lung reacted with the primers containing the noncoding exon (spleen-type), whereas the RNA from the stomach did not (stomach-type). The RNA from either type of tissues reacted with downstream primers containing exons 6–7. Numbers in each column indicate noncoding and coding exon. (b) Total RNAs were isolated from spleen and stomach of the wild type and cathepsin E-null mice. RT-PCR was performed with 1 μ g of total RNA. The amplified DNA was visualized on 2% agarose gel. W and KO indicate wild type and cathepsin E-null mice. (c) Binding sites of three possible transcription factors Sp1, Ap-1, and cEts-1 in cathepsin E gene. (d) A chromatin immunoprecipitation (ChIP) assay of the *in vivo* recruitment of Sp1, Ap-1, and cEts-1 to the promoter region of cathepsin E gene in human gastric adenocarcinoma (AGS) cells. DNA purified from the immunoprecipitated chromatin was amplified by PCR. (e) Western blot analysis of cathepsin E protein in wild-type and Sp1-overexpressing AGS cells. The protein from AGS cells and the cells stably transfected with Sp1 was collected and analyzed by a Western blotting analysis. β -Actin was used as a loading control. The density of each protein in Sp1-overexpressing AGS cells is expressed as relative ratios to that in the control cells, as shown below each column.

retains in the *trans*-Golgi region till the cells are activated in response to certain stimuli such as steroid hormone (Yamamoto *et al.*, 1994, 1997; Nishishita *et al.*, 1996). In Chinese hamster ovary cells and normal rat kidney cells expressing exogenous human and rat cathepsin E, respectively, procathepsin E was proteolytically processed to the intermediate form with high mannose-type oligosaccharides and stably retained in the ER and the *cis*- or *medial*-Golgi region (Tsukuba *et al.*, 1993; Yamamoto *et al.*, 1997; Yamamoto, 1999). In mouse L cells and monkey Cos 1 cells expressing human cathepsin E, the expressed protein was exclusively confined to the ER as the proenzyme bearing high mannose-type oligosaccharide chains and the first 48 amino acids of mature cathepsin E was responsible for its retention in the ER (Finley and Kornfeld, 1994). Studies with various cathepsin E mutants also revealed a vital role of propeptide in the correct folding, maturation, and targeting to its final destinations (Yasuda *et al.*, 2005a; Tsukuba *et al.*, 2006a). N-Glycosylation of cathepsin E was also important for its proper folding upon changes in temperature, pH, and redox state, and the complex-type oligosaccharides contribute to the completion of the tertiary structure to maintain its active conformation in the weakly acidic pH environments (Yasuda *et al.*, 1999a).

Autoactivation

Both natural and recombinant procathepsin E are synthesized as catalytically inactive enzymes, but they are rapidly

converted to the mature enzyme by a brief acid treatment, accompanying proteolytic removal of the propeptide (reviewed in, Yamamoto, 1999). Although procathepsin D (Conner, 1989) and pepsinogens (Athauda *et al.*, 1989) were activated through their intermediate forms, procathepsin E was directly converted to the mature form depending on its protein concentration and completely inhibited by pepstatin A (Athauda *et al.*, 1991a; Takeda-Ezaki and Yamamoto, 1993). The results thus indicate that the intermolecular reaction is predominantly involved in one-step activation of the enzyme (reviewed in, Yamamoto, 1995, 1999).

Substrates

Based on nonspecific degradation profiles of various protein substrates in acidic pH, cathepsin E has long been thought to serve as house-keeping enzymes for the acquisition of amino acids. On the other hand, cathepsin E is known to specifically cleave various natural proteins such as α 2-macroglobulin (Shibata *et al.*, 2003), reduced-carboxymethylated ribonuclease A (Athauda and Takahashi, 2002) and big endothelin (Lees *et al.*, 1990), and several biologically active peptides such as substance P, neurokinin (Kageyama, 1993; Kageyama *et al.*, 1995), and oxidized insulin B chain (Athauda *et al.*, 1991b), suggesting that it possesses more specific functions. On the basis of substrate cleavage specificity, many attempts have also been made to construct convenient and selective substrates for this enzyme. To date, several synthetic substrates of cathepsin E have been

Table 1 Kinetic parameters for hydrolysis by cathepsin E of peptide substrates

Enzyme	Substrate	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\mu M s^{-1}$)
MOCac-GSPAFLAK(Dnp)-D-R-NH ₂ ↑	Human cathepsin E	1.90	19.4	10.2 ^a
	Rat cathepsin E	1.79	20.1	11.2 ^a
	Rat cathepsin D	2.65	2.5	0.92 ^a
	Porcine pepsin A	1.44	1.9	1.30 ^a
MOCac-GSSAFLAFK(Dnp)-D-R-NH ₂ ↑	Rat cathepsin E	2.43	1.68	0.69 ^a
	Rat cathepsin D	2.12	0.75	0.35 ^a
	Porcine pepsin A	1.82	0.28	0.15 ^a
MOCac-GKPILFFRLK(Dnp)-D-R-NH ₂ ↑	Human cathepsin E	3.3	35.9	10.9 ^b
	Rat cathepsin D	3.7	57.8	15.6 ^b
MOCac-GKPIIFFRLK(Dnp)-D-R-NH ₂ ↑	Human cathepsin E	3.2	39.1	12.2 ^b
	Rat cathepsin D	3.7	59.6	16.3 ^b
KPIIF(Nph)RL ↑	Human cathepsin E	13 ± 4	98 ± 27	8.0 ± 3.0 ^c
KPIEF(Nph)RL ↑	Human cathepsin E	30 ± 6	130 ± 10	4.3 ^d
PPTIF(Nph)RL ↑	Human cathepsin E	60 ± 3	155 ± 5	2.5 ^d
Mca-AGFSLPAK(Dnp)-D-R-CONH ₂ ↑	Human cathepsin E	19.37	322.5	16.7 ^e

^aYasuda *et al.* (2005b).^bYasuda *et al.* (1999b).^cRao-Naik *et al.* (1995).^dValues were modified on the basis of the data of Jupp *et al.* (1988).^eAbd-Elgalil and Tung, 2010. The hydrolysis of each substrate was determined according to the methods described in the respective references. Dnp, dinitrophenyl; Mca, 7-methylcoumarin-4-acetic acid; MOCac, (7-methoxycoumarin-4-yl)acetyl; Nph, *p*-nitrophenylalanine. Arrows indicate preferred cleavage sites by cathepsin E.

synthesized (Rao-Naik *et al.*, 1995; Yasuda *et al.*, 1999b, 2005b; Abd-Elgalil and Tung, 2010). The kinetic parameters for the hydrolysis of various substrates by cathepsin E, as well as cathepsin D and pepsin A, are summarized in Table 1.

Activators

Although cathepsin E has only a little activity at pH values above 5.5, it was shown to exhibit virtually full activity on synthetic substrates and casein at pH 5.8 in the presence of ATP, where ATP is suggested to stabilize the molecule in such a way to maintain its active conformation (Thomas *et al.*, 1989a). Among nucleotides and inorganic phosphates, GTP and ATP were shown to most profoundly activate cathepsin E activity (about 3.5-fold that in the absence of each compound) (Yamamoto *et al.*, 1991). Recently, some peptide aptamers capable of enhancing cathepsin E activity at neutral pH have been developed by an *in vitro* evolution method based on an inverse substrate function-link-selection technique (Biyani *et al.*, 2011). Of these, the IGCEERSFPNIIIG peptide was shown to activate cathepsin E activity up to 2.6-fold that without the compound.

Inhibitors

Pepstatin is a peptide inhibitor isolated from cultures of *Actinomyces* by Umezawa *et al.* (1970) and is widely used as a potent inhibitor of cathepsin E. Because this inhibitor is a general inhibitor of aspartic peptidases of the A1 family, however, it is not specific for cathepsin E. Also, α 2-macroglobulin

was found to be a nonspecific protein inhibitor for cathepsin E (Thomas *et al.*, 1989b). The degradation of reduced and carboxymethylated ribonuclease A by cathepsin E at pH 5.5 was completely blocked by α 2-macroglobulin, but that of oxidized insulin B chain was not inhibited by this protein (Athauda *et al.*, 1993). The cleavage of peptide substrates such as Pro-Pro-Thr-Ile-Phe-Phe-(4-NO₂)-Arg-Leu by cathepsin E at pH 6.2 was also scarcely inhibited by α 2-macroglobulin (Thomas *et al.*, 1989b). In the course of the study to assess the involvement of cathepsin E in disruption of structural and functional integrity of α 2-macroglobulin in the endolysosome system, this inhibitor was found to be selectively and rapidly cleaved at the Phe811-Leu812 bond in about 100mer downstream of the bait region by cathepsin E below pH 5.0 (Shibata *et al.*, 2003). The natural protein inhibitor specific for cathepsin E, termed *Ascaris* pepsin inhibitor, was first isolated from the roundworm *Ascaris lumbricoides* (Abu-Ereish and Peanasky, 1974; Keilová and Tomášek, 1972). Recently, Kitamura *et al.* (2009) isolated a series of inhibitory peptide aptamers for cathepsin E by using the *in vivo* evolution method and identified the peptide SCGGIIISICIA (Pi101) as the most potent inhibitor of the enzyme (*K_i*; 5 nM) without affecting cathepsin D.

Pathophysiological Functions

Roles in immune responses

Proteolytic processing of the invariant chain (Ii) to class II-associated Ii peptide (CLIP) is known to be essential for the MHC class II antigen presentation. The cleavage of Ii has been

shown to be initiated by cathepsins D and/or E (Maric *et al.*, 1994; Kageyama *et al.*, 1996) and subsequently processed to generate CLIP by cysteine cathepsins L and S in the thymus and the peripheral lymphoid organs, respectively (reviewed in, Chapman, 2006; Colbert *et al.*, 2009). The cysteine cathepsin F is also implicated in CLIP generation in peripheral macrophages (Shi *et al.*, 2000). On the other hand, exogenous antigens, such as ovalbumin (Bennett *et al.*, 1992; Nishioku *et al.*, 2002; Chain *et al.*, 2005), myoglobin (Moss *et al.*, 2005), and tetanus toxin (Burster *et al.*, 2008), were shown to be processed by cathepsin E in various APC. An equivocal role of cathepsin E in class II ovalbumin presentation in macrophages was further substantiated by studies using cathepsin E-deficient macrophages (Kakehashi *et al.*, 2007). Interestingly, although the class II ovalbumin presentation in macrophages was markedly decreased by cathepsin E deficiency, that was enhanced in cathepsin E-deficient dendritic cells, partly due to enhanced phagocytic activity and increased expression of co-stimulatory molecules such as CD86, CD80, and CD40 in these cells (Kakehashi *et al.*, 2007).

A close link between cathepsin E function and immune responses was also supported by the increased expression of cathepsin E in microglia by treatment with interferon- γ without affecting the expression of cathepsins D, B, L, and S (Nishioku *et al.*, 2002). Consistently, the increased expression, maturation, and secretion of cathepsin E in macrophages were also induced by treatment with interferon- γ and lipopolysaccharide (Yanagawa *et al.*, 2006). Notably, cathepsin E-null mice spontaneously developed atopic dermatitis-like skin lesions when kept under conventional conditions (Tsukuba *et al.*, 2003). These mice, like atopic dermatitis patients, showed the strong polarization of naïve T cells to Th2 cells and systemic accumulation of IL-18 and IL-1 β accompanied by enhanced production of IL-4, IL-5, and IgE. In this regard, the significantly decreased cathepsin E expression was also observed in both atopic dermatitis patients and the animal model of NC/Nga mice. Together, considering that the expression of MHC class II molecules in dendritic cells is lower in NC/Nga mice than in BALB/c and DBA mice (Sakai *et al.*, 2012), cathepsin E deficiency is likely to cause aberrant Th2-type immune responses. Additionally, cathepsin E deficiency dramatically increased susceptibility to bacterial infection (Tsukuba *et al.*, 2006b). This is likely to be associated with decreased expression of multiple cell surface Toll-like receptors in macrophages. Similarly, the cell surface expression of chemotactic receptors (CCR-2 and FPRs) and adhesion receptors (CD18 and CD29) was significantly decreased by cathepsin E-deficient macrophages (Tsukuba *et al.*, 2009), thereby leading to their decreased chemotactic and adhesion abilities. In this regard, it has been shown that the impaired adipose tissue development in cathepsin E-null mice fed on a high-fat diet, accompanying reduced body weight and defective development of white and brown adipose tissues, is associated partly with the decreased infiltration of macrophages into white adipose tissues (Kadowaki *et al.*, 2014).

Importantly, cathepsin E deficiency was also shown to induce a novel form of lysosomal storage disorder in macrophages, but not in dendritic cells, which is accompanied by accumulation of major lysosomal membrane sialoglycoproteins such as LAMP-1 and LAMP-2 and elevation of lysosomal

pH (Yanagawa *et al.*, 2007). Additionally, cathepsin E deficiency caused the impairment of the autophagosome-lysosome fusion event in macrophages, accompanying accumulation of aberrant mitochondria and increased oxidative stress (Tsukuba *et al.*, 2013). A schematic representation of the effects of cathepsin E deficiency on the nature and functions of macrophages is shown in Figure 2.

It is noteworthy that cathepsin E expression at protein and message levels in hemopoietic cells, but not in gut, from C57BL/6 J mice is apparently lower than those from 129S2/Sv or Balb/c mice and causes the impairment of ovalbumin presentation in dendritic cells (Tulone *et al.*, 2007). The results suggested that C57BL/6 J mice exhibit tissue-specific cathepsin E deficiency and have poor immune responses to some antigens. Cathepsin E-null mice described here were generated using 129S2/Sv-derived embryonic stem cells and then backcrossed onto the C57BL/6 J strain. Hence, a question arose as to whether the pleiotropic phenotype of cathepsin E-null mice might be autonomous or secondary to other *in vivo* defects. However, considering that, both cathepsin E-null 129S2/Sv and C57BL/6 J mice similarly developed AD-like skin lesions and exhibited the enhanced susceptibility to bacterial infection and that ovalbumin presentation in dendritic cells, but not macrophages, was significantly enhanced by cathepsin E deficiency, it is likely that the observed pleiotropic phenotypes of cathepsin E-null mice are independent of mouse strains.

Roles in Cell Growth and Differentiation

Role in skin

It is generally accepted that dermatological disorders ranging from minor cosmetic problems to life-threatening conditions are commonly due to abnormal differentiation of keratinocytes, a major component of a stratified and keratinized epithelium. In human and rat epidermis, cathepsin E was shown to be localized mainly in keratinocytes and in close vicinity to the inner root sheath of hair follicles (Kawakubo *et al.*, 2011). Remarkably, cathepsin E deficiency caused abnormal keratinocyte differentiation in the epidermis and hair follicle of mice, accompanying the significant expansion of corium and the reduction of subcutaneous tissue and hair follicle. Additionally, cathepsin E deficiency induced significantly reduced expression and altered localization of keratinocyte differentiation-induced proteins, such as keratin 1 and loricrin. The role of cathepsin E in regulation of the epidermal differentiation marker proteins was further verified by experiments with primary cultures of keratinocytes from three different genotypes of mice, wild-type, cathepsin E-null, and cathepsin E-overexpressing transgenic mice. Only cathepsin E-null mice caused delayed differentiation accompanying the reduced expression or the ectopic localization of the differentiation marker proteins. On the contrary, cathepsin E-overexpressing animals enhanced the keratinocyte terminal differentiation process. Cathepsin E is thus likely to be associated with proper epidermis formation and homeostasis.

Roles in gastric development and cancer

The role of cathepsin E in early developmental stages was first demonstrated in various fetal rat tissues (Kageyama *et al.*, 1998).

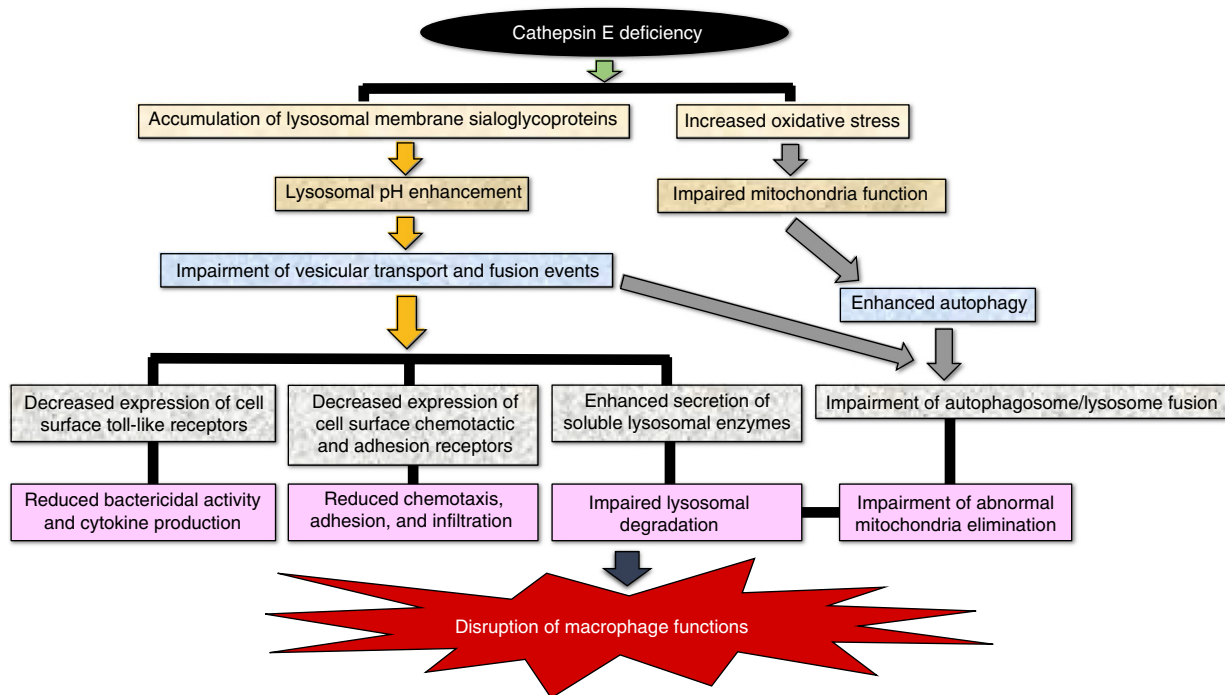


Figure 2 A schematic representation for the conditions manifested in macrophages by cathepsin E deficiency.

Although cathepsin D increased gradually in all the tissues during fetal development, cathepsin E was differentially expressed in various tissues at fetal developmental stages. Cathepsin E expression in liver increased rapidly from a 13-day gestation fetal stage and decreased gradually at later fetal stages, whereas that in stomach, spleen, and thymus began to increase at later fetal stage or the infant stage. Cathepsin E is thus likely to serve as a growth factor in fetal liver and gastrointestinal tissues and to stimulate the development of these tissues. On the other hand, cathepsin E is known to be intensely expressed in human and rat gastric foveolar epithelial cells (Shiraishi *et al.*, 1988; Sakai *et al.*, 1989) and its expression is altered in various gastric cancer types (Saku *et al.*, 1990; Fiocca *et al.*, 1990; Konno-Shimizu *et al.*, 2013). Cathepsin E was expressed mostly in poorly differentiated adenocarcinoma and signet-ring carcinoma, whereas it was barely detectable in mucinous adenocarcinoma, papillary adenocarcinoma, and dysplasia. A considerable part of specimens for moderately differentiated adenocarcinoma, papillary adenocarcinoma, and intestinal metaplasia exhibited cathepsin E expression. Therefore, cathepsin E is suggested to serve as a marker for gastric differentiation and signet-ring carcinoma.

Diagnostic and prognostic values in breast cancer

Cathepsin E is known to be up-regulated and secreted in certain types of normal cells (e.g., activated macrophages and mammary epithelial cells) and abnormal cancer cells. Recent studies have suggested the clinical utility of cathepsin E as a potential diagnostic or prognostic marker in various human cancers, although its clinical significance is controversial (reviewed in, Yamamoto *et al.*, 2012). This conflation is likely to arise from differences in expression levels of some

nonfunctional cathepsin E proteins between cancer cell types, because its expression levels were determined mostly by immunological methods such as immunohistochemistry, tissue microarray, and enzyme-linked immunosorbent assay (ELISA) or by gene expression profiling analysis. In fact, the recent study with patients with breast cancer (BC) showed that serum levels of cathepsin E determined by ELISA were not consistent with those by activity measurement using the synthetic substrate KYS-1 (Kawakubo *et al.*, 2014). The activity levels alone were negatively associated with the stages and progression of this disease. Univariate and multivariate analyses also revealed that the serum levels of cathepsin E activity were significantly correlated with favorable prognostic outcomes of the patients, indicating that the serum cathepsin E activity must serve as an independent biomarker for prognosis of BC. The functional link of cathepsin E expression with mammary carcinogenesis was further verified by *in vivo* and *in vitro* studies with female mice exhibiting different levels of cathepsin E expression. Surprisingly, multiparous cathepsin E-null mice spontaneously developed mammary tumors at a high frequency in an age-dependent manner, concomitant with morphological transformation, and altered growth characteristics of the mammary glands. However, virgin cathepsin E-null mice, like wild-type and cathepsin E-overexpressing transgenic mice, did not develop mammary tumors at all. The observed alterations in mammary glands of multiparous cathepsin E-null mice were associated with epithelial-mesenchymal transition and the activation of canonical Wnt/ β -catenin pathway in mammary cells, where the nuclear translocation and accumulation of noncanonical Wnt5a, a tumor suppressor in mammary carcinogenesis, were strongly induced, thereby resulting in the aberrant trafficking, maturation and secretion of Wnt5a, and

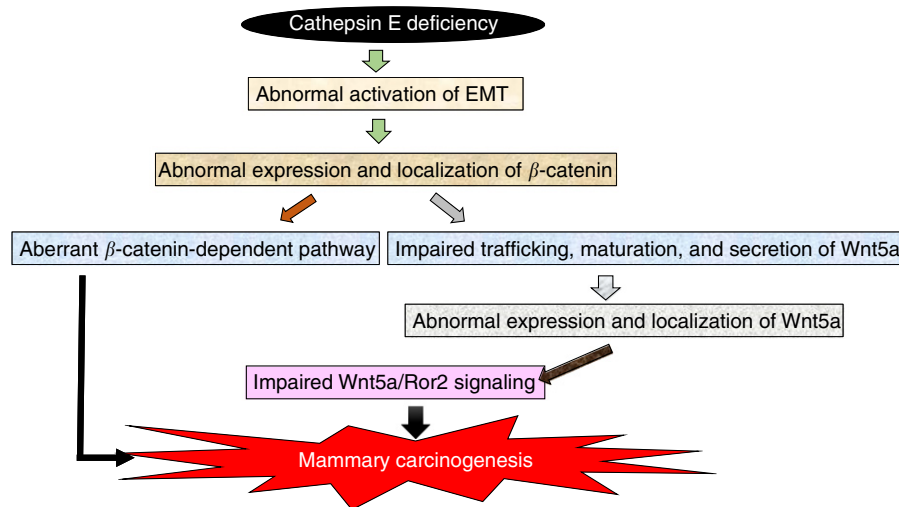


Figure 3 A schematic representation of the possible mechanisms for mammary carcinogenesis by cathepsin E deficiency.

the impairment of the related Wnt5a/Ror2 signaling. Accordingly, cathepsin E is likely to contribute to normal growth and development of mammary glands through proper trafficking and secretion of Wnt5a (Figure 3).

Suppression of cancer progression

It has recently demonstrated that cathepsin E, unlike other peptidases such as cathepsin D and cysteine cathepsins, has the anticancer ability through its manifold functions. These include the induction of tumor cell-specific apoptosis through the generation of tumor necrosis factor-related apoptosis ligand (TRAIL) from the surface of cancer cells, the cytotoxic effects of the infiltrated and activated immune cells, and the production of angiogenesis inhibitor such as endostatin in the vicinity of tumors (reviewed in, Yamamoto *et al.*, 2012). Cathepsin E is also shown to enhance the anticancer activity of doxorubicin on human prostate cancer cells showing resistance to TRAIL-mediated apoptosis by down-regulation of c-FLIP, a FLICE inhibitory protein, and decreased cancer cell proliferation (Yasukochi *et al.*, 2010).

Roles in the neuronal system

Cathepsin E is barely detectable in any brain regions of young rats and mice and is markedly increased in the cerebral cortex and neostriatum of aged rats and in the brain stem of aged mice (Nakanishi *et al.*, 1994; Amano *et al.*, 1995). Up-regulation of cathepsin E in both neurons and microglial cells in the central nervous system was also observed during neuropathological changes associated with transient forebrain ischemia (Nakanishi *et al.*, 1993) and kainate-induced epilepsy (Tominaga *et al.*, 1998), as well in aging. The ischemia-induced up-regulation was mainly observed in degenerating neurons and activated microglia in the hippocampal CA1 region. In aged neurons, highly expressed cathepsin E was co-localized with lipofuscin and APP C-terminal fragments, suggesting its association with lipofuscinosis and altered intracellular APP metabolism (Nakanishi *et al.*, 1997). Additionally, cathepsin E in the brains of patients with AD was observed in senile plaques, besides cerebral microvessels and

microglia (Bernstein and Wiederanders, 1994). Consequently, it is more likely that cathepsin E is involved in the process of neurodegeneration in AD. Interestingly, although cathepsin E-null mice were behaviorally normal when housed communally, they became more aggressive compared with the wild-type littermates when housed individually in a single cage, which was accompanied by an increase in substance P in amygdala, hypothalamus, and periaqueductal gray (Shigematsu *et al.*, 2008). The increased aggressive response was reduced by pretreatment with a neurokinin-1-specific antagonist, suggesting that cathepsin E is associated with the substance P/neurokinin-1 signaling system and thereby regulates the aggressive response of the animals to stressors.

Conclusion Remarks

The review describes the basic information and major functional roles of cathepsin E. The unique and strategic expression and localization of cathepsin E seem likely to be associated with specific functions of each cell type. Cathepsin E has been shown to be involved in a variety of physiological functions, including processing of secretory proteins in various cell types, degradation of biologically active peptide, exogenous antigen processing and presentation, host defense against bacterial infection, and chemotaxis and adhesion of the immune system cells. Growing evidence also indicate that cathepsin E has a potent anticancer activity through multiple mechanisms. In this regard, it has been demonstrated that regression of cathepsin E expression increases the risk of mammary carcinogenesis and links to poor prognosis in BC. Furthermore, cathepsin E is shown to be increased in both neuron and microglia during neuropathological changes associated with aging, transient forebrain ischemia, and epilepsy.

Acknowledgments

Work in the author's laboratory is supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

See also: Protein Synthesis/Degradation: Protein Degradation – Pathological Aspects: Inhibitors of HIV Protease; Protein Synthesis/Degradation: Protein Degradation – Protease Classes: Aspartic Proteases of Alzheimer's Disease: β - and γ -Secretases

References

- Abd-Elgaliel, W.R., Tung, C.H., 2010. Selective detection of cathepsin E proteolytic activity. *Biochimica Et Biophysica Acta* 1800, 1002–1008.
- Abu-Ereish, G.M., Pearsall, R.J., 1974. Pepsin inhibitor from *Ascaris lumbricoides*: isolation, purification, and some properties. *Journal of Biological Chemistry* 249, 1555–1565.
- Amano, T., Nakanishi, H., Oka, M., Yamamoto, K., 1995. Increased expression of cathepsins E and D in reactive microglial cells associated with spongiform degeneration in the brain stem of senescence-accelerated mouse. *Experimental Neurology* 136, 171–182.
- Athauda, S.B.P., Arakawa, H., Nishigai, M., *et al.*, 1993. Inhibition of cathepsin E by α 2-macroglobulin and the resulting structural changes in the inhibitor. *Journal of Biochemistry* 113, 526–530.
- Athauda, S.B.P., Takahashi, K., 2002. Distinct cleavage specificity of human cathepsin E at neutral pH with special preference for Arg-Arg bonds. *Protein and Peptide Letters* 9, 15–22.
- Athauda, S.B.P., Takahashi, T., Inoue, H., Ichinose, M., Takahashi, K., 1991b. Proteolytic activity and cleavage specificity of cathepsin E at the physiological pH as examined towards the B chain of oxidized insulin. *FEBS Letters* 292, 53–56.
- Athauda, S.B.P., Takahashi, T., Kageyama, T., Takahashi, K., 1991a. Autocatalytic processing of procathepsin E to cathepsin E and their structural differences. *Biochemical and Biophysical Research Communications* 175, 152–158.
- Athauda, S.B.P., Tanji, M., Kageyama, T., Takahashi, K., 1989. A comparative study on the NH₂-terminal amino acid sequences and some other properties of six isozymic forms of human pepsinogens and pepsins. *Journal of Biochemistry* 106, 920–927.
- Azuma, T., Liu, W., Vander Laan, D.J., *et al.*, 1992. Human gastric cathepsin E gene. Multiple transcripts result from alternative polyadenylation of the primary transcript of a single gene locus at 1q31-q32. *Journal of Biological Chemistry* 267, 1609–1614.
- Azuma, T., Pals, G., Mohandas, T.K., Couvreur, J.M., Taggart, R.T., 1989. Human gastric cathepsin E. Predicted sequence, localization to chromosome 1, and sequence homology with other aspartic proteinases. *Journal of Biological Chemistry* 264, 16748–16753.
- Bennett, K., Levine, T., Ellis, J.S., *et al.*, 1992. Antigen processing for presentation by class II major histocompatibility complex requires cleavage by cathepsin E. *European Journal of Immunology* 22, 1519–1524.
- Bernstein, H.G., Wiederanders, B., 1994. An immunohistochemical study of cathepsin E in alzheimer-type dementia brains. *Brain Research* 667, 287–290.
- Biyani, M., Futakami, M., Kitamura, K., *et al.*, 2011. *In vitro* selection of cathepsin E-activity-enhancing peptide aptamers at neutral pH. *International Journal of Peptides*. doi:10.1155/2011/834525.
- Burster, T., Reich, M., Zaidi, N., *et al.*, 2008. Cathepsin E regulates the presentation of tetanus toxin C-fragment in PMA activated primary human B cells. *Biochemical and Biophysical Research Communications* 377, 1299–1303.
- Chain, B.M., Free, P., Medd, P., *et al.*, 2005. The expression and function of cathepsin E in dendritic cells. *Journal of Immunology* 174, 1791–1800.
- Chapman, H.A., 2006. Endosomal proteases in antigen presentation. *Current Opinion in Immunology* 18, 78–84.
- Chlabicz, M., Gacko, M., Worowska, A., Łapiński, R., 2011. Cathepsin E (EC 3.4.23.34) – A review. *Folia Histochemica Et Cytobiologica* 49, 547–557.
- Colbert, J.D., Matthews, S.P., Miller, G., Watts, C., 2009. Diverse regulatory roles for lysosomal proteases in the immune response. *European Journal of Immunology* 39, 2955–2965.
- Conner, G.E., 1989. Isolation of procathepsin D from mature cathepsin D by pepstatin affinity chromatography. Autocatalytic proteolysis of the zymogen form of the enzyme. *Biochemical Journal* 263, 601–604.
- Cook, M., Caswell, R.C., Richards, R.J., Kay, J., Tatnell, P.J., 2001. Regulation of human and mouse procathepsin E gene expression. *European Journal of Biochemistry* 268, 2658–2668.
- Finley, E., Kornfeld, S., 1994. Subcellular localization and targeting of cathepsin E. *Journal of Biological Chemistry* 269, 31259–31266.
- Finzi, G., Cornaggia, M., Capella, C., *et al.*, 1993. Cathepsin E in follicle associated epithelium of intestine and tonsils: Localization to M cells and possible role in antigen processing. *Histochemistry* 99, 201–211.
- Fiocca, R., Villani, L., Tenti, P., *et al.*, 1990. The foveolar cell component of gastric cancer. *Human Pathology* 21, 260–270.
- Fowler, S.D., Kay, J., Dunn, B.M., Tatnell, P.J., 1995. Monomeric human cathepsin E. *FEBS Letters* 366, 72–74.
- Hara, K., Fukuyama, K., Sakai, H., Yamamoto, K., Epstein, W.L., 1993. Purification and immunohistochemical localization of aspartic proteinases in rat epidermis. *Journal of Investigative Dermatology* 100, 394–399.
- Henningson, F., Yamamoto, K., Saftig, P., *et al.*, 2005. A role for cathepsin E in the processing of mast-cell carboxypeptidase A. *Journal of Cell Science* 118, 2035–2042.
- Jupp, R.A., Richards, A.D., Kay, J., *et al.*, 1988. Identification of the aspartic proteinases from human erythrocyte membranes and gastric mucosa (slow-moving proteinase) as catalytically equivalent to cathepsin E. *Biochemical Journal* 254, 895–898.
- Kadowaki, T., Kido, M.A., Hatakeyama, J., *et al.*, 2014. Defective adipose tissue development associated with hepatomegaly in cathepsin E-deficient mice fed a high-fat diet. *Biochemical and Biophysical Research Communications* 446, 212–217.
- Kageyama, T., 1993. Rabbit procathepsin E and cathepsin E. Nucleotide sequence of cDNA, hydrolytic specificity for biologically active peptides and gene expression during development. *European Journal of Biochemistry* 216, 717–728.
- Kageyama, T., Ichinose, M., Tsukada, S., *et al.*, 1992. Gastric procathepsin E and progastricsin from guinea pig. Purification, molecular cloning of cDNAs, and characterization of enzymatic properties, with special reference to procathepsin E. *Journal of Biological Chemistry* 267, 16450–16459.
- Kageyama, T., Ichinose, M., Yonezawa, S., 1995. Processing of the precursors to neurotensin and other bioactive peptides by cathepsin E. *Journal of Biological Chemistry* 270, 19135–19140.
- Kageyama, T., Takahashi, K., 1980. A cathepsin D-like acid proteinase from human gastric mucosa. Purification and characterization. *Journal of Biochemistry* 87, 725–735.
- Kageyama, T., Tatematsu, M., Ichinose, M., *et al.*, 1998. Development-dependent expression of cathepsins D and E in various rat tissues, with special reference to the high expression of cathepsin E in fetal liver. *Zoological Science* 15, 517–523.
- Kageyama, T., Yonezawa, S., Ichinose, M., Miki, K., Moriyama, A., 1996. Potential sites for processing of the human invariant chain by cathepsins D and E. *Biochemical and Biophysical Research Communications* 223, 549–553.
- Kakehashi, H., Nishioku, T., Tsukuba, T., *et al.*, 2007. Differential regulation of the nature and functions of dendritic cells and macrophages by cathepsin E. *Journal of Immunology* 179, 5728–5737.
- Kawakubo, T., Okamoto, K., Iwata, J., *et al.*, 2007. Cathepsin E prevents tumor growth and metastasis by catalyzing the proteolytic release of soluble TRAIL from tumor cell surface. *Cancer Research* 67, 10869–10878.
- Kawakubo, T., Yasukochi, A., Okamoto, K., *et al.*, 2011. The role of cathepsin E in terminal differentiation of keratinocytes. *Biological Chemistry* 392, 571–585.
- Kawakubo, T., Yasukochi, A., Toyama, T., *et al.*, 2014. Regression of cathepsin E expression increases the risk of mammary carcinogenesis and links to poor prognosis in breast cancer. *Carcinogenesis* 35, 714–726.
- Keilová, H., Tomášek, V., 1972. Effect of pepsin inhibitor from *Ascaris Lumbricoides* on cathepsin D and E. *Biochimica Et Biophysica Acta* 284, 461–464.
- Kitamura, K., Yoshida, C., Kinoshita, Y., *et al.*, 2009. Development of systemic *in vitro* evolution and its application to generation of peptide-aptamer-based inhibitors of cathepsin E. *Journal of Molecular Biology* 387, 1186–1198.
- Konno-Shimizu, M., Yamamichi, N., Inada, K., *et al.*, 2013. Cathepsin E is a marker of gastric differentiation and signet-ring cell carcinoma of stomach: A novel suggestion on gastric tumorigenesis. *PLoS One* 8, e56766. doi:10.1371/journal.pone.0056766.
- Lapresle, C., Webb, T., 1962. The purification and properties of a proteolytic enzyme, rabbit cathepsin E, and further studies on rabbit cathepsin D. *Biochemical Journal* 84, 455–462.
- Lees, W.E., Kalinka, S., Meech, J., *et al.*, 1990. Generation of human endothelin by cathepsin E. *FEBS Letters* 273, 99–102.
- Louis, J.M., Ishima, R., Torchia, D.A., Weber, I.T., 2007. HIV-protease: structure, dynamics, and inhibition. *Advances in Pharmacology* 55, 261–298.
- Maric, M.A., Taylor, M.D., Blum, J.S., 1994. Endosomal aspartic proteinases are required for invariant-chain processing. *Proceedings of the National Academy of Sciences of the United States of America* 91, 2171–2175.

- Moss, C.X., Villadangos, J.A., Watts, C., 2005. Destructive potential of the aspartyl protease cathepsin D in MHC class II-restricted antigen processing. *European Journal of Immunology* 35, 3442–3451.
- Muto, N., Murayama-Arai, K., Tani, S., 1983. Purification and properties of a cathepsin D-like acid proteinase from rat gastric mucosa. *Biochimica Biophysica Et Acta* 745, 61–69.
- Muto, N., Yamamoto, M., Tani, S., 1987. Comparative studies of two types of acid proteases from rat gastric mucosa and spleen. *Journal of Biochemistry* 101, 1069–1075.
- Nakanishi, H., Amano, T., Sastradipura, D.F., *et al.*, 1997. Increased expression of cathepsins E and D in neurons of the aged rat brain and their colocalization with lipofuscin and carboxy-terminal fragments of alzheimer amyloid precursor protein. *Journal of Neurochemistry* 68, 739–749.
- Nakanishi, H., Tominaga, K., Amano, T., *et al.*, 1994. Age-related changes in activities and localizations of cathepsins D, E, B, and L in the rat brain tissues. *Experimental Neurology* 126, 119–128.
- Nakanishi, H., Tsukuba, T., Kondou, T., Tanaka, T., Yamamoto, K., 1993. Transient forebrain ischemia induces increased expression and specific localization of cathepsins E and D in rat hippocampus and neostriatum. *Experimental Neurology* 121, 215–223.
- Navia, M.A., MaKeever, B.M., 2006. A role for the aspartyl protease from the human immunodeficiency virus T type 1 (HIV-1) in the orchestration of virus assembly. *Annals of the New York Academy of Sciences* 616, 73–85.
- Nishioku, T., Hashimoto, K., Yamashita, K., *et al.*, 2002. Involvement of cathepsin E in exogenous antigen presentation in primary cultured murine microglia. *Journal of Biological Chemistry* 277, 4816–4822.
- Nishishita, K., Sakai, H., Sakai, E., Kato, Y., Yamamoto, K., 1996. Age-related and dexamethasone-induced changes in cathepsins E and D in rat thymic and splenic cells. *Archives of Biochemistry and Biophysics* 338, 349–358.
- Okamoto, K., Okamoto, Y., Kawakubo, T., *et al.*, 2012. Role of the transcription factor Sp1 in regulating the expression of the murine cathepsin E gene. *Journal of Biochemistry* 151, 263–272.
- Okamoto, K., Yu, H., Misumi, Y., Ikehara, Y., Yamamoto, K., 1995. Isolation and sequencing of two cDNA clones encoding rat spleen cathepsin E and analysis of the activation of purified procathepsin E. *Archives of Biochemistry and Biophysics* 322, 103–111.
- Patel, M., Mandava, N.K., Vadlapatla, R.K., Mitra, A.K., 2013. Recent patents and emerging therapeutics for HIV infections: A focus on protease inhibitors. *Pharmaceutical Patent Analyst* 2, 513–538.
- Rao-Naik, C., Curuprasad, K., Batley, B., *et al.*, 1995. Exploring the binding preferences/specificity in the active site of human cathepsin E. *Proteins: Structure, Function and Genetics* 22, 168–181.
- Sakai, H., Kato, Y., Yamamoto, K., 1992. Synthesis and intracellular distribution of cathepsins E and D in differentiating murine friend erythroleukemia cells. *Archives of Biochemistry and Biophysics* 294, 412–417.
- Sakai, H., Saku, T., Kato, Y., Yamamoto, K., 1989. Quantitation and immunohistochemical localization of cathepsins E and D in rat tissues and blood cells. *Biochimica Et Biophysica Acta* 991, 367–375.
- Sakai, T., Shuto, E., Taki, T., *et al.*, 2012. Decreased expression of MHC class II and cathepsin E in dendritic cells might contribute to impaired induction of antigen-specific T cell response in NC/Nga mice. *Journal of Medical Investigation* 59, 95–101.
- Saku, T., Sakai, H., Shibata, Y., Kato, Y., Yamamoto, K., 1991. An immunocytochemical study on distinct intracellular localization of cathepsin E and cathepsin D in human gastric cells and various rat cells. *Journal of Biochemistry* 110, 956–964.
- Saku, T., Sakai, H., Tsuda, N., *et al.*, 1990. Cathepsins D and E in normal, metastatic, dysplastic, and carcinomatous gastric tissue: An immunohistochemical study. *Gut* 31, 1250–1255.
- Samloff, I.M., Taggart, R.T., Shiraiishi, T., *et al.*, 1987. Slow moving proteinase. isolation, characterization, and immunohistochemical localization in gastric mucosa. *Gastroenterology* 93, 77–84.
- Sastradipura, D.F., Nakanishi, H., Tsukuba, T., *et al.*, 1998. Identification of cellular compartments involved in processing of cathepsin E in primary cultures of rat microglia. *Journal of Neurochemistry* 70, 2045–2056.
- Shi, G.-P., Bryant, B.A.R., Riese, R., *et al.*, 2000. Role for cathepsin F in invariant chain processing and major histocompatibility complex class II peptide loading by macrophages. *Journal Experimental of Medicine* 191, 1177–1186.
- Shibata, M., Sakai, H., Sakai, E., *et al.*, 2003. Disruption of structural and functional integrity of α 2-macroglobulin by cathepsin E. *European Journal of Biochemistry* 270, 1189–1198.
- Shigematsu, N., Fukuda, T., Yamamoto, T., *et al.*, 2008. Association of cathepsin E deficiency with the increased territorial aggressive response of mice. *Journal of Neurochemistry* 105, 1394–1404.
- Shiraiishi, T., Samloff, I.M., Taggart, R.T., Stemmermann, G.N., 1988. Slow moving proteinase in gastric cancer and its relationship to pepsinogens I and II. *Digestive Diseases and Sciences* 33, 1466–1472.
- Takeda, M., Ueno, E., Kato, Y., Yamamoto, K., 1986. Isolation, and catalytic and immunochemical properties of cathepsin D-like acid proteinase from rat erythrocytes. *Journal of Biochemistry* 100, 1269–1277.
- Takeda-Ezaki, M., Yamamoto, K., 1993. Isolation and biochemical characterization of procathepsin E from human erythrocyte membranes. *Archives of Biochemistry and Biophysics* 304, 352–358.
- Tarasova, N.I., Szecsi, P.B., Foltmann, B., 1986. An aspartic proteinase from human erythrocytes is immunochemically indistinguishable from a non-pepsin, electrophoretically slow moving proteinase from gastric mucosa. *Biochimica Et Biophysica Acta* 880, 96–100.
- Tatnell, P., Lee, W.E., Kay, J., 1997. Cloning, expression and characterization murine procathepsin E. *FEBS Letters* 408, 62–66.
- Tatnell, P.J., Roth, W., Deussing, J., Peters, C., Kay, J., 1998. Mouse procathepsin E gene: Molecular organization and chromosomal localization. *Biochimica Et Biophysica Acta* 1398, 57–66.
- Thomas, D.J., Richards, A.D., Kay, J., 1989b. Inhibition of aspartic proteinases by α 2-macroglobulin. *Biochemical Journal* 259, 905–907.
- Thomas, D.J., Richards, R.A., Jupp, R., *et al.*, 1989a. Stabilization of Cathepsin E by ATP. *FEBS Letters* 243, 145–148.
- Tominaga, K., Nakanishi, H., Yasuda, Y., Yamamoto, K., 1998. Excitotoxin-induced neuronal death is associated with response of a unique intracellular aspartic proteinase, cathepsin E. *Journal of Neurochemistry* 71, 2574–2584.
- Tsukada, S., Ichinose, M., Miki, K., *et al.*, 1992. Tissue- and cell-specific control of guinea pig cathepsin E gene expression. *Biochemical and Biophysical Research Communications* 187, 1401–1408.
- Tsukuba, T., Hori, H., Azuma, T., *et al.*, 1993. Isolation and characterization of recombinant human cathepsin E expressed in chinese hamster ovary cells. *Journal of Biological Chemistry* 268, 7276–7282.
- Tsukuba, T., Ikeda, S., Okamoto, K., *et al.*, 2006a. Characterization of rat cathepsin E and mutants with changed active-site residues and lacking propeptides and N-glycosylation, expressed in human embryonic kidney 293 T cells. *FEBS Journal* 273, 219–229.
- Tsukuba, T., Okamoto, K., Okamoto, Y., *et al.*, 2003. Association of cathepsin E deficiency with development of atopic dermatitis. *Journal of Biochemistry* 134, 893–902.
- Tsukuba, T., Sakai, H., Yamada, M., *et al.*, 1996. Biochemical properties of the monomeric mutant of human cathepsin E expressed in chinese hamster ovary cells: Comparison with dimeric forms of the natural and recombinant cathepsin E. *Journal of Biochemistry* 119, 126–134.
- Tsukuba, T., Yamamoto, S., Yanagawa, M., *et al.*, 2006b. Cathepsin E-deficient mice show increased susceptibility to bacterial infection associated with the decreased expression of multiple cell surface toll-like receptors. *Journal of Biochemistry* 140, 57–66.
- Tsukuba, T., Yanagawa, M., Kadowaki, T., *et al.*, 2013. Cathepsin E deficiency impairs autophagic proteolysis in macrophages. *PLoS One* 8, e82415.
- Tsukuba, T., Yanagawa, M., Okamoto, K., *et al.*, 2009. Impaired chemotaxis and cell adhesion due to decrease in several cell-surface receptors in cathepsin E-deficient macrophages. *Journal of Biochemistry* 145, 565–573.
- Tulone, C., Tsang, J., Prokopowicz, Z., Grosvenor, N., Chain, B., 2007. Natural cathepsin E deficiency in the immune system of C57BL/6 J mice. *Immunogenesis* 59, 927–935.
- Umezawa, H., Aoyagi, T., Morishima, H., *et al.*, 1970. Pepstatin, a new pepsin inhibitor produced by actinomycetes. *Journal of Antibiotics* 23, 259–262.
- Yago, N., Bowers, W., 1975. Unique cathepsin D-type proteases in rat thoracic duct lymphocytes and in rat lymphoid tissues. *Journal of Biological Chemistry* 250, 4749–4754.
- Yamamoto, K., 1995. Cathepsin E and cathepsin D: Biosynthesis, processing and subcellular location. In: Takahashi, K. (Ed.), *Aspartic Proteinases: Structure, Function, Biology and Biomedical Implications*. New York, NY: Plenum Press, pp. 223–229.
- Yamamoto, K., 1999. Cathepsin E and cathepsin D. In: Turk, V. (Ed.), *Proteases: New Perspectives*. Basel/Switzerland: Birkhäuser Verlag, pp. 59–71.
- Yamamoto, K., Kamata, O., Katsuda, N., Kato, K., 1980. Immunochemical difference between cathepsin D and cathepsin E-like enzyme from rat spleen. *Journal of Biochemistry* 87, 511–516.

- Yamamoto, K., Katsuda, N., Himeno, M., Kato, K., 1979. Cathepsin D of rat spleen: Affinity purification and properties of two types of cathepsin D. *European Journal of Biochemistry* 95, 459–467.
- Yamamoto, K., Katsuda, N., Kato, K., 1978. Affinity purification and properties of cathepsin E-like acid proteinase from rat spleen. *European Journal of Biochemistry* 92, 499–508.
- Yamamoto, K., Kawakubo, T., Yasukochi, A., Tsukuba, T., 2012. Emerging roles of cathepsin E in host defense mechanisms. *Biochimica Et Biophysica Acta* 1824, 105–112.
- Yamamoto, K., Marchesi, V.T., 1984. Purification and characterization of acid proteinase from human erythrocyte membranes. *Biochimica Biophysica Acta* 790, 208–218.
- Yamamoto, K., Nakanishi, H., Tsukuba, T., *et al.*, 1997. Biosynthesis and trafficking of cathepsin E. In: Hopsu-Have, V.K. (Ed.), *Proteolysis in Cell Functions*. Amsterdam: IOS Press, pp. 215–222.
- Yamamoto, K., Sakai, H., Ezaki, M., Kato, Y., 1994. Biosynthesis, processing and localization of cathepsin E. In: Katunuma, N., Suzuki, K., Travis, J., Fritz, H. (Eds.), *Biological Functions of Proteases and Inhibitors*. Tokyo: Japan Scientific Society Press, pp. 97–107.
- Yamamoto, K., Sakai, H., Ueno, E., Kato, Y., 1991. Biological significance and activity control of cathepsin E. In: Dunn, B.M. (Ed.), *Structure and Function of the Aspartic Proteinases*. New York, NY: Plenum Press, pp. 297–306.
- Yamamoto, K., Ueno, E., Uemura, H., Kato, Y., 1987. Biochemical and immunochemical similarity between erythrocyte membrane aspartic proteinase and cathepsin E. *Biochemical and Biophysical Research Communications* 148, 267–272.
- Yanagawa, M., Tsukuba, T., Nishioku, T., *et al.*, 2007. Cathepsin E deficiency induces a novel form of lysosomal storage disorder showing the accumulation of lysosomal membrane sialoglycoproteins and the elevation of lysosomal pH in macrophages. *Journal of Biological Chemistry* 282, 1851–1862.
- Yanagawa, M., Tsukuba, T., Okamoto, K., *et al.*, 2006. Up-regulation, enhanced maturation, secretion of cathepsin E in mouse macrophages treated by interferon- γ or lipopolysaccharide. *Journal of Oral Biosciences* 48, 218–225.
- Yasuda, Y., Ikeda, S., Sakai, H., *et al.*, 1999a. Role of N-glycosylation in cathepsin E. A comparative study of cathepsin E with distinct N-linked oligosaccharides and its nonglycosylated mutants. *European Journal of Biochemistry* 266, 383–391.
- Yasuda, Y., Kageyama, T., Akamine, A., *et al.*, 1999b. Characterization of new fluorogenic substrates for the rapid and sensitive assay of cathepsins E and D. *Journal of Biochemistry* 125, 1137–1143.
- Yasuda, Y., Kohmura, K., Kadowaki, T., Tsukuba, T., Yamamoto, K., 2005b. A new selective substrate for cathepsin E based on the cleavage site sequence of α 2-macroglobulin. *Biological Chemistry* 386, 299–305.
- Yasuda, Y., Tsukuba, T., Okamoto, K., Kadowaki, T., Yamamoto, K., 2005a. The role of the cathepsin E propeptide in correct folding, maturation and sorting to the endosome. *Journal of Biochemistry* 138, 621–630.
- Yasukochi, A., Kawakubo, T., Nakamura, S., Yamamoto, K., 2010. Cathepsin E enhances anticancer activity of doxorubicin on human prostate cancer cells showing resistance to TRAIL-mediated apoptosis. *Biological Chemistry* 391, 947–958.
- Yee, C.S.K., Yao, Y., Li, P., *et al.*, 2004. Cathepsin E: Novel target for regulation by class II transactivator. *Journal of Immunology* 172, 5528–5534.
- Yoshimine, Y., Tsukuba, T., Isobe, R., *et al.*, 1995. Specific immunocytochemical localization of cathepsin E at the ruffled border membrane of active osteoclasts. *Cell and Tissue Research* 281, 85–91.
- Zaidi, N., Kalbacher, H., 2008. Cathepsin E: A mini review. *Biochemical and Biophysical Research Communications* 367, 517–522.

Relevant Website

<http://merops.sanger.ac.uk>
MEROPS Database.

Metalloproteases Meprin α and Meprin β in Health and Disease

C Tredup and C Becker-Pauly, University of Kiel, Kiel, Germany

© 2016 Elsevier Inc. All rights reserved.

Introduction

Meprins were first described by Judith Bond and colleagues in 1980 who discovered strong proteolytic activity in mouse and rat kidney that could not be assigned to any known protease (Beynon *et al.*, 1981). The term 'meprin' was created meaning metalloprotease from renal tissue.

At this time an interesting clinical observation was made when patients after pancreatic surgery still exhibited proteolytic activity against the so-called PABA-peptide (N-benzoyl-L-tyrosyl-p-aminobenzoic acid), a diagnostic marker for chymotrypsin activity (Sterchi *et al.*, 1983). Thereafter, the responsible protease was called PPH (PABA-peptide hydrolase). Later on, based on sequence similarities, PPH was renamed human meprin.

In mouse and man, two related proteases meprin α and meprin β exist that are encoded by different genes. The meprin α and meprin β genes in mice are located on chromosomes 17 and 18, respectively (Bond *et al.*, 1984), whereas human meprin α and meprin β are located on chromosomes 6 and 18, respectively (Bond *et al.*, 1995). Interestingly, the meprin α genes in rodents and mammals are in close proximity to the MHC H2 complex, which might be a link to immunological functions of these enzymes, discussed later in this article.

All phylogenetic analyses so far provided evidence that meprin metalloproteases evolved in vertebrates only (Becker-Pauly *et al.*, 2009). Although other MAM (meprin 5A protein tyrosine phosphatase μ) domain containing astacins are existing already in invertebrates it was demonstrated by Becker-Pauly and colleagues that these proteases, for example, HMP2 from *Hydra vulgaris* (Yan *et al.*, 2000), did not cluster in the same branch with meprins when comparing the sequences of many astacins from different species. Thus independent achievement of MAM domains by astacin proteases probably occurred several times. However, recent sequencing projects revealed that two different meprins already exist in *Ciona intestinalis*, also called vase tunicate, an animal that is a phylogenetic ancestor of vertebrates (Figure 1). These proteins contain not only a MAM but also a TRAF (tumor-necrosis-factor-receptor-associated factor) domain, indicating that this combination is the minimal requirement for the group of meprin metalloproteases. This finding opens new ideas about the biological functions of meprins, which may be associated with immunological processes and tissue development, as described later on.

Structural Features of Meprin α and Meprin β

Meprins belong to the astacin family of metalloproteases and the metzincin superfamily (Stocker *et al.*, 1995). As all members of the astacin family, meprin α and meprin β contain the characteristic zinc-binding motif HExxHxxGxxHxxxRxDR that stabilizes the central zinc ion. A conserved sequence in close

proximity to the active site cleft, namely the Met-turn, includes a tyrosine residue that serves as an additional zinc ligand. However, among astacins meprins exhibit unique structural properties.

Both meprin subunits are expressed as highly glycosylated, disulfide linked multi-domain proteins comprising an amino-terminal signal peptide, which directs the protease into the endoplasmatic reticulum (Figure 2(a)). The signal peptide is followed by the propeptide that keeps meprins in the inactive state. Therefore, upon secretion meprin α and meprin β undergo posttranslational modification to gain their full activity. The enzymatic removal of the propeptide can be carried out by several serine proteases, such as the members of the kallikrein-related peptidases (KLKs) (Ohler *et al.*, 2010). C-terminal of the propeptide there is an astacin-like protease domain, a MAM domain, a TRAF domain, an EGF (epidermal growth factor)-like domain, a transmembrane region, and finally a cytosolic tail follows. In contrast to meprin β , meprin α contains an additional domain, the so-called inserted domain, which is cleaved by furin during the secretory pathway (Figure 2(a)). This leads to the release of the soluble ectodomain of meprin α into the extracellular space. Interestingly, as demonstrated in Figure 1, such an inserted domain is missing in meprins from *Ciona intestinalis*. However, this is probably due to incomplete sequencing, since the signal peptides are not entirely annotated as well.

Human meprin α and meprin β are secreted as homo-dimers. Until recently, it was unrevealed how exactly these homo-dimers are stabilized. The crystal structure of meprin β nicely uncovered this question by showing that a single disulfide-bridge between MAM domains is sufficient to form a meprin α or a meprin β homo-dimer (Arolas *et al.*, 2012) (Figures 2(b) and 2(c)). Upon secretion, meprin α homo-dimers oligomerize non-covalently to huge complexes, which can reach a size of up to 6 MDa (Bertenshaw *et al.*, 2003). The lack of the 'inserted' domain leads to the expression of membrane-bound meprin β homo-dimers, which can be released from the cell surface by ectodomain shedding through the metalloproteases ADAM10 or ADAM17 (Hahn *et al.*, 2003; Jefferson *et al.*, 2013).

Inhibition of Meprin Metalloproteases

The mechanisms for regulation of protease's activity involve endogenous activators and inhibitors and an imbalance consequently leads to a range of pathologic events. Less is known about endogenous inhibitors of meprin α and meprin β . Fetuin-A, a negative acute-phase plasma protein has been identified as endogenous inhibitor for meprin α and meprin β activity (Hedrich *et al.*, 2010). Additionally, nephrosin which is a fetuin-A homologue from carp is able to inhibit the catalytic activity of both human meprins. Further, cystatin C has been shown to regulate the activity of meprin α but not of

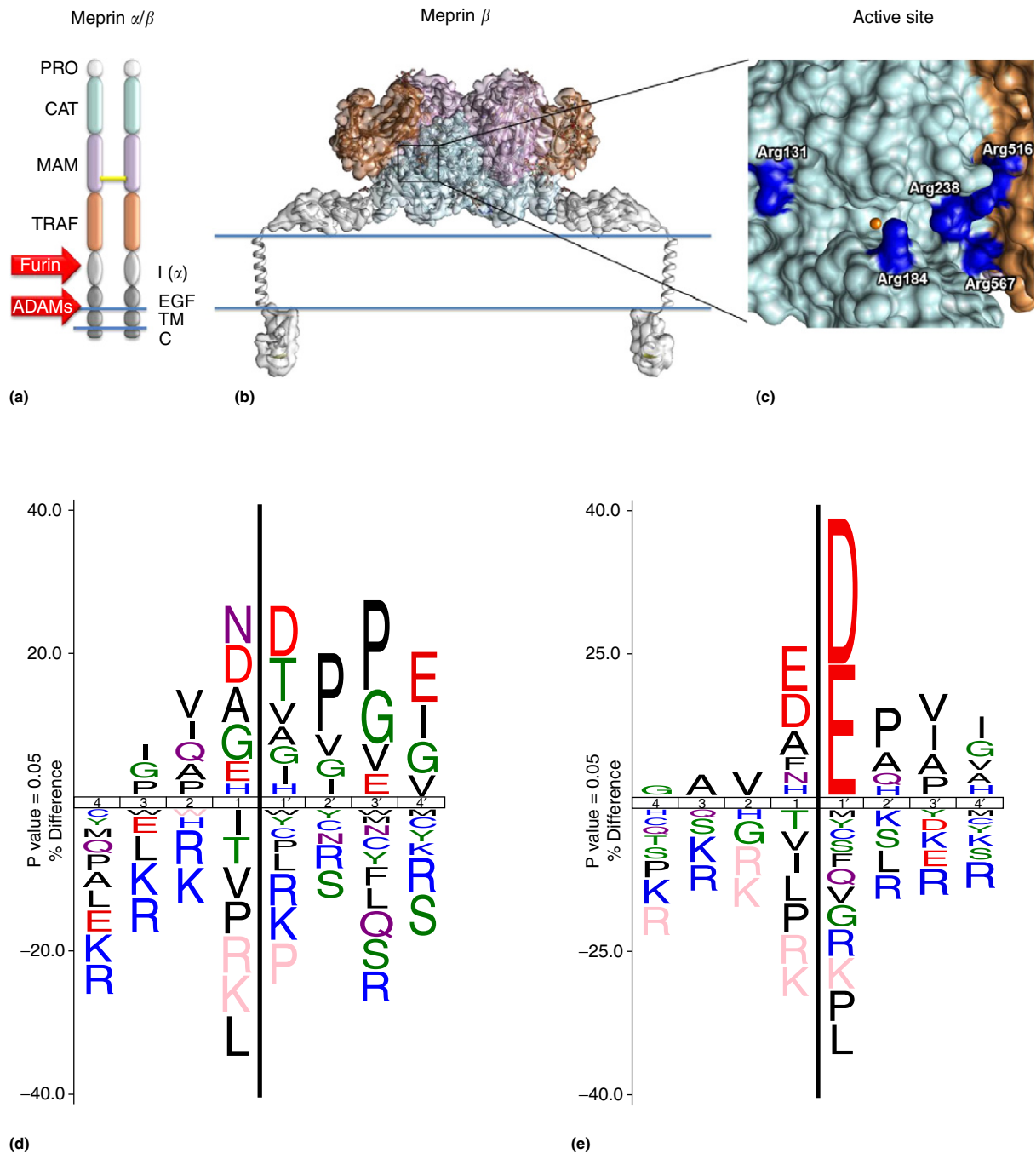


Figure 2 Structure and cleavage specificity of meprins. (a) Cartoon of domain organization of meprins before proteolytic processing. Homodimers are linked by a disulfide-bridge between the MAM domains. Only meprin α/α contains an Inserted (I) domain, which is cleaved by furin in the secretory pathway. Meprin β/β can be shed from the cell surface by ADAM. PRO (propeptide), CAT (catalytic), MAM, TRAF, EGF-like, TM (transmembrane region), C (cytosolic). (b) Model of membrane human meprin β/β dimer based on the crystal structure of the ectodomain (Arolas *et al.*, 2012). Color code according to (a). (c) Magnification of the active site cleft of meprin β reveals several positively charged residues critical for the cleavage specificity. Cleavage site specificities of human meprin α (d) and meprin β (e) based on peptide library cleavage assays (Becker-Pauly *et al.*, 2011). Amino acids are presented in the one letter code and frequency of certain residues in prime and non-prime positions is indicated by the size. Figures were generated using icelogo.

meprin β (Hedrich *et al.*, 2010). Interestingly, fetuin-A and cystatin C belong to the cystatin superfamily of protease inhibitors, indicating a general mechanism of inhibition by this group of proteins.

In addition to endogenous inhibitors, several small compounds were identified capable of reducing meprin activity (Kruse *et al.*, 2004). Most of these molecules are hydroxamate derivatives, which exhibit strong zinc-chelating

capacity, therefore often regulating a broad range of metalloproteases. However, targeting meprins under certain pathological conditions need the development of highly specific compounds to avoid cross-reactivity with other metalloproteases, which will be future challenge in the field (Madoux *et al.*, 2014).

Cleavage Specificity and Identification of Substrates by Proteomics

Meprins display broad expression pattern in humans, including brain, skin, intestine, and kidney (Broder and Becker-Pauly, 2013). However, only few substrates have been described and little was known about the cleavage specificity until recently, when the proteomics approaches PICS (proteomic identification of cleavage site specificity) and TAILS (terminal amine isotopic labeling of substrates) were applied for meprin α and meprin β . Using PICS a strong preference for negatively charged amino acid residues in the P1' position was identified, although more prominent for meprin β than for meprin α (Becker-Pauly *et al.*, 2011; Figures 2(d) and 2(e)). This cleavage specificity was found to be unique for the astacin family and, to date, it seems to be exclusive among extracellular proteases. Interestingly, most of the substrates of meprin β contain a glutamate or aspartate in the P1' position. This striking cleavage specificity becomes even more evident as meprin β is capable of efficiently cleaving completely negatively charged sequence motifs (Broder and Becker-Pauly, 2013). The recently solved crystal structure of meprin β helps to understand this special feature. Several positively charged arginine residues within the active site cleft of meprin β perfectly interact with the negatively charged residues of the substrate leading to the striking cleavage specificity (Figure 2(c)). This unique property combined with the knowledge of the structure of meprin β might help to design highly specific inhibitors.

The application of the mass spectrometry-based proteomics approach TAILS was a step forward in meprin research. This technique enabled the identification of 151 new substrates of meprin α and meprin β , and revealed complex networks of proteolytic events in the protease web (Jefferson *et al.*, 2013). This knowledge contributes to a deeper understanding of the physiologic and pathologic role of meprin α and meprin β and might be beneficial for the development of specific drugs.

Physiologic and Pathologic Role of Meprin α and Meprin β

Meprin in Inflammation

Meprins are highly expressed in rodent kidney and represent up to 5% of total brush border membrane protein (Craig *et al.*, 1987). Therefore changes in their expression levels often lead to pathologies. The first indication for meprin's role in acute renal failure was given when mice of the strains C3H/He and CBA/Ca developed less severe forms of kidney disease after injury, compared to other mouse strains (Trachtman *et al.*, 1995). Interestingly, the mouse strains C3H/He and

CBA/Ca were shown to express meprin β , but no meprin α in proximal tubular cells (Beynon and Bond, 1983). In healthy kidney, meprin α is secreted into the lumen of the proximal tubule, while meprin β is tethered to the apical plasma membrane (Walker *et al.*, 1998). In pathological situations such as ischemia-reperfusion injury or a cisplatin-induced acute kidney injury (AKI), meprin α and meprin β undergo redistribution from apical to basolateral tubular basement membranes (Oneda *et al.*, 2008; Bylander *et al.*, 2008; Herzog *et al.*, 2007). This leads to an elevated meprin activity at the basement membrane, and consequently to tissue damage by degradation of components of the basement membrane (Walker *et al.*, 1998; Kaushal *et al.*, 1994; Kruse *et al.*, 2004; Ambort *et al.*, 2010).

The role of the metalloprotease meprin α in inflammation is further supported by the observation that the gene encoding for meprin α , MEP1A (gene encoding for human meprin α) is genetically associated with inflammatory bowel disease. SNPs (single nucleotide polymorphisms) for MEP1A were found in ulcerative colitis patients and a down-regulation for the mRNA expression of MEP1A in the epithelium has been observed during intestinal inflammation (Banerjee *et al.*, 2009; Boyd *et al.*, 2010; Coskun *et al.*, 2012). In a DSS (dextran sodium sulfate)-induced colitis model, meprin α knock-out mice express a more severe inflammation and intestinal injury compared to wild type animals, supporting the anti-inflammatory role of this metalloprotease (Banerjee and Bond, 2008). However, it remains elusive which proteolytic event is responsible for this phenotype. Several studies proposed meprins to be modulators of the immune environment by processing of pro-inflammatory cytokines such as interleukin (IL)-1 β , IL-6, IL-18, or tumor necrosis factor (TNF) α (Banerjee and Bond, 2008; Herzog *et al.*, 2009; Minder *et al.*, 2012; Keiffer and Bond, 2014). Although both meprins process IL-1 β , only meprin β has been shown to activate IL-18. According to this, in DSS-treated meprin β knock-out mice the activation of IL-18 is reduced and these animals are less susceptible to intestinal inflammation compared to wild type mice.

Alternatively, several studies demonstrated meprin α being involved in epidermal growth factor receptor (EGFR) signaling by generation of soluble transforming growth factor α (TGF α) and EGF and consequently leading to an activation of the ERK1/2 signaling cascade (Minder *et al.*, 2012; Bergin *et al.*, 2008). This results in increased cell proliferation and cell migration in response to the meprin α induced EGFR activation, indicating meprin α to be involved in the spreading of cancer cells. Further, in human bronchial epithelial cells, soluble TGF α generated by meprin α binds and activates the EGFR/TLR4 (epidermal growth factor receptor/Toll-like receptor 4) signaling cascade (Bergin *et al.*, 2008). Interestingly, the activation of the EGFR/TLR 4 signaling leads to an increased secretion of IL-8, a neutrophil chemo-attractant implicated in chronic inflammatory diseases such as cystic fibrosis (CF). Cosgrove *et al.* (2011) demonstrated an elevated expression of meprin α in mice suffering from CF, pointing to the role of this protease in inflammatory lung disease.

The role of meprin α in inflammation is supported by the observation that meprin α knock-out mice exhibit less renal damage compared to wild type mice. Additionally, meprin β deficient mice show lower levels of the inflammatory marker

IL-6 and a decreased leukocyte infiltration after renal injury (Yura *et al.*, 2009; Bylander *et al.*, 2008).

The influence of meprin α in the pathology of renal injury could be confirmed by the administration of the meprin-specific hydroxamate inhibitor actinonin (Carmago *et al.*, 2002; Takayama *et al.*, 2008). The use of actinonin in a sepsis-induced AKI mouse model led to the inhibition of the initial peritubular capillary dysfunction, as well as the subsequent tubular injury suggesting meprins as possible therapeutic targets in renal injury and sepsis (Wang *et al.*, 2010). A number of studies addressed the role of meprins as diagnostic markers, which could possibly be a target for therapeutic treatment in renal diseases (Berthier *et al.*, 2006).

Meprin in Angiogenesis and Cancer

The pro-angiogenic activity of meprin α was firstly discovered in zebrafish, where a knock-down of meprin α mRNA expression led to severe abnormalities of the vascular system (Schutte *et al.*, 2010). Zebrafish embryos failed to develop intersegmental vessels, even at the age of 48 hpf. Only the large dorsal aorta was generated, however, the blood circulation was interrupted leading to enormous developmental deficits and an early death of the embryos. This morpholino knock-down of meprin α mimics in many details the phenotype of vascular endothelial growth factor-A (VEGF-A) morphants (Nasevicius *et al.*, 2000), pointing out a general pro-angiogenic role of meprin α . Interestingly, using the proteomics technique TAILS, VEGF-A was identified as one of the substrates of meprin α . *In vitro* it has been shown that meprin α generates the same cleavage products of VEGF-A as found in wild type zebrafish (Schutte *et al.*, 2010; Jefferson *et al.*, 2013).

Further, employing TAILS identified connective tissue growth factor (CTGF), another key player in the development of blood vessels, as a substrate of meprin α (Jefferson *et al.*, 2013). During angiogenesis, CTGF forms a complex with VEGF-A, consequently immobilizing VEGF-A in the extracellular matrix (ECM) and preventing the binding of VEGF-A to its receptor. *In vitro* experiments confirmed that enzymatic cleavage of CTGF by meprin α triggers angiogenesis. This leads to a meprin α induced release of VEGF-A from the CTGF/VEGF-A complex. Soluble VEGF-A is then able to bring to its receptor and to induce angiogenesis. Interestingly, meprin α knock-out mice contain VEGF-A clusters due to the lack of the proteolytic activity of meprin α in those animals. In hens' egg chorioallantoic membrane (CAM) assay of chicken embryos meprin α showed pro-angiogenic effects by modifications of the CTGF-VEGF-A complex. While VEGF-A had no influence on the vasculature in complex with CTGF, the proteolytic processing of CTGF by meprin α significantly induced an increase in number and growth of blood vessels. By analogy, in a rat aortic ring assay, meprin α was able to promote the outgrowth, length, and branching of capillaries (Lottaz *et al.*, 2011).

The pro-angiogenic role of meprin α was further described in cardiovascular homeostasis. The metalloprotease cleaves the 32-amino acid B-type natriuretic peptide (BNP1-32) to BNP 8–32 or BNP 7–32 *in vitro* and *in vivo*, respectively, resulting in reduced natriuretic bioactivity of this protein (Boerrigter *et al.*, 2009). Thus, meprin α modulates the bioactivity of BNP 1–32

and might be a potential therapeutic target in cardiovascular diseases.

Additionally, enzymatic activity of meprin α has been found in several tumors such as breast and colon carcinoma (Lottaz *et al.*, 1999; Matters *et al.*, 2005). Although the expression of meprin α varies between different carcinomas, with quite low levels in ovarian cancer compared to gastrointestinal carcinomas (Heinzelmann-Schwarz *et al.*, 2007; Lottaz *et al.*, 2011), it contributes to the breakdown of stromal structure as it is secreted into the tumor stroma. This might lead to proteolytic modulation of ECM and consequently to the proliferation and migration of tumor cells into the surrounding tissue. In cell culture experiments it has been shown that meprin α directly promotes cell migration and treatment of human breast carcinoma cells with the specific meprin inhibitor actinonin resulting in decreased invasiveness *in vitro* (Lottaz *et al.*, 2011; Matters *et al.*, 2005). However, the pro-angiogenic effect of meprin α and its role in collagen deposition might be even more important for cancer pathology. Using proteomics for the identification of substrates it becomes more and more evident that the role of meprin α in cancer might be more complex than the ECM degrading function known for certain matrix metalloproteases (MMPs). The pro-angiogenic function of meprin α and its crucial role in collagen deposition (see section Meprin in Skin and Fibrosis) displays the relevance of this protease and its important role in drug development.

Meprin in Skin and Fibrosis

In human skin, meprin α and meprin β are constitutively expressed in the epidermis (Becker-Pauly *et al.*, 2007). While meprin α is exclusively expressed in the *Stratum basale*, where it regulates cell proliferation and migration, meprin β is only expressed in the *Stratum granulosum* where it induces terminal differentiation and might be involved in the formation of the cornified envelope. Meprins are expressed as zymogens and require the removal of their propeptides to gain proteolytic activity. Ohler *et al.* (2010) demonstrated that serine proteases KLK 4, 5, and 8 might be responsible for this activation process in skin.

Recently it became evident that meprins are involved in the assembly of collagen fibrils and in collagen deposition and are therefore important for the integrity of connective tissue. Increased collagen deposition often leads to fibrosis. Therefore, it was not surprising that meprins were found to be highly overexpressed in human fibrotic skin and in fibrotic lung tissue (Kronenberg *et al.*, 2010; Biasin *et al.*, 2014). After secretion of the collagen precursor molecule procollagen, it undergoes proteolytic modifications, which lead to the formation of collagen fibrils. Procollagen contains N- and C-terminal propeptides that have to be removed by enzymatic cleavage to allow the collagen fibril to assemble. For a long time the procollagen N-proteinase activity has been referred to a disintegrin and metalloprotease with thrombospondin motifs (ADAMTS)-2, -3, and -14, and the only known C-proteinases were bone morphogenetic protein (BMP)-1 and tolloid-like proteases (Moali and Hulmes, 2009). In contrast to other procollagen proteinases which cleave off only C- or N-propeptides, meprin α as well as meprin β are unique in the

removal of both propeptides from type I and type III procollagen, thereby inducing *de novo* fibrillogenesis (Broder *et al.*, 2013; Kronenberg *et al.*, 2010). The physiological relevance of meprin α and meprin β in collagen assembly became evident by analyzing meprin deficient mice. Due to the lack of the procollagen proteinase activity of meprins, these mice exhibit greater amounts of the unprocessed procollagen I compared to wild type mice and a significantly reduced collagen deposition in skin. This results in a markedly decreased tensile strength of the skin, reminiscent of the fragile skin of Ehlers-Danlos syndrome VII patients. This phenotype is the result of a deletion mutation in the N-proteinase cleavage site in the Col1A1 or Col1A2 genes, or alternatively, lack of ADAMTS-2 activity, both leading to the retention of the N-propeptide of type I procollagen (Colige *et al.*, 2004). Interestingly, investigation of the cleavage site generated by meprins revealed exactly the same site, which is deleted in Ehlers-Danlos syndrome VII patients (Broder *et al.*, 2013). The distinct phenotype of meprin α and meprin β knock-out mice indicates a strong physiological relevance of meprins in collagen assembly and in the maintenance of the integrity of the skin.

On the other hand, both meprin α and meprin β were found being overexpressed in the dermis of human fibrotic skin, where they might be responsible for strong collagen deposition (Kronenberg *et al.*, 2010). Additionally, micro-array studies identified meprin β as an important factor in fibrotic lungs from patients suffering from pulmonary hypertension. It was not only highly overexpressed in diseased lungs, but also the most up-regulated gene in the corresponding mouse model (Biasin *et al.*, 2014). An acute overexpression of meprin consequently leads to an increased collagen deposition in this tissue resulting in fibrosis. Thus, inhibition of meprin activity under these pathologic conditions might be beneficial for the therapeutic treatment of fibrosis.

Meprin in Neurobiology

The role of meprin β in neurobiology was not known until the amyloid precursor protein (APP) was identified as a substrate for meprin β . Using TAILS, APP was found being shed from the cell surface by the membrane-bound meprin β , releasing the soluble APP fragment (sAPP β) and the membrane-bound C-terminal 99 amino acids of APP (C99) similarly to the β -secretase BACE1 (β -site APP-cleaving enzyme 1) (Bien *et al.*, 2012). C99 is further cleaved by γ -secretase, an event that leads to the generation of a small intracellular fragment and the neurotoxic C-terminal A β peptide (Selkoe, 2001). Upon cleavage A β peptides form plaques which are thought to have a pathologic effect on the progression of the Alzheimer's disease. In general, the majority of the released A β peptides is 40–42 amino acids in length. Interestingly, the β -secretase activity of meprin β leads to the generation of different A β species as it cleaves not only at the β -secretase cleavage site but also in close proximity. Bien and colleagues (2012) demonstrated that overexpression of meprin β results in the release of N-terminally truncated A β variants. Strikingly, exactly the same A β variant is found in human cerebrospinal fluids indicating that meprin β might be responsible for this proteolytic event *in vivo* (Wiltfang *et al.*, 2001). On the other hand, the application of TAILS identified ADAM10 as a substrate for meprin β and it

was demonstrated *in vitro* and *in vivo* that meprin β activates the α -secretase (Jefferson *et al.*, 2013). ADAM10 is known to cleave APP within the A β region, resulting in the release of the soluble ectodomain sAPP α . This proteolytic event consequently prevents the formation and aggregation of neurotoxic A β peptides (Kuhn *et al.*, 2010).

Meprin β is expressed as a membrane-tethered homodimer. However, it can be shed from the cell surface by ADAM (a disintegrin and metalloprotease) proteases (Hahn *et al.*, 2003; Jefferson *et al.*, 2013). Jefferson and colleagues have shown that soluble meprin β is able to process the N-terminus of APP *in vitro* and *in vivo* releasing N-terminal fragments of 11 and 22 kDa (APP11 and APP22) (Jefferson *et al.*, 2011). The latter was previously found in human brain lysates (Portelius *et al.*, 2010), once again suggesting the physiologic relevance of meprin β in neurobiology.

TAILS enabled the discovery of a proteolytic web of meprin β in a pathological context and to decipher meprin β 's role in neurodegeneration. Although the exact mechanism behind the generation of A β peptides remains elusive, it is unambiguous that meprin β is involved in the modulation of APP *in vivo* and might therefore be a novel therapeutic target in Alzheimer's disease.

Conclusion

Meprin metalloproteases exhibit distinctive structural and functional features. Revealing the unique cleavage specificity of these proteases with strong preference for negatively charged amino acid residues in combination with degradomic approaches enabled the identification of novel substrates important for health and disease. For instance, meprin α is a pro-angiogenic enzyme and important for certain tumor progression. Meprin β cleaves the APP at distinct sites, thereby releasing A β , a hallmark in the development of Alzheimer's disease. Both meprin α and meprin β are important for collagen fibrillogenesis, which correlates with increased meprin expression in fibrotic conditions such as keloids or pulmonary hypertension. Hence, regulation of these metalloproteases may be a suitable strategy for the treatment of associated pathological conditions.

See also: Protein Synthesis/Degradation: Protein Degradation – Protease Classes: Matrix Metalloproteinases

References

- Amort, D., Brellier, F., Becker-Pauly, C., *et al.*, 2010. Specific processing of tenascin-C by the metalloprotease meprin β neutralizes its inhibition of cell spreading. *Matrix Biology* 29, 31–42.
- Arolas, J.L., Broder, C., Jefferson, T., *et al.*, 2012. Structural basis for the sheddase function of human meprin β metalloproteinase at the plasma membrane. *Proceedings of the National Academy of Sciences of the United States of America* 109, 16131–16136.
- Banerjee, S., Bond, J.S., 2008. Prointerleukin-18 is activated by meprin β *in vitro* and *in vivo* in intestinal inflammation. *Journal of Biological Chemistry* 283, 31371–31377.
- Banerjee, S., Oneda, B., Yap, L.M., *et al.*, 2009. MEP1A allele for meprin A metalloprotease is a susceptibility gene for inflammatory bowel disease. *Mucosal Immunol* 2, 220–231.

- Becker-Pauly, C., Barre, O., Schilling, O., *et al.*, 2011. Proteomic analyses reveal an acidic prime side specificity for the astacin metalloprotease family reflected by physiological substrates. *Molecular & Cellular Proteomics* 10, M111.009233.
- Becker-Pauly, C., Bruns, B.C., Damm, O., *et al.*, 2009. News from an ancient world: Two novel astacin metalloproteases from the horseshoe crab. *Journal of Molecular Biology* 385, 236–248.
- Becker-Pauly, C., Howel, M., Walker, T., *et al.*, 2007. The alpha and beta subunits of the metalloprotease meprin are expressed in separate layers of human epidermis, revealing different functions in keratinocyte proliferation and differentiation. *Journal of Investigative Dermatology* 127, 1115–1125.
- Bergin, D.A., Greene, C.M., Sterchi, E.E., *et al.*, 2008. Activation of the epidermal growth factor receptor (EGFR) by a novel metalloprotease pathway. *Journal of Biological Chemistry* 283, 31736–31744.
- Bertenshaw, G.P., Norcum, M.T., Bond, J.S., 2003. Structure of homo- and hetero-oligomeric meprin metalloproteases. Dimers, tetramers, and high molecular mass multimers. *Journal of Biological Chemistry* 278, 2522–2532.
- Berthier, C.C., Lods, N., Joosten, S.A., *et al.*, 2006. Differential regulation of metzincins in experimental chronic renal allograft rejection: Potential markers and novel therapeutic targets. *Kidney International* 69, 358–368.
- Beynon, R., Shannon, J., Bond, J., 1981. Purification and characterization of a metallo-endoproteinase from mouse kidney. *Biochemical Journal* 199, 591–598.
- Beynon, R.J., Bond, J.S., 1983. Deficiency of a kidney metalloproteinase activity in inbred mouse strains. *Science* 219, 1351–1353.
- Biasin, V., Marsh, L.M., Egemnazarov, B., *et al.*, 2014. Meprin beta, a novel mediator of vascular remodelling underlying pulmonary hypertension. *Journal of Pathology* 233, 7–17.
- Bien, J., Jefferson, T., Causevic, M., *et al.*, 2012. The metalloprotease meprin beta generates amino terminal-truncated amyloid beta peptide species. *Journal of Biological Chemistry* 287, 33304–33313.
- Boerrigter, G., Costello-Boerrigter, L.C., Harty, G.J., *et al.*, 2009. B-type natriuretic peptide 8–32, which is produced from mature BNP 1–32 by the metalloprotease meprin A, has reduced bioactivity. *American Journal of Physiology – Regulatory, Integrative and Comparative Physiology* 296, R1744–R1750.
- Bond, J.S., Beynon, R.J., Reckelhoff, J.F., David, C.S., 1984. Mep-1 gene controlling a kidney metalloendopeptidase is linked to the major histocompatibility complex in mice. *Proceedings of the National Academy of Sciences of the United States of America* 81, 5542–5545.
- Bond, J.S., Rojas, K., Overhauser, J., Zoghbi, H.Y., Jiang, W., 1995. The structural genes, MEPIA and MEP1B, for the alpha and beta subunits of the metalloendopeptidase meprin map to human chromosomes 6p and 18q, respectively. *Genomics* 25, 300–303.
- Boyd, M., Hansen, M., Jensen, T.G., *et al.*, 2010. Genome-wide analysis of CDX2 binding in intestinal epithelial cells (Caco-2). *Journal of Biological Chemistry* 285, 25115–25125.
- Broder, C., Arnold, P., Vadon-Le Goff, S., *et al.*, 2013. Metalloproteases meprin alpha and meprin beta are C- and N-procollagen proteinases important for collagen assembly and tensile strength. *Proceedings of the National Academy of Sciences of the United States of America* 110, 14219–14224.
- Broder, C., Becker-Pauly, C., 2013. The metalloproteases meprin alpha and meprin beta: Unique enzymes in inflammation, neurodegeneration, cancer and fibrosis. *Biochemical Journal* 450, 253–264.
- Bylander, J., Li, Q., Ramesh, G., *et al.*, 2008. Targeted disruption of the meprin metalloproteinase beta gene protects against renal ischemia-reperfusion injury in mice. *American Journal of Physiology Renal Physiology* 294, F480–F490.
- Carmago, S., Shah, S.V., Walker, P.D., 2002. Meprin, a brush-border enzyme, plays an important role in hypoxic/ischemic acute renal tubular injury in rats. *Kidney International* 61, 959–966.
- Colige, A., Nuytinck, L., Hausser, I., *et al.*, 2004. Novel types of mutation responsible for the dermatosparic type of Ehlers-Danlos syndrome (Type VIIC) and common polymorphisms in the ADAMTS2 gene. *Journal of Investigative Dermatology* 123, 656–663.
- Cosgrove, S., Chotirmall, S.H., Greene, C.M., McElvaney, N.G., 2011. Pulmonary proteases in the cystic fibrosis lung induce interleukin 8 expression from bronchial epithelial cells via a heme/meprin/epidermal growth factor receptor/Toll-like receptor pathway. *Journal of Biological Chemistry* 286, 7692–7704.
- Coskun, M., Olsen, A.K., Holm, T.L., *et al.*, 2012. TNF-alpha-induced down-regulation of CDX2 suppresses MEPIA expression in colitis. *Biochimica et Biophysica Acta* 1822, 843–851.
- Craig, S.S., Reckelhoff, J.F., Bond, J.S., 1987. Distribution of meprin in kidneys from mice with high- and low-meprin activity. *American Journal of Physiology* 253, C535–C540.
- Hahn, D., Pischitz, A., Roesmann, S., *et al.*, 2003. Phorbol 12-myristate 13-acetate-induced ectodomain shedding and phosphorylation of the human meprin beta metalloprotease. *Journal of Biological Chemistry* 278, 42829–42839.
- Hedrich, J., Lottaz, D., Meyer, K., *et al.*, 2010. Fetuin-A and cystatin C are endogenous inhibitors of human meprin metalloproteases. *Biochemistry* 49 (39), 8599–8607.
- Heinzelmann-Schwarz, V.A., Scolyer, R.A., Scurry, J.P., *et al.*, 2007. Low meprin alpha expression differentiates primary ovarian mucinous carcinoma from gastrointestinal cancers that commonly metastasise to the ovaries. *Journal of Clinical Pathology* 60, 622–626.
- Herzog, C., Haun, R.S., Kaushal, V., *et al.*, 2009. Meprin A and meprin alpha generate biologically functional IL-1beta from pro-IL-1beta. *Biochemical and Biophysical Research Communications* 379, 904–908.
- Herzog, C., Seth, R., Shah, S.V., Kaushal, G.P., 2007. Role of meprin A in renal tubular epithelial cell injury. *Kidney International* 71, 1009–1018.
- Jefferson, T., Auf Dem Keller, U., Bellac, C., *et al.*, 2013. The substrate degradome of meprin metalloproteases reveals an unexpected proteolytic link between meprin beta and ADAM10. *Cellular and Molecular Life Sciences* 70, 309–333.
- Jefferson, T., Causevic, M., Auf Dem Keller, U., *et al.*, 2011. Metalloprotease meprin beta generates nontoxic N-terminal amyloid precursor protein fragments *in vivo*. *Journal of Biological Chemistry* 286, 27741–27750.
- Kaushal, G.P., Walker, P.D., Shah, S.V., 1994. An old enzyme with a new function: Purification and characterization of a distinct matrix-degrading metalloproteinase in rat kidney cortex and its identification as meprin. *Journal of Cell Biology* 126, 1319–1327.
- Keiffer, T.R., Bond, J.S., 2014. Meprin metalloproteases inactivate interleukin 6. *Journal of Biological Chemistry* 289, 7580–7588.
- Kronenberg, D., Bruns, B.C., Moali, C., *et al.*, 2010. Processing of procollagen III by meprins: new players in extracellular matrix assembly? *Journal of Investigative Dermatology* 130 (12), 2727–2735.
- Kruse, M.N., Becker, C., Lottaz, D., *et al.*, 2004. Human meprin alpha and beta homo-oligomers: Cleavage of basement membrane proteins and sensitivity to metalloprotease inhibitors. *Biochemical Journal* 378, 383–389.
- Kuhn, P.H., Wang, H., Dislich, B., *et al.*, 2010. ADAM10 is the physiologically relevant, constitutive alpha-secretase of the amyloid precursor protein in primary neurons. *EMBO Journal* 29, 3020–3032.
- Lottaz, D., Maurer, C.A., Hahn, D., Buchler, M.W., Sterchi, E.E., 1999. Nonpolarized secretion of human meprin alpha in colorectal cancer generates an increased proteolytic potential in the stroma. *Cancer Research* 59, 1127–1133.
- Lottaz, D., Maurer, C.A., Noel, A., *et al.*, 2011. Enhanced activity of meprin-alpha, a pro-migratory and pro-angiogenic protease, in colorectal cancer. *PLoS One* 6, e26450.
- Madoux, F., Tredup, C., Spicer, T.P., *et al.*, 2014. Development of high throughput screening assays and pilot screen for inhibitors of metalloproteases meprin alpha and beta. *Biopolymers* 102, 396–406.
- Matters, G.L., Manni, A., Bond, J.S., 2005. Inhibitors of polyamine biosynthesis decrease the expression of the metalloproteases meprin alpha and MMP-7 in hormone-independent human breast cancer cells. *Clinical and Experimental Metastasis* 22, 331–339.
- Minder, P., Bayha, E., Becker-Pauly, C., Sterchi, E.E., 2012. Meprin alpha transactivates the epidermal growth factor receptor (EGFR) via ligand shedding, thereby enhancing colorectal cancer cell proliferation and migration. *Journal of Biological Chemistry* 287, 35201–35211.
- Moali, C., Hulmes, D.J., 2009. Extracellular and cell surface proteases in wound healing: New players are still emerging. *European Journal of Dermatology* 19 (6), 552–564.
- Nasevicius, A., Larson, J., Ekker, S.C., 2000. Distinct requirements for zebrafish angiogenesis revealed by a VEGF-A morphant. *Yeast (Chichester, England)* 17, 294–301.
- Ohler, A., Debela, M., Wagner, S., Magdolen, V., Becker-Pauly, C., 2010. Analyzing the protease web in skin: Meprin metalloproteases are activated specifically by KLK4, 5 and 8 vice versa leading to processing of proKLK7 thereby triggering its activation. *Biological Chemistry* 392 (7), 653–664.
- Oneda, B., Lods, N., Lottaz, D., *et al.*, 2008. Metalloprotease meprin beta in rat kidney: Glomerular localization and differential expression in glomerulonephritis. *PLoS One* 3, e2278.
- Portelius, E., Brinkmalm, G., Tran, A., *et al.*, 2010. Identification of novel N-terminal fragments of amyloid precursor protein in cerebrospinal fluid. *Experimental Neurology* 223, 351–358.
- Schutte, A., Hedrich, J., Stocker, W., Becker-Pauly, C., 2010. Let it flow: Morpholino knockdown in zebrafish embryos reveals a pro-angiogenic effect of the metalloprotease meprin alpha2. *PLoS One* 5, e8835.
- Selkoe, D.J., 2001. Alzheimer's disease results from the cerebral accumulation and cytotoxicity of amyloid beta-protein. *Journal of Alzheimer's Disease* 3, 75–80.

- Sterchi, E.E., Green, J.R., Lentze, M.J., 1983. Nonpancreatic hydrolysis of *N*-benzoyl-L-tyrosyl-*p*-aminobenzoic acid (PABA peptide) in the rat small intestine. *Journal of Pediatric Gastroenterology and Nutrition* 2, 539–547.
- Stocker, W., Grams, F., Baumann, U., *et al.*, 1995. The metzincins—topological and sequential relations between the astacins, adamalysins, serralsins, and matrixins (collagenases) define a superfamily of zinc-peptidases. *Protein Science* 4, 823–840.
- Takayama, J., Takaoka, M., Yamamoto, S., *et al.*, 2008. Actinonin, a meprin inhibitor, protects ischemic acute kidney injury in male but not in female rats. *European Journal of Pharmacology* 581, 157–163.
- Trachtman, H., Valderrama, E., Dietrich, J.M., Bond, J.S., 1995. The role of meprin A in the pathogenesis of acute renal failure. *Biochemical and Biophysical Research Communications* 208, 498–505.
- Walker, P.D., Kaushal, G.P., Shah, S.V., 1998. Meprin A, the major matrix degrading enzyme in renal tubules, produces a novel nidogen fragment *in vitro* and *in vivo*. *Kidney International* 53, 1673–1680.
- Wang, Z., Herzog, C., Kaushal, G.P., Gokden, N., Mayeux, P.R., 2010. Actinonin, A Meprin A Inhibitor, protects the renal microcirculation during sepsis. *Shock* 35 (2), 141–147.
- Wiltfang, J., Esselmann, H., Cupers, P., *et al.*, 2001. Elevation of beta-amyloid peptide 2–42 in sporadic and familial Alzheimer's disease and its generation in PS1 knockout cells. *Journal of Biological Chemistry* 276, 42645–42657.
- Yan, L., Fei, K., Zhang, J., Dexter, S., Sarra, M.P., 2000. Identification and characterization of hydra metalloproteinase 2 (HMP2): a meprin-like astacin metalloproteinase that functions in foot morphogenesis. *Development (Cambridge, England)* 127, 129–141.
- Yura, R.E., Bradley, S.G., Ramesh, G., Reeves, W.B., Bond, J.S., 2009. Meprin A metalloproteases enhance renal damage and bladder inflammation after LPS challenge. *American Journal of Physiology Renal Physiology* 296, F135–F144.

Kallikrein

RA Bradshaw, University of California, Irvine, CA, USA

RL Lundblad, Chapel Hill, NC, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Angiogenesis The process of the development of new blood vessels. Angiogenesis is considered critical for tumor growth.

Extracellular matrix Refers to a substance consisting of proteins, glycoproteins, and polysaccharide that provide support as a structural scaffold for cells and can be a barrier between tissues.

Insulin-like growth factor-1 Originally known as somatomedin C and is a member of a group of proteins with homology to insulin. Insulin-like growth factor-1 is important for normal growth and is also involved in cancer cell development. Insulin-like growth factor-1 binds to an insulin-like growth factor binding protein where it is inactive; degradation of insulin-like growth factor binding protein results in the release of active insulin-like growth factor-1.

Kinin A biologically active peptide derived from precursor proteins, kininogens. Kinins are best known for vasodilation

and for the contraction of smooth muscle. Recent work has suggested that kinins have more diverse activities.

Missense polymorphism A mutation in DNA which results in a change in the amino acid sequence of proteins. A missense polymorphism results from the change in a single base in a codon; the change from CGC which codes for arginine to CAC (substitution of adenine to guanine) results in histidine in place of arginine in the amino acid sequence of the expressed protein.

Nitric oxide Originally described as endothelium derived relaxing factor (EDRF) which caused smooth muscle relaxation. Nitric oxide can react with peroxide to form peroxynitrite which has a variety of biological functions.

Pleiotropic Describing the ability of an object to have multiple functions.

Urokinase A plasminogen activator originally isolated from urine. Now known as urokinase-type plasminogen activator or urokinase plasminogen activator.

Introduction

Proteases can be separated into two groups: digestive proteases and regulatory proteases (Neurath, 1985). Trypsin and chymotrypsin are examples of digestive proteases while the coagulation enzymes and complement proteases are considered to be of the regulatory type. Thirty years ago, plasma kallikrein (KLK) and tissue/glandular KLK were two regulatory proteases defined by their ability to liberate kinins from kininogens. There was emerging evidence to suggest a role for KLK activity in processing growth factor precursors (Blabber *et al.*, 1989) and prostate-specific antigen (PSA) was a KLK-related protein identified early on as a biomarker for prostate cancer (Wang *et al.*, 1981). The past 20 years has seen a dramatic increase in the number of KLK-related peptidases identified (Table 1).

Plasma KLK

KLK is a functional activity defined by the cleavage kininogens to form kinins, which are vasoactive peptides (Schachter, 1979). Plasma KLK and tissue (glandular) KLK differ in distribution, size, and specificity. Plasma KLK has a molecular weight of 80–90 kDa depending on the extent of glycosylation (Björkqvist *et al.*, 2013) while tissue KLK has a molecular weight of approximately 40 kDa (Pathak *et al.*, 2013). Both are synthesized in precursor or zymogen form; prekallikrein (preKLK) in the case of plasma KLK and prokallikrein (proKLK) in the case of tissue KLK. The plasma moiety liberates kinin activity from high-molecular weight kininogen (HMWK) while tissue KLK

releases kinin activity from either HMWK or low-molecular weight kininogen (LMWK). Plasma preKLK is thought to participate in the contact phase of blood coagulation and it has been shown to circulate as a complex with HMWK (Saito and Ratnoff, 1979). HMWK, which has been cleaved by plasma KLK to a two-chain protein, HMWk_a, inhibits angiogenesis (Tesfay *et al.*, 2012). The activation of preKLK in plasma involves factor XIIIa, HMWK, endothelial cell surface, including proteoglycans, and prolylcarboxypeptidase. A sequential series of proteolytic events, occurring as a cascade, has been proposed for the KLK/bradykinin system (Figure 1; Kaplan and Joseph, 2010).

Tissue KLK

Tissue KLK (also KLK1) is a secretory product of the salivary glands and the pancreas and is found in the kidney. As with plasma preKLK, tissue KLK is synthesized as a zymogen, which is activated by limited proteolysis. Classic tissue kallikrein (KLK1) may be partially activated during the secretory process (Jenzano *et al.*, 1988; Peters *et al.*, 1992). Several proteases have been shown to activate proKLK, including trypsin, thermolysin, and clostripain but a physiological activator has not been identified (Mauer and Kemme, 2002). The failure to detect significant active KLK1 in the various biological fluids is likely a consequence of reaction with various inhibitors including α_2 -2-macroglobulin and α_1 -antitrypsin (Shimamoto *et al.*, 1984; Rahman *et al.*, 1994). Kallistatin was identified as a specific serpin inhibitor of tissue KLK (Chao *et al.*, 1996). Kallistatin has also been shown to have a negative effect on

Table 1 KLK-related peptidases

<i>KLK-related peptidase</i>	<i>Other names and tissue source</i>	<i>References</i>
KLK1	Classical tissue KLK or glandular KLK; primary sites of synthesis in pancreas, kidney, and salivary glands; wide tissue distribution	Chao (2013)
KLK2	Human glandular kallikrein 1; exclusive in the prostate	Veneris-Lowe <i>et al.</i> (2007)
KLK3	Prostate-specific antigen; exclusive in prostate	Mattsson <i>et al.</i> (2008)
KLK4	Protease; wide distribution including prostate and enamel; expressed in ovarian cancer	Stephenson <i>et al.</i> (1999)
KLK5	Human stratum corneum tryptic enzyme (SCTE); found in skin and breast tissue; ovarian cancer	Yousef <i>et al.</i> (2003)
KLK6	Protease M; found in brain, pancreas; proposed a biomarker for ovarian cancer, brain trauma	Bayés <i>et al.</i> (2004)
KLK7	Human stratum corneum chymotryptic enzyme (SCCE); found in skin, esophagus, heart, pancreas	Ramani <i>et al.</i> (2011)
KLK8	Neuropsin; found in brain, skin, and breast	Kishi <i>et al.</i> (2003)
KLK9	Wide tissue distribution; ovarian tumors, breast tumors	Memari <i>et al.</i> (2006)
KLK10	Normal epithelial cell-specific 1; NES1 serine protease; wide tissue distribution	Liu <i>et al.</i> (1996)
KLK11	Hippostasin, trypsin-like serine protease; TLSP; wide tissue distribution with various isoforms	Yoshida <i>et al.</i> (1998)
KLK12	KLK-like 5; first identified by sequence analysis of the chromosomal region encoding other members of the kallikrein-related peptidase family; it is suggested to have isoforms showing a wide tissue distribution	Yousef <i>et al.</i> (2000b)
KLK13	KLK-L4	Petraki <i>et al.</i> (2003)
KLK14	Wide tissue distribution	Yousef <i>et al.</i> (2000a)
KLK15	Prostin; prostinogen; ACO protease, hK15; found in prostate, colon, thyroid	Batra and Clements (2013)
Plasma KLK	Formed from plasma prekKLK (Fletcher factor), synthesized in liver and found in liver and interstitial fluid)	Schmaier and McCrae (2007)

Note: This table is a listing of human kallikrein-related peptidases. The rat and mouse KLK families are somewhat more complex (see Lundwall *et al.*, 2006)

wound healing through inhibition of Wnt/ β -catenin signaling pathway (McBride *et al.*, 2014).

Tissue KLK is suggested to have both kinin-dependent (Zaika *et al.*, 2011) and kinin-independent (Chambrey and Picard, 2011) functions in the kidney. Tissue KLK is not constitutively expressed in the kidney (Cumming *et al.*, 1994). Use of AAV (adeno-associated virus)-mediated gene therapy with human tissue KLK has been shown to protect renal function in acute renal failure in rats (Tu *et al.*, 2008). Urine has been a source for the purification of tissue KLK (Girolami *et al.*, 1989) and tissue proKLK (Takada *et al.*, 1985).

Tissue KLK is also expressed in pancreatic secretions and in saliva and its mRNA has also been shown to be present in heart tissue in rats (Campbell *et al.*, 2011). There are several variants of tissue KLK (Chan *et al.*, 1998; Fiedler *et al.*, 1999). A missense polymorphism, which results in the insertion of a histidine for arginine at position 53 in human tissue (urinary) KLK (R53H) is associated with the loss of kinin-generating activity (Slim *et al.*, 2002). The crystal structure of urinary KLK1 suggested that Arg53 is critical for the binding of substrate. Subsequent work (Azizi *et al.*, 2005) showed this mutation was associated with arterial dysfunction. More recent work with a murine model system indicated that tissue KLK is critical for arterial repair. Kinin-B2 receptor interaction by tissue KLK and the activation of MMP-9 are important for increased angiogenesis (Stone *et al.*, 2009).

Kinins

Much of the physiological action of plasma KLK and tissue KLK is mediated through the production of biologically active

peptides known as kinins (Roche e Silva *et al.*, 1949; Schachter and Thain, 1954). Bradykinin is produced by plasma KLK acting on HMWK while lysylbradykinin (kallidin) is produced by tissue KLK from either HMWK or LMWK. Lysylbradykinin can be converted to bradykinin through the action of an aminopeptidase. Kinins were originally described as vasoactive peptides, which lower blood pressure and cause uterine and intestinal contractions in an action mediated by nitric oxide (Moncada *et al.*, 1988). It is now apparent that kinins are pleiotropic (Maurer *et al.*, 2011) and they are of importance in cancer (da Costa *et al.*, 2014).

Kallikrein-Related Peptidases

The KLK family is comprised of 15 proteins (Harvey *et al.*, 2000; Lundwall *et al.*, 2006; Table 1). The majority of these have activity unrelated to the classic definition of a KLK as a kininogenase (Kraut *et al.*, 1930). The KLK-related peptidases function as secretory factors in the extracellular space where they have diverse functions including the processing of growth factors, cleavage of extracellular matrix proteins, and activation of other proteases. They are also involved in tumorigenesis (Dong *et al.*, 2014).

KLK2 and KLK3 (PSA) are exclusive products of the prostate and have been used for prostate cancer screening (Lövgren *et al.*, 1995; Voigt *et al.*, 2014). KLK-related peptidase 2 (KLK2, hK2) has been shown to have kininogenase activity (Deperthes *et al.*, 1997; Charlesworth *et al.*, 1999) while PSA has little (Andrade *et al.*, 2010) or no kininogenase activity (Deperthes

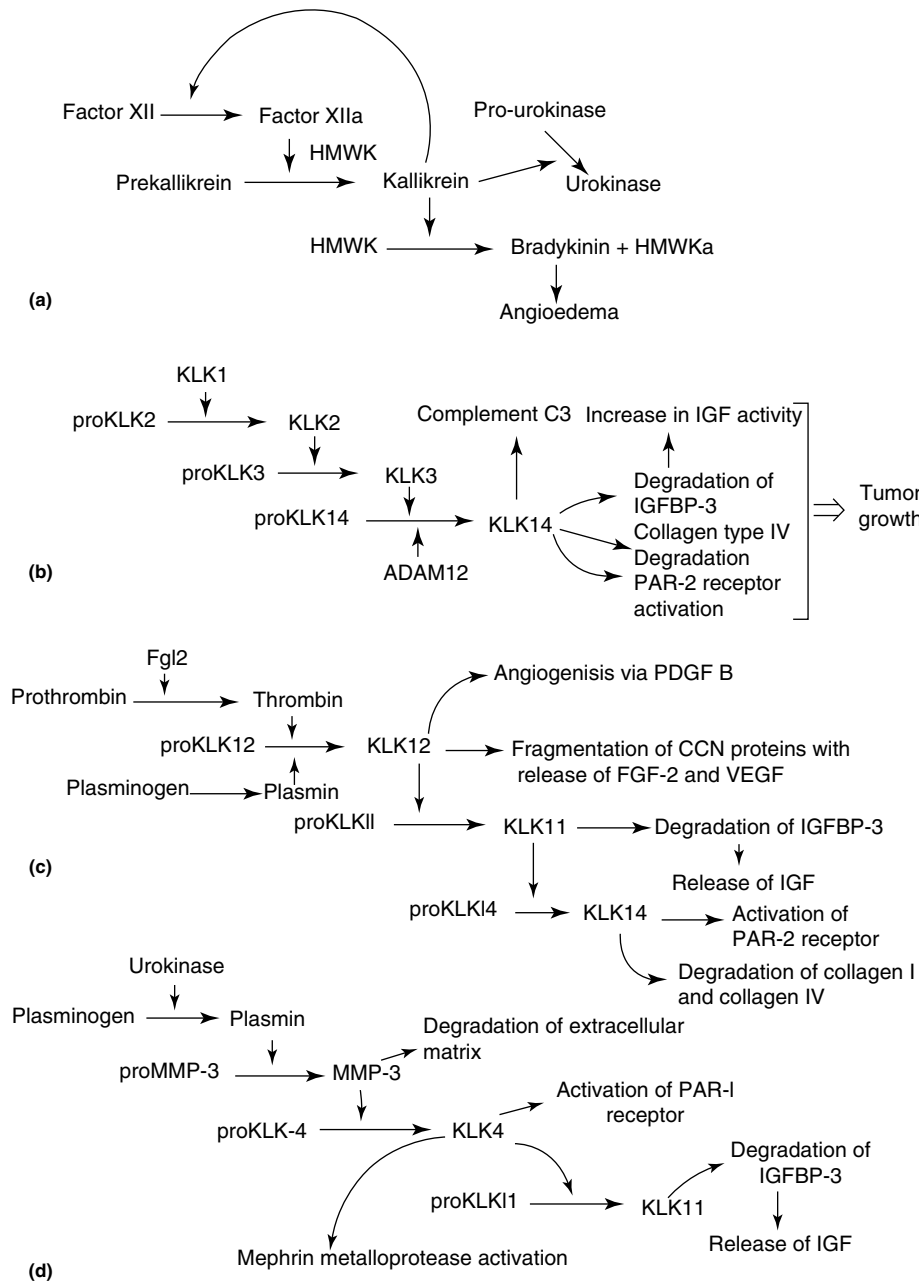


Figure 1 Cascade systems for the expression of kallikrein-like peptidase activity. (a) A putative scheme for the interaction of contact activation (Factor XII) and plasma prekallikrein to form kallikrein which can then activate pro-urokinase to form urokinase and form bradykinin from HMWK (Kaplan and Ghebrehiwet, 2010). (b) A putative scheme by which KLK1 could initiate a cascade of KLK peptidases resulting in formation of KLK14 which would degrade collagen type IV thus permitting tumor metastasis, would degrade IGF-BP with the liberation of IGF which would foster tumor cell growth, and cleave PAR-2 receptor also stimulating tumor cell growth (Gratio *et al.*, 2011; Albrechtsen *et al.*, 2013). KLK14 may also form biologically active peptides from complement C3 (Oikonomopoulou *et al.*, 2013). (c) An example of the interaction of the thrombostasis proteases with the KLK peptidases where thrombin activates proKLK12 resulting in the sequential activation of proKLK11 and proKLK14 (Blaber *et al.*, 2010; Guillon-Munos *et al.*, 2011; Kryza *et al.*, 2014). (d) A KLK cascade initiated by a matrix metalloproteinase which had been activated by urokinase. Urokinase activates pro-MMP12 which then activates proKLK4 which then activates proKLK11 (Debela *et al.*, 2008; Ohler *et al.*, 2010). This cascade has the potential to degrade extracellular matrix (MMP-3), activate PAR-1 receptors on various target cells including tumor cells (Wang *et al.*, 2010).

et al., 1997). Both KLK2 and PSA are considered primary secretory products of the prostate and, as with other KLK-related peptidases, they are products derived by alternative splicing (David *et al.*, 2002).

KLK-related peptidases which include tissue kallikrein (Table 1) have been shown to be present in a large number of tissues and fluids by mRNA analysis and highly sensitive immunoassays (Shaw and Diamandis, 2007). The presence of

Table 2 Cell receptor targets for KLK-related peptidases

<i>KLK-related peptidase</i>	<i>Cell receptor targets</i>	<i>References</i>
KLK1	PAR-1 on several cell types; suggested transactivation of EGF receptor	Gao <i>et al.</i> (2010), Oikonomopoulou <i>et al.</i> (2006), Lu <i>et al.</i> (2014)
KLK4	PAR-1 on colon cancer cells	Gratio <i>et al.</i> (2010)
KLK5	PAR-2 on keratinocytes	Stefansson <i>et al.</i> (2008)
KLK6	PAR-1 and PAR-2 on neurons	Yoo <i>et al.</i> (2013)
KLK8	Cannot activate PAR-2 but does inactivate PAR-1 in human embryonic kidney (HEK) cells; can activate Kristen transformed rat kidney (KTRK) cells expressing PAR-2	Ramachandran <i>et al.</i> (2012)
KLK14	KLK14 can cleave PAR-2 on HEK cells and in KTRK cells. KLK14 was able to activate PAR-2 in colon cancer cells	Gratio <i>et al.</i> (2011)
Plasma KLK	Cleavage of PAR-1/PAR-2 in vascular smooth muscle cells results in transactivation of the EGF receptor	Abdullah <i>et al.</i> (2010)

KLK-related peptidases 5, 6, 7, 10, 11, 12, 13, and 14 in the cytoplasm of glandular epithelial cells and in cellular secretions has been established with immunohistochemistry (Petraki *et al.*, 2006).

There are several substrates of KLKs that are notable. Proteolysis of insulin-like growth factor binding protein (IGF-BP) by several members of the KLK family (Figure 1) results in increased unbound insulin-related growth factor-1 (IGF-1) in the interstitial space; increased IGF-1 is known to enhance tumor growth (Hekim *et al.*, 2010). Various other proteins have been suggested to serve as substrates for the several KLK-like peptidases (Borgono *et al.*, 2007; Debela *et al.*, 2008; Sotiropoulou and Pampalakis, 2010).

The relationship between various KLKs and other proteases does suggest that functional cascades exist (Figure 1), such as observed with complement and coagulation proteases and that these are important regulatory processes (Blaber *et al.*, 2010). These processes not only involve zymogen activation but also interactions with various cells mediated by protease-activated receptors (PARs). These putative cascades involve the stepwise interaction of proteases where the zymogen form of one enzyme is converted to the activated form by limited proteolysis. Once activated, this protease can activate the next zymogen. Unlike the coagulation cascades where there is an amplification of the initiation signal, it is suggested that the kallikrein cascades are regulatory in nature and likely provide temporal control of events such as tissue remodeling.

KLK–Cell Interactions

There has been limited work on the interaction of KLK with cells (Table 2). Many of the KLK family members have a specificity for peptide bond cleavage, which would result in activation of PARs. PARs have tethered ligands where limited proteolysis in the amino terminal domain of the extracellular peptide chain results in an interaction with the G-protein coupled receptor region (Vu *et al.*, 1991; Gieseler *et al.*, 2013). There are examples of KLK-mediated transactivation of epidermal growth factor (EGF) receptors via PAR-2 cleavage (Michel *et al.*, 2014). In some cases, KLK-related peptidases can

inactivate or disarm a PAR when proteolysis occurs downstream of the activation-related scissile peptide bond (Dulon *et al.*, 2003). This would appear to be the case for KLK8, which has been shown to disarm PAR-1 on human embryonic kidney cells (HEK cells) (Ramachandran *et al.*, 2012). KLK8 was able to activate PAR-2 receptors in a transformed rat kidney cell line. KLK14 has been shown to activate PAR-2 receptors in a human colon cancer cells line and it is suggested that there is ectopic expression of KLK14 in colon cancer cells (Gratio *et al.*, 2011) inducing Erk1/Erk2 expression. The ectopic expression of KLK14 in colon cancer cells and PAR-2 activation would provide an autocrine pathway for cancer cell proliferation.

KLK and Cancer

An association of KLK with cancer has been suggested for some time dating at least to an observation on kinin release in carcinoid syndrome (Oates *et al.*, 1964). The use of KLK2 and KLK3 as biomarkers for prostate cancer is well known; there is evaluation of other KLK-like peptidases as biomarkers in other malignancies (Kashuba *et al.*, 2013; Dorn *et al.*, 2013; Vakrakou *et al.*, 2014). More recently, aprotinin has been shown to inhibit tumor growth, possibly on the basis of inhibition of KLKs (Vaporciyan *et al.*, 2004). KLK-like peptidases have a broad range of biological activities which would be useful in tumor growth and metastasis. The fragmentation of IGF-BP-3 by several members of the KLK family results in the increase of IGF activity fostering tumor growth. Tumor metastasis may be influenced directly by the proteolysis of collagen type IV by KLK-related peptidases or by the activation of other protease systems such as the matrix metalloproteinases (Pezzato *et al.*, 2004). There is considerable interest in the relationship of IGF-1 and KLK3 in prostate cancer (Holmberg and Van Hemelrijck, 2014). While there is no data to support a direct effect, KLK3 will activate KLK14, which can degrade IGF-BP.

See also: Protein Synthesis/Degradation: Proteolytic Pathways: Overview of Blood Coagulation and the Pathophysiology of Blood Coagulation Disorders

References

- Abdullah, R.T., Keum, J.-S., Le, M.-H., *et al.*, 2010. Plasma kallikrein promotes epidermal growth factor receptor transactivation and signaling in vascular smooth muscle through direct activation of protease-activated receptors. *Journal of Biological Chemistry* 285, 35206–35215.
- Albrechtsen, R., Kveiborg, M., Staut, D., *et al.*, 2013. ADAM12 redistributes and activate MMP-14, resulting in gelatin degradation, reduced apoptosis and increased tumor growth. *Journal of Cell Science* 126, 4707–4720.
- Andrade, D., Assis, D.M., Lima, A.R., *et al.*, 2010. Substrate specificity and inhibition of human kallikrein-related peptidase 3 (KLK3 or PSA) activated with sodium citrate and glycosaminoglycans. *Archives of Biochemistry and Biophysics* 498, 74–82.
- Azizi, M., Boutouyrie, P., Bissery, A., *et al.*, 2005. Arterial and renal consequences of partial genetic deficiency in tissue kallikrein activity in humans. *Journal of Clinical Investigation* 115, 780–787.
- Batra, J., Clements, J., 2013. Kallikrein-related peptidase 15 (prostinogen). In: Rawlings, N.D., Salvensen, G. (Eds.), *Handbook of Proteolytic Enzymes*, third ed., vol. 3. Amsterdam: Elsevier, pp. 2814–2817.
- Bayés, A., Tsetsenis, T., Ventura, S., *et al.*, 2004. Human kallikrein 6 activity is regulated via an autoprolytic mechanism of activation/inactivation. *Biological Chemistry* 385, 517–524.
- Björkqvist, J., Jämsä, A., Renné, T., 2013. Plasma kallikrein: The bradykinin-producing enzyme. *Thrombosis and Haemostasis* 110, 399–407.
- Blabber, M., Isackson, P.J., Marsters, Jr. J.C., Burnier, J.P., Bradshaw, R.A., 1989. Substrate specificities of growth factor associated kallikreins of the mouse submandibular gland. *Biochemistry* 28, 7813–7819.
- Blaber, M., Yoon, H., Juliano, M.A., Scarisbrick, I.A., Blaber, S.I., 2010. Functional intersection of the kallikrein-related peptidases (KLKs) and thrombostasis axis. *Biological Chemistry* 391, 311–320.
- Borgono, C.A., Gavigan, J.-A., Alves, J., *et al.*, 2007. Defining the extended substrate specificity of kallikrein 1-related peptidases. *Biological Chemistry* 388, 1215–1225.
- Campbell, D.J., Zhang, Y., Kelly, D.J., *et al.*, 2011. Alliskiren increases bradykinin and tissue kallikrein mRNA levels in the heart. *Clinical and Experimental Pharmacology & Physiology* 38, 623–631.
- Chambrey, R., Picard, N., 2011. Role of tissue kallikrein in regulation of tubule function. *Current Opinion in Nephrology* 20, 523–528.
- Chan, H., Springman, E.B., Clark, J.M., 1998. Expression and characterization of human tissue kallikrein variants. *Protein Expression and Purification* 12, 361–370.
- Chao, J., 2013. Human kallikrein 1, tissue kallikrein. In: Rawlings, N.D., Salvensen, G. (Eds.), *Handbook of Proteolytic Enzymes*, vol. 3. Amsterdam: Elsevier, pp. 2751–2761.
- Chao, J., Chai, K.X., Chao, L., 1996. Tissue kallikrein inhibitors in mammals. *Immunopharmacology* 32, 67–92.
- Charlesworth, M.C., Young, C.Y.F., Miller, V.M., Tindall, D.J., 1999. Kininogenase activity of prostate-derived human glandular kallikrein (hk2) purified from seminal fluid. *Journal of Andrology* 20, 220–229.
- da Costa, P.L.N., Sirois, P., Tannock, I.F., Chammas, R., 2014. The role of kinin receptors in cancer and therapeutic opportunities. *Cancer Letters* 345, 27–38.
- Cumming, A.D., Walsh, T., Wojtacha, D., *et al.*, 1994. Expression of tissue kallikrein in human kidney. *Clinical Science (London)* 87, 5–11.
- David, A., Mabeesh, N., Azar, I., *et al.*, 2002. Unusual alternative splicing within the human kallikrein genes KLK2 and KLK3 gives rise to novel prostate-specific proteins. *Journal of Biological Chemistry* 277, 18084–18090.
- Debela, M., Beaufort, N., Magdolen, V., *et al.*, 2008. Structures and specificity of the human kallikrein-related peptidases KLK 4, 5, 6, and 7. *Biological Chemistry* 389, 623–632.
- Deperthes, D., Marceau, F., Frenette, G., *et al.*, 1997. Human kallikrein hK2 has low kininogenase activity while prostate-specific antigen has none. *Biochimica et Biophysica Acta* 1343, 102–106.
- Dong, Y., Loessner, D., Irving-Rodgers, H., *et al.*, 2014. Metastasis of ovarian cancer is mediated by kallikrein related peptidases. *Clinical and Experimental Metastasis* 31, 135–147.
- Dorn, J., Bayani, J., Yousef, G.M., *et al.*, 2013. Clinical utility of kallikrein-related peptidases (KLK) in urogenital malignancies. *Thrombosis and Haemostasis* 110, 408–422.
- Dulon, S., Candé, C., Bunnett, N.W., *et al.*, 2003. Proteinase-activated receptor-2 and human lung epithelial cells. Disarming by neutrophil serine proteinases. *American Journal of Respiratory Cell and Molecular Biology* 28, 339–346.
- Fiedler, F., Betz, G., Hinz, H., *et al.*, 1999. Not more than three tissue kallikreins identified from organs of the guinea pig. *Biological Chemistry* 380, 63–73.
- Gao, L., Smith, R.S., Chen, L.M., *et al.*, 2010. Tissue kallikrein promotes prostate cancer cell migration and invasion via a protease-activated receptor-1-dependent signaling pathway. *Biological Chemistry* 391, 803–812.
- Gieseler, F., Ungefroren, H., Settmacher, U., Hollenberg, M.D., Kaufmann, R., 2013. Proteinase-activated receptors (PARs) – Focus on receptor-receptor-interactions and their physiological and pathophysiological impact. *Cell Communication and Signalling* 11, 86. doi: 10.1186/1478-811X-11-86.
- Girolami, J.P., Pecher, C., Bascands, J.L., *et al.*, 1989. Purification of human active urinary kallikrein: Comparative inhibition studies of kininogenase and amidolytic activities. *Preparative Biochemistry* 19, 175–190.
- Gratio, V., Beaufort, N., Selz, L., *et al.*, 2010. Kallikrein-related peptidase 4: A new activator of the aberrantly expressed protease-activated receptor 1 in colon cancer cells. *American Journal of Pathology* 176, 1452–1461.
- Gratio, V., Lolriot, C., Virca, G.D., *et al.*, 2011. Kallikrein-related peptidase 14 acts on proteinase-activated receptor 2 to induce signaling pathway in colon cancer cells. *American Journal of Pathology* 179, 2625–2636.
- Guillon-Munos, A., Oikonomopoulou, K., Michel, N., *et al.*, 2011. Kallikrein-related peptidase 12 hydrolyzes matricellular proteins of the CCN family and modifies interactions of CNN1 and CCN5 with growth factor. *Journal of Biological Chemistry* 286, 25505–25518.
- Harvey, T.J., Hooper, J.D., Myers, S.A., *et al.*, 2000. Tissue-specific expression patterns and fine mapping of the human kallikrein (KLK) locus on proximal 19q13.4. *Journal of Biological Chemistry* 275, 37397–37406.
- Hekim, C., Riipi, T., Weisell, J., *et al.*, 2010. Identification of IGFBP-3 fragments generated by KLK2 and prevention of fragmentation by KLK2-inhibiting peptides. *Biological Chemistry* 391, 475–479.
- Holmberg, L., Van Hemelrijck, M., 2014. The biology and natural history of prostate cancer: A short introduction. *Recent Results in Cancer Research* 202, 1–7.
- Jenzano, J.W., Coffey, J.C., Heizer, W.D., Lundblad, R.L., Scicli, A.G., 1988. The assay of glandular kallikrein and prekallikrein in human mixed saliva. *Archives of Oral Biology* 33, 641–644.
- Kaplan, A., Ghebrehewet, B., 2010. The plasma bradykinin-forming pathways and its interrelationships with complement. *Molecular Immunology* 47, 2161–2165.
- Kaplan, A.P., Joseph, K., 2010. The bradykinin-forming cascade and its role in hereditary angioedema. *Annals of Allergy, Asthma and Immunology* 104, 193–204.
- Kashuba, E., Bailey, J., Allsup, D., Cawkwell, L., 2013. The kinin-kallikrein system: Physiological roles, pathophysiology and its relationship to cancer biomarkers. *Biomarkers* 18, 279–296.
- Kishi, T., Grass, L., Soossapillai, A., *et al.*, 2003. Human kallikrein 8: Immunoassay development and identification in tissue extracts and biological fluids. *Clinical Chemistry* 49, 87–96.
- Kraut, H., Frey, E.K., Werle, E., 1930. Der Nachweis eines Kreislaufhormons in der Pankreasdrüse. *Hoppe-Seyler's Zeitschrift für Physiologische Chemie* 189, 97–106.
- Kryza, T., Achard, C., Parent, C., *et al.*, 2014. Angiogenesis stimulated by human kallikrein-related peptidase 12 acting via a platelet-derived growth factor B-dependent paracrine pathway. *FASEB Journal* 28, 740–751.
- Liu, X.-L., Wazer, D.E., Watanabe, K., Band, V., 1996. Identification of a novel serine protease-like gene, the expression of which is down-regulated during breast cancer progression. *Cancer Research* 56, 3371–3379.
- Lövgren, J., Piironen, T., Overmo, C., *et al.*, 1995. Production of recombinant PSA and HK2 and analysis of their immunologic cross-reactivity. *Biochemical and Biophysical Research Communications* 213, 888–895.
- Lu, Z., Yang, Q., Cui, M., *et al.*, 2014. Tissue kallikrein induces SH-SY5Y cell proliferation via epidermal growth factor receptor and extracellular signal-regulated kinase 1/2 pathway. *Biochemical and Biophysical Research Communication* 446, 25–29.
- Lundwall, A., Band, V., Blaber, M., *et al.*, 2006. A comprehensive nomenclature for serine proteases with homology to tissue kallikreins. *Biological Chemistry* 387, 637–641.
- Mattsson, J.M., Valmu, L., Laakkonen, P., Stenman, U.H., Koistinen, H., 2008. Structural characterization and anti-angiogenic properties of prostate-specific antigen isoforms in seminal fluids. *Prostate* 68, 945–954.
- Mauer, S., Kemme, M., 2002. Magnetic beads capture assay for the study of zymogen activation. *Analytical Biochemistry* 305, 284–287.
- Maurer, M., Bader, M., Bas, M., *et al.*, 2011. New topics in bradykinin research. *Allergy* 66, 1397–1406.
- McBride, J.D., Jenkins, A.J., Liu, X., *et al.*, 2014. Elevated circulation levels of an antiangiogenic SERPIN in patients with diabetic microvascular complications

- impaired wound healing through suppression of Wnt signaling. *Journal of Investigative Dermatology* 134, 1725–1734.
- Memari, N., Grass, L., Nakamura, T., *et al.*, 2006. Human tissue kallikrein 9: Production of recombinant proteins and specific antibodies. *Biological Chemistry* 387, 733–740.
- Michel, N., Heuzé-Vourch, N., Lavergne, E., *et al.*, 2014. Growth and survival of lung cancer cells: Regulation by kallikrein-related peptidase 6 via activation of proteinase-activated receptor 2 and the epidermal growth factor receptor. *Biological Chemistry* 395, 1015–1025.
- Moncada, S., Palmer, R.M., Higgs, E.A., 1988. The discovery of nitric oxide as the endogenous nitrovasodilator. *Hypertension* 12, 365–372.
- Neurath, H., 1985. Proteolytic enzymes, past and present. *Federation Proceedings* 44, 2907–2913.
- Oates, J.A., Melmon, K., Sjoerdsma, A., Gillespie, L., Mason, D.T., 1964. Release of a kinin peptide in the carcinoid syndrome. *Lancet* 283 (7332), 514–517.
- Ohler, A., Debela, M., Wagner, S., Magdolen, V., Becker-Pauly, C., 2010. Analyzing the protease web in skin: Mephrin metalloproteases are activated specifically by KLK4, 5, and 8 vice versa leading to the processing of proKLK7 thereby triggering its activation. *Biological Chemistry* 391, 455–460.
- Oikonomopoulou, K., DeAngelis, R.A., Chen, H., *et al.*, 2013. Induction of complement C3a receptor responses by kallikrein-related peptidase 14. *Journal of Immunology* 191, 3858–3866.
- Oikonomopoulou, K., Hansen, K.K., Saifeddine, M., *et al.*, 2006. Proteinase-activated receptors, targets for kallikrein signaling. *Journal of Biological Chemistry* 281, 32095–32112.
- Pathak, M., Wong, S.S., Dreveny, I., Emsley, J., 2013. Structure of plasma and tissue kallikreins. *Thrombosis and Haemostasis* 110, 423–433.
- Peters, J., Takahashi, S., Tada, M., Miyake, Y., 1992. mGK-6-derived true tissue kallikrein is synthesized, processed, and targeted through a regulated secretory pathway in mouse pituitary AtT-20 cells. *Journal of Biochemistry* 111, 643–648.
- Petraki, C.D., Karavana, V.N., Diamandis, E.P., 2003. Human kallikrein 13 expression in normal tissues: An immunohistochemical study. *Journal of Histochemistry and Cytochemistry* 51, 493–501.
- Petraki, C.D., Papanastasiou, P.A., Karavana, V.N., Diamandis, E.P., 2006. Cellular distribution of human tissue kallikreins: Immunohistochemical localization. *Biological Chemistry* 387, 653–663.
- Pezzato, E., Sarter, L., Dell'Aica, L., *et al.*, 2004. Prostate carcinoma and green tea: PSA-triggered basement membrane degradation and MMP-2 activation are inhibited by (-)epigallocatechin-3-gallate. *International Journal of Cancer* 112, 787–792.
- Rahman, M.M., Worthy, K., Elson, C.J., *et al.*, 1994. Inhibitor regulation of tissue kallikrein activity in the synovial fluid of patients with rheumatoid arthritis. *British Journal of Rheumatology* 33, 215–233.
- Ramachandran, R., Eissa, A., Mihara, K., *et al.*, 2012. Proteinase-activated receptors (PARs): Differential signalling by kallikrein-related peptidases KLK8 and KLK14. *Biological Chemistry* 395, 421–427.
- Ramani, V.C., Kaushal, G.P., Haun, R.S., 2011. Proteolytic action of kallikrein-related peptidase 7 produces unique active matrix metalloproteins-9 lacking the C-terminal hemopexin domains. *Biochimica et Biophysica Acta* 1813, 1525–1531.
- Roche e Silva, M., Beraldo, W.T., Rosenfield, G., 1949. Bradykinin, a hypotensive and smooth muscle stimulating factor released from plasma globulin by snake venoms and by trypsin. *American Journal of Physiology* 156, 261–273.
- Saito, H., Ratnoff, O.D., 1979. Interactions among Hageman factor (HG, factor XII), plasma thromboplastin antecedent (PTA, factor XI), plasma prekallikrein (PK, Fletcher factor) and high molecular weight kininogen (HMWK, Fitzgerald factor) in blood coagulation. *Advances in Experimental Medicine and Biology* 120B, 61–70.
- Schachter, M., 1979. Kallikreins (kininogenases) — A group of serine proteases with bioregulatory actions. *Pharmacology Reviews* 31, 1–17.
- Schachter, M., Thain, E.M., 1954. Chemical and pharmacological properties of the potent, slow contracting substance (kinin) in wasp venom. *British Journal of Pharmacology* 9, 352–359.
- Schmaier, A.H., McCrae, K.R., 2007. The plasma kallikrein-kinin system: Its evolution from contact activation. *Journal of Thrombosis and Haemostasis* 5, 2323–2329.
- Shaw, J.L.V., Diamandis, E.P., 2007. Distribution of 15 human kallikreins in tissues and biological fluids. *Clinical Chemistry* 53, 1423–1432.
- Shimamoto, K., Mayfield, R.K., Margolius, H.S., *et al.*, 1984. Immunoreactive tissue kallikrein in human serum. *Journal of Laboratory and Clinical Medicine* 103, 731–738.
- Slim, R., Torremocha, F., Moreau, T., *et al.*, 2002. Loss-of-function polymorphism of the human kallikrein gene with reduced urinary kallikrein activity. *Journal of the American Society of Nephrology* 13, 968–976.
- Sotiropoulou, G., Pampalakis, G., 2010. Kallikrein-related peptidases: Bridges between immune functions and extracellular matrix degradation. *Biological Chemistry* 391, 321–331.
- Stefansson, K., Brattsand, M., Roosterman, D., *et al.*, 2008. Activation of proteinase-activated receptor-2 by human kallikrein-related peptidases. *Journal of Investigative Dermatology* 128, 18–25.
- Stephenson, S.A., Verity, K., Ashworth, L.K., *et al.*, 1999. Localization of a new prostate-specific antigen-related serine protease gene, KLK4, is evidence for an expanded human kallikrein gene family cluster on chromosome 19q13.3-13.4. *Journal of Biological Chemistry* 274, 23210–23214.
- Stone, O.A., Richer, C., Emanuel, C., *et al.*, 2009. Critical role of tissue kallikrein in vessel formation and maturation: Implications for therapeutic revascularization. *Arteriosclerosis Thrombosis and Vascular Biology* 29, 657–664.
- Takada, Y., Skidgel, R.A., Erdös, E.G., 1985. Purification of human urinary prokallikrein. Identification of the site of activation by the metalloproteinase thermolysin. *Biochemical Journal* 232, 851–858.
- Teslay, L., Huhn, A.J., Hatcher, H., Torti, F.M., Torti, S.V., 2012. Ferritin block inhibitory effects of two-chain high molecular weight kininogen (HKa) on adhesion and survival signaling in endothelial cells. *PLoS One* 7 (7), e40030.
- Tu, L., Xu, X., Wan, H., *et al.*, 2008. Delivery of recombinant adeno-associated virus-mediated tissue kallikrein for therapy of chronic renal failure in rats. *Human Gene Therapy* 19, 318–330.
- Vakrakou, A., Devetzi, M., Papachristopoulou, G., *et al.*, 2014. Kallikrein-related peptidase 6 (KLK6) expression in the progression of colon adenoma to carcinoma. *Biological Chemistry* 395, 1105–1117.
- Vaporciyan, A.A., Putnam Jr., J.B., Smythe, W.R., 2004. The potential role of aprotinin in the perioperative management of malignant tumors. *Journal of the American College of Surgeons* 198, 266–278.
- Veneris-Lowe, T.L., Kruger, S.J., Walsh, T., *et al.*, 2007. Seminal fluid characterization for male fertility and prostate cancer: Kallikrein-related serine proteases and whole proteome approaches. *Seminars in Thrombosis and Hemostasis* 33, 87–99.
- Voigt, J.D., Zappala, S.M., Vaughn, E.D., Wein, A.J., 2014. The kallikrein panel for prostate cancer screening: Its economic impact. *Prostate* 74, 250–259.
- Vu, T.K., Hung, D.T., Wheaton, V.I., Coughlin, S.R., 1991. Molecular cloning of a functional thrombin receptors reveals a novel proteolytic mechanism of receptor activation. *Cell* 64, 1057–1068.
- Wang, M.C., Papsidero, L.D., Kuriyama, M., *et al.*, 1981. Prostate antigen: A new potential marker for prostatic cancer. *Prostate* 2, 89–96.
- Wang, W., Mize, G.J., Zhang, X., Takayama, T.K., 2010. Kallikrein-related peptidase-4 initiates tumor-stroma interactions in prostate cancer through protease-activated receptor 1. *International Journal of Cancer* 126, 599–610.
- Yoo, H., Radulovic, M., Wu, J., *et al.*, 2013. Kallikrein 6 signals through PAR1 and PAR2 to promote neuron injury and exacerbate glutamate neurotoxicity. *Journal of Neurochemistry* 127, 283–298.
- Yoshida, S., Taniguchi, M., Suemoto, T., *et al.*, 1998. cDNA cloning and expression of a novel serine protease, TLSP. *Biochimica et Biophysica Acta* 1399, 225–228.
- Yousef, G.M., Chang, A., Diamandis, E.P., 2000a. Identification and characterization of KLK-L4, a new kallikrein-related gene that appears to be down-regulated in breast cancer tissues. *Journal of Biological Chemistry* 275, 11891–11898.
- Yousef, G.M., Kapadia, C., Polymeris, M.E., *et al.*, 2003. The human kallikrein protein 5 (hK5) is enzymatically active, glycosylated and forms complexes with two protease inhibitors in ovarian cancer fluids. *Biochimica et Biophysica Acta* 1628, 88–96.
- Yousef, G.M., Magklara, A., Diamandis, E.P., 2000b. KLK12 is a novel serine protease and a new member of the human kallikrein gene family — Differential expression in breast cancer. *Genomics* 69, 331–341.
- Zaika, O., Memenko, M., O'Neil, R.G., Pochynyuk, O., 2011. Bradykinin acutely inhibits activity of the epithelial Na⁺ channel in mammalian aldosterone-sensitive distal nephron. *American Journal of Physiology Renal Physiology* 300, F1105–F1115.

Further Reading

Björkqvist, J., Jämsä, A., Renné, T., 2013. Plasma kallikrein: The bradykinin-producing enzyme. *Thrombosis and Haemostasis* 110, 399–407.

- Clements, J.A., Hooper, J.D., Dong, Y., 2013. The human tissue kallikrein and kallikrein-related peptidase family. In: Rawlings, N.D., Salvesen, G. (Eds.), *Handbook of Proteolytic Enzymes*, third ed. Amsterdam: Elsevier, pp. 2809–2813 (Chapter 606).
- Dong, Y., Brittney, S., Harrington, M.N., *et al.*, 2014. Activation of membrane-bound proteins and receptor systems: A link between tissue kallikrein and the KLK-related peptidases. *Biological Chemistry* 395, 971–990.
- Eden, G., Archinti, M., Furlan, F., Murphy, R., Degryse, B., 2011. The urokinase receptor interactome. *Current Pharmaceutical Design* 17, 1874–1889.
- Garg, V., Gu, N.Y., Borrego, M.E., Raisch, D.W., 2013. A literature review of cost-effectiveness analyses of prostate-specific antigen test in prostate cancer screening. *Expert Reviews of Pharmacoeconomic and Outcomes Research* 13, 327–342.
- Kontos, C.K., Scorilas, A., 2012. Kallikrein-related peptidases (KLKs): A gene family of novel cancer biomarkers. *Clinical Chemistry and Laboratory Medicine* 50, 1877–1891.
- Lundwall, A., Brattsand, M., 2008. Kallikrein-related peptidases. *Cellular and Molecular LifeSciences* 65, 2019–2038.
- Ma, J.X., Song, Q., Hatcher, H.C., *et al.*, 1996. Expression and cellular localization of the kallikrein-kinin system in human ocular tissues. *Experimental Eye Research* 63, 19–26.
- Pampalakis, G., Sotiropoulou, G., 2007. Tissue kallikrein proteolytic cascade pathways in normal physiology and cancer. *Biochimica et Biophysica Acta* 1776, 22–31.
- Rawlings, N.D., Barrett, A.J., 2013. Introduction: Serine peptidases and their clans. In: Rawlings, N.D., Salvesen, G. (Eds.), *Handbook of Proteolytic Enzymes*, third ed. Amsterdam: Elsevier, pp. 2491–2525 (Chapter 559).
- Rennè, T., 2012. The procoagulant and proinflammatory plasma contact system. *Seminars in Immunopathology* 34, 31–41.
- Rothmeier, A.S., Ruf, W., 2012. Protease-activated receptor 2 signaling in inflammation. *Seminars in Immunopathology* 34, 133–149.
- Schmaier, A.H., McCrae, K.R., 2007. The plasma kallikrein-kinin system: Its evolution from contact activation. *Journal of Thrombosis and Haemostasis* 5, 2323–2329.
- Yoon, H., Laxmikanthan, G., Lee, J., *et al.*, 2007. Activation profiles and regulatory cascades of the human kallikrein-related peptidases. *Journal of Biological Chemistry* 282, 31852–31864.

Naturally-Occurring Polypeptide Inhibitors: Cystatins/Stefins, Inhibitors of Apoptosis (IAPs), Serpins, and Tissue Inhibitors of Metalloproteinases (TIMPs)

K Brew and M Lizotte-Waniewski, Florida Atlantic University, Boca Raton, FL, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Apollon The human homolog of *Drosophila* BRUCE, a high molecular weight inhibitor of apoptosis (IAP), also known as BIRC6 protein.

Apoptosome A multimeric protein complex that mediates the activation of a caspase that initiates apoptosis; the human complex includes cytochrome c, Apaf-1 (apoptotic protease activating protein-1) and deoxy-ATP.

Cathelicidins A family of antimicrobial peptides found in the secondary granules of neutrophils and in some other cells.

Cathepsins A family of cysteine proteases related to papain, members of family C1 in the MEROPS database of proteolytic enzymes.

Fetuin (α 2-HS-glycoprotein) A multifunctional cystatin found in serum.

Inflammasome A multi-protein complex containing NOD-like receptors that promote the proteolytic activation of pro-inflammatory cytokines, IL1 β and IL18.

Latexin A mammalian inhibitor of metallocarboxypeptidases A, a member of MEROPS inhibitor family I47, clan IH.

Neddylation The process of covalent coupling of proteins with the ubiquitin-like protein, NEDD8.

Proteasome A large protein complex that degrades ubiquitin-tagged proteins in the cytoplasm and nucleus.

Serinopathy A disease caused by a misfolded serpin.

Survivin An IAP and regulator of mitosis, expressed in tumors and fetal cells, also known as BIRC5.

Tazarotene A retinoid pro-drug that is de-esterified to form tazarotenic acid, which binds to members of the retinoic acid receptor (RAR) family; used in treatments for acne and psoriasis.

Thrombophilia A disorder in which there is an increased tendency to form blood clots.

Ubiquitination The covalent coupling of proteins with ubiquitin.

General Introduction

Proteases catalyze peptide bond hydrolysis, a reaction underlying protein digestion and turnover. However, most proteases have restricted specificities reflecting roles in specific physiological processes, for example, the release of biologically active peptides from larger protein precursors, the activation of proenzymes and receptors, removal of localization signals, or cell signaling. In these processes, protease activity is needed for a limited duration, requiring the regulated production of active proteases through transcription, localization and zymogen activation and termination by endogenous inhibitors, structural change or degradation. Protease inhibition, which generally involves the action of specific protein inhibitors, is essential for avoiding the potentially pathogenic effects of unregulated proteolysis. Typically, the production of protein inhibitors accompanies the production of their cognate proteases; however, some pathogens produce inhibitors of specific proteases made by their hosts.

Polypeptide protease inhibitors vary in specificity, structure, and mechanism of action. Homologous inhibitors tend to target a particular class of protease, but that is not universal. Some families include members with diverse functions that inhibit different protease classes or lack inhibitory activity. This article focuses on some key families of protease inhibitors that vary in origins, chemistry, and biology. This article is not comprehensive; omissions include canonical inhibitors of serine proteases (e.g., pancreatic trypsin inhibitor), α ₂-macroglobulin, and an array of bacterial, yeast, and invertebrate inhibitors. Also, the inhibition of two medically important proteases: angiotensin I

converting enzyme (ACE) or HIV protease is not discussed because they lack protein inhibitors. A more comprehensive and detailed account of protease inhibitors is provided in MEROPS database (see 'Relevant Website' section).

Cystatins/Stefins

The cystatins (CSTs) are high affinity, reversible inhibitors of cysteine proteases including plant proteases related to papain and mammalian cathepsins. They are found in eukaryotes (animals, protozoa, plants, and algae) and bacteria, but not in archaea. CSTs fall into three groups; type 1: intracellular proteins (stefins A and B) that have about 100 residues and lack disulfide bonds and a secretion signal peptide; type 2: secreted CSTs (C, D, E, F, S, SA, and SN) with about 120 amino acids that have two conserved disulfide crosslinks and a secretion signal sequence; type 3: kininogens which are multidomain proteins containing three type-2 CST domains of which one is inactive as a protease inhibitor. CST domains are also present in other proteins that do not function as protease inhibitors including fetuin-A (histidine-rich glycoprotein or HRG). CSTs S, SA, and SN are closely similar in sequence (90% identical). Molecular phylogeny studies indicate that the CSTs originated in an ancestor of the eukaryotes from an intracellular form; a gene duplication subsequently generated two lineages: stefins (type 1) and CSTs (types 2 and 3). The CST line has diverged functionally and structurally during the evolution of multicellular eukaryotes through multiple domain and gene duplications (Kordiš and Turk, 2009).

Structure and Inhibition Mechanism

The structures of a number of type 1 and type 2 CSTs have been determined including human CSTs A, B, C, D, and F, chicken, tick, and plant CSTs, together with CST-protease complexes. Their shared CST fold consists of a 5-stranded antiparallel β -sheet that is wrapped around a 5-turn α -helix (Figure 1). CSTs inhibit cysteine proteases by blocking the active site; the regions that interact with the protease are the N-terminus and two loops connecting the second and third and fourth and fifth β -strands. This forms a wedge-like structure that inserts into the active site but does not interact with the catalytic cysteine (see Figure 1).

Tissue Distribution

Stefin A (type 1 CST) has a restricted tissue distribution, being found mainly in epidermal cells whereas stefin B is expressed in a wide range of cells. Among the type 2 CSTs, C is widely distributed in different tissues while D, S, SA, and SN are largely found in glandular tissues. CST F is targeted to endosomes and lysosomes and is selectively expressed in immune cells. CST 8 (CRES, CST-related epididymal spermatogenic protein), 9 (testatin), 11 and 12 (CST T) are localized in the male reproductive tract, while CST13 (CLM, cystatin-like

molecule) is in bone marrow and CST14 is a secreted phosphoprotein, also known as ssp24.

Biological Functions

The stefins (Type 1 CSTs) retain the intracellular cysteine protease inhibitory function of the CST ancestor, while proteins encoded by the lineage leading to type 2 and 3 CSTs acquired a secretion signal sequence and disulfide bonds, expanding through multiple gene duplications to produce proteins with diverse extracellular functions. Type 1 and 2 CSTs inhibit papain-like proteases and the lysosomal cysteine proteases, including cathepsins B, L, H, and S, but with varying affinities. Some type 2 CSTs are found in body fluids, including saliva and tears (S, SA, and SN). CRES (CST8) has been reported to inhibit the serine protease, prohormone convertase and is important for fertility of spermatozoa.

Type 3 CSTs are more variable in structure and function. Fetuin-A (histidine-rich glycoprotein, HRG) in plasma and platelets binds heparin. CST14 is an abundant serum phosphoprotein that binds cytokines of the bone morphogenetic protein/transforming growth factor- β (BMP/TGF β) superfamily. Cathelicidin has a variable C-terminal antimicrobial domain attached to a conserved CST region. Processing by skin proteases enables the antimicrobial activity. Tazarotene-induced gene 1 (TIG1) is known to function as a cell adhesion molecule, which leads to better cell–cell contacts and reduced proliferation of target cells. Latexin inhibits Zn²⁺-carboxypeptidase A4 in mammals and may be involved in the homeostasis of hematopoietic stem cells. Kininogen contains three type 2 CST domains, in tandem, of which two inhibit cysteine proteases in the circulation (lysosomal and microbial). Its cleavage by plasma kallikrein releases bradykinin, which reduces blood pressure through vasodilation.

CSTs and Disease

Mutations in stefin B, resulting from expansion of a nucleotide repeat, cause myoclonic epilepsy of Unverricht and Lundborg (EPM1), an autosomal recessive disorder that results in neurodegeneration at 6–13 years of age. The disease progresses during adolescence but tends to stabilize in adulthood. Two different mutations in CST C have been linked to diseases. A Leu68 to Gln mutation causes cerebral ‘icelandic-type’ amyloidosis, accompanied by intracerebral hemorrhage in the elderly (Ghiso *et al.*, 1986) while the Ala25 to Thr mutation causes a form of age-related macular degeneration (AMD). The latter appears to result from intracellular mislocalization to mitochondria (Zurdel *et al.*, 2002). Mutations in HRG resulting from missense mutations can result in low circulating levels, arising from reduced secretion and a thrombophilia. Paradoxically, elevated levels of HRG is also linked to an increased risk of thrombosis (OMIM 613116).

Inhibitors of Apoptosis

Inhibitors of apoptosis (IAPs) are a family of highly conserved proteins, initially discovered in baculovirus, that are now known to include representatives in both vertebrates and



Figure 1 Ribbon structure of the complex of stefin A (green) with cathepsin H (orange). The N- and C-terminal residues of each protein are labeled as well as two loops (L1, L2) of stefin A that, together with the N-terminal region, form most of the interaction interface with the protease (Jenko *et al.*, 2003). The image was generated from chains A and I of PDB 1NB5 using Chimera (Pettersen *et al.*, 2004).

invertebrates. The IAPs were initially categorized as caspase inhibitors that prevented apoptosis in insect cells, thereby allowing the baculovirus to persist (Crook *et al.*, 1993). Subsequently, it was found that IAPs mediate other cell-signaling processes that are independent of apoptosis. These activities include regulation of inflammatory responses and innate immune signaling, the control of cell motility/migration, invasion, metastasis, and differentiation. In humans, there are eight members of the IAP family: cellular IAP1 (c-IAP1/BIRC2), cellular IAP2 (c-IAP2/BIRC3), X-chromosome linked IAP (X-IAP/BIRC4), neuronal inhibitor apoptosis protein (NIAP/BIRC1), Survivin (TIAP/BIRC5), Apollon (BRUCE/BIRC6), melanoma IAP (ML-IAP/BIRC7), and IAP-like protein 2 (ILP-2/BIRC8). Interestingly, only four are direct regulators of apoptosis (c-IAP, c-IAP2, X-IAP, and ML-IAP), while the other proteins have more indirect, if any, effects on this process (Berthelet and Dubrez, 2013).

Structure

IAPs range in size from 142 amino acids (Survivin) to 4858 amino acids (Apollon). The defining feature of an IAP family member is the presence of at least one baculovirus IAP repeat (BIR) domain near the N-terminus (Figure 2). This zinc-binding domain is comprised of 70–80 amino acids containing a conserved cysteine and histidine motif that mediates protein-protein interactions. The BIR domain has an α/β fold;



Figure 2 The structure of XIAP_HUMAN BIR2 Domain Chain A PDB 4J3Y. Secondary structure elements are displayed as ribbons and the side chains of three Zn-binding cysteines and one histidine are also shown in purple and the Zn^{2+} ion is a purple sphere. The image was rendered using Chimera (Pettersen *et al.*, 2004).

when more than one BIR is present, the domains are arranged in tandem. Type-II BIR domains have a hydrophobic groove that binds a conserved four amino acid protein motif called the IAP binding motif (IBM) found in some processed/activated caspases and IAP antagonists. Type I BIR domains act on proteins mainly involved in cell-signaling pathways in an IBM-independent manner.

In addition to the signature BIR domain, five mammalian IAPs (c-IAP1, c-IAP2, X-IAP, ILP-2, and ML-IAP) contain a C-terminal RING (really interesting new gene) domain. This zinc finger-like domain confers E3 ubiquitin-ligase activity, which facilitates ubiquitination of targets with varying effects. Finally some members of the IAP family share a caspase recruitment domain (CARD) and/or ubiquitin-associated domain (UBA) that mediate the enlistment of caspases or binding to ubiquitin moieties, respectively.

Biological Functions

IAPs are many and varied; some are multifunctional, whereas others have limited roles. The effects of IAPs on numerous pathways elicit responses that include cell death, inflammation, cell migration, cell division, and development.

Inhibition of apoptosis

Programmed cell death is a key mechanism for terminating cells under conditions of stress or at key points during development. There are two mechanisms (intrinsic and extrinsic) for initiating apoptosis. Activation of the intrinsic pathway involves the release of mitochondrial pro-apoptotic molecules (such as cytochrome *c* and second mitochondria-derived activator of caspases – SMAC/DIABLO) in response to stimuli such as free radicals, radiation or viral infections. Cytochrome *c* then initiates formation of an apoptosome, typically followed by activation of caspase 9. The SMAC protein acts by neutralizing IAPs. Conversely, the extrinsic pathway is activated by the binding of specific ligands to cell surface ‘death’ receptors of the tumor necrosis factor receptor (TNFR) family, followed by the activation of an initiator caspase (primarily caspases 8 or 10). In either mechanism, the signaling pathways rely on the serial activation of initiator followed by effector caspases (such as 3, 6, and 7). The effector caspases subsequently facilitate apoptosis. Of the human IAPs, X-IAP, ML-IAP, c-IAP1, and c-IAP2 can bind to caspases -3, -7, and -9 via their BIR domains. Their RING domains then allow for ubiquitination (or neddylation) of caspases. Interestingly, only X-IAP is able to act to directly inhibit caspases -3 and -7, by blocking the substrate binding sites, as well as -9, by preventing homo-dimerization (Berthelet and Dubrez, 2013). Other IAPs such as survivin and NIAP also have roles in apoptosis but their primary functions may be in other pathways. ILP2 protein (a tissue-specific homolog of X-IAP) is unstable and has not yet been shown to inhibit apoptosis under physiologically relevant conditions.

Regulation of inflammation and immunity

X-IAP, c-IAP1, and c-IAP2 function in the initial immune responses to tissue injury or infection, principally through their ubiquitin-ligase activity. Inflammation is initiated by the

binding of innate immune pattern recognition receptors (PRRs) to stress signals such as microbial components (pathogen-associated molecular patterns – PAMPs) or endogenous molecules released by dying cells (danger/damage-associated molecular patterns – DAMPs). Key PPRs include Nod-like receptor (NLR), Toll-like Receptor (TLR), TNFR1, and RIG-1-like receptor (RLR). The outcome of the activation of certain PPRs is initiation of the mitogen-activated protein kinase (MAPK) and/or the traditional and alternate nuclear factor- κ B (NF- κ B) signaling pathways in which the IAPs are key effectors (Estornes and Bertrand, 2014). The downstream effect of activating these pathways is often the upregulation of genes for mediators of inflammation and immunity, such as cytokines. However, the c-IAPs, under appropriate signaling can also repress one of the NF- κ B pathways. IAPs can also impact the innate immune system through NIAP, c-IAP1, and c-IAP2 control of the inflammasome (a caspase 1 activating multi-protein molecular platform). This complex is ultimately responsible for producing and secreting interleukin 1 (IL-1). c-IAP1 and c-IAP2 also affect the adaptive immune response through their regulation of TNF superfamily signaling. Both proteins are negative regulators of B lymphocyte proliferation and survival through the NF- κ B alternate pathway (Beug *et al.*, 2012).

Regulation of cell migration, invasion, metastasis, and development

The activation of the NF- κ B and/or MAP Kinase signaling pathways, that involve IAPs, affect other cellular activities besides immune responses. In higher eukaryotes, it appears that the IAPs can be positive or negative effectors of migration, invasion and metastasis, or both, under varying circumstances (Kenneth and Duckett, 2012). For example, some data suggest that the c-IAPs, upon activation of the traditional (canonical) arm of the NF- κ B pathway, can ultimately lead to an increase in interleukin-8, which will then promote migration. Conversely, other evidence suggests that c-IAP signaling in the alternate (noncanonical) arm of the NF- κ B pathway can have a negative effect on these phenotypes (Fulda, 2014). In addition to these actions of IAPs, X-IAP, and c-IAP1 are known to interact directly with the Rho GTPases that are the molecular switches modulating the dynamics of the cytoskeleton. These proteins cycle between the active (GTP-bound) form and the inactive (GDP-bound form) and interact with downstream effectors that can exert either positive or negative effects on cell movement. In colorectal carcinoma cells, data from one study suggest that X-IAP has a positive effect on actin polymerization and cell motility whereas in HeLa cells, it appears that c-IAP1 and X-IAP can limit cell migration by targeting regulators of the actin cytoskeleton for proteasomal degradation (Kenneth and Duckett, 2012).

Appropriate cell movement and apoptosis are crucial for development in higher eukaryotes. Since IAPs modulate these developmental processes, their activity during embryogenesis is vital. Studies with model organisms, including *Drosophila* and *Xenopus*, show that disruption of various IAP genes can lead to embryonic death or incorrect cell migration leading to ventralized embryos (Kenneth and Duckett, 2012). With regard to mammals, experiments in mice have shown that knockouts of either X-IAP, c-IAP1, or c-IAP2 permit relatively

normal development. However, X-IAP/c-IAP1 or c-IAP1/c-IAP2 double knockouts do not survive past day 10 of development. This suggests that the activities of these proteins may overlap, allowing mutual compensation during embryonic development.

IAPs involved in cell division and cytokinesis

The smallest (Survivin) and largest (Apollon) of the IAP proteins are implicated in cell division and cytokinesis, respectively. Survivin transcription has been shown to increase during the G2 to mitosis transition in the cell cycle. As a subunit of the chromosome passenger complex (CPC), Survivin acts by targeting the CPC to the inner centromeres (Jeyaprakash *et al.*, 2007). The presence of the CPC generates a spindle checkpoint signal that prevents the onset of anaphase until all spindle fibers are appropriately attached to the centromeres. At a later point in the cell cycle, it is thought that Apollon may interact with Survivin at the end stages of cytokinesis. Recently it has also been shown that Apollon is a novel regulator of mitotic Cyclin A protein degradation in early mitosis (Kikuchi *et al.*, 2014).

IAPs in Disease

There is considerable evidence that IAP proteins play a role in human cancers. Increased expression of certain IAP proteins (particularly X-IAP, c-IAP1, and c-IAP2) in tumors results in a worse prognosis but in other kinds of cancers, increases in IAPs can lead to a more positive prognosis. In addition, Survivin and ML-IAP (which are usually at undetectable levels in normal adult cells) are highly expressed in most tumors and specifically in melanomas, respectively. These and other studies, suggest that the IAP proteins may represent promising targets for cancer researchers. Because certain IAP proteins prevent cell death, inhibitors of these proteins may facilitate apoptosis producing a more effective treatment and improved prognosis. Currently several small molecule inhibitors of IAPs, that are SMAC mimics, are being tested as potential therapeutic tools for cancer (Dubrez *et al.*, 2013).

Serpins

Members of this ubiquitous family are found in eukaryotes, bacteria, archaea and viruses. The acronym 'SERPIN' (serine protease inhibitors) reflects the functions of the first characterized members of the family as specific inhibitors of serine proteases. There are 36 genes for serpins in humans and 60 in mice; while most function as serine protease inhibitors, some inhibit cysteine proteases (caspases and papain-like cysteine proteases) and others have a broad array of functions, including molecular chaperone activity, DNA binding, or hormone transport (Heit *et al.*, 2013). Serpins are classified in family I4 of the MEROPS database; two major divisions of the serpins are the extracellular and intracellular groups (clades A and B, respectively). The inhibitory serpins inactivate their target proteases through an unusual 'suicide' mechanism that exploits the acyl-enzyme intermediate formed during covalent catalysis by serine and cysteine proteases.

Structure

Compared with other protease inhibitors, serpins are relatively large molecules, containing 330 to 500 residues. Their structures, exemplified by the well-studied α_1 protease inhibitor (α_1 PI) have been described as containing two domains, an N-terminal α -helical domain and a C-terminal β -barrel domain. Two key structural features are a mixed (part antiparallel and part antiparallel) β -sheet, designated β -sheet A, that runs between the two domains, and a loop that extends from the β -barrel domain, that is named the reactive center loop (RCL) (Figure 3(a)). Mixed β -sheets like the A sheet are less common in proteins because the H-bonding arrangements in antiparallel and parallel β -sheets require different positioning of peptide groups.

Mechanism of Inhibition

Serine proteases

The inhibitory mechanism of α_1 PI (and some other serpins, including antithrombin III (AT)) has been established by structural and biochemical studies. The protease recognizes a substrate sequence in the RCL, and binds to this to form a Michaelis complex. During cleavage of the peptide bond between Met 358 and Ser 359, an acyl-enzyme intermediate is formed between the carboxyl group of Met 358 on the N-terminal side of the bond and the active site serine of α_1 PI, with release of the polypeptide chain of the serpin C-terminal

to the cleaved bond. This triggers a large conformational change in the section of the RCL N-terminal to the cleaved bond, which moves 70 Å and inserts into the A sheet of α_1 PI between the two parallel strands, generating a 5-stranded antiparallel β -sheet (Figure 3(b)). The RCL carries the covalently attached protease, which is pulled against the body of the serpin, disrupting the catalytic triad by drawing away the catalytic serine (Huntington *et al.*, 2000). While both the RCL and A β -sheet play key roles in the protease inhibitory activity of serpins, other substructures are also important for the conformational change, including the 'hinge' (located in the N-terminal region of the RCL) and the 'breach' (the point in β -sheet A where the cleaved RCL is inserted). The conformational change generated by RCL cleavage increases the stability of the serpin molecule, reflecting a transition from a 'stressed' to 'relaxed' state. This reflects the metastable nature of the active (native) serpin structure, which is kinetically stabilized but is thermodynamically unstable.

Cysteine proteases

These proteases have a covalent mechanism, similar to serine proteases, involving the formation of a thioester bond between the active site cysteine and the carboxyl group of the cleaved peptide bond. The structure of the cleaved form of cytokine response modifier A (CrmA, an anti-apoptotic caspase inhibitor produced by cowpox virus) shows the RCL inserted into sheet A, suggesting that the mechanism of caspase

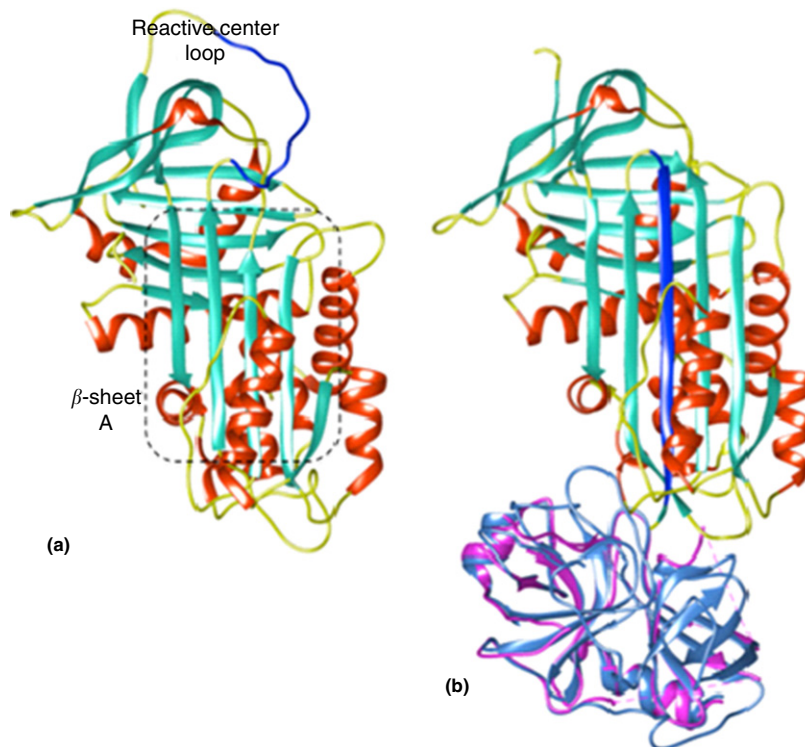


Figure 3 Ribbon structures of bovine α_1 PI (a) and its inhibitory complex with trypsin (b), showing the structural changes produced during inhibition. β -Strands are colored cyan, α -helices in orange, and other regions in yellow. The region of α_1 PI that is colored in dark blue, initially part of the reactive center loop (RCL), is inserted as an extra strand in β -sheet A. Changes in trypsin are highlighted by superimposing native trypsin (blue) on inhibited trypsin in the complex (pink). Image (a) was generated from PDB file 1QLP and image (b) from PDB 1EZK and 4I8G using Chimera (Pettersen *et al.*, 2004).

inhibition is closely similar to that of serine proteases by other serpins (Renatus *et al.*, 2000).

Activation and Latency

The activities of some serpins, including AT, are modulated by sulfated oligosaccharides, such as heparin. Native AT is a poor inhibitor of thrombin because the key residue of the RCL, Arginine 394, points toward the main body of the inhibitor. The binding of heparin induces a conformational change that releases the arginine and greatly enhances inhibitory activity. In some serpins, the carbohydrate bridges the inhibitor and its target protease. Also, most serpins can undergo a spontaneous structural change to a relaxed, inactive, structure in which the RCL is inserted into the A sheet without cleavage; an extra strand can be added also by rearranging a different part of the serpin structure (e.g., an α -helix). This limits the biological lifetime of some serpins as active inhibitors.

Biological Functions

The irreversible inactivation of proteases by serpins is singularly important in controlling the activity of protease cascades that require stringent regulation, such as blood coagulation and fibrinolysis, and cytosolic proteolysis, that can lead to apoptosis. AT has a crucial role in terminating the coagulation process, being an inhibitor of the intrinsic pathway, in which it inhibits activated forms of thrombin (Factor II), Factor VII, Factor X, Factor IX, Factor XI, and Factor XII. Although α 1PI can inhibit trypsin and other serine proteases, its biological target is leukocyte elastase, particularly in the lung, where excess proteolytic activity can lead to tissue degradation and emphysema. Plasminogen activator inhibitor-2 (PAI-2) has both extracellular and intracellular functions. In the extracellular environment it functions in regulating fibrinolysis through its inhibitory action on plasminogen activators, and has also been reported to have roles in tumor cell metastasis and embryo implantation. The cytoplasmic form of PAI-2 has been shown to interact with and inhibit the proteasome, thereby promoting apoptosis. Other cytoplasmic serpins have been shown to regulate necrosis and apoptosis.

Disease Aspects

The metastable character of the native serpin structure and its tendency to undergo conformational changes are linked to misfolding diseases (serpinopathies) caused by mutations in several human serpins (Roussel *et al.*, 2011). Homozygotes with mutations in α 1PI (S and Z variants) have low circulating levels of the inhibitor and a predisposition to emphysema because of lung damage produced by unregulated leukocyte elastase. The Z variant polymerizes in the endoplasmic reticulum of the liver, its site of synthesis, through a domain-swapping mechanism in which the RCL of one molecule inserts into β -sheet A of another. The accumulation of aggregates of the protein causes hepatocyte cell death and cirrhosis. Mutations in neuroserpin cause familial encephalopathy involving seizures and dementia through a similar polymerization mechanism that generates inclusions in neurons (Miranda *et al.*, 2008). Polymerization is also a feature of

thrombosis arising from mutations in AT. Hereditary C1 esterase inhibitor deficiency is an autosomal dominant disorder caused by mutations in the serpin, C1 esterase inhibitor, that variously lead to low levels of the protein in the circulation or a dysfunctional inhibitor. The symptoms, mostly attributed to inadequate regulation of kallikrein and overproduction of bradykinin, include swelling of skin and mucous membranes including face, abdominal organs and the upper airways; laryngeal edema can produce asphyxiation.

Tissue Inhibitors of Metalloproteinases

These extracellular metalloproteinase inhibitors are widely distributed in animals. Humans have four tissue inhibitors of metalloproteinases (TIMPs), numbered based on their order of discovery, and about 40% identical in sequence. TIMP-1 was first identified as an erythroid potentiating protein, and subsequently purified as an inhibitor of fibroblast collagenase. TIMPs are high affinity, slow-binding inhibitors of the matrix metalloproteinases (MMPs), members of family M10 in the MEROPS database. TIMPs show little discrimination between different MMPs, but TIMP-1 is a relatively weak inhibitor of the membrane-type MMPs (MMP-14, MMP-16, and MMP-24) and MMP-19. TIMP-3 inhibits more metalloproteinases than the other TIMPs, being a potent inhibitor of some important disintegrin-metalloproteinases (MEROPS family 12), particularly a disintegrin and metalloproteinase domain-17 (ADAM-17) (TNF- α -converting enzyme or tumor necrosis factor- α converting enzyme (TACE)), the principle sheddase of mammalian cells, and a disintegrin and metalloproteinase domain with thrombospondin type 1 domains (ADAM-TS- 4 and -5), (aggrecanases-1 and -2).

Structures

Mammalian TIMPs have two domains with larger (~125 residue) N-terminal domains and smaller (~65 residue) C-terminal domains; each domain contains six cysteines and is stabilized by three disulfide bonds (Figure 4).

The N-terminal domains can be expressed separately and fold to form fully active MMP inhibitors. The fact that TIMP-like molecules from nematodes and bacteria do not have C-terminal domains suggests that TIMPs evolved from single-domain ancestors. TIMP C-domains mediate interactions with pro-MMPs in which they interact with the hemopexin domains of the zymogens. Crystallographic structures have been determined for full-length TIMP-1 and TIMP-2 as well as N-TIMP-1 and N-TIMP-3 in complexes with active MMPs. The structure of free TIMP-2 is also known. The N-domains consist of a five-stranded twisted β -barrel with three α -helices, an OB (oligonucleotide and oligosaccharide binding) fold that also found in netrins, agrin, and some complement proteins. TIMP C-domains have a pair of parallel β -strands and two antiparallel strands separated by an α -helix and are closely packed against the N-domains.

Mechanism of Inhibition of MMPs

TIMPs interact directly with the active sites of MMPs in a quasi-substrate mode, denying access to substrates. The core of the

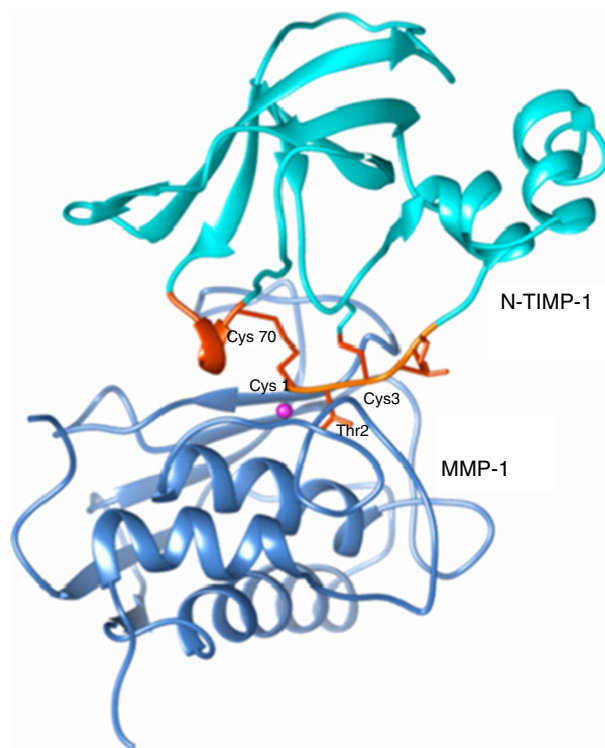


Figure 4 Ribbon structure of the complex of N-TIMP-1 with the catalytic domain of MMP-1. The core of the TIMP interaction site around the Cys1-Cys70 disulfide bond is colored orange and the catalytic Zn^{2+} is pink. The image was generated from PDB: 2JOT using Chimera (Pettersen *et al.*, 2004).

reactive site in the inhibitor is formed by the N-terminal five residues of the N-domain (Cys¹-Thr-Cys-Val-Pro⁵ in TIMP-1) and a five-residue loop (Met⁶⁶-Glu-Ser-Val-Cys⁷⁰ in TIMP-1) between the third and fourth β -strands that is attached to the N-terminal Cys by a disulfide bond. These two regions form a surface ridge that inserts into the MMP active site and is the core of the interaction site, making about 70% of the interactions between the two proteins. The N-terminal Cys sits above the catalytic Zn^{2+} and interacts with it through the α -amino group and carbonyl groups, displacing the water molecule that acts in peptide bond cleavage, while residue 2 (Ser or Thr in mammalian TIMPs) interacts with the P1' specificity pocket of the protease. Residues 3-5 of the TIMP interact with the S2', S3', and S4' subsites of the protease, while residues 68 and 69 (TIMP-1 numbering) interact with the S2 and S3 subsites (Gomis-Ruth *et al.*, 1997; Li *et al.*, 2005). These interactions are conserved in all structurally characterized inhibitory TIMP/MMP complexes. Additional interactions involve loops between the first and second β -strands, a region around the fifth Cys in the protein that is fastened to the core by a disulfide bond with Cys3, and contacts involving the C-domain; these vary in extent and significance in different TIMP/MMP complexes, partly owing to structural differences between TIMPs.

Interactions with pro-MMPs

These interactions are more specific than those with active forms of MMPs: TIMP-1 and TIMP-3 interact with pro-MMP-9

and TIMP-2, TIMP-3, and TIMP-4 can bind to pro-MMP-2. Complexes of TIMPs with pro-MMPs involve the C-domain of TIMP, so that the TIMP N-domains remain available to inhibit active MMPs. The structure of the TIMP-2/pro-MMP-2 complex shows two interactions with the 4-bladed MMP-14 β -propeller hemopexin domain. One involves a loop between the two β -strands of TIMP-2 and the C-terminal blade while the second is between the tail of the TIMP-2 molecule, C-terminal to the last cysteine, and the region between two blades (Morgunova *et al.*, 2002).

Biological Functions

Metalloproteinase inhibition is important for preventing unregulated cleavage of extracellular matrices and uncontrolled shedding of the extracellular domains of cell membrane proteins, including soluble forms of tumor necrosis factor α (TNF- α). Uncontrolled turnover of the proteoglycan, aggrecan, and type-II collagen results in cartilage degradation, leading to osteoarthritis, while excess levels of shed TNF- α can cause inflammatory diseases, notably rheumatoid arthritis. Disruptions of the balance between MMP activity and TIMP production are linked to other disease processes arising from excessive matrix degradation including cancer metastasis, cardiovascular, neurological and kidney diseases, and tissue ulceration.

Functions of TIMPs other than protease inhibition include a role of TIMP-2 in the activation of pro-MMP-2 by MMP-14. Initially, TIMP-2 binds the C-terminal hemopexin domain of pro-MMP-2. The N-domain of TIMP-2 in this complex interacts with an active MMP-14 anchored on the cell membrane and the pro-MMP-2 molecule in this ternary complex is activated by cleavage of its pro-domain by a second MMP-14 molecule. This process is facilitated by interactions between the hemopexin domains of the two MMP-14 molecules. Both TIMP-3 and TIMP-4 can also bind to pro-MMP-2 but inhibit its activation.

Angiogenesis involves cell migration and blood vessel formation, both of which require MMP proteolysis; consequently, TIMPs have anti-angiogenic properties. However, the mechanisms of these effects do not simply involve protease inhibition but also signaling by TIMP-2/integrin interactions and TIMP/VEGF (vascular endothelial growth factor) interactions. TIMP-3 has pro-apoptotic effects on cancer and other cells that are thought to arise from its inhibition of sheddases, including ADAM-17, resulting in the stabilization of death receptors (TNF receptor 1, FAS and TNF-related apoptosis-inducing ligand and receptor 1 (TRAIL-R1)). However, the availability of free TIMP-3 is limited because of its tight binding to the matrix and sequestration by endocytosis (Brew and Nagase, 2010). TIMPs have multiple activities that appear to be unrelated to their interactions with metalloproteinases (reviewed by Stetler-Stevenson, 2008) including effects of TIMP-1 and TIMP-2 on cell growth and differentiation, and angiogenesis. These appear to be mediated by interactions with cellular receptors including TIMP-1 interactions with the tetraspanin, CD63, and TIMP-2 with $\alpha_3\beta_1$ integrin.

Disease Aspects

Mutations in the C-domain of TIMP-3 cause of Sorsby fundus dystrophy (SFD) (Weber *et al.*, 1994), the only human disease

definitively linked to TIMP mutations. SFD is an autosomal dominant disorder that results in early-onset macular degeneration with a pathology similar to AMD. The SFD mutations in TIMP-3 do not affect its inhibitory activity but are localized in the non-inhibitory C-domain. They include missense and truncation mutations that result in unpaired cysteines, suggesting reduced stability. Various mechanisms have been proposed for the disease including increased TIMP-3 accumulation in Bruch's membrane, an extracellular matrix between the retina and choroid. Thickening of this matrix, also seen in AMD, may disrupt nutrient transport.

See also: Cell Division/Death: Apoptosis: Apoptosis; Caspases. Protein Synthesis/Degradation: Protein Degradation – Pathological Aspects: Alpha-1-Antitrypsin Deficiency: A Misfolded Secretory Glycoprotein Damages the Liver by Proteotoxicity and Its Reduced Secretion Predisposes to Emphysematous Lung Disease Because of Protease-Inhibitor Imbalance. Protein Synthesis/Degradation: Protein Degradation – Protease Classes: Matrix Metalloproteinases. Protein Synthesis/Degradation: Proteolytic Pathways: Overview of Blood Coagulation and the Pathophysiology of Blood Coagulation Disorders

References

- Berthelet, J., Dubrez, L., 2013. Regulation of apoptosis by inhibitors of apoptosis (IAPs). *Cells* 2, 163–187.
- Beug, S.T., Cheung, H.H., LaCasse, E.C., Korneluk, R.G., 2012. Modulation of immune signaling by inhibitors of apoptosis. *Trends in Immunology* 33, 535–545.
- Brew, K., Nagase, H., 2010. The tissue inhibitors of metalloproteinases (TIMPs): An ancient family with structural and functional diversity. *Biochimica et Biophysica Acta* 1803, 55–71.
- Crook, T.E., Clem, R.J., Miller, L.K., 1993. An apoptosis-inhibiting baculovirus gene with a zinc-finger-like motif. *Journal of Virology* 67, 2168–2174.
- Dubrez, L., Berthelet, J., Glorian, V., 2013. IAP proteins as targets for drug development in oncology. *Oncology Targets and Therapy* 6, 1285–1304.
- Estornes, Y., Bertrand, M.J.M., 2014. IAPs, regulators of innate immunity and inflammation. *Seminars in Cell and Developmental Biology*. doi:<http://dx.doi.org/10.1016/j.semcdb.2014.03.035>.
- Fulda, S., 2014. Regulation of cell migration, invasion, and metastasis by IAP proteins and their antagonists. *Oncogene* 33, 671–676.
- Ghiso, J., Jenson, O., Frangione, B., 1986. Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Icelandic type is a variant of gamma-trace basic protein (cystatin C). *Proceedings of the National Academy of Sciences of the United States of America* 83, 2974–2978.
- Gomis-Ruth, F.-X., Maskos, K., Betz, M., *et al.*, 1997. Crystal structure of the stromelysin 1-timp-1 complex reveals a novel proteinase inhibition mechanism. *Nature* 389, 77–81.
- Heit, C., Jackson, B.C., McAndrews, M., *et al.*, 2013. Update of the human and mouse SERPIN gene superfamily. *Human Genomics* 7, 22.
- Huntington, J.A., Read, R.J., Carrell, T.W., 2000. Structure of a serpin-protease complex shows inhibition by deformation. *Nature* 407, 923–926.
- Jenko, S., Dolenc, I., Guncar, G., *et al.*, 2003. Crystal structure of Stefin A in complex with cathepsin H: N-terminal residues of inhibitors can adapt to the active sites of endo- and exopeptidases. *Journal of Molecular Biology* 326, 875–885.
- Jeyaprakash, A.A., Klein, U.R., Lindner, D., *et al.*, 2007. Structure of a Survivin–Borealin–INCENP core complex reveals how chromosomal passengers travel together. *Cell* 131, 271–285.
- Kenneth, N.S., Duckett, C.S., 2012. IAP proteins: Regulators of cell migration and development. *Current Opinion in Cell Biology* 24, 871–875.
- Kikuchi, R., Ohata, H., Ohoka, N., Kawabata, A., Naito, M., 2014. Apollon protein promotes early mitotic Cyclin A degradation independent of the spindle assembly checkpoint. *Journal of Biological Chemistry* 289, 3457–3467.
- Kordiš, D., Turk, V., 2009. Phylogenomic analysis of the cystatin family in eukaryotes and prokaryotes. *BMC Evolutionary Biology* 9, 266.
- Li, Z., Clarke, M.P., Barker, M.D., McKie, N., 2005. TIMP3 mutation in Sorsby's fundus dystrophy: Molecular insights. *Expert Reviews in Molecular Medicine* 7, 1–15.
- Miranda, E., MacLeod, I., Davies, M.J., *et al.*, 2008. The intracellular accumulation of polymeric neuroserpin explains the severity of the dementia FENIB. *Human Molecular Genetics* 17, 1527–1539.
- Morgunova, E., Tuuttila, A., Bergmann, U., Tryggvason, K., 2002. Structural insight into the complex formation of latent matrix metalloproteinase 2 with tissue inhibitor of metalloproteinase 2. *Proceedings of the National Academy of Sciences of the United States of America* 99, 7414–7419.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., *et al.*, 2004. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25, 1605–1612.
- Renatus, M., Zhou, Q., Stennicke, H.R., *et al.*, 2000. Crystal structure of the apoptotic suppressor Crma in its cleaved form. *Structure* 8, 789–797.
- Roussel, B.D., Irving, J.A., Ekeowa, U.I., *et al.*, 2011. Unravelling the twists and turns of the serpinopathies. *FEBS Journal* 278, 3859–3867.
- Stetler-Stevenson, W.G., 2008. Tissue inhibitors of metalloproteinases in cell signaling: Metalloproteinase-independent biological activities. *Science Signaling* 1, 6.
- Weber, B.H.F., Vogt, G., Pruetz, R.C., Stohr, H., Felber, U., 1994. Mutations in the tissue inhibitor of metalloproteinases-3 (TIMP3) in patients with Sorsby's fundus dystrophy. *Nature Genetics* 8, 352–356.
- Zurdel, J., Finckh, U., Menzer, G., Nitsch, R.M., Richard, G., 2002. CST3 genotype associated with exudative age-related macular degeneration. *British Journal of Ophthalmology* 86, 214–219.

Relevant Website

<http://merops.sanger.ac.uk/inhibitors/>
MEROPS.

PROTEIN SYNTHESIS/DEGRADATION: PROTEOLYTIC PATHWAYS

Contents

Overview of Blood Coagulation and the Pathophysiology of Blood Coagulation Disorders

Molecular Mechanisms Underlying the Actions of the Complement System

Digestive Proteases: Roles in the Human Alimentary Tract

Overview of Blood Coagulation and the Pathophysiology of Blood Coagulation Disorders

JC Cardenas, University of North Carolina School of Medicine, Chapel Hill, NC, USA and University of Texas Health Science Center at Houston, Houston, TX, USA

CM Rein-Smith, University of North Carolina School of Medicine, Chapel Hill, NC, USA and Whitsell Innovations Inc., Chapel Hill, NC, USA

FC Church, University of North Carolina School of Medicine, Chapel Hill, NC, USA

© 2016 Elsevier Inc. All rights reserved.

Overview of Blood Coagulation

Hemostasis is the process that maintains blood in a fluid state and confines it to the circulatory system. The normal hemostatic system limits blood loss by highly regulated interactions between components of the vessel wall, circulating platelets, and plasma proteins ([Figure 1](#); [Table 1](#)).

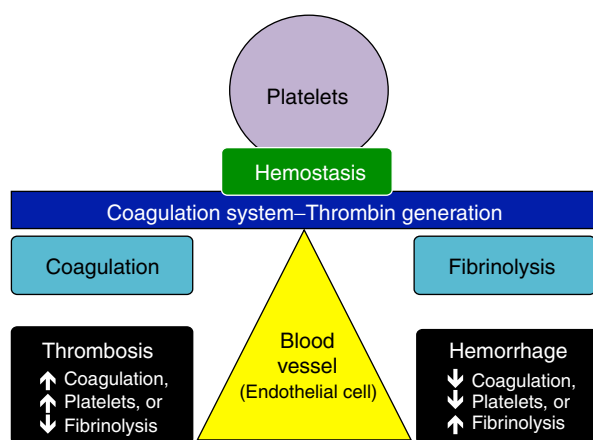


Figure 1 Balance of hemostasis: the interaction of blood platelets, vascular endothelial cells, and blood plasma-coagulation factors are needed for successfully achieving a blood clot. Many different circumstances that are either acquired/environmental or inherited lead to perturbation of hemostasis and pathological conditions to promote hemorrhage or thrombosis.

Historically, primary hemostasis described the role of the cells (platelets and endothelial cells) in blood coagulation, secondary hemostasis described the role of the blood coagulation system, and tertiary hemostasis described the dissolution of the clot by fibrinolysis. There are numerous diseases in which the hemostatic system becomes dysregulated, leading either to hemorrhage or intravascular thrombosis and embolism.

Primary Hemostasis

Platelet Production

Following injury or damage to a blood vessel, several physiological pathways converge to prevent the loss of excessive amounts of blood and to allow the vessel to be repaired. The platelet is one of the first responders to the site of vascular injury, and it plays a significant role in the formation of a plug to stem the loss of blood from the vessel. Platelets are formed within the bone marrow from their precursors, megakaryocytes. The coordinated effort of several cytokines and growth factors, most importantly thrombopoietin (TPO), leads to the maturation of hematopoietic stem cells into megakaryocytes. TPO plays several important roles in the formation of the megakaryocyte, including increasing the production of megakaryocyte precursors, inhibiting megakaryocyte apoptosis, and promoting megakaryocyte maturation ([Vainchenker et al., 2014](#)). Mature megakaryocytes extend projections into the sinusoids of the bone marrow where small pieces of the megakaryocytes are sheared off by flowing blood and become anucleate, granular platelets.

Table 1 Hemostatic and fibrinolytic proteins, characteristics and physiological roles/pathological processes

Protein	Plasma (mcg ml ⁻¹)	Chromosome loci	Physiological pathway	Pathological process (deficiency or mutation)
α_2 -Antiplasmin	70	18p11.1-q11.2	Anti-fibrinolysis	Hemorrhage
Antithrombin	150–400	1q23-q25	Anticoagulation	Thrombosis
Endothelial protein C receptor	–	20q11.22	Anticoagulation	Thrombosis
Factor V	5–10	1q21-q25	Coagulation	Hemorrhage ^a
Factor VII	0.5	13q34	Coagulation	Hemorrhage
Factor VIII	0.1–0.2	Xq28	Coagulation	Hemorrhage ^b
Factor IX	4–5	Xq27.1-q27.2	Coagulation	Hemorrhage ^b
Factor X	8–10	13q34	Coagulation	Hemorrhage
Factor XI	5	4q32-q35	Coagulation	Hemorrhage
Factor XII	30	5q33	Inflammation	–
Factor XIII	10–22	6p24-p25 and 1q31-q32.1	Coagulation	Hemorrhage
Fibrinogen	2000–4000	4q23-q32	Coagulation	Hemorrhage ^b
High-molecular weight kininogen	70	3q26	Inflammation	–
Plasminogen	10–16	6q26-27	Fibrinolysis	Thrombosis
Plasminogen activator inhibitor-1	0.012–0.014	q21.3-q22	Anti-fibrinolysis	Hemorrhage ^b
Prekallikrein (PK)	20	4q35	Inflammation	–
Protein C	4–5	2q13-q14	Anticoagulation	Thrombosis
Protein S	25	3p11.1-q11	Anticoagulation	Thrombosis
Protein Z	2–3	13q34	Anticoagulation	Thrombosis
Protein Z-dependent protease inhibitor (ZPI)	1–1.6	14q32	Anticoagulation	Thrombosis
Prothrombin	100–150	11p11-q12	Coagulation	Hemorrhage ^c
Thrombin activatable fibrinolysis inhibitor (TAFI)	6	13q14	Anti-fibrinolysis	Hemorrhage
Thrombomodulin	–	20p12-cen	Anticoagulation	Thrombosis
Tissue factor	–	1p21-p22	Coagulation	Hemorrhage ^b
Tissue factor pathway inhibitor	0.1	2q31-q32.1	Anticoagulation	Thrombosis
Tissue-type plasminogen activator	0–0.012	8p11.21	Fibrinolysis	Thrombosis
von Willebrand factor	10	12p13.2	Coagulation	Hemorrhage ^b

^aFV_{Leiden} promotes thrombosis.^bIncreased expression is associated with thrombosis.^cProthrombin G20210A promotes thrombosis.

Platelets contain two types of granules: alpha and dense granules. Alpha granules contain chemokines and growth factors, such as platelet-derived growth factor, and proteins involved in primary and secondary hemostasis, including von Willebrand factor (vWF), fibrinogen, and factor V (Blair and Flaumenhaft, 2009). Dense granules contain mediators of primary hemostasis, such as ADP and ATP, serotonin and calcium, which is a required cofactor for coagulation (McNicol and Israels, 1999). Since platelets are anucleate and are mostly incapable of undergoing protein synthesis, the contents of their granules are predetermined by the megakaryocyte from which they bud off. However, there are certain proteins that are taken up into granules via receptor-mediated endocytosis (Machlus *et al.*, 2014; Thon and Italiano, 2012).

Platelet Structure and Function in Primary Hemostasis

Primary hemostasis culminates in the formation of a platelet plug at the site of vascular injury. The proper formation of the platelet plug requires the function of several platelet surface glycoprotein receptors (Varga-Szabo *et al.*, 2008; Cimmino and Golino, 2013). These receptors serve to initially adhere platelets to the vessel wall at the site of injury, as well as to recruit

and aggregate nearby platelets to aid in the growth of the platelet plug.

In the absence of vessel injury, the endothelium of the blood vessel provides a protective anticoagulant surface and synthesizes anticoagulant proteins to prevent aberrant thrombosis. Upon injury, subendothelial proteins, such as type IV collagen, are exposed to flowing blood. Platelets adhere to the exposed subendothelium via vWF, a large multimeric glycoprotein produced by endothelial cells and degranulating platelets. vWF tethers platelets to the damaged endothelium through binding of the vWF A1 domain to glycoprotein Ib–glycoproteinV–glycoprotein IX complex on the surface of the platelet. Additionally, platelets can bind directly to exposed subendothelial collagen via glycoprotein VI on their surface. Through these two interactions, platelets are able to adhere to the site of vessel injury and begin the formation of a hemostatic plug (Varga-Szabo *et al.*, 2008; Andrews and Berndt, 2008; Firbas *et al.*, 2010; Lenting *et al.*, 2010). Once bound at the site of injury, the platelet undergoes an activation characterized by the rearrangement of cytoskeletal elements such as actin and myosin, which leads to a change in the platelet shape, along with prominent filopodia formation, and the release of granule contents from both the alpha and dense granules. Several mediators, including ADP

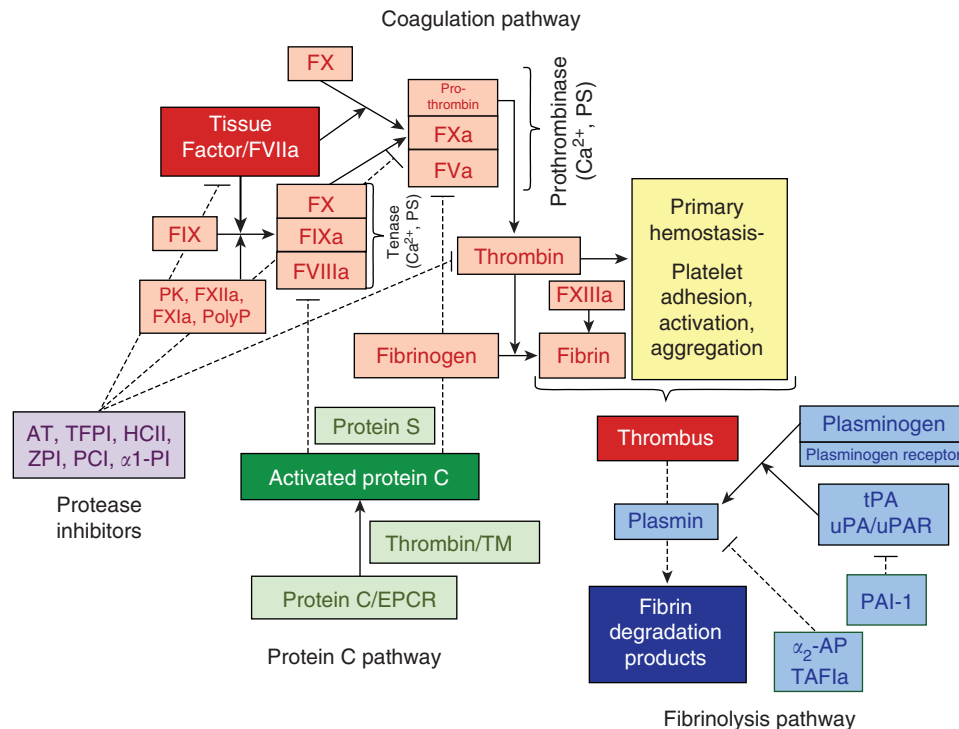


Figure 2 Blood coagulation is typically initiated following vascular injury by platelet adhesion, activation, and aggregation (primary hemostasis). The coagulation pathway (secondary hemostasis) is composed of a cascade of proteases that further activate platelets and ultimately generates thrombin to cleave fibrinogen to fibrin. Serine protease inhibitors (Serpins and Kunitz-type inhibitors) target activated coagulation proteases. The protein C pathway inhibits the further propagation of thrombin by inactivating FVa and FVIIIa. The sustained presence of a thrombus is regulated by the fibrinolysis pathway, which controls fibrin degradation. Accompanying feedback loops and other cofactors are not shown to maintain the simplicity of the schematic.

and serotonin, are released from the activated platelet and aid in the recruitment of additional platelets to the site of injury by positive feedback mechanisms (Cimmino and Golino, 2013). Once activated, glycoprotein IIb/IIIa (GP_{IIb/IIIa}) becomes exposed on the platelet surface. The exposure of GP_{IIb/IIIa} allows for the platelet to bind to circulating fibrinogen, another key coagulation protein. Fibrinogen is a bivalent molecule, and each fibrinogen molecule is capable of binding to two activated platelets. This binding effectively bridges activated platelets at the site of injury and increases the area and volume of the platelet plug. Finally, the exposure of phosphatidyl serine (PS) on the aggregated platelet surface catalyzes the conversion of fibrinogen to fibrin through the coagulation pathway, termed secondary hemostasis.

Secondary Hemostasis

Initiation Phase

Upon perturbation or injury to the vessel wall, subendothelial tissue factor (TF) expressed by cells such as fibroblasts and smooth muscle cells is exposed to the flowing blood, becoming available for complex formation with circulating activated factor VII (FVIIa) (Figure 2; Mann, 2003a; Butenas and Mann, 2002; Owens and Mackman, 2010; Williams and Mackman, 2012). This TF–FVIIa complex initiates the coagulation pathway by activation of both factor X (FX) to activated

factor X (FXa) and factor IX (FIX) to activated factor IX (FIXa). FXa bound to its cofactor, activated factor V (FVa), then cleaves prothrombin to its active form, thrombin. This is the initiation phase of coagulation (Hoffman and Monroe, 2001; Monroe and Hoffman, 2006).

The relatively small amount of thrombin generated during the initiation phase can activate factor XI, factor VIII, and factor V, and can activate platelets by cleavage of the protease-activated receptors on the platelet surface. In addition, thrombin stimulates microparticle (MP) release from platelets and circulating monocytes (Owens and Mackman, 2011). Rich in TF and PS, MPs support FXa generation and are critical in promoting thrombin generation and further platelet activation. Furthermore, inorganic polyphosphate (polyP) is secreted upon platelet activation and potentially augments several of the initiation phase components of blood coagulation (Smith *et al.*, 2006; Smith and Morrissey, 2014).

Propagation Phase

Greater activation of FIX by FXIa, and of FXa occurs on the PS-rich platelet surface by the FIXa and FVIIIa (factor VIII (activated)) complex with its substrate FX, also known as the ‘tenase complex’ (Figure 2; Table 1). Also on the platelet surface, FXa binds with its cofactor, FVa and its substrate prothrombin (all together referred to as the ‘prothrombinase complex’), generating a large volume burst of thrombin (Figure 2; Hoffman and Monroe, 2001; Monroe and Hoffman,

2006; Mann, 2003b). This is termed the propagation phase of coagulation and this process cycles to produce large amounts of thrombin that cleaves circulating fibrinogen to fibrin, which is stabilized and made insoluble through fibrin cross-linking by the transglutaminase, activated factor XIII (FXIIIa) (Aleman *et al.*, 2014).

Regulation of Blood Coagulation

As aberrant or sustained clotting can be pathological, several important anticoagulant and fibrinolytic pathways also regulate blood coagulation and clot dissolution. The vascular endothelium is an important tissue in both venous and arterial circulation, providing the barrier between the luminal blood and intima of the vessel wall. As the only noncirculating cell type in direct contact with the flowing blood, endothelial cells are pivotal in preventing thrombosis while maintaining hemostasis (Verhamme and Hoylaerts, 2006; Sagripanti and Carpi, 2000). Anticoagulant in nature, endothelial cells produce an array of proteins and substances designed to regulate thrombin production.

The protein C pathway is an essential anticoagulant regulatory pathway designed to inhibit the propagation of thrombin (Figure 2; Esmon, 2000, 2003). As part of their antithrombotic surface, vascular endothelial cells produce thrombomodulin (TM), a transmembrane glycoprotein that binds to thrombin in the presence of endothelial protein C receptor (EPCR), which binds zymogen protein C. The TM–thrombin complex cleaves protein C bound to EPCR to activated protein C (APC) (Figure 2). APC (a serine protease) can then interact with its cofactor, protein S, to proteolytically inactivate FVa and FVIIIa, down-regulating the further generation of thrombin.

Vascular endothelial cells also produce heparan sulfate and dermatan sulfate, which increase the efficiency of thrombin inhibition by antithrombin and other serine protease inhibitors (Serpins) (Figure 2; Tovar *et al.*, 2005; Rau *et al.*, 2007; Rein *et al.*, 2011). Endothelial cells synthesize tissue factor pathway inhibitor (TFPI), which is a Kunitz-type protease inhibitor that inhibits FXa and, in a FXa-dependent manner, the FVIIa–TF complex (Figure 2; Broze, 1995). Endothelial cells also produce key platelet inhibitors, including prostacyclin and nitric oxide (NO), limiting the availability of lipids for coagulation factor complex formation (Pate *et al.*, 2010).

In addition to pathways that control clot formation and prevent the generation of occlusive thrombi, fibrinolysis must also be regulated. The fibrinolytic pathway dictates whether clots are sustained or dissolved (Figure 2; Rijken and Lijnen, 2009; Fay *et al.*, 2007; Rein-Smith and Church, 2014). Endothelial cells and trapped leukocytes release tissue and urokinase plasminogen activator (tPA (tissue-type plasminogen activator) and uPA (urokinase plasminogen activator), respectively), proteases which activate plasminogen to plasmin. Plasmin is a serine protease that cleaves cross-linked fibrin to produce fibrin degradation products, which are easily swept away and degraded in flowing blood. Plasmin is inhibited by α_2 -antiplasmin (α_2 -AP) and plasminogen activator inhibitor-1 (PAI-1) is the primary inhibitor of tPA and uPA. PAI-1 and α_2 -AP are Serpins that bind tightly to the active site of these target proteases, therefore inhibiting fibrinolysis.

Coagulation Disorders Overview

There are a variety of pathologies secondary to genetic or acquired risk factors that cause dysregulation of coagulation pathways. The most common coagulation disorders that cause either bleeding or thrombotic phenotypes will be discussed.

Coagulation Disorders – Bleeding

Defects in platelet production

Decreases in platelet number, or thrombocytopenia, can be caused by defects in several systems in the body. Drugs, chemotherapy, hematological malignancy of the bone marrow, or vitamin deficiencies may inhibit the production of platelets in the bone marrow. Normal platelet counts generally range from 150 to 450×10^9 cells/ μ l. Epistaxis, mucocutaneous bleeding, bruising, petechiae, and heavy menses in females are commonly observed symptoms for patients with thrombocytopenia.

Alternatively, platelets may be produced normally, but destroyed prematurely by an immune system malfunction, as in the case of immune thrombocytopenic purpura (ITP). In ITP, antibodies against platelet surface antigens are produced, leading to immune-mediated destruction of the platelets in the spleen. ITP may be asymptomatic or may manifest with symptoms including purpura, petechiae, and mucocutaneous bleeding. Treatment for ITP includes the use of corticosteroids to inhibit autoantibody formation, and splenectomy to decrease antibody formation and platelet destruction (Cines *et al.*, 2014).

Thrombocytopenias can also be caused by defects in the processing of vWF, a specific thrombocytopenia known as thrombotic thrombocytopenic purpura (TTP). The etiology of TTP is known to be defects in ADAMTS13, a protease responsible for the cleavage of ultra-long vWF (ULVWF) secreted from the endothelium into smaller, less adhesive fragments. Deficiency, dysfunction, or antibodies against ADAMTS13 lead to the accumulation of ULVWF in the circulation. As ULVWF is a 'sticky' protein, its buildup leads to aberrant platelet adhesion and activation in the microvasculature, microvascular thrombosis, and consumption of platelets. Treatment for TTP is plasmapheresis in order to remove autoantibodies to ADAMTS13 (Kremer Hovinga and Lammle, 2012).

Defects in platelet function

Symptomatically, platelet dysfunction may resemble defects in platelet number, and may include epistaxis, mucocutaneous bleeding, easy bruising, and petechiae. Numerous qualitative platelet defects have been described, most of which affect the adhesion of, secretion from, or aggregation of platelets (Nurden and Nurden, 2014). Disorders of platelet adhesion can arise due to several factors, most notably abnormal or absent vWF (von Willebrand disease (vWD)), defects in GpIb (Bernard-Soulier Syndrome), or abnormalities of the sub-endothelial connective tissues (Ehlers–Danlos Syndrome). vWD is a particularly interesting disorder in that not only do patients have clinical symptoms consistent with a primary hemostatic defect due to vWF's role in platelet adhesion, but they also display symptoms associated with secondary

hemostatic defects due to vWF's role as a carrier protein for FVIII. Several types of hereditary and acquired vWD have been described, and their classification is based upon the nature of the defect. Type 1 vWD accounts for approximately 60–80% of all vWD and is defined as decreased levels of circulating vWF. However, many of these patients will lead a nearly normal life, requiring intervention only when undergoing surgery. When treatment is necessary, the most commonly used therapies are DDAVP (desmopressin) to stimulate the release of vWF and FVIII from endothelial cells, and the administration of cryoprecipitate to increase circulating FVIII and vWF (Lillicrap, 2013).

Hemophilia

Among the hereditary bleeding disorders, the best known are hemophilia A and B, which are X-linked recessive traits that affect the production of FVIII and FIX, respectively. Normal hemostasis requires between 25% and 33% of functional FVIII and FIX, and symptomatic patients typically have less than 5% of functional FVIII or FIX, with a close clinical correlation between severity of symptoms and factor levels (Hoyer, 1994). Patients with hemophilia typically bleed into muscles (hematomas), joints (hemarthroses), and body cavities following injury, as opposed to the mucocutaneous bleeding seen in the setting of primary hemostatic defects. Additionally, hemophilia A and B have clinically indistinguishable symptoms; therefore, direct measurements of FVIII and FIX activity should be undertaken to determine which type of hemophilia is present. Several treatment options are available for patients with hemophilia A and B including recombinant FVIII and FIX, plasma-purified products enriched in FVIII and FIX, factor concentrates, and hopefully in the near future, gene therapy (Schramm, 2014; Hoyer, 1994).

Vitamin K deficiency

Vitamin K is a fat-soluble vitamin that is absorbed from the diet in the small intestine and is stored in the liver. Vitamin K can also be synthesized by endogenous bacterial flora of the small intestine and colon. The three most recognized causes of vitamin K deficiency result from inadequate dietary intake, malabsorption in the gut, or impaired uptake and storage by the liver due to hepatic disease (Burke, 2013).

The process of sequential oxidation and reduction of vitamin K, known as the vitamin K cycle, plays a key role in the formation of functional coagulation proteases and cofactors, including prothrombin, FVII, FIX, FX, and proteins C and S. These proteins contain posttranslationally modified γ -carboxylated glutamate residues (Gla), commonly known as Gla residues. Due to their ability to bind calcium, Gla residues are vital to the proper folding of these proteins as they participate in intrachain linkages. Additionally, Gla residues are required for the binding of these coagulation factors to the surface of phospholipid membranes, namely platelets, where they can form protein complexes. The oxidation of vitamin K provides the energy necessary for the formation of Gla residues from glutamic acid residues. As such, vitamin K deficiency leads to deficiencies of prothrombin, FVII, FIX, FX, and proteins C and S (Greer, 2010). Manifestations of vitamin K deficiency may include easy bruising, mucocutaneous bleeding, and these

symptoms can be ameliorated by administration of parenteral vitamin K (McNinch, 2010).

Trauma-induced coagulopathy

Among the acquired risk factors for bleeding disorders, injury is the most common. Traumatic injuries cause 10% of deaths worldwide, with hemorrhage accounting for 40% of these fatalities (Norton and Kobusingye, 2013; Brohi *et al.*, 2007). Trauma-induced coagulopathy (TIC) is characterized by excessive and sustained bleeding, sometimes disproportionate to injury severity, affecting 25% of trauma patients. Patients suffering from TIC are at five-fold greater risk of death and require massive transfusions to replace lost blood volume (Brohi *et al.*, 2007; D'Angelo and Dutton, 2010). The mechanisms driving TIC are more complex than hemodilution or anatomic disruption. The TIC pathology is multifactorial and arises secondary to hemorrhage, tissue injury, hypoperfusion, acidosis, and hypothermia. It subsequently results in endothelial activation, platelet dysfunction, activation of the protein C system, and severe hyperfibrinolysis. Such simultaneous dysregulation of coagulation pathways results in substantial bleeding (Cardenas *et al.*, 2014). Current research in TIC is focused on methods of early and accurate diagnosis of TIC and identifying the most beneficial points of pharmacologic intervention.

Coagulation Disorders – Thrombosis

Arterial thrombosis – Overview

Coronary heart disease (CHD), including both myocardial infarction and stroke, is the leading cause of death among Americans. Approximately 720 000 Americans suffer from CHD annually and results in 380 000 deaths each year (Centers for Disease Control and Prevention, 2015). Both diseases are complications associated with arterial thrombosis, typically following atherosclerotic plaque rupture. Upon exposure of plaque contents to flowing blood, occlusive and migratory clots form very rapidly and are highly dependent on platelet activation. This results in ischemia of heart or brain tissue and is potentially fatal if not treated immediately (Massberg *et al.*, 2003). Many risk factors for arterial thrombosis are acquired and include age, diet, exercise, smoking, blood pressure, diabetes, and cholesterol levels. Further discussion of genetic and acquired risk factors is provided below.

Deep vein thrombosis – Overview

Deep vein thrombosis (DVT) is characterized by the formation of occlusive or semi-occlusive thrombi typically in the deep veins of the limbs. The most serious and potentially fatal complication of a DVT is pulmonary embolism (PE). DVT and PE are collectively referred to as venous thromboembolism (VTE). The incidence of VTE is approximately 1 in 1000, with between 350 000 and 600 000 individuals affected by this disease each year in the United States (Branchford *et al.*, 2012; Office of the Surgeon General (US); National Heart, Lung, and Blood Institute (US), 2008). Virchow's triad can be used to explain the etiology of thrombosis: changes in blood vessel due to endothelial injury, secondary to diseases like atherosclerosis, trauma, or chronic inflammation; changes in blood

constituents due to primary hypercoagulability or acquired hypercoagulability; and changes in blood flow such as caused by stasis or vessel turbulence especially behind venous valve pockets (Wolberg *et al.*, 2012). There are many risk factors for VTE, both genetic and acquired/environmental, which will be discussed in further detail with particular emphasis on age, as the population of elderly individuals in the United States is unique and rapidly growing.

Factor V Leiden

The most common genetic risk factor for venous thrombosis is the Factor V Leiden mutation. Factor V (FV) is a key component of the prothrombinase complex and regulates coagulation through its role in thrombin generation. The FV Leiden mutation results in resistance to APC-induced anticoagulation. FV Leiden results from a single point mutation in the factor V gene (guanine to adenine at nucleotide 1691), which leads to a single amino acid change (FV Arg506Gln). The FV Leiden mutation not only slows the inactivation of FVa by APC, but also alters a cofactor generated during APC-mediated cleavage of unactivated FV. Thus, resistance to APC inactivation results in sustained and uninhibited thrombin generation by these proteins. Approximately 3–7% of individuals in the United States carry the FV Leiden mutation and it predominantly affects Caucasians (Shaheen *et al.*, 2012). Heterozygotes are at 2–7 times greater risk of VTE compared to individuals without the FV Leiden mutation, while homozygotes are at a striking 40- to 80-fold higher risk of VTE (De Stefano *et al.*, 2002).

Prothrombin G20210A

The second most common polymorphism predisposing patients to VTE is the prothrombin G20210A mutation. As the precursor to thrombin, prothrombin is a procoagulant zymogen essential for effective coagulation. Similar to Factor V Leiden, the prothrombin G20210A mutation predominantly affects Caucasians. Approximately 2–3% of the Caucasian population carries the prothrombin mutation, which is a single substitution of adenine for guanine at the G20210 position resulting in reduced prothrombin degradation and a subsequent increase in circulating plasma levels of prothrombin. This mutation confers a three-fold greater risk of developing VTE (Bauduer and Lacombe, 2005; Johnson *et al.*, 2005).

Plasminogen activator inhibitor-1 (PAI-1)

PAI-1 blood levels are sensitive to many different pathophysiological factors; thus, increased synthesis of PAI-1 contributes to several cardiovascular disease states (Yamamoto *et al.*, 2005; Booth, 1999; Gramling and Church, 2010; Vaughan, 2001, 2005). As an acute phase reactant, PAI-1 blood levels increase in response to vessel injury and inflammatory signaling. Increased PAI-1 levels are associated with coronary artery disease and myocardial infarction (Tjarnlund-Wolf *et al.*, 2012). However, studies examining the association of thrombosis with a polymorphism in the PAI-1 promoter region (4G/5G) that increases PAI-1 expression remain controversial (Ridker *et al.*, 1997; Segui *et al.*, 2000). Use of statins, a class of drugs used to lower cholesterol levels, is associated with a decrease in PAI-1 expression and increase in tPA expression. Use of drug-eluting stents (coated with rapamycin or

paclitaxel) placed in an artery during percutaneous coronary intervention is associated with an increased risk of arterial thrombosis due to an up-regulation of PAI-1 (Muldowney *et al.*, 2007).

Aging

While there are a variety of acquired risk factors for thrombosis including obesity, diabetes, hypertension, injury, pregnancy, and hormone therapy, the strongest and most unchangeable risk factor for both arterial and venous thrombosis is age. Population-based and basic science research has uncovered a variety of age-related changes in blood proteins, cells, and the vasculature that could predispose individuals to thrombosis.

A number of large-scale studies have measured the changes in coagulation factors with respect to age. One of the coagulation proteins most commonly reported to increase with age is fibrinogen, with levels ranging from 250 mg/dL in young, 20-year-old subjects to 300 mg/dL in subjects over 65 years of age (Mari *et al.*, 2008; Meade *et al.*, 1977). As the precursor to fibrin, fibrinogen availability is a key determinant in the formation of stable thrombi.

FVIII levels are also known to increase with age. FVIII, with its cofactor FIXa, is a potent activator of FX. Thus, elevated FVIII has long been believed to be a risk factor for VTE, and as of recently, also for arterial thrombosis (Machlus *et al.*, 2011). A prospective cohort study showed an 11-fold increase in the risk of VTE in individuals with FVIII levels above the 90th percentile (Kytle *et al.*, 2000).

FVIII circulates bound to vWF and as such, age-related increases in vWF have also been observed (Toffler *et al.*, 2005). It was thought that vWF only contributed to the risk of VTE through its association with, or effect on levels of FVIII. However, the 2002 LITE study showed that FVIII and vWF are in fact independent risk factors for VTE (Tsai *et al.*, 2002). Mouse models have recently shown that both vWF and FVIII are independently critical for venous thrombus formation.

The zymogen FIX, a member of the tenase complex with cofactor FVIII, also undergoes an age-dependent increase (Sweeney and Hoernig, 1993). Given its importance in the activation of FX, elevated FIX levels are thought to be a risk factor for VTE. Data from the LETS study concluded that those with FIX levels above the 90th percentile had an odds ratio of 2.3 and were at elevated risk of VTE (van Hylckama Vlieg *et al.*, 2000).

PAI-1 is an important inhibitor of fibrinolysis and elevated levels of this protein have been reported in aging and a variety of other disease states. Data from the Framingham cohort found that PAI-1 antigen levels in plasma went from 19.4 ng/mL in men under 40 years to 24.6 ng/mL in those 70 years and above, with similar increases found in women (Toffler *et al.*, 2005). Elevated PAI-1 levels are positively correlated with the risk of thrombotic events (Prisco *et al.*, 1993; Hamsten *et al.*, 1987).

Interestingly, aging in humans is associated with a decrease in circulating platelet counts, with an estimated decrease of $6 \times 10^9/L$ for every 10 years of age (Biino *et al.*, 2012). It remains unclear whether this decrease in platelet count is due to reduced hematopoiesis, decreased production from megakaryocytes, or increased clearance by the spleen.

Despite the reduction in platelet number, aging is associated with differences in platelet function. While there is very little current data in this area, historical data has given insight into how platelet biology changes with age. Increased platelet aggregation in response to ADP, collagen, and adrenaline has been observed in platelets from aged individuals, in addition to increased ATP release upon collagen stimulation (Kasjanovova and Balaz, 1986; Vilen *et al.*, 1989). Platelets from older individuals also display reduced ADP thresholds for aggregation compared to those from the young, further suggesting they are more sensitive to certain agonists (Gleerup and Winther, 1995). It is thought that this hyperreactivity is due to increased platelet expression of receptors for these agonists (Kasjanovova and Balaz, 1986; Vilen *et al.*, 1989). Platelets from the elderly also produce more thromboxane A₂, which is involved in platelet activation and formation of platelet aggregates (Suehiro *et al.*, 1995; Gleerup and Winther, 1995). This could have clinical implications for treating elderly patients with antiplatelet agents such as clopidogrel, an inhibitor of the platelet P2Y₁₂ receptor for ADP. A recent study found that the linear age-related increase in ADP-induced platelet aggregation was associated with reduced sensitivity to initial clopidogrel treatment (Gremmel *et al.*, 2010).

Aside from biological changes that occur with age in platelets, age-associated changes in endothelial function could also promote platelet hyperreactivity. These include increased endothelial production of vWF, which could result in enhanced platelet tethering to endothelial cell surfaces (Muller *et al.*, 2002). Also, decreased production of endothelial nitric oxide synthase (eNOS) results in down-regulation of NO generation. NO is a potent inhibitor of platelet activation (Yoon *et al.*, 2010). Decreased levels of this important inhibitor could result in dysregulation of platelet activation and the inappropriate release of procoagulant material from platelet granules.

Endothelial dysfunction is a widely accepted consequence of aging and can turn the naturally anticoagulant endothelial cell surface into one that is more procoagulant. Many age-related changes have been reported to occur in the vascular endothelium, one of which is increased vascular permeability. The loss of cell-cell junctions with age results in vascular leakage, exposure of subendothelial proteins (i.e., TF), and leukocyte adhesion (Lum and Roebuck, 2001). Aged endothelial cells also produce more anti-fibrinolytic PAI-1 (Comi *et al.*, 1995) and, in animal models, less anticoagulant TM upon treatment with endotoxin (Starr *et al.*, 2010). There is also increased production and secretion of pro-inflammatory cytokines interleukin-6 (IL-6) and interleukin-1 β (IL-1 β) from aged endothelium, promoting recruitment and binding of leukocytes to the cell surface (Csiszar *et al.*, 2003). Aging is also associated with a loss of eNOS production, resulting in reduced availability of NO to regulate vessel dilation and platelet activation (Yoon *et al.*, 2010).

Conclusions

The hemostatic system is a highly regulated protease pathway involving intracellular and extracellular proteins, circulating

blood cells, and the dynamic vascular endothelium (Figures 1 and 2). Efficient and appropriate hemostasis is crucial to maintaining blood flow and perturbations in this system can cause life-threatening bleeding or thrombotic complications.

See also: Protein Synthesis/Degradation: Protein Degradation – Protease Classes: Kallikrein

References

- Aleman, M.M., Byrnes, J.R., Wang, J.G., *et al.*, 2014. Factor XIII activity mediates red blood cell retention in venous thrombi. *Journal of Clinical Investigation* 124, 3590–3600.
- Andrews, R.K., Berndt, M.C., 2008. Platelet adhesion: A game of catch and release. *Journal of Clinical Investigation* 118, 3009–3011.
- Bauduer, F., Lacombe, D., 2005. Factor V Leiden, prothrombin 20210A, methylenetetrahydrofolate reductase 677T, and population genetics. *Molecular Genetics and Metabolism* 86, 91–99.
- Biino, G., Gasparini, P., D'Adamo, P., *et al.*, 2012. Influence of age, sex and ethnicity on platelet count in five Italian geographic isolates: Mild thrombocytopenia may be physiological. *British Journal of Haematology* 157, 384–387.
- Blair, P., Flaumenhaft, R., 2009. Platelet alpha-granules: Basic biology and clinical correlates. *Blood Reviews* 23, 177–189.
- Booth, N.A., 1999. Fibrinolysis and thrombolysis. *Baillieres Best Practice & Research Clinical Haematology* 12, 423–433.
- Branchford, B.R., Mourani, P., Bajaj, L., *et al.*, 2012. Risk factors for in-hospital venous thromboembolism in children: A case-control study employing diagnostic validation. *Haematologica* 97, 509–515.
- Brohi, K., Cohen, M.J., Davenport, R.A., 2007. Acute coagulopathy of trauma: Mechanism, identification and effect. *Current Opinion in Critical Care* 13, 680–685.
- Broze Jr., G.J., 1995. Tissue factor pathway inhibitor and the revised theory of coagulation. *Annual Review of Medicine* 46, 103–112.
- Burke, C.W., 2013. Vitamin K deficiency bleeding: Overview and considerations. *Journal of Pediatric Health Care* 27, 215–221.
- Butenas, S., Mann, K.G., 2002. Blood coagulation. *Biochemistry (Moscow)* 67, 3–12.
- Cardenas, J.C., Wade, C.E., Holcomb, J.B., 2014. Mechanisms of trauma-induced coagulopathy. *Current Opinion in Hematology* 21, 404–409.
- Centers for Disease Control and Prevention, 2015. Deaths: Final data for 2013. *National Vital Statistics Report*, vol. 64, no. 2. Atlanta, GA: Centers for Disease Control and Prevention.
- Cimmino, G., Golino, P., 2013. Platelet biology and receptor pathways. *Journal of Cardiovascular Translational Research* 6, 299–309.
- Cines, D.B., Cuker, A., Semple, J.W., 2014. Pathogenesis of immune thrombocytopenia. *Presse Medicale* 43, e49–e59.
- Comi, P., Chiaramonte, R., Maier, J.A., 1995. Senescence-dependent regulation of type 1 plasminogen activator inhibitor in human vascular endothelial cells. *Experimental Cell Research* 219, 304–308.
- Csiszar, A., Ungvari, Z., Koller, A., Edwards, J.G., Kaley, G., 2003. Aging-induced proinflammatory shift in cytokine expression profile in coronary arteries. *FASEB Journal* 17, 1183–1185.
- D'Angelo, M.R., Dutton, R.P., 2010. Management of trauma-induced coagulopathy: Trends and practices. *AANA Journal* 78, 35–40.
- De Stefano, V., Rossi, E., Paciaroni, K., Leone, G., 2002. Screening for inherited thrombophilia: Indications and therapeutic implications. *Haematologica* 87, 1095–1108.
- Esmon, C.T., 2000. Regulation of blood coagulation. *Biochimica et Biophysica Acta* 1477, 349–360.
- Esmon, C.T., 2003. The protein C pathway. *Chest* 124, 26S–32S.
- Fay, W.P., Garg, N., Sunkar, M., 2007. Vascular functions of the plasminogen activation system. *Arteriosclerosis, Thrombosis, and Vascular Biology* 27, 1231–1237.
- Firbas, C., Siller-Matula, J.M., Jilma, B., 2010. Targeting von Willebrand factor and platelet glycoprotein Ib receptor. *Expert Review of Cardiovascular Therapy* 8, 1689–1701.

- Gleerup, G., Winther, K., 1995. The effect of ageing on platelet function and fibrinolytic activity. *Angiology* 46, 715–718.
- Gramling, M.W., Church, F.C., 2010. Plasminogen activator inhibitor-1 is an aggregate response factor with pleiotropic effects on cell signaling in vascular disease and the tumor microenvironment. *Thrombosis Research* 125, 377–381.
- Greer, F.R., 2010. Vitamin K the basics — What's new? *Early Human Development* 86 (Suppl. 1), 43–47.
- Gremmel, T., Steiner, S., Seidinger, D., *et al.*, 2010. Adenosine diphosphate-inducible platelet reactivity shows a pronounced age dependency in the initial phase of antiplatelet therapy with clopidogrel. *Journal of Thrombosis and Haemostasis* 8, 37–42.
- Hamsten, A., de Faire, U., Walldius, G., *et al.*, 1987. Plasminogen activator inhibitor in plasma: Risk factor for recurrent myocardial infarction. *Lancet* 2, 3–9.
- Hoffman, M., Monroe III, D.M., 2001. A cell-based model of hemostasis. *Thrombosis and Haemostasis* 85, 958–965.
- Hoyer, L.W., 1994. Hemophilia A. *New England Journal of Medicine* 330, 38–47.
- van Hylckama Vlieg, A., van der Linden, I.K., Bertina, R.M., Rosendaal, F.R., 2000. High levels of factor IX increase the risk of venous thrombosis. *Blood* 95, 3678–3682.
- Johnson, C.M., Mureebe, L., Silver, D., 2005. Hypercoagulable states: A review. *Vascular and Endovascular Surgery* 39, 123–133.
- Kasjanovova, D., Balaz, V., 1986. Age-related changes in human platelet function in vitro. *Mechanisms of Ageing and Development* 37, 175–182.
- Kremer Hovinga, J.A., Lammle, B., 2012. Role of ADAMTS13 in the pathogenesis, diagnosis, and treatment of thrombotic thrombocytopenic purpura. *American Society of Hematology. American Society of Hematology. Education Program* 2012, 610–616.
- Kyle, P.A., Minar, E., Hirsch, M., *et al.*, 2000. High plasma levels of factor VIII and the risk of recurrent venous thromboembolism. *New England Journal of Medicine* 343, 457–462.
- Lenting, P.J., Pegon, J.N., Groot, E., De Groot, P.G., 2010. Regulation of von Willebrand factor-platelet interactions. *Thrombosis and Haemostasis* 104, 449–455.
- Lillicrap, D., 2013. von Willebrand disease: Advances in pathogenetic understanding, diagnosis, and therapy. *American Society of Hematology. Education Program* 2013, 254–260.
- Lum, H., Roebuck, K.A., 2001. Oxidant stress and endothelial cell dysfunction. *American Journal of Physiology — Cell Physiology* 280, C719–C741.
- Machlus, K.R., Cardenas, J.C., Church, F.C., Wolberg, A.S., 2011. Causal relationship between hyperfibrinogenemia, thrombosis, and resistance to thrombolysis in mice. *Blood* 117, 4953–4963.
- Machlus, K.R., Thon, J.N., Italiano Jr., J.E., 2014. Interpreting the developmental dance of the megakaryocyte: A review of the cellular and molecular processes mediating platelet formation. *British Journal of Haematology* 165, 227–236.
- Mann, K.G., 2003a. Thrombin formation. *Chest* 124, 4S–10S.
- Mann, K.G., 2003b. Thrombin: Can't live without it; probably die from it. *Chest* 124, 1S–3S.
- Mari, D., Coppola, R., Provenzano, R., 2008. Hemostasis factors and aging. *Experimental Gerontology* 43, 66–73.
- Massberg, S., Schulz, C., Gawaz, M., 2003. Role of platelets in the pathophysiology of acute coronary syndrome. *Seminars in Vascular Medicine* 3, 147–162.
- McNicol, A., Israels, S.J., 1999. Platelet dense granules: Structure, function and implications for haemostasis. *Thrombosis Research* 95, 1–18.
- McNinch, A., 2010. Vitamin K deficiency bleeding: Early history and recent trends in the United Kingdom. *Early Human Development* 86 (Suppl. 1), 63–65.
- Meade, T.W., North, W.R., Chakrabarti, R., Haines, A.P., Stirling, Y., 1977. Population-based distributions of haemostatic variables. *British Medical Bulletin* 33, 283–288.
- Monroe, D.M., Hoffman, M., 2006. What does it take to make the perfect clot? *Arteriosclerosis, Thrombosis, and Vascular Biology* 26, 41–48.
- Muldowney III, J.A., Stringham, J.R., Levy, S.E., *et al.*, 2007. Antiproliferative agents alter vascular plasminogen activator inhibitor-1 expression: A potential prothrombotic mechanism of drug-eluting stents. *Arteriosclerosis, Thrombosis, and Vascular Biology* 27, 400–406.
- Muller, A.M., Skrzynski, C., Nessler, M., Skipka, G., Muller, K.M., 2002. Correlation of age with in vivo expression of endothelial markers. *Experimental Gerontology* 37, 713–719.
- Norton, R., Kobusingye, O., 2013. Injuries. *New England Journal of Medicine* 368, 1723–1730.
- Nurden, A.T., Nurden, P., 2014. Congenital platelet disorders and understanding of platelet function. *British Journal of Haematology* 165, 165–178.
- Office of the Surgeon General (US); National Heart, Lung, and Blood Institute (US), 2008. The Surgeon General's Call to Action to Prevent Deep Vein Thrombosis and Pulmonary Embolism.
- Owens III, A.P., Mackman, N., 2010. Tissue factor and thrombosis: The clot starts here. *Thrombosis and Haemostasis* 104, 432–439.
- Owens III, A.P., Mackman, N., 2011. Microparticles in hemostasis and thrombosis. *Circulation Research* 108, 1284–1297.
- Pate, M., Damarla, V., Chi, D.S., Negi, S., Krishnaswamy, G., 2010. Endothelial cell biology: Role in the inflammatory response. *Advances in Clinical Chemistry* 52, 109–130.
- Prisco, D., Chiarantini, E., Boddi, M., *et al.*, 1993. Predictive value for thrombotic disease of plasminogen activator inhibitor-1 plasma levels. *International Journal of Clinical and Laboratory Research* 23, 78–82.
- Rau, J.C., Beaulieu, L.M., Huntington, J.A., Church, F.C., 2007. Serpins in thrombosis, hemostasis and fibrinolysis. *Journal of Thrombosis and Haemostasis* 5 (Suppl. 1), 102–115.
- Rein, C.M., Desai, U.R., Church, F.C., 2011. Serpin-glycosaminoglycan interactions. *Methods in Enzymology* 501, 105–137.
- Rein-Spang, C.M., Church, F.C., 2014. Emerging pathophysiological roles for fibrinolysis. *Current Opinion in Hematology* 21, 438–444.
- Ricker, P.M., Hennekens, C.H., Lindpaintner, K., Stampfer, M.J., Miletich, J.P., 1997. Arterial and venous thrombosis is not associated with the 4G/5G polymorphism in the promoter of the plasminogen activator inhibitor gene in a large cohort of US men. *Circulation* 95, 59–62.
- Rijken, D.C., Lijnen, H.R., 2009. New insights into the molecular mechanisms of the fibrinolytic system. *Journal of Thrombosis and Haemostasis* 7, 4–13.
- Sagripanti, A., Carpi, A., 2000. Antithrombotic and prothrombotic activities of the vascular endothelium. *Biomedicine & Pharmacotherapy* 54, 107–111.
- Schramm, W., 2014. The history of haemophilia — A short review. *Thrombosis Research* 134, S4–S9.
- Segui, R., Estelles, A., Mira, Y., *et al.*, 2000. PAI-1 promoter 4G/5G genotype as an additional risk factor for venous thrombosis in subjects with genetic thrombophilic defects. *British Journal of Haematology* 111, 122–128.
- Shaheen, K., Alraies, M.C., Alraies, A.H., Christie, R., 2012. Factor V Leiden: How great is the risk of venous thromboembolism? *Cleveland Clinic Journal of Medicine* 79, 265–272.
- Smith, S.A., Morrissey, J.H., 2014. Polyphosphate: A new player in the field of hemostasis. *Current Opinion in Hematology* 21, 388–394.
- Smith, S.A., Mutch, N.J., Baskar, D., *et al.*, 2006. Polyphosphate modulates blood coagulation and fibrinolysis. *Proceedings of the National Academy of Sciences of the United States of America* 103, 903–908.
- Starr, M.E., Ueda, J., Takahashi, H., *et al.*, 2010. Age-dependent vulnerability to endotoxemia is associated with reduction of anticoagulant factors activated protein C and thrombomodulin. *Blood* 115, 4886–4893.
- Suehiro, A., Uedaa, M., Suehiro, M., Ohe, Y., Kakishita, E., 1995. Evaluation of platelet hyperfunction in aged subjects using spontaneous platelet aggregation in whole blood. *Archives of Gerontology and Geriatrics* 21, 277–283.
- Sweeney, J.D., Hoernig, L.A., 1993. Age-dependent effect on the level of factor IX. *American Journal of Clinical Pathology* 99, 687–688.
- Thon, J.N., Italiano, J.E., 2012. Platelets: Production, morphology and ultrastructure. *Handbook of Experimental Pharmacology* 210, 3–22.
- Tjarnlund-Wolf, A., Brogren, H., Lo, E.H., Wang, X., 2012. Plasminogen activator inhibitor-1 and thrombotic cerebrovascular diseases. *Stroke* 43, 2833–2839.
- Toffer, G.H., Massaro, J., Levy, D., *et al.*, 2005. Relation of the prothrombotic state to increasing age (from the Framingham offspring study). *American Journal of Cardiology* 96, 1280–1283.
- Tovar, A.M., de Mattos, D.A., Stelling, M.P., *et al.*, 2005. Dermatan sulfate is the predominant antithrombotic glycosaminoglycan in vessel walls: Implications for a possible physiological function of heparin cofactor II. *Biochimica et Biophysica Acta* 1740, 45–53.
- Tsai, A.W., Cushman, M., Rosamond, W.D., *et al.*, 2002. Coagulation factors, inflammation markers, and venous thromboembolism: The longitudinal investigation of thromboembolism etiology (LITE). *American Journal of Medicine* 113, 636–642.
- Vainchenker, W., Besancenot, R., Favale, F., 2014. Megakaryopoiesis: Regulation of platelet production by thrombopoietin. *Bulletin de l'Académie Nationale de Médecine* 197, 395–406.
- Varga-Szabo, D., Pleines, I., Nieswandt, B., 2008. Cell adhesion mechanisms in platelets. *Arteriosclerosis, Thrombosis, and Vascular Biology* 28, 403–412.
- Vaughan, D.E., 2001. Angiotensin, fibrinolysis, and vascular homeostasis. *American Journal of Cardiology* 87, 18C–24C.
- Vaughan, D.E., 2005. PAI-1 and atherothrombosis. *Journal of Thrombosis and Haemostasis* 3, 1879–1883.
- Verhamme, P., Hoylaerts, M.F., 2006. The pivotal role of the endothelium in hemostasis and thrombosis. *Acta Clinica Belgica* 61, 213–219.

- Vilen, L., Jacobsson, S., Wadenvik, H., Kutti, J., 1989. ADP-induced platelet aggregation as a function of age in healthy humans. *Thrombosis and Haemostasis* 61, 490–492.
- Williams, J.C., Mackman, N., 2012. Tissue factor in health and disease. *Frontiers in Bioscience (Elite Edition)* 4, 358–372.
- Wolberg, A.S., Aleman, M.M., Leiderman, K., Machlus, K.R., 2012. Procoagulant activity in hemostasis and thrombosis: Virchow's triad revisited. *Anesthesia & Analgesia* 114, 275–285.
- Yamamoto, K., Takeshita, K., Kojima, T., Takamatsu, J., Saito, H., 2005. Aging and plasminogen activator inhibitor-1 (PAI-1) regulation: Implication in the pathogenesis of thrombotic disorders in the elderly. *Cardiovascular Research* 66, 276–285.
- Yoon, H.J., Cho, S.W., Ahn, B.W., Yang, S.Y., 2010. Alterations in the activity and expression of endothelial NO synthase in aged human endothelial cells. *Mechanisms of Ageing and Development* 131, 119–123.

Molecular Mechanisms Underlying the Actions of the Complement System

RN Pike and LC Wijeyewickrema, Monash University, Clayton, VIC, Australia

© 2016 Elsevier Inc. All rights reserved.

The Complement System

The complement system is a set of 35 defensive proteins that protect our bodies against pathogens and altered self cells (Carroll, 2004). The vital role that the system plays is revealed by the many diseases that we become susceptible to when deficient in one of the proteins from the system (Ricklin *et al.*, 2010). The most evident of these is that humans become highly prone to infection when the complement system is compromised (Degn *et al.*, 2011). Equally, complement plays a major role in causing disease, as the system is involved in most diseases with an inflammatory component to them (Mollnes *et al.*, 2002), giving rise to a broad range of diseases affecting systems ranging from joints to teeth and the cardiovascular system. This has led to a range of complement molecules being targeted in order to control diseases such as atypical hemolytic urinary syndrome, reperfusion injury after acute myocardial infarction, arthritis, and graft versus host disease.

The complement system is initiated by the classical, lectin, and alternative pathways (Duncan *et al.*, 2008; Figure 1). Each of these pathways involves an initial recognition or binding event, in which a pathogen or altered self cell is 'tagged' by a molecule of the complement system. Subsequent to the initial binding event, protease molecules that are either pre-bound to the molecule that mediated the initial binding event or become bound to it, cleave substrates that amplify the initial binding signal (Gros *et al.*, 2008). These amplification events have both local and systemic effects. In local terms, the amplified system is able to induce death to pathogen or self cells by lysing their membranes using the C5–C9 components of the system to perforate the membranes by assembling these components into pore-like structures, causing death of the recognized cell. These cells are also decorated with other molecules assembled onto their surfaces along the way, particularly the cleaved C3b molecule, which flags the cell to receptors on professional phagocytic white blood cells and thus ensures its internalization into these cells and subsequent digestion. This not only removes the cell from the site, but also begins the antigen presentation process by these white blood cells, which provides an extremely important link to adaptive immunity (Dunkelberger and Song, 2010). In systemic terms, chemotactic molecules, particularly C5a, are liberated from tissue in which complement has been activated and taken up by the bloodstream, thus providing a chemotactic gradient to the site of infection or injury. White blood cells are thus alerted to the occurrence and migrate down the chemotactic gradient to the site to carry out functions ranging from phagocytosis (macrophages) to pathogen killing (neutrophils) and induction of an adaptive immune response (dendritic cells and T-cells). Thus the activated complement system potentially brings the entire force of the host's immune system to bear on pathogenic or altered self cells. Unfortunately, this same potent response can become a misguided missile, causing injury to the host and thus disease.

Classical Pathway

The classical pathway of complement (Figure 2) is activated by the C1 complex (Wallis *et al.*, 2010), comprised of six C1q molecules in association with the C1r and C1s serine proteases (Formeris *et al.*, 2012). Binding of C1q molecules to their ligands, classically antibody–antigen complexes, causes auto-activation of C1r, which in turn cleaves C1s to activate it. The activated C1s cleaves the C4 and C2 complement components in turn, allowing C4b and C2a to assemble into the C3 convertase complex. The proteolytic C2a component then cleaves C3 attracted to the convertase complex by a binding site on the C4b component. The cleaved C3b component remains bound to the complex to form the C5 convertase complex, in which C2a cleaves C5 recruited to the complex via binding sites on C4bC3b. The resultant C5b product is able to scaffold the assembly of the C6–C9 components into the MAC as noted above. It is worth noting that the C3 and C4 complement components belong to the group of molecules containing thioester groups (Law and Dodds, 1997), which are activated upon their cleavage and facilitate binding of these molecules to membrane proteins. Interestingly, while C5 belongs to this group of molecules as judged by sequence homology, it has lost the thioester group pivotal to the functions of C3 and C4 (Discipio, 1981). This binding mechanism of C3 and C4 results in their localization to the point of activation of the system, a crucial mechanism to prevent mislocalisation of the activated system. The proteolytic cleavage of the C3–C5 components also gives rise to the C3a, C4a, and C5a peptide components which play a key role in activating the rest of the immune system as noted above.

The C1 complex that initiates the classical pathway is composed of six C1q molecules, each in turn composed of three chains, associated with two C1r and two C1s protease molecules (Gaboriaud *et al.*, 2004). Electron microscopy and structural studies of the C1 complex and components have given some insights into the presumed structure of C1, with these studies revealing a structure somewhat reminiscent of an inverted bunch of tulips, consisting essentially of three regions: the neck, stalks, and heads. From the N-terminus, the C1q molecules are composed of collagen-like triple helices wound together into a super-helix in the neck and stalk regions and three globular heads. The neck regions of all six molecules are bound together, through a combination of covalent disulfide bonds and non-covalent bonds. The super-helices then splay away from each other in the stalk region to yield 6 individual stalks. Finally, the globular heads, responsible for binding to ligands, are found at the C-terminal end of the structure (Gaboriaud *et al.*, 2003). Recent new data indicates that full activation of this hexameric arrangement of C1q molecules is achieved when the C1 complex binds to the constant fragment 'tails' of six immunoglobulin molecules hexamerically arranged due to non-covalent bonding that occurs between their Fc segments after binding to their antigenic targets (Diebolder *et al.*, 2014).

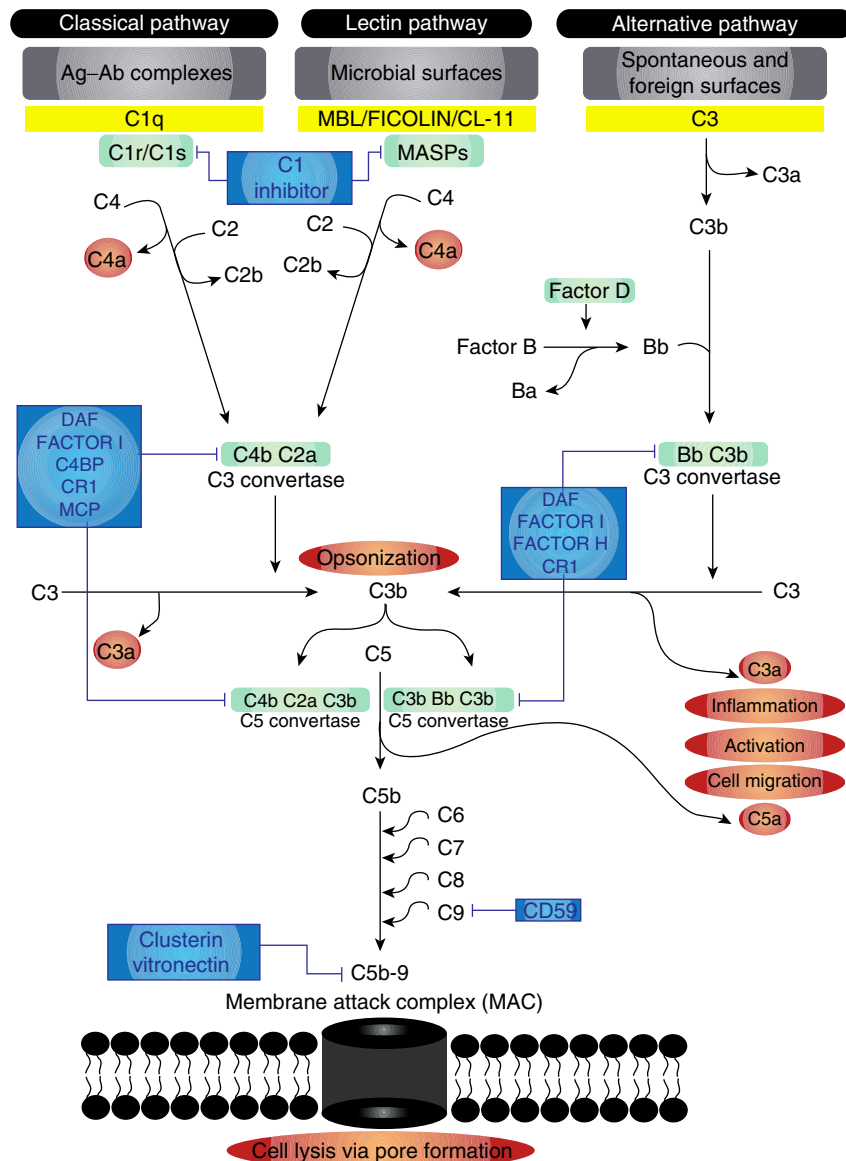


Figure 1 The three major complement activation pathways: classical, lectin, and alternative. The classical and lectin pathways are activated when recognition complexes bind to cognate targets, such as antibody–antigen complexes (C1 complex) or carbohydrate arrays on microbial surfaces (lectins). They use associated serine proteases, such as C1s and MASP-2, to cleave C2 and C4 to generate the C3 convertase (C4bC2a) and which in turn cleaves C3 to yield the C5 convertase (C4bC2aC3b). The alternative complement pathway begins with the spontaneous activation of C3 and requires fB and factor D to form the C3 convertase (C3bBb) and then the C5 convertase (C3bBbC3b). All three pathways converge at C3, which is converted into C3a and C3b. The further formed C5 convertase cleaves C5 into C5a and C5b. C5b scaffolds the assembly of the membrane-attack complex (MAC), composed of C5b, C6, C7, C8, and many copies of C9. The activation of the system thus results in the killing of the pathogen via the MAC, opsonization of the organism for phagocytosis and the released C3a and C5a result in activation of white blood cells, including those controlling the adaptive immune response. The complement system is controlled by many inhibitors (blue boxes).

The protease component of the C1 complex is composed of two molecules each of C1r and C1s, formed into a heterotetramer. C1r and C1s are both serine proteases that are composed of six domains arranged into two modules: (1) the non-catalytic module, composed of the CUB1-EGF-CUB2 domains, and (2) the catalytic module made up of the complement control protein-like (CCP1-CCP2-serine protease (SP)) domains. The non-catalytic module of each enzyme binds calcium ions and is vital for the binding of the enzymes

to each other and to C1q (Venkatraman Girija *et al.*, 2013). Recent data suggests that the CUB1-EGF-CUB2 modules of the two C1s molecules form the core of the molecular arrangement of the proteases in the inactive C1 complex, with the same regions of the two C1r molecules then arranged on the outside of this C1s core (Brier *et al.*, 2010; Bally *et al.*, 2009; Phillips *et al.*, 2009). It seems likely that the CUB1 domains of the C1r and C1s molecules are responsible for binding to four of the ‘stalks’ of the C1q molecule, while the remaining two

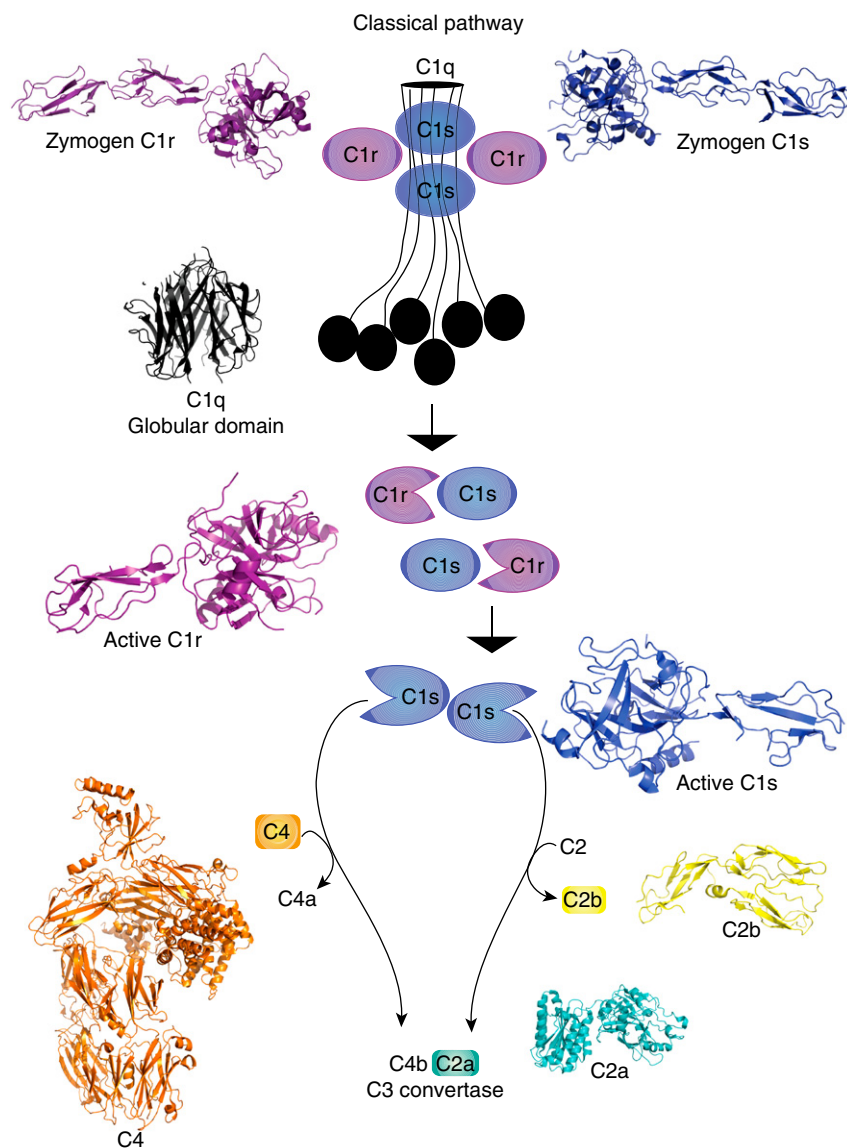


Figure 2 A schematic representation of the classical complement pathway. The C1 complex is composed of C1q (structure of globular head domain shown (PDB identifier:1PK6)) and two molecules each of zymogen C1r (1MD7) and C1s (4JIY). The classical pathway is triggered by binding of C1q to antibody–antigen complexes. C1r autoactivates (1MD8) and then cleaves and activates C1s (1ELV), which in turn cleaves C4 (4FXG) into C4a and C4b and then C2 into C2a (2I6S) and C2b (3ERB) fragments. C4b binds covalently to the target surface and C2a coordinates with it to form an enzymatic complex termed the C3 convertase (C4bC2a).

stalks bind at the interface of the CUB1-EGF-CUB2 domains of C1r (Gingras *et al.*, 2011; Forneris *et al.*, 2012). The resting C1 complexes contain zymogens or inactive forms of the proteases, while binding of the globular heads to ligands is thought to induce further splaying of the stalks, thus inducing a ‘mechanical’ change to the position of the proteases interwoven with the stalks and so allowing C1r molecules to catalyze their autoactivation (Brier *et al.*, 2010). Activated C1r in turn catalyzes the activation of C1s.

The structures of the zymogen and activated forms of C1r (Budayova-Spano *et al.*, 2002a,b; Kardos *et al.*, 2001) and C1s (Gaboriaud *et al.*, 2000; Perry *et al.*, 2013) reveal that the C1s zymogen is particularly well regulated prior to activation, with

multiple mechanisms used to occlude the active site of the protease and thus maintain it in the inactive zymogenic state (Perry *et al.*, 2013). Recent studies have shown that the C1r active site is highly specific for the P2 Gln residue of substrates in particular (Wijeyewickrema *et al.*, 2013) and that the CCP domains of the two enzymes are important for their recognition during autoactivation and activation of C1s by C1r (Kardos *et al.*, 2008).

Once activated, C1s is responsible for the cleavage of C4 and C2. Only the catalytic module (CCP1-CCP2-SP) is required for the cleavage of its substrates (Rossi *et al.*, 1998). A positively charged exosite on the SP domain of C1s interacts with a highly negatively charged region of C4, centered upon

three sulfotyrosine residues, to effect efficient cleavage of C4 (Duncan *et al.*, 2012). The exosite on the SP of C1s is not fully formed in the zymogen form of the molecule and only switches all four positively charged residues of the region into the correct positions for binding to C4 upon activation by C1r (Perry *et al.*, 2013). Recently, the structure of MASP-2 in complex with C4 has been solved (Kidmose *et al.*, 2012) and a similar SP exosite interacting with sulfotyrosine residues on C4 has been identified. Additionally, the location of a CCP exosite at the junction of the CCP1 and CCP2 was elucidated for MASP-2. By analogy, C1s has similar residues at this position in its CCP domains (Perry *et al.*, 2013) and therefore it is likely that this CCP exosite will also be required for efficient cleavage of C4 by C1s.

C1-inhibitor or serpin G1 has been shown to be the endogenous inhibitor of both C1r and C1s (Caliezi *et al.*, 2000). The inhibitor interacts with C1r with a k_{ass} value of $1 \times 10^3 \text{ M}^{-1} \text{ s}^{-1}$ (Sim *et al.*, 1980) and with C1s at a slightly higher rate of $4 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ (Wuillemin *et al.*, 1997). The rate of interaction between C1r and C1-inhibitor is enhanced about 2.5-fold by heparin (Sim *et al.*, 1980), while that between C1s and C1-inhibitor is enhanced 60-fold by heparin and 130-fold by polyanions, such as dextran sulfate (Wuillemin *et al.*, 1997). The rate enhancement brought about by heparin in this context is most likely physiologically relevant, although the greater enhancement induced by non-physiological polyanions, such as dextran sulfate, suggests that other physiologically relevant cofactors that are able to bring about a greater effect are yet to be identified.

The C2 component of complement is produced as a five domain protein with a number of glycosylation sites included in its primary structure. From the N-terminus, the protein has three CCP domains, a von Willebrand factor A-type (VWA) domain and serine protease domain (Forneris *et al.*, 2012). After complexing with the C4b produced by C1s or MASP-2 cleavage, C2 is cleaved by the same enzymes to produce the C2a and C2b fragments. C2a remains bound to C4b and forms the proteolytic component of the C3 and C5 convertase complexes, while the C2b fragment has no known explicit function.

C2 activation by proteolytic cleavage is atypical relative to other serine proteases, which are generally activated by cleavage to create a new N-terminal residue, generally Ile. This new N-terminal residue binds into a hydrophobic cavity close to the active site to effect completion of the so-called oxyanion hole within the active site that is required to stabilize highly charged intermediate species formed during the catalytic cycle of the enzyme. In contrast, C2 is cleaved between its CCP3 and VWA domains, thus the serine protease domain is not activated in the 'normal' manner for serine proteases. The solution of the structure of C2a (Milder *et al.*, 2006) provided clues as to how this worked for C2a. It was found that the new N-terminus created when C2 was cleaved to form C2a nestled into a cavity within the VWA domain located close to the interface between the VWA and SP domains of C2a. This new conformation then causes a change in the active site of C2a such that the enzyme is in a conformation close to that required for catalytic competence. The authors speculated that full catalytic competence would most likely be attained following the binding of the cognate substrate (C3) to the

enzyme. Thus, the creation of the new N-terminus for this enzyme upon activation does indeed appear to be crucial to its activation. Cofactor interactions with C4b and the incoming C3 substrate are also likely to play a major role, suggestive of an enzyme with several levels of regulation and control built into its structure.

Lectin Pathway

The lectin pathway of complement activation (Figure 3) is mediated by the recognition of arrays of molecules on foreign surfaces that are not present on host surfaces. The most typical arrays recognized by molecules of this pathway are glycan arrays (Ohta *et al.*, 1990), hence the term for the pathway as a whole.

The first lectin recognition molecule discovered was mannose-binding lectin (MBL) (Kawasaki *et al.*, 1989). This molecule has been shown to be somewhat similar in architecture to the C1q molecule, with the collagen-like stalks providing attachment points for the MBL-associated serine proteases (MASPs) (Gingras *et al.*, 2011; Forneris *et al.*, 2012) and the globular heads recognizing high-mannose glycan arrays on pathogen surfaces (Turner and Hamvas, 2000). The molecule is composed of a base trimer unit, with no intrinsic ability to activate complement, with disulfide bonding furthering the assembly into higher order oligomers that have been shown to be able to efficiently activate the system. More recently, a number of other recognition molecules have been identified for this pathway, such as the three ficolin molecules and collectin-11 (Endo *et al.*, 2011; Yongqing *et al.*, 2013; Ma *et al.*, 2013). These different recognition molecules appear to allow the system to recognize an overlapping, yet diverse spectrum of pathogens and thus mediate activation of the complement system.

The MASPs have very similar domain architecture to C1r and C1s of the classical pathway, but only two genes encode the 5 proteins that are produced in this system, with gene splicing providing a mechanism to generate the protein level complexity (Yongqing *et al.*, 2012). This fact alone makes study of these molecules by 'normal' genetic manipulation of model organisms more challenging. The *MASP1/3* gene encodes three proteins (Yongqing *et al.*, 2012): MASP-1, MASP-3, and Map44. MASP-1 and MASP-3 share the same CUB1-EGF-CUB2-CCP1-CCP2 region, but differ in their SP domains, while Map44 has the same CUB1-EGF-CUB2-CCP1 domains as the two enzymes, followed by a unique 17 amino acid segment C-terminal to the CCP1 domain. The *MASP-2* gene gives rise to two proteins: MASP-2 and sMAP. Once again, sMAP is a truncated version of MASP-2, consisting of the CUB1-EGF domains of the protein. The role of the truncated versions of the MASP proteases is not clear. It has been postulated that they compete for binding to recognition molecules and thus regulate the activity of the system (Iwaki *et al.*, 2006) since they are essentially inactive in every other sense thus far tested, but their precise roles are not clear at this point. It is evident that the MASPs and their truncated versions are able to bind as homo- and heterodimers to the recognition molecules, thus giving rise to enormous complexity in the composition of these complexes in an individual.

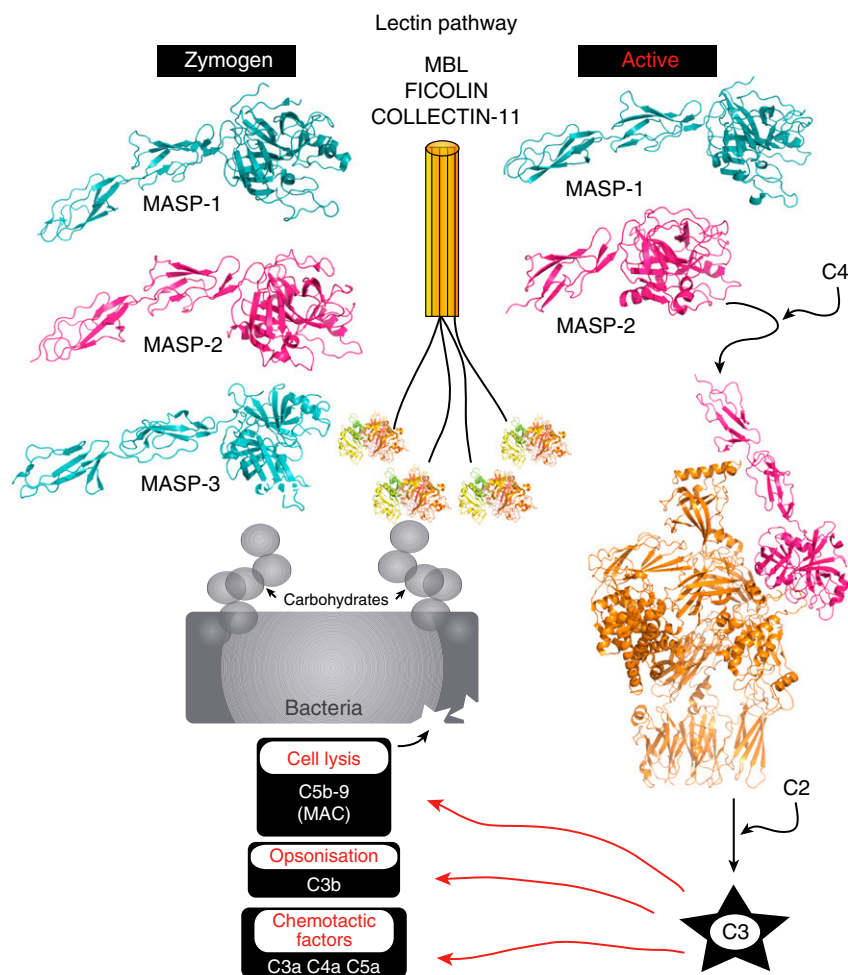


Figure 3 A schematic representation of the lectin complement pathway. The lectin pathway is activated when either MBL, collectin-11 or ficolins (structure for globular head of M-ficolin shown (PDB identifier: 2JHH)) bind to pathogen-specific carbohydrate arrays. This activates MASP-1 and MASP-2, converting them from zymogen (structures shown for MASP-1 (4IGD) and MASP-2 (1ZJK)) into active forms (structures shown for MASP-1 (3GOV) and MASP-2 (1Q3X)). Zymogen MASP-3 (4KKD) is also included. Activated MASP-2 binds and then cleaves C4 (structure for MASP-2/C4 shown: 4FXG) and C2, which combine to form the C3 convertase (C4bC2a) from whence the pathway culminates in cell lysis, opsonisation, and the release of chemotactic peptides.

MASP-1 was the first protease of the system identified (Ji *et al.*, 1993) and was thought to play a lone hand in activating the system until the discovery of MASP-2. Both molecules are capable of autoactivating and, following the discovery that MASP-2 was capable of cleaving both C4 and C2 (Vorup-Jensen *et al.*, 2000), while MASP-1 was inefficient at cleaving C4 in particular (Rossi *et al.*, 2001), it was thought for some time that MASP-2 was the sole activator of the lectin pathway. More recently, it has been shown that MASP-1 is necessary for efficient activation of MASP-2 and additionally contributes to the activation of the pathway further by playing a part in activating a fraction of the C2 activated by this pathway (Heja *et al.*, 2012). It is therefore clear, however, that MASP-2 is solely responsible for cleaving C4 and this becomes the pivotal reaction in pathway activation once MASP-2 has been activated. Initial studies showed that the CCP domains of MASP-2 play an important role in the efficient interaction between the protease and C4 (Rossi *et al.*, 2001; Ambrus *et al.*, 2003), while

the structure of the MASP-2/C4 complex has shown that both the CCP domains and the SP domains harbor exosites for C4 (Kidmose *et al.*, 2012). The precise details of how MASP-2 and MASP-1 cleave C2 efficiently have been less studied, aside from noting that the SP domain of MASP-2 appears to contain all of the required determinants to cleave C2 efficiently (Rossi *et al.*, 2001; Ambrus *et al.*, 2003).

When MASP-3 was first identified, initial experiments suggested that the enzyme played a role in regulating the complement system since it had low proteolytic activity, but could compete for binding sites on the lectin recognition molecules with MASP-1 and MASP-2 (Dahl *et al.*, 2001). Later, it was shown that murine MASP-3 was able to activate mouse factors D and B and thus activates the alternative pathway of complement in mice (Iwaki *et al.*, 2011). The relevance of this in the human system is, however, uncertain since similar results were not demonstrated in samples derived from human patients (Degn *et al.*, 2012). The role of MASP-3 in

complement is therefore unknown at this point, yet it is found in all vertebrates, including birds, which do not have MASP-1, for example, thus it is likely to have an important role. Most recently, it has been shown that mutants of MASP-3 are associated with the 3MC syndrome, characterized by skeletal and cranial deformities (Sirmaci *et al.*, 2010; Rooryck *et al.*, 2011). It is interesting to note here that mutants of the recognition molecule, CL-11, are also associated with this disease, suggesting that perhaps complexes of MASP-3 with CL-11 may play an important role in development. Genetic studies indicated that MASP-3 might be playing a role in neural crest migration in development of zebrafish embryos, suggesting a role for the enzyme in development (Rooryck *et al.*, 2011). Most mutants of the enzyme could clearly be predicted to be inactive, but our recent study (Yongqing *et al.*, 2013) has shown that two point mutants with uncertain outcomes with regard to the activity of the enzyme are indeed inactive and the solution of the first structure of MASP-3 (in a zymogen form) showed the mechanism whereby such a mutation causes the enzyme to be inactive. Following on directly from this, the structure of the active form of the enzyme in complex with an inhibitor, ecotin, completed the initial structural picture for this enzyme, although much remains to be learned about its structure and biology (Gaboriaud *et al.*, 2013). It was previously shown that enzyme could cleave insulin growth factor binding protein-5 (Cortesio and Jiang, 2006) and one study has suggested that cleavage of this molecule might lead to the normal development afforded by the release of insulin growth factor from the binding protein (Sirmaci *et al.*, 2010), but no studies have been carried out to test this hypothesis and thus the crucial substrate for the enzyme in this context remains obscure.

It has been shown that MASP-1 and MASP-2 are both reasonably efficiently inhibited by C1-inhibitor in the absence of glycosaminoglycans (GAGs) (Dobo *et al.*, 2009; Kerr *et al.*, 2008). The presence of GAGs increases the rate of interaction of MASP-2 with C1-inhibitor by approximately 10-fold (Parej *et al.*, 2013), but it also increases the rate of interaction between antithrombin and the enzymes to a level where this serpin can also be considered a physiologically significant inhibitor of these proteases. Interestingly, α -2-macroglobulin has been shown to be irrelevant as a MASP inhibitor under physiological conditions. MASP-3 is not inhibited by any inhibitor identified to date.

The Alternative Pathway

The alternative pathway of complement activation (Figure 4) functions not only as a means of activating the system, it has also been postulated to be an amplification mechanism for the classical and lectin pathways.

The mechanism of activation of the alternative pathway is based upon the intrinsic instability of the thioester bond in C3 compared to its homologous counterparts, C4 and C5 (Pangburn and Rawal, 2002). The instability of the thioester bond in C3 means that this molecule spontaneously attaches itself to surfaces at a much greater rate than C4 and C5 (Pangburn *et al.*, 1981a,b). Unless the molecule is cleared, its attachment triggers the formation of a C3b-like conformation,

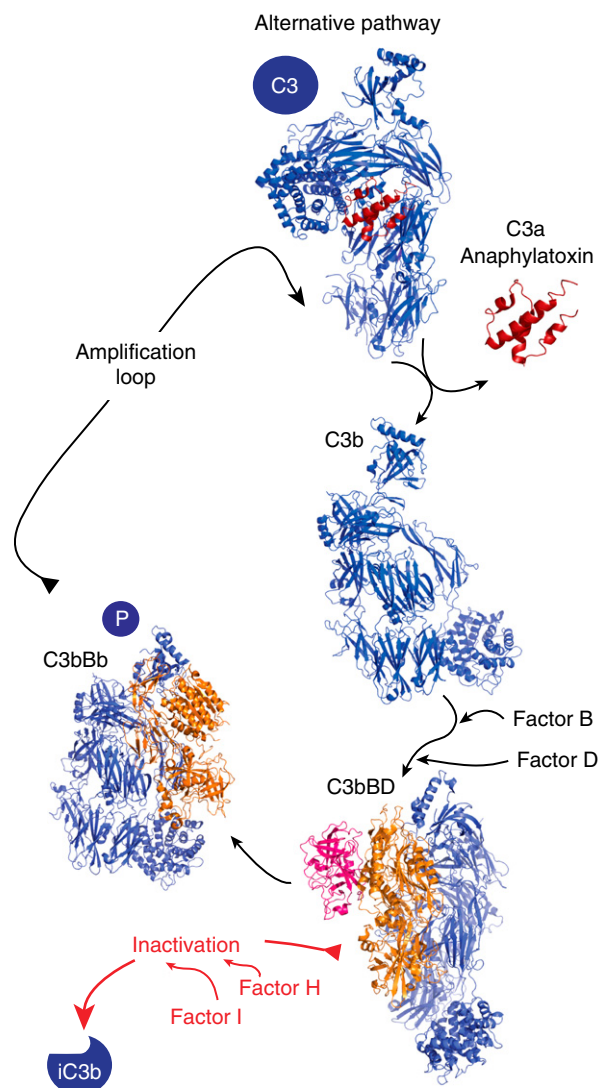


Figure 4 A schematic representation of the alternative complement pathway. Spontaneous C3 (PDB identifier: 2A73) turnover in blood or generated by the lectin or classical pathways results in the formation of C3a and C3b. C3b bound to a substrate binds the serine protease factor fB. This low-affinity complex is converted by the serine protease factor D (2XWB) into a C3-cleaving enzyme, C3bBb (2XWJ). This enzyme complex is stabilized by properdin (P). C3bBbP then cleaves C3 to C3a and C3b to complete the loop and lead to amplification. Since activation of the alternative pathway is spontaneous, regulators are particularly critical for preventing its excessive activation. Factor H, by binding to C3b, promotes the cleavage of the latter by the serine protease factor I to iC3b.

which in turn allows the attachment of factor B (fB) from the plasma. fB is homologous to component C2 of the classical and lectin pathways, with similar domain architecture of three N-terminal CCP domains linked to VWA and SP domains. Once bound to surface localized C3 or C3b, fB is activated by circulating factor D via cleavage between the third CCP domain and the VWA domain, much as occurs for the activation of C2 by C1s or MASP-2. Activated fB is referred to

as Bb and the C3bBb complex becomes a C3-convertase. Once stabilized by the binding of properdin, this C3 convertase is capable of cleaving further C3 molecules and thus allowing the surface of initial attachment to become decorated with C3b molecules that can attract further fB molecules. This mechanism therefore allows for activation of the alternative pathway or amplification of the other two pathways due to the central involvement of bound C3b molecules. Mammalian cells have a host of factors, such as decay-accelerating factor and complement receptor 1, that prevent or clear the initial attachment of C3 before it can bind the fB required to activate this pathway. This mechanism therefore provides an innate difference between foreign and host surfaces.

The C3bBb C3 convertase complex not only gives rise to further C3b molecules that can independently bind to the foreign surface, but a further molecule of C3b can associate with this complex. Since this extra C3b molecule provides a binding site for C5, the C3b₂Bb complex becomes a C5 convertase complex and thus can cleave C5 to form the C5b required to assemble the C6–C9 components into the MAC.

The mechanisms involved in the formation and action of the alternative pathway convertase complexes have been more studied than their classical/lectin counterparts (Forneris *et al.*, 2012). Much of the reason for this lies in the availability of molecules from bacteria and snakes that stabilize the alternative pathway complexes (e.g., Staphylococcal complement inhibitor (SCIN) molecule from *Staphylococcus aureus*) or bind to components of the complex in the same way as the human counterparts (cobra venom factor). The structure of the SCIN-stabilized C3bBb complex, for example, Rooijakkers *et al.* (2009) gave important insights into the assembly and molecular mechanisms underlying the operation of this complex in the activation of the alternative pathway. It was found that the Bb molecule bound to the C345 domain of the C3b via the N-terminal section of the VWA domain of Bb. The tight interactions between C345 and the VWA domain were not mirrored elsewhere, resulting in the Bb molecule essentially being free to ‘swing’ around this axis on the C3b molecule. This would allow it to be able to reach the scissile bonds in incoming C3 and eventually C5 molecules. The formation of a dimer between the C3 molecules in the unit cell of the crystals also pointed the way toward mechanisms whereby additional C3 molecules could bind to the C3b in the complex and thus be cleaved by the Bb component. Later work showed how factor D, which circulates in human plasma as a cleaved and activated form, was held in a disordered, inactive state until it bound to the C3bBb complex via an exosite on the factor D (Forneris *et al.*, 2010). The binding of factor D via its exosite caused an ordering of its active site such that it was able to cleave fB to yield activated Bb, which in turn allowed the C3bBb complex to act as a C3 convertase. Later work involving the solution of a structure of a complex between the C3-like cobra venom factor and C5 (Laursen *et al.*, 2011) showed how C5 was likely to be bound by C3b molecules in the convertase complexes and thus be presented for cleavage by Bb. The work on the alternative pathway complexes naturally also shed light on the likely mechanisms operative in the classical/lectin pathway convertase complexes.

Regulation of Complement Activation by Proteases and Binding Factors

Aside from the inhibition of C1r, C1s, MASP-1, and MASP-2 by C1-inhibitor, the complement system is highly regulated by a number of molecules. These molecules are particularly important to prevent unwarranted complement activation on host surfaces. The protease, factor I, associates with other regulators, such as factor H, bound to C3 and C4 and then cleaves them into inactive iC3b and iC4b forms (additional smaller forms, such as C3c and C4c can subsequently be formed). The enzyme consists of five domains: the N-terminal fl membrane-attack complex (FIMAC) domain, a scavenger receptor cysteine-rich (SRCR) domain, two low-density lipoprotein receptor (LDLRA) domains, and the serine protease domain. The enzyme is secreted in a cleaved form, but is inactive due to extensive disorder in its serine protease domain (Roversi *et al.*, 2011) that is maintained by the binding of the N-terminal region to the C-terminal serine protease domain. Binding of the enzyme to a C3b-factor H complex, for example, releases the N-terminal heavy region from the SP domain, allowing formation of the active conformation of the enzyme and thus promoting the cleavage of the C3b into inactive forms by the factor I.

See also: Protein Synthesis/Degradation: Proteolytic Pathways: Overview of Blood Coagulation and the Pathophysiology of Blood Coagulation Disorders

References

- Ambrus, G., Gal, P., Kojima, M., *et al.*, 2003. Natural substrates and inhibitors of mannan-binding lectin-associated serine protease-1 and -2: A study on recombinant catalytic fragments. *Journal of Immunology* 170, 1374–1382.
- Bally, I., Rossi, V., Lunardi, T., *et al.*, 2009. Identification of the C1q-binding sites of human C1r and C1s: A refined three-dimensional model of the C1 complex of complement. *Journal of Biological Chemistry* 284, 19340–19348.
- Brier, S., Pflieger, D., Le Mignon, M., *et al.*, 2010. Mapping surface accessibility of the C1r/C1s tetramer by chemical modification and mass spectrometry provides new insights into assembly of the human C1 complex. *Journal of Biological Chemistry* 285, 32251–32263.
- Budayova-Spano, M., Grabarse, W., Thielens, N.M., *et al.*, 2002a. Monomeric structures of the zymogen and active catalytic domain of complement protease C1r: Further insights into the C1 activation mechanism. *Structure* 10, 1509–1519.
- Budayova-Spano, M., Lacroix, M., Thielens, N.M., *et al.*, 2002b. The crystal structure of the zymogen catalytic domain of complement protease C1r reveals that a disruptive mechanical stress is required to trigger activation of the C1 complex. *The EMBO Journal* 21, 231–239.
- Caliezi, C., Willemin, W.A., Zeerleder, S., *et al.*, 2000. C1-Esterase inhibitor: An anti-inflammatory agent and its potential use in the treatment of diseases other than hereditary angioedema. *Pharmacological Reviews* 52, 91–112.
- Carroll, M.C., 2004. The complement system in regulation of adaptive immunity. *Nature Immunology* 5, 981–986.
- Cortésio, C.L., Jiang, W., 2006. Mannan-binding lectin-associated serine protease 3 cleaves synthetic peptides and insulin-like growth factor-binding protein 5. *Archives of Biochemistry and Biophysics* 449, 164–170.
- Dahl, M.R., Thiel, S., Matsushita, M., *et al.*, 2001. MASP-3 and its association with distinct complexes of the mannan-binding lectin complement activation pathway. *Immunity* 15, 127–135.
- Degn, S.E., Jensen, L., Hansen, A.G., *et al.*, 2012. Mannan-binding lectin-associated serine protease (MASP)-1 is crucial for lectin pathway activation in human serum, whereas neither MASP-1 nor MASP-3 is required for alternative pathway function. *Journal of Immunology* 189, 3957–3969.

- Degn, S.E., Jensenius, J.C., Thiel, S., 2011. Disease-causing mutations in genes of the complement system. *American Journal of Human Genetics* 88, 689–705.
- Diebolder, C.A., Beurskens, F.J., De Jong, R.N., *et al.*, 2014. Complement is activated by IgG hexamers assembled at the cell surface. *Science* 343, 1260–1263.
- Discipio, R.G., 1981. The conversion of human complement component C5 into fragment C5b by the alternative-pathway C5 convertase. *Biochemical Journal* 199, 497–504.
- Dobo, J., Harmat, V., Beinrohr, L., *et al.*, 2009. MASP-1, a promiscuous complement protease: Structure of its catalytic region reveals the basis of its broad specificity. *Journal of Immunology* 183, 1207–1214.
- Duncan, R.C., Mohlin, F., Taleski, D., *et al.*, 2012. Identification of a catalytic exosite for complement component C4 on the serine protease domain of C1s. *Journal of Immunology* 189, 2365–2373.
- Duncan, R.C., Wijeyewickrema, L.C., Pike, R.N., 2008. The initiating proteases of the complement system: Controlling the cleavage. *Biochimie* 90, 387–395.
- Dunkelberger, J.R., Song, W.C., 2010. Complement and its role in innate and adaptive immune responses. *Cell Research* 20, 34–50.
- Endo, Y., Matsushita, M., Fujita, T., 2011. The role of ficolins in the lectin pathway of innate immunity. *International Journal of Biochemistry & Cell Biology* 43, 705–712.
- Forneris, F., Ricklin, D., Wu, J., *et al.*, 2010. Structures of C3b in complex with factors B and D give insight into complement convertase formation. *Science* 330, 1816–1820.
- Forneris, F., Wu, J., Gros, P., 2012. The modular serine proteases of the complement cascade. *Current Opinion in Structural Biology* 22, 333–341.
- Gaboriaud, C., Gupta, R.K., Martin, L., *et al.*, 2013. The serine protease domain of MASP-3: Enzymatic properties and crystal structure in complex with ecotin. *PLoS One* 8, e67962.
- Gaboriaud, C., Juanhuix, J., Gruez, A., *et al.*, 2003. The crystal structure of the globular head of complement protein C1q provides a basis for its versatile recognition properties. *Journal of Biological Chemistry* 278, 46974–46982.
- Gaboriaud, C., Rossi, V., Bally, I., Arlaud, G.J., Fontecilla-Camps, J.C., 2000. Crystal structure of the catalytic domain of human complement C1s: A serine protease with a handle. *The EMBO Journal* 19, 1755–1765.
- Gaboriaud, C., Thielens, N.M., Gregory, L.A., *et al.*, 2004. Structure and activation of the C1 complex of complement: Unraveling the puzzle. *Trends in Immunology* 25, 368–373.
- Gingras, A.R., Girija, U.V., Keeble, A.H., *et al.*, 2011. Structural basis of mannan-binding lectin recognition by its associated serine protease MASP-1: Implications for complement activation. *Structure* 19, 1635–1643.
- Gros, P., Milder, F.J., Janssen, B.J., 2008. Complement driven by conformational changes. *Nature Reviews Immunology* 8, 48–58.
- Heja, D., Harmat, V., Fodor, K., *et al.*, 2012. Monospecific inhibitors show that both mannan-binding lectin-associated serine protease-1 (MASP-1) and -2 are essential for lectin pathway activation and reveal structural plasticity of MASP-2. *Journal of Biological Chemistry* 287, 20290–20300.
- Iwaki, D., Kanno, K., Takahashi, M., *et al.*, 2006. Small mannose-binding lectin-associated protein plays a regulatory role in the lectin complement pathway. *Journal of Immunology* 177, 8626–8632.
- Iwaki, D., Kanno, K., Takahashi, M., *et al.*, 2011. The role of mannose-binding lectin-associated serine protease-3 in activation of the alternative complement pathway. *Journal of Immunology* 187, 3751–3758.
- Ji, Y.H., Fujita, T., Hatsuse, H., *et al.*, 1993. Activation of the C4 and C2 components of complement by a proteinase in serum bactericidal factor, Ra reactive factor. *Journal of Immunology* 150, 571–578.
- Kardos, J., Gal, P., Szilagyi, L., *et al.*, 2001. The role of the individual domains in the structure and function of the catalytic region of a modular serine protease, C1r. *Journal of Immunology* 167, 5202–5208.
- Kardos, J., Harmat, V., Pallo, A., *et al.*, 2008. Revisiting the mechanism of the autoactivation of the complement protease C1r in the C1 complex: Structure of the active catalytic region of C1r. *Molecular Immunology* 45, 1752–1760.
- Kawasaki, N., Kawasaki, T., Yamashina, I., 1989. A serum lectin (mannan-binding protein) has complement-dependent bactericidal activity. *Journal of Biochemistry* 106, 483–489.
- Kerr, F.K., Thomas, A.R., Wijeyewickrema, L.C., *et al.*, 2008. Elucidation of the substrate specificity of the MASP-2 protease of the lectin complement pathway and identification of the enzyme as a major physiological target of the serpin, C1-inhibitor. *Molecular Immunology* 45, 670–677.
- Kidmose, R.T., Laursen, N.S., Dobo, J., *et al.*, 2012. Structural basis for activation of the complement system by component C4 cleavage. *Proceedings of the National Academy of Sciences of the United States of America* 109, 15425–15430.
- Laursen, N.S., Andersen, K.R., Braren, I., *et al.*, 2011. Substrate recognition by complement convertases revealed in the C5-cobra venom factor complex. *The EMBO Journal* 30, 606–616.
- Law, S.K., Dodds, A.W., 1997. The internal thioester and the covalent binding properties of the complement proteins C3 and C4. *Protein Science* 6, 263–274.
- Ma, Y.J., Skjoedt, M.O., Garred, P., 2013. Collectin-11/MASP complex formation triggers activation of the lectin complement pathway—the fifth lectin pathway initiation complex. *Journal of Innate Immunity* 5, 242–250.
- Milder, F.J., Raaijmakers, H.C., Vandeputte, M.D., *et al.*, 2006. Structure of complement component C2A: Implications for convertase formation and substrate binding. *Structure* 14, 1587–1597.
- Mollnes, T.E., Song, W.C., Lambris, J.D., 2002. Complement in inflammatory tissue damage and disease. *Trends in Immunology* 23, 61–64.
- Ohta, M., Okada, M., Yamashina, I., Kawasaki, T., 1990. The mechanism of carbohydrate-mediated complement activation by the serum mannan-binding protein. *Journal of Biological Chemistry* 265, 1980–1984.
- Pangburn, M.K., Rawal, N., 2002. Structure and function of complement C5 convertase enzymes. *Biochemical Society Transactions* 30, 1006–1010.
- Pangburn, M.K., Schreiber, R.D., Muller-Eberhard, H.J., 1981a. Formation of the initial C3 convertase of the alternative complement pathway. Acquisition of C3b-like activities by spontaneous hydrolysis of the putative thioester in native C3. *Journal of Experimental Medicine* 154, 856–867.
- Pangburn, S.H., Newton, R.S., Chang, C.M., Weinstein, D.B., Steinberg, D., 1981b. Receptor-mediated catabolism of homologous low density lipoproteins in cultured pig hepatocytes. *Journal of Biological Chemistry* 256, 3340–3347.
- Parej, K., Dobo, J., Zavodsky, P., Gal, P., 2013. The control of the complement lectin pathway activation revisited: Both C1-inhibitor and antithrombin are likely physiological inhibitors, while alpha2-macroglobulin is not. *Molecular Immunology* 54, 415–422.
- Perry, A.J., Wijeyewickrema, L.C., Wilmann, P.G., *et al.*, 2013. A molecular switch governs the interaction between the human complement protease C1s and its substrate, complement C4. *Journal of Biological Chemistry* 288, 15821–15829.
- Phillips, A.E., Toth, J., Dodds, A.W., *et al.*, 2009. Analogous interactions in initiating complexes of the classical and lectin pathways of complement. *Journal of Immunology* 182, 7708–7717.
- Ricklin, D., Hajishengallis, G., Yang, K., Lambris, J.D., 2010. Complement: A key system for immune surveillance and homeostasis. *Nature Immunology* 11, 785–797.
- Roijakkers, S.H., Wu, J., Ruyken, M., *et al.*, 2009. Structural and functional implications of the alternative complement pathway C3 convertase stabilized by a staphylococcal inhibitor. *Nature Immunology* 10, 721–727.
- Rooryck, C., Diaz-Font, A., Osborn, D.P., *et al.*, 2011. Mutations in lectin complement pathway genes COLEC11 and MASP1 cause 3MC syndrome. *Nature Genetics* 43, 197–203.
- Rossi, V., Bally, I., Thielens, N.M., Esser, A.F., Arlaud, G.J., 1998. Baculovirus-mediated expression of truncated modular fragments from the catalytic region of human complement serine protease C1s. Evidence for the involvement of both complement control protein modules in the recognition of the C4 protein substrate. *Journal of Biological Chemistry* 273, 1232–1239.
- Rossi, V., Cseh, S., Bally, I., *et al.*, 2001. Substrate specificities of recombinant mannan-binding lectin-associated serine proteases-1 and -2. *Journal of Biological Chemistry* 276, 40880–40887.
- Roversi, P., Johnson, S., Caesar, J.J., *et al.*, 2011. Structural basis for complement factor I control and its disease-associated sequence polymorphisms. *Proceedings of the National Academy of Sciences of the United States of America* 108, 12839–12844.
- Sim, R.B., Arlaud, G.J., Colomb, M.G., 1980. Kinetics of reaction of human C1-inhibitor with the human complement system proteases C1r and C1s. *Biochimica et Biophysica Acta* 612, 433–449.
- Sirmaci, A., Walsh, T., Akay, H., *et al.*, 2010. MASP1 mutations in patients with facial, umbilical, coccyeal, and auditory findings of Carnevale, Malpuech, OSA, and Michels syndromes. *American Journal of Human Genetics* 87, 679–686.
- Turner, M.W., Hamvas, R.M., 2000. Mannose-binding lectin: Structure, function, genetics and disease associations. *Reviews in Immunogenetics* 2, 305–322.
- Venkatraman Girija, U., Gingras, A.R., Marshall, J.E., *et al.*, 2013. Structural basis of the C1q/C1s interaction and its central role in assembly of the C1 complex of complement activation. *Proceedings of the National Academy of Sciences of the United States of America* 110, 13916–13920.
- Vorup-Jensen, T., Petersen, S.V., Hansen, A.G., *et al.*, 2000. Distinct pathways of mannan-binding lectin (MBL)- and C1-complex autoactivation revealed by reconstitution of MBL with recombinant MBL-associated serine protease-2. *Journal of Immunology* 165, 2093–2100.

- Wallis, R., Mitchell, D.A., Schmid, R., Schwaebler, W.J., Keeble, A.H., 2010. Paths reunited: Initiation of the classical and lectin pathways of complement activation. *Immunobiology* 215, 1–11.
- Wijeyewickrema, L.C., Yongqing, T., Tran, T.P., *et al.*, 2013. Molecular determinants of the substrate specificity of the complement-initiating protease, C1r. *Journal of Biological Chemistry* 288, 15571–15580.
- Wuillemin, W.A., Te Velhuis, H., Lubbers, Y.T., *et al.*, 1997. Potentiation of C1 inhibitor by glycosaminoglycans: Dextran sulfate species are effective inhibitors of *in vitro* complement activation in plasma. *Journal of Immunology* 159, 1953–1960.
- Yongqing, T., Drentin, N., Duncan, R.C., Wijeyewickrema, L.C., Pike, R.N., 2012. Mannose-binding lectin serine proteases and associated proteins of the lectin pathway of complement: Two genes, five proteins and many functions? *Biochimica et Biophysica Acta* 1824, 253–262.
- Yongqing, T., Wilmann, P.G., Reeve, S.B., *et al.*, 2013. The X-ray crystal structure of mannose-binding lectin-associated serine proteinase-3 reveals the structural basis

for enzyme inactivity associated with the Carnevale, Mingarelli, Malpuech, and Michels (3MC) syndrome. *Journal of Biological Chemistry* 288, 22399–22407.

Relevant Websites

<http://www.ecomplement.org>

European Complement Network.

<http://www.fh-hus.org>

FH (Factor H) aHUS (atypical haemolytic uremic syndrome) Mutation Database.

<http://www.complement.org>

International Complement Society.

<http://www.protease.org>

International Proteolysis Society.

Digestive Proteases: Roles in the Human Alimentary Tract

SG Bradley, Virginia Commonwealth University, Richmond, VA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Alimentary tract The mucus-lined muscular tube extending from the mouth to the anus; often used interchangeably with digestive tract or gastrointestinal tract.

Biofilm A thin layer of biological origin, usually containing bacteria, adhering to a surface that is exposed to an aqueous environment.

Bolus Food that has been chewed, mixed with saliva, and is ready to be swallowed.

Chyme Viscous fluid consisting of food partially digested in the stomach, and is ready to be passed into the small intestine.

Colon Major segment of the large intestine, including the ascending colon, transverse colon, descending colon and sigmoid colon but not the cecum, rectum and anal canal.

Endopeptidase A proteolytic enzyme that cleaves peptide bonds internally in a peptide rather than from the N- or C-terminus.

Fermentor A chamber for digestion and production of biological materials (such as antibiotics), often in the absence of air.

Heterodimer A molecule composed of two unlike subunits.

Homodimer A molecule composed of two like subunits.

Peristalsis Progressive waves of contractions that move digested food through the intestines during digestion.

Introduction

The digestive tract consists of several connected chambers, each serving unique functions. The various segments and chambers of this biological digester or fermentor are the sites of characteristic sequential processes that involve both enzymatic and physical activities (Vardakou *et al.*, 2011). Food enters the oral cavity and is processed for transfer to the stomach. The stomach is the first site of extensive proteolytic activity. The small intestine consists of three distinct segments anatomically: the duodenum, the jejunum and the ileum. The two segments of the small intestine nearer to the stomach, or proximal small intestine, are the sites of intense proteolysis and nutrient absorption. The distal small intestine, the ileum, continues the process of proteolysis, especially hydrolyzing peptides to free amino acids for absorption. The large intestine serves as a reclamation chamber, recovering water, and salts. The overall design and operation of the digestive tract resemble industrial processes used to produce antibiotics and therapeutic drugs. In both an industrial fermentor and the digestive tract there is a chamber where starting materials are mixed and prepared for the reaction chamber. Thereafter the product is recovered, and finally the waste materials are collected and eliminated. In human digestion of protein, the starting materials consist of exogenous substrates that are ingested periodically, supplemented by a steady addition of endogenous proteins as a result of release or secretion of proteinacious materials and cell shedding. The outputs of this integrated operation are the release and uptake of nutrients, the detoxification and elimination of waste materials, and the maintenance of the integrated system (Table 1). Proteolysis in the digestive tract releases bioactive peptides that act in the lumen or are absorbed (Chabance *et al.*, 1998) as well as peptides as a source of nutrients. In this article, the emphasis is on the key enzymes that release nutrients in foodstuffs. In addition, selected examples are described briefly of proteases that alter the activities of physiological active molecules such

as latent enzymes, chemokines, antigens and toxins and of proteases that serve to control biofilms interfering with absorption from or secretion into the lumen of the intestine (Johansson *et al.*, 2013). In addition to digestion by host proteases, microorganisms in the colon continue the hydrolysis of glycoproteins, polypeptides, oligopeptides, and nonprotein compounds with peptide bonds that have passed through the small intestine.

Role of the Oral Cavity in Proteolysis

Three processes critical for digestion of exogenous food occur in the oral cavity: (1) the teeth cut and grind the heterogeneous texture of ingested materials into smaller particles; (2) water and mucus are added as salivary secretions to make a somewhat viscous fluid that serves as a lubricant facilitating flow through the esophagus into the stomach; and (3) the digestive process is initiated by degrading starches and lipids that often coat food as a result of the preparation of meats or as an inherent part of the raw materials of vegetable proteins (Table 1). The 0.5–1.5 l of saliva secreted each day contains more than a thousand different proteins, most notably a number of amylases and lipases. Secreted saliva is slightly basic, pH 8, because of its sodium bicarbonate content, but the pH of masticated food decreases to pH 6 as carbon dioxide is released. The tongue acts as a mixing paddle and also has a role in moving the masticated food (bolus) into the pharynx and on to the esophagus. It is here that the respiratory tract and the digestive tract cross, creating the opportunity for food to be misdirected into the bronchus leading to choking.

The contents of the oral cavity include endogenous proteolytic enzymes and those produced by the oral microbial flora, an array of other proteins that serves as substrates, protease inhibitors, and antimicrobial products. The parotid gland secretions contain proline-rich proteins and secretory

Table 1 General description of the adult human digestive tract with an emphasis on protein digestion and amino acid absorption^a

Mouth	Stomach	Proximal small intestine	Distal small intestine	Colon/large intestine	Rectum
Maceration	Churning	Emulsification	Peristalsis	Resorption (H ₂ O)	Defecation
Lubrication	Acidification	Neutralization	Mucin shedding	Microbial flora	Microbial flora
2/3 cup volume	¼ cup to 1 quart	6–8 feet long	10–12 feet long	4.5 feet long	4 in. long
Held ca. 1 min	Retained 1–3 h	Flow through 2–3 h	Flow through 2–3 h	Held 15–20 h	Held 15–20 h
pH 6.5	pH 2	pH 6	pH 7.5	pH 5.7	pH 6.7
Mucosa lining	Protective mucus	Villi and microvilli	Villi and Microvilli	Invaginations/crypts	Mucus/muscular
Salivary glands	Parietal cells/HCl	Pancreatic enzymes/	Lymphoid Peyer's	K ⁺ /Cl ⁻ flux, soluble	Reflux to colon
amylases and	aspartic proteases	bile serine	patches	meprin A and shed	bacterial proteases
protease inhibitors	e.g., pepsin	proteases e.g.,	metalloproteases	B, bacterial	
		trypsin	e.g., meprins	proteases	
Proteins in food	Peptides released	Amino acids released	Amino acid, vitamin	Absorption vitamin	Undigested fiber and
particles exposed	and limited	and absorbed	B ₁₂ and bile salts	K, bile acids and	bacterial mass
	absorption		absorption	biotin	

^aData vary greatly according to age, gender, health status, and composition of the food.

Table 2 Important proteases in the digestive tract of humans

Enzymes	Enzyme types	Locations
Pepsin EC 3.4.23.1	Aspartic endopeptidase	Stomach
Trypsin isoforms EC 3.4.21.4	Serine endopeptidases	Duodenum primarily
Chymotrypsin isoforms EC 3.4.21.1	Serine endopeptidases	Duodenum primarily
Elastase 2A and 2B EC 3.4.21.71; 3A & 3B EC 2.4.21.70	Serine endopeptidases	Duodenum primarily
Carboxypeptidases 2A, 2B, 3B EC 3.4.17.1, EC 3.4.17.2	Metallo-exopeptidases	Duodenum primarily
Aminopeptidases EC 3.4.11.X; leucylaminopeptidase X=1	Metallo-exopeptidases	Jejunum and ileum
Dipeptidyl dipeptidase IV EC 3.4.14.5	Serine dipeptidase	Lymphatic vessels, ileum, etc.
Meprins A and B EC 3.4.24.18 and EC 3.4.24.63	Metallo-endopeptidase	Ileum and colon

immunoglobulin A in addition to amylases and lipases. The metalloprotease meprin has been identified on the luminal surfaces of parotid and submandibular glands (Craig *et al.*, 1991). The submandibular/sublingual gland secretions contain mucous glycoproteins, and the cysteine protease inhibitor cystatin (Zakhary *et al.*, 2007). About 2–4 g of salivary protein enter stomach each day, representing 4–8% of all proteins digested (Helmerhorst *et al.*, 2008). Salivary proteins that have a number of biological activities are the histatins, which are 3–4 kDa histidine-rich proteins secreted by both parotid and submandibular glands (Edgerton *et al.*, 2000). Histatins inhibit matrix metalloproteases, presumably by binding zinc and are potent inhibitors of aerobic bacteria and the pathogenic yeast *Candida albicans* (Sun *et al.*, 2009). Histatin 1 has been found to stimulate re-growth of epithelia and wound healing, thereby contributing to the integrity of the lining of the oral cavity (Oudhoff *et al.*, 2009).

Role of the Stomach in Proteolysis

Three processes critical for the digestion of endogenous and exogenous protein occur in the stomach: (1) the parietal cells produce hydrochloric acid bringing the food/saliva mixture to pH 2; (2) the 42 kDa proenzyme pepsinogen, produced by chief cells, partially unfolds in the acidic environment of the stomach allowing limited self-digestion that generates active

pepsin; and (3) the churning action of the stomach continues the process of homogenizing the stomach contents, which are periodically released into the first segment of the small intestine, the duodenum, as chyme (Gritti *et al.*, 2000). A thick mucus layer lines the stomach to protect epithelial cells from damage by HCl (Ermund *et al.*, 2013). At pH 2, the terminus of pepsinogen becomes accessible to the proteolytic active site of the molecule allowing the cleavage and release of a 44 amino acid sequence (Richter *et al.*, 1998). The resulting 34 kDa active pepsin is an aspartic protease that preferentially cleaves the peptide bond between hydrophobic and aromatic amino acid residues, resulting in polypeptides and oligopeptides rather than free amino acids. Pepsin was labeled as EC3.4.4.1 in 1961 and re-labeled as EC3.4.23.1 in 1972 (Table 2). An EC number is a system for identifying enzymes based upon the chemical reaction that they catalyze; for example, enzymes denoted as EC3.4 are hydrolases that act on peptide bonds (Rawlings and Salvesen, 2013). Pepsin is most active at pH 2 and relatively inactive above pH 6.5. In most young mammals, milk casein is initially processed by the aspartic protease chymosin (rennin), but human infants lack this enzyme even though the human genome contains a chymosin-related sequence (Kolmer *et al.*, 1991). In humans, casein in milk is precipitated by stomach acidity and then degraded by pepsin. There is limited absorption of oligopeptides generated in the stomach, such as kappa-caseinoglycopeptide, which is an inhibitor of platelet aggregation.

Role of the Small Intestine in Proteolysis

The duodenal epithelium produces bicarbonate to neutralize the hydrochloric acid in the chyme entering from the stomach. This protects the duodenal epithelium from direct damage and also makes the lumen favorable for the abundant digestive serine proteases, such as trypsin, released by the pancreas into the lumen (Rune and Viskum, 1969). The bile duct also empties into the duodenum, facilitating dispersion and emulsification of particles in the chyme. Bicarbonate concentration in the intestinal lumen determines pH, which is a key variable in enzyme activity, and also affects absorption and transport of nutrients (Davis *et al.*, 1980). The duodenum is only 10-in. long, but the epithelium is folded and covered with villi and microvilli, which amplify the effective absorptive surface by 20- to 50-fold. Digestion by serine proteases of oligopeptides to amino acids and absorption continue throughout the duodenum and jejunum (proximal small intestine). The transient time in the 6–8 foot proximal small intestine is about 2–3 h (Table 1).

Enteropeptidase (EC3.4.21.9) is a membrane-bound serine endopeptidase of the duodenal and jejunal brush borders (Zheng *et al.*, 2009). Enteropeptidase is produced as a proenzyme that is activated by other serine proteases such as duodenase (EC 3.4.21.23) or trypsin (Zamolodchikova *et al.*, 1995). Active enteropeptidase cleaves the specific trypsinogen activation peptide, Asp-Asp-Asp-Asp-Lys-Ile, to generate active trypsin. Following activation, trypsin cleaves and activates other proenzymes in the duodenum, including chymotrypsinogen, proelastase, and procarboxypeptidases (Lu *et al.*, 1997). Infants deficient in enteropeptidase develop diarrhea and edema, and fail to thrive (Holzinger *et al.*, 2002).

Trypsin (EC 3.4.21.4) is a 24 kDa serine endopeptidase, produced in the pancreas and secreted into the lumen of the duodenum. Actually three isoforms of trypsinogen are produced and each is activated by cleaving the carboxyl side of arginine and lysine residues (Szmola and Sahin-Tóth, 2007). Trypsin is the most abundant protease, constituting 19% of the protein in pancreatic secretions (Whitcomb and Lowe, 2007). Trypsin has a deep pocket with negatively charged amino acid residues at the bottom that favor binding of positively charged substrates. The catalytic amino acid residues of trypsin are His57... Asp102 ... Ser195, which are widely separated in sequence but conformationally brought together (Polgár, 2005). A proline residue on the carboxyl side of Lys or Arg prevents hydrolysis and an acidic residue on either side of Lys or Arg reduces the rate of cleavage. Trypsin undergoes autolysis to yield several active derivatives. Calcium is required for activation of trypsin (Cliffe and Grant, 1983) and for preventing autolysis leading to loss of proteolytic activity (Abbott *et al.*, 1975). The pH optimum for trypsin is 8, which is slightly more basic than its milieu in the duodenum and jejunum.

Chymotrypsin (EC 3.4.21.1) is a 26 kDa serine carboxypeptidase that preferentially cleaves the amide bond (the P₁ position) of an aromatic amino acid residues such as tyrosine, tryptophan and phenylalanine. Chymotrypsinogen constitutes 9% of the protein in pancreatic secretions (Whitcomb and Lowe, 2007). Chymotrypsin is produced as a 42 kDa proenzyme by acinar cells of the pancreas and released in vesicles, along with trypsin and a trypsin inhibitor, into the duodenum.

There are several isoforms of chymotrypsinogen having different substrate specificities and kinetics (Hudáky *et al.*, 1999) each of which is cleaved by trypsin into three segments that remain connected by S–S bonds. One of the isoforms, chymotrypsin C, degrades trypsinogen in the calcium-rich lumen of the duodenum and degrades trypsin in the calcium-poor lumen of the lower intestine (Szmola and Sahin-Tóth, 2007). The pH optimum and active site of chymotrypsin are the same as those of trypsin; that is, pH 8 and His...Asp...Ser, respectively (Polgár, 2005). Chymotrypsin has a deep cavity formed by hydrophobic amino acid residues, which accounts for the preference of the enzyme for peptide bonds containing tyrosine, phenylalanine or tryptophan (Hedstrom *et al.*, 1992).

Elastases 2A, 2B and 3B (EC 3.4.21.-) are endopeptidases produced in the pancreas as proenzymes that are activated by trypsin (Kawashima *et al.*, 1987). Humans have six loci coding for similar elastases that are differentially expressed throughout the body (Whitcomb and Lowe, 2007). Elastase 2A and 2B, which have 90% amino acid homology and similar conformations, cleave elastin-type proteins after Ala, Gly, Ser, Met and Phe residues. Elastase 3B has less proteolytic activity than elastases 2A and 2B. Consistent with the pH of the intestinal lumen, the optimum pH for elastases is 8.

Carboxypeptidases (CP) are zinc-containing exopeptidases that remove single amino acids from the carboxyl end of oligopeptides, many of which resulted from digestion of dietary proteins by pepsin, trypsin and chymotrypsin. Carboxypeptidases are produced in the pancreas as proenzymes that are activated by trypsin. Carboxypeptidase A (CPA) shows preference for terminal aliphatic or aromatic amino acids and has little or no activity on C-terminal Asp, Glu, Arg, Lys or Pro (Reverter *et al.*, 1997). In contrast, carboxypeptidase B (CPB) shows preference for basic C-terminal amino acid residues such as Lys and Arg. The differences in specificities between CPA and CPB can be attributed to the residues Ser205, Gly241, and Asp253 in CPB as compared to Gly207, Ile243, and Ile255 in CPA (Coll *et al.*, 1991). Carboxypeptidase O (CPO) shows preference for acidic C-terminal residues. CPA and CPB are found in the duodenum whereas CPO is located in the ileum.

Aminopeptidases (AP; EC subgroup 3.4.11.-) remove single amino acids (aminoacyl-peptide hydrolases) and dipeptides (dipeptidyl-peptide hydrolases) from the amino end of polypeptides. Many but not all aminopeptidases are zinc-dependent metallo-exopeptidases. Aminopeptidases are found in the cytoplasm and brush-border membrane of the intestinal epithelial and are released into the lumen as a result of cellular shedding. The optimum pH for aminopeptidases is between 7.5 and 8.5. As much as 30 g of mucosal cells are sloughed each day into the intestinal lumen (Freeman and Kim, 1978). The aminopeptidases in the jejunum originate in the cytoplasm of mucosal cells whereas the aminopeptidases in the ileum come from the brush-border membrane. Alanine aminopeptidase (EC 3.4.11.2), for example, is a glycoprotein in the distal ileum with a molecular mass of 118 kDa for the monomer, of which 12–21% of the mass is carbohydrate (McClellan and Garner, 1980; Sanderink *et al.*, 1988). Aminopeptidases are widely distributed throughout the body, but those in the intestinal tract are important in completing the digestion of proteins and peptides, allowing absorption of amino acids. This final processing of oligopeptides occurs both

at the cell surface and by cytosolic aminopeptidases (Sterchi and Woodley, 1980).

Dipeptidyl peptidase IV (EC 3.4.14.5; DDP-IV) is expressed on the surfaces of most cells in the body; however, it is enriched in the microvillus fraction of the distal ileum (Norén *et al.*, 1980). DDP-IV is a membrane-bound glycoprotein, acting as a serine protease, cell-surface marker and binding molecule (Misumi and Ikehara, 2004). DDP-IV is produced as an 88 kDa single chain of 766 amino acids that must form homodimers for optimum proteolytic activity. DDP-IV cleaves dipeptides from the N-terminus of peptides and proteins with a preference for cleaving Xaa-Pro dipeptides from polypeptides consisting of about 30 amino acids. DPP-IV has a wide range of substrates, including growth factors, chemokines, neuropeptides and vasoactive peptides (Kahne *et al.*, 1999). DDP-IV degrades vasoactive intestinal peptide, gastrin-releasing peptide and neuropeptide Y (Lambeir *et al.*, 2001), proteins such as gelatin and collagen (Green *et al.*, 2006) and atrial natriuretic peptide but not brain natriuretic peptide (Brandt *et al.*, 2006). DPP-IV cleaves glucose-independent insulinotropic polypeptide and glucagon-like peptide-1, incretins that are critical for glucose metabolism. DPP-IV is a key factor in the modulation of immune responses in the gut and other tissues. It is identical to the T cell activation antigen CD26 and to the adenosine deaminase binding protein (Drucker, 2007). DDP-IV is implicated in medical conditions as diverse as diabetes mellitus, obesity, tumor growth, and HIV infection (Engel *et al.*, 2003) and celiac disease (Martí *et al.*, 2005). Gluten, a protein in wheat and related cereal grains, when digested by pepsin/trypsin, generates proline-rich peptides with 10–50 amino acid residues as well as free amino acids and small peptides. Several of these proline-rich peptides have been found to elicit an inflammatory response in the small intestine. Indeed, about 1% of the American population develops this autoimmune disease in which the villi of the small intestine are damaged and absorption impaired. Another 6% of the American population experience symptoms similar to celiac disease but lack the antibodies and intestinal damage characteristic of this autoimmune disease; these individuals are said to be gluten sensitive or intolerant. DDP-IV has been implicated as the rate-limiting enzyme in the further degradation of immunogenic peptides that activate T cell responses in celiac patients (Hausch *et al.*, 2002). The only effective management of celiac disease and gluten sensitivity to date is elimination of gluten in the diet (Penagini *et al.*, 2013).

Meprin A and B (EC 3.4.24.18 and EC 3.4.24.63 respectively; PABA peptide hydrolase) are zinc metalloproteases that are abundant in the normal human small intestine, especially the brush border of the ileum and the intestinal lumen. Meprins are expressed by enterocytes as a dimer of alpha subunits lacking a membrane anchor and beta subunits with an anchor, linked by disulfide bridges (Leuenberger *et al.*, 2003). The alpha/beta dimers (heteromeric meprin A) and the dimers of beta subunits (meprin B) are retained at the cell surface. Both heteromeric meprin A and meprin B concentrate on the villi of the ileal brush border; however, meprin B can be shed from the cell surface. In contrast, the beta subunit is not expressed in the colon and intracellular proteolytic removal of the membrane anchor from the alpha subunit results in its secretion as homomeric soluble meprin A. The differential

expression of meprin alpha and beta subunits directs the proteolytic activity of the alpha subunit either to the brush-border membrane in the ileum or to the lumen in the colon (Lottaz *et al.*, 1999). The meprin subunits are expressed as 90 kDa glycoproteins that form inactive dimeric proenzymes that must be activated by a serine protease such as trypsin. The meprins act at neutral to basic pH values, and have the ability to digest a wide range of substrates found in the human diet and shed from the intestinal lining. For example, meprin A and B degrade extracellular matrix proteins, basement membrane proteins, cell junction proteins, azocasein, procollagen, gastrin, protein kinases, chemokines, and growth factors, with distinctive as well as overlapping substrate specificities. In addition, meprin activity has been correlated with monocyte migration and impaired epithelial barrier function (Bao *et al.*, 2013). Meprin B preferentially cleaves peptide bonds with negatively charged side chains of aspartic acid and glutamic acid, whereas meprin A cleaves peptide bonds with neutral aliphatic and aromatic side chains (Sterchi *et al.*, 2008). There are about 10^8 bacteria/gram of distal ileum tissue, kept at this low level compared to the 10^{12} bacteria/gram of colon tissue (Canny and McCormick, 2008) by the sloughing of biofilm and flushing action of the flow of fluid through the intestinal lumen. MUC2 mucin in the mouse small intestine is cleaved at the N-terminal end by meprin B, releasing the previously attached mucin (Schütte *et al.*, 2014). Meprins also degrade type 1 pili of *Escherichia coli*, thereby impairing its ability to invade the intestinal epithelium (Vazeille *et al.*, 2011). Celiac disease patients undergo a shift in expression of meprin from the duodenal epithelial cells to the lamina propria leukocytes. This shift correlates with development of marked duodenal inflammation of the mucosa after ingestion of gluten-containing cereals. (Lottaz *et al.*, 2007).

Role of the Large Intestine in Proteolysis

By the time ingested food reaches the large intestine about 5–8 h after ingestion, the proteins have been extensively digested and the amino acids and a few oligopeptides absorbed, leaving a watery indigestible sludge. Water is reabsorbed from the lumen of the small intestine to concentrate the digestive wastes. In addition, the large intestine contains a large and diverse array of facultative aerobic and anaerobic microorganisms that digest the waste material and produce B vitamins, including biotin, and vitamin K, which are absorbed (Canny and McCormick, 2008). The microbiome, which accounts for about half of the dry weight of feces, also produces gas (flatulence) by bacterial fermentation of residual nutrients. Epithelial sodium channels (ENaC) are abundant in the polarized epithelial cells of the colon, and are critical for maintaining the balance of Na^+ and K^+ balance in the body. The intestinal metalloprotease meprin B binds to, and stimulates ENaC activity (García-Caballero *et al.*, 2011). Meprins have been implicated in a number of intestinal diseases. Decreased meprin-alpha expression is associated with intestinal inflammation in inflammatory bowel disease (IBD) patients and in a mouse experimental model of IBD. Genetic analysis corroborate that MEP1A is a susceptibility gene for ulcerative colitis (Banerjee *et al.*, 2009). Meprin A is overexpressed in

colorectal cancer cells, leading to the suggestion that the pro-migratory and pro-angiogenic activity of meprin A may promote tumor progression (Lottaz *et al.*, 2011). The meprin-beta subunit is not expressed in the mouse colon but meprin B is present in the effluent from the small intestine and from macrophages entering the colon wall and lumen. Human meprin B sheds membrane-bound cytokines and growth factors, and also enhances inflammation and tumor progression (Arolas *et al.*, 2012).

Summary

The first stage of protein digestion occurs in the mouth where food is macerated and the protein made available for subsequent proteolysis. The Recommended Dietary Allowance for protein intake is 0.8 g/kg or 2 ounces for a 155 pound adult (Campbell *et al.*, 2002). In the stomach, the nutrient paste is exposed to hydrochloric acid that denatures and precipitates the proteins making them vulnerable to attack by pepsin liberating oligopeptides. The site of most active proteolysis is in the duodenum where bile is added to the chyme facilitating dispersion of particulate material, and trypsin and chymotrypsin extensively cleave proteins and polypeptides into oligopeptides and amino acids. Trypsin and chymotrypsin have complementary substrate preferences with the former cleaving Arg-Leu and Lys-X with X being any amino acid except proline, and the latter cleaving the carboxyl side of amide bonds of large hydrophobic amino acids. Carboxylases and aminopeptidases complete the digestion of proteins into amino acids, dipeptides and tripeptides that are readily absorbed in the duodenum and jejunum.

Digestive proteolysis is quite rapid with half of food leaving the stomach within 30 min and absorption nearly complete within 2–4 h after intake (Adibi, 1976; Akimov and Bezuglov, 2012). Free amino acids in the intestinal lumen are absorbed by the appropriate group-specific transport system: (1) neutral amino acids, (2) dibasic amino acids and cysteine, and (3) dicarboxylic amino acids (Silk *et al.*, 1985). Absorption of free amino acids is greater in the jejunum than in the ileum, which is the site of bile salt and vitamin B₁₂ absorption. Abundant Peyer's patches link the ileum to the immune system, sometimes with adverse consequences (celiac disease and ulcerative colitis). The large intestine serves as the waste collection and processing chamber, with reclamation of water, sodium ions, and bile acids.

References

- Abbott, F., Gomez, J.E., Birnbaum, E.R., Darnall, D.W., 1975. The location of the calcium ion binding site in bovine alpha-trypsin and beta-trypsin using lanthanide ion probes. *Biochemistry* 14, 4935–4943.
- Adibi, S.A., 1976. Intestinal phase of protein assimilation in man. *American Journal of Clinical Nutrition* 29, 205–215.
- Akimov, M., Bezuglov, V., 2012. Methods of protein digestive stability assay – State of the art. In: Brzozowski, T. (Ed.), *New Advance in the Basic and Clinical Gastroenterology*. Rijeka, Croatia: Intech, pp. 211–234.
- Arolas, J.L., Broder, C., Jefferson, T., *et al.*, 2012. Structural basis for the sheddase function of human meprin β metalloproteinase at the plasma membrane. *Proceedings of the National Academy of Sciences United States of America* 109, 16131–16136.
- Banerjee, S., Oneda, B., Yap, L.M., *et al.*, 2009. MEP1A allele for meprin A metalloprotease is a susceptibility gene for inflammatory bowel disease. *Mucosal Immunology* 2, 220–231.
- Bao, J., Yura, R.E., Matter, G.L., *et al.*, 2013. Meprin A impairs epithelial barrier function, enhances monocyte migration and cleaves the tight-junction protein occludin. *American Journal of Physiology Renal Physiology* 305, F714–F726.
- Brandt, I., Lambeir, A.-M., Ketelslegers, J.-M., *et al.*, 2006. Dipeptidyl-peptidase IV converts intact B-type natriuretic peptide into its des-SerPro form. *Clinical Chemistry* 52, 82–87.
- Campbell, W.W., Trappe, T.A., Jozsi, A.C., *et al.*, 2002. Dietary protein adequacy and lower body versus whole body resistive training in older humans. *Journal of Physiology* 542, 631–642.
- Canny, G.O., McCormick, B.A., 2008. Bacteria in the intestine, helpful residents or enemies from within? *Infection and Immunity* 76, 3360–3373.
- Chabance, B., Marteau, P., Rambaud, J.C., *et al.*, 1998. Casein peptide release and passage to the blood in humans during digestion of milk or yogurt. *Biochimie* 80, 155–165.
- Cliffe, S.G., Grant, D.A., 1983. Calcium-binding of trypsin. A reassessment. *Biochemical Journal* 193, 655–658.
- Coll, M., Guasch, A., Avilés, F.X., Huber, R., 1991. Three-dimensional structure of porcine procarboxypeptidase B: A structural basis for its inactivity. *EMBO Journal* 10, 1–9.
- Craig, S.S., Mader, C., Bond, J.S., 1991. Immunohistological localization of the metalloprotease meprin in salivary glands of male and female mice. *Journal of Histochemistry and Cytochemistry* 39, 123–129.
- Davis, G.R., Sanata Ana, C.A., Morawski, S., Fordtran, J.S., 1980. Active chloride secretion in the normal jejunum. *Journal of Clinical Investigation* 66, 1326–1333.
- Drucker, D.J., 2007. Dipeptidyl peptidase-4 inhibition and the treatment of Type 2 diabetes: Preclinical biology and mechanisms of action. *Diabetes Care* 30, 1335–1343.
- Edgerton, M., Koshlukova, S.E., Araujo, M.W.B., *et al.*, 2000. Salivary histatin 5 and human neutrophil defensin 1 kill *Candida albicans* via shared pathways. *Antimicrobial Agents and Chemotherapy* 44, 3310–3316.
- Engel, M., Hoffmann, T., Wagner, L., *et al.*, 2003. The crystal structure of dipeptidyl peptidase IV (CD26) reveals its functional regulation and enzymatic mechanism. *Proceedings of the National Academy of Sciences United States of America* 100, 5063–5068.
- Ermund, A., Schütte, A., Johansson, M.E.V., Gustafsson, J.K., Hansson, G.C., 2013. Studies of mucus in mouse stomach, small intestine and colon. I. Gastrointestinal mucus layers have different properties depending on location as well as over Peyer's patches. *American Journal of Physiology Gastroenterology and Liver Physiology* 305, G341–G347.
- Freeman, H.J., Kim, Y.S., 1978. Digestion and absorption of proteins. *Annual Reviews in Medicine* 29, 99–116.
- Garcia-Caballero, A., Ishmael, S.S., Dang, Y., *et al.*, 2011. Activation of the epithelial sodium channel by the metalloprotease meprin β subunit. *Channels (Austin)* 5, 14–22.
- Green, B.D., Flatt, P.R., Bailey, C.J., 2006. Dipeptidyl peptidase IV (DPP IV) inhibitors: A newly emerging class for the treatment of type 2 diabetes. *Diabetes and Vascular Disease Research* 3, 159–165.
- Gritti, T., Banfi, G., Roi, G.S., 2000. Pepsinogens: Physiology, pharmacology, pathophysiology and exercise. *Pharmacological Reviews* 41, 265–281.
- Hausch, F., Shan, L., Santiago, N.A., Gray, G.M., Khosia, C., 2002. Intestinal digestive resistance of immunodominant gliadin peptides. *American Journal of Physiology Gastrointestinal and Liver Physiology* 283, G996–G1003.
- Hedstrom, L., Szilagyi, L., Rutter, W.J., 1992. Converting trypsin to chymotrypsin: The role of surface loops. *Science* 255, 1249–1253.
- Helmerhorst, E.J., Sun, X., Smith, W., Oppenheim, F.G., 2008. Identification of Lys-Pro-Gln as a novel cleavage site specificity of salivary-associated proteases. *Journal of Biological Chemistry* 283, 19957–19966.
- Holzinger, A., Maier, E.M., Bück, C., *et al.*, 2002. Mutations in the proenteropeptidase gene are the molecular cause of congenital enteropeptidase deficiency. *American Journal of Human Genetics* 70, 20–25.
- Hudaky, P., Kasik, G., Venekei, I., Gráf, L., 1999. The differential specificity of chymotrypsin A and B is determined by amino acid 226. *European Journal of Biochemistry* 259, 528–533.
- Johansson, M.E.V., Sjövall, H., Hansson, G.C., 2013. The gastrointestinal mucus system in health and disease. *Nature Reviews: Gastroenterology, Hepatology* 10, 352–361.
- Kahne, T., Lendeckel, U., Wrenger, S., *et al.*, 1999. Dipeptidyl peptidase IV: a cell surface peptidase involved in regulating T cell growth. *International Journal of Molecular Medicine* 4, 3–15.

- Kawashima, I., Tani, T., Shimoda, I., Takiguchi, Y., 1987. Characterization of pancreatic elastase II cDNAs: two elastase II mRNAs are expressed in human pancreas. *DNA* (Mary Ann Liebert, Inc.) 6, 163–172.
- Kolmer, M., Ord, T., Alhonen, L., *et al.*, 1991. Assignment of human prochymosin pseudogene to chromosome 1. *Genomics* 10, 496–498.
- Lambeir, A.M., Durinx, C., Proost, P., *et al.*, 2001. Kinetic study of the processing by dipeptidyl-peptidase IV/CD26 of neuropeptides involved in pancreatic insulin secretion. *FEBS Letters* 507, 327–330.
- Leuenberger, B., Hahn, D., Pischitzis, A., Hansen, M.K., Sterchi, E.E., 2003. Human meprin beta: O-linked glucans in the intervening region of the type I membrane protein protect the C-terminal region from proteolytic cleavage and diminish its secretion. *Biochemical Journal* 369, 659–665.
- Lottaz, D., Buri, C., Monteleone, G., *et al.*, 2007. Compartmentalised expression of meprin in small intestinal mucosa: Enhanced expression in lamina propria in coeliac disease. *Biological Chemistry* 388, 337–341.
- Lottaz, D., Hahn, D., Müller, S., Sterchi, E.E., 1999. Secretion of human meprin from intestinal epithelial cells depends on differential expression of the alpha and beta subunits. *European Journal of Biochemistry* 259, 496–504.
- Lottaz, D., Maurer, C.A., Noel, S., *et al.*, 2011. Enhanced activity of meprin-a, a pro-migratory and pro-angiogenic protease, in colorectal cancer. *PLoS One* 6 (11), e26450. Epub 2011 Nov 11. doi:10.1371/journal.pone.0026450.
- Lu, D., Yuan, X., Zheng, X., Sadler, J.E., 1997. Bovine proenteropeptidase is activated by trypsin, and the specificity of enteropeptidase depends on the heavy chain. *Journal of Biological Chemistry* 272, 31293–31300.
- Marti, T., Molberg, O., Li, Q., *et al.*, 2005. Prolyl endopeptidase-mediated destruction of T cell epitopes in whole gluten: Chemical and immunological characterization. *Journal of Pharmacology and Experimental Therapeutics* 312, 19–26.
- McClellan Jr., J.B., Garner, C.W., 1980. Purification and properties of human intestine alanine aminopeptidase. *Biochimica Biophysica Acta* 613, 160–167.
- Misumi, Y., Ikehara, Y., 2004. Dipeptidyl-peptidase IV. In: Barrett, A.J., Rawlings, N.D., Woessner, J.F. (Eds.), *Handbook of Proteolytic Enzymes*, Vol. 2, Cysteine, Serine and Threonine Peptidases. San Diego, CA: Elsevier Academic Press, pp. 1905–1909.
- Norén, O., Sjöström, H., Gudmand-Hoyer, E., Skovbjerg, H., Andersen, B., 1980. Peptidase activities in the functioning jejunum and ileum before and after jejunoileal bypass in morbid obesity. *Scandinavian Journal of Gastroenterology* 15, 825–832.
- Oudhoff, M.J., Kroeze, K.L., Nazmi, K., *et al.*, 2009. Structure-activity analysis of histatin, a potent wound-healing peptide from human saliva: Cyclization of histatin potentiates molar activity 1000-fold. *FASEB Journal* 23, 3928–3935.
- Penagini, F., Dilillo, D., Meneghin, F., *et al.*, 2013. Gluten-free diet in children: An approach to a nutritionally adequate and balanced diet. *Nutrients* 5, 4553–4565.
- Polgár, L., 2005. The catalytic triad of serine peptidases. *Cellular and Molecular Life Sciences* 62, 2161–2172.
- Rawlings, N.D., Salvesen, G., 2013. *Handbook of Proteolytic Enzymes*, third ed. New York, NY: Academic Press, pp. 4104.
- Reverter, D., Garcia-Sáez, I., Catasús, L., *et al.*, 1997. Characterisation and preliminary X-ray diffraction analysis of human pancreatic procarboxypeptidase A2. *FEBS Letters* 420, 7–10.
- Richter, C., Tanaka, T., Yada, R.Y., 1998. Mechanism of activation of the gastric aspartic proteinases: Pepsinogen, progastricsin and prochymosin. *Biochemical Journal* 335, 481–490.
- Rune, S.J., Viskum, K., 1969. Duodenal pH values in normal controls and in patients with duodenal ulcer. *Gut* 10, 569–571.
- Sanderink, G.-J., Artur, Y., Siest, G., 1988. Human aminopeptidase: A review of the literature. *Journal of Clinical Chemistry and Clinical Biochemistry* 26, 795–807.
- Schütte, A., Ermund, A., Becker-Paul, C., *et al.*, 2014. Microbial-induced meprin beta cleavage in MUC2 mucin and a functional CFTR channel are required to release anchored small intestinal mucus. *Proceedings of the National Academy of Science United States of America* 111, 12396–12401.
- Silk, D.B., Grimble, G.K., Rees, R.G., 1985. Protein digestion and amino acid and peptide absorption. *Proceedings of the Nutritional Society* 44, 63–72.
- Sterchi, E.E., Stöcker, W., Bond, J.S., 2008. Meprins, membrane-bound and secreted astacin metalloproteinases. *Molecular Aspects of Medicine* 29, 309–328.
- Sterchi, E.E., Woodley, J.F., 1980. Peptide hydrolases of the human small intestinal mucosa: Distribution of activities between brush border membranes and cytosol. *Clinica Chimica Acta* 102, 49–56.
- Sun, X., Smith, W., Oppenheim, F.G., Helmerhorst, E.J., 2009. Kinetics of histatin proteolysis in whole saliva and the effect on bioactive domains with metal-binding, antifungal, and wound-healing properties. *FASEB Journal* 23, 2691–2701.
- Szmola, R., Sahin-Tóth, M., 2007. Chymotrypsin C (caldecrin) promotes degradation of human cationic trypsin: Identity with Rinderknecht's enzyme Y. *Proceedings of the National Academy of Science United States of America* 104, 11227–11232.
- Vardakou, M., Mercuri, A., Barker, S.A., *et al.*, 2011. Achieving antral grinding forces in biorelevant *in vitro* models: Comparing the USP dissolution apparatus II and the dynamic gastric model with human *in vivo* data. *American Association of Pharmaceutical Sciences PharmSciTech* 12, 620–626.
- Vazeille, E., Bringer, M.A., Gardarin, A., *et al.*, 2011. Role of meprins to protect ileal mucosa of Crohn's disease patients from colonization by adherent-invasive *E. coli*. *PLoS One* 6 (6), e21199. doi:10.1371/journal.pone.0021199.
- Whitcomb, D.C., Lowe, M.E., 2007. Human pancreatic digestive enzymes. *Digestive Diseases and Science* 52, 1–17.
- Zakhary, G.M., Clark, R.M., Bidichandani, S.I., *et al.*, 2007. Acidic proline-rich Db and caries in young children. *Journal of Dental Research* 86, 1176–1180.
- Zamolodchikova, T.S., Vorotyntseva, T.I., Antonov, V.K., 1995. Duodenase, a new serine protease of unusual specificity from bovine duodenal mucosal. Purification and properties. *European Journal of Biochemistry* 227, 866–872.
- Zheng, X.L., Kitamoto, Y., Sadler, J.E., 2009. Enteropeptidase, a type II transmembrane serine protease. *Frontiers of Bioscience* 1, 242–249.

Relevant Websites

- <http://www.brenda-enzymes.info/>
BRENDA.
- <http://www.cafa.org/what-are-crohns-and-colitis/what-is-ulcerative-colitis/>
Crohn's & Colitis Foundation of America.
- <http://www.celiaccentral.org/celiac-disease/facts-and-figures/>
National Foundation for Celiac Awareness.
- <http://digestive.niddk.nih.gov/ddiseases/pubs/yrdd/>
National Institute of Diabetes and Digestive and Kidney Diseases.
- <http://digestive.niddk.nih.gov/ddiseases/pubs/celiac/>
National Institute of Diabetes and Digestive and Kidney Diseases.

PROTEIN SYNTHESIS/DEGRADATION: PROTEIN DEGRADATION – PATHOLOGICAL ASPECTS

Contents

Inhibitors of HIV Protease

Blood Pressure, Proteases and Inhibitors

Cancer – Proteases in the Progression and Metastasis

Lysosomal Diseases

Alpha-1-Antitrypsin Deficiency: A Misfolded Secretory Glycoprotein Damages the Liver by Proteotoxicity and Its Reduced Secretion Predisposes to Emphysematous Lung Disease Because of Protease-Inhibitor Imbalance

Inhibitors of HIV Protease

A Wlodawer, National Cancer Institute, Frederick, MD, USA

M Jaskolski, Adam Mickiewicz University, Poznan, Poland and Polish Academy of Sciences, Poznan, Poland

© 2016 Elsevier Inc. All rights reserved.

Glossary

Allosteric site A site other than a protein's (usually enzyme's) active site, which affects its activity when an effector molecule binds to it.

Aspartic protease An enzyme cleaving peptide bonds in proteins that has an active site consisting of two closely spaced, coplanar aspartate residues that activate a water molecule for the hydrolysis reaction. Enzymes belonging to the AA clan of aspartic peptidases, typified by pepsin found in gastric juice, are present in most organisms higher than bacteria, as well as in retroviruses, where they are called retlopepsins and function as symmetric homodimers.

Bioavailability The fraction of an administered dose of a drug that reaches the systemic circulation, usually less than 100% in oral administration.

Diastereomer A stereoisomer of an optically active compound, in which the configuration of one or several chiral centers has been inverted (e.g., from *R* to *S*). Diastereomers should not be confused with (a pair of) enantiomers, in which the configuration of all chiral centers has been inverted. Consequently, enantiomers are mirror symmetric while diastereomers are not, and for this reason display different physical and chemical properties.

HIV (-1,-2) Human immunodeficiency virus (type 1 or 2), the causative agent of AIDS. The most common and virulent variant is designated HIV-1, whereas HIV-2 is more closely related to simian immunodeficiency virus and is less virulent.

HTLV Human T-cell leukemia virus, the first retrovirus to be associated with human disease (lymph node leukemia).

Integrase A virally encoded enzyme responsible for the incorporation of the double-stranded DNA copy of the retroviral genome into host cell genome. After integration, the infection is permanent and the infected cell cannot be cured.

Lipodystrophy A pathological condition characterized by abnormal or degenerative conditions or distribution of the body's fat.

M-PMV Mason–Pfizer monkey virus, a retrovirus similar to HIV, causing acquired immunodeficiency syndrome (AIDS)-like syndrome in macaque (rhesus) monkeys.

Nucleophile A chemical entity (such as a water molecule or hydroxyl group) rich in electrons, that is capable (usually after additional activation) of attacking another chemical group (such as the C atom of a peptide bond (O=C–N) with depleted electron density, during a chemical transformation, such as hydrolysis.

Pepstatin A universal inhibitor of aspartic proteases, originally isolated from cultures of various species of *Actinomyces*.

Retrovirus A virus with a single-stranded mRNA genome, which is reverse-transcribed into DNA in a process that inverts the normal flow (DNA→RNA) of genetic information.

Reverse transcriptase (RT) A virally encoded enzyme that produces a double-stranded DNA copy of a retroviral RNA genome. RT is an RNA-dependent DNA polymerase, with ribonuclease activity.

RSV Rous sarcoma virus, a retrovirus causing sarcoma in chickens. The first retrovirus to be described, over a century ago, by Peyton Rous. The currently preferred name is avian sarcoma virus (ASV).

Introduction

The emergence of the acquired immunodeficiency syndrome (AIDS) epidemic in the early 1980s and the subsequent identification of the human immunodeficiency virus (HIV) as its causative agent brought into focus the need to accelerate research on retroviruses, in order to assist the efforts to create anti-HIV drugs. No such drugs were available during the decade after the first reports of the new disease started appearing and initially the disease itself was invariably lethal since the retrovirus, which infects T4 leukocytes, devastates the immune system and leads to its complete failure. With lack of any treatment options, the outlook was obviously very grim.

Retroviruses have been known for over a century, since the identification of an infectious agent causing cancer in chickens, later named after its discoverer Rous sarcoma virus (RSV) (Rous, 1911). However, the exact nature of the agent, of its life cycle and the mode of infectivity were not established until much later. It is now known that the genetic material of retroviruses consists of single-stranded RNA of positive polarity (mRNA) that becomes transcribed into DNA in a reverse (or retro) transcription reaction that is catalyzed by a retrovirus-specific and virally encoded reverse transcriptase (RT). Another retroviral enzyme, integrase (IN), incorporates the resulting double-stranded viral DNA into the host genome, thus making the genetic material of the retrovirus (provirus) a permanent part of the infected cell.

One of the unusual characteristics of many viruses, including all retroviruses, is that their proteins are not translated as individual final units, but rather as one or more large polypeptides that need to be processed (cleaved) into the mature viral enzymes and structural proteins. In retroviruses, the enzyme responsible for such an activity is retroviral protease (PR). Analyzing the retrotransposon and retroviral genomic sequences, including human T-cell leukemia virus (HTLV) (closely related to HIV) and RSV, Toh *et al.* (1985) found a single copy of a signature sequence D(S/T)G (aspartate, serine or threonine, and glycine) in the translated proteins that could be matched with the active-site motif of aspartic proteases from the pepsin family. Pepsin and similar cell-derived aspartic proteases, however, contain this motif in two copies in their pseudo-twofold-symmetric active site, each contributed by a separate domain of a single polypeptide chain. To reconcile these puzzling observations, Toh *et al.* postulated the presence of a pepsin-like protease in retroviruses, with the caveat that the retroviral proteases would need to be symmetric homodimers composed of two identical subunits. This hypothesis was consistent with an earlier speculation by Tang *et al.* (1978) that cell-derived aspartic proteases have arisen from smaller proteins by gene duplication and divergent evolution. The postulate about the homodimeric nature of retroviral proteases (retropepsins) was experimentally confirmed when the crystal structure of RSV protease was determined (Miller *et al.*, 1989a).

Without a functional protease (which could be inactivated by mutation or by inhibitors) the retrovirus is still capable of replicating, but only immature viral particles can be formed, which are noninfectious, i.e., cannot infect other T4 cells or other patients. Although inhibition of HIV

PR does not cure the infection *per se*, it can be hoped that with prolonged treatment, inhibition of HIV PR will allow sufficient time for the elimination of the pool of the infected T4 cells.

Structure and Enzymatic Mechanism of HIV Protease

The structure of HIV-1 (type 1) PR itself, in unliganded state, was independently determined in 1989 by three groups (Navia *et al.*, 1989; Wlodawer *et al.*, 1989; Lapatto *et al.*, 1989). As in the case of RSV PR, HIV-1 PR is also a homodimer of two protein chains, 99 residues each. The protein fold of one subunit topologically resembles a truncated version of a single domain of pepsin, with the inter-subunit interface formed by four intertwined β strands from all termini (residues 1–4 and 95–99 from each subunit). The N-terminal strand is followed by two more β strands comprising residues 9–15 and 18–25, the latter including the catalytic Asp25. The next strand consists of residues 30–35 and is followed by a broad loop 36–42. Two more β strands, 43–49 and 52–58 form a flexible ‘flap’ loop that acts as a gating element for the active site. The second half of the HIV-1 PR subunit is topologically related to the first half and consists of β strands 52–66, 69–78, and 83–85. A prominent helix comprising residues 86–94 (absent in the N-terminal fold) has a clear counterpart in pepsins.

In the catalytic mechanism of aspartic proteases, a nucleophilic water molecule is activated by hydrogen bonding between the two (hemiprotonated) aspartate carboxylic groups. Such a catalytic water molecule was also found in the active sites of RSV PR (Miller *et al.*, 1989a) and HIV-1 PR (Wlodawer *et al.*, 1989) and its involvement in the enzymatic mechanism was gleaned from structural data (Jaskólski *et al.*, 1991).

Inhibitors of Retroviral Proteases

Strategies aimed at deactivation of HIV PR are focusing on relatively small chemical molecules that could inhibit its catalytic function. In a broad classification, the inhibitors can be covalent (irreversible) or non-covalent (reversible). Covalent inhibitors, such as molecules with a reactive oxirane (epoxy) group (Ro *et al.*, 1999) that attaches itself to the active-site aspartate, are not good candidates, even though they would block the enzyme permanently, because they can also modify many important host proteins. Non-covalent inhibitors typically compete with substrates for the active site, but bind more strongly and are resistant to cleavage. In the simplest approach, a competitive inhibitor is designed using as template the amino acid sequence of a good peptidic substrate (with some optimization), and replacing the scissile bond by a non-cleavable surrogate. Non-peptidic inhibitors use a similar strategy but the sequences of chemical moieties designed for docking in the respective enzyme binding sites do not have amino acid or peptidic character.

Whereas the structures of ligand-free retroviral proteases were crucial for delineating their structural and evolutionary relationship to pepsins, they were by themselves not sufficient to guide the design of potent inhibitors of HIV-1 PR.

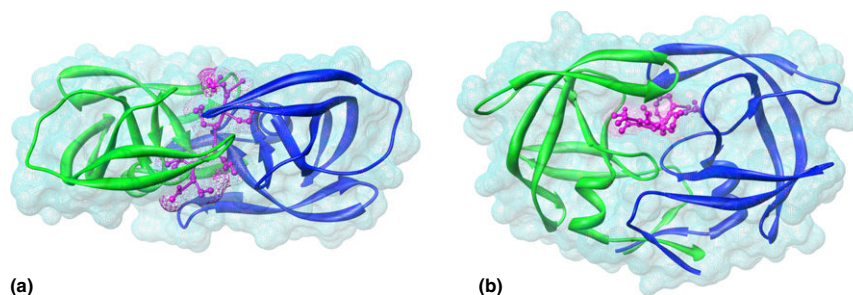


Figure 1 Two canonical views of HIV-1 PR complexed with MVT-101, the first inhibitor cocrystallized with this enzyme. The main-chain traces of the two subunits of the enzyme are colored green and blue, and the ball-and-stick model of the inhibitor is in magenta. The molecular surface of the protein is shown in light blue, and the surface of the inhibitor in pink. In (a) the homodimeric enzyme is viewed along its twofold axis. In (b) this axis is vertical. Figure prepared by Dr. Jiri Vondrasek.

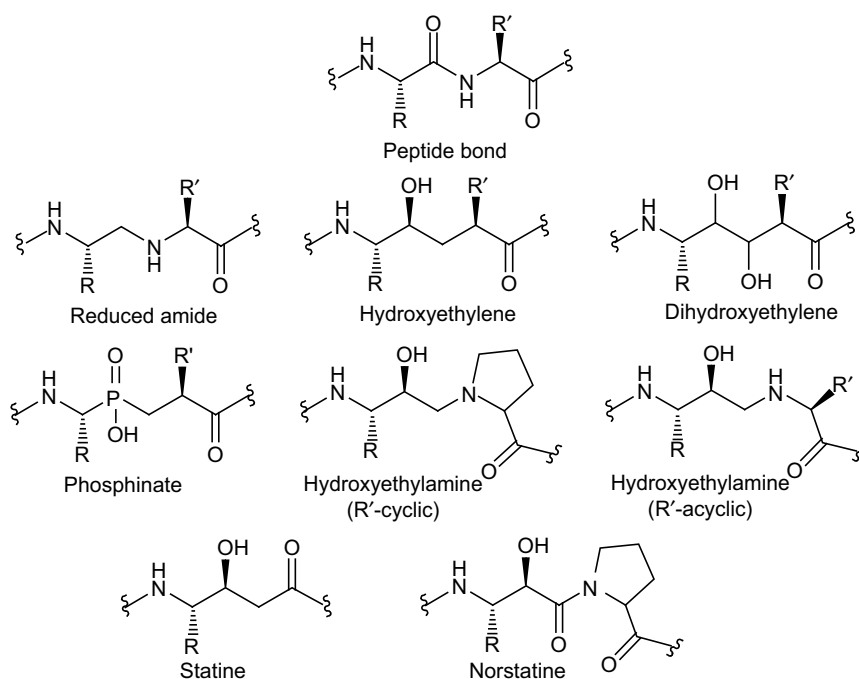


Figure 2 Chemical formulas of the peptide bond and its selected non-cleavable replacements.

Nevertheless, a model of an inhibitor complex was created on the basis of the structure of the apoenzyme of RSV PR superimposed on rhizopuspepsin, a fungal pepsin-like enzyme, in complex with a peptidic inhibitor with a reduced peptide replacing the cleavage site, His-Pro-Phe-His-Phe- ψ [CH-NH]Phe-Val-Tyr (where ψ denotes the reduced bond between the two Phe residues). Analysis of this model allowed successful delineation of seven subsites for the interactions between the side chains of the inhibitor and HIV-1 PR (Weber *et al.*, 1989).

Progress toward full understanding of the details of the interactions between the inhibitors of HIV-1 PR and the enzyme was boosted by the elucidation of the crystal structures of numerous complexes, now most likely going into thousands (although it is not possible to ascertain the exact count). The first such structure involved a complex with MVT-101 (Miller *et al.*, 1989b), a peptidic inhibitor created on the basis of a good natural substrate of HIV-1 PR with the sequence Ac-Thr-Ile-Met-Met-Gln-Arg.amide (K_m 1.4 mM) (Figure 1). In an

approach established previously for the synthesis of peptidic inhibitors of human renin (which is a pepsin-like aspartic protease involved in the blood pressure control cascade), both methionines in this inhibitor were replaced by norleucine isosteres, and the scissile peptide between them was replaced by a reduced analog. The inhibition constant K_i for MVT-101 was 0.78 μ M, indicating that this compound acted as a comparatively weak inhibitor and could not be considered as a good drug candidate, although it was powerful enough to form a stable complex with the protease during crystallization experiments.

Binding of MVT-101 led to quite substantial rearrangements of the structure of the enzyme, with the tips of the flaps, locked over the active-site-bound inhibitor, moving as much as 7 Å away from their positions in the free protease (where they were most likely fixed by crystal packing). In addition, in the complex with MVT-101 and all other peptidic inhibitors, a tightly bound water molecule was found at the interface

between the locked flap arms and the central part of the bound inhibitor, in variance with analogous pepsin-like complexes, where essentially only one flap is long enough to gate access to the active site. Although the HIV PR dimer is symmetric in the absence of a ligand with unique directionality, its binding induces some asymmetry which, if not transmitted to the surface during crystallization, leads to apparent twofold orientational disorder of the inhibitor that was noted in a number of structures. This binding, in turn, allowed delineation of the pockets that would accommodate the side chains of MVT-101. In the convention of [Schechter and Berger \(1967\)](#), in which the N-terminal (N-Pn...P1-) and C-terminal (-P1'...Pn'-C) substrate/inhibitor residues (linked by the P1-P1' scissile bond/analog) are docked in the corresponding Sn...S1...S1'...Sn' subsites of the enzyme, subsite S3 of HIV-1 PR included Arg8', Asp29', and Gly48'; subsite S2 was lined by Ala28', Ile47', Ile50', and Ile84'; and subsite S1 included Leu23', Asp25', Ile50', Pro81', Val82', and Ile84'. (Primed residues refer to the second subunit of the homodimer.) The subsites on the other side of the non-scissile P1-P1' bond were generally similar with few exceptions. Subsite S1' included Leu23', Gly27', Asp25', Ile50', Pro81', Val82', and Ile84'; subsite S2' consisted of Val32', Ile47', Gly48', and Ile50'; and subsite S3' was surrounded by Arg8', Gly27', Asp29', Gly48', and Val82'.

A large number of structures of complexes with other peptide-based inhibitors that included non-scissile peptide mimetics of different chemistry ([Figure 2](#)) followed, providing a wealth of information about the plasticity of the enzyme subsites. The structure of a complex with acetyl-pepstatin, in which the scissile bond was replaced by the unnatural amino acid statine, elucidated the binding of this standard inhibitor of cell-derived aspartic proteases to HIV-1 PR ([Fitzgerald et al., 1990](#)). Another early structure included a complex with JG-365, a hydroxyethylamine-containing peptide analog ([Swain et al., 1990](#)). The structural data on a large number of inhibitors and inhibitor complexes of HIV PR were previously summarized in detailed reviews ([Wlodawer and Erickson, 1993](#); [Fitzgerald, 1993](#); [Wlodawer and Vondrasek, 1998](#); [Qiu and Liu, 2011](#)).

First generation of HIV-1 PR Inhibitors Approved as AIDS Drugs

A complete summary of the status of the inhibitors of HIV-1 PR that have been approved for clinical use by the food and drug administration (FDA) can be found in Wikipedia. The first such inhibitor, approved on 6 December 1995, was saquinavir ([Figure 3](#)), developed by Hoffmann-La Roche. This compound, originally designated as Ro 31-8959, was the subject of intensive biochemical, biological, and structural studies ([Craig et al., 1991](#)). It has a molecular weight (MW) of 671. In common with JG-365, saquinavir utilizes hydroxyethylamine as a non-cleavable peptide isostere. However, early crystallographic studies ([Krohn et al., 1991](#)) suggested that the more potent version of this compound should be the *S* diastereomer of the -C*(OH)- chiral center, rather than the *R* form found in JG-365. Another major difference between these two compounds is the replacement of Pro, located in the P1' position of JG-365, by DIQ ((*S,S,S*)-decahydroisoquinoline-3-carbonyl) in saquinavir. The structure of a saquinavir

complex reported in 1991 ([Krohn et al., 1991](#)) was deposited in the Protein Data Bank only in 1996 and released in 1997 (PDB ID 1HXB), but a number of other structures using the native and mutant forms of HIV-1 PR became available since then, including the atomic-resolution (0.97 Å) structure ([Figure 4\(a\)](#)) of a complex with the V82A mutant (PDB ID 2NMZ; ([Tie et al., 2007](#))). The main problem with the original formulation of the drug (sold under the name of Invirase) was poor bioavailability, as low as 3–4%. However, the later re-formulation (brand name Fortovase) increased the bioavailability very significantly.

Three more inhibitors of HIV-1 PR gained FDA approval in 1996 and 1997 ([Wlodawer and Vondrasek, 1998](#)). Abbott Laboratories (now AbbVie) developed ritonavir (brand name Norvir), a relatively large (MW 721) peptidomimetic inhibitor ([Figure 3](#)) that was designed as a result of testing the concept of making the inhibitors fully, or nearly, *C*₂-symmetric. Even though ultimately such symmetric inhibitors turned out to be more difficult to manipulate for improved solubility and bioavailability, the concept resulted in the inclusion of a -CH₂-C(OH)- group as a non-cleavable linker between two phenylalanines that occupy the S1 and S1' subsites of the enzyme. Ritonavir has much better bioavailability than saquinavir, but it quickly became apparent that it had an unanticipated ability to act as a very potent inhibitor of cytochrome P450 (particularly of its isoform Cyp3A4), slowing down the removal of other pharmaceutical agents from circulation ([Lea and Faulds, 1996](#)). For that reason, ritonavir is currently used only as a booster for other protease inhibitors (see below).

Another peptidomimetic inhibitor of HIV-1 PR is indinavir ([Figure 3](#); brand name Crixivan), developed by Merck ([Vacca et al., 1994](#)). Slightly smaller (MW 614) than either saquinavir or ritonavir, it utilizes a hydroxyethylamine insert as a non-cleavable group. The compound is highly selective for HIV-1 PR (K_i 0.52 nM) and HIV-2 PR (K_i 3.3 nM) while it shows no inhibitory activity against any mammalian (including human) aspartic proteases.

An inhibitor that finally shed any peptidic character is viracept ([Figure 3](#); brand name Nelfinavir). Although this drug utilizes the same DIQ group at its 'C terminus' as saquinavir, it does not retain any actual peptide bonds. Viracept was developed by Agouron Pharmaceuticals ([Kaldor et al., 1997](#)), a company, now part of Pfizer, especially created with the task of introducing structure-based drug design ([Appelt et al., 1991](#)). Viracept is smaller (MW 568) than the previous approved inhibitors of HIV-1 PR and was the first protease inhibitor approved for the treatment of pediatric AIDS.

The next inhibitor to reach the market, amprenavir ([Figure 3](#); brand name Agenerase), was originally discovered at Vertex Pharmaceuticals, another venture capital company explicitly created to apply rational structure-based drug design in practice. The company was later acquired by GlaxoSmithKline, which became the distributor of this drug. Amprenavir is smaller still (MW 506) than the previous drugs but it was withdrawn from the market in 2004 after introduction of its prodrug form, fosamprenavir (brand name Lexiva (US) or Telzir (Europe)), in which a main-chain hydroxyl group was modified by phosphorylation. The prodrug is converted to amprenavir by host enzymes, thus slowing the release (and excretion) of the active form.

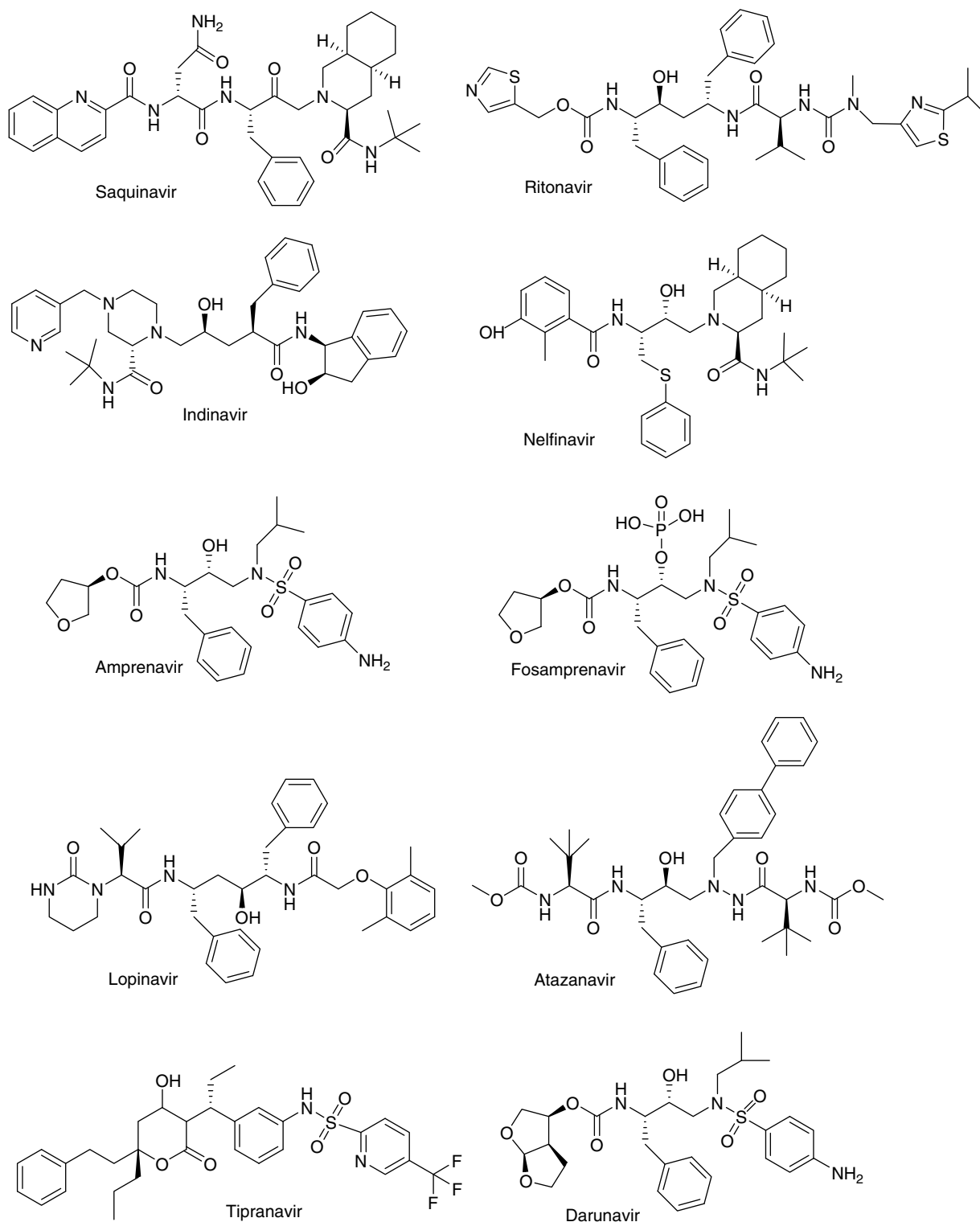


Figure 3 Chemical formulas of the 10 FDA-approved inhibitors of HIV-1 PR.

Emergence of Resistance to Clinical Inhibitors of HIV-1 PR

Since the RT enzymes of retroviruses lack proofreading mechanism, the error rates in transcription are very high, estimated at

1:10 000 for HIV-1 (Katz and Skalka, 1990). Such an elevated level of mutations indicates that every possible variant of the retroviral genome has already been transcribed in an infected patient treated with protease inhibitors (or, for that matter, with any other anti-AIDS therapy). Thus, development of resistance is

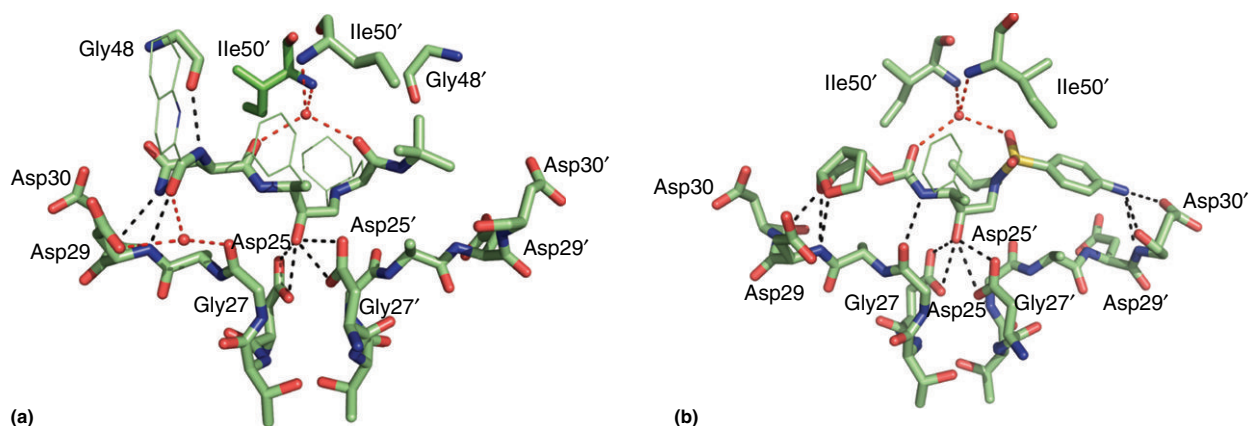


Figure 4 Hydrogen-bond interactions between HIV-1 PR and selected inhibitors. Hydrogen bonds are indicated by dashed lines. Interactions mediated by water molecules are shown in red. (a) Hydrogen-bond interactions between HIV-1 PR and saquinavir, based on the atomic-resolution structure of the complex (PDB ID 2NMZ). (b) Hydrogen-bond interactions between HIV-1 PR and darunavir, seen in an atomic-resolution crystal structure (PDB ID 2HS1). Figure prepared by Dr. Jerry Alexandratos.

just a matter of days, since what is required is only selection of preexisting variants of the virus, rather than mutagenesis subsequent to exposure of the virus to challenging drugs. This problem was noticed early on and became clearly evident in the early clinical studies. Indeed, [Condra *et al.* \(1995\)](#) showed that patients subjected to experimental monotherapy using only indinavir, developed resistance not only to this particular inhibitor, but also to other drugs that inhibit HIV-1 PR, including saquinavir, amprenavir, and ritonavir. The authors stated that “These observations reflect the common nature of the inhibitors’ target and suggest that combination therapy with multiple protease inhibitors may yield multiply resistant variants. In addition, initial therapy with any one inhibitor may limit the benefit of subsequent treatment with the remaining compounds.” It is significant that these observations were published even before the approval of the first of these drugs by the FDA, thus the need for multidrug therapy combining different classes of targets (such as RT and IN) and for the development of novel protease inhibitors that would be less effective in eliciting viral resistance, was apparent already early on.

One of the most common mutations in HIV-1 PR exposed to the first-generation inhibitors is replacement of Val82 by either threonine or alanine. The structure of a complex of HIV-1 PR with ritonavir was used at Abbott to guide synthesis of a new inhibitor, lopinavir ([Figure 3](#); [Sham *et al.*, 1998](#); [Hurst and Faulds, 2000](#)) that would not interact with residue 82 at all despite its comparatively larger size (MW 629 – still smaller, though, than the parent compound ritonavir). The problem with lopinavir, however, was its very short half-life in plasma, no longer than 2–4 h. However, combination of therapeutic amounts of lopinavir with subtherapeutic quantities of ritonavir extended the half-life of the drug very significantly. The form of lopinavir approved by the FDA is such a combination (brand name Kaletra).

Atazanavir ([Figure 3](#); brand name Reyataz), developed at Bristol-Myers Squibb, is another FDA-approved inhibitor of predominantly peptidic character ([Robinson *et al.*, 2000](#)), but with unexpectedly long half-life in plasma, allowing for once-a-day dosing. Although not explicitly designed to combat resistance, it is nevertheless quite successful in this regard, especially if

boosted with ritonavir. An advantage of atazanavir over other inhibitors of HIV-1 PR is that it causes fewer side effects, in particular it is less likely to cause lipodystrophy and elevate cholesterol levels.

Tipranavir ([Figure 3](#); brand name Aptivus), developed at Boehringer-Ingelheim, is a non-peptidic protease inhibitor originally synthesized at Upjohn as part of a structure-based discovery program ([Thaisrivongs and Strohbach, 1999](#)). It is administered in combination therapy with ritonavir to treat particularly difficult HIV infections. Tipranavir is successful in inhibiting the replication of viruses that are resistant to other protease inhibitors and it is recommended for patients not responding to other treatments. Development of resistance to tipranavir itself requires several simultaneous mutations of the HIV-1 PR sequence. In one study, as many as 10 mutations were required to decrease tipranavir’s inhibitory potency by two orders of magnitude ([Doyon *et al.*, 2005](#)).

The most recent among the FDA-approved inhibitors of HIV-1 PR and most often used in current clinical practice is darunavir ([Figure 3](#); brand name Prezista), developed at Tibotec/Johnson and Johnson. It is a non-peptidic compound with MW 548 and the K_d value of 4.5 pM, the latter significantly lower than for other inhibitors of HIV-1 PR ([Koh *et al.*, 2003](#)). It is also commonly administered in combination with ritonavir, in order to increase its plasma half-life. The hydrogen bonded interactions between darunavir and HIV-1 PR are shown in [Figure 4\(b\)](#).

Roads Not Taken – Inhibitors of HIV-1 PR That Did Not Become Drugs

Whereas the clinically relevant inhibitors of HIV-1 PR represent a triumph of rational drug design, many other compounds have been created as either intermediate steps of their development, or as testing grounds for the utilization of innovative chemical principles. Thus, for example, many fully twofold symmetric inhibitors were found to be quite potent, although they never became drugs ([Kempf *et al.*, 1994](#)). Some of the symmetric designs included an additional retropepsin-specific feature, namely a mimic of the water molecule that mediates the interactions

between the flap loops and the central region of the peptidic inhibitors or substrates. A way to accomplish this dual goal was to utilize cyclic urea derivatives as the central part of the inhibitors. Several very potent inhibitors, such as DMP-450 and DMP-323 resulted from these approaches, but for a variety of reasons, some of them commercial rather than scientific, they did not progress to become clinical drugs (Lam *et al.*, 1994).

One of the earliest attempts to use structural information for rational design of HIV-1 PR inhibitors was a computational experiment that led to the identification of haloperidol, a compound related to a known antipsychotic agent bromoperidol, as a potential inhibitor of this enzyme (DesJarlais *et al.*, 1990). This was an interesting early discovery of a non-peptidic compound which, however, did not lead to its practical application as an antiviral agent, mainly because the required concentration of haloperidol would exceed by several orders of magnitude the doses of bromoperidol administered in psychiatric treatment.

Considerable efforts went into modification of the successful inhibitors of HIV-1 PR that would lead to easier and more efficient synthesis. An example is provided by the modification of amprenavir, in which the non-cleavable isostere of the peptide bond was replaced by 1,2,3-triazole. Such compounds could be produced by fast and high-yield azide-alkyne click chemistry and were shown to be excellent inhibitors of both the native and mutant enzymes (Brik *et al.*, 2005).

A very attractive chemical procedure that used solid-state synthesis of 4-aryl-1,4-dihydropyridines as precursors of C₂-symmetric cage compounds (Hilgeroth *et al.*, 1999, 2002) with inhibitory properties against HIV PR, has never been developed beyond preliminary tests. Also other novel compounds that were not only non-peptidic, but even non-carbon, called carboranes, have not been developed in practice, although they appeared promising in preliminary studies (Cigler *et al.*, 2005).

Where Will the HIV-1 PR Inhibitors Be in the Future?

Detailed studies of not only inhibitor complexes cocrystallized with active HIV-1 PR, but also of substrate peptides bound to inactive enzyme (usually variants with the catalytic Asp25 mutated to asparagine) led to the definition of the concept of 'substrate envelope' – the minimum volume occupied by parts of the majority of efficiently processed substrates of the enzyme (Prabu-Jeyabalan *et al.*, 2002). The postulate was that an inhibitor that does not extend much beyond such an envelope is unlikely to be resistance-prone, since any mutations interfering with its binding would also prevent efficient processing of legitimate substrates. Based on this principle, a number of new inhibitors with nanomolar potencies against patient-derived wild-type viruses from HIV-1 clades A, B, and C, as well as their multidrug resistance (MDR) variants, were synthesized and tested (Parai *et al.*, 2012). In that work, introduction of cyclic and acyclic P2 carbamates with multiple hydrogen-bond donor and acceptor centers resulted in the creation of highly potent inhibitors of HIV-1 PR.

Another important concept applicable to the design of 'resistance-resistant' inhibitors is targeting the protein backbone. It was speculated that inhibitors that maximize their interactions in the HIV-1 PR active site, particularly through multiple hydrogen bonds with the protein backbone of wild-type HIV-1

protease, would maintain those contacts with mutant proteases as well. Targeting the sequence-independent protein backbone rather than the sequence-specific side chains, should therefore curtail the emergence of drug-resistant variants of HIV, as mutations that alter the backbone conformation would most likely reduce or abrogate the catalytic capacity as well (Ghosh *et al.*, 2012).

Fragment screening is a method of searching for compounds capable of even weak binding to the target proteins, with the aim of ultimately increasing their avidity through connecting two or more such 'fragments' covalently. This approach has also been applied to HIV PR, mostly with the aim of finding molecules that would bind to allosteric sites distant from the active site, interfering with such obligatory property as dimerization. Binding of components of suitable cocktails (i.e., chemical libraries) of such test compounds could be monitored by crystallography or, for example, by surface plasmon resonance. Although these studies yielded some interesting preliminary results, they did not lead so far to practical applications (Bauman *et al.*, 2012).

Most of the approaches to inhibitor design treat the target enzyme as a rigid object. However, consideration of the inherent flexibility of protein chains is becoming increasingly important and has already been utilized in the design of inhibitors of HIV PR (Lukman *et al.*, 2014). Another novel concept involved replacement of some non-exchangeable hydrogen atoms in the inhibitor molecule by deuterium. Deuterated compounds are metabolized more slowly, thus allowing prolonged maintenance of their effective blood levels. Such an approach was tested on a modification of atazanavir named CTP-518 (GlaxoSmithKline/Concert Pharmaceuticals). Whereas preclinical studies demonstrated that this modification retained the antiviral potency and increased the half-life and plasma trough levels, the results of Phase I clinical trials, initiated in 2009, have not been reported.

Interference with dimerization is a very attractive approach to HIV PR inhibition because it would prevent the formation of a viable enzyme in the first place (Boggetto and Reboud-Ravaux, 2002). Dimerization inhibitors would have to be extremely potent because the dissociation constant of HIV-1 PR is very low, which also hinders structural studies of this protein in monomeric form. The latter obstacle could be circumvented by looking at similar enzymes from other retroviruses (such as Mason-Pfizer monkey virus (M-PMV)), which can be crystallized as monomers (Khatib *et al.*, 2011; Gilski *et al.*, 2011). Interestingly, it appears that darunavir is capable of acting not only as a very potent competitive inhibitor of HIV-1 PR, but is also able to interfere with the dimerization equilibrium of the enzyme (Hayashi *et al.*, 2014).

Whereas it is not certain which of the novel approaches to the design of inhibitors of HIV-1 PR will lead to new and more potent AIDS drugs, the multitude of experimental approaches bodes well for the future success in creating compounds with novel mode of action and lower potential for triggering resistance. The last word has clearly not yet been said.

See also: Intracellular Infectiology: Infectious Agents: HIV – The Cell Biology of Virus Infection and Replication

References

- Appelt, K., Bacquet, R.J., Bartlett, C.A., *et al.*, 1991. Design of enzyme inhibitors using iterative protein crystallographic analysis. *Journal of Medicinal Chemistry* 34, 1925–1934.
- Bauman, J.D., Patel, D., Arnold, E., 2012. Fragment screening and HIV therapeutics. *Topics in Current Chemistry* 317, 181–200.
- Boggetto, N., Reboud-Ravaux, M., 2002. Dimerization inhibitors of HIV-1 protease. *Biological Chemistry* 383, 1321–1324.
- Brik, A., Alexandratos, J., Lin, Y.C., *et al.*, 2005. 1,2,3-triazole as a peptide surrogate in the rapid synthesis of HIV-1 protease inhibitors. *ChemBioChem* 6, 1167–1169.
- Cigler, P., Kozisek, M., Rezacova, P., *et al.*, 2005. From nonpeptide toward noncarbon protease inhibitors: Metallacarboranes as specific and potent inhibitors of HIV protease. *Proceedings of the National Academy of Sciences of the United States of America* 102, 15394–15399.
- Condra, J.H., Schleif, W.A., Blahy, O.M., *et al.*, 1995. *In vivo* emergence of HIV-1 variants resistant to multiple protease inhibitors. *Nature* 374, 569–571.
- Craig, J.C., Duncan, I.B., Hockley, D., *et al.*, 1991. Antiviral properties of Ro 31-8959, an inhibitor of human immunodeficiency virus (HIV) proteinase. *Antiviral Research* 16, 295–305.
- DesJarlais, R.L., Seibel, G.L., Kuntz, I.D., *et al.*, 1990. Structure-based design of nonpeptide inhibitors specific for the human immunodeficiency virus 1 protease. *Proceedings of the National Academy of Sciences of the United States of America* 87, 6644–6648.
- Doyon, L., Tremblay, S., Bourgon, L., Wardrop, E., Cordingley, M.G., 2005. Selection and characterization of HIV-1 showing reduced susceptibility to the non-peptidic protease inhibitor tipranavir. *Antiviral Research* 68, 27–35.
- Fitzgerald, P.M.D., 1993. HIV protease-ligand complexes. *Current Opinion in Structural Biology* 3, 868–874.
- Fitzgerald, P.M.D., McKeever, B.M., VanMiddlesworth, J.F., *et al.*, 1990. Crystallographic analysis of a complex between human immunodeficiency virus type 1 protease and acetyl-pepstatin at 2.0 Å resolution. *Journal of Biological Chemistry* 265, 14209–14219.
- Ghosh, A.K., Anderson, D.D., Weber, I.T., Mitsuya, H., 2012. Enhancing protein backbone binding – A fruitful concept for combating drug-resistant HIV. *Angewandte Chemie International Edition in English* 51, 1778–1802.
- Gilski, M., Kazmierczyk, M., Krzywda, S., *et al.*, 2011. High-resolution structure of a retroviral protease folded as a monomer. *Acta Crystallographica D67*, 907–914.
- Hayashi, H., Takamune, N., Nirasawa, T., *et al.*, 2014. Dimerization of HIV-1 protease occurs through two steps relating to the mechanism of protease dimerization inhibition by darunavir. *Proceedings of the National Academy of Sciences of the United States of America* 111, 12234–12239.
- Hilgeroth, A., Tykarska, E., Jaskolski, M., 2002. Crystal structure of a novel synthetic inhibitor of HIV-1 protease. *Journal of Molecular Structure* 605, 63–70.
- Hilgeroth, A., Wiese, M., Billich, A., 1999. Synthesis and biological evaluation of the first N-alkyl cage dimeric 4-aryl-1,4-dihydropyridines as novel nonpeptidic HIV-1 protease inhibitors. *Journal of Medicinal Chemistry* 42, 4729–4732.
- Hurst, M., Faulds, D., 2000. Lopinavir. *Drugs* 60, 1371–1379.
- Jaskólski, M., Tomasselli, A.G., Sawyer, T.K., *et al.*, 1991. Structure at 2.5 Å resolution of chemically synthesized human immunodeficiency virus type 1 protease complexed with a hydroxyethylene-based inhibitor. *Biochemistry* 30, 1600–1609.
- Kaldor, S.W., Kalish, V.J., Davies, J.F., *et al.*, 1997. Viracept (nelfinavir mesylate, AG1343): A potent, orally bioavailable inhibitor of HIV-1 protease. *Journal of Medicinal Chemistry* 40, 3979–3985.
- Katz, R.A., Skalka, A.M., 1990. Generation of diversity in retroviruses. *Annual Review of Genetics* 24, 409–445.
- Kempf, D.J., Marsh, K.C., Fino, L.C., *et al.*, 1994. Design of orally bioavailable, symmetry-based inhibitors of HIV protease. *Bioorganic & Medicinal Chemistry* 2, 847–858.
- Khatib, F., DiMaio, F., Cooper, S., *et al.*, 2011. Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nature Structural & Molecular Biology* 18, 1175–1177.
- Koh, Y., Nakata, H., Maeda, K., *et al.*, 2003. Novel bis-tetrahydrofuranylurethane-containing nonpeptidic protease inhibitor (PI) UIC-94017 (TMC114) with potent activity against multi-PI-resistant human immunodeficiency virus *in vitro*. *Antimicrobial Agents and Chemotherapy* 47, 3123–3129.
- Krohn, A., Redshaw, S., Ritchie, J.C., Graves, B.J., Hatada, M.H., 1991. Novel binding mode of highly potent HIV-proteinase inhibitors incorporating the (R)-hydroxyethylamine isostere. *Journal of Medicinal Chemistry* 34, 3340–3342.
- Lam, P.Y., Jadhav, P.K., Eyermann, C.J., *et al.*, 1994. Rational design of potent, bioavailable, nonpeptide cyclic ureas as HIV protease inhibitors. *Science* 263, 380–384.
- Lapatto, R., Blundell, T., Hemmings, A., *et al.*, 1989. X-ray analysis of HIV-1 proteinase at 2.7 Å resolution confirms structural homology among retroviral enzymes. *Nature* 342, 299–302.
- Lea, A.P., Faulds, D., 1996. Ritonavir. *Drugs* 52, 541–546.
- Lukman, S., Verma, C.S., Fuentes, G., 2014. Exploiting protein intrinsic flexibility in drug design. *Advances in Experimental Medicine and Biology* 805, 245–269.
- Miller, M., Jaskólski, M., Rao, J.K.M., Leis, J., Wlodawer, A., 1989a. Crystal structure of a retroviral protease proves relationship to aspartic protease family. *Nature* 337, 576–579.
- Miller, M., Schneider, J., Sathyanarayana, B.K., *et al.*, 1989b. Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 Å resolution. *Science* 246, 1149–1152.
- Navia, M.A., Fitzgerald, P.M., McKeever, B.M., *et al.*, 1989. Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. *Nature* 337, 615–620.
- Parai, M.K., Huggins, D.J., Cao, H., *et al.*, 2012. Design, synthesis, and biological and structural evaluations of novel HIV-1 protease inhibitors to combat drug resistance. *Journal of Medicinal Chemistry* 55, 6328–6341.
- Prabu-Jeyabalan, M., Nalivaika, E., Schiffer, C.A., 2002. Substrate shape determines specificity of recognition for HIV-1 protease: Analysis of crystal structures of six substrate complexes. *Structure* 10, 369–381.
- Qiu, X., Liu, Z.P., 2011. Recent developments of peptidomimetic HIV-1 protease inhibitors. *Current Medicinal Chemistry* 18, 4513–4537.
- Ro, S., Baek, S.G., Lee, B., Ok, J.H., 1999. Conformational studies of irreversible HIV-1 protease inhibitors containing cis-epoxide as an amide isostere. *Journal of Peptide Research* 54, 242–248.
- Robinson, B.S., Riccardi, K.A., Gong, Y.F., *et al.*, 2000. BMS-232632, a highly potent human immunodeficiency virus protease inhibitor that can be used in combination with other available antiretroviral agents. *Antimicrobial Agents and Chemotherapy* 44, 2093–2099.
- Rous, P., 1911. A sarcoma of the fowl transmissible by an agent separable from the tumor cells. *Journal of Experimental Medicine* 13, 397–411.
- Schechter, I., Berger, A., 1967. On the size of the active site in proteases. I. Papain. *Biochemical and Biophysical Research Communications* 27, 157–162.
- Sham, H.L., Kempf, D.J., Molla, A., *et al.*, 1998. ABT-378, a highly potent inhibitor of the human immunodeficiency virus protease. *Antimicrobial Agents and Chemotherapy* 42, 3218–3224.
- Swain, A.L., Miller, M.M., Green, J., *et al.*, 1990. X-ray crystallographic structure of a complex between a synthetic protease of human immunodeficiency virus 1 and a substrate-based hydroxyethylamine inhibitor. *Proceedings of the National Academy of Sciences of the United States of America* 87, 8805–8809.
- Tang, J., James, M.N.G., Hsu, I.N., Jenkins, J.A., Blundell, T.L., 1978. Structural evidence for gene duplication in the evolution of the acid proteases. *Nature* 271, 618–621.
- Thaisrivongs, S., Strohbach, J.W., 1999. Structure-based discovery of Tipranavir disodium (PNU-140690E): A potent, orally bioavailable, nonpeptidic HIV protease inhibitor. *Biopolymers* 51, 51–58.
- Tie, Y., Kovalevsky, A.Y., Boross, P., *et al.*, 2007. Atomic resolution crystal structures of HIV-1 protease and mutants V82A and I84V with saquinavir. *Proteins* 67, 232–242.
- Toh, H., Ono, M., Saigo, K., Miyata, T., 1985. Retroviral protease-like sequence in the yeast transposon Ty1. *Nature* 315, 691.
- Vacca, J.P., Dorsey, B.D., Schleif, W.A., *et al.*, 1994. L-735,524: An orally bioavailable human immunodeficiency virus type 1 protease inhibitor. *Proceedings of the National Academy of Sciences of the United States of America* 91, 4096–4100.
- Weber, I.T., Miller, M., Jaskólski, M., *et al.*, 1989. Molecular modeling of the HIV-1 protease and its substrate binding site. *Science* 243, 928–931.
- Wlodawer, A., Erickson, J.W., 1993. Structure-based inhibitors of HIV-1 protease. *Annual Review of Biochemistry* 62, 543–585.
- Wlodawer, A., Miller, M., Jaskólski, M., *et al.*, 1989. Conserved folding in retroviral proteases: Crystal structure of a synthetic HIV-1 protease. *Science* 245, 616–621.
- Wlodawer, A., Vondrasek, J., 1998. Inhibitors of HIV-1 protease: A major success of structure-assisted drug design. *Annual Review of Biophysics and Biomolecular Structure* 27, 249–284.

Relavant Website

http://en.wikipedia.org/wiki/Hiv_protease_inhibitors
Wikipedia.

Blood Pressure, Proteases and Inhibitors

NE Clarke and AJ Turner, University of Leeds, Leeds, UK

© 2016 Elsevier Inc. All rights reserved.

Glossary

Angioedema The rapid, and potentially fatal, swelling (edema) of the dermis, subcutaneous tissue, and mucosa and submucosal tissues. It can be caused through an allergic reaction or be inherited. It is a rare side effect of ACE inhibitor therapy in the treatment of hypertension.

Treatment can range from use of antihistamines or steroids, through to adrenaline in severe cases.

Baroreflex The baroreflex or baroreceptor reflex is a homeostatic mechanism used by the body to maintain a constant blood pressure.

Endopeptidase An endopeptidase is a proteolytic enzyme that cleaves a peptide bond within a peptide or protein sequence but not at the termini. Examples would include trypsin, chymotrypsin, and neprilysin.

Endothelium The endothelium is a cell layer lining the interior surface of blood and lymphatic vessels. The

endothelium is composed of endothelial cells and those in direct contact with the blood are referred to as vascular endothelial cells.

Exocytosis A highly regulated, cellular mechanism for the secretion of protein and other cellular constituents contained in secretory vesicles within the cell. It can either be a constitutive process or triggered by extracellular signals.

Natriuretic A natriuretic agent is one that causes sodium excretion in the urine via the action of the kidneys.

Zymogen A zymogen is an inactive enzyme or hormone precursor. An example of a proenzyme zymogen would be trypsinogen (converted to active trypsin), and of a prohormone, proinsulin (converted to Insulin). The activation is carried out by proteolytic enzymes.

Introduction

Blood pressure homeostasis is maintained principally by circulating small peptides generated by a highly ordered cascade of proteases. Disruption of this cascade is a major pathological step in the development of cardiovascular disease, which is the leading cause of death in the developed world. In a healthy individual, blood pressure maintenance is a finely tuned balance. Simply standing causes a drop in blood pressure that must be counteracted to avoid fainting. The cardiovascular system responds to this change by constricting the peripheral vessels and increasing heart rate. In pathological conditions this balance is chronically disrupted; with hypotension the organs do not receive sufficient oxygen or nutrients and as a result have compromised function which can, over sustained periods, lead to permanent damage. Conversely, in hypertension, the circulating pressure causes vascular stress and kidney dysfunction. It is estimated that 1 billion people worldwide suffer from hypertension and 13.5% of premature deaths globally are attributed to it (Lawes *et al.*, 2008). It is generally recognized that around 30% of the normal variation in blood pressure is attributable to genetic factors with around 50% being due to environmental factors.

For the most part, the sympathetic nervous system controls blood pressure. A fall in arterial pressure is detected by the baroreceptors. The baroreflex acts to compensate for the change in pressure by stimulating an increase in sympathetic and a decrease in parasympathetic nerve activity, as well as the release of hormonal mediators of blood pressure such as the peptide, vasopressin, which reduces urine output in the kidney and causes vasoconstriction. Sympathetic nerve activity stimulates the release of adrenaline. Adrenaline acts through the β_1 adrenergic receptors in cardiac tissue to affect cardiac

output and ventricular contractility. Under most circumstances adrenaline causes an increase in blood pressure partly through vasoconstriction of the vascular beds and vasodilation of the vessels in skeletal and cardiac muscle. The baroreflex therefore works to maintain a constant blood pressure under normal conditions. Hypertension develops slowly over a long period of time, resulting from a loss of baroreceptor sensitivity, such that the receptors reset at a new higher pressure. The baroreflex still works to maintain 'normal' pressure; however, normal now refers to a hypertensive state.

Proteolytic Regulation of Blood Pressure

The Renin-Angiotensin System (RAS)

Overview of the system

The key proteolytic cascade regulating blood pressure is the RAS. The RAS links brain and kidney control of blood pressure by controlling fluid and electrolyte balance. The RAS is composed of peptides derived from precursor peptides/proteins which undergo successive cleavage by the enzymes (proteases) within the system. As such, the RAS is able to fine tune blood pressure via a highly ordered endocrine cascade. The RAS acts to increase blood pressure, primarily by producing the vaso-pressor octapeptide, angiotensin (Ang) II, in a pathway which has been thoroughly characterized over more than 60 years. Ang II also promotes tissue remodeling in various organs, especially heart and kidney. The classical pathway begins with the precursor protein angiotensinogen, from which the inert decapeptide, Ang I, is cleaved and this, in turn, is then converted to Ang II by angiotensin-converting enzyme (ACE). ACE inhibitors have been the first line of treatment against hypertension for decades now and their success has served to place

ACE and its biologically active product, Ang II, as central regulators of the RAS.

Our knowledge of this over-simplistic linear RAS pathway has evolved in recent years into a far more complex and regulated system with multiple biologically active peptide components. Two principal discoveries drove these developments: the emergence of a role for the vasodilatory peptide Ang-(1–7) and the discovery of its biosynthesis from Ang II by the ACE homolog, ACE2, together providing a counterbalance to the actions of Ang II (Campagnole-Santos *et al.*, 1989; Tipnis *et al.*, 2000). Ang-(1–7) can also be generated from Ang I by another zinc peptidase, neprilysin (NEP) (Rice *et al.*, 2004).

The major actions of Ang II are effected through the angiotensin II type 1 receptor, AT₁R, which is a G-protein coupled receptor that mediates increases in blood pressure through a variety of mechanisms, which include vasoconstriction, renal sodium reabsorption, aldosterone and vasopressin secretion, and activation of hypothalamic neurons to stimulate thirst and hence fluid uptake. The second, less widely expressed receptor for Ang II is the related AT₂R, which mediates generally opposing effects to those of the AT₁R causing vasodilation, inhibition of hypertrophy and fibrosis (Carey *et al.*, 2000). The most well-established roles of Ang-(1–7) are thought to be mediated through its own specific receptor, termed Mas (Santos *et al.*, 2003) opposing the actions of Ang II that are mediated through the AT₁R. While AT₁R antagonists such as losartan are widely used in the treatment of hypertension and vascular disease, AT₂R or Mas receptor agonists may provide alternative therapeutic strategies. Likewise, up-regulation of ACE2 levels or activity could be cardioprotective.

Another fork in the pathway mediated by aminopeptidase A generates the peptide Ang III directly from Ang II. Ang III interacts with both AT₁R and AT₂R with similar affinity to Ang II and appears to have a key role in the brain RAS controlling blood pressure (Marc and Llorens Cortes, 2011). The action of aminopeptidase N on Ang III generates Ang IV which increases renal blood flow, decreases tubular sodium transport and also has central nervous system actions, having an influence on memory and acting through a receptor which is itself a protease – the insulin-regulated aminopeptidase (IRAP), inhibitors of which have actions as cognitive enhancers (Albiston *et al.*, 2011). Most recently, two new functional peptides in the RAS have been identified: angiotensin A, derived by decarboxylation of Ang II (Jankowski *et al.*, 2007), and alamandine produced from either Ang A or Ang-(1–7) (Lautner *et al.*, 2013). These interlinked pathways are summarized in Figure 1.

RAS proteases

From angiotensinogen to Ang I: renin

Angiotensinogen, a glycoprotein secreted by the liver into the bloodstream, is the immediate precursor of Ang I which is cleaved proteolytically from its N-terminus by renin, in the rate-limiting step of the RAS cascade. Renin is initially synthesized in a zymogen form, prorenin, and is stored in secretory granules in the kidney. Prorenin is both constitutively secreted and subject to regulated exocytosis in response to stimuli and it accounts for 70–90% of circulating renin. The

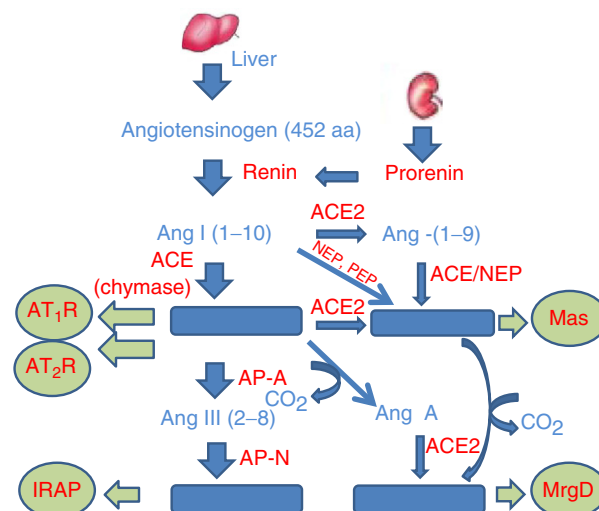


Figure 1 The complex network of the RAS. The precursor protein, angiotensinogen, secreted from the liver, is cleaved by the renal enzyme renin to generate the inactive decapeptide Ang I released from its N-terminus. ACE converts Ang I to the vasoregulatory octapeptide Ang II which acts through two types of G-protein coupled receptors, AT₁R and AT₂R. Ang II can also be converted to the counter-regulatory heptapeptide, Ang-(1–7), predominantly by ACE2, although an alternative but minor pathway exists from Ang I and Ang-(1–9). Other Ang products can also be formed, including Ang III, Ang IV, Ang-(1–7), and alamandine, which act through their own respective receptors (AT₁R/AT₂R for Ang III in brain (not shown), IRAP for Ang IV, Mas for Ang-(1–7), and MrgD for alamandine (see text for details)). ACE, angiotensin-converting enzyme; Ang, angiotensin; AP-A, aminopeptidase A; AP-N, aminopeptidase N; NEP, neprilysin; PEP, proline endopeptidase.

conversion of prorenin to renin involves the removal of a 43 amino acid peptide from its N-terminus that obstructs the active site. This activation can be facilitated by multiple pathways: enzymic cleavage, which results in irreversible activation of renin; propeptide unfolding that occurs at low temperature or pH resulting in reversible conversion, or via the action of the prorenin receptor (PRR) that results in reversible activation. Proteases capable of cleaving the prorenin propeptide include proconvertase I, kallikrein and cathepsin B. Activation of prorenin resulting from binding to the PRR allows spatial control of the conversion at the cell-surface (Nguyen *et al.*, 2002). However, the main function of the PRR may not relate to the RAS but to other quite distinct signalling pathways (Nguyen, 2011).

A big question in the RAS field has been whether the generation of Ang II *in vivo* in humans can involve enzymes or pathways other than those employing renin and ACE. The observation that renin and prorenin knockout mice lack detectable plasma levels of Ang I, and the effects of renin/prorenin deficiency in humans, establish a predominant role for renin in Ang I generation. Renin inhibitors also effectively suppress levels of Ang I and Ang II. However, ACE-null mice do have detectable circulating Ang II levels implying other enzymes may be able to compensate. One enzyme that could contribute to this process is the serine protease, chymase, and some studies indicate that chymase may be important in

Ang II generation, particularly in the human heart (Urata *et al.*, 1990). However, most clinical evidence currently supports the concept that renin, together with ACE, predominate in human Ang II generation although, at some sites and in certain pathological circumstances, other enzymes and particularly chymase, may be relevant (Campbell, 2012).

From Ang I to Ang II: ACE

The ACE and ACE2 are both zinc metallopeptidases. This metal ion activates a water molecule which then acts as a nucleophile to hydrolyze the peptide bond of their respective substrates. Members of this peptidase family are primarily exopeptidases, acting near the C-terminus of their oligopeptide substrates. These two enzymes have additional noncatalytic roles that will not be focused on in this article but have been well described elsewhere (Clarke *et al.*, 2012; Gonzalez-Villalobos *et al.*, 2013; Kuba *et al.*, 2013).

The ACE protein exists in two forms, the somatic form (sACE), expressed on the surface of vascular endothelial, epithelial and neuroepithelial cells, and the testicular form (tACE) present only in germinal cells and which regulates male fertility. The two isoforms are the result of tissue-specific promoter sequences within the ACE gene (Soubrier *et al.*, 1988; Hubert *et al.*, 1991). sACE has two catalytic sites (N- and C-domains) thought to be the result of a gene duplication, whereas the tACE contains only one of these sites (Hubert *et al.*, 1991).

ACE has a broad substrate specificity acting both as a peptidyl dipeptidase (i.e., releasing a C-terminal dipeptide from its substrates) as well as an endopeptidase and, in addition to Ang I, it can hydrolyze bradykinin, substance P, neurotensin, and a number of other peptide substrates *in vitro*, including the hemoregulatory and anti-fibrotic tetrapeptide, acetyl-Ser-Asp-Lys-Pro (AcSDKP; Rieger *et al.*, 1993). The two catalytic domains of ACE have differing substrate specificities (Gonzalez-Villalobos *et al.*, 2013). Angiotensin I and the counter-regulatory vasodilator bradykinin are substrates for both catalytic sites of ACE, although Ang II is more efficiently formed by the C-terminal domain. In contrast, AcSDKP is a substrate only for the N-terminal site of ACE and thereby affects inflammation and organ fibrosis. The activities of ACE are chloride-dependent, which is due to the presence of two distinct chloride binding sites (CL1 and CL2) located approximately 20 Å apart in ACE (Yates *et al.*, 2014).

ACE is subject to posttranslational regulation being released from the plasma membrane by proteolytic shedding (Hooper *et al.*, 1987). In fact, ACE was initially isolated from plasma in its shed form before the identification of its primary existence as a transmembrane, cell-surface enzyme (Skeggs *et al.*, 1956). The constitutive shedding of ACE is performed by an as yet unidentified zinc-dependent secretase or sheddase (Parvathy *et al.*, 1997). Such shedding events are commonly mediated by a disintegrin and metalloproteinase (ADAM) proteases, particularly ADAM10 or 17 (van Goor *et al.*, 2009) but this is not the case for ACE (Allinson *et al.*, 2004). However, under certain stimulated conditions ADAM9 may contribute to the shedding (English *et al.*, 2012). One regulator of the shedding event is the calcium-binding protein calmodulin, which binds to the cytoplasmic tail of ACE, and calmodulin

inhibitors stimulate the shedding of ACE (Chattopadhyay *et al.*, 2005).

ACE inhibitors as antihypertensive drugs have become one of the most successful rational therapeutic developments in biomedical science. Peptide inhibitors of ACE were first isolated from snake venom in 1965 and their cardiovascular effects demonstrated (Ferreira and Rocha e Silva, 1965). The first non-peptide inhibitor of ACE based on these studies, captopril, was developed in 1977 (Ondetti *et al.*, 1977) and entered clinical use in 1981. Unfortunately, due to the presence of a zinc coordinating thiol group, captopril had significant side effects, including development of a rash and taste disturbances. Other inhibitors were subsequently developed based on the structure of captopril but used different zinc coordinating groups. For example, lisinopril and enalaprilat, in which the thiol group of captopril is replaced by a carboxylate group. These too, although very effective in blood pressure control, do still have some side effects, partly because of their lack of catalytic domain specificity, most notably a persistent cough in 5–20% of patients and, more seriously, angioedema in 0.1–0.5% of patients (Morimoto *et al.*, 2004). The occurrence of angioedema is probably due to inhibition of bradykinin catabolism by ACE, resulting in the persistence of bradykinin and its vasodilatory and vascular permeability properties. The 3D structure of ACE was solved over a decade ago and subsequent structural studies are now allowing the development of a new generation of highly selective ACE inhibitors targeted against either the N- or C-domain of the enzyme (Natesh *et al.*, 2003; Douglas *et al.*, 2014). ACE inhibitor development has been widely reviewed elsewhere (see e.g., Anthony *et al.*, 2012).

While specific renin inhibitors reduce Ang II levels they do not affect the levels of other ACE substrates such as bradykinin and substance P which contribute to some of the side effects of ACE inhibitor use. Hence, inhibiting renin as the first and rate-limiting step would at first sight appear a better therapeutic option to block Ang II formation selectively than the use of ACE inhibitors. Furthermore, renin inhibition avoids ACE-independent pathways from Ang I to Ang II (e.g., chymase). Renin is an aspartic protease and, as such, designing specific and selective renin inhibitors has proved extremely difficult. However, aliskiren is proving its efficacy as the newest antihypertensive drug and the first orally active, clinically available renin inhibitor (Friedrich *et al.*, 2013).

Despite intensive investigation of the role of ACE in the development of hypertension and cardiovascular disease, relatively little is known about the signals that regulate vascular ACE expression levels. One such factor is that of fluid-induced shear stress, the drag force acting on the endothelium as a result of blood flow, and which plays a critical role in the development of endothelial atherosclerosis. This effect is mediated via the adenosine monophosphate (AMP)-activated protein kinase (AMPK) and involves posttranscriptional regulation by certain microRNAs (miR-143/145) (Kohlstedt *et al.*, 2013). There is additional evidence that the translation of ACE is regulated by these and other microRNAs in vascular smooth muscle cells and elsewhere (Fernandes *et al.*, 2011). There appears to be a strong influence of epigenetic mechanisms on sACE expression involving both DNA methylation and histone acetylation (Rivière *et al.*, 2011). Probably the

most studied aspect of ACE genetics and expression is the occurrence of an insertion/deletion (I/D) polymorphism in the ACE gene which is a major locus that determines plasma and tissue ACE levels (Rigat *et al.*, 1990). The deletion-allele has been proposed as a renal and cardiovascular risk factor (Staessen *et al.*, 1997).

From Ang II to Ang-(1–7): ACE2

ACE2 is a close homolog of ACE discovered in 2000 by two independent molecular approaches (Tipnis *et al.*, 2000; Donoghue *et al.*, 2000). The ACE2 gene is located on chromosome Xp22, a locus that was already known to be linked to hypertension when ACE2 was discovered, which suggested that ACE2 could be a key gene contributing to this link. Gene deletion of ACE2 in mice revealed it to be an essential regulator of heart function counter-balancing ACE (Crackower *et al.*, 2002). Even partial (heterozygote) loss of ACE2 is sufficient to enhance the susceptibility to heart disease (Wang *et al.*, 2014). ACE and ACE2 are single span, transmembrane proteins with their N-terminal catalytic sites being exposed on the extracellular surface, facilitating the metabolism of circulating peptides. The ACE2 protein resembles a chimera composed of a single ACE-like catalytic ectodomain fused to a homolog of the small transmembrane protein, collectrin. Despite having high similarity in sequence to ACE, particularly around the active site, ACE2 acts as a carboxypeptidase removing just a single amino acid from the C-terminus of susceptible substrates (Tipnis *et al.*, 2000). The distinct binding specificity of ACE2 was demonstrated by the failure of ACE inhibitors to block its activity. Comparative substrate specificities of ACE and ACE2 for angiotensin and bradykinin peptides are shown in Figure 2. Like ACE, the

activities of ACE2 are chloride-dependent, which is due to the presence of a comparable CL1 site in the ectodomain as revealed by structural and modeling studies (Rushworth *et al.*, 2008).

Soon after the discovery of ACE2, potent and specific inhibitors were rationally developed, for example, the compound MLN-4760 (Dales *et al.*, 2002), which has been used in numerous studies of ACE2 action *in vivo* and *in vitro*. MLN-4760 is a non-peptide inhibitor with which the inhibitor-bound crystal structure was subsequently solved (Towler *et al.*, 2004). MLN-4760 (also known as GL1001) was originally, but unsuccessfully, developed as a potential antiobesity drug and more recently is being examined as an anti-inflammatory agent (Byrnes *et al.*, 2009). Several other selective and potent ACE2 inhibitors have been developed, including a peptide inhibitor, DX600, (Huang *et al.*, 2003) and, more recently, a phosphinic-peptide (416F2), which mimics the enzyme transition state (Mores *et al.*, 2008).

ACE and ACE2 appear to be co-regulated with ACE2 being down-regulated in hypertension an effect that is mediated, at least in part, by Ang II (Gallagher *et al.*, 2008) although the mechanism is unclear. Some other peptide and steroid hormones modulate ACE2 expression levels as well as hypoxia and action of some cytokines. ACE2, like ACE, appears to be regulated by AMP kinase activation through epigenetic control via the histone deacetylase, SIRT1 (Clarke *et al.*, 2014). At the posttranscriptional level, miRNAs are also reported to regulate ACE2 expression (miRNA-143 and 421) (Fernandes *et al.*, 2011; Lambert *et al.*, 2014). These various regulatory sites may provide opportunities for therapeutic modulation of ACE2 whose up-regulation appears to be cardioprotective. ACE2 is shed from the membrane through proteolytic cleavage by

Comparative substrate specificity of ACE and ACE2

	ACE	ACE2
Angiotensin I	Asp Arg Val Tyr Ile His Pro Phe ↑ His Leu	Asp Arg Val Tyr Ile His Pro Phe His ↑ Leu
Angiotensin II	Not cleaved	Asp Arg Val Tyr Ile His Pro ↑ Phe
Angiotensin-(1–9)	Asp Arg Val Tyr Ile His Pro ↑ Phe His	Not cleaved
Angiotensin-(1–7)	Asp Arg Val Tyr Ile ↑ His Pro	Not cleaved
Angiotensin A	Not cleaved	Ala Arg Val Tyr Ile His Pro ↑ Phe
Bradykinin	Arg Pro Pro Gly Phe Ser Pro ↑ Phe Arg	Not cleaved
(des-Arg ⁹) Bradykinin	Not cleaved	Arg Pro Pro Gly Phe Ser Pro ↑ Phe

Figure 2 Comparative substrate specificity of ACE and ACE2. The hydrolysis of angiotensin and bradykinin peptides by ACE and ACE2 is illustrated. ↑ indicates the site of enzyme cleavage (peptidyl dipeptidase action for ACE) and carboxypeptidase action for ACE2. The heptapeptide product of angiotensin A cleavage by ACE2 is alamandine (Lautner *et al.*, 2013).

ADAM17 and, like ACE, this process is regulated by calmodulin (Lambert *et al.*, 2008). Furthermore, it has been shown that Ang II induces ACE2 shedding by promoting ADAM17 activity through signaling by reactive oxygen species (Patel *et al.*, 2013). This creates a positive feedback mechanism whereby Ang II facilitates the loss of its negative regulator, ACE2. For more detail on the regulation of ACE2, see Clarke and Turner, 2012.

The Natriuretic Peptide Family and Their Proteases (Corin, Furin, and NEP)

The family of natriuretic peptides is an evolutionarily, highly conserved cardiac endocrine system that regulates salt and body fluid balance, and hence blood pressure. The natriuretic peptide system can be viewed as a counterbalance to the RAS. Originally discovered in the early 1980s (de Bold *et al.*, 1981), the natriuretic peptides are stored in and secreted from granules within atrial tissue cardiocyte cells. This group of peptides comprises three closely related members: atrial natriuretic peptide (ANP), brain or B-type natriuretic peptide (BNP), both mainly synthesized in the heart, and C-type natriuretic peptide (CNP) which has a much broader tissue distribution and is principally of endothelial cell origin. They are synthesized as large prepropeptides which are consecutively processed proteolytically by signal peptidase to form the inactive propeptide and then by a proprotein convertase to form the active peptide. The active, vasodilatory ANP is a 28-amino acid peptide containing a 17-amino acid disulfide-bridged ring structure which typifies this family and which is shared by BNP and CNP. In the case of ANP and partly for BNP, the proprotein convertase is a transmembrane, cardiac serine protease known as corin, which has substrate specificity similar to trypsin, cleaving at a basic amino acid. BNP can also be processed by the intracellular, Golgi-enriched serine protease, furin, and CNP is exclusively processed by this enzyme (Wu *et al.*, 2009).

The natriuretic peptides act through natriuretic peptide receptors NPR-A (both ANP and BNP) and NPR-B (CNP), both of which are linked to guanylate cyclase. A third receptor (NPR-C) was originally thought to act as a clearance receptor to remove the peptides but is now known to act through inhibition of adenylate cyclase and phospholipase C activation (Mouawad *et al.*, 2004). Knockout mice lacking ANP, NPR-A, or corin all display a hypertensive phenotype and, while BNP-null mice are not hypertensive, they display cardiac fibrosis showing BNP to act as a local regulator of ventricular remodeling. More recent data also imply that CNP and NPR-B can play important roles in regulating cardiac hypertrophy and remodeling (Del Ry, 2013).

Modulating natriuretic peptide levels can provide therapeutic opportunities in hypertension and heart failure and the discovery that the zinc peptidase NEP is the principal protease inactivating them (Kenny *et al.*, 1993) led to a profusion of activity into developing NEP inhibitors as potential antihypertensives, and even dual 'vasopeptidase inhibitors' inhibiting both ACE and NEP (Worthley *et al.*, 2004). However, the discovery that NEP inactivates the Alzheimer's disease-related and neurotoxic amyloid β -peptide (A β) (Iwata *et al.*, 2000) has diminished enthusiasm for such

developments since such inhibitors might potentiate the cognitive and behavioral problems associated with the disease. Potentiating other features of the natriuretic peptide systems may still provide novel therapeutics and the peptides and their precursor forms can also serve as valuable biomarkers of heart disease. These peptide systems have also been suggested to play a more general integrating role in lifestyle-related metabolic and cardiovascular disorders (Zois *et al.*, 2014).

The Endothelin Peptide Family and Their Proteases

A highly potent vasoconstrictor peptide produced and secreted by endothelial cells, endothelin-1 (ET-1), was first isolated and characterized in 1988 (Yanagisawa *et al.*, 1988) and shown to be synthesized from its large precursor, proET-1, through two successive proteolytic steps, firstly catalyzed by the proprotein convertase furin to form the inactive big ET-1 and then by a protease shown to be inhibited by the NEP inhibitor, phosphoramidon. However, the biosynthetic enzyme turned out not to be NEP itself but a subsequently cloned close homolog, now termed endothelin-converting enzyme-1 (ECE-1) (Xu *et al.*, 1994). Rather, NEP turned out to be the primary protease inactivating ET-1 (Vijayaraghavan *et al.*, 1990). This provides another evolutionary example in which two homologous proteases function successively to synthesize a vasoactive peptide hormone and to metabolize it (cf. ACE and ACE2 above). In total there are three endothelin peptides (ET-1, -2, and -3) synthesized from separate genes, although ET-1 is the most abundant in the vasculature. ET-1 is a 21-amino acid peptide consisting of two rings formed by two intra-chain disulfide bonds and a 6-residue C-terminal tail and the other ETs have similar structures. The ETs act through related receptor subtypes (ET-A and -B) and ET-1 receptor antagonists (e.g., bosentan) are licensed for use in pulmonary arterial hypertension. ECE-1 exhibits similar topology to NEP being a zinc metallopeptidase composed of a large extracellular or luminal C-terminal catalytic domain, a transmembrane region and a small N-terminal cytoplasmic domain. It exists in four isoforms (designated ECE-1a, b, c, and d), differing only in part of their N-terminal cytoplasmic domains, which are derived from a single gene through the action of alternative promoters. They have similar catalytic properties but distinct subcellular localization and tissue distribution (Turner *et al.*, 1998). A second ECE-like gene encodes an ECE-2 protein, which is mainly neuronally located and likely involved in endosomal processing of neuropeptide precursors. ECE-1 inhibitors have been developed and may have therapeutic applications in hypertension, chronic heart or renal failure, and cerebral vasospasm (Jeng *et al.*, 2002). However, their clinical use is somewhat compromised by the involvement of ECE-1 (and ECE-2), like NEP, in metabolism of the Alzheimer's amyloid β -peptide (Eckman *et al.*, 2001).

Summary

In summary, the complex biology of blood pressure regulation relies on the counter-balancing effects of a host of proteases mediating interconversions of several distinct peptide families

(angiotensins, bradykinin, vasopressin, natriuretic, and endothelin peptides). Several of these proteases are the targets for successful antihypertensive and cardioprotective drugs, as are the peptide receptors. Newer discoveries in the biology of blood pressure regulation (ACE2, mas receptor, Ang-(1-7), and other peptides) highlight novel opportunities for intervention and treatment of cardiovascular disease.

References

- Albiston, A.L., Diwakarla, S., Fernando, R.N., *et al.*, 2011. Identification and development of specific inhibitors for insulin-regulated aminopeptidase as a new class of cognitive enhancers. *British Journal of Pharmacology* 164, 37–47.
- Allinson, T.M., Parkin, E.T., Condon, T.P., *et al.*, 2004. The role of ADAM10 and ADAM17 in the ectodomain shedding of angiotensin converting enzyme and the amyloid precursor protein. *European Journal of Biochemistry* 271, 2539–2547.
- Anthony, C.S., Masuyer, G., Sturrock, E.D., Acharya, K.R., 2012. Structure based drug design of angiotensin-I converting enzyme inhibitors. *Current Medicinal Chemistry* 19, 845–855.
- de Bold, A.J., Borenstein, H.B., Veress, A.T., Sonnenberg, H., 1981. A rapid and potent natriuretic response to intravenous injection of atrial myocardial extract in rats. *Life Sciences* 28, 89–94.
- Byrnes, J.J., Gross, S., Ellard, C., *et al.*, 2009. Effects of the ACE2 inhibitor GL1001 on acute dextran sodium sulfate-induced colitis in mice. *Inflammation Research* 58, 819–827.
- Campagnole-Santos, M.J., Diz, D.I., Santos, R.A., *et al.*, 1989. Cardiovascular effects of angiotensin-(1-7) injected into the dorsal medulla of rats. *American Journal of Physiology* 257, H324–H329.
- Campbell, D.J., 2012. Angiotensin II generation *in vivo*: Does it involve enzymes other than renin and angiotensin-converting enzyme? *Journal of the Renin-Angiotensin-Aldosterone System* 13, 314–316.
- Carey, R.M., Wang, Z.Q., Siragy, H.M., 2000. Role of the angiotensin type 2 receptor in the regulation of blood pressure and renal function. *Hypertension* 35, 155–163.
- Chattopadhyay, S., Santhamma, K.R., Sengupta, S., *et al.*, 2005. Calmodulin binds to the cytoplasmic domain of angiotensin-converting enzyme and regulates its phosphorylation and cleavage secretion. *Journal of Biological Chemistry* 280, 33847–33855.
- Clarke, N.E., Fisher, M.J., Porter, K.E., Lambert, D.W., Turner, A.J., 2012. Angiotensin converting enzyme (ACE) and ACE2 bind integrins and ACE2 regulates integrin signalling. *PLoS One* 7, e34747.
- Clarke, N.E., Turner, A.J., 2012. Angiotensin-converting enzyme 2: The first decade. *International Journal of Hypertension* 2012, 307315.
- Clarke, N.E., Belyaev, N.D., Lambert, D.W., Turner, A.J., 2014. Epigenetic regulation of angiotensin-converting enzyme 2 (ACE2) by SIRT1 under conditions of cell energy stress. *Clinical Science* 126, 507–516.
- Crackower, M., Sarao, R., Oudit, G., *et al.*, 2002. Angiotensin-converting enzyme 2 is an essential regulator of heart function. *Nature* 417, 822–828.
- Dales, N.A., Gould, A.E., Brown, J.A., *et al.*, 2002. Substrate-based design of the first class of angiotensin-converting enzyme-related carboxypeptidase (ACE2) inhibitors. *Journal of the American Chemical Society* 124, 11852–11853.
- Del Ry, S., 2013. C-type natriuretic peptide: A new cardiac mediator. *Peptides* 40, 93–98.
- Donoghue, M., Hsieh, F., Baronas, E., *et al.*, 2000. A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. *Circulation Research* 87, E1–E9.
- Douglas, R.G., Sharma, R.K., Masuyer, G., *et al.*, 2014. Fragment-based design for the development of N-domain-selective angiotensin-1-converting enzyme inhibitors. *Clinical Science* 126, 305–313.
- Eckman, E.A., Reed, D.K., Eckman, C.B., 2001. Degradation of the Alzheimer's amyloid beta peptide by endothelin-converting enzyme. *Journal of Biological Chemistry* 276, 24540–24548.
- English, W.R., Corvol, P., Murphy, G., 2012. LPS activates ADAM9 dependent shedding of ACE from endothelial cells. *Biochemical and Biophysical Research Communications* 421, 70–75.
- Fernandes, T., Hashimoto, N.Y., Magalhaes, F.C., *et al.*, 2011. Aerobic exercise training-induced left ventricular hypertrophy involves regulatory MicroRNAs, decreased angiotensin-converting enzyme-angiotensin ii, and synergistic regulation of angiotensin-converting enzyme 2-angiotensin (1-7). *Hypertension* 58, 182–189.
- Ferreira, S.H., Rocha e Silva, M., 1965. Potentiation of bradykinin and eledoisin by BPF (bradykinin potentiating factor) from *Bothrops jararaca* venom. *Experientia* 21, 347–349.
- Friedrich, S., Schmieder, R.E., 2013. Review of direct renin inhibition by aliskiren. *Journal of the Renin-Angiotensin-Aldosterone System* 14, 193–196.
- Gallagher, P.E., Ferrario, C.M., Tallant, E.A., 2008. MAP kinase/phosphatase pathway mediates the regulation of ACE2 by angiotensin peptides. *American Journal of Physiology Cell Physiology* 295, C1169–C1174.
- Gonzalez-Villalobos, R.A., Shen, X.Z., Bernstein, E.A., *et al.*, 2013. Rediscovering ACE: Novel insights into the many roles of the angiotensin-converting enzyme. *Journal of Molecular Medicine* 91, 1143–1154.
- van Goor, H., Melenhorst, W.B., Turner, A.J., Holgate, S.T., 2009. Adamalysins in biology and disease. *Journal of Pathology* 219, 277–286.
- Hooper, N.M., Keen, J., Pappin, D.J., Turner, A.J., 1987. Pig kidney angiotensin converting enzyme. Purification and characterization of amphipathic and hydrophilic forms of the enzyme establishes C-terminal anchorage to the plasma membrane. *Biochemical Journal* 247, 85–93.
- Huang, L., Sexton, D.J., Skogerson, K., *et al.*, 2003. Novel peptide inhibitors of angiotensin-converting enzyme 2. *Journal of Biological Chemistry* 278, 15532–15540.
- Hubert, C., Houot, A.M., Corvol, P., Soubrier, F., 1991. Structure of the angiotensin I-converting enzyme gene. Two alternate promoters correspond to evolutionary steps of a duplicated gene. *Journal of Biological Chemistry* 266, 15377–15383.
- Iwata, N., Tsubuki, S., Takaki, Y., *et al.*, 2000. Identification of the major Abeta1-42-degrading catabolic pathway in brain parenchyma: Suppression leads to biochemical and pathological deposition. *Nature Medicine* 6, 143–150.
- Jankowski, V., Vanholder, R., van der Giet, M., *et al.*, 2007. Mass-spectrometric identification of a novel angiotensin peptide in human plasma. *Arteriosclerosis, Thrombosis, and Vascular Biology* 27, 297–302.
- Jeng, A.Y., Mulder, P., Kwan, A.L., Battistini, B., 2002. Nonpeptidic endothelin-converting enzyme inhibitors and their potential therapeutic applications. *Canadian Journal of Physiology and Pharmacology* 80, 440–449.
- Kenny, A.J., Bourne, A., Ingram, J., 1993. Hydrolysis of human and pig brain natriuretic peptides, urodilatin, C-type natriuretic peptide and some C-receptor ligands by endopeptidase-24.11. *Biochemical Journal* 291, 83–88.
- Kohlstedt, K., Trouvain, C., Boettger, T., *et al.*, 2013. AMP-activated protein kinase regulates endothelial cell angiotensin-converting enzyme expression via p53 and the post-transcriptional regulation of microRNA-143/145. *Circulation Research* 112, 1150–1158.
- Kuba, K., Imai, Y., Penninger, J.M., 2013. Multiple functions of angiotensin-converting enzyme 2 and its relevance in cardiovascular diseases. *Circulation Journal* 77, 301–308.
- Lambert, D.W., Clarke, N.E., Hooper, N.M., Turner, A.J., 2008. Calmodulin interacts with angiotensin-converting enzyme-2 (ACE2) and inhibits shedding of its ectodomain. *FEBS Letters* 582, 385–390.
- Lambert, D.W., Lambert, L.A., Clarke, N.E., *et al.*, 2014. Angiotensin-converting enzyme 2 is subject to post-transcriptional regulation by miR-421. *Clinical Science* 127, 243–249.
- Lautner, R.Q., Villela, D.C., Fraga-Silva, R.A., *et al.*, 2013. Discovery and characterization of alamandine: A novel component of the renin-angiotensin system. *Circulation Research* 112, 1104–1111.
- Lawes, C.M., Vander Hoorn, S., Rodgers, A., 2008. Global burden of blood-pressure-related disease, 2001. *Lancet* 371, 1513–1518.
- Marc, Y., Llorens-Cortes, C., 2011. The role of the brain renin-angiotensin system in hypertension: Implications for new treatment. *Progress in Neurobiology* 95, 89–103.
- Mores, A., Matziari, M., Beau, F., *et al.*, 2008. Development of potent and selective phosphinic peptide inhibitors of angiotensin-converting enzyme 2. *Journal of Medicinal Chemistry* 51, 2216–2226.
- Morimoto, T., Gandhi, T.K., Fiskio, J.M., *et al.*, 2004. An evaluation of risk factors for adverse drug events associated with angiotensin-converting enzyme inhibitors. *Journal of Evaluation in Clinical Practice* 10, 499–509.
- Mouawad, R., Li, Y., Anand-Srivastava, M.B., 2004. Atrial natriuretic peptide-C receptor-induced attenuation of adenylyl cyclase signaling activates phosphatidylinositol turnover in A10 vascular smooth muscle cells. *Molecular Pharmacology* 65, 917–924.
- Natesh, R., Schwager, S.L., Sturrock, E.D., Acharya, K.R., 2003. Crystal structure of the human angiotensin-converting enzyme-lisinopril complex. *Nature* 421, 551–554.
- Nguyen, G., 2011. Renin, (pro)renin and receptor: An update. *Clinical Science* 120, 169–178.

- Nguyen, G., Delarue, F., Burckle, C., *et al.*, 2002. Pivotal role of the renin/prorenin receptor in angiotensin II production and cellular responses to renin. *Journal of Clinical Investigation* 109, 1417–1427.
- Ondetti, M.A., Rubin, B., Cushman, D.W., 1977. Design of specific inhibitors of angiotensin-converting enzyme: New class of orally active antihypertensive agents. *Science* 196, 441–444.
- Parvathy, S., Oppong, S.Y., Karran, E.H., *et al.*, 1997. Angiotensin-converting enzyme secretase is inhibited by zinc metalloprotease inhibitors and requires its substrate to be inserted in a lipid bilayer. *Biochemical Journal* 327, 37–43.
- Patel, V.B., Clarke, N.E., Wang, Z., *et al.*, 2013. Angiotensin II induced proteolytic cleavage of myocardial ACE2 is mediated by TACE/ADAM-17: A positive feedback mechanism in the RAS. *Journal of Molecular and Cellular Cardiology* 66, 167–176.
- Rice, G.I., Thomas, D.A., Grant, P.J., Turner, A.J., Hooper, N.M., 2004. Evaluation of angiotensin-converting enzyme (ACE), its homologue ACE2 and neprilysin in angiotensin peptide metabolism. *Biochemical Journal* 383, 45–51.
- Rieger, K.J., Saez-Servent, N., Papet, M.P., *et al.*, 1993. Involvement of human plasma angiotensin I-converting enzyme in the degradation of the haemoregulatory peptide N-acetyl-seryl-aspartyl-lysyl-proline. *Biochemical Journal* 296, 373–378.
- Rigat, B., Hubert, C., Alhenc-Gelas, F., *et al.*, 1990. An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. *Journal of Clinical Investigation* 86, 1343–1346.
- Rivière, G., Lienhard, D., Andrieu, T., *et al.*, 2011. Epigenetic regulation of somatic angiotensin-converting enzyme by DNA methylation and histone acetylation. *Epigenetics* 6, 478–489.
- Rushworth, C.A., Guy, J.L., Turner, A.J., 2008. Residues affecting the chloride regulation and substrate selectivity of the angiotensin-converting enzymes (ACE and ACE2) identified by site-directed mutagenesis. *FEBS Journal* 275, 6033–6042.
- Santos, R.A., Simoes e Silva, A.C., Maric, C., *et al.*, 2003. Angiotensin-(1–7) is an endogenous ligand for the G protein-coupled receptor Mas. *Proceedings of the National Academy of Sciences the United States of America* 100, 8258–8263.
- Skeggs Jr., L.T., Kahn, J.R., Shumway, N.P., 1956. The preparation and function of the hypertensin-converting enzyme. *Journal of Experimental Medicine* 103, 295–299.
- Soubrier, F., Alhenc-Gelas, F., Hubert, C., *et al.*, 1988. Two putative active centers in human angiotensin I-converting enzyme revealed by molecular cloning. *Proceedings of the National Academy of Sciences USA* 85, 9386–9390.
- Staessen, J.A., Wang, J.G., Ginocchio, G., *et al.*, 1997. The deletion/insertion polymorphism of the angiotensin converting enzyme gene and cardiovascular-renal risk. *Journal of Hypertension* 15, 1579–1592.
- Tipnis, S.R., Hooper, N.M., Hyde, R., *et al.*, 2000. A human homolog of angiotensin-converting enzyme. Cloning and functional expression as a captopril-insensitive carboxypeptidase. *Journal of Biological Chemistry* 275, 33238–33243.
- Towler, P., Staker, B., Prasad, S.G., *et al.*, 2004. ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *Journal of Biological Chemistry* 279, 17996–18007.
- Turner, A.J., Barnes, K., Schweizer, A., Valdenaire, O., 1998. Isoforms of endothelin-converting enzyme: Why and where? *Trends in Pharmacological Sciences* 19, 483–486.
- Urala, H., Kinoshita, A., Misono, K.S., Bumpus, F.M., Husain, A., 1990. Identification of a highly specific chymase as the major angiotensin II-forming enzyme in the human heart. *Journal of Biological Chemistry* 265, 22348–22357.
- Vijayaraghavan, J., Scicli, A.G., Carretero, O.A., *et al.*, 1990. The hydrolysis of endothelins by neutral endopeptidase 24.11 (enkephalinase). *Journal of Biological Chemistry* 265, 14150–14155.
- Yates, C.J., Masuyer, G., Schwager, S.L., *et al.*, 2014. Molecular and thermodynamic mechanisms of the chloride-dependent human angiotensin-I-converting enzyme (ACE). *Journal of Biological Chemistry* 289, 1798–1814.
- Wang, W., Patel, V.B., Parajuli, N., *et al.*, 2014. Heterozygote loss of ACE2 is sufficient to increase the susceptibility to heart disease. *Journal of Molecular Medicine* 92, 847–858.
- Worthley, M.I., Corti, R., Worthley, S.G., 2004. Vasoepitidase inhibitors: Will they have a role in clinical practice? *British Journal of Clinical Pharmacology* 57, 27–36.
- Wu, Q., Xu-Cai, Y.O., Chen, S., Wang, W., 2009. Corin: New insights into the natriuretic peptide system. *Kidney International* 75, 142–146.
- Xu, D., Emoto, N., Giaid, A., *et al.*, 1994. ECE-1: A membrane-bound metalloprotease that catalyzes the proteolytic activation of big endothelin-1. *Cell* 78, 473–485.
- Yanagisawa, M., Kurihara, H., Kimura, S., *et al.*, 1988. A novel potent vasoconstrictor peptide produced by vascular endothelial cells. *Nature* 332, 411–415.
- Zois, N.E., Bartels, E.D., Hunter, I., *et al.*, 2014. Natriuretic peptides in cardiometabolic regulation and disease. *Nature Reviews Cardiology* 11, 403–412.

Further Reading

- Bernstein, K.E., Ong, F.S., Blackwell, W.L., *et al.*, 2012. A modern understanding of the traditional and nontraditional biological functions of angiotensin-converting enzyme. *Pharmacological Reviews* 65, 1–46.
- de Bold, A.J., 1987. On the shoulders of giants: The discovery of atrial natriuretic factor. *Canadian Journal of Physiology and Pharmacology* 65, 2007–2012.
- Clarke, N.E., Turner, A.J., 2012. Angiotensin-converting enzyme 2: The first decade. *International Journal of Hypertension* 2012, 307315.
- Davenport, A.P., Maguire, J.J., 2006. Endothelin. *Handbook of Experimental Pharmacology* 176 (Pt 1), 295–329.
- Dive, V., Chang, C.F., Yiotakis, A., Sturrock, E.D., 2009. Inhibition of zinc metallopeptidases in cardiovascular disease – from unity to trinity, or duality? *Current Pharmaceutical Design* 15, 3606–3621.
- Villela, D.C., Passos-Silva, D.G., Santos, R.A., 2014. Alamandine: A new member of the angiotensin family. *Current Opinion in Nephrology and Hypertension* 23, 130–134.
- Zois, N.E., Bartels, E.D., Hunter, I., *et al.*, 2014. Natriuretic peptides in cardiometabolic regulation and disease. *Nature Reviews Cardiology* 11, 403–412.

Cancer – Proteases in the Progression and Metastasis

K Ji and BF Sloane, Wayne State University School of Medicine, Detroit, MI, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

In 1946 Fischer proposed that tumor-associated proteolytic activity could be responsible for the degradation of extracellular matrix (ECM) and the subsequent invasion of the tumor into surrounding normal tissue (Fischer, 1946). Indeed proteases have been confirmed to play pro-tumor roles throughout malignant progression, starting with carcinogenesis (Edwards *et al.*, 2008; Mason and Joyce, 2011; Sloane *et al.*, 2013); however, some proteases are protective and antitumor (Lopez-Otin and Matrisian, 2007; Edwards *et al.*, 2008; Mason and Joyce, 2011; Sloane *et al.*, 2013). Interestingly, what we refer to as tumor proteases often are derived from stromal or immune cells that have infiltrated into the tumor rather than from the tumor cells (Edwards *et al.*, 2008; Mason and Joyce, 2011; Sloane *et al.*, 2013), thus broadening the challenge of targeting tumor proteases. The connectivity among the many proteases causally linked to malignant progression and what are the critical substrates cleaved by those proteases *in vivo* are not yet well understood (Doucet *et al.*, 2008; Doucet and Overall, 2008; Rodriguez *et al.*, 2010; Fortelny *et al.*, 2014). Substrates are not simply ECM proteins, but include numerous other proteins, for example, growth factors such as hepatocyte growth factor (HGF) (Lee *et al.*, 2000) and vascular endothelial growth factor (Lee *et al.*, 2005), adhesion molecules such as E-cadherin (David and Rajasekaran, 2012) and tenascin C (Mai *et al.*, 2002), other proteases (Edwards *et al.*, 2008; Mason and Joyce, 2011; Sloane *et al.*, 2013), signaling molecules (Lopez-Otin and Hunter, 2010), and endogenous protease inhibitors (Declercq *et al.*, 1997; Mason and Joyce, 2011; Laurent-Matha *et al.*, 2012), i.e., substrates that are activated or inhibited by limited proteolysis. In this chapter, we will briefly review data on the links between cancer progression and metastasis and a few selected proteases from the following five proteolytic clans: aspartic, cysteine, metallo, serine, and threonine (for further information see the MEROPS database: Rawlings *et al.*, 2014).

Aspartic Proteases

Cathepsin D

Cathepsin D, a lysosomal aspartic protease that is secreted by tumor cells, cleaves structural and functional proteins and peptides; inactivates chemokines such as macrophage inflammatory protein-1 alpha (Chemokine (C-C motif) ligand (CCL) 3), CCL4 and CCL21 (Wolf *et al.*, 2003); and cleaves prolactin (Piwnicka *et al.*, 2004) and osteopontin (Christensen *et al.*, 2010), thereby modulating their functions. Cathepsin D also enhances cysteine cathepsin activity in tumors by inactivating their endogenous inhibitor cystatin C (Laurent-Matha *et al.*, 2012). Cathepsin D, which is highly expressed in human breast carcinomas, is an independent marker of poor

prognosis (Foekens *et al.*, 1999; Fitzgibbons *et al.*, 2000; Rochefort *et al.*, 2000). Increases in levels and the secretion of cathepsin D are linked to size, grade, aggressiveness, metastasis and prognosis of endometrial, ovarian and breast cancers (Garcia *et al.*, 1996; Cunat *et al.*, 2004; Ioachin, 2005; Liaudet-Coopman *et al.*, 2006). All of these may be due to the stimulation of tumor cell proliferation by cathepsin D (Garcia *et al.*, 1990; Liaudet *et al.*, 1994, 1995; Glondou *et al.*, 2001), an effect that is not mediated by its proteolytic activity, but may be mediated by binding to yet unidentified cell surface receptors (Glondou *et al.*, 2001; Masson *et al.*, 2010). Thus, targeting the proteolytic activity of cathepsin D would not totally abrogate its pro-tumorigenic functions.

Cysteine Proteases

Calpains

Calpains are a family of intracellular calcium-dependent cysteine proteases, which cleave a broad spectrum of substrates including the cytoskeletal proteins talin, vinculin and desmin; membrane proteins (growth factor receptors, adhesion molecules, ion transporters); transcription factors (p53, c-fos); and enzymes (caspases, protein kinase C, RhoA/Rac) (Goll *et al.*, 2003; Anagli *et al.*, 2013). Calpains are involved in cancer-related processes such as cell motility, apoptosis and cell cycle regulation (Carafoli and Molinari, 1998; Ono *et al.*, 1998; Zatz and Starling, 2005). Increased expression of calpain 1 has been observed in schwannomas and meningiomas (Kimura *et al.*, 1998), renal cell carcinomas (Braun *et al.*, 1999) and breast carcinomas (Shiba *et al.*, 1996); calpain 2 in colorectal carcinomas (Lakshmikuttyamma *et al.*, 2004) and prostate carcinomas (Rios-Doria *et al.*, 2003); and calpain 6 in uterine sarcomas and carcinosarcomas (Lee *et al.*, 2007) and cervical cancers (Lee *et al.*, 2008). In contrast, calpain 9 is decreased in gastric cancer (Yoshikawa *et al.*, 2000). This altered expression of calpain family members in tumors would be consistent with a role for these enzymes in tumor progression (Storr *et al.*, 2011).

Caspases

One of the best-known proteolytic pathways is apoptosis in which caspases initiate and execute programmed cell death (Flutsch and Grutter, 2013). Indeed caspases are classified as either initiator caspases (caspase-2, -8, -9, and -10) or executioner caspases (caspase-3, -6, and -7). Once activated by initiator caspases, executioner caspases cleave a broad spectrum of cellular proteins, ultimately resulting to cell death (Nicholson, 1999). Failure of apoptosis could allow the survival of transformed cells; indeed, experimental evidence from transgenic mice and human cell lines shows that caspases-2, -3, and -8 act as tumor suppressors in colorectal, gastric, breast,

and lung cancers (Jager and Zwacka, 2010; Olsson and Zhivotovsky, 2011; Puccini *et al.*, 2013). Mutations of caspases-3 and -6 are observed in human gastric, colorectal and non-small cell lung cancers and multiple myeloma (Soung *et al.*, 2004; Lee *et al.*, 2006). In addition, gene polymorphisms for caspase-3, -7, -8, and -9 that alter caspase activity are associated with risks for lung (Son *et al.*, 2006; Lee *et al.*, 2010) and breast cancers (Theodoropoulos *et al.*, 2012). Thus, genetic alterations that disrupt caspase activity appear to enhance cancer development:

Cysteine Cathepsins

The human cysteine cathepsin family is comprised of 11 members: cathepsins B, C, F, H, K, L, O, S, V, W, X/Z (Mohamed and Sloane, 2006). They are often secreted from tumor cells and may associate with binding partners on the cell surface (Mohamed and Sloane, 2006; Victor and Sloane, 2007). Causal roles for lysosomal cysteine cathepsins in human and murine cancers are well documented; however, there is a need to define the critical substrates cleaved (Mohamed and Sloane, 2006; Gocheva and Joyce, 2007; Victor and Sloane, 2007; Turk *et al.*, 2012). There is functional redundancy and cooperativity among the cysteine cathepsins. For example, redundancy has been observed for cathepsins B and X: ablation of cathepsin B in breast tumor cells that express this enzyme on their surface is compensated for by a relocalization of cathepsin X to the cell surface with either of the two cysteine cathepsins supporting invasion of the breast tumor cells (Vasiljeva *et al.*, 2006). Cooperativity in breast cancer progression and metastasis has been shown for cathepsins B and X/Z in a cell type known to contribute to malignant progression, i.e., carcinoma-associated fibroblasts (Orimo and Weinberg, 2006; Shimoda *et al.*, 2010). Ablation of both enzymes increases secretion of matrix metalloproteinase (MMP)-2, N-cadherin and neural cell adhesion molecule and decreases secretion of periostin by fibroblasts, changes that are consistent with the two enzymes playing causal and cooperative roles in cell adhesion and ECM remodeling. Individual ablation of cathepsin B or L has a more limited impact on the fibroblast secretome (Tholen *et al.*, 2014). Cysteine cathepsins exhibit cooperation with other proteases as well. For example, cathepsin B can trigger the plasminogen activation cascade (Kobayashi *et al.*, 1991; Guo *et al.*, 2002) and activate MMP2 and 9 in tumor-shed microvesicles or exosomes (Giusti *et al.*, 2008). Most cysteine cathepsins require acidic pH for optimal activity, raising a question as to their roles pericellularly; however, solid tumors acidify their microenvironment and cysteine cathepsins have been demonstrated to degrade ECM surrounding tumors *in vitro* and *in vivo* (Robey *et al.*, 2009; Estrella *et al.*, 2013; Rothberg *et al.*, 2013). The activation of MMP2 and 9 in tumor exosomes by cathepsin B also is increased at acidic pH (Giusti *et al.*, 2008).

Elevated expression and activity of cysteine cathepsins have been shown to be associated with cancer progression in a wide variety of tumors (Jedezsko and Sloane, 2004; Mohamed and Sloane, 2006). In transgenic mouse models, multiple cysteine cathepsins have been confirmed to play causal roles in malignant progression of pancreatic islet cell and mammary tumors (Joyce *et al.*, 2004; Reinheckel *et al.*, 2012; Withana

et al., 2012) and cysteine cathepsin inhibitors shown to reduce tumor growth, invasiveness, and metastasis (Reinheckel *et al.*, 2012; Withana *et al.*, 2012), either alone or in combination with chemotherapy (Bell-McGuinn *et al.*, 2007; Shree *et al.*, 2011; Withana *et al.*, 2012). There are examples of tissue specificity: cathepsin B enhances progression of pancreatic islet cell tumors, pancreatic ductal adenocarcinomas and mammary carcinomas (Joyce *et al.*, 2004; Vasiljeva *et al.*, 2006; Gopinathan *et al.*, 2012), but not squamous cell carcinomas (Ruffell *et al.*, 2013). In contrast, cathepsin C regulates squamous cell carcinogenesis, but not mammary carcinogenesis (Ruffell *et al.*, 2013). Cysteine cathepsins may also be protective, as shown by the ablation of cathepsin L resulting in promotion of squamous cell carcinogenesis (Dennemark *et al.*, 2010). ‘Tumor’ cysteine cathepsins that contribute to malignant progression may come from tumor cells, infiltrating immune cells such as macrophages, infiltrating fibroblasts and/or endothelial cells (Vasiljeva *et al.*, 2006; Gounaris *et al.*, 2008; Joyce and Pollard, 2009; Herroon *et al.*, 2013; Small *et al.*, 2013). Since cysteine cathepsins from either tumor cells or other cells in the tumor microenvironment can modulate malignant progression, this increases the difficulty of targeting these enzymes in cancer.

Metalloproteases

ADAM and AMAMTS Proteases

The metalloproteases known as a disintegrin and metalloproteinase (ADAM) and ADAM with thrombospondin motifs (ADAMTS) are named after their domains: A Disintegrin And Metalloprotease and ADAM with ThromboSpondin motifs. The 21 human ADAM proteases are transmembrane and secreted proteins and have both an adhesion domain and a protease domain (Wolfsberg *et al.*, 1995; Przemyslaw *et al.*, 2013). In contrast, the 19 human ADAMTS proteases have in addition multiple thrombospondin 1-like repeats and are secreted proteins (Tang, 2001; Przemyslaw *et al.*, 2013). Dysregulated expression of ADAMs and ADAMTSs has been reported in a wide range of human cancers, where they are implicated as regulators of cancer progression that function in ECM remodeling, shedding of adhesion molecules and regulation of activity of growth factors, integrins and cytokines (Rocks *et al.*, 2008; Mentlein *et al.*, 2012; Przemyslaw *et al.*, 2013). Their functions include the shedding of tumor necrosis factor (TNF)- α by ADAM17 that is associated with the invasiveness of triple negative breast cancers (Giricz *et al.*, 2013) and the shedding of epidermal growth factor receptor (EGFR) ligands by ADAM17 in breast cancer (Kenny and Bissell, 2007). ADAM17 (Merchant *et al.*, 2008) and ADAMTS1 (Lind *et al.*, 2006), which is epigenetically deregulated, are implicated in colon cancer.

Antitumor roles for ADAMs/ADAMTSs are also observed: ADAMTS15 is genetically inactivated in colon tumor cells (Viloria *et al.*, 2009) and ADAMTS12 epigenetically silenced in colon tumor cells and transcriptionally activated in carcinoma-associated fibroblasts (Moncada-Pazos *et al.*, 2009), which is linked to an anti-proliferative effect on the colon tumor cells. ADAMTS1 and 18 exhibit anti-angiogenic and anti-metastatic

properties in several cancer types including breast and pancreatic cancer (Masui *et al.*, 2001; Porter *et al.*, 2004). ADAM28 has been proposed to be an oncogene in asbestos-related lung cancers (Wright *et al.*, 2010) and high expression of active ADAM28 has been observed in non-small cell lung carcinomas that metastasize to lymph nodes (Ohtsuka *et al.*, 2006). Cooperativity with other proteases may be required: MMP7 has been shown to activate ADAM28 (Mochizuki *et al.*, 2004). There are ongoing clinical trials with selective ADAM inhibitors in breast cancer (Duffy *et al.*, 2011), but such small molecule inhibitors are selective rather than specific and thus anticipated to have off-target effects (Coussens *et al.*, 2002) for discussion of problems encountered in use of broad-spectrum MMP inhibitors). Murphy and colleagues (Murphy, 2010; Richards *et al.*, 2012) are testing an alternative approach to inhibition of ADAM17 using blocking antibodies. Although the antibodies are specific, compensation by other proteases is observed when these are used *in vivo*, once again indicating the difficulty of targeting proteases therapeutically. More extensive discussions of ADAMs and ADAMTSs in cancer can be found in the following reviews (Murphy, 2008; Kataoka, 2009; Turner *et al.*, 2009; Klein and Bischoff, 2011; Wagstaff *et al.*, 2011; Kumar *et al.*, 2012; Grotzinger and Rose-John, 2013).

Matrix Metalloproteinases

The original link of MMPs to cancer came from studies that demonstrated their ability to degrade type IV collagen in the basement membrane (Liotta *et al.*, 1979, 1980). Indeed MMPs are often classified by their specificity for ECM substrates, for example, collagenases (MMP1, 8, 13) that degrade native fibrillar collagen; stromelysins (MMP3, 10, 11) that degrade proteoglycans and glycoproteins such as laminin and fibronectin; and gelatinases (MMP2, 9) that degrade type IV collagen and denatured fibrillar collagen. Nonetheless, MMPs cleave many other substrates that are critical to malignant progression, for example, growth factors, cytokines, receptors, protease inhibitors and adhesion molecules (Egeblad and Werb, 2002; Sloane *et al.*, 2013). Six of the 23 human MMPs are associated with the cell membrane: four are transmembrane and two are glycosylphosphatidylinositol (GPI)-anchored (Hernandez-Barrantes *et al.*, 2002; Sohail *et al.*, 2008). Expression of many MMPs is highly increased in human cancers and correlates with invasiveness, metastasis and poor prognosis (Coussens *et al.*, 2002; Egeblad and Werb, 2002). Causal roles in malignant progression have been established for a wide variety of MMPs by studies in which murine transgenic models for cancers are crossed into mice deficient in individual MMPs. Such studies have demonstrated causal roles at stages that span the spectrum from tumor initiation through metastasis. These pro-tumorigenic activities of MMPs have been recently reviewed in detail (Sloane *et al.*, 2013) and are also delineated in numerous review articles (see e.g., Seiki, 1999; Fillmore *et al.*, 2001; Edwards *et al.*, 2008; Sohail *et al.*, 2008; Kessenbrock *et al.*, 2010; Khokha and Werb, 2011; Sloane *et al.*, 2013). Pro-tumor activities depend not only on an individual MMP, but may depend on the cell that is expressing the MMP. This was shown for colon tumors using a model in which the mice are predisposed to develop intestinal adenomas; in these mice development of intestinal tumors is

linked to MMP7 in the tumor cells (Fingleton *et al.*, 1999) and to MMP9 in tumor-associated neutrophils (Sinnamon *et al.*, 2008), but not to MMP2, 12, or 19 expressed in either tumor cells or tumor-associated cells. Lung metastasis of murine mammary carcinomas has been linked to ablation of MMP9 in tumor-associated neutrophils, but whether this does or does not occur depends on the genetic background of the mice (Martin *et al.*, 2008). Thus, even the propensity of some MMPs to induce tumors and enhance progression is not straightforward, making therapeutic targeting of these enzymes difficult.

The literature reporting pro-tumorigenic roles for MMPs is extensive, yet some MMPs act as tumor suppressors (Lopez-Otin and Matrisian, 2007). Of the 24 murine MMPs, 10 have antitumor or anti-inflammatory properties (Dufour and Overall, 2013). Ablation of either MMP3 or MMP8 strongly increases skin tumorigenesis in mice (Balbin *et al.*, 2003; Mccawley *et al.*, 2004), an effect that in one study occurred in male mice, but not female mice, unless they were ovariectomized or treated with tamoxifen (Balbin *et al.*, 2003). MMP8 inactivates pro-inflammatory cytokines and chemokines, a result that is consistent with the MMPs that are protective being expressed primarily in inflammatory cells associated with the skin tumors. Loss-of-function mutations that abrogate MMP8 activity have been found in human melanomas where they are associated with enhanced progression (Lopez-Otin *et al.*, 2009), whereas increases in MMP8 expression in human breast and oral carcinomas are associated with better prognosis (Lopez-Otin *et al.*, 2009). MMP3, as referenced above, can be antitumor in the skin, yet also can be pro-tumor as in murine mammary carcinomas and pancreatic ductal adenocarcinomas where it plays a role in tumor initiation (Sternlicht *et al.*, 1999; Lopez-Otin *et al.*, 2009). MMP3 also has been shown to be protective (Witty *et al.*, 1995), in this case in another mouse mammary carcinoma that is chemically-induced. Other MMPs shown to have dual roles in cancer are MMP9, 11, and 19 (Lopez-Otin and Matrisian, 2007). These findings indicate the complexities of targeting MMPs in cancer progression and metastasis and may account in part for the failure of broad-spectrum MMP inhibitors in clinical trials (Coussens *et al.*, 2002) and indicate the need to design inhibitors so that they do not target MMPs that have antitumor and anti-inflammatory roles (Dufour and Overall, 2013).

MMPs like cysteine cathepsins (see above) and serine proteases (see below) exhibit redundancy and cooperativity. Ablation of individual MMP genes, with the exception of MMP14 or MT1-MMP (Holmbeck *et al.*, 1999), does not result in an overt phenotype, findings consistent with there being functional redundancy among these enzymes (Page-Mccaw *et al.*, 2007). Proteases other than MMPs have been found to compensate when there is a deficiency in an MMP. For example, a surprising finding is that two mouse models of pancreatic neuroendocrine carcinomas become more invasive when MMP9 is ablated. This was shown to be due to an increase in the cysteine protease cathepsin B as a result of infiltration of a cathepsin B-expressing subpopulation of inflammatory cells at the invasive front of the tumors (Shchors *et al.*, 2013). Proteases of different clans, in this case MT1-MMP and the cysteine protease cathepsin K, cooperate to

promote invasion through collagen I by breast cancer cells (Montgomery *et al.*, 2012). This is not surprising given that these two enzymes have been shown to function in mesenchymal cells and macrophages in complementary pathways of collagen I degradation (Madsen *et al.*, 2007; Madsen and Bugge, 2013; Fonovic and Turk, 2014). However, it does not agree with the contention by Weiss and colleagues that only MT1-MMP is required for collagen I invasion by human mesenchymal stem cells (Lu *et al.*, 2010) and for the focal collagenolysis that supports the invasive phenotype of cancer cells (Sabeh *et al.*, 2009). These disparate results may reflect differences in the conformation of the collagen matrices used by investigators in *in vitro* analyses of collagen degradation as proposed by Weiss and colleagues; however, the two complementary pathways are observed *in vivo*. Yet another intriguing cooperativity has been observed among MMPs (MMP3, 10, and 13) in regard to degradation of collagen, in this case in a rat mammary carcinoma cell line. This cooperativity involves crosstalk between cells that retain epithelial characteristics and ones that have undergone an epithelial-mesenchymal transition (Martorana *et al.*, 1998). To understand the roles of MMPs in malignant progression, we will need to define the *in vivo* substrates that are critical to progression and the cellular sources of the MMPs that degrade those substrates as well as pathways for cooperativity with other MMPs or with proteases of other clans (see also Concluding Remarks).

Serine Proteases

There are 176 putative serine proteases (Puente *et al.*, 2003) in the human genome, a large majority of which have been implicated in cancer (see Netzel-Arnett *et al.*, 2003; Emami and Diamandis, 2008; Oikonomopoulou *et al.*, 2010; Webb *et al.*, 2011; Batra *et al.*, 2012; Liu *et al.*, 2012 for reviews). Due to space limitations, we will limit our review of serine proteases in cancer to the plasminogen activation system and one of the transmembrane serine proteases matriptase, which has oncogenic activity (see also Sloane *et al.*, 2013 for a more detailed discussion).

Plasminogen Activation System

The plasminogen activation system has been demonstrated to play a causal role in virtually all cancers (Dano *et al.*, 2005), including in their metastasis (Sidenius and Blasi, 2003; Dano *et al.*, 2005; Sloane *et al.*, 2013). The penultimate protease is plasmin, which is generated from plasminogen by one of two plasminogen activators: tissue plasminogen activator or urokinase plasminogen activator (uPA). The latter has been implicated in cancer and is often considered to be a therapeutic target that is preferable to plasmin (Hildenbrand *et al.*, 2010). uPA is itself a serine protease, associates with the cell surface by binding to the GPI-anchored receptor urokinase plasminogen activator receptor (uPAR), is inhibited by the plasminogen activator inhibitors-1 (PAI-1) or PAI-2 and is activated from pro-uPA by plasmin (Bekes *et al.*, 2011) via a feedback loop or other proteases such as the cysteine protease cathepsin B (Guo *et al.*, 2002). Bound uPA activates plasminogen more

efficiently than soluble uPA and focuses the proteolytic activity of plasmin to regions adjacent to the tumor cell surface. Plasmin has a broad substrate specificity and can degrade ECM proteins such as laminin and fibronectin directly (Andreasen *et al.*, 1997) as well as indirectly by activating MMP3, 9, 12, and 13 (Carmeliet *et al.*, 1997; Mason and Joyce, 2011). The various components of the plasminogen activation system are not necessarily expressed by tumor cells; stromal cells are for example the source of uPA in colon tumors and in addition which cells are the source of the various components of this system differs from one type of tumor to another (Dano *et al.*, 2005).

Studies in which mice deficient in plasminogen activation system components are crossed with transgenic models for cancer, reviewed in (Sloane *et al.*, 2013), have confirmed roles for uPA and plasminogen in invasion and metastasis of mammary carcinomas, but not in growth of the primary tumors; furthermore, neither uPAR or PAI-1 deficiency was found to have an effect. In contrast, ablation of uPA did not affect pancreatic islet cell tumorigenesis or the angiogenic switch that accompanies progression of these tumors. Ablation of PAI-1 does not affect skin carcinogenesis, but on the other hand inhibition of PAI-1 by small molecule inhibitors suppresses intestinal tumorigenesis (Hildenbrand *et al.*, 2010), indicating differences from one tumor type to another. The apparent lack of effects in regard to angiogenesis are somewhat surprising given the key roles established for the plasminogen activation system in the neoangiogenesis that is associated with prostate carcinoma cells *in vivo* (Conn *et al.*, 2009). Quigley and colleagues have suggested that the most effective therapeutic approach would be to target activation of pro-uPA at the apex of the plasminogen cascade through use of activation blocking antibodies rather than through use of protease inhibitors (Bekes *et al.*, 2011). In a recent review they discuss whether plasmin exerts its pro-tumor effects primarily through cleavage of cell surface proteins rather than ECM remodeling and ways that these cell surface proteins might be targeted to prevent cleavage and thereby reduce metastasis (Deryugina and Quigley, 2012). Thus, despite the extensive literature on the plasminogen activation system in cancer, reagents that will successfully target this system remain elusive.

Type II Transmembrane Trypsin-like Serine Proteases: Matriptase

Proteases of the TTSP family have been implicated in cancer initiation and progression (Sloane *et al.*, 2013). There are 17 human TTSPs (Netzel-Arnett *et al.*, 2003). The member that has been most studied is matriptase (also known as MT-SP1, TADG-15, epithin, and ST14), which was originally defined as a secreted gelatinase (Shi *et al.*, 1993). Matriptase is expressed in epithelial cells in breast, skin, prostate, ovary, stomach, and lung (Kim *et al.*, 1999; Takeuchi *et al.*, 1999; Oberst *et al.*, 2001; Tanimoto *et al.*, 2001) and plays a key role in the formation and maintenance of epithelial integrity (Carney *et al.*, 2007; List *et al.*, 2009). Substrates cleaved by matriptase include HGF, uPA (Lee *et al.*, 2000), MMP3 (Jin *et al.*, 2006), protease-activated receptor-2 (Takeuchi *et al.*, 2000), prostasin, macrophage stimulating protein-1, epithelial sodium channels (Andreasen *et al.*, 2006), acid-sensing ion

channel-1 (Clark *et al.*, 2010), fibronectin, and laminin (Satomi *et al.*, 2001), accounting for the role of this enzyme in cell growth, motility, survival, ECM remodeling, and skin barrier function under both physiological and pathological conditions.

A causal role for matriptase in initiation and progression of cancer has been demonstrated using a transgenic mouse model in which overexpression of matriptase is targeted to the skin. These mice develop spontaneous squamous cell carcinomas, consistent with an oncogenic function for matriptase, and are susceptible to carcinogen-induced tumor formation (List *et al.*, 2005). This oncogenic activity of matriptase can be abrogated by engineering expression of the cognate matriptase inhibitor, hepatocyte growth factor activator inhibitor (HAI)-1 in the basal layer of the mouse skin (List *et al.*, 2005). Matriptase can contribute to metastasis of gastric cancer cells: cells that have been engineered to express high levels of active matriptase exhibit an increase in metastasis to the lymph nodes (Ihara *et al.*, 2002). The pro-oncogenic and pro-metastatic activity of matriptase is proposed to be due to its activation of pathways linked to malignant progression, for example, HGF (Szabo *et al.*, 2011) and uPA (Ried *et al.*, 1999; Moriyama *et al.*, 1999; Lee *et al.*, 2000; Suzuki *et al.*, 2004). Indeed, c-Met-induced squamous cell carcinogenesis is initiated by matriptase (Szabo *et al.*, 2011). There is an exquisite balance between matriptase and HAI-1 that correlates with the grade of malignancy and is predictive of survival in human cancers including breast, cervical, ovarian, pancreatic, and prostate (Oberst *et al.*, 2002; Kang *et al.*, 2003; Uhland, 2006; List, 2009). In some tissues, matriptase appears to have an antitumor effect as downregulation of both matriptase and HAI-1 has been reported in gastrointestinal cancer (Zeng *et al.*, 2005). Targeting matriptase or downstream signaling pathways is being pursued, yet the possible antitumor effects in some tissues indicate a need to further define these pathways.

Threonine Proteases

Proteasomes

Proteasomes are multi-catalytic protease complexes engaged in recycling of intracellular proteins of short life span. These protein degradation machines play a role in elimination of cellular proteins that have been tagged for degradation through a complex modification termed polyubiquitination. More than 80% of cellular proteins are degraded through this pathway including those involved in an array of processes such as cell cycle, apoptosis, transcription, DNA repair, protein quality control, and antigen presentation (Dahlmann, 2007; Demasi and Laurindo, 2012). The 26S proteasome consists of a 20S proteolytic and a 19S regulatory subunit; there are three catalytic activities in the 20S subunit: caspase-like, trypsin-like, and chymotrypsin-like. The ubiquitin-proteasome pathway regulates signal transduction pathways that contribute to initiation and progression of cancer (Fuchs, 2002); thus, targeting the ubiquitin-proteasome pathway would appear to be a potential anti-cancer strategy. For more detailed information on the association of proteasomes with cancer and the potential of targeting proteasomes for cancer chemotherapy (please see Almond and Cohen, 2002; Sloane *et al.*, 2013).

Bortezomib, dipeptide boronic acid, was the first synthetic proteasome inhibitor (Adams *et al.*, 1998) and is approved by

the FDA for the treatment of multiple myeloma, pancreatic, and thyroid cancer (Kane *et al.*, 2003; Nawrocki *et al.*, 2005; Wunderlich *et al.*, 2012). Bortezomib targets the chymotryptic activity of the 20S proteasome. More recently compounds such as Carfilzomib (PR-171), Salinosporamide A (NP-0052), and MLN9708 have shown promising results in clinical trials (Chari *et al.*, 2010; Demo *et al.*, 2007). Moreover, many such compounds are also being tested in preclinical studies on cancer cell lines (Milacic *et al.*, 2008; Piva *et al.*, 2008; Yang *et al.*, 2009; Dai *et al.*, 2010). These proteasome inhibitors inhibit cell proliferation, induce cell cycle arrest, autophagy, and apoptosis.

One of the well-known functions of the proteasome is regulation of the transcription factor nuclear factor kappa B (NF κ B). NF κ B regulates a variety of inflammatory responses, including activation of pro-inflammatory cytokines that promote tumorigenesis (Cortes Sempere *et al.*, 2008; Grivennikov and Karin, 2010; Ben-Neriah and Karin, 2011). Inhibition of the proteasome by bortezomib induces accumulation of the inhibitor of NF κ B (I κ B), resulting in downregulation of transcriptional targets of NF κ B. However, in endometrial carcinoma cells and primary explants proteasome inhibitors activate NF κ B instead of blocking its transcription (Dolcet *et al.*, 2006). In multiple myeloma cells, bortezomib has been shown to activate NF κ B by triggering two upstream NF κ B activating kinases (I κ B kinase beta and receptor-interacting protein 2) in multiple myeloma cells (Hideshima *et al.*, 2009). These conflicting results may result in off-target effects of proteasome inhibitors like bortezomib and may contribute to development of resistance to proteasome inhibitors. Nonetheless the proteasome remains a viable target for cancer therapy and proteasome inhibitors an important example of protease inhibitors that have shown efficacy in the clinic.

Concluding Remarks

There is considerable evidence that proteases play critical roles in cancer progression, targeting a diversity of substrates and participating in steps from tumor initiation to metastasis. Here we have selected a few proteases for discussion and tried to indicate the connectivity between proteases of a single clan as well as between proteases of multiple clans. There are data on substrate specificity for many of the human proteases, but these data are for the most part from *in vitro* analyses, often carried out under conditions that would not be encountered by proteases secreted from tumor cells or associated with the tumor cell surface such as hypoxia and acidosis (Estrella *et al.*, 2013; Rothberg *et al.*, 2013). Thus there is clearly a need to identify the substrates cleaved by proteases *in vivo*. Some investigators have emphasized the importance of using protein substrates that are in their native conformation to the accurate identification of the proteases that cleave those substrates, for example, native type I collagen (Sabeh *et al.*, 2009). More than one protease is likely to cleave any given substrate and a deletion of one protease is likely to result in compensatory changes in other proteases that can cleave that substrate. Furthermore, proteases cleave more than one substrate, as is particularly true for the promiscuous lysosomal cysteine cathepsins.

Overall and colleagues (Fortelny *et al.*, 2014) have assembled curated and publically-available protease cleavage

and inhibition data from 4 of the 5 clans of proteases discussed here (aspartic, cysteine, metallo, and serine) into a computational representation that they term the protease web. This representation establishes that proteases do not act alone or in isolated hierarchical cascades such as the plasminogen activation cascade. Instead they operate in interactive networks that are multi-directional. Of note is that endogenous protease inhibitors are key components of the protease web and that a limited number of proteases form nodes in the web. This may reflect a larger literature on proteases that have been studied for years, rather than a critical role for those proteases. From the protease web, we can see that proteases can regulate the activity of proteases of a different clan; for example, metalloproteases were found to cleave serine proteases or their cognate inhibitors, the serpins. Other factors that need to be considered in regard to the protease web are temporal and dynamic changes that occur in protease expression and localization, changes that may influence the activation or inhibition of the proteases. These computational analyses are as of yet incomplete. Much of the information on substrates and inhibition was obtained in studies performed with purified proteases *in vitro* so there was a preselection of potential substrates rather than an unbiased global approach to substrate identification. Furthermore, the studies were based primarily on analysis of cleavage of single substrates rather than cleavage of a substrate in a milieu in which cleavage might be influenced by the presence of competing substrates. Cellular and subcellular sources of the proteases and inhibitors and how that might alter the connectivities observed also needs to be addressed. When targeting proteases therapeutically, we need to consider the protease web and how targeting one or a family of proteases may lead to unanticipated results. Nonetheless, proteases are viable therapeutic targets due to the alterations in their expression and localization that accompany tumor development and progression.

See also: Interorganellar Communication: Interplay and Processes: Endoplasmic Reticulum Stress in Disease. Protein Synthesis/Degradation: Protein Degradation – Intracellular: Role of Lysosomes in Intracellular Degradation; Ubiquitin, Ubiquitin-Like Proteins, and Proteasome-Mediated Degradation. Protein Synthesis/Degradation: Protein Degradation – Protease Classes: ADAMs Regulate Cell–Cell Interactions by Controlling the Function of the EGF-Receptor, TNF α and Notch; ADAMTS Proteases: Mediators of Physiological and Pathogenic Extracellular Proteolysis; Extracellular: Plasma Membrane Proteases – Serine Proteases; Matrix Metalloproteinases; Naturally-Occurring Polypeptide Inhibitors: Cystatins/Stefins, Inhibitors of Apoptosis (IAPs), Serpins, and Tissue Inhibitors of Metalloproteinases (TIMPs); The Calpain Proteolytic System

References

- Adams, J., Behnke, M., Chen, S., *et al.*, 1998. Potent and selective inhibitors of the proteasome: Dipeptidyl boronic acids. *Bioorganic & Medicinal Chemistry Letters* 8, 333–338.
- Almond, J.B., Cohen, G.M., 2002. The proteasome: A novel target for cancer chemotherapy. *Leukemia* 16, 433–443.
- Anagli, J., Wang, K., Ono, Y., Sorimachi, H., 2013. Calpains in health and disease. In: Brix, K., Stocker, W. (Eds.), *Proteases: Structure and Function*. Vienna, Austria: Springer-Verlag GmbH (Chapter 12).
- Andreasen, D., Vuagniaux, G., Fowler-Jaeger, N., Hummler, E., Rossier, B.C., 2006. Activation of epithelial sodium channels by mouse channel activating proteases (MCAP) expressed in xenopus oocytes requires catalytic activity of MCAP3 and MCAP2 but not MCAP1. *Journal of the American Society of Nephrology* 17, 968–976.
- Andreasen, P.A., Kjoller, L., Christensen, L., Duffy, M.J., 1997. The urokinase-type plasminogen activator system in cancer metastasis: A review. *International Journal of Cancer* 72, 1–22.
- Balbin, M., Fueyo, A., Tester, A.M., *et al.*, 2003. Loss of collagenase-2 confers increased skin tumor susceptibility to male mice. *Nature Genetics* 35, 252–257.
- Batra, J., O'mara, T., Patnala, R., Lose, F., Clements, J.A., 2012. Genetic polymorphisms in the human tissue kallikrein (KLK) locus and their implication in various malignant and non-malignant diseases. *Biological Chemistry* 393, 1365–1390.
- Bekes, E.M., Deryugina, E.I., Kupriyana, T.A., *et al.*, 2011. Activation of pro-uPA is critical for initial escape from the primary tumor and hematogenous dissemination of human carcinoma cells. *Neoplasia* 13, 806–821.
- Bell-McGuinn, K.M., Garfall, A.L., Bogoy, M., Hanahan, D., Joyce, J.A., 2007. Inhibition of cysteine cathepsin protease activity enhances chemotherapy regimens by decreasing tumor growth and invasiveness in a mouse model of multistage cancer. *Cancer Research* 67, 7378–7385.
- Ben-Neriah, Y., Karin, M., 2011. Inflammation meets cancer, with NF-kappaB as the matchmaker. *Nature Immunology* 12, 715–723.
- Braun, C., Engel, M., Seifert, M., *et al.*, 1999. Expression of calpain I messenger RNA in human renal cell carcinoma: Correlation with lymph node metastasis and histological type. *International Journal of Cancer* 84, 6–9.
- Carafoli, E., Molinari, M., 1998. Calpain: A protease in search of a function? *Biochemical and Biophysical Research Communications* 247, 193–203.
- Carmeliet, P., Moons, L., Lijnen, R., *et al.*, 1997. Urokinase-generated plasmin activates matrix metalloproteinases during aneurysm formation. *Nature Genetics* 17, 439–444.
- Carney, T.J., Von Der Hardt, S., Sonntag, C., *et al.*, 2007. Inactivation of serine protease Matriptase1a by its inhibitor Hai1 is required for epithelial integrity of the zebrafish epidermis. *Development* 134, 3461–3471.
- Chari, A., Mazumder, A., Jagannath, S., 2010. Proteasome inhibition and its therapeutic potential in multiple myeloma. *Biologics* 4, 273–287.
- Christensen, B., Schack, L., Klanning, E., Sorensen, E.S., 2010. Osteopontin is cleaved at multiple sites close to its integrin-binding motifs in milk and is a novel substrate for plasmin and cathepsin d. *Journal of Biological Chemistry* 285, 7929–7937.
- Clark, E.B., Jovov, B., Rooj, A.K., Fuller, C.M., Benos, D.J., 2010. Proteolytic cleavage of human acid-sensing ion channel 1 by the serine protease matriptase. *Journal of Biological Chemistry* 285, 27130–27143.
- Conn, E.M., Botkjaer, K.A., Kupriyana, T.A., *et al.*, 2009. Comparative analysis of metastasis variants derived from human prostate carcinoma cells: Roles in intravasation of VEGF-mediated angiogenesis and uPA-mediated invasion. *American Journal of Pathology* 175, 1638–1652.
- Cortes Sempere, M., Rodriguez Fanjul, V., Sanchez Perez, I., Perona, R., 2008. The role of the NFkappaB signalling pathway in cancer. *Clinical and Translational Oncology* 10, 143–147.
- Coussens, L.M., Fingleton, B., Matrisian, L.M., 2002. Matrix metalloproteinase inhibitors and cancer: Trials and tribulations. *Science* 295, 2387–2392.
- Cunat, S., Hoffmann, P., Pujol, P., 2004. Estrogens and epithelial ovarian cancer. *Gynecologic Oncology* 94, 25–32.
- Dahlmann, B., 2007. Role of proteasomes in disease. *BMC Biochemistry* 8 (Suppl. 1), S3.
- Dai, Y., Desano, J., Tang, W., *et al.*, 2010. Natural proteasome inhibitor celastrol suppresses androgen-independent prostate cancer progression by modulating apoptotic proteins and NF-kappab. *PLoS One* 5, e14153.
- Dano, K., Behrendt, N., Hoyer-Hansen, G., *et al.*, 2005. Plasminogen activation and cancer. *Thrombosis and Haemostasis* 93, 676–681.
- David, J.M., Rajasekaran, A.K., 2012. Dishonorable discharge: The oncogenic roles of cleaved E-cadherin fragments. *Cancer Research* 72, 2917–2923.
- Declerck, Y.A., Imren, S., Montgomery, A.M., *et al.*, 1997. Proteases and protease inhibitors in tumor progression. *Advances in Experimental Medicine and Biology* 425, 89–97.
- Demasi, M., Laurindo, F.R., 2012. Physiological and pathological role of the ubiquitin-proteasome system in the vascular smooth muscle cell. *Cardiovascular Research* 95, 183–193.

- Demo, S.D., Kirk, C.J., Aujay, M.A., *et al.*, 2007. Antitumor activity of PR-171, a novel irreversible inhibitor of the proteasome. *Cancer Research* 67, 6383–6391.
- Dennekamp, J., Lohmuller, T., Mayerle, J., *et al.*, 2010. Deficiency for the cysteine protease cathepsin L promotes tumor progression in mouse epidermis. *Oncogene* 29, 1611–1621.
- Deryugina, E.I., Quigley, J.P., 2012. Cell surface remodeling by plasmin: A new function for an old enzyme. *Journal Biomedicine Biotechnology* 2012 (2012), Article ID 564259, 21 pp. Available at: <http://dx.doi.org/10.1155/2012/564259> (accessed 22.12.14).
- Dolcet, X., Llobet, D., Encinas, M., *et al.*, 2006. Proteasome inhibitors induce death but activate NF-kappaB on endometrial carcinoma cell lines and primary culture explants. *Journal of Biological Chemistry* 281, 22118–22130.
- Doucet, A., Butler, G.S., Rodriguez, D., Prudova, A., Overall, C.M., 2008. Metadegradomics: Toward *in vivo* quantitative degradomics of proteolytic post-translational modifications of the cancer proteome. *Molecular & Cellular Proteomics* 7, 1925–1951.
- Doucet, A., Overall, C.M., 2008. Protease proteomics: Revealing protease *in vivo* functions using systems biology approaches. *Molecular Aspects of Medicine* 29, 339–358.
- Duffy, M.J., Mullooly, M., O'donovan, N., *et al.*, 2011. The ADAMs family of proteases: New biomarkers and therapeutic targets for cancer? *Clinical Proteomics* 8, 9.
- Dufour, A., Overall, C.M., 2013. Missing the target: Matrix metalloproteinase antitargets in inflammation and cancer. *Trends in Pharmacological Sciences* 34, 233–242.
- Edwards, D., Hoyer-Hansen, G., Blasi, F., Sloane, B. (Eds.), 2008. *The Cancer Degradome: Protease and Cancer Biology*. New York, NY: Springer.
- Egeblad, M., Werb, Z., 2002. New functions for the matrix metalloproteinases in cancer progression. *Nature Reviews Cancer* 2, 161–174.
- Emami, N., Diamandis, E.P., 2008. Utility of kallikrein-related peptidases (KLKS) as cancer biomarkers. *Clinical Chemistry* 54, 1600–1607.
- Estrella, V., Chen, T., Lloyd, M., *et al.*, 2013. Acidity generated by the tumor microenvironment drives local invasion. *Cancer Research* 73, 1524–1535.
- Fillmore, H.L., Vanmeter, T.E., Broadus, W.C., 2001. Membrane-type matrix metalloproteinases (MT-MMPs): Expression and function during glioma invasion. *Journal of Neuro-Oncology* 53, 187–202.
- Fingleton, B.M., Heppner Goss, K.J., Crawford, H.C., Matrisian, L.M., 1999. Matrilysin in early stage intestinal tumorigenesis. *APMIS* 107, 102–110.
- Fischer, A., 1946. Mechanism of the proteolytic activity of malignant tissue cells. *Nature* 157, 442.
- Fitzgibbons, P.L., Page, D.L., Weaver, D., *et al.*, 2000. Prognostic factors in breast cancer. College of american pathologists consensus statement 1999. *Archives of Pathology and Laboratory Medicine* 124, 966–978.
- Flutsch, A., Grutter, M., 2013. Proteases in death pathways. In: Brix, K., Stocker, W. (Eds.), *Proteases: Structure and Function*. Vienna, Austria: Springer-Verlag GmbH (Chapter 8).
- Foekens, J.A., Look, M.P., Bolt-De Vries, J., *et al.*, 1999. Cathepsin-D in primary breast cancer: Prognostic evaluation involving 2810 patients. *British Journal of Cancer* 79, 300–307.
- Fonovic, M., Turk, B., 2014. Cysteine cathepsins and extracellular matrix degradation. *Biochimica et Biophysica Acta* 1840, 2560–2570.
- Fortelny, N., Cox, J.H., Kappelhoff, R., *et al.*, 2014. Network analyses reveal pervasive functional regulation between proteases in the human protease web. *PLoS Biology* 12, e1001869.
- Fuchs, S.Y., 2002. The role of ubiquitin-proteasome pathway in oncogenic signaling. *Cancer Biology & Therapy* 1, 337–341.
- Garcia, M., Derocq, P., Pujol, P., Rochefort, H., 1990. Overexpression of transfected cathepsin D in transformed cells increases their malignant phenotype and metastatic potency. *Oncogene* 5, 1809–1814.
- Garcia, M., Platet, N., Liaudet, E., *et al.*, 1996. Biological and clinical significance of cathepsin D in breast cancer metastasis. *Stem Cells* 14, 642–650.
- Giricz, O., Calvo, V., Peterson, E.A., Abouzeid, C.M., Kenny, P.A., 2013. Tace-dependent TGFalpha shedding drives triple-negative breast cancer cell invasion. *International Journal of Cancer* 133, 2587–2595.
- Giusti, I., D'ascenzo, S., Millimaggi, D., *et al.*, 2008. Cathepsin B mediates the pH-dependent proinvasive activity of tumor-shed microvesicles. *Neoplasia* 10, 481–488.
- Glondou, M., Coopman, P., Laurent-Matha, V., *et al.*, 2001. A mutated cathepsin-D devoid of its catalytic activity stimulates the growth of cancer cells. *Oncogene* 20, 6920–6929.
- Gocheva, V., Joyce, J.A., 2007. Cysteine cathepsins and the cutting edge of cancer invasion. *Cell Cycle* 6, 60–64.
- Goll, D.E., Thompson, V.F., Li, H., Wei, W., Cong, J., 2003. The calpain system. *Physiological Reviews* 83, 731–801.
- Gopinathan, A., Denicola, G.M., Frese, K.K., *et al.*, 2012. Cathepsin B promotes the progression of pancreatic ductal adenocarcinoma in mice. *Gut* 61, 877–884.
- Gounaris, E., Tung, C.H., Restaino, C., *et al.*, 2008. Live imaging of cysteine-cathepsin activity reveals dynamics of focal inflammation, angiogenesis, and polyp growth. *PLoS One* 3, e2916.
- Grivennikov, S.I., Karin, M., 2010. Dangerous liaisons: STAT3 and NF-kappaB collaboration and crosstalk in cancer. *Cytokine & Growth Factor Reviews* 21, 11–19.
- Grotzinger, J., Rose-John, S., 2013. ADAM proteases in physiology and pathophysiology: Cleave to function in health or to cause disease. In: Brix, K., Stocker, W. (Eds.), *Proteases: Structure and Function*. Vienna, Austria: Springer-Verlag GmbH (Chapter 9).
- Guo, M., Mathieu, P.A., Linebaugh, B., Sloane, B.F., Reiners Jr., J.J., 2002. Phorbol ester activation of a proteolytic cascade capable of activating latent transforming growth factor-beta1 a process initiated by the exocytosis of cathepsin B. *Journal of Biological Chemistry* 277, 14829–14837.
- Hernandez-Barrantes, S., Bernardo, M., Toth, M., Fridman, R., 2002. Regulation of membrane type-matrix metalloproteinases. *Seminars in Cancer Biology* 12, 131–138.
- Herroon, M.K., Rajagurubandara, E., Rudy, D.L., *et al.*, 2013. Macrophage cathepsin k promotes prostate tumor progression in bone. *Oncogene* 32, 1580–1593.
- Hideshima, T., Ikeda, H., Chauhan, D., *et al.*, 2009. Bortezomib induces canonical nuclear factor-kappaB activation in multiple myeloma cells. *Blood* 114, 1046–1052.
- Hiidenbrand, R., Allgayer, H., Marx, A., Stroebel, P., 2010. Modulators of the urokinase-type plasminogen activation system for cancer. *Expert Opinion on Investigational Drugs* 19, 641–652.
- Holmbeck, K., Bianco, P., Caterina, J., *et al.*, 1999. MT1-MMP-deficient mice develop dwarfism, osteopenia, arthritis, and connective tissue disease due to inadequate collagen turnover. *Cell* 99, 81–92.
- Ihara, S., Miyoshi, E., Ko, J.H., *et al.*, 2002. Prometastatic effect of N-acetylglucosaminyltransferase V is due to modification and stabilization of active matriptase by adding beta 1-6 GlcNAc branching. *Journal of Biological Chemistry* 277, 16960–16967.
- Ioachin, E., 2005. Immunohistochemical tumour markers in endometrial carcinoma. *European Journal of Gynaecological Oncology* 26, 363–371.
- Jager, R., Zwacka, R.M., 2010. The enigmatic roles of caspases in tumor development. *Cancers (Basel)* 2, 1952–1979.
- Jedeszko, C., Sloane, B.F., 2004. Cysteine cathepsins in human cancer. *Biological Chemistry* 385, 1017–1027.
- Jin, X., Yagi, M., Akiyama, N., *et al.*, 2006. Matriptase activates stromelysin (MMP-3) and promotes tumor growth and angiogenesis. *Cancer Science* 97, 1327–1334.
- Joyce, J.A., Baruch, A., Chehade, K., *et al.*, 2004. Cathepsin cysteine proteases are effectors of invasive growth and angiogenesis during multistage tumorigenesis. *Cancer Cell* 5, 443–453.
- Joyce, J.A., Pollard, J.W., 2009. Microenvironmental regulation of metastasis. *Nature Reviews Cancer* 9, 239–252.
- Kane, R.C., Bross, P.F., Farrell, A.T., Pazdur, R., 2003. Velcade: U.S. FDA approval for the treatment of multiple myeloma progressing on prior therapy. *Oncologist* 8, 508–513.
- Kang, J.Y., Dolled-Filhart, M., Ocal, I.T., *et al.*, 2003. Tissue microarray analysis of hepatocyte growth factor/Met pathway components reveals a role for met, matriptase, and hepatocyte growth factor activator inhibitor 1 in the progression of node-negative breast cancer. *Cancer Research* 63, 1101–1105.
- Kataoka, H., 2009. EGFR ligands and their signaling scissors, ADAMs, as new molecular targets for anticancer treatments. *Journal of Dermatological Science* 56, 148–153.
- Kenny, P.A., Bissell, M.J., 2007. Targeting tace-dependent EGFR ligand shedding in breast cancer. *Journal of Clinical Investigation* 117, 337–345.
- Kessenbrock, K., Plaks, V., Werb, Z., 2010. Matrix metalloproteinases: Regulators of the tumor microenvironment. *Cell* 141, 52–67.
- Khokha, R., Werb, Z., 2011. Mammary gland reprogramming: Metalloproteinases couple form with function. *Cold Spring Harbor Perspectives in Biology* 3, a004333.
- Kim, M.G., Chen, C., Lyu, M.S., *et al.*, 1999. Cloning and chromosomal mapping of a gene isolated from thymic stromal cells encoding a new mouse type II membrane serine protease, epithin, containing four LDL receptor modules and two CUB domains. *Immunogenetics* 49, 420–428.
- Kimura, Y., Koga, H., Araki, N., *et al.*, 1998. The involvement of calpain-dependent proteolysis of the tumor suppressor NF2 (merlin) in schwannomas and meningiomas. *Nature Medicine* 4, 915–922.
- Klein, T., Bischoff, R., 2011. Active metalloproteases of the a disintegrin and metalloprotease (ADAM) family: Biological function and structure. *Journal of Proteome Research* 10, 17–33.

- Kobayashi, H., Schmitt, M., Goretzki, L., *et al.*, 1991. Cathepsin B efficiently activates the soluble and the tumor cell receptor-bound form of the proenzyme urokinase-type plasminogen activator (pro-uPA). *Journal of Biological Chemistry* 266, 5147–5152.
- Kumar, S., Rao, N., Ge, R., 2012. Emerging roles of ADAMTSs in angiogenesis and cancer. *Cancers (Basel)* 4, 1252–1299.
- Lakshmikuttyamma, A., Selvakumar, P., Kanthan, R., Kanthan, S.C., Sharma, R.K., 2004. Overexpression of M-calpain in human colorectal adenocarcinomas. *Cancer Epidemiology, Biomarkers & Prevention* 13, 1604–1609.
- Laurent-Matha, V., Huesgen, P.F., Masson, O., *et al.*, 2012. Proteolysis of cystatin C by cathepsin d in the breast cancer microenvironment. *FASEB Journal* 26, 5172–5181.
- Lee, J.W., Kim, M.R., Soung, Y.H., *et al.*, 2006. Mutational analysis of the CASP6 gene in colorectal and gastric carcinomas. *APMIS* 114, 646–650.
- Lee, S., Jilani, S.M., Nikolova, G.V., Carpizo, D., Iruela-Arispe, M.L., 2005. Processing of VEGF-A by matrix metalloproteinases regulates bioavailability and vascular patterning in tumors. *Journal of Cell Biology* 169, 681–691.
- Lee, S.J., Choi, Y.L., Lee, E.J., *et al.*, 2007. Increased expression of calpain 6 in uterine sarcomas and carcinosarcomas: An immunohistochemical analysis. *International Journal of Gynecological Cancer* 17, 248–253.
- Lee, S.J., Kim, B.G., Choi, Y.L., Lee, J.W., 2008. Increased expression of calpain 6 during the progression of uterine cervical neoplasia: Immunohistochemical analysis. *Oncology Reports* 19, 859–863.
- Lee, S.L., Dickson, R.B., Lin, C.Y., 2000. Activation of hepatocyte growth factor and urokinase/plasminogen activator by matriptase, an epithelial membrane serine protease. *Journal of Biological Chemistry* 275, 36720–36725.
- Lee, S.Y., Choi, Y.Y., Choi, J.E., *et al.*, 2010. Polymorphisms in the caspase genes and the risk of lung cancer. *Journal of Thoracic Oncology* 5, 1152–1158.
- Liaudet, E., Derocq, D., Rochefort, H., Garcia, M., 1995. Transfected cathepsin D stimulates high density cancer cell growth by inactivating secreted growth inhibitors. *Cell Growth & Differentiation* 6, 1045–1052.
- Liaudet, E., Garcia, M., Rochefort, H., 1994. Cathepsin D maturation and its stimulatory effect on metastasis are prevented by addition of KDEL retention signal. *Oncogene* 9, 1145–1154.
- Liaudet-Coopman, E., Beaujouin, M., Derocq, D., *et al.*, 2006. Cathepsin D: Newly discovered functions of a long-standing aspartic protease in cancer and apoptosis. *Cancer Letters* 237, 167–179.
- Lind, G.E., Kleivi, K., Meling, G.I., *et al.*, 2006. ADAMTS1, CRABP1, AND NR3C1 identified as epigenetically deregulated genes in colorectal tumorigenesis. *Cellular Oncology* 28, 259–272.
- Liotta, L.A., Abe, S., Robey, P.G., Martin, G.R., 1979. Preferential digestion of basement membrane collagen by an enzyme derived from a metastatic murine tumor. *Proceedings of the National Academy of Sciences of the United States of America* 76, 2268–2272.
- Liotta, L.A., Tryggvason, K., Garbisa, S., *et al.*, 1980. Metastatic potential correlates with enzymatic degradation of basement membrane collagen. *Nature* 284, 67–68.
- List, K., 2009. Matriptase: A culprit in cancer? *Future Oncology* 5, 97–104.
- List, K., Kosa, P., Szabo, R., *et al.*, 2009. Epithelial integrity is maintained by a matriptase-dependent proteolytic pathway. *American Journal of Pathology* 175, 1453–1463.
- List, K., Szabo, R., Molinolo, A., *et al.*, 2005. Deregulated matriptase causes ras-independent multistage carcinogenesis and promotes ras-mediated malignant transformation. *Genes & Development* 19, 1934–1950.
- Liu, R., Li, H., Liu, L., Yu, J., Ren, X., 2012. Fibroblast activation protein: A potential therapeutic target in cancer. *Cancer Biology & Therapy* 13, 123–129.
- Lopez-Otin, C., Hunter, T., 2010. The regulatory crosstalk between kinases and proteases in cancer. *Nature Reviews Cancer* 10, 278–292.
- Lopez-Otin, C., Matrisian, L.M., 2007. Emerging roles of proteases in tumour suppression. *Nature Reviews Cancer* 7, 800–808.
- Lopez-Otin, C., Palavalli, L.H., Samuels, Y., 2009. Protective roles of matrix metalloproteinases: From mouse models to human cancer. *Cell Cycle* 8, 3657–3662.
- Lu, C., Li, X.Y., Hu, Y., Rowe, R.G., Weiss, S.J., 2010. MT1-MMP controls human mesenchymal stem cell trafficking and differentiation. *Blood* 115, 221–229.
- Madsen, D.H., Bugge, T.H., 2013. Imaging collagen degradation *in vivo* highlights a key role for M2-polarized macrophages in extracellular matrix degradation. *Oncoimmunology* 2, e27127.
- Madsen, D.H., Engelholm, L.H., Ingvarsen, S., *et al.*, 2007. Extracellular collagenases and the endocytic receptor, urokinase plasminogen activator receptor-associated protein/Endo180, cooperate in fibroblast-mediated collagen degradation. *Journal of Biological Chemistry* 282, 27037–27045.
- Mai, J., Sameni, M., Mikkelsen, T., Sloane, B.F., 2002. Degradation of extracellular matrix protein tenascin-C by cathepsin B: An interaction involved in the progression of gliomas. *Biological Chemistry* 383, 1407–1413.
- Martin, M.D., Carter, K.J., Jean-Philippe, S.R., *et al.*, 2008. Effect of ablation or inhibition of stromal matrix metalloproteinase-9 on lung metastasis in a breast cancer model is dependent on genetic background. *Cancer Research* 68, 6251–6259.
- Martorana, A.M., Zheng, G., Crowe, T.C., O'grady, R.L., Lyons, J.G., 1998. Epithelial cells up-regulate matrix metalloproteinases in cells within the same mammary carcinoma that have undergone an epithelial-mesenchymal transition. *Cancer Research* 58, 4970–4979.
- Mason, S.D., Joyce, J.A., 2011. Proteolytic networks in cancer. *Trends in Cell Biology* 21, 228–237.
- Masson, O., Bach, A.S., Derocq, D., *et al.*, 2010. Pathophysiological functions of cathepsin D: Targeting its catalytic activity versus its protein binding activity? *Biochimie* 92, 1635–1643.
- Masui, T., Hosotani, R., Tsuji, S., *et al.*, 2001. Expression of METH-1 AND METH-2 in pancreatic cancer. *Clinical Cancer Research* 7, 3437–3443.
- McCawley, L.J., Crawford, H.C., King Jr., L.E., Mudgett, J., Matrisian, L.M., 2004. A protective role for matrix metalloproteinase-3 in squamous cell carcinoma. *Cancer Research* 64, 6965–6972.
- Mentlein, R., Hattermann, K., Held-Feindt, J., 2012. Lost in disruption: Role of proteases in glioma invasion and progression. *Biochimica et Biophysica Acta* 1825, 178–185.
- Merchant, N.B., Voskresensky, I., Rogers, C.M., *et al.*, 2008. Tace/ADAM-17: A component of the epidermal growth factor receptor axis and a promising therapeutic target in colorectal cancer. *Clinical Cancer Research* 14, 1182–1191.
- Milacic, V., Banerjee, S., Landis-Piwowar, K.R., *et al.*, 2008. Curcumin inhibits the proteasome activity in human colon cancer cells *in vitro* and *in vivo*. *Cancer Research* 68, 7283–7292.
- Mochizuki, S., Shimoda, M., Shiomi, T., Fujii, Y., Okada, Y., 2004. ADAM28 is activated by MMP-7 (matrilysin-1) and cleaves insulin-like growth factor binding protein-3. *Biochemical and Biophysical Research Communications* 315, 79–84.
- Mohamed, M.M., Sloane, B.F., 2006. Cysteine cathepsins: Multifunctional enzymes in cancer. *Nature Reviews Cancer* 6, 764–775.
- Moncada-Pazos, A., Obaya, A.J., Fraga, M.F., *et al.*, 2009. The ADAMTS12 metalloprotease gene is epigenetically silenced in tumor cells and transcriptionally activated in the stroma during progression of colon cancer. *Journal of Cell Science* 122, 2906–2913.
- Montgomery, N., Hill, A., McFarlane, S., *et al.*, 2012. CD44 enhances invasion of basal-like breast cancer cells by upregulating serine protease and collagen-degrading enzymatic expression and activity. *Breast Cancer Research* 14, R84.
- Moriyama, T., Kataoka, H., Hamasuna, R., *et al.*, 1999. Simultaneous up-regulation of urokinase-type plasminogen activator (uPA) and uPA receptor by hepatocyte growth factor/scatter factor in human glioma cells. *Clinical and Experimental Metastasis* 17, 873–879.
- Murphy, G., 2008. The ADAMs: Signalling scissors in the tumour microenvironment. *Nature Reviews Cancer* 8, 929–941.
- Murphy, G., 2010. Fell-muir lecture: Metalloproteinases: From demolition squad to master regulators. *International Journal of Experimental Pathology* 91, 303–313.
- Nawrocki, S.T., Carew, J.S., Dunner Jr., K., *et al.*, 2005. Bortezomib inhibits PKR-like endoplasmic reticulum (ER) kinase and induces apoptosis via ER stress in human pancreatic cancer cells. *Cancer Research* 65, 11510–11519.
- Netzel-Arnett, S., Hooper, J.D., Szabo, R., *et al.*, 2003. Membrane anchored serine proteases: A rapidly expanding group of cell surface proteolytic enzymes with potential roles in cancer. *Cancer and Metastasis Reviews* 22, 237–258.
- Nicholson, D.W., 1999. Caspase structure, proteolytic substrates, and function during apoptotic cell death. *Cell Death & Differentiation* 6, 1028–1042.
- Oberst, M., Anders, J., Xie, B., *et al.*, 2001. Matriptase and HAI-1 are expressed by normal and malignant epithelial cells *in vitro* and *in vivo*. *American Journal of Pathology* 158, 1301–1311.
- Oberst, M.D., Johnson, M.D., Dickson, R.B., *et al.*, 2002. Expression of the serine protease matriptase and its inhibitor HAI-1 in epithelial ovarian cancer: Correlation with clinical outcome and tumor clinicopathological parameters. *Clinical Cancer Research* 8, 1101–1107.
- Ohtsuka, T., Shiomi, T., Shimoda, M., *et al.*, 2006. ADAM28 is overexpressed in human non-small cell lung carcinomas and correlates with cell proliferation and lymph node metastasis. *International Journal of Cancer* 118, 263–273.
- Oikonomopoulou, K., Diamandis, E.P., Hollenberg, M.D., 2010. Kallikrein-related peptidases: Proteolysis and signaling in cancer, the new frontier. *Biological Chemistry* 391, 299–310.
- Olsson, M., Zhivotovsky, B., 2011. Caspases and cancer. *Cell Death & Differentiation* 18, 1441–1449.
- Ono, Y., Sorimachi, H., Suzuki, K., 1998. Structure and physiology of calpain, an enigmatic protease. *Biochemical and Biophysical Research Communications* 245, 289–294.

- Orimo, A., Weinberg, R.A., 2006. Stromal fibroblasts in cancer: A novel tumor-promoting cell type. *Cell Cycle* 5, 1597–1601.
- Page-Mccaw, A., Ewald, A.J., Werb, Z., 2007. Matrix metalloproteinases and the regulation of tissue remodelling. *Nature Reviews Molecular Cell Biology* 8, 221–233.
- Piva, R., Ruggeri, B., Williams, M., *et al.*, 2008. CEP-18770: A novel, orally active proteasome inhibitor with a tumor-selective pharmacologic profile competitive with bortezomib. *Blood* 111, 2765–2775.
- Piwnicka, D., Touraine, P., Struman, I., *et al.*, 2004. Cathepsin D processes human prolactin into multiple 16K-like N-terminal fragments: Study of their antiangiogenic properties and physiological relevance. *Molecular Endocrinology* 18, 2522–2542.
- Porter, S., Scott, S.D., Sassoon, E.M., *et al.*, 2004. Dysregulated expression of adamalysin-thrombospondin genes in human breast carcinoma. *Clinical Cancer Research* 10, 2429–2440.
- Przemyslaw, L., Boguslaw, H.A., Elzbieta, S., Malgorzata, S.M., 2013. ADAM and ADAMTs family proteins and their role in the colorectal cancer etiopathogenesis. *BMB Reports* 46, 139–150.
- Puccini, J., Dorstyn, L., Kumar, S., 2013. Caspase-2 as a tumour suppressor. *Cell Death & Differentiation* 20, 1133–1139.
- Puente, X.S., Sanchez, L.M., Overall, C.M., Lopez-Otin, C., 2003. Human and mouse proteases: A comparative genomic approach. *Nature Reviews Genetics* 4, 544–558.
- Rawlings, N.D., Waller, M., Barrett, A.J., Bateman, A., 2014. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research* 42, D503–D509.
- Reinheckel, T., Peters, C., Kruger, A., Turk, B., Vasiljeva, O., 2012. Differential impact of cysteine cathepsins on genetic mouse models of de novo carcinogenesis: Cathepsin B as emerging therapeutic target. *Frontiers in Pharmacology* 3, 133.
- Richards, F.M., Tape, C.J., Jodrell, D.I., Murphy, G., 2012. Anti-tumour effects of a specific anti-ADAM17 antibody in an ovarian cancer model *in vivo*. *PLoS One* 7, e40597.
- Ried, S., Jager, C., Jeffers, M., *et al.*, 1999. Activation mechanisms of the urokinase-type plasminogen activator promoter by hepatocyte growth factor/scatter factor. *Journal of Biological Chemistry* 274, 16377–16386.
- Rios-Doria, J., Day, K.C., Kuefer, R., *et al.*, 2003. The role of calpain in the proteolytic cleavage of E-cadherin in prostate and mammary epithelial cells. *Journal of Biological Chemistry* 278, 1372–1379.
- Robey, I.F., Baggett, B.K., Kirkpatrick, N.D., *et al.*, 2009. Bicarbonate increases tumor pH and inhibits spontaneous metastases. *Cancer Research* 69, 2260–2268.
- Rocheffort, H., Garcia, M., Glondou, M., *et al.*, 2000. Cathepsin D in breast cancer: Mechanisms and clinical applications, a 1999 overview. *Clinica Chimica Acta* 291, 157–170.
- Rocks, N., Paulissen, G., El Hour, M., *et al.*, 2008. Emerging roles of ADAM and ADAMTs metalloproteinases in cancer. *Biochimie* 90, 369–379.
- Rodriguez, D., Morrison, C.J., Overall, C.M., 2010. Matrix metalloproteinases: What do they not do? New substrates and biological roles identified by murine models and proteomics. *Biochimica et Biophysica Acta* 1803, 39–54.
- Rothberg, J.M., Bailey, K.M., Wojtkowiak, J.W., *et al.*, 2013. Acid-mediated tumor proteolysis: Contribution of cysteine cathepsins. *Neoplasia* 15, 1125–1137.
- Ruffell, B., Affara, N.I., Cottone, L., *et al.*, 2013. Cathepsin C is a tissue-specific regulator of squamous carcinogenesis. *Genes & Development* 27, 2086–2098.
- Sabeh, F., Shimizu-Hirota, R., Weiss, S.J., 2009. Protease-dependent versus-independent cancer cell invasion programs: Three-dimensional amoeboid movement revisited. *Journal of Cell Biology* 185, 11–19.
- Satomi, S., Yamasaki, Y., Tsuzuki, S., *et al.*, 2001. A role for membrane-type serine protease (MT-SP1) in intestinal epithelial turnover. *Biochemical and Biophysical Research Communications* 287, 995–1002.
- Seiki, M., 1999. Membrane-type matrix metalloproteinases. *APMIS* 107, 137–143.
- Shchors, K., Nozawa, H., Xu, J., *et al.*, 2013. Increased invasiveness of MMP-9-deficient tumors in two mouse models of neuroendocrine tumorigenesis. *Oncogene* 32, 502–513.
- Shi, Y.E., Torri, J., Yieh, L., *et al.*, 1993. Identification and characterization of a novel matrix-degrading protease from hormone-dependent human breast cancer cells. *Cancer Research* 53, 1409–1415.
- Shiba, E., Kambayashi, J.I., Sakon, M., *et al.*, 1996. Ca^{2+} -dependent neutral protease (calpain) activity in breast cancer tissue and estrogen receptor status. *Breast Cancer* 3, 13–17.
- Shimoda, M., Mellody, K.T., Orimo, A., 2010. Carcinoma-associated fibroblasts are a rate-limiting determinant for tumour progression. *Seminars in Cell & Developmental Biology* 21, 19–25.
- Shree, T., Olson, O.C., Elie, B.T., *et al.*, 2011. Macrophages and cathepsin proteases blunt chemotherapeutic response in breast cancer. *Genes & Development* 25, 2465–2479.
- Sidenius, N., Blasi, F., 2003. The urokinase plasminogen activator system in cancer: Recent advances and implication for prognosis and therapy. *Cancer and Metastasis Reviews* 22, 205–222.
- Sinamon, M.J., Carter, K.J., Fingleton, B., Matrisian, L.M., 2008. Matrix metalloproteinase-9 contributes to intestinal tumorigenesis in the adenomatous polyposis coli multiple intestinal neoplasia mouse. *International Journal of Experimental Pathology* 89, 466–475.
- Sloane, B., List, K., Fingleton, B., Matrisian, L., 2013. Proteases in cancer: Significance for invasion and metastasis. In: Brix, K., Stocker, W. (Eds.), *Proteases: Structure and Function*. Vienna, Austria: Springer-Verlag GmbH (Chapter 15).
- Small, D.M., Burden, R.E., Jaworski, J., *et al.*, 2013. Cathepsin s from both tumor and tumor-associated cells promote cancer growth and neovascularization. *International Journal of Cancer* 133, 2102–2112.
- Sohail, A., Sun, Q., Zhao, H., *et al.*, 2008. MT4-(MMP17) and MT6-MMP (MMP25), a unique set of membrane-anchored matrix metalloproteinases: Properties and expression in cancer. *Cancer and Metastasis Reviews* 27, 289–302.
- Son, J.W., Kang, H.K., Chae, M.H., *et al.*, 2006. Polymorphisms in the caspase-8 gene and the risk of lung cancer. *Cancer Genetics and Cytogenetics* 169, 121–127.
- Soung, Y.H., Lee, J.W., Kim, S.Y., *et al.*, 2004. Somatic mutations of CASP3 gene in human cancers. *Human Genetics* 115, 112–115.
- Sternlicht, M.D., Lochter, A., Symson, C.J., *et al.*, 1999. The stromal proteinase MMP3/stromelysin-1 promotes mammary carcinogenesis. *Cell* 98, 137–146.
- Storr, S.J., Carragher, N.O., Frame, M.C., Parr, T., Martin, S.G., 2011. The calpain system and cancer. *Nature Reviews Cancer* 11, 364–374.
- Suzuki, M., Kobayashi, H., Kanayama, N., *et al.*, 2004. Inhibition of tumor invasion by genomic down-regulation of matriptase through suppression of activation of receptor-bound pro-urokinase. *Journal of Biological Chemistry* 279, 14899–14908.
- Szabo, R., Rasmussen, A.L., Moyer, A.B., *et al.*, 2011. C-Met-induced epithelial carcinogenesis is initiated by the serine protease matriptase. *Oncogene* 30, 2003–2016.
- Takeuchi, T., Harris, J.L., Huang, W., *et al.*, 2000. Cellular localization of membrane-type serine protease 1 and identification of protease-activated receptor-2 and single-chain urokinase-type plasminogen activator as substrates. *Journal of Biological Chemistry* 275, 26333–26342.
- Takeuchi, T., Shuman, M.A., Craik, C.S., 1999. Reverse biochemistry: Use of macromolecular protease inhibitors to dissect complex biological processes and identify a membrane-type serine protease in epithelial cancer and normal tissue. *Proceedings of the National Academy of Sciences of the United States of America* 96, 11054–11061.
- Tang, B.L., 2001. ADAMs: A novel family of extracellular matrix proteases. *International Journal of Biochemistry & Cell Biology* 33, 33–44.
- Tanimoto, H., Underwood, L.J., Wang, Y., *et al.*, 2001. Ovarian tumor cells express a transmembrane serine protease: A potential candidate for early diagnosis and therapeutic intervention. *Tumor Biology* 22, 104–114.
- Theodoropoulos, G.E., Michalopoulos, N.V., Pantou, M.P., *et al.*, 2012. Caspase 9 promoter polymorphisms confer increased susceptibility to breast cancer. *Cancer Genetics* 205, 508–512.
- Tholen, S., Biniowski, M.L., Gansz, M., *et al.*, 2014. Double deficiency of cathepsins B and L results in massive secretome alterations and suggests a degradative cathepsin-MMP axis. *Cellular and Molecular Life Sciences* 71, 899–916.
- Turk, V., Stoka, V., Vasiljeva, O., *et al.*, 2012. Cysteine cathepsins: From structure, function and regulation to new frontiers. *Biochimica et Biophysica Acta* 1824, 68–88.
- Turner, S.L., Blair-Zajdel, M.E., Bunning, R.A., 2009. ADAMs and ADAMTs in cancer. *British Journal of Biomedical Science* 66, 117–128.
- Uhland, K., 2006. Matriptase and its putative role in cancer. *Cellular and Molecular Life Sciences* 63, 2968–2978.
- Vasiljeva, O., Papazoglou, A., Kruger, A., *et al.*, 2006. Tumor cell-derived and macrophage-derived cathepsin B promotes progression and lung metastasis of mammary cancer. *Cancer Research* 66, 5242–5250.
- Victor, B.C., Sloane, B.F., 2007. Cysteine cathepsin non-inhibitory binding partners: Modulating intracellular trafficking and function. *Biological Chemistry* 388, 1131–1140.
- Viloria, C.G., Obaya, A.J., Moncada-Pazos, A., *et al.*, 2009. Genetic inactivation of ADAMTs15 metalloproteinase in human colorectal cancer. *Cancer Research* 69, 4926–4934.
- Wagstaff, L., Kelwick, R., Decock, J., Edwards, D.R., 2011. The roles of ADAMTs metalloproteinases in tumorigenesis and metastasis. *Frontiers in Bioscience (Landmark Edition)* 16, 1861–1872.

- Webb, S.L., Sanders, A.J., Mason, M.D., Jiang, W.G., 2011. Type ii transmembrane serine protease (TTSP) deregulation in cancer. *Frontiers in Bioscience (Landmark Edition)* 16, 539–552.
- Withana, N.P., Blum, G., Sameni, M., *et al.*, 2012. Cathepsin B inhibition limits bone metastasis in breast cancer. *Cancer Research* 72, 1199–1209.
- Witty, J.P., Lempka, T., Coffey Jr., R.J., Matrisian, L.M., 1995. Decreased tumor formation in 7,12-dimethylbenzanthracene-treated stromelysin-1 transgenic mice is associated with alterations in mammary epithelial cell apoptosis. *Cancer Research* 55, 1401–1406.
- Wolf, M., Clark-Lewis, I., Buri, C., *et al.*, 2003. Cathepsin D specifically cleaves the chemokines macrophage inflammatory protein-1 alpha, macrophage inflammatory protein-1 beta, and SLC that are expressed in human breast cancer. *American Journal of Pathology* 162, 1183–1190.
- Wolfsberg, T.G., Primakoff, P., Myles, D.G., White, J.M., 1995. ADAM, a novel family of membrane proteins containing a disintegrin and metalloprotease domain: Multipotential functions in cell-cell and cell-matrix interactions. *Journal of Cell Biology* 131, 275–278.
- Wright, C.M., Larsen, J.E., Hayward, N.K., *et al.*, 2010. ADAM28: A potential oncogene involved in asbestos-related lung adenocarcinomas. *Genes Chromosomes & Cancer* 49, 688–698.
- Wunderlich, A., Roth, S., Ramaswamy, A., *et al.*, 2012. Combined inhibition of cellular pathways as a future therapeutic option in fatal anaplastic thyroid cancer. *Endocrine* 42, 637–646.
- Yang, H., Zhou, P., Huang, H., *et al.*, 2009. Shikonin exerts antitumor activity via proteasome inhibition and cell death induction *in vitro* and *in vivo*. *International Journal of Cancer* 124, 2450–2459.
- Yoshikawa, Y., Mukai, H., Hino, F., Asada, K., Kato, I., 2000. Isolation of two novel genes, down-regulated in gastric cancer. *Japanese Journal of Cancer Research* 91, 459–463.
- Zatz, M., Starling, A., 2005. Calpains and disease. *New England Journal of Medicine* 352, 2413–2423.
- Zeng, L., Cao, J., Zhang, X., 2005. Expression of serine protease SNC19/matriptase and its inhibitor hepatocyte growth factor activator inhibitor type 1 in normal and malignant tissues of gastrointestinal tract. *World Journal of Gastroenterology* 11, 6202–6207.

Further Reading

- Brix, K., Stocker, C.W. (Eds.), 2013. *Proteases: Structure and Function*. Vienna, Austria: Springer-Verlag GmbH.
- Edwards, D., Hoyer-Hansen, G., Blasi, F., Sloane, B. (Eds.), 2008. *The Cancer Degradome: Protease and Cancer Biology*. New York, NY: Springer.
- Lange, P.F., Huesgen, P.F., Overall, C.M., 2012. TopFIND 2.0-linking protein termini with proteolytic processing and modifications altering protein function. *Nucleic Acids Research* 40 (D1), D351–D361. Available at: <http://clipserve.clip.ubc.ca/topfind> (accessed 03.12.14).
- Rawlings, N.D., Waller, M., Barrett, A.J., Bateman, A., 2014. MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research* 42, D503–D509. Available at: <http://merops.sanger.ac.uk/index.shtml> (accessed 03.12.14).

Relevant Website

- <http://AtlasGeneticsOncology.org>
Atlas of Genetics and Cytogenetics in Oncology and Haematology.

Lysosomal Diseases

TM Cox, University of Cambridge, Cambridge, UK

© 2016 Elsevier Inc. All rights reserved.

Introduction

Cell Biology of the Lysosome

Christian de Duve, who discovered the lysosome, understood its dynamic nature almost from the outset: its distinctive acid micro-environment represented a destination for membrane structures and other organelles actively recycled within the cell (De Duve and Wattiaux, 1966; Klionsky, 2007). The nature of the lysosome and its biogenesis are reviewed elsewhere in this encyclopedia. Ultrastructural studies – that followed the first (biochemical) exploration of the cell – revealed peroxisomes, mitochondria, and exogenous particulate material within unit membrane-bound spaces staining positively for acid phosphatase activity (Ashford and Porter, 1962). These findings indicated that the lysosome is not only specialized for and a principal site of intracellular digestion, but constitutes part of an intracellular network which engages with other organelles and is in contact with the plasmalemma and its invaginations (Geuze *et al.*, 1985).

Recycling of other organelles and membrane structures by means of autophagy, together with fluid-phase pinocytosis, and receptor-mediated uptake of proteins by endocytosis, shows that the lysosomal compartment is exposed continually to the external environment (De Duve, 1964). The medical realization of this concept, which represents a potential access route for treating diseases affecting the lysosome, has already borne fruit in stem-cell transplantation and targeted biological therapies; it can also be exploited in therapeutic gene transfer (Cachón-González *et al.*, 2012).

Debt of Cell Science to Medicine

The Nobel prize for Medicine and Physiology, awarded in CE 1974 jointly to de Duve, Claude, and Palade, was a critical landmark in the new science of cell biology. Insights into the physiological rôle of the lysosome have consistently accrued from studies of inborn errors which affect the organelle. Scientific understanding of the dynamic interactions of the living cell has latterly emerged from exploration of the biochemical genetics of lysosomal diseases aided by multimodal techniques in microscopy combined with sophisticated immunological methods for molecular identification. In short, the systematic characterization of disorders affecting the lysosome is a potent source of mechanistic information in that branch of physiology represented by contemporary molecular cell biology (Xu and Ren, 2015). The former concept of the lysosome as the constitutive recycling organ of the cell with a simple destructive rôle in organizing macromolecular catabolism is no longer tenable: in the last years intensive studies based on observations in lysosomal diseases have revealed a biochemically and physically dynamic organelle that is at the center of a reactive control network which regulates cell metabolism and subject to input from countless exogenous and internal signals.

The Lysosomal Compartment

A complement of $\approx 10^3$ lysosomes is present in nearly all mammalian cells but notably absent in mature red blood cells. When the function of this organelle is defective, the range of manifestations is extreme but varies according to the particular lysosomal function which is disturbed and the extent to which the integrity of any given cell or tissue is dependent on that function. In the case of relative or absolute deficiency of lysosomal acid hydrolases (generally proteins found in the matrix of the unit-bound membrane space of the lysosome), manifestations of the genetic defect reflect the turnover rate of the macromolecular substrate(s) of the given enzyme system (Futerman and van Meer, 2004).

The lysosome and related membrane compartments of the cell, including the endosomal system, constitute an extensive network of functioning proteins continuously undergoing turnover and programmed interchanges at defined membrane-contact sites. These interchanges are subject to complex regulatory influences. After receptor-mediated endocytosis, endosomes form by internalization and fuse with early lysosomal vesicles as part of nutrient-gathering and transmission of signals initiated by ligand–receptor interactions. Simultaneously, the lysosomal compartment accepts the inflow of intracellular material – principally particulate or membranous – derived by autophagy. Breakdown products of the autophagosome and endolysosomal pathways are released for conservative resynthesis as part of the constitutive turnover of cellular components (Müller *et al.*, 2015). Dynamic recycling of polypeptides, glycosaminoglycans, and complex sphingolipids is accompanied by similar turnover of nucleic acids since the lysosome also harbors a complement of nucleases which participate in the turnover of mitochondrial, ribosomal, as well as exogenous polynucleotides. Lysosomes themselves are subject to active turnover as a result of fusion and rapid pH changes occurring within the common lumen after vesicular exchange. The membrane of the lysosome contains numerous highly glycosylated proteins forming an internal glycocalyx including integral lysosomal proteins (LIMPS) and the lysosomal-associated membrane proteins (LAMPs) (Safitig and Klumperman, 2009). Maintenance of a proton gradient with a concentration of protons ≈ 1000 -fold greater than that in the ambient cytosolic pH, depends on the presence of a vacuolar, (V-type) H^+ ATPase in the lysosomal membrane (Mindell, 2012). Other membrane transporters exist for the counter-exchange of ions, including the chloride channel, CLCN7 – mutations in which are responsible for dominantly inherited and severe autosomal recessive forms of osteopetrosis – in a recessive variant, severe neurodegeneration occurs. The CLCN7 channel is expressed in the endosome–lysosomal membranes of all cells but is also abundant on the ruffled border of the osteoclast. Other channels include those that transport Ca^{2+} (LAAT1), as well as the egress of small molecules generated by hydrolytic digestion of the macromolecular substrates in the

lysosome – for example, cystinosis (cystine), Cbl F (cobalamin), and sialin (sialic acid). Mutations in each of these are responsible for cognate lysosomal diseases.

Lysosomes acquire complex macromolecules for recycling and breakdown by (1) receptor-mediated endocytosis; (2) engulfment and fusion, for example in phagocytosis; and (3) autophagy – a variant of which, crinophagy, is involved in the maintenance of the contents of secretory granules, which if not discharged, are fused with lysosomes for reclamation of their contents. Latterly numerous molecular components which bring about trafficking and membrane-fusion events, for example Rab proteins and SNARE proteins, are localized to, and interact with the organelle – thus participating in the ebb-and-flow of substrates and digestive products as they move between compartments in the greater lysosomal network in the cell.

Endocytosis

Internalization of ligands for receptors by endocytosis at the cell surface facilitates the delivery of nutrients and molecules released within the cell after endocytic fusion with lysosomes to form the ‘endolysosome.’ Loss of certain membrane components and intensified acidification leads ultimately to formation of the phagolysosome. The acquisition of cholesterol bound to apolipoprotein B in low-density lipoproteins from the plasma and of iron bound to plasma transferrin enable the cognate membrane receptors to be returned to the cell surface, having despatched their complex cargo (Kornfeld and Mellman, 1989). Some receptor complexes are not retrieved but are degraded after fusion with mature lysosomes, for example, epidermal growth factor receptors. While the mechanism by which lysosomal diseases disrupt endosomal activity is not fully worked out, some lysosomal diseases are associated with abnormalities of serum lipoproteins and iron status, which may reflect disturbed endocytosis of receptors, for example, transferrin receptors.

Antigen presentation

In antigen-presenting cells, cysteine proteinases (cathepsins) cleave endocytosed protein antigens to generate peptide fragments of the cognate antigens that are ultimately presented by major histocompatibility complex (MHC) class II molecules. This is an essential step for induction of the adaptive immune response. Deficiency or inhibition of lysosomal acid proteinases, particularly cathepsin S, has been shown experimentally to attenuate the development of autoimmunity (Baugh *et al.*, 2011). Enhanced autoimmunity or disturbance of the adaptive immune response occur in several lysosomal diseases in humans, and may be partly explained by the coordinate expression of several lysosomal proteinases, including the cathepsins; this phenomenon may reflect expansion of the lysosomal network and enhanced activity of transcription factor EB (TFEB).

Phagocytosis

Lysosomes participate in phagocytosis, a key function of the innate immune system responsible for the degradation of microbes, as well as the physiological turnover of effete blood cells. First noted by I.I. Metchnikoff in CE 1882 (Metchnikoff, 1884), and widely distributed in nature, phagocytosis leads to

engulfment of particulates including bacteria, yeasts, and colloidal suspensions; this requires active membrane-fusion events (the term ‘phagolysosomes’ was introduced by de Duve) and is a particular feature of macrophages and dendritic cells (Aderem and Underhill, 1999). This process also represents an orchestrated interaction between innate immunity and the adaptive responses related to antigen presentation.

Autophagy

Three distinct intracellular activities participate in autophagy: microautophagy, chaperone-mediated autophagy, and macroautophagy. The latter was first identified as a result of appearance of ingested mitochondria, peroxisomes, and other membranous material within acid phosphatase-enriched unit membrane-bound vesicular structures, and the first insights into autophagy by de Duve and colleagues led to the coining of the term early in a long path of discovery (De Duve and Wattiaux, 1966; Klionsky, 2007; Ashford and Porter, 1962; De Duve, 1964).

The physiological significance of autophagy was clarified when excess glycogen was identified in the lysosomal compartment of patients with acid maltase deficiency due to Pompe disease. Accumulation of internal cellular debris representing the engulfment of relatively large-volume material, including other organelles and membrane fragments, occurs in a dynamic process of flow and fusion – a key function of the lysosome now termed macroautophagy (this includes ‘mitophagy,’ which refers specifically to the engulfment of mitochondria) (Yang and Klionsky, 2010). Macroautophagy is not only critical for nutrient retrieval through recycling of molecular building blocks in periods of starvation and other metabolic stress – it plays a constitutive rôle in developmental remodeling and maintenance of integrated cell function in health.

Microautophagy describes a continuous constitutive process in which cytosolic components trapped during invagination of endosomes occurs: cytosolic molecules are engulfed by invagination of lysosomal or endosomal membranes.

In chaperone-mediated autophagy, cytosolic proteins enter lysosomes principally from the cell surface through chaperone- and receptor-mediated internalization. This is known, partially through the genetic characterization of Danon disease, to involve protein unfolding and translocation mediated by the lysosomal membrane protein isoform, LAMP2a.

Autophagy, now more definitively termed macroautophagy, has been conserved throughout eukaryotic evolution; it involves sequestration of a cytosolic region and its particulate contents. The cytosol is at first surrounded and then sealed off with the formation of a phagophore – a double-layer membrane vesicle designated for recycling. The formation of the phagophore is initiated when a planar membrane cisterna appears to cordon off small volumes of cytoplasm and their contents to form a double-layered space. Sequestered cell cargo enters a stepwise process of compositional maturation and ultimately fusion of the vesicular structure with vacuoles and lysosomes. Intensive genetic studies in baker’s yeast have revealed cellular machinery that requires the products of more than 30 genes, termed essential autophagy (Atg) genes which include those encoding ubiquitin-like conjugation systems. Organelles and other content

engulfed as cargo in the new vesicular spaces include endoplasmic reticulum, ribosomes, mitochondria, peroxisomes, as well as recognizable glycogen: all are ultimately distributed to form autophagosomes – first named by de Duve. Subsequent ordered maturation also involves fusion with endosomal and lysosomal vesicles.

The stepwise processes with which macroautophagy engages are tightly regulated; but the mechanisms that control fusion and subsequent maturation are not fully understood. The terms accorded to the component structures are (1) the autophagosome (or early autophagic vacuole); (2) the amphisome (late or degradative autophagic vacuole); and, after fusion with the lysosome; (3) the autolysosome. Ambiguities arise because in ultrastructural studies the term autophagic vacuole is generally preferred for structures (2) and (3) since the molecular components were difficult to define before the introduction of dynamic high-resolution imaging studies. The maturation process involves increasing membrane cholesterol content, progressive acidification, and other changes in molecular composition – in the latter stages these include acquisition of lysosomal membrane proteins (Saftig and Klumperman, 2009; Eskelinen, 2005).

While systematic studies in yeast mutants have greatly advanced this field, clarification of the rôle of many of the molecular components of macroautophagy in mammalian cells still awaits systematic investigation in a pathophysiological context. Most of the genes and their products so far identified affect the early steps in induction of phagophores or autophagosomes. Important signaling cascades that induce or suppress autophagy are now known, but the dynamic forces and interrelationships between the pathways that drive autophagy are ill-understood. Several intracellular membrane domains including the endoplasmic reticulum, Golgi organ, mitochondria, and plasmalemma may be the true source of autophagic membranes, but trafficking requires components of the secretory pathway, as well as the SNARE fusion machinery including *N*-ethylmaleimide-sensitive attachment protein (SNAP) receptor (SNARE) proteins. Rab GTPase-activating proteins (Rab GAPs) are implicated in the spatial and temporal dynamics of the cellular endomembrane system and reconstitution studies confirm the requirement for SNAREs28. RAB5 and RAB7 are specifically involved in the tethering and docking processes in endo-lysosomal membrane trafficking (Popovic *et al.*, 2012). Formation of phagophores in mammalian cells requires an often-studied protein, LC3 (microtubule-associated protein 1 light chain 3), which is the homolog of the yeast autophagy protein, Atg8. Acidified late endosomes and lysosomes fuse with these nascent vacuoles to form the autolysosome; inhibitors of the $V\text{-(H}^+)$ ATPase, such as bafilomycin, inhibit fusion with both components. Deficiency of the E1 ubiquitin-activating enzyme also impairs clearance of the late autophagic vacuoles – critically implicating protein ubiquitination in autophagy. Continued acidification, with activation of the complement of lysosomal hydrolases initiates digestion of the segregated contents; these disintegrate, with formation of a multi-vesicular body containing electron-dense debris which is ultimately consolidated to an auto-fluorescent core or residual body (Piper and Luzio, 2001). Lysosomal diseases associated with deficient hydrolase activity or defective cofactors, or treatment with lysosomal

enzyme inhibitors, lead to the accumulation of late autophagic vacuoles stuffed with incompletely processed cytoplasmic contents and membranous debris.

Autophagy and the Lysosome in Energy Economy and Nutrient-Sensing

Autophagy

Autophagy provides nutrient precursors retrieved from prior biosynthetic investment under conditions of sufficiency; organelles and other cellular constituents serve as source that can be readily recovered as a working currency under conditions of metabolic stress and particularly when the supply of nutrients is limited. Autophagy is thus characteristically stimulated by starvation and allows the cell to adopt a parsimonious attitude to energy potential and nutrient stores.

Disuse of muscles, inflammatory conditions associated with fever, glucagon or thyroid hormone action, as well as sympathomimetic stimulation, also enhance autophagy in different tissues. Hence, what was formerly considered a 'housekeeping' function of the lysosome proves to be a highly regulated and adaptive process (Ravikumar *et al.*, 2010).

Perceptions of the rôle of the lysosome in metabolic control are changing rapidly. Over the last decade, a body of research, including discoveries in the laboratories of David Sabatini on the protean functions of mechanistic target of rapamycin complex (mTORC) (Sengupta *et al.*, 2010), and of Andrea Ballabio, has shown that the lysosome itself serves as a coordinating center which controls the clearance of cellular components in response to the demands for energy generation and biosynthesis of essential intracellular constituents. Identification of the controlling transcription factor TFEB for lysosomal protein expression and biogenesis is of signal importance (Sardiello *et al.*, 2009).

The multi-subunit serine/threonine protein kinase, mechanistic target of rapamycin complex 1 (mTORC1), regulates homeostasis by coordinating metabolism in response to nutrient, energy, and oxygen availability, as well as signaling for growth and proliferation from growth factor–receptor interactions. Energy-sensing of mTORC1 is effected through the ATP/AMP energy charge which if below a threshold will engage the kinase, 5'AMP-activated protein kinase (AMPK), that is activated under conditions of reduced ATP and elevated AMP availability. Autophagy is also orchestrated by the reduced energy charge which favors inhibition of mTORC1 by phosphorylation and allows coordination of overall nutrient input with the demands for formation *de novo* of molecules critical for cell survival. AMPK in turn influences protein synthesis; specific phosphorylation of a target serine residue inactivates TSC2 and also the protein raptor which prevents recruitment of mTORC1, stimulates autophagy, and arrests initiation of protein synthesis at the level of eIF-2 α (Sengupta *et al.*, 2010; Sardiello *et al.*, 2009; Settembre *et al.*, 2013; Proud, 2006).

Autophagy linked to nutrient supply is regulated by a signaling network susceptible to stimulatory as well as inhibitory inputs. The TFEB is induced by starvation and mediates an adaptive homeostatic response (Settembre *et al.*, 2012). Apart from mTORC1, several kinases have been shown to phosphorylate TFEB, including extracellular signal-related kinase 2

(ERK2) and an isoform of protein kinase C (PKC β 95). Phosphorylation of serine at position 142 by ERK2 and of serine 142 and serine 211 by mTORC1 determines the subcellular localization of the nuclear transcriptional factor; it is cytosolic when phosphorylated and translocates to the nucleus when dephosphorylated at these residues. When the supply of nutrients is adequate, TFEB interacts with a group of cellular regulatory molecules newly termed the lysosome nutrient sensing (LYNUS) machinery. This detects nutrient signals via the v-(H⁺) ATPase complex, and is subject to phosphorylation control mediated by (mTORC1) on the cytosolic surface of lysosome. In the presence of nutrients and growth factors, mTORC1 is a strong inhibitor of autophagy. When the mTORC1 kinase function is activated, autophagy is inhibited: this activation requires the interaction of numerous proteins, including RAG GTPases and the Ragulator complex which are assembled at the lysosomal membrane (Sengupta *et al.*, 2010).

Cell growth is an anabolic process also coordinated by mTORC1; this involves enhanced translation of mRNA in response to the availability of signaling nutrients such as amino acids (Proud, 2006). Not only does mTOR-mediated phosphorylation of the binding factor 4E-BP1 render eIF4E available to bind eIF4G and promotes formation of functional initiation complexes, but phosphorylation activates elongation factors and ribosome biogenesis. However it is also clear that at the same time mTORC1 suppresses the breakdown of cellular components by autophagy, thus coordinating biosynthesis and degradation (Sengupta *et al.*, 2010; Settembre *et al.*, 2013). While it has long been understood that an essential pathway for sensing nutrient status is the master growth regulator mTORC1, the mode by which it is coupled to the nutrient signals and the lysosome has not until recently been recognized. Recent discoveries provide a logistical point of interaction between the requirements for cell proliferative responses through growth factors and release of amino acids and other intracellular building blocks during autophagy.

Growth factor signaling

Autophagy is regulated by a signaling network susceptible to input from discrete and distinct metabolic influences. Binding of insulin and related hormones to their cognate receptors activate the mitogen-activated protein kinase (MAPK) signaling mechanism through the Raf-1–MEK1/2–ERK1/2 pathway. Distinct from mTORC1 pathway, this mechanism ensures that the master regulator of lysosomal biogenesis TFEB can be preferentially phosphorylated by separate mechanisms so that lysosomal protein synthesis and autophagy are not initiated above constitutive rates (Proud, 2006; Settembre *et al.*, 2011).

Mechanistically mTORC1 controls protein synthesis, coordinating signals related to nutrients and energy status – as well as those received from insulin and other growth factor receptors – to regulate metabolic fluxes involving biosynthesis and catabolism (Laplanche and Sabatini, 2013). Ultimately, this system controls cell growth and proliferation. Insulin-like growth factors may also act through the receptor tyrosine kinase Akt/PKB signaling to activate mTORC1. Many signals received by mTORC1 are orchestrated by a heterodimeric tumor suppressor complex TSC1/TSC2 (tuberous sclerosis complexes 1 and 2) that inhibits the guanosine triphosphatase, Rheb; this in turn activates mTORC1 protein

kinase activity and hence protein synthesis. Mitogens such as insulin-like growth factor 1 activate the mitogen activated protein kinase/extracellular kinase (MAPK/ERK) pathway, which can inhibit the TSC1/TSC2 complex, thus stimulating mTORC1 and inhibiting autophagy. Energy-sensing of mTORC1 is effected through the ATP/AMP energy charge which if below a threshold will activate the kinase, AMPK, that in turn influences protein synthesis and by means of a specific phosphorylation of a target serine residue, inactivates TSC2 and also the protein raptor. This prevents recruitment of mTORC1, stimulates autophagy and arrests initiation of protein synthesis at the level of the initiation factor 2 subunit, eIF-2 α (Sengupta *et al.*, 2010; Sardiello *et al.*, 2009; Proud, 2006; Settembre *et al.*, 2011, 2012, 2013; Laplanche and Sabatini, 2013).

Amino acid signaling

Autophagy is regulated by a signaling network susceptible to stimulatory and inhibitory inputs. mTORC1 integrates signaling by means of phosphoinositol 3-kinase (PI3 kinase) and pathways sensitive to amino acids. Amino acids are assembled into proteins but also serve as metabolic intermediates in other biosynthetic reactions; amino acids strongly influence an over-arching control switch that regulates the cellular supply of energy and key nutrients (Bar-Peled and Sabatini, 2014). This pathway ultimately involves the lysosome to enact the molecular directives of the multicomponent protein kinase and master growth regulator, the mTORC1.

Binding of insulin and related hormones to their cognate receptors activates the mTORC1 pathway, leading to inhibition of autophagy. In contrast to growth factor signaling, amino acids may signal to mTORC1 by inhibiting Raf-1–MEK1/2–ERK1/2 and so also inhibit autophagy. A critical factor for understanding metabolic control through the mTORC1 pathway resides in the observation that growth factors do not effectively activate mTORC in the absence of amino acids.

Amino acids influence the interaction between mTORC1 and a distinct family of four small GTPases. These so-called Rag GTPases promote the intracellular localization of mTORC1 to the proximity of its activator, Rheb, on the cytoplasmic surface of the lysosome. It has been recently shown that the vacuolar H⁺-adenosine triphosphatase is required for amino acids to activate mTORC1. The v-ATPase interacts with the Ragulator, a complex that serves as a scaffold to bind the Rag GTPases to the lysosome (Zoncu *et al.*, 2011).

Signaling via free amino acids, especially leucine and lysine (and possibly glutamine) occurs within the lysosomal lumen – a process dependent on the putative amino acid carrier, human member 9 of the solute carrier family 38 (SLC38A9) present in the lysosomal membrane and of which a unique cytoplasmic segment forms part of the conformation-specific Ragulator-Rag GTPases scaffold. Experimental mutagenesis and cellular expression of a defective amino acid carrier shown that SLC38A9 is an essential component of the nutrient-sensing machinery by which leucine and other intralysosomal amino acids released by proteolysis are able to permit activation of the mTORC1 kinase control pathway. Allosteric control exerted by the carrier protein on mTORC activity transduces signals indicating the abundance of key nutrients

released in the intralysosomal compartment – the system thus ensures permissive regulation of autophagy and organelle biogenesis as well as overall metabolic control by sensing at the core of the cell (Rebsamen *et al.*, 2015).

Lysosomal Diseases

Studies in several populations show that a range of 1 per 2315 – 7700 living infants are born with an inherited lysosomal disease (depending on the method of ascertainment (Mechtler *et al.*, 2012; Poupetová *et al.*, 2010)), representing a large burden to the individual, their carers, and society. Although the defect is present at birth, the diseases typically appear later. Clinically diverse, the rate of progression and effects on the individual vary greatly. While the manifestations of lysosomal diseases often include enlargement of the visceral organs, disturbed function and infiltration of the lungs and kidneys for example, often accompanied by skeletal effects. In about two-thirds of the lysosomal diseases, neurological deficits occur.

Clinical Conspectus

Arguably, the first disease in humans now known to be a lysosomal disorder was reported in four siblings from an isolated community in Norway by Otto Stengel, in 1826. Stengel gave a convincing description of impaired vision leading to blindness, cognitive decline, and early death. His description indicates a classical and most frequent form of juvenile neuronal ceroid lipofuscinosis (CLN3) – now seen as one of 14 genetically distinct disorders with common clinicopathological features including neurodegeneration with seizures, in the family of neuronal ceroid lipofuscinoses (Haltia *et al.*, 2011). Mutations in a protein, CLN3, particularly expressed on endosomal–lysosomal membranes, but with as yet unknown function, are responsible for this disease.

Pompe disease

Pompe disease in infants with massive cardiac hypertrophy and hypertonia due to profound skeletal muscle weakness was first described by Joannes Cassianus Pompe in CE 1932. The disease had a rapidly fatal outcome. Pompe disease (known also as glycogenosis II) has iconic significance since it was the first inborn error of the lysosome to be so identified. Large quantities of glycogen are deposited in muscle, including the tongue. Hepatic enlargement also occurs in affected infants. Acid maltase is profoundly deficient in affected infants with Pompe disease; patients with the disease beyond infancy usually have residual acid maltase activity and attenuated clinical variants characterized by a proximal skeletal myopathy and sparing of the heart but with weakness of the diaphragm leading to respiratory failure.

When manifest in children and adults, poor athletic performance and occasionally delayed achievement of developmental milestones are observed: the spine is hypermobile due to weakness of proximal muscles with inability to rise from the squatting position or maintain the upper limbs in an extended position; while the onset of symptoms varies

between childhood and up to the age of 60 years, few patients give a lifelong history. Ultimately, respiratory muscle weakness ensues with retention of carbon dioxide as a result of ventilatory failure. Inability to swallow solids is the result of impaired function of the voluntary pharyngeal muscles (van den Hout *et al.*, 2003).

Specific treatment has been introduced for Pompe disease in the form of recombinant acid maltase administered by infusion. In infants where acid maltase is absent, immune reactions to the infusions of the wild-type acid maltase occur and this may lead to treatment failure (Kishnani and Beckenmeyer, 2014). Otherwise survival is prolonged beyond infancy and the child may have normal growth, although non-muscle manifestations in heart, viscera, and kidney with prominent endothelial disease and late complications in the nervous system may ultimately prove fatal.

Tay-Sachs disease

Tay-Sachs disease was originally described as a rapidly fatal neurodegenerative condition that is also known as GM2 gangliosidosis, in a young child almost certainly of Ashkenazi origin from Whitechapel, East London, in CE 1881 (Tay, 1881). The disease remains a paradigmatic example of severity and also heterogeneity at many levels: Waren Tay noted progressive loss of sight, reduced muscle tone, poor movement, and loss of cognitive ability; he later reported on the familial nature of the condition. Owing to the near-specific retinal findings in the index patient with the ‘cherry red spot’ (representing pale retinal ganglion cells near the macula), Tay’s identification is unambiguous. Bernard Sachs, a neurologist in New York, reported the strong predilection for Ashkenazi Jews. An early indication of the lysosomal nature of Tay-Sachs disease was revealed by Sachs’s subsequent neuropathological studies of children with ‘amaurotic family idiocy,’ showing involvement of nearly all neurones by a process that distorts dendrites with cytopathic changes in neurones characterized by prominent intracytoplasmic bodies in the perinuclear region – electron microscopy of which, years later, revealed abnormal lysosomes (Terry and Korey, 1960). Tay-Sachs disease is a classical example of an extremely severe neurodegenerative lysosomal disease in infancy marked by strong ethnic associations (several in non-Jewish communities) and with effective prenatal screening programs established globally for at-risk individuals (Kaback and Desnick, 2011).

Gaucher disease

Gaucher disease, also over-represented in the Ashkenazim, is a sphingolipid disorder biochemically related but clinically distinct from Tay-Sachs disease. First reported in a doctoral thesis in the University of Paris (1882) by P.C. Ernest Gaucher in a 34-year-old woman who died in Paris in CE 1881, it causes massive enlargement of the liver and spleen with anemia and bleeding. These characteristic features are now readily ameliorated by functional complementation using the purified lysosomal glycoprotein (acid β -glucosylceramidase encoded by the GBA1 gene) that is deficient in the condition; the therapeutic enzyme is remodeled to display terminal mannose residues which mediate protein uptake by specific surface receptors on the pathological macrophages – a principal focus of the disease in viscera (Deegan and Cox, 2012).

Gaucher disease provides a spectacular example of the successful application of rational molecular understanding to therapeutic research with beneficial outcomes in an otherwise life-shortening and ultra-rare (orphan) disease: large biotechnology revenues have been generated – realizing the aims of the 1983 US Orphan Drug Act. However, after introduction of salutary targeted enzyme augmentation, the panethnic and protean nature of Gaucher disease has been revealed. Now viewed as a classical example of a multisystem lysosomal disease principally due to unitary genetic determinants, Gaucher disease presents challenging comorbidities. These include several unusual cancers (especially myeloma and B-cell lymphoma) (Weinreb and Lee, 2013) and complex and poorly understood neurological phenomena, including Parkinson disease and Lewy body dementia – late-onset neurological complications occurring in otherwise apparently healthy heterozygous carriers not only in pedigrees affected but in the general population, including those with Ashkenazi Jews but also in other ethnic groups, suggesting the influence of dominant-negative mutant GBA1 alleles (Shapira, 2015). With commercial introduction of effective treatments for this ultra-rare disorder, accretion of clinical experience based on cohorts of patients who attend expert clinical centers has, in an era of biopharmaceutical success, led in the treatment era to several unexpected mechanistic discoveries – beyond this, lies the expectation of advanced therapies based on fundamental understanding of the pathogenesis of lysosomal disorders.

Nomenclature

Classification of Inborn Errors of the Lysosome

At least 70 inborn errors of lysosomal biogenesis and function (including lysosome-related organelles) are reported. As powerful methods for ultra-rapid DNA sequencing are introduced for precise diagnosis of rare diseases and with intensive study of lysosomal membrane proteins (Saftig and Klumperman, 2009), this list is growing. Lysosomal diseases may be classified according to the nature of the biochemical consequences of the defect (usually pathological storage of a given macromolecular substrate and its precursors) or according to the defective molecular function induced by the given genetic defect; while this provides a better interpretation as to basic causation, for optimal practical understanding, a reconciliation of the two systems is desirable.

Inborn errors of lysosomal metabolism may be classified biochemically. While useful, this classification includes diverse conditions attributed to mutations at single loci but with multifarious clinico-pathological effects. The incomplete taxonomy of lysosomal diseases in part reflects the history of their individual discovery: it includes identification of the principal storage molecules and their relationship to the disrupted cellular biology. Each disorder reveals aspects of the many-faceted rôle of lysosomes in the remodeling of intracellular components through autophagy, recycling of exogenous materials as well as nutrient- and energy-sensing – functions linked to biological economy. While a more impartial genetic classification might appear attractive, defects in many non-lysosomal proteins that interact dynamically to sustain and

control the lysosomal network will be needed before a comprehensive taxonomy can be agreed.

- Major classes related to biochemical abnormalities include:
 - Mucopolysaccharidoses
 - Sphingolipidoses
 - Glycoproteinoses
 - Glycogenosis with or without autophagic debris
 - Conditions with multiple classes of storage material.
- Functional defects include:
 - Deficiency of a given acid hydrolase
 - Abnormal posttranslational modification of lysosomal proteins
 - Deficiency of an activator or protein cofactor
 - Defective lysosomal membrane or transporter protein
 - Disturbed biogenesis of lysosomes or lysosome-related organelles.

Identification of the First Inborn Disease of the Lysosome

Although the lysosome was a chance discovery by Christian de Duve, it was his former colleague, Henri-Géry Hers (also a Belgian physician-investigator) who identified the first inherited disease of the lysosome (Hers, 1963). Pompe disease is a fatal condition in which cardiomyopathy and skeletal myopathy are accompanied by excess glycogen deposition in the affected tissues. Notably, the histochemical properties of the excess glycogen appeared normal, including its chromogenic reaction with iodine/potassium iodide. Hers investigated the enzymatic activities related to the metabolism of glycogen that had been explored by Gerty and Carl Cori at Washington University School of Medicine. These investigators shared the 1947 Nobel prize for Medicine or Physiology. Of note, Carl Cori was a scientific mentor of de Duve; Gerty Cori, née Radnitz, was a former pediatrician and had encountered glycogen storage disorders – much later in the US identifying or correctly postulating the enzyme defects in four distinct types. This body of work provided the methodological platform for Hers's studies in tissue biopsy specimens obtained from patients with generalized glycogenosis, Pompe disease – previously classified by Gerty Cori as glycogen storage disease type II (notably without a biochemical explanation) (Cori, 1952).

Hers found that activities of branching enzyme, debranching enzyme, glucose 6-phosphatase, and glycogen phosphorylase were within the healthy reference range; however, in a series of experiments, in part to investigate glucosyltransferase activities related to the possible debranching enzyme postulated by W.H. Whelan, the activity of a previously unknown glycosidase in liver tissue was found to be deficient in patients with Pompe disease. This 'maltase' had a pH optimum of 4 and when control tissues obtained from patients with other glycogen storage diseases were examined, acid maltase activity was found to be selectively deficient in Pompe disease (Hers, 1983). Hers and colleagues subsequently showed that acid maltase present in rat liver had properties typical of a lysosomal hydrolase: activity in fresh tissue homogenates co-sedimented with the classical lysosomal marker enzyme, acid phosphatase; it was moreover found that after freeze-thawing or addition of mild non-ionic detergent (digitonin), the complement of lysosomal

acid maltase activity was liberated into 'soluble' supernatant fractions (Lejeune *et al.*, 1963). Pierre Baudhuin, an electron microscopist, reported the presence of intralysosomal and extralysosomal glycogen deposits in a biopsy sample obtained from the liver of a patient with Pompe disease (Baudhuin *et al.*, 1964).

Correlation of biochemical and structural abnormalities – organelle pathology

Prompted by the early introduction of fibroblast cultures obtained from skin biopsies as a technique for refined study of heritable disorders of connective tissue, Danes and Bearn showed that fibroblasts cultured from patients with Hurler syndrome (MPS I) showed cellular metachromasia on staining with various organic dyes (Danes and Bearn, 1965), with expression of the cellular defect in a manner compatible with the Lyon hypothesis of X-inactivation in fibroblasts from patients with X-linked disease (Hunter syndrome, MPS II). It is now known that these inherited metabolic diseases are associated with the lysosomal accumulation ('storage') of glycosaminoglycans. Hers later collaborated with François van Hoof, who had studied tissues obtained from patients with mucopolysaccharidosis (MPS) by electron microscopy. Having observed abnormal ultrastructure in the liver of patients with MPS I (Hurler disease), Hers and van Hoof determined the activities of lysosomal enzymes in mucopolysaccharidoses and later, the glycoprotein disorder, fucosidosis. At this time, the excess glycosaminoglycans were postulated to be the consequence of overactive biosynthesis. The biochemical findings were compatible with studies of ultrastructure by electron microscopy, which revealed prominent intracellular membrane-bound inclusion bodies representing the effects of the disease in the lysosomal compartment (Baudhuin *et al.*, 1964; Van Hoof and Hers, 1968).

Inborn errors of the lysosome

The concept of inborn errors of lysosomal function was formally enunciated by Henri-Géry Hers in CE 1965 (Hers, 1965). Thus numerous defects in the recycling of cellular macromolecules were classified as 'lysosomal storage diseases.' While the general potential for access of proteins to the lysosomal compartment had been realized earlier by de Duve, the enzyme defects soon to be found responsible for several lysosomal diseases have important consequences for diagnosis and understanding but also in the context of later discoveries by E.F. Neufeld, W.S. Sly, P.D. Stahl, R.O. Brady, and others that provide the experimental basis for specific therapy directed to the lysosome. De Duve later noted that the research by H.-G. Hers: "... elucidated, with important consequences for diagnosis, prevention and therapy, a vast chapter of pathology that had remained totally mysterious ..."

Pathogenesis of Lysosomal Diseases

Understanding of the molecular pathogenesis of lysosomal diseases is generally rudimentary, and despite increasing knowledge about the primary genetic defect, it is now clear that most lysosomal defects represent individual pathological domains with a few shared but many unique determinants

that defy general description. Here below, some basic aspects of pathogenesis and storage are set out but space does not permit comprehensive discussion; for more in-depth information please refer to recent reviews and references therein (Lieberman *et al.*, 2012; Klein and Futerman, 2013; Cox and Cachón-González, 2012; Platt *et al.*, 2012; Morgan *et al.*, 2011).

Lysosomal storage

A noteworthy feature of the lysosomal breakdown of cellular macromolecules is that the process usually occurs in a stepwise fashion and is initiated at terminal units, for example, glucose units in glycogen or nonreducing terminal sugars in the glycosaminoglycans and glycosphingolipids. Should the lysosome fail completely to digest the terminal unit, then subsequent sequential breakdown of serial units in the macromolecular substrate is arrested. However, at the same time, cellular biosynthesis of the macromolecule and its presentation to the lysosomal compartment via autophagy and endocytosis continues – thus inducing a buildup of the undegraded material as well as precursor molecules that have been generated in preceding enzymatic steps. Generally, the immediate consequence of the deficiency of any given acid hydrolase is the accumulation of material into an expanded lysosomal space which, as a result of unusual staining properties (e.g., metachromasia in the fibroblasts cultured from patients with MPS (Danes and Bearn, 1965)), ultimately alerts the microscopist. Electron microscopy reveals a distorted membranous compartment representing a greatly expanded lysosomal network within the ultrastructure of the cell (Baudhuin *et al.*, 1964).

Effects of Lysosomal Storage

Lysosomal storage may arise for reasons other than the direct or indirect impairment of digestive function: defective transport of the products of lysosomal digestion due to mutations in carrier proteins may be responsible. A classical example is cystinosis, where mutations in the transmembrane protein, cystinosin, lead to retention of cysteine – a product of intralysosomal proteolysis and which dimerizes to form crystalline oxidized cystine aggregates within the lysosomal compartment (Nesterova and Gahl, 2013). For reasons not completely understood, the condition usually first induces disease of the proximal renal tubule of the kidney but if untreated, proceeds rapidly to induce a multisystem disorder and renal failure; intracellular deposits of cystine are readily seen by light microscopy in many tissues including the thyroid gland, skeletal muscle, and heart – early on in the illness, crystalline deposits are readily visible in the cornea which, if untreated, cause ocular irritation and impaired vision.

Ultimately, once the capacity for retaining the array of cellular macromolecules is exceeded, many lysosomal functions become compromised. While simple observation of the diseased tissue may show distorted and hypertrophic cells with an expanded lysosomal space, often in a perinuclear distribution, vary rarely does storage material account directly for the effects on the manifestly diseased organ. For example, in Gaucher disease, where the spleen at 6 l may be enlarged more

than 40-fold above the normal adult volume of less than 150 mL, the additional tissue cannot be accounted for by the excess storage material – in this case principally β -glucosylceramides (which account generally for only a few tens of grams of the several kilograms of excess tissue mass). While this quantitative disparity has long been known, in the context of widespread inflammatory conditions with hyperplasia and multiple tissue adaptations related to an inflammatory response in the face of biochemical storage, it remains true that in most instances the consequential effects of the storage are poorly understood in biochemical or molecular terms; they are reported principally as part of the descriptive vocabulary of inflammatory and other histopathological changes, due to the release of cytokines, macrophage activation, edema, cell loss, fibrosis, and altered vascularity.

Lysosomal Diseases and Disturbances of Autophagy

Pompe disease demonstrates the rôle of autophagy in glycogen remodeling

Pompe disease, the first lysosomal disease to be biochemically characterized – known alternatively as glycogen storage disease Type II – affects a key cellular energy store in muscle, liver, brain, and endothelial cells. Glycogen is an intracellular metabolite and important repository of energy central to cellular anabolism and storage on the one hand, and energy-yielding catabolic pathways activated by nutrient crises on the other. Given the rôle of the lysosome in the continual remodeling of glycogen, as well as the recycling of protein and complex lipid components in the cell, accompanied by liberation of internal resources through autophagy, it is clear that the organelle plays a critical part in cellular homeostasis.

Glycogen is a highly arborized polymeric structure or dendrimer originating from a single tyrosine-linked protein core of glycogenin. Each linear stretch contains about 13 glucose residues and, except for the outermost layer, carries two branches of the same length that are attached 3–4 glucose units apart. The high density of the branches in each glycogen molecule favors ready access of the enzymes and regulatory proteins involved in fine metabolic control of energy and glucose supply at the termini. Crowding in the outer branches of the glycogen polymer ultimately leads to steric hindrance of branching enzyme, so that the molecule is limited to about 55 000 glucose units with a calculated MW $\approx 10^7$ Da (Cori, 1952).

Metabolic studies show that glycogen is degraded within a few hours of synthesis, but slight abnormalities in the molecular nanostructure due to faulty formation of its α -1,6 branches lead to condensed particles with over-packed polyglucose chains in the outer reaches which impede glycogenolysis – which is dependent on a coordinated multi-enzyme process. Thus, constitutive and random remodeling of glycogen by lysosomal acid maltase (Devos and Hers, 1980), with α -1,4 glucosidase, α -1,6 glucosidase as well as glucoamylase activities, removes nonfunctioning structural domains and maintains the integrity of cellular glycogen, thus preserving an accessible source of energy with a low osmotic burden in metabolically intense cells. When autophagic function is impaired in Pompe disease, the undegraded residues assume a characteristic pathological appearance: there are prominent

vacuolar spaces engorged with glycogen and other debris, including remnants of cellular organelles and membranes (Lieberman *et al.*, 2012) – features observed in Pompe and Danon disease but increasingly recognized as pathological consequences of many lysosomal diseases.

Detailed studies in Pompe disease immediately revealed the dynamic rôle of the lysosome in cellular biology and added further understanding of autophagy – itself an activity designated from the outset by de Duve. In Pompe disease, glycogen of a near-normal structure accumulates principally in a pathological intracellular network corresponding to the lysosomal compartment deficient in acid α -glucosidase (acid maltase) activity. Thus, not only does the lysosome serve as a recycling center for other effete organelles and membrane fragments, but under resting metabolic conditions, continually remodels cytosolic macromolecules with an apparently normal composition. In the case of glycogen, its physiological rôle is dependent on constitutive remodeling by acid maltase – without which its key function as a dynamic and accessible energy source is lost.

Coordination of Lysosomal Biogenesis and Autophagy – Revelations from Human Pathology

The existence of a process for coordinating transcriptional expression of lysosomal proteins had long been suspected, but identification of the responsible mechanism depended on a series of studies prompted by a clinico-pathological observation in human lysosomal diseases. For many years, those lysosomal diseases associated with extensive storage of partially or completely undegraded parent macromolecules have been known to have an increased tissue activities of lysosomal hydrolases and other proteins associated with pathological expansion or hypertrophy of the lysosomal compartment. The enhancement of acid hydrolase activity (as well as protein abundance) of lysosomal proteins is typically found to be two- to threefold in plasma and often much greater in diseased tissues.

The explanation for increased lysosomal protein expression had long been puzzling until a gene network regulating lysosomal biogenesis and function was identified by Sardiello and colleagues in the Ballabio laboratory. These investigators explored the possibility of a coordinated control system leading to enhanced transcriptional activity common to lysosomal proteins. Using pattern discovery analysis of the promoter region of genes encoding 96 lysosomal proteins, a palindromic decanucleotide motif (termed the coordinated lysosomal expression and regulation element, CLEAR) was identified. The CLEAR element is able to bind to a nuclear transcription factor, TFEB, which is a member of the basic helix-loop-helix leucine-zipper family, sub-family (MiT/TFE), and also a target for silencing by a micro RNA, miR-128 (Sardiello *et al.*, 2009).

Subsequent searches for genome-wide TFEB *cis*-acting motifs indicated that TFEB activates transcription of many target lysosomal genes. Cellular overexpression of TFEB induced biogenesis of lysosomes with expansion of the compartmental network. Enhanced expression of TFEB stimulates lysosomal degradation of glycosaminoglycans and other macromolecular substrates (Palmieri *et al.*, 2011). The transcription factor induced expansion of the pool of lysosomes in

the proximity of the plasma membrane and promoted their fusion at the membrane by raising intracellular Ca^{2+} concentrations with activation of the lysosomal Ca^{2+} channel, mucolipin 1, MCOLN1 which is mutated in mucopolipidosis type IV (Medina *et al.*, 2011).

Reciprocal Signaling between mTORC1 and TFEB on the Lysosome

Autophagy is stimulated by starvation or the withdrawal of growth factors; it is also stimulated by oxidative stress and protein aggregation. While inhibition of mTORC1 stimulates autophagy, the ensuing proteolysis with release of amino acids after autophagy will reactivate mTORC1, and lysosome numbers and compartmental volume will be restored (Yu *et al.*, 2010). As shown, many proteins downstream of mTORC1 contribute to the regulatory control program of autophagy. In the default state of nutrient starvation, the (mTORC1) is localized in the cytoplasm and is dispersed when amino acids are experimentally depleted. In the presence of supplemental amino acids, cellular mTORC1 has a punctate lysosomal localization and autophagy is inhibited (Settembre *et al.*, 2011). In the presence of nutrients, when mTORC1 is translocated to the lysosomal surface, it is able to phosphorylate critical serine residues (including ser 142) on TFEB which impede nuclear entry of the transcription factor. In this way co-localization of the transcription factor permits co-regulation with the master regulator of growth control (mTORC1) – thus simultaneously regulating lysosomal biogenesis and autophagy. In a reciprocal manner, when mTORC1 is inhibited by rapamycin (sirolimus) and everolimus, or by starvation, dephosphorylated TFEB is able to translocate to the nucleus – thereby increasing co-ordinate transcription of genes which encode key proteins involved in autophagy and lysosomal biogenesis. When after starvation, cultures are supplemented with serum, amino acids, insulin, or epidermal growth factor, TFEB nuclear translocation is inhibited compared with the location in cells maintained in starvation medium alone. This suggests that TFEB activation is controlled by a signaling mechanism which is sensitive to nutrient as well as growth factors.

Since starvation induces the transcription of several autophagy genes whereas inhibition of mTORC1 does not, Settembre and colleagues investigated TFEB as a mechanism for activation of autophagy during starvation (Settembre *et al.*, 2011). As explained, without nutrient deprivation, TFEB has a cytoplasmic location, but under conditions of starvation the factor is rapidly translocated to the nucleus; cytosolic TFEB from starved cells had a lower molecular size – a shift occurring transiently and reversibly after relief of starvation with relocation of TFEB from the nucleus and suggesting phosphorylation control. Nuclear localization and activity of TFEB is regulated by serine phosphorylation mediated by the extracellular signal-regulated kinase 2, ERK2, which in turn receives signaling influences from the mitogen-activated protein kinase (MAPK1) whose activity is determined by hormones and extracellular growth factors via the Raf-1–MEK1/2–ERK1/2 pathway. Overexpression of TFEB in HeLa cells and other mammalian cells increases autophagosome abundance and expression of the LC3-II protein, which forms part of the autophagosome membrane. Excess autophagosomes persist after treatment with the vacuolar $\text{v}-(\text{H}^+)$

ATPase inhibitor, bafilomycin, which inhibits autophagosome and LC3-II degradation – indicating that formation of autophagosomes was stimulated *de novo*.

When starved cells growing in normal medium were supplemented with agents, including rapamycin that inhibit either the mTORC, phosphatidylinositol 3-kinase–AKT, or MAPKs, only inhibition of MAPKs was found to induce translocation of the TFEB protein to the nucleus; inhibitors of AKT and mTOR had no effect. Latterly there is evidence that the fundamental mediator implicated in this process is the steady state concentration of Ca^{2+} ions (Medina *et al.*, 2015). Ca^{2+} release from the lysosomal compartment is induced by interactions of TFEB with the phosphatase, calcineurin; thus clinical inhibitors of calcineurin such as ciclosporin, may allow interdiction of these basic cell processes with possible benefit on lysosomal pathology. The master regulatory protein for lysosomal biogenesis, TFEB, also stimulates exocytosis of lysosomal contents by transcriptional effects. This process, which discharges lysosomal contents through docking and fusion of the organelle with the plasmalemma, is associated with dynamic enhancement of lysosomal trafficking and elevation of intracellular Ca^{2+} ions (Settembre *et al.*, 2013).

These recent studies, realized in mammalian cells, provide a convincing mechanistic link between nutrient supply, lysosomal expansion, and the initiation of autophagy under conditions of cellular stress or growth factor signaling. However, while many interacting components have been discovered, the regulatory hierarchies within the interrelated components assembled on large scaffolds in the cell present a puzzling intricacy; furthermore, the contribution of disturbed autophagy and nutrient-sensing to the energetics of the whole individual organism suffering from a particular lysosomal disease is incompletely understood (see Section ‘Gaucher disease’).

Danon disease – an intrinsic disturbance of macroautophagy

Danon disease was identified in CE 1981 in patients with abnormalities of glycogen storage in the heart and skeletal muscle previously thought to have Pompe disease; the patients were found to have normal acid maltase activity. Inherited as an X-linked disorder, heterozygous females die prematurely of progressive cardiomyopathy; in hemizygous males, cognitive impairment, skeletal- and cardio-myopathy is severe and death occurs usually by the age of 25 years. Danon disease is due to mutations in the lysosomal membrane protein LAMP-2 (Malicdan and Nishino, 2012), one splice variant of which (LAMP-2A) is a receptor implicated in chaperone-mediated autophagy. Histopathological examination shows vacuolar myopathy with large intracellular spaces filled with glycogen and other cellular debris, including mitochondria, peroxisomes, ribosomes, and membrane fragments in association with a normal or high acid maltase activity. The pathological features include expanded autophagic vacuoles with characteristic double-unit membranes. In LAMP-2A deficiency, the uptake and degradation of proteins in neurons via chaperone-mediated autophagy is defective: this leads to the accumulation of aggregated complexes of protein such as α -synuclein and huntingtin in the expanded autophagic vacuoles. Danon disease reveals the critical rôle of LAMP-2

protein within the nervous system, as well as skeletal and cardiac muscle.

Other Lysosomal Diseases

Functional Classification defines five groups of lysosomal diseases, paradigmatic examples of each category of which are given below:

Deficiency of a Specific Acid Hydrolase

Gaucher disease

Occurring at any age, this is one of the more frequent lysosomal diseases, usually due to a catalytic deficiency of an acid β -glucosidase, β -glucocerebrosidase (D-glucosyl-N-acyl-sphingosine glucosylhydrolase, EC 3.2.1.45), encoded by the human GBA1 gene (Cox and Schofield, 1997; Grabowski, 2008). The enzyme, a lysosomal glycoprotein with a range of natural substrates that are mixtures of N-acyl-sphingosyl-1-O- β -D glucosides with varying acyl and sphingosine moieties, including those like β -glucosylsphingosine devoid of fatty acids; the enzyme cleaves glucosylceramide into glucose and ceramide. A very few patients with deficiency of the cognate sphingolipid activator protein (SAP-C) required for activity of the enzyme toward its natural sphingolipid substrates have been described. The enzyme, a lysosomal glycoprotein, is an acid β -glucosidase. SAP-C deficiency usually causes a severe disorder intermediate between Gaucher disease and metachromatic leukodystrophy but exceptional patients with disease restricted to the non-neuronopathic manifestations of Gaucher disease are reported.

Numerous mutations responsible for the enzymatic deficiency have been identified in the human GBA1 gene that encodes the β -glucosidase. Very rarely, infants are born with an almost complete lack of β -glucocerebrosidase activity and die within a few days of birth, or are stillborn due to skeletal deformities and loss of skin integrity, with sloughing of layers of scaly skin to cause a collodion-like appearance 'icthyosis lamellaris' – the explanation for this lies in the requirement of the aqueous barrier function of epidermal keratinocytes for very long-chain (C26–C36) acyl ω -hydroxylated ceramides of unusual structure whose biosynthesis is dependent on lysosomal GBA1 activity.

Infantile Gaucher disease (acute neuronopathic) is a very rare condition associated with death in the first year of life; neuronopathic disease is characterized by bulbar palsy due to neuronal cell death in essential mid-brain and pontine nuclei, a decerebrate posture (opisthotonus), and visceral enlargement. The neuronal pathology is associated with loss of eye movement control, respiration, and other visceral functions; it does not respond to any existing therapy.

Chronic neuronopathic Gaucher disease describes less severe forms of the condition occurring in children, adolescents, and young adults, such as supranuclear gaze palsies, incoordinate movement, nerve deafness, and a form of epilepsy known as myoclonic epilepsy with spasmodic involuntary jerking movements of the limbs leading to generalized seizures. This form of Gaucher disease is often stable for many

years but may deteriorate slowly, particularly if the spleen is removed in attempts to ameliorate the accompanying visceral manifestations. This form occurs sporadically in all populations, but particularly in non-European populations where it is the most common form of Gaucher disease, for example, in Arab populations, China, and Southeast Asia. A particular variant of chronic neuronopathic disease has been reported from a province of Northern Sweden, where all individuals are homozygous for a single point transition, at position c. 1448 T->C in exon 10 of the GBA1 gene leading to the missense mutation L444P [p.L483P]; this appears to have arisen by descent from a common Swedish ancestor in the CE sixteenth century and is the most frequent mutant GBA1 allele in most populations.

The most reported but not necessarily most frequent adult form of Gaucher disease as first described by Ernest Gaucher occurs in all populations but is over-represented in Jews of Ashkenazi origin, where a single transition c.1226 A->G, leading to the missense mutation, N370S [p.409S], in the GBA1 gene is most prevalent. The condition does not typically affect the nervous system initially, but the visceral and skeletal manifestations are prominent due to infiltration of the bone marrow, spleen, liver, and occasionally other organs such as the lung, with pathological macrophages ('Gaucher cells') with an alternative form of macrophage activation ('M2'). Characteristically, the condition presents with the consequences of reduced platelet and leukocyte cell concentrations, as well as anemia, with bleeding and splenic enlargement due to pooling and increased destruction of the formed elements of the blood in this macrophage-rich organ. Episodes of acute pain occur in the bones, particularly during growth, followed by necrosis of the bone with later effects on the alignment and integrity of large joints including cartilage. The disease causes stunting of growth in children and delayed puberty with infantilism and lack of matrix and reduced bone mineral content – 'osteoporosis.' While the reduced abundance of circulating blood cells can be temporarily overcome by removing the spleen, this not only induces susceptibility to bacterial infections, but is associated with disabling and more frequent attacks of osteonecrosis.

Comorbid associations of Gaucher disease – the link to common disorders

Energy metabolism in lysosomal diseases

The extent to which the disruption of metabolic homeostasis accounts for cell death and tissue injury in the lysosomal storage diseases is unclear; however disturbed autophagy is likely to affect the metabolic economy of tissues as well as remodeling of glycogen.

An example of disturbed energy metabolism in a lysosomal disease has been well documented in one of the most common diseases, Gaucher disease – of note also, in milder variants that are not associated with neurological manifestations. Indirect calorimetry studies show greatly increased resting energy expenditure in these patients (increased up to 50% compared with that predicted for age, sex, and height and weight). Muscle mass is reduced. In one study, the abnormality was correlated with visceral enlargement, suggesting a relationship to disease bulk due to inflammatory changes and infiltration with pathological macrophages (Barton *et al.*,

1989). These metabolic changes may in part explain the prominent growth retardation observed in children with Gaucher disease; the relationship to the disease is demonstrated by salutary growth responses to introduction of enzyme augmentation therapy – and durable increases in body mass (often >10%) in similarly treated adults.

Disturbed immunity (B-cell lymphoma and myelomatosis)

Gaucher disease is a multisystem disorder, accompanied by numerous plasma and metabolic abnormalities. In the untreated state, the disease shortens life, even when splenectomy has not been done. Co-morbidities include increased frequency of autoimmune conditions; a polyclonal immunoglobulin proliferative response associated with monoclonal immunoglobulin paraproteins – and ultimately increased frequency of B-cell lymphoproliferative disorders leading to B-cell lymphoma and multiple myeloma, the cause of which is not fully understood (Mistry *et al.*, 2013). This may reflect the release of cytokines from the pathological macrophages, or the presentation of the pathological sphingolipids in B-cells via the CD1d surface receptor by Type II natural killer T-lymphocytes (NKT) follicular lymphocytes that secrete cytokines responsible for the alternative activation of macrophages (IL-4,6,10, and 13).

Parkinsonism

Other noteworthy comorbidities of Gaucher disease in affected homozygotes or compound heterozygotes for mutations in the GBA1 gene include Parkinsonian manifestations associated with cognitive impairment and Lewy body dementia (Sardi *et al.*, 2015; Beavan and Schapira, 2013). A striking further observation is that a high frequency of Parkinsonian manifestations occur in parents of patients with Gaucher disease and other heterozygous first-degree relatives (Goker-Alpan *et al.*, 2004); no overt gene-dosage effect of the GBA1 mutations has been established. A further striking consequence is the recent discovery that Gaucher-associated mutations at the GBA1 locus are strongly associated with Parkinsonism in all populations – indeed, these GBA1 mutations have proved to be the strongest genetic determinants of Parkinson's disease (Sardi *et al.*, 2015). Examination of tissues from patients with Gaucher disease and Parkinson disease, as well as those with Lewy body dementia, with or without Parkinson's disease in whom GBA1 mutations have been inherited in heterozygous form, show co-localization of ubiquitin and α -synuclein, together with a core of immuno-reactive glucocerebrosidase in the Lewy body (Goker-Alpan *et al.*, 2010). Protein aggregation of the mutant glucocerebrosidase, together with abnormalities of glycolipid metabolism in association with the primary germline defect, appear to interact unfavorably in nigro-striatal neurones that are the pathological focus of extrapyramidal disease.

These findings were initiated by careful study of the clinical narrative of Gaucher patients with Parkinson's disease: now applied to Parkinson's disease and Lewy body dementia generally, they offer remarkable insight into the neurobiological importance of the lysosome. Evidence that individuals heterozygous for other mutations in other lysosomal proteins, which when mutated are responsible for less frequent lysosomal diseases (e.g., neuronal ceroid lipofuscinosis and Niemann–Pick disease Type C) are pre-disposed to Parkinsonism,

offer potential insights for better understanding of common neurodegenerative diseases with large societal burdens. As a corollary, scientific investigators would wish also that other glycosphingolipid disorders which are very rare but disabling, such as neuronopathic Gaucher disease and GM2 gangliosidosis, will not be excluded from the excitement to develop innovative medicines which may have applications far beyond the narrow perspective of the lysosome (Gahl, 2012).

Defects of Posttranslational Processing and Organelle Targeting

I-cell disease (mucopolipidosis types II and III)

This disease resembles severe Hurler syndrome (MPS type 1), but excess glycosaminoglycans are not present in the urine and cultured fibroblasts from the patients show unusually large, phase-dense inclusion bodies – inclusion cells (I-cells). Almost identical inclusions occur in a milder clinical variant, formerly known as pseudo-Hurler polydystrophy. Multiple deficiencies of acid hydrolases are found in the tissues, but are combined with elevated activities of the cognate enzymes in plasma and other body fluids. The terms mucopolipidosis II and III were later used to reflect the hybrid biochemical features of mucopolysaccharidoses and the sphingolipidoses. In this classification, I-cell disease was called mucopolipidosis II (MLII) and pseudo-Hurler polydystrophy was assigned the term mucopolipidosis III (MLIII). Inherited as autosomal recessive traits, genetic studies in MLII and MLIII revealed interrelated complementation groups that are distinct from other known lysosomal disorders (Leroy *et al.*, 2012).

Although these are very rare inherited disorders (<1 in 200 000 live births), their scientific study lies at the crux of understanding the molecular basis for lysosomal targeting of acid hydrolases based on receptor-mediated uptake of glycoproteins harboring a specific ligand (mannose 6-phosphate). The experimental basis of this research is important and is set out elsewhere in this article (see section 'Functional Complementation of Lysosomal Diseases').

Action myoclonus-renal failure syndrome

Action myoclonus-renal failure syndrome (AMRF) is a distinctive form of progressive myoclonic epilepsy associated with renal failure. It has some biochemical features of Gaucher disease but unlike I-cell disease, affects targeting of nascent β -glucocerebrosidase to the lysosome that is independent of the mannose 6-phosphate ligand (Berkovic *et al.*, 2008). In AMRF, if the renal failure occurs, it is rapidly fatal with focal glomerulosclerosis and features of collapsing glomerulopathy; marked protein loss (nephrotic syndrome) may be present. Some patients develop dementia and peripheral neuropathy; neuropathological examination has showed accumulation of irregular and refractile autofluorescent pigmented granules in non-neuronal cells in the cerebral cortex, globus pallidus, and the cerebellum.

This unique disease is caused by mutations in the SCARB2 gene otherwise known as the gene encoding lysosomal intrinsic membrane protein (LIMP-2). This is another critical lysosomal membrane component, now known to be the chaperone molecule that brings about the 'mannose 6-phosphate-independent'

trafficking of nascent β -glucocerebrosidase to the lysosome (Reczek *et al.*, 2007). Patients with mutations in the LIMP-2 gene develop a clinical syndrome that was not predicted by knowledge of the molecular defect. Assays of fibroblasts obtained from patients show diminished ($\approx 10\%$) residual acid β -glucosidase due to lysosomal glucosylceramidase activity with an abnormal glycosylation pattern, suggesting reduced abundance of post-Golgi forms of the enzyme. However, leukocytes show acid β -glucosidase activity within the healthy reference range. Thus the condition is remarkable in that the deficiency of LIMP-2 function leads to a mosaic biochemical phenotype with normal activity of acid glucocerebrosidase due to satisfactory targeting to the lysosome in macrophages and leukocytes, but a deficiency of acid β -glucosidase activity in fibroblasts and other somatic cells in which LIMP-2 particularly exerts its chaperone rôle for delivery of nascent glucocerebrosidase to its lysosomal destination.

Deficiency of an Activator Protein

Sphingolipid activator proteins

Several proteins have been recognized as key accessories to the action of lysosomal proteins, and when deficient as a result of genetic mutation usually lead to loss of function affecting multiple systems (Kolter and Sandhoff, 2005).

In this category are deficiencies of sphingolipid activator proteins, which are cleaved products of a single gene (SAP). Each activator has distinct functions based on the activities of the discrete polypeptides, saposins A–D. Generally, inherited defects of saposins lead to complex disorders with functional deficiency of one or more lysosomal hydrolases toward their natural but not usually their artificial substrates – thus often leading to difficulties in diagnosis based on conventional biochemical assays in general use. For example, deficiency of Sap C usually leads to a complex syndrome in infants with features of metachromatic leukodystrophy and Gaucher disease due to loss of the ability of the saposin to present substrates (glucosylceramide and sulfatide) to β -glucosylceramidase and arylsulfatase A, respectively. Rare cases of saposin C deficiency have been associated with an attenuated juvenile form of Gaucher disease, without overt features of ‘metachromatic leukodystrophy.’ In these patients, administration of exogenous enzyme is unlikely to be beneficial because the activator protein is critical for its catalytic function. It is also notable that activity of acid β -glucosidase or arylsulfatase A (determined with fluorogenic or chromogenic substrates) is not deficient, even though many pathological features of disease are present and tissue biopsy specimens (and urine samples in the case of sulfatide) reveal characteristic lysosomal storage of sphingolipid.

The single-chain polypeptide, GM2 activator protein, plays a critical rôle in presenting membrane-bound ganglioside substrate to hexosaminidase A; GM2 activator and the heterodimeric hexosaminidase A are absolutely required for breakdown of this complex sphingolipid substrate. Very rare mutations in GM2 activator cause a disease that is clinically indistinguishable from late-infantile Tay-Sachs disease but, again, without any deficiency of hexosaminidases toward the fluorogenic substrates in routine diagnostic use. Moreover, other rare variants of Tay-Sachs disease, first identified within an Hispanic population, occur without apparent deficiency of

hexosaminidase A but with a mutation (R288H) in the HEXA gene rather than the GM2 activator gene: in this case, the single amino acid change in the α -subunit of hexosaminidase abrogates binding to the GM2 activator protein. Based on conventional biochemical testing, hexosaminidase activity toward artificial fluorogenic substrates is unremarkable, but the activity toward the natural GM2 substrate is impaired.

Intensive biochemical and genetic studies of a severe and disabling condition known as ‘multiple sulfatase deficiency,’ revealed the presence of a unique modifying factor, sulfatase modifying factor (SUMF). When SUMF is inactivated, deficiencies of numerous sulfatases occur in association with features of MPS, dry scaly skin (ichthyosis), and leukodystrophy (Sardiello *et al.*, 2005). The factor is an enzyme located on the endoplasmic reticulum that serves to modify all sulfatases present in mammals, thereby activating them. For enzymatic activity, a conserved cysteine residue present in a short peptide sequence in all eukaryotic sulfatases must be oxidized to formylglycine – a critical posttranslational modification. Patients with multiple sulfatase deficiency and harboring two inactivating mutations in the human SUMF gene have a variable phenotype with multiple deficiencies – including the numerous sulfatase activities implicated in sequential degradation and turnover of the many highly sulfated glycosaminoglycans. Neurological disease due to arylsulfatase A deficiency with features of metachromatic leukodystrophy and extensive demyelination is usually present; deficiency of steroid sulfatase appears to account for the ichthyosis.

Galactosialidosis

This monogenic disease is ultimately caused by deficient activity of three lysosomal acid hydrolases that are assembled into a stable functional complex (Caciotti *et al.*, 2013). The clinical features are shared in part by Hurler disease (MPS I): coarse facies, vertebral changes and storage foam cells in the bone marrow with vacuolated lymphocytes, as well as cherry-red spots in the macular region of the retina and visceromegaly. However the patients are deficient in β -galactosidase and neuraminidase, leading to the accumulation of sialyl oligosaccharides and glycopeptides in fluids and connective tissues. Mutations in the gene encoding what has come to be known as protective protein, otherwise designated cathepsin A, are responsible for galactosialidosis. The condition arises because the mutated protein, a serine carboxypeptidase, is unable to carry out its normal function by binding to β -galactosidase and neuraminidase, which would assure the physiological trafficking of the two newly formed glycosidases to the lysosome in a stable multi-enzyme complex.

Deficiency of a Lysosomal Membrane Protein or Transporter

Mutations in genes encoding this class of lysosomal proteins have been identified in several disorders and have proved scientifically informative (Mancini *et al.*, 2000).

Several of these defects show preferential expression in lysosome-related organelles of specialized cells and illustrative examples are provided in Section ‘Defects in the Biogenesis and Function of Lysosome-Related Organelles.’

Cystinosis

Cystinosis is one of the first lysosomal disorders to be accessed by a specific treatment (Nesterova and Gahl, 2013). The disease is a transport defect which causes systemic disease typified by the early development of renal tubular acidosis, with salt wasting and a Fanconi syndrome characterized by amino aciduria, phosphaturia, and glycosuria. Premature rickets and skeletal demineralization with episodes of dehydration and salt-wasting are followed by full-blown renal failure; thyroid failure, heart disease, and deposits of cysteine in the cornea, leading to impaired vision, are other characteristics. Patients with cystinosis usually require renal transplantation and other supportive measures, but the introduction of the aminothiols agent, cysteamine, to remove by complexation the accumulated intra-lysosomal cysteine is a crucial measure for delaying the advance of multisystem failure. Topical use of cysteamine has a notable effect on cystine crystals in the cornea and pre-empts the inflammation which causes ulceration effectively preserves vision. Cystinosis is due to mutations in CTNS encoding the transporter, cystosin, an integral protein of the lysosomal membrane; the disease arises as a result of a failure of the lysosome to export the products of intra-lysosomal proteolysis – cystine is formed from the oxidative dimerization of cysteine and as a result, non-protein free cystine accumulates in the lysosomes.

Clinically distinct inborn defects of lysosomal transport include the very rare ‘Salla disease,’ principally confined to persons of Finno-Ugric ancestry, in which free sialic and glucuronic acid accumulate in the lysosomal compartment due to a failure of egress after release from the breakdown of glycosaminoglycans and sialoglycoprotein digestion. The manifestations include hypotonia, cerebellar ataxia, and dementia; coarse features and enlargement of the liver and spleen occur with distended lysosomes identified by ultrastructural studies of tissue samples. Large amounts of free sialic acid is excreted in urine and a H⁺/anionic sugar symporter system for sialic acid and glucuronic acid export is impaired due to mutations in SLC17A5, encoding a protein (sialin), with a transport function typical of a family of anion/cation symporters.

There is also a very uncommon ‘cobalamin transport disorder,’ leading to intracellular deficiency of both active methyl- and adenosylcobalamin cofactors; this is classified as a cblF defect and has been shown to be due to mutations in the human LMBRD1 gene. The cognate lysosomal membrane protein apparently exports free cobalamins released by digestion of the transcobalamin II-receptor complex after internalization and transfer to the lysosomal compartment. Recently, genetic investigations into rare patients with a phenocopy of the cblF defect, have identified another cause of failure to release vitamin B12 from lysosomes: causal mutations in ABCD4, a gene that encodes an ABC transporter, have been identified (Coelho *et al.*, 2012). These genetic studies shed new light on the physiology of cobalamin metabolism, since ABCD4 was previously considered to be expressed on the peroxisomal membrane. Studies in fibroblasts from patients and healthy individuals show that ABCD4 co-localizes with the lysosomal proteins LAMP1 and LMBD1, the latter being deficient in the cblF defect. Since mutations affecting the putative ATPase domain of ABCD4 impair function, it seems likely that ATPase activity is a functional requirement for

cobalamin release during intracellular processing of vitamin B12.

Danon disease

Danon disease represents another lysosomal membrane defect not directly involving transporter activity, but LAMP-2, a protein, mapping to the X-chromosome and implicated in the control and activation of macro-autophagy. As a large highly glycosylated integral membrane glycoprotein, LAMP-2 serves a critical coordinating function between housekeeping lysosomal digestion and the disposal of large engulfed particles after formation of autophagic vacuoles – processes linked to cellular nutrient-sensing. LAMP-2 deficiency in Danon disease has principal effects in cardiac muscle, neurons, and skeletal muscle, tissues in which the constituent cells are particularly active in autophagy.

Defects in the Biogenesis and Function of Lysosome-Related Organelles

Abnormal biogenesis of lysosomes is illustrated by Hermansky–Pudlak diseases (Type 1–9); Chédiak–Higashi disease and the Griscelli syndromes Types I–III

The Hermansky–Pudlak syndrome describes a group of instructive conditions caused by genetic defects of lysosome-related organelles – these have far-reaching effects in the patient but also inform understanding of the biogenesis and specialized functions of several critical cellular constituents.

Lysosome-related organelles are found principally in specialized cell populations in which they serve to store and secrete proteins required for specific functions, for example, the platelet, the melanocyte, and T-lymphocyte (Luzio *et al.*, 2014). Diseases affecting lysosomal biogenesis and the formation of lysosome-related organelles point to essential functions of genes hitherto not recognized, but which have crucial rôles in pigmentation, innate and adaptive immunity, as well as blood coagulation; they provide, if it were needed, further examples of the debt of molecular cell science to the study of human disease. It appears that organelles such as the melanosome interact physically with mitochondria, endoplasmic reticulum at specific membrane-contact sites during biogenesis – a process mediated by mitofusin 2, a fibrillary protein which allows lipid transfer and regulates Ca²⁺ fluxes that affect membrane dynamics as well as their size, shape and motility (Daniele *et al.*, 2014).

Hermansky–Pudlak and Chédiak–Higashi syndromes are autosomal recessive conditions typically associated with oculo-cutaneous albinism and defective hemostatic function due to deficiency of the storage pool of coagulation factors in the dense δ granules of the blood platelet (Huizing *et al.*, 2008; Seward and Gahl, 2013; Introne *et al.*, 2015).

These diseases disturb the function of intracellular vesicles that are related to lysosomes, including the platelet dense body (containing serotonin, ADP and other nucleotides as well as Ca²⁺ ions) and the melanosome, in which melanin pigment is packaged. Hermansky–Pudlak syndrome appears to be a disorder of vesicle formation and components of an adaptor complex responsible for fashioning the structures from the trans-Golgi network.

Chédiak–Higashi syndrome is associated with defects in vesicle trafficking of critical intracellular cargo mediated by GTPase activity; in this condition, the *LYST* (lysosomal trafficking regulator gene), which encodes a large protein of ill-understood function, is mutated. Chédiak–Higashi syndrome causes not only albinism, but severe immunodeficiency and mild bleeding; abnormal inclusions in white blood cells and deficiency of platelet-dense bodies are associated with a mild bleeding tendency and recurrent bacterial, and occasionally viral or fungal infections. These microbial infections, which may be severe, typically affect the skin and respiratory tract. By the time they reach adult life, patients develop neurological symptoms affecting peripheral nerves, the extrapyramidal system, and cerebellum. Also, most of the patients enter an accelerated phase of lymphoproliferation which, without the availability of pre-emptive hematopoietic stem-cell transplantation, usually proves fatal – in a syndrome of hemophagocytic lymphohistiocytosis.

Hermansky–Pudlak syndrome occurs in nine distinct genetic forms, and although the function of those other than the adaptor protein AP-3 B1 are not thoroughly understood, these gene products interact in the biogenesis of lysosome-related organelle complexes. In the case of Hermansky–Pudlak syndrome, abnormal δ granules in the platelets and secretory granules in T-cells, as well as the lamellar bodies in Type II alveolar cells are found. Hermansky–Pudlak syndrome is characterized by a marked bleeding diathesis, particularly after surgery (Seward and Gahl, 2013). The albinism requires careful management to reduce the risk of light-induced cancers. In some of the variants, progressive pulmonary fibrosis occurs to the extent that lung transplantation may be required as a life-saving measure. Infections occur as a result of defective granulocyte function of peripheral white blood cells.

Related Conditions Affecting Organelle Biogenesis

Griscelli syndromes (Types 1–3) are also associated with albinism and immune deficiency: in Type 1, cognitive impairment and neurological function are frequent. In Griscelli syndrome Type 2, episodes of disseminated uncontrolled phagocytosis with gross hemolysis and fever develop due to uncontrolled activation of macrophages and T-lymphocytes, with widespread infiltration of lymph nodes and visceral organ as a result of unregulated natural-killer T-cell function. Griscelli syndrome occurs in three genetically distinct variants: Types 1 and 2 are caused by mutations in genes that localize to the same region of chromosome 15q (*Myosin-5a* and *RAB27a* genes) encoding proteins which bring about vesicular transport, especially exocytosis of cytotoxic granules (Krzewski and Coligan, 2012). Griscelli syndrome Type 3 is caused by mutations in *MLPH* – a gene that encodes melanophilin, a protein that complexes with Rab 27a and myosin Va implicated in the transport of melanosomes. The most severe of these syndromes is Type 2 Griscelli syndrome where the hemophagocytic crisis often proves fatal unless hematopoietic stem-cell transplantation is undertaken.

Lysosome-Related Organelles and the Osteoclast

As noted, lysosomal-related organelles participate in functions that are external to the cell. The modeling of bone is brought

about by specialized cells of mononuclear-phagocyte origin, termed osteoclasts. Digestion of mineralized osteoid occurs as this population of modified scavenger cells continuously digests bone matrix – a key function in the repair of fractures as well as maintenance of skeletal strength and integrity throughout life. The specialized sealing zone which encloses the ruffled cell border of the osteoclast, the so-called resorptive lacuna, is in effect a multi-exteriorized space formed by lysosomal fusion. The subcellular compartment is acidified through the action of the vacuolar V-(H⁺)-ATPase and mediated by the chloride conductance channel, *CLCN7*, which stabilizes the chemiosmotic gradient within the closed environment of the lacuna.

Failure to acidify this sub-osteoclastic resorptive vacuole as a result of mutations in the lysosomal V-(H⁺)-ATPase impairs diverse resorptive functions; this leads to formation of dense, un-remodeled bones ('osteopetrosis') (Frattoni *et al.*, 2000). Mutations in the gene, *CLCN7*, encoding the lysosomal chloride channel expressed on the lysosomal membrane but, in the adapted exteriorized lysosomal compartment of the osteoclast, on the ruffled border, also cause human osteopetrosis – as a result of failed acidification in the subcellular vacuole. This verifies the key function of the chloride counterion transporter in osteoclast activity (Kornak *et al.*, 2001). Mutations in cathepsin K, an acid hydrolase expressed by the osteoclast with selective activity toward collagen, elastin, and gelatin in cartilaginous and bone matrix, also impair bone modeling and cause a variant of osteopetrosis – 'osteopycnodysostosis' (Gelb *et al.*, 1996). This hereditary disorder describes mis-shaped hyper-dense bones associated with dwarfism.

Beyond their rôle in digestion, there are examples illustrated by diseases of specific function of lysosomal components in other specialized cells. In cells that contribute to immune function, and myeloid lineage shared with the osteoclast, a rare inherited deficiency of the iron-containing type 5 purple acid phosphatase (*Acp5*) has far-reaching effects. This lysosomal protein is a di-iron enzyme that is highly expressed in osteoclasts and, widely termed 'tartrate-resistant acid phosphatase (TRAP)', has long been used as an osteoclast 'marker' in histochemical studies. Type 5 acid phosphatase is also abundantly expressed in other cells of the mononuclear phagocytic phenotype, including dendritic cells and Langerhans cells in the dermis. The enzyme dephosphorylates the skeletal sialoprotein, osteopontin – thereby attracting attachment of osteoclasts to bone matrix where it also stimulates formation of their ruffled borders at the resorptive interface.

However, the pervasive function of this lysosomal protein has been revealed by studies of informative mutants informed by characterization of mice in which the gene has been mutated. Mutations in *ACP5* are responsible for a recessively inherited immuno-osseous dysplasia, 'spondyloenchondrodysplasia,' in mice and humans (Briggs *et al.*, 2011). Spondyloenchondrodysplasia affects ossification and bone modeling in vertebrae and long bones, causing short stature; in addition, the immune system is dysregulated with the development of autoimmunity, including systemic lupus erythematosus and Sjögren's syndrome. These disparate phenomena are in part explained by the accessory rôle of the soluble phosphorylated isoform of osteopontin as an early

T-lymphocyte activation factor (ETA-1) with effects on integrin binding that influence expression of other immuno-regulatory cytokines such as IL-10 and IL-12 (Bune *et al.*, 2001).

Functional Complementation of Lysosomal Diseases

In 1967, Elizabeth Neufeld, then working at the National Institutes of Health in Bethesda, USA, was aware that meta-chromatic deposits in the fibroblasts indicated the accumulation of glycosaminoglycans; at the time intrigued by the biosynthesis of uronic acids, she investigated fibroblasts cultured from skin biopsies obtained from patients with Hurler syndrome (MPS I). Excessive formation of glycosaminoglycans could not be demonstrated and thus with Joseph Fratantoni, Neufeld studied the incorporation of ^{35}S -labeled sulfate into TCA-insoluble material. In contrast to healthy fibroblasts, where the incorporation of sulfate leveled off after a few hours, radioactive incorporation into the glycosaminoglycans precipitated from the fibroblasts of a Hurler patient, continued to increase linearly over several days of culture. Similar behavior occurred in the clinically similar but genetically distinct (X-linked) condition, Hunter disease (MPS II). By accident, Hunter and Hurler cells were combined and, unexpectedly, sulfated glycosaminoglycan accumulated normally in the mixed culture; after pre-labeling with radioactive sulfate, the rate of degradation (retarded, as expected, in cells from patients with MPS syndromes), was restored to that observed in fibroblasts from healthy subjects (Fratantoni *et al.*, 1968). Eventually Neufeld and colleagues showed that a factor in the medium, later called a 'corrective factor,' was secreted from the diseased fibroblasts mutually to complement the degradative defects in fibroblasts obtained from patients with genetically and clinically distinct syndromes belonging to discrete complementation groups (Fratantoni *et al.*, 1969). In collaboration with Dr. Victor McKusick, and thus informed by scrupulously characterized biopsy subjects, Neufeld and colleagues defined specific corrective factors for MPSI, MPSII (Hunter and Hurler syndromes) and different complementation groups in Sanfilippo syndromes (MPS III, Sanfilippo A and B), as well as Maroteaux-Lamy disease (MPS VI) (Neufeld, 2011).

Investigations by William Sly and colleagues in St Louis showed that acid β -glucuronidase deficiency in fibroblasts obtained from patients with a cognate mucopolysaccharidosis (MPS VII – Sly syndrome) could be corrected by supplying the enzyme exogenously in the medium of the culture. Later Sly, shortly followed by Neufeld, Shapiro, and others, found that a particular high-molecular weight (phosphohexosyl) form of the respectively deficient enzyme, released into the medium from cultured fibroblasts, possessed the complementing property for therapeutic correction, *in vitro*. Eventually, Sly and colleagues showed that the structural feature required for saturable uptake (of acid β -glucuronidase), was mannose 6-phosphate – a carbohydrate modification previously unknown in mammalian cells (Natowicz *et al.*, 1979). The general applicability of this phenomenon was confirmed by Neufeld and colleagues who showed that excess mannose 6-phosphate in the medium inhibited the uptake of complementing α -L-iduronidase, the enzyme activity deficient in MPS I (Hurler disease).

I-Cell Disease

Leroy and DeMars also in CE 1967 reported striking abnormalities in fibroblasts cultured from patients suffering from a very rare condition with combined clinico-pathological features of a sphingolipidosis and mucopolysaccharide disorder – the condition was named inclusion-cell disease (I-cell disease) but its other name is mucopolipidosis (on the basis of clinical and genetic criteria, now sub-classified into types II and III). While the cells and tissues of the patient showed numerous deficiencies of lysosomal acid hydrolases and abundant lysosomal storage of glycosaminoglycans and sphingolipids, high activities of these enzymes (up to 20-fold greater) than healthy reference values could be detected in plasma and urine. Indeed, this behavior was mimicked by I-cell fibroblasts in culture: they remained intrinsically deficient with a 'storage' phenotype but released many enzymatically active acid hydrolases into the medium (Leroy and DeMars, 1967).

A decisive finding, in the 1970s however, was that complementation studies as first pioneered by Elizabeth Neufeld and colleagues showed that the acid hydrolases now present in excess in the medium were able neither to complement I-cell fibroblasts nor fibroblasts cultured from patients with various mucopolysaccharidoses. Hickman and Neufeld also made the critical observation that I-cell fibroblasts were capable of taking up acid hydrolases secreted into culture medium by healthy fibroblasts – leading to biochemical complementation. The findings suggested that lysosomal enzymes contain a recognition marker for uptake and transport to lysosomes; fibroblasts from patients with I-cell disease, which secrete abundant lysosomal acid hydrolase into the medium, are unable to generate the appropriate ligand (Hickman and Neufeld, 1972). This explanation was definitively confirmed in molecular terms by William Sly and colleagues with the identification of mannose 6-phosphate as the principal recognition moiety (Natowicz *et al.*, 1979). The conceptual insight that, as a result of intracellular mis-sorting, the enzymes were not trafficked to their lysosomal destination, led to the identification of a default pathway allowing active egress to the medium. Careful radio-tracer studies showed that numerous secreted lysosomal glycoproteins had escaped phosphorylation during passage through the *trans* Golgi system but were nonetheless actively transported to the plasma membrane and released. In contrast, cells from patients without I-cell disease synthesize acid hydrolases destined for the lysosome which harbor the mannose 6-phosphate (M6P) recognition signal; this is incorporated within oligosaccharide moieties as the nascent polypeptides traverse the Golgi network, there undergoing further posttranslational processing with acquisition and unmasking of the appropriate ligand – which is retained for intracellular sorting and delivery to the lysosomal compartment. This sorting pathway for newly formed acid hydrolases destined for the lysosomal matrix is not absolute, it is 'leaky.' About 15% of the complement of nascent enzymes generated by human cells in culture is directed to the plasma membrane and is released into the medium – from which it can be taken up after binding to other cell-surface M6P receptors. The so-called secretion-reuptake pathway appears not to be significant for most lysosomal hydrolases under healthy conditions *in vivo*.

Subsequent research by Stuart Kornfeld and Kurt Von Figura and their colleagues respectively in the US and Germany led to the characterization of mannose 6-phosphate receptors on the plasma membrane and of two intracellular enzyme complexes that catalyze formation of the mannose 6-phosphate recognition marker – UDP-*N*-acetylglucosamine: glycoprotein *N*-acetyl glucosamine-1 phosphotransferase and the second ('unmasking') enzyme, *N*-acetylglucosamine phosphoglycosidase (von Figura and Hasilik, 1986; Kornfeld, 1987). Patients with I-cell disease are deficient in the activity of the phosphotransferase – a multi-subunit polypeptide complex encoded by two genes. It is worth noting that in I-cell disease, the lysosomes in some tissues, such as the liver, certain macrophage populations, kidney, and brain have a near-normal complement of acid hydrolases although all are deficient in phosphotransferase activity, implying that distinct pathways for sorting the nascent hydrolases to the lysosomal compartment is used preferentially in some cell types and mannose 6-phosphate is not the only signal for delivery to lysosomes. Further, the tartrate-sensitive lysosomal acid phosphatase isozyme, Acp2, and β -glucocerebrosidase (deficient in Gaucher disease), do not employ the M6P signal (Reczek *et al.*, 2007). In Acp2, an exposed tyrosine residue signal in the cytoplasmic tail of this membrane polypeptide directs the nascent enzyme to the lysosome. Here in cooperation with the purple acid phosphatase, Acp5 (TRAP), the acid phosphatase has been shown to catalyze removal of the phosphate moiety from the M6P-linked glycoprotein hydrolases within the lysosomal matrix space (Makrypidi *et al.*, 2012). The mannose 6-phosphate-independent mechanism for targeting β -glucocerebrosidase to the lysosome, at least in some tissues, involves the membrane glycoprotein LIMP2 which serves a specific chaperone molecule. It now appears that there are several mannose 6-phosphate-independent pathways, some of which are not fully characterized (Kornfeld, 1987); it is nonetheless unclear as to why more than one sorting pathway is needed for lysosomal biogenesis.

Lysosomal Complementation *In Vivo*

The mannose 6-phosphate receptor system (cation-independent mannose 6-phosphate/insulin-like growth factor II receptor) system proved to be functionally 'leaky' *in vivo* – thus allowing a fraction of nascent lysosomal proteins to be released from cells, later to be taken up by cell surface receptors of other cells present in tissues at a distance. The term, 'secretion-recapture,' has been coined for this constitutive process.

Beyond experiments with cultured cells *in vitro*, clinical observation supports the concept of secretion-recapture for functional complementation of genetic defects of the lysosome. While most lysosomal diseases are inherited as autosomal recessive traits, three informative disorders are transmitted on the X chromosome – Hunter syndrome (MPS II, caused by mutations in the α -iduronate sulfatase gene), Anderson-Fabry disease (caused by mutations in the acid α -galactosidase A gene), and Danon disease (mutations in LAMP-2, lysosomal-associated membrane protein-2). Female heterozygotes with Danon disease are invariably symptomatic; however, unless unusual events have occurred on the X chromosome, with skewed Lyonization, female heterozygotes for MPS II do not show signs of this disease. Since, as a result of Lyonization, human females are

mosaics for their respective parental X chromosomes, this shows that secretion-recapture of appropriate molecular forms of iduronate sulfatase, which is deficient in affected males with Hunter disease, can be complemented *in vivo* by cells expressing the wild-type parental allele. While intercellular complementation of defective lysosomal membrane function *in vivo* will not readily occur, it is unclear why female heterozygotes for Fabry disease are symptomatic, although plasma, as opposed to tissue glycoforms of acid α -galactosidase A, are not readily taken up by human fibroblasts – a speculative possibility is that in this instance, inactivation of the requisite recognition marker occurs in the circulation. While the investigations of several informative human inherited disorders of lysosomal function have been very revealing for understanding the function of lysosomes and the nature of their biogenesis within the living cell, the laboratories of Elizabeth Neufeld and William Sly proved in the end not to be the first to deliver effective therapy involving enzymatic augmentation for lysosomal diseases (see Gaucher disease).

Treatment of Lysosomal Diseases

All these developments are predicated on careful experiments that show the central rôle of the lysosome in the cell and its critical functions in reclaiming and recycling of macromolecular constituents in metabolic homeostasis and are reviewed elsewhere (Cox, 2015).

Therapeutic stratagems specifically for lysosomal disease include:

- a. Dissolution and removal of intralysosomal storage material
- b. Enhanced exocytosis of pathological intralysosomal content
- c. Modulating substrate biosynthesis (substrate-reduction therapy)
- d. Complementation stratagems
 - i. Hematopoietic stem-cell (previously bone-marrow) allo-transplantation
 - ii. Reinfusion of vector-transduced autologous hematopoietic stem cells
 - iii. Direct gene transfer
 - iv. Enzymatic augmentation.

Dissolution and Removal of Intralysosomal Storage Material

Cystinosis

This condition is caused by genetic defects in the lysosomal cysteine/proton symporter, cystinosin. Systemic accumulation of the oxidized sulfamino acid causes widespread injury, particularly in the renal tubule, and also to the thyroid, heart, and other tissues related to the presence of crystalline aggregates of oxidized cystine dimers. A successful and simple biochemical approach to the matter of this storage has been long adopted and the aminothiol, cysteamine which interacts with the aggregates is capable of converting cystine to a mixed disulfide of cysteine-cysteamine which can leave the lysosomal via a cationic amino acid transporter system, which also transports the basic amino acid, lysine. If high doses of cysteamine are given early in the disease, there is a significant delay of progression to renal failure and topical therapy with this aminothiol can

prevent loss of vision and painful damage to the cornea due to local cystine (Nesterova and Gahl, 2013). This therapeutic approach, not only of clinical importance, shows that under some circumstances, the simple presence of excess material within the lysosomal compartment (crystalline cystine) is a credible therapeutic target, and moreover, a verifiable basis for the ensuing pathology and lysosomal dysfunction.

Niemann–Pick disease type C

Niemann–Pick disease Type C, is due to mutations in genes encoding cognate proteins (NPC1 and NPC2) that collaborate to bring about the egress of free cholesterol and other lipid molecules from the endolysosomal system. Experimental work in Niemann–Pick Type C in mice and cats shows that cyclodextrins, which are cyclic oligosaccharides with amphipathic properties when given intravenously have beneficial effects. Intrathecal cyclodextrin may cause deafness, but clears intracellular cholesterol that accumulates in Niemann–Pick disease Type C, prevents the onset of the phenotype in mice and cats – currently, and with high expectation of clinical benefit, this drug is currently undergoing Phase 1 and Phase 2 studies in human patients with NPC.

Enhanced Exocytosis of Pathological Intralysosomal Content

A further method directly to mitigate the effects of lysosomal storage is based on stimulating exocytosis and trials are underway in Niemann–Pick disease Type C. Experimental studies in several lysosomal disease models in which expression of heat-shock proteins that serve as molecular chaperones by exposure to the hydroxylamine co-inducer of the heat-shock response (arimoclomol) appears to be beneficial. Now licensed to the company Orphazyme, this agent is in clinical investigation for Niemann–Pick disease Type C – the results of this bold initiative are awaited, but promise an improved understanding of the rôle of excess lysosomal content in pathogenesis in this disease.

Experimentally enhanced TFEB activity in cell models of multiple sulfatase deficiency and Pompe disease rapidly induced exocytosis of lysosomal storage contents. Preliminary results showing glycogen clearance in muscles of mice that model Pompe disease reported by the Ballabio laboratory suggest that manipulation of the transcriptional control network for lysosomal biogenesis and autophagy through the TFEB may justify further clinical experimentation. However, since TFEB has been implicated in renal and other human cancers, therapeutic manipulation of its expression will clearly require cautious exploration.

Modulating Substrate Biosynthesis (Substrate-Reduction Therapy)

Understanding the biochemical pathology of metabolic diseases has a productive history in therapeutic development. Hypercholesterolaemia, hereditary tyrosinemia, and gout are metabolic disorders readily identified as over-production diseases due to the inability satisfactorily to degrade toxic metabolites. In the case of the sphingolipidoses, which are an important group of lysosomal diseases the identification of

inhibitors of biosynthesis has led recently to productive clinical experimentation. The stratagem remains under clinical development in Fabry disease, chronic neuronopathic Gaucher disease, and glycosphingolipidosis such as Niemann–Pick disease Type C and GM2 gangliosidosis. Early success in this field is also stimulating investment to discover tractable inhibitors to rebalance disordered glycosaminoglycan metabolism in the mucopolysaccharidoses. In the case of small molecules that may be chosen as selective inhibitors of biosynthetic processes, it is more likely that an orally active inhibitory molecule can be obtained with the required specificity, as well as the capacity directly to traverse the blood–brain barrier and exert its effect in intractable neurological diseases.

The pharmaceutical use of inhibitors of biosynthesis is most often successful when the target is an early committed step critical for the formation of the toxic products. In the case of the sphingolipid molecules, there are numerous critical functions: ceramides stimulate apoptosis and inhibit cell growth, as well as differentiation. At the same time, the chemically related sphingoid bases stimulate proliferation and control inflammatory processes, as well as the migration of cells. However, both nitisinone and the statin drugs prove to be highly successful agents in medical use, since they attenuate, rather than abrogate, the biosynthesis of the essential precursors in hereditary tyrosinemia type 1 and hypercholesterolaemia, respectively. Early success in this approach has occurred in Gaucher disease, a sphingolipidosis, where the putative toxins accumulate in the face of the lysosomal deficiency acid β -glucosidase are the many variant *N*-acyl-sphingosyl-1-*O*- β -D glucosides with differing acyl and sphingosine moieties, as well as their unacylated congenors, the glucosylsphingosines. In the adult-onset, and therefore milder clinical forms of Gaucher disease, β -glucosylceramides are derived from exogenous membrane components, particularly during the phagocytosis of blood cells by macrophages. Lysosomal recycling of these cellular glycosphingolipids is limiting, but where the genetic defect is more severe, in cells such as neurons, high rates of endogenous turnover of glycosphingolipids in ganglioside-rich neurones are associated with neurological injury.

Currently, two main classes of inhibitor have been explored therapeutically for the treatment of the sphingolipidoses. The target enzymatic reaction is catalyzed by UDP-glucosylceramide transferase, which leads to the formation of β -glucosylceramide that accumulates in Gaucher disease. This glycosphingolipid serves as a scaffold for the synthesis of higher-order glycosphingolipids, such as globosides and gangliosides. An acylated iminosugar, *N*-butyldeoxynojirimycin, first previously explored for an unrelated application in the control of human immunodeficiency virus, was chosen for clinical development in the sphingolipid disorders, including Gaucher disease, by the author working with FM Platt and T Butters in the University of Oxford. *N*-butyldeoxynojirimycin inhibits the biosynthesis of glucosylceramide in cultures of murine macrophages and in animals with the gangliosidosis, Sandhoff disease lacking β -hexosaminidases A and B, the agent decreases accumulation of GM2 ganglioside in the brain thereby increasing survival. Having already been administered to humans as a potential

inhibitor of the viral α -mannosidases required for the synthesis of the HIV coat glycoprotein, the drug was re-purposed for trial in Gaucher disease. It was found to have a modest therapeutic effect, reducing visceral enlargement and slowly improving hematological parameters, as well as macrophage molecules such as the enzyme chitotriosidase, which is a biomarker of the macrophage effects of Gaucher disease. Miglustat ultimately gained a license and was approved as a second-line agent for Gaucher disease in patients unable to tolerate or receive intravenous infusions of corrective enzyme. Miglustat has drawbacks, however, and although an already available small molecule it has many unwanted effects, for example on intestinal disaccharidases, thus inducing abdominal symptoms.

The iminosugar, designated miglustat, was able to enter the brain and was later shown to have modest effects on some clinical manifestations in patients with Niemann–Pick disease Type C, for which it has also been approved and is the only currently available drug for this neurodegenerative disease. While other congeners of the approved iminosugar for a lysosomal disease are in development, miglustat has not found great favor internationally, mainly because of its unwanted gastrointestinal effects. However, the ostensible success of the agent shortly after launching had a signal effect in stimulating further investment and competitive interest. Scientifically this was based on longstanding research by Norman Radin and his junior colleague James Shayman of the University of Michigan, Ann Arbor. This group pioneered the chemical development of inhibitors based on a ceramide structure, which unlike miglustat eventually led to the development of congeners that were potent highly specific inhibitors of UDP glucosylceramide with a K_i of approximately 25 nM. The lead agent appears to be a transition-state analogue and remain orally active, although it does not enter the brain. After 15 years of developmental work and several trials spanning many years, the final compound, termed eliglustat, has been approved for first-line oral use in adults with Type 1 Gaucher disease. Eliglustat has been reported to have salutary effects that resemble the efficacy recorded in naïve patients treated with enzyme therapy. The drug was found to be non-inferior in 2-year studies of approximately 160 patients worldwide who had previously been stabilized on imiglucerase enzyme therapy; more than 400 patients have now been exposed to the drug worldwide.

The burgeoning success of this developmental program, which required major investment from the Genzyme company, has led to further developments and the filing of a patent for chemically novel UDP-glucosylceramide transferase inhibitors that may be used in a range of metabolic conditions. Their particular interest for therapeutic evaluation will be particularly those devastating neurodegenerative disorders of sphingolipid metabolism which affect the brain including those which impair lysosomal recycling of glycosphingolipid constituents in neural cell membranes. Genzyme, now a Sanofi company, has obtained an innovative lead molecule that is chemically distinct from the prototypic inhibitors first developed by Radin, Shayman, and their colleagues. The entirely novel compound, which appears to have favorable biochemical and pharmacokinetic properties, is a high-affinity inhibitor, traverses the blood–brain barrier, and has been

tested on healthy human volunteers without ill effect. This in-house discovery program should allow the development of powerful agents to restrict formation of toxic sphingolipids and address the lexicon of disordered ganglioside and globoside metabolism, including systemic (as in Fabry disease) and central effects for example, for use in neuronopathic Gaucher disease and the GM2 gangliosidoses.

Hematopoietic Stem-Cell Transplantation

Hematopoietic stem-cell therapy was an early potentially curative treatment in lysosomal diseases; it was first applied in the treatment of mucopolysaccharidoses MPS I and II and, successful in 1982, for the severe systemic manifestations of Norrbottnian (chronic neuronopathic) Gaucher disease.

Hematopoietic stem-cell transplantation provides a complement of hematopoietic cells from the donor that are competent in metabolizing lysosomal substrates: in the case of Gaucher disease, in which pathological macrophages are a major component, safe and successful engraftment has the potential for systemic cure. In the mucopolysaccharide disorders, microglia and other cells of myeloid lineage which are competent in metabolizing glycosaminoglycans and may secrete active lysosomal hydrolases of donor origin which ultimately find their way to peripheral tissues and viscera. These may slowly replace the microglial cell complement in the brain to improve the capacity to remodel tissues in which glycosaminoglycans have accumulated.

Clearly, a minority of patients are suitable recipients for hematopoietic stem-cell transplantation from either HLA matched-unrelated or HLA-matched related donors, but the therapeutic position of the more severe variant of MPSI (Hurler disease) is moderately well-established to provide benefit in the first year or so of life. It is salutary to appreciate that the stem-cell transplantation does not cure the condition and is unable to reverse established brain injury, and most of the crippling skeletal manifestations.

Hematopoietic stem-cell transplantation has been used in attempts to treat lysosomal diseases with neurological effects since stem cells from suitable donors have the potential to generate progeny that can migrate in small numbers over time to the brain by traversing the physical blood–brain barrier. In the recipient's bone marrow and other niches, donor cells give rise to self-renewing blood cell lineages that are competent in autonomous functions (such as phagocytosis and other immune responses); in the brain they appear to migrate to regions of disease where they can participate in scavenging activity as healthy counterparts of the diseased microglia and other cells with impaired lysosomal function. In the case of disorders such as Krabbe disease, the Sanfilippo syndromes (MPS III A–D), as well as metachromatic leukodystrophy, the donor cells may also secrete relevant acid hydrolases to complement neighboring tissue via the secretion-recapture process.

Reported outcomes of transplantation in more than 60 European patients with lysosomal diseases, especially in Gaucher disease, indicate that timely replacement of the diseased macrophage population and complementation of other hematopoietic lineages may greatly improve the disease with

effective correction – or arrest – of systemic disease. Contemporary experience shows improved outcomes as well as tolerance and safety of the procedure but for Gaucher disease, with the advent of three approved and safe enzyme infusion therapies, as well as orally active substrate biosynthesis inhibitors, hematopoietic stem-cell transplantation has restricted use – except in resource-poor countries where a one-off procedure may be preferred on the grounds of practicality and lifetime expenditure for treatment. At present, while hematopoietic stem-cell therapy offers the prospect of benefit in several disorders, particularly the mucopolysaccharidoses and possibly in severe chronic neuronopathic Gaucher disease (type 3), practical limitations restrict its use. Typically, an ideal HLA-matched sibling donor is not available, and despite the best efforts of agencies to identify matched unrelated donors, confirmation of compatibility may be delayed beyond utility; in any event, the success of the procedure mandates early and accurate diagnosis of the disease in the potential recipient – a persistent challenge in all therapeutic settings for rare disorders. Use of umbilical cord blood donors as a rich source of unrelated donor stem cells for suitable recipients, as well as optimized conditioning and myeloablative techniques have nonetheless gradually improved outcomes for several lysosomal diseases with systemic manifestations. However, the results obtained in patients with progressive neurodegeneration are at best suboptimal and often very disappointing.

Despite these improvements, persistent morbidity and mortality ($\approx 10\%$) is associated with transplantation – leading to complex decision-making that seeks to balance the severity of the disease with likely or optimal therapeutic response. With the exception of Hurler disease (MPS1) hematopoietic stem-cell transplantation has usually been deferred in favor of enzyme therapy whenever this is available. However while enzyme therapy may improve engraftment, hematopoietic stem-cell transplantation improves the high incidence of neutralizing allo-antibodies observed in MPS I after pharmacological enzyme replacement therapy. Hematopoietic stem-cell therapy has been extended to the lysosomal diseases characterized by rapidly progressive white-matter neurodegeneration such as metachromatic leukodystrophy and Krabbe disease. Research in this demanding field of clinical practice seeks to ensure that indications for the intervention are matched to the stage of disease and predicted therapeutic response.

Enzymatic Complementation

Molecular therapy for Gaucher disease

Gaucher disease was the first lysosomal disease for which a definitive molecular therapy was developed. After more than two decades of study at the National Institutes of Health under the direction of R.O. Brady, it was found that the functional deficiency of β -glucocerebrosidase in the macrophages of the spleen, liver, and other organs could be corrected experimentally by supplying the exogenous enzyme protein in a form that allowed molecular targeting to the macrophage via the mannose receptor (Deegan and Cox, 2012). Unlike the mannose 6-phosphate receptor, identified through

contemporaneous studies as a key participant in the biogenesis and intracellular delivery of many lysosomal hydrolases to the lysosomal matrix, the avidity of the macrophage for protein ligands with terminal mannose residues, rather than the phosphohexosyl moiety, is high. Thus the critical discovery by P.D. Stahl and colleagues of the mannose receptor, which mediates the avid uptake of mannose-terminated glycoproteins in human alveolar macrophages (Stahl *et al.*, 1978) underpinned a large international therapeutic program. Subsequent *in vivo* studies using radio-labeled placental and recombinant human glucosylcerebrosidase in patients receiving enzyme therapy for Gaucher disease confirmed specific uptake and targeting of these agents – enzyme protein half-life was < 5 min after infusion with rapid and avid uptake in macrophage-rich viscera and sites of disease in the bone marrow (Mistry *et al.*, 1996).

Commercial development of this targeted enzyme therapy and introduction of the CE 1983 Orphan Drugs Act in the United States was fortuitous: efficacy of the first placental-derived acid β -glucocerebrosidase remodeled to expose critical mannose residues for uptake by the macrophage system (alglucerase) occurred in CE 1992, and the remodeled recombinant protein, imiglucerase, as a biologic expressed in Chinese hamster ovary cells in CE 1994, in effect introduced the first 'blockbuster' orphan therapy. High costs of this therapy were justified by the excellent safety profile and salutary effects on key clinical manifestations of Gaucher disease.

The introduction and marketing of an effective biological treatment for Gaucher disease, based on biotechnological advances and a tri-partite arrangement between academia (The National Institutes of Health), investors, and an influential patient organization (National Gaucher Foundation) brokered by the emerging small biotech company Genzyme, was a spectacular affirmation of the value of the Orphan Drug Act. Within a few years, targeted molecular therapy for Gaucher disease placed the Genzyme Corporation in the forefront of technology and enabled it to diversify its activities in an orphan disease context and thereby develop treatments for other ultra-rare diseases.

The Genzyme Corporation was first-in-class for lysosomal diseases and was one of few biotechnology companies that were almost immediately successful in the orphan disease field. Space does not permit full discussion of the Genzyme Corporation and events induced by viral contamination of bioreactors for enzyme manufacture. Regulatory shutdown of the production facility in the core biological manufacturing businesses led to takeover by Sanofi Aventis – a French pharmaceutical giant. While it would be unusual for a major biopharmaceutical companies to continue its own experimental medicine program, hitherto there has been sustained development by Genzyme as a distinct business unit within Sanofi.

After reports of the effect of miglustat in Gaucher disease, Genzyme developed an in-house program in association with Drs. Radin and Shayman (University of Michigan) to develop a portfolio of biosynthetic inhibitors for clinical application in the sphingolipidoses. It is unusual for a biopharmaceutical company to develop self-competitive products; but in the mature era of enzyme therapy for Gaucher disease, with competing drugs from Shire and the Israeli company Protalix, adopted for this purpose by the Pfizer company, introduction

of orally active agents with an innovative mode of action presented an unusual opportunity.

Mucopolysaccharidoses

As discussed elsewhere, intensive biochemical studies of fibroblasts cultured from genetically distinct forms of the MPS by Elizabeth Neufeld, William Sly, and later Stuart Kornfeld and Kurt von Figura and their colleagues revealed much through studies of functional complementation about receptor-mediated uptake molecules harboring a common lysosomal recognition marker and serving as corrective factors that define specific clinical syndromes for enzymatic therapy. After many years of scientific research to determine the molecular events leading to an understanding of the sorting of proteins destined for delivery to the lysosomal compartment, it is gratifying to see that the MPS are well represented in the canon of lysosomal disorders for which specific therapy has been adopted by biopharmaceutical companies.

These disorders form a convenient group of conditions where there is impaired handling of complex glycosaminoglycans which are widely distributed molecular components in connective tissues including joints, heart valves, skeleton, and in vascular and extra-vascular spaces throughout the body. Each defect is reflected in the accumulation of mucopolysaccharide fragments, particularly dermatan, keratin, heparin, and chondroitin that are constituents of macromolecular proteoglycans. These complex molecules are principally responsible for maintaining tissue integrity.

Patients with mucopolysaccharidoses have bony abnormalities and joint stiffness. Those MPS diseases associated with accumulation of heparan sulfate may be particularly associated with brain disease: thickening of the leptomeninges, and the accumulation of excess cognate glycosaminoglycans in neural cells. Patients typically have coarse features and short stature, with abnormalities of dentition and distorted cartilage within the upper respiratory tract. The heart valves may be distorted by outgrowth of calcified fibrous tissue containing excess pathological glycosaminoglycans; neural cells show cytoplasmic vacuolation and abundant storage granules.

Recombinant human α -L-iduronidase is given intravenously for MPS1 (Hurler, Hurler-Scheie, and Scheie syndromes) after years of intensive laboratory study. Not only does the enzyme therapy diminish the excretion of excess glycosaminoglycans in the urine, but enlargement of the liver is decreased with improved rate of skeletal growth and range of joint movement. The treatment also improves the disease in the upper airways, with a helpful improvement in the episodes of hypoventilation occurring during sleep. In this condition, although many patients develop antibodies against the infused protein, usually only transient immune reactions occur and a degree of efficacy appears to be maintained.

Other mucopolysaccharidoses

Enzyme therapy has been introduced for several MPS syndromes, most notably MPS2 (Hunter disease). The activity of recombinant iduronate sulfatase in which, as with other sulfatases, enzymatic activity is dependent on the posttranslational modification of an active-site cysteine to formylglycine, decorated by mannose 6-phosphate residues to allow specific binding to the cognate receptors on the cell surface. The

commercial preparation, idursulfase, has principally been studied in individuals with the attenuated form of Hunter disease, showing improvement in mobility and respiratory difficulties. Since proteins do not materially traverse the blood-brain barrier, enzyme treatment does not improve the cognitive and related features of MPS2 affecting the brain, and since the 1980s, following poor results of hematopoietic stem-cell (marrow) transplantation, there has been a moratorium on this latter modality.

In MPS6 (Maroteaux-Lamy disease), recombinant *N*-acetyl-galactosamine 4-sulfatase (arylsulfatase B – BioMarin Pharmaceuticals) has been approved for clinical use and improves the systemic manifestations of this condition in which cognitive impairment is unusual. Intravenous infusions improve walking and stair climbing capacity in patients with MPS6.

MPS4a (Morquio disease)

BioMarin Pharmaceuticals have used recombinant mannose 6-phosphate-decorated human *N*-acetylgalactosamine-6-sulfatase to clear keratan sulfate from the lysosomal compartment and arrest progression of this disease, which has very striking effects with abnormal thoracic bone structure and vertebral development that impede ventilatory capacity and distorted limbs due to short stature and lax joints. Together with subluxation of the atlantoaxial joint, impair mobility and may be fatal. Randomized placebo-controlled Phase 3 trials have shown that infusion of 2 mg/kg⁻¹ of the recombinant protein on a weekly basis improves six-minute walk tests and other secondary outcome measures, together with the coherent improvements in urine excretion of keratan sulfate. As with all the sulfatase enzymes, active product requires activation of the critical cysteine through the agency of the SUMF as an additional hurdle in the biopharmaceutical manufacturing of advanced biologic therapies. A summary of approved enzyme therapies currently authorized as biologics is available (Baldo, 2015).

Therapeutic Stratagems for Lysosomal Diseases with Neurological Manifestations

Most lysosomal diseases are associated with neurological effects, some of which, as in the lysosomal diseases such as Tay-Sachs, with primary defects in the metabolism of sphingolipids, prove fatal in infancy. As the nervous system is formed, the neurones and supporting glia only rarely undergo mitosis and thus depend heavily on lysosomes to recycle many critical components. These include large tracts of membrane, synaptic vesicles, an extensive endoplasmic reticulum, and abundant mitochondria and peroxisomes, durably throughout the lifetime of the cell. By the same token, neurodegenerative diseases have severe effects on the individual and remain a challenge for medical science. While delivery of complementing factors in the form of lysosomal hydrolases has proved relatively straightforward in the treatment of systemic disease, therapeutic targeting of lysosomes that are essential for the normal functioning of the nervous system has proved difficult. With their signaling functions and rich expression in neural cells, disorders of lysosomal breakdown of sphingolipids has been a focus of intense

research offering considerable promise that, sooner or later, definitive treatments will become available.

Principal therapeutic stratagems:

- i. Use of substrate biosynthesis inhibitors (substrate-reduction – or rebalancing – therapy)
- ii. Direct enzymatic augmentation
- iii. Gene transfer.

Enzymatic Augmentation

Attempts to ameliorate neurological aspects of lysosomal diseases by enzyme therapy

Recombinant enzymes to restore the deficient activities in Sanfilippo diseases A and B have been explored with mixed results. MPS3 (A–D) have rather inconspicuous systemic effects, compared with the florid progressive cognitive impairment associated with agitation and behavioral abnormalities. Attempts are being made to determine whether intrathecal administration of sulfamidase will have any benefit in either preventing cognitive decline or possibly improving some aspects of the neurological phenotype of MPS1. However, despite the heroic nature of such trials, the difficulties of administering recombinant human proteins safely in highly mobile and behaviorally disturbed children with these conditions cannot be underestimated. A similar trial has been underway in MPS2, in which heparin sulfate also accumulates. Characteristically there is progressive cognitive impairment in most boys and men with the X-linked Hunter syndrome.

Enzyme therapies have also been in development for a very rare disorder: α -mannosidosis with a recombinant preparation following longstanding preclinical animal studies which showed benefit of this therapy. In this very rare glycoproteinosis with a birth frequency of approximately 1:1 000 000, undegraded polysaccharide intermediates resulting from the failure to digest mannose-containing glycoproteins lead to mild skeletal abnormalities and progressive impairment of mental function, including speech. While hematopoietic stem-cell (bone marrow) transplantation has been reported in some cases to arrest the progression of the disease, due to current infection and other complications, as well as the rarity of the disorder, render this evidence rather tentative and the results of recent enzyme replacement trials with recombinant α -mannosidase supported by a European project are awaited.

Niemann–Pick disease types A and B

Niemann–Pick Disease Types A and B are caused by deficiency of acid sphingomyelinase – a sphingolipid disorder associated with hepatosplenomegaly. In the severe Type A form, with markedly disabling mutations in ASM, there is accumulation of sphingomyelin in the brain with the buildup of secondary lipids, including ceramide and sphingosine. Although the molecular pathogenesis is not fully understood, in both forms – A, and the less severe peripheral B subtype, B – the presence of large foam cells, which are abnormal macrophages is associated with inflammatory cell infiltration in the parenchymal organs. There is marked enlargement of the liver and spleen with infiltration of the lung. Splenic sequestration of the formed elements of the blood with bleeding tendencies due to reduced circulating platelets with anemia and reduced

neutrophilic leukocytes, causing susceptibility to microbial infection also occur. For many years, enzyme therapy has been studied in preclinical models of Niemann–Pick diseases type A and B, and although the potential release of ceramide by recombinant acid sphingomyelinase targeted to cells via the mannose 6-phosphate and mannose receptors therapeutic dosing has been undertaken cautiously. Increasing doses of the putative therapeutic acid sphingomyelinase has been infused to determine whether the expected benefit can be achieved in the parenchymal organs without inducing catastrophic activation of apoptosis in the presence of excess ceramide. Since the administered enzyme does not penetrate the blood–brain barrier, it is unlikely that the progressive cognitive defects and neuro-developmental abnormalities that characterize the severe Niemann–Pick A variant will never be susceptible to correction by systemic enzyme therapy in whatever dose is ultimately chosen for therapeutic application.

Substrate-Reduction

There has been increasing interest in the development of inhibitors of sphingolipid biosynthesis that can be used to ameliorate the over-production of toxic sphingolipids in conditions when their recycling in the lysosome is impaired – thus restoring a normal equilibrium between biosynthesis and degradation of these bioactive macromolecules. The substrate-reducing stratagem, which has led to the introduction of two approved small molecules for the treatment of the systemic manifestations of Gaucher disease and miglustat for the neurological manifestations of Niemann–Pick disease type C, is discussed above.

Gene Transfer

The transfer of human cDNAs encoding the cognate wild-type sequence corresponding to the appropriate protein with impaired function as a result of mutation has been effective in preclinical animal models of lysosomal diseases and is undergoing testing formally in Phase I and Phase II in clinical trials. Lysosomal diseases have the potential for successful correction, since the secretion-recapture system forms the basis of enzyme augmentation therapy. Thus *in situ* complementation of the lysosomal deficit if successful will either ameliorate, arrest or even reverse the evolution of the disease (Cachón-González *et al.*, 2012). In most cases, where this stratagem has proved successful in living animals, viral vector systems have been utilized. In this case, the transduced cells act as a source of functioning enzyme that is released to effect the therapeutic complementation by uptake at a distance (Cachón-González *et al.*, 2012) so that theoretically, at least, the more enzyme generated and released, the greater will be the extent of correction through receptor-mediated uptake. As in humans, naturally occurring models of lysosomal diseases occur frequently in animals, and these serve as authentic experimental systems for therapeutic exploration, since they are almost invariably coherent biochemically and genetically. Emphasis here is placed on lysosomal diseases that affect the nervous system, principally because untreated neurodegenerative conditions have extreme human and medical costs – they

moreover represent challenges to establish widespread and effective therapeutic complementation that is sustained. Thus, the main challenges are generating sufficient delivery and distribution of the therapeutic product in all tissues at risk from injurious pathology, as exemplified by the brain, and to some extent the spinal cord. A further requirement is the capacity to sustain expression of the corrective gene, which can be safely transduced. Two principal methods have been adopted:

(i) Direct administration of vectors to the brain

Lysosomal diseases with neurodegenerative effects are in the vanguard of clinical studies of gene therapy. The success of interventions for inherited retinal diseases and hemophilia based on many years work with vectors that proved suitable for long-term expression and stratagems that avoid immune shutdown provides confidence that has hitherto been lacking. Moreover, 'risk-adverse' biotechnology companies are investing in academic developments with cautious support and sometimes collaborative investment in early clinical trials. Currently, the most promising vectors for this application are based on recombinant adeno-associated viral vectors, which are delivered by intra-cerebral injection using neurosurgical procedures to various brain structures. This is a passenger virus which does not integrate frequently into the genome. These episomal vectors have many advantages and appear to be safe when injected into white matter, for example in primates. These lysosomal enzymes have a 'housekeeping' property beyond the productive focus of expression and uptake via self-service receptors of the mannose 6-class may well overcome the barriers to supply corrective proteins distributed over an extensive target field. Recombinant AAV vectors containing limited expression and regulatory elements, as well as the cDNA encoding the lysosomal hydrolase of interest, are capable of inducing long-term sustained expression of the enzymes at activity levels predicted to have useful therapeutic effects (Cachón-González *et al.*, 2012).

A clinical Phase 1/2 study using the rAAV-2 serotype vector has been undertaken without ill effect in infantile neuronal lipofuscinosis Type 2 (NCL2) and was administered to several patients at several cranial entry sites expeditiously and without overt ill effect. These studies emboldened later trials using high-expression vectors in this and other fatal lysosomal diseases with neurodegeneration. Targets for this approach with recombinant AAV vectors, principally using the serotype rh.10, have been underway with Sanfilippo A disease (Tardieu *et al.*, 2014), Sanfilippo B disease, and, at the time of writing, initiated in metachromatic leukodystrophy. It is notable that all of these three trials are being undertaken in Paris, France, where there are several leading groups pioneering gene therapy for lysosomal diseases, principally using rAAV vector technology. Further studies are planned in GM1 gangliosidosis and also GM2 gangliosidosis (Tay-Sachs disease) and are at the immediate non-clinical developmental stage.

While gene therapy is still in its clinical infancy, it has nonetheless been delivered so far safely as a regulated and approved entity in several lysosomal diseases with neurological effects. However, design of future trials will be critical to determine whether or not efficacy can be demonstrated in patients stricken with such severe diseases. This poses the

question of whether the procedure can be justified in rapidly fatal diseases or those with indolent effects on the nervous system since heterogeneous manifestations in the latter pose difficulties for quantifying outcome.

Direct administration of vectors to the brain may prove to be too laborious as well as dangerous: recent research has shown that members of the AAV9 serotype vector family have potential penetrating properties toward the blood-brain barrier; this applies to adult animals in which the barrier robustly excludes soluble proteins and viruses. In some experimental systems it has been possible to transduce the nervous system via a single peripheral injection administered intravenously. The recent report of a high proportion of tumors in mice with GM2 gangliosidosis that received systemic rAAV9 in the neonatal period raises some concern about oncogenicity of this virus system. However, it seems likely that the long-term experience with successful human rAAV therapy in hemophilia B (Factor IX deficiency) will supply the necessary confidence to persist in this field of therapeutic research.

(ii) Indirect use of vector-transduced autologous hematopoietic stem-cells which are re-infused

This latter technique is built upon early indications of efficacy, based on hematopoietic stem-cell transplantation, first exploited with bone marrow donors but latterly donor stem-cells in peripheral blood elicited by the administration of both factors, such as GCSF. However, a typically suitable HLA-matched sibling or unrelated donor is not available, and the mortality of this procedure remains appreciable.

Ex-vivo hematopoietic stem-cell gene (reinfusion) therapy enhances the weak therapeutic effects observed after cellular complementation delivered to the peripheral circulation, where the peripheral manifestations of disease may be corrected, but there are only limited salutary effects on the brain depending on the degree to which donor mononuclear phagocyte precursors leave the circulation to populate the brain as resident macrophages and microglia.

Over three decades since the first clinical attempts at retroviral transfer of genes into hematopoietic progenitor cells, the concept of a therapeutic Trojan horse has been developed. This has been explored clinically by the group of Naldini and colleagues in the San Raffaele University in Milan in the severe white-matter disease Infantile Metachromatic Leukodystrophy (due to mutations in arylsulfatase A). Unusually for a lysosomal disorder, the genetically modified murine model of arylsulfatase deficiency appears unrepresentative of metachromatic leukodystrophy; unlike human patients, the mice show limited if any demyelination and do not suffer acute neurodegeneration. Microscopic abnormalities that are present in the brain of the animals bear little resemblance to the human disease in infancy; hypomyelination and demyelination are not prominent.

Biffi and colleagues recently reported a Phase 1/2 clinical trial in which transfer of wild-type human arylsulfatase A into autologous CD34⁺ hematopoietic progenitor cells was harvested from three unrelated infants who had been diagnosed early in the course of the disease at the pre-symptomatic phase, as a consequence of metachromatic leukodystrophy in their older siblings (Biffi *et al.*, 2013). Transduction of the CD34⁺ cells was stimulated *ex-vivo* with early-acting cytokines, brought about by exposure to purified third-generation lentiviral vectors

expressing human arylsulfatase A cDNA under the control of the phosphoglycerate kinase promoter. After expansion *ex-vivo*, the transduced progenitor cells were re-infused into their hosts, who had undergone substantial myeloablation without immune-suppression. Biffi *et al.* confirmed the stable engraftment of the transduced cells, no evidence of aberrant clonal behavior or preferential integration of the vectors. Arylsulfatase A was expressed in excess of the normal reference range in hematopoietic lineage. In an earlier report, evidence that this intervention had salutary effects is provided by the delayed onset and attenuated course of metachromatic leukodystrophy in recipients compared with their untreated siblings and studies of the natural course of the disease. While very encouraging, formal proof that, when given in the early pre-symptomatic phase of a lysosomal disease, this approach provides unequivocally superior outcomes than use of untransduced donor hematopoietic stem-cells, is not yet absolute and the therapeutic position of this ingenious stratagem is yet to be established.

Advances in the Field

Future Prospects in Clinical Research

Interdisciplinary exploration of the pathophysiology and biology of the lysosome has entered an accelerated phase: molecular therapies with salutary effects have allowed the aims of orphan drug legislation to be realized – and with several effective biologics approved, commercial returns have in notable cases catalyzed collaborative discovery in academic and biopharmaceutical environments. At a practical level, enhanced awareness of individual lysosomal disorders and the need for expertise in the management of multisystem disease, together with effective treatments leading to improved life-quality and survival to adulthood, attracts patients to academic centers. Thus investigators are increasingly drawn into a field enriched by genuine opportunities for translational bioscience with burgeoning clinical application: these include gene transfer and cell-based therapies which exploit the lysosomal network as a route for functional complementation.

Gaucher Disease – Insights Emerging from Exploring Ultra-Rare Lysosomal Disorders

Commercial and scientific opportunities and the humanitarian need for orally active agents with the potential to improve the convenience of treatment for Gaucher patients internationally, and ultimately address disabling neurological manifestations, are attractive for scientists and physicians alike. Furthermore, continued study of Gaucher disease has revealed important and, at present, poorly understood comorbidities; these include Parkinsonism, arterio-vascular and cardiac disease, as well as cancers – preferentially myeloma and B-cell lymphoma. Thus development of inhibitors of glycosphingolipid biosynthesis with therapeutic potential and which enter the nervous system promises to open up fundamental aspects of sphingolipid metabolism to another level of understanding. Moreover, the ability of small molecules rather than the first effective cell-targeted biologics to access tissues beyond the macrophage compartment is expected to shed light on

carcinogenesis, particularly in B cells. In this way, a single lysosomal disease continues to serve as a discovery platform and an example of unexpected insights gained from investigation in an ultra-rare disease.

Manipulating Macromolecular Biosynthesis and Substrate Loading in the Lysosome

One avenue for advancement is the fusion of combinatorial biosyntheses with identification of selective therapeutic inhibitors that allow rebalancing of metabolic fluxes in the lysosome. The utility of this stratagem has been validated in the glycosphingolipidoses, such as Gaucher disease, in which accumulation of poorly recycled macromolecules initiates cascades in the cell which lead to the partially compensated pathological complex of clinical disease. The ultimate reward of this approach will be judged by the ability favorably to attenuate lysosomal diseases that affect the brain. Several promising ‘small-molecule’ inhibitors that traverse the blood–brain barrier are being examined in the severely disabling human sphingolipid diseases, including Fabry disease and the gangliosidoses.

Hitherto, investigators have proved less successful in identifying suitable inhibitors of other rate-limiting steps in the biosynthesis or intracellular delivery of macromolecular substrates such as glycoproteins or glycosaminoglycans to the lysosome for degradation but intensified screening of compounds against refined biochemical reaction targets in this field is likely to be rewarding. To address the specific consequences of pathological storage, further exploration of molecular chaperones, including heat-shock protein-inducers to enhance exocytosis of lysosomal contents is clearly warranted; and while manipulation of the transcriptional control of the master-regulator TFEB (reciprocally to interrogate lysosomal biogenesis and autophagy) may be ambitious because of the theoretical possibility of cell toxicity or uncontrolled nutrient signaling and growth, this would also represent a fundamental molecular target for therapeutic exploration.

Pathogenesis and Physiological Understanding

Discovery of the lysosome (and peroxisome) was essentially coincident with the inception of Cell Biology as a credible domain of experimental science. It is remarkable how many advances in understanding the physiological rôle of the lysosome have been elucidated from investigating inborn errors causing human diseases which affect the biogenesis and function of the organelle. With more than 70 distinct individual human phenotypes in the lexicon of hereditary lysosomal diseases and disorders of lysosome-related organelles and biogenesis, together with whole-exome sequencing and bioinformatic analysis for improved molecular diagnosis, the stage is set for a knowledge explosion. Principally this will be based on the critical insights obtained from the detailed characterization of human diseases. Widespread incursion of genome-wide studies into the study of pathogenesis has identified variation in human phenotypes; those that affect the function of genes encoding proteins implicated in lysosome biology are numerous. Thus while the central rôle of the

lysosomal compartment in the inner workings of the cell is no longer in doubt, the importance of the organelle in general pathology has yet to be fully realized.

Conclusion

The pathological sciences and cell biology have coevolved and will continue to advance hand in hand through common technologies. Ready availability of interventional genetics and powerful microscopes that combine immunological and confocal fluorescence-based methods allow the cell to be explored dynamically. Through this we are able to identify the molecular interactions of neo-prismatic structures identified in living cells. Thus a discovery-horizon founded on an old scientific tradition has been generated: the study of informative mutant phenotypes paradigmatically represented by lysosomal diseases.

Acknowledgments

The author is indebted to Elizabeth Morris and Begoña Cachón-González for outstanding contributions to clinical and experimental work in this field; Elizabeth M. Chatters has kindly provided sustained professional assistance.

The author's research is principally supported by grants from the UK Medical Research Council (MR/K025570 and MR/K015338/1) and the National Institutes of Health Research, Cambridge Biomedical Research Centre (Metabolism theme).

References

- Aderem, A., Underhill, D.M., 1999. Mechanisms of phagocytosis in macrophages. *Annual Review of Immunology* 17, 593–623.
- Ashford, T.P., Porter, K.R., 1962. Cytoplasmic components in hepatic cell lysosomes. *Journal of Cell Biology* 12, 198–202.
- Baldo, B.A., 2015. Enzymes approved for human therapy: Indications, mechanisms and adverse effects. *BioDrugs* 29, 31–55.
- Bar-Peled, L., Sabatini, D.M., 2014. Regulation of mTORC1 by amino acids. *Trends in Cell Biology* 24, 400–406.
- Barton, D.J., Ludman, M.D., Benkov, K., Grabowski, G.A., LeLeiko, N.S., 1989. Resting energy expenditure in Gaucher's disease type 1: Effect of Gaucher's cell burden on energy requirements. *Metabolism* 38, 1238–1243.
- Baudhuin, P., Hers, H.G., Loeb, H., 1964. An electron microscopic and biochemical study of type II glycogenosis. *Laboratory Investigation* 13, 1139–1152.
- Baugh, M., Black, D., Westwood, P., et al., 2011. Therapeutic dosing of an orally active, selective cathepsin S inhibitor suppresses disease in models of autoimmunity. *Journal of Autoimmunity* 36, 201–209.
- Beavan, M.S., Schapira, A.H., 2013. Glucocerebrosidase mutations and the pathogenesis of Parkinson disease. *Annals of Medicine* 45, 511–521.
- Berkovic, S.F., Dibbens, L.M., Oshlack, A., et al., 2008. Array-based gene discovery with three unrelated subjects shows SCARB2/LIMP-2 deficiency causes myoclonus epilepsy and glomerulosclerosis. *American Journal of Human Genetics* 82, 673–684.
- Biffi, A., Montini, E., Lorioli, L., et al., 2013. Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science* 341 (6148), 1233158.
- Briggs, T.A., Rice, G.I., Daly, S., et al., 2011. Tartrate-resistant acid phosphatase deficiency causes a bone dysplasia with autoimmunity and a type I interferon expression signature. *Nature Genetics* 43, 127–131.
- Bune, A.J., Hayman, A.R., Evans, M.J., Cox, T.M., 2001. Mice lacking tartrate-resistant acid phosphatase (Acp 5) have disordered macrophage inflammatory responses and reduced clearance of the pathogen, *Staphylococcus aureus*. *Immunology* 102, 103–113.
- Cachón-González, M.B., Wang, S.Z., McNair, R., et al., 2012. Gene transfer corrects acute GM2 gangliosidosis — Potential therapeutic contribution of perivascular enzyme flow. *Molecular Therapy* 20, 1489–1500.
- Caciotti, A., Catarzi, S., Tonin, R., et al., 2013. Galactosialidosis: Review and analysis of CTSA gene mutations. *Orphanet Journal of Rare Diseases* 8, 114. doi:10.1186/1750-1172-8-114.
- Coelho, D., Kim, J.C., Miousse, I.R., et al., 2012. Mutations in ABCD4 cause a new inborn error of vitamin B12 metabolism. *Nature Genetics* 44, 1152–1155.
- Cori, G.T., 1952. Glycogen Structure and Enzyme Deficiencies in Glycogen Storage Disease. The Harvey Lectures, 1952–1953, 48, 146 (Academic Press Inc., New York).
- Cox, T.M., 2015. Innovative treatments for lysosomal diseases. *Best Practice & Research Clinical Endocrinology & Metabolism* 29, 275–311.
- Cox, T.M., Cachón-González, M.B., 2012. The cellular pathology of lysosomal diseases. *Journal of Pathology* 226, 241–254.
- Cox, T.M., Schofield, J.P., 1997. Gaucher's disease: Clinical features and natural history. *Baillieres Clinical Haematology* 10, 657–689.
- Danes, B.S., Bearn, A.G., 1965. Hurler's syndrome: Demonstration of an inherited disorder of connective tissue in cell culture. *Science* 149 (3687), 987–989.
- Daniele, T., Hurbain, I., Vago, R., et al., 2014. Mitochondria and melanosomes establish physical contacts modulated by Mfn2 and involved in organelle biogenesis. *Current Biology* 24, 393–403.
- De Duve, C., 1964. From cytochromes to lysosomes. *Fed Proc* 23, 1045–1049.
- De Duve, C., Wattiaux, R., 1966. Functions of lysosomes. *Annual Review of Physiology* 28, 435–492.
- Deegan, P.B., Cox, T.M., 2012. Imiglucerase in the treatment of Gaucher disease: A history and perspective. *Journal of Drug Design, Development and Therapy* 6, 81–106.
- Devos, P., Hers, H.G., 1980. Random, presumably hydrolytic, and lysosomal glycogenolysis in the livers of rats treated with phlorizin and of newborn rats. *Biochemical Journal* 192, 177–181.
- Eskelinen, E.-L., 2005. Maturation of autophagic vacuoles in mammalian cells. *Autophagy* 1, 1–10.
- von Figura, K., Hasilik, A., 1986. Lysosomal enzymes and their receptors. *Annual Review of Biochemistry* 55, 167–193.
- Fratantoni, J.C., Hall, C.W., Neufeld, E.F., 1968. Hurler and Hunter syndromes: Mutual correction of the defect in cultured fibroblasts. *Science* 162 (3853), 570–572.
- Fratantoni, J.C., Hall, C.W., Neufeld, E.F., 1969. The defect in Hurler and Hunter syndromes. II. Deficiency of specific factors involved in mucopolysaccharide degradation. *Proceedings of the National Academy of Sciences (USA)* 64, 360–366.
- Frattoni, A., Orchard, P.J., Sobacchi, C., et al., 2000. Defects in TCIRG1 subunit of the vacuolar proton pump Defects in TCIRG1 subunit of the vacuolar proton pump are responsible for a subset of human autosomal recessive osteopetrosis. *Nature Genetics* 25, 343–346.
- Futerman, A.H., van Meer, G., 2004. The cell biology of lysosomal storage disorders. *Nature Reviews Molecular Cell Biology* 5, 554–565.
- Gahl, W.A., 2012. The battlefield of rare diseases: Where uncommon insights are common. *Science Translational Medicine* 4 (154), 154ed7.
- Gelb, B.D., Shi, G.-P., Chapman, H.A., Desnick, R.J., 1996. Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency. *Science* 273, 1236–1238.
- Geuze, H.J., Slot, J.W., Strous, G.J., Hasilik, A., von Figura, K., 1985. Possible pathways for lysosomal enzyme delivery. *Journal of Cell Biology* 101, 2253–2262.
- Goker-Alpan, O., Schiffmann, R., LaMarca, M.E., et al., 2004. Parkinsonism among Gaucher disease carriers. *Journal of Medical Genetics* 41, 937–940.
- Goker-Alpan, O., Stubblefield, B.K., Giasson, B.I., Sidransky, E., 2010. Glucocerebrosidase is present in α -synuclein inclusions in Lewy body disorders. *Acta Neuropathologica* 120, 641–649.
- Grabowski, G.A., 2008. Phenotype, diagnosis, and treatment of Gaucher's disease. *Lancet* 372 (9645), 1263–1271.
- Haltia, M., Elleder, M., Goebel, H.H., Lake, B.D., Mole, S.E., 2011. The NCLs: Evolution of the concept and classification. In: Mole, S.E., Williams, R.E., Goebel, H.H. (Eds.), *The Neuronal Ceroid Lipofuscinoses*, second ed. New York: Oxford University Press Inc., pp. 1–19.
- Hers, H.G., 1963. α -Glucosidase deficiency in generalized glycogen storage disease (Pompe's disease). *Biochemical Journal* 86, 11–16.
- Hers, H.G., 1965. Inborn lysosomal diseases. *Gastroenterology* 48, 625–633.
- Hers, H.G., 1983. From fructose to fructose-2, 6-bisphosphate with a detour through lysosomes and glycogen. In: Semenza, G. (Ed.), *Comprehensive Biochemistry* 35. Amsterdam: Elsevier, pp. 71–101.

- Hickman, S., Neufeld, E.F., 1972. A hypothesis for I-cell disease: Defective hydrolases that do not enter lysosomes. *Biochemical and Biophysical Research Communication* 49, 992–999.
- van den Hout, H.M., Hop, W., van Diggelen, O.P., *et al.*, 2003. The natural course of infantile Pompe's diseases: 20 Original cases compared with 133 cases from the literature. *Pediatrics* 112, 332–340.
- Huizing, M., Helip-Wooley, A., Westbroek, W., Gunay-Aygun, M., Gahl, W.A., 2008. Disorders of lysosome-related organelle biogenesis: Clinical and molecular genetics. *Annual Review of Genomics and Human Genetics* 9, 359–386.
- Introne, W.J., Westbroek, W., Golas, G.A., Adams, D., 2015. Chediak–Higashi Syndrome. In: Pagon, R.A. *et al.* (Eds.), *GeneReviews*® [Internet]. Seattle, WA: University of Washington, pp. 1993–2015. 2009 Mar 3 [updated].
- Kaback, M.M., Desnick, R.J., 2011. Hexosaminidase A deficiency. In: Pagon, R.A., Adam, M.P., Ardinger, H.H., *et al.* (Eds.), *GeneReviews*® [Internet]. Seattle, WA: University of Washington, pp. 1993–2015. 1999 Mar 11 [updated 2011 Aug 11].
- Kishnani, P.S., Beckenmeyer, A.A., 2014. New therapeutic approaches for Pompe disease: Enzyme therapy and beyond. *Pediatric Endocrinology Reviews* 12 (Suppl 1), 114–124.
- Klein, A.D., Futerman, A.H., 2013. Lysosomal storage disorders: Old diseases, present and future challenges. *Pediatric Endocrinology Reviews* 11 (Suppl 1), 59–63.
- Klionsky, D.J., 2007. Autophagy: From phenomenology to molecular understanding in less than a decade. *Nature Reviews Molecular Cell Biology* 8, 931–937.
- Kolter, T., Sandhoff, K., 2005. Principles of lysosomal membrane digestion: Stimulation of sphingolipid degradation by sphingolipid activator proteins and anionic lysosomal lipids. *Annual Review of Cell and Developmental Biology* 21, 81–103.
- Kornak, U., Kasper, D., Bo, M.R., *et al.*, 2001. Loss of the CIC-7 chloride channel leads to osteopetrosis in mice and man. *Cell* 104, 205–215.
- Kornfeld, S., 1987. Trafficking of lysosomal enzymes. *FASEB Journal* 1, 462–468.
- Kornfeld, S., Mellman, I., 1989. The biogenesis of lysosomes. *Annual Review of Cell Biology* 5, 483–525.
- Krzewski, K., Coligan, J.E., 2012. Human NK cell lytic granules and regulation of their exocytosis. *Frontiers in Immunology* 3, 335. 2015 Jan 15. doi:10.3389/fimmu.2012.00335.
- Laplante, M., Sabatini, D.M., 2013. mTOR signaling at a glance. *Journal of Cell Science* 122, 3589–3594.
- Lejeune, N., Thines-Sempoux, D., Hers, H.G., 1963. Tissue fractionation studies. 16. Intracellular distribution and properties of α -glucosidases in rat liver. *Biochemical Journal* 86, 16–21.
- Leroy, J.G., Cathey, S., Friez, M.J., 2012. Mucopolidosis II. In: Pagon, R.A., Adam, M.P., Ardinger, H.H., *et al.* (Eds.), *GeneReviews*® [Internet]. Seattle, WA: University of Washington, pp. 1993–2015.
- Leroy, J.G., DeMars, R.I., 1967. Mutant enzymatic and cytological phenotypes in cultured human fibroblasts. *Science* 157, 804–806.
- Lieberman, A.P., Puertollano, R., Raben, N., *et al.*, 2012. Autophagy in lysosomal storage disorders. *Autophagy* 8, 719–730.
- Luzio, J.P., Hackmann, Y., Dieckmann, N.M., Griffiths, G.M., 2014. The biogenesis of lysosomes and lysosome-related organelles. *Cold Spring Harbor Perspectives in Biology* 6 (9), a016840. Sep 2.
- Makrypidi, G., Damme, M., Müller-Loennies, S., *et al.*, 2012. Mannose 6 dephosphorylation of lysosomal proteins mediated by acid phosphatases Acp2 and Acp5. *Molecular and Cellular Biology* 32, 774–782.
- Malicdan, M.C., Nishino, I., 2012. Autophagy in lysosomal myopathies. *Brain Pathology* 22, 82–88.
- Mancini, G.M., Havelaar, A.C., Verheijen, F.W., 2000. Lysosomal transport disorders. *Journal of Inherited Metabolic Disease* 23, 278–292.
- Mechtler, T.P., Stary, S., Metz, T.F., *et al.*, 2012. Neonatal screening for lysosomal storage disorders: Feasibility and incidence from a nationwide study in Austria. *Lancet* 379 (9813), 335–341.
- Medina, D.L., Di Paola, S., Peluso, I., *et al.*, 2015. Lysosomal calcium signalling regulates autophagy through calcineurin and TFEB. *Nature Cell Biology* 17, 288–299.
- Medina, D.L., Fraldi, A., Bouche, V., *et al.*, 2011. Transcriptional activation of lysosomal exocytosis promotes cellular clearance. *Developmental Cell* 21, 421–430.
- Metchnikoff, E., 1884. Ueber eine Sporospizkrankheit der Daphnien. *Virchows Archiv* 96, 177–195.
- Mindell, J.A., 2012. Lysosomal acidification mechanisms. *Annual Review of Physiology* 74, 69–86.
- Mistry, P.K., Taddei, T., vom Dahl, S., Rosenbloom, B.E., 2013. Gaucher disease and malignancy: A model for cancer pathogenesis in an inborn error of metabolism. *Critical reviews in oncogenesis* 18, 235–246.
- Mistry, P.K., Wraight, E.P., Cox, T.M., 1996. Therapeutic delivery of proteins to macrophages: Implications for treatment of Gaucher's disease. *Lancet* 348 (9041), 1555–1559.
- Morgan, A.J., Platt, F.M., Lloyd-Evans, E., Galione, A., 2011. Molecular mechanisms of endolysosomal Ca^{2+} signalling in health and disease. *Biochemical Journal* 439, 349–374.
- Müller, M., Schmidt, O., Angelova, M., *et al.*, 2015. The coordinated action of the MVB pathway and autophagy ensures cell survival during starvation. *Elife* 22, 4. 2015 Apr.
- Natowicz, M.R., Chi, M.M.-Y., Lowry, O.H., Sly, W.S., 1979. Enzymatic identification of mannose 6-phosphate on the recognition marker for receptor-mediated pinocytosis of β -glucuronidase by human fibroblasts. *Proceedings of the National Academy of Sciences (USA)* 76, 4322–4326.
- Nesterova, G., Gahl, W.A., 2013. Cystinosis: The evolution of a treatable disease. *Pediatric Nephrology* 28, 51–59.
- Neufeld, E.F., 2011. From serendipity to therapy. *Annual Review of Biochemistry* 80, 1–15.
- Palmieri, M., Impey, S., Kang, H., *et al.*, 2011. Characterization of the CLEAR network reveals an integrated control of cellular clearance pathways. *Human Molecular Genetics* 20, 3852–3866.
- Piper, R.C., Luzio, J.P., 2001. Late endosomes: Sorting and partitioning in multivesicular Bodies. *Traffic* 2, 612–621.
- Platt, F.M., Bolland, B., van der Spoel, A.C., 2012. The cell biology of disease: Lysosomal storage disorders: The cellular impact of lysosomal dysfunction. *Journal of Cell Biology* 199, 723–734.
- Popovic, D., Akutsu, M., Novak, I., *et al.*, 2012. Rab GTPase-activating proteins in autophagy: Regulation of endocytic and autophagy pathways by direct binding to human ATG8 modifiers. *Molecular and Cellular Biology* 32, 1733–1744.
- Poupetová, H., Ledvinová, J., Berná, L., *et al.*, 2010. The birth prevalence of lysosomal storage disorders in the Czech Republic: Comparison with data in different populations. *Journal of Inherited Metabolic Disease* 33, 387–396.
- Proud, C.G., 2006. Regulation of protein synthesis by insulin. *Biochemical Society Transactions* 34, 213–216.
- Ravikumar, B., Sarkar, S., Davies, J.E., *et al.*, 2010. Regulation of mammalian autophagy in physiology and pathophysiology. *Physiological Reviews* 90, 1383–1435.
- Rebsamen, M., Pochini, L., Stasyk, T., *et al.*, 2015. SLC38A9 is a component of the lysosomal amino acid sensing machinery that controls mTORC1. *Nature* 519, 477–481.
- Reczek, D., Schwake, M., Schröder, J., *et al.*, 2007. LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase. *Cell* 131, 770–783.
- Saftig, P., Klumperman, J., 2009. Lysosome biogenesis and lysosomal membrane proteins: Trafficking meets function. *Nature Reviews Molecular Cell Biology* 10, 623–635.
- Sardi, S.P., Cheng, S.H., Shihabuddin, L.S., 2015. Gaucher-related synucleinopathies: The examination of sporadic neurodegeneration from a rare (disease) angle. *Progress in Neurobiology* 125, 47–62.
- Sardiello, M., Annunziata, I., Roma, G., Ballabio, A., 2005. Sulfatases and sulfatase modifying factors: An exclusive and promiscuous relationship. *Human Molecular Genetics* 14, 3203–3217.
- Sardiello, M., Palmieri, M., di Ronza, A., *et al.*, 2009. A gene network regulating lysosomal biogenesis and function. *Science* 325 (5939), 473–477.
- Sengupta, S., Peterson, T.R., Sabatini, D.M., 2010. Regulation of the mTOR complex 1 pathway by nutrients, growth factors, and stress. *Molecular Cell* 40, 310–322.
- Settembre, C., Di Malta, C., Polito, V.A., *et al.*, 2011. TFEB Links Autophagy to Lysosomal Biogenesis. *Science* 332 (6036), 1429–1433.
- Settembre, C., Fraldi, A., Medina, D.L., Ballabio, A., 2013. Signals from the lysosome: A control centre for cellular clearance and energy metabolism. *Nature Reviews Molecular Cell Biology* 14, 283–296.
- Settembre, C., Zoncu, R., Medina, D.L., *et al.*, 2012. A lysosome-to-nucleus signalling mechanism senses and regulates the lysosome via mTOR and TFEB. *EMBO Journal* 31, 1095–1108.
- Seward Jr., S.L., Gahl, W.A., 2013. Hermansky–Pudlak syndrome: Health care throughout life. *Pediatrics* 132, 153–160.
- Shapira, A.H., 2015. Glucocerebrosidase and Parkinson disease: Recent advances. *Molecular and Cellular Neuroscience* 66 (Pt A), 37–42.
- Stahl, P.D., Rodman, J.S., Miller, M.J., Schlesinger, P.H., 1978. Evidence for receptor-mediated binding of glycoproteins, glycoconjugates, and lysosomal glycosidases by alveolar macrophages. *Proceedings of the National Academy of Sciences (USA)* 75, 1399–1403.

- Tardieu, M., Zerah, M., Husson, B., *et al.*, 2014. Intracerebral administration of adeno-associated viral vector serotype rh.10 carrying human SGSH and SUMF1 cDNAs in children with mucopolysaccharidosis type IIIA disease: Results of a phase I/II trial. *Human Gene Therapy* 25, 506–516.
- Tay, W., 1881. Symmetrical changes in the region of the yellow-spot in each eye in an infant. *Transactions of the Ophthalmological Society* 1, 55–57.
- Terry, R.D., Korey, S.R., 1960. Membranous cytoplasmic granules in infantile amaurotic idiocy. *Nature* 188, 1000–1002.
- Van Hoof, F., Hers, H.G., 1968. The abnormalities of lysosomal enzymes in mucopolysaccharidoses. *European Journal of Biochemistry* 7, 34–44.
- Weinreb, N.J., Lee, R.E., 2013. Causes of death due to haematological and non-hematological malignancies in 57 US patients with type 1 Gaucher disease who were never treated with enzyme replacement therapy. *Critical Reviews in Oncogenesis* 18, 177–195.
- Xu, H., Ren, D., 2015. Lysosomal physiology. *Annual Review of Physiology* 77, 57–80.
- Yang, Z., Klionsky, D.J., 2010. Eaten alive: A history of macroautophagy. *Nature Cell Biology* 12, 814–822.
- Yu, L., McPhee, C.K., Zheng, L., *et al.*, 2010. Termination of autophagy and reformation of lysosomes regulated by mTOR. *Nature* 465, 942–946.
- Zoncu, R., Bar-Peled, L., Efeyan, A., *et al.*, 2011. mTORC1 senses lysosomal amino acids through an inside-out mechanism that requires the vacuolar H(+) -ATPase. *Science* 334 (6056), 678–683.

Alpha-1-Antitrypsin Deficiency: A Misfolded Secretory Glycoprotein Damages the Liver by Proteotoxicity and Its Reduced Secretion Predisposes to Emphysematous Lung Disease Because of Protease-Inhibitor Imbalance

S Chakraborty, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

DH Perlmutter University of Pittsburgh School of Medicine, Pittsburgh, PA, USA and Children's Hospital of Pittsburgh of UPMC, Pittsburgh, PA, USA

Amitava Mukherjee, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Autolysosome An acidic degradative hybrid organelle formed by a fusion event between a lysosome and an autophagosome.

Autophagosomes Double membrane bound vesicles that encapsulate cytoplasmic components which eventually fuse with lysosomes.

Chronic obstructive pulmonary disease (COPD) A type of obstructive lung disease characterized by airflow limitation that is not fully reversible and encompasses conditions often described with the terms emphysema and chronic bronchitis. The airflow limitation is typically progressive and is associated with an abnormal inflammatory response of the lungs to noxious particle or gases primarily caused by cigarette smoking and breathing polluted air.

Cirrhosis A medical condition associated with severe liver damage and is characterized by conversion of normal liver tissue by fibrosis and structurally abnormal regenerative nodules. These changes can often lead to loss of liver function.

Cytostatic Inhibition of cell growth.

Dysplasia A term used in pathology to describe the enlargement of an organ or tissue caused by deregulated proliferation of cells of an abnormal type.

Emphysema A term used in pathology which refers to the destruction of the gas-exchanging surfaces of the lung.

Endoplasmic reticulum-associated protein degradation (ERAD) A cellular protein quality control pathway which targets misfolded proteins from the endoplasmic reticulum to be degraded by the proteasome. The proteins targeted for degradation often undergo ubiquitination.

Fibrosis The formation or deposition of excess fibrous connective tissue (scar tissue) in an organ.

Hepatectomy Surgical excision of the liver.

Hepatocellular carcinoma Malignant tumor arising from hepatocytes as the primary source.

Hepatocyte An epithelial parenchymatous cell of the liver that synthesizes and secretes bile. Hepatocytes constitute ~70–85% of the cytoplasmic mass of liver.

High throughput Rapid automated analysis of various substances using robotics technology.

Huntington's disease A specific type of neurodegenerative disorder that affects muscle coordination and leads to progressive decline of cognitive abilities and results in neuropsychiatric manifestations.

Induced pluripotent stem cells (iPS) Genetically reprogrammed adult cells that mimic several properties of an embryonic stem cell. The reprogramming is performed by expressing specific genes and factors critical for maintaining the defining properties of embryonic stem cells.

Metalloprotease A protease which requires a metal for its catalytic function.

Peptidomimetic A small protein-like chain designed to mimic a biologically active peptide with precisely altered molecular structure to introduce additional properties such as enhanced stability or biological activity.

Protease Any enzyme that cleaves a protein by hydrolyzing the peptide bonds that link amino acids together in the polypeptide chain forming the protein.

Proteostasis Proteostasis or protein homeostasis is a complex regulatory process that maintains proper repertoire of proteins within the cell which is critical for the health of the cellular proteome as well as the organism itself. This involves complex interplay of integrated network of subcellular pathways that influence the fate of a protein from synthesis to degradation. Dysfunction of any of the proteostatic pathways could lead to various cellular disorders and eventually diseases.

Proteotoxicity Impairment of cell function due to toxicity caused by accumulation of misfolded proteins.

Steatosis Abnormal retention of lipids in liver cells.

Introduction

The classical form of α 1-ATD is an autosomal co-dominant disease affecting 1 in 3000 live births. It is characterized by a point mutation that renders a major hepatic glycoprotein prone to misfolding in the early compartments of the secretory pathway. Liver disease, including fibrosis and carcinoma, is

caused by the accumulation/proteotoxicity of mutant α 1-antitrypsin Z (ATZ) in liver cells. Chronic obstructive pulmonary disease (COPD) is caused by excess protease activity that impacts the lung when levels of α 1-antitrypsin (AT) in the blood and body fluids are decreased (Figure 1).

Genetic and environmental modifiers have a major effect on the incidence and severity of lung and liver disease among

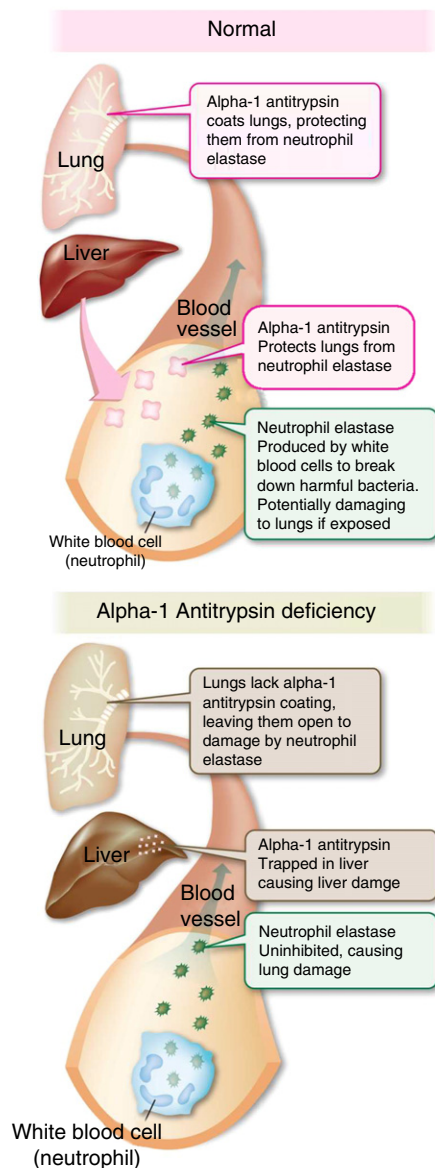


Figure 1 Pathophysiology of α 1-antitrypsin and its deficiency (ATD). Synthesis of AT in the liver and secretion into the circulation to protect the lung from neutrophil proteases in a normal individual is shown in the top panel. The classical form of ATD is depicted in the bottom panel with liver damage from accumulation of the mutant protein (gain-of-function) and lung damage due to protease excess (loss-of-function) shown in the bottom panel (image used with permission from The University of Utah Genetic Science Learning Center, <http://Learn.Genetics.utah.edu>).

ATD homozygotes. Multiple mechanisms have been implicated in the increased risk of COPD in affected ATD patients mediated by cigarette smoking, including active oxygen intermediates that abrogate proteinase inhibitory function of residual ATZ molecules. Variation in liver disease severity is thought to involve proteostasis mechanisms that mitigate the accumulation/proteotoxicity of the misfolded ATZ protein. Recent work has focused on the intracellular mechanisms by which the misfolded protein is degraded and specific adaptive

signaling pathways that are activated presumably to protect the cell (Ong and Kelly, 2011). This work has led to several novel therapeutic strategies which will be highlighted in this article.

Physiology and Function of AT

AT is a single-chain polypeptide with 394 amino acids and 3 asparagine-linked carbohydrate side-chains. It is encoded by a 12.2 Kb gene (SERPIN A1) located on human chromosome 14q31-32.2 (Long *et al.*, 1984). AT is synthesized as a 52-kDa partially glycosylated precursor and undergoes carbohydrate side-chain processing as it traverses the luminal secretory pathway from endoplasmic reticulum to Golgi and into the extracellular fluid with a half-time for secretion of 35–40 min (Lodish and Kong, 1987). The predominant site of synthesis of plasma AT is the liver, as shown by conversion of plasma AT to the donor protein isoform structure after orthotopic liver transplantation (Hood *et al.*, 1980). However, there is also evidence that AT is synthesized and secreted by extrahepatic cell types including monocytes and macrophages (Perlmutter *et al.*, 1985) and epithelial cells from a number of tissues (Koopman *et al.*, 1989; Carlson *et al.*, 1988; Perlmutter, 2011b).

Serum levels of AT increase by three- to five-fold during the host response to inflammation/tissue injury and so it is considered a typical 'acute phase reactant.' AT levels also increase during pregnancy and oral contraceptive therapy (reviewed in Teckman *et al.*, 1996). These changes in AT levels are believed to reflect regulation of gene expression at the transcriptional level by inflammatory cytokines, including IL-6, and by androgens (Perlmutter *et al.*, 1989). Expression of AT is also regulated by a feed-forward mechanism mediated through a receptor for AT-elastase complexes (Perlmutter *et al.*, 1990). Circulating AT diffuses into most tissues and body fluids at a concentration generally equivalent to that in plasma. It is cleared from the circulation with a half-time of ~ 5 days (Crystal, 1990). The daily production rate is estimated to be 34 mg kg^{-1} with 33% of the intravascular pool cleared and degraded daily.

The major physiological function of AT is inhibition of neutrophil elastase and several other serine proteases produced by neutrophils. It is the archetype of a family of structurally related proteins that are called SERPINS because most of the members are inhibitors of serine proteases. Although AT can inhibit multiple serine proteases *in vitro*, it is much more effective in inhibiting the neutrophil elastase, proteinase 3 and cathepsin G. Indeed the kinetics of its inhibitory activity for these neutrophil proteases are more favorable by several orders of magnitude and therefore these are believed to be its only physiological targets *in vivo* (Crystal, 1990). The protease inhibitory activity of AT may be modified by numerous factors but the effects of active oxygen intermediates and metalloproteases have been most well characterized. Active oxygen intermediate abrogate the elastase inhibitory properties of AT by oxidizing methionine at its reactive center (Hubbard *et al.*, 1987). Because cigarette smoking induces release of active oxygen intermediates by mononuclear phagocytes in the lung, this effect is believed to explain the increase in incidence and severity of destructive lung disease in ATD individuals that are

also smokers and is the essence of the 'protease-inhibitor imbalance,' or 'elastase-antielastase imbalance,' mechanism that causes the large majority of COPD cases in cigarette smokers who do not have ATD (Crystal, 1990). Metallo-proteases, either endogenous or expressed by invading microbes, are known to inactivate AT by proteolytic cleavage and therefore may also play a role in protease-inhibitor imbalance that contributes to the pathogenesis of COPD.

Mechanism of Deficiency in the Classical form of ATD

The mutant ATZ molecule is characterized by a single amino acid substitution, Lys for glu 342 (reviewed in Perlmutter,

2011a,b). This substitution is sufficient to decrease secretion with intracellular accumulation of the mutant protein. This alteration in biogenesis has been observed in multiple cell types and can explain the reduction in serum levels of AT that occurs in affected homozygotes (reviewed in Perlmutter, 2011a,b). The alteration in biogenesis and intracellular accumulation of mutant ATZ is selective in that the fate of other secretory proteins is not affected.

Studies by Lomas and colleagues and later by Huntington have shown that ATZ is prone to polymerization and aggregation (Lomas *et al.*, 1992; Belorgey *et al.*, 2007; Yamasaki *et al.*, 2008) and it now appears that the formation of insoluble polymers/aggregates play an important role in the pathobiology of liver disease (Figure 2). These investigators have also concluded that polymerization of ATZ is the cause of

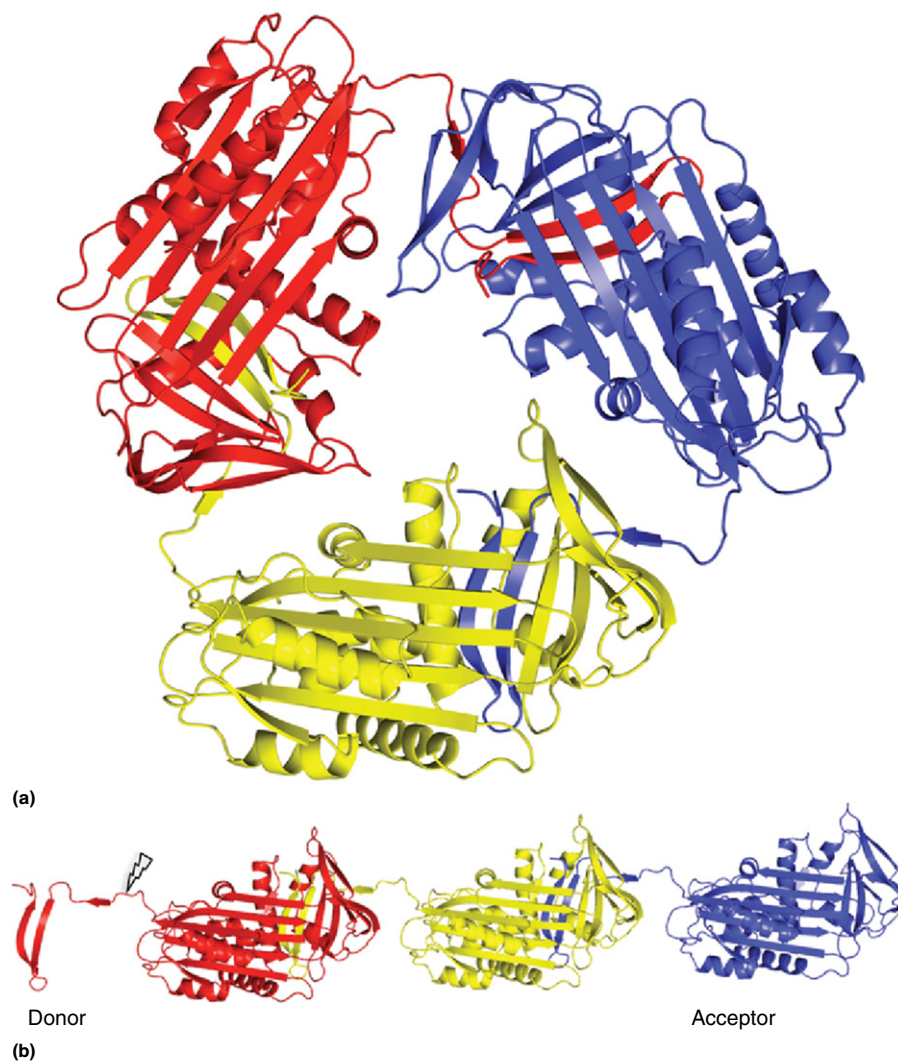


Figure 2 Ribbon diagram to show the structure of mutant ATZ in a trimer as depicted by Yamasaki *et al.*, 2008. Three crystallographically related monomers assemble into the trimer through C-terminal domain swap shown in panel (a). Each monomer is shown in a different color. A model of an open trimer (b), colored as in panel (a), was created by breaking one intermolecular contact and orienting the protomers in a head-to-tail manner. The red monomer is able to donate the C terminus (donor) and the blue monomer can accept a C terminus (acceptor) to elongate or self-terminate. The lightning bolt indicates a site of likely proteolytic susceptibility. This trimer is thought to be one of probably multiple types of polymers/aggregates that ATZ forms. This diagram was reproduced with permission from Yamasaki, M., Li, W., Johnson, D.J., Huntington, J.A., 2008. Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature* 455, 1255–1258.

defective secretion and intracellular accumulation. The most convincing evidence for this conclusion is a series of studies in which secretion of ATZ is partially corrected by introduction of a second mutation that suppresses polymerization (Sidhar *et al.*, 1995). However, these results could not exclude the possibility that intracellular retention is caused by a distinct abnormality in folding that is also partially corrected by the second experimentally introduced mutation. Furthermore, there is numerous observations that militate against polymerization as the cause of defective secretion. First, naturally occurring variants of AT that do not polymerize are retained in the ER (Lin *et al.*, 2001). Second, only ~18% of the intracellular pool of ATZ at steady state is in the form of polymers in model cell lines characterized by marked ER retention (Lin *et al.*, 2001; Schmidt and Perlmutter, 2005). Third, peptides and peptidomimetic drugs which interfere with polymerization of ATZ do not correct the secretory defect but rather increase intracellular degradation (Alam *et al.*, 2012). Taken together, we believe the data are most consistent with the notion that altered folding leads to accumulation in the secretory pathway and that polymerization and aggregation of retained ATZ occurs secondarily – i.e., the primary effect is protein misfolding. This is not a trivial distinction because it means that chemical or genetic interventions that interfere with polymerization will not correct the primary secretory defect.

Cellular Mechanisms of Liver Disease

Over the years it has become clear that the liver disease which occurs in some individuals with ATD is caused by a gain-of-function mechanism triggered by aberrant accumulation of misfolded ATZ in the secretory pathway, a mechanism that has come to be referred to as ‘gain-of-function proteotoxicity’ in diseases caused by misfolded proteins (reviewed in Perlmutter and Silverman, 2011). The strongest evidence for this mechanism comes from transgenic mouse models of ATD that have hepatic histological findings that recapitulate what is seen in humans with ATD (Figure 3), including fibrosis with minimal inflammation, mild steatosis, glycogen depletion and carcinogenesis (Hidvegi *et al.*, 2010; Rudnick *et al.*, 2004; Marcus *et al.*, 2010). These effects cannot be from loss-of-function because the mice still express endogenous anti-elastases. Recent studies have also demonstrated intracellular accumulation of ATZ with proteotoxic consequences in a transgenic *Caenorhabditis elegans* model of ATD (Long *et al.*, 2014).

The mechanisms by which ATZ accumulation leads to cellular injury are not known but studies in mammalian cell line models and the PiZ mouse model of ATD have provided evidence for mitochondrial damage with caspase-3 activation *in vivo* (Teckman *et al.*, 2004; Teckman *et al.*, 2002). Similar alterations were noted in the liver of α 1-AT-deficient patients. Perhaps even more important, treatment of the PiZ mouse

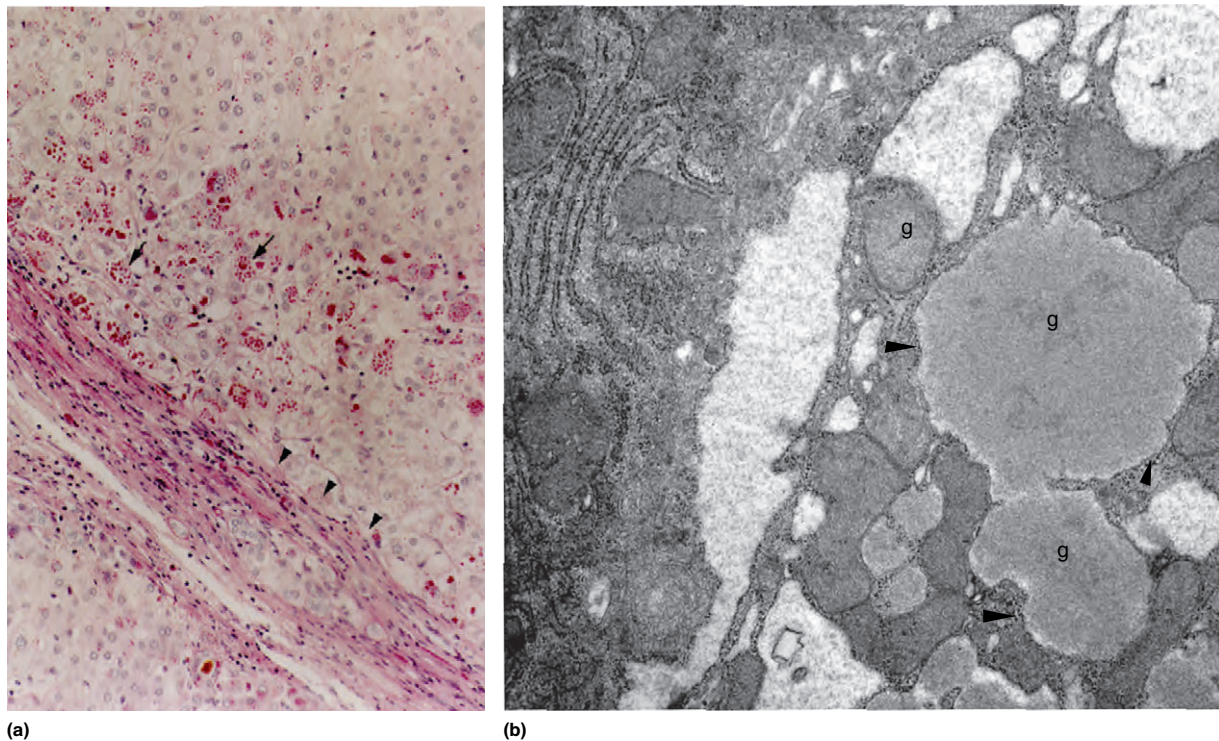


Figure 3 Hepatic histopathology in α 1-antitrypsin deficiency. (a) Micrograph of liver section stained with periodic acid–Schiff (PAS) followed by diastase treatment. The section shows characteristic PAS-positive, diastase-resistant globules (arrows) and a band of fibrosis (arrowheads). (b) Electron micrograph of a biopsy specimen showing globules in endoplasmic reticulum. Globules (g) and ribosomes (arrowheads) are indicated. Panel (a) has been reproduced Wang, Y., Perlmutter, D.H., 2014. Targeting intracellular degradation pathways for treatment of liver disease caused by α 1-antitrypsin deficiency. *Pediatric Research* 75(1–2), 133–139 with permission from Nature Publishing Group. Panel (b) was kindly provided by Donna Stolz, University of Pittsburgh.

model with cyclosporine A, but not tacrolimus, resulted in a significant decrease in mitochondrial dysfunction, caspase-3 activation and mortality. Because CsA is known to inhibit the mitochondrial permeability transition but tacrolimus does not have this effect, the results implicate mitochondrial damage in the hepatocellular toxicity of ATD.

Fibrosis and eventually cirrhosis are the major pathological features of the liver in ATD. Studies using a transgenic mouse model with liver-specific inducible expression of misfolded ATZ, ideal for identifying changes in gene expression in response to ATZ accumulation, have shown that there is increased TGF β signaling which is likely to at least in part play a role in the hepatic fibrotic characteristics of ATD (Hidvegi *et al.*, 2007).

Although ATZ is known to polymerize and form insoluble aggregates, it is still not clear whether it is the polymers/aggregates that are cytotoxic or if monomeric forms of the misfolded protein are responsible. In at least one other disease caused by a misfolded protein, Alzheimer's disease (AD), soluble oligomers of amyloid β rather than insoluble aggregates are believed to mediate proteotoxicity (Haass and Selkoe, 2007). Indeed, some evidence in mouse models of AD suggests that aggregation of amyloid β prevents its cytotoxicity (Cohen *et al.*, 2009). Further studies designed to identify whether monomers, soluble oligomers or polymers/aggregates of ATZ are responsible for its cellular toxicity will represent an important advance for understanding the pathobiology.

Autopsy studies have shown that ATD is strongly associated with primary liver cancer (Eriksson *et al.*, 1986). Hepatocellular carcinoma also increases in the PiZ mouse model of ATD (Marcus *et al.*, 2010). Utilizing BrdU labeling Rudnick *et al.* have shown hyperproliferation in the liver of the PiZ mouse model. Although the increase in hepatocellular proliferation was five- to tenfold above the control mice and highly statistically significant, there was still a relatively low number of BrdU positive hepatocytes at one time (2–3% detected over 72h of continuous labeling). These data indicate that liver injury in the mouse model is relatively mild and appropriately corresponds to the slowly progressing liver disease that is seen in most ATD patients. Almost all of the proliferating hepatocytes were globule-devoid. The origin of these globule-devoid hepatocytes is not known, but they are known to have less total α 1-ATZ per cell and less polymerized α 1-ATZ per cell compared with the globule-containing hepatocytes (An *et al.*, 2005). These studies of (Rudnick *et al.*, 2004) show that globule-devoid hepatocytes have a selective proliferative advantage over the neighboring globule-containing cells and because of that there are zones of globule-devoid hepatocytes in every liver biopsy from patients with ATD and in liver from the mouse models of ATD. The zones of globule-devoid hepatocytes also increase with age, occupying more of the liver area over time (Perlmutter, 2011a,b; Geller *et al.*, 1990; Geller *et al.*, 1994). Hepatic cancers arise predominantly from the zones of globule-devoid hepatocytes (Zhou and Fischer, 1998; Zhou *et al.*, 2000; Hadzic *et al.*, 2006). Furthermore, the degree of proliferation of the globule-devoid hepatocytes is directly proportional to the number of globule-containing hepatocytes (Rudnick *et al.*, 2004). We have speculated that the globule-devoid hepatocytes are younger cells, eventually destined to become globule-containing hepatocytes, but there is a conti-

nuous supply of younger cells as a part of the ongoing regenerative activity (Rudnick and Perlmutter, 2005). The globule-containing hepatocytes have more aggregated ATZ, activation of autophagy, and NF κ B activation. They are relatively impaired in regeneration and, although they are undergoing cell death (Lindblad *et al.*, 2007), it is at a relatively low rate and so we believe that they are also relatively impaired in death. The term 'relative' is used to describe the impairment in proliferation and death because these globule-containing cells will proliferate and die at higher rates after partial hepatectomy, a much more powerful regenerative stimulus than the proteotoxic state that is generated by chronic intracellular accumulation of a 1-ATZ.

Based on these observations, it has been theorized that the globule-containing hepatocytes are 'sick but not dead' and are responsible for regenerative signaling, and the globule-devoid hepatocytes are the predominant ones that can respond to these signals (Rudnick and Perlmutter, 2005). This hypothesis is supported by recent studies in which normal hepatocytes repopulate the livers of PiZ mice (Ding *et al.*, 2011). Wild-type donor hepatocytes spontaneously repopulated the livers of PiZ mice, replacing 20–98% of mutant host hepatocytes. The normal cells were shown to initially divide almost exclusively around globule-containing hepatocytes. Later they replaced globule-devoid hepatocytes presumably because these cells are 'sick' compared with the normal donor cells even though they are relatively 'less sick' than the original globule-containing hepatocytes. In essence, the theory (Rudnick and Perlmutter, 2005) envisions that cross-talk between the globule-containing and globule-devoid cells perpetuate chronic hepatocellular proliferation, eventually leading to carcinogenesis within globule-devoid zones. The relative impairment in cell death in the globule-containing hepatocytes could account for the small number of cancers that do arise in these cells (Zhou and Fischer, 1998; Zhou *et al.*, 2000; Hadzic *et al.*, 2006).

Variation in the incidence and severity of liver disease among individuals with classical form of ATD was first described by Sveger using a cohort of 127 homozygotes identified by a nationwide screening study in Sweden (Sveger, 1976; Piitulainen *et al.*, 2005). This represents the only unbiased analysis and the results over 40 years since then indicate that only 8–10% of the cohort have experienced clinically significant liver disease. We now know that some ATD homozygotes develop severe liver disease during infancy, childhood, adolescence or even at 50–65 years of age and others appear to completely escape liver disease. These observations have led us to hypothesize that genetic or environmental modifiers determine liver disease susceptibility among ATD homozygotes. Furthermore we theorized that such modifiers would target proteostasis mechanisms such as intracellular degradation pathways or signaling pathways that facilitate cellular adaptations to aberrant intracellular ATZ accumulation. The hypothesis was first supported in 1994 when Wu *et al.* showed a lag in degradation of misfolded ATZ in genetically engineered fibroblast cell lines from ATD patients with severe liver disease as compared to cell lines from ATD patients without liver disease (Wu *et al.*, 1994). Recently these results have been confirmed using iPS-derived hepatocytes (iHeps) (Tafaleng *et al.*, 2013). Intracellular degradation was significantly more efficient in iHeps from ATD patients without liver disease with

a half-time of ~2 h compared to a half-time of ~4 h in iHeps from ATD patients with severe liver disease requiring transplantation during childhood. These results indicate that iHeps can be used to identify proteostatic modifier mechanisms in other forms of ATD liver disease, infantile, adolescent and the adult age-dependent form of the disease.

There is still relatively limited information about specific genes that represent the putative modifiers. Pan *et al.* described a polymorphism in ER mannosidase I, an enzyme involved in the ERAD pathway, in infants with severe liver disease due to ATD (Pan *et al.*, 2009). The polymorphism is thought to lower the levels of ER mannosidase I and, in so doing, to diminish degradation of misfolded proteins by the ERAD pathway (Pan *et al.*, 2009; Iannotti *et al.*, 2014). This putative modifier mechanism is consistent with the concept that the modifiers will target proteostasis mechanisms such as ERAD in this case. However, it is difficult to imagine that a partial decrease in ERAD would have enough effect to predispose to severe liver disease in the first year of life because there are several lines of evidence that other degradation pathways can compensate for decreases in ERAD activity. For one thing studies in yeast suggest that autophagy can be activated to compensate for higher levels of ATZ accumulation when the ERAD pathway is impaired (Kruse *et al.*, 2006a,b). Second, studies in the *C. elegans* model of ATD indicate that previously uncharacterized degradation pathways, which are otherwise held in reserve, may be activated when autophagy is impaired (Long *et al.*, 2014; O'Reilly *et al.*, 2014).

Cellular Mechanisms of Lung Disease

AT is an inhibitor of serine proteases in general, but its most important targets are neutrophil elastase, cathepsin G, and proteinase 3, proteases released by activated neutrophils. Several lines of evidence suggest that inhibition of these neutrophil proteases is the major physiologic function of AT (Crystal, 1990). First, individuals with α 1-AT deficiency are susceptible to premature development of emphysema, a lesion that can be induced in experimental animals by instillation of excessive amounts of neutrophil elastase. Second, the kinetics of association for AT and neutrophil elastase are more favorable by several orders of magnitude than any other serine protease. Third, AT constitutes more than 90% of the neutrophil elastase inhibitory activity. These observations point to the notion that destructive lung disease may result from hampering the net balance of elastase and AT within the local environment of the lung, the so-called elastase-antielastase imbalance. Cigarette smoking markedly accelerates the destructive lung disease that is associated with ATD. It reduces the quality of life, and significantly shortens longevity of these individuals (Crystal, 1990). The effect of cigarette smoking is thought to be due to oxidative and proteolytic inactivation of residual AT function by active oxygen intermediates and metallo-proteases released from mononuclear phagocytes (Crystal, 1990; Stockley *et al.*, 2009). Furthermore, in the vast majority of cases of chronic obstructive lung disease in which there is a history of cigarette smoking and no ATD, the elastase-antielastase imbalance theory revolves around the concept that levels of functional AT are reduced exclusively by

oxidative and proteolytic inactivation as a result of cigarette smoking.

Nevertheless, variability in the development of lung disease among individuals with ATD appears to occur even when cigarette smoking is taken into consideration (DeMeo *et al.*, 2009). Genome-wide association studies and integrative genomic approaches in COPD have demonstrated significant associations in ATD patients for single-nucleotide polymorphism (SNPs) in the chromosome 15q region that includes *CHRNA3* (cholinergic nicotine receptor α 3) and *IREB2* (iron regulatory binding protein 2) (Kim *et al.*, 2012). Other potential modifiers of lung disease in ATD include *NOS3*, *GSTP1*, *TNF*, and *IL10* (Kim *et al.*, 2012).

Several observations have raised the possibility that gain-of-toxic function mechanisms can also contribute to the pathogenesis of COPD in ATD. For instance, deposits of ATZ, which are known to be chemotactic for neutrophils, have been found in the extracellular matrix of the lung next to neutrophils (Mahadeva *et al.*, 2005). These observations are consistent with early observations that COPD associated with AT deficiency had a much higher degree of neutrophilic infiltration and accelerated inflammatory destruction than other causes of COPD (Crystal, 1990; Stockley *et al.*, 2009). There is also evidence for synthesis of AT in lung epithelial cells (Hu and Perlmutter, 2002). If accumulation of ATZ in respiratory epithelial cells has the same 'cytostatic' consequences as it has in liver cells (see below) that could theoretically cause respiratory epithelial dysfunction that contributes to the progression of COPD. The fact that replacement therapy with purified AT has not completely arrested the progression of COPD in AT deficiency makes the notion that gain-of-toxic function mechanisms contribute to the tissue injury more appealing.

Cellular Responses to Misfolding of Mutant ATZ

Research in many laboratories has focused on how cells respond to aberrant accumulation of misfolded proteins. Our work has been designed to characterize how cells respond specifically to accumulation of mutant ATZ in the early compartments of the secretory pathway with the hypothesis that such responses will explain how some ATD individuals escape liver disease. Two general types of 'proteostatic' mechanisms were envisioned, intracellular degradation pathways and signaling pathways that permit the cell to adapt to the proteotoxic state.

The proteasomal pathway was the first to be implicated in intracellular degradation of mutant ATZ. Studies in cell-free microsomal, mammalian cell line and yeast model systems indicated involvement of proteasomal machinery in degradation of mutant ATZ (Qu *et al.*, 1996; Werner *et al.*, 1996; Teckman *et al.*, 2001). The participation of both classical ubiquitin-dependent and -independent pathways was implicated in removal of ATZ via proteasomes as a part of what later was termed the ER-associated degradation (ERAD) pathway (Qu *et al.*, 1996). It has been our assumption that ERAD and the proteasomal pathway predominantly mediate degradation of soluble forms of ATZ that accumulate in the ER but this point has never been definitively

proven. Although many of the molecules that mediate ERAD have been elucidated since then, the specific molecules involved in proteasomal degradation of ATZ and the sequence of molecular events involved are still not completely characterized.

In early studies it was also clear that the proteasomal pathway could not completely account for intracellular degradation of ATZ in model systems (Qu *et al.*, 1996; Werner *et al.*, 1996; Teckman *et al.*, 2001; Teckman *et al.*, 2000; Cabral *et al.*, 2000; Cabral *et al.*, 2002). A critical role for autophagy in intracellular disposal of ATZ was initially established by observations in genetically engineered human fibroblast cell line models of ATD (Teckman and Perlmutter, 2000). Autophagy is a catabolic process by which cytoplasmic components are sequestered within double-membrane vesicles called autophagosomes, which then fuse with lysosomes to form autolysosomes where its contents are degraded by acidic lysosomal hydrolases. Autophagy is involved in removing damaged organelles and misfolded or dysfunctional proteins and is critical for maintaining cell health (Rubinshtein *et al.*, 2011). In addition to the dramatic increase in number of autophagosomes seen in fibroblast cell lines a marked increase in autophagosomes was observed in the liver of PiZ mice and liver biopsy specimens from ATD patients (Teckman and Perlmutter, 2000; Teckman *et al.*, 2002). A significant delay in ATZ degradation in genetically engineered autophagy-deficient mammalian cell lines and yeast model systems provided definitive evidence in the role of autophagy in ATZ disposal (Kamimoto *et al.*, 2006; Kruse *et al.*, 2006a,b). Studies in yeast that are deficient for ATG6 and ATG16 showed that degradation of ATZ was affected only when its expression was very high, suggesting that of ERAD/proteasomal pathways was sufficient for disposing of ATZ at lower levels of expression whereas autophagy was needed when ATZ accumulated to a higher extent, presumably reflecting the importance of autophagy in degradation of polymerized and insoluble ATZ aggregates (Kruse *et al.*, 2006a,b). Alternatively, higher expression of ATZ may lead to additional uncharacterized post-translational modification(s) on ATZ that could render them inaccessible to the ERAD/proteasome system. Studies in *C. elegans* model system for ATD have recently provided evidence for the involvement of both proteasomal and autophagic pathways in ATZ disposal (Long *et al.*, 2014). Also, treatment of cell lines with chemical inhibitors of autophagy results in partial attenuation of ATZ degradation (Teckman and Perlmutter, 2000), while, activators of autophagy results in augmentation of ATZ degradation (Hidvegi *et al.*, 2010). Together, these studies strongly suggest autophagy as a critical pathway in mutant ATZ disposal (Figure 4).

Several studies have suggested that mechanisms other than the proteasomal and autophagic pathways contribute to the intracellular disposal of mutant ATZ. One such pathway described in a yeast model system involves targeting of ATZ from Golgi or TGN to lysosomes for degradation (Kruse *et al.*, 2006a). A recent study indicates involvement of sortilin in degradation of ATZ by the Golgi-lysosomal pathway in yeast and mammalian cell line model systems (Gelling *et al.*, 2012). It is speculated that any dysfunctional or aberrantly folded ATZ molecules that escape from ER into the Golgi are degraded by this pathway.

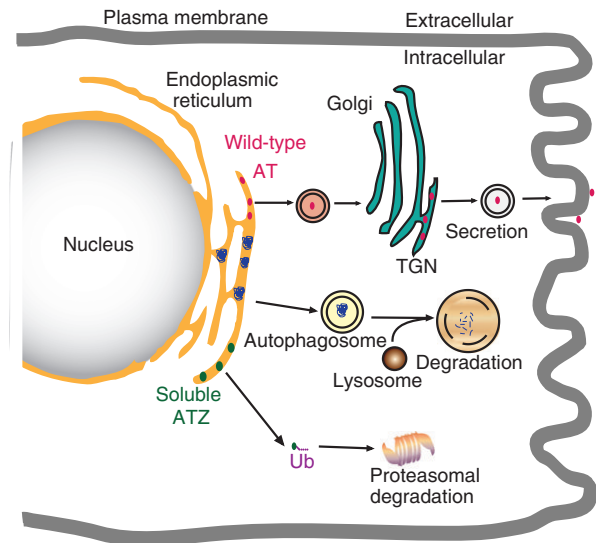


Figure 4 Subcellular itinerary of wild-type and mutant $\alpha 1$ -antitrypsin Z. Wild-type $\alpha 1$ -antitrypsin (AT; represented by red dots) is transported through the conventional secretory pathway as shown at the top of the figure. Mutant $\alpha 1$ -Antitrypsin Z (ATZ) accumulates in the ER in soluble (green dots) and insoluble polymers/aggregates (blue aggregates). The proteasomal pathway degrades soluble ATZ. The autophagic pathway is specialized for degradation of insoluble polymers/aggregates but may also play a role in degradation of soluble forms of ATZ.

There are likely to be still other mechanisms by which cells can degrade mutant ATZ and it is also likely that the different mechanisms of ATZ degradation can be induced to compensate for each other when stressed. These compensatory phenomenon will need to be taken into consideration as we further define the modifiers of proteotoxicity.

We have also envisioned that cells may respond to accumulation of misfolded ATZ by activating signaling pathways that are designed to facilitate adaptation and survival of cells under proteotoxic conditions. Studies in cell lines and mouse transgenic model systems with inducible expression of ATZ have been particularly useful in delineating signaling pathways activated as a primary response to ATZ accumulation, because the induction of these pathways would potentially be obscured in constitutively expressing model systems (Kamimoto *et al.*, 2006; Hidvegi *et al.*, 2005; Hidvegi *et al.*, 2007). In addition to autophagic response, activation of NF κ B and TGF β signaling pathways has been observed both in *in vitro* and *in vivo* model systems that were designed to express ATZ inducibly (Hidvegi *et al.*, 2005; Hidvegi *et al.*, 2007). The role of TGF β signaling pathway in hepatic fibrosis is well known (Wynn, 2007), and is presumed to be one of the mechanisms by which hepatic fibrosis occurs in ATD. Intriguingly, accumulation of ATZ in the ER does not appear to activate the unfolded protein response (UPR) pathway in most experimental systems (Hidvegi *et al.*, 2005). The lack of detectable UPR in ATD could be due to the inherent property of ATZ to polymerize and form insoluble aggregates as accumulation of other non-polymerogenic AT mutants in ER results in activation of UPR (Hidvegi *et al.*, 2007). Lack of UPR could be one of the reasons why globule-containing 'sick' cells survive and

could eventually contribute to pathogenesis of hepatocellular carcinoma.

Activation of NF κ B is believed to be a particularly important adaptive response to proteotoxicity in ATD patients (Hidvegi *et al.*, 2005). Interestingly, the assembled NF κ B complex in response to ATZ accumulation has a protein profile that is distinct from the complex which assembles upon treatment of cells with tunicamycin or tumor necrosis factor (Mukherjee and Perlmutter, unpublished data). One of the consequences of NF κ B activation is downregulation of Egr1, which is a transcription factor essential for hepatocyte proliferation and regenerative responses following partial hepatectomy (Liao *et al.*, 2004). Studies indicate that NF κ B activation could be involved in regulating liver cell proliferation and pathogenesis of hepatocellular carcinoma in AT deficiency (Maeda *et al.*, 2005). Intriguingly, mating PiZ mouse with conditional hepatocyte-specific NF κ B-deficient mouse gives rise to offsprings that exhibit more severe inflammation, fibrosis, dysplasia, and steatosis, and an increase in the number of globule-containing hepatocytes (Mukherjee and Perlmutter, unpublished data). This indicates a likely role of NF κ B signaling pathway acting as a protective mechanism against liver damage in AT deficiency. Ongoing studies are aimed at deeper understanding of the role of NF κ B signaling in ATD.

Genomic analyses of PiZ mouse liver tissue have identified other gene expression changes in response to hepatic accumulation of ATZ *in vivo*. Notably, upregulation of regulator of G protein signaling 16 (RGS16), could lead to activation of autophagy in the liver in ATD (Hidvegi *et al.*, 2007). The effect of RGS16 is thought to be mediated via its antagonistic effect on Gai3 which is known to inhibit hepatic autophagy (Gohla *et al.*, 2007). However, the exact molecular mechanism(s) by which RGS16 affects autophagy in the presence of ATZ is yet to be delineated.

Therapies for ATD Liver Disease

Orthotopic liver transplantation is the only treatment currently available for severe liver disease due to ATD and 5-year survival rates are greater than 90% for children and greater than 80% for adults (Kemmer *et al.*, 2008). Several novel strategies for treatment of ATD liver disease that would obviate the need for organ transplantation and extended immunosuppression are currently being investigated.

One of the newer strategies involves drugs that enhance autophagy. The rationale for this approach is based on studies in cell line and animal models showing that autophagy is specifically activated when ATZ accumulates in cells and that autophagy participates in intracellular disposal of ATZ (Kamimoto *et al.*, 2006; Kruse *et al.*, 2006a,b). It is also based on the observation that ATD liver disease is being recognized more frequently at 50–65 years of age, an age period which is known to be associated with a decline in autophagy function. The approach was first tested validated in studies which showed that carbamazepine (CBZ) enhanced autophagic degradation of ATZ in mammalian cell line models of ATD and that this drug could reduce hepatic ATZ load and hepatic fibrosis *in vivo* using the PiZ mouse model of ATD (Hidvegi

et al., 2010). Because CBZ is FDA approved and has been extensively used in humans as an anticonvulsant and mood stabilizer it is now being tested in a Phase II/III clinical trial for severe liver disease due to ATD.

High throughput automated screening of a drug library (LOPAC) using a recently developed *C. elegans* model system for ATD has identified five hit compounds that reduce cellular ATZ load in a dose dependent manner (Gosai *et al.*, 2010). Four of these five compounds have autophagy enhancer activity and are FDA approved for use in clinical practice. Importantly, two of the hit compounds belong to phenothiazine drug family that is structurally related to tricyclic antidepressants, including CBZ. The phenothiazines have been shown to be effective in enhancing autophagic disposal of aggregation-prone protein huntingtin, which causes Huntington's disease (Tsvetkov *et al.*, 2010; Sarkar *et al.*, 2007; Zhang *et al.*, 2007). One of these drugs, fluphenazine, has been tested in the *C. elegans* model of ATD as well as in mammalian cell line models and in the PiZ mouse. The results show that fluphenazine a positive effect in all of the models with reduction in hepatic ATZ load and hepatic fibrosis *in vivo* (Li *et al.*, 2014). This type of high throughput screening strategy and two novel strategies for chemical and computational-based drug discovery using the autophagy enhancer drug paradigm and the phenothiazine structure will become extremely helpful for discovering additional drugs that could be used for treating liver disease due to ATD (Long *et al.*, 2014; O'Reilly *et al.*, 2014).

Several research groups are developing other strategies for enhancing autophagy that could potentially be used for treatment of ATD liver disease. A novel autophagy-inducing peptide Tat-beclin 1 has been characterized which augments degradation of mutant huntingtin and several invasive bacterial and viral pathogens (Kyei *et al.*, 2009; Shoji-Kawata *et al.*, 2013). This peptide has been suggested to be useful in treating a broad range of human diseases caused by proteotoxicity mediated by aggregation-prone proteins and infectious pathogens. Viral-mediated gene transfer of transcription factor EB (TFEB), a master gene that regulates autophagy and lysosomal gene expression, has been reported to reduce hepatic ATZ load and liver fibrosis in animal models (Pastore *et al.*, 2013). *In vitro* studies show that TFEB reduces ATZ load in autophagy-dependent manner. Thus TFEB appears to be an excellent target for potential therapeutic drugs. A number of other drugs which have been shown to enhance autophagy are described in other reviews (Ghouse *et al.*, 2014; Wang and Perlmutter, 2014; Rubinshtein *et al.*, 2012).

Several gene therapy approaches currently under investigation are based on knocking down expression of mutant ATZ gene using vectors that will also harbor gene for wild-type AT and thereby could address both gain-of-function proteotoxicity and loss-of-function mechanisms of disease in ATD (Li *et al.*, 2011; Mueller *et al.*, 2012). In one approach, adeno-associated virus harboring short-hairpin RNA to knockdown endogenous ATZ expression was used together with a codon-optimized wild-type AT transgene cassette (Li *et al.*, 2011). In the other approach, an adeno-associated virus harboring microRNA to silence endogenous AT gene expression was used together with a microRNA-resistant wild-type AT gene (Mueller *et al.*, 2012). Both the approaches led to significant

reduction in hepatic ATZ load and high levels of human AT in the serum of a transgenic mouse model of ATD. However, the effect on liver fibrosis was not as extensive and so further studies are needed to test whether more potent and widespread silencing would be more effective. A more recent study using antisense oligonucleotides to silence ATZ expression had a more extensive effect in reducing hepatic fibrosis in the PiZ mouse by subcutaneous injection twice per week for 4–8 weeks (Guo *et al.*, 2014).

A ‘structure-based’ screening strategy has been employed by several groups that aim at designing and identifying peptides to prevent polymerization of the mutant ATZ and assist in its secretion. A small molecule compound engineered to target a lateral hydrophobic cavity in ATZ prevented its polymerization, however, when tested in a cell line model, it was found to promote increased intracellular degradation and had minimal effect on secretion (Mallya *et al.*, 2007). Another small molecule peptide targeting the reactive center loop of AT has been tested in cell line model systems, and was reported to promote increased secretion ATZ and reversed levels of polymerized ATZ. However, the efficiency of this peptide in reducing ATZ load and liver damage in animal models is yet to be tested (Alam *et al.*, 2012). Although promising, the use of this strategy in treatment of ATD will require rigorous studies to identify drugs based on the peptide sequence and thorough testing for their safety and efficacy.

The use of chemical chaperones that can facilitate folding of multiple misfolded proteins non-selectively have also been investigated as a potential therapeutic class for treatment of ATD. Among the chemical chaperones tested, glycerol and 4-phenylbutyric acid (PBA) were found to mediate a dramatic increase in the secretion of ATZ in mammalian cell culture model system (Burrows *et al.*, 2000). PBA, in particular, enhanced secretion of functionally active ATZ in cell culture model system. Importantly, oral administration of PBA in PiZ mice mediated an increase in blood levels of human AT reaching 20–50% of the levels present in PiM mice and normal humans. Because PBA has been used safely in clinical studies in humans, it was considered an excellent candidate for chemoprophylaxis of target organ injury in ATD. However, preliminary studies in ten patients with ATD-associated liver disease, provided no significant evidence of enhancement of serum levels of AT after 14 days of treatment with PBA. Lack of apparent effects of PBA could be due to insufficient duration of treatment time or inability of patients with chronic liver disease to tolerate large dose of this drug required for its function in humans (Teckman, 2004). As PBA selectively affects ATZ secretion and mediates its effect *in vivo*, it is required to further investigate its therapeutic potential if more favorable formulations become available. Recently, a drug similar to PBA, suberoylanilide hydroxamic acid (SAHA), has been reported to increase secretion of ATZ in cell line models of ATD (Bouchecareilh *et al.*, 2012). However, further studies are required to delineate whether this effect is due to increased synthesis of ATZ or due to its ability to reduce ATZ accumulation in cells. If the increased secretion of ATZ is primarily due to its increased synthesis, this may pose problems as treatment with SAHA may lead to increased cellular proteotoxicity in ATD patients. Moreover, rigorous studies remain to be

performed to study the effect of SAHA using *in vivo* models of ATD.

Hepatocyte transplantation therapy as an alternative to whole liver transplantation has been used successfully in treatment of several metabolic liver diseases (Fox *et al.*, 1998; Muraca *et al.*, 2002) and is now performed in at least 12 centers in 10 countries. Recent studies show that wild-type donor hepatocytes can repopulate almost the entire liver of the PiZ mouse model of ATD (Ding *et al.*, 2011). As the transplanted hepatocytes have a selective proliferative advantage over ATZ-containing endogenous hepatocytes and can replace the latter in a diseased liver, this type of therapy holds a potential to prevent ATD-associated COPD and possibly liver diseases.

In addition to these strategies, a combinatorial method of cell-based therapy and gene targeting could be a novel approach for treating COPD and liver diseases in ATD. Recent studies have shown that the mutation in the AT gene was corrected in human induced pluripotent stem (iPS) cells derived from an ATD patient homozygous for ATZ using a combination of zinc-finger nucleases and transposon techniques. Moreover, the corrected iPS cell lines could be successfully engrafted into the liver of a transgenic mouse model (Yusa *et al.*, 2011). This strategy, if adapted successfully for use in clinical application, has the potential to correct both the loss-of-function and gain-of-function mechanisms, and would have the advantage of personalized treatment options which will obviate the need for immunosuppression.

Therapies for ATD Lung Disease

Avoidance of cigarette smoking is a critical part of preventing lung disease and its progression. Replacement therapy with purified AT and lung transplantation are the two approaches currently used for treatment of severe COPD due to ATD. Replacement therapy has been shown to correct the decrease in AT levels and in neutrophil elastase inhibitory activity in bronchoalveolar lavage fluid (Crystal, 1990). Even though this therapeutic strategy has been used for over 20 years, there is still no evidence that it arrests the progression of ATD lung disease. Some studies have shown slowing of decline in forced expiratory volume in a subgroup of patients (Abboud *et al.*, 2005). The lack of clinical efficacy is likely due to multiple factors, including the fact that replacement therapy is only initiated late in the evolution of the disease and also that gain-of-function mechanisms contribute to the disease. Lung transplantation therapy has been used in over 2000 patients with ATD since 1995 with 50% of the recipients surviving after 8.6 years (Christie *et al.*, 2012).

See also: Cell Division/Death: Apoptosis: Autophagy. Molecular Principles, Components, Technology, and Concepts: Proteins: Diseases of Protein Folding: Huntington's Disease and Amyotrophic Lateral Sclerosis; Posttranslational Modifications: Key Players in Health and Disease. Protein Synthesis/Degradation: Protein Degradation – Intracellular:

Endoplasmic Reticulum-Associated Degradation and Protein Quality Control. Protein Synthesis/Degradation: Translation: Biogenesis of Secretory Proteins

References

- Abboud, R.T., Ford, G.T., Chapman, K.R., 2005. Emphysema in alpha1-antitrypsin deficiency: Does replacement therapy affect outcome? *Treatments in Respiratory Medicine* 4 (1), 1–8.
- Alam, S., Wang, J., Janciauskiene, S., Mahadeva, R., 2012. Preventing and reversing the cellular consequences of Z alpha-1 antitrypsin accumulation by targeting s4A. *Journal of Hepatology* 57, 116–124.
- An, J.K., Blomenkamp, K., Lindblad, D., Teckman, J.H., 2005. Quantitative isolation of alpha1AT mutant Z protein polymers from human and mouse livers and the effect of heat. *Hepatology* 41 (1), 160–167.
- Belorgey, D., Hägglöf, P., Karlsson-Li, S., Lomas, D.A., 2007. Protein misfolding and the serpinopathies. *Prion* 1 (1), 15–20.
- Boucheccarelh, M., Hutt, D.M., Szajner, P., Flotte, T.R., Balch, W.E., 2012. Histone deacetylase inhibitor (HDACi) suberoylanilide hydroxamic acid (SAHA)-mediated correction of α 1-antitrypsin deficiency. *Journal of Biological Chemistry* 287, 38265–38278.
- Burrows, J.A.J., Willis, L.K., Perlmutter, D.H., 2000. Chemical chaperones mediate increased secretion of mutant α 1-antitrypsin (α 1-AT) Z: A potential pharmacological strategy for prevention of liver injury and emphysema in α 1-AT deficiency. *Proceedings of the National Academy of Sciences of the United States of America* 97, 1796–1801.
- Cabral, C.M., Choudhury, P., Liu, Y., Sifers, R.N., 2000. Processing by endoplasmic reticulum mannosidases partitions a secretion-impaired glycoprotein into distinct disposal pathways. *Journal of Biological Chemistry* 275, 25015–25022.
- Cabral, C.M., Liu, Y., Moremen, K.W., Sifers, R.N., 2002. Organizational diversity among distinct glycoprotein endoplasmic reticulum-associated degradation programs. *Molecular Biology of the Cell* 13, 2639–2650.
- Carlson, J.A., Rogers, B.B., Sifers, R.N., *et al.*, 1988. Multiple tissues express alpha-1-antitrypsin in transgenic mice and man. *Journal of Clinical Investigation* 82, 26–36.
- Cohen, E., Paulsson, J.F., Blinder, P., *et al.*, 2009. Reduced IGF-1 signaling delays age-associated proteotoxicity in mice. *Cell* 139 (6), 1157–1169.
- Christie, J.D., Edwards, L.B., Kucheryavaya, A.Y., *et al.*, 2012. The Registry of the International Society for Heart and Lung Transplantation: 29th adult lung and heart-lung transplant report-2012. *Journal of Heart and Lung Transplantation* 10, 1073–1086.
- Crystal, R.G., 1990. Alpha-1-antitrypsin deficiency, emphysema and liver disease: Genetic basis and strategies for therapy. *Journal of Clinical Investigation* 95, 1343–1352.
- DeMeo, D.L., Campbell, E.J., Brantly, M.L., *et al.*, 2009. Heritability of lung function in severe alpha-1 antitrypsin deficiency. *Human Heredity* 67 (1), 38–45.
- Ding, J., Yannam, G.R., Roy-Chowdhury, N., *et al.*, 2011. Spontaneous hepatic repopulation in transgenic mice expressing mutant human α 1-antitrypsin by wild-type donor hepatocytes. *Journal of Clinical Investigation* 121 (5), 1930–1934.
- Eriksson, S., Carlson, J., Velez, R., 1986. Risk of cirrhosis and primary liver cancer in alpha 1-antitrypsin deficiency. *New England Journal of Medicine* 314 (12), 736–739.
- Fox, I.J., Chowdhury, J.R., Kaufman, S.S., *et al.*, 1998. Treatment of the Crigler-Najjar syndrome type I with hepatocyte transplantation. *New England Journal of Medicine* 338, 1422–1426.
- Geller, S.A., Nichols, W.S., Dyaico, M.J., Felts, K.A., Sorge, J.A., 1990. Histopathology of alpha 1-antitrypsin liver disease in a transgenic mouse model. *Hepatology* 12 (1), 40–47.
- Geller, S.A., Nichols, W.S., Kim, S., *et al.*, 1994. Hepatocarcinogenesis is the sequel to hepatitis in Z#2 alpha 1-antitrypsin transgenic mice: Histopathological and DNA ploidy studies. *Hepatology* 19 (2), 389–397.
- Gelling, C.L., Dawes, I.W., Perlmutter, D.H., Fisher, E.A., Brodsky, J.L., 2012. The endosomal protein sorting receptor sortilin has a role in trafficking α -1 antitrypsin. *Genetics* 192, 889–903.
- Ghouse, R., Chu, A., Wang, Y., Perlmutter, D.H., 2014. Mysteries of α 1-antitrypsin deficiency: Emerging therapeutic strategies for a challenging disease. *Disease Models & Mechanisms* 7 (4), 411–419.
- Gohla, A., Klement, K., Piekorz, R.P., *et al.*, 2007. An obligatory requirement for the heterotrimeric G protein G₁₃ in the autophagic action of insulin in the liver. *Proceedings of the National Academy of Sciences of the United States of America* 104, 3003–3008.
- Gosai, S.J., Kwak, J.H., Luke, C.J., *et al.*, 2010. Automated high-content live animal drug screening using *C. elegans* expressing the aggregation prone serpin α 1-antitrypsin Z. *PLoS One* 12, e15460.
- Guo, S., Booten, S.L., Aghajani, M., *et al.*, 2014. Antisense oligonucleotide treatment ameliorates alpha-1 antitrypsin-related liver disease in mice. *Journal of Clinical Investigation* 124, 251–261.
- Haass, C., Selkoe, D.J., 2007. Soluble protein oligomers in neurodegeneration: Lessons from the Alzheimer's amyloid beta-peptide. *Nature Reviews Molecular Cell Biology* 8, 101–112.
- Hadzic, N., Quaglia, A., Mieli-Vergani, G., 2006. Hepatocellular carcinoma in a 12-year-old child with PiZZ alpha1-antitrypsin deficiency. *Hepatology* 43 (1), 194.
- Hidvegi, T., Ewing, M., Hale, P., *et al.*, 2010. An autophagy-enhancing drug promotes degradation of mutant alpha1-antitrypsin Z and reduces hepatic fibrosis. *Science* 329 (5988), 229–232.
- Hidvegi, T., Mirnics, K., Hale, P., *et al.*, 2007. Regulator of G Signaling 16 is a marker for the distinct endoplasmic reticulum stress state associated with aggregated mutant alpha1-antitrypsin Z in the classical form of alpha1-antitrypsin deficiency. *Journal of Biological Chemistry* 282 (38), 27769–27780.
- Hidvegi, T., Schmidt, B.Z., Hale, P., Perlmutter, D.H., 2005. Accumulation of mutant alpha1-antitrypsin Z in the endoplasmic reticulum activates caspases-4 and -12, NF-kappaB, and BAP31 but not the unfolded protein response. *Journal of Biological Chemistry* 280, 39002–39015.
- Hood, J.M., Koep, L.J., Peters, R.L., *et al.*, 1980. Liver transplantation for advanced liver disease with alpha-1-antitrypsin deficiency. *New England Journal of Medicine* 302 (5), 272–275.
- Hu, C., Perlmutter, D.H., 2002. Cell-specific involvement of HNF-1beta in alpha(1)-antitrypsin gene expression in human respiratory epithelial cells. *American Journal of Physiology – Lung Cellular and Molecular Physiology* 282 (4), L757–L765.
- Hubbard, R.C., Ogushi, F., Fells, G.A., *et al.*, 1987. Oxidants spontaneously released by alveolar macrophages of cigarette smokers can inactivate the active site of alpha1-antitrypsin, rendering it ineffective as an inhibitor of neutrophil elastase. *Journal of Clinical Investigation* 80, 1289–1295.
- Iannotti, M.J., Figard, L., Sokac, A.M., Sifers, R.N., 2014. A Golgi-localized mannosidase (MAN1B1) plays a non-enzymatic gatekeeper role in protein biosynthetic quality control. *Journal of Biological Chemistry* 289 (17), 11844–11858.
- Kamimoto, T., Shoji, S., Mizushima, N., *et al.*, 2006. Intracellular inclusions containing mutant alpha1-antitrypsin Z are propagated in the absence of autophagic activity. *Journal of Biological Chemistry* 281, 4467–4476.
- Kemmer, N., Kaiser, T., Zacharias, V., Neff, G.W., 2008. Alpha-1-antitrypsin deficiency: Outcomes after liver transplantation. *Transplantation Proceedings* 40 (5), 1492–1494.
- Kim, W.J., Wood, A.M., Barker, A.F., *et al.*, 2012. Association of IREB2 and CHRNA3 polymorphisms with airflow obstruction in severe alpha-1-antitrypsin deficiency. *Respiratory Research* 22, 13–16.
- Koopman, P., Pavey, S., Lovel-Badge, R.H., 1989. Widespread expression of human alpha-1-antitrypsin in transgenic mice revealed by in situ hybridization. *Genes & Development* 3, 16–25.
- Kruse, K.B., Brodsky, J.L., McCracken, A.A., 2006a. Characterization of an *ERAD* gene as VPS30/ATG6 reveals two alternative and functionally distinct protein quality control pathways: One for soluble A1PIZ and another for aggregates of A1PIZ. *Molecular Biology of the Cell* 17, 203–212.
- Kruse, K.B., Brodsky, J.L., McCracken, A.A., 2006b. Autophagy: An ER protein quality control process. *Autophagy* 2, 135–137.
- Kyei, G.B., Dinkins, C., Davis, A.S., *et al.*, 2009. Autophagy pathway intersects with HIV-1 biosynthesis and regulates viral yields in macrophages. *Journal of Cell Biology* 186, 255–268.
- Li, C., Xiao, P., Gray, S.J., Weinberg, M.S., Samulski, R.J., 2011. Combination therapy utilizing shRNA knockdown and an optimized resistant transgene for rescue of diseases caused by misfolded proteins. *Proceedings of the National Academy of Sciences of the United States of America* 108, 14258–14263.
- Li, J., Pak, S.C., O'Reilly, L.P., *et al.*, 2014. Fluphenazine reduces proteotoxicity in *C. elegans* and mammalian models of alpha-1-antitrypsin deficiency. *PLoS One* 9 (1), e87260.
- Liao, Y., Shikapwashya, O.N., Shteyer, E., *et al.*, 2004. Delayed hepatocellular mitotic progression and impaired liver regeneration in early growth response-1-deficient mice. *Journal of Biological Chemistry* 279, 43107–43116.
- Lin, L., Schmidt, B., Teckman, J., Perlmutter, D.H., 2001. A naturally occurring nonpolymerogenic mutant of alpha 1-antitrypsin characterized by prolonged

- retention in the endoplasmic reticulum. *Journal of Biological Chemistry* 276 (36), 33893–33898.
- Lindblad, D., Blomenkamp, K., Teckman, J., 2007. Alpha-1-antitrypsin mutant Z protein content in individual hepatocytes correlates with cell death in a mouse model. *Hepatology* 46 (4), 1228–1235.
- Lodish, H.F., Kong, N., 1987. Glucose removal from N-linked oligosaccharides is required for efficient maturation of certain secretory glycoproteins from the rough endoplasmic reticulum to the Golgi complex. *Journal of Cell Biology* 98, 1720–1729.
- Lomas, D.A., Evans, D.L., Finch, J.T., Carrell, R.W., 1992. The mechanism of Z alpha 1-antitrypsin accumulation in the liver. *Nature* 357, 605–607.
- Long, G.L., Chandra, T., Woo, S.L.C., Davie, E.W., Kurachi, K., 1984. Complete sequence of the cDNA for human α 1-antitrypsin and the gene for the S variant. *Biochemistry* 23, 4828–4837.
- Long, O.S., Benson, J.A., Kwak, J.H., *et al.*, 2014. A *C. elegans* model of human α 1-antitrypsin deficiency links components of the RNAi pathway to misfolded protein turnover. *Human Molecular Genetics* 23, 5109–5122.
- Maeda, S., Kamata, H., Luo, J.L., Leffert, H., Karin, M., 2005. IKK β couples hepatocyte death to cytokine-driven compensatory proliferation that promotes chemical hepatocarcinogenesis. *Cell* 121 (7), 977–990.
- Mahadeva, R., Atkinson, C., Li, Z., *et al.*, 2005. Polymers of Z alpha1-antitrypsin co-localize with neutrophils in emphysematous alveoli and are chemotactic *in vivo*. *American Journal of Pathology* 166 (2), 377–386.
- Mallya, M., Phillips, R.L., Saldanha, S.A., *et al.*, 2007. Small molecules block the polymerization of Z α 1-antitrypsin and increase the clearance of intracellular aggregates. *Journal of Medicinal Chemistry* 50 (22), 5357–5363.
- Marcus, N.Y., Brunt, E.M., Blomenkamp, K., *et al.*, 2010. Characteristics of hepatocellular carcinoma in a murine model of alpha-1-antitrypsin deficiency. *Hepatology Research* 40 (6), 641–653.
- Mueller, C., Tang, Q., Gruntman, A., *et al.*, 2012. Sustained miRNA-mediated knockdown of mutant AAT with simultaneous augmentation of wild-type AAT has minimal effect on global liver miRNA profiles. *Molecular Therapy* 20, 590–600.
- Mukherjee, A., Perlmutter, D.H., unpublished data. Activation of a novel NF- κ B pathway as a protective response to the mutant alpha-1-antitrypsin Z accumulation in the liver.
- Muraca, M., Gerunda, G., Neri, D., *et al.*, 2002. Hepatocyte transplantation as a treatment for glycogen storage disease type 1a. *Lancet* 359 (9303), 317–318.
- Ong, D.S., Kelly, J.W., 2011. Chemical and/or biological therapeutic strategies to ameliorate protein misfolding diseases. *Current Opinion in Cell Biology* 23 (2), 231–238.
- O'Reilly, L.P., Long, O.S., Cobanoglu, M.C., *et al.*, 2014. A genome-wide RNAi screen identifies potential drug targets in a *C. elegans* model of α 1-antitrypsin deficiency. *Human Molecular Genetics* 23 (19), 5123–5132.
- Pan, S., Huang, L., McPherson, J., *et al.*, 2009. Single nucleotide polymorphism-mediated translational suppression of endoplasmic reticulum mannosidase I modifies the onset of end-stage liver disease in alpha-1-antitrypsin deficiency. *Hepatology* 50 (1), 275–281.
- Pastore, N., Blomenkamp, K., Annunziata, F., *et al.*, 2013. Gene transfer of master autophagy regulator TFEB results in clearance of toxic protein and correction of hepatic disease in alpha-1-antitrypsin deficiency. *EMBO Molecular Medicine* 5, 397–412.
- Perlmutter, D.H., 2011a. Alpha-1-antitrypsin deficiency: Importance of proteasomal and autophagic degradative pathways in disposal of liver disease-associated protein aggregates. *Annual Review of Medicine* 62, 333–345.
- Perlmutter, D.H., 2011b. Alpha-1-antitrypsin deficiency. In: Schiff, E.R., Sorrell, M.F., Maddrey, W.C. (Eds.), *Schiff's Diseases of the Liver*, eleventh ed. Oxford: Wiley-Blackwell, pp. 835–867.
- Perlmutter, D.H., Glover, G.I., Rivetna, M., Schastee, C.S., Fallon, R.J., 1990. Identification of a serpin-enzyme complex receptor on human hepatoma cells and human monocytes. *Proceedings of the National Academy of Sciences of the United States of America* 87 (10), 3753–3757.
- Perlmutter, D.H., Kay, R.M., Cole, F.S., *et al.*, 1985. The cellular defect in alpha 1-proteinase inhibitor (alpha 1-PI) deficiency is expressed in human monocytes and in *Xenopus* oocytes injected with human liver mRNA. *Proceedings of the National Academy of Sciences of the United States of America* 82, 6918–6921.
- Perlmutter, D.H., May, L.T., Sehgal, P.B., 1989. Interferon beta 2/interleukin 6 modulates synthesis of alpha 1-antitrypsin in human mononuclear phagocytes and in human hepatoma cells. *Journal of Clinical Investigation* 84 (1), 138–144.
- Perlmutter, D.H., Silverman, G.A., 2011. Hepatic fibrosis and carcinogenesis in α 1-antitrypsin deficiency: A prototype for chronic tissue damage in gain-of-function disorders. *Cold Spring Harbor Perspectives in Biology* 3 (3), a005801.
- Piitulainen, E., Carlson, J., Ohlsson, K., Sveger, T., 2005. Alpha1-antitrypsin deficiency in 26-year-old subjects: Lung, liver, and protease/protease inhibitor studies. *Chest* 128 (4), 2076–2081.
- Qu, D., Teckman, J.H., Omura, S., Perlmutter, D.H., 1996. Degradation of a mutant secretory protein, alpha1-antitrypsin Z, in the endoplasmic reticulum requires proteasome activity. *Journal of Biological Chemistry* 271, 22791–22795.
- Rubinshtein, D.C., Codogno, P., Levine, B., 2012. Autophagy modulation as a potential therapeutic target for diverse diseases. *Nature Reviews Drug Discovery* 11 (9), 709–730.
- Rubinshtein, D.C., Mariño, G., Kroemer, G., 2011. Autophagy and aging. *Cell* 146, 682–695.
- Rudnick, D.A., Liao, Y., An, J.K., *et al.*, 2004. Analyses of hepatocellular proliferation in a mouse model of alpha-1 antitrypsin deficiency. *Hepatology* 39 (4), 1048–1055.
- Rudnick, D.A., Perlmutter, D.H., 2005. Alpha-1-antitrypsin deficiency: A new paradigm for hepatocellular carcinoma in genetic liver disease. *Hepatology* 42 (3), 514–521.
- Sarkar, S., Perlstein, E.O., Imarisio, S., *et al.*, 2007. Small molecules enhance autophagy and reduce toxicity in Huntington's disease models. *Nature Chemical Biology* 3, 331–338.
- Schmidt, B.Z., Perlmutter, D.H., 2005. Grp78, Grp94, and Grp170 interact with alpha1-antitrypsin mutants that are retained in the endoplasmic reticulum. *American Journal of Physiology – Gastrointestinal and Liver Physiology* 289 (3), G444–G455.
- Shoji-Kawata, S., Sumpter, R., Leveno, M., *et al.*, 2013. Identification of a candidate therapeutic autophagy-inducing peptide. *Nature* 494, 201–206.
- Sidhar, S.K., Lomas, D.A., Carrell, R.W., Foreman, R.C., 1995. Mutations which impede loop/sheet polymerization enhance the secretion of human alpha 1-antitrypsin deficiency variants. *Journal of Biological Chemistry* 270, 8393–8396.
- Stockley, R.A., Mannino, D., Barnes, P.J., 2009. Burden and pathogenesis of chronic obstructive pulmonary disease. *Proceedings of the American Thoracic Society* 6 (6), 524–526.
- Sveger, T., 1976. Liver disease in alpha-1-antitrypsin deficiency detected by screening of 200,000 infants. *New England Journal of Medicine* 294 (24), 1316–1321.
- Tafaleng, E.N., Han, B., Hale, P., *et al.*, 2013. The rate of disappearance of intracellular α 1 antitrypsin correlates with liver disease severity in iPSC-derived hepatocytes generated from PIZZ α 1-antitrypsin deficiency. *Hepatology* 58 (4), 81A.
- Teckman, J.H., 2004. Lack of effect of oral 4-phenylbutyrate on serum alpha-1-antitrypsin in patients with alpha-1-antitrypsin deficiency: A preliminary study. *Journal of Pediatric Gastroenterology and Nutrition* 39, 34–37.
- Teckman, J.H., An, J.K., Blomenkamp, K., Schmidt, B., Perlmutter, D.H., 2004. Mitochondrial autophagy and injury in the liver in alpha 1-antitrypsin deficiency. *American Journal of Physiology – Gastrointestinal and Liver Physiology* 286 (5), G851–G862.
- Teckman, J.H., An, J.K., Loethen, S., Perlmutter, D.H., 2002. Fasting in alpha-1-antitrypsin deficient liver: Constitutive [correction of consultative] activation of autophagy. *American Journal of Physiology – Gastrointestinal and Liver Physiology* 283 (5), G1156–G1165.
- Teckman, J.H., Burrows, J., Hidvegi, T., *et al.*, 2001. The proteasome participates in degradation of mutant alpha 1-antitrypsin Z in the endoplasmic reticulum of hepatoma derived hepatocytes. *Journal of Biological Chemistry* 276, 44865–44872.
- Teckman, J.H., Gilmore, R., Perlmutter, D.H., 2000. Role of ubiquitin in proteasomal degradation of mutant α 1-antitrypsin Z in the endoplasmic reticulum. *American Journal of Physiology – Gastrointestinal and Liver Physiology* 278, G39–G48.
- Teckman, J.H., Perlmutter, D.H., 2000. Retention of mutant alpha(1)-antitrypsin Z in endoplasmic reticulum is associated with an autophagic response. *American Journal of Physiology – Gastrointestinal and Liver Physiology* 279, G961–G974.
- Teckman, J.H., Qu, D., Perlmutter, D.H., 1996. Molecular pathogenesis of liver disease in alpha1-antitrypsin deficiency. *Hepatology* 24 (6), 1504–1516.
- Tsvetkov, A.S., Miller, J., Arrasate, M., *et al.*, 2010. A small-molecule scaffold induces autophagy in primary neurons and protects against toxicity in a Huntington disease model. *Proceedings of the National Academy of Sciences of the United States of America* 107, 16982–16987.
- Wang, Y., Perlmutter, D.H., 2014. Targeting intracellular degradation pathways for treatment of liver disease caused by α 1-antitrypsin deficiency. *Pediatric Research* 75 (1–2), 133–139.
- Werner, E.D., Brodsky, J.L., McCracken, A.A., 1996. Proteasome-dependent endoplasmic reticulum-associated protein degradation: An unconventional route to a familiar fate. *Proceedings of the National Academy of Sciences of the United States of America* 93, 13797–13801.
- Wu, Y., Whitman, L., Molmenti, E., *et al.*, 1994. A lag in intracellular degradation of mutant alpha 1-antitrypsin correlates with the liver disease phenotype in

- homozygous PiZZ alpha 1-antitrypsin deficiency. Proceedings of the National Academy of Sciences of the United States of America 91 (19), 9014–9018.
- Wynn, T.A., 2007. Common and unique mechanisms regulate fibrosis in various fibroproliferative diseases. Journal of Clinical Investigation 117, 524–529.
- Yamasaki, M., Li, W., Johnson, D.J., Huntington, J.A., 2008. Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. Nature 455, 1255–1258.
- Yusa, K., Rashid, S.T., Strick-Marchand, H., *et al.*, 2011. Targeted gene correction of α 1-antitrypsin deficiency in induced pluripotent stem cells. Nature 478, 391–394.
- Zhang, L., Yu, J., Pan, H., *et al.*, 2007. Small molecule regulators of autophagy identified by an image-based high-throughput screen. Proceedings of the National Academy of Sciences of the United States of America 104, 19023–19028.
- Zhou, H., Fischer, H.P., 1998. Liver carcinoma in PiZ alpha-1-antitrypsin deficiency. American Journal of Surgical Pathology 22 (6), 742–748.
- Zhou, H., Ortiz-Pallardó, M.E., Ko, Y., Fischer, H.P., 2000. Is heterozygous alpha-1-antitrypsin deficiency type PiZ a risk factor for primary liver carcinoma? Cancer 88 (12), 2668–2676.

ENCYCLOPEDIA OF CELL BIOLOGY

ENCYCLOPEDIA OF CELL BIOLOGY

EDITORS-IN-CHIEF

RALPH A BRADSHAW

University of California, Irvine, CA, USA

PHILIP D STAHL

Washington University School of Medicine, St. Louis, MO, USA

VOLUME 2

ORGANIZATIONAL CELL BIOLOGY

VOLUME EDITORS

BRUNO GOUD

Institut Curie, Paris, France

GRAÇA RAPOSO

Institut Curie, Paris, France

SECTION EDITORS

ELLIOT ELSON

Washington University School of Medicine, St. Louis, Missouri, USA

GRAÇA RAPOSO

Institut Curie, Paris, France

PAUL GLEESON

The University of Melbourne, Melbourne, Victoria, Australia

ANNA AKHMANOVA

Utrecht University, Kruytgebouw, Utrecht, The Netherlands

PHILIP D STAHL

Washington University School of Medicine, St. Louis, MO, USA



AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
225 Wyman Street, Waltham, MA 02451, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

Copyright © 2016 Elsevier Inc. All rights reserved

Cover background image created by M W Goldberg and T D Allen.

Foreground images © Clifford Harding, John Heuser and Philip D Stahl, 1983, originally published in *The Journal of Cell Biology*, Volume 97, 329–339.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-394447-4

For information on all publications visit our
website at <http://store.elsevier.com>

Printed and bound in the United States of America

15 16 17 18 19 20 10 9 8 7 6 5 4 3 2 1



Working together
to grow libraries in
developing countries

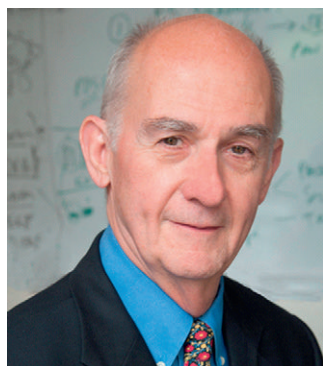
www.elsevier.com • www.bookaid.org

Publisher: Adam Gilbert
Acquisition Editor: Ginny Mills
Content Project Manager: Blerina Osmanaj
Associate Content Project Manager: Victoria Dyer
Cover Designer: Maria Inês Cruz

EDITORS-IN-CHIEF



Ralph A Bradshaw is Professor Emeritus in the Department of Physiology and Biophysics at the University of California, Irvine. Prior to that he was on the faculty of the Department of Biological Chemistry, Washington University School of Medicine in St. Louis, MO and was Professor and Chair of the Department of Biological Chemistry at University of California, Irvine. From 2006 to 2015, he was a member of the Mass Spectrometry Facility and Professor of Pharmaceutical Chemistry at the University of California, San Francisco. He holds degrees from Colby College and Duke University and was a post-doctoral fellow at Indiana University and the University of Washington. He has served as president of FASEB, was the founding president of the Protein Society and was the treasurer of the American Society for Biochemistry and Molecular Biology. His research has focused on protein chemistry and proteomics, with emphasis on the structure and function of growth factors and their receptors, particularly nerve growth factor and fibroblast growth factor, and the involvement of receptor tyrosine kinases in cell signaling. He has also studied the role of proteolytic processing and N-terminal modification in protein stability and turnover.



Philip D Stahl is E. Mallinckrodt Jr. Professor Emeritus at Washington University School of Medicine in St. Louis, MO. He was educated at West Virginia University with post-doctoral work at Vanderbilt University. He served as Head of the Department of Cell Biology and Physiology and Director of the Division of Biology and Biomedical Sciences at Washington University. He has been the recipient of many awards including a MERIT award from the NIH and the WICB Senior Recognition award given by the American Society for Cell Biology in recognition of his work supporting the advancement of women in science. Among StahlLab contributions are the discovery of the lysosomal enzyme clearance pathway now implemented in the treatment of lysosomal storage disease, discovery of the innate immune receptor, the mannose receptor, discovery of the exosome secretion pathway, and the role of Rab5 and Arf6 in endocytosis.

Currently his research focuses on endocytosis, signal transduction, and exosome biogenesis and secretion.

VOLUME EDITORS



Gerald Hart is the Director and DeLamar Professor of Biological Chemistry at Johns Hopkins Medical School. He began his research on glycoconjugates about 40 years ago as a graduate student, and he has been active in the field of Glycobiology ever since. Among his research accomplishments, he determined the minimal sequence requirements for N-linked glycosylation while a postdoc in William Lennarz's lab, and he collaborated with Paul Englund's group at Johns Hopkins to elucidate the biosynthetic pathway for GPI-anchors. In the early 1980s, while probing cells with glycosyltransferases, Hart's laboratory discovered cytoplasmic and nuclear protein glycosylation by O-linked N-acetylglucosamine (O-GlcNAc) (e.g., *Journal of Biological Chemistry* 259: 3308; *Journal of Biological Chemistry* 261: 8049). Since that time, the Hart laboratory has published nearly 200 papers on O-GlcNAcylation, identifying and cloning the enzymes controlling cycling, characterizing O-GlcNAcylation and its interplay with phosphorylation on hundreds of proteins, and they have developed many of the tools and methods in use today to study this modification. In

1989, Hart founded the leading journal in the field, *Glycobiology*, serving as Editor-in-Chief until 2001. Hart received the first International Glycoconjugate Organization (IGO) Award in 1997, the Karl Meyer Award from the Society for Glycobiology in 2006, and served as the 2009–2011 president of the IGO. Hart is currently an Associate Editor for *The Journal of Biological Chemistry* and an Associate Editor for *Molecular and Cellular Proteomics*. To date, Hart has published about 263 papers, all in the area of glycosciences. H factor=84.



Bruno Goud studied biochemistry and immunology at the Ecole Normale Supérieure de Cachan and the University of Paris. He received his PhD in 1981 under the mentorship of Jean-Claude Antoine and Stratis Avramias at Institut Pasteur and did postdoctoral work under the mentorship of Peter Novick at Yale University. In 1995, he established an independent research group in the Department of Cell Biology at the Institut Curie in Paris. The main focus of his studies is the regulation of intracellular transport and membrane trafficking in eukaryotic cells. In particular, his group is working on the functions of Rab GTPases, the mechanisms that sustain the global organization of intracellular compartments and the functions of myosins in membrane traffic. Since 2000, he has also developed original *in vitro* approaches to unravel physical parameters such as membrane tension or membrane curvature that underlie transport processes.

Bruno Goud is a recipient of an ERC (European Research Council) Advanced grant (2013). He is a member of the European Molecular Biology Organization (EMBO). He has been the Head of the Department of Cell Biology of Institut Curie since 2003.



Graça Raposo received her PhD in 1989 in Membrane Biology and Immunology at the University of Paris VII where she specialized in electron microscopy and membrane biology. From 1990 to 1995 she was a postdoctoral fellow at the Immunology Center in Marseille and then in the Department of Cell Biology, Utrecht University, The Netherlands. She is the deputy Director of the Department of Cell Biology at Institut Curie and Director of the Training unit. Her major research interests focus on the biogenesis and functions of exosomes and lysosome related organelles with implications in neurodegenerative disorders, lysosomal diseases, and cancer. Her group have started to unravel the cellular and molecular mechanisms regulating the biogenesis of melanosomes, the lysosome related organelles of epidermal melanocytes, studies that open a new avenue to modulate pigmentation in health and disease. In 2012 she was awarded the CNRS Silver Medal and in 2013 the Descartes Huygens Price from the Royal Dutch Academy of Sciences. She is a member of the European Molecular Biology Organization (EMBO).



Alan Ezekowitz, MBChB, DPhil, FAAP, is Co-Founder, President, and CEO of Abide Therapeutics, a company cofounded by Professors Dale Boger and Ben Cravatt from the Scripps Research Institute. He was previously Senior Vice President and Franchise Head, Bone, Respiratory, Immunology, Endocrine, Dermatology and Urology at Merck Research Laboratories. At Merck, he was responsible for the overall scientific direction of the drug discovery and development process. Prior to Merck, Dr. Ezekowitz was Chief of Pediatric Services at the Massachusetts General Hospital for Children and the Partners Healthcare System and the Charles Wilder Professor of Pediatrics at the Harvard Medical School. He chaired the committee that led to the establishment of the Academy at the Harvard Medical School where he served as a Scholar and Founding member. He served on the Board of Directors of the Partners Healthcare System and Massachusetts General Physicians Organization. In 2008 he was honored with the establishment R. Alan Ezekowitz Professorship in Pediatrics at the Harvard Medical School.

He is a member of the American Society of Clinical Investigation, the Association of American Physicians, the American Pediatric Society, and a Fellow of AAAS. He served on NIH Subcommittees on Biodefense and Vaccine Adjuvants; NAS panels on antibiotic resistance and academic pharmaceutical partnerships; and the ARISE 2 task force of American Academy of Arts & Science that is evaluating the impact of federal and industrial funding of science, engineering, and medicine on American universities.

Dr. Ezekowitz received his MBChB and DPhil from University of Cape Town and Oxford University, respectively. At Children's Hospital in Boston, he was a postdoctoral fellow in the Division of Hematology and Oncology with clinical training in pediatrics. He was on the staff of the Children's Hospital of Boston prior to moving to the Massachusetts General Hospital.

He is a pioneer in the field of innate immunity and has over 150 publications. In particular his group played a major role in defining the structure function of the mannose binding lectin and the macrophage mannose receptor.



Douglas A. Lauffenburger is Ford Professor of Bioengineering, and cofounder and Head of the Department of Biological Engineering at MIT, with affiliate appointment in the Department of Biology; he was a founding codirector of the MIT Computational and Systems Biology Initiative in 2002. His major research interests are in cell engineering, with central focus on cell-cell communication and intracellular signal transduction, emphasizing predictive computational models derived from quantitative experiment. Lauffenburger has coauthored a monograph entitled *Receptors: Models for Binding, Trafficking & Signaling* (Oxford University Press, 1993), has coedited the book entitled *Systems Biomedicine: Concepts and Perspectives* (Elsevier, 2010), and has supervised more than 100 doctoral and post-doctoral students. He is a member of the National Academy of Engineering and the American Academy of Arts and Sciences, and has served as President of the Biomedical Engineering Society, Chair of the AIMBE College of Fellows, and on the NIH NIGMS Advisory Council, and coauthored the NRC report on *A New Biology for the 21st Century*.

SECTION EDITORS



Kenneth E Neet received his PhD in Biochemistry in 1965 (with Dr. Frank W. Putnam) from the University of Florida. He was a postdoctoral fellow at the University of California, Berkeley (with Dr. Daniel E. Koshland, Jr.), and joined the faculty of Case Western Reserve University (CWRU) in 1967 as an Assistant Professor of Biochemistry.

Dr. Neet received a Faculty Research Award of the American Cancer Society from 1968 to 1973 and was on sabbatical leave at the National Institute for Medical Research, Mill Hill, England (with Dr. N. Michael Green) 1973–1974. From 1978 to 1990, he was Professor of Biochemistry at CWRU. During 1980–1981, he took a sabbatical year as a Josiah Macy Faculty Scholar in the Department of Neurobiology, Stanford University (with Dr. Eric M. Shooter).

Dr. Neet moved to Rosalind Franklin University of Medicine and Science/The Chicago Medical School in 1990 and was Professor and Chair of Biochemistry & Molecular Biology until 2005. Dr. Neet became Associate Dean for Research of the Chicago Medical School, RFUMS in 2004.

Dr. Neet has served on the Editorial Boards of the *Journal of Biological Chemistry*, *Protein Science*, and *Molecular & Cellular Proteomics* and as a member of study sections for the National Institutes of Health and the National Science Foundation. He was an Associate Editor of the *Journal of Biological Chemistry* (1996–2013).

Dr. Neet's research interests have been in the general areas of protein–protein interactions, allosteric interactions, protein conformational changes, and cell signaling. In his earlier years he studied theoretical and experimental aspects of slow transitions in enzymes, particularly the cooperativity of the monomeric enzyme glucokinase. Later, he studied ligand–receptor interactions of nerve growth factor, establishing a mutational system to study the interactions with its receptors (TrkA and p75), and ultimately the complex signaling emanating from these receptors within neuronal systems.

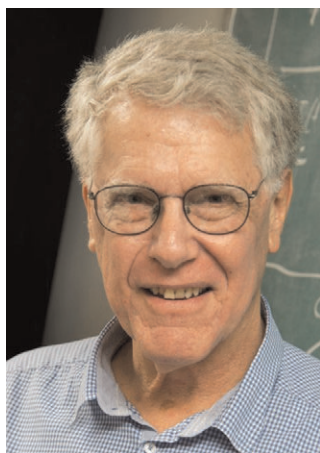


Sarah A Woodson is a biophysical chemist interested in how RNA molecules fold into specific three-dimensional structures and how they interact with proteins to turn genes on and off in the cell. She was born in Michigan, USA, and attended Kalamazoo College before receiving her PhD in Biophysical Chemistry from Yale University in 1987. After postdoctoral research in the laboratory of Tom Cech at the University of Colorado Boulder from 1987 to 1990, she joined the faculty of the University of Maryland in 1990 and the faculty of Johns Hopkins University in 1999, where she is currently the T.C. Jenkins Professor of Biophysics. Together with Mark Chance and Michael Brenowitz, she pioneered methods for visualizing how RNA molecules change shape in real time. She served on the board of the RNA Society and was elected an AAAS Fellow in 2011 and a Pew Scholar in the Biomedical Sciences in 1993. She is currently a reviewing editor of *Biopolymers* and *Journal of Molecular Biology* and serves on the editorial boards of *Nucleic Acids Research* and *RNA*.



Judith Bond, Evan Pugh Emeritus Professor of Biochemistry and Molecular Biology at the Pennsylvania State University College of Medicine, was President of the Federation of American Societies for Experimental Biology (2012–13) and an Associate Editor of the *Journal of Biological Chemistry* (1999–2013). She received her BS degree from Bennington College in Vermont, her PhD from Rutgers University, and did postdoctoral research at Vanderbilt University. She rose through the academic ranks at the Medical College of Virginia, Virginia Commonwealth University, became Professor and Head of Biochemistry and Nutrition at Virginia Polytechnic Institute and State University, and served at Penn State University College of Medicine as Professor and Chair of Biochemistry and Molecular Biology from 1992 to 2012. At Penn State, she also served as assistant dean for graduate studies, codirector of the All-University Interdisciplinary Biological Sciences Program and founding director of the Medical Scientist Training Program. She recently directed NIH-funded summer research programs for high school students and teachers and for undergraduate students (particularly underrepresented minority groups). Her work on metalloproteases has been funded continuously by the NIH for over 35 years. She has

trained 5 Master's, 21 PhD students, and 19 postdoctoral trainees. Her professional service included member and chair National Institutes of Health Study Sections, member of the National Institute of Diabetes, Digestive and Kidney Diseases Advisory Council, member of the American Association of Medical Colleges – Howard Hughes Medical Institute Committee on the Scientific Foundations for Future Physicians, and President of the American Society for Biochemistry and Molecular Biology.



Elliot Elson received his AB at Harvard, his PhD at Stanford under the mentorship of R.L. Baldwin and did postdoctoral work at University of California, San Diego, under Bruno Zimm. His first independent faculty position was in the Department of Chemistry, Cornell University, where he stayed for 11 years. He then moved to the Department of Biological Chemistry (now the Department of Biochemistry and Molecular Biophysics) at Washington University in St. Louis, School of Medicine. He has pursued research in several biophysical areas. One of these is the development of Fluorescence Correlation Spectroscopy and Fluorescence Photobleaching Recovery and their application to studies of diffusion in cell membranes and of phase separation in model membrane bilayers as well as of other cellular and noncellular phenomena. He has also worked in the areas of cell mechanics, cell motility, and the mechanical properties of engineered tissues. His laboratory has developed approaches for measuring mechanical properties of cells in monolayer culture and for determining the contributions of cells and extracellular matrix to the mechanics of engineered heart and connective tissue constructs.



Tamotsu Yoshimori received his PhD degree in 1989 at Osaka University. After working at several places including European Molecular Biology Laboratory (Prof. Kai Simons' lab) and National Institute of Basic Biology (Prof. Yoshinori Ohsumi's lab), he is now a distinguished professor of Osaka University (Graduate School of Medicine and of Frontier Biosciences). His research interests are focused on intracellular membrane trafficking, and especially for the last 18 years, on autophagy. He identified LC3 as an autophagosome-binding protein, which has been widely used as the gold standard in autophagy assays. The paper has been cited over 3000 times. He also provided new insights into membrane biogenesis in autophagy and the role of autophagy in pathogen defense and suppression of various diseases. He authored or coauthored over 220 journal articles and book chapters. He is an editor of *Journal Cell Science*, and on the editorial board of *Journal of Cell Biology*, *Molecular Biology of the Cell*, and so on. He is a member of the Faculty of 1000 and a vice president of Japan Society for Cell Biology. He was awarded Osaka University Presidential Award for 2012 and 2013, Prize for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology in 2013, Kakiuchi Saburo Memorial Prize by the Japanese Biochemical Society in 2014, and selected as Highly Cited Researchers 2014 by Thomson Reuters.



Paul A Gleeson is currently Head of the Department of Biochemistry and Molecular Biology at the University of Melbourne, located at Bio21 Institute. His research interests include the molecular mechanisms of intracellular membrane transport and the molecular basis of organ-specific autoimmune diseases. Paul Gleeson obtained his PhD in 1980 from the University of Melbourne and did postdoctoral research in the biosynthesis and function of glycoproteins at the Hospital for Sick Children, Toronto, National Institute for Medical research, Mill Hill London and Department of Biochemistry, La Trobe University, Melbourne. He established an independent laboratory at Monash University in 1986 where he defined the targeting signals of Golgi glycosyltransferases, identified golgins of the *trans*-Golgi network and along with his colleagues developed mouse models of autoimmune gastritis. In 2001 he moved to Department of Biochemistry and Molecular Biology at the University of Melbourne and has been Head of the Department since 2006. He has a number of international research collaborations and has been a visiting scientist at the EMBL, Heidelberg, and the Institut Curie, Paris.

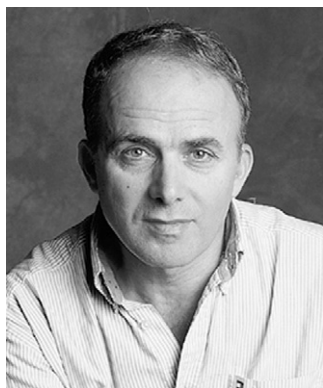


Anna Akhmanova studied biochemistry and molecular biology at the Moscow State University. She received her PhD in 1997 at the University of Nijmegen, the Netherlands. She worked as a postdoc at the Department of Microbiology and Evolutionary Biology at the University of Nijmegen and at the Department of Cell Biology at the Erasmus Medical Center in Rotterdam, the Netherlands. In 2001, she started her own research group at the Erasmus Medical Center. Since 2011, Anna Akhmanova is professor of Cell Biology at Utrecht University, the Netherlands.

Akhmanova studies cytoskeletal organization and trafficking processes, which contribute to cell polarization, differentiation, vertebrate development, and human disease. The main focus of her studies is the microtubule cytoskeleton. Her group has described key interactions between proteins associated with microtubule ends and determined their role in the regulation of microtubule organization and dynamics. She has also characterized mechanisms of bidirectional microtubule-based motility of membrane organelles such as cell nuclei and exocytotic vesicles. Her research relies on combining high-resolution live cell imaging and quantitative analysis of cytoskeletal remodeling, measurement of protein dy-

namics using advanced microscopic assays, *in vitro* reconstitution of dynamic cytoskeleton-based processes and different methods of identification of protein-protein interactions.

Anna Akhmanova is a recipient of the ALW Vernieuwingsimpuls VIDI (2001) and VICI awards (2007) (Netherlands Organization for Scientific Research), and an ERC Synergy grant (2013). She is the Chair of the board of the Netherlands Microscopy Society and a member of the European Molecular Biology Organization (EMBO).



Yosef (Yossi) Yarden is a Professor in the Department of Biological Regulation at Weizmann Institute of Science (Rehovot, Israel). Dr. Yarden obtained his PhD from Weizmann Institute (under the mentoring of Dr. Joseph Schlessinger) and later trained at both Genentech Inc., with Dr. Axel Ullrich, and at the Whitehead Institute (MIT), with Dr. Robert A. Weinberg. His group studies the roles played by growth factors and their receptors in cancer progression. They also explore strategies able to intercept growth factor signals in tumors.



Jason D Weber is an Associate Professor in the Departments of Internal Medicine and Cell Biology and Physiology at Washington University and is the Coleader of the Breast Cancer Research Program in the Siteman Comprehensive Cancer Center. Dr. Weber obtained his PhD from Saint Louis University and received postdoctoral training under the mentoring of Dr. Charles J. Sherr at St. Jude Children's Research Hospital in Memphis, TN. His group studies the molecular interplay between oncogenes and tumor suppressors, particularly focused on their role in regulating cellular growth processes.



Michael Dustin received his PhD in 1990 in Cell and Developmental Biology from Harvard University, where he worked in the laboratory of Timothy A. Springer, PhD. During his graduate work he was involved in the identification and cloning of ICAM-1 and ICAM-2, key ligands for the integrin LFA-1. He further demonstrated that LFA-1 dependent adhesion was stimulated by T cell receptor signaling. He did his postdoctoral training with Stuart Kornfeld at Washington University School of Medicine in St. Louis, focusing on mechanisms of lysosome biogenesis. He started his independent lab also at Washington University School of Medicine, and used supported lipid bilayers to define the first protein to modulate organization of the immunological synapse- CD2 associated protein- and the dynamics of immunological synapse formation. He moved to NYU School of Medicine in 2001 where he applied *in vivo* microscopy to tolerance induction and immune responses in effector sites including the liver, brain, and spleen. Continued work with the immunological synapse model resulted in discovery of signaling microclusters, the molecular basis of immunological synapse stability, and the direct budding of extracellular microvesicle enriched in T cell receptors in the immunological synapse. Professor Dustin recently took a position at the

University of Oxford with support of a Wellcome Trust Principal Research Fellowship. The focus of his new post is the translation of the immunological synapse. He received a Presidential Early Career Award in Science and Engineering and the DART-NYU Biotechnology Achievement Award.



David Sassoon directs a research team at Paris Sorbonne. He received his PhD from Columbia University (NYC, USA) in 1986 in the biological sciences and was a professor at Boston University Medical School and Mt. Sinai Medical School before relocating to Paris in 2006. His long-standing interests are in developmental and stem cell biology in a number of tissues including skeletal muscle, skin, CNS, and heart. Dr. Sassoon recently concluded the coordination of a EC-funded 15 partner international consortium (Endostem) designed to identify novel therapeutics for mobilizing endogenous stem/progenitor cells (<http://www.endostem.eu/>) and is a recent recipient as coordinator of a multipartner Transatlantic Network of Excellence grant from the Fondation Leducq focused on cardiac stem cell biology (<http://www.fondationleducq.org/nivel2.aspx?idsec=1360>).

A major focus of his research is on adult progenitor/stem cell biology and how these cells respond to stress. His team has identified a parentally imprinted gene, PW1/Peg3, which is involved in both p53 and inflammatory responses. PW1 is expressed in adult stem cells in all tissues identified to date. Loss of PW1 function in stem cells results in a loss of stem cell competence including a reduced capacity to undergo self-renewal and respond to hypoxic stress. We are currently evaluating a role for PW1 in postnatal and adult heart with a particular focus on its role in vessel precursors using both myocardial infarction and stress-induced myocardial hypertrophy.



Jason M Haugh is a Professor of Chemical and Biomolecular Engineering at North Carolina State University in Raleigh, NC, USA, where he has been on the faculty since 2000. His laboratory has been among those to pioneer the synthesis of quantitative experiments and modeling to study signal transduction in mammalian cells. Since the lab's inception, their approach has combined biochemical measurements, live-cell fluorescence microscopy, and computational modeling to elucidate signaling mechanisms by analyzing their kinetics and spatial organization in cells. The systems studied by the Haugh Laboratory include: regulation of the phosphoinositide 3-kinase, Ras/extracellular signal-regulated kinase, and phospholipase C pathways mediated by receptor tyrosine kinases, and cross talk between these pathways; dynamic organization of multimolecular complexes at cell membranes; signaling mediated by cytokine and chemokine receptors in immune cells; and integration of adhesion, signaling, and cytoskeletal dynamics that direct cell migration.



After graduating with an MA in Mathematics from the University of Cambridge, Helen Mary Byrne received her Master of Science and DPhil from the University of Oxford. She worked as a postdoctoral researcher at the Universities of Oxford and Bath, before taking up a lectureship at the UMIST in Manchester. She moved to the University of Nottingham in 1998 and was awarded a prestigious Advanced Research Fellowship from the United Kingdom's Engineering and Physical Sciences Research Council (2000–2006). She was promoted to Reader in 2002 and Professor in 2003. While in Nottingham, she established and then led the Centre for Mathematical Medicine and Biology until she returned to Oxford in 2011 where she is now based in the Mathematical Institute at the University of Oxford. She has over 20 years' experience of developing, analyzing, and simulating continuum and multiscale models of biomedical systems, and a particular interest in studying the growth and response to treatment of solid tumors.



Rune Linding completed his PhD at the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany, followed by postdoctoral training at EMBL. He then jointly trained with professors Tony Pawson and Mike Yaffe at the Lunenfeld at Mount Sinai Hospital in Toronto, Canada, and the Massachusetts Institute of Technology (MIT) in Cambridge, US, respectively. Dr. Linding then established his own laboratory of Cellular and Molecular Logic at the Institute of Cancer Research (ICR) in London, UK, before returning to Denmark to take a position as professor of cellular signal integration at the Technical University of Denmark. In 2014, Dr. Linding moved his laboratory to the Biotech Research and Innovation Centre (BRIC) at University of Copenhagen where he is currently professor of cellular signaling. His research group focuses on big data network biology, exploring biological systems by developing and deploying algorithms aimed to predict cell behavior, in particular looking at cellular signal processing and decision making. A strategic focus is to continue to develop computational tools (such as KinomeExplorer, NetworKIN, and NetPhorest) and to deploy these on genome-scale quantitative data obtained by, for example, mass spectrometry, genomic, and phenotypic screens to understand the principles of how spatio and temporal assembly of mammalian signaling networks transmit and process information at a systems level in order to alter cell behavior. His overarching aim is to advance network

medicine by identifying and targeting signaling networks associated with complex diseases. To this end Dr. Linding is currently leading high-level, strategic, multidisciplinary studies of signaling network dynamics driving cancer metastasis in collaboration with other labs at Harvard, Yale, The Jackson Laboratory, Memorial Sloan Kettering Cancer Center, MIT, and BRIC.

CONTRIBUTORS TO VOLUME 2

- A Aballay
Duke University, Durham, NC, USA
- BR Acharya
The University of Queensland, Brisbane, QLD, Australia
- RC Aguilar
Purdue University Center for Cancer Research, West Lafayette, IN, USA
- A Aguilera Gomez
Hubrecht Institute – KNAW and the UMC Utrecht, Utrecht, the Netherlands
- A Akhmanova
Utrecht University, Kruytgebouw, Utrecht, The Netherlands
- J Alanko
Turku Centre for Biotechnology, Turku, Finland
- L-AH Allen
University of Iowa, Coralville, IA, USA
- TD Allen
University of Manchester, Manchester, UK
- LA Amos
MRC Laboratory of Molecular Biology, Cambridge, UK
- E van Anken
Vita Salute San Raffaele University, Milan, Italy
- S Asrat
Tufts University School of Medicine, Boston, MA, USA
- D Axelrod
University of Michigan, Ann Arbor, MI, USA
- T Balla
Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD, USA
- M Barisic
Instituto de Biologia Molecular e Celular, Porto, Portugal
- W Becker
Becker & Hickl GmbH, Berlin, Germany
- I Bernascone
Centro de Biologia Molecular Severo Ochoa (CSIC-UAM), Madrid, Spain
- D Bliss
National Institutes of Health, Bethesda, MD, USA
- M Bornens
Institut Curie, Paris, France
- AD Bradshaw
Medical University of South Carolina, Charleston, SC, USA
- FM Brodsky
University College London, London, UK
- D Brough
University of Manchester, Manchester, UK
- C Bucci
University of Salento, Lecce, Italy
- B Cadot
Sorbonne Universités UPMC, Paris Cedex, France
- J Cancino
Universidad Andrés Bello, Viña del Mar, Chile
- A Capalbo
Institute of Protein Biochemistry, Naples, Italy
- S Caplan
The University of Nebraska Medical Center, Omaha, NE, USA
- M-F Carlier
Cytoskeleton Dynamics and Motility Group, CNRS, Gif-sur-Yvette, France
- RA Cerione
Cornell University, Ithaca, NY, USA
- CEJ Cheetham
Carnegie Mellon University, Pittsburgh, PA, USA
- W Chiu
Baylor College of Medicine, Houston, TX, USA
- ST Christensen
University of Copenhagen, Universitetsparken, Copenhagen OE, Denmark
- CA Cypranowska
University of California, Berkeley, CA, USA
- M Dahan
Institut Curie, Paris, France
- L Dolat
Drexel University, Philadelphia, PA, USA
- LA Earl
National Institutes of Health, Bethesda, MD, USA
- M El Beheiry
Institut Curie, Paris, France

EL Elson
Washington University School of Medicine, St. Louis, MO, USA

L Feo-Lucas
Universidad Autónoma de Madrid School of Medicine, Madrid, Spain

ML Ferrari
Institut Pasteur, Paris Cédex, France

M Fukuda
Tohoku University, Sendai, Miyagi, Japan

T Galli
Institut Jacques Monod, Paris, France; INSERM ERL U950, Paris, France; and Sorbonne Paris Cité, Paris, France

M Galvez-Santisteban
Centro de Biología Molecular Severo Ochoa (CSIC-UAM), Madrid, Spain

B Geiger
Weizmann Institute of Science, Rehovot, Israel

D Ghosal
MRC Laboratory of Molecular Biology, Cambridge, UK

S Giese
University College London, London, UK

PA Gleeson
The University of Melbourne, Melbourne, VIC, Australia

MW Goldberg
Durham University, Durham, UK

ER Gomes
Sorbonne Universités UPMC, Paris Cedex, France and Universidade de Lisboa, Lisbon, Portugal

KV Gorzelnik
Texas A&M University, College Station, TX, USA

B Goud
Institut Curie, Paris, France

J Gruenberg
University of Geneva, Geneva, Switzerland

R Gurung
Monash University, Clayton, VIC, Australia

H Hamidi
Turku Centre for Biotechnology, Turku, Finland

GRV Hammond
University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

PI Hanson
Washington University School of Medicine, St. Louis, MO, USA

LM Hartnell
National Institutes of Health, Bethesda, MD, USA

SM Heissler
National Institutes of Health, Bethesda, MD, USA

A Helenius
Institute for Biochemistry, Zürich, Switzerland

AD Hempstead
Tufts University School of Medicine, Boston, MA, USA

ST Hess
University of Maine, Orono, ME, USA

N Hirokawa
University of Tokyo, Tokyo, Japan and King Abdulaziz University, Jeddah, Saudi Arabia

W-C Hsieh
Purdue University Center for Cancer Research, West Lafayette, IN, USA

N Imamoto
RIKEN, Saitama, Japan

EY Isacoff
University of California Berkeley, Berkeley, CA, USA

RR Isberg
Tufts University School of Medicine, Boston, MA, USA

T Ishikawa
Paul Scherrer Institute, Villigen, Switzerland

J Ivaska
Turku Centre for Biotechnology, Turku, Finland

CE Jackson
Washington University School of Medicine, St. Louis, MO, USA

DE James
The University of Sydney, NSW, Australia

DA de Jesús
Tufts University School of Medicine, Boston, MA, USA

L Johannes
Institut Curie, Paris, France

J Kaplan
University of Utah School of Medicine, Salt Lake City, UT, USA

RJ Kaufman
Sanford Burnham Medical Research Institute, La Jolla, CA, USA

O Koren
Bar-Ilan University, Safed, Israel

PB Lazarow
Institut Pasteur, Paris, France

G Legname
*Scuola Internazionale Superiore di Studi Avanzati,
Trieste, Italy*

RE Leube
RWTH Aachen University, Aachen, Germany

G Li
*University of Oklahoma Health Sciences Center,
Oklahoma City, OK, USA*

M Li
*University of California Los Angeles, Los Angeles, CA,
USA*

AJ Lindsay
University College Cork, Corcaigh, Ireland

A Livne
Weizmann Institute of Science, Rehovot, Israel

J Löwe
MRC Laboratory of Molecular Biology, Cambridge, UK

R Lu
*University of California Los Angeles, Los Angeles, CA,
USA*

A Luini
Institute of Protein Biochemistry (IBP), Naples, Italy

MJ Lukey
Cornell University, Ithaca, NY, USA

K Madhivanan
*Purdue University Center for Cancer Research, West
Lafayette, IN, USA*

H Maiato
*Instituto de Biologia Molecular e Celular, Porto,
Portugal*

LL Marcotte
Assumption College, Worcester, MA, USA

M Marie
University of Bergen, Bergen, Norway

M Marsh
University College London, London, UK

F Martin-Belmonte
*Centro de Biología Molecular Severo Ochoa (CSIC-
UAM), Madrid, Spain*

M Mauthe
*University Medical Center Utrecht, Utrecht, The
Netherlands*

MW McCaffrey
University College Cork, Corcaigh, Ireland

V Mennella
University of Toronto, Toronto, ON, USA

K Mihara
Kyushu University, Fukuoka, Japan

EA Miller
Columbia University, New York, NY, USA

Y Mimura
RIKEN, Saitama, Japan

CA Mitchell
Monash University, Clayton, VIC, Australia

MJ Mlodzianoski
University of Maine, Orono, ME, USA

W Möbius
*Max-Planck-Institute of Experimental Medicine,
Göttingen, Germany*

P Montaville
*Cytoskeleton Dynamics and Motility Group, CNRS, Gif-
sur-Yvette, France*

C Montecucco
CNR Institute of Neuroscience, Padua, Italy

A Moran
National Institutes of Health, Bethesda, MD, USA

H-M Müller
*Heidelberg University Biochemistry Center, Heidelberg,
Germany*

M Munson
*University of Massachusetts Medical School, Worcester,
MA, USA*

RZ Murray
*Queensland University of Technology, Brisbane, QLD,
Australia*

CL Netherton
Institute for Animal Health, Pirbright, Woking, UK

H Neuman
Bar-Ilan University, Safed, Israel

W Nickel
*Heidelberg University Biochemistry Center, Heidelberg,
Germany*

F Niedergang
*Institut Cochin, Paris, France; Inserm, CNRS, Paris,
France; and Université Paris Descartes, Paris, France*

M Oka
*National Institute of Biomedical Innovation (NIBIO),
Osaka, Japan*

A Ortega-Carrion
*Universidad Autónoma de Madrid School of Medicine,
Madrid, Spain*

H Otera
Kyushu University, Fukuoka, Japan

S Pagant
Columbia University, New York, NY, USA

S Parashuraman
Institute of Protein Biochemistry (IBP), Naples, Italy

RG Parton
The University of Queensland, Brisbane, QLD, Australia

LB Pedersen
*University of Copenhagen, Universitetsparken,
Copenhagen OE, Denmark*

A Pelchen-Matthews
University College London, London, UK

A Penno
University of Bonn, Bonn, Germany

J Pernier
Cytoskeleton Dynamics and Motility Group, CNRS, Gif-sur-Yvette, France

RC Piper
University of Iowa, Iowa City, IA, USA

KE Pischke
*Scuola Internazionale Superiore di Studi Avanzati,
Trieste, Italy*

L Popolo
*Università degli Studi di Milano, Via Celoria, Milano,
Italy*

PP Powell
University of East Anglia, Norwich, UK

C Rabouille
*Hubrecht Institute – KNAW and the UMC Utrecht,
Utrecht, the Netherlands and University Medical Center
Utrecht, Utrecht, the Netherlands*

V Ramabhadran
Tufts University School of Medicine, Boston, MA, USA

G Raposo
Institut Curie, Paris, France

F Reggiori
*University Medical Center Utrecht, Utrecht, The
Netherlands*

A Reiner
*University of California Berkeley, Berkeley, CA, USA
and Ruhr University Bochum, Bochum, Germany*

M Rigoni
CNR Institute of Neuroscience, Padua, Italy

PJ Sansonetti
*Institut Pasteur, Paris Cédex, France and Collège de
France, Paris, France*

J Saraste
University of Bergen, Bergen, Norway

F Schmid
*University of Copenhagen, Universitetsparken,
Copenhagen OE, Denmark*

MF Schmid
Baylor College of Medicine, Houston, TX, USA

N Schwarz
RWTH Aachen University, Aachen, Germany

C Scott
University of Geneva, Geneva, Switzerland

MNJ Seaman
University of Cambridge, Cambridge, UK

JR Sellers
National Institutes of Health, Bethesda, MD, USA

M Sharma
*Indian Institute of Science Education and Research
Mohali, Manauli, Punjab, India*

DA Sheffield
Monash University, Clayton, VIC, Australia

R Shrestha
*University of Utah School of Medicine, Salt Lake City,
UT, USA*

J Sillibourne
*Commissariat à l'Energie Atomique et aux Energies
Alternatives (CEA), Grenoble, France*

V Singh
Indian Institute of Science, Bangalore, India

R Sitia
Vita Salute San Raffele University, Milan, Italy

P van der Sluijs
UMC Utrecht, Utrecht, The Netherlands

ET Spiliotis
Drexel University, Philadelphia, PA, USA

PD Stahl
*Washington University School of Medicine, St. Louis,
MO, USA*

P Stanley
Albert Einstein College Medicine, New York, NY, USA

M Stasi
University of Salento, Lecce, Italy

E Stefani
University of California Los Angeles, Los Angeles, CA, USA

JP Steringer
Heidelberg University Biochemistry Center, Heidelberg, Germany

J Stöckli
The University of Sydney, NSW, Australia

JL Stow
The University of Queensland, Brisbane, QLD, Australia

A Straube
University of Warwick, Coventry, UK

S Subramaniam
National Institutes of Health, Bethesda, MD, USA

J Sun
Washington State University, Spokane, WA, USA

TM Svitkina
University of Pennsylvania, Philadelphia, PA, USA

JA Swanson
University of Michigan Medical School, Ann Arbor, MI, USA

Y Tanaka
University of Tokyo, Tokyo, Japan

D Tareste
Institut Jacques Monod, Paris, France; INSERM ERL U950, Paris, France; and Sorbonne Paris Cité, Paris, France

KG Ten Hagen
Bethesda, MD, USA

C Thiele
University of Bonn, Bonn, Germany

L Toro
University of California Los Angeles, Los Angeles, CA, USA

LM Traub
University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

JC Trinidad
Indiana University, Bloomington, IN, USA

F Vacca
University of Geneva, Geneva, Switzerland

MM Valles
University of Maine, Orono, ME, USA

M Vicente-Manzanares
Universidad Autónoma de Madrid School of Medicine, Madrid, Spain

DM Ward
University of Utah School of Medicine, Salt Lake City, UT, USA

MHM Wauben
Utrecht University, Utrecht, The Netherlands

R Weigert
National Institutes of Health, Bethesda, MD, USA

T Wileman
University of East Anglia, Norwich, UK

RR Wise
University of Wisconsin Oshkosh, Oshkosh, WI, USA

X Wu
University of California Los Angeles, Los Angeles, CA, USA

Y Wu
University of California Los Angeles, Los Angeles, CA, USA

C Wunder
Institut Curie, Paris, France

Y Yamauchi
Institute for Biochemistry, Zürich, Switzerland

AS Yap
The University of Queensland, Brisbane, QLD, Australia

A Yildiz
University of California, Berkeley, CA, USA

Y Yoneda
National Institute of Biomedical Innovation (NIBIO), Osaka, Japan

S Yoshida
University of Michigan Medical School, Ann Arbor, MI, USA

T Yoshimori
Osaka University, Suita, Osaka, Japan

J Zhang
Texas A&M University, College Station, TX, USA

L Zhang
Bethesda, MD, USA

CONTENTS OF VOLUME 2

Preface	xxv
Cell Biology: An Overview <i>RA Bradshaw and PD Stahl</i>	xxvii

VOLUME 2 Organizational Cell Biology

Organizational Cell Biology: An Overview <i>EL Elson, G Raposo, PA Gleeson, A Akhmanova, T Yoshimori, PD Stahl, and B Goud</i>	1
--	---

Imaging the Cell

Electron Microscopy

Single-Particle CryoEM of Macromolecular Complexes <i>J Zhang, KV Gorzelnik, MF Schmid, and W Chiu</i>	5
Scanning Electron Microscopy in Cell Biology <i>TD Allen and MW Goldberg</i>	14
Electron Tomography <i>T Ishikawa</i>	22
Immunoelectron Microscopy: High-Resolution Immunocytochemistry <i>W Möbius</i>	32
Imaging Cellular Architecture with 3D SEM <i>LM Hartnell, LA Earl, D Bliss, A Moran, and S Subramaniam</i>	44

Light Microscopy

Genetically Encoded Fluorescent Probes and Live Cell Imaging <i>GRV Hammond, CEJ Cheetham, and T Balla</i>	51
Total Internal Reflection Fluorescence Microscopy <i>D Axelrod</i>	62
Super Resolution Fluorescence Localization Microscopy <i>MJ Mlodzianoski, MM Valles, and ST Hess</i>	70
Super-Resolution Light Microscopy: Stimulated Emission Depletion and Ground-State Depletion <i>Y Wu, X Wu, R Lu, M Li, L Toro, and E Stefani</i>	76
Structured Illumination Microscopy <i>V Mennella</i>	86
Fluorescence Correlation Spectroscopy: A Tool for Measuring Dynamic and Equilibrium Properties of Molecules in Cells <i>EL Elson</i>	99
Fluorescence Lifetime Imaging – Applications and Instrumental Principles <i>W Becker</i>	107
High-Speed Localization Microscopy and Single-Particle Tracking <i>M El Beheiry and M Dahan</i>	121
Intravital Microscopy in Mammalian Organisms: From Tissue Physiology to Cell Biology <i>R Weigert</i>	129
Optogenetics <i>EY Isacoff and A Reiner</i>	147

Organelles

Structure and Function

The Endoplasmic Reticulum	<i>E van Anken and R Sitia</i>	156
Intermediate Compartment: A Sorting Station between the Endoplasmic Reticulum and the Golgi Apparatus	<i>J Saraste and M Marie</i>	168
Golgi and TGN	<i>A Luini and S Parashuraman</i>	183
Early Endosomal Compartments	<i>P van der Sluijs</i>	192
The Late Endosome	<i>F Vacca, C Scott, and J Gruenberg</i>	201
Signaling from Endosomes	<i>J Alanko, H Hamidi, and J Ivaska</i>	211
Conventional and Secretory Lysosomes	<i>R Shrestha, J Kaplan, and DM Ward</i>	225
Lysosome-Related Organelles	<i>M Fukuda</i>	235
At the Center of Autophagy: Autophagosomes	<i>F Reggiori and M Mauthe</i>	243
Peroxisomes	<i>PB Lazarow</i>	248
Lipid Droplets	<i>C Thiele and A Penno</i>	273
Mechanisms and Functions of Mitochondrial Dynamics	<i>K Mihara and H Otera</i>	279
Synaptosomes and Synaptic Vesicles	<i>JC Trinidad</i>	293
Extracellular Vesicles	<i>MHM Wauben</i>	302
Nuclear Organization (Nuclear Structure and Dynamics)	<i>Y Mimura and N Imamoto</i>	311
Nuclear Pores	<i>M Oka and Y Yoneda</i>	319
Plastids: The Anabolic Factories of Plant Cells	<i>RR Wise</i>	324

Interorganellar Communication

Interplay and Processes

ER–Golgi Transport	<i>S Pagant and EA Miller</i>	331
N-Linked Glycans (N-Glycans)	<i>P Stanley</i>	339
Extracellular O-Glycans	<i>KG Ten Hagen and L Zhang</i>	347
Intra-Golgi Transport	<i>A Aguilera Gomez and C Rabouille</i>	354
Post-Golgi Transport – Cargo, Carriers, and Pathways	<i>JL Stow and RZ Murray</i>	363
Regulation of the Secretory Pathway	<i>A Capalbo, J Cancino, and A Luini</i>	371
Regulated versus Constitutive Secretion – A Major Form of Intercellular Communication	<i>J Stöckli, RC Piper, and DE James</i>	376
Clathrin and Clathrin-Dependent Endocytosis	<i>FM Brodsky</i>	384
Clathrin Independent Endocytosis	<i>RG Parton</i>	394
Rabs of the Endosomal Recycling Pathway	<i>AJ Lindsay and MW McCaffrey</i>	401
Endosome to Lysosome Transport	<i>C Bucci and M Stasi</i>	408
Endocytosis of Cargo Proteins: LDL	<i>LM Traub</i>	418
Retrograde Transport	<i>L Johannes and C Wunder</i>	433

Rabs and Other G Proteins	<i>G Li</i>	442
Components		
Adaptor Proteins: Inter-Organelle Traffic Controllers	<i>K Madhivanan, W-C Hsieh, and RC Aguilar</i>	450
SNAREs: Membrane Fusion and Beyond	<i>T Galli and D Tareste</i>	459
ESCRTing around the Cell	<i>PI Hanson and CE Jackson</i>	466
The Retromer Complex	<i>MNJ Seaman</i>	475
Vesicle Tethers	<i>LL Marcotte and M Munson</i>	485
BAR Domains and BAR Domain Superfamily Proteins	<i>M Sharma and S Caplan</i>	491
Interplay and Processes		
Role of Phosphoinositides in Membrane Traffic	<i>DA Sheffield, R Gurung, and CA Mitchell</i>	503
Unconventional Protein Secretion: Fibroblast Growth Factor 2 and Interleukin-1 β as Examples	<i>W Nickel, JP Steringer, H-M Müller, and D Brough</i>	520
Endoplasmic Reticulum Stress in Disease	<i>RJ Kaufman and L Popolo</i>	528
Cytoskeleton and Motors		
Cytoskeletal Components		
Microtubules and Microtubule-Associated Proteins (MAPs)	<i>A Straube</i>	539
Actin Assembly Dynamics and Its Regulation in Motile and Morphogenetic Processes	<i>J Pernier, P Montaville, and M-F Carlier</i>	548
Intermediate Filaments	<i>RE Leube and N Schwarz</i>	569
Septins: Cytoskeletal Filaments with Structural and Regulatory Functions	<i>ET Spiliotis and L Dolat</i>	579
Bacterial and Archaeal Cytoskeletons	<i>D Ghosal, J Löwe, and LA Amos</i>	588
Myosins	<i>SM Heissler and JR Sellers</i>	597
Kinesin Superfamily Proteins (KIFs) as a Fundamental Component of Life: Intracellular Transport and Beyond	<i>Y Tanaka and N Hirokawa</i>	608
Dyneins	<i>CA Cypranowska, A Yildiz, and T Ishikawa</i>	620
Complex Cytoskeletal Structures		
The Mitotic Spindle	<i>M Barisic and H Maiato</i>	637
Centrioles and the Centrosome	<i>J Sillibourne and M Bornens</i>	649
Cilia and Flagella	<i>F Schmid, ST Christensen, and LB Pedersen</i>	660
Skeletal Muscle	<i>B Cadot and ER Gomes</i>	677
Filopodia and Lamellipodia	<i>TM Svitkina</i>	683
The Extracellular Matrix	<i>AD Bradshaw</i>	694
Cell–Cell Adhesion and the Cytoskeleton	<i>BR Acharya and AS Yap</i>	704
Cell Adhesion to the Extracellular Matrix	<i>B Geiger and A Livne</i>	713

Cell Migration	<i>A Ortega-Carrion, L Feo-Lucas, and M Vicente-Manzanares</i>	720
Rho GTPases	<i>MJ Lukey and RA Cerione</i>	731
Cell Polarity	<i>F Martin-Belmonte, I Bernascone, and M Galvez-Santisteban</i>	741

Intracellular Infectiology

Cell Processes

Phagocytosis	<i>F Niedergang</i>	751
Macropinocytosis	<i>JA Swanson and S Yoshida</i>	758
Microbicidal Mechanisms	<i>L-AH Allen</i>	766
Bacterial Subversion of Phagocytic Killing	<i>RR Isberg, S Asrat, DA de Jesús, AD Hempstead, and V Ramabhadran</i>	776
Cellular Invasion by Bacterial Pathogens	<i>ML Ferrari and PJ Sansonetti</i>	784

Infectious Agents

Bacterial Protein Toxins as Tools in Cell Biology and Physiology	<i>M Rigoni and C Montecucco</i>	794
The Gut Microbiome	<i>H Neuman and O Koren</i>	799
Cell Biology of Virus Infection	<i>A Helenius and Y Yamauchi</i>	809
Virus Factories and Mini-Organelles Generated for Virus Replication	<i>T Wileman, CL Netherton, and PP Powell</i>	819
HIV – The Cell Biology of Virus Infection and Replication	<i>S Giese, A Pelchen-Matthews, and M Marsh</i>	828
Prions	<i>G Legname and KE Pischke</i>	839

Cell Processes

Cellular Responses to Infections in <i>Caenorhabditis elegans</i>	<i>J Sun, A Aballay, and V Singh</i>	845
---	--------------------------------------	-----

PREFACE

Cell biology is the study of cells, the integral unit that is the basis of all living tissues. In its earliest phases, this area of investigation was largely defined by microscopic-based observations of structure and organization that were, by their nature, largely descriptive. This also served to delineate cell biology from its nearest neighbor, biochemistry, which, dating from its post WWII renaissance, was driven in large part by reductionist approaches, where tissues, cells, and organelles were broken down in order to isolate and then characterize individual components. The melding force that bridged these two areas was molecular biology that allowed the cell biologist to transform pictures into functions and allowed the biochemist to reassemble their components into ensembles and subcellular entities. Not surprisingly the boundary between these two key biological sciences has become blurred beyond recognition. Today a considerable part of what passes as cell biology is by any other name biochemistry (or molecular biology – the distinction between these two areas being equally indistinct) and vice versa. Indeed, one of the first challenges we faced in organizing this Encyclopedia was how much molecular detail to include. It was our conclusion that for the sake of completeness, and to appropriately introduce the sections Organizational Cell Biology and Functional Cell Biology, we needed a description of the molecular components of the cell, particularly with respect to the synthesis and degradation of proteins and nucleic acids, along with outlines of important chemical principles and key methodology. We were guided to some degree in these decisions by input from undergraduate and graduate teachers and the contents of their introductory courses that they shared with us and by discussions with our section and subsection editors. The result was the first section, which we recognize is very 'biochemical,' but certainly lacking the more detailed coverage of the *Encyclopedia of Biological Chemistry* (Lennarz and Lane, 2004) or the many excellent biochemical texts that dot the educational landscape. Readers who find 'gaps' or incomplete treatments in this section will likely be able to find the additional detail they seek from these sources.

The second major decision that we had to confront was whether we would attempt to cover the full range of living organisms or whether we would place some limits on the scope of the material we included. In the final result, it was decided to place the emphasis on higher eukaryotes with only very prescribed treatments of lower eukaryotes and prokaryotes. Sections 2 and 3 thus deal with material that has perhaps been more traditionally thought of as cell biology. Again we were guided in our decisions about what to include by teaching considerations, brainstorming sessions with our editors, and a perceived need to separate cellular organization from cellular function. However, as even a quick perusal will reveal, this is not a sharp delineation either. Section 2 begins with a methodological treatment of imaging, which is basically divided between light and electron microscopic applications, and is followed by sections on organelles, interorganellar communication, and the cytoskeleton (including motors). The

last part of section 2 is devoted to intracellular infection and covers a variety of external agents and pathogens that impact cell function as well as human pathology. Section 3, in contrast, largely deals with major cell functions: signaling, cell-cycle control, and apoptosis. The last two parts are devoted to cells of the immune system and their responses and stem cells – two highly specialized niches characterized by unique cellular involvement.

The last part of the Encyclopedia deals with systems biology. This is really the newest area of cell biology and considers cellular involvement as part of a phenotypic response. As such, it strives to integrate all of the levels of the first three sections in a consistent manner to produce a seamless description of cellular function. This is clearly a newly evolving area and the most rapidly changing part of cell biology; indeed it is where we expect to see the greatest change in the next few years.

Assembling a treatment of a topic as large as cell biology had multiple challenges. Coverage was a significant problem: on the one hand it was almost impossible to avoid some redundancy in places while, on the other, there were inevitable gaps. Some of these arose from late cancellations; others from oversights on our part. We can only promise to fill these in future editions. We also note that as can be expected for a large multiauthor compilation the individual articles do differ in detail and treatment. We felt it was more important to allow our experts substantial latitude in deciding how to present their topics than to apply rigid guidelines. We did try to limit the number of references (to make it easier for people not familiar with a topic to make the transition to more extensive articles) but there is admittedly some considerable variation in the bibliographic material as well.

When we were well into the project, we were approached by Graham Nisbet (Elsevier) about including the Encyclopedia in a larger collection entitled the Reference Module in Biomedical Sciences. The Reference Module plan was grouped in single integrated work information from several other compendia and it was our view that this would greatly leverage the value of our efforts (and those of our contributors). While we were the 'new kid on the block' in that we were still compiling the Encyclopedia and the other collections to be included were already extant, we felt that it was too good an opportunity to pass up. The Reference Module has since been launched and although it does not yet contain the complete content from the Encyclopedia, this will be achieved in the not too distant future. We consider that this will be a particularly valuable resource for researchers that are just entering (or changing into) this field.

No work of this size could be completed without the help and advice of many people. In this regard we would like to thank a number of people who were instrumental in bringing this project to fruition. First and foremost we would like to recognize our section editors: Jerry Hart, Graça Raposo, Bruno Goud, Alan Ezekowitz, and Doug Lauffenburger; and our subsection editors: Ken Neet, Sarah Woodson, Judy Bond,

John Heuser, Elliot Elson, Tamotsu Yoshimori, Paul Gleeson, Anna Akhmanova, Yosef Yarden, Jason Weber, Michael Dustin, David Sassoon, Jason Haugh, Helen Byrne, and Rune Linding; and thank them for their invaluable input, enthusiasm, and hard work. They are truly a global group and have been a great team to work with. We also wish to thank the staff at Elsevier including Janice Audet, Peter Labella, Nicky Carter, Karen Maloney, Erin Hill-Parks, Rashmi Phadnis, Blerina Osmanaj, and Ginny Mills for their many efforts, perseverance (particularly with our idiosyncrasies), and skillful management of every aspect of this project.

In the course of planning this work, we gathered information from a number of sources. At the outset the publishers polled some 167 faculty from both undergraduate and graduate schools to ascertain their perceived interest in such a project, which also garnered considerable input about potential topics. The response was highly favorable. We also talked with many colleagues about experiences at their institutions and would like to particularly thank Catherine Bevier (Colby College), Robert Simoni (Stanford University), Larry Marsh (UC Irvine), and John Cooper (Washington University) for providing syllabi of germane courses taught in their institutions. Special thanks also go to Ruedi Abersold (ETH, Zurich) and Sergio Grinstein (University of Toronto) for their expertise that they readily shared with us, which was invaluable in planning certain sections.

Finally we would like to recognize all of our authors. We have been extraordinarily fortunate in attracting individuals from all around the globe to take time from their busy schedules to prepare this splendid set of contributions. Without these efforts there would be no Encyclopedia. Under ordinary circumstances it would be appropriate to dedicate the work to them and all the colleagues whose works they wrote about. However, during the preparation of this work, we sadly lost one of the most important contributors to cell biology, Tony Pawson. Tony was a true pioneer in elucidating basic mechanisms of how cells transmit information intracellularly, a fundamental knowledge that permeates all of cell biology. In an effort that was initiated by one of the subsection editors, Rune Linding, it was decided to dedicate the entire work to his memory as recognition of his contributions and a lasting tribute to a man who left us too soon.

We hope that the Encyclopedia will be of value to students and researchers alike and will become a useful tool in the training of future generations of scientists.

Ralph A Bradshaw and Philip D Stahl

References

Lennarz, W.J., Lane, M.D. (Eds.), 2004. Encyclopedia of Biological Chemistry. Oxford: Elsevier.

Cell Biology: An Overview

RA Bradshaw, University of California, Irvine, CA, USA

PD Stahl, Washington University School of Medicine, St. Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

All living organisms can be divided into three principal categories: archaeobacteria, eubacteria, and eukaryotes; although they differ in structure and organization, they are all composed of cells as the fundamental life unit. At the molecular level, there is also a great deal of similarity in the basic materials that make up these entities because they use the same kinds of molecules to store and reproduce information, to run the cellular metabolism and machinery and to provide the structural framework. Thus nucleic acids, proteins, lipid membranes, and carbohydrates – alone and in various combinations – are universally present, albeit in distinguishable forms, along with innumerable metabolites and ions. There are components that are apparently essential for life and are found in one form or another in all species and there are many unique moieties and associated activities that are highly specialized and are found in relatively few organisms. Indeed, the similarities have underpinned the development of our understanding of cellular function at a rudimentary level and the differences, basically engendered by evolution, have illustrated and delineated the complexity that speciation has introduced. Perhaps the largest of these differences is that which separates single cell organisms from multicellular organisms. The latter are exclusively eukaryotes while the former are composed of both eukaryotes and prokaryotes. The cellular organization and architecture that distinguishes these two major life forms is striking; although cell biology correctly embraces both, traditionally prokaryotic organisms have been the province of the microbiologists and the majority of cell biology research has been devoted to the eukaryotic world. In practical terms this translates for the most part into the study of human cells and those of easily maintained laboratory animals and selected paradigms, for example, fruit flies, worms, and zebra fish.

Human and animal cell biology is not a tightly proscribed science with well-defined borders. Basically it serves at the interface between biochemistry, molecular biology, and genetics, on the one hand, and anatomy and physiology, on the other. The continuum of these disciplines forms the core of the biomedical sciences, which also include the related but separate fields of pharmacology, microbiology, immunology, and pathology that provide the connections to disease and health. Cell biology has strong connections to all of these. There are also specialized areas, for example, neuroscience, that are of such importance that they warrant their own category and the cell biology associated with them is also highly specialized. Thus, cell biology is as complex as the enormous variety of cells that exist and achieving an accurate description of all of them in terms of their components and functions has long been a major part of the research in this field.

Imaging and Organelle Organization

Among the developments that propelled cell biology into the modern era are the introduction of the ultracentrifuge and

the rapid advances in microscopy both electron and light microscopy – the former allowing investigators to fractionate and characterize the components of the cell and the latter to literally see them – either *in situ* or in isolated form. This progress is illustrated by a series of Nobel Prizes starting in the early 1970s that chronicle the grand confluence of classical biochemistry and general physiology that created modern cell biology (Claude *et al.*, 1974) or the discovery and characterization of intracellular organelles – lysosomes and endoplasmic reticulum (ER) among others (Mitchell, 1978), for the chemiosmotic hypothesis (Brown and Goldstein, 1985), for endocytosis (Ciechanover *et al.*, 2004), for ubiquitination and protein degradation, and just recently (Rothman *et al.*, 2013), for vesicle trafficking, and (Betzig *et al.*, 2014) for advances in light microscopy. The early cell biologists insisted on quantitative application of these new techniques and laid the groundwork for modern cell biology. As with biochemistry, reductionism has been the keystone for the amazing success over these past several decades. Each organelle has its own research history – the nucleus, mitochondria, peroxisomes, the proteasome, the ER, and cytoskeleton. The Golgi apparatus/secretory pathway and the endosome–lysosome network have all been examined in detail both in isolation and in their relationships with each other. A new member, the exosome (aka extracellular vesicle) has emerged more recently and stimulated the imagination of aspiring young investigators (Harding *et al.*, 2013). The rise of genomics, the development of spectacular imaging modalities, and evolving biochemical techniques (proteomics) have led to the discovery of molecular motors, the identification of macromolecular complexes such as the exocyst, ESCRT, GARP, SNARES among others that choreograph complex intracellular pathways, including cell motility and cytokinesis, have brought cell biology into this new era where the whole is now seen as larger than the individual components. This creates the segue from Parts 1 and 2 of the *Encyclopedia (Molecular Cell Biology and Organizational Cell Biology)* to Part 3 (*Functional Cell Biology*) where signaling modalities will be seen as an integrative force for regulation and control.

Signaling

While all organisms can sense their environment and respond to cues from it, multicellular organisms must in addition coordinate their responses, which require intercellular communication at a sophisticated level (Bradshaw and Dennis, 2009). The higher the development, the more complex these communication systems become. Thus the cell biologist must focus not only on how molecular function is translated into cell organization and how these functions are coordinated from organelle to organelle but also on the external interactions and signals that control the larger functional responses of organs and ultimately organisms.

Intercellular communication is afforded by stimuli that can be transmitted across the cell membrane, either directly or via membrane bound molecules that in turn pass signals to the intracellular compartment. The origins of these stimuli may be quite diverse. Among other things, they can include cell–cell contacts, soluble factors/messengers, or foreign agents. Lipophilic molecules can cross the membrane by diffusion or by facilitated transport to recognize and bind to intracellular entities but most substances bind to membrane bound proteins and induce their signal by ‘activating’ them instead. These so-called receptors are capable of producing a variety of responses that invariably involve the generation of posttranslational modifications of preexisting proteins. Protein phosphorylation, which in eukaryotes mostly occurs on tyrosine, serine, and threonine residues, is highly prevalent and appears to be directly or indirectly involved in almost all transmembrane signaling (Gnad *et al.*, 2011). However, it is by no means the only vehicle for transmitting information as acetylation, methylation, monoglycosylation, and ubiquitination, to name a few, are both important and widespread. There are over three hundred different covalent modifications (Khoury *et al.*, 2011) that are introduced into proteins (and many more that have not yet been chemically defined) of both a transient and stable nature and most are likely to be involved to some extent in signal transduction processes. In addition, limited proteolysis can also be an important part of a pathway and this activity is a major player in cellular responses (Turk *et al.*, 2012).

The induction of an intracellular signal is basically perpetrated by the formation of new interactions between the modified proteins and other entities, which can range from small molecules to macromolecular protein complexes. One of the most important advances in understanding how cells process information induced from external stimuli was the appreciation that the newly formed sites produced by the protein modifications were recognized by other proteins with modules specifically designed for this purpose (Pawson, 1995). Thus, for example, phosphotyrosine residues could be recognized by other proteins containing a domain, termed SH2 for its relatedness to a similar domain found in the Src protein kinase, which could then be further modified allowing for additional interactions to occur. By these ‘docking’ events, signaling complexes could be assembled, often in multiple steps, that ultimately lead to the activation of key effectors. The end point of many of the pathways is the activation of transcription factors that lead to the modulation of gene expression of that cell by ultimately changing its protein expression profile (Bradshaw and Dennis, 2009). However, many molecules are usually modified and activated ‘along the way’ and these ‘new’ activities also add to the overall response to the original stimuli. Short term responses indeed require that the protein effectors necessary for it are already present; long term responses per force require new protein synthesis.

Signal transduction is thus dependent on two phenomena: posttranslational modifications, particularly of the readily reversible type, and protein–protein interactions. The extent to which these two activities take place, even in resting, unstimulated cells was greatly underestimated before the advent of proteomics and the introduction of mass spectrometric methodology into cell biological research (Bradshaw and

Burlingame, 2005). These high-throughput unbiased analyses of very complex mixtures, basically derived directly from cell lysates, revealed that essentially every protein had multiple interacting partners and that posttranslational modifications affected a very significant proportion of the proteins present. These were profound differences from what had been the prevailing wisdom and amounted to a paradigm shift in thinking about cellular organization and structure. As noted above, these same tools have gone on to cast new light on organelle organization in terms of resident proteins and have helped to disclose the structure of important cellular machines, such as the proteasome (Voges *et al.*, 1999), the nuclear pore (Routa *et al.*, 2000), and transcription complexes (Kornberg, 2007). They have also begun to elucidate the complexity of epigenetic regulation of gene transcription at the histone level (Cosgrove, 2012), which will also be essential in completing our understanding of signaling processes.

Concluding Remarks

Two of the most fundamental aspects of cells are their ability to reproduce themselves and, in higher organisms, to undergo changes that lead to new cell types and functions. These developmental processes leading to differentiation are what allow a single fertilized egg to form a complex adult organism and require the synthesis of all aspects of cell biology to understand. Knowing how cells go from one state to another in a timed and regulated fashion forms the core of developmental biology, which can be viewed as a sub discipline of cell biology. Also of major importance is the maintenance of viability and the associated turnover processes. The control of cell death is not only an essential part of development it is also key to managing situations that have gone awry and underlies many serious disease conditions.

It is clear that science is still a long way from understanding how even simple cells work. There is a long list of functions and structures that remain to be elucidated and integrating the various experimental approaches and the data they produce still lags far behind. This is the ‘Era of Big Data’ and vast amounts of new information that impact on our understanding of cells and their processes are collected every day. Genomic information is a good example – despite the rapidity that this sort of information can now be obtained, it still remains to be effectively mined in terms of what it can tell us about basic principles of how cells work. There has been an understandable pressure to apply this information to managing disease (Collins and Varmus, 2015), particularly those that are life threatening, but there certainly is information that is yet to be formulated that would apply to fundamental problems as well. The same can be said for transcriptomics, proteomics, and metabolomics. In addition to these powerful new technologies, singular advances in analyzing single cells promises to augment these molecular approaches significantly (Di Palma and Bodenmiller, 2015).

One aspect of cell biology that is looking to address the challenges of integrating these relatively newly minted studies is systems biology (Part 4). It is far too early to assess the value of these efforts in coordinating this flood of information but it certainly looks promising. It will be an elusive goal for some

time to come but also a stimulus for new generations of cell biologists that will be forthcoming.

References

- Betzig, E., Hell, S.W., Moerner, W.E., 2014. The Nobel Prize in chemistry — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2014/press.html (accessed 03.04.15).
- Bradshaw, R.A., Burlingame, A.L., 2005. From proteins to proteomics. *IUBMB Life* 57, 267–272.
- Bradshaw, R.A., Dennis, E.A. (Eds.), 2009. *Handbook of Cell Signaling*, second ed., vol. 1–3. San Diego, CA: Elsevier Academic Press, pp. 1–3047.
- Brown, M.S., Goldstein, J.L., 1985. Physiology or medicine — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1985/press.html (accessed 03.04.15).
- Ciechanover, A., Hershko, A., Rose, I., 2004. The Nobel Prize in chemistry — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2004/press.html (accessed 03.04.15).
- Claude, A., de Duve, C., Palade, G.E., 1974. Physiology or medicine — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1974/press.html (accessed 03.04.15).
- Collins, F.S., Varmus, H.A., 2015. New initiative on precision medicine. *New England Journal of Medicine* 372, 793–795.
- Cosgrove, M.S., 2012. Writers and readers: Deconvoluting the harmonic complexity of the histone code. *Nature Structural & Molecular Biology* 19, 739–740.
- Di Palma, S., Bodenmiller, B., 2015. Unraveling cell populations in tumors by single-cell mass cytometry. *Current Opinion in Biotechnology* 31, 122–129.
- Gnad, F., Gunawardena, J., Mann, M., 2011. PHOSIDA 2011: The posttranslational modification database. *Nucleic Acids Research* 39 (suppl. 1), D253–D260.
- Harding, C.V., Heuser, J.E., Stahl, P., 2013. Exosomes: Looking back three decades and into the future. *Journal of Cell Biology* 200, 367–371.
- Khoury, G.A., Baliban, R.C., Floudas, C.A., 2011. Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database. *Scientific Reports* 1, 90. doi:10.1038/srep00090.
- Kornberg, R.D., 2007. The molecular basis of eukaryotic transcription. *Proceedings of the National Academy of Sciences of the United States of America* 104, 12955–12961.
- Mitchell, P., 1978. The Nobel Prize in chemistry — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1978/press.html (accessed 03.04.15).
- Pawson, T., 1995. Protein modules and signaling networks. *Nature* 373, 573–580.
- Rothman, J.E., Schekman, R.W., Südhof, T.C., 2013. The 2013 Nobel Prize in physiology or medicine — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/2013/press.html (accessed 03.04.15).
- Rout, M.P., Aitchison, J.D., Supratka, A., *et al.*, 2000. The yeast nuclear pore complex — Composition, architecture, and transport mechanism. *Journal of Cell Biology* 148, 635–652.
- Türk, B., Türk, D., Türk, V., 2012. Protease signalling: The cutting edge. *EMBO Journal* 31, 1630–1643.
- Voges, D., Zwickl, P., Baumeister, W., 1999. The 26S proteasome: A molecular machine designed for controlled proteolysis. *Annual Review of Biochemistry* 68, 1015–1068.

Organizational Cell Biology: An Overview

EL Elson, Washington University School of Medicine, St. Louis, MO, USA

G Raposo, Institut Curie, Paris, France

PA Gleeson, The University of Melbourne, Melbourne, VIC, Australia

A Akhmanova, Utrecht University, Kruytgebouw, Utrecht, The Netherlands

T Yoshimori, Osaka University, Suita, Osaka, Japan

PD Stahl, Washington University School of Medicine, St. Louis, MO, USA

B Goud, Institut Curie, Paris, France

© 2016 Elsevier Inc. All rights reserved.

The cell is the building block of life and, as written by [Lewis Wolpert \(1999\)](#), can be considered as the evolution's most magnificent achievement. Cells, in particular animal cells (plants cells being surrounded by a wall that imposes a constraint), display all kinds of shapes and sizes, and many different types are present in the same living organism (for instance, more than 350 types of cells have been described in humans). Besides this huge diversity, all cells share the same fundamental feature, i.e., the capacity to behave as stable functional units. Such remarkable functional stability allows them to resist change over time and to respond to threats. The cell functional stability relies on its organization in compartments or organelles, which are themselves organized in cellular space by the cell cytoskeleton. Still this organization is not static and the development of live cell imaging approaches in the last 20 years has illustrated the extraordinary dynamics of cell compartments and the multiple connections that exist among them. On the other hand, the loss of stability at the cell level leads to a pathological state, a remarkable example being cancer progression. A loss of stability can also be provoked by pathogens and toxins, which in many cases, can subvert cellular functions to their own benefit.

Part II of the Encyclopedia is divided into five sections that describe the range of imaging approaches developed to date to understand the cell organization (see Section Imaging the Cell), the principle organelles (see Section Organelles), the communications between organelles (see Section Inter-organelle Communication), how cell architecture is maintained (see Section Cytoskeleton and Motors), and how cell organization can be altered by pathogens (see Section Intracellular Infection).

Imaging the Cell

The potential of optical methods to acquire a dynamic view of molecules, cells, tissues, and organisms is extraordinary indeed, microscopy is the Dean of Cell Biology. Both light and electron microscopy have moved cell biology into a new era. Yet subcellular structures and organelles, their specific morphological and compositional characteristics, their intimate details and possible connections can be seen only by high-resolution transmission and scanning electron microscopy (TEM and SEM, respectively). TEM has been used for more than 70 years to investigate the ultrastructure of cells and tissues in normal and pathological situations. In more recent times methods have been developed that allow for the analysis of molecular complexes, fine details and 3D analysis of cells,

tissues, small organisms, and subcellular structures including the localization of cellular components with an intraorganelle context. Single-particle cryoEM allows for a near-atomic ($\sim 3\text{--}4\text{ \AA}$) resolution of structures or macromolecular complexes such as virus particles, molecular machineries, and membrane proteins. Importantly, these can be viewed in close to native state, without contrasting agents and after vitrification by Cryo-EM ([Single-Particle CryoEM of Macromolecular Complexes](#)). Scanning EM has a lower resolution than TEM but it is the only method that permits one to delineate the surface of cells and small organisms and to appreciate structural details thanks to its depth of field and three-dimensional (3D) surface imaging characteristics ([Scanning Electron Microscopy in Cell Biology](#)). These classical SEM methods are powerful and have been improved lately by developments in electron microscope technology permitting visualization of isolated organelles and large complexes such as nuclear pores ([Scanning Electron Microscopy in Cell Biology](#)). Unfortunately, they do not provide 3D information on the internal characteristics of cells, tissues, and organisms. Over the past few years this limitation has been surpassed by combining SEM imaging with the use of a microtome or a focused ion beam (FIB) that progressively removes sections from the top of the object under study thus uncovering fine details of the interior structures. The acquired images can be stacked and a 3D model can be generated to appreciate the internal organization of a specimen that may be composed of several cells ([Imaging Cellular Architecture with 3D SEM](#)). In a very complementary manner and with a higher resolution, Electron Tomography allows one to resolve the 3D organization of molecular and subcellular structures by TEM of plastic embedded sections of cells and tissues and Cryo-TEM of unstained, vitrified specimens ([Electron Tomography](#)).

Visualization of molecules within their intracellular and intraorganelle contexts can be achieved using electron dense probes that can be viewed by TEM. Localization methods are most generally based on the recognition of molecules by specific antibodies (Immuno-EM). Such methods combine localization and preservation of ultrastructure and permit discrimination of seemingly identical compartments by their molecular composition in a quantitative manner ([Immuno-electron Microscopy: High-Resolution Immunocytochemistry](#)).

For many years fluorescence has been the dominant optical property used to study structural and functional properties of cells. Its advantageous qualities include chemical selectivity, i.e., the ability to measure specifically labeled molecular species at very low concentration in the presence of high concentrations of unlabeled molecules, relative ease and speed of measurement, high spatial resolution when used with a

confocal or 2-photon fluorescence microscope and, when used judiciously, lack of interference with normal cell function. The articles in this section of the Encyclopedia provide an introduction to methods that exploit recent advances in microscopy and labeling technology that have opened a new golden age of applications of fluorescence to cell biology. Several of the articles describe methods that enable imaging resolution beyond the Abbe/Rayleigh limit in the range of 200–300 nm for visible light. The oldest of these is total internal reflection microscopy (TIRFM) that provides sub-diffraction optical resolution along the optical axis and so has been a mainstay of studies of dynamic processes of single molecules bound to glass substrates. TIRFM does not, however, improve lateral resolution ([Total Internal Reflection Fluorescence Microscopy](#)). Localization methods such as PALM, STORM, and FPALM assemble high-resolution images of macromolecular structures by summing the contributions of component fluorophores whose individual positions are determined with accuracies near 10 nm ([Super Resolution Fluorescence Localization Microscopy](#)). As described in article [Super-Resolution Light Microscopy: Stimulated Emission Depletion and Ground-State Depletion](#), STED microscopy is scanning super-resolution approach based on restricting the radius of the excitation light to sub-diffraction dimensions. STED has the advantage of acquiring fluorescence signals rapidly enough to be useful in studies of dynamic molecular behavior, for example, via fluorescence correlation spectroscopy (FCS). Structured illumination microscopy (SIM) breaks the diffraction limit by imposing a periodic spatial pattern on the excitation light. Although SIM does not match the resolution of STED and localization methods, it has other advantages of flexibility, relative ease of use and high throughput ([Structured Illumination Microscopy](#)). At a higher structural level intravital microscopy ([Intravital Microscopy in Mammalian Organisms: From Tissue Physiology to Cell Biology](#)) is used to investigate tissue and cell architecture in living animals and cells. This method has recently benefitted from applications of molecular genetics and improvements in microscopy and fluorescent probes. Fluorescence lifetimes are sensitive to the environment of the emitting fluorophore. Therefore, fluorescence lifetime imaging microscopy (FLIM) provides a unique and popular method for enhancing the contrast of fluorescence images to reveal regional differences in a variety of cellular properties and functions, for example, metabolism, ionic concentrations, and molecular composition ([Fluorescence Lifetime Imaging – Applications and Instrumental Principles](#)). Optogenetics ([Optogenetics](#)) uses light to trigger molecular functions, and thereby to control activities in living cells. This powerful new approach has mainly been used to control cellular excitability through regulating ion channel function. Although FCS was invented in the 1970s, many variations and extensions have been developed, since then continue to be valuable tools to investigate dynamic molecular processes in cells. The family of FCS and closely related photobleaching (FRAP (fluorescence recovery after photobleaching)) methods yield important information about molecular transport as well as rates and extents of molecular interactions ([Fluorescence Correlation Spectroscopy: A Tool for Measuring Dynamic and Equilibrium Properties of Molecules in Cells](#)). While FCS and FRAP rely on dynamic changes in the numbers of fluorescent molecules

detected within a small measurement area that can be no smaller than the optical resolution limit to measure transport rates, single-particle tracking extended by current localization microscopy methods allows measurement of the trajectories of individual molecules with precision in the range of 10 s of nm ([High-Speed Localization Microscopy and Single-Particle Tracking](#)). This approach allows a unique analysis of the mechanisms of motion of the tracked molecules. Finally, the use of genetically encoded fluorescent probes coupled with live cell imaging opens a window of opportunity to view cellular events in spectacular detail ([Genetically Encoded Fluorescent Probes and Live Cell Imaging](#)).

Organelles

Acquisition of diverse intracellular compartments was truly revolutionary epoch in the long history of cellular evolution. Each compartment, known as an organelle, is enclosed by membrane, meaning that cells have a vast area of membrane surface at their disposal, in addition to the plasma membrane surrounding cells, which can act as a platform for biochemical reactions that sustain cellular activities. In addition, the interior space of organelles (the luminal space) sequestered by their limiting membranes also enables cells to perform distinct reactions efficiently. Indeed, organelles contain their own characteristic set of proteins and other molecules to achieve functions unique to each.

Study of organelles is undoubtedly one of major streams in cell biology, which started when Robert Brown observed a small dark structure inside plant cells and named it ‘nucleus’ in 1831. This seminal discovery was followed by identification (or naming) of mitochondria by C. Benda and the Golgi apparatus by C. Golgi in the same year 1898, and lysosomes by C. de Duve in 1955. An enormous amount of research from these times has revealed the amazing ‘true colors’ of organelles, but we have not yet completely understood their function and structure; study is still ongoing. This section of the encyclopedia comprehensively covers major organelles including recently discovered and cell-type-specific organelles.

Organization of the nucleus, a center for genome storage and for controlling its expression, is introduced in the article [Nuclear Organization \(Nuclear Structure and Dynamics\)](#). There is macromolecular traffic between nucleus and the cytoplasm via the nuclear pore, which is large protein complex inserted into the nuclear double membrane, as discussed in the article [Nuclear Pores](#). Van Anken and Sitia ([The Endoplasmic Reticulum](#)) describe the multifunctions of the endoplasmic reticulum; synthesis of lipids and proteins, Ca^{++} ion storage, signal integration, and interaction with other organelles. Otera and Mihara ([Mechanisms and Functions of Mitochondrial Dynamics](#)) focus on mitochondria dynamics; these cellular power plants are highly dynamic and fusion/fission of their membranes play a pivotal physiological role. Cargo internalized into cells from outside is transferred sequentially along the endocytic pathway consisting of early endosomes ([Early Endosomal Compartments; Endocytosis of Cargo Proteins: LDL](#)), late endosomes ([The Late Endosome](#)), and lysosomes ([Conventional and Secretory Lysosomes](#)). Lysosomes are final destination in the pathway, but can fuse with plasma membrane under certain

conditions (secretory lysosomes), as discussed in the article [Conventional and Secretory Lysosomes](#). In addition, endosomes also function as a platform of signaling, as described in the article [Signaling from Endosomes](#). Membrane traffic from the cytoplasm to lysosomes also exists (the autophagic pathway), which is mediated by autophagosomes ([At the Center of Autophagy: Autophagosomes](#)). On the other hand, lipids and proteins synthesized in the endoplasmic reticulum are transported to the Golgi (Golgi and TGN) via intermediate compartments ([Intermediate Compartment: A Sorting Station between the Endoplasmic Reticulum and the Golgi Apparatus](#)) to be sorted into each destination. Peroxisomes functioning in oxygen metabolism and fatty acid degradation ([Peroxisomes](#)) and lipid droplets specialized for lipid storage and metabolism ([Lipid Droplets](#)) are also discussed. Wauben ([Extracellular Vesicles](#)) discusses the most-recently discovered organelles, extracellular vesicles, which are released from cells and have diverse important physiological functions. While the above-mentioned organelles are ubiquitous in most cell types, there are some cell-type-specific organelles. Only plants and algae have chloroplasts for photosynthesis ([Plastids: The Anabolic Factories of Plant Cells](#)). Nerve cells in animals possess synaptic vesicles for neurotransmission, as described in the article [Synaptosomes and Synaptic Vesicles](#). Lastly, Fukuda ([Lysosome-Related Organelles](#)) discuss a group of organelles having very specialized functions, for example, pigmentation derived from lysosomes. These lysosome-related organelles exist in various highly differentiated cells.

Interorganellar Communication

The compartmental organization of the eukaryotic cell is essential for the development of multicellular organisms, for the organization of tissues and organs, and the coordination of physiological functions. The molecular processes which regulate the biogenesis of membrane compartments and the mechanisms by which information is shuttled back and forth between the different intracellular compartments represents fundamental processes in all eukaryotic cells. The pioneering work by Palade and colleagues, which integrated EM technologies with biochemical approaches, laid the foundation for understanding organelle dynamics and interorganellar communication. Subsequent work over the past 3–4 decades has revealed many of the molecular mechanisms which regulate these processes.

This section focuses on the molecular mechanisms underlying membrane transport pathways. ER export and ER stress responses are now well-understood and discussed in the articles [ER–Golgi Transport](#) and [Endoplasmic Reticulum Stress in Disease](#). Golgi organization and transport is a complex field that has invoked a variety of ingenious approaches in an attempt to define the underlying mechanisms of intracellular transport, as discussed in the article [Intra-Golgi Transport](#). The Golgi is also the site for the glycosylation, and the biosynthesis and role of N-linked glycans and O-glycans in cell biology and development that are discussed in the articles [N-Linked Glycans \(N-Glycans\)](#) and [Extracellular O-Glycans](#), respectively. An emerging field arises from the appreciation that signaling controls transport along the constitutive secretion pathway ([Regulation of the Secretory Pathway](#)). Post-Golgi transport involves

multiple pathways from the *trans*-Golgi network to the cell surface and constitutive and regulated post-Golgi transport are discussed in the articles [Post-Golgi Transport – Cargo, Carriers, and Pathways](#) and [Regulated versus Constitutive Secretion – A Major Form of Intercellular Communication](#). Endocytosis is also an important element of interorganellar transport, and this section discusses both the clathrin-mediated and non-clathrin-mediated pathways in the articles [Clathrin and Clathrin-Dependent Endocytosis](#) and [Clathrin Independent Endocytosis](#). Transport along the endosomal–lysosomal pathway is covered in the article [Endosome to Lysosome Transport](#), and pathways that carry cargo along the retrograde transport route from the endosomes to the Golgi are discussed in the article [Retrograde Transport](#). Machinery components of vesicle formation and vesicle fusion are essential for the regulation of these transport pathways, and in particular include small G proteins ([Rabs of the Endosomal Recycling Pathway; Rabs and Other G Proteins](#)), phosphoinositides ([Interorganellar Communication: Role of Phosphoinositides in Membrane Traffic](#)), Bar domain proteins ([BAR Domains and BAR Domain Superfamily Proteins](#)), SNAREs ([SNAREs: Membrane Fusion and Beyond](#)), ESCRTs ([ESCRTing around the Cell](#)), exocyst and tethering factors ([Vesicle Tethers](#)) and components of the vesicle coats that interact with cargo, such as adaptor complexes ([Adaptor Proteins: Inter-Organellar Traffic Controllers](#)) and retromer ([The Retromer Complex](#)). And finally also included are pathways of cell secretion that are independent of the pathway involving the ER and the Golgi, referred to as unconventional protein secretion ([Unconventional Protein Secretion: Fibroblast Growth Factor 2 and Interleukin-1 \$\beta\$ as Examples](#)).

Cytoskeleton and Motors

The ability of cells to regulate their shape, organize their components, move, and divide depends on the networks of filaments collectively called the cytoskeleton. In contrast to the human skeleton, which remodels slowly, most cytoskeletal filaments are dynamic, because they are composed of polymers of small building blocks that can rapidly assemble and disassemble, depending on the cellular needs. Growth and depolymerization of cytoskeletal filaments can generate forces that can drive cell protrusion and promote chromosome separation during cell division. Moreover, some cytoskeletal filaments can serve as rails for the movement of motor proteins, which use the energy released by ATP hydrolysis to perform mechanical work. The activity of such motors is responsible for the locomotion of single cells, cell contraction and intracellular transport of membrane organelles, RNA, and protein complexes.

This section describes the composition, structure, and dynamics of the major eukaryotic cytoskeletal filaments: actin ([Actin Assembly Dynamics and Its Regulation in Motile and Morphogenetic Processes](#)), microtubules ([Microtubules and Microtubule-Associated Proteins \(MAPs\)](#)), intermediate filaments ([Intermediate Filaments](#)), and septins ([Septins: Cytoskeletal Filaments with Structural and Regulatory Functions](#)). A separate article is dedicated to the structurally and functionally diverse cytoskeletal networks present in bacteria and archaea ([Bacterial and Archaeal Cytoskeletons](#)). Subsequent articles discuss the molecular mechanisms, motile properties,

and functions of the major cytoskeletal motor families. These include the actin-based motors, myosins ([Myosins](#)), and the microtubule-based motors, kinesins ([Kinesin Superfamily Proteins \(KIFs\) as a Fundamental Component of Life: Intracellular Transport and Beyond](#)) and dyneins ([Dyneins](#)).

An important property of the cytoskeletal components is their ability to form complex structures, which perform elaborate cellular functions. The article [Centrioles and the Centrosome](#) discusses centrioles and centrosomes, which are responsible for microtubule organization. The article [Cilia and Flagella](#) describes microtubule-based protrusions, cilia and flagella, which are required for motility or sensory and signaling functions. The intricate cytoskeletal machine that drives chromosome separation during cell division, the mitotic spindle, is discussed in article [The Mitotic Spindle](#). The motile and contractile properties of cells and whole organisms strongly depend on actin-based structures, such as filopodia and lamellipodia ([Filopodia and Lamellipodia](#)) or muscle fibers ([Skeletal Muscle](#)). Formation of complex cytoskeletal assemblies, which is strongly regulated by signaling factors, such as Rho GTPases ([Rho GTPases](#)), enables cells to adhere to the extracellular matrix ([The Extracellular Matrix; Cell Adhesion to the Extracellular Matrix](#)) or to each other ([Cell–Cell Adhesion and the Cytoskeleton](#)), polarize ([Cell Polarity](#)), and migrate ([Cell Migration](#)).

Intracellular Infectiology

The interaction between viral, bacterial, and protozoan pathogens and eukaryotic cells has been of interest to biologists and physicians for over two centuries, since Metchnikoff noted the internalization of bacteria by wandering mobile cells. With the advent of modern cell biology and the emergence of imaging and molecular biology technologies in the past three decades, the field of host defense against pathogens and its failure – resulting in intracellular infection – has advanced at a very rapid pace. Understanding how pathogens access the portals of entry normally reserved for nutrition and defense has intrigued a generation of cell biologists not only from the point of view of expanding our knowledge of normal cell behavior but also of getting a better handle on the molecular mechanisms that invasive organisms use to access intracellular niches where they proliferate. Moreover, understanding the molecular mechanisms of intracellular infection opens up a window of opportunity to develop therapeutics. This section of the Encyclopedia initiates a journey through the various pathways and processes that cells use to kill pathogens by laying out the mechanisms of phagocytosis ([Phagocytosis](#)), including a description of the various traffic events that bring the internalized microbes into compartments designed to kill and dismember the invader. Many small particles (and fluid) access the cell by macropinocytosis, a process first described by Lewis almost century ago. In article [Macropinocytosis](#) Swanson and Yoshida describe the role of macropinocytosis in cell growth, antigen processing, and infection by virus and

bacteria. Intracellular killing is mediated by reactive oxygen species carried out by complex oxidases, assisted by powerful cationic antimicrobial peptides. Allen ([Microbicidal Mechanisms](#)) lays out the basis for intracellular killing as an important component of host defense. Pathogenic bacteria and protozoa have evolved mechanisms to resist killing by host cells; the molecular details of these mechanisms are truly intriguing. Using examples from common pathogens such as *Brucella*, *Chlamydia*, *Legionella*, *Listeria*, *Salmonella*, and *Shigella*, Isberg and colleagues ([Bacterial Subversion of Phagocytic Killing](#)) lay out the pathways that intracellular pathogens use to evade detection and killing, thereby permitting residence in intracellular niches where they proliferate. Ferrari and Sansonetti ([Cellular Invasion by Bacterial Pathogens](#)) extend this analysis, focusing on the molecular mechanisms of pathogen uptake by nonphagocytic cells and how certain pathogens survive by subverting cellular functions. Lastly, Rigoni and Montecucco ([Bacterial Protein Toxins as Tools in Cell Biology and Physiology](#)) discuss potent toxins that increase the possibility of bacterial survival in the face of an effective microbicidal immune system, and how these toxins have become useful reagents in cell biology research. In addition to pathological interactions, the interaction between bacteria and host cells may operate in a saprophytic relationship. Together with the advent of NGS, this concept has opened up a new field called the microbiome ([The Gut Microbiome](#)).

A second large group of pathogens are the viruses, which need to infect host cells to replicate themselves. Helenius ([Cell Biology of Virus Infection](#)) describes the process by which viruses enter cells to exploit the cell's transport and synthetic pathways. Extending this analysis, Wileman and colleagues ([Virus Factories and Mini-Organelles Generated for Virus Replication](#)) describe how various viruses develop factories for efficient intracellular replication. One virus- HIV- has received intense scrutiny at the molecular, cellular, and biomedical/epidemiological levels leading to an impressive while ever improving understanding of its life cycle. Giese, Pelchen-Matthews and Marsh ([HIV – The Cell Biology of Virus Infection and Replication](#)) provide a article outlining in detail the Cell Biology of Virus Infection and replication.

Over the last two decades another type of disease transmission has emerged, following the discovery of prions. Legname and Pischke ([Prions](#)) discuss the cell biology and transmission of scrapie and other prion diseases. Lastly, model organisms being used to understand the relationship between pathogen and host are discussed. Sun and colleagues ([Cellular Responses to Infections in *Caenorhabditis elegans*](#)) describe the cellular responses to infection, using *Caenorhabditis elegans* as a model organism.

Reference

- Wolpert, L., 1999. From egg to adult to larva. *Evolution & Development* 1, 3–4.
Available at: <https://10.1046/j.1525-142x.1999.00111.x> (accessed 05.06.15).

IMAGING THE CELL: ELECTRON MICROSCOPY

Contents

Single-Particle CryoEM of Macromolecular Complexes

Scanning Electron Microscopy in Cell Biology

Electron Tomography

Immunoelectron Microscopy: High-Resolution Immunocytochemistry

Imaging Cellular Architecture with 3D SEM

Single-Particle CryoEM of Macromolecular Complexes

J Zhang and KV Gorzelnik, Texas A&M University, College Station, TX, USA

MF Schmid and W Chiu, Baylor College of Medicine, Houston, TX, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

One of the emerging structural methods in modern cell biology is single-particle electron cryo-microscopy (cryoEM). Recent advances in single-particle cryoEM have enabled near-atomic ($\sim 3\text{--}4\text{ \AA}$) resolution for structures of macromolecular complexes ranging from large icosahedral virus particles and molecular machines to moderate-sized enzymes and membrane proteins (e.g., Zhang *et al.*, 2010b; Bai *et al.*, 2013; Bartesaghi *et al.*, 2014; Liao *et al.*, 2013; Lu *et al.*, 2014). What makes this method unique is that the macromolecular complexes can be observed under nearly physiological or biochemical conditions, without artifacts possibly generated from either staining or crystallization.

Monodisperse particles of macromolecular complexes are frozen in vitreous ice. Thousands to hundreds of thousands of two-dimensional (2D) images of these particles are recorded in an electron microscope with a cryo-specimen stage. These particle images are computationally processed to generate a three-dimensional (3D) map and (with sufficient resolution) an atomic model of the macromolecular complex, which are deposited to the EMDData Bank and Protein Data Bank for public access (see Section Relevant Website).

Since numerous reviews have been written to describe the detailed protocols for each of these steps (Chang *et al.*, 2012; Baker *et al.*, 2010), we will not repeat them here. However, we will describe some cautionary notes about the hurdles often encountered in each of the steps. Then we will present some examples that illustrate the types of structural information that can be derived from cryoEM structures.

CryoEM Sample Consideration

For single-particle cryoEM, samples are diluted to a concentration of $\sim 0.1\text{--}1\text{ mg ml}^{-1}$ in a suitable buffer, in order to

ensure the particles are well separated from each other. Certain commonly used buffer constituents like glycerol and sucrose have an adverse effect on the image contrast and should be avoided or diluted out during the cryo-specimen preparation step. The molecular mass of the macromolecular complex can be as large as tens of MDa. Due to the weak contrast in the cryoEM images, there is generally a lower limit to the size of the macromolecules that can be recognized in images (Henderson, 1995). Recently, a cryoEM structure of a 170 kDa membrane protein was solved to 4.5 Å resolution (Lu *et al.*, 2014). For DNA/RNA, much smaller sizes can be imaged due to the stronger electron scattering power of the phosphate backbone. A low-resolution structure of a 50 kDa folding intermediate of a ribozyme domain has been reported using single-particle cryoEM (Baird *et al.*, 2010).

Compositional and conformational heterogeneity is commonly seen in biologically active complexes. Compositional heterogeneity occurs when the imaged particle is composed of multiple loosely connected components, or a ligand is weakly bound to its target. One successful biochemical method to minimize such sample heterogeneity is to mildly cross-link the subunits in the complex, then purify the cross-linked complex using ultracentrifugation or gel filtration techniques (Murakami *et al.*, 2013; Shukla *et al.*, 2014). Conformational heterogeneity could be large or small, continuous or discrete, but is commonly seen in single-particle cryoEM, since different parts of the macromolecules can have different dynamic motions in their biologically active states. For certain macromolecules, small ligands bound in the active site or other binding pocket may stabilize their structures and thus minimize the conformational heterogeneity. For example, non-hydrolyzable ATP analogues are commonly used to trap molecular chaperones in one of their ATP-dependent transition states, making the chaperones more uniform and sometimes more rigid in conformation (Zhang *et al.*, 2010a). We

should point out that cryoEM can deal with samples with compositional and/or conformational heterogeneity by various data sorting algorithms during image processing (Scheres *et al.*, 2007; Elad *et al.*, 2008; Penczek *et al.*, 2011; Elmlund and Elmlund, 2012; Chen *et al.*, 2006).

Frozen, Hydrated Specimen Preparation

An electron microscopy grid is equivalent to the glass slide in light microscopy. It provides a surface to support the specimen in the electron microscope. The standard grid is a thin copper or other metallic foil mesh with a diameter of 3.05 mm; it is covered by thin carbon film with periodically arranged holes (1–2 μm). Ideally, the specimens are suspended in these holes. In some cases, the particles may stick to the support net of the grid or do not stick to the grid at all. A thin carbon support film across the holes often works to recruit the particles but suffers from introducing a higher background to the images. Recently, innovative methods have been put forward to help anchor the particles on a cryoEM grid using graphene or specific affinity surface labeled molecules (Russo and Passmore, 2014; Kelly *et al.*, 2010).

Several semi-robotic devices (FEI Inc., Gatan Inc., Leica Inc.) have been developed to aid in preparing the frozen, hydrated specimen on a cryoEM grid. Many of these devices allow the users to control vitrification parameters such as the duration and force of the blotting, the temperature and humidity of the environment, etc. Standardization of these parameters improves the sample quality and the reproducibility of the experiment. To obtain single-particle cryoEM images of the specimen with good contrast, the ice must be as thin as possible while still fully embedding the specimen particles, keeping them in the frozen, hydrated state. Contamination of the sample by crystalline ice can occur when the processes of freezing or transferring the sample are not performed well. For example, the liquid ethane may not have been cooled down enough by the liquid nitrogen, or the grid may have been exposed to warm or humid air during the multiple grid transfers from the freezing apparatus to the microscope column. This type of contamination can be minimized by proper handlings of the frozen grids in various transfer steps and always keeping cryoEM room at low humidity.

An often-encountered frustration in cryoEM specimen preparation is referred to as the ‘preferred orientation’ of the particles embedded in ice. To obtain a 3D density map of the macromolecular complexes using cryoEM, 2D images of the particle must cover all orientations with respect to the incident electrons. Some particles have a preference to lie in certain orientations, generating insufficient data sampling for a 3D reconstruction. Such preferred particle orientation may be due to a surface-exposed patch of hydrophobic residues, for instance, which tends to be attracted to the air/ice interface on the grid. In this case, a small amount of mild detergent (e.g., 0.05% octyl glucoside) added to the buffer, may induce the particles to adopt a more random orientation on the grid before freezing (Zhang *et al.*, 2010a). One may also genetically mutate or remove the exposed hydrophobic residues, if the overall structure of the molecule is not changed (Zhang *et al.*, 2010a).

CryoEM Image Acquisition

Special care must be taken when imaging with electrons, as radiation damage will occur if the ice-embedded specimen is overexposed, leading to a loss of structural information. In order to minimize the radiation damage to the specimen, the total electron exposure applied to the specimen must be limited. About 15–25 electrons per $\text{\AA}^2$ are regarded as a safe specimen exposure for subnanometer resolution structure determination (Bammes *et al.*, 2010). The data collection can be automated (Nakamura *et al.*, 2010; Zhang *et al.*, 2003, 2009; Lei and Frank, 2005; Carragher *et al.*, 2000).

A recent exciting development in single-particle cryoEM is the availability of new generation direct electron detection cameras (Bammes *et al.*, 2012; Bai *et al.*, 2013; Li *et al.*, 2013). Compared with the traditional recording media such as charge coupled device (CCD) cameras, direct electron detection cameras omit the scintillator and thus avoid the random scattering of electrons when they travel through the scintillator, thereby minimizing the noise and enhancing the quantum detection efficiency. With their ultrafast refreshing rate, multiple frames of images from the same specimen area can be recorded and aligned computationally to correct for any image blurring due to the electron beam-induced movement or mechanical stage instability during image acquisition. Figure 1 shows an image of ice-embedded single particles of ribosomes before and after multiple frames of an image have been drift-corrected and averaged. It is clear that the power spectrum of the aligned average has more isotropic and higher resolution contents than the nonaligned average.

While the new digital cameras achieve cryoEM images with significantly better quality, the amount of raw digital data has dramatically increased from tens of megabytes for each micrograph to several gigabytes for a stack of frames. One stack of frames is taken from the same specimen location on the grid while fractionating the total electron exposure into each of the frames within the stack. Thus, it is feasible to acquire 0.5–1.0 TB of single-particle image data in a few hours of data collection. This poses a major challenge for data transfer between the microscopes and the data-processing computers and for the data archiving in the database; a much faster speed of the network connections and larger storage space are required.

CryoEM Image Processing

Image processing consumes a major block of time in the cryoEM workflow to convert 2D images into a 3D density map. Image processing consists of multiple steps, including image quality assessment, particle picking, determination and correction of image defocus effects, estimation of particle orientation parameters, map validation, model building and validation. There are multiple challenges in the workflow.

The first is to align the raw 2D particle images against the projection images of a 3D template, which are iteratively updated. There are many ways to generate an initial template. One way is to identify three orthogonal views of 2D images to generate a rough model (e.g., a barrel for a mammalian chaperonin (Cong *et al.*, 2011)). Another way is merely to generate a featureless model with an approximate size and shape (e.g.,

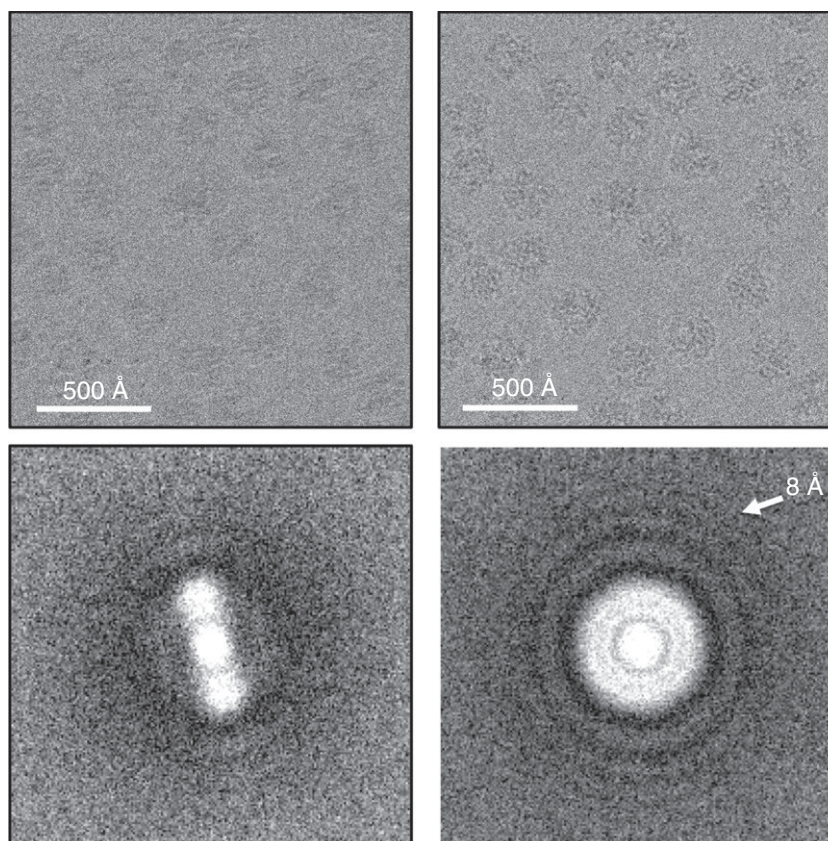


Figure 1 The cryoEM micrographs of the ribosome particles and their power spectra before (left) and after (right) drift correction. The ribosomes are directly frozen in the vitreous ice on a C-flatTM grid without continuous carbon support. The images were collected at a nominal magnification of $19\,000\times$ on a Gatan K2 Summit electron counting direct detection camera with an FEI Tecnai TF20 electron microscope.

spherical shell for a virus particle (Liu *et al.*, 2007)). Experimentally, one can also generate an initial model using random conical tilt or orthogonal tilt imaging (Radermacher *et al.*, 1987; Leschziner and Nogales, 2006). In these later methods, the same particles are imaged as pairs from two angles respectively. These are usually $0^\circ/55^\circ$ tilt pairs for random conical tilt and $-45^\circ/45^\circ$ tilt pairs for the orthogonal tilt methods. The resultant map can then be used as an initial template which is refined into a higher resolution map using the low-dose, untilted images. Particle orientation determination is an iterative process that gradually improves the accuracy and resolution of the map (Ludtke *et al.*, 1999; Tang *et al.*, 2007; Shaikh *et al.*, 2008; van Heel *et al.*, 1996; Grigorieff, 2007).

Another commonly encountered challenge in single-particle cryoEM image processing is the existence of multiple particle conformations in the population of particle images. With such data, one has to first separate the different conformations in the specimen, and ultimately reconstruct a density map from each of the data subsets (Chen *et al.*, 2006; Scheres *et al.*, 2007).

There are multiple image processing software packages which will carry out some or all of these steps. There is a current trend to hybridize different software to take advantage of the strengths of the algorithms in different software packages. In some cases, one may apply different image

reconstruction procedures to help validate the maps of specimens (Murray *et al.*, 2013). Other approaches have recently been proposed to build or refine 3D models directly against 2D cryoEM images (Zhang *et al.*, 2012; Velazquez-Muriel *et al.*, 2012) in order to address the problems due to structural heterogeneity or preferred orientations of the sample on the electron microscope grid.

Since cryoEM reconstruction can be completed at various resolutions ($40\text{--}3\text{ Å}$), the validation criteria of the map vary. One approach to validate the final refined map is called tilt-pair validation technique (Rosenthal and Henderson, 2003). A few particle image pairs of the same specimen are collected at 0° and $\sim 10^\circ$ tilt. The orientation of these image pairs can be determined by using the final refined map as an alignment reference. If the final map is correct, the determined orientations between the same particles in each tilt-pair image should differ by the 10° that was imposed experimentally.

Resolution of CryoEM Density Maps

For a long time, there was no uniform practice in reporting the resolution of a cryoEM map. Recently, a gold standard for this issue has emerged, and has been adopted in many publications since 2013 (Scheres and Chen, 2012; Henderson *et al.*,

2012). This resolution estimation is to split the raw particle images randomly into two groups and processing each of the groups independently. Fourier transformations of the two maps are then computed. These transformations are independently averaged in spherical shells, and then the two sets of shells are correlated, in the Fourier space, a procedure known as the Fourier Shell Correlation (FSC). The FSC shows how well the two maps agree with each other as a function of resolution. It has been shown analytically that the resolution at which the FSC value equals 0.143 is appropriate to report the resolution of the final map generated by the whole particle data set. A common fault in the map determination is over-refining the data so that the map resolution is over-estimated (DiMaio *et al.*, 2013). To avoid such occurrence, it has been suggested to run a test in which the image data has its phases randomized beyond 75% of the claimed resolution (Chen *et al.*, 2013b). The resolution of the maps computed from such data should fall off sharply at the resolution where the data are randomized, indicating no over-refinement (Wang *et al.*, 2014). The true resolution can be estimated accordingly.

Nevertheless, a single value at a certain threshold of the FSC gives only an overall quality of the density map. The resolution of a reconstructed map may not be uniform throughout the entire macromolecular complex. Therefore, efforts have been made to estimate the local resolution variations in a cryoEM density map (Kucukelbir *et al.*, 2014) or to compute the 3D variance map using a statistical bootstrapping approach (Penczek *et al.*, 2006; Chen *et al.*, 2008).

Needless to say, the ultimate judgment of the map resolution is the details of the structure features in the final density map: alpha-helices appear at ~ 10 Å resolution and become clearer beyond 7 Å resolution; the pitch of the alpha-helices

becomes evident at better than 5.5 Å resolution; the separation of beta-strands and the connectivity of the protein chain (in mostly alpha helical proteins) start to be resolved at 4.5 Å (Baker *et al.*, 2013). The resolvability of C-alpha backbone and many sidechains would become better as the map extends beyond 4 Å resolution.

CryoEM Model Building and Model Validation

Once a cryoEM density map is obtained, one can start to interpret its details by building the atomic coordinates of the protein or nucleic acid into the map. If a homologous structure is known for a component in a complex, it can be docked into the density by rigid body or flexible fitting. Several flexible model-fitting techniques based on different physics principles have been developed (Schröder *et al.*, 2007; Topf *et al.*, 2008; DiMaio *et al.*, 2009; Trabuco *et al.*, 2008; Hinsen *et al.*, 2005; Tama *et al.*, 2004). For cryoEM maps determined at sub-nanometer (< 10 Å) resolution, when many secondary structure elements are visible, the docking of molecular components becomes more reliable because the alpha-helices and beta-sheets serve as anchor points to guide the modeling process. At even higher resolutions, when many protein side chains are visible, one can build a *de novo* model restrained by the density, followed by model optimization using and modifying tools already existing for crystallography (Baker *et al.*, 2010).

Interpretation of low-resolution (lower than ~ 10 Å) density maps is nontrivial, especially when the map is a complex of multiple components. Methods like cross-linking mass spectrometry have been developed to facilitate the

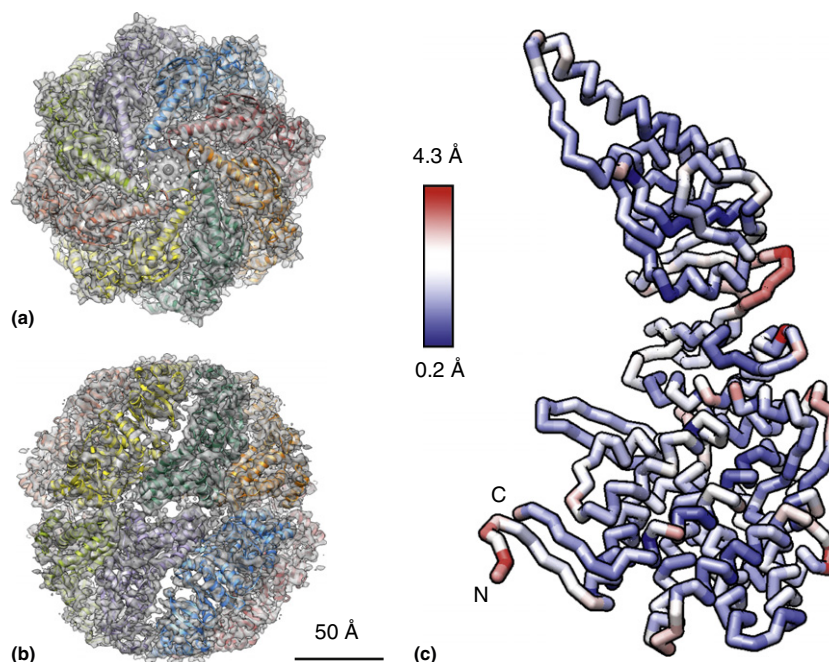


Figure 2 CryoEM map and model of the wild-type Mm-Cpn chaperonin in the top view (a) and side view (b). The pairwise C-alpha deviations between the cryoEM model and the crystal structure are less than 4.3 Å (c). The blue color indicates lower pairwise C-alpha deviations and the red color indicates higher pairwise C-alpha deviations.

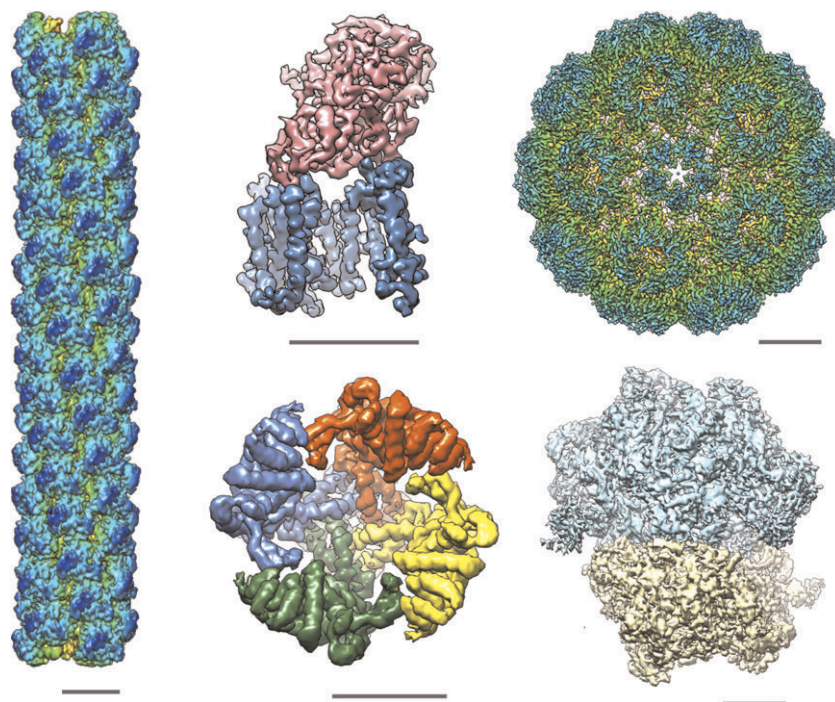


Figure 3 Gallery of a few cryoEM structures solved to 3–4 Å resolutions. PYD inflammasome filament (left, EMD-5830), human gamma secretase (top middle, EMD-2677), TRPV1 ion channel (bottom middle, EMD-5778), Brome Mosaic Virus (top right, EMD-6000), and archaeal 70S ribosome (bottom right, EMD-2277). The scale bars denote 50 Å.

annotation of different components in a cryoEM density map by providing their relative spatial information. Combining cross-linking mass spectrometry with cryoEM is also useful for deriving a model (Leitner *et al.*, 2012).

Model validation is an important step to assure the model is not overfitted with the density data. There are many ways to assure the model is properly optimized. A necessary check is to measure the FSC between the model and the experimental map. The resolution should be close to that of the claimed resolution derived from two independent maps. Another way is to compare independent models derived from each of the half data sets. An FSC of a model derived from one map can be quantitatively computed against the other map (DiMaio *et al.*, 2013). In addition, the RMS deviation of the two models can be quantified to estimate which regions of the model are more robustly determined than the others.

High-Resolution CryoEM Structures

In 2008, four cryoEM structures were reported to show backbone connectivity in individual components of a chaperonins and icosahedral viruses (Ludtke *et al.*, 2008; Jiang *et al.*, 2008; Yu *et al.*, 2008; Zhang *et al.*, 2008). A subsequent *de novo* cryoEM structure was determined for Mm-Cpn, an archeal chaperonin with two back-to-back rings of 16 homoprotomers (Zhang *et al.*, 2010a). Though the exact resolution of this map is lower than the published claim by today's gold standard criterion (DiMaio *et al.*, 2013), the reported backbone of Mm-Cpn was later confirmed to be correct by a subsequent X-ray structure (Pereira *et al.*, 2010), and the pairwise

C-alpha deviation is rather small between the two structures, with the higher deviations mostly in the loop regions of the subunits (Figure 2).

In the last 2 years, there have been an increasing number of atomic resolution maps (<4 Å) generated by cryoEM for icosahedral viruses, ribosomes, membrane proteins, enzymes, and helical filaments (Figure 3 for selected examples). These structures were made possible primarily due to the revolutionary development of the direct-electron detectors. All of these structures give rise to *de novo* models with clear backbone topology and sidechain densities (Figure 4).

Identifying Different Populations Within a Heterogeneous Sample

GroEL-ES is a molecular chaperone complex that helps other proteins fold correctly in the cell. Each GroEL complex consists of two back-to-back rings, each composed of seven subunits. GroES is a heptameric co-chaperonin that caps the GroEL after the substrate is encapsulated, concurrently with the ATP hydrolysis. Substrate binding and folding is a complex biochemical process. A biochemical mixture of GroEL, GroES, substrate, and ATP is bound to generate multiple reaction products (e.g., GroEL alone, GroEL with GroES capped, GroEL with the substrate, GroEL plus GroES with the substrate in the folding chamber, and the GroES can bind to either one ring or both rings). Such a heterogeneous sample presents challenges in sorting out the structures of the reaction products. Biochemists can carefully control the experimental conditions or use GroEL mutants to reduce the number of different

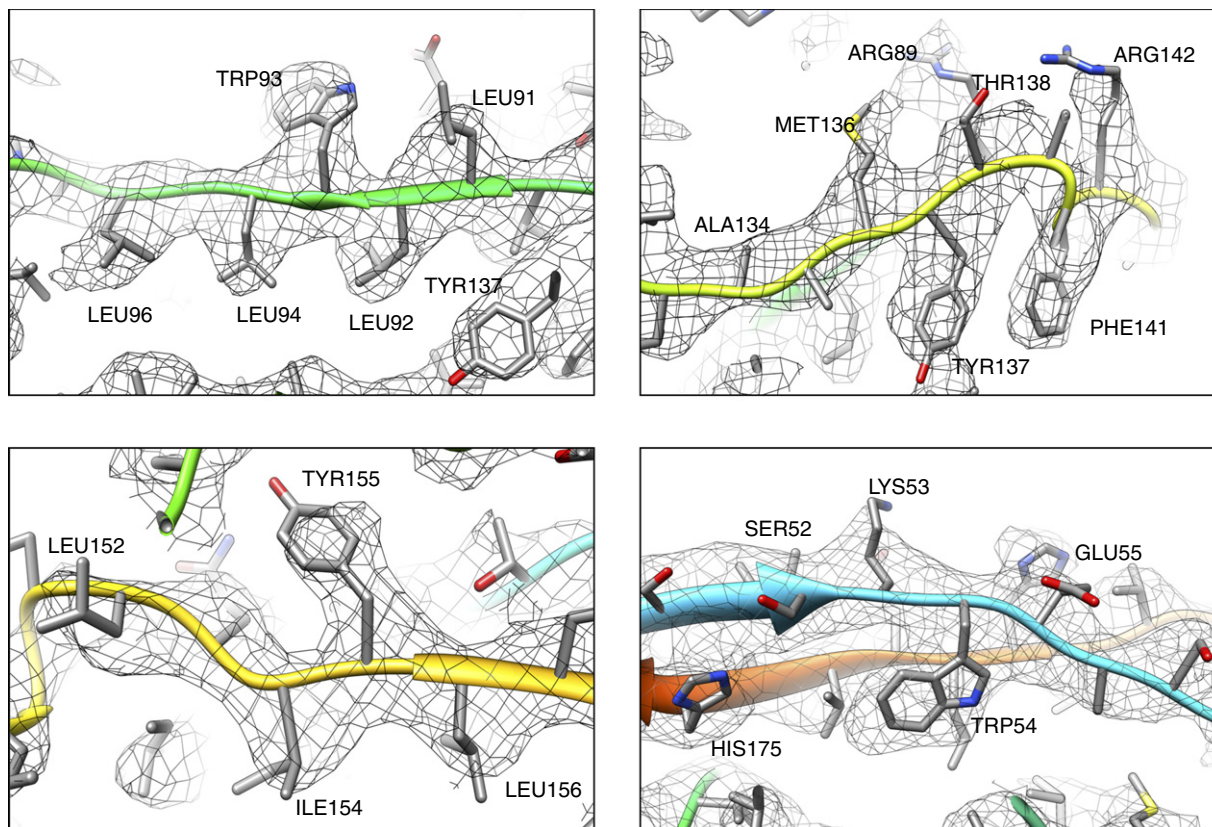


Figure 4 CryoEM maps and models of the Brome Mosaic Virus showing the sidechain densities. Reproduced from Wang, Z., Hryc, C.F., Bammes, B., *et al.*, 2014. An atomic model of brome mosaic virus using direct electron detection and real-space optimization. *Nature Communications* 5, 4808.

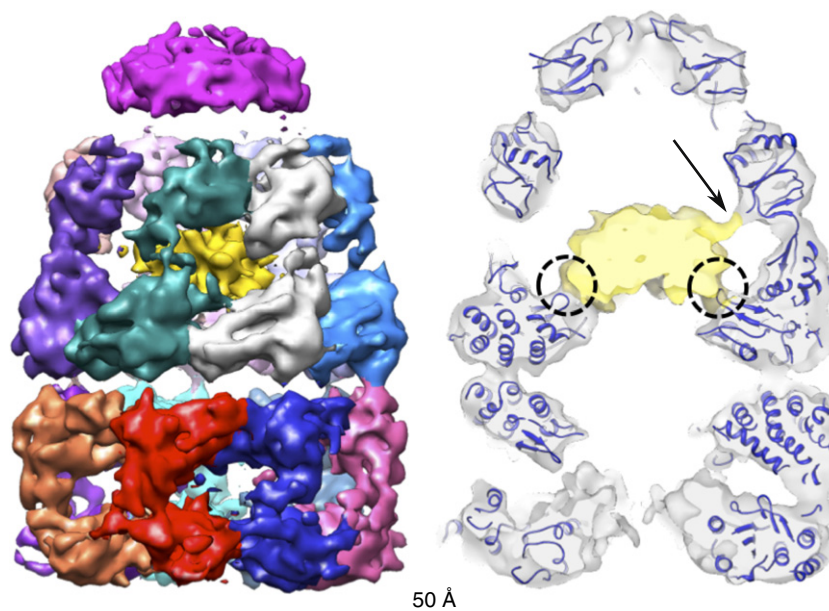


Figure 5 Front view (left) and side view (right) of the cryoEM density map of GroEL/ES complex with the substrate protein RuBisCO in the folding chamber reconstructed without any symmetry enforcement (EMD-2326). Each of the GroEL subunits is colored differently. GroES is colored in purple while the yellow color is the RuBisCO substrate directly interacting with the stem loops of the GroEL. The scale bar denotes 50 Å. Reproduced from Chen, D.H., Madan, D., Weaver, J., *et al.*, 2013a. Visualizing GroEL/ES in the act of encapsulating a folding protein. *Cell* 153, 1354–1365.

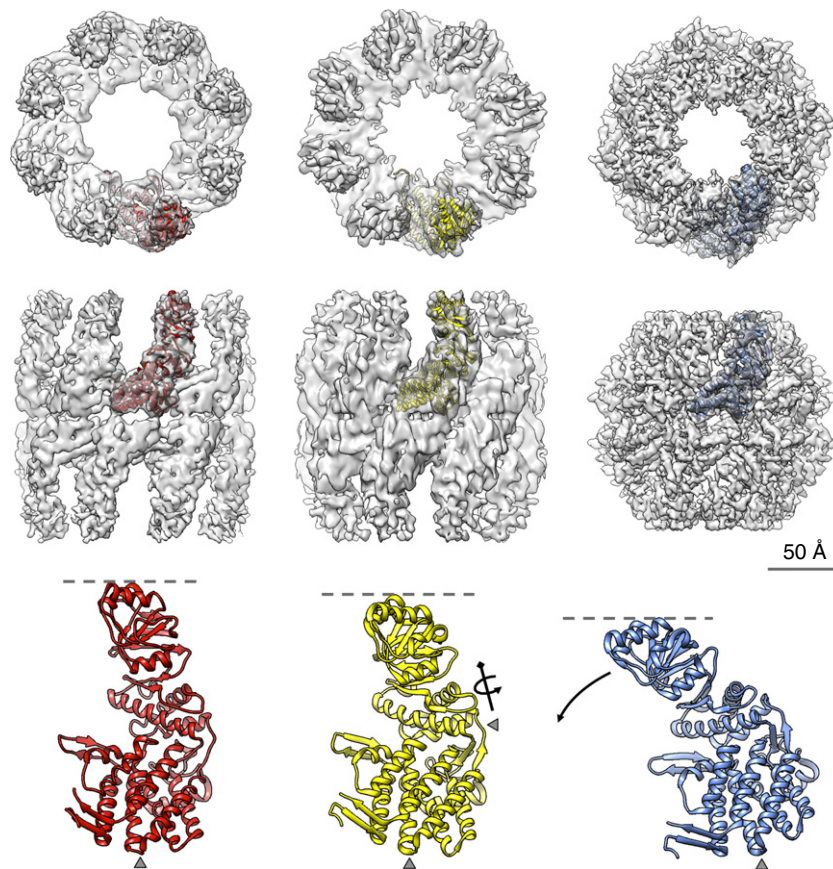


Figure 6 CryoEM maps and models of the lidless Mm-Cpn in the top views (top row), side views (center row), and the model of a single subunit (bottom row). From left to right, the Mm-Cpn is its ATP-free, ATP-bound and ATP-hydrolysis states, respectively. Domains in the Mm-Cpn subunit rotate around different hinges (labeled by the gray triangles) in these different ATP states to close the folding chamber.

conformations within the system (Chen *et al.*, 2013a). Even with many different biochemical approaches to control the chaperone's state there may still be a lot of distinct conformations within the sample solution. Single-particle cryoEM combined with image processing has been used to successfully separate the different conformations and visualize a native substrate RuBisCO inside the GroEL/GroES complex (Chen *et al.*, 2013a). The structure shows the breakdown of the seven- and twofold symmetry of GroEL-ES and the RuBisCO binding the equatorial domains of the GroEL (Figure 5).

Delineating the Conformational Changes of a Molecular Machine

Mm-Cpn is a different chaperonin that can fold proteins without the need of a co-chaperonin (Zhang *et al.*, 2010a). It also has two back-to-back rings but with eight identical subunits per ring. It opens and closes during its ATP-driven cycle. Using single-particle cryoEM, structures were determined for Mm-Cpn under the ATP-free, ATP-bound, and ATP-hydrolysis conditions. Figure 6 shows the cryoEM density maps and models of the Mm-Cpn under different ATP states. The Mm-Cpn slightly closes its folding chamber after ATP-binding and completely closes it upon ATP-hydrolysis (Zhang *et al.*, 2011).

These conformational changes also alter intra- and interring contacts. This mechanism of structural changes in response to ATP is entirely different from those found in GroEL and GroES complex.

Conclusion

The ability of relatively rapid determination of 3D density maps of proteins and protein complexes has led single-particle cryoEM to become an integral part of modern cell biology. It can be anticipated that resolution of cryoEM maps will continue to improve and more effective methodologies to handle particle heterogeneity will be emerged. Already it is supplanting established structural techniques for macromolecules that are not readily analyzed by NMR or X-ray crystallographic methods, such as the human gamma secretase (Lu *et al.*, 2014). It is certain that biochemical protocols to prepare suitable specimens for cryoEM will continue to be developed. Furthermore, the ability to visualize proteins in various conformations, as illustrated with GroEL and Mm-Cpn, enables researchers to elucidate the molecular mechanisms of cellular processes. The ability to model protein in multiple conformations facilitates the ultimate purposes of structural biology: determining its mechanisms of activities through visualizing

its structural intermediates and eventually identifying drug targets.

The recent technological advances in instrumentation, combined with new algorithms for image processing, have enabled single-particle cryoEM to yield reliable high-resolution maps and molecular models of large molecular machines, in which the protein backbone or even amino acid side chains can be identified and modeled. As a result, cryoEM will aid in the understanding of fundamental cellular processes. Structures established by cryoEM can foment drug discovery and lead to a greater comprehension of molecular interactions in biology.

Acknowledgments

This research has been supported by NIH grants (P41GM103832, PN2EY016525), and the Robert Welch Foundation grant (A-1863). JZ would like to acknowledge the Center for Phage Technology and the Department of Biochemistry and Biophysics at Texas A&M University for providing funding.

See also: Imaging the Cell: Electron Microscopy: Scanning Electron Microscopy in Cell Biology

References

- Bai, X.C., Fernandez, I.S., McMullan, G., Scheres, S.H., 2013. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *eLife* 2, e00461.
- Baird, N.J., Ludtke, S.J., Khant, H., *et al.*, 2010. Discrete structure of an RNA folding intermediate revealed by cryo-electron microscopy. *Journal of the American Chemical Society* 132, 16352–16353.
- Baker, M.L., Hryc, C.F., Zhang, Q., *et al.*, 2013. Validated near-atomic resolution structure of bacteriophage epsilon15 derived from cryo-EM and modeling. *Proceedings of the National Academy of Sciences of the United States of America* 110, 12301–12306.
- Baker, M.L., Zhang, J., Ludtke, S.J., Chiu, W., 2010. Cryo-EM of macromolecular assemblies at near-atomic resolution. *Nature Protocols* 5, 1697–1708.
- Bammes, B.E., Jakana, J., Schmid, M.F., Chiu, W., 2010. Radiation damage effects at four specimen temperatures from 4 to 100 K. *Journal of Structural Biology* 169, 331–341.
- Bammes, B.E., Rochat, R.H., Jakana, J., Chen, D.H., Chiu, W., 2012. Direct electron detection yields cryo-EM reconstructions at resolutions beyond 3/4 Nyquist frequency. *Journal of Structural Biology* 177, 589–601.
- Bartesaghi, A., Matthies, D., Banerjee, S., Merk, A., Subramaniam, S., 2014. Structure of beta-galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy. *Proceedings of the National Academy of Sciences of the United States of America* 111, 11709–11714.
- Carragher, B., Kisseberth, N., Kriegman, D., *et al.*, 2000. Leginon: An automated system for acquisition of images from vitreous ice specimens. *Journal of Structural Biology* 132, 33–45.
- Chang, J., Liu, X., Rochat, R.H., Baker, M.L., Chiu, W., 2012. Reconstructing virus structures from nanometer to near-atomic resolutions with cryo-electron microscopy and tomography. *Advances in Experimental Medicine and Biology* 726, 49–90.
- Chen, D.H., Luke, K., Zhang, J., Chiu, W., Wittung-Stafshede, P., 2008. Location and flexibility of the unique C-terminal tail of Aquifex aeolicus co-chaperonin protein 10 as derived by cryo-electron microscopy and biophysical techniques. *Journal of Molecular Biology* 381, 707–717.
- Chen, D.H., Madan, D., Weaver, J., *et al.*, 2013a. Visualizing GroEL/ES in the act of encapsulating a folding protein. *Cell* 153, 1354–1365.
- Chen, D.H., Song, J.L., Chuang, D.T., Chiu, W., Ludtke, S.J., 2006. An expanded conformation of single-ring GroEL-GroES complex encapsulates an 86 kDa substrate. *Structure* 14, 1711–1722.
- Chen, S., McMullan, G., Faruqi, A.R., *et al.*, 2013b. High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* 135, 24–35.
- Cong, Y., Schroder, G.F., Meyer, A.S., *et al.*, 2011. Symmetry-free cryo-EM structures of the chaperonin TRiC along its ATPase-driven conformational cycle. *EMBO Journal* 31, 720–730.
- DiMaio, F., Tyka, M.D., Baker, M.L., Chiu, W., Baker, D., 2009. Refinement of protein structures into low-resolution density maps using rosetta. *Journal of Molecular Biology* 392, 181–190.
- DiMaio, F., Zhang, J., Chiu, W., Baker, D., 2013. Cryo-EM model validation using independent map reconstructions. *Protein Science* 22, 865–868.
- Elad, N., Clare, D.K., Saibil, H.R., Orlova, E.V., 2008. Detection and separation of heterogeneity in molecular complexes by statistical analysis of their two-dimensional projections. *Journal of Structural Biology* 162, 108–120.
- Elmlund, D., Elmlund, H., 2012. SIMPLE: Software for ab initio reconstruction of heterogeneous single-particles. *Journal of Structural Biology* 180, 420–427.
- Grigorieff, N., 2007. FREALIGN: High-resolution refinement of single particle structures. *Journal of Structural Biology* 157, 117–125.
- van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., Schatz, M., 1996. A new generation of the IMAGIC image processing system. *Journal of Structural Biology* 116, 17–24.
- Henderson, R., 1995. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Quarterly Reviews of Biophysics* 28, 171–193.
- Henderson, R., Sali, A., Baker, M.L., *et al.*, 2012. Outcome of the first electron microscopy validation task force meeting. *Structure* 20, 205–214.
- Hinsen, K., Reuter, N., Navaza, J., Stokes, D.L., Lacapere, J.J., 2005. Normal mode-based fitting of atomic structure into electron density maps: Application to sarcoplasmic reticulum Ca-ATPase. *Biophysical Journal* 88, 818–827.
- Jiang, W., Baker, M.L., Jakana, J., *et al.*, 2008. Backbone structure of the infectious epsilon15 virus capsid revealed by electron cryomicroscopy. *Nature* 451, 1130–1134.
- Kelly, D.F., Dukovski, D., Walz, T., 2010. Strategy for the use of affinity grids to prepare non-His-tagged macromolecular complexes for single-particle electron microscopy. *Journal of Molecular Biology* 400, 675–681.
- Kucukelbir, A., Sigworth, F.J., Tagare, H.D., 2014. Quantifying the local resolution of cryo-EM density maps. *Nature Methods* 11, 63–65.
- Lei, J., Frank, J., 2005. Automated acquisition of cryo-electron micrographs for single particle reconstruction on an FEI Tecnai electron microscope. *Journal of Structural Biology* 150, 69–80.
- Leitner, A., Joachimiak, L.A., Bracher, A., *et al.*, 2012. The molecular architecture of the eukaryotic chaperonin TRiC/CCT. *Structure* 20, 814–825.
- Leschziner, A.E., Nogales, E., 2006. The orthogonal tilt reconstruction method: An approach to generating single-class volumes with no missing cone for ab initio reconstruction of asymmetric particles. *Journal of Structural Biology* 153, 284–299.
- Li, X., Mooney, P., Zheng, S., *et al.*, 2013. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nature Methods* 10, 584–590.
- Liao, M., Cao, E., Julius, D., Cheng, Y., 2013. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* 504, 107–112.
- Liu, X., Jiang, W., Jakana, J., Chiu, W., 2007. Averaging tens to hundreds of icosahedral particle images to resolve protein secondary structure elements using a Multi-Path Simulated Annealing optimization algorithm. *Journal of Structural Biology* 160, 11–27.
- Lu, P., Bai, X.C., Ma, D., *et al.*, 2014. Three-dimensional structure of human gamma-secretase. *Nature* 512, 166–170.
- Ludtke, S.J., Baker, M.L., Chen, D.H., *et al.*, 2008. De novo backbone trace of GroEL from single particle electron cryomicroscopy. *Structure* 16, 441–448.
- Ludtke, S.J., Baldwin, P.R., Chiu, W., 1999. EMAN: Semiautomated software for high-resolution single-particle reconstructions. *Journal of Structural Biology* 128, 82–97.
- Murakami, K., Elmlund, H., Kalisman, N., *et al.*, 2013. Architecture of an RNA polymerase II transcription pre-initiation complex. *Science* 342, 1238724.
- Murray, S.C., Flanagan, J., Popova, O.B., *et al.*, 2013. Validation of cryo-EM structure of IP(3)R1 channel. *Structure* 21, 900–909.
- Nakamura, N., Shimizu, Y., Shinkawa, T., *et al.*, 2010. Automated specimen search in cryo-TEM observation with DIFF-Defocus imaging. *Journal of Electron Microscopyscopy (Tokyo)* 59, 299–310.
- Penczek, P.A., Kimmel, M., Spahn, C.M., 2011. Identifying conformational states of macromolecules by eigen-analysis of resampled cryo-EM images. *Structure* 19, 1582–1590.

- Penczek, P.A., Yang, C., Frank, J., Spahn, C.M., 2006. Estimation of variance in single-particle reconstruction using the bootstrap technique. *Journal of Structural Biology* 154, 168–183.
- Pereira, J.H., Ralston, C.Y., Douglas, N.R., *et al.*, 2010. Crystal structures of a group II chaperonin reveal the open and closed states associated with the protein folding cycle. *Journal of Biological Chemistry* 285, 27958–27966.
- Radermacher, M., Wagenknecht, T., Verschoor, A., Frank, J., 1987. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *Journal of Microscopy* 146, 113–136.
- Rosenthal, P.B., Henderson, R., 2003. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of Molecular Biology* 333, 721–745.
- Russo, C.J., Passmore, L.A., 2014. Controlling protein adsorption on graphene for cryo-EM using low-energy hydrogen plasmas. *Nature Methods* 11, 649–652.
- Scheres, S.H., Chen, S., 2012. Prevention of overfitting in cryo-EM structure determination. *Nature Methods* 9, 853–854.
- Scheres, S.H., Gao, H., Valle, M., *et al.*, 2007. Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nature Methods* 4, 27–29.
- Schröder, G.F., Brunger, A.T., Levitt, M., 2007. Combining efficient conformational sampling with a deformable elastic network model facilitates structure refinement at low resolution. *Structure* 15, 1630–1641.
- Shaikh, T.R., Gao, H., Baxter, W.T., *et al.*, 2008. SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nature Protocols* 3, 1941–1974.
- Shukla, A.K., Westfield, G.H., Xiao, K., *et al.*, 2014. Visualization of arrestin recruitment by a G-protein-coupled receptor. *Nature* 512, 218–222.
- Tama, F., Miyashita, O., Brooks III, C.L., 2004. Normal mode based flexible fitting of high-resolution structure into low-resolution experimental data from cryo-EM. *Journal of Structural Biology* 147, 315–326.
- Tang, G., Peng, L., Baldwin, P.R., *et al.*, 2007. EMAN2: An extensible image processing suite for electron microscopy. *Journal of Structural Biology* 157, 38–46.
- Topf, M., Lasker, K., Webb, B., *et al.*, 2008. Protein structure fitting and refinement guided by cryo-EM density. *Structure* 16, 295–307.
- Trabuco, L.G., Villa, E., Mitra, K., Frank, J., Schulten, K., 2008. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* 16, 673–683.
- Velazquez-Muriel, J., Lasker, K., Russel, D., *et al.*, 2012. Assembly of macromolecular complexes by satisfaction of spatial restraints from electron microscopy images. *Proceedings of the National Academy of Sciences of the United States of America* 109, 18821–18826.
- Wang, Z., Hryc, C.F., Bammes, B., *et al.*, 2014. An atomic model of brome mosaic virus using direct electron detection and real-space optimization. *Nature Communications* 5, 4808.
- Yu, X., Jin, L., Zhou, Z.H., 2008. 3.88 Å structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy. *Nature* 453, 415–419.
- Zhang, J., Baker, M.L., Schroder, G.F., *et al.*, 2010a. Mechanism of folding chamber closure in a group II chaperonin. *Nature* 463, 379–383.
- Zhang, J., Ma, B., DiMaio, F., *et al.*, 2011. Cryo-EM structure of a group II chaperonin in the prehydrolysis ATP-bound state leading to lid closure. *Structure* 19, 633–639.
- Zhang, J., Minary, P., Levitt, M., 2012. Multiscale natural moves refine macromolecules using single-particle electron microscopy projection images. *Proceedings of the National Academy of Sciences of the United States of America* 109, 9845–9850.
- Zhang, J., Nakamura, N., Shimizu, Y., *et al.*, 2009. JADAS: A customizable automated data acquisition system and its application to ice-embedded single particles. *Journal of Structural Biology* 165, 1–9.
- Zhang, P., Borgnia, M.J., Mooney, P., *et al.*, 2003. Automated image acquisition and processing using a new generation of 4 K × 4 K CCD cameras for cryo electron microscopic studies of macromolecular assemblies. *Journal of Structural Biology* 143, 135–144.
- Zhang, X., Jin, L., Fang, Q., Hui, W.H., Zhou, Z.H., 2010b. 3.3 Å cryo-EM structure of a nonenveloped virus reveals a priming mechanism for cell entry. *Cell* 141, 472–482.
- Zhang, X., Settembre, E., Xu, C., *et al.*, 2008. Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction. *Proceedings of the National Academy of Sciences of the United States of America* 105, 1867–1872.

Relevant Website

<http://www.emdatabank.org/>
EMDatabank.

Scanning Electron Microscopy in Cell Biology

TD Allen, University of Manchester, Manchester, UK
MW Goldberg, Durham University, Durham, UK

© 2016 Elsevier Inc. All rights reserved.

Introduction

Long before the idea of scanning specimens with a small spot of light produced confocal light microscopy, the idea of using a small spot of electrons to scan surfaces had been around as long as electron microscopy (EM) itself. A surface demarcates the boundary of a solid and is the site of interaction with the surrounding environment, from a living cell to a ball bearing. In the mechanical world, adhesion, friction, wear, and corrosion are all dependent on surface properties. Living material, even those which have a solid nature such as bone, have surface properties infinitely more dynamic than the humble ball bearing. In a unicellular tissue such as blood, the white cells which are responsible for our immune defenses rely on surface molecules which are constantly recycled and redistributed to recognize and respond to infective agents such as bacteria and viruses. In multicellular tissues, information is incessantly passed between neighboring cells via their surfaces, and the progression of nerve impulses relies on electrical charge variation across the surface of neurons. Within the cell, most catalytic reactions are brought about by enzymes acting at the surface of membranes. All adhesion relies on surface properties, be it interactions between a tire and the road surface (enough friction to move the car or enough adhesion to prevent skidding?). Modifications of cell membranes such as formation of desmosomes in multicellular tissues provide remarkable mechanical strength and flexibility at the surfaces of both skin and internal organs. Although microscopy, in general, and scanning electron microscopy (SEM), in particular, are not the only ways of characterizing surfaces, they are the cornerstone of understanding their structure and properties.

Scanning Electron Microscopy 'Versus' Transmission Electron Microscopy

Rather than passing a 'flood' beam of electrons through a thin sliver of material as in transmission EM (TEM), SEMs move a small spot of electrons to and fro in a pattern of horizontal lines over the specimen surface (a raster). It is important to note that resolution in the SEM is not determined by the wavelength of electrons, as it is in the TEM, but by the diameter of the spot of electrons. It is this area (and the depth to which the beam penetrates) which generates the 'interaction volume' from which the signal is produced. As the beam diameter is reduced, so is the interaction volume, and resolution is increased. This constraint caused SEM to be considered 'low resolution' by TEM purists, and it took until the mid-1980s with the first commercial in-lens, field emission instrument from Hitachi, the S-900, for SEM to come close to the resolution of the TEM. However this type of instrument remained an expensive rarity, confined to only a few biology labs around the world and it is only relatively recently that

such instruments have become more widely available to cell biologists. Thus although SEM cannot claim the same massive strides forward in understanding of cell structure provided by Palade, Porter, Fawcett, and other TEM pioneers in the 1950s and 1960s, what it has done is to turn the 2D black line representation of membranes visualized in the TEM into a 3D surface representation of unimagined complexity (Figures 1(a) and 1(b)).

Construction of the SEM

From the top of the column to the specimen, there is little difference between a SEM and a TEM. Both have a source of electrons at the top of an evacuated column, and a set of condenser lenses to focus the electron beam onto the specimen. First-generation SEMs used the same source of electrons as TEMs, a 'V' shaped tungsten wire heated to around 1500 °C. Heating provides the energy required to remove electrons from the atoms in the tungsten, overcoming a barrier called the work function. If the amount of extracted electrons can be increased then a brighter, smaller spot can be used to scan the specimen, increasing resolution. The first improvement in electron sources was to use a crystal of Lanthanum hexaboride (LaB6), which has a lower work function than tungsten, increasing the available intensity tenfold. The next stage, which produced a further hundredfold improvement in intensity over LaB6, was to use a field emission gun (FEG). In field emission, electrons are extracted from an ultrafine needlepoint of tungsten by the application of a powerful electric field close to the tip. This overcomes the work function in a process called electron tunneling. This requires a vacuum of two orders of magnitude greater than conventional electron sources, requiring ion pumps. FEGs can be either 'cold' (room temperature), or heated. Cold FEGs are the most effective, but can become unstable as a single gas molecule can contaminate the tip, causing instability and a reduction in emission. A compromise with better stability and lower vacuum requirements, but resulting in slightly lower resolution, is thermal (or Schottky) field emission, in which the filament tip is both heated (which repels contaminants) as well as being subjected to an intense field. Most top-range SEMs (and many TEMs) have field emission sources, and are termed FEG-SEMs or FEG-TEMs (Figure 2).

Scanning the Specimen

Having generated an intense beam of electrons, the next stage in SEM is the point at which everything changes relative to TEM. Rather than illuminating the whole specimen in a 'flood beam,' electrons are concentrated into a small spot (demagnified) by two condenser lenses, and focused onto the surface

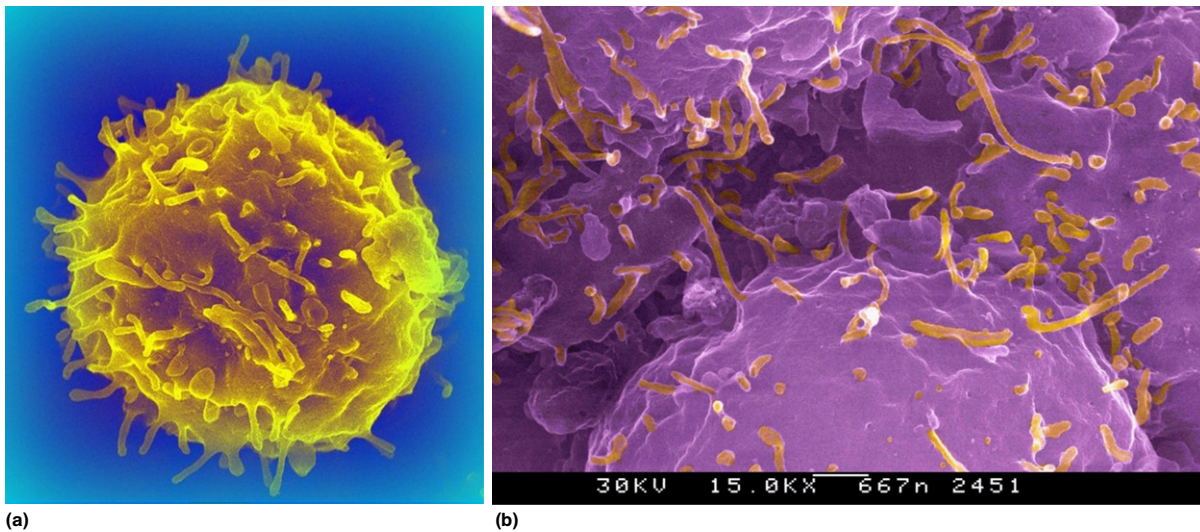


Figure 1 (a) Scanning EM visualization of a bone marrow stem cell (false colored). This typical visualization of a whole cell demonstrates both the surface and three-dimensional information available from this technique. (b) Higher magnification view of the surface of a human cell grown in tissue culture, showing the variation in surface projections termed microvilli, and the external membrane topography. The magnification ($15\,000\times$) shown is what the image was recorded at, and will vary according to the final size of the image, whereas the scale bar (667 nm) remains accurate.



Figure 2 A modern FEG-SEM instrument showing the column and the adjacent desk area, where the image is viewed on a monitor. The electron gun is mounted at the top of the column, evacuated by a series of ion pumps to produce a vacuum in which field emission of an electron beam is maintained.

of the specimen by the final lens directly above the specimen, sometimes (inappropriately) called the objective lens. The spot of electrons is moved across the specimen in a raster by a set of scanning coils which are usually contained within the final lens. These are four radially orientated electromagnets arranged so that their magnetic fields are perpendicular to the

axis of the beam. By varying the current, the beam can be controlled to scan across the chosen area of the specimen, building an image dot by dot and line by line. Magnification is determined by the chosen area of the scan as the final image is produced on a fixed size monitor. Thus if the scanned area is reduced by half, then the magnification will double. Most instruments will allow modest magnifications at the bottom end of their range (times 20 or so) up to 200 000 or even 3 million times in research grade microscopes. This massive range is useful with biological material, enabling selection of areas which have clean surfaces without physical damage, and to give confidence that the areas which are investigated are truly representative of the sample and to give an overall perspective. Image formation can be distorted by astigmatism, caused by lens imperfections when the beam becomes oval in cross-section rather than circular. This is corrected by the operator using a set of eight electromagnetic coils (stigmators). The final lens determines resolution by its formation of a minimum spot size (at best, 0.5–1 nm).

Beam – Specimen Interactions

If the scanned electron probe is ‘frozen’ at any one moment, then we can consider how the beam interacts with the specimen. Unlike TEM where the specimen has been suitably thinned to allow the beam to travel through, the bulk nature of specimens in SEM prevents this. The beam does penetrate the specimen surface to a depth that depends on the velocity of the electrons in the beam and the nature of the material in the sample but is generally around $1\,\mu\text{m}$, producing a volume of interaction shaped like a teardrop (Figure 3).

The collisions between the electrons from the beam and the atoms in the specimen surface lead to two types of interactions. Zero energy loss interactions result in elastic scattering

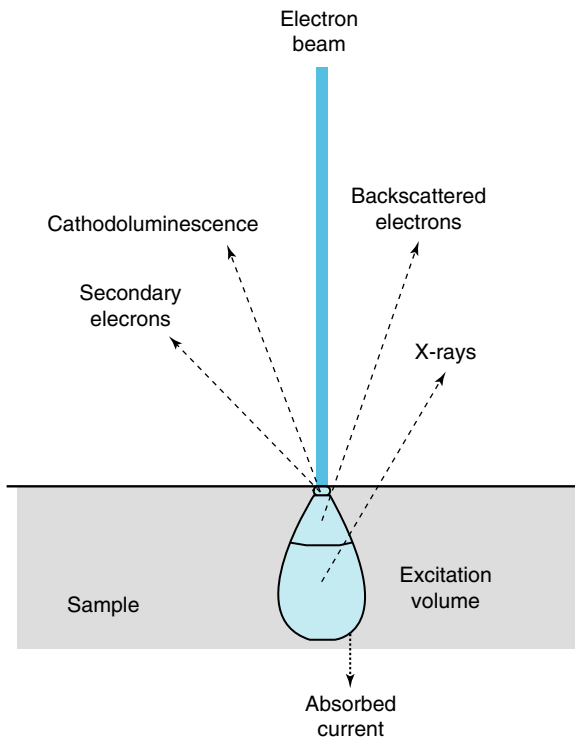


Figure 3 Interactions between the electron beam and the specimen surface. The beam penetrates the surface producing a 'teardrop' shaped interaction volume in which a variety of collisions take place.

where the primary beam electrons are simply deflected as they pass through the sample. In SEM, because the detectors are usually above the sample, only the most extremely elastically scattered electrons (by $\sim 180^\circ$), backscatter electrons, are of interest. If an electron transfers energy to the sample this is known as inelastic scattering. Inelastic scattering results in a change in velocity (and hence wavelength) of the electron, which in itself is more important for TEM, but the transfer of energy results in the production of secondary electrons, and various types of radiation, X-rays and cathodoluminescence (visible light), all of which can be used in the SEM to gain information about the sample.

Secondary electrons are formed as a result of inelastic interactions with atomic electron clouds in the specimen where substantial energy transfer results in the ejection of electrons from the sample atoms. Secondary electrons, which originate from the sample (not the high energy beam) have relatively low energies and are generally the most important signal produced for high-resolution surface imaging. Backscattered electrons, also important in biology, are the result of an elastic interaction with an atomic nucleus in the sample and are high-energy primary beam electrons that are scattered backwards, 'reflected,' due to attraction to the atomic nuclei in the specimen so that they reemerge from the sample surface with high energy.

Image Formation

Images are formed by modulating the brightness on a monitor (which is scanned in synchrony with the electron beam)

according to the interactions generated by the beam moving over the specimen. The secondary electron signal is collected by a detector which is positioned adjacent to the specimen in the chamber, surrounded by a wire grid kept at 200 V (in order to attract the secondary electrons) in front of a scintillator (kept at 10 000 V). Having passed through the grid, a secondary electron strikes the surface of the scintillator to produce a photon, which travels down a light pipe to a photomultiplier which in turn converts photons into an electrical signal. This signal determines the brightness of the spot on the monitor to create the image. Thus the more secondary electrons emitted from any one spot on the surface of the specimen, the greater the signal from the detector, and the brighter the area seen at that point in the image on the monitor.

Backscattered electrons also contribute to image formation, and can be imaged alone by either reducing the voltage on the grid in the detector (as backscattered electrons have sufficient energy to reach the scintillator regardless), or they can be collected more efficiently by a dedicated solid state detector, usually placed directly above the sample to ensure maximum collection. Backscattered electrons are useful, as they can be used to map the distribution of elements in the specimen because the number of backscattered electrons is proportional to the atomic number of the element which scattered them. This property therefore gives contrast dependent on the atomic number of the material. This is useful in biology because antibodies, which can be used to label specific proteins and other macromolecules, can be tagged with gold nanoparticles, which have a high atomic number and can easily be distinguished from the low atomic number atoms predominant in biological material. As they tend to arise from deeper within the surface of the specimen however, backscatter electrons are not as good as secondary electrons for resolving surface topography. Backscatter electrons can also undergo inelastic interactions as they emerge from the surface, producing type II secondary electrons which add to the noise in the secondary image. However, because they are much less affected by charge build up in nonconductive samples than secondary electrons, they may be preferred for imaging of some samples if enough electrons are backscattered to form an image on their own.

Another signal comes from 'Auger' electrons (after the person who discovered them), emitted a very short distance from the surface of the specimen, where they are mostly absorbed, but do allow chemical mapping of surfaces to a high resolution (with a dedicated detector). The X-ray emission (bremsstrahlung) from a scanned specimen can be used for energy dispersive elemental analysis similar to TEM. Finally, cathodoluminescence (the production of visible light) can also be detected and although rarely used in biology, is useful in imaging the internal structure of material science specimens such as semiconductors.

A Brief History of SEM

SEM was considered as a possibility by Ruska himself (the 'father' and Nobelist for TEM) in the very early development of EM, as methods for making thin specimens for TEM only emerged later. In 1935, Knoll invented a scanning instrument to study the targets of television camera tubes, to be followed by scanning instruments produced in 1938 and 1942 by von

Ardenne and Zworykin, respectively. Von Ardenne's instrument was destroyed in an air raid on Berlin in 1944. In the USA, the RCA corporation produced TEMs, and were working on scanning instruments. However, largely because of difficulties in the vacuum technology of the time, together with the success of surface imaging using replicas in TEM, they never brought a SEM to market. Consequently, we had to wait until 1965 for the first commercial instrument, produced by the Cambridge Instrument Company in the UK, as a result of research started in 1948 by Charles Oatley and his team. Six months later this was followed by an instrument produced by the Japanese Electron Optical Corporation (JEOL). From its earliest days, SEM has proved to be a source of images that everybody could relate to, regardless of microscopic experience or scientific background. From early images showing great detail of everyday objects and animals, for example, the edge of a scalpel or razor blade or the multiple eyes of a spider, the extra information provided by the high magnification was instantly apparent, grasping the attention of the general public in a way that TEM images did not. Today, images of bacteria, stem cells, and tumor cells are a regular sight in TV news, documentaries, newspapers and magazines, usually brightly colored. False or pseudo-coloring of SEM images is useful for highlighting specific features, as well as increasing the overall impact (albeit sometimes on the garish side).

Specimen Preparation for SEM

An ideal specimen for SEM is a coin: it is dry (no water content), has a hard surface, is made of heavy elements, and is an electrical conductor. These properties ensure that there is no problem with the vacuum in the specimen chamber, a good signal will be produced by the interaction of the electron beam at the surface, and that any electrical charge from the beam can run away to earth due to the conductive nature of the coin. Contrast this with biological tissue, which is ~70% water, which would evaporate and degrade the vacuum in the specimen chamber at the same time as distorting the specimen as it dries. The cell surface is soft, allowing the electron beam to penetrate deeper, masking surface detail and limiting resolution together with heat damage from the beam. On top of this, the light elemental composition of our specimens (made up of carbon, hydrogen, oxygen, and nitrogen) means there will be little useful signal to form images. Finally, because biological tissue is a poor conductor, electrons will build up within the specimen until they have sufficient energy to find a way to earth, resulting in a discharge which generates a burst of signal brightness (known as charging) which obliterates what image there might have been in the first place. All in all, not the most promising starting point for high-resolution imaging.

Critical Point Drying

These obstacles need to be circumvented to observe biological material. Firstly specimen water is removed. However, drying soft biological surfaces is tricky, as the surface tension of the receding liquid layer during drying will damage fine (and fragile) surface structures. Liquids, however, have a 'critical

point' where both vapor and liquid states coexist, allowing surface tension effects to be avoided if vapor is removed above this critical point. The critical point for water requires conditions that are not compatible with biological specimens, a pressure of ~200 atm and a temperature of 374 °C. The critical point for CO₂, however, is 31.1 °C, at a pressure of 100 atm, which is compatible. Samples must be fixed (in a similar way to TEM samples), which means introducing extensive chemical cross-linking between the molecules in the sample to prevent extraction and reorganization during the subsequent steps, followed by replacement of the water in the specimen by solvent (e.g., ethanol), which is in turn replaced by liquid CO₂ in a pressure vessel. As the temperature is raised in the pressure vessel, the critical point is reached, at which point the meniscus at the surface of the liquid CO₂ disappears as all the CO₂ is converted to vapor. This gaseous CO₂ is then gently vented, and the specimen is now dried, free from the distorting effects of surface tension.

Sputter Coating

The specimen has retained its fine structure, but still has a light elemental composition (C, H, O, N) which will only produce a poor signal. These limitations are overcome by the deposition of a very thin layer of metal over the surface. For surface replicas made for TEM, the metal is evaporated from a point source to produce a shadowing effect (see Heuser, 2011), but for SEM an even coating is required. This is achieved by a process called sputter coating, in which a target disc of the coating metal is bombarded under vacuum by ionized gas particles (usually argon) which strike the surface and remove atoms which are then deposited evenly over the surface of the specimen. Gold, or an alloy of gold and palladium, produces excellent secondary electron emission, as well as being permanent and noncorroding, allows long storage of samples for future reference. With the highest resolution instruments however, the actual grains of gold in the coating can be resolved – resulting in the imaging of the coating itself rather than the underlying surface of interest. In these situations chromium (Apkarian, 1994) or tungsten is used, which is deposited with a smaller grain size (although at a slightly reduced secondary electron emission). For studies involving elemental contrast (using backscattered electrons at low resolution) the specimen surface is made conductive by sputtering a thin layer of carbon, which appears transparent in the final image.

Labeling for SEM

As with TEM, small metal particles, such as colloidal gold, are used as markers to show specific binding sites on biological tissue (for instance by coupling the gold to an antibody). In SEM a gold label is visualized as a bright sphere (rather than the black dot seen in TEM), usually using a backscatter detector which produces a high contrast image due to the difference in the amount of backscattered electrons produced by the high atomic number of gold relative to the surrounding material. As gold labeling and gold coating would produce similar

backscatter signals, carbon coating is usually used to allow the colloidal gold signal to stand out. To retain the best surface imaging, chromium coating (atomic number 24) does allow visualization of gold colloid (atomic number 79). Also, as the backscattered electron signal from gold is particularly strong, the backscatter image can be collected simultaneously with the secondary electron signal, so that both can be merged in the computer to give the clearest labeling as well as optimal surface imaging.

Cryoscanning EM

An alternative way in which the soft, wet, and largely non-conductive nature of biological material can be made dry and hard is by freezing, which is best done as rapidly as possible, converting aqueous water to vitreous ice in milliseconds, and avoiding ice crystal damage/effects, as with Cryo TEM methods. For Cryo SEM the frozen specimen is mounted within the specimen chamber on a cold stage (cooled by liquid nitrogen) often with a facility to fracture the sample with a cold scalpel blade, and a sputter coater which allows the freshly fractured surface to be coated with metal whilst still in the microscope vacuum system. This approach allows SEM observation of material that has been neither chemically fixed nor dehydrated.

Environmental SEM

The preparative requirements for SEM would seem to be difficult to get round, but the ingenuity of instrument manufacturers should not be underestimated. The simplest solution, with respect to sample preparation, however, is to completely remove the requirement for a vacuum environment in the specimen chamber. Although an electron beam can only be generated in a vacuum, once the electrons have been accelerated, they will continue to travel in a directed manner, and can still interact with a specimen even if that specimen is at atmospheric pressure (i.e., no vacuum). The way that electrons can be used for imaging at 'high pressure' (up to 100% humidity) is to generate a vacuum gradient between the electron source and the specimen chamber. Although the electron beam will begin to scatter as soon as it encounters significant gaseous molecules on its journey down the column, there are still sufficient interactions with the specimen to generate secondary and backscattered electrons. Secondary electrons cannot be detected with standard detectors because of the presence of gas in the chamber. However the use of a positively charged detector results in acceleration of the secondary electrons toward the detector. These electrons do not reach the detector but collide with gas particles, displacing further electrons from the gas atoms, which themselves are accelerated toward the detector, undergoing more collisions and amplifying the signal, which eventually reaches the detector.

Environmental SEM therefore can be used to image non-conductive, uncoated, even 'wet' samples (Donald, 2003). Resolution in environmental microscopes suffers because of the nature of image formation in a 'wet' or gas-filled chamber, but this is offset by the ability to view wet specimens without

pretreatment. Future improvements in imaging quality may make this an attractive proposition for 'au natural' cellular imaging by SEM.

Serial Block Face Imaging

A major recent advance involving environmental SEMs, which takes advantage of the ability to suppress charging is serial block face imaging (Hughes *et al.*, 2014). A major limitation of thin section TEM is that it is essentially a 2D technique. Despite invaluable cross-sectional detail, translating this into the 3D context of a whole cell, tissue, or organism can be difficult and in the cases of serial section reconstruction and EM tomography, is very time consuming. Samples for thin section TEM are fixed, dehydrated with a solvent, in a similar way to SEM samples, but then instead of drying, they are infiltrated with a resin, which is then polymerized, to provide a hard support from which extremely thin sections can be cut and imaged. Such sections are fragile and display artifacts such as compression, altering the dimensions of biological material within, making reconstruction from a series of sections even more difficult. It was found, however, that TEM-like images could be obtained by imaging the cut face of the resin block (instead of the section obtained from it) with a backscatter detector in a SEM. Because the resin is nonconductive, the use of an environmental SEM is required to prevent electron charge build up in the specimen. In this technique an ultramicrotome is placed inside the SEM, so that sections can be trimmed off the block face and discarded, revealing a new face deeper in the sample, which is then imaged. This can be automated so that hundreds to thousands of block faces can be imaged progressively deeper into the specimen, effectively providing cross-sectional images from the whole volume of a cell or tissue. These can then be reconstructed into 3D structures using computer software.

Merging SEM and TEM

When a thin specimen is viewed in a SEM, the electron beam will pass through it. If the detector is placed underneath (rather than above as usual) an image can be gathered that is similar to that of a TEM. From the early years of SEMs, this was well known, but rarely used, as the resolution would be relatively poor. However, as the resolution and quality of imaging in SEMs improved, certain advantages to this type of imaging came into consideration. In a TEM, the projector lenses which form the image below the specimen can suffer from blurring due to inelastically scattered electrons causing chromatic aberration. This does not occur in scanning instruments as there are no lenses below the specimen. Also, images can be collected in the SEM with digital detectors of almost unlimited pixel numbers (an enormous 32 K by 32 K in the Zeiss Atlas system). Finally, when the magnification is changed in the SEM there is no image rotation as there is in the transmission microscope. All these factors have been utilized in biological samples to produce impressive results in tomographical (3D) reconstructions at the cellular level. However, in terms of absolute (atomic) levels of resolution, the TEM is still the

instrument of choice, although many TEMs will also incorporate a scanned beam illumination to take advantage of this mode of imaging, termed STEM (Scanning Transmission EM).

Exposing Subsurface Detail

Having collected surface details by SEM, the possibility to then strip back this layer and expose underlying structures was a thought that occurred to most microscopists. Simple approaches, such as the use of adhesive tape (applied after fixation, gentle detergent extraction, and refixation followed by dehydration and critical point drying), at which point the cell is in a friable condition, produce an effect rather like taking the pastry off a pie, and exposing various levels below (Figure 4).

Although electron beam lacks the power to remove any material in the microscope, if electrons are substituted with something of heavier mass, and enough energy provided, then surface material can be stripped away little by little (termed 'milling'). This is achieved by a technique which uses a beam of focused ions generated from a source in which a tungsten filament coated with gallium is heated under vacuum. The molten gallium flows to the tip of the tungsten needle where gallium atoms are subjected to an electric field which produces ionization and field emission of gallium atoms. An atom of gallium, compared to an electron, is like a cannonball compared to an air rifle pellet. This beam bombards the specimen

surface removing atoms and also producing secondary electrons, often sufficient to produce an image. The amount of material removed can be controlled by varying the beam current. At low currents, little is removed, but the high density of signal allows high-resolution imaging. At high beam currents, large amounts can be removed allowing precision milling of the surface micron by micron. This technique is mainly used in semiconductor production and research, using 'dual beam' instruments (FIB-SEMs) which combine SEM imaging with a focused ion beam (FIB) for micromachining. Specimens are imaged with secondary electrons, then 'stripped back' using the ion beam, and imaged again. This can be repeated over and over, often under remote control, until a series of images of successive 'slices' have been recorded (rather like photographing the Sunday joint after each slice has been cut). These images can then be reconstructed in the computer to form a 3D map. The ion beam itself can be used to produce images, which may lack the ultimate resolution of electrons, but can be extremely useful as they show better elemental contrast of metallic grains in alloys and in mechanical failure and corrosion. As well as material science specimens, blocks of fixed and resin-embedded biological tissues and also frozen biological material can be milled to reveal fresh surfaces (Kizilyaprak *et al.*, 2014). Muscle tissue has been reduced by milling to a thin section for TEM, with promising results. The data obtained can be considered as similar to serial block face imaging, although the process and some of the limitations and advantages differ. The introduction of various reactive gases into the specimen chamber also allows chemical modification of the environment in which material is etched.

SEM of Subcellular Detail

SEM of internal cell structure may appear nonintuitive, given the familiarity of SEM images of whole cells and tissues viewed from the surface, but separation of cell contents (largely for biochemical analyses) has been an integral part of cell biology since its inception. Isolated organelles such as mitochondria or other vesicular components can be purified and visualized by SEM, or even a molecular level of purification for cytoskeletal elements such as tubulin, which can then be repolymerized *in vitro* to form intact microtubules and viewed by SEM. The resolution limits of FEG-SEM were pushed in the early 1990s by Martin Müller and colleagues, who produced stunning images of bacteriophage substructure, DNA, Fab fragments, and ultra-small immuno-gold particles (Hermann *et al.*, 1991). One of the earliest candidates for the enhanced 3D imaging provided by SEM was the chromosome, a robust structure due to the packaging ratio of 10 000:1 for the length of chromosome compared to the length of DNA contained within it. Whole chromosomes proved fairly impenetrable to early TEM studies producing a dense silhouette with a few individual looped fibers at the periphery. SEM of chromosomes that had been originally prepared for human karyotype analysis showed the whole structure in good surface detail. SEM imaging also revealed structural alterations along the length of the chromosome when LM staining protocols such as Geimsa were used to produce banding patterns to aid chromosome analysis. In the case of the Fragile X syndrome (which causes learning disabilities and cognitive impairment),

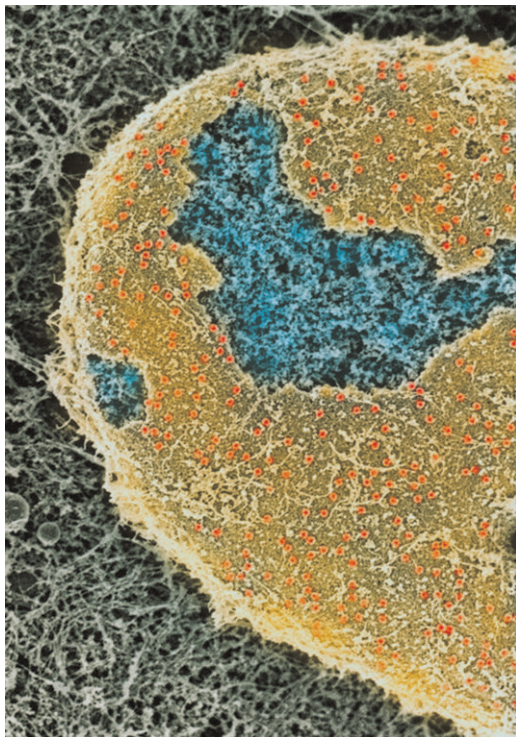


Figure 4 A simple method of visualizing internal cellular information by SEM. This tissue culture cell has been 'stripped' using adhesive tape. This procedure exposes both the nuclear surface (yellow) studded with nuclear pores (red) as well as internal nuclear chromatin (blue) on a base of cytoskeletal components after biochemical extraction of the remainder of the cytoplasm.

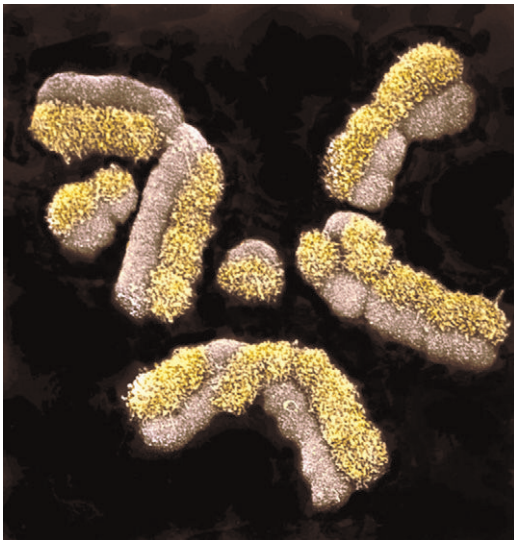


Figure 5 Chromosomes from a dividing muntjac cell in which the DNA was experimentally altered during replication, resulting in a structural difference between the chromatids which is clearly visible in SEM (courtesy C.J. Harrison).

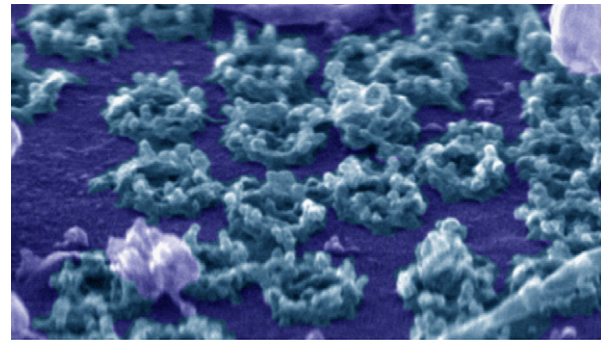
the potential break point was precisely located by SEM studies (Harrison *et al.*, 1983). As with whole cell images, SEM images of chromosomes have proved popular for the majority of textbooks (Figure 5).

Freeze Etching

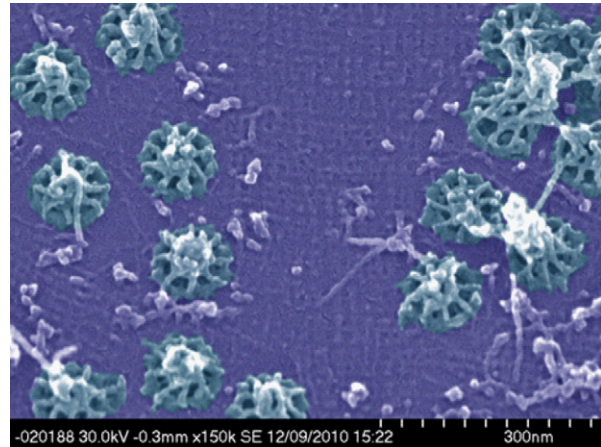
Another reason for the relatively late entrance of SEM into cell biology was the widespread use of surface replicas viewed in the TEM, which was (and still is) a popular technique for visualizing fine surface detail. These were made by evaporating a thin carbon film onto a surface of interest, and adding 'shadowing' contrast by evaporation of a heavy metal such as platinum at a low angle. The replica was floated off, and imaged in the TEM, providing images with depth perception not unlike SEM. Biological material was rapidly frozen, fractured with a cold razor blade, and then subjected to etching during which the surface water was removed by sublimation in a vacuum. At this point a replica was made, and the biological material thawed and removed – a technique known as freeze etching. As the fracture planes in frozen tissue would travel down membrane surfaces (which provide the easiest cleavage route), this technique, originally pioneered by Steere, and Moore and Mulethaler, and subsequently John Heuser, became widely used in cell biology (and still is). SEM was good for the more extreme topography of whole cell membranes (due to the very high depth of focus), but lacked equivalent resolution to the TEM, until the introduction of SEMs with FEGs and resolution of a few nanometers, which produce images with much the same resolution as TEM.

SEM and the Nuclear Envelope

Although early SEM images of the outer cell membrane were spectacular in their 3D detail, no new structure was discovered.



(a)



(b)

Figure 6 (a) High-resolution FEG-SEM image of the surface of an isolated nuclear membrane from the oocyte of *Xenopus laevis* (African clawed toad). Against the nuclear membrane itself (purple), each pore complex (gray) stands proud of the surface, bounded by short stubby projections. (b) Here the inside surface of the same material as (a) is imaged at a similar magnification (150 000 ×). The nuclear membrane (purple) shows an underlying orthogonal fibrous network, and each pore complex has a characteristic basket structure.

However, one area in which completely novel structure within the cell was shown by SEM was the nuclear envelope. This is a double membrane structure, regularly pierced by nuclear pore complexes (NPC), which are conduits of the continuously active exchange of material between nucleus and cytoplasm. For well over 30 years, NPCs appeared as little more than a break in the nuclear envelope with little associated structure when viewed in TEM thin sections, although Joe Gall had established an octagonal radial symmetry (Gall, 1967), extended into a 3D model by Hinshaw *et al.* (1992). Large areas of intact nuclear envelopes are easily dissected from the giant nuclei of amphibian oocytes, and when Hans Ris used FEG-SEM he found that the NPCs had a 'fishtrap' structure (subsequently the NPC basket) on their inside surface (Ris, 1997). Subsequent work by Goldberg and Allen, again using FEG-SEM went on to show further detail of the pore complex, and the presence of a fibrillar network attached to the baskets termed the nuclear envelope lattice (Goldberg and Allen, 1992; Figures 6(a) and 6(b)). Using a cell-free egg cytoplasmic extract system to incubate sperm heads, they showed stages in the formation of the NPC, and identified the roles of individual

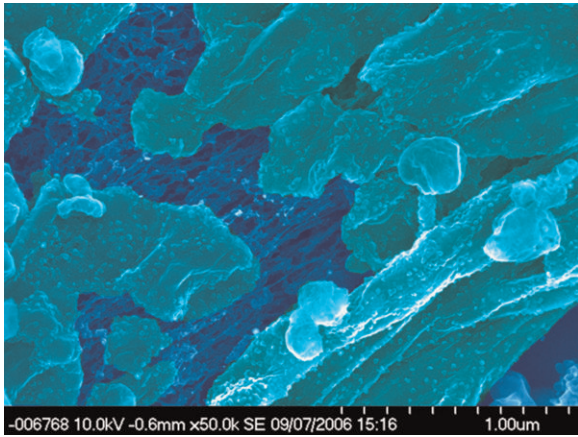


Figure 7 An isolated *Xenopus* sperm head is incubated *in vitro* in extracted cytoplasm from an oocyte and rapidly forms a new nuclear membrane (green) over the surface of the chromatin (blue) before going on to replicate its DNA, imaged here by FEG-SEM at 50 000 $\times$ magnification. This type of experimental protocol produces material which would be difficult to image by other means, but at the same time generates unique information in the stages of nuclear pore formation (Goldberg *et al.*, 1997).

proteins (Goldberg *et al.*, 1997; Figure 7). Further work was extended to *in situ* studies in tissue culture cells, including details of the attachment of the HIV virus and its entry into the nucleus via the pore (Arhel *et al.*, 2007).

SEM in Cell Biology

For all the involved preparation (particularly when compared with straightforward fast freezing), it might appear to some that SEM in cell biology has reached its peak, but it retains many advantages, not least of which is the sheer area of material that can be covered – particularly important when only one NPC amongst five thousand on the surface of a nucleus may have an HIV particle attached. Labeling is sophisticated, with different sized markers allowing identification of several proteins at once, and the stability of the preparations (unlike frozen material) allows multiple reexaminations. Experimental cell free systems of component formation do not easily lend themselves to cryo methods, and although the cryo-aficionados will always point out potential artifacts from fixation, dehydration, and coating, remarkable similarities have been shown for almost all samples between cell biological SEM and other methods. For medium resolution work, some have even

suggested that techniques such as serial block face imaging and FIB-SEM could make the TEM redundant, restricting it to high-end applications such as cryo-EM tomography and single particle reconstruction. The role of SEM in cell biology is one which should continue to produce novel and high-quality results for many years to come.

See also: Imaging the Cell: Electron Microscopy: Electron Tomography

References

- Apkarian, R.P., 1994. Analysis of high quality monatomic chromium films used in biological high resolution scanning electron microscopy. *Scanning Microscopy* 8, 289–299.
- Arhel, N., Souquere-Besse, S., Munier, S., *et al.*, 2007. HIV-1 DNA flap formation promotes uncoating of the pre-integration complex at the nuclear pore. *EMBO Journal* 26, 3025–3037.
- Donald, A.M., 2003. The use of environmental scanning electron microscopy for imaging wet and insulating materials. *Nature Materials* 2, 511–516.
- Gall, J.G., 1967. Octagonal nuclear pores. *Journal of Cell Biology* 32, 391–399.
- Goldberg, J.G., 1997. Octagonal nuclear pores. *Journal of Cell Science* 110, 409–420.
- Goldberg, M.W., Allen, T.D., 1992. High resolution scanning electron microscopy of the nuclear envelope: demonstration of a new, regular, fibrous lattice attached to the baskets of the nucleoplasmic face of the nuclear pores. *Journal of Cell Biology* 119, 1429–1440.
- Harrison, C.J., Jack, E.M., Allen, T.D., Harris, R., 1983. The fragile X: A scanning electron microscope study. *Journal of Medical Genetics* 20, 280–285.
- Herrmann, R., Schwarz, H., Müller, M., 1991. High precision immunoscanning electron microscopy using Fab fragments coupled to ultra-small colloidal gold. *Journal of Structural Biology* 107, 38–47.
- Heuser, J.E., 2011. The origins and evolution of freeze-etch electron microscopy. *Journal of Electron Microscopy (Tokyo)* 60 (Suppl. 1), S3–S29.
- Hinshaw, J., Carragher, B.O., Milligan, R.A., 1992. Architecture and design of the nuclear pore complex. *Cell* 69 (7), 1133–1144.
- Hughes, L., Hawes, C., Monteith, S., Vaughan, S., 2014. Serial block face scanning electron microscopy – The future of cell ultrastructure imaging. *Protoplasma* 251 (2), 395–401.
- Kizilyaprak, C., Daraspe, J., Humbel, B.M., 2014. Focused ion beam scanning electron microscopy in biology. *Journal of Microscopy* 254, 109–114.
- Ris, H., 1997. High-resolution field-emission scanning electron microscopy of nuclear pore complex. *Scanning* 19, 368–375.

Further Reading

- Allen, T.D. (Ed.), 2008. *Introduction to Electron Microscopy for Biologists*. Method in Cell Biology, vol. 88. Waltham, MA: Academic Press.
- Schatten, H. (Ed.), 2013. *Scanning Electron Microscopy for the Life Sciences*. Cambridge: Cambridge University Press.

Electron Tomography

T Ishikawa, Paul Scherrer Institute, Villigen, Switzerland

© 2016 Elsevier Inc. All rights reserved.

Glossary

Correlative microscopy Correlative microscopy is a hybrid method to apply two different types of microscopy to the same view area of a specimen. In electron tomography, combination with a fluorescent light microscopy is used to locate the target of interest, based on genetically tagged fluorescent proteins or antibodies.

Cryo-EM Electron microscopy of frozen-hydrated specimens. In this article, we describe transmission electron microscopy (cryo-TEM), while scanning electron microscopy (SEM) and scanning transmission electron microscopy (STEM) are also applied to frozen-hydrate specimens. Cryo-TEM is one of common approaches to analyze three-dimensional (3D) structure of biological macromolecules and organelles.

Electron tomography Electron tomography is a method to reconstruct 3D structure from electron micrographs (which are 2D images). In electron tomography, a number

of micrographs are recorded from the identical objects (macromolecules, organelles, cells, etc.) by tilting the stage. While single-particle analysis reconstruct 3D structure by merging images from different particles, assuming they share the same 3D structure, images used by electron tomography are from one identical object. Therefore electron tomography is a potential method to analyze structure with heterogeneity, but radiation damage due to multiple illuminations limits the resolution.

Subtomogram averaging Subtomogram averaging is a computational technique to improve signal-to-noise ratio of 3D reconstruction. 3D volumes which cover the target (normally macromolecules), which are called subtomograms, are computationally extracted, aligned, and averaged. The macromolecules in the subtomograms must share the identical conformation. When there is heterogeneity, subtomograms must be classified.

Introduction

X-ray crystallography, NMR, and single particle EM provide a detailed view of biological macromolecules, allowing us to understand biochemical phenomena *in vitro* at atomic resolution. Now we have started exploring biological phenomena *in vivo* in cells or tissues, mainly at molecular but, in some cases, also at atomic resolution. *In vivo* function usually involves more molecules than the corresponding biochemical reaction *in vitro*, and thus requires the analysis of complex molecular networks. Electron tomography is a suitable approach to analyze such systems, since it allows visualization of complex biological three-dimensional (3D) structures in a noninvasive way. Therefore electron tomography can cover *in vivo* structural analysis, dynamics, and particles with irregular shapes (for example, virus with envelope). The potential targets of electron tomography are organelles, cells, and tissues, if they are sectioned to be thin enough ($<1\ \mu\text{m}$) for transmission electron microscopy (TEM). Molecules with high flexibility can also be a good target of electron tomography for dynamic studies taking advantage of 3D reconstruction without interparticle averaging.

The principle of 3D reconstruction from a number of 2D micrographs is the central section theorem. In single particle analysis many randomly oriented particles are illuminated only once and view angles are calculated under the assumption that they share the identical 3D structure. On the other hand, what is specific to electron tomography is the recording of micrographs from a single object at different tilt angles, which means that the electron-sensitive biological specimen is illuminated at multiple times. As a result, the resolution of electron tomography is not as high as in single particle

analysis. However, in electron tomography, specimens with heterogeneity or flexibility are accessible to 3D structural analysis.

Compared with other recently emerging techniques at intermediate resolution, such as super-resolution microscopy and block-phase SEM, electron tomography will be characterized as an imaging technique that can be applied to more intact specimens, in a noninvasive way and at higher resolution, but with narrower view areas.

In this section, we review the technique of electron tomography as a method to reconstruct biological molecular and cellular structures.

Specimen Preparation Methods

Electron tomography of biological specimens was introduced as a way to visualize intracellular 3D morphology with chemically fixed and stained samples, as well as samples prepared by high-pressure freezing and freeze substitution, both at room temperature (Frank, 1995; Parsons *et al.*, 1990). Although the latter method is sometimes referred to as 'cryo-EM,' it must be emphasized that the specimen is in the end embedded in plastic and kept at room temperature, which is different from cryo-electron microscopy discussed here (frozen hydrated without staining, Figure 1(a)). One difficulty of these methods, which in the end embed samples in resin, is the 30–50% shrinkage of the plastic section induced by the beam (Luther, 1992). Once the specimen is prepared for TEM, electron micrographs are recorded at the same area from various view angles, by tilting the specimen (Figure 1(b)). As an example, micrographs of the same eukaryotic flagellum,

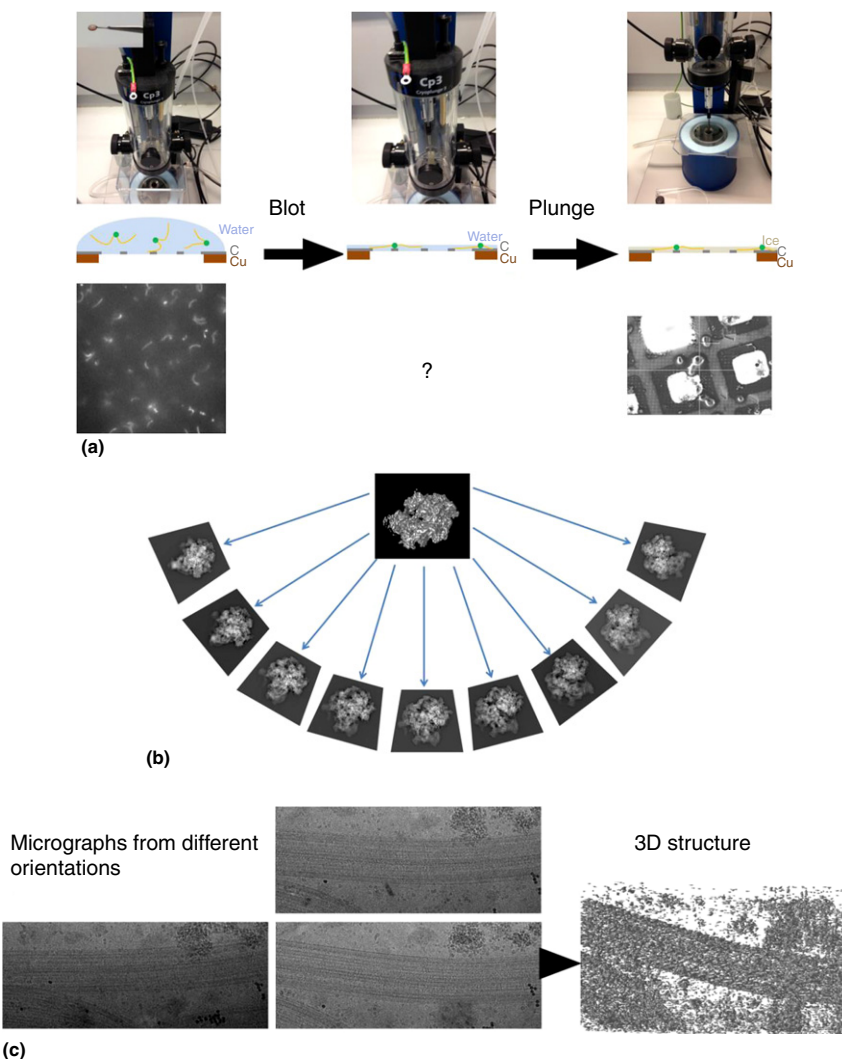


Figure 1 Flowchart of cryo-ET. (a) Plunge freezing. The specimen in buffer is mounted on an EM grid (left), blotted by filter paper (center) and plunged into cryogen (right). The grid is blotted in a chamber with temperature and humidity controlled (top). (b) Scheme of tomographic data acquisition. The same object is illuminated multiple times from different orientation. The specimen shown is a ribosome (EMD2914) filtered at 5 Å resolution. (c) 3D reconstruction. Micrographs of an object (cilium) from various orientations are merged to reconstruct a 3D structure. (a) and (c) Modified from Ishikawa, T., 2015. Cryo-electron tomography of motile cilia and flagella. *Cilia* 4, 3.

prepared by plunge freezing, viewed from different angles are shown in [Figure 1\(c\)](#). Typical tomograms from flagella prepared by three different methods are shown in [Figure 2](#). While tomograms from stained samples (top and middle of [Figure 2](#)) show high contrast, periodic structure of flagella is maintained better in cryo samples without chemical embedding and staining (bottom).

Cryo-electron tomography (cryo-ET) is relatively new technique. Although the possibility to reconstruct the 3D structure of unstained biological specimens by electron tomography was already predicted in 1968 ([Hart, 1968](#)), it was realized much later – in early 2000s. One of the major challenges was data acquisition. At each tilt angle, the motion of the stage must be detected and the electron beam must be shifted to track the object of interest (a process called ‘tracking’). To minimize radiation damage, tracking and

focusing must be done automatically at a certain distance from the area of interest, and then the beam must be shifted back precisely to the area of interest. This process was facilitated by the development of high sensitivity and large view area digital cameras, stable stages, and computer control of the lens/coil currents. Another difficulty was to enhance image contrast from unstained images under extremely low-dose conditions. Since the density of biological molecules and vitreous ice is not so different, image formation relies mainly on phase contrast ([Amos et al., 1982](#)). Moreover, since specimens used for tomography are relatively thick, inelastic scattering of electrons causes a high noise level (the mean free path of inelastic scattering is ~200 nm). Field emission guns (FEG) and energy filters enabled imaging of such thick unstained specimen. Therefore the first 3D map of the interior of an eukaryotic cells was published only in 2002 ([Medalia et al., 2002](#)). In



Figure 2 Specimen preparation methods. Micrographs of eukaryotic cilia prepared by three different methods are compared. (a) Chemical fixation, staining, and sectioning. Reproduced with permission from Pigino, G., Geimer, S., Lanzavecchia, S., *et al.*, 2009. Electron-tomographic analysis of Intraflagellar transport particle trains in situ. *Journal of Cell Biology* 187, 135–148. (b) High-pressure freezing and freeze substitution (Höög *et al.*, 2014). (c) Ice embedding (cryo-EM) by plunge freezing (our unpublished result). Despite low contrast, structure is preserved without severe distortion.

this section, we review mainly cryo-ET and related techniques, but also show some examples of tomography from samples at room temperature.

The maximum specimen thickness for TEM is 0.5 μm (200 kV) to 2 μm (300 kV). Thick specimens ($>2 \mu\text{m}$) must be sliced to allow the transmission of electrons. Cryo-ultramicrotome and cryo-FIB are methods to slice thick ice-embedded specimen. Cryo-section (CEMOVIS) (Al-Amoudi *et al.*, 2004; Figures 3(a) and 3(b)), although suffering from artifacts during sectioning by cryo-ultramicrotome such as compression, knife marks, crevasses, and chatter (Al-Amoudi *et al.*, 2005) (mainly at a cellular level, and not so much in molecular structures (Pierson *et al.*, 2011)), has the advantage of allowing serial sectioning, which could facilitate tomographic reconstruction of whole tissues. Cryo-FIB (focused ion beam) milling (Figures 3(c) and 3(d)) is an emerging technique that uses a gallium beam and is expected to cause fewer artifacts than a diamond knife, and, therefore, provides tomograms that are more straightforward to interpret.

Methods of Cryo-ET

Here the author describes a general protocol for cryo-ET. For other methods of specimen preparation, please refer to Heuser (1981) (freeze-fracture deep etch replica) and Vanhecke *et al.* (2008) (chemical fixation, high-pressure freezing, and freeze substitution). Process of cryo-ET data collection and analysis is summarized in the supplementary video.

Specimen Preparation

In cryo-ET, the specimen must be embedded in vitreous (amorphous) ice. There are two methods of vitrification: for thin specimen ($<1 \mu\text{m}$) such as macromolecules, macromolecular complexes *in vitro*, organelles, and small bacteria, the specimen can be vitrified quickly (within microseconds) by plunge freezing (Figure 1(a); Dubochet *et al.*, 1988). Aqueous solution of molecules or cells is mounted on EM grids covered by holey or continuous carbon support. After removing excess liquid, the grid is plunged into liquid ethane at liquid nitrogen temperature. By currently available freezing devices, it is difficult to control or examine samples during the freezing on EM grids, especially interaction between molecules and carbon support is hard to control (Figure 1(a) bottom).

Thick specimen, such as large eukaryotic cells and tissues must be frozen by high-pressure freezing, in which the specimen is frozen within milliseconds under ~ 2000 bar pressure and then cryo-sectioned using a cryo-ultramicrotome (Al-Amoudi *et al.*, 2004).

Microscopy

For electron tomography in general, the specimen stage must allow a tilt of 60° or more. Programs for automatic data acquisition are available (e.g., SerialEM (Mastronarde, 2005); UCSF Tomo (Zheng *et al.*, 2010); TOM Toolbox (Nickell *et al.*, 2005)). Thick ($>200 \text{ nm}$) specimens require an energy filter (in-column or post-column). An energy filter allows to perform spectroscopic analysis (EFTEM) for elemental mapping as well (Comolli *et al.*, 2006; Messaoudi *et al.*, 2013).

For cryo-ET, each dataset must be recorded at low cumulative dose, i.e., at an extremely low dose for each individual micrograph. A FEG is indispensable to acquire phase contrast. Since tomographic data acquisition is a time consuming process (1 h for acquisition of one dataset), the use of a specially designed polepiece (cryo-polepiece) to minimize ice contamination is desirable. Detector companies, such as TVIPS, FEI, and Gatan, also provide their own data acquisition software. It is also necessary for a low-dose system (or minimum dose system) to focus and track objects at an area distant from the object of interest by shifting the beam during the tilt series data collection. The direction of the beam shift must be parallel to the tilt axis to keep the same focus in both areas. The amount of the defocus (2–15 μm) depends on the specimen thickness and the desired resolution. To pursue high resolution, a few studies attempted contrast transfer function (CTF) correction despite the difficulty to evaluate defocus from low-dose images from highly tilted objects (Philippson *et al.*, 2007).

Since it is not possible to tilt the stage 90° , there is always some information missing, a circumstance called ‘missing wedge’ (single axis) or ‘missing pyramid’ (dual axis) (Lucić *et al.*, 2005). Dual-axis cryo-tomography became possible with modern high-end microscopes, which have a system to rotate grids horizontally inside the column (Myasnikov *et al.*, 2013). The higher the angle at which data is collected, the less information is missed. The optimum increment tilt angle depends also on the specimen thickness and the desired resolution (Crowther *et al.*, 1970). Although the movement of

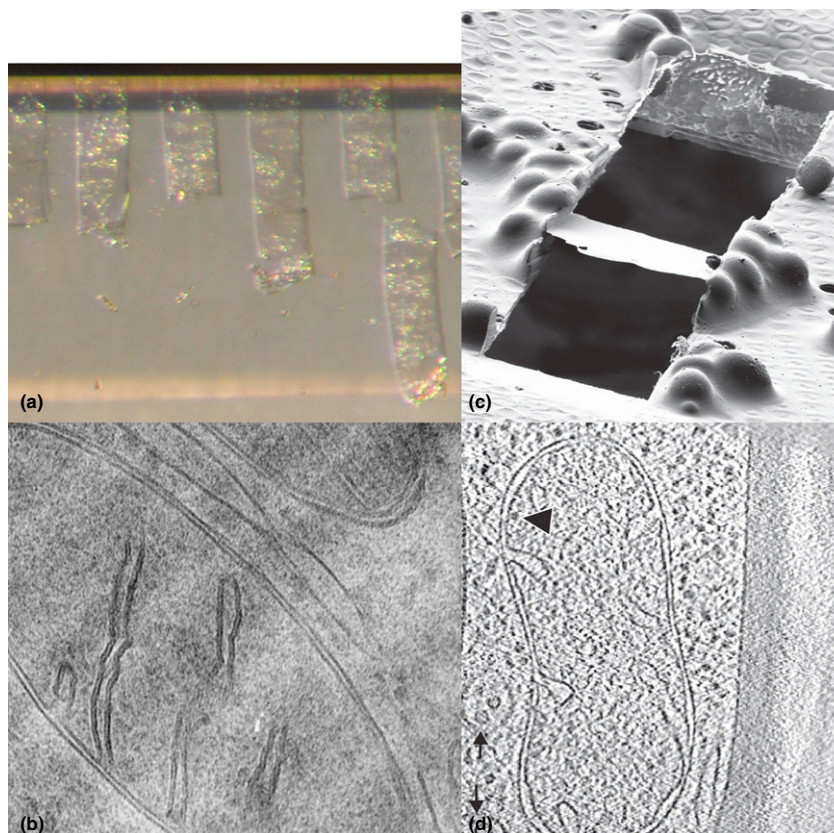


Figure 3 Two methods to slice frozen thick specimen. (a) and (b) Cryo-section (Bouchet-Marquis and Hoenger, 2011). (a) Sectioned frozen specimen (ribbon). (b) Electron micrograph. Artifact by compression is seen. (c) Cryo-FIB observed by SEM. (d) Electron micrograph of cryo-FIB milled cells (Villa *et al.*, 2013).

the stage during tilting is roughly tracked by the software, further alignment between micrographs within a tilt series is necessary. Normally, gold clusters (5–15 nm) are used as fiducial markers. Marker-free (fiducial-less) alignments have been successfully developed for stained sections (Winkler and Taylor, 2013). For cryo-ET, marker-free alignment is not common due to poor S/N, but there are attempts to develop algorithms of marker-free alignment for cryo-ET (Castaño-Díez *et al.*, 2007, 2010).

Tomogram Reconstruction and 3D Image Analysis

The aligned micrographs are merged into a 3D structure using R-weighted backprojection (Radermacher, 1988), SIRT (Zürner *et al.*, 2012), or ART (Messaoudi *et al.*, 2013). The reconstructed 3D tomogram suffers from poor signal-to-noise ratio (S/N), which is a direct evidence of the limitations of this technique. However, further structural information can be extracted by 3D image processing and the computational analysis of the tomograms. Denoising is a technique used to improve S/N and highlight structural features (Narasimha *et al.*, 2008; Figure 4(a)). Also, segmentation algorithms allow to distinguish two or more structures from a noisy tomogram (Frangakis and Hegerl, 2002; Figure 4(b)). For instance, synaptic vesicles and their connector and tethers have been

visualized by this method (Fernández-Busnadiego *et al.*, 2010; Figure 4(c)). Although these techniques can sometimes resolve features convincingly from noisy tomograms, they must be interpreted carefully.

Another approach to extract structural information from noisy cryo-electron tomograms is cross-correlation-based averaging. Volumes which involve target particles (called subtomograms), which supposedly share an identical 3D structure, are extracted, aligned in three-dimensions and averaged to improve S/N. The averaged structure from subtomograms extracted from even thick tomograms can reach ~25 Å resolution (Bui *et al.*, 2009). In this method, the main challenge is the proper extraction of subtomograms under noisy and heterogeneous environment, so they include the particles of interest. It is important to minimize the risks to average particles with different structures, or to extract at a wrong position or angle. This method is particularly suitable to analyze periodic structures inside cells or organelles, where subtomograms can be extracted based on periodicity, as shown for instance in eukaryotic axonemes (Bui *et al.*, 2012), and aligned with a roughly estimated Euler angle (Bui and Ishikawa, 2013). The technique of cross-correlation-based exhaustive search can be used to extract randomly distributed particles. To analyze heterogeneous environments (for example, cytoplasm with proteasomes and chaperones inside (Bohm *et al.*, 2000)), it is necessary to use a

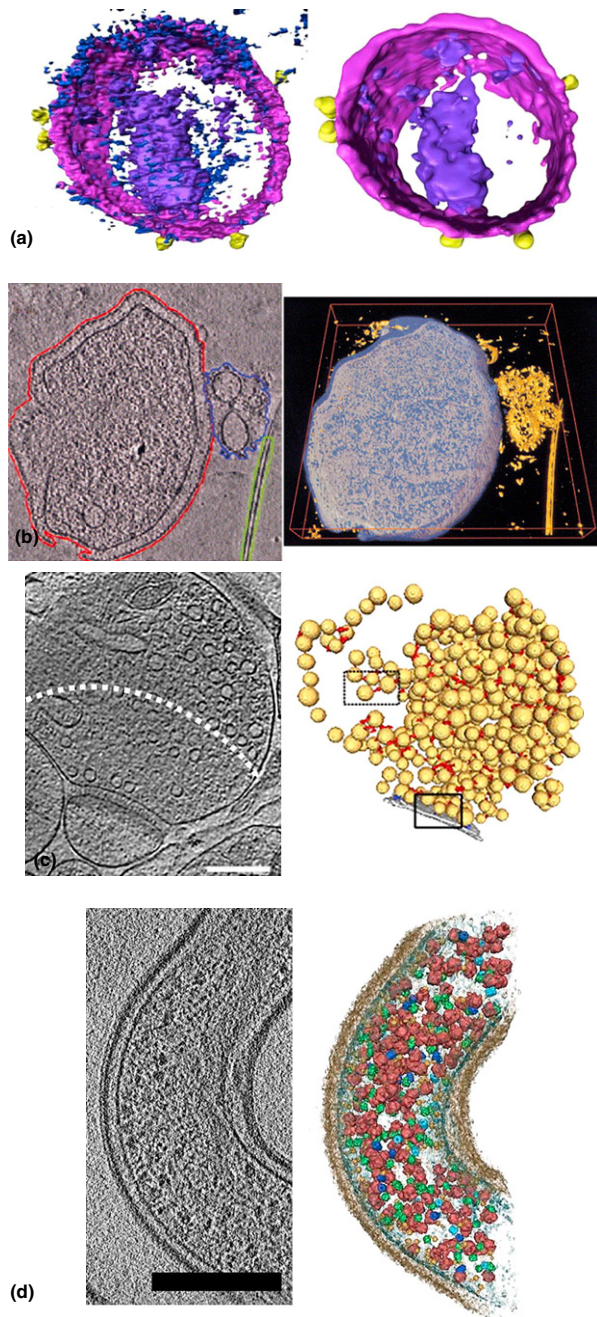


Figure 4 Various post-processing. (a) Denoising. Reproduced with permission from Narasimha, R., Aganj, I., Bennett, A.E., *et al.*, 2008. Evaluation of denoising algorithms for biological electron tomography. *Journal of Structural Biology* 164, 7–17. Left: before denoising. Right: after denoising. (b) Segmentation of cellular components from a tomogram of *Pyrodicticum abyssi*. Reproduced with permission from Frangakis, A.S., Hegerl, R., 2002. Segmentation of two- and three-dimensional data from electron microscopy using eigenvector analysis. *Journal of Structural Biology* 138, 105–113. (c) Presynaptic structure (left) (Fernández-Busnadiego *et al.*, 2010). Segmentation clarified presynaptic vesicles and their connectors and tethers. (d) Distribution of ribosomes, chaperones, and proteasomes revealed by template matching in a tomogram of *Leptospira interrogans*. Reproduced with permission from Beck, M., Malmström, J.A., Lange, V., *et al.*, 2009. Visual proteomics of the human pathogen *Leptospira interrogans*. *Nature Methods* 6, 817–823.

multi-template-matching-based search. For averaging, particles can be classified into subclasses by cross-correlation between subtomograms and multiple templates (subtomogram classification). Classification is not straightforward because of poor S/N and missing wedges/pyramids. For this process, algorithms as cross-correlation-based template matching and unsupervised multivariate statistical analysis are used (Beck *et al.*, 2009; Frangakis and Hegerl, 2002; Figures 4(d) and 4(e)). Unsupervised subtomogram classification often suffers from the missing wedge: subtomograms are classified based on the orientations of the missing wedge, instead of structure. Supervised classification must be carried out carefully to avoid bias by templates.

Techniques to Combine with Cryo-ET

While electron tomography has been successful to visualize 3D structures of highly complex biological macromolecules and organelles, limited resolution is still a significant drawback. In spite of a few reports of tomographic analysis at resolution below 1 nm (Schur *et al.*, 2013), in general it is challenging to reach 2 nm resolution. Especially for thick specimen, resolution is usually limited to ~ 5 nm. To extend the applications to molecular level, it is essential to combine this technique with other methods that allow us to identify the molecules corresponding to the densities observed in the tomogram.

The direct way to identify molecules is either to increase density in a region of the tomogram by specific labeling or genetic tagging, or to decrease the density by a deletion mutant. As in standard immunoelectron microscopy, antibodies can be used to locate target molecules in the tomogram. When they are engineered as a small derivatives, this helps to locate component proteins in a complex (Meyerson *et al.*, 2013). However, the access of the antibody to organelles is often sterically limited. Genetic tagging to enhance density has been carried out, for instance, with ferritin (Wang *et al.*, 2011), biotin-streptavidin (Oda and Kikkawa, 2013; Oda *et al.*, 2014), and metallothionein (Diestra *et al.*, 2009). The use of tags that induce metal cluster formation, when applied to the cell, could cause artifacts and thus must be treated carefully. Deletion mutants have been used, for instance, to locate isoforms of dynein motor proteins in cilia (Bui *et al.*, 2008). However, in highly complex biological systems, which cryo-ET often addresses, genetic deletion of one component sometimes causes the loss of other components, and therefore does not really help us locate these proteins. Expression of a tagged protein in the deletion mutant generates a structure with the density of the target protein amplified, which enables its localization (Oda *et al.*, 2014).

Cryo-correlative light/electron microscopy (cryo-CLEM) is a technique that allows localization at low resolution, by recording fluorescent micrographs of tagged or fluorescently labeled molecules on the frozen grid. The method of CLEM is widely applied to image specimens at room temperature. Similarly to the regular TEM, immunolabeling can be applied to map molecules (Mari *et al.*, 2014). The instrument to embed the cellular specimen, and observe it simultaneously by light microscopy, is commercially available. Although cryo-stages to mount frozen grids for light microscopy are available

and a few successful works have been reported (Zhang 2013; Koning *et al.*, 2014), cryo-CLEM still has some barriers to overcome. One problem is the long working distance: to protect the objective lens, cryo-specimens must be kept at a certain distance from the lens, which limits the numerical aperture and, thus, hinders the detection of weak fluorescent signal. Another problem is the resolution of fluorescent microscopy. To address it, a recent study proposed a method to reach a range of nanometer resolution using multiple dyes (Schellenberger *et al.*, 2014). An additional approach to identify molecules in tomograms is to fit external information, either high-resolution structures from X-ray crystallography, single particle EM and NMR, or biochemical information of molecular connectivity, such as crosslinking, pull-down, immune-precipitation, and mass spectroscopy (Bui *et al.*, 2013).

Examples

The first cryo-ET of vitrified eukaryotic cells showed networks of actin filaments, as well as the ribosome and other features in the cytoplasm (Medalia *et al.*, 2002; Figure 5(a)). The arrangement of actin networks has been extensively studied by cryo-ET – remodeling of actin networks induced by infection of *Listeria monocytogenes* (Jasnin *et al.*, 2013). Interestingly, in studies of lamellipodia, two groups claim different results: actin networks with (Yang and Svitkina, 2011) and without (Urban *et al.*, 2010) branches (Figures 5(b) and 5(c)).

Until recently it was believed that bacteria had no cytoskeleton. However, cryo-ET has allowed to clearly visualize cytoskeletons inside the bacteria *Spiroplasma melliferum* (Kürner *et al.*, 2005; Figure 6(a)). This filament was later proved to be MreB, an actin homologue (Trachtenberg *et al.*, 2008). Since then, various bacterial cytoskeletons have been found and characterized by cryo-ET, for example, BtubA/B proteins form microtubule-like structure in *Prostheco bacter* (Pilhofer *et al.*, 2011; Figure 6(b)).

In the long history of structural analysis of spherical viruses by cryo-EM single particle analysis, this method has been virtually limited to symmetric capsid. Other component of the virus, such as tegument, membrane, and nucleic acids were not targets of 3D reconstruction, due to structural heterogeneity and flexibility. However, recent cryo-ET studies offered the possibility of 3D reconstruction of the whole virus (Grünwald *et al.*, 2003; Figure 6(c)).

The strong-binding interface of actin and myosin has been studied by single particle analysis of filaments reconstituted *in vitro* (Holmes *et al.*, 2003). However, actin/myosin interactions during contraction result in highly heterogeneous structures, which require tomographic analysis. Cross-bridges between actin and myosin filaments were studied using freeze-substituted and sectioned contracting insect flight muscle (Wu *et al.*, 2009). Subtomogram averaging allowed to visualize transition from weak-binding to strong binding of myosin heads to actin filaments (Wu *et al.*, 2010; Figure 7(a)).

Cryo-ET has provided insights into microtubule networks. Here we overview one example of a microtubule-based system:

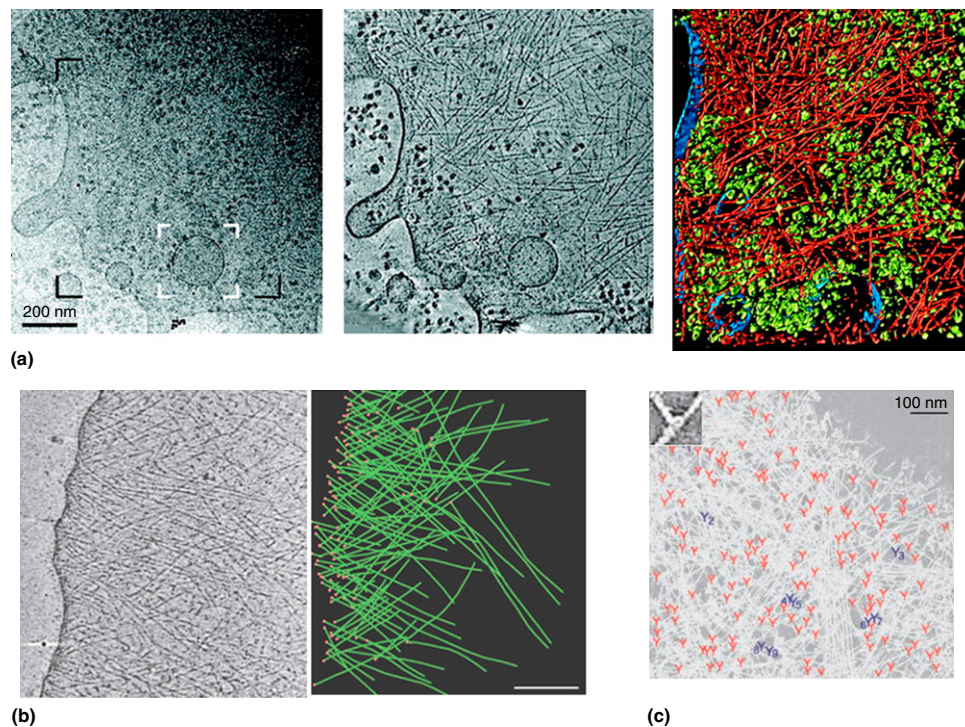


Figure 5 Cryo-electron tomography of eukaryotic cells. (a) Whole cell tomography of *Dictyostelium discoideum*. Reproduced with permission from Medalia, O., Weber, I., Frangakis, A.S., *et al.*, 2002. Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. Science 298, 1209–1213. (b) and (c) Actin networks inside lamellipodia: (b) without branches (Urban *et al.*, 2010) and (c) with branches (Yang and Svitkina, 2011). Inset of the right panel: subtomogram average to highlight branched actin filaments.

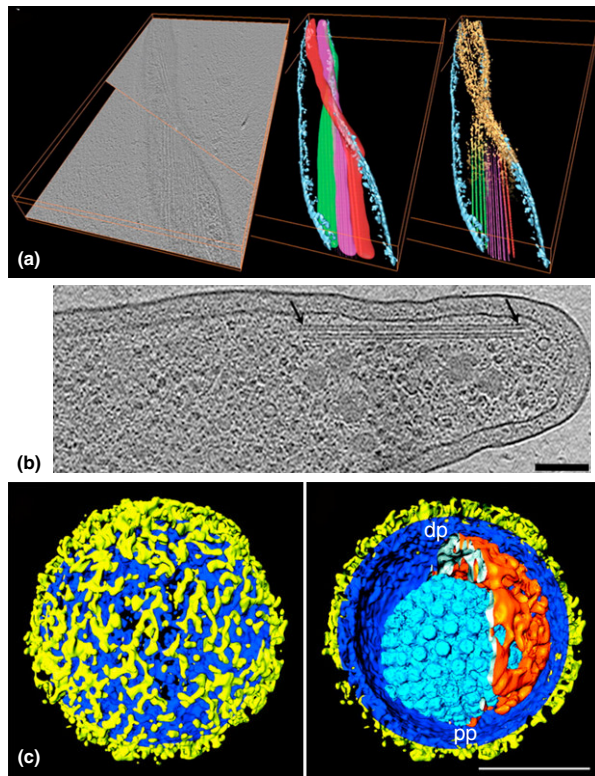


Figure 6 Cryo-electron tomography of bacteria and virus. (a) Whole cell tomography of bacteria *Spiroplasma melliferum* (Kürner *et al.*, 2005). Various structures of bacterial cytoskeleton are shown in colors. (b) Bacteria *Prosthecobacter* (Pilhofer *et al.*, 2011). Microtubule-like structure is shown by arrows. (c) Herpes simplex virus (Grünewald *et al.*, 2003). High-resolution structure of the capsid (cyan) was fitted to the tomography of the whole virus.

cilia. Ciliary (or eukaryotic flagellar) '9 + 2' axoneme structure has been known for a long time. In motile cilia, dynein, microtubule-based motor proteins are the driving force, together with many other regulatory protein complexes. While, various EM techniques (negative stain, freeze-fracture deep-etch (Goodenough and Heuser, 1985), and section (Mastrorade *et al.*, 1992)) have been applied to reveal the molecular architecture of cilia, the first 3D reconstruction of the entire cilia structure was obtained by cryo-ET and subtomogram averaging. Dynein and regulatory proteins follow either 96 nm or 24 nm periodicity along microtubules, which allowed to perform subtomogram averaging (Ishikawa *et al.*, 2007; Nicastro *et al.*, 2006), and to distinguish the dynein structure in pre- and post-power stroke states (Lin *et al.*, 2014; Movassagh *et al.*, 2010). Also, comparison of mutant structures allowed to identify different dynein isoforms (Bui *et al.*, 2008, 2012; Figure 7(b)). Finally, this technique has permitted to reconstruct the structure of regulatory proteins (Heuser *et al.*, 2009; Pigino *et al.*, 2011).

Another successful example of the combination of cryo-ET, subtomogram averaging and hybrid approaches is the structural analysis of the nuclear pore complex (NPC). cryo-ET allowed us to visualize the 3D structure of the NPC embedded

in the nuclear envelope (Figure 7(c), left). The NPC has apparent eightfold symmetry, but, actually, this symmetry is somewhat perturbed. Recent progress in image quality and image analysis techniques have facilitated a symmetry-independent alignment and averaging (Maimon *et al.*, 2012). The most recent tomographic analysis reached 32 Å resolution (Figure 7(c), center) and, combined with single particle analysis, chemical crosslink and mass spectroscopy, allowed the assignment of component molecules (Bui *et al.*, 2013; Figure 6(c), right). Furthermore gold-labeling demonstrated distribution of cargo-transporters (Figure 7(c) right bottom; Beck *et al.*, 2007).

Future Outlook

There are numerous on-going efforts to further develop electron tomography and expand it to explore new directions. One of the main issues is the improvement of resolution, especially for thick specimens. In the related technique of single particle analysis, another major method of electron microscopy, recently developed direct electron detectors, which replace negative and conventional charge-coupled device (CCD) or complementary metal-oxide-semiconductor (CMOS) cameras with scintillators, have represented a breakthrough that has enabled *in vitro* structural analysis at near-atomic resolution (Li *et al.*, 2013). Nevertheless, it is not yet clear how much contribution this type of detectors can bring to cryo-ET. On the one hand, improved S/N from the direct detector at high-resolution range will enable a more precise alignment, maybe at lower dose. However, drift correction, another advantage of direct detectors, may not be as effective as single particle analysis, since the exposure time of tomography is, due to extraordinary low dose, relatively short.

To compensate the limitation in resolution, there are attempts to further develop tagging and labeling techniques for molecular identification. These developments will be tightly linked to new methods of genetic engineering, such as knockout, knockdown by various RNAi, and protein expression techniques.

To handle eukaryotic cells and tissues noninvasively, i.e., without sectioning or milling, we need a breakthrough in the field. For thick samples (> 2 µm), TEM with higher accelerating voltage will be useful, once FEG and detectors for high voltage are developed. X-ray tomography is capable of non-invasive 3D visualization of whole cells and tissues. Soft X-ray tomography using zone-plate imaging technique reaches ~10 nm resolution of up to 5 µm thick specimen (Hagen *et al.*, 2012; McDermott *et al.*, 2012; Weiner *et al.*, 2013). Combination of X-ray tomography for the entire cells/tissues and electron tomography for sections could enable high-resolution imaging of the entire cells/tissues.

With the current imaging strategy using phase contrast, we need a large defocus to improve contrast of, especially, ice-embedded thick specimens, which degrade resolution. To overcome this problem and collect images at near-focus condition, the use of phase plates has allowed us to obtain good contrast without a large defocus. For instance, cryo-ET using Zernike phase plate provided amazingly high contrast (Fukuda

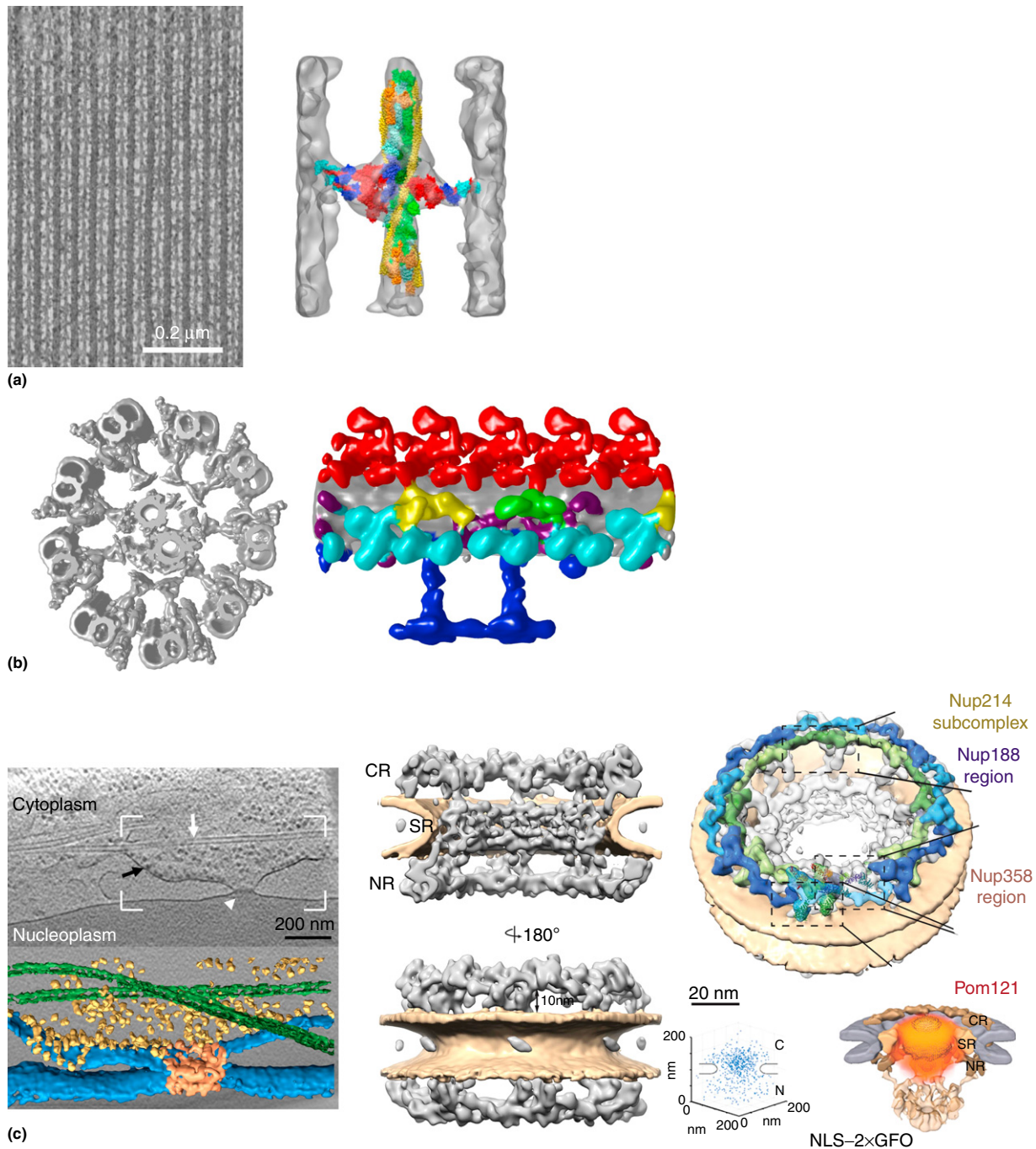


Figure 7 Various intracellular structures revealed by electron tomography. (a) Electron tomography of rapidly frozen, freeze-substituted, and sectioned insect flight muscle. Left: section of the tomogram (Wu *et al.*, 2009). Right: one 3D structure of cross-bridge after subtomogram classification and averaging. Atomic models of myosin, actin, and regulatory proteins are fitted (Wu *et al.*, 2010). (b) Axonemal structure from cilia/flagella by cryo-electron tomography (cryo-ET). Left: view of the whole '9 + 2' structure. Right: one microtubule doublet extracted from the '9 + 2' structure. Gray: microtubule. Red and cyan: dynein motor proteins (Bui *et al.*, 2008, 2012). Blue: radial spoke (Pigino *et al.*, 2011). Green: dynein regulator complex (Heuser *et al.*, 2009). (c) Nuclear pore complex by cryo-ET. Left: NPC embedded in nuclear envelope (Maimon *et al.*, 2012). Center: 32 Å resolution structure of human NPC (Bui *et al.*, 2013). Right top: molecular identification by fitting single particle analysis and mass spectroscopy data (Bui *et al.*, 2013). Right bottom: distribution of transporter revealed by gold-labeling and tomography (Beck *et al.*, 2007).

and Nagayama, 2012; Danev *et al.*, 2014). Another possibility is STEM tomography, with which all the data, including data from highly tilted specimen, could be collected without any defocus.

Future developments in electron tomography will fill the gap between X-ray crystallography or NMR and light microscopy, which will allow us to visualize detailed 3D structures of biological macromolecules *in vivo* at nanometer resolution.

Supplementary Video

Supplementary Video related to this article can be found online at doi:10.1016/B978-0-12-394447-4.20006-0

See also: Imaging the Cell: Electron Microscopy: Scanning Electron Microscopy in Cell Biology

References

- Amos, L.A., Henderson, R., Unwin, P.N., 1982. Three-dimensional structure determination by electron microscopy of two-dimensional crystals. *Progress in Biophysics and Molecular Biology* 39, 183–231.
- Al-Amoudi, A., Chang, J.-J., Leforestier, A., *et al.*, 2004. Cryo-electron microscopy of vitreous sections. *EMBO Journal* 23, 3583–3588.
- Al-Amoudi, A., Studer, D., Dubochet, J., 2005. Cutting artefacts and cutting process in vitreous sections for cryo-electron microscopy. *Journal of Structural Biology* 150, 109–121.
- Beck, M., Lucić, V., Förster, F., Baumeister, W., Medalia, O., 2007. Snapshots of nuclear pore complexes in action captured by cryo-electron tomography. *Nature* 449, 611–615.
- Beck, M., Malmström, J.A., Lange, V., *et al.*, 2009. Visual proteomics of the human pathogen *Leptospira interrogans*. *Nature Methods* 6, 817–823.
- Bohm, J., Frangakis, A.S., Hegerl, R., *et al.*, 2000. Toward detecting and identifying macromolecules in a cellular context: Template matching applied to electron tomograms. *Proceedings of the National Academy of Sciences of the United States of America* 97, 14245–14250.
- Bouchet-Marquis, C., Hoenger, A., 2011. Cryo-electron tomography on vitrified sections: A critical analysis of benefits and limitations for structural cell biology. *Micron (Oxford, England)* 1993 (42), 152–162.
- Bui, K.H., Ishikawa, T., 2013. 3D structural analysis of flagella/cilia by cryo-electron tomography. *Methods in Enzymology* 524, 305–323.
- Bui, K.H., Sakakibara, H., Movassagh, T., Oiwa, K., Ishikawa, T., 2008. Molecular architecture of inner dynein arms in situ in *Chlamydomonas reinhardtii* flagella. *Journal of Cell Biology* 183, 923–932.
- Bui, K.H., Sakakibara, H., Movassagh, T., Oiwa, K., Ishikawa, T., 2009. Asymmetry of inner dynein arms and inter-doublet links in *Chlamydomonas* flagella. *Journal of Cell Biology* 186, 437–446.
- Bui, K.H., von Appen, A., DiGiulio, A.L., *et al.*, 2013. Integrated structural analysis of the human nuclear pore complex scaffold. *Cell* 155, 1233–1243.
- Bui, K.H., Yagi, T., Yamamoto, R., Kamiya, R., Ishikawa, T., 2012. Polarity and asymmetry in the arrangement of dynein and related structures in the *Chlamydomonas axoneme*. *Journal of Cell Biology* 198, 913–925.
- Castaño-Díez, D., Al-Amoudi, A., Glynn, A.-M., Seybert, A., Frangakis, A.S., 2007. Fiducial-less alignment of cryo-sections. *Journal of Structural Biology* 159, 413–423.
- Castaño-Díez, D., Scheffer, M., Al-Amoudi, A., Frangakis, A.S., 2010. Alignator: A GPU powered software package for robust fiducial-less alignment of cryo tilt-series. *Journal of Structural Biology* 170, 117–126.
- Comolli, L.R., Kundmann, M., Downing, K.H., 2006. Characterization of intact subcellular bodies in whole bacteria by cryo-electron tomography and spectroscopic imaging. *Journal of Microscopy* 223, 40–52.
- Crowther, R., Derosier, D., Klug, A., 1970. Reconstruction of 3 dimensional structure from projections and its application to electron microscopy. *Proceedings of the Royal Society of London. Series B, Biological sciences* 317, 319.
- Danev, R., Buijsse, B., Khoshouei, M., Plitzko, J.M., Baumeister, W., 2014. Volta potential phase plate for in-focus phase contrast transmission electron microscopy. *Proceedings of the National Academy of Sciences of the United States of America* 111, 15635–15640.
- Diestra, E., Fontana, J., Guichard, P., Marco, S., Risco, C., 2009. Visualization of proteins in intact cells with a clonable tag for electron microscopy. *Journal of Structural Biology* 165, 157–168.
- Dubochet, J., Adrian, M., Chang, J.J., *et al.*, 1988. Cryo-electron microscopy of vitrified specimens. *Quarterly Reviews of Biophysics* 21, 129–228.
- Fernández-Busnadiego, R., Zuber, B., Maurer, U.E., *et al.*, 2010. Quantitative analysis of the native presynaptic cytomatrix by cryoelectron tomography. *Journal of Cell Biology* 188, 145–156.
- Frangakis, A.S., Hegerl, R., 2002. Segmentation of two- and three-dimensional data from electron microscopy using eigenvector analysis. *Journal of Structural Biology* 138, 105–113.
- Frank, J., 1995. Approaches to large-scale structures. *Current Opinion in Structural Biology* 5, 194–201.
- Fukuda, Y., Nagayama, K., 2012. Zernike phase contrast cryo-electron tomography of whole mounted frozen cells. *Journal of Structural Biology* 177, 484–489.
- Goodenough, U.W., Heuser, J.E., 1985. Outer and inner dynein arms of cilia and flagella. *Cell* 41, 341–342.
- Grünwald, K., Desai, P., Winkler, D.C., *et al.*, 2003. Three-dimensional structure of herpes simplex virus from cryo-electron tomography. *Science* 302, 1396–1398.
- Hagen, C., Guttman, P., Klupp, B., *et al.*, 2012. Correlative VIS-fluorescence and soft X-ray cryo-microscopy/tomography of adherent cells. *Journal of Structural Biology* 177, 193–201.
- Hart, R.G., 1968. Electron microscopy of unstained biological material: The polytropic montage. *Science* 159, 1464–1467.
- Heuser, J., 1981. Preparing biological samples for stereomicroscopy by the quick-freeze, deep-etch, rotary-replication technique. *Methods in Cell Biology* 22, 97–122.
- Heuser, T., Raytchev, M., Krell, J., Porter, M.E., Nicastro, D., 2009. The dynein regulatory complex is the nexin link and a major regulatory node in cilia and flagella. *Journal of Cell Biology* 187, 921–933.
- Holmes, K.C., Angert, I., Kull, F.J., Jahn, W., Schröder, R.R., 2003. Electron cryo-microscopy shows how strong binding of myosin to actin releases nucleotide. *Nature* 425, 423–427.
- Höög, J.L., Lacomble, S., O'Toole, E.T., *et al.*, 2014. Modes of flagellar assembly in *Chlamydomonas reinhardtii* and *Trypanosoma brucei*. *eLife* 3, e01479.
- Ishikawa, T., Sakakibara, H., Oiwa, K., 2007. The architecture of outer dynein arms in situ. *Journal of Molecular Biology* 368, 1249–1258.
- Jasnin, M., Asano, S., Gouin, E., *et al.*, 2013. Three-dimensional architecture of actin filaments in *Listeria monocytogenes* comet tails. *Proceedings of the National Academy of Sciences of the United States of America* 110, 20521–20526.
- Koning, R.I., Celler, K., Willemse, J., *et al.*, 2014. Correlative cryo-fluorescence light microscopy and cryo-electron tomography of *Streptomyces*. *Methods in Cell Biology* 124, 217–239.
- Kürner, J., Frangakis, A.S., Baumeister, W., 2005. Cryo-electron tomography reveals the cytoskeletal structure of *Spiroplasma melliferum*. *Science* 307, 436–438.
- Li, X., Mooney, P., Zheng, S., *et al.*, 2013. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nature Methods* 10, 584–590.
- Lin, J., Okada, K., Raytchev, M., Smith, M.C., Nicastro, D., 2014. Structural mechanism of the dynein power stroke. *Nature Cell Biology* 16, 479–485.
- Lucić, V., Förster, F., Baumeister, W., 2005. Structural studies by electron tomography: From cells to molecules. *Annual Review of Biochemistry* 74, 833–865.
- Luther, P.K., 1992. Sample shrinkage and radiation damage. In: Frank, J. (Ed.), *Electron Tomography*. New York, NY: Plenum, pp. 39–61.
- Maimon, T., Elad, N., Dahan, I., Medalia, O., 2012. The human nuclear pore complex as revealed by cryo-electron tomography. *Structure (London, England)* 20, 998–1006.
- Mari, M., Geerts, W.J.C., Reggiori, F., 2014. Immuno- and correlative light microscopy-electron tomography methods for 3D protein localization in yeast. *Traffic (Copenhagen, Denmark)* 15, 1164–1178.
- Mastrorade, D.N., 2005. Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology* 152, 36–51.
- Mastrorade, D.N., O'Toole, E.T., McDonald, K.L., McIntosh, J.R., Porter, M.E., 1992. Arrangement of inner dynein arms in wild-type and mutant flagella of *Chlamydomonas*. *Journal of Cell Biology* 118, 1145–1162.
- McDermott, G., Le Gros, M.A., Larabell, C.A., 2012. Visualizing cell architecture and molecular location using soft x-ray tomography and correlated cryo-light microscopy. *Annual Review of Physical Chemistry* 63, 225–239.
- Medalia, O., Weber, I., Frangakis, A.S., *et al.*, 2002. Macromolecular architecture in eukaryotic cells visualized by cryoelectron tomography. *Science* 298, 1209–1213.
- Messaoudi, C., Aschman, N., Cunha, M., *et al.*, 2013. Three-dimensional chemical mapping by EFTEM-TomoJ including improvement of SNR by PCA and ART reconstruction of volume by noise suppression. *Microscopy and Microanalysis: The Official Journal of Microscopy Society of America, Microbeam Analysis Society, Microscopical Society of Canada* 19, 1669–1677.
- Meyerson, J.R., Tran, E.E.H., Kuybeda, O., *et al.*, 2013. Molecular structures of trimeric HIV-1 Env in complex with small antibody derivatives. *Proceedings of the National Academy of Sciences of the United States of America* 110, 513–518.

- Movassagh, T., Bui, K.H., Sakakibara, H., Oiwa, K., Ishikawa, T., 2010. Nucleotide-induced global conformational changes of flagellar dynein arms revealed by in situ analysis. *Nature Structural & Molecular Biology* 17, 761–767.
- Myasnikov, A.G., Afonina, Z.A., Klaholz, B.P., 2013. Single particle and molecular assembly analysis of polyribosomes by single- and double-tilt cryo electron tomography. *Ultramicroscopy* 126, 33–39.
- Narasimha, R., Aganj, I., Bennett, A.E., *et al.*, 2008. Evaluation of denoising algorithms for biological electron tomography. *Journal of Structural Biology* 164, 7–17.
- Nicastro, D., Schwartz, C., Pierson, J., *et al.*, 2006. The molecular architecture of axonemes revealed by cryoelectron tomography. *Science* 313, 944–948.
- Nickell, S., Förster, F., Linaroudis, A., *et al.*, 2005. TOM software toolbox: Acquisition and analysis for electron tomography. *Journal of Structural Biology* 149, 227–234.
- Oda, T., Kikkawa, M., 2013. Novel structural labeling method using cryo-electron tomography and biotin-streptavidin system. *Journal of Structural Biology* 183, 305–311.
- Oda, T., Yanagisawa, H., Yagi, T., Kikkawa, M., 2014. Mechanosignaling between central apparatus and radial spokes controls axonemal dynein activity. *Journal of Cell Biology* 204, 807–819.
- Parsons, D.F., Marko, M., Leith, A., 1990. The relative merits of direct morphometry of reconstructions of whole cells, and statistical morphometry by stereology of random sections of cells. *Cell Biophysics* 17, 227–242.
- Philippson, A., Engel, H.-A., Engel, A., 2007. The contrast-imaging function for tilted specimens. *Ultramicroscopy* 107, 202–212.
- Pierson, J., Ziese, U., Sani, M., Peters, P.J., 2011. Exploring vitreous cryo-section-induced compression at the macromolecular level using electron cryo-tomography; 80S yeast ribosomes appear unaffected. *Journal of Structural Biology* 173, 345–349.
- Pigino, G., Bui, K.H., Maheshwari, A., *et al.*, 2011. Cryoelectron tomography of radial spokes in cilia and flagella. *Journal of Cell Biology* 195, 673–687.
- Pilhofer, M., Ladinsky, M.S., McDowell, A.W., Petroni, G., Jensen, G.J., 2011. Microtubules in bacteria: Ancient tubulins build a five-protofilament homolog of the eukaryotic cytoskeleton. *PLoS Biology* 9, e1001213.
- Radermacher, M., 1988. Three-dimensional reconstruction of single particles from random and nonrandom tilt series. *Journal of Electron Microscopy Technique* 9, 359–394.
- Schellenberger, P., Kaufmann, R., Siebert, C.A., *et al.*, 2014. High-precision correlative fluorescence and electron cryo microscopy using two independent alignment markers. *Ultramicroscopy* 143, 41–51.
- Schur, F.K.M., Hagen, W.J.H., de Marco, A., Briggs, J.A.G., 2013. Determination of protein structure at 8.5 Å resolution using cryo-electron tomography and sub-tomogram averaging. *Journal of Structural Biology* 184, 394–400.
- Trachtenberg, S., Dorward, L.M., Speransky, V.V., *et al.*, 2008. Structure of the cytoskeleton of *Spiroplasma melliferum* BC3 and its interactions with the cell membrane. *Journal of Molecular Biology* 378, 778–789.
- Urban, E., Jacob, S., Nemethova, M., Resch, G.P., Small, J.V., 2010. Electron tomography reveals unbranched networks of actin filaments in lamellipodia. *Nature Cell Biology* 12, 429–435.
- Vanhecke, D., Graber, W., Studer, D., 2008. Close-to-native ultrastructural preservation by high pressure freezing. *Methods in Cell Biology* 88, 151–164.
- Villa, E., Schaffer, M., Plitzko, J.M., Baumeister, W., 2013. Opening windows into the cell: Focused-ion-beam milling for cryo-electron tomography. *Current Opinion in Structural Biology* 23, 771–777.
- Wang, Q., Mercogliano, C.P., Löwe, J., 2011. A ferritin-based label for cellular electron cryotomography. *Structure (London, England)* 193 (19), 147–154.
- Weiner, A., Kapishnikov, S., Shimon, E., *et al.*, 2013. Vitrification of thick samples for soft X-ray cryo-tomography by high pressure freezing. *Journal of Structural Biology* 181, 77–81.
- Winkler, H., Taylor, K.A., 2013. Marker-free dual-axis tilt series alignment. *Journal of Structural Biology* 182, 117–124.
- Wu, S., Liu, J., Reedy, M.C., *et al.*, 2009. Methods for identifying and averaging variable molecular conformations in tomograms of actively contracting insect flight muscle. *Journal of Structural Biology* 168, 485–502.
- Wu, S., Liu, J., Reedy, M.C., *et al.*, 2010. Electron tomography of cryofixed, isometrically contracting insect flight muscle reveals novel actin-myosin interactions. *PLoS ONE* 5 (9), e12643.
- Yang, C., Svitkina, T., 2011. Visualizing branched actin filaments in lamellipodia by electron tomography. *Nature Cell Biology* 13, 1012–1013, author reply 1013–1014.
- Zhang, P., 2013. Correlative cryo-electron tomography and optical microscopy of cells. *Current Opinion in Structural Biology* 23, 763–770.
- Zheng, S.Q., Sedat, J.W., Agard, D.A., 2010. Automated data collection for electron microscopic tomography. *Methods in Enzymology* 481, 283–315.
- Zürner, A., Döblinger, M., Cauda, V., Wei, R., Bein, T., 2012. Discrete tomography of demanding samples based on a modified SIRT algorithm. *Ultramicroscopy* 115, 41–49.

Further Reading

- Briggs, J.A.G., 2013. Structural biology in situ — The potential of subtomogram averaging. *Current Opinion in Structural Biology* 23, 261–267.
- Diebold, C.A., Koster, A.J., Koning, R.I., 2012. Pushing the resolution limits in cryo electron tomography of biological structures. *Journal of Microscopy* 248, 1–5.
- Lucić, V., Förster, F., Baumeister, W., 2005. Structural studies by electron tomography: From cells to molecules. *Annual Review of Biochemistry* 74, 833–865.
- Lucić, V., Rigort, A., Baumeister, W., 2013. Cryo-electron tomography: The challenge of doing structural biology in situ. *Journal of Cell Biology* 202, 407–419.

Relevant Websites

- <http://bio3d.colorado.edu/imod/>
IMOD.
- <http://bio3d.colorado.edu/SerialEM/>
SerialEM.
- http://www.biochem.mpg.de/308743/Content_Software
Software in MPI Martinsried.
- http://www.biochem.mpg.de/278655/tom_e
TOM Toolbox.
- http://www.msg.ucsf.edu/Tomography/tomography_main.html
UCSF Tomography.

Immunoelectron Microscopy: High-Resolution Immunocytochemistry

W Möbius, Max-Planck-Institute of Experimental Medicine, Göttingen, Germany

© 2016 Elsevier Inc. All rights reserved.

Glossary

Freeze-fracture Cell or tissue samples are rapidly frozen by HPF or plunge-freezing and then fractured at low temperature. The fracture faces are then replicated by vacuum evaporation of platinum and carbon. After thawing, the replicas are cleaned and viewed in the TEM.

Freeze-substitution Frozen samples are dehydrated at 183 K (−90 °C) by removal of ice with acetone ('substitution'). The freeze-substitution cocktail contains fixing reagents such as uranyl acetate or glutaraldehyde. For purely morphological investigations, OsO₄ is often included in the substitution cocktail. Samples can be embedded at low temperature by resin polymerization with UV light in case Lowicryls are used, or at room temperature with heat-cured resins such as LR White.

High-pressure-freezing (HPF) Cryo-immobilization (also called cryo-fixation) procedure by which 2.045×10^8 Pa

(2045 bar) are applied simultaneous with the cryogen (LN₂) to suppress ice crystal nucleation and slow down ice crystal growth. Vitreous ice formation (ice without crystals) is only achieved in tiny samples (approximately 200 μm in height).

Progressive-lowering of temperature (PLT) The chemically fixed sample is cooled in a stepwise manner following the dehydration steps to 248 K (−35 °C), infiltrated with the resin and polymerized at low temperature by UV light. Suitable resins are LR Gold or Lowicryls.

Tokuyasu cryosectioning method Chemically fixed samples are infiltrated with 2.3 M sucrose, frozen in LN₂, and sectioned at 163 K (−110 °C) in a cryo-ultramicrotome. During pickup from the dry and cold knife and transfer to a TEM grid the sections are thawed and can be used for immunolabeling.

Introduction

For the understanding of cellular processes such as internalization of molecules, intracellular transport during endocytosis or in the biosynthetic pathway, knowledge about the site of action is crucial. Electron microscopy provides means to study the localization of molecules at high resolution in the context of the cellular environment. Here suitable probes such as antibodies are applied to the sample and then visualized with electron-dense markers. In principle, these particulate markers indicate the place where the specific probe is bound to the molecule of interest. This could be a plasma membrane receptor, an internalized ligand, an organelle-specific protein, or even a lipid. Since the marker is visualized on an ultrathin section in the transmission electron microscope (TEM), all the other structures are also visible by their contrast (Figure 1). Contrast formation in the TEM occurs by scattering the electrons of the incident beam while passing through the section, which can be increased by heavy metals introduced into the sample during the sample preparation. Therefore, samples for TEM undergo a sequence of preparation steps which starts with the immobilization/fixation of the living sample and end in a dry and very thin section in vacuum. In this article we will give an overview of the different approaches for sample preparation before discussing strategies in immunoelectron microscopy such as preembedding or on-section labeling using cryo- or plastic sections, whole mount, and replica labeling. Furthermore we will mention correlative light and electron microscopy (CLEM) methods and touch in the end aspects of immunogold label quantification.

Sample Preparation Strategies

The ultimate goal of sample preparation is the preservation of the structural and chemical composition of a cell as it is in the living state. Unfortunately, this is hardly possible, because any treatment of cells or organisms provokes changes which are difficult to control or assess. So our efforts rather produce approximations which have to be controlled by independent methods. For visualization by electron microscopy samples have to comply with electron bombardment under high vacuum conditions. As a consequence for most applications the sample needs to be stabilized or fixed and dehydrated. Both, the method of fixation as well as of dehydration, are potential sources of artefacts which are deviations from the natural state. It is widely accepted that cryo-immobilization by high-pressure-freezing (HPF) and subsequent dehydration/fixation by freeze-substitution and embedding at low temperature represents the best approach for structural preservation of small specimens (up to 200 μm sample thickness and 2 mm in diameter). But in many cases, for example for larger specimens, chemical fixation is a useful sample preparation tool for successful immunoelectron microscopy studies while keeping a critical eye on possible fixation artefacts. As depicted in Figure 2 several sample preparation strategies are applicable for immunoelectron microscopy. For further reading we recommend a recent article by McDonald (2014a).

Chemical Fixation

Aldehydes such as formaldehyde and glutaraldehyde react with amino groups in the sample and induce protein

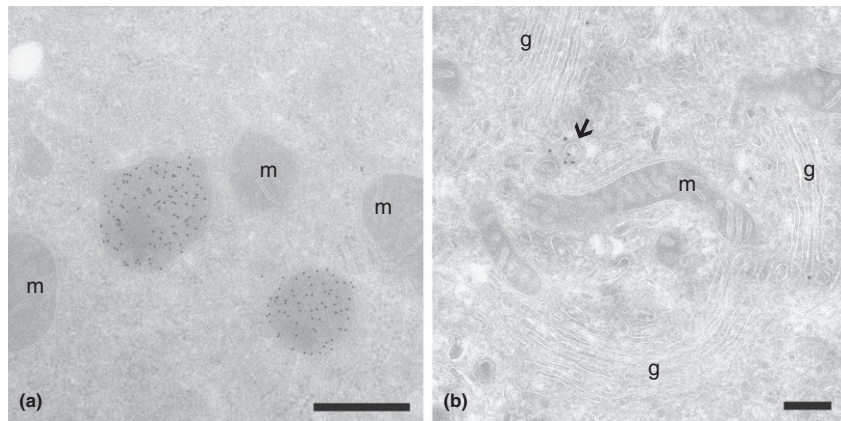


Figure 1 Immunoelectron microscopy provides information about the localization of antigens with high spatial resolution. The cellular context and the structure of the organelles, the so-called reference space, are visible. (a) Immunolabeling of catalase identifies peroxisomes in hepatocytes of rat liver. The primary rabbit antibody is detected with protein A-gold conjugates on ultrathin cryosections. m, Mitochondrium. Bar: 500 nm (b) Golgi complex in a mouse spinal cord neuron. Immunolabeling of a cryosection shows the localization of a γ -adaptin-positive vesicle (arrow) budding from the trans-Golgi network (TGN). g, Golgi complex. Bar: 200 nm.

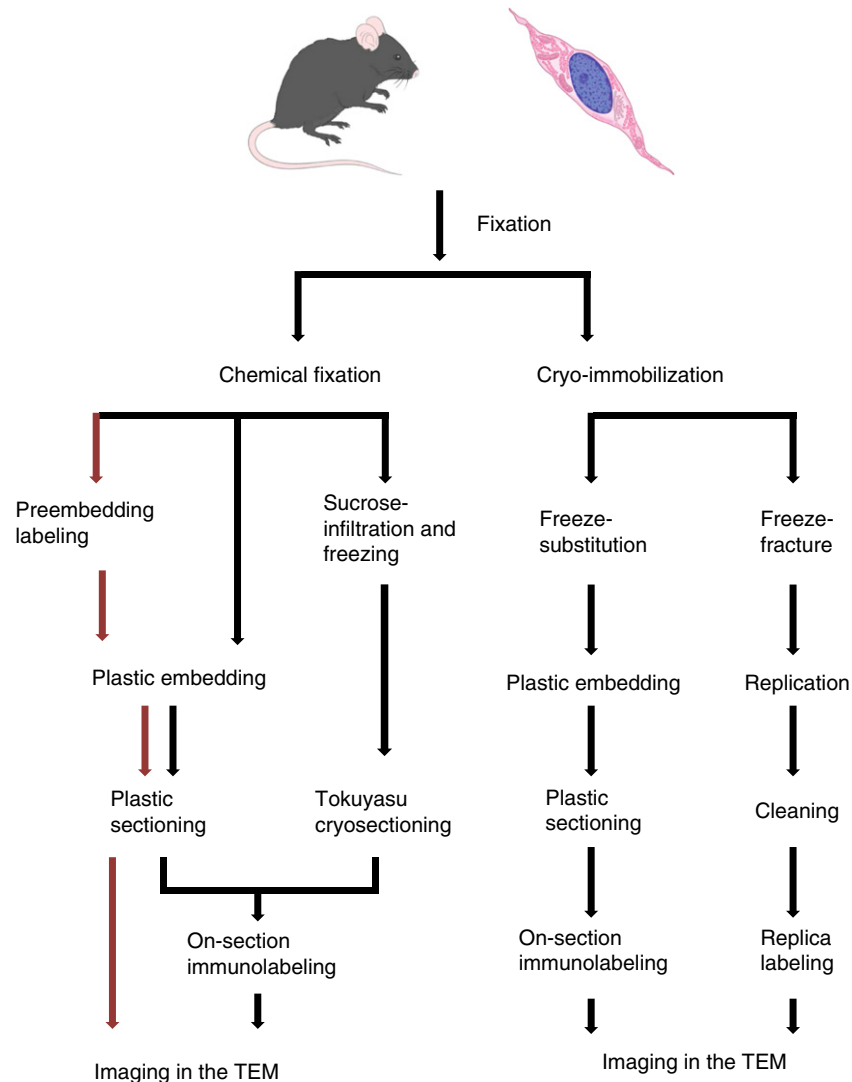


Figure 2 Sample preparation strategies for immunoelectron microscopy.

cross-linking. Higher concentrations of aldehydes result in stronger cross-linking and an improved structural preservation, but unfortunately the ability to localize antigens declines accordingly. To maintain antigenicity, for immunocytochemistry mostly low concentrations of glutaraldehyde (0.1–0.2%) in combination with formaldehyde (2–4%) are used at the cost of the structural preservation. Depending on the nature of the molecules in the sample and the sample size fixation takes between seconds or minutes for cultured cells to up to several hours for tissue. Some cellular constituents such as small proteins or molecules like neurotransmitters are difficult or impossible to fix. It is advisable to fix cultured cells by adding double-concentrated fixation solution directly into the culture dish with the cells in their warm medium to minimize disturbance of the cell's physiology before fixation. Then, after several minutes, the fixative-medium mixture is replaced by fresh fixation solution. For tissue immersion fixation or perfusion fixation are possible. To our experience some tissues like mouse optic nerve or peripheral nerves are very well preserved by immersion fixation. Other organs such as liver or pancreas require perfusion fixation for optimal results. In addition to this for every type of sample the suitable buffer vehicle needs to be determined. Since the reaction of aldehydes with proteins produces H^+ , the vehicle buffer should have a high buffer capacity. Ideally, the osmolality of the buffer is adjusted to the physiological demands of the tissue of interest. Unfortunately, this can only be a rough approximation, because the osmotic properties within a tissue or even a cell can vary significantly. For immunocytochemistry in general, the duration of the fixation should be minimized. The antigenicity of proteins declines with the incubation time. Therefore, in sensitive cases the duration of fixation can be even limited to 1 h. After fixation the samples can be stored in 1% formaldehyde at 4 °C until further processing. We kept samples under these conditions for more than 10 years without recognizable loss of antigenicity. For further reading about the chemistry and application of fixing agents the book by Griffiths, 1993 is recommended.

Plastic Embedding or Cryosectioning of Chemically Fixed Samples

Ultrathin sectioning requires hard blocks. This can be achieved by embedding in plastics which are hardened by polymerization. Freezing is another way of hardening a sample. This approach is applied for ultrathin cryosectioning according to Tokuyasu by embedding the samples in a frozen sucrose solution (Tokuyasu, 1980). The advantage of plastic embedding is the ease of sectioning and handling. In principle, the sample can be dehydrated at room temperature and embedded in plastic which is cured by heat or with an accelerator. One popular resin for this application is London resin (LR) White (Newman *et al.*, 1983). Another approach for plastic embedding of chemically fixed samples is the so-called progressive-lowering of temperature (PLT) protocol. Here, the sample is cooled in a stepwise manner following the dehydration steps to 248 K (–35 °C), infiltrated with the resin and polymerized at low temperature by UV light. Suitable resins are LR Gold or Lowicryl K4M and HM20 which were designed for low

temperature polymerization. The rationale here is the assumption that dehydration and resin polymerization at low temperatures induce fewer changes to antigens compared to processing at 4 °C or room temperature, since at low temperature less molecular movement occurs. Many of these protocols and resins are discussed in Newman and Hobot (1999). Detailed protocols describing PLT methods are also available (Cavalier and Spehner, 2009).

The Tokuyasu technique of ultrathin cryosectioning and immunolabeling does not require dehydration and embedding. Here, the chemically fixed samples are infiltrated with 2.3 M sucrose, mounted as tiny blocks on aluminum sample pins and frozen. These blocks are cut in a cryo-ultramicrotome at –110 °C, ultrathin sections are picked up with a small droplet of pickup solution (Liou *et al.*, 1996), thawed and transferred onto a carbon-formvar-coated TEM grid which could either be used directly for immunolabeling or stored at 4 °C (Griffith and Posthuma, 2002). In this approach any incubation with organic solvents and embedding in plastic resins are avoided and the sections on the TEM grid remain hydrated until the very last drying step after immunolabeling. Therefore, the Tokuyasu approach became popular and is applied successfully until today. For the experienced user of special equipment (cryo-ultramicrotome, cryo-immuno diamond knife, and other tools) this is a very quick method. But, apart from the pure handling also the interpretation of the cryo-immuno data at the TEM requires experience. As pictured in Figures 3(a) and 3(b) the contrast of cellular structures differs from plastic embedded samples (Figures 3(a) and 3(b): comparison of plastic and cryosection). This method is very well suited for membrane structures, but cytoskeletal elements are often insufficiently visible. Due to the lack of any embedding medium, the surface of cryosections is rougher and thereby provides good accessibility for the probes. Moreover, the antigens remain fully hydrated and immersed in the incubation solutions and are better recognized by the antibodies due to less denaturation. For further reading Peters and Piereson (2008) and Slot and Geuze (2007) are recommended.

Cryo-Immobilization and Freeze-Substitution

Fixation implies immobilization. By HPF the specimen is immobilized by purely physical means within milliseconds due to cooling with liquid nitrogen (LN_2). Ideally, the process of freezing preserves a snapshot of all cellular components in the live state. The application of approx. 2.045×10^8 Pa (2045 bar) simultaneous with the cryogen slows down the process of ice crystal nucleation and growth. In the best case, the freezing occurs quicker than the ice crystal formation resulting in so-called amorphous or vitrified ice in samples of approx. 200 μm thickness (Dubochet, 2007). Also at ambient pressure samples can be cryo-immobilized by other techniques such as plunge-freezing where thin samples are rapidly frozen by plunging into liquid propane at LN_2 temperature. In comparison to HPF the zone of vitrified ice is here significantly thinner and mostly confined to 10–20 μm . After obtaining a perfectly frozen and fully hydrated specimen at LN_2 temperature, several processing options are possible. The samples can be used for freeze-fracture cytochemistry as outlined below or

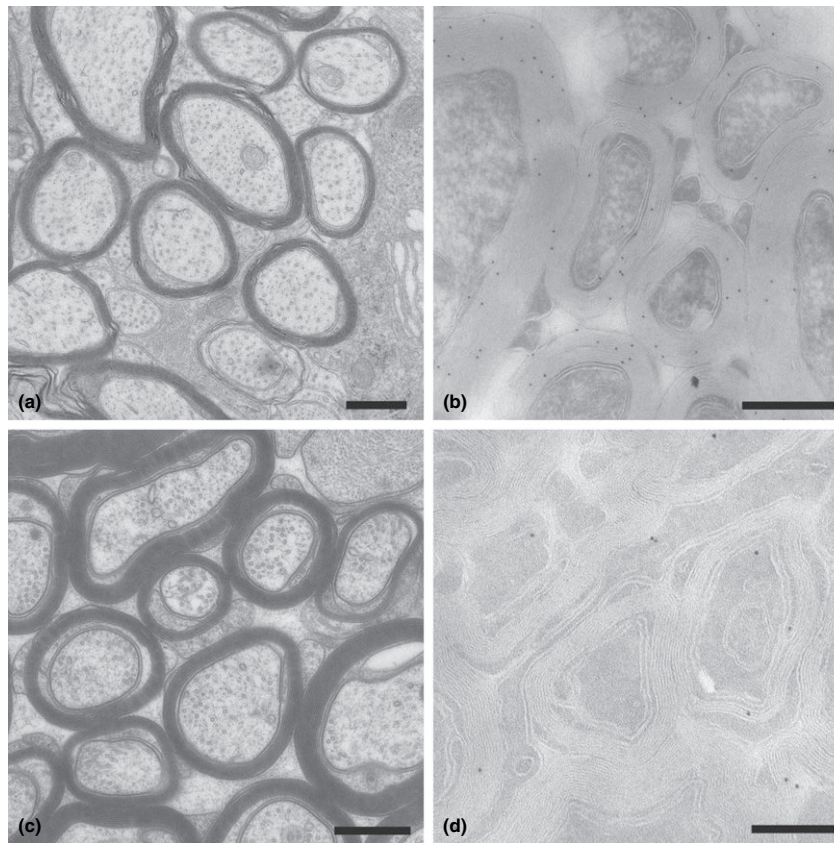


Figure 3 Influence of sample preparation method on the appearance of myelin in mouse optic nerve samples. In conventional plastic embedded samples membranes appear black and cytoskeletal elements like microtubules are clearly visible. In Tokuyasu cryosections membranes appear white. Axonal microtubules are hardly visible. Cryo-prepared samples have a well-preserved ultrastructure. This is especially true for myelin. Here, distortions of the tightly wrapped myelin membranes are avoided. However, cryosections of high-pressure frozen and rehydrated freeze-substituted optic nerve show only sparse labeling. Moreover, the myelin structure appears changed by the dehydration with acetone. (a) Mouse optic nerve fixed by immersion into 2.5% glutaraldehyde and 4% formaldehyde in phosphate buffer at pH 7.2 and subsequently embedded in Epoxy-resin following standard procedures. Bar: 500 nm. (b) Tokuyasu cryosection of a chemically fixed mouse optic nerve (4% formaldehyde and 0.2% glutaraldehyde). Immunogold labeling shows the presence of proteolipid protein (PLP) in the myelin sheath. Bar: 500 nm. (c) Mouse optic nerve prepared by high-pressure-freezing and freeze-substitution followed by Epoxy-resin embedding. Myelin structure is far better preserved, bar: 500 nm. (d) Tokuyasu cryosections after HPF, freeze-substitution and re-hydration for cryosectioning. Immunolabeling demonstrates PLP in compact myelin. Bar: 200 nm.

directly visualized by cryo-imaging techniques such as cryo-TEM after cryosectioning (cryo-electron microscopy of vitreous sections (CEMOVIS) (Al-Amoudi *et al.*, 2004)) or cryo-scanning electron microscopy (cryo-SEM, (Walther and Müller, 1999)) also in three dimensional (3D) using a focused ion beam (cryo-FIB-SEM, (Schertel *et al.*, 2013)).

For immunoelectron microscopy incubation with antibodies at room temperature is required, so the frozen samples have to be transferred somehow to ambient temperature. For this purpose they are stabilized chemically, dehydrated, and embedded in a resin for ultrathin sectioning. In a process called freeze-substitution the ice is removed at 183 K (-90°C) by organic solvents, mostly acetone. The freeze-substitution cocktail usually contains stabilizing and fixing reagents. This could be glutaraldehyde and also uranyl acetate. For Lowicryl-embedding OsO_4 is hardly used, but it has been shown in a study applying Tokuyasu cryosectioning after freeze-substitution that 1–2% OsO_4 in the substitution cocktail did not influence the immunolabeling but increased

membrane contrast (van Donselaar *et al.*, 2007). After a certain time span at 183 K the temperature is raised. It is not well studied how efficient fixatives react at low temperature. However, it was shown by Humbel *et al.* (1983) that glutaraldehyde cross-links bovine serum albumin (BSA) in solution within 1 h at 263 K (-10°C) and within 4 h at 243 K (-30°C).

In the end, the samples are infiltrated with the resin at low temperature (228 K/ -45°C) and polymerized by UV illumination. A typical resin for this procedure would be Lowicryl HM20. Alternatively, the temperature is allowed to reach 277 K (4°C) or room temperature for resin infiltration and polymerization by heat or accelerator, which can be done, for example, by using LR White. A plethora of all kinds of different published protocols is available. However, recent studies have shown that protocols can be simplified drastically for many samples in terms of procedure and equipment (McDonald, 2014a,b). These references are also recommended for further reading on sample preparation. Samples can also be

rehydrated after freeze-substitution and processed for Tokuyasu cryosectioning and immunolabeling if plastic embedding needs to be avoided (Ripper *et al.*, 2007; van Donseelaar *et al.*, 2007; Figure 3(d)).

In comparison to chemical fixation, conventional dehydration, and embedding, this cryopreparation methodology results in structural preservation that far better represents the natural state (compare Figure 3). This might be explained by several important characteristics of cryopreparation: First, the immobilization occurs very fast (in milliseconds) and purely physical without chemical modification of the samples. Second, the immobilization step is separated from the process of fixation or stabilization. This means, all movements have been stopped by freezing, so the fixation takes place in an immobilized sample. Related to this, the diffusion of the fixing reagents occurs prior to the fixation process, since the chemical reactions happen efficiently only during the warm-up phase after substitution and diffusion have been completed. Last, dehydration is performed at low temperature by removal of ice, which is a solid. This process is not well described, but it seems to differ fundamentally from extraction of water from a wet sample resulting in smaller aggregates than in conventional dehydration. Taken together, these properties help avoiding or at least limiting aggregation, shrinkage, extraction of material, and collapse of structures, which are common artefacts of conventional processing (Kellenberger, 1987).

Immunolabeling Strategies

At the beginning of the experiment several strategic decisions have to be made depending on the scientific question: What kind of antigen will be localized? Is it a cytoskeletal protein, a membrane receptor, a transmembrane or membrane-associated protein, a cytosolic antigen, or a lipid? Then, what kind of probe am I going to use? Is it an antibody or another probe like a recombinant binding protein? From which species is the antibody I am planning to use? How likely are the epitopes of the target molecule exposed *in situ*? Depending on all these aspects not only the most suited sample preparation method but also the labeling strategy has to be chosen: preembedding or on-section labeling, whole mount labeling, or freeze-fracture immunocytochemistry.

Preembedding- and On-Section Labeling

In the preembedding labeling approach the whole sample or a large piece of it is incubated with the labeling probes and then embedded, sectioned, and viewed in the TEM (Figure 4). For on-section labeling, also called post-embedding labeling, the sample is fixed and then further processed for sectioning. This could be embedding in Lowicryl or preparation for ultrathin cryosectioning according to Tokuyasu. Then the sectioning is performed and subsequently the sections are incubated with the probes (Figure 5). In comparison to preembedding labeling, the on-section labeling approach is more flexible, since it offers several options for the fixation method ranging from chemical aldehyde fixation to cryopreparation methods as described above.

For preembedding labeling the sample requires chemical fixation which consists usually of a combined formaldehyde

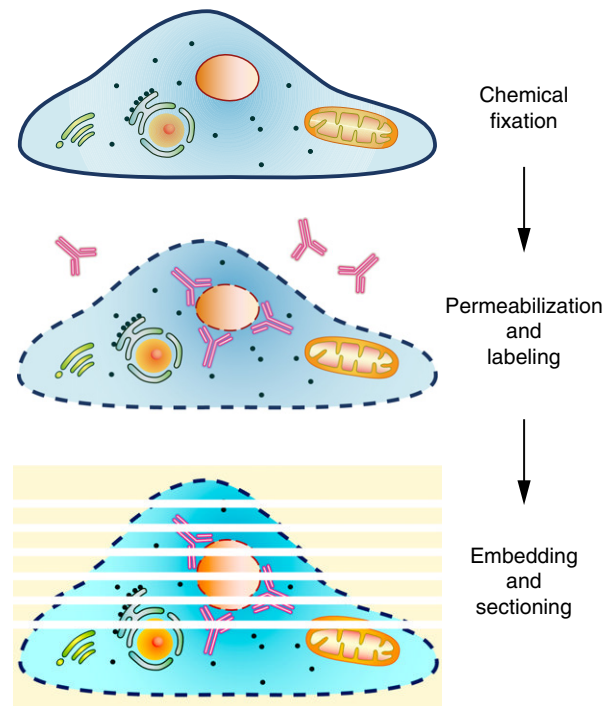


Figure 4 Preembedding labeling: samples are fixed chemically and permeabilized to allow diffusion of antibodies into the sample. After immunolabeling the sample is embedded in plastic, sectioned, and viewed in the electron microscope.

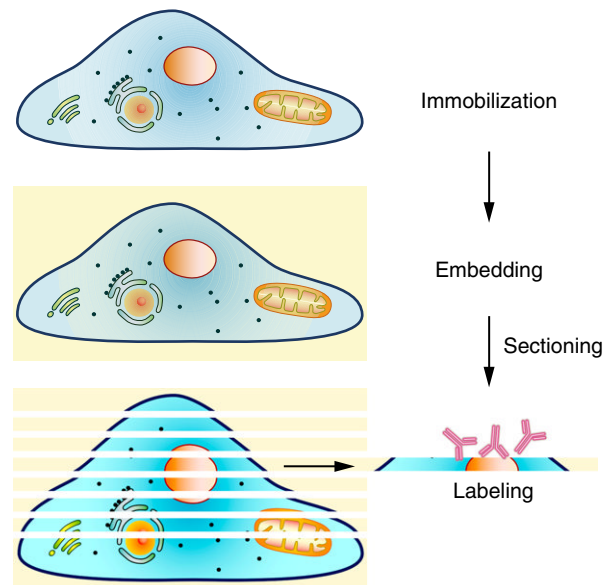


Figure 5 On-section labeling: samples are prepared following the sample preparation method of choice which could be chemical fixation or cryopreparation by HPF and freeze-substitution. After embedding and sectioning, immunolabeling is performed on the sections.

and glutaraldehyde solution in a suitable buffer at suitable pH. The amount of glutaraldehyde, if any, should be low (0.1–0.2%) since this di-aldehyde is cross-linking proteins to a

higher extent than formaldehyde, which is a mono-aldehyde. The extent of cross-linking negatively influences labeling efficacy, because antigens are masked, modified, or become inaccessible to the probe. Furthermore, preembedding labeling requires permeabilization of the sample to permit access of the probes. There are different permeabilization protocols ([Humbel et al., 1998](#)), for example, the use of detergent such as Triton X-100 or saponin, application of NaBH_4 which is a reducing agent used for antigen retrieval, because it breaks fixative-induced double bonds. Another possibility is freezing/thawing which is a purely physical method using ice crystals for permeabilization. The latter has the advantage that it avoids lipid extraction which could happen by application of detergents.

In the next step, the permeabilized sample is incubated with primary antibodies or other probes. Since the diffusion distance can be several micrometers, incubation times are long to permit sufficient infiltration of the probes into the sample. The incubation time with the primary antibody lasts typically several hours or takes even overnight. Accordingly the washing steps to remove unbound probe have to be long as well. Then the secondary antibody or probe is applied. This probe carries the electron-dense marker for visualization in the TEM. For preembedding labeling it is recommended to use antibodies/probes conjugated to ultra-small colloidal gold particles. These gold particles have a size of only a few nanometers in diameter and therefore they are small enough to allow efficient diffusion of the secondary probe into the sample. Because 2–3 nm gold particles are not well visible in standard TEM, these gold particles are enlarged by a so-called silver-enhancement procedure ([Bienz et al., 1986](#); [Yi et al., 2001](#)). After these procedures the sample is embedded in epoxy following a standard-embedding protocol, sectioned, and viewed in the TEM ([Figure 4](#)). Preembedding labeling is useful for surface labeling of whole organisms such as bacteria or living cells. In this case permeabilization and long incubation times are not required. These samples can also be visualized in the scanning electron microscope (SEM). Moreover, preembedding labeling can be applied for 3D immunolocalization ([Morphew, 2007](#)). This reference is also recommended for further reading. Except for pure surface labeling the disadvantages of this approach are the negative effect of permeabilization on the preservation of the ultrastructure and the practical limitations regarding sample preparation.

In on-section labeling antibodies or probes are applied to thin sections of the sample under investigation ([Figure 5](#)). Similar to preembedding labeling the sections are first incubated with the primary antibody/probe and then visualized with an electron-dense marker like IgG-gold or protein A-gold. Since the probes interact mostly only with the surface of the section, diffusion distances are short and therefore incubation times much shorter than in preembedding labeling. A typical labeling protocol is described in [Figure 6](#). The procedure starts with inactivation of residual aldehydes from the chemical fixation, continues with a blocking step to prevent nonspecific sticking of the probes to hydrophobic sites or charges in the sample using BSA, which is then followed by the primary antibody incubation in blocking buffer for typically 30 min to 1 h. Next, the secondary gold-conjugated probe is applied. On cryosections we often use protein A-gold which binds to rabbit IgG (immunoglobulin G) antibodies. Secondary probes are available coupled to different sizes of colloidal gold ranging

from 2 to 40 nm in diameter. Gold-conjugated secondary antibodies (IgGs) are preferred for labeling of Lowicryl sections. For the localization of two different antigens in the sample these two gold sizes allow double-labeling procedures. Also secondary antibodies coupled to ultra-small gold particles can be used. As already mentioned these require enlargement by silver enhancement. For further reading compare [Stierhof \(2009\)](#). In the end the sections are stained with heavy metals and dried before examination in the electron microscope. As already mentioned in the sample preparation section, on-section labeling can be performed on plastic sections or cryosections according to Tokuyasu. Plastic embedding requires dehydration, resin infiltration, and polymerization. During dehydration with organic solvents, extraction and displacement of cellular components can occur. Also membranes are modified, since the lipids are washed out as well. Plastic sections are rather stable, but the resin is also masking antigens. In comparison, Tokuyasu cryosections are resin-free. Therefore, antigens are better accessible. The lack of dehydration by organic solvents makes Tokuyasu cryosections suitable for lipid labeling ([Möbius et al., 2003](#)). But cryosections are also very fragile, because they consist only of the biological material. This can result in loss and also displacement of material in the critical moment when the cryosection is thawing after section pick-up. For the integrity of cryosections the composition of the pick-up solution matters ([Liou et al., 1996](#)).

Whole Mount Labeling

In some cases sectioning is not required for EM-immunocytochemistry. A whole mount preparation of a fibroblast-like cell from chicken embryos published in 1945 illustrates the early days of electron microscopy as a tool to study cell biology ([Porter et al., 1945](#)). The cell was grown on a thin film which was then transferred to a TEM grid and viewed in the electron microscope. Viewing parts of the cell as a whole hence not sectioned can be advantageous for studying complex organelles with branched tubular-vesicular morphology, since information about the 3D organization is often hard to retrieve from thin sections. A method for whole mount immunoelectron microscopy was developed and used successfully to study exit pathways from the endocytic system, which is characterized by a complex tubular morphology ([Stoorvogel et al., 1996](#)). For this purpose, cells are grown on formvar-coated golden TEM grids, incubated with transferrin coupled to horseradish peroxidase (Tf-HRP) for 1 h to specifically target endosomes followed by incubation with diaminobenzidine (DAB), H_2O_2 , and ascorbic acid. By this protocol, the freely diffusing DAB is polymerized only in the endocytic system. Polymerized DAB traps by cross-linking the protein content of the endosomes and makes the entire organelle visible by the electron-dense precipitate. Then, cells are permeabilized to remove the cytosol, immunolabeled, and prepared for TEM by critical point drying. As shown in [Figure 7](#) vesicular tubular endosomes of HeLa cells are visualized with the Tf-HRP DAB method and double labeled for clathrin and cholesterol. Clathrin, indicated by 15 nm gold particles, is mostly detectable on buds, while cholesterol, visualized by using biotinylated perfringolysin ([Möbius et al., 2002](#)) and 10 nm

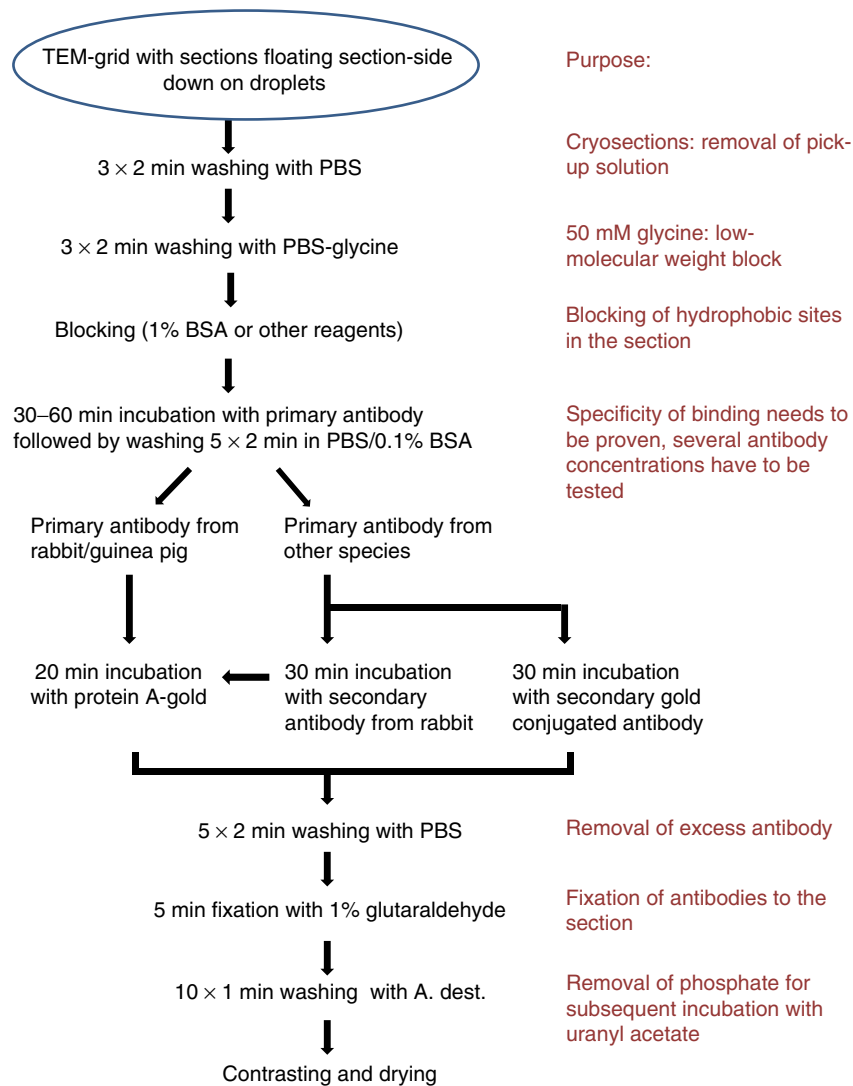


Figure 6 On-section immunolabeling protocol.

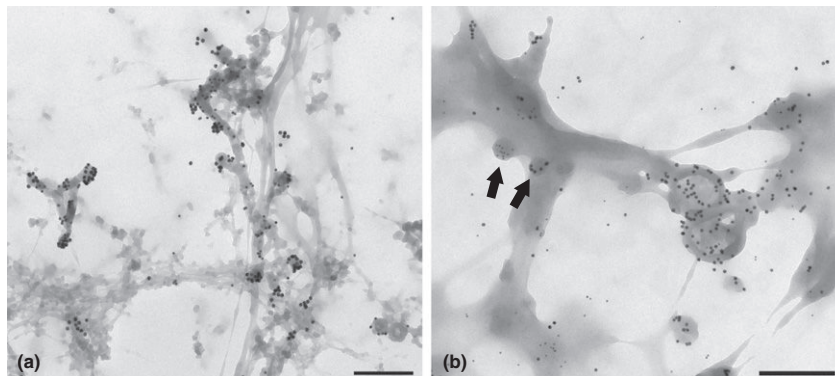


Figure 7 Whole mount labeling of vesicular tubular endosomes in Hela cells: The endocytic system was specifically filled by uptake of HRP-coupled TfR and subsequent DAB-reaction. After permeabilization double-immunolabeling of clathrin (15 nm gold) and cholesterol (10 nm gold) was performed. (a) Clathrin labeling is mostly found of buds, cholesterol labeling localizes also to tubules. (b) Sometimes adjacent buds are exclusively labeled for clathrin or cholesterol (arrows). Bar: 200 nm.

gold particles, is also localized to the tubular parts of the organelle. The whole mount immunolabeling method turned out to be especially useful to study processes which involve changes in the overall morphology of the endocytic system (Kleijmeer *et al.*, 2001). The method is also useful for the visualization of epitopes on viruses, bacteria, vesicles, or anything small that could be mounted directly onto a TEM grid.

Freeze-Fracture Replica Labeling

Another technique that localizes molecules to structures independent of thin sectioning is freeze-fracture cytochemistry. Here, samples are frozen, freeze-fractured and then prepared in such a way that a planar view on membrane fracture faces are created. In combination with immunolabeling this represents a powerful tool to study the size and shape of membrane domains and the lateral distribution of integral membrane proteins within a complex sample.

For freeze-fracture applications specimens can either be fixed chemically and then frozen after treatment with cryoprotectants such as glycerol or directly fresh frozen by HPF. The sample is then fractured in frozen state (at -100°C or lower). The fracture of the sample occurs usually at the weakest parts so that the fracture plane runs along the hydrophobic interior of membranes splitting them into the two leaflets and thereby exposing integral membrane proteins. This newly created fracture face is stabilized and replicated by vacuum-deposition of platinum and carbon. Before this coating an optional 'etching'-step could be included involving ice sublimation to expose more cellular structures which are buried in the ice. Without immunolabeling this replica is then cleaned by incubation with strong acids or sodium hypochloride solution so that the biological material is completely removed before examination in the TEM. For the description of the visible structures a specific freeze-fracture nomenclature was developed (Branton *et al.*, 1975; Figure 8). The lipid bilayer of a membrane consists of two membrane leaflets: the E-half facing the extracellular space and the P-half facing the protoplasm of the cell. After splitting the bilayer into the two leaflets by freeze-fracturing, the E-half has a side facing the extracellular space called E-surface (ES) and a side facing the now exposed interior of the membrane called E-face (EF). Correspondingly, P-surface (PS) describes the cytoplasmic surface of the membrane and P-face (PF) the artificially created fracture face of the inner side of the cytoplasmic leaflet. The P-half of the membranes faces the cytoplasm, the nucleoplasm, and

the matrix of mitochondria or chloroplasts. The E-half is in contact with the extracellular space and the topologically equivalent lumen of all single-membrane enclosed organelles.

Various techniques have been developed to combine freeze-fracturing with immunocytochemistry which can be grouped into three categories depending on the order of preparation steps:

'Fracture-label': immunolabeling is performed after freezing, freeze-fracturing, and thawing of the samples in aldehyde fixing solution. The exposed membrane halves are then accessible to the labeling probes. The resulting sample can be further processed by embedding and thin sectioning or by preparing replicas. For replica processing the labeled samples need to be dried either by critical point drying or freeze-drying before evaporating platinum-carbon on them. After the coating procedure, the biological material is removed leaving only the gold label on the replica. A disadvantage of this approach is that during thawing in aqueous fixation solution a reorganization of the fractured membrane halves into discontinuous bilayers occurs exposing more antigens than those in the original fracture plane. Also the typical freeze-fracture morphology is lost. Moreover, the effect of fixatives on the antigenicity has to be considered.

'Label-fracture': here a suspension of cells is immunogold labeled at the surface, then frozen, freeze-fractured, and replicated. The replica is then only washed with distilled water, so that biological material remains attached. Only in places where the sample was fractured concavely exposing the E-half with the label on the E-surface, the sample is thin enough for the electron beam to reveal the gold particles. Areas with convexly fractured cells still contain too much biological material for visualization of any detail in the TEM. An obvious limitation of this method is the restriction to cell surface labeling and therefore to studies of the E-half.

'Freeze-fracture replica labeling' (FRL): the invention of this technique by Kazushi Fujimoto (Fujimoto, 1997) was a major leap forward in freeze-fracture cytochemistry. Here, non-fixed samples are rapidly frozen by HPF or plunge-freezing with or without pre-treatment with glycerol, fractured, and replicated. Then the biological material is partially removed by prolonged treatment with 2.5% SDS. This leaves behind lipids and membrane proteins attached to the replicas (Figure 9). By this technique any membrane-associated antigen such as caveolin-1 (Figure 10) is well accessible for immunolabeling. The combination of structural information of large membrane regions from the replica with essential information about the

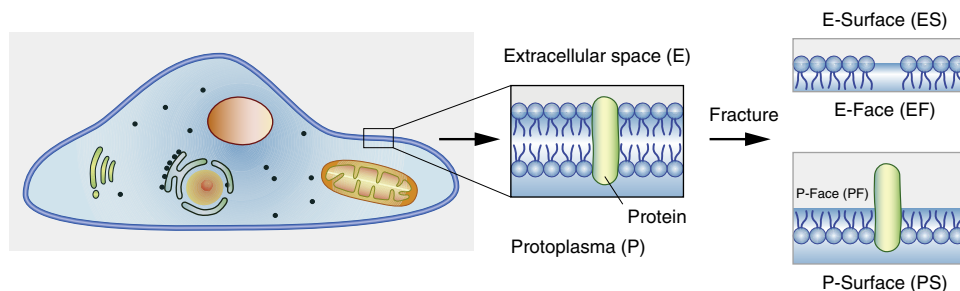


Figure 8 Freeze-fracture nomenclature: the fracture plane tends to follow the inner core of biomembranes splitting them into a cytoplasmic and an extracellular leaflet and exposing the inner side of each leaflet.

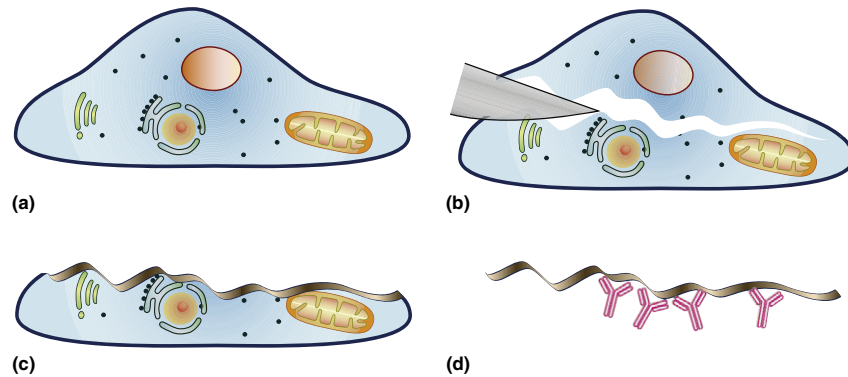


Figure 9 Freeze-fracture replica labeling: after rapid freezing by HPF or plunge-freezing samples are fractured ((a) and (b)). The fractured faces are then replicated by deposition of carbon/platinum layers (c). After reaching room temperature the biological material is partially removed by prolonged treatment with 2.5% SDS. This leaves behind lipids and membrane proteins embedded in the platinum-carbon cast which are well accessible for immunolabeling (d).

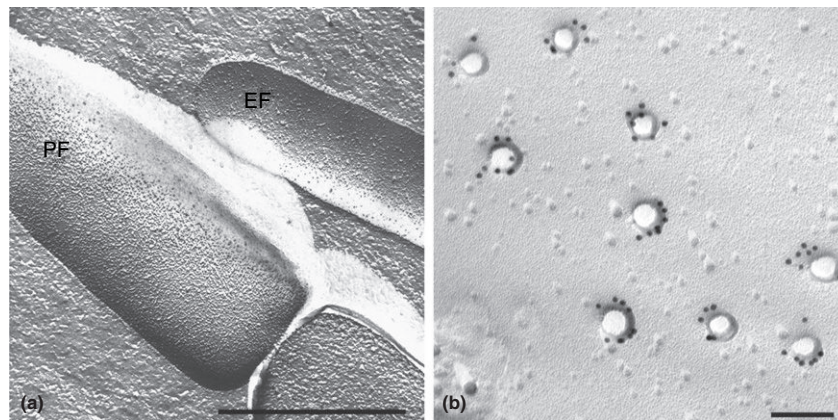


Figure 10 Freeze-fracture and freeze-fracture replica labeling: (a) freeze-fractured *Bacillus subtilis*. Fracturing exposed the membrane interior revealing the two inner faces of the membrane leaflets: the convex P-face (PF) of the cytoplasmic leaflet and the concave E-face (EF) of the extracellular leaflet on an adjacent bacterium. Bar: 1 μm . (b) Freeze-fracture replica labeling of caveolae in 3T3 mouse fibroblasts. Bar: 100 nm. Micrographs courtesy of Martin Westermann, University of Jena, Germany.

composition regarding proteins and lipids contributed substantially to our understanding of many aspects of cellular biology (Severs and Robenek, 2008; Möbius *et al.*, 2010).

CLEM

Electron microscopy offers high resolution imaging at sub-cellular scale, but light microscopy gives an overview of a complex tissue, visualizes processes in living samples, or helps localizing single rare events. The combination of different imaging modalities to examine the same sample in a correlative manner provides more information than any result achieved with only one imaging technique alone. The resolution limit of light microscopy has been overcome by the development of super-resolution fluorescence microscopy techniques. But this is only true for fluorescent signals. For a complete interpretation of localization data, the visualization of the so-called reference space is crucial. The reference space defines the unlabeled cellular context, which remains invisible

in fluorescence microscopy (Griffiths, 2001). This reference space is contributed by electron microscopy.

For CLEM different approaches and levels of complexity are possible. Not all of them necessarily require immunoelectron microscopy. In the simplest way, two similar samples are prepared for the respective microscopy modality and imaged independently. For a real correlation within the same sample, several options are possible: (1) In live imaging or *in vivo* imaging, the fluorescence signals from fluorescently-tagged proteins are recorded in the living specimen before the sample is prepared for electron microscopy for the purpose of either pure morphology or immunoelectron microscopy. The object that was imaged before by light microscopy is then re-localized and analyzed in the TEM. (2) During sample preparation, for example, in the dehydration step by freeze-substitution, fluorescent dyes can be included into the substitution cocktail to facilitate light microscopic observation prior to electron microscopy in order to identify the area of interest (Biel *et al.*, 2003). (3) A fluorescent probe can be used in the same specimen either on an adjacent section or on the same section by applying

on-section labeling. Here, the sample preparation has to fulfill the requirements of both, fluorescence- and electron microscopy (Karreman *et al.*, 2012). For visualization fluorescent- or gold-conjugated probes are applied (Schwarz and Humbel, 2007; Fabig *et al.*, 2012). A probe such as FluoronanogoldTM contains the fluorescent and electron-dense marker on the same antibody molecule (Takizawa and Robinson, 2000).

All these techniques either help finding the relevant spot in a large sample or choosing the right time point in a dynamic process. Depending on the question under investigation, a suitable workflow for correlative light- and electron microscopy needs to be worked out. For further reading Schwarz and Humbel (2009) and the books *Methods in Cell Biology*, volume 111, 2012 and volume 117, 2014 are recommended.

Immunogold Label Quantification and Labeling Specificity

Every experiment, also electron microscopic studies often regarded as 'descriptive,' has to provide quantitative data. A whole range of quantitative tools has been developed for the analysis of thin sectioned biological samples to provide so-called stereological estimations on number and size of structures. These take into account that a normal field of view in the TEM represents only a tiny fraction of the section, which in turn is only a tiny portion of the block, which is again derived from a much larger specimen or organ. The starting point of any quantitative immunoelectron microscopic study is therefore the correct sampling procedure. This means that for a faithful representation of the 'real' situation in the whole specimen, every part of it should have the same chance of being included in the final analyzed image. A tool called systematic uniform random sampling (SURS) applies a regular sampling array and makes sure that the sampling procedure is performed completely unbiased and efficient throughout the whole sampling cascade starting from the animal or cell culture dish, continuing with the selection of blocks and sections down to the level of the field of view in the TEM (Lucocq, 2012). SURS means that samples are selected in a regular systematic way by applying a pattern such as every third culture, block, and section. In the TEM analysis the selection of the starting point occurs at random (e.g., the upper left corner of the TEM-grid mesh), but the positioning of the successive fields of view is then made following a preset systematic pattern.

On randomly sampled fields of view colloidal gold labeling is easily quantified by its particulate nature. Label distribution can be assessed quickly by counting gold particles on organelles while scanning continuously in equally spaced parallel sweeps across the TEM-grid mesh. This label distribution can then be expressed in raw frequency data or as percentage of total number of gold. Often counting of 100–200 particles over 10–15 different compartments reveals a reproducible label distribution (Lucocq *et al.*, 2004). Gold counting in combination with stereological estimations of compartment area and membrane length provides readouts of the label concentration and density. Here profiles of cells or organelles on a TEM picture are overlaid with simple geometrical probes (arrays of test points, lines, rectangles) and encounters of these with structures of interest are counted. This could be points falling into an organelle profile area or

membrane-line intersections. Final values can be expressed as gold per μm^2 or per μm membrane length. Methods for statistical evaluation of gold counting allow determining whether organelles are preferentially or randomly labeled and whether the labeling differs between experimental groups. These employ simultaneous test point counting to create a random distribution. The observed labeling distribution can then be related to the 'expected' random distribution and evaluated by χ^2 statistics (Mayhew, 2005; Lucocq, 2008; Rabouille, 1999).

After having started with sample preparation and ending here with label quantification, lastly one important aspect has to be mentioned: the whole analysis comes down to the quality of the labeling probes, mostly antibodies. In immunoelectron microscopy we completely depend and rely on these tools. How can we be sure that the labeling is specific? Every researcher who has done immunocytochemistry has probably experienced that antibodies are very unreliable and variable tools. This is based on the nature of antibody structure and the biology of immunization. Therefore, specificity of the labeling needs to be proven. A recent review by Griffiths and Lucocq (2014) addresses this important question and is recommended for further reading. In brief, we summarize here several of their criteria which could be applied.

First of all, several antibody concentrations have to be applied. In the best case the labeling distribution is then evaluated by statistics to reveal whether it is random or preferential labeling. The labeling pattern should be in accordance with the published data about the expected localization of the antigen, so labeling of a plasma membrane protein should not be found in mitochondria, nucleus, or cytosol. It is also reassuring to have different antibodies recognizing different peptides from the same protein and giving the same result, although they could also possess the same cross-reactivity. Western blot data often show more than one band, but even a single clean band is no guarantee that the antibody will work specifically in immunoelectron microscopy, because epitopes treated with SDS for Western blots may differ significantly from aldehyde-fixed epitopes for immunocytochemistry. Another aspect is that epitopes on proteins can change during intracellular transport by posttranslational modifications which could affect the antibody–antigen interaction. Moreover, antibodies are able to cross-react with completely unrelated proteins. Knock-out (KO) controls of the protein of interest are helpful if they are available. So, if the antibody recognizes an epitope on a protein different from the target protein, the KO control will also be positive. But the KO control is not always conclusive: for example, under KO conditions related proteins might be up-regulated which could be detected by antibody cross-reaction. Related to the KO control, overexpression of the protein should increase the labeling. This control should be taken with caution, because artificial expression or overexpression can result in mis-localization of the protein and may induce changes of organelle structure. If functional antibodies are not available, often the protein is expressed as tagged version carrying hemagglutinin- (HA), myc- or, GFP-tags. Then tag-targeted antibodies are applied. But unfortunately these tags can change the property of the protein especially when the tag is rather large such as green fluorescent protein (GFP). If a pre-immune serum is available, this can be used to detect pre-existing immunity of the host due to infections. This is relevant for studies on infective pathogens. But in general, the

pre-immune serum is not indicative of the specificity of the antibody generated in the same host later. As another control, pre-absorption of the immunizing peptide should abolish the specific labeling. But this could also remove the cross-reactivity. In addition, the affinity of the antibody to the protein could be higher than to the peptide, so this cannot be detected by pre-absorption. Moreover, storage conditions have an impact on antibody properties. They could lose binding activity due to proteolysis, denaturation and aggregation. One should take care to optimize storage conditions as recommended by Griffiths and Lucocq (2014).

Thus, evaluation of antibody specificity requires a full set of controls, for example: KO control, application of several antibodies raised against different peptides of the same target protein, Western blot data with controls, pre-immune serum testing. Supporting data from independent experiments such as proteome analysis are highly valuable. A very good example of an elaborate evaluation of the specificity of immunolabeling can be found in Limbach *et al.* (2011) including the supplementary information.

Conclusions

- Depending on the question under investigation at the beginning of the experiment decisions have to be made regarding the sample preparation and labeling strategy.
- Samples can be fixed chemically with formaldehyde and low concentrations of glutaraldehyde or cryopreparation methods can be applied: cryo-immobilization by HPF and freeze-substitution.
- Immunolabeling can be performed on whole samples in preembedding labeling and whole mount labeling. Related to this, freeze-fracture immunolabeling is performed of replicas.
- On-section labeling is possible on samples after chemical fixation as well as cryopreparation. Embedding in resin suitable for immunolabeling or cryosectioning according to Tokuyasu can be applied. Cryosections are resin-free, so dehydration with organic solvents is omitted. But the lack of resin makes cryosections fragile.
- For the localization of rare events and for studies of dynamic processes methods for CLEM are available.
- Quantification of gold labeling is feasible by application of SURS and stereological tools for size and number estimations. Methods for statistical evaluation are available.
- Specificity of antibodies is a problem and needs to be demonstrated. Several criteria have to be applied, none of them alone is sufficient.

See also: Imaging the Cell: Electron Microscopy: Scanning Electron Microscopy in Cell Biology

References

Al-Amoudi, A., Norlen, L.P., Dubochet, J., 2004. Cryo-electron microscopy of vitreous sections of native biological cells and tissues. *Journal of Structural Biology* 148, 131–135.

- Biel, S.S., Kwaschinski, K., Wittern, K.P., Hintze, U., Wepf, R., 2003. From tissue to cellular ultrastructure: Closing the gap between micro- and nanostructural imaging. *Journal of Microscopy* 212, 91–99.
- Bienz, K., Egger, D., Pasamontes, L., 1986. Electron microscopic immunocytochemistry. Silver enhancement of colloidal gold marker allows double labeling with the same primary antibody. *Journal of Histochemistry and Cytochemistry* 34, 1337–1342.
- Branton, D., Bullivant, S., Gilula, N.B., *et al.*, 1975. Freeze-etching nomenclature. *Science* 190, 54–56.
- Cavalier, A., Spehner, D., 2009. Progressive lowering of temperature for immunolabeling and *in situ* hybridization. In: Cavalier, A., Spehner, D., Humbel, B.M. (Eds.), *Handbook of Cryo-Preparation Methods for Electron Microscopy*. Boca Raton, FL: CRC Press, pp. 433–465.
- van Donselaar, E., Posthuma, G., Zeuschner, D., Humbel, B.M., Slot, J.W., 2007. Immunogold labeling of cryosections from high-pressure frozen cells. *Traffic* 8, 471–485.
- Dubochet, J., 2007. The physics of rapid cooling and its implications for cryoimmobilization of cells. *Methods in Cell Biology* 79, 7–21.
- Fabig, G., Kretschmar, S., Weiche, S., *et al.*, 2012. Labeling of ultrathin resin sections for correlative light and electron microscopy. *Methods in Cell Biology* 111, 75–93.
- Fujimoto, K., 1997. SDS-digested freeze-fracture replica labeling electron microscopy to study the two-dimensional distribution of integral membrane proteins and phospholipids in biomembranes: Practical procedure, interpretation and application. *Histochemistry and Cell Biology* 107, 87–96.
- Griffith, J.M., Posthuma, G., 2002. A reliable and convenient method to store ultrathin thawed cryosections prior to immunolabeling. *Journal of Histochemistry and Cytochemistry* 50, 57–62.
- Griffiths, G., 1993. *Fine Structure Immunocytochemistry*. Berlin: Springer-Verlag.
- Griffiths, G., 2001. Bringing electron microscopy back into focus for cell biology. *Trends in Cell Biology* 11, 153–154.
- Griffiths, G., Lucocq, J.M., 2014. Antibodies for immunolabeling by light and electron microscopy: Not for the faint hearted. *Histochemistry and Cell Biology* 142, 347–360.
- Humbel, B.M., De Jong, M.D., Muller, W.H., Verkleij, A.J., 1998. Pre-embedding immunolabeling for electron microscopy: An evaluation of permeabilization methods and markers. *Microscopy Research and Technique* 42, 43–58.
- Humbel, B.M., Marti, T., Müller, M., 1983. Improved structural preservation by combining freeze substitution and low temperature embedding. *Beitr. elektronenmikroskop. Direktabb. Oberfl.* 16, 585–594.
- Karremans, M.A., Agronskaia, A.V., Van Donselaar, E.G., *et al.*, 2012. Optimizing immuno-labeling for correlative fluorescence and electron microscopy on a single specimen. *Journal of Structural Biology* 180, 382–386.
- Kellenberger, E., 1987. The response of biological macromolecules and supramolecular structures to the physics of specimen cryopreparation. In: Steinbrecht, R.A.Z.K. (Ed.), *Cryotechniques in Biological Electron Microscopy*. Berlin: Springer, pp. 35–63.
- Kleijmeer, M., Ramm, G., Schuurhuis, D., *et al.*, 2001. Reorganization of multivesicular bodies regulates MHC class II antigen presentation by dendritic cells. *Journal of Cell Biology* 155, 53–63.
- Limbach, C., Laue, M.M., Wang, X., *et al.*, 2011. Molecular *in situ* topology of Aczonin/Piccolo and associated proteins at the mammalian neurotransmitter release site. *Proceedings of the National Academy of Sciences of the United States of America* 108, E392–E401.
- Liou, W., Geuze, H.J., Slot, J.W., 1996. Improving structural integrity of cryosections for immunogold labeling. *Histochemistry and Cell Biology* 106, 41–58.
- Lucocq, J., 2008. Quantification of structures and gold labeling in transmission electron microscopy. *Methods in Cell Biology* 88, 59–82.
- Lucocq, J., 2012. Can data provenance go the full monty? *Trends in Cell Biology* 22, 229–230.
- Lucocq, J.M., Habermann, A., Watt, S., *et al.*, 2004. A rapid method for assessing the distribution of gold labeling on thin sections. *Journal of Histochemistry and Cytochemistry* 52, 991–1000.
- Mayhew, T., 2005. How to count your gold: A tutorial on TEM immunogold label quantification. *Microscopy and Analysis* 12, 9–12.
- McDonald, K.L., 2014a. Out with the old and in with the new: Rapid specimen preparation procedures for electron microscopy of sectioned biological material. *Protoplasma* 251, 429–448.
- McDonald, K.L., 2014b. Rapid embedding methods into epoxy and LR White resins for morphological and immunological analysis of cryofixed biological specimens. *Microscopy and Microanalysis* 20, 152–163.
- Möbius, W., Cooper, B., Kaufmann, W.A., *et al.*, 2010. Electron microscopy of the mouse central nervous system. *Methods in Cell Biology* 96, 475–512.

- Möbius, W., Ohno-Iwashita, Y., Van Donselaar, E.G., *et al.*, 2002. Immunoelectron microscopic localization of cholesterol using biotinylated and non-cytolytic perfringolysin O. *Journal of Histochemistry and Cytochemistry* 50, 43–55.
- Möbius, W., Van Donselaar, E., Ohno-Iwashita, Y., *et al.*, 2003. Recycling compartments and the internal vesicles of multivesicular bodies harbor most of the cholesterol found in the endocytic pathway. *Traffic* 4, 222–231.
- Morphew, M.K., 2007. 3D immunolocalization with plastic sections. *Methods in Cell Biology* 79, 493–513.
- Newman, G.R., Hobot, J.A., 1999. Resins for combined light and electron microscopy: A half century of development. *Histochemical Journal* 31, 495–505.
- Newman, G.R., Jasani, B., Williams, E.D., 1983. A simple post-embedding system for the rapid demonstration of tissue antigens under the electron microscope. *Histochemical Journal* 15, 543–555.
- Peters, P.J., Pierson, J., 2008. Immunogold labeling of thawed cryosections. *Methods in Cell Biology* 88, 131–149.
- Porter, K.R., Claude, A., Fullam, E.F., 1945. A study of tissue culture cells by electron microscopy: Methods and preliminary observations. *Journal of Experimental Medicine* 81, 233–246.
- Rabouille, C., 1999. Quantitative aspects of immunogold labeling in embedded and nonembedded sections. *Methods in Molecular Biology* 117, 125–144.
- Ripper, D., Schwarz, H., Stierhof, Y.D., 2007. Cryo-section immunolabelling of difficult to preserve specimens: Advantages of cryofixation, freeze-substitution and rehydration. *Biology of the Cell* 100 (2), 109–123.
- Schertel, A., Snaidero, N., Han, H.M., *et al.*, 2013. Cryo FIB-SEM: Volume imaging of cellular ultrastructure in native frozen specimens. *Journal of Structural Biology* 184, 355–360.
- Schwarz, H., Humbel, B.M., 2007. Correlative light and electron microscopy using immunolabeled resin sections. *Methods in Molecular Biology* 369, 229–256.
- Schwarz, H., Humbel, B.M., 2009. Correlative light and electron microscopy. In: Cavalier, A., Spehner, D., Humbel, B.M. (Eds.), *Handbook of Cryo-Preparation Methods for Electron Microscopy*. Boca Raton, FL: CRC Press.
- Severs, N.J., Robenek, H., 2008. Freeze-fracture cytochemistry in cell biology. *Methods in Cell Biology* 88, 181–204.
- Slot, J.W., Geuze, H.J., 2007. Cryosectioning and immunolabeling. *Nature Protocols* 2, 2480–2491.
- Stierhof, Y.D., 2009. Immunolabeling of ultra-thin sections with enlarged 1 nm gold or Qdots. In: Cavalier, A., Spehner, D., Humbel, B.M. (Eds.), *Handbook of Cryo-Preparation Methods for Electron Microscopy*. Boca Raton, FL: CRC Press, pp. 587–616.
- Stoorvogel, W., Oorschot, V., Geuze, H.J., 1996. A novel class of clathrin-coated vesicles budding from endosomes. *Journal of Cell Biology* 132, 21–33.
- Takizawa, T., Robinson, J.M., 2000. FluoroNanogold is a bifunctional immunoprobe for correlative fluorescence and electron microscopy. *Journal of Histochemistry and Cytochemistry* 48, 481–486.
- Tokuyasu, K.T., 1980. Immunocytochemistry on ultrathin frozen sections. *Histochemical Journal* 12, 381–403.
- Walther, P., Müller, M., 1999. Biological ultrastructure as revealed by high resolution cryo-SEM of block faces after cryo-sectioning. *Journal of Microscopy* 196, 279–287.
- Yi, H., Leunissen, J., Shi, G., Gutekunst, C., Hersch, S., 2001. A novel procedure for pre-embedding double immunogold-silver labeling at the ultrastructural level. *Journal of Histochemistry and Cytochemistry* 49, 279–284.

Further Reading

- Cavalier, A., Spehner, D., 2009. Progressive lowering of temperature for immunolabeling and *in situ* hybridization. In: Cavalier, A., Spehner, D., Humbel, B.M. (Eds.), *Handbook of Cryo-Preparation Methods for Electron Microscopy*. Boca Raton, FL: CRC Press, pp. 433–465.
- Griffiths, G., 1993. *Fine Structure Immunocytochemistry*. Berlin: Springer-Verlag.
- Griffiths, G., Lucocq, J.M., 2014. Antibodies for immunolabeling by light and electron microscopy: Not for the faint hearted. *Histochemistry and Cell Biology* 142, 347–360.
- Lucocq, J., 2008. Quantification of structures and gold labeling in transmission electron microscopy. *Methods in Cell Biology* 88, 59–82.
- Mayhew, T., 2005. How to count your gold: A tutorial on TEM immunogold label quantification. *Microscopy and Analysis* 12, 9–12.
- McDonald, K.L., 2014a. Out with the old and in with the new: Rapid specimen preparation procedures for electron microscopy of sectioned biological material. *Protoplasma* 251, 429–448.
- McDonald, K.L., 2014b. Rapid embedding methods into epoxy and LR white resins for morphological and immunological analysis of cryofixed biological specimens. *Microscopy and Microanalysis* 20, 152–163.
- Möbius, W., Cooper, B., Kaufmann, W.A., *et al.*, 2010. Electron microscopy of the mouse central nervous system. *Methods in Cell Biology* 96, 475–512.
- Morphew, M.K., 2007. 3D immunolocalization with plastic sections. *Methods in Cell Biology* 79, 493–513.
- Peters, P.J., Pierson, J., 2008. Immunogold labeling of thawed cryosections. *Methods in Cell Biology* 88, 131–149.
- Rabouille, C., 1999. Quantitative aspects of immunogold labeling in embedded and nonembedded sections. *Methods in Molecular Biology* 117, 125–144.
- Schwarz, H., Humbel, B.M., 2009. Correlative light and electron microscopy. In: Cavalier, A., Spehner, D., Humbel, B.M. (Eds.), *Handbook of Cryo-Preparation Methods for Electron Microscopy*. Boca Raton, FL: CRC Press.
- Severs, N.J., Robenek, H., 2008. Freeze-fracture cytochemistry in cell biology. *Methods in Cell Biology* 88, 181–204.
- Slot, J.W., Geuze, H.J., 2007. Cryosectioning and immunolabeling. *Nature Protocols* 2, 2480–2491.
- Stierhof, Y.D., 2009. Immunolabeling of ultra-thin sections with enlarged 1 nm gold or Qdots. In: Cavalier, A., Spehner, D., Humbel, B.M. (Eds.), *Handbook of Cryo-Preparation Methods for Electron Microscopy*. Boca Raton, FL: CRC Press, pp. 587–616.

Imaging Cellular Architecture with 3D SEM

LM Hartnell, LA Earl, D Bliss, A Moran, and S Subramaniam, National Institutes of Health, Bethesda, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Three-dimensional (3D) scanning electron microscopy (SEM) techniques provide information about the ultrastructure of cells and tissues, helping bridge the gap between light microscopy and transmission electron microscopy (TEM). Light microscopic techniques are typically limited in resolution to a few hundred nanometers, while TEM of resin-embedded fixed sections provides a resolution that can be 1–2 orders of magnitude higher, but is essentially limited to two dimensions (2D). To obtain information in the third dimension, a series of consecutive sections can be generated with an ultramicrotome, imaged, and then reconstructed into a 3D volume. Serial sectioning has been used effectively in some instances (Duman *et al.*, 2002; Marsh *et al.*, 1986; Sato *et al.*, 1999). Nevertheless, this approach is subject to limitations from missed sections, as well as artifacts arising from tears and folds in the thin sections imaged by TEM. There are well-established methods to obtain 3D structural information by combining TEM with electron tomography, but this is generally limited to sections that are <250 nm thick. Classical SEM methodologies enable visualization of the surface topology of whole objects, but do not provide 3D information about the internal features of the objects. This limitation can be overcome by combining SEM imaging with the use of a microtome or a focused ion beam (FIB) to progressively remove material from the surface of the imaged object. After each section step, a new surface is revealed; by stacking images generated from these newly created surfaces, it is thus possible to visualize the 3D density distribution in the imaged object. These new 3D SEM methods, which have been developed over the last decade, provide unprecedented tools for imaging cellular structure at nanoscale resolution.

A number of excellent articles have recently been published covering a number of aspects of 3D and serial section SEM imaging (Hughes *et al.*, 2014; Kizilyaprak *et al.*, 2014; Kuwajima *et al.*, 2013; Peddie and Collinson, 2014; Schauflinger *et al.*, 2013; Villinger *et al.*, 2014). In this article, our primary focus is on serial block face SEM (SBF-SEM) and FIB-SEM technologies, with a brief discussion of developments on the horizon of this emerging field.

Sample Preparation for 3D SEM

3D SEM requires that biological samples display a number of specific physical characteristics; these properties are dependent on the physics and mechanics of SEM backscattered imaging, the imaging modality most commonly employed for these technologies.

The first step in preparation of samples for most 3D SEM analyses requires fixation of the biological sample (using either chemical fixation or ultra-rapid freezing methods) (McDonald and Morphew, 1993; McDonald and Webb, 2011), followed by heavy metal staining, dehydration, and

resin embedding, similar in many respects to sample preparation for TEM imaging. This differs from sample preparation for topological SEM, which requires critical point drying and coating post dehydration, rather than embedding (Bray *et al.*, 1993). Although 3D SEM can be performed on samples without resin embedding (Heymann *et al.*, 2006), most delicate cellular ultrastructure is better supported, and there are fewer imaging artifacts, when the sample is resin-embedded (Drobne, 2013).

Because 3D SEM requires iterative imaging with the electron beam, it is important that samples be prepared in such a way that electrons are properly grounded. If the electrons are insufficiently grounded, a build up of negative charge (otherwise known as charging) repels the electron beam and produces imaging artifacts, degrading the quality of the image. Because the samples are resin-embedded (and thus electrically insulated), heavy metal treatment not only yields strong image contrast, but also increases the conductivity of the specimen, thereby reducing charging artifacts.

Orientation and Location of Region of Interest

3D SEM allows for imaging at higher spatial resolution than classical light and confocal microscopy. However, because the resolution is much higher, the fields of view are much smaller than that obtained using light microscopy. Thus, localization of the 3D volume or object of interest within the resin-embedded block of biological material is critical for effective imaging.

One method for locating an area of interest involves taking survey sections of the object within the resin-embedded block with an ultramicrotome (0.5–1 μm). These sections are stained with a histological stain such as Toluidine blue (Hayat, 2000), and correlated with the block so that one may locate the area of interest by SEM.

More recently, several groups have correlated 2D or 3D light microscopy with 3D SEM, allowing the localization of tagged elements within the architecture of the cell. In the study of adherent cells (Murphy *et al.*, 2011; Narayan *et al.*, 2014; Spiegelhalter *et al.*, 2010), cells expressing proteins tagged with fluorescent markers were grown on (or attached to) a marked surface, such as an etched, gridded coverslip. In the methods used for the studies listed above, confocal or standard fluorescent light microscopy is used to identify regions of interest or localize tagged proteins to specific areas of the cell; after light imaging, the samples are fixed, stained, and resin-embedded. The gridded coverslip is then removed, leaving a raised grid pattern on the face of the resin block. This pattern is then used to locate the region of interest on the prepared sample, and to later correlate the two sets of information.

While correlative light microscopy/electron microscopy allows one to generally localize tagged components to 3D SEM data, the recent development of clonable tags that can create electron dense markers offers a unique mechanism to directly image these elements by 3D SEM. APEX (Martell *et al.*, 2012)

and MiniSOG (Shu *et al.*, 2011) are both tags that are fluorescent in samples prior to resin embedding, but become electron dense after heavy metal staining, so that they may be directly observed both by light and electron microscopy.

3D SEM Techniques

A number of different methodologies have been developed for 3D SEM. Primarily, these techniques differ in the way that layers of material are removed (by abrasion or sectioning) and whether the newly exposed block face or the section itself is imaged.

One imaging approach commonly used for tissue imaging is SBF-SEM (Denk and Horstmann, 2004; Hughes *et al.*, 2014; 2013; Motskin *et al.*, 2011). This approach to imaging is commercially available from Gatan, Inc. as the '3View' system, which is a serial sectioning stage used in conjunction with an SEM microscope that collects backscattered electrons. The 3View system is designed with a diamond knife that is used to remove thin sections from the top of the block face, revealing a new face below after each cycle of operation (Figure 1(a)). The top of the block is imaged with SEM, followed sequentially with the removal of a new slice. SBF-SEM most commonly allows for a best x - y resolution of about 5 nm, and a z resolution (slice thickness) of 30 nm, provided that the sample has

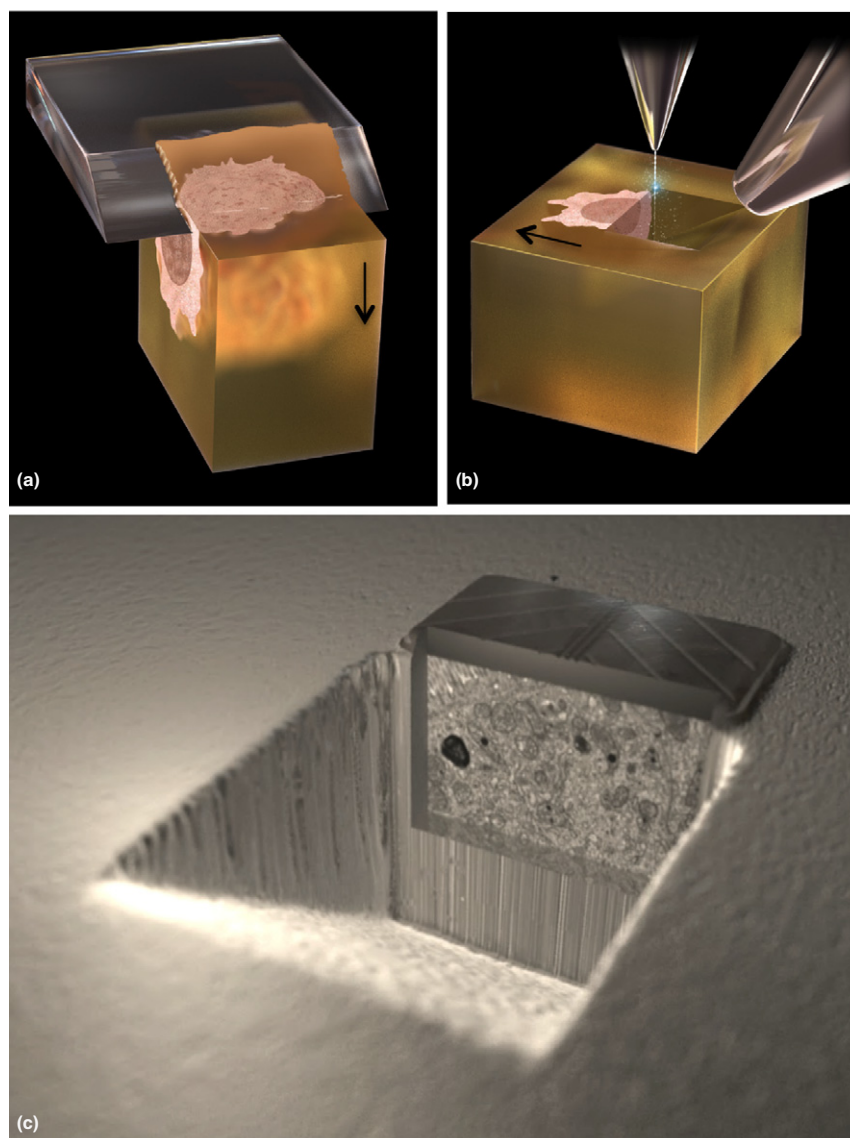


Figure 1 Methods for 3D scanning electron microscopy (SEM) that image the block face. (a) In serial block face SEM (SBF-SEM), a resin-embedded block of cells or tissue is sectioned with a diamond blade inside the SEM stage. The top face of the block is then serially imaged with a scanning electron beam, and backscattered electrons are detected to produce an image. (b) In focused ion beam SEM (FIB-SEM), the 'slices' are abraded with a focused ion beam, typically gallium, and imaged with the electron beam. Imaging is done orthogonally to the top of the block. (c) In preparation for FIB-SEM imaging, the FIB is used to mill a trench in front of the collection area. A pad of carbon and platinum, marked with a pattern for tracking z -spacing of slices, is deposited at the top of the block at the area of data collection.

been prepared optimally for this technique. This tool is especially useful for studying large sample areas (up to $\sim 800\ \mu\text{m}$ in x and y directions), limited primarily in the x - y plane by the size of the diamond knife. SBF-SEM instruments can acquire datasets with potentially thousands of sections (up to $600\ \mu\text{m}$ in the z plane) per imaging session.

In SBF imaging, the build up of charge during the data collection process can lower the quality of the images. One way to address this problem is to integrate a metal coating device into the SEM chamber, as reported by [Titze and Denk \(2013\)](#). Currently, however, the best ways to help dissipate charge and create stable imaging conditions are by coating the sides of the sample with metal, and by infusing additional heavy metals throughout the sample. This addition of metal stains also increases the number of backscattered electrons generated from the sample, thereby increasing the signal. Additionally, traditional epoxy resins may be too soft for SBF-SEM; the use of a harder resin, such as Durcupan ([Deerincx et al., 2010](#)), can be helpful in sample preparation protocols for SBF-SEM, especially when anticipating the need for better z resolution (i.e., thinner slices).

The most common use for SBF-SEM is for the analysis of large volumes. Some examples of these include imaging and following long processes in the central nervous system, and tracking of retinal neurons or long collagen fibrils ([Briggman et al., 2011](#); [Starborg et al., 2013](#); [Wilke et al., 2013](#)). In addition, some studies have also used SBF-SEM for statistical analyses where many cells or bacteria were required, as this method allows for very large datasets ([Hughes et al., 2013](#)).

An alternative approach for 3D SEM imaging is FIB-SEM (also referred to as ion abrasion (IA)-SEM) ([Giannuzzi and Stevie, 2005](#); [Subramaniam, 2005](#)). The instruments used for this method have two beams: a scanning electron beam for imaging the sample and an ion beam that abrades material from the exposed surface of the specimen ([Figure 1\(b\)](#)). While gallium is the material used most commonly in these instruments, other sources such as iridium or elemental gold are also used occasionally.

Because the current in the FIB can be controlled precisely within the pico-ampere range, it is possible to obtain significantly higher z resolution with the FIB-SEM as opposed to the SBF-SEM system. In addition, the closer placement of the sample surface to the pole piece of the electron beam gun (shorter working distance) in the FIB-SEM also allows a higher x - y pixel resolution. Thus, images can be recorded with $3\ \text{nm}$ pixel resolution in the x - y plane and $3\ \text{nm}$ z -slices, generating 3D volumes at significantly higher resolution than possible with SBF-SEM.

The use of two beams in FIB-SEM instruments and the mechanism by which the FIB operates require that the sample stage be tilted, so that both beams act on the sample at the same location. Once the object of interest is located, the stage is tilted ($\sim 54^\circ$ in presently available instruments) and the beams are aligned to be incident at the same region of the specimen. The FIB is oriented directly above the block face, and a pad of carbon and platinum is deposited on top of the imaging area ([Figure 1\(c\)](#)). This pad helps neutralize charging at the sample imaging area, reducing the need for the extra treatment of heavy metals that is required for SBF-SEM. This pad can also provide a platform for a pattern of marks that can

be used to keep track of the thickness of the ablation along the z axis, a process known as z -tracking ([Boergens and Denk, 2013](#); [Narayan et al., 2014](#)). The pad also protects the top of the imaging area from being damaged.

After the protective pad is deposited, a 'trench' ([Figure 1\(c\)](#)) is milled using the FIB, which exposes the imaging area. Whereas the top face of the block is imaged in SBF-SEM, the sample face to be imaged by the electron beam in FIB-SEM is orthogonal to the top face of the block ([Figure 1\(c\)](#)); this must be taken into account when orienting and trimming the sample prior to data collection, as the successive slices removed are from the interior, not the surface of the block.

Typically, FIB-SEM imaging is done on areas that are smaller than that accessible with SBF-SEM. With present-day technology, collection areas with dimensions of $\leq 50\ \mu\text{m}$ in the x direction and $\leq 30\ \mu\text{m}$ in the y direction can be achieved. This limitation derives mainly from the geometry of the imaging as noted above. FIB-SEM technology has been used to collect high-resolution 3D data, on cells and tissues, primarily for smaller volumes at higher resolution. Early FIB-SEM imaging focused on imaging of whole cells and subcellular structures, along with correlation with light microscopy ([Heymann et al., 2006](#); [Murphy et al., 2011](#)). However, several examples of tissue imaging using FIB-SEM have been reported, including 3D imaging of carbon nanotubes in lung tissue ([Kobler et al., 2014](#)), as well as tissues such as cardiac muscle ([Lafontant et al., 2013](#)), skeletal muscle ([Dahl et al., 2015](#)), bone ([Schneider et al., 2011](#)), and brain ([Blazquez-Llorca et al., 2013](#); [Bushby et al., 2011](#)). In tissues such as these and the duodenal epithelium shown in this article, where cellular function is spatially organized within a cell, FIB-SEM can be especially useful for studying the intricate interactions of cellular membranes and organelles.

The higher resolution of FIB-SEM makes it especially suitable for imaging 3D virus distribution in infected cells, as typified in studies of virological synapses formed between various cell types involved in HIV transmission ([Felts et al., 2010](#); [Bennett et al., 2009](#); [Do et al., 2014](#); [Nikolic et al., 2011](#)). Individual HIV cores within mammalian cells, which are $< 40\ \text{nm}$ across, can also be clearly visualized with FIB-SEM ([Narayan et al., 2014](#)), but would be challenging to resolve in 3D with SBF-SEM, even with sections as thin as 20 – $30\ \text{nm}$.

As with any imaging technique, artifacts and defects can be a problem for data acquired by either SBF-SEM or FIB-SEM. SBF-SEM is particularly prone to imaging artifacts caused by charging. FIB-SEM requires a very smooth top block face, as irregularities on the surface at the top of the imaging plane can introduce 'curtaining' defects (vertical marks that resemble a curtain through the volume). The FIB-SEM instrument is particularly useful for working with samples with uneven hardness that are difficult to section, such as bone, or materials that include silicate or carbonate along with biological material ([Reznikov et al., 2013](#); [Hildebrand et al., 2009](#)). In general, SBF-SEM is most useful for the collection of large volumes that need not be imaged at high resolution, while FIB-SEM is better suited to smaller volumes that are to be imaged at higher, isotropic resolution.

A few other specialized methodologies for 3D SEM have also been developed recently, although these tend to have more specific biological applications. One example is a technique developed primarily for imaging neural tissue, where the

sections (rather than the block face) are imaged by SEM. Fixed, stained, and resin-embedded biological material is sectioned with an ultramicrotome, and hundreds of 50–100 nm sections are collected on slides or silicon substrates, and then post-stained again for imaging in the SEM (Horstmann *et al.*, 2012; Kuwajima *et al.*, 2013). Another similar technique is array tomography, where the collected serial sections are first imaged with light microscopy, before being treated with heavy metal stains for SEM. This technique provides correlative 3D immunofluorescence imaging with SEM imaging. In the methodologies where the sections are imaged, the sections may be stored and re-imaged, or in some cases, serially labeled with several different probes (Micheva and Smith, 2007; Oberti *et al.*, 2011; Wacker and Schroeder, 2013).

Data Handling and Segmentation

Despite their differences, both SBF-SEM and FIB-SEM methods produce stacks of 2D images, which then must be aligned into a 3D volume and analyzed (Figure 2). Many algorithms are available that can correct for stage movement and beam instability to produce an aligned stack of images; these alignment procedures can be done with programs such as IMOD (Kremer *et al.*, 1996). In addition, although it is possible to acquire data with fully isotropic voxels, it is more common to have higher resolution in either the x - y or z directions, and these differences can be easily accommodated in preparation of the final aligned stack and in subsequent segmentation steps.

3D SEM methods are an especially powerful tool for visualizing the architecture and connectivity of intracellular organelles. An example of FIB-SEM imaging of murine duodenum epithelium is shown in Figure 2, revealing cellular components with the type of contrast usually observed for thin section imaging with TEM. Thus, within the cell, one can clearly see large branching mitochondria and neighboring endoplasmic reticulum studded with ribosomes. The interface between two epithelial cells can be seen, with desmosomes visible at the basolateral membrane interface. Microvilli can also be seen at the apical surface, with actin bundles extending down into cytoplasm of the cell.

After alignment, segmentation of the individual cellular components is required to analyze the 3D nature of the organelles or features of interest. Segmentation can sometimes be done automatically, using volume rendering tools supported in freely available packages such as IMOD (Kremer *et al.*, 1996), Slicer 3D (Fedorov *et al.*, 2012), ImageJ (Schneider *et al.*, 2012), or UCSF Chimera (Pettersen *et al.*, 2004), or in commercially available packages such as Avizo (Visualization Sciences Group, FEI). Automated tools for segmentation are particularly useful when there are significant differences in contrast between the object of interest and the surrounding morphology (e.g., see Figure 2(b)).

For analysis of more complex intracellular components, segmentation is frequently done manually, by tracing the object of interest on each image throughout the volume, in the xy plane, the xz plane, and the yz plane, to ensure continuity of the object in all directions. Although manual segmentation is

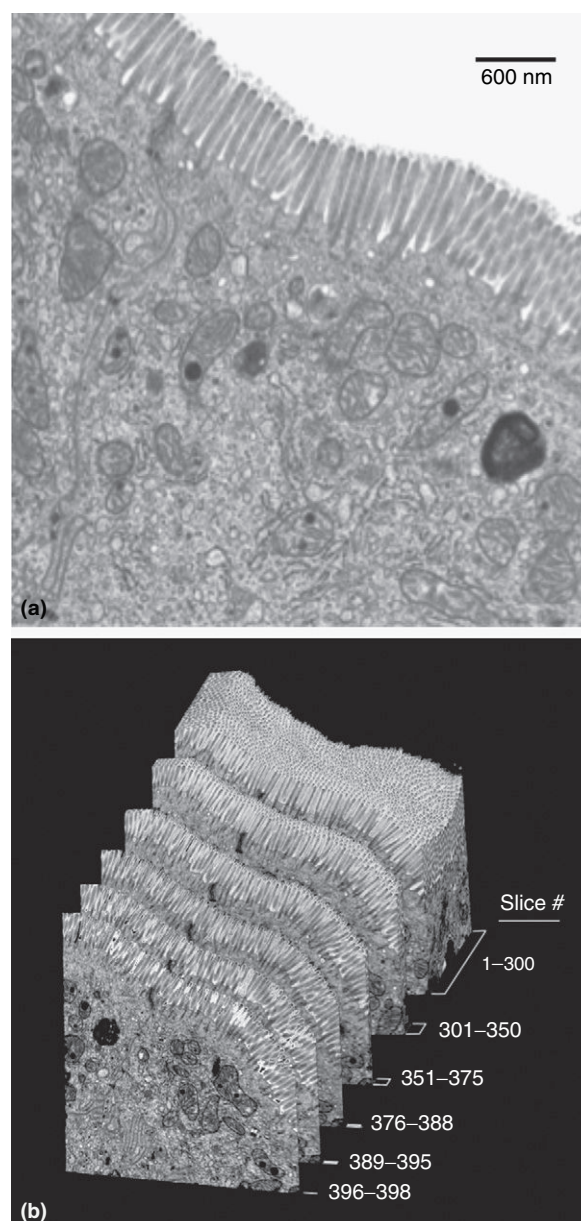


Figure 2 2D scanning electron microscopic (SEM) images are aligned into a 3D volume. (a) 2D FIB-SEM image of duodenum epithelia, showing the brush border and the basolateral cell boundary of two epithelial cells. (b) Individual 2D 'sections' from FIB-SEM imaging correspond to different z -thicknesses in the original tissue. These are computationally aligned and volume rendered. All gray scale images have been contrast inverted to aid interpretation.

time-consuming, it is often necessary when the contrast differences are low, or if the continuity of the object changes within the thickness of a section. One can also use customized procedures to bridge the gap between fully automated and fully manual segmentation; scripts can inform a program such as Slicer 3D (Fedorov *et al.*, 2012) to preferentially select objects of a certain shape or size. This is particularly useful when many similar objects within a cell, like regularly shaped virions or vesicles, need to be segmented and analyzed (Murphy *et al.*, 2011).

Manual segmentation of the duodenal image stack reveals the juxtaposed basolateral surfaces of both cells and the extensive winding and folding of the two membranes (Figures 3(a)

and 3(b)). Individual mitochondria were also segmented manually from this volume. Unexpectedly, 3D segmentation revealed that what appeared to be one large mitochondrion

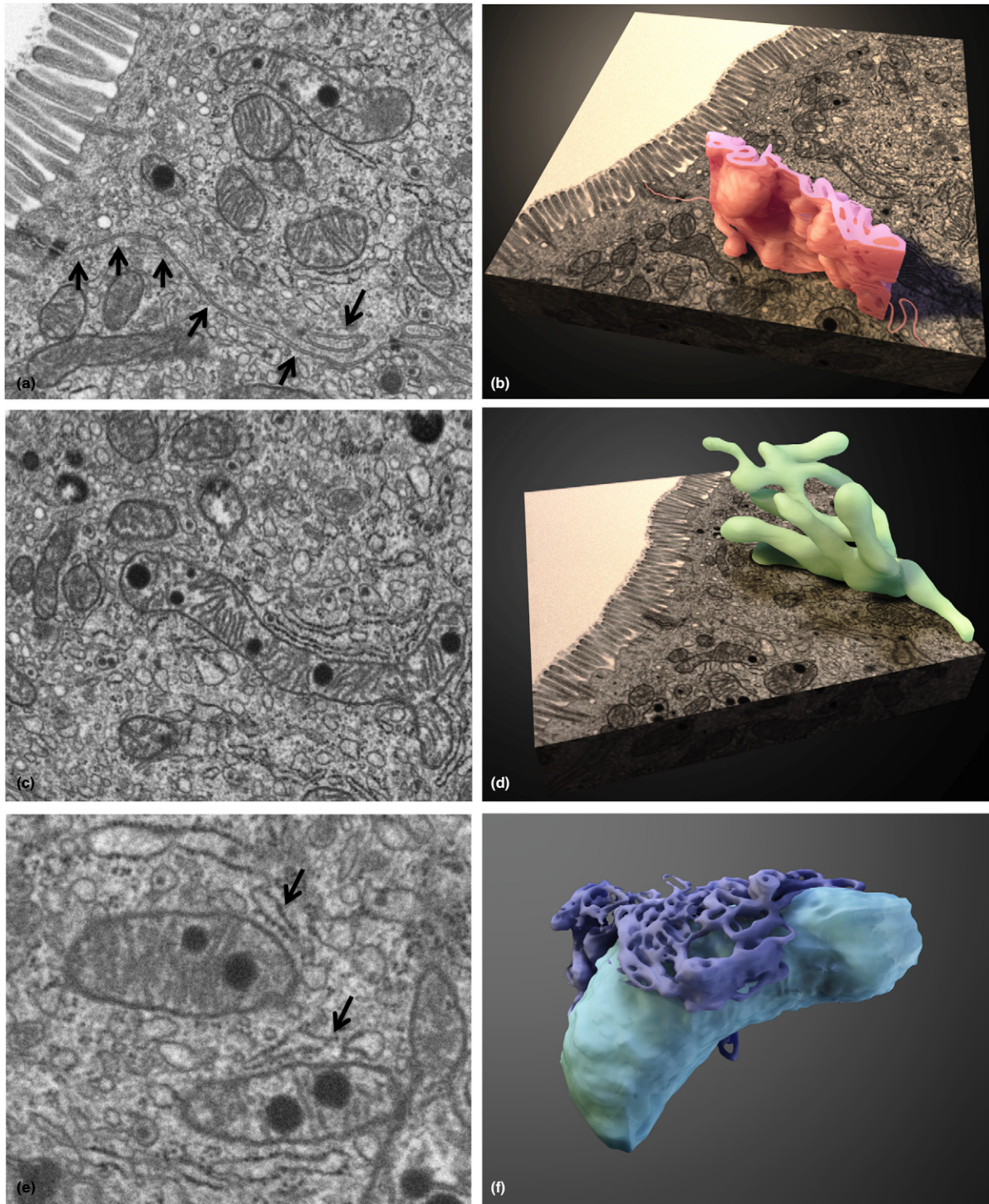


Figure 3 Segmentation of a 3D volume can illuminate the 3D architecture of individual subcellular components. (a) 2D SEM showing the basolateral surfaces between two epithelial cells. (b) 3D segmentation reveals complex folding of the two juxtaposed membranes. (c) 2D SEM image showing a mitochondrion with visible cristae. (d) 3D segmentation reveals the highly branched structure of this single mitochondrion. (e) 2D SEM image showing mitochondrion with neighboring rough endoplasmic reticulum. (f) 3D segmentation reveals lacy rough endoplasmic reticulum wrapping around the mitochondrion. Gray scale images have been contrast inverted to aid interpretation.

and other, small, oblong mitochondria in the 2D sections (Figure 2(a)) were in fact all part of one large mitochondrion with multiple branches (Figures 3(c) and 3(d)). Segmentation of the rough endoplasmic reticulum reveals the structure of fenestrated membranous network that blankets a lobe of mitochondrion, and the image resolution is adequate to visualize individual ribosomes that decorate the membrane (Figures 3(e) and 3(f)).

One of the key reasons to use 3D SEM techniques is that they directly reveal the 3D architecture of organelles and proximity to other objects within a volume. Illustrating these objects and drawing conclusions from the data is thus highly dependent on segmentation and analysis of the data volumes. Choosing whether to segment manually or automatically – or whether to write and use customized scripts to segment and analyze particular aspects of the data – depends on the type of question being investigated and the nature of the data. Provided there is sufficient contrast, a simple volume rendering can suffice for some projects; for precise measurements of objects in just one or two samples, manual segmentation is usually best. However, for statistical analysis of many similar objects from a number of samples, customized scripts can provide higher objectivity and reproducibility. As this technique becomes more common, such scripts may become available for use in open access software such as Slicer 3D (Fedorov *et al.*, 2012).

New Methods and Future Developments

One classic problem in imaging is the trade-off between resolution and data acquisition time. With SEM imaging in particular, a long image acquisition time can lead to imaging artifacts from the build up of charge. One unique advantage of SEM imaging is that it is possible to limit the regions of the block surface that are imaged to only those that are of interest, rather than imaging the entire surface. A recently developed method, ‘keyframe imaging,’ is a promising approach to achieve this type of control in imaging (Narayan *et al.*, 2014). In this approach, one can select specific regions of interest within a 3D volume, and image these at the desired resolution in x , y , and z . This can be combined with recording keyframe images of the entire surface every 30 or 40 slices to obtain a global view of the cell at lower resolution, which can often be of sufficient quality to provide spatial context for the local higher resolution images. This feature was recently developed in the Fibics Inc. software package Atlas3D, which runs the automated data collection process integrated into the Zeiss FIB-SEM instruments. The ‘key-frame’ feature not only allows for a lower resolution overview of the entire imaging area, but also allows the user to dynamically adjust the position of the region of interest during the run. A similar feature is also available for SBF-SEM imaging.

Another important recent technical advance for FIB-SEM imaging is the introduction of z -slice tracking (Narayan *et al.*, 2014). A carbon and platinum pad is deposited on the top face of the resin block, with a defined pattern of lines etched into the pad. As the data run progresses, the distance between the marks changes, and these distances can be used to define the precise z of the image being taken. The feature also allows for a significantly higher level of automation in setup and FIB stability monitoring, as well as automatic stigmation of the electron

beam. These focus-tracking solutions are not yet available for SBF-SEM. In both SBF-SEM and FIB-SEM approaches, the resolutions achievable by SEM do not yet match what can be obtained by TEM, but new detector developments may help to close this gap, and better techniques for reducing sample charging and imaging artifacts and for increasing run stability will likely lead to better quality data. And as 3D SEM becomes more widely used, more publicly available tools and scripts will likely be developed for automated analysis of cells and tissues.

Another area where advances are beginning to be made is in the use of unstained samples for FIB-SEM, where the imaging is carried out with frozen specimens maintained at temperatures below -150°C . When cells and tissues are frozen rapidly and kept below this temperature to prevent ice crystal formation, the cellular components are in a near-native state. This strategy has been difficult to implement mainly because of the absence of contrast-generating stains. In addition, there are difficulties applying a conductive coating to the frozen sample, but progress is being made (Huang *et al.*, 1994; Hayles *et al.*, 2007; Schertel *et al.*, 2013). Future developments that overcome these contrast limitations may enable cryo FIB-SEM to become a more commonly used method for 3D cellular imaging.

In summary, 3D SEM approaches provide unique access to cellular morphology at resolutions 1–2 orders of magnitude higher than that possible with confocal microscopy. Techniques that fall under this umbrella can deliver quantifiable information on the size, shape, and distribution of cellular components and organelles at nanometer resolution, without the need for specific tags, in a highly automated, reproducible, and unbiased procedure. With the new developments in correlative imaging and EM-compatible clonable tags, as well as the use of probes for array tomography, 3D SEM is poised to allow researchers to combine the best aspects of light and electron microscopy in 3D. For biological samples where this type of 3D information is necessary, 3D SEM thus bridges the gap between classical light and EM methodologies, and offers a unique glimpse of 3D cellular and tissue architecture.

Acknowledgments

This work was supported by the Intramural Program of the National Cancer Institute, Center for Cancer Research.

See also: Imaging the Cell: Electron Microscopy: Scanning Electron Microscopy in Cell Biology

References

- Bennett, A.E., Narayan, K., Shi, D., *et al.*, 2009. Ion-abrasion scanning electron microscopy reveals surface-connected tubular conduits in HIV-infected macrophages. *PLoS Pathogens* 5, e1000591.
- Blazquez-Llorca, L., Merchan-Perez, A., Rodríguez, J.R., Gascon, J., DeFelipe, J., 2013. FIB/SEM technology and Alzheimer's disease: Three-dimensional analysis of human cortical synapses. *Journal of Alzheimer's Disease* 34, 995–1013.
- Boergens, K.M., Denk, W., 2013. Controlling FIB-SEM slice thickness by monitoring the transmitted ion beam. *Journal of Microscopy* 252, 258–262. doi:10.1111/jmi.12086.

- Bray, D.F., Bagu, J., Koegler, P., 1993. Comparison of hexamethyldisilazane (HMDS), Peldri II, and critical-point drying methods for scanning electron microscopy of biological specimens. *Microscopy Research and Technique* 26, 489–495.
- Briggman, K.L., Helmstaedter, M., Denk, W., 2011. Wiring specificity in the direction-selectivity circuit of the retina. *Nature* 471, 183–188.
- Bushby, A.J., P'Ng, K.M., Young, R.D., *et al.*, 2011. Imaging three-dimensional tissue architectures by focused ion beam scanning electron microscopy. *Nature Protocols* 6, 845–858.
- Dahl, R., Larsen, S., Dohlmann, T.L., *et al.*, 2015. Three-dimensional reconstruction of the human skeletal muscle mitochondrial network as a tool to assess mitochondrial content and structural organization. *Acta Physiologica (Oxford)* 213 (1), 145–155.
- Deerinck, T.J., Bushong, E.A., Thor, A., Ellisman, M.H., 2010. A new protocol for preparation of biological specimens for serial block face scanning electron microscopy. *Microscopy and Microanalysis* 16 (Suppl. 2), 1138–1139.
- Denk, W., Horstmann, H., 2004. Serial block-face scanning electron microscopy to reconstruct three-dimensional tissue nanostructure. *PLoS Biology* 2, e329.
- Do, T., Murphy, G., Earl, L.A., *et al.*, 2014. Three-dimensional imaging of HIV-1 virological synapses reveals membrane architectures involved in virus transmission. *Journal of Virology* 88, 10327–10339.
- Drobne, D., 2013. 3D imaging of cells and tissues by focused ion beam/scanning electron microscopy (FIB/SEM). *Methods in Molecular Biology* 950, 275–292.
- Duman, J.G., Pathak, N.J., Ladinsky, M.S., McDonald, K.L., Forte, J.G., 2002. Three-dimensional reconstruction of cytoplasmic membrane networks in parietal cells. *Journal of Cell Science* 115, 1251–1258.
- Fedorov, A., Beichel, R., Kalpathy-Cramer, J., *et al.*, 2012. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging* 30, 1323–1341.
- Felts, R.L., Narayan, K., Estes, J.D., *et al.*, 2010. 3D visualization of HIV transfer at the virological synapse between dendritic cells and T cells. *Proceedings of the National Academy of Sciences of the United States of America* 107, 13336–13341.
- Giannuzzi, L.A., Stevie, F.A., 2005. *Introduction to Focused Ion Beams: Instrumentation, Theory, Techniques, and Practice*. New York, NY: Springer.
- Hayat, M.A., 2000. *Principles and Techniques of Electron Microscopy: Biological Applications*. Cambridge, New York, NY: Cambridge University Press.
- Hayles, M.F., Stokes, D.J., Phifer, D., Findlay, K.C., 2007. A technique for improved focused ion beam milling of cryo-prepared life science specimens. *Journal of Microscopy* 226, 263–269.
- Heymann, J.A.W., Hayles, M., Gestmann, I., *et al.*, 2006. Site-specific 3D imaging of cells and tissues with a dual beam microscope. *Journal of Structural Biology* 155, 63–73.
- Hildebrand, M., Kim, S., Shi, D., Scott, K., Subramaniam, S., 2009. 3D imaging of diatoms with ion-abrasion scanning electron microscopy. *Journal of Structural Biology* 166, 316–328.
- Horstmann, H., Korber, C., Satzler, K., Aydin, D., Kuner, T., 2012. Serial section scanning electron microscopy (SSEM) on silicon wafers for ultra-structural volume imaging of cells and tissues. *PLoS One* 7, e35172.
- Huang, C.X., Canny, M.J., Oates, K., McCully, M.E., 1994. Planing frozen hydrated plant specimens for SEM observation and EDX microanalysis. *Microscopy Research and Technique* 28, 67–74.
- Hughes, L., Hawes, C., Monteith, S., Vaughan, S., 2014. Serial block face scanning electron microscopy – The future of cell ultrastructure imaging. *Protoplasma* 251, 395–401.
- Hughes, L., Towers, K., Starborg, T., Gull, K., Vaughan, S., 2013. A cell-body groove housing the new flagellum tip suggests an adaptation of cellular morphogenesis for parasitism in the bloodstream form of *Trypanosoma brucei*. *Journal of Cell Science* 126, 5748–5757.
- Kizilyaprak, C., Bittermann, A.G., Daraspe, J., Humbel, B.M., 2014. FIB-SEM tomography in biology. *Methods in Molecular Biology* 1117, 541–558.
- Kobler, C., Saber, A.T., Jacobsen, N.R., *et al.*, 2014. FIB-SEM imaging of carbon nanotubes in mouse lung tissue. *Analytical and Bioanalytical Chemistry* 406 (16), 3863–3873.
- Kremer, J.R., Mastronarde, D.N., McIntosh, J.R., 1996. Computer visualization of three-dimensional image data using IMOD. *Journal of Structural Biology* 116, 71–76.
- Kuwajima, M., Mendenhall, J.M., Harris, K.M., 2013. Large-volume reconstruction of brain tissue from high-resolution serial section images acquired by SEM-based scanning transmission electron microscopy. *Methods in Molecular Biology* 950, 253–273.
- Lafontant, P.J., Behzad, A.R., Brown, E., *et al.*, 2013. Cardiac myocyte diversity and a fibroblast network in the junctional region of the zebrafish heart revealed by transmission and serial block-face scanning electron microscopy. *PLoS One* 8, e72388.
- Marsh, M., Griffiths, G., Dean, G.E., Mellman, I., Helenius, A., 1986. Three-dimensional structure of endosomes in BHK-21 cells. *Proceedings of the National Academy of Sciences of the United States of America* 83, 2899–2903.
- Martell, J.D., Deerinck, T.J., Sancak, Y., *et al.*, 2012. Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nature Biotechnology* 30, 1143–1148.
- McDonald, K., Morphew, M.K., 1993. Improved preservation of ultrastructure in difficult-to-fix organisms by high pressure freezing and freeze substitution: I. *Drosophila melanogaster* and *Strongylocentrotus purpuratus* embryos. *Microscopy Research and Technique* 24, 465–473.
- McDonald, K.L., Webb, R.I., 2011. Freeze substitution in 3 h or less. *Journal of Microscopy* 243, 227–233.
- Micheva, K.D., Smith, S.J., 2007. Array tomography: A new tool for imaging the molecular architecture and ultrastructure of neural circuits. *Neuron* 55, 25–36.
- Motskin, M., Muller, K.H., Genoud, C., Monteith, A.G., Skepper, J.N., 2011. The sequestration of hydroxyapatite nanoparticles by human monocyte-macrophages in a compartment that allows free diffusion with the extracellular environment. *Biomaterials* 32, 9470–9482.
- Murphy, G.E., Narayan, K., Lowekamp, B.C., *et al.*, 2011. Correlative 3D imaging of whole mammalian cells with light and electron microscopy. *Journal of Structural Biology* 176, 268–278.
- Narayan, K., Danielson, C.M., Lagarec, K., *et al.*, 2014. Multi-resolution correlative focused ion beam scanning electron microscopy: Applications to cell biology. *Journal of Structural Biology* 185, 278–284.
- Nikolic, D.S., Lehmann, M., Felts, R., *et al.*, 2011. HIV-1 activates Cdc42 and induces membrane extensions in immature dendritic cells to facilitate cell-to-cell virus propagation. *Blood* 118, 4841–4852.
- Oberli, D., Kirschmann, M.A., Hahnloser, R.H., 2011. Projection neuron circuits resolved using correlative array tomography. *Frontiers in Neuroscience* 5, 50.
- Peddie, C.J., Collinson, L.M., 2014. Exploring the third dimension: Volume electron microscopy comes of age. *Micron* 61C, 9–19.
- Petterson, E.F., Goddard, T.D., Huang, C.C., *et al.*, 2004. UCSF Chimera – A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* 25, 1605–1612.
- Reznikov, N., Almany-Magal, R., Shahar, R., Weiner, S., 2013. Three-dimensional imaging of collagen fibril organization in rat circumferential lamellar bone using a dual beam electron microscope reveals ordered and disordered sub-lamellar structures. *Bone* 52, 676–683.
- Sato, M., Henson, M.M., Henson Jr., O.W., Smith, D.W., 1999. The innervation of outer hair cells: 3D reconstruction from TEM serial sections in the Japanese macaque. *Hearing Research* 135, 29–38.
- Schaulflieder, M., Villinger, C., Walther, P., 2013. Three-dimensional visualization of virus-infected cells by serial sectioning: An electron microscopic study using resin embedded cells. *Methods in Molecular Biology* 1064, 227–237.
- Schertel, A., Snaidero, N., Han, H.M., *et al.*, 2013. Cryo FIB-SEM: Volume imaging of cellular ultrastructure in native frozen specimens. *Journal of Structural Biology* 184, 355–360.
- Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 Years of image analysis. *Nature Methods* 9, 671–675.
- Schneider, P., Meier, M., Wept, R., Muller, R., 2011. Serial FIB/SEM imaging for quantitative 3D assessment of the osteocyte lacuno-canalicular network. *Bone* 49, 304–311.
- Shu, X., Lev-Ram, V., Deerinck, T.J., *et al.*, 2011. A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms. *PLoS Biology* 9, e1001041.
- Spiegelhalter, C., Tosch, V., Hentsch, D., *et al.*, 2010. From dynamic live cell imaging to 3D ultrastructure: Novel integrated methods for high pressure freezing and correlative light-electron microscopy. *PLoS One* 5, e9014.
- Starborg, T., Kalson, N.S., Lu, Y., *et al.*, 2013. Using transmission electron microscopy and 3View to determine collagen fibril size and three-dimensional organization. *Nature Protocols* 8, 1433–1448.
- Subramaniam, S., 2005. Bridging the imaging gap: Visualizing subcellular architecture with electron tomography. *Current Opinion in Microbiology* 8, 316–322.
- Titze, B., Denk, W., 2013. Automated in-chamber specimen coating for serial block-face electron microscopy. *Journal of Microscopy* 250, 101–110.
- Villinger, C., Schaulflieder, M., Gregorius, H., *et al.*, 2014. Three-dimensional imaging of adherent cells using FIB/SEM and STEM. *Methods in Molecular Biology* 1117, 617–638.
- Wacker, I., Schroeder, R.R., 2013. Array tomography. *Journal of Microscopy* 252, 93–99.
- Wilke, S.A., Antonios, J.K., Bushong, E.A., *et al.*, 2013. Deconstructing complexity: Serial block-face electron microscopic analysis of the hippocampal mossy fiber synapse. *Journal of Neuroscience* 33, 507–522.

IMAGING THE CELL: LIGHT MICROSCOPY

Contents

Genetically Encoded Fluorescent Probes and Live Cell Imaging
Total Internal Reflection Fluorescence Microscopy
Super Resolution Fluorescence Localization Microscopy
Super-Resolution Light Microscopy: Stimulated Emission Depletion and Ground-State Depletion
Structured Illumination Microscopy
Fluorescence Correlation Spectroscopy: A Tool for Measuring Dynamic and Equilibrium Properties of Molecules in Cells
Fluorescence Lifetime Imaging – Applications and Instrumental Principles
High-Speed Localization Microscopy and Single-Particle Tracking
Intravital Microscopy in Mammalian Organisms: From Tissue Physiology to Cell Biology
Optogenetics

Genetically Encoded Fluorescent Probes and Live Cell Imaging

GRV Hammond, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

CEJ Cheetham, Carnegie Mellon University, Pittsburgh, PA, USA

T Balla, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction – Fluorescent Proteins

To coordinate and control the myriad molecular interactions that create life, cells exhibit a remarkable 3D organization. For example, DNA combines with proteins to create neatly packed chromosomes; lipids and proteins combine to form networks of membranes and organelles; and polymeric scaffold proteins crisscross the cell to form the exquisite yet robust cytoskeleton. Therefore, a crucial technique for understanding cell biology has always been the ability to image this intricate cellular architecture. Because so much of this structure is built over the cellular length scale of microns, the light microscope, with its submicron resolution, has always been the best friend of the cell biologist. Indeed, it is no accident that the field of cell biology was born out of the invention of this instrument.

The resolution limit for light microscopy, governed by Abbe's law, is roughly half the wavelength of light used – approximately 0.2 μm for blue light. Nevertheless, most cellular architecture larger than this remains invisible to the microscopist because the molecular assemblies that form it interact weakly with light; the cellular milieu therefore lacks contrast and appears largely amorphous after simple magnification. It is for this reason that staining techniques – the interaction of some light-absorbing substance with specific elements of the cell – are vital to generate contrast for specific architectural elements. In contemporary cell biology, fluorescence staining has become the dominant method, for two principal reasons: firstly, fluorescent probes can be easily and specifically targeted to individual molecular species; and secondly, the labeled

species emit a particular color of light that can be filtered out from all others. The resulting images are unparalleled in their contrast and, with modern confocal microscopy, the resolution attained is as high as Abbe's limit will permit.

Stunning as such fluorescence micrographs of cellular architecture are, they can nonetheless only tell part of the story: chromosomes are in a continuous state of flux, membranes restlessly fuse and bud off from each other, and the cytoskeleton constantly grows and shrinks to keep cells sturdy yet pliable. This incredible dynamism is fundamental to the business of life at the cellular level. For this reason, cell biologists need to be able to watch the living cell as its components move and interact, and fluorescent stains that can be applied to the living cell without perturbing these molecular components are essential tools.

For many years, the availability of such vital fluorescent dyes was limited to the imaging of dynamic cellular processes. However, all that changed with the discovery of the biochemical process that allowed the crystal jelly, *Aequorea victoria*, to generate its ethereal green luminescence. The crucial breakthrough came from the discovery, by Osamu Shimomura and colleagues, that the blue light emitted by the jellyfish's aequorin protein was converted to green by the eponymous green fluorescent protein, or GFP (Morise *et al.*, 1974). After the GFP gene was cloned, Martin Chalfie and colleagues realized that using molecular genetics, this gene could be introduced into cells from any organism. Not only would that cause the cells to fluoresce green, but through fusion of the GFP gene with others, individual fluorescent proteins could be made to

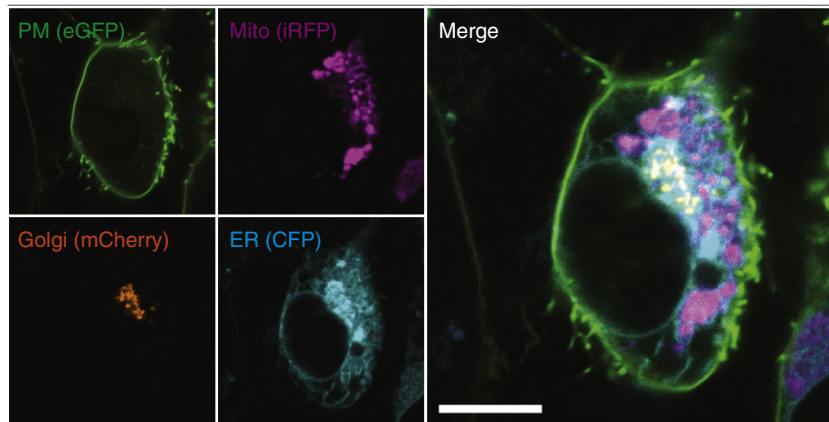


Figure 1 Imaging with multiple fluorescent proteins. A HEK-293 cell expressing enhanced green fluorescent protein (eGFP)-tagged plasma membrane protein (the prenylated polybasic peptide from K-Ras), monomeric cherry fluorescent protein (mCherry)-tagged Golgi protein (the transmembrane tail from Giantin), cyan fluorescent protein (CFP)-tagged ER marker (the transmembrane domains from Sacm11), and the infrared fluorescent protein (iRFP)-tagged mitochondrial resident protein (the mitochondrial targeting peptide sequence from AKAP1). Scale bar=10 μ m.

exhibit this fluorescence, lighting up specific cellular structures (Chalfie *et al.*, 1994). The uniquely compacted, independent folding of GFP, its lack of requirement for cofactors to exhibit fluorescence, and tireless optimization of the gene sequence to improve the efficiency of its production, stability, and brightness by Roger Tsien and others (Heim *et al.*, 1995; Zhang *et al.*, 1996), has yielded one of the most powerful tools in modern cell biology. GFP technology allows cell biologists to introduce fluorescent labeling in living cells to virtually any protein or structure of interest, and to observe this fluorescence in living cells and organisms. It thus represents the marriage of two revolutionary technologies: the historic, physical science of optics with the more recent molecular genetics. This journey from a relatively obscure jellyfish species to a technical tour-de-force in modern biomedicine is also a resounding endorsement of the power of curiosity-driven research in creating serendipitous discoveries and revolutionary new technologies. For bringing GFP to the scientific community, Shimomura, Chalfie, and Tsien were awarded the 2008 Nobel Prize in chemistry.

In the decades since the discovery of GFP, the field of genetically encoded fluorescent probes has exploded. Along with the aforementioned optimization of GFP itself, spectral mutants exhibiting, for example, cyan (CFP) and yellow (YFP) fluorescence have been devised (Heim and Tsien, 1996). Yet more organisms with fluorescent proteins have been discovered, such as the *Discoma* coral with its red fluorescent protein DsRed (Matz *et al.*, 1999), since carefully optimized to produce commonly used bright, stable orange, and red variants that are widely used today, such as mCherry (Shaner *et al.*, 2004). The recent utilization of bacterial phytochrome receptors for the generation of near infrared fluorescent proteins (Shcherbakova and Verkhusha, 2013) completes the visible spectrum of fluorescent proteins and extends it into the invisible infrared region. This makes it possible to introduce several different colored fluorescent proteins and thus to image multiple proteins and structures simultaneously in the same living cell (Figure 1).

The usefulness of fluorescent proteins has been even further extended by the discovery of proteins or mutants in which

fluorescence can be in some way actively controlled. First amongst these was the production of photoactivatable fluorescent proteins (PA-FP) such as PA-GFP, which proceeds from a nonfluorescent to the normal green fluorescent state after illumination with ultraviolet (UV) light (Patterson and Lippincott-Schwartz, 2002). In a similar vein, other proteins have been described that yield changes in fluorescence in response to UV light, for example, from green to red in the case of Eos (Wiedenmann *et al.*, 2004) and its optimized derivatives (McKinney *et al.*, 2009). These photoconvertible fluorescent proteins (PC-FPs) yield additional advantages for studying the dynamics of fluorescently labeled molecules, as we will see in following sections. Note that these fluorescent proteins and their derivatives are continually being discovered and optimized by researchers across the globe; the reader is encouraged to look for recent review articles describing the most up-to-date versions. Excellent recent overviews are listed here (Kremers *et al.*, 2011; Miyawaki, 2011).

Fluorescent Proteins as Probes for Function

As well as revealing the dynamic structural organization of proteins inside living cells, fluorescent proteins have also been adapted in many ingenious ways to enable probing of the dynamics and interactions of the protein molecules themselves. Perhaps the first illustration of this was fluorescence recovery after photobleaching (FRAP); with this method, fluorescent molecules in a defined subregion of the cell are photobleached with intense light (Axelrod *et al.*, 1976). Motion of the now nonfluorescing bleached molecules, combined with their replacement by unbleached fluorescent molecules, causes recovery of fluorescence into this bleached region (Figure 2). By following this process in real time in living cells, the observed kinetics can yield information about the physicochemical processes at work – such as the diffusion coefficient of the protein, or the off-rate of a binding reaction with an immobile partner. Information such as this can yield important insights into how a protein is able to function in its biochemical setting, such as defining the range over which

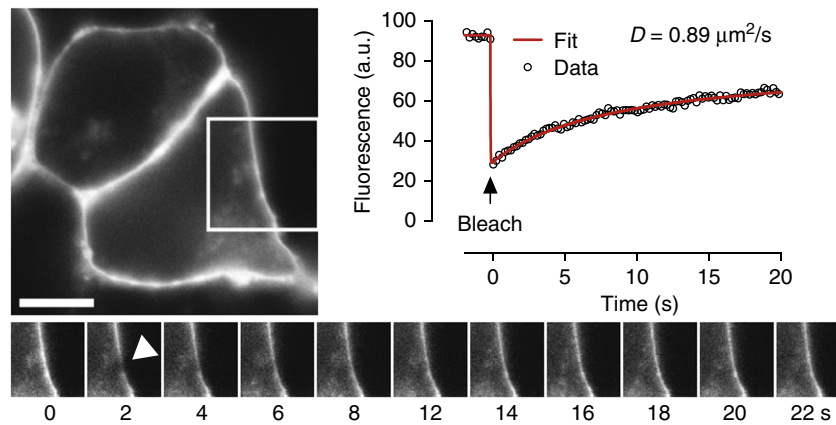


Figure 2 Fluorescence recovery after photobleaching (FRAP). A HEK-293 cell expressing eGFP-tagged plasma membrane protein (the myristoylated and palmitoylated peptide sequence from Lyn kinase) is observed with confocal microscopy. A small region of the membrane was photobleached at high laser power, and the recovery recorded by time-lapse imaging (see insets at bottom). The measured fluorescence intensity in this bleached region can be used to determine the kinetics of the recovery process and, in this case, determine an apparent lateral membrane diffusion coefficient of $\sim 0.9 \mu\text{m}^2 \text{s}^{-1}$. Scale bar = $10 \mu\text{m}$.

an activated protein can exert its biochemical effect (Teruel and Meyer, 2000), or the protein flux through a membrane compartment such as the Golgi apparatus (Patterson *et al.*, 2008). Similar experimental techniques now permit this process in reverse, through the use of PA-FPs or PC-FPs to selectively and acutely activate fluorescence of a sub-pool of fluorescent molecules, and then follow their redistribution in the cell. Although techniques such as these are not limited to fluorescent proteins – indeed, the first demonstrations used antibody-conjugated organic dyes (Axelrod *et al.*, 1976) – the same advantages of fluorescent proteins apply; that is, they can be used to tag virtually any protein of interest and used to study its dynamics.

Another powerful technique for assessing protein dynamics with fluorescent proteins is fluorescence correlation spectroscopy (FCS), wherein the fluctuation of fluorescence intensity in a specific illuminated region of the cell is measured (Bacia *et al.*, 2006). Since this fluctuation occurs due to the motion of fluorescent proteins into and out of the illuminated region, information on protein mobility can be extracted as for FRAP. However, FCS has the capacity to make measurements on a much faster timescale, and therefore to capture more dynamic events. The use of two different colored fluorescent molecules using fluorescence cross-correlation spectroscopy (FCCS) allows the investigation of their interactions (Bacia *et al.*, 2006).

In addition to the ability to image protein localization in living cells, molecular genetics makes it possible to engineer proteins in order to produce biosensors that can report on the presence and/or dynamics of other types of biologically active molecules. Prominent examples include the use of specific lipid-binding domains, fused to fluorescent proteins, to image the localization and dynamics of membrane lipids (Várnai and Balla, 1998; Meyer *et al.*, 1998). This approach has been particularly useful in the study of signaling lipids that are acutely synthesized or degraded in response to cellular stimulation.

GFP has also proved to be a robust scaffold for the generation of other types of biosensor. Mutagenesis of the wild-type GFP sequence led to the discovery of the so-called ‘pHluorins,’

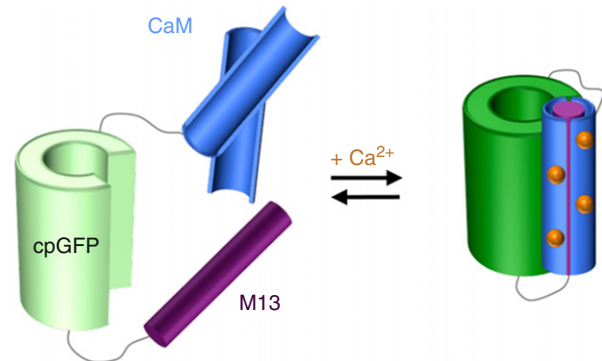


Figure 3 Calcium-dependent fluorescence using GCaMP. In the absence of calcium (left), circularly permuted GFP (cpGFP) is weakly fluorescent due to quenching by water molecules, which in this configuration can access the normally buried chromophore. Upon calcium binding to calmodulin (CaM; right), a conformational change in CaM ‘wraps’ the M13 peptide. This prevents solvent access to the chromophore, and results in a calcium-dependent increase in fluorescence.

whose fluorescence is strongly dependent on pH (Miesenböck *et al.*, 1998). Not only are they a valuable tool for probing the chemical environment of the cell, but they can also be used as a biosensor for regulated exocytosis and synaptic transmission, when acidic vesicles fuse with the plasma membrane exposing them to the neutral extracellular milieu. A strikingly elegant approach to the engineering of GFP into a biosensor was the generation of GCaMP, a GFP whose fluorescence intensity depends on calcium ions, and is thus an indicator of cellular calcium concentration (Nakai *et al.*, 2001). GCaMP consists of an N-terminal ‘M13’ peptide, the circularly permuted GFP (i.e., the C-terminal half of the protein followed by a short linker to the N-terminal half) fused at the C-terminus to the calmodulin protein (Figure 3). In this configuration, the GFP fold forms but is weakly fluorescent due to solvent exposure of the normally buried GFP chromophore. However, binding of calcium

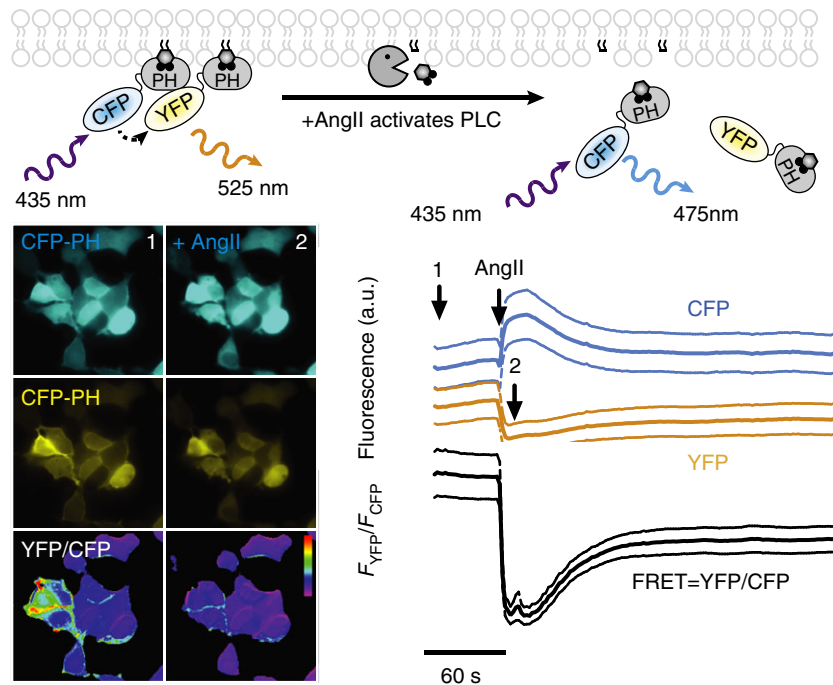


Figure 4 FRET. HEK-293 cells expressing cyan (CFP) or yellow (YFP)-conjugated pleckstrin homology (PH) domain proteins that bind to a membrane lipid, PIP_2 are examined with wide-field fluorescence microscopy. When bound to the lipid, the PH domains are concentrated on the membrane surface, leading to FRET between the CFP and YFP molecules, increasing the FRET ratio of YFP emitted relative to CFP when CFP is excited. However, after activation of a PIP_2 -cleaving enzyme, phospholipase C, the PH domain proteins are released from the membrane and become diluted, increasing the distance beyond the FRET range. Hence the FRET ratio YFP/CFP emission decreases. In this manner, FRET between the two probes acts as a sensitive biosensor for PIP_2 concentration.

by calmodulin causes it to wrap around the M13 peptide, 'plugging' the exposed chromophore (Figure 3) and dramatically increasing fluorescence (Akerboom *et al.*, 2009). Several iterations of GCaMP have since been produced, with dramatic increases in sensitivity. These enable calcium signaling to be monitored in whole cells and organisms with exquisite sensitivity, even allowing synaptic calcium signaling to be viewed *in vivo* (Chen *et al.*, 2013). Coupling of mutated GFPs to voltage-sensitive proteins has also formed the basis for voltage sensors (Siegel and Isacoff, 1997), which have recently been improved by the fusion of a voltage-sensing domain to a pHluorin (Jin *et al.*, 2012). Beyond the utilization of GFP variants, yet another sensitive voltage sensor has recently been engineered from an intrinsically voltage-sensitive and fluorescent archaeal rhodopsin (Kralj *et al.*, 2012). These voltage sensors now have the capacity to detect electrical activity associated with neuronal activity and show great promise for the study of circuit function in living brains.

A spectroscopic quirk of fluorescent molecules, including the fluorescent proteins, allows the nanometer-scale proximity of two such molecules to be probed via the phenomenon of Förster resonance energy transfer (FRET); this is the radiationless transfer of energy between two fluorescent molecules when the emission spectrum of one overlaps substantially with the absorption spectrum of the other (Stryer and Haugland, 1967). This energy transfer is highly dependent on molecular proximity and dipole orientation and can be monitored either

as a decrease and increase in the emission of the donor and acceptor fluorophore, respectively, (called sensitized emission, Figure 4) or as a change in the fluorescent lifetime of the donor (used in fluorescence lifetime imaging microscopy or FLIM (Sun *et al.*, 2012)). FRET between fluorescent proteins was first practically achieved with CFP and YFP (Miyawaki *et al.*, 1997), although FRET pairs have since been systematically improved and often utilize green and red fluorescent proteins (Lam *et al.*, 2012). As well as being a powerful probe for protein–protein interactions, FRET between two fluorescent proteins forms the basis of many biosensors. For example, in the case of lipid biosensors, the increased density of biosensors when bound to lipids in a membrane produces more FRET than with their more dilute unbound forms in the cytosol; hence, FRET efficiency correlates with lipid abundance (Figure 4). Alternatively, biosensors often take the form of a protein receptor whose conformation is shifted upon binding of a small molecule; this shift reorientates N- and C-terminally fused fluorescent proteins, moving them either into or out of FRET distance or orientation. Such tools are extremely useful in the monitoring of small messenger molecules produced during signal transduction, such as cyclic adenosine monophosphate (Ponsioen *et al.*, 2004) or inositol trisphosphate (Tanimura *et al.*, 2009). Coupling of a voltage-sensitive domain to two fluorescent proteins also forms the basis of 'Mermaid2,' a recently described sensitive FRET-based voltage sensor (Tsutsui *et al.*, 2013).

Imaging Cells in Whole Organisms

The cells that comprise complex organisms have a highly ordered 3D organization, forming tissues and organs. Furthermore, each cell interacts with many different cell types and responds to signals from these other cells, the surrounding extracellular environment, and the endocrine, nervous, and immune systems. Therefore, the ability to perform time-lapse imaging at the cellular and subcellular level in living organisms (intravital microscopy) is essential to understand complex processes such as vertebrate embryonic development, tumor metastasis, and brain plasticity. The fluorescent protein revolution, coupled with the development of microscopes capable of gentle deep tissue imaging, has made such ambitious goals a reality. Molecular genetic techniques enable both targeting of fluorescent proteins or biosensors to particular cell types, and temporal control of their expression. Imaging in whole organisms of course presents its own set of challenges; in particular, minimizing photo damage and achieving sufficient depth penetration in a highly light-scattering environment are key factors. In some more superficial tissues, specialized microscopes may not be required. For example, imaging superficial tumor cells (Pittet and Weissleder 2011) or tracking GFP-expressing metastatic cells relative to blood vessels is possible with a confocal microscope (Wyckoff *et al.*, 2000). But for deeper tissues, specialized microscopy techniques are required, and we will discuss the two major methods that have been developed and utilized for this purpose.

2-photon microscopy, as the name suggests, is based on the principle that a single molecule can absorb two photons of light near-simultaneously (Goeppert-Mayer, 1931); hence, a fluorophore can be excited by two lower-energy (longer wavelength) photons if they arrive in a sufficiently short-time window. Making use of newly developed high-powered lasers that could generate very rapid light pulses, Winfried Denk and colleagues built the first laser-scanning 2-photon microscope in 1990, demonstrating that 2-photon excitation of fluorescent dyes was possible in living cells (Denk *et al.*, 1990). 2-photon fluorescence microscopy provides major advantages for imaging in whole organisms. First, it uses infrared excitation, which can penetrate deeper, is less subject to scattering, and because of its lower energy, is less damaging to living cells. Second, fluorescence excitation is confined to a very small focal volume, which greatly reduces photobleaching and photodamage (see later), and generates an optical sectioning effect that permits 3D imaging. Hence, high resolution imaging deep within living tissue is now possible.

Long-term *in vivo* 2-photon imaging has found greatest utility in the mammalian brain (Figure 5). The first live animal 2-photon imaging study to use fluorescent proteins tracked tiny protrusions from neuronal dendrites over hours in the developing rat brain (Lendvai *et al.*, 2000). Even more exciting, the subsequent development of chronic imaging windows (Trachtenberg *et al.*, 2002; Grutzendler *et al.*, 2002) enabled similar longitudinal tracking of subcellular neuronal structures to be performed over weeks or even months even in the adult mouse brain. *In vivo* 2-photon microscopy has now become a widespread technique that enables changes in neuronal structure and function to be tracked during development,

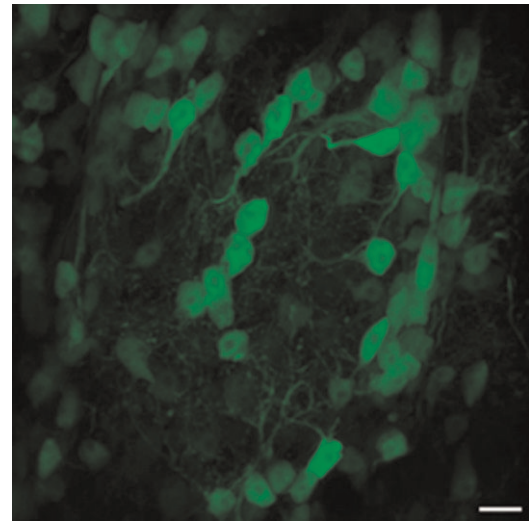


Figure 5 2-photon imaging of GFP-expressing neurons in the living mouse brain. Image shows GFP-expressing inhibitory periglomerular neurons in the olfactory bulb of a young adult mouse imaged through a cranial window. Scale bar 10 μm .

plasticity, learning, and behavioral tasks, even in awake mice (Holtmaat and Svoboda, 2009). 2-photon imaging of neuronal structure and function has also been extended to *Drosophila* (Wang *et al.*, 2003), and even primates (Stettler *et al.*, 2006). Outside of the brain, 2-photon microscopy of fluorescent proteins and biosensors has found utility in the fields of cancer biology and immunology, for example, to monitor gene expression and track cells deep within tumors (Brown *et al.*, 2001), and in developmental biology (Hardin, 2006). The recent development of 2-photon microendoscopy further extends the depth penetration of this already powerful technique (Barretto and Schnitzer, 2012).

More recently, a second class of microscopes that are ideally suited to rapid whole organism imaging have been developed. Light sheet fluorescence microscopy (LSFM), also known as single-plane illumination microscopy (SPIM), uses a thin sheet of light to illuminate only the focal plane of the detection optics. This produces high signal:noise ratio images and as for 2-photon microscopy, by spatially restricting fluorescence excitation, photobleaching and phototoxicity are minimized. Fast data acquisition is then achieved by simultaneous detection of emitted light from the entire illuminated sheet using a sensitive camera, in some cases with subcellular resolution. LSFM permits very gentle live imaging, enabling many thousands of images of entire organisms to be collected over long periods of time.

The first application of LSFM was to image the rhodamine-stained guinea pig cochlea after using a chemical clearing agent to render the tissue transparent (Voie *et al.*, 1993). But once again it was GFP that really brought LSFM to the forefront, when Ernst Stelzer's laboratory showed that it enabled GFP-labeled muscles throughout a transparent fish embryo to be visualized (Huisken *et al.*, 2004). This opened the door to continuous imaging of embryogenesis over many hours or even days, providing unprecedented insight into this dynamic process in fish and insect embryos (Keller and Dodt, 2011). A

type of LSM termed objective-coupled planar illumination has also been used to simultaneously track GCaMP2 calcium signals from hundreds of neurons in the living mouse brain (Holekamp *et al.*, 2008).

Where dynamic information is not required, the use of clearing techniques also allows opaque tissues and organisms to be studied with LSM: structures deep in the mammalian brain, whole mouse embryos and even entire adult *Drosophila melanogaster* fruit flies can be visualized (Dodt *et al.*, 2007). Whatever the tissue or organism of interest, the ability to sweep the illumination sheet now permits generation of 3D data sets without moving the specimen (Keller *et al.*, 2008). LSM technology is advancing all the time, and the range of applications it can be used for grows concomitantly. We direct the reader to recent review articles for a more complete description of the state-of-the-art, but note that LSM can now be combined with 2-photon excitation and even super-resolution techniques (see below) (Keller and Dodt, 2011).

Breaking the Diffraction Barrier

Although it was long accepted as an immutable law of physics, Abbe did not have fluorescence in mind when deriving his law of limiting resolution. As such, even though the light emitted from a fluorescent source – such as a fluorescent protein – is subject to blurring into a diffraction-limited spot by the microscope optics (its so called point spread function), higher resolution information is attainable if we only know the location in which that molecule was excited (Hell, 2008). Several ingenious methods for achieving this goal have recently been invented, all of which lend themselves to imaging fluorescent probes inside cells.

The first broad method to achieve super-resolution is to reduce the effective region in which fluorescent molecules become excited. Saturated structured illumination microscopy (SSIM) uses patterned illumination to limit the excitation pattern of the fluorescent molecules in the sample, leading to a doubling of the resulting resolution, and often revealing key structural features that are normally hidden just beneath Abbe's limit (Schermerle *et al.*, 2008; Dobbie *et al.*, 2011). Patterned illumination in an excitation plane orthogonal to the observation axis has also recently been demonstrated to provide this increased resolution, but with greatly improved speed, depth of penetration, and reduced phototoxicity (Gao *et al.*, 2012). Stimulated emission depletion (STED) microscopy similarly constrains excitation, but in this case in a point-scanning system: a 'donut' of light is used to deplete fluorescent molecules from their excited state around the periphery of the diffraction-limited excitation beam. This leads to the effective shrinking of the 'spot' of excited fluorescence from ~250 nm down to ~50 nm, and can now be achieved along the optical axis as well as laterally (Hell *et al.*, 2008). STED has been successfully applied to fluorescent protein samples in living cells, and has even allowed 3D imaging of structural dynamics in the endoplasmic reticulum that are ordinarily occluded by the diffraction limit (Hell *et al.*, 2008), as well as *in vivo* time-lapse imaging of dendritic spines in the living mouse brain (Berning *et al.*, 2012).

The second broad method involves imaging of individual fluorescent molecules; although the light emitted from a single protein produces a diffraction-limited spot, the predictable profile of this point spread function permits the localization of the molecule to be determined at a far higher precision – typically on the order of ~20 nm for the number of photons emitted by a fluorescent protein (Hell, 2008). Armed with this information, the precise position and dynamics of individual GFP-conjugated molecules can be calculated, allowing the molecule to be tracked (Figure 6), and parameters such as diffusion coefficients or membrane binding times to be calculated (Mashanov *et al.*, 2004). Of course, the trick with experiments such as this is to ensure that isolated single molecules can be observed (since two or more fluorescent molecules separated by less than the diffraction limit will appear as a single spot), requiring careful control of expression levels. However, the use of PA-FPs and PC-FPs allows limited activation of a sub-pool of fluorescent proteins, requiring far less stringent control of the expression level (Manley *et al.*, 2008). As an added bonus, once the first cohort of activated molecules has been bleached or has left the imaged field, further cohorts can be activated iteratively, greatly increasing the density of data that can be generated from a single cell. This makes it possible to produce a 'road map' of the mobility of thousands of individual molecules at the cell surface (Manley *et al.*, 2008).

Another offshoot of this capacity to localize isolated single molecules with high precision from a larger pool of inactive PC-FPs or PA-FPs is that, by repeating this process many thousands of times, a very high-density map of the protein distribution emerges, which effectively constitutes a super-resolution image (Figure 7). In a fixed specimen, or a live specimen where the structure of interest will not move substantially over the imaging period, this forms the basis of photo-activation assisted localization microscopy, or PALM (Betzig *et al.*, 2006). It should be noted that through a variety of technical modifications the same principle can be applied to nongenetically encoded fluorescent probes, with a variety of different monikers such as stochastic optical reconstruction microscopy (STORM) (Manley *et al.*, 2011). Although this is a relatively slow mode of imaging compared to STED or SSIM techniques, PALM and its brethren allow the probing of cellular architecture with levels of resolution only previously achievable with the electron microscope (Betzig *et al.*, 2006).

Technical Considerations for Live Imaging

So far, we have explored some of the many different ways in which genetically encoded probes can be used to observe and measure the intricate inner workings of the cell. Now we will discuss some of the technical challenges that must be overcome in order to use these tools effectively. Consider the requirements for a successful experiment on single cells expressing a fluorescent protein-conjugated probe: to begin, the genetically encoded probe must be introduced and expressed by the cells. Next, we need to observe the cells on the instrument that will record the data – almost always a microscope for the purposes of this discussion. This brings us to several requirements for collecting high quality data; firstly, we need

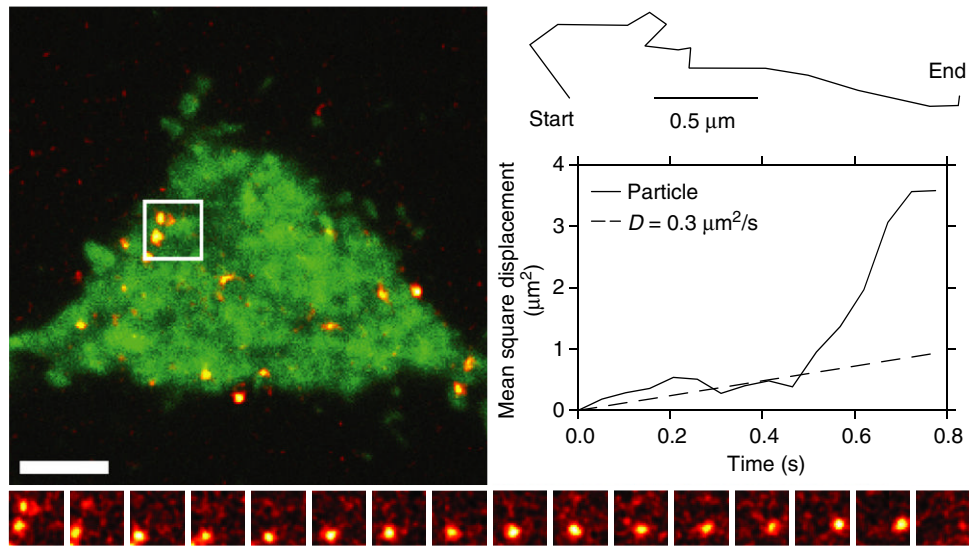


Figure 6 Single-particle tracking with fluorescent proteins. Here, a HEK-293 cell is viewed in the plane attached to the coverslip (TIRF microscopy) whilst expressing a plasma membrane-bound lipid binding protein tagged with the photo-convertible fluorescent protein mEos2. Before conversion, mEos2 exhibits green fluorescence, but after low intensity illumination with violet light, a subset of individual molecules shift to orange/red emission, as shown in the larger image. The inset shows time lapse imaging (at approximately 50 ms intervals) of a single molecule moving on the surface of the cell membrane. This particle's position was calculated and tracked to produce the trajectory shown at top right, and used to calculate the mean square displacement over time. The dashed line on this plot shows the average mean square displacement over time for many thousands of such tracks from this cell, corresponding to a lateral diffusion coefficient of $0.3 \mu\text{m}^2 \text{s}^{-1}$. Scale bar = $5 \mu\text{m}$.

to be able to detect the signal with sufficient separation from the background noise to produce a clear image and/or accurate measurement. Secondly, this signal must be detected with sufficient temporal and spatial resolution to cover the cellular structure or process under investigation. Finally, fluorescence and cell health must be preserved for the entire duration of the experiment. Unfortunately, many of these requirements make contrary demands on the instrument, and the optimum compromise has to be found. We will briefly describe some of the major considerations in the following section.

With modern molecular biology, stable or transient introduction of new or modified genetic material into most cell types and organisms is possible, and often trivial. A detailed analysis of the state of the art of these methods is beyond the scope of this article, but we will briefly touch on one vital and sometimes overlooked consideration: the control of expression level. At first glance, it may seem that maximal expression of any given probe is desirable to generate the strongest signal. However, over-expression can have its own problems. For example, an over-expressed probe for a small molecule may introduce buffering effects, and an over-expressed protein conjugated to a fluorescent protein may saturate its endogenous binding partners and occlude its normal cellular distribution. Construction of fusion proteins can also be problematic: fusion of a fluorescent tag at the N- or C-terminus of a protein may occlude signal sequences or cause steric hindrance of other interactions necessary for proper targeting. Many of these pitfalls can also block essential cellular activities and eventually lead to toxicity. When designing experiments with genetically encoded fluorescent probes, it is therefore vital to design appropriate control experiments to ensure both that the probe itself is not perturbing the cell, and

that expression levels are not so high as to cause artifacts. That being said, a study of almost 900 human proteins showed that over 80% localize appropriately when over-expressed as GFP fusion proteins (Stadler *et al.*, 2013). Whereas both N- and C-terminal tags should be tested, Stadler *et al.* (2013) did report that C-terminal tags were less likely to cause protein mislocalization compared to the endogenous population.

The next consideration is to select the appropriate instrument for the experiment. Measurements of whole cell intensity or population responses measured with a biosensor may well be achieved with fluorescence-activated cell sorting or spectrometers, respectively. However, if subcellular detail or single cell responses are required, the light microscope is the tool of choice. Modern high-resolution fluorescence microscope systems are extremely costly items, usually in the realms of ~US\$100 000 to US\$1 million. Therefore, the choice of instrument is often dominated by availability. Nonetheless, a good understanding of the relative strengths of different systems is useful. But before we can choose the best microscope for the job, we need to consider the requirements of our experiment.

Having optimized expression of the probe, we need to ensure that we can record a sufficiently clear signal to generate a useful image or measurement. The technical boundaries of this consideration are constantly being pushed back, both in terms of the sensitivity of microscope cameras and photomultipliers, and the development of ever-brighter fluorescent protein derivatives and more sensitive biosensors. Therefore, it may well be possible to improve an experiment with newer reagents or equipment. Once a system has been chosen and optimized, we next need to consider the appropriate imaging parameters: essentially, the aim is to maximize the collected signal, and obtain the optimal spatial and temporal

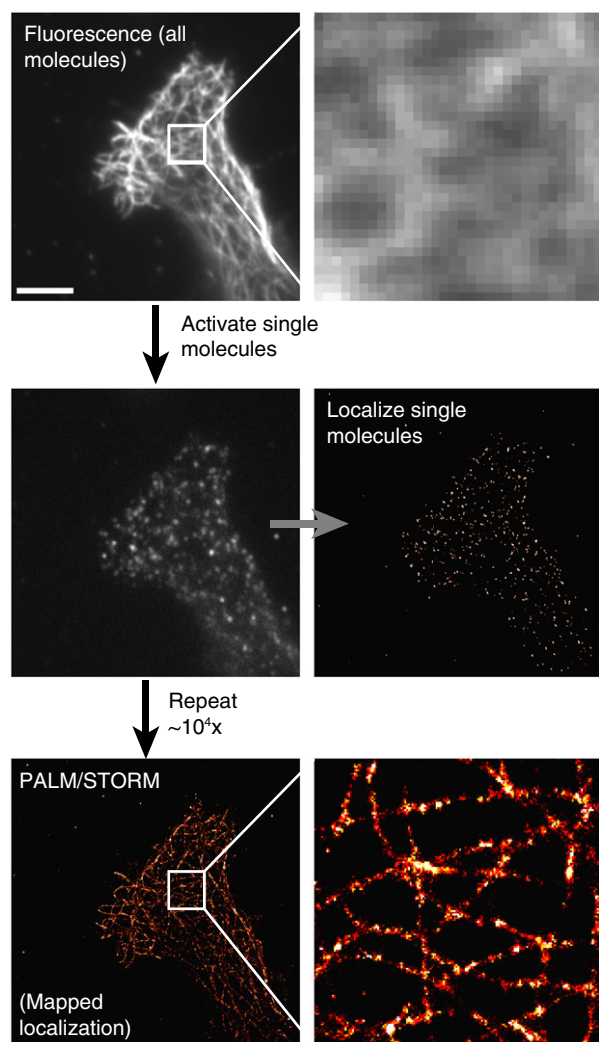


Figure 7 Super-resolution imaging via photo-activation localization microscopy (PALM)/stochastic optical reconstruction microscopy (STORM). Here, a HEK-293 cell has fluorescent labeling of its microtubule cytoskeleton. Individual molecules are induced to fluoresce, and their positions are calculated and mapped. The process is repeated $\sim 10^4$ times, and the result of these 10^5 – 10^6 mapped molecular positions is a super-resolution image. The increased resolution is illustrated in the insets, in which individual microtubules are resolved in the PALM/STORM images. Scale bar = 5 μm .

resolution. Unfortunately, these three parameters offset one another (**Figure 8**); for example, obtaining the maximum signal-to-noise ratio requires either a longer exposure time (reduced temporal resolution) or collection of signal from a greater focal volume (reduced spatial resolution). In practice, we have to empirically determine the best compromise between these three parameters for each experiment. This is also where the choice of microscope system (if available) becomes important.

Most academic research environments have access to a confocal laser-scanning microscope, which is the most widely used microscope for imaging single cells. The term ‘confocal’ refers to the use of a pinhole, which effectively rejects out of focus light and is able to achieve spatial resolution

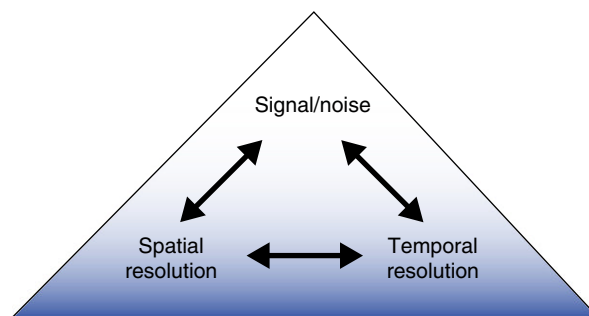


Figure 8 The competing parameters for optimal imaging. Increasing signal (more light collection) has to be balanced against temporal resolution (faster acquisition) and spatial resolution.

approaching Abbe’s limit. Modern systems are equipped with sensitive photomultiplier detectors and can easily achieve the maximum possible resolution with fluorescent protein probes at sufficient expression levels. However, if we refer back to **Figure 8** we can see that by rejecting out of focus light (improved spatial resolution), we must either sacrifice signal or else spend more time collecting light, reducing temporal resolution. Signal-to-noise or speed can be improved by opening the pinhole, which will lead to the collection of more light, with the offset that some axial resolution will be lost. Practically though, even with the brightest sample the physical speed with which the laser can be scanned across the specimen limits temporal resolution on these systems. Recently, resonant scanners have been incorporated into laser-scanning confocal microscopes, which greatly reduce scan-speed as a limit for imaging. Alternatively, spinning disk confocal microscopes sweep the entire field through an array of confocal pinholes mounted on the eponymous disk, sacrificing some axial resolution but with greatly enhanced sensitivity and speed of acquisition. A similar effect is achieved by sweeping a linear array of pinholes or a light sheet across the sample in swept field confocal microscopes. A special form of ‘confocal’ imaging can be performed by total internal reflection fluorescence (TIRF) microscopy ([Mattheyses et al., 2010](#)). In this case the illuminating laser light is directed to the specimen at an angle that yields an excitation only in a very thin layer (~ 100 nm) above the cover glass, resulting in fluorescence of the molecules that are in the plane of the cell attached to the cover slip. The advantage of this technique is its high signal to noise ratio to monitor processes at the ventral plasma membrane.

It should be noted that the popular and versatile laser-scanning confocal microscopes have the capacity to direct the diffraction-limited laser beam to anywhere in the sample. This makes FRAP and photoactivation-based experiments easy to perform on these systems. Furthermore, modern systems have such sensitive high-speed detectors that they are often capable of more advanced spectroscopic tasks, such as FCS/FCCS and FLIM.

An often cheaper but extremely sensitive option is to use a wide-field fluorescence microscope, where excitation light is focused onto the whole field of view, with no rejection of out-of-focus light. Despite the much poorer axial and lateral resolution compared to confocal microscopy, wide-field

has greatly improved sensitivity and temporal resolution (modern electron-multiplied charge-coupled device and scientific grade complimentary-metal-oxide-semiconductor cameras can achieve speeds exceeding 50 frames per second). This can make wide-field the choice for exceptionally dim and poorly expressing samples, or for many biosensors where sensitivity is the crucial parameter. Many FRET-based probes can exhibit fluorescence changes in the 5–10% range and the greatly increased sensitivity of wide-field systems produces much higher quality data.

Regardless of the microscope choice, the objective lens should be the best available. The higher the numerical aperture (NA), the more light will be collected, improving all three parameters listed in [Figure 8](#) – but with the caveat that the field of view and working distance will reduce as NA increases. The best NA objective available that can accommodate the sample should always be used!

Furthermore, the specimen must stay healthy, fluorescent and in focus for the duration of the experiment. High intensity light, especially laser light, can induce phototoxicity in cells. This becomes more severe the more blue-shifted (and hence, higher energy) the excitation light. Furthermore, repeated and/or high intensity illumination will cause fluorescent proteins to photobleach. Therefore, a balance has to be struck between maintaining the sample, and illuminating the sample enough to excite sufficient fluorescence. Again, compromise is necessary but as a rule, the minimum excitation intensity that produces satisfactory results should be used. Keeping cells healthy on the microscope stage involves maintaining them in a vessel bathed with appropriate medium. For experiments lasting a few minutes, balanced salt solutions are sufficient. For longer time periods, nutrient medium lacking phenol red (which is fluorescent and will occlude signal from the probe) can be used. The longer the incubation, the greater the care required with growth factors and serum, temperature and atmospheric conditions, and minimizing phototoxicity ([Coutu and Schroeder, 2013](#)). Maintaining focus can also be problematic with longer-term time-lapse imaging. Ensuring that the temperature of the microscope is equilibrated and the room is not subject to disruptive airflow from air conditioners or thoroughfare can help here. Conveniently, many modern microscopes are equipped with focal-feedback devices, which constantly measure the lens-specimen distance and automatically provide compensating adjustments.

When imaging in whole organisms, the above considerations are just as relevant; however, a whole host of additional technical issues must be taken into account. First, achieving optical access to the cells or structures of interest is often far from trivial. Precisely how this is best achieved depends on both the organism and cell type of interest. As mentioned above, LSM enables imaging of relatively transparent organisms without further intervention. In postnatal mammals, however, careful surgical exposure of the area to be imaged is typically required. For example, optical access to the rodent brain can be achieved either by thinning a small area of the skull to render it sufficiently transparent to image through ([Grutzendler et al., 2002](#)), or by replacing a small piece of the skull with a glass coverslip ([Trachtenberg et al., 2002](#)). Both procedures require extreme care and considerable technical expertise to avoid causing damage or inflammation, which could confound

experimental results. Similarly, dorsal skin-fold chambers provide optical access to implanted tumors ([Brown et al., 2001](#)). Where chronic imaging is to be performed over multiple days or weeks, animals must be carefully monitored for signs of infection and/or mechanical damage to any implanted imaging chamber. Furthermore, care must be taken to identify anatomical landmarks, such as surface blood vessel patterns, to enable the same cells to be relocated during subsequent imaging sessions. Where imaging in the live organism is not necessary, optical clearing techniques can provide an alternative approach. These chemical treatments reduce light scattering, usually by altering the refractive index of the specimen. Traditionally, clearing techniques have suffered from fluorescence quenching of fluorescent proteins and/or sample shrinkage or expansion, but newly developed methods are now circumventing these problems ([Chung et al., 2013](#); [Ke et al., 2013](#)).

Second, as mentioned above, mechanical stability is an absolute requirement for acquisition of high resolution and/or time-lapse images. In the living organism, motion artifacts from heartbeat, breathing and even pulsation of individual blood vessels can be problematic. In some cases these issues can be circumvented by mechanical stabilization, such as by head-fixation of rodents. In others, such as for soft tissue imaging, as is often the case in immunology or cancer biology studies, or where the experimental design necessitates the animal to move or perform behavioral tasks, it may be necessary to synchronize image acquisition to the heartbeat or respiration, and/or perform subsequent offline image correction. Covering or embedding the tissue of interest with agar can also help here; indeed, entire embryos are agar-embedded for long-term LSM.

Finally, great care must be taken to maintain the experimental subject in an optimal state, for both ethical and experimental reasons. Anesthetic agents must be carefully selected on the basis of their administration route, mode, and duration of action, and recoverability. Furthermore, a stable level of anesthesia must be maintained throughout the imaging session, and analgesics should be administered as required. Maintenance of body temperature and hydration are other key factors, especially for long-duration or repeated imaging sessions, and careful monitoring of physiological parameters may be necessary to produce optimal *in vivo* images.

Looking Forward

The use of genetically encoded fluorescent probes has come a very long way from the initial description of GFP in the 1970s. Even brighter fluorescent proteins are being developed with emission characteristics spanning the entire visible spectrum and beyond. These are being combined with ever more ingenious means of building them into biosensors for real-time, high-resolution reporting of molecular interactions and cellular activity. Combined with advances in the fields of molecular biology and optical physics, our capacity to image the molecular details of life inside cells is constantly improving. For an increasing number of problems in cell biology, the approach of simply watching how a process is driven by its protein components is now tenable. Whilst we anticipate that continued improvement along the lines discussed in this article will doubtless continue, we urge the reader to bear in

mind the relative obscurity of the crystal jelly that ultimately led to this explosion of progress in cell biology; in light of this, it would seem unwise to predict from where the next serendipitous technical revolution will emerge!

See also: Imaging the Cell: Light Microscopy: Fluorescence Correlation Spectroscopy: A Tool for Measuring Dynamic and Equilibrium Properties of Molecules in Cells; Fluorescence Lifetime Imaging — Applications and Instrumental Principles; High-Speed Localization Microscopy and Single-Particle Tracking; Intravital Microscopy in Mammalian Organisms: From Tissue Physiology to Cell Biology; Structured Illumination Microscopy; Super Resolution Fluorescence Localization Microscopy; Total Internal Reflection Fluorescence Microscopy

References

- Akerboom, J., *et al.*, 2009. Crystal structures of the GCaMP calcium sensor reveal the mechanism of fluorescence signal change and aid rational design. *Journal of Biological Chemistry* 284 (10), 6455–6464.
- Axelrod, D., *et al.*, 1976. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophysical Journal* 16 (9), 1055–1069.
- Bacia, K., Kim, S.A., Schwille, P., 2006. Fluorescence cross-correlation spectroscopy in living cells. *Nature Methods* 3 (2), 83–89.
- Barretto, R.P.J., Schnitzer, M.J., 2012. *In Vivo* microendoscopy of the hippocampus. *Cold Spring Harbor Protocols* 2012 (10), pdb.prot071472–pdb.prot071472.
- Berning, S., *et al.*, 2012. Nanoscopy in a living mouse brain. *Science* (New York, NY) 335 (6068), 551.
- Betzig, E., *et al.*, 2006. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* (New York, NY) 313 (5793), 1642–1645.
- Brown, E.B., *et al.*, 2001. *In vivo* measurement of gene expression, angiogenesis and physiological function in tumors using multiphoton laser scanning microscopy. *Nature Medicine* 7 (7), 864–868.
- Chalfie, M., *et al.*, 1994. Green fluorescent protein as a marker for gene expression. *Science* (New York, NY) 263 (5148), 802–805.
- Chen, T.-W., *et al.*, 2013. Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* 499 (7458), 295–300.
- Chung, K., *et al.*, 2013. Structural and molecular interrogation of intact biological systems. *Nature* 497 (7449), 332–337.
- Coutu, D.L., Schroeder, T., 2013. Probing cellular processes by long-term live imaging — historic problems and current solutions. *Journal of Cell Science* 126 (17), 3805–3815.
- Denk, W., Strickler, J.H., Webb, W.W., 1990. Two-photon laser scanning fluorescence microscopy. *Science* (New York, NY) 248 (4951), 73–76.
- Dobbie, I.M., *et al.*, 2011. OMX: A new platform for multimodal, multichannel wide-field imaging. *Cold Spring Harbor Protocols* 2011 (8), 899–909.
- Dodt, H.-U., *et al.*, 2007. Ultramicroscopy: Three-dimensional visualization of neuronal networks in the whole mouse brain. *Nature Methods* 4 (4), 331–336.
- Gao, L., *et al.*, 2012. Noninvasive Imaging beyond the diffraction limit of 3D dynamics in thickly fluorescent specimens. *Cell* 151 (6), 1370–1385.
- Goeppert-Mayer, M., 1931. Über Elementarakte mit zwei Quantensprüngen. *Annals of Physics* 9 (3), 273–295.
- Grutzendler, J., Kasthuri, N., Gan, W.-B., 2002. Long-term dendritic spine stability in the adult cortex. *Nature* 420 (6917), 812–816.
- Hardin, J., 2006. Confocal and multi-photon imaging of living embryos. In: Pawley, J. (Ed.), *Handbook of Biological Confocal Microscopy*. New York, NY: Plenum, pp. 746–768. *Cold Spring Harbor Protocols*.
- Heim, R., Tsien, R.Y., 1996. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Current Biology* 6 (2), 178–182.
- Heim, R., Cubitt, A.B., Tsien, R.Y., 1995. Improved green fluorescence. *Nature* 373 (6516), 663–664. doi:10.1038/373663b0.
- Hell, S., 2008. Microscopy and its focal switch. *Nature Methods* 6 (1), 24.
- Hell, S., Hein, B., Willig, K.I., 2008. Stimulated emission depletion (STED) nanoscopy of a fluorescent protein-labeled organelle inside a living cell. *Proceedings of the National Academy of Sciences of the United States of America* 105 (38), 14271–14276.
- Holekamp, T.F., Turaga, D., Holy, T.E., 2008. Fast three-dimensional fluorescence imaging of activity in neural populations by objective-coupled planar illumination microscopy. *Neuron* 57 (5), 661–672.
- Holtmaat, A., Svoboda, K., 2009. Experience-dependent structural synaptic plasticity in the mammalian brain. *Nature Reviews Neuroscience* 10 (9), 647–658.
- Huisken, J., *et al.*, 2004. Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science* (New York, NY) 305 (5686), 1007–1009.
- Jin, L., *et al.*, 2012. Single action potentials and subthreshold electrical events imaged in neurons with a fluorescent protein voltage probe. *Neuron* 75 (5), 779–785.
- Ke, M.-T., Fujimoto, S., Imai, T., 2013. SeeDB: A simple and morphology-preserving optical clearing agent for neuronal circuit reconstruction. *Nature Neuroscience. Technical Reports* 16 (8), 1154–1161.
- Keller, P.J., Dodt, H.-U., 2011. Light sheet microscopy of living or cleared specimens. *Current Opinion in Neurobiology* 22 (1), 138–143.
- Keller, P.J., *et al.*, 2008. Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science* (New York, NY) 322 (5904), 1065–1069.
- Kralj, J.M., *et al.*, 2012. Optical recording of action potentials in mammalian neurons using a microbial rhodopsin. *Nature Methods* 9 (1), 90–95.
- Kremers, G.-J., *et al.*, 2011. Fluorescent proteins at a glance. *Journal of Cell Science* 124 (2), 157.
- Stryer, L., Haugland, R.P., 1967. Energy transfer: A spectroscopic ruler. *Proceedings of the National Academy of Sciences of the United States of America* 58 (2), 719.
- Lam, A.J., *et al.*, 2012. Improving FRET dynamic range with bright green and red fluorescent proteins. *Nature Methods* 9 (10), 1005–1012.
- Lendvai, B., *et al.*, 2000. Experience-dependent plasticity of dendritic spines in the developing rat barrel cortex *in vivo*. *Nature* 404 (6780), 876–881.
- Manley, S., *et al.*, 2008. High-density mapping of single-molecule trajectories with photoactivated localization microscopy. *Nature Methods* 5 (2), 155–157.
- Manley, S., Gunzenhäuser, J., Olivier, N., 2011. A starter kit for point-localization super-resolution imaging. *Current Opinion in Chemical Biology* 15 (6), 813–821.
- Mashanov, G.I., *et al.*, 2004. The spatial and temporal dynamics of pleckstrin homology domain binding at the plasma membrane measured by imaging single molecules in live mouse myoblasts. *Journal of Biological Chemistry* 279 (15), 15274–15280.
- Matthyses, A.L., Simon, S.M., Rappoport, J.Z., 2010. Imaging with total internal reflection fluorescence microscopy for the cell biologist. *Journal of Cell Science* 123 (Pt 21), 3621–3628.
- Matz, M.V., *et al.*, 1999. Fluorescent proteins from nonbioluminescent Anthozoa species. *Nature Biotechnology* 17 (10), 969–973.
- McKinney, S.A., *et al.*, 2009. A bright and photostable photoconvertible fluorescent protein. *Nature Methods* 6 (2), 131–133.
- Meyer, T., Stauffer, T.P., Ahn, S., 1998. Receptor-induced transient reduction in plasma membrane PtdIns(4,5)P₂ concentration monitored in living cells. *Current Biology: CB* 8 (6), 343–346.
- Miesenböck, G., De Angelis, D.A., Rothman, J.E., 1998. Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins. *Nature* 394 (6689), 192–195.
- Miyawaki, A., 2011. Proteins on the move: Insights gained from fluorescent protein technologies. *Nature Reviews Molecular Cell Biology* 12 (10), 656–668.
- Miyawaki, A., *et al.*, 1997. Fluorescent indicators for Ca²⁺ based on green fluorescent proteins and calmodulin. *Nature* 388 (6645), 882–887.
- Morise, H., *et al.*, 1974. Intermolecular energy transfer in the bioluminescent system of *Aequorea*. *Biochemistry* 13 (12), 2656–2662.
- Nakai, J., Ohkura, M., Imoto, K., 2001. A high signal-to-noise Ca²⁺ probe composed of a single green fluorescent protein. *Nature Biotechnology* 19 (2), 137–141.
- Patterson, G., Lippincott-Schwartz, J., 2002. A photoactivatable GFP for selective photolabeling of proteins and cells. *Science* (New York, NY) 297 (5588), 1873–1877.
- Patterson, G., *et al.*, 2008. Transport through the Golgi apparatus by rapid partitioning within a two-phase membrane system. *Cell* 133 (6), 1055–1067.
- Pittet, M.J., Weissleder, R., 2011. Intravital imaging. *Cell* 147 (5), 983–991.
- Ponsioen, B., *et al.*, 2004. Detecting cAMP-induced Epac activation by fluorescence resonance energy transfer: Epac as a novel cAMP indicator. *EMBO Reports* 5 (12), 1176–1180.
- Schermelleh, L., *et al.*, 2008. Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* (New York, NY) 320 (5881), 1332–1336.
- Shaner, N.C., *et al.*, 2004. Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nature Biotechnology* 22 (12), 1567–1572.

- Shcherbakova, D.M., Verkhusha, V.V., 2013. Near-infrared fluorescent proteins for multicolor *in vivo* imaging. *Nature Methods* 10 (8), 751–754.
- Siegel, M.S., Isacoff, E.Y., 1997. A genetically encoded optical probe of membrane voltage. *Neuron* 19 (4), 735–741.
- Stadler, C., *et al.*, 2013. Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells. *Nature Methods* 10 (4), 315–323.
- Stettler, D.D., *et al.*, 2006. Axons and synaptic boutons are highly dynamic in adult visual cortex. *Neuron* 49 (6), 877–887.
- Sun, Y., *et al.*, 2012. Monitoring protein interactions in living cells with fluorescence lifetime imaging microscopy. *Methods in Enzymology* 504, 371–391.
- Tanimura, A., *et al.*, 2009. Use of fluorescence resonance energy transfer-based biosensors for the quantitative analysis of inositol 1,4,5-trisphosphate dynamics in calcium oscillations. *Journal of Biological Chemistry* 284 (13), 8910–8917.
- Teruel, M.N., Meyer, T., 2000. Translocation and reversible localization of signaling proteins: A dynamic future for signal transduction. *Cell* 103 (2), 181–184.
- Trachtenberg, J.T., *et al.*, 2002. Long-term *in vivo* imaging of experience-dependent synaptic plasticity in adult cortex. *Nature* 420 (6917), 788–794.
- Tsutsui, H., *et al.*, 2013. Improved detection of electrical activity with a voltage probe based on a voltage-sensing phosphatase. *Journal of Physiology* 591 (Pt 18), 4427–4437.
- Várnai, P., Balla, T., 1998. Visualization of phosphoinositides that bind pleckstrin homology domains: Calcium- and agonist-induced dynamic changes and relationship to myo-[3H]inositol-labeled phosphoinositide pools. *Journal of Cell Biology* 143 (2), 501–510.
- Voie, A.H., Burns, D.H., Spelman, F.A., 1993. Orthogonal-plane fluorescence optical sectioning: Three-dimensional imaging of macroscopic biological specimens. *Journal of Microscopy* 170, 229–236.
- Wang, J.W., *et al.*, 2003. Two-photon calcium imaging reveals an odor-evoked map of activity in the fly brain. *Cell* 112 (2), 271–282.
- Wiedenmann, J., *et al.*, 2004. EosFP, a fluorescent marker protein with UV-inducible green-to-red fluorescence conversion. *Proceedings of the National Academy of Sciences of the United States of America* 101 (45), 15905–15910.
- Wyckoff, J.B., *et al.*, 2000. A critical step in metastasis: *In vivo* analysis of intravasation at the primary tumor. *Cancer Research* 60 (9), 2504–2511.
- Zhang, G., Gurtu, V., Kain, S.R., 1996. An enhanced green fluorescent protein allows sensitive detection of gene transfer in mammalian cells. *Biochemical and Biophysical Research Communications* 227 (3), 707–711.

Total Internal Reflection Fluorescence Microscopy

D Axelrod, University of Michigan, Ann Arbor, MI, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Total internal reflection fluorescence (TIRF) provides a means to selectively excite fluorophores in an aqueous environment very near a solid surface (within ≤ 100 nm). When adapted to microscopy, TIRF produces wide-field images with very low background fluorescence from out-of-focus planes. The surface-selective illumination and the unique polarization features of TIRF, often in simultaneous combination with other fluorescence techniques, have enabled numerous applications in cell biology. This article discusses the principles of TIRF microscopy and various methods (both commercial and homemade) to implement it. The applications discussed here are to biological cells, not to artificial membranes or to single molecules, about which there is also an extensive literature.

For a brief history of the technique and a literature review, see Axelrod (2008).

Theoretical Basis

When a light beam (see Figure 1(a)) propagating through a transparent medium 3 of high index of refraction (e.g., glass) encounters a planar interface with medium 1 of lower index of refraction (e.g., water), it undergoes total internal reflection for incidence angles θ (measured from the normal to the interface) greater than the 'critical angle' θ_c given by:

$$\theta_c = \sin^{-1}(n_1/n_3) \quad [1]$$

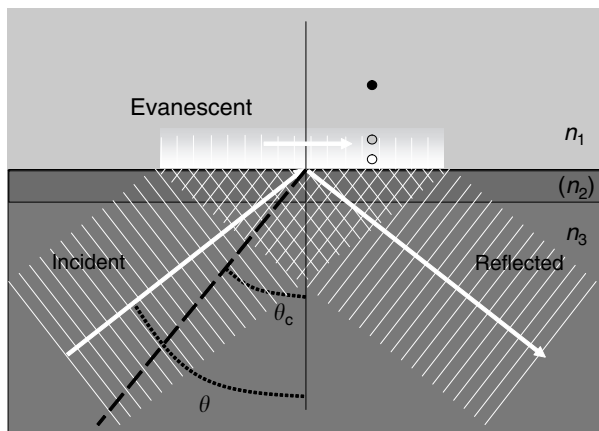


Figure 1 A TIR illumination. Refractive index n_3 must be greater than n_1 . The intermediate layer (consisting of metal or a dielectric material of refractive index n_2) is not necessary for TIR to occur, but can introduce some useful optical properties as explained in Axelrod (2013). Most applications of TIRF do not use an intermediate layer. The incidence angle θ must be larger than the critical angle θ_c for TIR to occur. The exponentially decaying evanescent field in the n_1 material is used to excite fluorophores in TIRF. The wavelength of the incident and reflected beams in the n_3 medium is λ_o/n_3 .

where n_1 and n_3 are the refractive indices of the liquid and the solid respectively (n_2 is reserved for intermediate layers, discussed in Axelrod, 2008). Ratio n_1/n_3 must be less than unity for TIR to occur. For 'subcritical' incidence angles $\theta < \theta_c$, most of the light propagates through the interface into the lower index material with a refraction angle (also measured from the normal) given by Snell's Law. (Some of the incident light also internally reflects back into the solid.) But for 'supercritical' incidence angles $\theta > \theta_c$, all of the light reflects back into the solid. Even in this case, some of the incident light energy penetrates through the interface and propagates parallel to the surface in the plane of incidence. The field in the liquid (sometimes called the 'evanescent wave') is capable of exciting fluorescent molecules that might be present near the surface.

The intensity I of the evanescent field at any position (measured as perpendicular distance z from the TIR interface) exponentially decays with z :

$$I(z) = I(0)e^{-z/d} \quad [2]$$

where,

$$\begin{aligned} d &= \frac{\lambda_o}{4\pi} (n_3^2 \sin^2 \theta - n_1^2)^{-1/2} \\ &= \frac{\lambda_o}{4\pi n_3} (\sin^2 \theta - \sin^2 \theta_c)^{-1/2} \end{aligned} \quad [3]$$

Parameter λ_o is the wavelength of the incident light in vacuum. Depth d is independent of the polarization of the incident light and decreases with increasing θ . Except for supercritical $\theta \rightarrow \theta_c$ (where $d \rightarrow \infty$), d is generally the order of λ_o or smaller.

Although the emission from a fluorophore excited by an evanescent field as viewed by a microscopic objective might be expected to follow an exponential decay with z according to eqn [2], this is not precisely true. Fluorescence emission near a dielectric interface (regardless of the means of excitation) is rather anisotropic and the degree of anisotropy is itself z -dependent (Hellen and Axelrod, 1987; Mertz, 2000; Axelrod, 2008, 2012). The fraction of the emission that can be captured by a fixed numerical aperture (NA) objective is therefore slightly z -dependent.

Applications and Combinations in Cell Biology

Cell/Substrate Contact Regions

TIRF can be used qualitatively to observe the position, extent, composition, and motion of contact regions even in samples in which fluorescence elsewhere would otherwise obscure the fluorescent pattern (Axelrod, 1981; Weis *et al.*, 1982; Joos *et al.*, 2006). The specific biochemistry of focal adhesions, using fluorescent proteins, has been studied with TIRF (Cohen *et al.*, 2006; Partridge and Marcantonio, 2006).

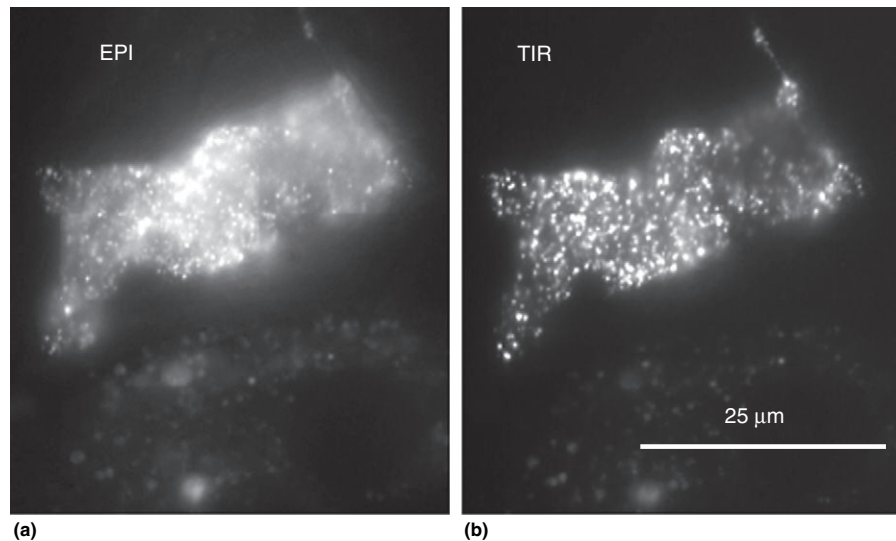


Figure 2 EPI vs. objective-based TIR fluorescence digital images, excited with an argon laser beam of wavelength 488 nm entering the side illumination port of an Olympus IX-70 microscope and viewed through an Olympus 1.45 NA 60X objective. This bovine chromaffin cells contains secretory granules marked with GFP-atrial natriuretic protein.

A variation of TIRF to identify cell–substrate contacts involves doping the solution surrounding the cells with a non-adsorbing and nonpermeable fluorescent volume marker; focal contacts then appear relatively dark (Gingell *et al.*, 1987; Todd *et al.*, 1988).

Secretory Granule Tracking and Exocytosis

Secretory granules often reside through the whole depth of a cell. When viewed with standard EPI-illumination, fluorescence-marked secretory granules are difficult to distinguish individually (Figure 2(a)). When viewed with TIRF, the granules closest to the membrane appear very distinctly (Figure 2(b)). The thin evanescent field allows small motions of individual fluorescence-marked secretory granules to manifest as small intensity changes arising from small motions in the direction normal to the substrate and plasma membrane (the ‘axial’ or ‘z-direction’). The precision of such axial tracking can be as small as 2 nm, much smaller than the light microscope resolution limit. Because of the high contrast and low background, the position of the centers of granules can be measured with an accuracy down to about 10 nm and motions much smaller than the granule diameter can be followed before and after exocytosis-inducing chemical stimulation (Allersma *et al.*, 2004, 2006). In some cases, dispersal of granule contents can be observed and interpreted as exocytosis and events immediately preceding and following exocytosis can be characterized (see Burchfield *et al.*, 2010 for a review; Weiss *et al.*, 2014a).

Most of the TIRF studies on the detailed mechanism of secretion have been on animal cells. However, the same techniques can be applied to plant cells, and in particular, have been applied to study secretion from pollen tubes (Wang *et al.*, 2006).

Cytoplasmic Filaments

Although TIRF cannot view deeply into thick cells, it can display with high contrast the fluorescence-marked sub-membrane filament structure at the substrate contact regions. Actin polymerization and its regulation, both *in vitro* and *in vivo*, have been studied by TIRF (see Manneville, 2006 for a review), as have microtubules (Shaw *et al.*, 2007) and microtubule-associated kinesin (Seitz and Surrey, 2006). The manner by which extracellular components such as fibronectin are organized through the membrane in concert with sub-membrane filaments can be investigated by TIRF (Yoneda *et al.*, 2007). In combination with very low levels of specific filament labeling, which produces fluorescent speckles, TIRF illumination can visualize filament subunit turnover (Adams *et al.*, 2004).

Calcium Channels and Transients

TIRF is completely compatible with standard EPI-fluorescence, bright field, dark field, or phase contrast illumination. These methods of illumination can be switched back and forth rapidly with TIRF by electro-optic devices to compare membrane-proximal ionic transients and deeper cytoplasmic transients (Omann and Axelrod, 1996). TIRF can be used to visualize single Ca^{2+} channels with good spatial and temporal resolution (Demuro and Parker, 2004).

TIR Combined with Polarization: Plasma Membrane Deformation

The polarization (i.e., the vector direction of the electric field) of the evanescent wave depends on the incident light polarization, which can be either ‘p-pol’ (polarized in the plane of incidence formed by the incident and reflected rays, denoted

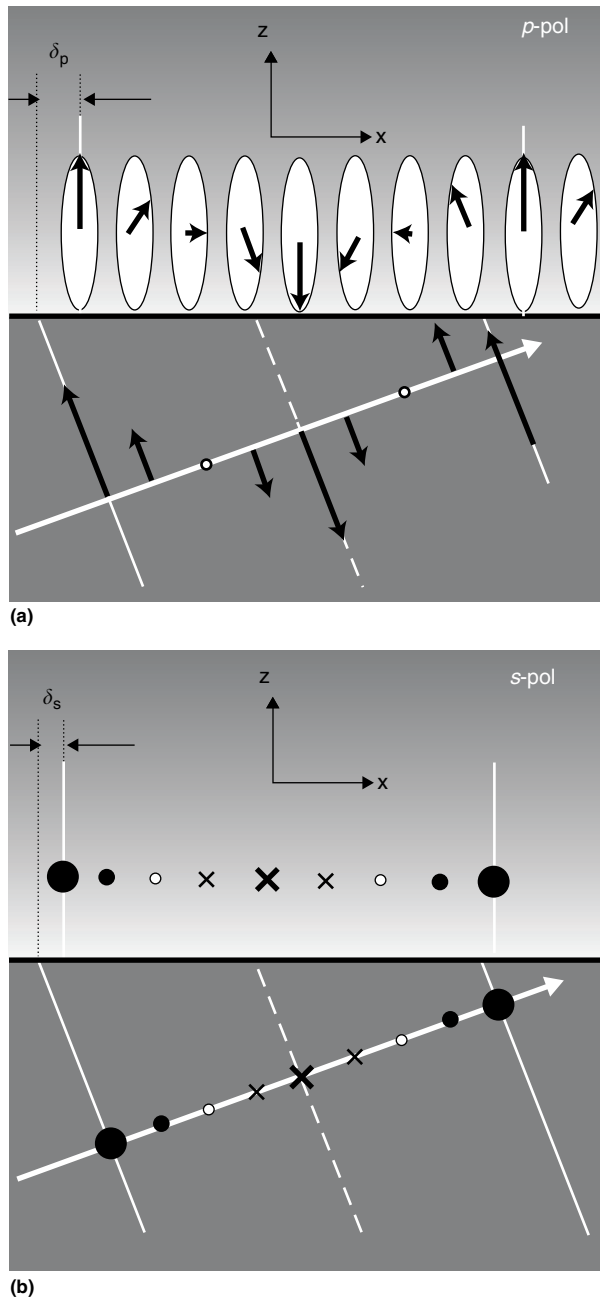


Figure 3 Polarization of the evanescent field. (a) P-pol incident field, giving rise to a p-pol evanescent field, with phase lag δ_p and coordinate system convention indicated. The p-pol evanescent field is elliptically polarized in the x-z plane as shown (primarily polarized in the z-direction with a weaker x-component at a relative phase of $\pi/2$). (b) S-pol incident field, giving rise to s-pol evanescent field, with phase lag δ_s indicated. The phase lags $\delta_{p,s}$ are not equal to each other.

here as the x-z plane) or 's-pol' (polarized normal to the plane of incidence, here the y-direction). In both polarizations, the evanescent wavefronts travel parallel to the surface in the x-direction.

For p-pol incident light (see [Figure 3\(a\)](#)), the evanescent electric field vector direction remains in the plane of incidence,

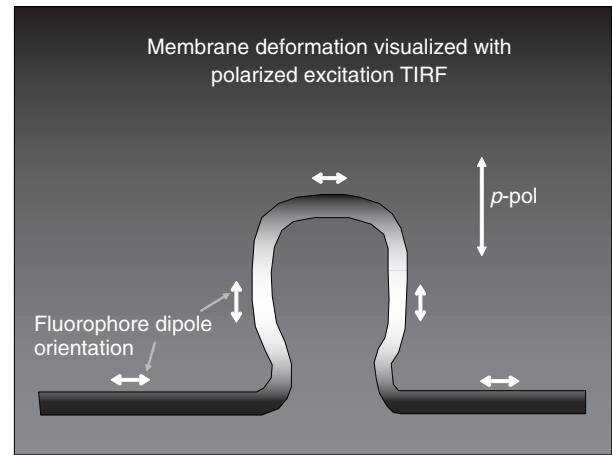


Figure 4 Polarized excitation TIRF at a membrane deformation. This is a schematic view of the excitation probability of oriented carbocyanine fluorophores embedded in a membrane in a z-polarized evanescent field (the dominant direction of a p-pol evanescent field). The membrane is depicted in cross section with a curved region corresponding to an inward deformation. The direction of absorption dipole of the fluorophores is shown with bidirectional arrows. Higher excitation probability is depicted by lighter shades. The z-component of the electric field selectively excites the regions of oblique membrane orientation.

mainly in the z-direction but with a small component in the x-direction depending on the incidence angle. For s-pol incident light (see [Figure 3\(b\)](#)), the evanescent electric field vector direction remains purely normal to the plane of incidence (the y-direction).

In standard normally incident EPI illumination, the incident light has no z-component. This unique z-direction found in p-pol TIR can be used to view submicroscopic deformations in the plasma membrane of living cells. For example, a membrane label might be chosen because its excitation dipole is always parallel to the membrane surface ([Axelrod, 1979](#)). By using excitation light that is polarized in the z-direction, perpendicular to the surface, only those regions of the plasma membrane that are not parallel to the substrate will be excited ([Figure 4](#)). To normalize against spatial-intensity variations that arise from actual place-to-place fluorophore-concentration differences, an 's-pol' TIR beam is also used to acquire an alternate image of each scene. The ratio of images taken with sequential p- and s-pol excitation is then calculated; the 'p/s' result reports only orientational effects and not local concentration effects.

With polarized excitation TIRF, along with a second label to mark the location of secretory granules, dynamic changes in membrane deformation can be tracked during the course of a secretion event ([Anantharam et al., 2010, 2012](#)).

TIR Combined with Fluorescence Recovery after Photobleaching and Fluorescence Resonance Energy Transfer (FRET)

TIR combined with fluorescence recovery after photobleaching (FRAP) can examine the specific or nonspecific binding kinetics of proteins to cellular membranes ([Hellen and Axelrod,](#)

1991; Stout and Axelrod, 1994; Sund and Axelrod, 2000). The post-bleach recovery can be quantitatively observed at each camera pixel to obtain a spatially-resolved 'image' of kinetic rates (Axelrod and Omann, 2006). TIR-FRAP has also been used to determine the diffusion coefficient of proteins inside single secretory granule in primary chromaffin cells (Weiss *et al.* (2014b)).

Interactions among channel or receptor proteins at surfaces can be studied by TIR/FRET (Khakh *et al.*, 2005).

TIR Combined with Interference Reflection Contrast

Interference reflection microscopy is a nonfluorescence method of visualizing the closeness of cell/substrate contact regions. By combining TIRF with interference reflection contrast (IRC), Ca^{2+} microdomains can be viewed simultaneously with exocytosis events in synapses (Beaumont *et al.*, 2005). Both TIR and IRC were used on the same sample to study the approach of neuronal growth cones to the substrate (Tatsumi *et al.*, 1999).

Optical Configurations and Setup

Many alternative optical arrangements for TIRF are available. Some configurations use a high NA ($\text{NA} > 1.4$) microscope objective for both TIR illumination and emission observation ('objective-based'), and others use a prism to direct the light toward the TIR interface with a separate objective for emission observation ('prism-based'). The evanescent illumination is not 'pure' with objective-based TIRF: a small fraction of the illumination of the sample results from excitation light scattered within the objective, and a small fraction of the observed fluorescence arises from luminescence of the objective's internal elements. Prism-based TIRF avoids these problems but geometrical constraints leave it most workable only for low- and medium-power objectives.

This section gives examples of these arrangements. We assume that the sample consists of fluorescence-labeled cells in culture adhered to a glass coverslip.

High Aperture Objective-Based TIR

By using an objective with a sufficiently high NA, supercritical angle incident light can be cast upon the sample by illumination through the objective (Stout and Axelrod, 1989; Axelrod, 2001). Although an arc lamp can be used as the light source, the general features are best explained with reference to a laser source.

The laser beam used for excitation is focused (by an external focusing lens) to a point at the back focal plane of the objective so that the light emerges from the objective in a collimated form (i.e., the 'rays' are parallel to each other). This insures that all the rays are incident upon the sample at the same angle θ with respect to the optical axis.

The point of focus in the back focal plane is adjusted to be off-axis. There is a one-to-one correspondence between the off-axis radial distance ρ and the angle θ . At a sufficiently high ρ , the critical angle for TIR can be exceeded. Further increases in ρ

serve to reduce the characteristic evanescent field depth d in a smooth and reproducible manner.

The NA of the objective must be greater than the refractive index of the sample to reach the critical angle and beyond. The 1.49NA objectives are probably the method of choice for TIR except for cells which have particularly dense organelles. Dense organelles tend to scatter the evanescent field, and this effect is less prominent with the higher angles of incidence accessible through higher aperture objectives.

The angle of convergence/divergence of the laser beam cone at the back focal plane is proportional to the diameter of the illuminated region subsequently created at the sample plane. Large angles (and consequent large illumination regions) can be produced by use of a beam expander placed just upstream from the focusing lens.

Custom-built TIRF systems generally use a free laser beam rather than fiber optics because this allows more accessible control of polarization and more rapid control of incidence direction with electrooptical devices. Figure 5 shows an example of such a system, suitable for both two-color TIRF and polarized excitation TIRF. Positioning of the laser beam focus at the desired off-axis location at the back focal plane is greatly facilitated by some fast electro-optical beam deflector system. For example, a pair of computer-controlled acousto-optical modulators or small galvanometer mirrors can adjust the focal point to any position in the two-dimensional back focal plane. Since altering the focus position at the back focal plane (to alter the incidence angle at the sample) should preferably leave the illumination region on the sample plane unmoved, it is desirable to place the deflector system at a plane in the optical system equivalent to the sample plane.

Commercial TIRF laser systems generally use enclosed fiber optics, need less expertise for setting up and maintaining alignment, and contain shut-off interlocks for laser safety. The tip of the optical fiber acts like a point source placed at an equivalent back focal plane (EBFP). The optical fiber tip itself is mounted on a micrometer translator to move from positions corresponding to subcritical incidence angles (EPI) to supercritical incidence angles (TIR). Figure 6 shows an example.

Commercial arc-lamp TIRF systems use an off-axis curved slit cutout at an EBFP. The slit is positioned off-axis to transmit only a thin-curved band of light at a radius that will produce supercritical angles. To optimize the throughput, the brightest part of the Hg arc is focused off-axis right upon the curved slit cutout, with a focusing aberration that curves the arc image to match well with the curved slit.

Prism-Based TIRF

Although a prism may restrict sample accessibility or choice of objectives in some cases, prism-based TIR is very inexpensive to set up and produces a 'cleaner' evanescent-excited fluorescence (i.e., less excitation light scattering in the optics) than objective-based TIR. Figure 7 shows several schematic drawings for setting up laser/prism-based TIR in an inverted or upright microscope.

A commercially available variation of prism-based TIR uses a high-aperture condenser fed with laser light from an optical fiber tip instead of a custom-installed prism fed by a free laser

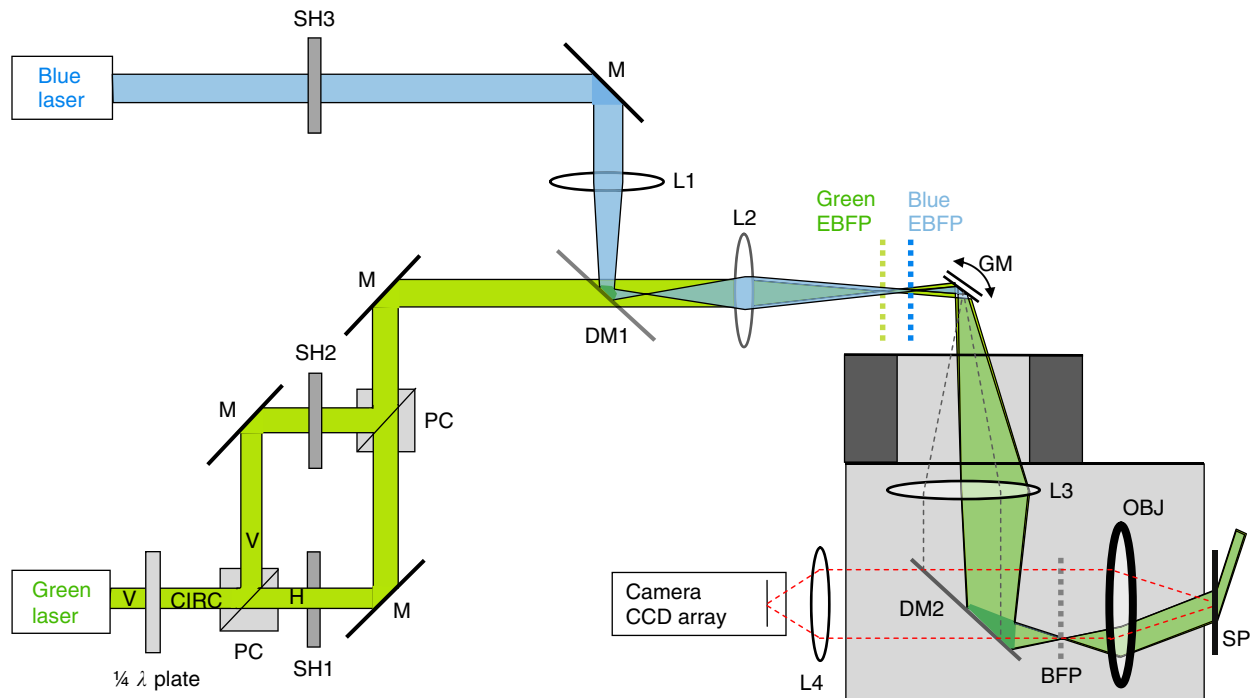


Figure 5 Optical ray diagram for custom-built TIRF with a free laser beam. This shows the basic setup and also optional add-ons for polarized excitation TIRF and for two excitation color TIRF. For the basic setup, the laser beam enters the microscope (lower right) through the rear hole in the microscope tower (dark gray) that was designed for arc illumination. For pictorial clarity, a collimated beam is shown here emerging from the sample plane at an oblique (but sub-TIR) angle, after having focused off axis at the objective (OBJ) BFP. Lens L3, placed in the small open space between the tower and the dichroic mirror (DM2) holder, ensures that the incoming beam focuses at the BFP with as large an angle of convergence as possible. (Large angles create a large illuminated region on the sample.) The off-axis focus position on the BFP (and the corresponding super- or subcritical incidence angle at the SP) is controlled by the angle of an electronically driven dual galvanometer mirror set (GM) positioned at an equivalent sample plane immediately behind the tower. Upbeam from the GM, the beam is focused by lens L2 to a point on an EBF that was formed there by the placement of lens L3. As usual for fluorescence microscopy, the sample plane is reimaged (see dashed red line) at the camera CCD array by lens L4 in the microscope/camera coupler. For the two-excitation color add-on, the second color is added via a DM1 and shuttered by SH3. The complication is that simple lenses L2 and L3 will have a chromatic aberration (shorter focal lengths) for blue vs green light. The appropriate EBF cast by lens L3 will thereby have slightly different positions for the two colors. Lens L1 placed upbeam in the blue beam is selected and positioned to ensure that both the colors are focused by L3 at the actual BFP. For the polarization control add-on option in the green beam, a quarter-wave plate converts the (presumably) vertical polarization emerging from the laser into circular polarization (CIRC). The first polarizing cube (PC1) divides the beam into its two polarization components, horizontal (H, straight ahead) and vertical (V, reflected 90°). Each is permitted to pass or be blocked by its own shutter (SH1 and SH2). The polarization paths are combined in a second polarizing cube (PC2).

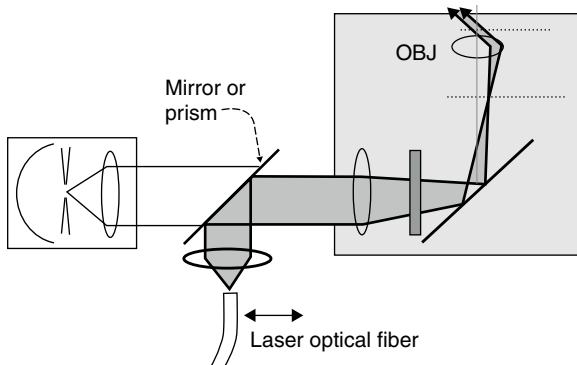


Figure 6 Laser illumination introduced by an optical fiber through the rear port normally used by the arc lamp. This scheme is employed by commercial TIRF systems.

beam. This system has the advantage of ready-made convenience and easy variation of both polar and azimuthal incidence angle adjustment, but is limited in maximum incidence angle.

In all the prism-based methods, the TIRF spot should be focused to a width no larger than the field of view. The larger the spot, the more the spurious scattering and out-of-focus fluorescence from the immersion oil layer between the prism and coverslip will increase the generally very low fluorescence background attainable by TIRF. Also, the incidence angle should exceed the critical angle by at least a couple of degrees. At incidence angles very near the critical angle, the cells cast a noticeable 'shadow' along the surface.

Additional Features

Suppression of Laser Interference Fringes

Laser illumination (whether TIRF or EPI) can produce annoying interference fringes because optical imperfections in the beam path can shift the local phase of coherent light. For critical applications, several approaches can alleviate the problem. One approach employs an optical fiber bundle in

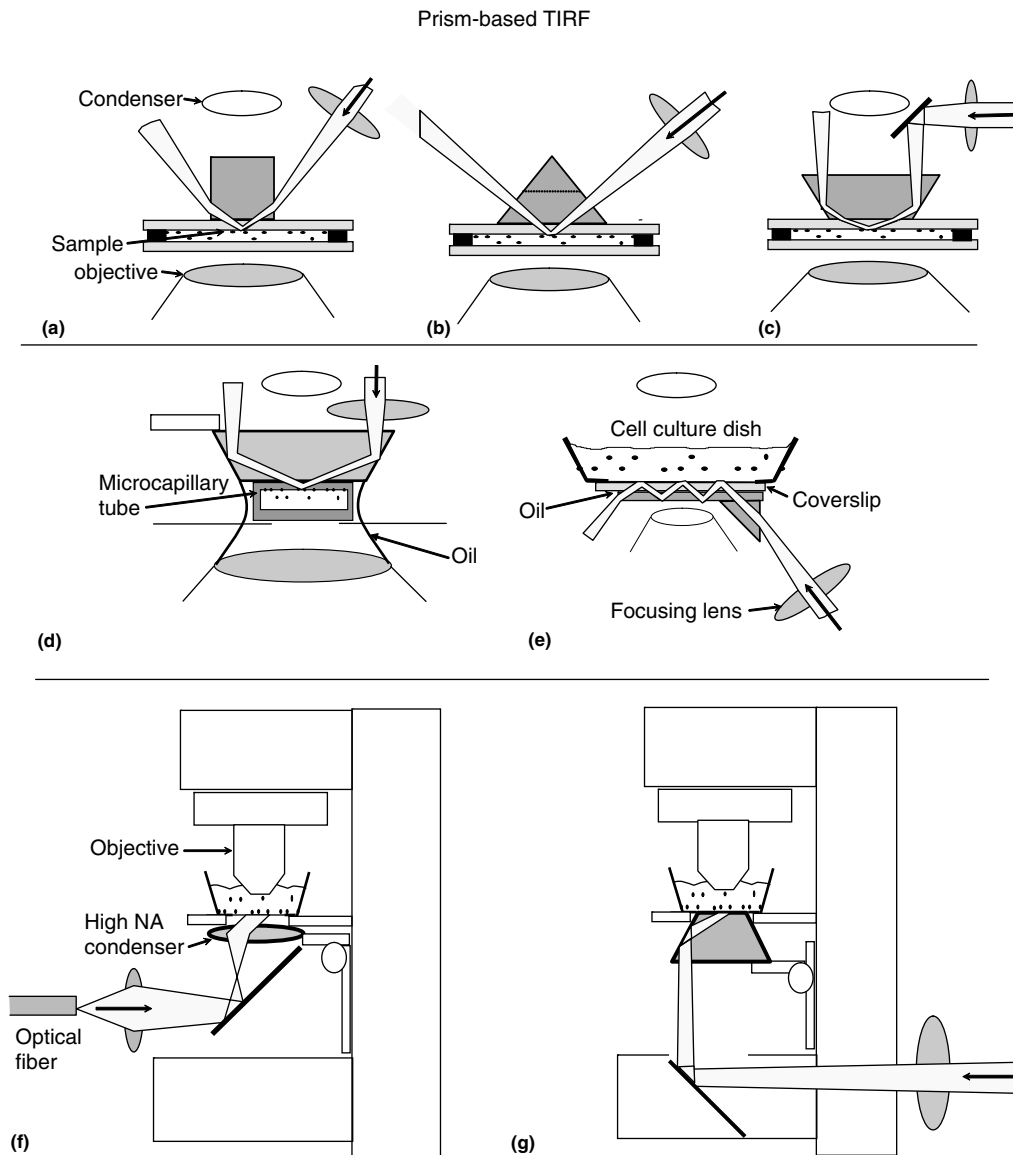


Figure 7 Schematic drawings for prism-based TIR, all using a laser as a light source. The vertical distances are exaggerated for clarity. The first four configurations (a)–(d) use a TIR prism above the sample. Configuration (e) places a small triangular prism below the sample off to the side and depends on multiple internal reflections in the substrate to bring the illumination into the field of view. Configurations (f) and (g) are for an upright microscope. In (f), a high-aperture condenser plays the role of an external prism. Configuration (g) utilizes the integral optics in the microscope base and a trapezoidal prism on the vertically movable condenser mount.

which individual fiber-to-fiber length differences exceed the laser light's coherence length. This produces an extended source at the output tip with point-to-point relative phases that are temporally randomized and defocused on the sample (Inoue *et al.*, 2001). This elegant system, with no moving parts, produces a speckle pattern that changes so rapidly that it appears uniform down to the nanosecond timescale. Commercially available mode scramblers and rapid flexing of optical fibers may also reduce some fringing.

Another set of methods is appropriate for a free laser beam rather than optical fiber system. For example, a finely frosted glass surface or a plastic plate (such as tissue culture dish bottom), spinning normally to the laser beam, temporally

randomizes the phase and produces a fringe pattern that fluctuates and can be averaged over the duration of a camera exposure (Kuhn and Pollard, 2005).

Interference fringes can be effectively suppressed by spinning the azimuthal angle of TIR incidence with electro-optical or mechano-optical devices such that the incident beam traces a circle where it focuses at the objective's back focal plane (BFP). Only one azimuthal direction illuminates the sample at any instant. However, the spinning period is much shorter than the camera's exposure time or retinal image retention time. Therefore, the interference fringes appear superimposed in their intensities, giving the illusion of a much more uniform illumination field (Mattheyses *et al.*, 2006).

Actual Evanescent Field Depth and Profile

The evanescent field characteristic depth in an actual sample may deviate from that predicted by eqn [3] for several reasons even if the incidence angle θ is well measured. Depth d depends on the refractive index of the liquid sample above the TIR surface, and this is not well known for samples such as local regions of cell cytoplasm. So far, no good way of accurately measuring the local evanescent field depth and profile in an actual cell has been implemented.

Measurements of evanescent depth in a homogeneous artificial sample with a refractive index that approximated the average in a heterogeneous sample can be done. In such a sample, the characteristic time for diffusion of fluorescent beads through the evanescent field (possibly by TIR/FRAP) can be measured and (given a known diffusion coefficient) converted to a corresponding characteristic distance. However, this method gives only a single number for effective depth but does not provide unambiguous information about the possibly non-exponential intensity profile as a function of z .

In samples with heterogeneous refractive index, regions with higher index may not support TIR, whereas lower index regions do. Clearly, it is best to illuminate with as high an incidence angle as possible to be sure that all regions support TIR. Even if the incidence angle is sufficiently high to support TIR everywhere, irregularities and heterogeneities in the sample give rise to scattering of the evanescent field. This scattered light can excite fluorescence as it penetrates much more deeply into the liquid than does the evanescent light. The effect is much less pronounced for higher TIR incidence angles.

The TIR optics itself, particularly the high aperture objectives used in objective-based illumination, can produce scattered light. Actual intensity profiles arising from objective-based TIR can be obtained by observing a large ($\sim 8\ \mu\text{m}$ diameter) spherical bead with a refractive index matched to the surrounding liquid to avoid disruption and scattering. The bead can be fluorescence-labeled on its surface. At the image of the point of contact between the bead and the TIR substrate, the fluorescence is the brightest. It becomes dimmer with increasing distance in the TIR plane from the contact point. Simple geometry then allows calculation of the fluorescence versus z . In an actual test of this method in 1.45NA and 1.65NA objective-based TIRF (Mattheyses and Axelrod, 2005), the evanescent field was found to account for about 90% of the intensity at $z=0$, and the exponential decay depth agreed well with expectations based on angle and refractive indices. However, a much slower decaying component, presumably from scattering within the objective, was also evident, and this component became dominant in the distant $z > 2d$ zone.

Apart from this dim scattering component, bright collimated shafts of scattered light, emanating from the edge of the top lens of the objective, can be seen in objective-based TIR. Fortunately, however, these shafts do not cross the field of view so they do not excite spurious fluorescence in the image.

Variable Incidence Angle TIR: Concentration Profiles

By decreasing the supercritical incidence angle back toward the critical angle, the depth of the evanescent field increases,

reaching out to more distal fluorophores. By measuring the fluorescence observed as a function of supercritical incidence angle, it is thereby possible to deduce what is the concentration profile $C(z)$ of fluorophores within the evanescent field, to a resolution much shorter than the evanescent field depth. This approach has been employed in practice to measure cell surface-to-substrate separation distances (Loerke *et al.*, 2000; Stock *et al.*, 2003) and the z -position of subcellular organelles (Loerke *et al.*, 2002).

Long-Term Fluorescence Movies of Cultured Cells

Since the cells are exposed to TIR excitation light only at their cell/substrate contact regions but not through their bulk, they tend to survive longer under observation, thereby enabling time-lapse recording of a week in duration. During this time, newly appearing cell surface receptors can be immediately marked by fluorescent ligand that is continually present in the full cell culture medium. Because TIRF is utilized, background fluorescence from this bath of unbound fluorophore is minimized (Wang and Axelrod, 1994).

TIRF versus Other Optical Sectioning Microscopies

TIRF is only one of several optical sectioning techniques enjoying wide use in microscopy, including confocal and multiphoton microscopies. Each has particular distinct advantages and drawbacks.

Confocal microscopy achieves axial selectivity by exclusion of out-of-focus emitted light with an image plane pinhole. Confocal microscopy has the clear advantage in versatility; its method of optical sectioning works at any plane of the sample, not just at an interface between dissimilar refractive indices. However, other differences exist which, in some special applications, can favor the use of TIRF.

- The depth of the optical section in TIRF is typically $\leq 0.1\ \mu\text{m}$, whereas in confocal microscopy it is relatively thick ($\sim 0.6\ \mu\text{m}$).
- In some applications (e.g., FRAP, fluorescence correlation spectroscopy, or on cells whose viability is damaged by light), illumination and not just detected emission is best restricted to a thin section; this is not possible with standard confocal.
- TIRF can be adapted to and made interchangeable with existing standard microscope optics, even with 'homemade' components. It is much less expensive than confocal microscopy, even for the commercial systems

Two-photon (or more generally, multiphoton) microscopy has many desirable features, including true optical sectioning, whereby the plane of interest is the only one that is actually excited, as in TIRF. Multiphoton microscopy is not restricted to the proximity of an interface, but its optical section depth is still several times thicker than that of TIRF. The setup expense of multiphoton microscopy for an infrared laser with sufficient pulse peak intensity can also be a consideration. Both multiphoton and confocal microscopy are necessarily scanning techniques; TIRF microscopy is a 'wide field' technique and is

thereby not limited in speed by the scanning system hardware or image reconstruction software.

See also: Imaging the Cell: Light Microscopy: Structured Illumination Microscopy; Super Resolution Fluorescence Localization Microscopy

References

- Adams, M.C., Matov, A., Yarar, D., *et al.*, 2004. Signal analysis of total internal reflection fluorescent speckle microscopy (TIR-FSM) and wide-field epi-fluorescence FSM of the actin cytoskeleton and focal adhesions in living cells. *Journal of Microscopy* 216, 138–152.
- Allersma, M.W., Bittner, M.A., Axelrod, D., Holz, R.W., 2006. Motion matters: Secretory granule motion adjacent to the plasma membrane and exocytosis. *Molecular Biology of the Cell* 17, 2424–2438.
- Allersma, M.W., Wang, L., Axelrod, D., Holz, R.W., 2004. Visualization of regulated exocytosis with a granule-membrane probe using total internal reflection microscopy. *Molecular Biology of the Cell* 15, 4658–4668.
- Anantharam, A., Axelrod, D., Holz, R.W., 2010. Polarized TIRFM reveals changes in plasma membrane topology before and during Granule Fusion. *Cellular and Molecular Neurobiology* 30, 1343–1349.
- Anantharam, A., Axelrod, D., Holz, R.W., 2012. Real-time investigation of fusion pore expansion by imaging of plasma membrane deformations. *Journal of Neurochemistry* 122, 661–671.
- Axelrod, D., 1979. Carbocyanine dye orientation in red cell membrane studied by microscopic fluorescence polarization. *Biophysical Journal* 26, 557–574.
- Axelrod, D., 1981. Cell surface contacts illuminated by total internal reflection fluorescence. *Journal of Cell Biology* 89, 141–145.
- Axelrod, D., 2001. Selective imaging of surface fluorescence with very high aperture microscope objectives. *Journal of Biomedical Optics* 6, 6–13.
- Axelrod, D., 2008. Total internal reflection fluorescence microscopy. In: Correia, J.J., Detrich, H.W. (Eds.), *Methods in Cell Biology: Biophysical Tools for Biologists 2: Methods In Vivo*. New York, NY: Academic Press, pp. 169–221.
- Axelrod, D., 2012. Fluorescence excitation and imaging of single molecules near dielectric-coated and bare surfaces: A theoretical study. *Journal of Microscopy* 247, 147–160.
- Axelrod, D., 2013. Evanescent excitation and emission in fluorescence microscopy. *Biophysical Journal* 104, 1401–1409. **Correction**, 104, 2321.
- Axelrod, D., Omann, G.M., 2006. Combinatorial microscopy. *Nature Reviews Molecular Cell Biology* 7 (12), 944–952.
- Baumont, V., Llobet, A., Lagnado, L., 2005. Expansion of calcium microdomains regulates fast exocytosis at a ribbon synapse. *Proceedings of the National Academy of Sciences of the United States of America* 102, 10700–10705.
- Burchfield, J.G., Lopez, J.A., Mele, K., Vallotton, P., Hughes, W.E., 2010. Exocytotic vesicle behaviour assessed by total internal reflection fluorescence microscopy. *Traffic* 11, 429–439.
- Cohen, M., Kam, Z., Addadi, L., Geiger, B., 2006. Dynamic study of the transition from hyaluronan- to integrin-mediated adhesion in chondrocytes. *EMBO Journal* 25, 302–311.
- Demuro, A., Parker, I., 2004. Imaging the activity and localization of single voltage-gated Ca²⁺ channels by total internal reflection fluorescence microscopy. *Biophysical Journal* 86, 3250–3259.
- Gingell, D., Heavens, O.S., Mellor, J.S., 1987. General electromagnetic theory of internal reflection fluorescence: The quantitative basis for mapping cell-substratum topography. *Journal of Cell Science* 87, 677–693.
- Hellen, E.H., Axelrod, D., 1987. Fluorescence emission at dielectric and metal-film interfaces. *Journal of the Optical Society of America B* 4, 337–350.
- Hellen, E., Axelrod, D., 1991. Kinetics of epidermal growth factor/receptor binding on cells measured by total internal reflection/fluorescence recovery after photobleaching. *Journal of Fluorescence* 1, 113–128.
- Inoue, S., Knudson, R.A., Goda, M., *et al.*, 2001. Centrifuge polarizing microscope. I. Rationale, design and instrument performance. *Journal of Microscopy* 201, 341–356.
- Joos, U., Biskup, T., Ernst, O., *et al.*, 2006. Investigation of cell adhesion to structured surfaces using total internal reflection fluorescence and confocal laser scanning microscopy. *European Journal of Cell Biology* 85, 225–228.
- Khakh, B.S., Fisher, J.A., Nashmi, R., Bowser, D.N., Lester, H.A., 2005. An angstrom scale interaction between plasma membrane ATP-gated P2X(2) and alpha(4)beta(2) nicotinic channels measured with fluorescence resonance energy transfer and total internal reflection fluorescence microscopy. *Journal of Neuroscience* 25, 6911–6920.
- Kuhn, J.R., Pollard, T.D., 2005. Real-time measurements of actin filament polymerization by total internal reflection fluorescence microscopy. *Biophysical Journal* 88, 1387–1402.
- Loerke, D., Preitz, B., Stuhmer, W., Oheim, M., 2000. Super-resolution measurements with evanescent-wave fluorescence excitation using variable beam incidence. *Journal of Biomedical Optics* 5, 23–30.
- Loerke, D., Stuhmer, W., Oheim, M., 2002. Quantifying axial secretory-granule motion with variable-angle evanescent-field excitation. *Journal of Neuroscience Methods* 119, 65–73.
- Manneville, J.B., 2006. Use of TIRF microscopy to visualize actin and microtubules in migrating cells. *Methods in Enzymology* 406, 520–532.
- Mattheyses, A., Axelrod, D., 2005. Direct measurement of evanescent field profile and depth in TIRF microscopy. *Journal of Biomedical Optics* 11, 014006/1–014006/7.
- Mattheyses, A., Shaw, K.D., Axelrod, D., 2006. Effective elimination of laser interference fringing in fluorescence microscopy by spinning azimuthal incidence angle. *Microscopy Research and Technique* 69, 642–647.
- Mertz, J., 2000. Radiative absorption, fluorescence, and scattering of a classical dipole near a lossless interface: A unified description. *Optics InfoBase: Journal of the Optical Society of America B* 17, 1906–1913.
- Omann, G.M., Axelrod, D., 1996. Membrane proximal calcium transients in stimulated neutrophils seen by total internal reflection fluorescence. *Biophysical Journal* 71, 2885–2891.
- Partridge, M.A., Marcantonio, E.E., 2006. Initiation of attachment and generation of mature focal adhesions by integrin-containing filopodia in cell spreading. *Molecular Biology of the Cell* 17, 4237–4248.
- Seitz, A., Surrey, T., 2006. Progressive movement of single kinesins on crowded microtubules visualized using quantum dots. *EMBO Journal* 25, 267–277.
- Shaw, R.M., Fay, A.J., Puthenveedu, M.A., *et al.*, 2007. Microtubule plus-end-tracking proteins target gap junctions directly from the cell interior to adherens junctions. *Cell* 128, 547–560.
- Stock, K., Sailer, R., Strauss, W.S.L., *et al.*, 2003. Variable-angle total internal reflection fluorescence microscopy (VA-TIRFM): realization and application of a compact illumination device. *Journal of Microscopy (Oxford)* 211, 19–29.
- Stout, A.L., Axelrod, D., 1989. Evanescent field excitation of fluorescence by epi-illumination microscopy. *Applied Optics* 28, 5237–5242.
- Stout, A.L., Axelrod, D., 1994. Reversible binding kinetics of a cytoskeletal protein at the erythrocyte submembrane. *Biophysical Journal* 67, 1324–1334.
- Sund, S.E., Axelrod, D., 2000. Actin dynamics at the living cell submembrane imaged by total internal reflection fluorescence photobleaching. *Biophysical Journal* 79, 1655–1669.
- Tatsumi, H., Katayama, Y., Sokabe, M., 1999. Attachment of growth cones on substrate observed by multi-mode light microscopy. *Neurosciences Research* 35, 197–206.
- Todd, I., Mellor, J.S., Gingell, D., 1988. Mapping of cell-glass contacts of *Dictyostelium amoebae* by total internal reflection aqueous fluorescence overcomes a basic ambiguity of interference reflection microscopy. *Journal of Cell Science* 89, 107–114.
- Wang, M.D., Axelrod, D., 1994. Time-lapse total internal reflection fluorescence video of acetylcholine receptor cluster formation on myotubes. *Developmental Dynamics* 201, 29–40.
- Wang, X.H., Teng, Y., Wang, Q.L., *et al.*, 2006. Imaging of dynamic secretory vesicles in living pollen tubes of *Picea meyeri* using evanescent wave microscopy. *Plant Physiology* 141, 1591–1603.
- Weis, R.M., Balakrishnan, K., Smith, B., McConnell, H.M., 1982. Stimulation of fluorescence in a small contact region between rat basophil leukemia cells and planar lipid membrane targets by coherent evanescent radiation. *Journal of Biological Chemistry* 257, 6440–6445.
- Weiss, A.N., Anantharam, A., Bittner, M.A., Axelrod, D., Holz, R.W., 2014a. Lumenal protein within secretory granules affects fusion pore expansion. *Biophysical Journal* 107, 26–33.
- Weiss, A.N., Bittner, M.A., Holz, R.W., Axelrod, D., 2014b. Protein mobility within secretory granules. *Biophysical Journal* 107, 16–25.
- Yoneda, A., Ushakov, D., Mulhaupt, H.A.B., Couchman, J.R., 2007. Fibronectin matrix assembly requires distinct contributions from rho kinases I and II. *Molecular Biology of the Cell* 18, 66–75.

Super Resolution Fluorescence Localization Microscopy

MJ Mlodzianoski, MM Valles, and ST Hess, University of Maine, Orono, ME, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Localization The two- or three-dimensional positional determination of a point source.

PSF The detection system's characteristic diffraction-limited image of a point source.

Resolution The smallest distinguishable distance between two point sources.

Introduction

Fluorescence microscopy is a popular imaging technique both for its relatively noninvasive properties and its ability to simultaneously image multiple species in living cellular specimens. Images are formed by the simultaneous observation of many fluorescing molecules to visualize cellular structures and processes. In conventional far-field fluorescence microscopy, the spatial resolution is diffraction-limited. Thus, the smallest resolvable details are given by the Rayleigh criterion, namely

$$r_0 \geq \frac{0.61\lambda}{NA}$$

where r_0 is the distance between point sources, λ is the wavelength of the photons detected and NA is the numerical aperture of the lens system (Born and Wolf, 1997). In fluorescence light microscopy, typical lateral and axial resolutions are ~ 200 nm and ~ 600 nm, respectively. Thus, in diffraction-limited microscopy, the size and spatial organization of cellular structures on length scales smaller than the diffraction limit remain unresolved.

Many important biological processes occur on length scales shorter than the diffraction limit, and the motivation to improve optical resolution in fluorescence microscopy has been strong. Fortunately, one can observe these phenomena through the use of localization-based super resolution techniques: stochastic optical reconstruction microscopy (STORM) (Rust *et al.*, 2006), photoactivatable localization microscopy (PALM) (Betzig *et al.*, 2006), and fluorescence photoactivation localization microscopy (FPALM) (Hess *et al.*, 2006). These and related techniques have recently become quite popular. Eric Betzig, inventor of PALM, Stefan W. Hell, and William E. Moerner, received the 2014 Nobel Prize in chemistry for their contributions to super resolution microscopy in biological imaging (Nobel Media AB, 2014).

Concept

Localization microscopy is made possible by labeling samples with specific fluorescent molecules that can be converted from a dark to bright state and eventually to another dark (bleached) state (Figure 1). Within the sample, an activation laser is used to potentiate the fluorescence of small subsets of single molecules. A second, readout laser excites the

fluorescence of those molecules so that they are imaged as diffraction-limited point-spread functions (PSFs) (Figures 1(a)–1(c)). During imaging, or, more typically after imaging, the molecular images are mathematically analyzed to determine the two- or three-dimensional coordinates of each identified molecule, which is called localization (Figures 1(e)–1(g)). One controls the molecular activation and bleaching rates to allow only a small number of molecules to be visible at once, and to cause those molecules to be separated by a distance greater than r_0 . Two-dimensional plots of this information constitute the image (Figure 1). Localization microscopy can improve the diffraction-limited resolution by about a factor of 10 or more (Figures 1(d) and 1(k)). This enables the elucidation of biological structures on sub-diffraction length scales.

Localization microscopy can image individual fluorophores as well as multiple fluorophores simultaneously (Bossi *et al.*, 2008; Testa *et al.*, 2010; Gunewardene *et al.*, 2011). Molecular orientations (Gould *et al.*, 2008), and localizations in two or three dimensions (Juette *et al.*, 2008; Huang *et al.*, 2008), can also be determined. Multicolor, molecular orientation and three-dimensional imaging require specific modifications to the detection path, involving the use of specialized optics (Figure 2) and methods (see Experimental Methods).

Super resolution imaging can be performed in live cells (Hess *et al.*, 2007) and fixed samples. Fixed cell imaging reveals static information about the size and spatial organization of the labeled structures at one instant in time. Live cell imaging reveals dynamics of cellular structures and enables the measurement of single molecule trajectories on millisecond timescales (Gudheti *et al.*, 2013; Manley *et al.*, 2008).

Experimental Methods

For particular functionalities, the microscope can be aligned to match one of the possible layouts in Figure 2. The activation laser, which frequently has a 405 nm wavelength, should be adjusted to spatially overlap with the readout laser. The readout wavelength can be one or several different wavelengths, chosen to best excite the fluorescent molecules in the sample. For example, the photoactivatable fluorescent protein Dendra2 is excited well by a 556 nm or 561 nm laser, whereas Alexa647 is optimally excited at 647 nm. Both lasers are

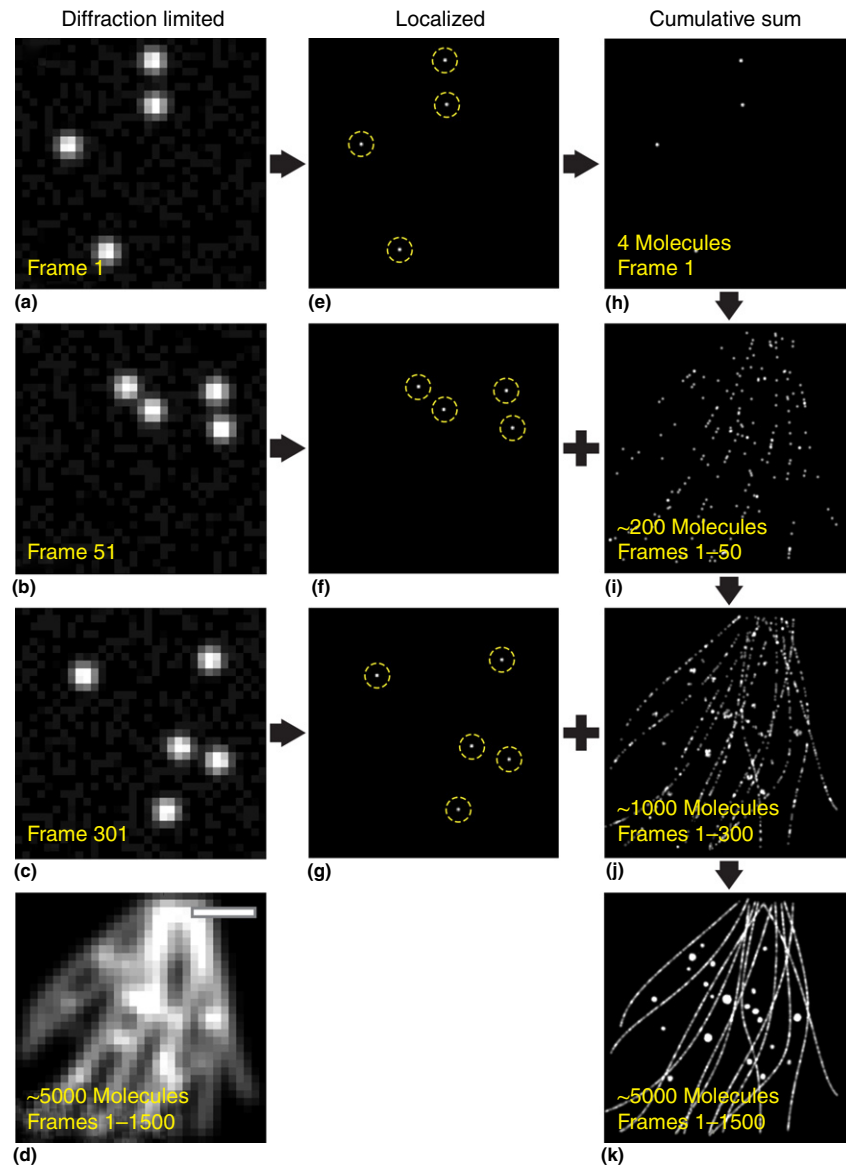


Figure 1 The concept of FPALM is shown using simulated single molecules. Diffraction-limited single molecules are sparsely activated, excited, and imaged before they photobleach (a). Photobleached molecules are replaced in each frame by newly activated single molecules (b,c). Together, all of the single molecules form a diffraction-limited image (d). The recorded single molecules in (a–c) are mathematically localized (e–g, enclosed in yellow circles). The cumulative sum of the localized molecules form super-resolved images (h–k). Sparse numbers of molecules (h,i) do not adequately define the structures as well as larger numbers of molecules (j,k). Scale bar in (d) is 1 μm and is the same for all images.

typically focused by a lens into the back aperture of the objective in order to achieve widefield illumination at the sample.

The lateral position of the laser as it enters the objective affects the excitation profile at the sample. In **Figure 2** panel I, the path of the laser lies directly on the optical axis, which results in excitation through the sample. By shifting the path of the laser laterally away from the optical axis, the laser beam becomes inclined in the sample, as seen in panel II. Further shifting of the laser from the optical axis puts the excitation into total internal reflection fluorescence (TIRF) mode, as seen in panel III. In TIRF, the evanescent electric field penetrates only a few hundred nanometers into the sample, enabling excitation of only molecules very close to the surface of the coverslip and greatly reducing out of focus background.

The objective collects a portion of the excited fluorescence, which then passes through the dichroic, and is then focused by the tube lens to form an image in an intermediate plane, where an aperture is placed to crop the area imaged. Two additional lenses form a telescope that magnifies and focuses the cropped image onto the camera. Modifications to the detection path increase the amount of information measured from the single molecules. In **Figure 2(a)**, the fluorescence signal is imaged directly onto the camera and provides two-dimensional spatial information.

Simultaneous multicolor imaging, i.e., the labeling of a sample with multiple spectrally distinct species of fluorescent molecules, can be done using the optical setup shown in **Figure 2(b)** (Bossi *et al.*, 2008; Testa *et al.*, 2010;

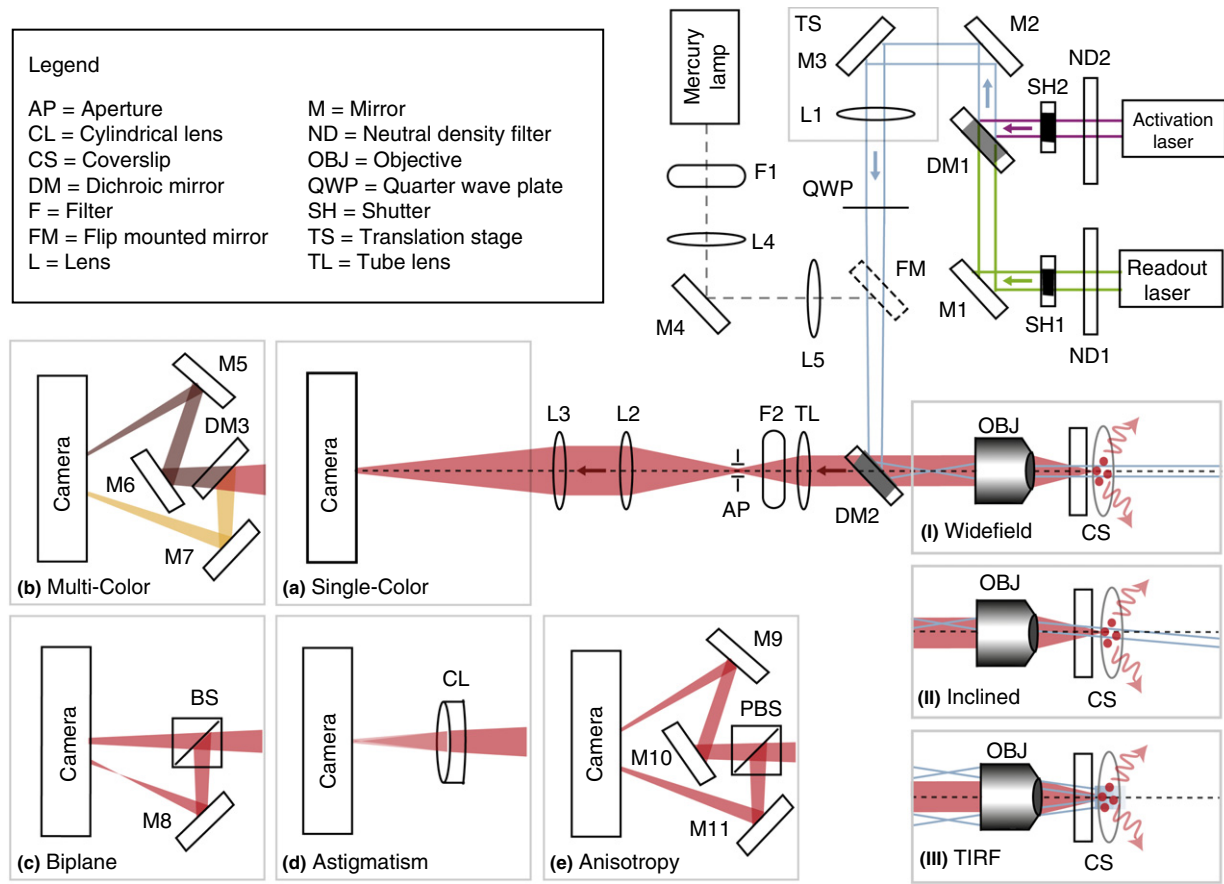


Figure 2 Basic layout of an FPALM setup and many modifications. The Readout and Activation laser on/off and powers are controlled by shutters, SH, and neutral density filters, ND, respectively. The lasers are combined at DM1 and reflected to the lens, L1, by mirrors M2 and M3. The lasers pass through a quarter wave plate, QWP, for conversion to circularly polarized light, then reflect off the filter cube dichroic, DM2, and focus at the back aperture of the objective, OBJ. The lateral position of the lasers at the back aperture of the objective controls the angle at which the lasers enter the sample: (I) centered for excitation perpendicular to the coverslip, (II) slightly offset for inclined beam in the sample, and (III) offset for TIRF. The lateral position of the laser is adjusted by shifting the translation stage, TS. Sample fluorescence is collected by the objective, passes through DM2, and is focused by the tube lens, TL. Outside the microscope, additional bandpass filters, F2, remove background. The aperture, AP, reduces the image region. The image is expanded by the telescope lenses, L2 and L3, before entering the detection box. Modifications to the detection path enable different modes of imaging. For single color imaging, fluorescence is measured directly by the camera (a). For simultaneous multicolor imaging, a dichroic, DM3, separates the fluorescence into two channels (b). Additional mirrors, M5 and M6, ensure the same focal plane is imaged. For three-dimensional imaging by the biplane method, the fluorescence is split by a 50:50 beam splitter, BS, with M8 redirecting reflected fluorescence to the camera (c). The different path lengths result in two different focal planes imaged at the camera. For three-dimensional images by astigmatism, a cylindrical lens is placed in the detection path (d). To measure anisotropy, a polarizing beam splitter, PBS, splits fluorescence depending on polarization, with M9 and M10 ensuring path-length equality (e). Light from the Mercury Lamp passes through a filter, F1, and expands with the telescope lenses L4 and L5. A flip mirror, FM, reflects the mercury lamp light into the microscope.

Gunewardene *et al.*, 2011). The dichroic splits the fluorescence into two color channels and the mirrors adjust the path lengths so that the same focal planes are imaged in side-by-side areas on the camera. The ratio of intensities between the two channels can be used to distinguish the different fluorescent species in the image.

Three-dimensional imaging requires a method to break the axial symmetry of the PSF. Breaking the symmetry can be done by a variety of methods (Juette *et al.*, 2008; Huang *et al.*, 2008; Pavani *et al.*, 2009; Shtengel *et al.*, 2009). Here, we show the two easiest methods to implement: Biplane (Juette *et al.*, 2008) in Figure 2(c) and astigmatism (Huang *et al.*, 2008) in Figure 2(d). The Biplane method splits the fluorescence

with a 50:50 beam splitter cube so that two different focal planes are imaged on the camera. This provides two separate measurements of the single molecule enabling axial localization. The fluorescence path lengths are different for each channel, resulting in two different focal planes being imaged on the camera. The differences in shape of the molecular images in the two channels enable the localization of the axial position of the molecule. The astigmatism method introduces a cylindrical lens into the detection path as seen in Figure 2(d). The cylindrical lens changes the focal length along only one of the lateral directions. The image of the single molecule will be stretched either horizontally or vertically depending on its axial position. This stretching is then

mathematically modeled to determine the molecular position in three dimensions.

The anisotropy, or dipole orientation, of the imaged single molecules can be measured using the setup in [Figure 2\(e\)](#) ([Gould *et al.*, 2008](#)). The fluorescence is split by a polarizing beam splitter cube and the mirrors adjust the path lengths so that the same focal planes are imaged on the camera. The ratio of intensities between the two channels enables determination of the azimuthal (in-plane) angle of the dipole of the single molecule.

Preparation of samples can follow standard protocols for fluorescence labeling, however, several crucial details need to be considered. Photoactivatable single molecules, by nature, are sensitive to light. Once labeled, samples should be shielded from stray light sources (room lights, incubator UV, etc) which may prematurely activate single molecules. Additionally, sample media and buffers often contain auto-fluorescent components which serve to increase background noise and worsen the localization precision. Sample media and buffers should be made such that auto-fluorescent components are removed or replaced with nonfluorescent alternatives.

For any multichannel or three-dimensional imaging modality, calibration data should be recorded. Multichannel calibration typically consists of bead images which can be used to create a mathematical transform to overlay the two channels. Three-dimensional calibration data requires precise positioning of beads in order to generate an axial calibration curve.

Data Analysis

Typical localization-based super resolution datasets consist of between a few hundred and several tens of thousands of image frames of single molecules. The crucial point of the technique is to measure spatially well-separated fluorescent single molecules in each frame, as seen in [Figures 1\(a\)–1\(c\)](#). As the single molecules photobleach and the pool of inactive molecules decreases, the activation intensity should be increased so that the density of active single molecules remains roughly constant during the imaging sequence.

Plots of localized single molecules are shown in [Figures 1\(e\)–1\(g\)](#). Localization algorithms read in raw data and output a list of localized particle positions. These involve background subtraction before each single molecule is identified and localized. The typical algorithm first involves a calculation of the effective pixel size by using a scale image to determine how many pixels fit in a known image area. Background noise is calculated by measuring the variation of detected fluorescence within a nearby region that does not contain any fluorescing molecules. Single molecules are then identified based on pixel intensity and mathematically localized either by using a least-squares fitting of a Gaussian, maximum-likelihood estimation of a theoretical PSF ([Mortensen *et al.*, 2010](#)) or fitting to the experimentally measured PSF ([Juetten *et al.*, 2008](#)). The identification and localization process repeats for all single molecules in the image. A first set of thresholds are applied to identify objects as bright as a single fluorophore. Objects that pass this first threshold are then localized. A second set of thresholds either discards, or sets aside for further analysis, the

PSFs that are substantially larger or smaller than the diffraction limit; or poorly localized based on goodness of fit and number of detected photons. A localization precision is then determined for each PSF to be analyzed, which depends on the number and emission wavelength of the number of photons detected from the fluorophore, the background noise per pixel and the effective pixel size.

For the detection setups in [Figures 2\(b\)–2\(e\)](#), the analysis algorithms must be modified to extract the appropriate and available information. For multichannel techniques ([Figures 2\(b\), 2\(c\), and 2\(e\)](#)), the two channels are overlaid using the calibration data. Single molecules are localized using the methods in the previous paragraph to determine the lateral localized positions. The ratio of intensities and/or the comparison of the shapes of the molecular images in the two channels reveals information such as the type, axial position, or dipole orientation of the single molecule. The images in [Figure 2\(d\)](#) form a single channel; the axial positions of the single molecules are extracted by measuring the horizontal and vertical widths of the PSF for each molecule within the image.

A final rendered image of all localized single molecules ([Figure 1\(k\)](#)) will have enhanced resolution in comparison to the diffraction-limited image ([Figure 1\(d\)](#)). Gaussian-based visualization was used to render the super resolution image in [Figure 1](#), but several other methods exist, such as scattergrams, quad-tree histograms, and Delaunay triangulation ([Baddeley *et al.*, 2010](#)).

For a more thorough discussion of the microscope setup, alignment, imaging, and analysis methods, see [Gould *et al.* \(2009\)](#), [Curthoys *et al.* \(2013\)](#).

Discussion

Localization microscopy can provide resolution at near molecular length scales, with time resolution of 0.1–0.03 s per rendered image ([Huang *et al.*, 2013](#); [Gunewardene *et al.*, 2011](#); [Shim *et al.*, 2012](#)). Because localization microscopy images single molecules, a number of different types of information can be obtained.

Live cell images can be obtained with frame rates of ~ 1 ms, allowing rendered images to be obtained within 0.1 s for fluorescent proteins (unpublished data) and 0.03 s for organic dyes ([Huang *et al.*, 2013](#)). Careful choice of excitation intensity can allow individual molecules to be imaged for multiple frames and thereby single molecule trajectories to be obtained ([Gudheti *et al.*, 2013](#); [Manley *et al.*, 2008](#)). From these trajectories, the diffusion or transport behavior of molecular species can be obtained for single species ([Hess *et al.*, 2007](#); [Manley *et al.*, 2008](#)), or combinations of species ([Gudheti *et al.*, 2013](#)). Multiple molecular species can be distinguished by splitting the detected fluorescence based on wavelength to form two or more simultaneous images, and using the ratio of intensities within those images to determine molecular type ([Bossi *et al.*, 2008](#); [Gunewardene *et al.*, 2011](#); [Testa *et al.*, 2010](#)). Alternatively, probes may also be distinguished based on activation or readout wavelength in addition to emission spectrum ([Bates *et al.*, 2007](#)). Finally, methods to quantify numbers of molecules or molecular

density are being developed (Manley *et al.*, 2008; Lee *et al.*, 2012; Wolter *et al.*, 2012; Avilov *et al.*, 2014).

Although a wide variety of information can be obtained from a biological sample, a number of parameters must be considered in order to optimize use of localization microscopy for such applications.

The molecular labeling density (Figures 1(h)–(k)) and localization have a profound effect on the ultimate quality (and resolution) of images obtained. The localization precision is strongly affected by number of detected photons per molecule, background noise per pixel, camera pixel size, and the width of the diffraction-limited PSF (Thompson *et al.*, 2002). Thus, minimization of fluorescence background and maximization of fluorescence detection efficiency are crucial. Minimization of other sources of image noise, such as pixel-to-pixel (spatial) (Pertsinidis *et al.*, 2010) and frame-to-frame (temporal) variation in camera sensitivity, are also important (Huang *et al.*, 2013). Based on the Nyquist sampling theorem, the density of labels should yield a nearest neighbor distance no larger than half the value of the desired resolution (Shroff *et al.*, 2007). More generally, the resolution can be defined in a number of ways (Betzig *et al.*, 2006; Hess *et al.*, 2006; Mukamel and Schnitzer, 2012; Nieuwenhuizen *et al.*, 2013). Recently, a resolution measure based on information content as a function of spatial frequency has led to a ‘Fourier ring correlation’ method, which quantifies the effects of localization precision and molecular density for a wide variety of conditions applicable to localization microscopy (Nieuwenhuizen *et al.*, 2013). Use of this method can be made through an ImageJ app.

To optimize image quality, molecular coordinates in localization microscopy datasets need to be corrected for axial and lateral sample drift, such as through the use of fiducials (Betzig *et al.*, 2006), cross-correlation of data subsets (obtained in fixed samples only) (Mlodzianoski *et al.*, 2011), or experimental modifications need to be made to prevent drift during acquisition, such as through autofocus and/or lateral stabilization of the microscope stage relative to the objective lens (Huang *et al.*, 2008).

For multicolor datasets, bleedthrough, defined as inadvertent identification of molecules of one species as another, can cause significant errors in interpretation of results (Kim *et al.*, 2013). Correction for misidentification rates of as little as a few percent is crucial if results are to be used for quantification of colocalization (Kim *et al.*, 2013).

In summary, localization microscopy is well-suited for imaging living and fixed cell specimens, and other biological samples, where fluorescent labels can be attached to the biomolecules of interest, where spatial resolution of tens of nanometers is desirable, and where a time resolution of 0.03 s or slower is sufficient. Molecular trajectories can also be obtained on millisecond timescales. Localization microscopy is to some extent limited by data analysis and the fluorescent labels used for imaging. Without considerable effort, immediate visualization is hindered by the data analysis required. Further development of the fluorescent labels used could further improve temporal and spatial resolution. However, the range of potential biological applications of the technology in its current state is extensive, and further insights into cell biology are likely to result.

Acknowledgments

This work was funded by NIH R15 GM094713, NIH R15 ES024593, NIH R01 AR054170, Maine Technology Institute MTA 1106 and 2061, the UMaine Office of the Vice President for Research, and the Maine Economic Improvement Fund.

See also: Imaging the Cell: Light Microscopy: Super-Resolution Light Microscopy: Stimulated Emission Depletion and Ground-State Depletion

References

- Avilov, S., Berardozi, R., Gunewardene, M.S., *et al.*, 2014. In cellulo evaluation of phototransformation quantum yields in fluorescent proteins used as markers for single-molecule localization microscopy. *Plos One* 9 (6), e98362.
- Baddeley, D., Cannell, M.B., Soeller, C., 2010. Visualization of localization microscopy data. *Microscopy and Microanalysis* 16, 64–72.
- Bates, M., Huang, B., Dempsey, G.T., Zhuang, X., 2007. Multicolor super-resolution imaging with photo-switchable fluorescent probes. *Science* 317, 1749–1753.
- Betzig, E., Patterson, G.H., Sougrat, R., *et al.*, 2006. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313, 1642–1645.
- Born, M., Wolf, E., 1997. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. New York, NY: Cambridge University Press.
- Bossi, M., Folling, J., Belov, V.N., *et al.*, 2008. Multicolor far-field fluorescence nanoscopy through isolated detection of distinct molecular species. *Nano Letters* 8, 2463–2468.
- Curthoys, N.M., Mlodzianoski, M.J., Kim, D., Hess, S.T., 2013. Simultaneous multicolor imaging of biological structures with fluorescence photoactivation localization microscopy. *Journal of Visualized Experiments* 9 (82), e50680.
- Gould, T.J., Gunewardene, M.S., Gudheti, M.V., *et al.*, 2008. Nanoscale imaging of molecular positions and anisotropies. *Nature Methods* 5, 1027–1030.
- Gould, T.J., Verkhusha, V.V., Hess, S.T., 2009. Imaging biological structures with fluorescence photoactivation localization microscopy. *Nature Protocols* 4, 291–308.
- Gudheti, M.V., Curthoys, N.M., Gould, T.J., *et al.*, 2013. Actin mediates the nanoscale membrane organization of the clustered membrane protein influenza hemagglutinin. *Biophysical Journal* 104, 2182–2192.
- Gunewardene, M.S., Subach, F.V., Gould, T.J., *et al.*, 2011. Superresolution imaging of multiple fluorescent proteins with highly overlapping emission spectra in living cells. *Biophysical Journal* 101, 1522–1528.
- Hess, S.T., Girirajan, T.P.K., Mason, M.D., 2006. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical Journal* 91, 4258–4272.
- Hess, S.T., Gould, T.J., Gudheti, M.V., *et al.*, 2007. Dynamic clustered distribution of hemagglutinin resolved at 40 nm in living cell membranes discriminates between raft theories. *Proceedings of the National Academy of Sciences of the United States of America* 104, 17370–17375.
- Huang, B., Wang, W.Q., Bates, M., Zhuang, X.W., 2008. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* 319, 810–813.
- Huang, F., Hartwich, T.M.P., Rivera-Molina, F.E., *et al.*, 2013. Video-rate nanoscopy using sCMOS camera-specific single-molecule localization algorithms. *Nature Methods* 10, 653–658.
- Juette, M.F., Gould, T.J., Lessard, M.D., *et al.*, 2008. Three-dimensional sub-100 nm resolution fluorescence microscopy of thick samples. *Nature Methods* 5, 527–529.
- Kim, D., Curthoys, N.M., Parent, M., Hess, S.T., 2013. Bleed-through correction for rendering and correlation analysis in multi-colour localization microscopy. *Journal of Optics* 15, 094011.
- Lee, S.H., Shin, J.Y., Lee, A., Bustamante, C., 2012. Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM). *Proceedings of the National Academy of Sciences of the United States of America* 109, 17436–17441.
- Manley, S., Gillette, J.M., Patterson, G.H., *et al.*, 2008. High-density mapping of single-molecule trajectories with photoactivated localization microscopy. *Nature Methods* 5, 155–157.

- Mlodzianoski, M.J., Schreiner, J.M., Callahan, S.P., *et al.*, 2011. Sample drift correction in 3D fluorescence photoactivation localization microscopy. *Optics Express* 19, 15009–15019.
- Mortensen, K.I., Churchman, L.S., Spudich, J.A., Flyvbjerg, H., 2010. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nature Methods* 7, 377–U59.
- Mukamel, E.A., Schnitzer, M.J., 2012. Unified resolution bounds for conventional and stochastic localization fluorescence microscopy. *Physical Review Letters* 109 (16), 168102.
- Nieuwenhuizen, R.P., Lidke, K.A., Bates, M., *et al.*, 2013. Measuring image resolution in optical nanoscopy. *Nature Methods* 10, 557–562.
- Nobel Media AB, 2014. The Nobel Prize in Chemistry 2014 — Press Release. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2014/press.html (accessed 08.11.14).
- Pavani, S.R.P., Thompson, M.A., Biteen, J.S., *et al.*, 2009. Three-dimensional, single-molecule fluorescence imaging beyond the diffraction limit by using a double-helix point spread function. *Proceedings of the National Academy of Sciences of the United States of America* 106, 2995–2999.
- Pertsinidis, A., Zhang, Y.X., Chu, S., 2010. Subnanometre single-molecule localization, registration and distance measurements. *Nature* 466, 647–U11.
- Rust, M.J., Bates, M., Zhuang, X.W., 2006. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods* 3, 793–795.
- Shim, S.H., Xia, C.L., Zhong, G.S., *et al.*, 2012. Super-resolution fluorescence imaging of organelles in live cells with photoswitchable membrane probes. *Proceedings of the National Academy of Sciences of the United States of America* 109, 13978–13983.
- Shroff, H., Galbraith, C.G., Galbraith, J.A., *et al.*, 2007. Dual-color superresolution imaging of genetically expressed probes within individual adhesion complexes. *Proceedings of the National Academy of Sciences of the United States of America* 104, 20308–20313.
- Shtengel, G., Galbraith, J.A., Galbraith, C.G., *et al.*, 2009. Interferometric fluorescent super-resolution microscopy resolves 3D cellular ultrastructure. *Proceedings of the National Academy of Sciences of the United States of America* 106, 3125–3130.
- Testa, I., Wurm, C.A., Medda, R., *et al.*, 2010. Multicolor fluorescence nanoscopy in fixed and living cells by exciting conventional fluorophores with a single wavelength. *Biophysical Journal* 99, 2686–2694.
- Thompson, R.E., Larson, D.R., Webb, W.W., 2002. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal* 82, 2775–2783.
- Wolter, S., Loschberger, A., Holm, T., *et al.*, 2012. rapidSTORM: Accurate, fast open-source software for localization microscopy. *Nature Methods* 9, 1040–1041.

Super-Resolution Light Microscopy: Stimulated Emission Depletion and Ground-State Depletion

Y Wu, X Wu, R Lu, M Li, L Toro, and E Stefani, University of California Los Angeles, Los Angeles, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Super-resolution light microscopy includes a set of techniques that are capable of reaching optical resolution beyond the classical diffraction limit. Diffraction is a fundamental property of light that restricted the resolution of light microscopy for over 300 years. The classical Abbe's law and the Rayleigh criterion state the lateral (The lateral plane is perpendicular to the optical axis. The optical axis is along the direction where light travels.) resolution of a light microscope d_{lat} is limited to:

$$d_{\text{lat}} \approx 0.61\lambda/\text{NA} \quad [1]$$

where λ is the wavelength of light, and NA is the numerical aperture of the microscope objective (Inoue, 1995). Within the visible spectrum (wavelength range from 400 nm to 700 nm), the lateral resolution is thus limited to ~ 170 – 300 nm using high NA oil immersion objectives. The diffraction limit remained a fundamental obstacle of optical microscopy until the invention of a number of 'super-resolution' fluorescence microscopy techniques since the new millennium. The most popular super-resolution techniques include stimulated emission depletion (STED) microscopy (Hell, 2003), scanning ground-state depletion (GSD) microscopy (Bretschneider *et al.*, 2007), photo-activated localization microscopy (PALM) (Betzig *et al.*, 2006), stochastic optical reconstruction microscopy (STORM) (Rust *et al.*, 2006), super-resolution optical fluctuation imaging (SOFI) (Dertinger *et al.*, 2009), and structured illumination microscopy (SIM) (Gustafsson, 2000). This contribution is mainly focused on STED microscopy, which has the following advantages: (1) it achieves super-resolution via a purely physics mechanism and is free from mathematical or numerical processing; (2) unlike SIM, it is in principle capable of reaching unlimited optical resolution (Nonlinear SIM is capable of unlimited optical resolution but lacks practical implementation (Gustafsson, 2005).); (3) unlike stochastic methods, such as PALM/STORM and SOFI, that need to collect thousands of images to obtain super-resolution, STED imaging is performed by point-scanning and can be completed with relatively short time; and (4) it is based on the confocal scheme which enables optical sectioning.

We will also briefly discuss scanning GSD microscopy, a technique very similar to STED microscopy, and ground-state depletion after individual molecule return imaging (GSDIM) (Also referred to as direct STORM or dSTORM (Heilemann *et al.*, 2008)). Scanning GSD microscopy and GSDIM are two very different techniques though based on the same physical mechanism.) (Folling *et al.*, 2008), a variant of STORM. STED and GSD were invented by Dr. Stefan W. Hell, for which he was awarded the 2014 Nobel Prize in Chemistry.

It is worth noting that diffraction is a fundamental property of light which the super-resolution techniques do not eliminate. Instead, all the super-resolution fluorescence microscopy techniques are based on tricks that switch the fluorophores between an on-state, where the fluorophores are fluorescent, and an off-state, where the fluorophores are dark. Two fluorophores are temporally differentiated by placing one of them in the on-state while the other one in the off-state at any given time. A diagram of the low-lying electronic states of a common fluorophore is shown in Figure 1. The ground state S_0 is the stable lowest-energy state, from which the fluorophore can absorb a photon and be excited to the first excited singlet state S_1 . From S_1 , the fluorophore may: (1) spontaneously decay to S_0 and emit fluorescence in a few nanoseconds; (2) undergo an intersystem crossing (ISC) and enter the lowest-lying triplet state T_1 with a long lifetime (depending on the type of dye and the microenvironment, T_1 lifetime ranges from several microseconds to hundreds of milliseconds); and (3) be forced to decay to the ground state S_0 immediately by stimulated emission. STED and GSD are two mechanisms that manipulate the fluorophore states differently. STED uses stimulated emission to quench fluorophores from the on-state S_1 to the off-state S_0 ; whereas GSD uses ISC to deplete fluorophores from the singlet on-states S_0 and S_1 to the dark triplet off-state T_1 .

In the following sections we will briefly overview various super-resolution techniques based on STED and GSD. We will emphasize the physical principles and the technical essentials of various STED and GSD implementations and discuss their pros and cons. We will also give an incomplete review of their numerous applications to the research of cell biology.

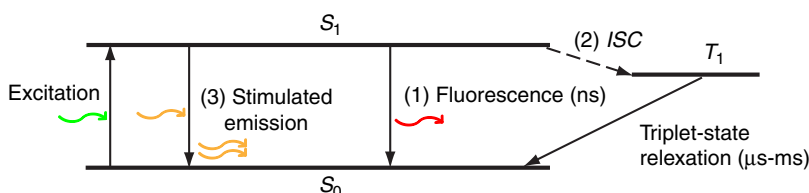


Figure 1 Low-lying electronic states of a typical fluorophore. After being excited by a photon from the ground state S_0 to the excited state (on-state) S_1 , the fluorophore has three fates: (1) fluorescence emission; (2) intersystem crossing; and (3) stimulated emission. STED and GSD uses S_0 and T_1 , respectively, as the off-state to switch the fluorophore.

STED and GSD for Super-Resolution: Principle and Technology

STED Microscopy: General Principles

STED microscopy is based on the reversible transitions between the ground state S_0 and first excited singlet state S_1 . The transition $S_0 \rightarrow S_1$ is induced by an excitation laser beam, just like in conventional fluorescence microscopy. The reversal transition $S_1 \rightarrow S_0$ is enforced by stimulated emission, induced by a depletion laser beam at a longer wavelength than the excitation beam, as shown in Figure 2. The depletion wavelength is outside the detection band, therefore stimulated emission is considered as resulting in a dark state. The probability of the reversal transition is controlled by the irradiance of the depletion laser beam: with a sufficiently bright depletion laser beam, the reversal transition dominates, the population at S_1 approaches zero, and the fluorophore is almost always in the off-state. Simply put, a sufficiently bright depletion laser beam will shut off fluorescence.

In confocal fluorescence microscopy, the focal spot of the excitation laser scans over the sample, and the optical resolution is determined by the size of this focal spot. In STED microscopy, the depletion laser beam is spatially modulated to have a doughnut-shaped (Other hollow shapes also work, but the doughnut-shape is most efficient for depletion (Keller *et al.*, 2007).) (hollow with a zero at the center) focal spot and aligned with the excitation laser beam. Assuming two fluorophores, a and b , are located so closely that they are both covered by the excitation focal spot, as shown in Figure 3. At a certain time during STED scanning, fluorophore a sits at the center and is differentiated from fluorophore b at the rim, because fluorophore a is only illuminated by the excitation light and is thus allowed to enter the on-state and emit fluorescence, whereas fluorophore b is also illuminated by the high-irradiance depletion light that effectively keeps it in the off-state S_0 (Figure 3(a)). As a result, at this particular time, only fluorophore a can be detected. When the focal spot moves to the next position, where fluorophore b is at the center and fluorophore a is at the rim (Figure 3(b)), the microscope only detects fluorophore b . Therefore, fluorophores a and b are temporally separated and super-resolution is thus achieved. In other words, the doughnut-shaped focal spot of

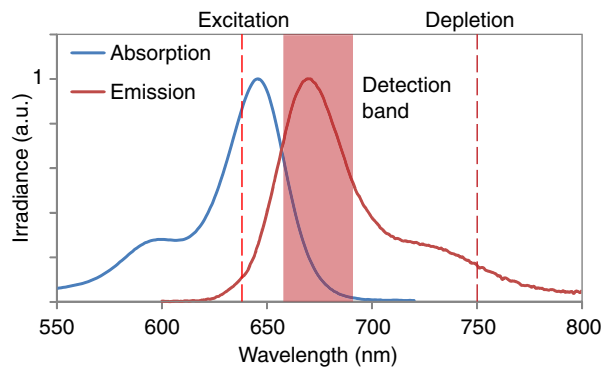


Figure 2 The depletion wavelength is chosen to be at the tail of the emission spectrum of the fluorophore, outside the detection band. Spectra of Atto 647 N are used as an example.

the depletion laser prohibits fluorescence emission from the rim of the excitation focal spot and effectively sharpens it, which enhances resolution.

Intuitively, the resolution of STED microscopy depends on how small the central hole of the depletion focal spot is. But the 'smallness' is not measured by the irradiance profile of the hole, which in fact is also subject to the diffraction limit and cannot be very small. Rather, it should be measured by the suppression factor η , defined as the fraction of the fluorescence emission being suppressed. η has a nonlinear dependence on the depletion irradiance I_{dep} : for example, with a pulsed STED configuration (see the next section for a detailed discussion) the suppression factor is

$$\eta(I) \approx 1 - \exp[-\ln 2 (I_{\text{dep}}/I_{\text{sat}})] \quad [2]$$

where I_{sat} is called saturation irradiance. When the depletion irradiance is at I_{sat} , stimulated emission and fluorescence emission are balanced, meaning that there is an equal probability for a fluorophore to be fluorescent or to be depleted (Harke *et al.*, 2008). Therefore, with increasing the depletion irradiance, the suppression factor profile will have a smaller central hole and the optical resolution of STED microscopy d_{STED} will in turn be enhanced:

$$d_{\text{STED}} \approx d_{\text{confocal}} / \sqrt{1 + a \cdot (I_{\text{dep}}/I_{\text{sat}})} \quad [3]$$

where d_{confocal} is the resolution of confocal microscopy, and a is a positive constant (Harke *et al.*, 2008). Figure 4 shows computer-simulated lateral profiles of the focal spots and the suppression factor. The depletion irradiance (red doughnut) has a fixed profile that cannot be controlled. However, when the maximum depletion irradiance I_{max} increases, the central hole of the suppression factor profile (magenta) gets smaller, and the effective focal spot (yellow) used for scanning diminishes, too. It is obvious that indefinitely increasing the

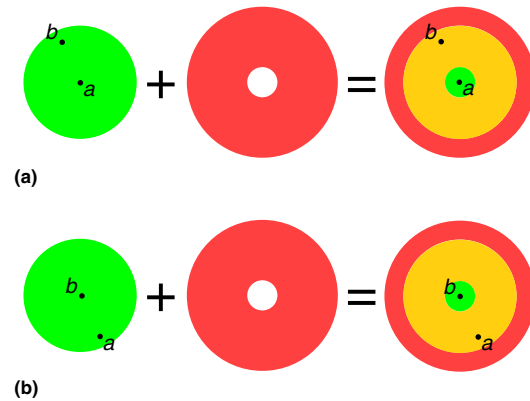


Figure 3 Schematic drawing demonstrating STED principle. Fluorophore a and fluorophore b cannot be resolved by conventional confocal microscopy as they are both covered by the excitation focal spot (green). When aligned with the hollow depletion focal spot (red): (a) at a certain time, fluorophore a is at the center and can emit fluorescence, whereas fluorophore b is at the rim (yellow = red + green, indicating illuminated by both excitation and depletion) and undetectable; (b) when scanning progresses to another time, fluorophore a and b switches positions and only b is detected.

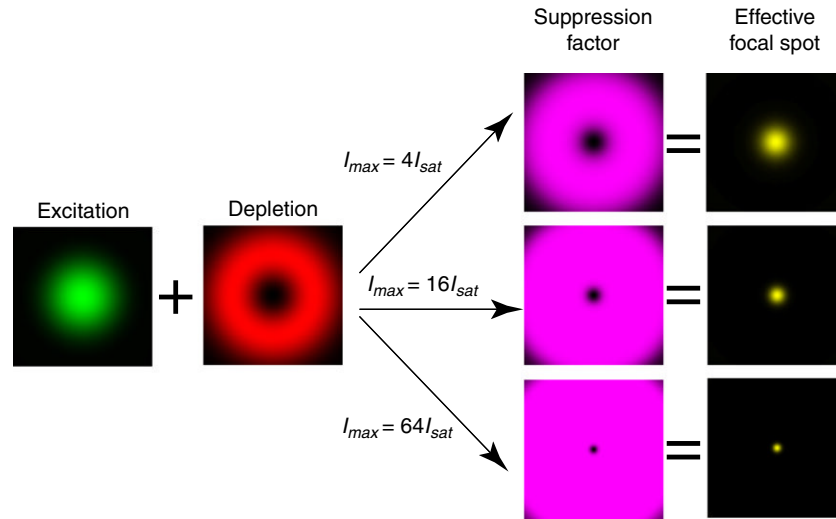


Figure 4 STED microscopy can theoretically reach unlimited resolution. When excitation focal spot (green) is aligned with depletion focal spot (red), the suppression factor profile (magenta) calculated by eqn [2] has a diminishing central hole with increasing depletion power, resulting in diminishing effective focal spot (yellow). The focal spot profiles are lateral and generated by computer simulations.

depletion irradiance can in theory result in unlimited optical resolution.

STED microscopy needs to create a central hole in the depletion beam that can be preserved during light propagation, diffraction, and focusing. The best way to achieve this is to twist the laser beam with a spatial light modulator (SLM) or a vortex phase plate (VPP) (Courtial and O’Holleran, 2007), which can add to the laser beam an angular momentum around the optical axis. Intuitively, a twisted laser beam is like a tornado traveling along its rotation axis. As long as the rotational symmetry is preserved, the light beam will always remain hollow, just like a tornado always preserves a central hole. The microscope objective strongly focuses the depletion beam, and the vector properties of light cannot be neglected in this tight focusing. To ensure the hole has a true zero at the center, the depletion beam must be circularly polarized with the same handedness as the angular momentum induced by spatial modulation (Bokor *et al.*, 2007; Hao *et al.*, 2010).

Pulsed STED Microscopy

Equation [2] indicates that a bright depletion laser beam is crucial to reach a high suppression factor leading to high STED resolution. In pulsed STED microscopy, the depletion light is a fast pulsed laser beam (‘fast’ means with short pulse duration) to reach required irradiance. A schematic drawing of a typical pulsed STED setup is shown in Figure 5. The excitation source must be temporally aligned with the depletion pulse train and therefore must also be pulsed. The excitation source can either be a standalone laser that is triggered by the depletion source, or use a portion of power of the depletion source to pump a nonlinear optics element (e.g., an optical parametric oscillator) to obtain the desired excitation wavelength. An optical delay line or an electronic time-delay circuit can be used to temporally synchronize the depletion and the excitation pulse trains. A depletion pulse follows immediately after an excitation pulse to deplete the fluorophores right after they are

excited. When a femtosecond laser source is used for depletion, the pulses need to be temporally stretched, for example, by a glass rod and a polarization-maintaining (PM) optical fiber as shown in Figure 5, to avoid multiphoton absorption and excessive photobleaching (Dyba and Hell, 2003). However, the depletion pulse duration needs to be much shorter than the fluorophore lifetime to ensure efficient depletion. For a typical fluorescent dye with a 3–5 ns lifetime, the optimal depletion pulse duration is about 0.2–1 ns. The excitation pulse duration needs to be as short as possible.

Pulsed STED microscopy is historically the earliest implementation of STED and it usually produces the best STED images. The primary challenge in pulsed STED microscopy is the availability of the depletion laser source. High power pulsed lasers are quite expensive. In the near-infrared wavelength range, a Ti:Sapphire laser source (e.g., MaiTai HP, Newport, USA) is a very good choice for depletion. In the visible wavelength range, parametric oscillators (Nagerl and Bonhoeffer, 2010), stimulated-Raman-scattering laser sources (Rankin and Hell, 2009), and super-continuum laser sources (Wildanger *et al.*, 2009) can be used for depletion.

Continuous-Wave STED Microscopy

Continuous-wave STED microscopy uses a powerful continuous-wave (CW) laser source for depletion (Willig *et al.*, 2007). The excitation source can be either CW or pulsed. It does not involve the complexity of temporal alignment in pulsed STED microscopy, and a CW laser source is more affordable than a pulsed laser source with similar optical power. The disadvantage of CW STED microscopy is that it requires higher average optical power to reach the same resolution as pulsed STED microscopy. This is obvious when the excitation light is pulsed and eqn [3] holds. The average irradiance of a CW depletion laser source needs to be I_{dep} constantly to reach the resolution d_{STED} ; whereas for a pulsed depletion source the irradiance I_{dep} only lasts for its pulse duration T_w , and the average irradiance

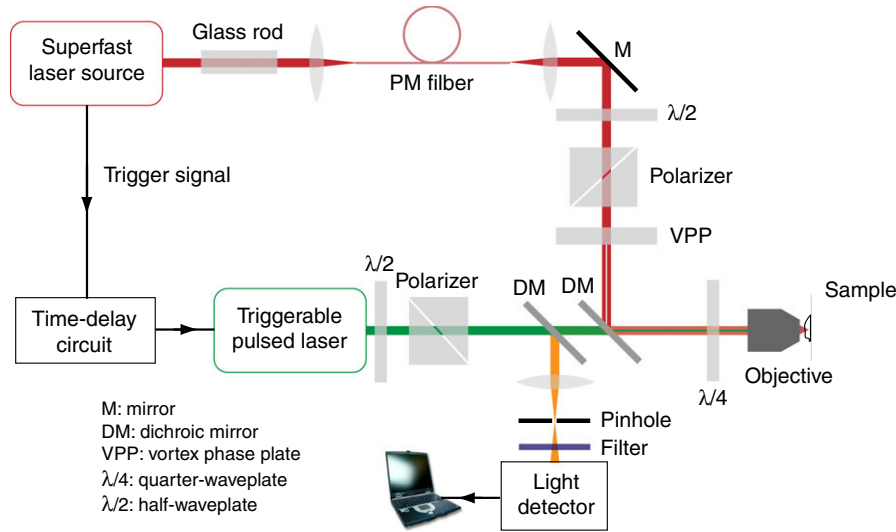


Figure 5 Simplified setup of a typical pulsed STED microscope. A superfast (pulse duration in femtoseconds) laser is used as the depletion source and triggers another pulsed laser source for excitation. A time-delay circuit synchronizes the depletion and the excitation pulse trains. The femtosecond pulses are stretched by a glass rod and a PM fiber to reach pulse duration of several hundred picoseconds. Depletion light is then modified by a VPP to have a hollow cross-section. Glan-laser polarizers and a quarter-waveplate ($\lambda/4$) ensure the light is circularly polarized before being focused by a microscope objective. Fluorescence is collected by the same objective and goes through a confocal pinhole to block out-of-focus light.

is $I_{\text{dep}} T_w R$, where R is the laser repetition rate. Therefore, the average depletion power needed in CW STED microscopy is $(T_w R)^{-1}$ times higher than in pulsed STED microscopy. Compared with using a common Ti:Sapphire laser with a repetition rate of 80 MHz laser and 400 ps pulse duration for depletion, CW STED microscopy needs 30 times higher depletion power to reach the same resolution. Such powerful CW laser sources are often not available, and therefore CW STED microscopes usually reach lower resolution than pulsed ones. Plus, high depletion power induces excessive photobleaching that can permanently destroy the fluorescent dyes. If the excitation is also CW, then excitation and depletion constantly compete against each other, which also results in higher depletion power requirement (Willig *et al.*, 2007).

To summarize, compared with pulse STED microscopy, CW STED microscopy has the advantage of being technically simpler and working with less-expensive laser sources. However, its resolution is usually lower and it suffers from rapid photobleaching.

Time-Gated STED Microscopy

Time-gated detection was invented to enhance the resolution and reduce photobleaching in CW STED microscopy (Vicidomini *et al.*, 2011). In time-gated STED microscopy, the excitation source is pulsed while the depletion source is CW and with relatively low power. Fluorescence detection only takes place within a time-window after a fixed time-delay T_d following each excitation pulse, as shown in Figure 6. The rationale is that, with the low depletion power, the suppression factor η is low and the fluorescence signal from the rim cannot be completely shut off. However, depletion effectively reduces the fluorescence lifetime τ to $\tau_{\text{STED}} < \tau$, because stimulated

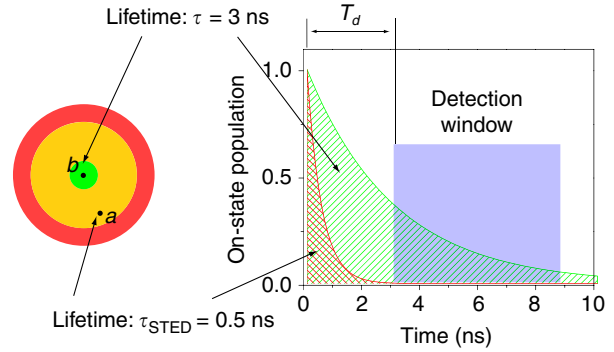


Figure 6 Schematic drawing demonstrating the principle of time-gated STED. Fluorophore *a* is located at the rim (yellow) and has a short lifetime $\tau_{\text{STED}} = 0.5$ ns (red shadow) due to depletion; fluorophore *b* is at the center (green) with normal fluorescence lifetime $\tau = 3$ ns (green shadow). Setting a detection window (light blue) with a delay $T_d = 3$ ns resulting in detecting fluorophore *b* only, because $T_d \gg \tau_{\text{STED}}$, and fluorophore *a* is effectively dark.

emission provides an extra pathway for the fluorophore to decay from the first excited state S_1 to the ground state S_0 . The difference in lifetime is exploited by the detection time-window to differentiate fluorophores. The time-delay T_d needs to be carefully chosen. On one hand, T_d should satisfy $T_d \gg \tau_{\text{STED}}$, so that fluoresce at the rim of the focal spot can hardly be detected; on the other hand, T_d should be as short as possible to detect more fluorescence signal at the center. In practice, one needs to work with these two conflicting conditions to find a compromise between the signal level and resolution.

The detection time-window can be either implemented by commercial time-correlated single photon counting (TCSPC) modules or custom-built gating electronics. TCSPC modules

are designed for lifetime microscopy, which are commercially available from PicoQuant or Becker and Hickl, GmbH. With these modules, the arrival time of each photon relative to the excitation pulses can be recorded, and the detection time-window is applied by software (Viciomini *et al.*, 2011). Alternatively, one can simply use a leading edge discriminator with a VETO input (e.g., TD2000 from Fast Comtec, GmbH) to build one's own gating electronics (Wu *et al.*, 2015).

To summarize, time-gated detection enhances resolution by differentiating fluorescence with usual lifetime τ from fluorescence with shortened lifetime τ_{STED} due to depletion. The trade-off is the lowered fluorescence signal level. We only recommend time-gated STED technique to be used for bright samples.

STED in the Axial Direction

The resolution in the axial direction, i.e., the direction where the laser beams travel, is also restricted by the diffraction limit. Similar to performing STED in the lateral direction, one can create a hollow depletion focal spot in the axial direction with SLM or VPP and use it to reach super-resolution (Klar *et al.*, 2000). Using two depletion beams, in which one is laterally hollow and the other axially hollow, super-resolution in three-dimensions (3D) can be achieved (Wildanger *et al.*, 2009). One can take a step further by combining STED with the 4Pi microscopy (Hell and Stelzer, 1992) technique and reach even higher resolution (Schmidt *et al.*, 2008). However, 4Pi-STED microscopy is highly complicated and requires a special sample preparation procedure, which limits its biological applications.

Multicolor STED Microscopy

Multicolor STED microscopy involves finding multiple dyes with well separable excitation and emission spectra. A simple way to perform multicolor STED is to use one depletion laser source to deplete two different dyes, one of which has a long Stokes shift (e.g., long-Stokes-shift dye Chromeo 494 from Active Motif, working together with Ato 647 N). Another way is to deplete two dyes with two different depletion laser lines (Donnert *et al.*, 2007). For example, Atto 647 N and Oregon Green 488 can be depleted at 750 nm and 592 nm, respectively (Wu *et al.*, 2010).

Two-Photon STED Microscopy

Two-photon excitation (2PE) is a widely used technology to image deeply into thick tissues or living animals. Apparently, it is rather straightforward to apply STED to 2PE as the depletion process is no different from single photon excitation (1PE) (Moneron and Hell, 2009). However, the challenge really lies in the fact that 2PE is useful to deep imaging in which the light penetrates into great depth. If the excitation light and the depletion light are at two different wavelengths, they will be scattered differently in the tissue and soon the alignment will be lost and the resolution will diminish. Luckily, it was found that for some organic dyes, at the depletion wavelength there happens to be a decent 2PE probability. Therefore, 2PE-STED

can be implemented with a single wavelength: the excitation laser pulses are very short (in femtoseconds) to induce 2PE, and the depletion laser pulses are greatly stretched to several hundred picoseconds to suppress excited fluorophores (Bianchini *et al.*, 2012). 2PE-STED microscopy faces two additional challenges. First, 2PE photobleaching is much more severe than 1PE (Patterson and Piston, 2000) and 1PE-STED photobleaching is already problematic; therefore 2PE-STED suffers greatly from photobleaching. Secondly, the depletion beam is likely not be able to maintain the hollow shape when penetrating the tissue due to light scattering, unless sophisticated technique such as adaptive optics is employed. Because of these challenges, 2PE-STED microscopy remains elusive and its application in biological research is scarce.

Photobleaching and Resonant Scanning in STED Microscopy

Photobleaching is arguably the biggest problem in STED microscopy. Photobleaching is an irreversible photochemical process that leads to the permanent destruction of the fluorescent dye. The bright depletion beam is the main cause of photobleaching. However, as the depletion pulses are stretched to several hundred picoseconds, they do not directly cause photobleaching (Dyba and Hell, 2003). Figure 7 shows that the primary STED photobleaching pathway is through the low-lying triplet state T_1 . Fluorophores stay in T_1 for a long time and may absorb additional depletion photons to enter the highly unstable high-energy triplet state, T_n , where they are quickly bleached. The long lifetime of T_1 (from microseconds to hundreds of milliseconds) is the key issue as the T_1 population can be accumulating during imaging (Deschenes and Bout, 2002).

Besides common practices such as using anti-fade mounting media and lower laser power, the efforts of reducing STED photobleaching is concentrated on eliminating the triplet-state (T_1) buildup. This can be done with oxygen removal from the microenvironment and using a reducing and oxidizing system (ROXS) as buffer that can quickly recover fluorophores from T_1 (Vogelsang *et al.*, 2008). Another idea is to avoid continuous illumination to give fluorophores enough time in dark for its recovery from T_1 . Triplet-state relaxation (T-Rex) technique is to use low repetition rate pulsed lasers for excitation and

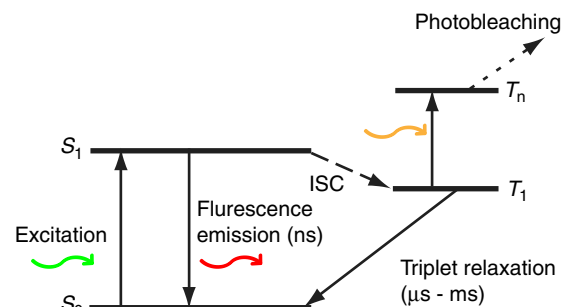


Figure 7 Primary photobleaching pathway in STED microscopy. Fluorophore is excited to the first excited singlet state S_1 , and then undergo ISC to enter the long-lived low-lying triplet state T_1 , where it may absorb a depletion photon to reach the high-energy triplet state T_n and then quickly bleaches.

depletion (Donnert *et al.*, 2007). For example, if the T_1 lifetime is $\sim 1 \mu\text{s}$, T-Rex will use lasers with < 1 MHz repetition rate, so that between consecutive laser pulses the fluorophores can relax from T_1 and go back to the ground state S_0 . This technique is very successful in reducing photobleaching but considerably slows down imaging and is often not practical for fluorophores with very long triplet-state lifetime. The bunched T-Rex technique is a modified version of T-Rex, which bundles 4–5 high repetition rate pulses together and separates the bunches with time periods that allow for T_1 relaxation (Donnert *et al.*, 2009). Another way to relax the triplet-state population is to use fast scanning speed, with which fluorophores are not constantly illuminated since the laser focal spot quickly scans them and then leave (Gardeazabal Rodriguez *et al.*, 2012). We showed that increasing scanning speed with resonant scanners and large field of view can significantly reduce photobleaching without sacrificing the imaging speed (Wu *et al.*, 2014).

Resonant scanning is crucial to implement vide-rate STED imaging, which has reached 28 frames per second (fps) for biological samples (Westphal *et al.*, 2008). Currently, imaging speed of live STED microscopy is not limited by scanning technology but by low acquired fluorescence signal at high frame rate. For bright enough specimens such as fluorescent bead samples, the imaging speed reached as high as 80 fps (Westphal *et al.*, 2007).

Scanning GSD Microscopy and GSDIM

GSD using the low-lying triplet state T_1 as the off-state. It is usually implemented by illuminating the fluorophores with saturated excitation laser power, causing them to be excited and gradually accumulate in the dark T_1 state, based on the long lifetime of T_1 . This process diminishes the S_0 population, which gives this technique its name. The advantage of GSD over STED is that GSD needs much lower depletion irradiance. In STED, depletion competes against spontaneous fluorescence emission, which has lifetime in nanoseconds; in GSD, depletion competes against relaxation from T_1 , which happens with much lower probability than fluorescence emission and has a lifetime in microseconds or even hundreds of milliseconds. Therefore, in GSD one can work with CW depletion laser sources with moderate optical power (Bretschneider *et al.*, 2007; Hell and Kroug, 1995).

There are two categories of methods that employ the GSD mechanism to achieve super-resolution: one is based on point-scanning and we call it scanning GSD microscopy; the other is GSDIM (dSTORM), a variant of PALM/STORM that is based on the wide-field technique using camera. Scanning GSD microscopy is very similar to STED microscopy, just replacing STED with GSD as its depletion mechanism. The advantage of scanning GSD microscopy over STED microscopy is its low need of depletion laser power. However, it has a serious issue of photobleaching due to constant illumination on fluorophores in the T_1 state. The imaging speed is also slower, because the transitions between the off-state (T_1) and the on-state (S_1) are much slower than stimulated emission. Scanning GSD microscopy was demonstrated to be very successful for nitrogen-vacancy centers in diamonds, which do not have the photobleaching problem (Han *et al.*, 2010). However, it is not

widely used in biological research due to the abovementioned disadvantages.

GSDIM is probably the better approach to apply GSD for super-resolution microscopy. After the fluorophores initially receive saturated excitation optical power and enter the dark T_1 state, they will randomly return back to the ground state S_0 , be excited and reenter the on-state S_1 , and release fluorescence due to spontaneous decay. The circle of $S_0 \rightleftharpoons S_1$ will repeat for a number of times before they enter the dark state T_1 again. In GSDIM, the fluorescence emission from the returning fluorophores is recorded by a camera. Although each return is corresponding to a diffraction-limited spot on the camera, the center of the spot can be mathematically determined which indicate the precise location of the fluorophore (Folling *et al.*, 2008). GSDIM is of course very similar to PALM/STORM. The difference is that GSDIM normally uses common organic dyes being switched between the dark triplet states and the fluorescent singlet state, whereas PALM/STORM uses specially engineered photo-switchable dyes or fluorescent proteins. In fact, GSDIM and similar techniques were independently invented by a number of groups and given different names, for example, direct STORM (dSTORM) (Heilemann *et al.*, 2008, 2009). GSDIM is also capable of achieve super-resolution in the axial direction by differentiating focal spots with astigmatism (Huang *et al.*, 2008). The major disadvantage of GSDIM is that it can only work with dyes with large triplet-state yield yet not too vulnerable to triplet-state photobleaching. The common practice is to use buffers that maximize the triplet-state yield of organic dyes (Olivier *et al.*, 2013).

Commercial STED and GSDIM Microscopes

Commercial STED and GSDIM microscopes are provided by Leica Microsystems, GmbH. The latest Leica TCS SP8 confocal microscope can perform STED in multiple channels with three depletion lasers. Leica SR GSD 3D is capable of three-dimensional super-resolution GSDIM imaging.

Biological Applications of STED and GSDIM

Biological Sample Preparation

For most researchers who perform super-resolution imaging, sample preparation is the single most important issue. Without properly prepared biological samples, it is impossible to acquire high quality STED and GSDIM images.

As a rule of thumb, an excellent biological sample for confocal microscopy is also excellent for STED microscopy. An ideal STED sample should be bright, photo-stable, and optically homogenous. In STED microscopy, a great part of fluorescence is suppressed and the image intensity is relatively low. In addition, STED photobleaching is fast due to high depletion power. Therefore, bright and photo-stable fluorescence dyes are highly desired. The choice of STED dyes is based on experiences. Popular STED dyes include Atto 647 N, Oregon Green 488, Chromeo 494, Atto 532, etc. A more complete list of STED dyes covering the entire visible spectrum is given in Reference Muller *et al.* (2012). STED imaging with fluorescent proteins (Willig *et al.*, 2006) is more difficult as

they are less bright and less stable than organic dyes. However, fluorescent proteins have the irreplaceable advantage of being much smaller in size compared to antibodies used for immunolabeling (The size of a typical fluorescent protein is ~ 4 nm. Antibodies are larger, for example, IgG is ~ 10 nm in length (Ehmann *et al.*, 2014). Conventional immunocytochemistry needs both the primary and the secondary antibody to label a biomacromolecule.) and being highly compatible with living cells (For organic dyes to work with live imaging, they need to somehow enter the living cell (e.g., uptake of dye-labeled lipid (Eggeling *et al.*, 2009))). Therefore, STED microscopy using fluorescent proteins can in principle reach higher resolution and is more versatile in live STED imaging than immunocytochemistry. Excellent optical homogeneity implies that the refractive index of the mounting media and immersion oil should be as close as possible to glass. The ProLongTM anti-fade reagent from Life Technologies, Inc. and 2,2'-thiodiethanol (TDE) are widely used as mounting media for STED imaging. Also, cover slips with excellent flatness and a thickness compatible to the microscope objective should be used. Scattering of the excitation and the depletion light by the specimen is more likely to compromise resolution at a higher imaging depth. Therefore, one should prepare the sample so that the area of interest is near the cover slip for best resolution.

Preparing samples for GSDIM is trickier than for STED. In GSDIM, one would want to choose fluorescent dyes with high triplet-state yield and high probability of blinking. Alexa 647, Alexa 532 and Alexa 488 are among the most widely used GSDIM dyes. Samples are often immersed in special buffers, which use reduction and oxidation (redox) agents to manipulate blinking (Steinhauer *et al.*, 2008; Vogelsang *et al.*, 2009). Recent work indicated that GSDIM working with a simple buffer is possible (Olivier *et al.*, 2013).

In all super-resolution techniques, it is important to minimize mechanical drifting and instability for both samples and the optical setup. Ideally, the environment should be free from vibration and temperature fluctuation. It is possible to correct some mechanical drifting with computer algorithms (McGorty *et al.*, 2013), but random vibration and drifting in the optical

components always diminish resolution without an easy remedy.

STED and GSDIM Applied to Biological Research

STED microscopy and GSDIM have been widely used in biological research (Blom and Brisma, 2014; Muller *et al.*, 2012). It is not our intention to provide a comprehensive review of these applications. Instead, we will just briefly overview some of them that we hope can inspire the reader to apply super-resolution fluorescence microscopy to their studies.

Because it works best with thin samples, many successful applications of STED microscopy have been done for isolated organelles and small (thin) cultured cells. An example is given by Figure 8, which shows an early STED image taken by our custom-built STED microscope (Gardeazabal Rodriguez *et al.*, 2012) for cultured neurons where actin was immune-labeled. The STED resolution was ~ 50 nm. Some organelles, such as mitochondria and endosome, are too small for conventional light microscopy to resolve. STED microscopy was used to image voltage-dependent anion channel 1 (VDAC1) and cytochrome c oxidase's subunit 2 (Cox2) clusters in purified mouse cardiac mitochondria, which found that the two proteins had very different cluster-size distributions (Singh *et al.*, 2012). 4Pi-STED microscopy was able to resolve the cristae of mitochondrial inner membrane in intact cultured cells (Schmidt *et al.*, 2009). Isolated early endosomes from PC12 cells were imaged by STED microscopy, which found that synaptic membrane proteins such as synaptophysin and some SNARE proteins form stable microdomains during endocytic trafficking (Geumann *et al.*, 2010). The helix structure of FtsZ, a cytokinetic protein that plays an important role in bacterial division, was revealed by STED microscopy to have an irregular structure (Jennings *et al.*, 2011).

Although it prefers thin samples, STED microscopy has also been extensively used to study large (thick) cells (e.g., cardiomyocytes), tissue sections and dissected samples with thickness up to ~ 100 μ m. STED imaging of membrane endothelial cell markers (e.g., CD146) showed that after acute ischemic cardiac injury fibroblasts can quickly adopt an

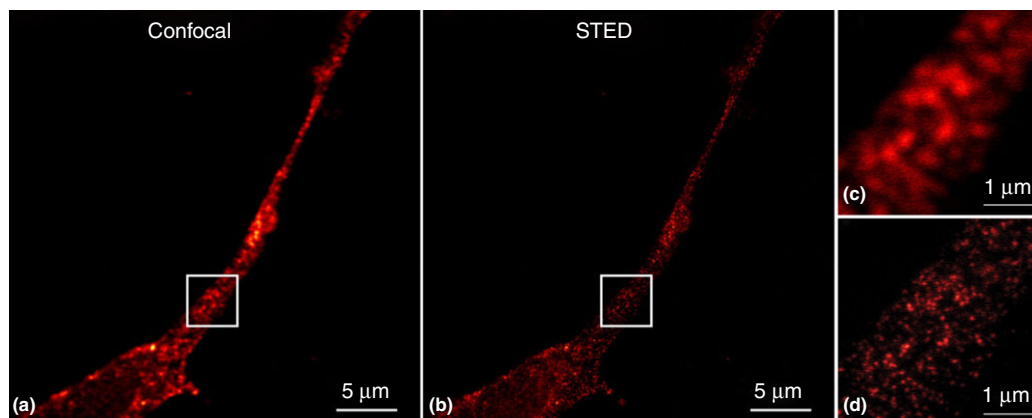


Figure 8 Comparison of conventional confocal microscopy (a) and STED microscopy (b) for neurons derived from human iPS cells where actin was immunolabeled with Atto 647 N. (c) and (d) magnify the enlarged boxes in (a) and (b), respectively. In (d), STED shows an axon with a punctuated actin distribution; whereas in the confocal image (c) there are no discrete clusters. Shown are raw images without any processing.

endothelial-cell-like fate (Ubil *et al.*, 2014). It also revealed a doughnut-shaped structure of active zones in larval *Drosophila* neuromuscular junctions, which were immunolabeled by an antibody specifically recognizing the Bruchpilot protein (Kittel *et al.*, 2006). The spatial distribution of histone H3 in neonatal rat ventricular myocytes was visualized by STED microscopy to study the chromatin structure (Mitchell-Jordan *et al.*, 2012). Sub-diffraction-limit imaging of the eGFP-tagged caveolin in Chinese hamster ovary cells (Li *et al.*, 2009) was one of the rare applications of 2PE-STED microscopy to biological samples.

Multicolor fluorescence microscopy and colocalization analysis are widely used in cell biology to study interactions among biomacromolecules. The accuracy of colocalization analysis is limited by the optical resolution of the microscope (Wu *et al.*, 2010). 'Protein *a* is colocalized to protein *b*' actually means that 'we tend to see protein *a* appears within a distance of *D* from protein *b*,' where *D* is at best the optical resolution of the microscope. It is obvious that by reducing this distance *D* from 200–300 nm to 20–60 nm with super-resolution microscopy greatly improve the reliability of colocalization analysis. In Figure 9 we give a concrete example of colocalization analysis in dual-color STED microscopy. We used our custom-built two-color STED microscopy (Wu *et al.*, 2015) to image an isolated cardiomyocyte, where L-type voltage-dependent calcium channel $\alpha 1C$ subunits (Cav1.2, red) and ryanodine

receptors (RyR, green) were independently immunolabeled by Atto 647N and Oregon green 488, respectively. In a small region as shown in Figures 9(c) and (d), the protein-proximity-index (PPI) (Wu *et al.*, 2010) measured by conventional confocal microscopy is 0.79; whereas the STED PPI value drastically drops to 0.38.

Dual-color STED microscopy was used to demonstrate and quantify the colocalization between proliferating cell nuclear antigen (PCNA) and replication protein A (RPA) in a human cell line to study DNA replication (Csereeny *et al.*, 2009). It was also used to quantify the colocalization between hexokinase-I, an important enzyme for phosphorylation, and VDAC in a human cell line. It was found that hexokinase-I has different degrees of colocalization with the three VDAC isoforms (Neumann *et al.*, 2010). At super-resolution, traditional colocalization analysis methods may not work, as the overlap between the signals in two channels greatly diminishes. Blom *et al.* used nearest neighbor analysis together with and dual-color STED microscopy to quantify the spatial density and association of dopamine receptor D1 and Na⁺/K⁺-ATPase in rat dendritic spines (Blom *et al.*, 2012).

STED microscopy has been successfully applied to image living organisms. Dendritic spines in living hippocampal slices from YFP-transgenic mice were visualized by a STED microscope using an optical parametric oscillator pumped

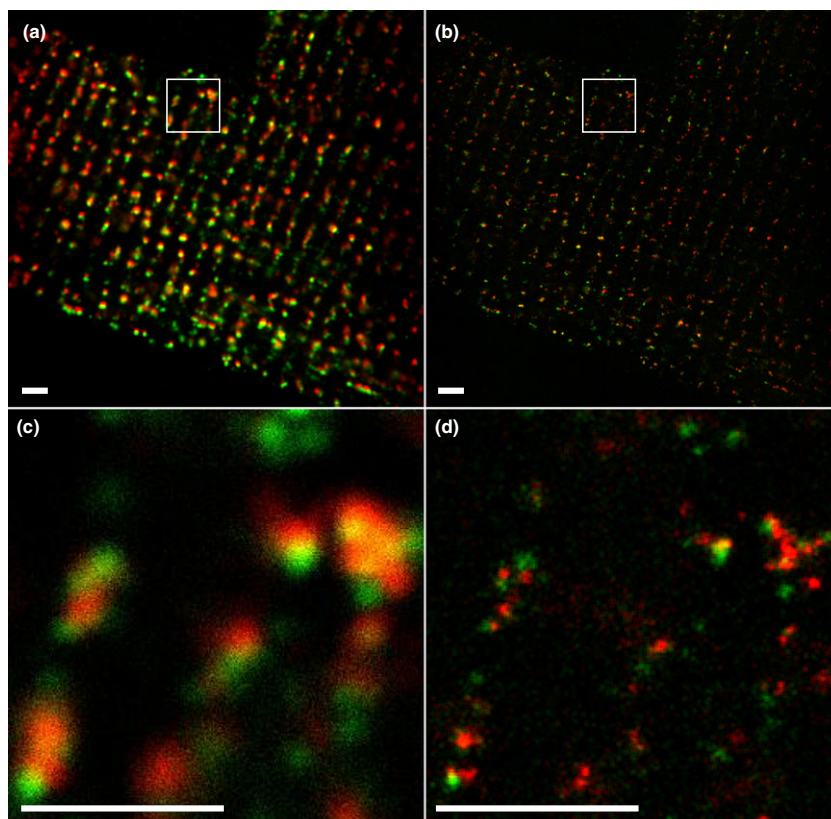


Figure 9 Dual-color STED images taken by our customized-built STED microscope. (a) Overlay of confocal images of an isolated cardiomyocyte where Cav1.2 (red) and RyR (green) were independently labeled. Field of view (FOV) is $32 \times 32 \mu\text{m}$. (b) Overlay of STED image in the same FOV as in A. (c) and (d) Blowup of a region (white square) in (a) and (b) respectively. Images were thresholded and denoised by a 3×3 median filter. The scale bars represent $2 \mu\text{m}$.

by a superfast Ti:Sapphire laser as the depletion source (Nagerl and Bonhoeffer, 2010). In a similar investigation, STED microscopy was able to reach up to 80 μm deep into the living hippocampal slice while maintaining <75 nm optical resolution (Urban *et al.*, 2011). The citrine-tagged (citrine is a variant of YFP that has higher photo-stability) endoplasmic reticulum in living Ptk2 cells was imaged with a 3D-STED microscope using two depletion beams (for the lateral and the axial direction, respectively) (Hein *et al.*, 2008). Video-rate STED microscopy using fast resonant scanners can visualize the living organisms in real time. STED microscopy using a 16 KHz resonant scanner was able to resolve single synaptic vesicles (immunolabeled by Atto 647 N) in cultured rat hippocampal neurons and record their movement at 28 fps (Westphal *et al.*, 2008). Besides fluorescent proteins and organic dyes, 'self-labeling' proteins have also been used for live STED imaging (Fitzpatrick *et al.*, 2009; Schroder *et al.*, 2009). Combined with fluorescence correlation spectroscopy, STED microscopy was able to directly observe the dynamics of fluorescently-labeled membrane lipids in living cultured cells (Eggeling *et al.*, 2009).

GSDIM (dSTORM) is most widely used to image fixed samples. For example, it resolved the distribution of Bruchpilot in active zone cytomatrix (Ehmann *et al.*, 2014) and human leukocyte antigen in natural killer cells (Pageon *et al.*, 2013). Dual-color GSDIM was used to visualize posttranslational tubulin isoforms in MDCK cells (Zink *et al.*, 2012) and to study extracellular vesicle secretion by the astrocyte (Hajj *et al.*, 2013). It was also demonstrated that GSDIM can be used to image living cell culture (Testa *et al.*, 2010). GSDIM (dSTORM) is a stochastic localization-based super-resolution technique that is very similar to PALM/STORM, which is covered elsewhere in this book. Therefore, we are satisfied by giving the above examples for its biological applications and refer the reader to other publications (Henriques *et al.*, 2011; Sengupta *et al.*, 2014) for more thorough reviews.

Conclusion

STED and GSD are two physical mechanisms that can switch common fluorescent molecules between an on-state and an off-state. Fluorescent molecules are temporally separated by this on-off switching, which results in super-resolution. STED microscopy is a point-scanning technology, in which a spatially-modulated focused light is used to sharpen the scanning focal spot to enhance resolution. In GSDIM, fluorescent molecules randomly return from the dominant dark triplet-state to the ground state, from which the single-molecule emission events are separately recorded and the precise location of the fluorescent emitters is numerically determined. STED microscopy has been implemented in various ways, including pulsed STED, CW STED, and time-gated STED. 3D-STED and 4Pi-STED extend super-resolution to three dimensions. 2PE-STED helps penetrate deeper into the sample. Resonant scanning reduces STED photobleaching and realizes video-rate STED imaging. The big family of technologies has enabled STED and GSDIM microscopy to be extensively applied to biological research.

Acknowledgment

This contribution was supported by NIH RO1 HL088640 (ES) and NIH RO1 HL107418 (ES & LT).

See also: Imaging the Cell: Light Microscopy: Super Resolution Fluorescence Localization Microscopy

References

- Betzig, E., Patterson, G.H., Sougrat, R., *et al.*, 2006. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313 (5793), 1642–1645.
- Bianchini, P., Harke, B., Galiani, S., Vicidomini, G., Diaspro, A., 2012. Single-wavelength two-photon excitation-stimulated emission depletion (SW2PE-STED) superresolution imaging. *Proceedings of the National Academy of Sciences of the United States of America* 109 (17), 6390–6393.
- Blom, H., Brisma, H., 2014. STED microscopy: Increased resolution for medical research? *Journal of Internal Medicine* 276 (6), 560–578.
- Blom, H., Ronnlund, D., Scott, L., *et al.*, 2012. Nearest neighbor analysis of dopamine D1 receptors and Na⁺ plus -K⁺ plus -ATPases in dendritic spines dissected by STED microscopy. *Microscopy Research and Technique* 75 (2), 220–228.
- Bokor, N., Iketaki, Y., Watanabe, T., *et al.*, 2007. On polarization effects in fluorescence depletion microscopy. *Optics Communications* 272 (1), 263–268.
- Bretschneider, S., Eggeling, C., Hell, S.W., 2007. Breaking the diffraction barrier in fluorescence microscopy by optical shelving. *Physical Review Letters* 98, 21.
- Courtial, J., O'Holleran, K., 2007. Experiments with twisted light. *European Physical Journal-Special Topics* 145, 35–47.
- Cseresnyes, Z., Schwarz, U., Green, C.M., 2009. Analysis of replication factories in human cells by super-resolution light microscopy. *BMC Cell Biology* 10, 88.
- Dertinger, T., Colyer, R., Iyer, G., Weiss, S., Enderlein, J., 2009. Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI). *Proceedings of the National Academy of Sciences of the United States of America* 106 (52), 22287–22292.
- Deschenes, L.A., Bout, D.A.V., 2002. Single molecule photobleaching: Increasing photon yield and survival time through suppression of two-step photolysis. *Chemical Physics Letters* 365 (5–6), 387–395.
- Donnert, G., Eggeling, C., Hell, S.W., 2007. Major signal increase in fluorescence microscopy through dark-state relaxation. *Nature Methods* 4 (1), 81–86.
- Donnert, G., Eggeling, C., Hell, S.W., 2009. Triplet-relaxation microscopy with bunched pulsed excitation. *Photochemical & Photobiological Sciences* 8 (4), 481–485.
- Donnert, G., Keller, J., Wurm, C.A., *et al.*, 2007. Two-color far-field fluorescence nanoscopy. *Biophysical Journal* 92 (8), L67–L69.
- Dyba, M., Hell, S.W., 2003. Photostability of a fluorescent marker under pulsed excited-state depletion through stimulated emission. *Applied Optics* 42 (25), 5123–5129.
- Eggeling, C., Ringemann, C., Medda, R., *et al.*, 2009. Direct observation of the nanoscale dynamics of membrane lipids in a living cell. *Nature* 457 (7233), 1159–1162.
- Ehmann, N., van de Linde, S., Alon, A., *et al.*, 2014. Quantitative super-resolution imaging of Bruchpilot distinguishes active zone states. *Nature Communications* 5, 4650.
- Fitzpatrick, J.A.J., Yan, Q., Sieber, J.J., *et al.*, 2009. STED nanoscopy in living cells using fluorogen activating proteins. *Bioconjugate Chemistry* 20 (10), 1843–1847.
- Folling, J., Bossi, M., Bock, H., *et al.*, 2008. Fluorescence nanoscopy by ground-state depletion and single-molecule return. *Nature Methods* 5 (11), 943–945.
- Gardeazabal Rodriguez, P.F., Wu, Y., Singh, R., *et al.*, 2012. Building a fast scanning stimulated emission depletion microscope: A step by step guide. In: Mendez-Vilas, A. (Ed.), *Current Microscopy Contributions to Advances in Science and Technology*, vol. 2. Formatex Research Center, pp. 791–800.
- Geumann, U., Schafer, C., Riedel, D., Jahn, R., Rizzoli, S.O., 2010. Synaptic membrane proteins form stable microdomains in early endosomes. *Microscopy Research and Technique* 73 (6), 606–617.
- Gustafsson, M.G., 2000. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy* 198 (Pt 2), 82–87.
- Gustafsson, M.G.L., 2005. Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences of the United States of America* 102 (37), 13081–13086.

- Hajji, G.N.M., Arantes, C.P., Dias, M.V.S., *et al.*, 2013. The unconventional secretion of stress-inducible protein 1 by a heterogeneous population of extracellular vesicles. *Cellular and Molecular Life Sciences* 70 (17), 3211–3227.
- Han, K.Y., Kim, S.K., Eggeling, C., Hell, S.W., 2010. Metastable dark states enable ground state depletion microscopy of nitrogen vacancy centers in diamond with diffraction-unlimited resolution. *Nano Letters* 10 (8), 3199–3203.
- Hao, X.A., Kuang, C.F., Wang, T.T., Liu, X., 2010. Effects of polarization on the de-excitation dark focal spot in STED microscopy. *Journal of Optics* 12, 11.
- Harke, B., Keller, J., Ullal, C.K., *et al.*, 2008. Resolution scaling in STED microscopy. *Optics Express* 16 (6), 4154–4162.
- Heilemann, M., Linde, S., Mukherjee, A., Sauer, M., 2009. Super-resolution imaging with small organic fluorophores. *Angewandte Chemie-International Edition* 48 (37), 6903–6908.
- Heilemann, M., van de Linde, S., Schuttpelz, M., *et al.*, 2008. Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angewandte Chemie-International Edition* 47 (33), 6172–6176.
- Hein, B., Willig, K.I., Hell, S.W., 2008. Stimulated emission depletion (STED) nanoscopy of a fluorescent protein-labeled organelle inside a living cell. *Proceedings of the National Academy of Sciences of the United States of America* 105 (38), 14271–14276.
- Hell, S., Stelzer, E.H.K., 1992. Properties of a 4Pi confocal fluorescence microscope. *Journal of the Optical Society of America A-Optics Image Science and Vision* 9 (12), 2159–2166.
- Hell, S.W., 2003. Toward fluorescence nanoscopy. *Nature Biotechnology* 21 (11), 1347–1355.
- Hell, S.W., Kroug, M., 1995. Ground-state-depletion fluorescence microscopy — A concept for breaking the diffraction resolution limit. *Applied Physics B-Lasers and Optics* 60 (5), 495–497.
- Henriques, R., Griffiths, C., Rego, E.H., Mhlanga, M.M., 2011. PALM and STORM: Unlocking live-cell super-resolution. *Biopolymers* 95 (5), 322–331.
- Huang, B., Wang, W.Q., Bates, M., Zhuang, X.W., 2008. Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. *Science* 319 (5864), 810–813.
- Inoue, S., 1995. Foundations of confocal scanned imaging in light microscopy. In: Pawley, J. (Ed.), *Handbook of Biological Confocal Microscopy*. New York, NY: Springer, pp. 1–17.
- Jennings, P.C., Cox, G.C., Monahan, L.G., Harry, E.J., 2011. Super-resolution imaging of the bacterial cytokinetic protein FtsZ. *Micron* 42 (4), 336–341.
- Keller, J., Schonle, A., Hell, S.W., 2007. Efficient fluorescence inhibition patterns for RESOLFT microscopy. *Optics Express* 15 (6), 3361–3371.
- Kittel, R.J., Wichmann, C., Rasse, T.M., *et al.*, 2006. Bruchpilot promotes active zone assembly, Ca²⁺ channel clustering, and vesicle release. *Science* 312 (5776), 1051–1054.
- Klar, T.A., Jakobs, S., Dyba, M., Egner, A., Hell, S.W., 2000. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proceedings of the National Academy of Sciences of the United States of America* 97 (15), 8206–8210.
- Li, Q.F., Wu, S.S.H., Chou, K.C., 2009. Subdiffraction-limit two-photon fluorescence microscopy for GFP-tagged cell imaging. *Biophysical Journal* 97 (12), 3224–3228.
- McGorty, R., Kamiyama, D., Huang, B., 2013. Active microscope stabilization in three dimensions using image correlation. *Optical Nanoscopy* 2, 3.
- Mitchell-Jordan, S., Chen, H.D., Franklin, S., *et al.*, 2012. Features of endogenous cardiomyocyte chromatin revealed by super-resolution STED microscopy. *Journal of Molecular and Cellular Cardiology* 53 (4), 552–558.
- Moneron, G., Hell, S.W., 2009. Two-photon excitation STED microscopy. *Optics Express* 17 (17), 14567–14573.
- Muller, T., Schumann, C., Kraegeloh, A., 2012. STED microscopy and its applications: New insights into cellular processes on the nanoscale. *Chemphyschem* 13 (8), 1986–2000.
- Nagerl, U.V., Bonhoeffer, T., 2010. Imaging living synapses at the nanoscale by STED microscopy. *Journal of Neuroscience* 30 (28), 9341–9346.
- Neumann, D., Buckers, J., Kastrop, L., Hell, S., Jakobs, S., 2010. Two-color STED microscopy reveals different degrees of colocalization between hexokinase-I and the three human VDAC isoforms. *PMC Biophysics* 3 (1), 1–15.
- Oliver, N., Keller, D., Rajan, V.S., Gonczy, P., Manley, S., 2013. Simple buffers for 3D STORM microscopy. *Biomedical Optics Express* 4 (6), 885–899.
- Paggon, S.V., Cordoba, S.P., Owen, D.M., *et al.*, 2013. Superresolution microscopy reveals nanometer-scale reorganization of inhibitory natural killer cell receptors upon activation of NKG2D. *Science Signaling* 6 (285), ra62.
- Patterson, G.H., Piston, D.W., 2000. Photobleaching in two-photon excitation microscopy. *Biophysical Journal* 78 (4), 2159–2162.
- Rankin, B.R., Hell, S.W., 2009. STED microscopy with a MHz pulsed stimulated-Raman-scattering source. *Optics Express* 17 (18), 15679–15684.
- Rust, M.J., Bates, M., Zhuang, X.W., 2006. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods* 3 (10), 793–795.
- Schmidt, R., Wurm, C.A., Jakobs, S., *et al.*, 2008. Spherical nanosized focal spot unravels the interior of cells. *Nature Methods* 5 (6), 539–544.
- Schmidt, R., Wurm, C.A., Punge, A., *et al.*, 2009. Mitochondrial cristae revealed with focused light. *Nano Letters* 9 (6), 2508–2510.
- Schroder, J., Benink, H., Dyba, M., Los, G.V., 2009. In vivo labeling method using a genetic construct for nanoscale resolution microscopy. *Biophysical Journal* 96 (1), L1–L3.
- Sengupta, P., van Engelenburg, S.B., Lippincott-Schwartz, J., 2014. Superresolution imaging of biological systems using photoactivated localization microscopy. *Chemical Reviews* 114 (6), 3189–3202.
- Singh, H., Lu, R., Rodriguez, P.F.G., *et al.*, 2012. Visualization and quantification of cardiac mitochondrial protein clusters with STED microscopy. *Mitochondrion* 12 (2), 230–236.
- Steinhauer, C., Forthmann, C., Vogelsang, J., Tinnefeld, P., 2008. Superresolution microscopy on the basis of engineered dark states. *Journal of the American Chemical Society* 130 (50), 16840.
- Testa, I., Wurm, C.A., Medda, R., *et al.*, 2010. Multicolor fluorescence nanoscopy in fixed and living cells by exciting conventional fluorophores with a single wavelength. *Biophysical Journal* 99 (8), 2686–2694.
- Ubil, E., Duan, J., Pillai, I.C.L., *et al.*, 2014. Mesenchymal-endothelial transition contributes to cardiac neovascularization. *Nature* 514 (7524), 585–590.
- Urban, N.T., Willig, K.I., Hell, S.W., Nagerl, U.V., 2011. STED nanoscopy of actin dynamics in synapses deep inside living brain slices. *Biophysical Journal* 101 (5), 1277–1284.
- Vicidomini, G., Moneron, G., Han, K.Y., *et al.*, 2011. Sharper low-power STED nanoscopy by time gating. *Nature Methods* 8 (7), 571–573.
- Vogelsang, J., Cordes, T., Forthmann, C., Steinhauer, C., Tinnefeld, P., 2009. Controlling the fluorescence of ordinary oxazine dyes for single-molecule switching and superresolution microscopy. *Proceedings of the National Academy of Sciences of the United States of America* 106 (20), 8107–8112.
- Vogelsang, J., Kasper, R., Steinhauer, C., *et al.*, 2008. A reducing and oxidizing system minimizes photobleaching and blinking of fluorescent dyes. *Angewandte Chemie-International Edition* 47 (29), 5465–5469.
- Westphal, V., Lauterbach, M.A., Di Nicola, A., Hell, S.W., 2007. Dynamic far-field fluorescence nanoscopy. *New Journal of Physics* 9, pii: pdb.prot069260.
- Westphal, V., Rizzoli, S.O., Lauterbach, M.A., *et al.*, 2008. Video-rate far-field optical nanoscopy dissects synaptic vesicle movement. *Science* 320 (5873), 246–249.
- Wildanger, D., Medda, R., Kastrop, L., Hell, S.W., 2009. A compact STED microscope providing 3D nanoscale resolution. *Journal of Microscopy-Oxford* 236 (1), 35–43.
- Willig, K.I., Harke, B., Medda, R., Hell, S.W., 2007. STED microscopy with continuous wave beams. *Nature Methods* 4 (11), 915–918.
- Willig, K.I., Kellner, R.R., Medda, R., *et al.*, 2006. Nanoscale resolution in GFP-based microscopy. *Nature Methods* 3 (9), 721–723.
- Wu, X., Toro, L., Stefani, E., Wu, Y., 2015. Ultrafast photon counting applied to resonant scanning STED microscopy. *Journal of Microscopy* 257 (1), 31–38.
- Wu, Y., Eghbali, M., Ou, J., *et al.*, 2010. Quantitative determination of spatial protein-protein correlations in fluorescence confocal microscopy. *Biophysical Journal* 98 (3), 493–504.
- Wu, Y., Wu, X., Lu, R., *et al.*, 2014. Resonant scanning with large field of view reduces photobleaching and enhances fluorescence yield in STED microscopy. *arXiv:1408.1507*. Available at: <http://arxiv.org/abs/1408.1507> (accessed 04.02.15).
- Zink, S., Grosse, L., Freikamp, A., *et al.*, 2012. Tubulin dephosphorylation promotes monolayer formation and apical trafficking in epithelial cells. *Journal of Cell Science* 125 (24), 5998–6008.

Relevant Websites

- <http://www.leica-microsystems.com/science-lab/quick-guide-to-sted-sample-preparation/>
Leica Science Lab.
- <http://www.leica-microsystems.com/science-lab/sample-preparation-for-gsdm-localization-microscopy-protocols-and-tips/>
Leica Science Lab.

Structured Illumination Microscopy

V Mennella, University of Toronto, Toronto, ON, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The recent advances in high-throughput genomic and proteomic analysis have helped compile the extensive, but finite list of components that define the function of a cell. A critical task ahead is to integrate this information into molecular assemblies and understand how their structure influences activity in cells (Alberts, 1998). Especially in cell biology, it is of fundamental importance to examine the function and biophysical properties of macromolecular assemblies in conditions closest to their native environment. To this end, fluorescence microscopy is a useful method because it provides molecular specificity through antibody labeling or genetically encoded tags; it is generally efficient at labeling and preserving ultrastructure; and importantly, it is compatible with live cell imaging.

However, fluorescence microscopy was considered until recently to be subject to the hard limit of the light diffraction barrier (Abbe, 1874; Heintzmann and Ficz, 2007, 2013). This theoretical limit was first established by Abbe in 1874 and states that resolution in an optical device is directly proportional to the wavelength of the observation and inversely proportional to the numerical aperture of the objective ($d = \lambda/2NA$; Abbe, 1874; Heintzmann and Ficz, 2007, 2013). In effect, in the best of circumstances, a microscopist can obtain an anisotropic resolution of ~ 250 nm in the in-plane (x - y) direction and ~ 500 nm in the axial (z) direction, when using a fluorescence microscope with a single, high numerical aperture (1.4 NA) objective, and when detecting fluorescence emission around 520 nm. The lack of tools capable of surpassing this optical limit has determined a substantial gap in our knowledge of cell organization, because in cells most macromolecules are organized in assemblies with sizes in the range of few nanometers.

Recently, sub-diffraction fluorescence microscopy methods – often referred to with the less precise but more effective name as super-resolution microscopies – have emerged as powerful tools in biology (Huang *et al.*, 2009; Hell, 2007a; Schermelleh *et al.*, 2010). When compared to conventional fluorescence microscopies (e.g., confocal and widefield microscopy), super-resolution microscopy methods provide an increase in resolving power from 2- to 10-fold (from 125 nm to ~ 10 –20 nm in the x - y plane). Super-resolution methods can surpass the diffraction barrier in two ways: by taking advantage of spatial patterning of the excitation illumination (as in SIM, RESOLFT, and STED) or by detecting single molecules separated in space and time to enable high-resolution localization (as in PALM, STORM, GSD, and related techniques). Since super-resolution methods differ in their principles, resolution capacity, and experimental flexibility, they provide different advantages and disadvantages to the experimental biologist (Huang *et al.*, 2009; Schermelleh *et al.*, 2010; Hell, 2007b).

In this article we present the principles and applications of structured illumination microscopy (SIM). SIM relies on

modulation of the excitation light by spatial patterning to break the diffraction barrier (Gustafsson, 2000, 2008). In its most practical implementation, SIM provides a relatively small resolution gain of a factor of two (Figure 1(a)), yet has the least stringent requirement for fluorophores selection and readily accommodates multiple labeling. This makes SIM the method of choice for multicolor labeling of cellular structures and high-throughput imaging due to its flexibility and relatively ease of sample optimization (Huang *et al.*, 2009; Schermelleh *et al.*, 2008a, 2010).

In the sections below, we first provide an introduction to the principles of linear 3DSIM, which is currently the most common implementation of this technology. We subsequently detail the impact of SIM on different fields of cell biology. We then conclude the article with the most recent developments in SIM imaging (e.g., 2D TIRF-SIM, nonlinear SIM illumination, Bessel-beam SIM, and multifocal SIM), which are leading to better live cell imaging capabilities, to an increase in resolution power and to more effective quantitative analysis.

Principles of SIM

SIM is a method that relies on a modified widefield microscope to collect light from a fluorescently labeled specimen. The defining characteristic of a SIM microscope is the sample illumination: while in a widefield microscope the sample is illuminated homogeneously, in SIM the sample is illuminated with a striped pattern – a ‘spatially structured light illumination.’ The purpose of this pattern is to create 3D interference fringes on the sample, which in turn reveal otherwise unobservable information about the sample fluorophore’s distribution. The net result is comparable to the moiré pattern, a phenomenon that occurs when two structural patterns are superimposed and features of the sample spatial frequencies are revealed that were previously unobservable under a homogeneous illumination pattern (Figure 1(b); Gustafsson *et al.*, 2008; Schermelleh *et al.*, 2008b; Gustafsson, 2000).

In a 3DSIM microscope, the sample illumination with a structured illumination pattern accomplishes two important objectives: (1) it shifts high-frequency sample information not normally observable by the microscope to the range of frequencies the microscope can observe (Figures 1(a) and 1(b)). In turn, these added frequencies lead to an increase in resolution of a factor of two both in the lateral (~ 125 nm) and axial directions (~ 250 nm); (2) it provides 3D optical sectioning by rejecting light from fluorescent objects that are not in the imaging focal plane by axial modulation of the illumination pattern.

Resolution Increase in SIM

To understand the principle of SIM it is useful to introduce the concept of Fourier transform of an image. When a point light

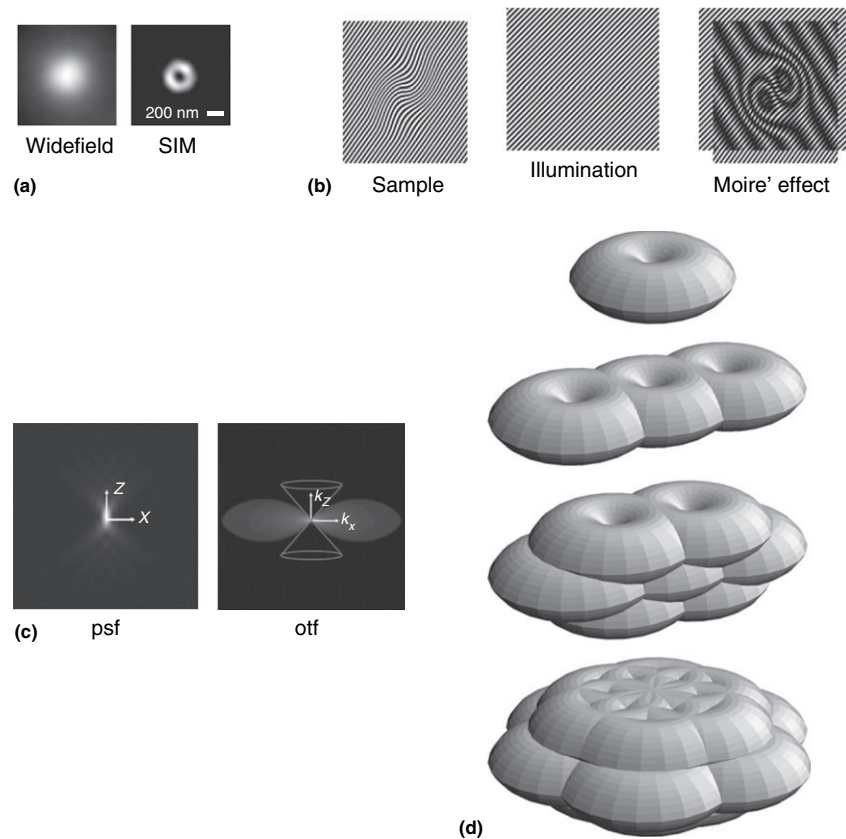


Figure 1 Principles of structured illumination microscopy. (a) Resolution increase in SIM. Widefield or 3DSIM microscopy maximum-intensity projection images along the z axis of centrioles from S2 cells labeled with antibodies against centriolar proteins. Scale bar, 200 nm. Adapted from Mennella, V., Keszthelyi, B., McDonald, K.L., *et al.*, 2012. Subdiffraction-resolution fluorescence microscopy reveals a domain of the centrosome critical for pericentriolar material organization. *Nature Cell Biology* 14 (11), 1159–1168. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23086239> (accessed 29.05.13). (b) Visualization of the moiré effect. Adapted from Gustafsson, M.G., 2000. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy* 198 (Pt 2), 82–87. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10810003> (accessed 01.04.15). (c) Cross section of a point spread function and optical transfer function of a widefield microscope. Note that the OTF image in (c) is a section of the volume shown in (d). Adapted from Fiolka, R., 2014. Seeing More with Structured Illumination Microscopy, first ed. Elsevier Inc. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24974034> (accessed 15.07.14). (d) Observable region in Fourier space of a widefield microscope, of a structured illumination experiment of a 2D image with three phases, in a structured illumination experiment of a 3D image with five phases and three angles rotations. Adapted from Gustafsson, M.G.L., Shao, L., Carlton, P.M., *et al.*, 2008. Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophysical Journal* 94 (12), 4957–4970. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2397368&tool=pmcentrez&rendertype=abstract> (accessed 02.11.12).

source is detected through a microscope, its image is degraded by the optical apparatus (lenses, objective, etc.). The result is a ‘real space’ blurred image that describes the optical properties of our specific microscope setup, which is referred to as the point spread function (PSF). The Fourier transform of the PSF is named the optical transfer function (OTF) (Figure 1(c); Pawley, 2006; Gustafsson *et al.*, 2008; Gustafsson, 2000; Schermelleh *et al.*, 2010). The OTF as any other image in ‘Fourier space,’ is represented by many sinusoidal curves with different amplitude and frequencies, which correspond to the intensities of each pixel in the ‘real space’ image. In ‘Fourier space,’ the limit of resolution defined by Abbe can be visualized by examining the OTF zone of support, a toroid region including all the spatial frequencies in which the amplitude is above zero – low frequencies lie close to the origin and high frequencies away from the zone (Figure 1(d)).

When we illuminate the sample with a structured illumination pattern we are effectively increasing the observable region in our microscope – the toroid of the OTF zone of support – to higher frequency areas that are normally not detectable by a conventional fluorescence microscope. In the real image those frequencies represent the features of the sample revealed by the moiré pattern (Figure 1(b); Gustafsson *et al.*, 2008; Gustafsson, 2000).

For example, in the case of a single 2D image acquisition in SIM, when we superpose the sample fluorophore distribution with the structured illumination pattern, we obtain the same original image summed with two other images shifted at two different frequency centers (Figures 1(d) and 2(d)). The new image detected is now a mix of information components from the normally observable frequencies of a widefield microscope and the new frequencies originated as in a moiré pattern (Figure 2(d)). The fundamental issue to solve

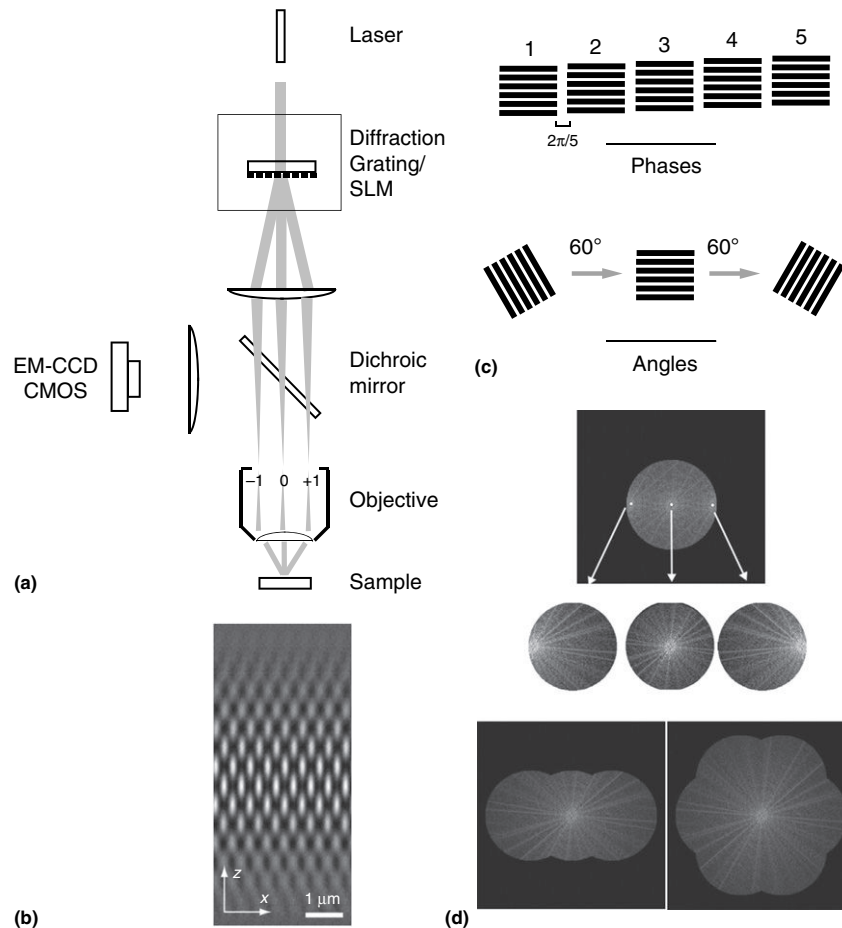


Figure 2 3DSIM microscope and data acquisition. (a) Diagram of the basic components of a 3DSIM microscope. The illumination path is highlighted in gray. (b) Image of the structured illumination pattern on the sample plane with lateral and axial structure. Adapted from Gustafsson, M.G.L., Shao, L., Carlton, P.M., *et al.*, 2008. Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophysical Journal* 94 (12), 4957–4970. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2397368&tool=pmcentrez&rendertype=abstract> (accessed 02.11.12). (c) Diagram of the 3DSIM acquisition strategy. Five phases of the sine wave pattern are recorded at each z position and three image stacks are recorded with the diffraction grating sequentially rotated into three positions 60° apart. Adapted from Schermelleh, L., Carlton, P.M., Haase, S., *et al.*, 2008a. Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* (New York, NY) 320 (5881), 1332–1336. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2916659&tool=pmcentrez&rendertype=abstract> (accessed 30.04.14); Schermelleh, L., Carlton, P.M., Haase, S., *et al.*, 2008b. Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* (New York, NY) 320 (5881), 1332–1336. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2916659&tool=pmcentrez&rendertype=abstract> (accessed 18.11.12). Note that it is possible to modify this acquisition scheme to acquire phases and angles of each z position. (d) Top. Fourier transform of the product of the illumination pattern with the fluorophore distribution. The three white dots represent the Fourier components of the illumination pattern. Convolution with these Fourier components superimposes three copies of the object spectrum into the observable region of the microscope (white arrows). Bottom, left, reconstructed observable region with enhanced support in the direction of pattern. Right, reconstructed observable region using three different orientations of the illumination pattern. Adapted from Fiolka, R., 2014. *Seeing More with Structured Illumination Microscopy*, first ed. Elsevier Inc. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24974034> (accessed 15.07.14).

at the core of the SIM microscope is to separate and correctly assign these new frequency components on the final ‘reconstructed’ image. This is possible by accumulating pixel intensity information derived from several raw images at each focal plane. In the case of a 2DSIM acquisition scheme, we collect three images, one at each of three lateral grating positions spaced by one-third of the grating’s period. Since translating the pattern by a known amount causes a known change to the phases of the sinusoidal illumination pattern, we refer to three collected images as three phases. The three phases are necessary for separating three information

components, in the same manner as solving three unknowns from three known equations. These three separated components can then be shifted back to where they originate and stitched together. The described procedure extends the resolution in one direction parallel to the image plane. To make the resolution extension more isotropic in the image plane, the procedure is repeated at three different orientations of the grating (Figure 2(c)). By imaging more than three orientations it is possible to obtain a more isotropic increase in the resolution, at the cost of additional images. In a linear 3DSIM setup, imaging three orientations allows

covering an area in Fourier space that is twice as large compared to a conventional microscope, which effectively increases our real image resolution by a factor of two. In the case of 3DSIM there are five information components within each pattern orientation – differently from the three required in 2DSIM – therefore images corresponding to five phases per orientation are collected (Figure 2(c)).

Optical Sectioning in SIM

In a widefield microscope, during image acquisition, the entire sample volume is illuminated every time the focus position is moved in the axial direction. As a consequence, each focal section emits light derived from fluorescent structures part in-focus and part out-of-focus. This type of illumination comes with benefits and costs. On one hand, light collection is very efficient, because all emitted photons from the fluorophores present in our sample are collected, and it is also fast because all the pixels are acquired simultaneously. On the other hand, in this type of illumination low spatial frequency components above and below the x - y plane of the imaged object are poorly detected because of the noise derived by out-of-focus light. This is represented clearly in Fourier space by the toroid shape of the widefield OTF, which has a depression in the middle because of these under-represented frequencies. Ultimately, this determines poor resolution in the axial direction. To improve optical sectioning (and resolution) in widefield, deconvolution algorithms are used to reject out-of-focus light by assigning out-of-focus signal back to the plane where it belongs through analysis of the OTF (Swedlow, 2013; Verdaasdonk *et al.*, 2014; Mertz, 2011).

Differently, in a linear 3DSIM microscope the interference light pattern created on the sample is used to create optical sectioning. Since light derived from in-plane structures is modulated by the interference pattern much more strongly than out-of-focus structures in our sample, we can compare several raw images of our sample illuminated with an interference light pattern and combine it with a computational algorithm to reject out-of-focus light. Because the diffraction pattern is also axially modulated the frequencies under-represented in widefield are better observed (Figure 1(d)). It is important to mention that the way optical sectioning is achieved in 3DSIM is different from another similar named technology optical-sectioning SIM (Mertz, 2011; Langhorst *et al.*, 2009). In this case, there is no axial modulation of the interference pattern, but only in-plane (Mertz, 2011; Langhorst *et al.*, 2009).

For a more detailed analysis of the principle of SIM the reader can refer to the initial theoretical contributions and the papers detailing the experimental proof of principles (Gustafsson *et al.*, 2008; Heintzmann *et al.*, 2002; Gustafsson, 2000).

Microscope Hardware, Software, and Data Acquisition

In this section we describe the main optical engineering elements and the fundamental steps of image acquisition and processing of a linear 3DSIM microscope.

Generation of the Structured Illumination Pattern

In a SIM microscope, the illumination is provided by a laser light source and delivered collimated by a fiber to the diffraction grating. This latter element is considered as the core of the SIM setup: its purpose is to generate the diffraction pattern that is then projected onto the back focal plane of the objective (Figure 2(a)). Of all the rays generated by the diffraction grating, the back aperture of the objective collects only three diffraction orders. These beam orders (-1 , 0 , $+1$, the undiffracted, and ± 1 diffraction maxima) interfere on the sample plane to produce the sinusoidal pattern that has both lateral and axial structure (Figure 2(b)). In linear 2DSIM, the 0th order is blocked, resulting in a purely lateral sinusoidal pattern with no axial structure, while in 3DSIM the 0th is unblocked. Altogether, this process can be visualized as if it produces a reduced image of the diffraction grating on the objective sample plane (Gustafsson *et al.*, 2008; Fiolka, 2014).

During SIM imaging of one plane, the grating is translated – to obtain five image phases – and rotated – to obtain three angles – for a total of 15 raw images. These raw images are essential to separate the different information components contained in the summed image, similar to the moiré pattern (Figure 2(c)). In one of the first implementation of 3DSIM a piezoelectric positioner moved the diffraction grating, but this device is relatively slow and requires long settling times (Dobbie *et al.*, 2011). In addition, a rotational stage is used to rotate the grating and the positioner for imaging the different angles. Recently, a spatial light modulator (SLM) has substituted these three elements (Kner *et al.*, 2009; Shao *et al.*, 2011). The SLM makes use of liquid crystals to generate the diffraction pattern very rapidly and reduce the need for moving mechanical parts, which in turn can introduce mechanical drift. While the illumination portion of the 3DSIM microscope is relatively complex, the detection path is very similar to any other widefield microscope and will not be discussed here in detail.

Image Preprocessing

As mentioned above, during a 3DSIM experiment a relatively high number of raw images per z stack are acquired – 5 images per each grating angle rotation for a total of 15 images for z slice. Because the images of the three angle rotations contribute equally to the final reconstructed image, it is then useful to compensate for sample photobleaching occurring during acquisition between different angles. While this correction might be already hardcoded in the reconstruction procedure, in the case of samples particularly sensitive to photobleaching ($>10\%$ between angle rotations), it might be necessary to adjust the exposure times for each angle in the acquisition scheme to obtain the best reconstruction quality. A step of background subtraction is also generally applied to the image stack before further processing, which takes care of homogeneous noise derived from sample preparation.

Band Separation/Unmixing

Once all the images derived from phase shifts and angle rotations are obtained, the SIM image stack now contains

high-frequency information summed with conventional image (Figure 2(d)). To obtain the final high-resolution 'reconstructed' image it is necessary to first separate these information components. This procedure – band separation/unmixing – is done by solving a series of linear equations. Either three (in 2D) or five independent (3D) images at different phases (3D) are necessary to separate the bands. This in turn determines the number of raw images acquired (9 or 15 for three rotation acquisitions) per individual z plane in the stack (Gustafsson *et al.*, 2008; Carlton, 2008; Righolt *et al.*, 2013, 2014). The next step is to estimate a scale factor that describes the contrast of each high-resolution band compared with the conventional band, a procedure that is referred to as parameter fitting (Figure 2(d)) and takes advantage of the overlapping region between the bands.

Inverse Wiener Filtering and Apodization

After band separation and parameter fitting, the bands are assigned to the correct locations in Fourier space and a generalized Wiener filter is applied. This is done by dividing the separate bands of the image in Fourier space by the OTF at the corresponding frequencies plus a constant (the Wiener filter). This constant is important because it prevents noise being amplified especially near the high-resolution region where the OTF value is low. The value of this Wiener constant is empirically determined: when it is set too high, it can have the effect of blurring the image by removing the fine details present at high spatial frequencies, while if it is too low, then low-frequency noise and honeycomb pattern can appear in the reconstructed image (Righolt *et al.*, 2013, 2014).

In the reconstruction algorithm, an apodization filter is applied to reduce the sharp transition to zero of frequencies near the edge of the observable region. The apodization step reduces ringing effects and ultimately improves the dynamic range of the reconstructed image. Since the parameters of the apodization function are empirically determined, there is an interest by several groups to better optimize their values to increase the robustness of the SIM processing algorithm (Righolt *et al.*, 2013, 2014). At last, after filtering the different bands corresponding to the high and low spatial frequencies the images are then assembled into a volume that has twice the resolution.

Imaging Artifacts: Causes and Manifestations

In this section we discuss the most common causes of artifacts and their manifestations in a linear 3DSIM experiment. While in certain cases the presence of some stereotypic image features can point to errors in particular steps during image acquisition and processing, it is important to underscore that multiple hardware and processing problems can cause a similar artifact in the final image.

Image Drift

In a 3DSIM experiment, sample movement is one of the most common causes of failure in obtaining a reconstructed image. The movement can be caused by the biological phenomenon

of interest, as in a live cell imaging experiment, or by mechanical or thermal drift. The SIM microscope has a certain tolerance to sample movement, which is directly related to its resolution. As a general rule, in a linear 3DSIM the fluorescence structure that is being imaged should not move a distance larger than the microscope resolution (~ 100 nm) during a single stack acquisition (Shao *et al.*, 2011; Kner *et al.*, 2009). This tolerance can be even larger in particular circumstances as discussed below (see section on live SIM). The manifestation of drift in the reconstructed image is the formation of a striped pattern and/or the formation of a deformed object due to the incorrect position assignment of the different sidebands.

Refractive Index Mismatch

When imaging deep into tissues or in an area in the cell where there is a local change in refractive index (i.e., around membrane regions), there is a significant scattering of light that can cause a reduction of the image signal to noise. The scattering can be accentuated by a mismatch of the refractive index between our sample, the imaging medium, and objective immersion material. This is a general problem that affects all microscopy techniques, but it has a more pronounced effect in widefield-based microscopes as SIM, where not only the fluorescence emission can be significantly altered by illumination from other focal planes, but also the quality of the SIM illumination pattern can be affected (Ball *et al.*, 2012; Gao *et al.*, 2014). In practice, in most experiments the refractive index mismatch becomes relevant only when imaging structures a few micrometers away from the coverslip. When present, refractive index mismatch results in the formation of an asymmetrical psf and can be easily detected by looking at x/z or y/z sections of a bead embedded in our sample and subsequently corrected by using oil of matching refractive index (Wallace *et al.*, 2001; Ball *et al.*, 2012).

Low Signal to Noise

The imaging of a fluorescent sample with high signal-to-noise is desirable in many types of microscopy, and SIM is no exception, but it is not always an option. When dealing with poor signal to noise it is important to understand its origin: it can be an inherent property of our sample – which is a complex problem to solve – or alternatively can be a result of the degradation of the sample modulation contrast in the interference pattern. This type of signal degradation can be caused, for example, by hardware misalignment or by light scattering present when imaging samples away from the coverslip. The problem of low modulation contrast has limited the application of linear 3DSIM to tissue specimens to depths not larger than 10–15 μm . As we will discuss in the Section Recent Developments, new SIM strategies are aimed at imaging samples at larger depths. When imaging samples with low signal to noise, the overall effect is generally a catastrophic failure of the reconstruction.

Pixel Negative Values

When imperfectly reconstructed, SIM images can show negative pixel values, especially at the edge of bright objects.

The negative pixel values are introduced by the inherent nature of complex number mathematics in reciprocal Fourier space, and are introduced when excessive low-frequency filtering is applied to the image. The presence of negative pixel values is a potential concern when quantitation of the image intensities is required (Righolt *et al.*, 2013, 2014). However, the problem of negative pixel values is generally readily addressed by an accurate assignment of the reconstruction algorithm parameters.

In conclusion, a quality control of the reconstructed SIM image is necessary to prevent or eliminate sources of artifacts. It is critical to carefully inspect the raw images and look in detail at the log files, which contain information on each step of the processing of the SIM image. The raw images should be examined for the presence of phase shifts, intensity differences between angles, as well as overall image modulation contrast. These elements give us an idea of the overall quality of our sample and its compatibility with SIM imaging. In the reconstructed image, the Fourier transform of the 3D volume should be analyzed for symmetry and homogeneity, which is lost due to incorrect sidebands assignment. Recently, an excellent user-friendly image j-based plug-in has been developed by Graeme Ball and Lothar Schermelleh at University of Oxford (See Relevant Website Section). SIM check is compatible with different formats of commercially available microscopes and provides an efficient way to assess the quality of the reconstructed SIM image.

Applications

Unlike other super-resolution techniques, SIM does not require any specialized fluorescent dyes or genetically encoded proteins and easily accommodates multicolor imaging. In addition the sample preparation is not different from any other conventional fluorescence microscopy technique (Ball *et al.*, 2012; Dobbie *et al.*, 2011). Altogether, these qualities of SIM microscopy greatly facilitate the direct transition to a super-resolution method from confocal or widefield microscopy.

Initially, the vast majority of SIM applications was mainly intended as proof of principle of the hardware and, as such, was focused on structures densely labeled like fluorescent beads, actin, or microtubule structures. Since the introduction of turn-key 3DSIM microscopes from several manufacturers (e.g., General Electric, Zeiss, and Nikon), we are now garnering a large amount of novel information in the recent literature on a variety of cellular structures. In this section we describe some of the most interesting applications with an emphasis on the experimental work that went beyond a purely descriptive analysis.

Nuclear, Chromatin Topography, and DNA-Protein Assemblies

One of the very first biological applications of SIM analyzed different protein components and DNA in the context of the nuclear pore complex (120 MDa in vertebrates) in mammalian tissue culture cells (Schermelleh *et al.*, 2008b). This work provided the first evidence of a zone of exclusion of chromatin from regions where nuclear pore complex (NPC) are positioned at the nuclear membrane. Recent studies have aimed at

further expanding these findings by combining 3DSIM with fluorescence *in situ* hybridization (FISH) to map in detail chromatin domains, areas of chromosomal interactions, and the positions of individual genes within specific protein assemblies (Markaki *et al.*, 2010, 2012). More recently, 3DSIM has been employed to characterize the number of foci of DNA replication (Baddeley *et al.*, 2010). By taking advantage of the increased axial resolution, 3DSIM allowed a more accurate and statistically significant quantitation ('quantification' is better) of the total number of foci in mouse cells compared to conventional fluorescence imaging. 3DSIM has also been successfully applied to examine the organization of chromosome structures as the synaptonemal complex both in *Caenorhabditis elegans* and maize chromosomes (Carlton, 2008). As a whole, the nuclear architecture studies highlight the power of 3DSIM for multicolor imaging and its flexibility to study both proteins and DNA when it is possible to use fluorescent probes as in FISH.

Centrosome Structure and Protein Domain Mapping

The centrosome – the major microtubule organizing center in animal cells – is an organelle whose dimensions are ~200 nm in diameter per 200 nm to several microns in length depending on the species. Its size at the boundary of the resolution limit makes it an ideal organelle for SIM imaging. The centrosome is composed of the centriole, a ninefold symmetrical core, surrounded by the pericentriolar material (PCM). While the centriole structure has been extensively studied by electron microscopy (EM), for many years the understanding of the architecture of the PCM has been a particularly difficult problem to tackle. The PCM contains numerous large coiled-coil containing proteins, which are very difficult to purify, and it appears as a relative homogeneous electron dense cloud under the electron microscope. Recent systematic SIM studies revealed the existence of multiple layers in the PCM by establishing the position and orientation of several proteins critical for centrosome maturation in a systematic manner (Lawo *et al.*, 2012; Mennella *et al.*, 2012; Sonnen *et al.*, 2013; Fu and Glover, 2012). Unexpectedly, SIM analysis showed that some of these components display distinctive distributions: some are organized as apparent molecular fibers or toroids extending away from the centriole wall (pericentrin/PLP, Cep152/Asl), while others are organized in a matrix (Cep192/Spd-2, Cep215/Cnn, and γ tubulin; Figures 3 and 4).

In addition to changing our understanding of the structural organization of the PCM, SIM analysis allowed for the construction of a positional map of specific components of the PCM using probes against different protein domains of the PCM coiled-coil proteins. These experiments demonstrated that one of the components of the PCM, pericentrin/PLP, adopts an extended conformation and that the molecules are anchored at the centriole wall region through their C-terminal PACT domain, while the N-terminal projects outwards the periphery. In addition, a cell cycle analysis of pericentrin-like protein (PLP) molecules showed a two-state PLP structure (open vs. closed) that correlates with different states during the centriole duplication pathway (Mennella *et al.*, 2013). Overall these experiments showed the potential of super-resolution

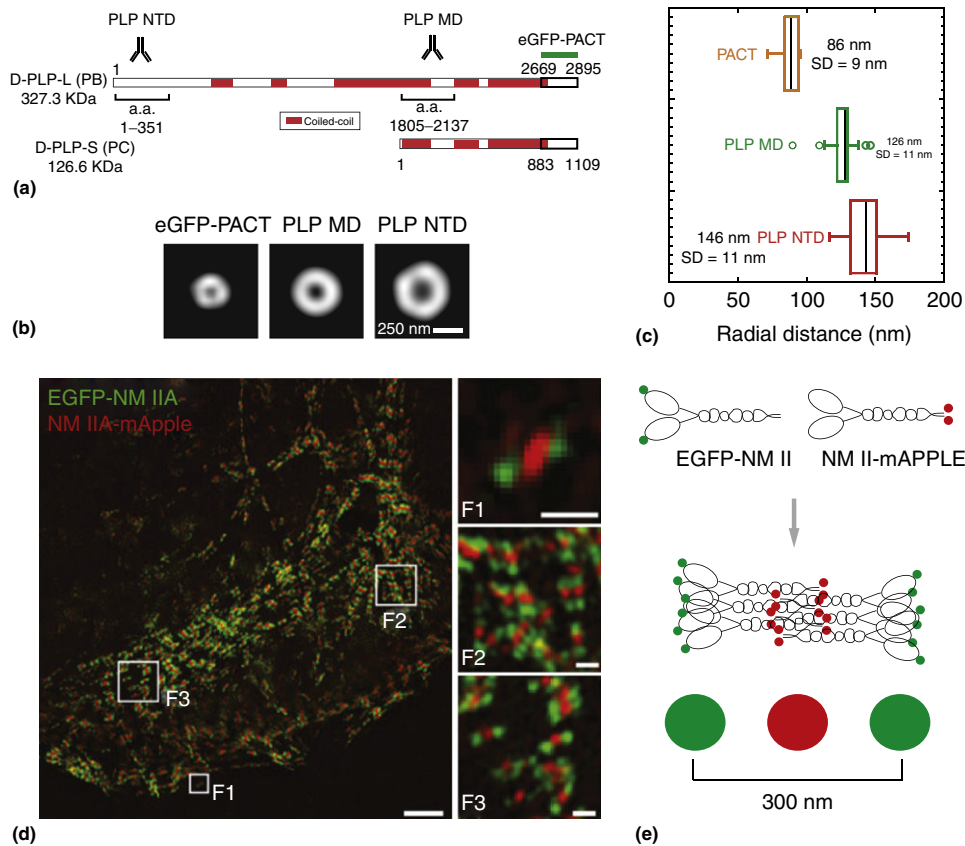


Figure 3 SIM and molecular architecture analysis. 3DSIM reveals the molecular architecture of pericentrin-like protein (PLP). (a) Amino-acid map of *Drosophila* PLP isoforms predicted from the Flybase database. (b) 2D projections of the average aligned volumes obtained either from S2 cells stained with rabbit anti-PLP MD primary antibodies (amino acids 1805–2137, anti-rabbit Alexa 488) or anti-PLP NTD primary antibodies (amino acids 1–381, anti-rabbit Alexa 488). S2 cells expressing eGFP-PACT were co-stained with mouse anti-eGFP monoclonal antibody (anti-mouse Alexa 488). (c) Radially averaged fluorescence intensity values obtained from individual centrosomal protein projected volumes were fitted to a Gaussian function to calculate the centre position and deviation of the distribution. Adapted from Mennella, V., Keszthelyi, B., McDonald, K.L., *et al.*, 2012. Subdiffraction-resolution fluorescence microscopy reveals a domain of the centrosome critical for pericentriolar material organization. *Nature Cell Biology* 14 (11), 1159–1168. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23086239> (accessed 29.05.13). (d) TIRF-SIM of cells expressing NM IIA with N- and C-terminal fluorescent tags allows identification of individual NM IIA bipolar filaments. TIRF-SIM image of a U2OS cell co-expressing EGFP-NM IIA and NMIIA-mApple. White numbered boxes correspond to the magnified insets to the right. Scale bars represent 2 μ m for (d) and 300 nm for insets. (e) Top. Cartoon of NM II with an N-terminal EGFP reporter, or with a C-terminal mApple reporter (light chains not depicted). Bottom. Cartoon of a NMII bipolar filament containing a combination of the NMII constructs. Below the filament is a cartoon depicting what one would expect to see when such a filament is imaged with TIRF-SIM. Adapted from Beach, J.R., Shao, L., Remmert, K., *et al.*, 2014. Nonmuscle myosin II isoforms coassemble in living cells. *Current Biology*: CB 24 (10), 1160–1166. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24814144> (accessed 16.07.14).

methods to study not only the ensemble distribution of molecules, but also the orientation of individual molecules, and their distinct large-scale conformational states, all in the context of cells.

Ciliogenesis and Structural Analysis of Organelles

3DSIM has been applied with impressive results to the analysis of the pathway of centriole duplication during ciliogenesis in epithelia from mouse trachea cells (Zhao *et al.*, 2013; Jord *et al.*, 2014). In these works, 3DSIM imaging allowed the identification of distinct stages of this process by detailing the specific protein composition and architecture of deuterosomes, while simultaneously allowing overall measurements of size and orientation of this relatively uncharacterized organelle. This is a paradigmatic example of how 3DSIM can

complement, enrich, and partially substitute the information derived by EM imaging, because it requires less time in terms of sample processing and optimization.

Cytoskeleton Structure and Motor Proteins Organization

The high density of labeling of cytoskeletal structures makes them a very attractive target for SIM imaging and super-resolution microscopy methods in general (Engel, 2014). The major advantage of using SIM for such biological application is to achieve a more effective separation of individual cytoskeletal elements, which in turn allows a better quantitation of the number of filaments (Miao *et al.*, 2013), their structural nature ((Brown *et al.*, 2011), and overall dynamics (Kner *et al.*, 2009; Komis *et al.*, 2014; Fiolka, 2014). Recently, the formation of assemblies composed of different isoforms of

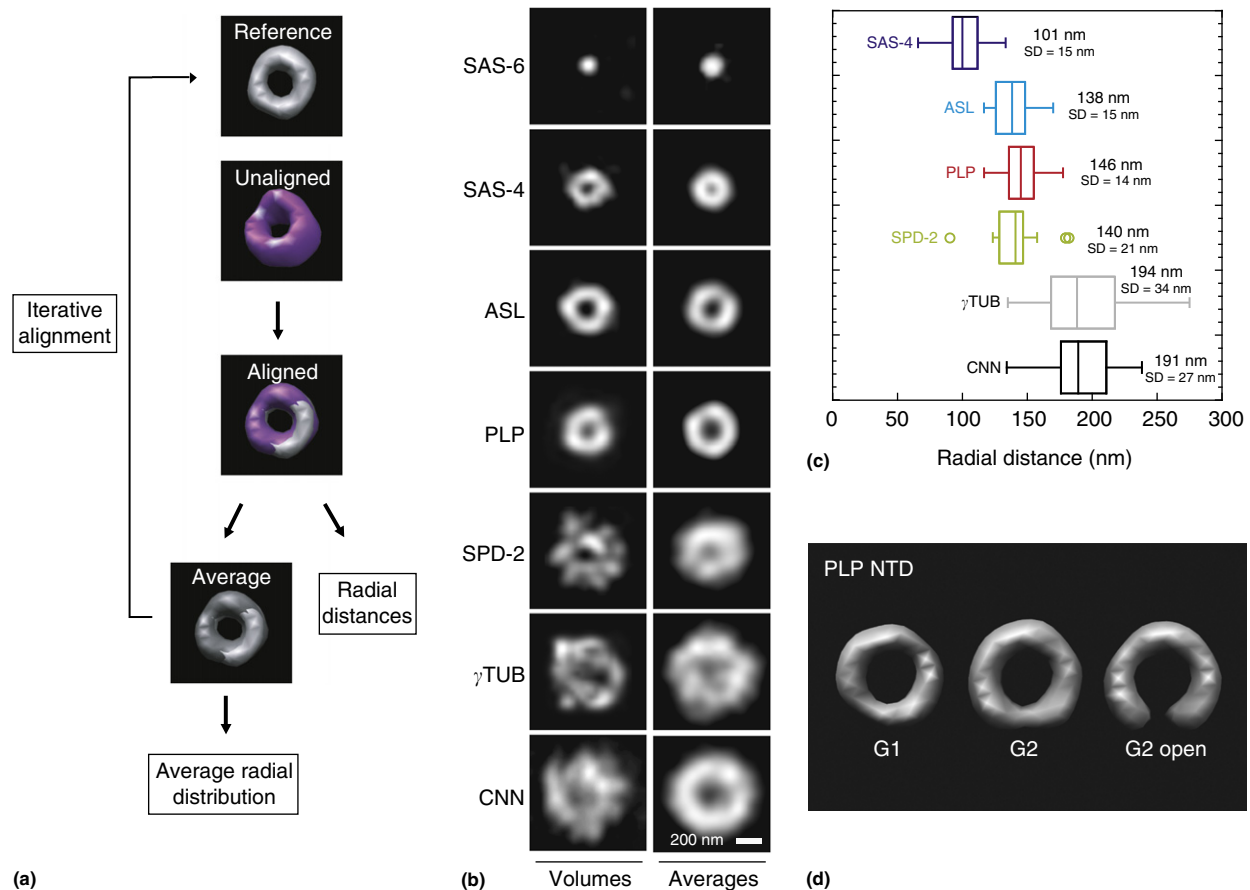


Figure 4 3DSIM volume alignment and averaging. (a) Summary schematic of the 3D subvolume iterative alignment strategy based on cross-correlation in real space used for alignment and analysis of experimental 3D volumes. (b) 2D projections of the average aligned volumes. Scale bar, 200 nm. (c) Radially averaged fluorescence intensity values obtained from individual centrosomal protein projected volumes were fitted to an offset Gaussian function to calculate the centre position and deviation of the distribution. (d) PLP fibrils associated with mother centrioles form a gap where the daughter centriole assembles. Rendering of volume averages of G1 and G2 centrosomes stained with anti-PLP NTD antibody. G2 centrosomes were divided into two separate populations for subclass averaging of centrioles with an open gap or partially open gap. Adapted from Mennella, V., Keszthelyi, B., McDonald, K.L., *et al.*, 2012. Subdiffraction-resolution fluorescence microscopy reveals a domain of the centrosome critical for pericentriolar material organization. *Nature Cell Biology* 14 (11), 1159–1168. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23086239> (accessed 29.05.13).

myosin II in cells has been demonstrated for the first time (Beach *et al.*, 2014). This was possible by taking advantage of the filamentous structure of myosin II and by labeling opposite ends of the heteromeric complexes that are separated by ~ 300 nm. This application highlights, similarly to the studies of the PCM, how SIM can be used to analyze the detailed organization of molecular assemblies in the cellular context. In a thorough analysis of the final steps of cell division in HeLa cells, SIM was critical for shedding light on the asymmetrical and dynamic organization of two machineries involved in complementary steps during abscission (Elia *et al.*, 2011, 2013; Guizetti *et al.*, 2011). In this work, SIM was used in combination with conventional live cell imaging and electron tomography reconstructions to define the different contribution of the ESCRT-III complex, which is required for cortical constriction, and the severing enzyme spastin, which is involved in microtubule bundle depolymerization in the process of abscission (Elia *et al.*, 2011, 2013; Guizetti *et al.*, 2011).

Bacteria Structural and Functional Organization

In a recent paper from Lesterlin *et al.* (2014), SIM was used to examine the organization and dynamics of bundles of Rec A, a protein involved in sister pair chromatid homology recombination repair in *Escherichia coli*. In this work, SIM imaging contributed to a better understanding of Rec A filament architecture and dynamics and their relationship to other components of the recombination machinery. By showing the existence of different populations of Rec A filaments, one stable and one highly dynamic, this work ultimately provided insights into the mechanism of homology repair in prokaryotes (Lesterlin *et al.*, 2014). Other studies also examined the role of filamentous structures in bacteria that are part of their shape formation machinery (as MreB filaments) (Olshausen *et al.*, 2013), the cellular architecture of the machinery involved in septum formation in *Bacillus subtilis* (Eswaramoorthy *et al.*, 2011), and the filamentous organization of tubulin and actin homologues involved in formation

of the z ring critical for cell division of bacteria (Coltharp and Xiao, 2012; Szwedziak *et al.*, 2012; Strauss *et al.*, 2012).

Recent Developments

Live SIM

When compared to other super-resolution microscopies as STORM–PALM, SIM has the considerable advantage of requiring relatively few raw images (5000–10 000 frames per STORM/PALM image, Winter and Shroff, 2014). SIM microscopy requires low illumination intensity compared to methods like STED (Klar *et al.*, 2000; Huang, 2010; Hell, 2009), which in turn allows minimal photobleaching and longer imaging time series. Moreover, because linear SIM does not rely on stimulated emission, it causes less phototoxicity and photodamage (Rego *et al.*, 2012). These properties suggest that SIM microscopy might be in the long run the super-resolution method best suited to perform live cell imaging experiments with long time series and large field of views.

However, with few exceptions SIM has been thus far applied to image samples chemically fixed and processed for immunostaining. The fast time scale of most biological processes, which are beyond the reach of linear SIM, has largely determined this limitation. To address this issue, in recent years hardware and software modifications have been implemented that significantly speed up image acquisition. These schemes improve the speed of interference pattern scanning (SLM), allow ultralow confined sample illumination (2D TIRF-SIM, Bessel beam SIM) or developed altogether different illumination schemes (multifocal SIM).

In the SIM implementation of Kner *et al.* SLMs were introduced for the first time. SLMs generate diffraction patterns with precise phase and intensities by varying the refractive properties of its liquid crystals. This allows rapid pattern switching and enables very fast acquisition rates (Table 1; Kner *et al.*, 2009; Fiolka *et al.*, 2012; Fiolka, 2014). This SIM microscope also implemented a TIRF illumination system which (1) improves the signal to noise, which allow further reduction of the illumination exposure time and (2) reduces the number of raw images to nine per z stack, because in a 2DTIRF scheme only three phase images per angle are necessary (Gustafsson *et al.*, 2008; Kner *et al.*, 2009). This strategy has been successfully applied to the live cell imaging of microtubules and kinesin motors (Kner *et al.*, 2009) and later to the study of the actin

cytoskeleton at the leading edge in the absence of cofilin (Vitriol *et al.*, 2013).

In a subsequent implementation, a new generation of faster SLM and a different acquisition scheme allowed impressive live imaging results of mitochondria dynamics in Hela cells (Table 1; Shao *et al.*, 2011). This strategy is based on the acquisition of first phases and angles for each z plane, which further reduces the overall acquisition time. As previously discussed, for effective live cell imaging in SIM mode, the fluorescence structure of interest should move slower than the speed of image volume acquisition – ~100 nm in a single stack. Interestingly, individual z planes did not suffer from motion artifacts of different planes due to the depth of focus of 400 nm in a 3DSIM experiment (Shao *et al.*, 2011).

In a scheme substantially different from linear 3DSIM, Schroff *et al.* implemented a multifocal live cell imaging SIM strategy that takes advantage of simultaneous point-based illumination pattern at multiple focal points and a final deconvolution step (York *et al.*, 2013, 2012). This combined approach produces an effect similar to a pinhole in a confocal microscope and it is capable of reaching resolution comparable to a linear 3DSIM (145 nm lateral and 350 nm axial). The main advantage of this approach is the very fast frame rate compared to other super-resolution microscopy techniques (two color 3 μm stack in 1.2 s (Winter and Shroff, 2014; York *et al.*, 2013, 2012)). This method has been applied successfully to image a wide range of processes as mitochondria, peroxisome, and ER dynamics (Figure 5; York *et al.*, 2013).

In another very recent implementation, the structured illumination pattern of SIM is combined with a confined illumination excitation into a defined focal plane achieved by planar bessel beam illumination (BBSIM, Gao *et al.*, 2012). While in BBSIM there is only a modest increase in resolution (Table 1), this scheme increases contrast and signal-to-noise and reduce phototoxicity, qualities particularly advantageous when imaging thick samples as in multicellular organisms. A direct comparison between SIM and BB3DSIM shows that the planar illumination introduced by Bessel beams increases image quality in thick samples at depths over 15 μm , which are generally outside the reach of SIM, because of low modulation contrast and increased sensitivity to aberration (Figure 6). Another important advantage of this method is that it determines an increase in the number of sections imaged per second to 10–27 planes sec^{-1} . This methodology has achieved impressive results in imaging individual chromosomes for karyotyping in mitosis, macropinocytosis

Table 1 Comparison of the properties of structured illumination microscopy methods for live cell imaging

Sim method	Publication	Resolution x/y/z (nm)	Depth of view (μm)	Frame rate (Hz)	Time points	Field of view (μm)	Power (W cm^{-2})
2DTIRF SIM	Kner <i>et al.</i> (2009)	112/112	0.1	3.7–11	480	32 $\times$ 32	5–10
3DSIM	Shao <i>et al.</i> (2011)	120/120/360	<15	0.2	50	200 $\times$ 200	5.5
SSIM	Rego <i>et al.</i> (2012)	40/40	0.1	NA	1	115 $\times$ 115	1–10
MMSIM	York <i>et al.</i> (2013)	145/145/356	>15	100 (2D); 5–10 (3D)	>100 (2D); 17 (3D)	6.5 $\times$ 3.8 (2D); 71 $\times$ 45 (3D)	5–50
BBSIM	Gao <i>et al.</i> (2012)	180/230/350	>15	10–20	50–100	Variable	NA

Note: Most parameters are dependent on the sample type.

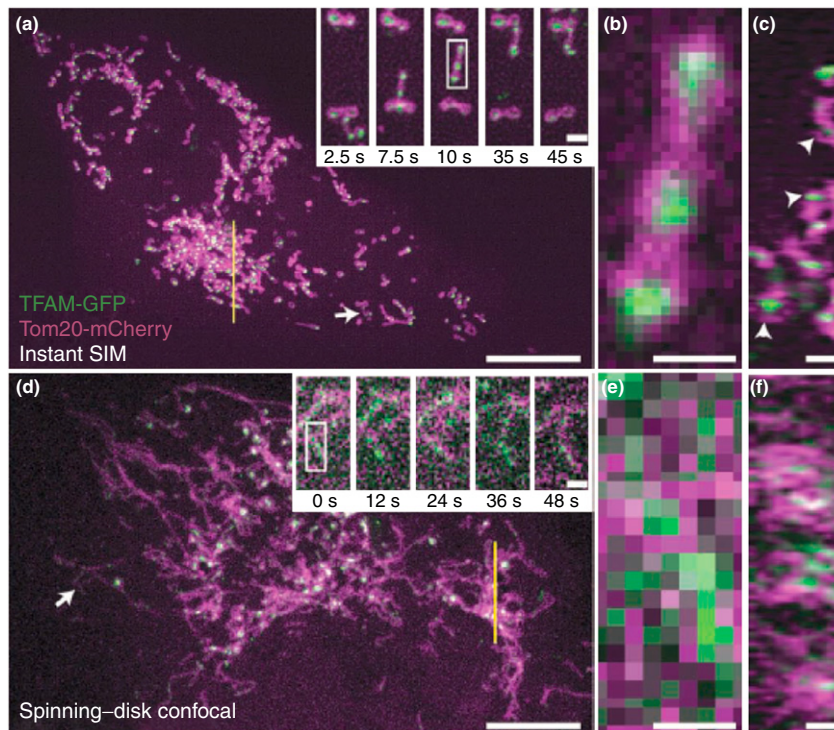


Figure 5 Multifocal SIM enables high-speed two-color super-resolution imaging. Live MRL-TR transformed human lung fibroblasts expressing TFAM-GFP (green) and Tom20-mCherry (magenta) were imaged in multifocal SIM (a–c) and spinning-disk confocal (d–f). Maximum-intensity projections (xy) of 3-μm-thick volumes are shown in (a), (d) with subregions (white arrows) at indicated time points in higher magnification (inset). Scale bars, 10 μm and 1 μm (insets). Higher magnification of mitochondria is shown in (b), (c), (e), (f) highlighted by white rectangles from insets in (a), (d). Scale bars, 0.5 μm. Axial (zy) views of ~270-nm-thick slices (yellow lines, (a), (d)) are shown in (c), (f). White arrowheads indicate TFAM enclosed by Tom20 in instant SIM. Scale bars, 1 μm.

and actin cytoskeleton dynamics, and more thick samples as *C. elegans* embryos (Gao *et al.*, 2012, 2014).

Another way to reduce the overall acquisition time is to find ways to reduce the exposure time by requiring less signals: this can be achieved with denoising algorithms that are capable of reducing the background noise introduced by digital processing. Denoising algorithms take advantage of the different properties of correlation between adjacent pixels between signal and noise to identify and reduce the impact on noise on raw images. As a consequence, a reduced exposure to laser light reduces phototoxicity and photobleaching thus allowing longer imaging experiments without compromising the integrity of the sample. So far these algorithms have been applied to widefield images, but efforts are under way to analyze SIM data sets (Buades *et al.*, 2005). It is likely that in the near future, the combination of fast and novel illumination patterning and noise reduction algorithms will allow the application of live SIM imaging to a large range of biological processes happening on faster time scales than $< 100 \text{ nm sec}^{-1}$.

Nonlinear SIM

In nonlinear structure illumination (NLSIM) experiments, resolution higher than linear 3DSIM can be derived from harmonics generated by nonlinear effects of fluorophore activation and emission. These higher order harmonics are

obtained either when high-energy light pulse is shined on the sample (Heintzmann *et al.*, 2002) or when photo-switchable fluorophores are used (Rego *et al.*, 2012). In principle, this method can achieve infinite resolution, in practice, the high laser intensities required to obtain nonlinear responses from the fluorophores decrease the sample photostability, leading to a rapid degradation of the signal to noise. This in turn limits the number of effective harmonics from which to derive the high-resolution information (Rego *et al.*, 2012). Experimentally, NLSIM has achieved resolution in the order of 40–50 nm in a 2D TIRF setup. By using the photoactivatable protein Dronpa NLSIM can achieve harmonics with significantly lower light intensities and it has been employed to study cellular structures as microtubule and actin. Unfortunately, the application of this strategy for live cell imaging or to analyze 3D volumes will require the development of photoactivatable proteins that are more photostable and that can sustain a large number of cycles with the same photon output (Rego *et al.*, 2012).

SIM Image Alignment and Averaging

As described in previous sections, SIM experiments can provide quantitative positional information, all inside the cell. The distribution of molecular assemblies, but also their domain localization and molecular orientation within individual molecules can be obtained providing an unprecedented

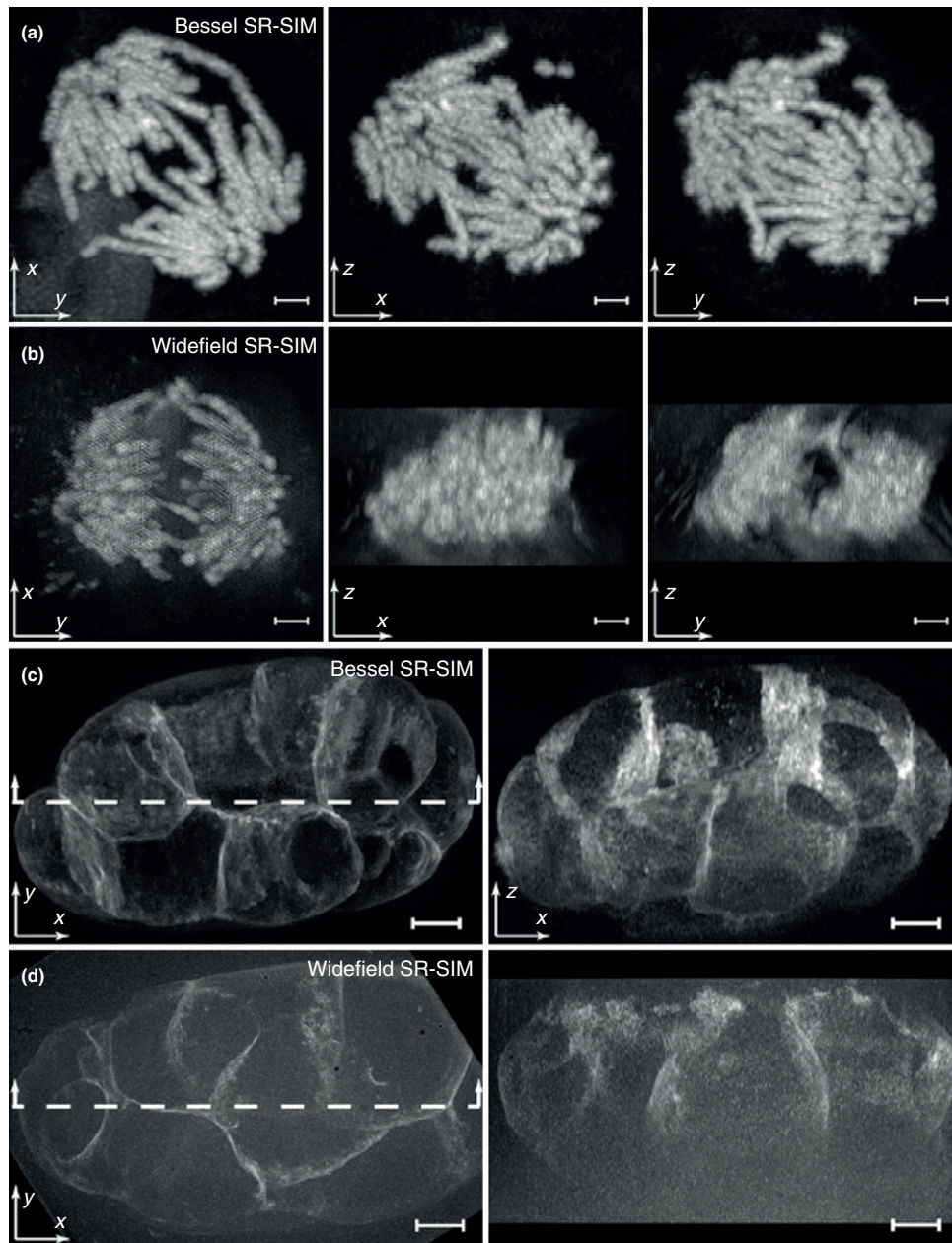


Figure 6 Comparison of Bessel plane SR-SIM with 3DSIM in densely fluorescent or thick specimens. (a) Three orthogonal maximum-intensity projections of chromosomes in a live LLC-PK1 cell in anaphase, as seen by Bessel plane SR-SIM. (b) Similar views in another live LLC-PK1 cell in anaphase, as seen by widefield 3D SR-SIM. (c) Lateral (left) and axial (right) maximum-intensity-projection views of a living early-stage *C. elegans* embryo, as seen by Bessel plane SR-SIM. (d) Similar views in another living *C. elegans* embryo, as seen by widefield 3D SR-SIM. Scale bars: 2 mm in (a) and (b); 5 mm in (c) and (d). Reproduced from Gao, L., Shao, L., Higgins, C.D., *et al.*, 2012. Noninvasive imaging beyond the diffraction limit of 3D dynamics in thickly fluorescent specimens. *Cell* 151 (6), 1370–1385. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3615549&tool=pmcentrez&rendertype=abstract> (accessed 10.07.14).

snapshot of the organization of supra-molecular complexes in cells (Lawo *et al.*, 2012; Mennella *et al.*, 2012; Beach *et al.*, 2014).

In order to fully extract the quantitative information from 3DSIM – and other super-resolution – images, rather than merely showing anecdotal examples, it is useful to develop image processing and analysis tools to derive image averages of cellular structures. The derivation of an average

image volume enhances the similarities between individual examples, largely deemphasizes stochastic variability and deformation introduced by fixation, while it still allows an analysis of the individual volume variance and heterogeneity (Mennella *et al.*, 2012, 2013).

Toward this goal, recently cryo-tomographic analyses have been applied for the first time in SIM imaging. In this method 3D sub-volumes are aligned by translating, tilting, and

in-plane rotating the volumes to a reference by cross-correlation of voxel intensities. One major advantage of this method is that it provides outstanding quantitative metrics of the spatial distribution of cellular components with error in the range of a few nanometers (up to 2.5 nm, personal communication). This in turn favors a precise and unambiguous protein domain position assignment. It is important to mention that high measurement precision does not necessarily mean high accuracy in determining macromolecules position. The distance of the antigen from the fluorophore (Ries *et al.*, 2012) still contributes to the overall accuracy of our measurement.

See also: Imaging the Cell: Light Microscopy: Super Resolution Fluorescence Localization Microscopy; Super-Resolution Light Microscopy: Stimulated Emission Depletion and Ground-State Depletion

References

- Abbe, E., 1874. A contribution to the theory of the microscope and the nature of microscopic vision. *Proceedings of the Bristol Naturalists' Society* 1, 200–261.
- Alberts, B., 1998. The cell as a collection of protein machines: Preparing the next generation of molecular biologists. *Cell* 92 (3), 291–294. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/9476889> (accessed 01.04.15).
- Baddeley, D., Chagin, V.O., Schermelleh, L., *et al.*, 2010. Measurement of replication structures at the nanometer scale using super-resolution light microscopy. *Nucleic Acids Research* 38 (2), e8. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2811013&tool=pmcentrez&rendertype=abstract> (accessed 14.07.14).
- Ball, G., Parton, R.M., Hamilton, R.S., Davis, I., 2012. A Cell Biologist's Guide to High Resolution Imaging, first ed. Elsevier Inc. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22264528> (accessed 16.07.14).
- Beach, J.R., Shao, L., Remmert, K., *et al.*, 2014. Nonmuscle myosin II isoforms coassemble in living cells. *Current Biology*: CB 24 (10), 1160–1166. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24814144> (accessed 16.07.14).
- Brown, A.C.N., Oddos, S., Dobbie, I.M., *et al.*, 2011. Remodelling of cortical actin where lytic granules dock at natural killer cell immune synapses revealed by super-resolution microscopy. *PLoS Biology* 9 (9), e1001152. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3172219&tool=pmcentrez&rendertype=abstract> (accessed 16.07.14).
- Buades, A., Coll, B., Morel, J.-M., 2005. A review of image denoising algorithms, with a new one. *Multiscale Modeling and Simulation* 4 (2), 490–530.
- Carlton, P.M., 2008. Three-dimensional structured illumination microscopy and its application to chromosome structure. *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology* 16 (3), 351–365. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18461477> (accessed 26.05.14).
- Coltharp, C., Xiao, J., 2012. Superresolution microscopy for microbiology. *Cellular Microbiology* 14 (12), 1808–1818. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3894611&tool=pmcentrez&rendertype=abstract> (accessed 10.07.14).
- Dobbie, I.M., King, E., Parton, R.M., *et al.*, 2011. OMX: A new platform for multimodal, multichannel wide-field imaging. *Cold Spring Harbor protocols* 2011 (8), 899–909. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21807861> (accessed 16.07.14).
- Elia, N., Ott, C., Lippincott-Schwartz, J., 2013. Incisive imaging and computation for cellular mysteries: Lessons from abscission. *Cell* 155 (6), 1220–1231. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24315094> (accessed 17.07.14).
- Elia, N., Sougrat, R., Spurlin, T.A., Hurlley, J.H., Lippincott-Schwartz, J., 2011. Dynamics of endosomal sorting complex required for transport (ESCRT) machinery during cytokinesis and its role in abscission. *Proceedings of the National Academy of Sciences of the United States of America* 108, 4846–4851.
- Engel, U., 2014. Structured Illumination Superresolution Imaging of the Cytoskeleton, first ed. Elsevier Inc. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24974035> (accessed 15.07.14).
- Eswaramoorthy, P., Erb, M.L., James, A., 2011. Cellular architecture mediates DivIVA ultrastructure and regulates min activity in *Bacillus subtilis*. *mBio* 2 (6), e00257–11.
- Fiolka, R., 2014. Seeing More with Structured Illumination Microscopy, first ed. Elsevier Inc. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24974034> (accessed 15.07.14).
- Fiolka, R., Shao, L., Rego, E.H., Davidson, M.W., Gustafsson, M.G., 2012. Time-lapse two-color 3D imaging of live cells with doubled resolution using structured illumination. *Proceedings of the National Academy of Sciences of the United States of America* 109 (14), 5311–5315. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3325651&tool=pmcentrez&rendertype=abstract> (accessed 26.05.14).
- Fu, J., Glover, D.M., 2012. Structured illumination of the interface between centriole and peri-centriolar material. *Open Biology* 2 (8), 120104. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3438536&tool=pmcentrez&rendertype=abstract> (accessed 01.05.15).
- Gao, L., Shao, L., Chen, B.C., Betzig, E., 2014. 3D live fluorescence imaging of cellular dynamics using Bessel beam plane illumination microscopy. *Nature Protocols* 9 (5), 1083–1101. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24722406> (accessed 10.07.14).
- Gao, L., Shao, L., Higgins, C.D., *et al.*, 2012. Noninvasive imaging beyond the diffraction limit of 3D dynamics in thickly fluorescent specimens. *Cell* 151 (6), 1370–1385. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3615549&tool=pmcentrez&rendertype=abstract> (accessed 10.07.14).
- Guizetti, J., Schermelleh, L., Mäntler, J., *et al.*, 2011. Cortical constriction during abscission involves helices of ESCRT-III-dependent filaments. *Science (New York, NY)* 331 (6024), 1616–1620. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21310966> (accessed 16.07.14).
- Gustafsson, M.G., 2000. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy* 198 (Pt 2), 82–87. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10810003> (accessed 01.04.15).
- Gustafsson, M.G.L., Shao, L., Carlton, P.M., *et al.*, 2008. Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophysical Journal* 94 (12), 4957–4970. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2397368&tool=pmcentrez&rendertype=abstract> (accessed 02.11.12).
- Heintzmann, R., Ficz, G., 2007. Breaking the resolution limit in light microscopy. *Methods Cell Biology* 81, 561–580.
- Heintzmann, R., Ficz, G., 2013. Breaking the resolution limit in light microscopy. *Methods Cell Biology* 114, 525–544. doi:10.1016/B978-0-12-407761-4.00022-1.
- Heintzmann, R., Jovin, T.M., Cremer, C., 2002. Saturated patterned excitation microscopy — A concept for optical resolution improvement. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision* 19 (8), 1599–1609. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12152701> (accessed 01.04.15).
- Hell, S.W., 2007a. Far-field optical nanoscopy. *Science (New York, NY)* 316 (5828), 1153–1158. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17525330> (accessed 11.07.14).
- Hell, S.W., 2007b. Far-field optical nanoscopy. *Science (New York, NY)* 316 (5828), 1153–1158.
- Hell, S.W., 2009. Microscopy and its focal switch. *Nature Methods* 6 (1), 24–32.
- Huang, B., 2010. Super-resolution optical microscopy: Multiple choices. *Current Opinion in Chemical Biology* 14 (1), 10–14. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19897404> (accessed 31.05.13).
- Huang, B., Bates, M., Zhuang, X., 2009. Super-resolution fluorescence microscopy. *Annual Review of Biochemistry* 78, 993–1016. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2835776&tool=pmcentrez&rendertype=abstract> (accessed 22.05.13).
- Jord, A.A.I., Lemaître, A.-I., Delgehyr, N., *et al.*, 2014. Centriole amplification by mother and daughter centrioles differs in multiciliated cells. *Nature*. Available at: <http://www.nature.com/doi/10.1038/nature13770> (accessed 13.11.14).
- Klar, T.A., Jakobs, S., Dyba, M., Egner, A., Hell, S.W., 2000. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proceedings of the National Academy of Sciences of the United States of America* 97 (15), 8206–8210. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=26924&tool=pmcentrez&rendertype=abstract> (accessed 01.04.15).
- Kner, P., Chhun, B.B., Griffis, E.R., Winoto, L., Gustafsson, M.G., 2009. Super-resolution video microscopy of live cells by structured illumination. *Nature Methods* 6 (5), 339–342.

- Komis, G., Mistrik, M., Samajová, O., *et al.*, 2014. Dynamics and organization of cortical microtubules as revealed by superresolution structured illumination microscopy. *Plant Physiology* 165 (1), 129–148. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24686112> (accessed 16.07.14).
- Langhorst, M.F., Schaffer, J., Goetze, B., 2009. Structure brings clarity: Structured illumination microscopy in cell biology. *Biotechnology Journal* 4 (6), 858–865. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19492328> (accessed 22.05.14).
- Lawo, S., Hasegan, M., Gupta, G.D., Pelletier, L., 2012. Subdiffraction imaging of centrosomes reveals higher-order organizational features of pericentriolar material. *Nature Cell Biology* 14 (11), 1148–1158. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23086237> (accessed 14.07.13).
- Lesterlin, C., Ball, G., Schermelleh, L., Sherratt, D.J., 2014. RecA bundles mediate homology pairing between distant sisters during DNA break repair. *Nature* 506 (7487), 249–253. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24362571> (accessed 26.05.14).
- Markaki, Y., Gunkel, M., Schermelleh, L., *et al.*, 2010. Functional nuclear organization of transcription and DNA replication: A topographical marriage between chromatin domains and the interchromatin compartment. *Cold Spring Harbor Symposia on Quantitative Biology* 75, 475–492. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21467142> (accessed 14.07.14).
- Markaki, Y., Smeets, D., Fiedler, S., *et al.*, 2012. The potential of 3D-FISH and super-resolution structured illumination microscopy for studies of 3D nuclear architecture: 3D structured illumination microscopy of defined chromosomal structures visualized by 3D (immuno)-FISH opens new perspectives for study. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 34 (5), 412–426. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22508100> (accessed 11.07.14).
- Mennella, V., Agard, D.A., Huang, B., Pelletier, L., 2013. Amorphous no more: Subdiffraction view of the pericentriolar material architecture. *Trends in Cell Biology*, 1–10. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24268653> (accessed 31.01.14).
- Mennella, V., Keszthelyi, B., McDonald, K.L., *et al.*, 2012. Subdiffraction-resolution fluorescence microscopy reveals a domain of the centrosome critical for pericentriolar material organization. *Nature Cell Biology* 14 (11), 1159–1168. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23086239> (accessed 29.05.13).
- Mertz, J., 2011. Optical sectioning microscopy with planar or structured illumination. *Nature Methods* 8 (10), 811–819. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21959136> (accessed 04.05.14).
- Miao, Y., Wong, C.C., Mennella, V., *et al.*, 2013. Cell-cycle regulation of formin-mediated actin cable assembly. *Proceedings of the National Academy of Sciences of the United States of America* 110 (47), E4446–E4455.
- Olshausen, P.V., Defeu Soufo, H.J., Wicker, K., *et al.*, 2013. Superresolution imaging of dynamic MreB filaments in *B. subtilis* — A multiple-motor-driven transport? *Biophysical Journal* 105 (5), 1171–1181. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24010660> (accessed 16.07.14).
- Pawley, J.B. (Ed.), 2006. *Handbook of Biological Confocal Microscopy*, second ed. New York: Springer.
- Rego, E.H., Shao, L., Macklin, J.J., *et al.*, 2012. Nonlinear structured-illumination microscopy with a photoswitchable protein reveals cellular structures at 50-nm resolution. *Proceedings of the National Academy of Sciences of the United States of America* 109 (3), E135–E143. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3271870&tool=pmcentrez&rendertype=abstract> (accessed 16.07.14).
- Ries, J., Kaplan, C., Platonova, E., Eghlidi, H., Ewers, H., 2012. A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. *Nature Methods* 9 (6), 582–584. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22543348> (accessed 11.07.14).
- Righolt, C.H., Mai, S., van Vliet, L.J., Stallinga, S., 2014. Three-dimensional structured illumination microscopy using Lukosz bound apodization reduces pixel negativity at no resolution cost. *Optics Express* 22 (9), 11215–11227.
- Righolt, C.H., Slotman, J.A., Young, I.T., *et al.*, 2013. Image filtering in structured illumination microscopy using the Lukosz bound. *Optics Express* 21 (21), 24431–24451.
- Schermelleh, L., Carlton, P.M., Haase, S., *et al.*, 2008a. Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* (New York, NY) 320 (5881), 1332–1336. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2916659&tool=pmcentrez&rendertype=abstract> (accessed 30.04.14).
- Schermelleh, L., Carlton, P.M., Haase, S., *et al.*, 2008b. Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* (New York, NY) 320 (5881), 1332–1336. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2916659&tool=pmcentrez&rendertype=abstract> (accessed 18.11.12).
- Schermelleh, L., Heintzmann, R., Leonhardt, H., 2010. A guide to super-resolution fluorescence microscopy. *Journal of Cell Biology* 190 (2), 165–175. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2918923&tool=pmcentrez&rendertype=abstract> (accessed 21.05.13).
- Shao, L., Kner, P., Rego, E.H., Gustafsson, M.G., 2011. Super-resolution 3D microscopy of live whole cells using structured illumination. *Nature Methods* 8 (12), 1044–1046. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22002026> (accessed 26.05.14).
- Sonnen, K.F., Gabryjonczyk, A.M., Anselm, E., Stierhof, Y.D., Nigg, E.A., 2013. Human Cep192 and Cep152 cooperate in Plk4 recruitment and centriole duplication. *Journal of Cell Science* 126 (Pt 14), 3223–3233.
- Strauss, M.P., Liew, A.T., Turnbull, L., *et al.*, 2012. 3D-SIM super resolution microscopy reveals a bead-like arrangement for FtsZ and the division machinery: Implications for triggering cytokinesis. *PLoS Biology* 10 (9), e1001389.
- Swedlow, J.R., 2013. *Quantitative Fluorescence Microscopy and Image Deconvolution*, fourth ed. Elsevier Inc. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23931516> (accessed 12.05.14).
- Szweziak, P., Wang, Q., Freund, S.M.V., Löwe, J., 2012. FtsA forms actin-like protofilaments. *EMBO Journal* 31 (10), 2249–2260. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3364754&tool=pmcentrez&rendertype=abstract> (accessed 22.07.14).
- Verdaasdonk, J.S., Stephens, A.D., Haase, J., Bloom, K., 2014. Bending the rules: Widefield microscopy and the Abbe limit of resolution. *Journal of Cellular Physiology* 229 (2), 132–138. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23893718> (accessed 07.05.14).
- Vitriol, E.A., Wise, A.L., Berginski, M.E., Bamburg, J.R., Zheng, J.Q., 2013. Instantaneous inactivation of cofilin reveals its function of F-actin disassembly in lamellipodia. *Molecular Biology of the Cell* 24 (14), 2238–2247. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3708729&tool=pmcentrez&rendertype=abstract> (accessed 15.07.14).
- Wallace, W., Schaefer, L.H., Swedlow, J.R., 2001. A working person's guide to deconvolution in light microscopy. *BioTechniques* 31 (5), 1076–1078, 1080, 1082 passim. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11730015> (accessed 01.04.15).
- Winter, P.W., Shroff, H., 2014. Faster fluorescence microscopy: Advances in high speed biological imaging. *Current Opinion in Chemical Biology* 20, 46–53. Figure 1. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24815857> (accessed 14.05.14).
- York, A.G., Chandris, P., Nogare, D.D., *et al.*, 2013. Instant super-resolution imaging in live cells and embryos via analog image processing. *Nature Methods* 10 (11), 1122–1126. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3898876&tool=pmcentrez&rendertype=abstract> (accessed 10.07.14).
- York, A.G., Parekh, S.H., Nogare, D.D., *et al.*, 2012. Resolution doubling in live, multicellular organisms via multifocal structured illumination microscopy. *Nature Methods* 9 (7), 749–754.
- Zhao, H., Zhu, L., Zhu, Y., *et al.*, 2013. The Cep63 paralogue Deup1 enables massive de novo centriole biogenesis for vertebrate multiciliogenesis. *Nature Cell Biology* 15 (12), 1434–1444. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24240477> (accessed 14.07.14).

Relevant Website

<http://www.micron.ox.ac.uk/software/SIMCheck.php>
University of Oxford.

Fluorescence Correlation Spectroscopy: A Tool for Measuring Dynamic and Equilibrium Properties of Molecules in Cells

EL Elson, Washington University School of Medicine, St. Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Observation volume or area The open volume (or area for two-dimensional systems) that is illuminated by the

excitation laser and from which fluorescence is detected. The observation volume is typically limited by a confocal aperture in the microscope optical system.

Introduction

To understand dynamic molecular processes of metabolism and regulation at the cellular level it is necessary to determine how specific molecules move and interact on the surfaces and within living cells. To accomplish this task requires methods that have the following characteristics:

1. Chemical selectivity – the ability to measure the concentrations of specified molecules present at low concentration within an environment, for example, cytoplasm, that contains high concentrations of many other kinds of molecules
2. A wide dynamic range from <microseconds to >minutes
3. Relatively short time required for the measurements (seconds)
4. High spatial resolution to enable measurements within dimensions <1 μm
5. Sensitivity, even down to the single molecule level
6. Lack of toxicity and interference with normal cell function
7. Ability to localize the measurement within identifiable cell structures and features
8. Applicability to a range of concentrations from nM to μM .

These specifications suggest the use of fluorescence, which has the required rapidity of measurement, sensitivity, and selectivity, and, when based on a confocal fluorescence microscope, the spatial resolution and ability to localize the measurement relative to cell structures. Furthermore, when used judiciously, fluorescence of labeled molecules can be measured with little damage to living cells. Two closely related and complementary methods, fluorescence correlation spectroscopy (FCS) and fluorescence photobleaching recovery (FPR, also known as fluorescence recovery after photobleaching, FRAP) take advantage of these properties of fluorescence to measure rates of molecular transport, for example, diffusion and convection, and of chemical reaction over a wide dynamic range of concentrations and time in living cells.

FCS was introduced in 1972 (Magde *et al.*, 1972). The history of its origins and development have been related from the perspectives of the two groups that first demonstrated the feasibility of using fluorescence fluctuations to measure translational diffusion and chemical kinetics (Elson, 2013a,b) and rotational motion (Rigler, 2009). FCS remained a specialized subject of study for a number of years until technological advances enabled more rapid and robust measurements and this in turn led to the production of

commercial instruments by a number of companies and widespread use of FCS in many laboratories. Important improvements in FCS measurements resulted from using confocal microscopy to limit the volume from which fluorescence is detected (Koppel *et al.*, 1976; Rigler *et al.*, 1993) and which allowed single molecule detection (Mets and Rigler, 1994; Rigler, 2009). Improvements in microscope technology (infinity-corrected optics) and stability of fluorophores and computation of correlation functions (Schatzel *et al.*, 1988) also were important contributions. Improvements in microscopy minimized the detection volume thereby diminishing interfering background fluorescence while increasing the detected number of photons emitted per fluorophore.

Basic Concepts and Theory

Correlation of Fluctuations of Concentration and Fluorescence

Most conventional methods to measure chemical rate constants determine their values by perturbing the molecular system from its equilibrium state, for example, by a temperature jump that displaces the system from equilibrium, and then observing the rate of relaxation back to equilibrium (Eigen and De Maeyer, 1963). Similarly, conventional measurements of diffusion are typically carried out by observing the dissipation of a concentration gradient that has been created in the system by the experimenter (Tanford, 1961). In contrast, FCS measures these rates without perturbing systems that remain in equilibrium or nonequilibrium steady states during the measurement. In an open volume of a chemical system defined by a focused laser beam, the numbers of molecules of the system components continually fluctuate due either to their diffusing into and out of the volume or to their participating in chemical reactions. The fluctuation of the concentration of chemical component, j , at position r and time t is $\delta C_j(r,t) = C_j(r,t) - \langle C_j \rangle$ where $\langle C_j \rangle$ is the equilibrium (or nonequilibrium) steady state value of the concentration. The concentration fluctuations, $\delta C_j(r,t)$, causes corresponding fluorescence fluctuations, $\delta F_j(t)$, that are measured directly, typically using a confocal fluorescence microscope (Koppel *et al.*, 1976; Qian and Elson, 1991). A fluorescence fluctuation is related to a concentration fluctuation by adding the contributions from different locations in the observation region weighted by the excitation intensity,

$I(r) : \delta F_j(t) = Q_j \int I(r) \delta C_j(r, t) dr$, where Q_j is a constant that accounts for the absorption coefficient and fluorescence quantum yield of the fluorophore (component j) and properties of the optical measurement system. The durations of the fluctuations depend on the size of the volume and the diffusion rate as well as on the chemical reaction rate(s) and for nonlinear chemical reactions, the concentrations of the reactants. It is not sufficient, however, simply to measure one or a few fluctuations to derive the diffusion coefficients and rate constants because the fluctuations are stochastic. Each fluctuation can differ substantially from the others. Therefore, it is necessary to carry out a statistical analysis of the time courses of many fluctuations to obtain accurate values of transport and reaction rate parameters. The statistical analysis of the fluorescence fluctuations is most often carried out by calculating a fluorescence fluctuation autocorrelation function:

$$G_{jj}(\tau) = \frac{\langle \delta F_j(t) \delta F_j(t + \tau) \rangle}{\langle F_j(t) \rangle^2},$$

where $\langle \dots \rangle$ denotes an average over many time points and the division by $\langle F_j(t) \rangle^2$ yields an expression independent of the excitation laser power and fluorophore brightness for a single component system. The fluorescence fluctuation at some time t is multiplied by a fluctuation at a later time $t + \tau$. Averaging many of these products obtained for different times, t , yields an indication of the extent to which the fluorescence/concentration fluctuations have relaxed toward the steady state value for the system during the delay time τ . Then, carrying out this operation for a range of τ values gives a functional form for $G_{jj}(\tau)$ that is characteristic for specific transport processes and chemical reactions (Elson, 2011).

Up to now the fluorescence fluctuations and the fluctuation autocorrelation function has been discussed from a conventional chemical perspective. It may also be of interest to frame the discussion from the point of view of contemporary single molecule studies. This discussion is presented in the Appendix. The final results of the derivations, whether from a chemical or a single molecule perspective, must, of course, yield the same equations for the fluorescence fluctuation autocorrelation function.

The form of $G_{jj}(\tau)$ also depends on the 'shape' of the volume or area from which the fluorescence is detected. This is frequently approximated as a Gaussian function with a characteristic width, w , in the focal $(x - y)$ plane and a width w_z along the optical $(z -)$ axis: $I = I_0 \exp\left[-\frac{2(x^2 + y^2)}{w^2}\right] \exp\left(-\frac{2z^2}{w_z^2}\right)$. As an example, for fluorescent compound j with diffusion coefficient D_j that is diffusing in a two-dimensional plane, for example, a biological membrane:

$$G_{jj}(\tau) = \frac{G_{jj}(0)}{1 + \frac{\tau}{\tau_{Dj}}}; \text{ the characteristic time is } \tau_{Dj} = \frac{w^2}{4D_j}.$$

For chemical reaction kinetics $G_{jj}(\tau)$ can have a number of exponential relaxation components. The amplitude of the correlation function depends only on the concentration and size of the observation area/volume defined by the focused laser. The relative size of the fluctuations is determined by a Poisson distribution that governs the number of molecules in an open volume. Thus, $G_{jj}(0) = \frac{1}{\langle N_j \rangle}$, where for a two-

dimensional system, for example, $\langle N_j \rangle = \pi w^2 \langle C_j \rangle$, the average number of fluorescent molecules of the j 'th chemical species in the observation area πw^2 (Elson, 2013a,b). This property of $G(0)$ provides an absolute measurement of the concentration of the fluorescent component and also the 'brightness' $\langle Q_j \rangle = \frac{\langle F_j \rangle}{\langle N_j \rangle} = G(0) \langle F \rangle$, the average number of fluorescence photons detected per unit time per fluorescent molecule. Many reviews and primary articles provide more details about the forms of correlation functions for different dynamic processes (e.g., Elson and Magde, 1974; Digman and Gratton, 2011; Elson, 2011, 2013a,b; Ries and Schwille, 2012). In systems with M fluorescent components, the contributions of all the fluorophores must be included in the correlation function. The total fluorescence fluctuation is $\delta F(t) = \sum_{j=1}^M \delta F_j(t)$ and the correlation function is

$$G(\tau) = \frac{\sum_{j=1}^M \sum_{k=1}^M \langle \delta F_j(t) \delta F_k(t + \tau) \rangle}{\left(\sum_{j=1}^M \langle F_j(t) \rangle \right)^2}.$$

For the simple example of a multicomponent system in which the only dynamic process is the independent diffusion of noninteracting molecules $\langle \delta F_j(0) \delta F_k(\tau) \rangle = 0$ for $j \neq k$. Then,

$$G(\tau) = \frac{\sum_{j=1}^M \frac{Q_j^2 \langle C_j \rangle}{1 + \frac{\tau}{\tau_{Dj}}}}{\pi w^2 \left(\sum_{j=1}^M Q_j \langle C_j \rangle \right)^2}.$$

If system components interact, as in chemical reactions, the situation is more complicated, especially if the kinetics of the reactions must be taken into account (Elson and Magde, 1974). For example, for the simplest reaction, $A \xrightleftharpoons[k_b]{k_f} B$, in a closed system the chemical relaxation can be described in simple terms $\delta C_B(t) = \delta C_B(0) \exp(-Rt) = -\delta C_A(t)$, where $R = k_f + k_b$. In contrast for an FCS measurement in the open observation region, the kinetics must be analyzed as a two component coupled reaction-diffusion system because the A and B molecules can independently diffuse into and out of the observation region (Elson and Magde, 1974). Although the reaction-diffusion systems are more complicated to analyze, there are standard methods for solving the equations describing them, and so the mathematic analysis does not present a fundamental problem. More important for measuring chemical kinetics by FCS is to find a substantial change in fluorescence that indicates the reaction progress.

Relationship between FCS and Fluorescence Photobleaching Measurements

Both the interpretive theory (Elson, 1985) and the implementation of measurements on a confocal fluorescence microscope (Koppel *et al.*, 1976) are very similar for FCS and fluorescence photobleaching. The main difference in principle between the two kinds of measurements is that FCS relies on spontaneous microscopic concentration (fluorescence) fluctuations in mesoscopic systems, i.e., systems in which there are so few molecules that the concentration fluctuations are readily measured compared to the mean concentrations. In contrast, photobleaching recovery uses a pulse of intense light to photolyze fluorophores irreversibly within a small

observation volume. This produces a localized macroscopic concentration gradient, i.e., a relatively large difference between the concentration in the observation volume and the mean concentration in the surrounding regions of the system. The measured rate of relaxation of the concentration in the observation volume back to its equilibrium value due to diffusion or chemical reaction then provides the desired kinetic parameters. For FCS many stochastic fluorescence fluctuations, no matter how accurately each is measured, must be analyzed statistically via the correlation function to obtain accurate values of phenomenological diffusion coefficients and rate constants. In contrast, in a photobleaching measurement, a single macroscopic relaxation is sufficient to provide the phenomenological coefficients to an accuracy determined by the accuracy of the single measurement. The theoretical basis for supposing that the coefficients determined by each approach are the same is Onsager's regression hypothesis (Onsager, 1931). During its early days FCS was difficult to use to measure processes in cells because the signals were relatively small and could be easily corrupted by cellular motions and other changes of cell properties during the time required to acquire a sufficient number of fluorescence fluctuations. Photobleaching was then more useful for measurements on live cells (e.g., Peters *et al.*, 1974; Edidin *et al.*, 1976; Schlessinger *et al.*, 1976a,b). For the last two decades, however, due to technological advances, especially the use of confocal microscopy to minimize the observation volume (Koppel *et al.*, 1976; Rigler, 1995), FCS has also proved to be useful for studies in living cells.

Initial FRAP measurements were carried out using a non-scanning confocal microscope. Fluorescence emission was measured from a single diffraction-limited spot in the sample defined by the incident laser beam and the confocal aperture. The duration of the bleach pulse, controlled by a fast mechanical or acousto-optical shutter could be very brief ($\sim < 1$ ms) and data could be essentially continually recorded from the single spot (Koppel *et al.*, 1976). More recently confocal laser scanning microscopes have been used for photobleaching measurements (Braga *et al.*, 2004; Seiffert and Oppermann, 2005; Kang *et al.*, 2012). Although this approach has the advantages of availability on a commercially available platform and of the possibility of bleaching and acquiring recovery data from arbitrarily shaped observation regions, very fast processes are less accessible by this approach due to its dependence on the scan speed of the microscope. Moreover, the interpretation of the recovery can be complicated by the need to account for the scanning rate of the microscope (e.g., Braeckmans *et al.*, 2003).

FCS as a Single Molecule Approach

FCS provides a bridge between population and single-molecule approaches. Simple diffusion measurements are instructive. Because FCS measurements are carried out in very dilute solutions, interactions among diffusing fluorophores are minimal, and so the solution is essentially ideal. As a result a molecule moving in this kind of system correlates only with itself. (In this respect FCS is a single molecule measurement.) It is also worth noting that increasing the sensitivity of

measurements to the level of detecting a single fluorophore by minimizing the observation volume (and therefore the interfering background fluorescence) was a major step in facilitating FCS measurements (Mets and Rigler, 1994; Rigler, 2009).

Experimental Considerations

Instrumentation and Observation Volume

The primary task in an FCS measurement is to record the number of photon counts measured during each of a sequence of fixed time intervals or 'bins.' Subtraction of the mean photon count from the counts for each bin yields the time sequence of count fluctuations, the digital representation of the fluctuations of the theoretical continuous fluorescence intensity. The experimental correlation function is calculated from these photon count fluctuations. Fitting the theoretical correlation function, calculated as described above, to the experimental correlation function tests the adequacy of the model and yields values for the model's parameters.

At a minimum an instrument for measuring FCS requires an excitation light source (typically a laser), an optical system to convey the excitation light to the sample and the emitted fluorescence to a detector such as a photomultiplier or avalanche photodiode, and the means to compute the correlation function from record of the fluctuating fluorescence intensity. The optical system is typically a confocal fluorescence microscope (Koppel *et al.*, 1976; Krichevsky and Bonnet, 2002). Very capable and versatile systems for FCS measurements are available commercially from suppliers such as Zeiss, Leica, Olympus, ISS, and others. A critical function of the optical system is to determine the shape of the observation volume or area, i.e., the region of the sample from which emitted fluorescence is detected. This volume is determined by the set of lenses and apertures that shape the excitation beam and by the confocal aperture that determines what emitted photons reach the detector (Koppel *et al.*, 1976; Qian and Elson, 1991). It is common to approximate the observation volume as a three-dimensional Gaussian function defined by beam width, w , in the focal plane and a different beam width, w_z , along the optical(z)-axis, as defined above. This yields a simple correlation function for diffusion of a single fluorophore component (Aragon and Pecora, 1976):

$$G(\tau) = \frac{1}{\langle N \rangle} \frac{1}{\left(1 + \frac{\tau}{\tau_D}\right)} \frac{1}{\left(1 + \frac{\tau}{\tau_{Dz}}\right)^{\frac{1}{2}}}.$$

As before $\tau_D = \frac{w^2}{4D}$ and $\tau_{Dz} = \frac{w_z^2}{4D}$. This approximation can lead to inaccurate measurements for observation volumes defined by diffraction limited laser illumination (Hess and Webb, 2002). One way to minimize this effect is to use two-photon excitation (2PE) to define the observation volume (Berland *et al.*, 1995; Schwille *et al.*, 1999). In addition to providing an observation volume that conforms more closely to a three-dimensional ellipsoidal Gaussian (Hess and Webb, 2002), 2PE also can penetrate deeper into cells while exposing them to less damaging low wavelength excitation light outside the observation volume.

Nanophotonic and Super-Resolution Approaches

The size of the observation region, defined by the values of w and w_z , determine the spatial resolution of an FCS measurement. Because of diffraction in the range of visible light $w > \sim 300 - 250$ nm and $w_z > \sim 500 - 700$ nm. It would be desirable to reduce the values of w and w_z to make possible measurements in small compartments such as endocytic vesicles, to reduce background fluorescence, to enable measurements at higher fluorophore concentration, and to diminish the time needed to measure very slow diffusion. For more than 30 years total internal reflection microscopy (TIRFM) has been used to confine the observation region to within very small distances from a glass substratum for both FCS and photobleaching measurements (Thompson *et al.*, 1981; Lieto *et al.*, 2003; Ohsugi *et al.*, 2006; Vobornik *et al.*, 2008). TIRFM does not, however, confine the observation region parallel to the optical axis. Among several methods developed to provide 'super resolution' (Toomre and Bewersdorf, 2010), stimulated emission depletion (STED) microscopy holds great promise for FCS measurements (Hell and Wichmann, 1994). The particular advantage of STED microscopy is that provides detection volumes with linear dimensions < 50 nm while retaining the ability to make rapid measurements (< 0.1 ms). For example, STED FCS measurements have revealed that cholesterol-rich nanodomains (< 20 nm) transiently trap proteins anchored to a cell membrane by glycerophosphatidylinositol tethers (Eggeling *et al.*, 2009). A general discussion of the application of STED FCS to membrane structure and an extension of the approach to raster scanning have recently appeared (Hedde *et al.*, 2013; Mueller *et al.*, 2013). Further extensions of FCS technology are also possible using nanophotonics (Wenger and Rigneault, 2010).

Errors

Systematic errors in FCS measurements are consistent from one measurement to another and can result from faults in the alignment of the optics, anomalies due to photobleaching and photophysical effects such as blinking, optical saturation, and triplet state decay, mischaracterization of the shape and size of the observation volume, detector artifacts, and other sources (Ries and Schwille, 2012). Random errors due to shot noise can be minimized by increasing the rate of fluorescence photon counts acquired. Eventually, however, there will be a point of diminishing returns when increasing excitation intensity causes unacceptable photobleaching of the fluorophores. Even if shot noise were negligible and each fluorescence fluctuation could be measured with high accuracy, the accurate determination of the phenomenological rate coefficients would nevertheless require measuring many fluctuations due to the stochastic behavior of the mesoscopic molecular system being observed. Evaluation of the random noise for FCS measurements can be based on theoretical calculation of the variance of the correlation function (Koppel, 1974; Qian, 1990; Kask *et al.*, 1997; Saffarian and Elson, 2003). Another approach is to use multiple measurements to estimate variance empirically (Wohland *et al.*, 2001). An apparently good fit of a measured to a theoretical correlation function guarantees neither that the model correctly represents the experimental system nor that

the derived parameters are accurate. It is important to test the model in as many ways as possible (e.g., Saffarian *et al.*, 2004). It is also important to evaluate systematic and random errors of FCS measurements both to optimize experimental accuracy and to have a quantitative sense of the accuracy of the evaluated phenomenological parameters.

Measurement of Molecular Association and Aggregation

FCS and its extensions provide a variety of ways to detect and measure the association of different molecular species with one another and the aggregation of molecules in and on cells. An elementary approach is through the diffusion rate, which depends on the size of the diffusing particle. In a simple solution the diffusion coefficient, D is given by the Einstein relation: $D = \frac{k_B T}{f}$, where k_B , T , and f are Boltzmann's constant, the absolute temperature and the frictional coefficient, respectively. For a spherical particle of radius a in a solution of viscosity η , $f = 6\pi\eta a$, and so the diffusion coefficient varies only as the cube root of the molecular weight. Then, if two spherical molecules of the same volume combine to form a particle, also regarded as spherical, of twice the volume, the diffusion coefficient will decrease by only $\sim 26\%$ ($2^{1/3} - 1$). Unfortunately, mainly because diffusion correlation functions even for single components are spread over a wide time range (due to the dependence on $\frac{G(\tau)}{G(0)} = \frac{1}{1 + \frac{\tau}{\tau_D}}$), it is difficult to resolve diffusion coefficients that are not very different from one another. It has been shown that even with good data (high photon count rates) it is possible to distinguish two components of comparable brightness by FCS only if their diffusion times differ by at least a factor of 1.6 (Meseth *et al.*, 1999). Thus a fourfold volume/weight change would be required to detect an aggregation or association process. Nevertheless, this approach can be useful for detecting the binding of a small fast diffusing molecule to a larger slow diffusing one. Several other approaches provide more sensitive indications of binding or aggregation.

Molecular Association Measured by Two-Color Cross-Correlation Spectroscopy

Fluorescence cross-correlation spectroscopy (FCCS) measures interactions among fluorophores that can be detected at different wavelengths. Suppose that the fluorescence of components A and B can be distinguishably measured at well-separated emission wavelengths λ_A and λ_B . The cross correlation function is

$$G_{AB}(\tau) = \frac{\langle \delta F_A(t) \delta F_B(t + \tau) \rangle}{\langle F_A(t) \rangle \langle F_B(t) \rangle}.$$

If A and B molecules move independently of each other, there is no correlation, and so $G_{AB}(\tau) = 0$. If, however, A and B are bound together, their fluorescence fluctuations are correlated. Then the amplitude $G_{AB}(0)$ indicates the fraction of A and B in complex and can also provide information about the stoichiometry of the complex (Kim *et al.*, 2005). To achieve accurate results, however, there are many practical issues that must be

addressed, including optimizing the overlap of the observation volumes for the two fluorescence wavelengths, saturation of the fluorophores at high excitation intensities, cross talk between the detected emission of the two fluorophores that can produce false cross-correlation, and the effects of photobleaching (Bacia *et al.*, 2006; Bacia and Schwille, 2007).

Aggregation Measured by the Photon Count Histogram

Formation of molecular aggregates, for example, of triskelions in the formation of coated pits (Pearse and Crowther, 1987), and of polymers such as actin microfilaments and tubulin microtubules (Frieden, 1985; Cooper, 1991), and also intermediate filaments that have tissue-specific monomers (Herrmann and Aebi, 2004), are essential processes in establishing cell structure and function. As a simple formal example, consider the aggregation reaction, $nA_1 \rightleftharpoons A_n$. In contrast to the weak dependence of diffusion coefficient on molecular size, the brightness scales linearly with the number of fluorescent particles in an aggregate. The brightness of A_n , containing n fluorescent molecules ('monomers') would be n -fold greater than that of the individual monomers, A_1 , as long as there were no electronic interactions among the monomers in the aggregate that quenched or enhanced their fluorescence. Therefore, A_1 and A_n can much more readily be distinguished by their brightness than by their diffusion coefficients. As we have seen the brightness of an individual component is readily determined as $\langle Q \rangle = \langle F(t) \rangle G(0)$. For typical aggregation or polymerization systems, however, assuming that the system contains only A_1 or A_n is a gross oversimplification. Rather there will typically be not only both A_1 and A_n but also intermediate species A_j , $j = 2 \dots n-1$. Therefore one ought to determine the set $\{\langle N_j \rangle, \langle Q_j \rangle\}$, the average molecule numbers and brightnesses of all species j in the system. This can be done in principle using the photon count histogram (PCH). The PCH is the probability distribution, $P(k)$, of finding k photons in any time bin. As with the correlation function the PCH can be calculated theoretically for different models of $\{\langle N_j \rangle, \langle Q_j \rangle\}$ and then fitted to the experimental PCH to determine the optimal values of $\langle N_j \rangle$ and $\langle Q_j \rangle$. There are several practical approaches to comparing experimental data to the model calculations including via the moments of the distribution (Qian and Elson, 1990) and via direct comparison with the PCH (Chen *et al.*, 1999) or its generating function (Kask *et al.*, 1999). PCH analysis has proved useful in a number of applications including among many others, studies of the formation of oligomeric complexes in cells (Chen *et al.*, 2003), the clustering of EGF receptors in the absence of EGF (Saffarian *et al.*, 2007), the formation of virus-like capsids in infected cells (Chen *et al.*, 2009).

Although the methods derived from the PCH can in principle yield values of $\langle N_j \rangle$ and $\langle Q_j \rangle$ for systems with many components with different brightnesses and concentrations, in practice there are limits on the numbers of components for which these parameters can be accurately evaluated (Pryse *et al.*, 2012). It appears that one of the underlying factors is a kind of ambiguity that results from the Gaussian laser excitation profile almost universally used in these kinds of measurements. A dim particle passing through the center of the

laser beam will contribute a fluorescence pulse that is comparable to a bright particle passing through the dimmer parts of the beam. Computational modeling shows that this apparent ambiguity can be resolved by acquiring a sufficient number of fluctuation measurements but the large number required may be impractical to obtain, especially on labile systems such as cells.

Aggregation by Image Correlation Spectroscopy

Imaging methods can measure fluctuation amplitudes that provide information about aggregation of fluorescent cellular constituents. Similar in principle to a time record of fluorescence fluctuations, spatial fluorescence fluctuations can be measured and auto-correlated from a two-dimensional image of a cell. This approach, called Image Correlation Spectroscopy, supplies a two-dimensional spatial fluorescence fluctuation autocorrelation function $G_{ICS}(x, y)$ (Petersen *et al.*, 1993). Then, (as for the temporal autocorrelation discussed above) for a one component system, $G_{ICS}(0, 0) = \frac{1}{\langle N \rangle}$ where $\langle N \rangle$ is the average number of particles in the observation area. For a one component aggregation system in which the aggregates contain m 'monomers,' the brightness of the aggregate is $Q_m = G_{ICS}(0, 0) \langle F \rangle$ and so the number of monomers in the aggregate is $m = \frac{Q_m}{Q_1}$ where Q_1 is the brightness of the monomer. One should note that while FCS measures fluctuations over time and so requires that the fluorescent particles of interest to move through the observation region, ICS is independent of dynamic properties and can measure aggregation of immobile structures. Both static and dynamic properties of aggregates are observable using more complex image correlation methods that include spatiotemporal ICS measurements carried out on a temporal sequence of images (Kolin and Wiseman, 2007). For example, this approach can show the direction and velocity of convective transport of fluorescent particles (Hebert *et al.*, 2005). A related method been cast in a form that readily takes into account fluctuations due to detector noise and autofluorescence to provide the number and brightness of fluorescent particles taken to comprise a one-component system (Digman *et al.*, 2008a,b). As an example, this approach has been applied to the determination of the clustering of ErbB1 (the EGF receptor) and ErbB2 in the presence and absence of EGF (Nagy *et al.*, 2010).

Summary

1. FCS can be used to measure both dynamic processes such as diffusion, convection, and chemical reaction kinetics and equilibrium properties such as concentrations and extents of aggregation of fluorescent molecules in very small systems. The laser-illuminated observation region can be of diffraction limited size (femtoliters) providing very high spatial resolution that allows comparison of properties of different regions and compartments of cells. The dynamic range is very broad, from microseconds to seconds.
2. FCS is based on measuring processes in mesoscopic systems, i.e., systems that contain small numbers of the fluorescent molecules of interest. Therefore it is necessary to

acquire a large number of fluctuations that must be averaged, usually by calculating a fluorescence fluctuation autocorrelation function. The time needed for the acquisition of these fluctuations determines how rapidly the measurements can be accomplished.

3. FCS is optimal for measurements of systems with fluorophores at very low concentration (nanomolar range) and with characteristic times, for example, τ_D , ranging from microseconds to seconds. Photobleaching recovery is suited to systems with higher concentrations and with times ranging from ~ 100 ms to many seconds. In contrast to the requirement by FCS for many fluctuations to achieve accuracy, a single photobleaching recovery measurement is sufficient to determine rate coefficients to the precision of the measurement of the recovery transient. Therefore, photobleaching is advantageous for measuring on moving cells. Also photobleaching is useful for measuring the time course of exchange reactions (McNally *et al.*, 2000). Nevertheless, FCS can provide information such as absolute fluorophore concentrations and extents of aggregation that are inaccessible to photobleaching recovery.
4. FCS has given rise to a large number of extensions that are useful for specific purposes, for example, two-color FCCS to measure association of fluorescent 'monomers' and image correlation spectroscopy to measure aggregation.
5. FCS has been applied to a wide range of cell biological questions, for example, among many others, detection and measurement of receptor clustering (Saffarian *et al.*, 2007; Nagy *et al.*, 2010), probing the endocytic pathway using two-color cross-correlation spectroscopy (Bacia *et al.*, 2002), measurement of the dynamics of paxillin molecules during cell adhesion (Digman *et al.*, 2008a,b).

Appendix

We have provided the conventional definition of the correlation function in terms of chemical concentrations, but it is instructive also to consider an equivalent definition from the perspective of contemporary single molecule studies. We will be concerned with mesoscopic systems in which there are a small number of fluorescent molecules of interest the positions of which fluctuate continuously over time. We therefore adopt a probabilistic point of view: $p(r, t)$ is the probability that a fluorescent particle is at position r at time t . The fluorescence intensity measured from a single molecule can be represented as $F(t) = Q \int I(r) p(r, t) d^3 r$ where $I(r)$ is the fluorescence intensity detected from position r , and Q is a brightness parameter, i.e., the emitted intensity measured from a single fluorescent molecule, as defined above. Then, for a system that contains a single fluorescent molecule the normalized and non-normalized correlation functions, $G_1(\tau)$ and $g_1(\tau)$, respectively, are

$$\begin{aligned} G_1(\tau) &= \frac{g_1(\tau)}{\langle F(t) \rangle^2} = \frac{\langle F(t) F(t + \tau) \rangle}{\langle F(t) \rangle^2} \\ &= \frac{Q^2 \int I(r_0) I(r) p(r_0, 0) W(r|r_0, \tau) d^3 r_0 d^3 r}{\langle F(t) \rangle^2} \end{aligned}$$

In this equation, and $W(r|r_0, \tau)$ is the probability that the molecule initially at r_0 will be at position r at a later time τ . A freely diffusing molecule is equally likely to be anywhere in the sample volume, and so $p(r_0, 0) = \frac{1}{V}$, where V is the total volume of the system. The transition probability for simple diffusion with diffusion coefficient D is

$$W(r|r_0, \tau) = \frac{1}{(4\pi D\tau)^{3/2}} \exp\left[-\frac{|r - r_0|^2}{4D\tau}\right].$$

For simplicity we now restrict ourselves to two-dimensional systems; the results are readily extended to three dimensions by taking into account the detected intensity along the z -axis of light propagation. FCS measurements commonly use a Gaussian laser intensity profile, that yields a detection function represented in the focal plane as $I(r) = \exp\left(-\frac{2r^2}{w^2}\right)$ with $r = (x, y)$ and w , the radius of the Gaussian detection area. This yields,

$$\begin{aligned} g_1(\tau) &= \frac{Q^2}{V} \int \exp\left(-\frac{2r_0^2}{w^2}\right) \exp\left(-\frac{2r^2}{w^2}\right) W(r|r_0, \tau) d^2 r_0 d^2 r \\ &= \frac{Q^2}{V} \int \exp\left(-\frac{2r_0^2}{w^2}\right) \exp\left(-\frac{2r^2}{w^2}\right) \frac{1}{(4\pi D\tau)} \\ &\quad \exp\left[-\frac{|r - r_0|^2}{4D\tau}\right] d^2 r_0 d^2 r. \end{aligned}$$

Here we have used $W(r|r_0, \tau)$ for a two-dimensional system.

Extension of the correlation function for a one-particle system to a system containing N_t fluorescent particles is complicated by the fact that the probability distribution function for N particles is the N -fold convolution of the distributions for single particles. It is useful to avoid convolutions by noting that $g_1(\tau)$ can be regarded as a moment of a bivariate distribution the two dimensions of which correspond to the fluctuating fluorescence intensity record and the same record delayed by lag time τ (Melnykov and Hall, 2009). Then, one may use the additive properties of cumulants, a related type of moment, so that $g_1(\tau) = \kappa_1(\tau)$ where $\kappa_1(\tau)$ is a cumulant (Qian and Elson, 1990; Wu *et al.*, 2008; Melnykov and Hall, 2009). Since cumulants are additive, for a system containing N molecules, $\kappa_N(\tau) = N\kappa_1(\tau)$. This leads to the following expression for the correlation function:

$$\begin{aligned} g_N(\tau) &= \frac{N_d Q^2}{A_d} \int \exp\left(-\frac{2r_0^2}{w^2}\right) \exp\left(-\frac{2r^2}{w^2}\right) \frac{1}{(4\pi D\tau)} \\ &\quad \exp\left[-\frac{|r - r_0|^2}{4D\tau}\right] d^2 r_0 d^2 r. \end{aligned}$$

Here, $A_d = \pi w^2$ is the area from which the fluorescence is detected, and N_d is the number of molecules in that area. (Thus the concentration of the fluorescent particles is $C = N_d/A_d$.) It can further be shown that carrying out the integrations in this equation yields the expected form of the correlation function for simple diffusion (Elson and Magde, 1974).

Acknowledgments

This work was financially supported by NIH via grant R01 HL 109505.

See also: Imaging the Cell: Light Microscopy: Fluorescence Lifetime Imaging – Applications and Instrumental Principles

References

- Aragon, S.R., Pecora, R., 1976. Fluorescence correlation spectroscopy as a probe of molecular-dynamics. *Journal of Chemical Physics* 64 (4), 1791–1803.
- Bacia, K., Kim, S.A., Schwille, P., 2006. Fluorescence cross-correlation spectroscopy in living cells. *Nature Methods* 3 (2), 83–89.
- Bacia, K., Majoul, I.V., Schwille, P., 2002. Probing the endocytic pathway in live cells using dual-color fluorescence cross-correlation analysis. *Biophysical Journal* 83 (2), 1184–1193.
- Bacia, K., Schwille, P., 2007. Practical guidelines for dual-color fluorescence cross-correlation spectroscopy. *Nature Protocols* 2 (11), 2842–2856.
- Berland, K.M., So, P.T., Gratton, E., 1995. Two-photon fluorescence correlation spectroscopy: Method and application to the intracellular environment. *Biophysical Journal* 68 (2), 694–701.
- Braeckmans, K., Peeters, L., Sanders, N.N., De Smedt, S.C., Demeester, J., 2003. Three-dimensional fluorescence recovery after photobleaching with the confocal scanning laser microscope. *Biophysical Journal* 85 (4), 2240–2252.
- Braga, J., Desterro, J.M., Carmo-Fonseca, M., 2004. Intracellular macromolecular mobility measured by fluorescence recovery after photobleaching with confocal laser scanning microscopes. *Molecular Biology of the Cell* 15 (10), 4749–4760.
- Chen, Y., Muller, J.D., So, P.T., Gratton, E., 1999. The photon counting histogram in fluorescence fluctuation spectroscopy. *Biophysical Journal* 77 (1), 553–567.
- Chen, Y., Wei, L.N., Muller, J.D., 2003. Probing protein oligomerization in living cells with fluorescence fluctuation spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* 100 (26), 15492–15497.
- Chen, Y., Wu, B., Musier-Forsyth, K., Mansky, L.M., Mueller, J.D., 2009. Fluorescence fluctuation spectroscopy on viral-like particles reveals variable gag stoichiometry. *Biophysical Journal* 96 (5), 1961–1969.
- Cooper, J.A., 1991. The role of actin polymerization in cell motility. *Annual Review of Physiology* 53, 585–605.
- Digman, M.A., Brown, C.M., Horwitz, A.R., Mantulin, W.W., Gratton, E., 2008a. Paxillin dynamics measured during adhesion assembly and disassembly by correlation spectroscopy. *Biophysical Journal* 94 (7), 2819–2831.
- Digman, M.A., Dalal, R., Horwitz, A.F., Gratton, E., 2008b. Mapping the number of molecules and brightness in the laser scanning microscope. *Biophysical Journal* 94 (6), 2320–2332.
- Digman, M.A., Gratton, E., 2011. Lessons in fluctuation correlation spectroscopy. *Annual Review of Physical Chemistry* 62, 645–668.
- Edidin, M., Zagayansky, Y., Lardner, T.J., 1976. Measurement of membrane protein lateral diffusion in single cells. *Science* 191 (4226), 466–468.
- Eggeling, C., Ringemann, C., Medda, R., *et al.*, 2009. Direct observation of the nanoscale dynamics of membrane lipids in a living cell. *Nature* 457 (7233), 1159–1162.
- Eigen, M., De Maeyer, L.C., 1963. In: Freiss, S.L., Lewis, E.S., Weissberger, A. (Eds.), *Techniques of Organic Chemistry*, vol. 8, Part II. New York, NY: Interscience, pp. 895.
- Elson, E., Magde, D., 1974. Fluorescence correlation spectroscopy. I. Conceptual basis and theory. *Biopolymers* 13, 1–27.
- Elson, E.L., 1985. Fluorescence correlation spectroscopy and photobleaching recovery. *Annual Review of Physical Chemistry* 36, 379–406.
- Elson, E.L., 2011. Fluorescence correlation spectroscopy: Past, present, future. *Biophysical Journal* 101 (12), 2855–2870.
- Elson, E.L., 2013a. 40 Years of FCS: How it all began. In: Tetin, S.Y. (Ed.), *Methods in Enzymology*, vol. 518. Burlington: Elsevier (Academic Press), pp. 1–10.
- Elson, E.L., 2013b. Brief introduction to fluorescence correlation spectroscopy. In: Tetin, S.Y. (Ed.), *Methods in Enzymology*, vol. 518. Burlington: Elsevier (Academic Press), pp. 11–41.
- Frieden, C., 1985. Actin and tubulin polymerization: The use of kinetic methods to determine mechanism. *Annual Review of Biophysics and Biomolecular Structure* 14, 189–210.
- Hebert, B., Costantino, S., Wiseman, P.W., 2005. Spatiotemporal image correlation spectroscopy (STICS) theory, verification, and application to protein velocity mapping in living CHO cells. *Biophysical Journal* 88 (5), 3601–3614.
- Hedde, P.N., Dörlich, R.M., Blomley, R., *et al.*, 2013. Stimulated emission depletion-based raster image correlation spectroscopy reveals biomolecular dynamics in live cells. *Nature Communications* 4, 2093.
- Hell, S.W., Wichmann, J., 1994. Breaking the diffraction resolution limit by stimulated emission: Stimulated-emission-depletion fluorescence microscopy. *Optics Letters* 19 (11), 780–782.
- Herrmann, H., Aebi, U., 2004. Intermediate filaments: Molecular structure, assembly mechanism, and integration into functionally distinct intracellular Scaffolds. *Annual Review of Biochemistry* 73, 749–789.
- Hess, S.T., Webb, W.W., 2002. Focal volume optics and experimental artifacts in confocal fluorescence correlation spectroscopy. *Biophysical Journal* 83 (4), 2300–2317.
- Kang, M., Day, C.A., Kenworthy, A.K., DiBenedetto, E., 2012. Simplified equation to extract diffusion coefficients from confocal FRAP data. *Traffic* 13 (12), 1589–1600.
- Kask, P., Gunther, R., Axhausen, P., 1997. Statistical accuracy in fluorescence fluctuations experiments. *European Biophysics Journal with Biophysics Letters* 25, 163–169.
- Kask, P., Palo, K., Ullmann, D., Gall, K., 1999. Fluorescence-intensity distribution analysis and its application in biomolecular detection technology. *Proceedings of the National Academy of Sciences of the United States of America* 96 (24), 13756–13761.
- Kim, S.A., Heinze, K.G., Bacia, K., Waxham, M.N., Schwille, P., 2005. Two-photon cross-correlation analysis of intracellular reactions with variable stoichiometry. *Biophysical Journal* 88 (6), 4319–4336.
- Kolin, D.L., Wiseman, P.W., 2007. Advances in image correlation spectroscopy: Measuring number densities, aggregation states, and dynamics of fluorescently labeled macromolecules in cells. *Cell Biochemistry and Biophysics* 49 (3), 141–164.
- Koppel, D., 1974. Statistical accuracy in fluorescence correlation spectroscopy. *Physical Review A* 10, 1938–1945.
- Koppel, D.E., Axelrod, D., Schlessinger, J., Elson, E.L., Webb, W.W., 1976. Dynamics of fluorescence marker concentration as a probe of mobility. *Biophysical Journal* 16 (11), 1315–1329.
- Krichevsky, O., Bonnet, G., 2002. Fluorescence correlation spectroscopy: The technique and its applications. *Reports on Progress in Physics* 65 (2), 251–297.
- Lieto, A.M., Cush, R.C., Thompson, N.L., 2003. Ligand-receptor kinetics measured by total internal reflection with fluorescence correlation spectroscopy. *Biophysical Journal* 85 (5), 3294–3302.
- Magde, D., Elson, E.L., Webb, W.W., 1972. Thermodynamic fluctuations in a reacting system – Measurement by fluorescence correlation spectroscopy. *Physical Review Letters* 29, 705–708.
- McNally, J.G., Muller, W.G., Walker, D., Wolford, R., Hager, G.L., 2000. The glucocorticoid receptor: Rapid exchange with regulatory sites in living cells. *Science* 287 (5456), 1262–1265.
- Melnykov, A.V., Hall, K.B., 2009. Revival of high-order fluorescence correlation analysis: Generalized theory and biochemical applications. *Journal of Physical Chemistry B* 113 (47), 15629–15638.
- Meseth, U., Wohland, T., Rigler, R., Vogel, H., 1999. Resolution of fluorescence correlation measurements. *Biophysical Journal* 76 (3), 1619–1631.
- Mets, U., Rigler, R., 1994. Submillisecond detection of single rhodamine molecules in water. *Journal of Fluorescence* 4 (3), 259–264.
- Mueller, V., Honigsmann, A., Ringemann, C., *et al.*, 2013. FCS in STED microscopy: Studying the nanoscale of lipid membrane dynamics. *Methods in Enzymology* 519, 1–38.
- Nagy, P., Claus, J., Jovin, T.M., Arndt-Jovin, D.J., 2010. Distribution of resting and ligand-bound ErbB1 and ErbB2 receptor tyrosine kinases in living cells using number and brightness analysis. *Proceedings of the National Academy of Sciences of the United States of America* 107 (38), 16524–16529.
- Ohsugi, Y., Saito, K., Tamura, M., Kinjo, M., 2006. Lateral mobility of membrane-binding proteins in living cells measured by total internal reflection fluorescence correlation spectroscopy. *Biophysical Journal* 91 (9), 3456–3464.
- Onsager, L., 1931. Reciprocal relations in irreversible processes. I. *Physical Review* 37, 405–426.
- Pearse, B.M., Crowther, R.A., 1987. Structure and assembly of coated vesicles. *Annual Review of Biophysics and Biomolecular Structure* 16, 49–68.
- Peters, R., Peters, J., Tews, K.H., Bahr, W., 1974. A microfluorimetric study of translational diffusion in erythrocyte membranes. *Biochimica et Biophysica Acta* 367 (3), 282–294.
- Petersen, N.O., Hoddellius, P.L., Wiseman, P.W., Seger, O., Magnusson, K.E., 1993. Quantitation of membrane receptor distributions by image correlation spectroscopy: Concept and application. *Biophysical Journal* 65 (3), 1135–1146.

- Pryse, K.M., Rong, X., Whisler, J.A., *et al.*, 2012. Confidence intervals for concentration and brightness from fluorescence fluctuation measurements. *Biophysical Journal* 103 (5), 898–906.
- Qian, H., 1990. On the statistics of fluorescence correlation spectroscopy. *Biophysical Chemistry* 38, 49–57.
- Qian, H., Elson, E.L., 1990. On the analysis of high order moments of fluorescence fluctuations. *Biophysical Journal* 57 (2), 375–380.
- Qian, H., Elson, E.L., 1991. Analysis of confocal laser-microscope optics for 3 – Fluorescence correlation spectroscopy. *Applied Optics* 30, 1185–1195.
- Ries, J., Schwille, P., 2012. Fluorescence correlation spectroscopy. *BioEssays* 34 (5), 361–368.
- Rigler, R., 1995. Fluorescence correlations, single molecule detection and large number screening. Applications in biotechnology. *Journal of Biotechnology* 41 (2–3), 177–186.
- Rigler, R., 2009. FCS and single molecule spectroscopy. In: Graslund, A., Rigler, R., Widengren, J. (Eds.), *Single Molecule Spectroscopy in Chemistry, Physics, and Biology*. Heidelberg: Springer-Verlag, pp. 77–103.
- Rigler, R., Mets, U., Widengren, J., Kask, P., 1993. Fluorescence correlation spectroscopy with high count rate and low background: Analysis of translational diffusion. *European Biophysics Journal* 22, 169–175.
- Saffarian, S., Collier, I.E., Marmer, B.L., Elson, E.L., Goldberg, G., 2004. Interstitial collagenase is a Brownian ratchet driven by proteolysis of collagen. *Science* 306 (5693), 108–111.
- Saffarian, S., Elson, E.L., 2003. Statistical analysis of fluorescence correlation spectroscopy: The standard deviation and bias. *Biophysical Journal* 84 (3), 2030–2042.
- Saffarian, S., Li, Y., Elson, E.L., Pike, L.J., 2007. Oligomerization of the EGF receptor investigated by live cell fluorescence intensity distribution analysis. *Biophysical Journal* 93 (3), 1021–1031.
- Schatzel, K., Drewel, M., Stimac, S., 1988. Photon-correlation measurements at large lag times – Improving statistical accuracy. *Journal of Modern Optics* 35 (4), 711–718.
- Schlessinger, J., Koppel, D.E., Axelrod, D., *et al.*, 1976a. Lateral transport on cell membranes: Mobility of concanavalin A receptors on myoblasts. *Proceedings of the National Academy of Sciences of the United States of America* 73 (7), 2409–2413.
- Schlessinger, J., Webb, W.W., Elson, E.L., Metzger, H., 1976b. Lateral motion and valence of Fc receptors on rat peritoneal mast cells. *Nature* 264 (5586), 550–552.
- Schwille, P., Haupts, U., Maiti, S., Webb, W.W., 1999. Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one- and two-photon excitation. *Biophysical Journal* 77 (4), 2251–2265.
- Seiffert, S., Oppermann, W., 2005. Systematic evaluation of FRAP experiments performed in a confocal laser scanning microscope. *Journal of Microscopy* 220 (Pt 1), 20–30.
- Tanford, C., 1961. *Physical Chemistry of Macromolecules*. New York, NY: John Wiley & Sons.
- Thompson, N.L., Burghardt, T.P., Axelrod, D., 1981. Measuring surface dynamics of biomolecules by total internal reflection fluorescence with photobleaching recovery or correlation spectroscopy. *Biophysical Journal* 33 (3), 435–454.
- Toomre, D., Bewersdorf, J., 2010. A new wave of cellular imaging. *Annual Review of Cell and Developmental Biology* 26, 285–314.
- Vobornik, D., Banks, D.S., Lu, Z.F., *et al.*, 2008. Fluorescence correlation spectroscopy with sub-diffraction-limited resolution using near-field optical probes. *Applied Physics Letters* 93 (16), 163904.
- Wenger, J., Rigneault, H., 2010. Photonic methods to enhance fluorescence correlation spectroscopy and single molecule fluorescence detection. *International Journal of Molecular Sciences* 11 (1), 206–221.
- Wohland, T., Rigler, R., Vogel, H., 2001. The standard deviation in fluorescence correlation spectroscopy. *Biophysical Journal* 80 (6), 2987–2999.
- Wu, B., Chen, Y., Muller, J.D., 2008. Fluorescence correlation spectroscopy of finite-sized particles. *Biophysical Journal* 94 (7), 2800–2808.

Fluorescence Lifetime Imaging – Applications and Instrumental Principles

W Becker, Becker & Hickl GmbH, Berlin, Germany

© 2016 Elsevier Inc. All rights reserved.

Glossary

Autofluorescence Fluorescence from endogenous compound in biological material.

Fluorescence decay function The waveform of the fluorescence signal emitted from an ensemble of fluorescent molecules after excitation with a short light pulse.

Fluorescence lifetime The average time a molecule stays in the excited state after being excited by absorbing a photon.

Förster resonance energy transfer (FRET) Energy transfer from an excited molecule (the donor) into another molecule (the acceptor) located in close proximity to it.

Phasor A pointer describing phase and the modulation degree of a periodic light signal.

Time-correlated single-photon counting A technique to record the waveform of a periodical optical signal by recording single photons, measuring the time differences of these photons to a reference signal synchronous with the optical signal, and building up the distribution of the number of photons over the time difference.

Fluorescence

Techniques based on fluorescence detection have found broad application in cell biology because they are extremely sensitive and able to deliver information about biochemical interactions on the molecular scale. Fluorescence techniques obtained a significant boost by the development of labeling and transfection techniques which permit the fluorophores to be placed into defined biological targets. Cells could now be genetically manipulated to express fluorophores in exactly defined target positions within the cell's intrinsic proteins (Chalfie *et al.*, 1994; Tsien, 1998). Another improvement was the introduction of confocal and multiphoton laser scanning microscopes (Denk *et al.*, 1990; Pawley, 2006; Wilson and Sheppard, 1984). By confining the detection to a well-defined focal plane, these instruments significantly improved the specificity of the data recorded.

Fluorescence is a process in which a molecule enters an excited electronic state by absorbing an optical photon and returns from this state by emitting a photon of longer wavelength. The mechanisms involved in fluorescence are summarized in Figure 1. For a comprehensive description of fluorescence, please see Lakowicz (2006).

The ground state is S0, the first excited state S1. By absorbing a photon with an energy higher than the gap between S1 and S0, the molecule transits into the S1 state. Higher excited states, S2, S3, do exist and can be excited, but decay at an

extremely rapid rate into the S1 state. Moreover, the electronic states of the molecules are broadened by vibration. Therefore, a molecule can be excited by a photon of almost any energy higher than the gap between S0 and S1.

A molecule can also be excited by absorbing two photons simultaneously (Göppert-Mayer, 1931). The sum of the energy of the photons must be larger than the energy gap between S1 and S0. Multiphoton absorption has long been only a theoretical possibility. However, it is easily achieved in the focus of a picosecond or femtosecond pulsed laser. Multiphoton excitation has therefore become a standard technique of laser scanning microscopy (Denk *et al.*, 1990; Diaspro, 2001; Pawley, 2006).

Without interaction with its environment, the molecule can return from the S1 state by emitting a photon or by internally converting the absorbed energy into heat. Moreover, the molecule can interact with its environment and return to the ground state by transferring the energy to another molecule, see Figure 1, left. Effects that let molecules lose their excitation energy to their environment are usually summarized by the term 'fluorescence quenching.'

Figure 1 shows that molecules that absorb light are, in principle, also able to emit fluorescence. How much fluorescence is emitted for a given amount of absorbed light depends on the balance of the radiative and non-radiative decay rates. Molecules in which the emission of fluorescence is the dominating decay path are called fluorophores.

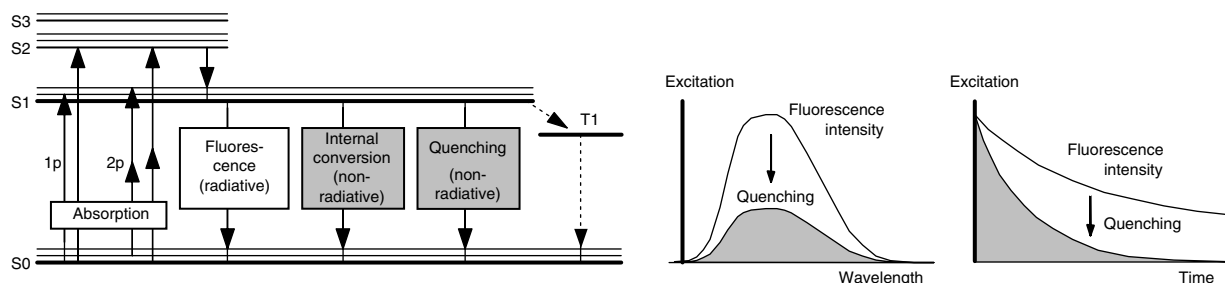


Figure 1 Left: General mechanisms involved in Fluorescence. Middle: fluorescence spectrum. Right: Fluorescence decay.

An excited molecule can also transit into a triplet state, T1. The 'intersystem crossing' rate is normally much smaller than the S1–S0 transition rate. Except for a few special cases intersystem crossing has therefore little influence on the fluorescence. However, the triplet state is a starting point for photochemical reactions. Excessive population of triplet states may therefore have an impact on the viability of cells and on the photostability of fluorophores.

If an ensemble of identical molecules is excited, individual molecules end up in different vibrational levels of the ground state. Thus, the fluorescence spectrum has a nonzero width, see [Figure 1](#), middle. The center wavelength and the shape of the spectrum are characteristic of the type of the molecule. Because emission, internal conversion, and quenching are competing, the fluorescence intensity changes with the rate constants.

If the same ensemble of molecules is excited with short light pulses the fluorescence intensity will rise with the excitation pulse and then decay at a rate equal to the sum of the radiative and non-radiative decay rates, see [Figure 1](#), right. The average time the molecules stay in the excited state is called the 'fluorescence lifetime.' Consequently, fluorescence signals are characterized by three parameters:

The 'intensity' of the fluorescence depends on the ratio of the radiative and non-radiative decay rates and thus on the molecular environment of the fluorophore. However, it also depends on the concentration of the fluorophore. Changes in the decay rates therefore cannot easily be distinguished from changes in the concentration. Images of the fluorescence intensity do, however, show where in a sample a fluorophore is located, and show the spatial structure of the specimen.

The 'fluorescence spectrum' is determined by the S1–S0 energy difference, and the width of the vibration bands. It is thus characteristic of the fluorophore. The shape of the fluorescence spectrum is, within reasonable limits, independent of the fluorophore concentration. Images containing spectral information thus allow the fluorophores in the individual pixels of an image to be identified.

The 'fluorescence lifetime' or, more exactly, the 'fluorescence decay function,' depends on the type of the fluorophore but not on its concentration. It depends also on the molecular environment of the fluorophore. It is influenced by the presence of fluorescence quenchers and their local concentration, by binding of the fluorophore to different biological targets, by the folding state of the binding targets, or by the presence of other optical absorbers which it may interact with by Förster energy transfer ([Förster, 1948, 2012](#)). Fluorescence decay functions and, consequently, fluorescence lifetime images, can therefore be used to obtain information on the molecular environment of the fluorophore molecules ([Becker, 2012](#); [Berezin and Achilefu, 2010](#); [Lakowicz, 2006](#)).

There are a number of different techniques to record fluorescence decay functions, and to combine fluorescence decay recording with imaging. The techniques can be classified into time-domain and frequency-domain techniques, photon counting and analog techniques, and point-scanning and wide-field imaging techniques. It also matters whether a technique acquires the signal waveform in a few time gates ([Buurman *et al.*, 1992](#)) or in a large number of time channels ([Becker *et al.*, 2004](#)), and whether this happens simultaneously ([Becker *et al.*, 2004](#); [Becker, 2005](#); [Buurman *et al.*, 1992](#)) or

sequentially ([Dowling *et al.*, 1997](#); [Straub and Hell, 1998](#)). Virtually all combinations are in use. This leads to a wide variety of instrumental principles. Different principles differ in their photon efficiency, i.e., in the number of photons required for a given lifetime accuracy ([Gerritsen *et al.*, 2002](#); [Köllner and Wolfrum, 1992](#); [Philip and Carlsson, 2003](#)), the acquisition time required to record these photons, the photon flux they can be used at, their time resolution, their ability to resolve the parameters of multi-exponential decay functions, multi-wavelength detection capability, optical sectioning capability, and compatibility with different imaging and microscopy techniques ([Becker and Bergmann, 2008](#)). The most frequently used fluorescence lifetime imaging (FLIM) techniques are described in section 'Technical Aspects of FLIM.' For the examples given in the section below Time-domain FLIM by time-correlated single-photon counting and laser scanning was used. The technique is not only the most sensitive one but also demonstrates most clearly the dependence of the fluorescence decay behavior on the molecular environment of the fluorophores.

FLIM Applications in Life Sciences

Measurement of Local Molecular Environment Parameters

There is a wide range of mechanisms for environment-induced fluorescence-lifetime changes ([Lakowicz, 2006](#)). The most relevant ones for biological imaging are fluorescence quenching by ions, influence of binding to DNA and RNA ([Van Zandvoort *et al.*, 2002](#)), binding to proteins, influence of the folding of the proteins the fluorophore is attached to, and Förster resonance energy transfer (FRET).

Another mechanism of lifetime changes is that a fluorophore has two forms with different fluorescence lifetimes. This is the case for large group of sensors for Ca^{2+} , Mg^{2+} , Na^{+} , and K^{+} ions. These fluorophores have a cation-bound and a cation-unbound form. The two forms have different fluorescence quantum yields, and thus different lifetimes. The apparent lifetime therefore depends on the concentration of the cation. pH sensors have a protonated and a deprotonated form. The forms have different fluorescence lifetimes. The apparent lifetime thus changes with the concentration ratio of the forms, and thus depends on the pH ([Hanson *et al.*, 2002](#); [Sanders *et al.*, 1995](#)).

There are also effects of local viscosity ([Kuimova *et al.*, 2008](#)), local refractive index ([Strickler and Berg, 1962](#); [Tregido *et al.*, 2008](#)), proximity to nanoparticles and metal surfaces ([Geddes *et al.*, 2003](#); [Muddana *et al.*, 2009](#); [Ritman-Meer *et al.*, 2007](#)), and aggregation of the fluorophore ([Kelbauskas and Dietel, 2002](#)).

[Figure 2](#), left, shows a sample of human skin stained with the pH sensor BCECF (2',7'-bis-(2-carboxyethyl)-5-(and-6)-carboxyfluorescein). The color of the image shows the fluorescence lifetime (mean lifetime of double-exponential decay, see 'FLIM Data Analysis'). Fluorescence decay curves in two selected pixels are shown on the left of the image.

[Figure 2](#), right, shows a barley root stained with the Ca^{2+} sensor Calcium Orange. A lifetime change due to different Ca^{2+} concentration is clearly visible. Fluorescence decay

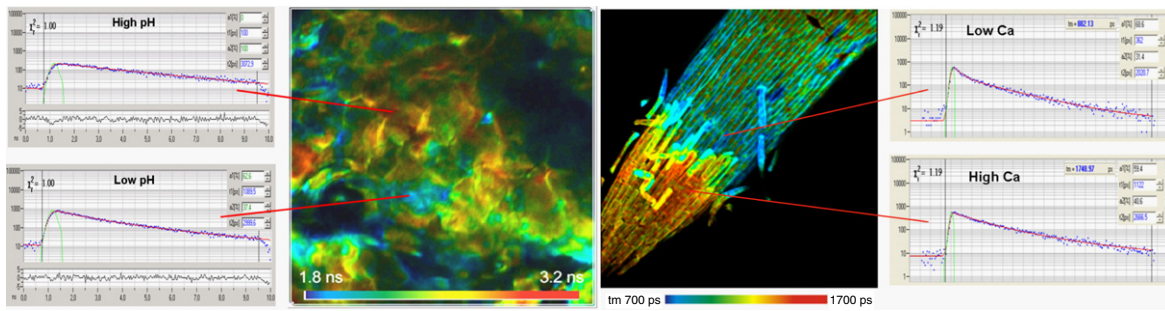


Figure 2 Left: Lifetime image of skin tissue stained with BCECF. The lifetime is an indicator of the pH. (Theodora Mauro, University of San Francisco, Zeiss LSM 510 NLO with bh SPC-830 TCSPC FLIM module.) Right: Barley root tip, stained with Calcium Orange. Courtesy of Feifei Wang, Zhonghua Chen and Anya Salih, Confocal Bioimaging Facility, University of Western Sydney, Australia. Leica SP5 MP with bh SPC-150 FLIM module.

curves in two selected spots of the image are shown on the right.

Other applications are Cl^- measurements in olfactory sensory neurons (Kaneko *et al.*, 2004), detection of changes in the Cl^- concentration during chloride homeostasis (Gilbert *et al.*, 2007), and on inflammation (Funk *et al.*, 2008). The application of FLIM to determine absolute Ca^{2+} resting concentrations in neurons has been described by Kuchibhotla *et al.* (2009). The detection of transient changes of the Ca^{2+} concentration in neurons on electrical stimulation has been described in Becker *et al.* (2014).

FRET Experiments

The most frequent application of FLIM is the measurement of protein interaction and protein folding by FRET (Förster, 1948, 2012). FRET occurs when the fluorescence emission band of one molecule (the donor) overlaps the absorption band of a second one (the acceptor), and the distance between donor and acceptor is smaller than a few nanometers. The energy absorbed by the donor is then transferred into the acceptor, and emitted via the acceptor fluorescence band. The efficiency of the energy transfer is an indicator of the distance between the donor and the acceptor.

Measurements of FRET by steady-state techniques are difficult: the concentration of the donor and the acceptor changes throughout the sample. The techniques therefore depend on ratios of the donor and acceptor intensities. However, the FRET-excited acceptor intensity is not directly available. There is 'donor bleedthrough' due to the overlap of the donor fluorescence into the acceptor emission band. Moreover, an unknown amount of the acceptor fluorescence comes from direct excitation. Steady-state FRET techniques therefore require careful calibration, including measurements of samples containing only the donor and only the acceptor. Another problem is that not all donor molecules interact with an acceptor, and there is usually a mixture of interacting and noninteracting donor. It therefore cannot be determined whether a variation in the classic 'FRET efficiency' is due to a variation in the distance between donor and acceptor or in the fraction of interacting proteins.

With FLIM, the energy transfer is directly observable as a decrease in the fluorescence lifetime of the donor. Donor

bleedthrough and directly excited acceptor fluorescence have no influence on FLIM-FRET measurements. The only reference value needed is the donor lifetime in absence of the acceptor (Bacskaï *et al.*, 2003; Biskup *et al.*, 2004; Chen and Periasamy, 2004; Duncan *et al.*, 2004; Gukassyan *et al.*, 2007; Periasamy, 2001; Periasamy and Clegg, 2009). It has been shown that even the reference lifetime is not necessarily needed: a part of the donor population does not interact with an acceptor. The lifetime of the noninteracting donor fraction can be extracted from the FLIM data and used as a reference lifetime (Becker *et al.*, 2004; Becker, 2005). The obvious advantage of this approach is that the reference lifetime is obtained from the same pixels of the same cell, and thus from the same molecular environment as the FRET data. Double-exponential FRET analysis is thus, within reasonable limits, independent of lifetime variations induced by variations in the local environment (Becker, 2015a).

An example of a FLIM FRET result is shown in Figure 3. The data were obtained from a cell expressing a green fluorescent protein (GFP) fusion protein and a Cy3-labeled antibody (Bastiaens and Squire, 1999). FRET occurs between the GFP (the donor) and the CY3 (the acceptor). A lifetime image was recorded at the emission wavelength of the GFP. The data were analyzed by a double-exponential decay model. An image of the amplitude-weighted lifetime is shown in Figure 3, left, decay curves in two characteristic pixels in Figure 3, right. The donor lifetime is significantly reduced at the cell membrane, indicating that this is the place where the protein interacts with the antibody.

It should be mentioned here that a number of pitfalls that can impair the results of FRET experiments. If the labeling density or the expression level is too low the fluorescence signals may be contaminated by autofluorescence. If the expression is too high energy can migrate between the donors, and donor molecules may interact with several acceptor molecules. The FRET results then no longer represent the distance between donor and acceptor and the relative amount of interacting donor (Koushik and Vogel, 2008; Vogel *et al.*, 2006). The results can be also biased by photobleaching or photoconversion: the noninteracting donor fraction photobleaches faster, resulting in an apparent increase of the FRET efficiency. Moreover, a part of the donor may be photoconverted and change its fluorescence lifetime (Hoffmann *et al.*,

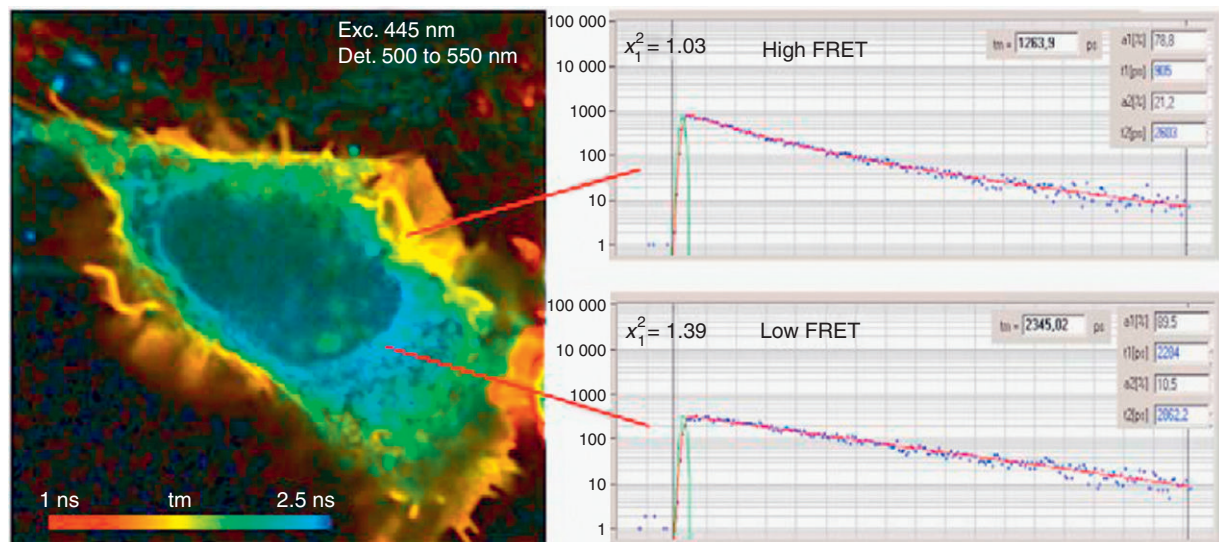


Figure 3 Cell expressing a GFP fusion protein (donor) and Cy3-labeled antibody (acceptor). Donor image, analysis by double-exponential decay model. Zeiss LSM 710 microscope with Becker and Hickl Simple-Tau 152 FLIM system.

2008). Massive changes in the fluorescence quantum efficiencies and the decay profiles of fluorescent proteins are introduced by fixation techniques (Becker, 2015a). FRET measurements should therefore always be performed on live cells. It should be noted that these problems are not specific of FLIM FRET experiments; they exist also for intensity-based techniques. In any case, FLIM measurements have the advantage that the nature of the problem can be more easily be identified.

An enormous number of FLIM-FRET papers have been published in the last ten years, most of them using time-correlated single-photon counting (TCSPC) FLIM. Please see Becker (2015a) for a review of the FLIM-FRET literature and for further details.

Autofluorescence FLIM

The use of autofluorescence for imaging of biological objects has the obvious benefit that the techniques can directly be transferred into clinical applications, see section below. The problem of steady-state autofluorescence measurement is that biological tissue contains a wide range of fluorophores (König and Riemann, 2003; Richards-Kortum *et al.*, 2003; Schweitzer *et al.*, 2004). The fluorescence spectra overlap, and their apparent shape can be changed by unknown absorbers in the tissue. The evaluation of the tissue state based on intensity imaging and spectral imaging is therefore difficult.

FLIM is often considered only a way to improve the contrast of separation of the different fluorophores. The most important argument for using lifetime imaging is, however, that the lifetimes of many endogenous fluorophores depend on local environment parameters, such as pH, oxygen saturation, on binding to proteins, and, most importantly, on metabolic activity. Autofluorescence lifetime detection therefore not only adds an additional separation parameter but also yields direct information about the metabolic state of the cells

and the microenvironment of the endogenous fluorophores (Bird *et al.*, 2005; Chorvat and Chorvatova, 2006, 2009; Chorvatova *et al.*, 2012; Gukassyan and Kao, 2009; Lakowicz *et al.*, 1992; Paul and Schneckenburger, 1996; Skala *et al.*, 2007a,b, 2010a,b; Szaszak *et al.*, 2011; Wang *et al.*, 2008).

An example is shown in Figure 4. It shows cultured human stem cells excited by two-photon excitation at 750 nm. The detection wavelength was 450–480 nm. Under these conditions, the fluorescence comes predominately from nicotinamide adenine dinucleotide (NADH). The protein-bound form of NADH has a fluorescence lifetime of about 2.5 ns, the unbound form a lifetime of about 500 ps (Lakowicz *et al.*, 1992). The two decay components can be separated by double-exponential decay analysis, see ‘FLIM Data Analysis’. The ratio of the amplitudes of the decay components, a_1 and a_2 , represent the ratio of unbound and bound NADH. It can clearly be seen in Figure 4 that the more differentiated cells have smaller a_1/a_2 (unbound/bound) ratios and longer mean lifetime, t_m . Please see König *et al.* (2011) and Guo *et al.* (2008) for further details.

Autofluorescence FLIM images can be surprisingly rich in detail. An example is shown in Figure 5. The images show a pig skin sample excited by two-photon excitation at 800 nm. The left image shows the wavelength channel below 480 nm. This channel contains both fluorescence and second harmonic generation (SHG) signals (Masters and So, 2008). The SHG signal comes from the collagen fibrils of the skin. The amount of collagen (and thus the relative intensity of SHG signals) can be used to quantify skin ageing (Koehler *et al.*, 2006). The SHG fraction of the signal has therefore been extracted from the FLIM data and displayed by color. The right image is from the channel >480 nm. It contains only fluorescence, the color corresponds to the amplitude-weighted mean lifetime of a double-exponential decay model. The fluorescence comes from a variety of compounds, mainly NADH, flavin adenine dinucleotide (FAD), melanin, and porphyrins, in the cells of the dermis.

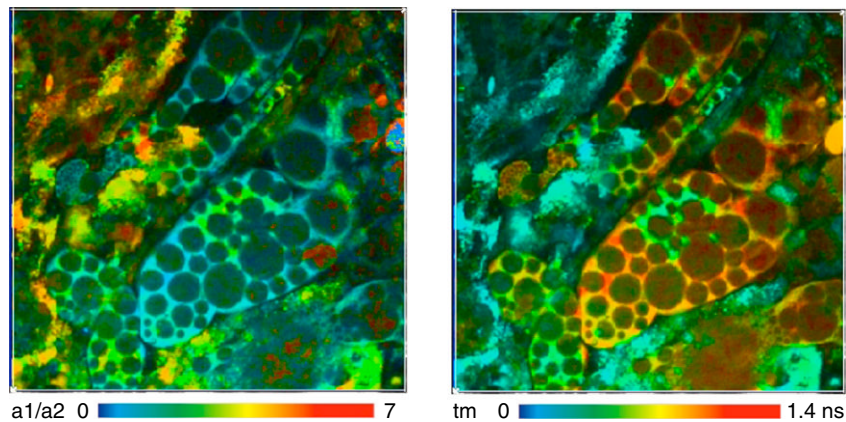


Figure 4 FLIM of human salivary gland stem cells. Differentiated cells have significantly lower a_1/a_2 ratio and longer mean lifetime. Data courtesy of Aisada Uchugonova and Karsten König, Saarland University, Saarbrücken.

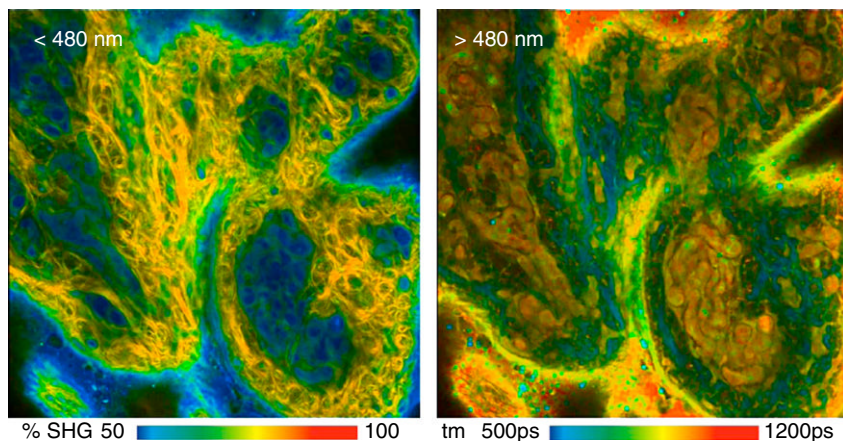


Figure 5 Two-photon FLIM of pig skin. Two-photon excitation, non-descanned detection. Left: Wavelength channel < 480 nm, color shows percentage of SHG in the recorded signal. Right: Wavelength channel > 480 nm, color shows amplitude-weighted mean lifetime. Microscope Zeiss LSM 710 NLO, with bh Simple-Tau 152 FLIM system.

Clinical FLIM Applications

Multiphoton Tomography of Skin

Multiphoton tomography of human skin uses scanning by a focused femtosecond laser beam, two-photon excitation, and non-descanned detection of the fluorescence signals. The three-dimensional tissue structure can be reconstructed from the data at subcellular resolution. The technique goes back to the work of Gratton, König, Masters, So, and Tromberg who showed that *in-vivo* two-photon autofluorescence imaging of cells and, especially, human skin is possible without impairing viability (König, 2008a; Masters *et al.*, 1997; Masters and So, 1999, 2008). Because the technique is based on fast scanning and pulsed excitation it can easily be combined with TCSPC FLIM. Instruments for clinical application have been developed by Jenlab GmbH, Jena, Germany. Figure 6 shows the stratum granulosum of a human patient recorded with a Jenlab Dermainspect system and a Becker and Hickl SPC-152 FLIM system. For interpretation of skin FLIM images please see Sanchez *et al.* (2015).

An impressive number of papers have appeared related to FLIM data obtained with these instruments, please see

Dimitrow *et al.* (2009), Geusens *et al.* (2010) Kantelhardt *et al.* (2007), König and Riemann (2003), König and Uchugonova (2009), Leppert *et al.* (2006), Lin *et al.* (2011), Patalay *et al.* (2011), Prow *et al.* (2011), Roberts *et al.* (2008, 2011), Sanchez *et al.* (2010), and Szaszak *et al.* (2011). An overview on clinical multiphoton tomography of skin is available in König (2008b) and in Sanchez *et al.* (2015).

Ophthalmic FLIM

FLIM of the human ocular fundus is obtained by combining an ophthalmic scanner with one or several ps diode lasers, and a TCSPC FLIM system. The problems of ophthalmic FLIM are that the excitation power is limited by eye safety considerations, and that motion of the eye during image acquisition must be taken into account. For technical details please see Becker (2015a) and Schweitzer *et al.* (2004, 2007). Ophthalmic FLIM is currently undergoing clinical trials.

Changes in the decay functions can result both from changes in the tissue constitution, and from changes in the local environment of the fluorophores. Distinct signatures in

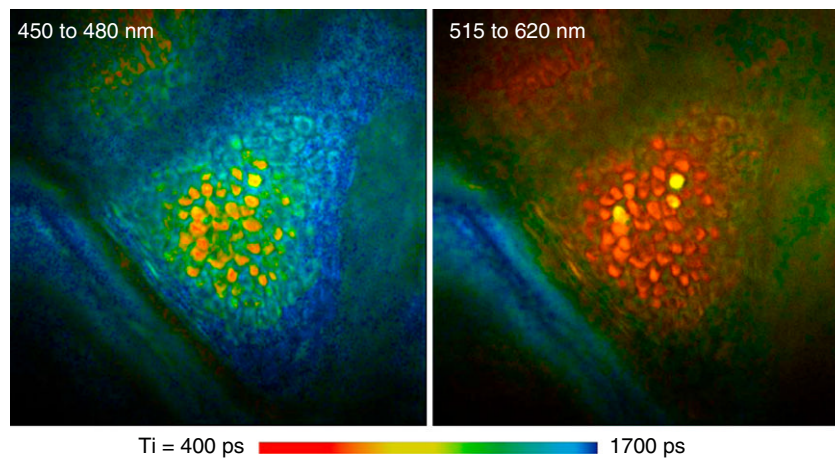


Figure 6 FLIM image of human skin recorded by a Jenlab Dermainspect Multiphoton Tomograph and a Becker and Hickl SPC-150 TCSPC FLIM system. Two-photon excitation at 750 nm. Data courtesy of Michael Roberts and Washington Sanchez, Brisbane, Australia.

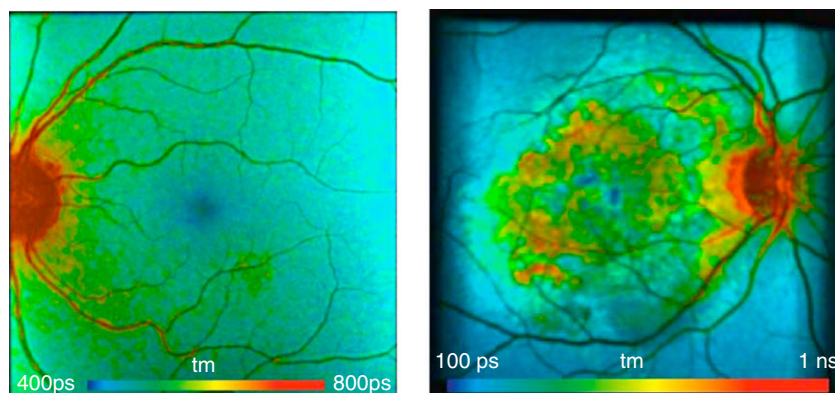


Figure 7 Left: lifetime image of the fundus of a healthy eye. Right: lifetime image of the fundus of an eye suffering from AMD. Amplitude-weighted lifetime of double-exponential decay. Data courtesy of Dietrich Schweitzer, Friedrich Schiller University Jena, Germany, and Heidelberg Engineering, Heidelberg, Germany.

the decay parameters were found for early and late age-related macula degeneration (AMD), for under-supplied regions after arterial branch occlusions, for metabolic alterations due to diabetes mellitus, and for extra-macular drusen (Schweitzer *et al.*, 2009, 2012; Schweitzer, 2009, 2010).

Typical results are shown in Figure 7. The images were scanned by a FLIO (fluorescence lifetime ophthalmoscope) laser ophthalmoscope of Heidelberg Engineering, Heidelberg, Germany. The system uses Becker and Hickl SPC-150 TCSPC FLIM modules to record the FLIM data.

Figure 7, left, shows the ocular fundus of a healthy eye. The image does not show significant pathological alterations. For comparison, a FLIM image of the ocular fundus of a patient with wet AMD is shown in Figure 7, right.

Technical Aspects of FLIM

Requirements to a FLIM Technique for Cell Biology

Photon efficiency

FLIM usually requires more photons than intensity imaging. The reason is not that FLIM is 'inefficient,' as it is sometimes

believed. The difference simply results from the fact that FLIM is aiming at recording small changes in the fluorescence behavior of a sample, whereas intensity imaging is used to record the spatial structure of a sample. Lifetime changes in FLIM experiments can be orders of magnitude smaller than (spatial) intensity changes in intensity images. The required signal-to-noise ratio is therefore much higher for FLIM. Both for lifetime recording and intensity recording, the signal-to-noise ratio is the square root of the number of photons per pixel (Gerritsen *et al.*, 2002; Köllner and Wolfrum, 1992). Consequently, typical FLIM experiments need more photons than spatial imaging. Remarkably, the 'efficiency' difference of FLIM and intensity imaging disappears when intensity data are used to derive molecular parameters, as has been demonstrated for FRET experiments (Pelet *et al.*, 2006).

The importance of photon efficiency becomes even more obvious if the samples are considered. Intensity imaging is usually performed on samples stained with high amounts of fluorophores. In contrast, most FLIM are performed on fluorophores attached to highly specific targets in the cells. That means the effective fluorophore concentration is low. The fluorophores have therefore to perform more

excitation-emission cycles, and thus photobleach faster. Moreover, FLIM measurements are usually performed on live specimens. High excitation power easily causes metabolic changes or even damage to the cells. To keep the exposure as low as possible it is essential that the FLIM technique obtains the best possible lifetime accuracy from a given number of photons detected. Photon efficiency (Geritsen *et al.*, 2002; Philip and Carlsson, 2003), i.e., the number of photons required to obtain a given lifetime accuracy, is therefore the most critical parameter of a lifetime imaging technique for biological application.

Single-photon capability

Biological samples, especially live specimens investigated under a microscope, deliver fluorescence intensities no higher than a few million photons per second. Under these conditions, the detector signal is a train of random pulses representing the detection of the single photons (Becker, 2005). The FLIM technique has to efficiently reconstruct the decay functions in the pixels of the image from the single-photon pulses of the detector.

Tolerance to dynamic effects in the sample

Fluorescence signals from live specimens cannot be expected to remain perfectly stable over the acquisition time of the image. A part of the fluorophore may be photobleaching, there may be motion in the samples, and the cells may respond to light-induced stress by changes in their metabolism.

Such changes should not induce direct artifacts, such as distortions in the decay curves, fluorescence spectra, or images recorded. Procedures like slow scanning of time gates or spectral scanning are therefore problematic. If spatial scanning is used, it will be shown that, it has to be faster than the changes in the sample.

Resolution of complex decay profiles

An ensemble of similar molecules in a homogeneous molecular environment delivers a single-exponential fluorescence decay function characterized by a single fluorescence lifetime. In cells and biological tissue the situation is different. The molecular environment of the fluorophore is rarely homogeneous, there may be several fluorophores of different fluorescence lifetime, or different fractions of the same fluorophore may be bound to different molecular targets. Examples are FRET experiments where donor molecules interacting with an acceptor have shorter fluorescence lifetimes than non-interacting ones. The decay profiles then become multi-exponential. Under these conditions, the lifetimes and amplitudes of the individual decay components provide better access to molecular parameters than an average 'lifetime' of the decay function. A FLIM technique for biological application should therefore be able to resolve complex decay profiles into their components.

It should be noted that the number of photons required for a multi-exponential decay analysis is higher than for single-exponential analysis. Köllner and Wolfrum (1992) estimated the number of photons required to resolve a double-exponential decay to be $N=400\,000$. A number of photons per pixel this high is, of course, entirely beyond the limits set by the photostability of a biological sample. Fortunately, the prospects of separating two lifetime components improve dramatically with the ratio of the two lifetimes and with the

amplitude factor of the short lifetime component. The lifetime components assumed by Köllner and Wolfrum (1992) were 10% of 2 ns and 90% of 4 ns. This is an extremely unfavorable situation which indeed requires an extremely high number of photons. Fortunately, the decay profiles encountered in FRET and autofluorescence measurements have a much more favorable composition. Usually the lifetime components are separated by a factor of 5–10, and the amplitude of the fast component is 50–90%. Under such conditions double or even triple exponential analysis is feasible with no more than a few 1000 photons per pixel, and very satisfactory results are obtained from 10 000 photon per pixel.

Time resolution

It is sometimes believed that FLIM does not require a particularly high-time resolution. It is certainly correct that the fluorescence lifetimes of typical fluorophores used as fluorescent labels for cell imaging are on the order of a few ns. However, this does not mean that there are no shorter lifetimes or shorter lifetime components. The lifetime of fast decay components of tissue autofluorescence can be 100 ps or shorter (Schweitzer *et al.*, 2007). The lifetime of the quenched donor fraction in FRET experiments can be as short as 300 ps. Lifetimes of dye aggregates in cells have been found as short as 50 ps (Kelbauskas and Dietel, 2002). The lifetime of fluorophores bound to metallic nanoparticles (Ceddes *et al.*, 2003) can be 100 ps and shorter. To resolve such lifetimes at reasonable efficiency from a multi-exponential decay an instrument response width of 100–200 ps full width at half maximum (FWHM) or shorter is desirable.

Suppression of out-of-focus signals

The optical system used in combination with the FLIM system should suppress the excitation and/or the detection of fluorescence outside the focal plane. It should also avoid the detection of laterally or longitudinally scattered light. Such unwanted signal components would not only blur the images and reduce the image contrast but also mix the decay profiles from different pixels and sample planes. Resolving several decay components present in the same pixel is difficult enough. Contamination of the decay data with out-of-focus signals would change the situation virtually hopeless. The only optical technique that efficiently rejects out-of-focus signals without compromising in-focus efficiency is laser scanning. Depth resolution is either obtained by confocal detection or by multiphoton excitation.

Fast scanning

As mentioned above, scanning must be fast enough to avoid artifacts caused by dynamic effects in the sample. There is also a practical requirement for fast scanning: live samples degrade quickly. There must be a possibility of fast live display for focusing, sample positioning, and adjusting the scan area to the region of interest. This requires a frame rate of at least one frame per second.

Live cell compatibility

Fixation techniques commonly used in microscopy coagulate the proteins or otherwise influence the molecular structure of the sample. In other words, fixation destroys exactly what is to

be investigated by FLIM. The vast majority of FLIM measurements are therefore performed in live cells or tissues. Live cell compatibility means minimizing sample exposure, and minimizing motion artifacts within the acquisition time.

Multiwavelength capability

For a number of applications FLIM has to be recorded in several wavelength intervals. Photobleaching and dynamic effects in the sample normally preclude to record data in several wavelength intervals sequentially. The FLIM technique should therefore be able to record data in several detection channels simultaneously.

FLIM Techniques

Time-domain FLIM by multidimensional TCSPC and laser scanning

Fluorescence decay measurement by TCSPC is based on the excitation of a sample with short laser pulses, the detection of single photons of the fluorescence light, the measurement of the detection times of the photons after the laser pulses, and the buildup of a photon distribution over the detection time (O'Connor and Phillips, 1984). TCSPC requires that the detector detects less than about 0.1 photons per excitation pulse. This was a limitation in early TCSPC implementations with low-repetition-rate light sources. With the excitation pulse rates of 50–80 MHz delivered by modern lasers this is no longer a problem (Becker, 2005).

A more serious limitation of classic TCSPC is that it is intrinsically one-dimensional: it delivers only a waveform of the optical signal. To apply classic TCSPC to FLIM a sample would be scanned with the laser beam, the beam position held steady in each pixel until a full fluorescence decay curve has been acquired, and the process repeated for all pixels of the image. Such implementations have indeed been used (Bugiel *et al.*, 1989) but are not compatible with the fast scan rates used in modern laser scanning microscopes.

Modern TCSPC FLIM uses a multidimensional TCSPC (Becker, 2005, 2015a). The principle is shown in Figure 8. The

technique builds up a photon distribution over several parameters: the arrival times of the photons after the laser pulses, t , and the position of the laser beam, x and y , in the sample. From these parameters, a photon distribution over the spatial coordinates, x , y , and the arrival times of the photons, t , is built up. The result is a three-dimensional data array that represents the pixels of the two-dimensional scan, with each pixel containing photons in a large number of time channels for consecutive times after the excitation pulses.

As classic TCSPC, the principle features near-ideal photon efficiency (Köllner and Wolfrum, 1992; Philip and Carlsson, 2003), extremely high-time resolution, and a sufficiently large number of time channels to resolve multi-exponential decay profiles. It has also a number of additional advantages. An obvious one is that it works at any scan rate. At high scan rates, the recording process becomes more or less random. When a photon is detected it is just stored in a pixel and a time channel of the data array according to its time in the fluorescence decay and the position of the laser beam in the scan area. The buildup of the images can directly be observed and the recording be continued over a collection time – or as many frames of the scan – as to obtain the desired signal-to-noise ratio. Fast scanning also avoids excessive triplet accumulation and heat concentration in the excited spot.

The biggest advantage is, however, that FLIM by multidimensional TCSPC can be extended to record more complex photon distributions. This is achieved by adding additional parameters of the photons to the recording process. These can be the wavelength of the emitted photons, several excitation wavelengths multiplexed at high rate, the time after an external stimulation of the sample, or the time within the period of an additional modulation of the laser. As a result, the technique delivers multiwavelength FLIM data (Becker, 2005; Becker *et al.*, 2007; Rück *et al.*, 2007), FLIM data for several excitation wavelengths (Becker, 2005, 2015a), FLIM data resolving dynamic changes in the sample in the millisecond range (Becker *et al.*, 2014; Becker, 2015b), and simultaneous fluorescence and phosphorescence lifetime imaging (FLIM and PLIM) data (Becker *et al.*, 2011b; Becker, 2015a).

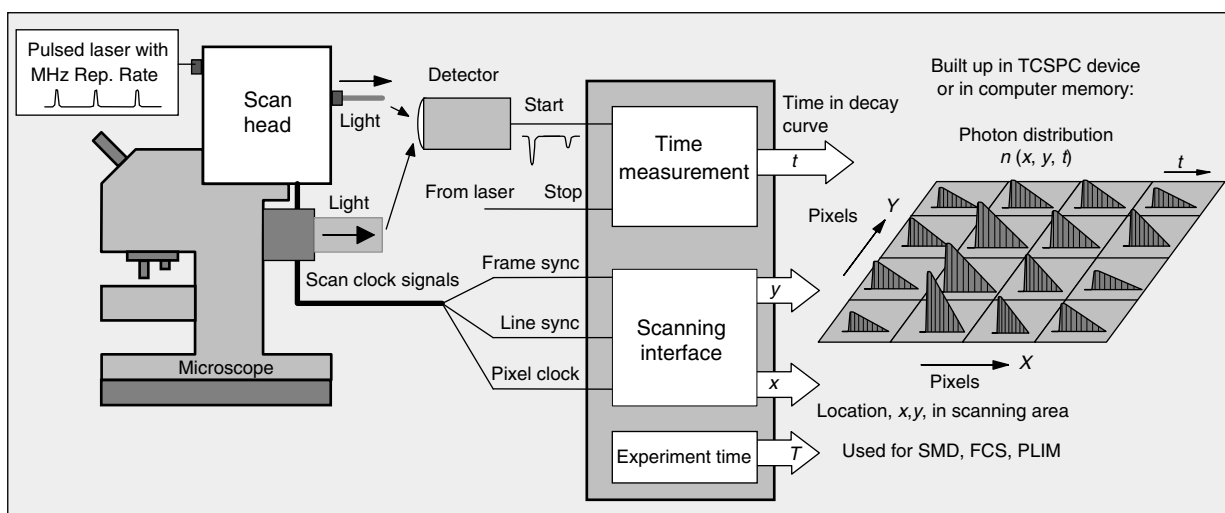


Figure 8 Principle of TCSPC FLIM.

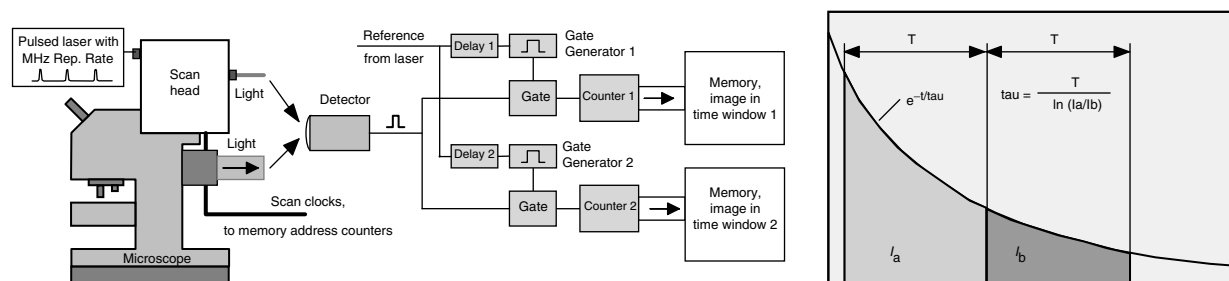


Figure 9 Gated photon counting. Two (or more) gated counters record photons in different time windows.

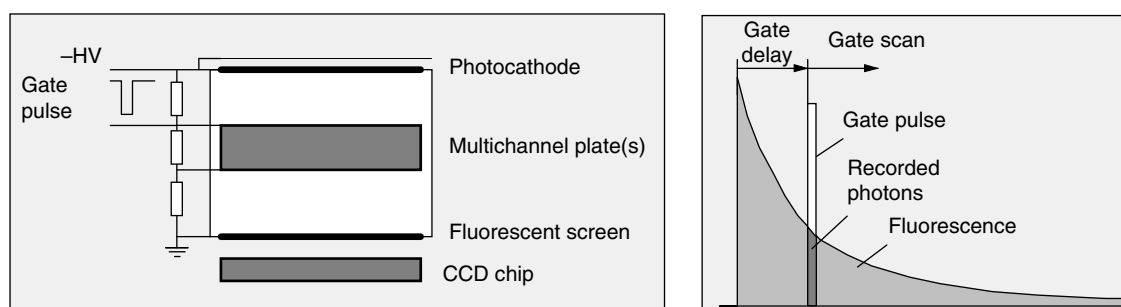


Figure 10 Gated image intensifier.

Moreover, TCSPC FLIM is perfectly compatible with confocal and multiphoton (Denk *et al.*, 1990; Diaspro, 2001; Pawley, 2006; Wilson and Sheppard, 1984) laser scanning microscopes. TCSPC FLIM thus takes advantage of the optical sectioning capability of confocal or multiphoton scanning: the data are obtained from an accurately defined plane of the sample, without contamination from out-of focus fluorescence or lateral scattering (Becker, 2015b).

In addition to building up FLIM data TCSPC makes the full information about the individual photons available. Such ‘parameter-tagged photon data’ can be used in various ways, for example, for single-molecule detection (SMD) (Prummer *et al.*, 2004; Widengren *et al.*, 2006), fluorescence correlation spectroscopy (FCS), (Becker *et al.*, 2006; Felekyan *et al.*, 2005), and PLIM (Becker *et al.*, 2011b). Please see Becker (2015a) for technical details.

As a disadvantage of TCSPC FLIM it is often stated that TCSPC FLIM is slow in terms of acquisition time. This is not quite correct, as has been shown by Becker *et al.* (2009) and Katsoulidou *et al.* (2007). Certainly, a TCSPC FLIM system cannot record an image faster than the scanner is able to scan the sample. Also, the count rate a TCSPC system can process is limited (Becker, 2005). However, the count rates achieved in typical TCSPC FLIM applications are rather limited by the photostability of the samples than by the counting capability of the FLIM system. Due to its near-ideal photon efficiency, TCSPC FLIM in these cases achieves the shortest acquisition time of all FLIM techniques based on sample scanning.

Gated photon counting

Fluorescence lifetime images can be obtained by scanning the sample with a pulsed laser beam, detecting single photons of

the fluorescence light, and recording the photons in several time-gated counters, see Figure 9, left (Buurman *et al.*, 1992). The fluorescence lifetime is calculated from the counting results in the different time windows, see Figure 9, right. A single-exponential lifetime can be calculated from the intensities, I_a and I_b , in two time windows. Up to four time windows are used to determine lifetime components of multi-exponential decay profiles.

FLIM by gated photon counting works up to very high count rates. Moreover, as long as the fluorescence lifetimes are in a reasonable relation to the width of the time windows, gated photon counting works at high photon efficiency (Gerritsen *et al.*, 2002). The problem of gated photon counting is that the decay profiles are under-sampled. The result therefore depends on the location of the excitation pulse in or before the first time window, and on the transit time spread in the detector. Both parameters cannot be determined directly. Gated photon counting therefore has to rely on calibration with fluorophores of known lifetimes.

Time-domain wide-field FLIM by gated image intensifiers

Image intensifiers are vacuum tubes containing a photocathode, a multiplication system for the photoelectrons, and a two-dimensional image detection system. Multichannel plates are used for electron multiplication, see Figure 10. One microchannel plate gives a typical multiplication factor of 10^3 , a gain of 10^6 can be achieved by two plates in series. With a gain this high, the device is able to detect single photons of the fluorescence signal.

Time resolution is obtained by gating the image intensifier. This is achieved by applying a gating pulse between the multichannel plate and the photocathode or a conductive grid

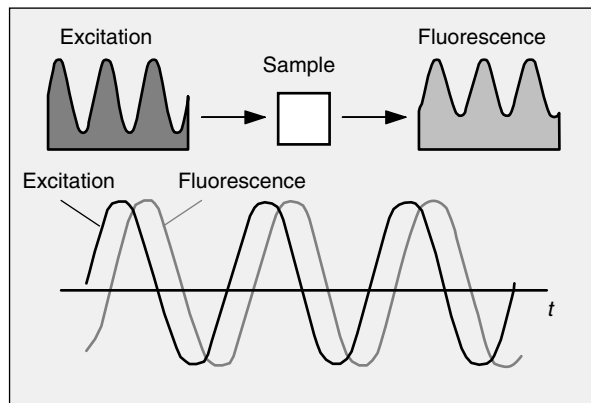


Figure 11 Principle of frequency-domain FLIM. The fluorescence lifetime is derived from the decrease in the modulation degree and the phase shift of the fluorescence compared to the excitation.

capacitively coupled to it. The fluorescence decay functions are measured by scanning the gate pulse in time, i.e., by taking consecutive images for different delay of the gate pulse, see [Figure 10](#), right.

Image intensifiers are normally used in combination with wide-field excitation ([Dowling et al., 1997](#)) or multiphoton multi-beam scanning ([Straub and Hell, 1998](#)). The advantage of wide-field imaging is that image information is acquired simultaneously for all pixels. Therefore the acquisition times can be made shorter than for techniques based on spatial scanning.

The disadvantage is that gating records only a fraction of the photons. The photon efficiency of the recording is therefore low, and the total sample exposure is high. The efficiency problem can be partially solved by recording only a few time gates. However, this results in under-sampled fluorescence decay data and problems to resolve multi-exponential decay profiles.

Moreover, wide-field excitation does not reject out-of-focus light. Unless the samples are very thin the decay data are thus contaminated by fluorescence from out-of-focus planes. In principle, out-of-focus signals can be removed from the data by structured illumination techniques ([Cole et al., 2001](#)). However, these techniques are based on calculating differences between images recorded for different illumination patterns. Although the result is an image containing only the signal from the focal plane this image contains the shot noise of the photons of all planes from which fluorescence is detected.

Frequency-domain FLIM by laser scanning

Frequency-domain FLIM records differences in the phase and the modulation degree between a modulated or pulsed excitation and the fluorescence signal, see [Figure 11](#). These values translate directly into the fluorescence lifetime.

A commonly used principle of frequency-domain FLIM with laser scanning and single-point detectors is shown in [Figure 12](#). The gain of detectors is modulated by an oscillator frequency, f_{osc} , slightly different from the laser pulse or modulation frequency, f_{laser} . The phase shift and the amplitude of the fluorescence signal then transfer into a signal at the difference frequency, $f_{laser} - f_{osc}$. This is the same ‘heterodyne’ principle that is used in a radio. The advantage is that the difference frequency can be made independent of and much

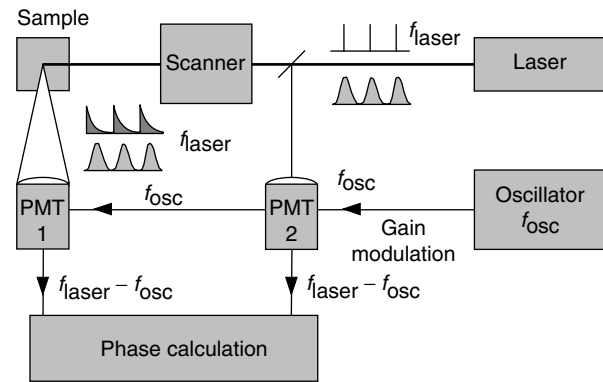


Figure 12 Frequency-domain FLIM by laser scanning and modulated point-detectors.

lower than the laser frequency. The phase and the amplitude are determined at the difference frequency where filtering and signal processing can be performed by digital techniques ([Gratton and Barbieri, 1986](#)).

Both the modulation of the excitation light and the modulation of the detectors need not necessarily be sinusoidal. Frequency-domain systems can therefore also be used with high-frequency pulsed lasers ([Gratton and Barbieri, 1986](#)).

Frequency-domain FLIM treats the fluorescence signal as an analog waveform, not as individual photon detection events. It is therefore able to work at extremely high intensities, where single-photon detection becomes impossible. Problems may occur at extremely low intensities, when the detector, on average, delivers less than one photon within the time interval of the phase calculation. What happens in this situation depends on the details of the electronics used.

The photon efficiency of frequency-domain FLIM depends on a number of instrumental details. The efficiency increases with the modulation degree in the excitation light source and in the detectors. The best efficiency is obtained by using short laser pulses and modulating the detectors by square-wave signals. Details can be found in [Philip and Carlsson \(2003\)](#).

Frequency-domain FLIM by modulated image intensifiers

Detection of frequency-domain FLIM data is possible by wide-field excitation and a gain-modulated camera ([Lakowicz and Berndt, 1991](#)). The electronic principle is shown in [Figure 13](#). An image intensifier is gain-modulated at a frequency slightly different from the modulation frequency of the laser. A series of images for different times in the period of the difference frequency is taken, from which the phase and the modulation degree are calculated. Alternatively, the phase of the camera modulation can be changed incrementally, and several images for different phase recorded.

FLIM Data Analysis

Time-domain data

Time-domain FLIM data (from TCSPC FLIM systems or gated camera systems) are arrays of pixels each of which contain photon numbers in a (usually large) number of time channels,

see Figure 14, middle. To obtain fluorescence decay parameters from such data an iterative convolution procedure is used. The function of a suitable decay model is convoluted with the instrument response function (IRF). The IRF is the response of the detection system to the excitation pulse itself. It is either measured or calculated from the fluorescence decay data. The fit procedure then optimizes the model parameters until the best fit to the photon numbers in the time channels is achieved.

Usually the decay profiles are modeled by sums of two or three exponential functions. These models are described by several decay times and amplitude coefficients. The final FLIM image is obtained by assigning the brightness to the total photon number in the pixel, and the color to a selected decay parameter (Figure 14, right). This can be the lifetime of a single-exponential decay, an amplitude- or intensity-weighted average of the lifetimes in a multi-exponential decay, or a lifetime or amplitude ratio.

Frequency-domain data

Frequency-Domain data contain two numbers in each pixel: a phase and a modulation degree. Such data can be analyzed elegantly by the 'Phasor' approach (Digman *et al.*, 2008). Phasor analysis does not explicitly aim on determining fluorescence lifetimes or decay components for the pixels. Instead, it uses the phase and the modulation degree directly. For each pixel, a pointer (the phasor) is defined and displayed in a polar plot. The phase is used as the angle of the pointer, the modulation degree as the amplitude. The phasors of pixels with single-exponential decay profiles end on a semicircle.

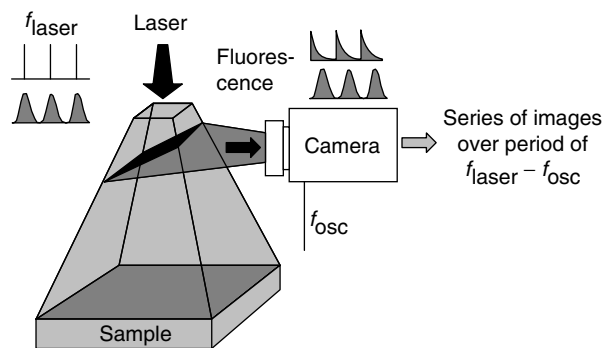


Figure 13 Frequency-domain FLIM by wide-field excitation and modulated image intensifier.

Phasors of pixels with multi-exponential decay profiles, i.e., sums of several decay components, end inside the semicircle.

Pixels with different signature in the phasor plot are back-annotated in the image by giving them different colors. An example is shown in Figure 15. It shows a phasor plot for several cells, B–F, shown at the top of the figure. The cells were transfected with different fluorescent proteins. (F) is an untransfected cell. As can be seen in the lower part of the figure, the fluorophores form different spots in the phasor plot, as does the autofluorescence of the untransfected cell.

Phasor analysis is not restricted to data obtained in frequency-domain systems. It can be applied to Fourier-transformed time-domain data as well (Digman *et al.*, 2008).

Latest Developments

In the last ten years, FLIM techniques have made impressive progress. The most obvious one is in the spatial

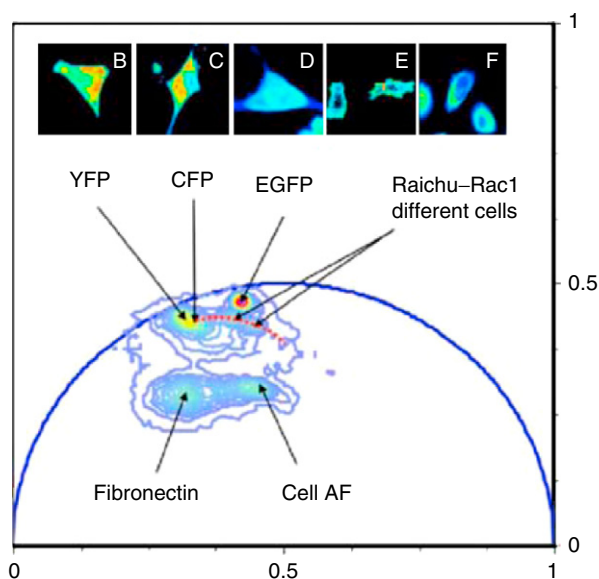


Figure 15 Phasor analysis. Five different cells, transfected with (B) YFP, (C) paxillin-CFP, (D) EGFP, and (E) RAICRU-Rac1. (F) is an untransfected cell emitting only autofluorescence. Reproduced from Digman, M.A., Caiolfa, V.R., Zama, M., Gratton, E., 2008. The phasor approach to fluorescence lifetime imaging analysis. *Biophysical Journal* 94, L14–L16.

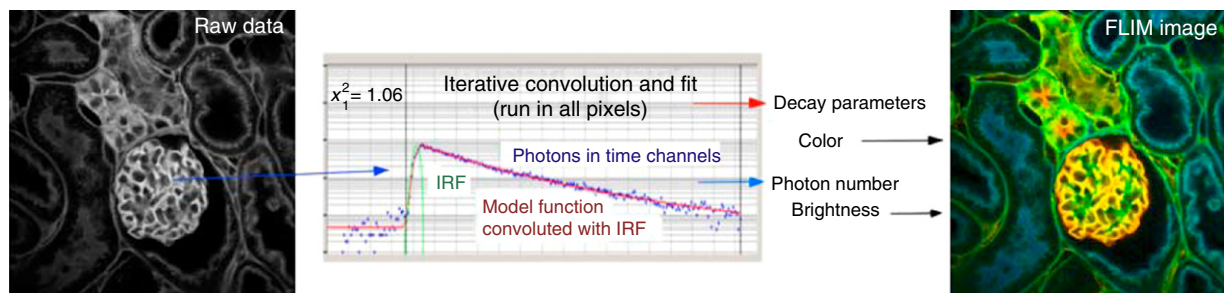


Figure 14 Time-domain FLIM data analysis.

resolution: early FLIM images were recorded with no more than 128×128 pixels. The reasons were low efficiency of the optical systems and detectors, and limitations in available memory size. Within the past ten years, the efficiency of the detection path of laser scanning microscopes has improved significantly. In TCSPC FLIM, conventional PMTs have been replaced with single-photon avalanche photodiodes and hybrid detectors. These detectors have considerably higher detection efficiency than conventional PMTs. Hybrid detectors also deliver cleaner signals (Becker *et al.*, 2011a). Thus, they not only detect more photons, but also allow the FLIM system to achieve a higher lifetime accuracy for a given number of photons.

The speed and memory size of computers has increased by more than an order of magnitude. As a result, the raw data (e.g., data for single photons) can be transferred directly into the computer and the FLIM data be built up in the computer memory. With the introduction of 64-bit operating systems memory size is no longer a problem. FLIM data can therefore be recorded at mega-pixel resolution. For example, TCSPC FLIM data have been recorded with a spatial resolution of 2048×2048 pixels, and with 256 time channels in each pixel. The total data size of one of these images is more than 2 GB (Studier and Becker, 2014).

Increased efficiency helps avoid artifacts induced by photobleaching, photodamage, or by photo-induced metabolic changes in the samples. As a result, FLIM can now be combined with multidimensional microscopy techniques, such as Z-stack imaging, time-series recording, mosaic imaging (Studier and Becker, 2014; Becker, 2015a,b), or excitation-wavelength scanning. Lasers of different wavelength can also be multiplexed at high rate, and lifetime images for several combinations of excitation and emission wavelength be recorded quasi-simultaneously (Becker, 2015a). The wavelength range of FLIM has recently been extended into the near infrared. The titanium-sapphire laser of a multiphoton microscope is used as a one-photon excitation source, and the fluorescence is detected by an NIR-sensitive hybrid detector (Becker and Shcheslavskiy, 2013).

FLIM, especially TCSPC FLIM, has been combined with STED (stimulated emission-depletion) microscopy (Auksorius *et al.*, 2008; Bückers *et al.*, 2011; Lesoine *et al.*, 2012) and with NSOM (near-field scanning optical microscopy) (Cadby *et al.*, 2005; Gaiduk *et al.*, 2007; Micic *et al.*, 2004) to obtain fluorescence lifetime images with optical super-resolution.

Instruments based on a combination of confocal or two-photon laser-scanning microscopes with TCSPC are, in principle, able to record also fluorescence correlation (FCS). In the early times of FLIM FCS was often prevented by low detection efficiency or the inability of the scanner to stop the beam in a defined spot. These problems have been solved by now so that almost any TCSPC FLIM system is able to record FCS. FCS measurement can be combined with fluorescence decay data to improve the separation between different fluorophores (Böhmer *et al.*, 2002; Anthony and Berland, 2013). FCS data can also be derived from scan data by RICS (raster image correlation spectroscopy) (Digman *et al.*, 2005).

TCSPC FLIM has been extended to simultaneously record fluorescence (FLIM) and PLIM (Becker *et al.*, 2011b; Becker, 2015a). This opens a way to determine oxygen concentrations

in cells and tissue and simultaneously obtain metabolic information via the fluorescence decay of NADH.

Progress has also been made in recording dynamic effects in the fluorescence lifetime of biological systems. With faster data recording and processing rates available, FLIM data can be acquired at a speed limited only by the maximum exposure the sample tolerates. Even TCSPC FLIM, which is usually considered slow in acquisition speed, delivers image rates of one image per second (Katsoulidou *et al.*, 2007). A technique to record extremely fast lifetime changes by TCSPC is based on the buildup of a photon distribution over the distance along a line scan, the arrival times of the photons after the laser pulse, and the arrival times after a periodic stimulation of the sample. The technique is called Fluorescence Lifetime-Transient Scanning (FLITS) and has been shown to record photochemical and non-photochemical chlorophyll transients in plants and transient changes in the Ca^{2+} concentration in cultured neurons (Becker *et al.*, 2014; Becker, 2015b).

See also: Imaging the Cell: Light Microscopy: Fluorescence Correlation Spectroscopy: A Tool for Measuring Dynamic and Equilibrium Properties of Molecules in Cells

References

- Anthony, N., Berland, K., 2013. Global analysis in fluorescence correlation spectroscopy and fluorescence lifetime microscopy. In: Tetin, S.Y. (Ed.), *Methods in Enzymology* 518, Fluorescence Fluctuation Spectroscopy (FFS), Part A. San Diego: Academic Press, pp. 145–173.
- Auksorius, E., Boruah, B.R., Dunsby, C., *et al.*, 2008. Stimulated emission depletion imaging microscopy with a super continuum source and fluorescence lifetime imaging. *Optics Letters* 33, 113–115.
- Bacskai, B.J., Skoch, J., Hickey, G.A., Allen, R., Hyman, B.T., 2003. Fluorescence resonance energy transfer determinations using multiphoton fluorescence lifetime imaging microscopy to characterize amyloid-beta plaques. *Journal of Biomedical Optics* 8, 368–375.
- Bastiaens, P.I.H., Squire, A., 1999. Fluorescence lifetime imaging microscopy: Spatial resolution of biochemical processes in the cell. *Trends in Cell Biology* 9, 48–52.
- Becker, W., 2005. *Advanced Time-Correlated Single-Photon Counting Techniques*. Berlin; Heidelberg; New York, NY: Springer.
- Becker, W., 2012. Fluorescence lifetime imaging – Techniques and applications. *Journal of Microscopy* 247, 119–136.
- Becker, W., 2015a. The bh TCSPC Handbook, sixth ed. Berlin: Becker & Hickl GmbH. Available at: www.becker-hickl.com (accessed 04.04.15).
- Becker, W. (Ed.), 2015b. *Advanced Time-Correlated Single Photon Counting Applications*. Berlin; Heidelberg; New York, NY: Springer.
- Becker, W., Bergmann, A., 2008. Lifetime-resolved imaging in nonlinear microscopy. In: Masters, B.R., So, P.T.C. (Eds.), *Handbook of Biomedical Nonlinear Optical Microscopy*. Oxford; New York, NY: Oxford University Press.
- Becker, W., Bergmann, A., Biskup, C., 2007. Multi-spectral fluorescence lifetime imaging by TCSPC. *Microscopy Research and Technique* 70, 403–409.
- Becker, W., Bergmann, A., Hausteiner, E., *et al.*, 2006. Fluorescence lifetime images and correlation spectra obtained by multi-dimensional TCSPC. *Microscopy Research and Technique* 69, 186–195.
- Becker, W., Bergmann, A., Hink, M.A., *et al.*, 2004. Fluorescence lifetime imaging by time-correlated single photon counting. *Microscopy Research and Technique* 63, 58–66.
- Becker, W., Shcheslavskiy, V., 2013. Fluorescence lifetime imaging with near-infrared dyes. *Proceedings of SPIE* 8588, 85880R.
- Becker, W., Shcheslavskiy, V., Frere, S., Slutsky, I., 2014. Spatially resolved recording of transient fluorescence lifetime effects by line-scanning TCSPC. *Microscopy Research and Technique* 77, 216–224.

- Becker, W., Su, B., Bergmann, A., 2009. Fast-acquisition multispectral FLIM by parallel TCSPC. *Proceedings of SPIE* 7183, 7183051–5–7183051-9.
- Becker, W., Su, B., Bergmann, A., Weisshart, K., Holub, O., 2011b. Simultaneous fluorescence and phosphorescence lifetime imaging. *Proceedings of SPIE* 7903, 790320-1–790320-7.
- Becker, W., Su, B., Weisshart, K., Holub, O., 2011a. FLIM and FCS detection in laser-scanning microscopes: Increased efficiency by GaAsP hybrid detectors. *Microscopy Research and Technique* 74, 804–811.
- Berezin, M.Y., Achilefu, S., 2010. Fluorescence lifetime measurement and biological imaging. *Chemical Reviews* 110, 2641–2684.
- Bird, D.K., Yan, L., Vrotsos, K.M., Eliceiri, K.E., Vaughan, E.M., 2005. Metabolic mapping of MCF10A human breast cells via multiphoton fluorescence lifetime imaging of coenzyme NADH. *Cancer Research* 65, 8766–8773.
- Biskup, C., Kelbauskas, L., Zimmer, T., *et al.*, 2004. Interaction of PSD-95 with potassium channels visualized by fluorescence lifetime-based resonance energy transfer imaging. *Journal of Biomedical Optics* 9, 735–759.
- Böhmer, M., Wahl, M., Rahn, H.-J., Erdmann, R., Enderlein, J., 2002. Time-resolved fluorescence correlation spectroscopy. *Chemical Physics Letters* 353, 439–445.
- Bückers, J., Wildanger, D., Vicidomini, G., Kastrop, L., Hell, S.W., 2011. Simultaneous multi-lifetime multi-color STED imaging for colocalization analysis. *Optics Express* 19, 3130–3143.
- Bugiel, I., König, K., Wabnitz, H., 1989. Investigations of cells by fluorescence laser scanning microscopy with subnanosecond time resolution. *Lasers in the Life Sciences* 3 (1), 47–53.
- Buurman, E.P., Sanders, R., Draaijer, A., *et al.*, 1992. Fluorescence lifetime imaging using a confocal laser scanning microscope. *Scanning* 14, 155–159.
- Cadby, A., Dean, R., Fox, A.M., Jones, R.A.L., Lidzey, D.G., 2005. Mapping the fluorescence decay lifetime of a conjugated polymer in a phase-separated blend using a scanning near-field optical microscope. *Nano Letters* 5, 2232–2237.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., Orashner, D.C., 1994. Green fluorescent protein as a marker for gene expression. *Science* 263, 802–805.
- Chen, Y., Periasamy, A., 2004. Characterization of two-photon excitation fluorescence lifetime imaging microscopy for protein localization. *Microscopy Research and Technique* 63, 72–80.
- Chorvat, D., Chorvatova, A., 2006. Spectrally resolved time-correlated single photon counting: A novel approach for characterization of endogenous fluorescence in isolated cardiac myocytes. *European Biophysics Journal* 36, 73–83.
- Chorvat, D., Chorvatova, A., 2009. Multi-wavelength fluorescence lifetime spectroscopy: A new approach to the study of endogenous fluorescence in living cells and tissues. *Laser Physics Letters* 6, 175–193.
- Chorvatova, A., Elzwiei, F., Mateasik, A., Chorvat, D., 2012. Effect of ouabain on metabolic oxidative state in living cardiomyocytes evaluated by time-resolved spectroscopy of endogenous NAD(P)H fluorescence. *Journal of Biomedical Optics* 17 (10), 101505-1–101505-7.
- Cole, M.J., Siegel, J., Webb, S.E.D., *et al.*, 2001. Time-domain whole-field lifetime imaging with optical sectioning. *Journal of Microscopy* 203, 246–257.
- Denk, W., Strickler, J.H., Webb, W.W., 1990. Two-photon laser scanning fluorescence microscopy. *Science* 248, 73–76.
- Diaspro, A. (Ed.), 2001. *Confocal and Two-Photon Microscopy: Foundations, Applications and Advances*. New York, NY: Wiley-Liss, Inc.
- Digman, M.A., Caiola, V.R., Zamai, M., Gratton, E., 2008. The phasor approach to fluorescence lifetime imaging analysis. *Biophysical Journal* 94, L14–L16.
- Digman, M.A., Sengupta, P., Wiseman, P.W., *et al.*, 2005. Fluctuation correlation spectroscopy with a laser-scanning microscope: Exploiting the hidden time structure. *Biophysical Journal* 88, L33–L36.
- Dimitrow, E., Riemann, I., Ehlers, A., *et al.*, 2009. Spectral fluorescence lifetime detection and selective melanin imaging by multiphoton laser tomography for melanoma diagnosis. *Experimental Dermatology* 18, 509–515.
- Dowling, K., Hyde, S.C.W., Dainty, J.C., French, P.M.W., Hares, J.D., 1997. 2-D fluorescence lifetime imaging using a time-gated image intensifier. *Optics Communications* 135, 27–31.
- Duncan, R.R., Bergmann, A., Cousin, M.A., Apps, D.K., Shipston, M.J., 2004. Multi-dimensional time-correlated single-photon counting (TCSPC) fluorescence lifetime imaging microscopy (FLIM) to detect FRET in cells. *Journal of Microscopy* 215, 1–12.
- Felekyan, S., Kühnemuth, R., Kudryavtsev, V., *et al.*, 2005. Full correlation from picoseconds to seconds by time-resolved and time-correlated single photon detection. *Review of Scientific Instruments* 76, 083104-1–083104-14.
- Förster, T.H., 1948. Zwischenmolekulare Energiewanderung und Fluoreszenz. *Annals of Physics* 2 (Serie 6), 55–75.
- Förster, T.H., 2012. Energy migration and fluorescence. Translated by Klaus Suhling. *Journal of Biomedical Optics* 17, 011002-1–011002-10.
- Funk, K., Woitecki, A., Franjic-Würtz, C., *et al.*, 2008. Modulation of chloride homeostasis by inflammatory mediators in dorsal ganglion neurons. *Molecular Pain* 4, 32–43.
- Gaiduk, A., Kühnemuth, R., Felekyan, S., *et al.*, 2007. Fluorescence detection with high time resolution: From optical microscopy to simultaneous force and fluorescence spectroscopy. *Microscopy Research and Technique* 70, 403–409.
- Geddes, C.D., Cao, H., Gryczynski, I., Fang, J., Lakowicz, J.R., 2003. Metal-enhanced fluorescence (MEF) due to silver colloids on a planar surface: Potential applications of indocyanine green to *in vivo* imaging. *Journal of Physical Chemistry A* 107, 3443–3449.
- Gerritsen, H.C., Asselbergs, M.A.H., Agronskaia, A.V., van Sark, W.G.J.H.M., 2002. Fluorescence lifetime imaging in scanning microscopes: Acquisition speed, photon economy and lifetime resolution. *Journal of Microscopy* 206, 218–224.
- Geusens, B., Van Gele, M., Braat, S., *et al.*, 2010. Flexible nanosomes (SECOsomes) enable efficient siRNA delivery in cultured primary skin cells and in the viable epidermis of ex vivo human skin. *Advanced Functional Materials* 20, 4077–4090.
- Gilbert, D., Franjic-Würtz, C., Funk, K., *et al.*, 2007. Differential maturation of chloride homeostasis in primary afferent neurons of the somatosensory system. *International Journal of Developmental Neuroscience* 25, 479–489.
- Göppert-Mayer, M., 1931. Über Elementarakte mit zwei Quantensprüngen. *Annals of Physics* 9, 273–294.
- Gratton, E., Barbieri, B.B., 1986. Multifrequency phase fluorometry using pulsed sources: Theory and applications. *Spectroscopy* 1, 28–36.
- Gukassyan, V., Hsu, Y.-Y., Kung, S.-H., Kao, F.-J., 2007. Application of fluorescence resonance energy transfer resolved by fluorescence lifetime imaging microscopy for the detection of enterovirus 71 infection in cells. *Journal of Biomedical Optics* 12, 024016-1–024016-8.
- Gukassyan, V., Kao, F.-J., 2009. Monitoring cellular metabolism with fluorescence lifetime of reduced nicotinamide adenine dinucleotide. *Journal of Physical Chemistry C* 113, 11532–11540.
- Guo, H.-W., Chen, C.-T., Wei, Y.-H., *et al.*, 2008. Reduced nicotinamide adenine dinucleotide fluorescence lifetime separates human mesenchymal stem cells from differentiated progenies. *Journal of Biomedical Optics* 13 (5), 050505-1–050505-3.
- Hanson, K.M., Behne, M.J., Barry, N.P., Mauro, T.M., Gratton, E., 2002. Two-photon fluorescence imaging of the skin stratum corneum pH gradient. *Biophysics Journal* 83, 1682–1690.
- Hoffmann, B., Zimmer, T.H., Klöcker, N., *et al.*, 2008. Prolonged irradiation of enhanced cyan fluorescent protein or Cerulean can invalidate Förster resonance energy transfer measurements. *Journal of Biomedical Optics* 13 (3), 031250-1–031250-9.
- Kaneko, H., Putzier, I., Frings, S., Kaupp, U.B., Gensch, T.H., 2004. Chloride accumulation in mammalian olfactory sensory neurons. *Journal of Neuroscience* 24 (36), 7931–7938.
- Kantelhardt, S.R., Leppert, J., Krajewski, J., *et al.*, 2007. Imaging of brain and brain tumor specimens by time-resolved multiphoton excitation microscopy ex vivo. *Neuro-Oncology* 9, 103–112.
- Katsoulidou, V., Bergmann, A., Becker, W., 2007. How fast can TCSPC FLIM be made? *Proceedings of SPIE* 6771, 67710B-1–67710B-7.
- Kelbauskas, L., Dietel, W., 2002. Internalization of aggregated photosensitizers by tumor cells: Subcellular time-resolved fluorescence spectroscopy on derivatives of pyropheophorbide-a ethers and chlorin e6 under femtosecond one- and two-photon excitation. *Photochemistry and Photobiology* 76, 686–694.
- Koehler, M.J., König, K., Elsner, P., Bückle, R., Kaatz, M., 2006. *In vivo* assessment of human skin aging by multiphoton laser scanning tomography. *Optics Letters* 31, 2879–2881.
- Köllner, M., Wolfrum, J., 1992. How many photons are necessary for fluorescence-lifetime measurements? *Journal of Physical Chemistry Letters* 200, 199–204.
- König, K., 2008a. Multiphoton-induced cell damage. In: Masters, B.R., So, P.T.C. (Eds.), *Handbook of Biomedical Nonlinear Optical Microscopy*. Oxford; New York, NY: Oxford University Press.
- König, K., 2008b. Clinical multiphoton tomography. *Journal of Biophotonics* 1, 13–23.
- König, K., Riemann, I., 2003. High-resolution multiphoton tomography of human skin with subcellular spatial resolution and picosecond time resolution. *Journal of Biomedical Optics* 8, 432–439.
- König, K., Uchugonova, A., 2009. Multiphoton fluorescence lifetime imaging at the dawn of clinical application. In: Periasamy, A., Clegg, R.M. (Eds.), *FLIM Microscopy in Biology and Medicine*. Boca Raton, FL; London; New York, NY: CRC Press.
- König, K., Uchugonova, A., Gorjup, E., 2011. Multiphoton fluorescence lifetime imaging of 3D-stem cell spheroids during differentiation. *Microscopy Research and Technique* 74, 9–17.
- Koushik, S.V., Vogel, S.S., 2008. Energy migration alters the fluorescence lifetime of Cerulean: Implications for fluorescence lifetime imaging Förster resonance energy transfer measurements. *Journal of Biomedical Optics* 13 (3), 031204-1–031204-9.

- Kuchibhotla, K.V., Lattarulo, C.R., Hyman, B., Bacskaï, B.J., 2009. Synchronous hyperactivity and intercellular calcium waves in astrocytes in Alzheimer mice. *Science* 323, 1211–1215.
- Kuimova, M.K., Yahioglu, G., Levitt, J.A., Suhling, K., 2008. Molecular rotor measures viscosity of live cells via fluorescence lifetime imaging. *Journal of the American Chemical Society* 130, 6672–6673.
- Lakowicz, J.R., 2006. *Principles of Fluorescence Spectroscopy*, third ed. New York, NY; Boston, MA; Dordrecht, London; Moscow: Springer.
- Lakowicz, J.R., Berndt, K., 1991. Lifetime-selective fluorescence lifetime imaging using an rf phase-sensitive camera. *Review of Scientific Instruments* 62, 1727–1734.
- Lakowicz, J.R., Szmanski, H., Nowaczyk, K., Johnson, M.L., 1992. Fluorescence lifetime imaging of free and protein-bound NADH. *Proceedings of the National Academy of Sciences of the United States of America* 89, 1271–1275.
- Leppert, J., Krajewski, J., Kantelhardt, S.R., *et al.*, 2006. Multiphoton excitation of autofluorescence for microscopy of glioma tissue. *Neurosurgery* 58, 759–767.
- Lesoin, M.D., Bose, S., Petrich, J.W., Smith, E.A., 2012. Supercontinuum stimulated emission depletion fluorescence lifetime imaging. *Journal of Physical Chemistry B* 116, 7821–7826.
- Lin, L.L., Grice, J.E., Butler, M.K., *et al.*, 2011. Time-correlated single photon counting for simultaneous monitoring of zinc oxide nanoparticles and NAD(P)H in intact and barrier-disrupted volunteer skin. *Pharmaceutical Research*. doi:10.1007/s11095-011-0515-5.
- Masters, B.R., So, P.T.C., 1999. Multi-photon excitation microscopy and confocal microscopy imaging of *in vivo* human skin: A comparison. *Microscopy and Microanalysis* 5, 282–289.
- Masters, B.R., So, P.T.C., 2008. *Handbook of Biomedical Nonlinear Optical Microscopy*. Oxford; New York, NY: Oxford University Press.
- Masters, B.R., So, P.T.C., Gratton, E., 1997. Multiphoton excitation fluorescence microscopy and spectroscopy of *in vivo* human skin. *Biophysics Journal* 72, 2405–2412.
- Micic, M., Hu, D., Suh, Y.D., *et al.*, 2004. Correlated atomic force microscopy and fluorescence lifetime imaging of live bacterial cells. *Colloids and Surfaces B, Biointerfaces* 34, 205–212.
- Muddana, H.S., Morgan, T.T., Adair, J.H., Butler, P.J., 2009. Photophysics of Cy3-encapsulated calcium phosphate nanoparticles. *Nano Letters* 9 (4), 1559–1566.
- O'Connor, D.V., Phillips, D., 1984. *Time-Correlated Single Photon Counting*. London: Academic Press.
- Patalay, R., Talbot, C., Alexandrov, Y., *et al.*, 2011. Quantification of cellular autofluorescence of human skin using multiphoton tomography and fluorescence lifetime imaging in two spectral detection channels. *Optics Express* 2, 3295–3308.
- Paul, R.J., Schneckenburger, H., 1996. Oxygen concentration and the oxidation-reduction state of yeast: Determination of free/bound NADH and flavins by time-resolved spectroscopy. *Naturwissenschaften* 83, 32–35.
- Pawley, J., 2006. *Handbook of Biological Confocal Microscopy*, third ed. New York, NY: Springer Science + Business Media.
- Pelet, S., Previte, M.J.R., So, P.T.C., 2006. Comparing the quantification of Förster resonance energy transfer measurement accuracies based on intensity, spectral, and lifetime imaging. *Journal of Biomedical Optics* 11, 034017-1–034017-11. (1006).
- Periasamy, A., 2001. *Methods in Cellular Imaging*. Oxford, NY: Oxford University Press.
- Periasamy, A., Clegg, R.M., 2009. *FLIM Microscopy in Biology and Medicine*. Boca Raton, FL; London; New York, NY: CRC Press.
- Philip, J.P., Carlsson, K., 2003. Theoretical investigation of the signal-to-noise ratio in fluorescence lifetime imaging. *Journal of Optical Society of America B* A20, 368–379.
- Prow, T.W., Grice, J.E., Lin, L.L., *et al.*, 2011. Nanoparticles and microparticles for skin drug delivery. *Advanced Drug Delivery Reviews* 63, 470–491.
- Prummer, M., Sick, B., Renn, A., Wild, U.P., 2004. Multiparameter microscopy and spectroscopy for single-molecule analysis. *Analytical Chemistry* 76, 1633–1640.
- Richards-Kortum, R., Dreze, R., Sokolov, K., Pavlova, I., Follen, M., 2003. Survey of endogenous biological fluorophores. In: Mycek, M.-A., Pogue, B.W. (Eds.), *Handbook of Biomedical Fluorescence*. New York, Basel: Marcel Dekker Inc., pp. 237–264.
- Ritman-Meer, T., Cade, N.I., Richards, D., 2007. Spatial imaging of modifications to fluorescence lifetime and intensity by individual Ag nanoparticles. *Applied Physics Letters* 91, 123122-1–123122-3.
- Roberts, M.S., Dancik, Y., Prow, T.W., *et al.*, 2011. Non-invasive imaging of skin physiology and percutaneous penetration using fluorescence spectral and lifetime imaging with multiphoton and confocal microscopy. *European Journal of Pharmaceutics and Biopharmaceutics* 77, 469–488.
- Roberts, M.S., Roberts, M.J., Robertson, T.A., *et al.*, 2008. *In vitro* and *in vivo* imaging of xenobiotic transport in human skin and in the rat liver. *Journal of Biophotonics* 1, 478–493.
- Rück, A., Hülshoff, C.H., Kinzler, I., Becker, W., Steiner, R., 2007. SLIM: A new method for molecular imaging. *Microscopy Research and Technique* 70, 403–409.
- Sanchez, W.Y., Pastore, M., Haridass, I., *et al.*, 2015. Fluorescence lifetime imaging of the skin. In: Becker, W. (Ed.), *Advanced Time-Correlated Single Photon Counting Applications*. Berlin; Heidelberg; New York, NY: Springer.
- Sanchez, W.Y., Prow, T.W., Sanchez, W.H., Grice, J.E., Roberts, M.S., 2010. Analysis of the metabolic deterioration of ex-vivo skin, from ischemic necrosis, through the imaging of intracellular NAD(P)H by multiphoton tomography and fluorescence lifetime imaging microscopy (MPT-FLIM). *Journal of Biomedical Optics* 15, 046008-1–046008-11.
- Sanders, R., Draaijer, A., Gerritsen, H.C., Houpt, P.M., Levine, Y.K., 1995. Quantitative pH imaging in cells using confocal fluorescence lifetime imaging microscopy. *Analytical Biochemistry* 227, 302–308.
- Schweitzer, D., 2009. Quantifying fundus autofluorescence. In: Lois, N., Forrester, J. V. (Eds.), *Fundus Autofluorescence*. Philadelphia, PA; Baltimore, MD; New York, NY; London; Buenos Aires; Hong Kong; Sydney, NSW; London: Wolters Kluwer, Lippincott Williams & Wilkins.
- Schweitzer, D., 2010. Metabolic mapping. In: Holz, F.G., Spaide, R.F. (Eds.), *Medical Retina, Essential in Ophthalmology*. Heidelberg, VIC; Dordrecht; London; New York, NY: Springer.
- Schweitzer, D., Gaillard, E.R., Dillon, J., *et al.*, 2012. Time-resolved autofluorescence imaging of human donor retina tissue from donors with significant extramacular drusen. *Investigative Ophthalmology & Visual Science* 53, 3376–3386.
- Schweitzer, D., Hammer, M., Schweitzer, F., *et al.*, 2004. *In vivo* measurement of time-resolved autofluorescence at the human fundus. *Journal of Biomedical Optics* 9, 1214–1222.
- Schweitzer, D., Quick, S., Schenke, S., *et al.*, 2009. Vergleich von Parametern der zeitaufgelösten Autofluoreszenz bei Gesunden und Patienten mit früher AMD. *Der Ophthalmologe* 8, 714–722.
- Schweitzer, D., Schenke, S., Hammer, M., *et al.*, 2007. Towards metabolic mapping of the human retina. *Microscopy Research and Technique* 70, 403–409.
- Skala, M.C., Fontanella, A., Lan, L., Izatt, J.A., Dewhirst, M.W., 2010a. Longitudinal optical imaging of tumor metabolism and hemodynamics. *Journal of Biomedical Optics* 15 (1), 011112-1–011112-8.
- Skala, M.C., Ramanujam, N., 2010b. Multiphoton redox ratio imaging for metabolic monitoring *in vivo*. *Methods in Molecular Biology* 594, 155–162.
- Skala, M.C., Richtig, K.M., Bird, D.K., *et al.*, 2007a. *In vivo* multiphoton fluorescence lifetime imaging of protein-bound and free nicotinamide adenine dinucleotide in normal and precancerous epithelia. *Journal of Biomedical Optics* 12, 02401-1–02401-10.
- Skala, M.C., Richtig, K.M., Gendron-Fitzpatrick, A., *et al.*, 2007b. *In vivo* multiphoton microscopy of NADH and FAD redox states, fluorescence lifetimes, and cellular morphology in precancerous epithelia. *Proceedings of the National Academy of Sciences of the United States of America* 104, 19494–19499.
- Straub, M., Hell, S.W., 1998. Fluorescence lifetime three-dimensional microscopy with picosecond precision using a multifocal multiphoton microscope. *Applied Physics Letters* 73, 1769–1771.
- Strickler, S.J., Berg, R.A., 1962. Relationship between absorption intensity and fluorescence lifetime of molecules. *Journal of Chemical Physics* 37, 814–822.
- Studier, H., Becker, W., 2014. Megapixel FLIM. *Proceedings of SPIE* 8948, 89481K.
- Szasz, M., Steven, P., Shima, K., *et al.*, 2011. Fluorescence lifetime imaging unravels C. Trachomatis metabolism and its crosstalk with the host cell. *PLoS Pathogens* 7, e1002108-1–e1002108-12.
- Tregido, C., Levitt, J.A., Suhling, K., 2008. Effect of refractive index on the fluorescence lifetime of green fluorescent protein. *Journal of Biomedical Optics* 13 (3), 031218-1–031218-8.
- Tsien, R.Y., 1998. The green fluorescent protein. *Annual Review of Biochemistry* 67, 509–544.
- Van Zandvoort, M.A.M.J., de Grauw, C.J., Gerritsen, H.C., *et al.*, 2002. Discrimination of DNA and RNA in cells by a vital fluorescent probe: Lifetime imaging of SYTO13 in healthy and apoptotic cells. *Cytometry* 47, 226–232.
- Vogel, S.S., Thaler, C., Koushik, S.V., 2006. Fanciful FRET. *Science's STKE: Signal Transduction Knowledge Environment* 2006 (re2), 1–8.
- Wang, H.-W., Ghukassyan, V., Chen, C.T., *et al.*, 2008. Differentiation of apoptosis from necrosis by dynamic changes of reduced nicotinamide adenine dinucleotide fluorescence lifetime in live cells. *Journal of Biomedical Optics* 13, 054011-1–054011-9.
- Widengren, J., Kudryavtsev, V., Antonik, M., *et al.*, 2006. Single-molecule detection and identification of multiple species by multiparameter fluorescence detection. *Analytical Chemistry* 78, 2039–2050.
- Wilson, T., Sheppard, C., 1984. *Theory and Practice of Scanning Optical Microscopy*. London: Academic Press.

High-Speed Localization Microscopy and Single-Particle Tracking

M El Beheiry and M Dahan, Institut Curie, Paris, France

© 2016 Elsevier Inc. All rights reserved.

Introduction

Fluorescence microscopy has long been the standard for cellular imaging. Its great advantage lies in its high specificity, which allows discriminate imaging of organelles and ensembles of specific proteins. The relatively recent ability to observe individual fluorescing molecules in living cells results from the convergence of innovative molecular labeling techniques and highly sensitive and fast imaging instruments. Indeed, through the culmination of these technological advances, researchers are now able to directly observe the dynamics and instantaneous distributions of single molecules of cellular systems. Moreover, these techniques offer the ability to study low-affinity, weak interactions that may otherwise be inaccessible with conventional biochemical techniques. The single-molecule imaging techniques described in this section have provided great insight into a host of cellular events, not limited to molecular target-searching, protein-scaffold interactions, and virus trafficking (Kusumi *et al.*, 2014; Ruthardt *et al.*, 2011; Levi and Gratton, 2007).

Conventionally, fluorescently labeled molecules may be imaged in an ensemble mode, in which captured images are the combination of the light emanating from all labeled molecules instantaneously. To image individual molecules, however, two accommodations are needed. First, the type of fluorescent label and its optical concentration must be selected judiciously. Second, a proper spatiotemporal resolution needs to be specified as determined by image acquisition and optical parameters.

The fundamental physical concept for imaging single molecules is the optical diffraction limit, first described by Abbe (1873):

$$d = \frac{\lambda}{2NA}$$

where d is the effective resolution, λ is the wavelength of emitted light, and NA is the numerical aperture of the imaging objective lens. This equation describes the minimal perceivable spatial resolution in a conventional optical imaging system. In the context of optical fluorescence imaging (for which the NA is typically ~ 1.3 – 1.4) this limit corresponds to roughly ~ 250 nm in practice. Furthermore, camera detectors that capture fluorescence emission events are discretized by pixels whose effective size is made to correspond to roughly half the diffraction limit. This signifies that any focused individually fluorescing molecule small enough to be considered a point source will appear as a discretized 250 nm diameter spot.

Localization microscopy (LM) comprises the experimental and computational techniques that permit the determination of the spatial coordinates of individual fluorescing molecules with high precision. Single-particle tracking (SPT) is a subfield of LM, which involves reconstructing single-molecule trajectories from sequential localization events in live cells.

In this article, the principles of LM will be overviewed, leading to a discussion of high-speed LM and SPT. Emphasis will be placed on the types of information these techniques yield and how they may contribute to the understanding of cellular processes.

The Localization Concept

The shape of light emitted from a single diffraction-limited fluorescing molecule is described by the point spread function (PSF). Consistent with the Abbe diffraction limit, the PSF has an approximately isotropic Gaussian profile in the lateral dimensions. The captured raw image of an emitting single molecule is therefore the intensity profile projected onto a pixelated detector array (e.g., a highly sensitive camera). This profile, in turn, may be numerically fitted by a two-dimensional (2D) Gaussian function, whose localization precision is inversely proportional to the square root of the number of photons captured from the PSF (Thompson *et al.*, 2002; Mortensen *et al.*, 2010). The 'localization' of the PSF in this manner provides a localization precision improvement by roughly an order of magnitude over the diffraction limit. This remark implies that LM may be used as a super-resolution microscopy technique; a sufficient density of localizations allows pointillist image reconstructions at sub-50 nm resolution.

Two major applications have emerged based on this form of microscopy: super-resolution LM and SPT. Early studies reported the use of functionalized quantum dots, beads, and dyes for localization (Pinaud *et al.*, 2010). Due to their long fluorescence lifetimes and narrow emission spectra, these probes remain strong candidates for LM studies, in particular for tracking. The main impediment for these types of exogenous markers is that they are often limited to studies of the cell membrane in which sampling of the surface is sparse.

Many studies in LM employ specialized fluorescent probes that are capable of stochastic activation. Under this configuration, only a small subpopulation of probes emits light in the image capture interval (typically of the order of tens of milliseconds). Probes that may be used in this fashion include the classes of photoactivatable, photoswitchable, and photoconvertible fluorescent proteins and dyes. Stochastic emission density is modulated by an activation illumination, which is used in conjunction with the read-out excitation illumination. These sparse emission events are captured in consecutive image frames, upon which their precise spatial coordinates are determined by fitting their PSF. Accumulation of these coordinates over a few thousand image frames eventually forms a pointillist reconstruction of the probe distribution inside the cell (Figure 1). Many freely available software packages are currently available for localization and visualization of PSFs from raw image acquisitions (Small and Stahlheber, 2014).

LM performed on fixed cells is possible with immuno-fluorescent labeling with the use of stochastically emitting

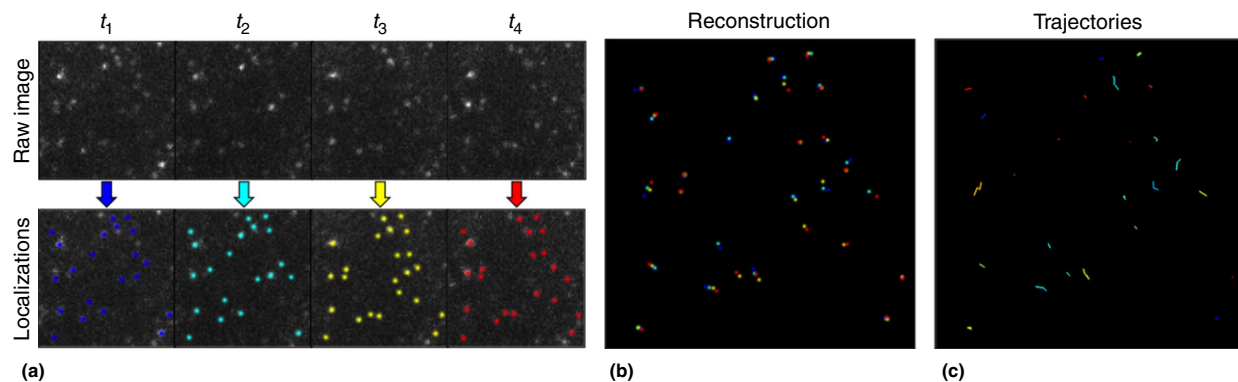


Figure 1 Sequence flow for LM in a live cell (over four image frames). Imaged are transmembrane domain proteins with an extracellular photoactivatable Dendra2 protein. (a) In the top row, multiple images with sparse fluorescent emissions are acquired. In the bottom row, corresponding localizations are shown. The color code corresponds to the images acquired at times t_1 through t_4 . (b) A reconstruction of the localizations is displayed by overlaying all the localizations from (a). The color code is conserved, and the individual spot sizes correspond to the approximate localization precision. (c) Localizations in sequential frames are connected together to form single-molecule trajectories (the color code corresponds to individualized trajectories).

Table 1 Different fluorescent probes used in LM and their relative properties

Probe	Size (nm)	Brightness	Lifetime
Fluorescent proteins	10–20	Low	Short
Fluorescent dyes	0.5–10	Medium–high	Medium
Nanoparticles	20–50	Medium–high	Long
Fluorescent beads	5–1000	High	Long
Quantum dots	10–20	High	Long

fluorescent dyes. Stochastic optical reconstruction microscopy (STORM) belong to a class of techniques that takes advantage of this approach and has been responsible for providing detailed images of intracellular organelles as well as insight to protein distributions, and morphologies of biological complexes (Rust *et al.*, 2006). With bright fluorescent dyes, localization precisions can reach sub-20 nm levels. Variants of the STORM technique include the possibility of imaging in live cells (Jones *et al.*, 2011).

Photoactivated localization microscopy (PALM) involves transgenically expressing photoactivatable fluorescent proteins for LM (Betzig *et al.*, 2006). As fluorescent proteins are not nearly as bright as synthetic dyes, PALM generally suffers from a weaker signal-to-noise ratio for single-molecule detections. However, it is less vulnerable to redundant localizations and is more amenable to live-cell imaging compared to STORM. Table 1 summarizes the properties of the most commonly used types of fluorescent probes used in LM.

There are persistent issues in LM, most of which can be addressed through post-processing steps. The first issue is that of false-positive detections. Inevitably, noise may be mischaracterized as an emitting probe during numerical localization. Many localization tools attempt to filter out these effects (e.g., by using a Gaussian filter), although it is often necessary to manually adjust localization tolerances as a function of the signal-to-noise ratio of the typical PSF. Second, many localization events may emanate from the same molecular target in consecutive image frames. This is prevalent

when photoswitchable probes are used but in the most general case it arises due to a mismatch between image integration time and the probe ‘on-time.’ As will be discussed, for high-speed LM and SPT, these are not necessarily drawbacks. Third, faithful reconstruction of imaged probes demands a mechanically stable imaging apparatus, as a drift over the course of an acquisition may easily surpass 100 nm. To address this challenge, piezo stages with closed-loop feedback controls can negate stage drift effects. The most accessible approach, however, is the use of fiduciary markers, such as diffraction-limited sized beads, which may be used to correct for stage drift after imaging. Finally, there is a fundamental relationship between localization density (which is dependent on the duration of the acquisition) and the resolution of final pointillist reconstruction that must not be overlooked. Effectively, this is a trade-off between temporal and spatial resolution; the smaller the feature size, the more time is required for imaging (Shroff *et al.*, 2008).

Experimental Requirements

LM typically involves the use of an inverted fluorescence microscope with multiple laser light sources (Figure 2). The relatively small number of photons emitted from individual fluorescent probes demands the use of both high-NA oil-immersion objective lenses (typically 60x or 100x magnification) and sensitive electron multiplying charge coupled device (EMCCD) cameras. To ensure that cells remain relatively unperturbed during experimentation, both the stage on which the sample is mounted and the imaging objective lens are heated to roughly 37 °C. Chambers to control levels of CO₂ and the humidity are also commonly used, especially for long duration experiments.

Wide-field (WF) microscopy is the simplest way of performing LM. Laser light is focused in the back focal plane (BFP) of the imaging objective to collimate light propagating toward the sample. Due to the short exposure times needed for LM, high laser powers of the order of 0.1–1 kW cm^{−2} at the

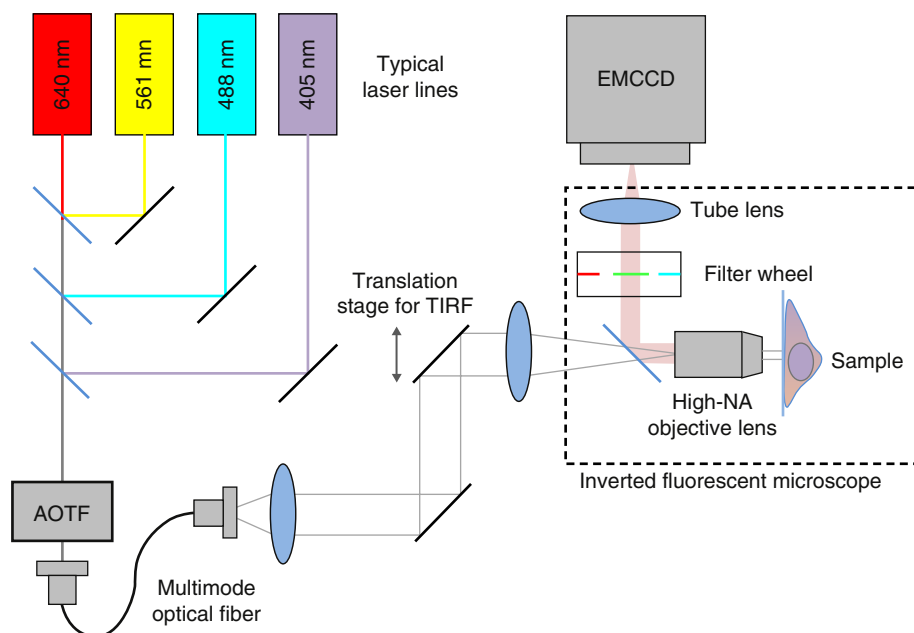


Figure 2 Schematic of the typical experimental setup used in SPT and high-speed LM microscopy. Laser lines are selected with an acousto-optic tunable filter (AOTF), the output of which is coupled into a multimode fiber. The output of the fiber is collimated and subsequently focused to the BFP of the high-NA objective lens. A mirror on a translation stage may be used to switch between WF and TIRF modes of illumination. The objective lens serves to illuminate the sample with excitation light and to capture resulting emission light, after which it passes through an emission filter before being imaged by EMCCD.

sample are required. Finally, effective pixel sizes are made to conform to the spatial Nyquist sampling criterion for a diffraction-limited spot (i.e., roughly ~ 120 nm).

Total internal reflection fluorescence (TIRF) microscopy is a popular configuration for LM. It involves focusing the excitation laser in the objective BFP at a position shifted from the optical axis. As a result, the excitation beam reaches the coverslip at an angle. When this incidence angle is sufficiently large, the laser beam is not transmitted but reflected by the coverslip, except for an evanescent field extending roughly ~ 200 nm in depth that illuminates only the region of the cell immediately adjacent to the coverglass. The consequence of this type of excitation is the suppression of background cytosolic signal, greatly enhancing the signal-to-noise ratio of imaged fluorophores.

High-Speed LM

The principle of high-speed LM is to capture successive super-resolution ‘snapshots’ of molecular targets inside a live cell. This type of imaging modality permits monitoring of dynamics of cellular complexes at the single-molecule level. More specifically, it offers the potential to monitor molecular assemblies that drive cellular processes in physiological conditions.

Both PALM and STORM may be used in this configuration, although distinct tradeoffs exist, as discussed in Section ‘The Localization Concept.’ Bright fluorescent dyes may be used for imaging structures on the outside of the cell membrane, however, imaging inside the cell requires intrusive techniques. Imaging inside the cell is greatly simplified through the use of

transgenically expressed fluorescent proteins. For this reason, PALM is a less restrictive choice for high-speed LM.

From the outset, however, there are significant challenges with high-speed LM. The dynamic environment of a live cell limits what targets may be imaged and accurately reconstructed. Desired targets must be chosen such that they remain relatively immobilized in the duration of image acquisition, which may last 25–60 s for a proper reconstruction snapshot (Shroff *et al.*, 2008). For example, cytosolic structures may be exceedingly difficult to image, while membrane rafts may prove a more amenable target. This leads to a second difficulty: a sufficient number of frames must be captured to ensure a proper localization density for super-resolution reconstruction. The judicious choice of bright probes and short acquisition times may alleviate the time required for acquisition. As imaging occurs intermittently over many intervals, the reserve of unbleached probes may eventually be depleted. Moreover, this implies that a loss of localization density (and subsequent resolution) over the course of an acquisition is to be expected in high-speed LM.

In practice, high-speed LM has been prevalently used in studies of membrane organization (in TIRF mode). The technique has also notably been used to monitor clustering events within cells and the nucleus, such as the formation of RNA polymerase II in human cells (in WF configuration) (Cisse *et al.*, 2013). Figure 3 shows changes in the actin distribution of a synaptic spine in the hippocampal neuron of a live rat after the activation of glutamate receptors (Izeddin *et al.*, 2011). In contrast to diffraction-limited imaging, this application reveals nanoscopic morphological differences over time.

STORM microscopy can also be made compatible with high-speed LM through the use of bright photoswitchable dyes

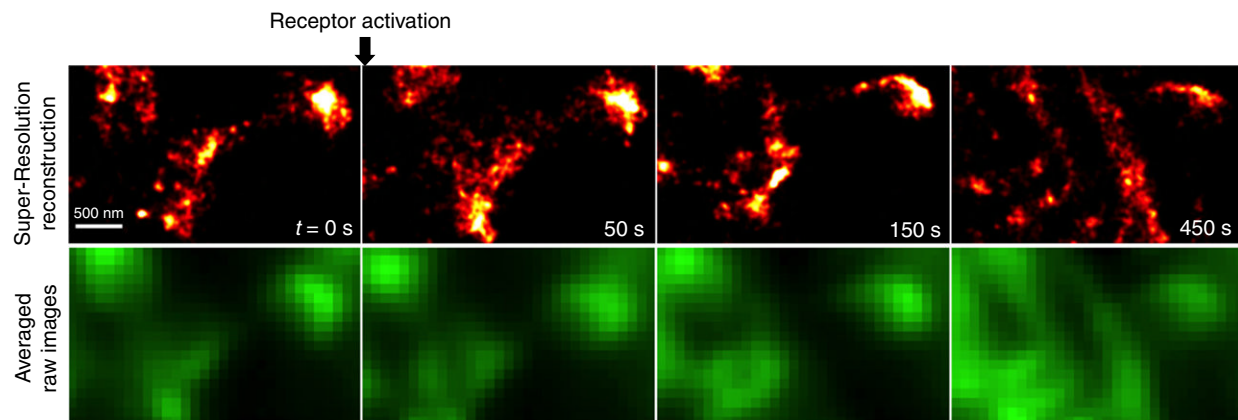


Figure 3 Structural dynamics of a synaptic spine visualized using high-speed LM in PALM mode (Izeddin *et al.*, 2011). Upper panels: super-resolution reconstruction of actin (ABP-EosFP construct) of a single spine at differently spaced time intervals. The spine head shrinks upon pharmacological application of the glutamate receptor agonist α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid receptor (AMPA). Bottom panels: corresponding raw image projections of frames used for super-resolution reconstruction.

and genetic fusion methods. Here, functionalized fluorescent dyes may be coupled with antibodies and bind to individual ligands at the cell membrane for imaging. For intracellular tagging, dyes may be functionalized and delivered inside the cell to bind to intracellularly expressed tags (e.g., SNAP tags or Halo tags). If photoswitchable dyes are used, there is the advantage that identical molecules may be monitored in successive super-resolution ‘snapshots.’ The higher brightness of dyes relative to fluorescent proteins also shortens the acquisition time required between snapshot intervals (Jones *et al.*, 2011).

SPT

Throughout the cell, the environment in which a single molecule translocates is generally highly heterogeneous and complex. Single-molecule motion may be affected by protein–protein and cytoskeletal interactions, but may also be influenced by active transport processes. SPT permits these events to be captured as they happen. As a subcategory of LM, SPT follows the same imaging and localization workflow, however with an additional step, that involves tracking of individual molecules based on their spatiotemporal coordinates (Saxton, 1993).

In comparison to diffraction-limited methods that measure molecule dynamics (such as fluorescence recovery after photobleaching microscopy) or molecule transiting events in confocal volumes (as in fluorescence correlation spectroscopy), SPT allows observation of different types of motion for the same molecule. Moreover, SPT has provided compelling evidence supporting the view that single molecules regulate cellular functions via a relentless ‘interplay’ between random motion and molecular interactions (Kusumi *et al.*, 2014).

There are a variety of probes that can be used for tracking single molecules in living cells. For example, long individual trajectories spanning hundreds of image frames may be acquired through the use of photostable-functionalized quantum dots (Dahan *et al.*, 2003). Fluorescent dyes may also be

coupled to ligands (e.g., antibodies or nanobodies) and bound to targets at the cell membrane to capture long and rich trajectories (Giannone *et al.*, 2010). Targets inside the cytoplasm, however, are difficult to reach with external fluorescent probes. Common approaches to introduce probes inside the cell may include microinjection, electroporation, and using fusogenic liposomes to transport probes inside the cell.

Single-particle tracking PALM (sptPALM) is a manner of tracking transgenically expressed photoactivatable fluorescent proteins (Manley *et al.*, 2008). Trajectories in this configuration are generally short, spanning only a few frames before fluorescence bleaching. The weak fluorescence signal of fluorescent proteins constitute another limitation for the applications of sptPALM. However, statistical techniques that consider only trajectory translocations (and not full-length trajectories) may stand to benefit from the high-density nature of sptPALM data.

SPT in many cases is a nontrivial and computationally expensive problem. A contributor to this difficulty is in cases of high localization density, which may render assignment of a specific localization to a given trajectory unclear. This task is additionally hampered when trajectories cross and when probes are prone to blinking. Finally, the diffusivity for the same tracked molecule may vary by orders of magnitude over distances as short as ~ 100 nm within the cell. Most tracking algorithms require an initial estimate of the diffusivity of the tracked species to evaluate the most probable connections between frames. For this reason, tracking of distinct diffusive populations of the same molecule imposes a challenge for single-molecule tracking. A result of the aforementioned challenges is that most high-density SPT methods are based on probabilistic notions of point connections within a trajectory, instead of simple nearest-neighbor approaches. A recent review describes in detail the performance of many commonly used tracking algorithms (Chenouard *et al.*, 2014).

Traditionally, trajectories may be analyzed by plotting how far a particle has traveled within a given time period. A plot of the mean square displacement (MSD) against the time period describes this relation and is commonly used to study single-particle trajectories, as described in Figure 4 and Table 2.

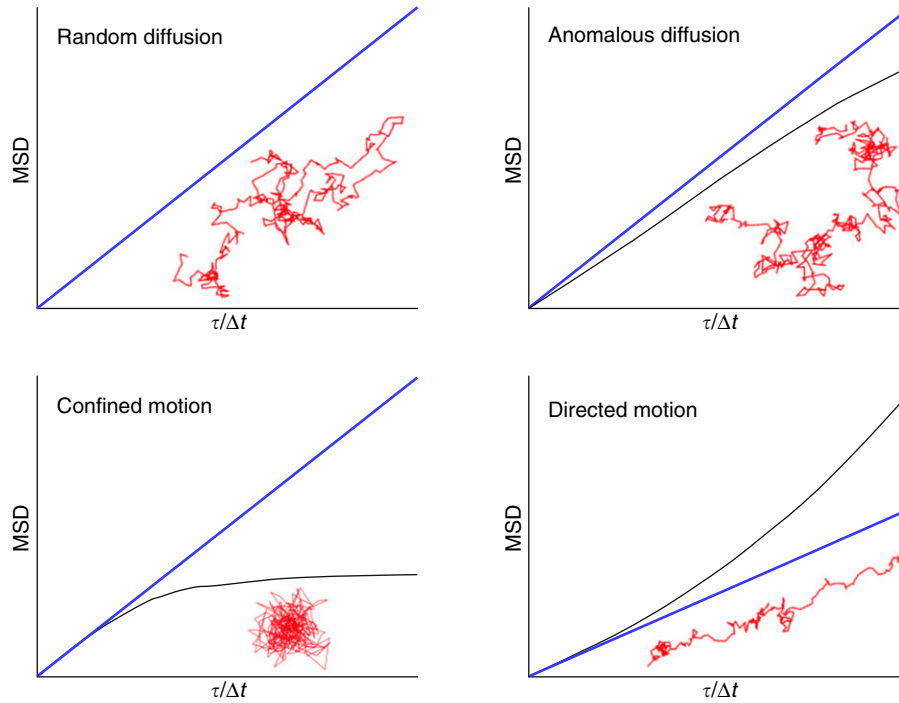


Figure 4 MSD against time lag divided by time step plots illustrating different types of motion observed in single-particle trajectories. The random diffusive trace is overlaid in blue on all plots.

Table 2 Different models of motion based on the mean square displacement where $X^2(\tau)$ is the lag time-dependent mean square displacement, τ is the lag time, D_{2D} is the 2D diffusion coefficient, α is the anomalous exponent, v is the magnitude of the particle velocity, R^2 is the size of the confinement area, and A_1 and A_2 are constants determined by the confinement geometry

Motion	Model	Possible explanations	References
Random diffusion	$X^2(\tau) = 4D_{2D}\tau$	Homogeneous environment	Saxton (1997)
Confined motion	$X^2(\tau) = R^2 \left(1 - A_1 e^{-A_2 \frac{4D_{2D}\tau}{R^2}} \right)$	Binding events Trapped particles Corralling Transient binding	Saxton and Jacobson (1997)
Anomalous diffusion	$X^2(\tau) = 4D_{2D}\tau^\alpha$	Hopping diffusion Obstacle presence	Feder <i>et al.</i> (1996)
Directed motion	$X^2(\tau) = 4D_{2D}\tau + (v\tau)^2$	Active/motor transport	Saxton (1994)

For purely diffusive (Brownian) motion where particle movement is due entirely to thermal agitations, this relationship is linear. In fact, a diffusion coefficient for a given trajectory can be estimated by taking the slope of this relation. For many biomolecules measured *in situ*, however, MSD relations are not linear for a variety of reasons. The media in which biomolecules are tracked are generally highly heterogeneous in composition and highly dynamic in their behavior. For example, tracked membrane proteins may interact with the actin cytoskeleton, focal adhesion points, or may undergo endocytosis. The diversity of models that exist to describe the cell membrane further justifies the consensus that the medium in which membrane proteins translocate is not homogeneous (Saxton and Jacobson, 1997). Theoretical expressions have been derived to describe different types of motion and how they may relate to these trends. Table 2 gives a summary of

these types of models of single-particle motion in the case of 2D trajectories.

Confined motion refers to cases in which a particle is immobilized within a certain region throughout the acquisition period. This may be the consequence of a binding event, such as a receptor becoming trapped in a lipid raft on the cell membrane, or due to the corralling of cytoskeletal polymers. A few ways of modeling confined trajectories have been proposed; in Table 2 we present the case for a circular 2D confinement area of size R^2 (Saxton and Jacobson, 1997).

Directed or active motion refers to cases where a tracked particle has a directional bias in a certain direction. This type of motion is regularly reported in actin- and microtubule-dependent active transport in the cytoplasm (Yildiz *et al.*, 2003).

Geometric effects additionally play an important role in the motion of single molecules. In environments such as the

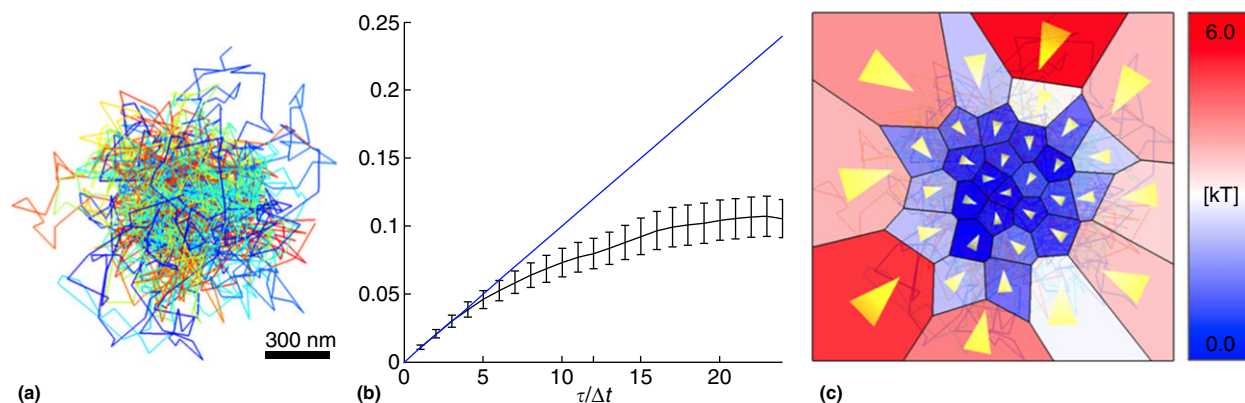


Figure 5 (a) Simulated trajectories with identical diffusion characteristics in the presence of a potential well. (b) In black, the MSD is plotted against the time lag (τ) divided by the time step (Δt) and demonstrates a trend consistent with that of a confined trajectory. In blue, the MSD of the purely random diffusive case without potential well. (c) Through spatial meshing and statistical inference, the interaction energy and directional bias of the trajectories may be resolved. The time step between trajectory points is 25 ms.

nucleus, there is evidence that the space explored by diffusing molecules possesses a fractal organization (Bancaud *et al.*, 2012). In this situation, obstacles at many different spatial scales affect the motion of diffusing molecules, resulting in an anomalous description of the motion. This type of behavior has been reported for transcription factors inside the nucleus (Récamiér *et al.*, 2014).

Additional effects on the MSD include the nonnegligible errors in localization of trajectory points. Acquisition times should generally be chosen such that tracked molecule displacements are greater than the localization precision. However, acquisition times greater than the characteristic time of molecule translocations may give rise to inaccurate estimates of the diffusivity (Montiel *et al.*, 2006). Nonetheless, altering the image acquisition time may permit observation of different diffusive populations.

MSD-based analysis may be subject to a number of drawbacks in a variety of contexts. First, it is particularly sensitive to noise, which may overestimate actual diffusion coefficients (Montiel *et al.*, 2006). Next, MSD values for large lag times are calculated with less data, hindering their accuracy. It is a common convention to calculate the MSD over τ values spanning less than 25% of the trajectory and measure the slope between the 2nd and 5th time lag to estimate a diffusion coefficient (Saxton and Jacobson, 1997). In the case of sptPALM, for example, trajectories are generally short (less than 10 translocation) and subject to position noise of the order of 30–50 nm. The net effect of these characteristics is a poor estimation of the MSD and, as a consequence, of the diffusion coefficient. Finally, while MSD plots may postulate an interaction via the presence of a trapping energy, there is no robust manner to extract information about the nature and strength of this type of interaction. Moreover, MSD analysis ostensibly models biomolecule motion as Brownian, upon which only deviations inform about the presence of an interaction.

Recent approaches have been introduced that rely on spatially partitioning the area explored by trajectories, and generating maps of the dynamics that describe the motion of tracked molecules (Masson *et al.*, 2009, 2013; Türkcan *et al.*, 2012). Such techniques are particularly well suited for

high-density trajectory data (such as in sptPALM) that invite the use of powerful statistical methods. The principle, in this case, is to infer a probability distribution for a dynamic parameter. The model of motion for the trajectory is chosen to account not only for the diffusion but also a possible interaction energy, which is systematically neglected in most MSD-based analyses. With this type of modeling, binding to molecular partners, crowding effects, and electrostatic interactions between molecules can be distinguished. Furthermore, an important advantage to this particular approach is that there is no restriction on the trajectory length. In other words, many short trajectories are statistically equivalent to a single long trajectory. Figure 5 shows how MSD analysis can evidence the presence of a potential energy trap for simulated confined trajectories. By using a statistical inference-based technique, the strength of the trap (as well as the directional bias it imposes) can be revealed and mapped with high accuracy and precision.

An important notion in SPT is that individual molecules switch between specific ‘states’ of motion at different moments in time. This may be the case of a molecule transitioning between an immobilized state (e.g., when bound to a scaffold) and a freely diffusing (unbound) state. This type of analysis requires tools that depart from typical MSD analysis that broadly categorizes trajectories (Figure 4). Specifically, this state description of single-molecule motion can be achieved by using a hidden Markov model (HMM) (Das *et al.*, 2009; Persson *et al.*, 2013). The HMM calculates the probability of a molecule assuming a state at a given time and may also be used to calculate transition rates between states to estimate molecular reaction kinetics.

Future Trends/Avenues

Single-molecule imaging garners significant interest in biological studies. The two methods discussed, high-speed LM and SPT, are critical tools in the study of modern cell biology and will benefit from new microscopy methods and data treatment methodologies in the future.

With regard to microscopy, progress is being made in three-dimensional (3D) single-particle imaging (Hajj *et al.*, 2014). Astigmatic imaging (Kao and Verkman, 1994), light-sheet microscopy (Chen *et al.*, 2014), adaptive optics (Izeddin *et al.*, 2012), biplane (Ram *et al.*, 2008) and multi-focus microscopy (Abrahamsson *et al.*, 2013) have already been used to this end. These techniques have the benefit of being able to capture the full range of motion of single molecules inside the cell, whereas 2D techniques are often limited to observing events on the membrane in TIRF mode illumination mode.

High-density SPT is of great interest, as the massive data it yields justifies the use of powerful statistical techniques for understanding whole-cell single-molecule dynamics with high spatial resolutions. Related, correlative studies with single-molecule imaging and observing gross cellular behavior will grant insight into the molecular basis of many cellular activities, such as polarization, mechanics, and migration.

The study of cell biology is inherently related to interactions between different types of molecules within the cell. To this end, multicolor SPT and high-density LM techniques permit the observation of interactions between different populations of molecules (Kapanidis *et al.*, 2004). It is worth noting, however, that this form of imaging is most amenable when fluorophores of narrow emission spectra are used (such as quantum dots) to avoid spectral mixing between channels.

On the data treatment end, much of the development in single-molecule research comes from techniques borrowed from statistical physics. Statistical inference methods, for example, enable mapping of the physical parameters that dictate particle motion. The application of HMMs to single-particle trajectories is also an important advance, and will further reinforce insight into the multiscale nature of diffusive processes inside the cell.

Conclusion

The introduction of single-molecule imaging in cellular biology has resulted in a paradigm shift in the way biological processes are understood (Ha, 2014). Applied to living cells, high-speed LM provides insight into the temporal evolution of biological complexes with nanoscale resolutions. SPT, in contrast, reveals unprecedented information of the environment experienced by individual biomolecules. Collectively, these are important tools that may be applied to a wide range of problems faced in modern cell biology.

See also: Imaging the Cell: Light Microscopy: Fluorescence Correlation Spectroscopy: A Tool for Measuring Dynamic and Equilibrium Properties of Molecules in Cells

References

- Abbe, E., 1873. Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung. *Archiv für Mikroskopische Anatomie* 9, 413–420.
- Abrahamsson, S., Chen, J., Hajj, B., *et al.*, 2013. Fast multifocus 3D imaging using aberration-corrected multifocus microscopy. *Nature Methods* 10 (1), 60–63.
- Bancaud, A., Lavelle, C., Huet, S., Ellenberg, J., 2012. A fractal model for nuclear organization: Current evidence and biological implications. *Nucleic Acids Research* 40 (18), 8783–8792.
- Betzig, E., Patterson, G.H., Sougrat, R., *et al.*, 2006. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313, 1642–1645.
- Chen, B.-C., Legant, W., Wang, K., *et al.*, 2014. Lattice light-sheet microscopy: Imaging molecules to embryos at high spatiotemporal resolution. *Science* 346, 1257998.
- Chenouard, N., Smal, I., de Chaumont, F., *et al.*, 2014. Objective comparison of particle tracking methods. *Nature Methods* 11 (3), 281–289.
- Cisse, I.I., Izeddin, I., Causse, S.Z., *et al.*, 2013. Real-time dynamics of RNA polymerase II clustering in live human cells. *Science* 341, 664–667.
- Dahan, M., Lévi, S., Luccardini, C., *et al.*, 2003. Diffusion dynamics of glycine receptors revealed by single-quantum dot tracking. *Science* 302, 442–445.
- Das, R., Cairo, C.W., Coombs, D., 2009. A hidden Markov model for single particle tracks quantifies dynamic interactions between LFA-1 and the actin cytoskeleton. *PLoS Computational Biology* 5 (11), e1000556.
- Feder, T.J., Brust-Mascher, I., Slattery, J.P., Baird, B., Webb, W.W., 1996. Constrained diffusion or immobile fraction on cell surfaces: A new interpretation. *Biophysical Journal* 70, 2767–2773.
- Giannone, G., Hosy, E., Levet, F., *et al.*, 2010. Dynamic superresolution imaging of endogenous proteins on living cells at ultra-high density. *Biophysical Journal* 99, 1303–1310.
- Ha, T., 2014. Single-molecule methods leap ahead. *Nature Methods* 11 (10), 1015–1018.
- Hajj, B., El Beheiry, M., Izeddin, I., Darzacq, X., Dahan, M., 2014. Accessing the third dimension in localization-based super-resolution microscopy. *Physical Chemistry Chemical Physics* 16 (31), 16340–16348.
- Izeddin, I., El Beheiry, M., Andilla, J., *et al.*, 2012. PSF shaping using adaptive optics for three-dimensional single-molecule super-resolution imaging and tracking. *Optics Express* 20 (5), 4957–4967.
- Izeddin, I., Specht, C.G., Lelek, M., *et al.*, 2011. Super-resolution dynamic imaging of dendritic spines using a low-affinity photoconvertible actin probe. *PLoS ONE* 6 (1), e15611.
- Jones, S.A., Shim, S.-H., He, J., Zhuang, X., 2011. Fast, three-dimensional super-resolution imaging of live cells. *Nature Methods* 8 (6), 499–505.
- Kao, H.P., Verkman, A.S., 1994. Tracking of single fluorescent particles in three dimensions: Use of cylindrical optics to encode particle position. *Biophysical Journal* 63 (3), 1291–1300.
- Kapanidis, A.N., Lee, N.K., Laurence, T.A., *et al.*, 2004. Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. *Proceedings of the National Academy of Sciences* 101 (4), 8936–8941.
- Kusumi, A., Tsunoyama, T.A., Hirose, K.M., Kasai, R.S., Fujiwara, T.K., 2014. Tracking single molecules at work in living cells. *Nature Chemical Biology* 10, 524–532.
- Levi, V., Gratton, E., 2007. Exploring dynamics in living cells by tracking single particles. *Cell Biochemistry and Biophysics* 48, 1–15.
- Manley, S., Gillette, J.M., Patterson, G.H., *et al.*, 2008. High-density mapping of single-molecule trajectories with photoactivated localization microscopy. *Nature Methods* 5 (2), 155–157.
- Masson, J.-B., Casanova, D., Türkcan, S., *et al.*, 2009. Inferring maps of forces inside cell membrane microdomains. *Physical Review Letters* 102, 048103.
- Masson, J.-B., Dionne, P., Salvatico, C., *et al.*, 2013. Mapping the energy and diffusion landscapes of membrane proteins at the cell surface using high-density single-molecule imaging and Bayesian inference: Application to the multiscale dynamics of glycine receptors in the neuronal membrane. *Biophysical Journal* 106 (1), 74–83.
- Montiel, D., Cang, H., Yang, H., 2006. Quantitative characterization of changes in dynamical behavior for single-particle tracking studies. *Journal of Physical Chemistry B* 110, 19763–19770.
- Mortensen, K.I., Churchman, L.S., Spudich, J.A., Flyvbjerg, H., 2010. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nature Methods* 7 (5), 377–381.
- Persson, F., Lindén, M., Unoson, C., Elf, J., 2013. Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nature Methods* 10 (3), 265–269.
- Pinaud, F., Clarke, S., Sittner, A., Dahan, M., 2010. Probing cellular events, one quantum dot at a time. *Nature Methods* 7 (4), 275–285.
- Ram, S., Prabhat, P., Chao, J., Ward, E.S., Ober, R.J., 2008. High accuracy 3D quantum dot tracking with multifocal plane microscopy for the study of fast intracellular dynamics in live cells. *Biophysical Journal* 95 (12), 6025–6043.

- Récamier, V., Izeddin, I., Bosanac, L., *et al.*, 2014. Single cell correlation fractal dimension of chromatin: A framework to interpret 3D single molecule super-resolution. *Nucleus* 5 (1), 75–84.
- Rust, M.J., Bates, M., Zhuang, X., 2006. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods* 3 (10), 793–795.
- Ruthardt, N., Lamb, D.C., Bräuchle, C., 2011. Single-particle tracking as a quantitative microscopy-based approach to unravel cell entry mechanisms of viruses and pharmaceutical nanoparticles. *Molecular Therapy* 19 (7), 1199–1211.
- Saxton, M.J., 1993. Lateral diffusion in an archipelago: Single-particle diffusion. *Biophysical Journal* 64, 1766–1780.
- Saxton, M.J., 1994. Single-particle tracking: Models of directed transport. *Biophysical Journal* 67, 2110–2119.
- Saxton, M.J., 1997. Single-particle tracking: The distribution of diffusion coefficients. *Biophysical Journal* 72, 1744–1753.
- Saxton, M.J., Jacobson, K., 1997. Single-particle tracking: Applications to membrane dynamics. *Annual Review of Biophysics and Biomolecular Structure* 26, 373–399.
- Shroff, H., Galbraith, C., Galbraith, J.A., Betzig, E., 2008. Live-cell photoactivated localization microscopy of nanoscale adhesion dynamics. *Nature Methods* 5 (5), 417–423.
- Small, A., Stahlheber, S., 2014. Fluorophore localization algorithms for super-resolution microscopy. *Nature Methods* 11 (3), 267–279.
- Thompson, R.E., Larson, D.R., Webb, W.W., 2002. Precise nanometer localization analysis for individual fluorescent probes. *Biophysical Journal* 82, 2775–2783.
- Türkcan, S., Alexandrou, A., Masson, J.-B., 2012. A Bayesian inference scheme to extract diffusivity and potential fields from confined single-molecule trajectories. *Biophysical Journal* 102 (10), 2288–2298.
- Yildiz, A., Forkey, J.N., McKinney, S.A., *et al.*, 2003. Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization. *Science* 300, 2061–2065.

Intravital Microscopy in Mammalian Organisms: From Tissue Physiology to Cell Biology

R Weigert, National Institutes of Health, Bethesda, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Brief Historical Background on Intravital Microscopy

Since the development of the first optical microscope in the late 1500s, biologists have been intrigued by the idea of imaging the dynamics of biological processes in live organisms. The first example of intravital microscopy (IVM) in a living animal is dated 1839, when Rudolph Wagner described the interaction of leukocytes with the walls of blood vessels in the webbed feet of a live frog by using bright field transillumination (Wagner, 1839). Since then, this approach has been extensively used particularly in thin transparent tissue to study vascular biology (Clark and Clark, 1932; Irwin and Macdonald, 1953; Zweifel, 1954) or cell migration (Thorogood and Wood, 1987; Wood and Thorogood, 1984). A major step forward in IVM came with the development of epifluorescence microscopy which enabled more detailed investigations of the dynamics of individual cells in the circulation (Nuttall, 1987), in tumors (MacDonald *et al.*, 1992), or in the immune system (von Andrian, 1996), and with the introduction of the confocal microscope, which has made possible to collect serial optical sections from a given specimen, thus significantly increasing the spatial resolution (Jester *et al.*, 1991; O'Rourke and Fraser, 1990; Villringer *et al.*, 1989). The main limitation of these techniques has been tissue light scattering, which limits imaging to either transparent tissue in small organisms or up to a few microns below the surface of mammalian organs. We had to wait for the early 1990s to extend the range of observations to deep tissue, thanks to the development of multiphoton microscopy (MPM) by the Webb's lab (Denk *et al.*, 1990; Zipfel *et al.*, 2003b). Initially, this technique was used in the neurosciences (Svoboda and Yasuda, 2006), and only later extended to other areas such as immunology and cancer biology (Amornphimoltham *et al.*, 2011; Pittet and Weissleder, 2011; Weigert *et al.*, 2010). Finally, in the last decade the development of strategies to minimize the motion artifacts due to the heartbeat and respiration has made possible to successfully image subcellular structures with spatial and temporal resolution comparable to those achieved in *in vitro* model systems, thus opening the door to the use of IVM to study cell biology (Weigert *et al.*, 2013).

Imaging Techniques Used to Perform IVM

IVM is performed primarily using laser-scanning confocal and MPM. Confocal microscopy is a well-established technique that is based on single photon excitation (Figure 1(a)), and relies on lasers emitting in the UV-visible range, as a light source for the excitation of the specimen. This technique uses pinholes to select the emitted light generated from a specific optical section, thus enhancing the spatial resolution and enabling the visualization of small structures in the submicron

range (Wilson, 2002). However, due to the high scattering of UV-visible light in biological samples, confocal microscopy can be used to samples up to 80–100 μm in depth (Figure 1(a); Masedunskas *et al.*, 2012a). Moreover, confocal microscopy is not suitable for long-term imaging (more than 30–60 min) due to out-of-focus photo-damage and photo-bleaching that may generate a pathological tissue response due to increased oxidative stress.

MPM is one of the nonlinear optical microscopy techniques that are based on generating contrast by higher-order interactions between light and the biological specimen, which involve the absorption or the scattering and recombination of two or more photons (Campagnola and Loew, 2003; Helmchen and Denk, 2005; Mertz, 2004; Oheim *et al.*, 2006; Rubart, 2004; Stutzmann and Parker, 2005; Svoboda and Yasuda, 2006; Zipfel *et al.*, 2003b; Figure 1(b)). The first multiphoton microscope (Denk *et al.*, 1990) was based on the principle of two-photon excitation postulated by Maria Goeppert-Mayer in her PhD thesis (Göppert-Mayer, 1931). Multiphoton excitation is based on the fact that a molecule of a fluorophore can be excited by the almost simultaneous absorption (within atto- or femto-seconds) of two (three) photons that have half (one-third) of the energy that is required to fill the gap between two of its energetic levels (Figure 1(b); Helmchen and Denk, 2005; Zipfel *et al.*, 2003b). This is achieved by using lasers that emit near-infrared/infrared (NIR/IR) light (from 690 to 1600 nm) in very short pulses (in the order of 100 fs) at high-repetition rates (80–100 MHz), and by using high numerical aperture lenses that focus the light to the excitation spot. In ideal conditions multiphoton absorption occurs in a very small volume (1 fl) (Helmchen and Denk, 2005; Zipfel *et al.*, 2003b). This results in the absence of off-focus emissions, thus eliminating the need for a pinhole and reducing significantly both photo-toxicity and photo-bleaching. Moreover, since NIR/IR light has a lower tissue scattering (i.e., deeper tissue penetration) than UV or visible light (Theer and Denk, 2006), MPM can resolve structures up to a depth of 300–500 μm in most of the tissues, and up to 1–1.5 mm in the brain (Masedunskas *et al.*, 2012a; Theer *et al.*, 2003).

Other nonlinear optical techniques have been used for IVM, and among them second harmonic generation (SHG) and third harmonic generation (THG), two processes that do not involve energy absorption since the incident photons are recombined into a single photon of a higher energy in a process without energy loss (Figure 1(c)). For this reason, they are suitable for imaging biological specimen with even lower photo-toxicity than MPM (Campagnola and Loew, 2003). Although the harmonic signals are produced in the forward direction and thus more suited for imaging slices, the backward scattering signal is still sufficient to generate details about the tissue architecture. Second harmonic signals are generated by specific molecules that are assembled in highly ordered

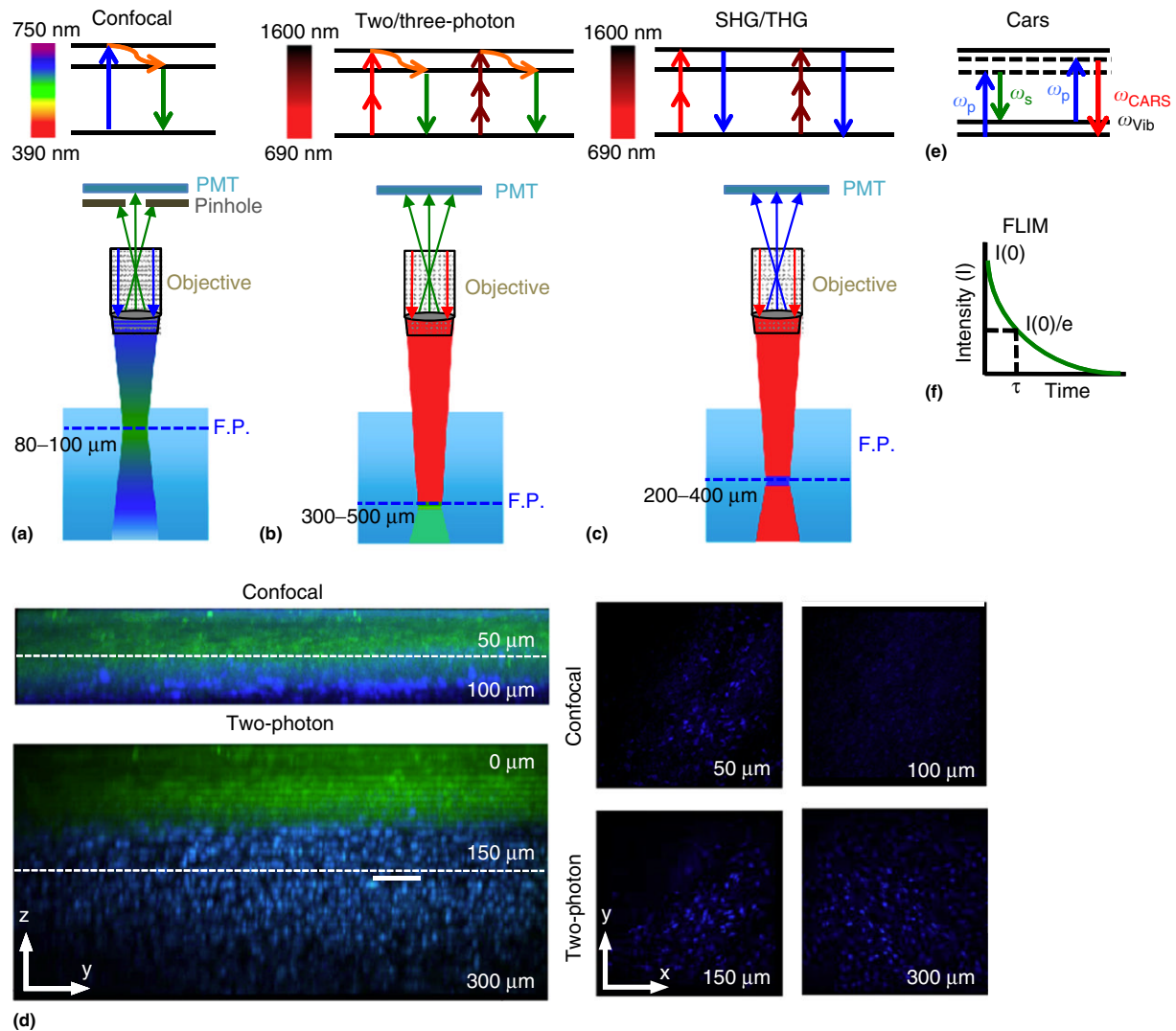
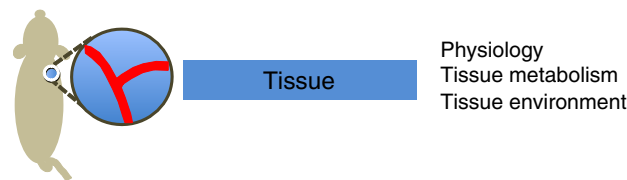


Figure 1 Fluorescent light microscopy imaging techniques used for IVM. (a) Confocal microscopy. In confocal microscopy a fluorophore absorbs a single photon with a wavelength in the UV-visible range of the spectrum (blue arrow). After a vibrational relaxation (orange curved arrow), a photon with a wavelength shifted toward the red is emitted (green arrow). In thick tissue, excitation and emission occur in a relative large volume around the focal plane (FP). The off-focus emissions are eliminated through a pinhole, and the signal from the FP is detected via a photomultiplier (PMT). Confocal microscopy enables imaging at a maximal depth to 80–100 μm . (b) Two- and three-photon microscopy. In this process, a fluorophore absorbs almost simultaneously two or three photons that have half (red arrow) or a third (dark red arrow) of the energy required for its excitation with a single photon. Two or three photon excitation typically require near IR or IR light (from 690 to 1600 nm). Emission and excitation occur only at the FP in a restricted volume (1.5 fl) and for this reason a pinhole is not required. Two- and three-photon microscopy enable imaging routinely at a maximal depth of 300–500 μm . (c) SHG and THG. In SHG and THG, photons interact with the specimen and combine to form new photons that are emitted with twice or three times their initial energy without any energy loss. These processes have similar features to those described for two- and three-photon microscopy and enable imaging at a maximal depth of 200–400 μm . (d) A mouse engineered to express a membrane-targeted peptide fused to GFP (m-GFP) in keratin 14-expressing cells was injected with the nuclear dye Hoechst, and the tongue was imaged by either confocal or two-photon microscopy. A z-scan shows that nuclei (blue) can be imaged up to 100 μm (confocal) or 300 μm (two-photon). (e) In CARS microscopy two beams are used: the pump (ω_p) and the Stokes (ω_s). When they are tuned to match a vibrational energy gap (ω_{vib}), a strong anti-stokes signal is generated (ω_{CARS}). (f) In FLIM, a fluorophore is excited and the decay of the fluorescence emitted is measured. The lifetime τ is calculated based on $I(t) = I(0)e^{-t/\tau}$. Adapted from Weigert, R., 2014. Imaging the dynamics of endocytosis in live mammalian tissues. *Cold Spring Harbor Perspectives in Biology* 6, a017012; Masedunskas, A., Milberg, O., Porat-Shliom, N. *et al.*, 2012a. Intravital microscopy: A practical guide on imaging intracellular structures in live animals. *BioArchitecture* 2, 143–157; Masedunskas, A., Porat-Shliom, N., Rechache, K., Aye, M.P., Weigert, R., 2012b. Intravital microscopy reveals differences in the kinetics of endocytic pathways between cell cultures and live animals. *Cells* 1, 1121–1132; and Weigert, R., Sramkova, M., Parente, L., Amornphimoltham, P., Masedunskas, A., 2010. Intravital microscopy: A novel tool to study cell biology in living animals. *Histochemistry and Cell Biology* 133, 481–491.

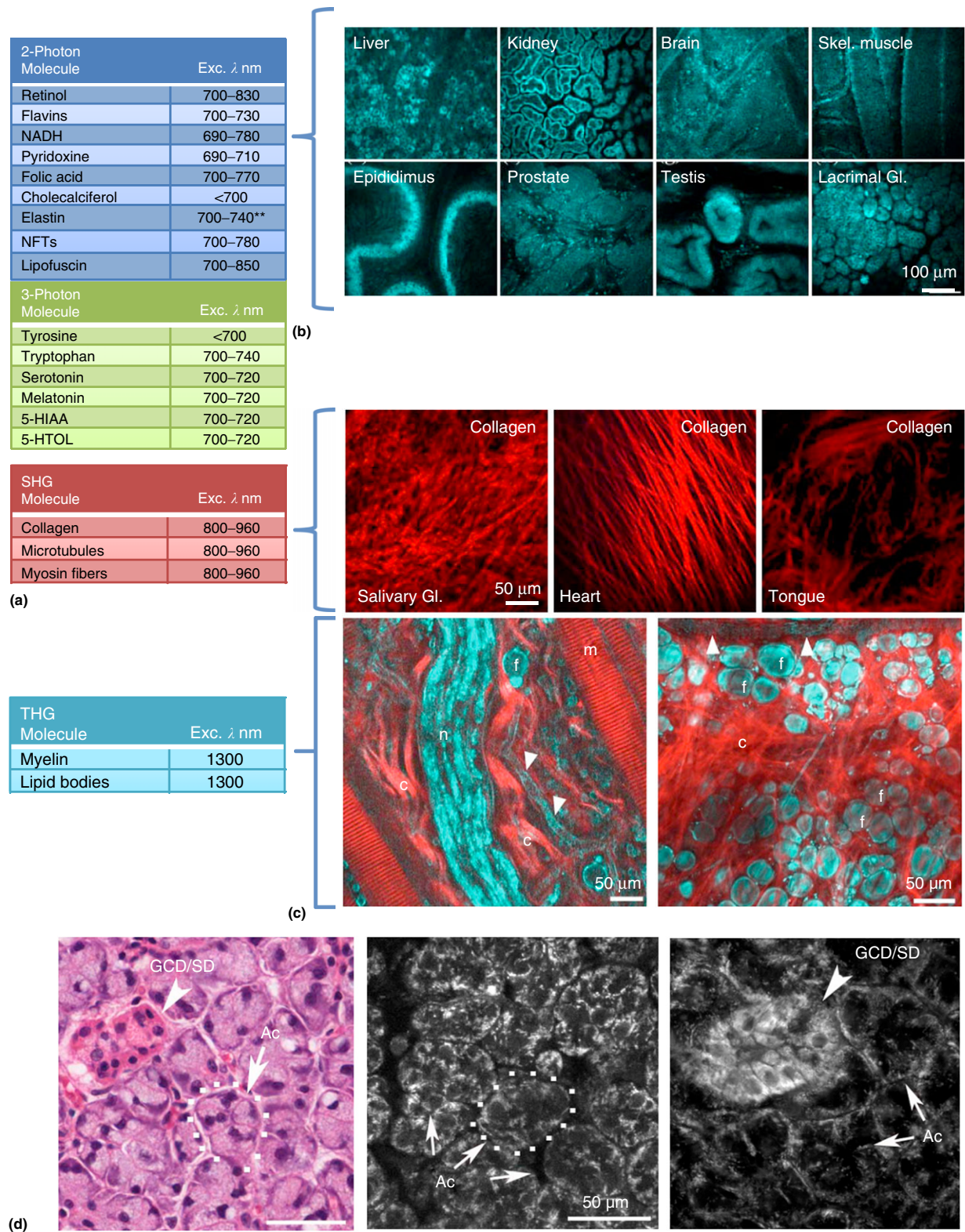
structures, such as collagen fibers, microtubules, and muscle myosin (Campagnola and Loew, 2003; Zipfel *et al.*, 2003a; Zoumi *et al.*, 2002). On the other hand, third harmonic signals are generated at the interfaces between tissue components with different optical properties. For example, THG have been used to image lipid bodies in various living organisms (Debarre *et al.*, 2006) and to visualize myelin fibers during cancer invasion (Weigelin *et al.*, 2012). Both SHG and THG are often combined with MPM, thus expanding the repertoire of information that can be acquired (Weigelin *et al.*, 2012), and recently, up to six intrinsic signals coming from both multiphoton and harmonic emissions were resolved in the skin of live nude mice using spectral unmixing techniques (Radosevich *et al.*, 2008). Three other techniques have been recently used for IVM with very promising results: Coherent anti-Stokes Raman scattering (CARS), fluorescence lifetime imaging (FLIM), and optical frequency domain imaging (OFDI). CARS utilizes two combined laser beams to match the energy gap between two vibrational levels of the molecule of interest (Figure 1(d)). Similarly to THG, CARS has been used to image lipids and myelin fibers *in vivo* (Fu *et al.*, 2008; Le

et al., 2010; Muller and Zumbusch, 2007). FLIM measures the lifetime that a molecule spends in the excited state (Figure 1(e)). Since the lifetime depends on the characteristics of the environment, this approach provides quantitative information on cellular parameters such as pH, oxygen levels, ions concentration, and the metabolic state of various biomolecules (Bakker *et al.*, 2012; Levitt *et al.*, 2009; Provenzano *et al.*, 2009). OFDI is based on focusing a laser beam into the tissue and combining the light reflected across all depths with a reference beam. The analysis of the interference signal is utilized to derive the optical scattering properties enabling to determine structural information on the tissue (Vakoc *et al.*, 2009). OFDI ensures a deeper tissue penetration (1 mm), a faster acquisition time, and does not require any exogenous labeling, although it exhibits a lower spatial resolution when compared with MPM. Recently, a high-resolution three-dimensional analysis of the tumor microenvironment and its response to a targeted therapy has been reported in the absence of any exogenous labeling highlighting the enormous potential of this technique for diagnostic purposes and basic research (Vakoc *et al.*, 2009).

Table 1 Intravital microscopy to image tissue architecture and tissue response



Event	Organ	Probes	References
Measurements of local blood flow and glial cell function	Brain	Dextran	Helmchen and Kleinfeld (2008)
Ischemia and reperfusion	Brain	Sulphorhodamine 101, Dextran	Masamoto <i>et al.</i> (2012); Zhang and Murphy (2007)
Glomerular filtration and tubular reabsorption	Kidney	Dextran, Albumin	Camirand <i>et al.</i> (2011); Kang <i>et al.</i> (2006); Yu <i>et al.</i> (2007)
Blood flow patterns	Pancreatic islets	Dextran	Nyman <i>et al.</i> (2008)
Capillary response and synaptic activation	Olfactory bulb	Dextran	Chaigneau <i>et al.</i> (2003)
Imaging angiogenesis during wound healing	Skullcap	Dextran	Holstein <i>et al.</i> (2011)
Pulmonary microvasculature activation	Lung	Dextran	Presson <i>et al.</i> (2011)
Heart microvasculature	Heart	Dextran	Vinegoni <i>et al.</i> (2012)
Retina vasculature	Eye	Dextran	Alt <i>et al.</i> (2012)
Morphology of blood vessels and permeability in tumors	Xenografts	Dextran, RGD quantum dots	Fukumura <i>et al.</i> (2010); Smith <i>et al.</i> (2008); Tozer <i>et al.</i> (2005); Vakoc <i>et al.</i> (2009)
Hepatic transport into the bile canaliculi	Liver	Carboxyfluorescein diacetate Rhodamine 123	Babbey <i>et al.</i> (2012); Liu <i>et al.</i> (2012)
Progression of amyloid plaques in Alzheimer's disease	Brain	Curcumin and metoxy-04	Garcia-Alloza <i>et al.</i> (2007); Spires <i>et al.</i> (2005)
Mitochondrial membrane potential	Liver Brain Kidney	Tetramethylrhodamine methyl ester Rhodamine 123, NADH	Theruvath <i>et al.</i> (2008); Zhong <i>et al.</i> (2008); Guan <i>et al.</i> (2009); Wu <i>et al.</i> (2007); Hall <i>et al.</i> (2013)
Oxygen consumption	Liver	Ru(phen3)2 +	Paxian <i>et al.</i> (2004)
Sarcomere contraction in humans	Skeletal muscle	Endogenous fluorescence	Llewellyn <i>et al.</i> (2008)



Current Applications of IVM

IVM is currently used in several areas of the biomedical field, which include physiology, neuroscience, immunology, tumor biology, stem cell biology, pathogen invasion, and cell biology. To properly illustrate the power of this approach the spatial resolution will be used as the main organizing criteria. First, the application of IVM to investigate the relationship between tissue architecture and function will be discussed; next, the developments in tracking individual cells and; finally, the most recent applications in studying the dynamics of subcellular structures *in vivo*.

Imaging Tissue Architecture and Function *In Vivo*

One of the main uses of IVM has been in the investigation of tissue response under both physiological and pathological conditions (Table 1). In this respect, the main advantage of this approach is the fact that various aspects of tissue physiopathology cannot be reconstituted in *in vitro* or *ex vivo* model systems (Weigert, 2014; Weigert *et al.*, 2013). The success of IVM in this arena has relied primarily on the availability and the development of a large series of tools and experimental approaches to contrast tissues *in vivo*. For example, tissue architecture has been highlighted by exploiting the fact that several endogenous molecules can be excited using either linear or nonlinear optical techniques (Zipfel *et al.*, 2003a; Figure 2(a)).

Among them is nicotinamide adenine dinucleotide (NADH), that emits in the visible range upon either single photon (360 nm) or two-photon excitation (710–760 nm). Although its two-photon cross section is very low, its abundance within the cell makes it a suitable endogenous label for both metabolic and structural studies (Figure 2(b)). Indeed, changes in the levels of NADH were measured in live mice during ischemia and reperfusion in the jejunum and the brain (Guan *et al.*, 2009), during microcirculatory failure in the liver (Paxian *et al.*, 2004) or in the kidney after lipopolysaccharide-induced sepsis (Wu *et al.*, 2007). These studies have provided novel insights on the acute and early changes in the metabolic state of tissues under various pathological conditions. Importantly, since NADH is distributed both in the cytoplasm and in the mitochondria, its emissions provide the contrast to highlight several details of the tissue architecture, thus

enabling the correlation between structure and metabolic activity. For example, in the brain, astrocytes can be distinguished from other cell populations based on the characteristic NADH pattern (Kasischke *et al.*, 2004); in skeletal muscles of both live animals and humans the architecture and the contractions of the sarcomere have been monitored in time-lapse modality by the NADH emissions from the mitochondria along the sarcomere Z-discs (Llewellyn *et al.*, 2008; Rothstein *et al.*, 2005); similarly, in the salivary glands of both live rats and mice, the architecture of the secretory acini and the fine details of the structure of the large ducts have also been imaged at a level of resolution almost comparable to that obtained by classical immunohistochemistry (Masedunskas and Weigert, 2008; Sramkova *et al.*, 2009; Figure 2(d)). Another molecule that has been extensively exploited for *in vivo* imaging is collagen, which generates a strong SHG signal when arranged in fibers (Campagnola and Loew, 2003; Zoumi *et al.*, 2002). In particular, detecting collagen has been extremely helpful in characterizing fibrotic tissues under various pathological conditions (Cox *et al.*, 2003; Megens *et al.*, 2007; Morishige *et al.*, 2006; Pena *et al.*, 2007; Schenke-Layland *et al.*, 2008; Figure 2(c)). Moreover, collagen fibers in live animals are a valuable reference point within the tissue, particularly in the context of tumor migration and invasion since they provide the means to correctly locate and orient tumors that are repeatedly imaged over long period of times (Brown *et al.*, 2003; Perentes *et al.*, 2009; Wyckoff *et al.*, 2007; see Figure 5).

Another approach to investigate tissue response *in vivo* is through the use of exogenously administered tracers. In particular, the vasculature is one of the main tissue components that have been visualized by IVM since it is relatively easy to label by systemic injections of fluorescently labeled probes (Movies 1–4; Figure 3(a)). This approach has provided not only the contrast to resolve the structure of the blood vessels but also a tool to measure the blood flow locally (Helmchen and Kleinfeld, 2008). For example, in mice and rat models for stroke, alterations in the vascular flow and in the size of the capillaries were monitored in the brain cortex during ischemia and reperfusion (Masamoto *et al.*, 2012). Systemic injections of fluorescently labeled bovine serum albumin (BSA) or dextrans have been used for a variety of applications such as, to measure glomerular filtration and tubular reabsorption in the kidney both in normal and ischemic conditions (Camirand *et al.*, 2011; Kang *et al.*, 2006; Yu *et al.*, 2007; Figure 3(b)), to determine

Figure 2 Excitation of intrinsic fluorescence to image tissue architecture *in vivo*. (a) List of the most common endogenous molecules excited by nonlinear microscopy techniques. (b) Excitation of NADH in rat tissues by 2P microscopy. Tissue architecture of various rat organs (liver, kidney, brain cortex, skeletal muscle, epididymus, bladder, prostate, and lacrimal glands) was observed by 2P microscopy using 740 nm as excitation wavelength to reveal cellular and mitochondrial NADH, which provides tissue contrast. (c) Second and third harmonic signals from rat and mouse tissues. Upper panels: Collagen fibers were revealed by SHG (excitation 930 nm) in the salivary gland and heart of rats, and in the tongue of a mouse. Lower panels: THG/SHG three-dimensional reconstruction of tissue interfaces in the mouse dermis (1180 nm excitation wavelength). THG (cyan) was generated by fat cells (f), nerve fibers (n), flowing erythrocytes labeling perfused blood vessels (arrowheads, left panel), and cell bodies interspersed in loose connective tissue (arrowheads, right panels). SHG (red) was confined to collagen bundles (c) and striated muscle fibers (m) (courtesy of Peter Friedl, adapted from Weigert, R., Sramkova, M., Parente, L., Amornphimoltham, P., Masedunskas, A., 2010. Intravital microscopy: A novel tool to study cell biology in living animals. *Histochemistry and Cell Biology* 133, 481–491). (d) Comparison between conventional H&E staining and 2P microscopy – H&E staining of rat salivary glands (left panel) and 2P microscopy (excitation wavelength 740 nm) of the exposed glands *in situ* (center and right panels). Acini (dotted line and arrow) and large ducts (arrowhead) are clearly visible. Adapted from Weigelin, B., Bakker, G., Friedl, P., 2012. Intravital third harmonic generation microscopy of collective melanoma cell invasion: Principles of interface guidance and microvesicle dynamics. *Intravital* 1, 32–43 and Sramkova, M., Masedunskas, A., Parente, L., Molinolo, A., Weigert, R., 2009. Expression of plasmid DNA in the salivary gland epithelium: Novel approaches to study dynamic cellular processes in live animals. *American Journal of Physiology Cell Physiology* 297, C1347–C1357.

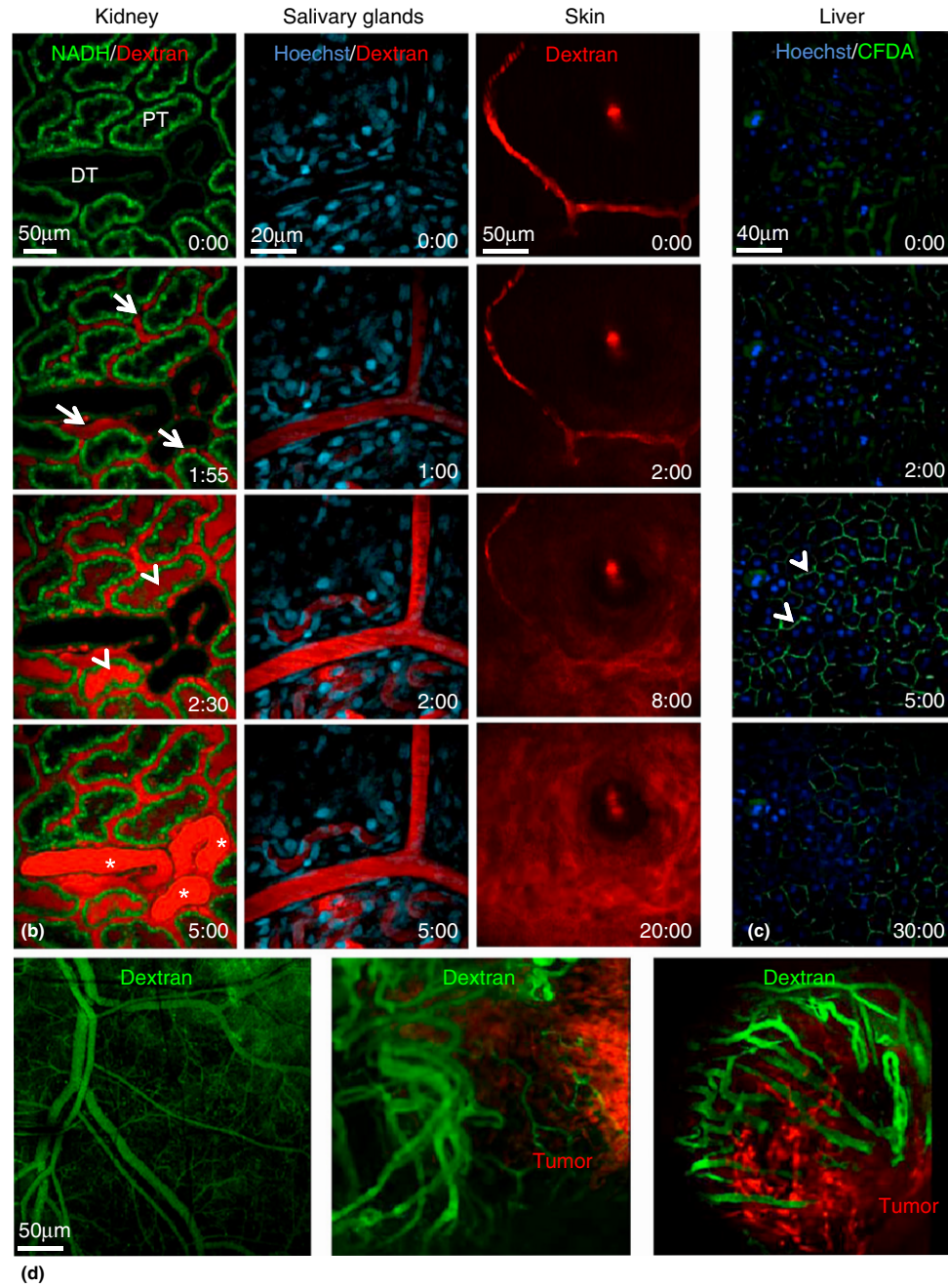
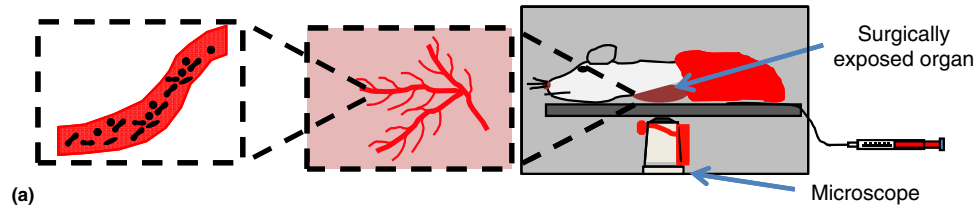
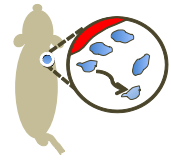


Table 2 Intravital microscopy to image the dynamics of individual cells


Individual cell

Neurobiology Cancer Biology
Immunology Stem Cells Biology

Event	Organ	Probe	References
Neuronal morphology of hippocampal neurons	Brain	Thy1-GFP mice, dextran	Barretto <i>et al.</i> (2011)
Neuronal circuitry	Brain	Brainbow mice	Livet <i>et al.</i> (2007)
Dendritic spine development in the cortex	Brain	YFP H-line mice	Pan and Gan (2008)
Calcium imaging in the brain	Brain	GCAMP	Zariwala <i>et al.</i> (2012)
Natural killer and cytotoxic T lymphocyte interactions with tumors	Xenograft	mCFP, mYFP	Deguine <i>et al.</i> (2010)
Neutrophil recruitment in beating heart	Heart	Dextran, CX3CR1-GFP mice	Li <i>et al.</i> (2012)
Immune cells in the central nervous system	Brain	Dextran, CX3CR1-GFP, LysM-GFP and CD11c-YFP mice	Nayak <i>et al.</i> (2012)
Dendritic cells migration	Skin	YFP, VE-cadherin RFP mice, dextran	Nitschke <i>et al.</i> (2008)
CD8 + T cells interaction with dendritic cells during viral infection	Lymph nodes	EGFP, Dextran, SHG	Hickman <i>et al.</i> (2008)
B cells and dendritic cells interactions outside lymph nodes	Lymph nodes	EGFP	Qi <i>et al.</i> (2006)
Change in gene expression during metastasis	Xenograft		Pinner <i>et al.</i> (2009)
Invasion and metastasis in head and neck cancer	Xenograft	YFP, RFP-lifect, dextran	Amornphimoltham <i>et al.</i> (2011)
Fibrosarcoma cell migration along collagen fibers	Dorsal skin chamber	SHG, EGFP, DsRed, Dextran	Alexander <i>et al.</i> (2008)
Long-term imaging mammary tumors and photo-switchable probes	Mammary window	Dendra-2	Kedrin <i>et al.</i> (2008); Gligorijevic <i>et al.</i> (2009)
Long-term imaging liver metastasis through abdominal window	Liver	SHG, Dendra-2, EGFP	Ritsma <i>et al.</i> (2012a,b)
Macrophages during intravasation in mammary tumors	Xenograft	EGFP, SHG, dextrans	Wang <i>et al.</i> (2007); Wyckoff <i>et al.</i> (2007)
Melanoma collective migration	Dorsal skin chamber	SHG, THG, EGFP, Dextran	Weigelin <i>et al.</i> (2012)
Hematopoietic stem cells and blood vessel	Skullcup	Dextran	Lo Celso <i>et al.</i> (2009)
Epithelial stem cells during hair regeneration	Skin	H2B-GFP mice	Rompolas <i>et al.</i> (2012)

blood flow patterns in the vasculature bed of pancreatic islets (Nyman *et al.*, 2008), to measure the vascular flow in the olfactory bulb (Chaigneau *et al.*, 2003; Zhang and Murphy,

2007), to study angiogenesis during wound healing (Holstein *et al.*, 2011), to unravel the signaling components mediating vasculature leakage (Gavard *et al.*, 2009; Figure 3(b)), and

Figure 3 Imaging tissue architecture and tissue response in live animals. (a) Diagram of systemic injections of fluorescent dyes in the rodent tail vein, and set up for IVM. (b) Dynamics of the vasculature in kidney, salivary glands, and skin revealed by systemic injections of fluorescent dextran. Fluorescently labeled dextran (red) were injected in the tail vein of anesthetized rats and the appearance in the vasculature was imaged by 2P microscopy. In the kidney (left panels) the proximal tubules (PT) and the distal (DT) tubules were revealed by endogenous fluorescence (green, excitation 740 nm). Dextran appears first in the microvasculature (arrows), then in the PT (arrowheads), and finally in the DT (asterisks) (Movie 1). In the salivary glands (center panels) the nuclei were revealed by injections of the nuclear dye Hoechst (blue) (Movie 2). In the skin (right panels), interleukin 8 (IL8) was injected locally to induce vasculature leakage, as shown by the diffusion of the dextran into the tissue (Movie 3). (c) CFDA was injected systemically and uptaken by the hepatocytes in the liver. After its processing and release into the bile canaliculi CFDA becomes fluorescent (green) and labels the bile canaliculi (arrowheads) providing a tool to analyze bile flux (Movie 4). (d) Change in morphology of the blood vessels in tumors *in situ*. The vasculature of an immunocompromised mouse was highlighted by the systemic injection of dextran (green) before (left panel) and after (center and right panel) the implant of breast cancer cells in the back (red). Note the change in shape and tortuosity of the blood vessels. Images were acquired by 2P microscopy (excitation wavelength 930 nm). Adapted from Gavard, J., Hou, X., Qu, Y., *et al.*, 2009. A role for a CXCR2/phosphatidylinositol 3-kinase gamma signaling axis in acute and chronic vascular permeability. Molecular and Cellular Biology 29, 2469–2480; Weigert, R., Sramkova, M., Parente, L., Amornphimoltham, P., Masedunskas, A., 2010. Intravital microscopy: A novel tool to study cell biology in living animals. Histochemistry and Cell Biology 133, 481–491; and Weigert, R., Porat-Shliom, N., Amornphimoltham, P., 2013. Imaging cell biology in live animals: Ready for prime time. Journal of Cell Biology 201, 969–979.

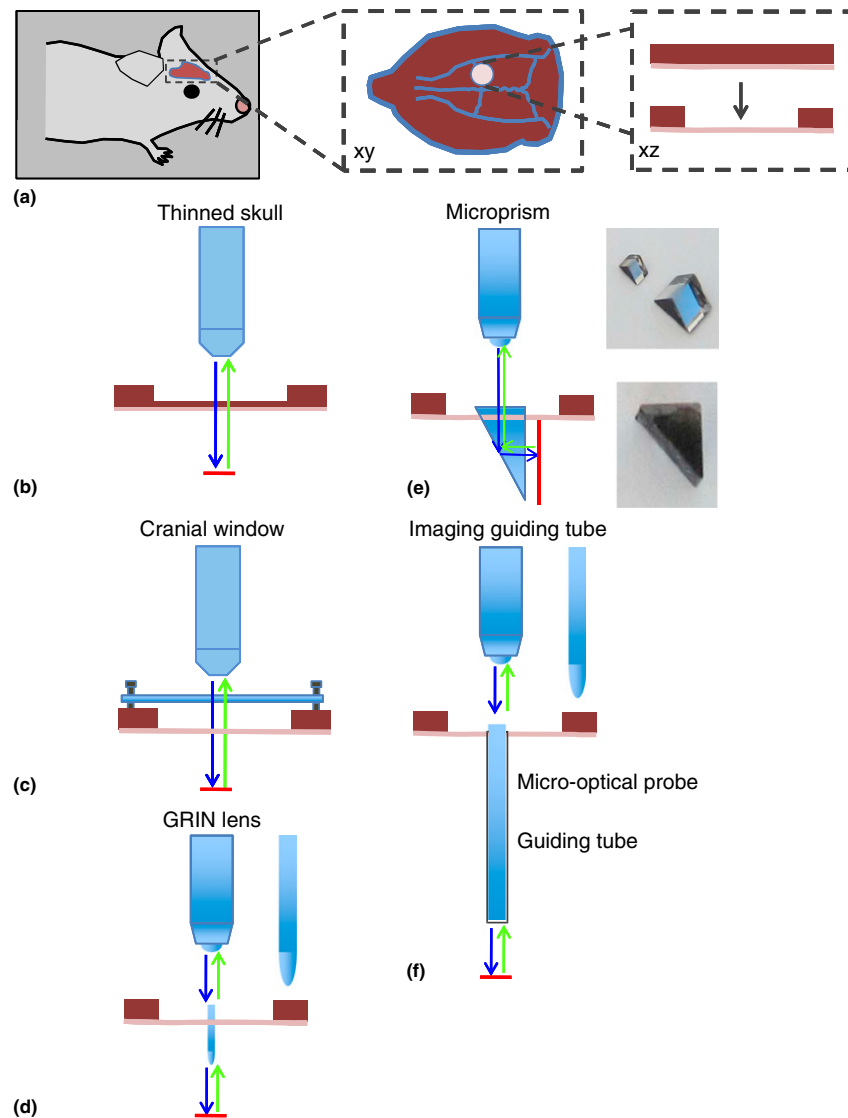


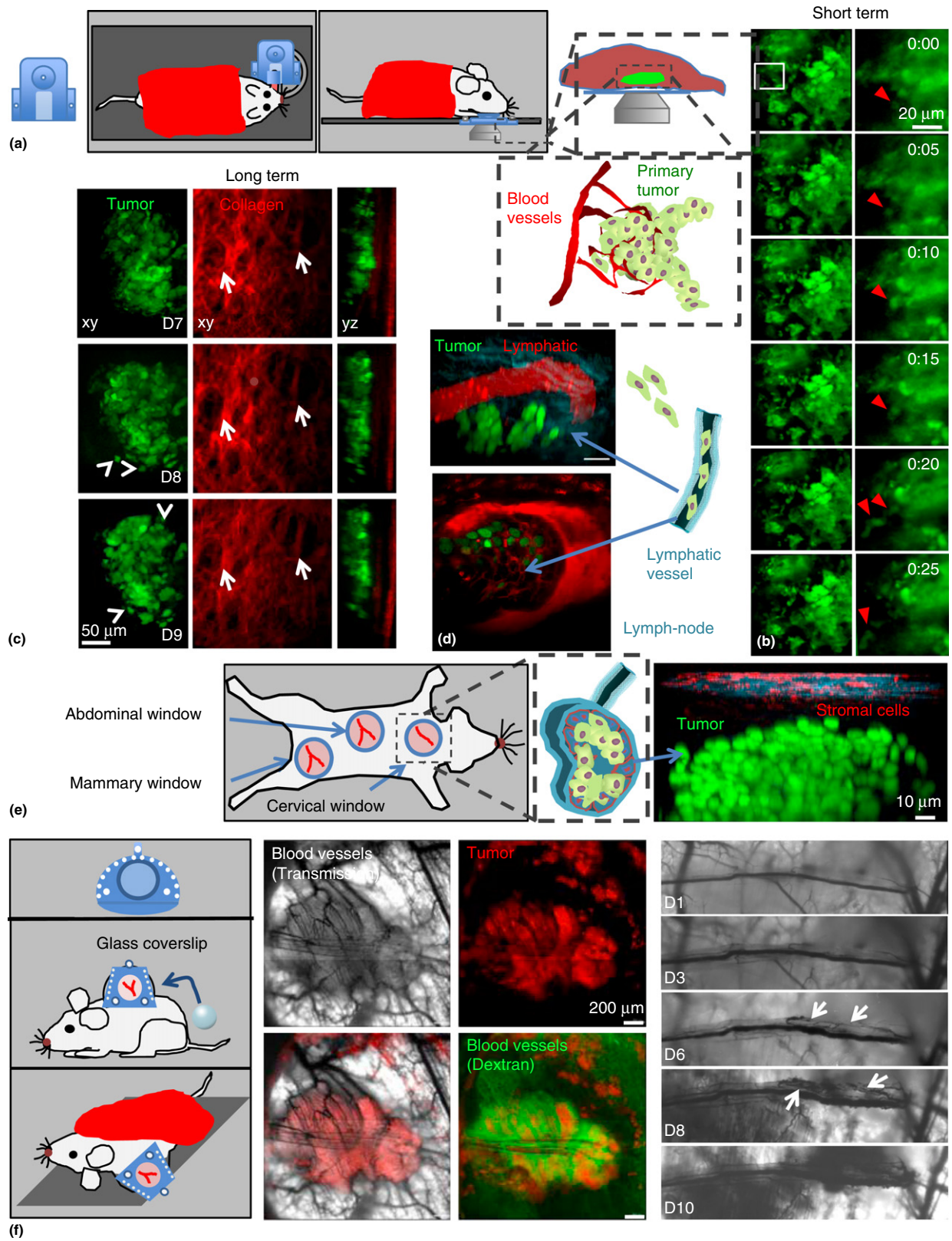
Figure 4 Approaches to perform long-term imaging in the brain – (a) Imaging the brain requires the surgical removal or the thinning of a small area of the skull. With the thinned skull chronic imaging is achieved by using a microscope lens (b). After removal of the skull and once the cortex is exposed several approaches can be used to image the area of interest such as, direct imaging, installation of a permanent cranial window (c), and the insertion of gradient Index (GRIN) lenses (d) or microprisms (e). Recently, permanently implanted optical guides have enabled imaging deep regions of the brain for periods up to 1 year (f).

recently to investigate the microvasculature in the lung (Presson *et al.*, 2011), and in the heart (Lee *et al.*, 2012; Vinegoni *et al.*, 2012).

Furthermore, the combination of dextran injections with procedures to perform long-term IVM has been instrumental in cancer biology to study tumor-mediated angiogenesis. In various murine tumor models, studies based on detailed three-dimensional analysis of the vascular morphology have clearly shown the dramatic changes in the structure and function of the blood vessels including their aberrant branching and increased permeability (Fukumura *et al.*, 2010; Smith *et al.*, 2008; Tozer *et al.*, 2005; Vakoc *et al.*, 2009; Figure 3(d)).

Specific tissue components can also be labeled by the selective delivery of small molecules to the organ of interest. For example, dextrans have been injected into the salivary duct of

live rats and mice to highlight the apical pole of the glandular epithelium and to visualize G protein-coupled receptor-stimulated fluid secretion (Masedunskas *et al.*, 2011a,b; Sramkova *et al.*, 2009). Other dyes, such as sodium fluorescein, carboxyfluorescein (CFDA) or Rhodamine 123, have been injected in live rats to visualize and to measure the dynamics of hepatic transport into the bile canaliculi (Babbey *et al.*, 2012; Liu *et al.*, 2012; Figure 3(c)). Curcumin and methoxy-04 have been used to label and monitor amyloid plaques in a mouse model for Alzheimer's disease (Garcia-Alloza *et al.*, 2007; Spires *et al.*, 2005). In addition, metabolic processes have been imaged *in vivo*, as well. For example, tissue energetics has been measured during ischemia and reperfusion in liver transplant using specific dyes such as, TMRM and Rhodamine 123, that detect changes in the mitochondrial



membrane potential (Theruvath *et al.*, 2008; Zhong *et al.*, 2008), or by measuring oxygen consumption using Ru (phen₃)²⁺ (Paxian *et al.*, 2004).

Imaging the Behavior of Individual Cells *In Vivo*

Imaging individual cells in live animals has revolutionized several fields such as neurobiology, immunology, cancer biology, and stem cell research (Table 2).

The neurosciences, in particular, have been the first discipline developing long-term *in vivo* imaging thanks to the establishment of surgical procedures to expose the brain cortex, and the implantation of chronic port of observations such as, cranial windows, imaging guide tubes for micro-optical probes, or micropiprims (Figure 4; Barretto *et al.*, 2011; Chia and Levene, 2009; Levene *et al.*, 2004; Svoboda and Yasuda, 2006; Xu *et al.*, 2007). Moreover, this field has tremendously benefited from the development of several transgenic mouse models expressing fluorescent reporters in specific neuronal populations (Barretto *et al.*, 2011; Chia and Levene, 2009; Levene *et al.*, 2004; Svoboda and Yasuda, 2006; Xu *et al.*, 2007). In particular, the so called 'brainbow' mouse, that is based on the random expression of various combinations of fluorescent variants of the green fluorescent protein (GFP), has opened the door to investigating neuronal interconnections *in vivo* (Livet *et al.*, 2007; Loulier *et al.*, 2014). This extensive toolbox has been used to investigate changes in neuronal morphology and function both during normal and pathological conditions such as, stroke (Zhang and Murphy, 2007), tumors (Barretto *et al.*, 2011), neurodegenerative diseases (Merlini *et al.*, 2012), and infections (McGavern and Kang, 2011).

In cancer biology, IVM has been instrumental in understanding the mechanisms of tumor progressions and the modality of invasion and metastasis of cancer cells within the tumor microenvironment (Beerling *et al.*, 2011). This has led to the discovery of novel mechanisms that were not appreciated in *in vitro* systems. For example, in mammary tumors, the intravasation of metastatic cells has been shown to require macrophages (Wang *et al.*, 2007; Wyckoff *et al.*, 2006, 2007). In highly invasive melanomas, the migratory ability of cells has been correlated with their differentiation state, as

determined by the expression of a reporter for melanin expression (Pinner *et al.*, 2009). In a recent study, melanoma cells have been shown to migrate using a combination of different modalities (single cell and collective migration) using various preformed tracks such as myelin and collagen fibers (Weigelin *et al.*, 2012). Also in this field, the ability to extend the length of the observations from a few hours to several days have provided a plethora of novel information about the relationship between metastatic behavior and the local micro-environment, particularly the vasculature and the lymphatic system. For example, squamous cell carcinomas in the oral cavity are an ideal system to perform chronic studies on invasion since they are accessible without the need for surgical procedures (Amornphimoltham *et al.*, 2013; Patel *et al.*, 2011; Movies 5 and 6; Figures 5(a–d)). On the other hand, for other less accessible tumors various optical windows installed in selected areas of the body such as, the back (Alexander *et al.*, 2008), the mammary glands (Kedrin *et al.*, 2008), and the liver (Ritsma *et al.*, 2012b, 2013) have been successfully established (Figures 5(e) and 5(f)).

The immune response is based on a complex array of very dynamic and specialized cells, which are quickly deployed to their site of action and often exhibit very complex behaviors. IVM has been a tremendous aid in unraveling novel qualitative and quantitative aspects of cellular immunity (Movies 7–9; Figures 6(a) and 6(b); Cahalan and Parker, 2008; Germain *et al.*, 2005; Nitschke *et al.*, 2008). For example, transferring specific fluorescently labeled populations of immune cells into recipient animals has allowed the investigation of the dynamics and interactions of B lymphocytes and T cells in lymphoid tissues (Qi *et al.*, 2006), to study T cells activation (Hickman *et al.*, 2008), migration of dendritic cells (Nitschke *et al.*, 2012), and in addition, to investigate the dynamics of the immune system outside lymphoid tissues such as, in the brain during pathogen infections (Nayak *et al.*, 2012), in the heart during inflammation (Li *et al.*, 2012), and in solid tumors (Deguine *et al.*, 2010). Moreover, IVM microscopy has been instrumental to study the biology of pathogen infection (Hickman *et al.*, 2009). In the last decade, early and late stages of the infective process have been characterized *in vivo* by using various approaches and focusing on different pathogens such as *Staphylococcus aureus* and *Borrelia burgdorferi*

Figure 5 Tumor progression and invasive behavior in live animals – (a–d) Intravital imaging of tumor progression in the mice tongue injected with fluorescently labeled metastatic head and neck cancer cell lines. (a) Schematic drawings of the tongue holding device, animal set up for intravital imaging, and primary tumor mass growing in the tongue, as described in Amornphimoltham *et al.*, 2013. (b) Short-term dynamics of tumor cells. The edge of the tumor was imaged by 2P microscopy for 2 h at 5 min intervals. Individual cells were observed leaving the tumor mass (insets, red arrowheads) (Movie 5). (c) Long-term dynamics of tumor cells. The tumor was imaged on a daily basis by using 2P microscopy to reveal collagen fibers (red) and the tumor mass (green). Z-stacks were acquired at day 7 (D7), 8 (D8), and 9 (D9) and three-dimensional reconstructions (right panels, yz view) and maximal projections of the xy view (center and left panels) are shown. The extracellular matrix at the surface of the tongue was used as a reference point (center panels, arrows). Individual tumor cells were observed migrating from the tumor mass (arrowheads). (d) Tumor cells inside lymphatic vessels. Dextran was injected in the tongue in order to map the lymphatic vessels, and Z-scans were performed. Maximal projections show tumor cells (green) either in proximity (upper panel, Movie 6) or inside lymphatic vessels (lower panel). Collagen fibers in cyan. (e) Diagram illustrating some of the optical windows used for long-term imaging of tumors. Right panel shows a lymph node invaded by metastatic cells (green) and visualized via a cervical chamber. The edge of the lymph node is revealed by highlighting stromal cells (red) and collagen fibers (cyan). (f) Dorsal skin chamber to image tumor progression and tumor-mediated angiogenesis (left panels). Colon cancer cells (red) were inoculated and after 7 days the overall view of tumor mass and blood vessels (green) observed at a low magnification using a combination of transmitted light and confocal microscope (center panel). Change in the morphology of the blood vessels from day1 to day10 (D1–D10) after tumor injection. Note the formation of new vessels and the branching out from an existing one (arrows). Adapted from Amornphimoltham, P., Rechache, K., Thompson, J., *et al.*, 2013. Rab25 regulates invasion and metastasis in head and neck cancer. Clinical Cancer Research: An Official Journal of the American Association for Cancer Research 19, 1375–1388.

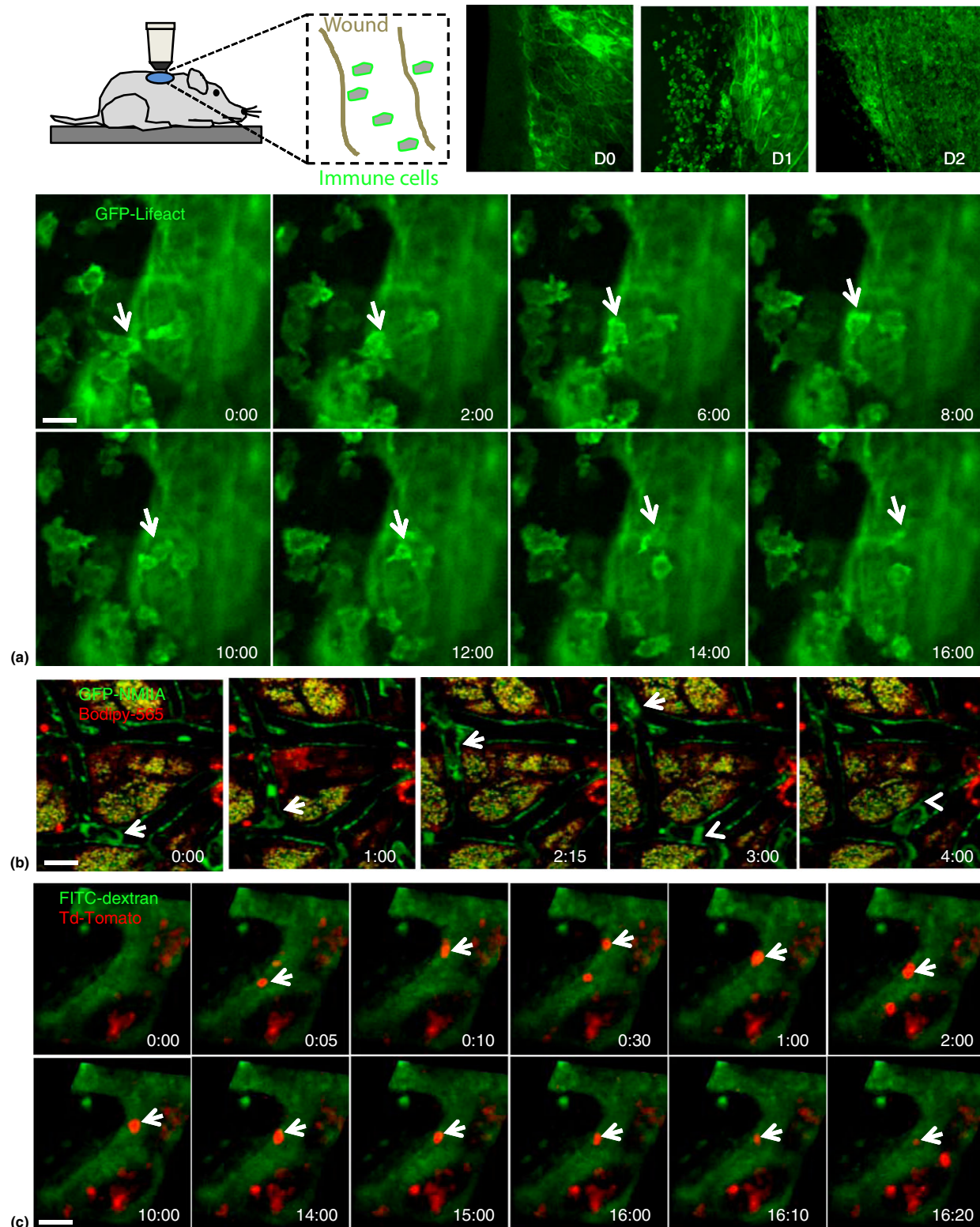


Figure 6 Imaging individual cells in live animals – (a–b) Dynamics of immune cells. (a) A wound was generated in the skin of a transgenic mouse expressing the actin reporter lifeact. After 1 day (D1), the dynamics of the immune cells recruited to the wound site was imaged by 2P microscopy (Movie 7). (b) A granulocyte moving inside a blood vessel in the mammary gland of a mouse expressing GFP-tagged myosin IIb (green) and labeled with mitotracker (red) was imaged in time-lapse mode by using confocal microscopy (Movie 8). (c) Dynamics of hematopoietic stem cells isolated from a transgenic mouse expressing the mTomato membrane marker (Masedunskas *et al.*, 2011a) and injected systemically in a recipient mouse. After 1 day the calvarial bone was exposed and the dynamics of the stem cells close to a large blood vessel (highlighted by injection of fluorescent dextran, green) were imaged by 2P microscopy. One of the stem cell surfs along the blood vessel and finally intravasates (Movie 9).

(Laschke *et al.*, 2005; Norman *et al.*, 2008) or *Leishmania major* (Filipe-Santos *et al.*, 2009). More recently, the advancement of the GFP technology has facilitated the generation of genetically engineered pathogens, thus increasing the number of tools to study infectious disease under physiological conditions. Two noticeable examples are in the kidney where the proliferation of a GFP-expressing uropathogenic *Escherichia coli* was studied in detail (Mansson *et al.*, 2007) and the role of resident macrophages in influencing the survival of *Candida albicans* was described for the first time (Lionakis *et al.*, 2013).

Finally, a few experimental approaches have been recently developed to track individual stem cells and to study their behavior in living mice (Figure 6(c)). A model for long-term imaging has been established in the calvarium bone marrow to examine the relationship between hematopoietic stem cells and blood vessel (Lo Celso *et al.*, 2009). Moreover, by using novel transgenic mice expressing specific reporters new paradigms have been recently established on stem cell maintenance either in the skin (Rompolas *et al.*, 2012, 2013) or in the intestinal crypts (Ritsma *et al.*, 2014).

Imaging Subcellular Structures *In Vivo*

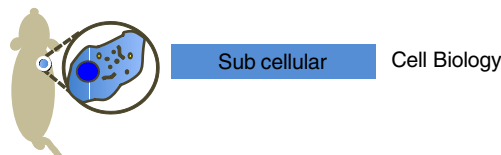
The main challenge in performing subcellular imaging in live animals is the minimization of the motion artifacts derived from the heartbeat and the respiration, while maintaining the integrity and the physiology of the tissue. Indeed, small shifts along the three axis make practically impossible to visualize structures whose sizes are in the micron or submicron range, whereas it marginally affects larger structures. Recently, this issue has been overcome by using a combination of strategies, which include: (1) the development of surgical procedures to

expose and to proper position the organ of interest (Masedunskas *et al.*, 2013), (2) the use of specific organ holders (Cao *et al.*, 2012; Masedunskas *et al.*, 2012a), (3) the synchronization of the imaging acquisition with the heartbeat and respiration (Li *et al.*, 2012; Presson *et al.*, 2011), and (4) novel software to correct for image distortion (Dunn *et al.*, 2014).

Nuclei have been easily imaged, thus making possible to study processes, such as cell division and apoptosis. For example, mitosis and the structure of the mitotic spindle were followed in a xenograft model of human cancer upon treatment with the anticancer drug Paclitaxel, revealing significant differences between *in vitro* and *in vivo* models (Orth *et al.*, 2011); cell division was followed in the hair-follicle stem cells of transgenic mice thus determining that epithelial–mesenchymal interactions are essential for stem cell activation and regeneration (Rompolas *et al.*, 2012); finally, apoptosis has been studied in the liver of live mice injected with the FAS ligand Jo2 by using confocal microendoscopy (Goetz *et al.*, 2011).

Another field where subcellular IVM has been utilized is membrane trafficking (Table 3). The first trafficking event successfully imaged *in vivo* is the endocytosis of various markers, such as fluorescently labeled dextrans, folate, or the aminoglycoside gentamicin, in the kidney proximal tubuli (Movies 10–12; Figure 7; Dunn *et al.*, 2002; Sandoval *et al.*, 2004). However, in these pioneering studies the residual motion artifacts limited the imaging to short periods of time. A major breakthrough was accomplished in rodents' salivary glands where various probes (dextrans, BSA, and transferrin) were observed to rapidly internalize in the stromal cells surrounding the salivary gland epithelium in an actin cytoskeleton-dependent fashion and traffic throughout the endo-lysosomal system (Masedunskas *et al.*, 2012b;

Table 3 Subcellular processes imaged by intravital microscopy



Event	Organ	Probe	References
Nuclear dynamics	Tumor	GFP-H2B	Orth <i>et al.</i> (2011)
	Skin	GFP-H2P	Rompolas <i>et al.</i> (2012)
Endocytosis	Kidney	Dextran	Dunn <i>et al.</i> (2002)
		Transferrin	Sandoval <i>et al.</i> (2004)
		SiRNA	Ohno <i>et al.</i> (2005)
	Salivary glands	Dextran	Masedunskas and Weigert (2008)
		Transferrin	Masedunskas <i>et al.</i> (2012a,b)
		Plasmid DNA	Sramkova <i>et al.</i> (2009, 2012)
	Immune cells		Chen <i>et al.</i> (2012)
	Tumor	EGF-nanotubes	Bhirde <i>et al.</i> (2009)
		Anti-EGFR antibodies	Amorphimoltham <i>et al.</i> (2011); Thurber and Weissleder (2011); Thurber <i>et al.</i> (2013)
Exocytosis	Kidney	Acridine orange	Toma <i>et al.</i> (2006)
	Salivary Glands	GFP	Masedunskas <i>et al.</i> (2011a); Masedunskas <i>et al.</i> (2012a,b)
		mTomato	
		GFP-lifect	
	Skeletal muscle	GFP-Glut4	Lauritzen <i>et al.</i> (2010)
Autophagy	Liver	GFP-LC3	Cao <i>et al.</i> (2012); Dupont <i>et al.</i> (2014)

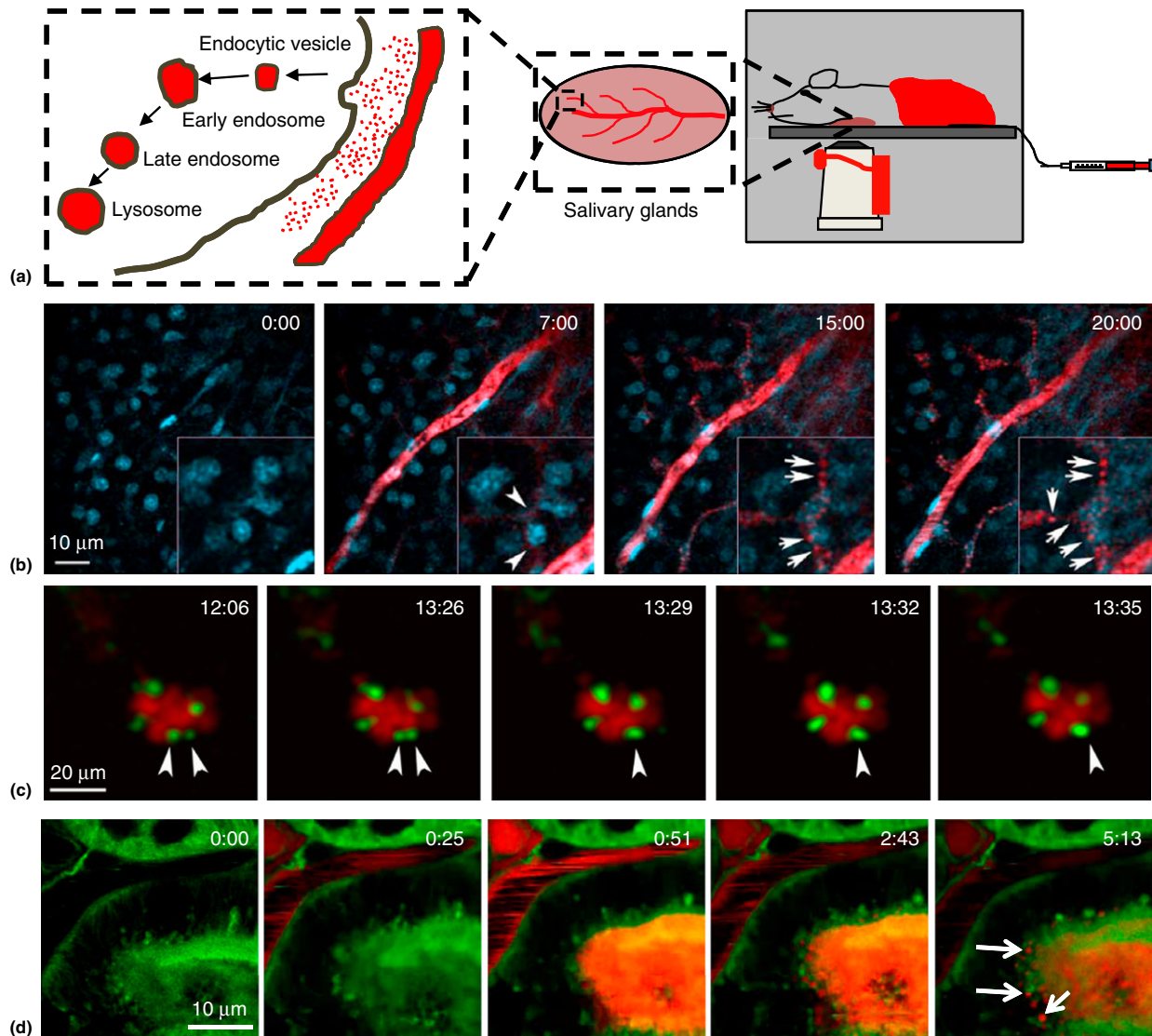


Figure 7 IVM of subcellular organelles (Endocytosis). (a) Set up for studying endocytosis *in vivo* in salivary glands. Fluorescently labeled molecules are injected into the tail vein. The probes reach the salivary glands, are released into the tissue through the fenestrated capillaries, and get internalized into the stromal cells. (b) Time-lapse imaging of dextran internalization in live animals – A rat was injected with Hoechst 33342, and followed by dextran injection into the tail artery, which highlights the vasculature almost instantly. A few minutes after the injection, small endocytic vesicles appeared inside cells adjacent to a blood vessel, which increased over time both in number and in size (Movie 10). (c) Dynamics of early endosomal fusion. Small endocytic vesicles (green) fuse to each other to form larger structures (arrowheads). Lysosomes are highlighted in red (arrowheads) (Movie 11). (d) Endocytosis of a systemically injected dextran into the kidney of a transgenic mouse expressing the membrane marker m-GFP. The dextran (red) is transported from the microvasculature into the proximal tubules, and then internalized in small endocytic vesicles (arrows) (Movie 12).

Masedunskas and Weigert, 2008; Figure 7). Interestingly, these studies revealed significant differences in the kinetics of internalization of transferrin and dextran when compared to cell cultures, indicating that the environment *in vivo* plays a major role in the regulation of membrane trafficking (Masedunskas et al., 2012b; Weigert, 2014). Endocytosis has also been studied in cancer models. For example, in head and neck cancer xenografts the uptake of a fluorescent epidermal growth factor (EGF) conjugated to carbon nanotubes has been characterized underscoring a potential role for this pathway as a therapeutic target (Bhirde et al., 2009). Moreover, the recycling of membranes controlled by the small GTPase Rab25 has been

shown to regulate the ability of tumor cells to invade and metastasize to cervical lymph nodes *in vivo* (Amornphimoltham et al., 2013).

The ability to almost completely minimize the motion artifacts has enabled performing rigorous quantitative analysis of membrane trafficking processes, such as regulated exocytosis (Masedunskas et al., 2011a). Indeed, in salivary glands, the use of transgenic mice expressing selected fluorescently tagged probes has revealed that the regulation and the modality of exocytosis differ between *in vivo* and *in vitro* systems (Movies 13 and 14; Figures 8(a–c)). In addition, it is important to emphasize that the ability to perform transient

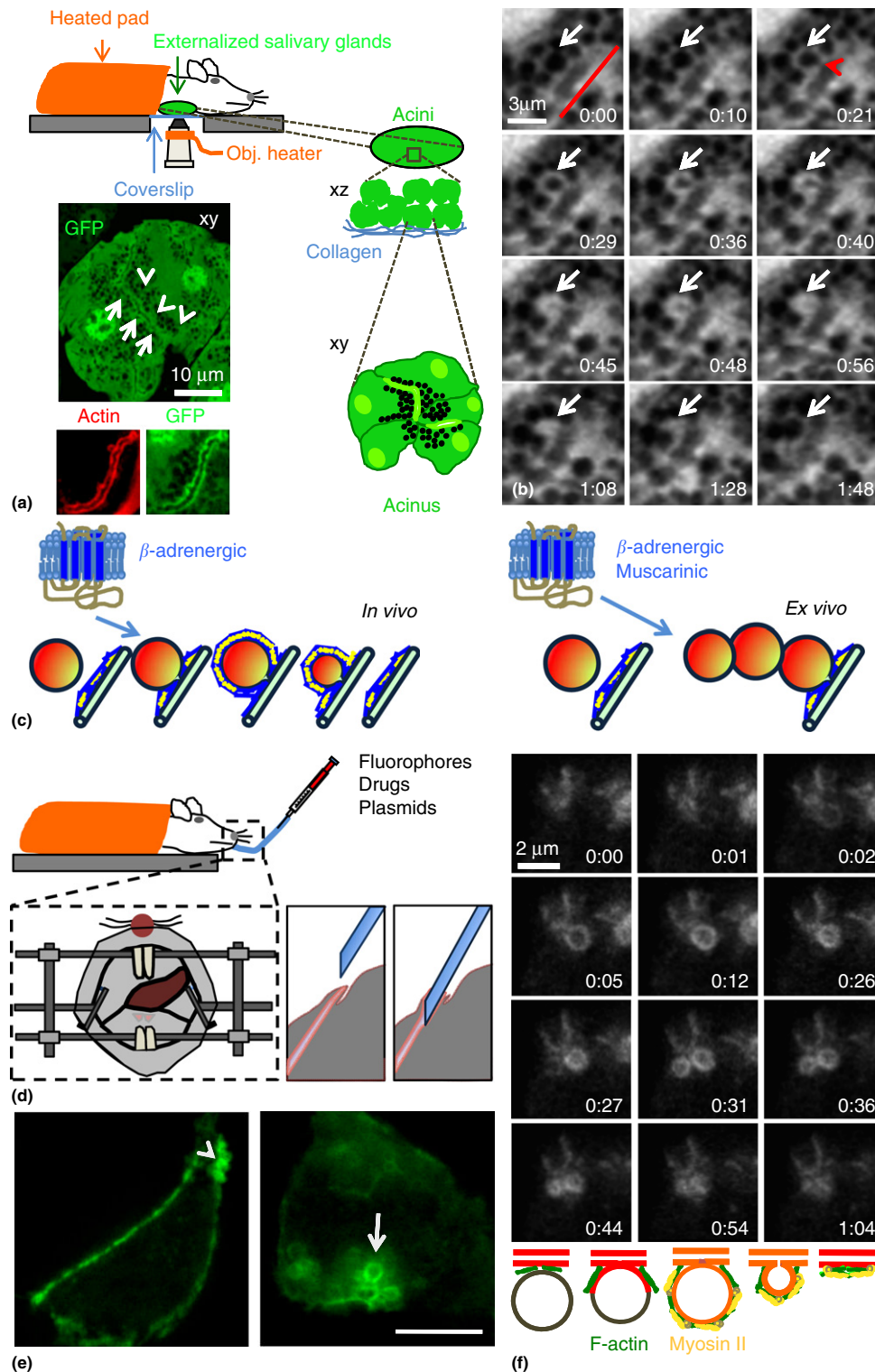


Figure 8 IVM of subcellular organelles (Exocytosis). (a) Diagram of the experimental set up to image regulated exocytosis in the salivary glands of transgenic mice expressing cytoplasmic GFP (green) that is excluded from the secretory granules (arrowheads) and it is enriched at the apical plasma membrane where exocytosis occurs (arrow). Note that the apical plasma membrane is enriched in GFP and actin as revealed by counter-staining with phalloidin (red) (Insets). (b) Time series of a single granule (white arrows) fusing with the APM (red line). A continuity between the lumen of the granule and the acinar canaliculus is visible (red arrowhead, [Movie 13](#)). (c) Comparison between regulated exocytosis *in vivo* and *ex vivo* in salivary glands. *In vivo*, regulated exocytosis is stimulated by β -adrenergic stimulation and involves the integration of the granular membrane into the apical plasma membrane (gradual collapse). In *ex vivo* preparations regulated exocytosis is elicited by either the β -adrenergic or the muscarinic receptors which trigger compound exocytosis. (d) Diagram of the cannulation of the salivary duct to deliver fluorescent molecules and plasmid DNA to perform transient *in vivo* transfections. (e) Single acinar cells showing GFP-lifect enriched at the apical plasma membrane in resting conditions (left panel, arrowhead) and recruited onto the secretory granules upon stimulation (right panel, arrow). (f) Time-lapse sequence of the recruitment of GFP-lifect onto the secretory granules (arrows) ([Movie 14](#)). Diagram illustrating the recruitment of F-actin and Myosin II onto the secretory granules. Adapted from Masedunskas, A., Sramkova, M., Parente, L., *et al.*, 2011a. Role for the actomyosin complex in regulated exocytosis revealed by intravital microscopy. *Proceedings of the National Academy of Sciences of the United States of America* 108, 13552–13557.

transfections *in vivo* (Sramkova *et al.*, 2009; Figure 8(d)) combined with a pharmacological approach has enabled the identification of specific components regulating this process. Specifically, these studies revealed the requirement for the assembly onto the secretory granules of a complex formed by F-actin and two isoforms of the motor protein myosin II, which are required to facilitate the completion of the exocytic process (Figures 8(e) and 8(f); Masedunskas *et al.*, 2011a,b, 2012c,d). These results underscore that the dynamics of the assembly of the actin cytoskeleton can be studied both qualitatively and quantitatively in live animals. In addition, this approach has shown that some of the mechanisms that contribute to regulate the apical plasma membrane homeostasis *in vivo* cannot be recapitulated in *in vitro* model systems (Masedunskas *et al.*, 2011b, 2012c). Similar approaches have begun to be used to investigate other models for protein secretion, such as the exocytosis of vesicles transporter type 4 (Glut4) and its insulin-dependent translocation to the plasma membrane (Lauritzen *et al.*, 2008, 2010). This represents a very powerful experimental model that bridges physiology and cell biology and has the potential to provide fundamental information on metabolic diseases.

The use of IVM to image intracellular organelles is rapidly expanding to other areas of cell biology. For example, the development of a mouse expressing the autophagic marker LC3-GFP has enabled imaging the dynamics of autophagy upon starvation in several organs such as liver, pancreas, and skeletal muscle (Cao *et al.*, 2012). This model system has been instrumental to reveal that lipid droplets in the liver contribute to the initiation of starvation-induced autophagy (Dupont *et al.*, 2014). Moreover, the dynamics of lipid droplets have been studied in the liver by using THG without the need for exogenous labeling (Debarre *et al.*, 2006) and the dynamics of mitochondria have been imaged in the salivary glands and in the liver by using various fluorescent probes (Masedunskas *et al.*, 2012a; Sun *et al.*, 2005; Weigert *et al.*, 2010). Finally, recent studies addressed the imaging of intracellular signaling in live animals using a variety of FRET-based probes specifically designed to measure the levels of intracellular calcium, or cAMP and the activity of proteins such as calpain (Ritsma *et al.*, 2012a; Rudolf *et al.*, 2006; Stockholm *et al.*, 2005).

Concluding Remarks

IVM is a powerful technique that has already begun to provide major contributions to several areas of biology and it is steadily broadening its range of application. The strength of this approach resides in the unique opportunity to study a given process in real time in the context of a fully developed tissue. Although *in vitro* and *ex vivo* model systems are still the most widely used tools to study biological processes at a molecular level, they are often inadequate to reconstitute the complexity of the signaling, metabolic cues, and structural organization that are the key features of multicellular organisms. In this respect, IVM is the obligatory choice to investigate tissue physiopathology, and the development of subcellular IVM has created new avenues to link tissue function to cellular and subcellular processes.

We foresee that the continuous development of novel tools specifically tailored to IVM, such as optics, imaging software, surgical procedures, fluorescent probes, and transgenic animal models will rapidly make IVM a 'standard' tool for studying biology.

Acknowledgments

This research was supported by the Intramural Research Program of the NIH, National Institute of Dental and Craniofacial Research.

References

- Alexander, S., Koehl, G.E., Hirschberg, M., Geissler, E.K., Friedl, P., 2008. Dynamic imaging of cancer growth and invasion: A modified skin-fold chamber model. *Histochemistry and Cell Biology* 130, 1147–1154.
- Alt, C., Runnels, J.M., Teo, G.S.L., Lin, C.P., 2012. In vivo tracking of hematopoietic cells in the retina of chimeric mice with a scanning laser ophthalmoscope. *Intravital* 1 (2), 132–140.
- Amornphimoltham, P., Masedunskas, A., Weigert, R., 2011. Intravital microscopy as a tool to study drug delivery in preclinical studies. *Advanced Drug Delivery Reviews* 63, 119–128.
- Amornphimoltham, P., Rechache, K., Thompson, J., *et al.*, 2013. Rab25 regulates invasion and metastasis in head and neck cancer. *Clinical Cancer Research: an Official Journal of the American Association for Cancer Research* 19, 1375–1388.
- von Andrian, U.H., 1996. Intravital microscopy of the peripheral lymph node microcirculation in mice. *Microcirculation* 3, 287–300.
- Babbey, C.M., Ryan, J.C., Gill, E.M., *et al.*, 2012. Quantitative intravital microscopy of hepatic transport. *Intravital* 1, 44–53.
- Bakker, G.J., Andresen, V., Hoffman, R.M., Friedl, P., 2012. Fluorescence lifetime microscopy of tumor cell invasion, drug delivery, and cytotoxicity. *Methods in Enzymology* 504, 109–125.
- Barretto, R.P., Ko, T.H., Jung, J.C., *et al.*, 2011. Time-lapse imaging of disease progression in deep brain areas using fluorescence microendoscopy. *Nature Medicine* 17, 223–228.
- Beerling, E., Ritsma, L., Vrisekoop, N., Derksen, P.W., van Rhee, J., 2011. Intravital microscopy: New insights into metastasis of tumors. *Journal of Cell Science* 124, 299–310.
- Bhirde, A.A., Patel, V., Gavard, J., *et al.*, 2009. Targeted killing of cancer cells *in vivo* and *in vitro* with EGF-directed carbon nanotube-based drug delivery. *ACS Nano* 3, 307–316.
- Brown, E., McKee, T., diTomaso, E., *et al.*, 2003. Dynamic imaging of collagen and its modulation in tumors *in vivo* using second-harmonic generation. *Nature Medicine* 9, 796–800.
- Cahalan, M.D., Parker, I., 2008. Choreography of cell motility and interaction dynamics imaged by two-photon microscopy in lymphoid organs. *Annual Review of Immunology* 26, 585–626.
- Camirand, G., Li, Q., Demetris, A.J., *et al.*, 2011. Multiphoton intravital microscopy of the transplanted mouse kidney. *American Journal of Transplantation: Official Journal of the American Society of Transplantation and the American Society of Transplant Surgeons* 11, 2067–2074.
- Campagnola, P.J., Loew, L.M., 2003. Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms. *Nature Biotechnology* 21, 1356–1360.
- Cao, L., Kobayakawa, S., Yoshiki, A., Abe, K., 2012. High resolution intravital imaging of subcellular structures of mouse abdominal organs using a microstage device. *PLOS One* 7, e33876.
- Chaigneau, E., Oheim, M., Audinat, E., Charpak, S., 2003. Two-photon imaging of capillary blood flow in olfactory bulb glomeruli. *Proceedings of the National Academy of Sciences of the United States of America* 100, 13081–13086.
- Chen, K., Nishi, H., Travers, R., *et al.*, 2012. Endocytosis of soluble immune complexes leads to their clearance by FcγRIIb but induces neutrophil extracellular traps via FcγRIIA *in vivo*. *Blood* 120, 4421–4431.
- Chia, T.H., Levene, M.J., 2009. *In vivo* imaging of deep cortical layers using a microprism. *Journal of Visualized Experiments* (30). pii: 1509. doi:10.3791/1509.

- Clark, E.R., Clark, E.L., 1932. Observations on living preformed vessels as seen in the transparent chamber inserted into a rabbit's ear. *American Journal of Anatomy* 49, 441.
- Cox, G., Kable, E., Jones, A., *et al.*, 2003. 3-dimensional imaging of collagen using second harmonic generation. *Journal of Structural Biology* 141, 53–62.
- Debarre, D., Supatto, W., Pena, A.M., *et al.*, 2006. Imaging lipid bodies in cells and tissues using third-harmonic generation microscopy. *Nature Methods* 3, 47–53.
- Deguine, J., Breart, B., Lemaitre, F., Di Santo, J.P., Bousso, P., 2010. Intravital imaging reveals distinct dynamics for natural killer and CD8(+) T cells during tumor regression. *Immunity* 33, 632–644.
- Denk, W., Strickler, J.H., Arbo, W.W., 1990. Two-photon laser scanning fluorescence microscopy. *Science* 248, 73–76.
- Dunn, K.W., Lorenz, K.S., Salama, P., Delp, E.J., 2014. IMART software for correction of motion artifacts in images collected in intravital microscopy. *Intravital* 3 (1). doi:10.4161/intv.28210.
- Dunn, K.W., Sandoval, R.M., Kelly, K.J., *et al.*, 2002. Functional studies of the kidney of living animals using multicolor two-photon microscopy. *American Journal of Physiology Cell Physiology* 283, C905–C916.
- Dupont, N., Chauhan, S., Arko-Mensah, J., *et al.*, 2014. Neutral lipid stores and lipase PNPLA5 contribute to autophagosome biogenesis. *Current Biology* 24, 609–620.
- Filipe-Santos, O., Pescher, P., Breart, B., *et al.*, 2009. A dynamic map of antigen recognition by CD4 T cells at the site of Leishmania major infection. *Cell Host and Microbe* 6, 23–33.
- Fu, Y., Huff, T.B., Wang, H.W., Wang, H., Cheng, J.X., 2008. *Ex vivo* and *in vivo* imaging of myelin fibers in mouse brain by coherent anti-Stokes Raman scattering microscopy. *Optics Express* 16, 19396–19409.
- Fukumura, D., Duda, D.G., Munn, L.L., Jain, R.K., 2010. Tumor microvasculature and microenvironment: Novel insights through intravital imaging in pre-clinical models. *Microcirculation* 17, 206–225.
- García-Alloza, M., Borrelli, L.A., Rozkalne, A., Hyman, B.T., Bacskaï, B.J., 2007. Curcumin labels amyloid pathology *in vivo*, disrupts existing plaques, and partially restores distorted neurites in an Alzheimer mouse model. *Journal of Neurochemistry* 102, 1095–1104.
- Gavard, J., Hou, X., Qu, Y., *et al.*, 2009. A role for a CXCR2/phosphatidylinositol 3-kinase gamma signaling axis in acute and chronic vascular permeability. *Molecular and Cellular Biology* 29, 2469–2480.
- Germain, R.N., Castellino, F., Chieppa, M., *et al.*, 2005. An extended vision for dynamic high-resolution intravital immune imaging. *Seminars in Immunology* 17, 431–441.
- Glgorijevic, B., Kedrin, D., Segall, J.E., Condeelis, J., van Rheenen, J., 2009. Dendra2 photoswitching through the Mammary Imaging Window. *Journal of Visualized Experiments* (28), pii: 1278.
- Goetz, M., Ansems, J.V., Galle, P.R., Schuchmann, M., Kiesslich, R., 2011. *In vivo* real-time imaging of the liver with confocal endomicroscopy permits visualization of the temporospatial patterns of hepatocyte apoptosis. *American Journal of Physiology Gastrointestinal and Liver Physiology* 301, G764–G772.
- Göppert-Mayer, M., 1931. Über elementarakte mit zwei quantensprungen. *Annalen der Physik (Leipzig)* 5, 273–294.
- Guan, Y., Worrell, R.T., Pritts, T.A., Montrose, M.H., 2009. Intestinal ischemia-reperfusion injury: Reversible and irreversible damage imaged *in vivo*. *American Journal of Physiology-Gastrointestinal and Liver Physiology* 297, G187–G196.
- Hall, A.M., Rhodes, G.J., Sandoval, R.M., Corridon, P.R., Molitoris, B.A., 2013. *In vivo* multiphoton imaging of mitochondrial structure and function during acute kidney injury. *Kidney International* 83, 72–83.
- Helmchen, F., Denk, W., 2005. Deep tissue two-photon microscopy. *Nature Methods* 2, 932–940.
- Helmchen, F., Kleinfeld, D., 2008. Chapter 10. *In vivo* measurements of blood flow and glial cell function with two-photon laser-scanning microscopy. *Methods in Enzymology* 444, 231–254.
- Hickman, H.D., Bennis, J.R., Yewdell, J.W., 2009. Caught in the act: Intravital multiphoton microscopy of host-pathogen interactions. *Cell Host and Microbe* 5, 13–21.
- Hickman, H.D., Takeda, K., Skon, C.N., *et al.*, 2008. Direct priming of antiviral CD8+ T cells in the peripheral interfollicular region of lymph nodes. *Nature Immunology* 9, 155–165.
- Holstein, J.H., Becker, S.C., Fiedler, M., *et al.*, 2011. Intravital microscopic studies of angiogenesis during bone defect healing in mice calvaria. *Injury* 42, 765–771.
- Irwin, J.W., Macdonald 3rd, J., 1953. Microscopic observations of the intrahepatic circulation of living guinea pigs. *Anatomical Record* 117, 1–15.
- Jester, J.V., Andrews, P.M., Petroll, W.M., Lemp, M.A., Cavanagh, H.D., 1991. *In vivo*, real-time confocal imaging. *Journal of Electron Microscopy Technique* 18, 50–60.
- Kang, J.J., Toma, I., Sipos, A., McCulloch, F., Peti-Peterdi, J., 2006. Quantitative imaging of basic functions in renal (patho)physiology. *American Journal of Physiology Renal Physiology* 291, F495–F502.
- Kasischke, K.A., Vishwasrao, H.D., Fisher, P.J., Zipfel, W.R., Webb, W.W., 2004. Neural activity triggers neuronal oxidative metabolism followed by astrocytic glycolysis. *Science* 305, 99–103.
- Kedrin, D., Glgorijevic, B., Wyckoff, J., *et al.*, 2008. Intravital imaging of metastatic behavior through a mammary imaging window. *Nature Methods* 5, 1019–1021.
- Laschke, M.W., Kerdudou, S., Herrmann, M., Menger, M.D., 2005. Intravital fluorescence microscopy: A novel tool for the study of the interaction of Staphylococcus aureus with the microvascular endothelium *in vivo*. *Journal of Infectious Diseases* 191, 435–443.
- Lauritzen, H.P., Galbo, H., Brandauer, J., Goodyear, L.J., Ploug, T., 2008. Large GLUT4 vesicles are stationary while locally and reversibly depleted during transient insulin stimulation of skeletal muscle of living mice: Imaging analysis of GLUT4-enhanced green fluorescent protein vesicle dynamics. *Diabetes* 57, 315–324.
- Lauritzen, H.P., Galbo, H., Toyoda, T., Goodyear, L.J., 2010. Kinetics of contraction-induced GLUT4 translocation in skeletal muscle fibers from living mice. *Diabetes* 59, 2134–2144.
- Le, T.T., Yue, S., Cheng, J.X., 2010. Shedding new light on lipid biology with CARS microscopy. *Journal of Lipid Research* 51 (11), 3091–3102.
- Lee, S., Vinegoni, C., Feruglio, P.F., *et al.*, 2012. Real-time *in vivo* imaging of the beating mouse heart at microscopic resolution. *Nature Communications* 3, 1054.
- Levene, M.J., Dombeck, D.A., Kasischke, K.A., Molloy, R.P., Webb, W.W., 2004. *In vivo* multiphoton microscopy of deep brain tissue. *Journal of Neurophysiology* 91, 1908–1912.
- Levitt, J.A., Matthews, D.R., Ameer-Beg, S.M., Suhling, K., 2009. Fluorescence lifetime and polarization-resolved imaging in cell biology. *Current Opinion in Biotechnology* 20, 28–36.
- Li, W., Nava, R.G., Bribiesco, A.C., *et al.*, 2012. Intravital 2-photon imaging of leukocyte trafficking in beating heart. *Journal of Clinical Investigation* 122, 2499–2508.
- Lionakis, M.S., Swamydas, M., Fischer, B.G., *et al.*, 2013. CX3CR1-dependent renal macrophage survival promotes Candida control and host survival. *Journal of Clinical Investigation* 123, 5035–5051.
- Liu, X., Thorling, C.A., Jin, L., Roberts, M.S., 2012. Intravital multiphoton imaging of rhodamine 123 in the rat liver after intravenous dosing. *Intravital* 1, 54–59.
- Livet, J., Weissman, T.A., Kang, H., *et al.*, 2007. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* 450, 56–62.
- Llewellyn, M.E., Barretto, R.P., Delp, S.L., Schnitzer, M.J., 2008. Minimally invasive high-speed imaging of sarcomere contractile dynamics in mice and humans. *Nature* 454, 784–788.
- Lo Celso, C., Fleming, H.E., Wu, J.W., *et al.*, 2009. Live-animal tracking of individual haematopoietic stem/progenitor cells in their niche. *Nature* 457, 92–96.
- Loulier, K., Barry, R., Mahou, P., *et al.*, 2014. Multiplex cell and lineage tracking with combinatorial labels. *Neuron* 81, 505–520.
- MacDonald, I.C., Schmidt, E.E., Morris, V.L., Chambers, A.F., Groom, A.C., 1992. Intravital videomicroscopy of the chorioallantoic microcirculation: A model system for studying metastasis. *Microvascular Research* 44, 185–199.
- Mansson, L.E., Melican, K., Boekel, J., *et al.*, 2007. Real-time studies of the progression of bacterial infections and immediate tissue responses in live animals. *Cellular Microbiology* 9, 413–424.
- Masamoto, K., Tomita, Y., Toriumi, H., *et al.*, 2012. Repeated longitudinal *in vivo* imaging of neuro-glio-vascular unit at the peripheral boundary of ischemia in mouse cerebral cortex. *Neuroscience* 212, 190–200.
- Masedunskas, A., Milberg, O., Porat-Shliom, N., *et al.*, 2012a. Intravital microscopy: A practical guide on imaging intracellular structures in live animals. *BioArchitecture* 2, 143–157.
- Masedunskas, A., Porat-Shliom, N., Rechache, K., Aye, M.P., Weigert, R., 2012b. Intravital microscopy reveals differences in the kinetics of endocytic pathways between cell cultures and live animals. *Cells* 1, 1121–1132.
- Masedunskas, A., Porat-Shliom, N., Weigert, R., 2012c. Linking differences in membrane tension with the requirement for a contractile actomyosin scaffold during exocytosis in salivary glands. *Communicative & Integrative Biology* 5, 84–87.
- Masedunskas, A., Porat-Shliom, N., Weigert, R., 2012d. Regulated exocytosis: Novel insights from intravital microscopy. *Traffic* 13, 627–634.
- Masedunskas, A., Sramkova, M., Parente, L., *et al.*, 2011a. Role for the actomyosin complex in regulated exocytosis revealed by intravital microscopy. *Proceedings*

- of the National Academy of Sciences of the United States of America 108, 13552–13557.
- Masedunskas, A., Sramkova, M., Parente, L., Weigert, R., 2013. Intravital microscopy to image membrane trafficking in live rats. *Methods in Molecular Biology* 931, 153–167.
- Masedunskas, A., Sramkova, M., Weigert, R., 2011b. Homeostasis of the apical plasma membrane during regulated exocytosis in the salivary glands of live rodents. *BioArchitecture* 1, 225–229.
- Masedunskas, A., Weigert, R., 2008. Intravital two-photon microscopy for studying the uptake and trafficking of fluorescently conjugated molecules in live rodents. *Traffic* 9, 1801–1810.
- McGavern, D.B., Kang, S.S., 2011. Illuminating viral infections in the nervous system. *Nature Reviews Immunology* 11, 318–329.
- Megens, R.T., Reitsma, S., Schiffrers, P.H., *et al.*, 2007. Two-photon microscopy of vital murine elastic and muscular arteries. Combined structural and functional imaging with subcellular resolution. *Journal of Vascular Research* 44, 87–98.
- Merlini, M., Davalos, D., Akassoglou, K., 2012. *In vivo* imaging of the neurovascular unit in CNS disease. *Intravital* 1, 87–91.
- Mertz, J., 2004. Nonlinear microscopy: New techniques and applications. *Current Opinion in Neurobiology* 14, 610–616.
- Morishige, N., Petroll, W.M., Nishida, T., Kenney, M.C., Jester, J.V., 2006. Noninvasive corneal stromal collagen imaging using two-photon-generated second-harmonic signals. *Journal of Cataract & Refractive Surgery* 32, 1784–1791.
- Muller, M., Zumbusch, A., 2007. Coherent anti-Stokes Raman scattering microscopy. *ChemPhysChem* 8, 2156–2170.
- Nayak, D., Zinselmeyer, B., Corps, K., McGavern, D., 2012. *In vivo* dynamics of innate immune sentinels in the CNS. *Intravital* 1 (2), 95–106.
- Nitschke, C., Garin, A., Kosco-Vilbois, M., Gunzer, M., 2008. 3D and 4D imaging of immune cells *in vitro* and *in vivo*. *Histochemistry and Cell Biology* 130, 1053–1062.
- Nitschke, M., Aebischer, D., Abadier, M., *et al.*, 2012. Differential requirement for ROCK in dendritic cell migration within lymphatic capillaries in steady-state and inflammation. *Blood* 120, 2249–2258.
- Norman, M.U., Moriarty, T.J., Dresser, A.R., *et al.*, 2008. Molecular mechanisms involved in vascular interactions of the Lyme disease pathogen in a living host. *PLOS Pathogens* 4, e1000169.
- Nuttall, A.L., 1987. Velocity of red blood cell flow in capillaries of the guinea pig cochlea. *Hearing Research* 27, 121–128.
- Nyman, L.R., Wells, K.S., Head, W.S., *et al.*, 2008. Real-time, multidimensional *in vivo* imaging used to investigate blood flow in mouse pancreatic islets. *Journal of Clinical Investigation* 118, 3790–3797.
- Oheim, M., Michael, D.J., Geisbauer, M., Madsen, D., Chow, R.H., 2006. Principles of two-photon excitation fluorescence microscopy and other nonlinear imaging approaches. *Advanced Drug Delivery Reviews* 58, 788–808.
- Ohno, Y., Birn, H., Christensen, E.I., 2005. *In vivo* confocal laser scanning microscopy and micropuncture in intact rat. *Nephron Experimental Nephrology* 99, e17–e25.
- O'Rourke, N.A., Fraser, S.E., 1990. Dynamic changes in optic fiber terminal arbors lead to retinotopic map formation: An *in vivo* confocal microscopic study. *Neuron* 5, 159–171.
- Orth, J.D., Kohler, R.H., Foijer, F., *et al.*, 2011. Analysis of mitosis and antimitotic drug responses in tumors by *in vivo* microscopy and single-cell pharmacodynamics. *Cancer Research* 71, 4608–4616.
- Pan, F., Gan, W.B., 2008. Two-photon imaging of dendritic spine development in the mouse cortex. *Developmental Neurobiology* 68, 771–778.
- Patel, V., Marsh, C.A., Dorsam, R.T., *et al.*, 2011. Decreased lymphangiogenesis and lymph node metastasis by mTOR inhibition in head and neck cancer. *Cancer Research* 71, 7103–7112.
- Paxian, M., Keller, S.A., Cross, B., Huynh, T.T., Clemens, M.G., 2004. High-resolution visualization of oxygen distribution in the liver *in vivo*. *American Journal of Physiology Gastrointestinal and Liver Physiology* 286, G37–G44.
- Pena, A.M., Fabre, A., Debarre, D., *et al.*, 2007. Three-dimensional investigation and scoring of extracellular matrix remodeling during lung fibrosis using multiphoton microscopy. *Microscopy Research and Technique* 70, 162–170.
- Perentes, J.Y., McKee, T.D., Ley, C.D., *et al.*, 2009. *In vivo* imaging of extracellular matrix remodeling by tumor-associated fibroblasts. *Nature Methods* 6, 143–145.
- Pinner, S., Jordan, P., Sharrock, K., *et al.*, 2009. Intravital imaging reveals transient changes in pigment production and Brn2 expression during metastatic melanoma dissemination. *Cancer Research* 69, 7969–7977.
- Pittet, M.J., Weissleder, R., 2011. Intravital imaging. *Cell* 147, 983–991.
- Presson Jr., R.G., Brown, M.B., Fisher, A.J., *et al.*, 2011. Two-photon imaging within the murine thorax without respiratory and cardiac motion artifact. *American Journal of Pathology* 179, 75–82.
- Provenzano, P.P., Eliceiri, K.W., Keely, P.J., 2009. Multiphoton microscopy and fluorescence lifetime imaging microscopy (FLIM) to monitor metastasis and the tumor microenvironment. *Clinical & Experimental Metastasis* 26, 357–370.
- Qi, H., Egen, J.G., Huang, A.Y., Germain, R.N., 2006. Extrafollicular activation of lymph node B cells by antigen-bearing dendritic cells. *Science* 312, 1672–1676.
- Radosevich, A.J., Bouchard, M.B., Burgess, S.A., Chen, B.R., Hillman, E.M., 2008. Hyperspectral *in vivo* two-photon microscopy of intrinsic contrast. *Optics Letter* 33, 2164–2166.
- Ritsma, L., Ellenbroek, S.I., Zomer, A., *et al.*, 2014. Intestinal crypt homeostasis revealed at single-stem-cell level by *in vivo* live imaging. *Nature* 507, 362–365.
- Ritsma, L., Ponsioen, B., van Rhee, J., 2012a. Intravital imaging of cell signaling in mice. *Intravital* 1, 2–10.
- Ritsma, L., Steller, E.J., Beerling, E., *et al.*, 2012b. Intravital microscopy through an abdominal imaging window reveals a pre-micrometastasis stage during liver metastasis. *Science translational medicine* 4, 158ra145.
- Ritsma, L., Steller, E.J., Ellenbroek, S.I., *et al.*, 2013. Surgical implantation of an abdominal imaging window for intravital microscopy. *Nature Protocols* 8, 583–594.
- Rompalas, P., Deschene, E.R., Zito, G., *et al.*, 2012. Live imaging of stem cell and progeny behaviour in physiological hair-follicle regeneration. *Nature* 487, 496–499.
- Rompalas, P., Mesa, K.R., Greco, V., 2013. Spatial organization within a niche as a determinant of stem-cell fate. *Nature* 502, 513–518.
- Rothstein, E.C., Carroll, S., Combs, C.A., Jobsis, P.D., Balaban, R.S., 2005. Skeletal muscle NAD(P)H two-photon fluorescence microscopy *in vivo*: Topology and optical inner filters. *Biophysical Journal* 88, 2165–2176.
- Rubart, M., 2004. Two-photon microscopy of cells and tissue. *Circulation Research* 95, 1154–1166.
- Rudolf, R., Magalhaes, P.J., Pozzan, T., 2006. Direct *in vivo* monitoring of sarcoplasmic reticulum Ca²⁺ and cytosolic cAMP dynamics in mouse skeletal muscle. *Journal of Cell Biology* 173, 187–193.
- Sandoval, R.M., Kennedy, M.D., Low, P.S., Molitoris, B.A., 2004. Uptake and trafficking of fluorescent conjugates of folic acid in intact kidney determined using intravital two-photon microscopy. *American Journal of Physiology Cell Physiology* 287, C517–C526.
- Schenke-Layland, K., Xie, J., Angelis, E., *et al.*, 2008. Increased degradation of extracellular matrix structures of lacrimal glands implicated in the pathogenesis of Sjogren's syndrome. *Matrix Biology* 27, 53–66.
- Smith, B.R., Cheng, Z., De, A., *et al.*, 2008. Real-time intravital imaging of RGD-quantum dot binding to luminal endothelium in mouse tumor neovasculature. *Nano Letters* 8, 2599–2606.
- Spires, T.L., Meyer-Luehmann, M., Stern, E.A., *et al.*, 2005. Dendritic spine abnormalities in amyloid precursor protein transgenic mice demonstrated by gene transfer and intravital multiphoton microscopy. *Journal of Neuroscience* 25, 7278–7287.
- Sramkova, M., Masedunskas, A., Parente, L., Molinolo, A., Weigert, R., 2009. Expression of plasmid DNA in the salivary gland epithelium: Novel approaches to study dynamic cellular processes in live animals. *American Journal of Physiology Cell Physiology* 297, C1347–C1357.
- Sramkova, M., Masedunskas, A., Weigert, R., 2012. Plasmid DNA is internalized from the apical plasma membrane of the salivary gland epithelium in live animals. *Histochemistry and Cell Biology* 138, 201–213.
- Stockholm, D., Bartoli, M., Sillon, G., *et al.*, 2005. Imaging calpain protease activity by multiphoton FRET in living mice. *Journal of Molecular Biology* 346, 215–222.
- Stutzmann, G.E., Parker, I., 2005. Dynamic multiphoton imaging: A live view from cells to systems. *Physiology (Bethesda)* 20, 15–21.
- Sun, C.K., Zhang, X.Y., Sheard, P.W., Mabuchi, A., Wheatley, A.M., 2005. Change in mitochondrial membrane potential is the key mechanism in early warm hepatic ischemia-reperfusion injury. *Microvascular Research* 70, 102–110.
- Svoboda, K., Yasuda, R., 2006. Principles of two-photon excitation microscopy and its applications to neuroscience. *Neuron* 50, 823–839.
- Theer, P., Denk, W., 2006. On the fundamental imaging-depth limit in two-photon microscopy. *Journal of the Optical Society of America A, Optics, Image Science, and Vision* 23, 3139–3149.
- Theer, P., Hasan, M.T., Denk, W., 2003. Two-photon imaging to a depth of 1000 microm in living brains by use of a Ti:Al₂O₃ regenerative amplifier. *Optics Letter* 28, 1022–1024.
- Theruvath, T.P., Snoddy, M.C., Zhong, Z., Lemasters, J.J., 2008. Mitochondrial permeability transition in liver ischemia and reperfusion: Role of c-Jun N-terminal kinase 2. *Transplantation* 85, 1500–1504.
- Thorogood, P., Wood, A., 1987. Analysis of *in vivo* cell movement using transparent tissue systems. *Journal of Cell Science* 8, 395–413.

- Thurber, G.M., Weissleder, R., 2011. Quantitating antibody uptake *in vivo*: Conditional dependence on antigen expression levels. *Molecular Imaging and Biology* 13, 623–632.
- Thurber, G.M., Yang, K.S., Reiner, T., *et al.*, 2013. Single-cell and subcellular pharmacokinetic imaging allows insight into drug action *in vivo*. *Nature Communications* 4, 1504. doi:10.1038/ncomms2506.
- Toma, I., Kang, J.J., Peti-Peterdi, J., 2006. Imaging renin content and release in the living kidney. *Nephron Physiology* 103 (2), 71–74.
- Tozer, G.M., Ameer-Beg, S.M., Baker, J., *et al.*, 2005. Intravital imaging of tumour vascular networks using multi-photon fluorescence microscopy. *Advanced Drug Delivery Reviews* 57, 135–152.
- Vakoc, B.J., Lanning, R.M., Tyrrell, J.A., *et al.*, 2009. Three-dimensional microscopy of the tumor microenvironment *in vivo* using optical frequency domain imaging. *Nature Medicine* 15, 1219–1223.
- Villringer, A., Haberl, R.L., Dirnagl, U., *et al.*, 1989. Confocal laser microscopy to study microcirculation on the rat brain surface *in vivo*. *Brain Research* 504, 159–160.
- Vinegoni, C., Lee, S., Gorbатов, R., Weissleder, R., 2012. Motion compensation using a suctioning stabilizer for intravital microscopy. *Intravital* 1, 115–121.
- Wagner, R., 1839. *Erläuterungstafeln zur physiologie und entwicklungsgeschichte*. Leipzig: Leopold Voss.
- Wang, W., Wyckoff, J.B., Goswami, S., *et al.*, 2007. Coordinated regulation of pathways for enhanced cell motility and chemotaxis is conserved in rat and mouse mammary tumors. *Cancer Research* 67, 3505–3511.
- Weigel, B., Bakker, G., Friedl, P., 2012. Intravital third harmonic generation microscopy of collective melanoma cell invasion: Principles of interface guidance and microvesicle dynamics. *Intravital* 1, 32–43.
- Weigert, R., 2014. Imaging the dynamics of endocytosis in live mammalian tissues. *Cold Spring Harbor Perspectives in Biology* 6, a017012.
- Weigert, R., Porat-Shliom, N., Amornphimoltham, P., 2013. Imaging cell biology in live animals: Ready for prime time. *Journal of Cell Biology* 201, 969–979.
- Weigert, R., Sramkova, M., Parente, L., Amornphimoltham, P., Masedunskas, A., 2010. Intravital microscopy: A novel tool to study cell biology in living animals. *Histochemistry and Cell Biology* 133, 481–491.
- Wilson, T., 2002. Confocal microscopy: Basic principles and architectures. In: Diaspro, A. (Ed.), *Confocal and Two-Photon Microscopy: Foundations, Applications and Advances*. New York: Wiley–Liss, pp. 19–38. ISBN: 978-0-471-40920-5.
- Wood, A., Thorogood, P., 1984. An analysis of *in vivo* cell migration during teleost fin morphogenesis. *Journal of Cell Science* 66, 205–222.
- Wu, L., Tiwari, M.M., Messer, K.J., *et al.*, 2007. Peritubular capillary dysfunction and renal tubular epithelial cell stress following lipopolysaccharide administration in mice. *American Journal of Physiology – Renal Physiology* 292, F261–F268.
- Wyckoff, J.B., Pinner, S.E., Gschmeissner, S., Condeelis, J.S., Sahai, E., 2006. ROCK- and myosin-dependent matrix deformation enables protease-independent tumor-cell invasion *in vivo*. *Current Biology* 16, 1515–1523.
- Wyckoff, J.B., Wang, Y., Lin, E.Y., *et al.*, 2007. Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. *Cancer Research* 67, 2649–2656.
- Xu, H.T., Pan, F., Yang, G., Gan, W.B., 2007. Choice of cranial window type for *in vivo* imaging affects dendritic spine turnover in the cortex. *Nature Neuroscience* 10, 549–551.
- Yu, W., Sandoval, R.M., Molitoris, B.A., 2007. Rapid determination of renal filtration function using an optical ratiometric imaging approach. *American Journal of Physiology Renal Physiology* 292, F1873–F1880.
- Zariwala, H.A., Borghuis, B.G., Hoogland, T.M., *et al.*, 2012. A Cre-dependent GCaMP3 reporter mouse for neuronal imaging *in vivo*. *Journal of Neuroscience* 32 (9), 3131–3141.
- Zhang, S., Murphy, T.H., 2007. Imaging the impact of cortical microcirculation on synaptic structure and sensory-evoked hemodynamic responses *in vivo*. *PLOS Biology* 5, e119.
- Zhong, Z., Ramshesh, V.K., Rehman, H., *et al.*, 2008. Activation of the oxygen-sensing signal cascade prevents mitochondrial injury after mouse liver ischemia-reperfusion. *American Journal of Physiology Gastrointestinal and Liver Physiology* 295, G823–G832.
- Zipfel, W.R., Williams, R.M., Christie, R., *et al.*, 2003a. Live tissue intrinsic emission microscopy using multiphoton-excited native fluorescence and second harmonic generation. *Proceedings of the National Academy of Sciences of the United States of America* 100, 7075–7080.
- Zipfel, W.R., Williams, R.M., Webb, W.W., 2003b. Nonlinear magic: Multiphoton microscopy in the biosciences. *Nature Biotechnology* 21, 1369–1377.
- Zoumi, A., Yeh, A., Tromberg, B.J., 2002. Imaging cells and extracellular matrix *in vivo* by using second-harmonic generation and two-photon excited fluorescence. *Proceedings of the National Academy of Sciences of the United States of America* 99, 11014–11019.
- Zweifach, B.W., 1954. Direct observation of the mesenteric circulation in experimental animals. *Anatomical Record* 120, 277–291.

Optogenetics

EY Isacoff, University of California Berkeley, Berkeley, CA, USA

A Reiner, University of California Berkeley, Berkeley, CA, USA and Ruhr University Bochum, Bochum, Germany

© 2016 Elsevier Inc. All rights reserved.

Glossary

Channelrhodopsin (ChR) ChR is a light-gated ion channel from algae, which contains a retinal cofactor and is opened by light, that is the channel becomes permeable upon illumination.

Excitability The ability of a cell to react to a stimulus with changes in membrane potential, for example, to generate an action potential.

Glutamate receptors (GluRs) GluRs are synaptic signaling proteins, which sense glutamate, an excitatory neurotransmitter that is released from the presynaptic terminal. Two classes exist: Ionotropic glutamate receptors (iGluRs), which are tetrameric ion channels that open upon glutamate binding and allow the flux of ions, and metabotropic glutamate receptors (mGluRs), which are

dimeric G protein-coupled receptors (GPCRs) that cause G protein activation.

Light-sensitive proteins Proteins which typically contain a light-absorbing cofactor that upon illumination triggers changes in protein function.

Optogenetics Experimental strategies that combine the advantages of optical manipulation with the abilities of genetic targeting.

Photoswitch A chemical compound that reversibly changes its conformation or chemical properties upon illumination with light.

Photoswitchable tethered ligand (PTL) A ligand that encompasses a photoswitch to light-control its binding and unbinding behavior, as well as a group for covalent tethering to the protein of interest.

Light As a Tool in Biology

Light is orthogonal to most biological systems and therefore fairly noninvasive. Moreover, light provides the resolution to resolve most biological processes in space and time. Not surprisingly, this made light microscopy a pivotal tool in biological research. But in the past decade, light has also become an increasingly valuable tool in another way. Next to the purpose of visualization in imaging-based approaches, the advantages of light are now also utilized to manipulate biological processes. Through the use of light-sensitive proteins and chemical photoswitches it has become possible to optically control numerous functions in living cells, which in combination with modern genetics has enabled a new, versatile method for biological experimentation, called optogenetics.

Optogenetics – Manipulating Molecules, Cells, and Organisms with Light

Optogenetic approaches allow for the optical control of a desired molecular or cellular function through targeting of light-sensitive (bio)molecules to a subset of genetically defined cells. The motivation for using such optogenetic strategies is twofold: (1) Light offers high spatial and temporal resolution for manipulating biological processes in living cells and organisms, with a precision and selectivity that is hard to achieve by other means. (2) Genetic tools enable the experimenter to define precisely, which cells are manipulated. Optogenetics thus gives a handle to probe in real time how specific molecules or cells contribute to complex cellular networks, to the function of organs or the behavior of whole organisms. Due to these advantages, optogenetic approaches

are now applied in many experimental paradigms (Miesenböck, 2009, 2011; Fenno *et al.*, 2011). As we will describe in the following sections, the key factors for optogenetic experimentation are the light-sensitive (bio)molecule used, the method of genetic targeting, and the means of illumination (Figure 1).

Light-Sensitive (Bio)molecules

Modulation with light can be achieved either through the expression of light-sensitive proteins or through the application of synthetic photoswitches that target specific proteins or biological processes. Examples of various optogenetic tools along with their key properties are given in Table 1. This list represents only a small selection of the available tools, also because the development of optogenetic methods continues to be an active area of research.

Naturally occurring light-sensitive proteins contain cofactors, which absorb light in certain wavelength ranges, and, which upon illumination modulate protein function. For instance, vision in vertebrates and invertebrates is mediated by rhodopsins ('visual pigments'), which are G protein-coupled photoreceptors that contain a covalently bound retinal cofactor. In the case of rhodopsins, the absorption of light induces an isomerization of the 11-*cis* form of retinal to an all-*trans* form, which leads to receptor activation and subsequent activation of G proteins and downstream signaling cascades (Palczewski, 2006). Indeed, invertebrate rhodopsins, which couple to the G_q class of G proteins and thereby indirectly increase cellular excitability, were used in the first optogenetic experiment (Zemelman *et al.*, 2002). However, invertebrate rhodopsins have found limited application, since their use in vertebrates requires the expression of additional G_q signaling components. On the other hand, an increasing number of

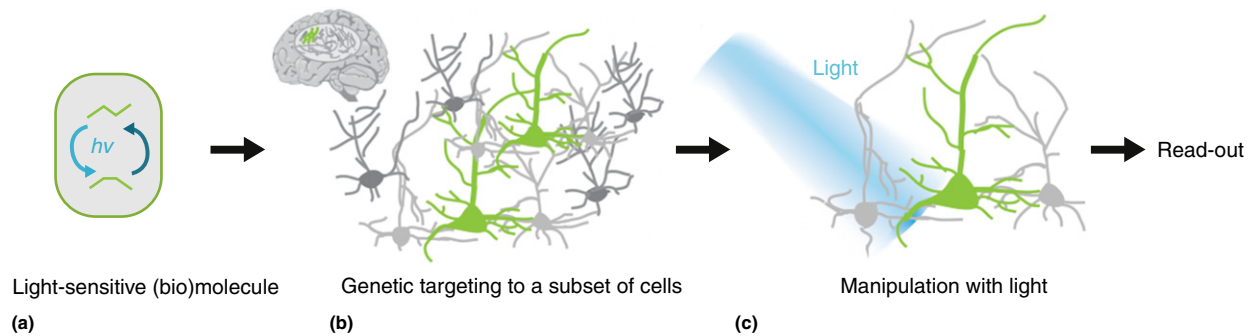


Figure 1 Scheme of optogenetic experimentation. (a) Light-sensitive or modified proteins (green box), which change their function in response to light, are used as tools to optically control specific cellular functions. (b) These light-sensitive proteins are expressed in defined regions or subsets of cells using genetic tools, such as viruses, driver lines, or transgenic animals. This can provide high specificity to which cells are manipulated, for example, which neurons in a certain brain region are targeted. (c) Illumination with light is used to control the desired molecular or cellular function, if necessary, with high spatial and/or temporal resolution. The observed read-out of an optogenetic experiment depends on the system and can be virtually anything, from changes in molecular processes, over cellular network functions to changes in animal behavior.

Table 1 Examples of light-sensitive proteins and chemical photoswitches used for optogenetic studies

Optogenetic tools	Function	Used for	References
<i>Rhodopsins (retinal cofactors)</i>			
Vertebrate rhodopsin	Light-activated G protein-coupled receptor (GPCR)	G _{i/o} signaling	Li <i>et al.</i> (2005); Masseck <i>et al.</i> (2011)
Channelrhodopsin (ChR)	Light-gated ion channel	Depolarization	Boyden <i>et al.</i> (2005)
Halorhodopsin (HR)	Light-driven Cl [−] pump	Hyperpolarization	Zhang <i>et al.</i> (2007)
Archaeorhodopsin (Arch)	Light-driven H ⁺ pump	Hyperpolarization	Chow <i>et al.</i> (2010)
<i>LOV/BLUF photoreceptors (flavin cofactors)</i>			
Adenylyl cyclases (PACs)	Light-driven cAMP production	cAMP signaling	Schröder-Lang <i>et al.</i> (2007); Raffelberg <i>et al.</i> (2013)
Isolated LOV/BLUF domains	Conformational changes	Engineering	Möglich and Moffat (2010)
<i>Phytochromes (bilin cofactors)</i>			
PhyB-PIF3	Light-controlled protein association	Dimerization, co-localization	Shimizu-Sato <i>et al.</i> (2002); Möglich and Moffat (2010)
<i>Synthetic photoswitches (azobenzenes)</i>			
Photoswitchable tethered ligands (PTLs)	Activation/inhibition of ligand-gated receptors	iGluRs mGluRs AChRs GABA _A Rs	Volgraf <i>et al.</i> (2006); Reiner <i>et al.</i> (2015) Levitz <i>et al.</i> (2013) Tochitsky <i>et al.</i> (2012) Lin <i>et al.</i> (2014)
Photoswitchable tethered blockers	Pore block of K ⁺ channels	K _v , SK, K _{2P}	Banghart <i>et al.</i> (2004); Sandoz <i>et al.</i> (2012)
Photoswitchable allosteric modulators	Allosteric modulation	P2X receptors	Lemoine <i>et al.</i> (2013); Browne <i>et al.</i> (2014)

vertebrate rhodopsin variants are utilized in optogenetic experiments. For instance, rat rhodopsin 4 (RO4), which is G_{i/o} coupled, can be used for light-induced hyperpolarization of cells through the activation of GIRK channels, while it also inhibits presynaptic Ca²⁺ channels (Li *et al.*, 2005). Importantly, the specificity of rhodopsin-mediated signaling can be increased by targeting the rhodopsins to specific cellular localizations by fusing C-termini from other G protein-coupled receptor (GPCR) to tap into local G protein signaling cascades (Oh *et al.*, 2010), or by replacing the intracellular loops to achieve coupling to G_s or G_q pathways (Kim *et al.*, 2005; Airan *et al.*, 2009). Recently, also melanopsin and cone opsins, the latter providing very high light sensitivity, have

been exploited as optogenetic tools (Masseck *et al.*, 2014; Spoida *et al.*, 2014).

Another, very versatile class of light-sensitive proteins is given by microbial rhodopsins, which like animal rhodopsins contain a retinal cofactor, but comprise other functionalities than G protein activation. Channelrhodopsins (ChRs), for instance, are light-gated ion channels controlling phototaxis in green algae (Nagel *et al.*, 2002, 2003), whereas halorhodopsins (HRs) and archaeorhodopsins (Archs), which are both of archaeal origin, are light-driven chloride and proton pumps, respectively. Photo-activation of ChRs or HRs/Archs directly affects the ionic distribution across the cell membrane, which allows for the rapid control of membrane potential and

excitability. We discuss this application in more detail in the section Light-Sensitive Proteins To Control Cellular Excitability (Boyden *et al.*, 2005; Li *et al.*, 2005; Han and Boyden, 2007; Zhang *et al.*, 2007; Chow *et al.*, 2010). Typically, microbial rhodopsins are further optimized by protein engineering, i.e., by amino acid substitutions or the generation of chimeric proteins, to improve their expression in the desired eukaryotic hosts and to achieve sensitive and/or fast activation with light of different wavelengths (Lin *et al.*, 2009; Gunaydin *et al.*, 2010; Berndt *et al.*, 2011; Mattis *et al.*, 2012; Lin *et al.*, 2013; Klapoetke *et al.*, 2014). Recently, also synthetic retinal analogs were explored for tuning the spectral and temporal characteristics of rhodopsin-based tools (AzimiHashemi *et al.*, 2014).

Bacteria and plants are also a rich source of photoreceptors with cofactors other than retinal, including xanthopsins, phytochromes, cryptochromes, and LOV and BLUF domain-containing photoreceptors (Table 1; Möglich *et al.*, 2010). Several naturally occurring photoreceptors are used as optogenetic tools, such as photo-activated adenylyl cyclases (PACs) that can be used to directly control cellular cAMP levels. PACs have been found in the flagellate *Euglena gracilis* and in bacteria and are either controlled by flavin-containing BLUF or LOV photoreceptor domains (Schröder-Lang *et al.*, 2007; Stierl *et al.*, 2011; Raffelberg *et al.*, 2013). The modular nature of plant photoreceptor also facilitates the engineering of new light-sensitive functionalities by combining different photoreceptor 'receiver' domains with different effector 'output' domains (Möglich and Moffat, 2010; Pathak *et al.*, 2013). A popular choice are LOV domains, in which illumination leads to the dissociation and unfolding of a C-terminal helix (Harper *et al.*, 2003). LOV domains have therefore been used as conformational switches to control other proteins, for example, DNA-binding proteins (Strickland *et al.*, 2008), enzymatic activities (Lee *et al.*, 2008), or signaling proteins such as small GTPases (Wu *et al.*, 2009) or receptor tyrosine kinases (Grusch *et al.*, 2014). Cryptochromes and plant phytochromes, which undergo blue- and red-light induced dimerization, respectively, are popular tools to control protein association and co-localization (Shimizu-Sato *et al.*, 2002; Kennedy *et al.*, 2010). Recently a bacterial phytochrome was used to construct a red-light-activated cAMP/cGMP-specific phosphodiesterase (Gasser *et al.*, 2014).

Another strategy to obtain light-control is to rely on chemical photoswitches to manipulate proteins that are not light-sensitive otherwise. A wide variety of photoswitches exists (Szymanski *et al.*, 2013), with azobenzene derivatives being the most popular choice to control biomolecules (Beharry and Woolley, 2011). Azobenzenes, which undergo a photo-induced *cis/trans* isomerisation, have excellent photostability and can be used to build photoswitchable ligands, blockers, or allosteric modulators (Table 1). In the case of these optochemical approaches the genetic specificity is ensured by the target molecule – typically a protein that carries a single amino acid mutation to covalently bind the photoswitch. We give a more detailed description of this approach in the section Optochemical Approaches To Control Specific Receptors and Ion Channels.

The example sections below provide more detailed descriptions of how light-sensitive proteins, as well as chemical photoswitches are used as optogenetic tools. When choosing a

tool for experiments, it is necessary to pay close attention to its characteristics, including the expression in the desired host system, the necessity to add cofactors, possible unspecific effects, dark activity, wavelength sensitivity, and so on. In all cases, it is mandatory to perform proper control experiments in the context of the system under investigation, in particular to establish that no unspecific effects arise from just expressing the optical actuator or from just applying light.

Genetic Targeting

While optical control on its own is powerful, the full strength of optogenetic approaches typically comes from the ability to limit the manipulation to certain types or subsets of cells (Zemelman and Miesenböck, 2001). Here optogenetics benefits from the numerous genetic tools that have been developed to drive expression of reporter proteins in specific cell types. These tools encompass gene delivery methods (e.g., biolistic, chemical, and viral transfection or chromosomal integration), as well as cell-type specific promoters and driver lines (e.g., GAL4/UAS or Cre-Lox systems) that have been identified or generated for various model organisms, such as worm, fly, fish and mouse. Applications in living rodents and primates often rely on a spatially restricted viral delivery using lenti- or adeno-associated viral (AAV) vectors in combination with specific promoters or enhancers. Moreover, a number of transgenic mice expressing rhodopsins directly or in a recombinase-dependent fashion have been developed (Arenkiel *et al.*, 2007; Fenno *et al.*, 2011; Madisen *et al.*, 2012). These genetic tools are extremely powerful, since they can be combined to achieve conditional labeling patterns, to get additional pharmacological control on developmental timescales, or to achieve retrograde labeling of neurons. New genome editing tools and new transgenic animals are likely to enhance these capabilities further.

Delivery of Light

The third factor in optogenetic experimentation is the delivery of an adequate light stimulus. Again numerous options exist, depending on the particular application. Lasers, diode lasers, and now more often inexpensive light emitting diodes (LEDs) are used to generate light of appropriate wavelengths and intensity. In cases, where high spatial resolution is paramount, for example to limit the manipulation to single cells or sub-cellular regions, a combination with an objective, typically a microscope, is used to deliver a small or patterned light stimulus. Laser scanning of large areas is possible, although this may decrease the sensitivity of optogenetic tools that are transiently activated. Alternative options are micromirror devices or holographic methods. Furthermore, to achieve high resolution in all dimensions, for example, in multi-layered tissues, it might be necessary to rely on two-photon methods (Papagiakoumou *et al.*, 2010; Vaziri and Emiliani, 2012).

In other cases spatial resolution is not a major concern, or, it is even desirable to manipulate cells in a wide region, for example, across some hundred microns in a layer of the cerebral cortex. In such cases, illumination of deeper regions can be provided through an implanted optical fiber, or, of more accessible areas through an implanted window.

Red-light activated photoreceptors can be advantageous, since red light penetrates deeper into tissue and in this way even transcranial activation has been achieved in mice (Lin *et al.*, 2013). In the case of smaller and transparent animals like fly, worm or zebrafish larvae light can be applied from the outside, either as full-field flash or again as a patterned light stimulus to activate individual cells.

Light-Sensitive Proteins to Control Cellular Excitability

An early example is the use of optogenetic tools to control the membrane potential and thereby neuronal excitability (Zemelman *et al.*, 2002; Banghart *et al.*, 2004). An important

breakthrough was the cloning and characterization of channelrhodopsin-2 (ChR2), a light-gated ion channel from the green algae *Chlamydomonas reinhardtii* that becomes conductive under illumination with blue light (Nagel *et al.*, 2003; Table 1). Importantly, ChR2 can be robustly expressed in mammalian cells, including neurons, and light-induced channel opening leads to a significant depolarization of the cell membrane thereby increasing neuronal excitability. Indeed, light can be used to induce action potential firing with high precision (Figures 2(a)–2(c); Boyden *et al.*, 2005; Li *et al.*, 2005). In most systems ChR2 works sufficiently well with endogenous levels of retinal so that no exogenous retinal has to be supplemented. The utility of ChR2 for activating neurons was quickly realized in different systems – next to cultured neurons it was used in the spinal cord and in chicken embryos (Li *et al.*, 2005), soon after in hippocampal slice

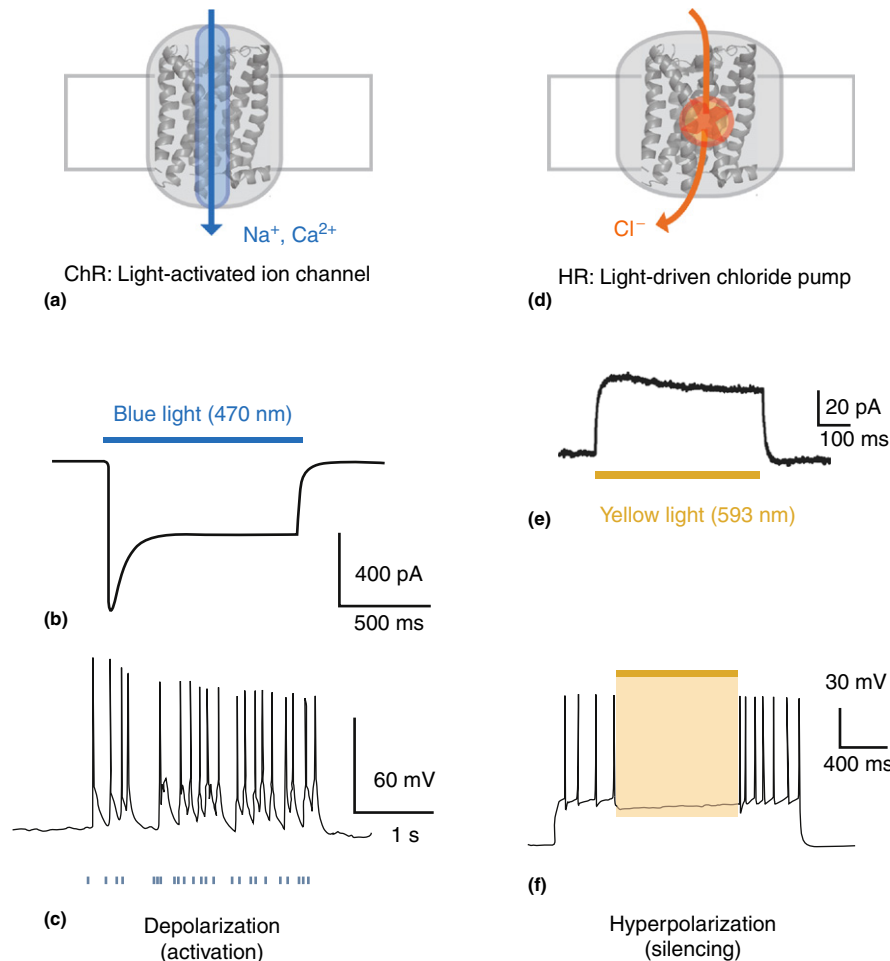


Figure 2 Optical control of neuronal excitability with ChR and HR. (a) ChR is a blue light-activated ion channel that, when expressed in mammalian cells, leads to a light-induced decrease in membrane potential (depolarization). (b) Voltage-clamp recording of a cell expressing ChR2. Illumination (blue bar) leads to an inward current, mainly through an influx of sodium ions. (c) Optical control of action potential firing in a ChR2 expressing neuron. Brief light pulses (blue bars) elicit single action potentials in a current-clamp recording. (b and c) Adapted with permission from Boyden, E.S., Zhang, F., Bamberg, E., Nagel, G., Deisseroth, K., 2005. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience* 8, 1263–1268, © 2005, Macmillan Publishers Ltd. (d) HR is a yellow light-driven chloride pump, which means that illumination causes an increase in membrane potential (hyperpolarization). (e) Voltage-clamp recording of a cell expressing *Natronomonas pharaonis* HR (NpHR). Light (yellow bar) drives NpHR to pump chloride ions into the cell, which results in an outward photocurrent. (f) Optical silencing of neuronal activity in a NpHR expressing neuron. Yellow light reversibly suppresses action potential firing. In this current-clamp recording firing had been induced through a current injection. (e and f) Adapted with permission from Zhang, F., Wang, L.P., Brauner, M., *et al.*, 2007. Multimodal fast optical interrogation of neural circuitry. *Nature* 446, 633–639, © 2007, Macmillan Publishers Ltd.

after stereotactic injection of ChR2-pseudovirions into mice (Ishizuka *et al.*, 2006), in the worm *Caenorhabditis elegans* to control behavior (Nagel *et al.*, 2005), and in the retina of mice models of inherited blindness to restore light sensitivity (Bi *et al.*, 2006). From this point on, ChR2 quickly became a standard tool to functionally probe connectivity, to study the role of particular subtypes of neurons or brain regions for circuit function, or to control the behavior of animals (Tye and Deisseroth, 2012; Reiner and Isacoff, 2013). Many improved variants of ChR are now available, including variants that are optimized for high frequency stimulation, more sensitive variants that turn off slowly after illumination, and variants for two-photon or red-light activation (Lin *et al.*, 2009; Gunaydin *et al.*, 2010; Berndt *et al.*, 2011; Mattis *et al.*, 2012; Lin *et al.*, 2013; Klapoetke *et al.*, 2014). Other variants have increased Ca^{2+} -permeability (Kleinlogel *et al.*, 2011), and based on recent structural and functional information (Eisenhauer *et al.*, 2012; Kato *et al.*, 2012) chloride permeable ChR variants were engineered, which can be used to shunt the membrane potential and thereby suppress excitations (Berndt *et al.*, 2014; Wietek *et al.*, 2014).

Soon after the introduction of ChR the first tools for directly hyperpolarizing cells were introduced in order to silence neurons and to prevent action potential firing. This was achieved with halorhodopsin from *Natronomonas pharaonis* (NpHR), a light-driven ion pump that upon illumination with yellow light pumps chloride ions into the cell (Figures 2(d)–2(f); Han and Boyden, 2007; Zhang *et al.*, 2007). Since ChR2 is activated by blue light and NpHR is driven by yellow light, these tools can be used in parallel to depolarize or hyperpolarize a given neuron, respectively. Somewhat later, archaeorhodopsin-3 (Arch), which is a light-driven outward proton pump from *Halorubrum sodomense*, was described as an alternative tool for silencing neurons (Chow *et al.*, 2010). Recently, a class of red-light driven cruxhalorhodopsin chloride pumps was characterized (Chuang *et al.*, 2014).

Optochemical Approaches to Control Specific Receptors and Ion Channels

The previous section gave examples of how the signaling state of cells can be controlled by the expression of light-sensitive proteins. In this section we focus on a second set of optogenetic tools that is concerned with controlling signaling proteins that are native to the cell. Such direct optical control is particularly helpful to delineate the physiological function of individual signaling components, and to test how and when these components cause changes within a given cell or within a given network. In the following example we summarize the use of photoswitchable tethered ligands (PTLs) to control ionotropic and metabotropic glutamate receptors (iGluRs and mGluRs), which are two important receptor families that sense glutamate, the major excitatory neurotransmitter in the central nervous system.

Light-Gated Glutamate Receptors – Principles

Glutamate receptors are cell surface receptors that, upon glutamate binding, generate postsynaptic currents and/or

intracellular signals. The spatio-temporal resolution of optical techniques bears great potential to understand how iGluRs and mGluRs contribute to synaptic transmission, homeostasis, and plasticity, especially since synaptic signaling happens on the timescale of few to hundreds of milliseconds and in compartments smaller than a micron in size (Reiner *et al.*, 2015). Equally desirable is a high molecular and cellular specificity of these manipulations in order to differentiate between different iGluR and mGluR family members, as well as to distinguish effects mediated by receptors located on the presynaptic neuron, the postsynaptic neuron, surrounding glia cells, or even neighboring neurons. The use of PTLs to optically control specific GluRs in real time is hence an important objective.

PTLs are ligands that are permanently linked to a specific receptor subunit and which contain an azobenzene photo-switch to control binding and unbinding of the ligand head-group to the receptor (Figures 3(a) and 3(b)). GluRs can be controlled by MAG ligands, which encompass a maleimide group for covalent attachment to an engineered cysteine residue on the receptor, an azobenzene group that serves as *cis/trans* photoswitch, and a glutamate derivative that is capable of activating the receptor (Figure 3(c)). The MAG ligand is attached to the receptor of interest by coupling to a cysteine residue, which is introduced close to the glutamate-binding site. The expression of this engineered receptor subunit is also the part that provides molecular and genetic specificity to this approach.

For instance, GluK2, a homotetrameric kainate receptor (a.k.a iGluR6), can be converted into a light-gated glutamate receptor (LiGluR) by introducing a cysteine substitution (L439C) at a position that is close to the glutamate-binding pocket, and attaching L-MAG0 to it (Figures 3(d) and 3(e)). As long as the azobenzene group is in its *trans* configuration, the glutamate headgroup of the PTL will not be able to reach into the binding pocket. Photo-isomerisation of MAG into its *cis* configuration, however, leads to rapid ligand binding and subsequent opening of the GluK2 ion channel pore (Volgraf *et al.*, 2006; Gorostiza *et al.*, 2007). A second wavelength of light can then be used to isomerize MAG back to *trans*, which reverses the effect and causes the channel to close. Photo-switching is fully reversible and azobenzenes have excellent photo-stability for hundreds of switching cycles. The efficacy and directionality of photoswitching depends on the ligand geometry (e.g., the linker lengths, which increases from MAG0 to MAG1 and MAG2) and the exact labeling position on the receptor, that is other labeling positions can result in a different outcome (Numano *et al.*, 2009).

The azobenzene photoswitch can be rationally modified, for instance, to change its wavelength sensitivity or the timescale on which the spontaneous *cis*-to-*trans* relaxation occurs: 'Regular' MAG is converted to *cis* by near UV light (~ 380 nm), and back to *trans* with blue light (~ 480 nm) (Figures 3(c)–3(e); Gorostiza *et al.*, 2007). It relaxes from *cis*-to-*trans* on the timescale of tens of minutes and thus displays bi-stable behavior on most experimental timescales. In contrast, MAG₄₆₀, which is built around a push-pull azobenzene, shows a considerably shifted action spectrum and allows for receptor activation with visible blue light (~ 460 nm) (Kienzler *et al.*, 2013). MAG₄₆₀ also has a faster *cis*-to-*trans* relaxation on the seconds timescale, i.e., the switch readily turns off in the dark. An important consequence of this behavior is, that within

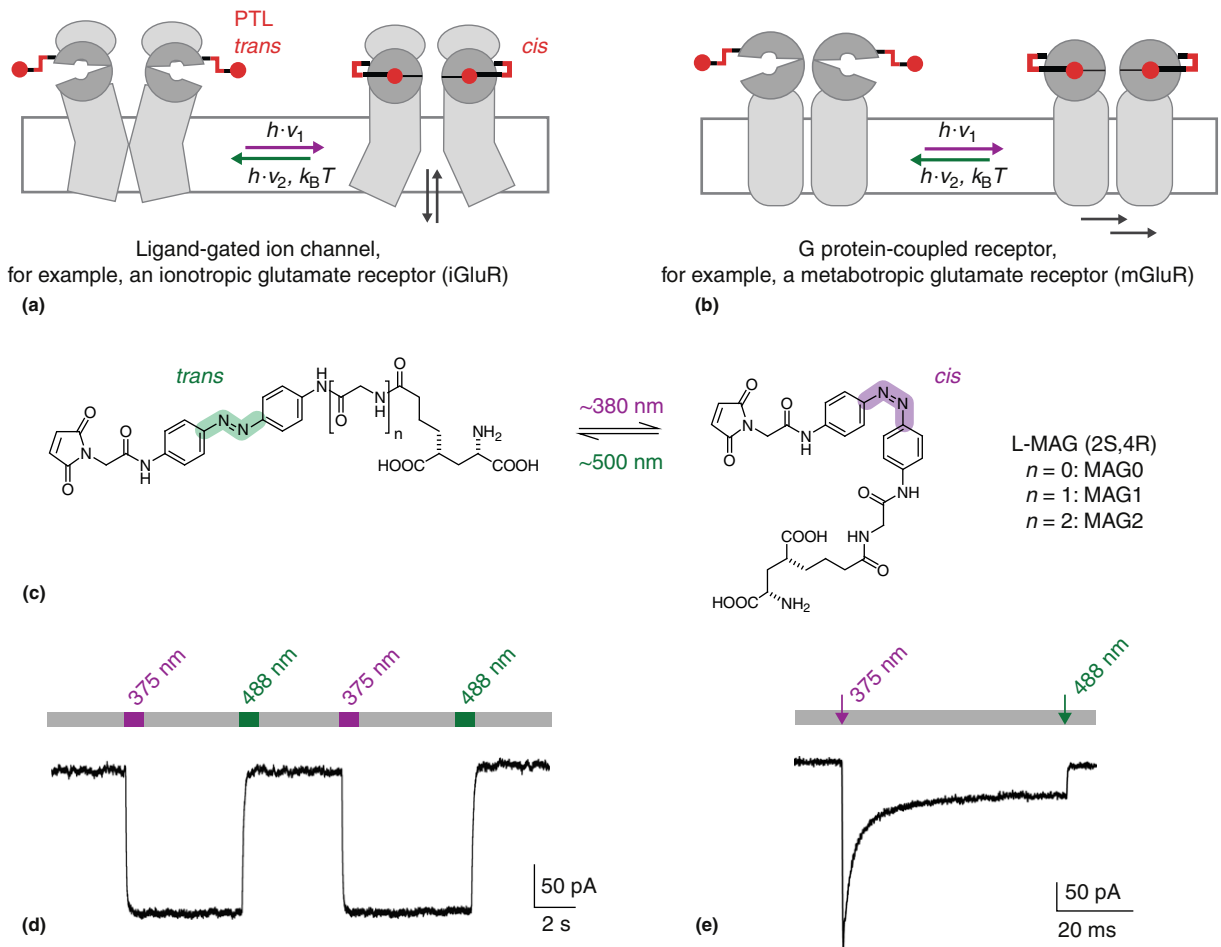


Figure 3 Photoswitchable tethered ligands (PTLs) and light-gated glutamate receptors. (a) PTLs (red) are ligands that are covalently attached to a receptor of interest, for example, a ligand-gated ion channel, or, (b) a G protein-coupled receptor to control receptor activation and deactivation with light of different wavelengths. (c) MAG-PTLs for the optical control of glutamate receptors encompass a maleimide for covalent attachment to an engineered cysteine residue, an azobenzene *cis/trans* photoswitch, and a glutamate derivative for receptor activation. Regular MAGs, as shown here, are switched from *trans*-to-*cis* with near UV light (~ 380 nm) and back to *trans* with blue light (~ 500 nm) (Gorostiza *et al.*, 2007). (d) Reversible LiGluR photoswitching after expression of GluK2(L439C) and labeling with L-MAG0. The voltage-clamp recording in the presence of desensitization blocker shows bi-stable activation with 375 nm light (violet bar) and deactivation with 488 nm light (green bar) (for details see Gorostiza *et al.*, 2007). (e) LiGluR photoswitching with high time resolution. A 100 μ s light pulse leads to rapid activation and subsequent desensitization on the milliseconds timescale, which enables the real-time manipulation of these synaptic signaling proteins. (d and e) Adapted from Reiner, A., Isacoff, E.Y., 2014. Tethered ligands reveal glutamate receptor desensitization depends on subunit occupancy. *Nature Chemical Biology* 10, 273–280.

certain intensity ranges, MAG₄₆₀ can be used to sense changes in light intensity, much like a photodiode (Gaub *et al.*, 2014). Current efforts are directed toward generating MAG photoswitches for efficient two-photon activation (Izquierdo-Serra *et al.*, 2014; Carroll *et al.*, 2015), or activation with red-light (Rullo *et al.*, 2014). Modifications can be made not only to the photoswitch, but also to the protein. An extreme example is the generation of HyLighter, which instead of an unselective ion pore typical for iGluRs, contains a potassium selective ion pore and thus leads to hyperpolarization rather than depolarization (Janovjak *et al.*, 2010).

Recent experiments also demonstrate the enormous temporal resolution that can be achieved with PTLs. The azobenzene photoisomerization is an extremely rapid process and tethering the ligand generates high local concentrations, which

circumvents slow diffusion steps. At higher light intensities, light pulses as short as 10 μ s are sufficient to induce rapid binding and receptor activation, which leads to opening of GluK2 receptors in less than a millisecond (Figure 3(e); Reiner and Isacoff, 2014). This fast activation is followed by receptor desensitization on the timescale of a few milliseconds, which is typical for glutamate-induced gating of kainate receptors. Rapid optical manipulation with short light pulses thus enables biophysical studies of the gating mechanism of homo- and heteromeric iGluR complexes and can be used to precisely mimic or shape their synaptic activation profiles.

Next to iGluRs, MAG photoswitches have been applied to control mGluRs, which are G protein-coupled receptors. In the case of light-gated mGluRs ('LimGluRs') the glutamate head-group of MAG has to be linked in D-stereochemistry. Labeling

of mGluR2 at L300C (and similarly of mGluR3 at Q306C) allows for reversible photo-activation ('photo-agonism') with D-MAG0 (Levitz *et al.*, 2013). Photo-activation of these $G_{i/o}$ -coupled receptors inhibits adenylyl cyclase, activates GIRK channels and inhibits presynaptic voltage-gated Ca^{2+} channels. PTLs can also be used for photo-inhibition ('photo-antagonism'), i.e., to block receptor activation by endogenous glutamate, which is observed if mGluR2 is labeled with D-MAG1 at S302C (Levitz *et al.*, 2013). Such light-induced antagonism can be particularly useful to acutely and reversibly block the function of specific receptor subtypes circumventing the disadvantages of genetic knock-outs or pharmacological inhibition.

Light-Gated Glutamate Receptors – Applications

Apart from molecular studies, LiGluRs, and more recently LimGluRs, have been used in a variety of applications and model organisms. For instance, using LiGluR to exclusively activate postsynaptic receptors in the fly neuromuscular junction, Kauwe and Isacoff (2013) found changes in the presynaptic release probability, indicating a novel plasticity mechanism. LiGluR is also somewhat permeable to Ca^{2+} ions and has therefore been used to trigger exocytosis from neuroendocrine cells (Izquierdo-Serra *et al.*, 2013), or to evoke gliotransmission through robust Ca^{2+} influxes into astrocytes (Li *et al.*, 2012). LimGluR2, for example, has been applied to control presynaptic release probability and short-term plasticity in hippocampal neurons (Levitz *et al.*, 2013).

Next to the synaptic context, LiGluR can be used to lower the membrane potential and to control firing on short and long timescales (Szobota *et al.*, 2007; Hou *et al.*, 2011), whereas HyLighter is a hyperpolarizing tool that is well suited for neuronal silencing (Janovjak *et al.*, 2010). As optogenetic tools, LiGluR and LimGluRs have also been used on a systems-wide level. Targeted expression of LiGluR in transgenic zebrafish lines allowed dissecting the role of different types of neurons in the spinal cord for the swimming behavior of zebrafish larvae (Wyart *et al.*, 2009). LimGluR2 has been shown to affect the acoustic startle response in zebrafish (Levitz *et al.*, 2013). Like other optogenetic tools, LiGluRs and LimGluRs may also have therapeutic potential. LiGluR, in particular, has been used to restore light sensitivity to retinæ, which have lost their photoreceptor cells due to inherited forms of blindness (Caporale *et al.*, 2011; Gaub *et al.*, 2014).

Other Targets

While MAGs are designed to control glutamate receptors, the same strategy of using tethered photoswitches can be applied to other receptors and channels residing on the cell surface. So far, PTLs have been used to control acetylcholine receptors (Lester *et al.*, 1980; Tochitsky *et al.*, 2012) and GABA_A receptors (Lin *et al.*, 2014), two other important classes of neurotransmitter-gated ion channels. A similar concept has been used to control a variety of native potassium channels. In that case, a photo-switchable pore blocker is installed instead of a ligand to control access to the ion permeation pathway in a light-dependent manner (Table 1; Banghart *et al.*, 2004; Sandoz and Levitz, 2013). This strategy has been applied to voltage-gated and

calcium-regulated potassium channels (Kramer *et al.*, 2013), and it was used to provide a pharmacological signature on two-pore-domain 'leak' potassium channels (Sandoz *et al.*, 2012). A third strategy, next to using photoswitches as ligands or blockers, is to employ photoswitches at allosteric sites, where changes of the electrostatics or local conformation affects the function, as used to optically control ATP-gated P2X receptors (Lemoine *et al.*, 2013; Browne *et al.*, 2014). Overall tethered photoswitches provide a versatile strategy to control protein function with genetic precision and more compounds will become available, given the wide success of soluble photoswitches (Fehrentz *et al.*, 2011; Kramer *et al.*, 2013).

Summary – From Molecules to Systems

The manipulation with light comes as a perfect complement to our well-established techniques to visualize and monitor biological systems. Optogenetic techniques get continuously refined and are actively used in a wide range of applications. Light-sensitive molecules, such as photoreceptors or chemical photoswitches, can be turned into powerful tools, when combined with modern genetic methods and state-of-the-art optical techniques. The versatility of these approaches is enormous, since the manipulations can span from molecules, cellular compartments, and cell-to-cell communication to larger networks and complex behaviors in living animals. Moreover, optogenetics can be combined with almost any read-out, such as imaging methods, electrical recording techniques, behavioral assays, or even magnetic resonance imaging (Lee *et al.*, 2010) or ultrastructural techniques (Watanabe *et al.*, 2013). Optogenetics are also actively used in disease related research and are explored for clinical applications, for example, for vision restoration or deep-brain stimulation. These tools, although mostly pioneered with studies in the central nervous system, also hold great promise for use in other areas of biological research (Arrenberg *et al.*, 2010; Jansen *et al.*, 2015).

Acknowledgments

Work in our lab was supported by the National Institutes of Health Nanomedicine Development Center for the Optical Control of Biological Function (2PN2EY018241). A.R. is currently funded by the Ministry of Innovation, Science and Research of the German State of North Rhine-Westphalia.

See also: Imaging the Cell: Light Microscopy: Super Resolution Fluorescence Localization Microscopy; Super-Resolution Light Microscopy: Stimulated Emission Depletion and Ground-State Depletion

References

- Airan, R.D., Thompson, K.R., Fenno, L.E., Bernstein, H., Deisseroth, K., 2009. Temporally precise *in vivo* control of intracellular signaling. *Nature* 458, 1025–1029.

- Arenkiel, B.R., Peca, J., Davison, I.G., *et al.*, 2007. *In vivo* light-induced activation of neural circuitry in transgenic mice expressing channelrhodopsin-2. *Neuron* 54, 205–218.
- Arrenberg, A.B., Stainier, D.Y., Baier, H., Huisken, J., 2010. Optogenetic control of cardiac function. *Science* 330, 971–974.
- AzimiHashemi, N., Erbguth, K., Vogt, A., *et al.*, 2014. Synthetic retinal analogs modify the spectral and kinetic characteristics of microbial rhodopsin optogenetic tools. *Nature Communications* 5, 5810.
- Banghart, M., Borges, K., Isacoff, E., Trauner, D., Kramer, R.H., 2004. Light-activated ion channels for remote control of neuronal firing. *Nature Neuroscience* 7, 1381–1386.
- Beharry, A.A., Woolley, G.A., 2011. Azobenzene photoswitches for biomolecules. *Chemical Society Reviews* 40, 4422–4437.
- Berndt, A., Lee, S.Y., Ramakrishnan, C., Deisseroth, K., 2014. Structure-guided transformation of channelrhodopsin into a light-activated chloride channel. *Science* 344, 420–424.
- Berndt, A., Schoenenberger, P., Mattis, J., *et al.*, 2011. High-efficiency channelrhodopsins for fast neuronal stimulation at low light levels. *Proceedings of the National Academy of Sciences of the United States of America* 108, 7595–7600.
- Bi, A., Cui, J., Ma, Y.P., *et al.*, 2006. Ectopic expression of a microbial-type rhodopsin restores visual responses in mice with photoreceptor degeneration. *Neuron* 50, 23–33.
- Boyden, E.S., Zhang, F., Bamberg, E., Nagel, G., Deisseroth, K., 2005. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience* 8, 1263–1268.
- Browne, L.E., Nunes, J.P., Sim, J.A., *et al.*, 2014. Optical control of trimeric P2X receptors and acid-sensing ion channels. *Proceedings of the National Academy of Sciences of the United States of America* 111, 521–526.
- Caporale, N., Kolstad, K.D., Lee, T., *et al.*, 2011. LiGluR restores visual responses in rodent models of inherited blindness. *Molecular Therapy* 19, 1212–1219.
- Carroll, E.C., Berlin, S., Levitz, J., *et al.*, 2015. Two-photon brightness of azobenzene photoswitches designed for glutamate receptor optogenetics. *Proceedings of the National Academy of Sciences of the United States of America* 112, E776–E785.
- Chow, B.Y., Han, X., Dobry, A.S., *et al.*, 2010. High-performance genetically targetable optical neural silencing by light-driven proton pumps. *Nature* 463, 98–102.
- Chuong, A.S., Miri, M.L., Busskamp, V., *et al.*, 2014. Noninvasive optical inhibition with a red-shifted microbial rhodopsin. *Nature Neuroscience* 17, 1123–1129.
- Eisenhauer, K., Kuhne, J., Ritter, E., *et al.*, 2012. In channelrhodopsin-2 Glu-90 is crucial for ion selectivity and is deprotonated during the photocycle. *Journal of Biological Chemistry* 287, 6904–6911.
- Fehrentz, T., Schönberger, M., Trauner, D., 2011. Optochemical genetics. *Angewandte Chemie* 50, 12156–12182.
- Fenno, L., Yizhar, O., Deisseroth, K., 2011. The development and application of optogenetics. *Annual Review of Neuroscience* 34, 389–412.
- Gasser, C., Taiber, S., Yeh, C.M., *et al.*, 2014. Engineering of a red-light-activated human cAMP/cGMP-specific phosphodiesterase. *Proceedings of the National Academy of Sciences of the United States of America* 111, 8803–8808.
- Gaub, B.M., Berry, M.H., Holt, A.E., *et al.*, 2014. Restoration of visual function by expression of a light-gated mammalian ion channel in retinal ganglion cells or ON-bipolar cells. *Proceedings of the National Academy of Sciences of the United States of America* 111, E5574–E5583.
- Gorostiza, P., Volgraf, M., Numano, R., *et al.*, 2007. Mechanisms of photoswitch conjugation and light activation of an ionotropic glutamate receptor. *Proceedings of the National Academy of Sciences of the United States of America* 104, 10865–10870.
- Grusch, M., Schelch, K., Riedler, R., *et al.*, 2014. Spatio-temporally precise activation of engineered receptor tyrosine kinases by light. *EMBO Journal* 33, 1713–1726.
- Gunaydin, L.A., Yizhar, O., Berndt, A., *et al.*, 2010. Ultrafast optogenetic control. *Nature Neuroscience* 13, 387–392.
- Han, X., Boyden, E.S., 2007. Multiple-color optical activation, silencing, and desynchronization of neural activity, with single-spike temporal resolution. *PLoS One* 2, e299.
- Harper, S.M., Neil, L.C., Gardner, K.H., 2003. Structural basis of a phototropin light switch. *Science* 301, 1541–1544.
- Hou, Q., Gilbert, J., Man, H.Y., 2011. Homeostatic regulation of AMPA receptor trafficking and degradation by light-controlled single-synaptic activation. *Neuron* 72, 806–818.
- Ishizuka, T., Kakuda, M., Araki, R., Yawo, H., 2006. Kinetic evaluation of photosensitivity in genetically engineered neurons expressing green algae light-gated channels. *Neuroscience Research* 54, 85–94.
- Izquierdo-Serra, M., Gascón-Moya, M., Hirtz, J.J., *et al.*, 2014. Two-photon neuronal and astrocytic stimulation with azobenzene-based photoswitches. *Journal of the American Chemical Society* 136, 8693–8701.
- Izquierdo-Serra, M., Trauner, D., Llobet, A., Gorostiza, P., 2013. Optical control of calcium-regulated exocytosis. *Biochimica et Biophysica Acta* 1830, 2853–2860.
- Janovjak, H., Szobota, S., Wyart, C., Trauner, D., Isacoff, E.Y., 2010. A light-gated, potassium-selective glutamate receptor for the optical inhibition of neuronal firing. *Nature Neuroscience* 13, 1027–1032.
- Jansen, V., Alvarez, L., Balbach, M., *et al.*, 2015. Controlling fertilization and cAMP signaling in sperm by optogenetics. *eLife* 4, e05161.
- Kato, H.E., Zhang, F., Yizhar, O., *et al.*, 2012. Crystal structure of the channelrhodopsin light-gated cation channel. *Nature* 482, 369–374.
- Kauwe, G., Isacoff, E.Y., 2013. Rapid feedback regulation of synaptic efficacy during high-frequency activity at the *Drosophila* larval neuromuscular junction. *Proceedings of the National Academy of Sciences of the United States of America* 110, 9142–9147.
- Kennedy, M.J., Hughes, R.M., Peteya, L.A., *et al.*, 2010. Rapid blue-light-mediated induction of protein interactions in living cells. *Nature Methods* 7, 973–975.
- Kienzler, M.A., Reiner, A., Trautman, E., *et al.*, 2013. A red-shifted, fast-relaxing azobenzene photoswitch for visible light control of an ionotropic glutamate receptor. *Journal of the American Chemical Society* 135, 17683–17686.
- Kim, J.M., Hwa, J., Garriga, P., *et al.*, 2005. Light-driven activation of beta 2-adrenergic receptor signaling by a chimeric rhodopsin containing the beta 2-adrenergic receptor cytoplasmic loops. *Biochemistry* 44, 2284–2292.
- Klapoetke, N.C., Murata, Y., Kim, S.S., *et al.*, 2014. Independent optical excitation of distinct neural populations. *Nature Methods* 11, 338–346.
- Kleinlogel, S., Feldbauer, K., Dempksi, R.E., *et al.*, 2011. Ultra light-sensitive and fast neuronal activation with the Ca²⁺-permeable channelrhodopsin CatCh. *Nature Neuroscience* 14, 513–518.
- Kramer, R.H., Mourot, A., Adesnik, H., 2013. Optogenetic pharmacology for control of native neuronal signaling proteins. *Nature Neuroscience* 16, 816–823.
- Lee, J., Natarajan, M., Nashine, V.C., *et al.*, 2008. Surface sites for engineering allosteric control in proteins. *Science* 322, 438–442.
- Lee, J.H., Durand, R., Gradinaru, V., *et al.*, 2010. Global and local fMRI signals driven by neurons defined optogenetically by type and wiring. *Nature* 465, 788–792.
- Lemoine, D., Habermacher, C., Martz, A., *et al.*, 2013. Optical control of an ion channel gate. *Proceedings of the National Academy of Sciences of the United States of America* 110, 20813–20818.
- Lester, H.A., Krouse, M.E., Nass, M.M., Wassermann, N.H., Erlanger, B.F., 1980. A covalently bound photoisomerizable agonist: comparison with reversibly bound agonists at Electrophorus electroplaques. *Journal of General Physiology* 75, 207–232.
- Levitz, J., Pantoja, C., Gaub, B., *et al.*, 2013. Optical control of metabotropic glutamate receptors. *Nature Neuroscience* 16, 507–516.
- Li, D., Héroult, K., Isacoff, E.Y., Oheim, M., Ropert, N., 2012. Optogenetic activation of LiGluR-expressing astrocytes evokes anion channel-mediated glutamate release. *Journal of Physiology* 590, 855–873.
- Li, X., Gutierrez, D.V., Hanson, M.G., *et al.*, 2005. Fast noninvasive activation and inhibition of neural and network activity by vertebrate rhodopsin and green algae channelrhodopsin. *Proceedings of the National Academy of Sciences of the United States of America* 102, 17816–17821.
- Lin, J.Y., Knutsen, P.M., Muller, A., Kleinfeld, D., Tsien, R.Y., 2013. ReaChR: A red-shifted variant of channelrhodopsin enables deep transcranial optogenetic excitation. *Nature Neuroscience* 16, 1499–1508.
- Lin, J.Y., Lin, M.Z., Steinbach, P., Tsien, R.Y., 2009. Characterization of engineered channelrhodopsin variants with improved properties and kinetics. *Biophysical Journal* 96, 1803–1814.
- Lin, W.C., Davenport, C.M., Mourot, A., *et al.*, 2014. Engineering a light-regulated GABAA receptor for optical control of neural inhibition. *ACS Chemical Biology* 9, 1414–1419.
- Madisen, L., Mao, T., Koch, H., *et al.*, 2012. A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. *Nature Neuroscience* 15, 793–802.
- Masseck, O.A., Rubelowski, J.M., Spoida, K., Herlitze, S., 2011. Light- and drug-activated G-protein-coupled receptors to control intracellular signaling. *Experimental Physiology* 96, 51–56.
- Masseck, O.A., Spoida, K., Dalkara, D., *et al.*, 2014. Vertebrate cone opsins enable sustained and highly sensitive rapid control of Gi/o signaling in anxiety circuitry. *Neuron* 81, 1263–1273.

- Mattis, J., Tye, K.M., Ferenczi, E.A., *et al.*, 2012. Principles for applying optogenetic tools derived from direct comparative analysis of microbial opsins. *Nature Methods* 9, 159–172.
- Miesenböck, G., 2009. The optogenetic catechism. *Science* 326, 395–399.
- Miesenböck, G., 2011. Optogenetic control of cells and circuits. *Annual Review of Cell and Developmental Biology* 27, 731–758.
- Möglich, A., Moffat, K., 2010. Engineered photoreceptors as novel optogenetic tools. *Photochemical & Photobiological Sciences* 9, 1286–1300.
- Möglich, A., Yang, X., Ayers, R.A., Moffat, K., 2010. Structure and function of plant photoreceptors. *Annual Review of Plant Biology* 61, 21–47.
- Nagel, G., Brauner, M., Liewald, J.F., *et al.*, 2005. Light activation of channelrhodopsin-2 in excitable cells of *Caenorhabditis elegans* triggers rapid behavioral responses. *Current Biology* 15, 2279–2284.
- Nagel, G., Ollig, D., Fuhrmann, M., *et al.*, 2002. Channelrhodopsin-1: A light-gated proton channel in green algae. *Science* 296, 2395–2398.
- Nagel, G., Szellas, T., Huhn, W., *et al.*, 2003. Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proceedings of the National Academy of Sciences of the United States of America* 100, 13940–13945.
- Numano, R., Szobota, S., Lau, A.Y., *et al.*, 2009. Nanosculpting reversed wavelength sensitivity into a photoswitchable iGluR. *Proceedings of the National Academy of Sciences of the United States of America* 106, 6814–6819.
- Oh, E., Maejima, T., Liu, C., Deneris, E., Herlitze, S., 2010. Substitution of 5-HT1A receptor signaling by a light-activated G protein-coupled receptor. *Journal of Biological Chemistry* 285, 30825–30836.
- Palczewski, K., 2006. G protein-coupled receptor rhodopsin. *Annual Review of Biochemistry* 75, 743–767.
- Papagiakoumou, E., Anselmi, F., Begue, A., *et al.*, 2010. Scanless two-photon excitation of channelrhodopsin-2. *Nature Methods* 7, 848–854.
- Pathak, G.P., Vrana, J.D., Tucker, C.L., 2013. Optogenetic control of cell function using engineered photoreceptors. *Biology of the Cell* 105, 59–72.
- Raffelberg, S., Wang, L., Gao, S., *et al.*, 2013. A LOV-domain-mediated blue-light-activated adenylate (adenylyl) cyclase from the cyanobacterium *Microcoleus chthonoplastes* PCC 7420. *Biochemical Journal* 455, 359–365.
- Reiner, A., Isacoff, E.Y., 2013. The Brain Prize 2013: The optogenetics revolution. *Trends in Neurosciences* 36, 557–560.
- Reiner, A., Isacoff, E.Y., 2014. Tethered ligands reveal glutamate receptor desensitization depends on subunit occupancy. *Nature Chemical Biology* 10, 273–280.
- Reiner, A., Levitz, J., Isacoff, E.Y., 2015. Controlling ionotropic and metabotropic glutamate receptors with light: Principles and potential. *Current Opinion in Pharmacology* 20, 135–143.
- Rullo, A., Reiner, A., Reiter, A., *et al.*, 2014. Long wavelength optical control of glutamate receptor ion channels using a tetra-*ortho*-substituted azobenzene derivative. *Chemical Communications* 50, 14613–14615.
- Sandoz, G., Levitz, J., 2013. Optogenetic techniques for the study of native potassium channels. *Frontiers in Molecular Neuroscience* 6, 6.
- Sandoz, G., Levitz, J., Kramer, R.H., Isacoff, E.Y., 2012. Optical control of endogenous proteins with a photoswitchable conditional subunit reveals a role for TREK1 in GABA(B) signaling. *Neuron* 74, 1005–1014.
- Schröder-Lang, S., Schwärzel, M., Seifert, R., *et al.*, 2007. Fast manipulation of cellular cAMP level by light *in vivo*. *Nature Methods* 4, 39–42.
- Shimizu-Sato, S., Huq, E., Tepperman, J.M., Quail, P.H., 2002. A light-switchable gene promoter system. *Nature Biotechnology* 20, 1041–1044.
- Spöida, K., Masseck, O.A., Deneris, E.S., Herlitze, S., 2014. Gq/5-HT2c receptor signals activate a local GABAergic inhibitory feedback circuit to modulate serotonergic firing and anxiety in mice. *Proceedings of the National Academy of Sciences of the United States of America* 111, 6479–6484.
- Stierl, M., Stumpf, P., Udvari, D., *et al.*, 2011. Light modulation of cellular cAMP by a small bacterial photoactivated adenylyl cyclase, bPAC, of the soil bacterium *Beggiatoa*. *Journal of Biological Chemistry* 286, 1181–1188.
- Strickland, D., Moffat, K., Sosnick, T.R., 2008. Light-activated DNA binding in a designed allosteric protein. *Proceedings of the National Academy of Sciences of the United States of America* 105, 10709–10714.
- Szobota, S., Gorostiza, P., Del Bene, F., *et al.*, 2007. Remote control of neuronal activity with a light-gated glutamate receptor. *Neuron* 54, 535–545.
- Szymanski, W., Beierle, J.M., Kistemaker, H.A., Velema, W.A., Feringa, B.L., 2013. Reversible photocontrol of biological systems by the incorporation of molecular photoswitches. *Chemical Reviews* 113, 6114–6178.
- Tochitsky, I., Banghart, M.R., Mouro, A., *et al.*, 2012. Optochemical control of genetically engineered neuronal nicotinic acetylcholine receptors. *Nature Chemistry* 4, 105–111.
- Tye, K.M., Deisseroth, K., 2012. Optogenetic investigation of neural circuits underlying brain disease in animal models. *Nature Reviews Neuroscience* 13, 251–266.
- Vaziri, A., Emiliani, V., 2012. Reshaping the optical dimension in optogenetics. *Current Opinion in Neurobiology* 22, 128–137.
- Volgraf, M., Gorostiza, P., Numano, R., *et al.*, 2006. Allosteric control of an ionotropic glutamate receptor with an optical switch. *Nature Chemical Biology* 2, 47–52.
- Watanabe, S., Rost, B.R., Camacho-Perez, M., *et al.*, 2013. Ultrafast endocytosis at mouse hippocampal synapses. *Nature* 504, 242–247.
- Wietek, J., Wiegert, J.S., Adeishvili, N., *et al.*, 2014. Conversion of channelrhodopsin into a light-gated chloride channel. *Science* 344, 409–412.
- Wu, Y.I., Frey, D., Lungu, O.I., *et al.*, 2009. A genetically encoded photoactivatable Rac controls the motility of living cells. *Nature* 461, 104–108.
- Wyart, C., Del Bene, F., Warp, E., *et al.*, 2009. Optogenetic dissection of a behavioral module in the vertebrate spinal cord. *Nature* 461, 407–410.
- Zemelman, B.V., Lee, G.A., Ng, M., Miesenböck, G., 2002. Selective photostimulation of genetically chARGed neurons. *Neuron* 33, 15–22.
- Zemelman, B.V., Miesenböck, G., 2001. Genetic schemes and schemata in neurophysiology. *Current Opinion in Neurobiology* 11, 409–414.
- Zhang, F., Wang, L.P., Brauner, M., *et al.*, 2007. Multimodal fast optical interrogation of neural circuitry. *Nature* 446, 633–639.

Further Reading

- Reiner, A., Isacoff, E.Y., 2014. Photoswitching of cell surface receptors using tethered ligands. *Methods in Molecular Biology* 1148, 45–68.
- Zhang, F., Gradinaru, V., Adamantidis, A.R., *et al.*, 2010. Optogenetic interrogation of neural circuits: Technology for probing mammalian brain structures. *Nature Protocols* 5, 439–456.

Relevant Websites

- <https://www.addgene.org/optogenetics>
Addgene Plasmid Repository.
- <http://syntheticneurobiology.org/protocols>
Boyden Lab, Optogenetics Resources.
- <http://web.stanford.edu/group/dlab/optogenetics/index.html>
Deisseroth Lab, Optogenetics Resources.
- <http://mcb.berkeley.edu/labs/isacoff/research2.html>
Isacoff Lab, Chemical Optogenetics Resources.
- <http://www.openoptogenetics.org>
Openoptogenetics Platform.

ORGANELLES: STRUCTURE AND FUNCTION

Contents

The Endoplasmic Reticulum

Intermediate Compartment: A Sorting Station between the Endoplasmic Reticulum and the Golgi Apparatus

Golgi and TGN

Early Endosomal Compartments

The Late Endosome

Signaling from Endosomes

Conventional and Secretory Lysosomes

Lysosome-Related Organelles

At the Center of Autophagy: Autophagosomes

Peroxisomes

Lipid Droplets

Mechanisms and Functions of Mitochondrial Dynamics

Synaptosomes and Synaptic Vesicles

Extracellular Vesicles

Nuclear Organization (Nuclear Structure and Dynamics)

Nuclear Pores

Plastids: The Anabolic Factories of Plant Cells

The Endoplasmic Reticulum

E van Anken and R Sitia, Vita Salute San Raffele University, Milan, Italy

© 2016 Elsevier Inc. All rights reserved.

Introduction

Robert Hooke discovered that tissues consist of cells already in 1665. Fellow microscopists at that time, Van Leeuwenhoek and Bauer, realized that smaller structures were discernable within the cell, in particular the vacuole and the nucleus. Yet, the terminology for these cell compartments stems from the nineteenth century, when several other subcellular structures were discovered. Möbius then coined the name 'organula' – a term that was later simplified to 'organelles' – to describe these subcellular compartments. Just before the turn of the century, Camillo Golgi identified the organelle that still bears his name.

At the very beginning of the twentieth century, a student of Golgi, Emilio Veratti, found yet another subcellular structure in muscle cells, which he claimed to be distinct from the Golgi apparatus (Veratti, 1961). Veratti's work failed to convince his contemporaries, however. About 50 years later, the advent of the electron microscope finally made the unambiguous identification of Veratti's elusive organelle possible when Keith Porter and colleagues observed a network (reticulum) inside the cytoplasm of eukaryotic cells (endoplasmic). He teamed up with George Palade to obtain high-resolution images (Figure 1) and in 1954 the endoplasmic reticulum (ER)

entered the canon of cell biology as a distinct organelle separated from the cytosol by virtue of a membrane and forming a complex tubular or sheet-like network (Palade and Porter, 1954).

The Palade lab then developed a method – which still is widely used today – to obtain ER-enriched cell fractions, called microsomes (Figure 1). The physical sheering of cells causes ER vesiculation and ER-derived vesicles are then purified by differential centrifugation (Palade and Siekevitz, 1956). The combination of microscopy and biochemical assays exploiting microsomes initiated a cascade of discoveries on the function and morphology of the ER and its specialized derivative in muscle cells: the sarcoplasmic reticulum (SR). In the early 1960s, Palade and coworkers proved by pulse chase analysis using radiolabeling that the ER is the first station along the secretory pathway (Siekevitz and Palade, 1960). Günther Blobel (a pupil of Palade) and colleagues elucidated in the 1970s that client proteins enter the ER by virtue of a signal peptide (Blobel and Dobberstein, 1975).

Another milestone in ER research came when genetic screens in yeast, pioneered by Randy Schekman and colleagues (Novick *et al.*, 1980), initiated the characterization of the molecular machinery of the secretory pathway. A long list of the so-called sec proteins was identified, including several that

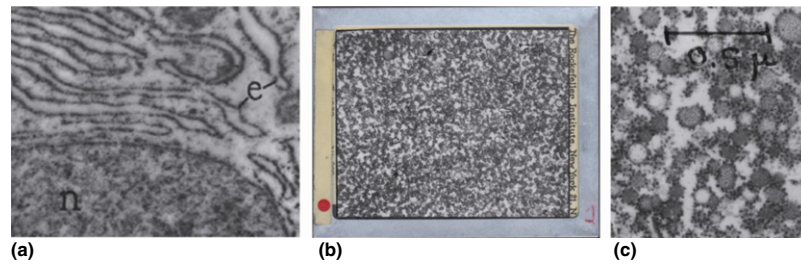


Figure 1 (a) One of the first sightings of the ER by electron microscopy at high resolution (n=nucleus, e=ER). Reproduced from Palade, G.E., Porter, K.R., 1954. Studies on the endoplasmic reticulum. I. Its identification in cells *in situ*. *Journal of Experimental Medicine* 100, 641–656. (b) An original slide produced in the 1960s in the Palade laboratory of an electron micrograph of microsomes, personal collection of Professor Jacopo Meldolesi, a former collaborator of Palade. (c) Magnification of the area around the handwritten size bar of the slide shown in the middle panel; note that the microsomes evidently are ribosome-studded.

are fundamental to ER function. Through the work of Ari Helenius and coworkers it became clear in the late 1980s/early 1990s that stringent quality control is exerted on protein folding in the ER (Hurtley *et al.*, 1989). At around the same time, pioneering work from Kazutoshi Mori and Peter Walter (a pupil of Blobel) initiated research on the unfolded protein response (UPR) (Cox *et al.*, 1993; Mori *et al.*, 1993), which is key for homeostasis of the ER. In this article we summarize the current understanding of this fascinating organelle.

Origin and Function of the ER

The hallmark of eukaryotic cells is the packing of DNA into the nucleus, a compartment that is bounded by the nuclear envelope. The nuclear envelope is continuous with the ER, which suggests that the ER was present in eukaryotic cells right from their ‘inception.’

As the first station along the exocytic pathway, the ER serves as an anticipation of the extracellular world, to prepare molecules to exert their function in a different environment. As such, it is functionally related to the periplasm, the space enclosed between the inner membrane and cell wall of gram-negative bacteria. In analogy, portions of the ER (the so-called cortical ER) are positioned directly below the cell surface also in eukaryotic cells. Both the periplasm and the ER serve to groom proteins that are destined to be part of the cell membranes or to be secreted. Instead of being translocated directly across the plasma membrane, these proteins enter the ER or periplasm first. Here, under the supervision of resident chaperones and enzymes, they attain their native structure before being deployed to the extracellular space. Altogether, these mechanisms warrant efficient production and transport, while simultaneously ensuring the fidelity of the protein secretome, which comprises a third of the eukaryotic proteome, equaling in humans approximately 18 000 gene products (Chen *et al.*, 2005).

Ribosomes that are translating proteins destined for the secretory pathway dock onto the cytosolic face of the ER membrane. These ribosomes give a ‘rough’ appearance to the ER, typical of protein secreting cells. Accordingly, the ribosome-studded parts of the ER therefore are referred to as ‘rough ER,’ while the portion devoid of ribosomes as ‘smooth ER.’ The majority of ‘ER client’ proteins enter co-translationally into the organelle. Some proteins can enter after their synthesis

is completed (posttranslationally). In either case, the proteins traverse the ER membrane through the so-called translocon complex. The latter’s main component, Sec61, has an evolutionary equivalent in gram-negative bacteria (Keenan *et al.*, 2001) that allows proteins to reach the periplasm, which supports the notion that the ER has a periplasmic-like origin.

Next to being the cradle of secretory and cell surface proteins, the ER is the most important site of (membrane) lipid synthesis in the eukaryotic cell (Fagone and Jackowski, 2009). Accordingly, a vast repertoire of enzymes devoted to lipid synthesis resides in the ER. Other organelles depend on membrane lipid supply from the ER, either through vesicular transport or through lipid transfer. The numerous contact points between ER and other organelles likely are sites of such lipid transfer, even though the membranes of the ER and the acceptor organelles are juxtaposed and not continuous. In other organelles the membrane lipids that are derived from the ER often are further derivatized or asymmetrically distributed between membrane leaflets according to need (Vance *et al.*, 1977).

In most eukaryotic cells, the ER also is the major store of intracellular Ca^{2+} ions and most ER resident proteins in fact are low-affinity/high-capacity calcium-binding proteins (Meldolesi and Pozzan, 1998). A wealth of signaling pathways exploits the Ca^{2+} gradient across the ER membrane (Berridge, 2002). Typically, signaling cascades drive the opening of Ca^{2+} channels to sustain a burst of Ca^{2+} through vicinal mitochondria and in the cytosol, which, in turn, serves to propagate the signal. In most cases, there is a rapid reuptake of Ca^{2+} back into the ER, ensuring that Ca^{2+} signaling can start anew. The most sophisticated Ca^{2+} -driven signaling cycles occur at the sarcoplasmic reticulum to sustain muscle contraction (Endo, 1977).

Morphology of the ER

The ER is the most extended organelle in many eukaryotic cells: on average it occupies 10% of the cellular volume, but this figure is larger in plasma cells, exocrine pancreas, and other professional secretory cells. The tubules and sheet-like structures (both rough and smooth) form a continuous membrane system that includes the nuclear envelope (Voeltz and Prinz, 2007). Tubules can be highly curved and branched to form a polygonal network extending throughout the whole

cytosol. ER sheets can be prominent around the nucleus, and may become predominant in some cell types dependent on their function (Figure 2).

Alterations in membrane curvature form the basis of the different morphologies of the ER. For instance, ER tubules are cylindrical with a diameter of 30–100 nm, while sheets are virtually flat. The width of the ER lumen in sheets is restricted to approximately 50–100 nm (Voeltz and Prinz, 2007), but occasionally sac-like structures may develop, bulging from the ER sheet and tubular network, in particular when cells sustain bulk protein secretion or when aberrant proteins aggregate into ER resident deposits (Kopito and Sitia, 2000).

In recent years insights into the mechanisms that are responsible for ER morphology have greatly improved. A number of proteins have been identified that determine its shape. Wedge-like proteins, reticulons and other similar proteins, insert into the ER membrane and dictate the typical curvature found in ER tubules and at the edges of ER sheets, including the nuclear envelope at the sites of nuclear pores (Voeltz *et al.*, 2006). Proteins driving formation of sheet-like structures

include CLIMP-63, a transmembrane protein whose ER luminal domains can homodimerize with another CLIMP-63 molecule protruding into the lumen from the opposing membrane in the ER sheet, and so determining the typical width (Klopfenstein *et al.*, 2001). Other proteins, for example, kinectin and p180, with bulky rod-like structures facing the cytosol seem to be required for stabilizing the flatness of sheets (Lin *et al.*, 2012; Figure 3).

While strictly organized in rough and smooth, tubular, and sheet-like, the ER is highly dynamic at the same time. The network is constantly reshaped through fission and fusion events, in which GTPases of the atlastin and Rab families are involved (Hu *et al.*, 2009; English and Voeltz, 2013). Reshaping of the ER occurs hand in hand with movements of ER elements along the cytoskeleton that are facilitated by motor proteins (Goyal and Blackstone, 2013).

Depending on the cell type and functional requirements, the shape, organization, and size of the ER varies considerably. In muscle cells, for instance, the sarcoplasmic reticulum is a highly organized form of smooth ER devoted to synchronize myofibril contraction (Endo, 1977), while the rough portion

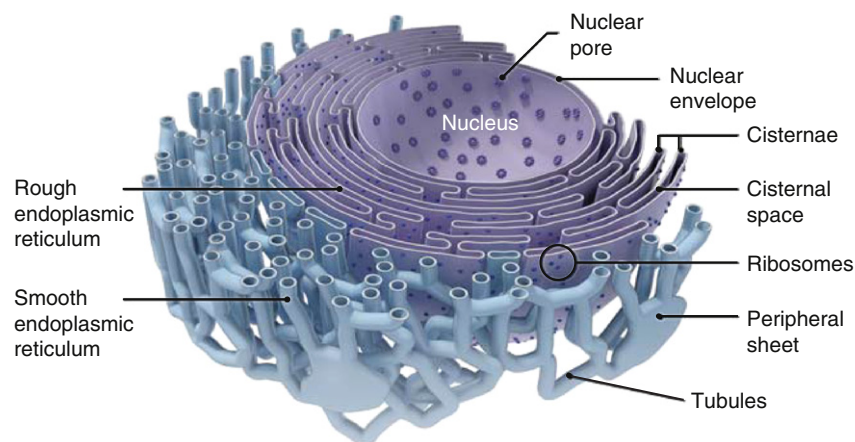


Figure 2 Schematic representation of the ER in animal cells. Reproduced from Goyal, U., Blackstone, C., 2013. Untangling the web: Mechanisms underlying ER network formation. *Biochimica et Biophysica Acta* 1833, 2492–2498.

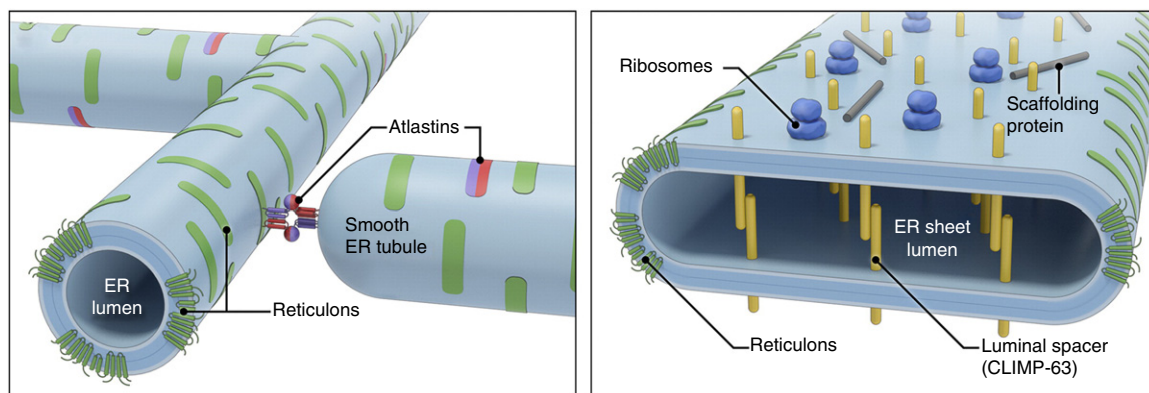


Figure 3 Wedge-shaped proteins such as reticulons facilitate membrane curvature in ER tubules or at edges of sheets. Spacers such as CLIMP-63 maintain thickness of sheets. Reproduced from Goyal, U., Blackstone, C., 2013. Untangling the web: Mechanisms underlying ER network formation. *Biochimica et Biophysica Acta* 1833, 2492–2498.

of the ER is marginal. In liver cells, the smooth ER is also extensively developed and a major site of detoxifying and synthetic enzymes. In professional secretory cells such as thyrocytes that produce thyroglobulin, acinar cells, which produce digestive enzymes, etc. the rough ER instead is pre-eminent. Antibody producing plasma cells are prototypes of cells with an expanded rough ER (Figure 4).

Membrane Contact Sites of the ER with Other Organelles

Subdomains of the ER make close contact with virtually all other cellular structures, foremost with the nucleus, as the nuclear envelope is continuous with the ER. Contacts of ER

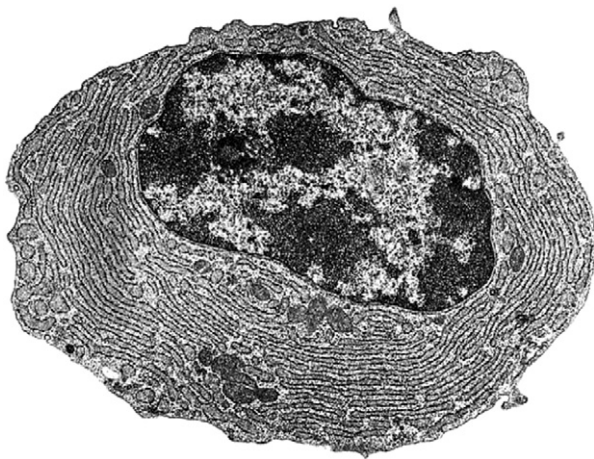


Figure 4 Electron micrograph of a plasma cell, which has an expanded ER to sustain antibody production and secretion. Courtesy of the late Dr. Carlo E. Grossi.

with other organelles are maintained by various tethering complexes that form a bridge between the apposing membranes. ER-plasma membrane contacts are predominant in yeast in which cortical ER represents up to 50% of the organelle. In metazoans ER-plasma membrane contacts are less abundant, but the interaction between STIM in the ER membrane and Orai in the plasma membrane is of particular interest. Together, STIM and Orai mediate store-operated Ca^{2+} entry from the extracellular space across both membranes into the ER lumen (Stefan *et al.*, 2013; Figure 5).

Likewise, protein-mediated tethering of ER to mitochondria sustains close contacts, which are referred to as mitochondria-associated membranes (MAM). In metazoans, these ER-mitochondria contact sites are key for Ca^{2+} transfer and, thus, for signaling: the opening of IP_3 receptor channels at the MAM leads to Ca^{2+} efflux, which is readily imported into the mitochondrial matrix through porins and mitochondrial calcium uniporters (MCUs) to initiate signaling. Excess Ca^{2+} is pumped back from the cytosol into the ER by SERCA (sarco-endoplasmic reticulum Ca^{2+} ATPase) or across the PM by PMCA (plasma membrane Ca^{2+} ATPase) (Figure 5). Ca^{2+} signaling activates some enzymes of the Krebs cycle and, thus, is key for overall metabolism. Yet, unrestrained Ca^{2+} transfer to mitochondria initiates apoptosis (Naon and Scorrano, 2014).

While Ca^{2+} shuttling through ER-mitochondria contacts seems less important in yeast, the contact sites are key for lipid transfer in all eukaryotic cells (Helle *et al.*, 2013). Mitochondria do not synthesize membrane lipids and there is no vesicular transport or any other membrane continuity between mitochondria and the endomembrane system. Thus, mitochondria must be furnished with membrane lipids through lipid transfer proteins. Yet, so far it remains unclear how the lipid transfer process at the MAM is organized.

The ER establishes tether-mediated contacts also with the Golgi, lysosome, and endosomes. Their main function seems

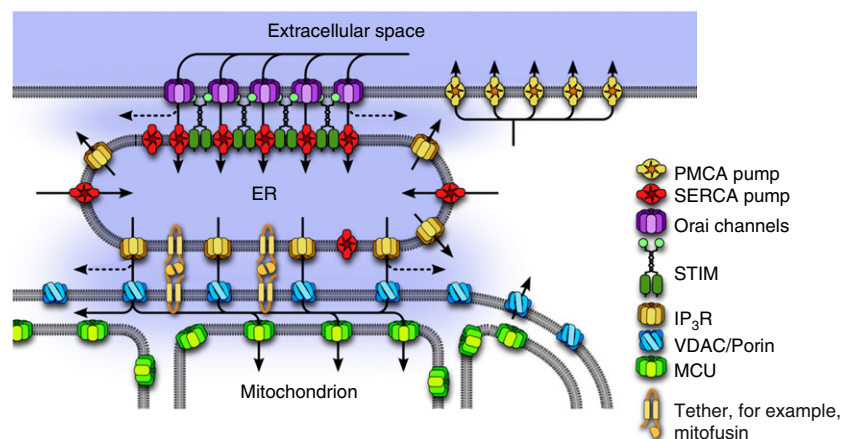


Figure 5 Organization of Ca^{2+} signaling through ER-PM and ER-mitochondria contact sites. Blue represents the local $[\text{Ca}^{2+}]$. A Ca^{2+} gradient is established by SERCA (red) and PMCA (yellow) pumps that pump Ca^{2+} ions from the cytosol into the ER or, respectively, into the extracellular space. ER-PM contacts facilitate replenishing of Ca^{2+} stores in the ER from the extracellular space. STIM (dark green) tethers the ER membrane to the PM at sites where Orai channels are, such that little Ca^{2+} leaks to the cytosol proper. Ca^{2+} is released from the ER through IP_3R channels (tan) mostly at sites, where tethers such as mitofusin (orange) maintain close ER-mitochondria contacts. Ca^{2+} is readily taken up by mitochondria via VDAC (Voltage-dependent anion channels)/Porin (blue) and MCU (light green) channels. Adapted from Helle, S.C., Kanfer, G., Kolar, K., *et al.*, 2013. Organization and function of membrane contact sites. *Biochimica et Biophysica Acta* 1833, 2526–2541.

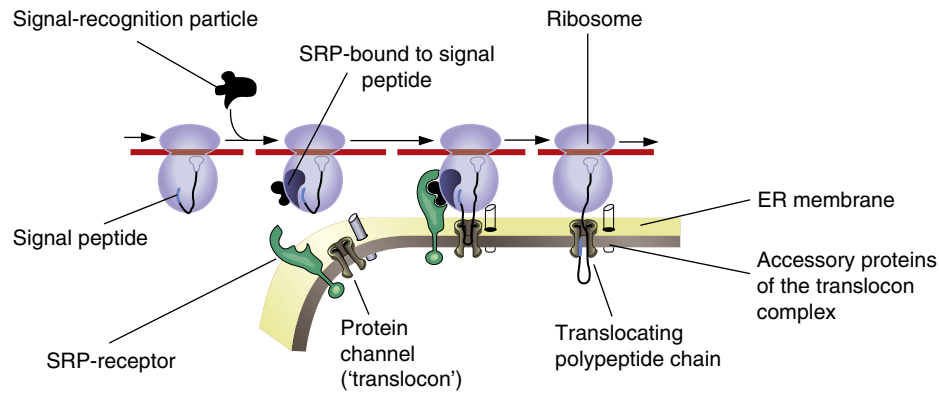


Figure 6 Proteins destined for the secretory pathway display an N-terminal signal peptide, which is recognized by SRP. The ternary complex of the ribosome, the nascent chain, and SRP is recognized by the SRP receptor and docks onto the translocon, whereupon co-translational translocation into the ER begins.

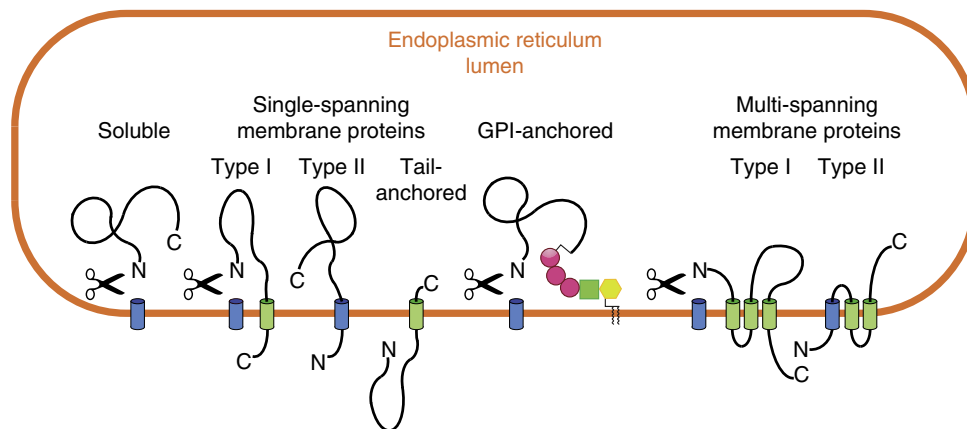


Figure 7 Topology of ER client proteins. From many ER client proteins the signal peptide (light blue) is removed by signal peptidase (scissors). If the protein has no other hydrophobic stretch, it will be a soluble protein. If the protein has one or more hydrophobic stretches (green) it will be membrane embedded. Tail-anchored proteins have negligible ectodomains at the C-terminus. Some proteins obtain a GPI-anchor, consisting of lipid moieties (wiggled lines) of phosphoinositol (yellow hexagon) that allow membrane embedding coupled via glucosamine (green square), three mannose moieties (pink) and phosphoethanolamine to the polypeptide chain.

to be the transfer from the ER of specific lipids that are in shortage in the target membranes (Helle *et al.*, 2013). Overall storage of lipids takes place in lipid droplets, and it comes as no surprise that lipid droplets are intimately linked with the ER as well (Xu *et al.*, 2012). The ER moreover is the source of membrane lipids for the peroxisome. While *de novo* peroxisome formation initiates with pre-peroxisomal structures derived from the ER (van der Zand *et al.*, 2012), it is unclear whether mature peroxisomes require lipid transfer from the ER through contact sites. Finally, it is still highly debated what is the source of autophagic membranes, which also may turn out to originate either directly or indirectly from the ER.

Translocation and Membrane Embedding of ER Client Proteins

Most client proteins enter the ER co-translationally by virtue of a hydrophobic signal peptide generally located at their

N-termini (Blobel and Dobberstein, 1975). As such, the signal peptide emerges first from the ribosome during translation, whereupon it binds to the signal recognition particle (SRP) (Walter *et al.*, 1981). SRP binding temporarily arrests elongation of the nascent polypeptide and guides the translating ribosome to the SRP receptor, which is associated with the translocon complex (Gilmore *et al.*, 1982). The net result is that translation is now coupled to translocation (Figure 6). The hydrophobic signal peptide is laterally released from the translocon into the ER membrane, while the remainder of the ER client nascent chain reels into the ER lumen with the assistance of resident chaperones (Tyson and Stirling, 2000). Often the signal peptide is then removed by signal peptidase (Blobel and Dobberstein, 1975), but when it is not cleaved off, the signal peptide serves as a membrane anchor (Figure 7).

Translocation continues unless the translocon encounters another hydrophobic stretch in the nascent chain. In that case, again, elongation slows down as the newly made hydrophobic segment is laterally released into the ER membrane to serve as

a transmembrane domain. The remainder of the nascent chain will then be cytosolic unless the translocon encounters yet again a hydrophobic stretch, in which case this segment also will be embedded into the ER membrane. The next portion of the nascent ER client will then enter the ER lumen again, etc. As such, multiple membrane spanning proteins are 'stitched' into the ER membrane and the topology is determined of cytosolic domains, transmembrane domains, and the so-called ectodomains inside the ER lumen. Depending on the 'topogenesis' membrane spanning proteins can be oriented with their N-terminus inside or outside the ER lumen. The first are referred to as type I membrane proteins and the latter as type II (Goder and Spiess, 2001; Figure 7).

A few examples are known of small proteins that traverse the translocon when they are fully synthesized, including the honeybee venom precursor (Zimmermann and Mollay, 1986). To sustain such posttranslational translocation, dedicated accessory proteins of the translocon complex, including Sec63, are required (Lang *et al.*, 2012). Also tail-anchored proteins are posttranslationally inserted into the ER membrane, because the hydrophobic domains that will serve as membrane anchors are at the C-terminus, and thus emerge from the ribosome only once these proteins are already fully synthesized (Figure 7). For tail-anchored protein insertion into the ER membrane TRC-40 (Get3 in yeast) is key to shuttle them from the cytosol to a receptor complex/insertion machinery in the ER membrane, consisting of WRB and CAML (Get1/Get2 in yeast) (Schuldiner *et al.*, 2008; Wang *et al.*, 2014).

Posttranslational Modifications of ER Client Proteins

Besides proteolytic cleavage of the signal peptide, three other common posttranslational modifications of client proteins (or rather their ectodomains) are special to the ER: N-linked glycosylation, insertion of disulfide bonds, and addition of glycolipid tail anchors. Client proteins may obtain in the ER a glycosylphosphatidylinositol (GPI) anchor covalently attached to their C-termini (Udenfriend and Kodukula, 1995). This glycolipid moiety is another means to ensure that the client protein remains associated with the membrane (Figure 7).

Like in the bacterial periplasm, free cysteine residues of client proteins are oxidized in the ER such that two cysteine residues are covalently linked to each other (Figure 8). Disulfide bonds can be formed within (intrachain disulfides) or between (interchain disulfides) client proteins. The latter are important in stabilizing multimeric proteins. Tightly regulated disulfide interchange reactions, driven by protein relays, regulate oxidative protein folding, a process intimately linked to redox signaling, in that reactive oxygen species can be generated as byproducts (Tu and Weissman, 2004).

ER client proteins moreover can be glycosylated when an asparagine (N) residue is followed by any residue but proline (X) and then a serine or threonine (NXS/T). The glycan, consisting of three glucoses, nine mannoses, and two *N*-acetyl glucosamines (Figure 9), is transferred en bloc from dolichol precursors. Glycosylation occurs most often co-translationally, as soon as a site emerges in the ER lumen and is recognized by oligosaccharyl transferase (Helenius and Aebi, 2004). Not all potential acceptor sites are always used, however. In the prion

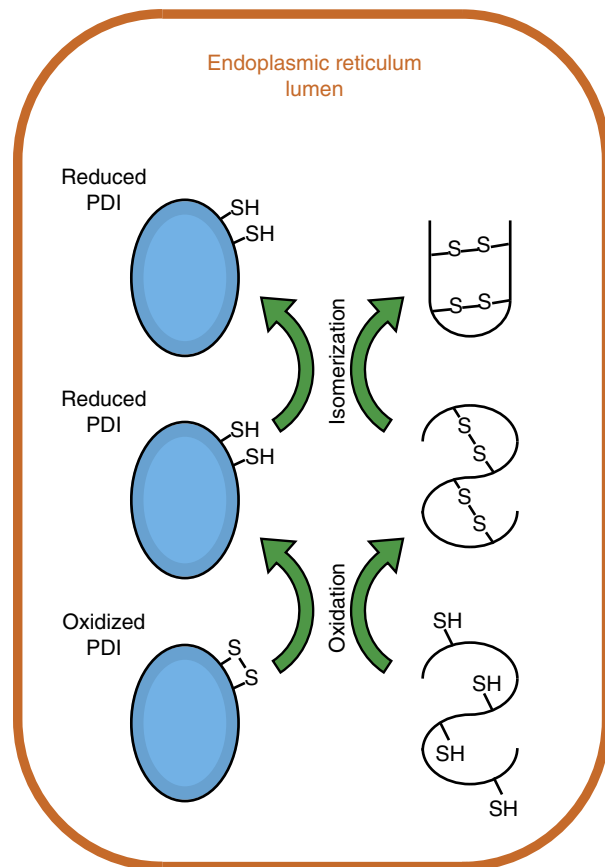


Figure 8 ER client proteins obtain disulfide bonds. PDI or PDI-like proteins catalyze the reaction, but also rearrange aberrant disulfide bonds.

protein, for instance, glycosylation competes with disulfide bonding and GPI addition, which increases the number of isoforms that can be produced (Orsi and Sitia, 2007). Soon after their en bloc addition to proteins in the ER, N-linked glycans are sequentially trimmed by resident glucosidases and mannosidases (Ellgaard and Helenius, 2003), providing important cues for quality control (see below).

Several other posttranslational modifications can occur in the ER. Some ER client proteins obtain mannose moieties that are covalently linked to serine or threonine residues. Such O-mannosylation occurs predominantly in serine/threonine-rich domains of ER client proteins (Praisman and Wells, 2014). Relevant to human pathology is hydroxylation of proline and lysine residues in collagens. The responsible ER resident hydroxylases use vitamin C as an obligatory cofactor, which nicely explains the pathogenesis of scurvy.

Protein Folding in the ER

Newly synthesized proteins progressively adopt conformations that are energetically more favorable, until the correctly folded 'native' conformation is reached. Anfinsen's axiom that folding information is contained in the protein sequence undoubtedly holds true (Anfinsen, 1973). Yet, metastable nonnative

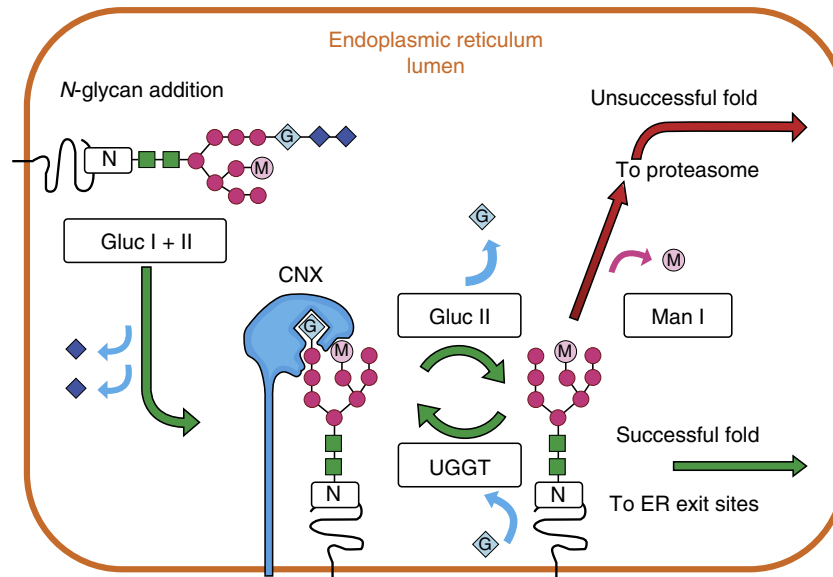


Figure 9 ER client proteins obtain *N*-glycans, consisting of 2 GlcNAc moieties (green squares), 9 mannose moieties (pink circles), and 3 glucose moieties (blue diamonds). Glucoses are trimmed by glucosidases I and II and monoglucosylated glycans allow association with lectin chaperones – calnexin (CNX) is shown here – Gluc II then removes the last glucose and the ER client will either go to ER exit sites (if successfully folded) or it will re-associate with CNX, because UGGT re-added a glucose. Repetitive (de-)glucosylation cycles may occur until mannose trimming leads to association with lectins like EDEM that target the client for ERAD by the proteasome. Adapted from Mitchell, W.B., Li, J., French, D.L., Collier, B.S., 2006. α IIb β 3 biogenesis is controlled by engagement of α IIb in the calnexin cycle via the N15-linked glycan. *Blood* 107, 2713–2719.

conformations may be adopted, often hindering productive completion of the folding process. This phenomenon is particularly evident when folding goes hand in hand with disulfide bond formation. Not surprisingly, therefore, Anfinsen identified protein disulfide isomerase (PDI) as a crucial catalyst through his pioneering work on RNase folding (Gold-enberger *et al.*, 1963). PDI can reshuffle disulfide bonds such that client proteins efficiently attain their native structure (Figure 8). The folding of ER client proteins often is particularly complex and, hence, the ER hosts an impressive repertoire of molecular chaperones and foldases.

Classical molecular chaperones in the ER are functionally similar to counterparts in other folding environments such as the HSP70 family member BiP (Kar2 in yeast). Similar to other HSP70s in the cytosol, mitochondria, etc., BiP transiently binds hydrophobic patches of ER client proteins and thereby protects them from undergoing inappropriate interactions with other proteins (Mayer and Bukau, 2005). BiP is an ATPase and cycles of client binding and release require ATP hydrolysis. Several J-domain containing co-chaperones (ERdj1–5), but also nucleotide exchange factors, modulate BiP's activity and help determine its clientele (Shen *et al.*, 2002; Steel *et al.*, 2004). A distant HSP70-like relative in the metazoan ER is GRP170 that often acts in tandem with BiP, most notably assisting translocation of client proteins (Steel *et al.*, 2004). Similarly, GRP94, the HSP90 equivalent in the metazoan ER, guides maturation of a limited repertoire of client proteins likely with assistance from co-chaperones of the CNPY family (Eletto *et al.*, 2010; Liu *et al.*, 2010). Surprisingly perhaps, the ER lacks members of small HSP (although they are present in plant ER) and the HSP60 families of chaperones.

Foldases are enzymes that catalyze folding steps. For the folding of many client proteins *cis/trans*-isomerization of

proline residues is the rate-limiting step. To smoothen the process of prolyl isomerization, the ER harbors a number of peptidylprolyl isomerases (PPIases), similar to other folding compartments. The most prominent PPIase in the ER is cyclophilin B, but also a few members of the FKBP family are present (Galat, 2003).

Unique to the ER are foldases of the PDI family that catalyze formation of disulfide bonds (Figure 8). Dependent on the cell type, various PDI family members are present, each likely with (slightly) distinct oxidative properties and clientele. Notable PDI family members include Erp57, Erp72, and P5 (Appenzeller-Herzog and Ellgaard, 2008). As in the case of RNase, PDIs can reshuffle disulfide bonds, exchanging aberrant for native disulfide bonds. The *de novo* disulfide bond formation requires transfer of disulfide bonds from PDIs to client proteins. This process entails that PDIs are constantly reoxidized. In yeast, reoxidation of PDI is the exclusive task of Ero1, which uses O₂ as terminal electron acceptor, thereby generating H₂O₂ (Tu and Weissman, 2004). In metazoans, ERO1 proteins share the burden of generating reoxidizing equivalents to sustain disulfide bond formation with peroxiredoxin IV, glutathione peroxidases 7 and 8, and vitamin K epoxide reductase (Bulleid and Ellgaard, 2011).

The glycans play a unique role in the folding process of ER client proteins. Once the first two glucose moieties are removed by glucosidases, lectin chaperones of the calnexin/calreticulin (CNX/CRT) family associate with the monoglucosylated glycans to promote folding. They release the client when also the final glucose is removed, but if the folding is incomplete, the glycan becomes reglucosylated by UGGT allowing another cycle of CNX/CRT association (Helenius and Aebi, 2004; Figure 9). Also the mannose moieties are progressively trimmed until the glycan structure is eventually recognized by lectins of the EDEM

and OS9 families and dispatched to the cytosol for proteasomal degradation (Hosokawa *et al.*, 2001; Christianson *et al.*, 2008). In this way, molecules that did not have enough time to fold are spared from degradation, while those client proteins that linger in the ER for too long because they fail to fold are retro-translocated to the cytosol. Here, they are deglycosylated, ubiquitinated, and, finally, degraded by the proteasome, a process referred to as ER-associated degradation (ERAD) (Werner *et al.*, 1996). Rarely, misfolded clients may escape ERAD and accumulate instead in the ER, sometimes even forming aggregates. Autophagy may serve as a backup removal mechanism. Yet, persistent accumulation of misfolded ER clients often leads to ER storage diseases (Anelli and Sitia, 2010).

In general, however, the combined efforts of the ER resident chaperones and foldases promote productive folding, and, at the same time, ensure through their persistent binding

that immature clients are retained in the ER. Only when clients are fully folded (and properly assembled in case of multimeric proteins) do all chaperones release and allow exit of clients to travel further along the secretory pathway. As such, stringent quality control is exerted on the secretory proteome (Ellgaard and Helenius, 2003; Anelli and Sitia, 2008).

Export from and Retrieval to the ER

Once client proteins are correctly folded, they are sorted to ER exit sites. Some clients team up with dedicated cargo receptors to promote this sorting process. Notable examples include the lectin ER–Golgi intermediate compartment (ERGIC-53) (Appenzeller *et al.*, 1999). At ER exit sites membrane curvature is promoted through the buildup of a coat. Key players in this

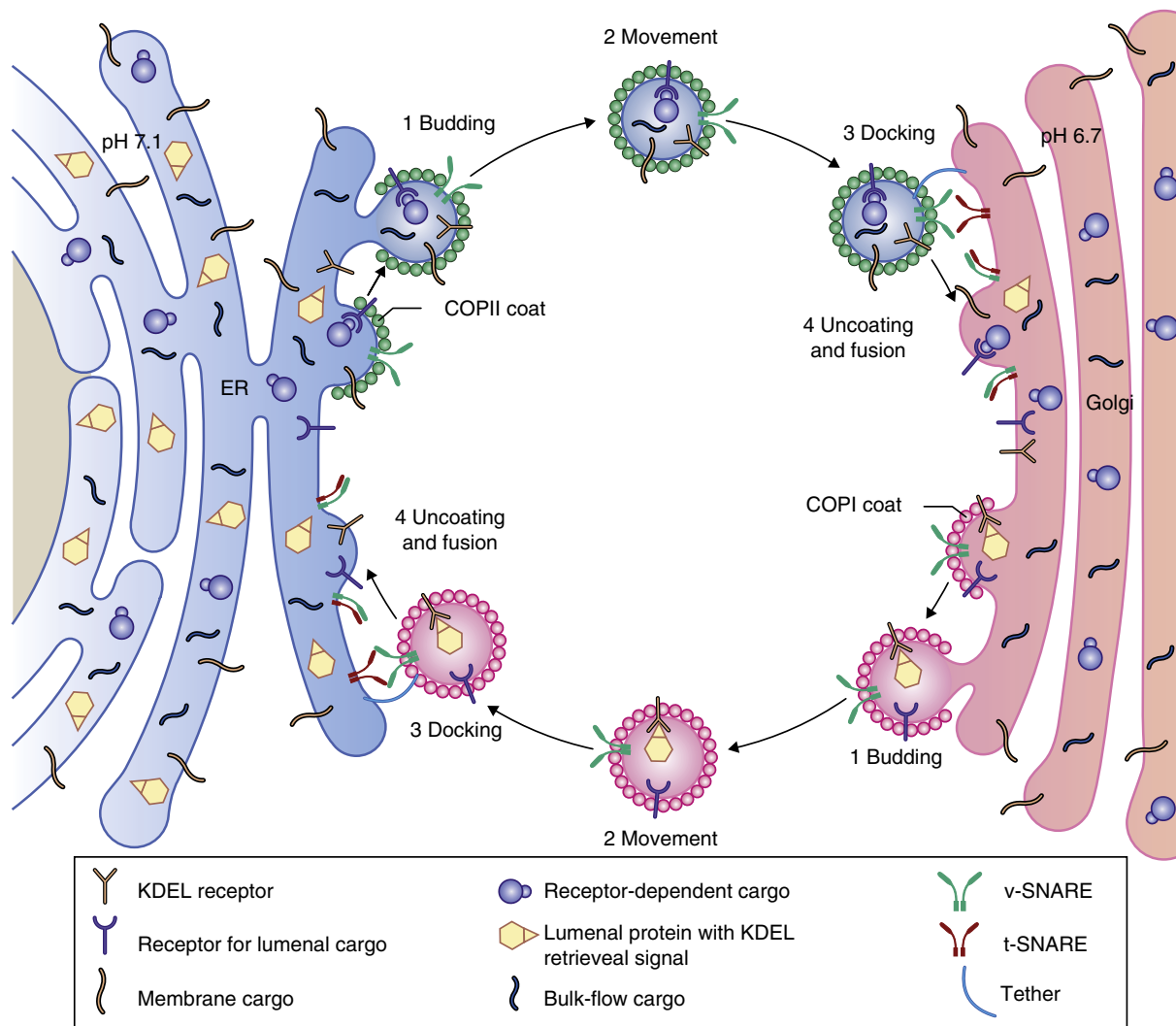


Figure 10 ER export and retrieval. Fully folded ER client proteins exit from the ER at ER exit sites, where they are packaged into COPII coated vesicles. Some clients are specifically recruited for exit by cargo receptors, while other clients enter these vesicles by bulk flow. The COPII vesicles bud from ER exit sites and travel through anterograde transport to the cis-Golgi. The vesicles dock through tethering of the two membranes and fuse through the action of SNAREs on both vesicle (v) and target (t) membranes. The cargo is released into the Golgi and the COPII coat dissociates. Retrograde transport is facilitated by COPI coated vesicles, that let cargo receptors and KDEL receptors travel back to the ER. KDEL receptors pick up ER chaperones and folding enzymes that display a KDEL (-like) tetrapeptide at their C-terminus.

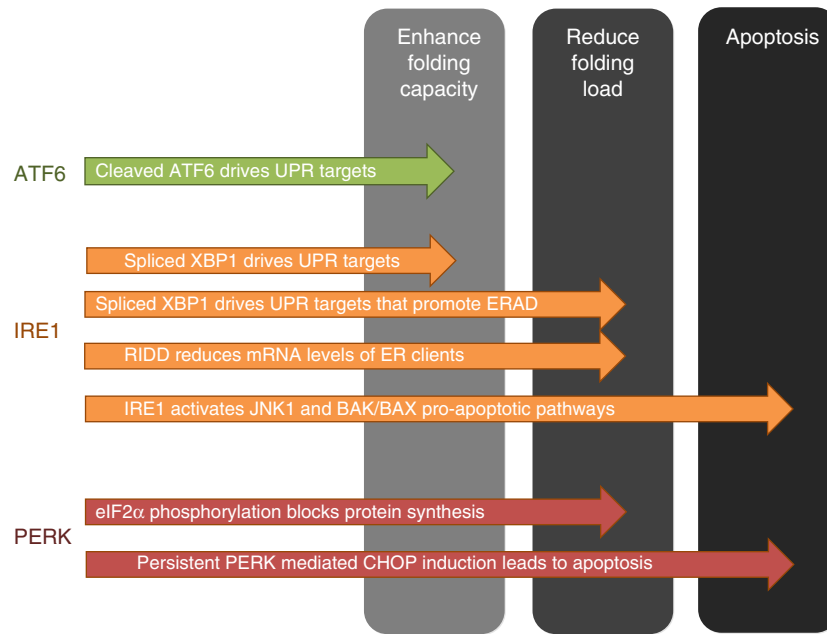


Figure 11 Overview of how UPR pathways aim for ER homeostasis or give way to apoptosis.

process are the components of coat protein complex II (COPII) and the GTPase Sar1 (Barlowe *et al.*, 1994). GTP hydrolysis guides both the installment of the coat and the eventual budding from the ER of COPII-coated vesicles that contain the client proteins (Kuehn *et al.*, 1998; Figure 10). The vesicles then undergo anterograde transport along the secretory pathway, traveling via the ERGIC, the Golgi apparatus, the *trans*-Golgi network before reaching the plasma membrane as the final destination. Through the endocytic pathway, i.e., via endosomes to lysosomes, proteins may travel back into the cell. The secretory and endocytic pathways therefore form a continuous endomembrane system that is interconnected through vesicular transport. Accordingly, almost all proteins of the endomembrane system originate from the ER.

Specific SNARE proteins mark transport vesicles to define their origin and they team up selectively with those SNARE proteins that mark the designated target membrane. Rab family members further assist in ensuring proper delivery of vesicles. The SNAREs on the two membranes combined drive fusion of the vesicle with the target organelle (ERGIC or Golgi) and release of the client protein cargo into the next station along the secretory pathway (Sudhof and Rothman, 2009; Figure 10). While most ER clients that are subunits of multimeric complexes need to oligomerize first in order to exit the ER, some assembly steps may take place in the ERGIC, as is the case for oligomerization of immunoglobulin M (Anelli *et al.*, 2007). In principle, folding and assembly are thus fully completed when clients arrive in the Golgi. In downstream compartments of the secretory pathway, N-linked glycans often are further modified in several ways, yielding varied complex glycan structures. In the *trans*-Golgi network, some clients are even proteolytically cleaved by furin(-like) endoproteases to reach their mature state.

If immature client proteins untimely exit the ER, they usually are retrieved from further stations along the secretory pathway through retrograde transport. The ER chaperones and foldases in small numbers do travel to the Golgi, where they can pick up escapee clients. The chaperones carry at their C-terminus a lysine aspartate glutamate leucine (KDEL) or related tetrapeptide (Munro and Pelham, 1987), which is recognized by KDEL receptors in the Golgi. The ternary client-chaperone-KDEL receptor complex is then sorted to COPI-coated vesicles that travel back to the ER for further client folding attempts (Lewis and Pelham, 1992; Figure 10). The pH gradient between the ER and Golgi contributes in regulating transport and quality control. ERGIC-53 captures glycoproteins in the ER releasing them in the more acidic pH in the Golgi, where instead the chaperone ERp44 and KDEL receptors are activated (Wilson *et al.*, 1993; Vavassori *et al.*, 2013). The ternary complex formed with clients is retrieved back to the ER, whose neutral pH favors dissociation. It is thought that KDEL receptors keep track of the protein flux exiting the ER and accordingly act as signaling devices (they are structurally and functionally similar to GPCRs), to drive adaptations in vesicular transport capacity to and from the Golgi (Pulvirenti *et al.*, 2008).

ER Homeostasis

To maintain homeostasis, the ER folding capacity or efficiency must be adjusted to the folding load, which may vary during differentiation or due to changing environmental cues (drugs or cofactor deficiencies). In essence, the ER perceives accumulation of aberrant or orphan proteins that cannot complete their folding schedule. The intricate sensing mechanisms that

monitor load and that trigger a response to 'ER stress' are collectively referred to as the UPR pathways (Ron and Walter, 2007). In metazoans, the UPR is not only key for mitigating ER stress, but also is essential for development and maintenance of professional secretory cells, for instance during plasma cell differentiation (Iwakoshi *et al.*, 2003) or in pancreatic β -cell physiology (Harding *et al.*, 2001).

Insufficient protein folding capacity in the ER is detected by UPR transducers of the IRE1, PERK, and ATF6 families (Ron and Walter, 2007). They all are transmembrane proteins with an ER luminal stress sensing domain and a cytosolic effector domain. ATF6 proteins – which are exclusive to metazoans – travel from the ER to the Golgi. Here, cleavage by a resident protease allows the ATF6 cytosolic domain to detach, such that it can relocate to the nucleus, where it activates transcription of target genes (Mori, 2009). IRE1 proteins – which are conserved in all eukaryotes – initiate the non-conventional splicing of *xbp1* mRNA (*HAC1* in yeast), whereupon the spliced mRNA is translated into another transcription activator of distinct UPR target genes (Ron and Walter, 2007). PERK – exclusive to animal cells – phosphorylates eIF2 α , which causes translational attenuation, thereby blocking further entry of ER clients. Paradoxically, synthesis of some other proteins that further control the response (e.g., ATF4 and CHOP) is enhanced when eIF2 α is phosphorylated (Ron and Walter, 2007).

UPR target genes include a full repertoire of ER chaperones and foldases, but also enzymes for membrane lipid synthesis that together sustain expansion of the ER and its folding capacity. Other UPR targets are ERAD components, which enhance the clearance of misfolded clients from the ER. IRE1 proteins may also contribute to the curtailing of the folding load in the ER through degradation of mRNAs other than *xbp1*, a process referred to as regulated Ire1-dependent decay (Hollien and Weissman, 2006). Finally, excess or damaged ER with its contents can be targeted for removal by autophagy (Bernales *et al.*, 2006). Such 'ERphagy' may also be key to downsize the ER when cells are recovering from ER stress or to sustain balanced ER expansion in professional secretory cells (Pengo *et al.*, 2013).

Yet, when all adaptive measures are exhausted, the UPR induces apoptosis, for instance through persistent activation of PERK and its downstream effector CHOP. Also IRE1 can invoke pro-apoptotic pathways (Ron and Walter, 2007). Maladaptive ER stress responses can lie at the basis of a growing list of pathologies. For example, excessive insulin synthesis and a corresponding 'overzealous' UPR in pancreatic β -cells lead to their demise and, hence, type 2 diabetes (Oyadomari *et al.*, 2002). A schematic overview of how the UPR orchestrates ER homeostasis is shown in Figure 11.

Perspectives

Many questions remain open concerning the ER. How do cells set and maintain an optimal ratio between ER subdomains? What are all the distinguishing components of subdomains and the plethora of ER contact sites? How is production in the ER factory coordinated with the flux of proteins through the secretory pathway? How do developmental cues determine

that the ER specializes in one of its many manifestations? Which are the cues that activate and direct ERphagy? Can we reconstruct the ER in a test tube? Can we create model cells with a 'tailor-made' ER for biotechnological purposes? It is clear that the ER has a multitude of surprising and exciting features in store that remain to be discovered.

Acknowledgments

The authors wish to thank the Associazione Italiana della Ricerca sul Cancro (EvA, RS), the Ministero della Salute (EvA, RS), the Armenise–Harvard Foundation (EvA), and Telethon (RS) for grant support.

See also: Organelles: Structure and Function: Intermediate Compartment: A Sorting Station between the Endoplasmic Reticulum and the Golgi Apparatus

References

- Anelli, T., Ceppi, S., Bergamelli, L., *et al.*, 2007. Sequential steps and checkpoints in the early exocytic compartment during secretory IgM biogenesis. *The EMBO Journal* 26, 4177–4188.
- Anelli, T., Sitia, R., 2008. Protein quality control in the early secretory pathway. *EMBO Journal* 27 (2), 315–327.
- Anelli, T., Sitia, R., 2010. Physiology and pathology of proteostasis in the early secretory compartment. *Seminars in Cell & Developmental Biology* 21, 520–525.
- Anfinsen, C.B., 1973. Principles that govern the folding of protein chains. *Science* 181, 223–230.
- Appenzeller, C., Andersson, H., Kappeler, F., Hauri, H.P., 1999. The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. *Nature Cell Biology* 1, 330–334.
- Appenzeller-Herzog, C., Ellgaard, L., 2008. The human PDI family: Versatility packed into a single fold. *Biochimica et Biophysica Acta* 1783, 535–548.
- Barlowe, C., Orci, L., Yeung, T., *et al.*, 1994. COPII: A membrane coat formed by sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* 77, 895–907.
- Bernales, S., McDonald, K.L., Walter, P., 2006. Autophagy counterbalances endoplasmic reticulum expansion during the unfolded protein response. *PLoS Biology* 4, e423.
- Berridge, M.J., 2002. The endoplasmic reticulum: A multifunctional signaling organelle. *Cell Calcium* 32, 235–249.
- Blobel, G., Dobberstein, B., 1975. Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *Journal of Cell Biology* 67, 835–851.
- Bulleid, N.J., Ellgaard, L., 2011. Multiple ways to make disulfides. *Trends in Biochemical Sciences* 36, 485–492.
- Chen, Y., Zhang, Y., Yin, Y., *et al.*, 2005. SPD – a web-based secreted protein database. *Nucleic Acids Research* 33, D169–D173.
- Christianson, J.C., Shaler, T.A., Tyler, R.E., Kopito, R.R., 2008. OS-9 and GRP94 deliver mutant alpha1-antitrypsin to the Hrd1-SEL1L ubiquitin ligase complex for ERAD. *Nature Cell Biology* 10, 272–282.
- Cox, J.S., Shamu, C.E., Walter, P., 1993. Transcriptional induction of genes encoding endoplasmic reticulum resident proteins requires a transmembrane protein kinase. *Cell* 73, 1197–1206.
- Eletto, D., Dersh, D., Argon, Y., 2010. GRP94 in ER quality control and stress responses. *Seminars in Cell & Developmental Biology* 21, 479–485.
- Ellgaard, L., Helenius, A., 2003. Quality control in the endoplasmic reticulum. *Nature Reviews Molecular Cell Biology* 4, 181–191.
- Endo, M., 1977. Calcium release from the sarcoplasmic reticulum. *Physiological Reviews* 57, 71–108.
- English, A.R., Voeltz, G.K., 2013. Rab10 GTPase regulates ER dynamics and morphology. *Nature Cell Biology* 15, 169–178.

- Fagone, P., Jackowski, S., 2009. Membrane phospholipid synthesis and endoplasmic reticulum function. *Journal of Lipid Research* 50 (Suppl.), S311–S316.
- Galat, A., 2003. Peptidylprolyl cis/trans isomerases (immunophilins): Biological diversity – targets – functions. *Current Topics in Medicinal Chemistry* 3, 1315–1347.
- Gilmore, R., Blobel, G., Walter, P., 1982. Protein translocation across the endoplasmic reticulum. I. Detection in the microsomal membrane of a receptor for the signal recognition particle. *Journal of Cell Biology* 95, 463–469.
- Goder, V., Spiess, M., 2001. Topogenesis of membrane proteins: Determinants and dynamics. *FEBS Letters* 504, 87–93.
- Goldenberger, R.F., Epstein, C.J., Anfinsen, C.B., 1963. Acceleration of reactivation of reduced bovine pancreatic ribonuclease by a microsomal system from rat liver. *Journal of Biological Chemistry* 238, 628–635.
- Goyal, U., Blackstone, C., 2013. Untangling the web: Mechanisms underlying ER network formation. *Biochimica et Biophysica Acta* 1833, 2492–2498.
- Harding, H.P., Zeng, H., Zhang, Y., *et al.*, 2001. Diabetes mellitus and exocrine pancreatic dysfunction in *perK*^{-/-} mice reveals a role for translational control in secretory cell survival. *Molecular Cell* 7, 1153–1163.
- Helenius, A., Aebi, M., 2004. Roles of N-linked glycans in the endoplasmic reticulum. *Annual Review of Biochemistry* 73, 1019–1049.
- Helle, S.C., Kanfer, G., Kolar, K., *et al.*, 2013. Organization and function of membrane contact sites. *Biochimica et Biophysica Acta* 1833, 2526–2541.
- Hollien, J., Weissman, J.S., 2006. Decay of endoplasmic reticulum-localized mRNAs during the unfolded protein response. *Science* 313, 104–107.
- Hosokawa, N., Wada, I., Hasegawa, K., *et al.*, 2001. A novel ER alpha-mannosidase-like protein accelerates ER-associated degradation. *EMBO Reports* 2, 415–422.
- Hu, J., Shibata, Y., Zhu, P.P., *et al.*, 2009. A class of dynamin-like GTPases involved in the generation of the tubular ER network. *Cell* 138, 549–561.
- Hurtley, S.M., Bole, D.G., Hoover-Litty, H., Helenius, A., Copeland, C.S., 1989. Interactions of misfolded influenza virus hemagglutinin with binding protein (BiP). *Journal of Cell Biology* 108, 2117–2126.
- Iwakoshi, N.N., Lee, A.H., Vallabhajosyula, P., *et al.*, 2003. Plasma cell differentiation and the unfolded protein response intersect at the transcription factor *xbp-1*. *Nature Immunology* 4, 321–329.
- Keenan, R.J., Freymann, D.M., Stroud, R.M., Walter, P., 2001. The signal recognition particle. *Annual Review of Biochemistry* 70, 755–775.
- Klopfenstein, D.R., Klumperman, J., Lustig, A., *et al.*, 2001. Subdomain-specific localization of climp-63 (p63) in the endoplasmic reticulum is mediated by its luminal alpha-helical segment. *Journal of Cell Biology* 153, 1287–1300.
- Kopito, R.R., Sitia, R., 2000. Aggresomes and Russell bodies. Symptoms of cellular indigestion? *EMBO Reports* 1, 225–231.
- Kuehn, M.J., Herrmann, J.M., Schekman, R., 1998. CopII-cargo interactions direct protein sorting into ER-derived transport vesicles. *Nature* 391, 187–190.
- Lang, S., Benedix, J., Fedele, S.V., *et al.*, 2012. Different effects of *sec61* alpha, *sec62* and *sec63* depletion on transport of polypeptides into the endoplasmic reticulum of mammalian cells. *Journal of Cell Science* 125, 1958–1969.
- Lewis, M.J., Pelham, H.R., 1992. Ligand-induced redistribution of a human KDEL receptor from the Golgi complex to the endoplasmic reticulum. *Cell* 68, 353–364.
- Lin, S., Sun, S., Hu, J., 2012. Molecular basis for sculpting the endoplasmic reticulum membrane. *International Journal of Biochemistry & Cell Biology* 44, 1436–1443.
- Liu, B., Yang, Y., Qiu, Z., *et al.*, 2010. Folding of toll-like receptors by the HSP90 paralogue gp96 requires a substrate-specific cochaperone. *Nature Communications* 1, 1–10.
- Mayer, M.P., Bukau, B., 2005. HSP70 chaperones: Cellular functions and molecular mechanism. *Cellular and Molecular Life Sciences* 62, 670–684.
- Meldolesi, J., Pozzan, T., 1998. The endoplasmic reticulum Ca^{2+} store: A view from the lumen. *Trends in Biochemical Sciences* 23, 10–14.
- Mori, K., 2009. Signalling pathways in the unfolded protein response: Development from yeast to mammals. *Journal of Biochemistry* 146, 743–750.
- Mori, K., Ma, W., Gething, M.J., Sambrook, J., 1993. A transmembrane protein with a *cdc2* + /CDC28-related kinase activity is required for signaling from the ER to the nucleus. *Cell* 74, 743–756.
- Munro, S., Pelham, H.R., 1987. A C-terminal signal prevents secretion of luminal ER proteins. *Cell* 48, 899–907.
- Naon, D., Scorrano, L., 2014. At the right distance: ER-mitochondria juxtaposition in cell life and death. *Biochimica et Biophysica Acta* 1843, 2184–2194.
- Novick, P., Field, C., Schekman, R., 1980. Identification of 23 complementation groups required for post-translational events in the yeast secretory pathway. *Cell* 21, 205–215.
- Orsi, A., Sitia, R., 2007. Interplays between covalent modifications in the endoplasmic reticulum increase conformational diversity in nascent prion protein. *Prion* 1, 236–242.
- Oyadomari, S., Araki, E., Mori, M., 2002. Endoplasmic reticulum stress-mediated apoptosis in pancreatic beta-cells. *Apoptosis* 7, 335–345.
- Palade, G.E., Porter, K.R., 1954. Studies on the endoplasmic reticulum. I. Its identification in cells in situ. *Journal of Experimental Medicine* 100, 641–656.
- Palade, G.E., Siekevitz, P., 1956. Liver microsomes; an integrated morphological and biochemical study. *Journal of Biophysical and Biochemical Cytology* 2, 171–200.
- Pengo, N., Scolari, M., Oliva, L., *et al.*, 2013. Plasma cells require autophagy for sustainable immunoglobulin production. *Nature Immunology* 14, 298–305.
- Prassman, J.L., Wells, L., 2014. Mammalian O-mannosylation pathway: Glycan structures, enzymes, and protein substrates. *Biochemistry* 53, 3066–3078.
- Pulvirenti, T., Giannotta, M., Capestrano, M., *et al.*, 2008. A traffic-activated Golgi-based signalling circuit coordinates the secretory pathway. *Nature Cell Biology* 10, 912–922.
- Ron, D., Walter, P., 2007. Signal integration in the endoplasmic reticulum unfolded protein response. *Nature Reviews. Molecular Cell Biology* 8, 519–529.
- Schuldiner, M., Metz, J., Schmid, V., *et al.*, 2008. The GET complex mediates insertion of tail-anchored proteins into the ER membrane. *Cell* 134, 634–645.
- Shen, Y., Meunier, L., Hendershot, L.M., 2002. Identification and characterization of a novel endoplasmic reticulum (ER) DnaJ homologue, which stimulates ATPase activity of BiP *in vitro* and is induced by ER stress. *Journal of Biological Chemistry* 277, 15947–15956.
- Siekevitz, P., Palade, G.E., 1960. A cytochemical study on the pancreas of the guinea pig. 5. *In vivo* incorporation of leucine-1-C14 into the chymotrypsinogen of various cell fractions. *Journal of Biophysical and Biochemical Cytology* 7, 619–630.
- Steel, G.J., Fullerton, D.M., Tyson, J.R., Stirling, C.J., 2004. Coordinated activation of HSP70 chaperones. *Science* 303, 98–101.
- Stefan, C.J., Manford, A.G., Emr, S.D., 2013. ER-PM connections: Sites of information transfer and inter-organelle communication. *Current Opinion in Cell Biology* 25, 434–442.
- Sudhof, T.C., Rothman, J.E., 2009. Membrane fusion: Grappling with SNARE and SM proteins. *Science* 323, 474–477.
- Tu, B.P., Weissman, J.S., 2004. Oxidative protein folding in eukaryotes: Mechanisms and consequences. *Journal of Cell Biology* 164, 341–346.
- Tyson, J.R., Stirling, C.J., 2000. Lhs1 and Sil1 provide a luminal function that is essential for protein translocation into the endoplasmic reticulum. *EMBO Journal* 19, 6440–6452.
- Udenfriend, S., Kodukula, K., 1995. How glycosylphosphatidylinositol-anchored membrane proteins are made. *Annual Review of Biochemistry* 64, 563–591.
- Vance, D.E., Choy, P.C., Farren, S.B., Lim, P.H., Schneider, W.J., 1977. Asymmetry of phospholipid biosynthesis. *Nature* 270, 268–269.
- Vavassori, S., Cortini, M., Masui, S., *et al.*, 2013. A pH-regulated quality control cycle for surveillance of secretory protein assembly. *Molecular Cell* 50, 783–792.
- Veratti, E., 1961. Investigations on the fine structure of striated muscle fiber read before the Reale Istituto Lombardo, 13 march 1902. *Journal of Biophysical and Biochemical Cytology* 10 (Suppl. 4), 1–59.
- Voeltz, G.K., Prinz, W.A., 2007. Sheets, ribbons and tubules - how organelles get their shape. *Nature reviews. Molecular Cell Biology* 8, 258–264.
- Voeltz, G.K., Prinz, W.A., Shibata, Y., Rist, J.M., Rapoport, T.A., 2006. A class of membrane proteins shaping the tubular endoplasmic reticulum. *Cell* 124, 573–586.
- Walter, P., Ibrahimi, I., Blobel, G., 1981. Translocation of proteins across the endoplasmic reticulum. I. Signal recognition protein (SRP) binds to in-vitro-assembled polysomes synthesizing secretory protein. *Journal of Cell Biology* 91, 545–550.
- Wang, F., Chan, C., Weir, N.R., Denic, V., 2014. The GET1/2 transmembrane complex is an endoplasmic-reticulum membrane protein insertase. *Nature* 512, 441–444.
- Werner, E.D., Brodsky, J.L., McCracken, A.A., 1996. Proteasome-dependent endoplasmic reticulum-associated protein degradation: An unconventional route

- to a familiar fate. *Proceedings of the National Academy of Sciences of the United States of America* 93, 13797–13801.
- Wilson, D.W., Lewis, M.J., Pelham, H.R., 1993. pH-dependent binding of KDEL to its receptor *in vitro*. *Journal of Biological Chemistry* 268, 7465–7468.
- Xu, N., Zhang, S.O., Cole, R.A., *et al.*, 2012. The FATP1-DGAT2 complex facilitates lipid droplet expansion at the ER-lipid droplet interface. *Journal of Cell Biology* 198, 895–911.
- van der Zand, A., Gent, J., Braakman, I., Tabak, H.F., 2012. Biochemically distinct vesicles from the endoplasmic reticulum fuse to form peroxisomes. *Cell* 149, 397–409.
- Zimmermann, R., Mollay, C., 1986. Import of honeybee preproinsulin into the endoplasmic reticulum. Requirements for membrane insertion, processing, and sequestration. *Journal of Biological Chemistry* 261, 12889–12895.
- Anfinsen, C.B., 1972. Studies on the Principles that Govern the Folding of Protein Chains. Nobel Lecture, December 11, 1972.
- Blobel, G., 1999. Protein Targeting. Nobel Lecture, December 8, 1999.
- Braakman, I., Hebert, D.N., 2013. Protein folding in the endoplasmic reticulum. *Cold Spring Harbor Perspectives in Biology* 5, a013201.
- Brodsky, J.L., 2012. Cleaning up: ER-associated degradation to the rescue. *Cell* 151, 1163–1167.
- Ellgaard, L., Molinari, M., Helenius, A., 1999. Setting the standards: Quality control in the secretory pathway. *Science* 286, 1882–1888.
- English, A.R., Voeltz, G.K., 2013. Endoplasmic reticulum structure and interconnections with other organelles. *Cold Spring Harbor Perspectives in Biology* 5, a013227.
- Palade, G.E., 1974. Of Protein Secretion. Nobel Lecture, December 12, 1974.
- Schekman, R.W., 2013. Genetic and Biochemical Dissection of the Secretory Pathway. Nobel Lecture, December 7, 2013.
- Walter, P., Ron, D., 2011. The unfolded protein response: From stress pathway to homeostatic regulation. *Science* 334, 1081–1086.

Further Reading

- Aebi, M., Bernasconi, R., Clerc, S., Molinari, M., 2010. N-glycan structures: Recognition and processing in the ER. *Trends in Biochemical Sciences* 35, 74–82.

Intermediate Compartment: A Sorting Station between the Endoplasmic Reticulum and the Golgi Apparatus

J Saraste and M Marie, University of Bergen, Bergen, Norway

© 2016 Elsevier Inc. All rights reserved.

Glossary

Anterograde transport Transport from endoplasmic reticulum (ER) toward the cell surface.

ER exit sites ER subdomains specialized in the export of newly synthesized molecules.

Golgi stacks Stacks of flat membrane-bound cisternae. The mammalian Golgi stacks contain 4–8 cisternae.

Lumen The space inside a membrane-bound structure.

Microtubules Cytoskeletal filaments that provide tracks for motor-dependent long-distance transport of organelles and transport carriers.

Pleomorphic Variable in size and shape.

Retrograde transport Transport from post-ER compartments toward the ER.

Saccule Membrane-bound structure with a large lumen.

Vesicle Membrane-bound structure with a small lumen.

Introduction

Newly synthesized proteins and lipids leave the endoplasmic reticulum (ER) at specialized transitional regions called ER exit sites (ERES) (Jamieson and Palade, 1967; Sesso *et al.*, 1994; Bannykh *et al.*, 1996; Hammond and Glick, 2000; Tang *et al.*, 2005) and enter the intermediate compartment (IC) that has been shown to operate as an obligatory a post-ER sorting station in the early biosynthetic-secretory trafficking of mammalian cells. From the IC they are typically transported to the cisternal stacks of the Golgi apparatus, prior to their delivery to the different organelles of the endomembrane system or secretion to the extracellular space. Bidirectional ER–Golgi trafficking involves the sequential operation of membrane-bound coat protein II (COPII) and COPI coats (Aridor and Balch, 1996; Scales *et al.*, 1997; Stephens *et al.*, 2000). ER-derived COPII vesicles mediate forward (anterograde) transport, while IC- and Golgi-derived COPI vesicles are thought to act in the opposite (retrograde) direction (Lee *et al.*, 2004; Rabouille and Klumperman, 2005). Retrograde transport also involves COPI-independent routes (Kano *et al.*, 2009).

Despite the conservation of the transport machineries (such as the COP coats), the organization of the ER–Golgi interface varies in different eukaryotic cells. For example, in plants, certain yeasts, and the fruit fly *Drosophila melanogaster*, the individual Golgi stacks lie next to the widespread ERES, establishing units for short range ER–Golgi communication. By contrast, in animal cells the Golgi stacks are linked together into a continuous ribbon around the microtubule-organizing center (MTOC)/centrosome, whereas the ER extends throughout the cytoplasm. Hence, a large proportion of the ERES reside at the cell periphery and ER–Golgi trafficking depends on the long distance movements of the IC elements along MT tracks (Saraste and Svensson, 1991; Presley *et al.*, 1997; Scales *et al.*, 1997; Brandizzi and Barlowe, 2013; Day *et al.*, 2013). Thus, it has been proposed that the IC represents a late evolutionary invention, which developed to meet the special sorting, transport and recycling requirements of the large-sized animal cells, but lacks lower eukaryotes. However, results showing that the IC constitutes an extensive membrane system

that can be compared with the endosomal network in complexity, are questioning this view.

Historical Perspective

15 °C Compartment

Electron microscopic (EM) studies using a temperature-sensitive mutant of Semliki Forest virus (SFV ts-1) to synchronize the transport of viral membrane glycoproteins from ER to the plasma membrane (PM) showed that when the cells are shifted from 39 °C to 15 °C the proteins exit the ER, but accumulate in vacuoles/saccules (up to 0.5 µm in diameter), tubules, and vesicles in the *cis*-Golgi region and more peripheral locations (Saraste and Kuusmanen, 1984). By light microscopy (LM) the proteins were localized at 15 °C to scattered globular structures in the cytoplasm, a pattern distinct from that of ER or Golgi (Kuusmanen and Saraste, 1989). The transport block was readily reversible, and the proteins entered the Golgi stacks and reached the PM following the transfer of cells from 15 °C to 28 °C, showing that these structures are normal transport intermediates.

Studies employing a similar mutant of vesicular stomatitis virus (VSV tsO45) showed that the '15 °C compartment' also acts as a way station during the transport of the VSV G glycoprotein (Balch *et al.*, 1986; Bonatti *et al.*, 1989; Schweizer *et al.*, 1990; Duden *et al.*, 1991; Lotti *et al.*, 1992). Like the SFV proteins, the G-protein was found to maintain its ER-type high-mannose glycans at 15 °C, indicating lack of processing by Golgi enzymes. Furthermore, cell fractionation experiments showed that newly synthesized secretory proteins are arrested in a post-ER location when pancreatic exocrine cells are kept at 16 °C (Saraste *et al.*, 1986). Live cell imaging of green fluorescent protein (GFP) equipped with an ER targeting signal (ssGFP) and EM studies of procollagen and growth hormone constructs verified that the transport of membrane and secretory proteins is similarly affected at 15–16 °C (Blum *et al.*, 2000; Volchuk *et al.*, 2000; Trucco *et al.*, 2004). In addition, the transport of virus glycoproteins and cholesterol appears to

be blocked at the same site at this temperature (Urbani and Simoni, 1990; Heino *et al.* 2000).

Coronavirus Budding Compartment

EM of mouse hepatitis virus (MHV)-infected cells showed that the budding of progeny virus initially takes place at tubulovesicular membranes located in the transitional region between the ER and the Golgi apparatus (Tooze *et al.*, 1984). The first step of O-glycosylation, the attachment of N-acetyl-galactosamine (GalNAc) to the viral membrane (M) glycoprotein, was suggested to occur in this compartment, causing its reactivity with the lectin Helix pomatia, which specifically binds GalNAc (Tooze *et al.*, 1988; Krijnse-Locker *et al.*, 1994). Subsequent studies showed that the intracellular maturation of various coronaviruses occurs at the same budding site, which corresponds to the IC (Klumperman *et al.*, 1994).

Salvage Compartment

The discovery of the lys-asp-glu-leu tetra-peptide (KDEL)-motif in abundant, luminal ER proteins lead to the proposal that the 15 °C/budding compartment is the post-ER site from which these proteins are retrieved to the ER (Munro and Pelham, 1987; Warren, 1987; Pelham, 1989). The first mammalian KDEL-receptor, a multispanning membrane protein, was identified and shown to predominantly localize to the IC and *cis*-Golgi (Lewis and Pelham, 1990; Tang *et al.*, 1993; Griffiths *et al.*, 1994; Orci *et al.*, 1997). Also, KDEL proteins, such as the immunoglobulin binding protein (BiP/GRP78), glucose-regulated protein of 94 kDa (GRP94), protein disulphide isomerase (PDI) and calreticulin (CR), are present in the IC (Griffiths *et al.*, 1994; Connolly *et al.*, 1994; Zuber *et al.*, 2001; Ying *et al.*, 2002; Breuza *et al.*, 2004). Following binding to their receptor, the proteins are retrieved to the ER in COPI vesicles (Orci *et al.*, 1997; Martínez-Menárguez *et al.*, 1999; Majoul *et al.*, 2001). Attachment of the KDEL-motif to lysosomal enzymes and the use of the 15 °C block suggested that the enzyme that initiates the formation of their lysosomal targeting signal (mannose-6-phosphate) resides in the IC (Pelham, 1988; Lazzarino and Gabel, 1988; Dittmer and von Figura, 1999). The cytoplasmic tails of certain type I integral membrane proteins were shown to contain dilysine (KKXX)-motifs, which by directly interacting with COPI coats result in their retrieval from the IC/*cis*-Golgi to the ER (Nilsson *et al.*, 1989; Jackson *et al.*, 1990, 1993; Cosson and Letourneur, 1997).

p58/ERGIC-53/LMAN1

The first endogenous IC markers rat p58 and human ER-Golgi intermediate compartment (ERGIC)-53 (89% homology) were identified by the generation of antibodies against the 16 °C post-ER fraction isolated from pancreatic acinar cells (Saraste *et al.*, 1987) and a Golgi fraction derived from epithelial Caco-2 cells (Schweizer *et al.*, 1988), respectively. The cytoplasmic C-terminal tails of these non-glycosylated, hexameric, type-1 integral membrane proteins (Schindler *et al.*,

1993; Lahtinen *et al.*, 1992, 1996; Neve *et al.*, 2005) contain a KKFF-motif, which interacts with COPII and COPI coats and gives rise to their continuous cycling between ER, IC, and *cis*-Golgi (Kappeler *et al.*, 1997; Tisdale *et al.*, 1997). At 15 °C the recycling of p58/ERGIC-53 to the ER is inhibited and the proteins pile up in the same pre-Golgi structures that contain the SFV or VSV membrane proteins (Schweizer *et al.*, 1990; Saraste and Svensson, 1991; Plutner *et al.*, 1992), verifying that the p58/ERGIC-53 compartment (Figures 1 and 4) is equivalent to the site where cargo is arrested at low temperature.

ERGIC-53 and the related VIP36 were shown to share homology with leguminous plant lectins (Fiedler and Simons, 1994) and to be identical with a mannose-binding protein MR60 of human myelomonocytic (HL60) cells (Arar *et al.*, 1995). The N-terminal domain of p58/ERGIC-53 binds mannose in a calcium-dependent manner (Itin *et al.*, 1996; Velloso *et al.*, 2003; Zheng *et al.*, 2013); hence the name lectin mannose-binding protein 1 (LMAN1). It is the best characterized mammalian cargo receptor, facilitating the COPII vesicle-mediated export of a subset of soluble glycoproteins from the ER (Nichols *et al.*, 1998; Appenzeller *et al.*, 1999; Hauri *et al.*, 2000).

Compositional Aspects

Most well-characterized IC proteins are components of the transport machinery. Many of them cycle at the ER-Golgi boundary and are also found in *cis*-Golgi, supporting the view of the IC as a transient structure (see below). In fact, EM studies showing the specific metal (osmium) staining of the IC and *cis*-Golgi provided the first indication of their compositional similarity (Friend and Murray, 1965; Rambourg *et al.*, 1974). However, the fungal compound brefeldin A (BFA) helps to discriminate between IC and Golgi components. It releases COPI coats, disassembles the Golgi stacks and redistributes Golgi components to the ER, whereas cycling proteins (such as p58/ERGIC-53/LMAN1 and the KDEL-receptor) are arrested in the drug-resistant IC elements (Lippincott-Schwartz *et al.*, 1990; Saraste and Svensson, 1991; Tang *et al.*, 1995b; Füllekrug *et al.*, 1997; Ward *et al.*, 2001; Marie *et al.*, 2009). The BFA resistance of the IC has been utilized for its proteomics analysis (Breuza *et al.*, 2004), and indicates its stability.

The p24 Protein Family

Like p58/ERGIC-53, the p24 family proteins contain motifs for COPII and COPI binding, resulting in their cycling between the ER, IC, and *cis*-Golgi (Rojo *et al.*, 1997; Dominguez *et al.*, 1998; Blum *et al.*, 1999; Gommel *et al.*, 1999; Strating and Martens, 2009). As abundant type-1 transmembrane proteins they participate in the biogenesis of COPI vesicles (Stamnes *et al.*, 1995; Majoul *et al.*, 2001; Beck *et al.*, 2009), but have also been implicated in tubulation (Simpson *et al.*, 2006) and the formation of membrane domains (Lavoie *et al.*, 1999; Emery *et al.*, 2003). Studies in yeast first suggested their function as receptors for the exit of

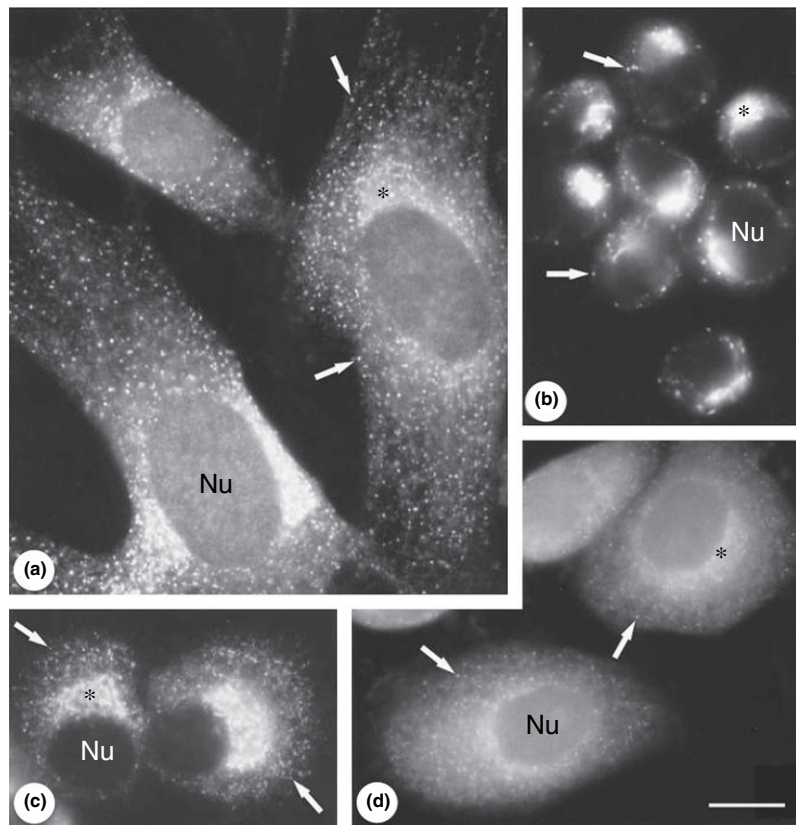


Figure 1 (a) Immunofluorescence LM localization of p58/ERGIC-53 in baby hamster kidney, (b) mouse myeloma, (c) rat neuroendocrine PC12, and (d) human HeLa cells. The protein marks the IC elements accumulating in the perinuclear Golgi region (asterisks), scattered throughout the cytoplasm, and located close to the PM (arrows). The reticular staining indicates the ER pool of p58/ERGIC-53, which varies in the different cell types. Nu, nucleus. Bar: 10 μ m.

glycosylphosphatidylinositol (GPI)-anchored proteins from the ER (Schimmöller *et al.*, 1995).

suggesting that the IC represents the main point for recycling (Martínez-Menárguez *et al.*, 1999).

COPI Coats

The COPI coats are mainly recruited to the cytoplasmic surface of IC and *cis*-Golgi membranes, but also associate with the lateral rims of the Golgi stacks (Duden *et al.*, 1991; Oprins *et al.*, 1993; Griffiths *et al.*, 1995; Orci *et al.*, 1997; Klumperman *et al.*, 1998). COPI vesicle budding depends on the activation of small GTPases of the ADP-ribosylation factor (Arf) family (Popoff *et al.*, 2011). The guanine nucleotide exchange factor (GEF) of these Arfs, GBF1, is the master regulator of COPI recruitment and the target of BFA (Niu *et al.*, 2005). It localizes to the IC and *cis*-Golgi and plays a key role in ER-to-Golgi trafficking (Kawamoto *et al.*, 2002; Garcia-Mata *et al.*, 2003; Zhao *et al.*, 2006; Szul *et al.*, 2007). GBF1 knock down or specific inhibition by Golgicide A releases COPI coats and arrests the VSV G-protein in the IC, indicating its involvement in anterograde IC-to-Golgi transport (Manolea *et al.*, 2008; Sáenz *et al.*, 2009). However, the prevailing view is that the main function of COPI vesicles is to recycle membrane and selected proteins to the ER. In pancreatic exocrine cells approximately 70% of the coats associate with the VTCs,

Rab1 and Rab2

The roles of two GTPases of the large Rab family, Rab1 and Rab2, in ER–Golgi trafficking are relatively well understood (Plutner *et al.*, 1991; Tisdale *et al.*, 1992). Both interact with multiple effectors suggesting that they coordinate successive transport steps (Barnekow *et al.*, 2009). The association of Rab1 with the cytoplasmic surface of IC and *cis*-Golgi membranes, and its absence from the ER, has been demonstrated by EM (Griffiths *et al.*, 1994; Saraste *et al.*, 1995; Satoh *et al.*, 2003; Marie *et al.*, 2012; Figure 2). The two Rab1 isoforms, Rab1A and RAB1B (93% homology), are recruited to IC membranes at ERES, show similar localizations by LM (Mochizuki *et al.*, 2013; Figure 3(a)), but seem to play distinct roles in tubular and vesicular (long and short range) trafficking within the IC. Accordingly, live cell imaging and cell fractionation have shown that Rab1A mainly associates with IC tubules (Sannerud *et al.*, 2006; Marie *et al.*, 2009), while Rab1B interacts with GBF1 and evidently modulates COPI recruitment to globular IC domains (Alvarez *et al.*, 2003; Monetta *et al.*, 2007; see below). Also, Rab1A is specifically

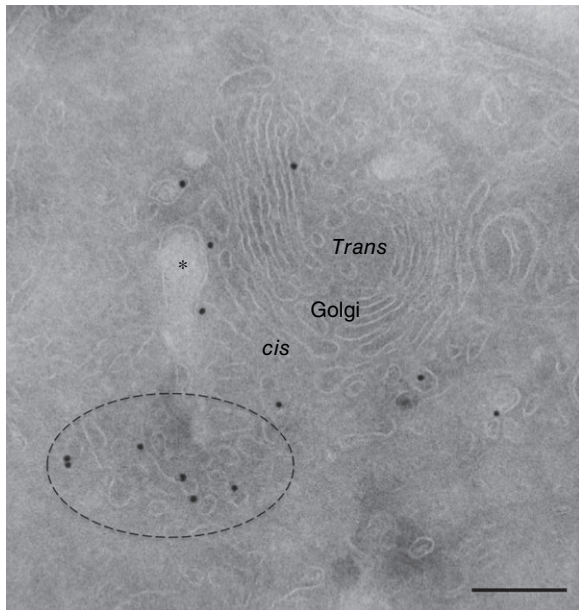


Figure 2 Immunogold EM localization of Rab1 in the Golgi region of a normal rat kidney (NRK) cell. The 15 nm gold particles predominantly label pleiomorphic IC elements at the *cis*-face of the Golgi stacks. A saccule (asterisk) and a vesicular tubular cluster (VTC; dashed area) are indicated. The micrograph was kindly provided by Dr. Karin Pernet-Gallay (University Joseph Fourier, Grenoble, France). Bar: 3 μ m.

phosphorylated at mitosis (Bailly *et al.*, 1991), correlating with the cessation of tubular IC dynamics, whereas COPI-mediated vesicular transport continues (Marie *et al.*, 2012). Rab1A can functionally replace Ypt1, which coordinates two-way ER–Golgi trafficking in the yeast *Saccharomyces cerevisiae* (Haubruck *et al.* 1989; Segev, 2001; Kamena *et al.*, 2008).

Rab2 has been localized to IC and *cis*-Golgi membranes (Chavrier *et al.*, 1990; Lotti *et al.*, 1992) and proposed to regulate the formation of retrograde COPI vesicles that contain p58/ERGIC-53 (Tisdale, 1999). It has also been implicated in the recruitment of the motor protein dynein to IC membranes and the association of IC elements with MTs (Tisdale *et al.*, 2009; see below).

Tethers and Soluble NSF Attachment Protein Receptors

Most of the known Rab1 and Rab2 effectors are cytoplasmically oriented, long coiled-coil proteins, which function in the tethering of vesicle and organelle membranes preceding their soluble NSF attachment protein receptor (SNARE)-mediated fusion (Barnekow *et al.*, 2009; Sztul and Lupashin, 2009). Besides the IC/*cis*-Golgi tethers GM130 and GMAP-210 (Rios *et al.*, 2004), Rab2 has been shown to interact with *medial*- and *trans*-Golgi tethers, suggesting a more widespread function (Short *et al.*, 2001; Sinka *et al.*, 2008). The Rab1 effector p115 is functional already at the peripheral ERES (Alvarez *et al.*, 1999), while GM130 (bound to its membrane anchor GRASP65) cycles between central IC and *cis*-Golgi (Marra *et al.*, 2001) and regulates membrane tethering at a later transport step (Alvarez *et al.*, 2001; Marra *et al.*, 2007). Rab1 also

interacts with the transmembrane tethers golgin-84 and giantin, which appear to act in COPI vesicle trafficking at the lateral rims of the Golgi stacks (Diao *et al.*, 2003; Malsam *et al.*, 2005; Barnekow *et al.*, 2009; Munro, 2011).

The membrane fusion machinery (SNARE proteins) operating in ER–Golgi trafficking in the yeast *Saccharomyces cerevisiae* has been well characterized (Cai *et al.*, 2007; Barlowe and Miller, 2013), whereas the fusion events that take place in mammalian cells remain enigmatic. The determination of the exact fusion steps is complicated by the continuous cycling of the SNAREs in COPI vesicles (Hay *et al.*, 1998; Martínez-Menárguez *et al.*, 2001). In analogy to yeast, Rab1 (Ypt1) has been suggested to recruit p115 (Uso1p) to COPII vesicles at ERES (Allan *et al.*, 2000), followed by the formation of a SNARE complex (Sec22B, membrin, Bet1 and syntaxin 5), which mediates either homotypic fusion of COPII vesicles or their heterotypic fusion with the stationary IC (Zhang *et al.*, 1997; Xu *et al.*, 2000; see below). Another SNARE complex (syntaxin 5, GS28, Bet1 and Ykt6) is involved in a later *cis*-Golgi transport step (Zhang and Hong, 2001).

Structure, Distribution, and Dynamics

By EM the IC elements can be readily distinguished from the ER and Golgi, but share structural similarity with endosomes. They reside close to peripheral and central ERES as clusters of vesicles and tubules (VTCs; Figure 4(f)) that frequently contain COPI coats (Balch *et al.*, 1994; Bannykh *et al.*, 1996; Martínez-Menárguez *et al.*, 1999). However, they display considerable size heterogeneity (Ying *et al.*, 2000) and also include large saccules that are found within the membrane clusters, free in the cytoplasm, or at the *cis*-face of the Golgi stacks (Saraste and Svensson, 1991; Lahtinen *et al.*, 1992; Connolly *et al.*, 1994; Stinchcombe *et al.*, 1995; Ladinsky *et al.*, 1999; Fan *et al.*, 2003; Figures 4(a)–4(e)). These pleiomorphic saccules can accommodate large-sized cargo, such as procollagen or luminal protein aggregates (Volchuk *et al.*, 2000; Trucco *et al.*, 2004; Zuber *et al.*, 2004), and based on correlative microscopy (LM/EM) correspond to many of the mobile carriers that are visible in living cells (Mironov *et al.*, 2003). Like endosomes, they extend narrow tubules and also bind COPI coats, indicating that they represent sites for vesicle budding (Saraste and Kuismanen, 1984; Volchuk *et al.*, 2000; Horstman *et al.*, 2002; Figures 4(c) and 4(d)). The hypertrophy of the saccular domains most likely gives rise to the pre-Golgi vacuoles seen in cells kept at 15 °C (Saraste and Kuismanen, 1992; Trucco *et al.*, 2004).

LM shows the division of the IC into globular and tubular domains (Presley *et al.*, 1998). The former contain anterograde cargo, cargo receptors and COPI coats, most likely corresponding to individual VTCs or free saccules, while the latter are highlighted by Rab1A (Sannerud *et al.*, 2006). The tubules extend from the globular domains (Figure 5). Some of them contain recycling proteins, but lack anterograde cargo, indicating that they function in retrograde transport (Palokangas *et al.*, 1998; Simpson *et al.*, 2006). However, under synchronized conditions cargo is also detected in the tubular domain, due to overload or the existence of different types of

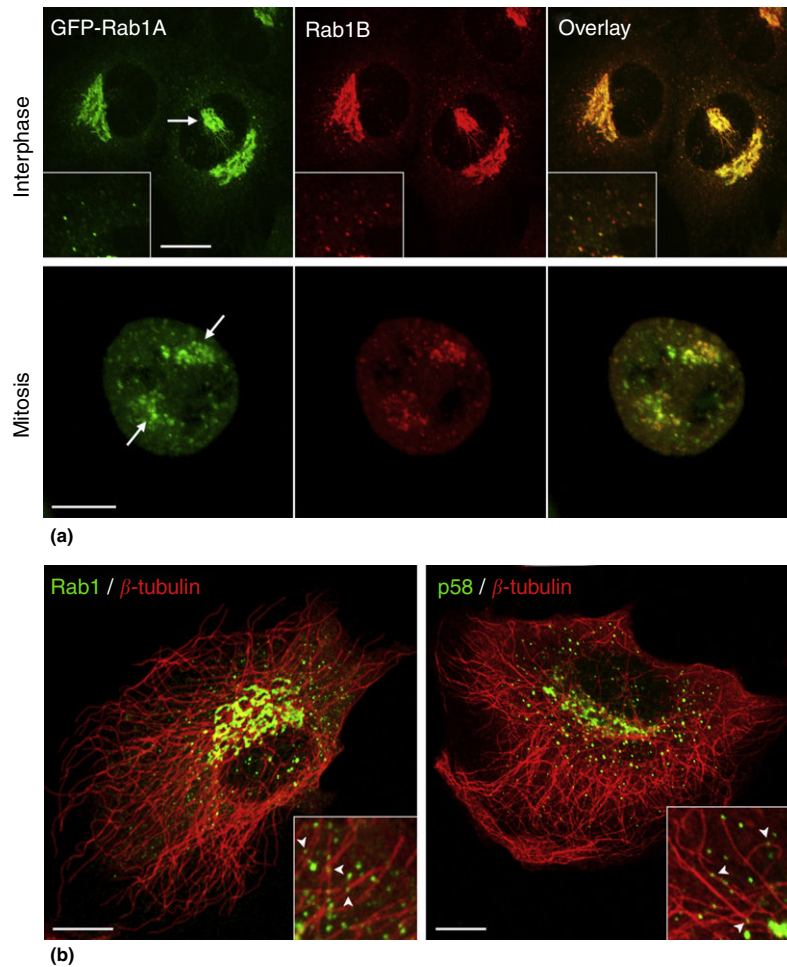


Figure 3 (a) Immunofluorescence LM localization of the two Rab1 isoforms at different phases of the cell cycle. NRK cells stably expressing GFP-Rab1A were stained with monoclonal antibodies against Rab1B (kindly provided by Prof. Angelika Barnekow, University of Münster, Germany). At interphase the proteins colocalize in the Golgi ribbon, the separated pcIC (arrow) and peripheral IC elements (insets). During mitosis (at metaphase), following Golgi disassembly, the isoforms maintain their pericentrosomal co-localization at the spindle poles (arrows). (b) Association of IC elements with MTs. NRK cells were stained for the IC markers Rab1 or p58 and β -tubulin (to visualize MTs). Only the image overlays are shown. The insets show coalignment of IC elements with MTs (arrowheads). See also Marie *et al.*, 2012. Bars: 10 μ m.

IC tubules (Figure 6). For example, incubation of cells at 15–16 °C inhibits the formation of IC tubules, but causes the expansion of the globular domain, while the shift of cells to 37 °C generates tubular networks containing both antero- and retrograde markers (Blum *et al.*, 2000; Ben-Takaya *et al.*, 2005; Simpson *et al.*, 2006). Proliferation of the tubules is induced when COPI function is compromised (Szul *et al.*, 2007; Marie *et al.*, 2009; Ben-Takaya *et al.*, 2010; Tomás *et al.*, 2010; Hamlin *et al.*, 2014), but also occurs under physiological conditions. In differentiating neuroendocrine PC12 cells the Rab1A-positive tubular IC domain expands and the tubules move from the cell body to the forming neurites accumulating in their growth cones (Sannerud *et al.*, 2006; Figure 5). An analogous pathway connects the IC with the leading edge of fibroblasts (Figure 6).

Live imaging of various fluorescent IC markers indicates that the tubules are highly dynamic, while the globular domains are typically more stationary. Due to their differential localization within these domains, the constructs highlight

different aspects of IC dynamics (see below). Long distance ER-to-Golgi transport involves the movement of IC elements from the cell periphery to the central *cis*-Golgi region, resulting in the division of the IC into spatially distinct early (ERES-adjacent) and late (*cis*-Golgi-adjacent) domains (Saraste and Svensson, 1991; Presley *et al.*, 1997; Scales *et al.*, 1997; Marra *et al.*, 2001). Two types of anterogradely moving IC elements can be resolved by LM, narrow tubules and large pleiomorphic structures. Some of the latter appear to represent saccular elements that transform into elongated structures ('thick tubules') (Presley *et al.*, 2002; Marie *et al.*, 2009). In addition, narrow tubules establish dynamic connections between the globular IC domains (Ben-Takaya *et al.*, 2005; Sannerud *et al.*, 2006).

Association with the Cytoskeleton

The long distance movements of the IC elements depend on MTs (Murshid and Presley, 2004; Palmer *et al.*, 2005; Figures 3

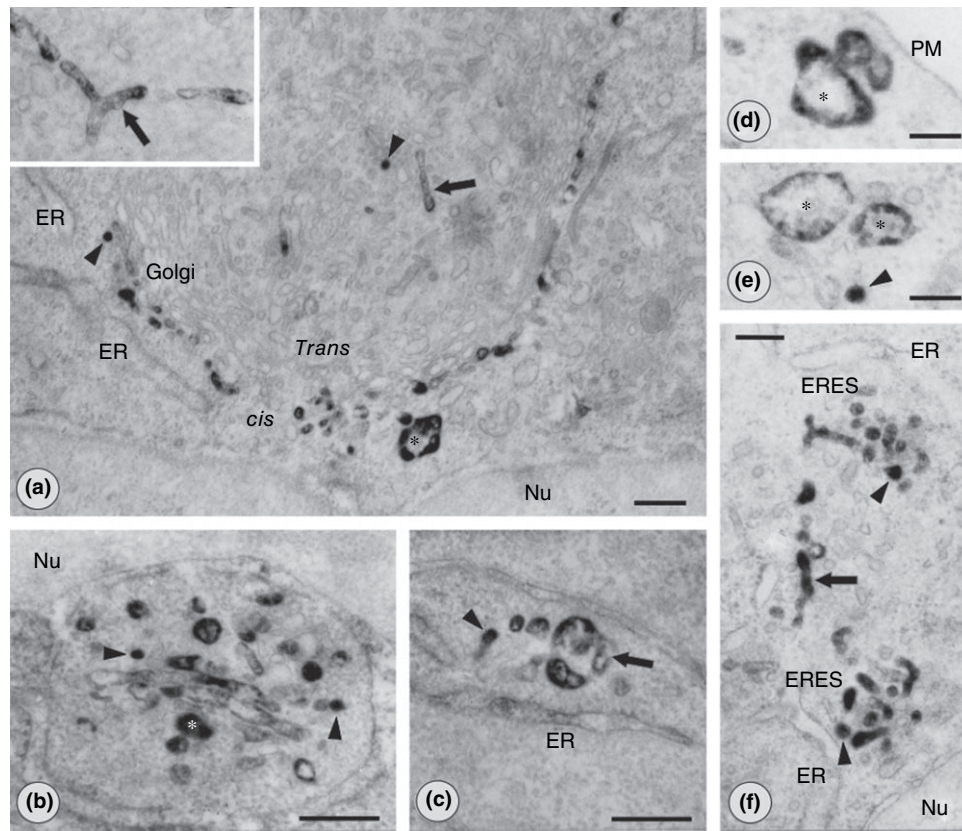


Figure 4 Immunoperoxidase EM analysis of the ultrastructure of p58-containing IC elements in NRK (a; inset) and mouse myeloma (b–f) cells. The protein is present in the *cis*-most Golgi cisterna (a) and in vesicular (arrowheads; ca. 80 nm in diameter), tubular (arrows) and saccular (asterisks; up to 0.5 μ m in diameter) IC elements. Panels (a–c) adapted from Ying, M., Flatmark, T., Saraste, J., 2000. The p58-positive pre-Golgi intermediates consist of distinct subpopulations of particles that show differential binding of COPI and COPII coats and contain vacuolar H^+ -ATPase. *Journal of Cell Science* 113, 3623–3638; (f) from Saraste, J., Palade, G.E., Farquhar, M.G., 1987. Antibodies to rat pancreas Golgi subfractions: Identification of a 58 kDa *cis*-Golgi protein. *Journal of Cell Biology* 105, 2021–2029. Nu; nucleus. Bars: 0.5 μ m (a–c), 0.25 μ m (d–f).

(b) and 6). The plus- and minus-end directed motor proteins kinesin and dynein associate with the IC elements (Lippincott-Schwartz *et al.*, 1995; Roghi and Allan, 1999; Stauber *et al.*, 2006), explaining their bidirectional movements even along the same MT tracks (Sannerud *et al.*, 2006). When the MTs in mammalian cells are depolymerized by nocodazole the IC elements become immobile and accumulate close to ERES (Saraste and Svensson, 1991; Hammond and Glick, 2000; Ben-Tekaya *et al.*, 2005; Sannerud *et al.*, 2006). Although the central Golgi ribbon breaks down, the formation of Golgi ministacks in the vicinity of ERES re-establishes ER–Golgi communication as a short range process (as in plants), explaining the ongoing Golgi modification and secretion of proteins (Saraste and Svensson, 1991; Cole *et al.*, 1996; Storrie *et al.*, 1998).

The Rho family GTPase cdc42, which regulates actin dynamics, interacts with COPI coats and affects dynein function, suggesting functional coupling between the actin filament system and MT-dependent motility of IC/*cis*-Golgi carriers (Luna *et al.*, 2002; Chen *et al.*, 2005). Similarly, WHAMM, which promotes actin nucleation and interacts with MTs, has been localized to the IC and implicated in the

formation and transport of pre-Golgi tubules (Campellone *et al.*, 2008).

Different IC Models

Transport Complexes

Based on imaging of fluorescent cargo, such as the VSV G-protein and procollagen, in living cells (Presley *et al.*, 1997; Scales *et al.*, 1997; Stephens and Pepperkok, 2002), it has been proposed that the IC represents a collection of large, pleiomorphic transport complexes (TCs) (Bannykh and Balch, 1997; Stephens and Pepperkok, 2001), which form via homotypic fusion of COPII vesicles, or protrude directly from the ER (Mironov *et al.*, 2003; Xu and Hay, 2004; Yu *et al.*, 2006). Thereafter, they move in a MT- and dynein-dependent (Burkhardt *et al.*, 1997) ‘stop-and-go’ fashion to the Golgi region, where they either fuse with or transform into *cis*-Golgi cisternae (Figure 7(a)). In other words, the IC elements are transient structures which are first formed *de novo* at ERES and then consumed at *cis*-Golgi. These mobile structures (speed of

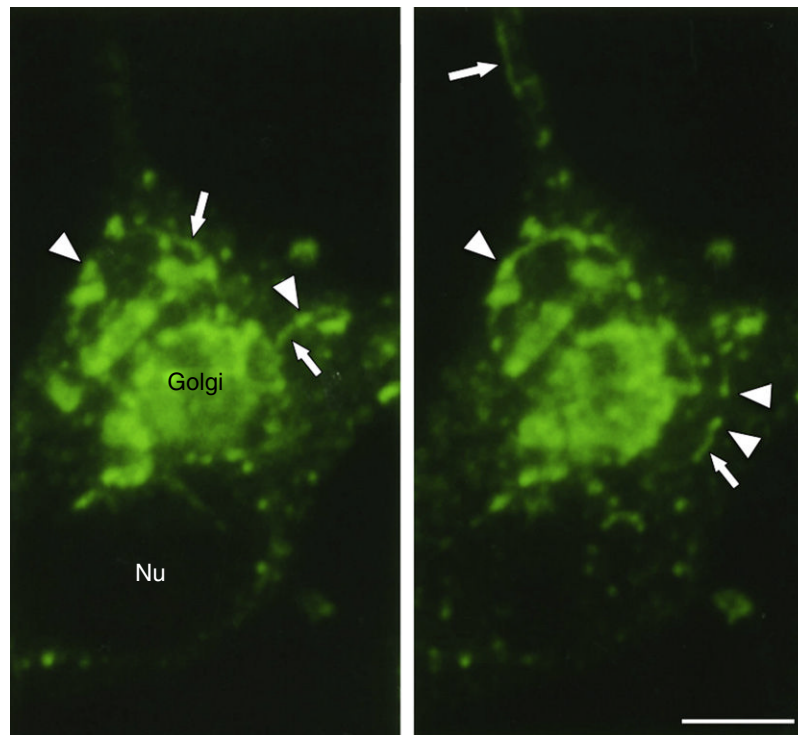


Figure 5 Immunofluorescence LM localization of Rab1 in differentiating PC12 cells, illustrating the interconnected globular (arrowheads) and tubular (arrows) domains of the IC. Two confocal sections of the same cell are shown. The IC has expanded in response to a 24 h treatment with the nerve growth factor (NGF) and the tubules are found both in the cell body and the neurite-like extensions (right panel). Nu, nucleus; Golgi, perinuclear Golgi region. Bar: 5 μ m.

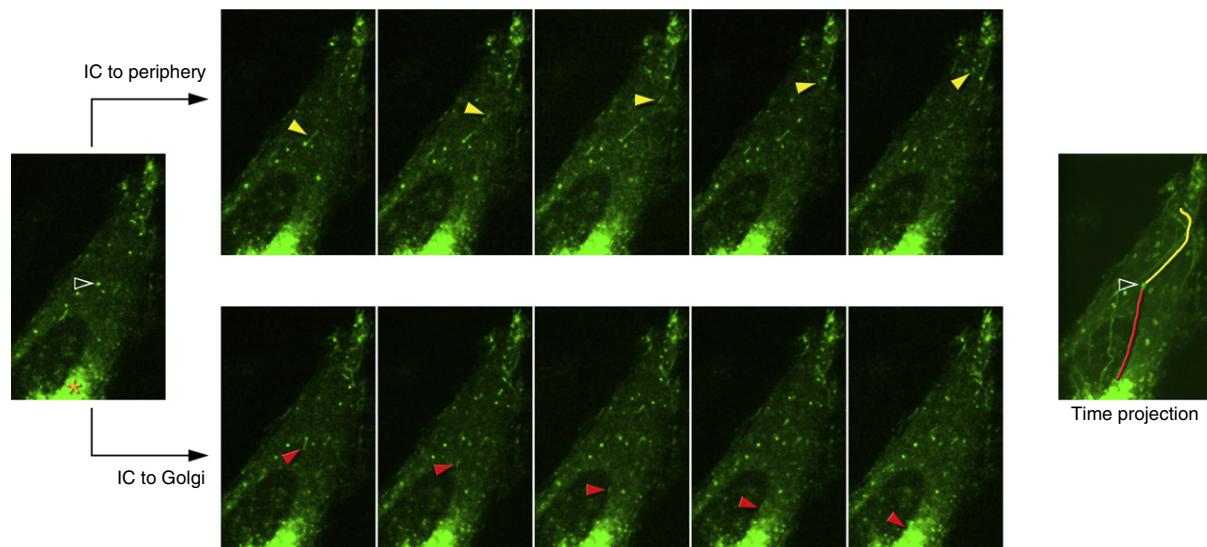


Figure 6 Bidirectional movements of GFP-Rab1A-positive tubules in a living NRK cell. Selected confocal images from a movie obtained by time-lapse microscopy are shown. Tubules extending from the same relatively stationary globular IC element (black arrowhead) pinch off and move either in the direction of the cell's leading edge (upper panels; yellow arrowheads), or toward the Golgi indicated by an asterisk (lower panels; red arrowheads). The paths taken by the tubules are highlighted in the time projection on the right. Adapted from Sannerud, R., Marie, M., Nizak, C., *et al.*, 2006. Rab1 defines a novel pathway connecting the pre-Golgi intermediate compartment with the cell periphery. *Molecular Biology of the Cell* 17, 1514–1526.

about 1 μ m/sec) corresponding to saccules or more complex, pleiomorphic elements (Mironov *et al.*, 2003) appear to consist of subdomains enriched in antero- or retrograde cargo, or

COPI (Shima *et al.*, 1999; Stephens *et al.*, 2000). They have also been shown to contain other machinery proteins, such as p23/24, p58, VIP36, membrin and Rab1A (Blum *et al.*, 1999;

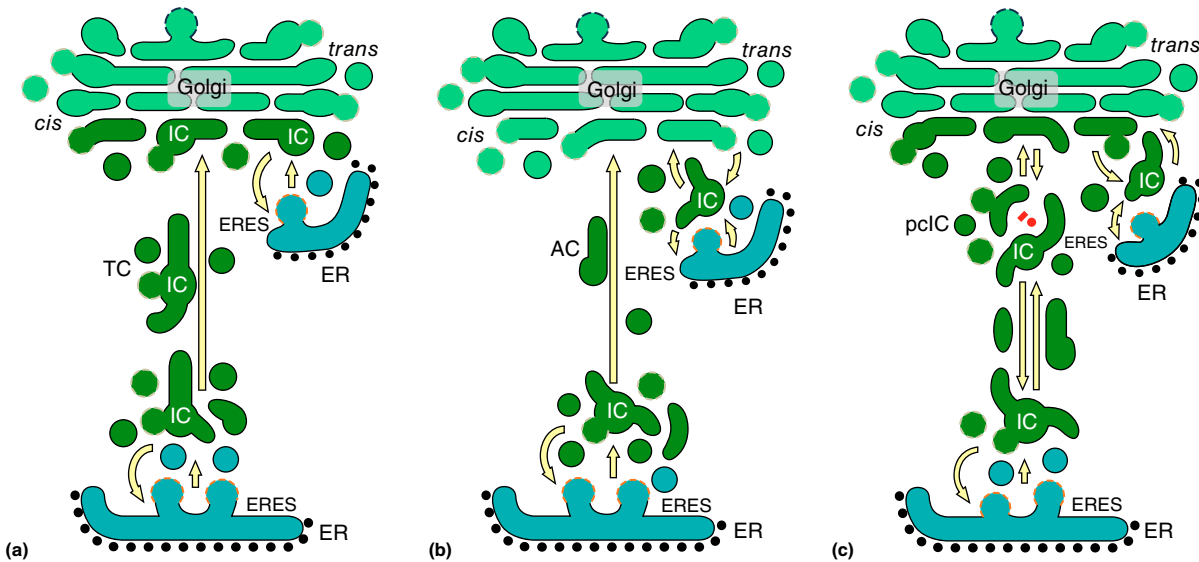


Figure 7 Three views of the IC. All models take into account the existence of two types of ERES in mammalian cells, peripheral and Golgi-adjacent. For simplicity, lateral communication between ERES-adjacent IC elements is not illustrated. (a) The IC elements form *de novo* at ERES (via homotypic fusion of COPII vesicles), giving rise to pleiomorphic transport complexes (TC) that move along MTs to the *cis*-Golgi region where they either transform into *cis*-Golgi cisternae (as shown here) or fuse with the stationary Golgi stacks. (b) The IC consists of stationary membrane clusters located close to ERES. In this case ER-to-Golgi trafficking involves two distinct transport steps, since the IC receives cargo from the ER via heterotypic fusion of COPII vesicles and forms special anterograde carriers (AC) that move along MTs to *cis*-Golgi. (c) The IC represents a stable interconnected network that is anchored both next to ERES and the centrosome. Bidirectional transport between these sites and between the central IC elements and the Golgi stacks involves vesicular, tubular or saccular carriers. ERES-IC communication via COPII vesicles occurs as in model b. COPI, COPII and clathrin coats are shown in gray, orange, and blue, respectively, and the centrosome in red.

Chao *et al.*, 1999; Dahm *et al.*, 2001; Sannerud *et al.*, 2006; Monetta *et al.*, 2007; Tomás *et al.*, 2010).

Stationary Compartment

Another model of the IC (Figure 7(b)) is largely based on the imaging of GFP-ERGIC-53 dynamics in living cells (Ben-Tekaya *et al.*, 2005). It proposes that the IC consists of stationary, long-lived membrane clusters, located close to the ERES, which communicate with the ER and Golgi via distinct transport carriers (Appenzeller-Herzog and Hauri, 2006). Thus, in a two-step transport process cargo is first transferred from ER to the IC via heterotypic fusion of COPII vesicles. The IC then generates novel (possibly COPI-coated) anterograde carriers (ACs) that move along MTs to the *cis*-Golgi. The recycling of GFP-ERGIC-53 to the ER evidently occurs mainly from the ERES-associated IC, since it was not detected in the ACs containing soluble, Golgi-destined cargo. The identification of a motif in the neuronal GABA transporter, that seems to be required for its exit from the IC, supports the stable nature of the IC (Farhan *et al.*, 2008).

Interconnected Membrane System

Visualization of IC dynamics using GFP-Rab1A showed that the pleiomorphic transport carriers arriving along MTs from the cell periphery do not move directly to *cis*-Golgi, but are targeted to the MTOC/centrosome that is normally positioned next to the Golgi ribbon. In cells displaying centrosome motility, for example, cells that are migrating or entering mitosis,

the centrosome-targeted membranes can be resolved as a separate compartment, called the pericentrosomal IC (pcIC), which is distinct from the *cis*-Golgi-adjacent IC domain (Marie *et al.*, 2009, 2012; Mochizuki *et al.*, 2013; Figure 3(a)). Live imaging further showed that the separated pcIC and the Golgi communicate via tubular and globular carriers. The pcIC contains its own pool of COPI coats, and mediates the BFA-induced, COPI-independent backflow of Golgi enzymes to the ER, suggesting that forward pcIC-to-Golgi transport depends on COPI coats, and thus is blocked by BFA (Marie *et al.*, 2009).

Both the peripheral IC elements and the pcIC persist upon Golgi disassembly by BFA and maintain their communication with via dynamic tubules (Marie *et al.*, 2009). Accordingly, the IC has been proposed to constitute a dynamic, but stable membrane network due to its anchoring next to the ERES and the centrosome (Saraste *et al.*, 2009; Figure 7(c)). This model is supported by studies of mitotic cells, showing that the IC persists despite Golgi breakdown and the rearrangement of the MTs, and maintains its spatial organization due to its ongoing association with the spindle MTs and the centrosomes at the spindle poles (Marie *et al.*, 2012; Figure 3(a)).

Functional Aspects

Sorting and Transport

In the endocytic pathway, soluble proteins and particles bound to lysosomes accumulate in the lumen of vacuolar endosomes, while many membrane proteins are sorted into narrow tubules for recycling to the PM. The observed major

concentration of soluble secretory proteins within the IC (Oprins *et al.*, 2001) could be explained by similar geometry-based sorting, namely, their exclusion from vesicles and tubules, which recycle lipids and membrane-bound proteins back to the ER (Martínez-Menárguez *et al.*, 1999). Membrane recycling is a major function of both the IC and endosomes in mammalian cells. Due to the similarity of the tubular domain of the IC with the endocytic recycling compartment (ERC), its 'mirror compartment' next to the centrosome, it has been designated as the biosynthetic recycling compartment (BRC; Saraste and Goud, 2007; Marie *et al.*, 2009).

Luminal conditions: Receptor-mediated retrieval of KDEL proteins from the IC requires that it is discontinuous with the ER and maintains special luminal conditions (Pelham, 1989). The finding of low pH-dependent binding of KDEL-ligands to their receptor *in vitro* suggested the existence of a pH gradient between the ER and *cis*-Golgi (Wilson *et al.*, 1993). The pH-sensitive interactions of the low density lipoprotein (LDL)-receptor-related protein (LRP) and procollagen with the chaperones RAP and Hsp47, respectively (Bu *et al.*, 1995; Satoh *et al.*, 1996), and partial co-localization of DAMP (a marker for acidic compartments) and p58/ERGIC-53 in central IC elements (Palokangas *et al.*, 1998), are in accordance with this idea. The pH of the ER and *cis*-Golgi has been estimated to be about 7.2 and 6.5–6.7, respectively (Paroutis *et al.*, 2004; Vavassori *et al.*, 2013).

Although the effect pH on the binding of mannose-containing cargo to p58/ERGIC-53/LMAN1 remains unclear, it appears that cargo release is caused by a drop in free calcium concentration between the ER and IC lumen (Appenzeller-Herzog *et al.*, 2004; Bentley *et al.*, 2010; Zheng *et al.*, 2013). On the other hand, the depletion of luminal calcium stores affects the morphology of the IC and the recycling of cargo receptors at the ER–Golgi boundary, and the IC has been reported to contain the calcium-ATPase SERCA, as well as the calcium-binding proteins BiP, GRP94, CR, and CALNUP (Ying *et al.*, 2002; Breuza *et al.*, 2004; Howe *et al.*, 2009; Bentley *et al.*, 2010), suggesting its role in intracellular calcium storage.

Role of COPI coats: There are at least three subtypes of COPI coats (Beck *et al.*, 2009), and four Arf GTPases that regulate their membrane binding (Popoff *et al.*, 2011). Three Arfs appear to act at the ER–Golgi boundary and two of these associate with membranes in a BFA-resistant manner (Volpicelli-Daley *et al.*, 2005; Chun *et al.*, 2008; Duijsings *et al.*, 2009; Ben-Tekaya *et al.*, 2010; Hamlin *et al.*, 2014), suggesting that different types of COPI vesicles mediate two-way trafficking at the level of the IC. The role of COPI in anterograde transport has been considered for some time (Hosobuchi *et al.*, 1992; Pepperkok *et al.*, 1993; Peter *et al.*, 1993; Orci *et al.*, 1997; Malsam *et al.*, 2005). In addition, although both p58/ERGIC-53 and KDEL-receptor employ COPI vesicles in their recycling, their transport itineraries differ (Tang *et al.*, 1995a; Marie *et al.*, 2009, 2012). In addition to acting in vesicle budding the COPI coats may form membrane domains (Presley *et al.*, 2002).

Golgi bypass: Besides mediating bidirectional ER–Golgi trafficking, the IC has been suggested to be involved in Golgi-independent pathway(s) (Marie *et al.*, 2008; Saraste *et al.*, 2009). The route between the IC and the leading edge or growth cone of motile fibroblasts or neurons, respectively (Figures 5 and 6), could participate in cholesterol and integrin

trafficking (Urbani and Simoni, 1990; Sannerud *et al.*, 2006; Wang *et al.*, 2010) or correspond to the BFA-resistant ER-IC-PM route that supports phagocytosis (Becker *et al.*, 2005; Saraste and Goud, 2007; see below). Direct pericentrosomal communication between the IC and the endosomal system seems to be used by the cystic fibrosis transmembrane conductance regulator (CFTR) during its Golgi-independent transport to the cell surface (Yoo *et al.*, 2002; Marie *et al.*, 2009; Gee *et al.*, 2011). Further, the presence of the IC in neuronal dendrites may provide a Golgi-independent satellite pathway for local dendritic trafficking (Krijnse-Locker *et al.*, 1995; Sannerud *et al.*, 2006; Ehlers, 2013).

Posttranslational Modification

Questioning the studies on coronavirus maturation, which suggested that O-glycosylation is initiated in the IC (see above), subsequent EM studies showed that the GalNAc-transferases are predominantly found in the Golgi. However, the activation of cells (by epidermal growth factor) causes their incorporation into COPI vesicles and relocation to the IC and ER, which consequently become positive for the lectin Helix pomatia (Gill *et al.*, 2010). The cycling of *cis*-Golgi proteins to the IC could be a constitutive event (Lin *et al.*, 1999; Marra *et al.*, 2001; Jarvela and Linstead, 2012). Moreover, the construction of the sugar chains on proteoglycans has been suggested to begin in the IC (Jönsson *et al.*, 2003).

Protein Maturation

The KDEL-containing molecular chaperones present in the IC (BiP, PDI, and GRP94; see above) could be cycling while still bound to their unfolded client proteins, opening for a post-ER level of protein folding and quality control. The presence of quality control machinery in the IC and the finding that the ER-associated degradation of certain proteins requires their ER export is in accordance with this idea (Zuber *et al.*, 2001; Anelli and Sitia, 2008). The PDI family member Erp44 and p58/ERGIC-53 cooperate in the IC/*cis*-Golgi in sequential assembly of IgM polymers (Anelli and Sitia, 2008). Unassembled IgM or T-cell antigen receptor subunits bound to Erp44 or BiP, respectively, can be retrieved to the ER by the KDEL-receptor in a pH-dependent manner (Yamamoto *et al.*, 2001; Vavassori *et al.*, 2013), similarly as the overexpressed, misfolded VSV G-protein (Hammond and Helenius, 1994) and mutant V2 vasopressin receptors (Hermosilla *et al.*, 2004). Additional proof for a post-ER checkpoint is provided by studies showing the accumulation of the deletion mutant Δ F508 of CFTR (Gilbert *et al.*, 1998), misfolded MHC class I proteins (Hsu *et al.*, 1991; Raposo *et al.*, 1995), and proinsulin (Zuber *et al.*, 2004) in expanded IC elements.

Signaling

The level of p58/ERGIC-53 mRNA is up-regulated by the unfolded protein response (UPR), which also requires yeast Ypt1 function, supporting a link between this signaling pathway and ER–Golgi trafficking (Nyfeler *et al.*, 2003; Tsvetanova *et al.*, 2012). Protein kinases Src, aPKC, and Scyl1 have been

implicated in COPI function at the level of the IC (Tisdale and Artalejo, 2006; Hamlin *et al.*, 2014) and PKC and its downstream effectors appear to control IC morphology (Ben-Tekaya *et al.*, 2010; Sugawara *et al.*, 2012). The activation of neuronal Trk receptor tyrosine kinases in the IC initiates signaling via the MEK pathway leading to Golgi fragmentation (Scheeterson *et al.*, 2010).

Autophagy

Membranes enriched in IC markers (p58/ERGIC-53/LMAN1, KDEL-receptor, and Sec22B) play a role in the biogenesis of autophagosomes by representing the primary membrane source for the lipidation of LC3, which triggers this process (Ge *et al.*, 2013). Moreover, Ypt1/Rab1 is a key regulator of autophagy in yeast and mammals (Lynch-Day *et al.*, 2010; Winslow *et al.*, 2010; Zoppino *et al.*, 2010; Huang *et al.*, 2011; Lipatova *et al.*, 2012). A phosphatidylinositol 3-phosphatase (MTMR6) acting in vesicle transport and autophagy is regulated by Rab1B and localizes predominantly to the pIC (Mochizuki *et al.*, 2013), further supporting the role of the IC in autophagy.

Phagocytosis and Antigen Presentation

Supporting the existence of an unconventional pathway that connects the ER with the PM-derived phagosome (Gagnon *et al.*, 2002), the IC-enriched SNARE Sec22B/ERS-24 (Zhang *et al.*, 1999) was shown to influence phagocytosis independently of its role in ER–Golgi trafficking (Becker *et al.*, 2005; Hatsuzawa *et al.*, 2009). In dendritic cells the delivery of certain proteins – such as the peptide transporter (TAP), CR and tapasin – from ER to the phagosome is required for antigen cross-presentation. This pathway depends on the interaction between Sec22B and the PM SNARE syntaxin 4 and involves efficient recruitment of the integral IC components sec22B and p58/ERGIC-53 to the phagosomes (Cebrian *et al.*, 2011), in accordance with the idea that the IC provides an important membrane source for their formation. The presence of CR, tapasin, and functional TAP in the IC/cis-Golgi support this possibility (Kleijmeer *et al.*, 1992; Howe *et al.*, 2009; Ghanem *et al.*, 2010).

Links to Disease

The IC plays a role in the Golgi-independent trafficking of CFTR (see above). Mutations in p58/ERGIC-53/LMAN1, or its partner MCFD2 (multiple coagulation factor deficiency protein 2) result in an autosomal recessive bleeding disorder called combined deficiency of coagulation factors V and VIII. They disrupt the lectin mannose-binding protein 1 (LMAN1)-MCFD2 receptor complex, thereby inhibiting the secretion of these factors and reducing their serum levels (Nichols *et al.*, 1998; Zheng and Zhang, 2013). Parkinson's disease-related cytotoxic protein, α -synuclein, has been suggested to interfere with ER–Golgi trafficking and to arrest autophagy by inhibiting Rab1 function (Cooper *et al.*, 2006; Winslow *et al.*, 2010). Many phagocytosed bacterial pathogens, such as

Legionella, hijack the ER-to-IC-to-phagosome pathway during their intracellular replication (Isberg *et al.*, 2009; Arasaki *et al.*, 2012; see above). In addition to coronaviruses, the IC has been implicated in the replication of bunya-, entero-, flavi-, picorna, and vacciniaviruses (Jääntti *et al.*, 1997; Risco *et al.*, 2002; Miller and Krijnse-Locker, 2008; Hsu *et al.*, 2010). Surprisingly, p58/ERGIC-53 interacts in a lectin-independent manner with fusion proteins of a number of membrane viruses and participates in different stages of their life cycle (Klaus *et al.*, 2013).

See also: Interorganellar Communication: Interplay and Processes: Endoplasmic Reticulum Stress in Disease; ER–Golgi Transport; Regulation of the Secretory Pathway; Unconventional Protein Secretion: Fibroblast Growth Factor 2 and Interleukin-1 β as Examples. Intracellular Infectiology: Cell Processes: Phagocytosis. Organelles: Structure and Function: At the Center of Autophagy: Autophagosomes; Golgi and TGN; The Endoplasmic Reticulum

References

- Allan, B.B., Moyer, B.D., Balch, W.E., 2000. Rab1 recruitment of p115 into a *cis*-SNARE complex: Programming budding COPII vesicles for fusion. *Science* 289, 444–448.
- Alvarez, C., Fujita, H., Hubbard, A., Sztul, E., 1999. ER to Golgi transport: Requirement for p115 at a pre-Golgi VTC stage. *Journal of Cell Biology* 147, 1205–1222.
- Alvarez, C., Garcia-Mata, R., Brandon, E., Sztul, E., 2003. COPI recruitment is modulated by a Rab1b-dependent mechanism. *Molecular Biology of the Cell* 14, 2116–2127.
- Alvarez, C., Garcia-Mata, R., Hauri, H.-P., Sztul, E., 2001. The p115-interactive proteins GM130 and giantin participate in endoplasmic reticulum-Golgi traffic. *Journal of Biological Chemistry* 276, 2693–2700.
- Anelli, T., Sitia, R., 2008. Protein quality control in the early secretory pathway. *EMBO Journal* 27, 315–327.
- Appenzeller, C., Andersson, H., Kappeler, F., Hauri, H.P., 1999. The lectin ERGIC-53 is a cargo transport receptor for glycoproteins. *Nature Cell Biology* 1, 330–334.
- Appenzeller-Herzog, C., Hauri, H.P., 2006. The ER-Golgi intermediate compartment (ERGIC): In search of its identity and function. *Journal of Cell Science* 119, 2173–2183.
- Appenzeller-Herzog, C., Roche, A.C., Nufer, O., Hauri, H.P., 2004. pH-induced conversion of the transport lectin ERGIC-53 triggers glycoprotein release. *Journal of Biological Chemistry* 279, 12943–12950.
- Arar, C., Carpentier, V., Le Caer, J.P., *et al.*, 1995. ERGIC-53, a membrane protein of the endoplasmic reticulum-Golgi intermediate compartment, is identical to MR60, an intracellular mannose-specific lectin of myelomonocytic cells. *Journal of Biological Chemistry* 270, 3551–3553.
- Arasaki, K., Toomre, D.K., Roy, C.R., 2012. The *Legionella pneumophila* effector DrrA is sufficient to stimulate SNARE-dependent membrane fusion. *Cell Host & Microbe* 11, 46–57.
- Aridor, M., Balch, W.E., 1996. Principles of selective transport: Coat complexes hold the key. *Trends in Cell Biology* 6, 315–320.
- Bailly, E., McCaffrey, M., Touchot, N., *et al.*, 1991. Phosphorylation of two small GTP-binding proteins of the Rab family by p34^{cdc2}. *Nature* 350, 715–718.
- Balch, W.E., Elliott, M.M., Keller, D.D., 1986. ATP-coupled transport of vesicular stomatitis virus G protein between the endoplasmic reticulum and Golgi. *Journal of Biological Chemistry* 261, 14681–14689.
- Balch, W.E., McCaffrey, J.M., Plutner, H., Farquhar, M.G., 1994. Vesicular stomatitis virus glycoprotein is sorted and concentrated during export from the endoplasmic reticulum. *Cell* 76, 841–852.
- Bannykh, S.I., Balch, W.E., 1997. Membrane dynamics at the endoplasmic reticulum-Golgi interface. *Journal of Cell Biology* 138, 1–4.
- Bannykh, S.I., Rowe, T., Balch, W.E., 1996. The organization of endoplasmic reticulum export complexes. *Journal of Cell Biology* 135, 19–35.

- Barlowe, C.K., Miller, E.A., 2013. Secretory protein biogenesis and traffic in the early secretory pathway. *Genetics* 193, 383–410.
- Barnekow, A., Thyrock, A., Kessler, D., 2009. Rab proteins and their interaction partners. *International Review of Cell and Molecular Biology* 274, 235–274.
- Beck, R., Rawet, M., Wieland, F.T., Cassel, D., 2009. The COPI system: Molecular mechanisms and function. *FEBS Letters* 583, 2701–2709.
- Becker, T., Volchuk, A., Rothman, J.E., 2005. Differential use of endoplasmic reticulum membrane for phagocytosis in J774 macrophages. *Proceedings of the National Academy of Sciences of the United States of America* 102, 4022–4026.
- Ben-Tekaya, H., Kahn, R.A., Hauri, H.P., 2010. ADP ribosylation factors 1 and 4 and group VIA phospholipase A₂ regulate morphology and intraorganellar traffic in the endoplasmic reticulum-Golgi intermediate compartment. *Molecular Biology of the Cell* 21, 4130–4140.
- Ben-Tekaya, H., Miura, K., Pepperkok, R., Hauri, H.P., 2005. Live imaging of bidirectional traffic from the ERGIC. *Journal of Cell Science* 118, 357–367.
- Bentley, M., Nycz, D.C., Joglekar, A., *et al.*, 2010. Vesicular calcium regulates coat retention, fusogenicity, and size of pre-Golgi intermediates. *Molecular Biology of the Cell* 21, 1033–1046.
- Blum, R., Pfeiffer, F., Feick, P., *et al.*, 1999. Intracellular localization and *in vivo* trafficking of p24A and p23. *Journal of Cell Science* 112, 537–548.
- Blum, R., Stephens, D.J., Schulz, I., 2000. Luminal targeted GFP, used as a marker of soluble cargo, visualises rapid ERGIC to Golgi traffic by a tubulo-vesicular network. *Journal of Cell Science* 113, 3151–3159.
- Bonatti, S., Migliaccio, G., Simons, K., 1989. Palmitoylation of viral membrane glycoproteins takes place after exit from the endoplasmic reticulum. *Journal of Biological Chemistry* 264, 12590–12595.
- Brandizzi, F., Barlowe, C., 2013. Organization of the ER-Golgi interface for membrane traffic control. *Nature Reviews Molecular Cell Biology* 14, 382–392.
- Breuz, L., Halbeisen, R., Jenö, P., *et al.*, 2004. Proteomics of endoplasmic reticulum-Golgi intermediate compartment (ERGIC) membranes from brefeldin A-treated HepG2 cells identifies ERGIC-32, a new cycling protein that interacts with human Erv46. *Journal of Biological Chemistry* 279, 47242–47253.
- Bu, G., Geuze, H.J., Strous, G.J., Schwartz, A.L., 1995. Binding and endocytosis of 39 kDa protein by MDBK cells. *EMBO Journal* 14, 2269–2280.
- Burkhardt, J.K., Echeverri, C.J., Nilsson, T., Vallee, R.B., 1997. Overexpression of the dynamin (p50) subunit of the dynactin complex disrupts dynein-dependent maintenance of membrane organelle distribution. *Journal of Cell Biology* 139, 469–484.
- Cai, H., Reinisch, K., Ferro-Novick, S., 2007. Coats, tethers, Rabs, and SNAREs work together to mediate the intracellular destination of a transport vesicle. *Developmental Cell* 12, 671–682.
- Campellone, K.G., Webb, N.J., Znameroski, E.A., Welch, M.D., 2008. WHAMM is an Arp2/3 complex activator that binds microtubules and functions in ER to Golgi transport. *Cell* 134, 148–161.
- Cebrian, I., Visentin, G., Blanchard, N., *et al.*, 2011. Sec22b regulates phagosomal maturation and antigen crosspresentation by dendritic cells. *Cell* 147, 1355–1368.
- Chao, D.S., Hay, J.C., Winnick, S., *et al.*, 1999. SNARE membrane trafficking dynamics *in vivo*. *Journal of Cell Biology* 144, 869–881.
- Chavrier, P., Parton, R.G., Hauri, H.-P., Simons, K., Zerial, M., 1990. Localization of low molecular weight GTP binding proteins to exocytic and endocytic compartments. *Cell* 62, 317–329.
- Chen, J.L., Cucini, R.V., Lacomis, L., *et al.*, 2005. Coatamer-bound Cdc42 regulates dynein recruitment to COPI vesicles. *Journal of Cell Biology* 169, 383–389.
- Chun, J., Shapovalova, Z., Dejgaard, S.Y., Presley, J.F., Melançon, P., 2008. Characterization of class I and class II ADP-ribosylation factors in live cells: GDP-bound class II Arfs associate with the ER-Golgi intermediate compartment independently of GBF1. *Molecular Biology of the Cell* 19, 3488–3500.
- Cole, N.B., Sciaky, N., Marotta, A., Song, J., Lippincott-Schwartz, J., 1996. Golgi dispersal during microtubule disruption: Regeneration of Golgi stacks at peripheral endoplasmic reticulum exit sites. *Molecular Biology of the Cell* 7, 631–650.
- Connolly, C.N., Futter, C.E., Gibson, A., Hopkins, C.R., Cutler, D.F., 1994. Transport into and out of the Golgi complex studied by transfecting cells with cDNAs encoding horseradish peroxidase. *Journal of Cell Biology* 127, 641–652.
- Cooper, A.A., Gitler, A.D., Cashikar, A., *et al.*, 2006. Alpha-synuclein blocks ER-Golgi traffic and Rab1 rescues neuron loss in Parkinson's models. *Science* 313, 324–328.
- Cosson, P., Letourneur, F., 1997. Coatamer (COPI)-coated vesicles: Role in intracellular transport and protein sorting. *Current Opinion in Cell Biology* 9, 484–487.
- Dahm, T., White, J., Grill, S., Füllekrug, J., Stelzer, E.H., 2001. Quantitative ER ↔ Golgi transport kinetics and protein separation upon Golgi exit revealed by vesicular integral membrane protein 36 dynamics in live cells. *Molecular Biology of the Cell* 12, 1481–1498.
- Day, K.J., Staehelin, L.A., Glick, B.S., 2013. A three-stage model of Golgi structure and function. *Histochemistry and Cell Biology* 140, 239–249.
- Diao, A., Rahman, D., Pappin, D.J.C., Lucocq, J., Lowe, M., 2003. The coiled-coil membrane protein golgin-84 is a novel Rab effector required for Golgi ribbon formation. *Journal of Cell Biology* 160, 201–212.
- Dittmer, F., von Figura, K., 1999. Phosphorylation of arylsulphatase A occurs through multiple interactions with the UDP-N-acetylglucosamine-1-phosphotransferase proximal and distal to its retrieval site by the KDEL receptor. *Biochemical Journal* 340, 729–736.
- Dominguez, M., Dejgaard, K., Füllekrug, J., *et al.*, 1998. gp25L/emp24/p24 protein family members of the cis-Golgi network bind both COPI and COPII coatamer. *Journal of Cell Biology* 140, 751–765.
- Duden, R., Griffiths, G., Frank, R., Argos, P., Kreis, T.E., 1991. Beta-COP, a 110 kD protein associated with non-clathrin-coated vesicles and the Golgi complex, shows homology to beta-adaptin. *Cell* 64, 649–655.
- Duijings, D., Lanke, K.H., van Dooren, S.H., *et al.*, 2009. Differential membrane association properties and regulation of class I and class II Arfs. *Traffic* 10, 316–323.
- Ehlers, M.D., 2013. Dendritic trafficking for neuronal growth and plasticity. *Biochemical Society Transactions* 41, 1365–1382.
- Emery, G., Parton, R.G., Rojo, M., Gruenberg, J., 2003. The transmembrane protein p25 forms highly specialized domains that regulate membrane composition and dynamics. *Journal of Cell Science* 116, 4821–4832.
- Fan, J.Y., Roth, J., Zuber, C., 2003. Ultrastructural analysis of transitional endoplasmic reticulum and pre-Golgi intermediates: A highway for cars and trucks. *Histochemistry and Cell Biology* 120, 455–463.
- Farhan, H., Reiterer, V., Kriz, A., *et al.*, 2008. Signal-dependent export of GABA transporter 1 from the ER-Golgi intermediate compartment is specified by a C-terminal motif. *Journal of Cell Science* 121, 753–761.
- Fiedler, K., Simons, K., 1994. A putative novel class of animal lectins in the secretory pathway homologous to leguminous lectins. *Cell* 77, 625–626.
- Friend, D.S., Murray, M.J., 1965. Osmium impregnation of the Golgi apparatus. *American Journal of Anatomy* 117, 135–149.
- Füllekrug, J., Sönnichsen, B., Schäfer, U., *et al.*, 1997. Characterization of brefeldin A-induced vesicular structures containing cycling proteins of the intermediate compartment/cis-Golgi network. *FEBS Letters* 404, 75–81.
- Gagnon, E., Duclos, S., Rondeau, C., *et al.*, 2002. Endoplasmic reticulum-mediated phagocytosis is a mechanism of entry into macrophages. *Cell* 110, 119–131.
- Garcia-Mata, R., Szul, T., Alvarez, C., Szul, E., 2003. ADP-ribosylation factor/COPI-dependent events at the endoplasmic reticulum-Golgi interface are regulated by the guanine nucleotide exchange factor GBF1. *Molecular Biology of the Cell* 14, 2250–2261.
- Ge, L., Melville, D., Zhang, M., Schekman, R., 2013. The ER-Golgi intermediate compartment is a key membrane source for the LC3 lipidation step of autophagosome biogenesis. *Elife* 2, e00947. doi:10.7554/eLife.00947.
- Gee, H.Y., Noh, S.H., Tang, B.L., Kim, K.H., Lee, M.G., 2011. Rescue of ΔF508-CFTR trafficking via a GRASP-dependent unconventional secretion pathway. *Cell* 146, 746–760.
- Ghanem, E., Fritzsche, S., Al-Balushi, M., *et al.*, 2010. The transporter associated with antigen processing (TAP) is active in a post-ER compartment. *Journal of Cell Science* 123, 4271–4279.
- Gilbert, A., Jadot, M., Leontieva, E., *et al.*, 1998. Delta F508 CFTR localizes in the endoplasmic reticulum-Golgi intermediate compartment in cystic fibrosis cells. *Experimental Cell Research* 242, 144–152.
- Gill, D.J., Chia, J., Senewiratne, J., Bard, F., 2010. Regulation of O-glycosylation through Golgi-to-ER relocation of initiation enzymes. *Journal of Cell Biology* 189, 843–858.
- Gommel, D., Orci, L., Emig, E.M., *et al.*, 1999. p24 and p23, the major transmembrane proteins of COPI-coated transport vesicles, form hetero-oligomeric complexes and cycle between the organelles of the early secretory pathway. *FEBS Letters* 447, 179–185.
- Griffiths, G., Ericsson, M., Krijnse-Locker, J., *et al.*, 1994. Localization of the Lys, Asp, Glu, Leu tetrapeptide receptor to the Golgi complex and the intermediate compartment in mammalian cells. *Journal of Cell Biology* 127, 1557–1574.
- Griffiths, G., Pepperkok, R., Locker, J.K., Kreis, T.E., 1995. Immunocytochemical localization of beta-COP to the ER-Golgi boundary and the TGN. *Journal of Cell Science* 108, 2839–2856.
- Hamlin, J.N., Schroeder, L.K., Fotouhi, M., *et al.*, 2014. Scyl1 scaffolds class II Arfs to specific subcomplexes of coatamer through the γ-COP appendage domain. *Journal of Cell Science* 127, 1454–1463.

- Hammond, A.T., Glick, B.S., 2000. Dynamics of transitional endoplasmic reticulum sites in vertebrate cells. *Molecular Biology of the Cell* 11, 3013–3030.
- Hammond, C., Helenius, A., 1994. Quality control in the secretory pathway: Retention of a misfolded viral membrane glycoprotein involves cycling between the ER, intermediate compartment, and Golgi apparatus. *Journal of Cell Biology* 126, 41–52.
- Hatsuzawa, K., Hashimoto, H., Hashimoto, H., *et al.*, 2009. Sec22b is a negative regulator of phagocytosis in macrophages. *Molecular Biology of the Cell* 20, 4435–4443.
- Haubruck, H., Prange, R., Vorgias, C., Gallwitz, D., 1989. The ras-related mouse ypt1 protein can functionally replace the YPT1 gene product in yeast. *EMBO Journal* 8, 1427–1432.
- Hauri, H.P., Kappeler, F., Andersson, H., Appenzeller, C., 2000. ERGIC-53 and traffic in the secretory pathway. *Journal of Cell Science* 113, 587–596.
- Hay, J.C., Klumperman, J., Oorschot, V., *et al.*, 1998. Localization, dynamics, and protein interactions reveal distinct roles for ER and Golgi SNAREs. *Journal of Cell Biology* 141, 1489–1502.
- Heino, S., Lusa, S., Somerharju, P., *et al.*, 2000. Dissecting the role of the Golgi complex and lipid rafts in biosynthetic transport of cholesterol to the cell surface. *Proceedings of the National Academy of Sciences of the United States of America* 97, 8375–8380.
- Hermosilla, R., Oueslati, M., Donalies, U., *et al.*, 2004. Disease-causing V(2) vasopressin receptors are retained in different compartments of the early secretory pathway. *Traffic* 5, 993–1005.
- Horstman, H., Ng, C.P., Tang, B.L., Hong, W., 2002. Ultrastructural characterization of endoplasmic reticulum-Golgi transport containers (EGTC). *Journal of Cell Science* 115, 4263–4273.
- Hosobuchi, M., Kreis, T., Schekman, R., 1992. SEC21 is a gene required for ER to Golgi protein transport that encodes a subunit of a yeast coatomer. *Nature* 360, 603–605.
- Howe, C., Garstka, M., Al-Balushi, M., *et al.*, 2009. Calreticulin-dependent recycling in the early secretory pathway mediates optimal peptide loading of MHC class I molecules. *EMBO Journal* 28, 3730–3744.
- Hsu, N.Y., Illytska, O., Belov, G., *et al.*, 2010. Viral reorganization of the secretory pathway generates distinct organelles for RNA replication. *Cell* 141, 799–811.
- Hsu, V.W., Yuan, L.C., Nuchtern, J.G., *et al.*, 1991. A recycling pathway between the endoplasmic reticulum and the Golgi apparatus for retention of unassembled MHC class I molecules. *Nature* 352, 441–444.
- Huang, J., Birmingham, C.L., Shahnazari, S., *et al.*, 2011. Antibacterial autophagy occurs at PI(3)P-enriched domains of the endoplasmic reticulum and requires Rab1 GTPase. *Autophagy* 7, 17–26.
- Isberg, R.R., O'Connor, T.J., Heidtman, M., 2009. The *Legionella pneumophila* replication vacuole: Making a cosy niche inside host cells. *Nature Reviews Microbiology* 7, 13–24.
- Itin, C., Roche, A.-C., Monsigny, M., Hauri, H.-P., 1996. ERGIC-53 is a functional mannose-selective and calcium-dependent human homologue of leguminous lectins. *Molecular Biology of the Cell* 7, 483–493.
- Jackson, M.R., Nilsson, T., Peterson, P.A., 1990. Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *EMBO Journal* 9, 3153–3162.
- Jackson, M.R., Nilsson, T., Peterson, P.A., 1993. Retrieval of transmembrane proteins to the endoplasmic reticulum. *Journal of Cell Biology* 121, 317–333.
- Jamieson, J.D., Palade, G.E., 1967. Intracellular transport of secretory proteins in the pancreatic exocrine cell. I. Role of the peripheral elements of the Golgi complex. *Journal of Cell Biology* 34, 577–596.
- Jäntti, J., Hildén, P., Rönkä, H., *et al.*, 1997. Immunocytochemical analysis of Uukuniemi virus budding compartments: Role of the intermediate compartment and the Golgi stack in virus maturation. *Journal of Virology* 71, 1162–1172.
- Jarvela, T., Linstedt, A.D., 2012. Irradiation-induced protein inactivation reveals Golgi enzyme cycling to cell periphery. *Journal of Cell Science* 125, 973–980.
- Jönsson, M., Eklund, E., Fransson, L.A., Oldberg, A., 2003. Initiation of the decorin glycosaminoglycan chain in the endoplasmic reticulum-Golgi intermediate compartment. *Journal of Biological Chemistry* 278, 21415–21420.
- Kamena, F., Diefenbacher, M., Kilchert, C., Schwarz, H., Spang, A., 2008. Ypt1p is essential for retrograde Golgi-ER transport and for Golgi maintenance in *S. cerevisiae*. *Journal of Cell Science* 121, 1293–1302.
- Kano, F., Yamauchi, S., Yoshida, Y., *et al.*, 2009. Yip1A regulates the COPI-independent retrograde transport from the Golgi complex to the ER. *Journal of Cell Science* 122, 2218–2227.
- Kappeler, F., Klopstein, D.R., Foguet, M., Paccaud, J.P., Hauri, H.-P., 1997. The recycling of ERGIC-53 in the early secretory pathway. ERGIC-53 carries a cytosolic endoplasmic reticulum-exit determinant interacting with COPII. *Journal of Biological Chemistry* 272, 31801–31808.
- Kawamoto, K., Yoshida, Y., Tamaki, H., *et al.*, 2002. GBF1, a guanine nucleotide exchange factor for ADP-ribosylation factors, is localized to the *cis*-Golgi and involved in membrane association of the COPI coat. *Traffic* 3, 483–495.
- Klaus, J.P., Eisenhauer, P., Russo, J., *et al.*, 2013. The intracellular cargo receptor ERGIC-53 is required for the production of infectious arenavirus, coronavirus, and filovirus particles. *Cell Host & Microbe* 14, 522–534.
- Kleijmeer, M.J., Kelly, A., Geuze, H.J., *et al.*, 1992. Location of MHC-encoded transporters in the endoplasmic reticulum and *cis*-Golgi. *Nature* 357, 342–344.
- Klumperman, J., Locker, J.K., Meijer, A., *et al.*, 1994. Coronavirus M proteins accumulate in the Golgi complex beyond the site of virion budding. *Journal of Virology* 68, 6523–6534.
- Klumperman, J., Schweizer, A., Clausen, H., *et al.*, 1998. The recycling pathway of protein ERGIC-53 and dynamics of the ER-Golgi intermediate compartment. *Journal of Cell Science* 111, 3411–3425.
- Krijnse-Locker, J., Ericsson, M., Rottier, P.J., Griffiths, G., 1994. Characterization of the budding compartment of mouse hepatitis virus: Evidence that transport from the RER to the Golgi complex requires only one vesicular transport step. *Journal of Cell Biology* 124, 55–70.
- Krijnse-Locker, J., Parton, R.G., Fuller, S.D., Griffiths, G., Dotti, C.G., 1995. The organization of the endoplasmic reticulum and the intermediate compartment in cultured rat hippocampal neurons. *Molecular Biology of the Cell* 6, 1315–1332.
- Kuismanen, E., Saraste, J., 1989. Low temperature-induced transport blocks as tools to manipulate membrane traffic. *Methods in Cell Biology* 32, 257–274.
- Ladinsky, M.S., Mastronarde, D.N., McIntosh, J.R., Howell, K.E., Staehelin, L.A., 1999. Golgi structure in three dimensions: Functional insights from the normal rat kidney cell. *Journal of Cell Biology* 144, 1135–1149.
- Lahtinen, U., Dahllöf, B., Saraste, J., 1992. Characterization of a 58 kDa *cis*-Golgi protein in pancreatic exocrine cells. *Journal of Cell Science* 103, 321–333.
- Lahtinen, U., Hellman, U., Wernstedt, C., Saraste, J., Pettersson, R.F., 1996. Molecular cloning and expression of a 58-kDa *cis*-Golgi and intermediate compartment protein. *Journal of Biological Chemistry* 271, 4031–4037.
- Lavoie, C., Paiement, J., Dominguez, M., *et al.*, 1999. Roles for alpha(2)p24 and COPI in endoplasmic reticulum cargo exit site formation. *Journal of Cell Biology* 146, 285–299.
- Lazzarino, D.A., Gabel, C.A., 1988. Biosynthesis of the mannose-6-phosphate recognition marker in transport-impaired mouse lymphoma cells. Demonstration of a two-step phosphorylation. *Journal of Biological Chemistry* 263, 10118–10126.
- Lee, M.C., Miller, E.A., Goldberg, J., Orci, L., Schekman, R., 2004. Bi-directional protein transport between the ER and Golgi. *Annual Review of Cell and Developmental Biology* 20, 87–123.
- Lewis, M.J., Pelham, H.R., 1990. A human homologue of the yeast HDEL receptor. *Nature* 348, 162–163.
- Lin, C.C., Love, H.D., Gushue, J.N., Bergeron, J.J., Ostermann, J., 1999. ER/Golgi intermediates acquire Golgi enzymes by brefeldin A-sensitive retrograde transport *in vitro*. *Journal of Cell Biology* 147, 1457–1472.
- Lipatova, Z., Belogortseva, N., Zhang, X.Q., *et al.*, 2012. Regulation of selective autophagy onset by a Ypt/Rab GTPase module. *Proceedings of the National Academy of Sciences of the United States of America* 109, 6981–6986.
- Lippincott-Schwartz, J., Cole, N.B., Marotta, A., Conrad, P.A., Bloom, G.S., 1995. Kinesin is the motor for microtubule-mediated Golgi-to-ER membrane traffic. *Journal of Cell Biology* 128, 293–306.
- Lippincott-Schwartz, J., Donaldson, J.G., Schweizer, A., *et al.*, 1990. Microtubule-dependent retrograde transport of proteins into the ER in the presence of brefeldin A suggests an ER recycling pathway. *Cell* 60, 821–836.
- Lotti, L.V., Torrisi, M.R., Pascale, M.C., Bonatti, S., 1992. Immunocytochemical analysis of the transfer of vesicular stomatitis virus G glycoprotein from the intermediate compartment to the Golgi complex. *Journal of Cell Biology* 118, 43–50.
- Luna, A., Matas, O.B., Martínez-Menárguez, J.A., *et al.*, 2002. Regulation of protein transport from the Golgi complex to the endoplasmic reticulum by CDC42 and N-WASP. *Molecular Biology of the Cell* 13, 866–879.
- Lynch-Day, M.A., Bhandari, D., Menon, S., *et al.*, 2010. Trs85 directs a Ypt1 GEF, TRAPP1, to the phagophore to promote autophagy. *Proceedings of the National Academy of Sciences of the United States of America* 107, 7811–7816.
- Majoul, I., Straub, M., Hell, S.W., Duden, R., Söling, H.D., 2001. KDEL-cargo regulates interactions between proteins involved in COPI vesicle traffic: Measurements in living cells using FRET. *Developmental Cell* 1, 139–153.
- Malsam, J., Satoh, A., Pelletier, L., Warren, G., 2005. Golgin tethers define subpopulations of COPI vesicles. *Science* 307, 1095–1098.

- Manolea, F., Claude, A., Chun, J., Rosas, J., Melancon, P., 2008. Distinct functions for Arf nucleotide exchange factors at the Golgi complex: GBF1 and BIGs are required for assembly and maintenance of the Golgi stack and TGN, respectively. *Molecular Biology of the Cell* 19, 523–535.
- Marie, M., Dale, H.A., Kouprina, N., Saraste, J., 2012. Division of the intermediate compartment at the onset of mitosis provides a mechanism for Golgi inheritance. *Journal of Cell Science* 125, 5403–5416.
- Marie, M., Dale, H.A., Sannerud, R., Saraste, J., 2009. The function of the intermediate compartment in pre-Golgi trafficking involves its stable connection with the centrosome. *Molecular Biology of the Cell* 20, 4458–4470.
- Marie, M., Sannerud, R., Dale, H.A., Saraste, J., 2008. Take the 'A' train: On fast tracks to the cell surface. *Cellular and Molecular Life Sciences* 65, 2859–2874.
- Marra, P., Maffucci, T., Daniele, T., *et al.*, 2001. The GM130 and GRASP65 Golgi proteins cycle through and define a subdomain of the intermediate compartment. *Nature Cell Biology* 3, 1101–1113.
- Marra, P., Salvatore, L., Mironov Jr., A., *et al.*, 2007. The biogenesis of the Golgi ribbon: The roles of membrane input from the ER and of GM130. *Molecular Biology of the Cell* 18, 1595–1608.
- Martínez-Menárguez, J.A., Geuze, H.J., Slot, J.W., Klumperman, J., 1999. Vesicular tubular clusters between the ER and Golgi mediate concentration of soluble secretory proteins by exclusion from COPI-coated vesicles. *Cell* 98, 81–90.
- Martínez-Menárguez, J.A., Prekeris, R., Oorschot, V.M., *et al.*, 2001. Peri-Golgi vesicles contain retrograde but not anterograde proteins consistent with the cisternal progression model of intra-Golgi transport. *Journal of Cell Biology* 155, 1213–1224.
- Miller, S., Krijnse-Locker, J., 2008. Modification of intracellular membrane structures for virus replication. *Nature Reviews Microbiology* 6, 363–374.
- Mironov, A.A., Mironov Jr., A.A., Bezoussenko, G.V., *et al.*, 2003. ER-to-Golgi carriers arise through direct en bloc protrusion and multistage maturation of specialized ER exit domains. *Developmental Cell* 5, 583–594.
- Mochizuki, Y., Ohashi, R., Kawamura, T., *et al.*, 2013. Phosphatidylinositol 3-phosphatase myotubularin-related protein 6 (MTMR6) is regulated by small GTPase Rab1B in the early secretory and autophagic pathways. *Journal of Biological Chemistry* 288, 1009–1021.
- Monetta, P., Slavin, I., Romero, N., Alvarez, C., 2007. Rab1B interacts with GBF1 and modulates both ARF1 dynamics and COPI association. *Molecular Biology of the Cell* 18, 2400–2410.
- Munro, S., 2011. The golgin coiled-coil proteins of the Golgi apparatus. *Cold Spring Harbor Perspectives in Biology*. doi:10.1101/cshperspect.a005256.
- Munro, S., Pelham, H.R., 1987. A C-terminal signal prevents secretion of luminal ER proteins. *Cell* 48, 899–907.
- Murshid, A., Presley, J.F., 2004. ER-to-Golgi transport and cytoskeletal interactions in animal cells. *Cellular and Molecular Life Sciences* 61, 133–145.
- Neve, E.P., Lahtinen, U., Pettersson, R.F., 2005. Oligomerization and interacellular localization of the glycoprotein receptor ERGIC-53 is independent of disulfide bonds. *Journal of Molecular Biology* 354, 556–568.
- Nichols, W.C., Seligsohn, U., Zivelin, A., *et al.*, 1998. Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. *Cell* 93, 61–70.
- Nilsson, T., Jackson, M., Peterson, P.A., 1989. Short cytoplasmic sequences serve as retention signals for transmembrane proteins in the endoplasmic reticulum. *Cell* 58, 707–718.
- Niu, T.K., Pfeiffer, A.C., Lippincott-Schwartz, J., Jackson, C.L., 2005. Dynamics of GBF1, a brefeldin A-sensitive Arf1 exchange factor at the Golgi. *Molecular Biology of the Cell* 16, 1213–1222.
- Nyfelér, B., Nufer, O., Matsui, T., Mori, K., Hauri, H.-P., 2003. The cargo receptor ERGIC-53 is a target of the unfolded protein response. *Biochemical and Biophysical Research Communications* 304, 599–604.
- Oprins, A., Duden, R., Kreis, T.E., Geuze, H.J., Slot, J.W., 1993. Beta-COP localizes mainly to the *cis*-Golgi side in exocrine pancreas. *Journal of Cell Biology* 121, 49–59.
- Oprins, A., Rabouille, C., Posthuma, G., *et al.*, 2001. The ER-to-Golgi interface is the major concentration site of secretory proteins in the exocrine pancreatic cell. *Traffic* 2, 831–838.
- Orci, L., Stannnes, M., Ravazzola, M., *et al.*, 1997. Bidirectional transport by distinct populations of COPI-coated vesicles. *Cell* 90, 335–349.
- Palmer, K.J., Watson, P., Stephens, D.J., 2005. The role of microtubules in transport between the endoplasmic reticulum and Golgi apparatus in mammalian cells. *Biochemical Society Symposium* 72, 1–13.
- Palokangas, H., Ying, M., Väänänen, K., Saraste, J., 1998. Retrograde transport from the pre-Golgi intermediate compartment and the Golgi complex is affected by the vacuolar H⁺-ATPase inhibitor bafilomycin A1. *Molecular Biology of the Cell* 9, 3561–3578.
- Paroutis, P., Touret, N., Grinstein, S., 2004. The pH of the secretory pathway: Measurement, determinants, and regulation. *Physiology (Bethesda)* 19, 207–215.
- Pelham, H.R., 1988. Evidence that luminal ER proteins are sorted from secreted proteins in a post-ER compartment. *EMBO Journal* 7, 913–918.
- Pelham, H.R., 1989. Control of protein exit from the endoplasmic reticulum. *Annual Review of Cell and Developmental Biology* 5, 1–23.
- Pepperkok, R., Scheel, J., Horstmann, H., *et al.*, 1993. Beta-COP is essential for biosynthetic membrane transport from the endoplasmic reticulum to the Golgi complex *in vivo*. *Cell* 74, 71–82.
- Peter, F., Plutner, H., Zhu, H., Kreis, T.E., Balch, W.E., 1993. β -COP is essential for transport of protein from the endoplasmic reticulum to the Golgi *in vitro*. *Journal of Cell Biology* 122, 1155–1167.
- Plutner, H., Cox, A.D., Pind, S., *et al.*, 1991. Rab1b regulates vesicular transport between the endoplasmic reticulum and successive Golgi compartments. *Journal of Cell Biology* 115, 31–43.
- Plutner, H., Davidson, H.W., Saraste, J., Balch, W.E., 1992. Morphological analysis of protein transport from the ER to Golgi membranes in digitonin-permeabilized cells: Role of the p58-containing compartment. *Journal of Cell Biology* 119, 1097–1116.
- Popoff, V., Langer, J.D., Reckmann, I., *et al.*, 2011. Several ADP-ribosylation factor (Arf) isoforms support COPI vesicle formation. *Journal of Biological Chemistry* 286, 35634–35642.
- Presley, J., Ward, T.H., Pfeffer, A.C., *et al.*, 2002. Dissection of COPI and Arf1 dynamics *in vivo* and role in Golgi membrane transport. *Nature* 417, 187–193.
- Presley, J.F., Cole, N.B., Schroer, T.A., *et al.*, 1997. ER-to-Golgi transport visualized in living cells. *Nature* 389, 81–85.
- Presley, J.F., Smith, C., Hirschberg, K., *et al.*, 1998. Golgi membrane dynamics. *Molecular Biology of the Cell* 9, 1617–1626.
- Rabouille, C., Klumperman, J., 2005. The maturing role of COPI vesicles in intra-Golgi transport. *Nature Reviews Molecular Cell Biology* 6, 812–817.
- Rambourg, A., Clermont, Y., Marraud, A., 1974. Three-dimensional structure of the osmium-impregnated Golgi apparatus as seen in the high voltage electron microscope. *American Journal of Anatomy* 140, 27–45.
- Raposo, G., van Santen, H.M., Leijendekker, R., Geuze, H.J., Ploegh, H.L., 1995. Misfolded major histocompatibility complex class I molecules accumulate in an expanded ER-Golgi intermediate compartment. *Journal of Cell Biology* 131, 1403–1419.
- Rios, R.M., Sanchis, A., Tassin, A.M., Fedriani, C., Bornens, M., 2004. GMAP-210 recruits γ -tubulin complexes to *cis*-Golgi membranes and is required for Golgi ribbon formation. *Cell* 118, 323–335.
- Risco, C., Rodríguez, J.R., López-Iglesias, C., *et al.*, 2002. Endoplasmic reticulum-Golgi intermediate compartment membranes and vimentin filaments participate in vaccinia virus assembly. *Journal of Virology* 76, 1839–1855.
- Roghi, C., Allan, V.J., 1999. Dynamic association of cytoplasmic dynein heavy chain 1a with the Golgi apparatus and intermediate compartment. *Journal of Cell Science* 112, 4673–4685.
- Rojó, M., Pepperkok, R., Emery, G., *et al.*, 1997. Involvement of the transmembrane protein p23 in biosynthetic protein transport. *Journal of Cell Biology* 139, 1119–1135.
- Sáenz, J.B., Sun, W.J., Chang, J.W., *et al.*, 2009. Golgicide A reveals essential roles for GBF1 in Golgi assembly and function. *Nature Chemical Biology* 5, 157–165.
- Sannerud, R., Marie, M., Nizak, C., *et al.*, 2006. Rab1 defines a novel pathway connecting the pre-Golgi intermediate compartment with the cell periphery. *Molecular Biology of the Cell* 17, 1514–1526.
- Saraste, J., Dale, H.A., Bazzocco, S., Marie, M., 2009. Emerging new roles of the pre-Golgi intermediate compartment in biosynthetic-secretory trafficking. *FEBS Letters* 583, 3804–3810.
- Saraste, J., Goud, B., 2007. Functional symmetry of endomembranes. *Molecular Biology of the Cell* 18, 1430–1436.
- Saraste, J., Kuismanen, E., 1984. Pre- and post-Golgi vacuoles operate in the transport of Semliki Forest virus membrane glycoproteins to the cell surface. *Cell* 38, 535–549.
- Saraste, J., Kuismanen, E., 1992. Pathways of protein sorting and membrane traffic between the rough endoplasmic reticulum and the Golgi complex. *Seminars in Cell Biology* 3, 343–355.
- Saraste, J., Lahtinen, U., Goud, B., 1995. Localization of the small GTP-binding protein Rab1 to early compartments of the secretory pathway. *Journal of Cell Science* 108, 1541–1552.
- Saraste, J., Palade, G.E., Farquhar, M.G., 1986. Temperature-sensitive steps in the transport of secretory proteins through the Golgi complex in exocrine pancreatic

- cells. *Proceedings of the National Academy of Sciences of the United States of America* 83, 6425–6429.
- Saraste, J., Palade, G.E., Farquhar, M.G., 1987. Antibodies to rat pancreas Golgi subfractions: Identification of a 58 kDa *cis*-Golgi protein. *Journal of Cell Biology* 105, 2021–2029.
- Saraste, J., Svensson, K., 1991. Distribution of the intermediate elements operating in ER to Golgi transport. *Journal of Cell Science* 100, 415–430.
- Satoh, A., Wang, Y., Malsam, J., Beard, M.B., Warren, G., 2003. Golgin-84 is a Rab1 binding partner involved in Golgi structure. *Traffic* 4, 153–161.
- Satoh, M., Hirayoshi, K., Yokota, S., Hosokawa, N., Nagata, K., 1996. Intracellular interaction of collagen-specific stress protein HSP47 with newly synthesized procollagen. *Journal of Cell Biology* 133, 469–483.
- Scales, S.J., Pepperkok, R., Kreis, T.E., 1997. Visualization of ER-to-Golgi transport in living cells reveals a sequential mode of action for COPII and COPI. *Cell* 90, 1137–1148.
- Schecterson, L.C., Hudson, M.P., Ko, M., *et al.*, 2010. Trk activation in the secretory pathway promotes Golgi fragmentation. *Molecular and Cellular Neuroscience* 43, 403–413.
- Schimmöller, F., Singer-Krüger, B., Schröder, S., *et al.*, 1995. The absence of Emp24p, a component of ER-derived COPII-coated vesicles, causes a defect in transport of selected proteins to the Golgi. *EMBO Journal* 14, 1329–1339.
- Schindler, R., Itin, C., Zerial, M., Lottspeich, F., Hauri, H.-P., 1993. ERGIC-53, a membrane protein of the ER-Golgi intermediate compartment, carries an ER retention motif. *European Journal of Cell Biology* 61, 1–9.
- Schweizer, A., Fransen, J.A., Bächli, T., Ginsel, L., Hauri, H.P., 1988. Identification, by a monoclonal antibody, of a 53-kD protein associated with a tubulo-vesicular compartment at the *cis*-side of the Golgi apparatus. *Journal of Cell Biology* 107, 1643–1653.
- Schweizer, A., Fransen, J.A., Matter, K., *et al.*, 1990. Identification of an intermediate compartment involved in protein transport from endoplasmic reticulum to Golgi apparatus. *European Journal of Cell Biology* 53, 185–196.
- Segev, N., 2001. Ypt and Rab GTPases: Insight into functions through novel interactions. *Current Opinion in Cell Biology* 13, 500–511.
- Sesso, A., de Faria, F.P., Iwamura, E.S., Corrêa, H., 1994. A three-dimensional reconstruction study of the rough ER-Golgi interface in serial thin sections of the pancreatic acinar cell of the rat. *Journal of Cell Science* 107, 517–528.
- Shima, D.T., Scales, S.J., Kreis, T.E., Pepperkok, R., 1999. Segregation of COPI-rich and anterograde-cargo-rich domains in endoplasmic-reticulum-to-Golgi transport complexes. *Current Biology* 9, 821–824.
- Short, B., Preisinger, C., Körner, R., *et al.*, 2001. A GRASP55-rab2 effector complex linking Golgi structure to membrane traffic. *Journal of Cell Biology* 155, 877–883.
- Simpson, J.C., Nilsson, T., Pepperkok, R., 2006. Biogenesis of tubular ER-to-Golgi transport intermediates. *Molecular Biology of the Cell* 17, 723–737.
- Sinka, R., Gillingham, A.K., Kondylis, V., Munro, S., 2008. Golgi coiled-coil proteins contain multiple binding sites for Rab family G proteins. *Journal of Cell Biology* 183, 607–615.
- Stamnes, M.A., Craighead, M.M., Hoe, M.H., *et al.*, 1995. An integral membrane component of coatamer-coated transport vesicles defines a family of proteins involved in budding. *Proceedings of the National Academy of Sciences of the United States of America* 98, 8011–8015.
- Stauber, T., Simpson, J.C., Pepperkok, R., Vernos, I., 2006. A role for kinesin-2 in COPI-dependent recycling between the ER and the Golgi complex. *Current Biology* 16, 2245–2251.
- Stephens, D.J., Lin-Marq, N., Pagano, A., Pepperkok, R., Paccard, J.P., 2000. COPI-coated ER-to-Golgi transport complexes segregate from COPII in close proximity to ER exit sites. *Journal of Cell Science* 113, 2177–2185.
- Stephens, D.J., Pepperkok, R., 2001. Illuminating the secretory pathway: When do we need vesicles? *Journal of Cell Science* 114, 1053–1059.
- Stephens, D.J., Pepperkok, R., 2002. Imaging of procollagen transport reveals COPI-dependent cargo sorting during ER-to-Golgi transport in mammalian cells. *Journal of Cell Science* 115, 1149–1160.
- Stinchcombe, J.C., Nomoto, H., Cutler, D.F., Hopkins, C.R., 1995. Anterograde and retrograde traffic between the rough endoplasmic reticulum and the Golgi complex. *Journal of Cell Biology* 131, 1387–1401.
- Storrie, B., White, J., Röttger, S., *et al.*, 1998. Recycling of Golgi-resident glycosyltransferases through the ER reveals a novel pathway and provides an explanation for nocodazole-induced Golgi scattering. *Journal of Cell Biology* 143, 1505–1521.
- Strating, J.R., Martens, G.J., 2009. The p24 family and selective transport processes at the ER-Golgi interface. *Biology of the Cell* 101, 495–509.
- Sugawara, T., Nakatsu, D., Kii, H., *et al.*, 2012. PKC δ and ϵ regulate the morphological integrity of the ER-Golgi intermediate compartment (ERGIC) but not the anterograde and retrograde transports via the Golgi apparatus. *Biochimica et Biophysica Acta* 1823, 861–875.
- Sztl, E., Lupashin, V., 2009. Role of vesicle tethering factors in ER-Golgi membrane traffic. *FEBS Letters* 583, 3770–3783.
- Szul, T., Grabski, R., Lyons, S., *et al.*, 2007. Dissecting the role of the ARF guanine nucleotide exchange factor GBF1 in Golgi biogenesis and protein trafficking. *Journal of Cell Science* 120, 3929–3940.
- Tang, B.L., Low, S.H., Hauri, H.-P., Hong, W., 1995a. Segregation of ERGIC-53 and the mammalian KDEL-receptor upon exit from the 15 °C compartment. *European Journal of Cell Biology* 68, 398–410.
- Tang, B.L., Low, S.H., Hong, W., 1995b. Differential response of resident proteins and cycling proteins of the Golgi to brefeldin A. *European Journal of Cell Biology* 68, 199–205.
- Tang, B.L., Wang, Y., Ong, Y.S., Hong, W., 2005. COPII and exit from the endoplasmic reticulum. *Biochimica et Biophysica Acta* 1744, 293–303.
- Tang, B.L., Wong, S.H., Qi, X.L., Low, S.H., Hong, W., 1993. Molecular cloning, characterization, subcellular localization and dynamics of p23, the mammalian KDEL receptor. *Journal of Cell Biology* 120, 325–338.
- Tisdale, E.J., 1999. A Rab2 mutant with impaired GTPase activity stimulates vesicle formation from pre-Golgi intermediates. *Molecular Biology of the Cell* 10, 1837–1849.
- Tisdale, E.J., Artalejo, C.R., 2006. Src-dependent atypical protein kinase C ι /lambda (aPKC ι /lambda) tyrosine phosphorylation is required for aPKC ι /lambda association with Rab2 and glyceraldehyde-3-phosphate dehydrogenase on pre-Golgi intermediates. *Journal of Biological Chemistry* 281, 8436–8442.
- Tisdale, E.J., Azizi, F., Artalejo, C.R., 2009. Rab2 utilizes glyceraldehyde-3-phosphate dehydrogenase and protein kinase C(ι) to associate with microtubules and to recruit dynein. *Journal of Biological Chemistry* 284, 5876–5884.
- Tisdale, E.J., Bourne, J.R., Khosravi-Far, R., Der, C.J., Balch, W.E., 1992. GTP-binding mutants of Rab1 and Rab2 are potent inhibitors of vesicular transport from the endoplasmic reticulum to the Golgi complex. *Journal of Cell Biology* 119, 749–761.
- Tisdale, E.J., Plutner, H., Matteson, J., Balch, W.E., 1997. p53/58 binds COPI and is required for selective transport through the early secretory pathway. *Journal of Cell Biology* 137, 581–593.
- Tomás, M., Martínez-Alonso, E., Ballesta, J., Martínez-Menárguez, J.A., 2010. Regulation of ER-Golgi intermediate compartment tubulation and mobility by COPI coats, motor proteins and microtubules. *Traffic* 11, 616–625.
- Tooze, J., Tooze, S., Warren, G., 1984. Replication of coronavirus MHV-A59 in sac-cells: Determination of the first site of budding of progeny virions. *European Journal of Cell Biology* 33, 281–293.
- Tooze, S.A., Tooze, J., Warren, G., 1988. Site of addition of N-acetyl-galactosamine to the E1 glycoprotein of mouse hepatitis virus-A59. *Journal of Cell Biology* 106, 1475–1487.
- Trucco, A., Polishchuk, R.S., Martella, O., *et al.*, 2004. Secretory traffic triggers the formation of tubular continuities across Golgi subcompartments. *Nature Cell Biology* 6, 1071–1081.
- Tsvetanova, N.G., Riordan, D.P., Brown, P.O., 2012. The yeast Rab GTPase Ypt1 modulates unfolded protein response dynamics by regulating the stability of HAC1 RNA. *PLoS Genetics* 8, e1002862.
- Urbani, L., Simoni, R.D., 1990. Cholesterol and vesicular stomatitis virus G protein take separate routes from the endoplasmic reticulum to the plasma membrane. *Journal of Biological Chemistry* 265, 1919–1923.
- Vavassori, S., Cortini, M., Masui, S., *et al.*, 2013. A pH-regulated quality control cycle for surveillance of secretory protein assembly. *Molecular Cell* 50, 783–792.
- Velloso, L.M., Svensson, K., Pettersson, R.F., Lindqvist, Y., 2003. The crystal structure of the carbohydrate-recognition domain of the glycoprotein sorting receptor p58/ERGIC-53 reveals an unpredicted metal-binding site and conformational changes associated with calcium ion binding. *Journal of Molecular Biology* 334, 845–851.
- Volchuk, A., Amherdt, M., Ravazzola, M., *et al.*, 2000. Megavesicles implicated in the rapid transport of intracisternal aggregates across the Golgi stack. *Cell* 102, 335–348.
- Volpicelli-Daley, L.A., Li, Y., Zhang, C.J., Kahn, R.A., 2005. Isoform-selective effects of the depletion of ADP-ribosylation factors 1–5 on membrane traffic. *Molecular Biology of the Cell* 16, 4495–4508.
- Wang, C., Yoo, Y., Fan, H., *et al.*, 2010. Regulation of Integrin β 1 recycling to lipid rafts by Rab1a to promote cell migration. *Journal of Biological Chemistry* 285, 29398–29405.

- Ward, T.H., Polishchuk, R.S., Caplan, S., Hirschberg, K., Lippincott-Schwartz, J., 2001. Maintenance of Golgi structure and function depends on the integrity of ER export. *Journal of Cell Biology* 155, 557–570.
- Warren, G., 1987. Signals and salvage sequences. *Nature* 327, 17–18.
- Wilson, D.W., Lewis, M.J., Pelham, H.R., 1993. pH-dependent binding of KDEL to its receptor *in vitro*. *Journal of Biological Chemistry* 268, 7465–7468.
- Winslow, A.R., Chen, C.W., Corrochano, S., *et al.*, 2010. α -Synuclein impairs macroautophagy: Implications for Parkinson's disease. *Journal of Cell Biology* 190, 1023–1037.
- Xu, D., Hay, J.C., 2004. Reconstitution of COPII vesicle fusion to generate a pre-Golgi intermediate compartment. *Journal of Cell Biology* 167, 997–1003.
- Xu, D., Joglegar, A.P., Williams, A.L., Hay, J.C., 2000. Subunit structure of a mammalian ER/Golgi SNARE complex. *Journal of Biological Chemistry* 275, 39631–39639.
- Yamamoto, K., Fujii, R., Toyofuku, Y., *et al.*, 2001. The KDEL receptor mediates a retrieval mechanism that contributes to quality control at the endoplasmic reticulum. *EMBO Journal* 20, 3082–3091.
- Ying, M., Flatmark, T., Saraste, J., 2000. The p58-positive pre-Golgi intermediates consist of distinct subpopulations of particles that show differential binding of COPI and COPII coats and contain vacuolar H^+ -ATPase. *Journal of Cell Science* 113, 3623–3638.
- Ying, M., Sannerud, R., Flatmark, T., Saraste, J., 2002. Colocalization of Ca^{2+} -ATPase and GRP94 with p58 and the effects of thapsigargin on protein recycling suggest the participation of the pre-Golgi intermediate compartment in intracellular Ca^{2+} -storage. *European Journal of Cell Biology* 81, 469–483.
- Yoo, J.S., Moyer, B.D., Bannykh, S., *et al.*, 2002. Non-conventional trafficking of the cystic fibrosis transmembrane conductance regulator through the early secretory pathway. *Journal of Biological Chemistry* 277, 11401–11409.
- Yu, S., Satoh, A., Pypaert, M., *et al.*, 2006. mBet3p is required for homotypic COPII vesicle tethering in mammalian cells. *Journal of Cell Biology* 174, 359–368.
- Zhang, T., Hong, W., 2001. Ykt6 forms a SNARE complex with syntaxin 5, GS28, and Bet1 and participates in a late stage in endoplasmic reticulum-Golgi transport. *Journal of Biological Chemistry* 276, 27480–27487.
- Zhang, T., Wong, S.H., Tang, B.L., *et al.*, 1997. The mammalian protein (rbet1) homologous to yeast Bet1p is primarily associated with the pre-Golgi intermediate compartment and is involved in vesicular transport from the endoplasmic reticulum to the Golgi apparatus. *Journal of Cell Biology* 139, 1157–1168.
- Zhang, T., Wong, S.H., Tang, B.L., Xu, Y., Hong, W., 1999. Morphological and functional association of Sec22b/ERS-24 with the pre-Golgi intermediate compartment. *Molecular Biology of the Cell* 10, 435–453.
- Zhao, X., Claude, A., Chun, J., *et al.*, 2006. GBF1, a *cis*-Golgi and VTCs-localized ARF-GEF, is implicated in ER-to-Golgi protein traffic. *Journal of Cell Science* 119, 3743–3753.
- Zheng, C., Page, R.C., Das, V., *et al.*, 2013. Structural characterization of carbohydrate binding by LMAN1 protein provides new insight into the endoplasmic reticulum export of factors V (FV) and VIII (FVIII). *Journal of Biological Chemistry* 288, 20499–20509.
- Zheng, C., Zhang, B., 2013. Combined deficiency of coagulation factors V and VIII: An update. *Seminars in Thrombosis and Hemostasis* 39, 613–620.
- Zoppino, F.C., Militello, R.D., Slavin, I., Alvarez, C., Colombo, M.I., 2010. Autophagosome formation depends on the small GTPase Rab1 and functional ER exit sites. *Traffic* 11, 1246–1261.
- Zuber, C., Fan, J., Guhl, B., *et al.*, 2001. Immunolocalization of UDP-glucose: glycoprotein glucosyltransferase indicates involvement of pre-Golgi intermediates in protein quality control. *Proceedings of the National Academy of Sciences of the United States of America* 98, 10710–10715.
- Zuber, C., Fan, J.Y., Guhl, B., Roth, J., 2004. Misfolded proinsulin accumulates in expanded pre-Golgi intermediates and endoplasmic reticulum subdomains in pancreatic beta cells of Akita mice. *FASEB Journal* 18, 917–919.

Further Reading

- Bannykh, S.I., Nishimura, N., Balch, W.E., 1998. Getting into the Golgi. *Trends in Cell Biology* 8, 21–25.
- Brownhill, K., Wood, L., Allan, V., 2009. Molecular motors and the Golgi complex: Staying put and moving through. *Seminars in Cell & Developmental Biology* 20, 784–792.
- Capitani, M., Sallese, M., 2009. The KDEL-receptor: New functions for an old protein. *FEBS Letters* 583, 3863–3871.
- Dancourt, J., Barlowe, C., 2010. Protein sorting receptors in the early secretory pathway. *Annual Review of Biochemistry* 79, 777–802.
- Farquhar, M.G., Palade, G.E., 1998. The Golgi apparatus: 100 years of progress and controversy. *Trends in Cell Biology* 8, 2–10.
- Klumperman, J., 2011. Architecture of the mammalian Golgi. *Cold Spring Harbor Perspectives in Biology* 3, a005181.
- Lippincott-Schwartz, J., Roberts, T.H., Hirschberg, K., 2000. Secretory protein trafficking and organelle dynamics in living cells. *Annual Review of Cell and Developmental Biology* 16, 557–589.
- Zanetti, G., Pahuja, K.B., Studer, S., Shim, S., Schekman, R., 2012. COPII and the regulation of protein sorting in mammals. *Nature Cell Biology* 14, 20–28.

Golgi and TGN

A Luini and S Parashuraman, Institute of Protein Biochemistry (IBP), Naples, Italy

© 2016 Elsevier Inc. All rights reserved.

The Golgi Apparatus – An Introduction

The Golgi apparatus and the closely associated *trans*-Golgi network (TGN) are the central transport stations of the eukaryotic secretory pathway, a fundamental positioning system through which a large fraction of the eukaryotic proteins (roughly 30% in mammals) are biochemically processed and conveyed from the site of synthesis, the endoplasmic reticulum (ER), toward their cellular destinations. The secretory pathway is, together with the endocytic apparatus, a major cellular organizer responsible for the location of every transported protein to the right cellular membrane or to the extracellular space, where these proteins must operate, as required for cell functionality.

Secretory traffic begins at the ER where cargo proteins are synthesized and inserted into the lumen or the membrane of this organelle. Cargo is then transferred, via the intermediate compartment (Appenzeller-Herzog and Hauri, 2006; Saraste and Kuismanen, 1992), to the Golgi complex, morphologically a stack of flat membranous cisternae enriched in glycosylating enzymes, and then through the *cis*-, *medial*- and *trans*-Golgi cisternae to the TGN, where it is sorted into membranous carriers for delivery to successive cellular destinations (De Matteis and Luini, 2008; Gu *et al.*, 2001). Along the way, and especially in the Golgi stack, cargo proteins (and lipids) are biochemically modified by glycosylation, the most abundant and complex eukaryotic posttranslational modification (Stanley, 2011); or they are broken down by degradative systems located in the ER (Vembar and Brodsky, 2008) and in the lysosomes (Luzio *et al.*, 2007). Moreover, at each transport stage, anterograde transport is balanced by regulated retrograde compensatory fluxes of membrane and proteins to maintain the morphological and compositional homeostasis of the system (Cancino *et al.*, 2014; Lorente-Rodriguez and Barlowe, 2011; Luini *et al.*, 2014). To carry out this enormous task, the transport apparatus relies on underlying molecular machinery that is estimated to comprise more than 2000 proteins, several hundreds of which reside in the Golgi/TGN (Gannon *et al.*, 2011; Gilchrist *et al.*, 2006).

The following description refers mostly to the mammalian Golgi complex, which has been extensively studied. The Golgi, however, is evolutionarily very ancient, most probably as ancient as the last common eukaryotic ancestor (LCEA), which appeared roughly 1.5 billion years ago (Klute *et al.*, 2011). Throughout evolution, the Golgi has varied significantly, especially in morphology, although its main functions, namely, glycosylation and transport and sorting, are likely to have been substantially preserved (Klute *et al.*, 2011).

The Structure of the Golgi Apparatus

The mammalian Golgi apparatus consists of hundreds of Golgi stacks that are typically made up of 3–6 flat cisternae of

about 1 μm in diameter (Klumperman, 2011; Figure 1). Cisternae are swollen at the rims, where transport takes place, and are decorated with perforations in the periphery, which may align across successive cisternae to form ‘wells’ (Rambourg and Clermont, 1990). Each Golgi stack is surrounded by numerous transport vesicles roughly 50–70 nm in diameter. The vesicles are especially concentrated near the rims and inside the wells (Ladinsky *et al.*, 1994). The Golgi stacks are linked together side by side to form a ribbon that is localized near the centrosome of the cell, by virtue of microtubule-based motors (Yadav and Linstedt, 2011). The links between stacks are constituted of tubules and perforated membrane areas (‘non-compact zones’) that typically join cisternae at the same level (i.e., *cis*-cisternae with *cis* and *medial* with *medial*) (Clermont *et al.*, 1994; Klumperman, 2011). In addition, connections may sometimes join different levels in adjacent stacks (Clermont *et al.*, 1994; Trucco *et al.*, 2004; Vivero-Salmeron *et al.*, 2008). These ‘heterologous’ connections, albeit infrequent, might be functionally important for transport through the Golgi stack (Beznoussenko *et al.*, 2014; Trucco *et al.*, 2004).

The *trans*-most and sometimes the penultimate cisternae extend tubules to form an interconnected tubular network at the *trans*-Golgi face, called *Trans*-Golgi Network or TGN (Gu *et al.*, 2001; Ladinsky *et al.*, 1994). The size of the TGN varies greatly with the secretory flux (Klumperman, 2011) and with transport conditions (Griffiths *et al.*, 1989). Sometimes, the *trans*-most cisterna can be seen to be ‘peeling off’ from the rest of the stack, again possibly as part of the transport process (Ladinsky *et al.*, 1994; Mogelsvang *et al.*, 2003; Rambourg and Clermont, 1990). Notably, though the TGN is a well-developed structure closely associated with the Golgi stack in mammalian cells, it appears to be an independent organelle in plants. The plant TGN receives endocytosed material from the plasma membrane, and acts as an early endosome. Indeed, TGN and early endosomes are generally considered a single organelle in plants, and distinctions, if any, between them have not been explored (Richter *et al.*, 2009). These and other observations (Lippincott-Schwartz *et al.*, 1991) suggest that the TGN might have evolved from endosomes early in evolution.

Transport through the Golgi

The Golgi is an obligate intracellular station for most of the secreted cargoes that move from the ER to other intracellular destinations. An exception is represented by a few cargo molecules that appear to bypass the entire secretory route and to follow an unconventional secretory pathway (Malhotra, 2013; Rabouille *et al.*, 2012). In the conventional route, secretory cargoes leave the ER via exit sites and reach the intermediate compartment before being transported to the Golgi via pleomorphic carriers (Appenzeller-Herzog and Hauri, 2006; Barlowe and Miller, 2013). The arrival at

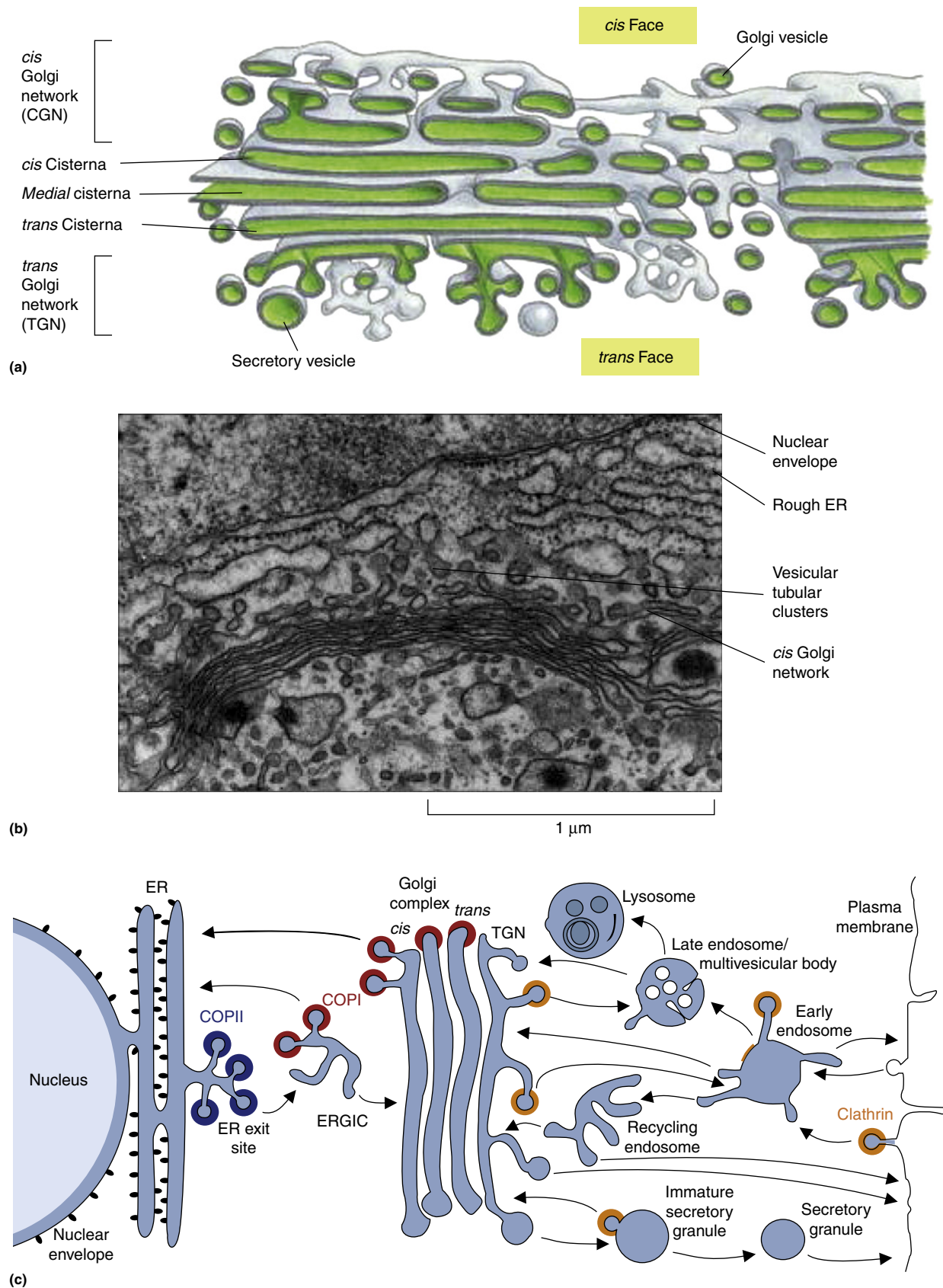


Figure 1 A representation of the Golgi stack depicting the transport pathways. (a) A cartoon depicting the 3D organization of the stack. (b) An electron micrograph showing the cross-section of a Golgi stack. (c) A schematic representation of the intracellular transport pathways and the central role played by the Golgi at the intersection of exocytic and endocytic pathways. Reprinted with permission from Alberts, B., 2002. *Molecular Biology of the Cell*, fourth ed. New York, NY: Garland Science; Adapted from Bonifacino, J.S., Glick, B.S., 2004. *The mechanisms of vesicle budding and fusion*. Cell 116, 153–166.

the Golgi of cargo-laden carriers stimulates a compensatory retrograde transport from the Golgi to the ER that maintains membrane homeostasis at the Golgi (Cancino *et al.*, 2014; Lorente-Rodriguez and Barlowe, 2011; Luini *et al.*, 2014).

The cargo is then transported to the *trans*-Golgi side. The mode of transport through the Golgi has been studied intensely since the birth of cell biology roughly 50 years ago (Farquhar and Palade, 1998). Despite major advances, definitive conclusions have not yet been reached, mostly because of the enormous technical difficulty of imaging dynamic transport events within the compact Golgi stack with sufficient resolution in time and space. Historically, the dominant concept until the mid-1980s had been that traffic is mediated by vesicles shuttling between stable cisternae. This 'vesicular' model was based essentially on classical electron microscopy as well as on genetic and biochemical evidence (Palade, 1975; Rothman and Wieland, 1996; Schekman and Orci, 1996). The successive advent of more sophisticated imaging technologies stimulated a reexamination of the traffic mechanisms resulting in the proposal of the cisternal maturation model (Bonfanti *et al.*, 1998; Losev *et al.*, 2006; Matsuura-Tokita *et al.*, 2006). This scheme proposes that the cargo remains always in the cisternal lumen and the cisternae progress from *cis* to *trans* positions in the stack while 'resident'

components (e.g., glycosylation enzymes) move backwards in the *trans* to *cis* direction in lockstep with cisternal progression. Today, both models are still being considered, though the present evidence favors the cisternal maturation model. They are concisely presented below (see Glick and Luini, 2011; Figure 2).

According to cisternal maturation, cargo-containing carriers arriving from the intermediate compartment fuse together and align with the *cis*-Golgi face. The Golgi resident proteins are then retrogradely transported into the newly arrived membrane to create a new *cis*-cisterna. Cargo proteins remain inside this cisterna and move forward in the *cis* to *trans* direction. Once arrived at the *trans* position, the cargo-laden cisterna disassembles into transport carriers. At the same time, resident proteins (mostly glycosylation enzymes) shift backward in synchrony with cargo progression (Papanikou and Glick, 2014), maintaining the compositional gradient of the stack. This retrograde transport of enzymes has been proposed to be mediated by the vesicles present in the periphery of the stack and also by intercisternal continuities (Lanoix *et al.*, 2001; Martinez-Menarguez *et al.*, 2001; Rizzo *et al.*, 2013; Trucco *et al.*, 2004). This model is supported by the observation that macromolecular cargoes such as procollagen in mammalian cells (Bonfanti *et al.*, 1998) or scales in certain algae

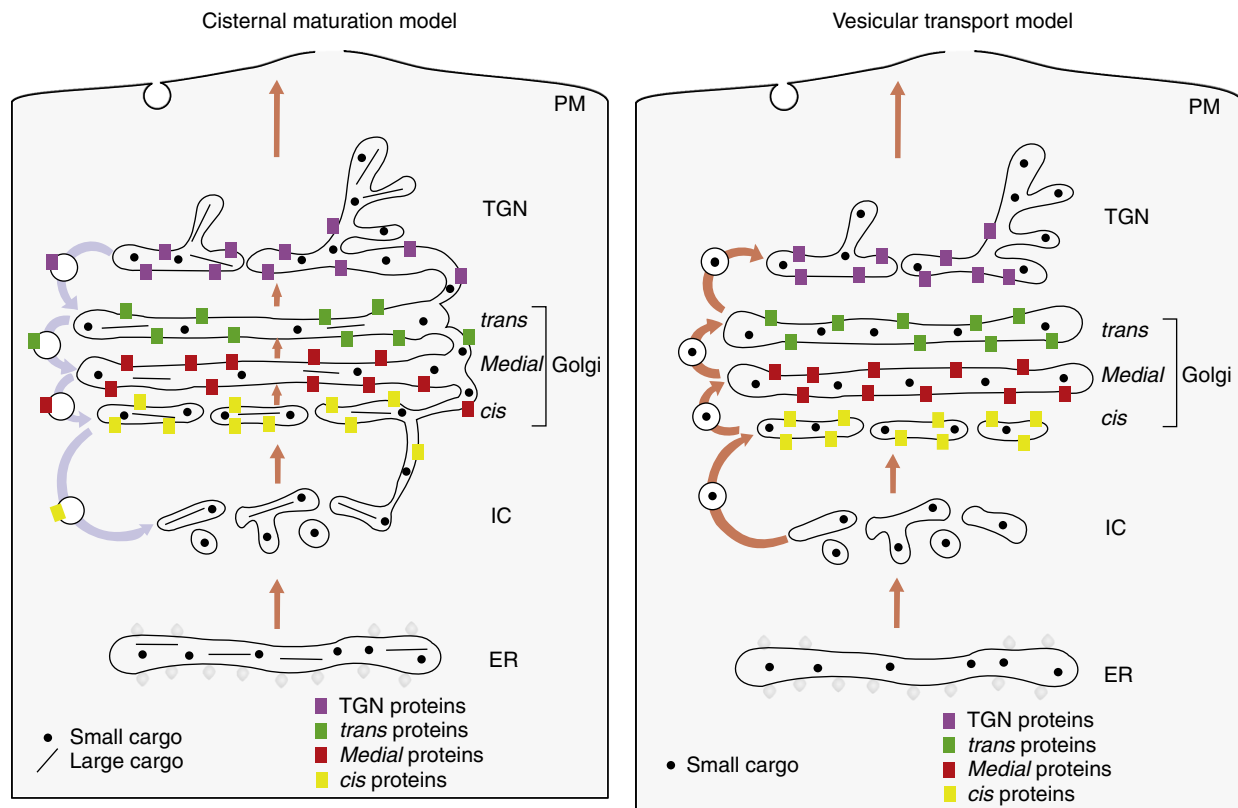


Figure 2 Models of intra-Golgi transport. The cisternal maturation model proposes that the cargo is stably retained in the cisterna while the residents are transported retrogradely by COPI derived vesicles or intercisternal continuities. The vesicular transport model proposes that the residents are stably retained in the cisterna while the cargo is transported in the anterograde direction by COPI derived vesicles. See text for a more detailed description of these models and for a list of variants of these basic schemes. Reprinted with permission from Glick, B.S., Luini, A., 2011. Models for Golgi traffic: A critical assessment. Cold Spring Harbor Perspectives in Biology 3, a005215. Copyright Cold Spring Harbor Laboratory Press.

(Becker *et al.*, 1995) move efficiently through the stack despite being too large to enter the vesicles and while remaining always inside the lumen of cisternae (Bonfanti *et al.*, 1998); and by direct visualization of Golgi components moving backwards from distal to proximal cisternae, in synchrony with cisternal progression. This observation was made in *Saccharomyces cerevisiae*, where the cisterna are unstacked, allowing for better microscopic resolution (Losev *et al.*, 2006; Matsuura-Tokita *et al.*, 2006).

The vesicular transport model instead proposes that cisternae are stable and the resident enzymes remain permanently in their lumen, while the cargo traverses the stack in vesicular carriers. This scheme was proposed by Palade and colleagues (Palade, 1975) and substantiated by the *in vitro* studies of Rothman and colleagues, who identified the molecular machinery of transport and proposed that the features of such machinery are consistent with vesicular transport (Rothman, 2002). Yeast genetic studies conducted in parallel by Schekman and colleagues reached the same conclusions (Schekman, 2002). The main evidence in support of this model is that cargo can be visualized by super-resolution microscopy inside transport vesicles, at least under certain transport conditions (Jamieson and Palade, 1967; Malhotra *et al.*, 1989; Pellett *et al.*, 2013). In addition, several variants of this stable cisternae model have been proposed and are in consideration. These include the rapid mixing-partitioning model (Patterson *et al.*, 2008), the cisternal progenitor model (Pfeffer, 2010), the kiss-and-run transport model (Beznoussenko and Mironov, 2002), and the rimmel maturation model (Lavieu *et al.*, 2013). Finally, recent studies on the transport of soluble cargoes across the cisterna have shown that this type of secretory cargoes can diffuse across the stack via inter-cisternal continuities while other cargoes like procollagen move by cisternal maturation (Beznoussenko *et al.*, 2014). Altogether, the available data suggest that multiple mechanisms of transport may coexist in the Golgi stack.

The cargoes arriving at the TGN are transported to multiple destinations including the basolateral domain, apical domain, endosomes, ER, and the Golgi stack itself (Anitei and Hoflack, 2011; De Matteis and Luini, 2008). The first step in the transport of cargoes to their appropriate destinations is their sorting into cargo domains at the TGN from where they depart. The segregation of these domains is due in most cases on sorting signals present on the cargo proteins that bind to appropriate adaptors leading to their concentration and segregation (Bonifacino, 2014), although other mechanisms have also been demonstrated (De Matteis and Luini, 2008). This step is followed by the formation of cargo-laden transport carriers. Irrespective of their destination, most carriers that transport cargo out of the TGN are large pleomorphic structures (Polishchuk *et al.*, 2000, 2006). They form by the extrusion of the cargo domains along cytoskeletal tracks by motor proteins. The tubules connecting the carriers to the TGN are then severed to release the carriers that follow their course to their appropriate destination using cytoskeletal tracks (Luini *et al.*, 2008).

In addition to playing a role in the sorting of cargoes out of the Golgi, TGN also acts as recipient for incoming traffic from the endosomes that allow for the recycling of receptors such as M6PR (Pfeffer, 2011).

Molecular Mechanism of Transport

Membrane traffic in eukaryotic cells uses a 'toolbox' of molecular processes that are involved in the transport of cargo from one compartment to another (De Matteis and Luini, 2011; Figure 3). These 'elementary' processes include cargo sorting and membrane bending and budding mediated by 'coat' proteins, targeting of carrier vesicles to the target membrane by Rab and tethering proteins, and finally membrane fusion mediated by soluble NSF attachment receptors (SNAREs), where NSF stands for N-ethylmaleimide sensitive fusion proteins. Each of these processes is mediated by a class of machinery proteins specialized for each of these 'elementary' events such as fission, fusion, etc. Here, we briefly mention the Golgi-specific variants of the proteins involved in these processes, thus supporting transport across the stack. The members of these families of proteins that act at the Golgi include the COPI coat involved in cargo sorting and carrier formation (Popoff *et al.*, 2011), adaptors for the coat formation like GOLPH3 in COPI vesicle formation (Banfield, 2011) and AP-1 in clathrin-mediated vesicle formation (Bonifacino, 2014), phospholipase D2 involved in vesicle scission (Hsu and Yang, 2009), the COG family of membrane tethers (Willett *et al.*, 2013) and SNARE proteins membrin, syntaxin 5, GOS 28 etc. (Malsam and Sollner, 2011) involved in vesicle fusion. In addition, the Golgi matrix proteins (Munro, 2011) maintain the architecture of the Golgi and are also involved in vesicle tethering (Shorter and Warren, 2002).

Glycosylation at the Golgi

The proteins and lipids passing through the Golgi are processed by the Golgi-localized enzymes. The Golgi-localized glycan-modifying enzymes are type II membrane proteins with a short cytoplasmic tail and the catalytic portion of the enzyme localized in the lumen of the organelle (Hirschberg and Snider, 1987; Rini *et al.*, 2009). Usually, the localization of the enzymes reflects their order of action in the glycosylation pathway, with the early acting enzymes present in the *cis*-side of the Golgi and the later acting enzymes in the *trans*-side (Dunphy and Rothman, 1985).

The Golgi glycosylation enzymes are involved in the processing of N-glycans, O-glycan biosynthesis, proteoglycan biosynthesis, and glycolipid biosynthesis. The protein N-glycosylation starts with the transfer of precursor polysaccharides *en bloc* to a conserved Asn residue of the protein by oligosaccharyltransferase in the ER. The mannose residues are then trimmed in the ER before reaching the *cis*-Golgi where α -mannosidases further trim the glycan. This glycan then becomes the substrate for *medial*-Golgi localized N-acetylglucosaminyl transferase that transfers a N-acetylglucosamine residue onto the N-glycan. The product of these reactions becomes the substrate for galactosyltransferases and sialyltransferases in the *trans*-Golgi/TGN, thus leading to the formation of complex N-glycan (for further details see Stanley *et al.*, 2009). In contrast to N-glycosylation, O-glycosylation starts in the Golgi with the addition of N-acetylgalactosamine on serine/threonine residue of the proteins by galactosamine transferases. The galactosamine residue is further modified in the *trans*-Golgi/TGN to

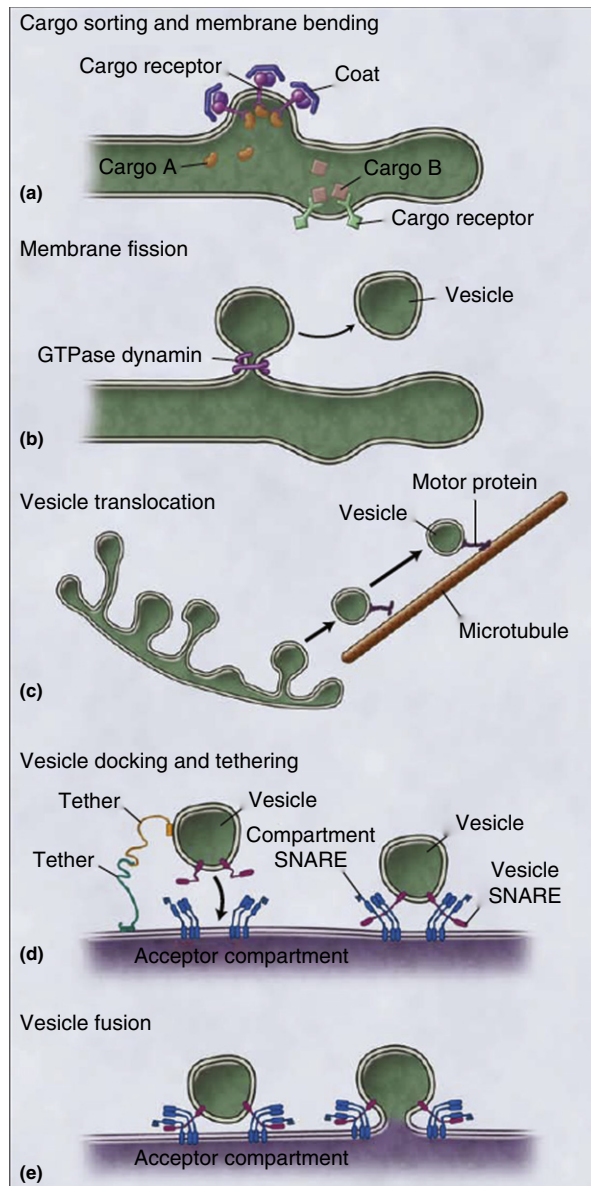


Figure 3 A schematic representation of the ‘toolbox’ of processes used in the intracellular transport. The vesicle formation involves cargo sorting followed by membrane bending initiated by coat proteins. (a) The mature vesicle is then scissioned by dedicated factors. (b) The vesicle then translocates via microtubule-based motors to the appropriate target compartment (c), where it is tethered by a Rab and a molecular tether-dependent process (d) that also regulates the specificity of this vesicle tethering followed by docking through specific SNAREs. (d) SNARE pairing results in membrane fusion (e) delivering the cargo to the target compartment. A specific temporal combination of these elementary processes is used in all the transport pathways and see text for the Golgi-specific molecules involved in these processes at the Golgi apparatus. Reprinted with permission from De Matteis, M.A., Luini, A., 2011. Mendelian disorders of membrane trafficking. *New England Journal of Medicine* 365, 927–938. Copyright 2014, Massachusetts Medical Society.

create complex O-glycan forms (Brockhausen *et al.*, 2009). Similar to O-glycosylation, proteoglycan synthesis starts with the addition of a monosaccharide, namely xylose, to the

serine/threonine residue. That is then extended to a tetrasaccharide containing one xylose, two galactoses, and a single glucuronic acid residue. This tetrasaccharide is then extended depending on the type of proteoglycan to be synthesized. The glycan residues can also get other modifications such as sulfation as in the case of proteoglycans (Esko *et al.*, 2009) or phosphorylation as in the case of mannose-6-phosphate synthesis.

Glycolipid biosynthesis starts in the ER with the production of ceramide (Schnaar *et al.*, 2009). A peculiar feature of the lipid biosynthesis is the use of lipid transfer proteins that transfer lipids to the TGN where most of the glycolipid biosynthesis takes place. The ceramide produced in the ER is transferred by CERT (ceramide transfer protein) to the TGN where it is converted into sphingomyelin by the action of sphingomyelin synthases 1 and 2 (Hanada *et al.*, 2009). The rest of the ceramide is transported to the *cis*-Golgi where it is first converted into glucosylceramide by glucosylceramide synthase and then processed in the TGN to produce complex glycosphingolipids (D’Angelo *et al.*, 2013).

The sugar residues needed for glycosylation reactions are transported into the lumen as nucleotide sugar units by transporters. While sialic acid is imported as CMP-sialic acid and fucose as GDP-fucose, most other sugars are imported as UDP-sugar conjugates (for further details, see Freeze and Elbein, 2009). The vacuolar-ATPase localized to the *trans*-side maintains an acidic pH required for the glycosylation reactions and the neutralization of pH impairs glycosylation (Rivinoja *et al.*, 2012).

Regulation of the Golgi Apparatus Structure and Function

Given the complex structural and functional features of the Golgi, it is not surprising that these features are controlled by specialized regulatory mechanisms. Screening studies have shown that the secretory pathway is under the regulation of many kinases and phosphatases (Chia *et al.*, 2012; Farhan *et al.*, 2010; Simpson *et al.*, 2012; Zacharogianni *et al.*, 2011). Here we highlight a few such processes where the significance of this kinase-mediated regulation is understood.

Some of these processes are involved in the maintenance of the homeostasis of the transport apparatus. Transport involves large membrane and protein fluxes across the transport stations, that are subjected to physiological and pathological perturbations as well as to spontaneous drifts away from equilibrium (Hirschberg *et al.*, 1998; Mironov *et al.*, 2001; Pulvirenti *et al.*, 2008; Trucco *et al.*, 2004). Such perturbations can be a serious threat to the homeostasis, as they might alter the composition and morphology of the transport compartments disrupting their function. An example of this possibility is transport at the interface between ER and the Golgi complex. The ER is roughly 10-fold that of the Golgi in size (Griffiths *et al.*, 1984) and, its membrane output toward the Golgi can represent a significant fraction of the Golgi volume/surface (Griffiths *et al.*, 1984; Klumperman, 2000; Martinez-Menarguez *et al.*, 1999; Thor *et al.*, 2009). Thus, even minor changes in export from the ER would cause major alterations in Golgi morphology and composition, if not compensated for by

corresponding changes in retrograde transport from the Golgi to the ER (Klumperman, 2000; Martinez-Menarguez *et al.*, 1999; Thor *et al.*, 2009). However, the Golgi always maintains or recovers its normal composition and morphology (Mironov *et al.*, 2001; Trucco *et al.*, 2004). At least some of the underlying compensatory mechanisms use control devices that sense the changes/perturbations and respond quickly, precisely, and appropriately to restore homeostasis.

One regulatory device that operates at the interface between ER and Golgi has been investigated fairly extensively, and concepts on the molecular composition and design of these devices have begun to emerge (Cancino *et al.*, 2014). Membrane fluxes leaving the ER for the Golgi result in the activation of a traffic sensor, or receptor (the KDEL), at the *cis*-Golgi. This acts as a signaling protein that activates transduction pathways that are similar in composition to canonical PM signaling pathways and kinase cascades. These kinases phosphorylate a large number of Golgi and cytosolic proteins, some of which are involved in anterograde and retrograde traffic. As a result, anterograde and retrograde traffic become activated, counteracting the accumulation of membrane/proteins that would result from a non-compensated continuous arrival of traffic to the Golgi and hence helping to maintain Golgi homeostasis. Whether this scheme applies at other stages of the trafficking pathways remains to be investigated.

Simpler molecular processes involved in the transport are also subjected to regulation. One particularly clear example is the regulation of uncoating of COPII-coated carriers. The COPII-coated carriers derived from the ER fuse with the Golgi apparatus to deliver the cargo and the regulation of uncoating of the carrier only at the appropriate compartment (Golgi) prevents the possibility of engagement with an inappropriate compartment. This is accomplished by the specific localization of casein kinase 1d at the Golgi that can phosphorylate the COPII coat components upon arrival of the anterograde carrier at the Golgi, leading to its uncoating and subsequent fusion with the Golgi cisternae (Lord *et al.*, 2011). Finally, the glycosylation processes at the Golgi are also subjected to regulation. Examples include the Src kinase-mediated regulation of the localization of N-acetylglucosamine transferases that initiate o-glycosylation (Chia *et al.*, 2012; Gill *et al.*, 2010) and the regulation of import of N-acetylgalactosamine into the Golgi lumen that determines branching of N-glycans (Lau *et al.*, 2007).

Role of Golgi in Disease

Given the central role played by the Golgi apparatus in the processing and secretion of proteins, a defect in Golgi function results in serious disorders/diseases (for a detailed discussion see Freeze and Ng, 2011). The disorders affecting specific glycosylation pathways lead to similar symptoms. For example, defects in O-mannose biosynthetic pathways lead to diseases involving muscular dystrophy probably due to the high levels of O-mannose present in α -dystroglycan while defects in proteoglycan biosynthesis mainly causes skeletal disorders since extracellular matrix abundant in proteoglycans plays a major role in bone development (Freeze and Ng, 2011). Golgi-associated diseases also occur due to mutations in the genes encoding proteins that form part of the structural/mechanistic

processes of the Golgi. Since a defect in structural or mechanistic processes affect many aspects of Golgi function, the disorders result in a plethora of symptoms. A classic example of such disorders is the congenital glycosylation disorders (CGD) caused due to mutations affecting the COG subunits involved in the recycling of Golgi residents, a process required for their appropriate localization (Zeevaert *et al.*, 2008). CDG disorders involving almost all the subunits of COG complex have been documented and the patients with these disorders show symptoms of impaired development, mental retardation, and hypotonia in many cases along with other subunit-specific symptoms (Freeze and Ng, 2011). In addition, mutations in a subunit of V-ATPase that is required for the maintenance of the pH inside the Golgi lumen essential for glycosylation result in connective tissue disorders with wrinkly skin (Cutis laxa) (Rosnoblet *et al.*, 2013).

Apart from the genetic diseases mentioned above, Golgi proteins have also been linked with various other diseases including cancer, infection, etc. (Hu *et al.*, 2007; Varki and Freeze, 2009; Varki *et al.*, 2009b). For instance, the Golgi-localized protein GOLPH3 has been recently shown to be overexpressed in solid tumors and to act as an oncogene (Scott *et al.*, 2009).

Evolution of the Golgi Apparatus

The secretory pathway is a central component of the eukaryotic endomembrane system and the Golgi apparatus plays an essential role in it. Nevertheless, and rather surprisingly, there are organisms that do not seem to have a recognizable Golgi apparatus and that led to the suggestion these organisms might represent the earliest branches of the eukaryotic tree before the evolution of a functional Golgi apparatus (Cavalier-Smith, 1987). Later phylogenetic studies indicated that this was not the case (reviewed in Mowbrey and Dacks, 2009) and suggested that the Golgi apparatus is a universal feature of the eukaryotic cells and probably the LCEA did have a functional Golgi apparatus.

The Golgi apparatus varies in shape from a ribbon-like Golgi in mammalian cells to scattered stacks in the plants or insect cells to unstacked fragmented Golgi cisterna in *S.cerevisiae*. This led to the question about the nature of Golgi apparatus present in the LCEA. Phylogenetic studies have shown that organisms that lacked Golgi stacks are embedded in clades that had stacked Golgi apparatus suggesting that the absence of a stack in certain organisms could be a secondary loss and there were at least eight such independent instances of Golgi unstacking during the eukaryotic evolutionary history (Mowbrey and Dacks, 2009). In addition, a recent phylogenomic study has shown that almost all the categories of Golgi proteins including SNAREs, Rabs, and matrix proteins are conserved across all the eukaryotes with secondary losses seen in some branches (Klute *et al.*, 2011). Given the importance of matrix proteins in the stacking of the Golgi (Seemann *et al.*, 2000; Shorter and Warren, 2002), this supports the idea that early eukaryotes probably had a stacked Golgi.

Apart from this conserved properties, evolutionary innovations have played a part in modifying and adapting the Golgi to the need of the organisms. A clear example is the evolution of ribbon-like organization in the vertebrates (Wei and Seemann, 2010). This structure is also present in sea urchin (Terasaki,

2000), suggesting that probably it is a deuterostome innovation. Cell-biological studies have pointed to a role for the Golgi ribbon in the efficiency of glycosylation (Puthenveedu *et al.*, 2006; Xiang *et al.*, 2013) and transition to deuterostomes is accompanied by huge increase in the diversity of glycosylation products made by the Golgi apparatus with a pronounced increase in sialylated products (Varki *et al.*, 2009a). Thus, the formation of the ribbon-like Golgi apparatus may have accompanied or contributed to the evolution of deuterostomes.

Another innovation during the course of evolution was the development of the TGN. While the TGN is rudimentary in most organisms, it is a well-developed structure in the mammalian cells (probably also in other vertebrates). Moreover, in some organisms the TGN and endosomes are identical (e.g., in plants) (Richter *et al.*, 2009) while the mammalian TGN is distinct from the endosomes, with extensive membrane traffic between them. The expansion of the TGN and development of the ribbon-like structure was also accompanied by expansion in the repertoire of membrane traffic regulators with the human genome encoding nearly 30 known Rab proteins (Diekmann *et al.*, 2011). Whether the evolution of TGN also parallels an increase in the number of sorting pathways exiting the TGN and how this may have contributed to the evolution of mammals remain open questions.

Perspectives

After more than a century after its discovery, our understanding of the functioning of the Golgi apparatus remains uneven. While there is as yet no consensus on aspects as basic as the mode of intra-Golgi transport or a definitive list of TGN exit pathways, the genomic and proteomic studies have led to a rather extensive catalog of proteins localized to the Golgi apparatus and to a satisfactory understanding of the molecular processes they execute. The progress in addressing the unresolved issues is intimately connected with advances in imaging technology and with the use of novel reconstitution approaches (Wakana *et al.*, 2012). Some of these aspects, including glycosylation, regulation of Golgi homeostasis, and cell-type specific variations of Golgi structure/function are yet to be explored in detail. The resolution of these basic questions of transport along with advances in the analysis of functional aspects of the Golgi at a molecular level is necessary to build a satisfactory systems understanding of the Golgi apparatus.

See also: Interorganellar Communication: Interplay and Processes: Post-Golgi Transport – Cargo, Carriers, and Pathways. ER–Golgi Transport. Organelles: Structure and Function: Early Endosomal Compartments; Intermediate Compartment: A Sorting Station between the Endoplasmic Reticulum and the Golgi Apparatus; Signaling from Endosomes; The Late Endosome

References

- Anitei, M., Hoflack, B., 2011. Exit from the *trans*-Golgi network: From molecules to mechanisms. *Current Opinion in Cell Biology* 23, 443–451.
- Appenzeller-Herzog, C., Hauri, H.P., 2006. The ER-Golgi intermediate compartment (ERGIC): In search of its identity and function. *Journal of Cell Science* 119, 2173–2183.
- Banfield, D.K., 2011. Mechanisms of protein retention in the Golgi. *Cold Spring Harbor Perspectives in Biology* 3, a005264.
- Barlowe, C.K., Miller, E.A., 2013. Secretory protein biogenesis and traffic in the early secretory pathway. *Genetics* 193, 383–410.
- Becker, B., Bolinger, B., Melkonian, M., 1995. Anterograde transport of algal scales through the Golgi complex is not mediated by vesicles. *Trends in Cell Biology* 5, 305–307.
- Beznoussenko, G.V., Mironov, A.A., 2002. Models of intracellular transport and evolution of the Golgi complex. *Anatomical Record* 268, 226–238.
- Beznoussenko, G.V., Parashuraman, S., Rizzo, R., *et al.*, 2014. Transport of soluble proteins through the Golgi occurs by diffusion via continuities across cisternae. *Elife* 3, e02009.
- Bonfanti, L., Mironov Jr., A.A., Martinez-Menarguez, J.A., *et al.*, 1998. Procollagen traverses the Golgi stack without leaving the lumen of cisternae: evidence for cisternal maturation. *Cell* 95, 993–1003.
- Bonifacino, J.S., 2014. Adaptor proteins involved in polarized sorting. *Journal of Cell Biology* 204, 7–17.
- Brockhausen, I., Schachter, H., Stanley, P., 2009. O-GalNAc glycans. In: Varki, A., Cummings, R.D., Esko, J.D. *et al.* (Eds.), *Essentials of Glycobiology*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 115–128.
- Cancino, J., Capalbo, A., Di Campli, A., *et al.*, 2014. Control systems of membrane transport at the interface between the endoplasmic reticulum and the Golgi. *Developmental Cell* 30, 280–294.
- Cavalier-Smith, T., 1987. Eukaryotes with no mitochondria. *Nature* 326, 332–333.
- Chia, J., Goh, G., Racine, V., *et al.*, 2012. RNAi screening reveals a large signaling network controlling the Golgi apparatus in human cells. *Molecular Systems Biology* 8, 629.
- Clermont, Y., Rambourg, A., Hermo, L., 1994. Connections between the various elements of the *cis*- and mid-compartments of the Golgi apparatus of early rat spermatids. *Anatomical Record* 240, 469–480.
- D'Angelo, G., Capasso, S., Sticco, L., Russo, D., 2013. Glycosphingolipids: Synthesis and functions. *The FEBS Journal* 280, 6338–6353.
- De Matteis, M.A., Luini, A., 2008. Exiting the Golgi complex. *Nature Reviews Molecular Cell Biology* 9, 273–284.
- De Matteis, M.A., Luini, A., 2011. Mendelian disorders of membrane trafficking. *New England Journal of Medicine* 365, 927–938.
- Diekmann, Y., Seixas, E., Gouw, M., *et al.*, 2011. Thousands of rab GTPases for the cell biologist. *PLoS Computational Biology* 7, e1002217.
- Dunphy, W.G., Rothman, J.E., 1985. Compartmental organization of the Golgi stack. *Cell* 42, 13–21.
- Esko, J.D., Kimata, K., Lindahl, U., 2009. Proteoglycans and sulfated glycosaminoglycans. In: Varki, A., Cummings, R.D., Esko, J.D. *et al.* (Eds.), *Essentials of Glycobiology*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (Chapter 16). Available at: <http://www.ncbi.nlm.nih.gov/books/NBK1900/#top> (accessed 11.12.14).
- Farhan, H., Wendeler, M.W., Mitrovic, S., *et al.*, 2010. MAPK signaling to the early secretory pathway revealed by kinase/phosphatase functional screening. *Journal of Cell Biology* 189, 997–1011.
- Farquhar, M.G., Palade, G.E., 1998. The Golgi apparatus: 100 years of progress and controversy. *Trends in Cell Biology* 8, 2–10.
- Freeze, H.H., Elbein, A.D., 2009. Glycosylation precursors. In: Varki, A., Cummings, R.D., Esko, J.D. *et al.* (Eds.), *Essentials of Glycobiology*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (Chapter 4). Available at: <http://www.ncbi.nlm.nih.gov/books/NBK1929/> (accessed 11.12.14).
- Freeze, H.H., Ng, B.G., 2011. Golgi glycosylation and human inherited diseases. *Cold Spring Harbor Perspectives in Biology* 3, a005371.
- Gannon, J., Bergeron, J.J., Nilsson, T., 2011. Golgi and related vesicle proteomics: Simplify to identify. *Cold Spring Harbor Perspectives in Biology* 3, a005421.
- Gilchrist, A., Au, C.E., Hiding, J., *et al.*, 2006. Quantitative proteomics analysis of the secretory pathway. *Cell* 127, 1265–1281.
- Gill, D.J., Chia, J., Senewiratne, J., Bard, F., 2010. Regulation of O-glycosylation through Golgi-to-ER relocation of initiation enzymes. *Journal of Cell Biology* 189, 843–858.
- Glick, B.S., Luini, A., 2011. Models for Golgi traffic: A critical assessment. *Cold Spring Harbor Perspectives in Biology* 3, a005215.
- Griffiths, G., Fuller, S.D., Back, R., *et al.*, 1989. The dynamic nature of the Golgi complex. *Journal of Cell Biology* 108, 277–297.
- Griffiths, G., Warren, G., Quinn, P., Mathieu-Costello, O., Hoppeler, H., 1984. Density of newly synthesized plasma membrane proteins in intracellular membranes. I. Stereological studies. *Journal of Cell Biology* 98, 2133–2141.

- Gu, F., Crump, C.M., Thomas, G., 2001. *Trans*-Golgi network sorting. *Cellular and Molecular Life Sciences: CMLS* 58, 1067–1084.
- Hanada, K., Kumagai, K., Tomishige, N., Yamaji, T., 2009. CERT-mediated trafficking of ceramide. *Biochimica et Biophysica Acta* 1791, 684–691.
- Hirschberg, C.B., Snider, M.D., 1987. Topography of glycosylation in the rough endoplasmic reticulum and Golgi apparatus. *Annual Review of Biochemistry* 56, 63–87.
- Hirschberg, K., Miller, C.M., Ellenberg, J., *et al.*, 1998. Kinetic analysis of secretory protein traffic and characterization of golgi to plasma membrane transport intermediates in living cells. *Journal of Cell Biology* 143, 1485–1503.
- Hsu, V.W., Yang, J.S., 2009. Mechanisms of COPI vesicle formation. *FEBS Letters* 583, 3758–3763.
- Hu, Z., Zeng, L., Huang, Z., Zhang, J., Li, T., 2007. The study of Golgi apparatus in Alzheimer's disease. *Neurochemical Research* 32, 1265–1277.
- Jamieson, J.D., Palade, G.E., 1967. Intracellular transport of secretory proteins in the pancreatic exocrine cell. I. Role of the peripheral elements of the Golgi complex. *Journal of Cell Biology* 34, 577–596.
- Klumperman, J., 2000. Transport between ER and Golgi. *Current Opinion in Cell Biology* 12, 445–449.
- Klumperman, J., 2011. Architecture of the mammalian Golgi. *Cold Spring Harbor Perspectives in Biology* 3, a005181.
- Klute, M.J., Melancon, P., Dacks, J.B., 2011. Evolution and diversity of the Golgi. *Cold Spring Harbor Perspectives in Biology* 3, a007849.
- Ladinsky, M.S., Kremer, J.R., Furcinitti, P.S., McIntosh, J.R., Howell, K.E., 1994. HVEM tomography of the *trans*-Golgi network: Structural insights and identification of a lace-like vesicle coat. *Journal of Cell Biology* 127, 29–38.
- Lanoix, J., Ouwendijk, J., Stark, A., *et al.*, 2001. Sorting of Golgi resident proteins into different subpopulations of COPI vesicles: A role for ArfGAP1. *Journal of Cell Biology* 155, 1199–1212.
- Lau, K.S., Partridge, E.A., Grigorian, A., *et al.*, 2007. Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. *Cell* 129, 123–134.
- Lavie, G., Zheng, H., Rothman, J.E., 2013. Stapled Golgi cisternae remain in place as cargo passes through the stack. *Elife* 2, e00558.
- Lippincott-Schwartz, J., Yuan, L., Tipper, C., *et al.*, 1991. Brefeldin A's effects on endosomes, lysosomes, and the TGN suggest a general mechanism for regulating organelle structure and membrane traffic. *Cell* 67, 601–616.
- Lord, C., Bhandari, D., Menon, S., *et al.*, 2011. Sequential interactions with Sec23 control the direction of vesicle traffic. *Nature* 473, 181–186.
- Lorente-Rodriguez, A., Barlowe, C., 2011. Entry and exit mechanisms at the *cis*-face of the Golgi complex. *Cold Spring Harbor Perspectives in Biology* 3, a005207.
- Losev, E., Reinke, C.A., Jellen, J., *et al.*, 2006. Golgi maturation visualized in living yeast. *Nature* 441, 1002–1006.
- Luini, A., Mavelli, G., Jung, J., Cancino, J., 2014. Control systems and coordination protocols of the secretory pathway. *F1000Prime Reports* 6, 88.
- Luini, A., Mironov, A.A., Polishchuk, E.V., Polishchuk, R.S., 2008. Morphogenesis of post-Golgi transport carriers. *Histochemistry and Cell Biology* 129, 153–161.
- Luzio, J.P., Pryor, P.R., Bright, N.A., 2007. Lysosomes: Fusion and function. *Nature Reviews Molecular Cell Biology* 8, 622–632.
- Malhotra, V., 2013. Unconventional protein secretion: An evolving mechanism. *EMBO Journal* 32, 1660–1664.
- Malhotra, V., Serafini, T., Orci, L., Shepherd, J.C., Rothman, J.E., 1989. Purification of a novel class of coated vesicles mediating biosynthetic protein transport through the Golgi stack. *Cell* 58, 329–336.
- Malsam, J., Sollner, T.H., 2011. Organization of SNAREs within the Golgi stack. *Cold Spring Harbor Perspectives in Biology* 3, a005249.
- Martinez-Menarguez, J.A., Geuze, H.J., Slot, J.W., Klumperman, J., 1999. Vesicular tubular clusters between the ER and Golgi mediate concentration of soluble secretory proteins by exclusion from COPI-coated vesicles. *Cell* 98, 81–90.
- Martinez-Menarguez, J.A., Prekeris, R., Oorschot, V.M., *et al.*, 2001. Peri-Golgi vesicles contain retrograde but not anterograde proteins consistent with the cisternal progression model of intra-Golgi transport. *Journal of Cell Biology* 155, 1213–1224.
- Matsuura-Tokita, K., Takeuchi, M., Ichihara, A., Mikuriya, K., Nakano, A., 2006. Live imaging of yeast Golgi cisternal maturation. *Nature* 441, 1007–1010.
- Mironov, A.A., Beznoussenko, G.V., Nicoziani, P., *et al.*, 2001. Small cargo proteins and large aggregates can traverse the Golgi by a common mechanism without leaving the lumen of cisternae. *Journal of Cell Biology* 155, 1225–1238.
- Mogelsvang, S., Gomez-Ospina, N., Soderholm, J., Glick, B.S., Staehelin, L.A., 2003. Tomographic evidence for continuous turnover of Golgi cisternae in *Pichia pastoris*. *Molecular Biology of the Cell* 14, 2277–2291.
- Mowbrey, K., Dacks, J.B., 2009. Evolution and diversity of the Golgi body. *FEBS Letters* 583, 3738–3745.
- Munro, S., 2011. The golgin coiled-coil proteins of the Golgi apparatus. *Cold Spring Harbor Perspectives in Biology* 3, a005256.
- Palade, G., 1975. Intracellular aspects of the process of protein synthesis. *Science* 189, 347–358.
- Papanikou, E., Glick, B.S., 2014. Golgi compartmentation and identity. *Current Opinion in Cell Biology* 29C, 74–81.
- Patterson, G.H., Hirschberg, K., Polishchuk, R.S., *et al.*, 2008. Transport through the Golgi apparatus by rapid partitioning within a two-phase membrane system. *Cell* 133, 1055–1067.
- Pellett, P.A., Dietrich, F., Bewersdorf, J., Rothman, J.E., Lavie, G., 2013. Inter-Golgi transport mediated by COPI-containing vesicles carrying small cargoes. *Elife* 2, e01296.
- Pfeffer, S.R., 2010. How the Golgi works: A cisternal progenitor model. *Proceedings of the National Academy of Sciences of the United States of America* 107, 19614–19618.
- Pfeffer, S.R., 2011. Entry at the *trans*-face of the Golgi. *Cold Spring Harbor Perspectives in Biology* 3, a005272.
- Polishchuk, R.S., Polishchuk, E.V., Marra, P., *et al.*, 2000. Correlative light-electron microscopy reveals the tubular-saccular ultrastructure of carriers operating between Golgi apparatus and plasma membrane. *Journal of Cell Biology* 148, 45–58.
- Polishchuk, R.S., San Pietro, E., Di Pentima, A., Tete, S., Bonifacino, J.S., 2006. Ultrastructure of long-range transport carriers moving from the trans Golgi network to peripheral endosomes. *Traffic* 7, 1092–1103.
- Popoff, V., Adolf, F., Brugger, B., Wieland, F., 2011. COPI budding within the Golgi stack. *Cold Spring Harbor Perspectives in Biology* 3, a005231.
- Pulvirenti, T., Giannotta, M., Capestrano, M., *et al.*, 2008. A traffic-activated Golgi-based signaling circuit coordinates the secretory pathway. *Nature Cell Biology* 10, 912–922.
- Puthenveedu, M.A., Bachert, C., Puri, S., Lanni, F., Linstedt, A.D., 2006. GM130 and GRASP65-dependent lateral cisternal fusion allows uniform Golgi-enzyme distribution. *Nature Cell Biology* 8, 238–248.
- Rabouille, C., Malhotra, V., Nickel, W., 2012. Diversity in unconventional protein secretion. *Journal of Cell Science* 125, 5251–5255.
- Rambourg, A., Clermont, Y., 1990. Three-dimensional electron microscopy: Structure of the Golgi apparatus. *European Journal of Cell Biology* 51, 189–200.
- Richter, S., Voss, U., Jurgens, G., 2009. Post-Golgi traffic in plants. *Traffic* 10, 819–828.
- Rini, J., Esko, J., Varki, A., 2009. Glycosyltransferases and Glycan-processing enzymes. In: Varki, A., Cummings, R.D., Esko, J.D. *et al.* (Eds.), *Essentials of Glycobiology*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (Chapter 5). Available at: <http://www.ncbi.nlm.nih.gov/books/NBK1921/> (accessed 11.12.14).
- Rivinoja, A., Pujol, F.M., Hassinen, A., Kellokumpu, S., 2012. Golgi pH, its regulation and roles in human disease. *Annals of Medicine* 44, 542–554.
- Rizzo, R., Parashuraman, S., Mirabelli, P., *et al.*, 2013. The dynamics of engineered resident proteins in the mammalian Golgi complex relies on cisternal maturation. *Journal of Cell Biology* 201, 1027–1036.
- Rosnoblet, C., Peanne, R., Legrand, D., Foulquier, F., 2013. Glycosylation disorders of membrane trafficking. *Glycoconjugate Journal* 30, 23–31.
- Rothman, J.E., 2002. Lasker Basic Medical Research Award. The machinery and principles of vesicle transport in the cell. *Nature Medicine* 8, 1059–1062.
- Rothman, J.E., Wieland, F.T., 1996. Protein sorting by transport vesicles. *Science* 272, 227–234.
- Saraste, J., Kuismanen, E., 1992. Pathways of protein sorting and membrane traffic between the rough endoplasmic reticulum and the Golgi complex. *Seminars in Cell Biology* 3, 343–355.
- Schekman, R., 2002. Lasker Basic Medical Research Award. SEC mutants and the secretory apparatus. *Nature Medicine* 8, 1055–1058.
- Schekman, R., Orci, L., 1996. Coat proteins and vesicle budding. *Science* 271, 1526–1533.
- Schnaar, R.L., Suzuki, A., Stanley, P., 2009. Glycosphingolipids. In: Varki, A., Cummings, R.D., Esko, J.D. *et al.* (Eds.), *Essentials of Glycobiology*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (Chapter 10). Available at: <http://www.ncbi.nlm.nih.gov/books/NBK1909/> (accessed 11.12.14).
- Scott, K.L., Kabbarah, O., Liang, M.C., *et al.*, 2009. GOLPH3 modulates mTOR signaling and rapamycin sensitivity in cancer. *Nature* 459, 1085–1090.
- Seemann, J., Jokitalo, E., Pypaert, M., Warren, G., 2000. Matrix proteins can generate the higher order architecture of the Golgi apparatus. *Nature* 407, 1022–1026.
- Shorter, J., Warren, G., 2002. Golgi architecture and inheritance. *Annual Review of Cell and Developmental Biology* 18, 379–420.

- Simpson, J.C., Joggerst, B., Laketa, V., *et al.*, 2012. Genome-wide RNAi screening identifies human proteins with a regulatory function in the early secretory pathway. *Nature Cell Biology* 14, 764–774.
- Stanley, P., 2011. Golgi glycosylation. *Cold Spring Harbor Perspectives in Biology* 3, a005199.
- Stanley, P., Schachter, H., Taniguchi, N., 2009. N-Glycans. In: Varki, A., Cummings, R.D., Esko, J.D. *et al.* (Eds.), *Essentials of Glycobiology*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (Chapter 8). Available at: <http://www.ncbi.nlm.nih.gov/books/NBK1917/> (accessed 11.12.14).
- Terasaki, M., 2000. Dynamics of the endoplasmic reticulum and golgi apparatus during early sea urchin development. *Molecular Biology of the Cell* 11, 897–914.
- Thor, F., Gautschi, M., Geiger, R., Helenius, A., 2009. Bulk flow revisited: Transport of a soluble protein in the secretory pathway. *Traffic* 10, 1819–1830.
- Trucco, A., Polishchuk, R.S., Martella, O., *et al.*, 2004. Secretory traffic triggers the formation of tubular continuities across Golgi sub-compartments. *Nature Cell Biology* 6, 1071–1081.
- Varki, A., Freeze, H.H., 2009. Glycans in acquired human diseases. In: Varki, A., Cummings, R.D., Esko, J.D. *et al.* (Eds.), *Essentials of Glycobiology*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (Chapter 43). Available at: <http://www.ncbi.nlm.nih.gov/books/NBK1946/> (accessed 11.12.14).
- Varki, A., Freeze, H.H., Gagneux, P., 2009a. Evolution of glycan diversity. In: Varki, A., Cummings, R.D., Esko, J.D. *et al.* (Eds.), *Essentials of Glycobiology*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press (Chapter 19). Available at: <http://www.ncbi.nlm.nih.gov/books/NBK1942/> (accessed 11.12.14).
- Varki, A., Kannagi, R., Toole, B.P., 2009b. Glycosylation changes in cancer. In: Varki, A., Cummings, R.D., Esko, J.D. *et al.* (Eds.), *Essentials of Glycobiology*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. (Chapter 44). Available at: <http://www.ncbi.nlm.nih.gov/books/NBK1963/> (accessed 11.12.14).
- Vembar, S.S., Brodsky, J.L., 2008. One step at a time: Endoplasmic reticulum-associated degradation. *Nature Reviews Molecular Cell Biology* 9, 944–957.
- Vivero-Salmeron, G., Ballesta, J., Martinez-Menarguez, J.A., 2008. Heterotypic tubular connections at the endoplasmic reticulum-Golgi complex interface. *Histochemistry and Cell Biology* 130, 709–717.
- Wakana, Y., van Galen, J., Meissner, F., *et al.*, 2012. A new class of carriers that transport selective cargo from the trans Golgi network to the cell surface. *EMBO Journal* 31, 3976–3990.
- Wei, J.H., Seemann, J., 2010. Unraveling the Golgi ribbon. *Traffic* 11, 1391–1400.
- Willett, R., Ungar, D., Lupashin, V., 2013. The Golgi puppet master: COG complex at center stage of membrane trafficking interactions. *Histochemistry and Cell Biology* 140, 271–283.
- Xiang, Y., Zhang, X., Nix, D.B., *et al.*, 2013. Regulation of protein glycosylation and sorting by the Golgi matrix proteins GRASP55/65. *Nature Communications* 4, 1659.
- Yadav, S., Linstedt, A.D., 2011. Golgi positioning. *Cold Spring Harbor Perspectives in Biology* 3, a005322.
- Zacharogianni, M., Kondylis, V., Tang, Y., *et al.*, 2011. ERK7 is a negative regulator of protein secretion in response to amino-acid starvation by modulating Sec16 membrane association. *EMBO Journal* 30, 3684–3700.
- Zeevaert, R., Foulquier, F., Jaeken, J., Matthijs, G., 2008. Deficiencies in subunits of the Conserved Oligomeric Golgi (COG) complex define a novel group of Congenital Disorders of Glycosylation. *Molecular Genetics and Metabolism* 93, 15–21.

Early Endosomal Compartments

P van der Sluijs, UMC Utrecht, Utrecht, The Netherlands

© 2016 Elsevier Inc. All rights reserved.

Endosomes

A hallmark of eukaryotic cells is the compartmentization of their cytoplasm which allows for the parallel processing of biochemical reactions with different requirements for optimal execution. All eukaryotic cells have endosomes which constitute a pleiomorphic interconnected membrane system important for many basic cellular processes including nutrient supply, signaling, and plasma membrane protein regulation. In multicellular organisms these features contribute in an essential manner to cellular phenotypes as shape, migration, motility, and division and in doing so are of importance in developmental processes and tissue homeostasis.

Endosomes and their role in biology have been firmly established and incorporated in the textbooks. This fact on itself actually bears testimony to the enormous development in imaging technologies that allowed us to rapidly accumulate a wealth of information on an organelle whose existence was unknown in the late 1970s. At that time the lysosome, an acidic degradative organelle was considered to be the destination of all endocytosed molecules and solutes. Observations of Hubbard, Steinman and Cohn, and Goldstein and Brown indicated that the synthesis rate of endocytosed membrane lipid, cell surface proteins, and receptors was insufficient to maintain the surface area of a cell and steady state pools of the proteins they studied (Steinman *et al.*, 1983; Goldstein *et al.*, 1985). Accumulating evidence converged on the notion that salvage pathways had to exist for the return of endocytosed materials back to the cell surface. Early EM studies revealed the existence of prelysosomal vacuoles that were more electronlucent than the typical acid hydrolase containing lysosomes and which could serve as the locale for escape from lysosomal degradation (Helenius *et al.*, 1983). In kinetic experiments on the mechanism of Semliki Forest virus penetration in tissue culture cells Helenius discovered that the low pH-mediated fusion occurred prior to localization of virions in lysosomes, which demonstrated a function for the acidic prelysosomal structures. The subsequent discovery of ATP-driven acidification activity in endosomal membranes cemented the notion that endosomes share the ability with lysosomes to acidify their lumen albeit to a lesser extent (Mellman *et al.*, 1986).

Endosome Heterogeneity Reflects Functional Compartmentalization

Microscopy studies showed that some endocytosed molecules like epidermal growth factor and low-density lipoprotein moved centripetally from small peripheral cytoplasmic structures close to the cell surface to more centrally located larger multivesicular organelles and subsequently 'true' lysosomes.

Other tracers of which Transferrin (Tf) and its receptor (TfR) are the archetypical representatives accumulated transiently near the microtubule organizing center before disappearing from the cell without being degraded. Biochemical correlates were derived from internalization experiments followed by cell fractionation where endocytosed tracers initially appeared in a mildly acidic light membrane fraction ('early endosomes') and next in more acidic, heavier membranes ('late endosomes') with different but overlapping protein composition which stand apart from lysosomes in terms of enrichment profiles for markers and isolation by density gradient centrifugation (Schmid *et al.*, 1988). These and many other microscopy and fractionation studies of increasing sophistication established the paradigm that the term endosome actually encompasses a heterogeneous class of structures both in terms of morphology and function that endow the animal cell with the ability to sort a subset of endocytosed cargo molecules for recycling and re-utilization, while other cargo is dispatched to lysosomes and disposed of by degradation (Geuze *et al.*, 1984; Figure 1).

Recent studies have begun to describe the macroscopic properties of the early endosomal system as an ensemble of individual elements that continuously exchange content through fusion and fission. The system analysis approach lends credence to the notion that the net effect of vesicle interactions is the continuous merger into larger ones with increased cargo content (Foret *et al.*, 2012). Interestingly, the quantitative description follows the same mathematical principles as those that were originally derived for the assembly of colloidal particles in larger clusters.

Epithelial Cells Have Spatially Separated, Connected Endosomal Circuits

The organization of endocytic organelles and pathways in epithelial cells is more complicated. The polarized phenotype of these cells is a consequence of the interplay between a epithelial polarity establishing program, positional sensors, and a polarized trafficking machinery (Weisz and Rodriguez-Boulan, 2009; Datta *et al.*, 2011). Maintenance of the chemically distinct apical and basolateral domains is of utmost importance for the barrier function and performance of transport in this cell type. This is most dramatically testified in the rare genetic disease 'microvillus inclusion disease' where inactivating mutations in MYO-Vb cause abnormal development of the apical cell surface, mislocalization of apical protein, and severe diarrhea (Müller *et al.*, 2008).

Since endocytosis can occur from both apical and basolateral plasma membrane, epithelia essentially duplicated early endosomal compartments. Endocytosed molecules first meet basolateral or apical early endosomes (AEE) that both feed into apical recycling endosomes. From these spatially

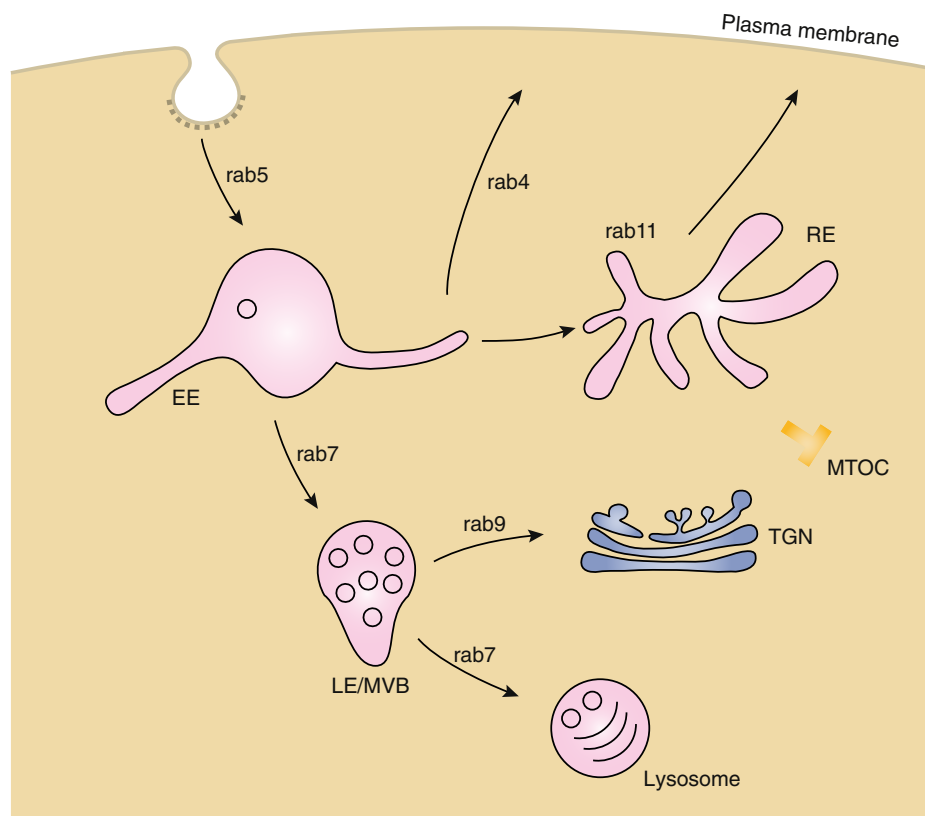


Figure 1 Overview of generic endosomal compartments. Simplified schematic of the endosomal system and its internal connectivity. Molecules that enter early endosomes (EE) are sorted in two major exit routes that either bring them back to the cell surface directly or via recycling endosomes (RE). The other pathway leads via late endosomes/multivesicular bodies (LE/MVB) to lysosomes or retrieval to the trans Golgi network (TGN). The tubular RE system is sometimes localized in the perinuclear region through interaction with the microtubule organizing center (MTOC). The core endosomal rab GTPases associated with regulation of transport through the conduits are indicated and further described in the main text.

separated early endosomes, material is either returned or transcytosed (for basolaterally internalized molecules) to the apical cell surface, or transported to a common recycling endosome above the nucleus. The common recycling endosome is the sorting station for delivery of cargo to the basolateral plasma membrane or targeting to common lysosomes. Specific sorting signals in cargo molecules are recognized by cytoplasmic machinery for incorporation in apically or basolaterally targeted carriers equipped with the appropriate SNARE proteins for fusion with apical and basolateral domains, respectively (Weisz and Rodriguez-Boulan, 2009; Figure 2).

(Ultra)structure and Specialization

Early endosomes (EE) consist of a vacuolar part from which one or more tubules emanate (Geuze *et al.*, 1984). The vacuolar portion serves to sort proteins in the degradative pathway and is also known as sorting endosome, which matures through sequential and cooperative recruitment of rab5 and rab7 effector networks into late endosomes (Rink *et al.*, 2005). The acidic endosomal pH causes dissociation of many ligands and receptors whereupon soluble ligands end up

in the vacuolar portion of the endosome (Maxfield and McGraw, 2004). The tubules may detach from the vacuolar part and oftentimes form a network, known as recycling endosomes. These typically contain a transient pool of intracellular receptors that are targeted for recycling to the plasma membrane. In some cells the tubular network has a very characteristic location adjacent to the microtubule organizing center, presumably through interaction with (peri)centriolar proteins (Maxfield and McGraw, 2004). A third exit conduit is used to recycle a subset of endosomal cargo molecules to the trans Golgi network (Bonifacino and Rojas, 2006). Finally, specialized cells have evolved unique endosome exit pathways that are vital to perform their dedicated functions, an example of which is provided by the presynaptic neuron. To prevent depletion of synaptic vesicles during neuronal stimulation, the rapid exocytosis of neurotransmitter is followed by an ultrafast endocytic pathway for retrieval of synaptic vesicles which involves the formation of new synaptic vesicles from early endosomes (Watanabe *et al.*, 2014).

A major fraction of internalized cell surface protein is recycled back from early and recycling endosomes to the plasma membrane. Because the surface to volume ratio of a thin tubule is much larger than of a spherical vesicle or the vacuolar

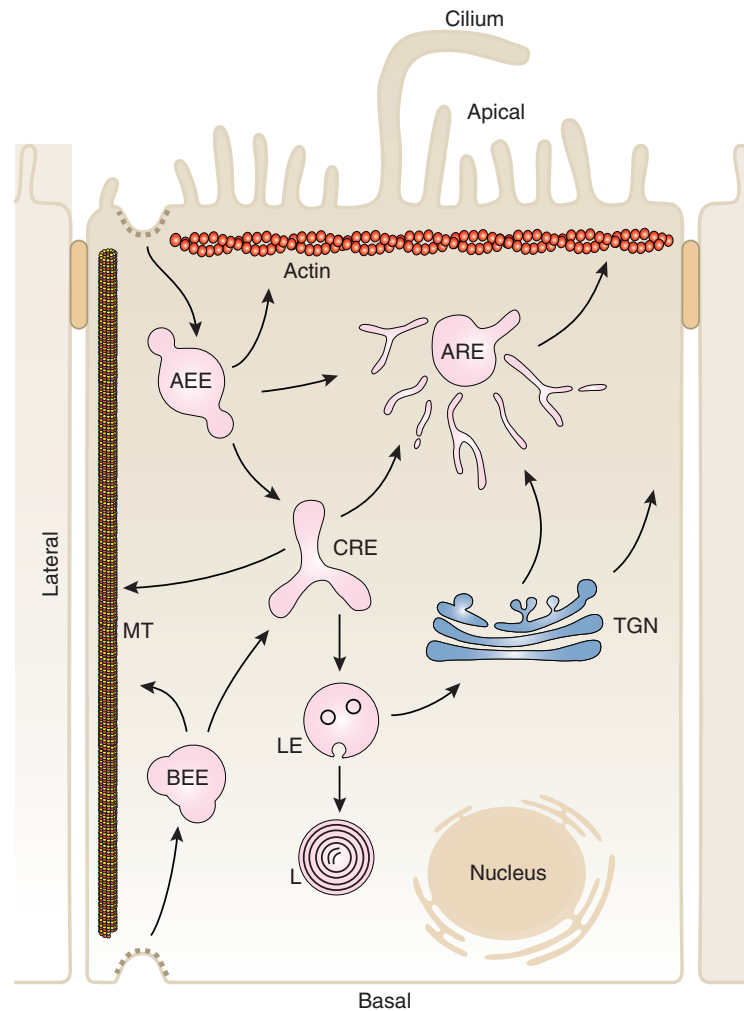


Figure 2 Endosomal organization in polarized epithelial cells. Molecules can endocytosed from the apical and basolateral cell surfaces into AEE and basolateral early endosomes (BEE), respectively. A large fraction is recycled directly, or via common recycling endosomes (CRE) or apical recycling endosomes (ARE) to the cells surface from where they originate. Depending on the sorting information present in cargo molecules, some are sorted for transport into the degradative pathway to late endosomes (LE)/lysosomes (L). Other cargo is sorted from the CRE to the cell surface opposite from where they are internalized, a pathway also known as transcytosis. Newly synthesized molecules are sorted from the TGN to the basolateral surface or via the endosomal compartments to various destinations.

part of the endosome, principles of calculus readily explain a passive endosomal sorting activity that directs receptors in tubules and ligand in the vacuolar part. For a long time the geometry-base paradigm was the driving force for the notion that sorting in the recycling pathway occurred by default.

Endosomal export has similarities with the process that delivers molecules from the cell surface to early endosomes. Although we generally refer with the single terms endocytosis and recycling to each of them, they actually represent multiple routes that end in, or leave from endosomes, respectively. From a logistic point of view transport pathways can be considered and in fact serve as a means of intracellular communication. Communication within the endosomal system and with the distinct destinations is brought about by vesicular transport pathways of unforeseen and increasing complexity. Vesicle formation is initiated by recruitment of cytoplasmic coat complexes to specific regions of the membrane, selection of cargo molecules, and reshaping the marked membrane

regions into vesicles of defined size between 40 and 80 nm (Bonifacino and Glick, 2004).

Regulation of Endosome Traffic

The architectural integrity of the endosomal system depends on sophisticated control systems for coordinating the transport routes for entry and exit of the lipids and proteins needed for maintenance of organelle identity. Two classes of molecules namely small GTPases of the rab and ADP ribosylation factor/Arf-like (Arf/Arl) families, and phosphoinositides (PIs) are thought to provide compartment specific tags and serve as key regulators of traffic and homeostasis within the endosomal system (Jean and Kiger, 2012). Endosomal membranes can be populated by more than one rab protein and may also display lateral heterogeneity in terms of phosphoinositide species which allows for the coupling of cargo entry and exit from that

membrane region. Although each in its own right can regulate independent processes, the simultaneous presence and concerted activities of a specific rab, Arf/Arl, and phosphoinositide in a membrane region is likely essential to achieve the requisite fidelity in endosomal trafficking pathways.

Rab Proteins

The best understood endosome-associated rab proteins are rab4 (early endosomes-recycling endosomes), rab5 (plasma membrane-recycling endosomes), rab7 (late endosomes-lysosomes), rab9 (late endosomes), and rab11 (recycling endosomes). Rab GTPases are molecular switches that cycle between an active guanosine triphosphate (GTP)-bound and inactive guanosine diphosphate (GDP)-bound state. GTPase Activating Proteins (GAPs) accelerate the intrinsic hydrolysis rate, while guanine nucleotide exchange factors (GEFs) catalyze the conversion of the GDP to GTP state. While the GTP-form of rab proteins is bound to the cytosolic leaflet of the endosomal membrane, the GDP form resides in the cytoplasm in a complex with the chaperone rab-GDP dissociation inhibitor (rab-GDI) that shields the hydrophobic geranylgeranyl groups in the C-terminal hypervariable region. The rab-rabGDI complex is needed for initial and reversible delivery of rabs to membranes. The specificity of which is provided by signals within the cognate GEF and the rab-rabGDI complex. Elegant studies with mitochondria-targeted cognate GEFs established that rab GEFs serve as the minimal machinery for targeting and activation of rab proteins including endosomal rab5 and rab35 (Barr, 2013). The active form of rabs binds to specific effectors and thereby creates oligomeric protein complexes that form nanodomains within the endosomal membrane. The rab-effector complexes provide spatiotemporal control of an array of processes including those that regulate shape and location of endosomes via interactions with cytoskeleton and motor proteins, phosphoinositide metabolizing enzymes, tethering factors, fusion and fission regulators, and signaling molecules (Wandinger-Ness and Zerial, 2014).

Some of the rab effectors pair with several rab proteins whose activity is sequentially organized along the endocytic route. On early endosomes, the rab5 effectors rabenosyn-5, rabaptin, and rabip4's each interact with, and have separate binding domains for rab5 and rab4 while rab4 and rab11 are coupled through such effectors as D-AKAP-2 and GRASP-1 in neurons (Wandinger-Ness and Zerial, 2014).

Bivalent effectors of endosomal rab proteins provide an attractive model to spatially link cargo delivery by incoming transport vesicles in the rab5 domain with carriers leaving endosomes from the adjacent but overlapping rab4 domain (Wandinger-Ness and Zerial, 2014). Conceptually this scenario is related to the coupling of exocytosis and re-intake of synaptic vesicles in neurons, or secretory lysosomes at the immunological synapse in cytotoxic lymphocytes. Mechanistic details of the molecular processes supporting this model however need to be established first.

Phosphoinositides

The endosomal localization of phosphoinositides just like that of rab proteins is also controlled by members of two classes of

enzymes (Jean and Kiger, 2012). Specific lipid kinases and phosphatases catalyze the formation and consumption of a substrate phosphatidylinositol. Their relative activity or location thereby drives the reversible localization of a given phosphatidylinositol (PI) in a membrane or organelle. The principal endosomal phosphatidylinositol is phosphatidylinositol 3-phosphate (PI(3)P) that to a large extent is produced by the type III kinase Vps34 on endocytic vesicles, where it controls docking and subsequent fusion of early endosomes (Wandinger-Ness and Zerial, 2014). PI(3)P binds to proteins containing a PX (after p40phox and p47phox subunits of NADPH oxidase) or FYVE domain (acronym of Fab1, YOTB, Vac1, EEA1, the founding members of this protein family) and serves as cue for endosomal localization. Some of the FYVE proteins like early endosome antigen (EEA1), rabankyrin, and rabenosyn-5 also interact with active rab5, and the simultaneous binding to PI(3)P and rab5 ensures the specificity and avidity required for their endosomal localization (Wandinger-Ness and Zerial, 2014).

PI(3)P generated by Vps34 serves as a docking site for phosphatidylinositol 3-phosphate 5-kinase PIKfyve which converts PI(3)P into PI(3,5)P₂, whose effectors are involved in sorting to the multivesicular body (Odorizzi *et al.*, 1998). PI that is funneled through Vps34 is primarily involved in initial stages of multivesicular body formation through the localization of PI(3)P effector hepatocyte growth factor-regulated tyrosine kinase substrate (Hrs) to a atypical clathrin-coated region on the vacuolar part of early endosomes (Sachse *et al.*, 2002). The adaptor that links endocytic cargo to the endosomal sorting complex required for transport (ESCRT) coat is Hrs, whose ubiquitin interacting motifs interact with ubiquitinated endosomal cargo molecules tagged for degradation (Umebayashi *et al.*, 2008). Through interactions with signal transducing adaptor molecule (STAM), Hrs assembles into a poly ubiquitin binding platform named ESCRT-0, which constitutes the first stage of ESCRT complex assembly (Bache *et al.*, 2003). A newly identified role of this PI(3)P pool relates to a localization in cytokinetic bridges during mammalian cytokinesis for recruitment of the FYVE-CENT protein (Sagona *et al.*, 2010). Together with the motor kinesin family 13A (KIF13A) and tetratricopeptide repeat domain protein 19 (TTC19), FYVE-CENT is thought to cooperate with the charged multivesicular body protein 4 (CHMP4) subunit of ESCRT-III in abscission.

Exocytic pathways leaving early endosomes do not seem to be controlled by Vps34 but instead depend on phosphoinositide 3-kinase C2α (Falasca *et al.*, 2007), a kinase that is far less sensitive to classical inhibitors LY294002 and wortmannin. The question then becomes whether early endosomes contain independent pools of PI(3)P that are involved in different aspects of early endosome function as is the case with PI(4,5)P₂ on the plasma membrane. Several recent studies now start to shed light on this question and reveal that the early endosomal system harbors distinct PI(3)P pools. Phosphatidylinositol 3-phosphate kinase (PI3K) C2α not only resides on the plasma membrane but has been found in clathrin-coated vesicles as well and in perinuclear recycling endosomes. The recycling endosome PI(3)P pool generated by PI3K C2α activates rab11 and via a rab11-rabin8-rab8 cascade regulates the function of primary cilia (Franco *et al.*, 2014).

Coordination of Phosphoinositides and Rab Proteins

PI kinases and rab proteins coordinately regulate endosome dynamics via multiple mechanisms. Often these involve the enzymatic activity of the PI kinases as with rab5 and Vps34 in endosome maturation (cf. below). In other instances the PI kinase serves a structural role, not directly related to its enzymatic activity. The function of rab11 in exocytic pathways depends on the interaction with a member of the PI4 kinases namely PI4KIII β . The two proteins co-localize on recycling endosomes and the trans Golgi network and the function of rab11 critically relies on PI4KIII β (De Graaf *et al.*, 2004). The kinase directly binds rab11 and functionally recruits it to membrane via a novel mechanism that is both dependent and independent of the lipid kinase activity (De Graaf *et al.*, 2004). In fruit fly, rab11 and the downstream effector Nuclear Fallout (ortholog of the human rab11-FIP3 (rab11 family interacting protein 3)) depend on Four Wheel Drive (PI4KIII β) for membrane localization, a prerequisite for cytokinesis (Polevoy *et al.*, 2009). Membrane ingression at the end of the asexual erythrocytic stage in the *Plasmodium falciparum* life cycle similarly depends on rab11 and PI4KIII β (McNamara *et al.*, 2013). In the rab11-PI4KIII β -FIP3 complex structure (Burke *et al.*, 2014), PI4KIII β only makes a single peripheral contact with rab11 switch region I, while molecular interactions with switch region II are absent. As a consequence the rab11 switch regions remain available for simultaneous FIP3 binding and by inference interactions with other effectors. The requirement for PI4KIII β enzyme activity likely reflects the role of its product PI(4)P in the recruitment of a DENND4 (differentially expressed in normal and neoplastic cells 4) type rab11 exchange factor (Xiong *et al.*, 2012) to activate rab11 as for the orthologs Ypt32 (yeast protein 32) and Pik1 (yeast 1-phosphatidylinositol 4-kinase) in yeast.

Early Endosome Maturation and Conversion to Late Endosomes

Endocytic internalization and early to late endosome conversion and maturation are the best understood trafficking pathways in terms of the interdependence of rab proteins and PI regulators. Rab5 and rab7 both interact with Vps34 and stimulate its enzyme activity. The increased PI(3)P levels in the endosomal rab5 domain serves to recruit proteins with FYVE signature and rab5 binding regions (cf. above) through coincidence detection. PI3K β on the cell surface and clathrin-coated vesicles is also activated by rab5 binding and generates PI(3,4,5)P₃ from PI(4,5)P₂ at these locations. Through downstream rab5 interactions with polyphosphate 4-phosphatase and inositol polyphosphate 5-phosphatase, PI(3,4,5)P₃ is converted to PI(3,4)P₂ and PI(3)P, which in parallel to the PI3K C2 α pathway contributes to the formation of a gradient of PI3P from the plasma membrane to early endosomes (Wandinger-Ness and Zerial, 2014).

The PI(3)P molecules on early endosomes are turned over via several mechanisms. First, recruitment of PIKfyve converts PI(3)P to PI(3,5)P₂. A second pathway involves activation of rab7 which sequesters Rubicon, a negative regulator of UV radiation resistance-associated gene (UVRAG) that binds the

Vps34-p150-beclin1 core complex for fusion on early endosomes (Sun *et al.*, 2010). As a consequence UVRAG activates the homotypic vacuolar protein sorting complex (HOPS) complex, a rab7 effector tethering late endosomes for fusion. The PI conversion pathways are superimposed on a rab cascade where rab5 is replaced by rab7 (Rink *et al.*, 2005). The essentials of the latter include the disruption of the positive cooperativity in the rab5 pathway via which rabex-5-rabaptin-5 exchange factor complex contributes to recruitment of additional rab5 effector molecules. Key events in this partially elucidated pathway are docking of the rab7 exchange factor Mon1-Ccz1 (Nordmann *et al.*, 2010) on rab5 and PI(3)P and removal of the rab5 exchange factor. GTP hydrolysis of rab5-GTP will cause dissociation of rab5GDP (Poteryaev *et al.*, 2010), recruitment and activation of rab7 and the assembly of a late endosomal domain populated by rab7-GTP and its own effectors such as the HOPS complex, RILP (Rab7-interacting lysosomal protein) and dynein motors (Johansson *et al.*, 2007), and retromer (Rojas *et al.*, 2008).

Proteins that do not contain the signals for targeting and degradation in late endosomes/lysosomes are either retrieved to the trans Golgi network or recycled to the plasma membrane. The transport carriers that are used in both exit ways are formed from tubular extensions on early endosomes. In the following sections, these pathways will be discussed in more detail.

Retrieval of Proteins from Endosomes to the Trans Golgi Network

Early work by Seaman and Emr (Seaman *et al.*, 1998) led to the identification of retromer as a coat complex required for the return of Vps10, a trans Golgi network (TGN) sorting receptor for the vacuolar enzyme carboxypeptidase Y and counterpart of the mammalian Mannose 6-Phosphate Receptor. The subunits of yeast retromer are organized in two interacting subcomplexes consisting of Vps26-Vps35-Vps29 trimer and a Vps5-Vps17 heterodimer. The first of which is highly conserved, interacts with cargo (Notwehr *et al.*, 2000) and is known as cargo selection complex, while Vps5 and Vps7 are member of the sorting nexin (SNX) family with more sequence diversity. Vps5 and Vps7 contain BAR (Bin-Amphiphysin-Rvs) domains that dimerize into a banana-shaped structure which can sense membrane curvature and generate membrane tubules (Peter *et al.*, 2004). In mammalian cells the cargo selection complex is associated with combinations of SNX1 and SNX2 with either SNX5 or SNX6 that all contain BAR and PX domains (SNX-BAR) proteins (Wassmer *et al.*, 2009). Docking of the cargo selection complex on membranes requires direct binding with rab7 (Rojas *et al.*, 2008) on early/maturing endosomes, and the affinity of the SNX-BAR subcomplex for PI(3)P in this endosomal membrane domain (Rojas *et al.*, 2008). The Vps29 subunit can recruit TBC1D5, a rab7-GAP whose activity converts rab7 to the inactive GDP form (Seaman *et al.*, 2009) that does not bind retromer.

More recent studies in *Caenorhabditis elegans* and *Drosophila melanogaster* uncovered a distinct retromer complex consisting of the cargo recognition complex and SNX3, an SNX without BAR domain. This functionally divergent retromer is required

for the retrieval of the wnt receptor wntless from endosomes to the TGN (Harterink *et al.*, 2011) and uses the PX domain of SNX3 for binding to membrane. In accordance with the lack of a BAR domain, the SNX3 retromer concentrates wntless cargo in membrane regions that are distinct from the tubular structures of the archetype SNX-BAR retromer and they represent earlier endocytic domains (Harterink *et al.*, 2011).

The existence of an independent second retromer pathway ferrying different cargo from endosomes back to the TGN is a conceptual advance in our understanding of the cellular organization of retrograde transport pathways. Conceivably, the combination of a conserved cargo selection complex with different SNX proteins could generate the requisite diversity for collecting and returning distinct cargo molecules from spatially separated endosomal domains back to the TGN.

To complicate matters further, not all traffic from endosomes back to the TGN involves retromer. Cargo molecules including shiga toxin and TGN38 are retrieved from endosomes to the TGN via mechanisms that operate without retromer deployment, and are only beginning to be understood recently. One of them invokes recruitment of the pleckstrin homology (PH) domain protein evectin-2 to recycling endosome domains enriched in phosphatidylserine, which is essential for the return of shiga toxin and TGN38 to the TGN (Uchida *et al.*, 2011). A more detailed description of this and the other salvage routes to the TGN is covered elsewhere in this volume.

Recycling from Early Endosomes to the Plasma Membrane

Early studies on transferrin recycling suggested that endosomal recycling occurs by default through geometric sorting into tubules which did not seem to need sorting signals in the cargo (Maxfield and McGraw, 2004). Conceptually this represents an oddity against the common and conserved principles that the cell elsewhere uses to sort proteins in transport carriers for vectorial transport (Bonifacino and Glick, 2004). Over the years, however, a plethora of studies converged on the notion that segregation and deviation from the path to lysosomes depends on signals in the cargo for transport from endosomes to such destinations as plasma membrane (Hsu *et al.*, 2012), TGN (Bonifacino and Rojas, 2006), and melanosomes (Theos *et al.*, 2005).

We now will address the recycling routes from endosomes to the plasma membrane. Morphological and biochemical tracer studies using TfR as a marker established the existence of a fast direct recycling route from early endosomes to the plasma membrane and a slow indirect pathway via recycling endosomes (Maxfield and McGraw, 2004). Several small GTPases including arf6, rab4, rab11, and rab35 localize to early endosomes/recycling endosomes and controls distinct stages and modalities of endosome recycling.

Arf6 was originally discovered as a regulator TfR recycling pathway (Peters *et al.*, 1995). It does so at multiple levels including the formation of a ACAP1 coat containing clathrin and the GAP protein ACAP1 that recognizes sorting sequence information in the cytoplasmic tail of cargo proteins (Hsu *et al.*, 2012). Arf6 also interacts with the sec10 subunit of the

octameric exocyst vesicle-tethering complex. Functional studies of this interaction suggest that arf6 specifies delivery of recycling endosome membrane through interaction with the exocyst complex (Prigent *et al.*, 2003).

Rab4 and rab11 activities are sequentially organized. Rab4 is a regulator of fast recycling, while rab11 distribution partially overlaps with rab4 and regulates the slow pathway (Wandinger-Ness and Zerial, 2014). Rab35 also regulates a fast recycling pathway (Allaire *et al.*, 2010) together with arf6 (Chesneau *et al.*, 2012). Rab4 acts upstream of rab11 since endocytosed Tf first meets rab4 and subsequently rab11 (Sönnichsen *et al.*, 2000). The function of rab4 is at the level of formation of recycling vesicles from early endosomes as found by *in vitro* reconstitution assays (Pagano *et al.*, 2004). It does so via recruitment of adaptor complexes of which adaptor protein complex 1 (AP-1) interacts directly (Perrin *et al.*, 2013) and via the effector rabaptin-5a (Deneka *et al.*, 2003) to rab4. The molecular mechanism for this pathway involves a GTPase cascade that is orchestrated by rab4, which results in association of Arl1, type I Arfs and the AP-1, adaptor protein complex 3 (AP-3), and GGA3 (Golgi-localized, γ -adapting ear-containing, ARF-binding protein) adapter complexes to endosomes (D'Souza *et al.*, 2014).

The mechanism of rab11 function in recycling occurs at multiple levels in the pathway. First, through interaction with the sec15 subunit of the octameric exocyst tethering complex (Wu *et al.*, 2005), rab11 can recruit other subunits of the complex to recycling vesicles and thereby regulate fusion at the plasma membrane (Takahashi *et al.*, 2012). Rab11 and its effector rab11-FIP also interact with eps15 homology domain (EHD) ATPase proteins (Naslavsky *et al.*, 2006) that regulate tubule formation and vesiculation (Cai *et al.*, 2013) and endosomal recycling (Pant *et al.*, 2009).

Other aspects of Rab11 function in recycling relate to a role in connecting endosomal membrane to the cytoskeleton through motor recruitment. Rab11 interacts in several ways with myosin V motors for motility of recycling vesicles in the peripheral cytoplasm (Welz *et al.*, 2014). Rab11 is essential for recycling endosome tubule morphogenesis and the distribution of the tubules toward the cell periphery through interaction with the microtubule motor KIF13A (Delevoeye *et al.*, 2014). Interestingly in melanocytes, KIF13A in complex with the AP-1 adapter is needed for the generation and positioning of peripheral recycling endosomal domains required for sorting of melanogenic enzymes to maturing melanosomes (Delevoeye *et al.*, 2009). The relation between recycling endosomes and melanosomes possibly reflects a more common principle that is used in specialized cells with lysosome-related organelles. In several cell types of the immune system including cytotoxic lymphocytes and mast cells, rab 11-positive recycling endosomes supply exocytic machinery to immature lysosome-related organelles that is needed for their maturation and fusion of the latter with the plasma membrane (Menager *et al.*, 2007; Elstak *et al.*, 2011).

Investigations into trafficking pathways of specific cargo molecules as LDL-related protein1 (LRP1) (Van Kerkhof *et al.*, 2005), potassium channels (Lunn *et al.*, 2007) and β 2-adrenergic receptors (Puthenveedu *et al.*, 2010) unveiled a unforeseen role for SNX17 and SNX27 and retromer in recycling of these ligands. SNX17 and SNX27 lack a BAR domain but

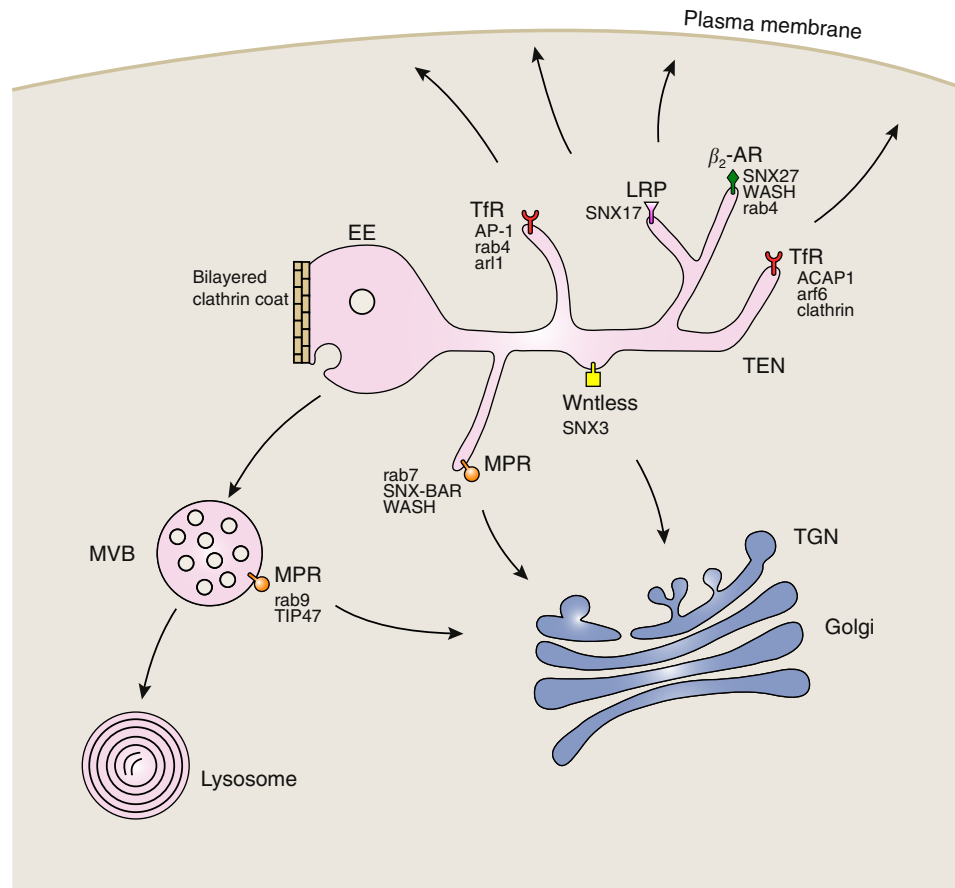


Figure 3 Sorting in endosomal exit pathways. Overview of sorting mechanisms within early endosomes (EE) and the connected tubular endosomal network (TEN). For the sake of simplicity the distinct recycling pathways for Transferrin receptor (TfR), LDL receptor-related protein (LRP) and β_2 -adrenergic receptor (β_2 -AR) are drawn from a continuous network of tubules emanating from the EE and from which recycling carriers bud. Sorting complexes for some of the recycling routes to the plasma membrane and retrieval to the trans Golgi network (TGN) are shown. Cargo destined for degradation such as EGFR is concentrated in the bilayered clathrin coat and subsequently sorted into regions from where intraluminal vesicles are formed and budding into the lumen of multivesicular bodies (MVB).

instead contain post synaptic density (PDZ) and FERM (4.1 protein, ezrin, radixin, moesin) domains. The PDZ domain of SNX27 interacts with signals in the cytoplasmic tail of the β_2 -adrenergic receptor (Lauffer *et al.*, 2010) and is needed together with the cargo selection complex for rapid retrieval of the receptor to the plasma membrane (Temkin *et al.*, 2011). Retromer-mediated recycling of β_2 -adrenergic occurs from rab4-associated endosomal tubules that are decorated with F-actin and distinct from the tubules primarily involved in TfR recycling (Puthenveedu *et al.*, 2010).

The actin patches are generated by the multisubunit Wiskott-Aldrich syndrome protein and SCAR homolog (WASH) complex, a nucleation promoting factor that binds Vps35 (Gomez and Billadeau, 2009) and SNX27 (Temkin *et al.*, 2011). WASH activates actin-related proteins 2/3 (Arp2/3) complex and drives actin branching on the endosomal tubules (Derivery *et al.*, 2009). WASH-regulated actin dynamics form a prerequisite for sorting β_2 -adrenergic receptor to the cell surface (Puthenveedu *et al.*, 2010). The extensive interactions between WASH and retromer suggest an attractive model in which actin and retromer-dependent remodeling of the endosomal surface cooperate for efficient recycling.

With this new role of the cargo selection complex in endosomal recycling, the functions of retromer have now diversified beyond those of the SNX-BAR and SNX3 – retromer complexes at the endosome TGN interface. In accord with a central role of SNX27-retromer in endosome recycling, many plasma membrane proteins have been identified that rely on SNX27 for recycling and bypassing of the degradative pathway (Steinberg *et al.*, 2013; Figure 3).

Conclusion and Perspectives

Nature's choice to allocate cellular functions to discrete compartments importantly shaped our thinking of the strategies to decipher the workings of a cell. By taking apart the cell part and studying individual elements, we made fast and great progress in understanding the ramifications of the endo-lysosomal system within the core fabric of life. We charted and partially delineated the complexity of pathways leading to and from this highly dynamic organelle. We also discovered many molecules that safeguard the uninterrupted flow of information within the endosomal system and from the upstream

and downstream neighboring organelles of the trafficking routes.

Given the massive flux of proteins and lipids that enter and leave the endosomal system we need to address now how compartmental integrity is maintained. In other words, how are input and output of the distinct pathways integrated, what is the nature and hierarchy of the control systems that coordinate and achieve homeostasis. These questions also ask for applying control and system theory to provide quantitative information needed for describing and predicting the behavior of the system in terms of response to, for instance, nutritional alterations.

Perhaps related to this problem is the need to explore whether and how the separate protein complexes that regulate endosomal recycling cooperate to sustain a specific pathway. We may ask which of the rab proteins co-organizes the SNX27-retromer pathway. Not only will this bring us closer to an understanding of the organizational principles of exit routes, we may also learn whether the rab7-SNX-BAR retromer co-incidence detection reflects an example of a general principle.

In recent years it has become evident that endosomes directly and functionally interact with the endoplasmic reticulum, the largest cellular compartment and store for ions and lipids. Signaling of the epidermal growth factor (EGF) receptor involves contacts between endosomes and endoplasmic reticulum membrane containing the protein tyrosine phosphatase PTP1B (Eden *et al.*, 2010). Importantly endoplasmic reticulum contacts were recently shown to determine the sites, timing, and efficiency of endosome fission events (Rowland *et al.*, 2014). Conceivably, such contacts may contribute to, or control sorting processes within the endosomal system. Elucidating the molecular principles underlying the contact sites may therefore be of great interest to understand the dynamics of the endosomal system as well as of the endoplasmic reticulum.

See also: Organelles: Structure and Function: The Endoplasmic Reticulum

References

Allaire, P.D., Marat, A.L., Dall'A Arm, C., *et al.*, 2010. The Connecdenn DENN domain: A GEF for rab35 mediating cargo-specific exit from early endosomes. *Molecular Cell* 37, 370–382.

Bache, K.G., Brech, A., Mehlum, A., Stenmark, H., 2003. Hrs regulates multivesicular body formation via ESCRT recruitment to endosomes. *Journal of Cell Biology* 162, 435–442.

Barr, F.A., 2013. Rab GTPases and membrane identity: Causal or inconsequential. *Journal of Cell Biology* 202, 191–199.

Bonifacino, J.S., Glick, B.S., 2004. The mechanisms of vesicle budding and fusion. *Cell* 116, 153–166.

Bonifacino, J.S., Rojas, R., 2006. Retrograde transport from endosomes to the trans Golgi network. *Nature Reviews Molecular Cell Biology* 7, 568–579.

Burke, J.E., Inglis, A.J., Peisic, O., *et al.*, 2014. Structures of PI4KIIb complexes show simultaneous recruitment of rab11 and its effectors. *Science* 344, 1035–1038.

Cai, B., Giridharan, S.S., Zhang, J., *et al.*, 2013. Differential roles of C-terminal Eps15 homology domain proteins as vesiculators and tubulators of recycling endosomes. *Journal of Biological Chemistry* 288, 30172–30180.

Chesneau, L., Dambournet, D., Machicoane, M., *et al.*, 2012. An ARF6/rab35 GTPase cascade for endocytic recycling and successful cytokinesis. *Current Biology* 22, 147–153.

D'Souza, R.S., Semus, R., Billings, E.A., *et al.*, 2014. Rab4 orchestrates a small GTPase cascade for recruitment of adaptor proteins to early endosomes. *Current Biology* 24, 1187–1198.

Datta, A., Bryant, D.M., Mostov, K.E., 2011. Molecular regulation of lumen morphogenesis. *Current Biology* 21, R126–R136.

De Graaf, P., Zwart, W.T., Van Dijken, R.A.J., *et al.*, 2004. Phosphatidylinositol 4-kinasebeta is critical for functional association of rab11 with the Golgi complex. *Molecular Biology of the Cell* 15, 2038–2047.

Delevoye, C., Hurbain, I., Tenza, D., *et al.*, 2009. AP-1 and KIF13A coordinate endosomal sorting and positioning during melanosome biogenesis. *Journal of Cell Biology* 187, 247–264.

Delevoye, C., Miserey-Lenkei, S., Montagnac, G., *et al.*, 2014. Recycling endosome tubule morphogenesis from sorting endosomes requires the kinesin motor KIF13A. *Cell Reports* 6, 445–454.

Deneka, M., Neef, M., Popa, I., *et al.*, 2003. Rabaptin-5a/rabaptin-4 serves as a linker between rab4 and γ 1-adaptin in membrane recycling from endosomes. *EMBO Journal* 22, 2645–2657.

Derivery, E., Sousa, C., Gautier, J.J., *et al.*, 2009. The Arp2/3 activator WASH controls the fission of endosomes through a large multiprotein complex. *Developmental Cell* 17, 712–723.

Eden, E.R., White, I.J., Tsapara, A., Futter, C.E., 2010. Membrane contacts between endosomes and ER provide sites for PTP1B-epidermal growth factor receptor interaction. *Nature Cell Biology* 12, 267–272.

Elstak, E.D., Neef, M., Nehme, N.T., *et al.*, 2011. Munc13-4 rab27 complex is specifically required for tethering secretory lysosomes at the plasma membrane. *Blood* 118, 1570–1578.

Falasca, M., Hughes, W.E., Dominguez, V., *et al.*, 2007. The role of phosphoinositide 3-kinase C2alpha in insulin signaling. *Journal of Biological Chemistry* 282, 28226–28236.

Foret, L., Dawson, J.E., Villaseor, R., *et al.*, 2012. A general theoretical framework to infer endosomal network dynamics from quantitative images. *Current Biology* 22, 1381–1390.

Franco, I., Gulluni, F., Campa, C.C., *et al.*, 2014. PI3 class II alpha controls spatially restricted endosomal PtdIns3P and rab11 activation to promote primary cilium function. *Developmental Cell* 28, 647–658.

Geuze, H.J., Slot, J.W., Strous, G.J., *et al.*, 1984. Intracellular receptor sorting during endocytosis: Comparative immuno electron microscopy of multiple receptors in rat liver. *Cell* 37, 195–204.

Goldstein, J.L., Brown, M.S., Anderson, R.G., Russell, D.W., Schneider, W.J., 1985. Receptor-mediated endocytosis: Concepts emerging from the LDL receptor system. *Annual Review of Cell and Developmental Biology* 1, 1–39.

Gomez, T.S., Billadeau, D.D., 2009. A FAM21 containing WASH complex regulates retromer dependent sorting. *Developmental Cell* 17, 699–711.

Harterink, M., Port, F., Lorenowicz, M.J., *et al.*, 2011. A SNX3-dependent retromer pathway mediates retrograde transport of the Wnt sorting receptor Wntless and is required for Wnt secretion. *Nature Cell Biology* 13, 914–923.

Helenius, A., Mellman, I., Wall, D., Hubbard, A.L., 1983. Endosomes. *Trends in Biochemical Sciences* 8, 45–50.

Hsu, V.W., Bai, M., Li, J., 2012. Getting active: Protein sorting in endocytic recycling. *Nature Reviews Molecular Cell Biology* 13, 323–328.

Jean, S., Kiger, A.A., 2012. Coordination between rab GTPase and phosphoinositide regulation and functions. *Nature Reviews Molecular Cell Biology* 13, 463–470.

Johansson, M., Rocha, N., Zwart, W.T., *et al.*, 2007. Activation of endosomal dynein motors by stepwise assembly of rab7-RILP-p150Glued, ORP1L and the receptor betalllspectrin. *Journal of Cell Biology* 176, 459–471.

Lauffer, B.E.L., Melero, C., Temkin, P., *et al.*, 2010. SNX27 mediates PDZ-directed sorting from endosomes to the plasma membrane. *Journal of Cell Biology* 190, 565–574.

Lunn, M.L., Nassirpour, R., Arrabit, C., *et al.*, 2007. A unique sorting nexin regulates trafficking of potassium channels via a PDZ domain interaction. *Nature Neuroscience* 10, 1249–1259.

Maxfield, F.R., McGraw, T.E.M., 2004. Endocytic recycling. *Nature Reviews Molecular Cell Biology* 5, 121–132.

McNamara, C.W., Lee, M.C., Lim, C.S., *et al.*, 2013. Targeting plasmodium PI(4)K to eliminate malaria. *Nature* 504, 248–253.

Mellman, I., Fuchs, R., Helenius, A., 1986. Acidification of the endocytic and exocytic pathways. *Annual Review of Biochemistry* 55, 712–722.

Menager, M.M., Menasche, G., Romao, M., *et al.*, 2007. Secretory cytotoxic granule maturation and exocytosis require the effector protein hMunc13-4. *Nature Immunology* 8, 257–267.

Müller, T., Hess, M.W., Schiefermeier, N., *et al.*, 2008. MYO VB mutations cause microvillus inclusion disease and disrupt epithelial cell polarity. *Nature Genetics* 40, 1163–1165.

- Naslavsky, N., Rahajeng, J., Sharma, M., Jovic, M., Caplan, S., 2006. Interactions between EHD proteins and rab11-FIP2: A role for EHD3 in early endosomal transport. *Molecular Biology of the Cell* 17, 163–177.
- Nordmann, M., Cabrera, M., Perz, A., *et al.*, 2010. The Mon1-CCz1 complex is the GEF of the late endosomal Rab7 homolog Ypt7. *Current Biology* 20, 1654–1659.
- Notwehr, S.F., Ha, S.A., Bruinsma, P., 2000. Sorting of yeast membrane proteins into a endosome to Golgi pathway involves direct interaction of their cytosolic domains. *Journal of Cell Biology* 151, 297–310.
- Odorizzi, G., Babst, M., Emr, S.D., 1998. Fab1p PtdIns(3)P5-kinase function essential for protein sorting in the multivesicular body. *Cell* 95, 847–858.
- Pagano, A., Crottet, P., Prescianotto-Baschong, C., Spiess, M., 2004. *In vitro* formation of recycling vesicles from endosomes requires AP-1/clathrin and is regulated by rab4 and the connector rabaptin-5. *Molecular Biology of the Cell* 15, 4990–5000.
- Pant, S., Sharma, M., Patel, K., *et al.*, 2009. AMPH-1/amphiphysin/Bin1 functions with RME-1/EHD1 in endocytic recycling. *Nature Cell Biology* 11, 1399–1410.
- Perrin, L., Lacas-Gervais, S., Gilleron, J., *et al.*, 2013. Rab4b controls an early endosome sorting event by interacting with the γ -subunit of the clathrin adaptor complex 1. *Journal of Cell Science* 126, 4950–4962.
- Peter, B.J., Kent, H.M., Mills, I.G., *et al.*, 2004. BAR domains as sensors of membrane curvature: The amphiphysin BAR structure. *Science* 303, 495–499.
- Peters, P., Hsu, V., Ooi, C.E., *et al.*, 1995. Overexpression of wild type and mutant ARF1 and ARF6: Distinct perturbations of nonoverlapping membrane compartments. *Journal of Cell Biology* 128, 1003–1017.
- Polevoy, G., Wei, H.C., Wong, R., *et al.*, 2009. Dual roles for the drosophila PI4-kinase four wheel drive in localizing rab11 during cytokinesis. *Journal of Cell Biology* 187, 847–858.
- Poteryaev, D., Datta, S., Ackema, K., Zerial, M., Spang, A., 2010. Identification of the switch in early to late endosome transition. *Cell* 141, 497–508.
- Prigent, M., Dubois, T., Raposo, G., *et al.*, 2003. ARF6 controls post-endocytic recycling through its downstream exocyst complex effector. *Journal of Cell Biology* 163, 1111–1121.
- Puthenveedu, M.A., Lauffer, B., Temkin, P.V.R., *et al.*, 2010. Sequence dependent sorting of recycling proteins by actin-stabilized endosomal microdomains. *Cell* 143, 761–773.
- Rink, J., Ghigo, E., Kalaidzidis, Y., Zerial, M., 2005. Rab conversion as a mechanism of progression from early to late endosomes. *Cell* 122, 735–749.
- Rojas, R., Van Vlijmen, T., Mardones, G., *et al.*, 2008. Regulation of retromer recruitment to endosomes by sequential action of rab5 and rab7. *Journal of Cell Biology* 183, 513–526.
- Rowland, A.A., Chitwood, P.J., Phillips, M.J., Voeltz, G.K., 2014. ER contact sites define the position and timing of endosome fission. *Cell* 159, 1027–1041.
- Sachse, M., Urbe, S., Oorschot, V., Strous, G.J., Klumperman, J., 2002. Bilayered clathrin coats on endosomal vacuoles are involved in protein sorting towards lysosomes. *Molecular Biology of the Cell* 13, 1313–1328.
- Sagona, A.P., Nezis, I.P., Pedersen, N.M., *et al.*, 2010. PtdIns(3 P) controls cytokinesis through KIF13A mediated recruitment of FYVE-CENT to the midbody. *Nature Cell Biology* 12, 362–371.
- Schmid, S.L., Fuchs, R., Male, P., Mellman, I., 1988. Two distinct subpopulations of endosomes involved in membrane recycling and transport to lysosomes. *Cell* 52, 73–83.
- Seaman, M.N., Harbour, M.E., Tattersall, D., Read, E., Bright, N., 2009. Membrane recruitment of the cargo-selective retromer subcomplex is catalysed by the small GTPase Rab7 and inhibited by the Rab-GAP TBC1D5. *Journal of Cell Science* 122, 2371–2382.
- Seaman, M.N.J., McCaffery, J.M., Emr, S.D., 1998. A membrane coat complex essential for endosome to Golgi retrograde transport in yeast. *Journal of Cell Biology* 142, 655–682.
- Sönnichsen, B., De Renzis, S., Nielsen, E., Rietdorf, J., Zerial, M., 2000. Distinct membrane domains on endosomes in the recycling pathway visualized by multicolor imaging of rab4, rab5, and rab11. *Journal of Cell Biology* 149, 901–913.
- Steinberg, F., Gallon, M., Winfield, M., *et al.*, 2013. A global analysis of SNX27-retromer assembly and cargo specificity reveals a function in glucose ad metal ion transport. *Nature Cell Biology* 15, 461–471.
- Steinman, R.M., Mellman, I., Muller, W.A., Cohn, Z.A., 1983. Endocytosis and the recycling of plasma membrane. *Journal of Cell Biology* 96, 1–27.
- Sun, Q., Westphal, W., Wong, K.N., Tan, I., Zhong, Q., 2010. Rubicon controls endosome maturation as a Rab7 effector. *Proceedings of the National Academy of Sciences* 107, 19338–19343.
- Takahashi, S., Kubo, K., Waguri, S., *et al.*, 2012. Rab11 regulates exocytosis of recycling vesicles at the plasma membrane. *Journal of Cell Science* 125, 4049–4057.
- Temkin, P., Lauffer, B., Jager, S., *et al.*, 2011. SNX27 mediates retromer tubule entry ad endosome to plasma membrane trafficking of signalling receptors. *Nature Cell Biology* 13, 715–721.
- Theos, A.C., Tenza, D., Martina, J.A., *et al.*, 2005. Functions of adaptor protein AP-3 and AP-1 in tyrosinase sorting from endosomes to melanosomes. *Molecular Biology of the Cell* 16, 5356–5372.
- Uchida, Y., Hasegawa, J., Chinnapen, D., *et al.*, 2011. Intracellular phosphatidylserine is essential for retrograde membrane traffic through endosomes. *Proceedings of the National Academy of Sciences of the United States of America* 108, 15846–15851.
- Umebayashi, K., Stenmark, H., Yoshimori, T., 2008. Ubc4/5 and c-abl continue to ubiquitinate EGF receptor after internalization to facilitate polyubiquitination and degradation. *Molecular Biology of the Cell* 19, 3454–3462.
- Van Kerkhof, B., Lee, J., McCormick, L., *et al.*, 2005. Sorting nexin 17 facilitates LRP recycling in the early endosome. *EMBO Journal* 24, 2851–2861.
- Wandinger-Ness, A., Zerial, M., 2014. Rab proteins and the compartmentalization of the endosomal system. *Cold Spring Harbor Perspectives in Biology* 6, a022509.
- Wassmer, T., Attar, N., Harterink, M., *et al.*, 2009. The retromer coat complex coordinates endosomal sorting and dynein-mediated transport with carrier recognition by the trans Golgi network. *Developmental Cell* 17, 110–122.
- Watanabe, S., Trimbuch, T., Camacho-Perez, M., *et al.*, 2014. Clathrin regenerates synaptic vesicles from endosomes. *Nature* 515, 228–233.
- Weisz, O.A., Rodriguez-Boulan, E., 2009. Apical trafficking in epithelial cells: Signals, clusters and motors. *Journal of Cell Science* 122, 4253–4266.
- Welz, T., Wellbourne-Wood, J., Kerkhoff, E., 2014. Orchestration of cell surface proteins by rab11. *Trends in Cell Biology* 24, 407–415.
- Wu, S., Mehta, S.Q., Pichaud, F., Bellen, H.J., Quiocho, F.A., 2005. Sec15 interacts with rab11 via a novel domain and affects rab11 localization *in vivo*. *Nature Structural & Molecular Biology* 12, 879–885.
- Xiong, T., Bayat, V., Jaiswal, M., *et al.*, 2012. CRAG is a GEF for rab11 required for rhodopsin trafficking and maintenance of adult photoreceptor cells. *PLoS Biology* 10, E1001438.

The Late Endosome

F Vacca, C Scott, and J Gruenberg, University of Geneva, Geneva, Switzerland

© 2016 Elsevier Inc. All rights reserved.

Introduction

The views on late endosomes have evolved quite significantly since the discovery of a prelysosome compartment (Straus, 1964) and of endosomes defined as a heterogeneous population of endocytic vacuoles (Helenius *et al.*, 1983). Endosomes were originally seen as a succession of transient steps along a gradual maturation process eventually leading to the lysosomes. It became apparent that endosomes along the pathway are morphologically and functionally distinct, with differences in their receptor content (Goldstein *et al.*, 1985; Geuze *et al.*, 1984; Griffiths *et al.*, 1988), acidification properties (Schmid *et al.*, 1988) and membrane dynamics (Bomsel *et al.*, 1990; Aniento *et al.*, 1993), and that early and late endosomes contain a selective set of components, notably RAB GTPases (Chavrier *et al.*, 1990). These observations eventually lead to the notions that early and late endosomes have selective sorting functions in the vacuolar apparatus, as a hub for recycling receptors and a last portal before the lysosomes, respectively. Today, it is becoming apparent that late endosomes function as a key sensing and signaling organelle that informs the cell about its nutrient and supply situation.

The Early/Recycling Endosome Network

The notions of late endosome, multivesicular body (MVB), and lysosome are often blurred, in part because of terminology, but probably also because of differences in the organization of endosomes in plants, fungi, and mammalian cells. To avoid any confusion, we will briefly introduce the general organization of the endocytic pathway across eukaryotes, before discussing the current notions on late endosome organization and functions.

In mammalian cells, endosomes play a central role in controlling the reutilization or degradation of membrane components, and thus regulate fundamental processes, including nutrient uptake, immunity, signaling, adhesion, membrane turnover, defense against pathogens, and development. Components that have been endocytosed by several pathways are delivered to a common early endosome (Figure 1). The V-ATPase, a multi-subunit proton pump (Marshansky and Futai, 2008), acidifies the endosome lumen to a $\text{pH} \approx 6.2$ in early endosomes. At this mildly acidic pH, receptors are uncoupled from their bound ligands so that ligands can be transported toward lysosomes and receptors reutilized via recycling pathways. Housekeeping receptors and other proteins are recycled back to the plasma membrane directly or indirectly via recycling endosome, while other molecules are routed toward the trans-Golgi network (TGN).

From a quantitative point of view, internalization into early endosomes and recycling back to the plasma membrane may well represent the major membrane traffic pathway in mammalian cells (Steinman *et al.*, 1983; Howes *et al.*, 2010). Early

endosomes and recycling endosomes contain selective subsets of lipids and proteins (Bohdanowicz and Grinstein, 2013; Hsu and Prekeris, 2010) and recycling endosomes are slightly less acidic than early endosomes (Yamashiro *et al.*, 1984). Beyond these differences, early and recycling endosomes form extensive networks of interconnected tubulo-cisternal elements that can exhibit a highly reticular organization in some cell types (Geuze *et al.*, 1983; Griffiths *et al.*, 1989; Marsh *et al.*, 1986; Stoorvogel *et al.*, 1996; Tooze and Hollinshead, 1991). Individual early endosome elements also exhibit a high homotypic fusion activity *in vitro* (Gruenberg *et al.*, 1989) and *in vivo* (Roberts *et al.*, 1999; Foret *et al.*, 2012), consistent with the notion that early endosomes form functional networks (Gruenberg and Howell, 1989).

To ensure that the pathways of membrane component reutilization are kept well separated from lysosome targeting – recycling receptors should not cycle through lysosomes – the early endosome functions as the key intracellular sorting station between recycling routes and transport toward lysosome for degradation. Activated signaling receptors and other proteins that need to be downregulated are sorted into luminal invaginations of the early endosome membrane, which are pinched off as free cargo-containing intraluminal vesicles (ILVs). These multivesicular regions detach and become multivesicular endosomes or bodies (Figure 1), which are often referred to as MVBs or ECVs (endosomal carrier vesicles). Eventually, ILVs and their cargo are delivered to lysosomes and degraded. Late endosomes function as a second trafficking hub in the endosomal system, at the crossroad with the autophagy pathway (Choi *et al.*, 2013) and as last sorting station in the membrane trafficking cycle to and from lysosomes, but also to other destinations, including the TGN and the plasma membrane (Bissig and Gruenberg, 2014; Huotari and Helenius, 2011; Raposo and Stoorvogel, 2013).

Biogenesis of the Multivesicular Body

ECV/MVBs function as intermediates during transport between the two-membrane networks formed by early/recycling endosomes and late endosome/lysosomes, respectively (Bissig and Gruenberg, 2013). In contrast to early or late endosomes, ECV/MVBs are not pleomorphic, and exhibit a characteristic and fairly uniform ultrastructure in most mammalian cell types. They appear as large spherical structures with a diameter of approximately $0.5 \mu\text{m}$, which contain a homogenous population of ILVs with a diameter of approximately 50 nm (Razi and Futter, 2006; Pons *et al.*, 2008); 25 nm in yeast cells (Wemmer *et al.*, 2011).

The biogenesis of ECV/MVBs from early endosome elements involves the deformation of the membrane toward the lumen during ILV formation and toward the cytoplasm during recycling tubule formation (Figure 1). Moreover, the membrane remodeling process also requires the expansion of the

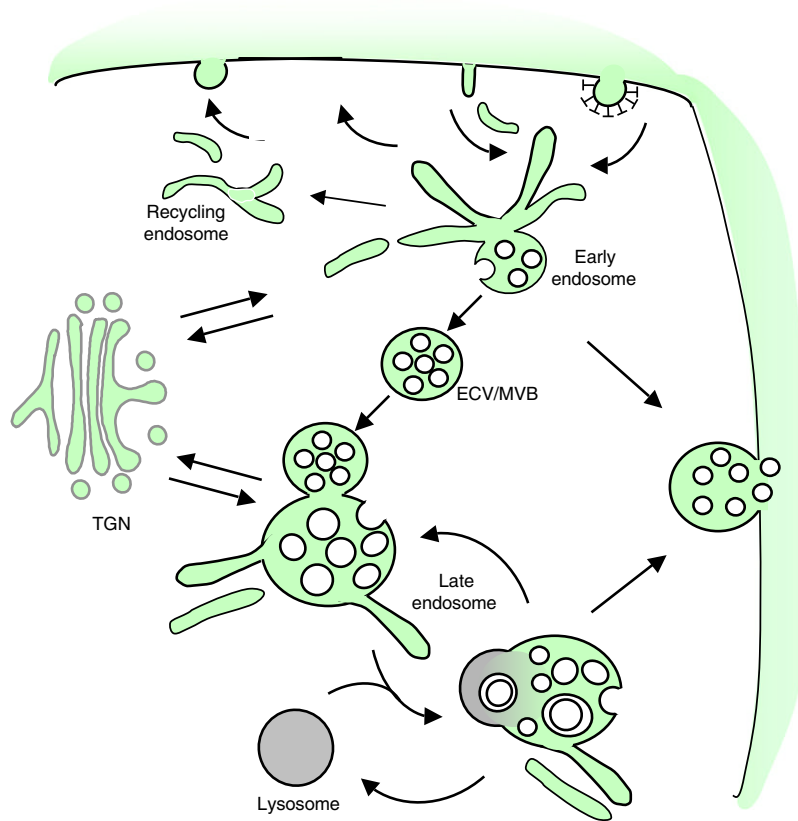


Figure 1 The figure outlines the main routes of endocytic membrane traffic. Solutes and membrane proteins, including receptors, and lipids are endocytosed by more than one pathway and these merge in a common early endosome. From there, some components, for example, housekeeping recycling receptors, are recycled to the cell surface directly or indirectly via recycling endosomes. Other proteins and lipids are transported to the trans-Golgi network (TGN), while downregulated signaling receptors are sorted into intraluminal vesicles (ILVs) that bud within the endosome lumen. These multivesicular regions then detach or mature from early endosomes and become free endosomal carrier vesicle (ECV) or multivesicular bodies (MVBs), which are then transported toward late endosomes with which they fuse. A transient intermediate, the endolysosome, is formed upon fusion of late endosomes and lysosomes, while ILVs and other components destined for degradation are eventually packaged in the lysosomes. The figure also illustrates possible pathways toward the plasma membrane followed by lysosome-related organelles (LROs) and multivesicular endosomes in nonspecialized cells.

forming ECV/MVB into a relatively large vacuole. Although this mechanism remains poorly understood, it seems that actin patches nucleated onto early endosomes via the lipid-binding protein AnxA2 drive the process (Mayran *et al.*, 2003; Morel *et al.*, 2009), and that ERM actin-binding proteins and the homotypic fusion and vacuole protein sorting (HOPS) complex may also be required (Chirivino *et al.*, 2011).

During ILV biogenesis, cargo sorting is mechanistically coupled to the membrane deformation process itself (Hurley and Stenmark, 2011), as is the case in other membrane transport steps. The sorting process is initiated by phosphatidylinositol 3-phosphate (PtdIns(3)P), a phosphoinositide generated by the PtdIns 3-kinase VPS34, which is itself an effector of the early endosome GTPase RAB5 (Shin *et al.*, 2005) – small GTPases of the RAB family are selectively associated to a subset of compartments (Galvez *et al.*, 2012) and control the function and identity of intracellular membranes. PtdIns(3)P, which can be hydrolyzed by 3-phosphatases or phosphorylated into PtdIns(3,5)P₂ by the PtdIns(3)P 5-kinase FAB1/PIKfyve

(Odorizzi *et al.*, 1998), serves as a recognition marker on endosomal membranes for PtdIns(3)P-binding proteins. These include in particular the Hrs subunit of the protein complex termed ESCRT-0 (endosomal sorting complex required for transport), which is recruited onto early endosomes via its FYVE PtdIns(3)P-binding domain and which in turn binds and clusters the ubiquitinated signaling receptor molecules. Then, through the sequential action of ESCRT-I, II and III complexes, receptors are collected into nascent ILVs (Hurley and Hanson, 2010; Henne *et al.*, 2011). The ESCRT-III complex directly contributes to the membrane deformation process leading to ILV formation via a mechanism that remains to be established (Hurley and Hanson, 2010; Henne *et al.*, 2011; Hurley and Stenmark, 2011). Plants do not express ESCRT-0 orthologs implying that as-of-yet unidentified adapter molecules must exist to fulfill an analogous cargo selecting function (Contento and Bassham, 2012). Consistent with this notion recent evidence indicates that adapters alternative to Hrs exist in the ESCRT sorting pathway of yeast and

mammalian cells (Bissig and Gruenberg, 2014). Moreover, other factors, in addition to ESCRTs and ESCRT-associated proteins are also involved in ILV biogenesis, including the PtdIns(3)P-binding proteins SNX3 and its homolog SNX12 (Pons *et al.*, 2008, 2011), as well as CD63 in melanosome biogenesis (van Niel *et al.*, 2011). ESCRT proteins are also involved in cytokinesis and virus budding at the plasma membrane, which are membrane deformation processes that share the same topology as ILV formation, and exhibit additional functions perhaps unrelated to membrane trafficking (Slagsvold *et al.*, 2006; Rusten *et al.*, 2012).

Once formed, ECV/MVBs rapidly acidify, eventually reaching pH=5.5, and mediate microtubule-dependent transport toward late endosomes, with which they can undergo fusion, thereby delivering their protein and lipid cargo (Bissig and Gruenberg, 2013; Huotari and Helenius, 2011; Figure 1). The switch from early to late endosomes along the pathway is driven by the conversion from RAB5 to RAB7, which epitomize early and late endosomes, respectively. This process, which is believed to occur on ECV/MVB membranes (Rink *et al.*, 2005), is regulated by the SAND1/MON1–CCZ1 complex (Poteryaev *et al.*, 2010; Kinchen and Ravichandran, 2010). This RAB conversion process is also involved in the mechanism that recruits the retromer to endosomal membranes (Rojas *et al.*, 2008; Seaman *et al.*, 2013) timing protein retrieval to endosome maturation. The membrane fusion specificity, which accompanies the maturation process, is ensured by substituting the class C core vacuole/endosome tethering factor (CORVET) by the late endosomal/lysosomal HOPS complex – a change concomitant with a change from the early to the late endosome SNARE docking/fusion proteins (Solinger and Spang, 2013). The change in membrane identity from early to late endosomes, by remaining well-defined in time and space, ensures that the early/recycling endosome-to-plasma membrane and late endosome-to-lysosome cycles remain topologically and functionally separated.

The Late Endosome–Lysosome Network

Late endosomes contain multivesicular regions, which unlike ECV/MVBs, exhibit highly pleomorphic and complex ultrastructures, including multilamellar or mixed multivesicular/multilamellar regions, depending on the cell type. Although the biophysical and biochemical principles intrinsic to such diverse membrane organization and packing are not known, this situation likely reflects specialized functions of intraluminal membranes in protein/lipid transport and metabolism (Bissig and Gruenberg, 2013). Moreover, in addition to endocytic membrane transport and delivery of Golgi-derived vesicles, autophagosomes represent an additional port of entry into the late endocytic stages of the pathway, and presumably also account for the pleomorphic nature of intraluminal membranes. Autophagy has been described in recent reviews (Choi *et al.*, 2013) and is introduced in another article, and will thus not be discussed further here.

Late endosomes, much like early endosomes, also exhibit high homotypic fusion activity *in vitro* (Bomsel *et al.*, 1989; Aniento *et al.*, 1993) and contain, in addition to multivesicular regions, dynamic tubulo-cisternal elements (Griffiths, 1988,

p. 415; Lebrand *et al.*, 2002; Zhang *et al.*, 2001; Ko *et al.*, 2001; Brankatschk *et al.*, 2011). The PX-containing protein SNX16 is selectively found on late endosome tubules and absent from multivesicular regions (Brankatschk *et al.*, 2011), consistent with the existence of different membrane domains in late endosomes (Barbero *et al.*, 2002) and with the mosaic organization of intracellular membranes (Gruenberg, 2001). Interestingly, J774.2 macrophages and phorbol ester-treated macrophages contain an abundant network of tubular lysosomes (Phaire-Washington *et al.*, 1980; Swanson *et al.*, 1987; Heuser, 1989), which may be of late endosomal origin (Rabinowitz *et al.*, 1992; Mrakovic *et al.*, 2012) perhaps under the control of the lipid kinase PI4KIIIb (Sridhar *et al.*, 2013). Presumably, tubular lysosomes in macrophages correspond to an extreme situation in the dynamic equilibrium between late endosomes and lysosomes.

Indeed, late endosomes and lysosomes rapidly exchange membrane components and solutes *in vivo* (Deng and Storrie, 1988; Jahraus *et al.*, 1994). These observations led to the predominant model that the two organelles undergo transient fusion and fission events, allowing an interchange of material (both cargo and enzymes). More recently with live imaging experiments, combined with ultra-structural analysis, transient (kiss-and-run) fusion of the organelles was proposed (Luzio *et al.*, 2007). The prevailing notion is that late endosomes and lysosomes form upon fusion a transient hybrid organelle, that can be referred to as endo-lysosome, which is converted into classical secondary or dense lysosomes (Mullock *et al.*, 1989; Griffiths *et al.*, 1990; Luzio *et al.*, 2007; Huotari and Helenius, 2011).

The distinction between late endosomes and lysosomes at the molecular level is further blurred by the fact that, to the best of our knowledge, there is no lysosome protein that is not also present in late endosomes. The main constituents of the membrane of both compartments are the highly sialylated membrane proteins LAMP1 and LAMP2 (Griffiths *et al.*, 1988), which form a protective glycocalyx-like layer on the luminal side of the membrane against lysosomal enzymes. While late endosomes and lysosomes are difficult to discriminate from each other at the molecular level, they can be distinguished by their physical properties and ultrastructure. In contrast to multivesicular and tubulo-cisternal elements of late endosomes, the classical secondary lysosomes are spherical electron-dense structures, originally referred to as dense bodies of variable size and content depending on the diet (Bainton, 1981; de Duve, 2005). Similarly, secondary lysosomes exhibit a high density on gradients and can be separated from endosomes by subcellular fractionation. In fact, before the discovery of late endosomes, newly synthesized lysosomal enzymes were known to transit via a compartment referred to as light lysosomes because of its low buoyancy on gradients before reaching secondary lysosomes – it is now apparent that light lysosomes are late endosomes (Brotherus and Renkonen, 1977; Sahagian and Neufeld, 1983; Aniento *et al.*, 1993; Kobayashi *et al.*, 1998; Tjelle *et al.*, 1996). Conversely, lysosomes can be defined biochemically as containing the bulk of the fully mature, dephosphorylated lysosomal enzymes.

Specialized cell types contain unique secretory compartments, collectively referred to as lysosome-related organelles

(LROs), which share some endosomal and lysosomal features and are involved in numerous processes, including pigmentation, hemostasis and immunity (Marks, 2013). In nonspecialized cells, conventional late endocytic and multivesicular compartments can acquire the capacity to undergo fusion with the plasma membrane upon a transient rise in cytosolic calcium, in a process that may depend on synaptotagmin VII (Rodriguez *et al.*, 1997; Rao *et al.*, 2004). This process may be involved in plasma membrane repair (Andrews, 2000), and is controlled by the master regulator of late endocytic functions, the transcription factor EB (TFEB) (Settembre *et al.*, 2013b). However, endo-lysosome exocytosis, which is reminiscent of LROs, may occur from more than one population of donor endosomes and may also fulfill other secretory functions (Laulagnier *et al.*, 2011), including the release of exosomes (Raposo and Stoorvogel, 2013) and storage materials in the lysosomes of mucopolidiosis type IV cells (LaPlante *et al.*, 2006).

Endosomes in Plants and Fungi

Late endosomes have sometimes been defined as MVBs in mammalian cells, but this notion is very confusing. Multivesicular regions are found at all stages of the endocytic pathway, and conversely, as we have seen above, late endosomes cannot be reduced strictly to their multivesicular elements. Part of the confusion also stems from marked differences in the organization of the endosomal system in plant, yeast and mammalian cells. In plant cells, the trans-Golgi network fulfills the role of the early endosome, being the first site for delivery of endocytosed material. Consequently, RABA1 GTPases, which belong to the large group of RAB11 homologs, are involved in TGN-to-plasma membrane transport in higher plants (Asaoka *et al.*, 2012), while RAB11 controls endosome-to-plasma recycling in mammalian cells. However, this hybrid organelle is not completely analogous to the early endosome of mammals as plants completely lack RAB4 and likely the rapid recycling pathway dependent on it (Saito and Ueda, 2009). It has been proposed that the trans-Golgi network then matures into an MVB via an annexin-dependent mechanism (Scheuring *et al.*, 2011; Viotti *et al.*, 2010) reminiscent of that found in mammalian cells (Mayran *et al.*, 2003; Morel *et al.*, 2009). Eventually, the MVB fuses with the plant vacuole, which plays the role of lysosomes in animal cells.

Unlike plants, yeast cells do not seem to contain a true hybrid early endosome/TGN-like organelle, but some endosomal functions in mammalian cells are linked to the Golgi in yeast cells. The class D VPS genes were characterized by the accumulation of Golgi-derived carrier vesicles laden with vacuolar transmembrane proteins (Raymond *et al.*, 1992). This class includes many of the canonical early endosomal proteins of mammalian cells including RAB5 (VPS21) and a VPS21 exchange factor VPS9, syntaxin 7 (PEP12/VPS6), the PtdIns 3-kinase VPS34 and its regulator VPS15, VPS19 (related to Rabenosyn-5), as well as the RAB5/VPS21 and PtdIns(3)P interacting protein VAC1 (Tall *et al.*, 1999). Consistent with this notion, several of these proteins have been localized to the TGN in yeast, implying that many of the sorting functions

ascribed to the early endosome may indeed take place in the Golgi – and earlier studies of the yeast RAB5 homologs already suggested the existence of an intersection of the endocytic and vacuolar sorting pathways (Singer-Kruger *et al.*, 1994).

Moreover, in contrast to plants, the delivery of endocytosed material to the prevacuolar compartment (late endosome) occurs via a Golgi-independent route in yeast cells (Riezman, 1985; Vida and Emr, 1995). Electron microscopy studies have shown that tracers internalized in yeast cells appear sequentially in structures that contain the tSNARE Tlg1p and then in the prevacuolar compartment that contains the tSNARE Pep12 (Prescianotto-Baschong and Riezman, 2002). However, the yeast and mammalian endosomal pathways are not equivalent and comparisons should be made with care. Although there is no doubt that basic machineries are conserved, and that studies in yeast cells have greatly helped unravel the mechanisms that regulate endocytic membrane transport, yeast cells seem to lack a compartment with the molecular or functional characteristics of early endosome in mammalian cells. In yeast, there is no evidence for the existence of direct recycling pathways that bypasses the Golgi, like the major mammalian routes of early endosome-to-plasma membrane recycling. In fact, recycling is regulated in mammalian cells by the small GTPases RAB11 and RAB4, with RAB4 controlling a faster route. Yeast cells do not express a RAB4 homolog, much like plant cells (Pereira-Leal and Seabra, 2001), and the RAB11 homologs Ypt31p/32p appear to control a pathway of protein recycling from the endosome to the Golgi (Chen *et al.*, 2005).

One should also bear in mind that recycling, secretion or exosome release can occur in mammalian cells from late multivesicular endosomes (Andrews, 2000; Laulagnier *et al.*, 2011; Marks *et al.*, 2013; Raposo and Stoorvogel, 2013) and that several lines of evidence suggest that mammalian cells contain more than one population of ILVs with different fates (discussed below). By contrast, yeast ILVs seem to have the vacuole as single destination and there is no evidence for direct pathways from the prevacuolar compartment containing the tSNARE Pep12 to the yeast plasma membrane.

Lysosome Enzymes and Lysosomal Membrane Proteins

It has long been appreciated that endosomes contain active lysosomal enzymes even at very early stages of the pathway (Diment and Stahl, 1985). Most newly synthesized lysosomal hydrolases carry a unique Man6P recognition motif and in the TGN bind their cognate Man6P receptors. From there, they are transported via coated vesicles presumably to early endosomes. In endosomes, the enzymes are released from the receptors at the low endosomal pH, and then packaged in lysosomes (Braulke and Bonifacino, 2009). The current view is that the Man6P receptors recycle from maturing endosomes back to the TGN via the retromer complex in a process that requires both RAB5 and RAB7. An additional route exists to transport the receptor from late endosomes to the Golgi, which involves the small GTPase RAB9 and its effectors (Ganley *et al.*, 2004, 2008) – whether proteins other than Man6P receptors can exploit this route is still not known. The role of this route is underscored by recent reports showing the salmonella protein sifA blocks the traffic of M6PR receptors by

sequestration of RAB9 (McGourty *et al.*, 2012). Man6P receptor retrieval to the TGN may thus occur by two sequential pathways: it begins in early endosomes via RAB5 and RAB7 and continues in late endosomes via RAB9 (Seaman, 2012; Rojas *et al.*, 2008; Johannes and Popoff, 2008; Bonifacino and Hurley, 2008; Pfeffer, 2011; Anitei and Hoflack, 2011). Lysosomal enzymes, in turn, undergo a maturation process, including dephosphorylation and often proteolytic processing, and become fully active at the low acidic pH of late endosomes and lysosomes (approximately 5.5–5.0).

The newly synthesized forms of the late endosome and lysosome transmembrane proteins, including LAMP1 and LAMP2, are transported from the TGN via a pathway different from that used by hydrolases, involving the adapter complex AP3 (Saftig and Klumperman, 2009). It has recently become apparent that lysosome biogenesis (Sardiello *et al.*, 2009), as well as autophagy (Settembre *et al.*, 2011), are under the control of the master regulator TFEB – a process that reveals how lysosomes can adapt to environmental cues (Settembre *et al.*, 2013b), and that will be discussed below. Numerous human diseases are in fact associated with defects in lysosome membrane proteins (Saftig and Klumperman, 2009; Ikonen and Holtta-Vuori, 2004) and enzymes (Futerman and van Meer, 2004; Sardiello and Ballabio, 2009; Schulze and Sandhoff, 2011). A defective enzyme causes a lysosomal storage disease because of the pathological accumulation of the corresponding substrate, which eventually jams the endolysosomal system. Some lipid storage diseases however are not caused by a defective enzyme, including the cholesterol storage disease Niemann-Pick type C, caused by mutations in the NPC1 or NPC2 genes whose functions are not clearly established (Ikonen and Holtta-Vuori, 2004; Vanier, 2010) and Mucopolipidosis type IV, which is caused by mutations in the cation channel mucopolipin-1 (Wakabayashi *et al.*, 2011).

Composition and Fate of Intraluminal Membranes

Several lines of evidence show that intraluminal membranes of late endosomes exhibit differences in their protein/lipid composition, fate and function. Typically, late endosomes of mammalian cells contain high amounts (approx. 15Mol%) of the unconventional phospholipid lysobisphosphatidic acid (LBPA) or bis(monoacylglycerol)phosphate (BMP) (Kobayashi *et al.*, 1998) – the lipid has not been detected in yeast. LBPA is abundant in ILVs of late endosomes after immunogold labeling of cryosections, and is not detected elsewhere in the cell, including early endosomes and ECV/MVBs (Kobayashi *et al.*, 1998; Bissig and Gruenberg, 2013). By contrast, PtdIns(3)P is generated on early endosomes, as we have seen above, where it controls the onset of the ESCRT pathway. Electron microscopy studies showed that PtdIns(3)P is also present in ILVs of ECV/MVBs and to a lesser extent late endosomes – where it does not seem to colocalize with LBPA (Gillooly *et al.*, 2000). LBPA and PtdIns(3)P thus appear to show mirror image-like distribution along membranes of the endocytic pathway.

LBPA is primarily present as the 2,2'-dioleoyl isoform, which is biologically active (Kobayashi *et al.*, 1998; Matsuo *et al.*, 2004; Chevallier *et al.*, 2008), but thermodynamically unstable as the fatty acid chains can migrate to form 3,3'-LBPA

(Chevallier *et al.*, 2000; Kobayashi *et al.*, 2002). The position of LBPA fatty acyl chains likely modulates its biological functions, by dramatically changing the overall structure (Goursot *et al.*, 2010). The pathway of LBPA synthesis and degradation is not clearly established (Bissig and Gruenberg, 2013), but it is a poor substrate for lipases and phospholipases presumably because it is stereoisomeric to other naturally occurring phospholipids (Brotherus *et al.*, 1974; Tan *et al.*, 2012). It has been shown that LBPA has the intrinsic capacity to deform membranes *in vitro* (Matsuo *et al.*, 2004), a situation which presumably reflects the role of LBPA during ILV formation within late endosomes *in vivo*. In addition, LBPA interacts directly with the ESCRT-associated protein ALIX (Bissig *et al.*, 2013), consistent with the notion that both LBPA and ALIX play a direct role in ESCRT-dependent ILV formation within late endosomes (Falguières *et al.*, 2008; Bissig and Gruenberg, 2014). It seems that, in addition to the ESCRT-0 subunit HRS, other adapters, including ALIX, may recruit cargo proteins into nascent ILVs presumably at sequential steps of the pathway (Bissig and Gruenberg, 2014).

Other Destinations: All Roads Do Not Lead to the Lysosomes

While signaling receptors sorted into ILVs are targeted to lysosomes for degradation, other molecules present in late endosomal intraluminal membranes are not rapidly turned over, like LBPA itself or proteins of the tetraspanin family that share four transmembrane segments and two conserved extracellular loops (Engering *et al.*, 2003). In addition, some proteins in transit through late endosomal intraluminal membranes including the Man6P receptors (Griffiths *et al.*, 1988; Kobayashi *et al.*, 1998) and the tetraspanin CD63 in endothelial cells (Kobayashi *et al.*, 2000) escape delivery to lysosomes (Braulke and Bonifacino, 2009) – again in contrast to signaling receptors. Finally, evidence shows that, in mammalian cells, some ILVs and their cargo are not destined for degradation. Microvesicles of ILV origin can be released as exosomes into the extracellular space after fusion of secretory endo-lysosomes with the plasma membrane (Raposo and Stoorvogel, 2013). While exosomes may (Denzer *et al.*, 2000) or may not (Wubbolts *et al.*, 2003) contain LBPA, ALIX itself is found in exosomes (Thery *et al.*, 2009; Mathivanan *et al.*, 2012). However, the role of ALIX in exosome biogenesis appears to be cell type dependent. In a mouse oligodendroglial cell line, exosome biogenesis was found to be ceramide-dependent and ALIX- and ESCRT-independent (Trajkovic *et al.*, 2008), while in MCF-7 cells, exosome biogenesis was promoted by syndecan-syntenin in an ALIX and ESCRT-dependent manner (Baietti *et al.*, 2012). Similarly, the formation of exosomes in human retinal pigmented epithelial 1 (RPE1) cells requires SNX3, consistent with its role in ILV formation (Pons *et al.*, 2008), as well as ALIX and the ESCRT-I subunit TSG101 (Abrami *et al.*, 2013).

In addition to ILV formation and exosome secretion, both ALIX and LBPA were also shown to be involved in the penetration of some pathogenic agents into the host cell cytoplasm, including anthrax toxin lethal factor (Abrami *et al.*, 2004, 2013), vesicular stomatitis virus (VSV) (Bissig *et al.*, 2013; Le Blanc

et al., 2005; Luyet *et al.*, 2008), Old World arenaviruses (Pasqual *et al.*, 2011). Although the mechanism remains to be elucidated (Bissig and Gruenberg, 2014), it has been proposed that anthrax toxin lethal factor and VSV RNA are released into the ILV lumen in order to reach late endosomes without encountering a degradative environment (Gruenberg and van der Goot, 2006). There, ILV back-fusion with the limiting membrane would ensure toxin or RNA release into the cytoplasm in a process controlled by LBPA, ALIX and ESCRTs. Proteins and lipids that transit through late endosomes, including low density lipoprotein (LDL)-derived cholesterol (Kobayashi *et al.*, 1999; Chevallier *et al.*, 2008) may utilize this ILV back-fusion as an escape route from lysosomes (Falguières *et al.*, 2012). It should be noted that ILV back-fusion has only been observed in LBPA-containing late endosomes, and consistent with this notion, there is no evidence for the existence of a similar escape pathway in yeast cells, which may not contain LBPA.

The Late Endosome: A Signaling Platform

Evidence is accumulating that late endosomes and lysosomes play an essential role as sensing and signaling organelles. In mammalian cells, late endosomes regulate the intracellular trafficking of LDL-derived cholesterol – LDL particles being the main source of cholesterol in the cells of most mammalian tissues. They are endocytosed by the LDL receptor, and, after uncoupling from the receptor, are transported to late endosomes, where hydrolysis of the ester by the lysosomal acidic lipase, releases free cholesterol (Goldstein *et al.*, 1985). Cholesterol then needs to reach its intracellular destinations, in particular the endoplasmic reticulum (ER) where the SREBP-dependent feedback regulation of cholesterol synthesis and uptake takes place (Brown and Goldstein, 1997).

Once de-esterified in the late endosome lumen, free cholesterol rapidly partitions into nearby membranes, presumably containing LBPA, due to its low partition coefficient. From there, it is exported to its cellular destinations via an export mechanism that remains to be elucidated. Mutations in the NPC1 and NPC2 proteins cause the cholesterol storage disorder Niemann-Pick C (Loftus *et al.*, 1997; Naureckiene *et al.*, 2000), but the precise functions of these proteins is being debated (Kwon *et al.*, 2009; Ikonen and Holttä-Vuori, 2004; Lloyd-Evans *et al.*, 2008). The disease can be phenocopied at the cellular level by conditions that interfere with the functions of LBPA (Kobayashi *et al.*, 1999). In fact, both cholesterol export from late endosomes and viral capsid release share the same LBPA-dependence and both processes also depend on ALIX, suggesting that cholesterol is exported to the endosome limiting membrane via ILV back-fusion (Bissig and Gruenberg, 2014), as discussed.

Experimental evidence supports the existence of ER-late endosome contact sites, which may be involved in signaling and regulatory functions. It has been proposed that ER-resident tyrosine phosphatase PTB1B contributes to dephosphorylate activated EGF receptor present in endosomes (Haj *et al.*, 2002; Eden *et al.*, 2010). More directly, the ER-resident protein oxysterol-binding protein ORP5 was shown to interact physically with NPC1, suggesting that this association may mediate cholesterol exchange (Du *et al.*, 2011). Another oxysterol-

binding protein ORP1L and its partner, the ER protein Vap-A, were proposed to regulate endosomal motility via a tripartite complex with RAB7 and its effector RILP, driving minus-end directed motility of endosomes mediated by the dynein/dynactin motor (Rocha *et al.*, 2009; Vihervaara *et al.*, 2011). These authors recently demonstrated a role for this complex in late endosomal tethering and fusion via interaction with the HOPS complex (van der Kant *et al.*, 2013). Interestingly, the same Vap-A protein has been shown to promote ER-LE contact via interaction with the proteins STARD3/MLN64 and STARD3NL/Mentho (Alpy *et al.*, 2013) – two proteins involved in cholesterol transport to mitochondria (Zhang *et al.*, 2002; Charman *et al.*, 2010). Recent studies, based on the cholesterol-modulating adenoviral protein RID α , which interacts with RAB7 and ORP1L, propose that ORP1L could indeed be essential for the transport of LDL-derived cholesterol to its regulatory pool in the ER (Cianciola *et al.*, 2013).

In addition to its central role in cholesterol transport, late endocytic compartments play a fundamental and evolutionary conserved role in nutrient sensing and the consequent adaptation of cell growth/differentiation and metabolism (Settembre *et al.*, 2013b). The master regulator is the Ser/Thr kinase mammalian target of rapamycin (mTOR), assembled in the multi-protein complex mTOR complex 1 (mTORC1) which integrates nutrient availability with other environmental cues, such as growth factors signals and cellular stress, at the surface of late endosomes (Jewell *et al.*, 2013). A series of signaling events converge to regulate mTORC1 (Jewell *et al.*, 2013; Dibble and Manning, 2013) but a crucial event – and prerequisite for activation – is mTORC1 translocation to late endosome/lysosomal membrane promoted by nutrient availability. Translocation is regulated by the RAG GTPases and its effector complex regulator and is promoted by the presence of both free amino acids (Sancak *et al.*, 2010) and glucose (Efe-yan *et al.*, 2013). Interestingly, the TORC1 complex in yeast stably resides in the vacuole, but the mechanism of regulation by GTPases in nutrient-dependent manner is otherwise very similar (Binda *et al.*, 2009; Kogan *et al.*, 2010). How amino acids in the endosome lumen are sensed by the Rag/regulator complex is still not perfectly clear. The V-ATPase is clearly essential for activation (Zoncu *et al.*, 2011; Bar-Peled *et al.*, 2012) but it is not established whether it directly transmits the signal, independently from its proton pump activity (Zoncu *et al.*, 2011) or whether the V-ATPase indirectly promotes the activity of the proton-dependent amino acid exporter PAT1 (Heublein *et al.*, 2010; Ogmundsdottir *et al.*, 2012). In addition to its role in the regulation of cellular metabolism (Duvel *et al.*, 2010), mTORC1 also plays a direct role in lipid metabolism through modulation of SREBP both at the transcriptional and post-translational level (Porstmann *et al.*, 2008; Peterson *et al.*, 2011; Owen *et al.*, 2012), representing another direct link between the endosomal system and the regulation of lipid metabolism. In addition, mTORC1 also regulates the nuclear translocation and activation of TFEB (Martina *et al.*, 2012; Roczniak-Ferguson *et al.*, 2012; Settembre *et al.*, 2012), the master regulator of lysosome biogenesis (Sardiello *et al.*, 2009), which also controls autophagy (Settembre *et al.*, 2011) and lipid homeostasis (Settembre *et al.*, 2013a). Conversely, mTORC1 also appears to regulate lysosome reformation after autophagy (Yu *et al.*, 2010).

See also: Organelles: Structure and Function: The Endoplasmic Reticulum

References

- Abrami, L., Brandi, L., Moayeri, M., *et al.*, 2013. Hijacking multivesicular bodies enables long-term and exosome-mediated long-distance action of anthrax toxin. *Cell Reports* 5, 986–996.
- Abrami, L., Lindsay, M., Parton, R.G., Leppla, S.H., van der Goot, F.G., 2004. Membrane insertion of anthrax protective antigen and cytoplasmic delivery of lethal factor occur at different stages of the endocytic pathway. *Journal of Cell Biology* 166, 645–651.
- Alpy, F., Rousseau, A., Schwab, Y., *et al.*, 2013. STARD3 or STARD3NL and VAP form a novel molecular tether between late endosomes and the ER. *Journal of Cell Science* 126, 5500–5512.
- Andrews, N.W., 2000. Regulated secretion of conventional lysosomes. *Trends in Cell Biology* 10, 316–321.
- Aniento, F., Emans, N., Griffiths, G., Gruenberg, J., 1993. Cytoplasmic dynein-dependent vesicular transport from early to late endosomes. *Journal of Cell Biology* 123, 1373–1387.
- Anitei, M., Hoflack, B., 2011. Exit from the trans-Golgi network: From molecules to mechanisms. *Current Opinion in Cell Biology* 23, 443–451.
- Asaoka, R., Uemura, T., Ito, J., *et al.*, 2012. *Arabidopsis* RABA1 GTPases are involved in transport between the trans-Golgi network and the plasma membrane, and are required for salinity stress tolerance. *Plant Journal* 73, 240–249.
- Baietti, M.F., Zhang, Z., Mortier, E., *et al.*, 2012. Syndecan-syntenin-ALIX regulates the biogenesis of exosomes. *Nature Cell Biology* 14, 677–685.
- Bainton, D.F., 1981. The discovery of lysosomes. *Journal of Cell Biology* 91, 66s–76s.
- Bar-Peled, L., Schweitzer, L.D., Zoncu, R., Sabatini, D.M., 2012. Ragulator is a GEF for the rag GTPases that signal amino acid levels to mTORC1. *Cell* 150, 1196–1208.
- Barbero, P., Bittova, L., Pfeffer, S.R., 2002. Visualization of Rab9-mediated vesicle transport from endosomes to the trans-Golgi in living cells. *Journal of Cell Biology* 156, 511–518.
- Binda, M., Peli-Gullii, M.P., Bonfils, G., *et al.*, 2009. The Vam6 GEF controls TORC1 by activating the EGO complex. *Molecular Cell* 35, 563–573.
- Bissig, C., Gruenberg, J., 2013. Lipid sorting and multivesicular endosome biogenesis. *Cold Spring Harbor Perspectives in Biology*. doi:10.1101/cshperspect.a016816.
- Bissig, C., Gruenberg, J., 2014. ALIX and the multivesicular endosome: ALIX in Wonderland. *Trends in Cell Biology* 24, 19–25.
- Bissig, C., Lenoir, M., Velluz, M.C., *et al.*, 2013. Viral infection controlled by a calcium-dependent lipid-binding module in ALIX. *Developmental Cell* 25, 364–373.
- Bohdanowicz, M., Grinstein, S., 2013. Role of phospholipids in endocytosis, phagocytosis, and macropinocytosis. *Physiological Reviews* 93, 69–106.
- Bomse, M., Parton, R., Kuznetsov, S.A., Schroer, T.A., Gruenberg, J., 1990. Microtubule- and motor-dependent fusion *in vitro* between apical and basolateral endocytic vesicles from MDCK cells. *Cell* 62, 719–731.
- Bomse, M., Prydz, K., Parton, R.G., Gruenberg, J., Simons, K., 1989. Endocytosis in filter-grown Madin-Darby canine kidney cells. *Journal of Cell Biology* 109, 3243–3258.
- Bonifacio, J.S., Hurley, J.H., 2008. Retromer. *Current Opinion in Cell Biology* 20, 427–436.
- Brankatschk, B., Pons, V., Parton, R.G., Gruenberg, J., 2011. Role of SNX16 in the dynamics of tubulo-cisternal membrane domains of late endosomes. *PLoS One* 6 (7), e21771. doi:10.1371/journal.pone.0021771.
- Braulik, T., Bonifacio, J.S., 2009. Sorting of lysosomal proteins. *Biochimica et Biophysica Acta* 1793, 605–614.
- Brotherus, J., Renkonen, O., 1977. Subcellular distributions of lipids in cultured BHK cells: Evidence for the enrichment of lysobisphosphatidic acid and neutral lipids in lysosomes. *Journal of Lipid Research* 18, 191–202.
- Brotherus, J., Renkonen, O., Herrmann, J., Fisher, W., 1974. Novel stereochemical configuration in lysobisphosphatidic acid of cultured BHK cells. *Chemistry and Physics of Lipids* 13, 178–182.
- Brown, M.S., Goldstein, J.L., 1997. The SREBP pathway: Regulation of cholesterol metabolism by proteolysis of a membrane-bound transcription factor. *Cell* 89, 331–340.
- Charman, M., Kennedy, B.E., Osborne, N., Karten, B., 2010. MLN64 mediates egress of cholesterol from endosomes to mitochondria in the absence of functional Niemann-Pick Type C1 protein. *Journal of Lipid Research* 51, 1023–1034.
- Chavrier, P., Parton, R.G., Hauri, H.P., Simons, K., Zerial, M., 1990. Localisation of low molecular weight GTP binding proteins to exocytic and endocytic compartments. *Cell* 62, 317–329.
- Chen, S.H., Chen, S., Tokarev, A.A., *et al.*, 2005. Ypt31/32 GTPases and their novel F-box effector protein Rcy1 regulate protein recycling. *Molecular Biology of the Cell* 16, 178–192.
- Chevallier, J., Chamoun, Z., Jiang, G., *et al.*, 2008. Lysobisphosphatidic acid controls endosomal cholesterol levels. *Journal of Biological Chemistry* 283, 27871–27880.
- Chevallier, J., Sakai, N., Robert, F., *et al.*, 2000. Rapid access to synthetic lysobisphosphatidic acids using P(III) chemistry. *Organic Letters* 2, 1859–1861.
- Chirivino, D., del maestro, L., Formstecher, E., *et al.*, 2011. The ERM proteins interact with the HOPS complex to regulate the maturation of endosomes. *Molecular Biology of the Cell* 22, 375–385.
- Choi, A.M., Ryter, S.W., Levine, B., 2013. Autophagy in human health and disease. *New England Journal of Medicine* 368, 1845–1846.
- Cianciola, N.L., Greene, D.J., Morton, R.E., Carlin, C.R., 2013. Adenovirus RIDalpha uncovers a novel pathway requiring ORP1L for lipid droplet formation independent of NPC1. *Molecular Biology of the Cell* 24, 3309–3325.
- Contento, A.L., Bassham, D.C., 2012. Structure and function of endosomes in plant cells. *Journal of Cell Science* 125, 3511–3518.
- Deng, Y.P., Storrie, B., 1988. Animal cell lysosomes rapidly exchange membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America* 85, 3860–3864.
- Denzer, K., van Eijk, M., Kleijmeer, M.J., *et al.*, 2000. Follicular dendritic cells carry MHC class II-expressing microvesicles at their surface. *Journal of Immunology* 165, 1259–1265.
- Dibble, C.C., Manning, B.D., 2013. Signal integration by mTORC1 coordinates nutrient input with biosynthetic output. *Nature Cell Biology* 15, 555–564.
- Diment, S., Stahl, P., 1985. Macrophage endosomes contain proteases which degrade endocytosed protein ligands. *Journal of Biological Chemistry* 260, 15311–15317.
- Du, X., Kumar, J., Ferguson, C., *et al.*, 2011. A role for oxysterol-binding protein-related protein 5 in endosomal cholesterol trafficking. *Journal of Cell Biology* 192, 121–135.
- de Duve, C., 2005. The lysosome turns fifty. *Nature Cell Biology* 7, 847–849.
- Duvel, K., Yecies, J.L., Menon, S., *et al.*, 2010. Activation of a metabolic gene regulatory network downstream of mTOR complex 1. *Molecular Cell* 39, 171–183.
- Eden, E.R., White, I.J., Tsapara, A., Futter, C.E., 2010. Membrane contacts between endosomes and ER provide sites for PTP1B-epidermal growth factor receptor interaction. *Nature Cell Biology* 12, 267–272.
- Efeyan, A., Zoncu, R., Chang, S., *et al.*, 2013. Regulation of mTORC1 by the Rag GTPases is necessary for neonatal autophagy and survival. *Nature* 493, 679–683.
- Engering, A., Kuhn, L., Fluittsma, D., Hoefsmit, E., Pieters, J., 2003. Differential post-translational modification of CD63 molecules during maturation of human dendritic cells. *European Journal of Biochemistry/FEBS* 270, 2412–2420.
- Falguieres, T., Castle, J.D., Gruenberg, J., 2012. Regulation of the MVB pathway by SCAMP3. *Traffic* 13, 131–142.
- Falguieres, T., Luyet, P.P., Bissig, C., *et al.*, 2008. *In vitro* budding of intraluminal vesicles into late endosomes is regulated by Alix and Tsg101. *Molecular Biology of the Cell* 19, 4942–4955.
- Foret, L., Dawson, J.E., Villaseñor, R., *et al.*, 2012. A general theoretical framework to infer endosomal network dynamics from quantitative image analysis. *Current Biology* 22, 1381–1390.
- Futerman, A.H., van Meer, G., 2004. The cell biology of lysosomal storage disorders. *Nature Reviews Molecular Cell Biology* 5, 554–565.
- Galvez, T., Gilleron, J., Zerial, M., O'sullivan, G.A., 2012. SnapShot: Mammalian Rab proteins in endocytic trafficking. *Cell* 151, 234–234.e2.
- Ganley, I.G., Carroll, K., Bittova, L., Pfeffer, S., 2004. Rab9 GTPase regulates late endosome size and requires effector interaction for its stability. *Molecular Biology of the Cell* 15, 5420–5430.
- Ganley, I.G., Espinosa, E., Pfeffer, S.R., 2008. A syntaxin 10-SNARE complex distinguishes two distinct transport routes from endosomes to the trans-Golgi in human cells. *Journal of Cell Biology* 180, 159–172.
- Geuze, H.J., Slot, J.W., Strous, G.J., *et al.*, 1984. Intracellular receptor sorting during endocytosis: Comparative immunoelectron microscopy of multiple receptors in rat liver. *Cell* 37, 195–204.
- Geuze, H.J., Slot, J.W., Strous, G.J., Lodish, H.F., Schwartz, A.L., 1983. Intracellular site of asialoglycoprotein receptor-ligand uncoupling: Double-label immunoelectron microscopy during receptor-mediated endocytosis. *Cell* 32, 277–287.

- Gillooly, D.J., Morrow, I.C., Lindsay, M., *et al.*, 2000. Localization of phosphatidylinositol 3-phosphate in yeast and mammalian cells. *EMBO Journal* 19, 4577–4588.
- Goldstein, J.L., Brown, M.S., Anderson, R.G., Russell, D.W., Schneider, W.J., 1985. Receptor-mediated endocytosis: Concepts emerging from the LDL receptor system. *Annual Review of Cell Biology* 1, 1–39.
- Gourso, A., Mineva, T., Bissig, C., Gruenberg, J., Salahub, D.R., 2010. Structure, dynamics, and energetics of lysobisphosphatidic acid (LBPA) isomers. *Journal of Physical Chemistry B* 114, 15712–15720.
- Griffiths, G., Back, R., Marsh, M., 1989. A quantitative analysis of the endocytic pathway in baby hamster kidney cells. *Journal of Cell Biology* 109, 2703–2720.
- Griffiths, G., Hoflack, B., Simons, K., Mellman, I., Kornfeld, S., 1988. The mannose 6-phosphate receptor and the biogenesis of lysosomes. *Cell* 52, 329–341.
- Griffiths, G., Matteoni, R., Back, R., Hoflack, B., 1990. Characterization of the cation-independent mannose 6-phosphate receptor-enriched prelysosomal compartment in NRK cells. *Journal of Cell Science* 95 (Pt 3), 441–461.
- Gruenberg, J., 2001. The endocytic pathway: A mosaic of domains. *Nature Reviews Molecular Cell Biology* 2, 721–730.
- Gruenberg, J., Griffiths, G., Howell, K.E., 1989. Characterization of the early endosome and putative endocytic carrier vesicles *in vivo* and with an assay of vesicle fusion *in vitro*. *Journal of Cell Biology* 108, 1301–1316.
- Gruenberg, J., Howell, K.E., 1989. Membrane traffic in endocytosis: Insights from cell-free assays. *Annual Review of Cell Biology* 5, 453–481.
- Gruenberg, J., van der Goot, F.G., 2006. Mechanisms of pathogen entry through the endosomal compartments. *Nature Reviews Molecular Cell Biology* 7, 495–504.
- Haj, F.G., Verveer, P.J., Squire, A., Neel, B.G., Bastiaens, P.J., 2002. Imaging sites of receptor dephosphorylation by PTP1B on the surface of the endoplasmic reticulum. *Science* 295, 1708–1711.
- Helenius, A., Mellman, I., Wall, D., Hubbard, A., 1983. Endosomes. *Trends in Biochemical Sciences* 8, 245–250.
- Henne, W.M., Buchkovich, N.J., EMR, S.D., 2011. The ESCRT pathway. *Developmental Cell* 21, 77–91.
- Heublein, S., Kazi, S., Ogmundsdottir, M.H., *et al.*, 2010. Proton-assisted amino acid transporters are conserved regulators of proliferation and amino-acid-dependent mTORC1 activation. *Oncogene* 29, 4068–4079.
- Heuser, J., 1989. Changes in lysosome shape and distribution correlated with changes in cytoplasmic pH. *Journal of Cell Biology* 108, 855–864.
- Howes, M.T., Kirkham, M., Riches, J., *et al.*, 2010. Clathrin-independent carriers form a high capacity endocytic sorting system at the leading edge of migrating cells. *Journal of Cell Biology* 190, 675–691.
- Hsu, V.W., Prekeris, R., 2010. Transport at the recycling endosome. *Current Opinion in Cell Biology* 22, 528–534.
- Huotari, J., Helenius, A., 2011. Endosome maturation. *EMBO Journal* 30, 3481–3500.
- Hurley, J.H., Hanson, P.I., 2010. Membrane budding and scission by the ESCRT machinery: It's all in the neck. *Nature Reviews Molecular Cell Biology* 11, 556–566.
- Hurley, J.H., Stenmark, H., 2011. Molecular mechanisms of ubiquitin-dependent membrane traffic. *Annual Review of Biophysics* 40, 119–142.
- Ikonen, E., Holtta-Vuori, M., 2004. Cellular pathology of Niemann-Pick type C disease. *Seminars in Cell & Developmental Biology* 15, 445–454.
- Jahraus, A., Storrie, B., Griffiths, G., Desjardins, M., 1994. Evidence for retrograde traffic between terminal lysosomes and the prelysosomal/late endosome compartment. *Journal of Cell Science* 107 (Pt 1), 145–157.
- Jewell, J.L., Russell, R.C., Guan, K.L., 2013. Amino acid signalling upstream of mTOR. *Nature Reviews Molecular Cell Biology* 14, 133–139.
- Johannes, L., Popoff, V., 2008. Tracing the retrograde route in protein trafficking. *Cell* 135, 1175–1187.
- van der Kant, R., Fish, A., Janssen, L., *et al.*, 2013. Late endosomal transport and tethering are coupled processes controlled by RILP and the cholesterol sensor ORP1L. *Journal of Cell Science* 126, 3462–3474.
- Kinchen, J.M., Ravichandran, K.S., 2010. Identification of two evolutionarily conserved genes regulating processing of engulfed apoptotic cells. *Nature* 464, 778–782.
- Ko, D.C., Gordon, M.D., Jin, J.Y., Scott, M.P., 2001. Dynamic movements of organelles containing Niemann-Pick C1 protein: NPC1 involvement in late endocytic events. *Molecular Biology of the Cell* 12, 601–614.
- Kobayashi, T., Beuchat, M.H., Chevallier, J., *et al.*, 2002. Separation and characterization of late endosomal membrane domains. *Journal of Biological Chemistry* 277, 32157–32164.
- Kobayashi, T., Beuchat, M.H., Lindsay, M., *et al.*, 1999. Late endosomal membranes rich in lysobisphosphatidic acid regulate cholesterol transport. *Nature Cell Biology* 1, 113–118.
- Kobayashi, T., Stang, E., Fang, K.S., *et al.*, 1998. A lipid associated with the antiphospholipid syndrome regulates endosome structure and function. *Nature* 392, 193–197.
- Kobayashi, T., Vischer, U.M., Rosnoblet, C., *et al.*, 2000. The tetraspanin CD63/lamp3 cycles between endocytic and secretory compartments in human endothelial cells. *Molecular Biology of the Cell* 11, 1829–1843.
- Kogan, K., Spear, E.D., Kaiser, C.A., Fass, D., 2010. Structural conservation of components in the amino acid sensing branch of the TOR pathway in yeast and mammals. *Journal of Molecular Biology* 402, 388–398.
- Kwon, H.J., Abi-Mosleh, L., Wang, M.L., *et al.*, 2009. Structure of N-terminal domain of NPC1 reveals distinct subdomains for binding and transfer of cholesterol. *Cell* 137, 1213–1224.
- LaPlante, J.M., Sun, M., Falardeau, J., *et al.*, 2006. Lysosomal exocytosis is impaired in mucopolidiosis type IV. *Molecular Genetics and Metabolism* 89, 339–348.
- Laulagnier, K., Schieber, N.L., Maritzen, T., *et al.*, 2011. Role of AP1 and Gakkin in the traffic of secretory endo-lysosomes. *Molecular Biology of the Cell* 22, 2068–2082.
- Le Blanc, I., Luyet, P.P., Pons, V., *et al.*, 2005. Endosome-to-cytosol transport of viral nucleocapsids. *Nature Cell Biology* 7, 653–664.
- Lebrand, C., Corti, M., Goodson, H., *et al.*, 2002. Late endosome motility depends on lipids via the small GTPase Rab7. *EMBO Journal* 21, 1289–1300.
- Lloyd-Evans, E., Morgan, A.J., He, X., *et al.*, 2008. Niemann-Pick disease type C1 is a sphingosine storage disease that causes deregulation of lysosomal calcium. *Nature Medicine* 14, 1247–1255.
- Loftus, S.K., Morris, J.A., Carstea, E.D., *et al.*, 1997. Murine model of Niemann-Pick C disease: Mutation in a cholesterol homeostasis gene [see comments]. *Science* 277, 232–235.
- Luyet, P.P., Falguieres, T., Pons, V., Pattnaik, A.K., Gruenberg, J., 2008. The ESCRT-I subunit Tsg101 controls endosome-to-cytosol release of viral RNA. *Traffic* 9, 2279–2290.
- Luzio, J.P., Pryor, P.R., Bright, N.A., 2007. Lysosomes: Fusion and function. *Nature Reviews Molecular Cell Biology* 8, 622–632.
- Marks, M.S., Heijnen, H.F., Raposo, G., 2013. Lysosome-related organelles: Unusual compartments become mainstream. *Current Opinion in Cell Biology* 25, 495–505.
- Marsh, M., Griffiths, G., Dean, G.E., Mellman, I., Helenius, A., 1986. Three-dimensional structure of endosomes in BHK-21 cells. *Proceedings of the National Academy of Sciences of the United States of America* 83, 2899–2903.
- Marshansky, V., Futai, M., 2008. The V-type H(+) -ATPase in vesicular trafficking: Targeting, regulation and function. *Current Opinion in Cell Biology* 20, 415–426.
- Martina, J.A., Chen, Y., Gucek, M., Puertollano, R., 2012. mTORC1 functions as a transcriptional regulator of autophagy by preventing nuclear transport of TFEB. *Autophagy* 8, 903–914.
- Mathivanan, S., Fahner, C.J., Reid, G.E., Simpson, R.J., 2012. ExoCarta 2012: Database of exosomal proteins, RNA and lipids. *Nucleic Acids Research* 40, D1241–D1244.
- Matsuo, H., Chevallier, J., Mayran, N., *et al.*, 2004. Role of LBPA and Alix in multivesicular liposome formation and endosome organization. *Science* 303, 531–534.
- Mayran, N., Parton, R.G., Gruenberg, J., 2003. Annexin II regulates multivesicular endosome biogenesis in the degradation pathway of animal cells. *EMBO Journal* 22, 3242–3253.
- McGourty, K., Thurston, T.L., Matthews, S.A., *et al.*, 2012. Salmonella inhibits retrograde trafficking of mannose-6-phosphate receptors and lysosome function. *Science* 338, 963–967.
- Morel, E., Parton, R.G., Gruenberg, J., 2009. Annexin A2-dependent polymerization of actin mediates endosome biogenesis. *Developmental Cell* 16, 445–457.
- Mrakovic, A., Kay, J.G., Furuya, W., Brumell, J.H., Botelho, R.J., 2012. Rab7 and Arl8 GTPases are necessary for lysosome tubulation in macrophages. *Traffic* 13, 1667–1679.
- Mullock, B.M., Branch, W.J., van Schaik, M., Gilbert, L.K., Luzio, J.P., 1989. Reconstitution of an endosome-lysosome interaction in a cell-free system. *Journal of Cell Biology* 108, 2093–2099.
- Naureckiene, S., Sleat, D.E., Lackland, H., *et al.*, 2000. Identification of HE1 as the second gene of niemann-pick C disease. *Science* 290, 2298–2301.
- van Niel, G., Charrin, S., Simoes, S., *et al.*, 2011. The tetraspanin CD63 regulates ESCRT-independent and -dependent endosomal sorting during melanogenesis. *Developmental Cell* 21, 708–721.
- Odorizzi, G., Babst, M., Emr, S.D., 1998. Fab1p PtdIns(3)P 5-kinase function essential for protein sorting in the multivesicular body. *Cell* 95, 847–858.
- Ogmundsdottir, M.H., Heublein, S., Kazi, S., *et al.*, 2012. Proton-assisted amino acid transporter PAT1 complexes with Rag GTPases and activates TORC1 on late endosomal and lysosomal membranes. *PLoS One* 7, e36616.

- Owen, J.L., Zhang, Y., Bae, S.H., *et al.*, 2012. Insulin stimulation of SREBP-1c processing in transgenic rat hepatocytes requires p70 S6-kinase. *Proceedings of the National Academy of Sciences of the United States of America* 109, 16184–16189.
- Pasqual, G., Rojek, J.M., Masin, M., Chatton, J.Y., Kunz, S., 2011. Old world arenaviruses enter the host cell via the multivesicular body and depend on the endosomal sorting complex required for transport. *PLoS Pathogens* 7, e1002232.
- Pereira-Leal, J.B., Seabra, M.C., 2001. Evolution of the Rab family of small GTP-binding proteins. *Journal of Molecular Biology* 313, 889–901.
- Peterson, T.R., Sengupta, S.S., Harris, T.E., *et al.*, 2011. mTOR complex 1 regulates lipin 1 localization to control the SREBP pathway. *Cell* 146, 408–420.
- Pfeffer, S.R., 2011. Entry at the trans-face of the Golgi. *Cold Spring Harbor Perspectives in Biology* 3. doi:10.1101/cshperspect.a005272.
- Phaire-Washington, L., Silverstein, S.C., Wang, E., 1980. Phorbol myristate acetate stimulates microtubule and 10-nm filament extension and lysosome redistribution in mouse macrophages. *The Journal of cell biology* 86, 641–655.
- Pons, V., Luyet, P.-P., Morel, E., *et al.*, 2008. Hrs and SNX3 functions in sorting and membrane invagination within multivesicular bodies. *PLoS Biology* 6, e214. doi:10.1371/journal.pbio.0060214.
- Pons, V., Ustunel, C., Rolland, C., *et al.*, 2011. SNX12 role in endosome membrane transport. *PLoS One* 7, e38949. doi:10.1371/journal.pone.0038949.
- Portsmann, T., Santos, C.R., Griffiths, B., *et al.*, 2008. SREBP activity is regulated by mTORC1 and contributes to Akt-dependent cell growth. *Cell Metabolism* 8, 224–236.
- Poteryaev, D., Datta, S., Ackema, K., Zerial, M., Spang, A., 2010. Identification of the switch in early-to-late endosome transition. *Cell* 141, 497–508.
- Prescianotto-Baschong, C., Riezman, H., 2002. Ordering of compartments in the yeast endocytic pathway. *Traffic* 3, 37–49.
- Rabinowitz, S., Horstmann, H., Gordon, S., Griffiths, G., 1992. Immunocytochemical characterization of the endocytic and phagolysosomal compartments in peritoneal macrophages. *Journal of Cell Biology* 116, 95–112.
- Rao, S.K., Huynh, C., Proux-Gillardeaux, V., Galli, T., Andrews, N.W., 2004. Identification of SNAREs involved in synaptotagmin VII-regulated lysosomal exocytosis. *Journal of Biological Chemistry* 279, 20471–20479.
- Raposo, G., Stoorvogel, W., 2013. Extracellular vesicles: Exosomes, microvesicles, and friends. *Journal of Cell Biology* 200, 373–383.
- Raymond, C.K., Howald-Stevenson, I., Vater, C.A., Stevens, T.H., 1992. Morphological classification of the yeast vacuolar protein sorting mutants: Evidence for a prevacuolar compartment in class E vps mutants. *Molecular Biology of the Cell* 3, 1389–1402.
- Razi, M., Futter, C.E., 2006. Distinct roles for Tsg101 and Hrs in multivesicular body formation and inward vesiculation. *Molecular Biology of the Cell* 17, 3469–3483.
- Riezman, H., 1985. Endocytosis in yeast: Several of the yeast secretory mutants are defective in endocytosis. *Cell* 40, 1001–1009.
- Rink, J., Ghigo, E., Kalaidzidis, Y., Zerial, M., 2005. Rab conversion as a mechanism of progression from early to late endosomes. *Cell* 122, 735–749.
- Roberts, R.L., Barbieri, M.A., Pryse, K.M., *et al.*, 1999. Endosome fusion in living cells overexpressing GFP-rab5. *Journal of Cell Science* 112 (Pt 21), 3667–3675.
- Rocha, N., Kuijl, C., van der Kant, R., *et al.*, 2009. Cholesterol sensor ORP1L contacts the ER protein VAP to control Rab7-RILP-p150 Glued and late endosome positioning. *Journal of Cell Biology* 185, 1209–1225.
- Roczniak-Ferguson, A., Petit, C.S., Froehlich, F., *et al.*, 2012. The transcription factor TFEB links mTORC1 signaling to transcriptional control of lysosome homeostasis. *Science Signaling* 5, ra42.
- Rodriguez, A., Webster, P., Ortego, J., Andrews, N.W., 1997. Lysosomes behave as Ca²⁺-regulated exocytic vesicles in fibroblasts and epithelial cells. *Journal of Cell Biology* 137, 93–104.
- Rojas, R., van Vlijmen, T., Mardones, G.A., *et al.*, 2008. Regulation of retromer recruitment to endosomes by sequential action of Rab5 and Rab7. *Journal of Cell Biology* 183, 513–526.
- Rusten, T.E., Vaccari, T., Stenmark, H., 2012. Shaping development with ESCRTs. *Nature Cell Biology* 14, 38–45.
- Saftig, P., Klumperman, J., 2009. Lysosome biogenesis and lysosomal membrane proteins: Trafficking meets function. *Nature Reviews. Molecular Cell Biology* 10, 623–635.
- Sahagian, G.G., Neufeld, E.F., 1983. Biosynthesis and turnover of the mannose 6-phosphate receptor in cultured Chinese hamster ovary cells. *Journal of Biological Chemistry* 258, 7121–7128.
- Saito, C., Ueda, T., 2009. Chapter 4: Functions of RAB and SNARE proteins in plant life. *International Review of Cell and Molecular Biology* 274, 183–233.
- Sancak, Y., Bar-Peled, L., Zoncu, R., *et al.*, 2010. Regulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. *Cell* 141, 290–303.
- Sardiello, M., Ballabio, A., 2009. Lysosomal enhancement: A CLEAR answer to cellular degradative needs. *Cell Cycle* 8, 4021–4022.
- Sardiello, M., Palmieri, M., Di Ronza, A., *et al.*, 2009. A gene network regulating lysosomal biogenesis and function. *Science* 325, 473–477.
- Scheuring, D., Viotti, C., Kruger, F., *et al.*, 2011. Multivesicular bodies mature from the trans-Golgi network/early endosome in *Arabidopsis*. *Plant Cell* 23, 3463–3481.
- Schmid, S.L., Fuchs, R., Male, P., Mellman, I., 1988. Two distinct subpopulations of endosomes involved in membrane recycling and transport to lysosomes. *Cell* 52, 73–83.
- Schulze, H., Sandhoff, K., 2011. Lysosomal lipid storage diseases. *Cold Spring Harbor Perspectives in Biology* 3. doi:10.1101/cshperspect.a004804.
- Seaman, M.N., 2012. The retromer complex - endosomal protein recycling and beyond. *Journal of Cell Science* 125, 4693–4702.
- Seaman, M.N., Gautreau, A., Billadeau, D.D., 2013. Retromer-mediated endosomal protein sorting: All WASHed up!. *Trends in Cell Biology* 23, 522–528.
- Settembre, C., de Cegli, R., Mansueto, G., *et al.*, 2013a. TFEB controls cellular lipid metabolism through a starvation-induced autoregulatory loop. *Nature Cell Biology* 15, 647–658.
- Settembre, C., Di Malta, C., Polito, V.A., *et al.*, 2011. TFEB links autophagy to lysosomal biogenesis. *Science* 332, 1429–1433.
- Settembre, C., Fraldi, A., Medina, D.L., Ballabio, A., 2013b. Signals from the lysosome: A control centre for cellular clearance and energy metabolism. *Nature reviews. Molecular Cell Biology* 14, 283–296.
- Settembre, C., Zoncu, R., Medina, D.L., *et al.*, 2012. A lysosome-to-nucleus signalling mechanism senses and regulates the lysosome via mTOR and TFEB. *EMBO Journal* 31, 1095–1108.
- Shin, H.W., Hayashi, M., Christoforidis, S., *et al.*, 2005. An enzymatic cascade of Rab5 effectors regulates phosphoinositide turnover in the endocytic pathway. *Journal of Cell Biology* 170, 607–618.
- Singer-Kruger, B., Stenmark, H., Dusterhoft, A., *et al.*, 1994. Role of three rab5-like GTPases, Ypt51p, Ypt52p, and Ypt53p, in the endocytic and vacuolar protein sorting pathways of yeast. *Journal of Cell Biology* 125, 283–298.
- Slagsvold, T., Pattni, K., Malerod, L., Stenmark, H., 2006. Endosomal and non-endosomal functions of ESCRT proteins. *Trends in Cell Biology* 16, 317–326.
- Solinger, J.A., Spang, A., 2013. Tethering complexes in the endocytic pathway: CORVET and HOPS. *FEBS Journal* 280, 2743–2757.
- Sridhar, S., Patel, B., Aphkharava, D., *et al.*, 2013. The lipid kinase PI4KIIbeta preserves lysosomal identity. *EMBO Journal* 32, 324–339.
- Steinman, R.M., Mellman, I.S., Muller, W.A., Cohn, Z.A., 1983. Endocytosis and the recycling of plasma membrane. *Journal of Cell Biology* 96, 1–27.
- Stoorvogel, W., Oorschot, V., Geuze, H.J., 1996. A novel class of clathrin-coated vesicles budding from endosomes. *Journal of Cell Biology* 132, 21–33.
- Straus, W., 1964. Occurrence of phagosomes and phago-lysosomes in different segments of the nephron in relation to the reabsorption, transport, digestion, and extrusion of intravenously injected horseradish peroxidase. *Journal of Cell Biology* 21, 295–308.
- Swanson, J., Bushnell, A., Silverstein, S.C., 1987. Tubular lysosome morphology and distribution within macrophages depend on the integrity of cytoplasmic microtubules. *Proceedings of the National Academy of Sciences of the United States of America* 84, 1921–1925.
- Tall, G.G., Hama, H., Dewald, D.B., Horazdovsky, B.F., 1999. The phosphatidylinositol 3-phosphate binding protein Vac1p interacts with a Rab GTPase and a Sec1p homologue to facilitate vesicle-mediated vacuolar protein sorting. *Molecular Biology of the Cell* 10, 1873–1889.
- Tan, H.H., Makino, A., Sudesh, K., Greimel, P., Kobayashi, T., 2012. Spectroscopic evidence for the unusual stereochemical configuration of an endosome-specific lipid. *Angewandte Chemie* 51, 533–535.
- Thery, C., Ostrowski, M., Segura, E., 2009. Membrane vesicles as conveyors of immune responses. *Nature Reviews. Immunology* 9, 581–593.
- Tjelle, T.E., Brech, A., Juvet, L.K., Griffiths, G., Berg, T., 1996. Isolation and characterization of early endosomes, late endosomes and terminal lysosomes: Their role in protein degradation. *Journal of Cell Science* 109 (Pt 12), 2905–2914.
- Tooze, J., Hollinshead, M., 1991. Tubular early endosomal networks in AtT20 and other cells. *Journal of Cell Biology* 115, 635–653.
- Trajkovic, K., Hsu, C., Chiantia, S., *et al.*, 2008. Ceramide triggers budding of exosome vesicles into multivesicular endosomes. *Science* 319, 1244–1247.
- Vanier, M.T., 2010. Niemann-Pick disease type C. *Orphanet Journal of Rare Diseases* 5, 16.

- Vida, T.A., Emr, S.D., 1995. A new vital stain for visualizing vacuolar membrane dynamics and endocytosis in yeast. *Journal of Cell Biology* 128, 779–792.
- Vihervaara, T., Uronen, R.L., Wohlfahrt, G., *et al.*, 2011. Sterol binding by OSBP-related protein 1L regulates late endosome motility and function. *Cellular and Molecular Life Sciences* 68, 537–551.
- Viotti, C., Bubeck, J., Stierhof, Y.D., *et al.*, 2010. Endocytic and secretory traffic in Arabidopsis merge in the trans-Golgi network/early endosome, an independent and highly dynamic organelle. *Plant Cell* 22, 1344–1357.
- Wakabayashi, K., Gustafson, A.M., Sidransky, E., Goldin, E., 2011. Mucopolipidosis type IV: An update. *Molecular Genetics and Metabolism* 104, 206–213.
- Wemmer, M., Azmi, I., West, M., *et al.*, 2011. Bro1 binding to Snf7 regulates ESCRT-III membrane scission activity in yeast. *Journal of Cell Biology* 192, 295–306.
- Wubbolts, R., Leckie, R.S., Veenhuizen, P.T., *et al.*, 2003. Proteomic and biochemical analyses of human B cell-derived exosomes. Potential implications for their function and multivesicular body formation. *Journal of Biological Chemistry* 278, 10963–10972.
- Yamashiro, D.J., Tycko, B., Fluss, S.R., Maxfield, F.R., 1984. Segregation of transferrin to a mildly acidic (pH 6.5) para-Golgi compartment in the recycling pathway. *Cell* 37, 789–800.
- Yu, L., McPhee, C.K., Zheng, L., *et al.*, 2010. Termination of autophagy and reformation of lysosomes regulated by mTOR. *Nature* 465, 942–946.
- Zhang, M., Dwyer, N.K., Love, D.C., *et al.*, 2001. Cessation of rapid late endosomal tubulovesicular trafficking in Niemann-Pick type C1 disease. *Proceedings of the National Academy of Sciences of the United States of America* 98, 4466–4471.
- Zhang, M., Liu, P., Dwyer, N.K., *et al.*, 2002. MLN64 mediates mobilization of lysosomal cholesterol to steroidogenic mitochondria. *Journal of Biological Chemistry* 277, 33300–33310.
- Zoncu, R., Bar-Peled, L., Efeyan, A., *et al.*, 2011. mTORC1 senses lysosomal amino acids through an inside-out mechanism that requires the vacuolar H(+) -ATPase. *Science* 334, 678–683.
- Aridor, M., Hannan, L.A., 2002. Traffic jams II: An update of diseases of intracellular transport. *Traffic* 3, 781–790.
- Collinet, C., Stoter, M., Bradshaw, C.R., *et al.*, 2010. Systems survey of endocytosis by multiparametric image analysis. *Nature* 464, 243–249.
- Contento, A.L., Bassham, D.C., 2012. Structure and function of endosomes in plant cells. *Journal of Cell Science* 125, 3511–3518.
- de Duve, C., 1991. *Blueprint for a Cell: The Nature and Origin of Life*. Burlington, NC: Neil Patterson.
- Mohr, S., Bakal, C., Perrimon, N., 2010. Genomic screening with RNAi: Results and challenges. *Annual Review of Biochemistry* 79, 37–64.
- Otegui, M.S., Spitzer, C., 2008. Endosomal functions in plants. *Traffic* 9, 1589–1598.
- Scott, C.C., Gruenberg, J., 2011. Ion flux and the function of endosomes and lysosomes: pH is just the start: The flux of ions across endosomal membranes influences endosome function not only through regulation of the luminal pH. *Bioessays* 33, 103–110.
- Snijder, B., Sacher, R., Ramo, P., *et al.*, 2012. Single-cell analysis of population context advances RNAi screening at multiple levels. *Molecular Systems Biology* 8, 579.

Relevant Websites

<http://www.heuserlab.wustl.edu/>
Heuser Lab.
<http://jenalib.fli-leibniz.de/>
Jena Library.
<https://www.cellimagelibrary.org/>
The Cell.
<http://www.signaling-gateway.org/>
UCSD Signaling Gateway.
<http://www.uniprot.org/>
UniProt.

Further Reading

Aridor, M., Hannan, L.A., 2000. Traffic jam: A compendium of human diseases that affect intracellular transport processes. *Traffic* 1, 836–851.

Signaling from Endosomes

J Alanko, H Hamidi, and J Ivaska, Turku Centre for Biotechnology, Turku, Finland

© 2016 Elsevier Inc. All rights reserved.

Endocytosis

Endocytosis is a process whereby cells uptake material from the surface by an inward fold of the plasma membrane (invagination) followed by internalization in membrane-bound vesicles called endosomes. These membrane-bound organelles are important carriers of plasma membrane components and extracellular ligands ingested by endocytosis and function to pass their cargo to lysosomes for degradation. However, not all endocytosed cargo is degraded and a major sorting step occurs in the early endosomes. Depending on the specific receptors and their ligands the endocytosed receptor can: (1) dissociate from the ligand (that will be degraded) and return to the plasma membrane, (2) remain ligand bound and become degraded, or (3) recycle with its ligand intact (e.g., transferrin receptor). Endocytosis can be a constitutive process whereby plasma membrane receptors such as integrins and transferrin receptors are internalized, sorted in early endosomes and returned back to the plasma membrane to engage new ligand molecules. Alternatively, receptor endocytosis can be actively triggered by ligand binding and this process is usually coupled to ligand-induced receptor activation and signaling. Receptor tyrosine kinases (RTKs) activated by external soluble growth factors typically undergo ligand-induced endocytosis and are subsequently degraded or recycled depending on the internalization route and the ligand.

Signaling Endosomes

The classical view of endosomes, in RTK signaling, regarded endocytosis solely as a means to terminate plasma membrane signals and to downregulate plasma membrane receptor availability through targeted receptor degradation. However, this traditional outlook is now being challenged as endocytosed receptors are frequently detected in endosomes in ligand-bound and active states suggesting persistent signaling post-endocytosis. Pioneering work from nearly two decades ago exploited a dynamin mutant that blocks all endocytosis and demonstrated that full activation of signaling pathways downstream of the epidermal growth factor receptor (EGFR) requires endocytosis (Vieira *et al.*, 1996). Today, a mounting body of evidence shows that endocytosis is linked to nearly all aspects of cellular signaling and that endosomes should be viewed as important intracellular platforms for signal transduction. In addition, the coordinated functions of the endocytic machinery and signaling pathways are critical in maintaining homeostasis in many different cell types and conditions.

Endosomes fulfill multiple functions to support and regulate cell signaling. These are depicted in Figure 1 and involve the following major regulatory mechanisms. Endocytosis functions to: (1) terminate plasma membrane signaling either via receptor degradation or via compartmentalization of the receptors away from plasma membrane-specific signaling

effectors, (2) limit bioavailability of cell surface receptors, (3) generate specific signals unique to the endosomal compartment, (4) target signaling in response to different ligands or ligand concentrations, (5) sustain signaling initiated at the plasma membrane, and (6) facilitate signal propagation over long distances that are beyond the limits of passive diffusion (Figure 1). In this article we will describe these general principles and some relevant biological examples related to these specific signaling subsets. For more detailed information we refer the reader to Table 1 (summary of endosome signaling pathways reported for different receptors) and to several in-depth and excellent reviews on this topic (Sigismund *et al.*, 2012; Benzing *et al.*, 2013; Scita and Di Fiore, 2010; Goh and Sorkin, 2013).

Different Cell Types Employ Signaling Endosomes for Distinct Purposes

The majority of our knowledge and evidence for the important biological role of signaling endosomes is derived from cell types like epithelial cells, fibroblasts, and cancer cells. In the remaining part of this article we will focus on describing examples from these cell types in more detail. However, other cell systems also rely on specific forms of endosomal signaling to facilitate their unique biological functions. In neurons signals initiated at the nerve terminals by neurotrophic receptors must be transmitted long distances to the neuronal body and nucleus. In fact the hypothesis of signaling endosomes was introduced based on the observation that nerve growth factor (NGF) induced endocytosis of its tyrosine kinase receptor (TrkA) and the resulting internalized ligand–receptor complex remained active in endosomal compartments (Grimes *et al.*, 1996). Endosomal NGF-TrkA undergo retrograde transport via microtubule and motor-protein based mechanisms to the cell body and signal to extracellular signal-regulated kinase isoforms 1/2 (ERK1/2) and phosphatidylinositol 3-kinase/protein kinase B (PI3K/Akt) pathways to promote cell survival (Howe and Mobley, 2005).

In T cell signaling endocytosis is strongly linked to T cell activation in such a way that the entire endocytic machinery is rearranged in these cells upon activation. T cell receptor (TCR) activation is triggered by the antigenic peptide displayed by the major histocompatibility complex (MHC) on antigen-presenting cells generating an immunological synapse between these cells. TCR is continuously endocytosed from the plasma membrane and returned to the cell surface via recycling endosomes (in a pathway overlapping with the transferrin receptor). This trafficking is critical for TCR accumulation in the immunological synapse (Das *et al.*, 2004) demonstrating that endocytosis contributes to T cell activation. In addition, T cells release cytoplasmic vesicular signaling components, such as Lat and Lck, to the forming immunological synapse in response to TCR activation to trigger efficient TCR signaling.

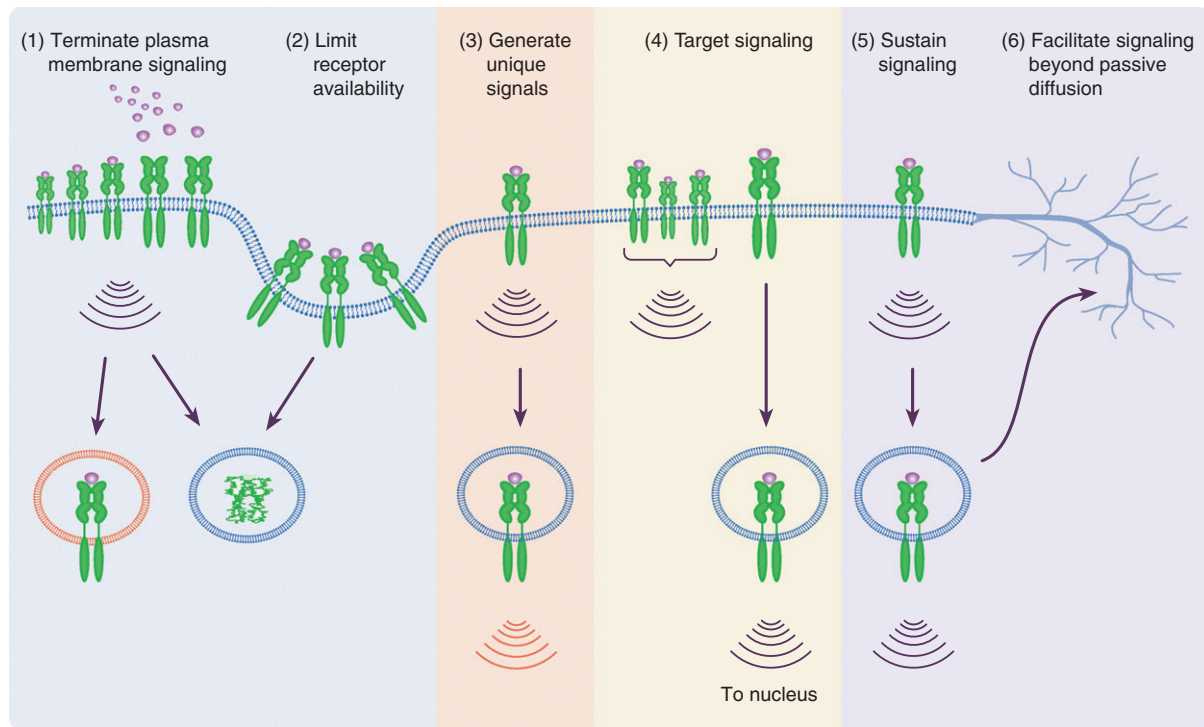


Figure 1 The various roles of endocytosis in cell signaling. Endocytosis of ligand-activated plasma membrane receptors can have different effects on signaling. (1) The ligand-induced receptor signal initiated at the plasma membrane can be terminated by endocytosis either by directing the receptor–ligand complex for degradation or due to a change in membrane composition, for example, PI(4,5)P₂ (PIP₂) (required for the plasma membrane signal) conversion to PI(3)P in endosomes. (2) Endocytosis can also limit the number of available receptors at the plasma membrane and thus reduce signaling. (3) The signal initiated at the plasma membrane can be altered or fine-tuned by ligand–receptor internalization resulting in the generation of a new signal. (4) In some cases if the ligand-induced plasma membrane signal is not strong enough, endocytosis and translocation of the signaling receptor closer to the nucleus will rescue the outcome by increasing the local signal intensity. (5) Perhaps the most common role for endocytosis is to sustain signaling after internalization and thus amplify the information. (6) Sustained growth factor signaling and active transport of endosomes, as opposed to random diffusion, particularly in the case of nerve cells, is crucial for transmitting signals from nerve terminals to distant nuclei. For more specific details and examples of endosome signaling for different receptors the reader is referred to [Table 1](#).

Subsequently, signaling endosomes containing active Lck, Lat, and phosphorylated receptors are generated to maintain TCR signaling (described in detail in [Benzing *et al.*, 2013](#)).

Endocytic Routes Specify Signaling Outcome

Endocytosis in mammalian cells is regulated by multiple classes of plasma membrane invaginations differing in their biological function, composition, and cargo recruitment. These can be divided into two major classes, clathrin-mediated endocytosis (CME) and non-clathrin endocytosis (NCE) ([Doherty and McMahon, 2009](#)). CME of plasma membrane cargo involves the specific recruitment of cargo receptors, through interaction with adapter proteins, to clathrin-coated pits (CCPs). Clathrin polymerization drives plasma membrane invagination to eventually form an endocytic vesicle in a dynamin-dependent manner. NCE is a more general term for a number of different pathways that are all clathrin-independent but dynamin dependent and here we will only present examples of caveolae-mediated endocytosis. Several plasma membrane receptor classes are known to undergo both CME and NCE depending on the context. Interestingly, the

internalization route can influence the endosomal signaling of the receptors and elicit differential biological signaling outcomes dependent on the external stimuli ([Figure 2](#)). Transforming growth factor beta receptor (TGFβR) can undergo both CME and NCE. Upon TGFβ stimulation the receptor traffics either to early endosome antigen 1 (EEA1)-positive endosomes via CME or to caveolin-positive structures via NCE. This defines the signaling outcome such that the clathrin-mediated endocytosed receptor couples to zinc finger domain found in Fab1 (FYVE) domain containing adapter protein SMAD anchor for receptor activation (SARA) and is routed to signaling-active endosomes, whereas the receptor entering via the NCE is targeted for ubiquitination and degradation ([Di Guglielmo *et al.*, 2003](#)). In the case of EGFR signaling the receptor is endocytosed only through CME if the ligand, EGF, is present in low concentrations. Following CME EGFR is trafficked to signaling endosomes and recycled to the plasma membrane to sustain signaling and minimize degradation. Conversely, at least in some cell types, high EGF levels induce additional NCE of EGFR resulting in receptor degradation ([Sigismund *et al.*, 2008](#)).

It is important to note that NCE does not necessarily lead to receptor degradation and CME to endosomal signaling.

Table 1 Summary of endosome signaling pathways reported for different receptors

Receptor type	Receptor	Ligand	Proposed outcome of endosome signaling	Proposed requirements for signaling	Notes and references
Receptor tyrosine kinases (RTKs)	Met	HGF	Sustained STAT3 phosphorylation and activation in the nucleus	Trafficking to perinuclear compartment	- Strength of STAT3 activation is receptor dependent - Endocytosis propagates weak STAT3 signal following Met activation (Kermorgant and Parker, 2008) - Biochemical fractionation revealed activated EGFR and downstream effectors (components of Ras activation pathway: Grb2, Shc, and SOS) in endosomes (Vieira et al., 1996) - RTK signaling proteins isolated from endosomes using magnetic EGF-coated microbeads (Li et al., 2005) - Targeting endocytosis blocks ERK activation by several RTKs (Vieira et al., 1996)
	EGFR	EGF	Full EGFR phosphorylation for activation of MAPK pathway	Rab5 and APPL positive endosomes	- Intracellular localization of activated EGFR, PIP₂, and lipid products of the PI3K pathway tracked in living cells using green fluorescent protein tag - PtdIns(3,4,5)P₃ and PtdIns(3,4)P₂ not detected in EGFR positive vesicles consistent with lack of accessible PIP₂ (Haugh and Meyer, 2002)
	EGFR	EGF	Attenuation of PI3K and PLC- γ signaling	Lack of PIP ₂ in endosome membranes	- TGFα-EGFR association is more pH-sensitive than EGF-EGFR and is disrupted in endosomes - Favors recycling of internalized EGFR rather than degradation (French et al., 1995; Ouyang et al., 1999; Sorkin and von Zastrow, 2002) - VEGFR is sequestered away from plasma membrane phosphotyrosine phosphatases (Lanahan et al., 2010)
	VEGFR	VEGF	VEGFR protected from dephosphorylation and PLC- γ /MAPK signaling prolonged	EEA1 positive endosomes	- Pool of ligand-bound internalized PDGFR is active and tyrosine phosphorylated (Sorkin et al., 1993) - Promotes cell survival and proliferation through Grb2, SHC, PLC-γ1, and p85α subunit of PI3K in endosomes (Wang et al., 2004) - Evidence for co-receptor function during endocytosis: - PDGFR and LRP (Newton et al., 2005) - PDGFR and ephrinB (Nakayama et al., 2013)
	PDGFR	PDGF	PDGFR signaling	CME and NCE endocytosis	

(Continued)

Table 1 Continued

Receptor type	Receptor	Ligand	Proposed outcome of endosome signaling	Proposed requirements for signaling	Notes and references
G-protein-coupled receptors (GPCRs)	TrkA receptor	NGF	Sustained MAPK signaling, activation of Rap1, and retrograde transport of survival signals	Rab5 and EEA1 positive endosomes	● PDGFR endocytosis is dependent on ligand concentration (De Donatis <i>et al.</i>, 2008): ○ Low PDGF – CME – favors cell migration ○ High PDGF – NCE – favors cell proliferation ● Neurotrophin–Trk complexes associate with endosomes and are necessary for signal transduction, for example, MAPK activation (Zweifelf <i>et al.</i>, 2005; Delcroix <i>et al.</i>, 2003; York <i>et al.</i>, 1998; York <i>et al.</i>, 2000; Wu <i>et al.</i>, 2001) ● Retrograde transport of complexes occurs in EEA1 and Rab5 endosomes ● Endosomal signaling promotes cell survival (Grimes <i>et al.</i>, 1996) ● Rapid formation of intracellular vesicles containing EphB–ephrinB complexes in cells ● Contact inhibition is reduced if endocytosis is inhibited or if either EPH or ephrin are missing cytoplasmic domains ● See Review (Egea and Klein, 2007) ● Phosphorylated insulin receptor found in endosomes (Baass <i>et al.</i>, 1995) ● RTK signaling proteins isolated from endosomes using magnetic insulin-coated microbeads (Li <i>et al.</i>, 2005) ● Overexpression of mutant dynamin inhibits insulin-mediated activation of ERK/MAPK 1/2 (Ceresa <i>et al.</i>, 1998) ● Arrestins have two other functions: 1. Arrestins 2 and 3 (main isoforms outside the visual system) act as regulated endocytic adaptors promoting GPCR-coated pit association 2. Bind various components of the MAPK modules and appear to act as scaffolds for MAPK signaling (Leikowitz and Shenoy, 2005; Moore <i>et al.</i>, 2007) ● Endocytosis of G-protein rather than the receptor itself is important for endosome signaling of GPCR (Garcia-Regalado <i>et al.</i>, 2008; Sorkin and von Zastrow, 2009)
	EPH	Ephrin	Activation of Rho GTPases, changes in actin cytoskeleton and contact-mediated repulsion of neighboring cells	Not indicated	
	Insulin receptor	Insulin	MAPK signaling	Not indicated	
	β 2AR	Agonist isoproterenol	Activation of ERK1, ERK2, and Jun N-terminal kinase	CME and requires β -arrestins	
	Specific GPCR not indicated	LPA	PI3K and Akt activation	G-protein $G\beta\gamma$ 2 translocation to endosomes and binding to Rab11A following GPCR activation	

Serine/threonine kinase receptors	Fz1 (distantly related to GPCRs)	Wnt	Stabilization of cytoplasmic β -catenin	CME	- Endocytosis of Fz1/Wnt/Arrow complexes into clathrin-coated pits is promoted by the receptor associated protein Dvl - Dvl binding to AP2 appears to be critical for receptor signaling (Yu <i>et al.</i>, 2007; Blitzler and Nusse, 2006) - Dvl also binds β-arrestins (Seto and Bellen, 2006) - Complexes recruited to early endosomes - Formation of protein scaffolds required for signaling (Di Guglielmo <i>et al.</i>, 2003; Hayes <i>et al.</i>, 2002)
	TGF β receptor	TGF β	Phosphorylation of SMAD transcription factors and translocation to the nucleus	CME and requires FYVE protein SARA	- Alternative route of endocytosis – caveolin-dependent for receptor turnover and degradation - TGFβ receptor binding to SMAD7–Smurf2, recruitment of E3-ligase and ubiquitination of TGFβ are required for degradation (Di Guglielmo <i>et al.</i>, 2003)
	TGF β receptor	TGF β	TGF β receptor turnover	NCE (caveolin-mediated)	
Cytokine receptors	TNFR	TNF α	Caspase 8 activation in endosomes and TNF α apoptotic signaling	Endosomal PI3P and proteolytic enzymes	- Promotes assembly of the death complex in endosomes - In contrast to the survival signals initiated at the plasma membrane, TNFR signaling in endosomes induces cell death (Schneider-Brachert <i>et al.</i>, 2004; Liao <i>et al.</i>, 2008)
	TLR4	LPS	Switch from surface to endosome-specific signaling and induction of INF β gene.	Lack of PIP ₂ in endosomes	- Other TLR family members are found to reside mostly in endosomes where they detect viral nucleic acids (Kagan <i>et al.</i>, 2008) - Sustained endosomal TLR9 shown to be necessary for a potent type I INF response (Honda <i>et al.</i>, 2005)
Notch signaling	NOTCH	DSL	NOTCH cleavage and translocation to the nucleus	Low pH, γ -secretase and dynamin	γ-secretase is present in both endosomes and plasma membrane, however, peak activity is at low pH (Pasternak <i>et al.</i>, 2003) - NOTCH-DSL interaction is favored by low pH (Pei and Baker, 2008) - Endocytosis is necessary for cleavage of NOTCH (Vaccari <i>et al.</i>, 2008) - Role for dynamin in NOTCH endocytosis and <i>Drosophila</i> neurogenesis (Seugnet <i>et al.</i>, 1997)

(Continued)

Table 1 Continued

Receptor type	Receptor	Ligand	Proposed outcome of endosome signaling	Proposed requirements for signaling	Notes and references
	NOTCH	DSL	NOTCH signaling	Endocytosis of DSL on signal-sending cell	- Alternative models of NOTCH signaling involve endocytosis of its ligand DSL: - Endocytosis of DSL physically strips NOTCH extracellular domain to release NOTCH cytoplasmic domain for cleavage (Windler and Bilder, 2010) - Recycling of DSL is necessary for DSL activation and high plasma membrane concentration (Rajan <i>et al.</i>, 2009) – <i>Drosophila</i> - Posttranslational modifications, for example, mono-ubiquitination in signaling or recycling endosomes is required to activate ligands, for example, DSL (remains to be clarified) (Fortini, 2009)

Note: Signaling from endosomes is key for propagation and enhanced receptor signaling (light green rows), attenuation of receptor signaling (pink rows) or for a change in the receptor signaling outcome (pale yellow rows).

Abbreviations: APPL, adapter protein, phosphotyrosine interaction, PH domain and leucine zipper-containing; β 2AR, beta 2-adrenergic receptor; CME, clathrin-mediated endocytosis; DSL, *delta/serrate/lag* family; Dvl, disheveled; EEA1, early endosome antigen 1; EGF, epidermal growth factor; EGFR, EGF receptor; ERK, extracellular signal-regulated kinase; FYVE, zinc finger domain found in Fab1; YOTB/ZK632.12, Vac1, and EEA1; Fz1, frizzled; GPCR, G-protein–coupled receptor; HGF, hepatocyte growth factor; INF β , interferon beta; LPA, lysophosphatidic acid; LPS, lipopolysaccharide; LRP, low density lipoprotein receptor-related protein; MAPK, mitogen-activated protein kinase; Met, HGF receptor; NCE, non-clathrin-mediated endocytosis; NGF, neuronal growth factor; PDGF, platelet-derived growth factor; PDGFR, PDGF receptor; PI3K, phosphatidylinositol 3-kinase; PIP $_2$, phosphatidylinositol (4,5) bisphosphate; PLC- γ , phospholipase C-gamma; RTK, receptor tyrosine kinase; SARA, SMAD anchor for receptor activation; STAT3, signal transducer and activator of transcription 3; TGF, transforming growth factor; TGFR, TGF receptor; TLR, toll-like receptor; TNF α , tumor necrosis factor alpha; TNFR, TNF receptor; TrkA, neurotrophic tyrosine kinase receptor type 1; VEGF, vascular endothelial growth factor; VEGFR, VEGF receptor; Wnt, wingless-type MMTV integration site family member.

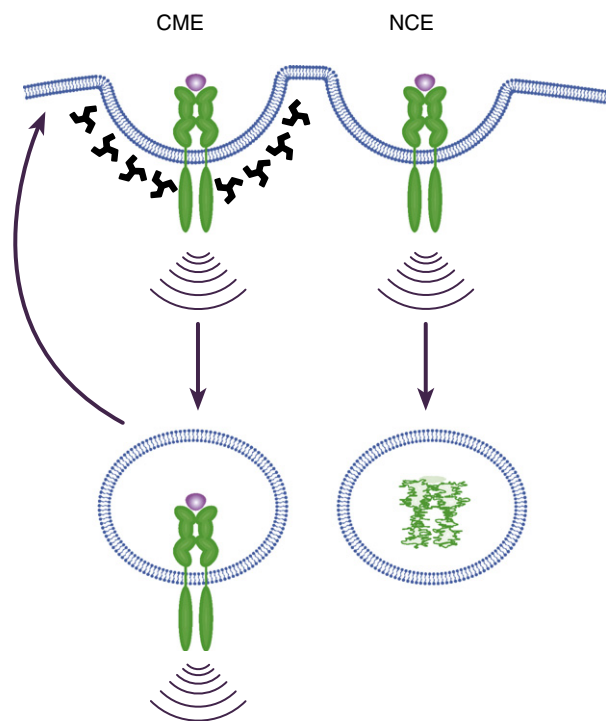


Figure 2 Different endocytosis routes lead to different biological outputs. EGFR and TGF β R can both be endocytosed via CME or NCE routes. Receptors endocytosed via CME are recycled back to the plasma membrane thus sustaining the signal, while receptors trafficked via the NCE route are directed to lysosomes for degradation. In the case of TGF β R, CME to EEA1 positive endosomes leads to receptor coupling with a SMAD anchor protein, SARA, which promotes TGF β R signaling from endosomes, whereas the caveolin-dependent internalization pathway leads to rapid receptor turnover.

Wnts constitute a large family of secreted ligands implicated in a wide array of developmental and physiological processes (Logan and Nusse, 2004; Yamamoto *et al.*, 2008). Wnt-induced low-density receptor-related protein 6 (LRP6) signaling results in NCE and relocalization of LRP6 to endosomes, where the Wnt-generated signal functions to stabilize β -catenin (Yamamoto *et al.*, 2008). In contrast, in response to Dickkopf family member 1 (Dkk1), a soluble ligand and a negative regulator of Wnt signaling, LRP6 undergoes CME and receptor degradation. In conclusion, the internalization routes of receptors can play a defining role on the signaling outcome upon ligand stimulation and interplay between different trafficking pathways facilitates fine-tuning of cellular signals.

Endocytosis Regulates Signal Termination and Cellular Response to Ligand Gradient

Ligand-induced endocytosis of RTKs is an important mechanism to downregulate cellular responsiveness to external stimuli. This may occur via sequestration of the receptor away from plasma membrane components that are critically important for signaling. For example, on the plasma membrane EGF-induced EGFR phosphorylation recruits SH2-domain containing lipid kinase PI3K or lipase phospholipase

C-gamma, that depend on plasma membrane enriched phospholipids as substrates. Following endocytosis the membrane lipid composition of endosomes no longer supports signaling by these molecules and results in signal termination (Haugh and Meyer, 2002).

Another important mechanism for endocytosis-regulated signal termination is lysosomal receptor degradation and subsequently reduced plasma membrane receptor levels. Lowered receptor levels can attenuate the overall signaling intensity within the cell but can also regulate cellular responsiveness in a localized and directional manner. An example of this localized receptor endocytosis and downregulation is demonstrated by the directionally migrating border cells during *Drosophila melanogaster* oogenesis (Jekely *et al.*, 2005; Duchek *et al.*, 2001). In this developmental process spatial localization of RTK signaling via EGFR and PVR (PDGF/vascular growth factor receptor) regulates directional cell migration. Receptor endocytosis is necessary in these cells to ensure a localized intracellular response to the growth factor-mediated guidance cues (Duchek *et al.*, 2001; Ducheka and Rorth, 2001). A similar mechanism is also involved in regulating cell migration within a growth factor gradient. Upon EGF stimulation the EGFR becomes sequestered at the advancing front of culture breast carcinoma cells. In this spatially restricted sub-cellular space endocytosis-mediated receptor downregulation gradually reduces cell responsiveness to chemotactic EGF and promotes directional cell migration (Bailly *et al.*, 2000; Figure 3).

Endosomes – A World of Their Own

What is special about endosomes? How are they capable of regulating specialized signaling events that are distinct from plasma membrane signals? What underlies their unique ability to allow vast diversification of signals within cells? As discussed above endosomes are capable of assembling novel signaling platforms, amplifying signals and differentiating between extended signaling and signal termination via degradation. Since the establishment of the signaling endosome concept several unique endosomal characteristics have been shown to contribute to endosome signaling. These unique physical characteristics of endosomes are: differential enrichment of membrane phosphoinositide phosphates (PIPs), the ability to recruit unique signaling adapter proteins, low acidic pH especially in the luminal compartment of late endosomes, high membrane curvature, confined size, and differential accessibility to cellular phosphatases (Figure 4).

PIPs

Endosomal membranes are enriched in phosphatidylinositol-3-phosphate (PI3P). This distinguishes them from the plasma membrane which is enriched in PI(4,5)P₂ (PIP₂) and PI(3,4,5)P₃ (PIP₃). In addition, the presence of PI3P defines different classes of endosomes by regulating the formation of endosome-specific signaling platforms. A population of early endosome precursors contain APPL1 and 2 (adapter protein, phosphotyrosine interaction, PH domain, and leucine zipper-containing) adapter proteins and lack PI3P. In the

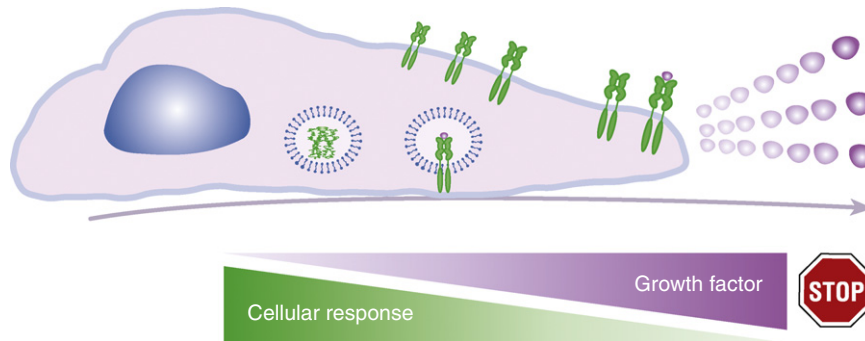


Figure 3 Localized endocytosis regulates the distance and directionality of cell migration. Several plasma membrane receptors function as motogenic sensors allowing cells to respond and migrate toward an increasing concentration of a soluble chemoattractant and receptor endocytosis appears to be crucial for this process. For example, in an EGF gradient, EGFR is evenly distributed on the cell surface. However, the endocytosis of the ligand–receptor complex is polarized and occurs primarily at the leading edge of migrating cells where the EGF concentration is highest. The internalized receptors are guided for degradation thus reducing cell responsiveness to increasing levels of EGF, finally leading to a halt in cell migration.

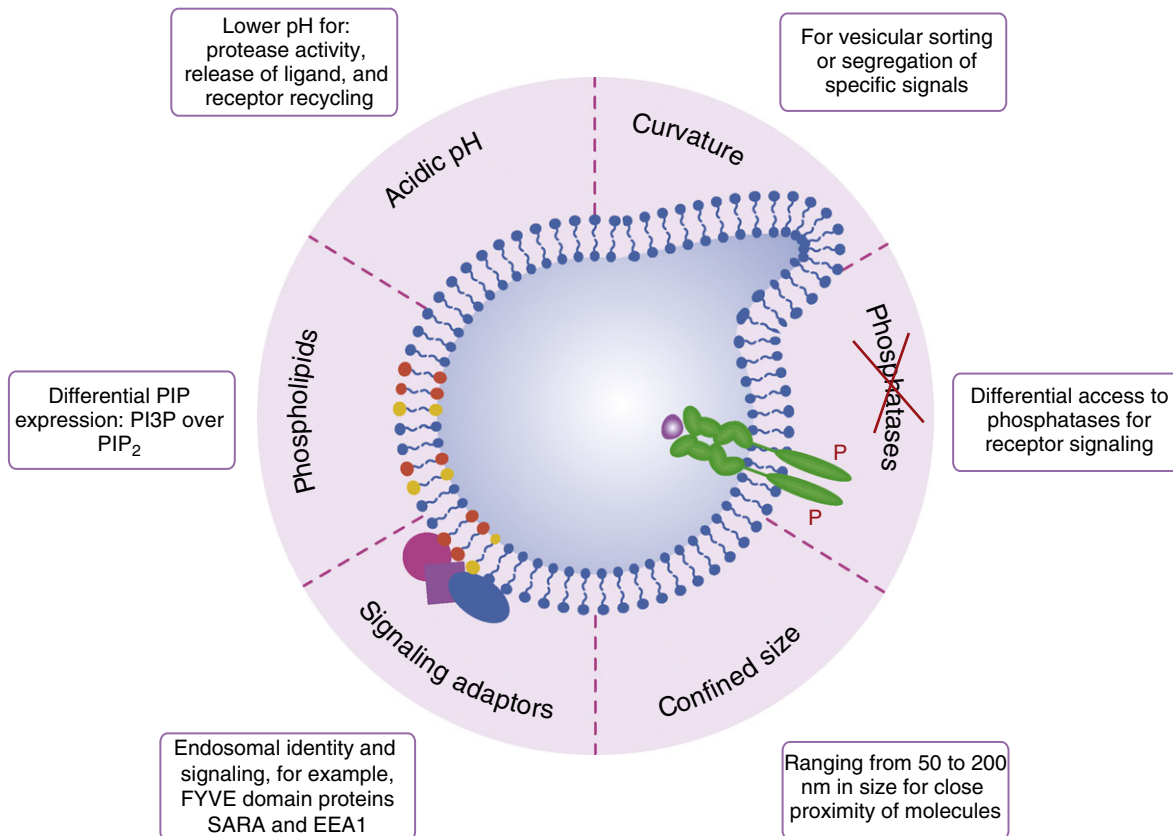


Figure 4 The unique signal promoting features of endosomes. Endosomes have several special characteristics that enable them to regulate, modify, and sustain signaling. Endosomal membranes are rich in PI(3)P allowing the recruitment of several PI(3)P-binding domain (such as FYVE) containing proteins to endosomes. Endosomes also exhibit lower pH than the cell cytoplasm, which is needed for some proteases to function. In addition, endosomal membranes are inevitably more curved compared to the plasma membrane, allowing endosomes to recruit specific curvature-sensing proteins such as those containing a BAR-domain. Endosomes also enable receptors to escape from the action of plasma membrane phosphatases leading to sustained phosphorylation-dependent signaling and finally the compact size of endosomes can facilitate ligand–receptor interactions by increasing proximity of the molecules.

APPL-endosomes, APPLs interact with endosomal membranes directly, and through Rab5-guanosine triphosphate (GTP) association. They also interact directly or indirectly with active

RTKs. On very early endosomes APPLs are important for survival signaling along the Akt and glycogen synthase kinase 3 beta (GSK3 β) axis in zebrafish (Schenck *et al.*, 2008). However,

there appear to be context-dependent variations on the pathways regulated by APPL-endosomes. In zebrafish, Akt signaling to mammalian target of rapamycin (mTOR) occurs independently of APPLs whereas in cultured HeLa cells APPL-endosomes are implicated in activation of ERK1 and ERK2 as well as Akt signaling to both GSK3 β and mTOR (Zoncu *et al.*, 2009).

Maturation of APPL-endosomes to EEA1-containing early endosomes involves acquisition of PI3P through the recruitment of the class 3 PI3K. APPL and EEA1 compete for binding to Rab5-GTP on endosomes and the generation of PI3P shifts the balance in favor of EEA1 (Zoncu *et al.*, 2009). This so-called phosphoinositide-switch induces endosomal maturation and terminates APPL-dependent signaling from the endosomes. Released APPL can translocate to the nucleus where it regulates chromatin remodeling in conjunction with endosomal sorting complexes required for transport III (ESCRT-III) proteins (Stauffer *et al.*, 2001). Rab5 and EEA1 additionally recruit specific signaling complexes to early endosomes to determine signal specificity. For example, active Akt interacts with EEA1 and EEA1 endosomes are crucial for angiotensin-II-induced Akt activation and downstream signaling to mTOR and S6K, but not to ERK1/2 (Nazarewicz *et al.*, 2011).

A third class of early endosomes consists of the SARA-endosomes. SARA is an FYVE-domain containing adapter protein that localizes to a subset of PIP₃-containing endosomes. SARA binds directly to active TGF β R steering the receptor toward SARA-endosomes via CME. SARA recruits the TGF β R target SMAD2 to the same endosomal membrane and therefore promotes efficient phosphorylation of SMAD2 by TGF β R. This induces dissociation of SMAD2 from the endosomes and the formation of SMAD2–SMAD4 complexes that translocate to the nucleus to regulate transcription (Tsukazaki *et al.*, 1998).

Endosome Maturation and Unique Signaling Adapters

Different endosomal compartments are defined by specific Rab proteins. Rab5 is a small GTPase and an important signaling molecule characteristic to early endosomes and as described above regulates signaling specificity in endosomes. However, Rab5 is also involved in regulating the duration of endosomal signals. Endosome maturation involves a transition from Rab5 early endosomes to Rab7 late endosomes and results in a termination of Rab5-dependent signaling (Rink *et al.*, 2005). Activation of P38/mitogen-activated protein kinase (MAPK) induced by EGFR signaling, stress, or cytotoxic drugs promotes the formation of Rab5-GDI-EEA1 complexes and slows down the conversion of Rab5 to Rab7 endosomes resulting in accelerated EGFR endocytosis (Cavalli *et al.*, 2001; Zwang and Yarden, 2006; Sorkin and von Zastrow, 2009; von Zastrow and Sorkin, 2007). Thus, extended residency of receptors in Rab5 endosomes is linked to prolonged signaling.

Specific signaling platforms also exist on late endosomes and play a role in the ability of RTKs like EGFR to fully activate ERK1/2. On late endosomes, endosome-specific lipid-raft adapter protein p18 binds to a scaffold consisting of MAPK and ERK kinase 1 (MEK1), MEK1 partner (MP1) and p14, an adapter protein. Depending on the type of signaling system p14 and p18 are necessary for maintenance of basal MEK-ERK

activity or efficient MEK activation following EGF stimulation (Teis *et al.*, 2002, 2006; Nada *et al.*, 2009). In addition to RTK signaling, the late endosomal p14–MP1 complex is also important for K-Ras, but not H-Ras or N-Ras, signaling. Like other members of the Ras family, K-Ras is a GTPase and an early player in many signal transduction pathways. Active K-Ras is recruited to late endosomes where together with the p14–MP1 complex signals to Raf1 kinase following EGF stimulation (Lu *et al.*, 2009). Thus, endosomal maturation functions to specify signaling and facilitates assembly of signaling scaffolds found uniquely on late endosomes.

Low pH

As internalized materials progress through the endocytic pathway, they are exposed to an increasingly acidic pH. The pH of early endosomes is typically near 6, late endosomes near 5, and lysosomes exhibit even lower ranges. There are several ways by which the low endosomal pH can influence cellular signaling. EGF and TGF α are both EGFR ligands. EGF has a higher affinity for the receptor and is thus more likely to remain receptor-bound in endosomes inducing receptor targeting for lysosomal degradation. Conversely, TGF α dissociates from EGFR at an early stage of endocytosis in the mildly acidic environment of the early endosomes. This induces rapid recycling of the receptor and enables renewed receptor activation and signaling on the plasma membrane by the recycled receptor. Consequently, EGF but not TGF α induces receptor downregulation and this is reflected by the fact that TGF α is a much more potent mitogen than EGF (Waterman *et al.*, 1998). In the NOTCH pathway, plasma membrane cleavage of NOTCH results in endosomal localization of its transmembrane and intracellular domains. In the endosomes the NOTCH domains are further cleaved by a low pH-dependent protease, γ -secretase, functional only in late endosomes. This step is necessary for the nuclear localization of NOTCH to activate transcription (Vaccari *et al.*, 2008).

Confined Size

The luminal volume on endosomes is small and the limited size of the endosomes may help concentrate receptors and their ligands in a manner enabling continued receptor–ligand occupancy and thus sustained signaling from the endosome (Sorkin and von Zastrow, 2009). In addition, many of the protein–protein and protein–lipid interactions between endosomal small GTPases and their effectors, scaffolding proteins and signaling molecules are typically of low affinity. The limited surface area of an endosome is likely to favor and promote the assembly of such multi-protein complexes (Pawson, 2007; Scita and Di Fiore, 2010). Furthermore, proteins like APPLs have membrane curvature-sensing domains which may also enhance specific interactions with endosomal membranes.

Accessibility to Phosphatases

Signaling by activated RTKs can be effectively attenuated by protein tyrosine phosphatases (PTPs) (Mattila *et al.*, 2005) and internalization to signaling-competent endosomes may function to sequester receptors away from PTPs. In the case of

the Met-receptor, endocytosis prolongs signaling to STAT3, possibly due to reduced accessibility to plasma membrane phosphatases (Kermorgant and Parker, 2008). In T cells, signaling of Lck kinase in endosomes is counteracted by CD45 phosphatase (Danielian *et al.*, 1992). However, at present it is not known whether CD45 is excluded from these endosomes. On the other hand, phosphatases can promote endosomal signaling. Endosomally localized PP2A serine-threonine phosphatase is essential for resensitization of G-protein-coupled receptor (GPCR) signaling enabling relocalization of the receptor to the plasma membrane in a functional form (Scita and Di Fiore, 2010). Additionally, a role for phosphatases in endosome regulation has been described downstream of the Ephrin B receptor. Activation of this receptor leads to the phosphorylation of synaptojanin, a phosphatidylinositol-5-phosphatase involved in uncoating of clathrin-coated vesicles following internalization, resulting in a subsequent decrease in phosphatase activity that influences the CME of the transferrin receptor and α -Amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) glutamate receptor (Irie *et al.*, 2005). Synaptojanin is also believed to be preferentially recruited to curved endosome membranes by interacting with BAR-domain proteins. This specific recruitment of synaptojanin promotes removal of PIP₂ (Chang-Ileto *et al.*, 2011) and therefore provides a differential phospholipid environment for protein interaction or chemical reactions.

Endocytosis Regulates Spatial Signaling and Signal Amplification

As described above, several RTKs and GPCRs remain in their ligand-bound active state after internalization in endosomes, thus enabling sustained receptor signaling. The relatively long residence time in endosomes compared to the generally fast internalization step from the plasma membrane allow signal amplification and ensure sufficient intensity of signaling. On the other hand, endosomes can protect phosphorylated receptor/protein signals from cytoplasmic phosphatases by transporting the membrane-associated signaling molecules rapidly inside the cell via directional microtubules. This concept was elegantly demonstrated in a mathematical model of transcription factor STAT3 signaling first presented by Howe (2005). STAT3 is activated by phosphorylation following engagement of different receptors at the plasma membrane after which STAT3 needs to be translocated into the nucleus. In the mathematical model, Howe compared a random diffusion of phosphorylated STAT3, combined with probabilistic modeling of dephosphorylation kinetics, to a direct membrane-associated microtubule-guided transportation system from the plasma membrane to the nucleus. Based on this model, if the distance traveled by phosphorylated signaling proteins exceeds 200 nm, the endosomal transport system along microtubules is more efficient and less information is lost due to dephosphorylation compared to a random diffusion (Howe, 2005).

In some cases, however, the signaling intensity is sufficiently high enough to reach the required threshold so that a simple diffusion through a phosphatase-rich cytoplasm can in fact be sufficient to relocate active signaling proteins to the nucleus. For example, in the case of STAT3, the efficiency and

amplitude of STAT3 activation seems to be receptor dependent. Cytokine oncostatin-M activates STAT3 from the plasma membrane after which STAT translocates to the nucleus by classical diffusion. However, the hepatocyte growth factor (HGF)-induced Met-receptor generated STAT-signal is weaker and thus the activated receptor itself needs to be transported by endocytosis closer to the nucleus to reach the same threshold of signaling and lead to nuclear accumulation of STAT3. Yet this is not the case for all Met signaling events as HGF-Met-induced nuclear accumulation of MAPK does not require perinuclear Met localization, although endocytosis is needed for full ERK1/2 activation.

A classic example of sustained and spatially regulated signaling mediated by endocytosis is the signaling of neurotrophin receptors. As already described earlier in this article, NGF-induced TrkA receptors need to transmit signals from nerve terminals to the nucleus over long distances (which can easily reach a meter). This would of course not be possible with random diffusion in a limited time and is only accomplished by endosomal transportation of the receptor–ligand complexes, thus sustaining the receptor-mediated ERK1/2 and PI3K/Akt signaling over long distances. Spatial restriction of signals is also critical for actin remodeling and thus for cell migration and invasion. Rac GTPase, an important signaling molecule involved in actin reorganization, has been shown to be activated on early endosomes after Rab5-mediated endocytosis. Activated Rac is then recycled to specific locations at the plasma membrane to provide localized signaling leading to actin polymerization-driven membrane protrusion and finally to increased cell migration and invasion (Palamidessi *et al.*, 2008).

An intriguing example of the importance of endocytosis in defining signaling outcome is described for the death receptor tumor necrosis factor receptor 1 (TNFR1). Ligand-induced activation of TNFR1 on the plasma membrane induces anti-apoptotic signals by activating the nuclear factor kappa B (NF- κ B) leading to cell protection. However, receptor internalization via CME leads to the recruitment of the ‘death-inducing signaling complex’ followed by termination of NF- κ B signaling that ultimately leads to apoptosis. In conclusion, the same receptor–ligand complex can induce both pro- and anti-apoptotic signaling, depending solely on the subcellular location (Schneider-Brachert *et al.*, 2004; Schutze *et al.*, 2008).

Signaling Regulates Endocytosis

The examples discussed so far highlight the important role of endosomes and endocytic pathways in determining the outcome of a membrane receptor-initiated signal. However, there is an extensive and reciprocal relationship between endocytosis and plasma membrane-derived signaling events. Indeed, accumulating evidence suggests that receptor-initiated signals can control endocytosis through the activation of pathways that impinge on endosome formation and maturation and on the route of internalization.

Transcription and Phosphorylation

Signals that control transcriptional programs represent the earliest stage at which endocytosis may be regulated.

For example, activation of hypoxia-induced transcription factor 1 α (HIF1 α) delays early endosome maturation by inhibiting Rabaptin-5 transcription (a Rab5 GTPase effector) leading to endosome retention of EGFR and sustained signaling (Wang *et al.*, 2009; Zhong *et al.*, 1999). Stress signals can also activate the tumor suppressor transcription factor p53 (TP53) resulting in suppressed cell growth through the regulation of many genes. In terms of endocytosis, TP53 activation mediates transcription of an ESCRT-II complex subunit, a lysosomal membrane protein and caveolin-1 resulting in simultaneous internalization of caveolin-1 and EGFR to multivesicular bodies to inhibit EGFR signaling and cell growth during stress conditions (Yu *et al.*, 2009).

Phosphorylation-dependent regulation of endocytosis can occur at several stages of endocytosis to influence the uptake of receptors and the formation and maturation of endosomes. Ligand-induced CME is usually dependent on receptor activation and autophosphorylation of specific residues within the receptor to recruit downstream effectors. In addition, ligand-induced activation of receptors, such as EGFR or TrkA, has been shown to trigger a rapid increase in the number and density of plasma membrane-associated CCPs that is dependent on Src kinase phosphorylation of the clathrin heavy chains and assembly of the clathrin triskelion (Wilde *et al.*, 1999; Beattie *et al.*, 2000; Sorkin and von Zastrow, 2009), clearly demonstrating that ligand-triggered signaling can initiate endocytosis.

The link between signaling and endocytosis has been investigated further in comprehensive RNA interference (RNAi) screens. In 2005, Pelkmans *et al.* revealed a role for a number of kinases in endocytosis in human culture cells. Interestingly, several kinases produced opposing effects on CME of the transferrin receptor and NCE of virus suggesting that distinct pathways can be coordinated by the same kinase (Pelkmans and Zerial, 2005). More recently, a genome-wide RNAi screen revealed that there is extensive cross talk between different molecular systems such that the components of signaling pathways including Wnt, integrins, TGF β , and NOTCH appear to regulate the endocytosis of other plasma membrane-derived signals including transferrin and EGF at different levels (Collinet *et al.*, 2010).

Actin in Regulation of Endocytosis – Actin Adapters, Stabilization of Tubules, and Vesicle Motility

The spatial and dynamic regulation of actin assembly and polymerization occurs downstream of several signaling events, including cell adhesion and RTK activation, and is fundamental for determining cell shape and the mode of cell migration. Tight regulation of actin dynamics is in turn crucial for several forms of cellular uptake, including phagocytosis, macropinocytosis, and some cases of NCE. Much of the early discussions for a role of actin in CME endocytosis have stemmed from studies in yeast demonstrating a complete dependency on actin dynamics for generation of membrane curvature, stabilization of endocytic vesicles, and vesicular scission. While it is apparent that this is not the case in mammalian cells where actin is dispensable for endocytosis, nevertheless an important regulatory role for actin in mediating endocytosis in mammalian cells has become evident. In

particular, live-cell imaging has revealed that several actin-binding/regulatory proteins are recruited to CCP (Sigismund *et al.*, 2012; Merrifield *et al.*, 2005; Yarar *et al.*, 2005) prior to detachment of the vesicles from the plasma membrane. In addition, several mammalian proteins have been shown to directly link endocytic pathways to the actin cytoskeleton. For example: dynamin, a GTPase involved in vesicle fission, binds directly to cortactin, an actin-related protein 2/3 (ARP2/3) activator and inducer of actin polymerization (Kessels *et al.*, 2001; McNiven *et al.*, 2000), HIP1R, a component of CCPs binds actin and cortactin (Le Clainche *et al.*, 2007) and intersectin binds to N-WASP to promote actin nucleation and facilitate CME (Hussain *et al.*, 2001). These interactions may serve to couple cellular processes that lead to extensive actin remodeling and cellular reorganization – such as changes in cell shape during cell migration – to the tight regulation of the endocytic pathway.

The requirement for actin in endocytosis has been further demonstrated in specific cell types. In neuronal cells, proteomic analysis of TrkA signaling endosomes revealed an association of TrkA complexes with actin-modulatory proteins such as cofilin. Importantly, cofilin-mediated dissociation of actin, in distal axons, was shown to be essential for NGF-stimulated retrograde transport of TrkA endosomes and thus potentiation of survival signals (Harrington *et al.*, 2011). Another study has highlighted the requirement for actin in stabilizing the formation of endosomal tubules and for concentrating B2AR (a GPCR) in these tubules for differential sorting (Puthenveedu *et al.*, 2010). Actin is further involved in both yeast and mammalian cells in endosome motility where actin, together with microtubules forms a crucial component of the transport machinery driving vesicular traffic (Ross *et al.*, 2008).

The examples described here are only a subset of the signaling pathways that regulate endocytosis. The reciprocal nature of regulation between signaling and the endocytic machinery is very complex and is crucial for constraining biological activities within cells that may otherwise lead to disease.

Endocytosis and Implications in Cancer

Considering the vast number of signaling pathways influenced by endocytosis and endosome signaling it is perhaps not surprising to find alterations in endocytic proteins and regulators in many human diseases. An interesting and comprehensive summary of altered endocytic proteins in human disease can be found in Sigismund *et al.*, 2012.

Increasing evidence also supports a role for the endocytic machinery, in particular endosome signaling, in cancer progression. For example, tumor hypoxia and HIF1 α overexpression, which has been shown to sustain endosomal EGFR signaling, are associated with aggressive tumor progression and poor prognosis in patients (Sigismund *et al.*, 2012; Wang *et al.*, 2009; Zhong *et al.*, 1999) and endosomal localization of EGFR appears to confer drug resistance to therapeutic antibodies (Nevo *et al.*, 2009).

Endocytosis also regulates several other oncogenic RTK signaling pathways. Met RTK and its ligand HGF are frequently

overexpressed in human cancer resulting in hyperactivation of the pathway and induction of cell functions such as proliferation, migration, and survival. Activated Met is rapidly internalized and continues to signal from endosomal compartments prior to degradation. Importantly, Met signaling from endosomes has recently been shown to be involved in tumorigenesis and endocytosis appears to be critical for active Met mutants to trigger anchorage-independent growth, transformation, and metastasis (Joffe *et al.*, 2011). In addition to RTK signaling, endocytosis is required for: activation of other oncogenic receptors (e.g., NOTCH) (Furthauer and Gonzalez-Gaitan, 2009), trafficking of key adhesion receptors (e.g., integrins) to regulate cell proliferation, cell migration, and cell survival and for spatial restriction of signals involved in cell migration (e.g., Rho GTPase signaling) (Caswell and Norman, 2008; Sigismund *et al.*, 2012).

These signaling and trafficking events are often misregulated in cancer and contribute to a highly motile and invasive cell phenotype. Additionally, there is now mounting evidence for a more direct role of endocytic proteins or regulators in human tumors (Mosesson *et al.*, 2008; Lanzetti and Di Fiore, 2008).

Conclusions and Future Perspectives

Joint efforts from cell biologists, molecular biologists, biochemists, and developmental biologists have firmly established the concept of signaling endosomes and demonstrated the functional relevance of endosome-derived signal transduction in a wide range of biological systems. Nevertheless, compared to the large body of work devoted to plasma membrane signaling our knowledge of endosomal signaling still remains rather underdeveloped. Distinguishing between plasma membrane and endosomal signals is challenging and requires careful experimental set up. Hopefully, future advances in techniques such as quantitative proteomics, improved cell fractionation protocols and use of specific kinase inhibitors to restrict signaling post-endocytosis will facilitate research in this field. RTKs and other cell surface receptors are the focus for designing targeted therapies for many human diseases including cancer. An in-depth understanding of the entire signaling system from the plasma membrane to the endosomes and lysosomes will be a key for the success of these translational efforts. Furthermore, as with any signaling pathway in cells, endosomal signaling appears to be highly context-dependent and is most likely intimately linked to adhesion, the cytoskeleton and micro environmental cues sensed by the cells. This is a largely unexplored area in endosomal signaling and increasing our understanding in this area will require joint efforts from different fields of biology.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: G Protein-Coupled Receptors. Interorganellar Communication: Interplay and Processes: Clathrin and Clathrin-Dependent Endocytosis; Clathrin Independent Endocytosis; Endosome to Lysosome Transport; Rabs and Other G Proteins; Rabs of the Endosomal Recycling Pathway. Organelles: Structure

and Function: Early Endosomal Compartments; The Late Endosome

References

- Baass, P.C., Di Guglielmo, G.M., Authier, F., Posner, B.I., Bergeron, J.J., 1995. Compartmentalized signal transduction by receptor tyrosine kinases. *Trends in Cell Biology* 5 (12), 465–470.
- Bailly, M., Wyckoff, J., Bouzahzah, B., *et al.*, 2000. Epidermal growth factor receptor distribution during chemotactic responses. *Molecular Biology of the Cell* 11 (11), 3873–3883.
- Beattie, E.C., Howe, C.L., Wilde, A., Brodsky, F.M., Mobley, W.C., 2000. NGF signals through TrkA to increase clathrin at the plasma membrane and enhance clathrin-mediated membrane trafficking. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 20 (19), 7325–7333.
- Benzing, C., Rossy, J., Gaus, K., 2013. Do signalling endosomes play a role in T cell activation? *FEBS Journal* 280 (21), 5164–5176.
- Blitzer, J.T., Nusse, R., 2006. A critical role for endocytosis in Wnt signaling. *BMC Cell Biology* 7, 28.
- Caswell, P., Norman, J., 2008. Endocytic transport of integrins during cell migration and invasion. *Trends in Cell Biology* 18 (6), 257–263.
- Cavalli, V., Vilbois, F., Corti, M., *et al.*, 2001. The stress-induced MAP kinase p38 regulates endocytic trafficking via the GDI: Rab5 complex. *Molecular Cell* 7 (2), 421–432.
- Ceresa, B.P., Kao, A.W., Santeler, S.R., Pessin, J.E., 1998. Inhibition of clathrin-mediated endocytosis selectively attenuates specific insulin receptor signal transduction pathways. *Molecular and Cellular Biology* 18 (7), 3862–3870.
- Chang-Ileto, B., Frere, S.G., Chan, R.B., *et al.*, 2011. Synaptojanin 1-mediated PI (4,5)P₂ hydrolysis is modulated by membrane curvature and facilitates membrane fission. *Developmental Cell* 20 (2), 206–218.
- Collinet, C., Stoter, M., Bradshaw, C.R., *et al.*, 2010. Systems survey of endocytosis by multiparametric image analysis. *Nature* 464 (7286), 243–249.
- Danielian, S., Alcover, A., Polissard, L., *et al.*, 1992. Both T cell receptor (TcR)-CD3 complex and CD2 increase the tyrosine kinase activity of p56lck. CD2 can mediate TcR-CD3-independent and CD45-dependent activation of p56lck. *European Journal of Immunology* 22 (11), 2915–2921.
- Das, V., Nal, B., Dujancourt, A., *et al.*, 2004. Activation-induced polarized recycling targets T cell antigen receptors to the immunological synapse; involvement of SNARE complexes. *Immunity* 20 (5), 577–588.
- De Donatis, A., Comito, G., Buricchi, F., *et al.*, 2008. Proliferation versus migration in platelet-derived growth factor signaling: The key role of endocytosis. *Journal of Biological Chemistry* 283 (29), 19948–19956.
- Delcroix, J.D., Valletta, J.S., Wu, C., *et al.*, 2003. NGF signaling in sensory neurons: Evidence that early endosomes carry NGF retrograde signals. *Neuron* 39 (1), 69–84.
- Di Guglielmo, G.M., Le Roy, C., Goodfellow, A.F., Wrana, J.L., 2003. Distinct endocytic pathways regulate TGF-beta receptor signalling and turnover. *Nature Cell Biology* 5 (5), 410–421.
- Doherty, G.J., McMahon, H.T., 2009. Mechanisms of endocytosis. *Annual Review of Biochemistry* 78, 857–902.
- Duchek, P., Rorth, P., 2001. Guidance of cell migration by EGF receptor signaling during *Drosophila* oogenesis. *Science (New York, NY)* 291 (5501), 131–133.
- Duchek, P., Somogyi, K., Jekely, G., Beccari, S., Rorth, P., 2001. Guidance of cell migration by the *Drosophila* PDGF/VEGF receptor. *Cell* 107 (1), 17–26.
- Egea, J., Klein, R., 2007. Bidirectional Eph-ephrin signaling during axon guidance. *Trends in Cell Biology* 17 (5), 230–238.
- Fortini, M.E., 2009. Notch signaling: The core pathway and its posttranslational regulation. *Developmental Cell* 16 (5), 633–647.
- French, A.R., Tadaki, D.K., Niyogi, S.K., Lauffenburger, D.A., 1995. Intracellular trafficking of epidermal growth factor family ligands is directly influenced by the pH sensitivity of the receptor/ligand interaction. *Journal of Biological Chemistry* 270 (9), 4334–4340.
- Furthauer, M., Gonzalez-Gaitan, M., 2009. Endocytic regulation of notch signalling during development. *Traffic (Copenhagen, Denmark)* 10 (7), 792–802.
- Garcia-Regalado, A., Guzman-Hernandez, M.L., Ramirez-Rangel, I., *et al.*, 2008. G protein-coupled receptor-promoted trafficking of Gbeta1gamma2 leads to AKT

- activation at endosomes via a mechanism mediated by Gbeta1gamma2-Rab11a interaction. *Molecular Biology of the Cell* 19 (10), 4188–4200.
- Goh, L.K., Sorkin, A., 2013. Endocytosis of receptor tyrosine kinases. *Cold Spring Harbor Perspectives in Biology* 5 (5), a017459. doi:10.1101/cshperspect.a017459.
- Grimes, M.L., Zhou, J., Beattie, E.C., *et al.*, 1996. Endocytosis of activated TrkA: Evidence that nerve growth factor induces formation of signaling endosomes. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 16 (24), 7950–7964.
- Harrington St, A.W., Hillaire, C., Zweifel, L.S., *et al.*, 2011. Recruitment of actin modifiers to TrkA endosomes governs retrograde NGF signaling and survival. *Cell* 146 (3), 421–434.
- Haugh, J.M., Meyer, T., 2002. Active EGF receptors have limited access to PtdIns (4,5)P(2) in endosomes: Implications for phospholipase C and PI 3-kinase signaling. *Journal of Cell Science* 115 (Pt 2), 303–310.
- Hayes, S., Chawla, A., Corvera, S., 2002. TGF beta receptor internalization into EEA1-enriched early endosomes: Role in signaling to Smad2. *Journal of Cell Biology* 158 (7), 1239–1249.
- Honda, K., Yanai, H., Takaoka, A., Taniguchi, T., 2005. Regulation of the type I IFN induction: A current view. *International Immunology* 17 (11), 1367–1378.
- Howe, C.L., 2005. Modeling the signaling endosome hypothesis: Why a drive to the nucleus is better than a (random) walk. *Theoretical Biology & Medical Modeling* 2, 43.
- Howe, C.L., Mobley, W.C., 2005. Long-distance retrograde neurotrophic signaling. *Current opinion in Neurobiology* 15 (1), 40–48.
- Hussain, N.K., Jenna, S., Glogauer, M., *et al.*, 2001. Endocytic protein intersectin-1 regulates actin assembly via Cdc42 and N-WASP. *Nature Cell Biology* 3 (10), 927–932.
- Irie, F., Okuno, M., Pasquale, E.B., Yamaguchi, Y., 2005. EphrinB-EphB signalling regulates clathrin-mediated endocytosis through tyrosine phosphorylation of synaptojanin 1. *Nature Cell Biology* 7 (5), 501–509.
- Jekely, G., Sung, H.H., Luque, C.M., Rorth, P., 2005. Regulators of endocytosis maintain localized receptor tyrosine kinase signaling in guided migration. *Developmental Cell* 9 (2), 197–207.
- Joffre, C., Barrow, R., Menard, L., *et al.*, 2011. A direct role for Met endocytosis in tumorigenesis. *Nature Cell Biology* 13 (7), 827–837.
- Kagan, J.C., Su, T., Horng, T., *et al.*, 2008. TRAM couples endocytosis of Toll-like receptor 4 to the induction of interferon-beta. *Nature Immunology* 9 (4), 361–368.
- Kermorgant, S., Parker, P.J., 2008. Receptor trafficking controls weak signal delivery: A strategy used by c-Met for STAT3 nuclear accumulation. *Journal of Cell Biology* 182 (5), 855–863.
- Kessels, M.M., Engqvist-Goldstein, A.E., Drubin, D.G., Qualmann, B., 2001. Mammalian Abp1, a signal-responsive F-actin-binding protein, links the actin cytoskeleton to endocytosis via the GTPase dynamin. *Journal of Cell Biology* 153 (2), 351–366.
- Lanahan, A.A., Hermans, K., Claes, F., *et al.*, 2010. VEGF receptor 2 endocytic trafficking regulates arterial morphogenesis. *Developmental Cell* 18 (5), 713–724.
- Lanzetti, L., Di Fiore, P.P., 2008. Endocytosis and cancer: An 'insider' network with dangerous liaisons. *Traffic (Copenhagen, Denmark)* 9 (12), 2011–2021.
- Le Clainche, C., Pauly, B.S., Zhang, C.X., *et al.*, 2007. A Hip1R-cortactin complex negatively regulates actin assembly associated with endocytosis. *EMBO Journal* 26 (5), 1199–1210.
- Lefkowitz, R.J., Shenoy, S.K., 2005. Transduction of receptor signals by beta-arrestins. *Science (New York, NY)* 308 (5721), 512–517.
- Li, H.S., Stolz, D.B., Romero, G., 2005. Characterization of endocytic vesicles using magnetic microbeads coated with signalling ligands. *Traffic (Copenhagen, Denmark)* 6 (4), 324–334.
- Liao, W., Xiao, Q., Tchikov, V., *et al.*, 2008. CARP-2 is an endosome-associated ubiquitin ligase for RIP and regulates TNF-induced NF-kappaB activation. *Current Biology* 18 (9), 641–649.
- Logan, C.Y., Nusse, R., 2004. The Wnt signaling pathway in development and disease. *Annual Review of Cell and Developmental Biology* 20, 781–810.
- Lu, A., Tebar, F., Alvarez-Moya, B., *et al.*, 2009. A clathrin-dependent pathway leads to KRas signaling on late endosomes en route to lysosomes. *Journal of Cell Biology* 184 (6), 863–879.
- Mattila, E., Pellinen, T., Nevo, J., *et al.*, 2005. Negative regulation of EGFR signalling through integrin-alpha1beta1-mediated activation of protein tyrosine phosphatase TCPTP. *Nature Cell Biology* 7 (1), 78–85.
- McNiven, M.A., Kim, L., Krueger, E.W., *et al.*, 2000. Regulated interactions between dynamin and the actin-binding protein cortactin modulate cell shape. *Journal of Cell Biology* 151 (1), 187–198.
- Merrifield, C.J., Perrais, D., Zenisek, D., 2005. Coupling between clathrin-coated-pit invagination, cortactin recruitment, and membrane scission observed in live cells. *Cell* 121 (4), 593–606.
- Moore, C.A., Milano, S.K., Benovic, J.L., 2007. Regulation of receptor trafficking by GRKs and arrestins. *Annual Review of Physiology* 69, 451–482.
- Mosesson, Y., Mills, G.B., Yarden, Y., 2008. Derailed endocytosis: An emerging feature of cancer. *Nature Reviews Cancer* 8 (11), 835–850.
- Nada, S., Hondo, A., Kasai, A., *et al.*, 2009. The novel lipid raft adapter p18 controls endosome dynamics by anchoring the MEK-ERK pathway to late endosomes. *EMBO Journal* 28 (5), 477–489.
- Nakayama, A., Nakayama, M., Turner, C.J., *et al.*, 2013. Ephrin-B2 controls PDGFRbeta internalization and signaling. *Genes & Development* 27 (23), 2576–2589.
- Nazarewicz, R.R., Salazar, G., Patrushev, N., *et al.*, 2011. Early endosomal antigen 1 (EEA1) is an obligate scaffold for angiotensin II-induced, PKC-alpha-dependent Akt activation in endosomes. *Journal of Biological Chemistry* 286 (4), 2886–2895.
- Nevo, J., Mattila, E., Pellinen, T., *et al.*, 2009. Mammary-derived growth inhibitor alters traffic of EGFR and induces a novel form of cetuximab resistance. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 15 (21), 6570–6581.
- Newton, C.S., Loukinova, E., Mikhailenko, I., *et al.*, 2005. Platelet-derived growth factor receptor-beta (PDGFR-beta) activation promotes its association with the low density lipoprotein receptor-related protein (LRP). Evidence for co-receptor function. *Journal of Biological Chemistry* 280 (30), 27872–27878.
- Ouyang, X., Gulliford, T., Huang, G., Epstein, R.J., 1999. Transforming growth factor-alpha short-circuits downregulation of the epidermal growth factor receptor. *Journal of Cellular Physiology* 179 (1), 52–57.
- Palamidessi, A., Frittoli, E., Garre, M., *et al.*, 2008. Endocytic trafficking of Rac is required for the spatial restriction of signaling in cell migration. *Cell* 134 (1), 135–147.
- Pasternak, S.H., Bagshaw, R.D., Guiral, M., *et al.*, 2003. Presenilin-1, nicastrin, amyloid precursor protein, and gamma-secretase activity are co-localized in the lysosomal membrane. *Journal of Biological Chemistry* 278 (29), 26687–26694.
- Pawson, T., 2007. Dynamic control of signaling by modular adaptor proteins. *Current Opinion in Cell Biology* 19 (2), 112–116.
- Pei, Z., Baker, N.E., 2008. Competition between Delta and the Abruptex domain of Notch. *BMC Developmental Biology* 8, 4. doi:10.1186/1471-213X-8-4.
- Pelkmans, L., Zerial, M., 2005. Kinase-regulated quantal assemblies and kiss-and-run recycling of caveolae. *Nature* 436 (7047), 128–133.
- Puthenveedu, M.A., Lauffer, B., Temkin, P., *et al.*, 2010. Sequence-dependent sorting of recycling proteins by actin-stabilized endosomal microdomains. *Cell* 143 (5), 761–773.
- Rajan, A., Tien, A.C., Haueter, C.M., Schulze, K.L., Bellen, H.J., 2009. The Arp2/3 complex and WASp are required for apical trafficking of Delta into microvilli during cell fate specification of sensory organ precursors. *Nature Cell Biology* 11 (7), 815–824.
- Rink, J., Ghigo, E., Kalaidzidis, Y., Zerial, M., 2005. Rab conversion as a mechanism of progression from early to late endosomes. *Cell* 122 (5), 735–749.
- Ross, J.L., Ali, M.Y., Warshaw, D.M., 2008. Cargo transport: Molecular motors navigate a complex cytoskeleton. *Current Opinion in Cell Biology* 20 (1), 41–47.
- Schenck, A., Goto-Silva, L., Collinet, C., *et al.*, 2008. The endosomal protein Ap1 mediates Akt substrate specificity and cell survival in vertebrate development. *Cell* 133 (3), 486–497.
- Schneider-Brachert, W., Tchikov, V., Neumeyer, J., *et al.*, 2004. Compartmentalization of TNF receptor 1 signaling: Internalized TNF receptosomes as death signaling vesicles. *Immunity* 21 (3), 415–428.
- Schutze, S., Tchikov, V., Schneider-Brachert, W., 2008. Regulation of TNFR1 and CD95 signalling by receptor compartmentalization. *Nature Reviews Molecular Cell Biology* 9 (8), 655–662.
- Scita, G., Di Fiore, P.P., 2010. The endocytic matrix. *Nature* 463 (7280), 464–473.
- Seto, E.S., Bellen, H.J., 2006. Internalization is required for proper Wingless signaling in *Drosophila melanogaster*. *Journal of Cell Biology* 173 (1), 95–106.
- Seugnet, L., Simpson, P., Haenlin, M., 1997. Requirement for dynamin during Notch signaling in *Drosophila* neurogenesis. *Developmental Biology* 192 (2), 585–598.
- Sigsmund, S., Argenzio, E., Tosoni, D., *et al.*, 2008. Clathrin-mediated internalization is essential for sustained EGFR signaling but dispensable for degradation. *Developmental Cell* 15 (2), 209–219.
- Sigsmund, S., Confalonieri, S., Ciliberto, A., *et al.*, 2012. Endocytosis and signaling: Cell logistics shape the eukaryotic cell plan. *Physiological Reviews* 92 (1), 273–366.
- Sorkin, A., Eriksson, A., Heldin, C.H., Westermark, B., Claesson-Welsh, L., 1993. Pool of ligand-bound platelet-derived growth factor beta-receptors remain activated and tyrosine phosphorylated after internalization. *Journal of Cellular Physiology* 156 (2), 373–382.

- Sorkin, A., von Zastrow, M., 2002. Signal transduction and endocytosis: Close encounters of many kinds. *Nature Reviews Molecular Cell Biology* 3 (8), 600–614.
- Sorkin, A., von Zastrow, M., 2009. Endocytosis and signalling: Intertwining molecular networks. *Nature Reviews Molecular Cell Biology* 10 (9), 609–622.
- Stauffer, D.R., Howard, T.L., Nyun, T., Hollenberg, S.M., 2001. CHMP1 is a novel nuclear matrix protein affecting chromatin structure and cell-cycle progression. *Journal of Cell Science* 114 (Pt 13), 2383–2393.
- Teis, D., Taub, N., Kurzbauer, R., *et al.*, 2006. p14-MP1-MEK1 signaling regulates endosomal traffic and cellular proliferation during tissue homeostasis. *Journal of Cell Biology* 175 (6), 861–868.
- Teis, D., Wunderlich, W., Huber, L.A., 2002. Localization of the MP1-MAPK scaffold complex to endosomes is mediated by p14 and required for signal transduction. *Developmental Cell* 3 (6), 803–814.
- Tsukazaki, T., Chiang, T.A., Davison, A.F., Attisano, L., Wrana, J.L., 1998. SARA, a FYVE domain protein that recruits Smad2 to the TGFbeta receptor. *Cell* 95 (6), 779–791.
- Vaccari, T., Lu, H., Kanwar, R., Fortini, M.E., Bilder, D., 2008. Endosomal entry regulates Notch receptor activation in *Drosophila melanogaster*. *Journal of Cell Biology* 180 (4), 755–762.
- Vieira, A.V., Lamaze, C., Schmid, S.L., 1996. Control of EGF receptor signaling by clathrin-mediated endocytosis. *Science (New York, NY)* 274 (5295), 2086–2089.
- Wang, Y., Pennock, S.D., Chen, X., Kazlauskas, A., Wang, Z., 2004. Platelet-derived growth factor receptor-mediated signal transduction from endosomes. *Journal of Biological Chemistry* 279 (9), 8038–8046.
- Wang, Y., Roche, O., Yan, M.S., *et al.*, 2009. Regulation of endocytosis via the oxygen-sensing pathway. *Nature Medicine* 15 (3), 319–324.
- Waterman, H., Sabanai, I., Geiger, B., Yarden, Y., 1998. Alternative intracellular routing of ErbB receptors may determine signaling potency. *Journal of Biological Chemistry* 273 (22), 13819–13827.
- Wilde, A., Beattie, E.C., Lem, L., *et al.*, 1999. EGF receptor signaling stimulates SRC kinase phosphorylation of clathrin, influencing clathrin redistribution and EGF uptake. *Cell* 96 (5), 677–687.
- Windler, S.L., Bilder, D., 2010. Endocytic internalization routes required for delta/notch signaling. *Current Biology: CB* 20 (6), 538–543.
- Wu, C., Lai, C.F., Mobley, W.C., 2001. Nerve growth factor activates persistent Rap1 signaling in endosomes. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 21 (15), 5406–5416.
- Yamamoto, H., Sakane, H., Yamamoto, H., Michiue, T., Kikuchi, A., 2008. Wnt3a and Dkk1 regulate distinct internalization pathways of LRP6 to tune the activation of beta-catenin signaling. *Developmental Cell* 15 (1), 37–48.
- Yarar, D., Waterman-Storer, C.M., Schmid, S.L., 2005. A dynamic actin cytoskeleton functions at multiple stages of clathrin-mediated endocytosis. *Molecular Biology of the Cell* 16 (2), 964–975.
- York, R.D., Molliver, D.C., Grewal, S.S., *et al.*, 2000. Role of phosphoinositide 3-kinase and endocytosis in nerve growth factor-induced extracellular signal-regulated kinase activation via Ras and Rap1. *Molecular and Cellular Biology* 20 (21), 8069–8083.
- York, R.D., Yao, H., Dillon, T., *et al.*, 1998. Rap1 mediates sustained MAP kinase activation induced by nerve growth factor. *Nature* 392 (6676), 622–626.
- Yu, A., Rual, J.F., Tamai, K., *et al.*, 2007. Association of Dishevelled with the clathrin AP-2 adapter is required for Frizzled endocytosis and planar cell polarity signaling. *Developmental Cell* 12 (1), 129–141.
- Yu, X., Riley, T., Levine, A.J., 2009. The regulation of the endosomal compartment by p53 the tumor suppressor gene. *FEBS Journal* 276 (8), 2201–2212.
- von Zastrow, M., Sorkin, A., 2007. Signaling on the endocytic pathway. *Current Opinion in Cell Biology* 19 (4), 436–445.
- Zhong, H., De Marzo, A.M., Laughner, E., *et al.*, 1999. Overexpression of hypoxia-inducible factor 1alpha in common human cancers and their metastases. *Cancer Research* 59 (22), 5830–5835.
- Zoncu, R., Perera, R.M., Balkin, D.M., *et al.*, 2009. A phosphoinositide switch controls the maturation and signaling properties of APPL endosomes. *Cell* 136 (6), 1110–1121.
- Zwang, Y., Yarden, Y., 2006. p38 MAP kinase mediates stress-induced internalization of EGFR: Implications for cancer chemotherapy. *EMBO Journal* 25 (18), 4195–4206.
- Zweifel, L.S., Kuruvilla, R., Ginty, D.D., 2005. Functions and mechanisms of retrograde neurotrophin signalling. *Nature Reviews Neuroscience* 6 (8), 615–625.

Conventional and Secretory Lysosomes

R Shrestha, J Kaplan, and DM Ward, University of Utah School of Medicine, Salt Lake City, UT, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Lysosomes were first described by Dr. Christian de Duve in 1955 as the major degradative compartment within eukaryotic cells (Appelmans *et al.*, 1955; de Duve, 1959). Structurally they appear as dense bodies that often are found associated with the microtubule-organizing center (MTOC) near the nucleus of the cell. They can vary in size, shape, and number depending upon the cell type. Recently, the quantity of lysosomes was found to be regulated by the transcription factor EB (TFEB) (Sardiello *et al.*, 2009). TFEB recognizes a DNA sequence (GTCACGTGAC) termed coordinated lysosomal expression and regulation or CLEAR sequence found in many genes that encode for lysosomal proteins (Settembre *et al.*, 2013). TFEB controls the biogenesis of lysosomal hydrolases and membrane proteins. It exists in the cytosol and under altered lysosomal conditions such as nutrient starvation or hyperaccumulation of substrates, TFEB translocates to the nucleus to activate lysosome biogenesis gene transcription, thus generating increased numbers of lysosomes. Lysosomes possess an acidic lumen containing many types of hydrolases responsible for the degradation of substrates. The high H⁺ content is driven by the presence of the vacuolar ATPase H⁺ pump. Lysosomes also contain high levels of calcium (Ghislat and Knecht, 2013), second in amount compared to the ER. The lysosomal membrane contains highly glycosylated proteins including lysosome-associated membrane proteins (LAMPs) and lysosomal integral membrane proteins (LIMPs). This high level of glycosylation is thought to protect the proteins from the degradative enzymes stored in lysosomes. Cytosolic facing lysosomal proteins are involved in regulating the transport of material to and from the lysosome (Safitig and Klumperman, 2009; Luzio *et al.*, 2014). Transport to the lysosome is primarily achieved through vesicular delivery of endocytosed or phagocytosed material. Along this delivery pathway to the lysosome lipids and proteins can be recycled from endosomal compartments or they can be targeted to lysosomes for degradation. Lysosomes can also obtain cytosolic components such as proteins, lipids, and organelles destined for degradation through regulated autophagy (Schneider and Cuervo, 2014; Parzych and Klionsky, 2014). The digestive nature of lysosomes allows for the reuse of amino acids and lipids as building 'material' for generating new nucleic acids, proteins, lipids, and organelles.

Along with their degradative functions, specialized lysosomes termed lysosome-related organelles (LROs) or secretory lysosomes can be found in most cells types. Examples of secretory lysosomes include melanosomes, lytic granules, platelet dense granules, and major histocompatibility complex class II (MHCII) compartments. They share many properties of lysosomes but also have cell-specific functions such as pigmentation, host defense, coagulation, and innate and adaptive immunity. This review will focus on the characteristics of lysosomes and cell-specific secretory lysosomes, their biogenesis

and functions including an emphasis on the importance of conventional and secretory lysosomes in human disease.

Conventional Lysosomes

All eukaryotes possess organelles that are ascribed to have lysosomal functions, the storage, and degradation of molecules. In plants and fungi such as the model organism *Saccharomyces cerevisiae*, lysosomes are called vacuoles. Much of what is known about lysosome biogenesis and delivery to lysosomes can be attributed to the exquisite genetic studies done in yeast (Schekman, 1985; Rothman *et al.*, 1989; Klionsky *et al.*, 1990; Bryant and Stevens, 1998; Wickner, 2010; Wickner and Haas, 2000; Efe *et al.*, 2005). Historically, the lysosomal compartment has been considered to be an endpoint organelle and many reviews on the endocytic system still depict this organelle as an endpoint. Over the past 20 years it has been appreciated that this organelle is not the endpoint of the endocytic pathway but rather is another membrane compartment where sorting occurs, pathogens can be degraded or persist, signaling can be initiated and energy metabolism can be assessed (Luzio *et al.*, 2014; Borlace *et al.*, 2011; Barrias *et al.*, 2013; Baxt *et al.*, 2013; Settembre *et al.*, 2013). Lysosomes and secretory lysosomes can be distinguished from late endosomes by the absence of mannose-6-phosphate receptors (M6PR) (Brown *et al.*, 1986), receptors responsible for the delivery of lysosomal hydrolases to the developing lysosome. The delivery of molecules to lysosomes can be mediated either by transport vesicles from the trans-Golgi network (TGN) fusing with late endosomes followed by the recycling of components required for further rounds of delivery from the TGN or by direct delivery of LAMPs from the TGN to lysosomes through an adaptor protein (AP). This AP-mediated delivery provides a mechanism for coated vesicles to sort cargo specific for lysosomal delivery. The delivery of molecules from late endosomes to lysosome was coined 'kiss-and-run,' in which late endosomes briefly fuse with lysosomes forming a hybrid organelle that fissions reforming lysosomes and recycling late endocytic vesicles (Bright *et al.*, 1997; Luzio *et al.*, 2007; Luzio *et al.*, 2005; Luzio *et al.*, 2000; Mullock *et al.*, 1998). Biochemical and microscopic evidence has also shown that lysosomes are capable of fusing with other lysosomes (homotypic fusion) and they can subsequently fission (Patterson and Lippincott-Schwartz, 2002; Ward *et al.*, 1997; Ward *et al.*, 2000; Durchfort *et al.*, 2012; Perou *et al.*, 1997; Michailat *et al.*, 2012; Michailat and Mayer, 2013; Peters *et al.*, 2004; Zieger and Mayer, 2012; Schulze *et al.*, 2013). This dynamic nature allows for the redistribution of substrates among lysosomes for efficient degradation and regulates the size of the lysosomes.

Cell Type-Specific Secretory Lysosomes

Secretory lysosomes are a subset of organelles that are indistinguishable from lysosomes, or they can be a distinct subset

of organelles that are present in cells that also contain 'traditional' lysosomes. The function of these LROs varies depending upon the cell type. Most cell types contain secretory lysosomes involved in plasma membrane repair (Jaiswal *et al.*, 2002; Bansal *et al.*, 2003; Keefe *et al.*, 2005; Zhang *et al.*, 2007; Laulagnier *et al.*, 2011; Xu *et al.*, 2012a; McNeil, 2002; McNeil and Kirchhausen, 2005). Plasma membrane injury can occur due to mechanical stress (e.g., muscle, skin, endothelial cells, and adipocytes) or by pathogens trying to gain access to host cells (e.g., pore-forming toxins generated by bacteria). Scientists have taken advantage of plasma membrane repair mechanisms to introduce reagents (DNA, immunoglobulins, and drugs) into cells by electroporation, glass-bead disruption, or scrape-loading methods. The repair of the plasma membrane occurs within seconds/minutes and is essential for survival. This process involves the influx of calcium, which binds to the phospholipid-binding protein synaptotagmin VII (Andrews *et al.*, 2014), triggering interaction with soluble N-ethylmaleimide-sensitive factor (NSF) attachment protein receptors (SNAREs) such as SNAP-23, which are necessary for membrane fusion and cytoskeletal rearrangement. The fusion of lysosomes at the wound site results in lysosomal hydrolases being released extracellularly and the presence of lysosomal membrane proteins including LAMPs, LIMPs and the vacuolar H^+ -ATPase localized at the cell surface. A lysosomal enzyme, acid sphingomyelinase, has been shown to be important for the restoration of plasma membrane integrity (Tam *et al.*, 2010). Data suggest that the transient plasma membrane localization of the vacuolar H^+ -ATPase may provide the necessary pH conditions for acid sphingomyelinase to function at the outer leaflet of the plasma membrane generating ceramide (Xu *et al.*, 2012b), which produces an inward bulging of the plasma membrane and the dynamin-independent internalization of the wound repair site (Idone *et al.*, 2008; Lariccia *et al.*, 2011; Sinha *et al.*, 2011; Jiang and Chen, 2009). The endocytosis of the wound patch allows for the recovery of the lysosomal membrane proteins. Defects in plasma membrane repair are associated with a broad range of human diseases including neurodegeneration, muscular dystrophy, lysosomal storage diseases (LSDs), heart disease, skin disorders, and obesity.

Hematopoietic-Lineage Secretory Lysosomes

Distinct subsets of secretory lysosomes are found in the cells that belong to hematopoietic lineages such as lymphocytes, granulocytes (eosinophils, neutrophils, basophils, mast cells), platelets, and macrophages. These vesicles contain cell type-specific molecules that are involved in disease-related processes including pathogen killing, lysing virally or bacterially infected cells and lysing tumourigenic cells. The secretory lysosomes are also involved in homeostatic processes including promoting coagulation and antigen presentation in innate and adaptive immune responses.

In cytotoxic T lymphocytes (CTLs), the LROs or lytic granules are the only lysosomes present in the cells. They store acid hydrolases to degrade endocytosed material but also contain secretory molecules such as perforin and granzymes, which function at neutral pH to insert into target cell

membranes at the immunologic synapse forming a pore that can result in target cell killing (Holt *et al.*, 2006; Luzio *et al.*, 2007; Pattu *et al.*, 2013; Figure 1). Another important molecule contained in small vesicles within these secretory lysosomes is Fas ligand (FasL), a transmembrane protein that is released in small vesicles when the lysosomes fuse with the plasma membrane at the immunologic synapse (Zuccato *et al.*, 2007). The small vesicles then fuse with the target antigen-presenting cell that contains Fas and this recognition triggers apoptosis. Loss of FasL gives rise to autoimmune lymphoproliferative syndrome or ALPS1B (Wu *et al.*, 1996). The release of the lytic granule contents at the immunologic synapse provides a restricted 'barrier' to protect against unwanted killing of neighboring cells.

Some of the mechanisms for lytic granule release have been elucidated through the use of inhibitory molecules (e.g., microtubule depolymerizing agents such as nocodazole or nonhydrolyzable nucleotides (GTP γ S or ATP γ S) and through the identification of human mutations in genes encoding required components. Two such components are the ras-related small GTPase (Rab) Rab27a, which is required for docking of the lytic granule at the plasma membrane, and Munc13-4, which is an effector molecule of Rab27a that regulates fusion with the plasma membrane. Mutations in the genes that encode Rab27a (Menasche *et al.*, 2000; Desnos *et al.*, 2003; Menasche *et al.*, 2003) or Munc13-4 (Feldmann *et al.*, 2003; Yamamoto *et al.*, 2004; Garrett *et al.*, 2012) result in defective cytotoxic T cell killing and human diseases (Griscelli syndrome and hemophagocytic lymphohistiocytosis, respectively). Several SNARE molecules (proteins on cognate donor and acceptor membranes that act as addressing molecules) required for lytic granule exocytosis in CTLs include Vamp8 (Loo *et al.*, 2009), Vti1b (Dressel *et al.*, 2010), Syntaxin7 (Pattu *et al.*, 2011), Syntaxin11 (Bryceson *et al.*, 2007; Halimani *et al.*, 2014), Munc18-2, and Synaptobrevin2 (Matti *et al.*, 2013). Mutations in many of these genes in humans give rise to type 3-5 familial hemophagocytic lymphohistiocytosis (FHL), a life-threatening disease where the immune system produces too many activated immune cells that can damage major organs of the body (Faitelson and Grunebaum, 2014; Sieni *et al.*, 2014). Similar regulatory mechanisms for LRO secretion appear to exist in natural killer (NK) cells although NK cells also have nonsecretory lysosomes that function to degrade endocytosed material (van der Sluijs *et al.*, 2013).

Neutrophils contain primary and secondary granules (LROs) that assist in the uptake and killing of bacteria. During phagosome formation the cell undergoes a change in the localization of the MTOC toward the forming phagosome (Tapper *et al.*, 2002). The primary granules can fuse with the phagosomal membrane and release of granule content assists in bacterial killing. The contents of the granules are very specific. Primary or azurophilic granules contain proteases such as elastase, Cathepsin G, antimicrobial peptides, and myeloperoxidase. Secondary granules contain lactoferrin and lysozyme, which are directed, respectively, at limiting bacterial iron acquisition and lysis of pathogen cell walls.

Basophils and mast cells have secretory lysosomes that act in inflammation often associated with allergic reactions. They contain acid hydrolases along with various inflammatory mediators such as histamine, heparin, serotonin, and neutral

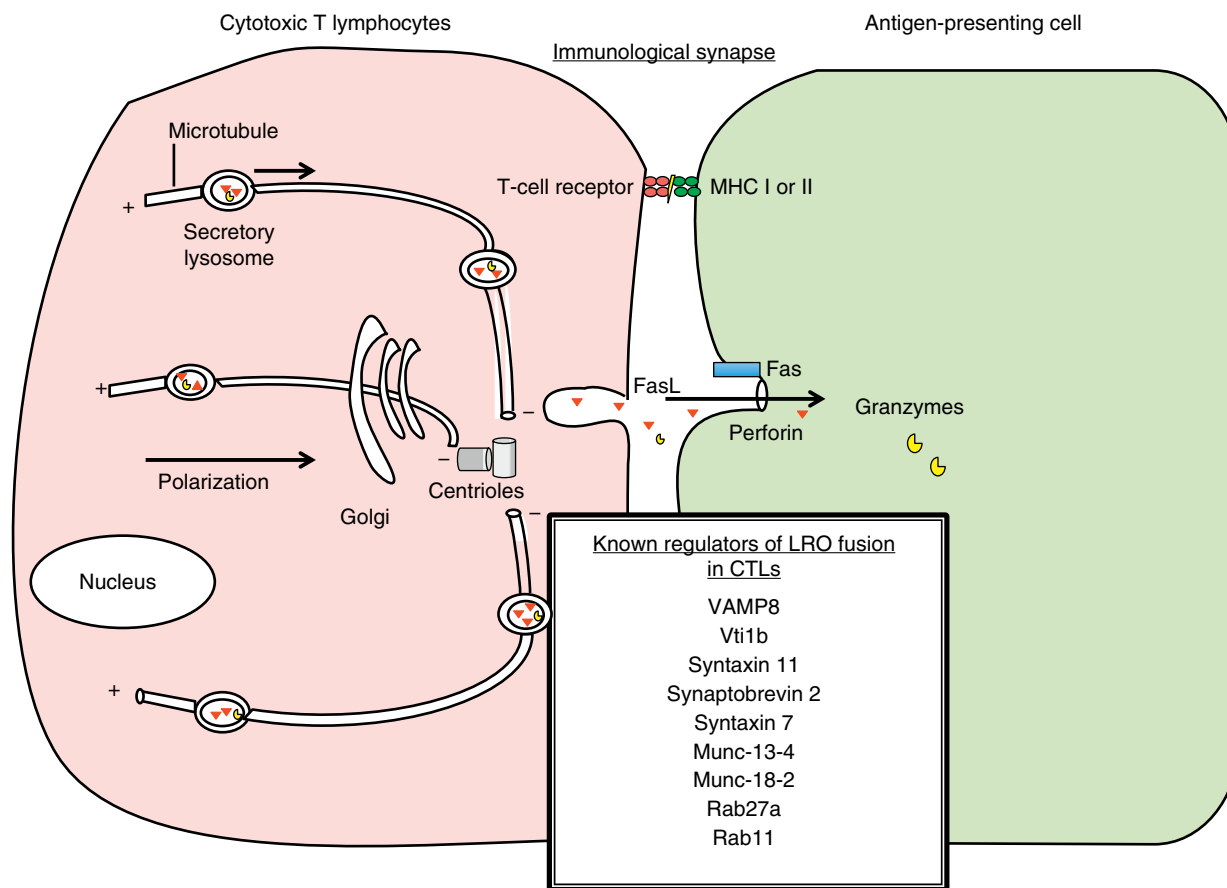


Figure 1 A model of the immunologic synapse in cytotoxic T lymphocytes (CTLs) secretory lysosome release. T-cell receptors on CTL recognizes antigen on antigen presenting cell in the context of MHC forming the immunologic synapse. This recognition sends a signal that causes the CTL centromere to polarize toward the immunologic synapse. Secretory lysosomes engage the microtubule machinery to move them toward the plasma membrane at the immunologic synapse. These lytic granules contain Fas ligand (FasL) that recognizes Fas being expressed on the antigen presenting cell. They also contain the pore-forming protein perforin and granzymes, which upon release at the synapse form a pore in the plasma membrane that induces apoptosis of the antigen presenting cell. Included is a list of known regulator of CTL secretory lysosome fusion.

proteases (Marone *et al.*, 1997). Eosinophilic secretory lysosomes are associated with host defense against parasitic helminthes, allergic reactions, and immunoregulation. The granules contain a variety of molecules including cationic proteins, growth factors, cytokines, receptors, and lytic molecules. The selective sorting of these molecules into eosinophil granules as well as their regulated release has identified different mechanisms that mediate release of granule contents (Muniz *et al.*, 2012). These mechanisms include: (1) compound exocytosis where the entire granule fuses with the plasma releasing its contents to kill target cells or parasites; (2) selective secretion of specific contents through tubular/vesicular structures that are released from the secretory lysosomes and fuse with the plasma membrane to release their contents; and (3) whole cell eosinophil lysis, which releases intact granules that may be directed to target cells or pathogens via membrane recognition molecules.

Macrophages contain conventional lysosomes as well as secretory lysosomes that are involved in membrane repair, pathogen killing, and immunomodulation. Specialized macrophages such as osteoclasts utilize these secretory

lysosomes to deliver acid hydrolases to sites of bone resorption (Teitelbaum, 2000). These secretory lysosomes contain a subset of acid hydrolases including cathepsin K and tartrate-resistant acid phosphatase or TRAP, two enzymes that play a role in bone resorption, which is critical for skeletal formation and remodeling (van Meel *et al.*, 2011). Some of the molecules that have been implicated in the release of secretory lysosomes in osteoclasts include Synaptotagmin VII (Zhao *et al.*, 2008), Syntaxin4, and Rab7 (Zhao *et al.*, 2001).

The megakaryocyte, the platelet-forming cell, contains secretory and conventional lysosomes. Platelets, which are only found in mammals, are formed by the shedding or breaking off of fragments of the megakaryocyte into the bloodstream. Platelets contain three kinds of secretory lysosomes: dense core granules containing clotting factors (fibrinogen, Von Willebrand Factor, and nucleotides), alpha granules containing vasoconstricting factors (serotonin and calcium), and lysosomes containing acid hydrolases and clot remodeling factors (plasminogen activator inhibitor-1 (PAI-1) and thrombin activatable fibrinolysis inhibitor (TAFI)) (Flaumenhaft *et al.*, 2009, Italiano *et al.*, 2008). These secretory lysosomes can also

release a number of chemokines and cytokines, which promote activation of leukocytes for immune response (Ren *et al.*, 2008). Secretory lysosomes must be moved to the periphery of the megakaryocyte prior to shedding of platelets. A failure to appropriately move vesicles will result in coagulation defects, which is a common phenotype associated with human diseases such as Chediak–Higashi syndrome (CHS), Hermansky–Pudlak syndrome (HPS), Gray Platelet syndrome (GPS), and Griscelli's syndrome (GS). Some details of these human diseases associated with secretory lysosome defects will be described later in this article. The basic machinery that moves and fuses vesicles with the plasma membrane is conserved for all secretory lysosome release, however, there are cell type-specific molecules that mediate this process. Molecules required for platelet dense granule release have been characterized through human disease gene identification, mouse knockouts and platelet dense granule proteomics (Lemons *et al.*, 1997; Flaumenhaft *et al.*, 1999; Chen *et al.*, 2000; Lemons *et al.*, 2000; Reed, 2004; Ren *et al.*, 2008). More recently syntaxin-binding protein 5 was shown to bind to core secretion machinery regulating when and where SNARE interactions occur, and the knockout mouse showed defective hemostasis and bleeding tendencies (Ye *et al.*, 2014). The general mechanisms of secretory lysosome fusion are discussed in detail below.

Nonhematologic Cell Secretory Lysosomes

While the specialization of secretory lysosomes is most pronounced in hematopoietic cell lineages, all cells have secretory lysosomes. The endothelial cell contains specialized secretory lysosomes termed Weibel–Palade bodies (Weibel and Palade, 1964), which play a key role in wound repair and coagulation (Hannah *et al.*, 2002). These secretory vesicles contain von Willebrand factor that is important in blood clotting (Sakariassen *et al.*, 1979; Ewenstein *et al.*, 1987; Ruggeri, 1999), as well as the membrane protein P-selectin (Bonfanti *et al.*, 1989; McEver *et al.*, 1989), which recruits leukocytes that play a role in initiating inflammation (Dore *et al.*, 1993; Mayadas *et al.*, 1993; Nolte *et al.*, 1994).

Astrocytes, neurons, and pulmonary type II cells have also been shown to contain specialized secretory lysosomes (Zhang *et al.*, 2007; Tribl *et al.*, 2006; Weaver *et al.*, 2002) that function in ATP, pigment and surfactant storage and regulated release.

The discovery that melanosomes are secretory lysosomes in pigment cells has provided much of the basis for understanding basic LRO biogenesis and plasma membrane release. This work was initiated by the observation that there were human diseases associated with pigmentation defects and that affected individuals often showed defects in hemostasis and immune function. Along with the more detailed characterization of these human diseases, mouse geneticists noticed that several spontaneously occurring mouse strains possessed varying levels of coat color defects and those mice often showed defects in other secretory lysosome functions (Jackson, 1991). The combination of these two observations provided the tools to identify those genes responsible for pigmentation and the biochemical pathways regulating LRO formation, movement, and function. During their development, melanosomes contain the enzymes and scaffolding proteins along

with several ion channels necessary for the synthesis of complex pigments. These pigments provide photoprotection from ionizing radiation and regulate retinal function in the eye (Raposo and Marks, 2007). The biogenesis of these organelles follows a highly specialized endocytic trafficking pathway (Wu and Hammer, 2014), which is more specifically detailed in the article on LROs.

Biogenesis of Lysosomes and Secretory Lysosomes

The generation of secretory lysosomes follows the same pathway as lysosome biogenesis. The delivery of internalized membrane, through endocytosis provides a pathway for selective sorting. Endosomes proceed through a series of sorting events where receptors and other plasma membrane components are recycled (e.g., transferrin receptor) as the endosome is 'matured.' Cargo destined for degradation remains in the endosome. The maturing endosome invaginates to form the multivesicular body (MVB) through a series of specific protein recruitment events that are regulated by the endosomal sorting complexes required for transport (ESCRT)-dependent machinery. Ubiquitin 'marked' membrane proteins are sorted into intraluminal vesicles that will be delivered to the lysosome for degradation. Again, much of this work was initially done in *S. cerevisiae* (Babst *et al.*, 2002a; Babst *et al.*, 2002b; Katzmann *et al.*, 2001; Babst and Odorizzi, 2013; Henne *et al.*, 2013), but a great deal of the knowledge on the mechanisms involved in MVB formation has come about through the study of virus release, specifically human immunodeficiency virus (HIV) (Votteler and Sundquist, 2013) and cell abscission (Carlton and Martin-Serrano, 2007; Morita *et al.*, 2007). There are currently over 20 ESCRT proteins in mammals and their roles in MVB formation, sorting, and trafficking and other cellular process such as abscission and exosome release are still being defined (McCullough *et al.*, 2013; Schuh and Audhya, 2014).

Hydrolases and other soluble lysosomal components, synthesized in the ER and trafficked through the Golgi for additional glycosylation, can be delivered to the now late endosome through M6PR, which release their cargo upon exposure to the acidic environment of the forming lysosome (Fischer *et al.*, 1980; Kornfeld and Mellman, 1989; Braulke and Bonifacino, 2009; Saftig and Klumperman, 2009). Other delivery mechanisms independent of M6PR are known to exist, as hydrolases can be delivered to lysosomes in the absence of M6P glycosylation as in the human disorder I-cell disease due to a deficiency in N-acetylglucosamine-phosphotransferase (Coutinho *et al.*, 2012). Membrane proteins destined for residence in lysosomes contain targeting motifs, YXX ϕ or [DE] XXXL[LI], which are recognized by adaptor proteins (AP) that sort them into coated cargo vesicles for delivery with the late endocytic machinery (Park and Guo, 2014). This can occur in the biosynthetic pathway via AP1 or AP3 or via an indirect route to the plasma membrane where the adaptor protein AP2 assists in the recognition and endocytosis of lysosomal proteins. There is also evidence that lysosomal membrane proteins such as LAMP1 and LAMP2 can be delivered to the forming lysosome in an AP-independent M6PR-independent manner (Pols *et al.*, 2013).

The mechanisms for sorting to conventional lysosomes also provide for the sorting of molecules to secretory lysosomes. There are some specialized sorting mechanisms in CTL secretory lysosome biogenesis that involve ubiquitination and phosphorylation of FasL and trafficking through the MVB for delivery to LROs (Park and Guo, 2014). Consistent with cell type-specific secretory lysosome sorting is the trafficking of elastase into neutrophils through the tetraspanin molecule CD63 (Harrison-Lavoie *et al.*, 2006). The sorting of molecules into the forming melanosomes is highly specific for melanocytes and deviates from typical secretory lysosome formation by sorting from the early rather than late endosomal compartment and there are no internal vesicles (Marks *et al.*, 2013).

Exocytosis of Secretory Lysosomes

The regulated secretion of LROs has some general principles that apply to all cell types (Figure 2). The first step in secretion is signaling to the cell to release secretory lysosomes. This

signal can be the influx of calcium as in wound repair or it may be a receptor ligand signal as in CTLs or NK cells where the formation of the immunologic synapse (TCR-antigen-MHC) signals the polarization of the cell and the movement or localization of 'mature' secretory lysosomes near the plasma membrane, however, not all cells polarize to release secretory lysosome contents. Secretion requires binding of the secretory lysosome to microtubules through the kinesin family of motor proteins providing long distance movements toward the cell periphery. The secretory lysosome is then released and may bind to actin through myosin motors for more subtle movements close to the plasma membrane. Once the secretory lysosome is near the plasma membrane docking molecules on the secretory lysosome and plasma membrane tether the vesicle to the plasma membrane. Finally, SNARE pairing brings the secretory lysosome membrane in close proximity with the plasma membrane, where the SNARE pairing and Rab molecules can affect a fusion event, thus releasing soluble secretory lysosome content as well as delivering membrane content to the plasma membrane. The mechanisms that regulate granule

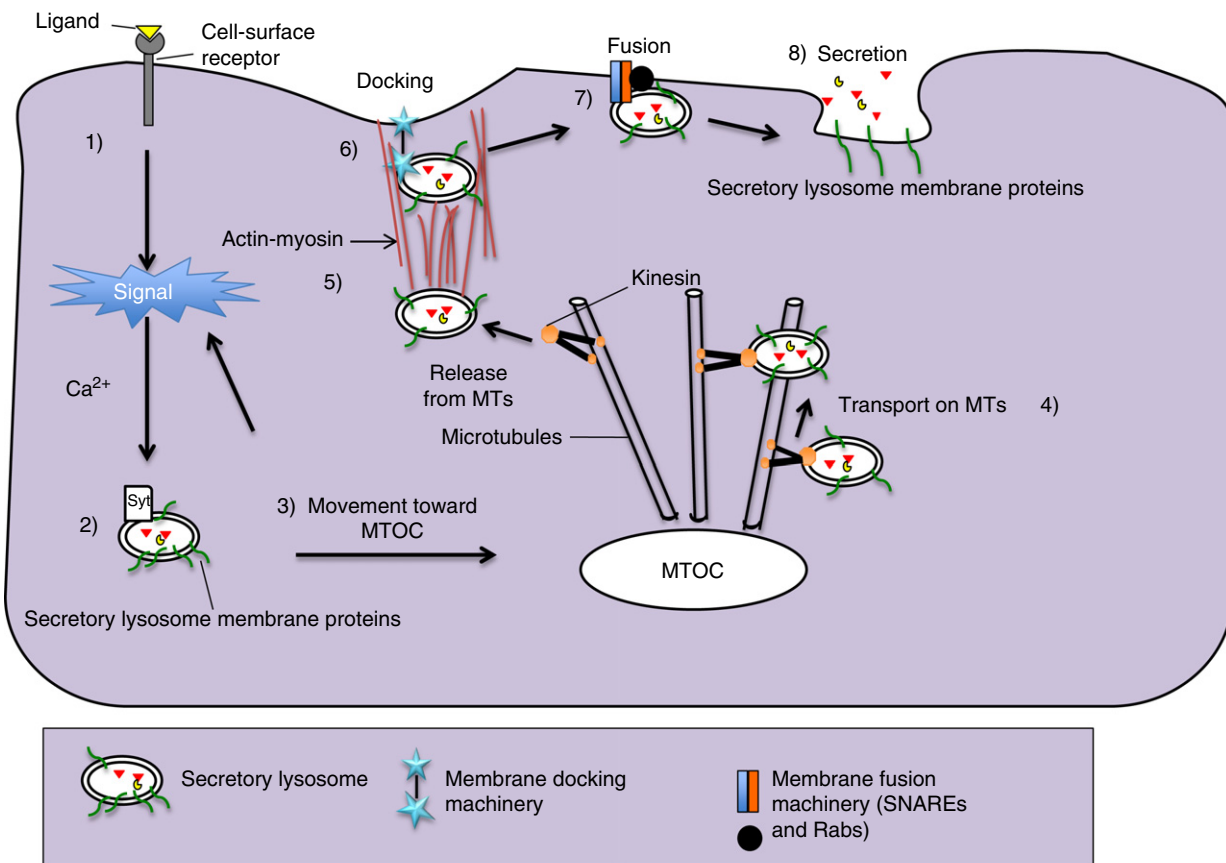


Figure 2 A model of secretory lysosome release. (1) A signal is transduced by receptor–ligand interaction, an intracellular signal or by wound formation at the plasma membrane triggering a calcium release from intracellular stores. (2) Synaptotagmin VII (Syn) binds to the secretory lysosome in response to calcium and (3) the vesicle moves toward the MTOC. (4) The secretory lysosome then engages the MT through the motor protein kinesin to move to the periphery of the cell where it is released. (5) For the short movement to the plasma membrane the secretory lysosome engages actin cytoskeletal elements through myosin motors. (6) When the vesicle is in close proximity to the plasma membrane, recognition membrane docking elements (stars) bind the docking elements on the vesicle. (7) The specific vesicular and plasma membrane v and t-SNAREs bind and bring the vesicle near the plasma membrane and fusion occurs (8) releasing soluble and membrane bound secretory lysosome content.

release have not been fully elucidated for each cell type, however, some calcium sensing molecules such as synaptotagmin VII and tethering molecules including synaptobrevin2, Munc13-4 and Munc18-2 along with SNARE molecules Vamp7, Vamp8, Vamp2, Syntaxin4, Syntaxin7, SNAP-23, have been suggested to have a role in cell-specific secretory lysosome release (Logan *et al.*, 2006; Tiwari *et al.*, 2008; Lacy *et al.*, 2001; Suzuki and Verma, 2008). Further, novel SNARE regulatory molecules have recently been identified for platelet granule release (Ye *et al.*, 2014).

Genetic Diseases Associate with Defects in Conventional or Secretory Lysosomes

There are many human diseases associated with defects in conventional lysosome or LRO function or biogenesis and patients have been identified and described for many decades. Some of the diseases associated with conventional lysosome defects include LSDs, which are classified according to the substrates they accumulate. These include mucopolipidosis, sphingolipidoses, oligosaccharidoses, lipoprotein storage disorders, mucopolysaccharidosis, and neuronal ceroid lipofuscinoses along with others, up to 50 have been described (Coutinho *et al.*, 2014; Samie and Xu, 2014). These disorders are most often due to mutations in lysosomally targeted enzymes that degrade their specific substrates or mutations in the proteins that regulate the delivery of these enzymes to the lysosome. The resulting phenotype commonly associated with LSDs is the abnormal accumulation of undegraded substrates. Examples of LSDs that are not due to mutations in the enzymes include mucopolipidosis type 4, where there is a defect in the cation channel mucolipin-1 and Niemann-Pick type C disease, which is due to mutation in either NPC1 or NPC2 (Ikonen and Holtta-Vuori, 2004; Vanier, 2010). It is unclear what ion is important in mucolipid degradation or what ion (s) mucolipin-1 transports (Dong *et al.*, 2008; Dong *et al.*, 2010; Eichelsdoerfer *et al.*, 2010). The exact roles of NPC1 and NPC2 in cholesterol metabolism have not been fully elucidated but they may play a role in sterol transport. Treatment for many of these LSDs is enzyme replacement therapy.

Other genetic diseases associated with the biogenesis, trafficking, or homeostasis of lysosomes and secretory lysosomes that are not due to substrate accumulation include: HPS, GS, CHS, GPS, and FHL. There are nine subtypes of HPS. Phenotypes associated with HPS include pigmentation defects of the hair, skin, and eyes, bleeding tendencies, neurologic disorders, and in some HPS cases pulmonary fibrosis. Mutations in HPS-associated genes encode for proteins that are involved in the biogenesis of secretory lysosomes (e.g., AP3, BLOCs, and geranylgeranyl transferases). As of yet there is no cure for HPS and there is a great deal of heterogeneity of HPS-associated defects (Seward and Gahl, 2013).

There are three types of GS, which can vary in phenotype with partial albinism being the common phenotype, but can include immunodeficiency or neurologic deficits. The genes associated with GS encode for proteins involved in the movement of LROs (e.g., Myosin-Va, Rab27a, and melanophilin). There is no cure for GS but the hematologic problems can be treated with bone marrow transplantation (Arico *et al.*, 2002; Cesaro *et al.*, 2008; Trottestam *et al.*, 2009).

CHS phenotypes include severe immune deficiency, hypopigmentation, bleeding tendency, frequent bacterial infection, oculocutaneous albinism, and progressive neurological dysfunction (Spritz, 1998). The hallmark characteristic of CHS is the presence of giant conventional and secretory lysosomes in all cells of the body (Introne *et al.*, 1999; Shiflett *et al.*, 2002; Stinchcombe *et al.*, 2004). The identification of the gene responsible for CHS was done through classic positional cloning of CHS patients and yeast artificial chromosome complementation using the *beige* mouse model (Perou *et al.*, 1996a; Perou *et al.*, 1996b; Barbosa *et al.*, 1996; Nagle *et al.*, 1996). The *CHS/LYST* gene encodes a protein of 3800 amino acids containing three to four defined domains or motif that have not informed on function (Kaplan *et al.*, 2008) although the WIDL and WD-40 repeat domains defined a family of protein called Beige and Chediak (BEACH) and all family members that have been characterized appear to be involved in membrane trafficking (Cullinane *et al.*, 2013). There are two competing hypotheses for the function of Lyst, one suggesting that Lyst functions as a negative regulator of fusion (Barrat *et al.*, 1999; Kypri *et al.*, 2007; Charette and Cosson, 2007; Rahman *et al.*, 2012; Kypri *et al.*, 2013) and the other suggesting that Lyst regulates lysosome fission (Perou *et al.*, 1997; Durchfort *et al.*, 2012). It is of interest to note that only CHS, as opposed to HPS, GS, or GPS, is associated with defects in conventional lysosomes as well as LROs suggesting that the Lyst protein regulates the maintenance or biogenesis of both types of lysosomes. Treatment for CHS is prophylactic antibiotics early in life and bone marrow transplant, which treats the life-threatening immunologic deficits but does not treat the neurologic defects.

GPS is a rare autosomal recessive bleeding disorder that is due to a reduction or the absence of alpha granules in platelets. The gene identified to be mutated in GPS is *NBEAL2*, which is a member of the BEACH family of proteins involved in vesicular trafficking (Kahr *et al.*, 2011). A mouse model of the disease demonstrated the importance of *Nbeal2* in hemostasis and thrombosis and tissue repair (Deppermann *et al.*, 2013).

FHL patients are immunodeficient but do not exhibit pigmentation defects. FHL is manifested in the first six months of life and is historically treated with bone marrow transplantation (Jabado *et al.*, 1997) or more recently with stem cell transplantation (Ohga *et al.*, 2010). It can be caused by mutations in the *PRF1* gene, which encodes for perforin, the protein that plays an important role in CTL-mediated killing or in the gene encoding Munc13-4 (Feldmann *et al.*, 2003). Munc13-4 protein is essential for the secretion of secretory vesicles from CTL but not from melanocytes or neuronal cells. FHL patient's secretory lysosomes are able to transport and dock to the plasma membrane but are not able to release their content in immunological synapse due to the mutation of Munc13-4. The loss of perforin-mediated cytotoxicity results in imbalanced immune homeostasis and deregulated activation and proliferation of CTLs during antiviral immune response (Lettau *et al.*, 2007).

Conclusions

The biosynthesis of lysosomes and secretory lysosomes involves multiple trafficking pathways including delivery of cargo to the forming lysosome from the endocytic and

biosynthetic pathways along with recycling of membrane constituents needed for further rounds of targeting to the lysosome. There is clear evidence that these organelles are capable of fusion and fission events and the machinery that regulates this process is being elucidated. One of the major limitations in understanding the biogenesis of lysosomes and secretory lysosomes is the ability to distinguish between late endosomes and lysosomes. Currently there are no clear markers of lysosomes, as late endosomes receive 'lysosomal' cargo to be delivered to the forming lysosome. Further, we do not understand many of the molecules that are involved in lysosome or secretory lysosome fission.

Many of the molecular components required for secretory lysosome biogenesis and secretion have been determined, however, it is clear that the detailed biochemical requirements for this process need to be better understood. That particular cell types have evolved mechanisms to generate and release secretory lysosomes further demonstrates the complexity of this process. The recent advances in live cell microscopy, protein tagging, and proteomics should provide some of the tools necessary to identify the roles of these and novel proteins and lipids involved in the biogenesis, maintenance, and secretion of lysosomes.

See also: Organelles: Structure and Function: Lysosome-Related Organelles

References

- Andrews, N.W., Almeida, P.E., Corrotte, M., 2014. Damage control: Cellular mechanisms of plasma membrane repair. *Trends in Cell Biology*. doi:10.1016/j.tcb.2014.07.008.
- Appelmans, F., Wattiaux, R., de Duve, C., 1955. Tissue fractionation studies. 5. The association of acid phosphatase with a special class of cytoplasmic granules in rat liver. *Biochemical Journal* 59, 438–445.
- Arico, M., Zecca, M., Santoro, N., et al., 2002. Successful treatment of Griscelli syndrome with unrelated donor allogeneic hematopoietic stem cell transplantation. *Bone Marrow Transplantation* 29, 995–998.
- Babst, M., Katzmann, D.J., Estepa-Sabal, E.J., Meerloo, T., Emr, S.D., 2002a. Escrt-III: An endosome-associated heterooligomeric protein complex required for mvb sorting. *Developmental Cell* 3, 271–282.
- Babst, M., Katzmann, D.J., Snyder, W.B., Wendland, B., Emr, S.D., 2002b. Endosome-associated complex, ESCRT-II, recruits transport machinery for protein sorting at the multivesicular body. *Developmental Cell* 3, 283–289.
- Babst, M., Odorizzi, G., 2013. The balance of protein expression and degradation: An ESCRTs point of view. *Current Opinion in Cell Biology* 25, 489–494.
- Bansal, D., Miyake, K., Vogel, S.S., et al., 2003. Defective membrane repair in dysferlin-deficient muscular dystrophy. *Nature* 423, 168–172.
- Barbosa, M.D., Nguyen, Q.A., Tchernev, V.T., et al., 1996. Identification of the homologous beige and Chediak-Higashi syndrome genes. *Nature* 382, 262–265.
- Barrat, F.J., Le Deist, F., Benkerrou, M., et al., 1999. Defective CTLA-4 cycling pathway in Chediak-Higashi syndrome: A possible mechanism for deregulation of T lymphocyte activation. *Proceedings of the National Academy of Sciences of the United States of America* 96, 8645–8650.
- Barrias, E.S., de Carvalho, T.M., de Souza, W., 2013. Trypanosoma cruzi: Entry into mammalian host cells and parasitophorous vacuole formation. *Frontiers in Immunology* 4, 186.
- Baxt, L.A., Garza-Mayers, A.C., Goldberg, M.B., 2013. Bacterial subversion of host innate immune pathways. *Science* 340, 697–701.
- Bonfanti, R., Furie, B.C., Furie, B., Wagner, D.D., 1989. PADGEM (GMP140) is a component of Weibel-Palade bodies of human endothelial cells. *Blood* 73, 1109–1112.
- Borlace, G.N., Jones, H.F., Keep, S.J., Butler, R.N., Brooks, D.A., 2011. Helicobacter pylori phagosome maturation in primary human macrophages. *Gut Pathogens* 3, 3.
- Braulke, T., Bonifacio, J.S., 2009. Sorting of lysosomal proteins. *Biochimica et Biophysica Acta* 1793, 605–614.
- Bright, N.A., Reaves, B.J., Mullock, B.M., Luzio, J.P., 1997. Dense core lysosomes can fuse with late endosomes and are re-formed from the resultant hybrid organelles. *Journal of Cell Science* 110 (Pt 17), 2027–2040.
- Brown, W.J., Goodhouse, J., Farquhar, M.G., 1986. Mannose-6-phosphate receptors for lysosomal enzymes cycle between the Golgi complex and endosomes. *Journal of Cell Biology* 103, 1235–1247.
- Bryant, N.J., Stevens, T.H., 1998. Vacuole biogenesis in *Saccharomyces cerevisiae*: Protein transport pathways to the yeast vacuole. *Microbiology and Molecular Biology Reviews* 62, 230–247.
- Bryceson, Y.T., Rudd, E., Zheng, C., et al., 2007. Defective cytotoxic lymphocyte degranulation in syntaxin-11 deficient familial hemophagocytic lymphohistiocytosis 4 (FHL4) patients. *Blood* 110, 1906–1915.
- Carlton, J.G., Martin-Serrano, J., 2007. Parallels between cytokinesis and retroviral budding: A role for the ESCRT machinery. *Science* 316, 1908–1912.
- Cesaro, S., Locatelli, F., Lanino, E., et al., 2008. Hematopoietic stem cell transplantation for hemophagocytic lymphohistiocytosis: A retrospective analysis of data from the Italian Association of Pediatric Hematology Oncology (AIEOP). *Haematologica* 93, 1694–1701.
- Charette, S.J., Cosson, P., 2007. A LYST/beige homolog is involved in biogenesis of Dictyostelium secretory lysosomes. *Journal of Cell Science* 120, 2338–2343.
- Chen, D., Lemons, P.P., Schraw, T., Whiteheart, S.W., 2000. Molecular mechanisms of platelet exocytosis: Role of SNAP-23 and syntaxin 2 and 4 in lysosome release. *Blood* 96, 1782–1788.
- Coutinho, M.F., Matos, L., Alves, S., 2014. From bedside to cell biology: A century of history on lysosomal dysfunction. *Gene*. doi:10.1016/j.gene.2014.09.054.
- Coutinho, M.F., Prata, M.J., Alves, S., 2012. A shortcut to the lysosome: The mannose-6-phosphate-independent pathway. *Molecular Genetics and Metabolism* 107, 257–266.
- Cullinane, A.R., Schaffer, A.A., Huizing, M., 2013. The BEACH is hot: A LYST of emerging roles for BEACH-domain containing proteins in human disease. *Traffic* 14, 749–766.
- de Duve, C., 1959. Lysosomes, a new group of cytoplasmic particles. In: Hayashi, T. (Ed.), *Subcellular Particles*. New York, NY: The Ronald Press Co., pp. 128–159.
- Deppermann, C., Cherpokova, D., Nurden, P., et al., 2013. Gray platelet syndrome and defective thrombo-inflammation in Nbeal2-deficient mice. *Journal of Clinical Investigation*. doi:10.1172/JCI69210.
- Desnos, C., Schonn, J.S., Huet, S., et al., 2003. Rab27A and its effector MyRIP link secretory granules to F-actin and control their motion towards release sites. *Journal of Cell Biology* 163, 559–570.
- Dong, X.P., Cheng, X., Mills, E., et al., 2008. The type IV mucopolidosis-associated protein TRPML1 is an endolysosomal iron release channel. *Nature* 455, 992–996.
- Dong, X.P., Shen, D., Wang, X., et al., 2010. PI(3,5)P2 controls membrane trafficking by direct activation of mucolipin Ca(2+) release channels in the endolysosome. *Nature Communications* 1, 38.
- Dore, M., Korthuis, R.J., Granger, D.N., Entman, M.L., Smith, C.W., 1993. P-selectin mediates spontaneous leukocyte rolling *in vivo*. *Blood* 82, 1308–1316.
- Dressel, R., Elsner, L., Novota, P., Kanwar, N., Fischer von Mollard, G., 2010. The exocytosis of lytic granules is impaired in Vti1b- or Vamp8-deficient CTL leading to a reduced cytotoxic activity following antigen-specific activation. *Journal of Immunology* 185, 1005–1014.
- Durchfort, N., Verhoef, S., Vaughn, M.B., et al., 2012. The enlarged lysosomes in beige j cells result from decreased lysosome fission and not increased lysosome fusion. *Traffic* 13, 108–119.
- Efe, J.A., Botelho, R.J., Emr, S.D., 2005. The Fab1 phosphatidylinositol kinase pathway in the regulation of vacuole morphology. *Current Opinion in Cell Biology* 17, 402–408.
- Eichelsdoerfer, J.L., Evans, J.A., Slaugenhaupt, S.A., Cuajungco, M.P., 2010. Zinc dyshomeostasis is linked with the loss of mucopolidosis IV-associated TRPML1 ion channel. *Journal of Biological Chemistry* 285, 34304–34308.
- Ewenstein, B.M., Warhol, M.J., Handin, R.L., Pober, J.S., 1987. Composition of the von Willebrand factor storage organelle (Weibel-Palade body) isolated from cultured human umbilical vein endothelial cells. *Journal of Cell Biology* 104, 1423–1433.
- Faitleson, Y., Grunebaum, E., 2014. Hemophagocytic lymphohistiocytosis and primary immune deficiency disorders. *Clinical Immunology* 155, 118–125.

- Feldmann, J., Callebaut, I., Raposo, G., *et al.*, 2003. Munc13-4 is essential for cytotytic granules fusion and is mutated in a form of familial hemophagocytic lymphohistiocytosis (FHL3). *Cell* 115, 461–473.
- Fischer, H.D., Gonzalez-Noriega, A., Sly, W.S., Morre, D.J., 1980. Phosphomannosyl-enzyme receptors in rat liver. Subcellular distribution and role in intracellular transport of lysosomal enzymes. *Journal of Biological Chemistry* 255, 9608–9615.
- Flaumenhaft, R., Croce, K., Chen, E., Furie, B., Furie, B.C., 1999. Proteins of the exocytotic core complex mediate platelet alpha-granule secretion. Roles of vesicle-associated membrane protein, SNAP-23, and syntaxin 4. *Journal of Biological Chemistry* 274, 2492–2501.
- Flaumenhaft, R., Dilks, J.R., Richardson, J., *et al.*, 2009. Megakaryocyte-derived microparticles: Direct visualization and distinction from platelet-derived microparticles. *Blood* 113, 1112–1121.
- Garrett, J.P., Fung, I., Rupon, J., *et al.*, 2012. Presentation of hemophagocytic lymphohistiocytosis due to a novel MUNC 13-4 mutation masked by partial therapeutic immunosuppression. *Pediatric Rheumatology Online Journal* 10, 13.
- Ghislat, G., Knecht, E., 2013. Ca²⁺(+)-sensor proteins in the autophagic and endocytic traffic. *Current Protein & Peptide Science* 14, 97–110.
- Halimani, M., Pattu, V., Marshall, M.R., *et al.*, 2014. Syntaxin11 serves as a t-SNARE for the fusion of lytic granules in human cytotoxic T lymphocytes. *European Journal of Immunology* 44, 573–584.
- Hannah, M.J., Williams, R., Kaur, J., Hewlett, L.J., Cutler, D.F., 2002. Biogenesis of Weibel-Palade bodies. *Seminars in Cell & Developmental Biology* 13, 313–324.
- Harrison-Lavoie, K.J., Michaux, G., Hewlett, L., *et al.*, 2006. P-selectin and CD63 use different mechanisms for delivery to Weibel-Palade bodies. *Traffic* 7, 647–662.
- Henne, W.M., Stenmark, H., Emr, S.D., 2013. Molecular mechanisms of the membrane sculpting ESCRT pathway. *Cold Spring Harbor Perspectives in Biology* 5, a016766.
- Holt, O.J., Gallo, F., Griffiths, G.M., 2006. Regulating secretory lysosomes. *Journal of Biochemistry* 140, 7–12.
- Idone, V., Tam, C., Goss, J.W., *et al.*, 2008. Repair of injured plasma membrane by rapid Ca²⁺-dependent endocytosis. *Journal of Cell Biology* 180, 905–914.
- Ikonen, E., Holta-Vuori, M., 2004. Cellular pathology of Niemann-Pick type C disease. *Seminars in Cell & Developmental Biology* 15, 445–454.
- Introne, W., Boissy, R.E., Gahl, W.A., 1999. Clinical, molecular, and cell biological aspects of Chediak-Higashi syndrome. *Molecular Genetics and Metabolism* 68, 283–303.
- Italiano, J.E., Richardson, J.L., Patel-Hett, S., *et al.*, 2008. Angiogenesis is regulated by a novel mechanism: Pro- and antiangiogenic proteins are organized into separate platelet alpha granules and differentially released. *Blood* 111, 1227–1233.
- Jabado, N., de Graeff-Meeder, E.R., Cavazzana-Calvo, M., *et al.*, 1997. Treatment of familial hemophagocytic lymphohistiocytosis with bone marrow transplantation from HLA genetically nonidentical donors. *Blood* 90, 4743–4748.
- Jackson, I.J., 1991. Mouse coat colour mutations: A molecular genetic resource which spans the centuries. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 13, 439–446.
- Jaiswal, J.K., Andrews, N.W., Simon, S.M., 2002. Membrane proximal lysosomes are the major vesicles responsible for calcium-dependent exocytosis in nonsecretory cells. *Journal of Cell Biology* 159, 625–635.
- Jiang, M., Chen, G., 2009. Ca²⁺ regulation of dynamin-independent endocytosis in cortical astrocytes. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29, 8063–8074.
- Kahr, W.H., Hinckley, J., Li, L., *et al.*, 2011. Mutations in NBEAL2, encoding a BEACH protein, cause gray platelet syndrome. *Nature Genetics* 43, 738–740.
- Kaplan, J., de Domenico, I., Ward, D.M., 2008. Chediak-Higashi syndrome. *Current Opinion in Hematology* 15, 22–29.
- Katzmann, D.J., Babst, M., Emr, S.D., 2001. Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex, ESCRT-I. *Cell* 106, 145–155.
- Keefe, D., Shi, L., Feske, S., *et al.*, 2005. Perforin triggers a plasma membrane-repair response that facilitates CTL induction of apoptosis. *Immunity* 23, 249–262.
- Klionsky, D.J., Herman, P.K., Emr, S.D., 1990. The fungal vacuole: Composition, function, and biogenesis. *Microbiological Reviews* 54, 266–292.
- Kornfeld, S., Mellman, I., 1989. The biogenesis of lysosomes. *Annual Review of Cell Biology* 5, 483–525.
- Kypri, E., Falkenstein, K., de Lozanne, A., 2013. Antagonistic control of lysosomal fusion by Rab14 and the Lyst-related protein LvsB. *Traffic* 14, 599–609.
- Kypri, E., Schmauch, C., Maniak, M., de Lozanne, A., 2007. The BEACH protein LvsB is localized on lysosomes and postlysosomes and limits their fusion with early endosomes. *Traffic* 8, 774–783.
- Lacy, P., Logan, M.R., Bablitz, B., Moqbel, R., 2001. Fusion protein vesicle-associated membrane protein 2 is implicated in IFN-gamma-induced piecemeal degranulation in human eosinophils from atopic individuals. *Journal of Allergy and Clinical Immunology* 107, 671–678.
- Lariccia, V., Fine, M., Magi, S., *et al.*, 2011. Massive calcium-activated endocytosis without involvement of classical endocytic proteins. *Journal of General Physiology* 137, 111–132.
- Laulagnier, K., Schieber, N.L., Maritzen, T., *et al.*, 2011. Role of AP1 and Gadkin in the traffic of secretory endo-lysosomes. *Molecular Biology of the Cell* 22, 2068–2082.
- Lemons, P.P., Chen, D., Bernstein, A.M., Bennett, M.K., Whiteheart, S.W., 1997. Regulated secretion in platelets: Identification of elements of the platelet exocytosis machinery. *Blood* 90, 1490–1500.
- Lemons, P.P., Chen, D., Whiteheart, S.W., 2000. Molecular mechanisms of platelet exocytosis: Requirements for alpha-granule release. *Biochemical and Biophysical Research Communications* 267, 875–880.
- Lettau, M., Schmidt, H., Kabelitz, D., Janssen, O., 2007. Secretory lysosomes and their cargo in T and NK cells. *Immunology Letters* 108, 10–19.
- Logan, M.R., Lacy, P., Odemuyiwa, S.O., *et al.*, 2006. A critical role for vesicle-associated membrane protein-7 in exocytosis from human eosinophils and neutrophils. *Allergy* 61, 777–784.
- Loo, L.S., Hwang, L.A., Ong, Y.M., *et al.*, 2009. A role for endobrevin/VAMP8 in CTL lytic granule exocytosis. *European Journal of Immunology* 39, 3520–3528.
- Luzio, J.P., Hackmann, Y., Dieckmann, N.M., Griffiths, G.M., 2014. The biogenesis of lysosomes and lysosome-related organelles. *Cold Spring Harbor Perspectives in Biology* 6, a016840.
- Luzio, J.P., Pryor, P.R., Bright, N.A., 2007. Lysosomes: Fusion and function. *Nature Reviews. Molecular Cell Biology* 8, 622–632.
- Luzio, J.P., Pryor, P.R., Gray, S.R., *et al.*, 2005. Membrane traffic to and from lysosomes. *Biochemical Society Symposia* 77–86.
- Luzio, J.P., Rous, B.A., Bright, N.A., *et al.*, 2000. Lysosome-endosome fusion and lysosome biogenesis. *Journal of Cell Science* 113 (Pt 9), 1515–1524.
- Marks, M.S., Heijnen, H.F., Raposo, G., 2013. Lysosome-related organelles: Unusual compartments become mainstream. *Current Opinion in Cell Biology* 25, 495–505.
- Marone, G., Casolaro, V., Patella, V., Florio, G., Triggiani, M., 1997. Molecular and cellular biology of mast cells and basophils. *International Archives of Allergy and Immunology* 114, 207–217.
- Matti, U., Pattu, V., Halimani, M., *et al.*, 2013. Synaptobrevin2 is the v-SNARE required for cytotoxic T-lymphocyte lytic granule fusion. *Nature Communications* 4, 1439.
- Mayadas, T.N., Johnson, R.C., Rayburn, H., Hynes, R.O., Wagner, D.D., 1993. Leukocyte rolling and extravasation are severely compromised in P selectin-deficient mice. *Cell* 74, 541–554.
- McCullough, J., Colf, L.A., Sundquist, W.I., 2013. Membrane fission reactions of the mammalian ESCRT pathway. *Annual Review of Biochemistry* 82, 663–692.
- McEver, R.P., Beckstead, J.H., Moore, K.L., Marshall-Carlson, L., Bainton, D.F., 1989. GMP-140, a platelet alpha-granule membrane protein, is also synthesized by vascular endothelial cells and is localized in Weibel-Palade bodies. *Journal of Clinical Investigation* 84, 92–99.
- McNeil, P.L., 2002. Repairing a torn cell surface: Make way, lysosomes to the rescue. *Journal of Cell Science* 115, 873–879.
- McNeil, P.L., Kirchhausen, T., 2005. An emergency response team for membrane repair. *Nature Reviews. Molecular Cell Biology* 6, 499–505.
- Menasche, G., Feldmann, J., Houdusse, A., *et al.*, 2003. Biochemical and functional characterization of Rab27a mutations occurring in Griscelli syndrome patients. *Blood* 101, 2736–2742.
- Menasche, G., Pastural, E., Feldmann, J., *et al.*, 2000. Mutations in RAB27A cause Griscelli syndrome associated with haemophagocytic syndrome. *Nature Genetics* 25, 173–176.
- Michaillat, L., Baars, T.L., Mayer, A., 2012. Cell-free reconstitution of vacuole membrane fragmentation reveals regulation of vacuole size and number by TORC1. *Molecular Biology of the Cell* 23, 881–895.
- Michaillat, L., Mayer, A., 2013. Identification of genes affecting vacuole membrane fragmentation in *Saccharomyces cerevisiae*. *PLoS One* 8, e54160.
- Morita, E., Sandrin, V., Chung, H.Y., *et al.*, 2007. Human ESCRT and ALIX proteins interact with proteins of the midbody and function in cytokinesis. *EMBO Journal* 26, 4215–4227.
- Mullock, B.M., Bright, N.A., Fearon, C.W., Gray, S.R., Luzio, J.P., 1998. Fusion of lysosomes with late endosomes produces a hybrid organelle of intermediate density and is NSF dependent. *Journal of Cell Biology* 140, 591–601.
- Muniz, V.S., Weller, P.F., Neves, J.S., 2012. Eosinophil crystalloid granules: Structure, function, and beyond. *Journal of Leukocyte Biology* 92, 281–288.

- Nagle, D.L., Karim, M.A., Woolf, E.A., *et al.*, 1996. Identification and mutation analysis of the complete gene for Chediak-Higashi syndrome. *Nature Genetics* 14, 307–311.
- Nolte, D., Schmid, P., Jager, U., *et al.*, 1994. Leukocyte rolling in venules of striated muscle and skin is mediated by P-selectin, not by L-selectin. *American Journal of Physiology* 267, H1637–H1642.
- Ohga, S., Kudo, K., Ishii, E., *et al.*, 2010. Hematopoietic stem cell transplantation for familial hemophagocytic lymphohistiocytosis and Epstein-Barr virus-associated hemophagocytic lymphohistiocytosis in Japan. *Pediatric Blood & Cancer* 54, 299–306.
- Park, S.Y., Guo, X., 2014. Adapter protein complexes and intracellular transport. *Bioscience Reports* 34, e00123.
- Parzych, K.R., Klionsky, D.J., 2014. An overview of autophagy: Morphology, mechanism, and regulation. *Antioxidants & Redox Signaling* 20, 460–473.
- Patterson, G.H., Lippincott-Schwartz, J., 2002. A photoactivatable GFP for selective photolabeling of proteins and cells. *Science* 297, 1873–1877.
- Pattu, V., Halimani, M., Ming, M., *et al.*, 2013. In the crosshairs: Investigating lytic granules by high-resolution microscopy and electrophysiology. *Frontiers in Immunology* 4, 411.
- Pattu, V., Qu, B., Marshall, M., *et al.*, 2011. Syntaxin7 is required for lytic granule release from cytotoxic T lymphocytes. *Traffic* 12, 890–901.
- Perou, C.M., Justice, M.J., Pryor, R.J., Kaplan, J., 1996a. Complementation of the beige mutation in cultured cells by episomally replicating murine yeast artificial chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* 93, 5905–5909.
- Perou, C.M., Leslie, J.D., Green, W., *et al.*, 1997. The Beige/Chediak-Higashi syndrome gene encodes a widely expressed cytosolic protein. *Journal of Biological Chemistry* 272, 29790–29794.
- Perou, C.M., Moore, K.J., Nagle, D.L., *et al.*, 1996b. Identification of the murine beige gene by YAC complementation and positional cloning. *Nature Genetics* 13, 303–308.
- Peters, C., Baars, T.L., Buhler, S., Mayer, A., 2004. Mutual control of membrane fission and fusion proteins. *Cell* 119, 667–678.
- Pols, M.S., van Meel, E., Oorschot, V., *et al.*, 2013. hVps41 and VAMP7 function in direct TGN to late endosome transport of lysosomal membrane proteins. *Nature Communications* 4, 1361.
- Rahman, M., Haberman, A., Tracy, C., Ray, S., Kramer, H., 2012. Drosophila mauve mutants reveal a role of LYST homologs late in the maturation of phagosomes and autophagosomes. *Traffic* 13, 1680–1692.
- Raposo, G., Marks, M.S., 2007. Melanosomes – Dark organelles enlighten endosomal membrane transport. *Nature Reviews. Molecular Cell Biology* 8, 786–797.
- Reed, G.L., 2004. Platelet secretory mechanisms. *Seminars in Thrombosis and Hemostasis* 30, 441–450.
- Ren, Q., Ye, S., Whiteheart, S.W., 2008. The platelet release reaction: Just when you thought platelet secretion was simple. *Current Opinion in Hematology* 15, 537–541.
- Rothman, J.H., Yamashiro, C.T., Kane, P.M., Stevens, T.H., 1989. Protein targeting to the yeast vacuole. *Trends in Biochemical Sciences* 14, 347–350.
- Ruggeri, Z.M., 1999. Structure and function of von Willebrand factor. *Thrombosis and Haemostasis* 82, 576–584.
- Saffitz, P., Klumperman, J., 2009. Lysosome biogenesis and lysosomal membrane proteins: Trafficking meets function. *Nature Reviews. Molecular Cell Biology* 10, 623–635.
- Sakariassen, K.S., Bolhuis, P.A., Sixma, J.J., 1979. Human blood platelet adhesion to artery subendothelium is mediated by factor VIII-Von Willebrand factor bound to the subendothelium. *Nature* 279, 636–638.
- Samie, M.A., Xu, H., 2014. Lysosomal exocytosis and lipid storage disorders. *Journal of Lipid Research* 55, 995–1009.
- Sardiello, M., Palmieri, M., Di Ronza, A., *et al.*, 2009. A gene network regulating lysosomal biogenesis and function. *Science* 325, 473–477.
- Schekman, R., 1985. Protein localization and membrane traffic in yeast. *Annual Review of Cell Biology* 1, 115–143.
- Schneider, J.L., Cuervo, A.M., 2014. Autophagy and human disease: Emerging themes. *Current Opinion in Genetics & Development* 26C, 16–23.
- Schuh, A.L., Audhya, A., 2014. The ESCRT machinery: From the plasma membrane to endosomes and back again. *Critical Reviews in Biochemistry and Molecular Biology* 49, 242–261.
- Schulze, R.J., Weller, S.G., Schroeder, B., *et al.*, 2013. Lipid droplet breakdown requires dynamin 2 for vesiculation of autolysosomal tubules in hepatocytes. *Journal of Cell Biology* 203, 315–326.
- Settembre, C., Fraldi, A., Medina, D.L., Ballabio, A., 2013. Signals from the lysosome: A control centre for cellular clearance and energy metabolism. *Nature Reviews. Molecular Cell Biology* 14, 283–296.
- Seward Jr, S.L., Gahl, W.A., 2013. Hermansky-Pudlak syndrome: Health care throughout life. *Pediatrics* 132, 153–160.
- Shiflett, S.L., Kaplan, J., Ward, D.M., 2002. Chediak-Higashi syndrome: A rare disorder of lysosomes and lysosome related organelles. *Pigment Cell Research* 15, 251–257.
- Sieni, E., Cetica, V., Hackmann, Y., *et al.*, 2014. Familial hemophagocytic lymphohistiocytosis: When rare diseases shed light on immune system functioning. *Frontiers in Immunology* 5, 167.
- Sinha, B., Koster, D., Ruez, R., *et al.*, 2011. Cells respond to mechanical stress by rapid disassembly of caveolae. *Cell* 144, 402–413.
- Spritz, R.A., 1998. Genetic defects in Chediak-Higashi syndrome and the beige mouse. *Journal of Clinical Immunology* 18, 97–105.
- Stinchcombe, J., Bossi, G., Griffiths, G.M., 2004. Linking albinism and immunity: The secrets of secretory lysosomes. *Science* 305, 55–59.
- Suzuki, K., Verma, I.M., 2008. Phosphorylation of SNAP-23 by IkappaB kinase 2 regulates mast cell degranulation. *Cell* 134, 485–495.
- Tam, C., Idone, V., Devlin, C., *et al.*, 2010. Exocytosis of acid sphingomyelinase by wounded cells promotes endocytosis and plasma membrane repair. *Journal of Cell Biology* 189, 1027–1038.
- Tapper, H., Furuya, W., Grinstein, S., 2002. Localized exocytosis of primary (lysosomal) granules during phagocytosis: Role of Ca2+ -dependent tyrosine phosphorylation and microtubules. *Journal of Immunology* 168, 5287–5296.
- Teitelbaum, S.L., 2000. Bone resorption by osteoclasts. *Science* 289, 1504–1508.
- Tiwari, N., Wang, C.C., Brochetta, C., *et al.*, 2008. VAMP-8 segregates mast cell-preformed mediator exocytosis from cytokine trafficking pathways. *Blood* 111, 3665–3674.
- Tribi, F., Marcus, K., Meyer, H.E., *et al.*, 2006. Subcellular proteomics reveals neuromelanin granules to be a lysosome-related organelle. *Journal of Neural Transmission* 113, 741–749.
- Trottestam, H., Beutel, K., Meeths, M., *et al.*, 2009. Treatment of the X-linked lymphoproliferative, Griscelli and Chediak-Higashi syndromes by HLH directed therapy. *Pediatric Blood & Cancer* 52, 268–272.
- van der Sluijs, P., Zibouche, M., van Kerkhof, P., 2013. Late steps in secretory lysosome exocytosis in cytotoxic lymphocytes. *Frontiers in Immunology* 4, 359.
- van Meel, E., Boonen, M., Zhao, H., *et al.*, 2011. Disruption of the Man-6-P targeting pathway in mice impairs osteoclast secretory lysosome biogenesis. *Traffic* 12, 912–924.
- Vanier, M.T., 2010. Niemann-Pick disease type C. *Orphanet Journal of Rare Diseases* 5, 16.
- Votteler, J., Sundquist, W.I., 2013. Virus budding and the ESCRT pathway. *Cell Host & Microbe* 14, 232–241.
- Ward, D.M., Leslie, J.D., Kaplan, J., 1997. Homotypic lysosome fusion in macrophages: Analysis using an *in vitro* assay. *Journal of Cell Biology* 139, 665–673.
- Ward, D.M., Pevsner, J., Scullion, M.A., Vaughn, M., Kaplan, J., 2000. Syntaxin 7 and VAMP-7 are soluble N-ethylmaleimide-sensitive factor attachment protein receptors required for late endosome-lysosome and homotypic lysosome fusion in alveolar macrophages. *Molecular Biology of the Cell* 11, 2327–2333.
- Weaver, T.E., Na, C.L., Stahlman, M., 2002. Biogenesis of lamellar bodies, lysosome-related organelles involved in storage and secretion of pulmonary surfactant. *Seminars in Cell & Developmental Biology* 13, 263–270.
- Weibel, E.R., Palade, G.E., 1964. New cytoplasmic components in arterial endothelia. *Journal of Cell Biology* 23, 101–112.
- Wickner, W., 2010. Membrane fusion: Five lipids, four SNAREs, three chaperones, two nucleotides, and a Rab, all dancing in a ring on yeast vacuoles. *Annual Review of Cell and Developmental Biology* 26, 115–136.
- Wickner, W., Haas, A., 2000. Yeast homotypic vacuole fusion: A window on organelle trafficking mechanisms. *Annual Review of Biochemistry* 69, 247–275.
- Wu, J., Wilson, J., He, J., *et al.*, 1996. Fas ligand mutation in a patient with systemic lupus erythematosus and lymphoproliferative disease. *Journal of Clinical Investigation* 98, 1107–1113.
- Wu, X., Hammer, J.A., 2014. Melanosome transfer: It is best to give and receive. *Current Opinion in Cell Biology* 29, 1–7.
- Xu, J., Toops, K.A., Diaz, F., *et al.*, 2012a. Mechanism of polarized lysosome exocytosis in epithelial cells. *Journal of Cell Science* 125, 5937–5943.
- Xu, M., Xia, M., Li, X.X., *et al.*, 2012b. Requirement of translocated lysosomal V1 H (+)-ATPase for activation of membrane acid sphingomyelinase and raft clustering in coronary endothelial cells. *Molecular Biology of the Cell* 23, 1546–1557.
- Yamamoto, K., Ishii, E., Sako, M., *et al.*, 2004. Identification of novel MUNC13-4 mutations in familial haemophagocytic lymphohistiocytosis and functional analysis of MUNC13-4-deficient cytotoxic T lymphocytes. *Journal of Medical Genetics* 41, 763–767.

- Ye, S., Huang, Y., Joshi, S., *et al.*, 2014. Platelet secretion and hemostasis require syntaxin-binding protein STXBP5. *Journal of Clinical Investigation* 124, 4517–4528.
- Zhang, Z., Chen, G., Zhou, W., *et al.*, 2007. Regulated ATP release from astrocytes through lysosome exocytosis. *Nature Cell Biology* 9, 945–953.
- Zhao, H., Ito, Y., Chappel, J., *et al.*, 2008. Synaptotagmin VII regulates bone remodeling by modulating osteoclast and osteoblast secretion. *Developmental Cell* 14, 914–925.
- Zhao, H., Laitala-Leinonen, T., Parikka, V., 2001. Downregulation of small GTPase Rab7 impairs osteoclast polarization and bone resorption. *Journal of Biological Chemistry* 276, 39295–39302.
- Zieger, M., Mayer, A., 2012. Yeast vacuoles fragment in an asymmetrical two-phase process with distinct protein requirements. *Molecular Biology of the Cell* 23, 3438–3449.
- Zuccato, E., Blott, E.J., Holt, O., *et al.*, 2007. Sorting of Fas ligand to secretory lysosomes is regulated by mono-ubiquitylation and phosphorylation. *Journal of Cell Science* 120, 191–199.
- Fitch-Tewfik, J.L., Flaumenhaft, R., 2013. Platelet granule exocytosis: A comparison with chromaffin cells. *Frontiers in Endocrinology* 4, 77.
- Luzio, J.P., Hackmann, Y., Dieckmann, N.M., Griffiths, G.M., 2014. The biogenesis of lysosomes and lysosome-related organelles. *Cold Spring Harbor Perspectives in Biology* 6, a016840.
- Muniz, V.S., Weller, P.F., Neves, J.S., 2012. Eosinophil crystalloid granules: Structure, function, and beyond. *Journal of Leukocyte Biology* 92, 281–288.
- Pattu, V., Halimani, M., Ming, M., *et al.*, 2013. In the crosshairs: Investigating lytic granules by high-resolution microscopy and electrophysiology. *Frontiers in Immunology* 4, 411.
- Saftig, P., Klumperman, J., 2009. Lysosome biogenesis and lysosomal membrane proteins: Trafficking meets function. *Nature Reviews. Molecular Cell Biology* 10, 623–635.
- Zhao, H., Ito, Y., Chappel, J., *et al.*, 2008. Synaptotagmin VII regulates bone remodeling by modulating osteoclast and osteoblast secretion. *Developmental Cell* 14, 914–925.
- van der Sluijs, P., Zibouche, M., van Kerkhof, P., 2013. Late steps in secretory lysosome exocytosis in cytotoxic lymphocytes. *Frontiers in Immunology* 4 (359),

Further Reading

- Andrews, N.W., Almeida, P.E., Corrotte, M., 2014. Damage control: Cellular mechanisms of plasma membrane repair. *Trends in Cell Biology*. doi:10.1016/j.tcb.2014.07.008.

Lysosome-Related Organelles

M Fukuda, Tohoku University, Sendai, Miyagi, Japan

© 2016 Elsevier Inc. All rights reserved.

Glossary

Geranylgeranylation A kind of lipid modification (prenylation) of proteins. A Rab geranylgeranyl transferase (GGT) complex, which is composed of α and β subunits, mediates the addition of C₂₀ geranylgeranyl group to a Cys residue(s) in the C-terminal region of Rab small GTPases. This modification is required for recruitment of Rab proteins to specific organelle membranes.

Melanogenic enzymes Three structurally related enzymes, tyrosinase (Tyr), tyrosinase-related protein 1 (Tyrp1), and dopachrome tautomerase (Dct, also called Tyrp2), catalyze different steps in the melanin synthesis pathway. Tyr is the rate-limiting enzyme that catalyzes the first two steps of melanin synthesis, and functional defects in Tyr cause the albinism of human oculocutaneous albinism type 1 (OCA1) patients and 'albino' mice.

Membrane trafficking A process by which proteins and lipids are transported from one organelle (donor) to another organelle (acceptor) by means of a transport vesicle.

Rab Rab family small GTPases are conserved intracellular molecular switches that regulate membrane trafficking in all eukaryotes. Rab proteins cycle between a GTP-bound active form and a GDP-bound inactive form, and active Rab proteins promote membrane trafficking by recruiting their specific effector molecules. Two enzymes, guanine nucleotide exchange factor (GEF=activator) and GTPase-activating protein (GAP=inactivator), spatiotemporally regulate Rab cycling.

SNARE complex Soluble N-ethylmaleimide-sensitive-factor attachment protein receptor (SNARE) complexes are general fusion machinery that mediate membrane fusion between transport vesicles and target organelle membranes. One vesicle-SNARE (v-SNARE) protein on the vesicle and two or three target SNARE (t-SNARE) proteins on the target membrane assemble into a parallel four-helix bundle (SNARE zippering) that mediates membrane fusion.

Morphology, Composition, and Function of Lysosome-Related Organelles

Lysosomes are major degradative organelles in eukaryotic cells. Their luminal pH is acidic (~5), and a variety of acid hydrolases in their lumen achieve their degradative function. Lysosomes also contain a unique set of highly glycosylated, lysosome-associated membrane proteins (LAMPs), for example, LAMP-1 and LAMP-2, in their limiting membrane (Saftig and Klumperman, 2009). Lysosome-related organelles (LROs) are a group of cell-type-specific membrane compartments in metazoans that share several of the above features with lysosomes, including an acidic pH and containing certain LAMPs (Dell'Angelica *et al.*, 2000). Despite the features they share with lysosomes, LROs are highly diversified in terms of morphology and composition and are present in only certain types of highly specialized cells (Raposo *et al.*, 2007) (representative cell-type-specific LROs are summarized in the first and second columns of Table 1). In contrast to the heterogeneous morphology of lysosomes, each LRO exhibits highly specialized morphology, for example, the dense granules and α -granules in platelets are spherical shaped, the same as the conventional secretory vesicles in endocrine cells; the melanosomes in melanocytes are rugby-ball-shaped; the Weibel-Palade bodies (WPBs) in endothelial cells are cigar-shaped; and the MHC (major histocompatibility complex) class II compartments in antigen-presenting cells and lamellar bodies in pulmonary alveolar type II cells have a multilamellar structure (Figure 1). The composition of each LRO is also highly specialized (Table 1, third column) to maintain its unique morphology and to perform specialized functions.

Melanosomes contain Pmel17 protein fibrils, on to which melanins are deposited, and WPBs contain von Willebrand factor (vWF) tubules (Figure 1). Because of the differences in structure and composition of LROs, each LRO possess a unique biological function(s) (Table 1, fourth column). The same as lysosomes undergo exocytosis under specialized conditions, for example, during plasma membrane repair, most LROs undergo exocytosis, that is, secretion of the luminal content of the LROs into the extracellular space by fusion between LROs and the plasma membrane, in response to extracellular stimuli, and so these secretory LROs are also called secretory lysosomes (Blott and Griffiths, 2002). The secretory functions of LROs are crucial for many aspects of biological activities in higher animals, including pigmentation (transfer of melanins from melanocytes to keratinocytes), hemostasis (secretion of clotting factors by platelets and endothelial cells), immunity (secretion of perforin and granzymes by cytotoxic T lymphocytes for target cell killing), and lung plasticity (surfactant secretion by alveolar type II cells).

Biogenesis and Transport of LROs

The mechanisms of biogenesis of each LRO are thought to differ considerably to maintain its unique morphology and composition, although several key factors appear to be involved in the biogenesis of all LROs (Marks *et al.*, 2013). However, all LROs are formed in a stepwise manner from immature LROs to mature LROs by acquiring specific cargos via membrane trafficking, a process that transports proteins

Table 1 Cell-specific LROs and their functions

Organelle	Cell type	Major content	Physiological function
Melanosomes	Melanocytes and RPE cells	Melanin and melanogenic enzymes	Melanin synthesis, storage, and pigmentation
Lytic granules ^a	Cytotoxic T lymphocytes and NK cells	Perforin and granzymes	Target cell killing
Dense granules ^a	Platelets and megakaryocytes	ATP, ADP, and serotonin	Blood clotting
α Granules ^a	Platelets and megakaryocytes	Fibrinogen, vWF, and P-selectin	Platelet adhesion and blood clotting
Basophilic granules ^a	Basophils and mast cells	Histamine and serotonin	Inflammation
MHC class II compartments ^a	Dendritic cells, B lymphocytes, and macrophages	MHC class II	Antigen presentation
Azurophilic/primary granules ^a	Neutrophils	Myeloperoxidase and defensins	Antimicrobial defense
Osteoclast granules ^a	Osteoclasts	Lysosomal hydrolyzes	Bone resorption and remodeling
Lamellar bodies ^a	Pulmonary alveolar type II cells	Surfactant phospholipids	Surfactant storage and secretion
Weibel–Palade bodies ^a	Endothelial cells	vWF and P-selection	Blood clotting and inflammation
Notochord vacuoles	Notochord	Not identified ^b	Morphogenesis of the body axis and spine
Pigment granules	<i>Drosophila melanogaster</i> eye cells	Eye pigments	Eye pigmentation
Fat storage organelles	<i>Caenorhabditis elegans</i> gut cells	Lipofuscin	Fat storage

^aBecause these lysosome-related organelles undergo exocytosis in a stimulus-dependent manner, they are also called secretory lysosomes. The stimulus for exocytosis varies with the cell types.

^bNotochord vacuoles were assumed to contain glycosaminoglycans, but recent evidence shows that zebrafish notochord vacuoles do not contain glycosaminoglycans.

Abbreviations: ADP, adenosine diphosphate; ATP, adenosine triphosphate; LROs, lysosome-related organelles; MHC, major histocompatibility complex; NK, natural killer; RPE, retinal pigment epithelial.

Source: Adapted from Raposo, G., Marks, M.S., Cutler D.F., 2007. Lysosome-related organelles: Driving post-Golgi compartments into specialization. *Current Opinion in Cell Biology* 19, 394–401; Blott, E.J., Griffiths, G.M., 2002. Secretory lysosomes. *Nature Reviews Molecular Cell Biology* 3, 122–131.

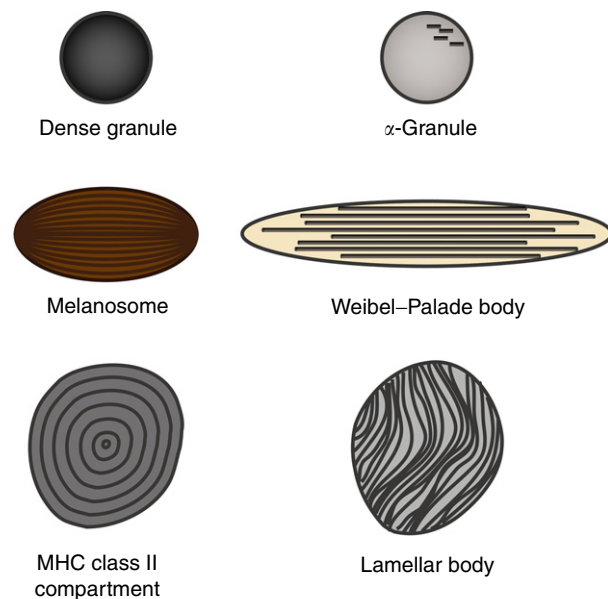


Figure 1 Unique morphology of LROs. Schematic representation of dense granule and α -granule (both spherical shaped) in platelets (top row); melanosome in melanocytes (rugby-ball-shaped) and WPB (cigar-shaped) in endothelial cells (middle row); and MHC class II compartment (multilamellar structure) in antigen-presenting cells and lamellar body (multilamellar structure) in pulmonary alveolar type II cells (bottom row).

and lipids from a donor organelle to an acceptor organelle by means of a vesicle carrier and generates a unique luminal environment. LROs may have different origins,

because four representative immature LROs, that is, immature melanosomes, immature WPBs, immature α -granules, and immature lytic granules, all originate from a different organelle sources (Marks *et al.*, 2013). Immature melanosomes are non-pigmented membrane compartments that contain few intraluminal vesicles and cytosolic coats, and they are thought to be derived from early endosomes (see next section for details). Immature WPBs originate from the *trans*-Golgi network (TGN) and assembly of vWF, a major content of WPBs, into tubules enables to form a cigar-shaped immature WPBs. α -Granules originate from the late endosomes or multivesicular endosomes in megakaryocytes. Immature lytic granules are likely to be transformed lysosomes that mature upon T-cell activation, because lytic granules and classical lysosomes are not simultaneously present in cytotoxic T lymphocytes. After the formation of immature LROs, their specific cargo molecules are transported by either the biosynthetic pathway, the endocytic pathway, or both pathways for their maturation. Mature LROs must be transported to the specific intracellular site where they function. Secretory LROs/lysosomes (and melanosomes) must be transported to just beneath the plasma membrane, where they exocytose, along two different cytoskeletons, microtubules and actin filaments. Specific kinesins (microtubule-based anterograde motors), dynein (a microtubule-based retrograde motor), and myosins (actin-based motors) (Akhmanova and Hammer, 2010) are thought to be involved in the transport of LROs, but the exact mechanism of transport of each LRO is not fully understood. Below is a description of the molecular basis of a representative example of the biogenesis and transport machineries of LROs, that is, melanosome biogenesis and transport machineries in epidermal melanocytes.

Molecular Basis of Melanosome Biogenesis and Transport

Melanosomes are the melanin-containing organelles that are responsible for pigmentation of the hair and skin of mammals. Because melanosomes are visible under a light microscope, they are often regarded as a representative model for analyzing LRO biogenesis and transport (Nascimento *et al.*, 2003; Wasmeier *et al.*, 2008). Melanosomes are classified into stages I to IV based on their morphology and degree of maturation (Figure 2, top). Stage I and stage II melanosomes are immature melanosomes (also called premelanosomes), and because of the absence of melanogenic enzymes they are not pigmented. Stage I melanosomes are non-pigmented endosomal compartments that contain few intraluminal vesicles and cytosolic coats and are also called coated endosomes. Their intraluminal vesicles serve as a scaffold for polymerizing melanocyte-specific Pmel17 amyloid fibrils into sheets, which help to form rugby-ball-shaped stage II melanosomes. Stage III melanosomes are less-pigmented melanosomes that contain three melanogenic enzymes, that is, tyrosinase (Tyr), tyrosinase-related protein 1 (Tyrp1), and dopachrome tautomerase (Dct), and other proteins required for melanosome maturation, including a putative transporter oculocutaneous albinism type 2 (OCA2, also called P protein), via post-Golgi trafficking or endosomal trafficking. These trafficking pathways are finely tuned by a set of molecules, including biogenesis of lysosome-related organelles complex-1 ~ 3 (BLOC-1 ~ 3) (Dell'Angelica, 2004), adaptor protein (AP) complex (Bonifacino and Traub, 2003), and Rab small GTPases (Stenmark, 2009). Melanins synthesized by the melanogenic enzymes are deposited on Pmel17 sheets, and mature stage IV melanosomes are formed.

Because melanosomes mature around the nucleus, they need to be transported to the cell periphery along two different cytoskeletons and they are eventually transferred from the dendrites of melanocytes to neighboring keratinocytes and hair matrix cells for pigmentation of mammalian skin and hair, respectively (Figure 2, middle). Mature melanosomes are first transported to the cell periphery along microtubules by a kinesin motor that powers anterograde movement on microtubules, then transferred to actin filaments, and then finally transported to just beneath the plasma membrane along actin filaments by an actin-based motor myosin-Va (Ohbayashi and Fukuda, 2012). Mature melanosomes anchored to the plasma membrane or actin filaments are eventually transferred from melanocyte dendrites to neighboring keratinocytes and hair matrix cells by mechanisms that are largely unknown (Wu and Hammer, 2014). In contrast to mammalian melanosomes, the melanosomes in fish and amphibian melanophores, which are equivalent to the melanocytes in mammals, are rapidly transported bidirectionally in response to extracellular stimuli and are not transferred to neighboring cells. Their bidirectional movements are regulated by coordination of two microtubule-based motors, kinesin and dynein. Rapid changes between two melanosome distribution states, melanosome aggregation around the nucleus and peripheral melanosome dispersion, in melanophores accounts for the changes in the body color of fishes and amphibians that serve to camouflage them.

The best characterized melanosome transport system is the actin-dependent melanosome transport system, which, in

vertebrates, is composed of three conserved molecules, Rab27A, Slac2-a (also called melanophilin), and myosin-Va (Fukuda, 2013). Rab27A is a member of the small GTPase Rab family and specifically localizes to mature melanosomes, when it binds GTP, that is, when Rab27A is activated. The active Rab27A on the melanosome then recruits its specific effector molecule Slac2-a via the N-terminal Slp homology domain (SHD=Rab27A effector domain). Slac2-a also contains a myosin-Va-binding domain in the middle of the molecule and functions as a linker protein between Rab27A on the melanosome and myosin-Va (Figure 2, bottom). The resulting tripartite protein complex composed of Rab27A, Slac2-a, and myosin-Va facilitates melanosome transfer from microtubules to actin filaments and regulates actin-dependent melanosome transport.

Although the above-described actin-dependent melanosome transport machinery is highly conserved both in mammalian melanocytes and in lower vertebrate melanophores, the microtubule-dependent melanosome transport systems of melanocytes and melanophores appear to differ, because heterotrimeric kinesin-2/Kif3 regulates microtubule-dependent anterograde melanosome transport in melanophores, whereas kinesin-1/Kif5 seems to be involved in microtubule-dependent anterograde melanosome transport in melanocytes. The precise mechanism of microtubule-dependent anterograde melanosome transport in melanocytes remains elusive, but it has been suggested that a specific Rab isoform may be involved in the process, possibly as a cargo receptor for the kinesin motor (Ishida *et al.*, 2012), in the same way that Rab27A functions as a cargo receptor for actin-dependent melanosome transport.

Genetic Diseases Associated with Defects in LRO Biogenesis and Transport

Because LROs play important roles in various biological activities, defects in LRO biogenesis or transport are known to cause human diseases, including albinism, immunodeficiency, and pulmonary fibrosis. Defects in melanosome biogenesis or transport cause pigmentary disorders in hair, skin, and/or eyes, for example, a functional defect in Tyr, a key enzyme in the melanin synthesis pathway in melanocytes, cause human oculocutaneous albinism type 1 (OCA1) (Tomita and Suzuki, 2004). Abnormalities in the biogenesis or transport of LROs in immune cells are associated with immunological disorders, for example, mutations of perforin, a major cargo of lytic granules in cytotoxic T lymphocytes that is responsible for their killing activity, cause familial hemophagocytic lymphohistiocytosis type 1 (FHL1). Lamellar body dysfunction in pulmonary alveolar type II cells is associated with pulmonary fibrosis. It is particularly noteworthy that the functions of several of the same LROs are impaired in patients with certain human genetic diseases, including Griscelli syndrome (GS), Hermansky-Pudlak syndrome (HPS), and Chédiak-Higashi syndrome (CHS). Patients with these syndromes have certain symptoms in common, including silvery hair (a symptom of melanosome dysfunction), pale blue eyes (a symptom of melanosome dysfunction), and bleeding tendency (a symptom of platelet granule dysfunction), thereby indicating that

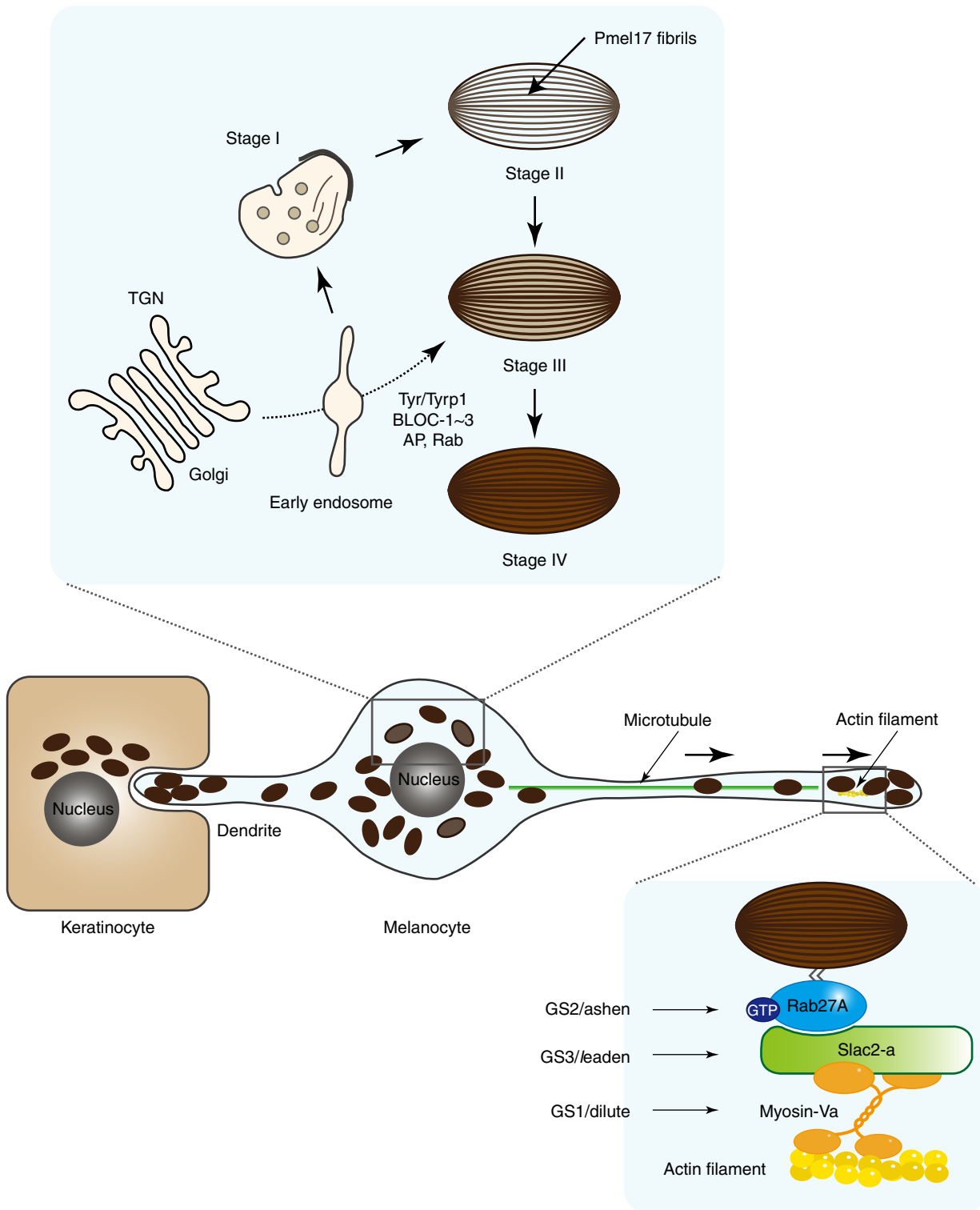


Figure 2 Biogenesis and transport of melanosomes in mammalian epidermal melanocytes. (Top) Schematic representation of the stepwise development of a mature melanosome. The stage I immature melanosome with intraluminal vesicles is derived from early endosomes. The stage II immature melanosome contains Pmel17 fibrils. Melanogenic enzymes, for example, tyrosinase (Tyr) and tyrosinase-related protein 1 (Typr1), are transported to the stage III melanosome, where melanin synthesis takes place. The stage IV melanosome is a mature melanosome. (Middle) Melanosomes form and mature around the nucleus. Mature melanosomes are transported to just beneath the plasma membrane along two different cytoskeletons, microtubules and actin filaments. Melanosomes are eventually transferred from the dendrites of melanocytes to neighboring keratinocytes. (Bottom) Mature melanosomes are transported to the cell periphery by motor protein complexes along different cytoskeletons, that is, microtubules and actin filaments. An actin-based melanosome transport complex composed of Rab27A, Slac2-a/melanophilin, and myosin Va, deficiency of each of which causes the hypopigmentation of GS2/ashen, GS3/leaden, and GS1/dilute, respectively.

Table 2 Genetic diseases associated with defects in a subset of LROs

Protein name	Human disease	Mouse model	Proposed function
Griscelli syndrome (GS)			
Myosin-Va	GS1	dilute	An actin-based melanosome transport complex composed of Rab27A, Slac2-a, and myosin-Va
Rab27A	GS2	ashen	
Slac2-a/melanophilin	GS3	leaden	
Chédiak–Higashi syndrome (CHS)			
CHS1/LYST	CHS	beige	Lysosome fission
Hermansky–Pudlak syndrome (HPS) ^a			
AP-3 β 3A	HPS2	pearl	A heterotetrameric adaptor complex (AP) that mediates cargo sorting from endosomes to LROs
AP-3 δ		mocha	
Pallidin	HPS9	pallid	Together with snapin, BLOS1, and BLOS2, these proteins form a BLOC-1 complex that contributes to t-SNARE regulation
Cappuccino		cappuccino	
Muted		muted	
Dysbindin	HPS7	sandy	
BLOS3a	HPS8	reduced pigmentation	
HPS3	HPS3	cocoa	A BLOC-2 complex composed of three subunits that functions downstream of BLOC-1 in cargo transport
HPS5	HPS5	ruby eye-2	
HPS6	HPS6	ruby eye	
HPS1	HPS1	pale ear	A heterodimeric guanine nucleotide exchange factor (HPS1–HPS4) for Rab32/38
HPS4	HPS4	light ear	
Rab38		chocolate	
Vps33a		buff	A subunit of the Vps-C complex
Rabggta		gunmetal	The α subunit of Rab GGT II complex, which mediates prenylation of Rab proteins

^aThe normal proteins related to HPS are classified into six groups based on their function: AP-3 complex, BLOC-1 ~3 complexes, Vps-C complex, and Rab GGT II complex (see [Figure 4](#)).

Abbreviations: GGT, geranylgeranyl transferase; LROs, lysosome-related organelles.

Source: Adapted from Raposo, G., Marks, M.S., Cutler D.F., 2007. Lysosome-related organelles: Driving post-Golgi compartments into specialization. *Current Opinion in Cell Biology* 19, 394–401; Wei, A.H., Li, W., 2013. Hermansky–Pudlak syndrome: Pigmentary and non-pigmentary defects and their pathogenesis. *Pigment Cell & Melanoma Research* 26, 176–192.

the same factors are involved in the biogenesis or transport of several LROs, despite the fact that the morphology and function of the LROs involved differ considerably. The mouse models (coat color mutant mice) of three well-known human syndromes whose manifestations are attributable to impaired biogenesis or function of several LROs, that is, of GS, HPS, and CHS, and the proposed functions of the gene products responsible for the clinical manifestations of these syndromes are described below ([Table 2](#)).

GS

GS is an extremely rare autosomal recessive disorder characterized by silvery hair (partial albinism) at birth. GS is classified into three types, that is, GS1 (mouse model: ‘dilute’), GS2 (mouse model: ‘ashen’), and GS3 (mouse model: ‘leaden’), based on difference between the genes responsible ([Table 2](#)) and in their symptoms ([Van Gele et al., 2009](#)). GS1 is caused by mutations in the *MYOSIN-VA* gene, which encodes an actin-based motor and is characterized by neurological

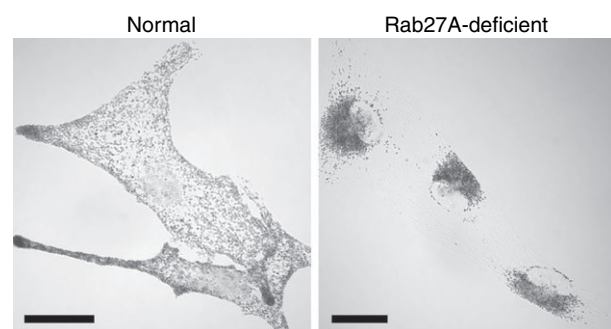


Figure 3 Normal melanocytes (left, mouse melan-a cells; [Bennett et al., 1987](#)), which have a peripheral melanosome distribution, and Rab27A-deficient melanocytes (right, mouse melan-ash cells), which are characterized by perinuclear melanosome aggregation (i.e., a typical Griscelli syndrome phenotype). Scale bar, 20 μ m. Images were taken by Morié Ishida. Melanosome movements in normal cells and in Slac2-a-deficient cells (equivalent to Rab27A-deficient cells) are shown in the form of Multimedia Annexes ([Video 1](#) and [Video 2](#), respectively).

disorders in addition to partial albinism. GS2 is caused by mutations in the *RAB27A* gene that encodes Rab-type small GTPase (Stenmark, 2009) and is characterized by immunodeficiency in addition to partial albinism. By contrast, GS3 is caused by mutations in the *SLAC2-A/MLPH* gene, which encodes an effector molecule of Rab27A, and it is characterized by partial albinism alone. Melanocytes from GS1~3 patients or their mouse models contain mature melanosomes and are characterized by melanosome aggregation in the perinuclear region, in contrast to the peripheral melanosome distribution of normal melanocytes (Figure 3). Thus, the hypopigmentation of the hair and skin of GS patients is attributable to a defect in melanosome transport, specifically in actin-dependent melanosome transport (Figure 2), and not to a defect in melanosome biogenesis. The gene products responsible for GS1, GS2, and GS3, that is, myosin-Va, Rab27A, and Slac2-a/melanophilin, respectively, cooperatively function in actin-

dependent melanosome transport by forming a tripartite protein complex on mature melanosomes. Rab27A localizes mature melanosomes by C-terminal geranylgeranylation, a lipid modification that anchors Rab27A to the organelle membrane. Slac2-a then directly interacts with Rab27A on the melanosome through its N-terminal SHD and recruits an actin-based motor myosin-Va to the melanosome (Figure 2, bottom). Interestingly, the tripartite protein complex appears to function in certain types of specialized cells, including epidermal melanocytes. For example, Rab27A is indispensable to exocytosis of cytotoxic granules, specifically to the docking step of cytotoxic granules to the plasma membrane, but neither myosin-Va nor Slac2-a is involved in this process, which is consistent with the fact that neither GS1 patients nor GS3 patients exhibit immunodeficiency (Van Gele *et al.*, 2009). Rab27A is thought to regulate the transport of LROs in different cell types through interaction with cell-type-specific or tissue-specific

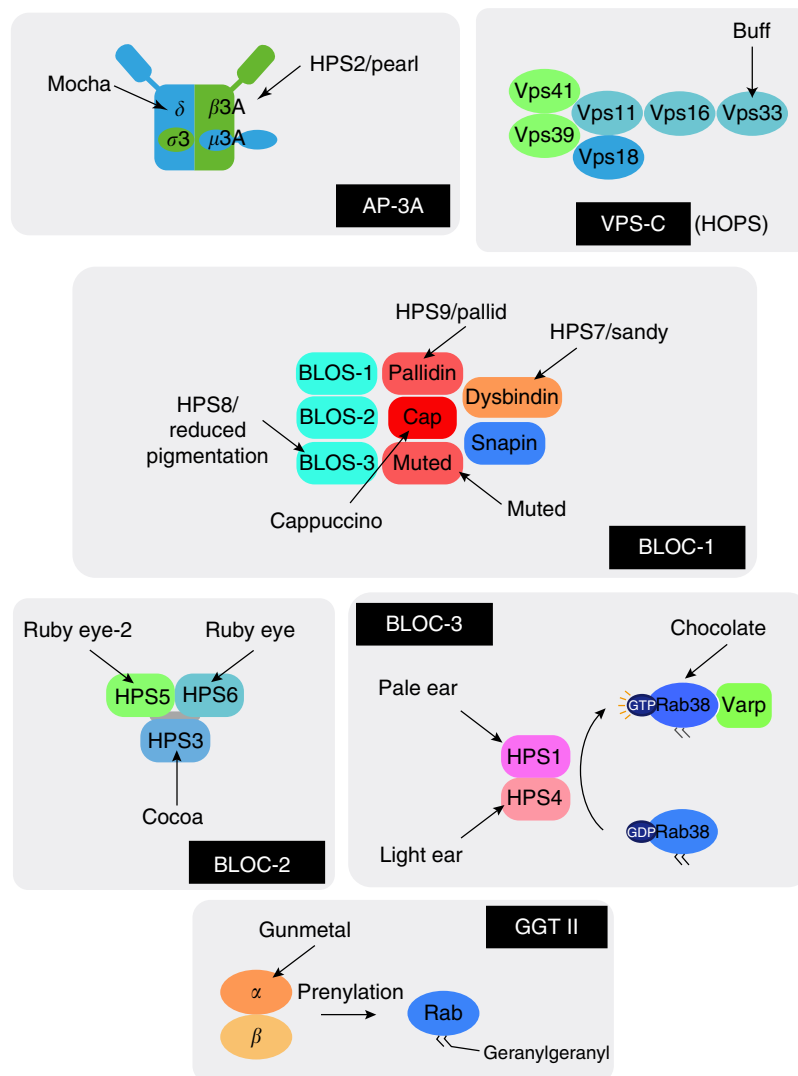


Figure 4 Schematic representation of six large cytoplasmic protein complexes associated with HPS: AP-3, VPS-C, BLOC-1, BLOC-2, BLOC-3, and Rab GGT II complex (see also Table 2). The VPS-C complex interacts with VPS18, VPS41, and VPS39, to form heterohexamers, named a HOPS complex, which mediates tethering between late endosomes and lysosomes (Solinger and Spang, 2013). Protein interactions between subunits of the BLOC-1 complex are complicated and their interaction networks are not depicted in detail here (see Wei and Li, 2013 for details).

Rab27A effector molecules (Fukuda, 2013): the Rab27A–Slac2-a–myosin-Va complex promotes actin-dependent melanosome transport in epidermal melanocytes, and the Rab27A–Slac2-c/MyRIP–myosin VIIa complex promotes actin-dependent melanosome transport in retinal pigment epithelial (RPE) cells.

HPS

HPS (HPS1 ~ 9) and its corresponding mouse models is a group of genetic disorders that are characterized by silvery hair and bleeding tendency at birth (Table 2) (Wei and Li, 2013). Several types of HPS are also characterized by pulmonary fibrosis due to lamellar body dysfunction in pulmonary alveolar type II cells. Unlike GS, the manifestations of HPS are caused by defects in the biogenesis of LROs, not in their transport, for example, malformation of melanosomes, specifically the transport of melanogenic enzymes to melanosomes, is the primary cause of the hair and skin hypopigmentation in HPS patients. The gene products responsible for HPS are subunits of large cytoplasmic protein complexes and have been classified into six different groups: adaptor protein complex (AP)-3, vacuolar protein sorting (VPS)-C, biogenesis of lysosome-related organelles complex (BLOC)-1, BLOC-2, and BLOC-3, and Rab geranylgeranyl transferase (GGT) II complex (Table 2 and Figure 4). All of these protein complexes are thought to be involved in sorting or transporting membrane cargo proteins to LROs, for example, melanogenic enzyme tyrosinase to melanosomes. AP-3 is a heterotetrameric adaptor that recognizes cytoplasmic signals of transmembrane cargo proteins and packs them into transport vesicles (Dell'Angelica, 2009). VPS-C is a heterotetrameric complex (VPS11, 16, 18, and 33) that is thought to regulate tethering and SNARE (soluble N-ethylmaleimide-sensitive-factor attachment protein receptor)-dependent fusion of transport vesicles to LROs (Solinger and Spang, 2013). BLOC-1, BLOC-2, and BLOC-3 are composed of eight subunits, three subunits, and two subunits, respectively, and are involved in the transport of transmembrane cargo proteins to LROs, although the exact functions of BLOC-1 and BLOC-2 remain largely unknown (Dell'Angelica, 2004). BLOC-3 functions as a heterodimer guanine nucleotide exchange factor (activator) for small GTPases Rab32 and Rab38, and a mutation in the mouse *rab38* gene causes the hypopigmentation phenotype of 'chocolate' mice (Marks *et al.*, 2013). Similarly, a functional defect in the Rab GGT II complex that mediates prenylation of Rab proteins results in the HPS-like phenotype of 'gunmetal' mice.

Rab32 and Rab38 redundantly function in the transport of melanogenic enzymes to melanosomes through interaction with their specific effector molecule Varp (VPS9-ankyrin-repeat protein), a scaffold protein that interacts with a variety of proteins, including a vesicle-SNARE VAMP7 (vesicle-associated membrane protein 7)/TI-VAMP (Ohbayashi and Fukuda, 2012). Varp maintains an inactive conformation of VAMP7 during the transport of transmembrane cargo proteins, for example, tyrosinase, after which VAMP7 forms a SNARE complex with two target-SNAREs, syntaxin-3 and SNAP23 (synapto-some-associated protein of 23 kDa), which promotes the transfer of melanogenic enzymes to melanosomes presumably

by mediating membrane fusion between the transport vesicle and the immature melanosome (Yatsu *et al.*, 2013).

CHS

CHS is a rare autosomal recessive disorder that is characterized by silvery hair at birth, and infections and neuropathy are frequently observed in CHS patients. Mutations in the *CHS1*/*LYST* gene are known to cause CHS, but the exact function of the large protein that contains a BEACH domain, whose function is unknown. *LYST* is likely to be involved in the biogenesis of lysosomes and LROs, because these organelles are enlarged in the cells of CHS patients. Thus, *LYST* may promote fission of lysosomes and LROs.

Multi-Media Annexes

The following are the supplementary data to this article: Videos 1 and 2.

See also: Organelles: Structure and Function: Conventional and Secretory Lysosomes

References

- Akhmanova, A., Hammer III, J.A., 2010. Linking molecular motors to membrane cargo. *Current Opinion in Cell Biology* 22, 479–487.
- Bennett, D.C., Cooper, P.J., Hart, I.R., 1987. A line of non-tumorigenic mouse melanocytes, syngeneic with the B16 melanoma and requiring a tumour promoter for growth. *International Journal of Cancer* 39, 414–418.
- Blott, E.J., Griffiths, G.M., 2002. Secretory lysosomes. *Nature Reviews Molecular Cell Biology* 3, 122–131.
- Bonifacino, J.S., Traub, L.M., 2003. Signals for sorting of transmembrane proteins to endosomes and lysosomes. *Annual Review of Biochemistry* 72, 395–447.
- Dell'Angelica, E.C., 2004. The building BLOC(k)s of lysosomes and related organelles. *Current Opinion in Cell Biology* 16, 458–464.
- Dell'Angelica, E.C., 2009. AP-3-dependent trafficking and disease: The first decade. *Current Opinion in Cell Biology* 21, 552–559.
- Dell'Angelica, E.C., Mullins, C., Caplan, S., Bonifacino, J.S., 2000. Lysosome-related organelles. *FASEB Journal* 14, 1265–1278.
- Fukuda, M., 2013. Rab27 effectors, pleiotropic regulators in secretory pathways. *Traffic* 14, 949–963.
- Ishida, M., Ohbayashi, N., Maruta, Y., *et al.*, 2012. Functional involvement of Rab1A in microtubule-dependent anterograde melanosome transport in melanocytes. *Journal of Cell Science* 125, 5177–5187.
- Marks, M.S., Heijnen, H.F., Raposo, G., 2013. Lysosome-related organelles: Unusual compartments become mainstream. *Current Opinion in Cell Biology* 25, 495–505.
- Nascimento, A.A., Roland, J.T., Gelfand, V.I., 2003. Pigment cells: A model for the study of organelle transport. *Annual Review of Cell and Developmental Biology* 19, 469–491.
- Ohbayashi, N., Fukuda, M., 2012. Role of Rab family GTPases and their effectors in melanosomal logistics. *Journal of Biochemistry* 151, 343–351.
- Raposo, G., Marks, M.S., Cutler, D.F., 2007. Lysosome-related organelles: Driving post-Golgi compartments into specialisation. *Current Opinion in Cell Biology* 19, 394–401.
- Saffitz, P., Klumperman, J., 2009. Lysosome biogenesis and lysosomal membrane proteins: Trafficking meets function. *Nature Reviews Molecular Cell Biology* 10, 623–635.
- Solinger, J.A., Spang, A., 2013. Tethering complexes in the endocytic pathway: CORVET and HOPS. *FEBS Journal* 280, 2743–2757.

- Stenmark, H., 2009. Rab GTPases as coordinators of vesicle traffic. *Nature Reviews Molecular Cell Biology* 10, 513–525.
- Tomita, Y., Suzuki, T., 2004. Genetics of pigmentary disorders. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics* 131C, 75–81.
- Van Gele, M., Dynoodt, P., Lambert, J., 2009. Griscelli syndrome: A model system to study vesicular trafficking. *Pigment Cell & Melanoma Research* 22, 268–282.
- Wasmeier, C., Hume, A.N., Bolasco, G., Seabra, M.C., 2008. Melanosomes at a glance. *Journal of Cell Science* 121, 3995–3999.
- Wei, A.H., Li, W., 2013. Hermansky-Pudlak syndrome: Pigmentary and non-pigmentary defects and their pathogenesis. *Pigment Cell & Melanoma Research* 26, 176–192.
- Wu, X., Hammer, J.A., 2014. Melanosome transfer: It is best to give and receive. *Current Opinion in Cell Biology* 29, 1–7.

- Yatsu, A., Ohbayashi, N., Tamura, K., Fukuda, M., 2013. Syntaxin-3 is required for melanosomal localization of Tyrp1 in melanocytes. *Journal of Investigation Dermatology* 133, 2237–2246.

Relevant Websites

- <http://www.espcr.org/micemut/>
International Federation of Pigment Cell Society (IFPCS).
- <http://www.sgul.ac.uk/depts/anatomy/pages/WTFGCB.htm>
The Wellcome Trust Functional Genomics Cell Bank.

At the Center of Autophagy: Autophagosomes

F Reggiori and M Mauthe, University Medical Center Utrecht, Utrecht, The Netherlands

© 2016 Elsevier Inc. All rights reserved.

Glossary

ATG proteins The AuTophagy-related (ATG) proteins are the key players regulating and mediating macroautophagy. On a note, there are also genes essential for the progression of autophagy that are not called ATG.

Autophagosomes They are vesicles with a double lipid bilayer, which are formed during macroautophagy. Autophagosomes sequester and deliver cytoplasmic material into the lysosomes for degradation.

Autophagy An evolutionary conserved lysosomal degradative pathway. The term encompasses 3 processes: macroautophagy, microautophagy, and chaperone-mediated autophagy. Basal autophagy constantly recycles the cytoplasmic components and thereby maintains the cellular homeostasis. When cells are exposed to stress, the enhancement of autophagy contributes to overcome them.

Autophagy receptors Autophagic receptors are a set of proteins that simultaneously bind specifically a structure targeted by autophagy and ATG proteins found in the

interior of autophagosomes, mainly LC3. This property provide the substrate specificity to the selective types of autophagy.

PAS The Phagophore Assembly Site or Pre-Autophagosomal Structure is the nucleation site where the ATG proteins assemble to first generate a phagophore and subsequently expand it into an autophagosome. In yeast there is one PAS per cell whereas in mammalian cells there are several per cell.

Phagophore (also known as isolation membrane) The phagophore is a membrane cistern that is the precursor structure of the autophagosome. It is formed at the PAS by the orchestrated actions of the ATG proteins, which also elongate it into an autophagosome.

Selective autophagy This term describes forms of macroautophagy in which a specific structure (i.e. an aggregate, organelle, or microbe) is specifically and exclusively sequestered by autophagosomes.

Macroautophagy and Autophagosomes

Macroautophagy can be distinct from the two other forms of autophagy, namely, chaperon-mediated autophagy and microautophagy, because it involved the formation of autophagosomes (Mizushima *et al.*, 2008; Kraft and Martens, 2012). Macroautophagy (hereafter autophagy) is an evolutionary conserved cellular pathway that is active at basal levels in every cell. Thus, cytoplasmic material including unfolded or obsolete proteins, and damaged organelles are continuously degraded into lysosomes even if the flux of autophagy can vary from tissue to tissue. The constant turnover of the cytoplasmic components by autophagy is crucial in maintaining cellular homeostasis and also provides metabolites such as amino acids, sugars, nucleotides, and lipids to generate new macromolecules (Kraft and Martens, 2012; Levine and Kroemer, 2008; Figure 1). Autophagy can be further stimulated when a cell is exposed to cellular stresses such as nutrient starvation and helps the cell to overcome them (Shintani and Klionsky, 2004). The physiological relevance of this type of responses was initially revealed by a study showing that mice unable to undergo autophagy are unable to survive the postnatal starvation period and die shortly after birth (Kuma *et al.*, 2004). In addition to stresses, specific developmental programs, immune responses and numerous other signals can also induce autophagy (Levine *et al.*, 2011; Choi *et al.*, 2013; Mizushima and Levine, 2010).

Together with the proteasome, autophagy is one of the two major protein degradation pathways of the cell. However, while the proteasome is mainly devoted to the turnover of short-lived proteins, autophagy degrades long-lived

proteins but also numerous other cell components including sugar chains, nucleotides and lipids. While in specific situations autophagosomes randomly sequester cytoplasmic material, often their cargo is a specific structure (Wong and Cuervo, 2010). Selective types of autophagy include the turnover of organelles (e.g., mitochondria, peroxisomes, ER, lysosomes, lipid droplets...), protein aggregates, large complexes (ribosomes, mid-bodies, inflammasome...) and invading pathogens (Choi *et al.*, 2013). These forms of autophagy rely on a series of the so-called autophagy receptors, which recognize the targeted structure and mediate the subsequent recruitment and assembly of the ATG machinery (Boyle and Randow, 2013; Isakson *et al.*, 2013). This results in the formation of the autophagosome around the selected material providing the specificity to this process (Shintani and Klionsky, 2004; Figure 2). Because of this ability to specifically eliminate unwanted structures, autophagy plays a crucial function in a multitude of physiological processes in eukaryotic organisms. This pathway is thus part of our innate immune response through its ability to target invading pathogens for degradation (Levine *et al.*, 2011). Moreover, the turnover of protein aggregates prevents an early onset of neurodegenerative disorders like Huntington's or Alzheimer's disease (Choi *et al.*, 2013). Autophagy also plays major roles in development and cell differentiation in multicellular organisms and contributes to their longevity (Yoshimori, 2007; Melendez *et al.*, 2003). As a result, an impairment or defect in autophagy causes a variety of human illnesses, including neurodegenerations, cancer, inflammatory diseases, and muscular dystrophies (Levine and Kroemer, 2008; Choi *et al.*, 2013).

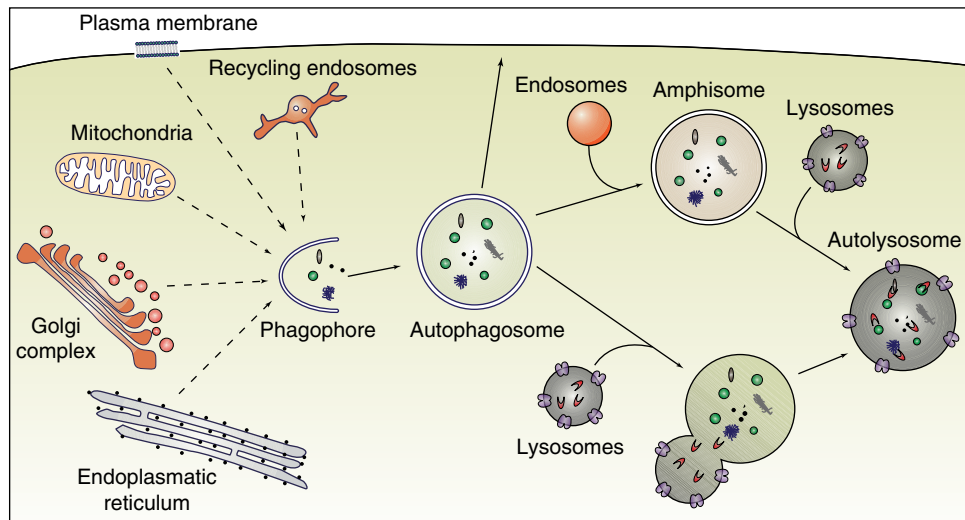


Figure 1 The mechanism of autophagy. Autophagy starts with the formation of a membrane cistern called the phagophore, where all the ATG proteins localize to. Often the initial phagophore is formed in close proximity to the ER. In addition to this organelle, other cellular compartments such as mitochondria, the plasma membrane, the Golgi complex and recycling endosomes have been postulated to contribute membranes to form the phagophore and then expand it into an autophagosome. During the formation of autophagosomes, the cytoplasmic material is sequestered within the forming vesicle. Complete autophagosomes fuse either directly with lysosomes to form autolysosomes or with endosomes to generate an organelle called amphisome. This latter eventually fuse with lysosomes too. In the lysosome lumen, resident hydrolases degrade the autophagosomal cargo and the resulting metabolites are transported back in to the cytoplasm by lysosomal permeases. The autophagosome can also fuse with the plasma membrane and thereby being one of the mechanisms for unconventional secretion.

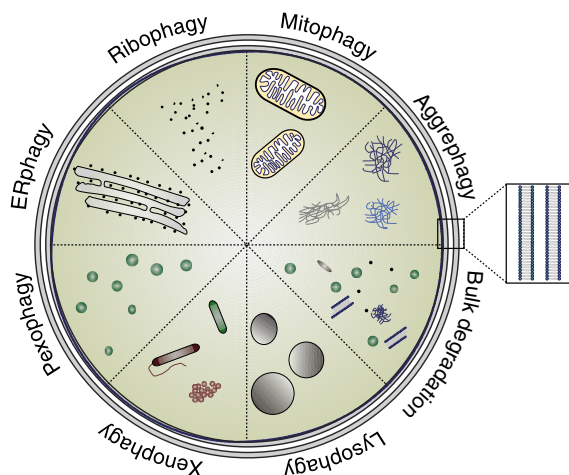


Figure 2 The different types of selective autophagy. In addition to the unspecific sequestration of cytoplasmic material (i.e., bulk degradation), autophagosomes can also specifically target a variety of cellular components. Depending on the conditions, autophagosomes can selectively and exclusively engulf mitochondria (i.e., mitophagy), ribosomes (i.e., ribophagy), peroxisomes (i.e., pexophagy), lysosomes (i.e., lysophagy), parts of the ER (i.e., ERphagy), protein aggregates (i.e., aggrephagy), pathogens (i.e., xenophagy), or lipid droplets (i.e., lipophagy; not shown here), and transport these cell components to the lysosomes for degradation.

General Characteristics of Autophagosomes

Autophagosomes were first described by the Nobel laureate Christian de Duve in 1966 (De Duve and Wattiaux, 1966). One of the peculiarities of these carriers in comparison to other

transport vesicles is their double lipid bilayer, which has also a low content in transmembrane proteins. Therefore, autophagosomes appear as very smooth vesicles when analyzed by freeze-fracture electron microscopy (Fengsrud *et al.*, 2000). Their size can vary from 500 nm up to 1.5 μ m and this flexibility allow them to sequester cargoes of different sizes (Fengsrud *et al.*, 2000; Klionsky *et al.*, 2012; Mizushima *et al.*, 2002). In contrast to other transport vesicles, an autophagosome is formed *de novo*, that is, not through budding from a pre-existing organelle, and it sequesters cytoplasmic cargo material during its completion. This different solution adopted by the cell to generate a vesicle is probably due to the topology of the cargo, which is cytoplasmic, and therefore cannot be engulfed by a single lipid bilayer with exposed extremities, something thermodynamically highly unfavorable (Reggiori and Klionsky, 2013).

The autophagic vesicles are categorized based on their morphology into early autophagosomes and late autophagosomes or autolysosomes. In early autophagosomes the sequestered cargo is morphological unchanged and whole organelles or part of organelles are found in the autophagosomal lumen. Late autophagosomes or autolysosomes are fused with lysosomes and the sequestered cargo is already being degraded (Yla-Anttila *et al.*, 2009b).

Generation of the Autophagosomes: The Machinery

The initial characterization of the molecular machinery mediating the biogenesis of autophagosomes has mostly been done in the yeast *Saccharomyces cerevisiae*. In particular the genes mediating the process of autophagy were identified through genetic screens in this organism and they are currently

called the AuTophagy-related (ATG) genes. This discovery was the breakthrough in the autophagy research because the high degree of conservation of numerous ATG genes allowed identifying their counterparts in high eukaryotes. To date 36 ATG genes have been characterized in yeast and many of them have human orthologs (Klionsky *et al.*, 2003; Yang and Klionsky, 2010). In high eukaryotes there are additional genes essential for autophagy that are not called ATG. It is now established that the generation of autophagosomes is regulated by a hierarchical interplay of the ATG proteins and it is divided in 3 discrete steps: (1) initiation, (2) nucleation, and (3) elongation and closure.

Initiation of autophagy is mediated by the ULK/ATG1 complex. The mammalian ULK kinase complex is composed of ULK1/2 (unc-51-like kinase), FIP200 (focal adhesion kinase family interacting protein of 200 kD), ATG13, and ATG101. The ULK/ATG1 complex is differentially regulated by TORC1 (mammalian target of rapamycin complex 1) and AMPK (AMP-activated protein kinase), and possibly by other signaling cascades. In particular, active TORC1 inhibits the ULK/ATG1 complex whereas active AMPK stimulates the ULK/ATG1 complex (He and Klionsky, 2009; Yang and Klionsky, 2010). TORC1, with its central serine/threonine kinase mTOR (mammalian target of rapamycin), acts as a nutrient sensor in the cell and is activated in presence of nutrients. Under these conditions, through phosphorylation, TORC1 inhibits both AMPK and the ULK/ATG1 complex, to which it is bound. In nutrient-deprived cells, in contrast, AMPK is activated and phosphorylates TORC1 leading to its inhibition and dissociation from the ULK/ATG1 complex. In addition, AMPK phosphorylates the ULK/ATG1 complex and thereby activates it. The activation by AMPK and simultaneously release of inhibition by the TORC1 stimulates the ULK/ATG1 complex kinase activity initiating the autophagic process (He and Klionsky, 2009; Yang and Klionsky, 2010).

After the initiation step, the ATG proteins assemble at cellular structures that has been called the phagophore assembly site or pre-autophagosomal structure (PAS). In mammalian cells numerous early autophagosomal structures can be observed upon autophagy induction whereas only one PAS is generated in yeast. The assembly of the ATG proteins, called nucleation, requires the local generation of the phosphoinositol-3-phosphate (PtdIns3P). PtdIns3P is generated by the PtdIns3 kinase class 3 (PtdIns3KC3), also called hVPS34. This protein is in complex with the regulatory factors BECLIN1/ATG6 and p150 (human VPS15), which together compose the PtdIns3KC3 complex (Simonsen and Tooze, 2009). The insertion of BECLIN1 into the PtdIns3KC3 complex is negatively regulated by BCL-2 protein family members and positively stimulated by the ULK complex (Russell *et al.*, 2013; Pattinger *et al.*, 2005). The PtdIns3KC3 complex fulfils several functions in autophagy and in endosomal vesicle transport pathways depending on which proteins are associated with the complex. Binding of ATG14L (or Barkor) and AMBRA1 (activating molecule in BECLIN1-regulated autophagy) to the PtdIns3KC3 complex recruits the complex to the PAS stimulating autophagosome formation (Kang *et al.*, 2011). When this complex is associated with UVRAG (ultraviolet irradiation resistant-associated gene), which is in turn bind to BIF-1 (Bax-interacting factor 1), it is engaged in autophagosome maturation as well

as endocytotic trafficking (Kang *et al.*, 2011). However, when UVRAG is interacting with RUBICON (RUN domain and cysteine-rich domain containing, BECLIN1-interacting protein), autophagosome maturation is blocked (He and Klionsky, 2009; Kang *et al.*, 2011).

While the precise role of PtdIns3P in autophagy is unclear, it is known that the phagophore is formed at PtdIns3P rich regions mostly adjacent to the ER. It is also clear that this lipid mediates the recruitment of PtdIns3P-binding proteins (also called PtdIns3P effector proteins) that are necessary for autophagosome biogenesis to the PAS (Mari *et al.*, 2011; Isakson *et al.*, 2013; Codogno *et al.*, 2012). Several of these PtdIns3P effector proteins are known that are involved in autophagosomal formation, including ALFY, DFCP1, and WIPI protein family members (homologs to the yeast ATG18). It was shown in yeast that ATG18 localization to the PAS depends on its interaction with ATG2 and the presence of PtdIns3P (Obara *et al.*, 2008; Rieter *et al.*, 2013). The ATG18-ATG2 complex interacts with ATG9, the only transmembrane ATG protein, at the PAS (Reggiori *et al.*, 2004). ATG9 thereby at least partially contributes membranes to the forming phagophore even if its interaction with the forming autophagosomes is very dynamic in both yeast and mammalian cells (Mari *et al.*, 2011; Isakson *et al.*, 2013; Codogno *et al.*, 2012; Chan and Tang, 2013; Orsi *et al.*, 2012; Yamamoto *et al.*, 2012).

The elongation and closure of the phagophore (also called isolation membrane), the third step in autophagosome biogenesis, is mainly executed by two ubiquitin-like conjugation systems, which have the ultimate result of forming the ATG16L1 complex and of lipidating the members of the ATG8/LC3 (microtubule associated protein 1 light chain 3) protein family (Shpilka *et al.*, 2012; Yang and Klionsky, 2010). The formation of the ATG16L1 complex as well as the LC3 lipidation are essential for the autophagosome biogenesis. The ATG16L1 complex contains the ATG12-ATG5 conjugate. The formation of this complex is mediated by the E1-like enzyme ATG7, which activates ATG12 and transfers it to ATG10, an E2-like enzyme, which finally covalently binds ATG12 to ATG5. The ATG12-ATG5 conjugate interact with ATG16L1 to form a multimeric complex, the ATG16L1 complex, which localizes to the phagophore and early autophagosome intermediates but dissociates from complete autophagosomes (Ohsumi and Mizushima, 2004; Yang and Klionsky, 2010).

In the second conjugation system, LC3 becomes linked to phosphatidylethanolamine (PE). Initially a preform of LC3 is posttranslationally cleaved at the C-terminus by the ATG4 proteases. The resulting cytoplasmic form of LC3, also called LC3-I, gets also activated by the E1-like enzyme ATG7 and transferred onto ATG3, another E2-like enzyme before being conjugated to PE. The ATG16L1 complex promotes this latter step by specifically bringing ATG8-ATG3 in proximity of the PE pool present in the autophagosomal membranes. The LC3-PE conjugate, also known as LC3-II, is present in the inner as well as at the outer surface of the expanding autophagosomal membrane. Prior to the autophagosome fusion with lysosomes, the LC3-II pool on the outer autophagosomal membrane is released from PE by ATG4 for reuse (Kabeya *et al.*, 2000; Yang and Klionsky, 2010; He and Klionsky, 2009). The fact that a subpopulation of LC3 is present in the interior

of autophagosomes makes this protein the ideal marker to monitor the generation of these carriers and their subsequent fusion with lysosomes (Kabeya *et al.*, 2000). In particular LC3 subcellular distribution can be analyzed by fluorescence microscopy and the redistribution of this protein from the cytoplasm into punctuate structures, for example, the autophagosomes, is an indicator of autophagy induction. Furthermore it allows assessing autophagy induction and progression by analyzing the LC3-II protein levels by Western blot (Klionsky *et al.*, 2012; Mizushima, 2004; Rubinsztajn *et al.*, 2009).

Generation of the Autophagosomes: The Membranes

While enormous progresses have been done regarding the characterization of the ATG machinery, the origin of the autophagosomal membranes is still not clear. Membranes from multiple sources have been shown to contribute to the formation and maturation of autophagosomes; therefore, this topic is the subject of intense debate and research. Amongst these there are several organelles, including the plasma membrane (Ravikumar *et al.*, 2010), early endosomes (Razi *et al.*, 2009), the mitochondria (Hailey *et al.*, 2010), recycling endosomes (Longatti and Tooze, 2012; Puri *et al.*, 2013), the Golgi complex (Mari *et al.*, 2011; Orsi *et al.*, 2010) and the ER (Axe *et al.*, 2008; Hayashi-Nishino *et al.*, 2009; Yla-Anttila *et al.*, 2009a; Figure 1). In fact, it was shown that most of the autophagosomes form in close association with the ER, in a structure called the omegasome (Axe *et al.*, 2008; Hayashi-Nishino *et al.*, 2009; Yla-Anttila *et al.*, 2009a). The omegasome is a PtdIns3P rich subdomain of the ER that is formed upon starvation. It is a dynamic structure where ATG proteins required for the nucleation step assemble and which gives rise to phagophores and then autophagosomes (Axe *et al.*, 2008; Hayashi-Nishino *et al.*, 2009; Yla-Anttila *et al.*, 2009a). The autophagosomes, however, do not bud or form directly from the ER but rather originate in close proximity of this organelle, possibly in regions that most likely are in close contact with mitochondrial membranes, that is, the so-called mitochondria-associated membranes (Hamasaki *et al.*, 2013).

Maturation and Consumption of Autophagosomes

Once complete, autophagosomes fuse either directly with lysosomes or with endosomes to form an intermediate organelle called amphisome. Amphisomes ultimately fuse with lysosomes too. These events lead to progressive maturation of autophagosomes into a degradative compartment, the autolysosome, through the acquisition of low pH and lysosomal hydrolases, which access the autophagosomal cargo and mediate its turnover (He and Klionsky, 2009; Figure 1). The precise mechanism of how the autophagosome fuses with the lysosomes has not yet been fully resolved. However, members of both the Rab GTPase and the SNARE (soluble N-ethylmaleimide-sensitive-factor attachment protein receptor) protein families as well as a low intralysosomal pH have been shown to be involved in fusion (Bento *et al.*, 2013; Moreau *et al.*, 2013; Yamamoto *et al.*, 1998). During the autophagosome fusion events, only the outer lipid bilayer of

these double-membrane carriers fuses with the lysosomal limiting membrane and as a result the inner autophagosomal vesicle is also degraded within the autolysosome. The resulting material is finally transported by lysosomal membrane permeases out of the autolysosome lumens into the cytoplasm, where it can be metabolized by the cell (He and Klionsky, 2009; Figure 1).

For more efficient fusion between them, autophagosomes and lysosomes have to come into close proximity. This is accomplished on the one hand through the active transport of the autophagosomes via microtubules from the cell periphery to the microtubule organization center (MTOC) (Bento *et al.*, 2013). On the other hand, lysosomal positioning within the cell is regulated in a nutrient-dependent manner by mTOR. That is, under growing conditions lysosomes bind to a kinesin complex that mediated their positioning at the cell periphery. Under starvation, an increase in the intracellular pH provokes an inhibition of the recruitment of the kinesin complex to the lysosomal membrane and thereby allowing lysosomes to move to the MTOC (Bento *et al.*, 2013).

Autophagosomes not only fuse with compartments of the endo-lysosomal system. Recent findings support the notion that these carriers could also fuse with the plasma membrane (Deretic *et al.*, 2012). This nondegradative function of autophagy plays a role in the unconventional secretion (i.e., secretion that does not involve transport through Golgi apparatus) of a subset of signal molecules (Deretic *et al.*, 2012). Diverting of autophagosomes from their conventional trajectory to the endo-lysosomal system is also something implemented by certain pathogens, mainly bacteria, to supply their intracellular inclusions with nutrients and thus fuel their replication and differentiation (Amer, 2013).

Acknowledgments

Fulvio Reggiori is supported by ECHO (700.59.003), ALW Open Program (821.02.017 and 822.02.014), DFG-NWO cooperation (DN82-303), and ZonMW VICI (016.130.606) grants.

See also: Cell Division/Death: Apoptosis: Autophagy. Cell Division/Death: Regulation of Cell Growth: mTORC1: Upstream and Downstream. Organelles: Structure and Function: Conventional and Secretory Lysosomes; Early Endosomal Compartments; Golgi and TGN; The Late Endosome

References

- Amer, A.O., 2013. The many uses of autophagosomes. *Autophagy* 9, 633–634.
- Axe, E.L., Walker, S.A., Maniava, M., *et al.*, 2008. Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum. *Journal of Cell Biology* 182, 685–701.
- Bento, C.F., Puri, C., Moreau, K., Rubinsztajn, D.C., 2013. The role of membrane-trafficking small GTPases in the regulation of autophagy. *Journal of Cell Science* 126, 1059–1069.
- Boyle, K.B., Randow, F., 2013. The role of 'eat-me' signals and autophagy cargo receptors in innate immunity. *Current Opinion in Microbiology* 16, 339–348.

- Chan, S.N., Tang, B.L., 2013. Location and membrane sources for autophagosome formation — from ER-mitochondria contact sites to Golgi-endosome-derived carriers. *Molecular Membrane Biology* 30, 394–402.
- Choi, A.M., Ryter, S.W., Levine, B., 2013. Autophagy in human health and disease. *New England Journal of Medicine* 368, 651–662.
- Codogno, P., Mehrpour, M., Proikas-Cezanne, T., 2012. Canonical and non-canonical autophagy: Variations on a common theme of self-eating? *Nature Reviews. Molecular Cell Biology* 13, 7–12.
- De Duve, C., Wattiaux, R., 1966. Functions of lysosomes. *Annual Review of Physiology* 28, 435–492.
- Deretic, V., Jiang, S., Dupont, N., 2012. Autophagy intersections with conventional and unconventional secretion in tissue development, remodeling and inflammation. *Trends in Cell Biology* 22, 397–406.
- Fengsrud, M., Erichsen, E.S., Berg, T.O., Raiborg, C., Seglen, P.O., 2000. Ultrastructural characterization of the delimiting membranes of isolated autophagosomes and amphisomes by freeze-fracture electron microscopy. *European Journal of Cell Biology* 79, 871–882.
- Hailey, D.W., Rambold, A.S., Satpute-Krishnan, P., et al., 2010. Mitochondria supply membranes for autophagosome biogenesis during starvation. *Cell* 141, 656–667.
- Hamasaki, M., Furuta, N., Matsuda, A., et al., 2013. Autophagosomes form at ER-mitochondria contact sites. *Nature* 495, 389–393.
- Hayashi-Nishino, M., Fujita, N., Noda, T., et al., 2009. A subdomain of the endoplasmic reticulum forms a cradle for autophagosome formation. *Nature Cell Biology* 11, 1433–1437.
- He, C., Klionsky, D.J., 2009. Regulation mechanisms and signaling pathways of autophagy. *Annual Review of Genetics* 43, 67–93.
- Isakson, P., Holland, P., Simonsen, A., 2013. The role of ALFY in selective autophagy. *Cell Death and Differentiation* 20, 12–20.
- Kabeya, Y., Mizushima, N., Ueno, T., et al., 2000. LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. *EMBO Journal* 19, 5720–5728.
- Kang, R., Zeh, H.J., Lotze, M.T., Tang, D., 2011. The Beclin 1 network regulates autophagy and apoptosis. *Cell Death and Differentiation* 18, 571–580.
- Klionsky, D.J., Abdalla, F.C., Abeliovich, H., et al., 2012. Guidelines for the use and interpretation of assays for monitoring autophagy. *Autophagy* 8, 445–544.
- Klionsky, D.J., Cregg, J.M., Dunn Jr, W.A., et al., 2003. A unified nomenclature for yeast autophagy-related genes. *Developmental Cell* 5, 539–545.
- Kraft, C., Martens, S., 2012. Mechanisms and regulation of autophagosome formation. *Current Opinion in Cell Biology* 24, 496–501.
- Kuma, A., Hatano, M., Matsui, M., et al., 2004. The role of autophagy during the early neonatal starvation period. *Nature* 432, 1032–1036.
- Levine, B., Kroemer, G., 2008. Autophagy in the pathogenesis of disease. *Cell* 132, 27–42.
- Levine, B., Mizushima, N., Virgin, H.W., 2011. Autophagy in immunity and inflammation. *Nature* 469, 323–335.
- Longatti, A., Tooze, S.A., 2012. Recycling endosomes contribute to autophagosome formation. *Autophagy* 8, 1682–1683.
- Mari, M., Tooze, S.A., Reggiori, F., 2011. The puzzling origin of the autophagosomal membrane. *F1000 Biology Reports* 3, 25.
- Melendez, A., Tallozy, Z., Seaman, M., et al., 2003. Autophagy genes are essential for dauer development and life-span extension in *C. elegans*. *Science* 301, 1387–1391.
- Mizushima, N., 2004. Methods for monitoring autophagy. *International Journal of Biochemistry & Cell Biology* 36, 2491–2502.
- Mizushima, N., Levine, B., 2010. Autophagy in mammalian development and differentiation. *Nature Cell Biology* 12, 823–830.
- Mizushima, N., Levine, B., Cuervo, A.M., Klionsky, D.J., 2008. Autophagy fights disease through cellular self-digestion. *Nature* 451, 1069–1075.
- Mizushima, N., Ohsumi, Y., Yoshimori, T., 2002. Autophagosome formation in mammalian cells. *Cell Structure and Function* 27, 421–429.
- Moreau, K., Renna, M., Rubinshtein, D.C., 2013. Connections between SNAREs and autophagy. *Trends in Biochemical Sciences* 38, 57–63.
- Obara, K., Sekito, T., Niimi, K., Ohsumi, Y., 2008. The Atg18-Atg2 complex is recruited to autophagic membranes via phosphatidylinositol 3-phosphate and exerts an essential function. *Journal of Biological Chemistry* 283, 23972–23980.
- Ohsumi, Y., Mizushima, N., 2004. Two ubiquitin-like conjugation systems essential for autophagy. *Seminars in Cell & Developmental Biology* 15, 231–236.
- Orsi, A., Polson, H.E., Tooze, S.A., 2010. Membrane trafficking events that partake in autophagy. *Current Opinion in Cell Biology* 22, 150–156.
- Orsi, A., Razi, M., Dooley, H.C., et al., 2012. Dynamic and transient interactions of Atg9 with autophagosomes, but not membrane integration, are required for autophagy. *Molecular Biology of the Cell* 23, 1860–1873.
- Pattingre, S., Tassa, A., Qu, X., et al., 2005. Bcl-2 antiapoptotic proteins inhibit Beclin 1-dependent autophagy. *Cell* 122, 927–939.
- Puri, C., Renna, M., Bento, C.F., Moreau, K., Rubinshtein, D.C., 2013. Diverse autophagosome membrane sources coalesce in recycling endosomes. *Cell* 154, 1285–1299.
- Ravikumar, B., Moreau, K., Jahreiss, L., Puri, C., Rubinshtein, D.C., 2010. Plasma membrane contributes to the formation of pre-autophagosomal structures. *Nature Cell Biology* 12, 747–757.
- Razi, M., Chan, E.Y., Tooze, S.A., 2009. Early endosomes and endosomal coatome are required for autophagy. *Journal of Cell Biology* 185, 305–321.
- Reggiori, F., Klionsky, D.J., 2013. Autophagic processes in yeast: Mechanism, machinery and regulation. *Genetics* 194, 341–361.
- Reggiori, F., Tucker, K.A., Stromhaug, P.E., Klionsky, D.J., 2004. The Atg1-Atg13 complex regulates Atg9 and Atg23 retrieval transport from the pre-autophagosomal structure. *Developmental Cell* 6, 79–90.
- Rieter, E., Vinke, F., Bakula, D., et al., 2013. Atg18 function in autophagy is regulated by specific sites within its beta-propeller. *Journal of Cell Science* 126, 593–604.
- Rubinshtein, D.C., Cuervo, A.M., Ravikumar, B., et al., 2009. In search of an “autophagometer”. *Autophagy* 5, 585–589.
- Russell, R.C., Tian, Y., Yuan, H., et al., 2013. Ulk1 induces autophagy by phosphorylating Beclin-1 and activating VPS34 lipid kinase. *Nature Cell Biology* 15, 741–750.
- Shintani, T., Klionsky, D.J., 2004. Autophagy in health and disease: A double-edged sword. *Science* 306, 990–995.
- Shpilka, T., Mizushima, N., Elazar, Z., 2012. Ubiquitin-like proteins and autophagy at a glance. *Journal of Cell Science* 125, 2343–2348.
- Simonsen, A., Tooze, S.A., 2009. Coordination of membrane events during autophagy by multiple class III PI3-kinase complexes. *Journal of Cell Biology* 186, 773–782.
- Wong, E., Cuervo, A.M., 2010. Integration of clearance mechanisms: The proteasome and autophagy. *Cold Spring Harbor Perspectives in Biology* 2, a006734.
- Yamamoto, A., Tagawa, Y., Yoshimori, T., et al., 1998. Bafilomycin A1 prevents maturation of autophagic vacuoles by inhibiting fusion between autophagosomes and lysosomes in rat hepatoma cell line, H-4-II-E cells. *Cell Structure and Function* 23, 33–42.
- Yamamoto, H., Kakuta, S., Watanabe, T.M., et al., 2012. Atg9 vesicles are an important membrane source during early steps of autophagosome formation. *Journal of Cell Biology* 198, 219–233.
- Yang, Z., Klionsky, D.J., 2010. Mammalian autophagy: Core molecular machinery and signaling regulation. *Current Opinion in Cell Biology* 22, 124–131.
- Yla-Anttila, P., Vihinen, H., Jokitalo, E., Eskelinen, E.L., 2009a. 3D tomography reveals connections between the phagophore and endoplasmic reticulum. *Autophagy* 5, 1180–1185.
- Yla-Anttila, P., Vihinen, H., Jokitalo, E., Eskelinen, E.L., 2009b. Monitoring autophagy by electron microscopy in Mammalian cells. *Methods in Enzymology* 452, 143–164.
- Yoshimori, T., 2007. Autophagy: Paying Charon's toll. *Cell* 128, 833–836.

Further Reading

- Choi, A.M., Ryter, S.W., Levine, B., 2013. Autophagy in human health and disease. *New England Journal of Medicine* 368, 651–662.
- Hamasaki, M., Shibutani, S.T., Yoshimori, T., 2013. Up-to-date membrane biogenesis in the autophagosome formation. *Current Opinion in Cell Biology* 25, 455–460.
- Kraft, C., Martens, S., 2012. Mechanisms and regulation of autophagosome formation. *Current Opinion in Cell Biology* 24, 496–501.
- Lamb, C.A., Yoshimori, T., Tooze, S.A., 2013. The autophagosome: Origins unknown, biogenesis complex. *Nature Reviews Molecular Cell Biology* 14, 759–774.

Relevant Websites

- <http://www.tanpaku.org/autophagy/>
Autophagy Database.
- <http://autophagy.lu/index.html>
Human Autophagy Database.
- <http://www.landesbioscience.com/journals/autophagy/>
Landes Bioscience.

Peroxisomes

PB Lazarow, Institut Pasteur, Paris, France

© 2016 Elsevier Inc. All rights reserved.

Introduction and Morphology

The peroxisome is a beautiful organelle to behold, whether glowing in the fluorescence microscope as a sea of green dots (Figure 1(A)), or showing off its finely granular matrix, sometimes with a paracrystalline core, in the electron microscope (Figure 1(B,C)), or visibly green to the human eye when purified from rat liver (green due to the high concentration of catalase). Most peroxisomes are spherical or ovoid, bounded by one membrane bilayer, and approximately 100–200 nm in diameter. In some cells they are 5 or 10 times bigger, for example, in hepatocytes, and sometimes they are tubular in form. A peroxisome reticulum of interconnected organelles was suggested from serial section analyses of peroxisomes with tails and tubular elements (Lazarow *et al.*, 1980) and has been seen in living COS-7 cells by real-time imaging (Schrader *et al.*, 2000) and in regenerating rat liver (Yamamoto and Fahimi, 1987). In exceptional cases, (e.g., the sebaceous preputial gland) the peroxisomes form large cup-shaped elements interconnected into elaborate branching structures (Gorgas, 1984; Figure 1(E)).

Behind its usually simple appearance, the peroxisome conceals a remarkable diversity of physiological roles and biochemical functions. These vary from plants to fungi to animals, from one tissue to another, and even from one cell type to another within a tissue. The organelle adapts its biochemical composition and its abundance to the needs of the cell and to changing environmental conditions. Nevertheless, there are some functional characteristics common to all peroxisomes (Wanders and Waterham, 2006a; Palma *et al.*, 2009; Islinger *et al.*, 2010; Hu *et al.*, 2012).

We describe its major physiological roles and discuss how it evolved its diverse biochemical abilities. We look at how peroxisomes multiply and adapt, how they are formed, remodeled, and degraded, and at the mechanisms that control this. We end with the human diseases that occur when the peroxisome has a defective enzyme or is not assembled correctly. For reasons of space, most references in this article are necessarily to reviews.

Some Important, Specialized Physiological Functions of Peroxisomes

In the Nervous System: Myelin Formation and Maintenance

Peroxisomes are required for the normal formation of myelin. The predominant phospholipids of myelin are plasmalogens, which contain an unsaturated long-chain alcohol in the place of one of the two fatty acids of the usual phospholipids (Figure 2(a)). Plasmalogens are thought to defend against reactive oxygen toxicity (thanks to their vinyl ether linkage) and to alter membrane flexibility and permeability. Plasmalogen synthesis begins obligatorily in peroxisomes (Braverman

and Moser, 2012; Figure 2(b)). Human patients without this function suffer defective or delayed myelination. Maintaining axonal integrity and myelin sheaths requires the function of oligodendrocyte peroxisomes (Baes and Aubourg, 2009). Plasmalogens are present in many other tissues as well.

Conversion of Fat to Sugar, Supplying Energy to Germinating Plant Seedlings

Many seeds contain a store of fat (in the form of triacylglycerols in lipid droplets) that serves the energy requirements of growing seedlings until they acquire photosynthetic capacity. To do so, fatty acids are cleaved from the triacylglycerols and undergo β -oxidation to form acetyl-CoA, which is then largely converted to glucose via the glyoxylate cycle. In plants, this β -oxidation occurs exclusively in peroxisomes, not in mitochondria. The acetyl-CoA end product formed within peroxisomes is used right where it is produced by glyoxylate cycle enzymes (see details and illustrations below) to form succinate, which is converted to fumarate by mitochondria and then to glucose in the cytosol. If one of the peroxisomal enzymes of β -oxidation or the glyoxylate cycle is defective, plant growth is greatly impaired (Hu *et al.*, 2012; Kunze *et al.*, 2006). Seedling peroxisomes also have been called 'glyoxysomes.'

Sperm Formation

Spermiogenesis depends upon peroxisomes. In the testis, peroxisomes are required in Sertoli cells, are abundant in Leydig cells (where they probably contribute to testosterone formation), and are present in germ cells but not in mature spermatozoa. The testis and spermatozoa are rich in plasmalogens and the testis contains an abundance of the peroxisomal enzymes required for plasmalogen synthesis. Mutations inactivating a single peroxisome enzyme required for either plasmalogen formation in mice or very long-chain fatty acid oxidation in humans cause impaired spermiogenesis and infertility. In fruit flies too, sperm development arrests prematurely without peroxisomes (Van Veldhoven and Baes, 2013).

Photorespiration in Green Leaves

Peroxisomes play a central role in photorespiration, a salvage pathway that rescues a 2-carbon by-product of photosynthesis. In the Calvin cycle of photosynthesis, the bifunctional chloroplast enzyme ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) uses its carboxylase activity to productively attach CO₂ to the 5-carbon substrate, ribulose-1,5-bisphosphate. The 6-carbon product is split into two 3-phosphoglycerate molecules, one of which goes to make glucose and the other continues through the Calvin cycle to regenerate ribulose-1,5-bisphosphate.

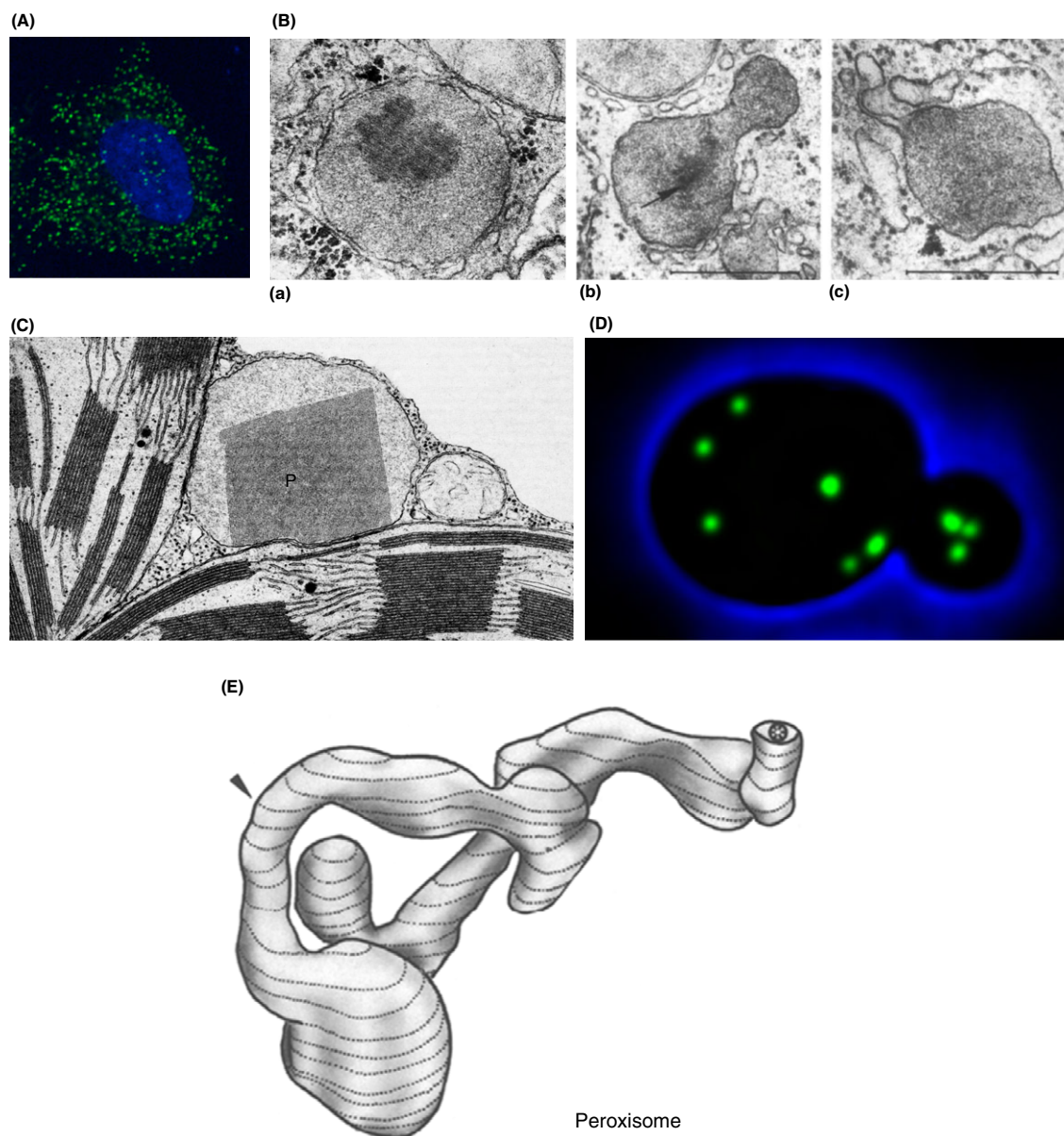


Figure 1 Images of peroxisomes, morphological variations. (A) HeLa cell immunofluorescence with green fluorescent protein (GFP)-labeled peroxisomes (by Gaelle Boncompain and Agathe Subtil, Institut Pasteur). (B) Rat liver electron microscopy (by Helen Shio, Rockefeller University). (a) A peroxisome with a paracrystalline core (de Duve, 1983); (b) a dumbbell-shaped peroxisome (dividing?) and (c) a peroxisome with a tubular tail (Lazarow and Fujiki, 1985). (C) Tobacco leaf electron microscopy: a peroxisome with a large trapezoidal paracrystalline core between two chloroplasts (de Duve, 1983). (D) *Saccharomyces cerevisiae* grown on glucose with GFP-labeled peroxisomes (by Richard Rachubinski, University of Alberta). (E) Mouse preputial gland, reconstruction of the elaborate peroxisome compartment within one cell from multiple serial sections (Gorgas, 1984).

However, the second activity of RubisCO, its oxygenase activity, makes mischief to varying extents, cleaving an unproductive 2-carbon molecule, glycolate phosphate, from ribulose-1,5-bisphosphate. The wasteful loss of this fragment is avoided by photorespiration, which, via an elaborate collaboration of chloroplasts, peroxisomes, mitochondria, and the cytosol, converts this glycolate phosphate into 3-phosphoglycerate to be productively metabolized in the Calvin cycle (Figure 3). Six of the key reactions, includ-

ing oxidation of glycolate to glyoxylate by an H_2O_2 -producing oxidase (see Figure 4) and transamination of the glyoxylate to glycine, occur in peroxisomes (Hu *et al.*, 2012).

Glycolysis in Trypanosomes

One of the most unusual functions sometimes found in peroxisomes is glycolysis. Glycolysis occurs in the cytosol in almost

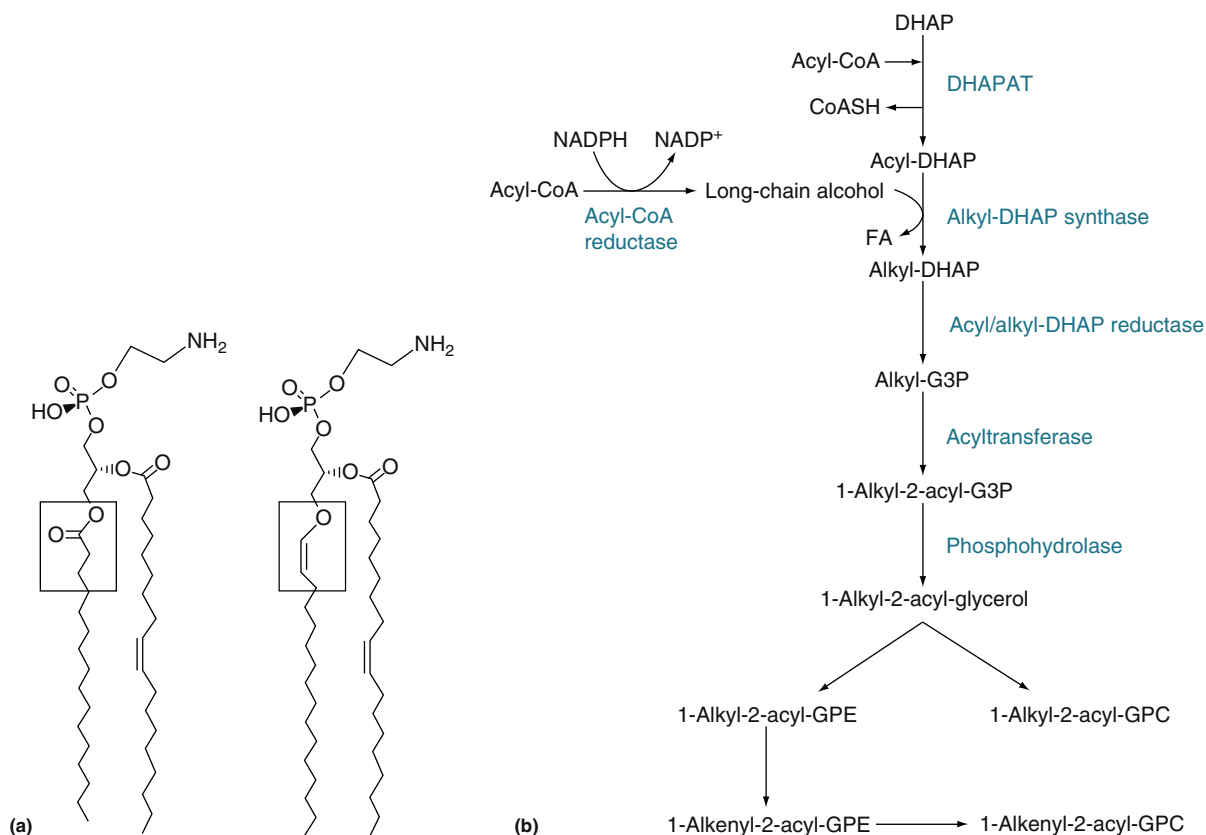


Figure 2 Plasmalogen structure and biosynthesis. (a) Structure. (Left) conventional phosphatidyl ethanolamine, (right) the plasmalogen version of phosphatidyl ethanolamine (Boncompain *et al.*, 2014). (b) Biosynthesis – a collaborative effort of peroxisomes and ER. The reactions in the top half occur in peroxisomes, those in the bottom half in the ER. DHAP, dihydroxyacetone phosphate; DHAPAT, DHAP acyltransferase; FA, fatty acid; G3P, glycerol-3-phosphate; GPC, glycerol phosphatidyl choline; GPE, glycerol phosphatidyl ethanolamine; NADP, nicotinamide adenine dinucleotide phosphate (Wanders and Waterham, 2006a).

all organisms. The exceptional compartmentation of the glycolytic enzymes within an organelle, first discovered in *Trypanosoma brucei* (Oppenheimer and Borst, 1977), allows this parasitic flagellated protozoan to efficiently metabolize the glucose that it absorbs from the blood of its host. Originally named a glycosome, and still referred to as such, this highly specialized peroxisome has subsequently been found in other trypanosomatids and in all kinetoplastids studied (Gualdron-Lopez *et al.*, 2012). Some of these organisms are human parasites, responsible for sleeping sickness, Chagas disease, and leishmaniasis; their glycolytic enzymes are interesting as potential drug targets. Moreover, some key enzymes of glycolysis are partially localized in peroxisomes in many fungi, thanks to cryptic peroxisome targeting sequences (Freitag *et al.*, 2012).

Plugging Holes in Fungi

The Woronin body is another highly specialized type of peroxisome, which seals the septal pores of filamentous fungi in response to wounding of the plasma membrane. This avoids cytoplasmic bleeding and allows time for the cell to reseal its plasma membrane. Thus the Woronin body plays a critical

function in the maintenance of cellular integrity. It is formed by budding from conventional peroxisomes (Liu *et al.*, 2008).

Conserved Biochemical Pathways Present in Most Peroxisomes

See reviews of conserved and specialized peroxisome functions (Wanders and Waterham, 2006a; Palma *et al.*, 2009; Islinger *et al.*, 2010; Hu *et al.*, 2012).

Reactive Oxygen and Nitrogen Metabolism

Hydrogen peroxide metabolism – A simple respiratory pathway for which peroxisomes were named by Christian de Duve

Peroxisomes were first isolated from rat liver and characterized by Christian de Duve and his collaborators. The enzymes that they identified then were oxidases and catalase (de Duve and Baudhuin, 1966). The oxidases use molecules of oxygen (O_2) to oxidize uric acid, D-amino acids, and L- α -hydroxyacids with the formation of hydrogen peroxide (H_2O_2), which is then decomposed by the catalase (Figure 4). Subsequently, many

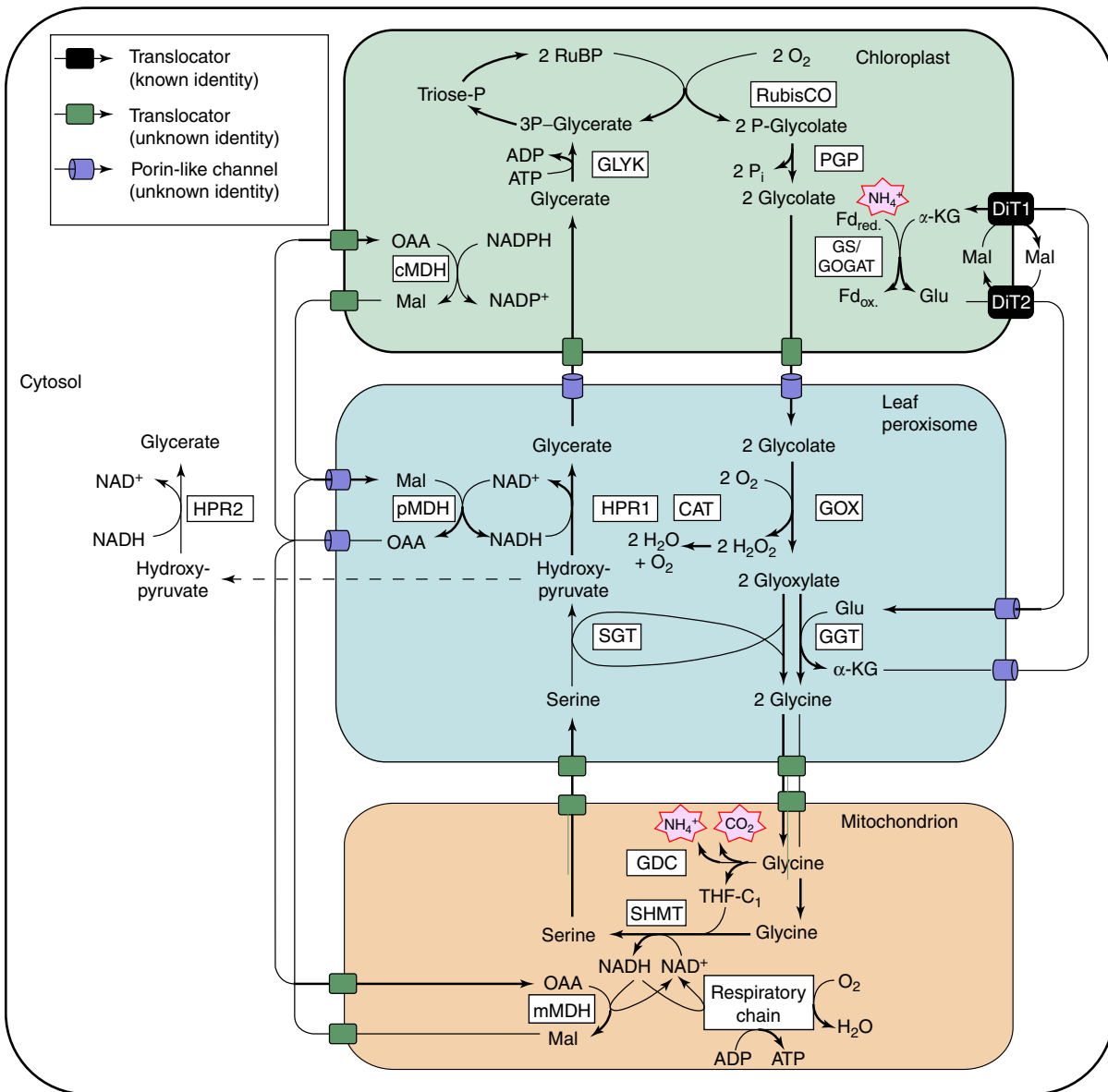


Figure 3 Photorespiration. Collaboration of four cell compartments. Metabolite abbreviations: Glu, glutamate; KG, ketoglutarate; Mal, malate; OAA, oxaloacetate; RuBP, ribulose-bisphosphate; THF, tetrahydrofolate. Enzyme abbreviations: CAT, catalase; GDC, glycine decarboxylase; GGT, glutamate:glyoxylate aminotransferase; GLYK, glycerate kinase; GOGAT, glutamate:oxoglutarate aminotransferase; GOX, glycolate oxidase; GS, glutamate synthase; HPR, hydroxypyruvate reductase; MDH, malate dehydrogenase; PGP, phosphoglycolate phosphatase; SGT, serine:glyoxylate aminotransferase; SHMT, serine hydroxymethyl transferase (Hu *et al.*, 2012).

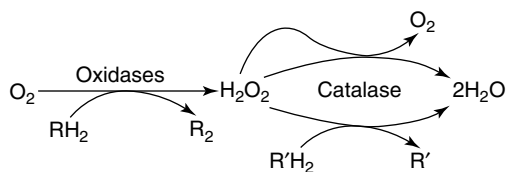


Figure 4 Peroxisomal respiration via H₂O₂. See text for R and R' substrates (de Duve and Baudhuin, 1966).

other oxidases have been discovered in peroxisomes, acting on diverse acyl-CoAs (including very long-chain and branched fatty acyl-CoAs), polyamines, pipercolic acid, sarcosine, and, in

plants, sulfite. A methanol oxidase permits some yeast, for example, *Hansenula polymorpha* and *Pichia pastoris*, to grow on methanol as sole carbon source.

Catalase decomposes the H₂O₂ by two alternative mechanisms: the 'catalatic' dismutation of 2H₂O₂ to 2H₂O + O₂ or via the peroxidatic oxidation of substrates such as ethanol: H₂O₂ + ethanol to acetaldehyde + H₂O. The peroxidatic reaction predominates at low rates of H₂O₂ production and the catalatic reaction takes over when H₂O₂ is produced more rapidly. These enzymes constitute a simple respiratory pathway. The energy of the oxidations is not conserved in the form of adenosine triphosphate (ATP) (or otherwise) but rather is dissipated as heat.

Protection against H_2O_2 /use of H_2O_2

H_2O_2 is toxic and can damage organic molecules. Thus cells are protected by putting the enzymes that make large amounts of H_2O_2 in the same compartment with catalase that destroys it. In peroxisomes, one molecule of H_2O_2 is produced with each turn of the fatty acid β -oxidation spiral, with each cycle of photorespiration and every time an oxidase functions.

Does H_2O_2 get out of peroxisomes? Although some experiments report a substantial leakage of H_2O_2 from rat liver peroxisomes (Boveris *et al.*, 1972), others predict that not more than 2% of the H_2O_2 escapes (Poole, 1975). Catalase has first order reaction kinetics, meaning the rate of reaction is proportional to the substrate concentration: it goes faster and faster as the generation of H_2O_2 increases, without reaching a plateau. These kinetics should avoid substantial leakage. In the case of germinating seedlings, which produce substantial H_2O_2 during early growth, the peroxisomes are equipped with a second defense system against H_2O_2 , which consists of ascorbate peroxidase and monodehydroascorbate reductase (SDP2) (Hu *et al.*, 2012).

On the other hand, some H_2O_2 is useful. H_2O_2 can modify the activity of signaling molecules, for example, by oxidizing cysteine residues, thereby influencing cell proliferation, migration, and differentiation (Rhee, 2006). Cell signaling by peroxisome-derived H_2O_2 is described in both plants (Del Rio, 2011; Mhamdi *et al.*, 2012) and animals (Fransen *et al.*, 2012). H_2O_2 is also employed in cellular defense, for example, in pathogen killing by neutrophils, where it is compartmentalized within phagosomes.

Other reactive oxygen and nitrogen species

Other reactive oxygen and nitrogen species are produced in peroxisomes. Xanthine oxidase in mammalian peroxisomes makes superoxide (O_2^-) as well as H_2O_2 and can reduce nitrates and nitrites to nitric oxide (NO). The inducible form of NO synthase is partly localized in peroxisomes and makes both O_2^- and NO. Hydroxyl radicals ($\bullet OH$) are formed non-enzymatically by the Fenton reaction between H_2O_2 and Fe^{2+} . In plants, NO and S-nitrosoglutathione are found in peroxisomes and may contribute to signaling, for example induction of defense genes by nitrosoglutathione. These reactive compounds affect the redox state within peroxisomes and, indirectly, that of other cell compartments.

Peroxisomes also contain enzymes that protect against reactive oxygen and nitrogen species, including superoxide dismutase, peroxiredoxin 5, glutathione S-transferases, and epoxide hydrolase as well as catalase. The balance between production and destruction of reactive oxygen and nitrogen molecules, in different cells and tissues, under different physiological and pathological conditions, determines the positive and negative results.

Fatty Acid β -Oxidation

A highly conserved pathway of fatty acid β -oxidation is found in almost all peroxisomes (Figure 5(a)). In plants, the peroxisome is the sole organelle capable of this catabolism. This is also true in many fungi (*Saccharomyces cerevisiae*, *Candida tropicalis*, *P. pastoris*, and *Yarrowia lipolytica*), which use these

peroxisomal reactions to grow on fatty acids. However, a bioinformatics analysis of 57 fungal species for which complete genome sequences were available showed that other fungal species also contain the genes for a second β -oxidation system in their mitochondria (Shen and Burger, 2009).

In animals, mitochondrial β -oxidation was well known long before peroxisomal β -oxidation was discovered. We now know that the pathway occurs in both organelles, but with different specificity, different regulation, and different gene products. Mammalian peroxisomes not only oxidize the usual long-chain fatty acids, but in addition oxidize a variety of substrates that the mitochondrial β -oxidation cannot: very long-fatty acids (i.e., C24, C26, and up), branched-chain fatty acids, and polyunsaturated fatty acids. Peroxisomal β -oxidation participates in the anabolic (biosynthetic) formation of bile acids by shortening the side chain of cholesterol. Peroxisomes oxidize about one-third of the palmitic acid in rat liver (Kondrup and Lazarow, 1985). If the peroxisomal pathway is defective, many of its substrates accumulate with toxic consequences.

Mitochondrial β -oxidation is coupled to oxidative phosphorylation and is well-adapted to generate ATP. The acetyl-CoA that is formed can be further oxidized in the Krebs cycle to make more ATP. In contrast, peroxisomal β -oxidation dissipates some energy as heat and the acetyl-CoA formed is likely used for biosynthetic purposes. In plants, protozoa and fungi, it is often used for gluconeogenesis via the glyoxylate cycle. In animals, it may be used for the synthesis of cholesterol and steroids and in fungi is used to make polyketide toxins.

The enzymes that catalyze the four reactions of the β -oxidation spiral in peroxisomes are distinct from their mitochondrial counterparts (Figure 5(a)). The first peroxisomal reaction produces H_2O_2 ; a variety of oxidases with different substrate specificities catalyze this reaction. The next two reactions are catalyzed by a bifunctional enzyme (there are two such enzymes with different stereochemistry) and the fourth reaction is catalyzed by several thiolases, again with differing substrate specificities. Additional enzymes are involved: ligases that conjugate substrates to Coenzyme A, a racemase, isomerases, carnitine transferases, and membrane transporters that bring in acyl-CoAs and export products (Figure 5(b)). Altogether, 17 gene products are involved in humans, not counting transporters (Wanders and Waterham, 2006a).

The Glyoxylate Cycle and Glyoxylate Metabolism

The glyoxylate cycle allows 2-carbon compounds to be converted to a 4-carbon molecule, succinate, which in turn can be converted to sugar or to amino acids or can replenish the Krebs cycle. This permits *de novo* gluconeogenesis from fatty acids, which occurs in some plant tissues (as described earlier for germinating seedlings) and in protozoa and fungi. It also allows the utilization of ethanol and acetate as sole carbon sources by yeasts. The two key enzymes of the cycle attach acetyl-CoA first to glyoxylate (to form malate, which is oxidized to oxaloacetate) and then to oxaloacetate (to form citrate). In the last step, isocitrate is cleaved to produce the product, succinate, and reform glyoxylate, for which the cycle is named (Figure 6(a)).

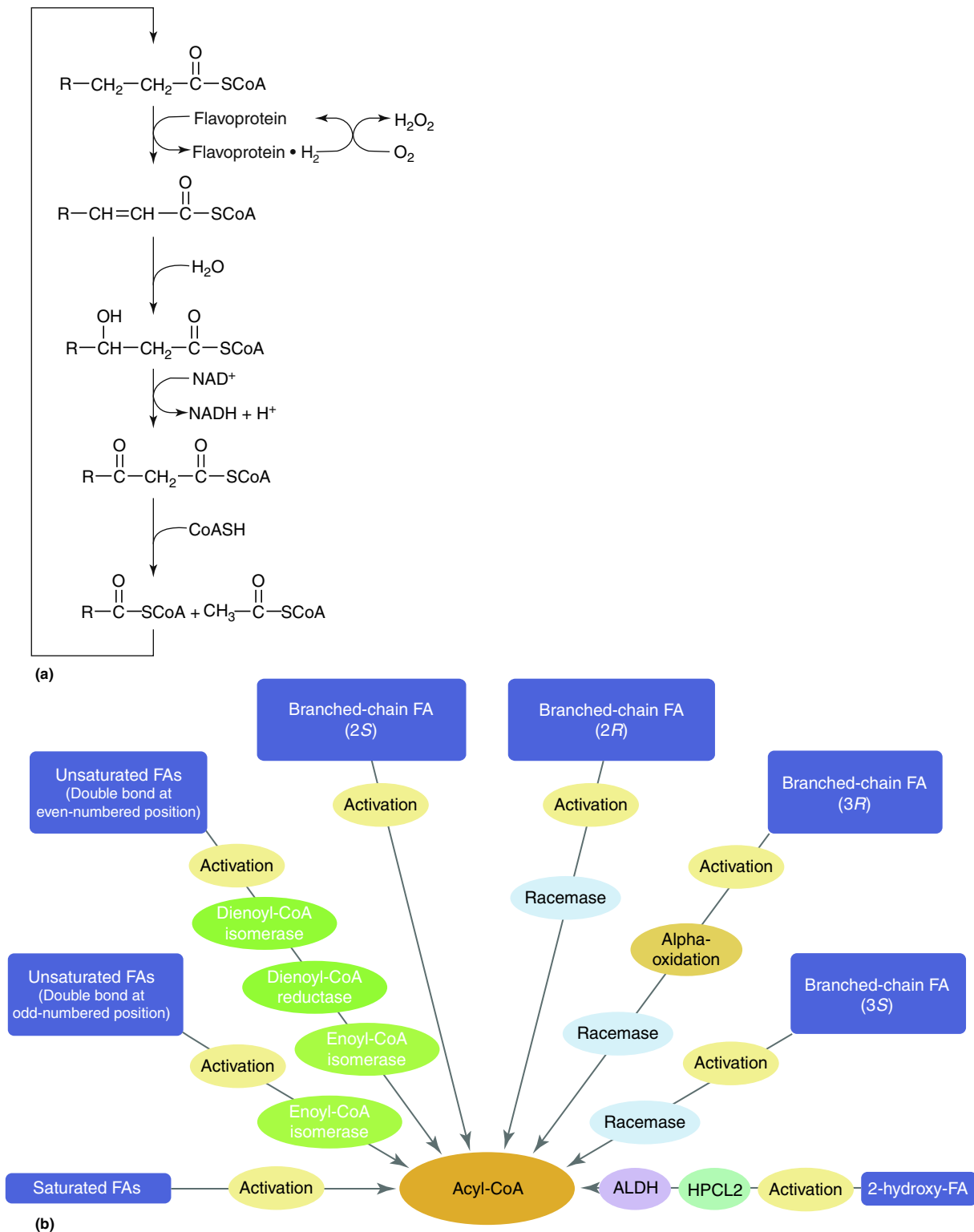


Figure 5 Peroxisomal fatty acid β -oxidation. (a) Principal reactions (Lazarow and de Duve, 1976). (b) Auxiliary reactions for diverse substrates (Wanders and Waterham, 2006a).

Formerly it was thought that all five enzymes of the cycle are located together inside peroxisomes (also called 'glyoxysomes'). In fact, only the two key enzymes, the malate and

citrate synthases, are always inside peroxisomes, which permits them to utilize directly the acetyl-CoA produced by β -oxidation (Kunze *et al.*, 2006). The malate dehydrogenase and

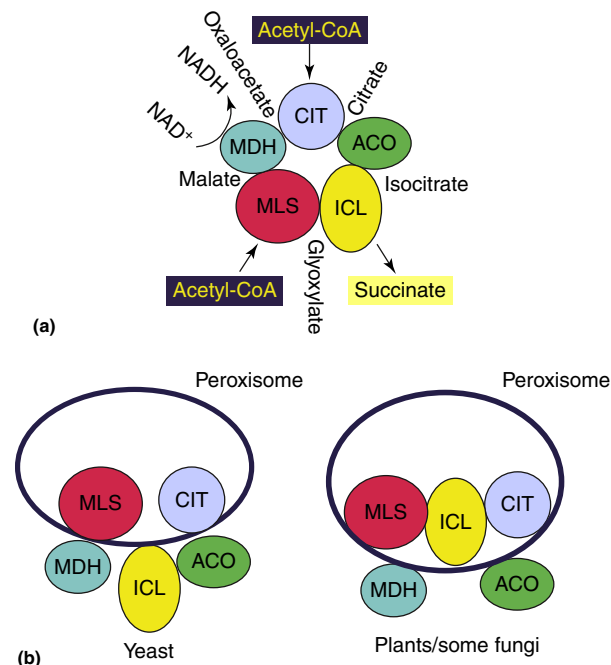


Figure 6 Glyoxylate cycle. (a) Reactions. Enzyme abbreviations: ACO, aconitase; CIT, citrate synthetase; ICL, isocitrate lyase; MDH, malate dehydrogenase; MLS, malate synthetase. (b) Compartmentation (Kunze *et al.*, 2006).

aconitase and sometimes the isocitrate lyase are found outside, in the cytosol (Figure 6(b)). This may serve to protect them from damage by H_2O_2 , to which some are sensitive, but it requires the efficient transport of the cycle intermediates into and out of the peroxisomes.

Mammals lack the glyoxylate cycle altogether (despite a few claims to the contrary) but glyoxylate is nevertheless produced by peroxisomal glyoxylate oxidase, and converted to glycine by an aminotransferase. Failure to do this conversion causes hyperoxaluria type 1 – details below.

Other Peroxisome Functions

Polyamine Catabolism

The conversion of spermine to spermidine and then to putrescine occurs in peroxisomes in both animals and plants.

Purine Catabolism

Most mammals and some plants degrade purines to allantoin (by means of xanthine oxidase and urate oxidase); other organisms degrade allantoin further to urea and glyoxylate. Urate oxidase is peroxisomal in both vertebrates and plants (e.g., *Arabidopsis thaliana*) but has been lost altogether in primates, making us susceptible to gout.

Fatty Acid α -Oxidation

Some fatty acids (e.g., with a methyl group on the 3rd carbon) require α -oxidation before entering the β -oxidation spiral. An

example is dietary phytanic acid (some of which is derived from chlorophyll). Peroxisomes are the unique site of this α -oxidation.

Bile Acid Synthesis

Peroxisomes catalyze the last steps in bile acid synthesis, including a β -oxidative shortening of the side chain of cholesterol.

Cholesterol, Steroids, Isoprenoids, and Wax Ester Synthesis

Peroxisomes are very abundant in cells and tissues that synthesize cholesterol, steroids, isoprenoids, and wax esters, suggesting the possibility that they participate in these metabolic pathways. There are differences of opinion on whether the enzymes of cholesterol biosynthesis are located in peroxisomes. It is certain that peroxisomes can make the long-chain alcohols that are found in wax esters.

Biotin Synthesis

Plants and some fungi synthesize biotin; the first steps require peroxisomes and the later steps occur in mitochondria (Maruyama *et al.*, 2012).

Fungal Sex

Great diversity is observed in the sexual reproduction of fungi. Mechanisms of meiosis, formation of reproductive structures, mating, germination, and dispersal of spores and mobilization of energy reserves to support these processes all vary. Peroxisomes are involved in many different ways (Peraza-Reyes and Berteaux-Lecellier, 2013).

Signaling

Peroxisomes contribute to the synthesis and/or degradation of signaling molecules in both plants and animals. For example, leukotrienes and prostaglandins are degraded by peroxisome β -oxidation.

Plant hormones (phytohormones)

Auxins (primarily indole-3-acetic acid) regulate cell growth and cell division and affect many aspects of plant development, including root system development and stamen elongation. Jasmonates (derived from fatty acids) regulate the response to stresses (e.g., wounding or infection by pathogens) and also affect plant development. Both jasmonates and auxins require peroxisomes for their synthesis, primarily utilizing peroxisomal enzymes of β -oxidation.

Platelet activating factor

Platelet activating factor (1-O-alkyl-2-acetyl-sn-glycero-3-phosphocholine) potentially mediates diverse intercellular interactions. It signals leukocyte activation in inflammation, is involved in allergic reactions and aggregates platelets to arrest bleeding. It is an alkyl phospholipid, similar to the plasmalogens described above, and thus requires peroxisomes for its

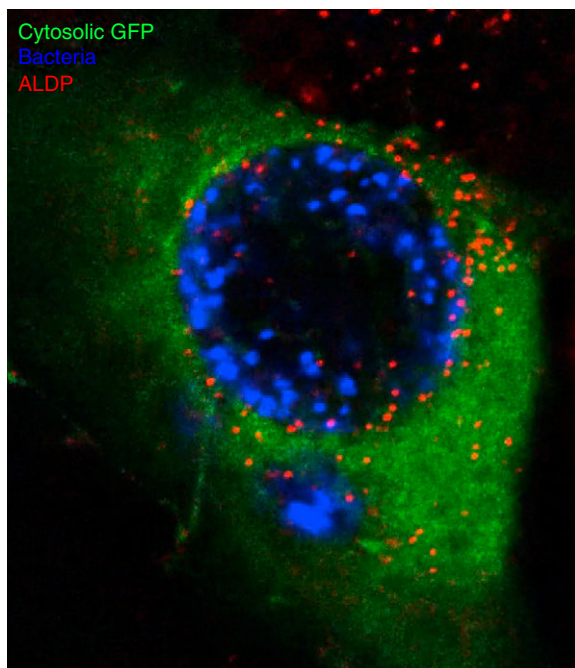


Figure 7 Chlamydia (blue) growing inside an inclusion body in human cells with some peroxisomes (red) imported into the inclusion; GFP in the cytosol (Boncompain *et al.*, 2014).

synthesis. However, most platelet activating factor is rapidly made when needed by the remodeling of preexisting alkyl phospholipids: an arachidonoyl moiety is removed by a specific phospholipase and replaced with an acetyl group; these reactions occur downstream from the initial formation in peroxisomes.

Pathogen Defense and Pathogen Offense

Exploitation of peroxisomes by a bacterium, *Chlamydia trachomatis*

Chlamydia trachomatis, an obligate intracellular pathogen that is responsible for loss of eyesight and sexually transmitted diseases, exploits many parts of the human host cell, including peroxisomes. The bacteria develop within a membrane-bounded inclusion; remarkably they cause peroxisomes to enter inside this inclusion (Figure 7). The growing bacteria incorporate plasmalogens and some of these plasmalogens contain bacteria-specific odd-chain fatty acids, indicating a metabolic collaboration of peroxisomes and bacteria for their formation (Boncompain *et al.*, 2014).

Antibiotic synthesis by fungi

Penicillin and cephalosporin, two β -lactam antibiotics that have been invaluable in defending humans against diverse pathogens, are synthesized in *Penicillium chrysogenum* by a collaboration of cytosolic enzymes with enzymes inside peroxisomes. The peroxisomal enzymes that help make penicillin, acyl-CoA ligases, and acyl-transferase, belong to the family of fatty acid metabolizing enzymes. The location of this part of the pathway within the organelle is important: in cells with

aberrant or missing peroxisomes, much less penicillin is produced. The fungal peroxisomal epimerases of cephalosporin synthesis have remarkable sequence similarity to mammalian peroxisomal enzymes involved in bile acid synthesis, and are related to peroxisomal enzymes of branched-chain fatty acid metabolism (Martin *et al.*, 2012).

Toxin synthesis by fungi

Aspergillus species aggress plants with polyketide mycotoxins, for example, aflatoxin, that are synthesized from acetyl-CoA produced in peroxisomes. The AK-toxin that attacks Japanese pears is made by peroxisomal enzymes.

Plant immune defense against fungi

In *Arabidopsis*, a peroxisomal glycosyl hydrolase (PEN2) is part of the innate immune response and is essential to the defense against noninvasive (nonadapted) fungal pathogens. Its activity and its localization within peroxisomes are both essential (Lipka *et al.*, 2005). Jasmonate signaling, mentioned above, mediates plant defense against necrotrophic pathogens.

Mammalian immune defense against viruses

Peroxisomes play an important role in mammalian immune defense against viruses. RIG-I-like receptors, which are soluble RNA helicases, detect viral RNA in the cytosol. These receptors then interact with the mitochondrial antiviral signaling (MAVS) protein that is located on both peroxisomal and mitochondrial membranes (Figure 8). The peroxisomal MAVS signals the cell to switch on an immediate and rapid antiviral response. In complementary fashion, the mitochondrial MAVS signals, via interferon production, a slower but sustained antiviral response. Both are necessary: for example, vesicular stomatitis virus can interfere with and escape the interferon-mediated response, but is susceptible to the fast peroxisome-mediated response. The peroxisomal response is accompanied by a striking change in morphology, with aggregation and the formation of peroxisomal tubules (Dixit *et al.*, 2010).

Exploitation of peroxisomes by viruses

Tombusviruses replicate their single-stranded RNA genome in invaginations of the peroxisomal membrane within plant cells. Rotavirus targets its spike protein, VP4, to peroxisomes. Other viral proteins interact with peroxisomal proteins, for example, HIV's Nef and influenza's NS1. It appears likely that the viruses exploit the lipid metabolism of peroxisomes, but further studies are necessary to elucidate these intriguing observations (Lazarow, 2011).

Metabolic Cooperation and Interaction with Other Organelles

For 40 years, peroxisomes have been observed in close physical proximity to other organelles, including lipid droplets, endoplasmic reticulum (ER), mitochondria, and chloroplasts (Figure 9). Apposition of the membranes facilitates transfer and exchange of metabolites and supports the multiple metabolic collaborations among these organelles: peroxisomes and lipid droplets for fatty acid metabolism in seedlings and adipocytes; peroxisomes and the ER to synthesize plasmalogens;

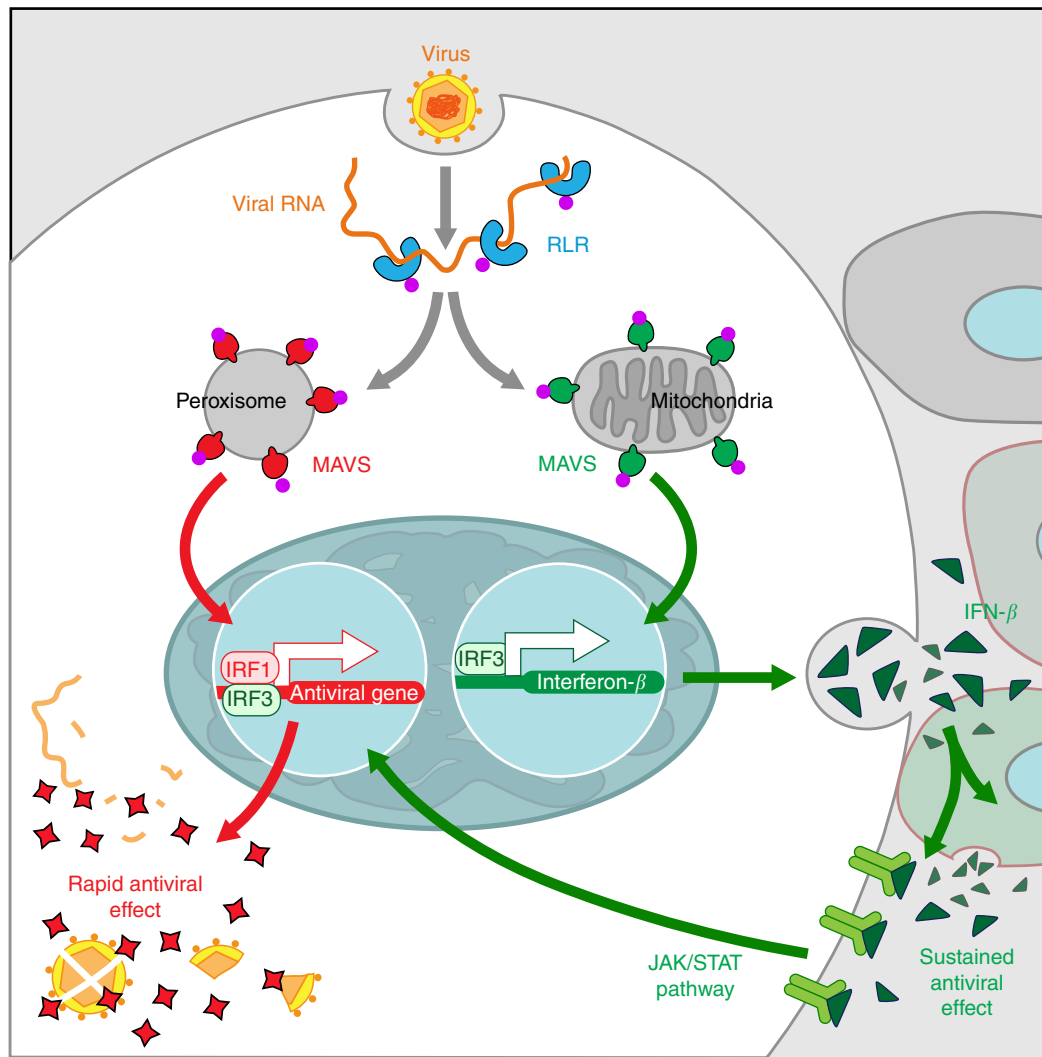


Figure 8 Peroxisome role in immune defense. Viral RNA in mammalian cells is detected by RIG-I and signaled to MAVS on peroxisomes and on mitochondria. The peroxisomal MAVS switches on a signaling pathway that produces a rapid antiviral response. Complementarily, mitochondrial MAVS uses other signaling pathways to produce interferon- β and a sustained antiviral response (Dixit *et al.*, 2010). JAK, janus kinase; STAT, signal transducer and activator of transcription.

peroxisomes, chloroplasts, and mitochondria in photorespiration; and peroxisomes, mitochondria, and the cytosol in gluconeogenesis. Peroxisome and mitochondria are further linked by shared functions in reactive oxygen metabolism and the acetyl-CoA formed in peroxisomes can be further oxidized via the Krebs cycle in mitochondria. Fission and inheritance of these two organelles share some common factors and vesicles produced from mitochondria can target peroxisomes. Moreover, there is reciprocal regulation: retrograde signaling of the mitochondrion to the nucleus causes a marked proliferation of peroxisomes in *S. cerevisiae* and defects in one have been shown to produce changes in the other (Thoms *et al.*, 2009). The old morphological observations of proximity have been refined recently by high-throughput genetic screens, which found that peroxisomes are located near two specific subdomains of mitochondria: one enriched in the pyruvate dehydrogenase complex and another that is the point of contact of mitochondria with ER (Cohen *et al.*, 2014). The screens also identified

mutants in which peroxisome membranes are assembled but matrix enzymes are not imported – these are the peroxisomal membrane ghosts first identified in Zellweger syndrome (Santos *et al.*, 1988b) – see human diseases below.

Peroxisome Abundance, Adaptability, and Regulation

The peroxisome population in cells is dynamic and its abundance and form is the result of multiple interacting processes: rate of formation (influenced by environmental factors via transcriptional regulation), division (affecting size, number, and shape), transport and segregation into daughter cells, and turnover by autophagy.

Abundance and Adaptability

One of the remarkable things about peroxisomes is their ability to adapt in abundance and enzymatic capability to

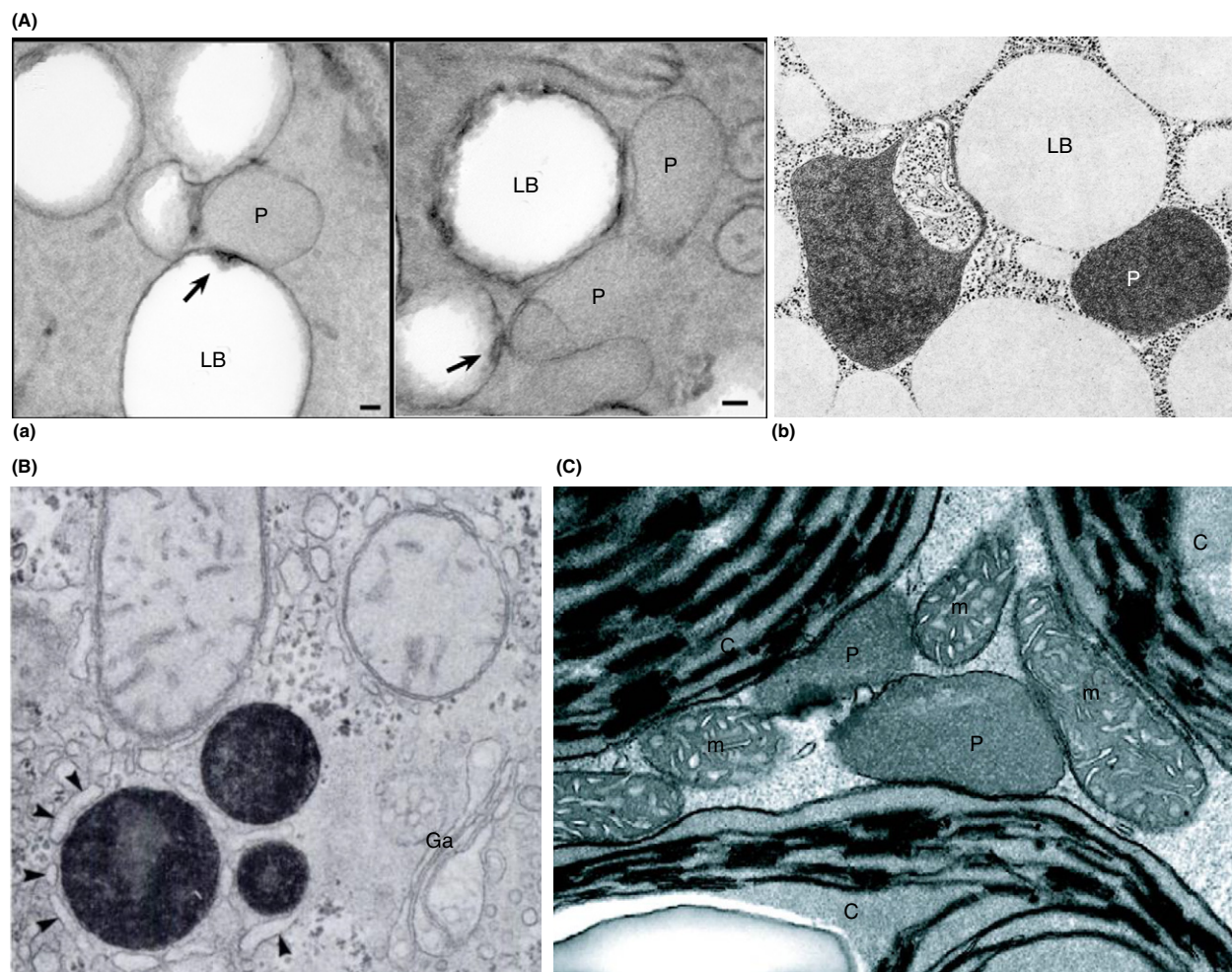


Figure 9 Peroxisome interactions with other organelles. (A) With lipid bodies. (a) In *S. cerevisiae*. LB, lipid droplet; P, peroxisome (Binns *et al.*, 2006); (b) in tomato seedling (de Duve, 1983). (B) With ER in rat liver. Peroxisomes are reacted with diaminobenzidine to visualize catalase. Arrowheads indicate ER. Ga, Golgi apparatus (Shio and Lazarow, 1981). (C) With mitochondria (m) and chloroplasts (c) in *Arabidopsis*. (Kaur *et al.*, 2009).

different circumstances. Diet, drugs, light, and temperature all influence the peroxisome population, along with the developmental needs of the organism.

An example is the yeast, *H. polymorpha*: it contains one or a few tiny peroxisomes when grown on glucose. Shift it to methanol medium and large peroxisomes develop with a new complement of enzymes adapted to methanol assimilation (Figure 10(A)). *Saccharomyces cerevisiae* adapts to a shift from glucose medium to oleic acid medium by a proliferation of peroxisomes containing upregulated β -oxidation enzymes. *Pichia pastoris* peroxisomes adapt appropriately to either methanol or fatty acids in the medium.

Light influences plant development and differentiation. When *Arabidopsis* is exposed to light and begins photosynthesis, peroxisomes that were oxidizing fat adapt to do photorespiration. Reactive oxygen species, UV radiation, and salt stress also induce plant peroxisome proliferation.

Fungi adapted to produce large amounts of penicillin have more peroxisomes than other strains (Figure 10(B)).

Cold temperatures induce peroxisomes in brown fat, a site of thermogenesis, and in some fishes, probably be-

cause peroxisomal oxidative metabolism dissipates energy as heat.

Many hypolipidemic drugs cause a striking proliferation of peroxisomes in rat liver (Figure 10(C)), an observation that led to the discovery of β -oxidation in animal peroxisomes (Lazarow and de Duve, 1976). Peroxisomal β -oxidation activity and peroxisome abundance are induced as much as 10-fold by these drugs.

Gene Regulation

Transcription factors control these responses (Schrader *et al.*, 2012; Figure 10(D)). In *S. cerevisiae*, two proteins with fatty acid-binding domains, Pip2 and Oaf1, form a heterodimer and switch on genes with oleate response elements. An additional zinc finger protein that senses carbon status, Adr1, and a kinase, Snf1, also participate. In a filamentous fungus, *Aspergillus nidulans*, FarA and FarB bind fatty acids and do the job.

In *Arabidopsis*, a photoreceptor, phytochrome A, translocates into the nucleus when stimulated by far red light and

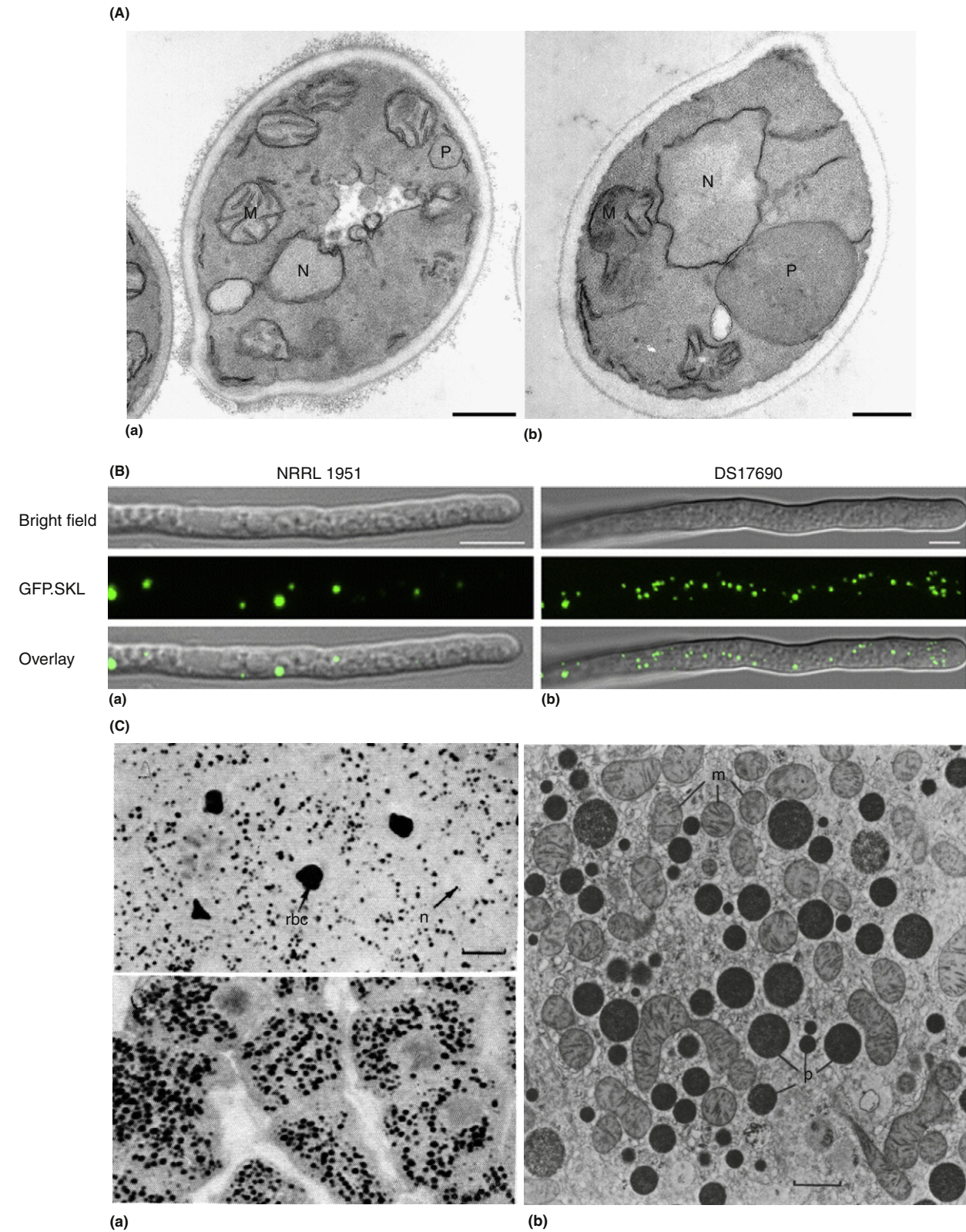


Figure 10 Peroxisome abundance induced. (A) *Hansenula polymorpha* grown on methanol (b) vs. glucose (a). M, mitochondrion; N, nucleus; P, peroxisome (Veenhuis and van der Klei, 2014). (B) *Penicillium chrysogenum* producing large amounts of penicillin (b) vs. normal strain (a) (Bartoszewski et al., 2011). (C) Rat liver after hypolipidemic drug administration. (a) Control (top) vs. Wy-14 643 (bottom), light microscopy, peroxisomes stained with the diaminobenzidine reaction. rbc, red blood cell (Lazarow, 1977); (b) after chlorophenoxyisobutyrate (CPIB) administration, electron microscopy. M, mitochondrion; P, peroxisome (Lazarow and de Duve, 1976). (D) Transcriptional regulation in (a) mouse, (b) *S. cerevisiae*, (c) *A. nidulans*, and (d) *Arabidopsis* (Schrader et al., 2012).

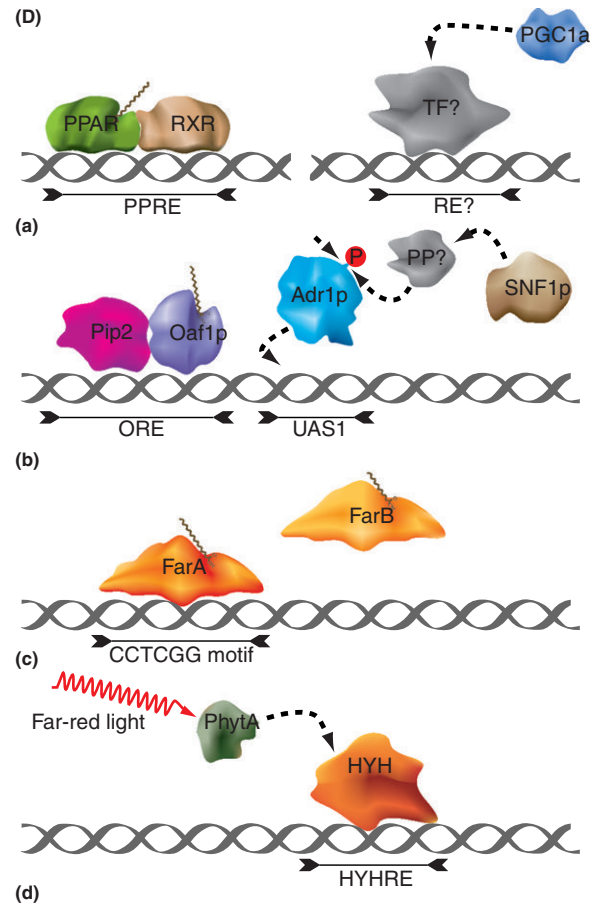


Figure 10 Continued.

interacts with transcription factors, for example, HYH, to switch on genes.

In mammals, the receptor for unsaturated long-chain fatty acids and hypolipidemic drugs is PPAR α (peroxisome proliferator activated receptor- α). It is in the class of steroid and retinoid receptors and dimerizes with RXR α (retinoid X receptor- α) to form the active DNA-binding transcription factor that upregulates a variety of genes, including the peroxisomal β -oxidation genes.

Peroxisome Division, Transport, and Inheritance

Peroxisome Division

Peroxisomes proliferate and divide by a conserved process, usually consisting of elongation or tubulation followed by constriction and fission (Saraya *et al.*, 2010; Schrader *et al.*, 2012; Figure 11(A)). The progeny may have the same size and composition or the daughter peroxisome may be smaller than the mother peroxisome (Figure 11(B)). Some membrane proteins may appear preferentially in the elongation and some matrix enzymes may be retained in the mother peroxisome, resulting in an asymmetric division of the content. An extreme example of asymmetric division is the formation of Woronin

bodies by budding from a conventional peroxisome (Figure 11(C)). Peroxisome division is controlled and catalyzed by two families of proteins, Pex11 proteins (specific to peroxisomes) and dynamin-related proteins (large GTPases with diverse roles in cell membrane fission and fusion).

Pex11 controls peroxisome proliferation and abundance: it is physiologically upregulated when peroxisomes are induced and its experimental overexpression produces many more, smaller peroxisomes. If knocked out, there is generally one or a few of the organelles. In *S. cerevisiae*, Pex11's activity is regulated by phosphorylation. Many organisms have several Pex11-related proteins with somewhat different functions. In the case of *Arabidopsis* there are five: overexpression of Pex11c causes tubulation; overexpression of Pex11e doubles the number of peroxisomes without visible elongation (Figure 11(D)).

Pex11(s) catalyze the first steps in division. It contains amphipathic helices that insert into the peroxisome membrane and cause bending that results in the elongation of the membrane. Pex11 then accumulates at or near the site where fission will occur and probably helps to recruit other proteins there, for example, Fis1, an integral membrane protein that anchors the protein(s) that catalyze fission.

The next steps, constriction and scission, are catalyzed by the dynamin-like proteins (DLP): Dlp1 in mammals, Dnm1

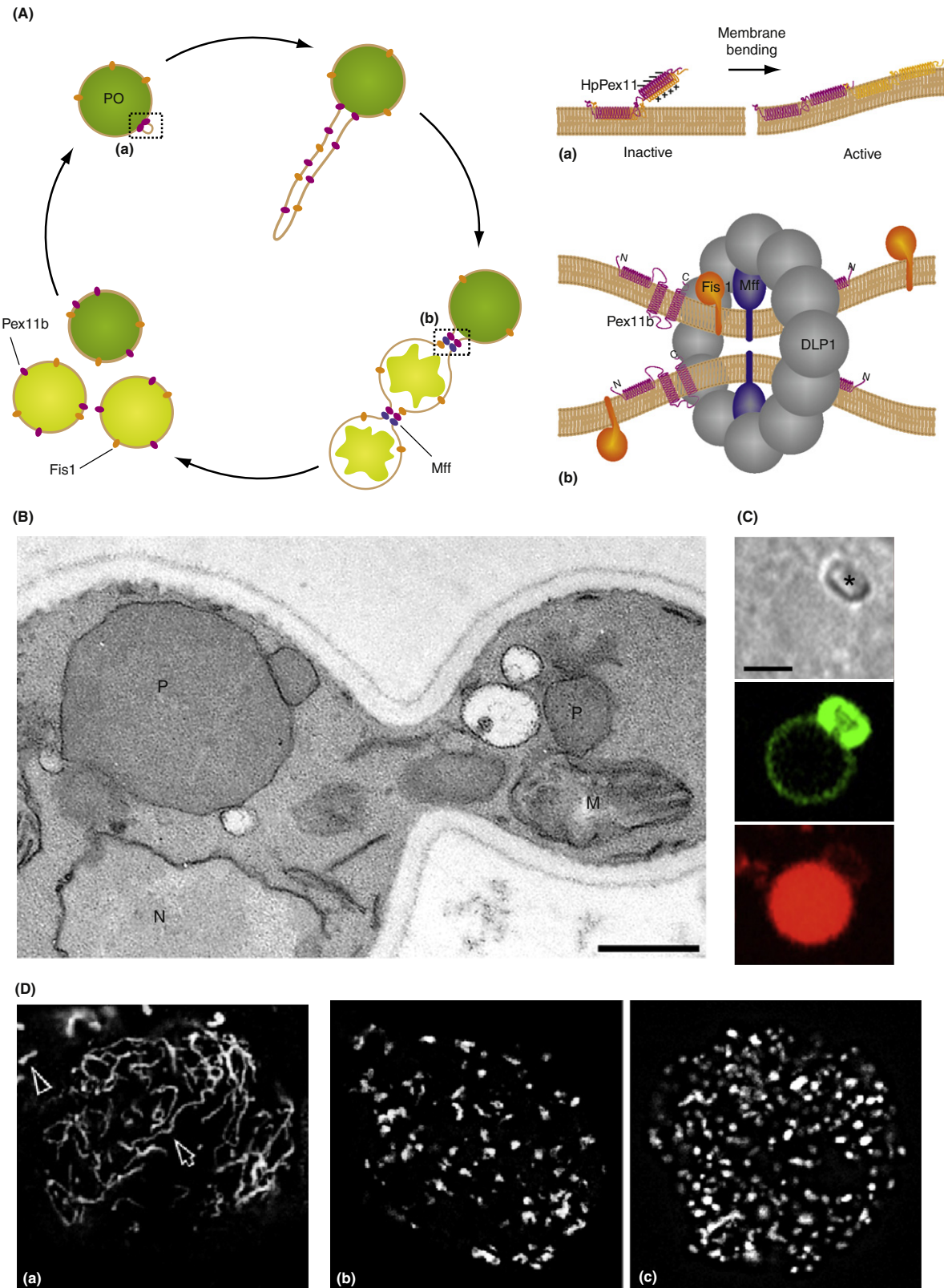


Figure 11 Peroxisome division. (A) Model of division. (a) membrane bending and elongation via Pex11 insertion; (b) constriction and scission by DLP with the energy of GTP hydrolysis (Schrader *et al.*, 2012). (B) Asymmetric division in *H. polymorpha* producing one large and one small daughter peroxisome (Veenhuis and van der Klei, 2014). (C) Woronin body budding. Woronin sorting complex protein (WSC, in green) in the bud (asterisk in the bright field image) and red fluorescent protein with a peroxisome targeting signal (in red) remaining in the original peroxisome (by Gregory Jedd, National University of Singapore). (D) Diverse Pex11 effects in *Arabidopsis*. (a) Tubulation by Pex11c; (b), control; (c), proliferation without visible elongation by Pex11e (Lingard and Trelease, 2006). (E) Peroxisome fission mutants in *S. cerevisiae* grown on oleic acid. (a) Wild type; (b) *vps1* mutant; (c) *dnm1/vps1* double mutant; (d) incomplete fission in a *dnm1/vps1* double mutant; and (e) a 3D reconstruction of incomplete fission from a stack of confocal images (Kuravi *et al.*, 2006).

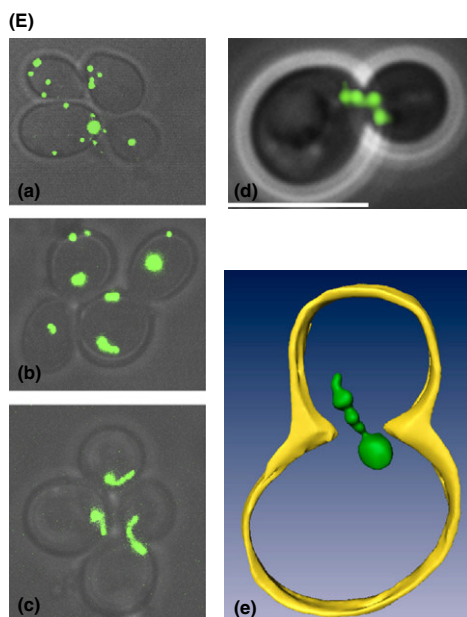


Figure 11 Continued.

and Vps1 in yeast, and several DRP isoforms in plants. They are recruited to the fission site by Fis1 and other auxiliary proteins and form a ring around the membrane elongation (Figure 11(A)). Remarkably, the same mechanism is used for other organelles: Vps1 also catalyzes mitochondrial fission and DRP5B also catalyzes chloroplast fusion. Defects in scission cause a variety of results, including fewer but larger peroxisomes or sometimes an elongated, partially divided peroxisome poised between mother and daughter in budding yeast (Figure 11(E)).

Peroxisome Fusion?

In contrast to mitochondria, peroxisomes appear not to fuse; certainly they lack the machinery used by mitochondria for this purpose. Experiments mixing *in vivo* peroxisomes labeled differently failed to detect any mixing of their contents (Huybrechts *et al.*, 2009). This suggests that the elaborate interconnected peroxisome networks in the preputial gland (Figure 1(E)) are due to a lack of fission rather than to fusion, or else that fusion is possible in certain cell types or under special metabolic circumstances (Schrader *et al.*, 2012).

Inheritance of Peroxisomes – Tethering and Guided Transport

When budding yeast divide, elaborate mechanisms ensure that the growing bud, and hence the daughter cell, receives all the cell organelles: at least one mitochondrion, some ER and Golgi, etc. Peroxisomes are transported along actin cables into the bud, pulled by the class V myosin motor, Myo2, which is attached to the peroxisomal membrane by the Inp2 protein (Fagarasanu *et al.*, 2010). Other peroxisomes remain in the

mother cell, firmly anchored in place by the Inp1 protein, which is attached to the peroxisomal integral membrane protein Pex3 (Figure 12). If transport is defective, cytokinesis is delayed. Coordination between anchoring and transport ensures appropriate partitioning of the organelle. In the case of impaired fission, one sometimes sees the partially divided peroxisome straddling the gap between the two cells (Figure 11(E)); eventually the peroxisome is pulled into two parts and the cell divides.

When mammalian cells divide, their large number is probably sufficient to partition the organelles into both daughter cells stochastically, but nevertheless there is also active transport along the cytoskeleton. One of the few big changes in peroxisome biology during evolution is that the peroxisome moves on microtubules in animals and on actin cables in yeast and plants.

Peroxisome Biogenesis

Historical Perspective

Peroxisome biogenesis has some unique features, not seen in other organelles. It has no DNA and synthesizes none of its own proteins. Our understanding of its biogenesis has changed several times and continues to evolve; some aspects remain uncertain and/or controversial.

Model 1: Formation from ER

In 1973, it was "almost textbook knowledge that peroxisomes arise as budding outgrowths of the ER. Their proteins are believed to be synthesized on membrane-bound ribosomes, to be delivered intraluminally and to be conveyed to the growing peroxisomes through ER channels" (de Duve,

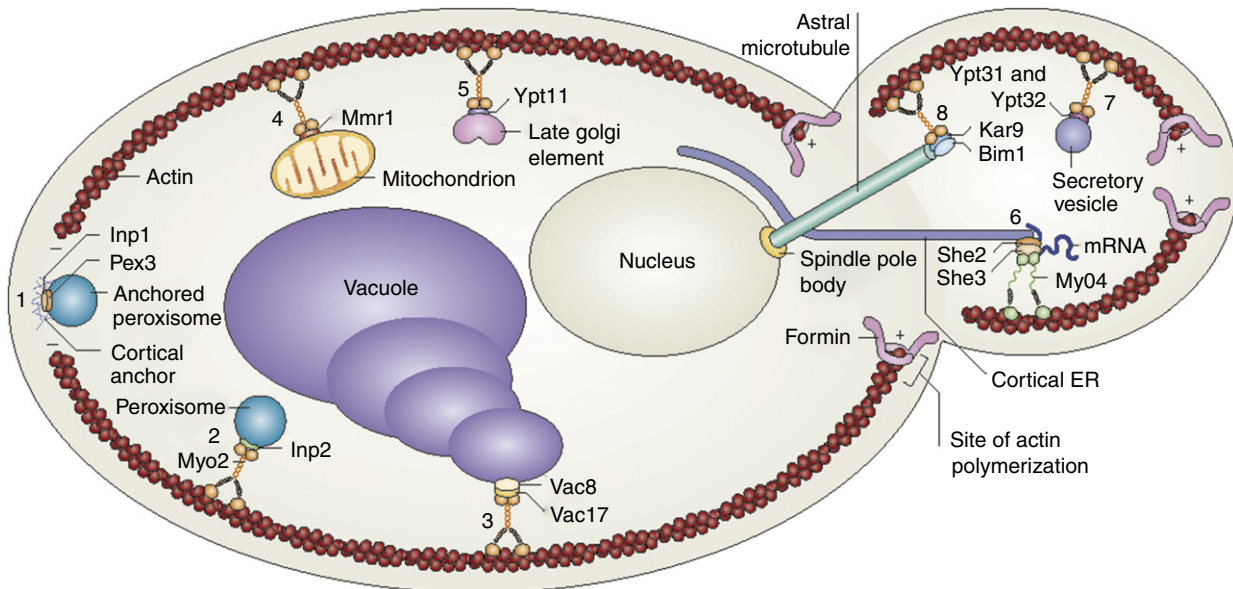


Figure 12 Model of organelle transport and inheritance via class V myosin motors in *S. cerevisiae*. Some peroxisomes travel to the bud while others remain anchored in the mother cell (Fagarasanu *et al.*, 2010).

1973; Figure 13(a)). This theory, that all peroxisome enzymes pass through the ER, as do secretory proteins, was based largely on electron microscopic studies of Novikoff and Shin (1964).

Model 2: Growth and division

The discovery that catalase never enters the ER but instead is imported directly into preexisting peroxisomes *in vivo* challenged this view (Lazarow and de Duve, 1973; Figure 13(b)). Subsequent experiments on peroxisomal proteins in animals, plants, and fungi gave similar results: many enzymes and a major membrane protein were made on free polyribosomes and released into the cytosol and could be imported into purified peroxisomes *in vitro*. Review of these data led to the hypothesis that peroxisomes form by growth and division, importing enzymes, and membrane proteins directly from the cytosol (Lazarow and Fujiki, 1985; Figure 13(b)). This model, with one important modification, a contribution of the ER to membrane growth, has been robustly confirmed by 25 years of molecular cell biology (Rucktaschel *et al.*, 2011; Nuttall *et al.*, 2011; Ma *et al.*, 2011; Dimitrov *et al.*, 2013; Smith and Aitchison, 2013; Fujiki *et al.*, 2014; Hettema *et al.*, 2014). The details continue to be refined.

Model 3: De novo membrane formation

The appearance of peroxisomes in mutant cells that appear to lack the organelle (see details below) implied the *de novo* formation of peroxisomes, with the membrane coming from the ER, followed by maturation as in model 2 (Figure 13(c)). Experiments testing the extent of the ER contribution to peroxisome biogenesis under physiological conditions have produced conflicting results and controversy, and research continues to address this question.

Current Model: Growth and Division with Some Help from the ER

Thirty-three 'PEX' genes have been identified whose products, 'peroxins,' (denoted here as Pex1, Pex2, etc.) catalyze peroxisome biogenesis (Table 1). Nineteen collaborate in the translocation of enzymes from the cytosol through the peroxisome membrane into the lumen of the organelle. Only three catalyze membrane protein assembly. The rest are involved in proliferation and division (Figures 14 and 15).

Targeting and import of soluble enzymes into peroxisomes

Two types of targeting sequence can direct peroxisome enzymes to their organelle: a C-terminal tripeptide (SKL and variants thereof) or an N-terminal nonapeptide (RLxxxxx(H/Q)I and variants thereof). These targeting sequences are recognized by soluble recycling receptors, Pex5 and Pex7, respectively, that bind the newly synthesized proteins in the cytosol and transport them to peroxisomes. The translocation machinery consists of a docking complex, a transient pore, and a complex to reexport the recycling receptor, driven by ubiquitinylation and ATP hydrolysis (Hettema *et al.*, 2014). Remarkably, the pore can accommodate folded and oligomerized proteins. It even allows a few proteins without targeting sequences to enter 'piggyback,' i.e., bound to a properly targeted protein. Pex5 forms part of the dynamic pore (Meinecke *et al.*, 2010), whereas Pex7 can pass inside the organelle and then be reexported (Nair *et al.*, 2004). The tripeptide targeting sequence is retained in the mature imported enzymes; the N-terminal sequence is often cleaved in plants but usually retained in animals and yeasts.

Direct targeting and insertion of peroxisome membrane proteins into peroxisomes

Most membrane proteins are targeted to existing peroxisomes by internal sequences consisting of a cluster of positively

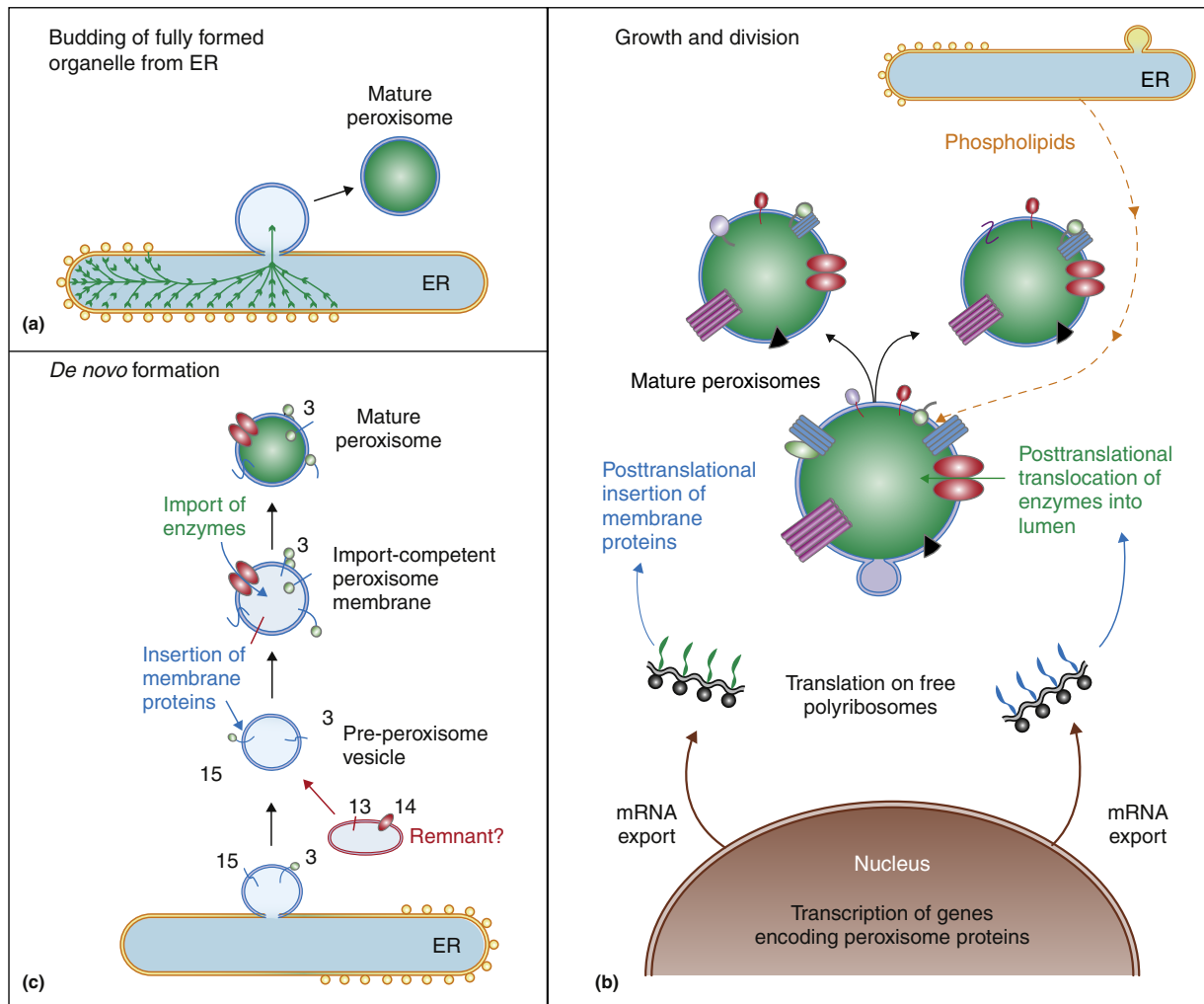


Figure 13 Peroxisome biogenesis – historical models. (a) Budding fully formed peroxisomes from the ER. (b) Growth and division. (c) *De novo* membrane formation in the ER. Numbers indicate Pex gene products.

charged amino acids near a transmembrane domain. These 'Class 1' proteins are recognized by a soluble, cytosolic, recycling receptor, Pex19, which also functions as a chaperone. Pex19, carrying its cargo, docks on the peroxisome integral membrane protein, Pex3, and the newly synthesized protein inserts into the membrane. In mammalian cells, a second integral membrane protein is required, Pex16, which recruits Pex3 to the membrane ('Class 2'). Thus far, no other components have been found necessary for membrane protein assembly. Membrane growth does not require functioning import machinery; when import is defective, membranes continue to grow and divide as empty peroxisome 'ghosts' (Santos *et al.*, 1988b; Figure 16(A)).

Modification of the model: ER contribution to peroxisome membrane growth

There is now good evidence for a second mechanism by which several membrane proteins reach peroxisomes: insertion in the ER followed by travel in vesicles to existing peroxisomes (Figure 14). In *S. cerevisiae*, epitope-tagged Pex3 expressed from a strong promoter moves from ER to

existing peroxisomes (as well as to new peroxisomes – see below) (Hoepfner *et al.*, 2005). Budding from the ER of vesicles carrying Pex3 and Pex15 has been demonstrated in a cell-free assay (Lam *et al.*, 2010). The budding reaction requires Pex19 and ATP and unknown cytosolic proteins, and is distinct from the budding of vesicles carrying secretory proteins. This vesicle traffic occurs physiologically: a multiply-tagged version of Pex15 with an engineered N-glycosylation site arrives in peroxisomes in wild type *S. cerevisiae* mostly glycosylated, indicating transit through the ER.

Pex3 also inserts directly into peroxisomes at Pex16 docking sites (Matsuzaki and Fujiki, 2008; Aranovich *et al.*, 2014), clearly indicating that it has two independent, redundant pathways to reach its organelle.

In several plant species, the normal physiological route of the APX membrane protein is first into ER and then by vesicle transport to peroxisomes, whereas other membrane proteins insert directly from the cytosol into peroxisomes (Mullen and Trelease, 2006). Thus, also in plants, peroxisome membrane growth occurs by two mechanisms.

Table 1 PEX genes, peroxin functions and human disease complementation groups

Process	Gene	Protein function	Notes	Human disease complementation group
Transport of newly synthesized enzymes through the cytosol to peroxisomes:				
PTS receptors	Pex5	Shuttling receptor for PTS1 proteins	Bifunctional amphipathic protein; coreceptor for Pex7 in mammals	US/EU 2
Co-receptors	Pex7	Shuttling receptor for PTS2 proteins	Recycles from peroxisome interior	11
	Pex18	Co-receptor for PTS2 proteins in <i>Saccharomyces cerevisiae</i>		
	Pex21	Co-receptor for PTS2 proteins in several yeasts	Structurally similar	R
	Pex20	Co-receptor for PTS2 proteins in other fungi		
Translocation of newly synthesized enzymes into peroxisomes:				
Receptor docking complex	Pex13	Site for shuttling receptors to dock on exterior of membrane	Membrane protein	13
	Pex14			
	Pex17			
Translocation through the membrane	Pex33	Functional homolog of Pex17	Membrane protein; also functions in autophagy in <i>Hansenula polymorpha</i>	H
	Pex5			
	Pex14	Formation of transient pore	Cytosolic face of membrane	2
	Pex8			
Linkage of docking and ubiquitination complexes	Pex2	Ubiquitin ligases	Integral membrane protein	15
	Pex10			
	Pex12			
	Pex4			
Ubiquitination	Pex22	Ubiquitin conjugating enzyme	Inner face of peroxisome membrane, both PTS1 and PTS2	K
	Pex1			
	Pex6			
	Pex15			
Receptor reexport	Pex26	Membrane anchor for Pex1 and Pex6 in <i>S. cerevisiae</i>	RING finger complex	F
		Membrane anchor for Pex1 and Pex6 in <i>H. sapiens</i>	Cytosolic face of membrane	B
		ATP-dependent export of receptor	Cytosolic face of membrane	
			AAA-type ATPases	
				E
				C
				A

Membrane protein assembly: Transport and insertion of newly synthesized proteins into the peroxisome membrane Transport of membrane proteins via the ER and vesiculation	PEX19 PEX3	Soluble chaperone and shuttling receptor Docking of Pex19 on the peroxisome membrane	Separate domains to bind mPTS and Pex3 Integral membrane protein, also functions in inheritance and degradation	14 12	J G
	PEX16 PEX16 PEX3 PEX19 PEX15	Insertion of Pex3 into the peroxisome membrane Recruits Pex3 in mammals May recruit other proteins Budding of vesicles containing Pex3 (and 16) unknown	Integral membrane protein, not present in yeasts	9 9 12	D D G
	In <i>S. cerevisiae</i>				
	PEX11	Proliferation and membrane elongation, recruitment of fission proteins	Amphipathic α helix	16	
	PEX25	Proliferation and elongation in yeasts; required for peroxisome reintroduction in <i>S. cerevisiae</i>	Pex11 family		
Proliferation	PEX27	Negative effector in <i>S. cerevisiae</i>			
	PEX34	Fission in <i>S. cerevisiae</i>			
	PEX23	Proliferation in <i>Yarrowia lipolytica</i>			
	PEX24				
	PEX28, 29	Proliferation in <i>S. cerevisiae</i>	Orthologs or family of PEX24 Orthologs or family of PEX23		
Omitted	PEX30, 31, 32	Proliferation, ER contact in <i>S. cerevisiae</i>			
	PEX9	None	Wrongly annotated orf		

There are conflicting data on the extent of the ER contribution and which proteins use this route. Affinities and kinetics may explain some of the differing results. Neither Pex3 nor Pex16 are seen in ER in wild type cells, but they appear there when over-expressed. If these proteins pass quickly

through the ER this could explain why they are not observed there physiologically. Or, if they normally enter peroxisomes, but the peroxisomes have a limited capacity to integrate them, this would explain the ER overexpression results. If vesicles coming from the ER transit rapidly to mature peroxisomes, they will simply donate membrane components to the peroxisomes. If the vesicles bearing Pex3 spend a long enough time in the cytoplasm, they could gradually acquire the other membrane proteins necessary for importing enzymes, and mature without fusion (see *de novo* formation below). If peroxisome membrane proteins reside long enough in the ER in some cell types, they could create specialized regions enriched in peroxisome proteins.

It is clear that there are two functional pathways for peroxisome membrane protein assembly. It is not certain whether they are fully redundant or if collaboration of the two pathways is required. More research on proteins expressed normally in wild type cells will clarify the physiological process.

Phospholipids

Since most phospholipids are synthesized in the ER, it was speculated that they are brought from the ER to growing peroxisomes by vesicular transport (Lazarow and Fujiki, 1985). This is indeed a function that can be provided by the vesicle traffic that has now been observed. This mechanism does not exclude additional vesicle-independent phospholipid transport at ER-peroxisome contact sites (Raychaudhuri and Prinz, 2008). Additional information on the quantitative contribution of each mechanism to lipid transport is desirable.

De novo Formation of Peroxisome Membranes from the ER

Evidence for de novo formation in mutants

A mutation in PEX3, PEX16, or PEX19 prevents the assembly of (most) membrane proteins, and causes the apparent disappearance of peroxisomes. Genetic complementation of the mutants results in the reappearance of peroxisomes. This process occurs by two distinct steps: first recognizable peroxisome membranes (similar to ghosts) gradually accumulate, and then they fill up with enzymes (Figure 13(c)); this takes several days in mammalian cells and several hours in yeast.

From where do the membranes arise in these mutants? In *S. cerevisiae*, Pex3 inserts into the ER and then appears in vesicles (Hoepfner *et al.*, 2005). In mammalian cells Pex16 inserts in the ER, recruits Pex3 and forms vesicles (Kim *et al.*,

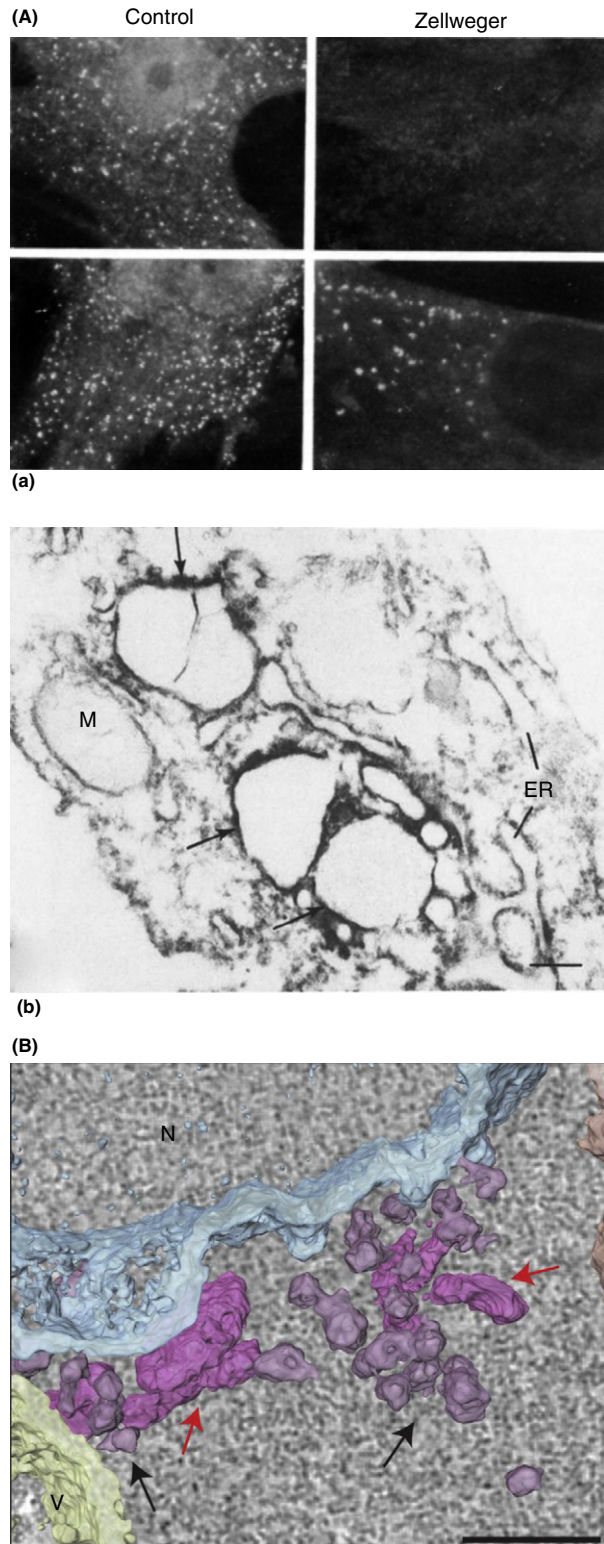


Figure 16 Peroxisomes in mutants. (A) Peroxisome membranes ('ghosts') in Zellweger patient cells. (a) Fluorescence microscopy: peroxisomes appear to be absent with an antibody against catalase, a luminal enzyme (top row); however, an antibody against peroxisome membrane proteins reveals the membranes, which are larger in size and fewer in number than normal (bottom row) (Santos *et al.*, 1988b). (b) Electron microscopy identifying peroxisome membrane ghosts with an electron dense cytochemical reaction product (arrows). M, mitochondrion (Santos *et al.*, 1988a). (B) Pre-peroxisome vesicles ('remnants') in *H. polymorpha* pex3/atg1 double mutant cells. 3D reconstruction of the vesicles (black arrows) and reticular structures (red arrows), which contain Pex8, Pex13, and Pex14 as well as a little Pex5 and alcohol oxidase enzyme. N, nucleus; V, vacuole (Knoops *et al.*, 2014).

2006). The consensus view is that these ‘pro-peroxisome’ vesicles then recruit the other peroxisome membrane proteins by the usual method from the cytosol, become translocation competent, and import the peroxisome enzymes, thus maturing into new peroxisomes. A minority view is that all the peroxisome membrane proteins pass through the ER and never fuse with older peroxisomes (van der Zand *et al.*, 2010, 2012).

De novo formation physiologically?

Whether and to what extent *de novo* peroxisome formation occurs in normal cells is the subject of ongoing debate and may be greater in mammals than in fungi or plants. Experiments with peroxisome-targeted photo-activatable green fluorescent protein (GFP) detected both peroxisome division and *de novo* peroxisome formation in cultured rat kidney epithelial cells, with about 70% *de novo* formation (Kim *et al.*, 2006). Subsequent work confirms the functioning of the two pathways, but shows that the exogenous expression of Pex16-mCerulean in the previous work may have exaggerated the ER role (Aranovich *et al.*, 2014). In contrast, peroxisomes are thought to form largely by asymmetric division in CHO cells (Huybrechts *et al.*, 2009). When peroxisome division is blocked in *H. polymorpha*, one large peroxisome is seen, implying that division is the major pathway in this species (Nagotu *et al.*, 2008). Pulse-chase and mating assays in wild type *S. cerevisiae* clearly showed that the ER supplies membrane components to existing peroxisomes, but no *de novo* formation occurred (Motley and Hettema, 2007). This conclusion was confirmed by fluorescence tagging experiments (Menendez-Benito *et al.*, 2013). Likewise, *de novo* formation has not been detected in any plant, but the ER does supply membrane proteins to existing organelles (Mullen and Trelease, 2006).

Doubts about de novo formation

The concept of *de novo* formation from the ER in yeast is challenged by the recent discovery using electron tomography that peroxisome-related structures are not, in fact, entirely

absent in PEX3 knockout strains of *H. polymorpha* and *S. cerevisiae* as had been thought; in fact small vesicles and tubules exist that contain Pex8, Pex13, and Pex14 as well as a little Pex8 and alcohol oxidase (Figure 16(B); Knoops *et al.*, 2014). Similar peroxisomal ‘remnants’ were detected previously in PEX3 and PEX19 knockouts of *P. pastoris* (Hazra *et al.*, 2002) but largely ignored. When Pex3 is reintroduced in *H. polymorpha*, it goes to these little structures, not to the ER, and peroxisomes are reformed gradually from them (Knoops *et al.*, 2014). The origin of these vesicles and tubules is not certain: if they are remnants of old peroxisomes from before the knockout, the peroxisomes have been rescued, as suggested previously (Lazarow, 2003), rather than made *de novo*. If they are not remnants, our understanding of *de novo* formation must change. These tiny structures are located adjacent to the ER and might not have been resolved by routine fluorescence microscopy in previous studies.

Remodeling, Random Turnover, and Selective Degradation

In most organisms, the steady growth in peroxisomes is counterbalanced by a steady, random autophagic degradation (Oku and Sakai, 2010; Saraya *et al.*, 2010). In rat liver, catalase and other peroxisome proteins turnover with exponential decay kinetics and a half-life of 1.5 days; these kinetics imply random destruction of the organelle (Poole, 1969). Forty years later, similar results (exponential decay, 2-day half-life) were obtained with entirely different methods on cultured CHO cells (Huybrechts *et al.*, 2009).

Peroxisomes are also selectively destroyed, when appropriate, by targeted autophagy, termed ‘pexophagy’ (Figure 17). For example, peroxisomes are induced in methylotrophic yeast by growth on methanol; withdrawal of this carbon source and addition of glucose causes these peroxisomes to be autophagocytosed and degraded. Similarly, withdrawal of a

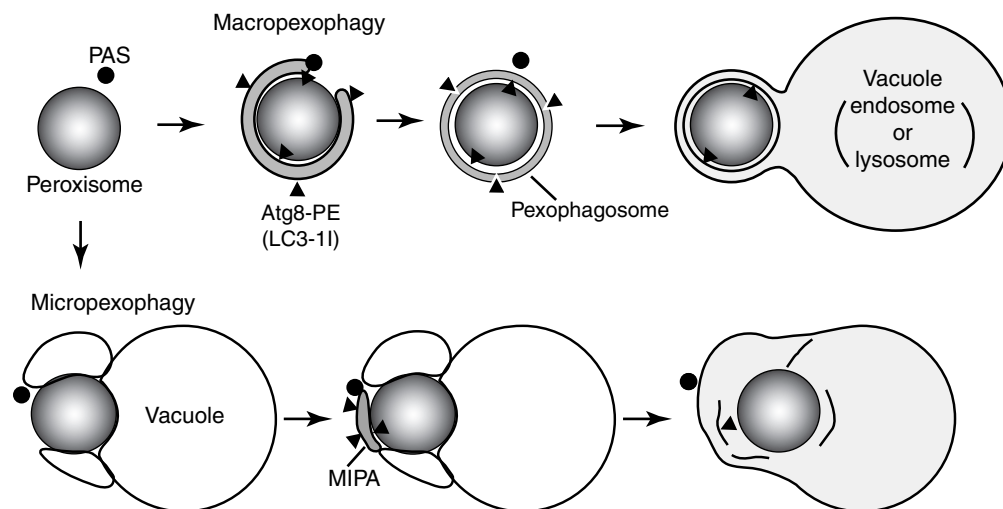


Figure 17 Autophagy of peroxisomes (‘pexophagy’). Top, macropexophagy, PAS=pre-autophagosomal structure; bottom, microautophagy, MIPA=micropexophagic membrane apparatus (Oku and Sakai, 2010).

hypolipidemic drug from mammals leads to pexophagy of the excess peroxisomes. In green leaves, peroxisomes damaged by H_2O_2 aggregate and are selectively degraded (Shibata *et al.*, 2014).

The functional remodeling of plant peroxisomes from fat metabolism to photorespiration upon greening is accomplished in part by a change in gene regulation (Mullen and Trelease, 2006): turn off one set of genes and turn on the other and in two half-lives three-fourth of the enzymes will be the new ones. In addition, plants help remodeling by selective autophagy (Kim *et al.*, 2014).

Evolution

Peroxisomes have been found in all five branches of the eukaryotic tree of life, although the vast majority of our knowledge concerns two branches, the Unikonts, containing animals and fungi, and the Plantae (Gabaldon, 2010). The mechanisms of peroxisome biogenesis and division are universally conserved, even among the peroxisomes with the most exceptional functions. Almost all peroxisomes harbor the enzymes of reactive oxygen metabolism and β -oxidation, strongly supporting a common evolutionary origin. Protection against H_2O_2 would have been useful to primitive aerobic organisms (de Duve and Baudhuin, 1966). Peroxisomes may be the descendants of an ancient endosymbiont that later lost all of its DNA while retaining a posttranslational type of protein import (topologically like mitochondria and chloroplasts but mechanistically different) (de Duve, 1969). However, the eukaryotic origin of the proteins that catalyze peroxisome biogenesis (Figure 18(A)) and the homology of some to the ER-associated decay (ERAD) pathway (Figure 18(B)) suggest an ER evolutionary origin for the membrane of the organelle (Schluter *et al.*, 2006; Gabaldon *et al.*, 2006). This would be consistent with new peroxisome membrane formation from the ER. The remarkable present diversity of peroxisomal enzymatic capabilities is likely due to the acquisition of enzymes and enzyme pathways retargeted from other cell compartments during evolution, which could have been accomplished by simply equipping them at their carboxy-terminus with the tripeptide that targets proteins to peroxisomes.

Human Disease and Peroxisomes

Defects in peroxisome biogenesis (Steinberg *et al.*, 2006) or function (Wanders and Waterham, 2006b) cause devastating human disease. Pathologists were aware of some of these diseases before much was known about peroxisomes. When (Goldfischer *et al.*, 1973) observed that peroxisomes were missing in the liver and kidney of two Zellweger syndrome neonates, the absence of the known enzymes could not explain the severe developmental neuropathology. This implied that the peroxisome had important other capabilities to be discovered and prompted a wave of productive research. The subsequent discovery that Zellweger cells contain peroxisome membrane ghosts (Santos *et al.*, 1988b) suggested that the

disease was due to a defect in biogenesis, specifically in the import of matrix enzymes, and stimulated further research.

Fourteen genetic complementation groups of Zellweger patients have been identified, each associated with a defective PEX gene, all autosomal recessive and all with defective peroxisome biogenesis (Steinberg *et al.*, 2006; Fujiki *et al.*, 2014). As a result of the impaired biogenesis, the major functions, including β -oxidation and plasmalogen synthesis, are defective. All organ systems are affected but the most striking effects are on the nervous system due to abnormal neural migration during prenatal development, with resulting hypotonia, lack of reflexes, seizures, and inability to feed. Milder variants of the syndrome, caused by less consequential mutations in the same set of genes, are called neonatal adrenoleukodystrophy and infantile Refsum disease. Defective individual enzymes of β -oxidation cause similar clinical results. Measurements of decreased enzyme activities and/or metabolite accumulations or deficits form the basis of clinical and prenatal diagnoses. Models of Zellweger syndrome in mice, worms, and fruit flies replicate many features of the human disease, including defects in the nervous system and testis, and are giving valuable insights into the roles of peroxisomes in development (Van Veldhoven and Baes, 2013).

Rhizomelic chondrodysplasia punctata is caused by a defect in one of the peroxisomal enzymes of plasmalogen synthesis or in Pex7, which is required for their import into peroxisomes. Ossification is impaired along with neurological and other abnormalities.

Adrenoleukodystrophy tragically appears in boys in late childhood with cerebral inflammatory demyelination leading to early death. It is caused by defects in the ABCD1 gene that encodes a peroxisomal membrane protein transporter of fatty acids; when defective or missing, oxidative stress occurs, apparently provoked by the accumulation of very long-chain fatty acids. A milder version of this X-linked disorder, adrenomyeloneuropathy, occurs in adults with damage to peripheral and spinal cord neurons. Both forms are seen within the same families, implying the involvement of additional genetic or environmental factors. Bone marrow transplantation was once the only treatment option for the childhood version, but this is now one of the first genetic diseases to be successfully treated by gene therapy. Autologous CD34+ cells received a good copy of the ABCD1 gene via a lentiviral vector and were infused back into patients, resulting in the gradual appearance of ABCD1-positive T and B lymphocytes, monocytes, and presumably microglia cells, implying transduction of hematopoietic stem cells. Clinically, disease progression stopped (Cartier *et al.*, 2009). A second type of therapy, a cocktail of antioxidants, is effective in a mouse model of adrenomyeloneuropathy, preventing axon degeneration and even permitting recovery of lost function. This indicates that oxidative stress is the underlying cause of the neuropathy, and has important implications for other human diseases including Alzheimer's disease and Parkinson's disease, in which oxidative stress likewise occurs (Galea *et al.*, 2012).

A second disease that develops in late childhood is Refsum disease, which is due to a failure in α -oxidation, and begins with deterioration in night vision and a progressive retinitis pigmentosa. It is treated by restricting dietary intake of phytanic acid.

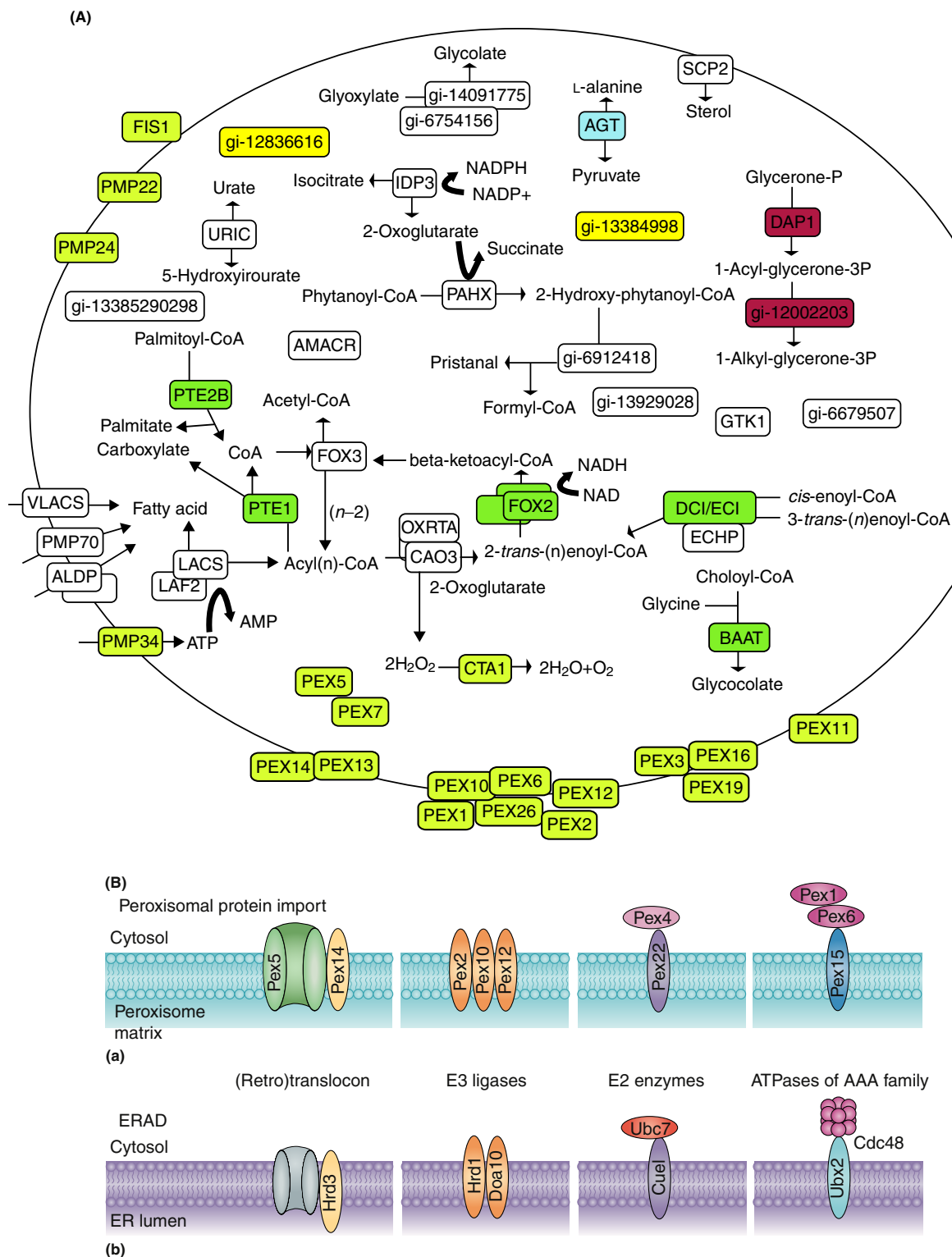


Figure 18 Evolution of the peroxisome. (A) Mammalian peroxisome proteome origins and metabolism, as inferred from proteomics data. The likely evolutionary origins of the proteins: yellow, eukaryotic; green, α proteobacterial; red, actinomycetales; blue, cyanobacterial; and white, undetermined (Gabaldon, 2010). (B) Molecular similarities between (a) the peroxisomal Pex5 import pathway and (b) the ERAD (ER-associated decay) pathway (Schliebs *et al.*, 2010).

Hyperoxaluria type 1 results from the failure of peroxisomal glyoxylate aminotransferase to remove glyoxylate. As a result, it is oxidized to oxalate, and insoluble calcium oxalate is deposited in the kidneys leading to kidney failure and death. This disease can result from mutations that inactivate the enzyme, but in many cases it is due to point mutations that cause a mistargeting of this peroxisomal enzyme to the mitochondria, where it fails to work properly.

Humans deficient in catalase are reported to be more susceptible to diabetes, cancer, and atherosclerosis and perhaps to alterations in aging as well (Fransen *et al.*, 2012).

Conclusion

Peroxisome biology turns out to be far more complex, diverse, and interesting than was imagined at its discovery. Almost 50 years later, we are still discovering new functions, new aspects of its biogenesis, and new pathologies caused by its defects. Some details remain controversial and need further elucidation. Future research is likely to produce more surprises about this fascinating organelle.

See also: Molecular Principles, Components, Technology, and Concepts: Lipids: Lipidomics; Synthesis and Structure of Glycerolipids

References

- Aranovich, A., Hua, R., Rutenberg, A.D., Kim, P.K., 2014. PEX16 contributes to peroxisome maintenance by constantly trafficking PEX3 via the ER. *Journal of Cell Science* 127, 3675–3686.
- Baes, M., Aubourg, P., 2009. Peroxisomes, myelination, and axonal integrity in the CNS. *Neuroscientist* 15, 367–379.
- Bartoszewski, M., Opalinski, L., Veenhuis, M., van der Klei, I.J., 2011. The significance of peroxisomes in secondary metabolite biosynthesis in filamentous fungi. *Biotechnology Letters* 33, 1921–1931.
- Binns, D., Januszewski, T., Chen, Y., *et al.*, 2006. An intimate collaboration between peroxisomes and lipid bodies. *Journal of Cell Biology* 173, 719–731.
- Boncompain, G., Muller, C., Meas-Yedid, V., *et al.*, 2014. The intracellular bacteria *Chlamydia* hijack peroxisomes and utilize their enzymatic capacity to produce bacteria-specific phospholipids. *PLoS One* 9, e86196.
- Boveris, A., Oshino, N., Chance, B., 1972. The cellular production of hydrogen peroxide. *Biochemical Journal* 128, 617–630.
- Braverman, N.E., Moser, A.B., 2012. Functions of plasmalogen lipids in health and disease. *Biochimica et Biophysica Acta* 1822, 1442–1452.
- Cartier, N., Hacein-Bey-Abina, S., Bartholomae, C.C., *et al.*, 2009. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science* 326, 818–823.
- Cohen, Y., Klug, Y.A., Dimitrov, L., *et al.*, 2014. Peroxisomes are juxtaposed to strategic sites on mitochondria. *Molecular BioSystems* 10, 1742–1748.
- de Duve, C., 1969. Evolution of the peroxisome. *Annals of the New York Academy of Sciences* 168, 369–381.
- de Duve, C., 1973. Biochemical studies on the occurrence, biogenesis and life history of mammalian peroxisomes. *Journal of Histochemistry & Cytochemistry* 21, 941–948.
- de Duve, C., 1983. Microbodies in the living cell. *Scientific American* 248, 74–84.
- de Duve, C., Baudhuin, P., 1966. Peroxisomes (microbodies and related particles). *Physiological Reviews* 46, 323–357.
- Del Rio, L.A., 2011. Peroxisomes as a cellular source of reactive nitrogen species signal molecules. *Archives of Biochemistry and Biophysics* 506, 1–11.
- Dimitrov, L., Lam, S.K., Schekman, R., 2013. The role of the endoplasmic reticulum in peroxisome biogenesis. *Cold Spring Harbor Perspectives in Biology* 5, a013243.
- Dixit, E., Boulant, S., Zhang, Y., *et al.*, 2010. Peroxisomes are signaling platforms for antiviral innate immunity. *Cell* 141, 668–681.
- Fagarasanu, A., Mast, F.D., Knoblach, B., Rachubinski, R.A., 2010. Molecular mechanisms of organelle inheritance: Lessons from peroxisomes in yeast. *Nature Reviews Molecular Cell Biology* 11, 644–654.
- Fransen, M., Nordgren, M., Wang, B., Apanasets, O., 2012. Role of peroxisomes in ROS/RNS-metabolism: Implications for human disease. *Biochimica et Biophysica Acta* 1822, 1363–1373.
- Freitag, J., Ast, J., Bolker, M., 2012. Cryptic peroxisomal targeting via alternative splicing and stop codon read-through in fungi. *Nature* 485, 522–525.
- Fujiki, Y., Okumoto, K., Mukai, S., Honsho, M., Tamura, S., 2014. Peroxisome biogenesis in mammalian cells. *Frontiers in Physiology* 5, 307.
- Gabalidon, T., 2010. Peroxisome diversity and evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 365, 765–773.
- Gabalidon, T., Snel, B., Van Zimmeren, F., *et al.*, 2006. Origin and evolution of the peroxisomal proteome. *Biol Direct* 1, 8.
- Galea, E., Launay, N., Portero-Otin, M., *et al.*, 2012. Oxidative stress underlying axonal degeneration in adrenoleukodystrophy: A paradigm for multifactorial neurodegenerative diseases? *Biochimica et Biophysica Acta* 1822, 1475–1488.
- Goldfischer, S., Moore, C.L., Johnson, A.B., *et al.*, 1973. Peroxisomal and mitochondrial defects in the cerebro-hepato-renal syndrome. *Science* 182, 62–64.
- Gorgas, K., 1984. Peroxisomes in sebaceous glands. V. Complex peroxisomes in the mouse preputial gland: Serial sectioning and three-dimensional reconstruction studies. *Anatomy and embryology* 169, 261–270.
- Gualdrón-López, M., Brennan, A., Hannaert, V., *et al.*, 2012. When, how and why glycolysis became compartmentalised in the Kinetoplastea. A new look at an ancient organelle. *International Journal for Parasitology* 42, 1–20.
- Hazra, P.P., Suriapranata, I., Snyder, W.B., Subramani, S., 2002. Peroxisome remnants in pex3delta cells and the requirement of Pex3p for interactions between the peroxisomal docking and translocation subcomplexes. *Traffic* 3, 560–574.
- Hettema, E.H., Erdmann, R., van der Klei, I., Veenhuis, M., 2014. Evolving models for peroxisome biogenesis. *Current Opinion in Cell Biology* 29, 25–30.
- Hoepfner, D., Schildknegt, D., Braakman, I., Philippsen, P., Tabak, H.F., 2005. Contribution of the endoplasmic reticulum to peroxisome formation. *Cell* 122, 85–95.
- Hu, J., Baker, A., Bartel, B., *et al.*, 2012. Plant peroxisomes: Biogenesis and function. *Plant Cell* 24, 2279–2303.
- Huybrechts, S.J., Van Veldhoven, P.P., Brees, C., *et al.*, 2009. Peroxisome dynamics in cultured mammalian cells. *Traffic* 10, 1722–1733.
- Islinger, M., Cardoso, M.J., Schrader, M., 2010. Be different – the diversity of peroxisomes in the animal kingdom. *Biochimica et Biophysica Acta* 1803, 881–897.
- Kaur, N., Reumann, S., Hu, J., 2009. Peroxisome biogenesis and function. *Arabidopsis Book* 7, e0123.
- Kim, J., Lee, H.N., Chung, T., 2014. Plant cell remodeling by autophagy: Switching peroxisomes for green life. *Autophagy* 10, 702–703.
- Kim, P.K., Mullen, R.T., Schumann, U., Lippincott-Schwartz, J., 2006. The origin and maintenance of mammalian peroxisomes involves a de novo PEX16-dependent pathway from the ER. *Journal of Cell Biology* 173, 521–532.
- Knoops, K., Manivannan, S., Cepinska, M.N., *et al.*, 2014. Preperoxisomal vesicles can form in the absence of Pex3. *Journal of Cell Biology* 204, 659–668.
- Kondrup, J., Lazarow, P.B., 1985. Flux of palmitate through the peroxisomal and mitochondrial beta-oxidation systems in isolated rat hepatocytes. *Biochimica et Biophysica Acta* 835, 147–153.
- Kunze, M., Pracharoenwattana, I., Smith, S.M., Hartig, A., 2006. A central role for the peroxisomal membrane in glyoxylate cycle function. *Biochimica et Biophysica Acta* 1763, 1441–1452.
- Kuravi, K., Nagotu, S., Krikken, A.M., *et al.*, 2006. Dynamin-related proteins Vps1p and Dnm1p control peroxisome abundance in *Saccharomyces cerevisiae*. *Journal of Cell Science* 119, 3994–4001.
- Lam, S.K., Yoda, N., Schekman, R., 2010. A vesicle carrier that mediates peroxisome protein traffic from the endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America* 107, 21523–21528.
- Lazarow, P.B., 1977. Three hypolipidemic drugs increase hepatic palmitoyl-coenzyme A oxidation in the rat. *Science* 197, 580–581.
- Lazarow, P.B., 2003. Peroxisome biogenesis: Advances and conundrums. *Current Opinion in Cell Biology* 15, 489–497.
- Lazarow, P.B., 2011. Viruses exploiting peroxisomes. *Current Opinion in Microbiology* 14, 458–469.
- Lazarow, P.B., de Duve, C., 1973. The synthesis and turnover of rat liver peroxisomes. V. Intracellular pathway of catalase synthesis. *Journal of Cell Biology* 59, 507–524.

- Lazarow, P.B., de Duve, C., 1976. A fatty acyl-CoA oxidizing system in rat liver peroxisomes; enhancement by clofibrate, a hypolipidemic drug. *Proceedings of the National Academy of Sciences of the United States of America* 73, 2043–2046.
- Lazarow, P.B., Fujiki, Y., 1985. Biogenesis of peroxisomes. *Annual Review of Cell Biology* 1, 489–530.
- Lazarow, P.B., Shio, H., Robbi, M., 1980. Biogenesis of peroxisomes and the peroxisome reticulum hypothesis. In: Bucher, T., Sebald, W., Weiss, H. (Eds.), 31st Mosbach Colloquium. *Biological Chemistry of Organelle Formation*. New York, NY: Springer-Verlag, pp. 187–206.
- Lingard, M.J., Trelease, R.N., 2006. Five Arabidopsis peroxin 11 homologs individually promote peroxisome elongation, duplication or aggregation. *Journal of Cell Science* 119, 1961–1972.
- Lipka, V., Dittgen, J., Bednarek, P., *et al.*, 2005. Pre- and postinvasion defenses both contribute to nonhost resistance in Arabidopsis. *Science* 310, 1180–1183.
- Liu, F., Ng, S.K., Lu, Y., *et al.*, 2008. Making two organelles from one: Woronin body biogenesis by peroxisomal protein sorting. *Journal of Cell Biology* 180, 325–339.
- Ma, C., Agrawal, G., Subramani, S., 2011. Peroxisome assembly: Matrix and membrane protein biogenesis. *Journal of Cell Biology* 193, 7–16.
- Martin, J.F., Ullan, R.V., Garcia-Estrada, C., 2012. Role of peroxisomes in the biosynthesis and secretion of beta-lactams and other secondary metabolites. *Journal of Industrial Microbiology and Biotechnology* 39, 367–382.
- Mariyama, J., Yamaoka, S., Matsuo, I., Tsutsumi, N., Kitamoto, K., 2012. A newly discovered function of peroxisomes: Involvement in biotin biosynthesis. *Plant Signaling & Behavior* 7, 1589–1593.
- Matsuzaki, T., Fujiki, Y., 2008. The peroxisomal membrane protein import receptor Pex3p is directly transported to peroxisomes by a novel Pex19p- and Pex16p-dependent pathway. *Journal of Cell Biology* 183, 1275–1286.
- Meinecke, M., Cizmowski, C., Schliebs, W., *et al.*, 2010. The peroxisomal importomer constitutes a large and highly dynamic pore. *Nature Cell Biology* 12, 273–277.
- Menendez-Benito, V., Van Deventer, S.J., Jimenez-Garcia, V., *et al.*, 2013. Spatiotemporal analysis of organelle and macromolecular complex inheritance. *Proceedings of the National Academy of Sciences of the United States of America* 110, 175–180.
- Mhamdi, A., Noctor, G., Baker, A., 2012. Plant catalases: Peroxisomal redox guardians. *Archives of Biochemistry and Biophysics* 525, 181–194.
- Motley, A.M., Hettema, E.H., 2007. Yeast peroxisomes multiply by growth and division. *Journal of Cell Biology* 178, 399–410.
- Mullen, R.T., Trelease, R.N., 2006. The ER-peroxisome connection in plants: Development of the "ER semi-autonomous peroxisome maturation and replication" model for plant peroxisome biogenesis. *Biochimica et Biophysica Acta* 1763, 1655–1668.
- Nagotu, S., Saraya, R., Otzen, M., Veenhuis, M., van der Klei, I.J., 2008. Peroxisome proliferation in *Hansenula polymorpha* requires Dnm1p which mediates fission but not de novo formation. *Biochimica et Biophysica Acta* 1783, 760–769.
- Nair, D.M., Purdue, P.E., Lazarow, P.B., 2004. Pex7p translocates in and out of peroxisomes in *Saccharomyces cerevisiae*. *Journal of Cell Biology* 167, 599–604.
- Novikoff, A.B., Shin, W.Y., 1964. The endoplasmic reticulum in the Golgi zone and its relations to microbodies, Golgi apparatus and autophagic vacuoles in rat liver cells. *Journal of Microscopy* 3, 187–206.
- Nuttall, J.M., Motley, A., Hettema, E.H., 2011. Peroxisome biogenesis: Recent advances. *Current Opinion in Cell Biology* 23, 421–426.
- Oku, M., Sakai, Y., 2010. Peroxisomes as dynamic organelles: Autophagic degradation. *FEBS Journal* 277, 3289–3294.
- Opperdoes, F.R., Borst, P., 1977. Localization of nine glycolytic enzymes in a microbody-like organelle in *Trypanosoma brucei*: The glycosome. *FEBS Letters* 80, 360–364.
- Palma, J.M., Corpas, F.J., del Rio, L.A., 2009. Proteome of plant peroxisomes: New perspectives on the role of these organelles in cell biology. *Proteomics* 9, 2301–2312.
- Peraza-Reyes, L., Berteaux-Lecellier, V., 2013. Peroxisomes and sexual development in fungi. *Frontiers in Physiology* 4, 244.
- Poole, B., 1969. Biogenesis and turnover of rat liver peroxisomes. *Annals of the New York Academy of Sciences* 168, 229–243.
- Poole, B., 1975. Diffusion effects in the metabolism of hydrogen peroxide by rat liver peroxisomes. *Journal of Theoretical Biology* 51, 149–167.
- Raychaudhuri, S., Prinz, W.A., 2008. Nonvesicular phospholipid transfer between peroxisomes and the endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America* 105, 15785–15790.
- Rhee, S.G., 2006. Cell signaling. H₂O₂, a necessary evil for cell signaling. *Science* 312, 1882–1883.
- Rucktaschel, R., Girzalsky, W., Erdmann, R., 2011. Protein import machineries of peroxisomes. *Biochimica et Biophysica Acta* 1808, 892–900.
- Santos, M.J., Imanaka, T., Shio, H., Lazarow, P.B., 1988a. Peroxisomal integral membrane proteins in control and Zellweger fibroblasts. *Journal of Biological Chemistry* 263, 10502–10509.
- Santos, M.J., Imanaka, T., Shio, H., Small, G.M., Lazarow, P.B., 1988b. Peroxisomal membrane ghosts in Zellweger syndrome—aberrant organelle assembly. *Science* 239, 1536–1538.
- Saraya, R., Veenhuis, M., van der Klei, I.J., 2010. Peroxisomes as dynamic organelles: Peroxisome abundance in yeast. *FEBS Journal* 277, 3279–3288.
- Schliebs, W., Girzalsky, W., Erdmann, R., 2010. Peroxisomal protein import and ERAD: Variations on a common theme. *Nature Reviews. Molecular Cell Biology* 11, 885–890.
- Schluter, A., Fourcade, S., Ripp, R., *et al.*, 2006. The evolutionary origin of peroxisomes: An ER-peroxisome connection. *Molecular Biology and Evolution* 23, 838–845.
- Schrader, M., Bonekamp, N.A., Islinger, M., 2012. Fission and proliferation of peroxisomes. *Biochimica et Biophysica Acta* 1822, 1343–1357.
- Schrader, M., King, S.J., Stroth, T.A., Schroer, T.A., 2000. Real time imaging reveals a peroxisomal reticulum in living cells. *Journal of Cell Science* 113 (Pt 20), 3663–3671.
- Shen, Y.Q., Burger, G., 2009. Plasticity of a key metabolic pathway in fungi. *Functional & Integrative Genomics* 9, 145–151.
- Shibata, M., Oikawa, K., Yoshimoto, K., *et al.*, 2014. Plant autophagy is responsible for peroxisomal transition and plays an important role in the maintenance of peroxisomal quality. *Autophagy* 10, 936–937.
- Shio, H., Lazarow, P.B., 1981. Relationship between peroxisomes and endoplasmic reticulum investigated by combined catalase and glucose-6-phosphatase cytochemistry. *Journal of Histochemistry & Cytochemistry* 29, 1263–1272.
- Smith, J.J., Aitchison, J.D., 2013. Peroxisomes take shape. *Nature Reviews Molecular Cell Biology* 14, 803–817.
- Steinberg, S.J., Dodd, G., Raymond, G.V., *et al.*, 2006. Peroxisome biogenesis disorders. *Biochimica et Biophysica Acta* 1763, 1733–1748.
- Thoms, S., Gronborg, S., Gartner, J., 2009. Organelle interplay in peroxisomal disorders. *Trends in Molecular Medicine* 15, 293–302.
- Van Veldhoven, P.P., Baes, M., 2013. Peroxisome deficient invertebrate and vertebrate animal models. *Frontiers in Physiology* 4, 335.
- Veenhuis, M., van der Klei, I.J., 2014. A critical reflection on the principles of peroxisome formation in yeast. *Frontiers in Physiology* 5, 110.
- Wanders, R.J., Waterham, H.R., 2006a. Biochemistry of mammalian peroxisomes revisited. *Annual Review of Biochemistry* 75, 295–332.
- Wanders, R.J., Waterham, H.R., 2006b. Peroxisomal disorders: The single peroxisomal enzyme deficiencies. *Biochimica et Biophysica Acta* 1763, 1707–1720.
- Yamamoto, K., Fahimi, H.D., 1987. Three-dimensional reconstruction of a peroxisomal reticulum in regenerating rat liver: Evidence of interconnections between heterogeneous segments. *Journal of Cell Biology* 105, 713–722.
- van der Zand, A., Braakman, I., Tabak, H.F., 2010. Peroxisomal membrane proteins insert into the endoplasmic reticulum. *Molecular Biology of the Cell* 21, 2057–2065.
- van der Zand, A., Gent, J., Braakman, I., Tabak, H.F., 2012. Biochemically distinct vesicles from the endoplasmic reticulum fuse to form peroxisomes. *Cell* 149, 397–409.

Lipid Droplets

C Thiele and A Penno, University of Bonn, Bonn, Germany

© 2016 Elsevier Inc. All rights reserved.

Glossary

Glycerolipid The most abundant class of lipids. They are built by esterification of fatty acyl-CoA to the hydroxyl groups of a glycerol molecule. They may also contain ether-linked fatty alcohols.

Lipid Naturally occurring hydrophobic molecules structurally derived from fatty acids or isoprenoids, biosynthetically derived from acetyl-CoA. The major mammalian lipid classes are glycerolipids, sphingolipids, and sterols.

Lipogenesis Cellular de novo synthesis of lipids. This term may be used to describe the de novo synthesis of fatty acyl-CoAs and their further metabolism to glycerolipids, sphingolipids, and sterols.

Lipolysis Term used to describe the degradation of triacylglycerol to glycerol and three fatty acids.

Monotopic membrane protein Proteins that integrate into membranes with a 'hairpin'-like structure. All hydrophilic parts of the proteins face the same side of the membrane.

Neutral lipid Highly hydrophobic lipids without polar/charged groups (phosphate, amine, sugar, or choline). The most common mammalian neutral lipids are triacylglycerol and sterol ester.

Phospholipid Lipids with a phosphate group. Due to the phosphate, they possess a polar, hydrophilic part in addition to the hydrophobic part. Phospholipids are the major components of biological membranes.

Polytopic membrane protein Proteins of which some hydrophilic parts are exposed to one side of the membrane while others face the other side.

Sphingolipid Lipids that differ in structure and biogenesis to the glycerol-based lipids. Their backbone is a long-chain amino alcohol instead of glycerol.

Sterol Hydrophobic polycyclic alcohols that are major components of membranes and precursors of many signaling substances. The most common mammalian sterol is cholesterol.

Basic Structure

Lipid droplets (LDs) are spherical organelles, which appear in the cytosol of some pro- and most eukaryotic cells upon energy surplus. They consist of a core of neutral lipids (NLs) surrounded by a phospholipid monolayer with embedded or loosely associated proteins (Tauchi-Sato *et al.*, 2002). This rather simple structure distinguishes LDs from all other intracellular organelles, which are all separated from the cytosol by membrane bilayers, and hence has some important biological consequences that will be discussed later in this article.

In prokaryotes, the neutral lipid core of LDs frequently contains poly-hydroxyalkanoates like poly-hydroxybutyrate, while storage of triacylglycerol (TAG) or wax esters is relatively rare (Waltermann and Steinbuchel, 2005). In yeasts, plants and animals, TAG is the predominant storage lipid, but other lipids such as sterol esters (SE), retinol esters, free sterols and ether lipids are also stored, often in specialized cell types. Some LDs, particularly in leukocytes, possess internal structures (Melo *et al.*, 2013) with membranes and proteins within the hydrophobic core. For mammalian LDs, the surrounding phospholipid monolayer is mainly composed of phosphatidylcholine (PC, up to 60%), phosphatidylethanolamine (PE, up to 24%), and phosphatidylinositol (PI, up to 8%). Other minor components of the phospholipid monolayer are lyso- and ether-linked forms of PC and PE (Bartz *et al.*, 2007).

The size of LDs ranges from approximately 100 nm in diameter in white adipose cells to the smallest visible LDs of

approximately 100 nm in mammalian cells not preferentially suited for lipid storage. They are usually perfectly round and thereby store a maximum of hydrophobic metabolites within a minimum of membrane layer: Considering the approximate surface area of PC in monolayers ($\sim 1 \text{ nm}^2$ per molecule) and the density of a NL core composed of triolein ($\sim 0.95 \text{ g l}^{-1}$), only 2% PC for small LDs (400 nm in diameter) and 0.5% PC for medium-sized LDs (2 μm in diameter) is needed to shield the hydrophobic core from the aqueous cytosol (Penno *et al.*, 2013).

LD Proteins

Proteome analyses of LDs of different species and cell types (for review see Yang *et al.*, 2012) have shown that LDs contain specific and relatively concise proteomes of about 30–80 different proteins. The unequivocal assignment of LD proteins is complicated due to methodological problems: LDs are isolated using density centrifugations in which the LDs float on top of sucrose gradients while other membranous organelles sink to the bottom fraction. To exclude contaminations of the LD fraction with other membranes, the centrifugation can be repeated several times. Nonetheless, purity is a major concern since large round LDs have relatively little surface compared to the tubular structures of the major contaminant, that is, the endoplasmic reticulum (ER). Hence the purity of an LD preparation may be evaluated by the lack of common ER membrane components such as calnexin. An additional complication, also for the identification of LD proteins via

microscopy, is the fact that LDs are often found partly enwrapped or in very close association with the ER. In some cases it may therefore be hard to discriminate between 'on the LD monolayer' or 'around the LD.' In light microscopy pictures of unfixed cells LD proteins are supposed to show rings rather than egg-cup like structures around the LDs.

Structurally, LD proteins may be divided into four groups:

1. Integral monotopic membrane proteins,
2. Proteins associated with the monolayer by a lipid-anchor,
3. Proteins inserted into the monolayer by an amphipathic helix, and
4. Loosely associated proteins, attached by protein–protein or lipid–protein interactions.

Polytopic integral membrane proteins are sometimes detected in LD proteomes but their presence likely indicates contaminations with closely associated membranes (mostly ER but also mitochondria, Golgi, plasma membrane, and vesicular proteins). This is due to the unique structural organization of LDs in which the hydrophobic core does not provide a luminal aqueous compartment to accommodate polytopic proteins.

Integral monotopic LD proteins (class a) have helix-strand breaking amino acids within a long hydrophobic stretch that provoke a 'hairpin'-like monotopic structure. The 'hairpin' allows for insertion into both, membrane bilayers or a monolayer with a hydrophobic core. This structure, however, is necessary but not sufficient for LD targeting. Additional positively charged amino acids surrounding the hydrophobic stretch seem to be required (Stevanovic and Thiele, 2013). These could directly interact with negative charges of the phospholipid monolayer or could be a sorting signal that selects proteins for transport into the LD monolayer. While integral ER proteins will be able to diffuse into the phospholipid monolayer of nascent LDs, it is less clear how these proteins would shuttle between bilayered membranes and mature cytoplasmic LDs. Constant or transient membrane continuities (hemifusion of the LD monolayer with the cytoplasmic leaflet of other organelles, see Wilfling *et al.*, 2013; Ohsaki *et al.*, 2008) as well as protein-supported transport mechanisms at organelle contact sites are plausible explanations. The protein classes (2)–(4) are at least transiently soluble in the cytoplasm, hence their localization may change more easily between cytosolic and LD-associated (see Figure 1).

The members of the perilipin protein family (see Braasmele, 2007; Bickel *et al.*, 2009) are widespread structural components of LDs and serve as the universal LD protein markers in higher eukaryotes. The five members (Plin1–5) associate with the LD monolayer by amphipathic helices. Some family members are ubiquitously expressed (Plin2–3) while others show tissue-dependent expression profiles (Plin1 and Plin4–5). Plin1–2 are exclusively found on LDs and protect the LD against lipolysis under basal conditions. They are rapidly degraded in the absence of LDs. In contrast to Plin1–2, Plin 3–5 are stable in the absence of LDs and relocate to the surface of LDs depending on the metabolic state of the cell. Plins have lipoprotein-like properties *in vitro*; they may shield lipids from an aqueous environment and hence potentially enable cells to form LD-like structures *de novo* in the cytosol (Bulankina *et al.*, 2009; Bartholomew *et al.*, 2012).

Adipocyte LDs

LDs in adipocytes are especially interesting with regard to widespread health threads such as overweight, obesity, and the metabolic syndrome. Many mammalian species, in particular hibernating animals and also obese humans can accumulate adipose tissue in amounts exceeding any other body tissue. Despite of their obvious importance, very little is actually known about bona fide adipocyte LDs. Currently, most LD research is performed in fibroblasts, hepatocytes, yeast and in differentiated 3T3-L1 cells, a mouse cell line that differentiates into adipocytes-like cells upon treatment with insulin and dexamethasone. 3T3-L1 cells are a major source for adipocytes used in cell biological experiments, and there is no doubt that many aspects of adipocyte biology are well represented by these cells. But it should also be noted that, compared to adipose tissue, 3T3-L1 cells have relatively small, multilocular LDs with a large surface to volume ratio. The reason for the prevalence of 3T3-L1 cells is the problem in handling adipose tissue and freshly isolated adipocytes. Although easily available from liposuction, adipocytes become mechanically extremely sensitive and disintegrate within minutes when separated from the connective tissue, releasing large amounts of fat that float on the surface of the culture system. One interesting aspect of white adipocytes is the mechanism by which they accomplish the transition of multilocular micrometer sized LDs to a single LD that fills the entire cytoplasm leaving only little space for nucleus and other cellular organelles. For a long time, research in this area attempted to find rapid fusion events similar to the SNARE-mediated sub-second time scale fusions of vesicular structures. While results of these attempts have not been convincing (Bostrom *et al.*, 2007; see also Murphy *et al.*, 2010; Ariotti *et al.*, 2012), more recent research has led to the discovery of a slow process of LD coalescence (Nishino *et al.*, 2008). Here, droplets form contact sites followed by slow transfer of material from the smaller to the larger droplet, probably driven by differences in surface tension between differently sized LDs (see Figure 1). At the end of the coalescence event, the smaller droplet has squeezed its entire content into the larger; the fate of its excess surface remains unclear. A specific protein, FSP27, localizes to these contact zones and is essential for the process. Since FSP27 is specifically present in adipocytes, droplet coalescence appears to be the mechanism leading to the formation of large unilocular LDs in adipocytes.

LD Subpopulations

The LD population within a given cell is not homogenous. Simple light microscopy using a lipophilic dye already reveals major differences in size and motility. They may also vary in their metabolic state – while one LD is shrinking due to lipolysis, the neighboring LD may grow because of sustained NL synthesis. It was furthermore suggested that LDs of the same cell differ in their lipid composition with certain droplets storing preferentially TAGs while others store preferentially SEs (Hsieh *et al.*, 2012). However, so far this is not proven and contradicts another report suggesting that SEs may form shell-like structures around TAGs in LDs (Czabany *et al.*, 2008). Without doubt, LDs differ in their protein composition even

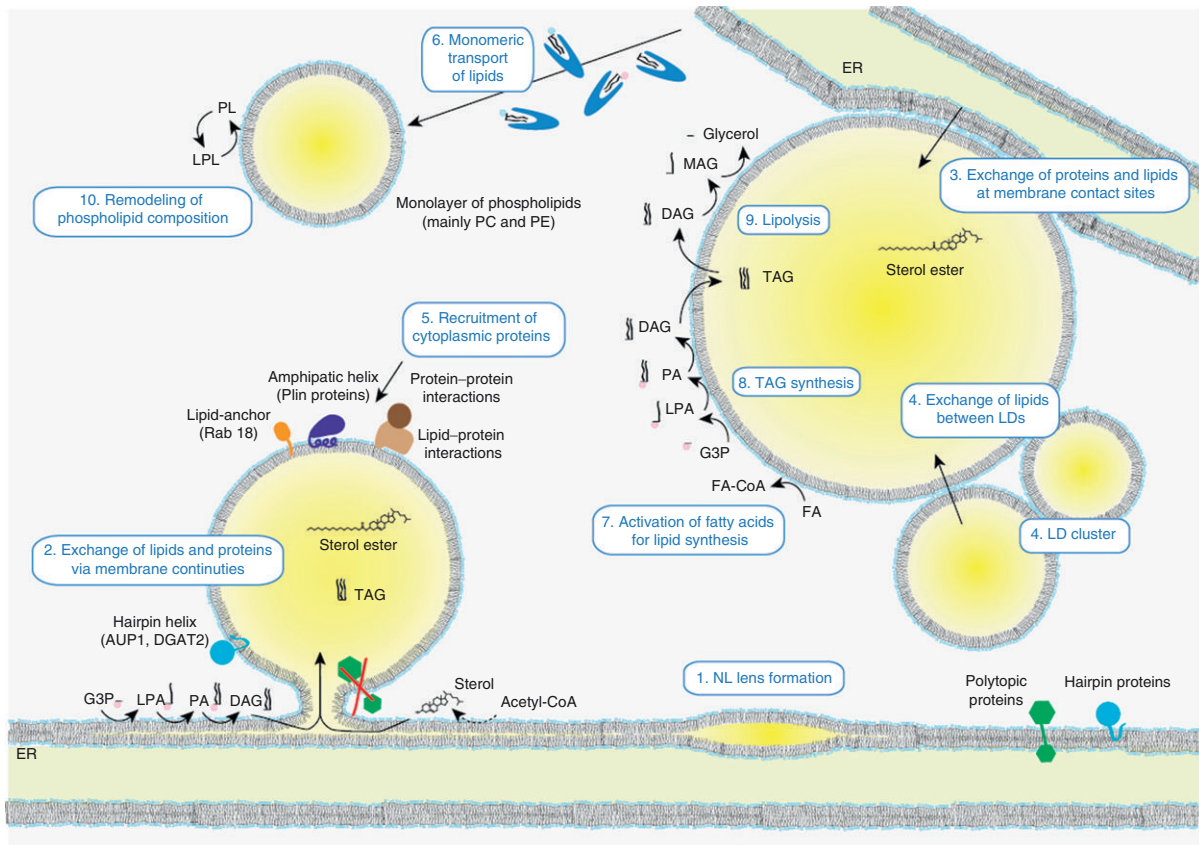


Figure 1 Schematic representation of cellular lipid droplets. Lipids droplets originate in the ER due to increased local NL synthesis leading to the formation of a NL aggregation/lens between the two leaflets of the ER (1). Upon further NL synthesis, the NL aggregation is thought to round to a nascent LD that would initially still be attached to the ER via a monolayer continuity, (2) before it would bud off into the cytoplasm. LDs form contact sides with other cellular organelles especially the ER and mitochondria (3) and other LDs (4) where proteins and lipids can be exchanged. LD proteins can diffuse into the LD monolayer via membrane continuities or be transported into the monolayer at organelle contact sites. Transiently soluble proteins are recruited from the cytoplasm and associated with the LD monolayer by lipid-anchors, amphipathic helices or lipid/proteins interactions (5). Apart from the exchange of lipids via membrane continuities or organelle contact sites, different lipids species are likely transported between the LD and other organelles via lipid binding proteins (monomeric transport) (6). LDs possess the enzymatic activities for activation of fatty acids (fatty acyl-CoA ligases (7)), for triacylglycerol synthesis (8), for lipolysis (deesterification of triacylglycerol or sterol ester (9)) and for the remodeling of their phospholipid composition (PLD, PLAs, LPCAT, (10)).

within the same cell. Plin1, for example, associates with large, Plin3 with small and Plin2 with medium sized LDs (Wolins *et al.*, 2006). The presence of one LD protein may even exclude another as it was observed for AUP1 and Rab18. The biological significance for these different subpopulations is an area of active research.

Biogenesis of LDs

The biogenesis of LDs is enigmatic but of high interest. In the prevalent model, LDs originate in the ER upon increased local NL synthesis. As NLs are highly hydrophobic and only soluble up to a few molar percent in phospholipid membranes, the increased local synthesis of NLs leads to an 'oiling out' of NLs between the two leaflets of the ER membrane. By a so far not further identified mechanism, the NL accumulation does not evenly spread within the bilayer but forms a lens-shaped inclusion (see Figure 1) that upon further growth rounds to a

nascent LD that buds off into the cytoplasm. Hence, the nascent LD is surrounded by a phospholipid monolayer originating from the cytosolic leaflet of the ER membrane. This model is based on the fact that the ER is the major site of cellular lipid synthesis and that early nascent LDs are found in close contact to the ER. Moreover, this model nicely explains the experimental findings that the LD monolayer resembles the cytosolic leaflet of the ER (mainly PC and PE, very little sphingolipids) and that integral ER proteins (AUP1, ACSL3, and ALDI) are already found on the first detectable LDs. Recent studies have also described proteins (seipin and FIT1/2) which locate to the assumed LD biogenesis sites of the ER and would regulate the 'oiling out' and budding of the nascent LDs (Szymanski *et al.*, 2007; Gross *et al.*, 2011). However, it was so far not possible to prove this biogenesis model beyond any doubt. This is due to methodological limitations as nascent LDs are thought to be approximately 10 nm in diameter (Zanghellini *et al.*, 2010), a size below the resolution of light microscopy and challenging for currently available electron

microscopy techniques. In addition and in spite of vast screening efforts, so far no single protein has been identified that would be obligatory for LD biogenesis. Only the combined loss of all NL synthesizing enzymes in yeast leads to a loss of LD generation upon fatty acid feeding (Sandager *et al.*, 2002; Oelkers *et al.*, 2002). This is in line with the budding hypothesis that is primarily based on biophysical principles of phase behavior rather than an active protein machinery.

Until new methods arise, which would enable the analysis of 10 nm-sized lipid aggregations, the ER-based budding of LDs explains most if not all experimental findings. Nonetheless other theoretical models may be discussed. LDs could in theory emerge by ‘cutting out’ or vesicular budding from the ER. In addition, they could be formed along the cytoplasmic side of the ER in an egg-cup manner. It is also possible that LDs or some LDs may originate from liposomes *de novo* in the cytoplasm or by fission events of preexisting LDs. It is tempting to speculate that redundant pathways for LD biogenesis exist, resulting in LDs of different origin, protein/lipid composition, and potential internal structure.

Neutral Lipid Synthesis and Lipolysis

The primary function of LDs is the storage of hydrophobic lipids as energy source or biosynthetic precursors. With the exception of free cholesterol in adipocyte LDs, all stored lipids are esters that are synthesized from their components for storage. Stored lipids need to be mobilized upon request, and consequently both NL synthesis and the reverse process of ester cleavage, called lipolysis, are key functions physically connected to LDs.

The classical pathway of TAG synthesis proceeds from glycerol-3-phosphate (G3P) via lysophosphatidic acid (LPA), phosphatidic acid (PA), and diacylglycerol (DAG) to TAG. All of the corresponding enzymatic activities (GPAT, LPAAT, PAP, and DGAT, respectively), as well as the fatty acyl-CoA synthetases, have been attributed to LDs by proteomics or by functional assays (see Figure 1). These assignments include a large range of different species and cell types but no single LD preparation has been shown to produce TAG from G3P and fatty acyl-CoA. Nonetheless, at least the ability to activate fatty acids to the respective CoAs and to synthesize TAG from DAG and fatty acyl-CoA appears to be common to LDs of different species and origin (see Figure 1). In contrast to the synthesis of TAG, LDs are unable to synthesize SEs but need to import them from the ER. LDs can modify their surface PC by exchange of the sn-2 fatty acid, using PLA₂ and LPCAT activities (see Figure 1). Lack of these activities leads to phenotypic alterations in the droplet pool; the physiological significance is unclear. LDs contain numerous sterol synthesizing or modifying enzymes. The presence of these activities on LDs is enigmatic since LDs are neither essential for sterol synthesis nor have been shown to contribute to it. It is tempting to speculate that droplets contain these enzymes in order to separate them physically from their ER-localized substrates to reduce sterol synthesis at high NL load. As a side effect of NL storage, LDs also accumulate other highly hydrophobic substances including xenobiotics and toxic compounds such as polychlorinated biphenyl (PCB) and tetrachlorodibenzodioxin

(TCDD). Sequestration into droplets reduces plasma concentrations of these substances, but adipose tissue itself may become a target for their toxic effects.

All LDs are able to cleave ester lipids to release their constituents (see Figure 1). This process of lipolysis depends on the interplay of its key components ATGL, CGI-58, HSL and, in adipocytes on Plin1. Plin1 localizes to very large droplets and shields the LD from lipolysis under basal conditions. Only upon phosphorylation via PKA, Plin1 releases CGI-58, an activator of the TAG lipase ATGL and recruits the SE and DAG lipase HSL to LDs. There are numerous other accessory proteins and lipases that further stimulate lipolysis and basal LD turnover, which are beyond the scope of this article (please see ‘Further Reading’).

Droplet Dynamics

With the exception of unilocular droplets in adipocytes, LDs are in principle motile organelles that move along microtubules (Spandl *et al.*, 2009). Motility is driven by the motor protein kinesin-1 with typical peak velocities around 1 $\mu\text{m s}^{-1}$, which is much slower than the rapid movement of endocytic vesicles. Such movement was thoroughly studied in syncytial *Drosophila* embryos, where it is part of a generalized cytoplasmic fluid dynamic (Welte, 2009). In other cell types, LD movement contributes to the frequently observed accumulation of LDs in perinuclear regions. These LD accumulations frequently form dense clusters of 10–100 LDs with cell type dependent large variations in cluster frequency and size (see Figure 1). Several LD proteins have been associated with cluster formation, but the physiological function and relevance of LD clustering is unclear (Lohmann *et al.*, 2013).

The size of the cellular LD pools varies not only with nutrient availability, but also during the cell cycle. Detailed studies in yeast have shown that LD breakdown is coordinated with daughter cell formation, mediated by phosphorylation of lipases by cyclin-dependent kinases (Kurat *et al.*, 2009). Much less is known about droplet dynamics in the mammalian cell cycle. During cell division, LDs are evenly distributed between the daughter cells in adipocytes (Nagayama *et al.*, 2007). LDs accumulate upon cell cycle arrest and in apoptotic or necrotic cells, which sometimes makes it difficult to distinguish phenotypes of lipid accumulation from cellular toxicity in screening experiments. Indeed, LD biogenesis is found to be a generalized response upon ER stress (Hapala *et al.*, 2011).

Lipid Droplets and Disease

Many disease conditions are linked to alterations in metabolism with subsequent accumulation or reduction of LDs. A detailed discussion of diseases like obesity and diabetes is beyond the scope of this article (see ‘Further Reading’), but some observations should be mentioned here. Some viruses have connected their life cycle to lipid synthesis and secretion and make use of LDs for their replication. The best-studied examples are the closely related Hepatitis C and Dengue viruses. HCV makes use of various different features of hepatic lipid metabolism (Alvisi *et al.*, 2011). It interferes with lipid

metabolism to increase LD accumulation and induces a structure rich in ER membranes and LDs, called the membranous web. There, several HCV proteins bind to LDs, promoting virus assembly and also virus release.

LDs are also closely related to cancer (Santos and Schulze, 2012). Many cancer types, with the renal clear cell carcinoma as a prominent example, accumulate large amounts of LDs (Straub *et al.*, 2010). This is apparently due to stimulation of growth factor and insulin signaling in cancer cells and supports the cells by supplying energy and protection against oxidative stress. On the other hand, loss of LDs is a hallmark of cancer cachexia, mediated by chronic activation of inflammatory signaling pathways that give rise to an imbalance between TAG synthesis and degradation (Das *et al.*, 2011).

See also: Cytoskeleton and Motors: Cytoskeletal Components: Kinesin Superfamily Proteins (KIFs) as a Fundamental Component of Life: Intracellular Transport and Beyond. Interorganellar Communication: Interplay and Processes: Endoplasmic Reticulum Stress in Disease. Molecular Principles, Components, Technology, and Concepts: Lipids: Cholesterol and Other Steroids; Synthesis and Structure of Glycerolipids. Molecular Principles, Components, Technology, and Concepts: Membranes: Composition, Physical Properties, and Curvature. Organelles: Structure and Function: The Endoplasmic Reticulum. Protein Synthesis/Degradation: Protein Degradation – Intracellular: Endoplasmic Reticulum-Associated Degradation and Protein Quality Control

References

- Alvisi, G., Madan, V., Bartenschlager, R., 2011. Hepatitis C virus and host cell lipids: An intimate connection. *RNA Biology* 8, 258–269.
- Ariotti, N., Murphy, S., Hamilton, N.A., *et al.*, 2012. Postlipolytic insulin-dependent remodeling of micro lipid droplets in adipocytes. *Molecular Biology of the Cell* 23, 1826–1837.
- Bartholomew, S.R., Bell, E.H., Summerfield, T., *et al.*, 2012. Distinct cellular pools of perilipin 5 point to roles in lipid trafficking. *Biochimica et Biophysica Acta* 1821, 268–278.
- Bartz, R., Li, W.H., Venables, B., *et al.*, 2007. Lipidomics reveals that adiposomes store ether lipids and mediate phospholipid traffic. *Journal of Lipid Research* 48, 837–847.
- Bickel, P.E., Tansey, J.T., Welte, M.A., 2009. PAT proteins, an ancient family of lipid droplet proteins that regulate cellular lipid stores. *Biochimica et Biophysica Acta* 1791, 419–440.
- Bostrom, P., Andersson, L., Rutberg, M., *et al.*, 2007. SNARE proteins mediate fusion between cytosolic lipid droplets and are implicated in insulin sensitivity. *Nature Cell Biology* 9, 1286–1293.
- Brasaemle, D.L., 2007. Thematic review series: Adipocyte biology. The perilipin family of structural lipid droplet proteins: Stabilization of lipid droplets and control of lipolysis. *Journal of Lipid Research* 48, 2547–2559.
- Bulankina, A.V., Deggerich, A., Wenzel, D., *et al.*, 2009. TIP47 functions in the biogenesis of lipid droplets. *Journal of Cell Biology* 185, 641–655.
- Czabany, T., Wagner, A., Zweytick, D., *et al.*, 2008. Structural and biochemical properties of lipid particles from the yeast *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 283, 17065–17074.
- Das, S.K., Eder, S., Schauer, S., *et al.*, 2011. Adipose triglyceride lipase contributes to cancer-associated cachexia. *Science* 333, 233–238.
- Gross, D.A., Zhan, C., Silver, D.L., 2011. Direct binding of triglyceride to fat storage-inducing transmembrane proteins 1 and 2 is important for lipid droplet formation. *Proceedings of the National Academy of Sciences of the United States of America* 108, 19581–19586.
- Hapala, I., Marza, E., Ferreira, T., 2011. Is fat so bad? Modulation of endoplasmic reticulum stress by lipid droplet formation. *Biology of the Cell/Under the Auspices of the European Cell Biology Organization* 103, 271–285.
- Hsieh, K., Lee, Y.K., Lontos, C., *et al.*, 2012. Perilipin family members preferentially sequester to either triacylglycerol-specific or cholesteryl-ester-specific intracellular lipid storage droplets. *Journal of Cell Science* 125, 4067–4076.
- Kurat, C.F., Wolinski, H., Petschnigg, J., *et al.*, 2009. Cdk1/Cdc28-dependent activation of the major triacylglycerol lipase Tgl4 in yeast links lipolysis to cell-cycle progression. *Molecular Cell* 33, 53–63.
- Lohmann, D., Spandl, J., Stevanovic, A., *et al.*, 2013. Monoubiquitination of ancient ubiquitous protein 1 promotes lipid droplet clustering. *PLoS One* 8, e72453.
- Melo, R.C., Paganoti, G.F., Dvorak, A.M., Weller, P.F., 2013. The internal architecture of leukocyte lipid body organelles captured by three-dimensional electron microscopy tomography. *PLoS One* 8, e59578.
- Murphy, S., Martin, S., Parton, R.G., 2010. Quantitative analysis of lipid droplet fusion: Inefficient steady state fusion but rapid stimulation by chemical fusogens. *PLoS One* 5, e15030.
- Nagayama, M., Uchida, T., Gohara, K., 2007. Temporal and spatial variations of lipid droplets during adipocyte division and differentiation. *Journal of Lipid Research* 48, 9–18.
- Nishino, N., Tamori, Y., Tateya, S., *et al.*, 2008. FSP27 contributes to efficient energy storage in murine white adipocytes by promoting the formation of unilocular lipid droplets. *Journal of Clinical Investigation* 118, 2808–2821.
- Oelkers, P., Cromley, D., Padamsee, M., Billheimer, J.T., Sturley, S.L., 2002. The DGA1 gene determines a second triglyceride synthetic pathway in yeast. *Journal of Biological Chemistry* 277, 8877–8881.
- Ohsaki, Y., Cheng, J., Suzuki, M., Fujita, A., Fujimoto, T., 2008. Lipid droplets are arrested in the ER membrane by tight binding of lipidated apolipoprotein B-100. *Journal of Cell Science* 121, 2415–2422.
- Penno, A., Hackenbroich, G., Thiele, C., 2013. Phospholipids and lipid droplets. *Biochimica et Biophysica Acta* 1831, 589–594.
- Sandager, L., Gustavsson, M.H., Stahl, U., *et al.*, 2002. Storage lipid synthesis is non-essential in yeast. *Journal of Biological Chemistry* 277, 6478–6482.
- Santos, C.R., Schulze, A., 2012. Lipid metabolism in cancer. *FEBS Journal* 279, 2610–2623.
- Spandl, J., White, D.J., Peychl, J., Thiele, C., 2009. Live cell multicolor imaging of lipid droplets with a new dye, LD540. *Traffic* 10 (11), 1579–1584.
- Stevanovic, A., Thiele, C., 2013. Monotopic topology is required for lipid droplet targeting of ancient ubiquitous protein 1. *Journal of Lipid Research* 54, 503–513.
- Straub, B.K., Herpel, E., Singer, S., *et al.*, 2010. Lipid droplet-associated PAT-proteins show frequent and differential expression in neoplastic steatogenesis. *Modern Pathology: An Official Journal of the United States and Canadian Academy of Pathology, Inc* 23, 480–492.
- Szymanski, K.M., Binns, D., Bartz, R., *et al.*, 2007. The lipodystrophy protein seipin is found at endoplasmic reticulum lipid droplet junctions and is important for droplet morphology. *Proceedings of the National Academy of Sciences of the United States of America* 104, 20890–20895.
- Tauchi-Sato, K., Ozeki, S., Houjou, T., Taguchi, R., Fujimoto, T., 2002. The surface of lipid droplets is a phospholipid monolayer with a unique Fatty Acid composition. *Journal of Biological Chemistry* 277, 44507–44512.
- Waltermann, M., Steinbuechel, A., 2005. Neutral lipid bodies in prokaryotes: Recent insights into structure, formation, and relationship to eukaryotic lipid depots. *Journal of Bacteriology* 187, 3607–3619.
- Welte, M.A., 2009. Fat on the move: Intracellular motion of lipid droplets. *Biochemical Society Transactions* 37, 991–996.
- Wilfling, F., Wang, H., Haas, J.T., *et al.*, 2013. Triacylglycerol synthesis enzymes mediate lipid droplet growth by relocating from the ER to lipid droplets. *Developmental Cell* 24, 384–399.
- Wolins, N.E., Brasaemle, D.L., Bickel, P.E., 2006. A proposed model of fat packaging by exchangeable lipid droplet proteins. *FEBS Letters* 580, 5484–5491.
- Yang, L., Ding, Y., Chen, Y., *et al.*, 2012. The proteomics of lipid droplets: Structure, dynamics, and functions of the organelle conserved from bacteria to humans. *Journal of Lipid Research* 53, 1245–1253.
- Zanghellini, J., Wodlei, F., von Grunberg, H.H., 2010. Phospholipid demixing and the birth of a lipid droplet. *Journal of Theoretical Biology* 264, 952–961.

Further Reading

- Brasaemle, D.L., Wolins, N.E., 2012. Packaging of fat: An evolving model of lipid droplet assembly and expansion. *Journal of Biological Chemistry* 287, 2273–2279.

- Digel, M., Ehehalt, R., Fullekrug, J., 2010. Lipid droplets lighting up: Insights from live microscopy. *FEBS Letters* 584, 2168–2175.
- Fei, W., Du, X., Yang, H., 2011. Seipin, adipogenesis and lipid droplets. *Trends in Endocrinology and Metabolism: TEM* 22, 204–210.
- Goodman, J.M., 2009. Demonstrated and inferred metabolism associated with cytosolic lipid droplets. *Journal of Lipid Research* 50, 2148–2156.
- Konige, M., Wang, H., Sztalryd, C., 2014. Role of adipose specific lipid droplet proteins in maintaining whole body energy homeostasis. *Biochimica et Biophysica Acta* 1842, 393–401.
- Kuhnlein, R.P., 2011. The contribution of the *Drosophila* model to lipid droplet research. *Progress in Lipid Research* 50, 348–356.
- Lass, A., Zimmermann, R., Oberer, M., Zechner, R., 2011. Lipolysis – a highly regulated multi-enzyme complex mediates the catabolism of cellular fat stores. *Progress in Lipid Research* 50, 14–27.
- Pol, A., Gross, S.P., Parton, R.G., 2014. Review: Biogenesis of the multifunctional lipid droplet: Lipids, proteins, and sites. *Journal of Cell Biology* 204, 635–646.
- Radulovic, M., Knittelfelder, O., Cristobal-Sarramian, A., *et al.*, 2013. The emergence of lipid droplets in yeast: Current status and experimental approaches. *Current Genetics* 59, 231–242.
- Suzuki, M., Shinohara, Y., Ohsaki, Y., Fujimoto, T., 2011. Lipid droplets: Size matters. *Journal of Electron Microscopy* 60 (Suppl 1), S101–S116.
- Walther, T.C., Farese Jr., R.V., 2012. Lipid droplets and cellular lipid metabolism. *Annual Review of Biochemistry* 81, 687–714.
- Xu, L., Zhou, L., Li, P., 2012. CIDE proteins and lipid metabolism. *Arteriosclerosis, Thrombosis, and Vascular Biology* 32, 1094–1098.
- Yang, H., Galea, A., Sytnyk, V., Crossley, M., 2012. Controlling the size of lipid droplets: Lipid and protein factors. *Current Opinion in Cell Biology* 24, 509–516.

Mechanisms and Functions of Mitochondrial Dynamics

K Mihara and H Otera, Kyushu University, Fukuoka, Japan

© 2016 Elsevier Inc. All rights reserved.

Glossary

Endophylin B1 It is also called Bif-1 and SH3GLB1; suppresses mitochondrial apoptosis; fatty acyl transferase; required for the regulation of mitochondrial outer membrane (MOM) dynamics; acting downstream of Drp1; cytoplasmic localization; targeted to MOM during apoptosis.

GDAP1 Ganglioside-induced differentiation-associated protein 1; mitochondrial fission factor; linked to Charcot-Marie-Tooth (CMT) disease; C-tail anchored MOM protein with two glutathione-S-transferase (GST) domains in the N- and C-terminal regions.

INF2 Inverted formin 2; stimulates both actin polymerization and depolymerization; exists as two isoforms differing in C-terminal sequence: the prenylated ER-bound isoform and nonprenylated cytoplasmic isoform.

MIB Mitofusin (Mfn)-binding protein; inhibits mitochondrial fusion; medium-chain dehydrogenase/reductase family protein.

MITOL/MARCH5 Membrane associated RING-CH E3 ubiquitin ligase; ubiquitin ligase of mitochondrial outer membrane (MOM); with four transmembrane domains (TMDs).

MTP18 Mitochondrial inner membrane (MIM) protein, 18 kDa; stimulates mitochondrial fission.

Prohibitin (PHB1/PHB2) MIM protein; proteins prohibiting cell-cycle progression; form multimeric ring complex in MIM; protein and lipid scaffold; PHB1 are PHB2 50% identical.

Tricoplein/mitostatin (TpMs) Weak homology to trichohyalin, plectin, and myosin; plectin/TpMs genome locus.

Introduction

Mitochondria are double membrane-bound organelles essential for numerous cellular processes, such as aerobic ATP generation, lipid biosynthesis, calcium signaling, and heme and iron-sulfur cluster biogenesis (Chan, 2012; Archer, 2013). Mitochondria are highly dynamic organelles that move along the cytoskeleton, continuously fusing and dividing in healthy cells, and their morphologic changes are essential for maintaining respiratory activity and mitochondrial DNA (mtDNA), and for controlling cellular processes such as embryonic development, cell-cycle progression, immune response, neuronal plasticity, calcium signaling, and apoptotic cell death. Mitochondrial fission contributes to the proper distribution of mitochondria in response to the local demand for ATP, as well as to the elimination of damaged mitochondrial fragments through mitophagy (autophagy for mitochondria) for mitochondrial quality control. Mitochondrial fusion, on the other hand, facilitates the exchange of mtDNA and other components between mitochondria to buffer transient defects in mitochondrial function (Figure 1). Mitochondrial fusion and fission are controlled by four high molecular weight GTPases conserved from yeast to mammals; mitofusins Mfn1 and Mfn2 (Fzo1 in yeast) in mitochondrial outer membrane (MOM) fusion (Figure 2); Opa1 (a causal gene product of optic atrophy type I; Mgm1 in yeast) in mitochondrial inner membrane (MIM) fusion and cristae organization (Figure 3); and cytoplasmic Drp1 (Dnm1 in yeast) in mitochondrial fission (Figure 4) indicating that the fundamental mechanisms controlling mitochondrial dynamics have been maintained during evolution. Abnormal mitochondrial dynamics cause neuronal synaptic loss and cell death in several human neurologic diseases, such as Charcot-Marie-Tooth neuropathy II (CMT2A), Alzheimer's disease, Parkinson's disease, Huntington's disease,

and amyotrophic lateral sclerosis. Furthermore, mitochondrial dynamics are suggested to be responsible for the pathogenesis of other complex diseases, including cancer, cardiovascular disease, and oxidative stress-associated disease, and have also been linked to aging (Chan, 2012; Archer, 2013).

Mitochondrial Fusion Machinery and Regulation

Mfn1 and Mfn2

Mitochondrial fusion of closely apposed mitochondria is a complex regulatory process involving multiple proteins that fuse the MOM and MIM of each mitochondrion, mediating two distinct membrane fusion events. Although fusion reactions between apposing MOMs and subsequent fusion between MIMs are normally highly synchronized, the two processes are functionally separable (Malka *et al.*, 2005). In cultured cells, the lack of mitochondrial fusion leads to defects in oxidative phosphorylation, and because fusion-deficient mitochondria cannot exchange contents, they are unable to restore and maintain the integrity of membranes, mtDNA, and the mtDNA-encoded proteins required for oxidative phosphorylation, and thus become respiration-defective (Chan, 2012).

As mentioned above, three large GTPase proteins, Mfn1, Mfn2, and Opa1 mediate mammalian mitochondrial fusion (Alexander *et al.*, 2000; Delettre *et al.*, 2000; Santel and Fuller, 2001). Mouse knockout (KO) studies clearly demonstrated that Mfn1 and Mfn2 are both essential for mitochondrial fusion (Chen *et al.*, 2003). They are anchored to the MOM through two transmembrane segments (TMs) with both an N-terminal GTPase followed by a coiled-coil heptad repeat (HR1), and a C-terminal HR2 exposing to the cytoplasm, and mediate MOM fusion in a GTPase-dependent manner (Figure 2). Mutations in Mfn2 cause CMT2A (Zuchner *et al.*,

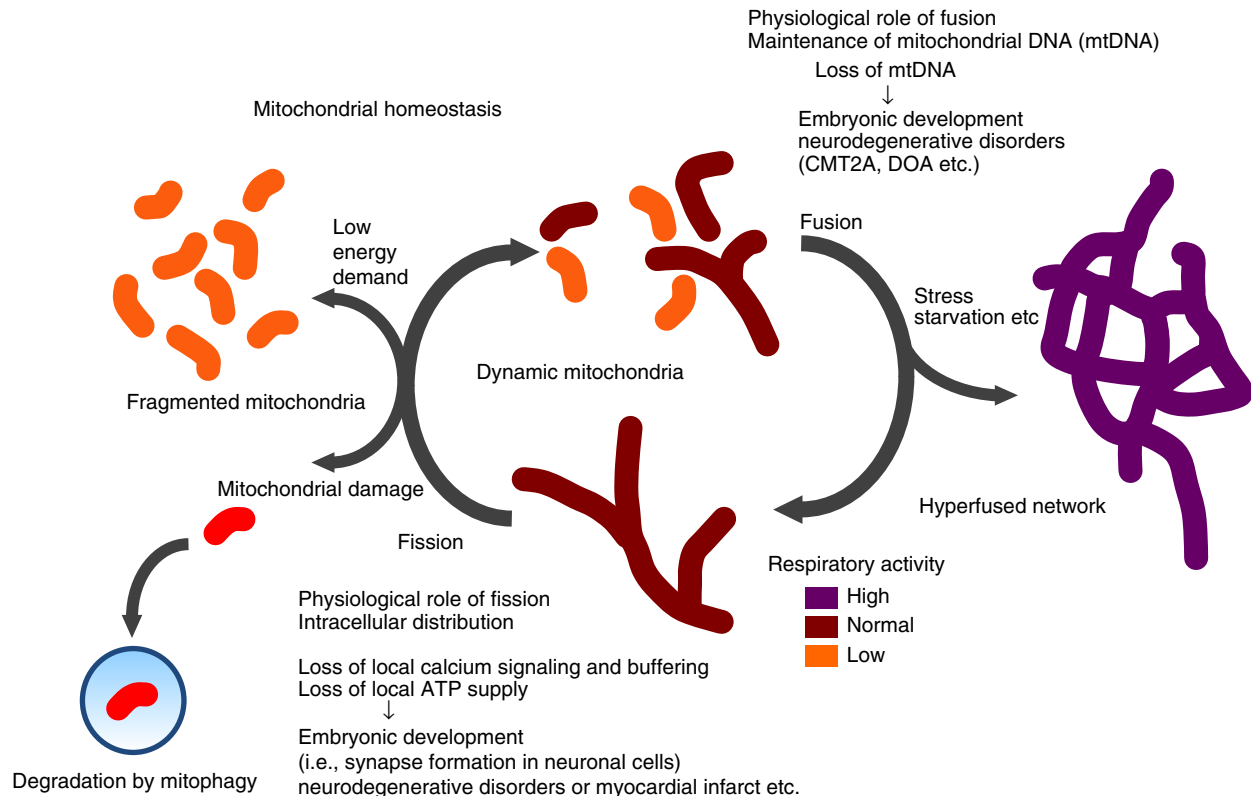


Figure 1 Physiological roles of mitochondrial fusion and fission.

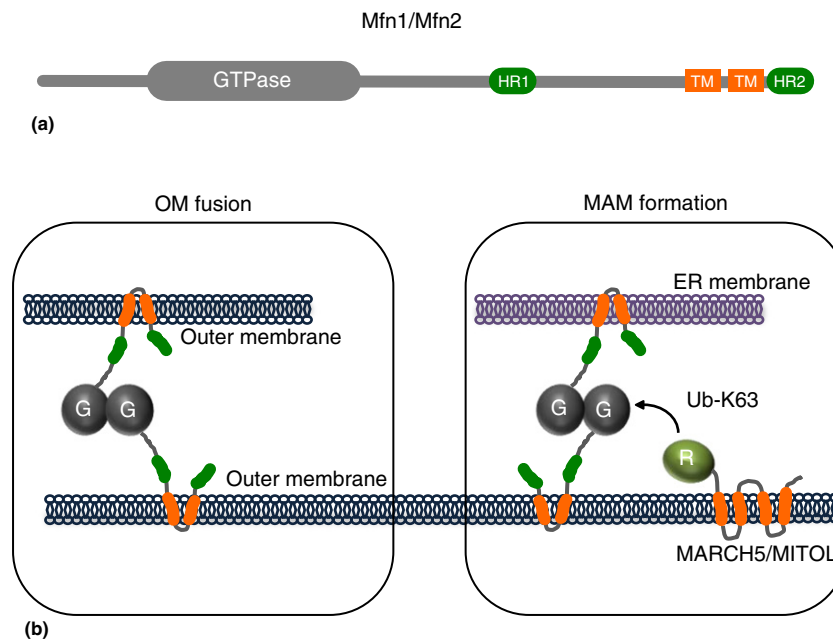


Figure 2 Mfn-mediated mitochondrial fusion and MAM formation. (a) Domain structure of Mfn1/Mfn2. HR1: heptad repeat 1, HR2: heptad repeat 2, TM: transmembrane domain, (b) Mfn proteins in mitochondrial fusion and MAM formation.

2004). The mechanism by which Mfn2 mutations cause CMT2A is not fully elucidated, but there are many loss-of-function Mfn2 mutations found within the GTPase domain in CMT2A. Studies of *in vivo* and *in vitro* membrane fusion

revealed that both Mfn1 and Mfn2 form homo- or hetero-protein complexes that are competent for fusion (Chen *et al.*, 2003; Ishihara *et al.*, 2004). Mfn1-KO results in small fragmented mitochondria that are broadly dispersed within the

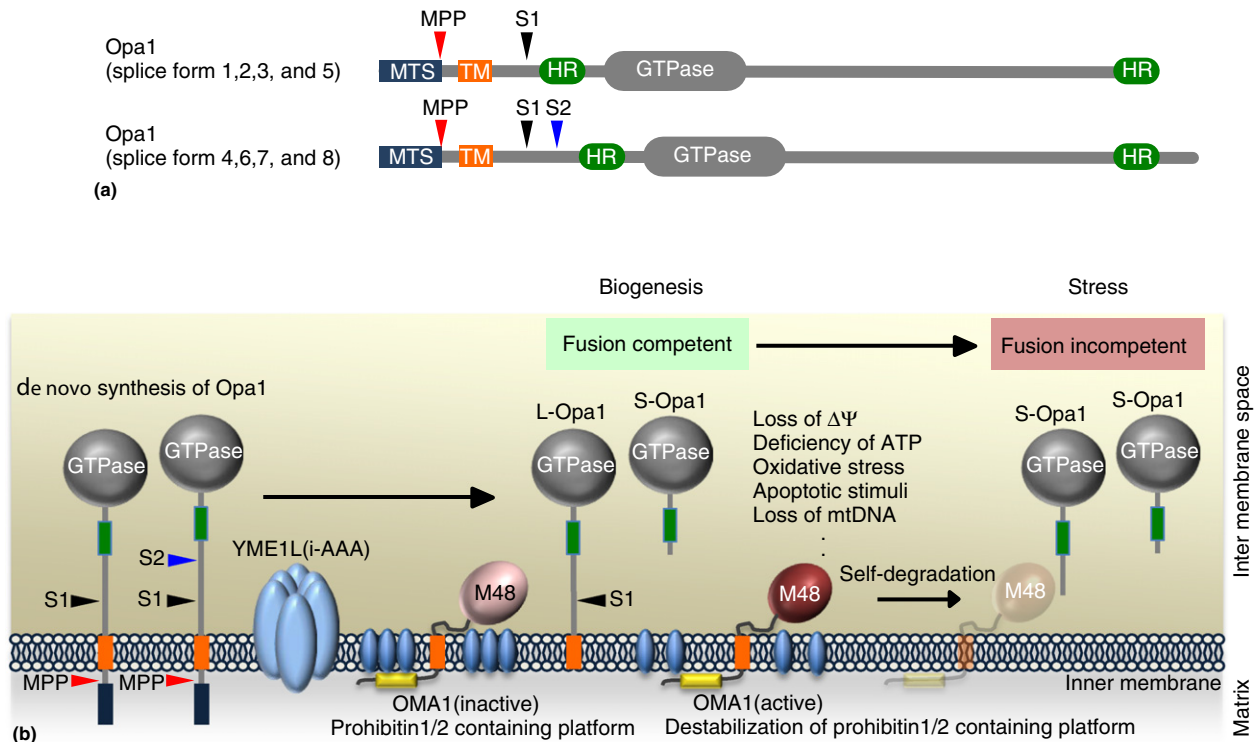


Figure 3 Processing of L-Opa1 to S-Opa1 under stable and stress conditions. (a) Domain structure of Opa1 splice variants. MTS: mitochondria-targeting sequence, MPP: mitochondrial processing peptidase, M48: M48-metalloprotease domain, (b) Constitutive and stress-induced processing of L-Opa1.

cell, whereas Mfn2-KO leads to large fragmented mitochondria concentrated near the nucleus (Chen *et al.*, 2003). Mfn1 and Mfn2 are functionally distinct within a hetero-oligomeric complex because the mitochondrial fusion defects associated with CMT2A mutations can be complemented by the expression of Mfn1, but not Mfn2 (Detmer and Chan, 2007). Consistent with these distinct KO phenotypes, purified recombinant Mfn1 exhibits higher GTPase activity than Mfn2 (Ishihara *et al.*, 2004). *In vitro* studies suggest that Mfn1 mediates mitochondrial tethering more efficiently than Mfn2 and thus homotypic *trans* Mfn1 interactions might be responsible for the initial GTP-dependent MOM tethering. How Mfn1 and Mfn2 promote MOM fusion beyond the initial tethering step is unclear. Atlastin (ATL)-mediated endoplasmic reticulum (ER) membrane fusion might represent the mechanistic background. ATL, an ER-bound GTP-dependent membrane fusion protein, is a causal gene product of hereditary spastic paraplegia in humans. GTP binding to ATL induces *trans*-association between the GTPase domains, tethering the membranes together, and GTP hydrolysis induces ATL conformational changes and rotation of the GTPase domain, which pulls the apposing membranes together, leading to membrane fusion (Bian *et al.*, 2011). How MOM and MIM fusion reactions are mechanistically coupled remains unknown.

Modulators of Mitofusin Proteins

In addition to its involvement in mitochondrial fusion, Mfn2 is enriched in the mitochondria-associated membranes

(MAM) of the ER, where Mfn1 and Mfn2 interact on the MOM to form interorganellar bridges (de Brito and Scorrano, 2008). Close contact between mitochondria and the ER is thought to be physiologically important for Ca^{2+} signaling, phospholipid transfer, metabolite exchange, and, therefore, regulation of mitochondrial metabolism and apoptosis (Friedman and Nunnari, 2014). Trichoplein/mitostatin, a keratin-binding protein that partly colocalizes with mitochondria, is involved in the regulation of ER-mitochondria tethering; its expression loosens tethering, whereas knockdown has the opposite effect (Cerqua *et al.*, 2010). Mitochondrial RING-CH E3-ubiquitin ligase, MARCH5/MITOL, regulates tight juxtaposition of mitochondria and ER through MAM by GTP-dependent K63-linked polyubiquitination of the mitochondria-localized form, but not the ER-associated form, of Mfn2 at K192. Therefore, MARCH5/MITOL-knockdown causes MAM dysfunction (Sugiura *et al.*, 2013).

It should be noted, however, that the ER-mitochondria contacts are still formed in mitofusin-deficient cells, however (Cosson *et al.*, 2012), and therefore other proteins may also participate in ER-mitochondria contacts. In yeast, the ER-mitochondria tethering complex, the ERMES complex comprising Mdm34, Mdm10, Mdm12, and Mmm1, forms the contacts involved in phospholipid exchange and mitophagy (Korrmann *et al.*, 2009). In contrast, however, the chemical entities of mammalian MAM, even single or multiple entities with distinct functions, have not been clarified, although many proteins with distinct functions are reported to be localized in MAM, such as PACS-2, Syntaxin17, MAVS, FACL4, PERK,

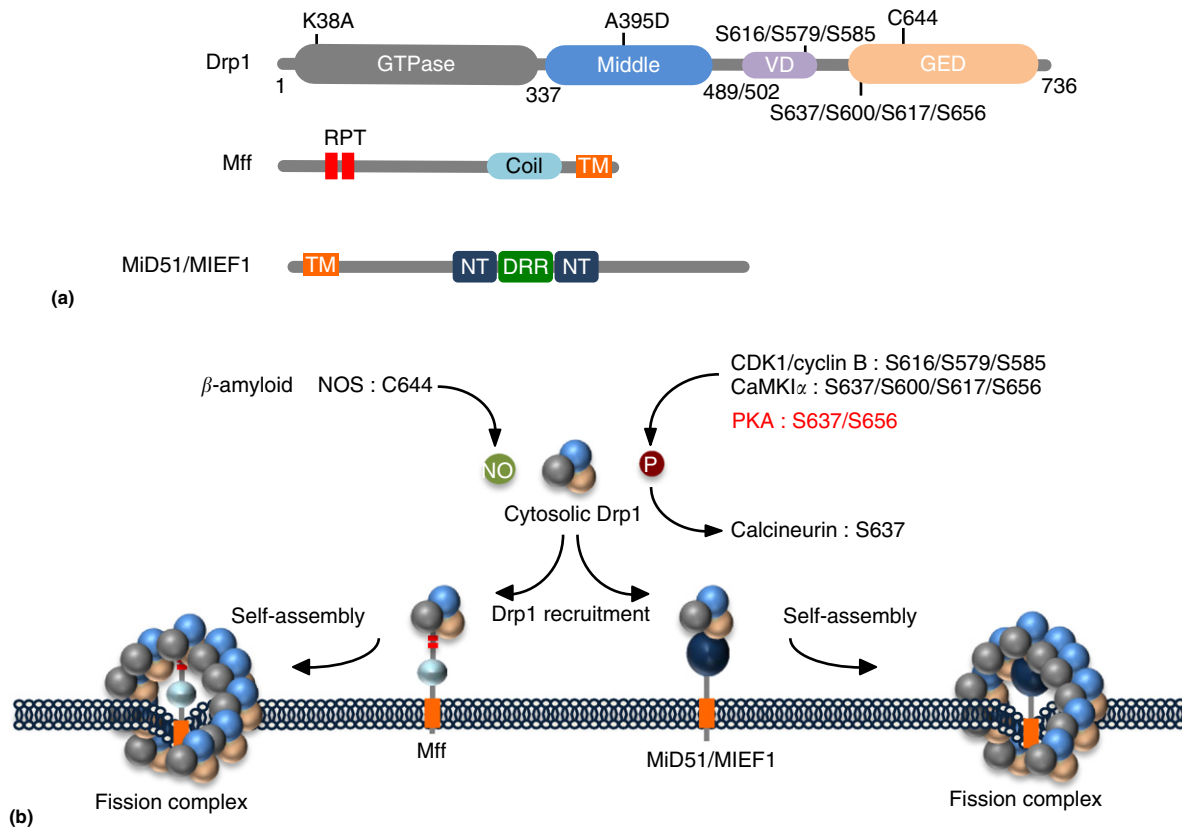


Figure 4 Mitochondrial recruitment of Drp1 and mitochondrial fission. (a) Domain structure of Drp1 and Drp1 receptors Mff and Mid 51/Mid49. Middle: middle domain, VD: variable domain, RPT: short amino acid repeats, coil: coiled-coil domain, NT: nucleotidyl transferase domain, DRR: Drp1 recruitment region, (b) regulation of mitochondrial recruitment of Drp1 by posttranslational modifications.

Rab32, and calnexin (Hamasaki *et al.*, 2013; Naon and Scorrano, 2014).

In addition to these players, a 55-kDa mitofusin binding protein (MIB) was identified as a fusion regulator. It is a member of the medium-chain dehydrogenase/reductase protein superfamily and has a conserved coenzyme-binding domain. MIB overexpression in HeLa cells induces mitochondrial fragmentation, whereas MIB silencing leads to the formation of elongated mitochondria and to significant defects in cell growth (Eura *et al.*, 2006). Therefore, MIB functions as a negative regulator of mitofusin proteins essential for cellular function.

Mitochondrial surface-localized phospholipase D (mito-PLD) functions downstream of mitofusin and promotes Mfn1/2-dependent mitochondrial fusion by hydrolysis of cardiolipin to generate phosphatidic acid. The thus-generated phosphatidic acid probably facilitates fusion by generating negative curvature in the opposing lipid bilayers through recruiting or activating other key proteins or through further conversion to other fusogenic lipids (Choi *et al.*, 2006). A recent *in vitro* mitochondrial fusion assay revealed that a Bcl-2 family proapoptotic protein, Bax (soluble monomeric form), specifically stimulates mitochondrial fusion through homotypic Mfn2 *trans* mitochondrial complexes, again indicating functional and mechanistic differences between Mfn1 and Mfn2, although the physiologic significance remains unclear (Hoppins

et al., 2011). Of note, Mfn2 ablation induces ER stress. The Zorzano laboratory recently reported that Mfn2, as an upstream modulator, interacts with an unfolded-response sensor PERK on the ER and coordinately regulates the ER unfolded protein response and mitochondrial function (Schneeberger *et al.*, 2013; Munoz *et al.*, 2013); the mechanism involved in the diet-dependent regulation of metabolism in specific neuronal circuits that affects the metabolic status of animals (see Section Mfn1-, Mfn2-, and Opa1-Dependent Mitochondrial Fusion).

Opa1, a Key Player in MIM Fusion

Opa1 is another key mediator essential for MIM fusion, cristae remodeling, and maintenance of mtDNA (Olichon *et al.*, 2003). Mutations in Opa1 cause autosomal dominant optic atrophy, a degenerative disease of the optic nerve, and Opa1-KO mice are embryonic lethal (E ~13.5) (Alexander *et al.*, 2000; Delettre *et al.*, 2000; Davis *et al.*, 2007). There are eight Opa1 splice variants that are all synthesized as precursor proteins with the mitochondrial localization signal in the N-termini and following hydrophobic stretches, which are responsible for sorting the protein into the MIM (Ishihara *et al.*, 2006). During mitochondrial import, the mitochondrial localization signal of Opa1 precursors is removed by mitochondrial processing peptidase (MPP) to form L-forms

(Ishihara *et al.*, 2006). They are anchored to the MIM as punctate structures exposing both a GTPase domain and a protease cleavable site S1 to the mitochondrial intermembrane space (IMS) (Olichon *et al.*, 2003; Figure 3). In four of the splice variants (variants 4, 6, 7, and 8), exon 5 provides an additional protease cleavage site S2 localized in IMS. The IMS AAA-protease (i-AAA protease) YME1L is responsible for constitutive cleavage of L-Opa1 at S2, whereas constitutive cleavage at the S1 site is mediated by m-AAA proteases (Afg3L1, Afg3L2, and Paraplegin for mice mitochondria, and Afg3L2 and Paraplegin for human mitochondria) and OMA1 (for overlapping activity with m-AAA protease. OMA1 is also involved in the stress-induced L-Opa1 processing; see below). These cleavages yield short isoforms (S-Opa1) that lack the TM segment (Figure 3; Ishihara *et al.*, 2006; Barker *et al.*, 2014). Of note, Chan and colleagues demonstrated using Opa1-KO cells that the L- and S-isoforms have little activity on their own, but the presence of both forms fully complements the function (Song *et al.*, 2007), indicating that mitochondrial function depends on a proper ratio of L- and S-Opa1 isoforms. As detailed below, the mitochondrial stress-induced processing of L-Opa1 by OMA1 occurs only at the S1 site (Song *et al.*, 2007; Baker *et al.*, 2014).

Prohibitins (PHB1 and PHB2), evolutionarily conserved MIM proteins that function as protein- and lipid-scaffolds and are essential for cell proliferation and development, regulate Opa1 processing (Merkwirth *et al.*, 2008). Interestingly, defective phenotypes of PHB1- or PHB2-KO cells, such as those with growth defects, mitochondrial fragmentation, susceptibility to apoptosis, and defective cristae morphogenesis, are all complemented by the exogenous expression of an uncleavable L-Opa1 mutant, indicating that the PHB complex is epistatic to Opa1 processing (Merkwirth *et al.*, 2008). Mutations in the GTPase domain lead to fragmented mitochondria, indicating that GTP hydrolysis is essential for mitochondrial fusion activity. Many Opa1 mutations found in dominant optic atrophy are detected in the GTPase domain. The mitochondrial IMS-exposed region locating next to the GTPase domain is required for the tetramerization and higher-order self-assembly of Opa1 (Ramachandran *et al.*, 2007). In yeast, it was demonstrated that l-Mgm1 functions as a membrane anchor for s-Mgm1 and the GTPase-defective l-Mgm1 mutant is functional (Zick *et al.*, 2009). In this relation, both l- and s-Mgm1 exist as inactive GTPase monomers in the absence of the membrane, but together in *trans* they form a functional dimer in a cardiolipin-dependent manner, providing the building block for higher-order assemblies (DeVay *et al.*, 2009). Cardiolipin is also important for stimulating GTPase and the assembly of Mgm1 and Opa1 (Rujiviphat *et al.*, 2009; Ban *et al.*, 2010).

Several types of mitochondrial dysfunction, such as loss of membrane potential, ATP deficiency, oxidative stress, apoptosis, or loss of mtDNA, induce remarkable fragmentation of mitochondria concomitant with a rapid conversion of L-Opa1 to S-Opa1 isoforms (Ishihara *et al.*, 2006; Chan, 2012). This 'induced' Opa1 processing is mediated by OMA1 (Ehse *et al.*, 2009; Baker *et al.*, 2014). OMA1 is a MIM-bound zinc protease with the N-terminal domain localized at the matrix side of the MIM and the large C-terminal M48 metalloprotease domain exposed to the IMS (Kaesler *et al.*, 2003; Baker *et al.*, 2014) that degrades mis-assembled membrane proteins in yeast

(Figure 3). It is synthesized as a precursor form with the N-terminal MTS. The MIM-integrated mature form of OMA1 is activated by various mitochondrial stimuli to cleave L-Opa1 at the S1 site and, interestingly, this activation is associated with its complete autocatalytic self-degradation. This is considered to be the mechanism ensuring the reversibility of the Opa1-mediated mitochondrial fusion response (Baker *et al.*, 2014; Zhang *et al.*, 2014). Mature OMA1 has an N-terminal domain composed of a hydrophobic region followed by a cluster of positively-charged amino acids; the former stretch is required for OMA1 protease activity, whereas the latter stretch functions as a stress sensor (Baker *et al.*, 2014; Figure 3). Neither knockdown nor overexpression of OMA1 affects mitochondrial morphology in mouse embryonic fibroblast cells, suggesting that OMA1 is dispensable for the balanced formation of L-Opa1 and S-Opa1 by constitutive cleavage under normal conditions. The regulation of Opa1 activities by OMA1 likely affects tissue development or the progression of neurodegenerative diseases. OMA1-KO mice develop and survive normally, although they exhibit marked upregulation of genes in lipid and glucose metabolic pathways, suggesting that the OMA1-Opa1 axis is essential for metabolic homeostasis (Quiros *et al.*, 2012).

Recent genome-wide RNA interference (RNAi) screening led to the identification of ROMO1 (for Reactive Oxygen species Modulator 1), which regulates Opa1-dependent mitochondrial fusion (Norton *et al.*, 2014). It is a 79-amino acid MIM protein that is inserted into the membrane by two TM helices with both N- and C-termini exposed to the IMS, and it promotes mitochondrial reactive oxygen species (ROS) production, cell proliferation, and senescence (Chung *et al.*, 2006; Na *et al.*, 2008). Of note, ROMO1 is also characterized as Higd-1a (for hypoxia-induced gene domain protein-1a), which is required for the morphologic and functional integrity of mitochondria (An *et al.*, 2013). ROM1/Higd-1a binds to Opa1 and stabilizes the Opa1 oligomeric complex and L-isoforms, and maintains the integrity of cristae junction; ROMO1/Higd-1a-depletion results in mitochondrial fragmentation, cristae disorganization, impaired mitochondrial respiration, and increased sensitivity to apoptotic stimuli (Norton *et al.*, 2014). ROMO1/Higd-1a regulates cristae dynamics in response to changes in the mitochondrial redox state and couples mitochondrial ROS to Opa1 oligomerization and cristae junction integrity.

Mitochondrial fusion depends on ATP (Schauss *et al.*, 2010). Chan and colleagues recently demonstrated using an *in vitro* fusion assay with isolated mitochondria that mitochondrial MIM fusion, but not MOM fusion, is sensitive to the state of oxidative phosphorylation or cellular metabolic states; the metabolic states regulate MIM fusion through the activation of Yme1L-dependent S2 cleavage of L-Opa1, suggesting a mechanism linking mitochondrial fusion with cellular metabolic states (Mishra *et al.*, 2014).

Langer and colleagues, using cell lines lacking either or both YME1L and OMA1, recently demonstrated that Opa1 processing is dispensable for mitochondrial fusion and further revealed an unexpected stimulatory role of S-Opa1 in mitochondrial fission under conditions in which membrane fusion normally proceeds (Anand *et al.*, 2014). Conceivably, S-Opa1 interacts with the fission machinery at the MOM or is involved

in MIM fission. Interestingly, S-Opa1 with the GTPase mutation partly localizes in the ER-mitochondria contact sites and mitochondrial fission machinery as punctate structures (Anand *et al.*, 2014).

Molecular Insights into Drp1 Actions During Mitochondrial Fission

Drp1, a Key Player

Drp1 is a member of the conserved dynamin GTPase superfamily, which includes a broad range of membrane fission proteins, and mediates mitochondria and peroxisome fission events. Drp1 is a cytosolic protein with an N-terminal GTPase domain thought to provide mechanical force, a dynamin-like middle domain, and a GTPase effector domain (GED) located in the C-terminal region (Figure 4). A dominant-negative middle domain mutation (A395D) in Drp1 was reported in a patient with a lethal disorder exhibiting microcephaly, abnormal brain development, optic atrophy, and hypoplasia (Waterham *et al.*, 2007). Cells derived from this patient had aberrant mitochondrial elongation. Drp1 mainly localizes in the cytosol, and during mitochondrial fission, translocates from the cytosol to prospective fission sites on the mitochondria (Figure 4) (Smirnova *et al.*, 2001). *In vitro* studies as well as studies of yeast Dnm1 revealed that Dnm1 or Drp1 assembles, like dynamin, into higher-order structures that wrap around the mitochondrial tubule (Smirnova *et al.*, 2001; Yoon *et al.*, 2001; Ingeman *et al.*, 2005). These spiral-shaped higher-order structures are thought to constrict and eventually sever the mitochondrial membrane by a GTP hydrolysis-dependent mechanism. Time-lapse imaging of GFP-tagged Drp1 shows that mitochondrial tubules divide at sites where these punctate structures are found (Smirnova *et al.*, 2001). The GTP-binding defective mutant (K38A) sequesters endogenous Drp1 into uncharacterized aggregated or dotted structures, thus inhibiting its localization on the mitochondrial fission sites to act as a dominant-negative mutant (Smirnova *et al.*, 2001; Yoon *et al.*, 2001). Intermolecular interactions between the N-terminal GTPase domain and C-terminal GED are also important for Drp1 self-assembly and functional regulation (Zhu *et al.*, 2004). The regulation of MIM scission as well as the mechanisms synchronizing scissions of the MIM and MOM are not yet known. The MIM-localized MTP18 is a transcriptionally regulated target of the phosphatidylinositol 3-kinase signaling and regulates mitochondrial fission coupled with the action of Drp1 at the MOM (Tondera *et al.*, 2005). Depletion and overexpression of MTP18 promote mitochondrial elongation and fragmentation, respectively, indicating its role in the regulation of mitochondrial fission. The identification of its interacting partners will help to understand the molecular mechanisms that synchronize MIM and MOM scission.

Key Players on the MOM in Drp1-Dependent Mitochondrial Fission: Mff, Mid49, and Mid51/MIEF1

Despite extensive studies, the mechanisms by which the cytoplasmically localized Drp1 is activated and recruited to the

prospective mitochondrial fission sites remain unclear. Cell biologic and genetic studies recently identified three MOM proteins involved in Drp1-dependent mitochondria fission, mitochondrial fission factor (Mff) and mitochondrial dynamics protein of 49- and 51-kDa (Mid49 and Mid51; also named mitochondrial elongation factor 1 (MIEF1) for Mid51) (Figure 4). Mff is a C-tail anchored protein on the MOM that was first identified in a *Drosophila* RNAi library search for mitochondrial morphology alterations (Gandre-Babbe and van der Bliek, 2008). Mammalian mitochondria contain an Mff ortholog and silencing this factor by RNAi induces mitochondrial elongation in mammalian cells. The specific role of Mff in mitochondrial fission, however, remains unknown. We demonstrated that (1) Mff RNAi affects the mitochondrial recruitment of Drp1. Endogenous Drp1, observed as dotted structures on mitochondria, was clearly decreased and dispersed in the cytoplasm in Mff RNAi cells, concomitant with mitochondrial network extension and acquisition of resistance to apoptotic stimuli. (2) In contrast, Mff overexpression induces mitochondrial fragmentation, concomitant with increased Drp1 recruitment to the mitochondria and increased apoptotic sensitivity. (3) Consistent with these observations, both *in vitro* and *in vivo* experiments demonstrated that Mff transiently interacts with Drp1 through its N-terminal cytoplasmic region. Furthermore, (4) Mff mostly colocalizes with the Drp1 foci on the MOM in marked contrast to the uniform localization of hFis1 in the MOM. (5) Of note, conditional knockout of hFis1 in colon carcinoma cells revealed that it is dispensable for mitochondrial fission. These observations indicate that Mff functions as a Drp1 receptor on the MOM to mediate mitochondrial fission (Otera *et al.*, 2010).

Mid49 and Mid51/MIEF1 were identified by random cellular localization screening of uncharacterized human proteins for unique changes in mitochondrial distribution and morphology (Palmer *et al.*, 2011; Zhao *et al.*, 2011). Mid49 and Mid51/MIEF share 45% sequence identity and specifically localize to mitochondria, but not to peroxisomes, and anchor to the MOM through N-terminal TMs exposing bulk C-terminal segments to the cytoplasm (Figure 4). Mid49 and Mid51 form foci and rings around mitochondria and can act independently of Mff in Drp1 recruitment and are specific for mitochondrial fission. Overexpression of Mid49 and Mid51 inhibits Drp1 function by acting in a dominant-negative manner to sequester Drp1 specifically at mitochondria, leading to unopposed fusion of mitochondria and peroxisomes. This elongation depends on the unopposed function of Mfn1 and Mfn2, thus indicating that the Mid proteins are not fusion factors; marked contrast to the suggested function for MIEF (Zhao *et al.*, 2011). In contrast to Fis1 and Mff, Mid49/Mid51 is not targeted to peroxisomes or lysosomes; although when artificially targeted using a peroxisomal- or lysosomal-targeting signal, Drp1 is specifically recruited to these organelles (Palmer *et al.*, 2013). The Drp1 recruitment activity appeared stronger than that of Mff. In contrast to the overexpression effect, different knockdown phenotypes were reported by two laboratories. Ryan and colleagues reported that knockdown of either Mid49 or Mid51 had no effect on mitochondrial morphology, but simultaneous knockdown of both induced mitochondrial elongation. In contrast, Nister and colleagues reported that knockdown of Mid51/MIEF induces

mitochondrial fragmentation, leading to the conclusion that Mid51/MIEF recruits Drp1 to the MOM and promotes mitochondrial fusion rather than fission. Experiments with receptor-specific KO cells will be required to resolve these inconsistencies.

The crystal structure of the cytoplasmic domain of human Mid51 was recently solved by X-ray crystallography (Richter *et al.*, 2014; Loson *et al.*, 2014). It contains a variant of the nucleotidyl transferase domain (NT) that binds ADP in the middle part of the molecule (Richter *et al.*, 2014; Loson *et al.*, 2014), and the surface-exposed Drp1 recruitment loop (DRR) (Figure 4). Mid51 is present as a dimer and ADP binding does not affect its oligomeric state or Drp1 binding. Strikingly, experiments both *in vivo* and *in vitro* with the ADP-binding mutant or dimerization mutant revealed that Mid51 is competent for Drp1 recruitment irrespective of the presence or absence of ADP, but its potential for ADP binding and dimerization is essential for Drp1 activation by stimulating oligomeric assembly of Drp1 on the mitochondrial surface in response to fission-stimulating signals (Loson *et al.*, 2014). Thus, ADP is essential for Mid51 function during mitochondrial fission. Of note, Mid49 has an only partially conserved nucleotidyl transferase domain and the function is not regulated by ADP (Loson *et al.*, 2014).

Drp1 might self-assemble via its ability to homo-oligomerize on Mff, Mid49, or Mid51 of the mitochondrial surface, forming spiral structures around the mitochondrial tubules. After this process, unidentified Mff- or Mids-interacting factors might affect the assembly and activation of the fission machinery, leading to membrane constrictions or lipid remodeling and eventually to membrane scission. In yeast, the MOM protein Fis1 functions as the fission receptor for Dnm1 via adapter proteins Mdiv1/Caf4. In contrast to the conservation of Fis1 through various species, there are no obvious homologs of yeast fission adapters Mdiv1/Caf4 in metazoans. Mammalian mitochondria seem to adopt fission mechanisms that are distinct from those of yeast or plants. As discussed in the following section, mammalian Fis1 does not appear to be directly involved in mitochondrial fission as a Drp1 receptor.

The capacity of the above-mentioned adapter proteins to recruit Drp1 was elegantly analyzed by Shaw's group in yeast strains lacking all fission proteins, which revealed that cytoplasmic domains of Mff or Mids targeted to the yeast MOM by the respective MOM-targeting signal are sufficient to recruit human Drp1 to yeast mitochondria, facilitating Drp1's oligomeric assembly around constriction sites to eventually catalyze mitochondrial fission, although human Fis1 failed to both recruit Drp1 and induce mitochondrial fission in this assay system (Koirara *et al.*, 2013).

What is the functional division of DRP1-daptor proteins Mff, Mid49, and Mid51 in the Drp1-dependent mitochondrial fission reaction? Studies to date have not identified significant differences in the mitochondrial fission events mediated by these adapters except for the overexpression phenotype; Mid proteins, but not Mff, exhibit a dominant-negative effect on mitochondrial fission (Otera *et al.*, 2010; Palmer *et al.*, 2011; Zhao *et al.*, 2011). Posttranslational modifications of Drp1, including phosphorylation, sumoylation, nitrosylation, and ubiquitination (see Section Regulation of Drp1 by Post-translational Modifications), could influence the identity of

the adapters, or these modifications of the adapters could regulate the function of the adapters themselves. Another possibility is that multiple adapters work together with Drp1 as a single division machinery.

How are future mitochondrial fission sites determined? Immuno-electron microscopy and time-lapse fluorescence microscopy in yeast and mammalian cells showed that Dnm1/Drp1 colocalizes with constrictions on the presumed future mitochondrial fission sites (Mozdy *et al.*, 2000; Smirnova *et al.*, 2001). Three-dimensional time-lapse fluorescence imaging revealed that constriction formation and Dnm1 recruitment as patches or ring structures are distinct processes (Legesse-Miller *et al.*, 2003). So, how are future fission sites selected? Friedman and colleagues addressed this problem with electron microscopy and live cell fluorescence microscopy in yeast and mammalian cells, and demonstrated that the ER wraps around mitochondria and mediates constrictions, where Mff accumulation and subsequent Drp1 recruitment proceed (Friedman *et al.*, 2011; Friedman and Nunnari, 2014). The ER-bound form of inverted formin (INF2) is reported to be involved in this process; INF2 accumulates at the ER-mitochondria contact sites, where INF2 promotes actin polymerization and myosin II accumulation to produce constrictions, thus enhancing the subsequent accumulation of Drp1 and fission (Korobova *et al.*, 2013; Korobova *et al.*, 2014; Figure 5).

Modulators of Mitochondrial Fission

Fis1

Fis1 is a C-tail anchored MOM protein with its N-terminal multiple tetratricopeptide repeat motif exposed to the cytoplasm (Yoon *et al.*, 2003). In yeast, Fis1 is required for Dnm1 recruitment. During mitochondrial fission, Fis1 transiently interacts via cytosolic adapter proteins Mdiv1/Caf4 with Dnm1 by its tetratricopeptide repeat motif, suggesting its function as a mitochondrial Dnm1 receptor (Lackner *et al.*, 2009). In mammals, Fis1 has also been identified in mitochondria (hFis1 for human Fis1) and is thought to be involved in recruiting Drp1 to mitochondria, as in yeast, through direct or indirect interactions (Yoon *et al.*, 2003). The actual function of hFis1, however, remains enigmatic, because Mdiv1/Caf4-like adapter proteins have not been identified, hFis1 localizes throughout the MOM unlike the punctate localization of Drp1, and mitochondrial recruitment of Drp1 is not affected by hFis1-knockdown (Otera *et al.*, 2010). TBC1D15, a Rab GTPase-activating protein (Rab-GAP), is a binding partner of hFis1, and furthermore TBC1D15 and another TBC family member TBC1D17 participate with hFis1 in autophagic encapsulation of damaged mitochondria in a later step of Parkin-PINK dependent mitophagy (Onoue *et al.*, 2013; Yamano *et al.*, 2014). Considering that Drp1 is involved in the early steps of mitophagy to segregate disintegrated mitochondria for degradation (Twig *et al.*, 2008), direct involvement of hFis1 in the regulation of Drp1 function seems to be very low.

Ganglioside-induced differentiation-associated protein 1 (GDAP1)

GDAP1 is another mitochondrial fission-related factor located on the MOM through a C-terminal hydrophobic

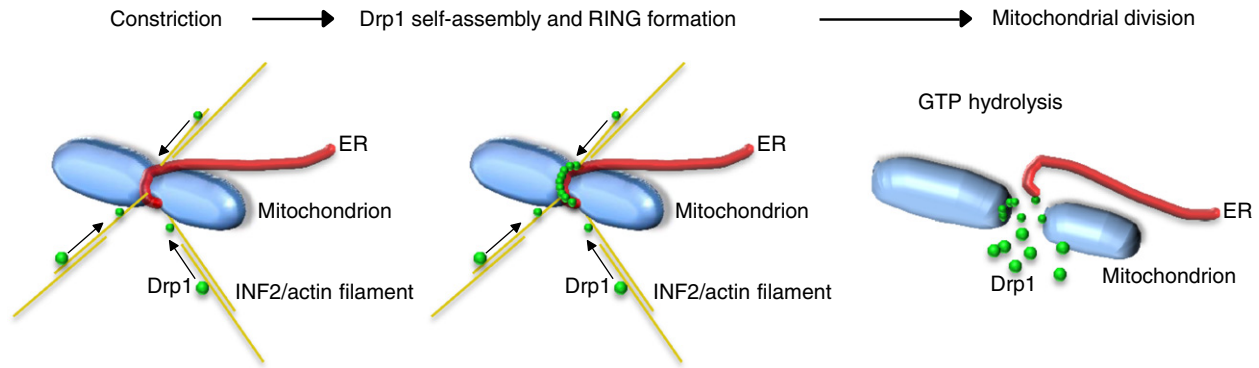


Figure 5 Endoplasmic reticulum (ER)-assisted mitochondrial fission. Speculative process of mitochondrial fission at the mitochondria-ER contact sites. The ER-bound form of inverted formin (INF2) wraps around mitochondrial tubules, and induces actin polymerization and myosin II accumulation to produce membrane constrictions, where adaptor proteins Mff and Mids (small green rods) accumulate to mediate Drp1 recruitment and its ring-like oligomeric assembly. GTP hydrolysis-mediated contractile force drives mitochondrial fission.

transmembrane domain extruding the bulk of the N-terminal domain into the cytoplasm (Niemann *et al.*, 2005). GDAP1 is involved in ganglioside maturation. Mutations in GDAP1 lead to the peripheral neuropathy Charcot-Marie-Tooth (CMT) disease, which affects Schwann cells, myelinating glia in the peripheral nervous system, and neurons (Cassereau *et al.*, 2011). Although it is not clear how GDAP1 is involved in mitochondrial fission, these data suggest the importance of membrane lipid compositions such as gangliosides in mitochondrial fission. GDAP1 mutants found in patients with CMT disease are not targeted to the mitochondria and lack mitochondrial fragmentation activity (Niemann *et al.*, 2005). GDAP1-induced mitochondrial fragmentation is inhibited by Drp1 knockdown or expression of the dominant-negative Drp1-K38A mutant, indicating that GDAP1 is a Drp1-dependent regulator of mitochondrial fission (Niemann *et al.*, 2005). Huber *et al.* (2013) recently reported that GDAP1 is also targeted to peroxisomes by the action of peroxisomal import receptor Pex19 and mediates Drp1- and Mff-dependent peroxisomal fission, as in mitochondrial fission. GDAP1 has two glutathione-S-transferase family domains and over-expression of wild-type GDAP1, but not CMT disease mutants, increases cellular glutathione levels and the mitochondrial membrane potential, although the functional relation to mitochondrial morphology regulation remains unknown (Noack *et al.*, 2012).

Endophilin B1/Bif-1

Endophilin B1 (also known as Bax-interacting factor 1, Bif-1) is a fatty acyl transferase mainly localizing in the cytoplasm that regulates mitochondrial morphology and MOM dynamics acting downstream of Drp1, probably through lipid remodeling of the MOM. Knockdown of endophilin B1 induces unusually interconnected or extended mitochondrial tubules, several of which have bleb-like structures (Karbowski *et al.*, 2004). This manipulation induces the dissociation of MOM and MIM compartments, suggesting that endophilin B1 is required for maintenance of MOM and MIM synchronous morphogenesis. Epistatic studies using cells subjected to single and double knockdown of Drp1 and endophilin B1 revealed

that Drp1 acts upstream of endophilin B1 in the maintenance of mitochondrial morphologic dynamics. Upon the induction of apoptosis, endophilin B1 is recruited from the cytoplasm to the mitochondria as punctate structures, where it interacts transiently with Bax oligomers, and promotes cytochrome c release from the mitochondria (Karbowski *et al.*, 2004; Rostovtseva *et al.*, 2009). The mechanism and physiologic function of Drp1-dependent mitochondrial recruitment of endophilin B1 require further analysis, although it is proposed to have a role in autophagosome formation (Takahashi *et al.*, 2009).

Sacsin

Sacsin is a 520-kDa protein encoded by the SACS gene, whose mutations cause childhood-onset autosomal recessive spastic ataxia of Charlevoix-Saguenay (ARSACS; Anderson *et al.*, 2010). At the N-terminus, Sacsin contains a ubiquitin-like domain, which is followed by three Sacsin Repeat Regions that possess ATPase activity, a 75-residue XPC-binding domain implicated in interactions with ubiquitin ligase Ube3A, a DNA J-like domain, and a HEPN (higher eukaryotes and prokaryotes nucleotide-binding) domain with an affinity for GTP. It is highly expressed in the central nervous system, and is also found in the skin, skeletal muscles, and, at low levels, in the pancreas. Recent work indicates that Sacsin is localized to the mitochondria, partially overlapping with the Drp1 foci (Giard *et al.*, 2012). Exogenously expressed Sacsin (1–1368 residue segment) interacts with endogenous Drp1. Interestingly, mitochondria in the cells of Sacsin-KO mice and in patients with the autosomal recessive spastic ataxia of Charlevoix-Saguenay display a hyperfused balloon-like mitochondrial morphology, mimicking the characteristic mitochondrial structures observed in Drp1- or Mff-knockdown cells or in dominant-negative Drp1-K38A-expressing cells. In sacsin-deficient neuronal cells, enlarged mitochondria are clustered and accumulate in the soma and proximal dendrites, and display fewer dendrites than control cells, a morphologic response observed in the neuronal cells of Drp1 KO mice (see below). Importantly, Sacsin-KO mice exhibit an age-dependent loss of cerebellar Purkinje cells, probably due to disturbances in the

mitochondrial delivery within neurites (Giard *et al.*, 2012). These findings suggest that Sacsin regulates Drp1 activity with a speculated function of chaperon activity.

Listeria toxin (LLO)

Of note, a Drp1-independent, atypical mitochondrial fission mode has been reported. *Listeria monocytogenes*, a food-borne pathogen capable of invading nonphagocytic cells, causes fragmentation of the host mitochondria by its pore-forming toxin LLO (55-kDa) before bacterial entry (Stavru *et al.*, 2013). A small dose of LLO efficiently induces mitochondrial fragmentation concomitant with transient dissipation of mitochondrial membrane potential, respiration, and cellular ATP levels. This seems to be an elaborate strategy of pathogens to induce transient shutdown of mitochondrial function to interfere with the host physiology. Interestingly, recombinant LLO induces the degradation of Mff concomitantly with the release of Drp1 from mitochondria to the cytoplasm and strong mitochondrial fragmentation. Moreover, it induces efficient mitochondrial fragmentation both in Drp1 KO MEF and Mff-KO HCT116 cells. Surprisingly, the ER and F-actin also have active roles in regulating LLO-induced mitochondrial fission, suggesting that this fission reaction partly shares the pathway with Drp1-dependent mitochondrial fission (Stavru *et al.*, 2013).

Regulation of Drp1 by Posttranslational Modifications

Various stressors outside or inside cells induce mitochondrial fission to remodel mitochondria and cellular functions. During apoptosis, cytoplasmic Drp1 is translocated to the mitochondria and induces mitochondrial fragmentation prior to caspase activation by the release of cytochrome *c* (Frank *et al.*, 2001). Such increased fission events are also important for autophagic clearance of depolarized (or dysfunctional) mitochondria or for proper segregation of mitochondria into daughter cells during mitosis (Twig *et al.*, 2008). Overexpression of wild-type Drp1 does not lead to mitochondrial fragmentation, suggesting that a simple alteration of Drp1 levels does not change mitochondrial fission but the regulation of Drp1 properties, such as mitochondrial recruitment, higher-order assembly, or GTPase activity is rather critical. In this context, posttranslational modifications are implicated as regulatory mechanisms during mitochondrial fission (Figure 4).

Phosphorylation

During mitosis, rat Drp1 is activated by Cdk1/cyclin B-mediated phosphorylation of Ser585 (Ser616 for human Drp1) in the GED. This mitotic phosphorylation promotes Drp1-dependent mitochondrial fission and facilitates proper distribution and segregation of mitochondria into the daughter cells (Taguchi *et al.*, 2007). The exact mechanisms linking Drp1 phosphorylation at Ser585 to increased fission activity remain to be determined. A serine residue within the GED (Ser637 for human Drp1; Ser656 for rat Drp1) is phosphorylated by protein kinase A in HeLa and PC12 cells. This phosphorylation inhibits mitochondrial fission through inhibition of the intramolecular interaction between the GTPase domain and the GED, GTPase activity, and eventually mitochondrial

recruitment of Drp1 (Chang and Blackstone, 2007). In this context, calcineurin dephosphorylates Ser637 and stimulates Drp1 translocation to the mitochondria (Cribbs and Strack, 2007; Cereghetti *et al.*, 2008). For example, poly-glutamine expansions in Huntingtin protein, the cause of Huntington's disease, superactivates calcineurin through enhanced calcium levels, and increases mitochondrial recruitment of Drp1 (Cereghetti *et al.*, 2008). Furthermore, mutant Huntingtin protein directly binds Drp1 and increases its GTPase activity, leading to mitochondrial fragmentation and defects in anterograde and retrograde mitochondrial transport (Song *et al.*, 2011). In contrast, Ca^{2+} signaling activates Ca^{2+} /calmodulin-dependent protein kinase $\text{I}\alpha$ to phosphorylate Ser637 and increases mitochondrial translocation of Drp1 by increasing the Drp1 binding affinity for hFis1 in hippocampal neurons (Han *et al.*, 2008). Thus, phosphorylation of Drp1 at the same GED residues is likely to have opposite effects on the mitochondrial fission activity in different cells or tissues. Ca^{2+} /calmodulin-dependent protein kinase 1a also mediates phosphorylation of Ser600 in human Drp1 isoform 3 (corresponding to Ser637 in human Drp1 isoform 1), which induces mitochondrial fragmentation during voltage-dependent Ca^{2+} signaling in cultured hippocampal neurons (Han *et al.*, 2008). The mechanistic background of these opposite effects remains to be investigated.

O-GlcNacylation

Gawlowski *et al.* (2012) recently demonstrated O-GlcNacylation at Thr-585 and Thr-586 in the insert B-domain (VD) of Drp1 in rat neonatal cardiac myocytes; the reaction was significantly augmented by the inhibition of N-acetyl-glucosaminidase, leading to elevated levels of GTP-bound active Drp1, its mitochondrial translocation, and the induction of mitochondrial fragmentation. The reaction was also stimulated upon shifting cells from low to high glucose medium, concomitant with an increase in the Drp1 expression level. Interestingly, this reaction paralleled a decrease in Ser637-phosphorylated Drp1, resulting in the mitochondrial recruitment of Drp1. The reaction was detected for cardiac Drp1 in a mouse model of type 2 diabetes, and therefore, might be linked to the development of diabetes-induced mitochondrial dysfunction and cardiovascular complications.

S-Nitrosylation

S-Nitrosylation is a ubiquitous protein modification in redox-based signaling. β -Amyloid protein, a key mediator of Alzheimer's disease, stimulates nitric oxide production to cause S-nitrosylation of human Drp1 at Cys644 within the GED, which enhances GTPase activity and Drp1 oligomer formation in association with excessive mitochondrial fission in neurons, leading to synaptic loss and neuronal damage in the brains of patients with Alzheimer's disease (Cho *et al.*, 2009), although this model has been challenged (Bossy *et al.*, 2010). A mutation of Cys644 prevents mitochondrial fragmentation and blocks the neurotoxicity induced by nitric oxide or β -amyloid protein (Cho *et al.*, 2009).

SUMOylation

The small ubiquitin-like modifier (SUMO) protein also affects Drp1 activity. Overexpression of SUMO1 stabilizes Drp1 in a

Bax/Bak-dependent manner on the mitochondrial membrane and induces mitochondrial fission, suggesting that sumoylation is a step in Drp1 regulation during early apoptosis progression (Wasiak *et al.*, 2007). Mitochondrial SUMO E3 ligase (MAPL) has been identified as SUMO E3 ligase for Drp1 (Braschi *et al.*, 2009). Conversely, overexpression of the Sumo protease SENP5 decreases Drp1 sumoylation and rescues SUMO1-induced mitochondrial division (Zunino *et al.*, 2007). Neuspiel *et al.* (2008) reported that MAPL is incorporated in previously uncharacterized mitochondria-derived vesicles (MDVs) that bud from mitochondria and are transported to peroxisomes. Communication with peroxisomal membranes might thus influence mitochondrial morphology or lipid biosynthesis (Braschi *et al.*, 2010).

Ubiquitination

In addition to SUMOylation, ubiquitination regulates Drp1 activity. A mitochondria-associated RING-finger E3 ubiquitin ligase, MARCH5 (also known as MITOL) ubiquitinates Drp1 on the MOM, although the effect of MARCH5/MITOL-dependent ubiquitination of Drp1 on mitochondrial dynamics remains controversial. MARCH5/MITOL silencing or overexpression of the MARCH5/MITOL mutant lacking ubiquitin ligase activity induces mitochondrial fragmentation (Nakamura *et al.*, 2006; Yonashiro *et al.*, 2006). Karbowski *et al.* (2007), however, later demonstrated that MARCH5/MITOL silencing, as well as overexpression of the RING-inactive MARCH5/MITOL mutant, induces abnormal mitochondrial accumulation of Drp1 in association with abnormal mitochondrial elongation and their interconnections. In addition, MARCH5/MITOL might have a more general role in the quality control of mitochondria by ubiquitination and removal of mutated, damaged, or misfolded protein accumulated in the MOM, as observed for a mutated version of superoxide dismutase or expanded poly-Q proteins (Sugiura *et al.*, 2011). Of note, as described already in part II-2, MARCH5/MITOL is also involved in the regulation of mitochondria-ER tethering by ubiquitinating the mitochondria-bound form of Mfn2. Thus, whether and how MARCH5/MITOL contributes to mitochondrial dynamics is an important issue for future studies. Different effects induced by the same modification might depend on circumstances such as where and when the effects occur within the cells or the cell type. Although the exact mechanism of Drp1 regulation via the above-mentioned modifications is unclear, it is likely that posttranslational modifications of Drp1 both positively and negatively regulate its function in mitochondrial dynamics.

Interaction with cardiolipin

Cardiolipin is the principal phospholipid of the mitochondrial membrane and is implicated in mitochondrial dynamics, affecting the MOM and MIM structures during various processes. In fact, coassembly of s- and l-Mgm1 proceeds in *trans* in a cardiolipin-dependent manner (DeVay *et al.*, 2009). In mammalian Opa1, a low basal GTPase activity is dramatically enhanced by an association with negatively charged phospholipids such as cardiolipin, which triggers the assembly of Opa1 into higher-order oligomers (Ban *et al.*, 2010). In a similar vein, *in vitro* analysis of Drp1 function in apoptosis revealed that Drp1 stimulates tBid-induced Bax oligomeri-

zation in a cardiolipin-dependent manner; Drp1 catalyzes the formation of a hemi-fusion/fission intermediates in a cardiolipin-dependent manner (Montessuit *et al.*, 2010). *In vitro*, cardiolipin potently binds to Drp1 and activates GTPase, leading to membrane tubulation to a diameter (100–150 nm) similar to that at the mitochondrial constriction site (Macdonald *et al.*, 2014). Considering that the cardiolipin is enriched at the MOM-MIM contact sites (Tatsuta *et al.*, 2014), the cardiolipin-enriched microdomains might constitute the sites of Drp1 recruitment. If this is the case, how cardiolipin shares receptor function with Mff and Mid49/Mid51 *in vivo* remains an important question.

Physiologic Importance of Mitochondrial Dynamics

Mfn1-, Mfn2-, and Opa1-Dependent Mitochondrial Fusion

Chan and colleagues generated Mfn1- and Mfn2-KO mice (Chen *et al.*, 2003). In contrast to the full viability of heterozygous animals, homozygous mutants are embryonic lethal. Mfn1-KO mice were embryonic lethal by embryonic day 12.5 (86% mutant embryos are resorbed). Similarly, 87% of Mfn2-KO embryos were resorbed by embryonic day 11.5. Mfn1 and Mfn2 double-KO embryos died earlier than either single mutants and showed greater developmental delay, indicating that Mfn1 and Mfn2 are both required and have a cooperative relationship in developmental events early in embryogenesis. The placental trophoblast giant cell layer was specifically and severely disrupted in Mfn2-KO mice, whereas they were normal in Mfn1-KO mice, suggesting a functional *in vivo* difference between Mfn1 and Mfn2. Mfn1-KO and Mfn2-KO mutant culture cells contained mitochondria with a depolarized membrane potential, indicating that while bulk cultures display respiratory activity, the function of individual mitochondria is compromised in the absence of either Mfn1 or Mfn2.

As discussed above, Mfn2 has critical roles not only in mitochondrial fusion, but also in the establishment of mitochondria-ER interactions (de Brito and Scorrano, 2008). In this context, Zorzano and colleagues, using liver-specific and skeletal muscle-specific Mfn2-KO mice, demonstrated that Mfn2 coordinates mitochondria and ER function to modulate metabolism; Mfn2 deficiency causes ER stress, mitochondrial dysfunction, and an increase in hydrogen peroxide, leading to altered insulin signaling and glucose intolerance in cells, which were ameliorated by chemical chaperones and the antioxidant N-acetylcysteine (Sebastian *et al.*, 2012). Similarly, Schneeberger *et al.* (2013) reported that mitochondrial network complexity and their association with the ER in anorexigenic pro-opiomelanocortin (POMC) neurons of the hypothalamus, major targets of leptin action, are decreased in mice with diet-induced obesity. Depletion of Mfn2, but not Mfn1, in POMC neurons results in various cellular defects, such as reduced POMC processing to α -melanocyte-stimulating hormone, reduced mitochondria-ER contacts, reduced mitochondrial respiration, enhanced ROS production, and development of leptin resistance, eventually causing severe obesity. Further, inhibition of ER stress in the hypothalamus by chemical chaperone treatments reverses these phenotypes

of POMC-Mfn2-KO mice. Thus, the mitochondria-ER contact regulated by Mfn2 in POMC neurons plays a crucial function in mediating ER-induced leptin resistance, and is an essential regulator of the whole-body energy balance.

On the other hand, [Dietrich et al. \(2013\)](#) analyzed mitochondrial dynamics in orexigenic agouti-related protein (Agrp) neurons during various metabolic states and assessed the role of Mfn1 and Mfn2 in these processes using Agrp Mfn1-KO and Agrp Mfn2-KO mice. They demonstrated that mitofusin deficiency impairs the action potential frequency of Agrp neurons in response to a high-fat diet and the KO mice gained less weight when fed a high-fat diet. Of note, in contrast to the case of anorexigenic POMC neurons, the depletion of Mfn proteins did not lead to a decrease in ER-mitochondria connections or an increase in ER stress, indicating divergent roles for Mfn proteins in a cell-type-specific manner ([Dietrich et al., 2013](#)). Together these experiments revealed critical roles of Mfn1 and Mfn2 in the regulation of the metabolic response of Agrp neurons. In addition to the role of mitochondrial fusion in tissue homeostasis, its role in the regulation of cellular differentiation has been clarified. Scorrano and colleagues demonstrated that heart-specific Mfn1-KO or Mfn2-KO mice, and gene-trapping of Mfn2 or Opa1 in mouse embryonic stem cells arrested mouse heart development and impaired differentiation of embryonic stem cells into cardiomyocytes, revealing the presence of a mitochondrial fusion-sensitive Ca^{2+} /calcineurin pathway that directs cardiomyocyte differentiation through Notch signaling ([Kasahara et al., 2013](#)). As discussed above, OMA1 mediates stress-induced processing of L-Opa1 at the S1 site. [Quiros et al. \(2012\)](#) generated OMA1-KO mice and analyzed their phenotype. The mice develop normally, with males and females being fertile, and their survival rates are indistinguishable from those of the wild-type littermates. At the cellular level, as confirmed in previous reports, stress-induced conversion of L-Opa1s to S-forms is completely blocked in OMA1-KO cells, and correspondingly, the cells have highly connected and less-fragmented mitochondria that are resistant to dissipation of the mitochondrial membrane potential by CCCP treatment ([Quiros et al., 2012](#)). Furthermore, mitochondria in OMA1-KO cells have a tight structural organization of cristae junctions and are resistant to proapoptotic reagents. Of note, the mutant cell mitochondria have decreased mtDNA levels, suggesting that Opa1 functions to maintain mtDNA. Interestingly, OMA1 ablation leads to a marked upregulation of genes related to lipid metabolism and metabolic pathways related to the obesity phenotype, and mutant mice exhibit a progressive increase in body weight compared with control mice; these effects are enhanced when the mice are maintained on a high-fat diet. These experiments revealed that the precise balance between L-Opa1 and S-OMA1 regulated by OMA1 is essential for metabolic homeostasis ([Quiros et al., 2012](#)).

Drp1-Dependent Mitochondrial Fission

To elucidate the detailed physiologic roles of mitochondrial fission *in vivo*, we and another group generated tissue-specific Drp1 KO mice using the Cre and loxP system ([Ishihara et al., 2009](#); [Wakabayashi et al., 2009](#)). Although Drp1 is dispensable

for murine embryonic fibroblast (MEF) viability, Drp1 KO mice die at around embryonic day 12.5 due to developmental abnormalities, particularly in the forebrain. In addition, a missense mutation in mouse Drp1 in the middle domain that is essential for intramolecular interactions (Python mice; C452F mutation) leads to cardiomyopathy ([Ashrafian et al., 2010](#)). The main developmental abnormalities observed in Drp1 KO mice are defects in the forebrain and synapse development, poorly developed livers, and compromised cardiac formation or function ([Ishihara et al., 2009](#); [Wakabayashi et al., 2009](#)), while heterozygous Python mice exhibit depletion of cardiac ATP and cardiomyopathy ([Ashrafian et al., 2010](#)). To date, only a dominant-negative middle domain mutation (A395D) in Drp1 has been reported in a lethal disorder with microcephaly, abnormal brain development, optic atrophy, hypoplasia, persistent lactic acidemia, and a mildly elevated plasma concentration of very-long-chain fatty acids ([Waterham et al., 2007](#)). Despite its severe clinical phenotype, respiratory-chain enzyme activities are normal in muscle and in cultured fibroblasts derived from the patient. Neurons are particularly vulnerable to mitochondrial dysfunction. Neuron-specific Drp1 KO mice are viable at birth, but quickly die due to neurodegeneration ([Ishihara et al., 2009](#); [Wakabayashi et al., 2009](#)). In primary cultured neural cells from Drp1 KO embryos, enlarged and aggregated mitochondria are sparsely distributed in the neurites, and the synaptic structures are lost ([Ishihara et al., 2009](#); [Wakabayashi et al., 2009](#)). These findings suggest that Drp1 deficiency causes an abnormal distribution of enlarged mitochondria in extremely polarized cells such as neurites; these spatiotemporal defects may inhibit the ATP supply and Ca^{2+} signaling, eventually preventing synapse formation. Taken together, these results suggest that Drp1 deficiency results in unusually shaped, large mitochondria with compromised intracellular movement, which leads to neuronal cell death. Mitochondrial fission also functions as a quality control system to suppress age-dependent oxidative damage and thus promote neuronal survival ([Twig et al., 2008](#); [Chan, 2012](#); [Archer, 2013](#)). Determining the physiologic relevance of Drp1 in other tissues that might underlie various human diseases remains to be addressed.

Mitochondrial fission is a physiologic response in brown adipose tissue. Brown adipose tissue is an important site for energy expenditure, and cold exposure and adrenergic stimulation induce heat generation preceded by rapid mitochondrial fragmentation. Shrihail and colleagues recently demonstrate that cold exposure or norepinephrine-stimulation activates Drp1 by PKA-dependent Ser600 phosphorylation to induce mitochondrial fragmentation as an essential step of heat generation ([Wikstrom et al., 2014](#)). This fragmentation enhances uncoupling of oxidative phosphorylation in mitochondria by free fatty acids that are generated through different pathways, leading to heat production in brown adipose tissue.

Perspective

Recent advances in live cell imaging revealed that mitochondria are highly dynamic through continuous fusion and

fission, as well as movement along the cellular cytoskeleton. Mitochondrial fusion and fission have important roles not only in the modulation of mitochondrial morphology, but also in other biologic processes, including bioenergetics, cellular metabolism, mitochondrial maintenance, synaptic integrity, and neuronal cell death. Over the past decade, despite the accumulation of a significant amount of data regarding the identification of proteins modulating mitochondrial dynamics and their molecular function, characterization of the coordination and regulation of these proteins remains in a preliminary stage. Several fundamental questions, such as the exact role of the mitochondrial shape proteins and the coordination of fusion and fission, are unsolved. A number of studies demonstrated that mitochondrial morphology and its physiologic functions differ depending on the cell-type or tissue. These variations and the cell-type specificity of mitochondrial dynamics might be related to specific cellular functions, such as in neurons. Therefore, learning how and the degree to which mitochondrial dynamics influence cell-specific functions in various tissues continues to be a challenging task. Future research to investigate the complex crosstalk between mitochondrial dynamics and their physiologic function are likely to provide exciting breakthroughs in the fields of cell biology and for clinical medicine.

Acknowledgments

The work of the authors was supported by grants from the Ministry of Education, Science, and Culture of Japan.

See also: Molecular Principles, Components, Technology, and Concepts: Membranes: Membrane Potential: Concepts

References

- Alexander, C., Votruba, M., Pesch, U.E., *et al.*, 2000. OPA1, encoding a dynamin-related GTPase, is mutated in autosomal dominant optic atrophy linked to chromosome 3q28. *Nature Genetics* 26, 211–215.
- Anand, R., Wai, T., Baker, M., *et al.*, 2014. The i-AAA protease YME1L and OMA1 cleave OPA1 to balance mitochondrial fusion and fission. *Journal of Cell Biology* 204, 919–929.
- An, H.-J., Cho, G., Lee, J.-O., *et al.*, 2013. Higd-1a interacts with Opa1 and is required for the morphological and functional integrity of mitochondria. *Proceedings of the National Academy of Sciences of the United States of America* 110, 13014–13019.
- Anderson, J.F., Siller, E., Barral, J.M., 2010. The saccin repeating region (SRR): A novel Hsp90-related supra-domain associated with neurodegeneration. *Journal of Molecular Biology* 400, 665–674.
- Archer, S.L., 2013. Mitochondrial dynamics — Mitochondrial fission and fusion in human disease. *New England Journal of Medicine* 369, 2236–2251.
- Ashrafian, H., Docherty, L., Leo, L.V., *et al.*, 2010. A mutation in the mitochondrial fission gene *Dnm1l* leads to cardiomyopathy. *PLoS Genetics* 6, e1001000.
- Baker, M.J., Lampe, P.A., Stojanovski, D., *et al.*, 2014. Stress-induced OMA1 activation and autocatalytic turnover regulate OPA1-dependent mitochondrial dynamics. *EMBO Journal* 33, 578–593.
- Ban, T., Heymann, J.A., Song, Z., Hinshaw, J.E., Chan, D.C., 2010. OPA1 disease alleles causing dominant optic atrophy have defects in cardiolipin-stimulated GTP hydrolysis and membrane tubulation. *Human Molecular Genetics* 19, 2113–2122.
- Barker, M.J., Lampe, P.A., Stojanovski, D., *et al.*, 2014. Stress-induced OMA1 activation and autocatalytic turnover regulate OPA1-dependent mitochondrial dynamics. *EMBO Journal* 33, 578–593.
- Bian, X., Klemm, R.W., Liu, T.Y., *et al.*, 2011. Structures of the atlastin GTPase provide insight into homotypic fusion of endoplasmic reticulum membranes. *Proceedings of the National Academy of Sciences of the United States of America* 108, 3976–3981.
- Bossey, B., Pettrilli, A., Klinglmayr, E., *et al.*, 2010. S-nitrosylation of Drp1 does not affect enzymatic activity and is not specific to Alzheimer's disease. *Journal of Alzheimer's Disease* 20, S513–S526.
- Braschi, E., Goyon, V., Zunino, R., *et al.*, 2010. Vps35 mediates vesicle transport between the mitochondria and peroxisomes. *Current Biology* 20, 1310–1315.
- Braschi, E., Zunino, R., McBride, H.M., 2009. MAPL is a new mitochondrial SUMO E3 ligase that regulates mitochondrial fission. *EMBO RepReports* 10, 748–754.
- de Brito, O.M., Scorrano, L., 2008. Mitofusin 2 tethers endoplasmic reticulum to mitochondria. *Nature* 456, 605–610.
- Cassereau, J., Chevrollier, A., Gueguen, N., *et al.*, 2011. Mitochondrial dysfunction and pathophysiology of Charcot-Marie-Tooth disease involving GDAPI mutations. *Experimental Neurology* 227, 31–41.
- Cereghetti, G.M., Stangherlin, A., de Brito, M., *et al.*, 2008. Dephosphorylation by calcineurin regulates translocation of Drp1 to mitochondria. *Proceedings of the National Academy of Sciences of the United States of America* 105, 15803–15808.
- Cerqua, C., Anesti, V., Pyakurel, A., *et al.*, 2010. Trichoplein/mitostatin regulates endoplasmic reticulum-mitochondria juxtaposition. *EMBO Reports* 11, 854–860.
- Chan, D.C., 2012. Fusion and fission: Interlinked processes critical for mitochondrial health. *Annual Review of Genetics* 46, 265–287.
- Chang, C.R., Blackstone, C., 2007. Cyclic AMP-dependent protein kinase phosphorylation of Drp1 regulates its GTPase activity and mitochondrial morphology. *Journal of Biological Chemistry* 282, 21583–21587.
- Chen, H., Detmer, S.A., Ewald, A.J., *et al.*, 2003. Mitofusins Mfn1 and Mfn2 coordinately regulate mitochondrial fusion and are essential for embryonic development. *Journal of Cell Biology* 160, 189–200.
- Cho, D.H., Nakamura, T., Fang, J., *et al.*, 2009. S-Nitrosylation of Drp1 mediates β -amyloid-related mitochondrial fission and neuronal injury. *Science* 324, 102–105.
- Choi, S.-Y., Huang, P., Jenkins, G.M., *et al.*, 2006. A common lipid links Mfn-mediated mitochondrial fusion and SNARE-regulated exocytosis. *Nature Cell Biology* 8, 1255–1262.
- Chung, Y.M., Kim, J.S., Yoo, Y.D., 2006. A novel protein Romo1, induces ROS production in the mitochondria. *Biochemical and Biophysical Research Communications* 347, 649–655.
- Cosson, P., Marchetti, A., Ravazzola, M., Orci, L., 2012. Mitofusin-2 independent juxtaposition of endoplasmic reticulum and mitochondria: An ultrastructural study. *PLoS One* 7, e46293.
- Cribbs, J.T., Strack, S., 2007. Reversible phosphorylation of Drp1 by cyclic AMP-dependent protein kinase and calcineurin regulates mitochondrial fission and cell death. *EMBO Reports* 8, 939–944.
- Davis, V.J., Hollins, A.J., Piechota, M.J., *et al.*, 2007. Opa1 deficiency in a mouse model of autosomal dominant optic atrophy impairs mitochondrial morphology, optic nerve structure and visual function. *Human Molecular Genetics* 16, 1307–1318.
- DeVay, R.M., Dominguez-Ramirez, L., Lackner, L.L., *et al.*, 2009. Coassembly of Mgm1 isoforms requires cardiolipin and mediates mitochondrial inner membrane fusion. *Journal of Cell Biology* 186, 793–803.
- Delettre, C., Lenaers, G., Griffoin, J.M., *et al.*, 2000. Nuclear gene OPA1, encoding a mitochondrial dynamin-related protein, is mutated in dominant optic atrophy. *Nature Genetics* 26, 207–210.
- Detmer, S.A., Chan, D.C., 2007. Functions and dysfunctions of mitochondrial dynamics. *Nature Reviews Molecular Cell Biology* 8, 870–879.
- Dietrich, M.O., Liu, Z.-W., Horvath, T.L., 2013. Mitochondrial dynamics controlled by mitofusins regulate Agpr neuronal activity and diet-induced obesity. *Cell* 155, 188–199.
- Ehse, S., Raschke, I., Mancuso, G., *et al.*, 2009. Regulation of OPA1 processing and mitochondrial fusion by m-AAA protease isoenzymes and OMA1. *Journal of Cell Biology* 187, 1023–1036.
- Eura, Y., Ishihara, N., Oka, T., Mihara, K., 2006. Identification of a novel protein that regulates mitochondrial fusion by modulating mitofusin (Mfn) protein function. *Journal of Cell Science* 119, 4913–4925.
- Frank, S., Gaume, B., Bergmann-Leitner, W.W., *et al.*, 2001. The role of dynamin-related protein 1, a mediator of mitochondrial fission, in apoptosis. *Developmental Cell* 4, 515–525.

- Friedman, J.R., Lackner, L.L., West, M., *et al.*, 2011. ET tubules mark sites of mitochondrial division. *Science* 334, 358–362.
- Friedman, J.R., Nunnari, J., 2014. Mitochondrial form and function. *Nature* 505, 335–343.
- Gandre-Babbe, S., van der Bliek, A.M., 2008. The novel tail-anchored membrane protein Mff controls mitochondrial and peroxisomal fission in mammalian cells. *Molecular Biology of the Cell* 19, 2402–2412.
- Gawlowski, T., Suarez, J., Scott, B., *et al.*, 2012. Modulation of dynamin-related protein 1 (DRP1) function by increased O-linked-b-N- acetylglucosamine modification (O-GlcNAc) in cardiac myocytes. *Journal of Biological Chemistry* 287, 30024–30034.
- Giard, M., Larivière, R., Parfitt, D.A., *et al.*, 2012. Mitochondrial dysfunction and Purkinje cell loss in autosomal recessive spastic ataxia of Charlevoix-Saguenay (ARSACS). *Proceedings of the National Academy of Sciences of the United States of America* 109, 1661–1666.
- Hamasaki, M., Furuta, N., Matsuda, A., *et al.*, 2013. Autophagosomes form at ER-mitochondria contact sites. *Nature* 495, 389–393.
- Han, X.J., Lu, Y.F., Li, S.A., *et al.*, 2008. CaM kinase I α -induced phosphorylation of Drp1 regulates mitochondrial morphology. *Journal of Cell Biology* 182, 573–585.
- Hoppins, S., Edlich, F., Cleland, M.M., *et al.*, 2011. The soluble form of Bax regulates mitochondrial fusion via Mfn2 homotypic complexes. *Cell* 141, 150–160.
- Huber, N., Guimaraes, S., Schrader, M., Suter, U., Niemann, A., 2013. Charcot-Marie-Tooth disease-associated mutants of GTPase dissociate its roles in peroxisomal and mitochondrial fission. *EMBO Reports* 14, 545–552.
- Ingerman, E., Perkins, E.M., Marino, M., *et al.*, 2005. Dnm1 forms spirals that are structurally tailored to fit mitochondria. *Journal of Cell Biology* 170, 1021–1027.
- Ishihara, N., Eura, Y., Mihara, K., 2004. Mitofusin 1 and 2 play distinct roles in mitochondrial fusion reactions via GTPase activity. *Journal of Cell Science* 117, 6535–6546.
- Ishihara, N., Fujita, Y., Oka, T., Mihara, K., 2006. Regulation of mitochondrial morphology through proteolytic cleavage of OPA1. *EMBO Journal* 25, 2966–2977.
- Ishihara, N., Nomura, M., Jofuku, A., *et al.*, 2009. Mitochondrial fission factor Drp1 is essential for embryonic development and synapse formation in mice. *Nature Cell Biology* 11, 958–966.
- Karbowski, M., Jeong, S.-Y., Youle, R., 2004. Endophilin B1 is required for the maintenance of mitochondrial morphology. *Journal of Cell Biology* 166, 1027–1039.
- Karbowski, M., Neutznier, A., Youle, R.J., 2007. The mitochondrial E3 ubiquitin ligase MARCH5 is required for Drp1 dependent mitochondrial division. *Journal of Cell Biology* 178, 71–84.
- Kasahara, A., Cipolat, S., Chen, Y., Dorn II, G.W., Scorrano, L., 2013. Mitochondrial fusion directs cardiomyocyte differentiation via calcineurin and notch signaling. *Science* 342, 734–737.
- Kaesler, M., Kambacheld, M., Kisters-Woike, B., Langer, T., 2003. Oma1, a novel membrane-bound metalloprotease with activities overlapping with the m-AAA protease. *Journal of Biological Chemistry* 278, 46414–46423.
- Kornmann, B., Currie, E., Collins, S.R., *et al.*, 2009. An ER-mitochondria tethering complex revealed by a synthetic biology screen. *Science* 325, 477–481.
- Koirala, S., Guo, Q., Kalia, R., *et al.*, 2013. Interchangeable adapters regulate mitochondrial dynamin assembly for membrane scission. *Proceedings of the National Academy of Sciences of the United States of America* 110, 1342–1351.
- Korobova, F., Ramabhadran, V., Higgs, H.N., 2013. An actin-dependent step in mitochondrial fission mediated by the ER-associated forming INF2. *Science* 339, 464–467.
- Korobova, F., Gauvin, T.J., Higgs, H.N., 2014. A role for myosin II in mammalian mitochondrial fission. *Current Biology* 24, 1–6.
- Lackner, L.L., Horner, J.S., Nunnari, J., 2009. Mechanistic analysis of a dynamin effector. *Science* 325, 874–877.
- Legesse-Miller, A., Massol, R.H., Kirchhausen, T., 2003. Constriction and Dnm1p recruitment are distinct processes in mitochondrial fission. *Molecular Biology of the Cell* 14, 1953–1963.
- Loson, O.C., Liu, R., Rome, M.E., *et al.*, 2014. The mitochondrial fission receptor Mid51 requires ADP as a cofactor. *Structure* 22, 1–11.
- Macdonald, P.J., Stepanyants, N., Mehrotra, N., *et al.*, 2014. A dimeric equilibrium intermediate nucleates Drp1 reassembly on mitochondrial membranes for fission. *Molecular Biology of the Cell* 25, 1905–1915.
- Malka, F., Guillery, O., Cifuentes-Diaz, C., *et al.*, 2005. Separate fusion of outer and inner mitochondrial membranes. *EMBO Reports* 6, 853–859.
- Merkwirth, C., Dargazanli, S., Tatsuta, T., *et al.*, 2008. Prohibitins control cell proliferation and apoptosis by regulating OPA1-dependent cristae morphogenesis in mitochondria. *Genes & Development* 22, 476–488.
- Mishra, P., Carelli, V., Manfredi, G., Chan, D.C., 2014. Proteolytic cleavage of Opa1 stimulates mitochondrial inner membrane fusion and couples fusion to oxidative phosphorylation. *Cell Metabolism* 19, 630–641.
- Montessuit, S., Somasekharan, S.P., Terrones, O., *et al.*, 2010. Membrane remodeling induced by the dynamin-related protein Drp1 stimulates Bax oligomerization. *Cell* 142, 889–901.
- Mozdy, A.D., McCaffery, J.M., Shaw, J.M., 2000. Dnm1 GTPase-mediated mitochondrial fission is a multi-step process requiring the novel integral membrane component Fis1p. *Journal of Cell Biology* 151, 367–380.
- Munoz, J.P., Ivanova, S., Sanchez-Wandelmer, J., *et al.*, 2013. Mfn2 modulates the UPR and mitochondrial function via repression of PERK. *EMBO Journal* 32, 2348–2361.
- Na, A.R., Chung, Y.M., Lee, S.B., *et al.*, 2008. A critical role for Romo1-derived ROS in cell proliferation. *Biochemical and Biophysical Research Communications* 369, 672–678.
- Nakamura, N., Kimura, Y., Tokuda, M., Honda, S., Hirose, S., 2006. MARCH-V is a novel mitofusin 2- and Drp1-binding protein able to change mitochondrial morphology. *EMBO Reports* 7, 1019–1022.
- Naon, D., Scorrano, L., 2014. At the right distance: ER-mitochondria juxtaposition in cell life and death. *Biochimica et Biophysica Acta – Molecular Cell Research* 1843, 2184–2194.
- Neuspiel, M., Schauss, A.C., Braschi, E., *et al.*, 2008. Cargo-selected transport from the mitochondria to peroxisomes is mediated by vesicular carriers. *Current Biology* 18, 102–108.
- Niemann, A., Ruegg, M., Padula, V.L., Schenone, A., Suter, U., 2005. Ganglioside-induced differentiation associated protein 1 is a regulator of the mitochondrial network: New implications for Charcot-Marie-Tooth disease. *Journal of Cell Biology* 170, 1067–1107.
- Norton, M., Ng, A.C.-H., Baird, S., *et al.*, 2014. ROMO1 is an essential redox-dependent regulator of mitochondrial dynamics. *Science Signaling* 7, 1–9.
- Noack, R., Frede, S., Albrecht, P., *et al.*, 2012. Charcot-Marie-Tooth disease CMT4A: GTPase increases cellular glutathione and the mitochondrial membrane potential. *Human Molecular Genetics* 21, 150–162.
- Olichon, A., Baricault, L., Gas, N., *et al.*, 2003. Loss of OPA1 perturbs the mitochondrial inner membrane structure and integrity, leading to cytochrome c release and apoptosis. *Journal of Biological Chemistry* 278, 7743–7746.
- Onoue, K., Jofuku, A., Ban-Ishihara, R., *et al.*, 2013. Fis1 acts as a mitochondrial recruitment factor for TBC1D15 that is involved in regulation of mitochondrial morphology. *Journal of Cell Science* 126, 176–186.
- Otera, H., Wang, C., Cleland, M.M., *et al.*, 2010. Mff is an essential factor for mitochondrial recruitment of Drp1 during mitochondrial fission in mammalian cells. *Journal of Cell Biology* 191, 1141–1158.
- Palmer, C.S., Elgass, K.D., Parton, R.G., *et al.*, 2013. Mid49 and Mid51 can act independently of Mff and Fis1 in Drp1 recruitment and are specific for mitochondrial fission. *Journal of Biological Chemistry* 288, 27584–27593.
- Palmer, C.S., Osellame, L.D., Laine, D., *et al.*, 2011. Mid49 and Mid51, new components of the mitochondrial fission machinery. *EMBO Reports* 12, 565–573.
- Quiros, P.M., Ramsay, A.J., Sala, D., *et al.*, 2012. Loss of mitochondrial protease OMA1 alters processing of the GTPase OPA1 and causes obesity and defective thermogenesis in mice. *EMBO Journal* 31, 2117–2133.
- Ramachandran, R., Surka, M., Chappie, J.S., *et al.*, 2007. The dynamin middle domain is critical for tetramerization and higher-order self-assembly. *EMBO Journal* 26, 559–566.
- Richter, V., Palmer, C., Osellame, L., *et al.*, 2014. Structural and functional analysis of Mid51, a dynamin receptor required for mitochondrial fission. *Journal of Cell Biology* 204, 477–486.
- Rostovtseva, T.K., Boukari, H., Antignani, A., *et al.*, 2009. Bax activates Endophilin B1 oligomerization and lipid membrane vesiculation. *Journal of Biological Chemistry* 284, 34390–34399.
- Ruijter, J., Meglei, G., Rubinstein, J.L., McQuibban, G.A., 2009. Phospholipid association is essential for the dynamin-related protein Mgm1 to function in mitochondrial membrane fusion. *Journal of Biological Chemistry* 284, 28682–28686.
- Santel, A., Fuller, M.T., 2001. Control of mitochondrial morphology by a human mitofusin. *Journal of Cell Science* 114, 867–874.
- Schauss, A.C., Huang, H., Choi, S.-Y., *et al.*, 2010. A novel cell-free mitochondrial fusion assay amenable for high-throughput screening of fusion modulators. *BMC Biology* 8, 100–1012.

- Schneeberger, M., Dietrich, M.O., Sebastian, D., *et al.*, 2013. Mitofusin 2 in POMC neurons connects ER stress with Leptin resistance and energy imbalance. *Cell* 155, 172–187.
- Sebastian, D., Hernandez-Alvarez, M.I., Segales, J., *et al.*, 2012. Mitofusin 2 (Mfn2) links mitochondrial and endoplasmic reticulum function with insulin signaling and is essential for normal glucose homeostasis. *Proceedings of the National Academy of Sciences of the United States of America* 109, 5523–5528.
- Smirnova, E., Griparic, L., Shurland, D.-L., van der Bliek, A.M., 2001. Dynamin-related protein Drp1 is required for mitochondrial division in mammalian cells. *Molecular Biology of the Cell* 12, 2245–2256.
- Song, W., Chen, J., Petrilli, A., *et al.*, 2011. Mutant huntingtin binds the mitochondrial fission GTPase dynamin-related protein-1 and increases its enzymatic activity. *Nature Medicine* 17, 377–383.
- Song, Z., Chen, H., Fiket, M., Alexander, C., Chan, D.C., 2007. A1 processing controls mitochondrial fusion and is regulated by mRNA splicing, membrane potential, and Yme1L. *Journal of Cell Biology* 178, 749–755.
- Stavru, F., Palmer, A.E., Wang, C., Youle, R.J., Cossart, P., 2013. Atypical mitochondrial fission upon bacterial infection. *Proceedings of the National Academy of Sciences of the United States of America* 110, 16003–16008.
- Sugiura, A., Nagashima, S., Tokuyama, T., *et al.*, 2013. MITOL regulates endoplasmic reticulum-mitochondria contacts via mitofusin 2. *Molecular Cell* 51, 20–34.
- Sugiura, A., Yonashiro, R., Fukuda, T., *et al.*, 2011. A mitochondrial ubiquitin ligase MITOL controls cell toxicity of polyglutamine-expanded protein. *Mitochondrion* 11, 139–146.
- Takahashi, Y., Meyerkord, C.L., Wang, H.-G., 2009. Bif-1/Endophilin B1: A candidate for crescent driving force in autophagy. *Cell Death and Differentiation* 16, 947–955.
- Tatsuta, T., Scharwey, M., Langer, T., 2014. Mitochondrial lipid trafficking. *Trends in Cell Biology* 24, 44–52.
- Taguchi, N., Ishihara, N., Jofuku, A., Oka, T., Mihara, K., 2007. Mitotic phosphorylation of dynamin-related GTPase Drp1 participates in mitochondrial fission. *Journal of Biological Chemistry* 282, 11521–11529.
- Tondera, D., Czauderna, F., Paulick, K., *et al.*, 2005. The mitochondrial protein MTP18 contributes to mitochondrial fission in mammalian cells. *Journal of Cell Science* 118, 3049–3059.
- Twig, A., Elorga, A., Molina, A.J.A., *et al.*, 2008. Fission and selective fusion govern mitochondrial segregation and elimination by autophagy. *EMBO Journal* 27, 433–446.
- Wakabayashi, J., Zhang, Z., Wakabayashi, N., *et al.*, 2009. The dynamin-related GTPase Drp1 is required for embryonic and brain development in mice. *Journal of Cell Biology* 186, 805–816.
- Wasiak, S., Zunino, R., McBride, H.M., 2007. Bax/Bak promote sumoylation of DRP1 and its stable association with mitochondria during apoptotic cell death. *Journal of Cell Biology* 177, 439–450.
- Waterham, H.R., Koster, J., van Roermund, C.W., *et al.*, 2007. A lethal defect of mitochondrial and peroxisomal fission. *New England Journal of Medicine* 356, 1736–1741.
- Wikstrom, J.D., Mahadaviani, K., Liesa, M., *et al.*, 2014. Hormone-induced mitochondrial fission is utilized by brown adipocytes as an amplification pathway for energy expenditure. *EMBO Journal* 33, 418–436.
- Yamano, K., Fogel, A.I., Wang, C., van der Bliek, A.M., Youle, R., 2014. Mitochondrial Rab GAPs govern autophagosome biogenesis during mitophagy. *eLife* 3, 1–24.
- Yonashiro, R., Ishido, S., Kyo, S., *et al.*, 2006. A novel mitochondrial ubiquitin ligase plays a critical role in mitochondrial dynamics. *EMBO Journal* 25, 3618–3626.
- Yoon, Y., Krueger, E.W., Oswald, B.J., McNiven, M.A., 2003. The mitochondrial protein hFis1 regulates mitochondrial fission in mammalian cells through an interaction with the dynamin-like protein DLP1. *Molecular and Cellular Biology* 23, 5409–5420.
- Yoon, Y., Pitts, K.R., McNiven, M.A., 2001. Mammalian dynamin-like protein DLP1 tubulates membranes. *Molecular Biology of the Cell* 12, 2894–2905.
- Zhang, K., Li, H., Song, Z., 2014. Membrane depolarization activates the mitochondrial protease OMA1 by stimulating self-cleavage. *EMBO Reports* 15, 576–585.
- Zick, M., Duvezin-Caubet, S., Schäfer, A., *et al.*, 2009. Distinct roles of the two isoforms of the dynamin-like GTPase Mgm1 in mitochondrial fusion. *FEBS Letters* 583, 2237–2243.
- Zhu, P.P., Patterson, A., Stadler, J., *et al.*, 2004. Intra- and intermolecular domain interactions of the C-terminal GTPase effector domain of the multimeric dynamin-like GTPase Drp1. *Journal of Biological Chemistry* 279, 35967–35974.
- Zhao, J., Kiu, T., Jin, S., *et al.*, 2011. Human MIEF1 recruits Drp1 to mitochondrial outer membranes and promotes mitochondrial fusion rather than fission. *EMBO Journal* 30, 2762–2778.
- Zuchner, S., *et al.*, 2004. Mutations in the mitochondrial GTPase mitofusin 2 cause Charcot-Marie-Tooth neuropathy type 2A. *Nature Genetics* 36, 449–451.
- Zunino, R., Schauss, A., Rippstein, P., Andrade-Navarro, M., McBride, H.M., 2007. The SUMO protease SENP5 is required to maintain mitochondrial morphology and function. *Journal of Cell Science* 120, 1178–1188.

Synaptosomes and Synaptic Vesicles

JC Trinidad, Indiana University, Bloomington, IN, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

This article describes the composition and cell biology of both synaptosomes and synaptic vesicles (SVs). Synaptosomes are biochemically isolated structures that consist of pinched-off nerve terminals, the juxtaposed postsynaptic density, and any associated extracellular matrix proteins that link the two. Synaptosomes are generated during homogenization of brain tissue under iso-osmotic conditions and can be considered as pseudo-organelles because the neck of the axon terminal spontaneously seals up upon homogenization. They were first isolated by Hebb and Whittaker in 1957 (Hebb and Whittaker, 1958), although a clear understanding of their nature was not known until their fine structure was determined using electron microscopy (Gray and Whittaker, 1960), and the term 'synaptosome' itself was not used to describe such structures until 1964 (Whittaker *et al.*, 1964). Because the isolated axon terminal is membrane-bound, synaptosomes are functionally active in the sense that they can respond to depolarizing conditions, which mimic invasion of an action potential into the terminal. In the presence of glucose and oxygen-containing media, they are able to maintain near physiological levels of ATP, which allows for kinase-based signal transduction. There are four criteria that synaptosomes must meet: the axon terminal portion must have resealed, this enclosed region must contain one or more mitochondria, a large fraction of the initial synaptic vesicle pool must be present, and a significant portion of the initial postsynaptic density must still be opposed to the presynaptic active zone. The initial iso-osmotic homogenization may leave the overall postsynaptic spine mostly intact, or substantially removed (Whittaker, 1965).

SVs are small, membrane-enclosed structures, which range from ten to hundreds of copies, that are present in axon terminals of chemical synapses. They are filled with neurotransmitter, and the particular neurotransmitter they contain varies as a function of synaptic type. During neurotransmission, an action potential invades the presynaptic terminal, which causes depolarization of the plasma membrane. This depolarization activates voltage-gated calcium channels that are positioned close to a subset of vesicles in a specialized region of the axon terminal known as the active zone. The resulting calcium influx is sensed by specific proteins on the surface of SVs. Calcium binding by these proteins acts as a trigger for a protein complex that mediates fusion of SVs with the plasma membrane, releasing neurotransmitter into the synaptic cleft. These neurotransmitter molecules diffuse across the cleft and activate specific receptors on the surface of the postsynaptic cell, thereby propagating the neuronal signal. SVs can be biochemically isolated from the brain, and synaptosome preparations are an intermediate step in their enrichment.

Synaptosomes

The synaptosome fraction is generally isolated using density centrifugation, and several approaches have been developed for its purification. The initial brain homogenate is first centrifuged in order to pellet the so-called 'P1' fraction, which consists of nuclei, tissue debris, and large fragments of myelin sheath. The supernatant 'P2' fraction is then subjected to another round of centrifugation using a continuous or discontinuous sucrose gradient, Ficoll or Percoll (Booth and Clark, 1978; Nagy and Delgado-Escueta, 1984). This yields a fraction highly enriched for synaptosomes and depleted of free mitochondria. Synaptosomes possess all those components present in the original axon terminal. These include SVs; large dense core vesicles, mitochondria, synaptic active zones, proteins of the postsynaptic density, and associated signaling machinery, as well as any cytoplasmic proteins present at the synapse.

While the efficiency of synaptosome generation will vary with terminal spine geometry as well as the degree to which shearing forces are applied during homogenization, synaptosomes have been isolated from most neural tissues with similar efficiencies, with the slight exception of large mossy-fiber nerve terminals in the cerebellum that generate synaptosomes at poor yield (Israel and Whittaker, 1965). Therefore, synaptosomes generated from the whole brain will contain a representation of all synaptic types present in the starting material, including both excitatory and inhibitory synapses. In general, multiple synapse types are tightly co-distributed throughout the brain, and therefore microdissection is unlikely to yield highly homogeneous populations of synaptosomes.

The protein composition of synaptosomes and SVs has been best characterized by mass spectrometry. Proteomic characterization of isolated proteins or protein complexes first involves enzymatic digestion of the proteins by specific proteases. Trypsin is the most common protease used for these analyses since it cuts C-terminal to arginine and lysine residues, thereby imparting a positive charge on this end of all resulting peptides (an important characteristic that aids in MS-based sequencing). As with any organelle or subcellular fraction, synaptosomes and SV preparations will contain some level of impurity, however small. As mass spectrometers become more sensitive, these impurities are increasingly detected, even if they are at relatively low levels. Therefore, a simple 'parts list' of any protein complex (no matter how well isolated) needs to be interpreted with a degree of caution, and it is absolutely essential to conduct subsequent cell-biological validation of a given protein's role in the functioning of that protein complex before accepting it as a bona fide component.

The synaptosome preparation is used as a starting point for the isolation of several sub-synaptic fractions. These include: postsynaptic densities; terminal-specific mitochondria

and cytoplasm; presynaptic active zones; and SVs (Whittaker *et al.*, 1964). SVs can be released from synaptosomes by osmotic shock, which lyses the synaptosome membrane while leaving SV membranes largely intact. Subsequent density gradient centrifugation or glass bead column chromatography can be used to isolate a SV pool.

Proteomic and Cell Biological Analysis of Synaptosomes

The first systematic mass spectrometry characterization of synaptosome-derived material was in 2000 by the Kennedy laboratory. These studies used a one dimensional SDS PAGE gel to fractionate PSD (postsynaptic density) proteins, and analysis of individual gel bands led to the identification of 31 proteins (Walikonis *et al.*, 2000). Advances in multi-dimensional fractionation and increasingly sensitive mass spectrometers have resulted in over 2000 components now being identified from isolated PSDs (Bayes *et al.*, 2011; Trinidad *et al.*, 2005). As mentioned previously, biochemically isolated subcellular fractions will never be completely pure, and undoubtedly, a significant percentage of identified components are not truly synaptic. To begin developing a molecular level map of synaptic function, the absolute copy number per synapse of several key proteins has been determined by comparison to isotopically labeled protein standards (Cheng *et al.*, 2006).

Because the axon terminal component of synaptosomes seals up during the initial homogenization, they possess several important biological characteristics. When incubated in media containing glucose and oxygen, they maintain membrane potential, translocate metabolites and ions across their plasma membranes, and can release neurotransmitter when they are depolarized. The release of neurotransmitter upon depolarization is due to activation of voltage-gated calcium channels, which mimics the arrival of an action potential at the axon terminal. Synaptosome membranes can be depolarized by increasing the extracellular K^+ concentration, electrical stimulation, or exposure to drugs that increase the membrane permeability of Na^+ (De Bellerche and Bradford, 1972; Blaustein and Goldring, 1975).

Such an experimental paradigm has been used to profile phosphorylation of the AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) receptor GluR1 (Munton *et al.*, 2007). Alternatively, synapses can be stimulated *in vivo* as has been done at the whole brain level using the muscarinic AChR antagonist pilocarpine (Trinidad *et al.*, 2013b). Quantitative analysis of protein dynamics at the synapse revealed the existence of a functionally associated core of synaptic proteins. This core, which included key neurotransmitter receptors as well as scaffolding proteins, displayed correlated changes in synaptic abundance in response to activity. Changes in synaptosome composition have also been characterized following chronic administration of morphine. It was demonstrated that the expression of a range of proteins were altered in response to drug exposure, most notably those regulating cell adhesion, G protein signaling, and vesicle trafficking (Abul-Husn *et al.*, 2011). Synaptosomes have been isolated from rats that had been trained to self-administer heroin. Changes in expression levels of AMPA receptors GluR2 and GluR3, as well as the NMDA (N-methyl-D-aspartate) receptor

NR2B were seen, resulting in altered AMPA/NMDA current ratio (Van den Oever *et al.*, 2008).

The protein composition of SVs and posttranslational modification states of these proteins remains an active field of investigation. To date, over 6000 proteins have been identified in synaptosome preparations (Trinidad *et al.*, 2012). Such a large number of components is perhaps not surprising given the complexity and heterologous nature of axon terminals and associated postsynaptic machinery. The mitochondria proteome alone, for example, contains over 1000 proteins (Pagliarini *et al.*, 2008). Synaptosome proteins span a wide range of functional classes. Importantly, they include a large number of posttranslational modifying proteins including kinases, methyltransferases, acetyltransferases as well as SUMO (small ubiquitin-like modifier), and ubiquitin ligases. Consistent with the central role of phosphorylation-based signal transduction, over 300 kinases have been identified in synaptosome preparations. The PTM (posttranslational modifications) content of synaptosomes has been studied at large scale for phosphorylation, O-GlcNAcylation as well as extracellular N- and O-linked glycosylation (Trinidad *et al.*, 2012, 2013a). Over 16 000 distinct sites of phosphorylation are currently known. O-GlcNAcylation is the addition of N-acetylglucosamine to serine and threonine residues on intracellular proteins. This is a dynamic and reversible modification that occurs at a relatively high frequency and the enzymes responsible for its regulation are present at the synapse (Akimoto *et al.*, 2003). To date, over 1700 sites of O-GlcNAcylation have been identified on synaptosomal proteins. This is at least an order of magnitude greater than the number of sites found in any other organ and is consistent with the high expression of OGT in the brain and at the synapse. Extracellular and transmembrane proteins are generally modified with branched N- and O-linked glycosylation (Varki and Sharon, 2009). This type of glycosylation is a challenging posttranslational modification to study because glycosylated peptides fragment extensively during tandem mass spectrometry. The resulting complex fragmentation patterns are difficult to interpret in an automated fashion. In addition, the addition of glycan structures is not template driven, as is the case for protein synthesis. Therefore, the exact sites of glycosylation, as well as nature of attached oligosaccharide cannot be determined *a priori* using genomic or bioinformatic analysis. Over 1500 unique glycopeptides are known on synaptosome proteins (Trinidad *et al.*, 2013a). Each of these instances represents a distinct glycan: glycosylation site combination.

SVs

The term 'synaptic vesicle' was first coined by De Robertis and Bennett in 1955 during a study of electron micrographs of frog sympathetic ganglia and earthworm nerve cord neuropile (De Robertis and Bennett, 1955). They observed vesicles with electron dense circumferences and relatively lighter centers that were in close spatial relationship to the presynaptic membrane. In the mammalian central nervous system, axon terminals contain approximately 10–500 SVs (Pocklington *et al.*, 2006; Schikorski and Stevens, 1997). The human brain

therefore contains an enormous number of SVs. A brain contains 10^{12} neurons, and a total of approximately 10^{15} synapses. Assuming an average of 100 SVs per synapse gives an estimate of 10^{17} SVs, a number that speaks to the centrality of SVs in transmission at chemical synapses. There are approximately 2000 neurotransmitter molecules packaged into a single vesicle giving an approximate concentration of 150 mM (Ryan *et al.*, 1993). In contrast to neurotransmitter vesicles, some aminergic neurotransmitters and neuropeptides are stored in dense core vesicles (Bauerfeind *et al.*, 1995). These vesicles are approximately 100 nm in diameter, and their trafficking and release are regulated by mechanisms that are distinct from those of SVs.

Within a couple hundred milliseconds of an action potential arriving at the axon terminal, SVs that are docked at the active zone fuse with the plasma membrane and release their neurotransmitter into the synaptic cleft (Sabatini and Regehr, 1996). SVs are docked within 100 nm of calcium channels to tightly couple calcium influx with vesicle fusion (Eggermann *et al.*, 2012). SVs at an axon terminal can be divided into three populations, which are defined based on functional differences (Alabi and Tsien, 2012). The first population is the readily releasable (i.e., docked) pool and it constitutes approximately five SVs per synapse (Dobrunz and Stevens, 1997). These vesicles are physically docked at the active zone in close proximity to voltage-gated calcium channels. The calcium-sensing proteins on these SV are thus optimally positioned to detect influx of calcium and initiate fusion. Synapses at the neuromuscular junction are of high fidelity in the sense that an incoming action potential induces fusion of multiple vesicles which results in muscle contraction with essentially 100% efficiency. In contrast, in the hippocampus the probability that an action potential will cause fusion of at least one SV is less than 30% (Murthy *et al.*, 1997). The fact that the release probability of central synapses is significantly less than 100% allows changes in a neural network to be encoded in part by increases or decreases in release probability.

The second population of vesicles is referred to as the recycling pool. This pool is not docked at the active zone, but is positioned to repopulate the readily releasable pool as those vesicles undergo fusion. The recycling pool is approximately three times larger than the readily releasable pool (Murthy and Stevens, 1999). A third pool exists consisting of vesicles which do not readily undergo fusion except during intense bouts of synaptic activity. This pool is called the resting or reserve pool and constitutes approximately 75% of all vesicles at the axon terminal (Rizzoli and Betz, 2005; Alabi and Tsien, 2012). Multiple functions for this pool of vesicles have been proposed, including a role in intracellular signal transduction, and acting to buffer the supply of soluble SV proteins during recycling so that they do not diffuse out of the synapse (Denker *et al.*, 2011). In addition, vesicles from this pool can travel between boutons, a phenomenon that may regulate synapse-specific changes in transmission efficiency (Herzog *et al.*, 2011). It is important to note that these three populations have been defined largely based on electrophysiological studies. In the case of docked vesicles, electron micrograph images clearly confirm a small subset of vesicles located proximal to the active zone, but differences between the recycling and

resting pools are not so distinct. In any case, vesicles must move between populations and at the molecular level; this likely occurs in stages rather than being a binary process.

Physical Characteristics of SVs

The average diameter of SVs varies as a function of organism and neuronal type, but it is within the range of 40–80 nm (Harris and Sultan, 1995). Within a given synaptic population, the size distribution is considerably narrower. For example, in axon terminals from the rat hippocampus the average synaptic diameter is 39 nm with a standard deviation of 5 nm (Qu *et al.*, 2009).

SVs contain numerous distinct membrane lipids. The most abundant of these are cholesterol, phosphatidylcholine, phosphatidylethanolamine, phosphatidylserine, and sphingomyelin. The phospholipid content (which does not include cholesterol) is expected to occupy ~50% of the SV surface (Takamori *et al.*, 2006). The elucidation of most of the principle SV proteins was achieved over the course of many years using a range of techniques, chief among them being genetic analysis (for a detailed review, see Sudhof (2004)). As mentioned above, SVs can be isolated from initial synaptosome preparations (Burre and Volkandt, 2007). A number of approaches exist for their enrichment, including density gradient centrifugation, immunoprecipitation with antibodies against SV proteins, and glass bead chromatography. The resulting preparations are composed of highly purified SV, which include their luminal components (e.g., neurotransmitters) as well as any peripheral membrane proteins associated with the outer surface of the SV. Initial proteomic analysis of SV preparations revealed the identity of 72 distinct proteins (Moricano *et al.*, 2005). A study from the Jahn laboratory has become the major point of reference for understanding SV composition (Takamori *et al.*, 2006). A total of 410 proteins were found highly enriched in SVs. The absolute protein copy number per SV was determined for 15 key proteins. In addition, the average physical characteristics of the SV population were determined, including density, diameter, mass, and volume. The amount of each principle membrane lipid was measured. These factors combined to allow the development of a molecular-level model of a typical SV. Key findings of this study with respect to protein composition were that the SNARE protein synaptobrevin was present at approximately 70 copies per vesicle; neurotransmitter transporters were only present at 10 copies per vesicle; the V-ATPase that provides the proton gradient for the neurotransmitter transporters occurred at approximately 1–2 copies per vesicle. Modeling the proteins and lipids onto a vesicle architecture demonstrated that approximately 20% of the vesicle surface was occupied by transmembrane domains.

Thus far, we have a good overall understanding of a population-average SV, but much remains to be determined about the distribution and characteristics of vesicle subpopulations, both with respect to SVs at different types of chemical synapses as well as different populations of SVs within a single type of synapse. Quantitative comparisons have been performed between glutamatergic and GABAergic SVs that were isolated through the use of antibodies (Pavlos *et al.*, 2010). The vast majority of the 450 proteins identified in

this study were of similar abundance in both classes of SVs. These results indicated that the neurotransmitter transporters themselves are likely the primary determinants of specific classes of SV.

The exact roles of most SV proteins are still poorly understood. While a model exists for the protein composition of SVs, we do not yet fully understand how this set of proteins physically interacts. The binding partners of SV2, a major SV protein, have been determined (Yao *et al.*, 2010). This approach relied on creating a fusion protein of SV2 with an enrichment tag, allowing both vesicle-associated and non-vesicle-associated binding proteins to be identified. Identified binding proteins include AP-1 and AP-2 complex subunits as well as amphiphysin and EPS15. A systematic approach for sequentially enriching and identifying partners of all SV proteins will likely be needed to further our understanding of SV physiology.

Key Proteins Associated with SVs

Broadly speaking, the key proteins that participate in SV function can be organized into three categories: those involved with exocytosis, those involved with endocytosis, and those involved with neurotransmitter packaging/SV localization. While most of the proteins that regulate these functions are not fully understood, the broad outline of this picture has been worked out in detail. Below is a brief description of the most well-characterized SV-associated proteins, which is followed by a more detailed description of their interactions during different phases of the SV life cycle.

Proteins mediating exocytosis

Synaptobrevin

Synaptobrevin, along with syntaxin-1 and SNAP-25 are the three members of the SNARE complex that provide the energy to drive fusion. Synaptobrevin is an 18 kDa protein and has two principle isoforms and is localized to SVs. It is cleaved and inactivated by clostridium toxins. Inactivation of any of the three SNARE proteins prevents SV fusion.

Syntaxin-1

Syntaxins are a family of membrane proteins primarily localized to the plasma membrane of the presynaptic active zone, of which syntaxin-1 is the primary form involved in vesicle fusion. In addition to participating in the SNARE complex, syntaxins bind synaptotagmin on SVs in response to calcium entry. While syntaxin's fusion-related activity depends on its localization in the plasma membrane, some percentage of syntaxin is present in SVs. This is likely due to syntaxin being associated with synaptobrevin when the latter is endocytosed into recycling SVs. It is cleaved and inactivated by botulinum toxin.

SNAP-25

SNAP-25 is a membrane-associated protein of the active zone. It binds to membranes via palmitoylation of multiple cysteine side chains. Like syntaxin, SNAP-25 can interact with synaptotagmin in a calcium dependent fashion. Two splice variants of SNAP-25 exist that vary in their developmental and tissue

distribution (Nagy *et al.*, 2005). It is present in SVs, likely in a similar fashion to syntaxins. Like syntaxin, SNAP-25 is cleaved by botulinum toxin, rendering it inactive.

Synaptotagmin

Synaptotagmins are a family of proteins, of which synaptotagmin-1 has a primary role in SVs. It is a single-pass transmembrane protein containing two cytosolic C-terminal C2 domains that bind calcium. Upon calcium binding, synaptotagmin binds to the SNARE complex and displaces complexin, thereby allowing SV fusion with the plasma membrane (Maximov *et al.*, 2009).

SV2

Synaptic vesicle glycoprotein 2 (SV2) occurs in two main isoforms in the brain, SV2A and SV2B. This protein can bind two adenine molecules and possibly links cellular energy state to vesicle release. Mice engineered to lack both SV2 isoforms possess a smaller number of readily releasable SVs, which implicates SV2 in vesicle priming events (Yao and Bajjalieh, 2008).

Rab3A

The Rab family of proteins contains over 70 proteins in humans and members of this family are monomeric G proteins of the Ras superfamily. Rabs are peripheral membrane proteins that function in essentially all membrane trafficking events in the cell. Rabs bind GDP/GTP and are active in their GTP-bound form and inactive when bound to GDP. Specific Rabs are localized to different membrane compartments. Rab3A is present on SVs, and in its active form it is involved in tethering vesicles to the active zone (Geppert *et al.*, 1997).

Proteins mediating endocytosis and vesicle recycling

NSF

NSF is an acronym for *N*-ethylmaleimide-sensitive factor, an AAA ATPase that exists as a homohexamer of NSF monomers (Glick and Rothman, 1987). NSF acts by coupling the energy of ATP hydrolysis to drive dissociation of SNARE complexes that are present in the *cis* conformation on the plasma membrane. These versions of the SNARE complex exist together in the plasma membrane after vesicle fusion. In order for synaptobrevin to be endocytosed and repackaged into SVs, it must be first separated from SNAP-25 and syntaxin-1 via the action of NSF.

Clathrin

Clathrin-mediated endocytosis is the primary mechanism for SV recycling at the synapse (Pearse, 1976). Two versions of clathrin light chain and two versions of clathrin heavy chain exist. Functional clathrin exists as a heterohexamer of three clathrin heavy chains and three light chains organized in a triskelion arrangement. As membrane is endocytosed from the plasma membrane to reconstitute a SV, clathrin polymerizes into a rounded lattice network to promote endocytosis.

Clathrin adaptor proteins

Clathrin by itself does not polymerize directly onto endocytosing membranes. Rather a family of adaptor proteins directly binds the membrane and mediates clathrin recruitment,

ultimately forming a layer between the clathrin lattice and the membrane. Examples of adaptor proteins include the AP2 complex, AP180, and epsin (Zhang *et al.*, 1998; Shih *et al.*, 1995).

Cysteine-string protein

Cysteine-string protein (CSP) interacts with both syntaxin-1 and synaptotagmin. It likely acts as a chaperone to mediate refolding of denatured SNARE proteins prior or subsequent to endocytosis (Fernandez-Chacon *et al.*, 2004).

Proteins regulating neurotransmitter packaging and SV localization

V-ATPase

The vacuolar ATPase acidifies the lumen of SVs by using the energy of ATP hydrolysis to pump protons into SVs. The ATPase consists of a V1 multiprotein complex that hydrolyzes ATP and a V0 complex that is responsible for translocating protons.

Neurotransmitter transporters

The exact transporter proteins present on an SV vary as a function of synapse type. In general, neurotransmitter transporters will be present not only on SVs, but also in the pre-synaptic plasma membrane as well as neighboring glial cells. The latter two are involved in reuptake of neurotransmitter and termination of synaptic signaling. SV neurotransmitters represent distinct proteins from their plasma membrane counterparts. Unique vesicle transporters exist for glutamate, GABA, glycine, monoamines, and adenosine. A given neurotransmitter may have more than one protein capable of mediating its transport. For example, there are three distinct glutamate transporters (VGLUT1-3).

Synapsin

Synapsins are a family of proteins with three genes in humans. These proteins play a role in localization of SVs, thereby controlling the number of vesicles that are at the active zone and ready for fusion. Synapsins link SVs to the cytoskeleton. The exact mechanism of regulation by synapsin remains unclear, but it is likely that depolarization of the nerve terminal activates the kinase PKA which phosphorylates synapsin. This phosphorylation event weakens the binding of synapsin to the cytoskeleton, thereby mobilizing SVs and allowing them to move to the active zone to replace recently fused vesicles (Bahler and Greengard, 1987).

Synaptophysin

Synaptophysin is highly abundant in SVs, and at 31 copies per vesicle is second only to synaptobrevin in molar amount. Despite its relative abundance, little is known about its biology. It can physically associate with synaptobrevin; however, it is not absolutely required for SV function. Mutant mice deficient in synaptophysin are largely normal (Eshkind and Leube, 1995) but display minor alterations in behavior (Schmitt *et al.*, 2009).

The Synaptic Vesicle Life Cycle

Different neuronal systems have been used to study SVs (Rizzoli and Betz, 2005). These include the calyx of Held and hippocampal nerve termini from rodents, motor nerve terminals of both *Drosophila* and frog, and the ribbon synapse from goldfish. Each of these systems has its relative experimental advantage such as large size (making it amenable to electrophysiological measurement), or ease of genetic manipulation (which allows individual components to be selectively reduced or eliminated). The assumption (generally found to be true) that comes with using a diverse array of biological systems is that the results obtained in a given system will be broadly applicable to most types of synapses.

The notion of SV recycling was first proposed by Heuser and Reese (1973) to describe the discrete stages in the life cycle of SVs. This stemmed from the observation that a mechanism must exist to regenerate SVs that have released their neurotransmitter contents and fused with the plasma membrane. More broadly, the steps below outline the basic stages that occur as vesicles are formed from the Golgi, migrate to axon terminals, dock with the active zone, then fuse with the plasma membrane, and are ultimately endocytosed and reloaded with neurotransmitter.

Formation in the ER/Golgi and trafficking to the synapse

Membrane components of SVs transit through the Golgi apparatus and cluster into domains. These domains subsequently bud off as vesicles. It is not known if these vesicles contain the full complement of SV proteins (and can therefore be considered 'mature') or if additional sorting/processing occurs upon arrival at axon terminals. The vesicles that bud from the Golgi are directed toward axons and vary in their size and shape (Tsukita and Ishikawa, 1980). However, a subpopulation of these vesicles does possess diameters consistent with SVs (Yonekawa *et al.*, 1998). Vesicles destined for axon terminals are transported in an anterograde fashion along axons using the motor protein kinesin (Okada *et al.*, 1995), although distinct populations of vesicle/motor complexes exist (Kaether *et al.*, 2000).

Loading with neurotransmitter

SVs contain transporters for their respective neurotransmitters. As a class, these pumps are anti-transporters and couple efflux of protons with influx of neurotransmitter. In the case of glutamate, GABA, and glycine, two protons are pumped out of the vesicle lumen for every molecule of neurotransmitter pumped in. In the case of catecholamine, histamine, serotonin, and acetylcholine, one proton is expelled for every neurotransmitter pumped into the vesicle. The proton gradient is supplied by the vacuolar ATPase complex, which uses the energy of ATP hydrolysis to pump protons across membranes (in this case from the cytosol into the vesicle lumen). It is estimated that 10 copies of a given neurotransmitter receptor are present on average per SV (Takamori *et al.*, 2006). There are approximately 2000 glutamate molecules per vesicle (Ryan *et al.*, 1993). Measurements in slice cultures give an estimate of 8.4 molecules s^{-1} pumped into a vesicle per neurotransmitter transporter (Hori and Takahashi, 2012).

Docking at the active zone and priming the fusion machinery

The bacteria *Clostridium botulinum* and *Clostridium tetani* secrete botulinum and tetanus toxin, respectively. These toxins were important to the initial elucidation of the SV fusion machinery. They act as proteases and cleave components of the SNARE complex, which inhibits vesicle fusion. This provided a handle to determine the identity of the SNARE proteins (Link *et al.*, 1992).

Vesicles dock at the active zone via interactions between proteins of the SNARE complex, namely synaptobrevin, SNAP-25, and syntaxin. The term 'SNARE' is derived from 'SNAP receptor.' SNAP in turn is an acronym for 'soluble NSF attachment protein.' Synaptobrevin (also known as VAMP) is a single-pass transmembrane protein located on SVs and anchored to the vesicles via its C-terminus. Syntaxin and SNAP-25 are localized to the plasma membrane via their C-terminus and palmitoylation of cysteine residue side chains, respectively. These proteins interact via a parallel four-helix bundle (Poirier *et al.*, 1998). Syntaxin and synaptobrevin each possess a single SNARE motif domain and consequently each contributes a single helix to the bundle. SNAP-25, in contrast, contains two SNARE motif domains, and contributes two helices to the binding complex (Sutton *et al.*, 1998).

While vesicles are docked at the active zone, partially assembled SNARE complexes also interact with SM proteins (for 'Sec1/Munc-18-like'), most notably Munc-18 and Munc13. While this interaction is required for SV fusion, the exact roles played by SM proteins remain unclear. Munc-18 is able to bind to an SM interaction motif on the N-terminus of syntaxin, which allows it to interact with a fully assembled SNARE complex. However, Munc-18 is also able to interact with a second site on syntaxin, primarily while syntaxin is in a closed formation and unable to interact with synaptobrevin and SNAP-25 to form the SNARE complex (Dulubova *et al.*, 1999). Overall, the interactions between Munc18 and syntaxin likely regulate the ability of syntaxin to interact with other SNARE proteins as well as any additional binding partners (Burkhardt *et al.*, 2008). The crystal structure of Mun domain of Munc18 has been solved and is shown to belong to the same structural family as other tethering factors such as the exocyst complex subunit Sec6p (Li *et al.*, 2011).

Fusion with the plasma membrane and release of neurotransmitter

Incoming action potentials depolarize axon terminals, which results in an influx of calcium via voltage-gated calcium channels. This leads to an overall increase in intracellular calcium, but more important than the overall concentration of calcium in the axon terminal is the local concentration of calcium in the active zone. SVs contain a calcium sensor, which is a single-pass transmembrane protein named synaptotagmin. Synaptotagmin possesses two C2 domains that bind calcium and are involved in regulation of SV fusion. Upon calcium binding, synaptotagmin-1 is able to bind simultaneously both SV membrane and plasma membrane (Arac *et al.*, 2006). This brings the SV and plasma membrane domains in closer proximity, promoting fusion.

When vesicles are docked at the plasma membrane, synaptobrevin, SNAP-25, and syntaxin interact in 'trans' fashion, with synaptobrevin in the SV membrane and SNAP-25

and syntaxin in the plasma membrane. Upon exocytosis, the vesicles fuse with the plasma membrane, and all three SNARE components remain associated. At this stage, the complex is then located entirely in the plasma membrane in a 'cis' conformation. The X-ray crystal structure of the SNARE complex composed of syntaxin 1A, SNAP-25, and synaptobrevin 2 has been determined. The complex assumes a helical bundle configuration with extensive interactions of amino acid residue side chains (Stein *et al.*, 2009).

Synapsin is a regulatory protein that is involved in calibrating the number of SVs that are able to undergo fusion at any given time. In its dephosphorylated state, it has been shown to tether SVs to the cytomatrix, thereby organizing the reserve pool of vesicles. Depolarization of the axon terminal activates a number of kinases that can phosphorylate synapsin, including CaMKII, PKA, and MAPK. Upon phosphorylation, synapsin dissociates from SVs (Greengard *et al.*, 1993). This dissociation promotes the migration of SVs from the reserve/recycling to readily releasable pool (Bloom *et al.*, 2003).

As little as two or three SNARE complexes may be needed to enable vesicle fusion (Mohrmann *et al.*, 2010). This applies mainly to fast-type SV fusion events and raises the possibility that variation in the number of SNARE complexes per vesicle may explain the range in vesicle release probabilities. With approximately 70 copies of synaptobrevin per SV, it is possible that vesicles can be docked to the active zone with a fairly broad range in their number of SNARE complexes. However spatial constraints make it unlikely that more than a small portion of these synaptobrevin molecules can simultaneously be involved in SNARE complexes at any given time. Continuous pulling by the SNARE complex is required throughout the exocytosis process, as has been demonstrated from mutagenesis studies on synaptobrevin (Kesavan *et al.*, 2007). Increasing the length of the juxtamembrane region on synaptobrevin decreases overall fusion efficacy.

Endocytosis and reformation of a fusion-competent synaptic vesicle

SV fusion is an exothermic process, and upon complete vesicle fusion, the three SNARE proteins convert from a *trans* to a *cis* conformation. NSF then acts as an ATPase and uses the energy of ATP hydrolysis to disassemble this trimeric complex, thereby releasing them into a higher-energy unbound state to be used again in subsequent rounds of exocytosis. Synaptobrevin is then repackaged into SVs, leaving the majority of syntaxin 1 and SNAP-25 on the plasma membrane.

Three related mechanisms have been proposed to account for the reconstitution of SVs after exocytosis. These mechanisms are 'kiss and run,' clathrin-mediated endocytosis, and bulk endocytosis (Figure 1). Experimental evidence exists to support all three models and the relative degree to which each occurs seems in large part a function of the synapse type studied and how vigorously the axon terminals are stimulated. That being said, clathrin-mediated endocytosis appears to account for the majority of SV recycling under most levels of activity.

Kiss and run fusion

The term kiss-and-run fusion was coined by Ceccarelli (Ceccarelli *et al.*, 1973) and refers to the notion that during some exocytosis events, SVs do not fully fuse with the plasma

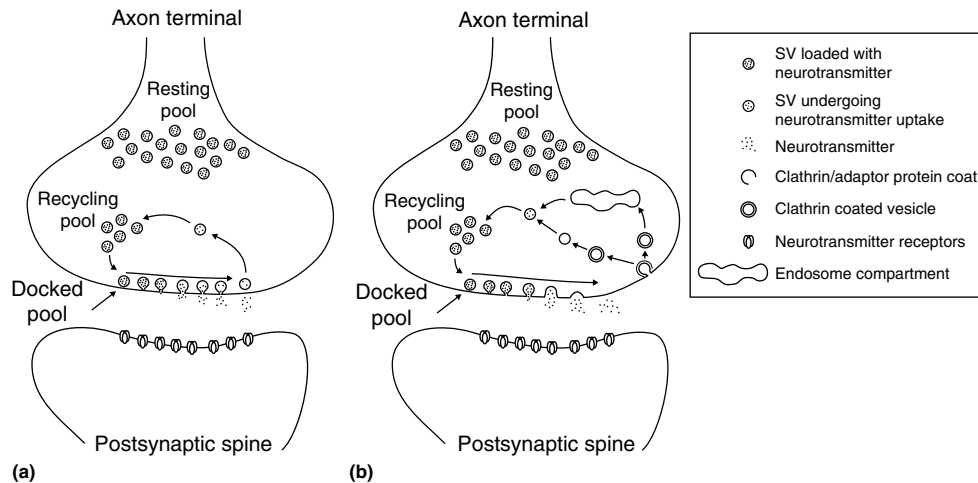


Figure 1 Two models of synaptic vesicle fusion and subsequent endocytosis. Shown in both figures are synapses consisting of an axon terminal (top) and its corresponding postsynaptic spine (bottom). Three populations of SVs are present in the axon terminal. The readily releasable docked pool is tethered to the plasma membrane at the active zone. Upon fusion, neurotransmitter is released into the cleft between the axon terminal and spine. Neurotransmitter diffuses across the cleft and is bound by ligand gated ion channels (neurotransmitter receptors) on the postsynaptic surface. The recycling pool of vesicles is positioned to replenish the docked pool as the latter is depleted by synaptic activity. The resting pool of vesicles does not generally exocytose directly, except perhaps during prolonged intense bouts of activity. (a) The ‘kiss and run’ model of vesicle recycling. The SVs in the active zone show the sequence of events during vesicle fusion. A small, transient pore opens between the SV and the plasma membrane. This allows most of the neurotransmitter to be released. The fusion pore then seals up, reconstituting a mature SV which then is refilled with neurotransmitter and becomes localized to the recycling pool. (b) The clathrin-mediated endocytosis model of vesicle recycling. In this model, SVs collapse completely or almost completely into the plasma membrane, largely mixing their protein and lipid contents and releasing all of their neurotransmitter. Endocytosis occurs proximal to the active zone. Endocytosing vesicles are coated in adaptor proteins and clathrin. Upon fission from the plasma membrane, the clathrin un-coats and vesicles either fill directly with neurotransmitter or may traffic first to an endosomal compartment for further protein sorting.

membrane (**Figure 1(a)**). Rather they form a small fusion pore between the SV and the plasma membrane, which allows neurotransmitter to diffuse through while retaining the overall spherical shape of the SV. This fusion pore is transient, and upon closure, an essentially intact SV is reformed, lacking only in luminal neurotransmitter. Such a mechanism provides a straightforward way to reconstitute SVs. However, from an energetic perspective, there is a strong drive on the part of curved membranes to flatten out subsequent to any initial pore formation, which would imply that the vesicles quickly collapse almost completely into the plasma membrane upon triggering of fusion. Therefore, a mechanism must exist to restrain this from happening, perhaps in the form of a molecular collar around the neck of the fusion pore.

Clathrin-mediated endocytosis

Clathrin-mediated endocytosis is a complex mechanism involving approximately 20 proteins. For a detailed review of the process, see [Saheki and De Camilli \(2012\)](#). It is believed to be the primary mechanism by which vesicles are retrieved after complete fusion with the plasma membrane. Clathrin-mediated endocytosis occurs after complete fusion of SVs to the plasma membrane (**Figure 1(b)**). Clathrin adaptor proteins mediate association between phospholipids and proteins on the plasma membrane and the associating clathrin coat that encloses the endocytosing vesicle. Major adaptor proteins include the AP2 complex, AP180, epsin and stonin. These adaptors are able to associate with each other as well as molecules of synaptotagmin, SV2, and synaptobrevin that are present on the plasma membrane after fusion ([Haucke and](#)

[De Camilli, 1999](#); [Nonet et al., 1999](#)). These multiple cross-linking interactions likely nucleate endocytosis. Multiple copies of the clathrin triskelion are then recruited to the adaptor proteins where they polymerize into a spherical lattice. This results in a mostly formed clathrin-coated vesicle that is attached by a thin stalk to the plasma membrane. Final fission of the vesicle is achieved by the action of a GTPase known as dynamin ([Koenig and Ikeda, 1989](#)). Quickly after vesicle fission, the clathrin coat depolymerizes via the action of hsc70 ([DeLuca-Flaherty et al., 1990](#); [Guan et al., 2010](#)). This process can generate a functional SV directly without the need for fusion of the clathrin-coated vesicle with an endosomal compartment ([Takei et al., 1996](#)).

Bulk endocytosis

Bulk membrane endocytosis has been observed to occur, generally after strong synaptic stimulation or under conditions where clathrin-mediated endocytosis is inhibited. Relatively large invaginations of the plasma membrane can be seen by electron microscopy. These infoldings break off from the plasma membrane and generate endosome-like vacuoles ([Miller and Heuser, 1984](#)). These vacuoles are not stable and presumably undergo fission events concurrent with the formation of new SVs ([Teng et al., 2007](#)).

See also: Molecular Principles, Components, Technology, and Concepts: Membranes: Composition, Physical Properties, and Curvature

References

- Abul-Husn, N.S., Annangudi, S.P., Ma'ayan, A., *et al.*, 2011. Chronic morphine alters the presynaptic protein profile: Identification of novel molecular targets using proteomics and network analysis. *PLoS One* 6, e25535.
- Akimoto, Y., Comer, F.I., Cole, R.N., *et al.*, 2003. Localization of the O-GlcNAc transferase and O-GlcNAc-modified proteins in rat cerebellar cortex. *Brain Research* 966, 194–205.
- Alabi, A.A., Tsien, R.W., 2012. Synaptic vesicle pools and dynamics. *Cold Spring Harbor Perspectives in Biology* 4, a013680.
- Arac, D., Chen, X., Khant, H.A., *et al.*, 2006. Close membrane-membrane proximity induced by Ca^{2+} -dependent multivalent binding of synaptotagmin-1 to phospholipids. *Nature Structural & Molecular Biology* 13, 209–217.
- Bahler, M., Greengard, P., 1987. Synapsin I bundles F-actin in a phosphorylation-dependent manner. *Nature* 326, 704–707.
- Bauerfeind, R., Jelinek, R., Hellwig, A., Huttner, W.B., 1995. Neurosecretory vesicles can be hybrids of synaptic vesicles and secretory granules. *Proceedings of the National Academy of Sciences of the United States of America* 92, 7342–7346.
- Bayes, A., van de Lagemaat, L.N., Collins, M.O., *et al.*, 2011. Characterization of the proteome, diseases and evolution of the human postsynaptic density. *Nature Neuroscience* 14, 19–21.
- Blaustein, M.P., Goldring, J.M., 1975. Membrane potentials in pinched-off presynaptic nerve terminals monitored with a fluorescent probe: Evidence that synaptosomes have potassium diffusion potentials. *Journal of Physiology* 247, 589–615.
- Bloom, O., Evergren, E., Tomilin, N., *et al.*, 2003. Colocalization of synapsin and actin during synaptic vesicle recycling. *Journal of Cell Biology* 161, 737–747.
- Booth, R.F., Clark, J.B., 1978. A rapid method for the preparation of relatively pure metabolically competent synaptosomes from rat brain. *Biochemical Journal* 176, 365–370.
- Burkhardt, P., Hattendorf, D.A., Weis, W.I., Fasshauer, D., 2008. Munc18a controls SNARE assembly through its interaction with the syntaxin N-peptide. *EMBO Journal* 27, 923–933.
- Burre, J., Volkmand, W., 2007. The synaptic vesicle proteome. *Journal of Neurochemistry* 101, 1448–1462.
- Ceccarelli, B., Hurlbut, W.P., Mauro, A., 1973. Turnover of transmitter and synaptic vesicles at the frog neuromuscular junction. *Journal of Cell Biology* 57, 499–524.
- Cheng, D., Hoogenraad, C.C., Rush, J., *et al.*, 2006. Relative and absolute quantification of postsynaptic density proteome isolated from rat forebrain and cerebellum. *Molecular & Cellular Proteomics* 5, 1158–1170.
- De Belleruche, J.S., Bradford, H.F., 1972. Metabolism of beds of mammalian cortical synaptosomes: Response to depolarizing influences. *Journal of Neurochemistry* 19, 585–602.
- DeLuca-Flaherty, C., Mckay, D.B., Parham, P., Hill, B.L., 1990. Uncoating protein (hsc70) binds a conformationally labile domain of clathrin light chain LCa to stimulate ATP hydrolysis. *Cell* 62, 875–887.
- Denker, A., Krohnert, K., Buckers, J., Neher, E., Rizzoli, S.O., 2011. The reserve pool of synaptic vesicles acts as a buffer for proteins involved in synaptic vesicle recycling. *Proceedings of the National Academy of Sciences of the United States of America* 108, 17183–17188.
- De Robertis, E.D., Bennett, H.S., 1955. Some features of the submicroscopic morphology of synapses in frog and earthworm. *Journal of Biophysical and Biochemical Cytology* 1, 47–58.
- Dobrunz, L.E., Stevens, C.F., 1997. Heterogeneity of release probability, facilitation, and depletion at central synapses. *Neuron* 18, 995–1008.
- Dulubova, I., Sugita, S., Hill, S., *et al.*, 1999. A conformational switch in syntaxin during exocytosis: Role of munc18. *EMBO Journal* 18, 4372–4382.
- Eggermann, E., Bucurenciu, I., Goswami, S.P., Jonas, P., 2012. Nanodomain coupling between Ca^{2+} channels and sensors of exocytosis at fast mammalian synapses. *Nature Reviews Neuroscience* 13, 7–21.
- Eshkind, L.G., Leube, R.E., 1995. Mice lacking synaptophysin reproduce and form typical synaptic vesicles. *Cell and Tissue Research* 282, 423–433.
- Fernandez-Chacon, R., Wolfel, M., Nishimune, H., *et al.*, 2004. The synaptic vesicle protein CSP alpha prevents presynaptic degeneration. *Neuron* 42, 237–251.
- Geppert, M., Goda, Y., Stevens, C.F., Sudhof, T.C., 1997. The small GTP-binding protein Rab3A regulates a late step in synaptic vesicle fusion. *Nature* 387, 810–814.
- Glick, B.S., Rothman, J.E., 1987. Possible role for fatty acyl-coenzyme A in intracellular protein transport. *Nature* 326, 309–312.
- Gray, E.G., Whittaker, V.P., 1960. The isolation of synaptic vesicles from the central nervous system. *Journal of Physiology* 153, P35–P37.
- Greengard, P., Valtorta, F., Czernik, A.J., Benfenati, F., 1993. Synaptic vesicle phosphoproteins and regulation of synaptic function. *Science* 259, 780–785.
- Guan, R., Dai, H., Harrison, S.C., Kirchhausen, T., 2010. Structure of the PTEN-like region of auxilin, a detector of clathrin-coated vesicle budding. *Structure* 18, 1191–1198.
- Harris, K.M., Sultan, P., 1995. Variation in the number, location and size of synaptic vesicles provides an anatomical basis for the nonuniform probability of release at hippocampal CA1 synapses. *Neuropharmacology* 34, 1387–1395.
- Haucke, V., De Camilli, P., 1999. AP-2 recruitment to synaptotagmin stimulated by tyrosine-based endocytic motifs. *Science* 285, 1268–1271.
- Hebb, C.O., Whittaker, V.P., 1958. Intracellular distributions of acetylcholine and choline acetylase. *Journal of Physiology* 142, 187–196.
- Herzog, E., Nadrigny, F., Silm, K., *et al.*, 2011. In vivo imaging of intersynaptic vesicle exchange using VGLUT1 Venus knock-in mice. *Journal of Neuroscience* 31, 15544–15559.
- Heuser, J.E., Reese, T.S., 1973. Evidence for recycling of synaptic vesicle membrane during transmitter release at the frog neuromuscular junction. *Journal of Cell Biology* 57, 315–344.
- Hori, T., Takahashi, T., 2012. Kinetics of synaptic vesicle refilling with neurotransmitter glutamate. *Neuron* 76, 511–517.
- Israel, M., Whittaker, V.P., 1965. The isolation of mossy fibre endings from the granular layer of the cerebellar cortex. *Experientia* 21, 325–326.
- Kaether, C., Skehel, P., Dotti, C.G., 2000. Axonal membrane proteins are transported in distinct carriers: A two-color video microscopy study in cultured hippocampal neurons. *Molecular Biology of the Cell* 11, 1213–1224.
- Kesavan, J., Borisovska, M., Bruns, D., 2007. v-SNARE actions during Ca^{2+} -triggered exocytosis. *Cell* 131, 351–363.
- Koenig, J.H., Ikeda, K., 1989. Disappearance and reformation of synaptic vesicle membrane upon transmitter release observed under reversible blockage of membrane retrieval. *Journal of Neuroscience* 9, 3844–3860.
- Li, W., Ma, C., Guan, R., *et al.*, 2011. The crystal structure of a Munc13 C-terminal module exhibits a remarkable similarity to vesicle tethering factors. *Structure* 19, 1443–1455.
- Link, E., Edelmann, L., Chou, J.H., *et al.*, 1992. Tetanus toxin action: Inhibition of neurotransmitter release linked to synaptobrevin proteolysis. *Biochemical and Biophysical Research Communications* 189, 1017–1023.
- Maximov, A., Tang, J., Yang, X., Pang, Z.P., Sudhof, T.C., 2009. Complexin controls the force transfer from SNARE complexes to membranes in fusion. *Science* 323, 516–521.
- Miller, T.M., Heuser, J.E., 1984. Endocytosis of synaptic vesicle membrane at the frog neuromuscular junction. *Journal of Cell Biology* 98, 685–698.
- Mohrmann, R., de Wit, H., Verhage, M., Neher, E., Sorensen, J.B., 2010. Fast vesicle fusion in living cells requires at least three SNARE complexes. *Science* 330, 502–505.
- Morciano, M., Burre, J., Corvey, C., *et al.*, 2005. Immunolocalization of two synaptic vesicle pools from synaptosomes: A proteomics analysis. *Journal of Neurochemistry* 95, 1732–1745.
- Munton, R.P., Tweedie-Cullen, R., Livingstone-Zatchej, M., *et al.*, 2007. Qualitative and quantitative analyses of protein phosphorylation in naive and stimulated mouse synaptosomal preparations. *Molecular & Cellular Proteomics* 6, 283–293.
- Murthy, V.N., Sejnowski, T.J., Stevens, C.F., 1997. Heterogeneous release properties of visualized individual hippocampal synapses. *Neuron* 18, 599–612.
- Murthy, V.N., Stevens, C.F., 1999. Reversal of synaptic vesicle docking at central synapses. *Nature Neuroscience* 2, 503–507.
- Nagy, A., Delgado-Escueta, A.V., 1984. Rapid preparation of synaptosomes from mammalian brain using nontoxic isoosmotic gradient material (Percoll). *Journal of Neurochemistry* 43, 1114–1123.
- Nagy, G., Milosevic, I., Fasshauer, D., *et al.*, 2005. Alternative splicing of SNAP-25 regulates secretion through nonconservative substitutions in the SNARE domain. *Molecular Biology of the Cell* 16, 5675–5685.
- Nonet, M.L., Holgado, A.M., Brewer, F., *et al.*, 1999. UNC-11, a *Caenorhabditis elegans* AP180 homologue, regulates the size and protein composition of synaptic vesicles. *Molecular Biology of the Cell* 10, 2343–2360.
- Okada, Y., Yamazaki, H., Sekine-Aizawa, Y., Hirokawa, N., 1995. The neuron-specific kinesin superfamily protein KIF1A is a unique monomeric motor for anterograde axonal transport of synaptic vesicle precursors. *Cell* 81, 769–780.
- Pagliarini, D.J., Calvo, S.E., Chang, B., *et al.*, 2008. A mitochondrial protein compendium elucidates complex I disease biology. *Cell* 134, 112–123.
- Pavlos, N.J., Gronborg, M., Riedel, D., *et al.*, 2010. Quantitative analysis of synaptic vesicle Rabs uncovers distinct yet overlapping roles for Rab3a and Rab27b in Ca^{2+} -triggered exocytosis. *Journal of Neuroscience* 30, 13441–13453.

- Pearse, B.M., 1976. Clathrin: A unique protein associated with intracellular transfer of membrane by coated vesicles. *Proceedings of the National Academy of Sciences of the United States of America* 73, 1255–1259.
- Pocklington, A.J., Armstrong, J.D., Grant, S.G., 2006. Organization of brain complexity – synapse proteome form and function. *Briefings in Functional Genomics & Proteomics* 5, 66–73.
- Poirier, M.A., Xiao, W., Macosko, J.C., *et al.*, 1998. The synaptic SNARE complex is a parallel four-stranded helical bundle. *Nature Structural & Molecular Biology* 5, 765–769.
- Qu, L., Akbergenova, Y., Hu, Y., Schikorski, T., 2009. Synapse-to-synapse variation in mean synaptic vesicle size and its relationship with synaptic morphology and function. *Journal of Comparative Neurology* 514, 343–352.
- Rizzoli, S.O., Betz, W.J., 2005. Synaptic vesicle pools. *Nature Reviews Neuroscience* 6, 57–69.
- Ryan, T.A., Reuter, H., Wendland, B., *et al.*, 1993. The kinetics of synaptic vesicle recycling measured at single presynaptic boutons. *Neuron* 11, 713–724.
- Sabatini, B.L., Regehr, W.G., 1996. Timing of neurotransmission at fast synapses in the mammalian brain. *Nature* 384, 170–172.
- Saheki, Y., De Camilli, P., 2012. Synaptic vesicle endocytosis. *Cold Spring Harbor Perspectives in Biology* 4, a005645.
- Schikorski, T., Stevens, C.F., 1997. Quantitative ultrastructural analysis of hippocampal excitatory synapses. *Journal of Neuroscience* 17, 5858–5867.
- Schmitt, U., Tanimoto, N., Seeliger, M., Schaeffel, F., Leube, R.E., 2009. Detection of behavioral alterations and learning deficits in mice lacking synaptophysin. *Neuroscience* 162, 234–243.
- Shih, W., Gallusser, A., Kirchhausen, T., 1995. A clathrin-binding site in the hinge of the beta 2 chain of mammalian AP-2 complexes. *Journal of Biological Chemistry* 270, 31083–31090.
- Stein, A., Weber, G., Wahl, M.C., Jahn, R., 2009. Helical extension of the neuronal SNARE complex into the membrane. *Nature* 460, 525–528.
- Sudhof, T.C., 2004. The synaptic vesicle cycle. *Annual Review of Neuroscience* 27, 509–547.
- Sutton, R.B., Fasshauer, D., Jahn, R., Brunger, A.T., 1998. Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature* 395, 347–353.
- Takamori, S., Holt, M., Stenius, K., *et al.*, 2006. Molecular anatomy of a trafficking organelle. *Cell* 127, 831–846.
- Takei, K., Mundigl, O., Daniell, L., De Camilli, P., 1996. The synaptic vesicle cycle: A single vesicle budding step involving clathrin and dynamin. *Journal of Cell Biology* 133, 1237–1250.
- Teng, H., Lin, M.Y., Wilkinson, R.S., 2007. Macroendocytosis and endosome processing in snake motor boutons. *Journal of Physiology* 582, 243–262.
- Trinidad, J.C., Barkan, D.T., Gullledge, B.F., *et al.*, 2012. Global identification and characterization of both O-GlcNAcylation and phosphorylation at the murine synapse. *Molecular & Cellular Proteomics* 11, 215–229.
- Trinidad, J.C., Schoepfer, R., Burlingame, A.L., Medzihradszky, K.F., 2013a. N- and O-glycosylation in the murine synaptosome. *Molecular & Cellular Proteomics* 12, 3474–3488.
- Trinidad, J.C., Thalhammer, A., Burlingame, A.L., Schoepfer, R., 2013b. Activity-dependent protein dynamics define interconnected cores of co-regulated postsynaptic proteins. *Molecular & Cellular Proteomics* 12, 29–41.
- Trinidad, J.C., Thalhammer, A., Specht, C.G., Schoepfer, R., Burlingame, A.L., 2005. Phosphorylation state of postsynaptic density proteins. *Journal of Neurochemistry* 92, 1306–1316.
- Tsukita, S., Ishikawa, H., 1980. The movement of membranous organelles in axons. Electron microscopic identification of anterogradely and retrogradely transported organelles. *Journal of Cell Biology* 84, 513–530.
- Van den Oever, M.C., Goriounova, N.A., Li, K.W., *et al.*, 2008. Prefrontal cortex AMPA receptor plasticity is crucial for cue-induced relapse to heroin-seeking. *Nature Neuroscience* 11, 1053–1058.
- Varki, A., Sharon, N., 2009. Historical background and overview. In: Varki, A., Cummings, R.D., Esko, J.D. *et al.* (Eds.), *Essentials of Glycobiology*, second ed. New York, NY: Cold Spring Harbor.
- Walikonis, R.S., Jensen, O.N., Mann, M., *et al.*, 2000. Identification of proteins in the postsynaptic density fraction by mass spectrometry. *Journal of Neuroscience* 20, 4069–4080.
- Whittaker, V.P., 1965. The application of subcellular fractionation techniques to the study of brain function. *Progress in Biophysics and Molecular Biology* 15, 39–96.
- Whittaker, V.P., Michaelson, I.A., Kirkland, R.J., 1964. The separation of synaptic vesicles from nerve-ending particles ('synaptosomes'). *Biochemical Journal* 90, 293–303.
- Yao, J., Bajjalieh, S.M., 2008. Synaptic vesicle protein 2 binds adenine nucleotides. *Journal of Biological Chemistry* 283, 20628–20634.
- Yao, J., Nowack, A., Kinsel-Hammes, P., Gardner, R.G., Bajjalieh, S.M., 2010. Cotrafficking of SV2 and synaptotagmin at the synapse. *Journal of Neuroscience* 30, 5569–5578.
- Yonekawa, Y., Harada, A., Okada, Y., *et al.*, 1998. Defect in synaptic vesicle precursor transport and neuronal cell death in KIF1A motor protein-deficient mice. *Journal of Cell Biology* 141, 431–441.
- Zhang, B., Koh, Y.H., Beckstead, R.B., *et al.*, 1998. Synaptic vesicle size and number are regulated by a clathrin adaptor protein required for endocytosis. *Neuron* 21, 1465–1475.

Further Reading

- Harris, K.M., Weinberg, R.J., 2012. Ultrastructure of synapses in the mammalian brain. *Cold Spring Harbor Perspectives in Biology* 4, 5.
- Ramakrishnan, N.A., Drescher, M.J., Drescher, D.G., 2012. The SNARE complex in neuronal and sensory cells. *Molecular and Cellular Neuroscience* 50, 58–69.
- Sudhof, T.C., Rizo, J., 2011. Synaptic vesicle exocytosis. *Cold Spring Harbor Perspectives in Biology* 3 (12), a005637.

Extracellular Vesicles

MMH Wauben, Utrecht University, Utrecht, The Netherlands

© 2016 Elsevier Inc. All rights reserved.

Extracellular Vesicles

Vesicles are small organelles consisting of fluid enclosed by a lipid bilayer. Their importance in intracellular transport and compartmentalization of the intracellular milieu has been appreciated with the 2013 Nobel Prize in Physiology or Medicine for Rothman, Schekman, and Südhof for their discoveries of machinery regulating vesicle traffic. Besides vesicular traffic inside cells, it is already known for several decades that cells also release vesicles into their environment. Nowadays, it has become clear that all living cells ranging from unicellular bacteria, archaea, and fungi to cells from parasites, plants, and animals have the capacity to release vesicles into the extracellular milieu. Vesicle release thus spans all kingdoms of life and is an evolutionary conserved process impacting organisms throughout the tree of life. The overall presence of extracellular vesicles (EVs) in all body fluids and the environment (e.g., EVs are abundantly present in oceans), further corroborates that these extracellular organelles are part of ecosystems.

The population of EVs is very heterogeneous and although cells constitutively release EVs, the amount of EVs and their cargo of bioactive macromolecules such as proteins, lipids, and nucleic acids are highly regulated and dependent on the cell type and cellular state. Interestingly, the composition of EVs is markedly different from the cell of origin. EVs from eukaryotic cells are enriched in specific lipids (e.g., cholesterol, sphingomyelin, and ceramide) and in specific proteins that are involved in the endosomal route and vesicle formation (e.g., tetraspanins, members of the Rab GTPase family and chaperones) (Kowal *et al.*, 2014). Also, the glycosylation pattern of EVs is distinct from the cell of origin. Nowadays, it is clear that besides selected sets of proteins and lipids, EVs also contain a wide variety of nucleic acids including all kinds of noncoding RNA biotypes, transposable elements, and extrachromosomal DNA (Villarroya-Beltri *et al.*, 2014). The demonstration that EV-associated RNAs were functionally transferred to recipient cells (Ratajczak *et al.*, 2006; Valadi *et al.*, 2007) further substantiated the concept that EVs are vehicles for intercellular communication involved in the horizontal transfer of combined biological signals to specific target cells.

EV-mediated intercellular communication has several advantages. The regulated release and composition of EVs allows that tailor-made nanoparticles composed of different physiologically relevant signal molecules are spread in a timely fashion. The incorporation into EVs protects luminal cargo for degradation and prevents undesired collateral effects. The presence of specific receptors on the surface contributes to specific targeting to destine cells. Finally, the processing of EVs after binding to the target cell determines how EV-associated components can signal to the target cell.

The Physiological Role of EVs: From Waste Bin to Versatile Transporter of Biological Signals

Although the process of microvesiculation is known for several decades, the physiological relevance of EVs is just emerging. Already in the 1960s, both prokaryotic cell-derived EVs from Gram-negative bacteria and eukaryotic cell-derived EVs from platelets (platelet dust) and bone matrix were identified. In historical perspective, the first EVs with a defined physiological function were bone matrix EVs having a role in bone mineralization and bone homeostasis (Golub, 2009) and procoagulant platelet-derived EVs in hemostasis (Owens and Mackman, 2011). In the 1980s, it was observed that during reticulocyte maturation, organelles filled with small membrane-enclosed vesicles (30–100 nm in size) were present in the cytosol which could fuse with the plasma membrane thereby releasing their intraluminal vesicles in the extracellular milieu (Harding *et al.*, 1983; Pan *et al.*, 1985). These so-called exosomes were regarded as waste bins dumping obsolete cellular products in the environment. The recent observations that EVs may be involved in the outsourcing of autophagy to other cells suggest that EVs play an active role in the recycling of cellular components. The findings in the 1990s that EVs derived from B lymphocytes and dendritic cells can activate adaptive immune responses were indicative for a physiological role of EVs in the immune system (Raposo *et al.*, 1996; Zitvogel *et al.*, 1998). Nowadays the picture emerges that EVs are an integral part of the innate and adaptive immune system involved in immune regulation (Robbins and Morelli, 2014). In mammals, EVs have additionally been shown to be involved in many physiological processes including embryonic development and homeostasis (Zhang and Wrana, 2014), reproduction and pregnancy (Ronquist, 2012; Arck and Hecher, 2013), communication in the nervous system (Frühbeis *et al.*, 2013), and tissue repair (Tetta *et al.*, 2011).

For unicellular organisms intercellular communication is essential to control population density and differentiation. In bacteria this so-called quorum sensing is a highly regulated process in which bacterial EVs play a role in the exchange of essential information between bacteria (Kulp and Kuehn, 2010; Deatherage and Cookson, 2012). Furthermore, bacterial EVs can interact with other bacterial species which can result in either negative effects, for example, lysis of bacteria, or positive effects such as the transfer of antibiotic-resistance enzymes or the transfer of enzymes for the acquisition of nutrients (Kulp and Kuehn, 2010; Deatherage and Cookson, 2012). Bacterial EVs are also involved in biofilm formation and as such play a general role in optimal niche formation. Under stress-conditions hypervesiculation has been observed in bacteria and it is assumed that EVs play a role in the defense mechanism to cope with harsh conditions (Kulp and Kuehn, 2010; Deatherage and Cookson, 2012). The findings that bacterial

EVs are very abundant in natural ecosystems, such as oceans, and that these EVs support the growth of heterotrophic bacterial cultures in these natural systems further substantiates the physiological role of bacterial EVs in natural ecosystems (Biller *et al.*, 2014).

Also parasites use quorum-sensing-like mechanisms in which EVs mediate the transmission of signals between parasites (Twu and Johnson, 2014). Parasite-derived molecules can be found on EVs derived from the parasite itself or, in case of intracellular parasites, on EVs derived from infected host cells. For example, malaria parasites inside red blood cells use EVs released by infected cells as vehicle to signal and sense malaria parasites present in other erythrocytes, thereby promoting a collective differentiation to sexual forms of the parasite population (Regev-Rudzki *et al.*, 2013). This communication and sensing mechanism are probably widespread among protozoa, as these populations often shift in a synchronized fashion from one state to another, reminiscent of quorum-sensing in bacteria. A recent finding in fruit flies indicates that EVs produced by cells of the accessory glands in the male reproductive tract can reprogram female behavior, thereby inducing reproductive advantage of the first male to mate (Corrigan *et al.*, 2014). These findings clearly demonstrate that EV-mediated regulation of intercellular communication reaches far beyond the confinements of the cell of origin and even of the organism in which the EVs are produced.

EVs in Cross-Kingdom Communication

In experimental science many studies are performed using isolated cells and different organisms are usually studied separately. In nature, however, different organisms continuously compete and collaborate in complex ecosystems. The fact that EV production is such an evolutionary conserved process, together with the abundance of EVs present in body fluids and the environment, triggered the idea that EVs play a role in cross-kingdom communication and even in the communication between the three domains of life: Eukaryota, and the prokaryotic Bacteria and Archaea (Figure 1).

Indeed it becomes apparent that prokaryotic cell-derived EVs not only play a role in the communication between prokaryotes but also between prokaryotes and their eukaryotic host. During host–pathogen interactions, pathogenic bacteria release EVs impacting pathogen survival and antibiotic resistance. These versatile EVs highly enriched in virulence factors are delivered intracellularly, for example, inside phagosomes, or extracellularly to host cells resulting in either immune activation or immune evasion. For example, bacterial EVs containing Toll-like receptor (TLR) and nucleotide-binding oligomerization domain receptor (NLR) activation ligands have been implicated in innate immune activation and induction of inflammation (Ferrand and Ferrero, 2013). Besides pathogenic effects, bacterial EVs can also contribute to

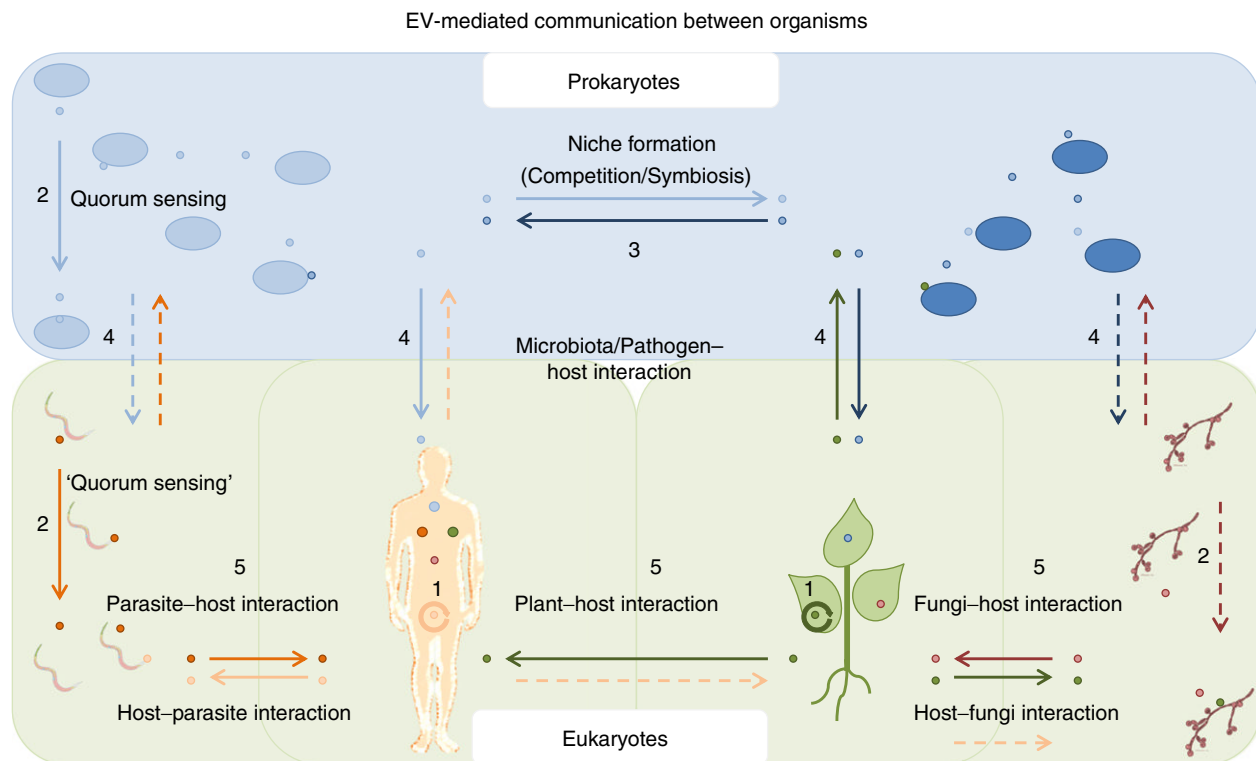


Figure 1 EV-mediated communication is an evolutionary conserved mechanism of intercellular communication. (1) In multicellular organisms EVs can signal to neighboring cells as well as to distant cells via body fluids or in plants via the phloem. (2) In unicellular organisms or intracellular living parasites, EV-mediated signal transmission plays a role in controlling population density and differentiation, i.e., quorum sensing. (3) Between different prokaryotes, EV-mediated communication plays a role in niche formation, competition, and symbiosis. (4) EV-mediated cross-kingdom communication between prokaryotes and eukaryotes, and (5) between different eukaryotic kingdoms is involved in pathogenicity and immune modulation. EVs are derived from blood cells and vessels.

symbiosis between bacteria and the host. For example, EVs from human commensal bacteria can communicate with immune cells of the host and induce immune tolerance (Shen *et al.*, 2012). Now there is clear evidence that prokaryotic cell-derived EVs can influence eukaryotic organisms in both the plant and mammalian world. The reciprocal interaction, influencing prokaryotes via eukaryotic cell-derived EVs, needs to be further substantiated. Yet, there is some evidence for this EV-pathway in higher plants. In higher plants, pathogens and microbiota typically do not enter the host cells. Hence the molecular interaction between these organisms and the host occur in the extracellular space, the apoplast. During environmental stress, such as pathogen attack by bacteria or fungi, the plant defense system releases specific proteins in the apoplast (Regente *et al.*, 2012; Kaffarnik *et al.*, 2009). Secretome analysis and direct evidence for fusion of multivesicular bodies (MVBs) with the plasma membrane resulting in the release of EVs indicate that plant-derived EVs participate in the defense response to pathogens (Delaunois *et al.*, 2014). Although no direct evidence is available on mammalian cell-derived EVs influencing prokaryotic cells, evidence exists for the cross talk with eukaryotic parasites. During protozoan parasite infections, gastrointestinal epithelial cell-derived EVs from the host containing antimicrobial peptides, can be released in the gut lumen and as such can be regarded as an arm of the mucosal immunity relevant to antimicrobial defense (Hu *et al.*, 2013). Also eukaryotic parasites release EVs, either intracellularly for intracellular living parasite stages or extracellularly for free-living stages. These EVs can promote growth and increase virulence of the parasites or can modulate the host immune response (Mantel and Marti, 2014). Also fungal EVs can interfere with the activity of host immune effector cells and can increase fungal pathogenesis (Rodrigues *et al.*, 2014). The observation that specific rice-derived microRNAs could be detected in blood after rice consumption and can regulate mammalian gene expression (Zhang *et al.*, 2012) has sparked the idea that diet-derived EVs can modulate host cells. Although it is still a matter of debate whether diet-derived microRNAs are associated with EVs and whether their amount in the host reaches high enough levels to be of any physiological relevance (Witwer *et al.*, 2013b), the idea that cross-kingdom communication is possible via EVs derived from plant-food components is intriguing and may represent another level of communication between different organisms.

Future studies addressing EV-mediated communication between prokaryotes and parasites or fungi, as well as between animals and plants, may uncover additional EV pathways between these biological systems (Figure 1). Altogether, the picture emerges that nano-sized EVs are master regulators in and between biological systems. Having the cells as confined building blocks for individual organisms, the sensing and signaling via EVs adds an extra level of regulation in biological systems.

EV Biogenesis

The ultimate physiological effect of EVs is dependent on their cargo. EVs contain functional proteins, lipids, and nucleic acids that are incorporated in EVs in a highly regulated

manner. The EV cargo is dependent on the cell of origin and the state of this cell. Additionally, the intracellular origin may determine the composition of EVs. Eukaryotic cells can release a plethora of different kinds of EVs that can be classified based on size and biogenesis route (Figure 2).

During programmed cell death, apoptotic cells release the so-called apoptotic bodies, which are relatively large (Raposo and Stoorvogel, 2013). Cancer cells can also release relatively large EVs; oncosomes (Morello *et al.*, 2013). However, the vast majority of EVs released by living cells are much smaller in size (<200 nm). These EVs can be formed either by budding of the plasma membrane, often referred to as microvesicles, microparticles, or ectosomes, or can be derived from MVBs, the so-called exosomes (Kowal *et al.*, 2014). Although the biogenesis, and thus cargo and function, may be diverse for small EVs (<200 nm) no discriminatory markers are available yet to distinguish plasma membrane-derived EVs from exosomes.

The EV biogenesis pathways, in particular the exosome pathway, are best studied in mammalian cells. Although originally thought that EVs always bud from the plasma membrane, studies on reticulocyte maturation in the 1980s described a more complex mode of EV release; the exosome pathway (Harding *et al.*, 1983; Pan *et al.*, 1985). Later, this pathway has been confirmed to exist in many eukaryotic cells. Exosomes are derived from the endosomal pathway in which intraluminal vesicles are formed by inward budding inside endosomes. Instead of degradation into lysosomes, the resulting MVBs fuse with the plasma membrane thereby releasing the intraluminal vesicles, exosomes, in the extracellular milieu. Different families of molecules are involved in intracellular vesicle formation and the subsequent release of exosomes, suggesting that even exosomes can be further divided in subtypes (Kowal *et al.*, 2014). In eukaryotic cells the endosomal sorting complex required for transport (ESCRT) machinery is involved in the formation of intraluminal vesicles and regulates membrane budding processes at the cell surface. ESCRT complexes (ESCRT 0-III) play a role in ubiquitin-dependent cargo selection and clustering, bud formation and vesicle scission. The involvement of the ESCRT machinery in exosome biogenesis is corroborated by the general identification of ESCRT components TSG101 and Alix in EV proteomics studies (Kowal *et al.*, 2014). An alternative ESCRT-independent pathway for the biogenesis of intraluminal vesicles appeared to be mediated by a lipid-dependent (ceramide-dependent) mechanism. The lipid composition of exosomes shares many properties with lipid rafts, for example, the enrichment in cholesterol, sphingolipids, and ceramides. Ceramides cluster in microdomains and favor budding processes (Raposo and Stoorvogel, 2013). Whether multiple mechanisms of intraluminal vesicle biogenesis occur in a single MVB or whether different mechanisms occur in different MVB types within an individual cell needs to be validated. For the specific sorting of proteins into exosomes also tetraspanins, four transmembrane domain family proteins, and chaperone HSC70 are involved. Although the packaging of nucleic acids in EVs has not yet been fully resolved, the finding that sumoylated protein hnRNP2B1 controls the selection of microRNAs through the recognition of sequence motifs indicates that incorporation of nucleic acids in EVs is a selective process (Villarroya-Beltri *et al.*, 2013). The subsequent

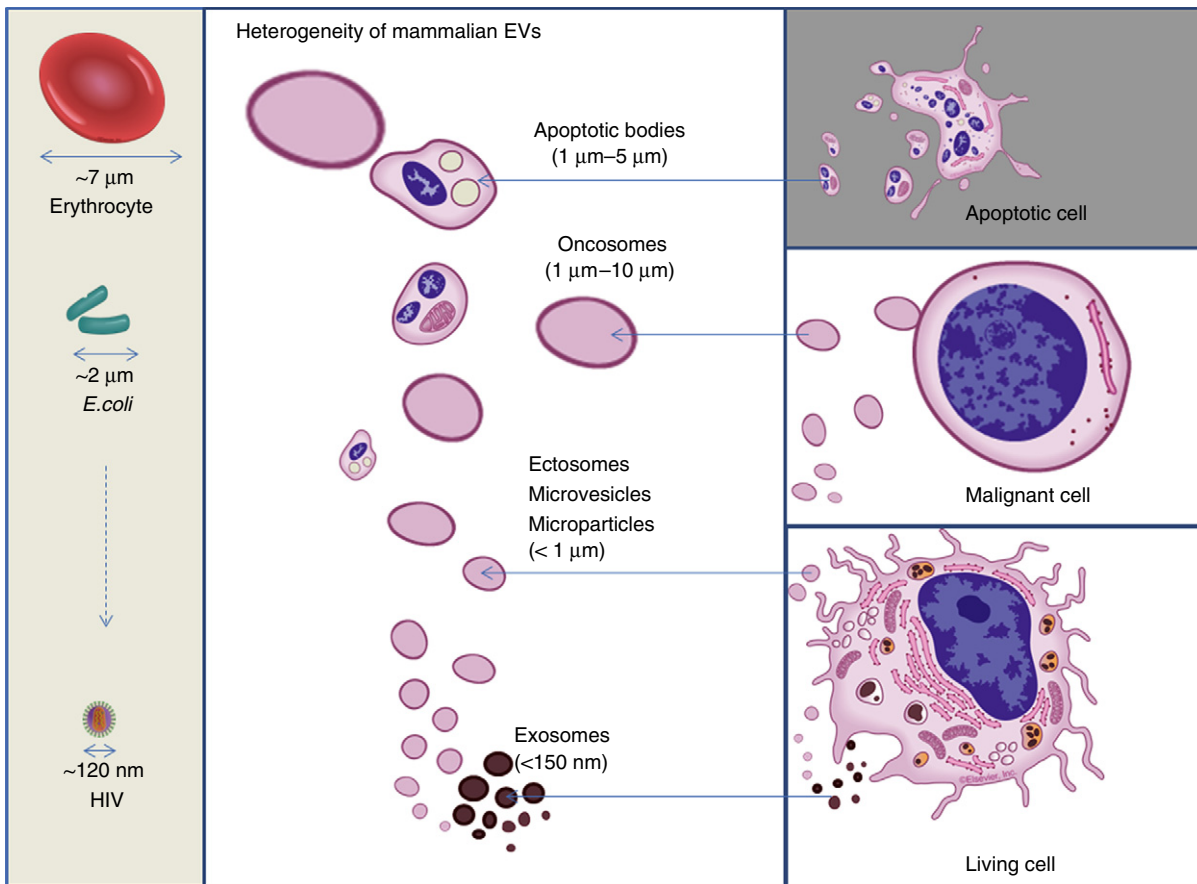


Figure 2 Heterogeneity of mammalian cell-derived EVs. Cells undergoing apoptosis release large EVs, apoptotic bodies. Also living cancer cells can produce EVs in a similar size range as apoptotic bodies. However, the majority of living cells produces much smaller EVs. In the size range of <200 nm it is difficult to discriminate between plasma membrane-derived EVs and EVs derived from the endosomal system, i.e., exosomes, because of the lack of unique markers to discriminate between these EV subsets.

mobilization of MVBs to and fusion with the plasma membrane is controlled by the RAB family of small GTPase proteins. For the release of exosomes RAB27A, RAB11, and RAB35 have been implicated (Kowal *et al.*, 2014). Little is known about the machinery that finally drives MVB fusion with the plasma membrane or direct EV budding from the plasma membrane, although increasing intracellular Ca^{2+} concentrations seem to be a central trigger (Raposo and Stoorvogel, 2013). Also for EVs budding from the plasma membrane, phospholipids and proteins form microdomains at the plasma membrane similar to lipid rafts. Budding of these plasma membrane microdomains, containing associated protein/RNA packages, requires local disassembly of the cytoskeleton. Translocation of phosphatidyl-serine to the outer leaflet can trigger this membrane budding and subsequent EV formation. In the formation of plasma membrane-derived EVs specific factors such as ARF-6 (GTP-binding protein ADP-ribosylation factor 6) interacting with cytoskeletal structures, TSG101, and the VPS4 ATPase are proposed to play a role (Raposo and Stoorvogel, 2013).

Although the formation of EVs by Gram-negative bacteria is already known for almost 50 years, vesiculation in bacteria is not well understood. EVs from Gram-negative bacteria are

spherical lipid bilayer-enclosed vesicles ranging in size from ~20 to 300 nm. Because these EVs consist mainly of typical outer membrane components of Gram-negative bacteria, for example, lipopolysaccharide (LPS), it is assumed that these EVs pinch off from the outer membrane and hence are often called outer membrane vesicles (OMVs). Since also cytoplasm-derived proteins consistently end up in these vesicles and certain proteins are enriched or excluded in these EVs, an active sorting mechanism during OMV biogenesis has been suggested (Kulp and Kuehn, 2010). Aside from lipids and proteins, RNA and DNA are also associated with bacterial EVs. There is growing evidence that also Gram-positive bacteria and *Archaea* produce EVs but their biogenesis is still obscure. Based on the differences in cell wall architecture between Gram-negative and Gram-positive bacteria it is unlikely that the biogenesis pathways of EVs are comparable (Deatherage and Cookson, 2012). Although taxonomically distant, EV release by fungi, parasites, and *Archaea* share some basic features and the EV biogenesis in *Archaea* shares similarities with both prokaryotic and eukaryotic mechanisms, for example, the involvement of ESCRT-III and VPS4 homologs (Deatherage and Cookson, 2012). Eukaryotic microbes, including fungi and parasites, can release EVs by direct budding from the cell

surface, as well as from the MVB route. These EVs contain microbial components (proteins, lipids, and nucleic acids) and many associated proteins lack the classical secretory signal sequences. However, many aspects of EV biogenesis of intracellular parasites are still elusive. For example, the biogenesis of EVs in malaria parasite-infected red blood cells which lack internal membranes and the machinery for exocytosis and endocytosis. Analysis of EVs released by infected red blood cells indicated that these EVs are enriched for host ARF-6 and VPS4 proteins, suggesting host machinery involved in EV biogenesis. Furthermore, these EVs contain parasite proteins derived from Maurer's Clefts structures; parasite-induced membrane platforms anchored to the host cell cytoskeleton involved in protein trafficking from the parasite to the red blood cell surface (Mantel and Marti, 2014). Also *Leishmania* parasite-infected macrophages release EVs with mammalian exosomal proteins, as well as *Leishmania* orthologues to ESCRT-III components. It is currently not clear how this parasite triggers EV formation, and how these EVs cross the phagolysosome and are eventually secreted in the environment (Mantel and Marti, 2014). EVs derived from the extracellular parasite *Trichomonas vaginalis* also show a high similarity with mammalian EVs including components of the ESCRT machinery, tetraspanin homologs and small RNA species of as yet unknown function (Mantel and Marti, 2014). The presence of EV-associated ESCRT components in different eukaryotic lineages and in *Archaea* supports the idea that EVs are an ancient mechanism for intercellular communication. The plasma membrane is in general the final barrier for EV release. Fungal and plant cells, however, are surrounded by a complex and dynamic cell wall which needs to be crossed before EVs are released in the environment. Despite this barrier, EVs have been isolated from culture supernatants indicating that transport across the cell wall occurs in fungi and plants (Rodrigues *et al.*, 2014; Regente *et al.*, 2012). The knowledge on the biogenesis of fungal and plant-derived EVs destined for release into the environment is still in its infancy. In fungal EVs, glycolipid components of the fungal cell wall are identified and many EV-associated proteins lack the characteristic signal peptide required for conventional protein secretion (Rodrigues *et al.*, 2014). Also in apoplast fractions, for example, the extracellular compartment of plant cells, EVs have been identified that contain leaderless proteins, as well as a homolog (68% identity) of the human Rab11a GTPase (Regente *et al.*, 2012).

EVs in Disease

Besides a role in normal physiology, EVs are also involved in many pathological processes. During bacterial and parasite infections, comprising both helminths and parasitic protozoa, pathogen-derived EVs are released and can influence host cells in many different ways to evade host defense mechanisms and to establish and propagate infection (Marcilla *et al.*, 2012; Twu and Johnson, 2014). Invasive intracellular pathogens (e.g., *Mycobacteria*, *Leishmania* sp., and *Plasmodium* sp.) can also influence the release and content of infected host cell-derived EVs. These altered host cell-EVs have been implicated in inflammatory responses and in the modulation of immune

responses to the pathogen (Mantel and Marti, 2014; Twu and Johnson, 2014). Also during viral infection with herpes viruses or retroviruses, amongst others, virus-infected cells have an altered EV production profile in which EVs can contain both viral and cellular components. This heterogeneous population of EVs is involved in infection, either by enhancing or blocking of infection, or cause dysregulation of the immune response (Wurdinger *et al.*, 2012). Furthermore, viruses can recruit several components from the EV biogenesis routes and in this way viruses exploit the EV machinery for viral spread (Gould *et al.*, 2003).

In cancer there is accumulating evidence that tumor cell-derived EVs play a role in tumorigenesis, metastasis, and chemotherapeutic resistance (Tickner *et al.*, 2014). During tumorigenesis tumor-derived EVs deregulate antitumor immune responses by impeding both adaptive and innate immune responses (Zhang and Grizzle, 2011). Furthermore, these EVs modulate and restructure the local tumor environment and travel to distant sites to promote the formation of premetastatic niches (Psaila and Lyden, 2009). Tumor-derived EVs have also been implicated in the transfer of oncogenic activity (Al-Nedawi *et al.*, 2008) and in chemotherapeutic resistance (Al-Nedawi *et al.*, 2008). Also in many chronic inflammatory processes such as diabetes, rheumatic diseases, inflammatory bowel diseases, and allergic diseases, as well as in (cardio)vascular disease and neurodegenerative disorders, the role of EVs becomes more and more evident (EL Andaloussi *et al.*, 2013).

Clinical Applications of EVs

EV-Based Biomarkers

During the last 4–5 years it has become evident that body fluids (e.g., blood, urine, cerebrospinal fluid, and milk) contain large amounts of nano-sized (<200 nm) EVs originating from a variety of cell types. Since cells concurrently release different types of EVs and the molecular composition of EVs depends on the state of the releasing cell, EVs in body fluids are not only very abundant and heterogeneous but they also reflect the state of the cell of origin at the time of release. Consequently, during disease processes EVs in body fluids can contain important disease-related information. Indeed, tumor-derived EV signatures, based on microRNA, protein, or lipid content, have been defined in body fluids of cancer patients. In cancer patients where biopsy taking is not straightforward, the potential of EVs as so-called liquid biopsies is now widely investigated (Revenfeld *et al.*, 2014). In other diseases the detection of particular EVs has been associated with the onset of disease. For example, the excessive release of placental vesicles into the blood during pregnancy as an underlying cause of vascular damage and inflammation leading to the development of preeclampsia (Mincheva-Nilsson and Baranov, 2014). In general there is now broad interest in the identification and comprehensive characterization (cellular origin, composition, and concentration) of disease-associated EVs in body fluids (Figure 3).

Currently, clinical trials are under way in which the presence of coagulation-promoting tumor-derived EVs in the

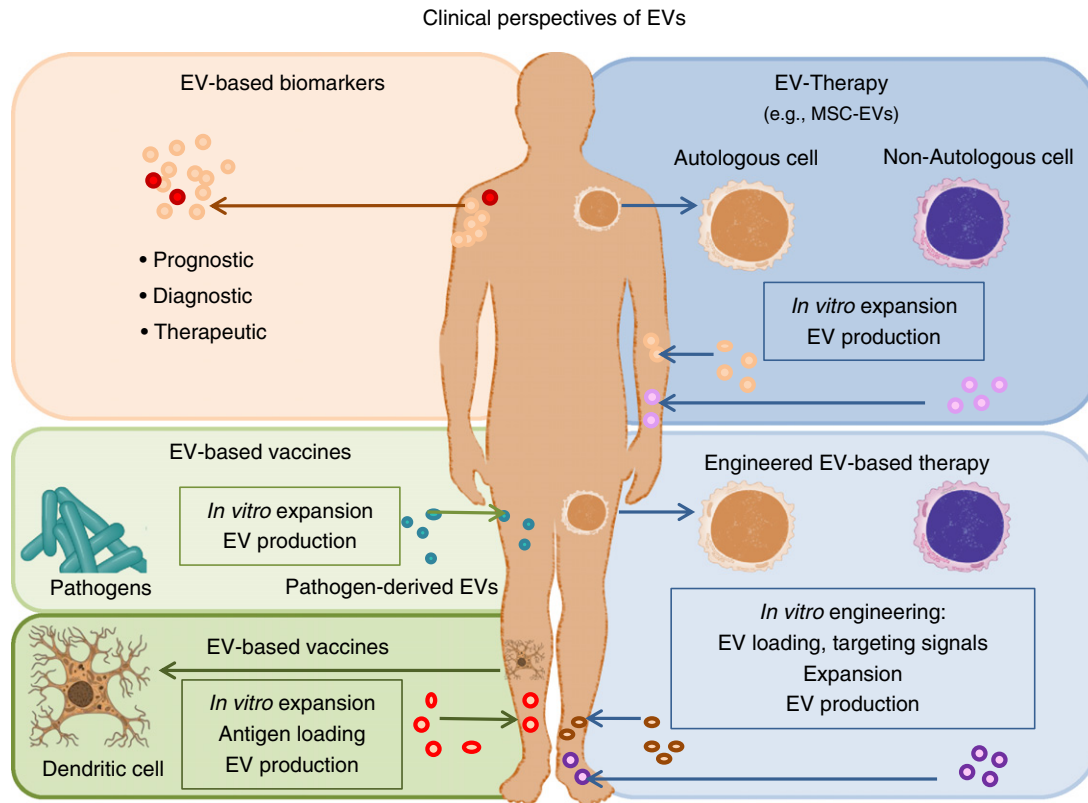


Figure 3 Clinical perspectives of EVs. Since EVs are present in body fluids and the EV population changes during different (patho)physiological conditions, EVs offer a great potential as multifactorial biomarkers. EVs are also in the limelight as therapeutics, for example, MSC-derived EVs in the treatment of graft-versus-host disease patients. The special features of EVs are also exploited to engineer EV-based nano-carriers for drug delivery. For vaccine development EVs from pathogens and EVs from dendritic cells loaded with antigens derived from pathogens or tumors are currently studied.

blood of cancer patients is studied as a potential risk score marker to identify cancer patients at risk of developing venous thromboembolism (Zwicker *et al.*, 2013). Also for infectious diseases the analysis of pathogen-derived or infected cell-derived EVs in blood offers possibilities to monitor the infection status. Besides the disease-specific signature, EV-based biomarkers also have a great advantage in dealing with the enormous dynamic range of molecules present in body fluids to detect the frequently low-abundant diagnostic molecules. Altogether, EV-based biomarkers hold a great promise as prognostic, diagnostic, and therapeutic biomarkers in a wide variety of diseases. However, the process of identification of EV-based biomarkers is very arduous. A major limitation in this emerging field is the complicated isolation of EVs from complex body fluids such as blood which also contains high concentrations of lipoprotein particles with overlapping physical characteristics. Currently, there is a strong need for specialized and standardized isolation methods (Witwer *et al.*, 2013a). Furthermore, specific disease-associated EVs, such as tumor-derived EVs, are relatively rare events in the total EV population in blood in which the majority of EVs is derived from blood cells and vessels. Consequently, high throughput EV analysis is needed to identify subpopulations of disease-associated EVs. Unfortunately, most commonly used techniques either lack the sensitivity and specificity for individual EV-based analysis (e.g., conventional flow cytometry with a

detection limit of ~ 300 nm) or are low-throughput (e.g., electron microscopy and atomic force microscopy) (Momen-Heravi *et al.*, 2012). However, the recent achievements in high-resolution flow cytometry-based EV analysis are promising (Van der Vlist *et al.*, 2012). The great challenge for the coming years is to standardize protocols for EV measurements in body fluids and to develop translational technologies for robust analysis methods in clinical settings.

EV-Based Vaccines and Therapeutics

EVs derived from Gram-negative bacteria, i.e., OMVs, have been used in vaccine development since the late 1980s, and they are considered a safe and effective prevention method against *Neisseria meningitidis* infections (Figure 3). However, widespread application of OMV-based vaccines is hampered by several problems such as scale of production, lack of broad immunogenicity, and toxicity due to the presence of LPS (Baker *et al.*, 2014). However, the possibility to engineer these EVs to optimize their immunogenicity via incorporation of selected sets of bacterial antigens and to lower the toxicity and optimize the self-adjunct properties by LPS modifications makes them interesting candidates for vaccine development (Baker *et al.*, 2014). The knowledge of Gram-positive bacteria-derived EVs is still limited and their application as vaccine

remains to be elucidated. Besides prokaryotic cell-derived EVs also eukaryotic cell-derived EVs are studied for vaccine development (Figure 3). In 1998, the immune stimulatory potential of dendritic cell-derived EVs loaded with tumor antigens was revealed (Zitvogel *et al.*, 1998), and resulted in the first clinical trials in cancer patients in which the potential of these EVs as cancer vaccine was investigated (Escudier *et al.*, 2005). Nowadays also the vaccine potential of parasite-derived EVs or EVs from parasite-infected cells is explored (Martin-Jaular *et al.*, 2011).

EVs are also studied as therapeutic entity in regenerative medicine. Based on the regenerative capacity, mesenchymal stem cell (MSC)-derived EVs hold a great promise for repairing, replacing, restoring, or regenerating damaged cells, tissues, and organs (EL Andaloussi *et al.*, 2013). Furthermore, the immune suppressive properties of MSC-derived EVs may be used to dampen the immune system (e.g., during Graft-versus-host disease) (Figure 3). The application of EVs can resolve major safety issues related to uncontrollable dissemination and differentiation of transplanted cells, and since EVs are relative stable and easy to store they can overcome limitations associated with the use of viable cells. A current challenge for therapeutic application of EVs is to develop scalable manufacturing processes while maintaining the critical quality parameters (functionality, purity, and safety) of the final therapeutic product and to establish robust quality control criteria.

EVs also receive attention as nano-carriers for therapeutic entities because of their bioavailability, their low immunogenicity if derived from the appropriate cells (e.g., MSCs or immature dendritic cells), their specific targeting potential, and their resistance to metabolic processes (EL Andaloussi *et al.*, 2013; Johnsen *et al.*, 2014; Figure 3). The fact that EVs can cross classical borders, such as the blood-brain barrier, and can deliver small RNAs to regulate gene expression in targeted cells demonstrated the potential of targeted delivery via EV-based nano-carriers (Alvarez-Erviti *et al.*, 2011). In a rapidly increasing number of studies there is evidence that the loading of siRNAs and microRNAs in EV-based delivery systems prevent these RNA molecules from rapid degradation in the systemic circulation and results in functional transfer to the preferred target cells. However, in most studies classical cell transfection strategies-like electroporation and chemical transfection have been used to load the EV-based carriers. Besides the questionable efficacy of RNA loading, these methods can induce adverse effects on EV integrity or therapeutic cargo (Johnsen *et al.*, 2014). The alternative approach in which EV-producing cells are transfected with the therapeutic RNA is generally more effective in EV-RNA loading, but the cell engineering required to express a selected targeting protein on the EV surface of the RNA-loaded EVs is more laborious (Johnsen *et al.*, 2014). Also for the clinical application in which patient's own cells are used as EV source this is problematic. To optimize and secure reproducibility of EV-RNA loading efficiency, more insight in the physiological RNA loading mechanism of EVs is essential. Furthermore the development of non-autologous EV factory cell lines generating loaded EVs expressing the proper target moiety is of great interest. Besides the loading of RNA molecules, other types of therapeutic cargo have been loaded in EVs,

such as doxorubicin for the development of chemotherapeutic-EVs for cancer treatment. The *in vivo* targeted delivery of engineered EVs derived from dendritic cells greatly enhanced the efficacy of doxorubicin (Tian *et al.*, 2014). Although promising results are obtained in preclinical studies using EV-based vectors derived from dendritic cells and MSCs, many properties remain elusive. The fact that contradictory results can be obtained with EVs and EV-based vectors from the same cell type urges for prudence. The next steps in the development of EV-based therapeutics comprise the selection of optimal; (1) cell types as EV factory and culture conditions, (2) loading strategies for the desired therapeutic, (3) specific targeting moieties, (4) standardized EV isolation methods, and (5) full characterization of therapeutic EVs. Furthermore, up-scaling procedures of EV production and development of control procedures are needed. Although it is expected that EV-based therapeutics will evolve as a new field for clinical application, the hurdles to be overcome are not trivial.

Concluding Remarks

The increasing interest in EVs led in 2012 to the creation of a scientific society fully dedicated to the subject: the International Society of Extracellular Vesicles. Convincing evidence is present that EVs are an evolutionary conserved mechanism of intercellular communication allowing the exchange of complex information within and beyond organisms. Furthermore EVs hold a great promise as biomarker and therapeutic agent.

See also: Cell Division/Death: Apoptosis: Apoptosis. Interorganellar Communication: Components: ESCRTing around the Cell. Interorganellar Communication: Interplay and Processes: Endosome to Lysosome Transport; Rabs of the Endosomal Recycling Pathway; Unconventional Protein Secretion: Fibroblast Growth Factor 2 and Interleukin-1 β as Examples. Molecular Principles, Components, Technology, and Concepts: Membranes: Lipid Rafts/Membrane Rafts. Nucleic Acid Synthesis/Breakdown: RNA Synthesis/Function: miRNAs/Small Noncoding RNAs. Organelles: Structure and Function: Conventional and Secretory Lysosomes; The Late Endosome

References

- Al-Nedawi, K., Meehan, B., Micallef, J., *et al.*, 2008. Intercellular transfer of the oncogenic receptor EGFRvIII by microvesicles derived from tumour cells. *Nature Cell Biology* 10 (5), 619–624.
- Alvarez-Erviti, L., Seow, Y., Yin, H., *et al.*, 2011. Delivery of siRNA to the mouse brain by systemic injection of targeted exosomes. *Nature Biotechnology* 29 (4), 341–345.
- Arck, P.C., Hecher, K., 2013. Fetomaternal immune cross-talk and its consequences for maternal and offspring's health. *Nature Medicine* 19 (5), 548–556.
- Baker, J.L., Chen, L., Rosenthal, J.A., *et al.*, 2014. Microbial biosynthesis of designer outer membrane vesicles. *Current Opinion in Biotechnology* 29C, 76–84.
- Biller, S.J., Schubotz, F., Roggensack, S.E., *et al.*, 2014. Bacterial vesicles in marine ecosystems. *Science* 343 (6167), 183–186.

- Corrigan, L., Redhai, S., Leiblich, A., *et al.*, 2014. BMP-regulated exosomes from *Drosophila* male reproductive glands reprogram female behavior. *Journal of Cell Biology* 206 (5), 671–688.
- Deatherage, B.L., Cookson, B.T., 2012. Membrane vesicle release in bacteria, eukaryotes, and archaea: A conserved yet underappreciated aspect of microbial life. *Infection and Immunity* 80 (6), 1948–1957.
- Delaunais, B., Jeandet, P., Clément, C., *et al.*, 2014. Uncovering plant-pathogen crosstalk through apoptotic proteomic studies. *Frontiers in Plant Science* 5, 249.
- El Andaloussi, S., Mager, I., Breakefield, X.O., Wood, M.J., 2013. Extracellular vesicles: Biology and emerging therapeutic opportunities. *Nature Reviews Drug Discovery* 12 (5), 347–357.
- Escudier, B., Dorval, T., Chaput, N., *et al.*, 2005. Vaccination of metastatic melanoma patients with autologous dendritic cell (DC) derived-exosomes: Results of the first phase I clinical trial. *Journal of Translational Medicine* 3 (1), 10.
- Ferrand, J., Ferrero, R.L., 2013. Recognition of extracellular bacteria by NLRs and its role in the development of adaptive immunity. *Frontiers in Immunology* 4, 344.
- Frühbeis, C., Fröhlich, D., Kuo, W.P., Krämer-Albers, E.-M., 2013. Extracellular vesicles as mediators of neuron-glia communication. *Frontiers in Cellular Neuroscience* 7, 182.
- Golub, E.E., 2009. Role of matrix vesicles in biomineralization. *Biochimica et Biophysica Acta* 1790 (12), 1592–1598.
- Gould, S.J., Booth, A.M., Hildreth, J.E.K., 2003. The Trojan exosome hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* 100 (19), 10592–10597.
- Harding, C., Heuser, J., Stahl, P., 1983. Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes. *Journal of Cell Biology* 97 (2), 329–339.
- Hu, G., Gong, A.-Y., Roth, A.L., *et al.*, 2013. Release of luminal exosomes contributes to TLR4-mediated epithelial antimicrobial defense. *PLoS Pathogens* 9 (4), e1003261.
- Johnsen, K.B., Gudbergsson, J.M., Skov, M.N., *et al.*, 2014. A comprehensive overview of exosomes as drug delivery vehicles — Endogenous nanocarriers for targeted cancer therapy. *Biochimica et Biophysica Acta* 1846 (1), 75–87.
- Kaffarnik, F.A.R., Jones, A.M.E., Rathjen, J.P., Peck, S.C., 2009. Effector proteins of the bacterial pathogen *Pseudomonas syringae* alter the extracellular proteome of the host plant, *Arabidopsis thaliana*. *Molecular & Cellular Proteomics: MCP* 8 (1), 145–156.
- Kowal, J., Tkach, M., Théry, C., 2014. Biogenesis and secretion of exosomes. *Current Opinion in Cell Biology* 29C, 116–125.
- Kulp, A., Kuehn, M.J., 2010. Biological functions and biogenesis of secreted bacterial outer membrane vesicles. *Annual Review of Microbiology* 64, 163–184.
- Mantel, P.-Y., Marti, M., 2014. The role of extracellular vesicles in Plasmodium and other protozoan parasites. *Cellular Microbiology* 16 (3), 344–354.
- Marcella, A., Trelis, M., Cortés, A., *et al.*, 2012. Extracellular vesicles from parasitic helminths contain specific excretory/secretory proteins and are internalized in intestinal host cells. *PLoS One* 7 (9), e45974.
- Martin-Jaular, L., Nakayasu, E.S., Ferrer, M., *et al.*, 2011. Exosomes from *Plasmodium yoelii*-infected reticulocytes protect mice from lethal infections. *PLoS One* 6 (10), e26588.
- Mincheva-Nilsson, L., Baranov, V., 2014. Placenta-derived exosomes and syncytiotrophoblast microparticles and their role in human reproduction: Immune modulation for pregnancy success. *American Journal of Reproductive Immunology* 72 (5), 440–457.
- Momen-Heravi, F., Balaj, L., Alian, S., *et al.*, 2012. Alternative methods for characterization of extracellular vesicles. *Frontiers in Physiology* 3, 354.
- Morello, M., Minciaccchi, V.R., de Candia, P., *et al.*, 2013. Large oncosomes mediate intercellular transfer of functional microRNA. *Cell Cycle (Georgetown, Tex.)* 12 (22), 3526–3536.
- Owens, A.P., Mackman, N., 2011. Microparticles in hemostasis and thrombosis. *Circulation Research* 108 (10), 1284–1297.
- Pan, B.T., Teng, K., Wu, C., Adam, M., Johnstone, R.M., 1985. Electron microscopic evidence for externalization of the transferrin receptor in vesicular form in sheep reticulocytes. *Journal of Cell Biology* 101 (3), 942–948.
- Psaila, B., Lyden, D., 2009. The metastatic niche: Adapting the foreign soil. *Nature Reviews. Cancer* 9 (4), 285–293.
- Raposo, G., Nijman, H.W., Stoorvogel, W., *et al.*, 1996. B lymphocytes secrete antigen-presenting vesicles. *Journal of Experimental Medicine* 183 (3), 1161–1172.
- Raposo, G., Stoorvogel, W., 2013. Extracellular vesicles: Exosomes, microvesicles, and friends. *Journal of Cell Biology* 200 (4), 373–383.
- Ratajczak, J., Wysoczynski, M., Hayek, F., Janowska-Wieczorek, A., Ratajczak, M.Z., 2006. Membrane-derived microvesicles: Important and underappreciated mediators of cell-to-cell communication. *Leukemia* 20 (9), 1487–1495.
- Regente, M., Pinedo, M., Elizalde, M., de la Canal, L., 2012. Apoptotic exosome-like vesicles: A new way of protein secretion in plants? *Plant Signaling & Behavior* 7 (5), 544–546.
- Regev-Rudzi, N., Wilson, D.W., Carvalho, T.G., *et al.*, 2013. Cell-cell communication between malaria-infected red blood cells via exosome-like vesicles. *Cell* 153 (5), 1120–1133.
- Revenfeld, A.L.S., Bæk, R., Nielsen, M.H., *et al.*, 2014. Diagnostic and prognostic potential of extracellular vesicles in peripheral blood. *Clinical Therapeutics* 36 (6), 830–846.
- Robbins, P.D., Morelli, A.E., 2014. Regulation of immune responses by extracellular vesicles. *Nature Reviews Immunology* 14 (3), 195–208.
- Rodrigues, M.L., Nakayasu, E.S., Almeida, I.C., Nimrichter, L., 2014. The impact of proteomics on the understanding of functions and biogenesis of fungal extracellular vesicles. *Journal of Proteomics* 97, 177–186.
- Ronquist, G., 2012. Prostatomes are mediators of intercellular communication: From basic research to clinical implications. *Journal of Internal Medicine* 271 (4), 400–413.
- Shen, Y., Giardino Torchia, M.L., Lawson, G.W., *et al.*, 2012. Outer membrane vesicles of a human commensal mediate immune regulation and disease protection. *Cell Host & Microbe* 12 (4), 509–520.
- Tetta, C., Bruno, S., Fonsato, V., Deregibus, M.C., Camussi, G., 2011. The role of microvesicles in tissue repair. *Organogenesis* 7 (2), 105–115.
- Tian, Y., Li, S., Song, J., *et al.*, 2014. A doxorubicin delivery platform using engineered natural membrane vesicle exosomes for targeted tumor therapy. *Biomaterials* 35 (7), 2383–2390.
- Tickner, J.A., Urquhart, A.J., Stephenson, S.-A., Richard, D.J., O'Byrne, K.J., 2014. Functions and therapeutic roles of exosomes in cancer. *Frontiers in Oncology* 4, 127.
- Twu, O., Johnson, P.J., 2014. Parasite extracellular vesicles: Mediators of intercellular communication. *PLoS Pathogens* 10 (8), e1004289.
- Valadi, H., Ekström, K., Bossios, A., *et al.*, 2007. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nature Cell Biology* 9 (6), 654–659.
- Villarroya-Beltri, C., Baixauli, F., Gutiérrez-Vázquez, C., Sánchez-Madrid, F., Mittelbrunn, M., 2014. Sorting it out: Regulation of exosome loading. *Seminars in Cancer Biology* 28C, 3–13.
- Villarroya-Beltri, C., Gutiérrez-Vázquez, C., Sánchez-Cabo, F., *et al.*, 2013. Sumoylated hnRNP A2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs. *Nature Communications* 4, 2980.
- Van der Vlist, E.J., Nolte-t Hoen, E.N., Stoorvogel, W., Arksteijn, G.J., Wauben, M.H., 2012. Fluorescent labeling of nano-sized vesicles released by cells and subsequent quantitative and qualitative analysis by high-resolution flow cytometry. *Nature Protocols* 7 (7), 1311–1326.
- Witwer, K.W., Buzás, E.I., Bemis, L.T., *et al.*, 2013a. Standardization of sample collection, isolation and analysis methods in extracellular vesicle research. *Journal of Extracellular Vesicles*. 2.
- Witwer, K.W., McAlexander, M.A., Queen, S.E., Adams, R.J., 2013b. Real-time quantitative PCR and droplet digital PCR for plant miRNAs in mammalian blood provide little evidence for general uptake of dietary miRNAs: Limited evidence for general uptake of dietary plant xenomiRs. *RNA Biology* 10 (7), 1080–1086.
- Wurdinger, T., Gatsen, N.N., Balaj, L., *et al.*, 2012. Extracellular vesicles and their convergence with viral pathways. *Advances in Virology* 2012, 767694.
- Zhang, H.G., Grizzle, W.E., 2011. Exosomes and cancer: A newly described pathway of immune suppression. *Clinical Cancer Research* 17 (5), 959–964.
- Zhang, L., Hou, D., Chen, X., *et al.*, 2012. Exogenous plant MIR168a specifically targets mammalian LDLRAP1: Evidence of cross-kingdom regulation by microRNA. *Cell Research* 22 (1), 107–126.
- Zhang, L., Wana, J.L., 2014. The emerging role of exosomes in Wnt secretion and transport. *Current Opinion in Genetics & Development* 27, 14–19.
- Zitvogel, L., Regnault, A., Lozier, A., *et al.*, 1998. Eradication of established murine tumors using a novel cell-free vaccine: Dendritic cell-derived exosomes. *Nature Medicine* 4 (5), 594–600.
- Zwicker, J.I., Liebman, H.A., Bauer, K.A., *et al.*, 2013. Prediction and prevention of thromboembolic events with enoxaparin in cancer patients with elevated tissue factor-bearing microparticles: A randomized-controlled phase II trial (the Microtec study). *British Journal of Haematology* 160 (4), 530–537.

Further Reading

- Deatherage, B.L., Cookson, B.T., 2012. Membrane vesicle release in bacteria, eukaryotes, and archaea: A conserved yet underappreciated aspect of microbial life. *Infection and Immunity* 80 (6), 1948–1957.
- EL Andaloussi, S., Mager, I., Breakefield, X.O., Wood, M.J., 2013. Extracellular vesicles: Biology and emerging therapeutic opportunities. *Nature Reviews Drug Discovery* 12 (5), 347–357.
- Kowal, J., Tkach, M., Théry, C., 2014. Biogenesis and secretion of exosomes. *Current Opinion in Cell Biology* 29C, 116–125.
- Mantel, P.-Y., Marti, M., 2014. The role of extracellular vesicles in *Plasmodium* and other protozoan parasites. *Cellular Microbiology* 16 (3), 344–354.
- Robbins, P.D., Morelli, A.E., 2014. Regulation of immune responses by extracellular vesicles. *Nature Reviews Immunology* 14 (3), 195–208.

Relevant Websites

- <http://www.evpedia.info/>
EVpedia.
- <http://www.isev.org/>
International Society for Extracellular Vesicles (ISEV).
- <http://microvesicles.org>
Vesiclepedia.

Nuclear Organization (Nuclear Structure and Dynamics)

Y Mimura and N Imamoto, RIKEN, Saitama, Japan

© 2016 Elsevier Inc. All rights reserved.

Introduction

In eukaryotic cells, unlike prokaryotic cells that contain naked DNA in their cytoplasm, chromosomes composed primarily of DNA and histones are wrapped with double lipid bilayers and are thus separated from the cytoplasm. This structure containing the genetic information is called the nucleus, and the double lipid bilayer wrapping the chromosomes is called the nuclear envelope (NE). The nucleus, in which essential cellular processes such as replication, transcription, and ribosome biogenesis are constantly occurring, is essential for basic cellular activities and for maintaining cellular homeostasis.

The NE is mainly composed of four components: the nuclear membrane, the nuclear pore complex (NPC), the nuclear lamina, and inner nuclear membrane (INM) proteins. The NE separates the nuclear contents from the cytoplasmic activities, which necessitates their mutual communication through the continuous exchange of molecules. The NPC, embedded in the NE, mediates nucleocytoplasmic molecular transport. The NE also plays a role in protecting genomic DNA from toxic metabolites produced in the cytoplasm, such as reactive oxygen species produced in mitochondria. Along with these functions, the NE is considered to regulate gene expression, DNA replication, and chromatin organization.

The nuclear interior is a well-ordered structure in which individual chromosomes are nonrandomly arranged into the particular nuclear space, termed as the chromosomal territory. Although the mechanism for the establishment of the chromosome territory remains elusive, the interaction between the chromosome and the NE is considered to play an essential role in the establishment of the chromosome territory. In addition, specialized domain structures, such as the nucleolus, Cajal body, and promyelocytic-leukemia nuclear body (PML-NB) are formed in the nuclear interior and dynamically interact with each other to coordinate nuclear functions.

Finally, it is also important to note that the nucleus is a dynamic structure. During the cell cycle, DNA in chromosomes must be replicated, and the replicated chromosomes must be precisely separated to transmit genomic information into daughter cells. Chromosome separation occurs during mitosis. In multicellular organisms, the nuclear structure is broken down at the mitotic entry. The NE–chromosome interaction is lost, and the NE is retracted into the endoplasmic reticulum (ER). The chromosomes are highly condensed, and a kinetochore is assembled on the centromere, which is a specific region in the chromosome. Centrosomes are duplicated in interphase, and mitotic microtubules form connections between the centrosome and kinetochore. Together with spatially and temporally well-orchestrated mitotic activities, such as kinases, phosphatases, nucleocytoplasmic transport machineries, and motor proteins, the chromosomes are segregated with the sophisticated machinery termed the mitotic spindle into the daughter cells. In late mitosis, the NE that was retracted into the ER is reassembled on the surface of

daughter chromosomes with the aid of protein–protein and protein–DNA interactions.

In this section, we briefly describe chromatin organization, the structure and function of the NE, and the mitotic dynamics of the nucleus.

Nuclear Structure and Function

Chromatin

All of the cells that compose our body contain the same base sequences of DNA in their nuclei. In humans, out of the entire genome, approximately 20 000 genes encode proteins. To generate the various organs such as brain or muscle from the fertilized egg that is a large single cell, gene expression must be regulated spatially and temporally through the developmental processes to build up specific gene expression patterns. To this end, cells potentially have many highly ordered and hierarchical mechanisms generating the diverse tissues via the structural alterations of the chromatin through histone exchange, posttranslational modifications of histones, and DNA modifications.

Eukaryotic DNA, whose length is approximately 2 m in human, is wrapped around cores of histone octamers that are composed of histone H2A, H2B, H3, and H4 (Luger *et al.*, 1997). In this way, all eukaryotic DNA is highly compacted in the nucleus. Histone H2A exclusively forms a heterodimer with histone H2B, while Histone H3 forms a heterodimer with Histone H4. A core histone octamer is assembled from two pairs of these heterodimers. The approximately 150 base pairs of DNA wrapped the core histone octamer is called a nucleosome, and it is the minimum module of chromatin structure. The nucleosome is formed in a stepwise manner. First, a pair of histone H3–H4 dimers wraps the DNA, and this intermediate is called a ‘tetrasome.’ Next, a pair of histone H2A–H2B dimers is further incorporated into the tetrasome and consequently matured into a complete nucleosome structure. These nucleosome assembly steps require the aid of a specific set of histone chaperones (De Koning *et al.*, 2007; Kurumizaka *et al.*, 2013).

Core histones without histone H4 contain many variants that are structurally distinct. Depending on the several cellular contexts (i.e., developmental processes, transcription, and DNA damage response), the core histones are dynamically and appropriately exchanged for other histone variants via specific histone chaperone activity. As a result of alterations of nucleosome structure via histone exchange, the cell can flexibly respond to internal and external stimuli (Volle and Dalal, 2014).

In addition to histone replacement, another important regulatory mechanism of chromatin for the response to many cellular processes is posttranslational modifications described as ‘epigenetics.’ Histone proteins can be structurally separated into the tail and core regions, and both regions contain many

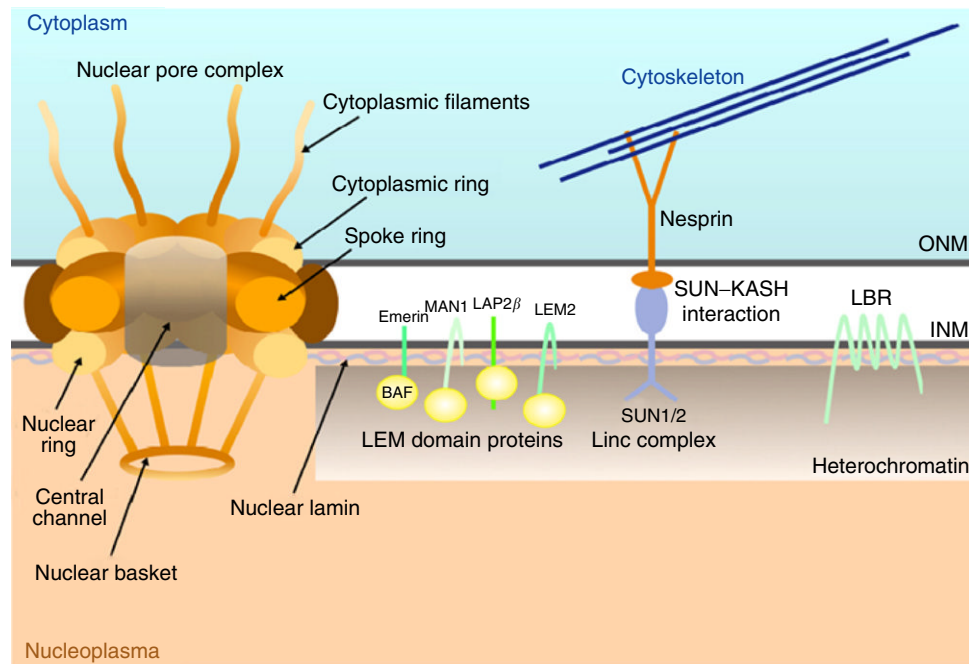


Figure 1 Structure of metazoan nuclear envelope (NE). Schematic representation of the overview of NE. The nuclear pore complex is embedded in the NE composed of the inner nuclear membrane (INM) and the outer nuclear membrane (ONM). INM proteins containing LEM proteins, SUN1/2 and Lamin B receptor (LBR) exclusively localized to the INM. SUN1/2 and Nesprin form linker of nucleoskeleton and cytoskeleton (LINC) complex via SUN–KASH interaction, and connect the nucleoskeleton to cytoskeleton through their interaction. Heterochromatin is accumulated underneath the INM via the interaction with INM proteins and the nuclear lamina.

sites that accept various posttranslational modifications, particularly phosphorylation, acetylation, and methylation. DNA, especially cytosine residues, is also methylated via DNA methyltransferase activities. Many of these modifications of histones are reversible, but it remains unknown how methylated cytosine is eliminated from DNA (Cantone and Fisher, 2013). Epigenetic regulation through the posttranslational modifications of histones and DNA are considered to affect gene expression by altering the interactions between histones, DNA, nucleosomes, and other proteins, including DNA-binding proteins. Accommodative functions with both processes, histone exchange and DNA/histone modifications, are required mechanisms for development, homeostasis and inheritance (Sasaki and Matsui, 2008). Conversely, misregulation of these mechanisms is closely associated with several critical diseases, such as cancer and leukemia, as well as with aging (Greer and Shi, 2012; Chi *et al.*, 2010).

NE

Nuclear membrane

The nuclear membrane is comprised of two phospholipid bilayers. The membrane facing the cytoplasm is termed the outer nuclear membrane (ONM), and the membrane facing the nucleoplasm is termed the INM. The ONM continuously connects to the ER, and its surface, like that of the ER, is decorated with ribosomes. The INM accumulates specific sets of proteins collectively termed INM proteins, and A-type and B-type lamins construct the nuclear lamina underneath the INM. Many INM

proteins, either directly or through association with lamins, interact with chromatin. The ONM is fused with the INM at sites termed pore membrane at which the NPC is present. The NE functions as the physical barrier that protects the nuclear interior from the cytoplasmic activities (Figure 1).

Nuclear lamina

In metazoans, the nuclear lamina, which is a meshwork structure, is composed of two types of lamin proteins termed A-type and B-type lamins. Both A-type and B-type lamins belong to the type V intermediate filaments (IFs), and the nuclear lamina confers physical strength to the NE. Interestingly, lamins are highly evolutionarily conserved in metazoans such as vertebrates but are not conserved in plants and yeast. Lamin A, Lamin C, Lamin C2, and Lamin AΔ10 belong to the A-type lamins and are encoded by the single *LMNA* gene (Burke and Stewart, 2013), while Lamin B1 and Lamin B2 belong to the B-type lamins and are encoded by separate genes, *LMNB1* and *LMNB2*, respectively. Lamin C is a C-terminus truncated form of Lamin A from an alternative splicing of the *LMNA* product. At least one form of B-type lamin is expressed in each cell, whereas the expression of A-type lamins is regulated in tissue-specific and development-dependent manners.

Both A-type and B-type lamins share the same structural features, including head and α -helical rod domains in the N-terminus. Although all tail regions in Lamin proteins contain the nuclear localization signal (NLS), their tail regions show structural differences. After translation, Lamin proteins mature through a series of posttranslational processes. The carboxy-terminals of Lamin A, Lamin B1, and Lamin B2, but

not Lamin C, contain the CAAX motif that is also conserved in the Ras superfamily (C: Cysteine, A: aliphatic amino acid, X: any amino acid). Initially, cysteine residues in these motifs of pre-Lamins are posttranslationally modified by farnesyl-transferase (FTase) that transfer the 15-carbon farnesyl lipid to cysteine, and the isoprenyl groups are transferred to its side chain (Roskoski, 2003). These posttranslational modifications require the presence of Zn^{2+} and Mg^{2+} ions. Following the isoprenylation, pre-Lamin A is cleaved at the carboxy-terminal end of the isoprenylcysteine residue in this motif by zinc metalloprotease ZMPSTE24, which is embedded in the NE/ER membrane via its 7-transmembrane regions and specifically recognizes the isoprenylcysteine residue (Quigley *et al.*, 2013; Pryor *et al.*, 2013). Meanwhile, pre-Lamin B is cleaved via farnesylated proteins-converting enzyme2 (FACE2) at the same site as pre-Lamin A. After removing the AAX tripeptide, isoprenylcysteine residues in the carboxyl terminal of a pre-Lamins becomes methylated via isoprenylcysteine carboxyl methyltransferase (ICMT) that also contain the multi-transmembrane regions and is embedded in the ER membrane (Winter-Vann and Casey, 2005). It is hypothesized that this isoprenyl group on Lamin B1 and B2 aids the Lamin B-phospholipid membrane attachment via a hydrophobic interaction. Only pre-Lamin A undergoes secondary digestion via ZMPSTE24 on the INM, leading to the removal of the 15 polypeptides in its carboxy terminal, including the isoprenylated cysteine. This series of processes produces mature Lamin A.

Both A-type and B-type lamins form homodimers as minimum modules for the construction of the nuclear lamina via α -helical coiled-coil regions within a rod domain. Similar to other IF family proteins, filament-like structures are formed by head-to-tail self-association of these homodimers both *in vivo* and *in vitro*. The precise overall structure and organization of the nuclear lamina is, however, an ongoing debate (Burke and Stewart, 2013). The nuclear lamina certainly provides the platform for a specific set of integral-membrane proteins, termed INM proteins, and specific DNA domains, termed lamin-associated domains (LAD).

Several human diseases associated with NE components, including both lamins and INM proteins, are collectively called 'laminopathies' (see also INM proteins section). Compared with *LMNB1* and *LMNB2*, a huge number of mutations in *LMNA* have been reported to be associated with laminopathies. Depending on the specific mutation, patients harboring mutations in *LMNA* associated with laminopathies show various symptoms, including cardiomyopathy, muscular dystrophy, lipodystrophy, mandibulofacial dysplasia, and progeria syndrome. One of the most famous and characterized *LMNA*-associated laminopathies is Hutchinson-Gilford progeria syndrome (HGPS) (Butin-Israeli *et al.*, 2012). Patients with HGPS show various clinical manifestations including premature aging, loss of subcutaneous fat, and atherosclerosis. Cells derived from HGPS patients express an abnormal Lamin A protein called progerin, which is an immature Lamin A protein containing an uncleaved farnesyl group in its C-terminus, due to a point mutation in the *LMNA* gene (C1828T). Although the progerin protein shows normal localization to the INM, cells from HGPS patients show nuclear blebbing and biological activity defects such as the loss of heterochromatin,

accumulation of DNA damage, and premature senescence. It is considered that these cellular defects are caused by the expression of 'farnesylated progerin' because FTase inhibitors can inhibit these cellular defects in HGPS patients.

NPC

The NPC, one of the largest protein complexes in the cell, has structurally and functionally well-conserved architecture in eukaryotic cells from yeast to mammals. The NPC is assembled from approximately 30 distinct proteins termed nucleoporins (Nups), and it has an eight-rotational asymmetric structure forming a tunnel between the cytoplasm and the nucleoplasm (Figure 2). The hollow space in the NPC structure is filled with many disordered/unstructured polypeptides containing hydrophobic amino acids, collectively called FG repeats (phenylalanine glycine repeats). This hollow is called the central channel, which is a site of the nucleocytoplasmic trafficking of molecules. Relatively small molecules of less than 50 kDa can freely pass through the NPC, while larger molecules above 50 kDa cannot do so. This size-dependent exclusion mechanism of the NPC is called a 'permeability barrier,' and the hydrophobicity in the central channel plays an essential role in the molecular transport that is a fundamental function of the NPC. In general, macromolecules containing specific amino-acid signal sequences, the NLS and the nuclear export signal (NES), can pass through the NPC in a nuclear transport receptor (NTR)-dependent manner.

Nups can be roughly categorized into four broad types by their structural roles: scaffold Nups, transmembrane Nups, central Nups, and peripheral Nups. Several Nups comprises subcomplexes throughout the cell cycle. For example, in mammalian cells, one of these subcomplexes, the NUP107-160 subcomplex, is composed of NUP107, NUP160, NUP133, NUP96, NUP75, NUP43, NUP37, Seh1, and Sec13, and the other, the NUP93-205 subcomplex, is composed of NUP93, NUP205, NUP188, NUP155, and NUP53. The NUP107-160 subcomplex has a Y-shaped structure, and it forms ring-like substructures at both the nucleoplasmic and cytoplasmic sides (termed the nucleoplasmic ring and the cytoplasmic ring, respectively). The NUP93-205 subcomplex is also assembled in a ring-like substructure, termed a spoke-ring complex, that adjoins the cytoplasmic and nucleoplasmic rings. The spoke-ring complex anchors the central Nups to form the central channel (Grossman *et al.*, 2012).

Transmembrane Nups, NDC1, POM121, and GP210 in mammalian cells, contain transmembrane regions and are embedded into the pore membrane. Because transmembrane Nups have an affinity for both NUP107-160 and NUP93-205 scaffold subcomplexes, this group of Nups is considered to anchor scaffold substructures to the pore membrane.

Central Nups, which consist of NUP54, NUP58, and NUP62, form a subcomplex and play a critical role in the formation of the central channel. These three Nups contain FG repeats and confer the relatively hydrophobic environment to the central channel. This hydrophobicity of the central channel plays the central role in the establishment of the permeability barrier and macromolecular transport.

Other Nups are categorized as peripheral Nups. This group of Nups is also associated with the establishment of the permeability barrier and molecular transport because of FG-repeat

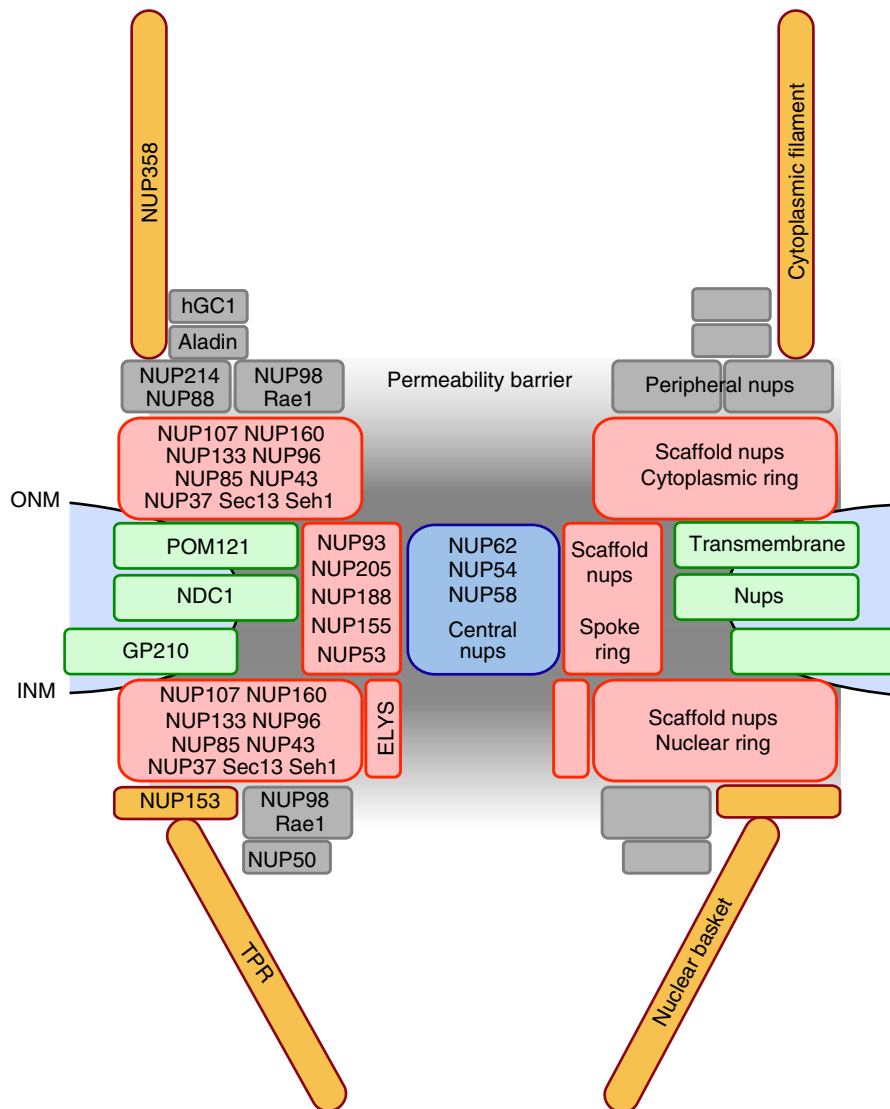


Figure 2 Metazoan nuclear pore complex (NPC) components. Schematic representation of the distribution of Nups within the NPC. Scaffold Nups and transmembrane Nups are indicated with red and green colors, respectively. Central Nups and peripheral Nups are indicated with blue and gray colors, respectively. The peripheral Nups that associate with cytoplasmic filament and nuclear basket are indicated with brown color. Note that this Nups distribution within the NPC is experimentally predicted model so far.

containing peripheral Nups. These Nups form two types of substructures on both the cytoplasmic and nucleoplasmic sides: cytoplasmic fibrils and nuclear baskets. Cytoplasmic fibrils are filamentous structures that are mainly composed of NUP358/RanBP2. Nuclear baskets are basket-like structures that are mainly composed of NUP153 and TPR. Both substructures participate in the regulation of transport receptor-dependent macromolecular exchange (Hoelz *et al.*, 2011).

In proliferating cells, NPCs are constantly assembling throughout the cell cycle. At the mitotic entry, NPCs are highly phosphorylated via several mitotic kinases, and they are consequently broken down into subcomplexes (see section NE disassembly and reassembly in Mitosis). After the onset of anaphase, Nups are dephosphorylated via protein phosphatases (PPases), and they then assemble into NPCs on the surface of daughter chromosomes. This process is termed

‘post-mitotic NPC assembly’ (Wandke and Kutay, 2013). By contrast, during interphase, NPCs are embedded in the growing and closed NE. For this reason, the INM and ONM must be fused via unknown mechanisms to form the pore membrane (Rothballer and Kutay, 2013).

INM proteins

INM proteins (also called as nuclear envelope transmembrane (NET) proteins) are defined as preferentially INM-localized proteins, and they contain several transmembrane regions. The INM proteins construct the nuclear lamina together with the nuclear lamins. The total number of INM proteins is reported as nearly 200 proteins, and the number of identified INM proteins continues to increase (Wilson and Berk, 2010; Berk *et al.*, 2013). Most INM proteins remain uncharacterized, and only a small number of INM proteins have been thoroughly

investigated. The INM proteins that have been investigated interact with either A-type lamin, B-type lamin or both, and they regulate processes including heterochromatin positioning, signal transduction, and the nucleo-cytoskeleton via protein–protein and protein–DNA interactions (Wilson and Foisner, 2010). The interaction between INM proteins and the nuclear lamina contribute, at least in part, to the retention of INM proteins at the INM.

The lamina-associated polypeptide (LAP) family was originally defined as integral-membrane proteins that localize to the INM and bind to the nuclear lamina (Senior and Gerace, 1988). Presently, LAP family proteins also include nuclear lamina-binding proteins that do not contain transmembrane regions, such as LAP2 α (Wilson and Foisner, 2010), but all LAP family proteins are localized to the INM.

Other characterized nuclear lamina-binding proteins contain LEM (LAP2, Emerin, MAN1) domains, which directly interact with a protein called Barrier to Autointegration Factor (BAF). BAF binds with DNA with high affinity, and it also binds with core histones, linker histones, nuclear lamina proteins, and BAF itself (Margalit *et al.*, 2007). LEM domain proteins therefore interact with chromatin through BAF, and they also interact with chromatin directly. Because of these properties, LEM domain-containing proteins play an important role in anchoring the INM to chromatin. Another chromatin-anchoring INM protein is Lamin B receptor (LBR). LBR is a transmembrane protein that contains an N-terminal nucleoplasmic region and eight membrane-spanning regions (Olins *et al.*, 2010). LBR constantly associates with B-type Lamins via its N-terminal nucleoplasmic region. In addition to its interaction with B-type Lamin, the N-terminal nucleoplasmic region also interacts with components of heterochromatin such as core histones, heterochromatin protein 1 (HP1), and heterochromatin methyl-binding protein 1 (MECP1).

Importantly, the nuclear lamina is linked with microtubules and actin filaments present in the cytoplasm. The skeletal structures of the nucleoplasm (nucleoskeleton) and cytoplasm (cytoskeleton) are linked through a protein complex called linker of nucleoskeleton and cytoskeleton (LINC) complex. This connection between the nucleoskeleton and the cytoskeleton is important for the intracellular positioning of the nucleus, centrosome tethering, the regulation of the distance between the INM and ONM, and the transduction of extracellular mechanical stress. The LINC complex is composed of SUN (Sad1 and UNC-84) domain-containing proteins on the INM and KASH (Klarsicht, ANC-1, Syne homology) domain-containing proteins on the ONM. In mammals, at least five SUN domain-containing proteins have been discovered: SUN1, SUN2, SUN3, SUN4/SPAG4, and SUN5/SPAG4L. Among these proteins, SUN3 and SUN4/SPAG4 and SUN5/SPAG4L show tissue-specific expression in the testis (Sosa *et al.*, 2013). SUN1 and SUN2 each contain a single transmembrane domain and a C-terminal SUN domain, which is located in the luminal region of the NE. SUN1 and SUN2 are exclusively localized to the INM, and they directly bind to Lamin A at the nucleoplasmic side. In mammals, six KASH domain-containing proteins have been reported; among them, NE spectrin-repeat proteins 1–4 (Nesprin1–4) are well characterized for LINC complex formation. Nesprin proteins

interact with cytoplasmic actin, microtubules, centrosomes, or IFs via the long N-terminal cytoplasmic region that contains a large number of spectrin repeats (Rajgor and Shanahan, 2013). After the spectrin repeats, all Nesprin proteins contain a single transmembrane region and a luminal KASH domain in their C-termini. In the luminal region in the NE, the SUN domain rigidly and stably binds with KASH domain, thereby forming the LINC complex (Horn, 2014).

Laminopathies are also caused by the mutation of genes encoding lamin-associated proteins. Patients harboring mutations in the *Emerin* gene show X-linked Emery-Dreifuss muscular dystrophy (X-EDMD), and mutations in *SYNE1* or *SYNE2*, encoding Nesprin1 and Nesprin2, respectively, also induce EDMD (Horn, 2014). Other mutations in the *Emerin* gene cause limb-girdle muscular dystrophy, cardiomyopathy, or familial atrial fibrillation (Wilson and Foisner, 2010). The LEM domain-binding protein BAF causes Nestor-Guillermo progeria syndrome (Berk *et al.*, 2013), and LBR mutations are associated with hydrops-ectopic calcification-moth-eaten/Greenberg skeletal dysplasia and Pelger–Huët anomaly.

NE Disassembly and Reassembly in Mitosis

‘Open’ and ‘closed’ mitosis

Mitosis is the most important cellular process for dividing cells. The genetic information in the eukaryotic nuclei must be accurately transmitted to daughter cells by the mitotic spindle, which is mainly composed of microtubules and centrosomes. The mitotic spindle is generally assembled such that the microtubules, extending from the centrosomes and other microtubule-organizing centers, capture the kinetochores formed at the centromeres of mitotic chromosomes (O’Connell and Khodjakov, 2007).

Eukaryotic cells have two general styles of mitosis. One is ‘closed mitosis,’ in which the NE is sustained through mitosis, and the other is ‘open mitosis,’ in which the NE is completely broken down (Fernandez-Martinez and Rout, 2009; Sazer *et al.*, 2014). Yeast cells exhibit closed mitosis. They embed the spindle pole body (SPB), the microtubule-organizing center, in the NE. This allows mitotic spindle assembly in yeast without NE breakdown (NEBD). By contrast, in multicellular organisms that exhibit open mitosis, the microtubule-organizing center, the centromere, resides in the cytoplasm. In these cells, the NE must be broken down at the onset of mitosis so that the kinetochores can be captured by the microtubules extending from the centromeres. In late mitosis, the NE that was broken down must be reassembled on the surface of segregated daughter chromosomes (Anderson and Hetzer, 2008; Wandke and Kutay, 2013).

RanGTP gradient in interphase and mitosis

During interphase, large numbers of macromolecules are bidirectionally transported between the cytoplasm and the nucleus by NTRs such as the importin β superfamily (Clarke and Zhang, 2008). This bidirectional macromolecular transport is driven by the small GTPase Ran. The GTP-bound form of Ran (RanGTP) is concentrated at the nuclear interior, where the nucleotide exchange factor regulator of chromosome condensation 1 (RCC1) exclusively binds to chromatin and

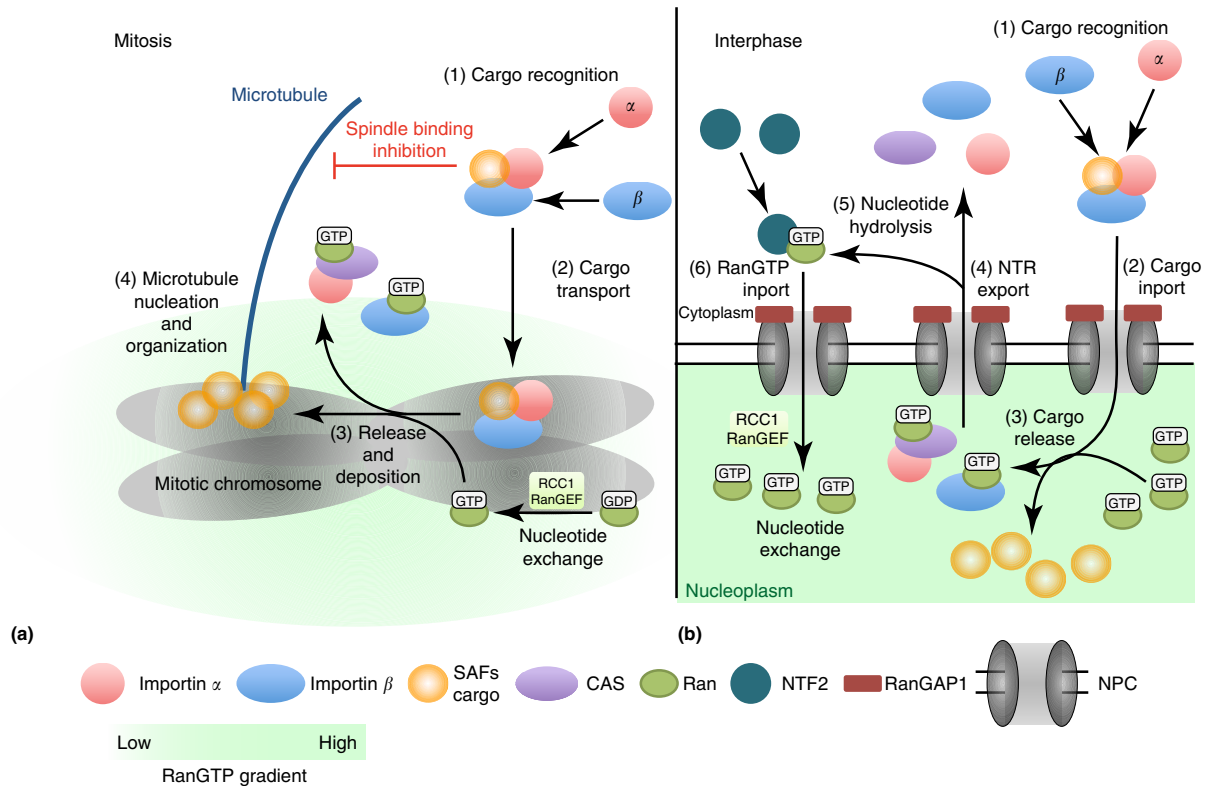


Figure 3 Importin α/β and Ran system in Interphase and Mitosis. (a) RanGTP and NTRs spatially regulate the spindle microtubule assembly via binding of spindle assembly factors (SAFs) to chromosome. (1) SAF harboring the nuclear localization signal (NLS) is recognized with nuclear transport receptor (NTR) such as Importin α/β , and they form a transport complex. The interaction of SAF (in this case, Tahara, K., Takagi, M., Ohsugi, M., *et al.*, 2008. Importin- β and the small guanosine triphosphatase Ran mediate chromosome loading of the human chromokinesin Kid. *Journal of Cell Biology* 180, 493–506) with NTR inhibits SAF's chromatin binding activity. (2) and (3), When the transport complex reaches mitotic chromosome, its disassembly is triggered by GTP-bound form of Ran (RanGTP) bound to NTR. (4) SAF released from transport complex accumulates on the chromosome, and promotes the spindle formation. RanGTP is constitutively generated via RCC1 residing on chromosomes, thus high concentration of RanGTP exists at the vicinity of chromosomes. (b) RanGTP and NTRs regulate the nuclear import of cargo proteins harboring the NLS. (1) By the same rule as (a), the cargo protein is recognized by NTRs, and they form a transport complex. (2) The transport complex passes through the nuclear pore complex (NPC), and enters the nuclear interior. (3) At the nuclear interior, RanGTP binds to the NTR and triggers disassembly of the transport complex. As a result, the cargo protein is released from the complex. (4) After releasing the cargo, the NTR binds to RanGTP and exits the nucleus through NPC. (5) In cytoplasm, GTPase activity of RanGTP bound to NTR is activated by the aid of RanGAP. Consequently, GTP on Ran is immediately hydrolyzed into GDP. (6) RanGDP generated by RanGAP in the cytoplasm is recognized by NTF2, and is imported into the nucleus. At the nuclear interior, GDP on Ran is replaced by GTP by RCC1 activity.

constantly replaces GDP with GTP on Ran. In the cytoplasm, RanGAP1 (RanGTPase-activating protein 1), which is exclusively localized to the cytoplasmic side of the NPC, facilitates the GTPase activity of Ran, and GTP on Ran is hydrolyzed to GDP. The GDP-bound form of Ran (RanGDP) in the cytoplasm is imported into the nucleus by nuclear transport factor 2 (NTF2). As a result, the concentration of RanGTP is higher in the nucleus than in the cytoplasm. When a cargo macromolecule containing an NLS is imported into the nucleus, it binds to a nuclear import receptor and then passes through the NPC. At the nuclear interior, RanGTP binds to the cargo-bound nuclear import receptor, thereby inducing cargo release into the nucleoplasm. Conversely, when a cargo macromolecule containing an NES is transported from the nucleus into the cytoplasm, the cargo interacts with the nuclear export receptor and RanGTP in the nucleus, forming the export complex. After passing through the NPC, RanGAP1 hydrolyzes

GTP on Ran to GDP in the cytoplasm, and the export complex consequently disassembles (Guttler and Gorlich, 2011; Figure 3).

During mitosis, although the NE, including the nuclear lamina and the NPC, is broken down, RCC1 remains to bind mitotic chromosomes. The action of RCC1 produces a high concentration of RanGTP around the mitotic chromosomes. The concentration of RanGTP gradually decreases toward the plasma membrane (Kalab *et al.*, 2006). As in interphase, the RanGTP gradient drives cargo release and binding to NTR during mitosis. Therefore, NLS-containing spindle assembly factors (SAFs) such as TPX2 and KID bind to nuclear import receptors in the mitotic cytoplasm. In the vicinity of mitotic chromosomes, SAFs are released from nuclear import receptors and become loaded onto mitotic chromosomes because of the high concentration of RanGTP. Many SAFs are associated with microtubule nucleation, stabilization and elongation; therefore,

the RanGTP gradient around mitotic chromosomes during mitosis forms a spatial context for regulating mitotic spindle assembly (Fuller, 2010; Kalab and Heald, 2008). In addition to mitotic spindle formation, the mitotic RanGTP gradient is also associated with the NE reassembly during late mitosis.

NE disassembly and reassembly

When the cell enters open mitosis, the entire NE structure, including the nuclear lamina, NPC, and nuclear membrane, must be disrupted for proper formation of the mitotic spindle. In this context, it is considered that phosphorylation by mitotic kinases is at least one of the key processes to disassemble the NE. The proteins forming the NE are directly phosphorylated via mitotic kinases including cyclin-dependent kinase1 (CDK1)-CyclinB, Polo-like kinase1 (PLK1), Protein kinase C (PKC), Aurora A and Aurora B. As initial phosphorylation targets in late prometaphase, peripheral FG repeat-containing Nups such as NUP98 become phosphorylated, before the phosphorylation of scaffold Nups, by multiple mitotic kinases including CDK1-CyclinB, disrupting the permeability barrier of NPCs (Laurell *et al.*, 2011). Consequently, cytoplasmic proteins leak into the nuclear interior and accelerate the NEBD. The lamins are also phosphorylated and depolymerized via CDK1-CyclinB activity at the mitotic entry. The LEM domain-binding protein BAF is phosphorylated via Vaccinia-related kinase (VRK) during mitosis. Phosphorylation of BAF by VRK causes the loss of its DNA-binding activity and reduces its interaction with Emerin (Margalit *et al.*, 2007). LBR is phosphorylated by mitotic kinases in its N-terminal region and loses its heterochromatin-binding activity. For proper progression of the NEBD, such phosphorylation-dependent attenuation of the interactions between INM proteins–chromatin might be essential. Furthermore, the cytoplasmic microtubules also support the destruction of the NE through their pulling force (Guttinger *et al.*, 2009).

After the onset of anaphase, PPases such as PP1 $\alpha/\beta/\gamma$ and PP2A/B are relatively active, whereas net kinase activities gradually decrease (Wurzenberger and Gerlich, 2011). The activated PP1 dephosphorylates the mitotically phosphorylated NE components. These dephosphorylation processes are associated with the NE reassembly. In late anaphase, the nucleosome-binding Nup ELYS is dephosphorylated, and it is then targeted to the surfaces of daughter chromosomes that are segregated by the mitotic spindle. Following the targeting of ELYS, the NUP107-160 subcomplex and POM121 are recruited to the chromosome surface. As a result, the complete NPC structure is built up on the daughter NE. BAF phosphorylated by VRK is dephosphorylated via PP2A, consequently restoring its ability to bind to both chromatin and LEM domain proteins (Asencio *et al.*, 2012). LBR can also become bound to components of heterochromatin upon its dephosphorylation. Such chromatin-binding activity of INM proteins through dephosphorylation is required for the re-formation of the NE in daughter cells (Clarke and Zhang, 2008).

Conclusion

In this section, we described the nuclear organization with a focus on the structures of chromatin and NE in interphase, as well as the mitotic dynamics of the NE. Through the cell

cycle, the chromatin structure is dynamically altered via post-translational modifications and histone exchange depending on both internal and external stimuli. In eukaryotic cells during interphase, the NE is the physical barrier protecting chromatin against numerous cytoplasmic toxic metabolites such as reactive oxygen species. The NE is also critical for the establishment of the macromolecular nucleocytoplasmic transport system that is operated by the NPC and diverse NTRs. The INM provides the sites in which heterochromatin is anchored via the interaction with the nuclear lamina, which is composed of lamins and INM proteins. By contrast, in cells demonstrating open mitosis, the NE is disassembled at the mitotic entry through mitotic kinase activities that allow chromatin segregation via microtubules. These cells reassemble the NE at late mitosis by the coordinated inactivation of mitotic kinases and activation of PPases. RanGTP and Importin β , which mediate nucleocytoplasmic transport in interphase, are required for assembling the mitotic spindle, chromosome segregation and NE formation, processes important for the inheritance of the mass of chromatin by daughter cells.

During the past decade, a number of NE proteins have been identified, along with their modifications and interactions (protein–protein, protein–DNA, protein–RNA), advancing our understanding of the structure and function of the NE itself. However, because the NE is tightly related to many cell biological processes involving cellular metabolism, cytoskeletons, and the nuclear genome, we require further clues to understand NE organization and dynamics in the context of its connections to a variety of cellular physiological processes.

See also: Organelles: Structure and Function: Nuclear Pores

References

- Anderson, D.J., Hetzer, M.W., 2008. The life cycle of the metazoan nuclear envelope. *Current Opinion in Cell Biology* 20, 386–392.
- Asencio, C., Davidson, I.F., Santarella-Mellwig, R., *et al.*, 2012. Coordination of kinase and phosphatase activities by Lem4 enables nuclear envelope reassembly during mitosis. *Cell* 150, 122–135.
- Berk, J.M., Tiffit, K.E., Wilson, K.L., 2013. The nuclear envelope LEM-domain protein emerlin. *Nucleus* 4, 298–314.
- Burke, B., Stewart, C.L., 2013. The nuclear lamins: Flexibility in function. *Nature Reviews Molecular Cell Biology* 14, 13–24.
- Butin-Israeli, V., Adam, S.A., Goldman, A.E., Goldman, R.D., 2012. Nuclear lamin functions and disease. *Trends in Genetics* 28, 464–471.
- Cantone, I., Fisher, A.G., 2013. Epigenetic programming and reprogramming during development. *Nature Structural & Molecular Biology* 20, 282–289.
- Chi, P., Allis, C.D., Wang, G.G., 2010. Covalent histone modifications—miswritten, misinterpreted and mis-erased in human cancers. *Nature Reviews Cancer* 10, 457–469.
- Clarke, P.R., Zhang, C., 2008. Spatial and temporal coordination of mitosis by Ran GTPase. *Nature Reviews Molecular Cell Biology* 9, 464–477.
- De Koning, L., Corpet, A., Haber, J.E., Almouzni, G., 2007. Histone chaperones: An escort network regulating histone traffic. *Nature Structural & Molecular Biology* 14, 997–1007.
- Fernandez-Martinez, J., Rout, M.P., 2009. Nuclear pore complex biogenesis. *Current Opinion in Cell Biology* 21, 603–612.
- Fuller, B.G., 2010. Self-organization of intracellular gradients during mitosis. *Cell Division* 5, 5.

- Greer, E.L., Shi, Y., 2012. Histone methylation: A dynamic mark in health, disease and inheritance. *Nature Reviews Genetics* 13, 343–357.
- Grossman, E., Medalia, O., Zwergner, M., 2012. Functional architecture of the nuclear pore complex. *Annual Review of Biophysics* 41, 557–584.
- Guttinger, S., Laurell, E., Kutay, U., 2009. Orchestrating nuclear envelope disassembly and reassembly during mitosis. *Nature Reviews Molecular Cell Biology* 10, 178–191.
- Guttler, T., Gorlich, D., 2011. Ran-dependent nuclear export mediators: A structural perspective. *EMBO Journal* 30, 3457–3474.
- Hoelz, A., Debler, E.W., Blobel, G., 2011. The structure of the nuclear pore complex. *Annual Review of Biochemistry* 80, 613–643.
- Horn, H.F., 2014. LINC complex proteins in development and disease. *Current Topics in Developmental Biology* 109, 287–321.
- Kalab, P., Heald, R., 2008. The RanGTP gradient — A GPS for the mitotic spindle. *Journal of Cell Science* 121, 1577–1586.
- Kalab, P., Pralle, A., Isacoff, E.Y., Heald, R., Weis, K., 2006. Analysis of a RanGTP-regulated gradient in mitotic somatic cells. *Nature* 440, 697–701.
- Kurumizaka, H., Horikoshi, N., Tachiwana, H., Kagawa, W., 2013. Current progress on structural studies of nucleosomes containing histone H3 variants. *Current Opinion in Structural Biology* 23, 109–115.
- Laurell, E., Beck, K., Krupina, K., *et al.*, 2011. Phosphorylation of Nup98 by multiple kinases is crucial for NPC disassembly during mitotic entry. *Cell* 144, 539–550.
- Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J., 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251–260.
- Margalit, A., Brachner, A., Gotzmann, J., Foisner, R., Gruenbaum, Y., 2007. Barrier-to-autointegration factor a BAFFling little protein. *Trends in Cell Biology* 17, 202–208.
- O'Connell, C.B., Khodjakov, A.L., 2007. Cooperative mechanisms of mitotic spindle formation. *Journal of Cell Science* 120, 1717–1722.
- Olins, A.L., Rhodes, G., Welch, D.B., Zwergner, M., Olins, D.E., 2010. Lamin B receptor: Multi-tasking at the nuclear envelope. *Nucleus* 1, 53–70.
- Pryor Jr., E.E., Horanyi, P.S., Clark, K.M., *et al.*, 2013. Structure of the integral membrane protein CAAX protease Ste24p. *Science* 339, 1600–1604.
- Quigley, A., Dong, Y.Y., Pike, A.C., *et al.*, 2013. The structural basis of ZMPSTE24-dependent laminopathies. *Science* 339, 1604–1607.
- Rajgor, D., Shanahan, C.M., 2013. Nesprins: From the nuclear envelope and beyond. *Expert Reviews in Molecular Medicine* 15, e5.
- Roskoski, R., 2003. Protein prenylation: A pivotal posttranslational process. *Biochemical and Biophysical Research Communications* 303, 1–7.
- Rothballe, A., Kutay, U., 2013. Poring over pores: Nuclear pore complex insertion into the nuclear envelope. *Trends in Biochemical Sciences* 38, 292–301.
- Sasaki, H., Matsui, Y., 2008. Epigenetic events in mammalian germ-cell development: Reprogramming and beyond. *Nature Reviews Genetics* 9, 129–140.
- Sazer, S., Lynch, M., Needleman, D., 2014. Deciphering the evolutionary history of open and closed mitosis. *Current Biology* 24, R1099–R1103.
- Senior, A., Gerace, L., 1988. Integral membrane proteins specific to the inner nuclear membrane and associated with the nuclear lamina. *Journal of Cell Biology* 107, 2029–2036.
- Sosa, B.A., Kutay, U., Schwartz, T.U., 2013. Structural insights into LINC complexes. *Current Opinion in Structural Biology* 23, 285–291.
- Volle, C., Dalal, Y., 2014. Histone variants: The tricksters of the chromatin world. *Current Opinion in Genetics & Development* 25, 8–14.
- Wandke, C., Kutay, U., 2013. Enclosing chromatin: Reassembly of the nucleus after open mitosis. *Cell* 152, 1222–1225.
- Wilson, K.L., Berk, J.M., 2010. The nuclear envelope at a glance. *Journal of Cell Science* 123, 1973–1978.
- Wilson, K.L., Foisner, R., 2010. Lamin-binding Proteins. *Cold Spring Harbor Perspectives in Biology* 2, a000554.
- Winter-Vann, A.M., Casey, P.J., 2005. Post-prenylation-processing enzymes as new targets in oncogenesis. *Nature Reviews Cancer* 5, 405–412.
- Wurzenberger, C., Gerlich, D.W., 2011. Phosphatases: Providing safe passage through mitotic exit. *Nature Reviews Molecular Cell Biology* 12, 469–482.

Nuclear Pores

M Oka and Y Yoneda, National Institute of Biomedical Innovation (NIBIO), Osaka, Japan

© 2016 Elsevier Inc. All rights reserved.

Structure of the Nuclear Pore Complex

The morphology of nuclear pore complex (NPC) has been studied intensively using electron microscopy (EM) (Callan and Tomlin, 1950; Gall, 1967; Maul, 1971) and cryo-electron microscopy (Akey and Radermacher, 1993; Yang *et al.*, 1998; Stoffer *et al.*, 2003; Beck *et al.*, 2004). Recent studies with cryo-electron tomography shows a 3D model of NPCs with ~6-nm resolution (Frenkiel-Krispin *et al.*, 2010; Beck *et al.*, 2007). The NPC shows an eightfold rotational symmetry to its central axis. There are three major octagonal rings connected by spokes, which form the core scaffold of the NPC. Specifically, they are called cytoplasmic ring, central spoke ring, and nuclear ring. The central channel located in the middle provides the NPC's function as a selective permeable barrier. There are elongated filamentous structures found asymmetrical in the cytoplasmic and nuclear sides of the NPC. Eight of these filaments extend from cytoplasmic ring into the cytoplasm, while the nuclear filaments extending from nuclear ring are connected at their ends forming a basket-like structure (Figure 1).

Even though the size of the NPC varies from ~66 MDa in yeast (Rout and Blobel, 1993) to ~125 MDa in vertebrates

(Reichelt *et al.*, 1990), its structural feature is conserved across species (Hinshaw *et al.*, 1992; Akey and Radermacher, 1993; Yang *et al.*, 1998; Grossman *et al.*, 2012).

NPCs are composed of approximately 30 different proteins, called nucleoporins (Nups), which are conserved among eukaryotes (Rout *et al.*, 2000; Cronshaw *et al.*, 2002; DeGrasse *et al.*, 2009; Tamura *et al.*, 2010).

Nucleoporins (Nups)

Nups can be classified according to their sequence motifs, structural folds, mobility, or relative localization within the NPC (Rout *et al.*, 2000; Tran and Went, 2006; Hoelz *et al.*, 2011; Grossman *et al.*, 2012). An example of Nup-classification is (1) membrane Nups, (2) scaffold Nups, and (3) peripheral Nups.

One characteristic sequence motif/structural folds of Nups is the tandem repeats of phenylalanine-glycine (FG repeats), which are unfolded in native protein (Denning *et al.*, 2003). These FG repeats are found in approximately one-third of Nups (also called as FG-Nups), mostly belonging to the peripheral Nups. Most of these FG repeats are further classified

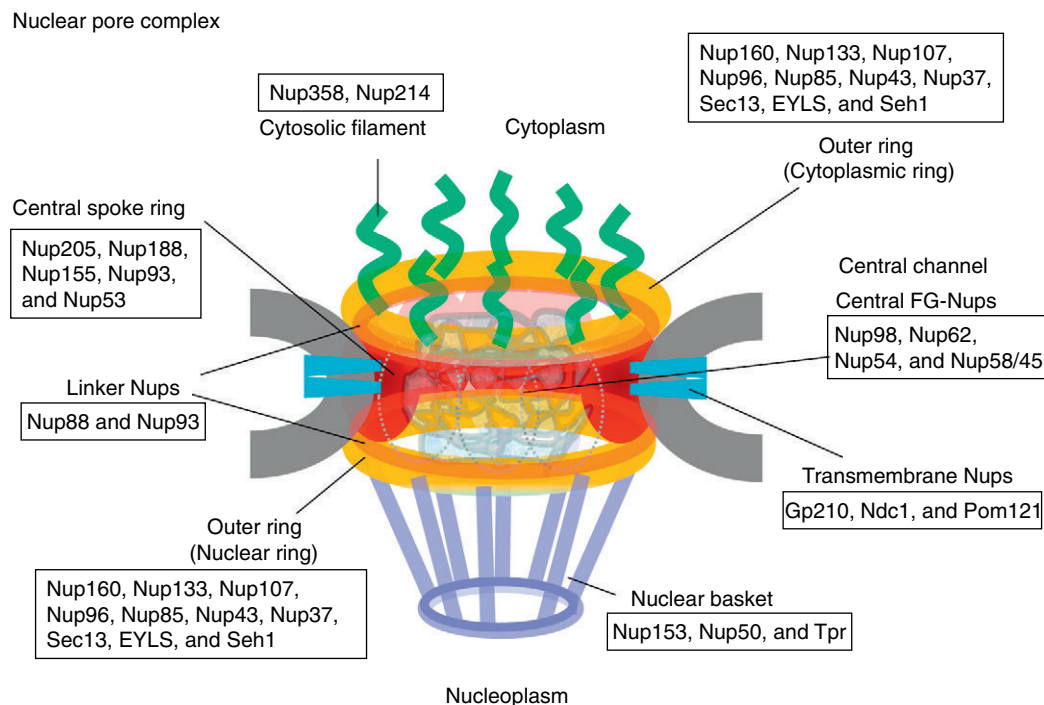


Figure 1 The structure of nuclear pore complex (NPC). Nups consist of each structural component are depicted in the box. There are three major rings (cytoplasmic ring, central spoke ring, and nuclear ring) forming the core scaffold, which is anchored to the nuclear membrane by transmembrane Nups. The cytoplasmic filaments are extending from the cytoplasmic ring into the cytoplasm, while the nuclear filaments extending from nuclear ring are connected at their ends forming a basket-like structure. The central channel is filled with central FG-Nups, which provides the function of NPC as a selective permeable barrier.

into either Phe-Gly (FG), Gly-Leu-Phe-Gly (GLFG), or Phe-any-Phe-Gly (FXFG) (Rout and Wente, 1994).

FG-Nups located in the central channel of NPC (central FG-Nups) are important for barrier formation, which inhibits the passive diffusion of macromolecules through the NPC. Furthermore, FG-Nups can bind to nuclear transport factors (Radu *et al.*, 1995; Rexach and Blobel, 1995; Hu *et al.*, 1996; Clarkson *et al.*, 1996; Bayliss *et al.*, 2000), which is important for the active transport process of selected molecules. Indeed, nanotube coated with FG-Nups can selectively allow the transport of karyopherin or karyopherin-cargo complexes (Jovanovic-Talisman *et al.*, 2009); moreover, some of these FG-Nups can form hydrogels *in vitro* that selectively allow the influx of transport factors (Frey and Gorlich, 2007; Hulsman *et al.*, 2012). Thus, FG-Nups are important for selective translocation of transport factor-cargo complexes through the NPC.

The property of this 'selective permeability barrier' of NPC is presumably attributed to either the highly dense unfolded FG domains which function as entropic bristles for macromolecules (Rout *et al.*, 2003; Lim *et al.*, 2006) which also can be collapsed upon karyopherin binding (Lim *et al.*, 2007) or the formation of cohesive meshwork through self- or inter-FG-Nup interaction (Ribbeck and Gorlich, 2001; Alber *et al.*, 2007) or hybrid of them (Patel *et al.*, 2007; Yamada *et al.*, 2010). Even though there are several sophisticated molecular models postulated to explain the mechanism of nuclear transport through the NPC, including those mentioned above, the exact molecular mechanism remains to be established (Walde and Kehlenbach, 2010; Grunwald *et al.*, 2011).

Structural Nups or scaffold Nups are important for the formation of the scaffold of NPC, and they can interact both with FG-Nups and transmembrane Nups. Almost half of Nups belongs to this group, and many of these structural Nups are contained at the inner, outer ring (nuclear ring and cytoplasmic ring) structures, central spoke ring, and linker connecting these rings. Structural Nups often contain α -solenoid fold and/or β -propeller fold. Of these, one structural complex termed Y-complex (known as Nup84 subcomplex in yeast and Nup107/160 subcomplex in vertebrates), named after its shape observed in EM (Siniosoglou *et al.*, 2000; Lutzmann *et al.*, 2002) has been well characterized. Interestingly, the structure of Y-complex shows similarity to COPII coated vesicle complexes, which also contain α -solenoid fold and β -propeller protein domains, thus, suggesting that NPC and coated vesicle might originate from the same ancestor (Devos *et al.*, 2004; Brohawn *et al.*, 2008).

Transmembrane nucleoporins anchor the NPC to the nuclear membrane. This group of Nups mainly consists of three members both in yeast (Pom34, Ndc1, and Pom152) and metazoa (gp210, Ndc1, and POM121). These transmembrane Nups were also suggested playing important roles in assembly/disassembly of NPCs (Guttinger *et al.*, 2009).

Dynamic Behavior of Nucleoporin

NPCs are not static structures playing a role as gateways in the nucleocytoplasmic transport system, but also show dynamic properties. Indeed, some of the peripheral Nups can get away

from NPC, as mobile Nups, they can actively participate in the nucleocytoplasmic transport (Tran and Wente, 2006) and be involved in other cellular processes, such as gene regulation, chromatin organization, mitotic spindle assembly, or DNA repair (Suntharalingam and Wente, 2003; Strambio-De-Castillia *et al.*, 2010).

A study showed that mammalian peripheral (mobile) Nups, such as Nup50 and Nup153, reside within the NPC for a short period of time (seconds scale), but stable scaffold Nups including Nup107 and Nup133, can remain within the NPC for days (Rabut *et al.*, 2004). Furthermore, these scaffold Nups are assumed to be one of the most stable proteins in a cell (Savas *et al.*, 2012).

These dynamic properties are also observed during open mitosis, in which the NPCs are disassembled into subcomplexes, which occurs in most multicellular organisms. At the end of mitosis, those subcomplexes assemble within the nuclear envelope formed around chromatin.

Nucleocytoplasmic Transport

NPCs are the sole gateway of nucleocytoplasmic transport. Small molecules, such as ions, metabolites, or peptides, can pass through the NPCs via passive diffusion; however, macromolecules, such as protein (larger than approximately 40 kDa) and RNA, are transported from cytoplasm into nucleus (import) or transported from nucleus into cytoplasm (export) in an energy-dependent process. Furthermore, even a macromolecule with a diameter of approximately 39 nm can pass through the NPC (Pante and Kann, 2002).

These active transports are generally mediated via transport receptors called karyopherin (or importin or exportin, termed their properties) (Hicks and Raikhel, 1995; Gorlich and Mattaj, 1996; Yoneda, 2000). The cargo protein macromolecules contain nuclear localization signal (NLS) or nuclear export signal (NES) or both, and those are recognized by specific subsets of karyopherin families, depending on their NLS or NES signal sequences. Karyopherins can bind to various FG-Nups with low affinity, and that causes the selective and efficient translocation of cargo in and out of the nucleus.

In case of RNAs, either karyopherin directly binds to RNAs (tRNA or miRNA) or indirectly via RNA-binding adapter proteins (mRNA in the form of messenger ribonucleoprotein particle) (Kohler and Hurt, 2007).

The direction of active nucleocytoplasmic transport is mainly determined by small GTPase RanGTP-GDP cycle. Ran exists as in GTP-bound form (RanGTP) or GDP-bound form (RanGDP). While a nucleotide exchange factor for Ran, RCC1, localizing in the nucleus catalyzes the formation of RanGTP (Bischoff and Ponstingl, 1991), a GTPase activating protein for Ran, RanGAP (Bischoff *et al.*, 1994), RanBP1 and RanBP2 localizing in the cytoplasm hydrolyze them into RanGDP (Gorlich *et al.*, 1996; Bischoff *et al.*, 1995). Thus, this compartmentalized distribution of RanGTP-GDP cycle regulators cause the creation of a keen asymmetric nucleocytoplasmic ratio of RanGTP/RanGDP, called Ran gradient. While importin-cargo complex can be disassembled in the presence of RanGTP in the nucleus, exportin-cargo complex can be formed only in the presence of RanGTP. Therefore, nuclear

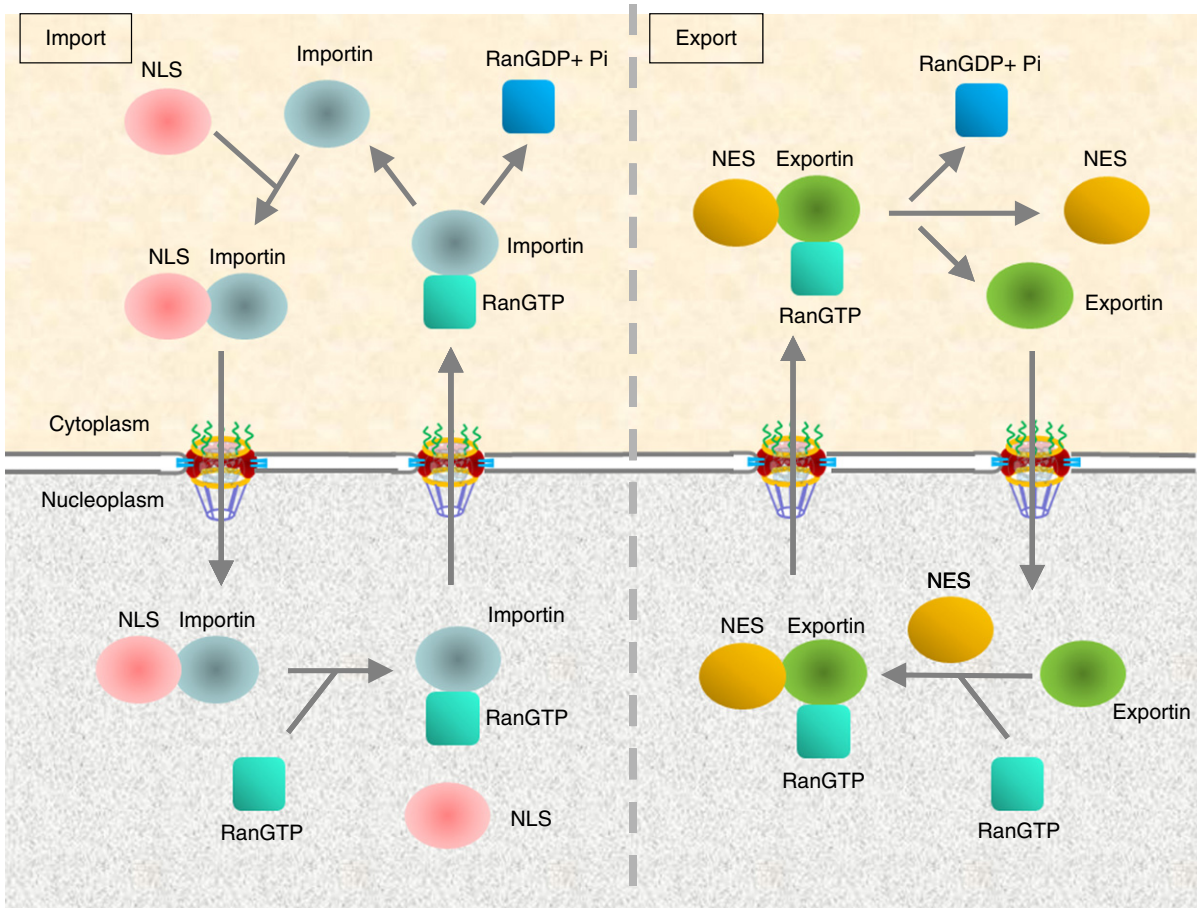


Figure 2 The mechanism of nucleocytoplasmic transport of proteins. Importin binds to the cargo proteins through NLS in the cytoplasm. After translocation into the nucleus, binding of RanGTP to importin cause the disassembly of the cargo (NLS)-importin complex. On the other hand, exportin binds to the cargo proteins through NES in the presence of RanGTP in the nucleus, then, carried into the cytoplasm. The hydrolysis of RanGTP into RanGDP in the cytoplasm cause the disassembly of cargo(NES)-exportin-RanGTP complex. Cargo-free importin or exportin can shuttle back into the cytoplasm or nucleus, respectively.

RanGTP directs the orientation of transport processes (Figure 2).

Nucleoporin and Disease

The malfunction of nucleoporin has been linked to diseases. A mutation that affects the subcellular localization of Nup155 or ALADIN is known to cause cardiac disease (Zhang *et al.*, 2008) or triple-A syndrome (a rare disease characterized by alacrima, achalasia of the esophageal cardia, and adrenal failure) (Tullio-Pelet *et al.*, 2000; Cronshaw and Matunis, 2003; Handschug *et al.*, 2001), respectively. Nucleoporins, such as Nup214, Nup98, Nup358, or Tpr, form Nup-fusion genes with various partner genes in malignancies, and those possess oncogenic activities (Cronshaw and Matunis, 2004; Capelson and Hetzer, 2009; Xu and Powers, 2009). Of these, Nup98 is known to fuse with ~30 different partner genes, whose physiological functions are diverse, as seen in leukemia (Gough *et al.*, 2011; Nakamura *et al.*, 1996). However, it remains unknown

whether these diverse Nup98-fusions share the same mechanism of pathogenesis.

Collectively, these facts suggest that unregulated nucleocytoplasmic traffic of specific subsets of cargo may induce the disease of various types.

See also: Organelles: Structure and Function: Nuclear Organization (Nuclear Structure and Dynamics)

References

- Akey, C.W., Radermacher, M., 1993. Architecture of the *Xenopus* nuclear pore complex revealed by three-dimensional cryo-electron microscopy. *Journal of Cell Biology* 122, 1–19.
- Alber, F., Dokudovskaya, S., Veenhoff, L.M., *et al.*, 2007. The molecular architecture of the nuclear pore complex. *Nature* 450, 695–701.
- Bayliss, R., Littlewood, T., Stewart, M., 2000. Structural basis for the interaction between FxFG nucleoporin repeats and importin-beta in nuclear trafficking. *Cell* 102, 99–108.

- Beck, M., Forster, F., Ecke, M., *et al.*, 2004. Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. *Science* 306, 1387–1390.
- Beck, M., Lucic, V., Forster, F., Baumeister, W., Medalia, O., 2007. Snapshots of nuclear pore complexes in action captured by cryo-electron tomography. *Nature* 449, 611–615.
- Bischoff, F.R., Klebe, C., Kretschmer, J., Wittinghofer, A., Ponstingl, H., 1994. RanGAP1 induces GTPase activity of nuclear Ras-related Ran. *Proceedings of the National Academy of Sciences of the USA* 91, 2587–2591.
- Bischoff, F.R., Krebber, H., Smirnova, E., Dong, W., Ponstingl, H., 1995. Co-activation of RanGTPase and inhibition of GTP dissociation by Ran-GTP binding protein RanBP1. *EMBO Journal* 14, 705–715.
- Bischoff, F.R., Ponstingl, H., 1991. Mitotic regulator protein RCC1 is complexed with a nuclear ras-related polypeptide. *Proceedings of the National Academy of Sciences of the USA* 88, 10830–10834.
- Brohawn, S.G., Leksa, N.C., Spear, E.D., Rajashankar, K.R., Schwartz, T.U., 2008. Structural evidence for common ancestry of the nuclear pore complex and vesicle coats. *Science* 322, 1369–1373.
- Callan, H.G., Tomlin, S.G., 1950. Experimental studies on amphibian oocyte nuclei. I. Investigation of the structure of the nuclear membrane by means of the electron microscope. *Proceedings of the Royal Society B: Biological Sciences* 137, 367–378.
- Capelson, M., Hetzer, M.W., 2009. The role of nuclear pores in gene regulation, development and disease. *EMBO reports* 10, 697–705.
- Clarkson, W.D., Kent, H.M., Stewart, M., 1996. Separate binding sites on nuclear transport factor 2 (NTF2) for GDP-Ran and the phenylalanine-rich repeat regions of nucleoporins p62 and Nsp1p. *Journal of Molecular Biology* 263, 517–524.
- Cronshaw, J.M., Krutchinsky, A.N., Zhang, W., Chait, B.T., Matunis, M.J., 2002. Proteomic analysis of the mammalian nuclear pore complex. *Journal of Cell Biology* 158, 915–927.
- Cronshaw, J.M., Matunis, M.J., 2003. The nuclear pore complex protein ALADIN is mislocalized in triple A syndrome. *Proceedings of the National Academy of Sciences of the USA* 100, 5823–5827.
- Cronshaw, J.M., Matunis, M.J., 2004. The nuclear pore complex: Disease associations and functional correlations. *Trends in Endocrinology & Metabolism* 15, 34–39.
- DeGrasse, J.A., Dubois, K.N., Devos, D., *et al.*, 2009. Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. *Molecular & Cellular Proteomics* 8, 2119–2130.
- Denning, D.P., Patel, S.S., Uversky, V., Fink, A.L., Rexach, M., 2003. Disorder in the nuclear pore complex: The FG repeat regions of nucleoporins are natively unfolded. *Proceedings of the National Academy of Sciences of the USA* 100, 2450–2455.
- Devos, D., Dokudovskaya, S., Alber, F., *et al.*, 2004. Components of coated vesicles and nuclear pore complexes share a common molecular architecture. *PLoS Biology* 2, e380.
- Frenkiel-Krispin, D., Maco, B., Aebi, U., Medalia, O., 2010. Structural analysis of a metazoan nuclear pore complex reveals a fused concentric ring architecture. *Journal of Molecular Biology* 395, 578–586.
- Frey, S., Gorlich, D., 2007. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell* 130, 512–523.
- Gall, J.G., 1967. Octagonal nuclear pores. *Journal of Cell Biology* 32, 391–399.
- Gorlich, D., Mattaj, I.W., 1996. Nucleocytoplasmic transport. *Science* 271, 1513–1518.
- Gorlich, D., Pante, N., Kutay, U., Aebi, U., Bischoff, F.R., 1996. Identification of different roles for RanGDP and RanGTP in nuclear protein import. *EMBO Journal* 15, 5584–5594.
- Gough, S.M., Slape, C.I., Aplan, P.D., 2011. NUP98 gene fusions and hematopoietic malignancies: Common themes and new biologic insights. *Blood* 118, 6247–6257.
- Grossman, E., Medalia, O., Zwerger, M., 2012. Functional architecture of the nuclear pore complex. *Annual Review of Biophysics* 41, 557–584.
- Grunwald, D., Singer, R.H., Rout, M., 2011. Nuclear export dynamics of RNA-protein complexes. *Nature* 475, 333–341.
- Guttinger, S., Laurell, E., Kutay, U., 2009. Orchestrating nuclear envelope disassembly and reassembly during mitosis. *Nature Reviews Molecular Cell Biology* 10, 178–191.
- Handschug, K., Sperling, S., Yoon, S.J., *et al.*, 2001. Triple A syndrome is caused by mutations in AAS, a new WD-repeat protein gene. *Human Molecular Genetics* 10, 283–290.
- Hicks, G.R., Raikhel, N.V., 1995. Protein import into the nucleus: An integrated view. *Annual Review of Cell and Developmental Biology* 11, 155–188.
- Hinshaw, J.E., Carragher, B.O., Milligan, R.A., 1992. Architecture and design of the nuclear pore complex. *Cell* 69, 1133–1141.
- Hoelz, A., Debler, E.W., Blobel, G., 2011. The structure of the nuclear pore complex. *Annual Review of Biochemistry* 80, 613–643.
- Hu, T., Guan, T., Gerace, L., 1996. Molecular and functional characterization of the p62 complex, an assembly of nuclear pore complex glycoproteins. *Journal of Cell Biology* 134, 589–601.
- Hulsmann, B.B., Labokha, A.A., Gorlich, D., 2012. The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. *Cell* 150, 738–751.
- Jovanovic-Talisman, T., Tetenbaum-Novatt, J., McKenney, A.S., *et al.*, 2009. Artificial nanopores that mimic the transport selectivity of the nuclear pore complex. *Nature* 457, 1023–1027.
- Kohler, A., Hurt, E., 2007. Exporting RNA from the nucleus to the cytoplasm. *Nature Reviews Molecular Cell Biology* 8, 761–773.
- Lim, R.Y., Fahrenkrog, B., Koser, J., *et al.*, 2007. Nanomechanical basis of selective gating by the nuclear pore complex. *Science* 318, 640–643.
- Lim, R.Y., Huang, N.P., Koser, J., *et al.*, 2006. Flexible phenylalanine-glycine nucleoporins as entropic barriers to nucleocytoplasmic transport. *Proceedings of the National Academy of Sciences of the USA* 103, 9512–9517.
- Lutzmann, M., Kunze, R., Buerer, A., Aebi, U., Hurt, E., 2002. Modular self-assembly of a Y-shaped multiprotein complex from seven nucleoporins. *EMBO Journal* 21, 387–397.
- Maul, G.G., 1971. On the octagonality of the nuclear pore complex. *Journal of Cell Biology* 51, 558–563.
- Nakamura, T., Largaespada, D.A., Lee, M.P., *et al.*, 1996. Fusion of the nucleoporin gene NUP98 to HOXA9 by the chromosome translocation t(7;11)(p15;p15) in human myeloid leukaemia. *Nature Genetics* 12, 154–158.
- Pante, N., Kann, M., 2002. Nuclear pore complex is able to transport macromolecules with diameters of about 39 nm. *Molecular Biology of the Cell* 13, 425–434.
- Patel, S.S., Belmont, B.J., Sante, J.M., Rexach, M.F., 2007. Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex. *Cell* 129, 83–96.
- Rabut, G., Doye, V., Ellenberg, J., 2004. Mapping the dynamic organization of the nuclear pore complex inside single living cells. *Nature Cell Biology* 6, 1114–1121.
- Radu, A., Moore, M.S., Blobel, G., 1995. The peptide repeat domain of nucleoporin Nup98 functions as a docking site in transport across the nuclear pore complex. *Cell* 81, 215–222.
- Reichelt, R., Holzenburg, A., Buhle Jr., E.L., *et al.*, 1990. Correlation between structure and mass distribution of the nuclear pore complex and of distinct pore complex components. *Journal of Cell Biology* 110, 883–894.
- Rexach, M., Blobel, G., 1995. Protein import into nuclei: Association and dissociation reactions involving transport substrate, transport factors, and nucleoporins. *Cell* 83, 683–692.
- Ribbeck, K., Gorlich, D., 2001. Kinetic analysis of translocation through nuclear pore complexes. *EMBO Journal* 20, 1320–1330.
- Rout, M.P., Aitchison, J.D., Magnasco, M.O., Chait, B.T., 2003. Virtual gating and nuclear transport: The hole picture. *Trends in Cell Biology* 13, 622–628.
- Rout, M.P., Aitchison, J.D., Suprapto, A., *et al.*, 2000. The yeast nuclear pore complex: Composition, architecture, and transport mechanism. *Journal of Cell Biology* 148, 635–651.
- Rout, M.P., Blobel, G., 1993. Isolation of the yeast nuclear pore complex. *Journal of Cell Biology* 123, 771–783.
- Rout, M.P., Wente, S.R., 1994. Pores for thought: Nuclear pore complex proteins. *Trends in Cell Biology* 4, 357–365.
- Savas, J.N., Toyama, B.H., Xu, T., Yates 3rd, J.R., Hetzer, M.W., 2012. Extremely long-lived nuclear pore proteins in the rat brain. *Science* 335, 942.
- Siniosoglou, S., Lutzmann, M., Santos-Rosa, H., *et al.*, 2000. Structure and assembly of the Nup84p complex. *Journal of Cell Biology* 149, 41–54.
- Stoffler, D., Feja, B., Fahrenkrog, B., *et al.*, 2003. Cryo-electron tomography provides novel insights into nuclear pore architecture: Implications for nucleocytoplasmic transport. *Journal of Molecular Biology* 328, 119–130.
- Strambio-De-Castilla, C., Niepel, M., Rout, M.P., 2010. The nuclear pore complex: Bridging nuclear transport and gene regulation. *Nature Reviews Molecular Cell Biology* 11, 490–501.
- Suntharalingam, M., Wente, S.R., 2003. Peering through the pore: Nuclear pore complex structure, assembly, and function. *Developmental Cell* 4, 775–789.
- Tamura, K., Fukao, Y., Iwamoto, M., Haraguchi, T., Hara-Nishimura, I., 2010. Identification and characterization of nuclear pore complex components in *Arabidopsis thaliana*. *Plant Cell* 22, 4084–4097.
- Tran, E.J., Wente, S.R., 2006. Dynamic nuclear pore complexes: Life on the edge. *Cell* 125, 1041–1053.

- Tullio-Pelet, A., Salomon, R., Hadj-Rabia, S., *et al.*, 2000. Mutant WD-repeat protein in triple-A syndrome. *Nature Genetics* 26, 332–335.
- Walde, S., Kehlenbach, R.H., 2010. The part and the whole: Functions of nucleoporins in nucleocytoplasmic transport. *Trends in Cell Biology* 20, 461–469.
- Xu, S., Powers, M.A., 2009. Nuclear pore proteins and cancer. *Seminars in Cell & Developmental Biology* 20, 620–630.
- Yamada, J., Phillips, J.L., Patel, S., *et al.*, 2010. A bimodal distribution of two distinct categories of intrinsically disordered structures with separate functions in FG nucleoporins. *Molecular & Cellular Proteomics* 9, 2205–2224.
- Yang, Q., Rout, M.P., Akey, C.W., 1998. Three-dimensional architecture of the isolated yeast nuclear pore complex: Functional and evolutionary implications. *Molecular Cell* 1, 223–234.
- Yoneda, Y., 2000. Nucleocytoplasmic protein traffic and its significance to cell function. *Genes Cells* 5, 777–787.
- Zhang, X., Chen, S., Yoo, S., *et al.*, 2008. Mutation in nuclear pore component NUP155 leads to atrial fibrillation and early sudden cardiac death. *Cell* 135, 1017–1027.

Plastids: The Anabolic Factories of Plant Cells

RR Wise, University of Wisconsin Oshkosh, Oshkosh, WI, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Chloroplasts are photosynthetic organelles found in plant cells. Indeed, the very definition of what it means to be a plant is largely predicated on their photoautotrophic abilities, which reside exclusively in the chloroplast. However, chloroplasts are only one member of the plastid family, a diverse, phylogenetically and ontogenetically related group of organelles (Kirk and Tilney-Bassett, 1967). Depending on the breadth of the definition, there are as many as 20 distinctly different types of plastids found in both plants and animals (Wise, 2006). The seven main plastid types found in green plants (members of the clade Viridiplantae; Figure 1) will be considered in this encyclopedia article to illustrate the wide diversity of anabolic activities of which plastids are capable – well beyond the basic, life-supporting process of photosynthesis.

Substantial evidence exists to support an endosymbiotic origin for plastids, with the original endosymbiont being a photosynthetic cyanobacterium (Sato, 2006; Stiller, 2014). Two things happened during the ensuing 1.6 billion years of evolution: (1) approximately 90% of the plastid genome was transferred to the nucleus, and (2) the proto-chloroplast radiated and evolved into the other plastid types (Keeling, 2010). Secondary and tertiary endosymbiotic events and plastid loss complicate this simple picture, but the plastid family has a common origin. Thus, all plastids share the same basic characters of a double-membrane envelope; a separate and third internal membrane system of greater (chromoplast and chloroplast) or lesser (proplastid and amyloplast) complexity; a complete prokaryotic-like genetic machinery consisting of organellar DNA, transcription factors and ribosomes; and division by fission (again, very prokaryotic-like) (Miya-gishima and Kuroiwa, 2006). Unlike mitochondria, which were also derived endosymbiotically (Gray *et al.*, 1999), plastids did not surrender the anabolic abilities of their free-living ancestors. In fact, the power to manufacture all the biomolecules needed to support every metabolic and developmental pathway has enabled plants to use plastids as the central organelle for anabolism, catabolism, energy regulation, and environmental sensing.

Plastids of the Viridiplantae

Proplastid

Plants have egg cells, much like animals, and the egg cell is the sole source of organelles for the zygote that leads, eventually, to an adult plant. The plastids found in the egg are called ‘proplastids’ and are the progenitors of all the other plastids (regardless of type) found in the mature plant (Wise, 2006). Subsequent mitosis and the generation of new tissues in plants take place in zones at the tips of shoots and roots called

meristems. Meristematic plant cells also contain proplastids. Most plants continue to grow throughout their entire life to support shoot and root expansion. Therefore, they may have hundreds to thousands of meristems at any given time, and thus a large supply of proplastids.

Unlike the other plastid types, proplastids as a group are defined by their appearance and location, and not by any specific metabolic function. Characteristically, they are small, relatively nondifferentiated with a few internal membranes and only found in young, undifferentiated cells (Figure 2(a)).

Etioplast

Etioplasts develop in complete darkness and are a transitional stage between the proplastid and the chloroplast during the process of greening (Figure 1; Willows, 2006). The reverse process – degreening (chloroplast to etioplast) – is also possible as evidenced by the pale color that results under any light barrier laid on a green lawn (board, tarp, garden hose, etc). In this process, the green grass undergoes chlorosis (the controlled, enzymatic loss of chlorophyll) and degreens upon the imposition of darkness. Removing the obstruction allows the greening process to resume. Thus the etioplast-to-chloroplast-to-etoplast transition is quite dynamic and under strict genetic control (Zhong *et al.*, 2014).

Etioplast ultrastructure is dominated by the prolamellar body (PLB), a large, crystalline lattice of membranes (Figure 2(b); Gunning, 2001). While the exact function of the PLB remains obscure, it probably represents a ‘holding pattern’ for the large amount of membrane and protein that will ultimately be utilized in formation of the thylakoid system (see chloroplast, below). Interestingly, PLBs only form in dark-grown cells of the leaves and stems, tissues that under normal conditions will eventually be exposed to light. They never form in root cells. The plant growth regulator gibberellic acid (GA) is synthesized in the etioplast (Solymosi *et al.*, 2004). Because GA helps direct many early developmental processes, such as those in an expanding and greening leaf, the etioplast probably contributes directly to the light-to-dark transition.

Elaioplast

Lipid synthesis in animal cells is localized to the smooth endoplasmic reticulum (SER). However, SER is rarely seen in plant cells because most plant lipids are made in the proplastid, chloroplast, or elaioplast (Dörmann, 2006; Hernandez-Pinzon *et al.*, 1999; Bates *et al.*, 2013). Elaioplasts (Figure 2(c)) are plastids that specialize in oil synthesis and storage and are found primarily in the layer of cells in the anther that surrounds developing pollen grains (celled the tapetum or tapetal layer) (Suzuki *et al.*, 2013). After meiosis and just prior to pollen release, the tapetal layer degrades and

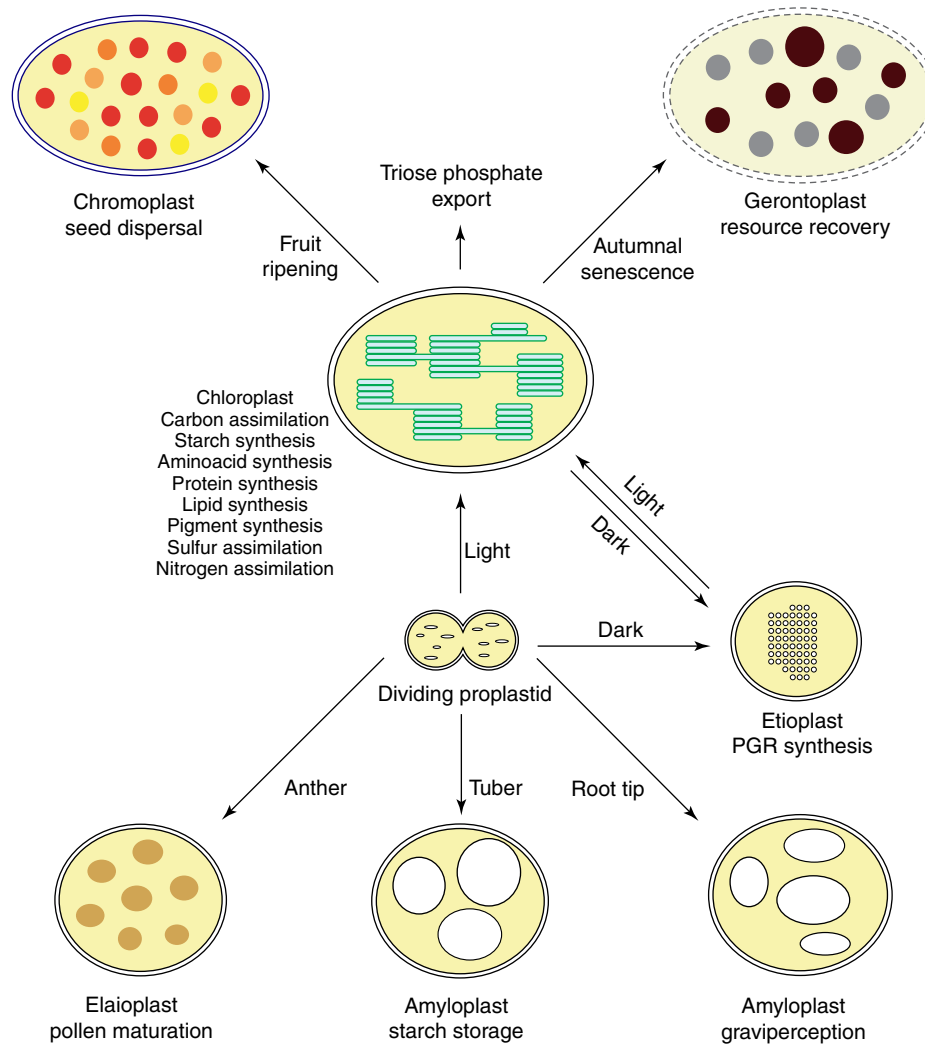


Figure 1 Plastids share a common progenitor and can be interconverted from one form to another. In the anther, proplastids develop into elaioplasts, oil-producing plastids. In the root, some amyloplasts accumulate starch for long-term storage while the amyloplasts in the root cap are involved in the mechanics of gravisensing. Proplastids in stems and shoots develop into etioplasts in the dark, but chloroplasts differentiate directly from proplastids if development takes place in the light. Chloroplasts perform a variety of important physiological and metabolic functions. Chloroplasts in immature tomato and peppers, for example, differentiate into brightly colored chromoplasts during ripening. Resource recovery from senescing leaves is tightly regulated as chloroplasts are converted into gerontoplasts, which are ultimately discarded with the abscised leaf.

releases the elaioplasts which contribute their oils to the exine, the outer-most, waterproof pollen wall. Oil synthesis in seeds is a more complicated process that utilizes plastids (fatty acid synthesis), the ER (lipid assembly and maturation) and special oil bodies (storage) and the trafficking of precursors and products between and among the various compartments (Bates *et al.*, 2013). Plant seed oils provide a significant supply of calories for the world's population.

Amyloplast

Starch is a macropolymer of several thousand repeating glucose units, and cannot be synthesized by animals. All starch synthesis in plants occurs inside a plastid and large starch granules can be found in proplastids, chloroplasts, and amyloplasts. Chloroplasts make short-term transitory starch with the entire pool being synthesized during the day and degraded

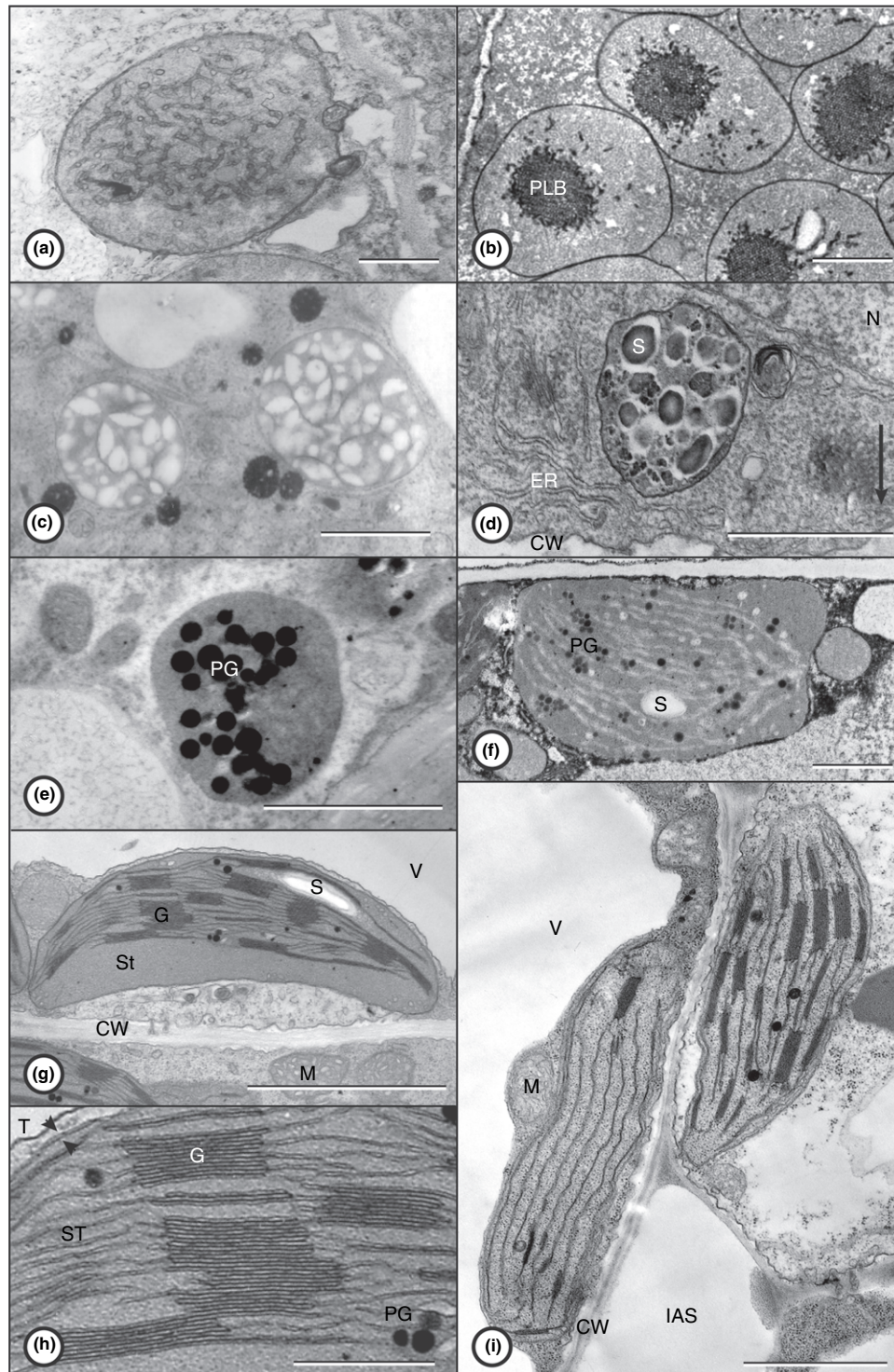
and exported at night. Proplastids and, in particular, amyloplasts make long-term storage starch that may persist for weeks, months or even years. Stored plant starches, like oils (see above), represent a major contribution to human dietary needs.

While most plastid types can contain starch, amyloplasts are unique because of, and named after, the copious amounts of starch they synthesize in the form of large grains (Figure 2(d)). In another example of plastid versatility, amyloplasts play two uniquely different roles in plant physiology. The function of amyloplasts in starch storage has already been noted. A second, separate function is in gravisensing (Palmieri and Kiss, 2006). Starch is heavy and starch-filled amyloplasts sediment on a minute time scale, even within plant cells. Amyloplasts in special tissues in the stem (the endodermis) and the root (the columella of the root cap) perform a mechanical, not a metabolic, function as they sink to the

bottom of the cell and signal an upper/lower cell polarity that initiates a gravitropic growth response (Toyota *et al.*, 2013). Shoots exhibit negative gravitropism and grow away from the signal, or up, while roots exhibit positive gravitropism and grow towards the signal, down.

Chromoplast

Chromoplasts contribute the bright red, orange, and yellow colors to many fruits, colors needed to attract and conscript animals to act as seed dispersers (Bouvier and Camara, 2006).



They originate as chloroplasts and the chloroplast-to-chromoplast transition (Figure 1) is tightly coordinated with the complex process of fruit ripening (Bramley, 2002; Kilcrease *et al.*, 2013).

The colors of chromoplasts come from large accumulations of carotenoid pigments of which there are two types, carotenes (carbon and hydrogen only) and xanthophylls (C, H plus oxygen) (Cuttriss *et al.*, 2006). In addition to functioning as animal visual attractants, the carotenoids are also precursors for Vitamin A biosynthesis in animals, thus rewarding seed dispersers with an essential nutrient. During the chloroplast-to-chromoplast transition, the chlorophyll-containing thylakoid membranes of the chloroplast are degraded and replaced with outsized pigment-protein droplets (Figure 2(e)). The protein fibrillin, which has many other functions in plant cells (Heinrich and Grossman, 2013), helps maintain the structure of the large hydrophobic accumulations (Singh and McNellis, 2011).

Gerontoplast

Leaf cells contain thousands of individual chloroplasts, and each chloroplast is rich in the proteins needed for photosynthesis. Autumnal senescence ('fall colors') is a genetically programmed, step-by-step dismantling of the chloroplast with the sole purpose of recovering that leaf protein. Cell walls and chloroplast lipids, including the lipid chlorophyll, are not recovered. As senescence proceeds, thylakoid membranes are degraded, lipids and pigments form large droplets and the chloroplast is slowly converted to a 'gerontoplast' (Figure 2(f); Harris and Arnott, 1973). Gerontoplasts have few internal membranes and large lipid/pigment droplets called plastoglobuli. They represent the catabolic end product of resource recovery (Krupinska, 2006).

Autumnal senescence has four steps. (1) Carbon accumulation: Warm autumn days promote photosynthesis and the production of leaf starch. Cool autumn nights, however, inhibit the translocation of that starch (as low-molecular-weight carbon compounds) out of the leaf via the phloem of the vasculature system. Therefore, in the fall leaves of deciduous species become overloaded with starch and photosynthesis is inhibited. (2) Carbon diversion (color formation): In order to remove the starch and restore the basal level of photosynthesis needed to support step 3, leaf carbon is converted into pig-

ments such as orange/red carotenoids. These lipids accumulate in large, hydrophobic droplets within the plastid. (3) Protein removal and chlorophyll detoxification: All chlorophyll in a chloroplast is safely bound to pigment-binding proteins, in order to direct and control energy absorption and processing. During senescence proteases and lipases degrade the thylakoid membrane, releasing chlorophyll. Free chlorophyll is a very potent photodynamic toxin and must be detoxified by chlorophyllases, with the resulting catabolites being shuttled into the vacuole. In short, the green pigments are removed and the thylakoid lipids join the carotenoid in the lipid droplets. (4) Leaf abscission: When all the proteins has been catabolised and translocated out of the leaf, a layer of cork cells forms where the leaf petiole attaches to the stem. Cork cells are weak, and they are waterproof. Eventually, the layer breaks, the leaf falls to the ground, and every hole that might have been left behind is plugged with a little bit of cork.

Chloroplast

Chloroplast structure

The typical C3 chloroplast is a round, plano-convex organelle approximately 5–10 μm in diameter (see below for definition of the C3 photosynthetic pathway). It has a double-membrane envelope and a third, internal system of membranes, thylakoids, that is suspended in a protein-rich fluid, the stroma (Figure 2(g)). In a functional chloroplast, thylakoids form a closed, flattened volume, with the stroma to the outside and the lumen to the inside. Thylakoids are of two varieties. Granal thylakoids are appressed together in stacks called grana; an individual granum may contain 2–30 or more thylakoids. Stromal thylakoids are unappressed and span the stroma to interconnect grana (Figure 2(h)). The three-dimensional architecture of this complicated membrane system appears to have a quasihelical structure (Mustárdy *et al.*, 2008), although due to the dynamic nature of the system, and technical difficulties, some refinement of that model may yet result (Daum and Kühlbrandt, 2011). Plants utilizing the C4 photosynthetic pathway have dimorphic chloroplasts (see below).

Chloroplast functions

Because chloroplasts have a constant supply of reduced carbon, NADPH, and ATP and originated from free-living

Figure 2 Common plastid types. Scale bar=5 μm in all images. (a) Proplastid (*Ipomea batatas* – sweet potato – root). This proplastid from a has only a few, disorganized internal membranes. It would probably have differentiated into an amyloplast, given it's location in the root. Scale bar=1 μm . (b) Etioplast (*Avena sativa* – oat – leaf). Large prolamellar bodies (PLB) occupy the majority of the etioplast interior. Scale bar=2 μm . (c) Elaioplast (*Brassica napus* – oilseed rape – tapetum). The plastid interior is almost completely filled with light-colored oil droplets. This oil reserve will be used to waterproof the developing pollen grains prior to their release. Scale bar=2 μm . (d) Amyloplast (*Vicia faba* – mung bean – root tip). An amyloplast from mung bean (*Vicia faba*) root tip. This gravisensing amyloplast has copious amounts of starch (S). Portion of the nucleus (N) is in the top right, and ER tubules lie between the amyloplast and cell wall (CW). The gravitational vector in this image is indicated by the arrow. Scale bar=2 μm . (e) Chromoplast (*Solanum esculentum* – tomato – fruit). Few internal membranes remain in this chromoplast and large pigment/protein/lipid droplets (PG) fill the chromoplast interior. Scale bar=2 μm . (f) Gerontoplast (*Nicotiana tabacum* – tobacco – leaf). A gerontoplast from tobacco has little starch (S), few internal membranes and numerous plastoglobuli (PG). Scale bar=2 μm . (g) C3 chloroplast (*Spinacia oleracea* – spinach – leaf). This typical C3 chloroplast is from a spinach (*Spinacia oleracea*) leaf. It contains grana (G), one starch grain (S) and an area of stroma (St). Two mitochondria (M) can be seen in the lower right of image, below the cell wall (CW). Scale bar=5 μm . (h) Grana (*Spinacia oleracea* – spinach – leaf). Several individual grana (G) are interconnected via stromal thylakoids (ST). Plastoglobuli (PG) can be seen in the lower right corner. The two membranes of the chloroplast envelope are indicated between the two arrowheads. T=tonoplast (vacuole membrane). Scale bar=1 μm . (i) Granal and agranal C4 chloroplasts (*Zea mays* – maize – leaf). Portions of two cells are shown, a bundle sheath cell to the left contains an agranal chloroplast and a mesophyll cell to the right contains a granal chloroplast. A cell wall (CW) separates the two cells. Scale bar=5 μm . IAS, intercellular air space, M=mitochondrion, V=vacuole.

prokaryotic endosymbionts, they have evolved over time to be capable of performing a large array of anabolic functions in a plant cell. Some of those functions are given here.

Photosynthesis and starch synthesis

Chloroplasts use photosynthesis to manufacture low-molecular-weight, reduced carbon compounds, commonly called sugars. In brief, photosynthesis can be divided into two distinct sets of reactions (Hoober, 2006). The 'light-dependent reactions' harvest light energy and use that energy to transport electrons through an electron transport chain embedded in the thylakoid membrane. Chlorophyll is the primary photosynthetic pigment; hence, thylakoid membranes are deep green in color. The light-dependent reactions synthesise ATP and the reductant NADPH. Subsequently, the 'light-independent reactions' use that NADPH and ATP to reduce and phosphorylate oxidized atmospheric carbon to the level of a sugar phosphate.

The 'light-dependent reactions' have an electron transport chain consisting of two photosystems, PSII and PSI, connected by a series of redox-active lipids, proteins, and metal cofactors (Figure 3). Each photosystem has as its core a reaction center (RC) surrounded by antenna complexes containing 300–400 molecules of chlorophyll, several molecules of carotenoid, and numerous pigment-binding proteins. When a chlorophyll molecule absorbs the energy of a photon, that energy is transferred from chlorophyll to chlorophyll through the antenna complex until it is ultimately absorbed by the reaction centers of PSII (P680) or PSI (P700). Note that the antenna does not contain a redox chain – electrons are not being transported, only the absorbed energy. The RC is made of two special chlorophyll molecules that are capable of performing a 'photoact,' i.e. undergoing oxidation and reducing the electron transport chain. The electron ejected from P680 of PSII ultimately reduces a molecule of the lipid plastoquinone (PQ). When doubly reduced with two electrons and two protons, the resulting PQH₂ debinds from PSII, diffuses to and reduces the

cytochrome *b₆/f* complex (Cyt *b₆/f*). The oxidized P680 becomes re-reduced by drawing an electron out of the adjacent water splitting apparatus (WSA). When four electrons have been extracted by PSII (requiring the absorbed energy of four photons), four protons and one molecule of O₂ are released by the WSA (Figure 3 is drawn based on a stoichiometry of two electrons per one NADPH). The absorption of a photon by a PSI antenna chlorophyll drives a second photoact, ultimately reducing NADP⁺ to NADPH. Oxidized PSI extracts an electron from plastocyanin (PC), which is re-reduced by electrons from the cytochrome *b₆/f* complex. Thus, whole-chain photosynthetic electron transport uses the energy of sunlight to oxidize low energy water, and reduce high-energy NADPH.

ATP is also synthesized by the light-dependent reactions, via chemiosmosis. The transthylakoid pH gradient needed to drive chemiosmosis is generated by two mechanisms: (1) the oxidation of one molecule of water produces two free protons, and (2) proton pumping via the PQ pool. When electrons coming from PSII reduce plastoquinone, the two protons needed to complete the reaction come from the stromal side of the thylakoid membrane. Then, when the resultant PQH₂ is oxidized, two electrons enter the cytochrome *b₆/f* complex the two protons are delivered to the thylakoid lumen. In this way, photosynthetic electron transport generates protons in the lumen and pumps protons across the thylakoid membrane, lowering the luminal pH from roughly 7 in the dark to approximately 5 in the light. This substantial pH gradient is used to synthesise ATP in the light.

The light-independent reactions are located in the stroma and use the ATP and NADPH produced by the thylakoid membrane to reduce and phosphorylate atmospheric CO₂ to the level of glyceraldehyde-3-phosphate (G3P). Photosynthesis does not make glucose as an end product. It makes G3P. The resulting triose phosphate can be exported from the chloroplast and be translocated to all parts of the plant. Alternatively, if photosynthesis is running faster than translocation,

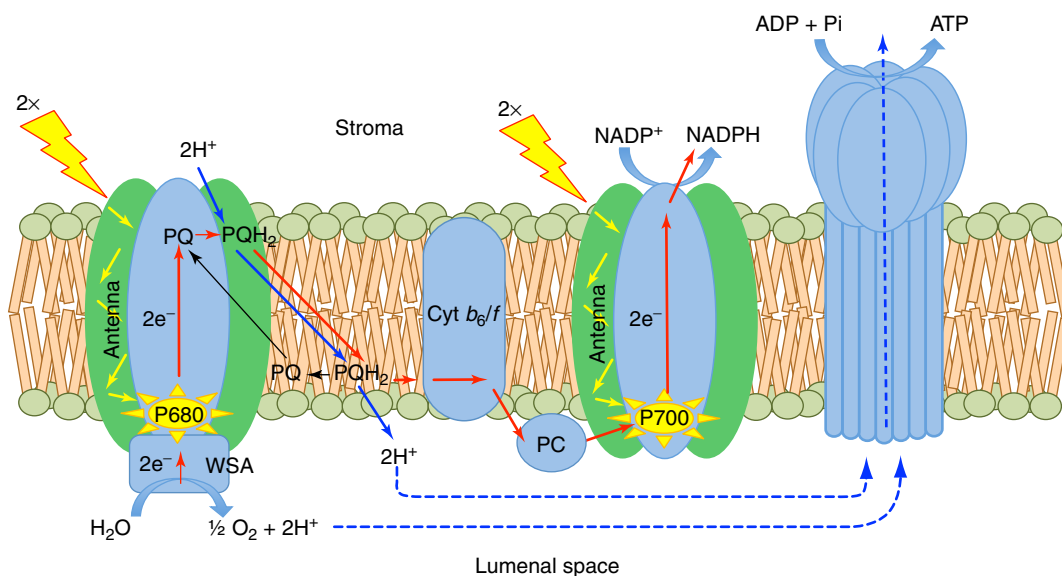


Figure 3 Photosynthetic electron transport. See text for explanation.

the G3P is temporarily stored within the chloroplast as starch. Starch levels are typically high at the end of the light period, and low at pre-dawn. 'Photosynthesis' literally means to use light to fuel an anabolic process. In that regard, photosynthesis makes carbohydrates, amino acids, proteins, lipids, and nucleic acids – every reduced-carbon molecule in the plant.

Nitrogen assimilation, sulfur assimilation, amino acid synthesis, and protein synthesis

Nitrogen and sulfur are 2 of the 20 essential elements plants need for growth and development. Both are rather scarce in the environment, at least in bioavailable forms. While different pathways are responsible for their uptake (and will not be discussed here), once in the plant both elements must be assimilated into safe, transportable, and usable forms such as amino acids (Lancien *et al.*, 2006; Pilon-Smits and Pilon, 2006).

As a class of chemical compounds, there are about 500 amino acids, of which only 20 are used to make proteins. These are the so-called proteinogenic (or proteinaceous) amino acids. Humans can only make 11 of the 20 proteinogenic amino acids. In contrast, plants can synthesize all 20 proteinogenic amino acids and plastids are the site for the synthesis for the nine essential amino acids (Lancien *et al.*, 2006). Lacking plastids, humans must acquire these nine 'essential' amino acids via their diet, by eating plants, or other animals that feed on plants. Glyphosate (Roundup®) is a potent inhibitor of a key plastid enzyme in the aromatic amino acid pathway (phenylalanine, tyrosine, and tryptophan) and, thus, a powerful and specific herbicide (Amrhein *et al.*, 1980). Synthesis of the branched chain amino acids (isoleucine, leucine, and valine) is also plastid localized and the subject of specific herbicide inhibition (Lancien *et al.*, 2006).

Lipid and pigment synthesis

Plastids, in one form or another, synthesize chlorophylls (Willows, 2006), carotenoids (Cuttriss *et al.*, 2006), fatty acids, phospholipids, sulfolipids and galctolipids (Dörmann, 2006), tocopherols (Lushchak and Semchuk, 2012), and quinone lipids (Jones *et al.*, 2013; Kim *et al.*, 2008). Plant cells only rarely possess smooth endoplasmic reticulum, the site of most lipid synthesis in animals.

Chloroplast variants

The dimorphic chloroplasts of C4 photosynthesis

The photosynthetic pathway described above results in the generation of a three-carbon compound, glyceraldehyde-3-phosphate. Hence, plants that use this pathway are called C3 plants. Other plants, most notably grasses such as maize and sugar cane, produce a four-carbon acid as the first stable product of photosynthesis. That form of photosynthesis is called the C4 pathway.

C4 plants differ from C3 plants on biochemical, physiological, anatomical, and ultrastructural bases (Sage, 2004). In C4 photosynthesis, carbon dioxide is initially fixed in chloroplasts in the mesophyll cells of the leaf. Mesophyll cell chloroplasts contain grana stacks, and are termed 'granal' (Figure 2(i), right side). The four-carbon compound resulting from the initial fixation, is transported to the bundle sheath

cells, where it is decarboxylated to a three-carbon molecule and CO₂. The released CO₂ is then fixed again in the bundle sheath chloroplast which lacks grana ('agranal') (Figure 2(i), left side). The different ultrastructures arise because the enzyme responsible for the secondary fixation in the bundle sheath chloroplast (ribulose-bis-phosphate carboxylase/oxygenase or RubisCO) is inhibited in the presence of oxygen. Unlike C3 chloroplasts and C4 mesophyll cell chloroplasts, the C4 agranal bundle sheath chloroplasts are non-oxygenic, thus the O₂ inhibition of RubisCO is alleviated.

Guard cell chloroplasts

Stomata (pl., stoma=sing.) are minute, adjustable pores on the leaf surface that allow for CO₂, H₂O, and O₂ gas exchange between the leaf and the atmosphere. A pair of epidermal cells called guard cells inflate with water, bend and open the stomatal pore. Subsequent loss of water causes the guard cells to deflate, and close the pore. In most plants stomata are closed at night (when photosynthesis cannot operate) and open during the day and the signals to open and close come from environmental cues and an internal circadian clock. The energy to drive the opening and closing comes from guard cell chloroplasts (GCC). The mechanism is based on the movement of osmolytes that influence cell water potential and osmotically driven water movements.

GCC are, in all respects, the same as C3 chloroplasts. They contain thylakoids, grana, chlorophyll, and all the components needed to perform both the light-dependent and the light-independent reactions of photosynthesis (Outlaw *et al.*, 1981; Fitzsimons and Weyers, 1983). They fix carbon and make starch (Gotow *et al.*, 1988). What sets GCCs apart from other chloroplasts is that they contribute nothing to the carbon economy of the leaf. GCCs function solely to provide the energy and osmotica needed to drive stomatal opening and closing.

GCC are an example of the evolutionary flexibility of plastids. Plants evolved from green algal ancestors and began land colonization during the Ordovician period (approximately 470 million years ago). One of the many adaptations needed for a terrestrial life was a waxy, water impervious outer cuticle which drove the need for stomatal openings in that cuticle. As guard cells evolved, plastids were enlisted to serve as the power plant for stomatal functioning. They are sensitive to light (and light is a major signal for stomatal opening) and can generate the low-molecular-weight osmolytes needed to support osmotically driven stomatal opening and closing.

Sun versus shade chloroplasts

Chloroplast development is light-dependent (Apel *et al.*, 1983) and different ultrastructures and physiologies can result as a consequence of the light environment during development (Lichtenthaler *et al.*, 1981). Because light is limiting in the shade, chloroplasts in leaves that develop in the interior of a tree canopy develop to favor the light-dependent reactions over the light-independent reactions. Shade-type chloroplasts are optimized for light capture with larger grana, more thylakoids per granum and more chlorophyll per antenna complex. Stromal volume is concomitantly reduced. Sun-type chloroplasts, which develop at the exterior of the canopy where light is not limiting, have smaller grana, less chlorophyll (reduced

capacity for the light-dependent reactions) and more stromal volume (increased capacity for the light-independent reactions). Both chloroplast types can be found on the same plant, thus these two different developmental outcomes are driven solely by the light environment experienced during leaf expansion.

References

- Amrhein, N., Dues, B., Gehrke, P., Steinrücken, H.C., 1980. The site of inhibition of the Shikimate pathway by glyphosate. 2. Interference of glyphosate with chorismate formation *in vivo* and *in vitro*. *Plant Physiology* 66, 830–834.
- Apel, L., Gollmer, I., Batschauer, A., 1983. The light-dependent control of chloroplast development in barley (*Hordeum vulgare* L.). *Journal of Cellular Biochemistry* 23, 181–189.
- Bates, P.D., Stymne, S., Ohlrogge, J., 2013. Biochemical pathways in seed oil synthesis. *Current Opinion in Plant Biology* 16, 358–364.
- Bouvier, F., Camara, B., 2006. The role of plastids in ripening fruits. In: Wise, R.R., Hooper, J.K. (Eds.), *The Structure and Function of Plastids*. Amsterdam: Springer, pp. 419–432.
- Bramley, P.M., 2002. Regulation of carotenoid formation during tomato fruit ripening and development. *Journal of Experimental Botany* 53, 2107–2113.
- Cuttriss, A.J., Mimica, J.L., Pogson, B.J., Howitt, C.A., 2006. Carotenoids. In: Wise, R.R., Hooper, J.K. (Eds.), *The Structure and Function of Plastids*. Amsterdam: Springer, pp. 315–334.
- Daum, B., Kühlbrandt, W., 2011. Electron tomography of plant thylakoid membranes. *Journal of Experimental Botany* 62, 2393–2402.
- Dörmann, P., 2006. Lipid synthesis, metabolism and transport. In: Wise, R.R., Hooper, J.K. (Eds.), *The Structure and Function of Plastids*. Amsterdam: Springer, pp. 335–353.
- Fitzsimons, P.J., Weyers, J.D.B., 1983. Separation and purification of proplastid types from *Commelina communis* L. leaf epidermis. *Journal of Experimental Botany* 34, 55–66.
- Gotow, K., Taylor, S., Zeiger, E., 1988. Photosynthetic carbon fixation in guard cell protoplasts of *Vicia faba* L. evidence from radiolabeled experiments. *Plant Physiology* 86, 700–705.
- Gray, M.W., Burger, G., Lang, B.F., 1999. Mitochondrial evolution. *Science* 283, 1476–1481.
- Gunning, B.E.S., 2001. Membrane geometry of 'open' prolamellar bodies. *Protoplasma* 215, 4–15.
- Harris, J.B., Arnott, H.J., 1973. Effects of senescence on chloroplasts of the tobacco leaf. *Tissue and Cell* 5, 527–544.
- Heinrich, M.L., Grossman, A.R., 2013. The GreenCut: Re-evaluation of physiological role of previously studied proteins and potential novel protein functions. *Photosynthesis Research* 116, 427–436.
- Hernandez-Pinzon, I., Ross, J.H.E., Barnes, K.A., Damant, A.P., Murphy, D.J., 1999. Composition and role of tapetal lipid bodies in the biogenesis of the pollen coat of *Brassica napus*. *Planta* 208, 588–598. Amsterdam: Springer.
- Hooper, J.K., 2006. Chloroplast development: Whence and whither. In: Wise, R.R., Hooper, J.K. (Eds.), *The Structure and Function of Plastids*. Amsterdam: Springer, pp. 27–51.
- Jones, M.O., Perez-Fons, L., Robertson, F.P., Bramley, P.M., Fraser, P.D., 2013. Functional characterization of long-chain prenyl diphosphate synthases from tomato. *Biochemical Journal* 449, 729–740.
- Keeling, P.J., 2010. The endosymbiotic origin, diversification and fate of plastids. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365, 729–748.
- Kilcrease, J., Collins, A.M., Richins, R.D., Timlin, J.A., O'Connell, M.A., 2013. Multiple microscopic approaches demonstrate linkage between chromoplast architecture and carotenoid composition in diverse *Capsicum annuum* fruit. *Plant Journal* 76, 1074–1083.
- Kim, H.U., van Oostende, C., Basset, G.J.C., Browse, J., 2008. The AAE14 gene encodes the Arabidopsis o-succinylbenzoyl-CoA ligase that is essential for phylloquinone synthesis and photosystem-I function. *Plant Journal* 54, 272–283.
- Kirk, J.T.O., Tilney-Bassett, R.A.E., 1967. *The Plastids, Their Chemistry, Structure, Growth and Inheritance*. London: WH Freeman and Co.
- Krupinska, K., 2006. Fate and activities of plastids during leaf senescence. In: Wise, R.R., Hooper, J.K. (Eds.), *The Structure and Function of Plastids*. Amsterdam: Springer, pp. 433–449.
- Lancien, M., Lea, P.J., Azevedo, R.A., 2006. Amino acid synthesis in plastids. In: Wise, R.R., Hooper, J.K. (Eds.), *The Structure and Function of Plastids*. Amsterdam: Springer, pp. 355–385.
- Lichtenthaler, H.K., Buschmann, C., Döll, M., *et al.*, 1981. Photosynthetic activity, chloroplast ultrastructure, and leaf characteristics of high-light and low-light plants and of sun and shade leaves. *Photosynthesis Research* 2, 115–141.
- Lushchak, V.I., Semchuk, N.M., 2012. Tocopherol biosynthesis: Chemistry, regulation and effects of environmental factors. *Acta Physiologia Plantarum* 34, 1607–1628.
- Miyagishima, S.-y., Kuroiwa, T., 2006. The mechanism of plastid division: The structure and origin of the plastid division apparatus. In: Wise, R.R., Hooper, J.K. (Eds.), *The Structure and Function of Plastids*. Amsterdam: Springer, pp. 103–121.
- Mustárdy, L., Buttle, K., Steinbach, G., Garab, G., 2008. The three-dimensional network of the thylakoid membranes in plants: Quasihelical model of the granum-stroma assembly. *Plant Cell* 20, 2552–2557.
- Outlaw, W.H., Mayne, B.C., Zenger, V.A., Manchester, J., 1981. Presence of both photosystems in guard cells of *Vicia faba* L.: Implications for environmental signal processing. *Plant Physiology* 67, 12–16.
- Palmieri, M., Kiss, J.Z., 2006. The role of plastids in gravitropism. In: Wise, R.R., Hooper, J.K. (Eds.), *The Structure and Function of Plastids*. Amsterdam: Springer, pp. 507–525.
- Pilon-Smits, E.A.H., Pilon, M., 2006. Sulfur metabolism in plastids. In: Wise, R.R., Hooper, J.K. (Eds.), *The Structure and Function of Plastids*. Amsterdam: Springer, pp. 387–402.
- Sage, R.F., 2004. The evolution of C4 photosynthesis. *New Phytologist* 161, 341–370.
- Sato, N., 2006. Origin and diversification of plastids: Genomic view on the unification and diversity of plastids. In: Wise, R.R., Hooper, J.K. (Eds.), *The Structure and Function of Plastids*. Amsterdam: Springer, pp. 75–102.
- Singh, D.K., McNellis, T.W., 2011. Fibrillin protein function: The tip of the iceberg? *Trends in Plant Science* 16, 432–441.
- Solymosi, K., Martinez, K., Kristof, Z., Sundqvist, C., Boddie, B., 2004. Plastid differentiation and chlorophyll biosynthesis in different leaf layers of white cabbage (*Brassica oleracea* cv capitata). *Physiologia Plantarum* 121, 520–529.
- Stiller, J.W., 2014. Toward an empirical framework for interpreting plastid evolution. *Journal of Phycology* 50, 462–471.
- Suzuki, T., Tsunekawa, S., Koizuka, C., *et al.*, 2013. Development and disintegration of tapetum-specific lipid-accumulating organelles, elaioplasts and tapetosomes, in *Arabidopsis thaliana* and *Brassica napus*. *Plant Science* 207, 25–36.
- Toyota, M., Ikeda, N., Sawai-Toyota, S., *et al.*, 2013. Amyloplast displacement is necessary for gravisensing in Arabidopsis shoots as revealed by a centrifuge microscope. *Plant Journal* 76, 648–660.
- Willows, R.D., 2006. Chlorophyll synthesis. In: Wise, R.R., Hooper, J.K. (Eds.), *The Structure and Function of Plastids*. Amsterdam: Springer, pp. 295–313.
- Wise, R.R., 2006. The diversity of plastid form and function. In: Wise, R.R., Hooper, J.K. (Eds.), *The Structure and Function of Plastids*. Amsterdam: Springer, pp. 3–26.
- Zhong, S., Shi, H., Xue, C., *et al.*, 2014. Ethylene-orchestrated circuitry coordinates a seedling's response to soil cover and etiolated growth. *Proceedings of the National Academy of Sciences of the United States of America* 111, 3913–3920.

Further Reading

Wise, R.R., Kenneth Hooper, J., 2006. *The Structure and Function of Plastids*. Amsterdam: Springer, 576 pp.

INTERORGANELLAR COMMUNICATION: INTERPLAY AND PROCESSES

Contents

ER–Golgi Transport

N-Linked Glycans (N-Glycans)

Extracellular O-Glycans

Intra-Golgi Transport

Post-Golgi Transport – Cargo, Carriers, and Pathways

Regulation of the Secretory Pathway

Regulated versus Constitutive Secretion – A Major Form of Intercellular Communication

Clathrin and Clathrin-Dependent Endocytosis

Clathrin Independent Endocytosis

Rabs of the Endosomal Recycling Pathway

Endosome to Lysosome Transport

Endocytosis of Cargo Proteins: LDL

Retrograde Transport

Rabs and Other G Proteins

ER–Golgi Transport

S Pagant and EA Miller, Columbia University, New York, NY, USA

© 2016 Elsevier Inc. All rights reserved.

Accurate and efficient trafficking of lipids and proteins is fundamental for cell growth, communication, and response to environmental conditions. Secretory and membrane proteins represent up to a third of the proteome in the model eukaryote, *Saccharomyces cerevisiae*. These proteins are synthesized and matured in the endoplasmic reticulum (ER), then subsequently captured into vesicles fashioned by COPII coat proteins in the first step of their journey to their final destination.

The COPII Coat

The formation of ER-derived COPII vesicles results from the sequential recruitment of the minimal machinery composed of five individual coat proteins: Sar1, Sec23, Sec24, Sec13, and Sec31 (Figure 1; Barlowe *et al.*, 1994). These components are sufficient to generate vesicles from synthetic membranes in the absence of other proteins (Matsuoka *et al.*, 1998), but in cells various interacting partners (e.g., Sec12 and Sec16) serve to regulate this process. Activation of the small G protein, Sar1, by the ER membrane localized guanine nucleotide exchange factor (GEF), Sec12, induces a structural change upon exchange of GDP for GTP that exposes an amphipathic α -helix that inserts into the ER membrane (Yoshihisa *et al.*, 1993).

Sar1.GTP initiates membrane curvature (Bielli *et al.*, 2005; Lee *et al.*, 2005) and recruits the coat inner layer (Antonny *et al.*, 2001), composed of the heterodimeric complex of Sec23 and Sec24, through direct interaction with Sec23 (Bi *et al.*, 2002). Finally, the heterotetrameric complex of Sec13 and Sec31, referred as the outer coat layer, is recruited via a complex set of interactions with the ‘prebudding’ complex Sar1/Sec23/Sec24 (Antonny *et al.*, 2001; Bi *et al.*, 2007). Purified Sec13/Sec31 is capable of forming spherical assemblies that resemble the diameter of COPII vesicles (Stagg *et al.*, 2006), suggesting that the outer coat contributes to membrane deformation and establishes the geometry of the vesicle.

Cargo Capture during Vesicle Formation

Efficient incorporation of cargo proteins into COPII vesicles is triggered by the cytosolic presentation of protein-based signals, referred as sorting signals or export motifs, that bind to discrete cargo-recognition sites on the surface of Sec24 (Miller *et al.*, 2003; Mossessova *et al.*, 2003). Secretory proteins that cannot bind directly to Sec24, because of lack of signals or topology constraints, make indirect contact with the COPII coat by interacting with membrane proteins that themselves present a binding motif for Sec24 (Dancourt and Barlowe,

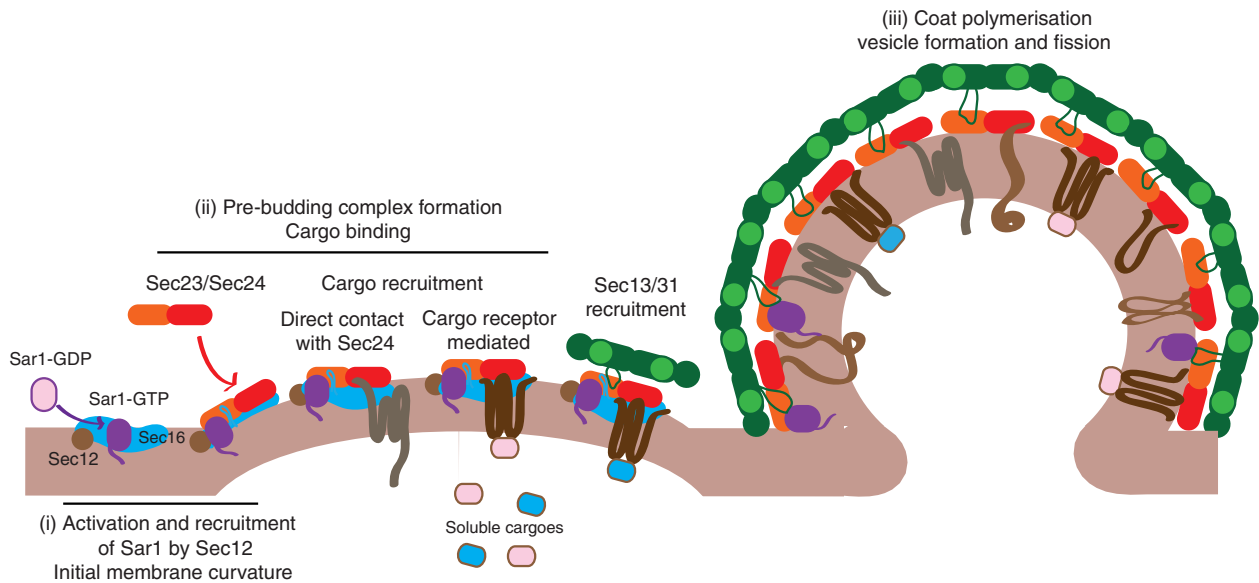


Figure 1 Hierarchical assembly of the COPII coat. Vesicle formation proceeds via defined steps: (i) The G-protein, Sar1, is locally activated by its GEF, Sec12, which drives membrane insertion of an amphipathic α -helix that initiates membrane curvature; (ii) GTP-loaded Sar1 recruits the cargo-binding adapter complex, Sec23/Sec24, which binds to specific sorting signals on cargo proteins (gray/brown proteins) or cargo receptors (blue proteins); (iii) this 'prebudding complex' recruits the outer coat layer, Sec13/Sec31, which polymerizes into a spherical structure thereby driving membrane curvature and completing vesicle formation.

2010). This class of proteins, known as receptor proteins, includes receptors such as yeast Erv29 (Belden and Barlowe, 2001) and mammalian ERGIC-53 (Appenzeller *et al.*, 1999) that bind to soluble luminal cargoes. Some classes of membrane proteins, like GPI-anchored proteins, are embedded in the bilayer by a luminal lipid anchor and thus are topologically unable to bind directly to the COPII coat. Such cargo proteins must also employ receptors, in the case of GPI-anchored proteins, the p24 family of proteins mediate efficient ER export (Muniz *et al.*, 2000). Additional cargo receptors, like Erv26 and Erv14 (Dancourt and Barlowe, 2009; Herzig *et al.*, 2012; Powers and Barlowe, 2002), mediate efficient export of transmembrane proteins even though such cargoes can potentially interact directly with Sec24. Recent work on Erv14 suggests that this receptor serves to strengthen the interaction of a cargo with Sec24 by providing a second signal that must be recognized simultaneously with a cargo-borne signal (Pagant *et al.*, 2015).

The relatively small subset of ER export motifs defined so far fail to conform to a canonical ER export signal, and instead range in character from simple 2–3 amino acid motifs to larger folded epitopes. Moreover, four independent cargo-binding sites on the surfaces of the yeast and mammalian Sec24 isoforms have been structurally and functionally characterized: the so-called yeast A-site binds an N-P-F motif on Sed5 (Miller *et al.*, 2005; Mossessova *et al.*, 2003); the yeast and mammalian B-site is responsible for the packaging of numerous proteins and can recognize different motifs, most notably simple acidic-based signals; the conserved C-site binds a folded epitope formed by the longin domain of the SNARE, Sec22 (Mancias and Goldberg, 2007; Miller *et al.*, 2003); an additional final site on mammalian Sec24 recognizes the mammalian SNARE proteins, membrin and syntaxin5 (Mancias

and Goldberg, 2008). The repertoire of binding site-export motif pairs is further expanded by the existence of multiple Sec24 isoforms (three in yeast and four in mammals), that each possess distinct cargo-binding specificities. As an example, the I-X-M motif of membrin and syntaxin5 binds a surface groove on the mammalian SEC24C and SEC24D isoforms that is occluded in the structure of the SEC24A and SEC24B proteins (Mancias and Goldberg, 2008). Similarly, the packaging of PCSK9 and Vangl2 is specifically mediated by SEC24A and SEC24B (Chen *et al.*, 2013; Merte *et al.*, 2010), whereas the transporters of the SCL6 family favor the SEC24C or SEC24D isoforms (Sucic *et al.*, 2011, 2013). Conversely, any of the four SEC24 isoforms are competent to promote ER export of ERGIC-53 (Merte *et al.*, 2010), a cargo receptor for coagulation factors. This level of specificity might have evolved to finely regulate the deployment of key players of various developmental processes.

Regulation of COPII Biogenesis: The GTPase Cycle

The GTPase cycle of Sar1 is the central regulator of COPII trafficking since it influences multiple steps of vesicle formation. Association of Sar1.GTP with the ER membrane occurs via an N-terminal amphipathic α helix that may initiate curvature on the membrane and/or alter the physical properties of the bilayer to make it more susceptible to deformation (Lee *et al.*, 2005; Loftus *et al.*, 2012; Settles *et al.*, 2010). Sar1.GTP also stimulates the local recruitment of the other COPII proteins at the ER exit sites (ERES). Moreover, vesicle release from the membrane also requires both the N-terminal helix and GTPase activity of Sar1 (Bielli *et al.*, 2005; Lee *et al.*, 2005). Finally, GTP hydrolysis on Sar1 is thought to

destabilize membrane association of the other coat proteins and thereby contribute to vesicle uncoating that precedes fusion with the cis-Golgi acceptor membrane (Antonny *et al.*, 2001; Barlowe *et al.*, 1994). Given these multiple roles of the GTPase cycle of Sar1 in COPII vesicles biogenesis, it is not surprising that a suite of effectors contributes to quantitative and local regulation.

Like most G-proteins, Sar1 possesses very low intrinsic GTPase activity, but basal hydrolysis is enhanced during polymerization of the coat; Sec23 is the GTPase-activating protein (GAP) for Sar1, stimulating GTP hydrolysis by inserting a catalytic arginine residue into the GTP-binding pocket of Sar1 (Bi *et al.*, 2002). Furthermore, binding of an unstructured loop of Sec31 to the Sec23–Sar1 interface repositions this arginine residue, resulting in a 10-fold increase in GTPase activity (Bi *et al.*, 2007). This results in a somewhat counterintuitive scenario whereby coat assembly, and in particular coupling between the inner and outer coat layers, should drive coat depolymerization. Clearly, additional layers of regulation must spatially and temporarily stabilize the coat to complete vesicle biogenesis.

One mechanism that contributes to modulation of the Sar1–GTP cycle is its GEF, Sec12, which catalyzes GTP loading on Sar1 with a turnover 10-fold higher than the coat-stimulated GAP activity (Futai *et al.*, 2004; Sato and Nakano, 2005). Since Sec12 is not packaged into COPII vesicles, its presence in the ER membrane could drive local COPII assembly sufficient for vesicle formation, then its absence from vesicles would shift the balance toward the Sec23-driven GAP activity and uncoating. Consistent with this model, Sec12 is concentrated at ERES through binding with Sec16 (discussed below) in the budding yeast, *Pichia pastoris* (Rossanese *et al.*, 1999; Soderholm *et al.*, 2004), and in mammalian cells (Montegna *et al.*, 2012). However, in *Saccharomyces cerevisiae*, Sec12 is dispersed across the entire ER, and ERES marked by COPII coat proteins are smaller and more numerous, raising questions about the distribution of Sec12 as a primary driver of ERES stability and organization (Rossanese *et al.*, 1999). Interestingly, mammalian cells show ERES recruitment of Sec12 independent of Sec16, but contingent on cTAGE5 (Saito *et al.*, 2014), an ER protein involved in transport of procollagen, a remarkably large cargo that requires carriers of much bigger size than canonical COPII vesicles (discussed further below). Together, the *in vitro* and *in vivo* data suggest that local GEF activity can counter Sar1 GTP hydrolysis, but that spatial restriction of this activity to discrete ERES is not absolutely required for secretion. Instead, ERES enrichment of Sec12 may be critical for packaging of ‘difficult’ cargoes that require larger structures and therefore prolonged coat association.

Multiple lines of evidence delineate a major role for the large ER membrane protein, Sec16, in promoting COPII vesicle biogenesis. However, its mechanism of action in this process is complex and remains to be fully understood. Purified Sec16 can bind *in vitro* to synthetic liposomes and in turn drive the stable recruitment of the COPII coat proteins through multiple interactions (Espenshade *et al.*, 1995; Gimeno *et al.*, 1996; Supek *et al.*, 2002). In *Pichia pastoris* and mammalian cells, Sec16 localizes at ERES with a lifetime higher than that observed for COPII coat proteins. Moreover, the observation that ERES formation is impaired in the absence of Sec16

strongly suggests a role in organization of the COPII coat at specific localizations on the ER membrane (Connerly *et al.*, 2005). Recent findings also indicate that the GTPase stimulatory effect of Sec31 can be attenuated by Sec16 (Kung *et al.*, 2012; Yorimitsu and Sato, 2012). Therefore, it remains unclear to what extent the capacity of Sec16 to maintain ERES structure results from its action in reducing Sar1 GTPase activity, from its binding with Sec12 or from its ability to stabilize coat polymerization by scaffolding the other COPII components.

Cellular Regulation of COPII Vesicle Biogenesis

Despite intricate knowledge of COPII coat structure and function *in vitro*, major questions remain about how the process of vesicle formation is regulated to satisfy specific cellular demands during development, differentiation or stress. These questions are now being explored using large-scale screens. Systematic knock-down analysis of human kinases and phosphatases identified 122 proteins that affect ER export, ERES formation, and/or the Golgi apparatus (Farhan *et al.*, 2010). Among the candidates identified, ERK2 impacted phosphorylation of Sec16 and depletion of this member of the Raf–MEK–ERK cascade led to a reduction in the number of ERES. In a separate study, overexpression of ERK7 negatively impacted ERES formation through the redistribution of Sec16 to the cytosol from ERES (Zacharogianni *et al.*, 2011). Taken together, these results place Sec16 as a major signal integrator, capable of modulating COPII vesicle formation sites in response to cellular conditions. Sec16 is also responsible for mediating changes in ERES number during increases in cargo load at the ER, although the signaling events that modulate these events are not known.

Proteome-wide analysis of posttranslational modifications have identified the COPII coat proteins as targets of phosphorylation, ubiquitylation, and acetylation, but how these changes affect COPII vesicle production and how they relate to specific cellular regulation remain largely unexplored. In a more targeted study, recombinant Akt, whose perturbation is linked to cancer and diabetes, directly phosphorylated mammalian SEC24, which resulted in decreased binding SEC23 (Sharpe *et al.*, 2011). These data suggest that one avenue for cellular signaling to regulate secretory flux is by preventing Sec24 integration into prebudding complexes. Sec23 and Sec24 are also phosphorylated by a CK1 kinase, Hrr25/CK1δ, although these phosphorylation events have been implicated in vesicle delivery rather than in more global regulation of secretory events (Lord *et al.*, 2011).

Additional layers of regulation likely come from a growing number of accessory proteins that interact with the various coat components but contribute to traffic in poorly understood ways. Such proteins include the Sec23- and Sec31-interacting protein p125 (Klinkenberg *et al.*, 2014; Ong *et al.*, 2010; Shimoi *et al.*, 2005), the Sec16-interactor TFG-1 (Witte *et al.*, 2011), and the Ca(2+)–binding protein ALG-2 (Helm *et al.*, 2014). ALG-2 binds specifically to SEC31 in a calcium-dependent manner and its depletion impairs the recruitment of the COPII outer coat at ERES. Moreover, addition of ALG-2 to *in vitro* budding reactions showed that ALG-2 significantly reduced the formation of COPII vesicles, and that its

inhibitory effect was calcium-dependent (la Cour *et al.*, 2013). The mechanistic effects of such accessory proteins and how they may also be points of cellular regulation remain to be more fully explored.

Regulation of Vesicle Size

COPII vesicles produced *in vitro* by the yeast minimal machinery (Sar1, Sec23/Sec24, Sec13/Sec31) range from 60 nm to 80 nm in diameter (Barlowe *et al.*, 1994). However, the secretion of unusually sizable cargoes, such as procollagen fibers and chylomicrons, requires the production of much larger carriers. Recent evidence from cryo-EM structures strongly suggests that the geometry of the outer Sec13/Sec31 complex contributes to shaping carriers from the ER membrane. Purified mammalian Sec13/Sec31 can self-assemble into spherical cages with cuboctahedral geometry, although larger and more irregularly shaped structures can also be observed, suggesting that the outer coat might be adaptable to different vesicle shapes (Noble *et al.*, 2013; Stagg *et al.*, 2006, 2008). Cage assembly in the prevailing cuboctahedral structure is driven by Sec31 oligomerization at two interfaces: tail-to-tail dimerization of two Sec31 molecules via their α -solenoid domains; and 4-fold interaction via N-terminal β -propeller domains (Fath *et al.*, 2007; Noble *et al.*, 2013). In support for a fundamental role of Sec31 in sculpting COPII carriers, monoubiquitylation of mammalian Sec31 induces the formation of distended COPII structures at ERES and is specifically essential for ER export of procollagen fibers (Jin *et al.*, 2012). Sec13 forms a partial β -propeller that inserts between the α -solenoid and β -propeller domains of Sec31, and is thought to provide rigidity to the cage. Deleting a flexible hinge region of yeast Sec31 that encompasses the Sec13 binding site complements the lethality associated with the absence of Sec13 in yeast (Copic *et al.*,

2012). Similarly, the essentiality of Sec13 could be rescued by inhibiting the ER export of GPI-anchored proteins, suggesting that cargo burden and/or ER membrane composition influence the cellular function of the outer coat.

The inner coat layer, composed of Sec23 and Sec24, has also been evoked to provide structural regulation that may determine vesicle shape and size. Cryo-EM analysis revealed that in the presence of Sec23/Sec24, Sec13/Sec31 assembly generates icosidodecahedral cages of ~ 100 nm in diameter (Stagg *et al.*, 2008). Furthermore, cryo-EM of the yeast COPII coat assembled on synthetic liposomes revealed long membrane tubules that were densely coated with oligomeric arrays of Sec23/Sec24 (Zanetti *et al.*, 2013). One emerging model is that the positioning of Sec23/Sec24 can influence the geometry of Sec13/Sec31 and thereby alter the architecture of the underlying membrane. In further support of a role for the inner coat in modulating vesicle size, vesicles produced *in vitro* with the yeast Sec24 isoform Sfb2/Lst1 are larger than the standard COPII vesicles (90 nm for Lst1 vs 60 nm for Sec24) (Shimoni *et al.*, 2000). Interestingly, Lst1 is required for traffic of Pma1 oligomers and GPI-anchored proteins, both of which associate with ceramide-rich lipids and may correspond to cargoes that require more spacious carriers (Lee *et al.*, 2002; Peng *et al.*, 2000; Roberg *et al.*, 1999). Together these data suggest that Sec24 isoforms might influence the diameter of the COPII vesicles and/or cages in response to the specific cargo with which they interact.

Sec23, the other inner coat protein, has been more directly implicated in the production of COPII carriers that transport large cargo proteins. Point mutations in conserved residues of human SEC23A have been linked with a craniofacial disorder in which bone abnormalities are thought to arise from defects in collagen secretion. The first mutation (F382L) linked with this disease changes a residue directly involved in the interaction between Sec23 and Sec31, and presumably leads to defective

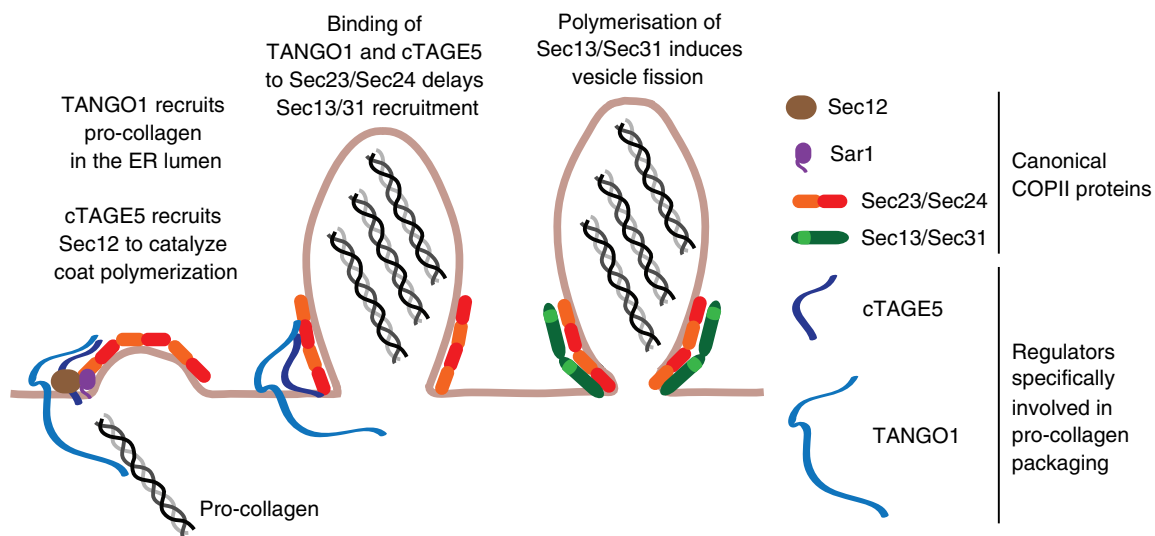


Figure 2 Packaging of noncanonical cargo. Some cargoes, like procollagen and lipid particles, are too large to fit into a standard COPII vesicle, so the budding process must be altered in some way. Specific accessory proteins like TANGO1 and cTAGE5 are thought to specifically alter COPII coat assembly to create conditions that permit procollagen traffic. cTAGE5 binds Sec12 to initiate coat recruitment locally; TANGO1 binds procollagen in the lumen and may link it to these regions of Sec12 activity. One current model is that accessory proteins prevent recruitment of Sec13/Sec31 and thereby prolong the coat assembly phase prior to vesicle release.

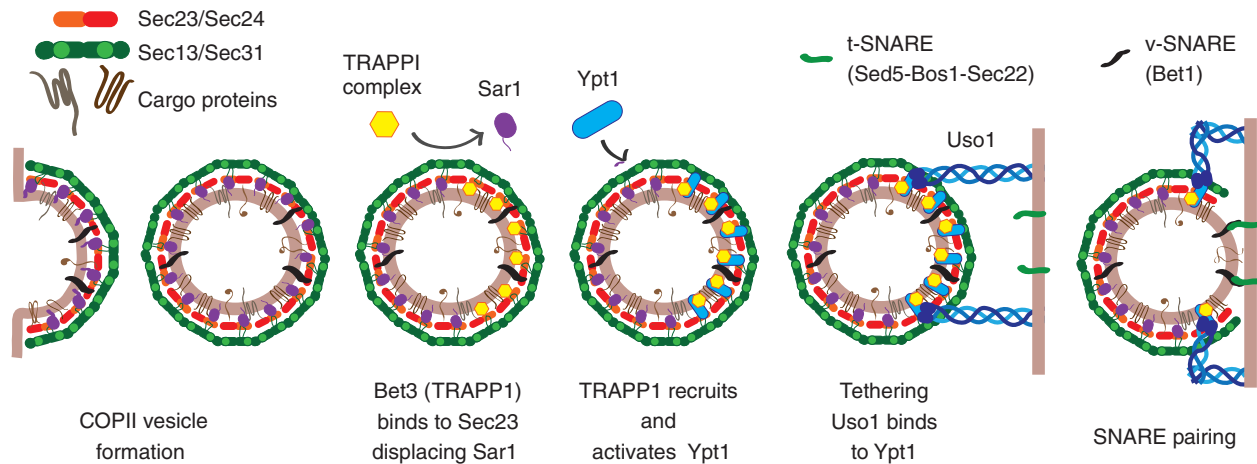


Figure 3 Directionality of vesicle delivery. Once a vesicle has been released from the ER it must be delivered to the appropriate target membrane. This process is driven by sequential interactions of Sec23 with various tethering and fusion regulators. When GTP hydrolysis is complete and Sar1 has been released from the vesicle, the TRAPP component, Bet3, binds to Sec23. TRAPP initiates local activation of Ypt1/Rab1, which in turn recruits coiled-coil tethers like Uso1/p115. Phosphorylation of Sec23 drives TRAPP release, and may prime the vesicle for uncoating, which would expose the SNARE machinery that ultimately drives membrane mixing.

vesicle fission since extended ER cisternae and uncoated tubular projections were observed in patient-derived fibroblasts (Boyadjiev *et al.*, 2006). The precise mechanism by which these mutations affect vesicle production remains to be determined, since the loss of Sec31/Sec13 recruitment might impact the physical capacity for the coat to deform ER membrane but also impact GTPase activity on Sar1. A second disease-associated SEC23A variant (M702V) is specifically defective for collagen transport, since vesicle formation and traffic of most cargo proteins remained unaffected (Kim *et al.*, 2012). Unlike SEC23-F382L, SEC23-M702V does not diminish association with Sec31 but results in elevated SAR1B GTPase activity. In support of a model in which the SAR1B GTPase cycle is regulated to allow the transport of large secretory cargo, efficient packaging of large lipid particles known as chylomicrons also relies specifically on SAR1B activity (Jones *et al.*, 2003).

In addition to the coat itself, further accessory proteins may also contribute to generation of large COPII carriers (Figure 2). One such component is the metazoan ER membrane protein, TANGO1, which interacts with procollagen in the ER lumen and with Sec23/Sec24. Unlike canonical receptors that promote ER exit of luminal cargoes, TANGO1 itself is not incorporated into COPII vesicles (Saito *et al.*, 2009). It has been proposed that binding of these proteins to the inner coat might interfere with recruitment of Sec31 onto Sec23 and therefore restrain the GAP activity associated with the full coat and prevent premature vesicle scission. Moreover, the TANGO1-related protein, cTAGE5 (Saito *et al.*, 2011), associates with Sec12 at ERES, which could contribute additional coat stabilization (Saito *et al.*, 2014). Conversely, Sedlin, another binding partner of TANGO1, acts directly on Sar1 to reduce its conversion to Sar1-GTP (Venditti *et al.*, 2012). Sedlin is also implicated in procollagen trafficking since its deficiency leads to the collagen-related disease, spondyloepiphyseal dysplasia tarda (SED). Therefore, the exit of collagen from the ER may rely on both the activation of the Sar1 GTP cycle by cTAGE5–Sec12 and its inactivation by

TANGO1–Sedlin interactions. How those events are regulated to allow the formation of carriers efficient for the transport of procollagen or other large cargo proteins awaits further investigation. Furthermore, direct evidence addressing a role for GTPase regulation by TANGO1/cTAGE5 has not yet emerged, and additional potential roles in altering the geometry of the inner and outer COPII coat layers or altering membrane curvature have not been ruled out.

Vesicle Tethering

Following their release from ER membranes, COPII vesicles must target to and fuse with Golgi acceptor membranes (Figure 3). In mammalian cells, ER-derived vesicles also undergo homotypic fusion, which likely generates the ER–Golgi Intermediate Compartment (ERGIC), an intermediary sorting station that is also marked with COPI coat proteins. ERGIC membranes undergo long-distance transport to the Golgi proper. The mechanisms of tethering and fusion of ER-derived vesicles are broadly conserved from yeast to mammals, using much of the same machinery. Initial tethering of COPII vesicles seems to involve interaction between Sec23 and the transport protein particle I (TRAPP) complex, which functions as a GEF for the Rab GTPase, Ypt1 (Rab1 in mammals) (Lord *et al.*, 2011). TRAPP is composed of six subunits (Bet3, Bet5, Trs20, Trs23, Trs31, and Trs33) that assemble onto the surface of the COPII vesicles via direct interaction between Bet3 and Sec23. GTP loading on Ypt1/Rab1 recruits the activated Rab1 to the vesicle where it in turn drives recruitment of the coiled-coil tethering protein Uso1 (p115 in mammals) (Cao *et al.*, 1998; Seog *et al.*, 1994a). Structural analysis shows that the amino-terminal domain of p115 forms a dimer between two globular regions that contain Armadillo-like α -helical repeats, termed tether repeats (TRs), which are also found in other tethering factors (Seog *et al.*, 1994b; Yamakawa *et al.*, 1996). The extended coiled-coil structures of both p115 and Uso1 yield an average length of ~ 154 nm, which may be

required for the capture of vesicles at a significant distance away from the Golgi. In addition, the coiled-coil regions of p115/Uso1 can undergo local bending, which may facilitate multidirectional collection of vesicles and/or may indicate conformational change that would physically bring the vesicles close to the acceptor membranes.

Additional cis-Golgi localized proteins, GRASP65 and GM130, are also implicated *in vitro* to contribute to COPII vesicle tethering (Alvarez *et al.*, 2001; Behnia *et al.*, 2007; Nakamura *et al.*, 1995; Seemann *et al.*, 2000), although since their respective yeast orthologs, Grh1 and Bug1, are not essential for cell viability their importance in secretion remains to be fully determined (Levi *et al.*, 2010). Nevertheless, the observation that these proteins physically interact with Uso1 and that their respective deletions display negative genetic interactions with thermosensitive alleles of both Ypt1 and Uso1 suggests that they participate in tethering events perhaps redundantly with the Ypt1/Uso1 pathway (Behnia *et al.*, 2007). Indeed, these tethers connect directly to the COPII coat, binding to the Sec23/Sec24 complex independently of Ypt1. Tethering events that mediate Golgi delivery seem to be coordinated by the kinase, Hrr25/CK1 δ , which phosphorylates Sec23 and triggers release of TRAPPI (Lord *et al.*, 2011). These phosphorylation events may also destabilize the coat to permit local uncoating immediately adjacent to the acceptor membrane, an event that is required for exposure of the vesicle-borne fusion machinery.

Vesicle Fusion

Subsequent to vesicle tethering, the membranes of the acceptor compartment and vesicle must undergo fusion. This process is mediated by type II membrane proteins called SNAREs that share a conserved domain defined as the SNARE motif. SNAREs embedded in the vesicle membrane specifically interact with those located in the acceptor membranes to form 'trans' complexes that bring the opposing lipid bilayers in close proximity, which catalyzes their mixing. SNAREs are sub-categorized as Q-SNAREs or R-SNAREs based on the presence of conserved glutamine or arginine residues within the SNARE motifs. Membrane fusion generally requires the zipping of three Q-SNAREs (Qa, Qb, and Qc) and one R-SNARE. Vesicle fusion can be reconstituted with synthetic proteoliposomes using Sed5 as the Qa-SNARE, Bos1 as the Qb-SNARE, Bet1 as the Qc-SNARE, and Sec22 as the R-SNARE and *in vivo* studies also implicate this set of SNAREs in ER–Golgi traffic (McNew *et al.*, 2000). This process is also under the control of the Sec/Munc protein family member, Sly1, that binds to Sed5 to increase the fidelity of assembly between Sed5, Bos1, Bet1, and Sec22 relative to noncognate SNARE complexes (Cao and Barlowe, 2000). Once fusion is completed, the SNAREs remain associated in a cis-SNARE complex at the target membrane. This cis-SNARE complex is disassembled by SNAP/Sec17 and the AAA+ ATPase NSF/Sec18 in an ATP-dependent manner to allow the SNAREs to proceed in new rounds of membrane fusion.

Conclusions

In vitro reconstitution and yeast genetics have been instrumental in defining the essential actors involved in COPII

vesicle biogenesis. The further biochemical characterization of this minimal machinery and of various binding partners allowed the functional assignments of these key players to the different steps involved in vesicle formation, cargo protein recruitment, vesicle tethering and fusion. Forthcoming challenges include a better understanding of the molecular determinants that regulate these processes during development and various environmental conditions as well as the strategies employed to accommodate the transport of the vast diversity presented by the entire complement of eukaryotic secretory proteins.

See also: Interorganellar Communication: Interplay and Processes: Post-Golgi Transport – Cargo, Carriers, and Pathways. Organelles: Structure and Function: Golgi and TGN

References

- Alvarez, C., Garcia-Mata, R., Hauri, H.P., Sztul, E., 2001. The p115-interactive proteins GM130 and giantin participate in endoplasmic reticulum–Golgi traffic. *Journal of Biological Chemistry* 276, 2693–2700.
- Antonny, B., Madden, D., Hamamoto, S., Orci, L., Schekman, R., 2001. Dynamics of the COPII coat with GTP and stable analogues. *Nature Cell Biology* 3, 531–537.
- Appenzeller, C., Andersson, H., Kappeler, F., Hauri, H.P., 1999. The lectin ERGIC-58 is a cargo transport receptor for glycoproteins. *Nature Cell Biology* 1, 330–334.
- Barlowe, C., Orci, L., Yeung, T., *et al.*, 1994. COPII: A membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* 77, 895–907.
- Behnia, R., Barr, F.A., Flanagan, J.J., Barlowe, C., Munro, S., 2007. The yeast orthologue of GRASP65 forms a complex with a coiled-coil protein that contributes to ER to Golgi traffic. *Journal of Cell Biology* 176, 255–261.
- Belden, W.J., Barlowe, C., 2001. Role of Erv29p in collecting soluble secretory proteins into ER-derived transport vesicles. *Science* 294, 1528–1531.
- Bi, X., Corpina, R.A., Goldberg, J., 2002. Structure of the Sec23/24–Sar1 pre-budding complex of the COPII vesicle coat. *Nature* 419, 271–277.
- Bi, X., Mancias, J.D., Goldberg, J., 2007. Insights into COPII coat nucleation from the structure of Sec23.Sar1 complexed with the active fragment of Sec31. *Developmental Cell* 13, 635–645.
- Bielli, A., Haney, C.J., Gabreski, G., *et al.*, 2005. Regulation of Sar1 NH2 terminus by GTP binding and hydrolysis promotes membrane deformation to control COPII vesicle fission. *Journal of Cell Biology* 171, 919–924.
- Boyadjev, S.A., Fromme, J.C., Ben, J., *et al.*, 2006. Cranio-lenticulo-sutural dysplasia is caused by a SEC23A mutation leading to abnormal endoplasmic-reticulum-to-Golgi trafficking. *Nature Genetics* 38, 1192–1197.
- Cao, X., Ballew, N., Barlowe, C., 1998. Initial docking of ER-derived vesicles requires Uso1p and Ypt1p but is independent of SNARE proteins. *EMBO Journal* 17, 2156–2165.
- Cao, X., Barlowe, C., 2000. Asymmetric requirements for a Rab GTPase and SNARE proteins in fusion of COPII vesicles with acceptor membranes. *Journal of Cell Biology* 149, 55–66.
- Chen, X.W., Wang, H., Bajaj, K., *et al.*, 2013. SEC24A deficiency lowers plasma cholesterol through reduced PCSK9 secretion. *eLife* 2, e00444.
- Connerly, P.L., Esaki, M., Montegna, E.A., *et al.*, 2005. Sec16 is a determinant of transitional ER organization. *Current Biology* 15, 1439–1447.
- Copic, A., Latham, C.F., Horlbeck, M.A., D'Arcangelo, J.G., Miller, E.A., 2012. ER cargo properties specify a requirement for COPII coat rigidity mediated by Sec13p. *Science* 335, 1359–1362.
- la Cour, J.M., Schindler, A.J., Berchtold, M.W., Schekman, R., 2013. ALG-2 attenuates COPII budding *in vitro* and stabilizes the Sec23/Sec31A complex. *PLoS one* 8, e75309.
- Dancourt, J., Barlowe, C., 2009. Erv26p-dependent export of alkaline phosphatase from the ER requires luminal domain recognition. *Traffic* 10, 1006–1018.
- Dancourt, J., Barlowe, C., 2010. Protein sorting receptors in the early secretory pathway. *Annual Review of Biochemistry* 79, 777–802.

- Espenshade, P., Gimeno, R.E., Holzmacher, E., Teung, P., Kaiser, C.A., 1995. Yeast SEC16 gene encodes a multidomain vesicle coat protein that interacts with Sec23p. *Journal of Cell Biology* 131, 311–324.
- Farhan, H., Wendeler, M.W., Mitrovic, S., et al., 2010. MAPK signaling to the early secretory pathway revealed by kinase/phosphatase functional screening. *Journal of Cell Biology* 189, 997–1011.
- Fath, S., Mancias, J.D., Bi, X., Goldberg, J., 2007. Structure and organization of coat proteins in the COPII cage. *Cell* 129, 1325–1336.
- Futai, E., Hamamoto, S., Orci, L., Schekman, R., 2004. GTP/GDP exchange by Sec12p enables COPII vesicle bud formation on synthetic liposomes. *EMBO Journal* 23, 4146–4155.
- Gimeno, R.E., Espenshade, P., Kaiser, C.A., 1996. COPII coat subunit interactions: Sec24p and Sec23p bind to adjacent regions of Sec16p. *Molecular Biology of the Cell* 7, 1815–1823.
- Helm, J.R., Bentley, M., Thorsen, K.D., et al., 2014. Apoptosis-linked gene-2 (ALG-2)/Sec31 interactions regulate endoplasmic reticulum (ER)-to-Golgi transport: A potential effector pathway for luminal calcium. *Journal of Biological Chemistry* 289, 23609–23628.
- Herzig, Y., Sharpe, H.J., Elbaz, Y., Munro, S., Schuldiner, M., 2012. A systematic approach to pair secretory cargo receptors with their cargo suggests a mechanism for cargo selection by Erv14. *PLoS Biology* 10, e1001329.
- Jin, L., Pahuja, K.B., Wickliffe, K.E., et al., 2012. Ubiquitin-dependent regulation of COPII coat size and function. *Nature* 482, 495–500.
- Jones, B., Jones, E.L., Bonney, S.A., et al., 2003. Mutations in a Sar1 GTPase of COPII vesicles are associated with lipid absorption disorders. *Nature Genetics* 34, 29–31.
- Kim, S.D., Pahuja, K.B., Ravazzola, M., et al., 2012. The [corrected] SEC23–SEC31 [corrected] interface plays critical role for export of procollagen from the endoplasmic reticulum. *Journal of Biological Chemistry* 287, 10134–10144.
- Klinkenberg, D., Long, K.R., Shome, K., Watkins, S.C., Aridor, M., 2014. A cascade of ER exit site assembly that is regulated by p125A and lipid signals. *Journal of Cell Science* 127, 1765–1778.
- Kung, L.F., Pagant, S., Futai, E., et al., 2012. Sec24p and Sec16p cooperate to regulate the GTP cycle of the COPII coat. *EMBO Journal* 31, 1014–1027.
- Lee, M.C., Hamamoto, S., Schekman, R., 2002. Ceramide biosynthesis is required for the formation of the oligomeric H+ -ATPase Pma1p in the yeast endoplasmic reticulum. *Journal of Biological Chemistry* 277, 22395–22401.
- Lee, M.C., Orci, L., Hamamoto, S., et al., 2005. Sar1p N-terminal helix initiates membrane curvature and completes the fission of a COPII vesicle. *Cell* 122, 605–617.
- Levi, S.K., Bhattacharyya, D., Strack, R.L., Austin 2nd, J.R., Glick, B.S., 2010. The yeast GRASP Grh1 colocalizes with COPII and is dispensable for organizing the secretory pathway. *Traffic* 11, 1168–1179.
- Loftus, A.F., Hsieh, V.L., Parthasarathy, R., 2012. Modulation of membrane rigidity by the human vesicle trafficking proteins Sar1A and Sar1B. *Biochemical and Biophysical Research Communications* 426, 585–589.
- Lord, C., Bhandari, D., Menon, S., et al., 2011. Sequential interactions with Sec23 control the direction of vesicle traffic. *Nature* 473, 181–186.
- Mancias, J.D., Goldberg, J., 2007. The transport signal on Sec22 for packaging into COPII-coated vesicles is a conformational epitope. *Molecular Cell* 26, 403–414.
- Mancias, J.D., Goldberg, J., 2008. Structural basis of cargo membrane protein discrimination by the human COPII coat machinery. *EMBO Journal* 27, 2918–2928.
- Matsuoka, K., Orci, L., Amherdt, M., et al., 1998. COPII-coated vesicle formation reconstituted with purified coat proteins and chemically defined liposomes. *Cell* 93, 263–275.
- McNew, J.A., Parlati, F., Fukuda, R., et al., 2000. Compartmental specificity of cellular membrane fusion encoded in SNARE proteins. *Nature* 407, 153–159.
- Merte, J., Jensen, D., Wright, K., et al., 2010. Sec24b selectively sorts Vangl2 to regulate planar cell polarity during neural tube closure. *Nature Cell Biology* 12, 41–46. ^{sup} 41–48.
- Miller, E.A., Beilharz, T.H., Malkus, P.N., et al., 2003. Multiple cargo binding sites on the COPII subunit Sec24p ensure capture of diverse membrane proteins into transport vesicles. *Cell* 114, 497–509.
- Miller, E.A., Liu, Y., Barlowe, C., Schekman, R., 2005. ER–Golgi transport defects are associated with mutations in the Sed5p-binding domain of the COPII coat subunit, Sec24p. *Molecular Biology of the Cell* 16, 3719–3726.
- Montegna, E.A., Bhav, M., Liu, Y., Bhattacharyya, D., Glick, B.S., 2012. Sec12 binds to Sec16 at transitional ER sites. *PLoS One* 7, e31156.
- Mossessova, E., Bickford, L.C., Goldberg, J., 2003. SNARE selectivity of the COPII coat. *Cell* 114, 483–495.
- Muniz, M., Nuoffer, C., Hauri, H.P., Riezman, H., 2000. The Emp24 complex recruits a specific cargo molecule into endoplasmic reticulum-derived vesicles. *Journal of Cell Biology* 148, 925–930.
- Nakamura, N., Rabouille, C., Watson, R., et al., 1995. Characterization of a cis-Golgi matrix protein, GM130. *Journal of Cell Biology* 131, 1715–1726.
- Noble, A.J., Zhang, Q., O'Donnell, J., et al., 2013. A pseudoatomic model of the COPII cage obtained from cryo-electron microscopy and mass spectrometry. *Nature Structural & Molecular Biology* 20, 167–173.
- Ong, Y.S., Tang, B.L., Loo, L.S., Hong, W., 2010. p125A exists as part of the mammalian Sec13/Sec31 COPII subcomplex to facilitate ER–Golgi transport. *Journal of Cell Biology* 190, 331–345.
- Pagant, S., Wu, A., Edwards, S., Diehl, F., Miller, E.A., 2015. Sec24 is a coincidence detector that simultaneously binds two signals to drive ER export. *Current Biology* 25, 403–412.
- Peng, R., De Antoni, A., Gallwitz, D., 2000. Evidence for overlapping and distinct functions in protein transport of coat protein Sec24p family members. *Journal of Biological Chemistry* 275, 11521–11528.
- Powers, J., Barlowe, C., 2002. Erv14p directs a transmembrane secretory protein into COPII-coated transport vesicles. *Molecular Biology of the Cell* 13, 880–891.
- Roberg, K.J., Crotwell, M., Espenshade, P., Gimeno, R., Kaiser, C.A., 1999. LST1 is a SEC24 homologue used for selective export of the plasma membrane ATPase from the endoplasmic reticulum. *Journal of Cell Biology* 145, 659–672.
- Rossanese, O.W., Soderholm, J., Bevis, B.J., et al., 1999. Golgi structure correlates with transitional endoplasmic reticulum organization in *Pichia pastoris* and *Saccharomyces cerevisiae*. *Journal of Cell Biology* 145, 69–81.
- Saito, K., Chen, M., Bard, F., et al., 2009. TANGO1 facilitates cargo loading at endoplasmic reticulum exit sites. *Cell* 136, 891–902.
- Saito, K., Yamashiro, K., Ichikawa, Y., et al., 2011. cTAGE5 mediates collagen secretion through interaction with TANGO1 at endoplasmic reticulum exit sites. *Molecular Biology of the Cell* 22, 2301–2308.
- Saito, K., Yamashiro, K., Shimazu, N., et al., 2014. Concentration of Sec12 at ER exit sites via interaction with cTAGE5 is required for collagen export. *Journal of Cell Biology* 206, 751–762.
- Sato, K., Nakano, A., 2005. Dissection of COPII subunit-cargo assembly and disassembly kinetics during Sar1p-GTP hydrolysis. *Nature Structural & Molecular Biology* 12, 167–174.
- Seemann, J., Jokitalo, E.J., Warren, G., 2000. The role of the tethering proteins p115 and GM130 in transport through the Golgi apparatus *in vivo*. *Molecular Biology of the Cell* 11, 635–645.
- Seog, D.H., Kito, M., Igarashi, K., Yoda, K., Yamasaki, M., 1994a. Molecular characterization of the USO1 gene product which is essential for vesicular transport in *Saccharomyces cerevisiae*. *Biochemical and Biophysical Research Communications* 200, 647–653.
- Seog, D.H., Kito, M., Yoda, K., Yamasaki, M., 1994b. Uso1 protein contains a coiled-coil rod region essential for protein transport from the ER to the Golgi apparatus in *Saccharomyces cerevisiae*. *Journal of Biochemistry* 116, 1341–1345.
- Settles, E.L., Loftus, A.F., McKeown, A.N., Parthasarathy, R., 2010. The vesicle trafficking protein Sar1 lowers lipid membrane rigidity. *Biophysical Journal* 99, 1539–1545.
- Sharpe, L.J., Luu, W., Brown, A.J., 2011. Akt phosphorylates Sec24: New clues into the regulation of ER-to-Golgi trafficking. *Traffic* 12, 19–27.
- Shimoi, W., Ezawa, I., Nakamoto, K., et al., 2005. p125 is localized in endoplasmic reticulum exit sites and involved in their organization. *Journal of Biological Chemistry* 280, 10141–10148.
- Shimoni, Y., Kurihara, T., Ravazzola, M., et al., 2000. Lst1p and Sec24p cooperate in sorting of the plasma membrane ATPase into COPII vesicles in *Saccharomyces cerevisiae*. *Journal of Cell Biology* 151, 973–984.
- Soderholm, J., Bhattacharyya, D., Strongin, D., et al., 2004. The transitional ER localization mechanism of *Pichia pastoris* Sec12. *Developmental Cell* 6, 649–659.
- Stagg, S.M., Gurkan, C., Fowler, D.M., et al., 2006. Structure of the Sec13/31 COPII coat cage. *Nature* 439, 234–238.
- Stagg, S.M., LaPointe, P., Razvi, A., et al., 2008. Structural basis for cargo regulation of COPII coat assembly. *Cell* 134, 474–484.
- Sucic, S., El-Kasaby, A., Kudlacek, O., et al., 2011. The serotonin transporter is an exclusive client of the coat protein complex II (COPII) component SEC24C. *Journal of Biological Chemistry* 286, 16482–16490.
- Sucic, S., Koban, F., El-Kasaby, A., et al., 2013. Switching the clientele: A lysine residing in the C terminus of the serotonin transporter specifies its preference for the coat protein complex II component SEC24C. *Journal of Biological Chemistry* 288, 5330–5341.
- Supek, F., Madden, D.T., Hamamoto, S., Orci, L., Schekman, R., 2002. Sec16p potentiates the action of COPII proteins to bud transport vesicles. *Journal of Cell Biology* 158, 1029–1038.
- Venditti, R., Scanu, T., Santoro, M., et al., 2012. Sedlin controls the ER export of procollagen by regulating the Sar1 cycle. *Science* 337, 1668–1672.

- Witte, K., Schuh, A.L., Hegemann, J., *et al.*, 2011. TFG-1 function in protein secretion and oncogenesis. *Nature Cell Biology* 13, 550–558.
- Yamakawa, H., Seog, D.H., Yoda, K., Yamasaki, M., Wakabayashi, T., 1996. Usa1 protein is a dimer with two globular heads and a long coiled-coil tail. *Journal of Structural Biology* 116, 356–365.
- Yorimitsu, T., Sato, K., 2012. Insights into structural and regulatory roles of Sec16 in COPII vesicle formation at ER exit sites. *Molecular Biology of the Cell* 23, 2930–2942.
- Yoshihisa, T., Barlowe, C., Schekman, R., 1993. Requirement for a GTPase-activating protein in vesicle budding from the endoplasmic reticulum. *Science* 259, 1466–1468.
- Zacharogianni, M., Kondylis, V., Tang, Y., *et al.*, 2011. ERK7 is a negative regulator of protein secretion in response to amino-acid starvation by modulating Sec16 membrane association. *EMBO Journal* 30, 3684–3700.
- Zanetti, G., Prinz, S., Daum, S., *et al.*, 2013. The structure of the COPII transport-vesicle coat assembled on membranes. *eLife* 2, e00951.

N-Linked Glycans (N-Glycans)

P Stanley, Albert Einstein College Medicine, New York, NY, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Glycan A group of sugars synthesized by the sequential action of glycosyltransferases that link sugars to each other to form a linear or branched structure termed a glycan.

Glycosidase An enzyme that hydrolyzes the chemical bond between two sugar residues. There are exo- and endo-glycosidases.

Glycosylation The transfer of a sugar or glycan to protein or lipid to make a glycoprotein or glycolipid.

Glycosyltransferase Enzyme that catalyzes the transfer of the sugar from a nucleotide sugar to a protein or lipid, or to another sugar, in a specific linkage.

N-linked glycan Specific type of glycan linked to asparagine (Asn, N) in the sequence Asn-X-Ser/Thr where X can be any amino acid except proline (Pro, P).

N-glycan Common name of an N-linked glycan.

Introduction

The first structures of N-glycans were determined in the 1970s using a combination of biochemical strategies to identify the sugars attached to Asn in certain glycoproteins, and their linkage arrangements. The glycans released from glycoproteins contained several monosaccharides, including N-acetylglucosamine (GlcNAc), mannose (Man), galactose (Gal), fucose (Fuc), and sialic acid (NeuAc). However, the composition gave only the slightest clue as to the origins of these complex structures. The N-glycans could be divided into three types – complex, hybrid and high mannose N-glycans (Figure 1). The clue to their origin was that each N-glycan type contains the same core region consisting of two GlcNAc and three Man residues. Determining how the common core of N-glycans is synthesized required much careful work because it was quickly revealed that precursors contain glucose (Glc), a

sugar almost never found in mature N-glycans of glycoproteins from a healthy cell. This article describes each stage of N-glycan synthesis from initiation on dolichol-phosphate (Dol-P) on the cytosolic face of the endoplasmic reticulum (ER); to the synthesis of a 14-sugar precursor attached to dolichol-pyrophosphate on the luminal face of the ER membrane; the en-bloc transfer of the mature oligosaccharide to Asn in a conserved amino acid sequence; the processing or trimming to remove sugars in an ordered progression that begins in the ER; the continued synthesis and processing in the *cis*- and *medial*-Golgi compartments; and final maturation in the *trans*-Golgi and *trans*-Golgi network (TGN) compartments. It is a complicated pathway involving all the basic building blocks of the cell – lipids, proteins, nucleic acids, and sugars. Therefore, defects in many different aspects of cell metabolism may cause altered synthesis of N-glycans, which may in turn lead to human disease. A large group of diseases included in the

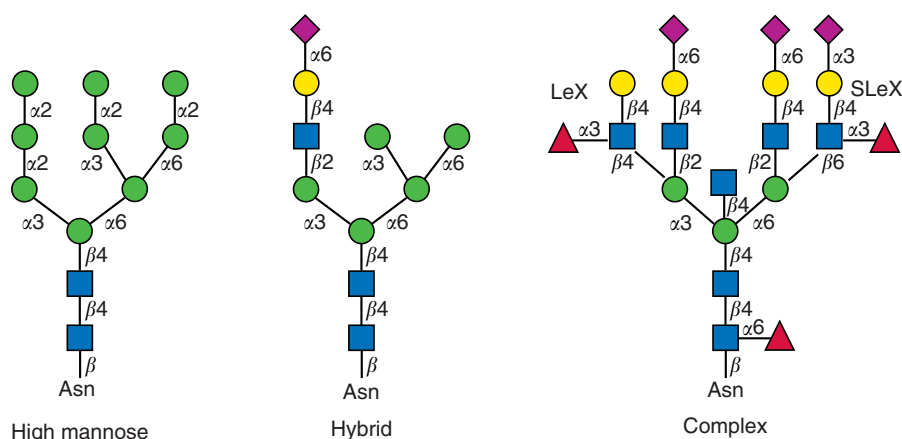


Figure 1 Typical N-glycans. Glycoproteins on the plasma membrane or secreted from the cell may carry high mannose, hybrid, or complex N-glycans. High mannose N-glycans may have fewer than 9 Mans, hybrid N-glycans may also contain only 3 Mans, and complex N-glycans may have varying numbers of branches, lactosamine units, and terminal modifications. Each N-glycan has a common core of Man₃GlcNAc₂. LeX and SLeX determinants are noted. Blue square, GlcNAc; green circle, Man; yellow circle, Gal; purple diamond, sialic acid; red triangle, Fuc. This Figure is modified from Figure 8.1, and reproduced with permission of the Consortium of Glycobiology editors from Chapter 8, Stanley, P., Schachter, H., Taniguchi, N., 2009. N-Glycans. In: Varki, A., Cummings, R.D., Esko, J.D., *et al.* (Eds.), *Essentials of Glycobiology*, second ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 101–114.

Congenital Disorders of Glycosylation (CDG) have been shown to arise from defective N-glycan synthesis.

Initiation of an N-Glycan on Dolichol-Phosphate

Contrary to expectations, the synthesis of an N-glycan begins on the cytosolic face of the ER membrane. Dolichol-phosphate (Dol-P), a polyprenol containing 15–20 unsaturated isoprene units, is located in the ER membrane with the monophosphate unit accessible to the cytosol. The first reaction of N-glycan synthesis is the transfer of phospho-GlcNAc from the nucleotide sugar UDP-GlcNAc to Dol-P, to form Dol-P-P-GlcNAc (Figure 2). All subsequent reactions in N-glycan synthesis involve the sequential addition of monosaccharides to form linear or branched sugar chains. The next sugar added is GlcNAc from UDP-GlcNAc and then three Man residues are added from GDP-Man (Figure 2). Each sugar is added at a precise location and each sugar transfer is catalyzed by a unique glycosyltransferase encoded in a single gene. Many of the enzymes are named ALG for Asparagine-Linked Glycosylation. The Dol-P-oligosaccharide with five sugars is 'flipped' across the ER membrane so that the sugars now face the lumen of the ER (Figure 3).

Maturation of the Dol-P-P-Oligosaccharide

After flipping of the Dol-P-P-oligosaccharide across the ER membrane, four additional Man residues and three Glc residues are added (Figure 3). Importantly, these sugars are all added from Dol-P-Man or Dol-P-Glc, respectively. This form of activated sugar substrate for ER glycosyltransferases is made on the cytosolic face of the ER membrane. Dol-P receives Man or Glc from GDP-Man or UDP-Glc, respectively, and the

Dol-P-sugar is 'flipped' across the ER membrane to become Dol-P-Man or Dol-P-Glc substrate. Each glycosyltransferase that adds the 4 Man and 3 Glc residues is encoded by a different gene. The mannosyltransferases and glucosyltransferases that catalyze these reactions are multi-membrane spanning proteins. The mature Dol-P-P-oligosaccharide contains 7 core sugars added in the cytosol and 7 additional sugars added in the ER to form 14 sugars linked to Dolichol (Dol-P-P-GlcNAc₂Man₉Glc₃) (Figure 3).

En-Bloc Transfer of an N-Glycan to Protein

The mature Glc₃Man₉GlcNAc₂ oligosaccharide is transferred, for the most part co-translationally, to certain Asn residues of a protein during its synthesis and after passage into the ER lumen via a translocon. The translocon is a large complex of proteins that includes the subunits of oligosaccharyltransferase (OST), the activity that recognizes Asn residues to receive an N-glycan. The minimal consensus sequence for the receipt of the Glc₃Man₉GlcNAc₂ oligosaccharide is Asn-X(not Pro)-Ser/Thr. Glycoproteomics studies have revealed Asn-X-Cys as a rare site of N-glycosylation and some non-consensus sites (Zielinska *et al.*, 2010). Asn residues in these sequons are termed putative N-glycan sites. This is because not all such sites are accessible to OST and thus not all putative sites receive an N-glycan. This may be due to conformational restrictions, or to the speed of passage past OST, or to blockage of access by other proteins. Attempts to broaden the consensus sequence in order to add N-glycosylation sites to proteins to aid in their folding and stability have identified enhanced aromatic sequons that may be efficiently N-glycosylated (Price *et al.*, 2012). Dol-P-P-oligosaccharides that contain fewer than 14 sugars may arise if glucose is limiting, due to altered expression of biosynthetic enzymes or due to ER stress. Such

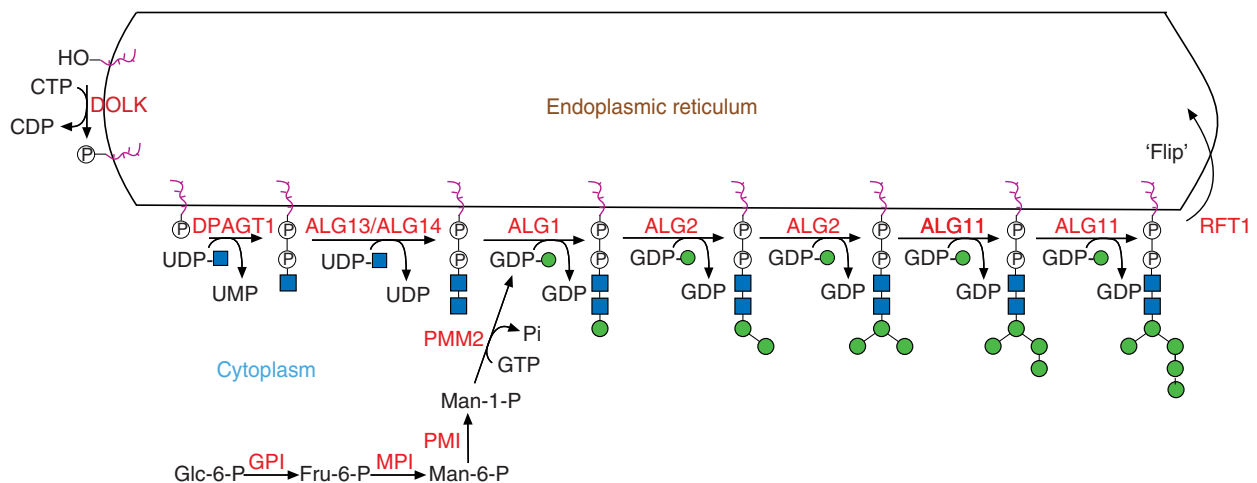


Figure 2 N-glycan synthesis begins on the cytosolic face of the ER. Dol-P is the substrate for DPAGT1 that initiates N-glycan synthesis by transferring P-GlcNAc from UDP-GlcNAc to generate Dol-P-P-GlcNAc. The subsequent addition of each sugar is catalyzed by the enzymes shown. Dol-P-P-GlcNAc₂Man₅ is flipped across the ER membrane and subsequent reactions occur on the lumenal face of the ER membrane. Blue square, GlcNAc; green circle, Man. Enzymes in red are named according to the recommendation of the Human Genome Nomenclature Committee. This Figure is modified from Figure 8.3, and reproduced with permission of the Consortium of Glycobiology editors from Chapter 8, Stanley, P., Schachter, H., Taniguchi, N., 2009. N-Glycans. In: Varki, A., Cummings, R.D., Esko, J.D., *et al.* (Eds.), *Essentials of Glycobiology*, second ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 101–114.

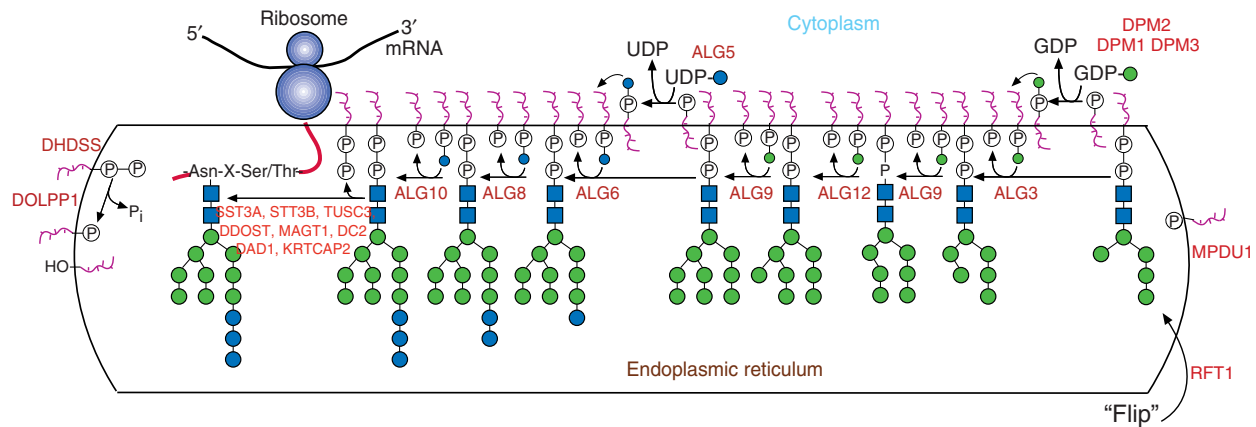


Figure 3 N-glycan transfer to protein. The Dol-P-P-GlcNAc₂Man₅ that is flipped across the ER membrane is the substrate of a mannosyltransferase that transfers Man from Dol-P-Man, which is also flipped across the ER membrane. Three more Mans are added before the addition of Glc from Dol-P-Glc by glucosyltransferases. This 14-sugar N-glycan Dol-P-P-GlcNAc₂Man₉Glc₃ is transferred 'en bloc' to accessible Asn residues in the sequon Asn-X(not Pro)-Ser/Thr during translocation of nascent protein into the ER. The oligosaccharyltransferase (OST) that catalyzes N-glycan transfer is a complex of many subunits. Blue square, GlcNAc; blue circle, Glc; green circle, Man. Enzymes in red are named according to the recommendation of the Human Genome Nomenclature Committee. This Figure is modified from Figure 8.3, and reproduced with permission of the Consortium of Glycobiology editors from Chapter 8, Stanley, P., Schachter, H., Taniguchi, N., 2009. N-Glycans. In: Varki, A., Cummings, R.D., Esko, J.D., *et al.* (Eds.), *Essentials of Glycobiology*, second ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 101–114.

truncated Dol-P-P-oligosaccharides may be transferred to Asn, but with reduced efficiency. Many empty N-glycan sites will arise if little or no extension of the core Dol-P-P-oligosaccharide occurs. In this case, proteins will be poorly folded and subject to ER degradation.

Processing and Folding of Glycoproteins in the ER

During folding of a nascent glycoprotein, the N-glycans and glycan binding chaperones that recognize N-glycans, play an important role in quality control. The first step in N-glycan processing is removal of the terminal Glc residue by a membrane-bound α -glucosidase, MOGS (Figure 4). The next Glc is removed by a soluble α -glucosidase (GANAB) that generates Glc₁Man₉GlcNAc₂ recognized by the membrane-bound chaperone calnexin or the soluble chaperone calreticulin (Figure 4). Protein disulfide isomerase (PDI) ERp57 is also in this complex. As folding continues, the last Glc residue is removed by GANAB. If folding is still incomplete following Glc removal, a glucosyltransferase (UGGT) binds to exposed hydrophobic patches on glycoproteins and reglucosylates Man₉GlcNAc₂ N-glycans to restore Glc₁Man₉GlcNAc₂ that may be chaperoned again by calnexin/ERp57 or calreticulin. This cycle of removal and re-addition of a Glc residue continues until folding is optimal. The Man₉GlcNAc₂ N-glycan on correctly folded glycoproteins is recognized by XTP3-B, a protective chaperone that may stabilize glycoproteins. Once folding is of high quality, an ER mannosidase (ERMAN1) removes the central, terminal Man residue to generate Man₈GlcNAc₂ that is recognized by ERGIC-53 in the complex LMAN1/MCFD2 and escorted from the ER to the ER-Golgi intermediate compartment (ERGIC). This occurs for both soluble and membrane-bound glycoproteins.

If appropriate folding fails to occur, several chaperones assist in the removal of misfolded glycoproteins from the ER.

For glycoproteins that are misfolded early after synthesis, removal of the first Glc generates a recognition determinant for a complex between membrane-bound malectin (MLEC) that recognizes Glc₂Man₉GlcNAc₂ and ribophorin 1 (RPN1), a subunit of OST that recognizes misfolded protein. The misfolded glycoprotein is translocated to the cytoplasm for degradation by the proteasome. If appropriate folding has not been achieved after the last Glc residue has been removed, the glycoprotein is subjected to ER-associated degradation (ERAD). Removal of two Man residues from Man₉GlcNAc₂ by ER degradation enhancing mannosidase-like protein 3 (EDEM3), results in binding of OS-9 to Man₇GlcNAc₂ N-glycans. OS-9 is in a complex with an ER ubiquitin ligase, and assists in escorting the unfolded glycoprotein from the ER for degradation by the proteasome. EDEM1 and EDEM2 also bind to N-glycans and assist in ERAD. NGLY1 is a cytosolic N-glycanase that removes N-glycans prior to degradation by the proteasome. All N-glycans in the ER are sensitive to bacterial endoglycosidase H (Endo H), which cleaves between the two core GlcNAc residues leaving a single GlcNAc attached to Asn. To date no mammalian gene encoding an Endo H activity has been reported.

Unique Modification of Lysosomal Hydrolases

Lysosomal hydrolases that will become residents of lysosomes are recognized in the *cis*-Golgi by a complex enzyme with phospho-GlcNAc-transferase activity. The enzyme is composed of two alpha and beta subunits (GNPTAB) and two gamma subunits (GNPTG), and transfers phospho-GlcNAc from UDP-GlcNAc to the high mannose N-glycans of lysosomal hydrolases (Figure 5). Following passage to the *trans*-Golgi, the terminal GlcNAc is removed by an N-acetylglucosaminidase (NAGPA) and the exposed Man-P is recognized by a Man-6-P

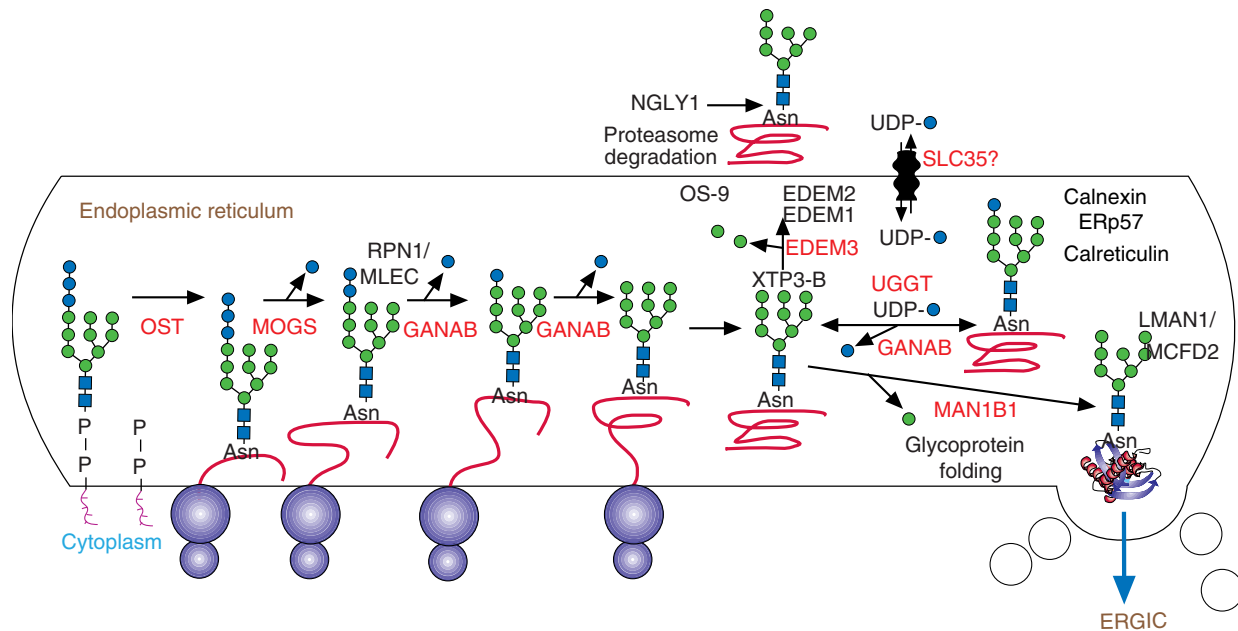


Figure 4 N-glycan processing and glycoprotein folding in the ER. Following transfer by OST to protein, the α -glucosidase MOGS binds to the N-glycan. Processing of the N-glycan begins when the terminal Glc is removed by MOGS, generating an N-glycan recognized by the chaperone complex RPN1/MLEC. Removal of the next Glc generates a recognition determinant for calnexin/calreticulin/ERp57 which assist in glycoprotein folding. When all 3 Glc residues have been removed, the partially folded glycoprotein is usually recognized by the unfolded glycoprotein glucosyltransferase UGGT, re-glucosylated and chaperoned by calnexin in a cycle until appropriate folding has been achieved. If that is never achieved, the glycoprotein is either degraded in the ER by ERAD, or in the cytoplasm by the proteasome. Properly folded glycoproteins have one Man removed and are escorted to the *cis*-Golgi network (CGN) by LMAN1/MCFD2. Blue square, GlcNAc; blue circle, Glc; green circle, Man. Enzymes in red and lectin chaperones in black are named according to the recommendation of the Human Genome Nomenclature Committee. This Figure is modified from Figure 8.4, and reproduced with permission of the Consortium of Glycobiology editors from chapter 8, Stanley, P., Schachter, H., Taniguchi, N., 2009. N-Glycans. In: Varki, A., Cummings, R.D., Esko, J.D., *et al.* (Eds.), *Essentials of Glycobiology*, second ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 101–114.

receptor (MPR) which escorts the lysosomal hydrolase to the lysosome (Figure 5). The Man-6-P receptor is released in the acidic environment of the endosome and cycles between the Golgi, endosomes and the plasma membrane. Mutations in enzymes that generate the Man-6-P recognition determinant lead to lysosomal storage diseases including I-cell disease (MPSII and MPSIIIA).

Maturation of N-Glycans During Transit Through Golgi Compartments

Processing of the N-glycans on glycoproteins continues in the *cis*-Golgi compartment (Figure 5). If high mannose N-glycans on a glycoprotein still carry one or more Glc residues that were not removed in the ER, a Golgi endo-mannosidase will cleave between Man residues to release Glc₁₋₃Man. Terminal Man residues are removed by class I Golgi mannosidase MAN1A2 to generate the Man₅GlcNAc₂ substrate for MGAT1 which transfers GlcNAc from UDP-GlcNAc, to generate a hybrid N-glycan in the *medial*-Golgi compartment. If not further processed, the hybrid N-glycan can be extended by the addition of Gal to GlcNAc and of sialic acid to Gal (Figure 1). More typically, the N-glycan product of MGAT1 is trimmed by removal of two Man residues by a class II mannosidase MAN2A1 which generates the substrate for MGAT2. In the

absence of MGAT2, the GlcNAc on the Man₃ hybrid N-glycan can be extended with Gal and sialic acid. Transfer of GlcNAc from UDP-GlcNAc by MGAT2 generates a biantennary, complex N-glycan that is a substrate for many glycosyltransferases. In the *medial*-Golgi, additional GlcNAc residues may be added via MGAT3 to generate a bisected N-glycan, or MGAT4, MGAT5, MGAT6 to generate branched, complex N-glycans (Figure 1). In the *trans*-Golgi, the GlcNAc linked to Asn may receive a Fuc transferred from GDP-Fuc by FUT8 (Figure 5). Terminal GlcNAc residues on complex N-glycans are the substrate for galactosyltransferases including B4GALT1. Gal is usually added to all terminal GlcNAcs of complex or hybrid N-glycans in the *trans*-Golgi. Gal residues on complex or hybrid N-glycans in the *trans*-Golgi or the TGN are substrates for sialyltransferases such as ST3GAL1, that transfer sialic acid from CMP-sialic acid. There are at least 5 sialyltransferases that act on Gal residues of complex N-glycans. Finally, there are fucosyltransferases in the *trans*-Golgi and TGN that transfer Fuc from GDP-Fuc to the GlcNAc in Gal β 1,4GlcNAc (lactosamine) units. The trisaccharide Gal β 1,4GlcNAc(α 1,3Fuc) is the Lewis X (LeX) oncofetal antigen, whereas Neu α 2,3Gal β 1,4GlcNAc(α 1,3Fuc) is the sialyl-Lewis X (SLeX) antigen (Figure 1). SLeX is a key recognition determinant for selectins that allow endothelial cells to recruit lymphocytes to roll on the endothelium prior to extravasation into tissues. In certain cell types, Gal and GlcNAc residues can be modified by sulfotransferases.

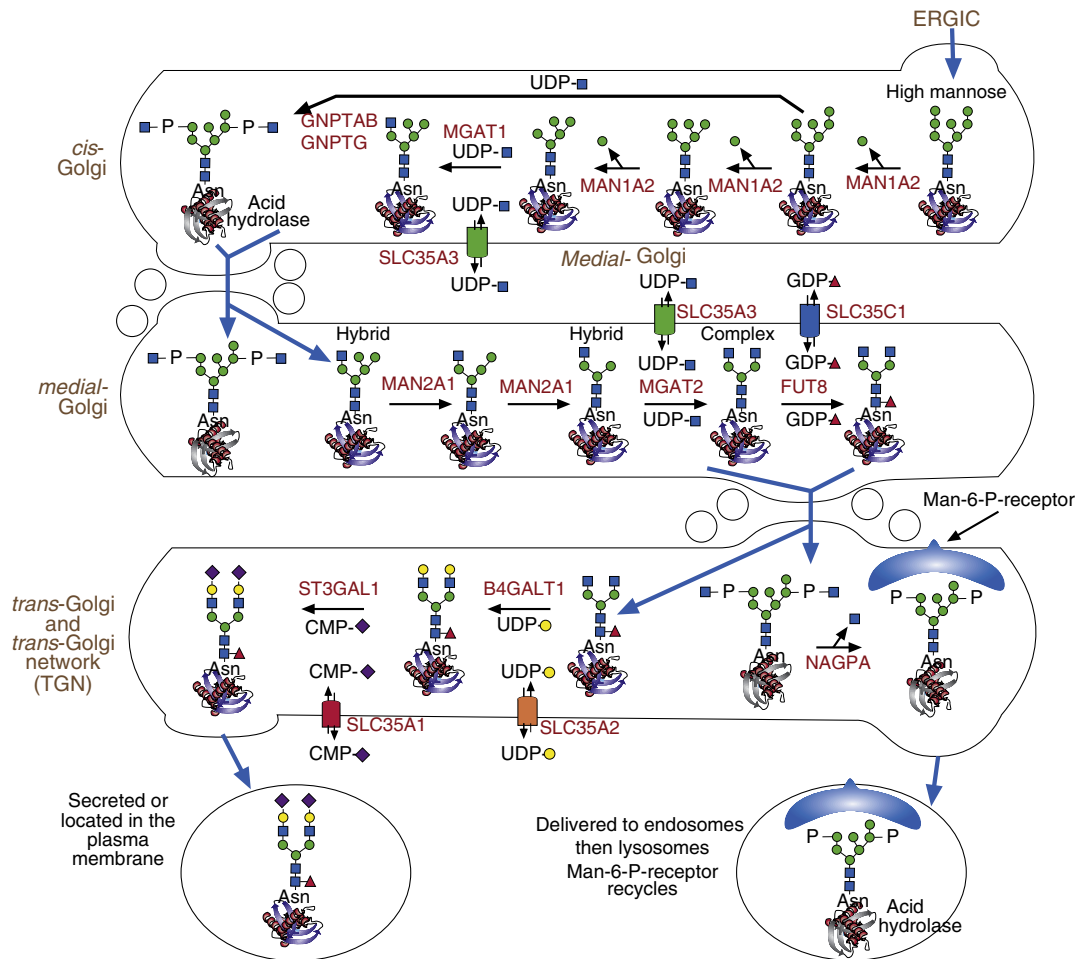


Figure 5 N-glycan maturation in the Golgi. For most glycoproteins, processing of the high mannose N-glycan continues in the *cis*-Golgi. However, for lysosomal hydrolases destined for the lysosome, the high mannose N-glycan is a substrate for a phospho-GlcNAc-transferase which transfers P-GlcNAc, precluding further processing on that branch. For most glycoproteins, 3 MANS are removed to form Man₅GlcNAc₂, and then complex and hybrid N-glycan synthesis is initiated by MGAT1. Subsequent removal of 2 MANS and elongation by the addition of GlcNAc, Gal and sialic acid (NeuAc) leads to the generation of complex N-glycans. Blue square, GlcNAc; blue circle, Glc; green circle, Man; yellow circle, Gal; purple diamond, sialic acid; red triangle, Fuc. Enzymes in red are named according to the recommendation of the Human Genome Nomenclature Committee. This Figure is modified from Figure 8.4, and reproduced with permission of the Consortium of Glycobiology editors from Chapter 8, Stanley, P., Schachter, H., Taniguchi, N., 2009. N-Glycans. In: Varki, A., Cummings, R.D., Esko, J.D., *et al.* (Eds.), *Essentials of Glycobiology*, second ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 101–114.

Functions of N-Glycans

Chaperones Recognize High Mannose N-Glycans on Unfolded Glycoproteins

In addition to contributing to the physical properties of a glycoprotein in terms of its solubility, resistance to proteases and half-life, particular N-glycan residues, or groups interactions with glycan-binding proteins (GBPs) that regulate the expression, activity and function of a glycoprotein. In the ER, chaperones and enzymes that bind to specific regions of high mannose N-glycans lead to optimal glycoprotein folding, or degradation by ERAD, or ubiquitination and translocation to the cytosol and degradation by the proteasome (Figure 4). Transit from the ER to ERGIC is facilitated by LMAN1/MCFD2 binding to a particular high mannose N-glycan (Figure 5). If a glycoprotein makes it to the *trans*-Golgi with high mannose

N-glycans and in an unfolded state, the mannose-binding lectin VIP36 will deliver it to the chaperone BiP in the ER.

Man-6-P Receptors

Lysosomal hydrolases destined to reside in lysosomes where they degrade proteins, lipids and carbohydrates are escorted through the Golgi compartments and into endosomes by Man-6-P receptors (MPRs) which recognize Man-6-P exposed following removal of the GlcNAc transferred by phospho-GlcNAc-transferase (Figure 5). There are two types of receptor that differ in their dependence on calcium for binding Man-6-P. The cation-independent receptor, CI-MPR, contains 15 Man-6-P receptor homology domains (MRH), each of 150 amino acids, and is a transmembrane glycoprotein that cycles to the cell surface where it binds Insulin-like growth factor II

and retinoic acid in addition to Man-6-P. The cation-dependent MPR (CD-MPR) is much smaller with one MRH domain, a transmembrane and a cytosolic domain. Both MPRs transport lysosomal hydrolases from the *trans*-Golgi to endosomes where the acidic pH causes release of MPRs which recycle to the *cis*-Golgi or go to the plasma membrane.

Functions of Complex N-Glycans at the Cell Surface

The functions of complex N-glycans were discovered using genetic and biochemical strategies. N-glycans are the primary ligands for mammalian lectins such as the galectins that recognize Gal residues. Siglecs bind to sialic acid residues and selectins bind to SLeX (Figure 1). Pathogens such as influenza virus bind to sialic acid, and cell-cell interactions in the nervous system are mediated by polysialic acid on N-CAM. Mutant Chinese hamster ovary (CHO) cells have been useful in determining functions for N-glycans (Table 1). Interestingly, deletion of the *Mgat1* gene in the mouse causes embryonic lethality, whereas cells lacking this enzyme grow well in culture. The severe effects of removing MGAT1 during development may be due to several factors, including the roles complex N-glycans play in regulating growth factor and cytokine signaling via their interactions with galectins. Thus, the branching of complex N-glycans and the length of lactosamine repeats (polylactosamine) on each branch dictates the nature and strength of galectin interactions with N-glycans attached to signaling receptors. Now that new gene targeting technology can readily be applied to mammalian cells, the generation of precise somatic cell mutants to determine sites of N-glycan modification, N-glycan functions and for use in glycosylation engineering can proceed at an enhanced pace. Similarly, the generation of mouse mutants defective in N-glycan synthesis continues to open many avenues for functional studies.

Functions of Complex N-Glycans on Secreted Glycoproteins

Complex N-glycans on secreted glycoproteins affect their overall properties, solubility, stability, half-life and cell-specific targeting.

Complex N-glycans may be used to target therapeutic glycoproteins to cells with specific glycan binding receptors. Thus, removal of sialic acid to expose terminal Gal residues causes glycoproteins to be rapidly cleared from the circulation by the Ashwell-Morrell receptor (ASGR) in the liver. This may be used as a strategy for targeting therapeutics to liver. Glycoprotein hormones are also targeted to liver for removal from the circulation by recognition of terminal sulfated GalNAc on their complex N-glycans by the mannose/GalNAc-4-SO(4) receptor (MANR) of hepatic endothelial cells. Targeting of lysosomal hydrolases to mannose receptors of reticuloendothelial cells for enzyme replacement therapy in patients with lysosomal storage diseases is enhanced by preparing enzymes that carry only high mannose N-glycans. Cytotoxic IgG antibodies are made ~1000 times more cytotoxic if the core Fuc residue of the complex N-glycan in the hinge region is not added. Immune IgG serum used in passive immunization of patients with auto-immune disease is more active as an anti-inflammatory treatment than if it is fully sialylated.

Summary

Since the discovery of dolichol-linked sugars in the late 1970s, the complexity of N-glycan synthesis has continued to be revealed. Activities still remain to be discovered, such as one or more flippases responsible for translocating Dol-P-sugars to face the lumen of the ER (Figures 2 and 3). Nevertheless, the last 35 years has led to enormous progress in the identification of a large number of genes that code for activities required for N-glycan synthesis. The enzymes and transporters noted in the figures represent the major activity known to catalyze each reaction. However, for some enzymes such as the mannosidases, more than one isozyme exists. N-glycan synthesis begins on the cytosolic face of the ER membrane (Figure 2), continues in the lumen of the ER (Figure 3), then moves to the *cis*-Golgi network (CGN) and processing and maturation proceeds during transit through the Golgi compartments.

Table 1 Somatic cell mutants for glycosylation engineering and determining glycan functions

Mutant line	Cell type	Biochemical block	Gene mutated	N-Glycans on glycoproteins	References
Lec1	CHO	No complex N-glycans	<i>Mgat1</i>	High mannose N-glycans	Chen and Stanley (2003)
Lec2	CHO	Little NeuAc on complex N-glycans	<i>Slc35a1</i>	N-glycans terminate in Gal	Eckhardt <i>et al.</i> (1998)
Lec4	CHO	No β 1,6 GlcNAc branch on complex N-glycans	<i>Mgat5</i>	Only bi- or tri-antennary complex N-glycans	Weinstein <i>et al.</i> (1996)
Lec8	CHO	Little Gal & NeuAc on complex N-glycans	<i>Slc35a2</i>	Complex N-glycans terminate in GlcNAc	Oelmann <i>et al.</i> (2001)
Lec10	CHO	β (1,4)GlcNAc added to N-glycan core	<i>Mgat3</i>	Complex N-glycans carry the bisecting GlcNAc	Stanley <i>et al.</i> (2005)
Lec11	CHO	α (1,3)Fuc added to lactosamine units in complex N-glycans	<i>Fut6B</i>	Complex N-glycans carry α (1,3)Fuc to form LeX or SLeX	Zhang <i>et al.</i> (1999)
Lec12	CHO	α (1,3)Fuc added to lactosamine units in complex N-glycans	<i>Fut9</i>	Complex N-glycans carry α (1,3)Fuc to form LeX	Patnaik <i>et al.</i> (2004)
Lec13	CHO	Little synthesis of GDP-Fuc	<i>Gmfs</i>	Complex N-glycans have little Fuc	Ohyama <i>et al.</i> (1998)

Abbreviation: GC, glycoconjugate.

Glycan binding chaperones with lectin domains assist folding, removal of unfolded proteins in the ER, and transport to the CGN. In the Golgi, lysosomal hydrolases are escorted by MPRs through the Golgi to endosomes and VIP36 recognizes unfolded glycoproteins with high mannose N-glycans and returns them to the ER. It seems likely that additional glycan binding chaperones will be found in the Golgi. The N-glycans that end up on a mature glycoprotein may be very heterogeneous. They may comprise all three classes of N-glycan and many variants of each. This complexity is being revealed by glycomics and glycoproteomics mass spectrometry. However, it is clear that specific functions of individual sugars or groups of sugars on N-glycans operate in a sea of background heterogeneity. An excellent example is the recognition of SLeX on P-selectin glycoprotein ligand 1 (PSGL-1) which requires SLeX to be present on an O-glycan attached to a specific Thr near the N-terminus of PSGL-1, and which also requires sulfated tyrosine residues nearby. The fact that many other N- and O-glycans on PSGL-1 may carry SLeX is of little or no benefit if the N-terminal recognition determinant is absent. For this reason, finding the functional glycan unit using a biochemical strategy alone is very difficult. Genetic strategies in combination with biochemistry are much more efficient.

See also: Interorganellar Communication: Interplay and Processes: Extracellular O-Glycans; Intra-Golgi Transport. Molecular Principles, Components, Technology, and Concepts: Proteins: Posttranslational Modifications: Key Players in Health and Disease. Organelles: Structure and Function: Intermediate Compartment: A Sorting Station between the Endoplasmic Reticulum and the Golgi Apparatus; The Endoplasmic Reticulum

References

- Chen, W., Stanley, P., 2003. Five Lec1 CHO cell mutants have distinct Mgat1 gene mutations that encode truncated N-acetylglucosaminyltransferase I. *Glycobiology* 13, 43–50.
- Eckhardt, M., Gotza, B., Gerardy-Schahn, R., 1998. Mutants of the CMP-sialic acid transporter causing the Lec2 phenotype. *Journal of Biological Chemistry* 273, 20189–20195.
- Oelmann, S., Stanley, P., Gerardy-Schahn, R., 2001. Point mutations identified in Lec8 chinese hamster ovary glycosylation mutants that inactivate both the UDP-galactose and CMP-sialic acid transporters. *Journal of Biological Chemistry* 276, 26291–26300.
- Ohyama, C., Smith, P.L., Angata, K., *et al.*, 1998. Molecular cloning and expression of GDP-D-mannose-4,6-dehydratase, a key enzyme for fucose metabolism defective in Lec13 cells. *Journal of Biological Chemistry* 273, 14582–14587.
- Patnaik, S.K., Potvin, B., Stanley, P., 2004. Lec12 and Lec29 gain-of-function chinese hamster ovary mutants reveal mechanisms for regulating VIM-2 antigen synthesis and E-selectin binding. *Journal of Biological Chemistry* 279, 49716–49726.
- Price, J.L., Culyba, E.K., Chen, W., *et al.*, 2012. N-glycosylation of enhanced aromatic sequons to increase glycoprotein stability. *Biopolymers* 98, 195–211.
- Stanley, P., Sundaram, S., Tang, J., Shi, S., 2005. Molecular analysis of three gain-of-function CHO mutants that add the bisecting GlcNAc to N-glycans. *Glycobiology* 15, 43–53.
- Weinstein, J., Sundaram, S., Wang, X., *et al.*, 1996. A point mutation causes mistargeting of Golgi GlcNAc-TV in the Lec4A Chinese hamster ovary glycosylation mutant. *Journal of Biological Chemistry* 271, 27462–27469.
- Zhang, A., Potvin, B., Zaiman, A., *et al.*, 1999. The gain-of-function Chinese hamster ovary mutant Lec11B expresses one of two Chinese hamster FUT6 genes due to the loss of a negative regulatory factor. *Journal of Biological Chemistry* 274, 10439–10450.
- Zielinska, D.F., Gnad, F., Wisniewski, J.R., Mann, M., 2010. Precision mapping of an *in vivo* N-glycoproteome reveals rigid topological and sequence constraints. *Cell* 141, 897–907.

Further Reading

- Aebi, M., 2013. N-linked protein glycosylation in the ER. *Biochimica et Biophysica Acta* 1833, 2430–2437.
- Boscher, C., Dennis, J.W., Nabi, I.R., 2011. Glycosylation, galectins and cellular signaling. *Current Opinion in Cell Biology* 23, 383–392.
- Caramelo, J.J., Parodi, A.J., 2008. Getting in and out from calnexin/calreticulin cycles. *Journal of Biological Chemistry* 283, 10221–10225.
- Cylwik, B., Naklicki, M., Chrostek, L., Gruszczyńska, E., 2013. Congenital disorders of glycosylation. Part I. Defects of protein N-glycosylation. *Acta Biochimica Polonica* 60, 151–161.
- Enns, G.M., Shashi, V., Bainbridge, M., *et al.*, 2014. Mutations in NGLY1 cause an inherited disorder of the endoplasmic reticulum-associated degradation pathway. *Genetics in Medicine* 16, 751–758.
- Freeze, H.H., 2013. Understanding human glycosylation disorders: Biochemistry leads the charge. *Journal of Biological Chemistry* 288, 6936–6945.
- Groisman, B., Shenkman, M., Ron, E., Lederkremer, G.Z., 2011. Mannose trimming is required for delivery of a glycoprotein from EDEM1 to XTP3-B and to late endoplasmic reticulum-associated degradation steps. *Journal of Biological Chemistry* 286, 1292–1300.
- Kanehara, K., Kawaguchi, S., Ng, D.T., 2007. The EDEM and Yos9p families of lectin-like ERAD factors. *Seminars in Cell & Developmental Biology* 18, 743–750.
- Kornfeld, R., Kornfeld, S., 1985. Assembly of asparagine-linked oligosaccharides. *Annual Review of Biochemistry* 54, 631–664.
- Needham, P.G., Brodsky, J.L., 2013. How early studies on secreted and membrane protein quality control gave rise to the ER associated degradation (ERAD) pathway: The early history of ERAD. *Biochimica et Biophysica Acta* 1833, 2447–2457.
- Olivari, S., Molinari, M., 2007. Glycoprotein folding and the role of EDEM1, EDEM2 and EDEM3 in degradation of folding-defective glycoproteins. *FEBS Letters* 581, 3658–3664.
- Patnaik, S.K., Potvin, B., Carlsson, S., *et al.*, 2006. Complex N-glycans are the major ligands for galectin-1, -3, and -8 on Chinese hamster ovary cells. *Glycobiology* 16, 305–317.
- Patnaik, S.K., Stanley, P., 2006. Lectin-resistant CHO glycosylation mutants. *Methods in Enzymology* 416, 159–182.
- Paulson, J.C., Macauley, M.S., Kawasaki, N., 2012. Siglecs as sensors of self in innate and adaptive immune responses. *Annals of the New York Academy of Sciences* 1253, 37–48.
- Ripka, J., Adamany, A., Stanley, P., 1986. Two Chinese hamster ovary glycosylation mutants affected in the conversion of GDP-mannose to GDP-fucose. *Archives of Biochemistry and Biophysics* 249, 533–545.
- Shental-Bechor, D., Levy, Y., 2009. Folding of glycoproteins: Toward understanding the biophysics of the glycosylation code. *Current Opinion in Structural Biology* 19, 524–533.
- Stanley, P., Schachter, H., Taniguchi, N., 2009. N-Glycans. In: Varki, A., Cummings, R.D., Esko, J.D., *et al.* (Eds.), *Essentials of Glycobiology*, second ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 101–114.
- Steenfott, C., Vakhrushev, S.Y., Joshi, H.J., *et al.*, 2013. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO Journal* 32, 1478–1488.
- Williams, D.B., 2006. Beyond lectins: The calnexin/calreticulin chaperone system of the endoplasmic reticulum. *Journal of Cell Science* 119, 615–623.
- Yamamoto, K., 2014. Intracellular lectins are involved in quality control of glycoproteins. *Proceedings of the Japan Academy. Series B, Physical and Biological Sciences* 90, 67–82.
- Zarbock, A., Ley, K., McEver, R.P., Hidalgo, A., 2011. Leukocyte ligands for endothelial selectins: Specialized glycoconjugates that mediate rolling and signaling under flow. *Blood* 118, 6743–6751.
- Zhang, P., Chan, K.F., Haryadi, R., Bardor, M., Song, Z., 2013. CHO glycosylation mutants as potential host cells to produce therapeutic proteins

with enhanced efficacy. *Advances in Biochemical Engineering/Biotechnology* 131, 63–87.

Relevant Websites

<http://www.cazy.org/>

Carbohydrate Active enZymes Database (CAZY).

<http://www.ncbi.nlm.nih.gov/books/NBK1908/term=essentials>

Essentials of Glycobiology.

<http://www.functionalglycomics.org/>

Functional Glycomics Gateway (CFG).

<http://www.genenames.org/>

Human Genome Nomenclature Committee.

<http://www.genome.jp/kegg/pathway.html>

KEGG Pathway Database.

<http://www.cdgs.com/>

The CDG Family Network.

Extracellular O-Glycans

KG Ten Hagen and L Zhang, Bethesda, MD, USA

Published by Elsevier Inc.

Glossary

Delta Cell surface-anchored ligand that binds to the Notch receptor, thereby initiating conserved signaling cascades.

Glycosylation The act of modifying proteins or other macromolecules through the covalent attachment of carbohydrates.

Hyperphosphatemia Abnormally high levels of phosphate in the blood.

Integrins Transmembrane proteins that mediate cell attachment and conserved signaling events that regulate cell shape, adhesion, and migration.

Notch Transmembrane receptor protein that initiates a conserved signaling cascade upon binding to the ligands Serrate or Delta.

Osteochondroma A benign tumor consisting of bone and cartilage.

Proteoglycans Proteins with glycosaminoglycan chains attached.

Serrate Cell surface-anchored ligand that binds to the Notch receptor, thereby initiating conserved signaling cascades.

Spondyloepiphyseal dysplasia Inherited disease characterized by short stature and skeletal abnormalities.

Temtamy preaxial brachydactyly syndrome Autosomal recessive disorder characterized by shortened digits, growth retardation, facial dysmorphism, delayed motor, and mental development.

Introduction

The addition of sugars to the hydroxyl groups of proteins (O-glycosylation) as they transit the secretory apparatus affects the function of diverse membrane-bound and secreted proteins. O-linked glycosylation can occur in many forms and is characterized by the first sugar attached to the protein backbone (Figure 1 and Table 1). The following is a description of the major types of extracellular O-linked glycans, the enzymes responsible for their synthesis, and the known biological functions of these diverse sugar chains.

Mucin-Type O-Glycans/O-GalNAc Glycans

Mucin-type O-glycosylation is characterized by the initial addition of an *N*-acetylgalactosamine (GalNAc) sugar onto the hydroxyl group of either serine or threonine in protein substrates (Figure 2). This type of glycosylation received its name because of its abundance on mucins, proteins that contain repeating domains rich in serine, threonine, and proline. The initial addition of GalNAc is catalyzed by an evolutionarily conserved family of enzymes, known as the UDP-GalNAc: polypeptide *N*-acetylgalactosaminyltransferases (ppGalNAc-Ts in mammals, PGANTs in *Drosophila*) (Bennett *et al.*, 2012; Tran and Ten Hagen, 2013). These enzymes are type II transmembrane proteins present in the Golgi apparatus. There are as many as 20 family members in humans, 19 in mice, and 12 in *Drosophila*. While each family member is capable of transferring GalNAc to protein/peptide substrates, *in vitro* studies have revealed that certain family members have preferences for particular substrates, suggesting that each enzyme may be responsible for modifying specific proteins *in vivo*. Additionally, gene expression studies have revealed that certain family members are expressed in specific temporal and spatial

patterns, suggesting that they may have unique roles in certain tissues during eukaryotic development (Kingsley *et al.*, 2000; Tian and Ten Hagen, 2006).

The further addition of other sugars, such as galactose (Gal), *N*-acetylglucosamine (GlcNAc), or GalNAc, forms extended core structures, some of which are shown in Figure 2. Core 1 or T antigen is the most common O-GalNAc glycan in mammals and is formed by the Core 1 β 1,3-galactosyltransferase (C1GalT1 or T-synthase). The core structures can be further extended and modified by the addition of GlcNAc, Gal, and fucose (Fuc). Finally, the addition of charged sugars (sialic acid in mammals and glucuronic acid (GlcA) in *Drosophila*) at the end of each chain can confer properties to the glycans and their associated proteins reflective of an increased negative charge.

Mucin-type O-glycans have been found to have diverse biological roles. Early *in vitro* and cell culture work demonstrated that O-GalNAc glycans play important roles in maintaining the physical and chemical properties of proteins (Tabak, 1995). O-GalNAc glycans serve as important ligands

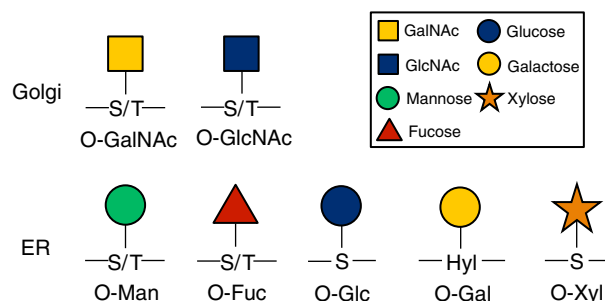
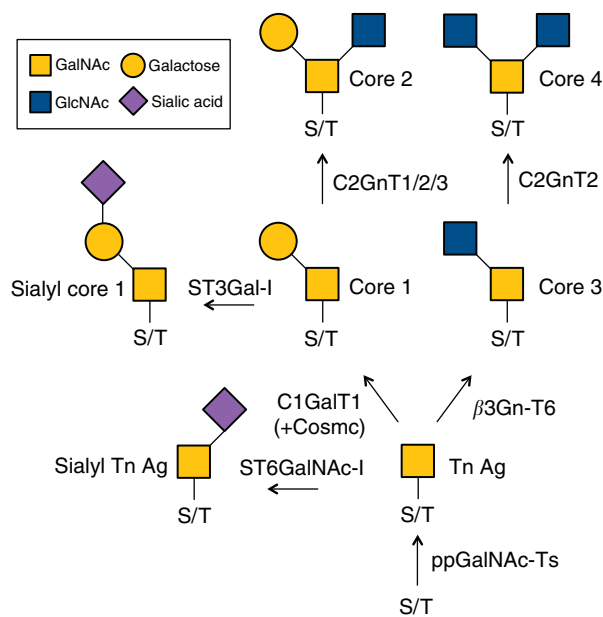


Figure 1 Extracellular O-Glycans. The major types of O-linked glycosylation initiated in the Golgi apparatus or endoplasmic reticulum (ER) are shown. S, serine; T, threonine; Hyl, hydroxylysine.

Table 1 Different types of extracellular O-glycans

Type of O-glycan	Core structure and linkage	Enzyme responsible
O-GalNAc	GalNAc α 1-Ser/Thr	ppGalNAcTs
O-Mannose	Man α 1-Ser/Thr	POMT1/2
O-Fucose	Fuc α 1-Ser/Thr	POFUT1/2
O-Glucose	Glc β 1-Ser	POGLUT
O-GlcNAc	GlcNAc β 1-Ser/Thr	EOGT
O-Galactose	Gal β 1-Hyl	GLT25D1/2
O-Xylose	Xyl β 1-Ser	XYLT1/2

Abbreviations: GalNAc, *N*-acetylgalactosamine; Man, mannose; Fuc, fucose; Glc, glucose; Gal, galactose; Xyl, xylose; ppGalNAcTs, UDP-GalNAc:polypeptide *N*-acetylgalactosaminyltransferases; POMT1/2, protein O-mannosyltransferase 1 and 2; POFUT1/2, protein O-fucosyltransferase 1 and 2; POGLUT, protein O-glucosyltransferase; EOGT, EGF repeat-specific O-GlcNAc transferase; GLT25D1/2, glycosyltransferase 25 family member 1/2; XYLT1/2, xylosyltransferase 1 and 2; Ser, serine; Thr, threonine; Hyl, hydroxylysine.

**Figure 2** Mucin-Type O-Glycans/O-GalNAc Glycans. The enzymes responsible for the synthesis of mucin-type O-glycans and the common core structures are shown.

for molecular recognition involved in lymphocyte homing (Yago *et al.*, 2010; Tenno *et al.*, 2007; Block *et al.*, 2012). Many studies in model organisms have identified crucial roles for O-glycans during development as well. Studies from *Drosophila* demonstrated that O-glycans are essential for development (Tran *et al.*, 2012; Bennett *et al.*, 2012; Tran and Ten Hagen, 2013). Further, O-glycans are essential for epithelial tube formation in the respiratory tract (Tian and Ten Hagen, 2007), nervous system development (Lin *et al.*, 2008), and integrin-mediated cell adhesion (Zhang *et al.*, 2008). It was further determined that the loss of mucin-type O-glycosylation affected the secretion of extracellular matrix (ECM) proteins during *Drosophila* development, leading to a disruption in integrin-mediated cell adhesion (Zhang *et al.*, 2010). The effects of O-GalNAc glycans on secretion of ECM proteins during development was also verified in mice, where loss of

ppGalNAcT-1 resulted in a disruption of basement membrane secretion, resulting in reduced integrin and FGF signaling, induction of ER stress, and reduced organ growth (Tian *et al.*, 2012). Most recently, studies in *Drosophila* demonstrated that O-glycosylation influences secretion by protecting an essential component of the secretory apparatus (Tango1) from proteolysis (Zhang *et al.*, 2014). Other studies suggest that O-GalNAc glycans may influence TGF β signaling (Herr *et al.*, 2008) and also Notch signaling, where they influence left-right patterning during development (Boskovski *et al.*, 2013).

Core 1-derived O-glycans are involved in angiogenesis (Xia *et al.*, 2004), separation of blood and lymphatic vessels during embryonic and postnatal development (Fu *et al.*, 2008), and proper platelet biogenesis and function (Alexander *et al.*, 2006; Wang *et al.*, 2012). Loss of core 3 O-glycans resulted in increased susceptibility to colitis and colon cancer (An *et al.*, 2007). Loss of core 1-derived O-glycans lead to spontaneous colitis in mice, supporting a protective function for O-glycans for intestinal health (Fu *et al.*, 2011; Bergstrom and Xia, 2013).

Recent genetic linkage and genome wide association studies have found associations between the enzymes involved in the synthesis of mucin-type O-glycans and various human diseases and predispositions to disease. Diseases resulting from mutations in genes involved in glycan biosynthesis are collectively known as congenital disorders of glycosylation (CDGs). For example, mutations in GALNT3 in humans are responsible for the CDG known as familial tumoral calcinosis, characterized by hyperphosphatemia, bone density changes, and subdermal calcified tumors (Topaz *et al.*, 2004). It is hypothesized that GALNT3 normally glycosylates the phosphate-regulating hormone fibroblast growth factor 23 (FGF23) to protect it from proteolysis. Other studies have suggested links between the genes initiating mucin-type O-glycosylation and other human conditions, including blood cholesterol levels (GALNT2); acute coronary syndrome (GALNT4); heterotaxy (GALNT11); breast and colorectal cancers (GALNT5 and GALNT12), suggesting crucial roles for these glycans in many aspects of human health and disease (Bennett *et al.*, 2012; Tran and Ten Hagen, 2013). Future studies will focus on identifying the molecular mechanism by which these glycans function in both developing systems as well as in adult tissues.

O-Mannose Glycans

O-linked mannose (O-Man) glycans are found on a limited number of proteins, but have been extensively studied on the extracellular protein α -dystroglycan. α -Dystroglycan comprises the extracellular portion of the dystrophin-glycoprotein complex, which is responsible for linking the extracellular matrix to the cytoskeleton and thus maintaining the integrity of muscle cells. O-linked mannose glycans present on α -dystroglycan are required for binding ECM proteins, such as laminin, agrin, and perlecan. Loss or disruption of α -dystroglycan glycosylation results in various CDGs known as congenital muscular dystrophies (CMDs) (also known as dystroglycanopathies), highlighting the importance of these glycans in human biology (Wells, 2013). Human diseases such as Walker-Warburg syndrome (WWS), muscle-eye-brain disease (MEB), and Fukuyama congenital muscular dystrophy

(FCMD), which present with varying degrees of neurological defects and muscle degeneration, are associated with mutations in the enzymes responsible for forming O-linked mannose chains on α -dystroglycan (Wells, 2013; Nakamura *et al.*, 2010; Hennet, 2012).

The initial addition of mannose to the hydroxyl groups of either serine or threonine residues occurs through the action of protein O-mannosyltransferase 1 (POMT1) and 2 (POMT2). POMT1 and POMT2 are integral membrane proteins that are localized to the ER, where mannose addition takes place. Both enzymes are required for the initial addition of mannose, suggesting that they function as a heterocomplex. Mutations in these genes cause CMDs, including WWS, one of the most severe forms of muscular dystrophy. The *Drosophila* orthologues of POMT1 and POMT2 similarly add mannose to *Drosophila* dystroglycan and result in muscle abnormalities when mutated (Nakamura *et al.*, 2010).

Diverse chain extensions can occur from the initial mannose sugar and are produced by the action of a number of different enzymes that add sugars in a step-wise fashion. Chain extension is carried out by POMGnT1, a β 1,2-N-acetylglucosaminyltransferase. POMGnT1 is a Golgi-resident type II transmembrane protein that adds GlcNAc to the extant mannose in a β 1,2 linkage. Mutations in this enzyme are also responsible for various forms of congenital muscular dystrophies (CMD) in humans (Wells, 2013). Formation of branched O-mannose glycans occurs through the activity of another Golgi-resident enzyme GlcNAcT-Vb, which adds GlcNAc in a β 1,6 linkage to mannose. Branched mannose

chains (and GlcNAcT-Vb activity) are primarily confined to the brain. Recent studies have identified another N-acetylglucosaminyltransferase (POMGnT2) that forms a β 1,4 linked GlcNAc to the extant mannose (Yoshida-Moriguchi *et al.*, 2013). This disaccharide can then be extended by the β 1,3-N-acetylgalactosaminyltransferase2 (B3GALNT2) to form a linear trisaccharide (Yoshida-Moriguchi *et al.*, 2013). It is believed that this trisaccharide serves as the substrate for phosphorylation by the atypical kinase, SGK196 or POMK, to form a phosphorylated O-mannose-based glycan. Subsequent extension of some form of this phosphorylated glycan by the enzyme LARGE further elaborates O-mannose chains (Inamori *et al.*, 2013). LARGE is unique in that it is a bi-functional enzyme with 2 independent glycosyltransferase domains. One domain is responsible for the addition of xylose to GlcA and the other is responsible for the addition of GlcA to xylose, thus making LARGE capable of adding repeating disaccharides to the phosphorylated O-mannose chains (Figure 3). Studies suggest that these repeating disaccharides are responsible for mediating binding to ECM proteins. Mutations in LARGE result in CMDs. Finally, mutations in two additional proteins responsible for the formation of O-mannose glycans, Fukutin and FKRP, also result in CMDs. While they likely contribute to the elaboration of the O-mannose chains, their biochemical activities and specific functions are currently unknown (Wells, 2013).

In addition to α -dystroglycan, recent studies have identified O-mannose glycans on additional proteins, including the cell adhesion proteins, cadherin, and plexin (Vester-Christensen

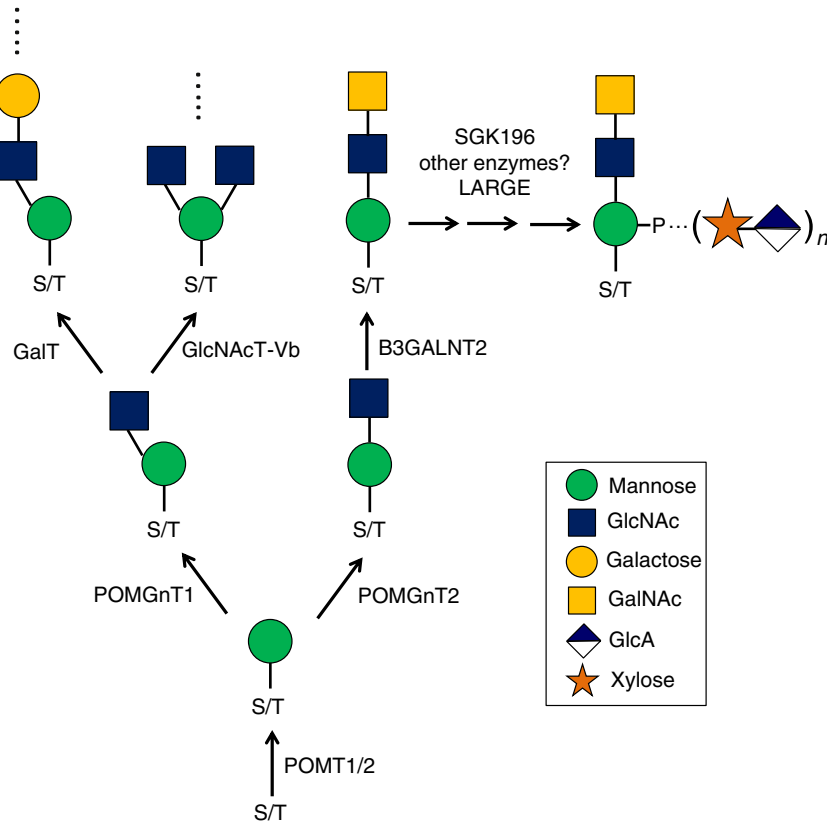


Figure 3 O-Mannose glycans. The enzymes involved in the synthesis of O-linked mannose structures are shown.

et al., 2013). It has further been found that the O-mannose glycans present on cadherin are essential for cadherin-mediated cell adhesion during early mouse embryogenesis (Lommel *et al.*, 2013).

O-Fuc Glycans

There are two distinct enzymes responsible for the addition of O-linked fucose to proteins: O-fucosyltransferase 1 (Pofut1 in mammals or Ofut1 in *Drosophila*) and O-fucosyltransferase 2 (Pofut2 in mammals or Ofut2 in *Drosophila*) (Luo *et al.*, 2006). Ofut1 is a soluble ER-localized enzyme that transfers O-linked fucose to the EGF repeats of Notch and its ligands, Delta and Serrate, recognizing the consensus motif, $C_2X_{4-5}(S/T)C_3$ (Rana and Haltiwanger, 2011). After the addition of O-fucose, the Golgi-localized $\beta 1,3$ -N-acetylglucosaminyltransferase Fringe (Fng in *Drosophila*) adds GlcNAc to the O-fucose (Stanley and Okajima, 2010). Further elongation of the GlcNAc $\beta 1$ -3Fuc can occur in mammals through the addition of galactose and sialic acid (Figure 4). Genetic studies in both the *Drosophila* and mammals have revealed the biological importance of the O-fucose glycans in conserved aspects of Notch signaling and function. Mutations in POFUT1 have been found in patients with Dowling–Degos disease, a pigmentation disorder thought to be due to defects in Notch function (Li *et al.*, 2013). Loss of O-Fuc glycans on mouse Notch1 leads to defects in T cell development and embryogenesis (Ge and Stanley, 2008). Likewise, loss of O-fucose in the fly affects Notch signaling during development. Interestingly, Ofut1 in the fly may also serve in a chaperone capacity for Notch folding or cell-surface expression, independent of its ability to transfer O-linked fucose to Notch (Rana and Haltiwanger, 2011). Studies in both mammals and *Drosophila* identified Fng as a regulator of Notch signaling. Fng expression enhances activation of Notch by the ligand Delta while inhibiting activation mediated by the ligand Serrate (Stanley and Okajima, 2010; Rana and Haltiwanger, 2011). These effects are thought to occur by Fng altering the binding between Notch and its ligands. Additionally, it was found that O-Fuc glycans, especially the structure elongated by Fringe, can affect the structure of Notch (Rana and Haltiwanger, 2011).

Pofut2 adds O-fucose to the consensus motif, $CX_2-3(S/T)CX_2G$ in the thrombospondin type 1 repeats (TSRs), which are

found in many extracellular matrix proteins, such as thrombospondin 1, properdin, f-spondin, ADAMTS-like 1, and ADAMTS-13 (Luo *et al.*, 2006; Heinonen and Maki, 2009). The O-Fuc modification catalyzed by Pofut2 can then be extended by a $\beta 1,3$ -glucosyltransferase that adds glucose to form the disaccharide structure, Glc $\beta 1$ -3Fuc. These glycans are involved in protein interactions and are required for the proper protein folding (Luo *et al.*, 2006). Mutations in this glucosyltransferase result in abnormal O-Fuc glycans on TSRs and lead to Peters Plus syndrome (Hess *et al.*, 2008), a human CDG characterized by short stature, irregular facial features, and mental retardation. Defects in O-fucosylation lead to impaired secretion of the proteases ADAMTS-like 1/punctin 1 (Wang *et al.*, 2007) and ADAMTS-13 (Ricketts *et al.*, 2007). Another recent study found that mutations in Pofut2 caused impaired epithelial to mesenchymal transition, cell movement, and signaling during mouse gastrulation, suggesting the O-Fuc glycans are essential for the formation of a proper extracellular environment (Du *et al.*, 2010).

O-Glc Glycans

O-glucose glycans were recently identified on the EGF repeats of Notch. Protein O-glucosyltransferase 1 (Poglut1 in mammals or Rumi in *Drosophila*) (Acar *et al.*, 2008), a soluble ER-localized enzyme, catalyzes the transfer of glucose to serine (Ser) in the conserved sequence $C_1XSX(P/A)C_2$ (Rana and Haltiwanger, 2011). The O-Glc can be elongated by the addition of a xylose sugar, through the activity of the Golgi-resident type II transmembrane glucoside $\alpha 1$ -3xylosyltransferases, GXYLT1, and GXYLT2 (Sethi *et al.*, 2010). The Xyl $\alpha 1$ -3Glc disaccharide can then be further elongated by the addition of another xylose by the xylosyltransferase 1, XXYL1 (Sethi *et al.*, 2012).

The *rumi* gene was originally identified in a screen for genes that affect Notch signaling (Acar *et al.*, 2008). Mutations in *rumi* resulted in developmental defects related to disruption of Notch signaling (Acar *et al.*, 2008). Similar to what was seen for POFUT1 in humans, mutations in POGLUT1 were also found in patients suffering from Dowling–Degos disease, suggesting a role in Notch function. In mice, loss of Poglut1 resulted in early embryonic lethality and severe growth retardation along with defects in cardiovascular and neural development (Fernandez-Valdivia *et al.*, 2011). Recent studies have shown that mouse Notch1 activity is greatly affected by loss of the O-Glc modification at specific positions. Interestingly, loss of O-Glc does not appear to affect Notch cell-surface expression or ligand binding. Rather, O-Glc affected Notch proteolysis that is required for proper signaling. These studies suggest that O-Glc on the extracellular domain of Notch influences conformational changes that are required within that domain to mediate efficient proteolysis (Acar *et al.*, 2008; Fernandez-Valdivia *et al.*, 2011). In contrast to the positive effects of O-Glc on Notch function, the addition of xylose appears to inhibit Notch function (Lee *et al.*, 2010).

While Notch is known to contain O-Glc on its EGF repeats, bioinformatic analysis based on the consensus sequences of O-Glc addition suggests there are other proteins that contain this sugar modification. This thought is supported by studies

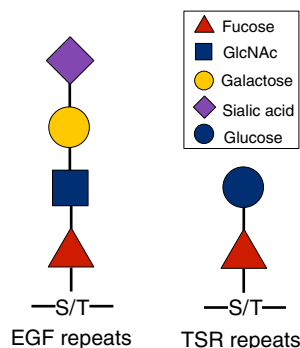


Figure 4 O-Fucose glycans. O-fucose structures initiated by Pofut1 (EGF repeats) and Pofut2 (TSR repeats) are shown.

showing that loss of Poglut1/Rumi results in phenotypes that are more severe than those seen upon loss of Notch (Fernandez-Valdivia *et al.*, 2011; Rana and Haltiwanger, 2011). Future studies will elucidate additional O-Glc containing proteins and how this sugar modification affects their biological functions.

O-GlcNAc Glycans

The addition of O-linked GlcNAc to cytoplasmic and nuclear proteins has been well documented (Hart *et al.*, 2011). However, recent studies have demonstrated that O-linked GlcNAc is also present on proteins destined for the extracellular environment. This modification is catalyzed by the ER-resident enzyme Eogt (EGF repeat-specific O-GlcNAc transferase), which generates β -linked GlcNAc attached to serine or threonine on substrate proteins (Sakaidani *et al.*, 2012). This modification was originally found on the EGF repeats of the Notch receptor and subsequently, on the ligands, Delta and Serrate in *Drosophila* (Muller *et al.*, 2013). The mouse orthologue of Eogt has a similar enzymatic activity and can glycosylate the mammalian Notch1 EGF domain (Sakaidani *et al.*, 2012).

The functional role of O-GlcNAc addition by Eogt is under investigation. The effects of O-GlcNAc modification of Notch in mammals is still unknown. However, mutations in the human EOGT have been found in patients with Adams–Oliver syndrome, an autosomal recessive disorder characterized by defects in the cranium and limbs (Shaheen *et al.*, 2013). In *Drosophila*, Eogt has also been found to glycosylate a membrane-anchored extracellular protein, Dumpy, and regulate Dumpy-dependent epithelial cell–matrix interaction (Sakaidani *et al.*, 2011), suggesting that extracellular O-GlcNAc glycans may be involved in cell–matrix or cell–cell interactions. However, other studies in *Drosophila* suggest that Eogt activity on Dumpy and Notch may influence their ability to regulate pyrimidine biosynthesis within cells (Muller *et al.*, 2013).

O-Gal Glycans

O-galactose glycans are found predominantly on collagens, one of the major groups of ECM proteins in eukaryotes. These glycans are also found on other proteins, including complement factor C1q, pulmonary surfactant proteins, and mannan binding proteins. What is unique about this type of glycosylation is that it requires hydroxylation of lysine by lysyl hydroxylase (LH1 and LH2) enzymes before galactose addition to this amino acid can occur. After lysine hydroxylation, the β -galactosyltransferases, (GLT25D1 or GLT25D2) add galactose and then the α 1,2-glucosyltransferase (LH3) adds glucose to the galactose, forming the Glc α 1-2Gal disaccharide (Schegg *et al.*, 2009). This modification occurs before collagen is packaged into its triple-helix formation to be secreted.

It is believed that O-Gal glycans are required for many aspects of collagen biosynthesis, secretion, and function (reviewed in Yamauchi and Sricholpech, 2012). Altering collagen O-glycosylation causes developmental defects, including many connective tissue disorders. In *C. elegans*, mutations in the gene

encoding lysyl hydroxylase are embryonic lethal, due to defects in collagen IV secretion. In humans, mutations in the lysyl hydroxylases causes Ehlers–Danlos syndrome type VIa, characterized by defects in the joints and skin; and mutations in LH2 causes Bruck syndrome type 2, characterized by bone fragility and joint contractures. In mice, mutations in LH3 cause embryonic lethality, due to abnormal distribution and aggregation of collagens IV and VI, and impaired basement membrane formation. In humans, the mutations in the ortholog of LH3 (PLOD3) cause congenital connective tissue disorders and epidermolysis bullosa simplex. Altogether, these studies highlight the importance of glycosylation in collagen biosynthesis and function.

Proteoglycans/O-Xyl Glycans

O-xylose addition is the first step in the formation of proteoglycans, proteins to which long linear glycosaminoglycan (GAG) chains are attached. A complete description of their synthesis and biological functions is beyond the scope of this article but is discussed in detail in another article. Additional references are also provided for a comprehensive overview (*Essentials of Glycobiology*, 2nd edition). Briefly, initial xylose addition to the serine residue of substrate proteins is catalyzed by xylosyltransferases (XYLT1 and XYLT2 in mammals and Oxt in *Drosophila*). The xylose is then extended to form the core structure GlcA β 1-3Gal β 1-3Gal β 1-4Xyl. This core structure is extended by the addition of repeating disaccharides, and these chains can then be further modified by the addition of sulfate at specific positions in the constituent sugars to form chondroitin sulfate (CS), dermatan sulfate (DS), heparin sulfate (HS), or heparin.

Proteoglycans are produced by most cells and can exist as either secreted or membrane-bound proteins. Their attached GAGs provide them with unique structural and functional properties, including the ability to bind water and form highly hydrated matrices. Proteoglycan matrices can absorb pressure; serve as scaffolds for ECM and basement membrane organization; regulate permeability of basement membranes; and regulate protease activity. Proteoglycans are also involved in protein–protein and protein–matrix interactions to regulate concentration gradients and endocytosis of bound ligands, thereby affecting cell adhesion, proliferation, differentiation, and motility (*Essentials of Glycobiology*, 2nd edition). Proteoglycans have profound effects on conserved signaling pathways during development. Many studies in model organisms have documented their effects on Hedgehog, Wnt/Wingless, and FGF signaling. There are a number of models that describe the mechanistic role of proteoglycans in signaling events, including: (1) acting as co-receptors to aid in receptor ligand interactions, (2) facilitating dimerization of ligands to aid in receptor binding, and (3) creating developmental gradients of ligands or morphogens by influencing their diffusion, transport, secretion, or stability. Mutations in the genes responsible for proteoglycan biosynthesis are responsible for a number of human diseases and congenital disorders. For example, mutations in β 4GalT-7 causes defective GAG formation of decorin and biglycan, resulting in altered cell migration, adhesion, and wound healing seen in the Ehlers–Danlos

syndrome (Seidler *et al.*, 2006). Additionally, mutations in the enzymes involved in the synthesis of HS, such as EXT1 and EXT2, cause the autosomal dominant disease hereditary multiple exostosis, characterized by bone malformations and osteochondromas (Zak *et al.*, 2002). Alterations in the synthesis of CS cause the CDGs temtamy preaxial brachydactyly syndrome and spondyloepiphyseal dysplasia (Hennet, 2012). Thus, proteoglycans and their associated polysaccharide chains perform crucial roles in diverse cellular and developmental events, and are intimately involved in human health.

Conclusions

O-glycans on extracellular proteins are uniquely positioned to mediate diverse events regulating communication between cells and other cells as well as cells and their surrounding environment. The information described herein highlights the importance of these types of protein modifications in many cellular and developmental processes. Further research in this field will uncover the myriad roles these sugars play in human development, homeostasis, and disease.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: Receptor Tyrosine Kinases and Their Ligands. Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell Adhesion to the Extracellular Matrix; Cell–Cell Adhesion and the Cytoskeleton. Interorganellar Communication: Interplay and Processes: Endoplasmic Reticulum Stress in Disease. Interorganellar Communication: Interplay and processes: N-Linked Glycans (N-Glycans). Intracellular Infectiology: Cell Processes: Bacterial Subversion of Phagocytic Killing. Molecular Principles, Components, Technology, and Concepts: Carbohydrates: Proteoglycans. Molecular Principles, Components, Technology, and Concepts: Proteins: Posttranslational Modifications: Key Players in Health and Disease. Organelles: Structure and Function: Golgi and TGN; The Endoplasmic Reticulum

References

- Acar, M., Jafar-Nejad, H., Takeuchi, H., *et al.*, 2008. Rumi is a CAP10 domain glycosyltransferase that modifies Notch and is required for Notch signaling. *Cell* 132, 247–258.
- Alexander, W.S., Viney, E.M., Zhang, J.G., *et al.*, 2006. Thrombocytopenia and kidney disease in mice with a mutation in the C1galT1 gene. *Proceedings of the National Academy of Sciences of the United States of America* 103, 16442–16447.
- An, G., Wei, B., Xia, B., *et al.*, 2007. Increased susceptibility to colitis and colorectal tumors in mice lacking core 3-derived O-glycans. *Journal of Experimental Medicine* 204, 1417–1429.
- Bennett, E.P., Mandel, U., Clausen, H., *et al.*, 2012. Control of mucin-type O-glycosylation: A classification of the polypeptide GalNAc-transferase gene family. *Glycobiology* 22, 736–756.
- Bergstrom, K.S., Xia, L., 2013. Mucin-type O-glycans and their roles in intestinal homeostasis. *Glycobiology* 23, 1026–1037.
- Block, H., Ley, K., Zarbock, A., 2012. Severe impairment of leukocyte recruitment in ppGalNAcT-1-deficient mice. *Journal of Immunology* 188, 5674–5681.
- Boskovski, M.T., Yuan, S., Pedersen, N.B., *et al.*, 2013. The heterotaxy gene GALNT11 glycosylates Notch to orchestrate cilia type and laterality. *Nature* 504, 456–459.
- Du, J., Takeuchi, H., Leonhard-Melief, C., *et al.*, 2010. O-fucosylation of thrombospondin type 1 repeats restricts epithelial to mesenchymal transition (EMT) and maintains epiblast pluripotency during mouse gastrulation. *Developmental Biology* 346, 25–38.
- Fernandez-Valdivia, R., Takeuchi, H., Samarghandi, A., *et al.*, 2011. Regulation of mammalian Notch signaling and embryonic development by the protein O-glucosyltransferase Rumi. *Development* 138, 1925–1934.
- Fu, J., Gerhardt, H., McDaniel, J.M., *et al.*, 2008. Endothelial cell O-glycan deficiency causes blood/lymphatic misconnections and consequent fatty liver disease in mice. *Journal of Clinical Investigation* 118, 3725–3737.
- Fu, J., Wei, B., Wen, T., *et al.*, 2011. Loss of intestinal core 1-derived O-glycans causes spontaneous colitis in mice. *Journal of Clinical Investigation* 121, 1657–1666.
- Ge, C., Stanley, P., 2008. The O-fucose glycan in the ligand-binding domain of Notch1 regulates embryogenesis and T cell development. *Proceedings of the National Academy of Sciences of the United States of America* 105, 1539–1544.
- Hart, G.W., Slawson, C., Ramirez-Correa, G., Lagerlof, O., 2011. Cross talk between O-GlcNAcylation and phosphorylation: Roles in signaling, transcription, and chronic disease. *Annual Review of Biochemistry* 80, 825–858.
- Heinonen, T.Y., Maki, M., 2009. Peters-plus syndrome is a congenital disorder of glycosylation caused by a defect in the beta1,3-glucosyltransferase that modifies thrombospondin type 1 repeats. *Annals of Medicine* 41, 2–10.
- Hennet, T., 2012. Diseases of glycosylation beyond classical congenital disorders of glycosylation. *Biochimica et Biophysica Acta* 1820, 1306–1317.
- Herr, P., Korniyuchuk, G., Yamamoto, Y., Grubisic, K., Oelgeschlager, M., 2008. Regulation of TGF-(beta) signalling by N-acetylgalactosaminyltransferase-like 1. *Development* 135, 1813–1822.
- Hess, D., Keusch, J.J., Oberstein, S.A., Hennekam, R.C., Hofsteenge, J., 2008. Peters Plus syndrome is a new congenital disorder of glycosylation and involves defective Omicron-glycosylation of thrombospondin type 1 repeats. *Journal of Biological Chemistry* 283, 7354–7360.
- Inamori, K., Hara, Y., Willer, T., *et al.*, 2013. Xylosyl- and glucuronyltransferase functions of LARGE in alpha-dystroglycan modification are conserved in LARGE2. *Glycobiology* 23, 295–302.
- Kingsley, P.D., Hagen, K.G., Maltby, K.M., Zara, J., Tabak, L.A., 2000. Diverse spatial expression patterns of UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase family member mRNAs during mouse development. *Glycobiology* 10, 1317–1323.
- Lee, T.V., Takeuchi, H., Jafar-Nejad, H., 2010. Regulation of notch signaling via O-glycosylation insights from Drosophila studies. *Methods in Enzymology* 480, 375–398.
- Li, M., Cheng, R., Liang, J., *et al.*, 2013. Mutations in POFUT1, encoding protein O-fucosyltransferase 1, cause generalized Dowling-Degos disease. *American Journal of Human Genetics* 92, 895–903.
- Lin, Y.R., Reddy, B.V., Irvine, K.D., 2008. Requirement for a core 1 galactosyltransferase in the Drosophila nervous system. *Developmental Dynamics* 237, 3703–3714.
- Lommel, M., Winterhalter, P.R., Willer, T., *et al.*, 2013. Protein O-mannosylation is crucial for E-cadherin-mediated cell adhesion. *Proceedings of the National Academy of Sciences of the United States of America* 110, 21024–21029.
- Luo, Y., Nita-Lazar, A., Haltiwanger, R.S., 2006. Two distinct pathways for O-fucosylation of epidermal growth factor-like or thrombospondin type 1 repeats. *Journal of Biological Chemistry* 281, 9385–9392.
- Muller, R., Jenny, A., Stanley, P., 2013. The EGF repeat-specific O-GlcNAc-transferase Eogt interacts with notch signaling and pyrimidine metabolism pathways in Drosophila. *PLoS One* 8, e62835.
- Nakamura, N., Stalnaker, S.H., Lyalin, D., *et al.*, 2010. Drosophila Dystroglycan is a target of O-mannosyltransferase activity of two protein O-mannosyltransferases, rotated abdomen and twisted. *Glycobiology* 20, 381–394.
- Rana, N.A., Haltiwanger, R.S., 2011. Fringe benefits: Functional and structural impacts of O-glycosylation on the extracellular domain of Notch receptors. *Current Opinion in Structural Biology* 21, 583–589.
- Ricketts, L.M., Dlugosz, M., Luther, K.B., Haltiwanger, R.S., Majerus, E.M., 2007. O-fucosylation is required for ADAMTS13 secretion. *Journal of Biological Chemistry* 282, 17014–17023.
- Sakaidani, Y., Ichiyanagi, N., Saito, C., *et al.*, 2012. O-linked-N-acetylglucosamine modification of mammalian Notch receptors by an atypical O-GlcNAc transferase Eogt1. *Biochemical and Biophysical Research Communications* 419, 14–19.
- Sakaidani, Y., Nomura, T., Matsuura, A., *et al.*, 2011. O-linked-N-acetylglucosamine on extracellular protein domains mediates epithelial cell-matrix interactions. *Nature Communications* 2, 583.
- Schegg, B., Hulsmeier, A.J., Rutschmann, C., Maag, C., Hennet, T., 2009. Core glycosylation of collagen is initiated by two beta(1-0)galactosyltransferases. *Molecular and Cellular Biology* 29, 943–952.

- Seidler, D.G., Faiyaz-Ul-Haque, M., Hansen, U., *et al.*, 2006. Defective glycosylation of decorin and biglycan, altered collagen structure, and abnormal phenotype of the skin fibroblasts of an Ehlers-Danlos syndrome patient carrying the novel Arg270Cys substitution in galactosyltransferase I (beta4GalT-7). *Journal of Molecular Medicine* (Berlin, Germany) 84, 583–594.
- Sethi, M.K., Buettner, F.F., Ashikov, A., *et al.*, 2012. Molecular cloning of a xylosyltransferase that transfers the second xylose to O-glucosylated epidermal growth factor repeats of notch. *Journal of Biological Chemistry* 287, 2739–2748.
- Sethi, M.K., Buettner, F.F., Krylov, V.B., *et al.*, 2010. Identification of glycosyltransferase 8 family members as xylosyltransferases acting on O-glucosylated notch epidermal growth factor repeats. *Journal of Biological Chemistry* 285, 1582–1586.
- Shaheen, R., Aglan, M., Keppler-Noreuil, K., *et al.*, 2013. Mutations in EOGT confirm the genetic heterogeneity of autosomal-recessive Adams-Oliver syndrome. *American Journal of Human Genetics* 92, 598–604.
- Stanley, P., Okajima, T., 2010. Roles of glycosylation in Notch signaling. *Current Topics in Developmental Biology* 92, 131–164.
- Tabak, L.A., 1995. In defense of the oral cavity: Structure, biosynthesis, and function of salivary mucins. *Annual Review of Physiology* 57, 547–564.
- Tenno, M., Ohtsubo, K., Hagen, F.K., *et al.*, 2007. Initiation of protein O glycosylation by the polypeptide GalNAcT-1 in vascular biology and humoral immunity. *Molecular and Cellular Biology* 27, 8783–8796.
- Tian, E., Hoffman, M.P., Ten Hagen, K.G., 2012. O-glycosylation modulates integrin and FGF signalling by influencing the secretion of basement membrane components. *Nature Communications* 3, 869.
- Tian, E., Ten Hagen, K.G., 2006. Expression of the UDP-GalNAc: Polypeptide N-acetylgalactosaminyltransferase family is spatially and temporally regulated during *Drosophila* development. *Glycobiology* 16, 83–95.
- Tian, E., Ten Hagen, K.G., 2007. A UDP-GalNAc:polypeptide N-acetylgalactosaminyltransferase is required for epithelial tube formation. *Journal of Biological Chemistry* 282, 606–614.
- Topaz, O., Shurman, D.L., Bergman, R., *et al.*, 2004. Mutations in GALNT3, encoding a protein involved in O-linked glycosylation, cause familial tumoral calcinosis. *Nature Genetics* 36, 579–581.
- Tran, D.T., Ten Hagen, K.G., 2013. Mucin-type O-glycosylation during development. *Journal of Biological Chemistry* 288, 6921–6929.
- Tran, D.T., Zhang, L., Zhang, Y., *et al.*, 2012. Multiple members of the UDP-GalNAc: Polypeptide N-acetylgalactosaminyltransferase family are essential for viability in *Drosophila*. *Journal of Biological Chemistry* 287, 5243–5252.
- Vester-Christensen, M.B., Halim, A., Joshi, H.J., *et al.*, 2013. Mining the O-mannose glycoproteome reveals cadherins as major O-mannosylated glycoproteins. *Proceedings of the National Academy of Sciences of the United States of America* 110, 21018–21023.
- Wang, L.W., Dlugosz, M., Somerville, R.P., *et al.*, 2007. O-fucosylation of thrombospondin type 1 repeats in ADAMTS-like-1/punctin-1 regulates secretion: Implications for the ADAMTS superfamily. *Journal of Biological Chemistry* 282, 17024–17031.
- Wang, Y., Jobe, S.M., Ding, X., *et al.*, 2012. Platelet biogenesis and functions require correct protein O-glycosylation. *Proceedings of the National Academy of Sciences of the United States of America* 109, 16143–16148.
- Wells, L., 2013. The o-mannosylation pathway: Glycosyltransferases and proteins implicated in congenital muscular dystrophy. *Journal of Biological Chemistry* 288, 6930–6935.
- Xia, L., Ju, T., Westmuckett, A., *et al.*, 2004. Defective angiogenesis and fatal embryonic hemorrhage in mice lacking core 1-derived O-glycans. *Journal of Cell Biology* 164, 451–459.
- Yago, T., Fu, J., McDaniel, J.M., *et al.*, 2010. Core 1-derived O-glycans are essential E-selectin ligands on neutrophils. *Proceedings of the National Academy of Sciences of the United States of America* 107, 9204–9209.
- Yamauchi, M., Sricholpech, M., 2012. Lysine post-translational modifications of collagen. *Essays in Biochemistry* 52, 113–133.
- Yoshida-Moriguchi, T., Willer, T., Anderson, M.E., *et al.*, 2013. SGK196 is a glycosylation-specific O-mannose kinase required for dystroglycan function. *Science* 341, 896–899.
- Zak, B.M., Crawford, B.E., Esko, J.D., 2002. Hereditary multiple exostoses and heparan sulfate polymerization. *Biochimica et Biophysica Acta* 1573, 346–355.
- Zhang, L., Syed, Z.A., van Dijk Härd, I., *et al.*, 2014. O-glycosylation regulates polarized secretion by modulating Tango1 stability. *Proceedings of the National Academy of Sciences of the United States of America* 111, 7296–7301.
- Zhang, L., Tran, D.T., Ten Hagen, K.G., 2010. An O-glycosyltransferase promotes cell adhesion during development by influencing secretion of an extracellular matrix integrin ligand. *Journal of Biological Chemistry* 285, 19491–19501.
- Zhang, L., Zhang, Y., Hagen, K.G., 2008. A mucin-type O-glycosyltransferase modulates cell adhesion during *Drosophila* development. *Journal of Biological Chemistry* 283, 34076–34086.

Further Reading

- Varki, A., Cummings, R.D., Esko, J.D. (Eds.), 2009. *Essentials of Glycobiology*, second ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. ISBN-13: 9780879697709.

Intra-Golgi Transport

A Aguilera Gomez, Hubrecht Institute – KNAW and the UMC Utrecht, Utrecht, the Netherlands

C Rabouille, Hubrecht Institute – KNAW and the UMC Utrecht, Utrecht, the Netherlands and University Medical Center Utrecht, Utrecht, the Netherlands

© 2016 Elsevier Inc. All rights reserved.

Glossary

Anterograde transport The movement of newly synthesized proteins from the site of synthesis to their final destination.

Cisterna Flattened membrane-bound compartment to which a defined set of Golgi resident enzymes localizes and functions.

COPI-coated vesicle 50–70 nm small circular carriers forming at the Golgi that are covered by a complex of COPI subunits forming the COPI coat.

Golgi apparatus One of the key organelle of the secretory pathway made of stacked cisternae (Golgi stack) and two membrane networks on each side of the stack.

Golgi stack A stack of Golgi cisternae.

Secretory pathway The series of membrane-bound compartments allowing the anterograde transport of proteins from the endoplasmic reticulum to the plasma membrane.

Introduction

Cell survival and communication largely depends on the secretion of circulating protein ligands (such as insulin and epithelial growth factor among many others) and plasma membrane receptors that bind these ligands, thus eliciting appropriate signaling pathways triggering growth, differentiation, proliferation, release of biological materials etc. (Farhan and Rabouille, 2011).

Although these ligands and receptors are usually synthesized in different cell types, they have in common to have reached their final destination by trafficking through the secretory pathway. This pathway is made of discrete membrane-bound organelles. The secretory cargo (proteins to be transported through the secretory pathway) are *de novo* synthesized in the endoplasmic reticulum (ER), packaged into COPII-coated carriers at ER exit sites, and transported to the Golgi complex, the central sorting and processing station of secretory cargo proteins of the secretory pathway. The passage through the Golgi ensures that they become biologically active after they are properly glycosylated and processed before being packaged for transport to their final destinations. Note that these destinations are not only the plasma membranes but also the multiple internal membrane-bound compartments, including those of the secretory pathway itself. The Golgi has the capacity of processing grams of proteins per day, such as the digestive enzymes secreted by the exocrine pancreas.

Despite this enormous protein trafficking, the Golgi complex keeps its structural integrity. At the structural level, the most basic organization of the Golgi complex found in most organisms comprises three sub-compartments, the *cis*, medial, and *trans* cisternae, flattened membrane-bound compartments. In mammalian cells, they are of 1 μ m cross-sectional diameter and stacked to form a polarized compact structure, the Golgi stack (Figure 1). A noticeable exception to the stacked organization is the Golgi in *Saccharomyces cerevisiae*, which displays only single cisternae, at least in normal lab conditions (Rambourg *et al.*, 1993).

This Golgi stack is flanked on each side by a tubular membrane network: The *cis* Golgi network (CGN) at the *cis* side of the stack where proteins that have just exited the ER enter the Golgi complex, and the *trans* Golgi network (TGN), where they exit. In mammalian cells, the stacks are connected together by tubules to form the so-called mammalian Golgi ribbon capping the nucleus (Figure 1). In non-mammalian species, the Golgi stacks are not connected and remain discrete in the cytoplasm, for instance in *Drosophila* (Kondylis and Rabouille, 2009). In addition to the stack's *cis* to *trans* polarity and to the 3D ribbon organization, there is also lateral anisotropy within each cisterna: the cisterna thin core seems to be functionally different from its rims that are often dilated and highly fenestrated (Weidman *et al.*, 1993; Figure 1).

Recognized 50 years ago by Nobel Prize winner George Palade (Palade, 1975), *cis* to *trans* anterograde transport through the Golgi is a prerequisite for almost all proteins that reach the plasma membrane and the extracellular medium. There is now a consensus about the molecular machinery mediating intra-Golgi transport, as recognized by the Nobel Prize awarded to James E. Rothman in 2013, a prize that also recognized Randy W. Schekman and Thomas S. Südhof for the elucidation of the machinery required for ER exit and neurotransmitter release, respectively. However, a general agreement on how intra-Golgi transport is mechanistically sustained is yet to be reached. We are now down to six models and not closer to reach a consensus, as groups using almost identical technology continue to reach opposite conclusions.

A Crash Course in the Machinery Involved in Intra-Golgi Transport

In order to provide a framework to place the different models by which proteins are thought to undergo intra-Golgi transport, we briefly mention here a number of key molecules required in the process (reviewed in Mellman and Warren, 2000).

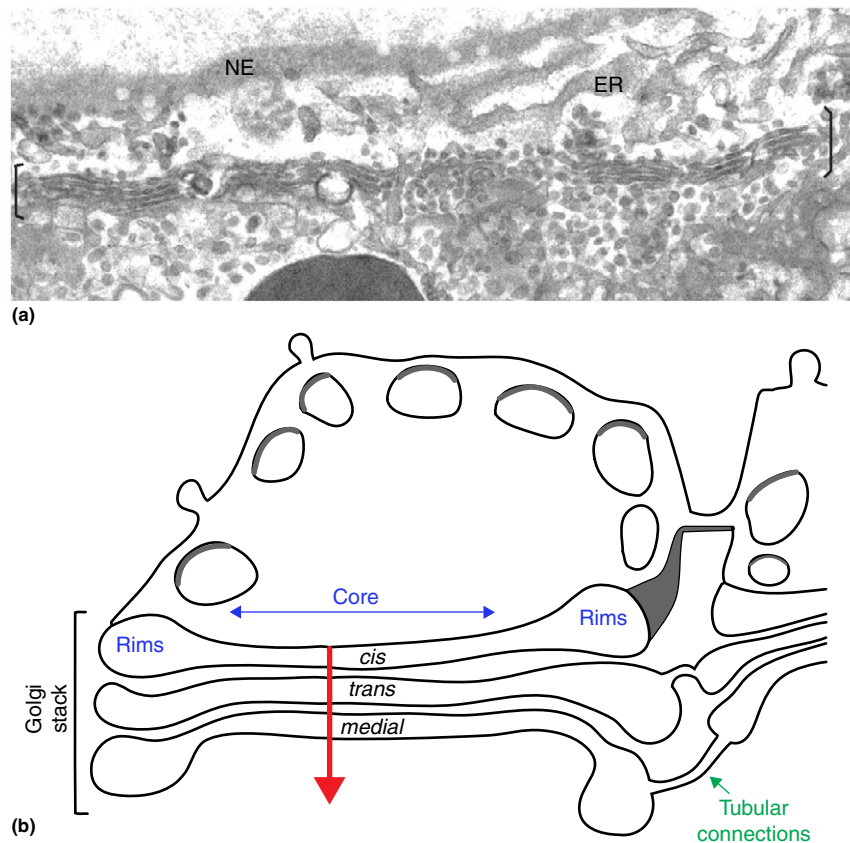


Figure 1 The Golgi complex. (a) Electron micrograph of a Golgi stack of a mouse epithelial fibroblast. A portion of the Golgi ribbon is displayed between brackets. NE, nuclear envelope; ER, endoplasmic reticulum. (b) Cartoon representation of a Golgi stack comprising 3 stacked cisternae (*cis*, medial, and *trans*), each with a core (center) and more dilated (and fenestrated) rims. Each cisterna is also laterally linked to an equivalent cisterna by tubular connections, thus forming the Golgi ribbon.

COPI-coated vesicles

Net forward movement of proteins through the secretory pathway is mediated by formation of small membrane carriers that specifically fuse with receptor compartments. One type that buds at the Golgi are COPI (coatamer protein complex I)-coated vesicles that were identified in the late 80s by the group of Jim Rothman (Malhotra *et al.*, 1989). COPI is a cytosolic 700-kDa protein complex comprising seven stoichiometric subunits, α , β , β' , γ , δ , ϵ , and ζ -COP conserved from yeast to mammals (Gaynor *et al.*, 1998) and preassembled into two cytoplasmic sub-complexes in the cytoplasm (β , γ , δ and ζ -COP subunits on one hand, and α , β and ϵ -COP subunits, on the second hand) (Watson *et al.*, 2004). Recruited upon activation of the activated small GTPase ADP-ribosylation factor 1 (Arf1), the coatamer complex is sufficient for COPI-coated vesicle formation (reviewed in Hsu and Yang, 2009).

Note that in addition to COPI-coated vesicles, two (and perhaps more) other types of coated vesicles are present in the secretory pathway. COPII-coated vesicles bud from the ER and form ER exit sites (reviewed in Lee and Schekman, 2004), whereas clathrin-coated vesicles bud from the TGN to mediate transport of newly synthesized lysosomal hydrolases to the lysosomes (as well as from the plasma membrane to mediate endocytosis of transmembrane receptors).

Tethers and Rabs

The fusion of small carriers to their receptor compartment requires at least two types of complexes. The first is made of tethering factors and their associated small Rab GTPases. In the Golgi, there are two types of tethers: (1) long coiled-coil peripheral proteins called golgins, except Giantin that spans the membrane, and GRASP65/55 that are not coiled-coils and (2) multimeric complexes, such as the COG (conserved oligomeric Golgi) (Lees *et al.*, 2010) and TRAPP (transport protein particle) (Hughson and Reinisch, 2010). All specifically associate to one (or more) of the 60 small Rab GTPases (reviewed in Pfeffer, 2013a; Ramirez and Lowe, 2009; Short *et al.*, 2005; Sinka *et al.*, 2008 for the golgins and GRASPs; Cottam and Ungar, 2012 for the COG; and Barrowman *et al.*, 2010 for the TRAPP).

SNARE complexes

The second complex required for the fusion of small carriers to their receptor compartment is made of SNARE complexes and their 'recycling/priming' machinery (trimeric triple ATPase NSF and its SNARE adapter SNAP). SNAREs are type II transmembrane proteins with most of their mass in the cytoplasm. They are functionally divided into v-SNAREs for the SNAREs associated with the vesicle and t-SNAREs for those at target compartments (Fasshauer *et al.*, 1998). Overlapping this

functional definition, SNAREs are also classified as Qa, Qb, Qc and R subfamilies. For intra-Golgi transport, the Qa t-SNARE is Syntaxin 5, Qb v-SNARE, Gos28, the Qc t-SNARE Gs15 and the R-tSNARE, Ykt6, that are assembled, as other SNAREs complexes in a parallel helical bundle (Malsam and Sollner, 2011; Sollner *et al.*, 1993; Sudhof and Rothman, 2009).

Upon budding, a coated vesicle incorporates a v-SNARE, a tether and a Rab. When the coat is shed, the vesicle is captured near the target membrane via the interaction of tethers at their surface and tether at the surface of the target compartment. This allows the formation of functional SNARE complexes and mediates bilayer fusion. When fusion is achieved, the SNARE complex must disassemble to recycle the SNARE molecules, and this is mediated by NSF and SNAP (for details, see excellent review by Malsam and Sollner, 2011).

The Vesicular Anterograde Transport Model

At the beginning of the 1980s, the group of Jim Rothman at Stanford University, CA, designed a biochemical assay aimed to dissect the different steps of intra-Golgi transport. The concept behind this assay was to follow the acquisition of endoglycosidase H (endoH) resistance of a sensitive VSV-G protein contained in a 'donor' Golgi fraction enriched from the mutant CHO cell line that lacks the N-acetylglucosaminyl-transferase. This fraction was mixed with an acceptor Golgi fraction from wild-type CHO cells. Vesicles derived from the donor Golgi fraction (thus encapsulating endoH sensitive VSV-G) would fuse with wild-type acceptor Golgi, allowing the VSV-G oligosaccharide chain to be processed and become endoH-resistant (Balch *et al.*, 1984; Fries and Rothman, 1980). This assay allowed the identification of key molecules, such as NSF (Malhotra *et al.*, 1988), SNAP (Clary *et al.*, 1990) and COPI (Malhotra *et al.*, 1989; Orci *et al.*, 1989; Waters *et al.*, 1991) and SNAREs (Sollner *et al.*, 1993).

Note that at the same time and only a few miles away, at the University of Berkeley, CA, the Schekman's group designed a forward genetic screen in yeast (Novick and Schekman, 1979) that identified 23 genes required for secretion, including the COPII components (Barlowe *et al.*, 1994; Novick *et al.*, 1980). Both breakthroughs rang the advent of molecular membrane traffic.

The 'vesicular anterograde transport model' was the first attempt to make sense of the molecular data obtained with the Golgi transport assay, and has paved the first roadmap of intra-Golgi transport. In this model, the 2–4 cisternae of the Golgi stack (each defined by a set of Golgi resident enzymes, such as glycosylation enzymes and proteases) are proposed to be static over the time scale of protein flow across it (typically 5–25 min, depending on the cargo). In this model, what moves is the secretory cargo. Upon delivery from the ER to the *cis*-most cisterna, the cargo undergoes a series of modifications specific to this cisterna. The cargo is then repackaged into COPI-coated vesicles that deliver it to the medial cisternae where it is further modified. It is then repackaged again in COPI-coated vesicles to reach the *trans* cisterna where it undergo final modifications before exiting the Golgi (reviewed in Rothman, 2010; Figure 2(a)).

This model is very much in line with the general consensus that small vesicular carriers are the means by which proteins

reach their destination, as it is the case for the first step of the secretory pathway, the exit of the ER by COPII-coated vesicles (Barlowe *et al.*, 1994). Furthermore, the use of vesicles as anterograde vehicles conceptually allows the Golgi to maintain its structural integrity.

The Cisternal Maturation Model

In the mid-late 1990s, however, the cisternal maturation model, that was prevalent in the pre-Rothman era (Morre *et al.*, 1979) started to reemerge as an alternative for intra-Golgi transport (Glick *et al.*, 1997). One of the motivations for challenging the vesicular transport model was the size that some of the cargo adopts, such as long algal scales (Becker *et al.*, 1995) and rod shape trimeric collagen complex of 300 nm, both way too large to fit into a 50–100 nm COPI-coated vesicles.

The cisternal maturation model has two tenets: Tenet 1 proposes that the cargo is static in one membrane compartment that undergoes changes of its identity. In this model, the part that moves is the processing machinery, not the secretory cargo. Exiting the ER, the cargo is present in a compartment at the *cis* side of the most *cis* Golgi cisterna. This compartment acquires *cis* Golgi resident enzymes and therefore becomes a *bonafide cis* cisterna. Tenet 2 proposes that the acquisition of cisternal identity is mediated by the retrograde movement of these enzymes by COPI-coated vesicles from the preexisting *cis* cisterna that becomes devoid of *cis* markers. But in the same time frame, this cisterna acquires medial markers from the preexisting medial cisterna and becomes a medial cisterna (Figure 2(b)), whereas the medial cisterna becomes a *trans* cisterna by the same retrograde movement of *trans* markers.

Overall, the features that best distinguish the two models (cisternal maturation vs. vesicular anterograde transport) are the content and directionality of the COPI-coated vesicles that bud from the rims of Golgi cisternae. In the vesicular transport model, COPI vesicles carry cargo and move from one Golgi cisterna to the next in an anterograde *cis* > medial > *trans* fashion. In the cisternal maturation model, the COPI-coated vesicles move in a retrograde *trans* > medial > *cis* fashion and function as a retrieving device that is used by Golgi enzymes to maintain their specific and differential localization over the Golgi stack (Rabouille and Klumperman, 2005).

The evidence for the cisternal maturation model is to be found in four separate sets of results:

- Large cargo, such as collagen has been shown to move through the Golgi stack without entering vesicular structures (Bonfanti *et al.*, 1998; Mironov *et al.*, 2001).
- Purified COPI-coated vesicles are shown to contain a higher concentration of Golgi enzymes/proteins than secretory cargo (Lanoix *et al.*, 1999, 2001), supporting the role of COPI vesicles in cisternal maturation (Gannon *et al.*, 2011; Gilchrist *et al.*, 2006). Conversely, abundant cargo, such as albumin appears depleted in these vesicles when compared to the Golgi stack (Gilchrist *et al.*, 2006).
- COPI-coated vesicles observed *in vivo* by immuno-electron microscopy are shown to have 1.5-fold more Mannosidase II, a medial Golgi resident enzyme, than the adjacent cisternae (Martinez-Menarguez *et al.*, 2001). Furthermore,

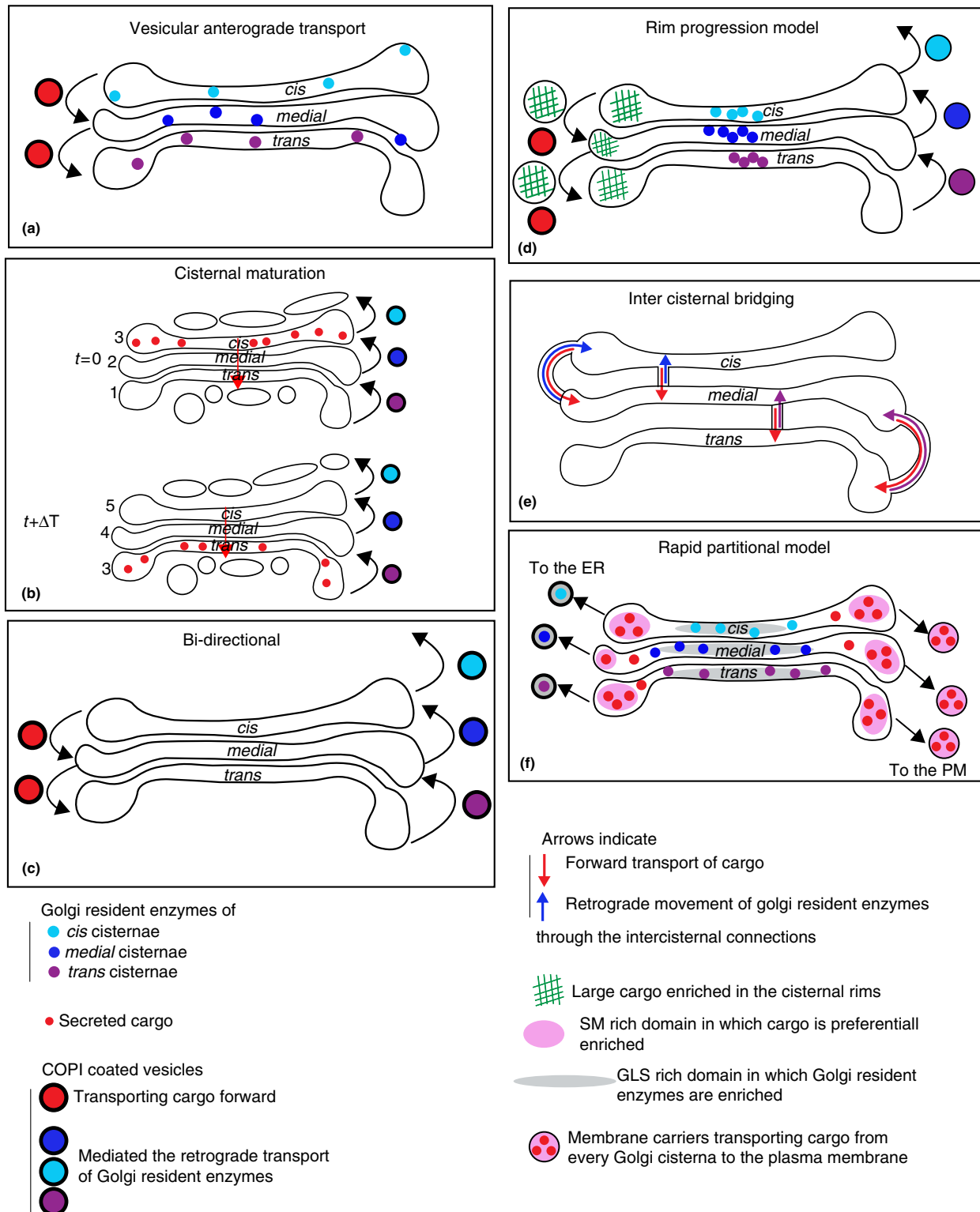


Figure 2 The different models by which proteins are proposed to undergo intra-Golgi transport. (a) Vesicular anterograde transport model: The Golgi stack and its resident proteins are static while the secretory cargo moves through the cisternae in a *cis* to *trans* fashion via COPI-coated vesicles. (b) Cisternal maturation model: The cargo remains in one cisterna that acquires in turn *cis*, medial and *trans* identity by the COPI-coated vesicle mediated retrograde movement of Golgi resident enzymes. (c) Bidirectional model: COPI-coated vesicles mediate the anterograde transport of cargo and the retrograde transport of Golgi enzymes. (d) Rim progression model: Large cargo is transported forward by rearrangement/budding of one cisternal rim and fusion to the next. In this model, the Golgi enzymes are proposed to remain in the cisternal core. (e) Intercisternal continuities: Vesicles are not involved in the cargo anterograde transport that is mediated by intercisternal tubular bridges. Golgi resident enzymes are also moving through the tubules as well as in COPI-coated vesicles. (f) Rapid partitioning model. The cargo is able to leave the Golgi from any cisternae after being enriched in sphingolipid domains whereas the Golgi enzymes are retained in the stack in glycerophospholipids.

GFP tagged GalNAc-Transferase 2 is shown to be present in COPI-coated vesicles identified by 3D-electron tomography after GFP-triggered DAB precipitation following illumination (Grabenbauer *et al.*, 2005).

- State of the art live cell imaging techniques in yeast has shown the conversion of a *cis* cisterna into a *trans* one as predicted by the model (Losev *et al.*, 2006; Matsuura-Tokita *et al.*, 2006).

The COPI-mediated Bidirectional Transport Model

Although many compelling evidences (listed above) have been gathered in support of cisternal maturation, the debate remained open. The 'first strategy' to support the vesicular transport model has been to get evidence that COPI-coated vesicles also contain secretory cargo (moving in an anterograde manner) in addition to Golgi enzymes (moving in a retrograde manner; Figure 2(c), tenet 2). Using the same technique of immuno-electron microscopy as the Klumperman's group (Martinez-Menarguez *et al.*, 2001), the Orci's group showed that in insulin producing cells, COPI-coated vesicles budding at every level of the Golgi stack contain both the KDEL receptor (retrograde-directed) and proinsulin (anterograde-directed) (Orci *et al.*, 1997).

The Warren's group then showed the coexistence of two classes of COPI-coated vesicles: The first class contains Golgin 84 and CASP, lacks cargo receptor of the p24 family and is enriched in Golgi enzymes. Manipulation of these golgins inhibits the delivery of Golgi enzymes to the ER, supporting the role of these first class of COPI vesicles in retrograde transport. The second class of COPI-coated vesicles contains p115 and giantin, and uses GM130 and Rab1 to be tethered to the CGN in an anterograde manner (Malsam *et al.*, 2005). These results suggest that COPI-coated vesicles can mediate the differential retrograde and anterograde movement of two different types of molecules, whether this is related to the existence of multiple coatomer complexes isolated by the group F. Wieland (Wegmann *et al.*, 2004) remains to be established.

Recently, a study aiming to monitor Golgi transport by live cell imaging in mammalian cells supported the ability of COPI-coated vesicles to move both secretory cargo and Golgi resident enzymes. This study uses fused cells containing two types of Golgi, one marked by a small secretory cargo coupled to a fluorophore and the other marked by a Golgi resident enzyme coupled to a different fluorophore. Both cargo and Golgi enzymes are transported in small carriers with properties consistent with COPI-coated vesicles, suggesting that these vesicles have the ability to incorporate both types of molecules (Pellett *et al.*, 2013). However, this assay monitors material exchange between two Golgis (inter-Golgi), not intra-Golgi transport. It is therefore possible that COPI delivers cargo (and Golgi enzymes) from one given cisterna in one Golgi to an equivalent cisterna in the second Golgi. It does not directly address how cargo progresses through the stack.

In summary, COPI-coated vesicles may contain anterograde cargo as well as Golgi enzymes, and they might mediate anterograde as well as retrograde transport.

The 'second strategy' to support the vesicular transport model has been to directly test and disprove one critical prediction of the cisternal maturation model, according to which,

once a cargo is trapped in a given Golgi cisterna, it progresses forward without leaving this cisterna (tenet 1). The group of Rothman at Yale developed a new method (Lavieau *et al.*, 2013) exactly designed to follow forward movement of cargo of different size within the Golgi stack. The first cargo used in this study is a chimeric type I membrane protein made of CD8 fused to 4 copies of the dimerization motif FM at its N-terminal luminal domain (4FM-CD8). The presence of these FM domains makes this protein to spontaneously self-oligomerize. Conversely, addition to a specific small molecule (AP21998) triggers the release of monomers (Rollins *et al.*, 2000). When expressed in any compartments of the secretory pathway, the self-aggregated protein is trapped. For instance in the ER, it forms small aggregates (staples) that spans and compacts the width of the ER tubules, as if 'stapled.'

By manipulating the culture conditions in the presence and absence of the disaggregating drug, staples were allowed to form either in the *cis* or in the *trans* cisternae of the Golgi stack. Once formed, for instance in the *cis* cisterna, the staples were found not to move, and therefore did not traverse the Golgi stack. As controls, monomeric 4FM-CD8 moved to the cell surface (Lavieau *et al.*, 2013). The fact that 4FM-CD8 staples remain where they are generated supports a model in which Golgi compartments are static.

Staple technology has nevertheless limitations that are subjected to interpretation. Indeed, A. Luini's group (Napoli, Italy) used an almost identical staple strategy and reached opposite results (Rizzo *et al.*, 2013). In this study, instead of 4FM-CD8 (see above), the type II protein Golgi mannosidase I was appended by 3 FM domains at the C-terminal (MannI-3FM). Again it self-aggregates and forms plaques in the *cis* Golgi cisterna as expected, but the main difference is that these plaques are later found in the *trans* cisterna. This suggests that they have moved forward, contrary to the 4FM-CD8 staples, but in agreement with the prediction of the cisternal maturation model.

Whether the sharp difference in these results arises from using a cargo (CD8) instead of a resident enzyme (MannI), a type I instead of the type II protein, 3 FM domains instead of 4, at high or low expression, is not elucidated, and we are left with no clearer answers (Morriswood and Warren, 2013).

The Rim Progression Model Specifically Addresses Large Cargo Transport

As mentioned above, the original motivation to challenge the vesicular anterograde transport model was that large cargo could not fit into the small 50–70 nm COPI-coated vesicles. So, the 'third strategy' to validate this model was to explain how large cargo moves from cisterna to cisterna across the Golgi stack. The existence of mega vesicles moving large cargo forward (Volchuk *et al.*, 2000) has recently been reinvestigated using the staple technology (see above). Indeed, large aggregates of a cargo can be generated, but this time with the distinct difference that these aggregates are not attached to the Golgi membrane by a transmembrane domain as in the case of 4FM-CD8 (Lavieau *et al.*, 2013). Remarkably, these aggregates move through the Golgi stacks through cisternal rims. This has led to the so-called 'Rim progression model' that proposes that large cargo are enriched at the cisternal rims and that the rims

progress forward by rearrangement/budding from one cisterna and fusion to the next, very much alike what is proposed for COPI vesicles enriched in small cargo (that is independently packaged) (Figure 2(d)).

There are, however, two major difference with the COPI-coated vesicles. The fission of COPI-coated vesicles is well documented and largely depends on the coat itself (Adolf *et al.*, 2013), whereas nothing is known on the fission of these rim-derived mega vesicles. Second, it appears that a Golgi ribbon is necessary for the intra-Golgi transport of large cargo. When the mammalian Golgi ribbon is broken into ministacks, the transport of large cargo is significantly slowed down whereas small cargo (such as VSV-G) is unaffected (Lavieu *et al.*, 2014). This suggests that large vesicles generated from one cisterna in one given Golgi stack cannot fuse with the next cisterna of the same stack (see below VI for a possible mechanism).

The rim progression model also proposes that Golgi resident enzymes are enriched in the cisternal core (Cosson *et al.*, 2002) and that they do not diffuse to the rims. How the Golgi enzymes are retained in the cisterna core in a manner that also allows their dynamic incorporation in COPI-coated vesicles for recycling to the ER but not their diffusion in the rims remains to be investigated. In this regard, the kin-oligomerisation model proposed by the Warren's group (Nilsson *et al.*, 1994) could be useful in explaining both the retention/localization and the dynamics of Golgi resident enzymes. The stalks and transmembrane domains of Golgi enzymes localizing in the same set of cisternae have been shown to form oligomers (kin-oligomer), helping in their retention. Modulation of the size of this oligomer (through mechanisms that are still to be defined) could allow the enzyme recycling to the ER (between 5% and 10%) via COPI-coated vesicles.

The Cisternal Progenitor Model

How can rims participate in forward movement large cargo and more generally of secretory cargo?

The first consideration put forward by the Pfefer' group (Pfeffer, 2013a) is that in mammalian cells, the rims are very dynamics and the site of many lateral fusion/fission events. For instance, the depolymerization of microtubules, the depletion of Golgin 84 (Diao *et al.*, 2003), and the double depletion of GRASP65 and 55 (Jarvela and Linstedt, 2014) leads to the dispersion of Golgi ribbon in ministacks. Second, it is now well established from research in the endocytic field that transport vectoriality is provided by small Rab GTPases. They decorate the surfaces of almost all membrane compartments (including the Golgi) where they catalyze the formation of membrane microdomains (particularly clear on the endosomal system (Sonnichsen *et al.*, 2000). The vectoriality is provided by Rab cascade of Rab activation by specific guanine nucleotide exchange factors (GEFs) and inactivation by specific GTPase-activating proteins (GAPs) (Chesneau *et al.*, 2012; Pfeffer, 2013b; Poteryaev *et al.*, 2010; Rink *et al.*, 2005; Rivera-Molina and Novick, 2009).

Applied to the Golgi, this Rab cascade has led to the following model: A Golgi stack comprises three tightly stacked cisternae. The *cis* cisterna is marked by RabA, the medial by RabB, and the *trans* by RabC. RabA defines a domain

containing the machinery facilitating homologous fusion between RabA containing cisterna. RabB and RabC would have the same property for the medial and *trans* cisterna. This would account for, for instance, the fusion of ministacks into a Golgi ribbon. RabA can also recruit the GEF for RabB. As a result, a RabA domain leads to the formation of an adjacent RabB domain but because of the ministack dynamics, the two domains could be separated. This new RabB compartment deriving for the *cis* cisterna can then fuse with the compartment in which RabB is the most enriched, the medial cisterna. If this new RabB compartment is a rim enriched with large cargo, it would essentially equate with the fusion of the *cis* cisterna's rim with the medial cisterna, exactly what the rim progression proposes.

At the same time, RabB could recruit the GEF for RabC and the same scenario could repeat itself so that a new RabC compartment deriving from the medial cisterna would fuse with the *trans* cisterna (Pfeffer, 2013b). In conclusion, one or several Rab cascades could in principle explain the forward progression of the cisternal rims. It would also explain why a Golgi ribbon is conducive for rim progression and (see above).

In conclusion, many evidences have been gathered to support either the cisternal maturation model or the vesicular anterograde transport model in its extended version that includes rim progression. As stated above, the issue is, however, not settled and is further complicated by the description of two new models.

Intercisternal Continuities. Is COPI No Longer Involved?

At the beginning of the new century, in what seemed a complete departure from the involvement of COPI-coated vesicle in intra-Golgi transport, a new model emerged. This was based on observations of intercisternal bridges that seemed to form under high cargo loading, such as VSV-G transfection of HeLa cells (Trucco *et al.*, 2004) or glucose stimulation of pancreatic islets (Marsh *et al.*, 2004). Till this time, the Golgi cisternae were considered to be discrete sub-compartments of the Golgi stack, and although each cisterna is linked laterally to an equivalent adjacent cisterna to form the Golgi ribbon (at least in mammalian cells, see Figure 1), intercisternal bridges were not readily (to almost never) observed.

This lack of evidence might be due to the fact that these bridges cannot be resolved by light microscopy. Furthermore, due to their small size and tubular nature, they would mostly be missed in EM using ultrathin sections. In this regard, 3D-electron tomography and its application on cellular structures provided the first evidence for their existence. In the pancreatic islet, few T-shaped tubules are found to connect the flat core part of two cisternae (Marsh *et al.*, 2004). In VSV-G expressing HeLa cells, the tubules are reported to be more numerous and tended to preferentially connect the rims (Trucco *et al.*, 2004).

The relevance of these tubules has become topical to Golgi transport as they are shown to play a role during intra-Golgi trafficking of several secretory cargo (Figure 2(e)). This would therefore suggest that COPI-coated vesicles are no longer involved in transport. Indeed, intra-Golgi tubulation (as well inter-Golgi tubulation generating the Golgi ribbon) appears to be induced by the enzyme cPLA2 (San Pietro *et al.*, 2009).

cPLA2 inhibition prevents both types of Golgi tubulation and results in the inhibition of anterograde transport (San Pietro *et al.*, 2009). A recent article from the same group now reports that albumin uses these intercisternal continuities for its secretion, but not pro-collagen, suggesting that multiple types of transport might occur in the same stack (Beznoussenko *et al.*, 2014). Golgi enzymes also seem to also move from cisterna to cisterna via these tubules (Yang *et al.*, 2011).

Is the COPI coat out of the equation with the involvement of intercisternal continuities in intra-Golgi transport? A membrane bud coated with COPI is meant to become a COPI-coated vesicle but under certain circumstances, it can also become an intercisternal tubule. This bud fate decision is mediated by the interplay between two enzymes. When lyso-phosphatidic acid (LPA) acyltransferase type γ (LPAAT- γ) that promotes COPI-coated vesicle formation is inhibited, tubules grow. cPLA2 inhibits LPAAT- γ , thus promoting tubule growth at the expense of vesicles. As expected, the COPI complex is necessary for the formation of the vesicles (that according to the authors carry Golgi enzymes in the retrograde fashion), but unexpectedly, also the tubules through which small secretory cargo and enzymes move (Yang *et al.*, 2011).

In conclusion, the role for COPI vesicles in anterograde is put under scrutiny but not the role of the COPI coat itself that is here shown to be instrumental in the formation of intercisternal continuities.

The Rapid Partitioning Model

The last model described here has been put forward in 2008 by the group of J. Lippincott-Schwartz (NIH, Bethesda) (Patterson *et al.*, 2008). This group studied the dynamics of small secretory cargo (VSV-G-GFP) through the secretory pathway and got results that did not fit the cisternal maturation model. This model predicts that once a cargo enters the Golgi, it would remain in this organelle a given time to allow the maturation to take place, and leave the Golgi at the TGN with a linear kinetics, in line with the consumption of a single compartment. Instead, the Lippincott-Schwartz's group found no lag time and an exponential kinetics. From this, they concluded that a cargo is able to leave the Golgi from any cisterna, not only from the TGN. Their second observation was that Golgi resident enzymes and cargo did not completely overlap within a single Golgi cisterna, suggesting the existence of domains within cisternae.

In an attempt to integrate the kinetic data mentioned above and the existence of these two domains, lipid-trafficking pathways were introduced as the driving force for protein trafficking. The Golgi contains two major lipids, the glycerophospholipids (GPLs) and sphingolipids (SLs) that have significantly different synthetic and trafficking routes. On one hand, GPLs are synthesized in the ER and are transported to the Golgi in vesicles together with proteins. On the other hand, SLs are uniquely synthesized in the Golgi (Levine *et al.*, 2000; van Meer *et al.*, 2008) from ceramide that is itself produced in the ER but transferred directly to the trans Golgi by a lipid carrier protein, CERT (D'Angelo *et al.*, 2007; Halter *et al.*, 2007).

The rapid partitioning model proposes that each cisterna (whether *cis*, medial or *trans*) contains two domains: The first is enriched in SLs, and a large proportion of secretory cargo (that has a higher affinity for this lipid) partitions in this

domain. This is consistent with observations that SLs are delivered to the PM from the Golgi by transport carriers that are formed at all levels of the Golgi (Helms and Zurzolo, 2004; Rodriguez-Boulán *et al.*, 2005; van Meer *et al.*, 2008) and preferably at the rims. The second domain is enriched in GPLs in which Golgi processing enzymes (that have a higher affinity for this lipid) partition. This has been proposed to be part of the localization/retention of glycosylation enzymes within the Golgi by S. Munro's group (Bretscher and Munro, 1993).

Thus, a SM-rich domain containing the partitioned cargo could be incorporated in transport carriers, the nature of which needs to be determined. Concomitantly, the glycosylation enzymes would be retained in the GLM domain of the Golgi, thus both retained and retrieved back to the ER- in GLM-rich carriers (Figure 2(f)), possibly in COPI-coated vesicles.

How mechanisms of Golgi enzyme retention in the Golgi that have been proposed in the past (Nilsson *et al.*, 1994; Opat *et al.*, 2001; Schmitz *et al.*, 2008; Tu *et al.*, 2008) fit in the rapid partitioning is, however, not addressed. For instance, could kin-oligomerisation of Golgi enzymes (see above, Nilsson *et al.*, 1994) be modulated by GLM?

Conclusions/perspectives

The journey ends here with an entire library of evidences in support of each model, ranging from live cell imaging to EM to 3D-electron tomography, transport assays, biochemical assays, mathematical modeling, and mass spectrometry, in yeast and several mammalian cells. In 2009, the community agrees to disagree in an amiable fashion (Emr *et al.*, 2009).

Whether this reflects fundamental biological differences in cargo types, cell types, growth conditions, tissue-cultured conditions versus *in vivo*, is starting to be addressed. In particular, the departure from VSV-G as a typical anterograde secretory cargo has been initiated, for instance in the flow cytometry-based assay for measuring constitutive secretion from the Peden' lab (Gordon *et al.*, 2010), and the RUSH system from the Perez's lab (Boncompain *et al.*, 2012). The latter allows controlled studies of different cargo molecule trafficking and might reveal interesting differences that might reconcile aspects of the proposed models. The next step will be to study intra-Golgi transport *in vivo* with endogenous proteins tagged in the genome.

Acknowledgment

AA is supported by an open ALW-NWO grant (822-020-016).

See also: Interorganellar Communication: Components: Vesicle Tethers. Interorganellar Communication: Interplay and Processes: ER–Golgi Transport. Organelles: Structure and Function: Golgi and TGN

References

- Adolf, F., Herrmann, A., Hellwig, A., *et al.*, 2013. Scission of COPI and COPII vesicles is independent of GTP hydrolysis. *Traffic (Copenhagen, Denmark)* 14, 922–933.

- Balch, W.E., Dunphy, W.G., Braell, W.A., Rothman, J.E., 1984. Reconstitution of the transport of protein between successive compartments of the Golgi measured by the coupled incorporation of N-acetylglucosamine. *Cell* 39, 405–416.
- Barlowe, C., Orci, L., Yeung, T., *et al.*, 1994. COPII: A membrane coat formed by Sec proteins that drive vesicle budding from the endoplasmic reticulum. *Cell* 77, 895–907.
- Barrowman, J., Bhandari, D., Reinisch, K., Ferro-Novick, S., 2010. TRAPP complexes in membrane traffic: Convergence through a common Rab. *Nature Reviews Molecular Cell Biology* 11, 759–763.
- Becker, B., Bolinger, B., Melkonian, M., 1995. Anterograde transport of algal scales through the Golgi complex is not mediated by vesicles. *Trends in Cell Biology* 5, 305–307.
- Beznoussenko, G.V., Parashuraman, S., Rizzo, R., *et al.*, 2014. Transport of soluble proteins through the Golgi occurs by diffusion via continuities across cisternae. *eLife* e02009.
- Boncompain, G., Divoux, S., Gareil, N., *et al.*, 2012. Synchronization of secretory protein traffic in populations of cells. *Nature Methods* 9, 493–498.
- Bonfanti, L., Mironov Jr., A.A., Martinez-Menarguez, J.A., *et al.*, 1998. Procollagen traverses the Golgi stack without leaving the lumen of cisternae: Evidence for cisternal maturation. *Cell* 95, 993–1003.
- Bretscher, M.S., Munro, S., 1993. Cholesterol and the Golgi apparatus. *Science* 261, 1280–1281.
- Chesneau, L., Dambournet, D., Machicoane, M., *et al.*, 2012. An ARF6/Rab35 GTPase cascade for endocytic recycling and successful cytokinesis. *Current Biology: CB* 22, 147–153.
- Clary, D.O., Griff, I.C., Rothman, J.E., 1990. SNAPS, a family of NSF attachment proteins involved in intracellular membrane fusion in animals and yeast. *Cell* 61, 709–721.
- Cosson, P., Amherdt, M., Rothman, J.E., Orci, L., 2002. A resident Golgi protein is excluded from peri-Golgi vesicles in NRK cells. *Proceedings of the National Academy of Sciences of the United States of America* 99, 12831–12834.
- Cottam, N.P., Ungar, D., 2012. Retrograde vesicle transport in the Golgi. *Protoplasma* 249, 943–955.
- D'Angelo, G., Polishchuk, E., Di Tullio, G., *et al.*, 2007. Glycosphingolipid synthesis requires FAPP2 transfer of glucosylceramide. *Nature* 449, 62–67.
- Diao, A., Rahman, D., Pappin, D.J., Lucocq, J., Lowe, M., 2003. The coiled-coil membrane protein golgin-84 is a novel rab effector required for Golgi ribbon formation. *Journal of Cell Biology* 160, 201–212.
- Emr, S., Glick, B.S., Linstedt, A.D., *et al.*, 2009. Journeys through the Golgi-taking stock in a new era. *Journal of Cell Biology* 187, 449–453.
- Farhan, H., Rabouille, C., 2011. Signalling to and from the secretory pathway. *Journal of Cell Science* 124, 171–180.
- Fasshauer, D., Sutton, R.B., Brunger, A.T., Jahn, R., 1998. Conserved structural features of the synaptic fusion complex: SNARE proteins reclassified as Q- and R-SNAREs. *Proceedings of the National Academy of Sciences of the United States of America* 95, 15781–15786.
- Fries, E., Rothman, J.E., 1980. Transport of vesicular stomatitis virus glycoprotein in a cell-free extract. *Proceedings of the National Academy of Sciences of the United States of America* 77, 3870–3874.
- Gannon, J., Bergeron, J.J., Nilsson, T., 2011. Golgi and related vesicle proteomics: Simplify to identify. *Cold Spring Harbor Perspectives in Biology* 3. doi:10.1101/cshperspect.a005421.
- Gaynor, E.C., Graham, T.R., Emr, S.D., 1998. COPI in ER/Golgi and intra-Golgi transport: Do yeast COPI mutants point the way? *Biochimica et Biophysica Acta* 1404, 33–51.
- Gilchrist, A., Au, C.E., Hiding, J., *et al.*, 2006. Quantitative proteomics analysis of the secretory pathway. *Cell* 127, 1265–1281.
- Glick, B.S., Elston, T., Oster, G., 1997. A cisternal maturation mechanism can explain the asymmetry of the Golgi stack. *FEBS Letters* 414, 177–181.
- Gordon, D.E., Bond, L.M., Sahlender, D.A., Peden, A.A., 2010. A targeted siRNA screen to identify SNAREs required for constitutive secretion in mammalian cells. *Traffic* 11, 1191–1204.
- Grabenbauer, M., Geerts, W.J., Fernandez-Rodriguez, J., *et al.*, 2005. Correlative microscopy and electron tomography of GFP through photooxidation. *Nature Methods* 2, 857–862.
- Halter, D., Neumann, S., van Dijk, S.M., *et al.*, 2007. Pre- and post-Golgi translocation of glucosylceramide in glycosphingolipid synthesis. *Journal of Cell Biology* 179, 101–115.
- Helms, J.B., Zurzolo, C., 2004. Lipids as targeting signals: Lipid rafts and intracellular trafficking. *Traffic* 5, 247–254.
- Hsu, V.W., Yang, J.S., 2009. Mechanisms of COPI vesicle formation. *FEBS Letters* 583, 3758–3763.
- Hughson, F.M., Reinisch, K.M., 2010. Structure and mechanism in membrane trafficking. *Current Opinion in Cell Biology* 22, 454–460.
- Jarvela, T., Linstedt, A.D., 2014. Isoform-specific tethering links the Golgi ribbon to maintain compartmentalization. *Molecular Biology of the Cell* 25, 133–144.
- Kondylis, V., Rabouille, C., 2009. The Golgi apparatus: Lessons from Drosophila. *FEBS Letters* 583, 3827–3838.
- Lanoix, J., Ouwendijk, J., Lin, C.C., *et al.*, 1999. GTP hydrolysis by arf-1 mediates sorting and concentration of Golgi resident enzymes into functional COP I vesicles. *EMBO Journal* 18, 4935–4948.
- Lanoix, J., Ouwendijk, J., Stark, A., *et al.*, 2001. Sorting of Golgi resident proteins into different subpopulations of COPI vesicles: A role for ArfGAP1. *Journal of Cell Biology* 155, 1199–1212.
- Lavie, G., Dunlop, M.H., Lerich, A., *et al.*, 2014. The golgi-ribbon structure facilitates anterograde transport of large cargoes. *Molecular Biology of the Cell* 25, 3028–3036.
- Lavie, G., Zheng, H., Rothman, J.E., 2013. Stapled Golgi cisternae remain in place as cargo passes through the stack. *eLife* 2, e00558.
- Lee, M.C.S., Schekman, R., 2004. Cell biology. BAR domains go on a bender. *Science (New York, N.Y.)* 303, 479–480.
- Lees, J.A., Yip, C.K., Walz, T., Hughson, F.M., 2010. Molecular organization of the COG vesicle tethering complex. *Nature Structural & Molecular Biology* 17, 1292–1297.
- Levine, T.P., Wiggins, C.A., Munro, S., 2000. Inositol phosphorylceramide synthase is located in the Golgi apparatus of *Saccharomyces cerevisiae*. *Molecular Biology of the Cell* 11, 2267–2281.
- Losev, E., Reinke, C.A., Jellen, J., *et al.*, 2006. Golgi maturation visualized in living yeast. *Nature* 441, 1002–1006.
- Malhotra, V., Orci, L., Glick, B.S., Block, M.R., Rothman, J.E., 1988. Role of an N-ethylmaleimide-sensitive transport component in promoting fusion of transport vesicles with cisternae of the Golgi stack. *Cell* 54, 221–227.
- Malhotra, V., Serafini, T., Orci, L., Shepherd, J.C., Rothman, J.E., 1989. Purification of a novel class of coated vesicles mediating biosynthetic protein transport through the Golgi stack. *Cell* 58, 329–336.
- Malsam, J., Satoh, A., Pelletier, L., Warren, G., 2005. Golgin tethers define subpopulations of COPI vesicles. *Science* 307, 1095–1098.
- Malsam, J., Sollner, T.H., 2011. Organization of SNAREs within the Golgi stack. *Cold Spring Harbor Perspectives in Biology* 3, a005249.
- Marsh, B.J., Volkman, N., McIntosh, J.R., Howell, K.E., 2004. Direct continuities between cisternae at different levels of the Golgi complex in glucose-stimulated mouse islet beta cells. *Proceedings of the National Academy of Sciences of the United States of America* 101, 5565–5570.
- Martinez-Menarguez, J.A., Prekeris, R., Oorschot, V.M., *et al.*, 2001. Peri-Golgi vesicles contain retrograde but not anterograde proteins consistent with the cisternal progression model of intra-Golgi transport. *Journal of Cell Biology* 155, 1213–1224.
- Matsuura-Tokita, K., Takeuchi, M., Ichihara, A., Mikuriya, K., Nakano, A., 2006. Live imaging of yeast Golgi cisternal maturation. *Nature* 441, 1007–1010.
- van Meer, G., Voelker, D.R., Feigenson, G.W., 2008. Membrane lipids: Where they are and how they behave. *Nature Reviews Molecular Cell Biology* 9, 112–124.
- Mellman, I., Warren, G., 2000. The road taken: Past and future foundations of membrane traffic. *Cell* 100, 99–112.
- Mironov, A.A., Beznoussenko, G.V., Nicoziani, P., *et al.*, 2001. Small cargo proteins and large aggregates can traverse the Golgi by a common mechanism without leaving the lumen of cisternae. *Journal of Cell Biology* 155, 1225–1238.
- Morre, D.J., Kartenbeck, J., Franke, W.W., 1979. Membrane flow and interconversions among endomembranes. *Biochimica et Biophysica Acta* 559, 52–71.
- Morriswood, B., Warren, G., 2013. Cell biology. Stalemate in the Golgi battle. *Science* 341, 1465–1466.
- Nilsson, T., Hoe, M.H., Slusarewicz, P., *et al.*, 1994. Kin recognition between medial Golgi enzymes in HeLa cells. *EMBO Journal* 13, 562–574.
- Novick, P., Field, C., Schekman, R., 1980. Identification of 23 complementation groups required for post-translational events in the yeast secretory pathway. *Cell* 21, 205–215.
- Novick, P., Schekman, R., 1979. Secretion and cell-surface growth are blocked in a temperature-sensitive mutant of *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* 76, 1858–1862.
- Opat, A.S., Houghton, F., Gleeson, P.A., 2001. Steady-state localization of a medial-Golgi glucosyltransferase involves transit through the trans-Golgi network. *Biochemical Journal* 358, 33–40.
- Orci, L., Malhotra, V., Amherdt, M., Serafini, T., Rothman, J.E., 1989. Dissection of a single round of vesicular transport: Sequential intermediates for intercisternal movement in the Golgi stack. *Cell* 56, 357–368.
- Orci, L., Starnes, M., Ravazzola, M., *et al.*, 1997. Bidirectional transport by distinct populations of COPI-coated vesicles. *Cell* 90, 335–349.
- Palade, G., 1975. Intracellular aspects of the process of protein synthesis. *Science* 189, 867.

- Patterson, G.H., Hirschberg, K., Polishchuk, R.S., *et al.*, 2008. Transport through the Golgi apparatus by rapid partitioning within a two-phase membrane system. *Cell* 133, 1055–1067.
- Pellett, P.A., Dietrich, F., Bewersdorf, J., Rothman, J.E., Lavieu, G., 2013. Inter-Golgi transport mediated by COPI-containing vesicles carrying small cargoes. *eLife* 2, e01296.
- Pfeffer, S.R., 2013a. Hopping rim to rim through the Golgi. *eLife* 2, e00903.
- Pfeffer, S.R., 2013b. Rab GTPase regulation of membrane identity. *Current Opinion in Cell Biology* 25, 414–419.
- Poteryaev, D., Datta, S., Ackema, K., Zerial, M., Spang, A., 2010. Identification of the switch in early-to-late endosome transition. *Cell* 141, 497–508.
- Rabouille, C., Klumperman, J., 2005. Opinion: The maturing role of COPI vesicles in intra-Golgi transport. *Nature Reviews Molecular Cell Biology* 6, 812–817.
- Rambourg, A., Clermont, Y., Kepes, F., 1993. Modulation of the Golgi apparatus in *Saccharomyces cerevisiae* sec7 mutants as seen by three-dimensional electron microscopy. *Anatomical Record* 237, 441–452.
- Ramirez, I.B., Lowe, M., 2009. Golgins and GRASPs: Holding the Golgi together. *Seminars in Cell & Developmental Biology* 20, 770–779.
- Rink, J., Ghigo, E., Kalaidzidis, Y., Zerial, M., 2005. Rab conversion as a mechanism of progression from early to late endosomes. *Cell* 122, 735–749.
- Rivera-Molina, F.E., Novick, P.J., 2009. A Rab GAP cascade defines the boundary between two Rab GTPases on the secretory pathway. *Proceedings of the National Academy of Sciences of the United States of America* 106, 14408–14413.
- Rizzo, R., Parashuraman, S., Mirabelli, P., *et al.*, 2013. The dynamics of engineered resident proteins in the mammalian Golgi complex relies on cisternal maturation. *Journal of Cell Biology* 201, 1027–1036.
- Rodriguez-Boulan, E., Kreitzer, G., Musch, A., 2005. Organization of vesicular trafficking in epithelia. *Nature Reviews Molecular Cell Biology* 6, 233–247.
- Rollins, C.T., Rivera, V.M., Woolfson, D.N., *et al.*, 2000. A ligand-reversible dimerization system for controlling protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America* 97, 7096–7101.
- Rothman, J.E., 2010. The future of Golgi research. *Molecular Biology of the Cell* 21, 3776–3780.
- San Pietro, E., Capestrano, M., Polishchuk, E.V., *et al.*, 2009. Group IV phospholipase A(2)alpha controls the formation of inter-cisternal continuities involved in intra-Golgi transport. *PLoS Biology* 7, e1000194.
- Schmitz, K.R., Liu, J., Li, S., *et al.*, 2008. Golgi localization of glycosyltransferases requires a Vps74p oligomer. *Developmental Cell* 14, 523–534.
- Short, B., Haas, A., Barr, F.A., 2005. Golgins and GTPases, giving identity and structure to the Golgi apparatus. *Biochimica et Biophysica Acta* 1744, 383–395.
- Sinka, R., Gillingham, A.K., Kondylis, V., Munro, S., 2008. Golgi coiled-coil proteins contain multiple binding sites for Rab family G proteins. *Journal of Cell Biology* 183, 607–615.
- Sollner, T., Whiteheart, S.W., Brunner, M., *et al.*, 1993. SNAP receptors implicated in vesicle targeting and fusion. *Nature* 362, 318–324.
- Sonnichsen, B., De Renzis, S., Nielsen, E., Rietdorf, J., Zerial, M., 2000. Distinct membrane domains on endosomes in the recycling pathway visualized by multicolor imaging of Rab4, Rab5, and Rab11. *Journal of Cell Biology* 149, 901–914.
- Sudhof, T.C., Rothman, J.E., 2009. Membrane fusion: Grappling with SNARE and SM proteins. *Science* 323, 474–477.
- Trucco, A., Polishchuk, R.S., Martella, O., *et al.*, 2004. Secretory traffic triggers the formation of tubular continuities across Golgi sub-compartments. *Nature Cell Biology* 6, 1071–1081.
- Tu, L., Tai, W.C., Chen, L., Banfield, D.K., 2008. Signal-mediated dynamic retention of glycosyltransferases in the Golgi. *Science* 321, 404–407.
- Volchuk, A., Amherdt, M., Ravazzola, M., *et al.*, 2000. Megavesicles implicated in the rapid transport of intracisternal aggregates across the Golgi stack. *Cell* 102, 335–348.
- Waters, M.G., Serafini, T., Rothman, J.E., 1991. 'Coatomer': A cytosolic protein complex containing subunits of non-clathrin-coated Golgi transport vesicles. *Nature* 349, 248–251.
- Watson, P.J., Frigerio, G., Collins, B.M., Duden, R., Owen, D.J., 2004. Gamma-COP appendage domain — Structure and function. *Traffic* 5, 79–88.
- Wegmann, D., Hess, P., Baier, C., Wieland, F.T., Reinhard, C., 2004. Novel isotopic subunits reveal three coatomer complexes in mammals. *Molecular and Cellular Biology* 24, 1070–1080.
- Weidman, P., Roth, R., Heuser, J., 1993. Golgi membrane dynamics imaged by freeze-etch electron microscopy: Views of different membrane coatings involved in tubulation versus vesiculation. *Cell* 75, 123–133.
- Yang, J.S., Valente, C., Polishchuk, R.S., *et al.*, 2011. COPI acts in both vesicular and tubular transport. *Nature Cell Biology* 13, 996–1003.

Post-Golgi Transport – Cargo, Carriers, and Pathways

JL Stow, The University of Queensland, Brisbane, QLD, Australia

RZ Murray, Queensland University of Technology, Brisbane, QLD, Australia

© 2016 Elsevier Inc. All rights reserved.

Introduction

The bulk of newly synthesized proteins transported through the secretory pathway, having traversed the endoplasmic reticulum (ER) and Golgi complex, will arrive at the trans-Golgi network (TGN) where they will be directed into one or more post-Golgi pathways, *en route* to their final destinations. Upon leaving the neat confines of the Golgi complex, this protein cargo will find itself in the labyrinthine world of post-Golgi transport pathways, where carriers of different guises move cargo along criss-crossing routes to endosomal hubs, to other organelles, or to the various domains of the plasma membrane for residence or release.

The cargo of newly synthesized proteins moving through the ER, COPII vesicles, and the Golgi cisternae is concentrated and moving mostly in the same direction, making it possible to label and experimentally track cargo molecules through the early secretory pathway in cells (Farquhar and Palade, 1998). Some cells then have large distinctive secretory granules for the later stages of transport where the cargo is also concentrated. However, the bulk of transport beyond the TGN occurs in many fast-moving, transient carriers where the cargo is dilute and in small volumes. Thus the carriers and cargo of these constitutive transport pathways have traditionally been much more difficult to trace and study, especially since these pathways are often varied and customized in different cell types.

With the advent of GFP-tagged proteins and the ability to image trafficking in live cells and model organisms (Lippincott-Schwartz *et al.*, 2000), together with genomics and proteomics, there is growing understanding of the TGN and of post-Golgi transport. Increasingly we learn from the defective trafficking that results from gene mutations – experimentally introduced mutations and those in model organisms and in human disease (De Matteis and Luini, 2011). The molecular machinery of the TGN and its carriers, at the ‘business end’ of the secretory pathway – close to the functional sites of proteins, can offer unique opportunities for experimental or clinical modulation of protein transport and function.

Overview of Pathways

TGN-Derived Constitutive Transport

The TGN is an active transport hub positioned at the trans-side of the Golgi stack. Ultrastructurally it is distinguished as a large pleiomorphic reticulum with multiple budding profiles of clathrin- and non-clathrin-coated vesicles (Farquhar and Palade, 1998). The morphology of vesicles, coats, and relationship between the TGN and Golgi cisternae have been further revealed through elegant high voltage electron tomographic studies (Mogelsvang *et al.*, 2004; Storrie *et al.*, 2012). Imaging in living (unfixed) cells has revealed that the TGN

also gives rise to many dynamic tubular extensions coursing along microtubules (De Matteis and Luini, 2008; Goud and Gleeson, 2010). The TGN is positioned close to the microtubule-organizing center and it has associated actin arrays. From early on (Farquhar and Palade, 1998), close connections between the TGN and ER were noted, which may serve in lipid transfer (Ngo and Ridgway, 2009).

In all eukaryotic cells there is constitutive (continuous) transport of newly synthesized and sometimes recycling proteins from the TGN to the cell surface and multiple other destinations. Cargo transport through these pathways typically involves the continuous movement of small amounts of cargo in carriers that form through budding at the TGN and then fuse with their target membrane upon cargo delivery. Multiple types of carriers and transport routes are needed for constitutive transport (Figure 1). These pathways service major cell functions such as secretion, receptor recycling and signaling, cell adhesion, cell movement, and cell proliferation. The post-Golgi pathways feed into endosomal hubs, autophagic organelles, and also serve for retrograde transport. The variety of cargo and the volume of constitutive transport performed in different cell types can vary enormously. Minimal transport is needed, for instance, at a maintenance level in mature parenchymal cells, which may be just replenishing their own membranes and organelles. At the other extreme, protein synthesis and constitutive trafficking in mature plasma cells can be almost solely dedicated to the secretion of large amounts of immunoglobulin. Indeed, for this reason, this was one of the first secretory pathways where the role of the Golgi complex was first revealed (Zagury *et al.*, 1970).

In some endo- and exocrine cell types there are additional pathways for the regulated secretion of special cell products. These pathways differ from constitutive transport in several ways, namely that cargo leaving the TGN is packaged and stored in large granules or lysosome-related organelles (LROs) for release in response to specific stimuli. Regulated secretion is described elsewhere. The details of constitutive trafficking are worthy of further consideration here.

Modulating Constitutive Transport

The level of constitutive transport can be ‘dialed up’ to cope with periods of high demand for secretion or for stimulus-coupled translocation of proteins to the cell surface. This is exemplified in macrophages, which, upon activation by pathogen contact, synthesize and secrete inflammatory cytokines. To accommodate this sudden need for cytokine release, the expression of trafficking machinery is increased and the number of constitutive carriers budding off the TGN is enhanced (Lieu *et al.*, 2008; Stow and Murray, 2013). The levels of secretion of inflammatory mediators in macrophages and epithelial cells can also be modulated by siphoning off cargo from the Golgi into autophagy-related organelles (Narita *et al.*,

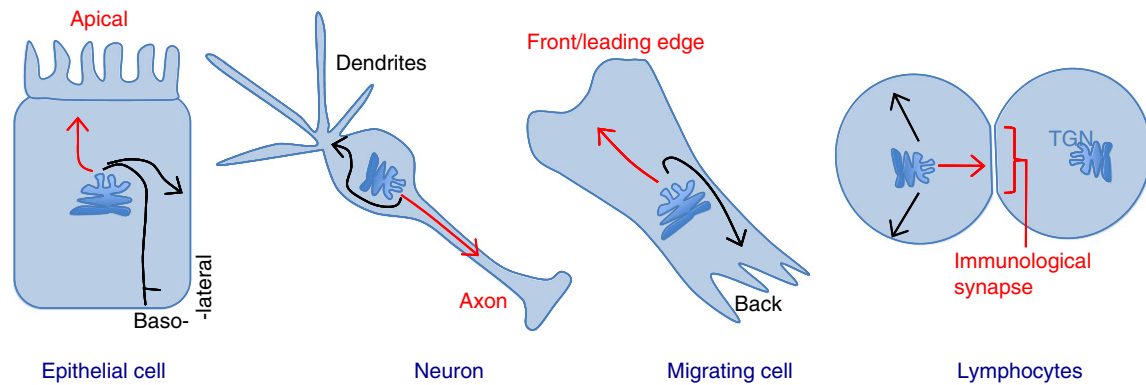


Figure 1 Transport in polarized cells. Polarity is a permanent feature of differentiated epithelial cells and neurons; cells undergoing migration or lymphocytes forming an immunological synapse with a target cell have contextual polarity. In all cases the TGN has the added complication of having to direct different cargo and carriers to the basolateral pole or the apical pole of epithelial cells, the dendrites or axon of neurons, the front or back of migrating cells, and the synapse or rest of cell surface in lymphocytes. Sorting, carriers, and routes for polarized transport are described herein.

2011). Another example of up-regulated trafficking is in adipocytes, which, upon activation of the insulin receptor, must translocate Glut4 glucose transporters to the cell surface from a 'holding pattern' in special post-Golgi vesicles (Stockli *et al.*, 2011). These and other situations demonstrate that constitutive trafficking in post-Golgi pathways can be modulated and tailored to the needs of the cell.

Polarized Transport

Transport from the TGN to the cell surface instantly takes on a new level of complexity in cells where distinct cargo has to be routed in different directions for delivery to separate cell domains (Mellman and Nelson, 2008; Figure 1). Mature epithelial cells and neurons are classical examples of differentiated, polarized cell types with morphologically and functionally different poles. Cell polarity can also be temporary or contextual, for instance, in migrating cells which develop a leading edge for forward movement (Miller *et al.*, 2009; Veale *et al.*, 2010), and in lymphocytes that come into contact with a target cell at specialized contact zones or immunological synapses (Angus and Griffiths, 2013). In all of these cases, the cells must have the ability to direct a subset of cargo toward a particular 'end' or 'side' of the cell (Figure 1).

Both the TGN and endosomes are sorting stations for polarized trafficking and at these sites a number of mechanisms are used to sort cargo into carriers destined for different polarized domains (Apodaca *et al.*, 2012; Weisz and Rodriguez-Boulan, 2009). As described below, lipid-based sorting and carriers often service apical transport in epithelia, while protein-based sorting, including sorting signals, specific adapters, and both clathrin- and non-clathrin-mediated carriers service basolateral transport.

The cytoskeleton – both microtubule and actin networks – plays ubiquitous roles in TGN function and in all pathways of post-Golgi transport (Goud and Gleeson, 2010). However, microtubules have very demonstrable roles in directing carriers to specific surface domains in polarized cells (Musch, 2004). In both permanent and contextual cell polarity, microtubules and microtubule motors offer very directed transport to

specific membrane domains. This can be seen in cytotoxic T cells where, upon forming an immunological synapse with a target cell, the Golgi complex and TGN rotate to face this contact zone and interconnecting microtubules directly carry secretory vesicles and LROs to this highly specific polarized site for cell–cell killing (Angus and Griffiths, 2013). Similarly, cilia emerging from the apical surface of epithelial cells have a special post-Golgi transport loop directed by microtubules to shuttle cargo to the base of the cilia for their formation and function (Deretic, 2013).

Constitutive Transport via Endosomal Hubs

While constitutive traffic was originally conceived as solely direct transport from the TGN to the cell surface, it is now recognized that many post-Golgi pathways intersect with endosomes, especially recycling endosomes – as intervening hubs on the way to the cell surface or other destinations (Hsu and Prekeris, 2010; Figure 2). Recycling endosomes have a rich panoply of trafficking GTPases and their effectors which include scaffolds, motors, and kinases, for trafficking through this compartment. While there are multiple recycling endosome-associated Rabs, the Rab11 subfamily (11a, 11b, and 25) and Rabs8 (8a and 8b) have been widely implicated in traffic in and out of recycling endosomes (Apodaca *et al.*, 2012; Baetz and Goldenring, 2013; Hsu and Prekeris, 2010).

Recycling endosomes themselves can form a large, dynamic tubular reticulum wherein recycling surface receptors and other proteins converge with newly synthesized exocytic cargo arriving from the TGN. Proteins can also be directed in retrograde fashion from endosomal hubs back to the TGN. Hubs may be required in exocytic pathways to perform additional sorting or rearrangement of cargo and to manage the timing and direction of final cargo deployment. For instance, the recycling endosome is a key hub for further cargo sorting for polarized transport in epithelial cells; basolaterally directed proteins first transit the recycling endosome after leaving the TGN (Ang *et al.*, 2004; Lock and Stow, 2005) and in 3-dimensional epithelial cysts, recycling endosomes and their

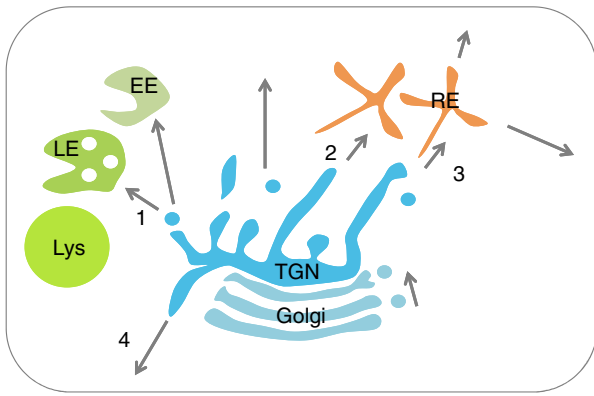


Figure 2 Routes, cargo, and carriers. The TGN gives rise to transport routes (arrows) that connect to different organelles and the cell surface. Examples 1–4 illustrate the types of cargo and TGN-derived carriers moving along some of these routes: (1) AP-1/GGA and clathrin-coated vesicles transport lysosomal hydrolases to late endosomes (Rouille *et al.*, 2000). (2) Rab6 and myosinII-associated tubular carriers undergo fission from the TGN to transport the experimental cargo protein VSV-G from TGN to recycling endosomes (Miserey-Lenkei *et al.*, 2010). (3) Tubular carriers decorated with golgin-245, syntaxin6, and Vti1b form at the TGN to transport the cytokine TNF to recycling endosomes which then carry the cytokine to the surface for release (Manderson *et al.*, 2007; Murray *et al.*, 2005). (4) The so-called ‘CART’ carriers transport selective cargo including pancreatic adenocarcinoma up-regulated factor, but not VSVG or collagen, from TGN to cell surface. CARTs use PKD for fission and are decorated with Rab6, Rab8, and myosin II (Wakana *et al.*, 2012).

machinery have primary roles in establishing luminal and basolateral cell domains (Apodaca *et al.*, 2012). In macrophages, characterization of the soluble NSF attachment protein receptor (SNARE) machinery showed that the transmembrane precursor of the cytokine TNF is trafficked from the TGN to recycling endosomes which perform the final step of delivering TNF to a polarized site on the cell surface at the outer edge of phagosomes (Murray *et al.*, 2005). Furthermore, in a variation of sorting roles, TNF may be retained at the recycling endosome by AP-1 and clathrin complexes and then trafficked in carriers with AP-1 alone to reach the cell surface (Braun *et al.*, 2007). Thus the TGN can work intimately with other organelles for multistage transport.

Cargo Sorting and Packaging at the TGN

A complex array of cargo, including soluble and transmembrane proteins, arrive at the TGN for sorting and packaging into different carriers. Multiple mechanisms are used for cargo sorting in this location, including protein-based sorting for proteins that extend into or are connected to the cytoplasm, or alternatively there are lipid-based sorting mechanisms. Protein clustering with special packaging proteins is also utilized in some cells to direct cargo into granules. The extension of TGN-derived tubules also facilitates the spatial segregation of cargo for sorting. It is clear that cargo – both proteins and lipids – along with membrane subdomains, can be separated with

high fidelity for packaging into a variety of different carriers for transport to different destinations (Figure 2).

Lipid Rafts in Sorting and Transport

Membrane lipids are also important determinants of membrane trafficking in post-Golgi pathways. The concept of glycosphingolipid and protein-rich domains, the so-called ‘lipid rafts,’ as islands in the fluid mosaic environment of biological membranes, is now embedded as a fundamental principle for clustering and organizing groups of proteins for function (e.g., receptor-mediated signaling) and for transport (Lingwood and Simons, 2010). Lipid rafts are recognized as dynamic, mobile, domains of varying sizes with a high intrinsic concentration of cholesterol and sphingolipids, which attract lipophilic proteins such as glycosylphosphatidyl-inositol (GPI)-anchored proteins, along with receptors, and molecules for signaling and cytoskeletal attachment (Lingwood and Simons, 2010; Surma *et al.*, 2012). At the TGN, lipid rafts form with the purpose of clustering specific cargo, such as GPI-anchored proteins, which cannot access protein-based sorting mechanisms on the cytoplasmic leaflet of membranes. These GPI-anchored proteins are delivered to specific destinations, most notably the apical surface of polarized cells or the bud plasma membrane of the yeast, *Saccharomyces cerevisiae*. The apical membrane of polarized epithelial cells is relatively enriched in glycosphingolipids and cholesterol compared to the basolateral membrane. Lipid rafts, from early on, were thus recognized for their role in delivering GPI-anchored proteins, some viral proteins and other cargo to the apical membrane, either through direct transport routes, or via endosomes (reviewed in Weisz and Rodriguez-Boulton, 2009). Lipid rafts not only help sort cargo but also actively help to form the carriers for such transport. Isolating populations of post-Golgi vesicles in yeast with different profiles of cholesterol or sphingolipids helped to establish that the TGN can actually sort membrane lipids to form different carriers (Klemm *et al.*, 2009). Budding of such carriers involves the coalescence of individual rafts into larger patches or domains, which create biophysical forces, including line tension (Baumgart *et al.*, 2003), which favor membrane curvature and budding of lipid raft-rich vesicles or tubules as cargo carriers (Julicher and Lipowsky, 1993; Roux *et al.*, 2005; van Meer and Vaz, 2005). The influenza proteins hemagglutinin and neuraminidase are both raft-associated proteins and the apical transport of both proteins is accelerated when they are coexpressed due to clustering of disparate rafts (Ohkura *et al.*, 2014).

Lipid rafts participate not only in sorting cargo for transport, but also in clustering the protein machinery that is important for trafficking and secretion, including elements such as SNAREs, Rabs, adapters, and cytoskeletal linkers (Weisz and Rodriguez-Boulton, 2009). The adaptor annexin II is commonly found in rafts and is required for apical transport in polarized epithelial cells (Jacob *et al.*, 2004). The secretion of the cytokine TNF in macrophages (Kay *et al.*, 2006) and of mast cell mediators (Puri and Roche, 2006) depend on lipid rafts to orchestrate the plasma membrane SNAREs, syntaxin 4, and SNAP23, for membrane fusion at the cell surface. Pathogens make ready use of lipid rafts for trafficking during cell

entry and exit (Manes *et al.*, 2003), including targeting some of the raft-associated trafficking machinery with pathogen-derived effectors. Thus membrane lipids are intimately involved in sorting, and in the differential packaging and movement of cargo in bidirectional transport between the TGN and cell surface.

Protein-Based Sorting in the TGN

Multiple mechanisms and a vast array of cytoplasmic proteins are used to sort cargo based on contact with the cytoplasmic tails of transmembrane proteins. These sorting mechanisms package cargo into a variety of vesicular and tubular carriers, although how specificity is achieved is not fully understood for all types of intrinsic cargo, given that many studies to date have used model systems or exogenous cargo. Various aspects of protein-based sorting are considered as follows.

Sorting signals

One common principle for sorting transmembrane proteins is through sorting signals encoded in the cytoplasmic tails of cargo proteins. These amino acid motifs are recognized by cytoplasmic adaptors and coat proteins to sort this cargo into specific carriers. At the TGN, tyrosine- or dileucine-based sorting signals are known for trafficking proteins to endosomes or to the basolateral domain of epithelial cells (Rodríguez-Boulán *et al.*, 2005). Protein phosphorylation and ubiquitination can also direct trafficking, and luminal O-glycosylation can act as an apical sorting signal (Scott *et al.*, 2004; Yeaman *et al.*, 1997).

Phosphatidylinositol 4-phosphate and ADP-ribosylation factors

The TGN membrane is enriched in phosphatidylinositol 4-phosphate (PI4P), which forms the basis for recruiting some of the budding machinery (as PI4P effectors) and for deforming the membranes for budding itself (De Matteis and Luini, 2008). The PI4-kinases (Pik1p in yeast or PI4KIII β in mammals) are also located on the TGN to produce PI4P for trafficking, and loss or mutation of PI4-kinases on the TGN blocks secretion of PI4P effectors usually bound to PI4P, in addition to binding one or more of the GTPases on the TGN membrane (reviewed in Santiago-Tirado and Bretscher, 2011).

One of the first proteins recruited to PI4P is the small GTPase ADP-ribosylation factor 1 (ARF1), activated on the TGN by one of its TGN-resident ARF-GEFs, Sec7 in yeast or BIG1/2 in mammals (Casanova, 2007; Jones *et al.*, 2005; Lowery *et al.*, 2013). Parenthetically, the ARF-GEFs are recruited with help from ARF and ARL (another GTPase) on the TGN (Christis and Munro, 2012; Richardson *et al.*, 2012). Thus ARFs on the TGN work in one of the striking examples of GTPase cascades that operate on several membranes throughout the cell (Mizuno-Yamasaki *et al.*, 2012). The coupling of cargo with PI4P and ARF1 forms the 'first layer' of sorting and packaging, underpinning the recruitment of adaptor complexes for multiple types of carriers assembling on the TGN.

Adaptors and clathrin-coats on the TGN

PI4P and ARF1 complexes on the TGN recruit specific sets of adaptors, the GGAs and AP-1 together with clathrin for formation of clathrin-coated vesicles (CCVs) around clustered sorting signal-selected cargo for transport to endosomes or other destinations. In yeast the Gg2 protein also recruits the PI4P kinase Pik1p to produce more local PI4P, in a positive feedback loop that is needed to support secretion (Daboussi *et al.*, 2012). The clathrin adaptor AP-1 on the TGN emulates the formation of CCV supported by AP-2 at the plasma membrane, and AP-4 is also located on the TGN although it has not yet been shown to form CCVs (Robinson, 2004). It is not clear in all cells whether GGAs and AP-1 assemble on separate or overlapping CCVs, nor ultimately how many different types of CCVs are generated.

One of the best known roles for CCVs at the TGN is in the transport of lysosomal contents to one or more stages of the endolysosomal pathway (Saftig and Klumperman, 2009). Classically, lysosomal enzymes tagged with mannose-6-phosphate (M6P) are bound by M6P receptors in the TGN, loaded into CCV, and delivered to late endosomes *en route* to lysosomes (Ghosh *et al.*, 2003). Transport of lysosomal enzymes through endosomal compartments can invoke both AP-1 and AP-3 mediated steps (Peden *et al.*, 2004). Interestingly, other components of lysosomes can be carried via different endosomal, or even direct trajectories, from the TGN to lysosomes in CCV or in non-clathrin-mediated carriers (Saftig and Klumperman, 2009). Thus, sphingosin activator protein is carried by the multipurpose receptor sortilin into CCV for transport directly from the TGN to lysosomes, bypassing M6P receptors (Lefrançois *et al.*, 2003). The enzyme, β -glucocerebrosidase, also bypasses M6P receptors, instead being carried to lysosomes with the lysosomal integral protein, LIMP-2 (Reczek *et al.*, 2007). The dire consequences of disrupting these transport steps through gene mutation or loss-of-function are evident in resulting lysosomal storage diseases (Parkinson-Lawrence *et al.*, 2010; Saftig and Klumperman, 2009).

The clathrin adaptors also participate directly in sorting for polarized transport. This was highlighted after AP-1B was discovered as an epithelial specific AP-1 isoform required for basolateral targeting (Traub and Apodaca, 2003). Genetic ablation of an AP-1B subunit in mice results in intestinal epithelial hyperplasia and loss of cell polarity (Hase *et al.*, 2013). It is now evident that AP-1B normally acts at the level of recycling endosomes, rather than in the TGN, for basolateral sorting (Ang *et al.*, 2004), but it works in concert with the ubiquitous AP-1A at the TGN for fully efficient basolateral transport (Gravotta *et al.*, 2012).

Ca²⁺-dependent sorting and transport at the TGN

The cytoplasmic actin-modeling proteins ADF/cofilin or their homologs in flies, are recruited to the TGN where they are required for sorting a subset of cargo in concert with actin and the Ca²⁺ transporter, SPCA1 (von Blume *et al.*, 2011). SPCA1 regulates Ca²⁺ influx into the TGN. The cargo sorted in this manner include specific subsets of Ca²⁺ binding and non-binding proteins. The sorting mechanism proposed involves the severing of actin filaments to release or activate SPCA1 for Ca²⁺ pumping to create local ionic environments for sorting cargo in the TGN lumen.

Tubules and Carriers – Assembly and Fission

Ultrastructural studies have shown an abundance of CCVs, COP-coated vesicles, and uncoated vesicles associated with the TGN and as budding profiles on these membranes (Farquhar and Palade, 1998; Mogelsvang *et al.*, 2004; Storrie *et al.*, 2012). In addition to vesicles, long tubular carriers have also been evident since live imaging of fluorescently tagged components has been possible (Lippincott-Schwartz *et al.*, 2000). Tubules can be seen emerging from the TGN, undergoing fission and microtubule-directed transport, and they can be selectively loaded with soluble or membrane-bound cargo (De Matteis and Luini, 2008; Goud and Gleeson, 2010; Manderson *et al.*, 2007; Miserey-Lenkei *et al.*, 2010). How many types of tubular carriers are there is not known, nor yet is the full complement of proteins engaged in sorting cargo into them, or controlling their fission and transport.

Golgins

Coiled-coil membrane tethers or scaffolds abound across the Golgi stack and TGN. On the TGN, the GRIP-golgin family members golgin-245, golgin-97, GCC185, and GCC88 act as multifunctional scaffolds for carrier formation and retrieval and for tethering organelles and protein complexes for trafficking (Goud and Gleeson, 2010). GRIP-golgins are recruited and bound to the TGN membrane via their GRIP domain and often, but not always, by interaction with members of the Arf-like ARL and the Rab families of GTPases (Goud and Gleeson, 2010). Notably, individual golgins usually occupy different subdomains of the TGN and they selectively participate in trafficking of different cargo. GTPase ARF-related protein 1 (ARFRP1) regulates the recruitment of ARL1 and golgin-245 to the TGN. Liver-specific knockout of ARFRP1 in mice interferes with the secretion of IGF1 but not IGFBP2 from hepatocytes and also disrupts sorting of Glut2 (Hesse *et al.*, 2012), indicative of a role for golgin-245 in the transport of only a subset of hepatocyte cargo. In macrophages, golgin-245 transports TNF out of the TGN for secretion (Lieu *et al.*, 2008). Golgin-97 transports E-cadherin out of the TGN in epithelial cells (Lock and Stow, 2005). Working in the other transport direction, golgin-97 on the TGN binds to the Rab11 effector FIP1/RCP to regulate incoming traffic from endosomes (Jing *et al.*, 2010). Golgins, in general, bind to multiple families of small GTPases, particularly Rabs, to cytoskeletal proteins including linkers for both actin and microtubule networks and to SNAREs (Goud and Gleeson, 2010; Munro, 2011). This has led to one view of golgins forming a net around the TGN for capturing trafficking machinery (Munro, 2011), and certainly of being key players in post-Golgi transport.

Modifying lipids for tubulation

A variety of lipid modifying enzymes and lipid transfer proteins have been implicated in tubule formation and fission at the TGN, together with protein assemblies. FAPP2 is an adaptor that interacts preferentially with PI4P via its PH domain and with ARF1 on tubular domains of the TGN, it is found only at the tubule junction and not after fission of VSV-G-containing carriers (Godi *et al.*, 2004). Lipid transfer proteins such as CERT and OSBP1 are similarly recruited, potentially regulating the nature of the lipid microenvironment for tubulation

(De Matteis and Luini, 2008). Phospholipase A2 (PLA2) causes tubulation of the TGN membranes by regulation acetylation of phospholipids (Ha *et al.*, 2012). Various PI3Ks, including PI3K δ , VPS-34, and PI3KC2 α , have been implicated in budding of constitutive carriers, CCV and neurosecretory granules from the TGN, anticipating the additional involvement of 3-phosphoisositides at the TGN for membrane fission (Domin *et al.*, 2000; Jones and Howell, 1997; Low *et al.*, 2010).

Membrane fission

Protein kinase D (PKD) has become well-known as a membrane fission agent at the TGN to generate vesicular or pleomorphic carriers, following the implication of this kinase as the target of a marine inhibitor ilimaquinone which vesiculates the entire Golgi (Takizawa *et al.*, 1993). The trimeric G proteins, G $\beta\gamma$ induces PLC-driven diacylglycerol (DAG) production at the TGN and this in turn recruits PKD to activate the protein kinase, PKC η , which phosphorylates and activates PKD (reviewed in Campelo and Malhotra, 2012).

PKD also binds to ARF1 and interacts with a key set of GTPase effectors, the Bin/amphiphysin/Rvs (BAR)-domain containing family of ARFaptins to regulate fission (Gehart *et al.*, 2012). The banana-shaped BAR-domains of several families of BAR-domain containing proteins on the TGN, sense membrane curvature and regulate fission (McMahon and Gallop, 2005). BAR-domain-containing proteins work together with the GTPase dynamin for membrane fission (Meinecke *et al.*, 2013). The ubiquitous dynamin2 isoform performs fission of carriers at the TGN, for a selective subset of apical carriers in polarized transport (Jones *et al.*, 1998; Kreitzer *et al.*, 2000). Other fission proteins like CtBP3/BARS, drive fission in dynamin-independent pathways (Bonazzi *et al.*, 2005). Following fission, loaded carriers are at the behest of microtubules and motors such as KIFs for their transport and delivery.

Actin, myosins, and tubules

Short actin filaments surround the Golgi complex (Percival *et al.*, 2004) and actin, paired with various myosin motors, is intimately involved in the deformation of membranes, carrier fission, and movement. GOLPH3 bridges the PI4P-rich TGN membrane and actin filaments via the motor Myo18a to exert tensile forces for pulling membranes into buds and tubules (Dippold *et al.*, 2009). MyosinIIA is found associated with several types of post-Golgi carriers emerging from the TGN. MyosinII is an effector of Rab6 in which guise it regulates post-Golgi transport and maintenance of Golgi cisternae and stack morphology (Miserey-Lenkei *et al.*, 2010; Storrie *et al.*, 2012). MyosinII is found on subpopulations of tubules during budding at the TGN and on some carriers as they move into the cell periphery (Ikonen *et al.*, 1997; Wakana *et al.*, 2012). Myosin1b working with actin foci, is also implicated in the formation of tubules for post-Golgi transport; disrupting myosin1b induces shape changes in tubules and the TGN (Almeida *et al.*, 2011). These and many other examples give testament to the close links between actin, myosins, and membranes at the TGN.

Rabs and SNAREs

Of the 60 or more mammalian Rab family members, about one-quarter of them are associated with the Golgi or its

emergent transport pathways. These include Rabs 6, 33B, 1, 2, 43, 18, 32, 13, 27, 30, 37, 26, 22, and 3 and relevant yeast family members, Ypt1, 6 and 32 (Goud and Gleeson, 2010; Liu and Storie, 2012). The Rab6 splice variants Ra6A and 6A' are the best known of the TGN-associated Rabs. In addition to roles exerted through myosinII (above), Rab6 has roles in trafficking cargo out of the TGN, for instance in complex with Rab8 and the scaffold MICAL3 (Grigoriev *et al.*, 2007; Grigoriev *et al.*, 2011). Rab 6A and 6A' also regulate TNF trafficking out of the TGN in macrophages (Micaroni *et al.*, 2013). Many of the other Rabs associated with the TGN are implicated through roles on exiting or incoming carriers.

Various SNAREs and fusogenic complexes have been found on the Golgi complex, the best known of which are involved in transport between the ER and early Golgi in mammalian cells and yeast (Malsam and Sollner, 2011). SNAREs must be loaded into carriers for transport to the fusion site. During budding the SNAREs can assemble with cargo, Rabs and Rab-GEFs via tethering proteins, such as GARP at the TGN (Perez-Victoria and Bonifacio, 2009). Two SNARE complexes identified as residents on the TGN for fusion of incoming transport from different endosomes are syntaxin16/Vti1a/syntaxin 6/VAMP4 together with Rab6 and the complex syntaxin16/Vti1a/syntaxin10/VAMP3 (Ganley, 2008; Mallard *et al.*, 2002).

Concluding Remarks

The number of molecules required for sorting and transport at the TGN grows ever larger, unsurprisingly, given the vast and complex postal service run from this organelle. Common principles governing the sorting and transport of proteins and lipids continue to emerge from studies in model organisms and human cells and from insights gained from human disease. Future studies will benefit from the power of genomics and bioinformatics to reveal complete gene networks and from new innovations in microscopy and imaging to visualize these fascinating cell processes.

Acknowledgments

The authors are supported by research and fellowship funding from the National Health and Medical Research Council of Australia.

See also: Interorganellar Communication: Interplay and Processes: Regulated versus Constitutive Secretion – A Major Form of Intercellular Communication; Retrograde Transport. Organelles: Structure and Function: Golgi and TGN; Lysosome-Related Organelles

References

- Almeida, C.G., Yamada, A., Tenza, D., *et al.*, 2011. Myosin 1b promotes the formation of post-Golgi carriers by regulating actin assembly and membrane remodelling at the trans-Golgi network. *Nature Cell Biology* 13, 779–789.
- Ang, A.L., Taguchi, T., Francis, S., *et al.*, 2004. Recycling endosomes can serve as intermediates during transport from the Golgi to the plasma membrane of MDCK cells. *Journal of Cell Biology* 167, 531–543.
- Angus, K.L., Griffiths, G.M., 2013. Cell polarisation and the immunological synapse. *Current Opinion in Cell biology* 25, 85–91.
- Apodaca, G., Gallo, L.I., Bryant, D.M., 2012. Role of membrane traffic in the generation of epithelial cell asymmetry. *Nature Cell Biology* 14, 1235–1243.
- Baetz, N.W., Goldenring, J.R., 2013. Rab11-family interacting proteins define spatially and temporally distinct regions within the dynamic Rab11a-dependent recycling system. *Molecular Biology of the Cell* 24, 643–658.
- Baumgart, T., Hess, S.T., Webb, W.W., 2003. Imaging coexisting fluid domains in biomembrane models coupling curvature and line tension. *Nature* 425, 821–824.
- von Blume, J., Alleaume, A.M., Cantero-Recasens, G., *et al.*, 2011. ADF/cofilin regulates secretory cargo sorting at the TGN via the Ca²⁺ ATPase SPCA1. *Developmental Cell* 20, 652–662.
- Bonazzi, M., Spano, S., Turacchio, G., *et al.*, 2005. ClB3/BARS drives membrane fission in dynamin-independent transport pathways. *Nature Cell Biology* 7, 570–580.
- Braun, V., Deschamps, C., Raposo, G., *et al.*, 2007. AP-1 and ARF1 control endosomal dynamics at sites of FcR mediated phagocytosis. *Molecular Biology of the Cell* 18, 4921–4931.
- Campelo, F., Malhotra, V., 2012. Membrane fission: The biogenesis of transport carriers. *Annual Review of Biochemistry* 81, 407–427.
- Casanova, J.E., 2007. Regulation of Arf activation: The Sec7 family of guanine nucleotide exchange factors. *Traffic* 8, 1476–1485.
- Christis, C., Munro, S., 2012. The small G protein Arl1 directs the trans-Golgi-specific targeting of the Arf1 exchange factors BIG1 and BIG2. *Journal of Cell Biology* 196, 327–335.
- Daboussi, L., Costaguta, G., Payne, G.S., 2012. Phosphoinositide-mediated clathrin adaptor progression at the trans-Golgi network. *Nature Cell Biology* 14, 239–248.
- De Matteis, M.A., Luini, A., 2008. Exiting the Golgi complex. *Nature Reviews. Molecular Cell Biology* 9, 273–284.
- De Matteis, M.A., Luini, A., 2011. Mendelian disorders of membrane trafficking. *New England Journal of Medicine* 365, 927–938.
- Deretic, D., 2013. Crosstalk of Arf and Rab GTPases en route to cilia. *Small GTPases* 4, 70–77.
- Dippold, H.C., Ng, M.M., Farber-Katz, S.E., *et al.*, 2009. GOLPH3 bridges phosphatidylinositol-4-phosphate and actomyosin to stretch and shape the Golgi to promote budding. *Cell* 139, 337–351.
- Domin, J., Gaidarov, I., Smith, M.E., Keen, J.H., Waterfield, M.D., 2000. The class II phosphoinositide 3-kinase PI3K-C2alpha is concentrated in the trans-Golgi network and present in clathrin-coated vesicles. *Journal of Biological Chemistry* 275, 11943–11950.
- Farquhar, M.G., Palade, G.E., 1998. The Golgi apparatus: 100 years of progress and controversy. *Trends in Cell Biology* 8, 2–10.
- Ganley, I.G., Espinosa, E., Pfeffer, S.R., 2008. A syntaxin 10-SNARE complex distinguishes two distinct transport routes from endosomes to the trans-Golgi in human cells. *Journal of Cell Biology* 180, 159–172.
- Gehart, H., Goginashvili, A., Beck, R., *et al.*, 2012. The BAR domain protein Arlaptin-1 controls secretory granule biogenesis at the trans-Golgi network. *Developmental Cell* 23, 756–768.
- Ghosh, P., Dahms, N.M., Kornfeld, S., 2003. Mannose 6-phosphate receptors: New twists in the tale. *Nature Reviews. Molecular Cell Biology* 4, 202–212.
- Godi, A., Di Campli, A., Konstantakopoulos, A., *et al.*, 2004. FAPPs control Golgi-to-cell-surface membrane traffic by binding to ARF and PtdIns(4)P. *Nature Cell Biology* 6, 393–404.
- Goud, B., Gleeson, P.A., 2010. TGN golgins, Rabs and cytoskeleton: Regulating the Golgi trafficking highways. *Trends in Cell Biology* 20, 329–336.
- Gravotta, D., Carvajal-Gonzalez, J.M., Mattera, R., *et al.*, 2012. The clathrin adaptor AP-1A mediates basolateral polarity. *Developmental Cell* 22, 811–823.
- Grigoriev, I., Splinter, D., Keizer, N., *et al.*, 2007. Rab6 regulates transport and targeting of exocytotic carriers. *Developmental Cell* 13, 305–314.
- Grigoriev, I., Yu, K.L., Martinez-Sanchez, E., *et al.*, 2011. Rab6, Rab8, and MICAL3 cooperate in controlling docking and fusion of exocytotic carriers. *Current Biology* 21, 967–974.
- Ha, D.K., Clarke, B.A., Brown, W.J., 2012. Regulation of the Golgi complex by phospholipid remodelling enzymes. *Biochimica et Biophysica Acta (BBA) – Molecular and Cell Biology of Lipids* 1821, 1078–1088.
- Hase, K., Nakatsu, F., Ohmae, M., *et al.*, 2013. AP-1B-mediated protein sorting regulates polarity and proliferation of intestinal epithelial cells in mice. *Gastroenterology* 145, 625–635.

- Hesse, D., Jaschke, A., Kanzleiter, T., *et al.*, 2012. GTPase ARFRP1 is essential for normal hepatic glycogen storage and insulin-like growth factor 1 secretion. *Molecular and Cellular Biology* 32, 4363–4374.
- Hsu, V.W., Prekeris, R., 2010. Transport at the recycling endosome. *Current Opinion in Cell Biology* 22, 528–534.
- Ikonen, E., de Almeida, J.B., Fath, K.R., *et al.*, 1997. Myosin II is associated with Golgi membranes: Identification of p200 as nonmuscle myosin II on Golgi-derived vesicles. *Journal of Cell Science* 110 (Pt 18), 2155–2164.
- Jacob, R., Heine, M., Eikemeyer, J., *et al.*, 2004. Annexin II is required for apical transport in polarized epithelial cells. *Journal of Biological Chemistry* 279, 3680–3684.
- Jing, J., Junutula, J.R., Wu, C., *et al.*, 2010. FIP1/RCP binding to Golgin-97 regulates retrograde transport from recycling endosomes to the trans-Golgi network. *Molecular Biology of the Cell* 21, 3041–3053.
- Jones, H.D., Moss, J., Vaughan, M., 2005. BIG1 and BIG2, brefeldin A-inhibited guanine nucleotide-exchange factors for ADP-ribosylation factors. *Methods in Enzymology* 404, 174–184.
- Jones, S.M., Howell, K.E., 1997. Phosphatidylinositol 3-kinase is required for the formation of constitutive transport vesicles from the TGN. *Journal of Cell Biology* 139, 339–349.
- Jones, S.M., Howell, K.E., Henley, J.R., Cao, H., McNiven, M.A., 1998. Role of dynamin in the formation of transport vesicles from the trans-Golgi network. *Science* 279, 573–577.
- Julicher, F., Lipowsky, R., 1993. Domain-induced budding of vesicles. *Physical Review Letters* 70, 2964–2967.
- Kay, J.G., Murray, R.Z., Pagan, J.K., Stow, J.L., 2006. Cytokine secretion via cholesterol-rich lipid raft-associated SNAREs at the phagocytic cup. *Journal of Biological Chemistry* 281, 11949–11954.
- Klemm, R.W., Ejising, C.S., Surma, M.A., *et al.*, 2009. Segregation of sphingolipids and sterols during formation of secretory vesicles at the trans-Golgi network. *Journal of Cell Biology* 185, 601–612.
- Kreitzer, G., Marmorstein, A., Okamoto, P., Vallee, R., Rodriguez-Boulant, E., 2000. Kinesin and dynamin are required for post-Golgi transport of a plasma-membrane protein. *Nature Cell Biology* 2, 125–127.
- Lefrançois, S., Zeng, J., Hassan, A.J., Canuel, M., Morales, C.R., 2003. The lysosomal trafficking of sphingolipid activator proteins (SAPs) is mediated by sortilin. *EMBO Journal* 22, 6430–6437.
- Lieu, Z.Z., Lock, J.G., Hammond, L.A., *et al.*, 2008. A trans-Golgi network golgin is required for the regulated secretion of TNF in activated macrophages *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America* 105, 3351–3356.
- Lingwood, D., Simons, K., 2010. Lipid rafts as a membrane-organizing principle. *Science* 327, 46–50.
- Lippincott-Schwartz, J., Roberts, T.H., Hirschberg, K., 2000. Secretory protein trafficking and organelle dynamics in living cells. *Annual Review of Cell and Developmental Biology* 16, 557–589.
- Liu, S., Storrie, B., 2012. Are Rab proteins the link between Golgi organization and membrane trafficking? *Cellular and Molecular Life Sciences* 69, 4093–4106.
- Lock, J.G., Stow, J.L., 2005. Rab11 in recycling endosomes regulates the sorting and basolateral transport of E-cadherin. *Molecular Biology of the Cell* 16, 1744–1755.
- Low, P.C., Misaki, R., Schroder, K., *et al.*, 2010. Phosphoinositide 3-kinase delta regulates membrane fission of Golgi carriers for selective cytokine secretion. *Journal of Cell Biology* 190, 1053–1065.
- Lowery, J., Szul, T., Styers, M., *et al.*, 2013. The Sec7 guanine nucleotide exchange factor GBF1 regulates membrane recruitment of BIG1 and BIG2 guanine nucleotide exchange factors to the trans-Golgi network (TGN). *Journal of Biological Chemistry* 288, 11532–11545.
- Mallard, F., Tang, B.L., Galli, T., *et al.*, 2002. Early/recycling endosomes-to-TGN transport involves two SNARE complexes and a Rab6 isoform. *Journal of Cell Biology* 156, 653–664.
- Malsam, J., Sollner, T.H., 2011. Organization of SNAREs within the Golgi stack. *Cold Spring Harbor Perspectives in Biology* 3, a005249.
- Manderson, A.P., Kay, J.G., Hammond, L.A., Brown, D.L., Stow, J.L., 2007. Subcompartments of the macrophage recycling endosome direct the differential secretion of IL-6 and TNF α . *Journal of Cell Biology* 178, 57–69.
- Manes, S., del Real, G., Martinez, A.C., 2003. Pathogens: Raft hijackers. *Nature Reviews. Immunology* 3, 557–568.
- McMahon, H.T., Gallop, J.L., 2005. Membrane curvature and mechanisms of dynamic cell membrane remodelling. *Nature* 438, 590–596.
- van Meer, G., Vaz, W.L., 2005. Membrane curvature sorts lipids. Stabilized lipid rafts in membrane transport. *EMBO Reports* 6, 418–419.
- Meinecke, M., Boucrot, E., Camdere, G., *et al.*, 2013. Cooperative recruitment of dynamin and BIN/amphiphysin/Rvs (BAR) domain-containing proteins leads to GTP-dependent membrane scission. *Journal of Biological Chemistry* 288, 6651–6661.
- Mellman, I., Nelson, W.J., 2008. Coordinated protein sorting, targeting and distribution in polarized cells. *Nature Reviews. Molecular Cell Biology* 9, 833–845.
- Micaroni, M., Stanley, A.C., Khromykh, T., *et al.*, 2013. Rab6a/a' are important Golgi regulators of pro-inflammatory TNF secretion in macrophages. *PLoS One* 8, e57034.
- Miller, P.M., Folkmann, A.W., Maia, A.R., *et al.*, 2009. Golgi-derived CLASP-dependent microtubules control Golgi organization and polarized trafficking in motile cells. *Nature Cell Biology* 11, 1069–1080.
- Miserey-Lenkei, S., Chalancon, G., Bardin, S., *et al.*, 2010. Rab and actomyosin-dependent fission of transport vesicles at the Golgi complex. *Nature Cell Biology* 12, 645–654.
- Mizuno-Yamasaki, E., Rivera-Molina, F., Novick, P., 2012. GTPase networks in membrane traffic. *Annual Review of Biochemistry* 81, 637–659.
- Mogelsvang, S., Marsh, B.J., Ladinsky, M.S., Howell, K.E., 2004. Predicting function from structure: 3D structure studies of the mammalian Golgi complex. *Traffic* 5, 338–345.
- Munro, S., 2011. The golgin coiled-coil proteins of the Golgi apparatus. *Cold Spring Harbor Perspectives in Biology* 3, a005256.
- Murray, R.Z., Kay, J.G., Sangermani, D.G., Stow, J.L., 2005. A role for the phagosome in cytokine secretion. *Science* 310, 1492–1495.
- Murray, R.Z., Wylie, F.G., Khromykh, T., Hume, D.A., Stow, J.L., 2005. Syntaxin 6 and Vti1b form a novel SNARE complex, which is up-regulated in activated macrophages to facilitate exocytosis of tumor necrosis factor- α . *Journal of Biological Chemistry* 280, 10478–10483.
- Musch, A., 2004. Microtubule organization and function in epithelial cells. *Traffic* 5, 1–9.
- Narita, M., Young, A.R., Arakawa, S., *et al.*, 2011. Spatial coupling of mTOR and autophagy augments secretory phenotypes. *Science* 332, 966–970.
- Ngo, M., Ridgway, N.D., 2009. Oxysterol binding protein-related Protein 9 (ORP9) is a cholesterol transfer protein that regulates Golgi structure and function. *Molecular Biology of the Cell* 20, 1388–1399.
- Ohkura, T., Momose, F., Ichikawa, R., Takeuchi, K., Morikawa, Y., 2014. Influenza A virus hemagglutinin and neuraminidase mutually accelerate their apical targeting through clustering of lipid rafts. *Journal of Virology* 88 (17), 10039–10055.
- Parkinson-Lawrence, E.J., Shandala, T., Prodoehl, M., *et al.*, 2010. Lysosomal storage disease: Revealing lysosomal function and physiology. *Physiology* 25, 102–115.
- Peden, A.A., Oorschot, V., Hesser, B.A., *et al.*, 2004. Localization of the AP-3 adaptor complex defines a novel endosomal exit site for lysosomal membrane proteins. *Journal of Cell Biology* 164, 1065–1076.
- Percival, J.M., Hughes, J.A., Brown, D.L., *et al.*, 2004. Targeting of a tropomyosin isoform to short microfilaments associated with the Golgi complex. *Molecular Biology of the Cell* 15, 268–280.
- Perez-Victoria, F.J., Bonifacio, J.S., 2009. Dual roles of the mammalian GARP complex in tethering and SNARE complex assembly at the trans-golgi network. *Molecular and Cellular Biology* 29, 5251–5263.
- Puri, N., Roche, P.A., 2006. Ternary SNARE complexes are enriched in lipid rafts during mast cell exocytosis. *Traffic* 7, 1482–1494.
- Reczek, D., Schwake, M., Schroder, J., *et al.*, 2007. LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase. *Cell* 131, 770–783.
- Richardson, B.C., McDonald, C.M., Fromme, J.C., 2012. The Sec7 Arf-GEF is recruited to the trans-Golgi network by positive feedback. *Developmental Cell* 22, 799–810.
- Robinson, M.S., 2004. Adaptable adaptors for coated vesicles. *Trends in Cell Biology* 14, 167–174.
- Rodriguez-Boulant, E., Kreitzer, G., Musch, A., 2005. Organization of vesicular trafficking in epithelia. *Nature Reviews. Molecular Cell Biology* 6, 233–247.
- Rouille, Y., Rohn, W., Hoflack, B., 2000. Targeting of lysosomal proteins. *Seminars in Cell & Developmental Biology* 11, 165–171.
- Roux, A., Cuvelier, D., Nassoy, P., *et al.*, 2005. Role of curvature and phase transition in lipid sorting and fission of membrane tubules. *EMBO Journal* 24, 1537–1545.
- Saftig, P., Klumperman, J., 2009. Lysosome biogenesis and lysosomal membrane proteins: Trafficking meets function. *Nature Reviews. Molecular Cell Biology* 10, 623–635.
- Santiago-Tirado, F.H., Bretscher, A., 2011. Membrane-trafficking sorting hubs: Cooperation between PI4P and small GTPases at the trans-Golgi network. *Trends in Cell Biology* 21, 515–525.

- Scott, P.M., Bilodeau, P.S., Zhdankina, O., *et al.*, 2004. GGA proteins bind ubiquitin to facilitate sorting at the trans-Golgi network. *Nature Cell Biology* 6, 252–259.
- Stockli, J., Fazakerley, D.J., James, D.E., 2011. GLUT4 exocytosis. *Journal of Cell Science* 124, 4147–4159.
- Storrie, B., Micaroni, M., Morgan, G.P., *et al.*, 2012. Electron tomography reveals Rab6 is essential to the trafficking of trans-Golgi clathrin and COPI-coated vesicles and the maintenance of Golgi cisternal number. *Traffic* 13, 727–744.
- Stow, J.L., Murray, R.Z., 2013. Intracellular trafficking and secretion of inflammatory cytokines. *Cytokine & Growth Factor Reviews* 24, 227–239.
- Surma, M.A., Klose, C., Simons, K., 2012. Lipid-dependent protein sorting at the trans-Golgi network. *Biochimica et Biophysica Acta* 1821, 1059–1067.
- Takizawa, P.A., Yucel, J.K., Veit, B., *et al.*, 1993. Complete vesiculation of Golgi membranes and inhibition of protein transport by a novel sea sponge metabolite, ilimaquinone. *Cell* 73, 1079–1090.
- Traub, L.M., Apodaca, G., 2003. AP-1B: Polarized sorting at the endosome. *Nature Cell Biology* 5, 1045–1047.
- Veale, K.J., Offenhauser, C., Whittaker, S.P., Estrella, R.P., Murray, R.Z., 2010. Recycling endosome membrane incorporation into the leading edge regulates lamellipodia formation and macrophage migration. *Traffic* 11, 1370–1379.
- Wakana, Y., van Galen, J., Meissner, F., *et al.*, 2012. A new class of carriers that transport selective cargo from the trans Golgi network to the cell surface. *EMBO Journal* 31, 3976–3990.
- Weisz, O.A., Rodriguez-Boulán, E., 2009. Apical trafficking in epithelial cells: Signals, clusters and motors. *Journal of Cell Science* 122, 4253–4266.
- Yeaman, C., Le Gall, A.H., Baldwin, A.N., *et al.*, 1997. The O-glycosylated stalk domain is required for apical sorting of neurotrophin receptors in polarized MDCK cells. *Journal of Cell Biology* 139, 929–940.
- Zagury, D., Uhr, J.W., Jamieson, J.D., Palade, G.E., 1970. Immunoglobulin synthesis and secretion. II. Radioautographic studies of sites of addition of carbohydrate moieties and intracellular transport. *Journal of Cell Biology* 46, 52–63.

Regulation of the Secretory Pathway

A Capalbo, Institute of Protein Biochemistry, Naples, Italy

J Cancino, Universidad Andrés Bello, Viña del Mar, Chile

A Luini, Institute of Protein Biochemistry (IPB), Naples, Italy

© 2016 Elsevier Inc. All rights reserved.

Constitutive and Regulated Secretion

The eukaryotic cell relies heavily on spatiotemporal organization for optimal function. Constitutive secretory transport is a fundamental cellular positioning system by which a large fraction of the eukaryotic proteome (roughly 30% in mammals) is biochemically processed and transported from the site of synthesis, the endoplasmic reticulum (ER) to various appropriate final cellular destinations in its chemically mature form. Membrane transport is thus responsible for controlling the size, shape, and molecular composition of most cellular organelles, including the plasma membrane, and for mediating the secretion of thousands of cargo species, including hormones, growth factors, antibodies, matrix and serum proteins, digestive enzymes, and many more. Other eukaryotic positioning systems that operate in the cytoplasm (Lee *et al.*, 2013) and in the nucleus (Cavalli and Misteli, 2013) exist in the eukaryotic cell; however, they are much less well understood than membrane transport so far.

The regulation of secretory transport is a central issue in cell biology. Historically, there has been a long-standing distinction in the field of membrane transport between constitutive and regulated secretion (Burgess and Kelly, 1987). For several years, it has been held that, while constitutive transport is a general housekeeping process (see below), regulated secretion is restricted to a limited number of cell types and to the last stages of secretion, namely the fusion of a transport carrier with the plasma membrane (Burgess and Kelly, 1987). Cells such as, for instance, chromaffin or pituitary cells can generate granules containing large amounts of secretory material from the last station of the Golgi complex, the *trans*-Golgi network (TGN) (Burgess and Kelly, 1987; Tooze, 1991). These granules remain quiescent in the cytoplasm until cells are challenged by extracellular signals through plasma membrane (PM) receptors, resulting in signaling cascades that rapidly trigger the fusion of the granules with the PM and the release of their content outside the cell. Other cells do not possess secretory granules and secrete their cargo in an apparently continuous fashion. It was therefore thought that constitutive transport is not regulated, and this has been indeed the prevailing view for a long time. In the last decade, however, it has become clear that constitutive secretion also is potentially regulated by different kinases and signaling pathways at multiple stages of the transport pathway, with important physiological consequences (Simpson *et al.*, 2012; Cancino *et al.*, 2014; Chia *et al.*, 2012; Farhan *et al.*, 2010; Pulvirenti *et al.*, 2008; Mitrovic *et al.*, 2013). The regulation of constitutive transport is the central theme of this section.

The Regulation of Constitutive Secretion

Constitutive secretion begins with the translocation of soluble secretory cargo or membrane proteins into the lumen or the

membrane of the ER, respectively. Transport then continues with the transfer of cargo to the Golgi complex via the intermediate compartment and through the *cis*-, *medial*- and *trans*-Golgi sub-compartments to the TGN, the last Golgi station. At the TGN, cargo proteins are sorted into specialized membranous carriers for delivery to their appropriate cellular destinations, such as the apical or basolateral PM or the endolysosomal system, where cargo continues to be transported until it reaches a state of dynamic equilibrium at a suitable location. Each of these transport stages is balanced by suitable retrograde compensatory fluxes of membrane and proteins in order to keep the morphological and compositional homeostasis of the system. Finally, proteins that need to be degraded are delivered to two degradative systems by which cells regulate the concentration of misfolded and unwanted proteins, one located in the ER (endoplasmic reticulum-associated degradation (ERAD)), and another in the lysosome. This very large and complex apparatus is continuously challenged by exogenous signals or perturbations and by changing internal cellular states.

There are three general reasons for which the transport apparatus requires a sophisticated regulatory system to supervise its activities. One is to respond to environmental signals and demands; another is to support the homeostasis and the robustness of the transport apparatus itself (where robustness is the ability to function optimally under variable and uncertain conditions); and a third is to coordinate the activity of the transport apparatus with those of other cellular modules, such as energy production or motility. Indeed, there is increasing evidence that a large number of signaling molecules reside in the secretory organelles to control transport itself as well as other cell functions (Sallese *et al.*, 2006; Farhan and Rabouille, 2011; Cancino and Luini, 2013; Mayinger, 2011; Figure 1). The mechanisms and principles of this regulation are just beginning to be understood.

Extent of the Regulation Network of the Secretory Pathway

There are two main lines of evidence based on screening and on focused approaches that indicate that the cellular potential for regulation of the secretory machinery is very large.

A number of recent screening studies based on systematic depletion of the kinases and phosphatases as well as other human genes, have sought to unravel the signaling components that control cargo processing and transport as well as the morphology of the secretory pathway (Pelkmans *et al.*, 2005; Bard *et al.*, 2006; Chia *et al.*, 2012; Farhan *et al.*, 2010; Simpson *et al.*, 2012; Mitrovic *et al.*, 2013). What emerges from these investigations is that hundreds of kinases and phosphatases are involved in the regulation of the transport

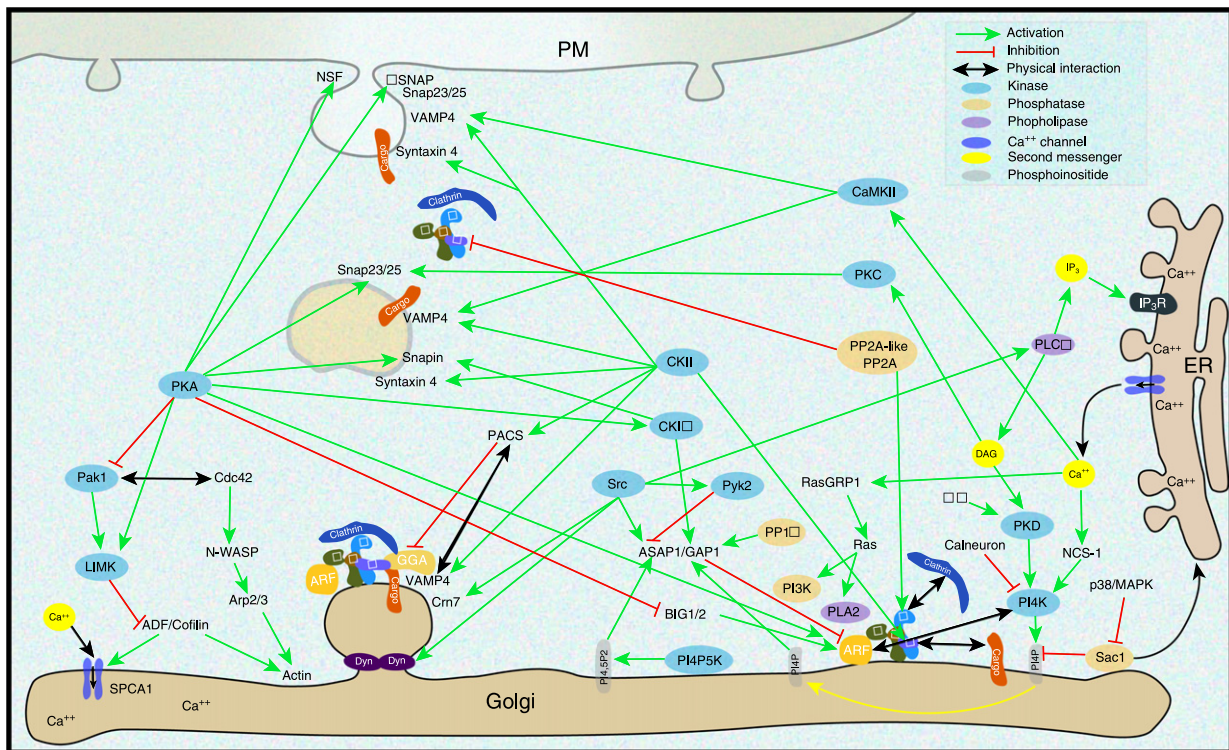


Figure 1 Golgi-based signaling molecules that regulate Golgi transport. Kinases, phosphatases, phospholipases, phosphoinositides, and second messengers that are involved in Golgi signaling and in controlling the traffic machinery. Reproduced with permission from Cancino, J., Luini, A., 2013. Signaling circuits on the Golgi complex. *Traffic* 14 (2), 121–134. doi: 10.1111/tra.12022. Epub 2012 Nov 15. Review. PMID: 23078632.

machinery affecting protein traffic as well as organelles morphology and organization. A large fraction of the signaling molecules controls cargo transport at the level of the ER and affects Golgi transport. Several kinases and phosphatases also affect the Golgi morphology (Chia *et al.*, 2012; Farhan *et al.*, 2010) as well as the glycosylation of cargo proteins that occurs in the Golgi, with potentially major functional consequences (Chia *et al.*, 2012). These dramatic effects are associated with the presence of many different phospho-proteins on the Golgi membranes (Farhan and Rabouille, 2011; Farhan *et al.*, 2010; Zacharogianni *et al.*, 2011).

In addition, a number of focused studies have shown that the flux of protein trafficking at specific stages of the secretory pathway is under the control of kinases (De Matteis *et al.*, 1993; Cancino and Luini, 2013; Farhan and Rabouille, 2011; Mayinger, 2011). There are well-characterized examples to support this case. One is the effect of protein kinase D (PKD) (Bossard *et al.*, 2007; Malhotra and Campelo, 2011) on transport out of the Golgi. Others, among many, are the effects of the tyrosine kinase Src (Pulvirenti *et al.*, 2008) on different phases of intra-Golgi transport, or of the mitogen activated protein kinases (MAPK) (Farhan *et al.*, 2010; Zacharogianni *et al.*, 2011) and of protein kinase A (PKA) (Cancino *et al.*, 2014; Cabrera *et al.*, 2003) on transport at the interface between the ER and the Golgi. PKA also regulates the constitutive TGN-PM transport (Muniz *et al.*, 1996, 1997). Several reviews of these and other similar studies are available (Cancino *et al.*, 2013; Cancino and Luini, 2013; Farhan and Rabouille, 2011; Mayinger, 2011).

This type of data, however, remains fragmentary at this stage. While these observations do indicate that, as noted, constitutive traffic can be extensively regulated, they do not provide a clear comprehension of the physiological significance of this regulation, as such comprehension would require that the information flow through these regulatory pathways be identified, with a definition of the specific activatory inputs, signaling cascades and effector molecules inducing specific cellular responses. This type of information has been so far partial and limited to a small number of cases, some of which are discussed below.

Traffic Regulation in Response to Environmental Signals and Demands

In a few cell types that have been studied, it has been shown that the secretory pathway is controlled by membrane receptors and related signaling pathways. This is presumably a part of physiological cellular responses to external inputs, although the significance of such responses is not always clear.

The first reported example of regulation of constitutive secretion by a PM sensor was concerned with the activation of the IgE receptor in mastocytes. The activation of the protein kinase C (PKC), which regulates the assembly of the COP-I vesicle involved in the intra-Golgi transport is required for this response (De Matteis *et al.*, 1993). The IgE receptor accelerates the secretion of glycans in tumoral mastocytic line and modulates the interaction of different vesicle components with Golgi membranes *in vitro* (Buccione *et al.*, 1996).

In other cell types, external and internal stimuli such as nutrients and mitogens are sensed by receptors impinging on the MAPK pathway, and result in the regulation of the number of endoplasmic reticulum exit site (ERES) as well as of the rate of cargo flux leaving the ER (Farhan *et al.*, 2010; Zacharogianni *et al.*, 2011). In particular, the activation of the EGFR activates ERK2 through the MAPK signaling cascade. ERK2 then modulates the formation of ERESs via phosphorylation of Sec16, an important player in the biogenesis of the ERESs and COPII (Farhan *et al.*, 2010). In this case, the activation of the secretory pathway is probably part of a general trophic cellular response to nutrients and growth signals.

Signaling Pathways that Support Homeostasis and Coordination of the Secretory Apparatus with Other Cellular Modules

The transport regulatory systems can also respond to internal cellular states, rather than to external inputs. Transport involves membrane fluxes across the transport compartments and is exposed to internal and external perturbations (Mironov *et al.*, 2001; Pulvirenti *et al.*, 2008; Trucco *et al.*, 2004). Such perturbations have the potential to disrupt the transport apparatus, particularly if they occur at the interface between the ER and the Golgi complex. The ER is up to 10-fold larger than the Golgi (Griffiths *et al.*, 1984), and changes in its membrane output can profoundly alter Golgi morphology and composition if not compensated for by corresponding adaptive changes in the retrograde or anterograde traffic out of the Golgi (Griffiths *et al.*, 1984; Klumperman, 2000; Martinez-Menarguez *et al.*, 1999; Thor *et al.*, 2009; Wieland *et al.*, 1987; Pulvirenti *et al.*, 2008). The compensatory mechanisms can be of two kinds. One relies on the self-assembly properties (Gong *et al.*, 2008; Heinrich and Rapoport, 2005) of the secretory machinery itself. Elegant models have been proposed to explain the behavior of segments of the secretory pathway based simply on the properties and the stoichiometry of the transport machinery molecules involved in those steps (Gong *et al.*, 2008; Heinrich and Rapoport, 2005; Sengupta and Linstedt, 2011). And a second major type of compensatory mechanisms consists of devices that sense the changes/ perturbations in the transport rates and respond quickly and precisely to restore homeostasis, though the activation of signaling and actuating systems (Cancino *et al.*, 2013, 2014).

Thus, at the ER–Golgi interface, membrane fluxes leaving the ER for the Golgi result in the activation of a traffic sensor, or receptor, called KDEL receptor (KDEL-R) at the *cis*-Golgi (Cancino *et al.*, 2014; Giannotta *et al.*, 2012; Pulvirenti *et al.*, 2008). The KDEL-R is a seven-transmembrane domain protein belonging to the PQ-loop protein family (Saudek, 2012) that is distantly related to the G-protein-coupled receptor (GPCR) superfamily (Yee *et al.*, 2013; Zhai *et al.*, 2001) and resembles the GPCRs in topology and fold of the transmembrane helices (Yee *et al.*, 2013; Zhai *et al.*, 2001). The activation of the KDEL-R by transport is most likely mediated by a family of ER resident proteins (Cancino *et al.*, 2014; Giannotta *et al.*, 2012; Pulvirenti *et al.*, 2008), the chaperones, which carry a KDEL signal at their C-terminus and cycle between the ER and the Golgi. Of note, the KDEL-R has been discovered and

characterized as molecules involved in salvaging chaperones leaked out of the ER (Lewis and Pelham, 1990; Munro and Pelham, 1987; Pelham, 1988). We now know that when export from the ER brings chaperones to the Golgi, these proteins bind the KDEL-R (Semenza *et al.*, 1990), which then activates a Golgi pool of the G proteins Gq (stimulatory heterotrimeric G-protein) and Gs (stimulatory heterotrimeric G-protein) and the relative transduction pathways. Gs then activates a signaling cascade at the *cis*-Golgi resulting in the phosphorylation of a large number of Golgi and cytosolic proteins and in the consequent activation of retrograde membrane transport (Cancino *et al.*, 2014). Gq, instead, appears to act by inducing a local release of calcium (Micaroni, 2010) (possibly via stimulation of a PLC and IP3 production) and the activation of the kinase SRC. SRC phosphorylates a large number of proteins, which activate anterograde traffic. Thus, a device based on the KDEL-R and two signaling pathways appears to sense incoming traffic into the Golgi and to activate anterograde and retrograde (Figure 2) traffic steps. This reaction counteracts the accumulation of membrane/proteins that would result from a non-compensated continuous arrival of traffic to the Golgi (Klumperman, 2000; Martinez-Menarguez *et al.*, 1999; Wieland *et al.*, 1987) and helps to maintain Golgi homeostasis.

In addition to regulating the Golgi acutely, the KDEL-R exerts a long-term control on the availability of the transport machinery components. Thus, the KDEL-R regulates the expression of a large number of genes, many of which code for transport-related proteins that localize both in biosynthetic and in endo-lysosomal organelles such as SNAREs, adapter complexes, GTPases, chaperones, Golgi and lysosomal enzymes (Cancino *et al.*, 2014). Most likely, a prolonged activation of the KDEL-R is decoded by the cell as a signal of chronic transport overload, which needs to be compensated by the expansion of the transport pathways.

Also of note, the KDEL-R up-regulate other genes that belong to functional modules that are different from transport, such as mitochondrial and energy metabolism, peroxisomal and lipid metabolism, protein degradation and autophagy (Cancino *et al.*, 2014). These responses indicate the presence of mechanisms that coordinate membrane transport with other cellular functions.

These KDEL-R-centered mechanisms are probably just an example of the regulatory devices that operate at different levels of the transport pathways.

The Conceptual Framework

A useful paradigm to understand the above regulatory processes is the theory of control of complex machines. Cells are composed of modules, or subsystems, that execute functions such as protein synthesis or folding, metabolism, membrane transport, autophagy, and apoptosis (Hartwell *et al.*, 1999; Sontag, 2004; Stelling *et al.*, 2004; Csete and Doyle, 2002). These functional modules and the related organelles must be able to preserve their homeostasis and to operate in a complex, coordinated, sensitive, and robust fashion (Waldherr *et al.*, 2008).

Robustness, complexity and sensitivity are features that modern engineers have incorporated into manmade machines

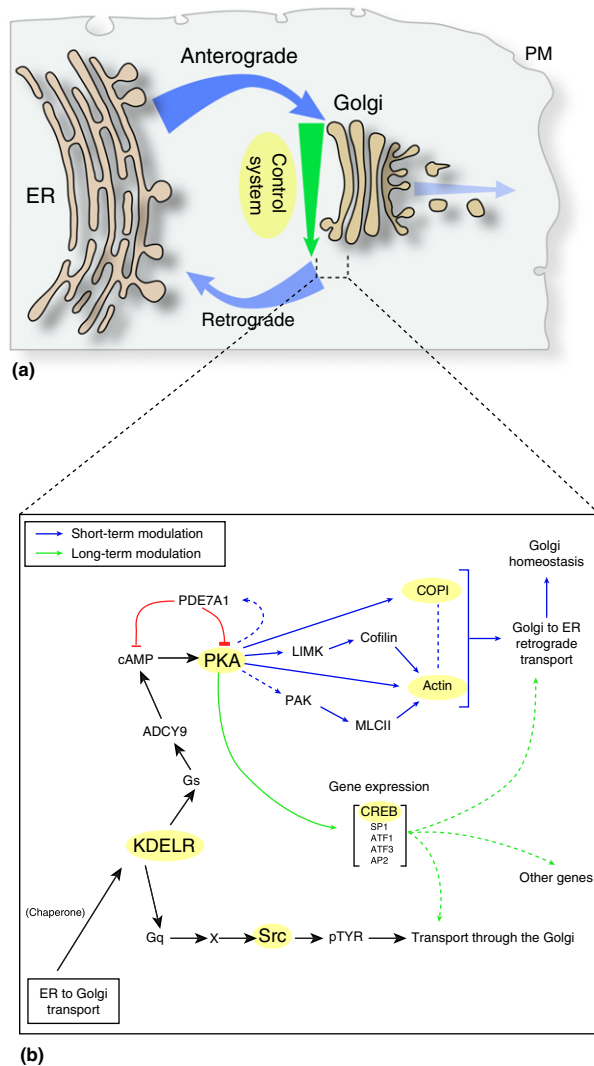


Figure 2 A control system balances the membrane fluxes between the ER and the Golgi and helps to maintain the homeostasis of secretory compartments. (a) A control system (green arrow) is activated at the *cis*-Golgi face by anterograde traffic from ER to Golgi and accelerates the retrograde fluxes from Golgi to ER as well as anterograde fluxes through the Golgi. This helps to balance the exchange of membranes at the ER–Golgi interface and maintains the homeostasis of the Golgi complex. (b) Scheme of the molecular mechanism involved in the Golgi control system. Transport fluxes that reach the Golgi activate the KDEL-R, which promotes the activation of the Gq and Gs heterotrimeric GTPases at the *cis*-Golgi face. Gs leads to the activation of AC9 to increase cAMP levels. cAMP then activates PKA, which increases retrograde transport via the phosphorylation of many proteins including components of the actin machinery and of the COPI complex (Cancino *et al.*, 2014). Moreover, the KDEL-R activates Gq, which induces the acceleration of anterograde traffic through the Golgi (Pulvirenti *et al.*, 2008). The prolonged activation of the KDEL-R results in the activation of several transcription factors including the cAMP/PKA-regulated CREB1. This in turn results in the upregulation of a number of transport machinery genes as well as of genes involved in other functions (Cancino *et al.*, 2014). Modified with permission from Cancino, J., Capalbo, A., Di Campli, A., *et al.*, 2014. Control systems of membrane transport at the interface between the endoplasmic reticulum and the Golgi. *Developmental Cell* 30, 280–294. doi: 10.1016/j.devcel.2014.06.018. PMID: 25117681.

(Hartwell *et al.*, 1999; Sontag, 2004; Stelling *et al.*, 2004; Csete and Doyle, 2002). For this, they have developed the theory of control. The control system is a central concept of the control theory. Such systems are mechanisms that regulate some basic process executed by a machine, or a factory, often by maintaining the homeostasis of the process despite internal or external perturbations that otherwise might have disruptive effects (Stelling *et al.*, 2004; Kitano, 2007; Csete and Doyle, 2002). Another important concept is that the activities of different modules must be coordinated. Such coordination is provided by specific devices, which manage the interface between modules, that is, in physical terms, the operation of the molecular pathways that coordinate separate modules.

Within this framework, the KDEL-R-based mechanism that senses the membrane fluxes that reach the Golgi from the ER and tends to compensate such fluxes by inducing retrograde as well as anterograde transport out of the Golgi can be viewed as a part of a control system that maintains the homeostasis of the Golgi complex. Similarly, the property of the KDEL-R to up-regulate the expression of genes belonging to different functional modules such as energy metabolism, peroxisomal and lipid metabolism, etc., when subjected to prolonged stimulation, can be seen as part of a coordination device that harmonizes the activity of membrane transport with that of other cellular functions.

Summary and Conclusions

Constitutive secretion has been described for a long time as a protein transport process for cellular housekeeping with no needs for signals. Relatively recently, however, the secretory pathway has been shown to be tightly regulated by a complex network of signaling machineries, PKA (Cabrera *et al.*, 2003; Cancino *et al.*, 2014; Martin *et al.*, 2000; Muniz *et al.*, 1996), PKD (Bossard *et al.*, 2007; Diaz Anel and Malhotra, 2005; Liljedahl *et al.*, 2001; Malhotra and Campelo, 2011), MAPK (Farhan *et al.*, 2010; Simpson *et al.*, 2012; Zacharogianni *et al.*, 2011), SRC (Giannotta *et al.*, 2012; Pulvirenti *et al.*, 2008). One of the most characterized control systems resides at the level of the Golgi apparatus (Cancino *et al.*, 2014; Giannotta *et al.*, 2012; Pulvirenti *et al.*, 2008). The Golgi control system operates by sensing the traffic input into the Golgi through the binding of the KDEL-R (here acting as a traffic sensor) with chaperones that escaped the ER. The KDEL-R then activates a controller; the Gs–PKA pathway, which ‘calculates’ the response and triggers the actuators of retrograde traffic, including the actin-based molecular machinery mentioned above (Cancino *et al.*, 2014). This compensates for the membrane input from the ER, and helps to maintain a suitable size and composition of the involved organelles. It is very likely that other internal control points exist and interact with exogenous signaling networks for optimal tuning of the cellular transport and distribution processes.

Altogether, the mechanisms that regulate constitutive secretion appear to be designed to sense and respond to both endogenous states and exogenous signals or perturbations in a harmonic and complex fashion while maintaining the homeostasis of the transport system.

See also: Interorganellar Communication: Interplay and Processes: Regulated versus Constitutive Secretion – A Major Form of Intercellular Communication

References

- Bard, F., Casano, L., Mallabiarrena, A., *et al.*, 2006. Functional genomics reveals genes involved in protein secretion and Golgi organization. *Nature* 439, 604–607.
- Bossard, C., Bresson, D., Polishchuk, R.S., Malhotra, V., 2007. Dimeric PKD regulates membrane fission to form transport carriers at the TGN. *Journal of Cell Biology* 179, 1123–1131.
- Buccione, R., Bannykh, S., Santone, I., *et al.*, 1996. Regulation of constitutive exocytic transport by membrane receptors. A biochemical and morphometric study. *Journal of Biological Chemistry* 271, 3523–3533.
- Burgess, T.L., Kelly, R.B., 1987. Constitutive and regulated secretion of proteins. *Annual Review of Cell and Developmental Biology* 3, 243–293.
- Cabrera, M., Muniz, M., Hidalgo, J., *et al.*, 2003. The retrieval function of the KDEL receptor requires PKA phosphorylation of its C-terminus. *Molecular Biology of the Cell* 14, 4114–4125.
- Cancino, J., Capalbo, A., Di Campli, A., *et al.*, 2014. Control systems of membrane transport at the interface between the endoplasmic reticulum and the Golgi. *Developmental Cell* 30, 280–294.
- Cancino, J., Jung, J.E., Luini, A., 2013. Regulation of Golgi signaling and trafficking by the KDEL receptor. *Histochemistry and Cell Biology* 140, 395–405.
- Cancino, J., Luini, A., 2013. Signaling circuits on the Golgi complex. *Traffic* 14, 121–134.
- Cavalli, G., Misteli, T., 2013. Functional implications of genome topology. *Nature Structural & Molecular Biology* 20, 290–299.
- Chia, J., Goh, G., Racine, V., *et al.*, 2012. RNAi screening reveals a large signaling network controlling the Golgi apparatus in human cells. *Molecular Systems Biology* 8, 629.
- Csete, M.E., Doyle, J.C., 2002. Reverse engineering of biological complexity. *Science* 295, 1664–1669.
- De Matteis, M.A., Santini, G., Kahn, R.A., Di Tullio, G., Luini, A., 1993. Receptor and protein kinase C-mediated regulation of ARF binding to the Golgi complex. *Nature* 364, 818–821.
- Diaz Anel, A.M., Malhotra, V., 2005. PKCeta is required for beta1gamma2/beta3gamma2- and PKD-mediated transport to the cell surface and the organization of the Golgi apparatus. *Journal of Cell Biology* 169, 83–91.
- Farhan, H., Rabouille, C., 2011. Signalling to and from the secretory pathway. *Journal of Cell Science* 124, 171–180.
- Farhan, H., Wendeler, M.W., Mitrovic, S., *et al.*, 2010. MAPK signaling to the early secretory pathway revealed by kinase/phosphatase functional screening. *Journal of Cell Biology* 189, 997–1011.
- Giannotta, M., Ruggiero, C., Grossi, M., *et al.*, 2012. The KDEL receptor couples to Galphq11 to activate Src kinases and regulate transport through the Golgi. *EMBO Journal* 31, 2869–2881.
- Gong, H., Sengupta, D., Linstedt, A.D., Schwartz, R., 2008. Simulated de novo assembly of golgi compartments by selective cargo capture during vesicle budding and targeted vesicle fusion. *Biophysical Journal* 95, 1674–1688.
- Griffiths, G., Warren, G., Quinn, P., Mathieu-Costello, O., Hoppeler, H., 1984. Density of newly synthesized plasma membrane proteins in intracellular membranes. I. Stereological studies. *Journal of Cell Biology* 98, 2133–2141.
- Hartwell, L.H., Hopfield, J.J., Leibler, S., Murray, A.W., 1999. From molecular to modular cell biology. *Nature* 402, C47–C52.
- Heinrich, R., Rapoport, T.A., 2005. Generation of nonidentical compartments in vesicular transport systems. *Journal of Cell Biology* 168, 271–280.
- Kitano, H., 2007. Towards a theory of biological robustness. *Molecular Systems Biology* 3, 137.
- Klumperman, J., 2000. Transport between ER and Golgi. *Current Opinion in Cell Biology* 12, 445–449.
- Lee, C.F., Brangwynne, C.P., Gharakhani, J., Hyman, A.A., Julicher, F., 2013. Spatial organization of the cell cytoplasm by position-dependent phase separation. *Physical Review Letters* 111, 088101.
- Lewis, M.J., Pelham, H.R., 1990. A human homologue of the yeast HDEL receptor. *Nature* 348, 162–163.
- Liljedahl, M., Maeda, Y., Colanzi, A., *et al.*, 2001. Protein kinase D regulates the fission of cell surface destined transport carriers from the trans-Golgi network. *Cell* 104, 409–420.
- Malhotra, V., Campelo, F., 2011. PKD regulates membrane fission to generate TGN to cell surface transport carriers. *Cold Spring Harbor Perspectives in Biology* 3, a005280.
- Martin, M.E., Hidalgo, J., Rosa, J.L., Crottet, P., Velasco, A., 2000. Effect of protein kinase A activity on the association of ADP-ribosylation factor 1 to golgi membranes. *Journal of Biological Chemistry* 275, 19050–19059.
- Martinez-Menarguez, J.A., Geuze, H.J., Slot, J.W., Klumperman, J., 1999. Vesicular tubular clusters between the ER and Golgi mediate concentration of soluble secretory proteins by exclusion from COPI-coated vesicles. *Cell* 98, 81–90.
- Mayinger, P., 2011. Signaling at the Golgi. *Cold Spring Harbor Perspectives in Biology* 3 (8), a005264.
- Micaroni, M., 2010. The role of calcium in intracellular trafficking. *Current Molecular Medicine* 10, 763–773.
- Mironov, A.A., Beznoussenko, G.V., Nicoziani, P., *et al.*, 2001. Small cargo proteins and large aggregates can traverse the Golgi by a common mechanism without leaving the lumen of cisternae. *Journal of Cell Biology* 155, 1225–1238.
- Mitrovic, S., Nogueira, C., Cantero-Recasens, G., *et al.*, 2013. TRPM5-mediated calcium uptake regulates mucin secretion from human colon goblet cells. *eLife* 2, e00658.
- Muniz, M., Alonso, M., Hidalgo, J., Velasco, A., 1996. A regulatory role for cAMP-dependent protein kinase in protein traffic along the exocytic route. *Journal of Biological Chemistry* 271, 30935–30941.
- Muniz, M., Martin, M.E., Hidalgo, J., Velasco, A., 1997. Protein kinase A activity is required for the budding of constitutive transport vesicles from the trans-Golgi network. *Proceedings of the National Academy of Sciences of the United States of America* 94, 14461–14466.
- Munro, S., Pelham, H.R., 1987. A C-terminal signal prevents secretion of luminal ER proteins. *Cell* 48, 899–907.
- Pelham, H.R., 1988. Evidence that luminal ER proteins are sorted from secreted proteins in a post-ER compartment. *EMBO Journal* 7, 913–918.
- Pelkmans, L., Fava, E., Grabner, H., *et al.*, 2005. Genome-wide analysis of human kinases in clathrin- and caveolae/raft-mediated endocytosis. *Nature* 436, 78–86.
- Pulvirenti, T., Giannotta, M., Capestrano, M., *et al.*, 2008. A traffic-activated Golgi-based signalling circuit coordinates the secretory pathway. *Nature Cell Biology* 10, 912–922.
- Sallese, M., Pulvirenti, T., Luini, A., 2006. The physiology of membrane transport and endomembrane-based signalling. *EMBO Journal* 25, 2663–2673.
- Saudek, V., 2012. Cystinosin, MPDU1, SWEETs and KDELr belong to a well-defined protein family with putative function of cargo receptors involved in vesicle trafficking. *PLoS One* 7, e30876.
- Semenza, J.C., Hardwick, K.G., Dean, N., Pelham, H.R., 1990. ERD2, a yeast gene required for the receptor-mediated retrieval of luminal ER proteins from the secretory pathway. *Cell* 61, 1349–1357.
- Sengupta, D., Linstedt, A.D., 2011. Control of organelle size: The Golgi complex. *Annual Review of Cell and Developmental Biology* 27, 57–77.
- Simpson, J.C., Joggerst, B., Laketa, V., *et al.*, 2012. Genome-wide RNAi screening identifies human proteins with a regulatory function in the early secretory pathway. *Nature Cell Biology* 14, 764–774.
- Sontag, E.D., 2004. Some new directions in control theory inspired by systems biology. *Systematic Biology (Stevenage)* 1, 9–18.
- Stelling, J., Sauer, U., Szallasi, Z., Doyle 3rd, F.J., Doyle, J., 2004. Robustness of cellular functions. *Cell* 118, 675–685.
- Thor, F., Gautschi, M., Geiger, R., Helenius, A., 2009. Bulk flow revisited: Transport of a soluble protein in the secretory pathway. *Traffic* 10, 1819–1830.
- Tooze, S.A., 1991. Biogenesis of secretory granules. Implications arising from the immature secretory granule in the regulated pathway of secretion. *FEBS Letters* 285, 220–224.
- Trucco, A., Polishchuk, R.S., Martella, O., *et al.*, 2004. Secretory traffic triggers the formation of tubular continuities across Golgi sub-compartments. *Nature Cell Biology* 6, 1071–1081.
- Waldherr, S., Eissing, T., Allgöwer, F., 2008. Analysis of feedback mechanisms in cell-biological systems. In: Chung, M.J., Misra, P. (Eds.), *Proceedings of the 17th World Congress, The International Federation of Automatic Control*, Seoul, Korea, July 6–11, pp. 15861–15866. Seoul: Elsevier.
- Wieland, F.T., Gleason, M.L., Serafini, T.A., Rothman, J.E., 1987. The rate of bulk flow from the endoplasmic reticulum to the cell surface. *Cell* 50, 289–300.
- Yee, D.C., Shlykov, M.A., Vastermark, A., *et al.*, 2013. The transporter-opsin-G protein-coupled receptor (TOG) superfamily. *FEBS Journal* 280, 5780–5800.
- Zacharogianni, M., Kondylis, V., Tang, Y., *et al.*, 2011. ERK7 is a negative regulator of protein secretion in response to amino-acid starvation by modulating Sec16 membrane association. *EMBO Journal* 30, 3684–3700.
- Zhai, Y., Heijne, W.H., Smith, D.W., Saier Jr., M.H., 2001. Homologues of archaeal rhodopsins in plants, animals and fungi: Structural and functional predictions for a putative fungal chaperone protein. *Biochimica et Biophysica Acta* 1511, 206–223.

Regulated versus Constitutive Secretion – A Major Form of Intercellular Communication

J Stöckli, The University of Sydney, NSW, Australia
RC Piper, University of Iowa, Iowa City, IA, USA
DE James, The University of Sydney, NSW, Australia

© 2016 Elsevier Inc. All rights reserved.

Glossary

Action potential An electrical signal that travels down a neuron or other electrically excitable cell.

Endocytosis A process that allows cells to engulf and internalize material from the outside.

Hormone A protein or substance that is released into the blood from one cell and which has an effect on distant cells.

Receptor A protein that binds a 'ligand.' Ligands include proteins such as hormones.

Secretory pathway A series of membrane compartments through which proteins travel on their way to be excreted from the cell inside to the outside.

SNARE A particular class of membrane-embedded proteins that catalyzes fusion of biological membranes. It stands for Soluble NSF Attachment Protein Receptor.

Transport vesicle Small membrane compartment that carries select cargo (proteins or metabolites) from one place in the secretory pathway to another.

Summary

The concept of regulated versus constitutive secretion was first described almost 30 years ago. The constitutive pathway is a system where certain proteins synthesized by the cell transit through an intricate lipid membrane network, which ultimately fuses with the cell surface membrane releasing the constitutive protein cargo outside the cell. In this pathway, protein cargo moves unrestrained to the plasma membrane. The regulated system shares many features in common with the constitutive system except in the regulated system as the cargo moves through the membrane network it is diverted into a special storage compartment. This compartment can be triggered to fuse with the cell surface in response to external signals. Many types of regulated secretory pathways, their varied cargos and types of storage compartments housing them have been identified. Differences among these systems are found in how they are made, what cargo they store, how they look under the microscope, how rapidly they release their content to the external environment and what triggers release of these stores. Here we survey constitutive secretion and the different types of regulated secretion and briefly overview the mechanisms that dictate their formation and function.

Overview

The ability of cells to release factors into their external environment serves as a major route for cell-to-cell communication. This fundamental process known as secretion has been adapted for a number of biological processes, which allow cells to respond to changes in their environment, principally to either modulate the external environment or to alert both distant or nearby cells to this change (Derby and Gleeson, 2007). Hormones, for example, are secreted into the bloodstream and travel through the body until they find their specific receptor found on distant cells. The specific interaction

between a hormone and its receptor initiates intracellular signaling events that evoke a particular response.

Of course cells had to develop different kinds of secretory mechanisms to accommodate different environmental conditions. This called for the development of secretory systems that could release factors only under the right circumstances. In addition, the timescale of secretion needed to be tailored for specific functions – some extremely rapid on a millisecond timescale for purposes such as neurotransmission in the brain (Südhof, 2013), and others far slower that could sustain longer-term responses, such as release of antibodies to combat pathogens (Borghesi and Milcarek, 2006). In the case of the slower more tonic forms of secretion it is feasible for cells to match the secretory demand simply by manufacturing the required amount of protein as needed. However, in the case of rapid secretion, protein synthesis is too slow to keep up with this demand and so different kinds of release mechanisms are needed. To match this biological need, cells evolved the ability to store newly synthesized material until required, in some cases for very long periods of time (Burgess and Kelly, 1987). The former type of secretion is commonly referred to as 'constitutive secretion' and the latter as 'regulated secretion' (Figure 1).

The average cell makes more than 10 000 different kinds of proteins and each protein is endowed with a unique function, and often a unique location within the cell. Proteins enter the secretory pathway upon synthesis on the endoplasmic reticulum (ER), whereupon they are imported into the lumen if they are soluble secreted proteins, or embedded into the ER membrane if they function as integral membrane proteins (Figure 1). The ER has a unique chemical environment, which allows membrane or soluble proteins to fold into their correct 3-dimensional structure (Derby and Gleeson, 2007). Protein folding is a highly complicated process that sometimes goes awry. Because of this the ER utilizes a range of quality control safety checks on newly synthesized proteins to determine if they are correctly folded. If proteins are improperly folded, they are pumped out of the ER and degraded (Vembar and

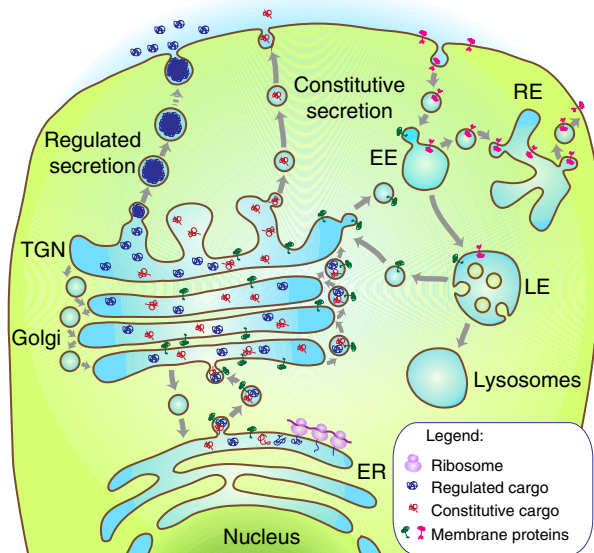


Figure 1 Exocytosis and endocytosis. Soluble protein cargo and membrane protein cargo is synthesized at the endoplasmic reticulum (ER) and transported through the secretory pathway. At the trans-Golgi network (TGN) cargo is sorted into constitutive or regulatory secretory vesicles that are either targeted for exocytosis and fusion with the cell surface or are sorted into intracellular transport vesicles destined for the endocytic pathway, such as early endosomes (EE), late endosomes (LE), or lysosomes. Membrane protein cargo is endocytosed into EE and either recycled to the cell surface via the recycling endosomes (RE) or moved into the degradative pathway through LE to lysosomes.

Brodsky, 2008). Examples include mutant forms of CFTR, which are depleted by ER-associated degradation leading to cystic fibrosis (Lukacs and Verkman, 2012). If a protein passes the safety check in the ER it is packaged into lipid bilayer-enclosed transport vesicles that are delivered to the next compartment in the secretory pathway, the Golgi apparatus. Here proteins can undergo an array of posttranslational modifications such as the addition of sugar residues. Once proteins have navigated their way through the Golgi to the trans-Golgi network (TGN) they can be ushered to one of several possible routes. Proteins destined for constitutive secretion are packaged into post-TGN vesicle carriers that move and fuse with the cell surface in an unrestrained manner, thus releasing cargo into the extracellular environment. Cargos of the regulated secretion pathway are packaged into different carriers stored inside the cell for days or even weeks until they receive an instruction to fuse with the cell surface (Burgess and Kelly, 1987; Figure 1).

Cargo of the regulated secretory pathway includes components of the endocrine (e.g., peptide hormones such as insulin, glucagon, neuropeptides, and adrenocorticotrophic hormone (ACTH)) and exocrine systems (e.g., digestive enzymes manufactured by the pancreas). While these molecules are all soluble proteins, regulated secretory cargo can also include membrane proteins such as nutrient transporters. Here, however, the regulated secretory process is generally used to insert those membrane proteins into the cell surface

membrane rather than secrete them from the cell. Other regulated secretory cargo includes metabolites, such as the amino acid glutamate, which acts as a neurotransmitter.

A major function of the regulated secretory pathway is to rapidly deliver a packet of cargo to the plasma membrane that can be released in a burst. While examples of regulated secretion can be found amongst single cell eukaryotes, the system is more widely used in multicellular organisms, which rely on the ability to respond in various ways to intra- and inter-cellular stimuli with both speed and functional purpose. As such, the regulated secretory process plays a central role in mediating various forms of cell communication ranging from endocrine signaling to neurotransmission.

How Do We Know about Regulated versus Constitutive Secretion?

The first observations that distinguished regulated secretion from constitutive secretion were made by simultaneously following the release of two different sets of cargo – viral glycoproteins versus ACTH – from cells cultured in the laboratory. ACTH is produced in the cell from a much larger precursor protein called Pro-opiomelanocortin (POMC). POMC is cut or processed into a discrete set of smaller peptides by enzymes known as convertases or peptidases during the maturation of regulated secretory vesicles (Zhou *et al.*, 1999). These smaller peptides are released via the regulated secretory system and each one has a unique receptor on a specific target cell to which it can bind. This enables one single cell, in this case the POMC producing cell, to transmit different kinds of information to multiple cells within the same multicellular environment.

When comparing the secretion of ACTH and the viral glycoprotein, both proteins were found to pass through the Golgi apparatus as determined by their ability to undergo Golgi-specific glycosylation events (Gumbiner and Kelly, 1982). Yet, the secretion of ACTH was greatly stimulated when cells were exposed to chemicals (secretagogues) that mimicked the second messenger of a stimulatory signal transduction pathway. In contrast, the viral glycoprotein was released from cells at a constant rate. In addition, using physical methods to separate intracellular compartments, it could be demonstrated that ACTH accumulated in a membrane-enclosed compartment, which did not contain the viral glycoprotein. This storage compartment represents the key feature of the regulated secretory pathway, housing a variety of constituents for release only when the appropriate stimulus is received.

Tour of Regulated Secretion

To exemplify the form and function of the regulated secretory system, let us consider the consumption of a meal, a process common to all animals that involves a number of regulated secretory systems (Figure 2). Note that a disruption to any one of these functions contributes to many common diseases in humans including anorexia nervosa, diabetes, and obesity.

Synaptic transmission is responsible not only for thinking about food, but also for triggering the motor activity required

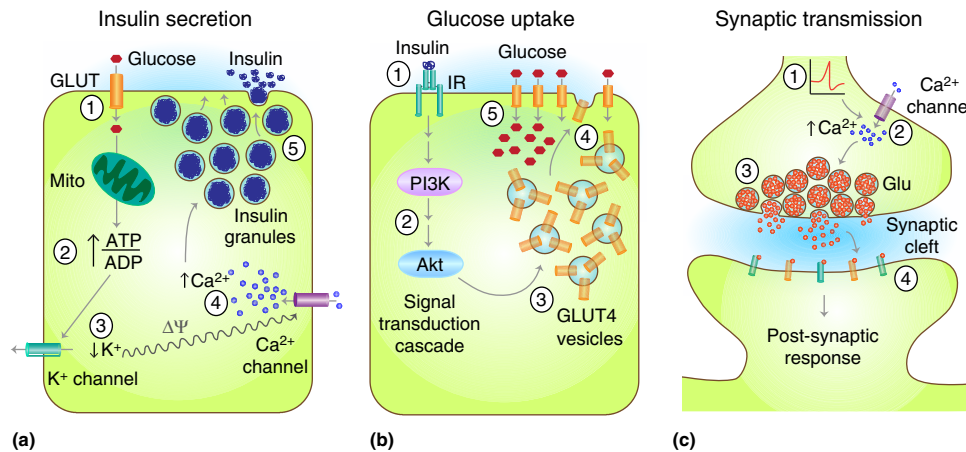


Figure 2 Regulated Secretion. A, Glucose-stimulated insulin secretion in beta cells exemplifies hormone protein secretion. Glucose enters the beta cell via a glucose transporter (GLUT) (1) and is metabolized via the mitochondrion (mito) and ATP is produced. This increases the ATP to ADP ratio (2), which activates the potassium (K^+) channel that exports K^+ from the cell (3). This leads to a voltage change ($\Delta\Psi$) that opens the calcium (Ca^{2+}) channel (4), triggering fusion of insulin granules with the cell surface and release of insulin. B, Insulin-stimulated translocation of GLUT4 to the cell surface in fat and muscle cells exemplifies exocytosis of membrane proteins. Insulin-binding to the insulin receptor (IR, 1) activates the PI3K/Akt signal transduction pathway (2) that triggers movement of GLUT4 containing vesicles to the cell surface (3) and fusion with the cell surface (4). This embeds GLUT4 in the cell surface membrane and allows glucose uptake into the cell (5). C, Glutamine release at the synapse in neurons is displayed as an example of metabolite cargo secretion. An action potential (1) results in Ca^{2+} influx (2) that triggers fusion of synaptic vesicles containing the neurotransmitter glutamate (Glu) with the synaptic cell surface (3) and release of glutamate into the synaptic cleft, where glutamate binds to cell surface receptors to trigger a postsynaptic response (4).

for finding it. At the center of neurotransmission is the generation and fusion of synaptic vesicles with the neuronal cell surface membrane (Figure 2(c)). This releases chemical neurotransmitters into the synaptic cleft, the small gap between adjacent cells, to evoke electrical changes in downstream neurons or muscle cells (De Camilli and Jahn, 1990; Südhof, 2004). Most synaptic vesicles contain small chemical neurotransmitters such as glutamate that are pumped into the lumen of the synaptic vesicle by a transport protein. Most cells possess large reservoirs of glutamate since it is obtained in the diet or via protein breakdown and so this kind of secretion mechanism does not require the cell to manufacture the secretory factor. Synaptic vesicles are induced to fuse with the cell surface in response to increased levels of the intracellular second messenger calcium, which enters through channels found in the surface membrane of the cell.

The entry of food into the digestive system triggers numerous secretory events essential for breaking down food into its individual digestible components, such as sugars, fats, and amino acids. This includes the release of enzymes by the exocrine pancreas to aid digestion. These enzymes are released by pancreatic acinar cells, which synthesize and store proteases and lipases, to breakdown large proteins and fat molecules in the lumen of the intestine (Messenger et al., 2014). Individuals lacking proper exocrine function suffer from pancreatic insufficiency, cannot digest food correctly, and develop nutrient malabsorption and malnutrition. Secretion of pancreatic exocrine factors is triggered by several stimuli including parasympathetic neurons, which release acetylcholine (ACh) via their own regulated secretory process. ACh receptors on acinar cells evoke production of a different kind of second messenger called IP3 (inositol-triphosphate). IP3 then binds to a specific IP3 binding protein on the membrane of the ER and triggers

the release of calcium from this organelle. This calcium signal causes enzyme-containing storage compartments (zymogen granules) to fuse with the cell surface and empty their contents into the gut.

Once food is broken down, its minimal components are absorbed across the intestinal wall for entry into the bloodstream. Specific cells in the pancreas sense some of the incoming dietary components, such as glucose (Aguilar-Bryan et al., 1995). These insulin-producing pancreatic beta cells become more energized by higher blood glucose leading to an increase in the ratio of ATP/ADP inside the cell (Figure 2(a)). This inhibits a potassium (K^+) channel that ultimately results in decreased voltage across the plasma membrane of the pancreatic beta cell. This voltage change ($\Delta\Psi$) causes the opening of another channel, a calcium (Ca^{2+}) channel (Henquin, 2000). The resulting elevation in intracellular calcium leads to fusion of insulin-containing granules with the cell surface releasing insulin into the bloodstream (Hou et al., 2009). Defective insulin secretion or insulin action contributes to diabetes in humans (see Box 1).

A major role of insulin is to stimulate glucose uptake into muscle and adipose tissue where it can be stored either as glycogen or triglyceride for subsequent use during times of energy demand. The ability of insulin to regulate glucose uptake in these organs is accomplished by triggering another regulated secretory pathway in these tissues that results in the mobilization of a glucose transporter known as GLUT4 from intracellular stores to the cell surface (Figure 2(b)). In contrast to the dense-core granules that release their cargo such as pancreatic zymogens or insulin outside the cell, GLUT4 storage compartments fuse with the cell surface and the glucose transporter is incorporated into the surface membrane (Bryant et al., 2002; Stöckli et al., 2011). This exposes GLUT4 to

Box 1 Diabetes – a disease of defective regulated secretion

There are two kinds of diabetes – Type 1 and Type 2. Type 1 diabetes, which commonly occurs in younger children, is due to autoimmune destruction of the pancreatic insulin-producing cells (Yoon and Jun, 2005). This involves beta cell autoantigens on antigen presenting cells that in turn activate cytotoxic T cells that then kill the body's own beta cells. Individuals with Type 1 diabetes therefore lack the ability to produce and release insulin after a meal and if untreated results in death due to an inability to store and utilize energy. For this reason, the discovery of insulin represents one of the greatest medical discoveries of modern times. Type 2 diabetes develops from an impaired response to insulin. Normally, insulin causes glucose to be absorbed by muscle and fat cells in the body, in part by using the regulated secretory system to move GLUT4 glucose transporters from intracellular vesicles to the plasma membrane (Bryant *et al.*, 2002; Stöckli *et al.*, 2011). This response becomes defective in certain people who are at risk of developing Type 2 diabetes. This defect, often referred to as insulin resistance, is associated with obesity and lack of physical activity. The body's response to insulin resistance is to secrete even more insulin from the pancreas to compensate for this defect. While this helps in the short-term, over time this additional burden on the pancreas may in certain individuals ultimately compromise the normal function of the pancreatic insulin-producing cells again leading to destruction of these cells and to diabetes (Ogihara and Mirmira, 2010).

glucose on the outside of the cell enabling it to ferry glucose into the cell. This process plays an essential role in removing nutrients, such as glucose, from the bloodstream after a meal. When it has fulfilled its task, GLUT4 moves back into the cell and is repacked into storage vesicles so that it can repeat this process many times. Defective GLUT4 exocytosis is a hallmark of insulin resistance and type 2 diabetes (see **Box 1**).

Most meals are also accompanied by a beverage, which results in production of glomerular filtrate by the kidney filling the bladder with urine. The bladder responds by using a regulated secretory response. Whereas for the above described processes the objective is to secrete proteins or small molecules outside the cell, the bladder uses the secretory pathway to add extra membrane to the cell surface so that the bladder can expand (Apodaca, 2004). This is accomplished by the unique properties of the umbrella cells lining the bladder. These cells possess a large intracellular reservoir of membranes that can be rapidly added to the surface membrane facilitating a doubling of the bladder surface area. This process is activated by stretch and pressure due to increased urine volume. Two other secretory mechanisms guard against the release of too much fluid into the bladder because this could result in dehydration. Vasopressin is a peptide hormone stored in dense-core secretory granules in posterior pituitary cells in the brain. Loss of fluid from the body results in increased blood pressure and increased osmolarity due to the increased salt concentration. Both of these changes trigger calcium entry into these pituitary secretory cells resulting in vasopressin secretion. Vasopressin increases water permeability of the distal tubule and collecting duct cells of the kidney allowing water to be reabsorbed into the blood. This is mediated by the mobilization of aquaporin

water channels from intracellular stores to the cell surface in these kidney cells (Ward *et al.*, 1999). As in the case of GLUT4, aquaporin is a membrane protein stored inside the cell in storage vesicles that fuse with the cell surface making aquaporin part of the surface membrane enabling it to transport water either in or out of the cell.

The Steps of Regulated Exocytosis

As indicated from the above examples, all regulated secretory processes share a set of conserved common principles (Burgess and Kelly, 1987; Derby and Gleeson, 2007):

Making the Regulated Secretory Compartment

The first specialized feature of the regulated exocytic process is the biosynthesis of the secretory compartment inside the cell – typically a discrete vesicle, that can house and store cargo for later release. This means cargo, such as hormone proteins or other cargo, has to be delivered to these compartments while other proteins such as constitutive secretory cargo need to be excluded from those compartments. Because the average cell produces in the order of 10–20 000 different proteins, sorting the secretory cargo from this sea of protein requires considerable regulation. The regulated secretory systems of multicellular organisms have evolved at least three different mechanisms to achieve this specificity and this is based upon the type of secretory cargo. Here we consider soluble protein hormones, membrane proteins and small molecules such as amino acids, as these exemplify these three separate pathways.

Protein hormones

Protein hormones such as insulin or neuropeptides are synthesized in the form of a pro-hormone in the ER and trafficked through the Golgi to the TGN (example in **Figure 2(a)**). In the TGN these pro-hormones are sorted away from the constitutive secretory pathway into regulated secretory vesicles. Two models have been suggested, 'sorting by entry,' where sorting motifs on the pro-hormones interact with sorting receptors to enter the budding secretory vesicle, or 'sorting by retention,' where pro-hormones as well as other cargo bud off the TGN as immature regulated secretory vesicles that after removal of nonregulated cargo become mature regulated secretory vesicles and it is likely that the biogenesis of regulated secretory vesicles involves a combination of both models (Arvan and Castle, 1998; Dikeakos and Reudelhuber, 2007; Gondre-Lewis *et al.*, 2012; Hou *et al.*, 2009). Several sorting motifs have been identified in pro-hormones, including paired basic amino acids that form a cleavage site for pro-hormone convertases or a 3-dimensional motif on the pro-hormone exposed surface that interacts with the 'sorting receptor' carboxypeptidase E (CPE) and can act as a sorting signal or retention signal (Dikeakos and Reudelhuber, 2007; Gondre-Lewis *et al.*, 2012; Hou *et al.*, 2009). CPE and the pro-hormone convertases contain a short α -helix that binds to cholesterol- and sphingolipid-rich lipid rafts in the TGN, thereby sorting them into regulated secretory vesicles (Tooze *et al.*, 2001). Another sorting mechanism involves aggregation and many regulated

secretory cargo proteins have the capability to aggregate at mildly acidic pH in the presence of Ca^{2+} in the TGN (Dannies, 1999). For example, the hormone insulin forms hexameric arrays and this is coordinated by a direct interaction between insulin and zinc. During maturation of the secretory vesicle pro-hormones such as pro-insulin are cleaved and processed by convertases and CPE that are activated due to the acidified milieu inside the vesicle thus generating mature hormones (Zhou *et al.*, 1999). In the case of insulin the processing leads to insolubility and the formation of large insoluble crystals. This ability to condense into very dense aggregates of protein is reflected morphologically by transmission electron microscopy where they appear as ‘dense-core granules’ owing to the concentrated proteins within them that absorb the stains often used in electron microscopy studies. They are also by definition much larger than most other intracellular organelles, typically of 300 nm diameter. This innate potential to form aggregates serves not only to help guide secretory proteins to their correct destination in the cell but it provides the hormone tremendous stability, protecting it from potential degradation processes. Given that secretory granules exist for many days inside the cell, this property is quite essential. Most importantly, this feature also means that the fusion of one single secretory granule with the cell surface membrane gives rise to the release of a large bolus of bioactive hormone.

Small synaptic vesicles

These vesicles are found in considerable abundance in neurons and are approximately 10 times smaller than protein hormone containing regulated secretory compartments (De Camilli and Jahn, 1990; Südhof, 2004) (example in Figure 2(c) and Box 2). To understand their biology a brief tour of neuronal anatomy and function is necessary. Neurons or nerve cells are excitable cells that may be considered as information transfer units. They sense information from one cell and convey it to others using exocytosis of synaptic vesicles. This process occurs specifically at the neuronal synapse, a specialized region where two neurons almost touch one another. Neurons are made up of a cell body like all cells and this is where the nucleus, the ER and Golgi apparatus are found. In addition, neurons possess two other specialized structures-axons and dendrites. The axons are finger like projections and they make the synaptic connections with other neurons. In some cases axons can be very long, the best example being the laryngeal nerve in the giraffe where a single axon runs for several meters from the brain, all the way down the neck around the heart in the chest and back up to the voice box or larynx in the throat. In fact, Richard Dawkins often defers to this poor design feature as a prime example of evolutionary principle. Synaptic vesicles are enriched at the extreme tips of axons where they make synaptic connections. This design feature made it impossible for this cell to use the same mechanism as those of the soluble protein secretory family because it would simply not be feasible to constantly produce protein in the cell body and transport it all the way to these axonal tips in time for rapid secretory demand. So synaptic vesicles rely on the use of very small molecules such as acetylcholine, serotonin, glutamate, or adrenaline that can be readily synthesized or metabolized in axonal tips. Synaptic vesicles possess transport proteins that

pump these small molecules inside the synaptic vesicle where they are stored until the synaptic vesicle is instructed to fuse with the membrane. This design feature means that synaptic vesicles can be rapidly recharged with their secretory cargo almost upon demand. In some cases neurons reabsorb any unused neurotransmitter from the synaptic cleft after a secretory event either directly or via neighboring glia cells so that it can be reemployed for the same purpose. One of the most commonly used class of antidepressant medications known as serotonin reuptake inhibitors block this process thus increasing the concentration of serotonin in the synaptic cleft thus potentiating the signal between the two cells.

Membrane protein exocytosis

Although the terms secretion and exocytosis were first devised to account for the physical release of soluble cargo from the cell to the external environment, this system is also used to selectively place membrane proteins onto the cell surface in certain cells to enable them to rapidly exchange information in a different way (example in Figure 2(b)). As noted above, two examples of this are the deployment of glucose transporters to the cell surface in adipocytes and muscle cells or water channels in bladder or kidney cells (Stöckli *et al.*, 2011; Ward *et al.*, 1999). The intracellular secretory storage vesicles are often small (50–100 nm in diameter) and once the membrane protein cargo has been delivered to the cell surface membrane they are not released from the cell like soluble cargo but rather they can be reabsorbed back into the cell in a process known as endocytosis (see Box 2). In contrast to soluble cargo, which is found entirely inside the secretory organelle, membrane proteins span the membrane and often possess protein domains on both sides of the membrane. The domains that are located on the cytosolic side often possess address tags or targeting signals that are recognized by sorting machinery in the cytosol whose job is to escort these proteins back to their correct destination – the small recycling secretory vesicles.

Storage Inside the Cell

Once made, the regulated secretory compartments, whether dense-core granules, vesicles or other carriers, are stored within the cell. Depending on the cell and the cargo, these compartments can be stored for long periods of time. In addition, many of these mature storage compartments behave as though they are largely isolated from the rest of the secretory and endocytic pathways since little material is exchanged between the storage compartments and other intracellular compartments. At some point, these compartments will fuse with the cell surface, and often these compartments are positioned within the cell to optimize this latter activity. For some dense-core granules, this involves transporting and positioning the granules near the cell surface, and this is achieved by linking the granule to the cytoskeleton. Positioning of synaptic vesicles is also a key aspect for their function since they must fuse quickly with the cell surface upon the stimulus and as noted above they must fuse in the right place (the synapse) within the cell surface membrane so that the neurotransmitter is secreted into the cleft where two neurons are almost touching (Südhof, 2013). This process is often referred to as pre-docking

Box 2 Synaptic transmission – the fastest regulated secretory system

One of the most complex yet instructive examples of regulated secretion takes place at the synapse and controls the release of neurotransmitters (Burgoyne and Morgan, 2003; De Camilli and Jahn, 1990). Here synaptic vesicles filled with neurotransmitter fuse with the pre-synaptic plasma membrane to release their contents so they can diffuse across to nearby neurons that have neurotransmitter receptors on their surface postsynaptic membrane (Figure 2(c)). All of the basic processes and cellular machinery for regulated secretion are in place for synaptic transmission. However, on top of these basic mechanisms are additional features that allow synaptic vesicles to do something both critical and rare across all the different types of regulated secretory pathways – that is to respond incredibly fast. This is done by using a very fast sensor for calcium and by poising the storage compartment (the synaptic vesicle) at a late stage in the overall fusion process such that only a few molecular rearrangements need to be made before the completion of fusion.

Figure 3 shows an overview of the process and revolves around the action of synaptotagmin, a calcium-binding protein that binds calcium with the same fast kinetics that is shown by the overall process of synaptic vesicle fusion. The storage compartment for most neurotransmission is derived continuously from the endocytic pathway as many of the components of discrete intracellular neurotransmitter vesicles are retrieved from the plasma membrane. These components can be directly endocytosed together following what is known as a 'kiss-and-run fusion' event where the synaptic vesicle remains largely intact after fusion. Alternatively, synaptic vesicle components can be more dispersed whereby specialized synaptic vesicles form after trafficking more deeply through the endocytic pathway.

Neurotransmitters are transported into synaptic vesicles through the combined action of an exchanger, which concentrates neurotransmitters such as glycine, dopamine, and glutamate into the vesicle lumen using a proton gradient, and a nearby proton ATPase pump that establishes the proton gradient. Once filled, the synaptic vesicles are immobilized in specialized areas of the membrane where fusion can take place. Here synaptic vesicles undergo 'docking' with the plasma membrane and a 'priming' step making them ready to undergo the last stages of fusion when the appropriate signal, a spike in intracellular calcium, is given (Figure 3). This primed state is thought to involve a state in which SNARE complexes are already partially formed and may also position the lipid bilayers of synaptic vesicles and the cell surface in a state where membranes have an intermediate stage of fusion called hemifusion, a state where the outer but not inner leaflet of the membranes has fused (Chernomordik and Kozlov, 2005). The ability to arrest fusion at this precarious state is afforded by a number of specialized proteins including synaptotagmin and complexin. In the absence of calcium, synaptotagmin binds SNARE complexes and prevents them from catalyzing full fusion. An action potential traveling down the axon will activate voltage-gated calcium channels allowing calcium to enter the cell near the docked synaptic vesicle. Synaptotagmin binds calcium via special calcium binding domains that allow synaptotagmin to bury itself into the surface membrane and also bind stronger to SNARE complexes – now helping to complete the formation of SNARE complexes and complete the fusion process (Rizo and Rosenmund, 2008; Südhof, 2013).

and this enables the fusion process the capability to occur with incredible speed since arguably only one single biochemical step is required to achieve secretion. Exocytic vesicles containing membrane protein cargo also have a very slow rate of exchange with the cell surface under non-stimulated conditions and a specific signal like insulin or vasopressin triggers their fusion with the cell surface (Stöckli *et al.*, 2011; Ward *et al.*, 1999). The fusion of these vesicles is not restricted to a specialized part of the cell surface, unlike synaptic vesicles.

Responding to a Signal

Each different kind of secretory cell responds to a unique signal empowering the system with enormous capacity to respond to a myriad of factors present in the outside world. This might be heat, stretch, pH, nutrients, or the presence of a foreign invader. Cells are designed with the unique ability to respond to these different triggers but usually each individual cell is designed to respond to just one trigger. For example, pancreatic beta cells respond to nutrients but not to heat or stretch. Immune cells respond to foreign material but importantly not to other materials that might be deemed as part of ourselves otherwise this might trigger an undesirable self-attack as in the case of Type 1 diabetes (see Box 1) or other autoimmune diseases such as celiac disease, arthritis, or rheumatic Fever. The sensing of some kind of environmental change triggers a signal transduction pathway inside the cell. As noted above, calcium is a very common intracellular transmitter for many different kinds of regulated secretion but others include cyclic AMP, IP₃, changes in protein phosphorylation, nitric oxide, or specific kinds of lipids. The common elements that make each of these suitable as signal transducers is that each of them can be rapidly and readily acquired because of their abundance or the ability of cells to rapidly manufacture them, they are short lived so that the signal once generated can easily be extinguished. In the case of regulated secretion, the ultimate outcome is to accelerate the steps that enable vesicles to fuse with the surface membrane.

The Secretory Process

Südhof, Schekman, and Rothman won the 2013 Nobel Prize for Physiology or Medicine for their contributions to the discovery of the machinery that catalyzes the movement and fusion of vesicles in eukaryotic cells (Bonifacino, 2014). Central to their work was the discovery of the SNARE proteins, which are embedded in the membrane and catalyze fusion of vesicles with the cell surface (Rizo and Südhof, 2002; Rothman, 1996; Figure 3). The family of SNARE proteins has many members and different members localize to different compartments in the cell and catalyze distinct fusion events. This is an important aspect of SNARE proteins, since it helps explain the specificity of membrane fusion and why a regulated secretory dense-core granule fuses with the cell surface and not, for instance, with the ER or the Golgi. The second important aspect of SNARE proteins is that those on the vesicle and on the cell surface bind to each other, inducing a very energetically favored conformational change that brings membranes close enough to fuse and override the energy barrier that would

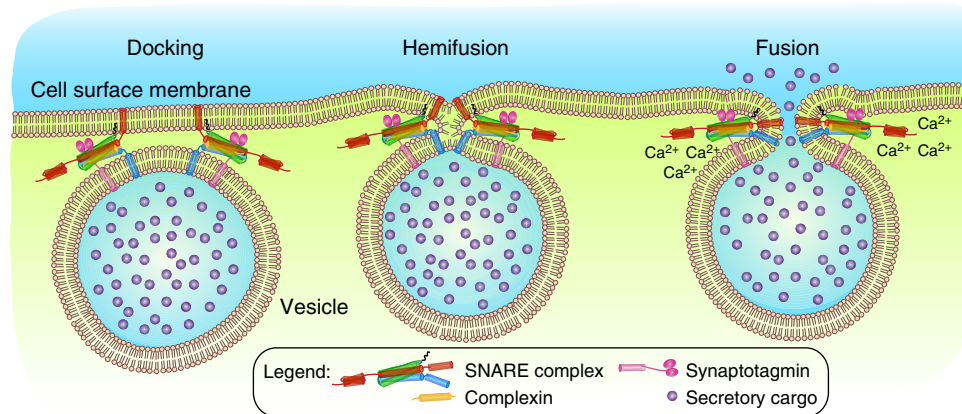


Figure 3 Molecular mechanism of synaptic fusion. Often synaptic vesicles are pre-docked or hemifused under unstimulated conditions and this enables very fast release of cargo once the signal is received. During docking, the SNARE complex is formed and complexin and synaptotagmin are bound to the SNARE complex. This can lead to hemifusion, where the inner leaflet of the membrane bilayer is fused but not the outer leaflet. Once Ca^{2+} is channeled into the synapse, Ca^{2+} binds to the calcium-binding domains in synaptotagmin that triggers binding and insertion of those domains into the cell surface membrane, thus generating a membrane curvature that results in fusion of the vesicle with the cell surface membrane and release of the cargo.

ordinarily keep membranes apart. Essentially, a SNARE protein, which is embedded in the membrane on one end, starts to wrap around other SNARE proteins, which are embedded in the opposing membrane. They start interacting on their free end and then continue to zipper toward the membrane anchors in a process that brings the two membranes close enough to force the melding of the lipid bilayers. For many regulated secretory systems, nature has added additional calcium-binding proteins onto this basic template, which allows fusion to be under the control of a surge in calcium. These calcium-sensitive proteins can work by holding SNARE proteins in an intermediate state during the fusion process, which only proceeds when calcium is present. Many additional proteins act prior to the formation of the SNARE complex, and since their activity can also be controlled, they too can contribute to trigger the fusion of regulated secretory vesicles. These additional proteins include the so-called ‘tethering’ factors, which help attach vesicles to the membranes they are targeted to for fusion. These tethering factors can respond to a variety of queues to help instigate the next steps of SNARE complex formation and thus can serve as trigger points that can induce regulated secretory vesicles to fuse to the plasma membrane.

Constitutive Secretion

Regulated secretion enables cells to respond rapidly, in some cases extremely rapidly, to changes in the outside world (Burgess and Kelly, 1987). This is appropriate as with eating, for example, numerous secretory processes need to be evoked with temporal precision and defects in this timing often lead to disease. Individuals with Type 1 diabetes are required to inject insulin prior to meals. If the timing of this is delayed this can result in major perturbations in blood glucose. Another important feature of regulated secretion is that cells can deliver pre-made contents without having to wait for the biosynthetic

process to synthesize and carry content all the way from the ER. There are many other kinds of secretion, however, for which these requirements are unnecessary and increased synthesis from the ER is an appropriate mechanism to increase the release of particular content into the extracellular environment.

The serum/plasma or the noncellular part of blood is comprised of many proteins with some found at extremely high concentrations – in the order of low millimolar levels. In many cases, these proteins serve what could be deemed as ‘house keeping’ functions. Albumin, one of the most abundant of all, maintains the osmotic pressure of blood and escorts other factors such as lipids and hormones from place to place in the body. Globulins assist in immune function and clotting factors are constantly at the ready to seal any breaks in the vascular system. Here it is not necessary to maintain a regulated secretory system but rather a system that constantly manufactures these proteins and then deposits them into the bloodstream in a completely unrestricted manner, hence ‘constitutive secretion.’ Another example of constitutive secretion is the secretion of antibodies from B lymphocytes, so named because they develop in the bone marrow. Each B lymphocyte can produce a specific antibody that might be directed against a particular pathogen. When that antibody is needed, the appropriate B cell is stimulated to produce large amounts of its particular antibody, which coincides with proliferation of the entire constitutive secretory pathway (Borghesi and Milcarek, 2006). Over the course of days, secretion by stimulated B cells will result in high levels (titers) of those antibodies in the bloodstream. While the production of the antibodies is stimulated, the subsequent release or secretion of the antibodies occurs constitutively. When the pathogen has been destroyed, the B cells are no longer stimulated. They ramp down their production of antibody and diminish the size of their secretory pathway. These resulting ‘memory’ B cells lie in wait to respond if the pathogen is detected in the future.

Conclusion

Secretory systems have evolved to become highly specialized to meet the numerous demands of multicellular life. These systems defend us against foreign attack, they underpin our ability to think, to contract our muscles, to smell and to see. They are crucial for eating and processing food into energy and they aid our ability to grow, fight stress, and to regulate every homeostatic system in our body. There are many different diseases and abnormalities that arise from defects in these secretory systems. All of these systems share a common set of rules and once you have mastered the principles of one, it is very easy to understand others. One of the most powerful aspects of these secretory systems for the future is that almost every cell in the human body secretes factors into the bloodstream. Hence, in a sense the blood is a blueprint for what is going on in every cell in the body. Therefore, one of the hopes of the future is that we can tap this resource to obtain hidden information about the health of different organs and use this information productively to treat disease before it has reached a point where it is no longer treatable. So in many ways both the regulated and constitutive secretory systems could hold the secret to our future long-term health in ways we have not even begun to imagine.

See also: Interorganellar Communication: Interplay and Processes: Regulation of the Secretory Pathway

References

- Aguilar-Bryan, L., Nichols, C.G., Wechsler, S.W., *et al.*, 1995. Cloning of the beta cell high-affinity sulfonylurea receptor: A regulator of insulin secretion. *Science* 268, 423–426.
- Apodaca, G., 2004. The uroepithelium: Not just a passive barrier. *Traffic* 5, 117–128.
- Arvan, P., Castle, D., 1998. Sorting and storage during secretory granule biogenesis: Looking backward and looking forward. *Biochemical Journal* 332 (Pt 3), 593–610.
- Bonifacio, J.S., 2014. Vesicular transport earns a Nobel. *Trends in Cell Biology* 24, 3–5.
- Borghesi, L., Milcarek, C., 2006. From B cell to plasma cell: Regulation of V(D)J recombination and antibody secretion. *Immunologic Research* 36, 27–32.
- Bryant, N.J., Govers, R., James, D.E., 2002. Regulated transport of the glucose transporter GLUT4. *Nature Reviews. Molecular Cell Biology* 3, 267–277.
- Burgess, T.L., Kelly, R.B., 1987. Constitutive and regulated secretion of proteins. *Annual Review of Cell Biology* 3, 243–293.
- Burgoyne, R.D., Morgan, A., 2003. Secretory granule exocytosis. *Physiological Reviews* 83, 581–632.
- Chernomordik, L.V., Kozlov, M.M., 2005. Membrane hemifusion: Crossing a chasm in two leaps. *Cell* 123, 375–382.
- Dannies, P.S., 1999. Protein hormone storage in secretory granules: Mechanisms for concentration and sorting. *Endocrine Reviews* 20, 3–21.
- De Camilli, P., Jahn, R., 1990. Pathways to regulated exocytosis in neurons. *Annual Review of Physiology* 52, 625–645.
- Derby, M.C., Gleeson, P.A., 2007. New insights into membrane trafficking and protein sorting. *International Review of Cytology* 261, 47–116.
- Dikeakos, J.D., Reudelhuber, T.L., 2007. Sending proteins to dense core secretory granules: Still a lot to sort out. *Journal of Cell Biology* 177, 191–196.
- Gondre-Lewis, M.C., Park, J.J., Loh, Y.P., 2012. Cellular mechanisms for the biogenesis and transport of synaptic and dense-core vesicles. *International Review of Cell and Molecular Biology* 299, 27–115.
- Gumbiner, B., Kelly, R.B., 1982. Two distinct intracellular pathways transport secretory and membrane glycoproteins to the surface of pituitary tumor cells. *Cell* 28, 51–59.
- Henquin, J.C., 2000. Triggering and amplifying pathways of regulation of insulin secretion by glucose. *Diabetes* 49, 1751–1760.
- Hou, J.C., Min, L., Pessin, J.E., 2009. Insulin granule biogenesis, trafficking and exocytosis. *Vitamins and Hormones* 80, 473–506.
- Lukacs, G.L., Verkman, A.S., 2012. CFTR: Folding, misfolding and correcting the DeltaF508 conformational defect. *Trends in Molecular Medicine* 18, 81–91.
- Messenger, S.W., Falkowski, M.A., Groblewski, G.E., 2014. Ca²⁺(+)-regulated secretory granule exocytosis in pancreatic and parotid acinar cells. *Cell Calcium* 55, 369–375.
- Ogihara, T., Mirmira, R.G., 2010. An islet in distress: Beta cell failure in type 2 diabetes. *Journal of Diabetes Investigation* 1, 123–133.
- Rizo, J., Rosenmund, C., 2008. Synaptic vesicle fusion. *Nature Structural & Molecular Biology* 15, 665–674.
- Rizo, J., Südhof, T.C., 2002. Snares and Munc18 in synaptic vesicle fusion. *Nature Reviews Neuroscience* 3, 641–653.
- Rothman, J.E., 1996. The protein machinery of vesicle budding and fusion. *Protein Science* 5, 185–194.
- Stöckli, J., Fazakerley, D.J., James, D.E., 2011. GLUT4 exocytosis. *Journal of Cell Science* 124, 4147–4159.
- Südhof, T.C., 2004. The synaptic vesicle cycle. *Annual Review of Neuroscience* 27, 509–547.
- Südhof, T.C., 2013. Neurotransmitter release: The last millisecond in the life of a synaptic vesicle. *Neuron* 80, 675–690.
- Tooze, S.A., Martens, G.J., Huttner, W.B., 2001. Secretory granule biogenesis: Rafting to the SNARE. *Trends in Cell Biology* 11, 116–122.
- Vembar, S.S., Brodsky, J.L., 2008. One step at a time: Endoplasmic reticulum-associated degradation. *Nature Reviews. Molecular Cell Biology* 9, 944–957.
- Ward, D.T., Hammond, T.G., Harris, H.W., 1999. Modulation of vasopressin-elicited water transport by trafficking of aquaporin2-containing vesicles. *Annual Review of Physiology* 61, 683–697.
- Yoon, J.W., Jun, H.S., 2005. Autoimmune destruction of pancreatic beta cells. *American Journal of Therapeutics* 12, 580–591.
- Zhou, A., Webb, G., Zhu, X., Steiner, D.F., 1999. Proteolytic processing in the secretory pathway. *Journal of Biological Chemistry* 274, 20745–20748.

Relevant Websites

- <http://www.nobelprize.org/mediaplayer/index.php?id=1977>
Nobelprize.
- <http://www.nobelprize.org/mediaplayer/index.php?id=1975>
Nobelprize.
- <http://www.nobelprize.org/mediaplayer/index.php?id=1986>
Nobelprize.

Clathrin and Clathrin-Dependent Endocytosis

FM Brodsky, University College London, London, UK

© 2016 Elsevier Inc. All rights reserved.

Introduction

Clathrin is an intracellular protein present in the cytosol of eukaryotic cells. When recruited to the cytosolic surface of the plasma and intracellular membranes, clathrin can self-assemble into a polyhedral lattice. Polymerization of clathrin organizes the adaptor proteins responsible for clathrin recruitment to membrane surfaces, causing adaptor-associated cargo molecules to concentrate within the clathrin-coated membrane, thereby selecting cargo for transport or sorting (Figure 1; Brodsky, 2012). In this capacity, clathrin-coated vesicles (CCVs) play a key role in endocytosis of many cell surface receptors, internalizing both constitutively active and ligand-gated receptors. CCVs also sort proteins in the trans-Golgi network (TGN) that are destined for lysosomes or for storage granules that undergo regulated secretion and clathrin is required for some recycling pathways. Thus clathrin contributes to selective membrane traffic to and from the plasma

membrane (PM), controlling how integral membrane proteins and their ligands are exposed to the different microenvironments enclosed by intracellular membrane-bound organelles. Patches of clathrin-coated membrane on endosomes and at the PM further mediate segregation of integral membrane proteins to create signaling domains (Kim *et al.*, 2013) or to target proteins for recycling (Majeed *et al.*, 2014) or degradation (Raiborg *et al.*, 2002). Finally, in a less-conserved function, clathrin plays a non-membrane traffic role by participation in microtubule cross-linking during mitosis and stabilizing the mitotic spindle.

Biochemistry of Clathrin

Clathrin is formed by trimerization of three clathrin heavy chain (CHC) subunits near their C-terminus. This forms a triskelion or three-legged shape that readily self-assembles into

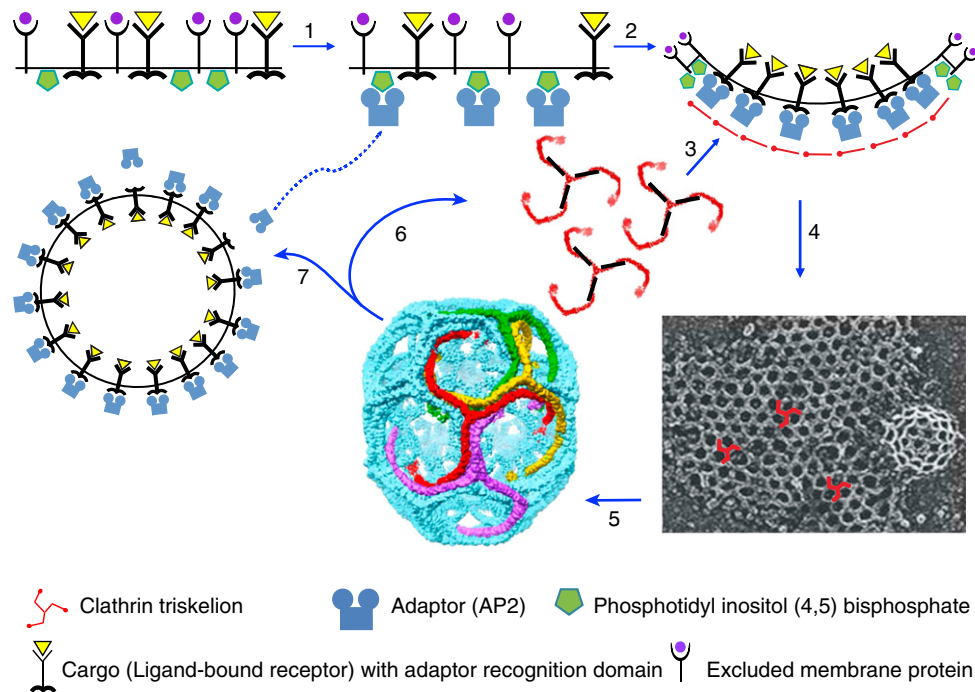


Figure 1 Clathrin function in cargo selection during receptor-mediated endocytosis. Step 1: Adaptor complexes (AP2) are recruited to the PM by interaction with the phosphoinositide PI(4,5)P, which causes a conformational change exposing the cargo-binding site. Step 2: Adaptors interact with cargo, which causes a conformational change exposing their clathrin-binding site. Step 3: Clathrin triskelia from the cytosol interact with adaptors and associated cargo. Step 4: Adaptors stimulate clathrin polymerization into a lattice, which can deform membranes into buds and generate coated vesicles containing selected adaptor-binding cargo. Step 5: CCVs are fully formed and excised from the membrane by dynamin polymerization at the vesicle neck. Step 6: The activity of Hsc70 ATPase and auxilin disassembles the clathrin lattice. Step 7: The adaptors are liberated from the uncoated vesicle by phosphoinositide phosphatase activity and the uncoated vesicle fuses into endosomes to deliver its contents. This figure is reproduced from Brodsky, F.M., 2012. Diversity of clathrin function: New tricks for an old protein. *Annual Review of Cell and Developmental Biology* 28, 309–336. The micrograph of the clathrin lattice is reproduced from Heuser, J.E., Keen, J.H., Amende, L.M., Lippoldt, R.E., Prasad, K., 1987. Deep-etch visualization of 27S clathrin: A tetrahedral tetramer. *Journal of Cell Biology* 105, 1999–2009. The model of triskelia on the clathrin coat is reproduced with permission from Fotin, A., Cheng, Y., Sliz, P., *et al.*, 2004b. Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* 432, 573–579. © 2004, Macmillan Publishers, Ltd.

Table 1 Clathrin subunits – genes and proteins

Clathrin subunit	Gene name	Human chromosome location	Length of human protein (aa)	Apparent molecular mass by SDS-PAGE (kDa)
Clathrin heavy chain 17 (CHC17)	CLTC	17q.23.2	1675	180
Clathrin light chain A (CLCa)	CLTA	9p.13.3	218-No inserts	34
		8p.22 (Pseudogene)	220-Splice variant s2	35
		12p.13.31 (Pseudogene)	236-Splice variant s1	36
			248-Neuronal splice variant	39
Clathrin light chain B (CLCb)	CLTB	5q.35.2	211-No insert 229-Neuronal splice variant	32 36
Clathrin heavy chain 22 (CHC22)	CLTCL1	22q11.21	1640	170

a polyhedral lattice (Heuser *et al.*, 1987). The triskelion is puckered at the vertex (Ferguson *et al.*, 2008), providing an orientation that favors formation of curved lattices and closure of assembled clathrin into a cage structure. Clathrin was discovered in 1975 by Barbara Pearse and named for the 'latticed' cage that it forms (Pearse, 1975). The triskelion shape was first described 6 years later (Ungewickell and Branton, 1981; Kirchhausen and Harrison, 1981; Crowther and Pearse, 1981).

CHCs

A gene duplication that occurred around the time that vertebrates evolved gave rise to two genes encoding CHCs, CLTC and CLTCL1 (formerly CLTD) (Wakeham *et al.*, 2005). These encode two distinct CHCs, CHC17 and CHC22, named for their human chromosome locations (Table 1). CHC17 binds clathrin light chain (CLC) subunits (see below) and performs the conventional membrane traffic functions associated with clathrin. CLTCL1 has become a pseudogene in mice, so mice are missing CHC22 protein. The status of CHC22 protein in other vertebrates is not known, though CLTCL1 genes appear to be intact in some other vertebrate genomes. CHC22 does not associate with CLCs in cells (Liu *et al.*, 2001). Non-vertebrate species have a single CHC (Wakeham *et al.*, 2005) and plants have two (Baisa *et al.*, 2013; Kania *et al.*, 2014), which appear to be CLC-binding.

The structures of the proximal leg domain and the CHC17 terminal domain have been determined to 2–3 Å resolution by X-ray crystallography (Figure 2; ter Haar *et al.*, 1998; Ybe *et al.*, 1999; von Kleist *et al.*, 2011). Contained within the proximal leg domain is a repeated motif of 10–12 short alpha helices, called a clathrin heavy chain repeat (CHCR). The triskelion leg comprises eight CHCRs (numbered 0–7) and the trimerization domain is partially formed by interactions of three CHCR7s (Ybe *et al.*, 2007), which extend into a longer triple helix coil (Fotin *et al.*, 2004b; Figure 2). CHCR motifs appear singly and multiply in other proteins (Ybe *et al.*, 1999) and for the Vps41 coat protein, the CHCR motif has been implicated in self-assembly, similar to its role in clathrin (Asensio *et al.*, 2013).

CHC17 is recruited to membranes through interaction of its N-terminal domain with the tetrameric adaptors AP1, AP2, and AP3 and the monomeric adaptors GGA1, GGA2, GGA3, and several other adaptors. The terminal domain has a seven-bladed beta-propeller structure, with mapped-binding sites for adaptor proteins between the blades (Lemmon and Traub, 2012; Figure 2). In cells, CHC22 interacts with AP1 and GGA2 and does not bind AP2 (Liu *et al.*, 2001). The protein sequences of human CHC17 and CHC22 are 85% identical, with sequence differences spread throughout (Wakeham *et al.*, 2005), and CHC22 is 35 amino acids smaller (Figure 2). No obvious sequences account for the difference in adaptor preference, thus CHC22 may interact with unknown adaptors that bias its membrane localization. The apparent molecular masses by SDS-PAGE are 180 kDa for CHC17 and 170 kDa for CHC22.

CLCs

Concomitant with vertebrate emergence, there was a gene duplication that gave rise to two genes, CLTA and CLTB, encoding CLCa and CLCb proteins in vertebrates (Wakeham *et al.*, 2005; Table 1). In humans, CLCa and CLCb are ~60% identical in protein sequence, and all vertebrate CLCa sequences are more homologous to each other than to the CLCb sequences from the same species. While the ratio of CLCa to CLCb is characteristic for specific tissues (Acton and Brodsky, 1990), their differential function is presently not known. CLCs associate exclusively with CHC17 in cells, though can interact with CHC22 in yeast two-hybrid assays. Yeast and flies have a single CLC (Wakeham *et al.*, 2005), while some plants have more than two types of CLCs (Baisa *et al.*, 2013).

Biochemical properties have been assigned to linear domains of the CLCs, although the biological functions of some of these domains are not yet elucidated. Target phosphorylation sites for casein kinase 2 unique to CLCb are present at the N-terminus (Hill *et al.*, 1988), and shared phosphorylation sites for GRK (Ferreira *et al.*, 2012), implicated in uncoating, are present at the C-terminus. CLCa has a unique sequence

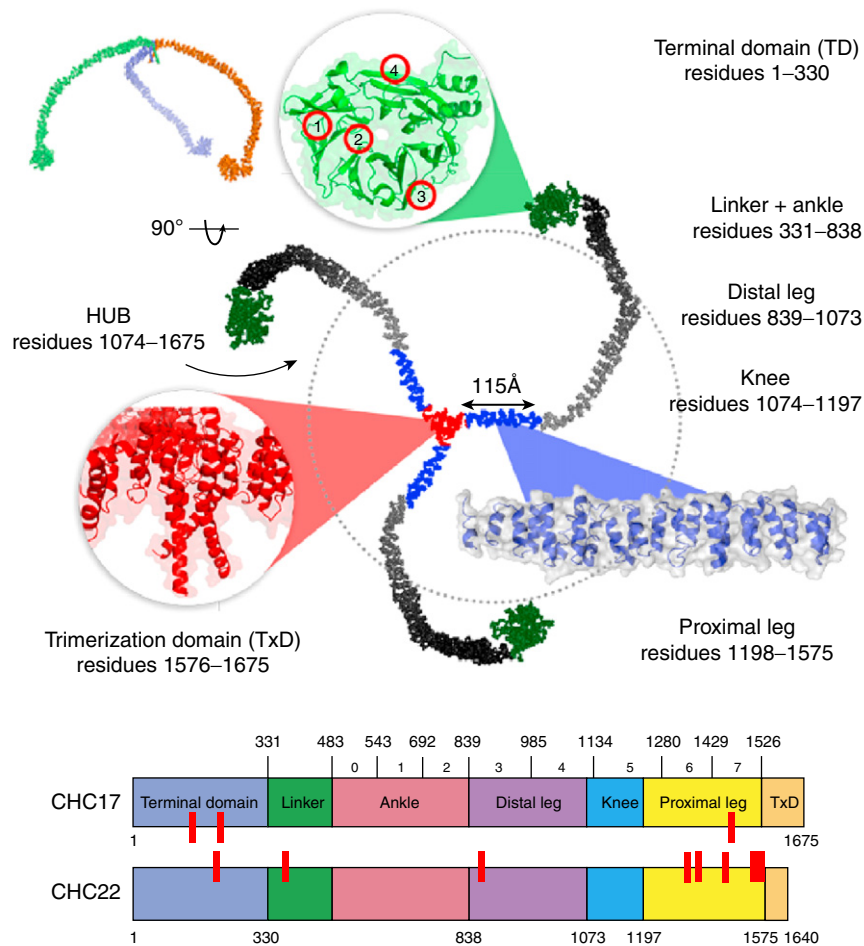


Figure 2 CHCs – structure and comparison. The location of CHC17 sequences and domains with solved structures are shown in the context of the triskelion. The structures shown are from PyMol (triskelion-accession number 3IYV; TxD-accession number 3LVH; TD-accession number 2XZG; proximal leg-accession number 1B89) and are based on Fotin *et al.* (2004b), ter Haar *et al.* (1998), Ybe *et al.* (1999). The numbering of the domain boundaries is based on human CHC17. The numbered sites on the TD structure are binding sites for adaptors, as detailed in Lemmon and Traub (2012). In the upper left is a side view of the triskelion demonstrating its characteristic pucker (PyMol accession number 3IYV). Below is a linear map of the amino acid boundaries of the eight CHCR structural repeats (0–7) relative to the structural domains of CHC17. This is aligned with the predicted domain structure of CHC22. The red marks indicate characteristic amino acid differences that are conserved in each type of CHC, as determined by DIVERGE (adapted from Wakeham *et al.*, 2005). This figure is reproduced from Brodsky, F.M., 2012. Diversity of clathrin function: New tricks for an old protein. *Annual Review of Cell and Developmental Biology* 28, 309–336.

that stimulates uncoating *in vitro* (DeLuca-Flaherty *et al.*, 1990) and both CLCs have a central calcium-binding domain (Näthke *et al.*, 1990), a sequence that binds the leucine-rich repeat kinase 2 (LRRK2) implicated in Parkinson's disease (Schreij *et al.*, 2015) and a C-terminal calmodulin-binding domain (Pley *et al.*, 1995; Figure 3). Additional diversity in vertebrate CLC protein is generated by RNA splicing (Figure 4; Jackson *et al.*, 1987). The CLTA gene has two exons encoding 12 and 18 residues that are retained in neuronal clathrin, but removed by splicing in the majority of non-neuronal tissues. Splicing that occurs during murine heart development generates an intermediate variant of CLCa with an insert of 12 residues and in brain (Giudice *et al.*, 2014), a fourth splice variant with an insert of 18 residues is detected (Kirchhausen *et al.*, 1987). The CLTB gene comprises an exon homologous to the longer spliced exon in CLTA, generating CLCb with an 18 residue insert in neurons and CLCb with no insert in non-

neuronal tissue. Thus there are four forms of CLCa and two of CLCb, with neurons expressing predominantly the longest forms of both. CLCb has only two cysteine residues, which are located near the C-terminus and readily form an internal disulfide bond under oxidizing conditions *in vitro* (not *in vivo*) (Figure 4). The smallest splice variant of CLCa has only one C-terminal cysteine, but the 12-residue insert introduces a second cysteine to the protein, allowing an internal disulfide to form *in vitro*. The formation of internal disulfide bonds cause a mobility shift of the CLC proteins such that molecular mass can be unreliable for distinguishing between some forms of CLCa and CLCb (Parham *et al.*, 1989). However, specific monoclonal antibodies are available that allow these forms to be distinguished (Brodsky *et al.*, 1987).

CLCs have an extended interaction with CHC17 residues from the trimerization domain at the triskelion vertex to the bend in the triskelion knee (Wilbur *et al.*, 2010; Figure 5). This

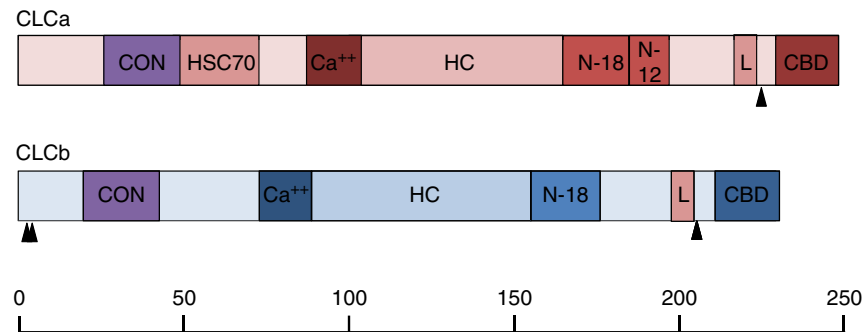


Figure 3 CLC protein domains. Domain maps of the vertebrate CLCs, CLCa and CLCb scaled to number of amino acids. Common functional domains indicated include the consensus sequence (CON) shared by all vertebrate CLCs, the calcium-binding sequence (Ca⁺⁺), the heavy chain-binding region (HC), the neuronal inserts of 18 (N-18) and 12 (N-12) residues, the LLRK2-binding domain (L), and the calmodulin-binding domain (CBD). Unique to CLCa is a region that can stimulate the uncoated ATPase, *in vitro* (HSC70). The black arrowheads show identified phosphorylation sites, serines 11 and 13 unique to CLCb (Hill *et al.*, 1988) and serine 205 (human CLCb) and its equivalent in CLCa (Ferreira *et al.*, 2012). Note that numbering varies slightly between species.

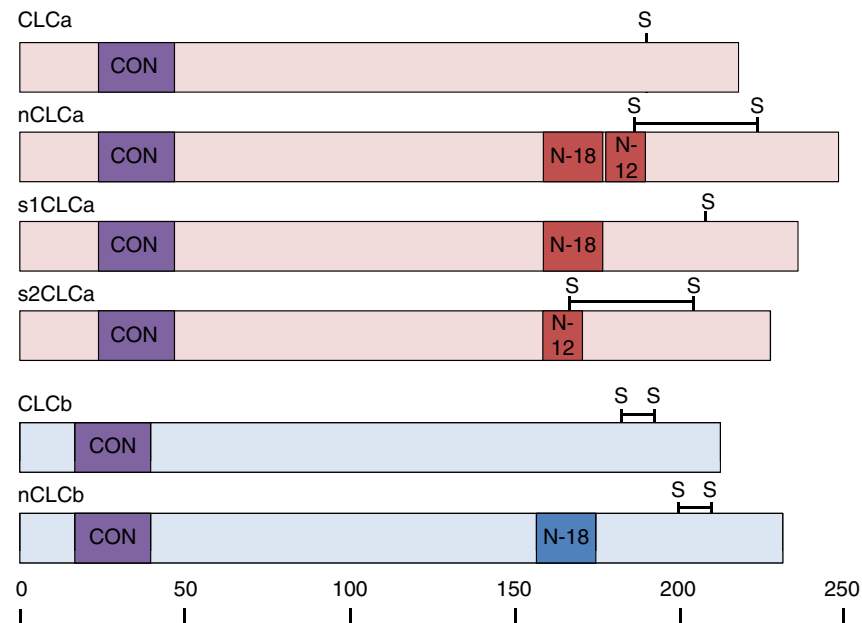


Figure 4 CLC diversity due to alternate splicing and internal disulfide bond formation. Alternate splicing produces four forms of CLCa and two forms of CLCb. Amino acid inserts of 18 and 12 residues can be encoded by two exons in *CLTA* and a homologous (but not identical) insert of 18 residues can be encoded by a separate exon in *CLTB*. The 12-residue insert includes a cysteine residue at positions 187 in nCLCa and 169 in s2CLCa, allowing an internal disulfide bond spanning 31 residues that can form with a conserved cysteine at positions 218 and 200 respectively, under oxidizing conditions after purification and during electrophoresis. CLCb and nCLCb have the conserved cysteine at positions 191 and 209 respectively that can form an internal disulfide bond spanning 10 residues with a CLCb-specific cysteine at positions 181 or 199.

stabilizes the trimerization domain (Ybe *et al.*, 2007) and buttresses the whole C-terminal third of the CHC, affecting triskelion flexibility and pucker (Dannhauser *et al.*, 2015). The presence of CLCs increases the tensile strength of the clathrin lattice, which could be influenced by all three of these properties (Dannhauser *et al.*, 2015). Interestingly, the spliced inserts are localized near the CLC C-terminal interaction with the trimerization domain, where they may affect these physical properties of clathrin. CLCs bind members of the Huntingtin-interaction protein (Hip) family (HIP1 and HIP1R in vertebrates, Sla2p in yeast) (Gottfried *et al.*, 2010), via 20 conserved residues near the CLC N-terminus (Chen and

Brodsky, 2005; Legendre-Guillemin *et al.*, 2005; Figure 3). Hips recruit actin and affect actin dynamics, thereby linking the CHC17 clathrin lattice to the actin cytoskeleton. Thus structurally and functionally, CLCs contribute strength to the clathrin lattice.

Clathrin Assembly and Disassembly

Assembly of CHC17 into a polyhedral lattice has been characterized biochemically, and visualized at 8–12 Å resolution by cryo-electron microscopy (Figure 1; Fotin *et al.*, 2004b).

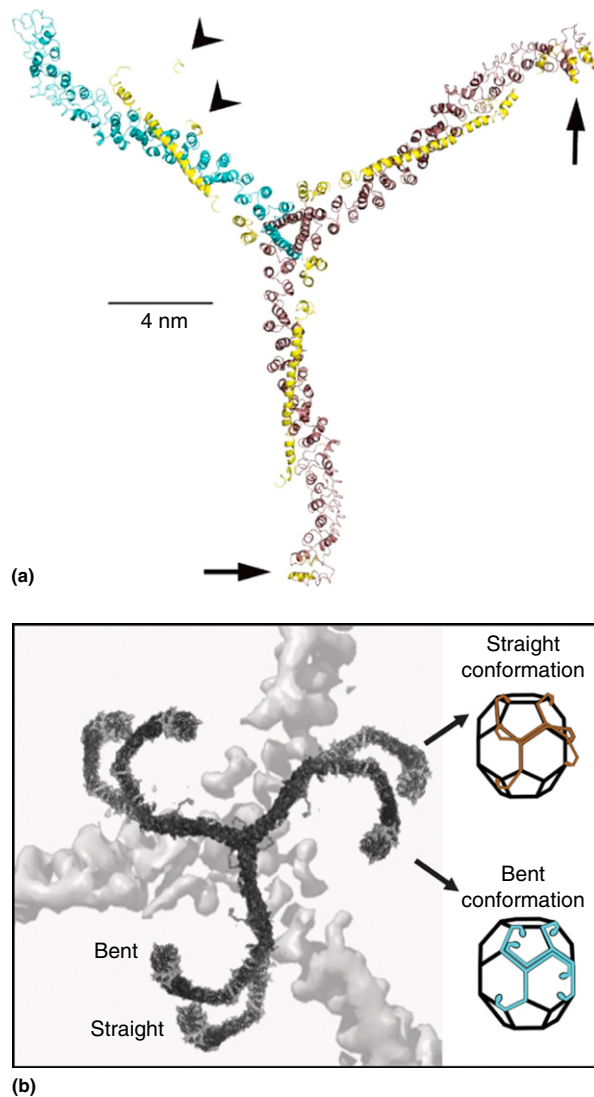


Figure 5 Low resolution structure (7.9 Å) of CLCb bound to CHC17 showing CLCb conformational changes and their effects on assembly. (a) Structure of CLCb bound to the triskelion hub (residues 1074–1675) (Wilbur *et al.*, 2010). The CLCb density is shown as α -helices (yellow). The established structure of CHC17 (Ybe *et al.*, 1999), as fitted to crystallographic density is shown. The brown CHC proximal legs have extended (straight) knees and CLCb density is associated with these knees (arrows). The blue CHC proximal leg has a bent knee, permitted by CLCb bound in the second, retracted conformation (arrowheads). (b) Models of triskelia with all straight or all bent knees are shown. Projection onto clathrin lattices at the right shows that the knee must bend to allow for assembly, facilitated by a conformational change in the CLCb. Both panels are reproduced with permission from Wilbur, J., Hwang, P.K., Ybe, J.A., *et al.*, 2010. Conformation switching of clathrin light chain regulates clathrin lattice assembly. *Developmental Cell* 18, 841–848.

CHC17 can form a lattice in the absence of the CLCb subunits, with key assembly interactions occurring between the proximal and distal leg domains (Wakeham *et al.*, 2003; Bocking *et al.*, 2014). CLCb interaction with the triskelion knee is sensitive to pH changes and can regulate bending, providing a

'brake' on spontaneous self-assembly of CHC17 into a lattice, by extending the triskelion knee (Wilbur *et al.*, 2010; Figure 5). In cells, assembly is promoted by CHC17 interaction with adaptors localized on membranes and is energetically favorable.

During CHC17 disassembly, a disordered C-terminal segment is recognized by the Hsc70 ATPase (Rapoport *et al.*, 2008) and its J-domain-containing cofactor auxilin (Unge-wickell *et al.*, 1995). ATP hydrolysis by Hsc70 is required for clathrin disassembly under biological conditions (Merrifield and Kaksonen, 2014). In spite of available cryo-electron microscopy structures of auxilin fragments and/or Hsc70 bound to CHC17 lattices (Fotin *et al.*, 2004a; Young *et al.*, 2013) and analysis of uncoating kinetics (Rothnie *et al.*, 2011; Bocking *et al.*, 2014), there is still debate about the exact mechanism by which disassembly is triggered by the Hsc70–auxilin complex. Concomitant with clathrin lattice disassembly, disruption of adaptor interaction with their recruiting phospholipids by lipid phosphatases (e.g., synaptojanin, endophilin, and OCRL) contributes to vesicle uncoating (Posor *et al.*, 2014).

It is notable that purified protein or recombinant fragments of CHC17 will self-assemble into a lattice (Greene *et al.*, 2000). Early biochemical studies of purified clathrin assembly described a variety of ionic and nonionic conditions that influence *in vitro* assembly of clathrin (Keen *et al.*, 1979; Nandi and Edelhofer, 1984), an interest that has reemerged with consideration of clathrin's potential as a biological nanomaterial (Hom *et al.*, 2012). CHC22 has a triskelion shape, similar to CHC17 and assembles into larger structures, yet to be characterized. CHC22 is missing the peptide sequence recognized by Hsc70, so will have a different disassembly mechanism.

Cellular Function and Regulation of Clathrin

Clathrin plays a sorting and transport function in membrane traffic for all eukaryotic cells. Its roles in vertebrate, and particularly mammalian cells, are more clearly defined than for yeast and plants. Here the mammalian roles, which correspond to similar pathways observed in *Drosophila* and nematodes, are described first. Then the known clathrin-mediated membrane traffic pathways in yeast and plants will be addressed, followed by discussion of clathrin function independent of membrane traffic.

PM Roles

Receptor-mediated endocytosis

The classic function ascribed to clathrin is to facilitate internalization of ligands bound to integral membrane receptors by receptor-mediated endocytosis (RME) (Figure 1). This function was postulated by Roth and Porter (1964) from observation of clathrin-coated pits (depressions) and vesicles during the uptake of yolk protein in mosquito oocytes (Roth and Porter, 1964). Clathrin's role in RME has since been functionally verified by many approaches. This role is restricted to CHC17. CHC22 clathrin does not naturally interact with the AP2 adaptor, when CHC17 is present (Liu *et al.*, 2001;

Esk *et al.*, 2010). The process of clathrin lattice formation at the PM involves cooperative interactions between receptors and the AP2 adaptor to nucleate clathrin assembly (Cocucci *et al.*, 2012). AP2 is recruited to the PM by specific recognition of phosphatidylinositol (4,5)-bisphosphate (PIP2). PIP2 binding to AP2 causes a conformational change that enhances cargo interaction, as does AP2 phosphorylation (Canagarajah *et al.*, 2013). Receptor (cargo) interaction with AP2 adaptors then induces a conformational change that reveals their clathrin-binding sites (Kelly *et al.*, 2014). Purified CHC17 bound by CLCs does not self-assemble at physiological pH (Ungewickell and Ungewickell, 1991; Liu *et al.*, 1995). Adaptor interaction with CHC17 overcomes this inhibition, presumably by alignment of triskelion legs such that the CLCs are dislodged from their inhibitory position at the knee (Greene *et al.*, 2000). Exposure of clathrin-binding sites upon receptor engagement by adaptors thereby directs clathrin assembly at sites where cargo is present, and presence of the CLCs prevents assembly of empty clathrin cages in the cytosol. For some signaling receptors, ligand binding enhances adaptor binding through receptor phosphorylation, such that RME is inefficient in the absence of ligand binding. Other receptors constitutively interact with AP2 adaptors. Early studies showed that adaptors in the cytosol that are not clathrin-associated are phosphorylated (Wilde and Brodsky, 1996), suggesting an additional regulatory mechanism for preventing nonproductive clathrin-adaptor interactions in the cytosol.

The minimal interaction of receptor domains with AP2 and clathrin at the surface of a PIP2-containing liposome is sufficient for the formation of a clathrin-coated membrane bud (Kelly *et al.*, 2014). In cells, other clathrin-associated proteins may contribute to membrane curvature, such as BAR-domain containing proteins, which are naturally curved and associate with invaginating membranes (Merrifield and Kaksonen, 2014; McMahon and Boucrot, 2011). Successful detachment of a CCV from the PM depends on self-assembly of dynamin GTPases into a protein 'collar' at the site of bud extrusion (Schmid and Frolov, 2011; Ferguson and De Camilli, 2012). Dynamins also contribute regulatory functions at earlier stages of CCV formation (Schmid and Frolov, 2011).

RME of transferrin receptor, low-density lipoprotein receptor and epidermal growth factor receptor can proceed in the absence of CLCs (Huang *et al.*, 2004; Poupon *et al.*, 2008). However, RME of certain G-protein-coupled receptors (GPCRs) requires CLCs (Ferreira *et al.*, 2012). Given that CLCs contribute to lattice rigidity (Dannhauser *et al.*, 2015), they may be required for RME of cargo whose properties increase resistance to membrane bending (Stachowiak *et al.*, 2013). Analogously, the Sec13 subunit of the COPII coat, which enhances coat rigidity, is required for GPI-linked cargo, but dispensable for other vesicle cargo (Copic *et al.*, 2012).

Alternatively, CLCs could be required for RME of selected cargo because some cargo alter membrane properties such that CLC recruitment of actin is needed for effective budding (Brodsky, 2012). RME of selected cargo in some mammalian tissue culture cells can occur in the absence of actin. However, morphological analysis shows that actin is present at the necks of the majority of coated pits in cultured mouse melanoma cells (Collins *et al.*, 2011). RME at the apical surface of polarized cells requires CLCs but RME of the same cargo at the

baso-lateral surface does not (Boulant *et al.*, 2011). In the case of apical endocytosis, the interaction of CLCs with Hip family proteins is required, implicating the CLC–Hip–actin connection in this process. Additionally, the CLC requirement can be reproduced by experimentally increasing membrane surface tension (Boulant *et al.*, 2011). Thus, the absolute requirement for actin recruitment by CLCs during mammalian RME is variable, depending on the force needed for vesicle detachment from the PM. However, it may be that actin participates in budding of most CCVs, whether it is needed or not (Collins *et al.*, 2011).

Particle uptake and domain organization

The size of CCVs can range from 60 to 200 nm in diameter (Pearse and Crowther, 1987). At the neuronal synapse, where clathrin is involved in recapture of synaptic vesicle proteins, CCV size is restricted to the smaller dimensions in this range by the neuronal assembly protein AP180 (Morris *et al.*, 1993). In several instances, clathrin is involved in the uptake of particles and microbes that are larger than 200 nm. In these cases, clathrin assembles at the interface between the microbe and actin. For uptake of *Listeria* bacteria, the CLC–Hip interaction, as well as CHC17 phosphorylation is required at internalization site (Bonazzi *et al.*, 2011). In the extreme instance of enteropathic *Escherichia coli*–host cell contact, clathrin is recruited at the bacterial contact interface to nucleate actin polymerization and pedestal formation. This latter process involves CHC17 phosphorylation and CLC–Hip interaction but no internalization, utilizing polymerized clathrin as a template for actin recruitment (Bonazzi *et al.*, 2011). A similar process accompanies formation of adherens junctions between mammalian cells (Bonazzi *et al.*, 2012). Clathrin-AP2 assemblies are also involved in stabilizing signaling domains (signalosomes) for the developmentally important Wnt complex (Kim *et al.*, 2013), in which the clathrin lattice appears to play an organizational role, as opposed to cargo capture for vesicle formation.

Endosome Roles

Two functions for clathrin on endosomes have been established. Cargo can recycle from sorting endosomes to the PM by many pathways. One of these pathways is clathrin-mediated and involves structures called G (gyrating)-clathrin, which are observed as highly motile clathrin-coated structures in the cell periphery (Zhao and Keen, 2008). G-clathrin mediates rapid recycling of internalized transferrin receptor and is not detected in the absence of CLCs (Majeed *et al.*, 2014). CLCs are also needed for recycling internalized β 1-integrin to the PM and for cell migration, implicating G-clathrin in directed recycling of cargo from endosomes to the leading edge of the PM. G-clathrin utilizes the GGA1 clathrin adaptor. The small GTPases ARF1 and ARF6 are involved in G-clathrin recruitment (Luo *et al.*, 2013).

Patches of clathrin-associated with the ESCRT0 component (clathrin adaptor Hrs) are observed on late endosomes during formation of multivesicular bodies. These patches, which have an unusual double-layered morphology by electron microscopy, are involved in concentrating ubiquitinated cargo for

efficient transfer to the ESCRT1-3 degradation machinery (Raiborg *et al.*, 2002).

Organelle Biogenesis

The clathrin-mediated pathways that generate lysosomes, secretory granules, dense core granules, and the storage compartment for the GLUT4 glucose transporter arise from the TGN of mammalian cells. These pathways are nucleated by the AP1 and GGA adaptors, and some involve retrograde transport from late endosomes as well as anterograde transport from the secretory pathway. The AP1 and GGA adaptors are recruited to membranes containing PI(4)P via interaction with ARF1. In the case of AP1, ARF1 binding causes a conformational change that facilitates cargo interaction, similar to that observed upon AP2-PI(4,5)P₂ interaction (Canagarajah *et al.*, 2013). Whether the GGAs function independently or in conjunction with AP1 to direct intracellular clathrin-mediated cargo traffic is not yet established.

Formation of lysosomes, secretory granules, and dense core vesicles

The majority of synthesized lysosomal enzymes are modified with mannose-6-phosphate, which is recognized by two specific receptors in mammalian cells (cation-dependent and cation-independent). These receptors are cargo for CHC17 CCVs at the TGN, which deliver the modified lysosomal enzymes to late endosomes, where they dissociate to remain in the forming lysosome, while the receptor is recycled by clathrin back to the TGN (Braulke and Bonifacino, 2009). Clathrin's role in the maturation of granules containing hormones for regulated secretion, such as insulin, is not active delivery of contents but removal of non-granule membrane components from the maturing granule (Klumperman *et al.*, 1998; Molinete *et al.*, 2001). With regard to formation of dense core vesicles containing neurotransmitters, CHC17 clathrin recruited by the AP3 adaptor may play a similar removal role in maturation. However, the self-assembling Vps41 coat protein (CHCR-containing) interacting with AP3 has been implicated in active sorting of components to the dense core granule (Asensio *et al.*, 2013).

Formation of the GLUT4 storage compartment (CHC22 clathrin)

GLUT4 is the insulin-responsive glucose transporter responsible for postprandial glucose clearance. In muscle and adipose cells, GLUT4 is packaged in vesicles that are released upon stimulation of the insulin receptor and, in the case of muscle, also in response to exercise (Bryant *et al.*, 2002). The formation of intracellular GLUT4 storage vesicles provides a reservoir of transporters that, when delivered to the PM, can generate a rapid metabolic response resulting in tissue uptake of glucose. From well-studied murine adipocyte and rat myoblast models, it has been established that CHC17 clathrin plays a role in formation of the GLUT4 storage compartment through membrane traffic from late endosomes and the TGN (Bogan, 2012). In humans, this pathway is mediated by CHC22 clathrin instead of CHC17 (Vassilopoulos *et al.*, 2009). CHC17 in human GLUT4 membrane traffic controls only endocytosis

and not intracellular sequestration. CHC22 sorting functions have been mapped to retrograde transport from late endosomes to the TGN (Esk *et al.*, 2010) and CHC22 expression levels are highest in human muscle, where this pathway contributes to GLUT4 sequestration. In muscle of insulin-resistant type 2 diabetic patients, CHC22 accumulates at the expanded GLUT4 storage compartment that results from an impaired insulin response (Vassilopoulos *et al.*, 2009). The different biochemical properties of CHC22 clathrin compared to CHC17 define specialized features of human glucose metabolism.

Clathrin Roles in Yeast and Plants

Clathrin-mediated endocytosis in yeast requires CLC and actin, as well as the Hip family protein Sla2p (Boettner *et al.*, 2012). One explanation for this requirement is that the turgor pressure of yeast cells generates an outward force on the PM that resists endocytosis, requiring CLC and actin to provide strength to the clathrin lattice (Aghamohammadzadeh and Ayscough, 2009). A known cargo for clathrin-mediated endocytosis in yeast is the pheromone receptor Ste2p, which is a GPCR, and might require CLC participation for the same reasons as needed for uptake of some mammalian GPCRs. Clathrin in yeast is also responsible for retention of proteins in the TGN, a function established by analysis of one of the first CHC mutants (Payne and Schekman, 1989). The cargo identified was the Kex2p endoprotease, needed for processing the alpha factor mating pheromone.

Arabidopsis plants have two CHC proteins and three CLCs (Baisa *et al.*, 2013). Clathrin plays a key role in maintaining the polarity of plant cells. The polarized localization of the PIN efflux carriers for auxin depends on their clathrin-mediated endocytosis and targeted recycling to the correctly polarized domain (Kania *et al.*, 2014). Although plants perform endocytosis in the context of a cell wall, the role of actin in their clathrin-mediated endocytosis is not clearly established (Baisa *et al.*, 2013).

Non-membrane Traffic Roles for Mammalian CHC17 Clathrin

Clathrin is observed associated with the mammalian mitotic spindle (Okamoto *et al.*, 2000). While initially thought to play a role in membrane traffic to and from the spindle-associated membrane-containing matrix, it is now established that clathrin's role at the spindle is independent of a membrane traffic function for clathrin (Royle, 2012). During mitosis, phosphorylation of the TACC3 protein by Aurora kinase A creates a binding site that interacts with the CHC17 terminal domain and incorporates clathrin (with associated CLCs) into a complex with the TACC3 and chTOG proteins. This protein complex binds kinetochore (K-fiber) microtubules in the mitotic spindle and, due to the trimeric nature of clathrin, stabilizes these bundled microtubules by cross-linking them. Depletion of clathrin thereby induces chromosomal dysjunction and mitotic defects. Clathrin's role at the spindle can be restored using a mutant truncated clathrin that contains the trimerized terminal domain, but is missing the proximal leg domains needed for clathrin lattice formation (Booth *et al.*, 2011).

Thus, clathrin functions as a stabilizing cross-linker in this instance. Early in mitosis the TACC3–chTOG–clathrin complex also plays a role in centrosome integrity, so that clathrin depletion induces fragmented centrosomes and formation of multipolar spindles (Foraker *et al.*, 2012). It is not yet clear how conserved is the participation of clathrin in a microtubule-stabilizing function. Clathrin has been observed at the mitotic spindle in vertebrates, *Drosophila*, and plants, but its role in yeast mitosis is yet to be clarified. It has been proposed that participation in the TACC3–chTOG complex is an evolutionarily recent function for clathrin compared to its role in membrane traffic (Royle, 2012). The discovery of this novel function for clathrin that arises from its cross-linking as opposed to polymerization capacity suggests that other such specialized roles for clathrin may be uncovered in the future.

Conclusions

1. The triskelion shape and ability of CHCs to polymerize into a lattice are the key features that enable the protein to perform its function as a vesicle coat that mediates cargo selection for transport.
2. The CLCs contribute stability and strength to the clathrin lattice through interaction with the central portion of the triskelion and through binding the Hip family proteins that connect clathrin to the actin cytoskeleton.
3. The properties conferred by CLCs expand the type and size of cargo that clathrin-coated membranes can transport and allow clathrin to play an actin-organizing function during microbial infection and cell–cell contact.
4. CHC22 clathrin, the second isoform present in humans but missing from mice, plays a key role in targeting the GLUT4 glucose transporter to its insulin-responsive intracellular storage compartment, defining a novel aspect of human glucose metabolism.
5. Participation of CHC17 clathrin (conventional isoform) in a microtubule-binding complex during mitosis contributes to stability of centrosomes and the mitotic spindle, exploiting clathrin's cross-linking properties and independent of its membrane traffic role.

See also: Interorganellar Communication: Interplay and Processes: Clathrin Independent Endocytosis; Rabs of the Endosomal Recycling Pathway

References

Acton, S.L., Brodsky, F.M., 1990. Predominance of clathrin light chain LCb correlates with the presence of a regulated secretory pathway. *Journal of Cell Biology* 111, 1419–1426.

Aghamohammadzadeh, S., Ayscough, K.R., 2009. Differential requirements for actin during yeast and mammalian endocytosis. *Nature Cell Biology* 11, 1039–1042.

Asensio, C.S., Sirkis, D.W., Maas Jr., J.W., *et al.*, 2013. Self-assembly of VPS41 promotes sorting required for biogenesis of the regulated secretory pathway. *Developmental Cell* 27, 425–437.

Baisa, G.A., Mayers, J.R., Bednarek, S.Y., 2013. Budding and braking news about clathrin-mediated endocytosis. *Current Opinion in Plant Biology* 16, 718–725.

Bocking, T., Aguet, F., Rapoport, I., *et al.*, 2014. Key interactions for clathrin coat stability. *Structure* 22, 819–829.

Boettner, D.R., Chi, R.J., Lemmon, S.K., 2012. Lessons from yeast for clathrin-mediated endocytosis. *Nature Cell Biology* 14, 2–10.

Bogan, J.S., 2012. Regulation of glucose transporter translocation in health and diabetes. *Annual Review of Biochemistry* 81, 507–532.

Bonazzi, M., Kuhbacher, A., Toledo-Arana, A., *et al.*, 2012. A common clathrin-mediated machinery co-ordinates cell-cell adhesion and bacterial internalization. *Traffic* 13, 1653–1666.

Bonazzi, M., Vasudevan, L., Mallet, A., *et al.*, 2011. Clathrin phosphorylation is required for actin recruitment at sites of bacterial adhesion and internalization. *Journal of Cell Biology* 195, 525–536.

Booth, D.G., Hood, F.E., Prior, I.A., Royle, S.J., 2011. A TACC3/ch-TOG/clathrin complex stabilises kinetochore fibres by inter-microtubule bridging. *EMBO Journal* 30, 906–919.

Boulant, S., Kural, C., Zeeh, J.C., Ubelmann, F., Kirchhausen, T., 2011. Actin dynamics counteract membrane tension during clathrin-mediated endocytosis. *Nature Cell Biology* 13, 1124–1131.

Braulke, T., Bonifacino, J.S., 2009. Sorting of lysosomal proteins. *Biochimica et Biophysica Acta* 1793, 605–614.

Brodsky, F.M., 2012. Diversity of clathrin function: New tricks for an old protein. *Annual Review of Cell and Developmental Biology* 28, 309–336.

Brodsky, F.M., Galloway, C.J., Blank, G.S., *et al.*, 1987. Localization of clathrin light-chain sequences mediating heavy-chain binding and coated vesicle diversity. *Nature* 326, 203–205.

Bryant, N.J., Govers, R., James, D.E., 2002. Regulated transport of the glucose transporter GLUT4. *Nature Reviews Molecular Cell Biology* 3, 267–277.

Canagarajah, B.J., Ren, X., Bonifacino, J.S., Hurley, J.H., 2013. The clathrin adaptor complexes as a paradigm for membrane-associated allostery. *Protein Science* 22, 517–529.

Chen, C.Y., Brodsky, F.M., 2005. Huntingtin-interacting protein 1 (Hip1) and Hip1-related protein (Hip1R) bind the conserved sequence of clathrin light chains and thereby influence clathrin assembly *in vitro* and actin distribution *in vivo*. *Journal of Biological Chemistry* 280, 6109–6117.

Cocucci, E., Aguet, F., Boulant, S., Kirchhausen, T., 2012. The first five seconds in the life of a clathrin-coated pit. *Cell* 150, 495–507.

Collins, A., Warrington, A., Taylor, K.A., Svitkina, T., 2011. Structural organization of the actin cytoskeleton at sites of clathrin-mediated endocytosis. *Current Biology* 21, 1167–1175.

Copic, A., Latham, C.F., Horlbeck, M.A., D'Arcangelo, J.G., Miller, E.A., 2012. ER cargo properties specify a requirement for COPII coat rigidity mediated by Sec13p. *Science* 335, 1359–1362.

Crowther, R.A., Pearse, B.M.F., 1981. Assembly and packing of clathrin into coats. *Journal of Cell Biology* 91, 790–797.

Dannhauser, P.N., Platen, M., Boning, H., *et al.*, 2015. Effect of clathrin light chains on the stiffness of clathrin lattices and membrane budding. *Traffic* (in press).

DeLuca-Flaherty, C., McKay, D.B., Parham, P., Hill, B.L., 1990. Uncoating protein (hsc70) binds a conformationally labile domain of clathrin light chain LC_a to stimulate ATP hydrolysis. *Cell* 62, 875–887.

Esk, C., Chen, C.Y., Johannes, L., Brodsky, F.M., 2010. The clathrin heavy chain isoform CHC22 functions in a novel endosomal sorting step. *Journal of Cell Biology* 188, 131–144.

Ferguson, M.L., Prasad, K., Boukari, H., *et al.*, 2008. Clathrin triskelia show evidence of molecular flexibility. *Biophysical Journal* 95, 1945–1955.

Ferguson, S.M., De Camilli, P., 2012. Dynamin, a membrane-remodelling GTPase. *Nature Reviews Molecular Cell Biology* 13, 75–88.

Ferreira, F., Foley, M., Cooke, A., *et al.*, 2012. Endocytosis of G protein-coupled receptors is regulated by clathrin light chain phosphorylation. *Current Biology* 22, 1361–1370.

Foraker, A.B., Camus, S.M., Evans, T.M., *et al.*, 2012. Clathrin promotes centrosome integrity in early mitosis through stabilization of centrosomal ch-TOG. *Journal of Cell Biology* 198, 591–605.

Fotin, A., Cheng, Y., Grigorieff, N., *et al.*, 2004a. Structure of an auxilin-bound clathrin coat and its implications for the mechanism of uncoating. *Nature* 432, 649–653.

Fotin, A., Cheng, Y., Sliz, P., *et al.*, 2004b. Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* 432, 573–579.

Giudice, J., Xia, Z., Wang, E.T., *et al.*, 2014. Alternative splicing regulates vesicular trafficking genes in cardiomyocytes during postnatal heart development. *Nature Communications* 5, 3603.

Gottfried, I., Ehrlich, M., Ashery, U., 2010. The Sla2p/HIP1/HIP1R family: Similar structure, similar function in endocytosis? *Biochemical Society Transactions* 38, 187–191.

- Greene, B., Liu, S.-H., Wilde, A., Brodsky, F.M., 2000. Complete reconstitution of clathrin basket formation with recombinant protein fragments: Adaptor control of clathrin self-assembly. *Traffic* 1, 69–75.
- ter Haar, E., Musacchio, A., Harrison, S.C., Kirchhausen, T., 1998. Atomic structure of clathrin: A β propeller terminal domain joins an α zigzag linker. *Cell* 95, 563–573.
- Heuser, J.E., Keen, J.H., Amende, L.M., Lippoldt, R.E., Prasad, K., 1987. Deep-etch visualization of 27S clathrin: A tetrahedral tetramer. *Journal of Cell Biology* 105, 1999–2009.
- Hill, B.L., Drickamer, K., Brodsky, F.M., Parham, P., 1988. Identification of the phosphorylation sites of clathrin light chain LC_b. *Journal of Biological Chemistry* 263, 5499–5501.
- Hom, N., Mehta, K.R., Chou, T., *et al.*, 2012. Anisotropic nanocrystal arrays organized on protein lattices formed by recombinant clathrin fragments. *Journal of Materials Chemistry* 22, 23335–23339.
- Huang, F., Khvorova, A., Marshall, W., Sorkin, A., 2004. Analysis of clathrin-mediated endocytosis of epidermal growth factor receptor by RNA interference. *Journal of Biological Chemistry* 279, 16657–16661.
- Jackson, A.P., Seow, H.-F., Holmes, N., Drickamer, K., Parham, P., 1987. Clathrin light chains contain brain-specific insertion sequences and a region of homology with intermediate filaments. *Nature* 326, 154–159.
- Kania, U., Fendrych, M., Friml, J., 2014. Polar delivery in plants: commonalities and differences to animal epithelial cells. *Open Biology* 4, 140017.
- Keen, J.H., Willingham, M.C., Pastan, I.H., 1979. Clathrin-coated vesicles: Isolation, dissociation, and factor-dependent reassociation of clathrin baskets. *Cell* 16, 303–312.
- Kelly, B.T., Graham, S.C., Liska, N., *et al.*, 2014. Clathrin adaptors. AP2 controls clathrin polymerization with a membrane-activated switch. *Science* 345, 459–463.
- Kim, I., Pan, W., Jones, S.A., *et al.*, 2013. Clathrin and AP2 are required for PtdIns (4,5)P₂-mediated formation of LRP6 signalosomes. *Journal of Cell Biology* 200, 419–428.
- Kirchhausen, T., Harrison, S.C., 1981. Protein organization in clathrin trimers. *Cell* 23, 755–761.
- Kirchhausen, T., Scarmato, P., Harrison, S.C., *et al.*, 1987. Clathrin light chains LC_a and LC_b are similar, polymorphic, and share repeated heptad motifs. *Science* 236, 320–324.
- von Kleist, L., Stahlschmidt, W., Bulut, H., *et al.*, 2011. Role of the clathrin terminal domain in regulating coated pit dynamics revealed by small molecule inhibition. *Cell* 146, 471–484.
- Klumpperman, J., Kuliawat, R., Griffith, J.M., Geuze, H.J., Arvan, P., 1998. Mannose 6-phosphate receptors are sorted from immature secretory granules via adaptor protein AP-1, clathrin, and syntaxin 6-positive vesicles. *Journal of Cell Biology* 141, 359–371.
- Legendre-Guillemin, V., Metzler, M., Lemaire, J.F., *et al.*, 2005. Huntingtin interacting protein 1 (HIP1) regulates clathrin assembly through direct binding to the regulatory region of the clathrin light chain. *Journal of Biological Chemistry* 280, 6101–6108.
- Lemmon, S.K., Traub, L.M., 2012. Getting in touch with the clathrin terminal domain. *Traffic* 13 (4), 511–519.
- Liu, S.H., Towler, M.C., Chen, E., *et al.*, 2001. A novel clathrin homolog that co-distributes with cytoskeletal components functions in the trans-Golgi network. *EMBO Journal* 20, 272–284.
- Liu, S.H., Wong, M.L., Craik, C.S., Brodsky, F.M., 1995. Regulation of clathrin assembly and trimerization defined using recombinant triskelion hubs. *Cell* 83, 257–267.
- Luo, Y., Zhan, Y., Keen, J.H., 2013. Arf6 regulation of gyrating-clathrin. *Traffic* 14, 97–106.
- Majeed, S.R., Vasudevan, L., Chen, C.Y., *et al.*, 2014. Clathrin light chains are required for the gyrating-clathrin recycling pathway and thereby promote cell migration. *Nature Communications* 5, 3891.
- McMahon, H.T., Boucrot, E., 2011. Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nature Reviews Molecular Cell Biology* 12, 517–533.
- Merrifield, C.J., Kaksonen, M., 2014. Endocytic accessory factors and regulation of clathrin-mediated endocytosis. *Cold Spring Harbor Perspectives in Biology* 6 (11), a016733.
- Molinet, M., Dupuis, S., Brodsky, F.M., Halban, P.A., 2001. Role of clathrin in the regulated secretory pathway of pancreatic beta-cells. *Journal of Cell Science* 114, 3059–3066.
- Morris, S.A., Schröder, S., Plessmann, U., Weber, K., Ungewickell, E., 1993. Clathrin assembly protein AP180: Primary structure, domain organization and identification of a clathrin binding site. *EMBO Journal* 12, 667–675.
- Nandi, P.K., Edelhofer, H., 1984. The effects of lyotropic (hofmeister) salts on the stability of clathrin coat structure in coated vesicles and baskets. *Journal of Biological Chemistry* 259, 11290–11296.
- Näthke, I., Hill, B.L., Parham, P., Brodsky, F.M., 1990. The calcium binding site of clathrin light chains. *Journal of Biological Chemistry* 265, 18621–18627.
- Okamoto, C.T., McKinney, J., Jeng, Y.Y., 2000. Clathrin in mitotic spindles. *American Journal of Physiology – Cell Physiology* 279, C369–C374.
- Parham, P., Brodsky, F.M., Drickamer, K., 1989. The occurrence of disulfide bonds in purified clathrin light chains. *Biochemical Journal* 257, 775–781.
- Payne, G.S., Schekman, R., 1989. Clathrin: A role in the intracellular retention of a Golgi membrane protein. *Science* 245, 1358–1365.
- Pearse, B.M.F., 1975. Coated vesicles from pig brain: Purification and biochemical characterization. *Journal of Molecular Biology* 97, 93–98.
- Pearse, B.M.F., Crowther, R.A., 1987. Structure and assembly of coated vesicles. *Annual Review of Biophysics and Biophysical Chemistry* 16, 49–68.
- Pley, U.M., Hill, B.L., Alibert, C., Brodsky, F.M., Parham, P., 1995. The interaction of calmodulin with clathrin-coated vesicles, triskelions, and light chains. Localization of a binding site. *Journal of Biological Chemistry* 270, 2395–2402.
- Posor, Y., Eichhorn-Grunig, M., Hauke, V., 2014. Phosphoinositides in endocytosis. *Biochimica et Biophysica Acta* 14, 00192–00199.
- Poupon, V., Girard, M., Legendre-Guillemin, V., *et al.*, 2008. Clathrin light chains function in mannose phosphate receptor trafficking via regulation of actin assembly. *Proceedings of the National Academy of Sciences of the United States of America* 105, 168–173.
- Raiborg, C., Bache, K.G., Gillooly, D.J., *et al.*, 2002. Hrs sorts ubiquitinated proteins into clathrin-coated microdomains of early endosomes. *Nature Cell Biology* 4, 394–398.
- Rapoport, I., Boll, W., Yu, A., Bocking, T., Kirchhausen, T., 2008. A motif in the clathrin heavy chain required for the hsc70/auxilin uncoating reaction. *Molecular Biology of the Cell* 19, 405–413.
- Roth, T.F., Porter, K.R., 1964. Yolk protein uptake in the oocyte of the mosquito *Aedes aegypti*. *L. Journal of Cell Biology* 20, 313–332.
- Rothnie, A., Clarke, A.R., Kuzmic, P., Cameron, A., Smith, C.J., 2011. A sequential mechanism for clathrin cage disassembly by 70-kDa heat-shock cognate protein (Hsc70) and auxilin. *Proceedings of the National Academy of Sciences of the United States of America* 108, 6927–6932.
- Royle, S.J., 2012. The role of clathrin in mitotic spindle organisation. *Journal of Cell Science* 125, 19–28.
- Schmid, S.L., Frolov, V.A., 2011. Dynamin: Functional design of a membrane fission catalyst. *Annual Review of Cell and Developmental Biology* 27, 79–105.
- Schreij, A.M., Chaineau, M., Ruan, W., *et al.*, 2015. LRRK2 localizes to endosomes and interacts with clathrin-light chains to limit Rac1 activation. *EMBO Reports* 16 (1), 79–86. doi:10.15252/embr.201438714.
- Stachowiak, J.C., Brodsky, F.M., Miller, E.A., 2013. A cost-benefit analysis of the physical mechanisms of membrane curvature. *Nature Cell Biology* 15, 1019–1027.
- Ungewickell, E., Branton, D., 1981. Assembly units of clathrin coats. *Nature* 289, 420–422.
- Ungewickell, E., Ungewickell, H., 1991. Bovine brain clathrin light chains impede heavy chain assembly *in vitro*. *Journal of Biological Chemistry* 266, 12710–12714.
- Ungewickell, E., Ungewickell, H., Holstein, S.E.H., *et al.*, 1995. Role of auxilin in uncoating clathrin-coated vesicles. *Nature* 378, 632–635.
- Vassilopoulos, S., Esk, C., Hoshino, S., *et al.*, 2009. A role for the CHC22 clathrin heavy-chain isoform in human glucose metabolism. *Science* 324, 1192–1196.
- Wakeham, D.E., Abi-Rached, L., Towler, M.C., *et al.*, 2005. Clathrin heavy and light chain isoforms originated by independent mechanisms of gene duplication during chordate evolution. *Proceedings of the National Academy of Sciences of the United States of America* 102, 7209–7214.
- Wakeham, D.E., Chen, C.Y., Greene, B., Hwang, P.K., Brodsky, F.M., 2003. Clathrin self-assembly involves coordinated weak interactions favorable for cellular regulation. *EMBO Journal* 22, 4980–4990.
- Wilbur, J., Hwang, P.K., Ybe, J.A., *et al.*, 2010. Conformation switching of clathrin light chain regulates clathrin lattice assembly. *Developmental Cell* 18, 841–848.
- Wilde, A., Brodsky, F.M., 1996. *In vivo* phosphorylation of adaptors regulates their interaction with clathrin. *Journal of Cell Biology* 135, 635–645.
- Ybe, J.A., Brodsky, F.M., Hofmann, K., *et al.*, 1999. Clathrin self-assembly is mediated by a tandemly repeated superhelix. *Nature* 399, 371–375.
- Ybe, J.A., Perez-Miller, S., Niu, Q., *et al.*, 2007. Light chain C-terminal region reinforces the stability of clathrin heavy chain trimers. *Traffic* 8, 1101–1110.
- Young, A., Stoilova-McPhie, S., Rothnie, A., *et al.*, 2013. Hsc70-induced changes in clathrin-auxilin cage structure suggest a role for clathrin light chains in cage disassembly. *Traffic* 14, 987–996.

Zhao, Y., Keen, J.H., 2008. Gyrating clathrin: Highly dynamic clathrin structures involved in rapid receptor recycling. *Traffic* 9, 2253–2264.

Further Reading

Baisa, G.A., Mayers, J.R., Bednarek, S.Y., 2013. Budding and braking news about clathrin-mediated endocytosis. *Current Opinion in Plant Biology* 16, 718–725.

Boettner, D.R., Chi, R.J., Lemmon, S.K., 2012. Lessons from yeast for clathrin-mediated endocytosis. *Nature Cell Biology* 14, 2–10.

Bogan, J.S., 2012. Regulation of glucose transporter translocation in health and diabetes. *Annual Review of Biochemistry* 81, 507–532.

Brodsky, F.M., 2012. Diversity of clathrin function: New tricks for an old protein. *Annual Review of Cell and Developmental Biology* 28, 309–336.

Canagarajah, B.J., Ren, X., Bonifacio, J.S., Hurley, J.H., 2013. The clathrin adaptor complexes as a paradigm for membrane-associated allostery. *Protein Science* 22, 517–529.

Ferguson, S.M., De Camilli, P., 2012. Dynamin, a membrane-remodelling GTPase. *Nature Reviews Molecular Cell Biology* 13, 75–88.

Gottfried, I., Ehrlich, M., Ashery, U., 2010. The Sla2p/HIP1/HIP1R family: Similar structure, similar function in endocytosis? *Biochemical Society Transactions* 38, 187–191.

Kania, U., Fendrych, M., Friml, J., 2014. Polar delivery in plants: commonalities and differences to animal epithelial cells. *Open Biology* 4, 140017.

Lemmon, S.K., Traub, L.M., 2012. Getting in touch with the Clathrin terminal domain. *Traffic* 13, 511–519.

McMahon, H.T., Boucrot, E., 2011. Molecular mechanism and physiological functions of clathrin-mediated endocytosis. *Nature Reviews Molecular Cell Biology* 12, 517–533.

Merrifield, C.J., Kaksonen, M., 2014. Endocytic accessory factors and regulation of clathrin-mediated endocytosis. *Cold Spring Harbor Perspectives in Biology* 6 (11), a016733.

Pearse, B.M.F., Crowther, R.A., 1987. Structure and assembly of coated vesicles. *Annual Review of Biophysics and Biophysical Chemistry* 16, 49–68.

Posor, Y., Eichhorn-Grunig, M., Haucke, V., 2014. Phosphoinositides in endocytosis. *Biochimica et Biophysica Acta* 14, 00192–00199.

Royle, S.J., 2012. The role of clathrin in mitotic spindle organisation. *Journal of Cell Science* 125, 19–28.

Schmid, S.L., Frolov, V.A., 2011. Dynamin: Functional design of a membrane fission catalyst. *Annual Review of Cell and Developmental Biology* 27, 79–105.

Stachowiak, J.C., Brodsky, F.M., Miller, E.A., 2013. A cost-benefit analysis of the physical mechanisms of membrane curvature. *Nature Cell Biology* 15, 1019–1027.

Clathrin Independent Endocytosis

RG Parton, The University of Queensland, Brisbane, QLD, Australia

© 2016 Elsevier Inc. All rights reserved.

Introduction

Clathrin-mediated endocytosis is one of the best-studied processes in mammalian cell biology and is now understood in submolecular detail. Endocytosis also occurs without clathrin and in recent years it has become apparent that clathrin-independent (CI) endocytic processes are of high capacity, have crucial physiological roles, and have been exploited by many pathogens. Here the major clathrin-independent pathways will be reviewed focusing on macropinocytosis, caveolar endocytosis, and the other major clathrin-independent endocytic pathways, particularly the CLIC/GEEC (clathrin-independent carrier/glycosylphosphatidylinositol (GPI)-anchored protein-enriched early endosomal compartment) pathway.

Macropinocytosis

Definition

Macropinocytosis is the uptake of large volumes of fluid through actin-dependent protrusions of the cell surface. Macropinocytosis has been defined operationally as a pathway sensitive to the drug, amiloride, which involves the formation of large (diameter greater than 0.2 μm) vacuoles in the region of plasma membrane ruffles (Kerr and Teasdale, 2009; Swanson, 2008; Figure 1). Constitutive macropinocytic activity appears to vary greatly between different cell types, with particularly high activity in dendritic cells (Norbury, 2006). In other cells macropinocytosis can be stimulated by growth factors and active oncogenes.

Molecules

Machinery: Macropinocytosis involves the generation of circular ruffles from actin-rich plasma membrane ruffles that close to engulf large volumes of fluid. The small GTPases, Rho, Rac, and Cdc42 work together to initiate actin polymerization (Mooren *et al.*, 2012). The importance of this step can be demonstrated experimentally by the sensitivity of macropinocytosis to disruption of actin. In addition, the macropinocytosis inhibitor, amiloride, inhibits Na^+H^+ exchange, which would normally counteract a submembrane drop in pH on growth factor stimulation. In the presence of amiloride the low pH inactivates the small GTPases required for actin rearrangement (Koivusalo *et al.*, 2010). Branching actin filaments required for formation of the macropinosome are formed through the action of Arp2/3. Other molecules linked to macropinosome formation include BAR domain-containing sorting nexins (Wang *et al.*, 2010), Rabankyrin5 (Schnatwinkel *et al.*, 2004), p21-activated kinase (PAK1), and CtBP1/BARS (Liberali *et al.*, 2008). PAK1 phosphorylates CtBP1/BARS which is essential for the final closure step (i.e., scission from the plasma membrane) of the macropinosome. Mercer and

Helenius (Mercer and Helenius, 2009, 2012) have suggested a macropinocytosis 'toolbox,' a set of inhibitors and molecular targets, that can be used to test whether uptake of a particular agent uses macropinocytosis.

Cargo: By definition macropinocytosis is a pathway for uptake of large volumes of nonspecific solutes. Macropinocytosis allows for efficient sampling of large volumes of extracellular media and rapid uptake of solutes destined for degradation. In addition to these physiological roles, a number of viruses and bacteria exploit this pathway, stimulating macropinocytosis as a means to enter the cell. These include pathogenic bacteria such as *Salmonella*, *Shigella*, and mycobacteria as well as viruses such as vaccinia virus, coxsackie virus, and adenovirus (Kerr and Teasdale, 2009; Mercer and Helenius, 2009, 2012). Stimulation of macropinocytosis is an advantageous strategy for the pathogen by causing its encapsulation within a macropinosome. Interestingly, macropinocytosis appears to also be an important pathway for uptake of many artificial agents, particularly those designed for therapeutic delivery to cells. Macropinocytosis is known to contribute to the uptake of cell-penetrating peptides such as transactivator of transcription (TAT) and to a number of nanoparticle delivery systems (Gilleron *et al.*, 2013).

Carriers

Macropinosomes form when surface ruffles fuse back with the plasma membrane to encapsulate a large volume of extracellular fluid and can be over 5 μm in diameter (Figure 2). Carrier formation occurs as ruffles, sheet-like extensions of the cell surface, form crater-like structures termed cups (Swanson, 2008). To form the macropinosome the cup must close through contraction of the distal region of the cup followed by membrane fission and fusion to generate the macropinosome (for review see Swanson, 2008). Although macropinosome formation is intimately linked to and dependent on membrane ruffle formation, macropinocytosis can be inhibited without affecting membrane ruffling (Gilleron *et al.*, 2013; West *et al.*, 2000) showing that macropinosome formation is a regulated process. The importance of phosphoinositide lipids in macropinocytosis has been known for some time. Inhibition of phosphoinositide 3-kinases inhibits macropinosome formation. More recently the sequential breakdown of $\text{PI}(3,4,5)\text{P}_3$ to PI via $\text{PI}(3,4)\text{P}_2$ and $\text{PI}(3)\text{P}$ was shown to be essential for completion of macropinosome formation (Maekawa *et al.*, 2014). One effector for $\text{PI}(3)\text{P}$ was identified as the calcium-activated potassium channel, KCa3.1 , which was also shown to be required for macropinocytosis. An interesting hypothesis for the function of the channel in macropinocytosis involves $\text{PI}(3)\text{P}$ activating the channel to increase K^+ efflux, decrease cytosolic solute concentration, induce cell shrinkage and stimulate macropinocytosis (Maekawa *et al.*, 2014).

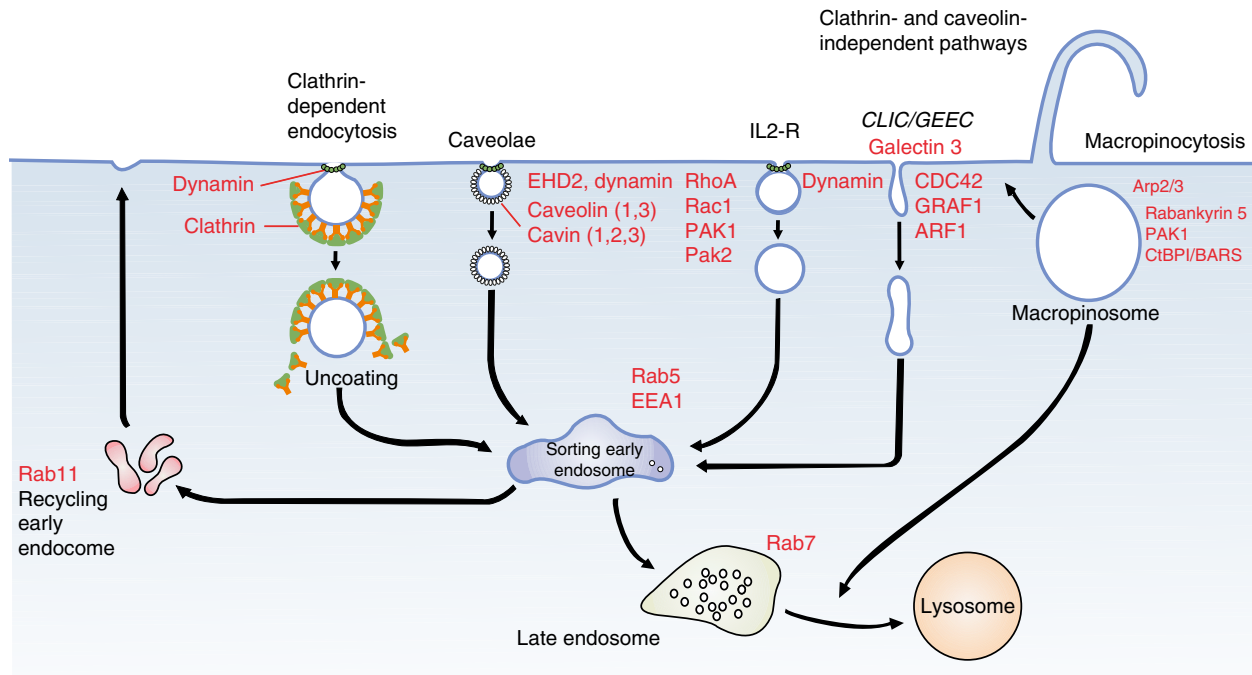


Figure 1 Clathrin-dependent and independent endocytic pathways. Scheme showing major pathways in mammalian cells together with some of the key molecules. Modified from Mayor, S., Pagano R.E., 2007. Pathways of clathrin-independent endocytosis. *Nature Reviews Molecular Cell Biology* 8 (8), 603–612.

Pathway

The majority of microscopic studies following macropinocytosis have shown that macropinosomes fuse with endosomal compartments and deliver their contents to lysosomes. This is consistent with the role of macropinocytic trafficking in the degradation of endocytosed proteins in the kidney and in immunological cells (Kerr and Teasdale, 2009). However, in other cell types separation from the endocytic pathway and refusion with the plasma membrane has been described (Hewlett *et al.*, 1994).

The trafficking pathways have been particularly well characterized in EGF-stimulated HEK293 cells. The large size of macropinosomes and the ability to stimulate formation in a synchronized manner has allowed a detailed quantitative description of their formation and maturation (Kerr *et al.*, 2006; Hamilton *et al.*, 2009). After a few minutes of stimulation macropinosomes have been formed and are associated with tubules coated in sorting nexins. These tubules rapidly depart from the forming macropinosome to fuse with more centrally located early endosomes in a microtubule-dependent fashion. Loss of these tubules, which have a high membrane area to volume ratio, causes a rapid loss of macropinosome surface area, without a major change in volume increasing the internal pressure in the macropinosome. This process is required for maturation of the macropinosome including acquisition of Rab7, a small GTPase associated with late endosomal trafficking. Concomitant with Rab7 acquisition are further changes in phosphoinositide composition with the transition from PI(3)P to PI(3,5)P₂ essential for fusion with late endosomes/lysosomes (Kerr *et al.*, 2010).

Quantitative Aspects and Dynamics

Macropinosomes can engulf a large amount of the plasma membrane as they form: a 5 μ m diameter macropinosome clearly removes a significant portion of the plasma membrane. Recycling from the macropinosome back to the plasma membrane therefore becomes an important issue and sorting nexin-labeled tubules have been estimated to remove over 80% of the limiting membrane of the macropinosome as it matures (Hamilton *et al.*, 2009). Detailed kinetics and quantitative parameters have been provided for macropinosome formation and maturation and correlated to specific endosomal markers in growth factor-treated cultured cells.

Functions

Macropinocytosis is a crucial feature of dendritic cells and macrophages. Engulfment of large volumes of the extracellular medium allows sampling of the immediate environment for foreign agents. Antigens taken up by macropinocytosis can be presented on both class I and class II major histocompatibility complex (MHC) molecules and so macropinocytosis has a crucial role in the immune system (Norbury, 2006; Sallusto *et al.*, 1995; Lim and Gleeson, 2011). Macropinocytosis has also been suggested to be a major pathway in the reabsorption of proteins from the glomerular filtrate by proximal tubule cells in the kidney based on studies of Rabankyrin5 (Schnatwinkel *et al.*, 2004). There is also increasing evidence for a role of macropinocytosis in cell migration (Gu *et al.*, 2011).

Macropinocytosis has been known for some time to be induced in cells transformed by oncogenes such as mutant

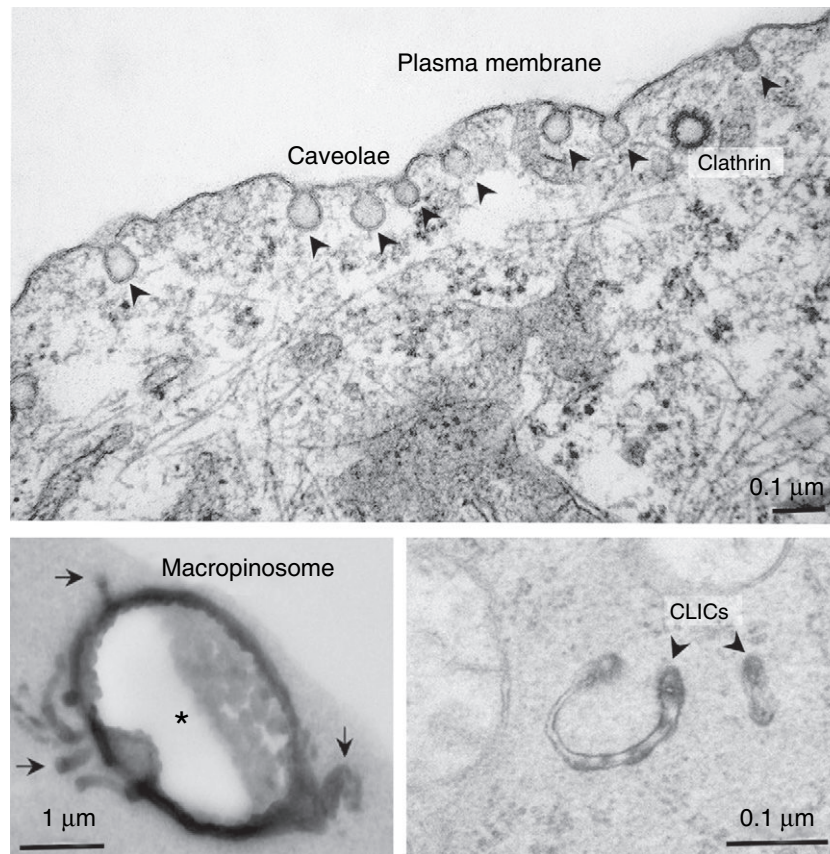


Figure 2 Morphology of caveolae, macropinosomes, and early carriers on the CLIC/GEEC pathway. Upper panel: caveolae (arrowheads) are abundant flask-shaped pits without an evident coat (compare with clathrin coated profile). Lower left panel: macropinosome (asterisk) labeled by HRP as visualized in a thick section. Adapted from [Figure 2\(c\)](#); Kerr, M.C., *et al.*, 2006. Visualisation of macropinosome maturation by the recruitment of sorting nexins. *Journal of Cell Science* 119 (Pt 19), 3967–3980. Lower right panel: CLICs labeled with an internalized toxin-peroxidase construct. Adapted from part of [Figure 1\(g\)](#); Howes, M.T., *et al.*, 2010. Clathrin-independent carriers form a high capacity endocytic sorting system at the leading edge of migrating cells. *Journal of Cell Biology* 190 (4), 675–691.

forms of K-Ras ([Bar-Sagi and Feramisco, 1986](#)). Transformed cells change their metabolic properties and become dependent on supplied amino acids for growth. Upregulation of macropinocytosis increases uptake of extracellular proteins. The intracellular degradation of the internalized proteins provides an essential amino acid supply in transformed cells ([Commisso *et al.*, 2013](#)). Thus, macropinocytosis of extracellular proteins provides an amino acid supply route for the tumor cells, making macropinocytosis a potential therapeutic target.

Caveolae

Definition

Caveolae are 50–80 nm cup-shaped pits in the plasma membrane of many vertebrate cells ([Figure 2](#)). Caveolae are extremely abundant in adipocytes, skeletal muscle cells, endothelia, and fibroblasts but undetectable in some other cell types. Caveolae are generally classified as uncoated structures as they do not possess the prominent coat structure characteristic of clathrin-coated pits. However, with specialized

electron microscopic techniques a characteristic striated coat can be visualized on the cytoplasmic face of caveolae. Caveolae are defined by their morphology and by their association with two classes of proteins, caveolins and cavins.

Molecules

Machinery: Caveolins are oligomeric integral membrane proteins that have been widely used as markers of caveolae. Three gene products exist in mammalian cells; caveolin-1 (CAV1), caveolin-2 (CAV2), and caveolin-3 (CAV3). CAV1 and CAV3 are essential for caveolar formation in non-muscle tissues and striated muscle respectively whereas CAV2 plays a facilitatory role in combination with CAV1 ([Parton and Simons, 2007](#)). Cavins are cytoplasmic proteins that work together with caveolins to generate stable caveolae. Four cavin proteins exist in mammalian cells, cavin1 (PTRF), cavin2 (SDPR), cavin3 (PRKCDP) and the muscle-specific cavin4 (MURC), forming a complex within the cytosol of mammalian cells which is recruited to the cell surface to form caveolae in conjunction with caveolins ([Hansen and Nichols, 2010](#); [Ariotti and Parton,](#)

2013). Cavins can be released into the cell in response to specific stimuli (Sinha *et al.*, 2011) but during endocytic trafficking of caveolae stay associated with the budded caveolae (Boucrot *et al.*, 2011). Budding of caveolae is dependent on the mechanical enzyme dynamin which has been localized to the neck of caveolae (Oh *et al.*, 2012, 1998; Henley *et al.*, 1998) or shown to be recruited in response to specific cargoes (Pelkmans *et al.*, 2002). The large ATPase, EHD2, negatively regulates the dynamin-dependent budding of caveolae (Morén *et al.*, 2012; Stoeber *et al.*, 2012).

Cargo: Considerable controversy remains about the specific cargo internalized through caveolae. Proposed cargo include the uncoated virus simian virus 40 (Anderson *et al.*, 1996; Stang *et al.*, 1997), bacterial toxins such as cholera toxin, and membrane proteins including the TGF-beta receptor (Di Guglielmo *et al.*, 2003), insulin receptor (Fagerholm *et al.*, 2009), integrin beta1 (Kanai *et al.*, 2014), and LRP1 (Kanai *et al.*, 2014). In very few cases has the dependence of these agents on caveolae for uptake (e.g., by comparison with cells lacking caveolae) been demonstrated. Where this has been examined in detail the agent of interest was generally still efficiently internalized in cells lacking caveolae (e.g., SV40; Damm *et al.*, 2005).

Carriers

Budded caveolae have been described in ultrastructural studies which carefully distinguished surface-connected caveolae from budded caveolae (Kirkham *et al.*, 2005). Budded caveolae showed the predicted size, vesicular morphology, and CAV1 labeling consistent with scission of individual caveolae from the plasma membrane. In addition, larger complex multicaveolar structures have been described in many cell types (Parton, 1994).

Pathway

A consensus has now emerged to suggest that caveolae fuse with the Rab5 and EEA1-positive early endosome (Boucrot *et al.*, 2011; Pelkmans *et al.*, 2004). Earlier studies suggested that a novel endosomal compartment, the caveosome, was the destination for budded caveolae (Pelkmans *et al.*, 2001). This is now thought to represent an artefactual situation due to overexpression of GFP-tagged CAV1 (Hayer *et al.*, 2010) and until further evidence is available the caveosome can be removed from caveolae-trafficking schemes (Parton and Howes, 2010). Caveolae can recycle back to the plasma membrane from the early endosome. During this cycle and after fusion with the early endosome, caveolar domains defined by a specific quantum of caveolin molecules can be detected on the limiting membrane of the early endosome (Pelkmans *et al.*, 2004). After releasing their cargo of cholera toxin, these domains can cycle back to the cell surface.

Caveolin can also be turned over via recruitment into the intraluminal vesicles of multivesicular bodies and degradation (Hayer *et al.*, 2010). This may be particularly important in cells in which there is an imbalance of caveolin and cavins, for example if caveolin is overexpressed or cavin levels are low (Hayer *et al.*, 2010; Hill *et al.*, 2008).

Quantitative Aspects and Dynamics

Caveolae can occupy over 40% of the cell surface area in adipocytes. In other cells, such as many lymphocytes (Fra *et al.*, 1994), many neurons, and kidney proximal tubule cells (Zhuang *et al.*, 2011) caveolae are undetectable. This should be taken into account when attempting to understand the endocytic functions of caveolae. In fibroblastic cell lines caveolae generally occupy between 1% and 7% of the plasma membrane surface area. In these cells, in which caveolar dynamics have been studied in some detail, total internal reflection fluorescence microscopy (TIRF) microscopy has been used to follow the dynamics of GFP CAV1 puncta (considered to be caveolae). The lifetime of individual caveolae has been shown to vary within one population with caveolar budding within seconds up to several minutes (Boucrot *et al.*, 2011). Electron microscopic studies estimated 0.5–1% of caveolae to bud within 1 minute in cultured cells (Kirkham *et al.*, 2005; Le Lay *et al.*, 2006).

Functions

Patients lacking CAV1 (and thus non-muscle caveolae) show a severe lipodystrophy, characterized by lack of adipose tissue (Garg and Agarwal, 2008). In addition, a number of muscle diseases have been linked to mutations in CAV3. The underlying cause of these conditions may be due to a loss of mechanoprotection mediated by caveolae (Sinha *et al.*, 2011). Cavins have also been linked to a number of diseases including cardiomyopathy and cancer. Caveolar endocytosis has been implicated in transcellular transport across the endothelial monolayer (Simionescu *et al.*, 2002), endocytosis of junctional components (Marchiando *et al.*, 2010; Orlichenko *et al.*, 2009), uptake of lipid microdomain components as cells detach from the substratum (del Pozo *et al.*, 2005), uptake of specific surface receptors and membrane repair (Corrotte *et al.*, 2013). The dynamic cycle of caveolar endocytosis may also serve to regulate levels of surface caveolae. For example, caveolar density is modulated as cells progress through the cell cycle by regulation of the caveolar endocytic/exocytic cycle (Boucrot *et al.*, 2011).

Clathrin- and Caveolin-Independent Micropinocytic Pathways

Definition

Unlike macropinocytosis and caveolar endocytosis, micropinocytic clathrin-independent pathways have been difficult to define in terms of morphology and specific markers. Instead, nonspecific endocytic markers, such as fluid phase markers and membrane-bound toxins, that are taken up in a clathrin- and caveolin-independent manner have generally defined clathrin-independent endocytic pathways. However, a number of proteins have now been identified as specific cargo proteins and components of the molecular machinery of these endocytic pathways as outlined below.

Molecules

Machinery: The CLIC/GEEC pathway is now the best understood clathrin-independent endocytic pathway. Endocytosis is

dynamamin-independent but is dependent on the small GTPase Cdc42 (Sabharanjak *et al.*, 2002; Chadda *et al.*, 2007), the coat protein ARF1, actin, the Rho-GAP-domain and BAR domain protein, GRAF1 (Doherty *et al.*, 2011; Lundmark *et al.*, 2008), and the lectin, galectin3 (Lakshminarayan *et al.*, 2014). Cholesterol and glycosphingolipids are integral to the function of the CLIC/GEEC pathway (Chadda *et al.*, 2007; Lakshminarayan *et al.*, 2014; Chaudhary *et al.*, 2014).

Uptake of the interleukin 2-receptor (IL2R) defines a separate endocytic pathway which differs from the above by dependence on dynamin. Key machinery proteins include RhoA, Rac1, PAK1, and PAK2 (Grassart *et al.*, 2008; Lamaze *et al.*, 2001). Endocytosis of MHC Class I appears to define a third pathway which is inhibited by ARF6 mutants (Radhakrishna and Donaldson, 1997) although recent studies have questioned the clathrin independence of this pathway (Stahlschmidt *et al.*, 2014).

Cargo: Proteins lacking signals for uptake via clathrin coated pits are internalized through clathrin-independent pathways. GPI-anchored proteins were one of the first identified CLIC/GEEC cargo proteins (Sabharanjak *et al.*, 2002) but whether there is specific concentration in clathrin-independent carriers at the plasma membrane is unclear (Bhagatji *et al.*, 2009). More recently the transmembrane hyaluronic acid receptor, CD44, was shown to be internalized via the CLIC/GEEC pathway (Lakshminarayan *et al.*, 2014; Howes *et al.*, 2010). CD44 is also associated with clathrin-independent endocytosis via the pathway involved in MHC Class I uptake and which is affected by dominant acting mutants of ARF6 (Eyster *et al.*, 2009). The boundaries between these two pathways are therefore somewhat unclear at present as other cargo proteins, such as CD98, an amino acid transporter, have also been shown to associate with both pathways (Howes *et al.*, 2010; Eyster *et al.*, 2009). Further characterization of the uptake of these and other cargo proteins is needed to resolve whether they are completely distinct or represent the same basic pathway but with slightly different characteristics in different cell types.

Recent studies have lead to a model for the formation of the CLIC/GEEC carriers and recruitment of cargo by galectin3. Galectin3 on the exterior surface of the plasma membrane binds both glycosphingolipids and cargo proteins such as CD44 and these interactions shape the plasma membrane into the tubular forms required for generating the early carrier (Lakshminarayan *et al.*, 2014). These studies demonstrate an intriguing mechanism by which endocytosis can occur without the cytoplasmic clathrin coat; early carrier formation relies on a novel extracellular membrane-shaping and cargo recruitment mechanism.

Carriers

The endocytic carriers on the CLIC/GEEC endocytic pathway were characterized through the development of an electron microscopy technique designed to identify the earliest budded structures and distinguished surface and internal vesicles (Kirkham *et al.*, 2005). This has been combined with electron tomography to yield 3D information (Howes *et al.*, 2010). Early carriers are tubular and cisternal (Figure 2) and even at the earliest stages studied (15 s) intraluminal vesicles, characteristic of later endosomal compartments, were occasionally observed. Tubular elements are less

than 50 nm in diameter and thus have a high surface area: volume ratio.

Less information is currently available on the carriers involved in the MHC class I uptake pathway shown to be disrupted by ARF6 mutants and the pathway mediating IL2R uptake.

Pathway

The CLIC/GEEC pathway converges with the clathrin-dependent endocytic pathway at the Rab5-positive early endosome (Sabharanjak *et al.*, 2002; Howes *et al.*, 2010). Early carriers acquire Rab5 and fuse with EEA1-positive early endosomes in a PI3 kinase-dependent step (Kalia *et al.*, 2006). A number of cargo proteins internalized through clathrin-independent endocytosis then diverge from the conventional endocytic cycling pathway to enter a novel recycling route. This is best studied for CD44, CD98, and CD147 (also called emmprin or basigin). These proteins enter cells via clathrin-independent endocytosis but then rapidly enter specialized recycling tubules devoid of the transferrin receptor and EEA1 (Eyster *et al.*, 2009). CD44, CD98, and CD147 avoid trafficking to lysosomes and degradation (Eyster *et al.*, 2011) and this is mediated by sorting determinants in their cytoplasmic domains. These domains interact with the microtubule and cargo binding protein Hook1 in coordination with Rab22a and microtubules (Maldonado-Báez *et al.*, 2013).

Quantitative Aspects and Dynamics

Until recently the magnitude of clathrin-independent endocytosis was difficult to ascertain because of the lack of specific markers for the pathway and still remains a controversial area (Bitsikas *et al.*, 2014). However, in immortalized mouse embryonic fibroblasts uptake through the CLIC/GEEC pathway appears to surpass uptake through clathrin coated pits (Howes *et al.*, 2010). Using a number of independent techniques including inhibition of other endocytic pathways with specific dynamin inhibitors and EM quantitation using strict morphological criteria in the absence of inhibitors it was estimated that an area equivalent to the entire plasma membrane of the cell could be endocytosed in 12–17 min in these cells, with an estimated 3000 carriers per minute (Howes *et al.*, 2010). This endocytosis must be balanced by insertion of membrane into the plasma membrane through a high capacity recycling pathway. In one study it was estimated that 40% of endocytosed fluid is regurgitated in 5 min (Chadda *et al.*, 2007).

Functions

The magnitude of the CLIC/GEEC pathway and relative nonspecificity suggests that this pathway could allow rapid remodeling of the cell surface. This is consistent with findings that when cells detach, round up and then spread when replated trafficking is via an intracellular compartment enriched in GPI-anchored proteins (Gauthier *et al.*, 2009). Ablation of CI-endocytosis or recycling on this pathway inhibits the spreading process.

One striking feature of the CLIC pathway is its polarization to the leading edge of migrating cells (Howes *et al.*, 2010). Ablation of the CLIC/GEEC pathway reversibly inhibits cell

migration. Undoubtedly with the recent tools that have just become available other cellular processes will be found to be dependent on this universal pathway. For example, recent work has linked CI-endocytosis to lumen formation in endothelia (Porat-Shliom *et al.*, 2013).

Conclusions and Perspectives

Recent work has given insights into the novel molecular mechanisms involved in endocytic carrier formation without clathrin. Caveola formation involves integral membrane proteins working together with a complex of cytoplasmic coat proteins. Formation of CLIC/GEEC carriers involves a novel mechanism of extracellular lipid and membrane protein binding on the external side of the membrane in a process that couples cargo recruitment and membrane carrier formation. Macropinocytosis differs from the above in that the carrier engulfs fluid and uses actin rearrangement as the driving force for the formation of the characteristic large macropinosome.

Recent years have seen great advances in understanding these three types of endocytic pathways but the actual number of distinct pathways is still unclear. While galectin-3 is associated with and functional in the CLIC/GEEC pathway, galectin-4 is associated with novel clathrin-independent early carriers (Lakshminarayan *et al.*, 2014). The boundaries between distinct pathways are also indistinct. Both macropinocytosis and CLIC/GEEC endocytosis can occur in regions of membrane ruffling and share some common components. There is also crosstalk between different pathways with caveolae negatively regulating CLIC/GEEC endocytosis (Chaudhary *et al.*, 2014).

Finally, the three pathways described here have been extensively characterized in cultured cells. Far less is known about these pathways in tissues *in vivo*. What is the magnitude of these pathways in different tissues in a living animal? Analysis of these pathways *in vivo* in health and in disease is a great challenge for the future but the effort will undoubtedly reap great rewards.

Acknowledgments

R. Parton is funded by a Senior Principal Research Fellowship from the National Health and Medical Research council of Australia (NHMRC), NHMRC project and program grants, and the ARC Centre of Excellence in Convergent Bio-Nano Science and Technology. I am grateful to Markus Kerr, Paul Gleeson, Tom Hall, and Susan Nixon for helpful comments.

See also: Interganellar Communication: Interplay and Processes: Clathrin and Clathrin-Dependent Endocytosis

References

Anderson, H.A., Chen, Y., Norkin, L.C., 1996. Bound simian virus 40 translocates to caveolin-enriched membrane domains, and its entry is inhibited by drugs that selectively disrupt caveolae. *Molecular Biology of the Cell* 7 (11), 1825–1834.

- Ariotti, N., Parton, R.G., 2013. SnapShot: Caveolae, caveolins, and cavinins. *Cell* 154 (3), 704–704.e1.
- Bar-Sagi, D., Feramisco, J.R., 1986. Induction of membrane ruffling and fluid-phase pinocytosis in quiescent fibroblasts by ras proteins. *Science* 233 (4768), 1061–1068.
- Bhagatji, P., Leventis, R., Comeau, J., *et al.*, 2009. Steric and not structure-specific factors dictate the endocytic mechanism of glycosylphosphatidylinositol-anchored proteins. *Journal of Cell Biology* 186 (4), 615–628.
- Bitsikas, V., Correa, I.R., Nichols, B.J., 2014. Clathrin-independent pathways do not contribute significantly to endocytic flux. *eLife* 3, e03970.
- Boucrot, E., Howes, M.T., Kirchhausen, T., Parton, R.G., 2011. Redistribution of caveolae during mitosis. *Journal of Cell Science* 124 (Pt 12), 1965–1972.
- Chadda, R., Howes, M.T., Plowman, S.J., *et al.*, 2007. Cholesterol-sensitive Cdc42 activation regulates actin polymerization for endocytosis via the GEEC pathway. *Traffic* 8 (6), 702–717.
- Chaudhary, N., Gomez, G.A., Howes, M.T., *et al.*, 2014. Endocytic crosstalk: Cavinins, caveolins, and caveolae regulate clathrin-independent endocytosis. *PLoS Biology* 12 (4), e1001832.
- Commisso, C., Davidson, S.M., Soydaner-Azeloglu, R.G., *et al.*, 2013. Macropinocytosis of protein is an amino acid supply route in Ras-transformed cells. *Nature* 497 (7451), 633–637.
- Corrotte, M., Almeida, P.E., Tam, C., *et al.*, 2013. Caveolae internalization repairs wounded cells and muscle fibers. *eLife (Cambridge)* 2, e00926.
- Damm, E.M., Pelkmans, L., Kartenbeck, J., *et al.*, 2005. Clathrin- and caveolin-1-independent endocytosis: Entry of simian virus 40 into cells devoid of caveolae. *Journal of Cell Biology* 168 (3), 477–488.
- Di Guglielmo, G.M., Le Roy, C., Goodfellow, A.F., Wrana, J.L., 2003. Distinct endocytic pathways regulate TGF-beta receptor signalling and turnover. *Nature Cell Biology* 5 (5), 410–421.
- Doherty, G.J., Ahlund, M.K., Howes, M.T., *et al.*, 2011. The endocytic protein GRAF1 is directed to cell-matrix adhesion sites and regulates cell spreading. *Molecular Biology of the Cell* 22 (22), 4380–4389.
- Eyster, C.A., Cole, N.B., Petersen, S., *et al.*, 2011. MARCH ubiquitin ligases alter the itinerary of clathrin-independent cargo from recycling to degradation. *Molecular Biology of the Cell* 22 (17), 3218–3230.
- Eyster, C.A., Higginson, J.D., Huebner, R., *et al.*, 2009. Discovery of new cargo proteins that enter cells through clathrin-independent endocytosis. *Traffic* 10 (5), 590–599.
- Fagerholm, S., Örtengren, U., Karlsson, M., *et al.*, 2009. Rapid insulin-dependent endocytosis of the insulin receptor by caveolae in primary adipocytes. *PLoS One* 4 (6), e5985.
- Fra, A.M., Williamson, E., Simons, K., Parton, R.G., 1994. Detergent-insoluble glycolipid microdomains in lymphocytes in the absence of caveolae. *Journal of Biological Chemistry* 269 (49), 30745–30748.
- Garg, A., Agarwal, A.K., 2008. Caveolin-1: A new locus for human lipodystrophy. *Journal of Clinical Endocrinology and Metabolism* 93 (4), 1183–1185.
- Gauthier, N.C., Rossier, O.M., Mathur, A., Hone, J.C., Sheetz, M.P., 2009. Plasma membrane area increases with spread area by exocytosis of a GPI-anchored protein compartment. *Molecular Biology of the Cell* 20 (14), 3261–3272.
- Gillieron, J., Querbes, W., Zeigerer, A., *et al.*, 2013. Image-based analysis of lipid nanoparticle-mediated siRNA delivery, intracellular trafficking and endosomal escape. *Nature Biotechnology* 31 (7), 638–646.
- Grassart, A., Dujeancourt, A., Lazarow, P.B., Dautry-Varsat, A., Sauvonnnet, N., 2008. Clathrin-independent endocytosis used by the IL-2 receptor is regulated by Rac1, Pak1 and Pak2. *EMBO Reports* 9 (4), 356–362.
- Gu, Z., Noss, E.H., Hsu, V.W., Brenner, M.B., 2011. Integrins traffic rapidly via circular dorsal ruffles and macropinocytosis during stimulated cell migration. *Journal of Cell Biology* 193 (1), 61–70.
- Hamilton, N.A., Wang, J.T.H., Kerr, M.C., Teasdale, R.D., 2009. Statistical and visual differentiation of subcellular imaging. *BMC Bioinformatics* 10, 94.
- Hansen, C.G., Nichols, B.J., 2010. Exploring the caves: Cavinins, caveolins and caveolae. *Trends in Cell Biology* 20 (4), 177–186.
- Haye, A., Stoeber, M., Ritz, D., *et al.*, 2010. Caveolin-1 is ubiquitinated and targeted to intraluminal vesicles in endolysosomes for degradation. *Journal of Cell Biology* 191 (3), 615–629.
- Henley, J.R., Krueger, E.W., Oswald, B.J., McNiven, M.A., 1998. Dynamin-mediated internalization of caveolae. *Journal of Cell Biology* 141 (1), 85–99.
- Hewlett, L.J., Prescott, A.R., Watts, C., 1994. The coated pit and macropinocytic pathways serve distinct endosome populations. *Journal of Cell Biology* 124 (5), 689–703.
- Hill, M.M., Bastiani, M., Luetterforst, R., *et al.*, 2008. PTRF-Cavin, a conserved cytoplasmic protein required for caveola formation and function. *Cell* 132 (1), 113–124.

- Howes, M.T., Kirkham, M., Riches, J., *et al.*, 2010. Clathrin-independent carriers form a high capacity endocytic sorting system at the leading edge of migrating cells. *Journal of Cell Biology* 190 (4), 675–691.
- Kalia, M., Kumari, S., Chadda, R., *et al.*, 2006. Arf6-independent GPI-anchored protein-enriched early endosomal compartments fuse with sorting endosomes via a Rab5/phosphatidylinositol-3'-kinase-dependent machinery. *Molecular Biology of the Cell* 17 (8), 3689–3704.
- Kanai, Y., Wang, D., Hirokawa, N., 2014. KIF13B enhances the endocytosis of LRP1 by recruiting LRP1 to caveolae. *Journal of Cell Biology* 204 (3), 395–408.
- Kerr, M.C., Lindsay, M.R., Luetterforst, R., *et al.*, 2006. Visualisation of macropinosome maturation by the recruitment of sorting nexins. *Journal of Cell Science* 119 (Pt 19), 3967–3980.
- Kerr, M.C., Wang, J.T., Castro, N.A., *et al.*, 2010. Inhibition of the PtdIns(5) kinase PIKfyve disrupts intracellular replication of Salmonella. *EMBO Journal* 29 (8), 1331–1347.
- Kerr, M.C., Teasdale, R.D., 2009. Defining macropinocytosis. *Traffic* 10 (4), 364–371.
- Kirkham, M., Fujita, A., Chadda, R., *et al.*, 2005. Ultrastructural identification of uncoated caveolin-independent early endocytic vehicles. *Journal of Cell Biology* 168 (3), 465–476.
- Koivusalo, M., Welch, C., Hayashi, H., *et al.*, 2010. Amiloride inhibits macropinocytosis by lowering submembranous pH and preventing Rac1 and Cdc42 signaling. *Journal of Cell Biology* 188 (4), 547–563.
- Lakshminarayan, R., Wunder, C., Becken, U., *et al.*, 2014. Galectin-3 drives glycosphingolipid-dependent biogenesis of clathrin-independent carriers. *Nature Cell Biology* 16, 592–603.
- Lamaze, C., Dujancourt, A., Baba, T., *et al.*, 2001. Interleukin 2 receptors and detergent-resistant membrane domains define a clathrin-independent endocytic pathway. *Molecular Cell* 7 (3), 661–671.
- Le Lay, S., Hajdud, E., Lindsay, M.R., *et al.*, 2006. Cholesterol-induced caveolin targeting to lipid droplets in adipocytes: A role for caveolar endocytosis. *Traffic* 7 (5), 549–561.
- Liberali, P., Kakkonen, E., Turacchio, G., *et al.*, 2008. The closure of Pak1-dependent macropinosomes requires the phosphorylation of CIBP1/BARS. *EMBO Journal* 27 (7), 970–981.
- Lim, J.P., Gleeson, P.A., 2011. Macropinocytosis: An endocytic pathway for internalising large gulps. *Immunology and Cell Biology* 89 (8), 836–843.
- Lundmark, R., Doherty, G.J., Howes, M.T., *et al.*, 2008. The GTPase-activating protein GRAF1 regulates the CLIC/GEEC endocytic pathway. *Current Biology* 18 (22), 1802–1808.
- Maekawa, M., Terasaka, S., Mochizuki, Y., *et al.*, 2014. Sequential breakdown of 3-phosphorylated phosphoinositides is essential for the completion of macropinocytosis. *Proceedings of the National Academy of Sciences of the United States of America* 111 (11), E978–E987.
- Maldonado-Báez, L., Cole, N.B., Krämer, H., Donaldson, J.G., 2013. Microtubule-dependent endosomal sorting of clathrin-independent cargo by Hook1. *Journal of Cell Biology* 201 (2), 233–247.
- Marchiando, A.M., Shen, L., Graham, W.V., *et al.*, 2010. Caveolin-1-dependent occludin endocytosis is required for TNF-induced tight junction regulation in vivo. *Journal of Cell Biology* 189 (1), 111–126.
- Mercer, J., Helenius, A., 2009. Virus entry by macropinocytosis. *Nature Cell Biology* 11 (5), 510–520.
- Mercer, J., Helenius, A., 2012. Gulping rather than sipping: Macropinocytosis as a way of virus entry. *Current Opinion in Microbiology* 15 (4), 490–499.
- Mooren, O.L., Galletta, B.J., Cooper, J.A., 2012. Roles for actin assembly in endocytosis. *Annual Review of Biochemistry* 81, 661–686.
- Morén, B., Shah, C., Howes, M.T., *et al.*, 2012. EHD2 regulates caveolar dynamics via ATP-driven targeting and oligomerization. *Molecular Biology of the Cell* 23 (7), 1316–1329.
- Norbury, C.C., 2006. Drinking a lot is good for dendritic cells. *Immunology* 117 (4), 443–451.
- Oh, P., Horner, T., Witkiewicz, H., Schnitzer, J.E., 2012. Endothelin induces rapid, dynamin-mediated budding of endothelial caveolae rich in ET-B. *Journal of Biological Chemistry* 287 (21), 17353–17362.
- Oh, P., McIntosh, D.P., Schnitzer, J.E., 1998. Dynamin at the neck of caveolae mediates their budding to form transport vesicles by GTP-driven fission from the plasma membrane of endothelium. *Journal of Cell Biology* 141 (1), 101–114.
- Orlichenko, L., Weller, S.G., Cao, H., *et al.*, 2009. Caveolae mediate growth factor-induced disassembly of adherens junctions to support tumor cell dissociation. *Molecular Biology of the Cell* 20 (19), 4140–4152.
- Parton, R.G., 1994. Ultrastructural localization of gangliosides; GM1 is concentrated in caveolae. *Journal of Histochemistry & Cytochemistry* 42 (2), 155–166.
- Parton, R.G., Howes, M.T., 2010. Revisiting caveolin trafficking: The end of the caveosome. *Journal of Cell Biology* 191 (3), 439–441.
- Parton, R.G., Simons, K., 2007. The multiple faces of caveolae. *Nature Reviews Molecular Cell Biology* 8 (3), 185–194.
- Pelkmans, L., Bürl, T., Zerial, M., Helenius, A., 2004. Caveolin-stabilized membrane domains as multifunctional transport and sorting devices in endocytic membrane traffic. *Cell* 118 (6), 767–780.
- Pelkmans, L., Kartenbeck, J., Helenius, A., 2001. Caveolar endocytosis of simian virus 40 reveals a new two-step vesicular-transport pathway to the ER. *Nature Cell Biology* 3 (5), 473–483.
- Pelkmans, L., Puntener, D., Helenius, A., 2002. Local actin polymerization and dynamin recruitment in SV40-induced internalization of caveolae. *Science* 296 (5567), 535–539.
- Porat-Shliom, N., Weigert, R., Donaldson, J.G., 2013. Endosomes derived from clathrin-independent endocytosis serve as precursors for endothelial lumen formation. *PLoS One* 8 (11), e81987.
- del Pozo, M.A., Balasubramanian, N., Alderson, N.B., *et al.*, 2005. Phospho-caveolin-1 mediates integrin-regulated membrane domain internalization. *Nature Cell Biology* 7 (9), 901–908.
- Radhakrishna, H., Donaldson, J.G., 1997. ADP-ribosylation factor 6 regulates a novel plasma membrane recycling pathway. *Journal of Cell Biology* 139 (1), 49–61.
- Sabharanjak, S., Sharma, P., Parton, R.G., Mayor, S., 2002. GPI-anchored proteins are delivered to recycling endosomes via a distinct cdc42-regulated, clathrin-independent pinocytic pathway. *Developmental Cell* 2 (4), 411–423.
- Sallusto, F., Cella, M., Danieli, C., Lanzavecchia, A., 1995. Dendritic cells use macropinocytosis and the mannose receptor to concentrate macromolecules in the major histocompatibility complex class II compartment: Downregulation by cytokines and bacterial products. *Journal of Experimental Medicine* 182 (2), 389–400.
- Schnatwinkel, C., Christoforidis, S., Lindsay, M.R., *et al.*, 2004. The Rab5 effector Rabankyrin-5 regulates and coordinates different endocytic mechanisms. *PLoS Biology* 2 (9), E261.
- Simionescu, M., Gafencu, A., Antohe, F., 2002. Transcytosis of plasma macromolecules in endothelial cells: A cell biological survey. *Microscopy Research and Technique* 57 (5), 269–288.
- Sinha, B., Köster, D., Ruez, R., *et al.*, 2011. Cells respond to mechanical stress by rapid disassembly of caveolae. *Cell* 144 (3), 402–413.
- Stahlschmidt, W., Robertson, M.J., Robinson, P.J., McCluskey, A., Haucke, V., 2014. Clathrin terminal domain-ligand interactions regulate sorting of mannose 6-phosphate receptors mediated by AP-1 and GGA adaptors. *Journal of Biological Chemistry* 289 (8), 4906–4918.
- Stang, E., Kartenbeck, J., Parton, R.G., 1997. Major histocompatibility complex class I molecules mediate association of SV40 with caveolae. *Molecular Biology of the Cell* 8 (1), 47–57.
- Stoeber, M., Stoeck, I.K., Hänni, C., *et al.*, 2012. Oligomers of the ATPase EHD2 confine caveolae to the plasma membrane through association with actin. *EMBO Journal* 31 (10), 2350–2364.
- Swanson, J.A., 2008. Shaping cups into phagosomes and macropinosomes. *Nature Reviews Molecular Cell Biology* 9 (8), 639–649.
- Wang, J.T., Kerr, M.C., Karunaratne, S., *et al.*, 2010. The SNX-PX-BAR family in macropinocytosis: The regulation of macropinosome formation by SNX-PX-BAR proteins. *PLoS One* 5 (10), e13763.
- West, M.A., Prescott, A.R., Eskelinen, E.L., Ridley, A.J., Watts, C., 2000. Rac is required for constitutive macropinocytosis by dendritic cells but does not control its downregulation. *Current Biology* 10 (14), 839–848.
- Zhuang, Z., Marshansky, V., Breton, S., Brown, D., *et al.*, 2011. Is caveolin involved in normal proximal tubule function? Presence in model PT systems but absence in situ. *American Journal of Physiology – Renal Physiology* 300 (1), F199–F206.

Further Reading

- Doherty, G.J., McMahon, H.T., 2009. Mechanisms of endocytosis. *Annual Review of Biochemistry* 78, 857–902.
- Lim, J.P., Gleeson, P.A., 2011. Macropinocytosis: An endocytic pathway for internalising large gulps. *Immunology and Cell Biology* 89 (8), 836–843.
- Mayor, S., Parton, R.G., Donaldson, J.G., 2014. Clathrin-independent pathways of endocytosis. *Cold Spring Harbor Perspectives in Biology* 6, a016758.
- Parton, R.G., del Pozo, M.A., 2013. Caveolae as plasma membrane sensors, protectors and organizers. *Nature Reviews Molecular Cell Biology* 14 (2), 98–112.

Rabs of the Endosomal Recycling Pathway

AJ Lindsay and MW McCaffrey, University College Cork, Corcaigh, Ireland

© 2016 Elsevier Inc. All rights reserved.

Overview

The balance between the internalization and recycling of proteins and lipids is crucial for control of plasma membrane composition and contributes to many cellular processes such as nutrient uptake, cell signaling, cell adhesion and junction formation, cell migration, cytokinesis, and cell polarity. The endosomal recycling pathway is composed of two main compartments, the early sorting endosome (ESE) and the endosomal recycling compartment (ERC). Early endosomes are tubulo-vesicular structures that typically assume a peripheral location in the cell, while the ERC is a collection of tubular membranes that are typically condensed around the centrosome, also known as the microtubule organizing center near the nucleus (Maxfield and McGraw, 2004). Proteins present in endosomes may be cargo such as receptors and their ligands, or they may be proteins that contribute to the shape and function of the vesicle or organelle, for example, EEA1. Indeed, the ESE matures into late endosomes (LE).

There are several mechanisms by which molecules are internalized from the plasma membrane. These mechanisms fall into two main categories: clathrin-mediated endocytosis and clathrin-independent endocytosis (CIE). In clathrin-mediated endocytosis the cytoplasmic domains of cell surface proteins are recognized by adapter proteins and packaged into clathrin-coated vesicles that are transported into the cell. CIE is an umbrella term for several modes of internalization including phagocytosis, macropinocytosis, and caveolae-mediated internalization. Regardless of the mechanism of internalization most pathways converge at the ESE, where cargo is sorted for onward transport to the degradative or recycling pathways (Grant and Donaldson, 2009).

ESEs have a pH of ~ 6.0 that promotes the uncoupling of many ligands from their receptors. This represents the first sorting step, in which the receptors remain embedded in the membrane bilayer while the ligand is released into the lumen. Most ESE are capable of accepting internalized cargo for 5–10 min, after which they undergo ‘maturation’ into LE. Maturation of ESE involves their microtubule-dependent translocation toward the cell interior, the lumen becomes more acidic, they stop fusing with newly endocytosed vesicles and they accumulate acid hydrolases. Thus cargo destined for degradation does not leave the lumen of the ESE. This is an efficient means of ‘delivering’ cargo to LE, eliminating the need for transport intermediates.

Cargo that is to be recycled back to the plasma membrane, which accounts for the majority of internalized molecules, is removed from the maturing ESE during the period of time before it translocates to the cell center. This occurs by the pinching off of narrow membrane tubules from the ESE. These tubules have a greater surface-to-volume ratio than the vesicular portion of the ESE, therefore recycled membrane is preferentially sorted away from the soluble molecules in the lumen. This is a bulk transport pathway, i.e., a membrane-

bound protein that has been recently internalized will be transported away from LE in recycling tubules. Membrane-bound proteins that are to be degraded must have specific targeting information, usually ubiquitin covalently linked to lysine residues in the cytoplasmic domain, which targets them to LE (Jovic *et al.*, 2010).

There are two main pathways for returning cargo to the plasma membrane (Figure 1). The ‘short-loop’ recycling pathway involves transport of molecules directly from the ESE back to the plasma membrane. The halftime for this pathway is ~ 2 min. The other route, the so-called ‘long-loop’ recycling pathway, involves transport back to the plasma membrane via the ERC. The ERC is a longer-lived organelle which appears to have several distinct return routes back to the cell surface. Cargos, such as the transferrin receptor, that are internalized by clathrin-mediated endocytosis are recycled back to the plasma membrane in transport carriers that are distinct from the tubular recycling endosomes that carry cargo molecules, such as sphingomyelin, which undergo CIE. Despite the fact that these two pathways are regulated by different machineries they have similar kinetics, with a $T_{1/2}$ for return to the cell surface of ~ 10 min for both routes. Glycosylphosphatidylinositol-anchored proteins are recycled back to the plasma membrane by a much slower pathway, with a $T_{1/2}$ of ~ 30 min. Cargos that enter the ERC can also be sorted for onward transport to the *trans*-Golgi network (TGN). Shiga toxin and TGN38 are examples of such cargo (Grant and Donaldson, 2009).

Polarized epithelial cells have a more complex endosomal recycling pathway. Molecules internalized from the basolateral plasma membrane enter basolateral early endosomes while molecules from the apical plasma membrane enter apical early endosomes. Cargo that is not trafficked onwards to LE, or directly recycled back to the plasma membrane of origin, are sent to a common recycling endosome (CRE), which is mildly acidic. Proteins destined for the apical plasma membrane are further sorted into an apical recycling endosome (ARE). It is unclear whether the ARE is a distinct organelle or a sub-domain of the CRE, however the ARE can be distinguished by its cup-shaped morphology, the absence of basolateral cargo such as the transferrin receptor, and a neutral pH (Thompson *et al.*, 2007).

Many ‘nonclassical’ endosomal recycling pathways can be found in specialized cells. In mature neurons recycling endosomes localize to the base of dendritic spines and play a role in long-term potentiation (LTP). LTP is one of the major cellular mechanisms that underpin learning and memory and relies on the rapid insertion of glutamate receptors into the spine and spine growth. The recycling endosomes serve as local stores of membrane and glutamate receptors and stimuli that induce LTP result in the rapid translocation of recycling endosomes into spines (Kelly *et al.*, 2011).

In fat and muscle cells GLUT4 is the major glucose transporter, and controlling the amount of GLUT4 on the surface of these cells regulates the amount of glucose they internalize.

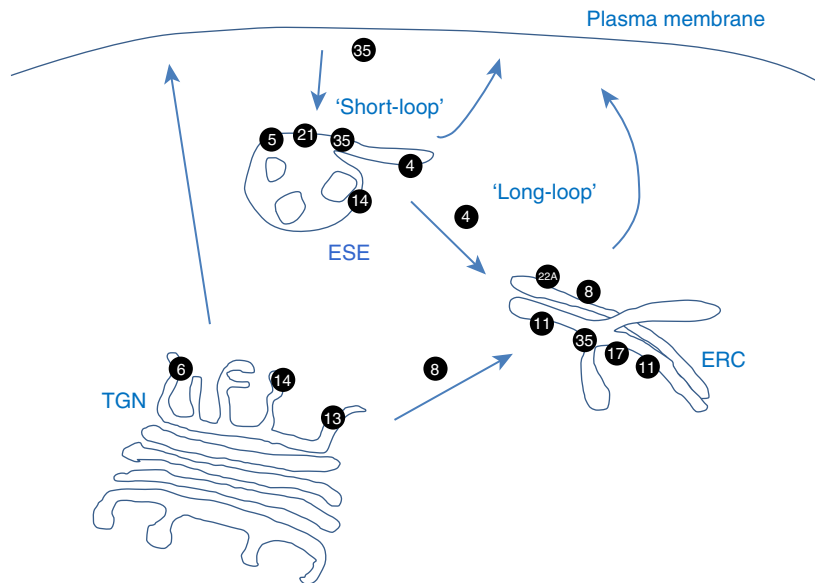


Figure 1 Schematic depiction of the endosomal recycling pathway. The localization of the Rab GTPases that function along this pathway is indicated. ESE, early sorting endosome; ERC, endosomal recycling compartment; TGN, *trans*-Golgi network.

GLUT4 is found in the ERC and TGN, but in the absence of insulin the majority of GLUT4 is stored in specialized vesicles called GLUT4 storage vesicles (GSV), or insulin responsive vesicles. In low insulin conditions GLUT4 idly cycles between these intracellular compartments with a small fraction reaching the plasma membrane at any point in time. Following a meal, the pancreas secretes insulin into the bloodstream which binds to its receptor on the surface of myocytes and adipocytes activating the PI-3-kinase/Akt pathway, leading to a rapid translocation of GLUT4-containing GSVs to the plasma membrane. This results in a 10- to 30- fold increase in the amount of glucose taken up by these cells. It is currently believed that GLUT4 undergoes clathrin-mediated endocytosis and is transported from the ESE to the ERC. In the absence of insulin GLUT4 is sorted from the ERC to GSVs, with a small fraction going directly from the ERC to the plasma membrane. However, in the continual presence of insulin it recycles back to the plasma membrane in vesicles that also contain the transferrin receptor and are distinct from GSVs (Stockli *et al.*, 2011).

Endosomal recycling is a highly regulated process and requires large protein machineries to ensure that cargo reaches the right destination at the right time. Defects in components of these machineries can have severe consequences for the cell and lead to a wide range of diseases including absorption disorders such as hemochromatosis, many neurological diseases, cancer, intestinal disorders such as microvillus inclusion disease, metabolic disorders such as Niemann–Pick disease, osteoporosis, and many others. Furthermore, many bacterial and viral pathogens subvert components of the endosomal recycling pathway during their infection cycle (Puertollano, 2006).

Key regulators of the endosomal recycling pathways are members of the Rab family of small GTPases. Humans have over sixty Rabs and they constitute the largest branch of the Ras superfamily of small GTPases. Some are expressed

ubiquitously while others are tissue-specific. Each membrane-bound compartment is typically associated with one or more Rabs. They are molecular switches that control all aspects of intracellular transport including vesicle budding, uncoating, transport, tethering, and fusion. As their name suggests they bind to guanine-nucleotides and cycle between a GTP-bound active conformation during which they recruit a diverse range of ‘effector’ proteins, and a GDP-bound inactive conformation which is predominantly cytosolic. They mediate, or ‘effect,’ their transport functions through the effector proteins they recruit when in their active conformation. Effectors include motor proteins, and proteins involved in vesicle formation, tethering and membrane fusion. GTP-bound Rabs can also recruit proteins involved in phospholipid synthesis and can directly bind to cargo proteins (Kelly *et al.*, 2012).

Several accessory proteins regulate the Rab GTPase cycle. A GTPase-activating protein stimulates the intrinsic GTPase activity of the Rab to promote the hydrolysis of bound GTP to GDP. A GDP-dissociation inhibitor (GDI) binds to the GDP-bound Rab and extracts it from membranes by interacting with the geranylgeranyl lipid moieties that are attached to two cysteines at the carboxy-terminus of the Rab. The GDI also prevents the dissociation of the GDP. A RabGEF (GDP/GTP exchange factor) displaces the GDI and catalyzes the exchange of GDP for GTP allowing the cycle to start again.

In this article we will discuss in detail the Rab GTPases that have been implicated in the transport of cargo via the ESE and the ERC to the plasma membrane (see Figure 1).

Rab4 and the ‘Short-Loop’ Recycling Pathway

The majority of Rab4 resides on the ESE, with a minor pool also localizing to the ERC (Sonnichsen *et al.*, 2000). There are two Rab4 isoforms, A and B, which are both ubiquitously expressed and have been implicated in the recycling of cargo

directly to the cell surface from the ESE in the 'short-loop' recycling pathway (van der Sluijs *et al.*, 1992; McCaffrey *et al.*, 2001; Perrin *et al.*, 2013). The trafficking of the β_2 -adrenergic receptor, membrane-type 1 matrix metalloprotease (MT1-MMP), the oxytocin receptor, the mu-opioid receptor, the transferrin receptor and the Fc receptor (FcRn) have all been shown to involve Rab4. Both Rab4 isoforms have also been implicated in the transport of GLUT4 in adipocytes (Mari *et al.*, 2006; Kaddai *et al.*, 2009). Several Rab4 effector proteins have been identified such as the dynein light intermediate chain (Bielli *et al.*, 2001) and syntaxin 4 (Li *et al.*, 2001). Rab4 also shares effectors with several other Rabs including Rabaptin-5 (Vitale *et al.*, 1998), Rabenosyn-5 (de Renzis *et al.*, 2002), and RUFY1 (Fouraux *et al.*, 2004), which it shares with Rab5. D-AKAP2 and RCP interact with Rab4 and Rab11 (Eggers *et al.*, 2009; Lindsay *et al.*, 2002), and RUFY1 also interacts with Rab14 (Yamamoto *et al.*, 2010).

Rab4 has recently been reported to coordinate the formation of the tubular subdomains of the ESE that give rise to carrier vesicles targeted to various intracellular locations including the plasma membrane and the ERC. It does this by orchestrating the formation of a GTPase cascade on the tubules. Rab4 initiates the recruitment of Arl, another class of small GTPase, which in turn recruits the Arf GEFs BIG1 and BIG2. These GEFs then locally activate Arf1 and Arf3, which in turn recruit adapter proteins that are required for cargo capture and carrier vesicle formation (D'Souza *et al.*, 2014).

Rab11 and the Endosomal Recycling Compartment

Rab11 is the best-characterized Rab that functions in the endosomal recycling pathway. The Rab11 subfamily has three members, Rab11A, Rab11B and Rab25 (also known as Rab11C). Rab11A is abundantly expressed in most tissues, Rab11B is expressed widely but not as abundantly as Rab11A, and Rab25 expression is restricted to epithelial tissues. Rab11 is primarily localized to the ERC in most cell types and it regulates recycling from this compartment (Ullrich *et al.*, 1996). Rab11 has also been observed at the TGN and on post-Golgi vesicles suggesting that it functions at the interface between the endosomal and exocytic pathways. It has been implicated in the transport of a wide variety of cell surface proteins including the EGF receptor, E-cadherin, a number of G-protein coupled receptors such as rhodopsin and the β_2 -adrenergic receptor, the AMPA receptor, the Toll-like receptor 4, and the $\alpha 5 \beta 1$ integrin. It influences many cellular processes including neurotransmission, cytokinesis, ciliogenesis, directed cell migration, and development.

Rab11 Effector Proteins

GTP-bound Rab11 recruits a number of effector proteins, key among these are the members of the Rab11-FIPs (Rab11-Family of Interacting Proteins). The Rab11-FIPs are a family of five proteins that all have a highly conserved Rab11-binding domain at their carboxy-termini. The FIPs can be subdivided into two classes. The class I FIPs include RCP (also known as FIP1C), FIP2, and FIP5 (also known as Rip11) and have a C2 phospholipid-binding domain near their amino-termini

which binds phosphatidic acid and PI(3,4,5)P₃ (Lindsay and McCaffrey, 2004). The class II FIPs, FIP3 and FIP4, lack the C2 domain but have EF-hand calcium-binding motifs, suggesting that their function is regulated via calcium signaling. Another key Rab11 effector is the exocyst, an eight subunit complex (Sec15, Sec10, Sec8, Sec84, Sec5, Sec6, Sec3, and Exo70) that tethers transport vesicles to specific regions of the plasma membrane. Exo70 and Sec3 (by analogy to its yeast homolog) bind PI(4,5)P₂ and the TC10 small G protein at the plasma membrane, thereby providing the means to target the exocyst to specific subdomains of the plasma membrane. Rab11-GTP binds directly to Sec15 (Zhang *et al.*, 2004) and this interaction may nucleate the assembly of the exocyst complex on Rab11-positive transport carriers. The exocyst is involved in most, if not all, cellular processes influenced by Rab11.

Rab11 Binds Motor Proteins

Rab11 also binds to a number of motor protein complexes, thus facilitating its bidirectional transport along microtubules and actin filaments. Myosin Vb was the first motor protein identified to interact with Rab11 (Lapierre *et al.*, 2001). It is an actin-based motor protein that mediates short-range transport along actin filaments. Rab11-GTP binds directly to the carboxy-terminal globular tail domain of myosin Vb, and Rab11 has since been found to interact with the same domain of its close relative myosin Va (Roland *et al.*, 2009). The Rab11-myosin V interaction mediates the rapid translocation of recycling endosomes into dendritic spines during synaptic plasticity (Correia *et al.*, 2008; Wang *et al.*, 2008), and the tethering and possible short-range transport of Rab11-positive vesicles at the periphery of non-specialized cells (Lindsay *et al.*, 2013).

Rab11 has also been reported to bind to both plus-end and minus-end directed microtubule motor protein complexes. To date, such interactions have been found to be indirect and occur via an intervening or linker protein. For example, FIP3 binds to the dynein light intermediate chain and acts as a bridge between Rab11 and the dynein motor complex (Horgan *et al.*, 2010). Dynein is a minus-end directed motor and its association with Rab11 facilitates the transport of cargo from the ESE to the ERC. Rab11 also associates with the plus-end directed kinesin 5A (KIF5A) via an interaction with the neuronal protein protrudin (Matsuzaki *et al.*, 2011). The Rab11-protrudin-KIF5A complex plays a role in neurite formation in neurons, and overexpression of protrudin or KIF5A induces the formation of protrusions in non-neuronal cells. A unique feature of this complex is that protrudin interacts with the GDP-bound form of Rab11. Rab11 also interacts indirectly with kinesin-2 via FIP5/Rip11 (Schonteich *et al.*, 2008), and this complex delivers vesicles to the cleavage furrow during apical lumen formation, a key step in epithelial morphogenesis of tubular organs (Li *et al.*, 2014).

Rab11 in Cell Division

Rab11 and its effector FIP3 have been implicated in cytokinesis. Cytokinesis is the final stage of mitosis in which the cytoplasm of a parent cell is divided into two daughter cells. A necessary step for the completion of cytokinesis is the

delivery of membranes to the cleavage furrow, and recent work has demonstrated that an important source of this membrane is endosomes from the recycling pathway. Both Rab11 and FIP3 localize to the cleavage furrow during cytokinesis (Horgan *et al.*, 2004), and RNAi-mediated knockdown of FIP3 results in cytokinesis defects and a failure of cells to undergo abscission (Wilson *et al.*, 2005). FIP3 also binds to Arf6, another small GTPase that is enriched at the cleavage furrow and binds to the Sec10 subunit of the exocyst complex. It has been postulated that a combination of the Rab11 association with Sec15 and the FIP3 association with Arf6 serve to target Rab11: FIP3-endosomes to the cleavage furrow during cell division (Simon and Prekeris, 2008). Recent work has demonstrated that FIP3-endosomes contain several Rac/Rho and actin regulators. The fusion of these endosomes with the plasma membrane surrounding the intercellular bridge (ICB) causes localized actin depolymerization which results in the thinning of the ICB to less than 100 nm, a process called secondary ingression. This allows the recruitment of further machinery (the ESCRT complex) that facilitates the abscission of the two daughter cells (Schiel *et al.*, 2012).

The Role of Rab11 in Synaptic Transmission

Rab11 regulates neuronal migration and maturation through its role in transporting N-cadherin (Kawauchi *et al.*, 2010), and as mentioned above the Rab11-protrudin-KIF5A complex is involved in neurite formation. *Drosophila melanogaster* embryos with homozygous Rab11 mutants display disorganized and misrouted embryonic axons (Bhuin and Roy, 2009). Overexpression of Rab11 restores synaptic dysfunction due to the expression of mutant huntingtin (htt) protein in a *Drosophila* model of Huntington's disease (Richards *et al.*, 2011; Steinert *et al.*, 2012). Thus Rab11 has an important role in nervous system development, and the observation that Rab11 is present on synaptic vesicles and is involved in mediating the transport of recycling endosomes into stimulated dendritic spines indicates that it is also involved in synaptic transmission in mature neurons.

Rab11 and Cancer

Rab25 (Rab11C) has been implicated in cancer and is unique among the Rab GTPases in that it has a GTP-binding site that is likely to render it constitutively-active *in vivo*. It has been reported to be both an oncogene in aggressive breast and ovarian cancer (Cheng *et al.*, 2004), and a tumor-suppressor gene in colon cancer (Nam *et al.*, 2010). Cheng *et al.* reported that Rab25 was upregulated at the DNA, mRNA, and protein levels and that this correlates with reduced survival of breast and ovarian cancer patients. They also demonstrated that overexpression of Rab25 in ovarian cancer cell lines increased several of the hallmarks of cancer (cell proliferation, resistance to apoptosis, anoikis, anchorage-independent growth). Rab25 expression also correlates with histological grade in patients, making it a useful clinical biomarker for ovarian cancer (Sheach *et al.*, 2009). While Rab25 overexpression can clearly be correlated to ovarian cancer aggressiveness, the situation is not as clear for breast cancer. Rab25 loss has been observed in a number of breast cancer cell lines (Cheng *et al.*, 2006, 2010).

It has recently been proposed that Rab25 functions as an oncogene in hormone receptor-positive breast cancers and as a tumor-suppressor in estrogen receptor-negative breast cancers (Mitra *et al.*, 2012). Decreases in Rab25 expression have also been observed in colon cancer patients, and this loss was associated with poorer patient outcome (Nam *et al.*, 2010). These contradictory results suggest that the role of Rab25 in cancer progression is highly context dependent and is likely to depend on the repertoire of Rab25-interacting proteins expressed by different cell types. Furthermore, Rab25 itself has been reported to control the expression of a number of gene transcripts. Knockdown of Rab25 leads to the loss of α_5 -integrin mRNA (Krishnan *et al.*, 2013), suggesting that one of the ways Rab25 may function as a tumor suppressor is by impacting upon the expression of integrins.

Our knowledge of how Rab25 functions mechanistically in cancer has been aided greatly by insights from the laboratory of Jim Norman at the Beatson Institute, Glasgow. The key features of cancer are dysregulated cell proliferation and increased migratory behavior of cancer cells, which can lead to invasion of healthy tissue and metastatic dissemination. Jim Norman and coworkers have shown that Rab25 binds directly to the β_1 integrin and promotes cell migration by transporting $\alpha_5\beta_1$ integrin-containing vesicles to the pseudopodial-tip of migrating ovarian cancer cells (Caswell *et al.*, 2007), and they subsequently showed that the Rab11 effector, RCP, is also involved in the transport of $\alpha_5\beta_1$ integrin to the leading edge of migrating cells (Caswell *et al.*, 2008). Recently they demonstrated that Rab25 also retrieves integrins that have been targeted for degradation from LE/lysosomes and transports them to the plasma membrane (Dozynkiewicz *et al.*, 2012).

Other Rabs of the Endosomal Recycling Pathway

Rab4 and Rab11 are the best studied Rabs that participate in endosomal recycling, however several other Rabs have also been implicated in this pathway. Rab8 is involved in the transport of post-Golgi transport carriers from the TGN to the plasma membrane, however in polarized cells some cargo proteins are transported to the plasma membrane via the ERC. Rab8 localization sometimes overlaps with Rab11, suggesting that these two Rabs can function at the intersection of the biosynthetic and recycling pathways. Indeed, Rab8 and Rab11 share a number of effector proteins including myosin Va (Roland *et al.*, 2009; Lindsay *et al.*, 2013) and myosin Vb (Roland *et al.*, 2007). Rab8 and Rab22A are also involved in the recycling of cargo, that has entered the cell by CIE, in tubular recycling intermediates (Sharma *et al.*, 2009; Hattula *et al.*, 2006; Weigert *et al.*, 2004). Another Rab that bridges transport between the TGN and the ERC is Rab13 (Nokes *et al.*, 2008). One of the cargos that Rab13 transports is the tight junction protein occludin (Morimoto *et al.*, 2005). Rab21 localizes to the ESE and is involved in β_1 -integrin trafficking (Pellinen *et al.*, 2006). Rab17 is an epithelial-specific Rab and its expression is induced during cell polarization (Lutcke *et al.*, 1993). It localizes to the ARE and is involved in regulating the transcytosis of cargo from the basolateral to apical membranes (Zacchi *et al.*, 1998; Hunziker and Peters, 1998). Rab14 is ubiquitously expressed and localizes to the endosomal

recycling pathway and to some extent the TGN (Junutula *et al.*, 2004). Interestingly, as is the case for many Rabs involved in the regulation of endosomal trafficking Rab14 interacts with effectors of Rab4 and Rab11. Specifically Rab14 interacts with RUFY1, a Rab4 effector (Yamamoto *et al.*, 2010), and with the class I Rab11-FIPs (Kelly *et al.*, 2009; Lindsay *et al.*, 2002). A recent report has localized Rab14 to an intermediate compartment between the ESE and the ERC, from which it transports the ADAM10 protease to cell–cell junctions in migrating cells. ADAM10 cleaves N-cadherin, an integral component of adherens junctions. In Rab14-depleted cells ADAM10 transport to the adherens junctions is inhibited resulting in the inhibition of cell–cell junction disassembly, since N-cadherin cannot be cleaved, and this leads to reduced cell migration (Linford *et al.*, 2012).

Rab35 localizes to membranes throughout the endosomal recycling pathway including the plasma membrane, ESE, and tubular recycling compartment intermediates. It regulates a direct recycling pathway, and also plays a role in cytokinesis (Kouranti *et al.*, 2006). Rab35 also regulates the actin cytoskeleton (Zhang *et al.*, 2009), and localizes to synaptic vesicles in *Drosophila* (Uytterhoeven *et al.*, 2011).

Intracellular Pathogens Subvert the Function of Rab GTPases

The normal host defense for protecting cells during infection by bacterial pathogens is to transport pathogen-containing vacuoles (PCV) to lysosomes for degradation. This pathway requires many Rab GTPases to coordinate the transport of the PCVs from an early phagocytic compartment to a Rab5-positive endosomal compartment, their subsequent shuttling through a Rab7-positive late endosome and onward transport to the lysosome. Studies on latex bead-containing phagosomes in mouse macrophages have implicated up to 40 different Rabs in this process (Stein *et al.*, 2012).

In order to evade degradation many pathogens hijack Rab protein function. They do this by expressing bacterial effectors, a diverse range of proteins that can act as receptors for Rab GTPases, or as enzymes that posttranslationally modify Rab proteins or modify the endosomal membrane lipids required by the Rabs. Bacterial effectors subvert Rab function in order to direct pathogen transport to specific intracellular locations and to monopolize host vesicles in order to facilitate their growth and differentiation. Effector expression is temporally and spatially regulated and multiple factors may act in concert to hijack Rab function. Furthermore, bacterial pathogens can modulate the expression of Rab GTPases upon infection of their hosts (Seixas *et al.*, 2012).

Some intracellular pathogens infect more than one cell type in order to gain access to the host cell machinery to facilitate their growth and differentiation. Some pathogens utilize phagocytic cells such as macrophages, for example, *Legionella pneumophila* and *Mycobacterium tuberculosis*, while others infect non-phagocytic cells and enter the cell by receptor-mediated endocytosis, lipid raft mediated internalization, or by modulation of the actin cytoskeleton through Rho family GTPases. *Salmonella enterica* invades intestinal epithelia and *Helicobacter pylori* specifically infects mucosal epithelia.

Irrespective of the cell type the following intracellular pathogens all modulate the transport of their vacuoles by subverting the function of the membrane trafficking machinery in order to evade degradation in lysosomes – Chlamydiae, *H. pylori*, *L. pneumophila*, Mycobacteria, *Pseudomonas aeruginosa*, and *S. enterica*.

A multistep process, involving a series of host-bacterial protein interactions that modulate Rab activity, is required to establish an appropriate intracellular niche for an invading pathogen to grow and divide. These steps include the inhibition of PCV transport along the degradative pathway, transport of the PCV to appropriate intracellular sites for bacterial growth and differentiation, modulation of cell signaling, and the recruitment of host vesicles that contain the proteins and lipids required to modify the PCV and provide nutrients for growth.

Chlamydia trachomatis is a well studied bacterial pathogen that subverts the function of Rab GTPases for its own ends. It is the most common sexually transmitted bacterium in the Western world and left untreated *Chlamydia* infections can lead to blindness. *Chlamydia trachomatis* enters host mucosal epithelial cells or macrophages and establishes a membrane-bound vacuole called an inclusion, in which it spends its entire lifecycle. Several Rab proteins from the endosomal recycling and biosynthetic pathway (Rab1, Rab4, Rab6, Rab11, and Rab14) have been localized to the inclusion. Unsurprisingly, Rabs involved in the degradative pathway (Rab5, Rab7, and Rab9) are not found on the inclusion body. Rab interacting proteins such as FIP2, but not the other FIP family members, also localize to the inclusion (Leiva *et al.*, 2013). Chlamydiae express a group of coiled-coil proteins called Incs that are inserted into the inclusion membrane with a domain exposed to the cytosol and are believed to be key factors involved in subverting Rab function. Indeed, the Inc protein CT229 binds directly to GTP-bound Rab4. Another such protein is the immediate early gene product CT147. It is a 162 kDa protein that shares strong structural homology with EEA1, a Rab5 effector that mediates the tethering and Rab5-dependent fusion of endosomes. CT147 lacks the Rab5 binding domain that is present in EEA1 and it has been proposed that in chlamydial infected cells it may function to tether endosomes but prevent their fusion, thus providing a means for chlamydial endosomes to avoid fusion with lysosomes (Belland *et al.*, 2003). Rab6, Rab11, and Rab14 are believed to play a role in the transport of sphingolipids to the inclusion as their depletion hinders the maturation of *C. trachomatis* and reduces the yield of infectious particles (Damiani *et al.*, 2014). See Stein *et al.* (2012) for a comprehensive review of other bacterial pathogens that hijack Rab function.

A number of viruses have also been found to subvert the function of Rab GTPases belonging to the endosomal recycling pathway. Recently it was shown that the HIV-1 retrovirus utilizes Rab14, and its effector RCP, to transport the HIV envelope glycoprotein complex to microdomains of the plasma membrane where they will be incorporated into HIV-1 particles (Qi *et al.*, 2013). The herpes simplex virus-1 is endocytosed from the plasma membrane and transported to the ERC. Rab11-depletion reduces virus yield in infected cells indicating that it is involved in HSV-1 morphogenesis (Hollinshead *et al.*, 2012). Rab11 is also involved in Influenza A virus

morphogenesis and budding (Bruce *et al.*, 2010). Both the respiratory syncytial virus (the most common cause of lower respiratory tract illness in children) and HIV-1 'piggy back' on myosin Vb in order to be transported to the plasma membrane (Brock *et al.*, 2003; Varthakavi *et al.*, 2006).

Summary

The endosomal recycling pathway plays a critical role in determining how cells communicate with the 'outside world' by regulating the protein and lipid composition of the plasma membrane. A number of Rab GTPases have been found to play key roles in coordinating the spatiotemporal transport of cargo through this pathway.

It is becoming increasingly clear that endosomal recycling does not occur on a single autonomous transport route but rather involves trafficking along a pathway that is intersected by components of other pathways including the biosynthetic and degradative pathways.

Defects in endosomal recycling can have severe consequences for the cell, which may lead to the development of a diverse range of catastrophic diseases including cancer. Furthermore, intracellular pathogens have learned to subvert the function of components of the endosomal recycling pathway to facilitate their own growth and differentiation.

A greater understanding of the molecular mechanisms involved in regulating endosomal recycling will not only lead to further insights into the functioning of normal cells but will also lead to a greater understanding of the etiology of diseases ranging from neurodegenerative disorders to cancer, as well as pathogenic bacterial and viral diseases. Such findings will likely lead to the identification of new targets for therapeutic intervention.

Acknowledgments

This work was supported by a Science Foundation Ireland Programme Grant (09/IN1/B2629).

See also: Interorganellar Communication: Interplay and Processes: Endocytosis of Cargo Proteins: LDL. Organelles: Structure and Function: Early Endosomal Compartments

References

- Belland, R.J., Zhong, G., Crane, D.D., *et al.*, 2003. Genomic transcriptional profiling of the developmental cycle of *Chlamydia trachomatis*. *Proceedings of the National Academy of Sciences of the United States of America* 100, 8478–8483.
- Bhuin, T., Roy, J.K., 2009. Rab11 is required for embryonic nervous system development in *Drosophila*. *Cell and Tissue Research* 335, 349–356.
- Bielli, A., Thornqvist, P.O., Hendrick, A.G., *et al.*, 2001. The small GTPase Rab4A interacts with the central region of cytoplasmic dynein light intermediate chain-1. *Biochemical and Biophysical Research Communications* 281, 1141–1153.
- Brock, S.C., Goldenring, J.R., Crowe Jr., J.E., 2003. Apical recycling systems regulate directional budding of respiratory syncytial virus from polarized epithelial cells. *Proceedings of the National Academy of Sciences of the United States of America* 100, 15143–15148.
- Bruce, E.A., Digard, P., Stuart, A.D., 2010. The Rab11 pathway is required for influenza A virus budding and filament formation. *Journal of Virology* 84, 5848–5859.
- Caswell, P.T., Chan, M., Lindsay, A.J., *et al.*, 2008. Rab-coupling protein coordinates recycling of alpha5beta1 integrin and EGFR1 to promote cell migration in 3D microenvironments. *Journal of Cell Biology* 183, 143–155.
- Caswell, P.T., Spence, H.J., Parsons, M., *et al.*, 2007. Rab25 associates with alpha5beta1 integrin to promote invasive migration in 3D microenvironments. *Developmental Cell* 13, 496–510.
- Cheng, J.M., Ding, M., Aribi, A., Shah, P., Rao, K., 2006. Loss of RAB25 expression in breast cancer. *International Journal of Cancer* 118, 2957–2964.
- Cheng, J.M., Volk, L., Janaki, D.K., *et al.*, 2010. Tumor suppressor function of Rab25 in triple-negative breast cancer. *International Journal of Cancer* 126, 2799–2812.
- Cheng, K.W., Lahad, J.P., Kuo, W.L., *et al.*, 2004. The RAB25 small GTPase determines aggressiveness of ovarian and breast cancers. *Nature Medicine* 10, 1251–1256.
- Correia, S.S., Bassani, S., Brown, T.C., *et al.*, 2008. Motor protein-dependent transport of AMPA receptors into spines during long-term potentiation. *Nature Neuroscience* 11, 457–466.
- Damiani, M.T., Gambarte Tudela, J., Capmany, A., 2014. Targeting eukaryotic Rab proteins: A smart strategy for chlamydial survival and replication. *Cellular Microbiology* 16 (9), 1329–1338.
- Dozynkiewicz, M.A., Jamieson, N.B., Macpherson, I., *et al.*, 2012. Rab25 and CLIC3 collaborate to promote integrin recycling from late endosomes/lysosomes and drive cancer progression. *Developmental Cell* 22, 131–145.
- D'Souza, R.S., Semus, R., Billings, E.A., *et al.*, 2014. Rab4 orchestrates a small GTPase cascade for recruitment of adapter proteins to early endosomes. *Current Biology* 24, 1187–1198.
- Eggers, C.T., Schafer, J.C., Goldenring, J.R., Taylor, S.S., 2009. D-AKAP2 interacts with Rab4 and Rab11 through its RGS domains and regulates transferrin receptor recycling. *Journal of Biological Chemistry* 284, 32869–32880.
- Fouraux, M.A., Deneka, M., Ivan, V., *et al.*, 2004. Rabip4' is an effector of rab5 and rab4 and regulates transport through early endosomes. *Molecular Biology of the Cell* 15, 611–624.
- Grant, B.D., Donaldson, J.G., 2009. Pathways and mechanisms of endocytic recycling. *Nature Reviews Molecular Cell Biology* 10, 597–608.
- Hattula, K., Furuhielm, J., Tikkanen, J., *et al.*, 2006. Characterization of the Rab8-specific membrane traffic route linked to protrusion formation. *Journal of Cell Science* 119, 4866–4877.
- Hollinshead, M., Johns, H.L., Sayers, C.L., *et al.*, 2012. Endocytic tubules regulated by Rab GTPases 5 and 11 are used for envelopment of herpes simplex virus. *EMBO Journal* 31, 4204–4220.
- Horgan, C.P., Hanscom, S.R., Jolly, R.S., Futter, C.E., McCaffrey, M.W., 2010. Rab11-FIP3 links the Rab11 GTPase and cytoplasmic dynein to mediate transport to the endosomal-recycling compartment. *Journal of Cell Science* 123, 181–191.
- Horgan, C.P., Walsh, M., Zurawski, T.H., McCaffrey, M.W., 2004. Rab11-FIP3 localizes to a Rab11-positive pericentrosomal compartment during interphase and to the cleavage furrow during cytokinesis. *Biochemical and Biophysical Research Communications* 319, 83–94.
- Hunziker, W., Peters, P.J., 1998. Rab17 localizes to recycling endosomes and regulates receptor-mediated transcytosis in epithelial cells. *Journal of Biological Chemistry* 273, 15734–15741.
- Jovic, M., Sharma, M., Rahajeng, J., Caplan, S., 2010. The early endosome: A busy sorting station for proteins at the crossroads. *Histology and Histopathology* 25, 99–112.
- Junutula, J.R., De Maziere, A.M., Peden, A.A., *et al.*, 2004. Rab14 is involved in membrane trafficking between the Golgi complex and endosomes. *Molecular Biology of the Cell* 15, 2218–2229.
- Kaddai, V., Gonzalez, T., Keslair, F., *et al.*, 2009. Rab4b is a small GTPase involved in the control of the glucose transporter GLUT4 localization in adipocyte. *PLoS One* 4, e5257.
- Kawauchi, T., Sekine, K., Shikanai, M., *et al.*, 2010. Rab GTPases-dependent endocytic pathways regulate neuronal migration and maturation through N-cadherin trafficking. *Neuron* 67, 588–602.
- Kelly, E.E., Horgan, C.P., Adams, C., *et al.*, 2009. Class I Rab11-family interacting proteins are binding targets for the Rab14 GTPase. *Biology of the Cell* 102, 51–62.
- Kelly, E.E., Horgan, C.P., Goud, B., McCaffrey, M.W., 2012. The Rab family of proteins: 25 years on. *Biochemical Society Transactions* 40, 1337–1347.
- Kelly, E.E., Horgan, C.P., McCaffrey, M.W., Young, P., 2011. The role of endosomal-recycling in long-term potentiation. *Cellular and Molecular Life Sciences* 68, 185–194.

- Kourant, I., Sachse, M., Arouche, N., Goud, B., Echard, A., 2006. Rab35 regulates an endocytic recycling pathway essential for the terminal steps of cytokinesis. *Current Biology* 16, 1719–1725.
- Krishnan, M., Lapiere, L.A., Knowles, B.C., Goldenring, J.R., 2013. Rab25 regulates integrin expression in polarized colonic epithelial cells. *Molecular Biology of the Cell* 24, 818–831.
- Lapiere, L.A., Kumar, R., Hales, C.M., et al., 2001. Myosin vb is associated with plasma membrane recycling systems. *Molecular Biology of the Cell* 12, 1843–1857.
- Leiva, N., Capmany, A., Damiani, M.T., 2013. Rab11-family of interacting protein 2 associates with chlamydial inclusions through its Rab-binding domain and promotes bacterial multiplication. *Cellular Microbiology* 15, 114–129.
- Li, D., Kuehn, E.W., Prekeris, R., 2014. Kinesin-2 mediates apical endosome transport during epithelial lumen formation. *Cellular Logistics* 4, e28928.
- Li, L., Omata, W., Kojima, I., Shibata, H., 2001. Direct interaction of Rab4 with syntaxin 4. *Journal of Biological Chemistry* 276, 5265–5273.
- Lindsay, A.J., Hendrick, A.G., Cantalupo, G., et al., 2002. Rab coupling protein (RCP), a novel Rab4 and Rab11 effector protein. *Journal of Biological Chemistry* 277, 12190–12199.
- Lindsay, A.J., Jollivet, F., Horgan, C.P., et al., 2013. Identification and characterization of multiple novel Rab-myosin Va interactions. *Molecular Biology of the Cell* 24, 3420–3434.
- Lindsay, A.J., McCaffrey, M.W., 2004. The C2 domains of the class I Rab11 family of interacting proteins target recycling vesicles to the plasma membrane. *Journal of Cell Science* 117, 4365–4375.
- Linford, A., Yoshimura, S., Nunes Bastos, R., et al., 2012. Rab14 and its exchange factor FAM116 link endocytic recycling and adherens junction stability in migrating cells. *Developmental Cell* 22, 952–966.
- Lutcke, A., Jansson, S., Parton, R.G., et al., 1993. Rab17, a novel small GTPase, is specific for epithelial cells and is induced during cell polarization. *Journal of Cell Biology* 121, 553–564.
- Mari, M., Monzo, P., Kaddai, V., et al., 2006. The Rab4 effector Rabip4 plays a role in the endocytic trafficking of Glut 4 in 3T3-L1 adipocytes. *Journal of Cell Science* 119, 1297–1306.
- Matsuzaki, F., Shirane, M., Matsumoto, M., Nakayama, K.I., 2011. Protrudin serves as an adapter molecule that connects KIF5 and its cargoes in vesicular transport during process formation. *Molecular Biology of the Cell* 22, 4602–4620.
- Maxfield, F.R., McGraw, T.E., 2004. Endocytic recycling. *Nature Reviews Molecular Cell Biology* 5, 121–132.
- McCaffrey, M.W., Bielli, A., Cantalupo, G., et al., 2001. Rab4 affects both recycling and degradative endosomal trafficking. *FEBS Letters* 495, 21–30.
- Mitra, S., Cheng, K.W., Mills, G.B., 2012. Rab25 in cancer: A brief update. *Biochemical Society Transactions* 40, 1404–1408.
- Morimoto, S., Nishimura, N., Terai, T., et al., 2005. Rab13 mediates the continuous endocytic recycling of occludin to the cell surface. *Journal of Biological Chemistry* 280, 2220–2228.
- Nam, K.T., Lee, H.J., Smith, J.J., et al., 2010. Loss of Rab25 promotes the development of intestinal neoplasia in mice and is associated with human colorectal adenocarcinomas. *Journal of Clinical Investigation* 120, 840–849.
- Nokes, R.L., Fields, I.C., Collins, R.N., Folsch, H., 2008. Rab13 regulates membrane trafficking between TGN and recycling endosomes in polarized epithelial cells. *Journal of Cell Biology* 182, 845–853.
- Pellinen, T., Arjonen, A., Vuoriluoto, K., et al., 2006. Small GTPase Rab21 regulates cell adhesion and controls endosomal traffic of beta1-integrins. *Journal of Cell Biology* 173, 767–780.
- Perrin, L., Lacas-Gervais, S., Gilleron, J., et al., 2013. Rab4b controls an early endosome sorting event by interacting with the gamma-subunit of the clathrin adapter complex 1. *Journal of Cell Science* 126, 4950–4962.
- Puertollano, R., 2006. Endocytic trafficking and human diseases. Endosomes. New York, NY: Springer.
- Qi, M., Williams, J.A., Chu, H., et al., 2013. Rab11-FIP1C and Rab14 direct plasma membrane sorting and particle incorporation of the HIV-1 envelope glycoprotein complex. *PLoS Pathogens* 9, e1003278.
- de Renzis, S., Sonnichsen, B., Zerial, M., 2002. Divalent Rab effectors regulate the sub-compartmental organization and sorting of early endosomes. *Nature Cell Biology* 4, 124–133.
- Richards, P., Didszun, C., Campesan, S., et al., 2011. Dendritic spine loss and neurodegeneration is rescued by Rab11 in models of Huntington's disease. *Cell Death & Differentiation* 18, 191–200.
- Roland, J.T., Kenworthy, A.K., Peranen, J., Caplan, S., Goldenring, J.R., 2007. Myosin Vb interacts with Rab8a on a tubular network containing EHD1 and EHD3. *Molecular Biology of the Cell* 18, 2828–2837.
- Roland, J.T., Lapiere, L.A., Goldenring, J.R., 2009. Alternative splicing in class v myosins determines association with rab10. *Journal of Biological Chemistry* 284, 1213–1223.
- Schiel, J.A., Simon, G.C., Zaharris, C., et al., 2012. FIP3-endosome-dependent formation of the secondary ingression mediates ESCRT-III recruitment during cytokinesis. *Nature Cell Biology* 14, 1068–1078.
- Schonteich, E., Wilson, G.M., Burden, J., et al., 2008. The Rip11/Rab11-FIP5 and kinesin II complex regulates endocytic protein recycling. *Journal of Cell Science* 121, 3824–3833.
- Seixas, E., Ramalho, J.S., Mota, L.J., Barral, D.C., Seabra, M.C., 2012. Bacteria and protozoa differentially modulate the expression of Rab proteins. *PLoS One* 7, e39858.
- Sharma, M., Giridharan, S.S., Rahajeng, J., Naslavsky, N., Caplan, S., 2009. MICAL-1 links EHD1 to tubular recycling endosomes and regulates receptor recycling. *Molecular Biology of the Cell* 20, 5181–5194.
- Sheach, L.A., Adeney, E.M., Kucukmetin, A., et al., 2009. Androgen-related expression of G-proteins in ovarian cancer. *British Journal of Cancer* 101, 498–503.
- Simon, G.C., Prekeris, R., 2008. The role of FIP3-dependent endosome transport during cytokinesis. *Communicative & Integrative Biology* 1, 132–133.
- van der Sluijs, P., Hull, M., Webster, P., et al., 1992. The small GTP-binding protein rab4 controls an early sorting event on the endocytic pathway. *Cell* 70, 729–740.
- Sonnichsen, B., Renzis, D.E., Nielsen, S., Rietdorf, E., Zerial, M., J., 2000. Distinct membrane domains on endosomes in the recycling pathway visualized by multicolor imaging of Rab4, Rab5, and Rab11. *Journal of Cell Biology* 149, 901–914.
- Stein, M.P., Muller, M.P., Wandinger-Ness, A., 2012. Bacterial pathogens commandeer Rab GTPases to establish intracellular niches. *Traffic* 13, 1565–1588.
- Steinert, J.R., Campesan, S., Richards, P., et al., 2012. Rab11 rescues synaptic dysfunction and behavioural deficits in a *Drosophila* model of Huntington's disease. *Human Molecular Genetics* 21, 2912–2922.
- Stockli, J., Fazakerley, D.J., James, D.E., 2011. GLUT4 exocytosis. *Journal of Cell Science* 124, 4147–4159.
- Thompson, A., Nessler, R., Wisco, D., et al., 2007. Recycling endosomes of polarized epithelial cells actively sort apical and basolateral cargos into separate subdomains. *Molecular Biology of the Cell* 18, 2687–2697.
- Ullrich, O., Reinsch, S., Urbe, S., Zerial, M., Parton, R.G., 1996. Rab11 regulates recycling through the pericentriolar recycling endosome. *Journal of Cell Biology* 135, 913–924.
- Uytterhoeven, V., Kuenen, S., Kasprzewicz, J., Miskiewicz, K., Verstreken, P., 2011. Loss of skywalker reveals synaptic endosomes as sorting stations for synaptic vesicle proteins. *Cell* 145, 117–132.
- Varthakavi, V., Smith, R.M., Martin, K.L., et al., 2006. The pericentriolar recycling endosome plays a key role in Vpu-mediated enhancement of HIV-1 particle release. *Traffic* 7, 298–307.
- Vitale, G., Rybin, V., Christoforidis, S., et al., 1998. Distinct Rab-binding domains mediate the interaction of Rabaptin-5 with GTP-bound Rab4 and Rab5. *EMBO Journal* 17, 1941–1951.
- Wang, Z., Edwards, J.G., Riley, N., et al., 2008. Myosin Vb mobilizes recycling endosomes and AMPA receptors for postsynaptic plasticity. *Cell* 135, 535–548.
- Weigert, R., Yeung, A.C., Li, J., Donaldson, J.G., 2004. Rab22a regulates the recycling of membrane proteins internalized independently of clathrin. *Molecular Biology of the Cell* 15, 3758–3770.
- Wilson, G.M., Fielding, A.B., Simon, G.C., et al., 2005. The FIP3-Rab11 protein complex regulates recycling endosome targeting to the cleavage furrow during late cytokinesis. *Molecular Biology of the Cell* 16, 849–860.
- Yamamoto, H., Koga, H., Katoh, Y., et al., 2010. Functional cross-talk between Rab14 and Rab4 through a dual effector, RUFY1/Rabip4. *Molecular Biology of the Cell* 21, 2746–2755.
- Zacchi, P., Stenmark, H., Parton, R.G., et al., 1998. Rab17 regulates membrane trafficking through apical recycling endosomes in polarized epithelial cells. *Journal of Cell Biology* 140, 1039–1053.
- Zhang, J., Fonovic, M., Suyama, K., Bogoy, M., Scott, M.P., 2009. Rab35 controls actin bundling by recruiting fascin as an effector protein. *Science* 325, 1250–1254.
- Zhang, X.M., Ellis, S., Sriratanana, A., Mitchell, C.A., Rowe, T., 2004. Sec15 is an effector for the Rab11 GTPase in mammalian cells. *Journal of Biological Chemistry* 279, 43027–43034.

Endosome to Lysosome Transport

C Buccì and M Stasi, University of Salento, Lecce, Italy

© 2016 Elsevier Inc. All rights reserved.

Introduction

Endocytic cargo internalized from the plasma membrane to early endosomes (EEs) has to be transported to late endosomes (LEs) and then to lysosomes in order to be degraded. The endosome to lysosome pathway consists in a complex carefully regulated series of events important to control degradation of molecules. Different cellular pathways depend on this transport. For instance membrane signaling receptors, such as epidermal growth factor receptor (EGFR), are activated by ligand binding and, as they execute their function, their stimulation must be reduced in order to prevent aberrant signaling. Indeed, termination of these signaling pathways is often mediated by downregulation of membrane receptors accomplished by targeting them to degradation compartments (Ceresa and Schmid, 2000). Notably, alterations in one or more steps of endosome to lysosome transport determine deregulation of cellular functions and are causative of several diseases, thus demonstrating the importance of this pathway. For instance, alterations of protein sorting to lysosomes cause mucopolipidosis II, also called Inclusion-cell (I-cell) disease (Olkonen and Ikonen, 2006).

Endocytosed cargo is delivered to endosomes where proteins destined for degradation are sorted and delivered to lysosomes. Sorting mechanisms for correct cargo shipment to lysosomes take place both in EEs, where membrane proteins are internalized in intraluminal vesicles destined to lysosomes, and in LEs, where hydrolases incoming from Golgi compartment are delivered. Maturation of EEs is a crucial step, which occurs in a short time and involves several steps, including acidification and modification of endosome membrane and microenvironment composition.

Several key regulators of these processes have been identified and among them a crucial role is played by Rab proteins, that coordinate virtually all steps of this pathway, and by molecular motors, such as dynein and kinesins that are responsible for movement of vesicles and organelles. Although many players in this pathway have been identified and their functions unraveled, several mechanisms underlying this pathway remain poorly understood and even definitions of endocytic structures are not completely clear. This is due to the lack of sufficient evidences and difficulties in the interpretation of a wide range of heterogeneous data. Indeed, the mechanism of transport of materials between endosomes and lysosomes is still controversial and several models about the biogenesis and the maturation of endosomes and lysosomes have been proposed. These models include maturation of endosome to lysosome, vesicular transport or kiss and run fusion between endosomes and lysosomes, and complete fusion between endosomes and lysosomes in order to form a hybrid organelle (Luzio *et al.*, 2007).

Here we describe each step of the endosome to lysosome transport, illustrating the mechanisms that control these processes.

Endocytic Vesicles and Early Endosomes

Cellular internalization of fluids, solutes, proteins, and plasma membrane components such as receptors occurs via endocytosis. The most characterized mechanism that drives endocytosis is the clathrin-dependent pathway (Conner and Schmid, 2003). In this pathway, the clathrin adapter protein 2 (AP2) complex recognizes tyrosine or dileucine motifs present in the cytoplasmic domain of cell-surface proteins directing them into clathrin-coated pits that pinch off the plasma membrane to form clathrin-coated vesicles (Kirchhausen, 1999). Endocytic vesicles also arise from clathrin-independent pathways such as, for instance, caveolar, GPI-enriched early endosomal compartments GEEC- and ARF6-dependent pathways (Conner and Schmid, 2003; Mayor and Pagano, 2007). Incoming materials deriving from clathrin-dependent or independent endocytosis are then delivered to EEs (Conner and Schmid, 2003; Mayor and Pagano, 2007; Figure 1).

EEs are tubulovesicular structures involved in the sorting of internalized molecules and thus also called sorting endosomes. Their volume and tubular membrane extension are relatively small until they receive endocytic vesicles. In a short period of time (5–15 minutes) EEs accept and accumulate cargo coming from vesicles. Membranes, solutes and fluids arrived into EEs are then recycled to the plasma membranes directly or via perinuclear recycling endosomes or, alternatively, are delivered to LEs and lysosomes for degradation (Luzio *et al.*, 2007; Figure 1). Delivery of endocytosed materials to lysosomes for degradation occurs after progressive acidification of endosomal organelles, formation of multi-vesicular bodies (MVBs) and LEs, recruitment of lysosomal hydrolases from Golgi, positioning of LEs in the perinuclear region and sorting of proteins destined for degradation. Importantly, recent findings demonstrate that fission of recycling tubules in EEs is important for subsequent acidification and maturation, including endosome motility, perinuclear localization, and degradation of proteins (Mesaki *et al.*, 2011; Figure 1).

Biogenesis of MVBs and Cargo Sorting

MVBs are organelles characterized by the accumulation of vesicles in their lumen; material destined for degradation compartments is delivered to MVBs and then reaches lysosomes. MVBs are often considered as newly formed LEs derived from maturation of EEs, but, according to another model, they are defined as transporter intermediates between early and late endosomes (Gruenberg and Stenmark, 2004). MVBs move from the periphery to the center of the cell along microtubules, and mature in LE; when MVBs have lost all recycling proteins and have acquired all the molecules and the characteristics indispensable for fusion (directly or via kiss and

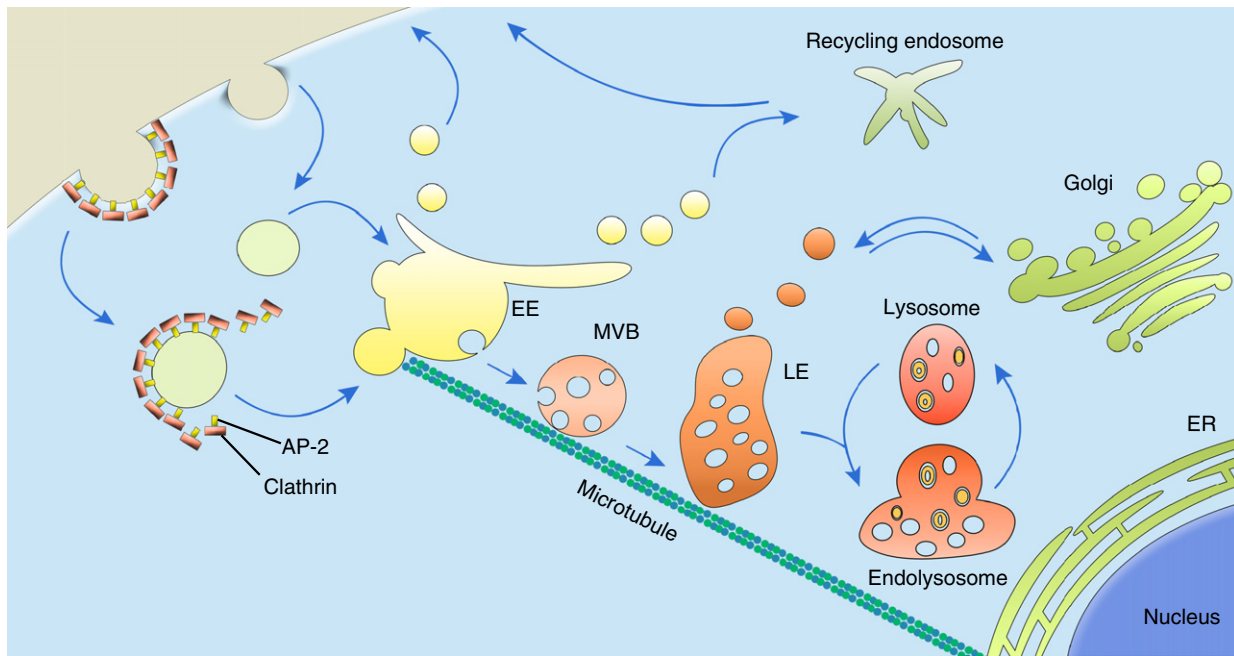


Figure 1 Endosome maturation. Vesicles derived from clathrin-dependent or independent endocytosis fuse with early endosomes (EEs). Membrane cargo destined for degradation is sorted in ILVs and delivered to lysosomes. Endosome to lysosome transport includes different maturation steps, comprising progressive acidification, MVBs formation, inclusion of material incoming from TGN and fusion of LEs with lysosome. During maturation, organelles move along microtubules from the periphery to the perinuclear region of the cell. Progressive acidification of organelles is represented by gradual color change from yellow (almost neutral pH, in EEs) to red (more acidic pH, in LEs, endolysosomes and lysosomes).

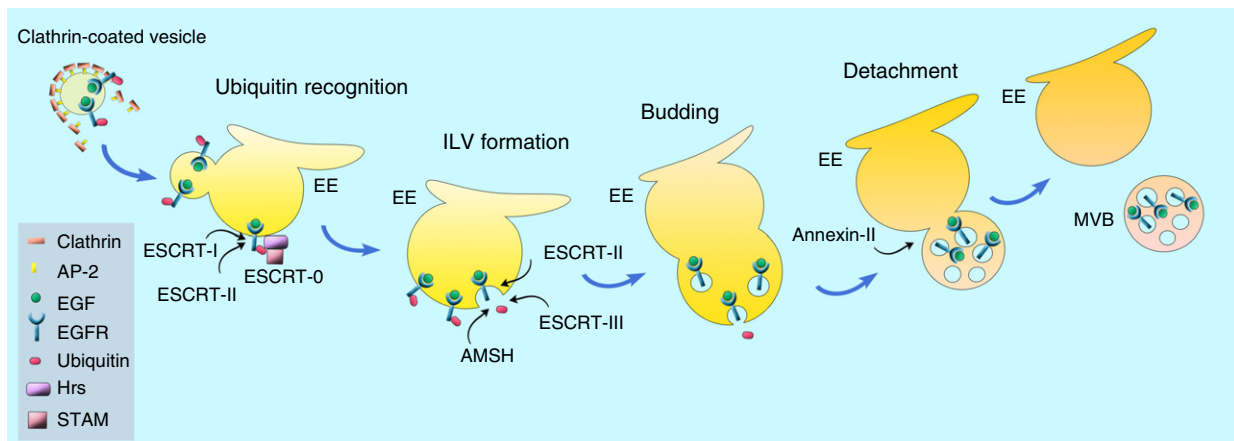


Figure 2 MVBs biogenesis. Endocytosed ubiquitylated receptors (e.g., EGFR) destined for degradation are delivered to EEs, where Hrs recognizes ubiquitin residues and promote ILV formation. When material destined for degradation is sorted in ILVs, MVBs bud from EEs and then detach by means of annexin-II. As described in the text, ESCRT machineries are responsible for this process.

run mechanisms) with lysosome, they are called LEs (Gruenberg and Stenmark, 2004; Woodman and Futter, 2008).

The mechanism that guides the biogenesis of MVBs is not yet completely clear and requires the sequential action of endosomal sorting complex required for transport (ESCRT) machineries (Figure 2). Membrane proteins destined for degradation in lysosomes are mono-ubiquitylated and then recognized by the hepatocyte growth factor receptor-regulated tyrosine kinase substrate (Hrs) that, acting with the signal-transducing adapter molecule (STAM) and binding to

phosphoinositol-3-phosphate (PI(3)P), sorts proteins into clathrin-coated microdomains of EEs delivering them to the degradative pathway (Jovic *et al.*, 2010). The events involving clathrin, the STAM-Hrs complex, also called ESCRT-0, and the signaling of PI(3)P represent the first step in the transition from EEs to MVBs involving invagination and formation of intraluminal vesicles (ILVs) (Gruenberg and Stenmark, 2004; Falguières *et al.*, 2009; Williams and Urbé, 2007). Hrs plays an essential role in recruiting ESCRT-I to endosomes (Bache *et al.*, 2003). ESCRT-I in mammalian cells is composed of TSG101,

VPS28, VPS37A-D, and of the orthologue of the yeast Mvb12 (Williams and Urbé, 2007). TSG101 regulates the maturation of EEs in MVBs controlling the formation of stable vacuolar domains (Razi and Futter, 2006). ESCRT-I recruits in turn ESCRT-II that, in mammalian cells, is composed of EAP30, EAP20 and EAP45 (Williams and Urbé, 2007). Both ESCRT-I and ESCRT-II, as well as Hrs, contain ubiquitin binding domains and recognize ubiquitylated proteins (Hurley, 2008). The last ESCRT complex is ESCRT-III and it is composed soluble monomers (CHMP-1, CHMP-2, CHMP-3, CHMP-4, CHMP-5, and CHMP-6) that assemble on the endosomal membrane forming an insoluble array (Williams and Urbé, 2007; Hurley, 2008). ESCRT-II recruits ESCRT-III by interaction between EAP-20 and CHMP-6 while CHMP-6 has been demonstrated to regulate cargo sorting (Yorikawa *et al.*, 2005). ESCRT-III complex recruits deubiquitinating enzymes (e.g., AMSH) that remove ubiquitin moieties after sorting and before degradation; this step is thought to be the last crucial step for the entry of cargoes in ILVs (Williams and Urbé, 2007; Stuffers *et al.*, 2009a). ESCRT-III also mediates membrane invagination, concentrates cargoes in MVBs and could play a role in fusion of MVBs with LEs (Babst *et al.*, 2002; Stuffers *et al.*, 2009a; Adell *et al.*, 2014). Afterwards this complex is disassembled by VPS4, an AAA + ATPases (Williams and Urbé, 2007).

Interestingly, degradation of membrane receptors is regulated by a feedback control mechanism mediated by the ESCRT itself. Indeed, ESCRT-I components are delivered together with the cargo to lysosomes after endosomal sorting in ILVs and this event prevents collateral degradation of surface receptors (Malerød *et al.*, 2011). Thus, stimulation of degradation of a specific membrane receptor causes also downregulation of ESCRT-I components inducing transient resistance to degradation of other membrane receptors (Malerød *et al.*, 2011).

Recent findings suggest that alternative ESCRT-independent mechanisms for formation of ILVs coexist in mammalian cells with the ESCRT-dependent one (Stuffers *et al.*, 2009b). These alternative pathways are driven by lipids comprising lysobisphosphatidic acid (LBPA), ceramide and phosphoinositol-3,5-bisphosphate (PI(3,5)P₂) that spontaneously trigger the formation of ILVs by acting with their interactors (Woodman and Futter, 2008).

Several other proteins contribute to the formation of MVBs. One example is represented by SNX-3 and SNX-12, two members of sorting nexins (SNXs) proteins that function in sorting and invagination of ILVs within MVBs (Pons *et al.*, 2008, 2012). Also the Rab-interacting lysosomal protein (RILP), an effector of the small GTPase Rab7, participates at the biogenesis of MVBs. Indeed, it was found to interact with the EAP-30 and EAP45 subunits of ESCRT-II and to be responsible for ILVs formation (Progida *et al.*, 2006, 2007; Wang and Hong, 2006). The detachment of newly formed MVBs is then regulated by annexin-II that is associated to the early endosomal membrane probably by a cholesterol-dependent mechanism (Gruenberg and Stenmark, 2004; Figure 2).

Late Endosomes

When MVBs form from EEs, they mature and move to the perinuclear region of the cell where they fuse with each other

to form larger bodies that in turn can fuse with lysosomes in order to assure the degradation of transported proteins (Gruenberg and Stenmark, 2004; Huotari and Helenius, 2011). MVBs and LEs are morphologically different: MVBs are spherical, with a diameter of 400–500 nm, while LEs are larger, more pleiomorphic than MVBs and have a different composition in lipids and proteins. LEs are characterized by the presence of numerous ILVs, of lysosomal-associated membrane protein 1 and 2 (LAMP1 and LAMP2), which are abundant also in lysosomes, of LBPA present mainly within internal membranes and of several acid hydrolases in the lumen. LEs have a pH between 6.0–4.8 and represent the sorting center from where the mannose-6-phosphate receptors (MPRs) are recycled to the *trans*-Golgi network (Gruenberg and Stenmark, 2004).

Lysosomal hydrolases are synthesized in the endoplasmic reticulum (ER) and are then transported to *cis*-Golgi network, where they are covalently modified by the addition of mannose-6-phosphate (M6P) groups and then transported in the *trans*-Golgi network (TGN) and delivered to the LEs (Coutinho *et al.*, 2012). Indeed, M6P-tagged proteins are recognized by MPRs in the TGN and then are sorted and delivered to the LEs by means of vesicles budding (Rouillé *et al.*, 2000; Coutinho *et al.*, 2012). There are two different MPRs: cation-dependent (CD) and cation-independent (CI). Even if they share similar sequences, these receptors are not synonymous components, but they bound different lysosomal enzymes (Ghosh *et al.*, 2003; Coutinho *et al.*, 2012).

Defects in targeting lysosomal enzymes cause mucopolisaccharidosis II (Olkonen and Ikonen, 2006).

Vesicles that work as intermediate carriers in this pathway are coated by a mechanism that involves adapter protein (AP) complexes, Golgi localized γ -ear-containing ARF-binding proteins (GGAs) and clathrin. While AP-1 mediates transport of MPR-bounded proteins from Golgi to LEs, transmembrane proteins that have to be delivered to LEs, including the glycoproteins LAMP-1 and LAMP-2, are sorted by AP-3 by recognition of tyrosine-based or dileucine based sorting signals. These proteins are then transported to LEs in vesicles that are different from those containing MPR-bounded proteins (Karlsson and Carlsson, 1998; Rouillé *et al.*, 2000). When vesicles containing MPR-bound proteins reach the LEs, the lower pH in the LEs causes the detachment of MPR; in this way hydrolases remain in the LEs and are then transported to lysosomes, while MPRs are recycled to the TGN by means of Rab9-TIP47 and a complex known as retromer. Other membrane proteins are not recycled but, instead, are routed to lysosomes (Rouillé *et al.*, 2000; Coutinho *et al.*, 2012). Even though acidic hydrolases coming from TGN seem to be mainly delivered to LEs, these proteins also arrive to EEs contributing to their maturation (Huotari and Helenius, 2011; Bonifacino and Rojas, 2006). Indeed, both EEs and LEs communicate with TGN in a bidirectional way, with acidic hydrolase receptors, transmembrane enzymes and SNAREs cycling between these compartments (Huotari and Helenius, 2011; Bonifacino and Rojas, 2006).

Maturation of endosomes requires changes in several regulatory proteins (Huotari and Helenius, 2011; Solinger and Spang, 2013). An important event is the so-called Rab5/Rab7 switch where Rab5, mostly expressed in EEs, is replaced by

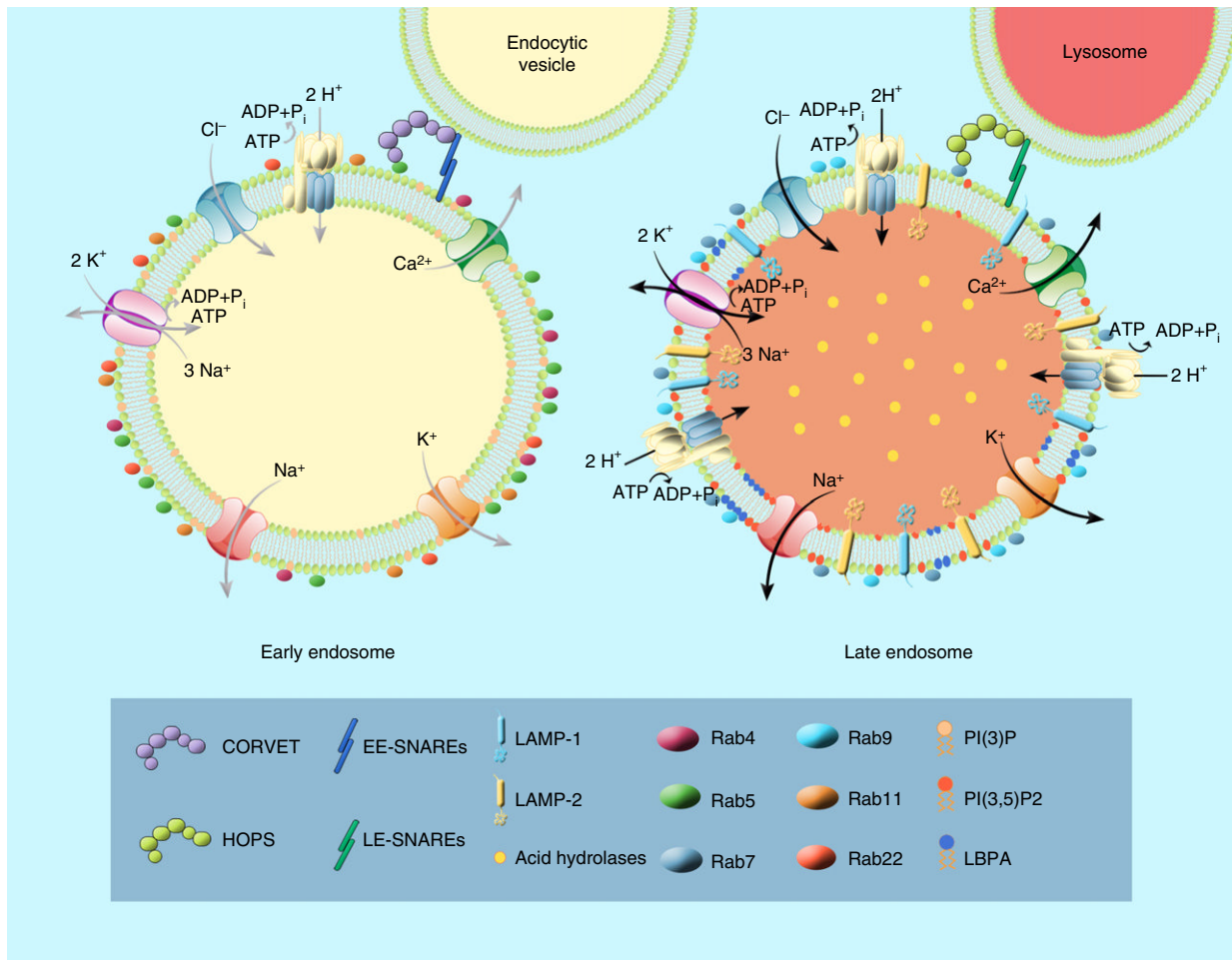


Figure 3 Differences between EEs and LEs. LEs are characterized by a more acidic environment compared with EEs. The acidic pH is assured by V-ATPase, which, hydrolyzing ATP, introduces protons in the LE lumen. The correct function of V-ATPase is allowed by maintenance of charge flow mediated by ion channels, which promote anion influx and cation efflux, and by Na^+/K^+ ATPase. These processes start in EEs (gray arrows), where pH range is comprised between 6.8 and 6.1, but are more intense in LEs (black arrows) whose pH is comprised in a range between 6.0 and 4.8. Early endosomal Rab proteins, membrane proteins and lipids, tethering and fusion machineries are replaced by late endosomal ones. Furthermore LEs are characterized by the inclusion of several acid hydrolases and transmembrane glycoproteins incoming from TGN. For clarity of representation in this figure ILVs are not included.

Rab7, mostly expressed in LEs (Figure 3). Also, PI(3)P abundant in early endosomal membranes is substituted by the PI(3,5)P2 and the class C core vacuole/endosome tethering factor (CORVET) is replaced by late endosomal/lysosomal homotypic fusion and vacuole protein sorting (HOPS) complex (Figure 3). In addition, early endosomal SNAREs (soluble NSF (N-ethylmaleimide (NEM)-sensitive factor) attachment protein receptors) important for fusion are substituted by late endosomal SNAREs (Figure 3).

Role of pH and Ion Concentrations in Endosomal Maturation

pH plays an essential role in the maturation of endosomes and in their fusion with lysosome. Acidification of endosomes is required for the release of lysosomal hydrolases incoming from TGN, for the genesis of ILVs, and for sorting and budding

of endosomal carrier vesicles that move from early to late endosomes (Cipriano *et al.*, 2008; Scott and Gruenberg, 2011). Furthermore acidic environment in lysosomes allows the activation of degradative enzymes and the transport of ions and small molecules across the lysosomal membrane (Cipriano *et al.*, 2008). pH in these organelles becomes progressively lower and this gradient plays an important role in the correct delivery of molecules. The progressive acidification of endosomes (starting from a 6.8–6.1 range in EEs and culminating in 6.0–4.8 range in LEs and about 4.5 in lysosomes), that drives molecules through the pathway, is assured by vacuolar (H^+) ATPases (V-ATPases) (Cipriano *et al.*, 2008). Alterations of pH in LEs and lysosomes have been detected in cancer cells (Cipriano *et al.*, 2008).

V-ATPases are composed of a cytoplasmic sector (V_1), that contains eight different subunits (A-H) and is involved in the binding and hydrolysis of ATP, and a membrane-integral sector (V_0), that contains six different subunits (a, c, c', c'', d, and e)

and regulates the proton transport into the lumen of endomembrane organelles (Cipriano *et al.*, 2008). These proton pumps are characterized by a rotary mechanism that starts in the V₁ domain after the hydrolysis of ATP and involves the D, F, and d subunits and the proteolipid ring (composed of c', c' and c'' subunits); when the proteolipid ring rotates, proton is transferred from the cytoplasm to the lumen, contributing to the acidification of the environment (Cipriano *et al.*, 2008).

Even though acidification is essential for organelle function, ion concentration is also important for maturation of endosomes (Figure 3). In the absence of a regulation of ion flux in the endosome, V-ATPase would activate a self-inhibiting mechanism because of the electrochemical gradient; ion channels (i.e., for Cl⁻ or other anion influx and Ca²⁺, Na⁺, K⁺ and other cation efflux), as well as Na⁺/K⁺ ATPase, existing in the endosomal membrane assure the correct maintenance of charge flow (Scott and Gruenberg, 2011; Huotari and Helenius, 2011). However, ion concentration itself can regulate endosomal maturation beyond acidification. Differences in ion concentrations have been detected in the transition from EE to LE and lysosome: Cl⁻, Ca²⁺, and K⁺ concentrations increase, while Na⁺ concentrations decrease (Scott and Gruenberg, 2011; Figure 3). In particular it is clear that Ca²⁺ release from endosomes and lysosomes is important for fission and fusion events occurring in the endolysosomal pathway; endosomes and lysosomes represent an important source of Ca²⁺ because of the high concentration of this cation in the lumen (~0.5 mM in lysosomes and between 0.003–2 mM in the endosome) (Shen *et al.*, 2011). Furthermore, an anion binding site was found in transferrin demonstrating that also Cl⁻ has an essential role: indeed, increasing concentrations of this anion, in addition to a low pH, play a role in the release of iron from transferrin in EEs (Byrne *et al.*, 2010).

Perinuclear Clustering of LE and Fusion Events

After maturation, LEs cluster in the perinuclear area, near the microtubule-organizing center (MTOC) and are subjected to homotypic (LE-LE) and heterotypic (LE-lysosome) fusion events (Luzio *et al.*, 2007; Huotari and Helenius, 2011).

There are three alternative hypotheses about the mixing between the endosome and lysosome content. The first one proposes a vesicular traffic between the two organelles; the second one, called “kiss and run”, considers that endosomes and lysosomes continuously undergo fusion and fission processes. The third hypothesis can be considered an extension of “kiss and run” and assumes that direct fusion between late endosomes and lysosomes occurs to form an hybrid organelle, called endolysosome, in which degradation takes place; after this fusion, lysosomes are re-formed from endolysosome and become storage organelle for lysosomal hydrolases and membrane components that can be reutilized (Storrie and Desjardins, 1996; Luzio *et al.*, 2000; Huotari and Helenius, 2011). The latter two hypotheses are the more validated and can coexist (Bright *et al.*, 2005).

Similar to other fusion events, fusion process is composed mainly of three steps: tethering, formation of *trans*-SNARE complex, that acts as a bridge between the two organelles, and,

finally, membrane fusion (Luzio *et al.*, 2007). Thus, LEs in order to fuse with each other, with lysosomes and probably with other organelles containing material destined for degradation, depend on NSF, SNAPs (soluble NSF-attachment proteins), SNAREs, tethering factors and Rabs (Huotari and Helenius, 2011).

Tethering is an essential event for fusion as it allows organelles to form links between each other (Luzio *et al.*, 2007). In particular, CORVET is a Rab5 effector required for fusion between Rab5 positive endosomes, while HOPS is involved in tethering at late-endocytic stages and acts as Rab7 effector after the Rab7/Rab5 switch (Bröcker *et al.*, 2010; Balderhaar and Ungermann, 2013). Indeed overexpression of Rab7, Vps18p, and Vam6p (which are HOPS components) causes clustering of late-endocytic structures, while dispersion of the organelles was observed after depletion of Vps18p (Bucci *et al.*, 2000; Caplan *et al.*, 2001; Poupon *et al.*, 2003). Tethering subunits of HOPS complex and Rab7 act together in order to recruit SNAREs and promote membrane fusion (Solinger and Spang, 2013). Thus SNAREs form a complex that is composed of three Q-SNAREs (Vamp7 for heterotypic and Vamp8 for homotypic fusion) on one membrane and one R-SNARE (Syntaxin-7 and -8) on the other membrane. After this stable four helix bundle complex is formed, the bilayer mixing reaction starts. Heterotypic fusion is not driven only by SNAREs, as also Rab effectors and Ca²⁺ released from endolysosomes are essential regulators (Luzio *et al.*, 2007; Shen *et al.*, 2011). After fusion the NSF ATPase and its cofactor SNAP cause the recycle of SNAREs (Balderhaar and Ungermann, 2013).

Role of Rab Proteins

Rab GTPases are key regulators of membrane traffic and thus play an essential role also in endosome to lysosome transport, being involved in several different steps of this pathway as described before. Rab5, in particular, regulates several steps in the EEs formation and maturation: it controls cargo selection in vesicles, generation of phosphatidylinositol-3-phosphate lipid on EEs, homotypic fusion, motility of EEs on actin and microtubules tracks and activates signaling pathways on EEs by recruiting its effectors. Instead, Rab4 and Rab11 promote recycling events (Jovic *et al.*, 2010). Rab7 is involved in cargo sorting at the level of EEs, but it plays its main role in late-endocytic stages being required for the maturation of EEs in LEs (the so-called Rab7/Rab5 switch) (Girard *et al.*, 2014). Rab7 also regulates transport from LEs to lysosomes and controls clustering and fusion of late-endocytic structures and lysosomes in the perinuclear region. In addition, several Rab7 effectors can regulate positively (RILP, component of the retromer and HOPS complexes, Rabring7) or negatively (Rubicon and PLEKHM1) maturation and late-endocytic events (Huotari and Helenius, 2011). Interestingly, Rab7, through its effector RILP, seems to regulate also the endosomal pH. Indeed, RILP controls assembly and function of the V-ATPase by interacting with the V1G1 subunit (De Luca *et al.*, 2014). In the transition step between EEs to LEs, Rab9, which plays an essential role in lysosome biogenesis and late endosome morphology, increases, while other Rabs, such as Rab4,

Rab11, and Rab22 are lost (Riederer *et al.*, 1994; Ganley *et al.*, 2004; Huotari and Helenius, 2011).

Role of Cytoskeletal Elements and of Molecular Motors

All cytoskeletal elements (microtubules, actin filaments and intermediate filaments) are involved in endosome to lysosome transport and myosin, kinesin and dynein motors cooperate to assure correct organelle's movements (Minin and Moldaver, 2008; Granger *et al.*, 2014). Table 1 provides a list of cytoskeletal elements and motors that acting with Rab proteins and other interactors, promote several functions in the endosomal-lysosomal transport.

During all maturation steps previously described, EEs and LEs move along microtubules, transported by molecular

motors (Figure 4). Kinesin motors mediate mainly transport from the minus to the plus-end of microtubules, while dynein promotes transport in the opposite direction (Hirokawa, 1998). Movement of EEs from the periphery to the MTOC is controlled by the dynein–dynactin complex, which can be recruited by the microtubule-binding protein CLIP-170 after docking of endosomes on microtubules (Valetti *et al.*, 1999). On the other hand transport of EEs in the opposite direction is regulated by a member of kinesin superfamily (KIF), KIF16B, in a process controlled by the Rab5 GTPase and by its effector the PI(3)P kinase VPS34 (Hoepfner *et al.*, 2005). On EEs Rab5 also acts by means of another effector: Huntingtin (Htt) associated protein (HAP40) that recruits Htt on EEs thus mediating the transport of these organelles. Moreover the Htt-HAP40-Rab5 complex mediates the switch from one type of filaments to another: indeed overexpression of HAP40, which limits recruitment of Htt on the membrane, causes the

Table 1 Role of cytoskeletal elements in endosome to lysosome transport

<i>Cytoskeletal elements</i>	<i>Organelle</i>	<i>Motor protein</i>	<i>Other proteins</i>	<i>Function</i>	<i>References</i>
Microtubules	EE	Dynein	Dynactin CLIP-170	Transport from minus to plus-end of microtubules	Valetti <i>et al.</i> (1999)
	EE	KIF16B	Rab5 VPS34	Transport from plus to minus-end of microtubules	Hoepfner <i>et al.</i> (2005)
	EE		Rab5 Htt HAP-40	Switch from microtubules to actin filaments	Pal <i>et al.</i> (2006)
	EE	KIF13A	Rab11	EE morphogenesis and motility	Delevoye <i>et al.</i> (2014)
	LE	Dynein	Dynactin Rab7 RILP ORP1L β III spectrin	Movement from the periphery to the perinuclear area	Aniento <i>et al.</i> (1993); Bananis <i>et al.</i> (2004); Jordens <i>et al.</i> (2001); Johansson <i>et al.</i> (2007)
	LE lysosome		RILP BLOC-3	Attachment on microtubule-dependent motors	Falcón-Pérez <i>et al.</i> (2005)
	LE lysosome		Htt	Facilitation of dynein-mediated trafficking and of association with microtubules	Caviston <i>et al.</i> (2011)
	LE	Kinesin	Rab7 FYCO1	Transport from the nucleus to the periphery	Pankiv <i>et al.</i> (2010)
Actin filaments	MVBs	Myosin α		Modulation of protein transport and probably of morphology	Salas-Cortes <i>et al.</i> (2005)
	Endosomes lysosomes	Myosin α		Tethering of organelles membrane; regulation of lysosomes movement acting with microtubules	Raposo <i>et al.</i> (1999); Cordonnier <i>et al.</i> (2001)
	EE	Myosin VI	LMTK2	Regulation of cargo delivery from EE to the endocytic recycling compartment	Chibalina <i>et al.</i> (2007)
	EE	Myosin Va	Rabs	Role in endocytic recycling events; maintenance of Rab11	Lindsay <i>et al.</i> (2013)

(Continued)

Table 1 Continued

Cytoskeletal elements	Organelle	Motor protein	Other proteins	Function	References
Intermediate filaments	LE lysosomes		AP-3	and Rab14-positive endosomes in the periphery compartments Peripherin, α -internexin and vimentin interact with AP-3; IFs provides positional information for AP-3, LE and lysosomes; IFs regulate AP-3 mediated sorting	Styers <i>et al.</i> (2004)
	LE lysosomes		Rab7	Peripherin and vimentin interacts with Rab7; possible role in positioning and transport of late-endocytic organelles	Cogli <i>et al.</i> (2013a,b)

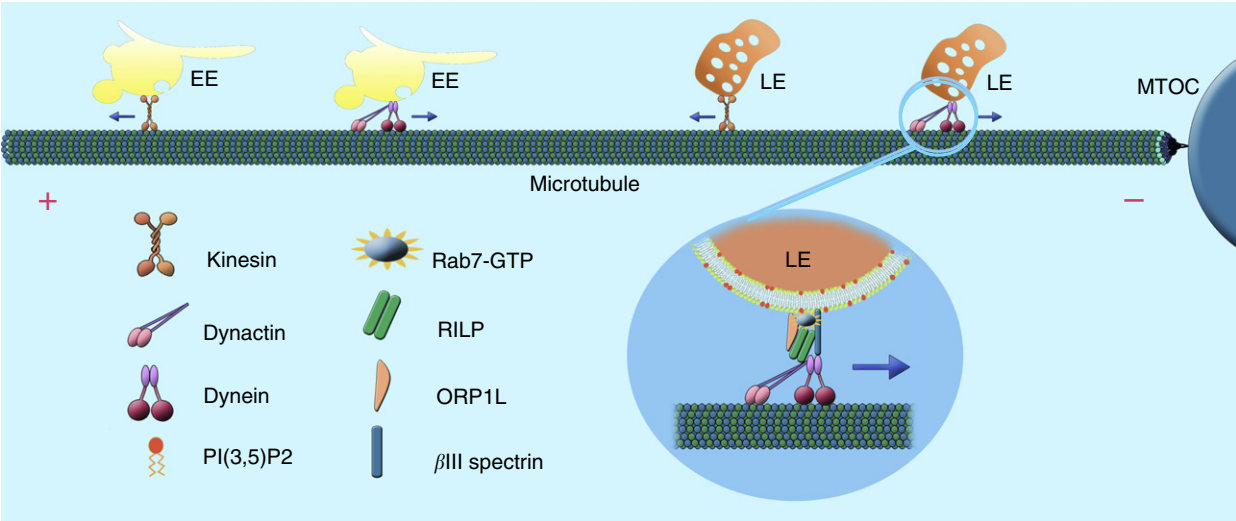


Figure 4 Transport on microtubules. Transport of EEs and LEs from the minus to the plus-end of microtubules is mediated by kinesin, whereas these organelles move to the microtubule-organizing center (MTOC) by means of dynein acting with dynactin. The transport of LEs to the MTOC mediated by Rab7 and RILP is represented in more details.

detachment of EEs from microtubules and the preferential binding of actin filaments (Pal *et al.*, 2006). Htt facilitates also dynein-mediated trafficking of LEs and lysosomes and regulates their association with microtubules (Caviston *et al.*, 2011). Htt mutations, causing defects in the movement of vesicles or organelles along cytoskeletal elements in late-endocytic transport, cause Huntington’s disease (Olkkonen and Ikonen, 2006). Recent findings demonstrated that another Rab protein, Rab11, regulates EEs morphogenesis and motility, interacting with the motor protein KIF13A (Delevoeye *et al.*, 2014).

LEs move to the perinuclear area where fusion events take place; this movement is mainly regulated by dynein (Aniento *et al.*, 1993; Bananis *et al.*, 2004). Attachment of LEs through dynein–dynactin complex is mediated by the Rab7 GTPase,

that, in its active form, cooperates with RILP, which is abundantly recruited on LEs membranes (Cantalupo *et al.*, 2001; Jordens *et al.*, 2001). Upon activation, Rab7 also interacts with ORP1L (oxysterol-binding protein-related protein 1 L) and thus a tripartite complex is formed. ORP1L transfers RILP to β III spectrin and this event is required for the translocation of late endosomes to MTOC (Johansson *et al.*, 2007). As well as RILP, BLOC-3 assures optimal attachment of late-endocytic organelles on the microtubules-dependent motors (Falcón-Pérez *et al.*, 2005). Recent evidences demonstrated that also an opposite-directed movement of LEs exists: indeed, interaction of Rab7 with FYVE and coiled-coil domain containing protein (FYCO) 1 regulates transport of LEs to the periphery (plus-end directed motility), probably by means of kinesin motors (Pankiv *et al.*, 2010).

Movement of endocytic organelles is mediated not only by microtubules, but also by actin cytoskeleton with myosin motors playing an essential role (DePina and Langford, 1999). Also, sorting of EGFR in MVBs destined for lysosomes degradation is mediated by its binding on actin filaments. This interaction could prevent the incorporation of receptor in the recycling pathway or could drive it in specific membrane compartments of endosomes destined for ILVs formation (Stoorvogel *et al.*, 2004). Traffic of protein cargo in MVBs is also mediated by myosin I that probably modulates the morphology of these organelles by interacting with actin (Salas-Cortes *et al.*, 2005). In fact, it was demonstrated that in mammalian cells endosomes and lysosomes are associated to myosin Ia that probably can tether organelle membranes to actin filaments, thus mediating the endosome-lysosome trafficking (Raposo *et al.*, 1999). Indeed this motor protein, working on actin filaments, was found to regulate the movement of lysosomes in cooperation with microtubules; these movement are rapid and direct along microtubules, while sometimes are short and random on actin filaments (Matteoni and Kreis, 1987; Prekeris *et al.*, 1999; Cordonnier *et al.*, 2001). Other myosins are instead important for endocytic recycling. Myosin VI, acting with Lemur tyrosine kinase2 (LMTK2), controls cargo delivery from EE to the endocytic recycling compartment while myosin Va interacts with several Rab proteins involved in recycling and it is required for the maintenance of Rab11 and Rab14-positive endosomes in the periphery of the cell (Lindsay *et al.*, 2013; Chibalina *et al.*, 2007).

As well as other cytoskeleton members, intermediate filament (IF) proteins control the degradative endocytic pathway (Styers *et al.*, 2004). Peripherin, α -internexin and vimentin bind to AP-3 complex, which is involved in transport between endosomes and lysosomes or related compartments (Styers *et al.*, 2004). IFs provide positional information for AP-3 and late-endocytic organelles. Furthermore, IFs regulate the sorting function of AP-3. In fact, AP-3 mediated sorting of LAMP1 and LAMP2 is prevented in absence of vimentin (Styers *et al.*, 2004). In support of these data we found that Rab7 interacts with and regulates the assembly of peripherin and vimentin (Cogli *et al.*, 2013a,b). As Rab7 is involved in endosome to lysosome transport, peripherin and vimentin, acting with this small GTPase, could play an important role in positioning and transport of late-endocytic organelles.

Conclusions and Perspectives

- Endocytosed cargo is delivered to EEs, where material is sorted and destined for degradation in lysosomes or recycled.
- Maturation of EEs in LEs is a complex process, not yet fully understood, comprising several modifications of the organelle environment. These events include formation of ILVs, progressive acidification, change of ion concentrations, substitution of membrane and luminal molecules, and inclusion of lysosomal proteins incoming from TGN.
- LEs cluster in the perinuclear area and fuse with each other or with lysosomes. These fusion events probably occur by “kiss and run” mechanisms that involve several proteins such as NSF, SNAPs and SNAREs.

- Rab proteins and their effectors play a key role, regulating tethering, maturation, transport, and fusion steps of endocytic organelles.
- Cytoskeletal elements are fundamental for organelle movement, organelle transport to the perinuclear region, and organelle positioning, and are regulated by Rab proteins.

Although in the past 15 years a huge amount of data have been produced on the molecular mechanism underlying the late steps of endocytosis (for instance unraveling mechanisms of sorting and degradation of membrane receptors by the ESCRT machineries), more work is required in order to better understand transport and relationships between endosomes and lysosomes. Indeed, several aspects of this transport including biogenesis of MVBs, LEs and lysosomes, role of cytoskeletal elements and signal transductions pathways need to be further investigated.

See also: Interorganellar Communication: Interplay and Processes: Endocytosis of Cargo Proteins: LDL

References

- Adell, M.A., Vogel, G.F., Pakdel, M., *et al.*, 2014. Coordinated binding of Vps4 to ESCRT-III drives membrane neck constriction during MVB vesicle formation. *Journal of Cell Biology* 205, 33–49.
- Aniento, F., Emans, N., Griffiths, G., Gruenberg, J., 1993. Cytoplasmic dynein-dependent vesicular transport from early to late endosomes. *Journal of Cell Biology* 123, 1373–1387.
- Babst, M., Katzmann, D.J., Estepa-Sabal, E.J., Meerloo, T., Emr, S.D., 2002. Escrt-III: An endosome-associated heterooligomeric protein complex required for MVB sorting. *Developmental Cell* 3, 271–282.
- Bache, K.G., Brech, A., Mehlum, A., Stenmark, H., 2003. Hrs regulates multivesicular body formation via ESCRT recruitment to endosomes. *Journal of Cell Biology* 162, 435–442.
- Balderhaar, H.J., Ungermann, C., 2013. CORVET and HOPS tethering complexes – coordinators of endosome and lysosome fusion. *Journal of Cell Science* 126, 1307–1316.
- Banakis, E., Nath, S., Gordon, K., *et al.*, 2004. Microtubule-dependent movement of late endocytic vesicles *in vitro*: Requirements for Dynein and Kinesin. *Molecular Biology of the Cell* 15 (8), 3688–3697.
- Bonifacio, J.S., Rojas, R., 2006. Retrograde transport from endosomes to the trans-Golgi network. *Nature Reviews Molecular Cell Biology* 7, 568–579.
- Bright, N.A., Gratian, M.J., Luzio, J.P., 2005. Endocytic delivery to lysosomes mediated by concurrent fusion and kissing events in living cells. *Current Biology* 15, 360–365.
- Bröcker, C., Engelbrecht-Vandré, S., Ungermann, C., 2010. Multisubunit tethering complexes and their role in membrane fusion. *Current Biology* 20, R943–R952.
- Bucci, C., Thomsen, P., Nicoziani, P., McCarthy, J., van Deurs, B., 2000. Rab7: A key to lysosome biogenesis. *Molecular Biology of the Cell* 11, 467–480.
- Byrne, S.L., Steere, A.N., Chasteen, N.D., Mason, A.B., 2010. Identification of a kinetically significant anion binding (KISAB) site in the N-lobe of human serum transferrin. *Biochemistry* 49, 4200–4207.
- Cantalupo, G., Alifano, P., Roberti, V., Bruni, C.B., Bucci, C., 2001. Rab-interacting lysosomal protein (RILP): The Rab7 effector required for transport to lysosomes. *EMBO Journal* 20, 683–693.
- Caplan, S., Hartnell, L.M., Aguilar, R.C., Naslavsky, N., Bonifacio, J.S., 2001. Human Vam6p promotes lysosome clustering and fusion *in vivo*. *Journal of Cell Biology* 154, 109–122.
- Caviston, J.P., Zajac, A.L., Tokito, M., Holzbaur, E.L., 2011. Huntingtin coordinates the dynein-mediated dynamic positioning of endosomes and lysosomes. *Molecular Biology of the Cell* 22, 478–492.
- Ceresa, B.P., Schmid, S.L., 2000. Regulation of signal transduction by endocytosis. *Current Opinion in Cell Biology* 12, 204–210.

- Chibalina, M.V., Seaman, M.N., Miller, C.C., Kendrick-Jones, J., Buss, F., 2007. Myosin VI and its interacting protein LMTK2 regulate tubule formation and transport to the endocytic recycling compartment. *Journal of Cell Science* 120, 4278–4288.
- Cipriano, D.J., Wang, Y., Bond, S., *et al.*, 2008. Structure and regulation of the vacuolar ATPases. *Biochimica et Biophysica Acta* 1777, 599–604.
- Cogli, L., Progidia, C., Bramato, R., Bucci, C., 2013b. Vimentin phosphorylation and assembly are regulated by the small GTPase Rab7a. *Biochimica et Biophysica Acta* 1833, 1283–1293.
- Cogli, L., Progidia, C., Thomas, C.L., *et al.*, 2013a. Charcot-Marie-Tooth type 2B disease-causing RAB7A mutant proteins show altered interaction with the neuronal intermediate filament peripherin. *Acta Neuropathologica* 125, 257–272.
- Conner, S.D., Schmid, S.L., 2003. Regulated portals of entry into the cell. *Nature* 422, 37–44.
- Cordonnier, M.N., Dauzonne, D., Louvard, D., Coudrier, E., 2001. Actin filaments and myosin I alpha cooperate with microtubules for the movement of lysosomes. *Molecular Biology of the Cell* 12, 4013–4029.
- Coutinho, M.F., Prata, M.J., Alves, S., 2012. Mannose-6-phosphate pathway: A review on its role in lysosomal function and dysfunction. *Molecular Genetics and Metabolism* 105, 542–550.
- De Luca, M., Cogli, L., Progidia, C., *et al.*, 2014. RILP regulates vacuolar ATPase through interaction with the V1G1 subunit. *Journal of Cell Science* 127, 2697–2708.
- Delevoe, C., Miserey-Lenkei, S., Montagnac, G., *et al.*, 2014. Recycling endosome tubule morphogenesis from sorting endosomes requires the kinesin motor KIF13A. *Cell Reports* 6, 445–454.
- DePina, A.S., Langford, G.M., 1999. Vesicle transport: The role of actin filaments and myosin motors. *Microscopy Research and Technique* 47, 93–106.
- Falcón-Pérez, J.M., Nazarian, R., Sabatti, C., Dell'Angelica, E.C., 2005. Distribution and dynamics of Lamp1-containing endocytic organelles in fibroblasts deficient in BLOC-3. *Journal of Cell Science* 118, 5243–5255.
- Falguieres, T., Luyet, P.P., Gruenberg, J., 2009. Molecular assemblies and membrane domains in multivesicular endosome dynamics. *Experimental Cell Research* 315, 1567–1573.
- Ganley, I.G., Carroll, K., Bittova, L., Pfeffer, S., 2004. Rab9 GTPase regulates late endosome size and requires effector interaction for its stability. *Molecular Biology of the Cell* 15, 5420–5430.
- Ghosh, P., Dahms, N.M., Kornfeld, S., 2003. Mannose 6-phosphate receptors: New twists in the tale. *Nature Reviews Molecular Cell Biology* 4, 202–212.
- Girard, E., Chmiest, D., Fournier, N., *et al.*, 2014. Rab7 is functionally required for selective cargo sorting at the early endosome. *Traffic* 15, 309–326.
- Granger, E., McNee, G., Allan, V., Woodman, P., 2014. The role of the cytoskeleton and molecular motors in endosomal dynamics. *Seminars in Cell and Developmental Biology* 31, 20–29.
- Gruenberg, J., Stenmark, H., 2004. The biogenesis of multivesicular endosomes. *Nature Reviews Molecular Cell Biology* 5, 317–323.
- Hirokawa, N., 1998. Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science* 279, 519–526.
- Hoepfner, S., Severin, F., Cabezas, A., *et al.*, 2005. Modulation of receptor recycling and degradation by the endosomal kinesin KIF16B. *Cell* 121, 437–450.
- Huotari, J., Helenius, A., 2011. Endosome maturation. *EMBO Journal* 30, 3481–3500.
- Hurley, J.H., 2008. ESCRT complexes and the biogenesis of multivesicular bodies. *Current Opinion in Cell Biology* 20, 4–11.
- Johansson, M., Rocha, N., Zwart, W., *et al.*, 2007. Activation of endosomal dynein motors by stepwise assembly of Rab7-RILP-p150Glued, ORP1L, and the receptor betall spectrin. *Journal of Cell Biology* 176, 459–471.
- Jordens, I., Fernandez-Borja, M., Marsman, M., *et al.*, 2001. The Rab7 effector protein RILP controls lysosomal transport by inducing the recruitment of dynein-dynactin motors. *Current Biology* 11, 1680–1685.
- Jovic, M., Sharma, M., Rahajeng, J., Caplan, S., 2010. The early endosome: A busy sorting station for proteins at the crossroads. *Histology and Histopathology* 25, 99–112.
- Karlsson, K., Carlsson, S.R., 1998. Sorting of lysosomal membrane glycoproteins lamp-1 and lamp-2 into vesicles distinct from mannose 6-phosphate receptor/gamma-adaptin vesicles at the trans-Golgi network. *Journal of Biological Chemistry* 273, 18966–18973.
- Kirchhausen, T., 1999. Adaptors for clathrin-mediated traffic. *Annual Review of Cell and Developmental Biology* 15, 705–732.
- Lindsay, A.J., Jollivet, F., Horgan, C.P., *et al.*, 2013. Identification and characterization of multiple novel Rab-myosin Va interactions. *Molecular Biology of the Cell* 24, 3420–3434.
- Luzio, J.P., Pryor, P.R., Bright, N.A., 2007. Lysosomes: Fusion and function. *Nature Reviews Molecular Cell Biology* 8, 622–632.
- Luzio, J.P., Rous, B.A., Bright, N.A., *et al.*, 2000. Lysosome-endosome fusion and lysosome biogenesis. *Journal of Cell Science* 113, 1515–1524.
- Malerød, L., Pedersen, N.M., Sem Wegner, C.E., *et al.*, 2011. Cargo-dependent degradation of ESCRT-I as a feedback mechanism to modulate endosomal sorting. *Traffic* 12, 1211–1226.
- Matteoni, R., Kreis, T.E., 1987. Translocation and clustering of endosomes and lysosomes depends on microtubules. *Journal of Cell Biology* 105, 1253–1265.
- Mayor, S., Pagano, R.E., 2007. Pathways of clathrin-independent endocytosis. *Nature Reviews Molecular Cell Biology* 8, 603–612.
- Masaki, K., Tanabe, K., Obayashi, M., Oe, N., Takei, K., 2011. Fission of tubular endosomes triggers endosomal acidification and movement. *PLoS One* 10, e19764.
- Minin, A.A., Moldaver, M.V., 2008. Intermediate vimentin filaments and their role in intracellular organelle distribution. *Biochemistry (Mosc)* 73, 1453–1466.
- Oikkonen, V.M., Ikonen, E., 2006. When intracellular logistics fails — Genetic defects in membrane trafficking. *Journal of Cell Science* 119, 5031–5045.
- Pal, A., Severin, F., Lommer, B., Shevchenko, A., Zerial, M., 2006. Huntingtin-HAP40 complex is a novel Rab5 effector that regulates early endosome motility and is up-regulated in Huntington's disease. *Journal of Cell Biology* 172, 605–618.
- Pankiv, S., Alemu, E.A., Brech, A., *et al.*, 2010. FYCO1 is a Rab7 effector that binds to LC3 and PI3P to mediate microtubule plus end-directed vesicle transport. *Journal of Cell Biology* 188, 253–269.
- Pons, V., Luyet, P.P., Morel, E., *et al.*, 2008. Hrs and SNX3 functions in sorting and membrane invagination within multivesicular bodies. *PLoS Biology* 6, e214.
- Pons, V., Ustunel, C., Rolland, C., *et al.*, 2012. SNX12 role in endosome membrane transport. *PLoS One* 7, e38949.
- Poupon, V., Stewart, A., Gray, S.R., Piper, R.C., Luzio, J.P., 2003. The role of mVps18p in clustering, fusion, and intracellular localization of late endocytic organelles. *Molecular Biology of the Cell* 14, 4015–4027.
- Prekeris, R., Foletti, D.L., Scheller, R.H., 1999. Dynamics of tubulovesicular recycling endosomes in hippocampal neurons. *Journal of Neuroscience* 19, 10324–10337.
- Progidia, C., Malerød, L., Stuffers, S., *et al.*, 2007. RILP is required for the proper morphology and function of late endosomes. *Journal of Cell Science* 120, 3729–3737.
- Progidia, C., Spinosa, M.R., De Luca, A., Bucci, C., 2006. RILP interacts with the VPS22 component of the ESCRT-II complex. *Biochemical and Biophysical Research Communications* 347, 1074–1079.
- Raposo, G., Cordonnier, M.N., Tenza, D., *et al.*, 1999. Association of myosin I alpha with endosomes and lysosomes in mammalian cells. *Molecular Biology of the Cell* 10, 1477–1494.
- Razi, M., Futter, C.E., 2006. Distinct roles for Tsg101 and Hrs in multivesicular body formation and inward vesiculation. *Molecular Biology of the Cell* 17, 3469–3483.
- Riederer, M.A., Soldati, T., Shapiro, A.D., Lin, J., Pfeffer, S.R., 1994. Lysosome biogenesis requires Rab9 function and receptor recycling from endosomes to the trans-Golgi network. *Journal of Cell Biology* 125, 573–582.
- Rouillé, Y., Rohn, W., Hoflack, B., 2000. Targeting of lysosomal proteins. *Seminars in Cell and Developmental Biology* 11, 165–171.
- Salas-Cortes, L., Ye, F., Tenza, D., *et al.*, 2005. Myosin Ib modulates the morphology and the protein transport within multi-vesicular sorting endosomes. *Journal of Cell Science* 118, 4823–4832.
- Scott, C.C., Gruenberg, J., 2011. Ion flux and the function of endosomes and lysosomes: pH is just the start: The flux of ions across endosomal membranes influences endosome function not only through regulation of the luminal pH. *Bioessays* 33, 103–110.
- Shen, D., Wang, X., Xu, H., 2011. Pairing phosphoinositides with calcium ions in endolysosomal dynamics: Phosphoinositides control the direction and specificity of membrane trafficking by regulating the activity of calcium channels in the endolysosomes. *Bioessays* 33, 448–457.
- Solinger, J.A., Spang, A., 2013. Tethering complexes in the endocytic pathway: CORVET and HOPS. *FEBS Journal* 280, 2743–2757.
- Stoorvogel, W., Kerstens, S., Fritzsche, I., *et al.*, 2004. Sorting of ligand-activated epidermal growth factor receptor to lysosomes requires its actin-binding domain. *Journal of Biological Chemistry* 279, 11562–11569.
- Storrie, B., Desjardins, M., 1996. The biogenesis of lysosomes: Is it a kiss and run, continuous fusion and fission process? *Bioessays* 18, 895–903.
- Stuffers, S., Brech, A., Stenmark, H., 2009a. ESCRT proteins in physiology and disease. *Experimental Cell Research* 315, 1619–1626.

- Stuffers, S., Sem Wegner, C., Stenmark, H., Brech, A., 2009b. Multivesicular endosome biogenesis in the absence of ESCRTs. *Traffic* 10, 925–937.
- Styers, M.L., Salazar, G., Love, R., *et al.*, 2004. The endo-lysosomal sorting machinery interacts with the intermediate filament cytoskeleton. *Molecular Biology of the Cell* 15, 5369–5382.
- Valetti, C., Wetzel, D.M., Schrader, M., *et al.*, 1999. Role of dynactin in endocytic traffic: Effects of dynamin overexpression and colocalization with CLIP-170. *Molecular Biology of the Cell* 10, 4107–4120.
- Wang, T., Hong, W., 2006. RILP interacts with VPS22 and VPS36 of ESCRT-II and regulates their membrane recruitment. *Biochemical and Biophysical Research Communications* 350, 413–423.
- Williams, R.L., Urbé, S., 2007. The emerging shape of the ESCRT machinery. *Nature Reviews Molecular Cell Biology* 8, 355–368.
- Woodman, P.G., Futter, C.E., 2008. Multivesicular bodies: Co-ordinated progression to maturity. *Current Opinion in Cell Biology* 20, 408–414.
- Yorikawa, C., Shibata, H., Waguri, S., *et al.*, 2005. Human CHMP6, a myristoylated ESCRT-III protein, interacts directly with an ESCRT-II component EAP20 and regulates endosomal cargo sorting. *Biochemical Journal* 387, 17–26.
- Babst, M., 2005. A protein's final ESCRT. *Traffic* 6, 2–9.
- Loubéry, S., Wilhelm, C., Hurbain, I., *et al.*, 2008. Different microtubule motors move early and late endocytic compartments. *Traffic* 9, 492–509.
- Metcalf, D., Isaacs, A.M., 2010. The role of ESCRT proteins in fusion events involving lysosomes, endosomes and autophagosomes. *Biochemical Society Transactions* 38, 1469–1473.
- Murray, J.W., Wolkoff, A.W., 2003. Roles of the cytoskeleton and motor proteins in endocytic sorting. *Advanced Drug Delivery Reviews* 55, 1385–1403.
- Pfeffer, S.R., 2013. Rab GTPase regulation of membrane identity. *Current Opinion in Cell Biology* 25, 414–419.
- Styers, M.L., Kowalczyk, A.P., Faundez, V., 2005. Intermediate filaments and vesicular membrane traffic: The odd couple's first dance? *Traffic* 6, 59–65.
- Toivola, D.M., Tao, G.Z., Habtezion, A., Liao, J., Omary, M.B., 2005. Cellular integrity plus: Organelle-related and protein-targeting functions of intermediate filaments. *Trends in Cell Biology* 15, 608–617.
- Zhang, M., Chen, L., Wang, S., Wang, T., 2009. Rab7: Roles in membrane trafficking and disease. *Bioscience Reports* 29, 193–209.

Further Reading

- Abe, K., Puertollano, R., 2011. Role of TRP channels in the regulation of the endosomal pathway. *Physiology (Bethesda)* 26, 14–22.

Endocytosis of Cargo Proteins: LDL

LM Traub, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Avidity A phenomenon often seen during assembly of large macromolecular complexes, whereby numerous weak-binding affinities simultaneously produce a strong but transient interaction. Also termed coincidence detection or a dual key strategy.

Clathrin-coated vesicle A class of membrane-bounded intracellular transport carriers that are temporarily encased within a polyhedral clathrin protein lattice. The coat functions to both deform the membrane into a ~100 nm-sized vesicle and to select the cargo packaged into the vesicle.

Integrins A set of heterodimeric transmembrane cell surface receptors that link the extracellular matrix to the intracellular actin-based cytoskeleton.

Interaction hub A designation for a protein or assembled protein complex that is able to interact with numerous distinct partner proteins or lipids, either one at a time or simultaneously.

Selectin A group of Ca^{2+} -dependent, carbohydrate recognizing transmembrane adhesion, and signaling proteins.

Introduction

Cholesterol is a fundamentally important constituent of animal cell lipid bilayers. Being the only inflexible heterocyclic (sterol) lipid in the membrane matrix, the pathway for cholesterol biosynthesis is distinct from that of phospholipids and glycosphingolipids (Holthuis and Menon, 2014). Cellular production of the 27-carbon-containing cholesterol molecule begins with acetyl-Coenzyme A and depends on 26 sequential enzymatic steps. In complex organisms, although each cell type generally has the intrinsic capability to assemble cholesterol from an acetate precursor, supply of cholesterol through the diet is often preferable. In fact, when cholesterol levels become limiting in cells, an elegant molecular response is initiated that not only induces coordinated transcription and translation of the enzymatic suite necessary for cholesterol biosynthesis, but also drives parallel upregulation of the principal transmembrane receptor required for internalization of dietary-derived cholesterol (Goldstein and Brown, 2009). This permits cells to scavenge any available extracellular cholesterol source as effectively as possible; in fact, this is a clinically important aspect of therapeutic statin administration.

This article deals with the process of external cholesterol uptake at the cellular level. The operation depends upon clathrin-mediated endocytosis, an evolutionarily ancient mode of internalization of extracellular macromolecules and surface membrane constituents within small (~100 nm), short-lived, vesicular shuttles. Characterization of the mechanistic basis of the internalization and intracellular mobilization of plasma-borne cholesterol mapped the basic features of the endocytic pathway (Brown and Goldstein, 1979; Goldstein and Brown, 2009), and provided the basis for further structural and functional dissection of numerous endocytic trafficking events. The focus here is primarily on the human low density lipoprotein (LDL) receptor; over 1000 distinct mutations in the *LDLR* gene have been characterized in familial hypercholesterolemia patients (Fokkema *et al.*, 2005; Hobbs *et al.*, 1990) and interesting new details about LDL receptor

trafficking have come recently from three clinically related human disease states (Soutar and Naoumova, 2007; Sharpe *et al.*, 2014).

Low Density Lipoprotein Biosynthesis

While the LDL receptor is ubiquitously expressed, production of plasma LDL is linked directly to the normal operation of the liver in placental mammals. Initially, fatty acids, in the form of triacylglycerols, and exogenous esterified cholesterol, derived largely from dietary chylomicron remnants, are assembled into a monolayer-bounded lipoprotein particle within the lumen of the hepatocyte endoplasmic reticulum. The particle is lipidated around a core of a single apolipoprotein B100 (apoB100) molecule, a necessary and very large (>5000 kDa) transport protein (Rutledge *et al.*, 2010). The smaller exchangeable apolipoprotein E (apoE) protein (36 kDa) is also co-translationally translocated into the lumen of the hepatocyte endoplasmic reticulum. The incorporation of apoE into nascent apoB100-positive lipoprotein particles enhances biosynthesis and secretion. Upon exocytic release into the circulation, the macromolecular assemblage (30–50 nm in diameter) is designated very low density lipoprotein (VLDL). VLDL supplies peripheral tissues (primarily muscle) with fatty acids as an energy source, and allows adipocytes to stockpile lipid from the hydrophobic triacylglyceride/esterified cholesterol store within intracellular droplets. This continuous remodeling of circulating VLDL particles by capillary-bed lipoprotein lipases, with concomitant loss of associated apoE and the exchangeable lipoprotein lipase cofactor apoC-II (both salvaged by plasma high density lipoprotein (HDL)), alters the density of the particle, ultimately generating LDL that is relatively depleted of triglycerides but substantially enriched with cholesteryl esters within the hydrophobic core. The now ~20-nm-sized LDL particle contains apoB100 as the only protein cofactor and binds exclusively to the LDL receptor. The circulating macromolecular particle gains access to the cell

interior from body fluids through the LDL receptor on the plasma membrane physically engaging apoB100 with nanomolar affinity, followed rapidly by receptor-mediated endocytosis (Goldstein and Brown, 2009).

The LDL Receptor

Receptor Structure

The LDL receptor is a type I transmembrane glycoprotein (Brown and Goldstein, 1979; Goldstein and Brown, 2009;

Hobbs *et al.*, 1990; Soutar and Naoumova, 2007). Because of the clinical significance of this protein, it represents the founding member of a larger LDL receptor superfamily of proteins, characterized by a shared modular architecture (Table 1). By far the bulk of the mass of these receptors is located extracellularly, and the LDL receptor is composed of multiple, tandemly arrayed functional luminal modules. An extreme N-terminal signal peptide begins the immature polypeptide but is removed in the endoplasmic reticulum. The mature protein contains a set of seven serially arranged, nonidentical cysteine-rich, Ca^{2+} -coordinating the so-called LDL receptor type A (LA) or complement-type ligand-binding

Table 1 Human genes and proteins related to LDL receptor biology

Gene	Protein	Function	UniProt accession number
<i>LDLR</i>	Low density lipoprotein receptor	Internalization of LDL	P01130
<i>APOB</i>	Apolipoprotein B100	Lipoprotein cofactor	P04114
<i>APOE</i>	Apolipoprotein E	Lipoprotein cofactor	P02649
<i>VLDLR</i>	Very low density lipoprotein receptor	Internalization of VLDL, Reelin signaling	P98155
<i>LRP8/APOER2</i>	Low density lipoprotein receptor-related protein 8/apoE receptor 2	Reelin signaling	Q14114
<i>MESDC2</i>	LDLR chaperone MESD/boca	Endoplasmic reticulum molecular chaperone	Q14696
<i>LRPAP1</i>	α 2-macroglobulin receptor-associated protein	Endoplasmic reticulum molecular chaperone	P30533
<i>SELP</i>	P-selectin	Cell adhesion receptor	P16109
<i>AMN1</i>	Protein AMN1 homolog/amnionless	Molecular chaperone/scavenger receptor subunit	Q8IY45
<i>CLTC</i>	Clathrin heavy chain 1	Coat protein	Q00610
<i>CLTA</i>	Clathrin light chain A	Coat protein subunit	P09496
<i>CLTB</i>	Clathrin light chain B	Coat protein subunit	P09497
<i>AP2A1</i>	AP-2 complex α -1/ α_A subunit	Coat component/sorting adaptor	O95782
<i>AP2A2</i>	AP-2 complex α -2/ α_C subunit	Coat component/sorting adaptor	O94973
<i>AP2B1</i>	AP-2 complex β 2 subunit	Coat component/sorting adaptor	P63010
<i>AP2M1</i>	AP-2 complex μ 2 subunit	Coat component/sorting adaptor	Q96CW1
<i>AP2S1</i>	AP-2 complex σ 2 subunit	Coat component/sorting adaptor	P53680
<i>LDLRAP1/ARH</i>	Low density lipoprotein receptor adapter protein 1/autosomal recessive hypercholesterolemia protein	Clathrin-associated sorting protein/cargo specific adaptor	Q5SW96
<i>DAB2</i>	Disabled homolog 2	Clathrin-associated sorting protein/cargo specific adaptor	P98082
<i>DAB1</i>	Disabled homolog 1	Signal transduction adapter	O75553
<i>ITGB1</i>	Integrin β 1	Extracellular matrix receptor	P05556
<i>SNX17</i>	Sorting nexin 17	Cargo-selective adaptor	Q15036
<i>NPC1</i>	Niemann-Pick C1 protein	Cholesterol transport transmembrane protein	O15118
<i>NPC2</i>	Niemann-Pick disease type C2 protein/Epididymal secretory protein E1	Luminal cholesterol transport protein	P61916
<i>NR1H3</i>	Oxysterols receptor LXR- α /Liver X receptor α	Transcription factor	Q13133
<i>NR1H2</i>	Oxysterols receptor LXR- β /Liver X receptor β	Transcription factor	P55055
<i>UBC</i>	Polyubiquitin C/ubiquitin	Posttranslational degradation signaling molecule	P0CG48
<i>MYLIP/IDOL</i>	E3 ubiquitin-protein ligase MYLIP/Inducible degrader of the LDL receptor	E3 ubiquitin ligase	Q8WY64
<i>EPN1</i>	Epsin-1/EPS-15-interacting protein 1	Clathrin-associated sorting protein/cargo specific adaptor	Q9Y6I3
<i>UBE2D1</i>	Ubiquitin-conjugating enzyme E2 D1	E2 ubiquitin-conjugating enzyme	UBE2D1
<i>PCSK9</i>	Proprotein convertase subtilisin/kexin type 9	Receptor degradation regulator	Q8NBP7

repeats, three epidermal growth factor (EGF) precursor-homology repeats (designated EGF-A–EGF-C) with a so-called YWTD domain, which folds into a six-bladed β -propeller, interposed between the second and third EGF repeat, preceding a serine/threonine-enriched juxtamembrane O-linked oligosaccharide domain. This extracellular assemblage is followed by the transmembrane segment of 22 nonpolar amino acids and, finally, by a 50-residue cytosolic domain (Figure 1). The VLDL receptor and the apoE receptor 2 (apoER2) are structurally highly homologous to the LDL receptor. Other members of the superfamily, including the LDL receptor-related protein 1 (LRP1) and the giant renal scavenger receptor megalin (also designated LRP2), have homologous but distinct combinations of the luminal repeats (Figure 1). In the

majority of cases, the precise role of each of the domain units within these mosaic receptor assemblages is not definitively established. Certainly, not all of the repetitive units contact the ligand(s) simultaneously. Evolutionarily related and functionally comparable receptors are also found in chordates and invertebrates; in insects (Figure 1) they are involved in yolk accumulation in developing oocytes and in shuttling the invertebrate hemolymph lipoprotein-like particle termed lipophorin (Rodenburg and Van Der Horst, 2005).

The complex ectodomain architecture of these receptors requires at least two dedicated endoplasmic reticulum molecular chaperones. The Boca/mesoderm development (MESD) protein is necessary for the proper folding of the numerous β -strands that make up the toroidal β -propeller

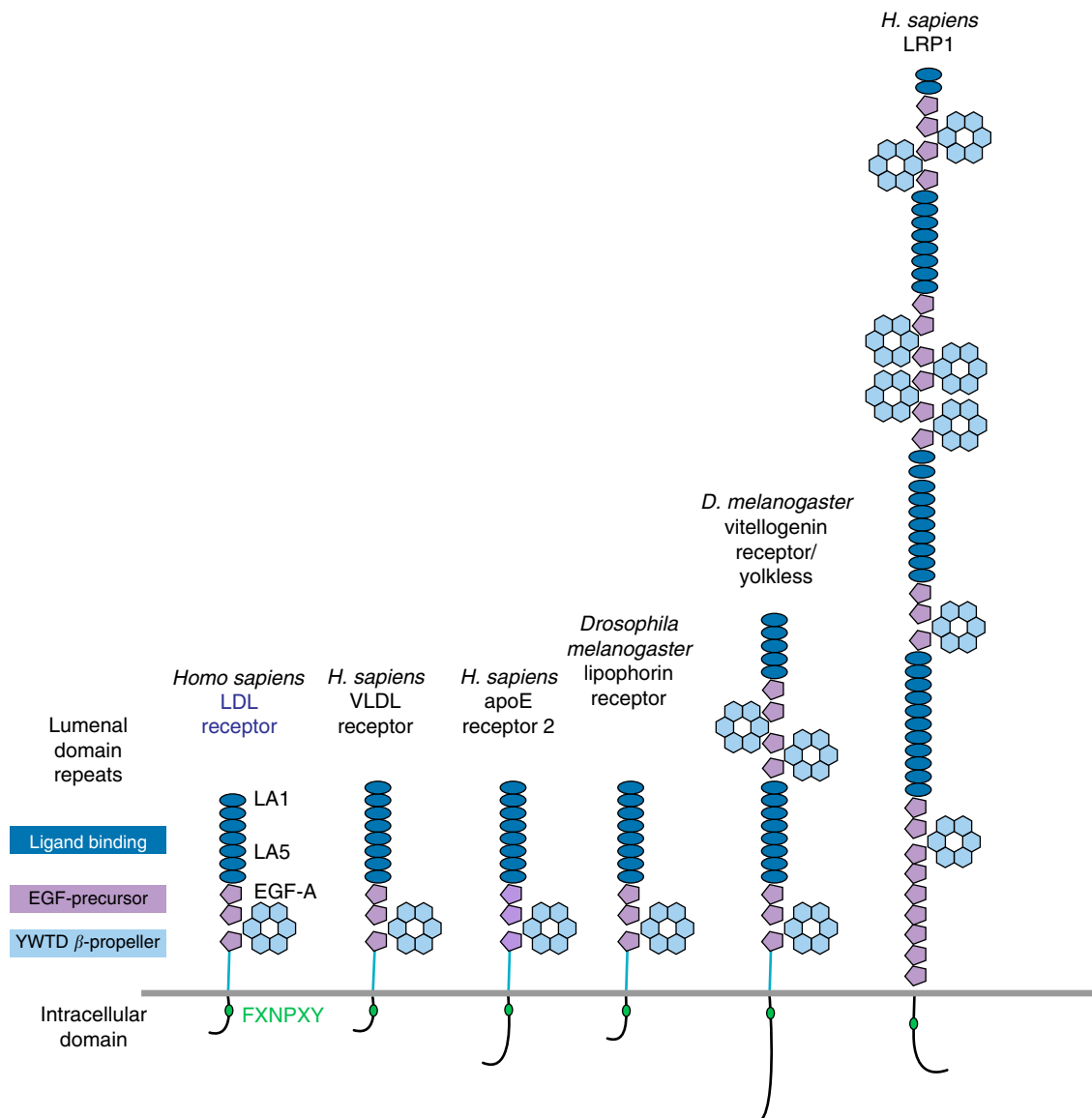


Figure 1 The LDL receptor superfamily. Schematic depiction of the domain organization of the human (*Homo sapiens*) mature LDL receptor and structurally related family members from both chordates and invertebrates (*Drosophila melanogaster*). The various serially arranged domains in the protein models are not drawn to scale, and the N-terminal signal sequence of each receptor, removed by proteolysis within the endoplasmic reticulum, is not shown.

domain(s) (Culi *et al.*, 2004), while receptor-associated protein (RAP) is responsible for efficient export of receptors from the endoplasmic reticulum (Bu, 2001) and appears to prevent inappropriate ligand engagement during passage along the biosynthetic secretory pathway by acting as a pseudoligand. This is important because LDL and VLDL receptors, as well as LRP1, are constantly being produced in hepatocytes, where VLDL is also being synthesized simultaneously within the lumen of the endoplasmic reticulum.

LDL Receptor Mutants

As indicated above, inherited defects in the LDL receptor gene lead to familial hypercholesterolemia, an autosomal-dominant disease characterized by high circulating LDL-cholesterol levels, premature atherosclerosis and increased risk of coronary artery disease. The huge assortment of *LDLR* mutations makes the prevalence of familial hypercholesterolemia allele heterozygosity about 1:200 worldwide (Nordestgaard *et al.*, 2013). Disease-associated mutations occur scattered over all 18 exons of the *LDLR* gene on chromosome 19, with the highest mapped density within exon 4 that encodes ligand-binding repeats LA3-6 (Fokkema *et al.*, 2005). The cellular consequences of *LDLR* mutations fall into five phenotypic/functional groups, designated Class 1–5 (Hobbs *et al.*, 1990). These range from deletions and missense, nonsense or frameshift mutations that

produce aberrant proteins that never fold satisfactorily or exit the endoplasmic reticulum (Class 1), to receptor forms that traverse the biosynthetic pathway and are successfully inserted into the plasma membrane, but fail to bind the exogenous LDL ligand (Class 3). Thus the various inherited pathologic mutations define numerous sequential stages during the biosynthesis and cellular functioning of the LDL receptor (Figure 2), but clinically all lead to an operationally defective phenotype and ensuing hypercholesterolemia.

Endocytic-sorting signals

Some C-terminally truncated forms of the defective LDL receptor do reach the cell surface and properly bind LDL (Class 4), but stagnate on the plasma membrane due to a lack of the intracellular segment (Hobbs *et al.*, 1990; Soutar and Naoumova, 2007). Dissimilar to the wild-type LDL receptor (Anderson *et al.*, 1982), mutant allele-encoded receptors without a cytosolic portion fail to cluster at clathrin-coated regions on the plasma membrane. This finding revealed definitively that a cytosolically disposed positive molecular determinant is necessary for selective concentration and import of LDL receptors into the cell interior within clathrin-coated vesicles. Among the numerous characterized clinical *LDLR* mutations within the cytosolic domain, arguably the most illuminating was from patient J.D. (Davis *et al.*, 1986). This unfortunate compound-heterozygous individual inherited a

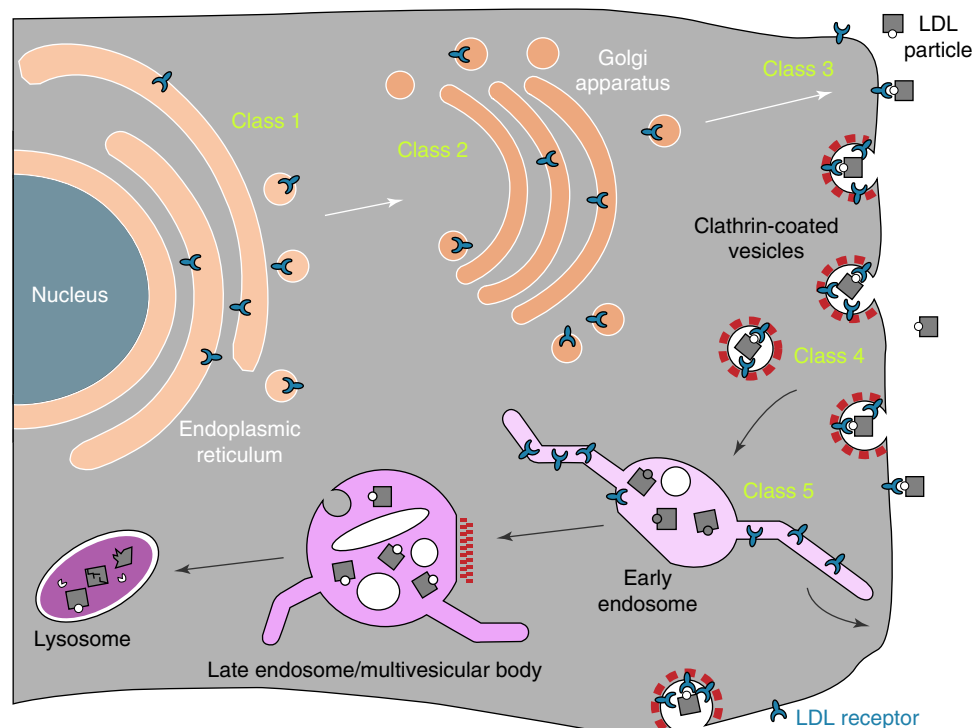


Figure 2 LDL receptor biosynthesis and trafficking itineraries as revealed by familial hypercholesterolemia mutations. A stylized cartoon of a mammalian cell, highlighting the major intracellular organelles involved in LDL receptor biosynthesis and endocytic trafficking. Organelles along the biosynthetic secretory pathway are bounded in white, while endocytic compartments have a black outline. The primary site affected by the different inherited familial hypercholesterolemia mutations in the *LDLR* gene are indicated. Class 1: null alleles with defective synthesis that never progress past the endoplasmic reticulum; Class 2: transport defective alleles that do not transit past the Golgi complex; Class 3: binding-defective alleles; Class 4: internalization-defective alleles that stagnate on the cell surface, and Class 5: recycling-deficient alleles (Hobbs *et al.*, 1990; Goldstein and Brown, 2009).

single point mutation from his father that resulted in a Y807C substitution in the mature LDL receptor. The maternally inherited *LDLR* allele was a Class 1 null (Brown and Goldstein, 1979). Just like receptors missing the cytosolic portion entirely, the tyrosine-to-cysteine mutation strands the liganded LDL receptor at the cell surface. Thus, a native aromatic tyrosine side chain at position 807 is a critical chemical requirement for LDL receptors to be able to cluster into specialized dimpled zones on the cell surface as a prelude to internalization (Davis *et al.*, 1986). Initially, the necessary signal seemed to be a tetrapeptide ⁸⁰⁴NPVY (Chen *et al.*, 1990), but detailed analysis deduced a more general consensus [YF]XNPX[YF], where '[]' indicate residues permissible at that position and 'X' denotes any amino acid.

The discovery of tyrosine-based internalization information within the primary sequence of an important transmembrane protein led to swift evaluation of numerous other well-studied transmembrane-spanning receptors. Gratifyingly, a vital tyrosine located within the cytosolic domain and positioned close to the membrane bilayer interface was found to govern the rapid uptake of several surface receptors. Perhaps the most widely recognized is the ²⁰YTRF sequence at the amino terminus of the transferrin receptor 1 (Jing *et al.*, 1990), a type II transmembrane protein, but the ²³⁵³YSKV in cation-independent mannose 6-phosphate receptor (Canfield *et al.*, 1991) and ²⁵⁵YRGV within the smaller cation-dependent mannose 6-phosphate receptor (Johnson *et al.*, 1990) cytosolic segments also led the way in defining the so-called YXXØ sorting signal, where 'Ø' designates a bulky hydrophobic residue (Traub, 2009).

The YXXØ signal is recognized directly by the heterotetrameric AP-2 clathrin adaptor complex. This makes intuitive sense as, after clathrin, AP-2 is the most abundant protein in purified clathrin-coated vesicles derived from the plasma membrane, and AP-2 is positioned between the cargo-containing membrane layer and the outer polymerized clathrin coat (Traub, 2009). Of the four co-assembled AP-2 chains (α , β 2, μ 2 and σ 2), it is the ~50-kDa μ 2 subunit that physically engages YXXØ signals (Ohno *et al.*, 1995). The X-ray crystal structure of a YQRF signal bound to μ 2 shows that the YXXØ signal forms what is termed a β -augmentation: the polypeptide signal forms a short antiparallel β -strand that hydrogen bonds temporarily to the edge of the β 16 strand of the immunoglobulin-like β -sandwich μ 2 C-terminal domain (Owen and Evans, 1998). In this binding configuration, the two anchor residues, the tyrosine and the hydrophobic 'Ø', each occupy spatially complementary, hydrophobic-lined pockets on the surface of μ 2 and, in the tyrosine-accepting pocket, Asp176 accepts a critical hydrogen bond from the signal tyrosine hydroxyl. This accounts for the tight chemical selectivity of this side chain in the YXXØ signal.

It was clear already at the outset, however, that there were fundamental biochemical distinctions between the FDNPVY signal delineated in the LDL receptor and the more common YXXØ signal. Mutation of the tyrosine ablates sorting to bulk-flow levels (about a 5- to 10-fold decrease) in altered YXXØ signals (Jadot *et al.*, 1992), even if the switch is to phenylalanine. Again, this is because the tyrosine hydroxyl is a critical hydrogen bond donor (Owen and Evans, 1998). Yet, in the LDL receptor, substitution of Tyr807 with phenylalanine

causes no meaningful alteration in trafficking, and a more bulky tryptophan side chain switch impedes internalization only modestly (Davis *et al.*, 1986, 1987). But other non-aromatic side chains cannot substitute in directing efficient LDL receptor internalization (Davis *et al.*, 1986, 1987). A chicken LDL receptor has an ⁸⁶⁵FGNPLF signal and operative [YF]XNPX[YF]-type signals in the *Homo sapiens* transmembrane proteins P-selectin (⁸²¹FTNAAF) and amnionless (⁴⁰⁶FRNPVF and ⁴⁴¹FVNPLF) also contain naturally occurring phenylalanine residues at the position corresponding to Tyr807 in the LDL receptor. The two chemically distinct 'tyrosine-based' signals also saturate distinct sorting machinery in cells; forced overexpression of the transferrin receptor competes with other YXXØ-displaying transmembrane cargo, but not with the LDL receptor, for internalization (Warren *et al.*, 1998). Yet, despite this information revealing a dichotomy between the two types of tyrosine-based internalization signals, there were reports of a general structural analogy between the FXNPXY and YXXØ signals (Collawn *et al.*, 1990) and even that, like YXXØ, the FXNPXY polypeptide bound physically to the AP-2 μ 2 subunit (Boll *et al.*, 2002).

Cargo-Selective Clathrin Adaptors for the LDL Receptor

Definitive identification of dedicated adaptors specific for the [FY]XNPX[YF] signal came from two different lines of investigation. In the first, the advent of facile RNA interference (RNAi) methodology allowed silencing the mRNA transcript (s) encoding the subunits of the heterotetrameric AP-2 complex. Somewhat unexpected at the time, because AP-2 μ 2 gene disruption in mice is lethal, this approach revealed that clathrin-mediated uptake of transferrin and LDL could be uncoupled (Motley *et al.*, 2003). Cells transfected with small interfering RNAs targeting the AP-2 μ 2- or α -subunit mRNA transcript still take up LDL (Keyel *et al.*, 2006; Motley *et al.*, 2003; Maurer and Cooper, 2006; Eden *et al.*, 2007), concordant with LDL receptor internalization not being affected by excess YXXØ-displaying cargo (Warren *et al.*, 1998). Internalization of transferrin, however, is strongly inhibited by μ 2 RNAi, as anticipated.

Disabled Homolog 2

A probable explanation for the continued internalization of LDL in AP-2 extinguished cells comes from the identification of Disabled-2 (Dab2) as a *bona fide* clathrin-coat component (Mishra *et al.*, 2002a; Morris and Cooper, 2001; Maurer and Cooper, 2006). Like AP-2, Dab2 is a steady-state component of clathrin-coated structures at the plasma membrane (Table 1). In fact, part of the reason why Dab2 is localized to endocytic clathrin zones is that the protein interacts physically with AP-2 (Morris and Cooper, 2001). Significantly, this interaction is not competitive with the YXXØ contacting site on μ 2 but, instead utilizes independently folded C-terminal appendage domains that project off the trunks of the large α and β 2 chains of AP-2 and operate as interaction hubs. Dab2 also engages clathrin directly, and the molecular information to bind the core clathrin coat proteins is displayed as a series of

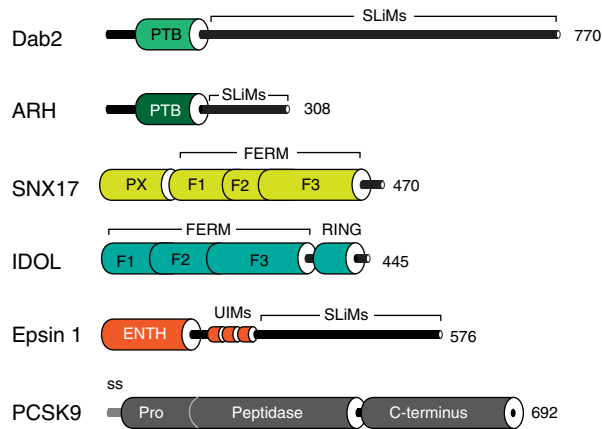


Figure 3 Key endocytic-sorting machinery involved in LDL receptor transport and turnover. Schematic illustration of the overall domain organization of LDL receptor binding/sorting proteins. The extended intrinsically disordered regions of Dab2, ARH, and epsin-1 are indicated by the narrow black lines and house the SLiMs. SS; signal sequence, which is removed by signal peptidase during co-translational translocation of PCSK9 into the endoplasmic reticulum.

serially arrayed short linear motifs (SLiMs) (Mishra *et al.*, 2002a; Morris and Cooper, 2001). These are located within a long C-terminal region of Dab2 that is predicted to be intrinsically disordered (Figure 3).

The architecture of Dab2 also reveals an N-terminal phosphotyrosine-binding (PTB) domain. The PTB domain is a folded ~150 amino acid module structurally related to the phospholipid binding pleckstrin-homology (PH) domain. But the domain name is clearly a misnomer because early on it was plainly established that the requirement for an anchor phosphotyrosine residue was not absolute; instead, a consensus binding sequence [FY]XNPXY was evident. Disabled homolog 1 (Dab1) has a phylogenetically related PTB domain but the C-terminal portion of the protein is very divergent, although both are predicted to be intrinsically unstructured. Dab1 participates in the process of cortical lamination in the developing brain and operates on a genetic pathway with the VLDL receptor and ApoER2 and the large extracellular matrix protein ligand Reelin. Both Dab1 and Dab2 utilize the PTB domain to bind physically to LDL receptor superfamily [FY]XNPXY signals. From a structural perspective, the mode of engagement of the [FY]XNPXY peptide is quite unlike YXXØ signals binding to the $\mu 2$ C domain, although the β -augmentation is common to both (Figure 4). Still, taken together, the functional characteristics of Dab2 broadly parallel those of AP-2: Both are able to bind to clathrin and to each other, and interact with phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P₂), a regulatory phospholipid that is enriched within the cytosolic leaflet of the plasma membrane. While membrane docked, both AP-2 and Dab2 have customized folded modular domains that can recognize and engage characterized sorting signals (Figure 4). RNAi-mediated suppression of Dab2 mRNA produces the opposite phenotype to AP-2 RNAi; transferrin uptake is rapid but LDL receptors, despite being normal, accumulate in a diffuse pattern on the cell surface, analogous to internalization-defective hypercholesterolemia mutants (Mishra *et al.*, 2002a; Morris and Cooper, 2001; Keyel *et al.*, 2006).

The Autosomal Recessive Hypercholesterolemia Protein

The case for PTB domain adaptors regulating LDL receptor internalization was closed with the identification of the gene responsible for another human disease, autosomal recessive hypercholesterolemia (ARH). Patients affected by this disorder present with very similar clinical features to familial hypercholesterolemia individuals, but do not have mutations in the LDL receptor or AP-2 (Eden *et al.*, 2001; Norman *et al.*, 1999). The relevant distinguishing feature in these patients is the recessive pattern of inheritance (Khachadurian and Uthman, 1973). Mapping the defective gene locus in ARH patients to a site on the short arm of chromosome 1 (Eden *et al.*, 2001; Garcia *et al.*, 2001) revealed a cytosolic 308 amino acid protein with a PTB domain at the N-terminus (Garcia *et al.*, 2001; Figure 3). Within the greater PTB domain protein superfamily, the ARH and Dab1/2 PTB domains cluster as close phylogenetic relatives, separate from the domains found in phosphotyrosine signaling components (Traub, 2009). Atomic structures of numerous PTB domains bound to peptide ligands, including those from Dab2 and ARH, are available. Generally, the models reveal the molecular basis for selectivity of tyrosine over phosphotyrosine and vice versa, and why in some PTB domains a tyrosine or phenylalanine are interchangeable within the cognate [FY]XNPXY peptide signal. Moreover, the structures provide an explanation for non-competitive binding of the Dab2/ARH PTB domains to both PtdIns(4,5)P₂ and the [FY]XNPXY sequence (Figure 4).

The C-terminal stretch of the ARH protein (also designated LDL receptor-associated protein 1 (LDLRAP1)) is again predicted to be intrinsically disordered, but contains SLiMs that are responsible for binding to clathrin or AP-2 (Mishra *et al.*, 2002b; He *et al.*, 2002), functionally analogous to Dab2. Retroviral transduction of ARH cDNA into immortalized ARH patient lymphocytes restores LDL uptake in these cells (Eden *et al.*, 2002).

Because of the increased clathrin-coat connectivity dictated by multiple SLiMs in the larger Dab2 protein, Dab2-mediated uptake of the LDL receptor is independent of AP-2 and ARH (Maurer and Cooper, 2006). Nevertheless, at a fundamental cellular level, Dab2 and ARH are functionally redundant; cultured fibroblasts from ARH patients depend on Dab2 for efficient LDL internalization (Eden *et al.*, 2007; Keyel *et al.*, 2006). Why then are ARH patients hypercholesterolemic? The answer is that in humans, about 70% of circulating LDL is cleared through the liver (Soutar and Naoumova, 2007; Sharpe *et al.*, 2014; Goldstein and Brown, 2009). In fact, the liver is the major site of LDL receptor operation. However, hepatocytes do not express high levels of Dab2 (Keyel *et al.*, 2006), making LDL internalization exquisitely sensitive to the ARH protein concentration. The relative paucity of Dab2 in hepatocytes is supported by RNA sequencing data (Uhlen *et al.*, 2010). In adult liver, the LDL receptor mRNA expression level is about 10-fold higher than in kidney, while ARH transcript abundance is roughly equivalent. Dab2 mRNA expression is approximately 10 times lower in liver than in kidney. In adrenal gland, the original tissue from which the LDL receptor was purified biochemically, LDLR mRNA expression is roughly equal to that of liver, yet Dab2 transcript abundance is about eightfold higher in adrenals. Lymphocytes

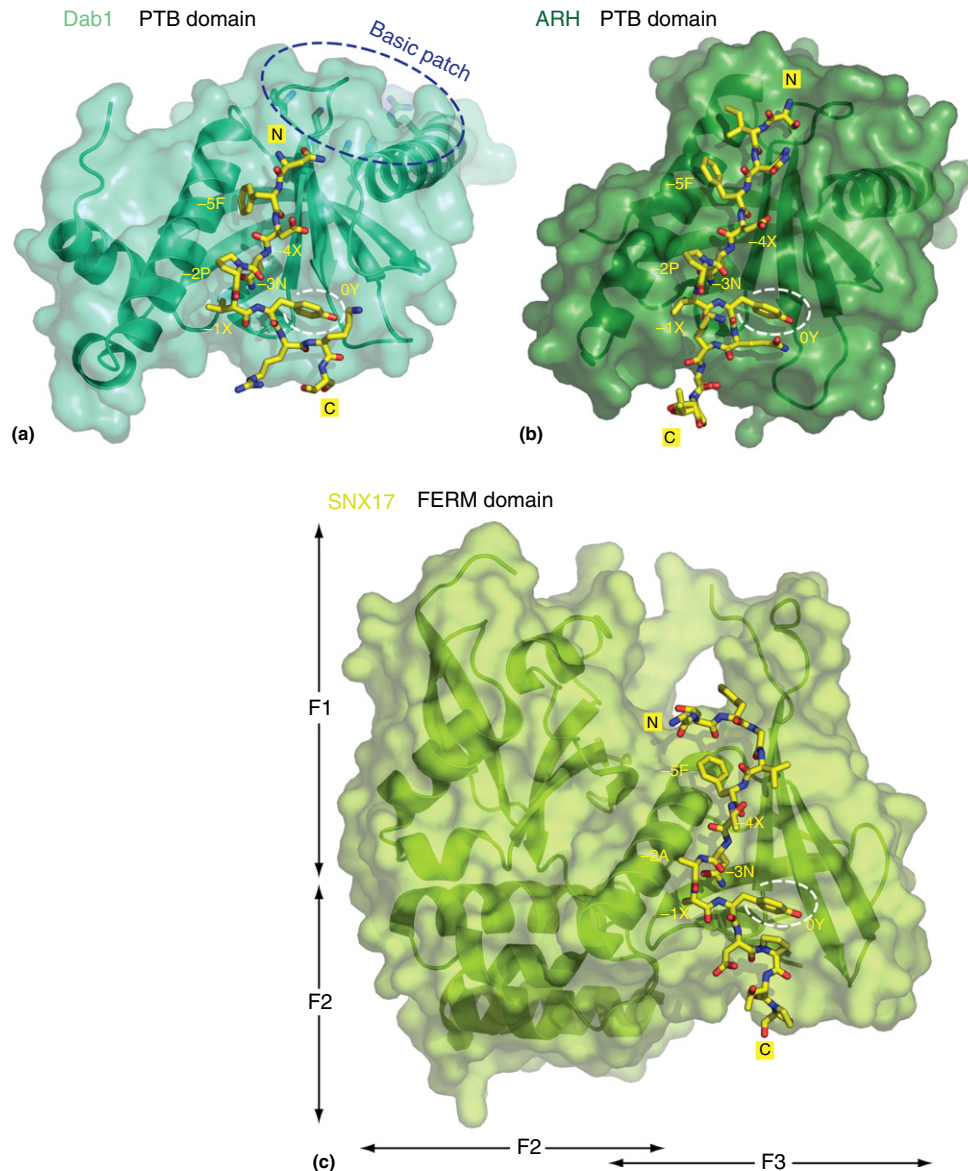


Figure 4 The structural basis for cargo recognition along the endocytic pathway. (a and b) Combination ribbon and molecular surface representations of the Dab1 (a; PDB ID: 1NTV) and ARH (b; PTB ID: 3S06) PTB domains co-crystallized with either the apoE receptor 2 (NFDPVYRKT) or the LDL receptor (SINFDPVYQKTT) sorting signal, respectively. The bound signals are shown colored yellow in stick representation, with nitrogen atoms shaded blue and oxygen red. The location of the -5 to -1 residues relative to the FXNPXY anchor tyrosine, designated residue 0 (0 Y), is indicated, as are the N and C termini of each peptide. The key features of the β -augmentation-based interaction are the extended conformation of the -5 to -2 residues hydrogen bonded via the main chain to the underlying β -strand, the tight β -turn highly favored by the -2 position proline, which redirects the main chain of the peptide perpendicular to the axis of the β -strand it contacts. This orients the anchor tyrosine (or phenylalanine) to fit into a spatially complementary acceptor pocket (white broken oval). The cluster of basic side chains in the Dab1 PTB domain (Lys 45, Arg 76, His81, Lys82, Arg124, and Lys142) are shown in stick representation and the cohesive basic surface patch they generate that is utilized to bind to PtdIns(4,5) P_2 is indicated. Although this structure shown is from Dab1, the Dab2 PTB domain structure is nearly identical (but of lower resolution; PTB ID: 1M7E) and has analogous molecular surfaces that bind to [YF]XNPX[YF] signals and PtdIns(4,5) P_2 . (c) Ribbon/surface representation of the SNX17 FERM domain illustrating the close structural relationship between the FERM F3 subdomain and the PTB domain, and the conserved mode of peptide interaction. The tyrosine-based GTYGVFTNAAYDPTP signal from P-selectin is shown in yellow stick representation and colored as in A and B. Note that the mode of engagement with the F3 fold is highly similar to the PTB domain. Although this peptide does not contain a -2 position proline residue, a different NPXY peptide similarly crystallized with the SNX17 FERM domain (PTB ID: 4TKN) verifies the orientation of the -2 proline relative to the standard PTB domain.

similarly express low levels of Dab2 and depend on ARH for LDL receptor internalization (Eden *et al.*, 2002). Why co-expression of the functionally redundant Dab2 has been

selected against in these tissues is not entirely clear at present. An obvious possibility is that ARH expression primarily allows regulatory inputs that would not be possible with the

combined presence of both Dab2 and ARH. Importantly, while Dab2 is a steady-state building block of surface clathrin-coated structures, ARH deposition at these zones is strongly affected by cargo (LDL receptor) availability (Mettlen *et al.*, 2010). In addition, the α -helical SLiM within ARH that permits AP-2 engagement contacts the appendage domain of the β 2 subunit (Mishra *et al.*, 2002b; He *et al.*, 2002). A critical residue upon the AP-2 β 2 appendage required for this interaction is Tyr888, which is positioned in the central region of the contact interface. This residue undergoes tyrosine phosphorylation, with the size and charge of the reversibly attached phosphate reshaping the interaction surface and sterically preventing ARH binding (Chakraborty *et al.*, 2014). The high incidence of the pTyr888 modification in lymphocytes may explain why LDL receptors do not congregate in clathrin-coated structures at the surface of these cells at steady state, unlike fibroblasts (Eden *et al.*, 2002) and other cell types.

The LDL receptor thus displays a structurally distinct type of internalization signal that meshes with modular PTB domain adaptors as opposed to the standard AP-2 sorting adaptor. Dab2 and ARH are typical examples of clathrin-associated sorting proteins (CLASPs), an expansive set of cargo-selective proteins that diversify the sorting possibilities of clathrin-coated structures on the cell surface beyond the cargo recognition capabilities of AP-2 (Traub, 2009). Being encoded by the primary structure of the LDL receptor, the [FY]XNPX[YF] signal directs constitutive internalization irrespective of whether the receptor is ligand bound or not. For ARH, the privileged interaction surface on the AP-2 β 2 appendage permits noncompetitive LDL receptor uptake under conditions of extensive internalization of other cargo within a single vesicle.

Endosomal Dynamics

Concomitant with cargo capture, the transition from a gently curved clathrin-coated patch to a deeply invaginated clathrin-coated bud occurs within only a minute or two at 37 °C. After pinching off from the plasma membrane in a GTP-requiring step, clathrin-coated vesicles are rapidly uncoated in an ATP-dependent fashion. The released coat protomers reenter the cytosolic pool to initiate new rounds of coat assembly. The uncoated vesicles, laden with LDL and other cargo, very rapidly fuse with one another to generate a so-called early endosome compartment, initially positioned in the cortical region of the cell. Proteolytic degradation products of internalized apoB100 become detectable at about 30 min post uptake, meaning that the LDL particle has reached the lysosome within this time interval, and is being catabolized to constituent amino acids (Brown and Goldstein, 1979; Goldstein and Brown, 2009). Lysosomal cathepsin D is the principal degradative protease (Muller *et al.*, 2012). On the same time scale, lysosomal acid lipase removes the long-chain fatty acid esterified to cholesterol in the lipoprotein particle core. Incorporation of the dietary-derived cholesterol into cellular membranes then depends on a pair of disease-related cholesterol-binding proteins, Niemann-Pick type C1 (NPC1) and NPC2, which are located in the late

endosome/lysosome compartment at steady state (Goldstein and Brown, 2009).

The biological half-life of the LDL receptor is, by comparison, in excess of 15 h (Brown and Goldstein, 1979; Goldstein and Brown, 2009). In the presence of cycloheximide to prevent new protein synthesis, rapid LDL particle uptake and breakdown persists for several hours. This large difference between the lifetimes of the receptor and its ligand dictates that each LDL receptor typically participates in many rounds of reinsertion into the plasma membrane, internalization and recycling, before undergoing degradation itself. It is estimated that the round trip is completed by an LDL receptor every 10 min-or-so (Brown *et al.*, 1997; Goldstein and Brown, 2009). A necessary physical requirement for this to occur is the detachment and separation of the LDL and LDL receptor within the early endosomal compartment.

Uncoupling of LDL from the Receptor in Endosomes

The moniker 'ligand-binding repeat' explicitly connotes the fact that these elements represent the principal part of the extracellular domain of the LDL receptor that engages lipoproteins. Over 40% of *LDLR* sequence variations map to exons 2–6, which encode the LA repeat array (Hobbs *et al.*, 1990; Soutar and Naoumova, 2007; Fokkema *et al.*, 2005). These are primarily Class 3 – binding defective – alleles. Deletion and truncation studies show that for LDL, apoB100 binds to a segment containing LA repeats 3–7. The LDL receptor can also interact with apoE (on VLDL), and this requires the LA5 repeat (Rudenko and Deisenhofer, 2003; Andersen and Moestrup, 2014). A natively folded LA5 repeat is therefore necessary for any lipoprotein to bind productively to the LDL receptor. At least one biologically important role for the set of flexible, tandemly arrayed LA repeats in the LDL receptor is to increase ligand affinity through avidity phenomena, making the critical process of ligand uncoupling within the endosome feasible (Andersen and Moestrup, 2014).

The lineup of compact LA repeats coordinate Ca^{2+} with measured equilibrium dissociation constants (K_D s) in the nanomolar to low micromolar range, depending on the individual repeat (Rudenko and Deisenhofer, 2003; Andersen and Moestrup, 2014). Three intermolecular disulfide bonds form within each ~ 40 -residue LA repeat that underpin the structural integrity of the Ca^{2+} -coordinating acidic surface, the so-called Ca^{2+} cage (Brown *et al.*, 1997). Indeed, the overall cohesion and ligand-binding capability of these repeats can be disrupted either by disulfide bond reduction or by Ca^{2+} chelation. Also, Class 3 disease-producing mutations affecting cysteine or other cage residues similarly diminish LDL binding because the protein becomes nonfunctional (Brown *et al.*, 1997; Hobbs *et al.*, 1990; Soutar and Naoumova, 2007). Thus, Ca^{2+} obviously participates in high affinity ligand binding at the neutral pH of the extracellular milieu, but whether Ca^{2+} physically dissociates from the LA repeats in a mildly acidic endosomal environment, generated through the action of the transmembrane proton pump, the vacuolar ATPase, as a dedicated mechanism for localized unfolding/destabilization to release bound LDL, is not yet entirely clear. An LDL receptor with an intact set of LA repeats but with the abutting C-terminal region, which displays homology

with the membrane-attached EGF precursor (the EGF homology module), deleted entirely still binds to and internalizes LDL (Hobbs *et al.*, 1990; Rudenko and Deisenhofer, 2003). Yet there is no separation within the early endosome compartment. This finding puts important mechanistic constraints on the idea that a binary Ca^{2+} -coordination toggle switch represents an unusual form of Ca^{2+} intracellular signaling governing LDL release (Andersen and Moestrup, 2014).

The Class 5 familial hypercholesterolemia mutations in the *LDLR* result in failure of the receptor–ligand complex to properly dissociate within endosomes (Brown and Goldstein, 1979; Goldstein and Brown, 2009; Figure 2). These inherited defects map primarily to the EGF homology module encoded by exons 7–14 of the *LDLR*, pinpointing these integrated folded domains as important regulators of apoB100 attachment. Coordinated Ca^{2+} also stabilizes the fold of the 40–50 amino acid EGF-A and EGF-B (but not EGF-C) repeats in the LDL receptor; Ca^{2+} is bound with low micromolar affinity (Rudenko and Deisenhofer, 2003; Andersen and Moestrup, 2014). Because it is firmly established that Ca^{2+} efflux accompanies proton entry during early endosome maturation, Ca^{2+} regulation could be operative at the EGF homology module as well (Andersen and Moestrup, 2014). The crystal structure of the almost complete luminal portion of the LDL receptor at acidic pH (5.3) reveals why ligand rebinding is prevented in the endosome (Rudenko *et al.*, 2002). The chief result is that the LA4 and LA5 repeats form direct intramolecular contacts with the β -propeller in the EGF homology module, folding the extended LA repeat series over the EGF homology module and, thereby, occluding the major ligand-binding region of the receptor. Correctly positioned lysine residues in the β -propeller interact electrostatically with the Ca^{2+} cage regions of LA4 and LA5, presumably mimicking, intramolecularly, the binding of apoB or RAP (Rudenko and Deisenhofer, 2003; Andersen and Moestrup, 2014). Clearly, in this structure, the Ca^{2+} ions remain coordinated by the ligand-binding repeats.

So what then drives these intramolecular contacts to occur particularly at the lower pH of the endosome? Protonation of luminal histidine side chains is implicated in pH-responsive ligand release by other endocytic receptors. There are three interfacial histidine residues in the LA4/LA5- β -propeller interaction surface that may confer pH sensitivity (Rudenko and Deisenhofer, 2003). Provocatively, two of these residues are represented among *LDLR* mutant alleles (H190Y on LA5 and H562Y on the β -propeller) (Huang *et al.*, 2010; Rudenko and Deisenhofer, 2003). In a current model to explain LDL discharge in the endosome compartment (Huang *et al.*, 2010), LDL particle binding to the receptor involves the LA repeat tract, as well as the β -propeller surface containing His562 and His586. The protonation of these histidine residues that accompanies endosomal acidification reverses contacts with LDL and favors rapid intramolecular closure of the receptor ligand-binding interface (Huang *et al.*, 2010). It seems clear that a combination of Ca^{2+} , pH and domain configuration changes combine to promote optimal release of LDL (and VLDL) from the LDL receptor (Andersen and Moestrup, 2014). Future studies will no doubt clarify the precise biological role of each of these processes.

Selection of the LDL Receptor for Recycling to the Cell Surface

Once the LDL particle is no longer able to bind tightly to the LDL receptor within the acidified endosome lumen, geometric considerations play a significant role in directing the differential movement of the two separated binding partners. Narrow tubules begin to form on early endosomes, which emanate from the central spherical vesicular portion (Bonifacino and Rojas, 2006). The high surface-to-volume ratio of the emerging tubules biases them toward an enriched load of transmembrane receptors while restricting bulk ingress of bulky luminal macromolecules, especially large assemblages like LDL (Figure 2). Failure of the stoichiometric LDL–LDL receptor complex to separate causes the entire macromolecular complex to travel to the lysosome and get degraded. The maturing tubulovesicular endosome is an important sorting station and several compositionally distinct tubules are produced to convey cargo to different intracellular acceptor compartments as well (Bonifacino and Rojas, 2006). The mechanistic implication of this intrinsic sorting complexity within a morphologically plastic endosomal compartment is that the LDL receptor must be sequestered into the correct inchoate tubules to facilitate rapid and efficient recycling back to the plasma membrane.

Current evidence indicates that a member of the sorting nexin family, sorting nexin 17 (SNX17), is important to actively segregate the LDL receptor into the proper recycling tubules during the process of early endosome maturation (Table 1). The sorting nexins are soluble proteins generally deposited on endosomes at steady state, but what sets SNX17 apart from many other members of the sorting nexin family is the presence of a band 4.1/ezrin/radixin/moesin (FERM) domain (Teasdale and Collins, 2012; Figure 3). The FERM domain is a roughly 300 amino acid, trefoil-like shaped globular domain. The founding members of the FERM domain family are peripheral membrane proteins that utilize the domain as a lipid-binding linker between the extracellular matrix/plasma membrane and the assembled cortical cytoskeleton (Chishti *et al.*, 1998). Specifically, integrins can engage the FERM domain to connect extracellular cues to actin microfilaments. The part of the β -integrin chain that engages the FERM domain displays an FXNPXY signal, a sequence tract originally noted to be related to the LDL receptor (Chen *et al.*, 1990).

SNX17 is implicated in endosomal sorting of numerous LDL receptor superfamily members, including the LDL, VLDL and apoER2 receptors and LRP1 (Stockinger *et al.*, 2002; Burden *et al.*, 2004), as well as the structurally unrelated P-selectin (Williams *et al.*, 2004), amyloid precursor protein and certain β -integrins (Lee *et al.*, 2008; Steinberg *et al.*, 2012). These proteins are all unified by sharing [FY]XNPX[YF]-type sorting signals located within the cytosolic domain (Chen *et al.*, 1990). For the LDL receptor, experimental manipulation of the cytosolic SNX17 concentration or structure directly affects rates of receptor recycling and intracellular positioning. Essentially, loss of SNX17 shortens the biologic half-life and therefore the number of duty cycles each LDL receptor can perform; normally SNX17 prevents the LDL (and other recognized) receptors from traveling to the lysosome.

FERM domains are typically composed of three interlinked folded lobes, called F1, F2, and F3 (Figure 3), with the F3 lobe

structurally related to the aforementioned PTB domain. Atomic-resolution data on the SNX17 FERM domain explains how the [YF]XNPX[YF] signal is recognized, which is remarkably similar to how Dab1/2 and ARH bind to the same sequence (Figure 4). The closest paralogue, SNX31, and possibly the more divergent SNX27, may act in a functionally redundant fashion with SNX17; SNX27 associates directly with the P-selectin FTNAAAY signal (Ghai *et al.*, 2013; Traub, 2009).

An overlapping tract of the LDL receptor cytosolic domain therefore directs both internalization and recycling, but by being decoded by different molecular machinery. Yet, if PTB and FERM domains both engage the [YF]XNPX[YF]-type signal, why do they not operate promiscuously? The answer is that Dab2/ARH and SNX17 are directed to separate membrane surfaces by physically engaging differentially phosphorylated variants of phosphatidylinositol; PtdIns(4,5) P_2 at the plasma membrane and phosphatidylinositol 3-phosphate (PtdIns3P) in the endosomal compartment. Recognition of PtdIns3P by SNX17 is specified by the N-terminal ~110-residue phox-homology (PX) domain (Teasdale and Collins, 2012; Figure 3). The typical affinity of the PTB or FERM domain for the [FY]XNPX[YF] signal, or of the PX domain for PtdIns3P, is in the low micromolar range. This range of K_D s display interaction half times that persist for seconds to minutes (Hammond *et al.*, 2009). If sorting depended solely on these bimolecular interactions, it would be highly susceptible to promiscuous binding mistakes (error prone). One mechanism to diminish this deleterious possibility is to temporally integrate several low affinity contacts to guard against biological noise, in the form of nonspecific engagements by related (or unrelated) proteins. The proper associations are thus strengthened and prolonged at the correct compartmental membrane surface by the phenomenon of avidity; synchronous contacts with other proteins and lipids that facilitate appropriate macromolecular assembly formation and sorting. Because avidity is related to valency, the capability of the LDL receptor to form functional dimers will also assist in the fidelity of sorting.

Other SNX family proteins that direct sorting operations at the early endosome cooperate with scaffolding and cytoskeletal proteins, to which they bind physically (Teasdale and Collins, 2012). The precise mechanism by which SNX17 partitions the LDL for efficient recycling to the cell surface is not fully revealed. It is also yet to be resolved whether a single lengthy endosomal tubule contains multiple SNX17-, SNX1/2+retromer- and EHD-positive zones that can fragment to produce differentially trafficked cargo carriers destined for distinct downstream acceptor compartments.

Dynamic Modulation of LDL Receptor Surface Levels

Cholesterol Overload Responses

The impressive strategies that have evolved in animal cells to assure ample cholesterol supply are counterpointed by recently uncovered, yet equally striking, pathways to limit cholesterol over-accumulation at the cellular level. Cholesterol in excess of cellular demand undergoes oxidative modification. Accumulating oxidized forms of the sterol relay biological homeostatic information by binding to the nuclear receptor

family factors liver X receptor α (LXR α) and LXR β (Zelcer *et al.*, 2009; Hong and Tontonoz, 2014). Oxysterol-stimulated derepression of the LXR transcription factors promptly turns on a battery of genes that manage cholesterol efflux from the cell, the so-called reverse cholesterol transport (Sharpe *et al.*, 2014; Sorrentino and Zelcer, 2012; Zelcer *et al.*, 2009). Overall, the process is aimed at relocating excess cholesterol from within peripheral tissues back to the liver within HDL particles, and then excreting the returned cholesterol as bile salts. Under these conditions, the cellular abundance of the LDL receptor decreases. One intriguing target of LXRs is a 445 amino acid protein (Zelcer *et al.*, 2009) originally dubbed MYLIP (myosin regulatory light chain interacting protein) (Table 1). The domain organization of this protein illustrates that it is essentially composed of two modular parts: an N-terminal FERM domain and a C-terminal really interesting new gene (RING) domain (Figure 3), which immediately suggests a possible mechanism for the observed down-regulation of the LDL receptor. Together with the RING domain, which imparts E3 ubiquitin ligase capability, this particular LXR target gene could first identify and then post-translationally modify the LDL receptor with ubiquitin chains (constructed by covalent linkage of the 76 amino acid protein ubiquitin) to reroute the protein for lysosome delivery and catabolism. This is indeed the case, prompting the redesignation of MYLIP as an inducible degrader of the LDL receptor (IDOL) (Zelcer *et al.*, 2009). Neither the FERM nor RING domain alone diminishes LDL receptor abundance in cells; the protein must be intact to function in this way (Sorrentino *et al.*, 2011; Scotti *et al.*, 2013). Further, an LDL receptor lacking the intracellular domain is impervious to IDOL-mediated action; the cytosolic domain is necessary both for recognition by the IDOL FERM/RING domains and to provide an acceptor lysine (or cysteine) residue for ubiquitin addition (Sorrentino and Zelcer, 2012).

Despite IDOL and SNX17 both containing FERM domains, there is little primary sequence identity between the two proteins. Unfortunately there is no direct atomic-resolution structural information currently available for the IDOL FERM domain, but sequence comparisons reveal that it has an insertion within the F3 lobe that is unusual (Calkin *et al.*, 2011). Notably, LRP1 is not an IDOL substrate, so the FERM domain in this E3 ligase cannot recognize the [FY]XNPX[YF] signal in an identical manner to SNX17. Also, the principal acceptor lysine residue in the LDL receptor (Lys809), to which the initial ubiquitin is attached via an isopeptide linkage, is separated from the anchor Tyr807 (the position 0 Y altered in the J.D. point mutation) by only a single amino acid (Zelcer *et al.*, 2009). Residues N-terminal to the [FY]XNPX[YF] sorting signal appear crucial for IDOL-mediated ubiquitin addition to the LDL receptor at position Lys809; a consensus sequence of WXXKNXXS[IM]XF[DE] (where the trailing 'F' is the -5 position F of the [FY]XNPX[YF] sequence) is needed to bind to the IDOL FERM domain (Sorrentino *et al.*, 2013a; Calkin *et al.*, 2011). This reliance on a preceding amino acid tract to provide specificity-determining side chains for FERM engagement is analogous to how the TALIN FERM domain binds to β -integrins (Wegener *et al.*, 2007), and likely explains why IDOL evolved to contain a FERM and not a PTB domain to engage LDL receptor family targets (Calkin *et al.*, 2011). The IDOL

FERM has undergone functional selection to, first, facilitate recognition of only a subset of the LDL receptor superfamily proteins and, second, permit docking of the sorting signal in such a way that the ϵ -amino group of the Lys809 side chain is accessible for ubiquitin ligation (Sorrentino *et al.*, 2013a; Calkin *et al.*, 2011).

Strikingly, population gene sequencing efforts show that a possible gain-of-function (N342S) and a loss-of-function (R266X) IDOL variant are associated, respectively, with high and low circulating LDL-cholesterol levels in humans (Sorrentino and Zelcer, 2012). This ubiquitin-dependent mechanism to turn over the LDL receptor therefore impacts systemic LDL clearance in animals. The ubiquitination of the LDL receptor by IDOL dimers, working in conjunction with the E2 ubiquitin-conjugating enzymes UBE2D (Zhang *et al.*, 2011) or UBE2N (Sorrentino *et al.*, 2011), generates Lys63-linked polyubiquitin chains at the cell surface (Sorrentino *et al.*, 2011; Scotti *et al.*, 2013). IDOL also ubiquitinates the phylogenetically closest LDL receptor relatives, the VLDL and apoER2 receptors (Sorrentino and Zelcer, 2012). Similarly positioned lysine acceptor residues are found flanking the FXNPXY signals in these receptors. The isopeptide-linked bulky ubiquitin adduct most likely obstructs normal recognition of the FXNPXY signal by the PTB domain and instead directs the receptor for lysosomal degradation. Preliminary results suggest that the IDOL-directed route of intracellular movement to the lysosome is distinct from the normal trafficking itinerary of LDL-LDL receptor complexes (Figure 2), and does not require clathrin-mediated endocytosis (Sharpe *et al.*, 2014; Sorrentino *et al.*, 2013b). Rather, decoding of the ubiquitin signal is attributed to epsin, a 576 amino acid CLASP architecturally reminiscent of, and functionally analogous to, Dab2 and ARH (Figure 3), and (curiously) also a steady-state constituent of cell surface clathrin-coated structures (Traub, 2009). Instead of a PTB domain, epsin contains an epsin N-terminal homology (ENTH) domain, which binds to PtdIns(4,5) P_2 , and SLiMS within the intrinsically unstructured C-terminal region that interact with AP-2 and clathrin directly. Trailing after the ENTH domain are a set α -helical of ubiquitin-interacting motifs (Figure 3) that recognize Lys63-linked polyubiquitin chains (Traub, 2009; Piper *et al.*, 2014). New careful analysis shows that almost all (>95%) of internalization events from the plasma membrane are clathrin dependent (Bitsikas *et al.*, 2014), making the assertion that IDOL drives a different process (Scotti *et al.*, 2013; Sorrentino *et al.*, 2013b) somewhat provisional and requiring systematic reevaluation. Why a separate sorting route is necessary is not immediately obvious, as many signaling and nutrient receptors are effectively targeted for lysosomal degradation from the cell surface by the standard endosomal pathway utilizing Lys63-linked polyubiquitin as the sorting signal (Piper *et al.*, 2014).

The discovery of IDOL deftly highlights that appended ubiquitin can potentially override the normal trafficking behavior of the LDL receptor. What this means operationally is that when the LDL receptor arrives in the endosomal compartment, the default recycling itinerary governed by the FDNPVY signal is thwarted by the nearby, posttranslationally added ubiquitin chain. How does this happen? It turns out that Lys63-linked conjugation is actually a general way to

direct selected transmembrane cargo to the lysosomal compartment (Piper *et al.*, 2014). It is vital to recognize that this ubiquitin-dependent degradation of cell surface transmembrane proteins/cargo within the endosomal system is a fundamentally different molecular mechanism from the ubiquitin-mediated degradation of soluble proteins by the cytosolic proteasome (Sharpe *et al.*, 2014). While they both use a common pool of ubiquitin, part of the way these two ubiquitin-governed catabolic processes are segregated is by using different E2 conjugating enzymes and E3 ubiquitin ligases that catalyze different chain linkages (principally Lys48 polyubiquitin for proteasomal degradation).

To decode the ubiquitin-encoded lysosomal sorting information during intracellular sorting, a collection of sequentially assembled macromolecular complexes, designated the endosomal sorting complex required for transport-0 to -III (ESCRT-0–III), form a sorting hub upon the early endosome surface by recognizing PtdIns3P and (Lys63-containing) ubiquitin simultaneously (Piper *et al.*, 2014). The same guiding mechanistic principles apply: these are again modular peripheral membrane proteins composed of a run of folded units that can interact with specific, compartmentally restricted lipids, ubiquitin conformations, and each other. Recalling the geometry-influenced sorting that occurs in the early endosome (Figure 2), assembly of the ESCRT machinery provides a mechanism to enrich/trap transmembrane proteins within the bulbous vesicular portion of the developing early endosome as the thin sorting tubules are drawn out. Fundamentally, the ESCRTs represent patch of membrane that displays a distinct composition, dynamic and density compared with the adjacent, non-assembled membrane, which permits retention based on ubiquitin-chain instructions. A detailed description of the ESCRT apparatus and multivesicular body biogenesis is beyond the scope of this article, but more information can be found in other articles in this Encyclopedia.

Ubiquitin-Independent Downregulation of the LDL Receptor

The genetic basis for yet another (rare) hypercholesterolemic phenotype discloses a molecularly distinct but parallel pathway for posttranscriptional LDL receptor breakdown. It has emerged recently that a very minor fraction of dominant hypercholesterolemias is due, not to *LDLR* mutations, but to gain-of-function mutations in the gene encoding a putative protease designated proprotein convertase subtilisin/kexin type 9 (PCSK9) (Soutar and Naoumova, 2007; Horton *et al.*, 2009; Sharpe *et al.*, 2014; Leren, 2014). PCSK9 is a 692 amino acid secreted serine protease that undergoes autocatalytic cleavage of the nascent zymogen form during biosynthesis and transport along the secretory pathway (Table 1 and Figure 3). As for VLDL, the liver seems to be the primary site of production of the PCSK9 coursing through the plasma (Horton *et al.*, 2009). Population studies, analogously to *IDOL*, illustrate that gain-of-function (D374Y/H, F216L, and S127R) or loss-of-function (missense and nonsense mutations, including rare homozygous null individuals) PCSK9 alleles occur in human subjects, which co-segregate well with plasma LDL-cholesterol levels (Soutar and Naoumova, 2007; Horton *et al.*, 2009; Sharpe *et al.*, 2014; Leren, 2014).

The way that PCSK9 affects plasma LDL levels is largely by driving hepatic clearance of normal LDL receptors. Yet this mode of regulation is obviously quite distinct from IDOL-mediated ubiquitination, because PCSK9 does not operate in the cytosolic compartment – it is released from cells in exocytic vesicles and binds to the extracellular region of cell-surface LDL receptors. Newly synthesized PCSK9 exhibits a very short (minutes) circulatory half-life, testifying to the fact that once bound to the LDL receptor, the 1:1 stoichiometric complex swiftly winds up in the lysosome compartment for degradation (Soutar and Naoumova, 2007; Horton *et al.*, 2009; Sharpe *et al.*, 2014; Leren, 2014). And despite the catalytic potential of PCSK9, in the plasma the protein acts as an enzymatically inactive regulator of LDL receptor abundance that reprograms the trafficking itinerary of the bound receptor. This is because the cleaved but folded prodomain portion, which remains non-covalently attached to the mature protein, is positioned to conceal the catalytic triad (Horton *et al.*, 2009; Leren, 2014). Some (but not all) of the pathophysiological gain-of-function PCSK9 alleles encode a protein that binds to the LDL receptor with considerably higher affinity. The beneficial loss-of-function alleles cause hypocholesterolemic profiles because apparently they extend the normal half-life of the LDL receptor in the liver, improving cholesterol elimination from the plasma.

The FDNPVY signal and the core coat machinery that oversees the routine internalization of the LDL receptor (Dab2, ARH, clathrin) are generally still required for LDL receptor internalization after PCSK9 engagement at the cell surface (Horton *et al.*, 2009; Sorrentino and Zelcer, 2012; Soutar and Naoumova, 2007). Importantly, PCSK9 does not trigger ubiquitination of the LDL receptor. A lysineless mutated cytosolic domain form of the LDL receptor is still down-regulated by bound PCSK9, and the proteasome inhibitor MG139 has negligible impact on the kinetics of degradation of the LDL receptor in the presence of PCSK9 (Wang *et al.*, 2012). In fact, once in the endosomal compartment, redirection of the LDL receptor to the degradative lysosome by bound PCSK9 unexpectedly appears to occur quite independently of the intracellular domain of the receptor (Leren, 2014). Moreover, delivery of the LDL receptor to the lysosome is not contingent upon the canonical ESCRT-0/-I machinery (Wang *et al.*, 2012) that is necessary for proper sorting of the IDOL ubiquitinated receptor (Scotti *et al.*, 2013; Sorrentino *et al.*, 2013b).

How then does an extracellularly bound PCSK9 alter the fate of the LDL receptor in the sorting endosome? A hint for understanding this conundrum comes from the reduced cellular abundance of the Class 5 *LDLR* mutants – these defective proteins traffic to the lysosome locked to their LDL particle ligand (Hobbs *et al.*, 1990), as do normal mature forms of the LDL receptor complexed with polyclonal anti-receptor antibodies (Anderson *et al.*, 1982). Circulating PCSK9 interacts directly with the EGF-precursor-homology region on the luminal side of the LDL receptor. Alone, the inactive catalytic domain of PCSK9 attaches specifically to the EGF-A module, and this interaction is sufficient for internalization of PCSK9 together with the LDL receptor. Yet, in this setting, the half-life of the LDL receptor is not reduced because the molecular interaction between the PCSK9 catalytic domain and EGF-A is

sensitive to pH-induced dissociation within the endosome lumen (Leren, 2014), just like the interaction with apoB100. The structural basis for the binding of PCSK9 to EGF-A reveals a similar Ca^{2+} dependence (Andersen and Moestrup, 2014), accounting for the acidification-associated release, as discussed for the LA repeats and apoB100 above.

Paradoxically, the affinity of mature, full-length PCSK9 for the extracellular portion of the LDL receptor increases (> 100 fold) in a mildly acidic environment. An appealing view is that it is the less highly conserved C-terminal domain of PCSK9, which binds to the LA repeats of the LDL receptor ectodomain (Leren, 2014), that functionally diverts the bound receptor from recycling toward degradation. The LA repeats are needed for PCSK9 to orchestrate the redirection of the LDL receptor, implying that they play a role in counteracting the acid-triggered detachment of the PCSK9 catalytic domain from the EGF-A repeat. Because the C-terminal domain of PCSK9 is the least phylogenetically conserved portion of the protein, general electrostatic attraction that is pH enhanced, as opposed to tailored binding of co-evolved interaction surfaces, seems to be involved. The C-terminus of PCSK9 is also rich in surface exposed histidine residues that can add pH-responsive positive charge for effective salt bridging with of the negatively charged LA repeat Ca^{2+} cages. Further evidence for an important role of the C-terminal domain of PCSK9 is that neutralizing antibodies directed against this region counteract the normal degradative activity of the convertase (Leren, 2014); this represents a promising new avenue for the treatment of hypercholesterolemia. Altogether, it seems highly likely that the interaction of PCSK9 with the LDL receptor in an acidic environment prevents the receptor ectodomain assuming the closed compact conformation that promotes recycling. While residing within the vesicular region of the sorting endosome, proteolytic cleavage of extracellular bulk of the LDL receptor, liberating the ectodomain into the lumen, may further favor passage to the late endocytic compartment at the expense of recycling (Leren, 2014).

One curious fact is that PCSK9 mRNA is upregulated, in parallel with the LDL receptor, when cellular cholesterol levels fall. At first glance, this seems counterintuitive because, once secreted from the liver, PCSK9 will drive the rerouting of the receptor, increasing LDL receptor degradation in the lysosome compartment, decreasing the steady-state level and restricting bulk LDL uptake capability. This genetically programmed transcriptional response actually complicates the effectiveness of cholesterol-lowering statin therapy. It may be, however, that the concurrent transcription and translation of the LDL receptor and PCSK9 driven by the sterol responsive element binding protein (Goldstein and Brown, 2009) reflects the fact that PCSK9 could be an important endoplasmic reticulum chaperone for the LDL receptor (Strom *et al.*, 2014), in addition to the aforementioned Boca and RAP proteins. Another possibility is that localized secretion of PCSK9 from hepatocytes biases just exocytosed nascent VLDL particles to enter the bloodstream instead of being immediately trapped and endocytosed by LDL receptors on the sinusoidal (basolateral) plasma membrane (Horton *et al.*, 2009). But if the latter mechanism is operative, it is not immediately apparent why PCSK9 evolved to guarantee the swift and irreversible destruction of the LDL receptors it engages.

Concluding Remarks

Since the seminal investigations in the Goldstein and Brown laboratory mapped out receptor-mediated endocytosis as a centerpiece of homeostatic sterol regulation in humans, the LDL–LDL receptor pathway has provided unrivaled insight into endocytosis and protein trafficking mechanisms. One is struck by the impressive mechanistic mutability that has evolved to precisely regulate the intracellular trafficking of this clearly clinically important nutrient receptor. The high degree of complexity of the regulatory interaction web is illustrated schematically in Figure 5. The network is characterized by multiple routes of feedback inhibition when cholesterol levels are elevated, and by positive feedback when cellular levels of the sterol drop. Opposing and parallel layers of transcriptional and posttranscriptional regulation of LDL receptor abundance in cells allow heightened homeostatic responsiveness, but with

different kinetic signatures. As has been the case over the past several decades, it is highly probable that the type of sterol-dependent regulatory mechanisms delineated for the LDL receptor pathway will be broadly applicable to other cell membrane proteins. To be anticipated is clearer appreciation of new and unexpected levels of further regulation. For example, there are over 150 independent reports of the vital Tyr807 residue being posttranslationally modified to phosphotyrosine in high-throughput mass spectrometry databases (Hornbeck *et al.*, 2012). This modification to the [FY]XNPX [YF] signal, like vicinal ubiquitin conjugation by IDOL, will alter the recognition and hence trafficking information hardwired into the primary sequence of the LDL receptor. The identity of the tyrosine kinase(s) that phosphorylates the LDL receptor at Tyr807 is unknown, but could be responsible for uncoupling LDL traffic flow during acute signaling responses from the plasma membrane.

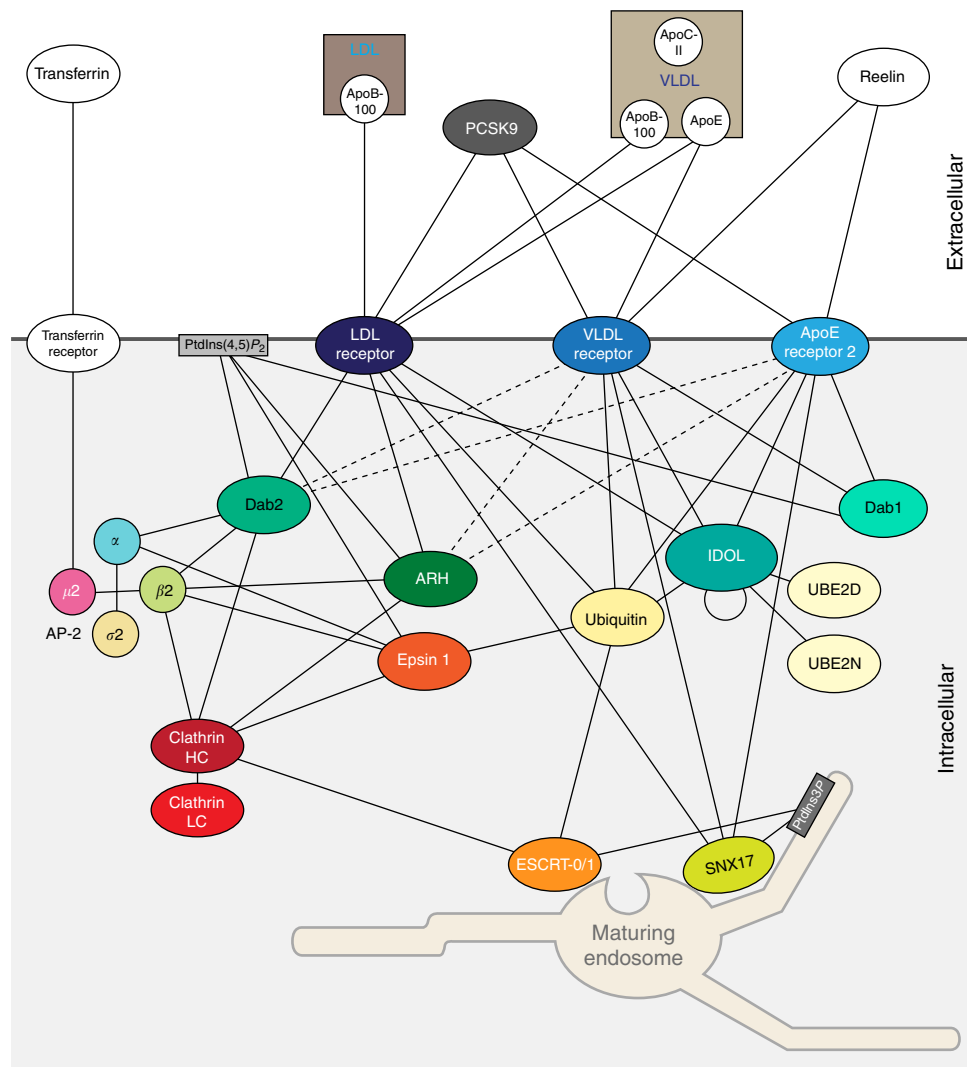


Figure 5 The LDL receptor protein interaction hub. An integrated hub-and-spoke depiction of the protein–protein and protein–lipid interactions that regulate LDL receptor trafficking operations and receptor half-life. Established molecular interactions are represented with a solid connecting line while presumed or likely interactions are shown with a broken line. HC: heavy chain; LC: light chain.

Acknowledgment

This work was supported by the National Institutes of Health grant R01 GM106963.

References

- Andersen, C.B., Moestrup, S.K., 2014. How calcium makes endocytic receptors attractive. *Trends in Biochemical Sciences* 39, 82–90.
- Anderson, R.G., Brown, M.S., Beisiegel, U., Goldstein, J.L., 1982. Surface distribution and recycling of the low density lipoprotein receptor as visualized with antireceptor antibodies. *Journal of Cell Biology* 93, 523–531.
- Bitsikas, V., Correa, I.R., Nichols, B.J., 2014. Clathrin-independent pathways do not contribute significantly to endocytic flux. *eLife* 3, e03970.
- Boll, W., Rapoport, I., Brunner, C., *et al.*, 2002. The $\mu 2$ subunit of the clathrin adaptor AP-2 binds to FDNVY and Ypp Φ sorting signals at distinct sites. *Traffic* 3, 590–600.
- Bonifacio, J.S., Rojas, R., 2006. Retrograde transport from endosomes to the trans-Golgi network. *Nature Reviews Molecular Cell Biology* 7, 568–579.
- Brown, M.S., Goldstein, J.L., 1979. Receptor-mediated endocytosis: Insights from the lipoprotein receptor system. *Proceedings of the National Academy of Sciences of the United States of America* 76, 3330–3337.
- Brown, M.S., Herz, J., Goldstein, J.L., 1997. LDL-receptor structure. Calcium cages, acid baths and recycling receptors. *Nature* 388, 629–630.
- Bu, G., 2001. The roles of receptor-associated protein (RAP) as a molecular chaperone for members of the LDL receptor family. *International Review of Cytology* 209, 79–116.
- Burden, J.J., Sun, X.M., Garcia, A.B., Soutar, A.K., 2004. Sorting motifs in the intracellular domain of the low density lipoprotein receptor interact with a novel domain of sorting nexin-17. *Journal of Biological Chemistry* 279, 16237–16245.
- Calkin, A.C., Goult, B.T., Zhang, L., *et al.*, 2011. FERM-dependent E3 ligase recognition is a conserved mechanism for targeted degradation of lipoprotein receptors. *Proceedings of the National Academy of Sciences of the United States of America* 108, 20107–20112.
- Canfield, W.M., Johnson, K.F., Ye, R.D., Gregory, W., Kornfeld, S., 1991. Localization of the signal for rapid internalization of the bovine cation-independent mannose 6-phosphate/insulin-like growth factor-II receptor to amino acids 24–29 of the cytoplasmic tail. *Journal of Biological Chemistry* 266, 5682–5688.
- Chakraborty, S., Umasankar, P.K., Preston, G.M., *et al.*, 2014. A phosphotyrosine switch for cargo sequestration at clathrin-coated buds. *Journal of Biological Chemistry* 289, 17497–17514.
- Chen, W.J., Goldstein, J.L., Brown, M.S., 1990. NPXY, a sequence often found in cytoplasmic tails, is required for coated pit-mediated internalization of the low density lipoprotein receptor. *Journal of Biological Chemistry* 265, 3116–3123.
- Chishti, A.H., Kim, A.C., Marfatia, S.M., *et al.*, 1998. The FERM domain: A unique module involved in the linkage of cytoplasmic proteins to the membrane. *Trends in Biochemical Sciences* 23, 281–282.
- Collawn, J.F., Stangel, M., Kuhn, L.A., *et al.*, 1990. Transferrin receptor internalization sequence YXRF implicates a tight turn as the structural recognition motif for endocytosis. *Cell* 63, 1061–1072.
- Culi, J., Springer, T.A., Mann, R.S., 2004. Boca-dependent maturation of β -propeller/EGF modules in low-density lipoprotein receptor proteins. *EMBO Journal* 23, 1372–1380.
- Davis, C.G., Lehrman, M.A., Russell, D.W., *et al.*, 1986. The J.D. mutation in familial hypercholesterolemia: Amino acid substitution in cytoplasmic domain impedes internalization of LDL receptors. *Cell* 45, 15–24.
- Davis, C.G., Van Driel, I.R., Russell, D.W., Brown, M.S., Goldstein, J.L., 1987. The low density lipoprotein receptor. Identification of amino acids in cytoplasmic domain required for rapid endocytosis. *Journal of Biological Chemistry* 262, 4075–4082.
- Eden, E.R., Naoumova, R.P., Burden, J.J., McCarthy, M.I., Soutar, A.K., 2001. Use of homozygosity mapping to identify a region on chromosome 1 bearing a defective gene that causes autosomal recessive homozygous hypercholesterolemia in two unrelated families. *American Journal of Human Genetics* 68, 653–660.
- Eden, E.R., Patel, D.D., Sun, X., *et al.*, 2002. Restoration of LDL-receptor function in cells from patients with autosomal recessive hypercholesterolemia by retroviral expression of *ARH1*. *Journal of Clinical Investigation* 110, 1695–1702.
- Eden, E.R., Sun, X.M., Patel, D.D., Soutar, A.K., 2007. Adaptor protein Disabled-2 modulates low density lipoprotein (LDL) receptor synthesis in fibroblasts from patients with autosomal recessive hypercholesterolemia. *Human Molecular Genetics* 16, 2751–2759.
- Fokkema, I.F., Den Dunnen, J.T., Taschner, P.E., 2005. LOVD: Easy creation of a locus-specific sequence variation database using an "LSDB-in-a-box" approach. *Human Mutation* 26, 63–68.
- Garcia, C.K., Wilund, K., Arca, M., *et al.*, 2001. Autosomal recessive hypercholesterolemia caused by mutations in a putative LDL receptor adaptor protein. *Science* 292, 1394–1398.
- Ghai, R., Bugarcic, A., Liu, H., *et al.*, 2013. Structural basis for endosomal trafficking of diverse transmembrane cargos by PX-FERM proteins. *Proceedings of the National Academy of Sciences of the United States of America* 110, E643–E652.
- Goldstein, J.L., Brown, M.S., 2009. The LDL receptor. *Arteriosclerosis, Thrombosis, and Vascular Biology* 29, 431–438.
- Hammond, G.R.V., Sim, Y., Lagnado, L., Irvine, R.F., 2009. Reversible binding and rapid diffusion of proteins in complex with inositol lipids serves to coordinate free movement with spatial information. *Journal of Cell Biology* 184, 297–308.
- He, G., Gupta, S., Yi, M., *et al.*, 2002. ARH is a modular adaptor protein that interacts with the LDL receptor, clathrin and AP-2. *Journal of Biological Chemistry* 277, 44044–44049.
- Hobbs, H.H., Russell, D.W., Brown, M.S., Goldstein, J.L., 1990. The LDL receptor locus in familial hypercholesterolemia: Mutational analysis of a membrane protein. *Annual Review of Genetics* 24, 133–170.
- Holthuis, J.C., Menon, A.K., 2014. Lipid landscapes and pipelines in membrane homeostasis. *Nature* 510, 48–57.
- Hong, C., Tontonoz, P., 2014. Liver X receptors in lipid metabolism: Opportunities for drug discovery. *Nature Reviews Drug Discovery* 13, 433–444.
- Hornbeck, P.V., Kornhauser, J.M., Tkachev, S., *et al.*, 2012. PhosphoSitePlus: A comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Research* 40, D261–D270.
- Horton, J.D., Cohen, J.C., Hobbs, H.H., 2009. PCSK9: A convertase that coordinates LDL catabolism. *Journal of Lipid Research* 50 (Suppl), S172–S177.
- Huang, S., Henry, L., Ho, Y.K., Pownall, H.J., Rudenko, G., 2010. Mechanism of LDL binding and release probed by structure-based mutagenesis of the LDL receptor. *Journal of Lipid Research* 51, 297–308.
- Jadot, M., Canfield, W.M., Gregory, W., Kornfeld, S., 1992. Characterization of the signal for rapid internalization of the bovine mannose 6-phosphate/insulin-like growth factor-II receptor. *Journal of Biological Chemistry* 267, 11069–11077.
- Jing, S.Q., Spencer, T., Miller, K., Hopkins, C., Trowbridge, I.S., 1990. Role of the human transferrin receptor cytoplasmic domain in endocytosis: Localization of a specific signal sequence for internalization. *Journal of Cell Biology* 110, 283–294.
- Johnson, K.F., Chan, W., Kornfeld, S., 1990. Cation-dependent mannose 6-phosphate receptor contains two internalization signals in its cytoplasmic domain. *Proceedings of the National Academy of Sciences of the United States of America* 87, 10010–10014.
- Keyel, P.A., Mishra, S.K., Roth, R., *et al.*, 2006. A single common portal for clathrin-mediated endocytosis of distinct cargo governed by cargo-selective adaptors. *Molecular Biology of the Cell* 17, 4300–4317.
- Khachadurian, A.K., Uthman, S.M., 1973. Experiences with the homozygous cases of familial hypercholesterolemia. A report of 52 patients. *Nutrition and Metabolism* 15, 132–140.
- Lee, J., Retamal, C., Cuitino, L., *et al.*, 2008. Adaptor protein sorting nexin 17 regulates amyloid precursor protein trafficking and processing in the early endosomes. *Journal of Biological Chemistry* 283, 11501–11508.
- Leren, T.P., 2014. Sorting an LDL receptor with bound PCSK9 to intracellular degradation. *Atherosclerosis* 237, 76–81.
- Maurer, M.E., Cooper, J.A., 2006. The adaptor protein Dab2 sorts LDL receptors into coated pits independently of AP-2 and ARH. *Journal of Cell Science* 119, 4235–4246.
- Mettlen, M., Loerke, D., Yarar, D., Danuser, G., Schmid, S.L., 2010. Cargo- and adaptor-specific mechanisms regulate clathrin-mediated endocytosis. *Journal of Cell Biology* 188, 919–933.
- Mishra, S.K., Keyel, P.A., Hawryluk, M.J., *et al.*, 2002a. Disabled-2 exhibits the properties of a cargo-selective endocytic clathrin adaptor. *EMBO Journal* 21, 4915–4926.
- Mishra, S.K., Watkins, S.C., Traub, L.M., 2002b. The autosomal recessive hypercholesterolemia (ARH) protein interfaces directly with the clathrin-coat machinery. *Proceedings of the National Academy of Sciences of the United States of America* 99, 16099–16104.
- Morris, S.M., Cooper, J.A., 2001. Disabled-2 colocalizes with the LDLR in clathrin-coated pits and interacts with AP-2. *Traffic* 2, 111–123.

- Motley, A., Bright, N.A., Seaman, M.N., Robinson, M.S., 2003. Clathrin-mediated endocytosis in AP-2-depleted cells. *Journal of Cell Biology* 162, 909–918.
- Muller, S., Dannemark, J., Reinheckel, T., 2012. Specific functions of lysosomal proteases in endocytic and autophagic pathways. *Biochimica et Biophysica Acta* 1824, 34–43.
- Nordestgaard, B.G., Chapman, M.J., Humphries, S.E., *et al.*, 2013. Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: Guidance for clinicians to prevent coronary heart disease: Consensus statement of the European Atherosclerosis Society. *European Heart Journal* 34, 3478–3490a.
- Norman, D., Sun, X.M., Bourbon, M., *et al.*, 1999. Characterization of a novel cellular defect in patients with phenotypic homozygous familial hypercholesterolemia. *Journal of Clinical Investigation* 104, 619–628.
- Ohno, H., Stewart, J., Fournier, M.-C., *et al.*, 1995. Interaction of tyrosine-based sorting signals with clathrin-associated proteins. *Science* 269, 1872–1875.
- Owen, D.J., Evans, P.R., 1998. A structural explanation for the recognition of tyrosine-based endocytotic signals. *Science* 282, 1327–1332.
- Piper, R.C., Dikic, I., Lukacs, G.L., 2014. Ubiquitin-dependent sorting in endocytosis. *Cold Spring Harbor Perspectives in Biology* 6, a016808.
- Rodenburg, K.W., Van Der Horst, D.J., 2005. Lipoprotein-mediated lipid transport in insects: Analogy to the mammalian lipid carrier system and novel concepts for the functioning of LDL receptor family members. *Biochimica et Biophysica Acta* 1736, 10–29.
- Rudenko, G., Deisenhofer, J., 2003. The low-density lipoprotein receptor: Ligands, debates and lore. *Current Opinion in Structural Biology* 13, 683–689.
- Rudenko, G., Henry, L., Henderson, K., *et al.*, 2002. Structure of the LDL receptor extracellular domain at endosomal pH. *Science* 298, 2353–2358.
- Rutledge, A.C., Su, Q., Adeli, K., 2010. Apolipoprotein B100 biogenesis: A complex array of intracellular mechanisms regulating folding, stability, and lipoprotein assembly. *Biochemistry and Cell Biology* 88, 251–267.
- Scotti, E., Calamai, M., Goulbourne, C.N., *et al.*, 2013. IDOL stimulates clathrin-independent endocytosis and MVB-mediated lysosomal degradation of the LDLR. *Molecular and Cellular Biology* 33, 1503–1514.
- Sharpe, L.J., Cook, E.C., Zelcer, N., Brown, A.J., 2014. The UPS and downs of cholesterol homeostasis. *Trends in Biochemical Sciences* 39, 527–535.
- Sorrentino, V., Fouchier, S.W., Motazacker, M.M., *et al.*, 2013a. Identification of a loss-of-function inducible degrader of the low-density lipoprotein receptor variant in individuals with low circulating low-density lipoprotein. *European Heart Journal* 34, 1292–1297.
- Sorrentino, V., Nelson, J.K., Maspero, E., *et al.*, 2013b. The LXR-IDOL axis defines a clathrin-, caveolae-, and dynamin-independent endocytic route for LDLR internalization and lysosomal degradation. *Journal of Lipid Research* 54, 2174–2184.
- Sorrentino, V., Scheer, L., Santos, A., *et al.*, 2011. Distinct functional domains contribute to degradation of the low density lipoprotein receptor (LDLR) by the E3 ubiquitin ligase inducible Degradator of the LDLR (IDOL). *Journal of Biological Chemistry* 286, 30190–30199.
- Sorrentino, V., Zelcer, N., 2012. Post-transcriptional regulation of lipoprotein receptors by the E3-ubiquitin ligase inducible degrader of the low-density lipoprotein receptor. *Current Opinion in Lipidology* 23, 213–219.
- Soutar, A.K., Naoumova, R.P., 2007. Mechanisms of disease: Genetic causes of familial hypercholesterolemia. *Nature Clinical Practice Cardiovascular Medicine* 4, 214–225.
- Steinberg, F., Heesom, K.J., Bass, M.D., Cullen, P.J., 2012. SNX17 protects integrins from degradation by sorting between lysosomal and recycling pathways. *Journal of Cell Biology* 197, 219–230.
- Stockinger, W., Sailer, B., Strasser, V., *et al.*, 2002. The PX-domain protein SNX17 interacts with members of the LDL receptor family and modulates endocytosis of the LDL receptor. *EMBO Journal* 21, 4259–4267.
- Strom, T.B., Tveten, K., Leren, T.P., 2014. PCSK9 acts as a chaperone for the LDL receptor in the endoplasmic reticulum. *Biochemical Journal* 457, 99–105.
- Teasdale, R.D., Collins, B.M., 2012. Insights into the PX (phox-homology) domain and SNX (sorting nexin) protein families: Structures, functions and roles in disease. *Biochemical Journal* 441, 39–59.
- Traub, L.M., 2009. Tickets to ride: Selecting cargo for clathrin-regulated internalization. *Nature Reviews Molecular Cell Biology* 10, 583–596.
- Uhlen, M., Oksvold, P., Fagerberg, L., *et al.*, 2010. Towards a knowledge-based Human Protein Atlas. *Nature Biotechnology* 28, 1248–1250.
- Wang, Y., Huang, Y., Hobbs, H.H., Cohen, J.C., 2012. Molecular characterization of proprotein convertase subtilisin/Kexin type 9 (PCSK9)-mediated degradation of the LDLR. *Journal of Lipid Research* 53, 1932–1943.
- Warren, R.A., Green, F.A., Stenberg, P.E., Enns, C.A., 1998. Distinct saturable pathways for the endocytosis of different tyrosine motifs. *Journal of Biological Chemistry* 273, 17056–17063.
- Wegener, K.L., Partridge, A.W., Han, J., *et al.*, 2007. Structural basis of integrin activation by talin. *Cell* 128, 171–182.
- Williams, R., Schluter, T., Roberts, M.S., *et al.*, 2004. Sorting nexin 17 accelerates internalization yet retards degradation of P-selectin. *Molecular Biology of the Cell* 15, 3095–3105.
- Zelcer, N., Hong, C., Boyadjian, R., Tontonoz, P., 2009. LXR regulates cholesterol uptake through Idol-dependent ubiquitination of the LDL receptor. *Science* 325, 100–104.
- Zhang, L., Fairall, L., Gault, B.T., *et al.*, 2011. The IDOL-UBE2D complex mediates sterol-dependent degradation of the LDL receptor. *Genes & Development* 25, 1262–1274.

Relevant Websites

<http://www.ucl.ac.uk/fh/>
British Heart Foundation.

<http://www.ebi.ac.uk/interpro/>
InterPro.

<http://www.rcsb.org/pdb/home/home.do>
PDB.

<http://www.phosphosite.org/homeAction.do>
PhosphoSitePlus.

<http://www.proteinatlas.org/index.php>
The Human Protein Atlas.

<http://www.uniprot.org/>
UniProt.

Retrograde Transport

L Johannes and C Wunder, Institut Curie, Paris, France

© 2016 Elsevier Inc. All rights reserved.

Introduction

Trafficking of vesicular carriers between the plasma membrane and endomembrane compartments is essential for transport of proteins and other macromolecules to diverse destinations inside and outside of the cell. Membrane trafficking is also required to assure the essential need of cells to maintain membrane homeostasis and to meet specific demands during development, signal perception, and transmission. Two major trafficking routes exist, comprising the endoplasmic reticulum (ER), the Golgi apparatus, and endosomes: the anterograde and retrograde pathways. Both are highly conserved in the eukaryotic kingdom, including plants and some protozoa, e.g., *Toxoplasma gondii*. The anterograde pathway delivers proteins and lipids in a vectorial manner from the ER to the plasma membrane. This outward flow of material is counterbalanced by various retrograde trafficking routes. Retrograde cargo proteins and lipids are endocytosed and transported to endosomes, before being delivered to the trans-Golgi network (TGN), and in some instances further to the ER.

The discovery of the retrograde trafficking pathway started in 1972 by studying the cellular intoxication process mediated by the plant toxin ricin (Olsnes and Pihl, 1972). Ricin is a highly toxic glycoprotein produced by the castor plant *Ricinus communis* and inhibits protein biosynthesis by modifying ribosomal ribonucleic acid (RNA) of toxin-exposed cells (Montanaro *et al.*, 1973). Later, Avrameas and colleagues found that even the Golgi apparatus was a trafficking station in the intoxication process of ricin (Gonatas *et al.*, 1975). Hence, the retrograde cargo protein ricin is binding to cell surface

glycoconjugates having nonreducing terminal galactose (ricin receptors), enters cells via endocytosis and reaches morphologically distinct Golgi membranes. Thus, the first step of the retrograde pathway was defined – trafficking from endosomes to the Golgi apparatus (Figure 1).

Scientists were searching for more than 10 years whether endogenous proteins would also use the retrograde pathway, as described by Olsnes for ricin. In 1985, a cellular protein – transferrin receptor – could be identified in leukemia cells by Snider and Rogers (1985), that also traffics from the plasma membrane to the TGN. This initial discovery was followed very soon by several other studies that used similar approaches based on Golgi-specific enzyme modifications. Of these, the characterization of mannose 6-phosphate receptor (MPRs) trafficking by Duncan and Kornfeld (1988) stood out. These authors provided first evidence for a model in which trafficking between endosomes and the TGN is part of the functional cycle of MPRs. Many lysosomal enzymes are post-translationally modified by mannose 6-phosphate residues. These are recognized by MPRs and shuttled from the TGN to endosomes. The unloaded receptors then return to the TGN for a new transport cycle (Figure 2).

Despite the discovery of several endogenous retrograde cargo molecules, toxins were used as discovery tools to study and understand this pathway further. Sandvig *et al.* (1992) showed for the first time in 1992 that another exogenous protein – the bacterial Shiga toxin (STx) – could reach the ER via the retrograde transport route, similar to ricin. As for ricin, the catalytic A-subunit of Shiga toxin is then translocated to the cytosol to inhibit protein biosynthesis by acting as a N-glycosidase, which cleaves a specific adenine nucleobase

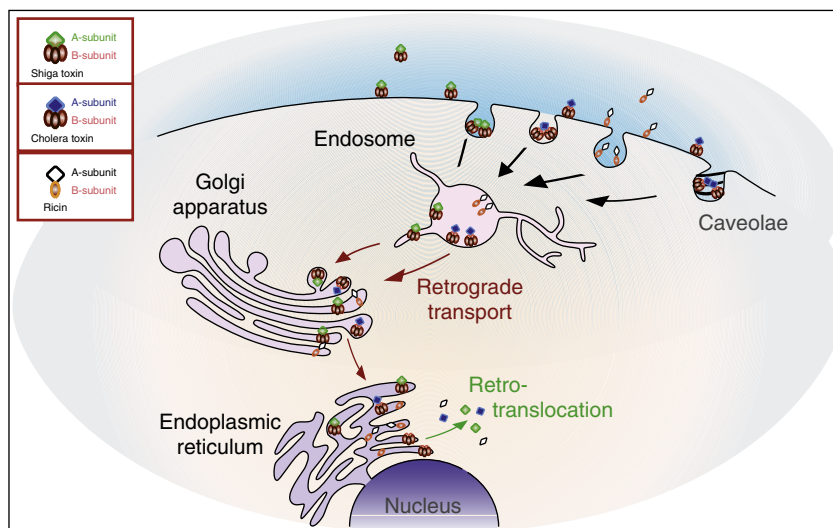


Figure 1: Toxin entry into cells. Via different uptake routes, cholera toxin, Shiga toxin, and ricin reach early endosomes, from where they then traffic to the ER, via the Golgi apparatus. The catalytic A-subunits of these toxins are then retrotranslocated across the ER membrane to the cytosol where they can meet their molecular targets.

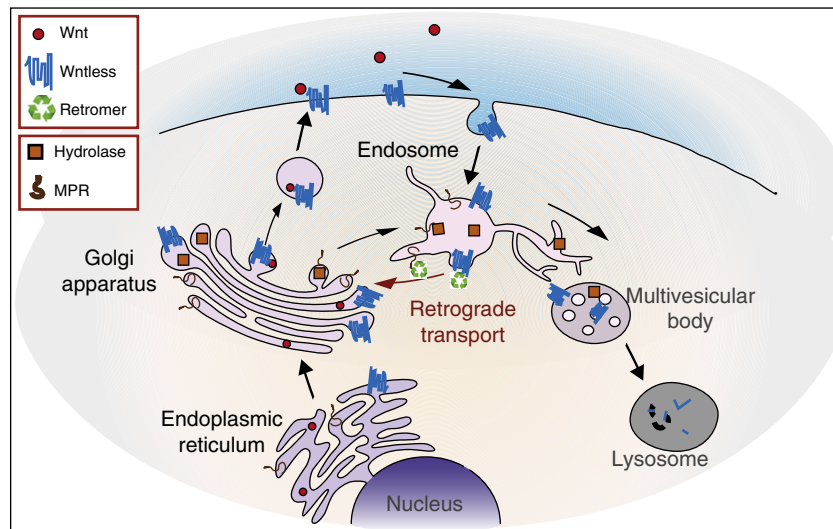


Figure 2 The role of retrograde transport in the cellular functions of Wntless and MPRs. Retromer is of critical importance for the trafficking of Wntless and MPRs from endosomes to the TGN as part of the cellular functions of these receptor molecules. See text for details.

from the 28S rRNA of the ribosomal 60S subunit (reviewed in Johannes and Römer, 2010).

These studies have laid the groundwork for the understanding of the trafficking behavior of a whole class of cytotoxic molecules, AB₅ toxins, for which no pore forming capacity could be demonstrated and that use the retrograde route to enter into cells (Figure 1). These toxins are named after their unique subunit composition, containing a single catalytic A-subunit and a homopentameric B-subunit composed of five B-fragments. In some cases, the A-subunit is covalently linked via a disulfide bridge to the B-subunit (Sandvig and van Deurs, 2002). The catalytically inactive B-subunit is binding to cell surface receptors and serves to mediate transport of the AB₅ protein complex from the cell surface to the lumen of the ER, via vesicular and tubular carriers and intracellular compartments. The B-subunit retains its binding capacity even in the absence of the A-subunit. However, the complete AB₅ holotoxin is required for mediating toxicity.

AB₅ toxins are a medically important class of bacterial exotoxins, which cause devastating human diseases. Prominent members of this class include the above-mentioned STx, which causes life-threatening diarrhea, dysentery, hemorrhagic colitis, and hemorrhagic uremic syndrome; *Escherichia coli* heat-labile enterotoxins and cholera toxin (CTx), which cause endemic and epidemic diarrhea and pertussis toxin (PTx), which is the causative agent for whooping cough (Beddoe et al., 2010). Each year, infections with bacteria that produce these toxins affect millions of individuals and cause more than a million deaths.

Many cellular functions in development and disease depend on the retrograde transport route, such as wnt morphogen gradient formation (Harterink et al., 2011; Yu et al., 2014; Figure 2), intracellular glucose (Esk et al., 2010), iron and copper ion homeostasis (Tabuchi et al., 2010), and glutamate receptor recycling (Zhang et al., 2012). Besides the poisonous effect of several toxins, the medical importance of the retrograde trafficking route was corroborated by many findings. Its dysfunction causes abnormal elevation of amyloid β -peptide

(Burgos et al., 2010); the nef protein of human immunodeficiency virus (HIV-1) favors immune-evasion by down modulating major histocompatibility complex (MHC) class I (Blagoveshchenskaya et al., 2002) and co-stimulatory molecules (Chaudhry et al., 2008). Recently, a new aspect in modulation of the retrograde pathway was discovered: the intracellular bacterium *Legionella pneumophila* is able to shut down the retrograde route in order to promote bacterial replication (Finsel et al., 2013). Taken together, understanding the retrograde trafficking pathway is important to develop efficient therapeutic strategies for a wide range of diseases.

Different Retrograde Pathways

The above-mentioned ground-breaking papers have proven the existence of a transport connection between plasma membrane/endosomes and TGN/Golgi/ER. The trafficking interface between endosomes and the TGN has become a subject of intense investigation in membrane biology research. The complexity of retrograde transport arises from different pathways between endosomes and the TGN. Endosomes are a complex membrane system of tubulo-vacuolar nature of which early endosomes and late endosomes/lysosomes are the major parts (Huotari and Helenius, 2011). There is an ongoing debate on whether these separate endosomal entities of different protein and lipid compositions are linked by vesicular carriers, or are transformed from one to the other by a process of maturation. Exactly how very early endosomes arise is not entirely clear, but it seems likely that membranes and volume are mainly derived from primary endocytic vesicles that fuse with each other. Very early endosomes receive cargo proteins through several endocytic pathways, including clathrin-mediated and clathrin-independent endocytic processes – caveolar, clathrin-independent carrier (CLIC) and ARF6-dependent pathways (Mayor and Pagano, 2007). Therefore very early endosomes are heterogeneous in terms of morphology, localization, composition, and function.

The identity and functional attributes of different endosomes is also defined by the association of cytosolic proteins to the cytoplasmic surface of the endosomal membrane (Huotari and Helenius, 2011).

Very early endosomes then converge into conventional early endosomes. Their limiting membrane contains a mosaic of tubular and vacuolar domains that differ in composition and function (Zerial and McBride, 2001). These include domains enriched in Rab5, Rab4, Rab11, Arf1/COPI, and retromer (Sonnichsen *et al.*, 2000; Bonifacino and Rojas, 2006). Domains located in the tubular extensions allow molecular sorting and generate transport carriers targeted to distinct organelles, including the plasma membrane, the recycling endosomes, and the TGN for retrograde trafficking (Bonifacino and Rojas, 2006). One of the cytosolic key players in retrograde trafficking is the retromer complex. This complex was shown to be involved in trafficking events from different endosomes to the TGN as well as for recycling to the plasma membrane (Temkin *et al.*, 2011). How specificity is achieved will be discussed later.

Pfeffer and colleagues were the first to pinpoint in 1993 an endosomal location from which the MPR recycles to the TGN: late endosomes (Lombardi *et al.*, 1993). In 1998, Johannes and colleagues reported that STx was not found on late endosomal membranes en route to the TGN, which led to the proposal that the toxin bypasses the late endocytic pathway altogether and directly traffics from early endosomes to the TGN (Mallard *et al.*, 1998). Recently, Lieu and Gleeson (2010) suggested that Shiga toxin might in part also traffic from recycling endosomes to the TGN. In summary, retrograde trafficking does occur from different endosomal sites (Figure 3). Routes from late, early, and recycling endosomes seem to operate in parallel. Below, we will discuss these three different scenarios individually. We list key molecular and morphological data that have helped to define them, and point out controversial matters whenever these exist.

Late Endosomes-to-TGN Trafficking

MPRs are the best characterized cargo proteins of the late endosomes-to-TGN retrograde transport process (Rohn *et al.*, 2000). More than 60 lysosomal hydrolases acquire a unique posttranslational modification in the Golgi complex: a mannose 6-phosphate residue, which enables them to bind to MPRs. Two types of MPRs have been identified, the cation-independent (CI-MPR) and the cation-dependent MPR (CD-MPR). Both carry newly synthesized precursor hydrolases from the Golgi to endosomes. MPRs release their ligands upon encountering the low pH of the late endocytic pathway. The receptors then return from late endosomes to the TGN to start another cycle of biosynthetic enzyme transport.

Cargo Sorting and Budding

Lysosomal hydrolase precursors are transported within the lumen of endocytic carrier vesicles to the lysosome for proteolytic activation. A multitude of sorting signals in the cytoplasmic tail of MPRs has been identified that ensure that MPRs are not lost into lysosomes where they would be degraded. A subset of these sorting signals is recognized by tail interacting protein of 47 kDa (TIP47), and strong arguments have been presented for a role of TIP47 in late endosomes-to-TGN trafficking of MPRs (Diaz *et al.*, 1997), even if some complexity has been noticed that may originate from TIP47's additional functions in lipid droplet formation (Bulankina *et al.*, 2009). TIP47 recognizes a phenylalanine-tryptophan motif in the cytoplasmic tail of CD-MPR (Diaz and Pfeffer, 1998), and a more complex domain of CI-MPR. TIP47 also binds Rab9•GTP, upon which its affinity for the CI-MPR is increased threefold (Carroll *et al.*, 2001). Rab9 is itself involved in the retrograde transport of MPRs (Riederer *et al.*, 1994). It has thus been suggested that TIP47 is a coat for MPR sorting into the retrograde pathway.

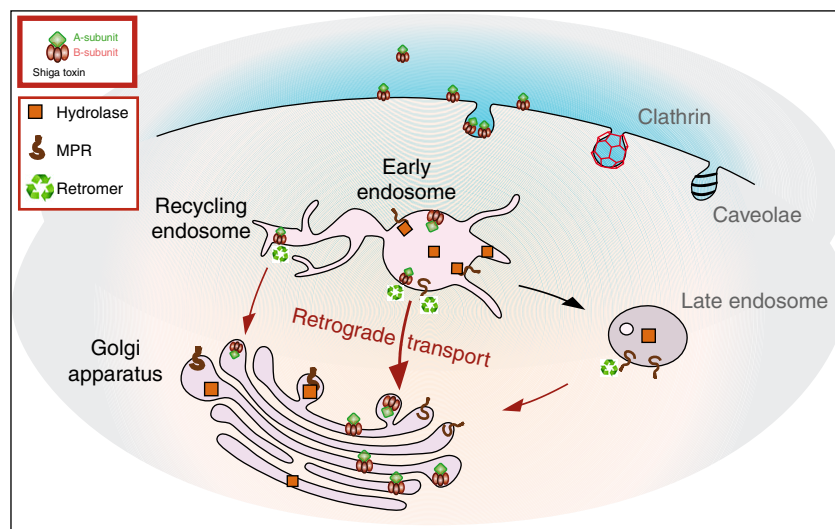


Figure 3 Sites of endosomal retrograde sorting in mammalian cells. Retrograde sorting can operate from different endosomal locations. See text for a detailed discussion of retrograde trafficking from early, recycling, and late endosomes.

Docking and Fusion

The Rab9 GTPase also binds to the GRIP domain golgin protein GCC185 (Golgi coiled-coil protein) (Reddy *et al.*, 2006) for which tethering functions have been suggested. Tethering is the process by which transport intermediates are captured on target membranes to allow for subsequent *N*-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) complex formation and fusion. GCC185 is required for retrograde transport of MPRs (Reddy *et al.*, 2006; Derby *et al.*, 2007) and interacts with the SNARE protein syntaxin 16 that equally has key functions for retrograde transport from late (Ganley *et al.*, 2008) and early endosomes (see below). For the fusion of late endosomes-derived retrograde transport intermediates, syntaxin 16 is in a SNARE complex with syntaxin 10, Vti1a, and vesicle-associated membrane protein (VAMP3) (Ganley *et al.*, 2008).

Early Endosomes-to-TGN Trafficking

Mallard *et al.* (1998) provided first evidence for a direct transport route from early endosomes to the TGN, using Shiga toxin as a model cargo. This transport interface is used by other exogenous (cholera toxin) and endogenous (TGN38) cargo proteins (Ghosh *et al.*, 1998; Amessou *et al.*, 2007; Lieu *et al.*, 2007). Even MPRs may traffic directly from early endosomes to the TGN, in addition to using the late endosomes-TGN interface. Indeed, MPRs accumulate in early endosomes under many conditions that block retrograde trafficking to the TGN, i.e., depletion for the following proteins: retromer (Seaman *et al.*, 2013), the tSNARE syntaxin 10 (Ganley *et al.*, 2008), the clathrin adapter AP1 (Meyer *et al.*, 2000), the phosphatidylinositol 3-phosphate 5-kinase PIKfyve (Rutherford *et al.*, 2006), and the GRIP domain golgin GCC88 (Lieu *et al.*, 2007). Probably the multitude of sorting signals in the cytosolic tails of MPRs has evolved to permit recycling to the TGN from any stage of the endocytic pathway.

Cargo Sorting and Budding

Clathrin is localized on early endosomes and was shown to be required for the retrograde sorting of Shiga toxin (Saint-Pol *et al.*, 2004). In clathrin-depleted cells Shiga toxin perfectly colocalizes with transferrin receptor at times when the molecule has reached the TGN/Golgi in control conditions, demonstrating that the toxin needs clathrin to leave early endosomes. Different adapters have been identified that function with clathrin in retrograde sorting of Shiga toxin and/or other retrograde cargo proteins: AP1 (Meyer *et al.*, 2000), epsinR (Saint-Pol *et al.*, 2004), and oculocerebrorenal syndrome of Lowe (inositol polyphosphate 5-phosphatase) (OCRL) (Choudhury *et al.*, 2005). For some cargo proteins with acidic cluster motifs in their cytosolic tail, the phosphofurin acidic cluster sorting protein 1 (PACS1) appears to stabilize binding to AP1, and the CI-MPR accumulates in early endosomes upon overexpression of a dominant negative PACS1 (Scott *et al.*, 2006).

The retromer complex was identified first in yeast and then in mammalian cells. It has been shown that retromer is localized on the cytosolic leaflet of early endosomes and

required for retrograde transport. Retromer comprises two subcomplexes, a vps protein heterotrimer (Vps26, Vps29, and Vps35) and a sorting nexin (SNX) dimer. The vps heterotrimer binds cargo proteins (Seaman *et al.*, 2013). However, different cargo proteins are recognized by different vacuolar protein sorting (VPS) proteins of the retromer heterotrimer (A-ALP and Vps10p bind to Vps35p, CI-MPR shows bipartite interactions with VPS35 and VPS29, sortilin to VPS35, divalent metal transporter 1 (DMT1-II) to VPS35, and sorLA to VPS26).

The SNX subcomplex senses and/or generates curvature. Latest work from the Cullen laboratory indeed suggests a contribution of the sorting nexins to membrane bending, as purified sorting nexins were shown to deform and tubulate lipid vesicles (van Weering *et al.*, 2012). As such, retromer could function as a coat. However, retromer does not form the same type of electron-dense coats as observed for classical coatamer protein complex-I (COPI), coatamer protein complex-II (COPII), and clathrin. In addition, the molecular mechanism by which retromer sorts cargo seems to be distinct from that of classical membrane coats (Seaman *et al.*, 2013), but somehow similar to other endosomal sorting complexes, like the endosomal sorting complexes required for transport (ESCRT). Retromer seems to regulate trafficking of cargo proteins on non-flat membranes, rather within morphologically distinct tubulovesicular membranes, suggesting a novel mechanism of action. Furthermore, when retromer vps heterotrimer proteins are depleted, retrograde cargo-containing tubules still form that fail to detach from early endosomes (Popoff *et al.*, 2007, 2009). These findings suggest that vps retromer heterotrimer may play a critical role for scission. Whether this function is fulfilled via amphipathic helices/BAR-domains of SNX proteins, or indirectly via recruitment of a scission factor still needs to be determined. One possibility how retromer fulfills a scission function might be via its capacity to recruit the actin nucleation promotion factor Wiskott–Aldrich syndrome protein and scar homolog protein (WASH) and to drive the polymerization of actin on early endosomal membranes (Seaman *et al.*, 2013). A second hypothesis might be an indirect effect of the WASH complex, via the recruitment of dynamin-2 (Derivery *et al.*, 2009). Whether and how dynamin contributes to membrane scission on early endosomes in human cells has not yet been convincingly documented. In yeast, a dynamin-related protein, termed Vps1, promotes fission of SNX-coated retromer tubules on early endosomes (Chi *et al.*, 2014).

Important differences exist between the prototypical retromer in yeast and retromer in higher eukaryotes. A key difference is that, in mammalian cells, retromer is not a stable heteropentamer as in yeast, but a much more transient association of the cargo-selective trimer (Vps35, Vps26, and Vps29) and the Snx proteins (Snx1 or Snx2 with Snx5 or Snx6) (Seaman *et al.*, 2013). In addition, retromer exists in different subcomplexes, depending on which SNX protein is involved. In mammalian cells, SNX proteins indeed seem to determine the trafficking route and thereby the specificity of retrograde trafficking. Furthermore, many accessory retromer proteins that were identified from studies in higher eukaryotes are not conserved in yeast. The yeast retromer complex seems to function independent of the WASH complex. Thus, the exact

scope of functions of retromer accessory proteins is not clear at this stage.

Lipids also play critical roles in retrograde sorting. Notably the phosphatidylinositol lipids phosphatidylinositol 3-phosphate (Skånland *et al.*, 2007) and phosphatidylinositol 3,5-bisphosphate (Rutherford *et al.*, 2006) are directly involved in the membrane recruitment of sorting nexins, and may also affect retrograde sorting via Hrs (Popoff *et al.*, 2009). Furthermore, a redistribution of retromer components has been observed in glycosphingolipid-depleted cells (Raa *et al.*, 2009), in agreement with earlier studies that had found a role for raft lipids in retrograde sorting (Falguières *et al.*, 2001). By what mechanism mesoscale raft lipid organization affects the retromer complex, and whether specialized lipid domains are required for retrograde trafficking from early endosomes has yet to be determined.

Docking and Fusion

Early endosome-derived vesicles are trafficking further to the TGN, and this transport is microtubule dependent (Hehnl *et al.*, 2006). The Golgi-associated retrograde protein (GARP) complex was identified for tethering of MPR positive vesicles to TGN membranes (Perez-Victoria *et al.*, 2008). The GARP complex is recruited at the level of early endosomes, and then functions at the TGN to promote SNARE complex formation in interaction with syntaxin 6, syntaxin 16, and VAMP4 (Perez-Victoria and Bonifacino, 2009), three SNARE proteins that have previously been implicated in early endosomes-to-TGN trafficking (Mallard *et al.*, 2002). Other tethering factors for early endosomal retrograde cargo proteins are GCC88 and GCC185 (Derby *et al.*, 2007; Lieu *et al.*, 2007).

Recycling Endosomes-to-TGN Trafficking

In a comparative study it was found that depletion of the tethering factor GCC185 inhibited retrograde transport of Shiga toxin, but not that of the endogenous cargo protein TGN38 (Derby *et al.*, 2007), while the opposite was observed for the depletion of the tethering factor GCC85 (Lieu *et al.*, 2007). Since Shiga toxin, but not TGN38, could be detected in recycling endosomes under different experimental conditions, the authors concluded that TGN38 was transported directly

from early endosomes to the TGN, while Shiga toxin could also transit via recycling endosomes (Lieu and Gleeson, 2010). This conclusion is in apparent contradiction with earlier studies that had described the retrograde trafficking of TGN38 via recycling endosomes (Ghosh *et al.*, 1998), and it is difficult to reconcile with the established function of early/maturing endosomal retromer in the trafficking of Shiga toxin to the TGN. Furthermore, as mentioned above, recycling endosomes remain poorly defined in molecular terms, and whether they have evolved a truly independent retrograde sorting machinery, possibly involving Rab11 (Chaudhry *et al.*, 2008; Wilcke *et al.*, 2000) remains to be established.

Biophysical Mechanisms for Retrograde Trafficking

All trafficking events of proteins and lipids between distinct membrane-enclosed organelles are following the same mechanistic principles. The course of action per se for generic trafficking processes is as follows: cargo sorting, membrane bending and bud formation, scission in the neck region of the invaginated bud or tubule, and finally fusion of transport intermediates with acceptor organelles. In the following, we have summarized biophysical mechanisms of each of these steps, based on the list of established retrograde trafficking factors.

Membrane Bending

The exact mechanism of membrane bending for retrograde trafficking from early endosomes is not known. However, different SNXs bearing BAR-domains and/or amphipathic helices were analyzed by *in vitro* studies (van Weering *et al.*, 2012). Therefore, two SNX-driven mechanisms for membrane bending can be envisioned (Figure 4): (1) asymmetric insertion of amphipathic helices and/or (2) scaffolding by BAR-domain containing SNXs. Whether SNXs are able to form scaffolds at physiological concentrations such as to drive membrane deformation (Sorre *et al.*, 2012), still needs to be determined. Recently the concerted action of retromer SNXs and the WASH complex on tubule formation was supported by Seaman and colleagues (Freeman *et al.*, 2014). This report indicated the binding of RME-8 to retromer SNXs and the WASH complex

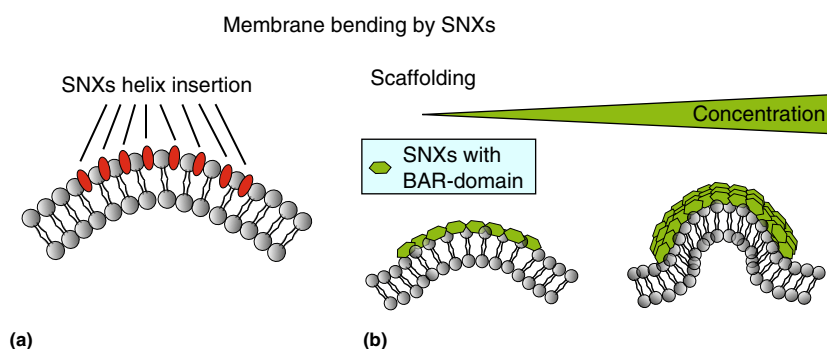


Figure 4 Membrane bending mechanisms in relation to SNX proteins. (a) Insertion of helices leads to asymmetric expansion of the cytosolic leaflet and to subsequent membrane bending. (b) BAR domain binding can serve to recognize membrane curvature (low BAR-domain protein concentration on membranes), or to drive membrane bending through a scaffolding effect (high concentration).

protein FAM21. Interestingly, clathrin could be localized to early endosomes (Popoff *et al.*, 2009; Esk *et al.*, 2010; Freeman *et al.*, 2014), and RME-8 also binds to the clathrin uncoating ATPase Hsc70 (Girard *et al.*, 2005). However, no reports currently exist on interactions between clathrin and components of the retromer machinery. Whether clathrin is directly involved in membrane deformation for other retrograde cargo protein requires further studies but seems likely when considering the function of the clathrin interactors AP1 (Meyer *et al.*, 2000; Natsume *et al.*, 2006), epsinR (Saint-Pol *et al.*, 2004), and OCRL (Choudhury *et al.*, 2005) in retrograde sorting on early endosomes.

Scission

The first indication for a molecular player for a scission process of endosomal tubules came from knock down experiments of vps retromer proteins (Popoff *et al.*, 2007, 2009). Endosomal tubules formed that failed to detach from early endosomes, pointing toward a function for retromer or retromer binding proteins in the membrane scission process. Recently, the WASH complex was discovered as a binding partner of retromer and a nucleation point for actin polymerization on early endosomes (Derivery *et al.*, 2009; Gomez and Billadeau, 2009). Pulling forces that are applied to the budding structure by molecular motors or polymerizing actin filaments are expected to reduce the energy barrier involved in neck or tubule fission (Liu *et al.*, 2006). However, even a reduced energy barrier may still require a dedicated scission factor to finalize the reaction. Several lines of evidence indicate that dynamin may provide this function. First, an interaction between WASH and dynamin was identified (Derivery *et al.*, 2009); second, the WASH complex was found at the base of extended tubules (Derivery *et al.*, 2009); third, dynamin inhibition induced extended long tubules on early endosomes (Mesaki *et al.*, 2011). A combined function of actin and dynamin on membrane scission in the retrograde route was recently postulated (Johannes *et al.*, 2014).

In summary, at least in human, a model can be proposed for the concerted action of different molecules within the same macromolecular complex. The VPS trimer recruits cargo and SNX proteins, which are deforming the membrane via a BAR domain-dependent scaffolding mechanism and assymmetric insertion of amphipathic helices. The RME-8 protein is recruited and to link the SNXs and the WASH subcomplexes for a concerted action in the scission process. WASH is able to initiate actin polymerization and to recruit dynamin.

Biomedical Applications of Retrograde Trafficking

Initially, bacterial toxins were used to eliminate cancer cells, or at reduced levels to alter cellular processes that control proliferation, apoptosis, and differentiation. Specificity for cancer cells was reached by coupling cytotoxic toxins to antibodies (immunotoxins) (Becker and Benhar, 2012). However, for most antibody-toxin conjugates clinical evaluation did not go beyond phase I trials for various reasons, mostly linked to toxicity and lack of efficacy.

Since several pathogens are using the retrograde pathway for the targeting of cytotoxic proteins into host cells, scientist

are developing tools to use the same pathway for pharmaceutical delivery (Tarrago-Trani and Storrie, 2007). Currently, three different applications of retrograde transport research can be identified (Tartour *et al.*, 2002; Moron *et al.*, 2004; Engedal *et al.*, 2011): the nontoxic B-subunits of AB₅ toxins can be used as carriers for cytotoxic drugs in cancer therapy, as carriers for imaging agents, or to deliver antigens to dendritic cells in order to trigger a therapeutic immune response.

Targeted Tumor Therapy

Altered glycosylation patterns of tumor cells are a universal feature in carcinogenesis, and may affect cell signaling, adherence, and motility (Hakomori, 1996). Remarkably, human cancers are often characterized by modified glycosphingolipid composition and metabolism, and several tumor-associated antigens have indeed been found to be glycosphingolipids (Hakomori and Zhang, 1997). The retrograde cargo protein STxB recognizes a specific glycosphingolipid, Gb3, which is highly expressed on human cancers (Garipey, 2001; Johannes and Decaudin, 2005; Engedal *et al.*, 2011). By following the retrograde route, STxB avoids recycling to the plasma membrane or degradation in the late endocytic pathway, which are features that may be of general interest in the development of innovative cancer targeting strategies (reviewed in Tarrago-Trani and Storrie, 2007). Since intratumoral injection of Shiga toxin inhibits tumor growth (Arab *et al.*, 1999; Ishitoya *et al.*, 2004), STxB might be an appropriate tool for targeted therapy.

Following intravenous injection or oral application in mice, STxB has indeed been shown to accumulate in adenocarcinomas of the digestive tract (Janssen *et al.*, 2006). A prodrug composition using the topoisomerase I inhibitor SN38 has been developed that is specifically activated in membranes of the secretory pathway that have been targeted from the outside of cells, via the retrograde route (El Alaoui *et al.*, 2007). Exploitation of this delivery modality may be particularly beneficial in therapeutic strategies requiring prolonged association of the prodrug with tumor cells to ensure efficient prodrug conversion to the active form.

Biomedical Imaging

Gb3 overexpressing cancer cells might be visualized by injecting nontoxic STxB molecules coupled to imaging agents (Janssen *et al.*, 2006; Viel *et al.*, 2008). The imaging agents are covalently coupled to the delivery tool and highlight Gb3 expressing cells. The following imaging agents were successfully coupled to STxB and delivered to tumors in mice: fluorescent STxB (Viel *et al.*, 2008), ¹⁸F-labeled STxB followed by Positron emission tomography (PET) imaging (Janssen *et al.*, 2006), and ultrasound contrast agents (Couture *et al.*, 2011).

Immunotherapy

Protein toxins have also been used to generate specific cytotoxic T lymphocytes (CTLs) against tumor antigens to which they have been linked chemically or genetically (Smith *et al.*, 2002). The function of the toxin-antigen conjugates is in this case to facilitate cytosolic entry of antigens, which are then degraded by

proteasomes and displayed on MHC class I molecules to prime a CTL response. This approach to generate therapeutic vaccines has been used with adenylate cyclase of *Bordetella pertussis* (Fayolle *et al.*, 1996), *Pseudomonas* exotoxin A (PEx A) (Donnelly *et al.*, 1993), *E. coli* bacterial heat-labile enterotoxin (Hearn *et al.*, 2004), and STxB (Haicheur *et al.*, 2000).

For example, antigens that are chemically coupled to STxB are selectively targeted to dendritic cells expressing the Gb3 receptor, and it has been shown that the MHC class I-restricted presentation of the vectorized antigens is proteasome- and transporter-associated with antigen processing (TAP) dependent (Haicheur *et al.*, 2000). This strongly suggests that the exogenous antigens are indeed delivered to the cytosolic compartment. In mice, the delivery of exogenous antigens to dendritic cells induces a potent cytotoxic T-lymphocyte response and protects the animals against experimentally induced tumor growth (Vingert *et al.*, 2006; Pere *et al.*, 2011; Badoual *et al.*, 2013; Sandoval *et al.*, 2013).

One could envisage the design of new synthetic molecules that interact with target cells similarly to the toxins discussed above, and mimic the toxin trafficking pathway to deliver molecules specifically to the Golgi or ER in order to correct enzymatic defects, or elicit toxicity for the treatment of cancer. Moreover, given its avoidance of the lysosome, perhaps this route could be of better use for gene delivery and enable prolonged survival of gene therapy vectors after targeted cell entry.

See also: Vertical Integration: Applications: Neurogenesis in the Adult Brain

References

- Amessou, M., Fradagrada, A., Falguières, T., *et al.*, 2007. Syntaxin 16 and syntaxin 5 control retrograde transport of several exogenous and endogenous cargo proteins. *Journal of Cell Science* 120, 1457–1468.
- Arab, S., Rutka, J., Lingwood, C., 1999. Verotoxin induces apoptosis and the complete, rapid, long-term elimination of human astrocytoma xenografts in nude mice. *Oncology Research* 11 (1), 33–39.
- Badoual, C., Hans, S., Merillon, N., *et al.*, 2013. PD-1-expressing tumor-infiltrating T cells are a favorable prognostic biomarker in HPV associated head and neck cancer. *Cancer Research* 73 (1), 128–138. doi:0008-5472.CAN-12-2606 [pii] 10.1158/0008-5472.CAN-12-2606 [doi].
- Becker, N., Benhar, I., 2012. Antibody-based immunotoxins for the treatment of cancer. *Antibodies* 1 (1), 39–69.
- Beddoe, T., Paton, A.W., Le Nours, J., Rossjohn, J., Paton, J.C., 2010. Structure, biological functions and applications of the AB(5) toxins. *Trends in Biochemical Sciences* 35 (7), 411–418. doi:10.1016/j.tibs.2010.02.003.
- Blagoveshchenskaya, A.D., Thomas, L., Feliciangeli, S.F., Hung, C.H., Thomas, G., 2002. HIV-1 Nef downregulates MHC-I by a PACS-1- and PI3K-regulated ARF6 endocytic pathway. *Cell* 111 (6), 853–866.
- Bonifacio, J.S., Rojas, R., 2006. Retrograde transport from endosomes to the trans-Golgi network. *Nature Reviews Molecular Cell Biology* 7 (8), 568–579.
- Bulankina, A.V., Deggerich, A., Wenzel, D., *et al.*, 2009. TIP47 functions in the biogenesis of lipid droplets. *Journal of Cell Biology* 185 (4), 641–655. doi: jcb.200812042. [pii] 10.1083/jcb.200812042 [doi].
- Burgos, P.V., Mardones, G.A., Rojas, A.L., *et al.*, 2010. Sorting of the Alzheimer's disease amyloid precursor protein mediated by the AP-4 complex. *Developmental Cell* 18 (3), 425–436. doi:S1534-5807(10)00096-1 [pii] 10.1016/j.devcel.2010.01.015 [doi].
- Carroll, K.S., Hanna, J., Simon, I., *et al.*, 2001. Role of Rab9 GTPase in facilitating receptor recruitment by TIP47. *Science* 292 (5520), 1373–1376.
- Chaudhry, A., Das, S.R., Jameel, S., *et al.*, 2008. HIV-1 Nef induces a Rab11-dependent routing of endocytosed immune costimulatory proteins CD80 and CD86 to the Golgi. *Traffic* 9 (11), 1925–1935. doi:TRA802 [pii] 10.1111/j.1600-0854.2008.00802.x [doi].
- Chi, R.J., Liu, J., West, M., *et al.*, 2014. Fission of SNX-BAR-coated endosomal retrograde transport carriers is promoted by the dynamin-related protein Vps1. *Journal of Cell Biology* 204 (5), 793–806. doi:jcb.201309084 [pii] 10.1083/jcb.201309084 [doi].
- Choudhury, R., Diao, A., Zhang, F., *et al.*, 2005. Lowe syndrome protein OCRL1 interacts with clathrin and regulates protein trafficking between endosomes and the trans-Golgi network. *Molecular Biology of the Cell* 16, 3467–3479.
- Couture, O., Dransart, E., Dehay, S., *et al.*, 2011. Ultrasound contrast agent delivery using Shiga toxin B-subunit. *Molecular Imaging* 10 (2), 135–143.
- Derby, M.C., Lieu, Z.Z., Brown, D., *et al.*, 2007. The trans-Golgi network golgin, GCC185, is required for endosome-to-Golgi transport and maintenance of Golgi structure. *Traffic* 8 (6), 758–773.
- Derivery, E., Sousa, C., Gautier, J.J., *et al.*, 2009. The Arp2/3 activator WASH controls the fission of endosomes through a large multiprotein complex. *Developmental Cell* 17 (5), 712–723. doi:S1534-5807(09)00392-X [pii] 10.1016/j.devcel.2009.09.010 [doi].
- Diaz, E., Pfeffer, S.R., 1998. TIP47: A cargo selection device for mannose 6-phosphate receptor trafficking. *Cell* 93 (3), 433–443.
- Diaz, E., Schimmoller, F., Pfeffer, S.R., 1997. A novel Rab9 effector required for endosome-to-TGN transport. *Journal of Cell Biology* 138 (2), 283–290.
- Donnelly, J.J., Ulmer, J.B., Hawe, L.A., *et al.*, 1993. Targeted delivery of peptide epitopes to class I major histocompatibility molecules by a modified *Pseudomonas* exotoxin. *Proceedings of the National Academy of Sciences of the United States of America* 90 (8), 3530–3540.
- Duncan, J.R., Kornfeld, S., 1988. Intracellular movement of two mannose 6-phosphate receptors: Return to the Golgi apparatus. *Journal of Cell Biology* 106 (3), 617–628.
- El Alaoui, A., Schmidt, F., Amessou, M., *et al.*, 2007. Shiga toxin-mediated retrograde delivery of a topoisomerase I inhibitor prodrug. *Angewandte Chemie International Edition in English* 46, 6469–6472.
- Engedal, N., Skotland, T., Torgersen, M.L., Sandvig, K., 2011. Shiga toxin and its use in targeted cancer therapy and imaging. *Microbial Biotechnology* 4 (1), 32–46.
- Esk, C., Chen, C.-Y., Johannes, L., Brodsky, F.M., 2010. The clathrin heavy chain isoform CHC22 functions in a novel endosomal sorting step. *Journal of Cell Biology* 188, 131–144.
- Falguières, T., Mallard, F., Baron, C.L., *et al.*, 2001. Targeting of Shiga toxin B-subunit to retrograde transport route in association with detergent resistant membranes. *Molecular Biology of the Cell* 12, 2453–2468.
- Fayolle, C., Sebo, P., Ladant, D., Ullmann, A., Leclerc, C., 1996. *In vivo* induction of CTL responses by recombinant adenylate cyclase of *Bordetella pertussis* carrying viral CD8 + T cell epitopes. *Journal of Immunology* 156 (12), 4697–4706.
- Finsel, I., Ragaz, C., Hoffmann, C., *et al.*, 2013. The Legionella effector RldL inhibits retromer function to promote intracellular replication. *Cell Host & Microbe* 14 (1), 38–50.
- Freeman, C.L., Hesketh, G., Seaman, M.N., 2014. RME-8 coordinates the activity of the RME complex with the function of the retromer SNX dimer to control endosomal tubulation. *Journal of Cell Science* 127 (Pt 9), 2053–2070. doi: jcs.144659 [pii] 10.1242/jcs.144659 [doi].
- Ganley, I.G., Espinosa, E., Pfeffer, S.R., 2008. A syntaxin 10-SNARE complex distinguishes two distinct transport routes from endosomes to the trans-Golgi in human cells. *Journal of Cell Biology* 180 (1), 159–172.
- Garipey, J., 2001. The use of Shiga-like toxin 1 in cancer therapy. *Critical Reviews in Oncology/Hematology* 39 (1-2), 99–106.
- Ghosh, R.N., Mallet, W.G., Soe, T.T., McGraw, T.E., Maxfield, F.R., 1998. An endocytosed TGN38 chimeric protein is delivered to the TGN after trafficking through the endocytic recycling compartment in CHO cells. *Journal of Cell Biology* 142 (4), 923–936.
- Girard, M., Poupon, V., Blondeau, F., McPherson, P.S., 2005. The DnaJ-domain protein RME-8 functions in endosomal trafficking. *Journal of Biological Chemistry* 280 (48), 40135–40143.
- Gomez, T.S., Billadeau, D.D., 2009. A FAM21-containing WASH complex regulates retromer-dependent sorting. *Developmental Cell* 17 (5), 699–711. doi:S1534-5807(09)00391-8 [pii] 10.1016/j.devcel.2009.09.009 [doi].
- Gonatas, N.K., Steiber, A., Kim, S.U., Graham, D.J., Avrameas, S., 1975. Internalization of neuronal plasma membrane ricin receptors into the Golgi apparatus. *Experimental Cell Research* 94 (2), 426–431.
- Haicheur, N., Bismuth, E., Bosset, S., *et al.*, 2000. The B-subunit of Shiga toxin fused to a tumor antigen elicits CTL and targets dendritic cells to allow MHC class I restricted presentation of peptides derived from exogenous antigens. *Journal of Immunology* 165, 3301–3308.

- Hakomori, S., 1996. Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism. *Cancer Research* 56 (23), 5309–5318.
- Hakomori, S., Zhang, Y., 1997. Glycosphingolipid antigens and cancer therapy. *Chemistry and Biology* 4 (2), 97–104.
- Harterink, M., Port, F., Lorenowicz, M.J., *et al.*, 2011. A SNX3-dependent retromer pathway mediates retrograde transport of the Wnt sorting receptor Wntless and is required for Wnt secretion. *Nature Cell Biology* 13 (8), 914–923. doi:ncb2281 [pii] 10.1038/ncb2281 [doi].
- Hearn, A.R., de Haan, L., Pemberton, A.J., Hirst, T.R., Rivett, A.J., 2004. Trafficking of exogenous peptides into proteasome-dependent major histocompatibility complex class I pathway following enterotoxin B subunit-mediated delivery. *Journal of Biological Chemistry* 279 (49), 51315–51322.
- Hehny, H., Sheff, D., Stammes, M., 2006. Shiga toxin facilitates its retrograde transport by modifying microtubule dynamics. *Molecular Biology of the Cell* 17 (10), 4379–4389.
- Huotari, J., Helenius, A., 2011. Endosome maturation. *EMBO Journal* 30 (17), 3481–3500. doi:emboj2011286 [pii] 10.1038/emboj.2011.286 [doi].
- Ishitoya, S., Kurazono, H., Nishiyama, H., *et al.*, 2004. Verotoxin induces rapid elimination of human renal tumor xenografts in SCID mice. *Journal of Urology* 171 (3), 1309–1313.
- Janssen, K.-P., Vignjevic, D., Boisgard, R., *et al.*, 2006. *In vivo* tumor targeting using a novel intestinal pathogen-based delivery approach. *Cancer Research* 66, 7230–7236.
- Johannes, L., Decaudin, D., 2005. Protein toxins: Intracellular trafficking for targeted therapy. *Gene Therapy* 16, 1360–1368.
- Johannes, L., Römer, W., 2010. Shiga toxins – From cell biology to biomedical applications. *Nature Reviews Microbiology* 8, 105–116.
- Johannes, L., Wunder, C., Bassereau, P., 2014. Bending “on the rocks” – A cocktail of biophysical modules to build endocytic pathways. *Cold Spring Harbor Perspectives in Biology* 6, 1.
- Lieu, Z.Z., Derby, M.C., Teasdale, R.D., *et al.*, 2007. The Golgin GCC88 is required for efficient retrograde transport of cargo from the early endosomes to the trans-Golgi network. *Molecular Biology of the Cell* 18 (12), 4979–4991.
- Lieu, Z.Z., Gleeson, P.A., 2010. Identification of different itineraries and retromer components for endosome-to-Golgi transport of TGN38 and Shiga toxin. *European Journal of Cell Biology* 89 (5), 379–393. doi:S0171-9335(10)00009-9 [pii] 10.1016/j.ejcb.2009.10.021 [doi].
- Liu, J., Kaksonen, M., Drubin, D.G., Oster, G., 2006. Endocytic vesicle scission by lipid phase boundary forces. *Proceedings of the National Academy of Sciences of the United States of America* 103, 10277–10282.
- Lombardi, D., Soldati, T., Riederer, M.A., *et al.*, 1993. Rab9 functions in transport between late endosomes and the trans Golgi network. *EMBO Journal* 12 (2), 677–682.
- Mallard, F., Tang, B.L., Galli, T., *et al.*, 2002. Early/recycling endosomes-to-TGN transport involves two SNARE complexes and a Rab6 isoform. *Journal of Cell Biology* 156, 653–664.
- Mallard, F., Tenza, D., Antony, C., *et al.*, 1998. Direct pathway from early/recycling endosomes to the Golgi apparatus revealed through the study of Shiga toxin B-fragment transport. *Journal of Cell Biology* 143, 973–990.
- Mayor, S., Pagano, R.E., 2007. Pathways of clathrin-independent endocytosis. *Nature Reviews Molecular Cell Biology* 8 (8), 603–612.
- Mesaki, K., Tanabe, K., Obayashi, M., Oe, N., Takei, K., 2011. Fission of tubular endosomes triggers endosomal acidification and movement. *PLoS ONE* 6 (5), e19764. [doi] PONE-D-11-00108 [pii] doi:10.1371/journal.pone.0019764.
- Meyer, C., Zizioli, D., Lausmann, S., *et al.*, 2000. μ 1 A-adaptin-deficient mice: Lethality, loss of AP-1 binding and rerouting of mannose 6-phosphate receptors. *EMBO Journal* 19 (10), 2193–2203.
- Montanaro, L., Sperti, S., Stirpe, F., 1973. Inhibition by ricin of protein synthesis *in vitro*. Ribosomes as the target of the toxin. *Biochemical Journal* 136 (3), 677–683.
- Moron, G., Dadaglio, G., Leclerc, C., 2004. New tools for antigen delivery to the MHC class I pathway. *Trends in Immunology* 25 (2), 92–97.
- Natsume, W., Tanabe, K., Kon, S., *et al.*, 2006. SMAP2, a novel ARF GTPase-activating protein, interacts with clathrin and clathrin assembly protein and functions on the AP-1-positive early endosome/trans-Golgi network. *Molecular Biology of the Cell* 17 (6), 2592–2603.
- Olsnes, S., Pihl, A., 1972. Ricin – A potent inhibitor of protein synthesis. *FEBS Letters* 20 (3), 327–329. doi:0014-5793(72)80098-X [pii].
- Pere, H., Montier, Y., Bayry, J., *et al.*, 2011. A CCR4 antagonist combined with protein or DNA based vaccines is efficient to break tolerance and elicit CD8 + T cells directed against various self antigens. *Blood* 118 (18), 4853–4862.
- Perez-Victoria, F.J., Bonifacio, J.S., 2009. Dual roles of the mammalian GARP complex in tethering and SNARE complex assembly at the trans-Golgi network. *Molecular Cell Biology* 29 (19), 5251–5263. doi:MCB.00495-09 [pii] 10.1128/MCB.00495-09 [doi].
- Perez-Victoria, F.J., Mardones, G.A., Bonifacio, J.S., 2008. Requirement of the human GARP complex for mannose 6-phosphate-receptor-dependent sorting of cathepsin D to lysosomes. *Molecular Biology of the Cell* 19 (6), 2350–2362. doi:E07-11-1189 [pii] 10.1091/mbc.E07-11-1189 [doi].
- Popoff, V., Mardones, G.A., Bai, S.K., *et al.*, 2009. Analysis of articulation between clathrin and retromer in retrograde sorting on early endosomes. *Traffic* 10, 1868–1880.
- Popoff, V., Mardones, G.A., Tenza, D., *et al.*, 2007. The retromer complex and clathrin define a post-early endosomal retrograde exit site. *Journal of Cell Science* 120, 2022–2031.
- Raa, H., Grimmer, S., Schwudke, D., *et al.*, 2009. Glycosphingolipid requirements for endosome-to-Golgi transport of Shiga toxin. *Traffic* 10 (7), 868–882. doi:TRA919 [pii] 10.1111/j.1600-0854.2009.00919.x[doi].
- Reddy, J.V., Burguete, A.S., Sridevi, K., *et al.*, 2006. A functional role for the GCC185 golgin in mannose 6-phosphate receptor recycling. *Molecular Biology of the Cell* 17 (10), 4353–4363.
- Riederer, M.A., Soldati, T., Shapiro, A.D., Lin, J., Pfeffer, S.R., 1994. Lysosome biogenesis requires Rab9 function and receptor recycling from endosomes to the trans-Golgi network. *Journal of Cell Biology* 125 (3), 573–582.
- Rohn, W.M., Rouille, Y., Waguri, S., Hoflack, B., 2000. Bi-directional trafficking between the trans-Golgi network and the endosomal/lysosomal system. *Journal of Cell Science* 113 (Pt 12), 2093–2101.
- Rutherford, A.C., Traer, C., Wassmer, T., *et al.*, 2006. The mammalian phosphatidylinositol 3-phosphate 5-kinase (PIKfyve) regulates endosome-to-TGN retrograde transport. *Journal of Cell Science* 119 (Pt 19), 3944–3957.
- Saint-Pol, A., Yélamos, B., Amessou, M., *et al.*, 2004. Clathrin adaptor epsinR is required for retrograde sorting on early endosomal membranes. *Developmental Cell* 6, 525–538.
- Sandoval, F., Terme, M., Nizard, M., *et al.*, 2013. Mucosal imprinting of vaccine-induced CD8 + T cells is crucial to inhibit the growth of mucosal tumors. *Science Translational Medicine* 5 (172), 172ra20.
- Sandvig, K., Garred, O., Prydz, K., *et al.*, 1992. Retrograde transport of endocytosed Shiga toxin to the endoplasmic reticulum. *Nature* 358 (6386), 510–512.
- Sandvig, K., van Deurs, B., 2002. Transport of protein toxins into cells: Pathways used by ricin, cholera toxin and Shiga toxin. *FEBS Letters* 529 (1), 49–53. doi: S0014579302031824 [pii].
- Scott, G.K., Fei, H., Thomas, L., Medigeshi, G.R., Thomas, G., 2006. A PACS-1, GGA3 and CK2 complex regulates Cl-MPR trafficking. *EMBO Journal* 25 (19), 4423–4435.
- Seaman, M.N., Gautreau, A., Billadeau, D.D., 2013. Retromer-mediated endosomal protein sorting: All WASHed up! *Trends in Cell Biology* 23 (11), 522–528. doi: S0962-8924(13)00083-4 [pii] 10.1016/j.tcb.2013.04.010 [doi].
- Skänland, S.S., Wälchli, S., Utarkerpen, A., Wandinger-Ness, A., Sandvig, K., 2007. Phosphoinositide-regulated retrograde transport of ricin: Crosstalk between hVps34 and sorting nexins. *Traffic* 8 (3), 297–309.
- Smith, D.C., Lord, J.M., Roberts, L.M., Tartour, E., Johannes, L., 2002. 1st class ticket to class I: Protein toxin as pathfinders for antigen presentation. *Traffic* 3, 697–704.
- Snider, M.D., Rogers, O.C., 1985. Intracellular movement of cell surface receptors after endocytosis: Resialylation of asialo-transferrin receptor in human erythroleukemia cells. *Journal of Cell Biology* 100 (3), 826–834.
- Sonnichsen, B., De Renzis, S., Nielsen, E., Rietdorf, J., Zerial, M., 2000. Distinct membrane domains on endosomes in the recycling pathway visualized by multicolor imaging of Rab4, Rab5, and Rab11. *Journal of Cell Biology* 149 (4), 901–914.
- Sorre, B., Callan-Jones, A., Manzi, J., *et al.*, 2012. Nature of curvature coupling of amphiphysin with membranes depends on its bound density. *Proceedings of the National Academy of Sciences of the United States of America* 109 (1), 173–178. doi:1103594108 [pii] 10.1073/pnas.1103594108 [doi].
- Tabuchi, M., Yanatori, I., Kawai, Y., Kishi, F., 2010. Retromer-mediated direct sorting is required for proper endosomal recycling of the mammalian iron transporter DMT1. *Journal of Cell Science* 123 (Pt 5), 756–766. doi:123/5/756 [pii] 10.1242/jcs.060574 [doi].
- Tarrago-Trani, M.T., Storrie, B., 2007. Alternate routes for drug delivery to the cell interior: Pathways to the Golgi apparatus and endoplasmic reticulum. *Advanced Drug Delivery Reviews* 59, 782–797.
- Tartour, E., Benchetrit, F., Haicheur, N., Adotevi, O., Fridman, W.H., 2002. Synthetic and natural non-live vectors: Rationale for their clinical development in cancer vaccine protocols. *Vaccine* 20 (Suppl. 4), A32–A39.

- Ternkin, P., Lauffer, B., Jager, S., *et al.*, 2011. SNX27 mediates retromer tubule entry and endosome-to-plasma membrane trafficking of signalling receptors. *Nature Cell Biology* 13, 715–721. doi:ncb2252 [pii] 10.1038/ncb2252 [doi].
- Van Weering, J.R.T., Sessions, R.B., Traer, C.J., *et al.*, 2012. Molecular basis for SNX-BAR-mediated assembly of distinct endosomal sorting tubules. *EMBO Journal* 31 (23), 4466–4480. doi:10.1038/emboj.2012.283.
- Viel, T., Dransart, E., Henry, E., *et al.*, 2008. *In vivo* tumor targeting by the B-subunit of Shiga toxin. *Molecular Imaging* 7, 239–247.
- Vingert, B., Adotevi, O., Amessou, M., *et al.*, 2006. The Shiga toxin B-subunit targets antigen *in vivo* to dendritic cells and elicits anti-tumor immunity. *European Journal of Immunology* 36, 1124–1135.
- Wilcke, M., Johannes, L., Galli, T., *et al.*, 2000. Rab11 regulates the compartmentalization of early endosomes required for efficient transport from early endosomes to the trans-Golgi network. *Journal of Cell Biology* 151, 1207–1220.
- Yu, J., Chia, J., Canning, C.A., *et al.*, 2014. WLS retrograde transport to the endoplasmic reticulum during Wnt secretion. *Developmental Cell* 29 (3), 277–291. doi:S1534-5807(14)00168-3 [pii] 10.1016/j.devcel.2014.03.016 [doi].
- Zerial, M., McBride, H., 2001. Rab proteins as membrane organizers. *Nature Reviews Molecular Cell Biology* 2, 107–117.
- Zhang, D., Isack, N.R., Glodowski, D.R., *et al.*, 2012. RAB-6.2 and the retromer regulate glutamate receptor recycling through a retrograde pathway. *Journal of Cell Biology* 196 (1), 85–101. doi:jcb.201104141 [pii] 10.1083/jcb.201104141 [doi].

Rabs and Other G Proteins

G Li, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

There are hundreds of proteins that regulate diverse cellular functions via binding and hydrolyzing GTP molecules and are termed GTP-binding proteins (or G proteins for short) or GTPases. By alternating between GTP- and GDP-bound conformations, G proteins serve as molecular switches for regulation of essential cellular functions such as signal transduction, cytoskeleton organization, membrane trafficking, nucleocytoplasmic transport, cytokinesis, protein translation and translocation, etc. The GTP-bound form is biologically active in the sense that it interacts, recruits and/or activates effector proteins to promote the cellular functions. GTP hydrolysis converts the GTP-bound form to inactive GDP-bound form, which can be re-activated by GDP dissociation and GTP binding. This universal GTPase cycle is facilitated by guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs), which accelerate GDP dissociation and GTP hydrolysis in the cell, respectively. Rab proteins represent a large family of such G proteins and function in intracellular membrane trafficking. This overview discusses the structure and function of Rab proteins and several other well-studied G proteins to illustrate how G proteins control cellular processes.

Rab Family of G Proteins and Membrane Trafficking

Rab proteins were discovered in the budding yeast *Saccharomyces cerevisiae* in the forms of Ypt1 and Sec4 involved in different stages of exocytosis (Segev *et al.*, 1988; Salminen and Novick, 1987). It is now clear that Rabs are master regulators of vesicular transport of membranes and proteins in all eukaryotes from the last eukaryotic common ancestor (LECA) to mammals (Diekmann *et al.*, 2011; Elias *et al.*, 2012; Kloppe *et al.*, 2012). Vesicular transport is a fundamental mechanism that governs intracellular membrane trafficking on endocytic and exocytic pathways and involves budding and formation of vesicles from a donor compartment, movement of vesicles on cytoskeleton toward a target compartment, and fusion with the target compartment with the delivery of membrane and luminal cargoes (Sudhof and Rothman, 2009; Guo *et al.*, 2014). Rab proteins regulate each of these vesicular transport steps, via temporal interaction and recruitment of multiple effectors, which include cargo proteins to be packaged into the vesicles, motor proteins that facilitate the movement of vesicles along actin and microtubule cytoskeletons, and tethering factors that dock vesicles to target compartments for membrane fusion.

Human cells contain 66 Rabs, and each Rab is localized to a particular membrane compartment and controls a specific vesicular transport event between a donor compartment and a target compartment (Hutagalung and Novick, 2011; Li and Segev, 2012; Pfeiffer, 2013). For example, Rab5 is localized to

early endosomes and Rab7 to late endosomes on the endocytic pathway, whereas Rab2 is localized to the transport vesicles between the endoplasmic reticulum (ER) and the Golgi on the exocytic pathway (Chavrier *et al.*, 1990; Figure 1). The aforementioned Rabs are ubiquitously expressed in all tissues and control essential housekeeping transport functions, but there are Rabs that are tissue-specific and regulate specialized transport functions in a particular cell type, e.g., Rab3a and Rab27 in the exocytosis of synaptic vesicles and secretory granules in neurons, immune cells and skin cells, respectively (Darchen and Desnos, 2012; Figure 1).

Biosynthesis and Membrane Targeting of Rab Proteins

Rabs are small 20–30 kDa monomeric proteins made in the cytoplasm and contain two Cys residues at or near the C-terminus. Both Cys residues are posttranslationally modified by isoprenylation with the 20-carbon geranylgeranyl group, which provides the hydrophobic lipid anchor for targeting to the membrane. The Rab geranylgeranyl transferase that catalyzes the Rab isoprenylation is a multisubunit enzyme consisting of a Rab escort protein (REP) and a catalytic

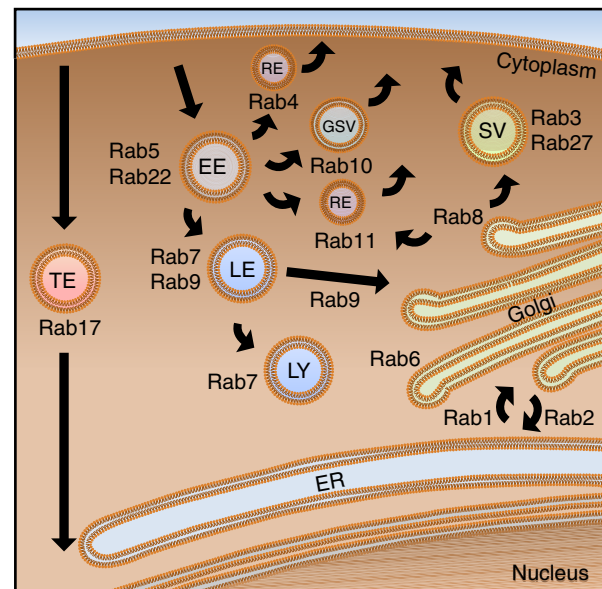


Figure 1 Rabs localizing to distinct organelles in the cell. Some Rabs are ubiquitously expressed in all eukaryotic cells, for example, Rab5, Rab7, Rab1, Rab2, etc. that are localized to endocytic and exocytic compartments, while other Rabs are expressed only in specific cell types, for example, Rab3a and Rab27 that are localized on secretory vesicles in professional exocytotic cells. EE, early endosome; GSV, Glut 4 storage vesicle; LE, late endosome; LY, lysosome; RE, recycling endosome; SV, secretory vesicle; TE, transcytotic endosome.

heterodimer (Leung *et al.*, 2006). The REP recognizes and binds to conserved sequences in Rabs and the catalytic complex transfers the geranylgeranyl group. Some Rabs that end with a CxC motif are further modified by methylation at the C-terminal Cys residue. Once modified, the Rabs are delivered to the membranes by the REP. REP prefers GDP-bound to GTP-bound Rab by three orders of magnitude, indicating that Rabs are prenylated and delivered to the membrane mainly as GDP-bound form. Upon GDP dissociation and GTP loading at the membrane, which is facilitated by a GEF (Blumer *et al.*, 2013; Soldati *et al.*, 1994; Ullrich *et al.*, 1994), the GTP-bound Rab is stably associated with the membrane and can then interact with effectors to promote vesicular transport (Figure 2).

The specificity of membrane targeting by each Rab may be determined by the specific membrane localization of a cognate GEF for the Rab, although the association of GEFs with distinct membrane compartments is not completely understood. In addition, a GDP displacement factor (GDF) is shown to facilitate Rab membrane targeting to endosomes (Sivars *et al.*, 2003; Figure 2). Upon membrane association and GTP loading, the activated GTP-bound Rab then interacts with multiple effectors to facilitate packaging of cargo proteins into transport vesicles, movement of vesicles on actin and microtubule cytoskeleton, and vesicle fusion with the target compartment (Figure 2). The Rab functional cycle is completed by GTP hydrolysis, which is catalyzed by a GAP on the target membrane (Barr and Lambright, 2010), and recycling back to the

donor compartment, which is mediated by the GDP-dissociation inhibitor (GDI) (Soldati *et al.*, 1993; Garrett *et al.*, 1993; Ullrich *et al.*, 1993) (Figure 2). GDI is a protein homologous to REP and dissociates GDP-bound Rab from the target membrane by forming a Rab-GDI complex and protecting the geranylgeranyl groups. The Rab-GDI complex recycles via cytosol back to the donor compartment where a cognate GEF is located for GTP loading and activation of the Rab and a new round of Rab action (Figure 2). Unlike REP, GDI cannot associate with the Rab geranylgeranyl transferase.

Structure and GTPase cycle of Rab Proteins

Like other Ras-like small G proteins, Rab proteins contain a canonical GTP-binding domain consisting of five conserved sequence motifs G1-5, followed by a hypervariable region terminating with the C-terminal Cys motif that is post-translationally geranylgeranylated for membrane association as described above. The GTP-binding domain is folded in such a way that the G1-5 motifs are brought into proximity to form a GTP-binding pocket supported by a six-stranded β sheet and five α helices (Dumas *et al.*, 1999; Zhu *et al.*, 2004). The G1 motif GxxxVGKS/T is in the so-called phosphate-binding loop and forms extensive hydrogen bonds with the β and γ phosphate groups of GTP via backbone amide groups and the side chains of the conserved Lys and Ser/Thr residues. The G2 motif is represented by a highly conserved Thr residue that

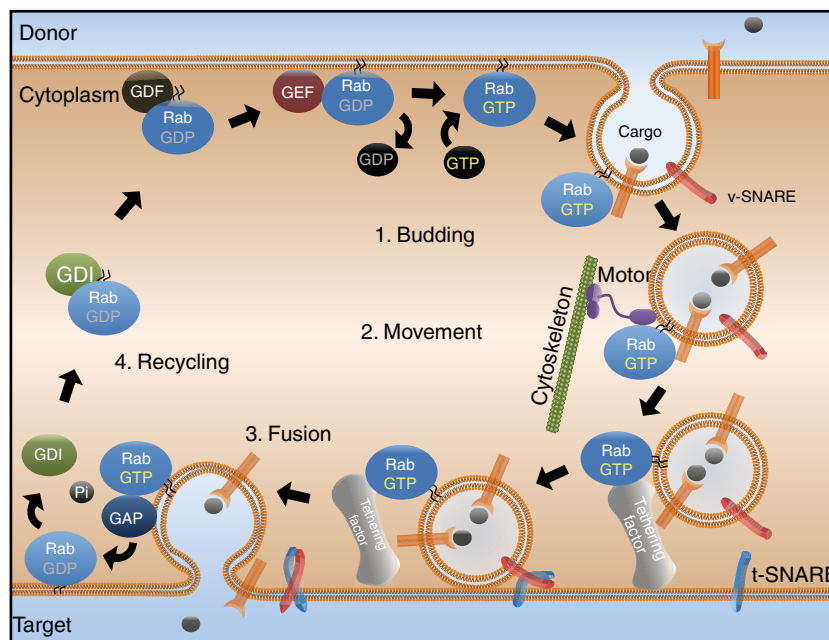


Figure 2 Rab GTPase cycle coupled with membrane association and dissociation. GDP-bound Rabs are complexed with GDI in the cytosol and targeted to the donor membrane compartment. At the membrane, the Rab-GDI complex may initially interact with a GDF to facilitate GDI dissociation or directly interact with a cognate GEF for the Rab to release GDP, reload GTP and insert into the membrane. The active GTP-bound Rab can interact with the cytoplasmic domain of transmembrane cargo proteins to package them into transport vesicles where Rabs may interact with myosin or kinesin motors to facilitate vesicle movement on actin or microtubule cytoskeleton and subsequently interact with tethering factors to dock the vesicle onto the target membrane compartment to facilitate SNARE-mediated membrane fusion, with the delivery of membrane and cargo proteins to the target compartment. Then GTP hydrolysis occurs, accelerated by a cognate Rab GAP on the target membrane, converts active Rab-GTP to inactive Rab-GDP, which is extracted by GDI and recycled back to the donor compartment. Thus through spatial and temporal interactions with multiple effectors, Rabs regulate every aspect of vesicular transport.

coordinates Mg^{2+} for GTP binding and hydrolysis and is in the effector loop (switch I region), which interacts with effectors in a GTP-dependent manner. The G3 motif WDTAGQE is in the switch II loop, interacts with the γ phosphate of GTP and positions the attacking H_2O molecule for GTP hydrolysis. The G4 and G5 motifs (NKxD and ETSAK) interact with the guanine base of GTP and determine the nucleotide-binding specificity, which prefers GTP to ATP by several orders of magnitude in affinity.

The importance of G1-5 motifs in GTP binding and hydrolysis is reflected by well-documented mutations that dramatically affect both processes. The Ser/Thr to Asn mutation in the G1 motif is shown to abolish GTP binding as well as reduce GDP binding. As a result, the mutant is either nucleotide-free or GDP-bound in the cell and is able to sequester cognate GEFs and block endogenous Rab activation. Such a mutant is called a dominant negative Rab mutant and is a useful tool in the study of Rab functions in cells. Likewise, the Asn to Ile mutation in the G4 motif abolishes GTP/GDP binding and the nucleotide-free mutant is inactive and manifests dominant negative phenotype. In contrast, the Gln to Leu mutation in the G3 motif does not change GTP-binding but blocks GTP hydrolysis, and consequently the mutant is locked in active GTP-bound conformation. Such a mutant is called a constitutively active Rab mutant.

The GTP hydrolysis or GTPase cycle allows Rab proteins to alternate between active GTP-bound and inactive GDP-bound conformations. However, Rabs show high affinity for GTP and GDP (Kd in the nanomolar range) and weak intrinsic GTPase activity, which means slow rates in GDP dissociation and GTP hydrolysis. In the cell, both reactions in the GTPase cycle are facilitated by catalyzing protein factors such as GEFs and GAPs, respectively (Barr and Lambright, 2010; Figure 2).

Rab GEFs include diverse groups of proteins that do not share sequence or structural homology but show a general mechanism by displacing the switch I region, disrupting Mg^{2+} coordination, and stabilizing the nucleotide-free form of Rab proteins (Barr and Lambright, 2010). As such, the GEFs facilitate GDP dissociation and GTP loading on Rabs in the cell where GTP concentration is two orders of magnitude higher. Rab GAPs, in contrast, contain a TBC (Tre-2/Bud2/Cdc16) domain for catalysis of GTP hydrolysis (Barr and Lambright, 2010). The TBC domain contains conserved catalytic motifs IxxDxxR and YxQ from which the Arg and Gln side chains insert into the GTP-binding site on the Rab to stabilize the transition state for GTP hydrolysis in a so-called 'dual finger' mechanism (Pan et al., 2006).

The Rab GTPase cycle is coupled to Rab membrane association/dissociation cycle in the sense that GDP-bound Rab is delivered to a donor membrane compartment by GDI and is activated by a cognate GEF localized on the membrane to become GTP-bound. The GTP-bound Rab then interacts with multiple effectors to promote transport vesicle formation, vesicle movement on cytoskeleton, and vesicle fusion with a target membrane compartment where a cognate GAP catalyzes GTP hydrolysis to convert it to inactive GDP-bound conformation (Figure 2). The GDP-bound Rab dissociates from the membrane by GDI and recycles back to the donor compartment for a new round of function. Thus the Rab GTPase cycle is highly regulated by strategically located GEFs and

GAPs. In some cases, the GEF of a downstream Rab happens to be an effector of an upstream Rab on the same trafficking pathway, forming a Rab cascade where the upstream Rab recruits the GEF for activation of the downstream Rab to establish distinct Rab membrane domains. This Rab cascade principle was discovered in the budding yeast *S. cerevisiae* where the GEF of an exocytic Rab (Sec4) is also an effector of the upstream Rab Ypt32, and is recruited by Ypt32-GTP to the Golgi membrane for activation of Sec4 in exocytosis (Ortiz et al., 2002). It is now clear that the GEF-linked Rab cascades also exist along the endocytic pathway and in mammalian systems (Mizuno-Yamasaki et al., 2012; Nottingham and Pfeffer, 2009). In addition, a GAP can also link a Rab cascade where a downstream Rab recruits the GAP for inactivation of an upstream Rab to establish the boundary between the two Rab membrane domains (Mizuno-Yamasaki et al., 2012; Nottingham and Pfeffer, 2009). As such, the GEFs and GAPs effectively promote the transition from an early to a late membrane compartment along a transport pathway, and may possibly generate ultrasensitivity and 'all-or-none' switch-like behavior in the Rab activity (Li and Qian, 2003; Barr, 2013).

Effectors and Functions of Rab Proteins

In the active GTP-bound form on the membrane, Rabs can interact with and recruit multiple effectors to exert a number of functions in vesicular transport, including selection of cargo proteins into vesicles, vesicle movement on actin and microtubule cables, and tethering of vesicles to target compartment for membrane fusion (Figure 2).

Packaging Cargo Proteins into Transport Vesicles

It is well documented that some cargo proteins are Rab effectors and are selectively packaged into transport vesicles via interactions with the Rab proteins (Smythe, 2002). These cargo proteins are usually receptors and other transmembrane proteins. Their cytoplasmic domains contain sequence motifs for binding to the Rabs and such interactions are essential for selective packaging of the cargo proteins into transport vesicles. For example, the α subunit of $\beta 1$ integrins is an effector of Rab5 and Rab21 and is selectively packaged into Rab5- or Rab21-positive endocytic vesicles for recycling and remodeling of the plasma membrane during cell migration and cytokinesis (Pellinen et al., 2006, 2008). In addition, the angiotensin II Type 1A receptor ($AT_{1A}R$) is also a Rab5 effector and the interaction facilitates $AT_{1A}R$ endocytosis (Seachrist et al., 2002). On the exocytic pathway, Rab3b binds directly to polymeric IgA receptor (pIgR) to modulate its transcytosis in epithelial cells (van et al., 2002). In some cases, the Rabs do not directly interact with the cargo proteins but can still package them into transport vesicles via adaptor proteins that are Rab effectors. Such examples include Rab9-mediated recycling of mannose-6-phosphate receptor (M6PR) from late endosomes to *trans*-Golgi network (TGN) via the Rab9 effector TIP47 (Carroll et al., 2001) and Rab5-mediated recruitment of transferrin receptor into endocytic vesicles (McLauchlan et al., 1998).

Vesicle Movement on Actin and Microtubule Cytoskeleton

Another type of Rab effectors is the motor proteins including myosins, kinesins, and dyneins, which are recruited by the Rabs to transport vesicles to promote their movement on actin or microtubule cytoskeleton (Figure 2). These Rabs are usually on exocytic and recycling vesicles and they can bind to the C-terminal cargo-binding globular tail domain (GTD) of class V myosins whose N-terminal actin-binding motor domain may attach to actin cytoskeleton to facilitate vesicle movement toward the plasma membrane (Lindsay *et al.*, 2013). In addition, the Rabs can recruit microtubule-based kinesin and dynein motors directly or indirectly via effectors for movement on microtubule cytoskeleton (Horgan and McCaffrey, 2011). The kinesins include KIF20A and KIF16B that are plus-end microtubule motors and direct the transport vesicles toward cell periphery, whereas dynein is a minus-end microtubule motor and moves the vesicles toward perinuclear region at the center of the cell.

Tethering and Docking Transport Vesicles to Target Membrane Compartment

SNARE-mediated membrane fusion between a transport vesicle and a target compartment completes a vesicular transport event with the delivery of luminal and membrane cargoes to the latter compartment. In the cell, this process is facilitated by active Rab proteins and a group of Rab effectors called tethering factors that tether and dock transport vesicles to target compartments to bring the v-SNAREs on vesicles and the t-SNAREs on target compartments into proximity to form *trans*-SNARE complexes for fusion (Figure 2). These tethering Rab effectors include long coiled-coil homodimers and large multi-subunit complexes (Kummel and Ungermann, 2014), and they are recruited by the Rabs to the membranes to link a vesicle to a target compartment and can also interact with the SM (Sec1-Munc18) proteins to facilitate the assembly of SNARE complexes. As such, Rab proteins are essential for the efficiency of membrane fusion in vesicular transport (Ohya *et al.*, 2009).

Other Rab Effectors

A Rab can interact with multiple effectors, which may be exemplified by the identification of over 20 Rab5 effectors that bind specifically to GTP-bound Rab5 in pull-down assays (Christoforidis *et al.*, 1999). In addition to the effectors that directly participate in vesicular transport, some effectors play a regulatory role. For example, some GEFs and GAPs are also effectors of upstream and downstream Rabs that get recruited to the membrane to establish Rab cascades as described above. Another example is the Rab5 effector Rabaptin-5, which forms a complex with and recruits Rabex-5 (a Rab5 GEF) to the membrane via binding to active Rab5-GTP. As a result, the Rabaptin-5 and Rabex-5 complex can activate more Rab5 molecules on the membrane in a positive feedback loop to recruit additional Rab5 effectors to create and maintain a functional Rab5 membrane domain (Horiuchi *et al.*, 1997; Zhu *et al.*, 2010). Furthermore, the Rab5 effectors APPL1 and APPL2 play an important role in signal transduction by

recruiting Akt and regulating its substrate specificity, which is critical in zebrafish development (Schenck *et al.*, 2008). Recently Rab effectors are also shown to regulate autophagy (Szatmari and Sass, 2014). One such example is the class III PI 3-kinase (Vps34) that is an effector of both Rab5 and Rab7 (Murray *et al.*, 2002; Stein *et al.*, 2003) and is recruited to endosomes to produce PI3P and promote autophagosome formation during autophagy. In addition, an effector of the exocytic Rab1/Ypt1, TRAPP1 complex, is well documented for its role in the initiation of autophagy by promoting the formation of preautophagosomal structure (PAS) (Barrowman *et al.*, 2010; Lipatova and Segev, 2012).

Other G Proteins

In addition to the Rab proteins, there are hundreds of G proteins, large and small, that utilize the universal GTP hydrolysis cycle to control the efficiency, fidelity and directionality of diverse cellular functions via specific intracellular localization and interaction with distinct effectors. Such G proteins are everywhere in the cell and some well-documented examples are described below.

Ras Superfamily of G Proteins

The Rab family represents the largest branch of the Ras superfamily of small G proteins, which consists of Ras proteins, Rho proteins, Arf proteins and Ran protein in addition to the Rab proteins (Cox and Der, 2010). These 20–30 kDa monomeric G proteins share a similar structure, with a conserved GTP-binding domain at the N-terminus followed by a hypervariable domain at the C-terminus. The Ras and Rho proteins, like the Rab proteins, contain a C-terminal Cys motif that is modified by isoprenylation. The Ras proteins are modified by the 15-carbon farnesyl group, whereas the Rho proteins are mostly modified by the 20-carbon geranylgeranyl group. H-Ras and N-Ras proteins but not K-Ras also contain palmitoylation at upstream Cys residues to increase further the hydrophobicity for association with the plasma membrane. The Arf proteins also contain lipid modification, but it occurs at the N-terminal Gly residue by myristoylation. These lipid modifications provide the hydrophobicity necessary for the G proteins to associate with the membranes. In contrast, The Ran protein contains no lipid modification, and it is soluble and plays a role in nucleocytoplasmic transport of proteins and RNAs through nuclear pores.

Ras proteins are localized to the plasma membrane of cells and function as a critical molecular switch in controlling growth factor-mediated signal transduction pathways and cell growth and proliferation (Colicelli, 2004; Cox and Der, 2010). In this case, growth factors such as epidermal growth factor (EGF) bind to their cognate receptors on the cell surface causing conformational changes and activating the tyrosine kinase in the cytoplasmic domain. The resulting phosphorylation of Tyr residues in the receptor cytoplasmic domain leads to recruitment of adaptors and a Ras GEF to activate Ras, i.e., conversion of Ras-GDP to Ras-GTP, and recruitment of Ras effectors for signal transduction. A well-known Ras-mediated signal transduction pathway is the so-called mitogen-activated

protein kinase (MAPK) cascade that ultimately leads to phosphorylation of nuclear transcription factors, activation of target gene expression and cell proliferation. The Ras effector that triggers this phosphorylation cascade is Raf protein, a MAPK kinase kinase (MAPKKK), which phosphorylates and activates MEK protein, a MAPK kinase (MAPKK), which in turn phosphorylates and activates a MAPK that subsequently moves to the nucleus to phosphorylate and activate the transcription factors. In addition, there are other Ras effectors such as PI 3-kinase, RalGDS, RIN, MEKK-1, PKC, Nore and AF6, which may temporally and/or spatially contribute to the regulation of cell proliferation and differentiation. Mutations in the GTP-binding domain of Ras can lead to defects in GTP binding or hydrolysis, and produce dominant negative or constitutive active Ras phenotype. For its important role in regulation of cell growth and proliferation, constitutive active Ras mutations are frequently found in a variety of cancers such as colon cancer, pancreatic cancer, lung cancer, etc. and indeed Ras is the first human oncogene discovered in the early 1980s.

Rho proteins are well known regulators of actin cytoskeleton, including Rho isoforms, Rac proteins, and Cdc42 (Ridley, 2012). Like other small G proteins, they are activated by GEFs and inactivated by GAPs, and their membrane association is regulated by Rho GDIs that keep the proteins in inactive GDP-bound state in cytosol. Once in the active GTP-bound state, Rho, Rac and Cdc42 promote the formation of distinct actin structures in the cell including stress fiber, lamellipodium, and filopodium, respectively. The Rho effectors, ROKs and mDia, may play a role in stress fiber formation by interaction with actin-binding proteins such as myosin and profilin, respectively. Lamellipodia are plasma membrane ruffles formed by *de novo* actin polymerization, and the Rac effector WAVE complex may regulate lamellipodia formation via the Arp2/3 protein complexes. Filopodia are long protrusions of plasma membrane caused by directed actin polymerization and cross-linking via actin-binding proteins, and the Cdc42 effectors, PAKs and WASPs, may regulate filopodia formation. Because of their important role in actin organization, Rho proteins in general are involved in integrin-mediated focal adhesion complex formation for interaction with extracellular matrix and cadherin-mediated cell-cell junctions for intercellular communication. As such, they contribute to cell polarization and migration. In addition, Rho proteins are known for regulation of endocytosis and phagocytosis as well as target gene expression and cell cycle progression.

Arf proteins are essential for recruitment of coat proteins and budding/formation of transport vesicles during vesicular transport (D'Souza-Schorey and Chavrier, 2006; Donaldson and Jackson, 2011; Gillingham and Munro, 2007). In the inactive GDP-bound conformation, the N-terminal myristyl group is buried inside and the protein is cytosolic. Upon activation by membrane-associated GEFs, the GTP-bound form releases the myristyl group and inserts into the membrane. Unlike Rab and Rho proteins, there is no Arf GDI to regulate membrane association and dissociation. The Arf function in coat assembly and vesicle formation may be exemplified by Sar1, a member of the Arf family discovered in yeast. Sar1-GTP on the ER membrane recruits the COPII coat consisting of Sec23p-Sec24p and Sec13p-Sec31p complexes, which induce

membrane curvature to promote vesicle budding off the membrane (Antonny *et al.*, 2001). Arf1 plays a similar role in vesicle formation on the Golgi membrane via recruiting the COPI coat. The biochemical activities of Arf proteins involve stimulation of ADP-ribosylation of the α subunit of heterotrimeric G proteins by cholera toxin (Kahn and Gilman, 1986), which is how Arfs were discovered, and activation of phospholipase D (Brown *et al.*, 1993), which produces phosphatidic acid and activates PI 5-kinase to alter membrane lipid composition in such a way that facilitates vesicle formation. In addition, the Arf family contains a number of Arf-like (Arl) proteins that show Arf-like structure but not the ADP-ribosylation activity (Kahn *et al.*, 2006). Arl proteins may also play an important role in membrane trafficking, especially at the Golgi complex. For example, Arl1 can recruit multiple effectors to the TGN, including GRIP domain-containing proteins, a GGA protein, an ArfGEF and a flippase (Singer-Kruger *et al.*, 2008; Tsai *et al.*, 2013; Gillingham and Munro, 2007), to facilitate the membrane trafficking process. In addition, several Arl proteins (Arl3, Arl6, Arl8) are suggested to be involved in regulation of microtubule-based functions such as cilogenesis (Kahn *et al.*, 2005; Donaldson and Jackson, 2011).

Dynamin

Parallel to Arf-mediated vesicle formation on the ER and Golgi membranes, dynamin is well documented for its role in vesicle formation on the plasma membrane as well as at the TGN (Jones *et al.*, 1998; Kessels *et al.*, 2006; Schmid and Frolov, 2011). Dynamin is a large multi-domain G protein, with a N-terminal GTP-binding domain followed by a α -helical middle domain, a pleckstrin homology (PH) domain, a GTPase effector domain (GED), and a C-terminal proline-arginine domain (PRD). Dynamin molecules can self-assemble into dimers and oligomers via the GED domain, and this self-assembly greatly accelerates dynamin GTPase activity via stabilization of G1 loop and switch regions and is essential for membrane fission at the plasma membrane to produce endocytic vesicles during clathrin-mediated endocytosis. Dynamin binds to the membrane via the lipid-binding PH domain and assemble to induce membrane curvature and invagination, followed by GTP hydrolysis to promote fission and vesicle formation and dissociation from the membrane. Recent studies suggest that dynamin is sufficient to produce vesicles from a lipid bilayer with lipid composition similar to the plasma membrane (Schmid and Frolov, 2011). For its function in membrane fission at the plasma membrane as well as at the TGN, dynamin may control the formation of both endocytic and exocytic vesicles and regulate multiple membrane trafficking pathways, including clathrin-mediated endocytosis, caveolae-mediated endocytosis, phagocytosis, macropinocytosis, and protein secretion (Praefcke and McMahon, 2004; Schmid and Frolov, 2011).

Heterotrimeric G Proteins

Heterotrimeric G proteins are well-known signal transducing proteins and consist of α , β and γ subunits, with α subunit being the actual G protein for GTP binding and hydrolysis (Tesmer, 2010; Hepler and Gilman, 1992). At resting state, the

α subunit is in inactive GDP-bound conformation complexed with β and γ subunits, and this trimeric complex is anchored to the plasma membrane via a number of lipid modifications, including myristoylation and/or palmitoylation of the α subunits and prenylation of the γ subunits. There is a large family of G protein-coupled receptors (GPCRs) with seven transmembrane domains that relay diverse extracellular signals such as hormones and growth factors to intracellular G proteins, which in turn activate various effectors and signal transduction pathways leading to physiological changes of cells. It starts with a ligand binding to a GPCR, and the ligand-bound GPCR functions as a GEF for activation of the G protein by stimulating the exchange of GDP for GTP. The GTP-bound α subunit undergoes conformational changes and dissociates from the GPCR as well as the $\beta\gamma$ complex. Both the α subunit and the $\beta\gamma$ complex can activate downstream effectors and promote signal transduction processes. GTP hydrolysis then converts the α subunit to inactive GDP-bound state, which reunite with the $\beta\gamma$ complex to form the trimeric G protein ready for a new round of GTPase cycle and signal transduction. The GTP hydrolysis reaction is accelerated by RGS (regulators of G-protein signaling) proteins (Berman and Gilman, 1998).

Based on α subunits, G proteins are classified into G_s , G_i , G_q , and G_{12} families that activate different effectors and signal transduction pathways (Neer, 1995). A well-known effector of G_s and G_i proteins is adenylyl cyclase, which is activated by the GTP-bound α subunits of G_s proteins but inhibited by those of G_i proteins. Adenylyl cyclase catalyzes the production of cAMP that is a second messenger for activation of cAMP-dependent protein kinases and induction of physiological changes in the cell. Among members of G_i proteins, the GTP-bound α subunit can activate cGMP phosphodiesterase for phototransduction in retinal rod cells, whereas the $\beta\gamma$ complex can activate a K^+ channel during acetylcholine-mediated signal transduction in cardiac cells. G_q proteins are known to activate phosphoinositide-specific phospholipase C, which produces inositol 1, 4, 5-triphosphate (IP_3) and diacylglycerol (DAG) as second messengers for mobilization of intracellular Ca^{2+} store and activation of protein kinase C, respectively, to induce physiological changes in the cell.

G Proteins in Protein Synthesis and Translocation

Protein synthesis occurs on ribosomes where triplet codons in mRNAs are translated into amino acid sequences in proteins. This is a complex process and involves initiation factors, elongation factors and releasing factors, many of which are G proteins, including bacterial initiation factor IF2, elongation factors Tu (EF-Tu) and EF-G, and releasing factor RF3 and their homologs in eukaryotes. Among them, EF-Tu is one of the best-characterized and forms a complex with aminoacyl-tRNA (aa-tRNA) in its active GTP-bound state to deliver the latter to the A site of the ribosome for addition of amino acid to the growing polypeptide chain (Maracci *et al.*, 2014). GTP hydrolysis by EF-Tu serves as a timer for proof-reading of incoming aa-tRNA so that only correct amino acid (cognate to the codon) is incorporated into the polypeptide chain (Thompson, 1988). In this regard, the GTPase activity of EF-Tu is greatly increased by the cognate aa-tRNA-codon recognition on the ribosome, which

thus functions as a GAP for EF-Tu. The resulting GDP-bound EF-Tu reduces affinity for aa-tRNA and ribosome and dissociates from the complex. EF-Tu is tightly bound to GDP and requires interaction with EF-Ts, a GEF, to release GDP and re-loading of GTP for a new round of delivery of aa-tRNA to the ribosome.

Protein translocation refers to the process of secretory proteins to move from the site of synthesis, i.e., ribosomes, into the lumen of the ER, which involves binding of a hydrophobic signal peptide on the nascent polypeptide to the signal recognition particle (SRP) that in turn binds to the SRP receptor (SR) in the ER membrane to facilitate the co-translational translocation process via the ER translocon (Egea *et al.*, 2005). Both SRP and SR are G proteins and protein translocation depends on their GTP hydrolysis cycle. In GTP-bound state, SRP docks onto SR to form a SRP-SR complex, and crystal structure of this complex suggests conformational changes to form a composite active site with twinned GTP molecules, which are hydrogen bonded to each other and contribute to each other's binding and hydrolysis. GTP hydrolysis triggers disassembly of the SRP-SR complex. In this interesting variation of G protein-mediated function, two G proteins in GTP-bound state form a complex, which is arranged in such a way that each G protein functions as a GAP for the other G protein to promote GTP hydrolysis leading to disassembly of this complex. Thus, the GTP binding and hydrolysis cycle of SRP and SR provide a spatial and temporal control for protein translocation into the ER.

Acknowledgement

The author thanks M. Caleb Marlin for figure illustration, and the NIH/NIGMS for research support (R01 GM074692).

See also: Organelles: Structure and Function: Early Endosomal Compartments; Golgi and TGN; Signaling from Endosomes; The Late Endosome

References

- Antonny, B., Madden, D., Hamamoto, S., Orci, L., Schekman, R., 2001. Dynamics of the COPII coat with GTP and stable analogues. *Nature Cell Biology* 3, 531–537.
- Barr, F., Lambright, D.G., 2010. Rab GEFs and GAPs. *Current Opinion in Cell Biology* 22, 461–470.
- Barr, F.A., 2013. Review series: Rab GTPases and membrane identity: Causal or inconsequential? *Journal of Cell Biology* 202, 191–199.
- Barrowman, J., Bhandari, D., Reinisch, K., Ferro-Novick, S., 2010. TRAPP complexes in membrane traffic: Convergence through a common Rab. *Nature Reviews Molecular Cell Biology* 11, 759–763.
- Berman, D.M., Gilman, A.G., 1998. Mammalian RGS proteins: Barbarians at the gate. *Journal of Biological Chemistry* 273, 1269–1272.
- Blumer, J., Rey, J., Dehmelt, L., *et al.*, 2013. RabGEFs are a major determinant for specific Rab membrane targeting. *Journal of Cell Biology* 200, 287–300.
- Brown, H.A., Gutowski, S., Moomaw, C.R., Slaughter, C., Sternweis, P.C., 1993. ADP-ribosylation factor, a small GTP-dependent regulatory protein, stimulates phospholipase D activity. *Cell* 75, 1137–1144.
- Carroll, K.S., Hanna, J., Simon, I., *et al.*, 2001. Role of Rab9 GTPase in facilitating receptor recruitment by TIP47. *Science* 292, 1373–1376.
- Chavrier, P., Parton, R.G., Hauri, H.P., Simons, K., Zerial, M., 1990. Localization of low molecular weight GTP binding proteins to exocytic and endocytic compartments. *Cell* 62, 317–329.

- Christoforidis, S., McBride, H.M., Burgoyne, R.D., Zerial, M., 1999. The Rab5 effector EEA1 is a core component of endosome docking. *Nature* 397, 621–625.
- Colicelli, J., 2004. Human RAS superfamily proteins and related GTPases. *Sci STKE* 2004, RE13.
- Cox, A.D., Der, C.J., 2010. Ras history: The saga continues. *Small GTPases* 1, 2–27.
- D'Souza-Schorey, C., Chavrier, P., 2006. ARF proteins: Roles in membrane traffic and beyond. *Nature Reviews Molecular Cell Biology* 7, 347–358.
- Darchen, F., Desnos, C., 2012. Masterclass with Rab3 and Rab27: Orchestrating regulated secretion. In: Li, G., Segev, N. (Eds.), *Rab GTPases and Membrane Trafficking*. Sharjah: Bentham Science Publishers, pp. 55–76.
- Diekmann, Y., Seixas, E., Gouw, M., *et al.*, 2011. Thousands of rab GTPases for the cell biologist. *PLoS Computational Biology* 7, e1002217.
- Donaldson, J.G., Jackson, C.L., 2011. ARF family G proteins and their regulators: Roles in membrane transport, development and disease. *Nature Reviews Molecular Cell Biology* 12, 362–375.
- Dumas, J.J., Zhu, Z., Connolly, J.L., Lambright, D.G., 1999. Structural basis of activation and GTP hydrolysis in Rab proteins. *Structure* 7, 413–423.
- Egea, P.F., Stroud, R.M., Walter, P., 2005. Targeting proteins to membranes: Structure of the signal recognition particle. *Current Opinion in Structural Biology* 15, 213–220.
- Elias, M., Brighouse, A., Gabernet-Castello, C., Field, M.C., Dacks, J.B., 2012. Sculpting the endomembrane system in deep time: High resolution phylogenetics of Rab GTPases. *Journal of Cell Science* 125, 2500–2508.
- Garrett, M.D., Kabacell, A.K., Zahner, J.E., *et al.*, 1993. Interaction of Sec4 with GDI proteins from bovine brain, *Drosophila melanogaster* and *Saccharomyces cerevisiae*. Conservation of GDI membrane dissociation activity. *FEBS Letters* 331, 233–238.
- Gillingham, A.K., Munro, S., 2007. The small G proteins of the Arf family and their regulators. *Annual Review of Cell and Developmental Biology* 23, 579–611.
- Guo, Y., Sirkis, D.W., Schekman, R., 2014. Protein sorting at the trans-Golgi network. *Annual Review of Cell and Developmental Biology* 30, 169–206.
- Hepler, J.R., Gilman, A.G., 1992. G proteins. *Trends in Biochemical Sciences* 17, 383–387.
- Horgan, C.P., McCaffrey, M.W., 2011. Rab GTPases and microtubule motors. *Biochemical Society Transactions* 39, 1202–1206.
- Horiuchi, H., Lippe, R., McBride, H.M., *et al.*, 1997. A novel Rab5 GDP/GTP exchange factor complexed to Rabaptin-5 links nucleotide exchange to effector recruitment and function. *Cell* 90, 1149–1159.
- Hutagalung, A.H., Novick, P.J., 2011. Role of Rab GTPases in membrane traffic and cell physiology. *Physiological Reviews* 91, 119–149.
- Jones, S.M., Howell, K.E., Henley, J.R., Cao, H., McNiven, M.A., 1998. Role of dynamin in the formation of transport vesicles from the trans-Golgi network. *Science* 279, 573–577.
- Kahn, R.A., Cherifis, J., Elias, M., *et al.*, 2006. Nomenclature for the human Arf family of GTP-binding proteins: ARF, ARL, and SAR proteins. *Journal of Cell Biology* 172, 645–650.
- Kahn, R.A., Gilman, A.G., 1986. The protein cofactor necessary for ADP-ribosylation of Gs by cholera toxin is itself a GTP binding protein. *Journal of Biological Chemistry* 261, 7906–7911.
- Kahn, R.A., Volpicelli-Daley, L., Bowzard, B., *et al.*, 2005. Arf family GTPases: Roles in membrane traffic and microtubule dynamics. *Biochemical Society Transactions* 33, 1269–1272.
- Kessels, M.M., Dong, J., Leibig, W., Westermann, P., Qualmann, B., 2006. Complexes of syndapin II with dynamin II promote vesicle formation at the trans-Golgi network. *Journal of Cell Science* 119, 1504–1516.
- Klopper, T.H., Kienle, N., Fasshauer, D., Munro, S., 2012. Untangling the evolution of Rab G proteins: Implications of a comprehensive genomic analysis. *BMC Biology* 10, 71.
- Kummel, D., Ungermann, C., 2014. Principles of membrane tethering and fusion in endosome and lysosome biogenesis. *Current Opinion in Cell Biology* 29C, 61–66.
- Leung, K.F., Baron, R., Seabra, M.C., 2006. Thematic review series: Lipid posttranslational modifications. Geranylgeranylation of Rab GTPases. *Journal of Lipid Research* 47, 467–475.
- Li, G., Qian, H., 2003. Sensitivity and specificity amplification in signal transduction. *Cell Biochemistry and Biophysics* 39, 45–59.
- Li, G., Segev, N., 2012. Ypt/Rab GTPases and intracellular membrane trafficking: An overview. In: Li, G., Segev, N. (Eds.), *Rab GTPases and Membrane Trafficking*. Sharjah: Bentham Science Publishers, pp. 3–17.
- Lindsay, A.J., Jollivet, F., Horgan, C.P., *et al.*, 2013. Identification and characterization of multiple novel Rab-myosin Va interactions. *Molecular Biology of the Cell* 24, 3420–3434.
- Lipatova, Z., Segev, N., 2012. A Ypt/Rab GTPase module makes a PAS. *Autophagy* 8, 1271–1272.
- Maracci, C., Peske, F., Dannies, E., Pohl, C., Rodnina, M.V., 2014. Ribosome-induced tuning of GTP hydrolysis by a translational GTPase. *Proceedings of the National Academy of Sciences of the United States of America* 111, 14418–14423.
- McLauchlan, H., Newell, J., Morrice, N., *et al.*, 1998. A novel role for Rab5-GDI in ligand sequestration into clathrin-coated pits. *Current Biology* 8, 34–45.
- Mizuno-Yamasaki, E., Rivera-Molina, F., Novick, P., 2012. GTPase networks in membrane traffic. *Annual Review of Biochemistry* 81, 637–659.
- Murray, J.T., Panaretou, C., Stenmark, H., Miaczynska, M., Backer, J.M., 2002. Role of Rab5 in the recruitment of hVps34/p150 to the early endosome. *Traffic* 3, 416–427.
- Neer, E.J., 1995. Heterotrimeric G proteins: Organizers of transmembrane signals. *Cell* 80, 249–257.
- Nottingham, R.M., Pfeffer, S.R., 2009. Defining the boundaries: Rab GEFs and GAPs. *Proceedings of the National Academy of Sciences of the United States of America* 106, 14185–14186.
- Ohya, T., Miaczynska, M., Coskun, U., *et al.*, 2009. Reconstitution of Rab- and SNARE-dependent membrane fusion by synthetic endosomes. *Nature* 459, 1091–1097.
- Ortiz, D., Medkova, M., Walch-Solimena, C., Novick, P., 2002. Ypt32 recruits the Sec4p guanine nucleotide exchange factor, Sec2p, to secretory vesicles; evidence for a Rab cascade in yeast. *Journal of Cell Biology* 157, 1005–1015.
- Pan, X., Eathiraj, S., Munson, M., Lambright, D.G., 2006. TBC-domain GAPs for Rab GTPases accelerate GTP hydrolysis by a dual-finger mechanism. *Nature* 442, 303–306.
- Pellinen, T., Arjonen, A., Vuoriluoto, K., *et al.*, 2006. Small GTPase Rab21 regulates cell adhesion and controls endosomal traffic of beta1-integrins. *Journal of Cell Biology* 173, 767–780.
- Pellinen, T., Tuomi, S., Arjonen, A., *et al.*, 2008. Integrin trafficking regulated by Rab21 is necessary for cytokinesis. *Developmental Cell* 15, 371–385.
- Pfeffer, S.R., 2013. Rab GTPase regulation of membrane identity. *Current Opinion in Cell Biology* 25, 414–419.
- Praefcke, G.J., McMahon, H.T., 2004. The dynamin superfamily: Universal membrane tubulation and fission molecules? *Nature Reviews Molecular Cell Biology* 5, 133–147.
- Ridley, A.J., 2012. Historical overview of Rho GTPases. *Methods in Molecular Biology* 827, 3–12.
- Salminen, A., Novick, P.J., 1987. A ras-like protein is required for a post-Golgi event in yeast secretion. *Cell* 49, 527–538.
- Schenck, A., Goto-Silva, L., Collinet, C., *et al.*, 2008. The endosomal protein App1 mediates Akt substrate specificity and cell survival in vertebrate development. *Cell* 133, 486–497.
- Schmid, S.L., Frolov, V.A., 2011. Dynamin: Functional design of a membrane fission catalyst. *Annual Review of Cell and Developmental Biology* 27, 79–105.
- Seachrist, J.L., Laporte, S.A., Dale, L.B., *et al.*, 2002. Rab5 association with the angiotensin II type 1 A receptor promotes Rab5 GTP binding and vesicular fusion. *Journal of Biological Chemistry* 277, 679–685.
- Segev, N., Mulholland, J., Botstein, D., 1988. The yeast GTP-binding YPT1 protein and a mammalian counterpart are associated with the secretion machinery. *Cell* 52, 915–924.
- Singer-Kruger, B., Lasic, M., Burger, A.M., *et al.*, 2008. Yeast and human Ysl2p/hMon2 interact with Gga adaptors and mediate their subcellular distribution. *EMBO Journal* 27, 1423–1435.
- Sivars, U., Aivazian, D., Pfeffer, S.R., 2003. Yip3 catalyses the dissociation of endosomal Rab-GDI complexes. *Nature* 425, 856–859.
- Smythe, E., 2002. Direct interactions between rab GTPases and cargo. *Molecular Cell* 9, 205–206.
- Soldati, T., Riederer, M.A., Pfeffer, S.R., 1993. Rab GDI: A solubilizing and recycling factor for rab9 protein. *Molecular Biology of the Cell* 4, 425–434.
- Soldati, T., Shapiro, A.D., Sveistrup, A.B., Pfeffer, S.R., 1994. Membrane targeting of the small GTPase Rab9 is accompanied by nucleotide exchange. *Nature* 369, 76–78.
- Stein, M.P., Feng, Y., Cooper, K.L., Welford, A.M., Wandinger-Ness, A., 2003. Human VPS34 and p150 are Rab7 interacting partners. *Traffic* 4, 754–771.
- Sudhof, T.C., Rothman, J.E., 2009. Membrane fusion: Grappling with SNARE and SM proteins. *Science* 323, 474–477.
- Szatmari, Z., Sass, M., 2014. The autophagic roles of Rab small GTPases and their upstream regulators: A review. *Autophagy* 10, 1154–1166.
- Tesmer, J.J., 2010. The quest to understand heterotrimeric G protein signaling. *Nature Structural & Molecular Biology* 17, 650–652.
- Thompson, R.C., 1988. EFTu provides an internal kinetic standard for translational accuracy. *Trends in Biochemical Sciences* 13, 91–93.

- Tsai, P.C., Hsu, J.W., Liu, Y.W., Chen, K.Y., Lee, F.J., 2013. Arl1p regulates spatial membrane organization at the trans-Golgi network through interaction with Arf-GEF Gea2p and flippase Drs2p. *Proceedings of the National Academy of Sciences of the United States of America* 110, E668–E677.
- Ullrich, O., Horiuchi, H., Bucci, C., Zerial, M., 1994. Membrane association of Rab5 mediated by GDP-dissociation inhibitor and accompanied by GDP/GTP exchange. *Nature* 368, 157–160.
- Ullrich, O., Stenmark, H., Alexandrov, K., *et al.*, 1993. Rab GDP dissociation inhibitor as a general regulator for the membrane association of rab proteins. *Journal of Biological Chemistry* 268, 18143–18150.
- van, I.S.C., Tuvim, M.J., Weimbs, T., Dickey, B.F., Mostov, K.E., 2002. Direct interaction between Rab3b and the polymeric immunoglobulin receptor controls ligand-stimulated transcytosis in epithelial cells. *Developmental Cell* 2, 219–228.
- Zhu, G., Zhai, P., Liu, J., *et al.*, 2004. Structural basis of Rab5-Rabaptin5 interaction in endocytosis. *Nature Structural & Molecular Biology* 11, 975–983.
- Zhu, H., Qian, H., Li, G., 2010. Delayed onset of positive feedback activation of Rab5 by Rabex-5 and Rabaptin-5 in endocytosis. *PLoS One* 5, e9226.

Further Reading

- Hall, A. (Ed.), 2000. *GTPases*. Oxford: Oxford University Press.
- Kelly, E.E., Horgan, C.P., Goud, B., McCaffrey, M.W., 2012. The Rab family of proteins: 25 years on. *Biochemical Society Transactions* 40, 1337–1347.
- Li, G., Segev, N. (Eds.), 2012. *Rab GTPases and Membrane Trafficking*. Bentham Science Publishers.
- Mizuno-Yamasaki, E., Rivera-Molina, F., Novick, P., 2012. GTPase networks in membrane traffic. *Annual Review of Biochemistry* 81, 637–659.

INTERORGANELLAR COMMUNICATION: COMPONENTS

Contents

Adaptor Proteins: Inter-Organelle Traffic Controllers

SNAREs: Membrane Fusion and Beyond

ESCRTing around the Cell

The Retromer Complex

Vesicle Tethers

BAR Domains and BAR Domain Superfamily Proteins

Adaptor Proteins: Inter-Organelle Traffic Controllers

K Madhivanan, W-C Hsieh, and RC Aguilar, Purdue University Center for Cancer Research, West Lafayette, IN, USA

© 2016 Elsevier Inc. All rights reserved.

Adaptor Proteins

Adaptors were originally defined as proteins able to link transmembrane cargo proteins to the membrane coat. This definition was implicitly expanded to also require binding to phospholipids and other elements of the trafficking machinery to incorporate cargo into and to facilitate the assembly of transport carriers. Whereas some adaptors have the ability to bind membrane, cargo, and coat proteins, others indirectly fulfill these requirements by interacting with bridging elements and consolidating the protein network (Table 1).

Heterotetrameric Adaptor Complexes

Assembly polypeptides (APs) or adaptins are a family of protein complexes involved in vesicle trafficking with five members (AP1–5) found in most eukaryotes (Bonifacino, 2004; Hirst *et al.*, 2013). APs are heterotetramers composed of two large subunits (~100 kDa, $\beta 1/\gamma$, $\beta 2/\alpha$, $\beta 3/\delta$, $\beta 4/\epsilon$, $\beta 5/\zeta$ in AP1–5 respectively), one medium (~50 kDa, $\mu 1$ –5), and one small subunit (~20 kDa, $\sigma 1$ –5) (Figure 1(a)). In addition, AP1–3 exhibit cell type-specific isoforms. While vertebrates and plants have all the 5 members of this family, only AP1–3 are present in flies, worms, and yeast (Hirst *et al.*, 2013). Surprisingly, while in multicellular organisms several APs are required for normal embryonic development, they are dispensable for survival and clathrin-mediated endocytosis (CME) in yeast (Mitsunari *et al.*, 2005; Huang *et al.*, 1999).

The large subunits consist of an N-terminal trunk domain, followed by an unstructured hinge domain and a C-terminal appendage domain (Collins *et al.*, 2002; Figure 1(a)). The N-terminal trunks together with the μ and σ subunits form the core of the AP complex. While the trunk domains of α , γ , δ , ϵ ,

and ζ subunits are involved in binding to target membranes, the μ and σ subunits are involved in cargo recognition. AP recruitment to membranes is facilitated by ADP-ribosylation factor (ARF) family of proteins in their membrane-bound, GTP-loaded state. While, Arf-6 primarily fulfils this function at the plasma membrane, Arf-1 does it at the *trans*-Golgi network (TGN) (Ren *et al.*, 2013). The hinge domains of the β subunits in AP1–3 have one or more L Φ X Φ [D/E] (L represents Leucine, X represents any amino acid, and Φ represents bulky hydrophobic amino acids, such as leucine, isoleucine, methionine, valine, and phenylalanine) motifs responsible for binding to clathrin (Owen *et al.*, 2004). However, $\beta 4$ and $\beta 5$ lack this sequence and thus AP4 and AP5 complexes were suggested to mediate clathrin-independent vesicle trafficking (Hirst *et al.*, 2013). The appendage domains serve as a platform for interacting with other adaptor and accessory proteins (Figure 1(b); Owen *et al.*, 2004). While the N-terminal domain of μ subunits contributes to the core, the C-terminal domain recognizes tyrosine-based sorting signals within cargoes.

AP complexes mediate the sorting of various transmembrane proteins, including type I, type II, and multi-span proteins. To recruit cargoes to vesicles, APs recognize short sorting signal sequences that reside in the cytosolic region, usually at 4–10 residues from the transmembrane domain. There are two well-characterized AP-targeting sorting signal sequences: tyrosine-based and leucine-based signals (Nakatsu and Ohno, 2003; Traub, 2009).

The tyrosine signal consists of four amino acids fitting the YXX Φ consensus, where Y stand for tyrosine. The medium μ subunits of AP complexes have been identified to play a central role in YXX Φ sequence recognition. The leucine-based signals recognized by APs fit into the [D/E]XXXL[L/ Φ] consensus, where the -4 position from the first leucine is either aspartic acid or glutamic acid, then followed by three non-conserved amino acids. This sequence motif precedes two

Table 1 Adaptor properties

Adaptor name	Clathrin (CBM)	AP2 (ABM)	Lipid – Adaptor binding domain	Example of cargo (signal) – Adaptor interacting domain/subunit	Other domains – binding partners	Localization
Classical heterotetrameric adaptor proteins recognizing YxxΦ and diLL signals						
AP1	+		PI4P – $\gamma, \beta 1$	CI-M6PR (YXXΦ)- $\mu 1$, (diLL)-[$\gamma/\sigma 1$]		TGN
AP2	+		PI(4,5)P ₂ , PI(3,4,5)P ₃ – $\alpha, \mu 2$	TfR (YXXΦ)- $\mu 2$, CD4 (diLL)-[$\alpha/\sigma 2$]		Plasma membrane
AP3	+		?	Lamp-1(YXXΦ)- $\mu 3$, Tyrosinase (diLL)-[$\delta/\sigma 3$]		Endosome
AP4	–		?	APP (YKFFE)- $\mu 4$		TGN
AP5	–		?	?		Late endosome
Adaptors for cargo carrying acidic DXXLL peptide						
GGA1,2,3	+		PI4P – GAT	M6PR(DXXLL)-VHS	GAT – Ub	TGN
Adaptors for ubiquitinated-cargo						
Epsin	+	+	PI(4,5)P ₂ – ENTH	VEGFR2 – UIM	NPF – Eps15	Plasma membrane
Eps15	–	+	PI(3,5)P ₂ – EH	EGFR – UIM	ESFGDGFADFSTLS – AP1	Plasma membrane
Eps15R	+		PI4P – ENTH	c-Met receptor – coiled coil Vtib-?	DFxD[F/W] – AP1	TGN
Adaptors for cargo carrying NPXY peptide						
Dab1	–	+	PI(4,5)P ₂ – PTB	ApoER2 (NPXY) – PTB		Plasma membrane
Dab2	+	+	PI(4,5)P ₂ – PTB	LDLR (NPXY) – PTB	NPF – Eps15	Plasma membrane
ARH	+	+	PI(4,5)P ₂ – PTB	LDLR (NPXY) – PTB		Plasma membrane
Numb	+	+	PI(4,5)P ₂ – PTB	APP (NPXY) – PTB	NPF – Eps15	Plasma membrane
Adaptors for GPCRs (Ser and Thr clusters)						
β -Arr	+	+	PI(4,5)P ₂	GPCRs – Arrestin fold		Plasma membrane
Adaptors for VAMPs						
AP180	+	+	PI(4,5)P ₂ – ANTH	Synaptobrevin2 – ANTH		Plasma membrane
CALM	+	+	PI(4,5)P ₂ – ANTH	Synaptobrevin2 – ANTH	NPF – Eps15	Plasma membrane

Notes: +, Motif present; –, Motif absent; ?, Not known; Φ, Hydrophobic amino acid; x, Any amino acid, single letter abbreviations of amino acids are used in the table above. Abbreviations: ANTH, AP180 N-terminal homology domain; ApoER2, Apolipoprotein E receptor 2; APP, Amyloid β precursor protein; CI-M6PR, Cation-independent mannose phosphate receptor; EGFR, Epidermal growth factor receptor; EH, Eps15 Homology; ENTH, Epsin N-terminal homology domain; GAT, GGA and Tom; GPCR, G-Protein coupled receptor; LDLR, Low-density lipoprotein 2; PIP, Phosphatidylinositol phosphate; PTB, Phosphotyrosine binding domain; TfR, Transferrin receptor; TGN, Trans-Golgi network; Ub, ubiquitin; VEGFR, Vascular endothelial growth factor; VHS – Vps27, Hrs and STAM.

leucines that lead to its 'dileucine signal' denomination. The first leucine cannot be replaced by any other amino acid, while the second can be substituted with a hydrophobic residue. In some proteins, the acidic residue at -4 position in the consensus can be replaced by a serine that upon phosphorylation acquires a negative charge and becomes functionally equivalent to the D/E residues. The dileucine signal is recognized by γ - $\sigma 1$, α - $\sigma 2$, and δ - $\sigma 3$ hemicomplexes of AP1–3 adaptors (Mattera *et al.*, 2011).

AP1

This adaptor is expressed ubiquitously and is essential during mammalian development (Meyer *et al.*, 2000). The AP1

gamma subunit interacts with the Golgi apparatus-enriched lipid phosphatidylinositol-4-phosphate (PI(4)P) and it is involved in cargo transport between the TGN and endosomal/lysosomal compartments (Wang *et al.*, 2003; Lefkir *et al.*, 2003). The localization of AP1 is also regulated by phosphorylation of its subunits, namely AP1 associates with membranes when the $\mu 1$ subunit is phosphorylated, whereas it detaches when $\beta 1$ is phosphorylated (Ghosh and Kornfeld, 2003). The cargoes of AP1 include furin, mannose-6-phosphate receptors (M6PR), and other proteins of the secretory pathway that are processed at the Golgi apparatus (Bonnemaizon *et al.*, 2013; Table 1). It has also been suggested that AP1 and Golgi-localized gamma-adaptin ear-containing Arf-binding proteins (GGAs) cooperate in cargo recruitment to clathrin-coated vesicles (CCVs) by recognition of sorting motifs on the

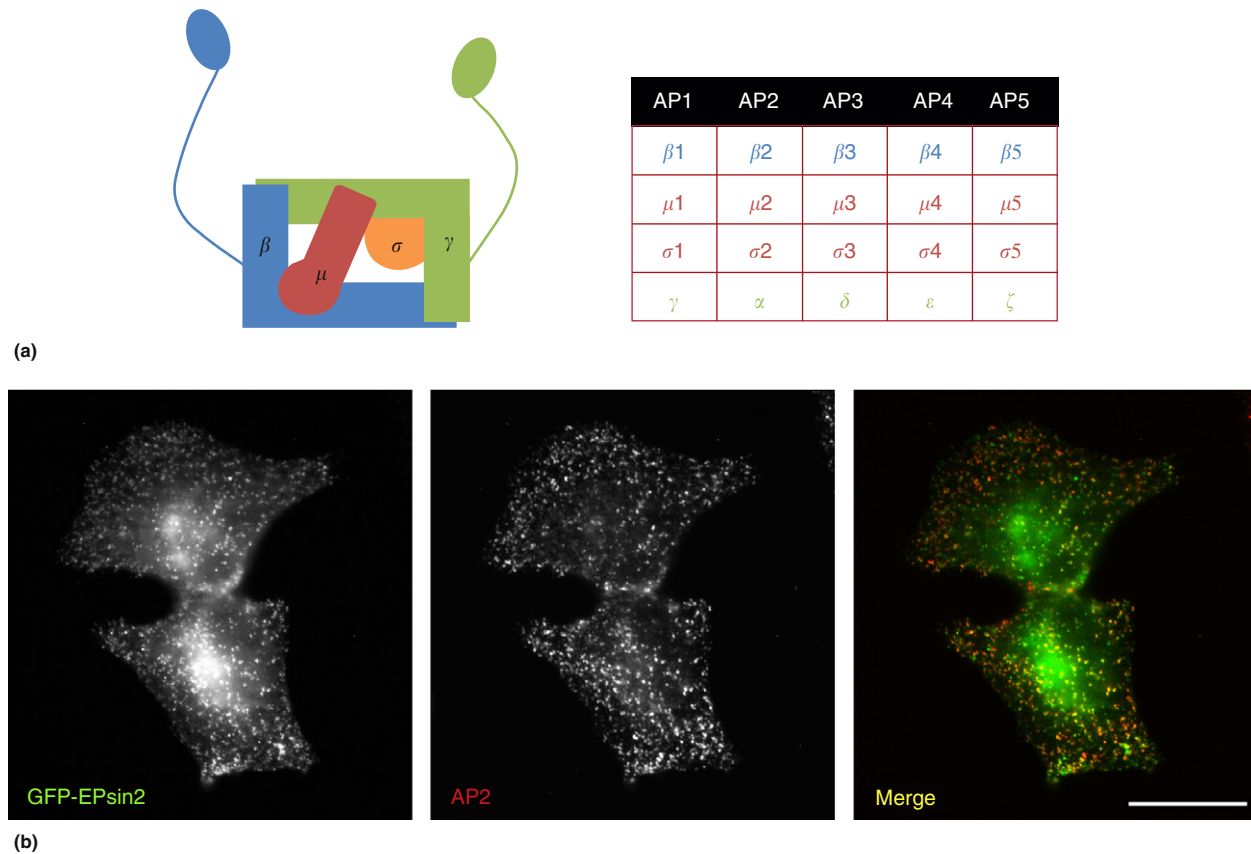


Figure 1 Heterotetrameric adaptor complexes. (a). Left panel: Cartoon representing the subunit organization of the AP1 complex as example of a heterotetrameric adaptin. Right panel: Subunit composition of known AP complexes. (b) Colocalization of AP2 with the monomeric adaptor Epsin2. HeLa cells expressing GFP-Epsin2 (green) were fixed and the presence of AP2 was revealed by indirect immunofluorescence using an anti-AP2 antibody (red). Scale bar: 20 μ m.

cytoplasmic tail of the cargoes (Bonnemaïson *et al.*, 2013). In fact, AP1 is believed to be a central player in CCV formation at the TGN affecting a wide variety of transmembrane proteins (Hirst *et al.*, 2012). Additionally, AP1 has also been linked to retrograde transport wherein several receptors were enriched in peripheral compartments upon AP1 depletion (Hirst *et al.*, 2012). The $\mu 1$ B subunit isoform specifically expresses in epithelial cells forming the AP1B complex which is involved in basolateral sorting of cargo such as the low-density lipoprotein receptor (LDLR) (Gonzalez and Rodriguez-Boulant, 2009). Mutations in AP1 have been linked to X-linked mental retardation (Tarpey *et al.*, 2006), MEDNIK syndrome (Montpetit *et al.*, 2008), and Pettigrew syndrome (Cacciagli *et al.*, 2014).

AP2

This adaptor is expressed ubiquitously and is essential for mammalian development (Mitsunari *et al.*, 2005). AP2 is capable of binding the plasma membrane-enriched lipids PI (4,5) P_2 and/or PI(3,4,5) P_3 and is a key element for CME. Upon recruitment to the plasma membrane, AP2 undergoes a conformational change from a 'lock' state to an 'open' state that exposes the clathrin-binding motif on the $\beta 2$ hinge (Collins *et al.*, 2002). The interaction of AP2 with cargo

stabilizes the open state, and thus facilitating formation and cargo inclusion in clathrin-coated pits (CCPs) (Kelly *et al.*, 2014). The α and β appendage domains interact with other clathrin adaptors and accessory proteins, such as Epsin, ARH, and AP180 to consolidate CCVs. The initial presumption that AP2 was essential for CME was challenged by the observation that although depletion of AP2 reduced CCV formation and specifically blocked transferrin-receptor internalization, epidermal growth factor receptor (EGFR), and LDLR internalization still progressed via CME (Motley *et al.*, 2003).

AP3

Organisms deficient in this adaptor survive to adulthood; however, they have defects in lysosome and lysosomal-related organelle biosynthesis (Yang *et al.*, 2000). AP3 membrane localization is similar to AP1, but with different degree of colocalization with clathrin, suggesting localization/functional differences (Peden *et al.*, 2004). Selective depletion of AP3 subunits resulted in indirect trafficking of lysosomal membrane proteins namely Lamp-1, Lamp-2, and CD63 via a plasma membrane-to-endosome pathway (Dell'Angelica *et al.*, 1999). The $\beta 3$ and $\mu 3$ subunits of the AP3 complex in metazoans have two isoforms: the isoform A is ubiquitously present, while the

isoform B is specifically expressed in neuronal and neuroendocrine tissues resulting in two AP3 complexes. The ubiquitous AP3A sorts its cargo (Lamp-1, Lamp-2, CD63) to the lysosome, while neuronal AP3B sorts its cargo (Chloride Channel-3, Vesicular glutamate transporter 1, vesicular GABA transporter, and many more synaptic proteins) to synaptic vesicles (Nakatsu and Ohno, 2003). Mutations in the ubiquitous AP3A complex results in Hermansky-Pudlak syndrome type-2, while mutations in the neuronal AP3B complex results in neurological symptoms such as epilepsy and hyperactivity (Badolato and Parolini, 2007; Nakatsu and Ohno, 2003).

AP4

This adaptor is also ubiquitously expressed in all tissues, but is not essential for development as evidenced by the existence of healthy knockout mice (Matsuda *et al.*, 2008). AP4 localizes to the TGN, but does not colocalize with AP1 perhaps due to the lack of CBMs (Hirst *et al.*, 2013). Notably, perturbation of CCV-assembly increased AP4 vesicle formation without changing the expression levels of AP4, suggesting a dynamic compensatory mechanism for TGN trafficking (Hirst *et al.*, 2012). Hence, AP4 is speculated to traffic at least some cargo also transported by CCV. Surprisingly, AP4 recognizes various types of sorting motifs on its cargoes, including canonical tyrosine motif on LDLR, YX[F/Y/L][F/L]E motifs on amyloid precursor protein (APP), FI (Phenylalanine-Isoleucine) motif on Furin, and unconventional basolateral sorting signal (EQFPHLAFWQDLG- (Bresciani *et al.*, 1997)) on M6PR (Simmen *et al.*, 2002). AP4 complex was required for the basolateral trafficking of these cargoes. Additionally, AP4 is also involved in basolateral trafficking in neurons where its cargoes are LDLR, AMPA (α -Amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid), and δ 2 glutamate receptors (Matsuda *et al.*, 2008). AP4-deletion showed aberrant accumulation of APP in TGN rather than in the endosomes implying a role for the AP4 complex in TGN-to-endosomes trafficking (Burgos *et al.*, 2010). Deficiency of AP4 in humans has been shown to cause intellectual disability and progressive spasticity that gradually develops to hereditary spastic paraplegia (HSP), suggesting that AP4 participates in human brain development (Hirst *et al.*, 2013).

AP5

This adaptor is the most recently identified, ubiquitous, heterotetrameric adaptor protein complex. Despite having less than 10% protein sequence identity with other APs, the secondary structure prediction of AP5 subunits is similar to the counterparts in other APs (Hirst *et al.*, 2013). AP5 does not associate with clathrin and was shown to localize on late endosomal compartments. Surprisingly, RNAi knock-down of AP5 resulted in cation-independent M6PR trapping in swollen early endosomes, albeit AP5 localization in late endosomes (Hirst *et al.*, 2013). AP5 endosomal trafficking route and cargoes are yet to be described. Mutations in AP5 ζ subunits, similar to AP4 subunits, lead to HSP. It is possible that the same cargo is transported by AP4 and AP5 at different points along the vesicle trafficking pathway.

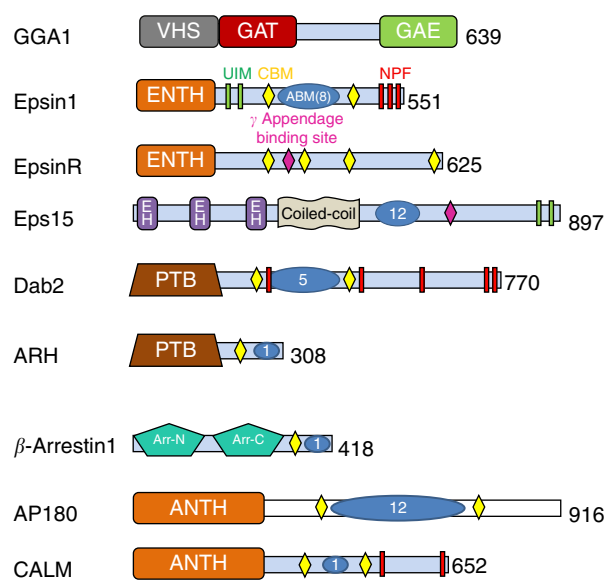


Figure 2 Monomeric adaptors. Domain/motif organization of several monomeric adaptor complexes is shown. ABM, AP2-Binding Motif; ANTH, AP180 N-terminal Homology domain; Arr, Arrestin fold; CBM, Clathrin-binding motif; EH, Eps15 homology domain; ENTH, Epsin N-terminal Homology domain; GAE, Gamma-adaptin ear homology domain; GAT, GGA and Tom; NPF, Asparagine (N)-Proline (P)-Phenylalanine (F) peptide; PTB, Phosphotyrosine binding domain; UIM: Ubiquitin interacting motif; VHS, Vps27, Hrs, and STAM.

GGAs

GGAs were identified in early 2000s, and as their name indicates, they are localized at the Golgi apparatus (TGN), but they can also be found on endosomes (Ghosh and Kornfeld, 2004). GGAs are ubiquitously expressed and evolutionarily conserved, with three GGAs in human and mouse, two in yeast, and one in nematode and fruit fly. The existence of distinct GGA functions of the three GGAs has been demonstrated in mouse models. GGA2 null mice are embryonic or neonatal lethal; double mutant of GGA1 and GGA3 result in neonatal lethal, indicating that GGAs are required for developmental process and not fully redundant (Govero *et al.*, 2012). There are four folding modules in GGAs, from N-terminal to C-terminal: VHS (Vps27, Hrs, and STAM), GAT (GGA and Tom), hinge, and GAE (gamma-adaptin ear homology) domain (Figure 2).

The GAT domain has an N-terminal α -helical hook sub-domain which interacts with Arf-1 and consequently is recruited to the TGN (Collins *et al.*, 2003; Shiba *et al.*, 2003). Like the hinge region of β subunits in AP1–3, the hinge region of GGAs also contains an L Φ X Φ [D/E] CBM, that allows the GGAs to bind clathrin (Bonifacino, 2004). This indicated that GGAs are involved in CM-trafficking from the TGN. The C-terminus of the GAT domain binds ubiquitin and has been implicated in sorting of ubiquitinated-(Ub)-cargo from TGN to the vacuole in yeast (Prag *et al.*, 2005). The role of GGAs in trafficking Ub-cargo in mammals has not been fully explored.

GGAs also recognize cargo via the VHS domain which binds a DXXLL motif located in the cytosolic tail of the transported proteins, called acidic-cluster-dileucine signal. The motif consists of an aspartate residue followed by two leucines

at the third downstream position (Puertollano *et al.*, 2001; Doray *et al.*, 2002). The D and dileucine bind to the sixth to eighth α -helices of VHS where an electropositive pocket and two hydrophobic pockets reside (Misra *et al.*, 2000; Shiba *et al.*, 2002). In addition, acidic-cluster-dileucine signals also have been found in the hinge region of human GGA1 and GGA3, suggesting the formation of intramolecular complexes competing with cargo recognition; i.e., with auto-inhibitory properties (Bonifacino, 2004). The cargoes trafficked by GGAs include M6PRs and Sortilin (Puertollano *et al.*, 2001; Hirst *et al.*, 2012). Notably GGAs and AP1 cooperate for the trafficking of M6PR from the TGN to the lysosome. The GAE domain homologous to the γ -adaptin ear of AP1, also serves as a platform for binding to other endocytic accessory proteins with a DFGX Φ motif, namely γ -synergins, p56, Rabaptin 5 and Clint/EpsinR in mammals (Bonifacino, 2004). In addition, evidence showed the involvement of GGA1 in the retrograde trafficking of β -site APP-cleaving enzyme 1 from endosomes to the TGN while GGA3 plays a role in the targeting of BACE1 to the lysosomes (Tan and Evin, 2012).

The Epsin and Eps15 Families

The Epsin Family

These endocytic adaptors including Epsin and EpsinR are highly conserved in yeast and most vertebrates (Sen *et al.*, 2012). There are three Epsin paralogs in humans (Rosenthal *et al.*, 1999; Chen *et al.*, 1999; Spradling *et al.*, 2001). While Epsin-1 and 2 are ubiquitously expressed, Epsin3 has a more restricted expression pattern being confined to migratory keratinocytes, stomach cells, and cancer cells (Rosenthal *et al.*, 1999; Chen *et al.*, 1999; Coon *et al.*, 2011; Spradling *et al.*, 2001).

The N-terminus of Epsins has a highly conserved Epsin N-terminal homology domain (ENTH) (Figure 2) that binds PI(4,5)P₂, a lipid enriched at endocytic and signaling sites (Sen *et al.*, 2012). This interaction triggers a conformational change in the ENTH domain resulting in the formation of an additional α -helix, helix 0 which inserts into the plasma membrane inducing membrane curvature (Ford *et al.*, 2002). This structural transformation of the ENTH domain is believed to constitute the initial step of membrane invagination required for the formation of CCPs (Horvath *et al.*, 2007; Legendre-Guillemin *et al.*, 2004).

C-terminal to the ENTH domain is an unstructured region containing multiple motifs required for engaging cargo and endocytic machinery. Each Epsin has two or three UIMs (ubiquitin interacting motifs) able to bind ubiquitinated-cargo (Madshus, 2006). Epsins also bind clathrin via 1 (e.g., yeast) or 2 (e.g., mammalian) clathrin-binding motifs (CBMs) with the consensus L Φ Z Φ Z, where L is leucine, Φ is any hydrophobic residue, and Z is any polar residue. In higher eukaryotes, the CBMs flank 3–8 DP[W/F] (aspartate, proline, and tryptophan/phenylalanine) repeats that bind to AP2 (Figure 1B). Lastly, the C-terminus houses 2–3 NPFs (asparagine, proline, phenylalanine) which bind to Eps15 homology (EH) domain containing proteins, namely Eps15 and REPS1/2. In fact, Eps15 and Epsin together have been implicated in EGFR internalization (Sigismund *et al.*, 2005).

EGFR is a receptor tyrosine kinase belonging to the ErbB family which plays a major role in proliferation and differentiation. Interestingly, two different routes have been proposed for Epsin-bound activated EGFR. Classical route entails EGFR ubiquitination and recruitment to CCPs for internalization (Bertelsen *et al.*, 2011; Hawryluk *et al.*, 2006). However, at high EGF concentrations, a clathrin-independent, Eps15- and Epsin-dependent internalization was also found (Sigismund *et al.*, 2005). Interestingly, it has been reported that Epsin binding to ubiquitin negatively regulates clathrin binding (Chen and De Camilli, 2005), supporting the existence of an Epsin-mediated clathrin-independent route. Notably, the fate of EGFR differs based on the internalization route, wherein the clathrin-dependent route and the clathrin-independent route favor recycling and degradation of the EGFR receptor, respectively (Sigismund *et al.*, 2005).

While Epsin can contribute to EGFR signaling termination, it is required for activation of the notch signaling pathway (Wang and Struhl, 2004). Elegant studies done in flies to study the juxtacrine notch signaling pathway has shown that Epsin-dependent Delta endocytosis in the signal-sending cell is required for signaling (Wang and Struhl, 2004). Supporting evidence comes from Epsin-deficient mutants in flies, worms, and mice where defects in notch signaling pathway such as cardiovascular development, germline, and neural tube differentiation were observed (Tian *et al.*, 2004; Chen *et al.*, 2009).

Other Epsin-specific cargoes that undergo UIM-mediated internalization are the epithelial sodium channel ENaC (Wang *et al.*, 2006) and the Vascular Endothelial Growth Factor Receptor-2 (Tessneer *et al.*, 2013).

Progression from an initiated pit into a vesicle depends on reaching a critical mass of cargo and endocytic components. Epsin-1 contributes to this checkpoint which when unsatisfied will result in the destabilization and abortion of pits (Sen *et al.*, 2012). In addition to serving as an endocytic adaptor protein, Epsins are proposed to function as stabilizers of the endocytic network. Yeast Epsins have also shown to play an analogous role in expanding and sustaining the endocytic network by mediating protein–protein interaction via their NPFs and UIMs (Dores *et al.*, 2010).

The Epsin family of adaptor proteins has been recognized as regulators of cancer progression (Tessneer *et al.*, 2013). Epsins are upregulated in skin, breast, prostate, pancreatic, and lung cancers (Tessneer *et al.*, 2013; Sen *et al.*, 2012). The enhanced invasion potential of Epsin-overexpressing cells is attributed to the nonclassical Epsin function in RhoGTPase signaling regulation that leads to enhanced cell migration and invasion (Coon *et al.*, 2010). In addition, Epsin's role in angiogenesis by regulating VEGFR2 levels has been demonstrated, and considered as a potential treatment of cancer. Knockout of Epsin in Lewis Lung Cancer model effectively reduced tumorigenesis by misregulation of VEGFR2 and angiogenesis (Tessneer *et al.*, 2013).

The EpsinR Family

These adaptors bear an ENTH domain similar to Epsin (Figure 2), but differs in lipid specificity as it binds PI(4)P and

therefore, localizes to the TGN. EpsinR (epsin-related protein) has an AP1-binding motif [D/E]FxD[F/W] and four identified CBM of the consensus DΦF. Some of the cargo proteins under control of EpsinR include M6PR and the SNARE protein Vtib (Hirst *et al.*, 2004; Mills *et al.*, 2003). This adaptor is believed to play a role in cargo recycling to the TGN.

Eps15 Family

The presence of UIMs in Eps15 (Figure 2) indicates that, similar to Epsin, this protein is capable of ubiquitinated-cargo recognition (although lacks coat-binding determinants). Eps15 contains NPF-binding, EH domains in the N-terminus which also interacts with PI(3,5)P₂ (Naslavsky *et al.*, 2007). The Eps15 coiled-coiled region following the EH domain is involved in protein dimerization (van Bergen En Henegouwen, 2009) and along with the UIMs have been shown to define distinct cargo-specificity (c-MET and EGFR) (Parachoniak and Park, 2009; Sigismund *et al.*, 2005; De Melker *et al.*, 2004; Polo *et al.*, 2002). The C-terminus also houses an array of AP2-interacting motifs, and an AP1-interacting motif. Other members of this family include Eps15R which is structurally similar to Eps15 and Eps15b, but lacks the N-terminal EH-domains (van Bergen En Henegouwen, 2009).

In addition to its involvement in endocytosis, Eps15 co-operates with AP1 in sorting cargo at the Golgi apparatus. Indeed, mutation of the AP1 binding site on Eps15 impaired the secretion of nascent secretory proteins and TGN-to-endosomal transport of mannose-6-phosphate receptor (van Bergen En Henegouwen, 2009).

AP180 and Clathrin Assembly Lymphoid Myeloid Leukemia

The ubiquitous clathrin assembly lymphoid myeloid leukemia (CALM) protein and its neuronal counterpart the assembly protein 180 (AP180) have a characteristic ANTH (AP180 N-terminal homology) domain (structurally related to the ENTH domain) (Figure 2). AP180/CALM binds to PI(4,5)P₂, but does not induce membrane deformation (Stahelin *et al.*, 2003). The C-terminus of AP180/CALM has a subtype of CBMs (DL[L/F] – Aspartate, Leucine[Leucine/Phenylalanine]), and several NPF motifs. Except for the yeast homologs, AP180s contain multiple AP2-binding motifs. AP180s are found conserved in yeast (Yap1801, Yap1802), *Caenorhabditis elegans* (unc-11), and in *Drosophila melanogaster* (Lap) while CALM is only present in mammalian cells.

CALM and AP180 play a conserved role in the endocytosis of SNAREs, specifically vesicle-associated membrane proteins (VAMP)/ R-SNAREs (Moshkanbaryans *et al.*, 2014). Depletion of AP180 and CALM in hippocampal neurons resulted in aberrant retention of synaptobrevin2/VAMP at the neuronal surface, while the abundant synaptic vesicular protein vGLUT1 was unaffected. CALM also plays a role in VAMP3 and VAMP8 internalization in non-neuronal cells. Study of the CALM–SNARE interface by X-Ray crystallography revealed that the N-terminal half of the SNARE domain binds the ANTH

domain in a manner similar to the SNARE complex formation (Moshkanbaryans *et al.*, 2014).

The involvement of CALM and AP180 in neurodegenerative diseases has gained enormous support in recent years. Genome wide-association study identified alterations in the *PICALM* gene encoding CALM as a genetic risk factor in the onset of Alzheimer's disease (AD) (Moshkanbaryans *et al.*, 2014). It has also been implicated in affecting cognitive function with aging. A recent study directly correlated the levels of CALM to the levels of pathogenic Aβ complexes in AD, by affecting the endocytosis and localization of Aβ-secretase involved in the production of toxic Aβ (Kanatsu *et al.*, 2014). On the contrary, cytotoxicity of Aβ peptide was reduced in yeast, worms, and in mice cortical neurons by overexpressing AP180 family of proteins (Moshkanbaryans *et al.*, 2014). Although these studies are contradictory, they emphasize the role of CALM in APP processing highlighting its therapeutic potential for treatment of AD.

PTB Domain Containing Family (DAB, ARH, and Numb)

This family of adaptor proteins is comprised of neuronal disabled-1 (Dab1), ubiquitous Dab2, autosomal recessive hypercholesterolemia (ARH), and numb which share a common phosphotyrosine binding domain (PTB) (Figure 2). This fold is commonly found in signaling molecules and binds to phosphorylated-tyrosines. However, the PTB domain of this adaptor family specifically recognizes the tyrosine-based signal (FX)NPXY on the cytoplasmic tail of receptor proteins (Mishra *et al.*, 2002). The PTB domain structure has been solved and shows simultaneous binding of cargo and PI(4,5)P₂ (Dvir *et al.*, 2012; Stolt *et al.*, 2003) and hence regulates vesicle budding from the plasma membrane. These adaptors have been linked to CME via their interaction with AP2 and/or clathrin directly. Dab2 and ARH interact with AP2 and clathrin, while numb and Dab1 only interact with AP2. Additionally, Dab2 has multiple NPF motifs capable of binding to EH domain containing proteins such as Eps15 (Figure 2).

Dab2 and ARH have been reported to be functionally redundant for CME of LDLR (Keyel *et al.*, 2006; Maurer and Cooper, 2006). However, while ARH-mediated LDLR internalization required AP2, Dab2-mediated LDLR internalization does not (Keyel *et al.*, 2006; Maurer and Cooper, 2006). The AP2-binding mechanism is very distinct for ARH as it engages AP2 by binding the β2 appendage while the others bind to the α-appendage of AP2 (Mishra *et al.*, 2005).

Dab2 was also required for albumin internalization (Maurer and Cooper, 2005) by recognizing the (ΦX)NPXY motif on the cytoplasmic tail of the albumin receptor, megalin (Gallagher *et al.*, 2004). PTB domain containing adaptor proteins bind to the amyloid precursor protein family as a (YX)NPXY motif is conserved among members (Howell *et al.*, 1999; Homayouni *et al.*, 1999). Both numb and Dab2 have been implicated in the endocytosis of integrins which also contains NPXY motifs in their cytoplasmic tails (Bridgewater *et al.*, 2012).

Mutations in ARH lead to defective endocytosis of LDLR and in consequence to hypercholesterolemia. Dab2 is unable to compensate for the loss of ARH, due to the low levels of

expression of Dab2 in hepatocytes. The role of Dab and Numb in APP processing and production of $A\beta$ has been exploited and depletion of these adaptor proteins was shown to reduce the levels of $A\beta$ production (Xie *et al.*, 2012).

Arrestin Family

The arrestin family of adaptor proteins shares a common arrestin-fold domain and consists of α -arrestins, a group of vacuolar protein sorting (Vps) 26 related proteins, visual and nonvisual arrestins or β -arrestins. The members of this family show evolutionary conservation from humans, flies, and worms to yeasts (Aubry *et al.*, 2009). Arrestin family in humans has 14 members: 6 α -arrestins, 4 visual and β -arrestins, and 4 Vps26 genes. Phylogenetic analysis suggests that β -arrestins, the best-studied members of this clan, branched from the α -arrestins and coevolved with G-protein coupled receptors (GPCRs) to mediate their internalization (Alvarez, 2008). The Arrestins were named for their scaffolding function in binding and terminating G-protein coupled signaling. However, many more functions of this protein family in assisting the internalization of cargoes including GPCRs have been described since.

In addition to functioning in signaling regulation of GPCRs by binding to their cytoplasmic tails, β -arrestins' ability to act as adaptor proteins became clear as they can bind clathrin and AP2 via motifs in their C-termini (Owen *et al.*, 2004; Figure 2). Also, arrestin interaction with phosphoinositides at the plasma membrane is required for recruitment of the cargo into endocytic vesicles (Gaidarov *et al.*, 1999). Structural studies have shed light on the mechanism of action of β -arrestins. In an inactive phosphorylated state, intramolecular interactions between the N- and C-domains of arrestins lead to the shielding of a core of polar residues involved in the recognition of cargo and binding other elements of the endocytic machinery. Arrestin binding to an activated phosphorylated GPCR and subsequent arrestin dephosphorylation (by an unknown phosphatase) disrupts the intramolecular interactions in arrestins and allows binding to the endocytic machinery (Kang *et al.*, 2014).

Arrestins internalize cargoes from the plasma membrane and traffic them to the early endosomes by binding phosphorylated Ser/Thr clusters in their C-termini (Kang *et al.*, 2014). The rates of receptor recycling and degradation depend on the strength of receptor–arrestin interaction. For example, Class-A receptors (β 2 Adrenergic receptor, vasopressin 1a) with weak interaction are recycled back quickly, whereas Class-B receptors (angiotensin II type 1A receptor, vasopressin 2 receptor) with strong interaction are recycled slowly (Kang *et al.*, 2014). Arrestins get ubiquitinated by the E3 ubiquitin-ligase Mdm2 and their ability to recruit deubiquitinases (DUBs) and undergo deubiquitination also leads to differential consequences. While Class-A receptor–arrestin complex recruits DUBs leading to receptor recycling, Class-B receptors do not have a conformation that supports DUB recruitment and hence leads to stable ubiquitination and degradation (Kang *et al.*, 2014). Subsequently, arrestin–GPCR interaction strength also affects signaling. Specifically Class-B receptors result in the formation of stable signalosomes due to a strong interaction.

Opsin/rhodopsin family of GPCRs serve as cargoes for the visual arrestins. In addition to classic GPCRs, β -arrestins also interact with activated/phosphorylated unconventional GPCRs (Dishevelled and Smoothened), Receptor Tyrosine Kinases (IGF-1R, TGF β RIIR), and ion channels (NHE1/5, Ca(v)1, TRPV4, Nicotinic cholinergic receptor) to mediate their internalization.

The vps26-related proteins are structurally more similar to β -arrestins than to other members of the family. Vps26A/B, 29, and 35 are four mammalian members of the Vps family part of the so-called retromer complex. The retromer is involved in endosomes-to-TGN cargo transport. However, Vps26 is not directly involved in cargo recognition and therefore, cannot be considered an adaptor.

α -Arrestins lack the helix 1 that forms the polar core of inactive β -arrestins. Another major difference is that the C-terminus of α -arrestins contains PPXY motifs instead of the clathrin- and AP2-binding sites found in β -arrestins. Members of the α -arrestin family include the 10 yeast ARTS (Arrestin related trafficking adaptors) and the 6 α -arrestin (ARRDC 1–5; Arrestin domain containing proteins and TXNIP Thioredoxin-interacting protein) mammalian proteins.

ARTS in yeast have been shown to function as cargo adaptors for the HECT ubiquitin-ligase Rsp5, i.e., facilitating cargo ubiquitination and internalization (Lin *et al.*, 2008). Similar to ARTS in yeast, the role for ARRDCs in mammalian cells as E3 ligase adaptors has been well-established linking GPCR to E3 HECT ubiquitin ligases Nedd4, AIP4, WWP1, and WWP2. In addition, ARTS have been shown to function as secondary adaptors to recruit cargo, β -arrestin-E3 ligase complex on early endosomes (Nabhan *et al.*, 2010). Further, they interact with components of the ESCRT complex and hence are implicated in multivesicular body/vacuolar sorting.

GPCRs represent the largest family of receptors (~1000 members) that localize on the cell surface and serve as target for about 40% of the drugs available. Therefore, arrestins are considered valuable targets to control GPCR signaling regulation. Further, arrestin involvement in pathological conditions has been substantially documented. Mutation of visual arrestins is associated with blindness following photoreceptor death due to impaired receptor desensitization. An alternative mechanism leading to blindness by stable association of visual arrestin to the photoreceptor rhodopsin has been suggested (Chen *et al.*, 2006). Overexpression of β -arrestins in AD contributes to the formation of extracellular neuritic plaques containing the amyloid- β ($A\beta$) peptide. This key pathogenic component of AD was reduced by disrupting the interaction between β -arrestin and β 2-adrenergic receptor which in turn lowers $A\beta$ production (Jiang *et al.*, 2013). β -Arrestins participate in signaling pathways mediated by Smoothened and Frizzled are required for the maintenance and onset of chronic myeloid leukemia (CML). Recently, targeting β -arrestin by RNA-apamers was effective in reducing tumorigenic growth in CML models and patient samples (Kotula *et al.*, 2014).

The Adaptor Protein Network

As described above, adaptors bear multiple-binding motifs that mediate several types of protein–protein interactions. This

promiscuous binding behavior has two general consequences: (1) it allows low-affinity (micromolar range) protein–protein interactions to sustain stable, but dynamic, multi-protein structures, and (2) it confers these massive complexes with classical network properties. For example, it is clear that AP2 at the plasma membrane (or AP1 at the TGN/endosomal level) along with clathrin constitute the ‘hubs’ (maximal connectivity centers) of the adaptor network. This observation explains why these key components are regulatory targets allowing the fast assembly or disassembly of the adaptor network.

See also: Interorganellar Communication: Interplay and Processes: Clathrin and Clathrin-Dependent Endocytosis

References

- Alvarez, C.E., 2008. On the origins of arrestin and rhodopsin. *BMC Evolutionary Biology* 8, 222.
- Aubry, L., Guetta, D., Klein, G., 2009. The arrestin fold: Variations on a theme. *Current Genomics* 10, 133–142.
- Badolati, R., Parolini, S., 2007. Novel insights from adaptor protein 3 complex deficiency. *Journal of Allergy and Clinical Immunology* 120, 735–741. quiz 742–743.
- van Bergen En Henegouwen, P.M., 2009. Eps15: A multifunctional adaptor protein regulating intracellular trafficking. *Cell Communication and Signaling* 7, 24.
- Bertelsen, V., Sak, M.M., Breen, K., *et al.*, 2011. A chimeric pre-ubiquitinated EGF receptor is constitutively endocytosed in a clathrin-dependent, but kinase-independent manner. *Traffic* 12, 507–520.
- Bonifacino, J.S., 2004. The GGA proteins: Adaptors on the move. *Nature Reviews Molecular Cell Biology* 5, 23–32.
- Bonnemaizon, M.L., Eipper, B.A., Mains, R.E., 2013. Role of adaptor proteins in secretory granule biogenesis and maturation. *Front Endocrinol (Lausanne)* 4, 101.
- Bresciani, R., Denzer, K., Pohlmann, R., Von Figura, K., 1997. The 46 kDa mannose-6-phosphate receptor contains a signal for basolateral sorting within the 19 juxtamembrane cytosolic residues. *Biochemical Journal* 327 (Pt 3), 811–818.
- Bridgewater, R.E., Norman, J.C., Caswell, P.T., 2012. Integrin trafficking at a glance. *Journal of Cell Science* 125, 3695–3701.
- Burgos, P.V., Mardones, G.A., Rojas, A.L., *et al.*, 2010. Sorting of the Alzheimer's disease amyloid precursor protein mediated by the AP-4 complex. *Developmental Cell* 18, 425–436.
- Cacciagli, P., Desvignes, J.P., Girard, N., *et al.*, 2014. AP1S2 is mutated in X-linked Dandy-Walker malformation with intellectual disability, basal ganglia disease and seizures (Pettigrew syndrome). *European Journal of Human Genetics* 22, 363–368.
- Chen, H., De Camilli, P., 2005. The association of epsin with ubiquitinated cargo along the endocytic pathway is negatively regulated by its interaction with clathrin. *Proceedings of the National Academy of Sciences of the United States of America* 102, 2766–2771.
- Chen, H., Ko, G., Zatti, A., *et al.*, 2009. Embryonic arrest at midgestation and disruption of Notch signaling produced by the absence of both Epsin 1 and Epsin 2 in mice. *Proceedings of the National Academy of Sciences of the United States of America* 106, 13838–13843.
- Chen, H., Slepnev, V.I., Di Fiore, P.P., De Camilli, P., 1999. The Interaction of Epsin and Eps15 with the clathrin adaptor AP-2 is inhibited by mitotic phosphorylation and enhanced by stimulation-dependent dephosphorylation in nerve terminals. *Journal of Biological Chemistry* 274, 3257–3260.
- Chen, J., Shi, G., Concepcion, F.A., Xie, G., Orian, D., 2006. Stable rhodopsin/arrestin complex leads to retinal degeneration in a transgenic mouse model of autosomal dominant retinitis pigmentosa. *Journal of Neuroscience* 26, 11929–11937.
- Collins, B.M., McCoy, A.J., Kent, H.M., Evans, P.R., Owen, D.J., 2002. Molecular architecture and functional model of the endocytic AP2 complex. *Cell* 109, 523–535.
- Collins, B.M., Watson, P.J., Owen, D.J., 2003. The structure of the GGA1-GAT domain reveals the molecular basis for ARF binding and membrane association of GGAs. *Developmental Cell* 4, 321–332.
- Coon, B.G., Burgner, J., Camonis, J.H., Aguilar, R.C., 2010. The Epsin family of endocytic adaptors promotes fibrosarcoma migration and invasion. *Journal of Biological Chemistry* 285, 33073–33081.
- Coon, B.G., Drenzo, D.M., Konieczny, S.F., Aguilar, R.C., 2011. Epsins' novel role in cancer cell invasion. *Communicative & Integrative Biology* 4, 95–97.
- De Melker, A.A., Van Der Horst, G., Borst, J., 2004. c-Cbl directs EGF receptors into an endocytic pathway that involves the ubiquitin-interacting motif of Eps15. *Journal of Cell Science* 117, 5001–5012.
- Dell'Angelica, E.C., Shotelersuk, V., Aguilar, R.C., Gahl, W.A., Bonifacino, J.S., 1999. Altered trafficking of lysosomal proteins in Hermansky-Pudlak syndrome due to mutations in the β 3A subunit of the AP-3 adaptor. *Molecular Cell* 3, 11–21.
- Doray, B., Bruns, K., Ghosh, P., Kornfeld, S., 2002. Interaction of the cation-dependent mannose 6-phosphate receptor with GGA proteins. *Journal of Biological Chemistry* 277, 18477–18482.
- Dores, M.R., Schnell, J.D., Maldonado-Baez, L., Wendland, B., Hicke, L., 2010. The function of yeast Epsin and Ede1 ubiquitin-binding domains during receptor internalization. *Traffic* 11, 151–160.
- Dvir, H., Shah, M., Girardi, E., *et al.*, 2012. Atomic structure of the autosomal recessive hypercholesterolemia phosphotyrosine-binding domain in complex with the LDL-receptor tail. *Proceedings of the National Academy of Sciences of the United States of America* 109, 6916–6921.
- Ford, M.G., Mills, I.G., Peter, B.J., *et al.*, 2002. Curvature of clathrin-coated pits driven by Epsin. *Nature* 419, 361–366.
- Gaidarov, I., Krupnick, J.G., Falck, J.R., Benovic, J.L., Keen, J.H., 1999. Arrestin function in G protein-coupled receptor endocytosis requires phosphoinositide binding. *EMBO Journal* 18, 871–881.
- Gallagher, H., Oleinikov, A.V., Fenske, C., Newman, D.J., 2004. The adaptor disabled-2 binds to the third psi xNPxY sequence on the cytoplasmic tail of megalin. *Biochimie* 86, 179–182.
- Ghosh, P., Kornfeld, S., 2003. AP-1 binding to sorting signals and release from clathrin-coated vesicles is regulated by phosphorylation. *Journal of Cell Biology* 160, 699–708.
- Ghosh, P., Kornfeld, S., 2004. The GGA proteins: Key players in protein sorting at the trans-Golgi network. *European Journal of Cell Biology* 83, 257–262.
- Gonzalez, A., Rodriguez-Boulán, E., 2009. Clathrin and AP1B: Key roles in basolateral trafficking through trans-endosomal routes. *FEBS Letters* 583, 3784–3795.
- Govero, J., Doray, B., Bai, H., Kornfeld, S., 2012. Analysis of Gga null mice demonstrates a non-redundant role for mammalian GGA2 during development. *PLoS One* 7, e30184.
- Hawryluk, M.J., Keyel, P.A., Mishra, S.K., *et al.*, 2006. Epsin 1 is a polyubiquitin-selective clathrin-associated sorting protein. *Traffic* 7, 262–281.
- Hirst, J., Borner, G.H., Antrobus, R., *et al.*, 2012. Distinct and overlapping roles for AP-1 and GGAs revealed by the “knocksidesways” system. *Current Biology* 22, 1711–1716.
- Hirst, J., Irving, C., Borner, G.H., 2013. Adaptor protein complexes AP-4 and AP-5: New players in endosomal trafficking and progressive spastic paraplegia. *Traffic* 14, 153–164.
- Hirst, J., Miller, S.E., Taylor, M.J., Von Mollard, G.F., Robinson, M.S., 2004. EpsinR is an adaptor for the SNARE protein Vti1b. *Molecular Biology of the Cell* 15, 5593–5602.
- Homayouni, R., Rice, D.S., Sheldon, M., Curran, T., 1999. Disabled-1 binds to the cytoplasmic domain of amyloid precursor-like protein 1. *Journal of Neuroscience* 19, 7507–7515.
- Horvath, C.A., Vanden Broeck, D., Boulet, G.A., Bogers, J., De Wolf, M.J., 2007. Epsin: Inducing membrane curvature. *International Journal of Biochemistry & Cell Biology* 39, 1765–1770.
- Howell, B.W., Lanier, L.M., Frank, R., Gertler, F.B., Cooper, J.A., 1999. The disabled 1 phosphotyrosine-binding domain binds to the internalization signals of transmembrane glycoproteins and to phospholipids. *Molecular and Cellular Biology* 19, 5179–5188.
- Huang, K.M., D'Hondt, K., Riezman, H., Lemmon, S.K., 1999. Clathrin functions in the absence of heterotetrameric adaptors and AP180-related proteins in yeast. *EMBO Journal* 18, 3897–3908.
- Jiang, T., Yu, J.T., Tan, M.S., Zhu, X.C., Tan, L., 2013. Beta-Arrestins as potential therapeutic targets for Alzheimer's disease. *Molecular Neurobiology* 48, 812–818.
- Kanatsu, K., Morohashi, Y., Suzuki, M., *et al.*, 2014. Decreased CALM expression reduces Abeta42 to total Abeta ratio through clathrin-mediated endocytosis of gamma-secretase. *Nature Communications* 5, 3386.
- Kang, D.S., Tian, X., Benovic, J.L., 2014. Role of beta-arrestins and arrestin domain-containing proteins in G protein-coupled receptor trafficking. *Current Opinions in Cell Biology* 27, 63–71.

- Kelly, B.T., Graham, S.C., Liska, N., *et al.*, 2014. Clathrin adaptors. AP2 controls clathrin polymerization with a membrane-activated switch. *Science* 345, 459–463.
- Keyel, P.A., Mishra, S.K., Roth, R., *et al.*, 2006. A single common portal for clathrin-mediated endocytosis of distinct cargo governed by cargo-selective adaptors. *Molecular Biology of the Cell* 17, 4300–4317.
- Kotula, J.W., Sun, J., Li, M., *et al.*, 2014. Targeted disruption of beta-arrestin 2-mediated signaling pathways by aptamer chimeras leads to inhibition of leukemic cell growth. *PLoS One* 9, e93441.
- Lefkir, Y., De Chasse, B., Dubois, A., *et al.*, 2003. The AP-1 clathrin-adaptor is required for lysosomal enzymes sorting and biogenesis of the contractile vacuole complex in Dictyostelium cells. *Molecular Biology of the Cell* 14, 1835–1851.
- Legendre-Guillemain, V., Wasiak, S., Hussain, N.K., Angers, A., McPherson, P.S., 2004. ENTH/ANTH proteins and clathrin-mediated membrane budding. *Journal of Cell Science* 117, 9–18.
- Lin, C.H., MacGurn, J.A., Chu, T., Stefan, C.J., Emr, S.D., 2008. Arrestin-related ubiquitin-ligase adaptors regulate endocytosis and protein turnover at the cell surface. *Cell* 135, 714–725.
- Madhus, I.H., 2006. Ubiquitin binding in endocytosis – How tight should it be and where does it happen? *Traffic* 7, 258–261.
- Matsuda, S., Miura, E., Wasiak, K., *et al.*, 2008. Accumulation of AMPA receptors in autophagosomes in neuronal axons lacking adaptor protein AP-4. *Neuron* 57, 730–745.
- Mattera, R., Boehm, M., Chaudhuri, R., Prabhu, Y., Bonifacio, J.S., 2011. Conservation and diversification of dileucine signal recognition by adaptor protein (AP) complex variants. *Journal of Biological Chemistry* 286, 2022–2030.
- Maurer, M.E., Cooper, J.A., 2005. Endocytosis of megalin by visceral endoderm cells requires the Dab2 adaptor protein. *Journal of Cell Science* 118, 5345–5355.
- Maurer, M.E., Cooper, J.A., 2006. The adaptor protein Dab2 sorts LDL receptors into coated pits independently of AP-2 and ARH. *Journal of Cell Science* 119, 4235–4246.
- Meyer, C., Zizioli, D., Lausmann, S., *et al.*, 2000. mu1A-adaptin-deficient mice: Lethality, loss of AP-1 binding and rerouting of mannose 6-phosphate receptors. *EMBO Journal* 19, 2193–2203.
- Mills, I.G., Praefcke, G.J., Vallis, Y., *et al.*, 2003. EpsinR: An AP1/clathrin interacting protein involved in vesicle trafficking. *Journal of Cell Biology* 160, 213–222.
- Mishra, S.K., Keyel, P.A., Edeling, M.A., *et al.*, 2005. Functional dissection of an AP-2 beta2 appendage-binding sequence within the autosomal recessive hypercholesterolemia protein. *Journal of Biological Chemistry* 280, 19270–19280.
- Mishra, S.K., Keyel, P.A., Hawryluk, M.J., *et al.*, 2002. Disabled-2 exhibits the properties of a cargo-selective endocytic clathrin adaptor. *EMBO Journal* 21, 4915–4926.
- Misra, S., Beach, B.M., Hurley, J.H., 2000. Structure of the VHS domain of human Tom1 (target of myb 1): Insights into interactions with proteins and membranes. *Biochemistry* 39, 11282–11290.
- Mitsunari, T., Nakatsu, F., Shioda, N., *et al.*, 2005. Clathrin adaptor AP-2 is essential for early embryonic development. *Molecular and Cellular Biology* 25, 9318–9323.
- Montpetit, A., Cote, S., Bruste, E., *et al.*, 2008. Disruption of AP1S1, causing a novel neurocutaneous syndrome, perturbs development of the skin and spinal cord. *PLoS Genetics* 4, e1000296.
- Moshkanbaryans, L., Chan, L.S., Graham, M.E., 2014. The biochemical properties and functions of CALM and AP180 in clathrin mediated endocytosis. *Membranes (Basel)* 4, 388–413.
- Motley, A., Bright, N.A., Seaman, M.N., Robinson, M.S., 2003. Clathrin-mediated endocytosis in AP-2-depleted cells. *Journal of Cell Biology* 162, 909–918.
- Nabhan, J.F., Pan, H., Lu, Q., 2010. Arrestin domain-containing protein 3 recruits the NEDD4 E3 ligase to mediate ubiquitination of the beta2-adrenergic receptor. *EMBO Reports* 11, 605–611.
- Nakatsu, F., Ohno, H., 2003. Adaptor protein complexes as the key regulators of protein sorting in the post-Golgi network. *Cell Structure and Function* 28, 419–429.
- Naslavsky, N., Rahajeng, J., Chenavas, S., Sorgen, P.L., Caplan, S., 2007. EHD1 and Eps15 interact with phosphatidylinositols via their Eps15 homology domains. *Journal of Biological Chemistry* 282, 16612–16622.
- Owen, D.J., Collins, B.M., Evans, P.R., 2004. Adaptors for clathrin coats: Structure and function. *Annual Review of Cell and Developmental Biology* 20, 153–191.
- Parachoniak, C.A., Park, M., 2009. Distinct recruitment of Eps15 via its coiled-coil domain is required for efficient down-regulation of the met receptor tyrosine kinase. *Journal of Biological Chemistry* 284, 8382–8394.
- Peden, A.A., Oorschot, V., Hesser, B.A., *et al.*, 2004. Localization of the AP-3 adaptor complex defines a novel endosomal exit site for lysosomal membrane proteins. *Journal of Cell Biology* 164, 1065–1076.
- Polo, S., Sigismund, S., Faretta, M., *et al.*, 2002. A single motif responsible for ubiquitin recognition and monoubiquitination in endocytic proteins. *Nature* 416, 451–455.
- Prag, G., Lee, S., Mattera, R., *et al.*, 2005. Structural mechanism for ubiquitinated-cargo recognition by the Golgi-localized, gamma-ear-containing, ADP-ribosylation-factor-binding proteins. *Proceedings of the National Academy of Sciences of the United States of America* 102, 2334–2339.
- Puertollano, R., Aguilar, R.C., Gorshkova, I., Crouch, R.J., Bonifacio, J.S., 2001. Sorting of mannose 6-phosphate receptors mediated by the GGAs. *Science* 292, 1712–1716.
- Ren, X., Farias, G.G., Canagarajah, B.J., Bonifacio, J.S., Hurley, J.H., 2013. Structural basis for recruitment and activation of the AP-1 clathrin adaptor complex by Arf1. *Cell* 152, 755–767.
- Rosenthal, J.A., Chen, H., Slepnev, V.I., *et al.*, 1999. The epsins define a family of proteins that interact with components of the clathrin coat and contain a new protein module. *Journal of Biological Chemistry* 274, 33959–33965.
- Sen, A., Madhivanan, K., Mukherjee, D., Aguilar, R.C., 2012. The epsin protein family: Coordinators of endocytosis and signaling. *Biomolecular Concepts* 3, 117–126.
- Shiba, T., Kawasaki, M., Takatsu, H., *et al.*, 2003. Molecular mechanism of membrane recruitment of GGA by ARF in lysosomal protein transport. *Nature Structural & Molecular Biology* 10, 386–393.
- Shiba, T., Takatsu, H., Nogi, T., *et al.*, 2002. Structural basis for recognition of acidic-cluster dileucine sequence by GGA1. *Nature* 415, 937–941.
- Sigismund, S., Woelk, T., Puri, C., *et al.*, 2005. Clathrin-independent endocytosis of ubiquitinated cargos. *Proceedings of the National Academy of Sciences of the United States of America* 102, 2760–2765.
- Simmen, T., Honing, S., Icking, A., Tikkanen, R., Hunziker, W., 2002. AP-4 binds basolateral signals and participates in basolateral sorting in epithelial MDCK cells. *Nature Cell Biology* 4, 154–159.
- Spradling, K.D., McDaniel, A.E., Lohi, J., Pilcher, B.K., 2001. Epsin 3 is a novel extracellular matrix-induced transcript specific to wounded epithelia. *Journal of Biological Chemistry* 276, 29257–29267.
- Stahelin, R.V., Long, F., Peter, B.J., *et al.*, 2003. Contrasting membrane interaction mechanisms of AP180 N-terminal homology (ANTH) and Epsin N-terminal homology (ENTH) domains. *Journal of Biological Chemistry* 278, 28993–28999.
- Stolt, P.C., Jeon, H., Song, H.K., *et al.*, 2003. Origins of peptide selectivity and phosphoinositide binding revealed by structures of disabled-1 PTB domain complexes. *Structure* 11, 569–579.
- Tan, J., Evin, G., 2012. Beta-site APP-cleaving enzyme 1 trafficking and Alzheimer's disease pathogenesis. *Journal of Neurochemistry* 120, 869–880.
- Tarpey, P.S., Stevens, C., Teague, J., *et al.*, 2006. Mutations in the gene encoding the Sigma 2 subunit of the adaptor protein 1 complex, AP1S2, cause X-linked mental retardation. *American Journal of Human Genetics* 79, 1119–1124.
- Tessner, K.L., Cai, X., Pasula, S., *et al.*, 2013. Epsin family of endocytic adaptor proteins as oncogenic regulators of cancer progression. *Journal of Cancer Research Updates* 2, 144–150.
- Tian, X., Hansen, D., Schedl, T., Skeath, J.B., 2004. Epsin potentiates Notch pathway activity in *Drosophila* and *C. elegans*. *Development* 131, 5807–5815.
- Traub, L.M., 2009. Tickets to ride: Selecting cargo for clathrin-regulated internalization. *Nature Reviews. Molecular Cell Biology* 10, 583–596.
- Wang, H., Traub, L.M., Weixel, K.M., *et al.*, 2006. Clathrin-mediated endocytosis of the epithelial sodium channel. Role of Epsin. *Journal of Biological Chemistry* 281, 14129–14135.
- Wang, W., Struhl, G., 2004. *Drosophila* Epsin mediates a select endocytic pathway that DSL ligands must enter to activate Notch. *Development* 131, 5367–5380.
- Wang, Y.J., Wang, J., Sun, H.Q., *et al.*, 2003. Phosphatidylinositol 4 phosphate regulates targeting of clathrin adaptor AP-1 complexes to the Golgi. *Cell* 114, 299–310.
- Xie, Z., Dong, Y., Maeda, U., Xia, W., Tanzi, R.E., 2012. RNAi-mediated knock-down of Dab and Numb attenuate Abeta levels via gamma-secretase mediated APP processing. *Translational Neurodegeneration* 1, 8.
- Yang, W., Li, C., Ward, D.M., Kaplan, J., Mansour, S.L., 2000. Defective organellar membrane protein trafficking in Ap3b1-deficient cells. *Journal of Cell Science* 113 (Pt 22), 4077–4086.

SNAREs: Membrane Fusion and Beyond

T Galli and D Tareste, Institut Jacques Monod, Paris, France; INSERM ERL U950, Paris, France; and Sorbonne Paris Cité, Paris, France

© 2016 Elsevier Inc. All rights reserved.

Introduction: A Brief Historical Perspective on Soluble *N*-Ethylmaleimide-Sensitive Factor Attachment Protein Receptors (SNAREs)

The genesis of the discovery of SNAREs and the molecular mechanisms of their mode of action have been recently covered in details in several reviews (Jahn and Scheller, 2006; Sotirakis and Galli, 2007; Hussain and Davanger, 2011; Mellman and Emr, 2013) including the press release of the 2013 Nobel Prize in Physiology and Medicine (Nobelprize.org, 2013). In brief, the initial SNARE hypothesis of membrane fusion was proposed in the seminal article by Rothman and coworkers (Söllner *et al.*, 1993). This hypothesis was based on their discovery of a complex composed of a synaptic vesicle protein (vesicle-associated membrane protein, VAMP2) and two plasma membrane proteins (Syntaxin 1 and SNAP-25), suggesting that this complex (named SNARE complex) allows vesicles to dock and fuse with their target membrane. The take-home message of 20 years of work on SNAREs, from 1993 to 2013, is certainly that SNAREs are both necessary and sufficient for membrane fusion. Their necessity was shown by genetic studies in several organisms (most notably amoebae, fly, worm, and mouse) which showed paralytic phenotype when VAMP1 or VAMP2 were inactivated (see Schoch *et al.*, 2001; Nystuen *et al.*, 2007 for mouse VAMP null mutants' phenotype) and the finding that clostridial neurotoxins, the most potent blockers of neurotransmitter release, a special case of vesicular traffic, are zinc-proteases which cleave and inactivate synaptic SNAREs (Galli and Haucke, 2004). The sufficiency of SNAREs for membrane fusion was essentially shown using artificial membranes devoid of any other protein (Weber *et al.*, 1998) and cells modified to express SNAREs on their surface (Hu *et al.*, 2003). In these systems, the formation of a complex between a vesicular SNARE (v-SNARE) on one membrane and its target SNARE (t-SNARE) on the other membrane was sufficient to drive the fusion of these membranes. Still the story of SNAREs was not fully written yet as much remained to be understood, notably on how SNAREs operate at the molecular level and how they organize intracellular membrane traffic both spatially and temporally. Here we will illustrate the current frontiers in 'SNARE science' with two main topics: (1) the biophysical properties of SNAREs as membrane fusion nanomachines and (2) their emergence as multitask proteins with a special focus on the case of the Longin SNARE proteins Sec22b and TI-VAMP/VAMP7.

SNAREs: Membrane Fusion Nanomachines

The first reconstitution of SNAREs in artificial membranes was carried out by the Rothman lab (Weber *et al.*, 1998). In this work, they generated a population of liposomes bearing the synaptic vesicle protein VAMP2 (v-liposomes) on one hand,

and another population of liposomes carrying the presynaptic plasma membrane t-SNARE composed of Syntaxin 1 and SNAP-25 (t-liposomes) on the other hand. *In vivo*, VAMP2 and Syntaxin 1 each contain a C-terminal transmembrane domain (TMD), whereas SNAP-25 lacks a TMD, and is targeted to the presynaptic plasma membrane *via* palmitoylation of its central cysteine residues. Here the trick consisted in using bacteria coproducing Syntaxin 1 and SNAP-25, thus generating high amounts of properly folded and active t-SNAREs. Fusion was measured using Förster resonance energy transfer (FRET) between two fluorochromes, NBD and rhodamine, linked to lipids of the v-liposome. Upon fusion with t-liposomes, the fluorescent lipids are diluted and the FRET signal decreases proportionally. While this initial reconstitution experiment supported the notion that SNAREs constitute the minimal machinery for membrane fusion, it could not tell how they operate. Membrane fusion is thermodynamically unfavorable mainly because repulsive forces between the hydration layers of two interacting membranes oppose their approach at distances less than ~ 2 nm. *In vitro*, close membrane apposition and fusion can be triggered artificially, in the absence of any proteins, upon the addition of divalent cations (Duzgunes *et al.*, 1981) or poly-ethylene glycol (Lentz, 2007). So how do SNAREs produce similar membrane-bridging effects? The crystal structure of the synaptic SNARE complex (Sutton *et al.*, 1998) revealed that an essential domain of the SNARE proteins, therefore referred to as the SNARE motif, was able to fold as a coiled-coil upon formation of the SNARE complex. In this structure, the SNARE motifs assemble as a four-helix bundle (VAMP2 and Syntaxin 1 each contribute one helix and SNAP-25 contributes two helices) with the four helices oriented in parallel (N-termini at the same end). Later studies then showed that v- and t-SNAREs assemble from their membrane-distal N-termini and toward their membrane-proximal C-termini, thus bringing the membranes close together as they fold-up (Melia *et al.*, 2002; Sorensen *et al.*, 2006; Figure 1).

The energy released during this zipper-like assembly was evaluated using a surface forces apparatus, which measures the interaction energy as a function of distance between two mica-supported lipid bilayers. In this experiment, one bilayer was functionalized with the cytoplasmic domain of VAMP2 (v-SNARE) and the other one with the cytoplasmic domain of Syntaxin 1 bound to SNAP-25 (t-SNARE); these v- and t-SNAREs were covalently attached to their respective bilayer through a lipidic anchor instead of their natural TMD. In this setup, v- and t-SNAREs started to assemble when the bilayers were ~ 10 nm apart and formed a partially zippered complex (unstructured in its membrane-proximal region) releasing ~ 35 k_BT (Li *et al.*, 2007), an amount of energy well compatible with the notion that SNAREs provide the energy for membrane fusion. Indeed, the energy required to fuse two lipid bilayers was estimated in the range of 50–100 k_BT (Cohen and Melikyan, 2004). Recent experiments at the single

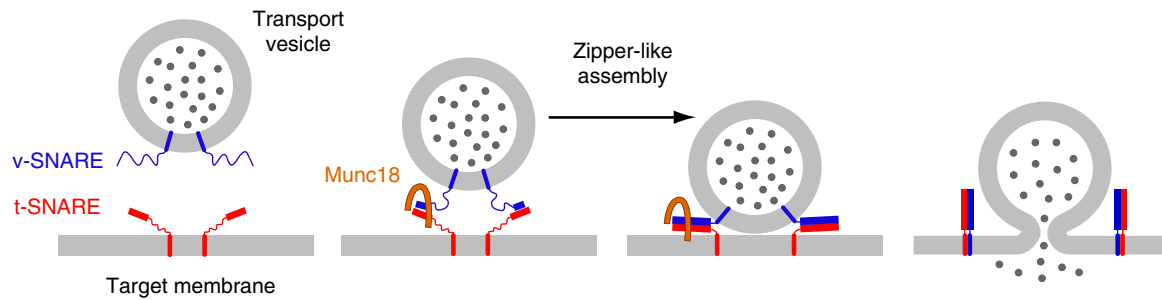


Figure 1 Core mechanism of the SNARE fusion nanomachine. The transport vesicle (v-) SNARE protein (in blue) forms a membrane-bridging complex with the target membrane (t-) SNARE protein (in red). This complex assembles like a zipper (from the N-terminal toward the C-terminal regions of the cognate v- and t-SNARE proteins), which brings the membranes into close apposition and triggers their fusion. Munc18 (in orange, shown only on the left for clarity) acts as a stabilizer of the SNARE complex; its U-shape geometry is appropriate to clasp the four-helix bundle of zippering t- and v-SNARE proteins.

SNARE complex level using optical or magnetic tweezers (Gao *et al.*, 2012; Min *et al.*, 2013) have confirmed the energy released by partial (membrane-distal) assembly of the SNARE complex and further found the occurrence of sequential zippering intermediates, with full zippering releasing ~ 65 k_BT. This is particularly interesting because it suggests that SNARE complexes do not necessarily wind up completely leading to fusion at once and opens the possibility for regulation of C-terminal zippering (and thus membrane fusion) by regulatory proteins. Several groups have then studied the number of SNARE complexes required to fuse membranes. Whilst one seemed to be enough in the case of liposomes fusing *in vitro* (van den Bogaart *et al.*, 2010), two to three appeared to be required for the exocytosis of secretory vesicles *in vivo* (Hua and Scheller, 2001; Mohrmann *et al.*, 2010; Sinha *et al.*, 2011). These apparently contradictory results were recently reconciled using a novel lipid bilayer nanodiscs system, which showed that although one SNARE complex was enough to trigger lipid bilayer mixing, at least three SNARE complexes were necessary for efficient content mixing (Shi *et al.*, 2012). Whether or not SNARE complexes are arranged geometrically (Megighian *et al.*, 2013) potentially by partners like Sec1/Munc18 (SM) proteins or complexins (Rizo *et al.*, 2006; Kummel *et al.*, 2011) to provide the energy required for fusion synergistically still needs to be further studied. From this, we are led to conclude that SNAREs operate as a mechanical device delivering energy to the membranes to bring them in close proximity until they fuse.

This then raises the questions of how the force is transmitted to the interacting bilayers and what is the role of SNARE TMDs in this mechanism. These questions have been studied using both *in vitro* liposome fusion assays (McNew *et al.*, 1999, 2000) and *in vivo* electrophysiological recordings in chromaffin or neuronal cells (Kesavan *et al.*, 2007; Zhou *et al.*, 2013) by extending the region between the SNARE motifs and the TMDs and/or replacing the TMDs with lipidic anchors. Introducing a flexible linker of ~ 10 amino acids between the SNARE motif and the TMD of VAMP2 or Syntaxin 1 did not affect liposome fusion *in vitro* (McNew *et al.*, 1999) or spontaneous vesicle fusion *in vivo* (Kesavan *et al.*, 2007), but it severely impaired calcium-mediated neurotransmitter release *in vivo* (Kesavan *et al.*, 2007; Zhou *et al.*, 2013). Addition of longer (~ 20 amino acids) flexible linkers completely

abolished fusion both *in vitro* and *in vivo*. In addition, helical continuity between SNARE motifs and TMDs is not required for fusion as shown by the moderate effect on liposome fusion when introducing two prolines in the juxtamembrane domain of VAMP2 or Syntaxin 1 (McNew *et al.*, 1999). Altogether these results show that conformational coupling between SNARE assembly and the TMDs is not important for fusion *per se*, and suggest that the force for fusion is rather transmitted by pulling on the membrane anchors as SNARE motifs zipper their membrane-proximal end. This notion is further supported by experiments in which TMDs could be functionally replaced by lipidic anchors *in vitro*, as long as the hydrophobic chain of the lipid was long enough to span both leaflets of the bilayer (McNew *et al.*, 2000) and *in vivo*, where both VAMP2 and Syntaxin 1 anchored *via* 'short' lipids could replace the endogenous molecules (Zhou *et al.*, 2013). Functional restoration *in vivo* appeared nevertheless incomplete for evoked release which suggests that TMDs may be involved in coupling SNAREs with some other components involved in calcium-triggered fusion events. The fact that *in vitro* fusion requires long hydrophobic anchors, while *in vivo* fusion apparently does not, can be explained by the absence of accessory proteins in minimal reconstitution systems. The membrane buckling effect resulting from pulling on both leaflets of the bilayer at once upon zippering of cognate SNAREs – when they are anchored to long hydrophobic anchors *in vitro* (McNew *et al.*, 2000) – could be compensated *in vivo* by regulatory proteins with membrane deformation properties, such as synaptotagmin (McMahon *et al.*, 2010; Bharat *et al.*, 2014). This is consistent with *in vitro* reconstitution experiments of yeast vacuolar fusion (Xu *et al.*, 2011), where fusion by lipid-anchored SNAREs was only possible in the presence of accessory proteins involved in membrane tethering and protein recycling.

The Regulators of the SNARE Nanomachine

Another fundamental question in the field of intracellular membrane traffic is how SNARE assembly is regulated to allow fusion to occur at the right place and at the right time under the control of specific signaling mechanisms. We will here only briefly present the main proteins that collaborate with SNAREs

to regulate membrane fusion events spatially and temporally. Sec1/Munc18 (SM) proteins are probably the most important partners of SNAREs, which are now even considered (together with the 4 SNARE motifs) as the fifth component of the core membrane fusion nanomachine (Sudhof and Rothman, 2009). Accordingly, *in vivo* deletion of SM proteins leads to a severe reduction of neurotransmission in various organisms including flies, nematodes and mice (Harrison *et al.*, 1994; Verhage *et al.*, 2000; Weimer *et al.*, 2003). The structure of SM proteins displays a U-shaped geometry, with the central cavity being involved in SNARE binding (Archbold *et al.*, 2014). The neuronal SM protein Munc18-1 binds to the closed (fusion incompetent conformation) of the Syntaxin 1 molecule (Misura *et al.*, 2000), but also to the SNARE complex (with Syntaxin 1 in its open conformation) (Dulubova *et al.*, 2007). The first mode of binding may protect Syntaxin 1 from undesirable interactions during its trafficking toward the plasma membrane, whereas the second mode of binding may stabilize SNARE complex and thus stimulate fusion. *In vitro*, the synaptic SM protein Munc18-1 was shown to exclusively activate membrane adhesion and fusion mediated by the synaptic SNARE complex, thus improving fusion specificity (Tareste *et al.*, 2008; Shen *et al.*, 2007). Interestingly, another *in vitro* reconstitution experiment has demonstrated that the HOPS complex (a multicomponent tethering complex of the endolysosomal system including the SM protein Vps33) protects membrane-bridging SNARE complexes from being disassembled by the Sec17p/Sec18p recycling machinery (Xu *et al.*, 2010). A conserved feature of the SM proteins thus seems to be their ability to clasp the four-helix bundle of zippering SNARE complexes, thus increasing their assembling energy (Figure 1).

Proteins from the complexin and synaptotagmin families constitute another class of SNARE regulators that are particularly important in calcium-triggered fusion events. These regulators are not necessary for fusion *per se* (Geppert *et al.*, 1994; Reim *et al.*, 2001), but they allow the accumulation, at the plasma membrane, of a pool of fusion-ready vesicles awaiting the signal (calcium entry) for rapid and synchronous release. Recent genetic and biophysical studies indicate that complexin traps the SNARE complex in a partially zippered, fusion incompetent state, and that calcium-bound synaptotagmin releases this clamp probably by physically displacing

complexin from the SNARE complex (Tang *et al.*, 2006; Schaub *et al.*, 2006; Giraudo *et al.*, 2006). Complexin exerts its clamping activity by binding to the membrane-proximal region of the t-SNARE, thus preventing its assembly with the C-terminal residues of the v-SNARE (Giraudo *et al.*, 2009; Li *et al.*, 2011). In addition to releasing the complexin clamp and thus allowing full SNARE complex zippering, calcium-bound synaptotagmin might activate fusion upon its insertion into the plasma membrane, which induces positive (fusion-promoting) curvature (McMahon *et al.*, 2010; Figure 2).

In conclusion, SNAREs appear as the basic nanomachine for intracellular membrane fusion in vesicular traffic and they operate most likely by mediating close apposition of membranes, which lowers the energy barrier for fusion. How exactly SNARE assembly is coupled to membrane fusion and what is the role of SNARE TMDs in the process still need to be clarified. The mechanics of the SNARE nanomachine is regulated by proteins from the SM, complexin, and synaptotagmin families, which facilitate or inhibit SNARE assembly to allow precise spatial and temporal organization of membrane fusion events.

SNAREs: More than Fusogenic Proteins? The Case of the Longin v-SNAREs Sec22b and TI-VAMP/VAMP7

After the discovery that SNAREs constitute the core machinery of membrane fusion, several studies have suggested that SNAREs could play additional functions linked to their interactions with other proteins involved in membrane traffic. Here we will review the particular case of the Longin v-SNAREs Sec22b, found to be involved in membrane contact sites and TI-VAMP/VAMP7, which we propose to view as a multitask protein.

Our recent data suggest a role for Sec22b in neurite growth through anchoring the endoplasmic reticulum (ER) to the plasma membrane, and generating a non-fusogenic SNARE complex involved in plasma membrane growth (Petkovic *et al.*, 2014). The role of the ER in plasma membrane expansion is well-established, because it is the major site of lipid synthesis in cells, and the source of secretory vesicles which are required for growth. Recent studies have also highlighted roles for membrane contacts between the ER and other organelles

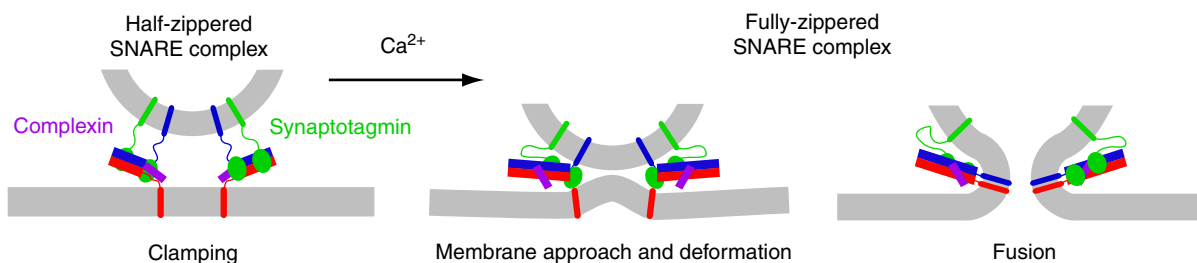


Figure 2 Model for complexin and synaptotagmin function in calcium-triggered fusion. Before calcium entry, complexin (in purple) clamps fusion by binding to the C-terminal region of the t-SNARE protein, thus preventing full assembly of the SNARE complex. This clamp is further stabilized by synaptotagmin (in green), which also binds to SNAREs in the absence of calcium. Upon calcium entry, synaptotagmin modifies its mode of binding to SNAREs, which displaces complexin and induces C-terminal assembly of the SNARE complex. Concomitantly, calcium-bound synaptotagmin inserts into the plasma membrane thus causing positive (fusion-promoting) curvature. Munc18 was removed from this scheme to better illustrate the role of complexin and synaptotagmin.

including the plasma membrane. ER-plasma membrane contact sites (MCS) are thought to be involved in (1) the exchange of molecules, such as lipids, between the two bilayers, (2) the function of ER-localized enzymes that act 'in trans' on the plasma membrane, and (3) the control of calcium homeostasis (Toulmay and Prinz, 2011; Stefan *et al.*, 2011; Hogan *et al.*, 2010; Friedman and Voeltz, 2011; Elbaz and Schuldiner, 2011). In cells of higher eukaryotes, MCS were first described by electron microscopy in the 1950s in the case of muscle cells (Porter and Palade, 1957), where they are extremely abundant and are responsible for depolarization-contraction coupling. MCS are also abundant in neurons (Rosenbluth, 1962) and have subsequently been described in many other cell types. Yeast cells are particularly rich in MCS, where approximately half of the ER localizes to the cell periphery. There are an estimated 1100 sites per yeast cell in which the distance between the ER and the plasma membrane is less than 30 nm (Pichler *et al.*, 2001). Mutants in six genes involved in MCS also lead to growth defects in yeast (Manford *et al.*, 2012), suggesting a clear link between MCS and growth. We found that the ER-localized v-SNARE Sec22b interacts with the plasma membrane t-SNARE Syntaxin 1 in neurons, and experiments in budding yeast also showed an interaction between Sec22 and the plasma membrane t-SNARE Sso1 (ortholog of Syntaxin 1), suggesting conservation of the interaction. We did not find co-immunoprecipitation of Sec22b with SNAP-23, -25 or -29, suggesting that Sec22b is not part of a fusogenic SNARE complex at ER-plasma membrane contact sites. Accordingly, the complex Syntaxin 1/Sec22b did not mediate fusion either in neurons *in vivo* or between liposomes *in vitro* (Petkovic *et al.*, 2014). Interestingly, a similar binary Syntaxin 4/Sec22b (without SNAP-23) complex was identified in mammalian dendritic cells (Cebrian *et al.*, 2011). Introduction of a long rigid spacer between the SNARE and the TMDs, increasing the distance between ER and plasma membrane, impaired neuronal growth but did not perturb the role of Sec22b in the secretory pathway. We further demonstrated that elimination of Sec22 function in yeast impaired the non-vesicular lipid transfer of phosphoinositides between the ER and the plasma membrane, and that Sec22 and Sso1 interact with Osh2 and Osh3, two lipid transfer proteins (LTPs). Interestingly, extended-Synaptotagmin 2 (E-Syt2), a protein closely related to the neuronal synaptotagmin partners of the synaptic SNARE complex, and which regulates the formation of ER-plasma membrane contact sites (Giordano *et al.*, 2013), was recently suggested to be a phospholipid transfer protein (Schauder *et al.*, 2014). Together, these results lead us to propose that the interaction between ER Sec22 and plasma membrane Syntaxin participates in MCS function (Figure 3), likely fostering non-vesicular lipid transfer by shortening the distance between the membranes, and that this highly conserved mechanism contributes to plasma membrane expansion (Petkovic *et al.*, 2014). While most ER-plasma membrane contact sites are 30 nm large, the Sec22b/Syntaxin 1 complex could possibly bring them as close as 10 nm (Figure 3). Interestingly, in yeast, Sec22 interacts with Sso1 and the SNAP25-like Sec9, and is involved in the transport of the autophagy-related protein Atg9 (Nair *et al.*, 2011), suggesting that Sec22b and Syntaxin 1 could coordinate lipid transfer when SNAP-25 is kept out of their complex (by an as yet

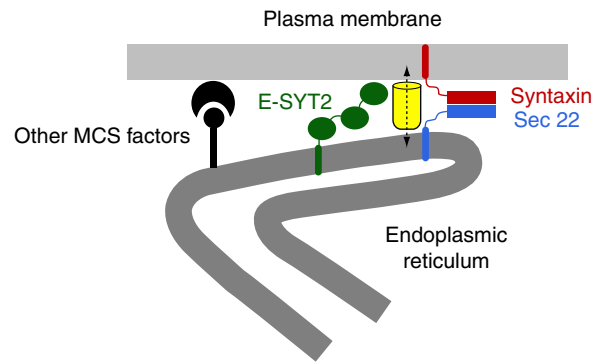


Figure 3 Putative role of Sec22 and Syntaxin at ER-plasma membrane contact sites. A complex of Sec22 and Syntaxin is shown to tether in close proximity the endoplasmic reticulum membrane and the plasma membrane and to interact with lipid transfer proteins (in yellow), as shown for oxysterol-binding protein Osh2/3 in yeast. Extended-synaptotagmin 2 (E-Syt2) is an ER protein which is involved in membrane contact sites and could potentially also interact with the non-fusogenic binary Syntaxin/Sec22 SNARE complex depicted. Other tethering factors (such as STIM1-Orai1) are depicted without detail.

unidentified mechanism/factor), and classical membrane fusion regulating autophagosome biogenesis when SNAP-25 or possibly another SNAP is present.

Our team has identified and characterized TI-VAMP/VAMP7, a SNARE protein localized to secretory vesicles and insensitive to tetanus and botulinum neurotoxins (Galli *et al.*, 1998). VAMP7 has the classical SNARE sequence with a SNARE motif and a TMD but also includes an N-terminal extension which we called Longin domain (Filippini *et al.*, 2001). The Longin domain plays an auto-inhibitory role through intramolecular interaction with the SNARE motif thus controlling the fusogenic activity of VAMP7 (Martinez-Arca *et al.*, 2000, 2003). We have characterized the role of VAMP7 and searched for its partners by yeast two-hybrid screens, an approach which proved to be very productive in that case. Screens performed with the cytoplasmic domain of VAMP7 as bait pulled out its already known SNARE partners: Syntaxin 1, Syntaxin 3, SNAP-23 and SNAP-25 (Burgo *et al.*, 2009; Martinez-Arca *et al.*, 2003; Figure 4). These molecules are present at the plasma membrane and the SNARE complexes VAMP7/Syntaxin 1/SNAP25 and VAMP7/Syntaxin 3/SNAP-23 are thought to allow for fusion of secretory vesicles with the plasma membrane (Chaineau *et al.*, 2009).

We further identified the interaction of VAMP7 with the δ -subunit of the molecular coat AP-3 in yeast two-hybrid screens using its Longin and cytoplasmic domains as baits in both fly and human libraries (Burgo *et al.*, 2009; Formstecher *et al.*, 2005; Martinez-Arca *et al.*, 2003; Figure 4). AP-3 is a clathrin adaptor involved in the intracellular targeting from the Golgi apparatus and endosomes, to lysosomes and secretory organelles, including secretory lysosome-related organelles and synaptic vesicles. Mutations in AP-3 subunits lead to coat dilution, bleeding, and neurological defects (Kantheti *et al.*, 1998). Recent work further suggests important function of AP-3 in the biogenesis of peptide-secreting granules at the Golgi apparatus (Sirkis *et al.*, 2013; Asensio *et al.*, 2010).

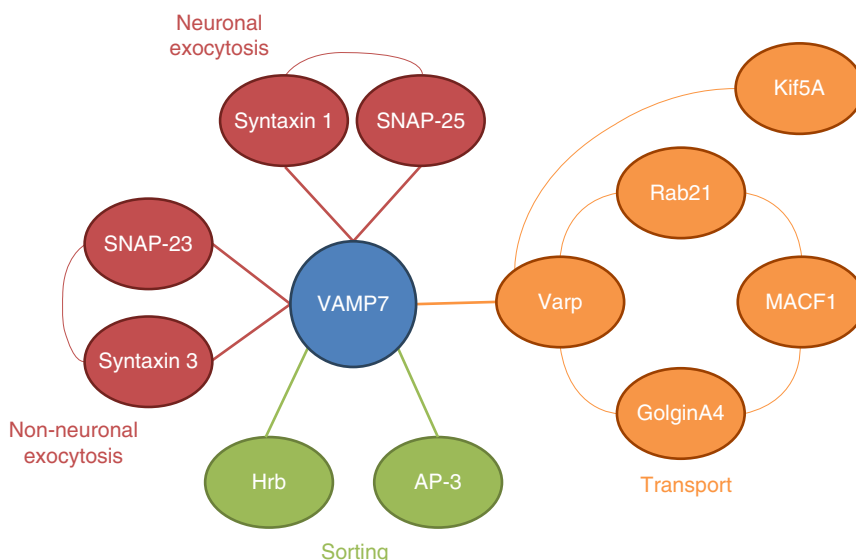


Figure 4 VAMP7 network of protein–protein interactions. Direct partners of VAMP7 include the neuronal t-SNAREs Syntaxin 1 and SNAP-25, the non-neuronal t-SNAREs Syntaxin 3 and SNAP-23, all involved in exocytosis, Hrb involved in clathrin-dependent endocytosis, AP-3 involved in sorting to late endosomes and synaptic vesicles, and Varp. Varp is a guanine-nucleotide exchange factor of Rab21. Varp interacts with the molecular motor Kif5A and GolginA4. One of Rab21's effectors is MACF1 a very large protein which binds to both actin and microtubules. GolginA4 interacts also with MACF1. In summary, VAMP7 network of interaction includes proteins involved in the three major vesicular trafficking processes: sorting, transport, and fusion. The present interactome is not exhaustive and solely illustrates the above-mentioned principle.

Accordingly, the VAMP7/AP-3 interaction was further found to be crucial for intracellular targeting of VAMP7 to late endosomes in non-neuronal cells and synaptic vesicles in neurons (Bennett *et al.*, 2008; Proux-Gillardeaux *et al.*, 2007; Salazar *et al.*, 2006; Martinez-Arca *et al.*, 2003). This interaction was then confirmed at the structural level and was shown to involve Longin residues interacting with the SNARE motif (Kent *et al.*, 2012). Whether VAMP7 is sorted *in solo* by AP-3 as originally suggested by the interaction with both Longin and cytoplasmic domains (Martinez-Arca *et al.*, 2003) or in complexes with t-SNAREs (Kent *et al.*, 2012) is not yet clear.

Another partner identified in yeast two-hybrid screens using the cytoplasmic domain of VAMP7 as bait was HIV Rev binding protein (Hrb) (Pryor *et al.*, 2008; Chaîneau *et al.*, 2008; Figure 4). We showed that Hrb, which is a clathrin accessory protein (Schmid and McMahon, 2007), is involved in clathrin-dependent endocytosis (Chaîneau *et al.*, 2008) in good agreement with the phenotype of the Hrb knockout mouse (Khawaja *et al.*, 2010) demonstrating that Hrb loss leads to inefficient transferrin uptake, a typical clathrin-dependent mechanism, and in contrast with the proposal of a role of Hrb restricted to VAMP7 clathrin-dependent endocytosis (Pryor *et al.*, 2008).

We then went on to show that VAMP7 interacts with Vps9 and Ankyrin repeat protein (Varp) (Burgo *et al.*, 2009), an exchange factor for the small GTPase Rab21 (Zhang *et al.*, 2006; Figure 4) and effector of Rab32/38 (Wang *et al.*, 2008). Varp interacts with the closed conformation of VAMP7 (Schafer *et al.*, 2012). More recently, again using yeast two-hybrid screens with Varp and Rab21 as baits, we showed that VAMP7 is the starting point of a molecular network that combines proteins belonging to the main classes involved in vesicular trafficking (Figure 4): Varp, kinesin 1 (Kif5A), a

molecular motor partner of Varp, GolginA4, a Golgi attachment factor partner of Varp, and the spectraplakine MACF1, an effector of Rab21 (Burgo *et al.*, 2012), which binds both actin and microtubules. We found that this network can send VAMP7 vesicles to the cell periphery along microtubules, thus allowing exocytosis.

In conclusion, these studies altogether show that VAMP7, on top of its SNARE fusogenic activities with several t-SNAREs, is also able to interact with molecules involved in sorting (AP-3 and Hrb) and transport (the Varp-Kif5A-GolginA4-Rab21-MACF1 molecular network) (Figure 4). Whether VAMP7 plays an active role in recruiting AP-3, Hrb and Varp or is just a follower remains to be determined as we do not know if there are any defects of AP-3, Hrb and/or Varp function in the absence of VAMP7. The idea that SNAREs are more than mere fusogenic proteins but are also connected with proteins involved in sorting has also been illustrated for VAMP2 which interacts with AP180 and Calm, both involved in clathrin-dependent endocytosis (Miller *et al.*, 2011; Koo *et al.*, 2011) and for VAMP4 which interacts with AP-1, involved in clathrin-dependent sorting at the *trans*-Golgi network (Hinnens *et al.*, 2003; Peden *et al.*, 2001). This would further suggest that sorting, fusion, and possibly transport are closely connected at the molecular level.

Conclusion

In conclusion, 20 years of work on SNAREs have convincingly led to the principle that they constitute the core nanomachines for intracellular membrane fusion. SNAREs also form non-fusogenic complexes playing a role in the generation of membrane contact sites, and interact with multiple proteins

involved in sorting and transport, thus making them emerge as general organizers of intracellular membrane adhesion and fusion events.

See also: Cytoskeleton and Motors: Cytoskeletal Components: Kinesin Superfamily Proteins (KIFs) as a Fundamental Component of Life: Intracellular Transport and Beyond. Dynamic Integration: Dynamics: Theory of Cargo and Membrane Trafficking

References

- Archbold, J.K., Whitten, A.E., Hu, S.H., Collins, B.M., Martin, J.L., 2014. SNARE-ing the structures of Sec1/Munc18 proteins. *Current Opinion in Structural Biology* 29C, 44–51.
- Asensio, C.S., Sirkis, D.W., Edwards, R.H., 2010. RNAi screen identifies a role for adaptor protein AP-3 in sorting to the regulated secretory pathway. *Journal of Cell Biology* 191, 1173–1187.
- Bennett, N., Letourneur, F., Ragno, M., Louwagie, M., 2008. Sorting of the v-SNARE VAMP7 in Dictyostelium discoideum: A role for more than one Adaptor Protein (AP) complex. *Experimental Cell Research* 314, 2822–2833.
- Bharat, T.A., Malsam, J., Hagen, W.J., *et al.*, 2014. SNARE and regulatory proteins induce local membrane protrusions to prime docked vesicles for fast calcium-triggered fusion. *EMBO Reports*.
- van den Bogaart, G., Holt, M.G., Bunt, G., *et al.*, 2010. One SNARE complex is sufficient for membrane fusion. *Nature Structural & Molecular Biology* 17, 358–364.
- Burgo, A., Proux-Gillardeaux, V., Sotirakis, E., *et al.*, 2012. A Molecular Network for the Transport of the TI-VAMP/VAMP7 Vesicles from Cell Center to Periphery. *Developmental Cell* 23, 166–180.
- Burgo, A., Sotirakis, E., Simmler, M., *et al.*, 2009. Role of Varp, a Rab21 exchange factor and TI-VAMP/VAMP7 partner, in neurite growth. *EMBO Reports* 10, 1117–1124.
- Cebrian, I., Visentin, G., Blanchard, N., *et al.*, 2011. Sec22b regulates phagosomal maturation and antigen crosspresentation by dendritic cells. *Cell* 147, 1355–1368.
- Chaineau, M., Danglot, L., Galli, T., 2009. Multiple roles of the vesicular-SNARE TI-VAMP in post-Golgi and endosomal trafficking. *FEBS Letters* 583, 3817–3826.
- Chaineau, M., Danglot, L., Proux-Gillardeaux, V., Galli, T., 2008. Role of HRB in Clathrin-dependent Endocytosis. *Journal of Biological Chemistry* 283, 34365–34373.
- Cohen, F.S., Melikyan, G.B., 2004. The energetics of membrane fusion from binding, through hemifusion, pore formation, and pore enlargement. *Journal of Membrane Biology* 199, 1–14.
- Dulubova, I., Khvotchev, M., Liu, S., *et al.*, 2007. Munc18-1 binds directly to the neuronal SNARE complex. *Proceedings of the National Academy of Sciences of the United States of America* 104, 2697–2702.
- Duzgunes, N., Nir, S., Wilschut, J., *et al.*, 1981. Calcium- and magnesium-induced fusion of mixed phosphatidylserine/phosphatidylcholine vesicles: Effect of ion binding. *Journal of Membrane Biology* 59, 115–125.
- Elbaz, Y., Schuldiner, M., 2011. Staying in touch: The molecular era of organelle contact sites. *Trends in Biochemical Science* 36, 616–623.
- Filippini, F., Rossi, V., Galli, T., *et al.*, 2001. Longins: A new evolutionary conserved VAMP family sharing a novel SNARE domain. *Trends in Biochemical Science* 26, 407–409.
- Formstecher, E., Aresta, S., Collura, V., *et al.*, 2005. Protein interaction mapping: A *Drosophila* case study. *Genome Research* 15, 376–384.
- Friedman, J.R., Voeltz, G.K., 2011. The ER in 3D: A multifunctional dynamic membrane network. *Trends in Cell Biology* 21, 709–717.
- Galli, T., Haucke, V., 2004. Cycling of synaptic vesicles: How far? How fast! *Science's STKE*. 2004, re19.
- Galli, T., Zahraoui, A., Vaidyanathan, V.V., *et al.*, 1998. A novel tetanus neurotoxin-insensitive vesicle-associated membrane protein in SNARE complexes of the apical plasma membrane of epithelial cells. *Molecular Biology of the Cell* 9, 1437–1448.
- Gao, Y., Zorman, S., Gundersen, G., *et al.*, 2012. Single reconstituted neuronal SNARE complexes zipper in three distinct stages. *Science* 337, 1340–1343.
- Geppert, M., Goda, Y., Hammer, R.E., *et al.*, 1994. Synaptotagmin I: A major Ca^{2+} sensor for transmitter release at a central synapse. *Cell* 79, 717–727.
- Giordano, F., Saheki, Y., Idevall-Hagren, O., *et al.*, 2013. PI(4,5)P(2)-dependent and Ca^{2+} -regulated ER-PM interactions mediated by the extended synaptotagmins. *Cell* 153, 1494–1509.
- Giraudo, C.G., Eng, W.S., Melia, T.J., Rothman, J.E., 2006. A Clamping Mechanism Involved in SNARE-Dependent Exocytosis. *Science* 313, 676–680.
- Giraudo, C.G., Garcia-Diaz, A., Eng, W.S., *et al.*, 2009. Alternative zippering as an on-off switch for SNARE-mediated fusion. *Science* 323, 512–516.
- Harrison, S.D., Broadie, K., van de Goor, J., Rubin, G.M., 1994. Mutations in the *Drosophila* Rop gene suggest a function in general secretion and synaptic transmission. *Neuron* 13, 555–566.
- Hinners, I., Wendler, F., Fei, H., *et al.*, 2003. AP-1 recruitment to VAMP4 is modulated by phosphorylation-dependent binding of PACS-1. *EMBO Reports* 4, 1182–1189.
- Hogan, P.G., Lewis, R.S., Rao, A., 2010. Molecular basis of calcium signaling in lymphocytes: STIM and ORAI. *Annual Review of Immunology* 28, 491–533.
- Hu, C., Ahmed, M., Melia, T.J., *et al.*, 2003. Fusion of cells by flipped SNAREs. *Science* 300, 1745–1749.
- Hua, Y., Scheller, R.H., 2001. Three SNARE complexes cooperate to mediate membrane fusion. *Proceedings of the National Academy of Sciences of the United States of America* 98, 8065–8070.
- Hussain, S., Davanger, S., 2011. The discovery of the soluble N-ethylmaleimide-sensitive factor attachment protein receptor complex and the molecular regulation of synaptic vesicle transmitter release: The 2010 Kavli Prize in neuroscience. *Neuroscience* 190, 12–20.
- Jahn, R., Scheller, R.H., 2006. SNAREs—engines for membrane fusion. *Nature Reviews Molecular Cell Biology* 7, 631–643.
- Kanethi, P., Qiao, X., Diaz, M.E., *et al.*, 1998. Mutation in AP-3 delta in the mocha mouse links endosomal transport to storage deficiency in platelets, melanosomes, and synaptic vesicles. *Neuron* 21, 111–122.
- Kent, H.M., Evans, P.R., Schafer, I.B., *et al.*, 2012. Structural basis of the intracellular sorting of the SNARE VAMP7 by the AP3 adaptor complex. *Developmental Cell* 22, 979–988.
- Kesavan, J., Borisovska, M., Bruns, D., 2007. v-SNARE actions during Ca^{2+} -triggered exocytosis. *Cell* 131, 351–363.
- Khwaja, S.S., Liu, H., Tong, C., *et al.*, 2010. HIV-1 Rev-binding protein accelerates cellular uptake of iron to drive Notch-induced T cell leukemogenesis in mice. *Journal of Clinical Investigation* 120, 2537–2548.
- Koo, S.J., Markovic, S., Puchkov, D., *et al.*, 2011. SNARE motif-mediated sorting of synaptobrevin by the endocytic adaptors clathrin assembly lymphoid myeloid leukemia (CALM) and AP180 at synapses. *Proceedings of the National Academy of Sciences of the United States of America* 108, 13540–13545.
- Kummel, D., Krishnakumar, S.S., Radoff, D.T., *et al.*, 2011. Complexin cross-links peripheral SNAREs into a zigzag array. *Nature Structural & Molecular Biology* 18, 927–933.
- Lentz, B.R., 2007. PEG as a tool to gain insight into membrane fusion. *European Biophysics Journal* 36, 315–326.
- Li, F., Pincet, F., Perez, E., *et al.*, 2007. Energetics and dynamics of SNAREpin folding across lipid bilayers. *Nature Structural & Molecular Biology* 14, 890–896.
- Li, F., Pincet, F., Perez, E., *et al.*, 2011. Complexin activates and clamps SNAREpins by a common mechanism involving an intermediate energetic state. *Nature Structural & Molecular Biology* 18, 941–946.
- Manford, A.G., Stefan, C.J., Yuan, H.L., Macgurn, J.A., Emr, S.D., 2012. ER-to-plasma membrane tethering proteins regulate cell signaling and ER morphology. *Developmental Cell* 23, 1129–1140.
- Martinez-Arca, S., Alberts, P., Zahraoui, A., Louvard, D., Galli, T., 2000. Role of tetanus neurotoxin insensitive vesicle-associated membrane protein (TI-VAMP) in vesicular transport mediating neurite outgrowth. *Journal of Cell Biology* 149, 889–899.
- Martinez-Arca, S., Rudge, R., Vacca, M., *et al.*, 2003. A dual mechanism controlling the localization and function of exocytic v-SNAREs. *Proceedings of the National Academy of Sciences of the United States of America* 100, 9011–9016.
- McMahon, H.T., Kozlov, M.M., Martens, S., 2010. Membrane curvature in synaptic vesicle fusion and beyond. *Cell* 140, 601–605.
- McNew, J.A., Weber, T., Engelman, D.M., Sollner, T.H., Rothman, J.E., 1999. The length of the flexible SNAREpin juxtamembrane region is a critical determinant of SNARE-dependent fusion. *Molecular Cell* 4, 415–421.
- McNew, J.A., Weber, T., Parlati, F., *et al.*, 2000. Close is not enough: SNARE-dependent membrane fusion requires an active mechanism that transduces force to membrane anchors. *Journal of Cell Biology* 150, 105–117.

- Meghian, A., Zordan, M., Pantano, S., *et al.*, 2013. Evidence for a radial SNARE super-complex mediating neurotransmitter release at the *Drosophila* neuromuscular junction. *Journal of Cell Science* 126, 3134–3140.
- Melia, T.J., Weber, T., McNew, J.A., *et al.*, 2002. Regulation of membrane fusion by the membrane-proximal coil of the t-SNARE during zippering of SNAREpins. *Journal of Cell Biology* 158, 929–940.
- Mellman, I., Emr, S.D., 2013. A Nobel Prize for membrane traffic: Vesicles find their journey's end. *Journal of Cell Biology* 203, 559–561.
- Miller, S.E., Sahlender, D.A., Graham, S.C., *et al.*, 2011. The molecular basis for the endocytosis of small R-SNAREs by the clathrin adaptor CALM. *Cell* 147, 1118–1131.
- Min, D., Kim, K., Hyeon, C., *et al.*, 2013. Mechanical unzipping and rezippering of a single SNARE complex reveals hysteresis as a force-generating mechanism. *Nature Communications* 4, 1705.
- Misura, K.M.S., Scheller, R.H., Weis, W.I., 2000. Three-dimensional structure of the neuronal-Sec1-syntaxin 1a complex. *Nature* 404, 355–362.
- Mohrmann, R., de Wit, H., Verhage, M., Neher, E., Sorensen, J.B., 2010. Fast vesicle fusion in living cells requires at least three SNARE complexes. *Science* 330, 502–505.
- Nair, U., Jotwani, A., Geng, J., *et al.*, 2011. SNARE proteins are required for macroautophagy. *Cell* 146, 290–302.
- Nobelprize.org, 2013. The 2013 Nobel Prize in Physiology or Medicine – Advanced Information. Nobelprize.org. Nobel Media AB 2013. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/2013/advanced.html (accessed 24.11.14).
- Nystuen, A.M., Schwendinger, J.K., Sachs, A.J., Yang, A.W., Haider, N.B., 2007. A null mutation in VAMP1/syntaxin is associated with neurological defects and prewean mortality in the lethal-wasting mouse mutant. *Neurogenetics* 8, 1–10.
- Peden, A.A., Park, G.Y., Scheller, R.H., 2001. The di-leucine motif of vesicle-associated membrane protein 4 is required for its localization and AP-1 binding. *Journal of Biological Chemistry* 276, 49183–49187.
- Petkovic, M., Jemaiel, A., Daste, F., *et al.*, 2014. The SNARE Sec22b has a non-fusogenic function in plasma membrane expansion. *Nature Cell Biology* 16, 434–444.
- Pichler, H., Gaigg, B., Hrastnik, C., *et al.*, 2001. A subfraction of the yeast endoplasmic reticulum associates with the plasma membrane and has a high capacity to synthesize lipids. *European Journal of Biochemistry* 268, 2351–2361.
- Porter, K.R., Palade, G.E., 1957. Studies on the endoplasmic reticulum. III. Its form and distribution in striated muscle cells. *Journal of Biophysical and Biochemical Cytology* 3, 269–300.
- Proux-Gillardeaux, V., Raposo, G., Irinopoulou, T., Galli, T., 2007. Expression of the Longin domain of TI-VAMP impairs lysosomal secretion and epithelial cell migration. *Biology of the Cell* 99, 261–271.
- Pryor, P.R., Jackson, L., Gray, S.R., *et al.*, 2008. Molecular basis for the sorting of the SNARE VAMP7 into endocytic clathrin-coated vesicles by the ArfGAP Hrb. *Cell* 134, 817–827.
- Reim, K., Mansour, M., Varoqueaux, F., *et al.*, 2001. Complexins regulate a late step in Ca²⁺-dependent neurotransmitter release. *Cell* 104, 71–81.
- Rizo, J., Chen, X., Arac, D., 2006. Unraveling the mechanisms of syntaxin and SNARE function in neurotransmitter release. *Trends in Cell Biology* 16, 339–350.
- Rosenbluth, J., 1962. Subsurface cisterns and their relationship to the neuronal plasma membrane. *Journal of Cell Biology* 13, 405–421.
- Salazar, G., Craige, B., Styers, M.L., *et al.*, 2006. BLOC-1 complex deficiency alters the targeting of adaptor protein complex-3 cargoes. *Molecular Biology of the Cell* 17, 4014–4026.
- Schaefer, I.B., Hesketh, G.G., Bright, N.A., *et al.*, 2012. The binding of Varp to VAMP7 traps VAMP7 in a closed, fusogenically inactive conformation. *Nature Structural & Molecular Biology* 19, 1300–1309.
- Schaub, J.R., Lu, X., Doneske, B., Shin, Y.K., McNew, J.A., 2006. Hemifusion arrest by complexin is relieved by Ca(2+) -syntaxin I. *Nature Structural & Molecular Biology* 13, 748–750.
- Schneider, C.M., Wu, X., Saheki, Y., *et al.*, 2014. Structure of a lipid-bound extended syntaxin indicates a role in lipid transfer. *Nature* 510, 552–555.
- Schmid, E.M., McMahon, H.T., 2007. Integrating molecular and network biology to decode endocytosis. *Nature* 448, 883–888.
- Schoch, S., Deak, F., Königstorfer, A., *et al.*, 2001. SNARE function analyzed in syntaxin/syntaxin knockout mice. *Science* 294, 1117–1122.
- Shen, J., Tareste, D.C., Paumet, F., Rothman, J.E., Melia, T.J., 2007. Selective activation of cognate SNAREpins by Sec1/Munc18 proteins. *Cell* 128, 183–195.
- Shi, L., Shen, Q.T., Kiel, A., *et al.*, 2012. SNARE proteins: One to fuse and three to keep the nascent fusion pore open. *Science* 335, 1355–1359.
- Sinha, R., Ahmed, S., Jahn, R., Klingauf, J., 2011. Two syntaxin molecules are sufficient for vesicle fusion in central nervous system synapses. *Proceedings of the National Academy of Sciences of the United States of America* 108, 14318–14323.
- Sirkis, D.W., Edwards, R.H., Asensio, C.S., 2013. Widespread dysregulation of peptide hormone release in mice lacking adaptor protein AP-3. *PLoS Genetics* 9, e1003812.
- Söllner, T., Bennett, M.K., Whiteheart, S.W., Scheller, R.H., Rothman, J.E., 1993. A protein assembly-disassembly pathway *in vitro* that may correspond to sequential steps of synaptic vesicle docking, activation, and fusion. *Cell* 75, 409–418.
- Sorensen, J.B., Wiederhold, K., Muller, E.M., *et al.*, 2006. Sequential N- to C-terminal SNARE complex assembly drives priming and fusion of secretory vesicles. *EMBO Journal* 25, 955–966.
- Sotirakis, E., Galli, T., 2007. Exocytosis: Lessons from SNARE mutants and friends. In: Regazzi, R. (Ed.), *Molecular Mechanisms of Exocytosis*. Georgetown: Landes Bioscience and Springer Science + Business Media.
- Stefan, C.J., Manford, A.G., Baird, D., *et al.*, 2011. Osh proteins regulate phosphoinositide metabolism at ER-plasma membrane contact sites. *Cell* 144, 389–401.
- Sudhof, T.C., Rothman, J.E., 2009. Membrane fusion: Grappling with SNARE and SM proteins. *Science* 323, 474–477.
- Sutton, R.B., Fasshauer, D., Jahn, R., Brunger, A.T., 1998. Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 angstrom resolution. *Nature* 395, 347–353.
- Tang, J., Maximov, A., Shin, O.H., *et al.*, 2006. A complexin/syntaxin 1 switch controls fast synaptic vesicle exocytosis. *Cell* 126, 1175–1187.
- Tareste, D., Shen, J., Melia, T.J., Rothman, J.E., 2008. SNAREpin/Munc18 promotes adhesion and fusion of large vesicles to giant membranes. *Proceedings of the National Academy of Sciences of the United States of America* 105, 2380–2385.
- Toulmay, A., Prinz, W.A., 2011. Lipid transfer and signaling at organelle contact sites: The tip of the iceberg. *Current Opinion in Cell Biology* 23, 458–463.
- Verhage, M., Maia, A.S., Plomp, J.J., *et al.*, 2000. Synaptic assembly of the brain in the absence of neurotransmitter secretion. *Science* 287, 864–869.
- Wang, F., Zhang, H., Zhang, X., *et al.*, 2008. Varp interacts with Rab38 and functions as its potential effector. *Biochemical and Biophysical Research Communications* 372, 162–167.
- Weber, T., Zemelman, B.V., McNew, J.A., *et al.*, 1998. SNAREpins: Minimal machinery for membrane fusion. *Cell* 92, 759–772.
- Weimer, R.M., Richmond, J.E., Davis, W.S., *et al.*, 2003. Defects in synaptic vesicle docking in unc-18 mutants. *Nature Neuroscience* 6, 1023–1030.
- Xu, H., Jun, Y., Thompson, J., Yates, J., Wickner, W., 2010. HOPS prevents the disassembly of trans-SNARE complexes by Sec17p/Sec18p during membrane fusion. *EMBO Journal* 29, 1948–1960.
- Xu, H., Zick, M., Wickner, W.T., Jun, Y., 2011. A lipid-anchored SNARE supports membrane fusion. *Proceedings of the National Academy of Sciences of the United States of America* 108, 17325–17330.
- Zhang, X., He, X., Fu, X.Y., Chang, Z., 2006. Varp is a Rab21 guanine nucleotide exchange factor and regulates endosome dynamics. *Journal of Cell Science* 119, 1053–1062.
- Zhou, P., Bacaj, T., Yang, X., Pang, Z.P., Sudhof, T.C., 2013. Lipid-anchored SNAREs lacking transmembrane regions fully support membrane fusion during neurotransmitter release. *Neuron* 80, 470–483.

Further Reading

- Vicogne, J., Pessin, J., 2008. SNARE-mediated fusion of liposomes. In: Vancura, A. (Ed.), *Membrane Trafficking* 457. Totowa, NJ: Humana Press, pp. 241–251.
- Xu, D., Hay, J.C., 2009. *Intracellular membrane fusion. Trafficking Inside Cells: Pathways, Mechanisms and Regulation*. New York, NY: Landes Bioscience and Springer Science + Business Media, pp. 283–311.

Relevant Websites

- <http://www.youtube.com/watch?v=JMjurQ5i4PI>
James E Rothman Conference at NIH, November 14, 2012.
- <http://www.youtube.com/watch?v=zHewifC9Kc>
Randy Schekman iBio Seminar 'Synaptic Vesicle Fusion in Nerve Cells'.

ESCRTing around the Cell

PI Hanson and CE Jackson, Washington University School of Medicine, St. Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

History

Proteins that comprise the endosomal sorting complexes required for transport (ESCRT) machinery were identified in screens for vacuolar protein sorting (vps) mutants in the yeast *Saccharomyces cerevisiae* that share a common 'class E phenotype' in which newly synthesized vacuolar proteins as well as recycled Golgi components accumulate in a prominent 'class E compartment' built of stacked cisternae adjacent to an otherwise normal yeast vacuole (Raymond *et al.*, 1992). Endocytosed plasma membrane proteins also concentrate in this class E compartment (Davis *et al.*, 1993), identifying it as a central sorting station 'en route' to the vacuole. Additional studies demonstrated that class E gene function is required to deliver transmembrane cargo into the endosomal interior (Odorizzi *et al.*, 1998), connecting the class E compartment and failed multivesicular body (MVB) biogenesis.

Molecular analysis of class E gene products identified Vps4 as an AAA + ATPase (Babst *et al.*, 1997) that controls cycling of other class E proteins on and off the endosomal membrane (Babst *et al.*, 1998). Studies in the early 2000s established that most of the remaining class E gene products form heterooligomeric complexes that were named endosomal sorting complexes required for transport (ESCRTs), including ESCRT-0 (Bilodeau *et al.*, 2002; Katzmann *et al.*, 2003), ESCRT-I (Katzmann *et al.*, 2001), ESCRT-II (Babst *et al.*, 2002a), and ESCRT-III (Babst *et al.*, 2002b). ESCRT complexes and Vps4 are conserved across evolution (Babst, 2005; Leung *et al.*, 2008) and a large body of literature now supports a broad and general role for this machinery in endolysosomal sorting and function.

The discovery that ESCRTs play an essential role in the life cycle of the human immunodeficiency virus (HIV-1) and related retroviruses by enabling release from infected cells solidified the idea that ESCRTs are centrally involved in membrane scission when vesicles (or viruses) bud away from the cytoplasm (Garrus *et al.*, 2001). The finding that ESCRTs participate in cytokinesis and cell division even in Archaeobacteria that lack internal membranes lent additional support to this concept (Carlton and Martin-Serrano, 2007; Morita *et al.*, 2007; Samson *et al.*, 2008; Lindas *et al.*, 2008) and also suggested that cell division might be the primordial function of the ESCRT machinery. Successful reconstitution of ESCRT-driven vesicle formation in giant unilamellar vesicles in 2010 solidified a broad and still current framework for understanding ESCRT function (Wollert *et al.*, 2009a; Wollert and Hurley, 2010).

While the intervening years have seen advances in structural and functional analyses of the ESCRT machinery in many contexts and pathways, fundamental questions remain about how components of the ESCRT machinery – and especially ESCRT-III and Vps4 – operate to drive membrane scission. Models abound, but definitive answers await new approaches and systems. Recent excitement in the field comes from

expanded appreciation of non-endosomal roles for components of the ESCRT machinery. These support an emerging concept of ESCRT complexes as modular machinery that functions to remodel cellular membranes (Figure 1) in cooperation with other factors highlighted elsewhere in this volume.

ESCRT Machinery

ESCRT proteins and the complexes they form are soluble and associate transiently with membranes, similar to other membrane trafficking and remodeling factors. In their best known role – delivering cargo protein into the vacuole or lysosome – ESCRT complexes can be divided into those that cooperatively bind ubiquitin, phosphatidylinositol 3-phosphate, and each other to recruit cargo and initiate membrane deformation and vesicle formation on the endosomal membrane (ESCRT-0, -I, and -II) and those that complete the release of vesicles into the MVB (ESCRT-III and Vps4). The numbering of the ESCRT complexes is based on the order in which they are recruited for vacuolar protein sorting in yeast although, as will be apparent in the following discussion, these are modular complexes that function in different combinations depending on the system involved. A brief description of ESCRT complexes follows, and their protein components in well-studied model organisms are shown in Table 1. Excellent in-depth reviews discussing ESCRT complex structure and organization can be found in (Hurley, 2010; Henne *et al.*, 2011; Wollert *et al.*, 2009b; Hanson and Cashikar, 2012).

Early Acting ESCRTs

ESCRT complexes implicated in cargo recognition and early stages of intraluminal vesicle (ILV) formation include ESCRT-0 (composed of Vps27/Hrs and Hse1/STAM), ESCRT-I (composed of Vps23/Tsg101, Vps28, Mvb12, and Vps37), and ESCRT-II (composed of Vps36/EAP45, Snf8/EAP30, and Vps25/EAP20). Each of these complexes is a constitutive heterooligomeric complex and depletion of any single subunit destabilizes the whole complex. Important features of these early acting complexes include a number of low-affinity ubiquitin-binding motifs for cargo recognition (all three), phosphatidylinositol 3-phosphate-binding sites for endosomal targeting (ESCRT-0 and ESCRT-II), and specific motifs that connect the complexes to each other and in the case of ESCRT-II to ESCRT-III. Other factors able to engage and activate ESCRT-III can replace some or all of these early acting ESCRT complexes, expanding the repertoire of molecular pathways into the MVB. The best studied of these are Bro1-domain containing proteins such as Alix that (like the early acting ESCRT complexes) bind both ubiquitin and CHMP4 family ESCRT-III proteins (Katoh *et al.*, 2003; Dowlatshahi *et al.*, 2012; Pashkova *et al.*, 2013).

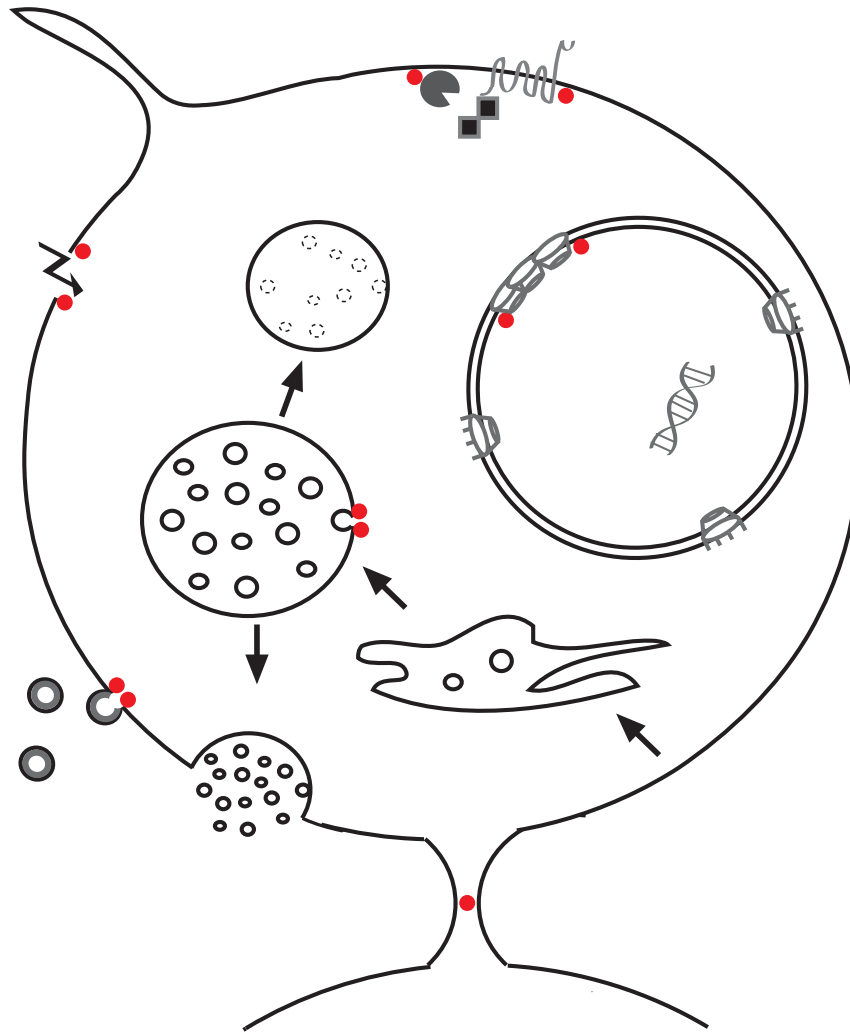


Figure 1 Sites of ESCRT function around the cell. Schematic drawing that highlights the many places around the cell where ESCRT machinery is thought to function. Red dots represent the ESCRT machinery. Well-studied roles in multivesicular body (MVB) biogenesis and virus budding are shown in detail, and highlight the unique role of this machinery in budding events directed away from the cytoplasm. Once formed, MVBs fuse with lysosomes for degradation and with the plasma membrane to release exosomes. At the midbody, ESCRTs separate dividing cells. Not shown is a comparable role in severing neurites during neuronal development. At the cell surface, ESCRTs organize signaling platforms for pH sensing and resolve membrane lesions (indicated by a lightning bolt). At the nuclear envelope, ESCRTs are implicated in nuclear pore complex quality control. Finally, ESCRTs are present in cilia, where details of their function have yet to be uncovered.

ESCRT-III and Vps4

ESCRT-III and its AAA ATPase Vps4 are the central players in ESCRT-driven membrane scission. The overall structure of individual ESCRT-III proteins (Muziol *et al.*, 2006; Bajorek *et al.*, 2009; Xiao *et al.*, 2009) and the filamentous assemblies they form (Hanson *et al.*, 2008; Cashikar *et al.*, 2014; Henne *et al.*, 2012) is established (Figure 2). A number of models describing how rings and spirals on the membrane might facilitate membrane remodeling and scission have been proposed and are discussed further below. There are 7 non-redundant ESCRT-III or ESCRT-III related proteins in the yeast *S. cerevisiae* and 12 in humans, and a varying but generally similar complement in other eukaryotes. All are predicted to share a core molecular structure in which a helical hairpin is surrounded and regulated by short C-terminal helices (Muziol

et al., 2006; Bajorek *et al.*, 2009). ESCRT-III proteins cycle between a closed or autoinhibited conformation in the cytoplasm and an open or activated conformation on the membrane that is primed for polymer assembly (Lin *et al.*, 2005; Zamborlini *et al.*, 2006; Shim *et al.*, 2007). ESCRT-II and Alix bind ESCRT-III proteins to initiate ESCRT-III polymer assembly. ESCRT-III then grows in a process for which mechanisms controlling composition and final size have yet to be defined. Assembled ESCRT-III polymers are intrinsically stable but because of Vps4-driven disassembly are short-lived in the cell (Cashikar *et al.*, 2014; Babst *et al.*, 2002b). Studies using RNAi to examine roles for individual ESCRT-III proteins support a general conclusion that particular ESCRT-III proteins are more or less essential in different settings. A general requirement for one protein related to Snf7/CHMP4 and one related to Vps2/CHMP2 appears to be conserved in most

Table 1 ESCRT protein nomenclature

ESCRT	<i>Saccharomyces cerevisiae</i>	Human	<i>Drosophila melanogaster</i> ^a	<i>Caenorhabditis elegans</i> ^a	<i>Arabidopsis thaliana</i> ^b
ESCRT-0	Vps27 Hse1	HRS STAM1/2	Hrs Stam	HGRS-1 STAM-1	
ESCRT-I	Vps23 Vps37 Vps28 Mvb12	Tsg101 Vps37A,B,C,D Vps28 Mvb12A,B	Tsg101 Vps37 Vps28	TSG101 VPS37 VPS28 MVB-12	VPS23A, 23B VPS37-1, 37-2 VPS28-1, 28-2
ESCRT-II	Vps36 Snf8/Vps22 Vps25	VPS36/EAP45 SNF8/EAP30 VPS25/EAP20	Vps36 Isn/Vps22 Vps25	VPS36 VPS22 VPS25	VPS36 VPS22 VPS25
ESCRT-III	Vps20 Snf7 Vps24 Vps2 Did2 Vps60 Ist1 Yji049w	CHMP6 CHMP4A, 4B, 4C CHMP3 CHMP2A, 2B CHMP1A, 1B CHMP5 IST1 CHMP7	Vps20 Shrb/Snf7 Vps24 Vps2 Chmp1 Vps60	VPS20 VPS32.1, VPS32.2 VPS24 VPS-2 DID2 VPS60	VPS20 SNF7.1, 7.2 VPS24-1, 24-2 VPS2-1, 2-2, 2-3 CHMP1A, 1B VPS60-1, 60-2
Vps4	Vps4 Vta1	VPS4A, 4B LIP5	Vps4 CG7967	VPS4 T23G11.7, CBG00308	VPS4 LIP5/EXT-like
Bro1 proteins	Bro1 Rim20	ALIX/AIP1 BROX HDPTP	ALiX	ALX-1 B0507.2 EGO-2	

^aRoxrud *et al.* (2010) and nomenclature from <http://flybase.org> and <http://wormbase.org>

^bRichardson and Mullen (2011) and Shahriari *et al.* (2011) and nomenclature from <https://arabidopsis.org>

membrane remodeling processes (Hanson *et al.*, 2009). The reader is referred to original descriptions of ESCRT-III polymers cited above and previous reviews (McCullough *et al.*, 2013; Peel *et al.*, 2011; Henne *et al.*, 2013) for additional details.

ESCRTs and Other Cellular Machinery

In addition to the interactions that connect ESCRT complexes to each other and to cargo, there are a wide range of other factors that bind to different ESCRT proteins. A list compiled in 2010 identified 31 interactors for ESCRT-0 components, 25 for ESCRT-I, 6 for ESCRT-II, and 12 for ESCRT-III (Roxrud *et al.*, 2010), and many more have emerged since that time. Interesting examples include connections between Hrs in ESCRT-0 and clathrin or retromer that are important for controlling entry into the degradative pathway (Raiborg *et al.*, 2002; Strohlic *et al.*, 2008), potential regulators of ESCRT-III assembly (CC2D1A is proposed to regulate incorporation of CHMP4 into ESCRT-III (Martinelli *et al.*, 2012)), and connections to the cytoskeleton (RILP bridges ESCRT-II to dynein, promoting late endosomal transport (Wang and Hong, 2006; Progida *et al.*, 2007)). Surveys of known and suspected protein–protein interactions involving ESCRTs suggest that many functional partners are waiting in the wings for analysis as insight into ESCRT function expands.

Models of ESCRT Function

ESCRT complexes are modular but interconnected entities that most clearly act at the MVB to recognize and concentrate cargo, facilitate membrane deformation to initiate vesicle formation, and finally constrict vesicle necks to drive membrane scission. The modular nature of the ESCRT machinery provides flexibility in how their function can be harnessed to accomplish a range of activities. Models for how ESCRTs operate in MVB biogenesis as well as the topologically related processes of viral particle release and cytokinesis are explored in a number of reviews written over the past few years (Hurley and Hanson, 2010; McCullough *et al.*, 2013; Hanson and Cashikar, 2012; Henne *et al.*, 2013).

Common to all ESCRT-dependent reactions is the need for membrane scission, and despite general agreement that ESCRT-III filaments are centrally involved much remains to be learned about how they actually facilitate this critical endpoint for ESCRT pathway function. Current thinking centers on ways in which ESCRT-III filaments with their known propensity to spiral inside membrane tubes or necks of varying diameter could promote membrane scission, either on their own or with assistance from the disassembling ATPase Vps4. Most models build on the general idea that ESCRT-III spirals create a dome that promotes scission when the attached membrane is drawn into a sufficiently narrow neck (Fabrikant *et al.*, 2009). Variations on this invoke a more dynamic role for ESCRT-III

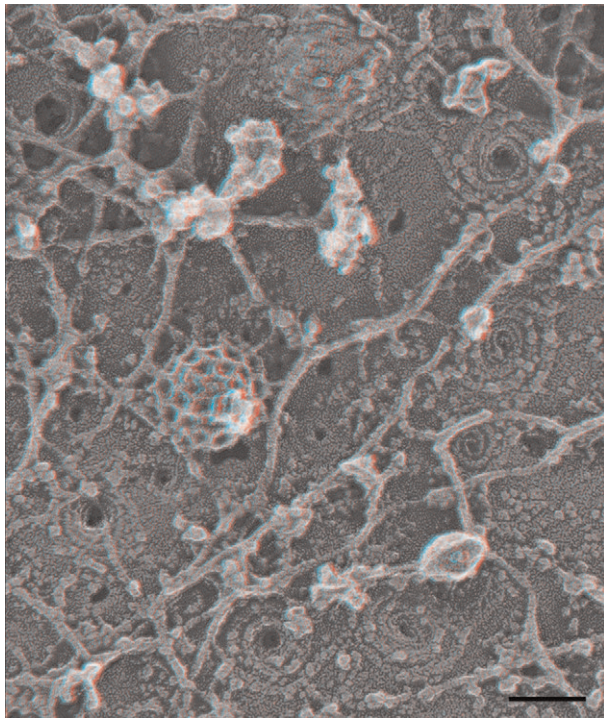


Figure 2 ESCRT-III filament spirals. Deep-etch electron micrograph showing endogenous ESCRT-III filament spirals on the plasma membrane of a HeLa cell treated with siRNA to deplete Vps4a and Vps4b. Scale bar represents 100 nm.

filaments in neck constriction and membrane scission, often in conjunction with Vps4. Dome-based models of membrane scission are supported by computational analyses confirming energetic plausibility (Fabrikant *et al.*, 2009; Elia *et al.*, 2012), but many aspects remain difficult to conceptualize and new approaches with greater spatial and temporal resolution are needed to clarify the driving principles responsible for this reaction and its regulation.

ESCRTs in the Endolysosomal System

ESCRTs are responsible for delivering membrane and membrane proteins into the endosomal interior for degradation in the lysosome or vacuole. This broad and essential role places them at the heart of the endolysosomal pathway, with control over the fate of membrane proteins that enter the endosomal system. Comprehensive insight into which ESCRTs are required for ILV biogenesis is available in *S. cerevisiae*, where mutation of components across the entire pathway including ESCRT-0, ESCRT-I, ESCRT-II, four 'core' ESCRT-III proteins, and Vps4 all lead to formation of multilamellar class E compartments that lack ILVs (Raymond *et al.*, 1992; Nickerson *et al.*, 2006). Importantly, temperature sensitive class E mutant strains reform ILVs upon return to permissive temperature (Russell *et al.*, 2012).

Ubiquitination is both necessary and sufficient for entry of membrane proteins into the MVB (Piper *et al.*, 2014) and is the trigger for initiating recruitment of ESCRTs to create an ILV.

In the typical pathway, ESCRT-0 is the first point of contact and in turn facilitates sequential engagement of downstream ESCRT complexes. A key transition point occurs upon recruitment of the Snf7 component of ESCRT-III. Mutations 'upstream' of Snf7 (in components of ESCRT-0, -I, and -II as well as Vps20) lead to enhanced recycling of proteins to the cell surface while mutations downstream of Snf7 do not (Teis *et al.*, 2010). This observation supports a working model in which upstream ESCRT complexes are important for collecting and concentrating ubiquitinated cargo that is then encircled by ESCRT-III filament(s) prior to release into the endosomal interior on an ILV.

A large body of work in other systems supports the general concept that the ESCRT machinery is essential for efficiently delivering membrane proteins to the lysosome for degradation (Rusten *et al.*, 2012). ESCRT mutations typically result in enlarged endosomes that contain few or no ILVs and do not successfully sort and degrade membrane proteins, including a wide variety of signaling receptors that then accumulate on endosomal membranes and continue to signal. One or another ESCRT component has been shown to be essential for downregulating receptor tyrosine kinases such as the epidermal growth factor receptor, a number of G protein coupled receptors, Wnt receptors, integrins, transforming growth factor beta (TGF- β) receptors, and more. Other membrane proteins such as gap junctions similarly depend on the ESCRT machinery for degradation (Leithe *et al.*, 2009). Interestingly, mutation or inactivation of individual ESCRT components is typically not lethal at the cellular level but leads instead to varying phenotypes that can usually be ascribed to dysregulated signaling. These can have dramatic effects at the organismal level, and in many cases are lethal during development. Overall, while it is clear that ESCRT function is key in the endosomal system, the complexity of phenotypes arising from perturbations in what is clearly an intricate and multistep set of events has made a full understanding of the ESCRT pathway and its connections with other machinery in the endolysosomal system challenging. There is clearly much yet to be learned.

ESCRTs in Virus Budding

HIV was the first virus shown to usurp the ESCRT machinery for budding and release from the plasma membrane of infected host cells (Garrus *et al.*, 2001; Martin-Serrano *et al.*, 2001; VerPlank *et al.*, 2001) in what has emerged as a general mechanism underlying release of most membrane-enclosed viruses (Votteler and Sundquist, 2013). In the case of HIV, the virally encoded Gag protein assembles on the plasma membrane to sequester the viral genome into an incomplete spherical particle. Viral Gag proteins use well-defined 'late domain' motifs to engage the ESCRTs and thereby release the virus. In the case of HIV Gag there are two late domain motifs, PT/SAP and LYPX_nL, that bind specific domains of Tsg101 (ESCRT-I) (Garrus *et al.*, 2001; Martin-Serrano *et al.*, 2001; VerPlank *et al.*, 2001) and Alix (Strack *et al.*, 2003), respectively. Different retroviruses use one or the other of these connections to the ESCRT machinery (Votteler and Sundquist, 2013) to directly or indirectly engage ESCRT-III and Vps4

(Baumgartel *et al.*, 2011; Jouvenet *et al.*, 2011). As in the case of ILV biogenesis, the prevailing view is that ESCRT-III filaments spiral on the cytosolic side of a narrowing neck to drive membrane scission (Bleck *et al.*, 2014; Ladinsky *et al.*, 2014; Cashikar *et al.*, 2014; Peel *et al.*, 2011; McCullough *et al.*, 2013). Other ideas including a proposal that ESCRT-III might instead assemble and act from within the viral particle (Van Engelenburg *et al.*, 2014) are also being considered. Open questions about the relationship between ESCRT-III, Vps4, and viral structural proteins such as HIV Gag include: (1) how is the extent of Gag assembly coordinated with ESCRT engagement to produce viral particles with an appropriate but incomplete (Carlson *et al.*, 2008) Gag shell?, (2) what molecule(s) connects Gag to ESCRT-III and how are these connections controlled?, (3) what role does Vps4 play and does it act at one, two, or more stages?, and finally (4) how do ESCRT-III filament spirals promote membrane scission?

ESCRTs in Cytokinesis

An evolutionarily ancient role of the ESCRT machinery is in cytokinesis. In organisms ranging from Archaea to mammalian cells, ESCRTs are important for successful abscission at the end of cytokinesis. Perhaps the clearest demonstration that their function is required is in certain Archaeobacteria that lack the typical fission-promoting GTPase FtsZ and instead use a cell division (cdv) system consisting of two ESCRT-III proteins and Vps4 (Bernander and Ettema, 2010). In the context of a multicellular organism, a recent study demonstrates that cytokinetic abscission during female germline stem cell division in *Drosophila* requires ESCRT proteins Shrub and Alix (Eikenes *et al.*, 2015).

In cultured mammalian cells where involvement of ESCRTs in cytokinesis has been most thoroughly explored, comprehensive siRNA screens defined the subset of ESCRT proteins required for successful cell division (Morita *et al.*, 2010; Carlton and Martin-Serrano, 2007). Interestingly, these include ESCRT-I, a subset of ESCRT-III and cooperating factors including Vps4 but not ESCRT-0 or ESCRT-II. This parallels requirements for virus release (Morita *et al.*, 2011) and may define a common set of scission promoting factors. Live cell imaging of fluorescently tagged proteins demonstrated that the centrosomal protein CEP55 arrives at the midbody early in cytokinesis and then recruits ALIX and ESCRT-I followed by two bursts of ESCRT-III (Elia *et al.*, 2011; Guizetti *et al.*, 2011). One proposal is that ESCRT-III filaments help to constrict the intercellular bridge by virtue of their decreasing spiral diameter and tight association with the membrane (Elia *et al.*, 2012). Vps4 may facilitate this reaction by freeing a sub-pool of ESCRT-III to generate an abscission-promoting spiral (Elia *et al.*, 2011). Many aspects of this process need further clarification, including the significance and function of ESCRT-III at two sites, the relationship between ~17 nm filaments seen by thin section electron microscopy (EM; Guizetti *et al.*, 2011) and ESCRT-III or other cytoskeletal components, the relative timing and role(s) of spastin versus Vps4 (Guizetti and Gerlich, 2012), and finally the role of additional proteins implicated in cytokinesis including MITD1, Brox, and others.

Recently described regulatory control of ESCRT-III ensures that abscission is delayed until chromatin segregation is complete. Termed the abscission checkpoint (NoCut), Aurora B kinase phosphorylates CHMP4C to delay abscission when chromatin lags within the intercellular bridge (Carlton *et al.*, 2012). Another protein, ANCHR (abscission/NoCut checkpoint regulator, previously known as ZFYVE19), also participates in this pathway, trapping Vps4 at the midbody together with CHMP4C in cells that contain chromatin bridges (Thoresen *et al.*, 2014). ANCHR may thus separate Vps4 from ESCRT-III in order to delay abscission. These studies provide strong support for the central role of the ESCRT machinery in cytokinesis.

ESCRTs in Extracellular Vesicle Biogenesis

Aside from delivery to lysosomes, MVBs can also fuse with the plasma membrane to release their contents from the cell. This phenomenon was first observed by EM in studies of transferrin receptor trafficking in maturing reticulocytes (Harding *et al.*, 1983). Vesicles released into the cell culture medium were termed 'exosomes' (Pan and Johnstone, 1983). Virtually all cells are now thought to produce exosomes, and these vesicles comprise one class of the broad collective of extracellular vesicles (EVs) (Raposo and Stoorvogel, 2013). An increasing body of evidence points to the ubiquity of EV release and implicates EVs as signaling units capable of eliciting phenotypic responses in cells with which they come into contact, as discussed elsewhere in this volume.

MVBs thus have two potential fates: degradation or secretion. While degradation predominates in most situations, it is not known what controls the fate of individual MVBs. Furthermore, MVB fate may change in response to cellular or environmental cues. There are indications that heterogeneity among MVBs may correlate with their eventual fate; for example, lysobisphosphatidic acid is concentrated in ILVs targeted for degradation (Matsuo *et al.*, 2004) but is largely absent from exosomes (Kowal *et al.*, 2014; Trajkovic *et al.*, 2008). Regulated recruitment of exocytosis-enabling Rab GTPases such as Rab27a (Ostrowski *et al.*, 2010), Rab35 (Hsu *et al.*, 2010), or Rab11 (Savina *et al.*, 2002) and their effectors is likely to be one factor correlated with secretion.

How involved the ESCRT machinery is in exosome (and more broadly EV) biogenesis is still an open question.

Because of the topology of EV formation (budding away from the cytoplasm), an early and logical hypothesis was that EV formation would utilize the same machinery responsible for creating ILVs within MVBs 'en route' to degradation, i.e., the ESCRTs. Proteomic studies abound in EV research, and many such studies identify ESCRT proteins as components of the EV proteome (Simpson *et al.*, 2009; Choi *et al.*, 2013). In fact, Tsg101 (ESCRT-I) and Alix are among the most common markers used to identify EVs (Kowal *et al.*, 2014). Despite interest in the role of the ESCRT machinery in EV biogenesis, it has remained unclear if ESCRT function is generally required for EV biogenesis. Seemingly straightforward experimental tests in different systems provide contradictory results (Trajkovic *et al.*, 2008; van Niel *et al.*, 2011; Baietti *et al.*, 2012; Colombo *et al.*, 2013) leading to the idea that at least some

EVs may be created without the help of ESCRTs (Babst, 2011; Buschow *et al.*, 2009). However, while lipids clearly affect EV biogenesis (Trajkovic *et al.*, 2008), ESCRT-independent pathways for controlling EV formation have not emerged. In fact, proteins recently implicated in EV biogenesis have each revealed new connections to the ESCRT machinery. In one example, the adaptor protein syntenin-1 binds syndecans and other transmembrane proteins and promotes their release in exosomes by recruiting Alix and ESCRT-III (Baietti *et al.*, 2012; Hurley and Odorizzi, 2012). In another example, the arrestin domain protein ARRDC1 interacts with Tsg101 and ubiquitinated proteins at the plasma membrane to stimulate EV formation and release (Nabhan *et al.*, 2012). A number of reports indicate that perturbing ESCRT function can affect EV number, content, or size (Colombo *et al.*, 2013; Baietti *et al.*, 2012). EM of artificial immunologic synapses showed that dominant-negative Vps4 inhibits release of ectosomes at the plasma membrane (Choudhuri *et al.*, 2014). Our finding that ESCRT spirals accumulate on the plasma membrane in Vps4 depleted cells suggests the possibility of similar release from other cultured cells (Cashikar *et al.*, 2014). Further understanding of how the ESCRT machinery contributes to generating EVs will aid in developing tools to manipulate EV biogenesis.

Non-Endosomal Roles for the ESCRT Machinery

It has been clear for some time that ESCRTs have non-endosomal roles that extend beyond viral budding and cytokinesis. Functions ranging from cell cycle control to gene silencing to cell polarity were highlighted in a 2006 review of the ESCRT machinery (Slagsvold *et al.*, 2006), but because the underlying mechanisms were not defined, these observations did not change the consensus view that ESCRTs operate primarily to drive membrane budding and scission at the MVB. Findings over the past few years highlighting a new set of non-endosomal functions for the ESCRTs – discussed below – are likely to rejuvenate interest in a broader consideration of ESCRT function throughout the cell.

pH Sensing

Sensing and responding to the surrounding environment is an ancient requirement for cell survival. A prime example is the conserved and well-studied alkaline pH-sensing pathway in fungi, which is important for both viability and virulence (Peñalva *et al.*, 2008; Davis, 2009). A receptor at the cell surface senses alkaline extracellular pH (Negrete-Urtasun *et al.*, 1999; Herranz *et al.*, 2005) and induces proteolytic activation of a transcription factor, which in turn increases expression of genes important for maintaining the proton gradient, promoting nutrient absorption and iron homeostasis (Davis, 2009; Peñalva *et al.*, 2008). Mutational analyses established the necessity of ESCRTs for pathway activation (Calcagno-Pizarelli *et al.*, 2011). Early efforts to connect ESCRT function in pH sensing to the endosomal network have given way to experiments demonstrating that ESCRTs operate at the plasma membrane in this pathway. Imaging studies show that ESCRTs localize to the cell cortex, rather than endosomes, for function

in pH sensing (Galindo *et al.*, 2007, 2012). One proposal is that ESCRTs function in pH signaling by corraling signaling components at the plasma membrane to bring the responsible calpain protease into contact with its substrate to release the activating transcription factor (Peñalva *et al.*, 2014). This highlights what may be a general role for membrane-associated ESCRT-III filaments in creating functional domains along the membrane.

Plasma Membrane Resealing

All cells possess mechanisms to defend against loss of plasma membrane integrity. Plasma membrane repair is triggered by Ca^{2+} influx (Idone *et al.*, 2008) and has been connected with lysosomal exocytosis (Reddy *et al.*, 2001), endocytosis of damaged domains (Corrotte *et al.*, 2012), and most recently ESCRT-III mediated repair potentially via shedding of damaged membrane from the cell surface (Jimenez *et al.*, 2014). This novel role for ESCRTs emerged from the observation that Alix and ESCRT-III (but not ESCRT-0, -I, or -II) are recruited to lesions generated by exposing cells to detergent, pore forming toxins, or laser ablation (Jimenez *et al.*, 2014). Experiments using RNAi to deplete ESCRT-III proteins further demonstrated that these are required for cell viability after damage (Jimenez *et al.*, 2014). This process is modeled to involve the shedding of membrane vesicles from the cell surface (Jimenez *et al.*, 2014). Direct constriction by spiraling ESCRT-III filaments could also drive resealing (Andrews *et al.*, 2014).

Neurite Severing in Neuronal Development

A recent report establishes that ESCRTs mediate membrane scission during developmental pruning of axons and dendrites in *Drosophila* (Loncle *et al.*, 2015). As in cytokinesis, a subset of ESCRT proteins is required near the site of scission, in this case for efficient pruning or severing of dendrites. Critical among the required ESCRT proteins are *Drosophila* Snf7/CHMP4 (known as shrub) and its interacting Bro1-domain protein myopic (HDPTP in humans). The reaction appears to parallel that of abscission.

Activity at the Nuclear Envelope

Perhaps the least expected non-endosomal role for ESCRTs emerged last year with the description of a new and surprising role for ESCRT-III at the nuclear envelope in yeast. Mutational analyses revealed genetic interactions between genes involved in nuclear pore complex (NPC) assembly and a subset of ESCRTs (Webster *et al.*, 2014). Inactivating ESCRTs caused incomplete NPC clusters to accumulate and led to the idea that ESCRTs function in removing misassembled NPC oligomers. The lack of genetic interactions between NPC genes and those encoding ESCRT-0, -I, -II indicate that these complexes are not involved. The ESCRT-III protein Snf7 binds directly to Heh2 (related to emerin) at the inner nuclear membrane providing insight into how ESCRT-III may be recruited. However, how ESCRT-III detects malformed NPCs is unknown. Intriguingly, genetic analyses of spindle pole body duplication in *Schizosaccharomyces pombe* (but not *S. cerevisiae*)

defined a clear role for Vps4, Vps32, and Vps24 in this nuclear envelope-based event (Frost *et al.*, 2012) further highlighting potential roles for ESCRT-III at the nuclear envelope.

Organization and Membrane Dynamics in Cilia

Primary cilia are sophisticated sensory structures that project from the plasma membrane to mediate a variety of signaling functions including mechanosensation, osmosensation, and many developmental signaling events. Studies over past decade have uncovered cytoskeletal and trafficking pathways responsible for biogenesis and maintenance of the primary cilium, and ESCRT proteins have emerged as candidates involved in different aspects of ciliary function. Cilia are microtubule-based projections present in at least a single copy on most human cells (Satir *et al.*, 2010). Cilia are conserved throughout evolution including in single-celled organisms such as the green algae *Chlamydomonas reinhardtii* that possess a single flagellar cilium (Kim and Dynlacht, 2013). Assembly and turnover of cilia is frequently studied in *Chlamydomonas* (Kim and Dynlacht, 2013). Disassembly of cilia during cell division requires vesicular encapsulation of the ciliary transition zone, a structure that regulates protein import and export into and from the cilium (Cavalier-Smith, 1974; Diener *et al.*, 2015). ESCRTs, Vps4, and Alix were recently identified in proteomic studies of purified transition zones (Diener *et al.*, 2015). This raises interesting questions about how ESCRT might function at this specialized compartment. Both functional and imaging studies of ESCRT during the course of the cilium life cycle will be important to answering these exciting new questions.

Summary and Perspectives

The combination of increasingly incisive experiments designed to explore the mechanisms by which the ESCRT machinery promotes membrane remodeling and the emergence of new roles for the ESCRTs all around the cell promises that the next several years will see continued growth and excitement in the study of this fascinating and elaborate machinery.

See also: Molecular Principles, Components, Technology, and Concepts: Membranes: Composition, Physical Properties, and Curvature. Organelles: Structure and Function: Extracellular Vesicles; The Late Endosome

References

- Andrews, N.W., Almeida, P.E., Corrotte, M., 2014. Damage control: Cellular mechanisms of plasma membrane repair. *Trends in Cell Biology* 24, 734–742.
- Babst, M., 2005. A protein's final ESCRT. *Traffic* 6, 2–9.
- Babst, M., 2011. MVB vesicle formation: ESCRT-dependent, ESCRT-independent and everything in between. *Current Opinion in Cell Biology* 23, 452–457.
- Babst, M., Katzmann, D., Snyder, W.B., Wendland, B., Emr, S.D., 2002a. Endosome-associated complex, ESCRT-II, recruits transport machinery for protein sorting at the multivesicular body. *Developmental Cell* 3, 283–289.
- Babst, M., Katzmann, D.J., Estepa-Sabal, E.J., Meerloo, T., Emr, S.D., 2002b. Escrt-III: An endosome-associated heterooligomeric protein complex required for mvb sorting. *Developmental Cell* 3, 271–282.
- Babst, M., Sato, T.K., Banta, L.M., Emr, S.D., 1997. Endosomal transport function in yeast requires a novel AAA-type ATPase, Vps4p. *EMBO Journal* 16, 1820–1831.
- Babst, M., Wendland, B., Estepa, E.J., Emr, S.D., 1998. The Vps4p AAA ATPase regulates membrane association of a Vps protein complex required for normal endosome function. *EMBO Journal* 17, 2982–2993.
- Baietti, M.F., Zhang, Z., Mortier, E., *et al.*, 2012. Syndecan-syntenin-ALIX regulates the biogenesis of exosomes. *Nature Cell Biology* 14, 677–685.
- Bajorek, M., Schubert, H.L., McCullough, J., *et al.*, 2009. Structural basis for ESCRT-III protein autoinhibition. *Nature Structural & Molecular Biology* 16, 754–762.
- Baumgartel, V., Ivanchenko, S., Dupont, A., *et al.*, 2011. Live-cell visualization of dynamics of HIV budding site interactions with an ESCRT component. *Nature Cell Biology* 13, 469–474.
- Bernander, R., Ettema, T.J., 2010. FtsZ-less cell division in archaea and bacteria. *Current Opinion in Microbiology* 13, 747–752.
- Bilodeau, P.S., Urbanowski, J.L., Winistorfer, S.C., Piper, R.C., 2002. The Vps27p Hse1p complex binds ubiquitin and mediates endosomal protein sorting. *Nature Cell Biology* 4, 534–539.
- Bleck, M., Itano, M.S., Johnson, D.S., *et al.*, 2014. Temporal and spatial organization of ESCRT protein recruitment during HIV-1 budding. *Proceedings of the National Academy of Sciences of the United States of America* 111, 12211–12216.
- Buschow, S.I., Nolte-t Hoen, E.N., van Niel, G., *et al.*, 2009. MHC II in dendritic cells is targeted to lysosomes or T cell-induced exosomes via distinct multivesicular body pathways. *Traffic* 10, 1528–1542.
- Calcagno-Pizarelli, A.M., Hervas-Aguilar, A., Galindo, A., *et al.*, 2011. Rescue of *Aspergillus nidulans* severely debilitating null mutations in ESCRT-0, I, II and III genes by inactivation of a salt-tolerance pathway allows examination of ESCRT gene roles in pH signalling. *Journal of Cell Science* 124, 4064–4076.
- Carlson, L.A., Briggs, J.A., Glass, B., *et al.*, 2008. Three-dimensional analysis of budding sites and released virus suggests a revised model for HIV-1 morphogenesis. *Cell Host & Microbe* 4, 592–599.
- Carlton, J.G., Caballe, A., Agromayor, M., Kloc, M., Martin-Serrano, J., 2012. ESCRT-III governs the Aurora B-mediated abscission checkpoint through CHMP4C. *Science* 336, 220–225.
- Carlton, J.G., Martin-Serrano, J., 2007. Parallels between cytokinesis and retroviral budding: A role for the ESCRT machinery. *Science* 316, 1908–1912.
- Cashikar, A.G., Shim, S., Roth, R., *et al.*, 2014. Structure of cellular ESCRT-III spirals and their relationship to HIV budding. *eLife*. e02184.
- Cavalier-Smith, T., 1974. Basal body and flagellar development during the vegetative cell cycle and the sexual cycle of *Chlamydomonas reinhardtii*. *Journal of Cell Science* 16, 529–556.
- Choi, D.-S., Kim, D.-K., Kim, Y.-K., Gho, Y.S., 2013. Proteomics, transcriptomics and lipidomics of exosomes and ectosomes. *Proteomics* 13, 1554–1571.
- Choudhuri, K., Llodrá, J., Roth, E.W., *et al.*, 2014. Polarized release of T-cell-receptor-enriched microvesicles at the immunological synapse. *Nature* 507, 118–123.
- Colombo, M., Moita, C., van Niel, G., *et al.*, 2013. Analysis of ESCRT functions in exosome biogenesis, composition and secretion highlights the heterogeneity of extracellular vesicles. *Journal of Cell Science* 126, 5553–5565.
- Corrotte, M., Fernandes, M.C., Tam, C., Andrews, N.W., 2012. Toxin pores endocytosed during plasma membrane repair traffic into the lumen of MVBs for degradation. *Traffic* 13, 483–494.
- Davis, D.A., 2009. How human pathogenic fungi sense and adapt to pH: The link to virulence. *Current Opinion in Microbiology* 12, 365–370.
- Davis, N.G., Horecka, J.L., Sprague, G.F., 1993. Cis- and trans-acting functions required for endocytosis of the yeast pheromone receptors. *Journal of Cell Biology* 122, 53–65.
- Diener, D.R., Lupetti, P., Rosenbaum, J.L., 2015. Proteomic analysis of isolated ciliary transition zones reveals the presence of ESCRT proteins. *Current Biology* 25, 379–384.
- Dowlatabadi, D.P., Sandrin, V., Vivona, S., *et al.*, 2012. ALIX is a Lys63-specific polyubiquitin binding protein that functions in retrovirus budding. *Developmental Cell* 23, 1247–1254.
- Eikenes, Å.H., Malerød, L., Christensen, A.L., *et al.*, 2015. ALIX and ESCRT-III coordinately control cytokinetic abscission during germline stem cell division *in vivo*. *PLoS Genetics* 11, e1004904.
- Elia, N., Fabrikant, G., Kozlov, M.M., Lippincott-Schwartz, J., 2012. Computational model of cytokinetic abscission driven by ESCRT-III polymerization and remodeling. *Biophysical Journal* 102, 2309–2320.
- Elia, N., Sougrat, R., Spurlin, T.A., Hurley, J.H., Lippincott-Schwartz, J., 2011. Dynamics of endosomal sorting complex required for transport (ESCRT) machinery during cytokinesis and its role in abscission. *Proceedings of*

- the National Academy of Sciences of the United States of America 108, 4846–4851.
- Fabrikant, G., Lata, S., Riches, J.D., *et al.*, 2009. Computational model of membrane fission catalyzed by ESCRT-III. *PLOS Computational Biology* 5, e1000575.
- Frost, A., Elgort, M.G., Brandman, O., *et al.*, 2012. Functional repurposing revealed by comparing *S. pombe* and *S. cerevisiae* genetic interactions. *Cell* 149, 1339–1352.
- Galindo, A., Calcagno-Pizarelli, A.M., Arst, H.N.J., Penalva, M.A., 2012. An ordered pathway for the assembly of fungal ESCRT-containing ambient pH signalling complexes at the plasma membrane. *Journal of Cell Science* 125, 1784–1795.
- Galindo, A., Hervás-Aguilar, A., Rodríguez-Galan, O., *et al.*, 2007. PalC, one of two Bro1 domain proteins in the fungal pH signalling pathway, localizes to cortical structures and binds Vps32. *Traffic* 8, 1346–1364.
- Garrus, J.E., von Schwedler, U.K., Pornillos, O., *et al.*, 2001. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* 107, 55–65.
- Guizetti, J., Gerlich, D.W., 2012. ESCRT-III polymers in membrane neck constriction. *Trends in Cell Biology* 22 (3), 133–140.
- Guizetti, J., Schermelleh, L., Mantler, J., *et al.*, 2011. Cortical constriction during abscission involves helices of ESCRT-III-dependent filaments. *Science* 331, 1616–1620.
- Hanson, P.I., Cashikar, A., 2012. Multivesicular body morphogenesis. *Annual Review of Cell and Developmental Biology* 28, 337–362.
- Hanson, P.I., Roth, R., Lin, Y., Heuser, J.E., 2008. Plasma membrane deformation by circular arrays of ESCRT-III protein filaments. *Journal of Cell Biology* 180, 389–402.
- Hanson, P.I., Shim, S., Merrill, S.A., 2009. Cell biology of the ESCRT machinery. *Current Opinion in Cell Biology* 21, 568–574.
- Harding, C., Heuser, J., Stahl, P., 1983. Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes. *Journal of Cell Biology* 97, 329–339.
- Henne, W.M., Buchkovich, N.J., Emr, S.D., 2011. The ESCRT pathway. *Developmental Cell* 21, 77–91.
- Henne, W.M., Buchkovich, N.J., Zhao, Y., Emr, S.D., 2012. The endosomal sorting complex ESCRT-II mediates the assembly and architecture of ESCRT-III helices. *Cell* 151, 356–371.
- Henne, W.M., Stenmark, H., Emr, S.D., 2013. Molecular mechanisms of the membrane sculpting ESCRT pathway. *Cold Spring Harbor Perspectives in Biology* 5, a016766.
- Herranz, S., Rodríguez, J.M., Bussink, H.-J., *et al.*, 2005. Arrestin-related proteins mediate pH signaling in fungi. *Proceedings of the National Academy of Sciences of the United States of America* 102, 12141–12146.
- Hsu, C., Morohashi, Y., Yoshimura, S., *et al.*, 2010. Regulation of exosome secretion by Rab35 and its GTPase-activating proteins TBC1D10A-C. *Journal of Cell Biology* 189, 223–232.
- Hurley, J.H., 2010. The ESCRT complexes. *Critical Reviews in Biochemistry and Molecular Biology* 45, 463–487.
- Hurley, J.H., Hanson, P.I., 2010. Membrane budding and scission by the ESCRT machinery: It's all in the neck. *Nature Reviews Molecular Cell Biology* 11, 556–566.
- Hurley, J.H., Odorizzi, G., 2012. Get on the exosome bus with ALIX. *Nature Cell Biology* 14, 654–655.
- Idone, V., Tam, C., Goss, J.W., *et al.*, 2008. Repair of injured plasma membrane by rapid Ca^{2+} -dependent endocytosis. *Journal of Cell Biology* 180, 905–914.
- Jimenez, A.J., Maiuri, P., Lafaurie-Janvore, J., *et al.*, 2014. ESCRT machinery is required for plasma membrane repair. *Science* 343, 1247136.
- Jouvenet, N., Zhadina, M., Bieniasz, P.D., Simon, S.M., 2011. Dynamics of ESCRT protein recruitment during retroviral assembly. *Nature Cell Biology* 13, 394–401.
- Katoh, K., Shibata, H., Suzuki, H., *et al.*, 2003. The ALG-2-interacting protein Alix associates with CHMP4b, a human homologue of yeast Snf7 that is involved in multivesicular body sorting. *Journal of Biological Chemistry* 278, 39104–39113.
- Katzmann, D., Babst, M., Emr, S.D., 2001. Ubiquitin-dependent sorting into the multivesicular body pathway requires the function of a conserved endosomal protein sorting complex ESCRT-I. *Cell* 106, 145–155.
- Katzmann, D.J., Stefan, C.J., Babst, M., Emr, S.D., 2003. Vps27 recruits ESCRT machinery to endosomes during MVB sorting. *Journal of Cell Biology* 162, 413–423.
- Kim, S., Dynlacht, B.D., 2013. Assembling a primary cilium. *Current Opinion in Cell Biology* 25, 506–511.
- Kowal, J., Tkach, M., Théry, C., 2014. Biogenesis and secretion of exosomes. *Current Opinion in Cell Biology* 29, 116–125.
- Ladinsky, M.S., Kieffer, C., Olson, G., *et al.*, 2014. Electron tomography of HIV-1 infection in gut-associated lymphoid tissue. *PLOS Pathogens* 10, e1003899.
- Leithe, E., Kjensteth, A., Simes, S., *et al.*, 2009. Ubiquitylation of the gap junction protein connexin-43 signals its trafficking from early endosomes to lysosomes in a process mediated by Hrs and Tsg101. *Journal of Cell Science* 122, 3883–3893.
- Leung, K.F., Dacks, J.B., Field, M.C., 2008. Evolution of the multivesicular body ESCRT machinery; retention across the eukaryotic lineage. *Traffic* 9, 1698–1716.
- Lin, Y., Kimpler, L.A., Naismith, T.V., Lauer, J.M., Hanson, P.I., 2005. Interaction of the mammalian endosomal sorting complex required for transport (ESCRT) III protein hSnf7-1 with itself, membranes, and the AAA+ ATPase SKD1. *Journal of Biological Chemistry* 280, 12799–12809.
- Lindas, A.C., Karlsson, E.A., Lindgren, M.T., Ettema, T.J., Bernander, R., 2008. A unique cell division machinery in the Archaea. *Proceedings of the National Academy of Sciences of the United States of America* 105, 18942–18946.
- Loncle, N., Agromayor, M., Martin-Serrano, J., Williams, D.W., 2015. An ESCRT module is required for neuron pruning. *Scientific Reports* 5, 8461.
- Martin-Serrano, J., Zang, T., Bieniasz, P.D., 2001. HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. *Nature Medicine* 7, 1313–1319.
- Martinelli, N., Hartlieb, B., Usami, Y., *et al.*, 2012. CC2D1A is a Regulator of ESCRT-III CHMP4B. *Journal of Molecular Biology* 419, 75–88.
- Matsuo, H., Chevallier, J., Mayran, N., *et al.*, 2004. Role of LBP1 and Alix in multivesicular liposome formation and endosome organization. *Science* 303, 531–534.
- McCullough, J., Colf, L.A., Sundquist, W.I., 2013. Membrane fission reactions of the mammalian ESCRT pathway. *Annual Review of Biochemistry* 82, 663–692.
- Morita, E., Colf, L.A., Karren, M.A., *et al.*, 2010. Human ESCRT-III and VPS4 proteins are required for centrosome and spindle maintenance. *Proceedings of the National Academy of Sciences of the United States of America* 107, 12889–12894.
- Morita, E., Sandrin, V., Chung, H.Y., *et al.*, 2007. Human ESCRT and ALIX proteins interact with proteins of the midbody and function in cytokinesis. *EMBO Journal* 26, 4215–4227.
- Morita, E., Sandrin, V., McCullough, J., *et al.*, 2011. ESCRT-III protein requirements for HIV-1 budding. *Cell Host & Microbe* 9, 235–242.
- Muziol, T., Pineda-Molina, E., Ravelli, R.B., *et al.*, 2006. Structural basis for budding by the ESCRT-III factor CHMP3. *Developmental Cell* 10, 821–830.
- Nabhan, J.F., Hu, R., Oh, R.S., Cohen, S.N., Lu, Q., 2012. Formation and release of arrestin domain-containing protein 1-mediated microvesicles (ARMMs) at plasma membrane by recruitment of TSG101 protein. *Proceedings of the National Academy of Sciences of the United States of America* 109, 4146–4151.
- Negrete-Urtasun, S., Reiter, W., Diez, E., *et al.*, 1999. Ambient pH signal transduction in *Aspergillus*: Completion of gene characterization. *Molecular Microbiology* 33, 994–1003.
- Nickerson, D.P., West, M., Odorizzi, G., 2006. Did2 coordinates Vps4-mediated dissociation of ESCRT-III from endosomes. *Journal of Cell Biology* 175, 715–720.
- van Niel, G., Charrin, S., Simoes, S., *et al.*, 2011. The tetraspanin CD63 regulates ESCRT-independent and -dependent endosomal sorting during melanogenesis. *Developmental Cell* 21, 708–721.
- Odorizzi, G., Babst, M., Emr, S.D., 1998. Fab1p PtdIns(3)P 5-kinase function essential for protein sorting in the multivesicular body. *Cell* 95, 847–858.
- Ostrowski, M., Carmo, N.B., Krumeich, S., *et al.*, 2010. Rab27a and Rab27b control different steps of the exosome secretion pathway. *Nature Cell Biology* 12, 19–30.
- Pan, B.T., Johnstone, R.M., 1983. Fate of the transferrin receptor during maturation of sheep reticulocytes *in vitro*: Selective externalization of the receptor. *Cell* 33, 967–978.
- Pashkova, N., Gakhar, L., Winistorfer, S.C., *et al.*, 2013. The yeast alix homolog Bro1 functions as a ubiquitin receptor for protein sorting into multivesicular endosomes. *Developmental Cell* 25, 520–533.
- Peel, S., Macheboeuf, P., Martinelli, N., Weissenhorn, W., 2011. Divergent pathways lead to ESCRT-III-catalyzed membrane fission. *Trends in Biochemical Sciences* 36, 199–210.
- Peñalva, M.A., Lucena-Agell, D., Arst, H.N., 2014. Liaison alcaline: Pals entice non-endosomal ESCRTs to the plasma membrane for pH signaling. *Current Opinion in Microbiology* 22C, 49–59.
- Peñalva, M.A., Tilburn, J., Bignell, E., Arst, H.N., 2008. Ambient pH gene regulation in fungi: Making connections. *Trends in Microbiology* 16, 291–300.
- Piper, R.C., Dikic, I., Lukacs, G.L., 2014. Ubiquitin-dependent sorting in endocytosis. *Cold Spring Harbor Perspectives in Biology* 6, a016808.
- Progidia, C., Malerod, L., Stuffers, S., *et al.*, 2007. RILP is required for the proper morphology and function of late endosomes. *Journal of Cell Science* 120, 3729–3737.

- Raiborg, C., Bache, K.G., Gillooly, D.J., *et al.*, 2002. Hrs sorts ubiquitinated proteins into clathrin-coated microdomains of early endosomes. *Nature Cell Biology* 4, 394–398.
- Raposo, G., Stoorvogel, W., 2013. Extracellular vesicles: Exosomes, microvesicles, and friends. *Journal of Cell Biology* 200, 373–383.
- Raymond, C.K., Howald-Stevenson, I., Vater, C.A., Stevens, T.H., 1992. Morphological classification of the yeast vacuolar protein sorting mutants: Evidence for a prevacuolar compartment in class E vps mutants. *Molecular Biology of the Cell* 3, 1389–1402.
- Reddy, A., Caler, E.V., Andrews, N.W., 2001. Plasma membrane repair is mediated by Ca(2+)-regulated exocytosis of lysosomes. *Cell* 106, 157–169.
- Richardson, L.G.L., Mullen, R.T., 2011. Meta-analysis of the expression profiles of the *Arabidopsis* ESCRT machinery. *Plant Signaling and Behavior* 6, 1897–1903.
- Roxrud, L., Stenmark, H., Malerod, L., 2010. ESCRT & Co. *Biology of the Cell* 102, 293–318.
- Russell, M.R.G., Shideler, T., Nickerson, D.P., West, M., Odorizzi, G., 2012. Class E compartments form in response to ESCRT dysfunction in yeast due to hyperactivity of the Vps21 Rab GTPase. *Journal of Cell Science* 125, 5208–5220.
- Rusten, T.E., Vaccari, T., Stenmark, H., 2012. Shaping development with ESCRTs. *Nature Cell Biology* 14, 38–45.
- Samson, R.Y., Obita, T., Freund, S.M., Williams, R.L., Bell, S.D., 2008. A role for the ESCRT system in cell division in Archaea. *Science* 322, 1710–1713.
- Satir, P., Pedersen, L.B., Christensen, S.T., 2010. The primary cilium at a glance. *Journal of Cell Science* 123, 499–503.
- Savina, A., Vidal, M., Colombo, M.I., 2002. The exosome pathway in K562 cells is regulated by Rab11. *Journal of Cell Science* 115, 2505–2515.
- Shahriari, M., Richter, K., Keshavaiah, C., *et al.*, 2011. The *Arabidopsis* ESCRT protein–protein interaction network. *Plant Molecular Biology* 76, 85–96.
- Shim, S., Kimpler, L.A., Hanson, P.J., 2007. Structure/function analysis of four core ESCRT-III proteins reveals common regulatory role for extreme C-terminal domain. *Traffic* 8, 1068–1079.
- Simpson, R.J., Lim, J.W., Moritz, R.L., Mathivanan, S., 2009. Exosomes: Proteomic insights and diagnostic potential. *Expert Review of Proteomics* 6, 267–283.
- Slagsvold, T., Pattai, K., Malerod, L., Stenmark, H., 2006. Endosomal and non-endosomal functions of ESCRT proteins. *Trends in Cell Biology* 16, 317–326.
- Strack, B., Calistri, A., Craig, S., Popova, E., Göttinger, H.G., 2003. AIP1/ALIX is a binding partner for HIV-1 p6 and EIAV p9 functioning in virus budding. *Cell* 114, 689–699.
- Strochlic, T.I., Schmiedekamp, B.C., Lee, J., Katzmann, D.J., Burd, C.G., 2008. Opposing activities of the Snx3-retromer complex and ESCRT proteins mediate regulated cargo sorting at a common endosome. *Molecular Biology of the Cell* 19, 4694–4706.
- Teis, D., Saksena, S., Judson, B.L., Emr, S.D., 2010. ESCRT-II coordinates the assembly of ESCRT-III filaments for cargo sorting and multivesicular body vesicle formation. *EMBO Journal* 29, 871–883.
- Thoresen, S.B., Campsteijn, C., Vietri, M., *et al.*, 2014. ANCHR mediates Aurora-B-dependent abscission checkpoint control through retention of VPS4. *Nature Cell Biology* 16, 550–560.
- Trajkovic, K., Hsu, C., Chiantia, S., *et al.*, 2008. Ceramide triggers budding of exosome vesicles into multivesicular endosomes. *Science* 319, 1244–1247.
- Van Engelenburg, S.B., Shtengel, G., Sengupta, P., *et al.*, 2014. Distribution of ESCRT machinery at HIV assembly sites reveals virus scaffolding of ESCRT subunits. *Science* 343, 653–656.
- VerPlank, L., Bouamr, F., LaGrassa, T.J., *et al.*, 2001. Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55Gag. *Proceedings of the National Academy of Sciences of the United States of America* 98, 7724–7729.
- Votteler, J., Sundquist, W.I., 2013. Virus Budding and the ESCRT Pathway. *Cell Host & Microbe* 14, 232–241.
- Wang, T., Hong, W., 2006. RILP interacts with VPS22 and VPS36 of ESCRT-II and regulates their membrane recruitment. *Biochemical and Biophysical Research Communications* 350, 413–423.
- Webster, B.M., Colombi, P., Jäger, J., Lusk, C.P., 2014. Surveillance of nuclear pore complex assembly by ESCRT-III/Vps4. *Cell* 159, 388–401.
- Wollert, T., Hurley, J.H., 2010. Molecular mechanism of multivesicular body biogenesis by ESCRT complexes. *Nature* 464, 864–869.
- Wollert, T., Wunder, C., Lippincott-Schwartz, J., Hurley, J.H., 2009a. Membrane scission by the ESCRT-III complex. *Nature* 458, 172–177.
- Wollert, T., Yang, D., Ren, X., *et al.*, 2009b. The ESCRT machinery at a glance. *Journal of Cell Science* 122, 2163–2166.
- Xiao, J., Chen, X.-W., Davies, B.A., *et al.*, 2009. Structural basis of Ist1 function and Ist1-Did2 interaction in the multivesicular body pathway and cytokinesis. *Molecular Biology of the Cell* 20, 3514–3524.
- Zamborlini, A., Usami, Y., Radoshitzky, S.R., *et al.*, 2006. Release of autoinhibition converts ESCRT-III components into potent inhibitors of HIV-1 budding. *Proceedings of the National Academy of Sciences of the United States of America* 103, 19140–19145.

The Retromer Complex

MNJ Seaman, University of Cambridge, Cambridge, UK

© 2016 Elsevier Inc. All rights reserved.

Introduction

Endosomes serve as a point of convergence between the secretory and endocytic pathways and as such operate to sort membrane proteins (often referred to as cargo) to three distinct destinations. A cargo protein in an endosome can be sorted for retrieval to the Golgi, recycling to the cell surface or delivered to a lysosome for degradation. Sophisticated sorting mechanisms therefore exist to ensure cargo proteins reach their correct destination (Seaman, 2008). A key element of the endosomal sorting machinery is the retromer complex. The retromer complex is ubiquitously conserved in eukaryotes and functions to sort cargo, tubulate membranes and also recruits various accessory proteins to endosomes that contribute to endosomal protein sorting and provide regulation of retromer. Aberrant endosomal protein sorting and dysregulation of the retromer complex is linked to many pathological processes including neurodegenerative diseases such as Alzheimer disease and Parkinson's disease and retromer function has been linked bacterial and viral pathogenesis. In this article the molecular cell biology of the retromer complex and its function in endosomal protein sorting will be discussed.

A Brief History of Retromer

The mechanisms that govern sorting and delivery to the vacuole in yeast are very similar to the sorting and trafficking of lysosomal hydrolase receptors such as the cation-independent mannose 6-phosphate receptor (CIMPR) in mammals and other metazoan eukaryotes and hence yeast makes an excellent model system. The retromer complex was first identified in the yeast *Saccharomyces cerevisiae* through the analysis of vacuolar protein sorting (VPS) mutants. Genetic screens for *vps* mutants in yeast identified approximately 40 genes that are required for the efficient delivery of the yeast vacuolar hydrolase, carboxypeptidase Y (CPY) to the vacuole (reviewed in Bowers and Stevens, 2005). A key protein required for sorting CPY to the vacuole is a type-I membrane protein encoded by the *VPS10* gene (Marcusson *et al.*, 1994). The Vps10 protein is a receptor that binds to CPY in the late-Golgi and is concentrated into clathrin-coated vesicles which are then delivered to a prevacuolar endosome (Deloche *et al.*, 2001). The dissociation of CPY from Vps10p is induced by the luminal environment in the endosome, CPY then proceeds to the vacuole whilst Vps10p is recycled back to the late-Golgi.

Analysis of *vps* mutants that exhibit similar vacuolar protease sorting phenotypes to the *vps10* mutant revealed the requirement for Vps35p, Vps29p, and Vps30p in mediating the retrieval of the Vps10 protein from endosomes to the Golgi (Seaman *et al.*, 1997). Biochemical experiments using chemical crosslinking agents subsequently demonstrated that Vps35p and Vps29p form a complex that also contains Vps26p, Vps5p, and Vps17p – the complex was called retromer

to denote its role in retrieval from endosomes (Seaman *et al.*, 1998).

The biochemical analysis of yeast retromer showed that the pentameric complex could be dissected into two distinct subcomplexes; a trimer of Vps35p, Vps29p, and Vps26p that was proposed to sort cargo; and a dimer of the Vps5 and Vps17 proteins that could assemble into large complexes and was suggested to fulfill a structural role of deforming the endosomal membrane to produce a vesicle or tubule. The Vps5 and Vps17 proteins are members of the sorting nexin (SNX) family and both contain a phox homology (PX) domain that binds to phosphatidylinositol 3-phosphate (PtdIns 3P) (Horazdovsky *et al.*, 1997; Yu and Lemmon, 2001). Although *vps30* mutants are phenotypically similar to *vps29* and *vps35* mutants, the Vps30 protein was not detected as part of the retromer complex and was later shown to regulate the retromer complex by modulating the activity of Vps34p, a phosphatidylinositol kinase that generates PtdIns 3P on endosomal membranes. This is discussed in detail later.

The Assembly and Structure of Retromer

Subsequent to the identification of retromer in yeast, much has been learned about how the trimer of the VPS35, VPS29, and VPS26 proteins assembles to create the cargo-selective complex. The crystal structures of VPS26 and VPS29 have been solved, the latter whilst bound to VPS35. The VPS26 protein is an arrangement of two domains comprising only β -sheets that together create a structure that is similar to arrestins (Shi *et al.*, 2006; Collins *et al.*, 2008). As arrestins have been shown to be able to interact with membrane proteins, in particular G-protein coupled receptors (GPCRs) (see Kang *et al.*, 2014 for review) to mediate their endocytosis, the structural similarity between VPS26 and the arrestin proteins hints that VPS26 may also be able to directly interact with cargo proteins.

The VPS29 protein is a mixture of α -helices and β -sheets that fold to form a globular protein that is structurally similar to the PPP family of protein phosphatases but VPS29 does not appear to exhibit phosphatase activity (Collins *et al.*, 2005; Swarbrick *et al.*, 2011). In contrast to VPS26 and VPS29, the structure of full length VPS35 has not yet been determined but it is believed that VPS35 forms a flexible rod-like structure made up of multiple helical repeats (Hierro *et al.*, 2007). Studies in mammalian cells using siRNA to silence expression of individual components of the retromer cargo-selective complex has shown that loss of one subunit leads to the instability and degradation of the other proteins (Arighi *et al.*, 2004; Seaman, 2004).

The N-terminal region of VPS35 binds to VPS26, an association that involves a conserved motif in VPS35 comprising Pro-Arg-Leu-Tyr-Leu (PRLYL) residues where mutation of the first leucine to proline abolishes the interaction (Haft *et al.*, 2000; Gokool *et al.*, 2007a). The importance of these residues

was discovered by analysis of a yeast mutant of Vps35p that exerts a dominant negative effect on Vps10p retrieval (Seaman *et al.*, 1998). The VPS35-VPS26 interaction is high affinity with a dissociation constant (K_D) of 4 nM (Collins *et al.*, 2008; Norwood *et al.*, 2011).

The VPS29 protein binds to the C-terminal region of VPS35 which forms a horse shoe-shaped arrangement of helices around the globular VPS29 protein (Hiero *et al.*, 2007). On the surface of VPS29 there are two conserved hydrophobic patches which lie on opposing faces. The patch that contains Val and Ile at positions 90 and 91 respectively is required for binding to VPS35 (Collins *et al.*, 2005). Mutation of Val90 to Asp or Ile91 to Ser respectively weakens or abolishes the VPS29-VPS35 interaction. The other conserved hydrophobic patch on VPS29 faces away from the VPS35-VPS29 interface and has been shown to be important for binding to the SNX proteins Vps5p and Vps17p in yeast but in mammalian cells other proteins bind at this site that regulate retromer function and these will be discussed later.

Within the horse shoe-shaped fold of VPS35, a key His residue at position 675 is required for binding to VPS29 forming a hydrogen bond with the Ile91 residue of VPS29 (Helfer *et al.*, 2013). The VPS35-VPS29 interaction is substantially weaker than the VPS35-VPS26 with a K_D of approximately 250 nM (Swarbrick *et al.*, 2011). With VPS26 and VPS29 respectively binding at the N- and C-termini of VPS35, the retromer cargo-selective complex is similar to a dumbbell where the central region is flexible possibly so that VPS26 and VPS29 are able to move relative to each other on mobile and dynamic endosomal membranes. This intrinsic flexibility of VPS35 may be an important aspect of the function of the retromer cargo-selective complex in selecting cargo proteins for trafficking out of endosomes.

The initial identification of retromer in yeast indicated that the complex is a relatively stable heteropentamer, albeit one with the property that it could be biochemically split into two distinct subcomplexes (Seaman *et al.*, 1998). In higher eukaryotes such as mammals, however, it appears that the retromer cargo-selective complex and SNX dimer associate only transiently (Seaman *et al.*, 2009; Harbour *et al.*, 2010; Swarbrick *et al.*, 2011). Why this should be is not clear although it has been suggested that evolutionary variation of the VPS29 protein which has resulted in yeast Vps29p containing sequences not present in other eukaryotes may determine the degree to which the retromer cargo-selective complex associates with the SNX dimer or other proteins that regulate retromer function and are not conserved in yeast (Harbour and Seaman, 2011) – these retromer accessory proteins are discussed later.

The SNX dimer functions to tubulate the endosomal membrane to generate the transport intermediates required for endosome-to-Golgi retrieval (Carlton *et al.*, 2004). The production of tubules from endosomal membranes by the SNX dimer is driven by the Bin-Amphiphysin-Rvs (BAR) domains in the C-terminal regions of the SNXs along with an intrinsic self-assembly activity (Peter *et al.*, 2004). The BAR domains comprise predicted coiled-coils regions that form homo- or heterodimers. In yeast Vps5p and Vps17p heterodimerise via their respective coiled-coil C-termini (Seaman and Williams, 2002) and in higher eukaryotes a similar arrangement exists. For example, in humans there are two Vps5p homologs, SNX1

and SNX2 and there are two orthologs of Vps17p, SNX5 and SNX6. Thus the retromer SNX dimer in humans is a heterodimer of SNX1 or SNX2 with SNX5 or SNX6 (Wassmer *et al.*, 2007).

The BAR domains adopt a curved banana-shaped structure that can self-assemble to drive membrane tubulation (Peter *et al.*, 2004). Not all members of the SNX family have BAR domains and therefore the SNX proteins with BAR domains are distinguished from the others by being described as SNX-BAR proteins (for review see van Weering *et al.*, 2010). Even among the SNX-BAR proteins there is much variation regarding the ability of individual BAR domains to form tubules – variation that stems from the presence of amphipathic sequences in certain BAR domains that promote tubule formation. For example, the BAR domains of SNX1 and SNX2 can readily promote tubule formation whilst recombinant SNX5 or SNX6 BAR domains do not exhibit this property (van Weering *et al.*, 2012). The interactions and assembly of the retromer proteins to form the cargo-selective complex and SNX dimer is shown schematically in Figure 1.

How Membrane Tubules are Stabilized

The tubules observed in mammalian cells that are positive for SNX1 are often several microns long. Maintaining a long membrane tubule over this distance will present a biological challenge, for without some form of stabilization or scaffolding a tubule will tend to break apart into vesicles. Fortunately the interaction of the retromer SNX dimer with the microtubule cytoskeleton can provide the means to attach the membrane tubule to a large scaffold such as a microtubule. The SNX5 and SNX6 proteins can bind to the p150glued protein (also known as dynactin) thereby linking the SNX dimer to microtubules and dynein-mediated movement along microtubules (Wassmer *et al.*, 2009). As dynein is a minus-end directed motor that will travel along microtubules toward the microtubule organizing center (MTOC), a structure that is located close to the Golgi complex, it is perfectly suited to linking with retromer and thereby participating in the endosome-to-Golgi pathway. Yeast do not appear to employ microtubules for membrane trafficking pathways and there is no evidence that either Vps5p or Vps17p can link to the microtubule cytoskeleton. Possibly the small size of yeast negate the requirement for the membrane tubules to be stabilized by the cytoskeleton as is required in much larger cells such as mammalian cells.

In higher eukaryotes such as mammals additional machinery appears to also be required for stabilizing the tubules formed by the retromer SNX dimer. It has been shown that the Eps15-homology domain protein-1 (EHD1) is required to stabilize tubules generated by SNX1 (Gokool *et al.*, 2007b). Interestingly, however, EHD1 does not appear to associate with SNXs but rather can interact with the retromer cargo-selective complex (Zhang *et al.*, 2012). The EHD family of proteins can assemble to form ring-like structures around membrane tubules in-vitro and, through an intrinsic ATP-hydrolyzing ability, may generate force to induce membrane tubule scission or else could operate as an additional scaffold to stabilize membrane tubules (Daumke *et al.*, 2007). The

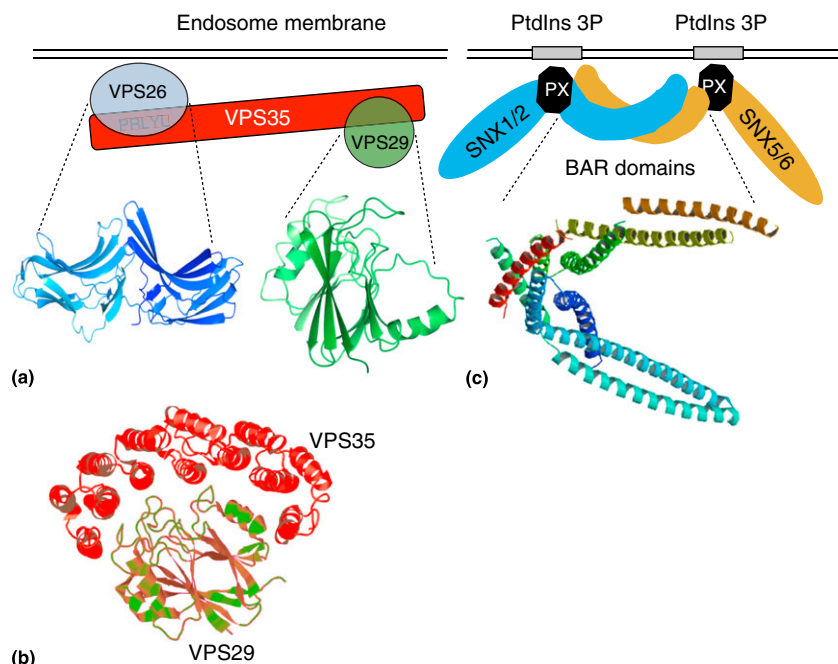


Figure 1 Key elements of the retromer complex. (a) The VPS35, VPS29 and VPS26 assemble into a trimer that can bind to membrane proteins (cargo) and hence is known as the retromer cargo-selective complex. The structures of VPS26 and VPS29 have been solved and are shown in blue and green, respectively, and are based on the pdb entries 2R51 and 1Z2X (<http://www.rcsb.org/pdb/home/home.do>). (b) The C-terminal region of VPS35 forms a horse shoe-shaped fold around VPS29. Residues critical for the VPS35-VPS29 interaction, (e.g., VPS35 His675 and VPS29 Ile91) are located in the cleft between the two molecules. The structure shown is based on pdb entry 2R17. (c) The sorting nexin dimer associates with the membrane by binding to PtdIns 3P and can tubulate membranes through the action of the C-terminal BAR domains. The structure shown is pdb entry 4FZS.

precise details of how EHD proteins contribute to the endosome-to-Golgi pathway have yet to be determined.

Recruitment of Retromer

Before retromer can sort cargo or tubulate the membrane it must be recruited to the endosome. For the SNX dimer the process of membrane recruitment is mediated by the production of PtdIns 3P on the endosome by a PtdIns 3-kinase (Cozier *et al.*, 2002). In yeast the activity of the PtdIns 3-kinase is itself regulated by the Vps30 protein which forms a complex with Vps38 (Burda *et al.*, 2002). In higher eukaryotes the Vps30 homologue is known as Beclin and also acts to promote production of PtdIns 3P (Kihara *et al.*, 2001). This lipid is specifically enriched on endosomal membranes and the PX domains present within the SNXs bind to this lipid (Cheever *et al.*, 2001). The presence of PtdIns 3P is regulated not only by the kinase, but also by phosphatases of the myotubularin family (Cao *et al.*, 2008). Additionally, PtdIns 3P can be converted to PtdIns 3,5P₂ by the PIKFyve protein, another PtdIns kinase. As the SNXs bind PtdIns 3,5P₂ with generally lower affinity than PtdIns 3P this restricts the amount of PtdIns 3P available for binding by SNXs (Rutherford *et al.*, 2006). These mechanisms can provide spatial and temporal control of the amount of PtdIns 3P on endosomes and thereby regulate the membrane association of the SNX proteins.

For the retromer cargo-selective complex, endosomal membrane recruitment occurs by mechanisms distinct from

the SNXs. None of the retromer cargo-selective complex proteins can bind lipids directly and therefore recruitment is instead mediated by binding to proteins. In mammalian cells the small GTPase Rab7a plays a key role in regulating retromer cargo-selective complex recruitment and the association with Rab7a is conserved across many eukaryotes (Rojas *et al.*, 2008; Seaman *et al.*, 2009). In addition to Rab7a, the SNX SNX3 is also required (Harterink *et al.*, 2011; Vardarajan *et al.*, 2012). Unlike most SNX proteins, the SNX3 protein comprises a PX domain but little else and has been shown to bind to PtdIns 3P with high affinity (Yu and Lemmon, 2001). Using recombinant proteins it has been recently demonstrated that Rab7a and SNX3 together promote the membrane association of the retromer cargo-selective complex by binding to the VPS35 protein within the N-terminal region of VPS35 that also binds to VPS26 (Harrison *et al.*, 2014).

The requirement for both Rab7a and SNX3 to mediate the recruitment of the retromer cargo-selective complex may define the region of the endocytic system where retromer operates, namely, a point where Rab7a and SNX3 are sufficiently enriched to enable recruitment of the retromer cargo-selective complex. The typical punctate co-localization of GFP-tagged SNX3 and the retromer VPS26 protein is shown in Figure 2. Interestingly a mutation in Rab7a that causes Charcot-Marie-Tooth (CMT) disease, a peripheral neuropathy, reduces the binding of Rab7a to the retromer cargo-selective complex although it is not yet known how this contributes to the pathology of CMT disease (Seaman *et al.*, 2009; Harrison *et al.*,

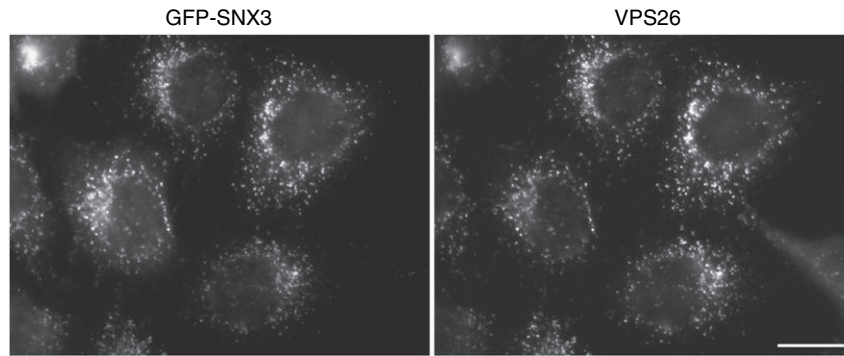


Figure 2 Typical localization of the retromer cargo-selective complex. HeLa cells stably transfected with GFP-SNX3 were fixed and stained with monoclonal anti-GFP and polyclonal anti-VPS26 along with appropriate fluorescently labeled secondary antibodies. The proteins colocalise on numerous punctate structures that exhibit typical endosomal morphology and distribution. Scale bar = 20 μ m.

2014). The role of SNX3 in promoting the recruitment of the retromer cargo-selective complex, although potentially conserved in yeast where SNX3 is known as Grd19p and functions with retromer (Strochlic *et al.*, 2007), is not universal as plants do not appear to express a SNX3 homologue.

Recognition of Cargo

The role of SNX3 in retromer-mediated endosomal protein sorting goes beyond recruiting the cargo-selective complex to the membrane. Studies, initially conducted in yeast, have shown that SNX3 can also interact directly with membrane proteins ('cargo') that traffic in a retromer-dependent manner. In yeast, SNX3/Grd19p is required for the sorting of the Ftr1 protein, a protein required for iron homeostasis (Strochlic *et al.*, 2007), and dipeptidyl aminopeptidase A (DPAPA), a late-Golgi protease involved in the processing of the yeast mating pheromone, α -factor (Nothwehr *et al.*, 1996). Other studies conducted in yeast have shown that Vps35p interacts with the cytoplasmic tail of the CPY receptor, Vps10p (Nothwehr *et al.*, 2000). The finding that Vps35p can directly interact with cargo proteins was replicated for mammalian VPS35 and the cation-independent mannose 6-phosphate receptor (CIMPR) using the yeast two-hybrid system to investigate the interaction (Arighi *et al.*, 2004). Loss of retromer function in mammalian cells results in mislocalization and degradation of the CIMPR and the secretion of immature precursor forms of lysosomal hydrolases such as cathepsin D (Arighi *et al.*, 2004; Seaman, 2004).

Another cargo protein for retromer in mammalian cells is the iron transporter SLC11a2, also known as DMT1-II and Nramp2 (Tabuchi *et al.*, 2010) – mirroring the role of retromer in sorting the iron transporter protein, Ftr1p, in yeast. Studies conducted *in vitro* were able to demonstrate direct binding of the retromer cargo-selective complex to the C-terminal cytoplasmic domain of SLC11a2 through a hydrophobic motif, YLL (Tyr-Leu-Leu). This sequence is similar to motifs identified in the cytoplasmic tails of the CIMPR and sortilin, a mammalian homologue of the yeast Vps10 protein – WLM (Trp-Leu-Met) for CIMPR and FLV (Phe-Leu-Val) for sortilin (Seaman, 2007) – suggesting that retromer-mediated sorting

operates through a motif with a consensus of aromatic-Leu-hydrophobic. In worms and flies, developmental cues controlled by Wnt signaling require the action of the Wntless protein which is a cargo for retromer that also contains a hydrophobic aromatic-containing sequence in its cytoplasmic domain (Eaton, 2008).

Whilst the experimental evidence shows that VPS35 can bind to retromer cargo proteins, it is not the only component that can interact with cargo. The VPS26 protein can also bind to cargo proteins and has been shown to recognize the cytoplasmic tail of the SorL1 protein (also known as SorLA) (Fjorback *et al.*, 2012), a transmembrane protein that, like sortilin, is related to yeast Vps10p. The VPS26-SorL1 interaction is mediated by the motif FANSHY, (Phe-Ala-Asn-Ser-His-Tyr) in the SorL1 cytoplasmic tail. Presently, however, the precise molecular details of how VPS35 or VPS26 bind to cargo proteins remains to be determined but an interaction between VPS26 and cargo is consistent with structural similarities with arrestin proteins described previously. Some of the best characterized retromer cargo proteins are listed in Table 1.

Accessory Proteins that Modulate Retromer Function

Since its discovery in yeast and characterization in mammalian cells, a number of accessory proteins that associate with retromer to modulate its activity have been identified. In many cases there is a link to neurodegenerative disease through the accessory proteins that interact with retromer. The interaction of the SNX dimer with p150glued and the association of the retromer cargo-selective complex with EHD1 have been described previously as has the interaction of the retromer cargo-selective complex with Rab7a and SNX3.

The immunoprecipitation of tagged versions of the retromer cargo-selective complex has led to the identification of additional accessory proteins that modulate retromer activity and contribute to efficient endosomal protein sorting. The TBC1D5 protein associates with the retromer cargo-selective complex by binding to VPS29 (Seaman *et al.*, 2009; Harbour *et al.*, 2010). Interestingly, TBC1D5 binds to a hydrophobic patch on VPS29 that, in yeast, mediates the interaction between the retromer cargo-selective complex and the SNX dimer (Collins *et al.*,

Table 1 Notable cargo proteins that require retromer for their localization

Cargo protein	Function	Physiological process	Organism	Reference(s)
Vps10p	Vacuole hydrolase sorting	Vacuole biogenesis	Yeast	Seaman <i>et al.</i> (1997, 1998); Nothwehr <i>et al.</i> (2000)
Ftr1p	Component of iron transporter	Iron homeostasis	Yeast	Strochlic <i>et al.</i> (2007)
CIMPR	Lysosome hydrolase sorting	Lysosome biogenesis	Humans/metazoans	Seaman (2004); Arighi <i>et al.</i> (2004)
SLC11a2/DMT1-II	Iron transporter	Iron homeostasis	Humans/metazoans	Tabuchi <i>et al.</i> (2010)
Sortilin	Lysosome hydrolase sorting	Lysosome biogenesis	Humans/metazoans	Seaman (2004); Seaman (2007)
SorL1/SorLA	Associates with APP	Regulates APP processing	Humans	Fjorback <i>et al.</i> (2012)
plgAR	Binds polymeric IgA	Transcytosis of IgA	Humans	Vergés <i>et al.</i> (2004)
Crumbs	Marker of apical domain	Establishment of polarity	Humans/flies	Pocha <i>et al.</i> (2011)
Wntless	Chaperone for Wnt	Wnt-based signaling	Humans/flies	Summarized in Eaton (2008)
TGN38/46	Unknown	Unknown	Humans/metazoans	Lieu and Gleeson (2010)

2005). TBC1D5 is a member of the TBC family of proteins that regulate rab GTPases by promoting the hydrolysis of GTP by the respective rab and as such are referred to as rab GTPase activating proteins (GAPs) (see Fukuda, 2011 for review). An obvious potential target for TBC1D5 is the Rab7a protein that the retromer cargo-selective complex requires for its membrane association and indeed the nematode homologue of TBC1D5, the Rbg-3 protein, does promote hydrolysis of GTP by Rab7 (Mukhopadhyay *et al.*, 2007). In mammalian cells, overexpression of TBC1D5 causes the retromer cargo-selective complex to dissociate from endosomal membranes similar to the siRNA knockdown of Rab7a (Seaman *et al.*, 2009). Mutation of conserved Arg and Glu residues in the TBC domain of TBC1D5 that, based on the structure of TBC domains (Pan *et al.*, 2006), are predicted to be required for GAP activity renders TBC1D5 inactive with respect to affecting the membrane association of the retromer cargo-selective complex. Thus, TBC1D5 acts to control the membrane association of the retromer cargo-selective complex by regulating the activity of Rab7a.

Curiously, the retromer cargo-selective complex has recently been shown to associate with a protein called VARP that functions as a guanine nucleotide exchange factor (GEF) for Rab21 and an effector of Rab32/38 – Rab GTPases that are respectively linked with endosome-to-cell surface transport and melanosome biogenesis. What is intriguing about the retromer-VARP interaction is that it occurs through the same hydrophobic patch on VPS29 responsible for binding TBC1D5 (Hesketh *et al.*, 2014). Thus the association of the retromer cargo-selective complex with either VARP or TBC1D5 may provide the means for retromer to regulate sorting into multiple pathways leading away from endosomes.

Along with TBC1D5 and VARP, the retromer cargo-selective complex also interacts with the WASH complex (Harbour *et al.*, 2010) which functions to promote formation of branched actin patches on endosomes by activating the Arp2/3 complex. The WASH complex comprises five proteins; WASH1, KIAA1033, strumpellin, CCDC53, and FAM21 that form a stable complex (Derivery *et al.*, 2009). The strumpellin and KIAA1033 subunits of the WASH complex have respectively been linked with hereditary spastic paraplegia (HSP) and autosomal recessive intellectual disability (ARID) indicating

that WASH complex function is especially important in neuronal tissue (Valdmanis *et al.*, 2007; Ropers *et al.*, 2011). Additionally, the retromer cargo-selective complex also interacts with a protein called FKBP15 (also known as WAFL) that associates with the WASH complex and may modulate actin dynamics on endosomes through further interactions with actin and the WIP (WAS/WASL-interacting protein family member 1) protein (Viklund *et al.*, 2008). The precise role of FKBP15/WAFL in retromer-mediated endosomal protein sorting remains to be determined.

The interaction between the retromer cargo-selective complex and the WASH complex occurs through VPS35 binding to FAM21, an association that requires VPS29 to be bound to VPS35 and is mediated by repeated sequences called LFa motifs in the long unstructured tail of FAM21 (Harbour *et al.*, 2010, 2012; Jia *et al.*, 2012; Helfer *et al.*, 2013). The retromer cargo-selective complex is required for the WASH complex to localize to endosomes and overexpression of just the FAM21 tail can cause endogenous WASH complex to be displaced from the endosome due to competition for binding to VPS35. Recently it has been reported that a mutation in VPS35 that causes an inherited form of Parkinson's disease (Vilariño-Güell *et al.*, 2011; Zimprich *et al.*, 2011) impairs the association of the WASH complex with the retromer cargo-selective complex and results in reduced WASH complex recruitment to endosomes (Zavodszky *et al.*, 2014). The sequential recruitment of the retromer cargo-selective complex and the WASH complex to the endosomal membrane, and the role of retromer in sorting cargo proteins into multiple pathways, is depicted in Figure 3.

The association of the retromer cargo-selective complex with the WASH complex and the interaction of the SNX dimer with the microtubule cytoskeleton means that, in mammals, retromer-mediated endosomal protein sorting is linked to both short range actin-based movement to regulate protein localization within the endosomal membrane and long range microtubule-mediated movement to control the location of the endosome within the cell.

In addition to binding to the retromer cargo-selective complex, the FAM21 tail also associates with a protein called RME-8 (Freeman *et al.*, 2014). The RME-8 protein is a member of the DNAJ domain protein family that often associate

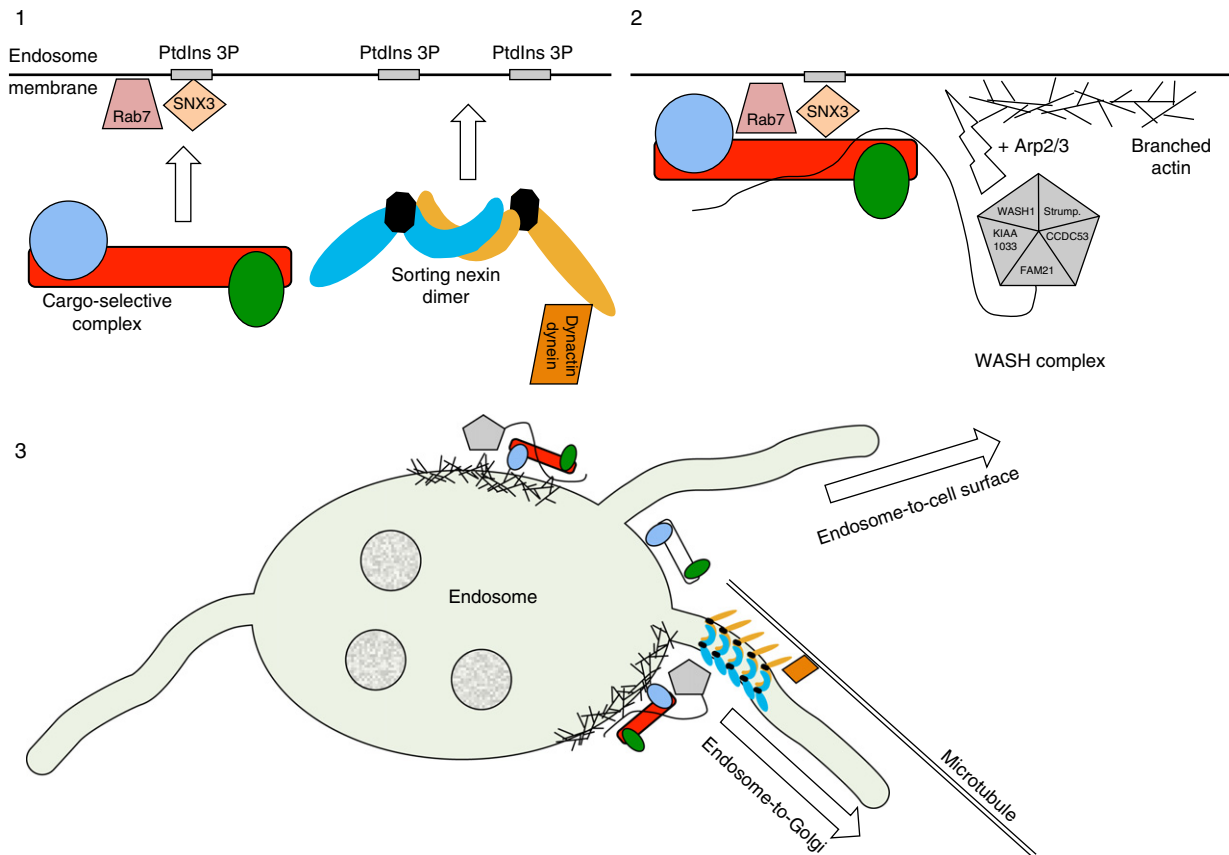


Figure 3 Recruitment and action of the retromer complex. In the three panels the action of the retromer complex is shown sequentially: (a) The retromer cargo-selective complex is recruited to the membrane through binding to Rab7 and SNX3 whilst the sorting nexin dimer binds to PtdIns 3P and can also associate with the dynactin–dynein complex. In (b), the retromer cargo-selective complex recruits the WASH complex through VPS35 binding to the extended unstructured ‘tail’ of FAM21. This enables the WASH complex to promote Arp2/3-mediated branched actin production on endosomal membranes. In (c), actin-stabilized microdomains facilitate sorting of cargo proteins into multiple pathways exiting the endosome.

with chaperone proteins and have roles in regulating assembly of macromolecular complexes. Interestingly, RME-8 (also known as DNAJC13) has been shown to interact with SNX1 (Shi *et al.*, 2009) and loss of RME-8 function alters the kinetics of SNX1 association with endosomes and results in pronounced tubulation of endosomal membranes (Freeman *et al.*, 2014). It has recently been demonstrated that RME-8/DNAJC13 is mutated in an inherited form of Parkinson’s disease (PD) (Vilar-iño-Güell *et al.*, 2014). As both VPS35 and RME-8/DNAJC13 have now been linked to PD and both can associate with the FAM21 tail, it seems possible that the protein–protein interaction network that controls retromer–WASH complex function at endosomes is especially important in the pathology of PD. The interaction of the retromer cargo-selective complex with various accessory proteins, many of which are implicated in neurological disease is shown in Figure 4.

Studies of the role of the WASH complex in endosomal protein sorting have shown that numerous cargo proteins require the WASH complex for their correct localization. The WASH complex has been implicated in the retrieval of the CIMPR from endosomes to the Golgi (Gomez and Billadeau, 2009) and has also recently been linked to trafficking the multi-pass membrane protein ATG9 to autophagosomes (Zavodszky

et al., 2014). Interestingly, however, most of the cargo proteins shown to require WASH complex function traffic in an endosome-to-cell surface pathway (reviewed in Seaman *et al.*, 2013). These proteins include the glucose transporter GLUT-1, the β -adrenergic receptor, which is a G-protein coupled receptor (GPCR) and the $\alpha 5 \beta 1$ integrin protein required for cell spreading. Thus, through the requirement for the retromer cargo-selective complex to mediate the endosomal recruitment of the WASH complex, retromer is involved in sorting cargo proteins for recycling to the cell surface.

Recently another SNX protein, SNX27 has been shown to associate with both the WASH complex and also the retromer cargo-selective complex (Steinberg *et al.*, 2013). The SNX27 protein binds to VPS26 and may participate in cargo selection through the FERM (4.1/ezrin/radixin/moesin) domain present in SNX27 that can recognize, NPxY (Asn-Pro-x-Tyr) motifs present in a number of cargo proteins within the endocytic system (Ghai *et al.*, 2013). Unlike the retromer SNX dimer, SNX27 does not possess a BAR domain and therefore cannot mediate membrane tubulation.

It is important to note that whilst the retromer complex is ubiquitously conserved, many of the accessory proteins that associate with retromer in complex metazoans such as

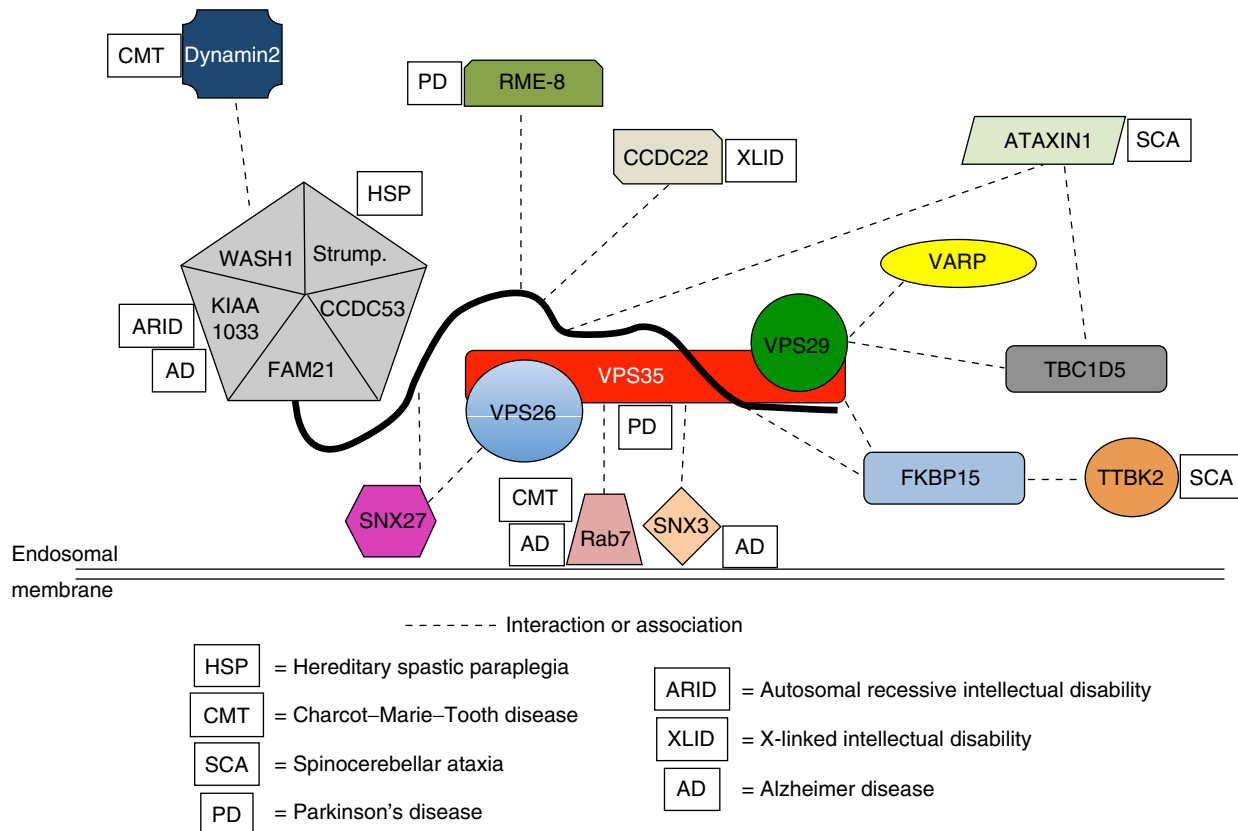


Figure 4 The retromer cargo-selective complex is the nexus of a protein–protein interaction network that is linked to many neurological disorders. The retromer cargo-selective complex and WASH complex associate with a number of proteins linked to disease, for example, the WASH complex interacts with the Parkinson's disease protein, RME-8 via FAM21 (Freeman *et al.*, 2014). Proteins shown here that are not discussed in the text include Dynamin-2 that associates with the WASH complex (Derivery *et al.*, 2009), Ataxin-1 that can interact with TBC1D5 and FAM21 (Lim *et al.*, 2006), CCDC22 that binds to FAM21 (Harbour *et al.*, 2012) and TTBK2 that is reported to interact with FKBP15 (Pan *et al.*, 2010).

mammals are less well conserved. The TBC1D5 protein is not conserved in yeast, nor is the WASH complex or EHD1 suggesting that retromer can perform its essential function(s) without these accessory proteins. The retromer accessory proteins may therefore provide additional layers of regulation required for 'fine tuning' retromer-mediated endosomal protein sorting – adaptations that have been lost in simple organisms such as yeast.

The Physiological Processes That Depend on Retromer

The sorting of a diverse array of cargo proteins by retromer sets the scene for retromer to contribute to a great many physiological processes. The requirement for retromer in sorting vacuolar and lysosomal hydrolase receptors for retrieval to the Golgi demonstrates the importance of retromer in lysosome biogenesis (Arighi *et al.*, 2004; Seaman, 2004). Developmental pathways that utilize the morphogen Wnt also depend on retromer to ensure the Wnt chaperone protein Wntless is retrieved from endosomes to pick up more Wnt in the Golgi (see Eaton, 2008 for summary). The SorL1 protein associates

with amyloid precursor protein (APP) and controls APP localization within the endocytic system thereby regulating the processing of APP – processing that determines whether APP is cleaved to form the neurotoxic pro-aggregatory $A\beta$ peptide that is causal in Alzheimer disease. SorL1 is a cargo for retromer and therefore retromer has become recognized as a key player in the pathogenesis of Alzheimer disease (for review see Willnow and Andersen, 2013). Indeed studies of neuronal tissue from Alzheimer disease (AD) patients have revealed reduced retromer expression and RNAi-mediated silencing of retromer proteins results in increased production of the $A\beta$ peptide (Muhammad *et al.*, 2008). The key role that retromer-mediated sorting of SorL1 plays in regulating the localization and processing of APP is shown schematically in Figure 5.

In neuronal cells, retromer has been shown to mediate sorting and transport of proteins from endosomes to the cell surface at localized sites thereby contributing to the generating of dendritic spikes – structures that are often associated with synapse formation and synaptic plasticity (Choy *et al.*, 2014). It seems likely that retromer-associated proteins such as the WASH complex will, through an interaction with retromer, play a major role in these processes and hence this may

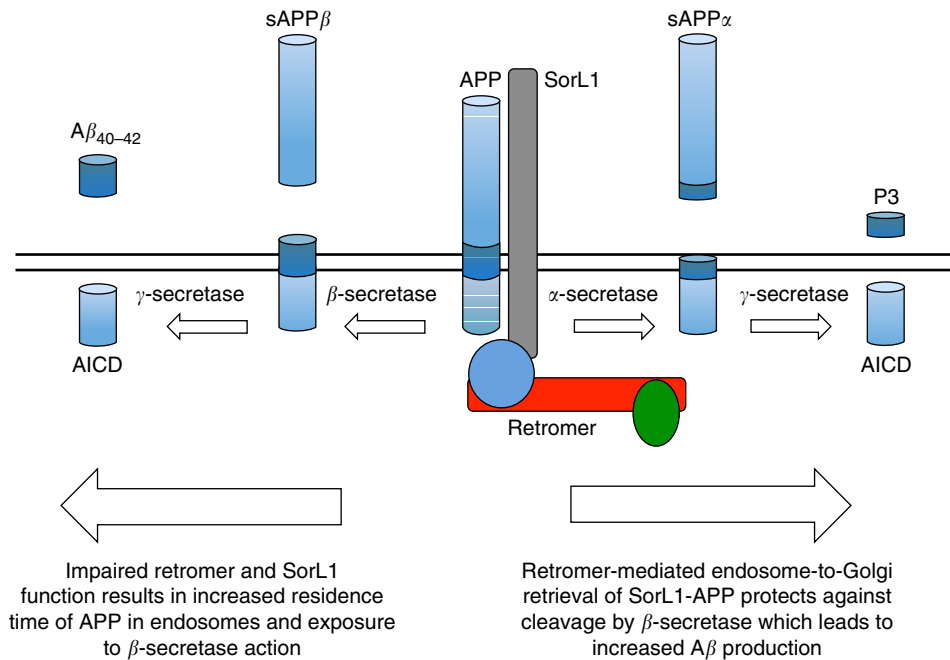


Figure 5 Retromer function regulates APP processing and therefore is important for Alzheimer disease pathogenesis. The retromer cargo-selective complex can associate with the SorL1 protein that regulates APP localization. Sorting of SorL1-APP by retromer reduces APP residence time in endosomes and thereby protects APP from cleavage by the β -secretase (BACE). The diagram is adapted from Andersen, O.M., Willnow, T.E., 2006. Lipoprotein receptors in Alzheimer's disease. *Trends in Neurosciences* 29 (12), 687–694.

provide some means to explain the neuronal phenotypes (e.g., HSP and ARID) associated with WASH complex dysfunction.

Retromer has been shown to control the localization of the membrane protein Crumbs that is required for the establishment of cell polarity (Pocha *et al.*, 2011). Additionally, in polarized cells, retromer is required for the transcytosis of the polymeric IgA receptor – a protein that mediates the transfer of IgA from blood into bile (Vergés *et al.*, 2004). In addition to regulating protein localization, retromer has also been shown to control the signaling activity of certain GPCRs. Upon ligand binding, the parathyroid hormone receptor (PTHr) is endocytosed in an arrestin-dependent manner. After arrival at the endosome, the signaling activity of the PTHr is down regulated – a process that involves the displacement of the arrestin by retromer, most likely with the VPS26 subunit playing a key role (Feinstein *et al.*, 2011).

A recent proteomic study of retromer-mediated recycling of proteins to the cell surface identified more than 100 membrane proteins that require retromer function for their localization to the plasma membrane including many receptors and signaling molecules (Steinberg *et al.*, 2013). These proteins are likely to participate in a great many physiological processes and therefore retromer is likely to be required for additional processes and mechanisms that have yet to be identified and reported.

Retromer and Disease

The importance of retromer-mediated endosomal protein sorting to the pathogenesis of AD has already been described.

Indeed certain polymorphisms of the gene that encodes the retromer cargo SorL1 have been shown to cause late-onset AD (Rogaeva *et al.*, 2007). Additionally variants of the Rab7a and SNX3 genes have also been linked to late-onset AD (Vardarajan *et al.*, 2012). A rare point mutation in VPS35 that causes Parkinson's disease (PD) has been mentioned previously. Mutants of the WASH complex proteins, strumpellin, and KIAA1033, respectively, cause HSP and ARID. These studies all show how retromer-mediated endosomal protein sorting is key to many neurodegenerative diseases although much remains unknown regarding the patho-physiological processes that result from defective or deficient retromer function.

Various pathogens exploit retromer-mediated endosomal protein sorting for their own ends. For example, the Shiga toxin protein produced by Shigella bacteria utilizes retromer-mediated endosome-to-Golgi retrieval in order to traverse the endocytic system en route to firstly the Golgi and then the endoplasmic reticulum (Bujny *et al.*, 2007). Invasive bacteria such as Salmonella and Legionella also interfere with retromer-mediated endosomal protein sorting as part of the process of pathogenesis (McGourty *et al.*, 2012; Finsel *et al.*, 2013). In the case of Legionella, once inside a cell, the bacteria express a protein called RidL that can bind to VPS29 to inhibit the function of retromer.

Viral pathogens can also interact with or subvert retromer function as recently demonstrated by the finding that the human papilloma virus (HPV) interacts with retromer post-infection (Lipovsky *et al.*, 2013).

Whilst the precise mechanisms that underlie these various pathologies remain to be determined, it seems likely that

interfering in retromer function, or the function of retromer-associated proteins, is likely to lead to mislocalization of many different membrane proteins linked to various physiological processes. There is, therefore, much yet to learn regarding how retromer functions in endosomal protein sorting and how these mechanisms contribute to health and disease and no doubt many exciting discoveries will emerge in the coming years.

Acknowledgments

MNJS is funded by a grant from the Medical Research Council.

See also: Interorganellar Communication: Interplay and Processes: Rabs of the Endosomal Recycling Pathway. Organelles: Structure and Function: Early Endosomal Compartments

References

- Arighi, C.N., Hartnell, L.M., Aguilar, R.C., Haft, C.R., Bonifacio, J.S., 2004. Role of the mammalian retromer in sorting of the cation-independent mannose 6-phosphate receptor. *Journal of Cell Biology* 165 (1), 123–133.
- Bowers, K., Stevens, T.H., 2005. Protein transport from the late Golgi to the vacuole in the yeast *Saccharomyces cerevisiae*. *Biochimica et Biophysica Acta* 1744 (3), 438–454.
- Bujny, M.V., Popoff, V., Johannes, L., Cullen, P.J., 2007. The retromer component sorting nexin-1 is required for efficient retrograde transport of Shiga toxin from early endosome to the trans Golgi network. *Journal of Cell Science* 120 (Pt 12), 2010–2021.
- Burda, P., Padilla, S.M., Sarkar, S., Emr, S.D., 2002. Retromer function in endosome-to-Golgi retrograde transport is regulated by the yeast Vps34 PtdIns 3-kinase. *Journal of Cell Science* 115 (Pt 20), 3889–3900.
- Cao, C., Backer, J.M., Laporte, J., Bedrick, E.J., Wandering-Ness, A., 2008. Sequential actions of myotubularin lipid phosphatases regulate endosomal PI(3)P and growth factor receptor trafficking. *Molecular Biology of the Cell* 19 (8), 3334–3346.
- Carlton, J., Bujny, M., Peter, B.J., et al., 2004. Sorting nexin-1 mediates tubular endosome-to-TGN transport through coincidence sensing of high-curvature membranes and 3-phosphoinositides. *Current Biology* 14 (20), 1791–1800.
- Cheever, M.L., Sato, T.K., de Beer, T., et al., 2001. Phox domain interaction with PtdIns(3)P targets the Vam7 t-SNARE to vacuole membranes. *Nature Cell Biology* 3 (7), 613–618.
- Choy, R.W., Park, M., Temkin, P., et al., 2014. Retromer mediates a discrete route of local membrane delivery to dendrites. *Neuron* 82 (1), 55–62.
- Collins, B.M., Norwood, S.J., Kerr, M.C., et al., 2008. Structure of Vps26B and mapping of its interaction with the retromer protein complex. *Traffic* 9 (3), 366–379.
- Collins, B.M., Skinner, C.F., Watson, P.J., Seaman, M.N.J., Owen, D.J., 2005. Vps29 has a phosphoesterase fold that acts as a protein interaction scaffold for retromer assembly. *Nature Structural & Molecular Biology* 12 (7), 594–602.
- Cozier, G.E., Carlton, J., McGregor, A.H., et al., 2002. The phox homology (PX) domain-dependent, 3-phosphoinositide-mediated association of sorting nexin-1 with an early sorting endosomal compartment is required for its ability to regulate epidermal growth factor receptor degradation. *Journal of Biological Chemistry* 277 (50), 48730–48736.
- Daumke, O., Lundmark, R., Vallis, Y., et al., 2007. Architectural and mechanistic insights into an EHD ATPase involved in membrane remodelling. *Nature* 449, 923–927.
- Deloche, O., Yeung, B.G., Payne, G.S., Schekman, R., 2001. Vps10p transport from the trans-Golgi network to the endosome is mediated by clathrin-coated vesicles. *Molecular Biology of the Cell* 12 (2), 475–485.
- Derivery, E., Sousa, C., Gautier, J.J., et al., 2009. The Arp2/3 activator WASH controls the fission of endosomes through a large multiprotein complex. *Developmental Cell* 17 (5), 712–723.
- Eaton, S., 2008. Retromer retrieves wntless. *Developmental Cell* 14 (1), 4–6.
- Feinstein, T.N., Wehbi, V.L., Ardura, J.A., et al., 2011. Retromer terminates the generation of cAMP by internalized PTH receptors. *Nature Chemical Biology* 7 (5), 278–284.
- Finsel, I., Ragaz, C., Hoffmann, C., et al., 2013. The Legionella effector RldL inhibits retrograde trafficking to promote intracellular replication. *Cell Host & Microbe* 14 (1), 38–50.
- Fjorback, A.W., Seaman, M., Gustafsen, C., et al., 2012. Retromer binds the FANSHY sorting motif in SorLA to regulate amyloid precursor protein sorting and processing. *Journal of Neuroscience* 32 (4), 1467–1480.
- Freeman, C.L., Hesketh, G., Seaman, M.N.J., 2014. RME-8 coordinates the activity of the WASH complex with the function of the retromer SNX dimer to control endosomal tubulation. *Journal of Cell Science* 127, 2053–2070.
- Fukuda, M., 2011. TBC proteins: GAPs for mammalian small GTPase Rab? *Bioscience Reports* 31 (3), 159–168.
- Ghai, R., Bugarcic, A., Liu, H., et al., 2013. Structural basis for endosomal trafficking of diverse transmembrane cargos by PX-FERM proteins. *Proceedings of the National Academy of Sciences of the United States of America* 110 (8), E643–E652.
- Gokool, S., Tattersall, D., Reddy, J.V., Seaman, M.N.J., 2007a. Identification of a conserved motif that is required for Vps35p/Vps26p interaction and assembly of the retromer complex. *Biochemical Journal* 408 (2), 287–295.
- Gokool, S., Tattersall, D., Seaman, M.N.J., 2007b. EHD1 interacts with retromer to stabilize SNX1-positive tubules and facilitate endosome-to-Golgi retrieval. *Traffic* 8, 1873–1886.
- Gomez, T.S., Billadeau, D.D., 2009. A FAM21-containing WASH complex regulates retromer-dependent sorting. *Developmental Cell* 17 (5), 699–711.
- Haft, C.R., de la Luz Sierra, M., Bafford, R., et al., 2000. Human orthologs of yeast vacuolar protein sorting proteins Vps26, 29, and 35: Assembly into multimeric complexes. *Molecular Biology of the Cell* 11 (12), 4105–4116.
- Harbour, M.E., Breusegem, S.Y., Antrobus, R., et al., 2010. The cargo-selective retromer complex is a recruiting hub for protein complexes that regulate endosomal tubule dynamics. *Journal of Cell Science* 123, 3703–3717.
- Harbour, M.E., Breusegem, S.Y., Seaman, M.N.J., 2012. Recruitment of the endosomal WASH complex is mediated by the extended 'tail' of Fam21 binding to the retromer protein Vps35. *Biochemical Journal* 442 (1), 209–220.
- Harbour, M.E., Seaman, M.N.J., 2011. Evolutionary variations of VPS29, and their implications for the pentameric model of retromer. *Communicative & Integrative Biology* 4 (5), 619–622.
- Harrison, M.S., Hung, C.S., Liu, T.T., et al., 2014. A mechanism for retromer endosomal coat complex assembly with cargo. *Proceedings of the National Academy of Sciences of the United States of America* 111 (1), 267–272.
- Harterink, M., Port, F., Lorenowicz, M.J., et al., 2011. A SNX3-dependent retromer pathway mediates retrograde transport of the Wnt sorting receptor Wntless and is required for Wnt secretion. *Nature Cell Biology* 13 (8), 914–923.
- Heifer, E., Harbour, M.E., Henriot, V., et al., 2013. Endosomal recruitment of the WASH complex: Active sequences and mutations impairing interaction with the retromer. *Biology of the Cell* 105 (5), 191–207.
- Hesketh, G., Pérez-Dorado, I., Jackson, L.P., et al., 2014. VARP is recruited onto endosomes by direct interaction with retromer where together they function in export to the cell surface. *Developmental Cell* 29 (5), 591–606.
- Hierro, A., Rojas, A.L., Rojas, R., et al., 2007. Functional architecture of the retromer cargo-recognition complex. *Nature* 449, 1063–1067.
- Horazdovsky, B.F., Davies, B.A., Seaman, M.N., et al., 1997. A sorting nexin-1 homologue, Vps5p, forms a complex with Vps17p and is required for recycling the vacuolar protein-sorting receptor. *Molecular Biology of the Cell* 8, 1529–1541.
- Jia, D., Gomez, T.S., Billadeau, D.D., Rosen, M.K., 2012. Multiple repeat elements within the FAM21 tail link the WASH actin regulatory complex to the retromer. *Molecular Biology of the Cell* 23 (12), 2352–2361.
- Kang, D.S., Tian, X., Benovic, J.L., 2014. Role of β -arrestins and arrestin domain-containing proteins in G protein-coupled receptor trafficking. *Current Opinion in Cell Biology* 27C, 63–71.
- Kihara, A., Kabeya, Y., Ohsumi, Y., Yoshimori, T., 2001. Beclin-phosphatidylinositol 3-kinase complex functions at the trans-Golgi network. *EMBO Reports* 2 (4), 330–335.
- Lieu, Z.Z., Gleeson, P.A., 2010. Identification of different itineraries and retromer components for endosome-to-Golgi transport of TGN38 and Shiga toxin. *European Journal of Cell Biology* 89 (5), 379–393.
- Lim, J., Hato, T., Shaw, C., et al., 2006. A protein-protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. *Cell* 125 (4), 801–814.

- Lipovsky, A., Popa, A., Pimienta, G., *et al.*, 2013. Genome-wide siRNA screen identifies the retromer as a cellular entry factor for human papillomavirus. *Proceedings of the National Academy of Sciences of the United States of America* 110 (18), 7452–7457.
- Marcusson, E.G., Horazdovsky, B.F., Cereghino, J.L., Gharakhanian, E., Emr, S.D., 1994. The sorting receptor for yeast vacuolar carboxypeptidase Y is encoded by the VPS10 gene. *Cell* 77 (4), 579–586.
- McGourty, K., Thurston, T.L., Matthews, S.A., *et al.*, 2012. Salmonella inhibits retrograde trafficking of mannose-6-phosphate receptors and lysosome function. *Science* 338, 963–967.
- Muhammad, A., Flores, I., Zhang, H., *et al.*, 2008. Retromer deficiency observed in Alzheimer's disease causes hippocampal dysfunction, neurodegeneration, and Abeta accumulation. *Proceedings of the National Academy of Sciences of the United States of America* 105 (20), 7327–7332.
- Mukhopadhyay, A., Pan, X., Lambright, D.G., Tissenbaum, H.A., 2007. An endocytic pathway as a target of tubby for regulation of fat storage. *EMBO Reports* 8 (10), 931–938.
- Norwood, S.J., Shaw, D.J., Cowieson, N.P., *et al.*, 2011. Assembly and solution structure of the core retromer protein complex. *Traffic* 12 (1), 56–71.
- Nothwehr, S.F., Bryant, N.J., Stevens, T.H., 1996. The newly identified yeast GRD genes are required for retention of late-Golgi membrane proteins. *Molecular and Cellular Biology* 16 (6), 2700–2707.
- Nothwehr, S.F., Ha, S.A., Bruinsma, P., 2000. Sorting of yeast membrane proteins into an endosome-to-Golgi pathway involves direct interaction of their cytosolic domains with Vps35p. *Journal of Cell Biology* 151 (2), 297–310.
- Pan, X., Eathiraj, S., Munson, M., Lambright, D.G., 2006. TBC-domain GAPs for Rab GTPases accelerate GTP hydrolysis by a dual-finger mechanism. *Nature* 442, 303–306.
- Pan, Y.F., Viklund, I.M., Tsai, H.H., Pettersson, S., Maruyama, I.N., 2010. The ulcerative colitis marker protein WAFL interacts with accessory proteins in endocytosis. *International Journal of Biological Sciences* 6 (2), 163–171.
- Peter, B.J., Kent, H.M., Mills, I.G., *et al.*, 2004. BAR domains as sensors of membrane curvature: The amphiphysin BAR structure. *Science* 303, 495–499.
- Pocha, S.M., Wassmer, T., Niehage, C., Hofflack, B., Knust, E., 2011. Retromer controls epithelial cell polarity by trafficking the apical determinant Crumbs. *Current Biology* 21 (13), 1111–1117.
- Rogaeva, E., Meng, Y., Lee, J.H., *et al.*, 2007. The neuronal sortilin-related receptor SORL1 is genetically associated with Alzheimer disease. *Nature Genetics* 39 (2), 168–177.
- Rojas, R., van Vlijmen, T., Mardones, G.A., *et al.*, 2008. Regulation of retromer recruitment to endosomes by sequential action of Rab5 and Rab7. *Journal of Cell Biology* 183 (3), 513–526.
- Ropers, F., Derivery, E., Hu, H., *et al.*, 2011. Identification of a novel candidate gene for non-syndromic autosomal recessive intellectual disability: The WASH complex member SWIP. *Human Molecular Genetics* 20 (13), 2585–2590.
- Rutherford, A.C., Traer, C., Wassmer, T., *et al.*, 2006. The mammalian phosphatidylinositol 3-phosphate 5-kinase (PIKfyve) regulates endosome-to-TGN retrograde transport. *Journal of Cell Science* 119 (Pt 19), 3944–3957.
- Seaman, M.N.J., 2004. Cargo-selective endosomal sorting for retrieval to the Golgi requires retromer. *Journal of Cell Biology* 165 (1), 111–122.
- Seaman, M.N.J., 2007. Identification of a novel conserved sorting motif necessary for retromer-mediated endosome-to-Golgi retrieval. *Journal of Cell Science* 120, 2378–2389.
- Seaman, M.N.J., 2008. Endosome protein sorting: Motifs and machinery. *Cellular and Molecular Life Sciences* 65, 2842–2858.
- Seaman, M.N.J., Gautreau, A., Billadeau, D.D., 2013. Retromer-mediated endosomal protein sorting: All WASHed up!. *Trends in Cell Biology* 23 (11), 522–528.
- Seaman, M.N.J., Harbour, M.E., Tattersall, D., Read, E., Bright, N., 2009. Membrane recruitment of the cargo-selective retromer subcomplex is catalysed by the small GTPase Rab7 and inhibited by the Rab-GAP TBC1D5. *Journal of Cell Science* 122, 2371–2382.
- Seaman, M.N.J., Marcusson, E.G., Cereghino, J.-L., Emr, S.D., 1997. Endosome to Golgi retrieval of the vacuolar protein sorting receptor, Vps10p, requires the function of the VPS29, VPS30, and VPS35 gene products. *Journal of Cell Biology* 137 (1), 79–92.
- Seaman, M.N.J., McCaffrey, J.M., Emr, S.D., 1998. A membrane coat complex essential for endosome to Golgi retrograde transport in yeast. *Journal of Cell Biology* 142 (3), 665–681.
- Seaman, M.N.J., Williams, H.P., 2002. Identification of the functional domains of the yeast Sorting Nexins, Vps5p and Vps17p. *Molecular Biology of the Cell* 13, 2840–2856.
- Shi, A., Sun, L., Banerjee, R., *et al.*, 2009. Regulation of endosomal clathrin and retromer-mediated endosome to Golgi retrograde transport by the J-domain protein RME-8. *EMBO Journal* 28 (21), 3290–3302.
- Shi, H., Rojas, R., Bonifacio, J.S., Hurley, J.H., 2006. The retromer subunit Vps26 has an arrestin fold and binds Vps35 through its C-terminal domain. *Nature Structural & Molecular Biology* 13 (6), 540–548.
- Steinberg, F., Gallon, M., Winfield, M., *et al.*, 2013. A global analysis of SNX27-retromer assembly and cargo specificity reveals a function in glucose and metal ion transport. *Nature Cell Biology* 15 (5), 461–471.
- Strochlic, T.I., Setty, T.G., Sitaram, A., Burd, C.G., 2007. Grd19/Snx3p functions as a cargo-specific adapter for retromer-dependent endocytic recycling. *Journal of Cell Biology* 177 (1), 115–125.
- Swarbrick, J.D., Shaw, D.J., Chhabra, S., *et al.*, 2011. VPS29 is not an active metallo-phosphatase but is a rigid scaffold required for retromer interaction with accessory proteins. *PLoS One* 6 (5), e20420.
- Tabuchi, M., Yanatori, I., Kawai, Y., Kishi, F., 2010. Retromer-mediated direct sorting is required for proper endosomal recycling of the mammalian iron transporter DMT1. *Journal of Cell Science* 123 (Pt 5), 756–766.
- Valdmanis, P.N., Meijer, I.A., Reynolds, A., *et al.*, 2007. Mutations in the KIAA0196 gene at the SPG8 locus cause hereditary spastic paraplegia. *American Journal of Human Genetics* 80 (1), 152–161.
- Vardarajan, B., Bruesegem, S.Y., Harbour, M.E., *et al.*, 2012. Identification of Alzheimer disease associated variants in genes that regulate retromer function. *Neurobiology of Aging* 33 (9), 2231.e15–2231.e30.
- Vergès, M., Luton, F., Gruber, C., *et al.*, 2004. The mammalian retromer regulates transcytosis of the polymeric immunoglobulin receptor. *Nature Cell Biology* 6 (8), 763–769.
- Viklund, I.M., Aspenström, P., Meas-Yedid, V., *et al.*, 2008. WAFL, a new protein involved in regulation of early endocytic transport at the intersection of actin and microtubule dynamics. *Experimental Cell Research* 315 (6), 1040–1052.
- Vilariño-Güell, C., Rajput, A., Milnerwood, A.J., *et al.*, 2014. DNAJC13 mutations in Parkinson disease. *Human Molecular Genetics* 23 (7), 1794–1801.
- Vilariño-Güell, C., Wider, C., Ross, O.A., *et al.*, 2011. VPS35 mutations in Parkinson disease. *American Journal of Human Genetics* 89 (1), 162–167.
- Wassmer, T., Attar, N., Bujny, M.V., *et al.*, 2007. A loss-of-function screen reveals SNX5 and SNX6 as potential components of the mammalian retromer. *Journal of Cell Science* 120 (Pt 1), 45–54.
- Wassmer, T., Attar, N., Harterink, M., *et al.*, 2009. The retromer coat complex coordinates endosomal sorting and dynein-mediated transport, with carrier recognition by the trans-Golgi network. *Developmental Cell* 17 (1), 110–122.
- van Weering, J.R., Sessions, R.B., Traer, C.J., *et al.*, 2012. Molecular basis for SNX-BAR-mediated assembly of distinct endosomal sorting tubules. *EMBO Journal* 31 (23), 4466–4480.
- van Weering, J.R., Verkade, P., Cullen, P.J., 2010. SNX-BAR proteins in phosphoinositide-mediated, tubular-based endosomal sorting. *Seminars in Cell & Developmental Biology* 21 (4), 371–380.
- Willnow, T.E., Andersen, O.M., 2013. Sorting receptor SORLA — A trafficking path to avoid Alzheimer disease. *Journal of Cell Science* 126 (Pt 13), 2751–2760.
- Yu, J.W., Lemmon, M.A., 2001. All phox homology (PX) domains from *Saccharomyces cerevisiae* specifically recognize phosphatidylinositol 3-phosphate. *Journal of Biological Chemistry* 276 (47), 44179–44184.
- Zavodszky, E., Seaman, M.N., Moreau, K., *et al.*, 2014. Mutation in VPS35 associated with Parkinson's disease impairs WASH complex association and inhibits autophagy. *Nature Communications* 5, 3828.
- Zhang, J., Reiling, C., Reinecke, J.B., *et al.*, 2012. Rabankyrin-5 interacts with EHD1 and Vps26 to regulate endocytic trafficking and retromer function. *Traffic* 13 (5), 745–757.
- Zimprich, A., Benet-Pagès, A., Struhal, W., *et al.*, 2011. A mutation in VPS35, encoding a subunit of the retromer complex, causes late-onset Parkinson disease. *American Journal of Human Genetics* 89 (1), 168–175.

Vesicle Tethers

LL Marcotte, Assumption College, Worcester, MA, USA

M Munson, University of Massachusetts Medical School, Worcester, MA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

GTPase Protein that binds to and hydrolyzes the nucleotide GTP.

Membrane fusion The process by which two initially distinct lipid bilayers merge to form one interconnected structure.

Multisubunit tethering complexes Large protein complexes consisting of 3–10 polypeptides, that are used for tethering vesicles to target membranes.

SNARE complex A long four-helical-bundle complex consisting of 3–4 distinct polypeptides, which forms the membrane fusion machinery.

Tethering The process by which a protein or protein complex interacts with two distinct membranes to bridge them at long distances and bring them into closer proximity. Tethering proteins or complexes may also facilitate SNARE complex assembly and membrane fusion.

Vesicular transport Trafficking of proteins and lipids from one cellular compartment to another by means of membrane-enclosed intermediates such as vesicles.

Introduction

Vesicular traffic is the principle mechanism in eukaryotic cells by which proteins and other cargo are transported from one membrane-bounded organelle to another, as well as to the plasma membrane for growth and secretion. A vesicle that has budded and been released for transport will diffuse through the cell or travel along a cytoskeletal track to its target membrane. After a vesicle reaches the target membrane, proteins, and protein complexes known as tethering factors are proposed to facilitate the initial linkage of the membranes (for review, see Whyte and Munro, 2002; Sztul and Lupashin, 2006; Cai *et al.*, 2007; Yu and Hughson, 2010; Heider and Munson, 2012; Brocker *et al.*, 2010; Hong and Lev, 2014).

There are two conserved structural classes of tethering complexes: long coiled-coil tethers (Gillingham and Munro, 2003) and multisubunit tethering complexes (MTCs); (Koumandou *et al.*, 2007; Figure 1). Many coiled-coil tethers are associated with trafficking in the early secretory pathway at the ER and the Golgi complex, including p115/Uso1, GMAP210, GCC185, and giantin, or with endosomal trafficking (e.g., EEA1). The MTCs are also associated with ER and Golgi trafficking events (e.g., Dsl1, transport protein particle (TRAPP), and conserved oligomeric Golgi (COG)); additionally, they serve as tethers for endocytic trafficking pathways (homotypic fusion and vacuole protein sorting (HOPS) class C core vacuole/endosome tethering (CORVET), and Golgi-associated retrograde protein (GARP)), as well as for vesicles traveling to the plasma membrane for exocytosis (exocyst). Beyond the canonical secretory pathway, tethers have also been implicated in other dynamic membrane processes, including cytokinesis, autophagy, and ciliogenesis (Heider and Munson, 2012; Solinger and Spang, 2013; Thellmann *et al.*, 2010; Farkas *et al.*, 2003; Rybak *et al.*, 2014). Furthermore, mutations in several of these complexes are associated with a variety of genetic diseases (Ungar, 2009; Miller and Ungar, 2012; Bonifacino and Hierro, 2011; Lobingier and Merz, 2012; Tornieri *et al.*, 2013; Brunet

and Sacher, 2014b; Huang and Lipschutz, 2014 and references therein). The precise functions of these complexes are still not clear; it is thought that, through their interactions with small GTPases, lipids, and soluble NSF attachment protein receptor (SNARE) proteins, they would bridge the vesicle and target membranes to facilitate tethering and promote vesicle fusion at the correct locations. This broad idea has gained much popularity, although tethering has only been directly observed for a few of these complexes (Chia and Gleeson, 2014; Heider and Munson, 2012; Brunet and Sacher, 2014a; Cao *et al.*, 1998; Sacher *et al.*, 2001; Stroupe *et al.*, 2009).

Interactions with Small GTPases and SNARE Proteins

Regulation or recruitment of tethers is modulated by interactions with a number of small GTPases, including those of the Rab, Arf/Arl, and Rho/Ral families (Stenmark, 2009; Barr and Lambright, 2010; Pfeffer, 2013; Donaldson and Jackson, 2011; Wu *et al.*, 2008). These are small prenylated GTP-binding proteins that interact reversibly with membranes. They are activated by guanine nucleotide exchange factors (GEFs) that catalyze the release of GDP and promote the binding of GTP. GTPase-activating proteins (GAPs) accelerate the hydrolysis of bound GTP to GDP to inactivate the small GTPase function. Thus, Rab GTPases act as molecular switches that are active when bound to GTP and deactivated upon GTP hydrolysis.

Best studied are the connections between active Rab GTPases on vesicles, the tethering complexes, and the SNARE proteins on target membranes to mediate accurate vesicle docking and fusion. The tethers specifically recognize vesicles through their interactions with distinct Rab-GTP proteins. Many tether-Rab interactions have been identified, including the interactions between the HOPS complex and the Rab GTPases Ypt7 (Rab7 homolog) and Vps21 (Rab5) (Plemel *et al.*, 2011; Ostrowicz *et al.*, 2010; Brocker *et al.*, 2012), the

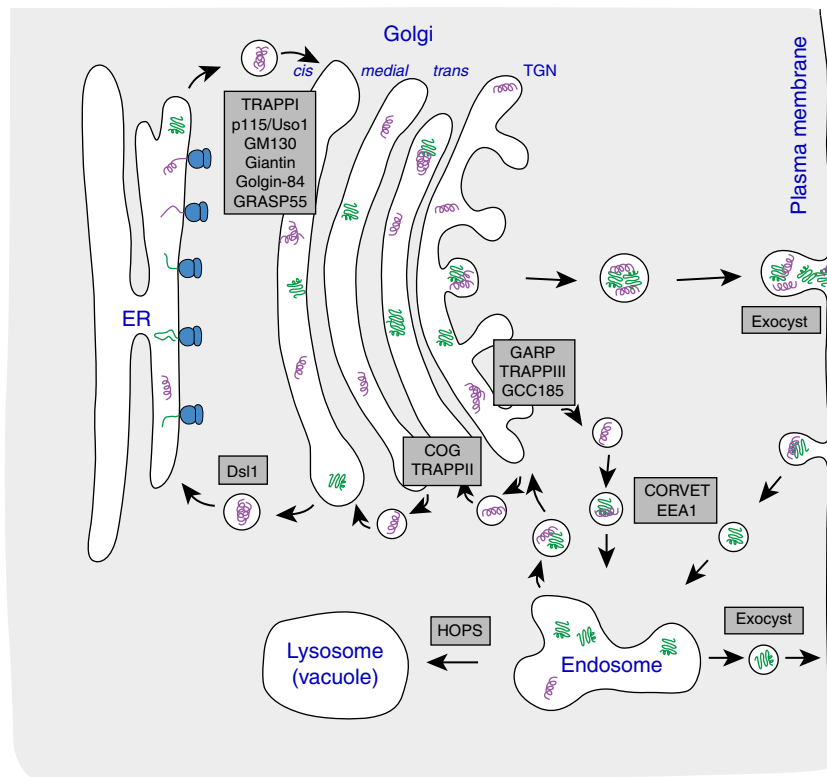


Figure 1 Location of many of the known tethering factors in the eukaryotic membrane trafficking pathways. The names of the coiled-coil proteins and multisubunit tethering complexes are located nearby the trafficking step with which they are associated. The TRAPPIII and exocyst complexes are also implicated in autophagy (not shown).

interaction between GARP and Ypt6 (Siniosoglou and Pelham, 2001), the interaction between the Sec15 subunit of the exocyst complex and the Sec4 GTPase (Rab11) (Guo *et al.*, 1999; Wu *et al.*, 2005), and the COG complex binds multiple Rabs (Miller *et al.*, 2013). The TRAPP complexes also interact with the Rab proteins Ypt1 and Ypt31/32, but appear to function as GEFs to activate these Rabs (Barrowman *et al.*, 2010; Wang *et al.*, 2000; Jones *et al.*, 2000).

To coordinate vesicle recognition with membrane fusion, tethers bind specific SNARE proteins and may regulate SNARE complex assembly and fusion (Hong and Lev, 2014). SNAREs are the core components of the membrane fusion apparatus – SNARE proteins on a transport vesicle (v-SNAREs) pair with cognate SNAREs on the target membrane (t-SNAREs), forming a four-helix bundle complex that brings the two membranes into close apposition for membrane fusion (Sutton *et al.*, 1998; Weber *et al.*, 1998; Sudhof and Rothman, 2009). Although many SNARE-SNARE interactions appear to be specific (McNew *et al.*, 2000), precise temporal and spatial regulation of membrane fusion requires additional layers of regulation ((Furukawa and Mima, 2014) and references therein). Tethering complexes may provide this specificity, as many have been shown to directly interact with partner SNAREs and affect the stability and the rates of SNARE assembly. These include p115 (Shorter *et al.*, 2002), the Dsl1 complex (Diefenbacher *et al.*, 2011; Ren *et al.*, 2009), exocyst (Sivaram *et al.*, 2005), COG (Laufman *et al.*, 2011; Willett *et al.*, 2013), GARP (Perez-Victoria and Bonifacio, 2009), and HOPS and CORVET (Kramer and

Ungermann, 2011; Balderhaar *et al.*, 2013). Regulation of SNARE assembly and fusion may be performed in cooperation with the critical SNARE regulatory Sec1/Munc18 (SM) proteins, which also bind tethering proteins (Morgera *et al.*, 2012; Hong and Lev, 2014). Moreover, several tethers interact with vesicle coat components, suggesting an additional role of tethers in providing specificity prior to SNARE complex assembly (Sztul and Lupashin, 2006; Cai *et al.*, 2007).

Coiled-Coil Tethers

Coiled-coil tethers are large dimeric proteins containing globular heads and long, coiled-coil tails (Gillingham and Munro, 2003). The length and structure of coiled-coil tethers allow them to span distances of more than 200 nm, which is several times the diameter of a typical vesicle. Many coiled-coil tethers, called golgins, associate with the Golgi apparatus and facilitate the trafficking of vesicles traveling from the ER to Golgi or vesicles moving between the cisternae of the Golgi complex. Exemplary members of the golgin family of coiled-coil tethers include p115 (and its yeast homolog Uso1) (Waters *et al.*, 1992; Cao *et al.*, 1998) and its binding partners GM130 (Nakamura *et al.*, 1997) and giantin (Sonnichsen *et al.*, 1998), GCC185 (Luke *et al.*, 2003; Burguete *et al.*, 2008), golgin-84 (Bascom *et al.*, 1999), and GRASP55 (Shorter *et al.*, 1999).

Endosomal trafficking requires coiled-coil tethers such as early endosomal antigen 1 (EEA1); (Simonsen *et al.*, 1998;

Christoforidis *et al.*, 1999; Dumas *et al.*, 2001; Ohya *et al.*, 2009 and references therein). Like other coiled-coil tethers, EEA1 forms a homodimer, and the EEA1 protein contains many different structural domains that impart specific binding properties to the EEA1 complex. For example, the EEA1 homodimer contains N-terminal C₂H₂ zinc finger domains and C-terminal FYVE domains. The FYVE domains bind to phosphatidylinositol 3-phosphate (PI3P); this FYVE-EEA1 interaction recruits EEA1 to the early endosome. The C₂H₂ zinc finger domain of the homodimer binds to the Rab GTPase Rab5, which regulates homotypic tethering (i.e., tethering between endosomes) and fusion of endosomes. Endosomal tethering via EEA1 requires many additional factors, however, and the development of synthetic endosomes has provided new insights into the complexity of tethering and fusion at the endosome. In addition to EEA1 and Rab5, a Rab5 GEF, numerous Rab effectors, lipids, and a variety of SNARE and SNARE-associated proteins are required for efficient homotypic fusion of synthetic endosomes (Ohya *et al.*, 2009).

The structures and multiple binding partners of coiled-coil tethers makes them attractive mediators of tethering. One hypothesis is that coiled-coil proteins of the Golgi may associate to form a 'tentacular cytomatrix' that encapsulates the Golgi apparatus and creates binding sites for vesicles and other molecules that need to associate with the Golgi (Munro, 2011). While some coiled-coil tethers appear to form a vesicle-capturing meshwork, the length of other coiled-coil tethers may facilitate longer-range tethering. For example, p115/Uso1 has a long tail with intervening proline-rich sequences; these proline-rich sequences are relatively flexible and hinge-like, enabling p115/Uso1 to capture vesicles from afar. The p115/Uso1 tether may then bend or collapse upon itself to bring the vesicle and target membrane into close apposition (Yamakawa *et al.*, 1996). Furthermore, the ability of golgins to mediate specificity of vesicle trafficking at the Golgi was recently demonstrated by ectopically targeting different golgins to the mitochondrial membrane – these golgins were able to specifically reroute and capture vesicles bound for the Golgi (Wong and Munro, 2014).

Multisubunit Tethering Complexes

Similar to coiled-coil tethers, MTCs regulate vesicular trafficking at many different stages within the secretory pathway. Nine MTCs (exocyst, COG, CORVET, Dsl1, GARP, HOPS, TRAPPI, TRAPPII, and TRAPPIII) have been identified to date (Koumandou *et al.*, 2007; Brockner *et al.*, 2010; Yu and Hughson, 2010; Perez-Victoria *et al.*, 2010). As the name suggests, each MTC contains several (3–10) different subunits of varying sizes. In general, the individual subunits are more compactly folded than the coiled-coil tethers, which may limit how far the MTCs can reach; EM studies show that several of the MTCs are ~25–30 nm in length (Hsu *et al.*, 1998; Ren *et al.*, 2009; Kim *et al.*, 2006; Yip *et al.*, 2010; Lees *et al.*, 2010; Brockner *et al.*, 2012).

CATCHR Complexes

The exocyst, Dsl1p, COG, and GARP are complexes required for membrane fusion with organelles of the secretory pathway and the plasma membrane. Crystallographic studies of these

complexes showed that subunits from all four complexes have elongated α -helical rod-like structures; these complexes are collectively known as the complexes associated with tethering containing helical rods (CATCHR) family (Yu and Hughson, 2010). The structures of the GARP subunits Vps53 and Vps54 share similar structures to Dsl1, partial structures of Cog2 and Cog4, and four of the eight exocyst subunits, and may have divergently evolved from an ancient ancestor protein (Sivaram *et al.*, 2006; Munson, 2009; Yu and Hughson, 2010). Curiously, similar α -helical bundle structures have been observed for the cargo-binding domain of the type V myosin motors that transport vesicles and RNA (Myo2, Myo4, Myosin V; (Jin *et al.*, 2011)) and the neuronal SNARE regulator Munc13 (Li *et al.*, 2011).

The exocyst is an octomeric complex that regulates many cellular processes, including trafficking and fusion of Golgi-derived vesicles to the plasma membrane, endocytosis, cytokinesis, autophagy, ciliogenesis, and cell migration (for review, see Munson and Novick, 2006; Hertzog and Chavrier, 2011; Liu and Guo, 2011; Heider and Munson, 2012). All of the subunits (Sec5, Sec6, Sec8, Sec10, Sec15, Exo70, and Exo84), except Sec3, are essential in budding yeast, and null mutants of several subunits result in lethality in *Drosophila* and mice, indicating the exocyst's critical role in growth and development. The subunits are localized to sites of polarized growth and secretion. The exocyst subunit Sec15 is proposed to recognize post-Golgi vesicles through interactions with the Rab GTPase Sec4 (Rab11), and the Sec3 and Exo70 subunits bind the plasma membrane through a series of multivalent interactions with phosphatidylinositol 4,5-bisphosphate [PI(4,5)P₂] and the GTPases Cdc42, Rho1, and Rho3. This leads to a model that the assembly of exocyst complexes drives the tethering of secretory vesicles to the plasma membrane (Boyd *et al.*, 2004). Subsequent interactions with the exocytic SNARE proteins and the regulatory protein Sec1 would lead to fusion of the vesicle at proper sites on the plasma membrane.

The Dsl1 complex is located at the ER membrane and binds to Golgi-derived COPI-coated vesicles, and is proposed to function in vesicle tethering, uncoating and SNARE complex assembly (Hughson and Reinisch, 2010; Spang, 2012). Its architecture is simple, as it consists of just three helical subunits – Dsl1, Dsl3/Sec39, and Tip20. A model based on structures of the subunits suggests that the Dsl1 complex is tower-shaped; the base of the tower is anchored to the ER membrane via interactions between the subunits Sec39 and Tip20 with the ER t-SNAREs Sec20 and Use1 (Ren *et al.*, 2009). At the top of the tower is a flexible 'lasso' that can interact with the COPI coat complex of incoming vesicles and tethering these vesicles to the ER.

The COG complex functions as a tether within the Golgi and mediates the retrograde transport of vesicles traveling between the Golgi cisternae (Miller and Ungar, 2012). Mutations in any of its eight subunits (Cog1 to Cog8) cause glycosylation defects and disease in patients and model organisms. The eight subunits are organized into two sets of subunits, or lobes. The first lobe includes Cog2, Cog3, and Cog4; the second lobe consists of Cog5, Cog6, and Cog7. The two lobes are linked by a heterodimer of Cog1 and Cog8 subunits. EM studies show COG as an elongated complex with the subunits assembling into long, thin legs. Although direct

evidence of its role as a tether is still lacking. subunits of the COG complex interact with the Rab GTPases Rab1 (Ypt1) and Rab30, SNARE proteins (e.g., yeast Sed5 and mammalian syntaxin 5), the COPI coat, and SNARE regulators (e.g., the Sec1/Munc18 protein Sly1). COG may also function with coiled-coil tethers of the Golgi to mediate retrograde transport, as interactions between the COG complex and two golgins, p115 and golgin-84, have been observed.

The GARP complex mediates retrograde trafficking from endosomes to the *trans*-Golgi network (TGN) (Bonifacino and Hierro, 2011). GARP consists of four vacuolar protein sorting (Vps) subunits: Vps51, Vps52, Vps53, and Vps54; Vps53, and Vps54 have similar α -helical bundle structures to the other CATCHR family members. GARP, like the other MTCs, also associates with SNARE proteins and small GTPases. In particular, the Vps51 subunit binds to the yeast t-SNARE Tlg1, and the mammalian GARP complex interacts with a host of SNARE proteins including syntaxin 6, syntaxin 16, and Vamp4. GARP also binds to GTP-bound forms of the yeast Rab GTPase Ypt6 and the Arl1 GTPase. The interactions between GARP and the GTPases may mediate GARP binding to the TGN, as depletion of Ypt6 results in the dissociation of GARP from the TGN; however, the functional interplay between GARP and the GTPases has not yet been elucidated.

HOPS and CORVET

The HOPS and CORVET complexes regulate membrane trafficking events at endosomes, vacuoles, and lysosomes (Nickerson *et al.*, 2009; Wickner, 2010; Balderhaar and Ungermann, 2013). Structurally, both complexes are heterohexamers and share a common set of four Class C vacuolar protein sorting (Vps) proteins: Vps11, Vps16, Vps18, and the SM family member Vps33. Each complex has two additional homologous subunits, suggesting that the overall architecture of the HOPS and CORVET complexes is similar. Structurally, all of the Vps-C subunits (with the exception of Vps33) are predicted to contain an amino-terminal β -propeller domain and a carboxy-terminal α -solenoid domain. The Vps11 and Vps18 subunits also contain RING finger zinc-binding domains, potentially mediating protein-protein interactions between the HOPS and CORVET complexes and their binding partners. Interestingly, the predicted Vps-C structures resemble known structures of nuclear pore complex subunits and coat proteins, which also assemble into multisubunit complexes on membranes (Devos *et al.*, 2004).

In addition to the Vps proteins it shares in common with CORVET, the HOPS complex also includes Vps39 and Vps41. It uses these additional subunits, in part, to mediate homotypic fusion of vacuoles. To determine additional factors required for vacuolar fusion, Wickner and colleagues developed a proteoliposome fusion assay to mimic the fusion of two vacuoles *in vitro* (Stroupe *et al.*, 2009). They found that stable tethering and fusion between proteoliposomes required a minimal subset of components: the HOPS complex, the Rab GTPase Ypt7, the vacuolar SNARE proteins (Nyv1, Vti1, Vam3, and Vam7), and the SNARE disassembly proteins Sec17 (α -SNAP) and Sec18 (NSF)/ATP for priming. HOPS appears to coordinate these vacuolar fusion factors by organizing and proofreading SNARE assembly at sites of membrane fusion

and also protects the SNARE complexes from the disassembly proteins Sec17 and Sec18.

The Class C Vps subunits were initially only thought to be part of the HOPS complex; however, these subunits were later identified in association with two non-HOPS subunits, Vps3 and Vps8 (which are homologous to the HOPS subunits Vps39 and Vps41). This alternative complex, CORVET, is required for the fusion of Golgi-derived vesicles with the endosome and homotypic endosomal fusion. CORVET interacts with the early endosomal GTPase Vps21 (Rab5) via both Vps3 and Vps8 and may facilitate membrane fusion by coupling the binding of Vps21 with endosomal SNARE assembly.

TRAPP Complexes

The TRAPP complexes are similar to each other in their composition, but have different roles in vesicular trafficking events (Cai *et al.*, 2007; Sacher *et al.*, 2008; Hughson and Reinisch, 2010; Barrowman *et al.*, 2010). There are three forms of the complex: TRAPPI, TRAPPII, and TRAPPIII. All TRAPP complexes have a common protein core consisting of the proteins Bet3, Bet5, Trs20, Trs23, Trs31, and Trs33. TRAPPI additionally contains the Trs65, Trs120, Trs130, and Tca17/TRAPPC2L subunits, while the TRAPPIII complex is comprised of the core proteins and Trs85. TRAPPI and TRAPPII both function in the secretory pathway in yeast; TRAPPI is required to tether ER-derived vesicles to Golgi membranes, while TRAPPII is postulated to serve as an intra-Golgi complex tether. The recently identified yeast TRAPPIII complex appears to coordinate a membrane tethering event required for autophagy (Lynch-Day *et al.*, 2010).

TRAPPI tethers ER-derived COPII-coated vesicles to the Golgi complex through a direct interaction between the Bet3 subunit of the TRAPPI complex and the vesicle coat protein Sec23. Even though Sec23 is part of the inner layer of the COPII coat, electron microscopy images show that there are gaps in the outer Sec13/31 cage that could allow TRAPPI access to Sec23. The Bet3-Sec23 interaction appears to be essential for vesicle tethering, and this specific interaction allows for the TRAPPI complex to distinguish between COPII-coated vesicles and all other coated vesicles. In addition to their roles as tethers, the TRAPP complexes also act as Rab GEFs for the Rab GTPases Ypt1 and Ypt31/32. Structural data suggests that the central subunits of the TRAPP complexes (Bet3, Bet5, and Trs23) interact with Ypt1 and confer GEF activity. Activation of Ypt1 and Ypt31/32 by the TRAPP complexes may promote tethering and, subsequently, membrane fusion.

Conclusions

The purpose of vesicle tethering is to bring together transport vesicles with their target membranes, perhaps to provide quality control, to prevent diffusion of the vesicle, and to help stimulate membrane fusion. The two major classes of tethers, long coiled-coils and MTCs, share many similar properties, including localization to precise sites of vesicle fusion, binding to specific small GTPases and lipids, and interactions with the SNARE membrane fusion proteins. Many years of genetic and

cell biological experiments identified the players and the relevant membrane trafficking pathways; however, the mechanisms of action of tethers at the molecular level remain to be determined. Indeed, whether or not most of these complexes are in fact tethers needs to be confirmed through direct observation. Higher resolution crystallographic and EM studies are beginning to reveal the structures of these complexes, and our understanding of tethers will benefit greatly from functional reconstitution experiments using purified or biochemically characterized components.

See also: Interorganellar Communication: Components: SNAREs: Membrane Fusion and Beyond. Interorganellar Communication: Interplay and Processes: Endosome to Lysosome Transport; Post-Golgi Transport — Cargo, Carriers, and Pathways

References

- Balderhaar, H.J., Lachmann, J., Yavavli, E., *et al.*, 2013. The CORVET complex promotes tethering and fusion of Rab5/Vps21-positive membranes. *Proceedings of the National Academy of Sciences of the United States of America* 110, 3823–3828.
- Balderhaar, H.J., Ungermann, C., 2013. CORVET and HOPS tethering complexes — Coordinators of endosome and lysosome fusion. *Journal of Cell Science* 126, 1307–1316.
- Barr, F., Lambright, D.G., 2010. Rab GEFs and GAPs. *Current Opinion in Cell Biology* 22, 461–470.
- Barrowman, J., Bhandari, D., Reinisch, K., Ferro-Novick, S., 2010. TRAPP complexes in membrane traffic: Convergence through a common Rab. *Nature Reviews Molecular Cell Biology* 11, 759–763.
- Bascom, R.A., Srinivasan, S., Nussbaum, R.L., 1999. Identification and characterization of golgin-84, a novel Golgi integral membrane protein with a cytoplasmic coiled-coil domain. *Journal of Biological Chemistry* 274, 2953–2962.
- Bonifacio, J.S., Hierro, A., 2011. Transport according to GARP: Receiving retrograde cargo at the trans-Golgi network. *Trends in Cell Biology* 21, 159–167.
- Boyd, C., Hughes, T., Pypaert, M., Novick, P., 2004. Vesicles carry most exocyst subunits to exocytic sites marked by the remaining two subunits, Sec3p and Exo70p. *Journal of Cell Biology* 167, 889–901.
- Brockner, C., Engelbrecht-Vandre, S., Ungermann, C., 2010. Multisubunit tethering complexes and their role in membrane fusion. *Current Biology* 20, R943–R952.
- Brockner, C., Kuhlee, A., Gatsogiannis, C., *et al.*, 2012. Molecular architecture of the multisubunit homotypic fusion and vacuole protein sorting (HOPS) tethering complex. *Proceedings of the National Academy of Sciences of the United States of America* 109, 1991–1996.
- Brunet, S., Sacher, M., 2014a. Are MTCs tethering complexes or trafficking complexes that may act as tethers? *Traffic* 15, 1282–1287.
- Brunet, S., Sacher, M., 2014b. In sickness and in health: The role of TRAPP and associated proteins in disease. *Traffic* 15, 803–818.
- Burguete, A.S., Fenn, T.D., Brunger, A.T., Pfeffer, S.R., 2008. Rab and Arl GTPase family members cooperate in the localization of the golgin GCC185. *Cell* 132, 286–298.
- Cai, H., Reinisch, K., Ferro-Novick, S., 2007. Coats, tethers, Rabs, and SNAREs work together to mediate the intracellular destination of a transport vesicle. *Developmental Cell* 12, 671–682.
- Cao, X., Ballew, N., Barlowe, C., 1998. Initial docking of ER-derived vesicles requires Uso1p and Ypt1p but is independent of SNARE proteins. *EMBO Journal* 17, 2156–2165.
- Chia, P.Z., Gleeson, P.A., 2014. Membrane tethering. *F1000Prime Reports* 6, 74.
- Christoforidis, S., McBride, H.M., Burgoyne, R.D., Zerial, M., 1999. The Rab5 effector EEA1 is a core component of endosome docking. *Nature* 397, 621–625.
- Devos, D., Dokudovskaya, S., Alber, F., *et al.*, 2004. Components of coated vesicles and nuclear pore complexes share a common molecular architecture. *PLoS Biology* 2, e380.
- Diefenbacher, M., Thorsteinsdottir, H., Spang, A., 2011. The Dsl1 tethering complex actively participates in soluble NSF (N-ethylmaleimide-sensitive factor) attachment protein receptor (SNARE) complex assembly at the endoplasmic reticulum in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 286, 25027–25038.
- Donaldson, J.G., Jackson, C.L., 2011. ARF family G proteins and their regulators: Roles in membrane transport, development and disease. *Nature Reviews Molecular Cell Biology* 12, 362–375.
- Dumas, J.J., Merithew, E., Sudharshan, E., *et al.*, 2001. Multivalent endosome targeting by homodimeric EEA1. *Molecular Cell* 8, 947–958.
- Farkas, R.M., Giansanti, M.G., Gatti, M., Fuller, M.T., 2003. The *Drosophila* Cog5 homologue is required for cytokinesis, cell elongation, and assembly of specialized Golgi architecture during spermatogenesis. *Molecular Biology of the Cell* 14, 190–200.
- Furukawa, N., Mima, J., 2014. Multiple and distinct strategies of yeast SNAREs to confer the specificity of membrane fusion. *Scientific Reports* 4, 4277.
- Gillingham, A.K., Munro, S., 2003. Long coiled-coil proteins and membrane traffic. *Biochimica et Biophysica Acta* 1641, 71–85.
- Guo, W., Roth, D., Walch-Solimena, C., Novick, P., 1999. The exocyst is an effector for Sec4p, targeting secretory vesicles to sites of exocytosis. *EMBO Journal* 18, 1071–1080.
- Heider, M.R., Munson, M., 2012. Exorcising the exocyst complex. *Traffic* 13, 898–907.
- Hertzog, M., Chavrier, P., 2011. Cell polarity during motile processes: Keeping on track with the exocyst complex. *Biochemical Journal* 433, 403–409.
- Hong, W., Lev, S., 2014. Tethering the assembly of SNARE complexes. *Trends in Cell Biology* 24, 35–43.
- Hsu, S.C., Hazuka, C.D., Roth, R., *et al.*, 1998. Subunit composition, protein interactions, and structures of the mammalian brain sec6/8 complex and septin filaments. *Neuron* 20, 1111–1122.
- Huang, L., Lipschutz, J.H., 2014. Cilia and polycystic kidney disease, kith and kin. *Birth Defects Research Part C: Embryo Today* 102, 174–185.
- Hughson, F.M., Reinisch, K.M., 2010. Structure and mechanism in membrane trafficking. *Current Opinion in Cell Biology* 22, 454–460.
- Jin, Y., Sultana, A., Gandhi, P., *et al.*, 2011. Myosin V transports secretory vesicles via a Rab GTPase cascade and interaction with the exocyst complex. *Developmental Cell* 21, 1156–1170.
- Jones, S., Newman, C., Liu, F., Segev, N., 2000. The TRAPP complex is a nucleotide exchanger for Ypt1 and Ypt31/32. *Molecular Biology of the Cell* 11, 4403–4411.
- Kim, Y.G., Raunser, S., Munger, C., *et al.*, 2006. The architecture of the multisubunit TRAPP I complex suggests a model for vesicle tethering. *Cell* 127, 817–830.
- Koumandou, V.L., Dacks, J.B., Coulson, R.M., Field, M.C., 2007. Control systems for membrane fusion in the ancestral eukaryote; evolution of tethering complexes and SM proteins. *BMC Evolutionary Biology* 7, 29.
- Kramer, L., Ungermann, C., 2011. HOPS drives vacuole fusion by binding the vacuolar SNARE complex and the Vam7 PX domain via two distinct sites. *Molecular Biology of the Cell* 22, 2601–2611.
- Laufman, O., Hong, W., Lev, S., 2011. The COG complex interacts directly with Syntaxin 6 and positively regulates endosome-to-TGN retrograde transport. *Journal of Cell Biology* 194, 459–472.
- Lees, J.A., Yip, C.K., Walz, T., Hughson, F.M., 2010. Molecular organization of the COG vesicle tethering complex. *Nature Structural & Molecular Biology* 17, 1292–1297.
- Li, W., Ma, C., Guan, R., *et al.*, 2011. The crystal structure of a Munc13 C-terminal domain exhibits a remarkable similarity to vesicle tethering factors. *Structure* 19, 1443–1455.
- Liu, J., Guo, W., 2011. The exocyst complex in exocytosis and cell migration. *Protoplasma* 249, 587–597.
- Lobingier, B.T., Merz, A.J., 2012. Sec1/Munc18 protein Vps33 binds to SNARE domains and the quaternary SNARE complex. *Molecular Biology of the Cell* 23, 4611–4622.
- Luke, M.R., Kjer-Nielsen, L., Brown, D.L., Stow, J.L., Gleeson, P.A., 2003. GRIP domain-mediated targeting of two new coiled-coil proteins, GCC88 and GCC185, to subcompartments of the trans-Golgi network. *Journal of Biological Chemistry* 278, 4216–4226.
- Lynch-Day, M.A., Bhandari, D., Menon, S., *et al.*, 2010. Trs85 directs a Ypt1 GEF, TRAPPIII, to the phagophore to promote autophagy. *Proceedings of the National Academy of Sciences of the United States of America* 107, 7811–7816.
- McNew, J.A., Parlati, F., Fukuda, R., *et al.*, 2000. Compartmental specificity of cellular membrane fusion encoded in SNARE proteins. *Nature* 407, 153–159.
- Miller, V.J., Sharma, P., Kudlyk, T.A., *et al.*, 2013. Molecular insights into vesicle tethering at the Golgi by the conserved oligomeric Golgi (COG) complex and the golgin TATA element modulatory factor (TMF). *Journal of Biological Chemistry* 288, 4229–4240.
- Miller, V.J., Ungar, D., 2012. Re'COG'nition at the Golgi. *Traffic* 13, 891–897.

- Morgera, F., Sallah, M., Dubuke, M., *et al.*, 2012. Regulation of exocytosis by the exocyst subunit Sec6 and the SM protein Sec1. *Molecular Biology of the Cell* 23, 337–346.
- Munro, S., 2011. The golgin coiled-coil proteins of the Golgi apparatus. *Cold Spring Harbor Perspectives in Biology* 3, a005256.
- Munson, M., 2009. Tip20p reaches out to Dsl1p to tether membranes. *Nature Structural & Molecular Biology* 16, 100–102.
- Munson, M., Novick, P., 2006. The exocyst defrocked, a framework of rods revealed. *Nature Structural & Molecular Biology* 13, 577–581.
- Nakamura, N., Lowe, M., Levine, T.P., Rabouille, C., Warren, G., 1997. The vesicle docking protein p115 binds GM130, a cis-Golgi matrix protein, in a mitotically regulated manner. *Cell* 89, 445–455.
- Nickerson, D.P., Brett, C.L., Merz, A.J., 2009. Vps-C complexes: Gatekeepers of endolysosomal traffic. *Current Opinion in Cell Biology* 21, 543–551.
- Ohya, T., Miaczynska, M., Coskun, U., *et al.*, 2009. Reconstitution of Rab- and SNARE-dependent membrane fusion by synthetic endosomes. *Nature* 459, 1091–1097.
- Ostrowicz, C.W., Brocker, C., Ahnert, F., *et al.*, 2010. Defined subunit arrangement and rab interactions are required for functionality of the HOPS tethering complex. *Traffic* 11, 1334–1346.
- Perez-Victoria, F.J., Abascal-Palacios, G., Tascon, I., *et al.*, 2010. Structural basis for the wobbler mouse neurodegenerative disorder caused by mutation in the Vps54 subunit of the GARP complex. *Proceedings of the National Academy of Sciences of the United States of America* 107, 12860–12865.
- Perez-Victoria, F.J., Bonifacio, J.S., 2009. Dual roles of the mammalian GARP complex in tethering and SNARE complex assembly at the trans-Golgi network. *Molecular and Cellular Biology* 29, 5251–5263.
- Pfeiffer, S.R., 2013. Rab GTPase regulation of membrane identity. *Current Opinion in Cell Biology* 25, 414–419.
- Plemel, R.L., Lobingier, B.T., Brett, C.L., *et al.*, 2011. Subunit organization and Rab interactions of Vps-C protein complexes that control endolysosomal membrane traffic. *Molecular Biology of the Cell* 22, 1353–1363.
- Ren, Y., Yip, C.K., Tripathi, A., *et al.*, 2009. A structure-based mechanism for vesicle capture by the multisubunit tethering complex Dsl1. *Cell* 139, 1119–1129.
- Rybak, K., Steiner, A., Synek, L., *et al.*, 2014. Plant cytokinesis is orchestrated by the sequential action of the TRAPP II and exocyst tethering complexes. *Developmental Cell* 29, 607–620.
- Sacher, M., Barrowman, J., Wang, W., *et al.*, 2001. TRAPP I implicated in the specificity of tethering in ER-to-Golgi transport. *Molecular Cell* 7, 433–442.
- Sacher, M., Kim, Y.G., Lavie, A., Oh, B.H., Segev, N., 2008. The TRAPP complex: Insights into its architecture and function. *Traffic* 9, 2032–2042.
- Shorter, J., Beard, M.B., Seemann, J., Dirac-Svejstrup, A.B., Warren, G., 2002. Sequential tethering of Golgins and catalysis of SNAREpin assembly by the vesicle-tethering protein p115. *Journal of Cell Biology* 157, 45–62.
- Shorter, J., Watson, R., Giannakou, M.E., *et al.*, 1999. GRASP55, a second mammalian GRASP protein involved in the stacking of Golgi cisternae in a cell-free system. *EMBO Journal* 18, 4949–4960.
- Simonsen, A., Lippe, R., Christoforidis, S., *et al.*, 1998. EEA1 links PI(3)K function to Rab5 regulation of endosome fusion. *Nature* 394, 494–498.
- Siniouoglou, S., Pelham, H.R., 2001. An effector of Ypt6p binds the SNARE Tlg1p and mediates selective fusion of vesicles with late Golgi membranes. *EMBO Journal* 20, 5991–5998.
- Sivaram, M.V., Furgason, M.L.M., Brewer, D.N., Munson, M., 2006. The structure of the exocyst subunit Sec6p defines a conserved architecture with diverse roles. *Nature Structural & Molecular Biology* 13, 555–556.
- Sivaram, M.V., Saporita, J.A., Furgason, M.L.M., Boettcher, A.J., Munson, M., 2005. Dimerization of the exocyst protein Sec6p and its interaction with the t-SNARE Sec9p. *Biochemistry* 44, 6302–6311.
- Solinger, J.A., Spang, A., 2013. Tethering complexes in the endocytic pathway: CORVET and HOPS. *FEBS Journal* 280, 2743–2757.
- Sonnichsen, B., Lowe, M., Levine, T., *et al.*, 1998. A role for giantin in docking COPI vesicles to Golgi membranes. *Journal of Cell Biology* 140, 1013–1021.
- Spang, A., 2012. The DSL1 complex: The smallest but not the least CATCHR. *Traffic* 13, 908–913.
- Stenmark, H., 2009. Rab GTPases as coordinators of vesicle traffic. *Nature Reviews Molecular Cell Biology* 10, 513–525.
- Stroupe, C., Hickey, C.M., Mima, J., Burfeind, A.S., Wickner, W., 2009. Minimal membrane docking requirements revealed by reconstitution of Rab GTPase-dependent membrane fusion from purified components. *Proceedings of the National Academy of Sciences of the United States of America* 106, 17626–17633.
- Sudhof, T.C., Rothman, J.E., 2009. Membrane fusion: Grappling with SNARE and SM proteins. *Science* 323, 474–477.
- Sutton, R.B., Fasshauer, D., Jahn, R., Brunger, A.T., 1998. Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature* 395, 347–353.
- Sztl, E., Lupashin, V., 2006. Role of tethering factors in secretory membrane traffic. *American Journal of Physiology – Cell Physiology* 290, C11–C26.
- Theilmann, M., Rybak, K., Thiele, K., Wanner, G., Assaad, F.F., 2010. Tethering factors required for cytokinesis in Arabidopsis. *Plant Physiology* 154, 720–732.
- Tornieri, K., Zlatić, S.A., Mullin, A.P., *et al.*, 2013. Vps33b pathogenic mutations preferentially affect VIPAS39/SPE-39-positive endosomes. *Human Molecular Genetics* 22, 5215–5228.
- Ungar, D., 2009. Golgi linked protein glycosylation and associated diseases. *Seminars in Cell and Developmental Biology* 20, 762–769.
- Wang, W., Sacher, M., Ferro-Novick, S., 2000. TRAPP stimulates guanine nucleotide exchange on Ypt1p. *Journal of Cell Biology* 151, 289–296.
- Waters, M.G., Clary, D.O., Rothman, J.E., 1992. A novel 115-kD peripheral membrane protein is required for intercisternal transport in the Golgi stack. *Journal of Cell Biology* 118, 1015–1026.
- Weber, T., Zemelman, B.V., McNew, J.A., *et al.*, 1998. SNAREpins: Minimal machinery for membrane fusion. *Cell* 92, 759–772.
- Whyte, J.R., Munro, S., 2002. Vesicle tethering complexes in membrane traffic. *Journal of Cell Science* 115, 2627–2637.
- Wickner, W., 2010. Membrane fusion: Five lipids, four SNAREs, three chaperones, two nucleotides, and a Rab, all dancing in a ring on yeast vacuoles. *Annual Review of Cell and Developmental Biology* 26, 115–136.
- Willett, R., Kudlyk, T., Pokrovskaya, I., *et al.*, 2013. COG complexes form spatial landmarks for distinct SNARE complexes. *Nature Communications* 4, 1553.
- Wong, M., Munro, S., 2014. Membrane trafficking. The specificity of vesicle traffic to the Golgi is encoded in the golgin coiled-coil proteins. *Science* 346, 1256898.
- Wu, S., Mehta, S.Q., Pichaud, F., Bellen, H.J., Quirocho, F.A., 2005. Sec15 interacts with Rab11 via a novel domain and affects Rab11 localization *in vivo*. *Nature Structural & Molecular Biology* 12, 879–885.
- Wu, H., Rossi, G., Brennwald, P., 2008. The ghost in the machine: Small GTPases as spatial regulators of exocytosis. *Trends in Cell Biology* 18, 397–404.
- Yamakawa, H., Seog, D.H., Yoda, K., Yamasaki, M., Wakabayashi, T., 1996. Uso1 protein is a dimer with two globular heads and a long coiled-coil tail. *Journal of Structural Biology* 116, 356–365.
- Yip, C.K., Berscheminski, J., Walz, T., 2010. Molecular architecture of the TRAPP II complex and implications for vesicle tethering. *Nature Structural & Molecular Biology* 17, 1298–1304.
- Yu, I.M., Hughson, F.M., 2010. Tethering factors as organizers of intracellular vesicular traffic. *Annual Review of Cell and Developmental Biology* 26, 137–156.

Further Reading

- Balderhaar, H.J., Ungermann, C., 2013. CORVET and HOPS tethering complexes – Coordinators of endosome and lysosome fusion. *Journal of Cell Science* 126, 1307–1316.
- Barrowman, J., Bhandari, D., Reinisch, K., Ferro-Novick, S., 2010. TRAPP complexes in membrane traffic: Convergence through a common Rab. *Nature Reviews Molecular Cell Biology* 11, 759–763.
- Brocker, C., Engelbrecht-Vandre, S., Ungermann, C., 2010. Multisubunit tethering complexes and their role in membrane fusion. *Current Biology* 20, R943–R952.
- Cai, H., Reinisch, K., Ferro-Novick, S., 2007. Coats, tethers, Rabs, and SNAREs work together to mediate the intracellular destination of a transport vesicle. *Developmental Cell* 12, 671–682.
- Heider, M.R., Munson, M., 2012. Exorcising the exocyst complex. *Traffic* 13, 898–907.
- Hong, W., Lev, S., 2014. Tethering the assembly of SNARE complexes. *Trends in Cell Biology* 24, 35–43.
- Hughson, F.M., Reinisch, K.M., 2010. Structure and mechanism in membrane trafficking. *Current Opinion in Cell Biology* 22, 454–460.
- Miller, V.J., Ungar, D., 2012. ReCOGNition at the Golgi. *Traffic* 13, 891–897.
- Munro, S., 2011. The golgin coiled-coil proteins of the Golgi apparatus. *Cold Spring Harbor Perspectives in Biology* 3, 1–10.
- Perez-Victoria, F.J., Bonifacio, J.S., 2009. Dual roles of the mammalian GARP complex in tethering and SNARE complex assembly at the trans-Golgi network. *Molecular and Cellular Biology* 29, 5251–5263.
- Solinger, J.A., Spang, A., 2013. Tethering complexes in the endocytic pathway: CORVET and HOPS. *FEBS Journal* 280, 2743–2757.
- Sztl, E., Lupashin, V., 2006. Role of tethering factors in secretory membrane traffic. *American Journal of Physiology – Cell Physiology* 290, C11–C26.
- Yu, I.M., Hughson, F.M., 2010. Tethering factors as organizers of intracellular vesicular traffic. *Annual Review of Cell and Developmental Biology* 26, 137–156.

BAR Domains and BAR Domain Superfamily Proteins

M Sharma, Indian Institute of Science Education and Research Mohali, Manauli, Punjab, India
S Caplan, The University of Nebraska Medical Center, Omaha, NE, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Eukaryotic cells contain multiple membrane-bound organelles that often require great flexibility in their shape. Accordingly, proteins that are capable of inducing changes in membrane shape are crucial for the organization and function of intracellular organelles, and thus influence important cellular events including endocytic and membrane trafficking, cell cycle, division, and migration. The BIN-Amphiphysin-Rvs (BAR) domain is a sequence of ~200 residues found in proteins that directly influence membrane curvature and shaping. Over 750 BAR domain-containing proteins have been identified so far, in species as diverse as yeast and mammals. However, the primary sequence of BAR domains can vary significantly so that crystal structures were needed initially to confirm that certain (BAR) domains indeed belonged to the same BAR domain family (Tarricone *et al.*, 2001; Peter *et al.*, 2004).

BAR domains typically contain a three-helix coiled-coil region that facilitates dimerization and a positively charged module that is capable of both membrane curvature-sensing as well as membrane binding (Peter *et al.*, 2004). Indeed, loss of dimerization ability renders BAR domain-containing proteins incapable of inducing membrane curvature (Dislich *et al.*, 2011).

Given the wide assortment of membranes that BAR domains must remodel, it is not surprising that these domains have been categorized into subsets, largely dependent on their dimeric structure and ability to induce varying degrees of membrane curvature (reviewed in Frost *et al.*, 2009). In general, the (regular) BAR domains have been divided into 'Classical BAR' domains, 'N-BAR' (amphipathic N-terminal helix) domains, 'BAR-PH' (BAR + pleckstrin homology) domains, and 'PX-BAR' (phox domain-containing) domains (see Figure 1). Whereas the classical BAR domains have the highest level of intrinsic curvature, the N-BAR, pleckstrin homology-containing BAR proteins (PH-BAR), and PX-BAR domains all contain an additional membrane binding domain (e.g., the phosphatidylinositol-4,5-bisphosphate-binding PH domain) (Figure 2).

The largest family of BAR domains is the F-BAR (FCH-BAR-Fes/CIP4 homology-BAR) domain group (Figure 3). These F-BAR domains induce positive curvature (similar to BAR domains), but individual members differ greatly in their degree of curvature induction. In contrast, the I-BAR proteins (Figure 4) generate negative curvature, allowing the cell to extrude membranes (Ahmed *et al.*, 2010), whereas the PinkBAR (planar intestinal- and kidney-specific BAR) domain has no intrinsic curvature and is thought to serve in scaffolding other proteins on flat membrane surfaces (Pykäläinen *et al.*, 2011).

Key BAR Domain-Containing Proteins and Their Cellular Functions

Many BAR domain superfamily proteins contain additional domains that regulate their intracellular functions. For example,

the SH3 domain recognizes proline-rich domains (PRDs) found in actin nucleation promoting factors (NPFs) and in the endocytic regulatory GTPase, dynamin (Suetsugu and Gautreau, 2012). Some BAR domain proteins (endophilin, amphiphysin, and syndapin/PACSIN) also bind to the lipid phosphatase synaptojanin via their SH3 domain (Milosevic *et al.*, 2011; Qualmann *et al.*, 1999). The interaction of BAR domain proteins with NPFs induces membrane protrusions and lamellipodia/filopodia generation, or invaginations leading to clathrin-coated pits (CCPs) and/or caveolae formation. Other combinations of functional domains present in the BAR domain superfamily include the PH and PX domains (as noted above), catalytic domains such as Rho family GTPase guanine exchange factor (GEF) and GTPase activation protein (GAP) domain, tyrosine (Tyr)-kinase domain as well as many protein-protein interaction domains (see Table 1 for a list of well-characterized BAR domain superfamily proteins).

Role of BAR Proteins in Regulating NPF-Mediated Actin Assembly

A dynamic actin cytoskeleton is required to generate forces and induce cell membrane deformation to facilitate various developmental processes, including collective cell migration and tissue formation. Moreover, the fundamental process of endocytosis also requires force generated by the actin cytoskeleton to drive membrane invagination, constricting the neck of these newly formed vesicles and driving the vesicles away from their site of formation (Romer *et al.*, 2010). The rate-limiting step for actin filament polymerization is nucleation or formation of stable actin multimers. This requires the presence of nucleation factors such as the actin-related proteins 2/3 (Arp2/3) complex, formins, and tandem-monomer-binding nucleators (Pollard, 2007). However, in isolation the Arp2/3 complex is an ineffective nucleator that requires binding to preexisting actin filaments and to NPFs. Mammalian cells express several NPFs, such as Wiskott-Aldrich syndrome protein (WASP), neuronal WASP (N-WASP), WASP and verprolin homologs (WAVEs), WASP homolog-associated with actin, membranes and microtubules (WHAMM), WASP and Scar homolog (WASH), and junction mediating regulatory (JMY) protein (Takenawa and Suetsugu, 2007).

NPFs of the WASP and WAVE families contain a conserved C-terminal verprolin-homology domain-cofilin-homology domain-acidic domain (VCA) that is critical for their role in actin filament assembly. Actin nucleators such as Arp2/3 complex are activated by binding to the CA region of NPFs and to the side of actin filaments (Padrick and Rosen, 2010). Tissue-specific NPFs, such as WASP (which is expressed in hematopoietic cells), and ubiquitously expressed N-WASP normally adopt an auto-inhibited conformation in which the VCA region displays an intramolecular interaction with its central GTPase-binding domain rendering it inaccessible to the

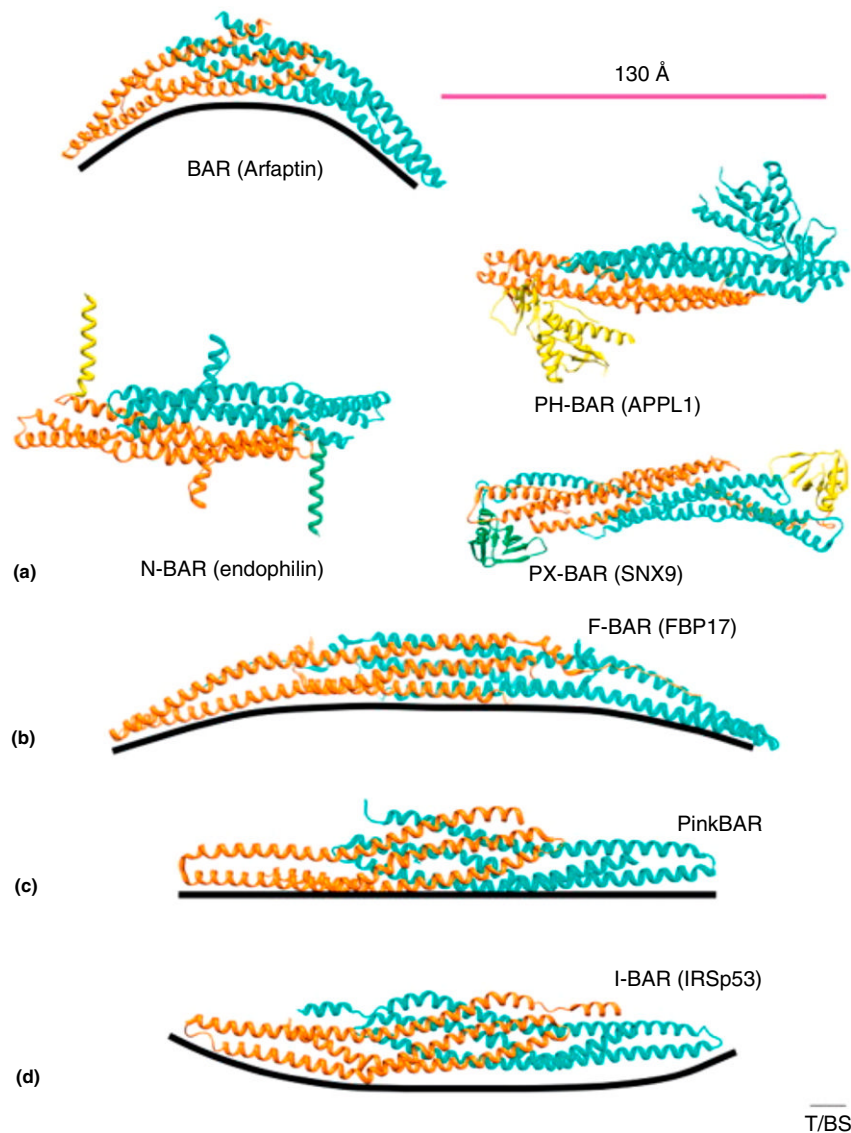


Figure 1 Structures of BAR domains from different subfamilies. All structures are depicted as dimers (cyan and orange) (scale bar 130 Å). The black line underneath each structure was introduced to show the intrinsic curvature of each domain. (a) Left, above: side view of the BAR domain of arfaptin 1 (pdb ID 1I49). BAR domains exhibit the highest positive curvature within the BAR superfamily. Classic BAR domains are further subdivided based on additional membrane binding domains. For clarity, different BAR domains are shown as top view. The N-BAR domain contains a N-terminal helix (yellow and green) and a second set of amphipathic helices known as 'insert helices' (orange and cyan). APPL1 (right, above) is an example of a PH-BAR domain that contains an additional pleckstrin homology domain (PH) (pdb ID 2Z00). Lastly, PX-BAR proteins (right, below), such as sorting nexin 9 (SNX9, pdb ID 2RAI), are comprised of a phox-homology domain (PX) and a BAR domain. (b) Side view of the F-BAR domain from FBP (pdb ID 2EFL) exemplifies a shallow, positive curvature that is typical for this subfamily. (c) The Pinkbar is the only instance of a BAR protein that promotes zero curvature; its BAR domain shows no intrinsic curvature (pdb ID 3OK8). (d) I-BAR domain proteins are the only subfamily that can bend membranes to generate structures with negative curvature (IRSp53, pdb ID 1Y20). Reprinted with permission from Mim, C., Unger V.M., 2012. Membrane curvature and its generation by BAR proteins. Trends in Biochemical Sciences 37 (12), 526–533.

Arp2/3 complex (Padrick and Rosen, 2010). This inactive conformation is further stabilized by interaction of WASP with WASP-interacting protein (WIP) family members (de la Fuente *et al.*, 2007). Interaction of WASP/N-WASP with phosphoinositides and cell division control protein 42 homolog (Cdc42), a member of the Rho GTPase family, is crucial to relieving this auto-inhibition and in turn for stimulate Arp2/3 complex-mediated actin polymerization (Rohatgi *et al.*, 1999).

To further delineate how factors such as Cdc42 are physiologically regulated, Ho and colleagues identified a novel protein, Toca-1 (transducer of Cdc42-dependent actin assembly), that was required for the full Cdc42-induced activation of the WIP–WASP complex (Ho *et al.*, 2004). Toca-1 is a member of the F-BAR family of proteins that has two other homologs, CIP4 (cell division cycle 42-interacting protein 4) and FBP17 (formin-binding protein 17) (Itoh *et al.*, 2005).

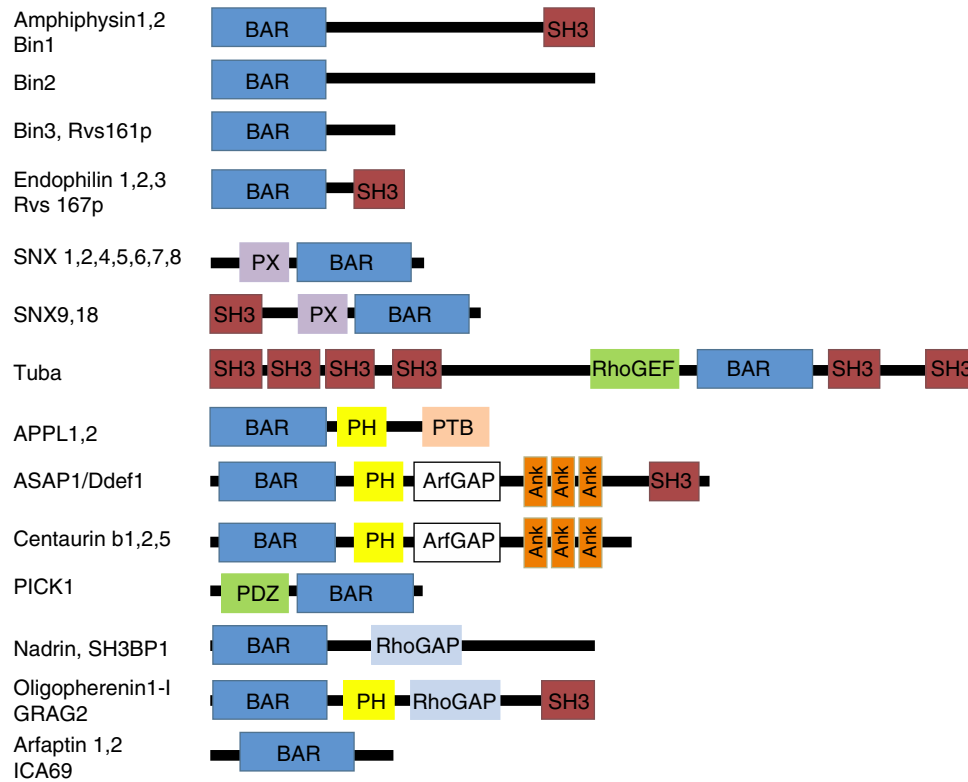


Figure 2 Bar domain-containing proteins. The domain organization of various BAR subfamily proteins is depicted. BAR, Bin-Amphiphysin-Rvs domain; SH3, Src homology 3 domain; RhoGAP, Rho GTPase activating protein domain; PX, Phox-homology domain; PH, Pleckstrin Homology domain; ArfGAP, Arf Rho GTPase activating protein domain; PTB, Phosphotyrosine binding domain; PDZ, Ptd-95, Dlg and Z01 domain; RhoGEF, Rho guanine-nucleotide exchange factors domain; Ank, Ankyrin. Reprinted with permission from Safari, F., Suetsugu S., 2012. The BAR domain superfamily proteins from subcellular structures to human diseases. *Membranes* 2 (1), 91–117.

The presence of the F-BAR domain is the hallmark of this family of proteins and CIP4 was one of the first vertebrate F-BAR proteins identified (Gyuris *et al.*, 1993). Besides their N-terminal F-BAR domains, these proteins contain a central HR1 (protein kinase C-related kinase homology region 1) domain that binds to Rho GTPases, and a C-terminal SH3 domain that recognizes PRD regions in NPFs, including WASP/N-WASP (Ho *et al.*, 2004; Aspenstrom 1997). The importance of NPF binding to SH3 domains of F-BAR proteins is reflected by the fact that mutations of the PRDs, cause X-linked neutropenia by disrupting the auto-inhibition of WASP (Devriendt *et al.*, 2001).

Based on recent studies, it has been suggested that the F-BAR-mediated NPF activation is strongly enhanced in the presence of liposomes with specific curvature, and surprisingly activation can occur even in the absence of Cdc42 (Takano *et al.*, 2008). This leads to the notion that the primary role of F-BAR proteins is to recruit the NPFs to membranes of specific curvature rather than playing a direct role in their activation. EFC/FCH-BAR homologs from *Drosophila melanogaster*, *Caenorhabditis elegans*, and vertebrates, also bind and recruit WASP, as well as the related WASP family WAVE/suppressor of cAR (SCAR), to membranes (Fricke *et al.*, 2009; Giuliani *et al.*, 2009). Several studies have demonstrated the role of the CIP4 family in diverse cellular processes, including cell migration, actin pedestals induced by intracellular bacteria, formation

of podosomes, filopodia, and phagocytic cup biogenesis in macrophages (Tsuboi *et al.*, 2009; Pichot *et al.*, 2010; Campellone *et al.*, 2012). The role of this family in clathrin-mediated endocytosis is discussed below. Several other F-BAR proteins, including PACSIN, srGAPs, PSTPIPs, and nostrin, bind WASP family members via their SH3 domains and regulate actin remodeling (Suetsugu and Gautreau, 2012). PACSIN (protein kinase C and casein kinase 2 substrate in neurons)/syndapin was initially identified as a dynamin interaction partner; however later studies showed that its SH3 domain can bind to both dynamin and N-WASP (Qualmann *et al.*, 1999; Qualmann and Kelly, 2000). Syndapin has three isoforms that are expressed in a tissue-specific manner and regulate endocytosis (Modregger *et al.*, 2000). Syndapin-N-WASP complexes can drive local cortical actin polymerization, as well as induce filopodia formation at the cell surface. The primary function of syndapins is thought to couple dynamin-mediated vesicle fission with Arp2/3-complex-dependent F-actin nucleation, thus linking endocytosis and actin dynamics (as discussed below). Syndapins have also been linked to other endocytic pathways and trafficking events in the cell, including caveolae-mediated endocytosis, transport at the trans-Golgi network and endocytic recycling (Senju *et al.*, 2011; Kessels *et al.*, 2006; Giridharan *et al.*, 2013).

Unlike CIP4 family members that activate N-WASP by recruiting both N-WASP and Cdc42 in close proximity to each

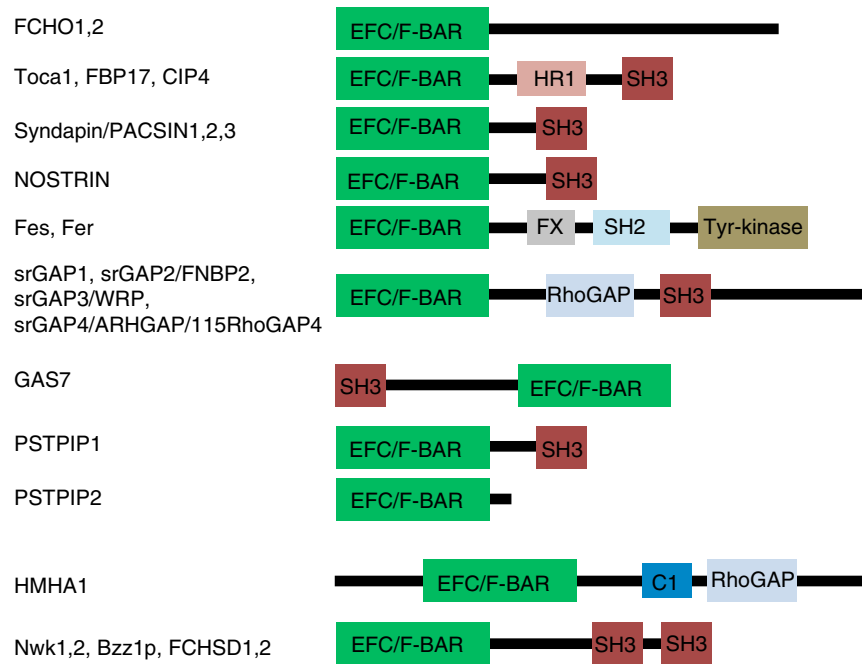


Figure 3 F-BAR domain-containing proteins. The domain organization of various F-BAR subfamily proteins is depicted. Domain organization of F-BAR proteins: C1, protein kinase C conserved region 1; EFC, extended FER-CIP4 homology (EFC) or FCH and BAR (F-BAR) domain; FX, F-BAR extension domain; HR1, Protein kinase C-related kinase homology region 1 domain; RhoGAP, Rho GTPase activating protein domain; SH2, Src homology 2 domain; SH3, Src homology 3 domain; Tyr-kinase, Tyrosine kinase domain. Reprinted with permission from Safari, F., Suetsugu S., 2012. The BAR domain superfamily proteins from subcellular structures to human diseases. *Membranes* 2 (1), 91–117.



Figure 4 I-BAR domain-containing proteins. The domain organization of various I-BAR subfamily proteins is depicted. Schematic of domain organization of I-BAR proteins. CRIB: Cdc42-Rac interactive binding region; IRSp53-MIM homology domain (IMD)/inverse-BAR (I-BAR); SH3: Src homology 3 domain; WH2: Wasp-homology 2 (verprolin-homology) domain. Reprinted with permission from Safari, F., Suetsugu S., 2012. The BAR domain superfamily proteins from subcellular structures to human diseases. *Membranes* 2 (1), 91–117.

other, members of the srGAP (Slit-Robo Rho GTPase activating protein) family of F-BAR proteins negatively regulate actin polymerization by inactivating the Rho GTPases due to the presence of the RhoGAP domain. The srGAPs were originally identified as proteins interacting with the intracellular domain of Robo, receptor for the neuronal repellent factor, Slit (Wong *et al.*, 2001). The role of the Slit-Robo system has been demonstrated in repulsion of motor and olfactory bulb axons in mammals, repulsion of migratory neurons, and inhibition of leukocyte chemotaxis (Wu *et al.*, 2001). The srGAP family consists of three members, srGAP1, srGAP2, and srGAP3 (Wong *et al.*, 2001). *In vitro* studies indicate that srGAP1 modulates the Slit-Robo-dependent repulsive cues in neuronal

migration through inactivation of Cdc42 (Wong *et al.*, 2001). SrGAP2 negatively regulates neuronal migration through the ability of its I-BAR domain to induce filopodia formation (Guerrier *et al.*, 2009). SrGAP3 (also known as mental-disorder associated GAP protein) can regulate Rac and Cdc42 activity and inhibit Rac-dependent neurite outgrowth and may be involved in severe X-linked mental retardation (Endris *et al.*, 2002).

Typically, activation of WASP/WAVE family member NPFs requires their binding to Rho GTPases via their central GTPase-binding domain. However, in T lymphocytes, T cell antigen receptor-induced actin polymerization occurred normally in WASP^{-/-} mice expressing a WASP transgene lacking

Table 1 Well-characterized BAR domain proteins and their functions

BAR domain type	BAR protein	SH3 domain	Other functional domains (besides BAR and SH3)	Binding to dynamin	Binding to actin NPFs	Cellular function	References
BAR/N-BAR	Amphiphysin I, amphiphysin II/Bin1	Yes		Yes	Yes (N-WASP)	Clathrin-mediated endocytosis, T-tubule biogenesis, endocytic recycling	Peter <i>et al.</i> (2004); Qualmann <i>et al.</i> (2000); David <i>et al.</i> (1996); McMahon <i>et al.</i> (1997); Farsad <i>et al.</i> (2003); Shupliakov <i>et al.</i> (1997); Pant <i>et al.</i> (2009); Wigge <i>et al.</i> (1997); Di Paolo <i>et al.</i> (2002)
	Endophilin 1, 2, 3	Yes		Yes	Yes (N-WASP)	Clathrin-mediated endocytosis	Peter <i>et al.</i> (2004); David <i>et al.</i> (1996); Gad <i>et al.</i> (2000); Verstreken <i>et al.</i> (2003); Schuske <i>et al.</i> (2003)
	SNX9, 18, 33	Yes	PX	Yes	Yes (N-WASP)	Clathrin-mediated endocytosis	Lundmark and Carlsson (2004); Worby and Dixon (2002); Yasar <i>et al.</i> (2007)
	SNX1, 2, 3, 4, 5, 6, 7, 8	No	PX	Yes	Yes (WASH)	Endosome trafficking	Worby and Dixon (2002); Kurten <i>et al.</i> (1996); Griffin <i>et al.</i> (2005)
	APPL1, APPL2	No	PH, PTB	No	N.D.	Signal transduction, regulates Akt substrate specificity, cell survival	Miaczynska <i>et al.</i> (2004)
	PICK1	No	PDZ	No	N.D. but directly binds to Arp2/3 complex	Trafficking and Synaptic targeting of neuronal receptors	Staudinger <i>et al.</i> (1995); Höhne <i>et al.</i> (2011); Jin <i>et al.</i> (2006); Hanley (2008)
F-BAR	ASAP1	Yes	PH, ArfGAP, Ankyrin	No	N.D. shown to bind cortactin	Podosome and invadopodia formation	Kam <i>et al.</i> (2000); Brown <i>et al.</i> (1998)
	Tuba	Yes	DH/RhoGEF	Yes	Yes (N-wasp, Ena/VASP)	Apical junction formation in epithelial cells	Leibfried <i>et al.</i> (2008); Salazar <i>et al.</i> (2003); Otani <i>et al.</i> (2006); Kovacs <i>et al.</i> (2006)
	FCHo1, 2	No		No	No	Clathrin-mediated endocytosis	Henne <i>et al.</i> (2007); Henne <i>et al.</i> (2010); Umasankar <i>et al.</i> (2012)
	CIP4, Toca-1, FBP17	Yes	HR1	Yes	Yes (N-WASP)	Clathrin-mediated endocytosis, filopodia formation	Ho <i>et al.</i> (2004); Itoh <i>et al.</i> (2005); Gyuris <i>et al.</i> (1993); Aspenstrom (1997); Takano <i>et al.</i> (2008); Fricke <i>et al.</i> (2009); Giuliani <i>et al.</i> (2009); Tsuboi <i>et al.</i> (2009); Pichot <i>et al.</i> (2010)
	PACSIN/ Syndapin I, II, III	Yes		Yes	Yes (N-WASP)	Clathrin- and caveolin-mediated endocytosis, filopodia formation, endocytic recycling	Qualmann <i>et al.</i> (1999); Qualmann <i>et al.</i> (2000); Modregger <i>et al.</i> (2000); Senju <i>et al.</i> (2011); Kessels <i>et al.</i> (2006); Giridharan <i>et al.</i> (2013); Andersson <i>et al.</i> (2008); Braun <i>et al.</i> (2005)
	srGAP1, 2, 3	Yes	RhoGAP	N.D.	Yes (WAVE, N-WASP)	Filopodium and lamellipodia formation	Wong <i>et al.</i> (2001); Wu <i>et al.</i> (2001); Guerrier <i>et al.</i> (2009); Endris <i>et al.</i> (2002)

(Continued)

Table 1 Continued

BAR domain type	BAR protein	SH3 domain	Other functional domains (besides BAR and SH3)	Binding to dynamin	Binding to actin NPFs	Cellular function	References
	Fes/Fps, Fer	Yes	FX, SH2, Tyr-kinase	N.D.		Non receptor tyrosine kinase, implicated in cellular transformation	Beemon (1981); Greer (2002); Brooks-Wilson <i>et al.</i> (1989); Zirnigbi <i>et al.</i> (2001); Piedra <i>et al.</i> (2003); Iwanishi <i>et al.</i> (2000); Filippakopoulos <i>et al.</i> (2008); Maru <i>et al.</i> (1995)
	Nostrin	Yes		Yes	Yes (N-WASP)	eNOS trafficking, angiogenesis	Icking <i>et al.</i> (2005); Kovacevic <i>et al.</i> (2012)
	GAS7	Yes		N.D.	Yes (N-WASP)	Neurite outgrowth and formation of cellular protrusions	You and Lin-Chao (2010); She <i>et al.</i> (2002)
	PSTPIP1	Yes		N.D.	Yes (WASP)	Inhibits WASP-mediated actin polymerization in immune cells, filopodia formation	Badour <i>et al.</i> (2004); Côté <i>et al.</i> (2002); Chitu <i>et al.</i> (2005); Wise <i>et al.</i> (2002)
I-BAR	IRSp53	Yes	CRIB, WH2-like	Yes	Yes (WAVE, N-WASP, VASP, Mena)	Filopodium and lamellipodia formation	Suetsugu <i>et al.</i> (2006); Miki <i>et al.</i> (2000); Lim <i>et al.</i> (2008) Scita <i>et al.</i> (2008)
	PINKBAR	Yes	WH2-like	N.D.		Formation of planar membrane sheets	Pykäläinen <i>et al.</i> (2011)
	MIM and ABBA	No	WH2-like	No	Cortactin	Filopodium formation, cell polarization	Mattila <i>et al.</i> (2003); Lee <i>et al.</i> (2002); Xie <i>et al.</i> (2011)

Abbreviations: ABBA, actin-bundling protein with BAIAP2 homology; ASAP1, ArfGAP with SH3 domain, ankyrin repeat and PH domain; BAR, BIN-Amphiphysin-Rvs; CIP4, cell division cycle 42-interacting protein 4; DH, Dbl homology; EFC/FCH-BAR, extended FCH homology-BAR; F-BAR, Fes/CIP4 homology-BAR; FBP17, formin-binding protein 17; FCHO1/2, Fer/CIP4 homology domain-only proteins 1 and 2; Fer, Fes-related protein; Ies, teline sarcoma; Fps/Fes, Fujinami poultry sarcoma; FX, F-BAR extension domain; GAP, GTPase activation protein; GAS7, growth arrest-specific protein; GEF, guanine exchange factor; I-BAR, inverted-BAR; IRSp53, insulin receptor substrate protein of 53 kDa; MIM, missing-in-metastasis; N-BAR, N-terminal amphipathic helix-BAR; NPFs, nucleation promoting factors; N-WASP, neuronal WASP; PACSIN, protein kinase C and casein kinase 2 substrate in neurons; PH, pleckstrin Homology; PICK1, protein interacting with C-Kinase 1; PSTPIPs, proline-serine-threonine phosphatase interacting protein 1; PX, Phox-domain; SCAR, suppressor of cAr; SH3, Src homology 3 domain; SNX, sorting nexin; sGAP, SLIT-ROBO Rho GTPase activating protein; Toca-1, transducer of Cdc42-dependent actin assembly; VASP, vasodilator-stimulated phosphoprotein; VCA, verprolin-homology domain—cofilin-homology domain—acidic domain; WASH, WASP and Scar Homolog; WASP, Wiskott-Aldrich syndrome protein; WAVES, WASP and verprolin homologs.

the cdc42-binding domain. In contrast, the phosphorylation of a tyrosine residue of WASP by Fyn kinase was absolutely required for WASP effector activities in T cells (Badour *et al.*, 2004). The exclusively hemopoietic F-BAR family member, PSTPIP1 (proline-serine-threonine phosphatase interacting protein 1), interacts with WASP via its SH3 domain and regulates phosphorylation-dependent activation of WASP in T lymphocytes (Côté *et al.*, 2002). PSTPIP1 is a cytoskeletal scaffold protein that binds to protein tyrosine phosphatase (PTP)-PEST and allows the phosphatase to dephosphorylate WASP, thus inhibiting WASP-mediated actin polymerization and formation of the immunological synapse (Badour *et al.*, 2004; Côté *et al.*, 2002). PSTPIP1 also regulates motility of neutrophil-like cells and its ectopic expression drives filopodia-like membrane extensions (Chitu *et al.*, 2005). Mutations in PSTPIP1 cause PAPA (pyogenic sterile arthritis, pyoderma gangrenosum, and acne) syndrome, an autoinflammatory disease with severe skin and joint involvement (Wise *et al.*, 2002). The molecular basis of this disease and how PSTPIP1 is linked to inflammatory signaling is not understood. Other F-BAR proteins such as GAS7 (growth arrest-specific protein) and Nostrin also contain SH3 domain that bind to N-WASP and regulate actin dynamics in the cell (Suetsugu and Gautreau, 2012). GAS7 uniquely contains an N-terminal SH3 domain and is required for neurite outgrowth in cerebellar neurons and can induce formation of cellular protrusions in NIH3T3 cells (You and Lin-Chao, 2010; She *et al.*, 2002). Nostrin was identified as an interaction partner for endothelial nitric oxide synthase (eNOS) and facilitates its internalization by coordinating dynamin and N-WASP activity (Icking *et al.*, 2005). Recently, it was demonstrated that Nostrin is required for Rac1 activation and it mediates endothelial cell migration to promote angiogenesis in zebrafish (Kovacevic *et al.*, 2012).

In addition to the F-BAR proteins that bind WASP/WAVE family members, BAR, N-BAR, PX-BAR, and I-BAR proteins also bind and regulate actin polymerization at specific cellular locations. For example, proteins such as amphiphysin, endophilin, and SNX9 are involved in clathrin-mediated endocytosis (CME) and regulate actin polymerization at the neck of CCPs, allowing the branched actin network to assist dynamin during vesicle scission (Qualmann *et al.*, 2011). The role of these proteins in CME will be discussed below. Members of the I-BAR protein family, such as IRSp53 (insulin receptor substrate protein of 53 kDa), also bind to WASP/WAVE NPFs and are primarily implicated in inducing negative curvature leading to formation of lamellipodial and filopodial membrane protrusions (Suetsugu *et al.*, 2006). IRSp53 can bind to both N-WASP and WAVE NPFs via its SH3 domain (Takenawa and Suetsugu, 2007; Miki *et al.*, 2000). Overexpression of full-length IRSp53 induces filopodia formation by binding and recruiting N-WASP. However, this seems to be independent of the VCA domain of N-WASP, as N-WASP deleted of its VCA domain is still able to mediate IRSp53-dependent induction of filopodia (Lim *et al.*, 2008). The SH3 domain of IRSp53 can also bind to the formin, mDia1, and to the actin elongation factors, mammalian enabled protein (MENA) and VASP, which promote actin polymerization without branching, and actin bundling required for filopodia formation (Scita *et al.*, 2008). IRSp53 also has a cell division control protein 42 homolog (CDC42)-Rac interactive binding

region (CRIB) that is required for Cdc42-induced filopodia formation (Scita *et al.*, 2008).

Other I-BAR proteins such as MIM (missing-in-metastasis) and ABBA (actin-bundling protein with BAIAP2 homology) lack an SH3 domain, but instead have a C-terminal WH2 (WASP-homology 2) domain that mediates direct binding to actin monomers (Mattila *et al.*, 2003). MIM was first identified as a tumor suppressor that is down-regulated in bladder carcinoma cell lines (Lee *et al.*, 2002). Overexpression of MIM is associated with changes in cell shape, driving lamellipodia formation, an increase in membrane ruffles and promoting metastatic ability of cancer cells by engaging Rho GTPases and cofilin in this process (Xie *et al.*, 2011). Another BAR domain protein involved in formation of membrane protrusions such as invadopodia and podosomes is ASAP1 (ArfGAP with SH3 domain, ankyrin repeat and PH domain), a tandem BAR-PH domain-containing protein with an ArfGTPase activating domain (ArfGAP) that catalyzes hydrolysis of GTP-bound Arf proteins (Kam *et al.*, 2000). ASAP1 interacts with Src kinase, which phosphorylates a crucial tyrosine residue required for ASAP1 function in podosome formation (Brown *et al.*, 1998). PinkBAR, an epithelial tissue-specific I-BAR protein has a unique flat BAR domain and thus cannot induce curvature but rather deforms phosphoinositide-containing membranes into planar structures (Pykäläinen *et al.*, 2011). PinkBAR localizes to specific cellular locations such as the tight junctions in kidney and intestinal epithelial cells and is implicated in maintaining the integrity of these intercellular connections (Pykäläinen *et al.*, 2011).

BAR Domain Proteins as Regulators of Intracellular Transport

One of the most well-studied functions of BAR domain proteins is during clathrin-mediated endocytosis. During CME, the plasma membrane of the cell undergoes sequential stages of remodeling with increasing membrane curvature. It has been suggested that BAR domain proteins can induce/sense this membrane curvature, and create macromolecular platforms that recruit different binding partners. The recruitment of actin, dynamin, and synaptojanin thus coordinate the different stages of CME such as membrane invagination, clathrin coat formation, actin nucleation, fission, and uncoating (Suetsugu and Gautreau, 2012). It is interesting to speculate that the spatiotemporal recruitment of different BAR proteins to these increasingly curved membranes is predetermined by the intrinsic curvature of their BAR domains. Recent studies have revealed that for some, but not all BAR domain proteins, this correlation holds true (Taylor *et al.*, 2011). For instance, members of F-BAR family, Fer/Cip4 homology domain-only proteins 1 and 2 (FCHo1/2), which have a very shallow curvature, are recruited early during CME and are thought to induce/stabilize the shallow invagination of the plasma membrane marking the sites for CCP formation (Henne *et al.*, 2007). In agreement with this, overexpression of FCHo protein increases nucleation of CCPs while knockdown of FCHo leads to reduced formation of CCPs (Henne *et al.*, 2010). However, compelling studies in zebrafish demonstrate that FCHo activity is not required for clathrin-coat nucleation and that adaptor protein-2 (AP-2) depletion leads to a broader and earlier

developmental defect that is independent of FCHo1/2 (Umasankar *et al.*, 2012).

N-BAR family proteins, amphiphysin and endophilin as well as the F-BAR protein Syndapin are recruited to late stages of clathrin-coated vesicle formation prior to scission, as they have a very narrow curvature which fits with that of the deeply invaginated coated pits (Taylor *et al.*, 2011). Amongst these three BAR-domain proteins, endophilin has the sharpest curvature followed by amphiphysin. Syndapins have a shallower curvature of ~ 50 nm circle diameter (Peter *et al.*, 2004). However, their near simultaneous recruitment to coated invaginations is surprising and suggests that other factors besides intrinsic BAR domain curvature guide their temporal association with CCPs (Taylor *et al.*, 2011). Similarly, the biphasic recruitment of the F-BAR subfamily CIP4/FBP17/Toca-1 before and after vesicle scission (Taylor *et al.*, 2011) does not correlate with their BAR domain structure alone, suggesting roles for other factors regulating recruitment of BAR domain proteins during CME. It is interesting to note that recruitment of these BAR domain proteins correlates with the recruitment of their common binding partners, the small GTPase dynamin and N-WASP (Qualmann *et al.*, 1999, 2000; Ho *et al.*, 2004; Itoh *et al.*, 2005; David *et al.*, 1996).

The prototype N-BAR proteins, endophilin and amphiphysin, regulate receptor-mediated endocytosis and synaptic vesicle recycling (Gad *et al.*, 2000; Verstreken *et al.*, 2003; Schuske *et al.*, 2003). Endophilin interacts with dynamin, N-WASP, and synaptojanin via its SH3 domain while amphiphysin can bind to clathrin, the AP-2 adaptor complex, dynamin, and synaptojanin (Verstreken *et al.*, 2003; McMahon *et al.*, 1997). Initial loss-of-function studies of endophilin A in worms and flies revealed the presence of CCPs devoid of dynamin rings, suggesting a role for endophilin A in the recruitment of dynamin to later stages CCP life spans (Gad *et al.*, 2000; Schuske *et al.*, 2003). However, results from knockout mice of all three endophilin A isoforms demonstrate that endophilins are involved in synaptic vesicle recycling and uncoating of clathrin-coated vesicles rather than dynamin-dependent scission of clathrin-coated vesicles. Uncoating of clathrin-coated vesicles (CCVs) is regulated by dephosphorylation of phosphatidylinositol-4,5-bisphosphate to phosphatidylinositol-4-monophosphate by the phosphatase synaptojanin, whose recruitment to CCPs and control of enzymatic activity is regulated by its interaction with the endophilin A SH3 domain (Milosevic *et al.*, 2011). It is now generally accepted that the main role of endophilin A during clathrin-mediated endocytosis is to regulate synaptojanin, but not dynamin recruitment to CCPs.

Amphiphysin I, the first structurally characterized BAR domain protein, has an N-terminal BAR domain, a central alanine- and glutamate-rich domain harboring a proline-rich motif and hydrophobic region, and a C-terminal SH3 domain (Peter *et al.*, 2004). Initial studies had identified dynamin and synaptojanin as two major binding partners of amphiphysin's SH3 domain from brain extracts (David *et al.*, 1996). Interestingly, amphiphysin I SH3 domain is also able to mediate an intramolecular interaction with its own central proline-rich motif, which is important for regulating its interaction with other binding partners such as dynamin and synaptojanin (Farsad *et al.*, 2003). Amphiphysin I also interacts with the

α subunit of AP-2 and clathrin heavy chain via an interaction interface known as the CLAP (clathrin/AP-2 binding) domain (McMahon *et al.*, 1997). Initial studies on the SH3 domain of amphiphysin revealed its essential role in recruiting dynamin to endocytic pits, both in neuronal and non-neuronal cells (Shupliakov *et al.*, 1997). Expression of the amphiphysin I SH3 domain (but not SH3 domains from other proteins that bind dynamin via other motifs distinct from the amphiphysin 1 SH3 domain) inhibits receptor-mediated endocytosis. This effect was specifically due to inhibition of dynamin, as its overexpression neutralized the effect of the amphiphysin1 SH3 domain and restored endocytosis (Wigge *et al.*, 1997). Similarly, microinjection of the GST-CLAP domain of amphiphysin I blocked the early stages of clathrin-mediated endocytosis, suggesting that the association of amphiphysin with both AP-2 and clathrin is essential for endocytosis (McMahon *et al.*, 1997). Interestingly, analysis of brain extracts from amphiphysin I knockout mice revealed reduced membrane associations between AP-2, clathrin, and synaptojanin, but surprisingly not dynamin (Di Paolo *et al.*, 2002). This contradicted earlier work suggesting that amphiphysin is critical for dynamin recruitment to CCPs and inducing dynamin's GTPase activity (Shupliakov *et al.*, 1997). However, further analysis revealed reduced association of the dynamin1 PRD with the clathrin heavy chain, and AP-2 in the amphiphysin I knockout mice, but these interactions were restored upon expression of recombinant amphiphysin I (Di Paolo *et al.*, 2002). Thus while amphiphysin was not required for dynamin's association with membranes, it was important to link dynamin to clathrin and AP-2 on CCPs.

In non-neuronal cells, the PX-BAR protein SNX9 is critical for dynamin recruitment to CCP (Lundmark and Carlsson, 2004). SNX9 is a member of sorting nexin family of proteins, which has 33 members in mammals, several of which are involved in membrane trafficking (Worby and Dixon, 2002). While most SNX family members contain BAR and PX domain, the SNX9 subfamily (which includes SNX18 and SNX33) contains an additional SH3 domain that binds to dynamin and N-WASP and is implicated in clathrin-mediated endocytosis (Worby and Dixon, 2002). SNX9 knockdown in HeLa cells leads to the reduced internalization of transferrin receptor. Moreover, in non-neuronal cells, SNX9 binds and recruits dynamin to the plasma membrane and stimulates its GTPase activity, thus suggesting a role for SNX9 during CME vesicle scission (Lundmark and Carlsson, 2004). SNX9 also participates in N-WASP activation leading to Arp2/3 complex-mediated actin nucleation in the presence of liposomes (Yarar *et al.*, 2007). Other SNX family members are involved in the later stages of endocytic transport as discussed below.

Similar to the BAR and N-BAR proteins mentioned above, syndapin I and closely related F-BAR family members CIP4, FBP17 and Toca-1 also interact with dynamin and N-WASP to regulate CME (Itoh *et al.*, 2005). Syndapin I is predominantly expressed in neurons and is thought to play a critical role in recruiting dynamin to the plasma membrane, thereby regulating its function during synaptic vesicle recycling (Qualmann *et al.*, 1999). Both proteins are simultaneously recruited to CCPs, as revealed by TIRF analyses (Taylor *et al.*, 2011). Microinjection of antibodies either against full-length syndapin, or its SH3 domain, blocks scission of synaptic vesicles, an

effect similar to blocking dynamin function (Andersson *et al.*, 2008). Syndapin I knockout mice phenocopy several defects found in the dynamin I knockout mice, further reinforcing the notion that syndapin I regulates dynamin function (Ferguson *et al.*, 2007). The F-BAR subfamily CIP4/FBP17/Toca-1 proteins regulate endocytosis of transferrin receptor, as demonstrated by single, double, or triple knockdown of the proteins in cultured cells (Itoh *et al.*, 2005). Targeted deletion or knockdown of these proteins in flies and worms has revealed their role in endocytosis of E-cadherin and the yolk receptor respectively (Fricke *et al.*, 2009; Giuliani *et al.*, 2009; Leibfried *et al.*, 2008). Interaction of CIP4/FBP17/Toca-1 and syndapin I with N-WASP is thought to link actin cytoskeleton dynamics with clathrin-coated vesicle detachment and driving the vesicle away from the plasma membrane after its scission (Suetsugu and Gautreau, 2012; Itoh *et al.*, 2005; Qualmann *et al.*, 2011).

There are several other multiple domain-containing BAR proteins that apparently function in endocytic transport. One example is Tuba, which has four N-terminal SH3 domain that bind to dynamin, a central Dbl homology (DH) domain that functions as a GTPase exchange factor for Cdc42, a BAR domain and two C-terminal SH3 domains that bind to N-WASP and Ena/VASP, proteins that regulate actin polymerization (Salazar *et al.*, 2003). The DH domain of Tuba is unique amongst the proteins containing DH domains as it is followed by a BAR domain and not a PH domain, suggesting that Tuba might provide a functional link between dynamin, Rho GTPase signaling, the actin cytoskeleton, and endocytosis. Tuba also localizes to and stimulates dorsal ruffle formation in a melanoma cell line and controls junctional organization of simple epithelial cells (Otani *et al.*, 2006; Kovacs *et al.*, 2006). Another example is the PDZ and BAR domain protein PICK1 (protein interacting with C-kinase 1), which is highly expressed in the central nervous system and was identified as a protein interacting with PKC α (Staudinger *et al.*, 1995). PICK1 also interacts with several critical neuronal receptors, including α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)-type ionotropic glutamate receptors (AMPA receptors), the metabotropic glutamate receptor mGluR7 as well as various neurotransmitter transporters including the dopamine transporter (DAT), the norepinephrine transporter, and the Glt1b glutamate transporter (Höhne *et al.*, 2011; Jin *et al.*, 2006). For several of these receptors, it has been suggested that PICK1 regulates their endocytosis by direct binding, or by regulating their PKC α -dependent phosphorylation, which in turn controls their neuronal functions important for synaptic plasticity (Hanley, 2008).

While the BAR domain proteins discussed thus far regulate initial stages of endocytic transport at the plasma membrane, and their BAR domains have affinity toward the phospholipid phosphatidylinositol-4,5-bisphosphate, BAR domain proteins with either a PH or PX functional module (such as APPL1 and sorting nexins) display affinity toward phospholipid phosphatidylinositol-3-monophosphate, a lipid enriched on early endosomes. Early endosome localization of APPL1 and its isoform APPL2 is regulated by their interaction with the endocytic Rab GTPase, Rab5, which regulates vesicle fusion at the early endosome (Miaczynska *et al.*, 2004). APPL1 and APPL2 are unique endocytic adaptor that shuttle between endosomes and nucleus in response to extracellular stimuli

such as EGF. In the nucleus APPL proteins interact with the nucleosome remodeling and histone deacetylase complex (NuRD/MeCP1) and thus regulate cell proliferation (Miaczynska *et al.*, 2004). Sorting nexins were originally identified as proteins interacting with EGFR (Kurten *et al.*, 1996). Of the 33 members, only nine have the BAR domain described as SNX-BAR proteins (Worby and Dixon, 2002). SNX1, the most well-characterized member of the SNX-BAR subfamily is part of large multiprotein complex known as the retromer and it is implicated in cargo retrieval along the endosome-to-Golgi pathway, a function that is conserved from yeast to mammals (Worby and Dixon, 2002). Endogenous SNX1 is greatly enriched on highly tubular membranes that emanate from early endosomes (Griffin *et al.*, 2005). Both the PX domain and the curvature-sensing BAR domain are involved in correctly targeting SNX1 to these intracellular membranes and regulating its function in retrograde trafficking (Worby and Dixon, 2002; Griffin *et al.*, 2005). There is growing evidence suggesting that other members of the SNX family (i.e., SNX4 and SNX9) regulate membrane trafficking from the plasma membrane to endosomes by recognizing lipids enriched at the plasma membrane (Worby and Dixon, 2002).

Finally, several recent studies of ubiquitously expressed isoforms of the brain-specific BAR domain proteins amphiphysin I and syndapin I, amphiphysin II, and syndapin II, have revealed their roles in the transport of cargo from the endocytic recycling compartment to the plasma membrane. These proteins associate with C-terminal Eps15 homology domain (EHD) family of proteins to regulate receptor recycling (Giridharan *et al.*, 2013; Pant *et al.*, 2009; Braun *et al.*, 2005).

Role of BAR Proteins in Regulating Signaling

While BAR domain-containing proteins are generally associated with regulating cytoskeletal dynamics and endocytic trafficking, a subgroup of BAR-domain proteins such as Fps/Fes (fps, Fujinami poultry sarcoma; fes, feline sarcoma) and Fer (Fes-related protein) play a crucial role in regulating signaling pathways. Fps/Fes and Fer are structurally unique members of the non-receptor protein tyrosine kinase family and were first isolated as oncogenes from avian and feline retroviruses (Beemon, 1981). Subsequently, the cellular homolog of the Fps/Fes oncogene was identified in several species, including humans and mice, where they play key roles in inflammatory pathways, hemostasis, and innate immunity (Greer, 2002). Fes and Fer proteins contain an N-terminal F-BAR domain, an F-BAR extension (FX) domain that binds to phosphatidic acid, a SH2 domain that recognizes phosphotyrosine motifs, and a C-terminal kinase domain. Early studies on the viral Fps pointed to an important role of its N-terminal region in regulating transforming activity and membrane localization (Brooks-Wilson *et al.*, 1989). Moreover, the F-BAR domains of Fes and Fer are necessary for their oligomerization, which potentiates their *trans*-autophosphorylation activity. However, whether oligomerization is required for their catalytic activity is not known (Greer, 2002). The Fes F-BAR domain can bind phosphoinositides and induce liposome tubulation *in vitro* (Zirngibl *et al.*, 2001). The F-BAR domain in Fer interacts with the adherens-junction protein p120catenin,

and overexpression of Fer results in increased phosphorylation of p120catenin and dissolution of adherens junctions and cell–cell adhesion (Piedra *et al.*, 2003). The SH2 domain of Fps/Fes and Fer is thought to mediate phosphotyrosine-dependent interactions with several receptors such as EGFR and PDGFR, as well as signaling molecules such as PI3K and insulin receptor substrate-1 (Iwanishi *et al.*, 2000). Moreover structural insights into the SH2 domain from Fes have suggested that the SH2 domain tightly interacts with the kinase domain, and activation of Fes kinase activity is closely coupled to SH2–kinase–ligand interactions (Filippakopoulos *et al.*, 2008). The kinase domain of Fps/Fes phosphorylates, among other substrates, the protein BCR (breakpoint cluster region) (Maru *et al.*, 1995), which has GEF activity toward Rho GTPase and RhoA and GAP activity toward Rac1, suggesting that Fps/Fes can regulate cytoskeletal dynamics by controlling BCR phosphorylation (Chuang *et al.*, 1995). Other functions of Fes and Fer include vesicular trafficking and receptor internalization-mediated phosphorylation of protein substrates such as cortactin, platelet/endothelial cell adhesion molecule1 (Pecam-1), and β -catenin, as well as regulating cell migration, platelet activation, and leukocyte extravasation at sites of inflammation (Greer, 2002).

See also: Molecular Principles, Components, Technology, and Concepts: Membranes: Composition, Physical Properties, and Curvature

References

- Ahmed, S., Goh, W.I., Bu, W., 2010. I-BAR domains, IRSp53 and filopodium formation. *Seminars in Cell and Developmental Biology* 21 (4), 350–356.
- Andersson, F., Jakobsson, J., Löw, P., Shupliakov, O., Brodin, L., 2008. Perturbation of syndapin/PACSIN impairs synaptic vesicle recycling evoked by intense stimulation. *Journal of Neuroscience* 28 (15), 3925–3933.
- Aspenstrom, P., 1997. A Cdc42 target protein with homology to the non-kinase domain of FER has a potential role in regulating the actin cytoskeleton. *Current Biology* 7 (7), 479–487.
- Badour, K., Zhang, J., Shi, F., *et al.*, 2004. Fyn and PTP-PEST-mediated regulation of Wiskott-Aldrich syndrome protein (WASP) tyrosine phosphorylation is required for coupling T cell antigen receptor engagement to WASP effector function and T cell activation. *Journal of Experimental Medicine* 199 (1), 99–112.
- Beemon, K., 1981. Transforming proteins of some feline and avian sarcoma viruses are related structurally and functionally. *Cell* 24 (1), 145–153.
- Braun, A., Pinyol, R., Dahlhaus, R., *et al.*, 2005. EHD proteins associate with syndapin I and II and such interactions play a crucial role in endosomal recycling. *Molecular Biology of the Cell* 16 (8), 3642–3658.
- Brooks-Wilson, A.R., Ball, E., Pawson, T., 1989. The myristylation signal of p60v-src functionally complements the N-terminal fps-specific region of P130gag-fps. *Molecular Biology of the Cell* 9 (5), 2214–2219.
- Brown, M.T., Andrade, J., Radhakrishna, H., *et al.*, 1998. ASAP1, a phospholipid-dependent arf GTPase-activating protein that associates with and is phosphorylated by Src. *Molecular and Cellular Biology* 18 (12), 7038–7051.
- Campellone, K.G., Siripala, A.D., Leong, J.M., Welch, M.D., 2012. Membrane-deforming proteins play distinct roles in actin pedestal biogenesis by enterohemorrhagic *Escherichia coli*. *Journal of Biological Chemistry* 287 (24), 20613–20624.
- Chitu, V., Pixley, F.J., Macaluso, F., *et al.*, 2005. The PCH family member MAYP/PSTPIP2 directly regulates F-actin bundling and enhances filopodia formation and motility in macrophages. *Molecular Biology of the Cell* 16 (6), 2947–2959.
- Chuang, T.H., Xu, X., Kaartinen, V., *et al.*, 1995. Abr and Bcr are multifunctional regulators of the Rho GTP-binding protein family. *Proceedings of the National Academy of Sciences of the United States of America* 92 (22), 10282–10286.
- Côté, J.F., Chung, P.L., Théberge, J.F., *et al.*, 2002. PSTPIP is a substrate of PTP-PEST and serves as a scaffold guiding PTP-PEST toward a specific dephosphorylation of WASP. *Journal of Biological Chemistry* 277 (4), 2973–2986.
- David, C., McPherson, P.S., Mundigl, O., de Camilli, P., 1996. A role of amphiphysin in synaptic vesicle endocytosis suggested by its binding to dynamin in nerve terminals. *Proceedings of the National Academy of Sciences of the United States of America* 93 (1), 331–335.
- Devriendt, K., Kim, A.S., Mathijs, G., *et al.*, 2001. Constitutively activating mutation in WASP causes X-linked severe congenital neutropenia. *Nature Genetics* 27 (3), 313–317.
- Di Paolo, G., Sankaranarayanan, S., Wenk, M.R., *et al.*, 2002. Decreased synaptic vesicle recycling efficiency and cognitive deficits in amphiphysin 1 knockout mice. *Neuron* 33 (5), 789–804.
- Dislich, B., Than, M.E., Lichtenthaler, S.F., 2011. Specific amino acids in the BAR domain allow homodimerization and prevent heterodimerization of sorting nexin 33. *Biochemical Journal* 433 (1), 75–83.
- Endris, V., Wogatzky, B., Leimer, U., *et al.*, 2002. The novel Rho-GTPase activating gene MEGAP/ srGAP3 has a putative role in severe mental retardation. *Proceedings of the National Academy of Sciences of the United States of America* 99 (18), 11754–11759.
- Farsad, K., Slepnev, V., Ochoa, G., *et al.*, 2003. A putative role for intramolecular regulatory mechanisms in the adaptor function of amphiphysin in endocytosis. *Neuropharmacology* 45 (6), 787–796.
- Ferguson, S.M., Brasnjo, G., Hayashi, M., *et al.*, 2007. A selective activity-dependent requirement for dynamin 1 in synaptic vesicle endocytosis. *Science* 316 (5824), 570–574.
- Filippakopoulos, P., Kofler, M., Hantschel, O., *et al.*, 2008. Structural coupling of SH2-kinase domains links Fes and Abl substrate recognition and kinase activation. *Cell* 134 (5), 793–803.
- Fricke, R., Gohl, C., Dharmalingam, E., *et al.*, 2009. Drosophila Cip4/Toca-1 integrates membrane trafficking and actin dynamics through WASP and SCAR/WAVE. *Current Biology* 19 (17), 1429–1437.
- Frost, A., Unger, V.M., De Camilli, P., 2009. The BAR domain superfamily: Membrane-molding macromolecules. *Cell* 137 (2), 191–196.
- de la Fuente, M.A., Sasahara, Y., Calamito, M., *et al.*, 2007. WIP is a chaperone for Wiskott-Aldrich syndrome protein (WASP). *Proceedings of the National Academy of Sciences of the United States of America* 104 (3), 926–931.
- Gad, H., Ringstad, N., Löw, P., *et al.*, 2000. Fission and uncoating of synaptic clathrin-coated vesicles are perturbed by disruption of interactions with the SH3 domain of endophilin. *Neuron* 27 (2), 301–312.
- Giridharan, S.S., Cai, B., Vitale, N., Naslavsky, N., Caplan, S., 2013. Cooperation of MICAL-L1, syndapin2, and phosphatidic acid in tubular recycling endosome biogenesis. *Molecular Biology of the Cell* 24 (11), 1776–1790. S1–15.
- Giuliani, C., Troglio, F., Bai, Z., *et al.*, 2009. Requirements for F-BAR proteins TOCA-1 and TOCA-2 in actin dynamics and membrane trafficking during *Caenorhabditis elegans* oocyte growth and embryonic epidermal morphogenesis. *PLoS Genetics* 5 (10), e1000675.
- Greer, P., 2002. Closing in on the biological functions of Fps/Fes and Fer. *Nature Reviews Molecular Cell Biology* 3 (4), 278–289.
- Griffin, C.T., Trejo, J., Magnuson, T., 2005. Genetic evidence for a mammalian retromer complex containing sorting nexins 1 and 2. *Proceedings of the National Academy of Sciences of the United States of America* 102 (42), 15173–15177.
- Guerrier, S., Coutinho-Budd, J., Sassa, T., *et al.*, 2009. The F-BAR domain of srGAP2 induces membrane protrusions required for neuronal migration and morphogenesis. *Cell* 138 (5), 990–1004.
- Gyuris, J., Golemis, E., Chertkov, H., Brent, R., 1993. Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* 75 (4), 791–803.
- Hanley, J.G., 2008. PICK1: A multi-talented modulator of AMPA receptor trafficking. *Pharmacology & Therapeutics* 118 (1), 152–160.
- Henne, W.M., Boucrot, E., Meinecke, M., *et al.*, 2010. FCHO proteins are nucleators of clathrin-mediated endocytosis. *Science* 328 (5983), 1281–1284.
- Henne, W.M., Kent, H.M., Ford, M.G., *et al.*, 2007. Structure and analysis of FCHO2 F-BAR domain: A dimerizing and membrane recruitment module that effects membrane curvature. *Structure* 15 (7), 839–852.
- Ho, H.Y., Rohatgi, R., Lebensohn, A.M., *et al.*, 2004. Toca-1 mediates Cdc42-dependent actin nucleation by activating the N-WASP-WIP complex. *Cell* 118 (2), 203–216.
- Höhne, M., Lorscheider, J., von Bardeleben, A., *et al.*, 2011. The BAR domain protein PICK1 regulates cell recognition and morphogenesis by interacting with Neph proteins. *Molecular Biology of the Cell* 31 (16), 3241–3251.
- Icking, A., Matt, S., Opitz, N., *et al.*, 2005. NOSTRIN functions as a homotrimeric adaptor protein facilitating internalization of eNOS. *Journal of Cell Science* 118 (Pt 21), 5059–5069.

- Itoh, T., Erdmann, K.S., Roux, A., *et al.*, 2005. Dynamin and the actin cytoskeleton cooperatively regulate plasma membrane invagination by BAR and F-BAR proteins. *Developmental Cell* 9 (6), 791–804.
- Iwanishi, M., Czech, M.P., Cherniack, A.D., 2000. The protein-tyrosine kinase *fer* associates with signaling complexes containing insulin receptor substrate-1 and phosphatidylinositol 3-kinase. *Journal of Biological Chemistry* 275 (50), 38995–39000.
- Jin, W., Ge, W.-P., Xu, J., *et al.*, 2006. Lipid binding regulates synaptic targeting of PICK1, AMPA receptor trafficking, and synaptic plasticity. *Journal of Neuroscience* 26 (9), 2380–2390.
- Kam, J.L., Miura, K., Jackson, T.R., *et al.*, 2000. Phosphoinositide-dependent activation of the ADP-ribosylation factor GTPase-activating protein ASAP1. Evidence for the pleckstrin homology domain functioning as an allosteric site. *Journal of Biological Chemistry* 275 (13), 9653–9663.
- Kessels, M.M., Dong, J., Leibig, W., Westermann, P., Qualmann, B., 2006. Complexes of syndapin II with dynamin II promote vesicle formation at the trans-Golgi network. *Journal of Cell Science* 119 (Pt 8), 1504–1516.
- Kovacevic, I., Hu, J., Siehoff-Icking, A., *et al.*, 2012. The F-BAR protein NOSTRIN participates in FGF signal transduction and vascular development. *EMBO Journal* 31 (15), 3309–3322.
- Kovacs, E.M., Makar, R.S., Gertler, F.B., 2006. Tuba stimulates intracellular N-WASP-dependent actin assembly. *Journal of Cell Science* 119 (Pt 13), 2715–2726.
- Kurten, R.C., Cadena, D.L., Gill, G.N., 1996. Enhanced degradation of EGF receptors by a sorting nexin, SNX1. *Science* 272 (5264), 1008–1010.
- Lee, Y.G., Macoska, J.A., Korenchuk, S., Pienta, K.J., 2002. MIM, a potential metastasis suppressor gene in bladder cancer. *Neoplasia* 4 (4), 291–294.
- Leibfried, A., Fricke, R., Morgan, M.J., Bogdan, S., Bellaiche, Y., 2008. Drosophila Cip4 and WASp define a branch of the Cdc42-Par6-aPKC pathway regulating E-cadherin endocytosis. *Current Biology* 18 (21), 1639–1648.
- Lim, K.B., Bu, W., Goh, W.I., *et al.*, 2008. The Cdc42 effector IRSp53 generates filopodia by coupling membrane protrusion with actin dynamics. *Journal of Biological Chemistry* 283 (29), 20454–20472.
- Lundmark, R., Carlsson, S.R., 2004. Regulated membrane recruitment of dynamin-2 mediated by sorting nexin 9. *Journal of Biological Chemistry* 279 (41), 42694–42702.
- Maru, Y., Peters, K.L., Afar, D.E., *et al.*, 1995. Tyrosine phosphorylation of BCR by FPS/FES protein-tyrosine kinases induces association of BCR with GRB-2/SOS. *Molecular Biology of the Cell* 15 (2), 835–842.
- Mattila, P.K., Salminen, M., Yamashiro, T., Lappalainen, P., 2003. Mouse MIM, a tissue-specific regulator of cytoskeletal dynamics, interacts with ATP-actin monomers through its C-terminal WH2 domain. *Journal of Biological Chemistry* 278 (10), 8452–8459.
- McMahon, H.T., Wigge, P., Smith, C., 1997. Clathrin interacts specifically with amphiphysin and is displaced by dynamin. *FEBS Letters* 413 (2), 319–322.
- Miaczynska, M., Christoforidis, S., Giner, A., *et al.*, 2004. APPL proteins link Rab5 to nuclear signal transduction via an endosomal compartment. *Cell* 116 (3), 445–456.
- Miki, H., Yamaguchi, H., Suetsugu, S., Takenawa, T., 2000. IRSp53 is an essential intermediate between Rac and WAVE in the regulation of membrane ruffling. *Nature* 408 (6813), 732–735.
- Milosevic, I., Dorion, H., Ricchiuti, D., 2011. Incarcerated retroperitoneal hernia following total extraperitoneal laparoscopic radical nephrectomy. *Journal of the Society of Laparoendoscopic Surgeons* 15 (3), 424–426.
- Milosevic, I., Giovedi, S., Lou, X., *et al.*, 2011. Recruitment of endophilin to clathrin-coated pit necks is required for efficient vesicle uncoating after fission. *Neuron* 72 (4), 587–601.
- Mim, C., Unger, V.M., 2012. Membrane curvature and its generation by BAR proteins. *Trends in Biochemical Sciences* 37 (12), 526–533.
- Modregger, J., Ritter, B., Witter, B., Paulsson, M., Plomann, M., 2000. All three PACSIN isoforms bind to endocytic proteins and inhibit endocytosis. *Journal of Cell Science* 113 (Pt 24), 4511–4521.
- Otani, T., Ichii, T., Aono, S., Takeichi, M., 2006. Cdc42 GEF Tuba regulates the junctional configuration of simple epithelial cells. *Journal of Cell Biology* 175 (1), 135–146.
- Padrick, S.B., Rosen, M.K., 2010. Physical mechanisms of signal integration by WASP family proteins. *Annual Review of Biochemistry* 79, 707–735.
- Pant, S., Sharma, M., Patel, K., *et al.*, 2009. AMPH-1/Amphiphysin/Bin1 functions with RME-1/Ehd1 in endocytic recycling. *Nature Cell Biology* 11 (12), 1399–1410.
- Peter, B.J., Kent, H.M., Mills, I.G., *et al.*, 2004. BAR domains as sensors of membrane curvature: The amphiphysin BAR structure. *Science* 303 (5657), 495–499.
- Pichot, C.S., Arvanitis, C., Hartig, S.M., *et al.*, 2010. Cdc42-interacting protein 4 promotes breast cancer cell invasion and formation of invadopodia through activation of N-WASP. *Cancer Research* 70 (21), 8347–8356.
- Piedra, J., Miravet, S., Castaño, J., *et al.*, 2003. p120 Catenin-associated *Fer* and *Fyn* tyrosine kinases regulate beta-catenin Tyr-142 phosphorylation and beta-catenin-alpha-catenin interaction. *Molecular Biology of the Cell* 23 (7), 2287–2297.
- Pollard, T.D., 2007. Regulation of actin filament assembly by Arp2/3 complex and formins. *Annual Review of Biophysics and Biomolecular Structure* 36, 451–477.
- Pykäläinen, A., Boczkowska, M., Zhao, H., *et al.*, 2011. Pinkbar is an epithelial-specific BAR domain protein that generates planar membrane structures. *Nature Structural & Molecular Biology* 18 (8), 902–907.
- Qualmann, B., Kelly, R.B., 2000. Syndapin isoforms participate in receptor-mediated endocytosis and actin organization. *Journal of Cell Biology* 148 (5), 1047–1062.
- Qualmann, B., Kessels, M.M., Kelly, R.B., 2000. Molecular links between endocytosis and the actin cytoskeleton. *Journal of Cell Biology* 150 (5), F111–F116.
- Qualmann, B., Koch, D., Kessels, M.M., 2011. Let's go bananas: Revisiting the endocytic BAR code. *EMBO Journal* 30 (17), 3501–3515.
- Qualmann, B., Roos, J., DiGregorio, P.J., Kelly, R.B., 1999. Syndapin I, a synaptic dynamin-binding protein that associates with the neural Wiskott-Aldrich syndrome protein. *Molecular Biology of the Cell* 10 (2), 501–513.
- Rohatgi, R., Ma, L., Miki, H., *et al.*, 1999. The interaction between N-WASP and the Arp2/3 complex links Cdc42-dependent signals to actin assembly. *Cell* 97 (2), 221–231.
- Romer, W., Pontani, L.L., Sorre, B., *et al.*, 2010. Actin dynamics drive membrane reorganization and scission in clathrin-independent endocytosis. *Cell* 140 (4), 540–553.
- Safari, F., Suetsugu, S., 2012. The BAR domain superfamily proteins from subcellular structures to human diseases. *Membranes* 2 (1), 91–117.
- Salazar, M.A., Kwiatkowski, A.V., Pellegrini, L., *et al.*, 2003. Tuba, a novel protein containing bin/amphiphysin/Rvs and Dbl homology domains, links dynamin to regulation of the actin cytoskeleton. *Journal of Biological Chemistry* 278 (49), 49031–49043.
- Schuske, K.R., Richmond, J.E., Matthies, D.S., *et al.*, 2003. Endophilin is required for synaptic vesicle endocytosis by localizing synaptotagmin. *Neuron* 40 (4), 749–762.
- Scita, G., Confalonieri, S., Lappalainen, P., Suetsugu, S., 2008. IRSp53: Crossing the road of membrane and actin dynamics in the formation of membrane protrusions. *Trends in Cell Biology* 18 (2), 52–60.
- Senju, Y., Itoh, Y., Takano, K., Hamada, S., Suetsugu, S., 2011. Essential role of PACSIN2/syndapin-II in caveolae membrane sculpting. *Journal of Cell Science* 124 (Pt 12), 2032–2040.
- She, B.R., Liou, G.G., Lin-Chao, S., 2002. Association of the growth-arrest-specific protein Gas7 with F-actin induces reorganization of microfilaments and promotes membrane outgrowth. *Experimental Cell Research* 273 (1), 34–44.
- Shupliakov, O., Löw, P., Grabs, D., *et al.*, 1997. Synaptic vesicle endocytosis impaired by disruption of dynamin-SH3 domain interactions. *Science* 276 (5310), 259–263.
- Staudinger, J., Zhou, J., Burgess, R., Elledge, S.J., Olson, E.N., 1995. PICK1: A perinuclear binding protein and substrate for protein kinase C isolated by the yeast two-hybrid system. *Journal of Cell Biology* 128 (3), 263–271.
- Suetsugu, S., Gautreau, A., 2012. Synergistic BAR-NPF interactions in actin-driven membrane remodeling. *Trends in Cell Biology* 22 (3), 141–150.
- Suetsugu, S., Murayama, K., Sakamoto, A., *et al.*, 2006. The RAC binding domain/IRSp53-MIM homology domain of IRSp53 induces RAC-dependent membrane deformation. *Journal of Biological Chemistry* 281 (46), 35347–35358.
- Takano, K., Toyooka, K., Suetsugu, S., 2008. EFC/F-BAR proteins and the N-WASP-WIP complex induce membrane curvature-dependent actin polymerization. *EMBO Journal* 27 (21), 2817–2828.
- Takenawa, T., Suetsugu, S., 2007. The WASP-WAVE protein network: Connecting the membrane to the cytoskeleton. *Nature Reviews Molecular Cell Biology* 8 (1), 37–48.
- Tarricone, C., Xiao, B., Justin, N., *et al.*, 2001. The structural basis of Arfaptin-mediated cross-talk between Rac and Arf signalling pathways. *Nature* 411 (6834), 215–219.
- Taylor, M.J., Perrais, D., Merrifield, C.J., 2011. A high precision survey of the molecular dynamics of mammalian clathrin-mediated endocytosis. *PLoS Biology* 9 (3), e1000604.
- Tsuboi, S., Takada, H., Hara, T., *et al.*, 2009. FBP17 mediates a common molecular step in the formation of podosomes and phagocytic cups in macrophages. *Journal of Biological Chemistry* 284 (13), 8548–8556.

- Umasankar, P.K., Sanker, S., Thieman, J.R., *et al.*, 2012. Distinct and separable activities of the endocytic clathrin-coat components Fcho1/2 and AP-2 in developmental patterning. *Nature Cell Biology* 14 (5), 488–501.
- Verstreken, P., Koh, T.W., Schulze, K.L., *et al.*, 2003. Synaptotagmin is recruited by endophilin to promote synaptic vesicle uncoating. *Neuron* 40 (4), 733–748.
- Wigge, P., Vallis, Y., McMahon, H.T., 1997. Inhibition of receptor-mediated endocytosis by the amphiphysin SH3 domain. *Current Biology* 7 (8), 554–560.
- Wise, C.A., Gillum, J.D., Seidman, C.E., *et al.*, 2002. Mutations in CD2BP1 disrupt binding to PTP PEST and are responsible for PAPA syndrome, an autoinflammatory disorder. *Human Molecular Genetics* 11 (8), 961–969.
- Wong, K., Ren, X.R., Huang, Y.Z., *et al.*, 2001. Signal transduction in neuronal migration: Roles of GTPase activating proteins and the small GTPase Cdc42 in the Slit-Robo pathway. *Cell* 107 (2), 209–221.
- Worby, C.A., Dixon, J.E., 2002. Sorting out the cellular functions of sorting nexins. *Nature Reviews Molecular Cell Biology* 3 (12), 919–931.
- Wu, J.Y., Feng, L., Park, H.-T., *et al.*, 2001. The neuronal repellent Slit inhibits leukocyte chemotaxis induced by chemotactic factors. *Nature* 410 (6831), 948–952.
- Xie, F., Ye, L., Ta, M., *et al.*, 2011. MTSS1: A multifunctional protein and its role in cancer invasion and metastasis. *Frontiers in bioscience (Scholar edition)* 3, 621–631.
- Yarar, D., Waterman-Storer, C.M., Schmid, S.L., 2007. SNX9 couples actin assembly to phosphoinositide signals and is required for membrane remodeling during endocytosis. *Developmental Cell* 13 (1), 43–56.
- You, J.J., Lin-Chao, S., 2010. Gas7 functions with N-WASP to regulate the neurite outgrowth of hippocampal neurons. *Journal of Biological Chemistry* 285 (15), 11652–11666.
- Zirngibl, R., Schulze, D., Mirski, S.E., Cole, S.P., Greer, P.A., 2001. Subcellular localization analysis of the closely related Fps/Fes and Fer protein-tyrosine kinases suggests a distinct role for Fps/Fes in vesicular trafficking. *Experimental Cell Research* 266 (1), 87–94.

INTERORGANELLAR COMMUNICATION: INTERPLAY AND PROCESSES

Contents

Role of Phosphoinositides in Membrane Traffic

Unconventional Protein Secretion: Fibroblast Growth Factor 2 and Interleukin-1 β as Examples

Endoplasmic Reticulum Stress in Disease

Role of Phosphoinositides in Membrane Traffic

DA Sheffield, R Gurung, and CA Mitchell, Monash University, Clayton, VIC, Australia

© 2016 Elsevier Inc. All rights reserved.

Glossary

14-3-3 proteins Family of proteins with protean functions, including acting as a scaffold for tubulation and fission from the Golgi complex.

Adaptor protein 180 (AP180) AP180 N-terminal homology (ANTH)-domain containing protein which binds phosphoinositide 4,5 bisphosphate (PtdIns(4,5)P₂) and clathrin to promote endocytosis.

Adaptor protein 2 (AP2) Multimeric protein complex, which interacts with phosphoinositides by positively charged domains and clathrin to promote endocytosis.

Adaptor protein containing pleckstrin homology (PH) domain phosphotyrosine binding (PTB)-domain and leucine zipper motif (APPL) Rab5-binding proteins which are on a subset of very early endosomes and are a precursor to phosphoinositide-3-phosphate (PtdIns(3)P)-positive early endosomes.

Adenosine diphosphate(ADP)-ribosylation factor 1 (Arf1) Small guanosine 5'-triphosphatase (GTPase) protein which mediates coat protein dynamics and participates in coincidence binding of effector proteins at the Golgi together with phosphoinositide 4 phosphate (PtdIns(4)P).

Atg12-Atg5-Atg16L complex It regulates autophagosome maturation by promoting the lipidation of microtubule associated protein light chain 3 (LC3).

Calcium-dependent activator protein for secretion (CAPS) PtdIns(4,5)P₂-binding protein involved in plasma membrane exocytosis by priming calcium-dependent activator vesicles and organizing SNARE proteins.

Ceramide trafficking protein (CERT) PH-domain containing protein which binds PtdIns(4)P and is involved in endoplasmic reticulum to Golgi trafficking of ceramide.

Class I PI3-kinase (PI3KC) Phosphoinositide kinase which catalyzes PtdIns(4,5)P₂ to phosphatidylinositol (3,4,5)-trisphosphate (PtdIns(3,4,5)P₃).

Class II PI3-kinase (PI3KC2A/PI3KC2 α) Phosphoinositide kinase which catalyzes PtdIns to PtdIns(3)P and PtdIns(4)P to phosphatidylinositol (3,4)-bisphosphate (PtdIns(3,4)P₂).

Class III PI3-kinase (PI3KC3) Phosphoinositide kinase which catalyzes PtdIns to PtdIns(3)P.

Class III PI3-kinase complex Autophagy initiation complex of Class III PI3-kinase (Vps34)-Beclin-1(Atg6)-Vps15-Atg14L.

Class II PI4-kinase II α Phosphoinositide kinase which mediates formation of PtdIns(4)P at the Golgi best described in Golgi to endosome trafficking.

Class III PI4-kinase III β Phosphoinositide kinase which mediates formation of PtdIns(4)P at the Golgi best described in Golgi to plasma membrane trafficking.

Clathrin Polymerized clathrin molecules form a coat to facilitate vesicle formation.

CtBP1-S/BARS Protein present in the nucleus and cytosol with functions in vesicle fission from membranes.

Double FYVE containing protein 1 (DFCP1) FYVE-domain containing protein involved in autophagosome genesis and found on omegasomes.

Dynamin PH-domain containing protein, which has GTPase activity and is involved in membrane scission for vesicle formation.

Early endosomal antigen 1 (EEA1) FYVE-domain containing protein present at early endosomes and involved in endosomal fusion which interacts with PtdIns(3)P.

Endosomal sorting complexes required for transport complexes 0-III (ESCRT0-III) Multiprotein subunit complexes involved in maturation and ubiquitination of endosomes.

Epidermal growth factor receptor (EGFR) Cell surface receptor for signaling via epidermal growth factor ligands.

Four phosphate adaptor proteins (FAPP) PH-domain containing protein effectors of PtdIns(4)P which also

bind Arf1 and mediate Golgi to plasma membrane secretion.

FYVE and coiled-coil domain containing 1 (FYCO1) FYVE-domain containing protein which regulates autophagosome motility away from the perinuclear region 1.

Golgi localized gamma ear containing ADP-ribosylation factor binding (GGA) Multidomain coat proteins which interact with PtdIns(4)P and coordinate traffic from the Golgi to endosomes.

GTPase activating protein (GAP) Protein which facilitate Rab GTP to GDP transition to promote Rab membrane dissociation.

Guanine nucleotide displacement inhibitor (GDI) Protein which maintains Rab proteins in the cytosol by masking the prenylated tail moiety of the Rab.

Guanine nucleotide exchange factor (GEF) Protein which facilitates Rab GDP to GTP transition to promote Rab recruitment to membrane sites.

Hepatocyte growth regulated tyrosine kinase substrate (Hrs) FYVE-domain containing protein recruited to endosomes and involved in the genesis of multivesicular bodies as part of the ESCRT0 complex.

Homotypic fusion and vacuole sorting complex (HOPS complex) Multisubunit complex containing Vps11, Vps16, Vps18, and Vps33 which coordinates late endosome homotypic fusion.

Inpp5b (75-kDa 5-phosphatase) Phosphoinositide 5-phosphatase with activity against multiple phosphoinositides which forms a complex with a number of endocytic Rab proteins.

Inpp5e Phosphoinositide 5-phosphatase with activity against multiple phosphoinositides which has a role in phagocytosis and when mutated results in ciliopathies.

LC3/GABARAP/GATE16 (Atg8) It remains bound to the maturing autophagosome to form the autophagolysosome.

Lysosomal-associated membrane protein 1 (lamp1) Type-1 transmembrane protein present on lysosomes and late endosomes.

Mammalian target of rapamycin complex 1 (mTORC1) Protein complex which senses nutrient state and acts in growth signaling pathways and regulates autophagy.

MUNC13 PtdIns(4,5)P₂-binding protein involved in plasma membrane exocytosis by priming vesicles and organizing SNARE proteins.

Myotubularin (MTM1) Phosphoinositide 3-phosphatase with roles in muscle function, autophagy, and myotubularin endosomal trafficking.

Myotubularin-related proteins 1-14 (MTMR1-14) Phosphoinositide 3-phosphatases with varying subcellular distributions and roles in endocytic trafficking and autophagy.

Neuronal calcium sensor-1 (NCS-1) Calcium sensor protein associated with vesicle secretion at the Golgi and in neurons.

Oxysterol-binding protein 1 (OSBP1) PH-domain containing protein which binds PtdIns(4)P and Arf1 and regulates sphingomyelin synthesis.

Phaphin2 FYVE-domain containing protein recruited to early endosomes with a role in endosomal fusion.

Phosphatidylinositol (3)-phosphate (PtdIns(3)P) Signaling phosphoinositide present on early and late endosomes with roles in coordinating endosomal trafficking.

Phosphatidylinositol (3,4)-bisphosphate (PtdIns(3,4)P₂) Signaling phosphoinositide present at the plasma membrane and endocytic vesicles with roles in growth factor signaling and emerging role in endocytosis .

Phosphatidylinositol (3,4,5)-trisphosphate (PtdIns(3,4,5)P₃) Signaling phosphoinositide present at low levels in resting cells and rapidly synthesized on stimulation with it roles including growth factor signaling macropinocytosis and phagocytosis.

Phosphatidylinositol (3,5)-bisphosphate (PtdIns(3,5)P₂) Signaling phosphoinositide present on late endosomes with roles in endosomal phosphatidylinositol maturation (3,5)-bisphosphate.

Phosphatidylinositol (4)-phosphate (PtdIns(4)P) Signaling phosphoinositide present at the Golgi with roles in coordinating vesicle phosphatidylinositol trafficking (4)-phosphate.

Phosphatidylinositol (4,5)-bisphosphate (PtdIns(4,5)P₂) Signaling phosphoinositide present at the plasma membrane, Golgi, and nuclear membranes important in many processes including endocytosis, exocytosis, and actin dynamics.

Phosphatase and tensin homolog (PTEN) Phosphoinositide 3-phosphatase with activity against PtdIns(3,4,5)P₃ and regulates cell growth, differentiation, and polarity as a tumor suppressor.

PIKfyve/PIP5K3 Phosphoinositide kinase which converts PtdIns(3)P to PtdIns(3,5)P₂ and PtdIns to PtdIns(5)P.

Rab4 Small GTPase which is recruited to membrane sites when in GTP-bound form and important in recycling endosome function.

Rab5 Small GTPase which is recruited to membrane sites when in GTP-bound form and important in both early and late stages of endocytosis as well as early endosome function.

Rab11 Small GTPase which is recruited to membrane sites when in GTP-bound form and important in recycling endosome function.

Rabenosyn5 FYVE-domain containing protein effector which is recruited to early endosomes and involved in endosome fusion.

Retromer Multisubunit complex containing Vps26p, Vps29p, Vps35p, Vps5p (SNX1 and SNX2), and Vps17p (SNX4 and SNX5) which is involved in endosome to Golgi transport.

SAND-1/Mon1 *Caenorhabditis elegans* protein which directs a switch from early to late endosomes.

SHIP2 Phosphoinositide 5-phosphatase with activity against multiple phosphoinositides and a role in endocytosis, phagocytosis, and insulin signaling.

Soluble NSF attachment protein receptors (SNARES) Family of proteins which mediate vesicle fusion with membranes.

Sorting nexin 3 (SNX3) PX-domain containing protein which acts at endosomes to direct recycling and maturation.

Sorting nexin 4 (SNX4) BAR- and PX-domain containing protein which is involved in the trafficking of transferrin receptor from early endosomes to the recycling compartment.

Sorting nexin 5 (SNX5) BAR- and PX- domain containing protein which induces membrane curvature and regulates macropinocytosis and retrograde trafficking.

Sorting nexin 9 (SNX9) BAR- and PX-domain containing protein which induces membrane curvature and has a role in the late stages of endocytosis.

Synaptojanin-1 Phosphoinositide 5-phosphatase with activity against multiple phosphoinositides and a role in late stages of endocytosis facilitating PtdIns(4,5)P₂ turnover.

Type-I inositol 4-phosphatase

(Inpp4a) Phosphoinositide 4-phosphatase which

degrades PtdIns(3,4)P₂ to generate endosomal PtdIns(3)P and regulates AKT signaling.

Type-II inositol 4-phosphatase

(Inpp4b) Phosphoinositide 4-phosphatase with a role in turnover of PtdIns(3,4)P₂ and endocytosis to endosome transition.

Uncoordinated 51-like kinase (ULK1/ULK2) Serine/threonine protein kinase involved in autophagy induction (Atg1) which forms a complex with mTORC1-Atg13-FIP200 (Atg17)-Atg101 and feeds back to the Class III PI3-kinase complex to increase PtdIns(3)P.

WD-repeat domain containing phosphoinositide interacting proteins (WIPI proteins) WIPI 1/2 (Atg18) and WIPI 3/4 are PtdIns(3)P effectors which can inhibit autophagy (WIPI1) and promote autophagosome maturation (WIPI2b).

Phosphoinositides at the Membrane–Cytosol Interface

Phosphoinositides are glycerophospholipids that contain a diacylglycerol (DAG) lipid membrane-embedded backbone, and a myoinositol head group, that protrudes from membranes into the cytosol. The inositol ring can be phosphorylated at any or all of three possible hydroxyl groups, permitting the generation of seven different possible phosphoinositide species (Figure 1; Balla, 2013). Phosphoinositides constitute only a small component of the total phospholipid within eukaryote cells, however play a significant role in both signaling and trafficking events. The phosphorylation status of the inositol head group allows different phosphoinositides to interact with high specificity with a variety of effector proteins, and it is this selectivity in protein binding that confers the ability of phosphoinositides to act as master organelle traffic controllers (Rusten and Stenmark, 2006; Table 1). The lipid soluble tail localizes phosphoinositides to membrane compartments, and the variably phosphorylated head group lies at the interface between the membrane and the cytosol, a specialized locale containing a complex array of proteins, including enzymes, cytoskeletal structures, and trafficking motors. Contingent on the appearance of a specific phosphoinositide within the cytoplasmic leaflet of an organelle membrane region, there is recruitment of dedicated protein group scaffolds which translate into a plethora of phosphoinositide-dependent functional outcomes.

Organelle Compartments the Spatial Restriction of Phosphoinositide Species and Organelle Identity

A number of generalized membrane compartments can be recognized within the cell, and there is extensive vesicle traffic between these organelles, carrying lipids, membrane-bound proteins such as receptors, and cargo, such as growth factors. There is a distinct subcellular distribution of different

phosphoinositide species, with enrichment at specific organelles, contributing to the identity of that organelle (Figure 2). Membrane trafficking functions include vesicle budding, fission, fusion, and motility. Flow occurs from the plasma membrane, enriched in phosphoinositide 4,5 biphosphate (PtdIns(4,5)P₂), into vesicles, which communicate with the early endosomes, an initial way station for processing. Early endosomes are enriched in phosphoinositide-3-phosphate (PtdIns(3)P), and communicate with the recycling network, by sorting directly back to the plasma membrane, or via a more delayed route via the endocytic recycling compartment (ERC), which exhibits a perinuclear distribution in many cells. Alternatively, vesicles can exit the sorting endosomes in a pathway destined for degradation, and become late endosomes or multivesicular endosomes decorated by PtdIns(3)P and phosphoinositide 3,5 biphosphate (PtdIns(3,5)P₂) to ultimately fuse with acidic and proteolytic lysosomes. Vesicles can also transit from sorting endosomes back to the Golgi in the retrograde pathway, a process which requires PtdIns(3,5)P₂. The endoplasmic reticulum (ER) is the site of protein synthesis, and communicates with both the nucleus and the Golgi compartment, enriched in phosphoinositide 4 phosphate (PtdIns(4)P). Proteins sorted in the Golgi exit the *trans*-Golgi network (TGN) and are sorted via vesicles.

Organelle identity, conferred by the presence of one or more specific phosphoinositide species, is also determined by specific small guanosine 5'-triphosphatase (GTPase) proteins, including the Rab and Arf proteins. These proteins are part of the Ras superfamily, and can exist in an active GTP-bound state, or an inactive guanosine 5'-diphosphatase (GDP)-bound form. Accessory proteins include guanine nucleotide exchange factors (GEFs), that facilitate conversion of a Rab to its GTP-bound, membrane-associated state, and GTPase activating proteins (GAPs) which facilitate conversion back to the GDP-bound state promoting membrane dissociation (Mizuno-Yamasaki *et al.*, 2012). Rabs are chaperoned by guanine nucleotide displacement inhibitors (GDIs), which act to

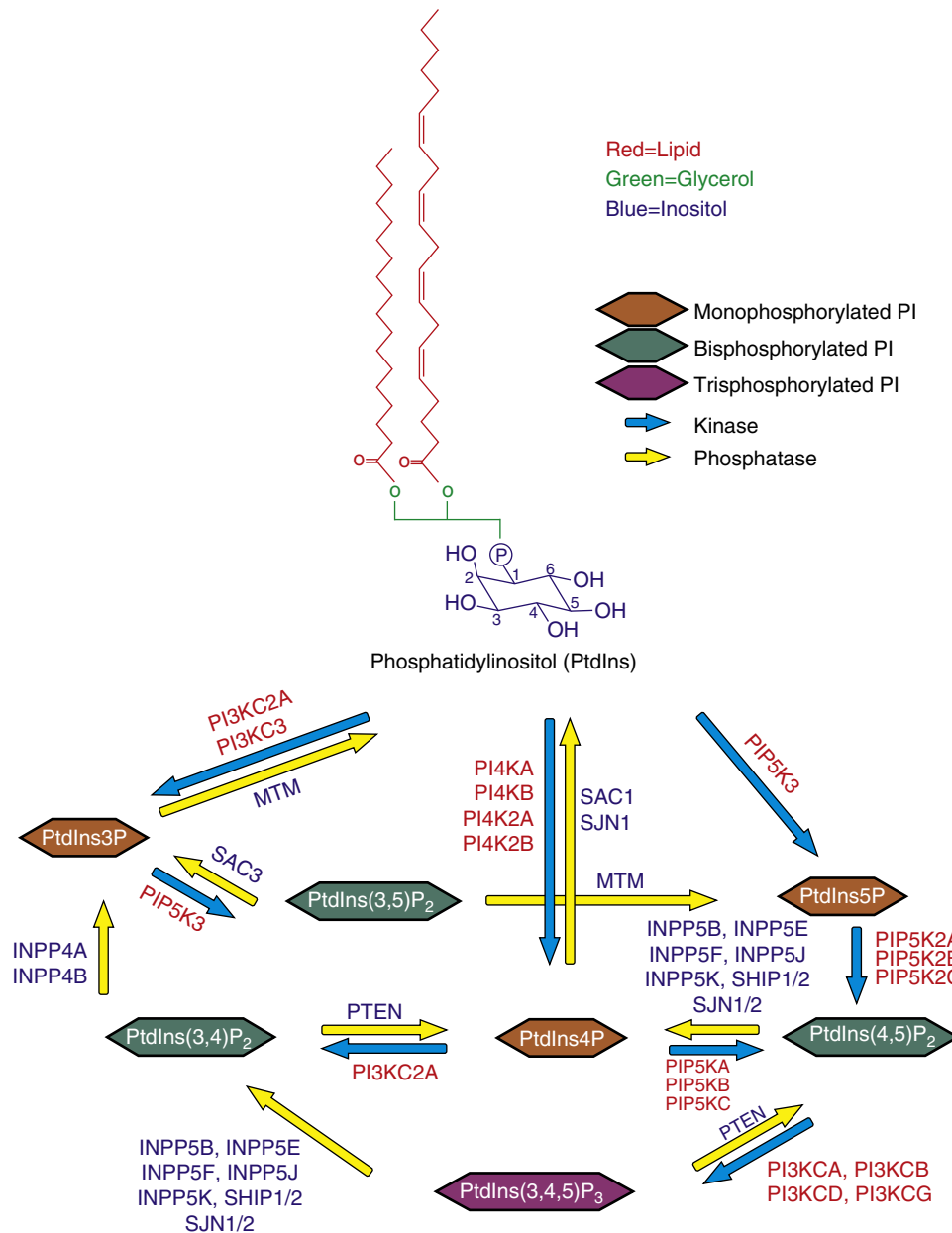


Figure 1 Phosphoinositide turnover by kinases and phosphatases. Seven different species of phosphoinositides can be formed from the differential phosphorylation of hydroxyl groups on the myoinositol head group of phosphatidylinositol (PtdIns). Different phosphoinositide kinases and phosphatases present at the membrane/cytosolic interface are able to convert one phosphoinositide into another, by the addition or removal of phosphate groups to the inositol head group, respectively. The subcellular distribution, and activation status of these kinases and phosphatases thus dictates the pool of phosphoinositide species present on individual organelles.

maintain Rabs in the cytosol by masking their prenylated tail moiety, but permit cycling into the membrane where they can be stabilized by GEF function into their GTP-bound form (Mizuno-Yamasaki *et al.*, 2012).

Organelle function is dependent on the requirement for the combined presence of both a specific phosphoinositide as well as a Rab or ARF protein to coordinately interact in a process termed coincidence detection, or cooperative binding, with membrane recruited protein effectors (Di Paolo and De Camilli, 2006). The phosphoinositide and small GTPase regulator need to be present and cooperate to enable the

switch to induce the correct functional protein scaffold in order to signal the subsequent membrane trafficking and thereby contribute to essential organelle function.

The regulated turnover of individual phosphoinositides in a controlled fashion is also necessary for coordinated organelle trafficking. The highly specific nature of the interaction between individual phosphoinositides and dedicated binding proteins depends upon the number and configuration of phosphate groups on the myoinositol head group. Thus the addition or removal of a phosphate group, by phosphoinositide kinases or phosphatases, respectively (Figure 1), dictates

Table 1 Phosphoinositide binding domains

Domains		Phosphoinositides bound	Protein examples	References
C2	Protein kinase C conserved region 2	PtdIns(4,5)P ₂	PI3KC2 α , CAPS, PKC α , synaptotagmin	Corbalan-Garcia and Gómez-Fernández (2014)
PTB	Phosphotyrosine binding	PtdIns(4,5)P ₂	Dab2, Mint1	Alajlouni <i>et al.</i> (2011); Weber-Boyvat <i>et al.</i> (2013)
BAR	Bin/amphiphysin/Rvs	PtdIns(4,5)P ₂ , PtdIns(3,4,5)P ₃	Amphiphysin, endophilin1	Yoon <i>et al.</i> (2012)
	Fab1/YOTB/ZK632/Vac1p/EEA1	PtdIns(3)P	EEA1, Hrs, DFCEP1	Kutateladze (2007)
PX	Phox homology	PtdIns(3)P, PtdIns(3,4)P ₂ , PtdIns(4,5)P ₂ , PtdIns(3,4,5)P ₃	SNXs, p40 _{phox} , PI3KC2 α	Kutateladze (2007)
ML1N		PtdIns(3,5)P ₂	TRPML	Li <i>et al.</i> (2013)
Proppin		PtdIns(3,5)P ₂ and PtdIns(3)P	Atg18p	Dove <i>et al.</i> (2009)
ENTH	Epsin N-terminal homology	PtdIns(4,5)P ₂	Epsin	Hom <i>et al.</i> (2007)
ANTH	AP180 N-terminal homology	PtdIns(4,5)P ₂	AP180	Hom <i>et al.</i> (2007)
PH	Pleckstrin homology	PtdIns(4,5)P ₂ , PtdIns(3,4)P ₂ , PtdIns(3,4,5)P ₃ , PtdIns(4)P	PLC δ , FAPP1, OSBP1	Lemmon (2007)
FERM	4.1/ezrin/radixin/moesin	PtdIns(4,5)P ₂	Ezrin, moesin	Maniti <i>et al.</i> (2012)
GLUE	GRAM-like ubiquitin binding in EAP45	PtdIns(3,4,5)P ₃ , PtdIns(3)P	Eap45, Vps36	Slagsvold <i>et al.</i> (2005); Teo <i>et al.</i> (2006)
PDZ	PSD-95/Discs-large/ZO-1	PtdIns(4,5)P ₂ , PtdIns(3,4,5)P ₃ , PtdIns(3)P, PtdIns(3,5)P ₂	Zona occludens, Par-3, syntenin-1	Wawrzyniak <i>et al.</i> (2013)
PH-GRAM	Pleckstrin homology glucosyltransferases, Rab-like GTPase activators and myotubularins	PtdIns(5)P, PtdIns(3,5)P ₂	Myotubularins	Choudhury <i>et al.</i> (2006)
EHD	Eps15 homology domain	PtdIns4P and PtdIns(4,5)P ₂	EHD1-4	Jovic <i>et al.</i> (2009)

Note: Protein effectors are recruited to phosphoinositides and bind via specific domains, as shown. AP180, clathrin assembly protein 180; CAPS, calcium-dependent activator protein for secretion; dab2, disabled homolog 2; DFCEP1, Double FYVE containing protein 1; EEA1, Early endosomal antigen 1; EHD, Eps15 homology domain; FAPP1, four-phosphate-adaptor protein 1; Hrs, Hepatocyte growth factor-regulated tyrosine kinase substrate; OSBP 1, oxysterol-binding protein 1; PI3KC2 α , Class II phosphoinositide 3 kinase α ; PKC α , protein kinase C α ; PLC δ , phospholipase C δ ; SNXs, sorting nexins; TRPML, mucolipin subfamily of transient receptor potential 1.

the ability to recruit protein effectors to a membrane region. Turnover of phosphoinositides via conversion of one species to another by phosphoinositide phosphatases and kinases represents a switch to commence or cease a trafficking or signaling event. In many situations efficient progression of complex trafficking steps requires phosphoinositide inter-conversion of one phosphoinositide species into another (Cullen and Carlton, 2012).

Phosphoinositides in Plasma Membrane Trafficking

Endocytosis

PtdIns(4,5)P₂ plays a major role in plasma membrane trafficking pathways, including endocytosis, the internalization of extracellular components into intracellular membrane-bound vesicles, and exocytosis, the egress of material from intracellular vesicles to the extracellular environment. Endocytosis

occurs by the invagination of membrane buds containing extracellular material, and includes coat protein clathrin-mediated endocytosis, non-coat protein-mediated endocytosis as well as specialized internalization mechanisms such as macropinocytosis and phagocytosis. The critical role PtdIns(4,5)P₂ plays in clathrin-mediated endocytosis illustrates many of the key themes representative of phosphoinositide membrane trafficking (Figure 3). PtdIns(4,5)P₂ is enriched at clathrin-coated pits as well as caveolae (Fujita *et al.*, 2009), and regulates the initial stages of endocytosis, coordinating pit formation. Coated pit initiation requires the PtdIns(4,5)P₂-dependent recruitment of effector proteins to the plasma membrane, including AP2 and AP180 (Godlee and Kaksonen, 2013). The N-terminal regions of AP2 and AP180 contain positively charged domains that interact specifically with PtdIns(4,5)P₂, and AP180 contains a phosphoinositide binding AP180 N-terminal homology (ANTH) domain (Gaidarov and Keen, 1999; Hom *et al.*, 2007). Experimental manipulation to reduce the levels of PtdIns(4,5)P₂ inhibits endocytosis,

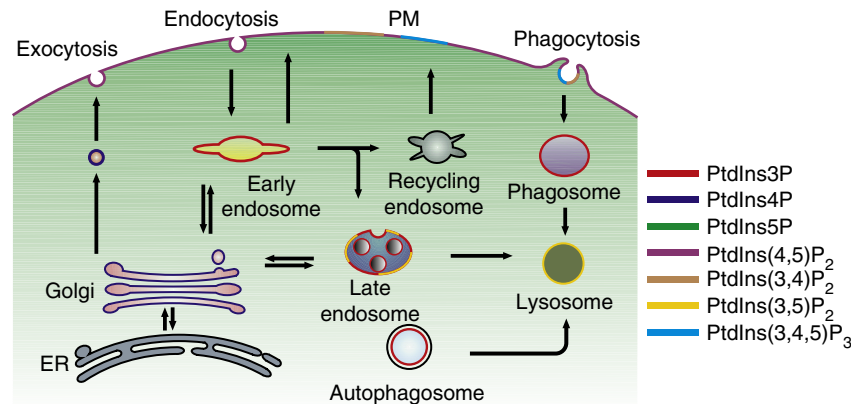


Figure 2 Phosphoinositide subcellular distribution to membrane compartments. Different phosphoinositide species are selectively enriched on distinct membrane organelle regions, a phosphoinositide code for organelle identity. There is flow of cargo between the organelles, and turnover of different phosphoinositide species is required for coordinating the orderly function of many membrane processes, as illustrated. Membrane identity is conferred by the individual phosphoinositide species in combination with small GTPases and specific effector proteins (see detailed figures below).

disrupts the recruitment of many of the proteins to the complex endocytosis machinery, and impairs pit invagination (Abe *et al.*, 2008; Sun and Drubin, 2012; Zoncu *et al.*, 2007).

Rab5 is a GTPase protein present at the plasma membrane and on early endosome membranes. *In vitro* Rab5 is required for the early steps of clathrin-coated pit formation, where it is necessary for ligand internalization (McLauchlan *et al.*, 1998). In *Drosophila*, Rab5 is required for endocytosis, and its recruitment to the cortical region under the plasma membrane depends on the presence of PtdIns(4,5)P₂ (Compagnon *et al.*, 2009).

Turnover of PtdIns(4,5)P₂ is required for normal endocytosis progression. Cells with increased PtdIns(4,5)P₂ levels have enlarged pits (Antonescu *et al.*, 2011; Sun and Drubin, 2012). Inositol polyphosphate 5-phosphatases, such as synaptojanin-1 are recruited to clathrin-coated pits, and promote maturation of pits, in part via the degradation of PtdIns(4,5)P₂. This leads to a phosphoinositide switch, whereby PtdIns(4,5)P₂ is converted to PtdIns(4)P via 5-phosphatase activity and PtdIns(3,4)P₂ is formed at clathrin-coated pits by the activity of Class II PI3-kinase C2α (PIK3C2A), as recently suggested by one report (Posor *et al.*, 2013). This switch is important for the normal maturation of clathrin-coated pits and is required for the recruitment of the effector sorting nexin 9 (SNX9) (Posor *et al.*, 2013). Fission to form an intracellular vesicle is mediated by the PtdIns(4,5)P₂-effector dynamin which contains a PH domain (Balla, 2013). In neurons, synaptojanin-1 is required for efficient endocytosis and re-availability of synaptic vesicles for neurotransmission, postulated to be a result of the need for removal of PtdIns(4,5)P₂ to allow the uncoating of internalized vesicles (Mani *et al.*, 2007). Synaptojanin-1 knockout mice exhibit an accumulation of clathrin-coated vesicles (Hayashi *et al.*, 2008). Endophilin is a BAR domain containing protein which recruits synaptojanin-1 to clathrin-coated pits, and thus facilitates PtdIns(4,5)P₂ degradation prior to dynamin-mediated fission (Milosevic *et al.*, 2011). Additionally, the small GTPase Rab5, and with its activating protein (GEF) Rabex 5 are recruited to clathrin-coated pits, to promote the exclusion of PtdIns(4,5)P₂ from the formed vesicle and vesicle uncoating (Compagnon *et al.*, 2009; Semerdjieva *et al.*, 2008). Thus endocytosis

requires both the enrichment and localization of PtdIns(4,5)P₂ and its progression involves PtdIns(4,5)P₂ regulated turnover.

In contrast to the above events, in some stimulated forms of endocytosis there is a prominent regulatory role for phosphoinositide 3,4,5 trisphosphate PtdIns(3,4,5)P₃, the product of Class I PI3-kinase action on PtdIns(4,5)P₂. Endocytosis of the G-protein coupled β-adrenergic receptor is important for down-regulation of the adrenergic signal and requires PtdIns(3,4,5)P₃ (Naga Prasad *et al.*, 2002). Agonist occupation of the β-adrenergic receptor results in the recruitment of the Class I PI3-kinase via β-adrenergic receptor kinase, which enhances AP2 recruitment and clathrin-coated pit formation (Naga Prasad *et al.*, 2002).

Phagocytosis, the engulfment of large particles such as apoptotic cells or microorganisms and macropinocytosis, the internalization of extracellular fluid, are specialized forms of receptor-mediated endocytosis (Figure 4). Following receptor stimulation, extensive changes in the actin cytoskeleton occur, to produce pseudopods, in the case of phagocytosis, and ruffling, in the case of macropinocytosis. Both of these processes are characterized by the appearance of PtdIns(3,4,5)P₃ in the membrane of pseudopods and late-stage ruffles (Swanson, 2008). PtdIns(3,4,5)P₃ binds WAVE, an actin nucleation promoting factor, and signals via the Rho GTPases Rac and Cdc42 to contribute to the dramatic actin dynamics observed during phagocytosis (Padrick and Rosen, 2010). Phosphoinositide 5-phosphatases can degrade PtdIns(3,4,5)P₃ to form PtdIns(3,4)P₂, at phagocytic cups and macropinosomes (Horan *et al.*, 2007; Maekawa *et al.*, 2014). The SNX5 may be recruited by PtdIns(3,4)P₂ at the macropinosome, where it regulates the number and size of macropinosomes and tubules (Lim *et al.*, 2012). Turnover of PtdIns(3,4)P₂ is also essential for the completion of macropinocytosis, with sequential actions of the 4-phosphatase Inpp4b, and the 3-phosphatases MTMR6 and MTMR9 controlling the phosphoinositide flux in *Caenorhabditis elegans* macropinocytosis (Maekawa *et al.*, 2014).

However, further work is needed to establish a clear role for PtdIns(3,4,5)P₃ in the regulation of other clathrin- and non-clathrin-mediated endocytocytic events. Knockdown of

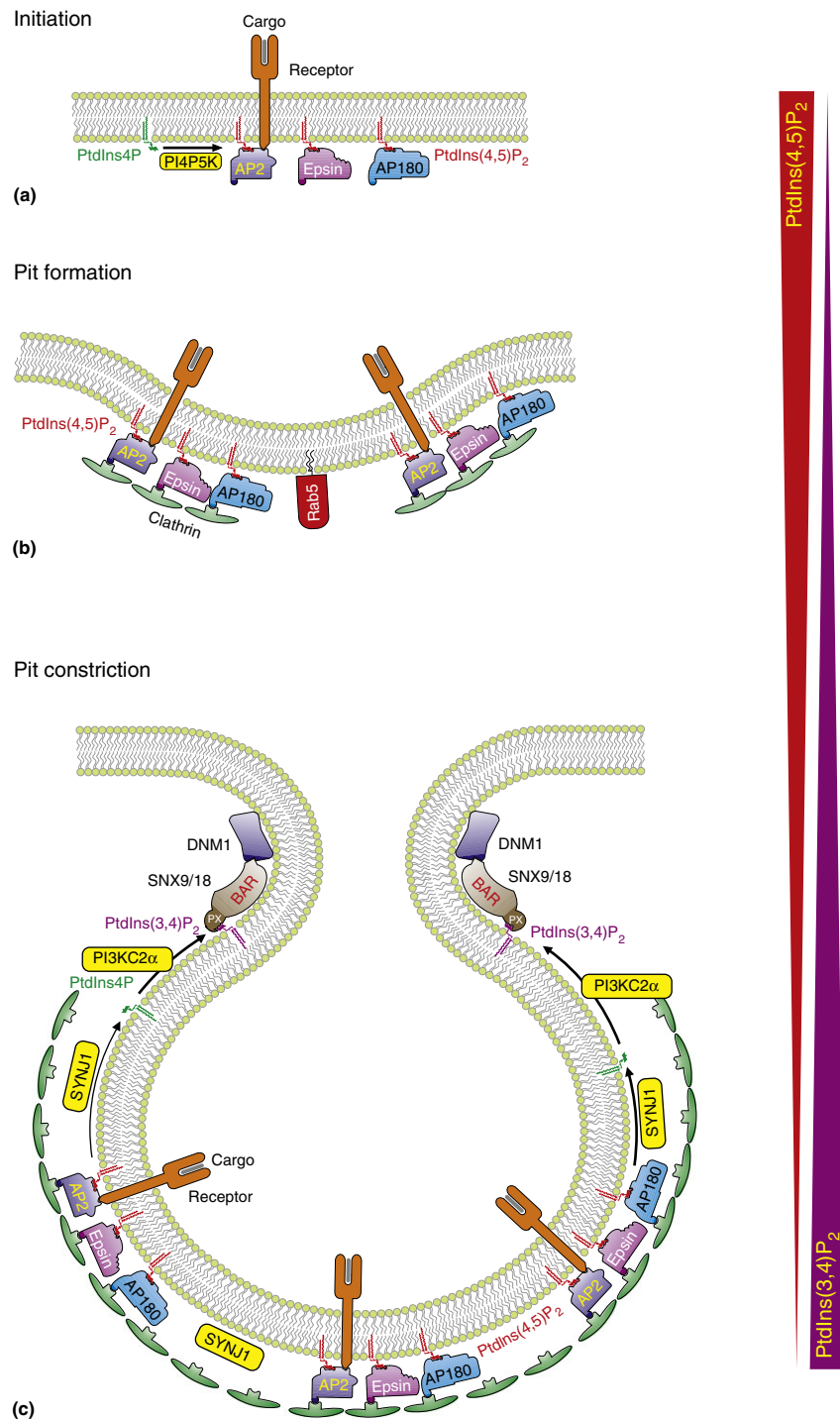


Figure 3 Endocytic steps and phosphoinositide engagement. PtdIns(4,5)P₂ and PtdIns(3,4)P₂ are sequentially involved in endocytosis. Shown is a schematic for clathrin-mediated endocytosis. Multiple adaptor and accessory proteins assemble at the site of the receptor, relying on combined coincidence detection to overcome otherwise weak interactions. AP2 and AP180 binding require PtdIns(4,5)P₂ present in membrane patches, essential for the initiation of the clathrin-coated pit. This results in pit invagination and progression of clathrin-coat formation. Later stages of endocytosis require turnover of PtdIns(4,5)P₂, mediated by the sequential recruitment of the 5-phosphatase synaptojanin-1 (SYNJ1), and then clathrin-dependent PI3KC2α to generate PtdIns(3,4)P₂. Rab5 is also present on clathrin-coated vesicles and may also contribute to loss of PtdIns(4,5)P₂. Membrane curvature and PtdIns(3,4)P₂ may recruit the PX/BAR domain containing sorting nexin 9 (SNX9) and possibly SNX18 and subsequently dynamin (DNM1), an additional PtdIns(4,5)P₂ effector, to mediate fission of the clathrin-coated vesicle. This then undergoes uncoating prior to entering the early endosomal system.

the inositol polyphosphate Type 1 α 4-phosphatase inhibits Rab5-dependent transferrin uptake (Shin *et al.*, 2005). SHIP2 5-phosphatase, which has a preference for PtdIns(3,4,5)P₃ as a substrate, is recruited to clathrin-coated pits to negatively regulate their growth, however SHIP2 also has activity against PtdIns(4,5)P₂ (Nakatsu *et al.*, 2010). Experimental manipulation to increase plasma membrane PtdIns(3,4,5)P₃ levels by directly activating Class I PI3-kinase results in shorter clathrin-coated pit lifetimes, either a direct result of PtdIns(3,4,5)P₃ signals, or possibly representing the impact of turnover of PtdIns(4,5)P₂ (Nakatsu *et al.*, 2010). However, overexpression of PTEN, which results in reduced plasma membrane PtdIns(3,4,5)P₃, has no effect on endocytosis of transferrin (Posor *et al.*, 2013).

Exocytosis

During exocytosis, vesicles mobilize to the plasma membrane, where they fuse to intermix and deliver their cargo to the extracellular space. Exocytosis requires the coordination of motor machinery such as myosins and kinesins, the cytoskeletal network including actin microtubules together with membrane fusion apparatus including SNAREs. Exocytosis has been studied extensively in neuroendocrine cells, specifically exocytosis of dense core vesicles, which is calcium dependent, and to a lesser degree in neuronal synaptic vesicle exocytosis, as well as constitutive exocytosis from the Golgi (Martin, 2012). Examples of specialized exocytosis occur in neurotransmitter transmission and leukocyte cytotoxicity. PtdIns(4,5)P₂ localizes to sites of exocytosis (James *et al.*, 2008). Increasing or decreasing PtdIns(4,5)P₂ signals results in corresponding increased or decreased exocytosis, respectively (Martin, 2012). Effector proteins which bind PtdIns(4,5)P₂ include CAPS and MUNC13 (Kabachinski *et al.*, 2014). Paradoxically, it has been suggested that PtdIns(4,5)P₂ alone can inhibit vesicle fusion, until additional factors such as the effector CAPS are present (James *et al.*, 2008). CAPS requires PtdIns(4,5)P₂ for its activity

in priming, the process by which vesicles docked at a recipient membrane site become fusion ready (Kabachinski *et al.*, 2014). MUNC13 is recruited to vesicles by PtdIns(4,5)P₂ at the plasma membrane in a calcium-dependent manner. Under high agonist stimulation conditions, phospholipase C η 2 hydrolyzes PtdIns(4,5)P₂ into DAG which itself can regulate fusion (Kabachinski *et al.*, 2014). PtdIns(4,5)P₂ also reorganizes actin to mobilize secretory vesicles for exocytosis (Wen *et al.*, 2011). In yeast, PtdIns(4,5)P₂ is required for secretion from the Golgi to the plasma membrane. PtdIns(4,5)P₂ interacts to recruit components of the exocyst, a multisubunit complex that facilitates the recruitment and fusion of vesicles with the plasma membrane (Martin, 2012).

There is an additional role for PtdIns(3)P in specialized exocytosis. Class II PI3-kinase-derived PtdIns(3)P is required for the priming of large dense core vesicles for exocytosis in neurons (Meunier *et al.*, 2005). This is negatively regulated by the turnover of PtdIns(3)P via the PIKfyve kinase, converting PtdIns(3)P to PtdIns(3,5)P₂ (Osborne *et al.*, 2008). Class II PI3-kinase also regulates insulin secretion from granules in pancreatic beta cells, and insulin stimulated GLUT4 vesicle exocytosis in skeletal muscle (Dominguez *et al.*, 2011; Falasca and Maffucci, 2007).

Phosphoinositides in the Endosomal Compartment – Early Endosomes, Recycling Endosomes, and Late Endosomes

Endocytosed cargo within their membrane vesicles become part of a way station known as the early, or sorting endosomal compartment (Figure 5). Early endosomes in many cells are located in the cell periphery, exhibit a characteristic tubule-vesicular morphology, and represent a site for the direction of vesicle traffic via a number of routes. In the recycling pathway, cargo and membrane, can branch off in tubules from sorting endosomes, to either recycle directly to the plasma membrane,

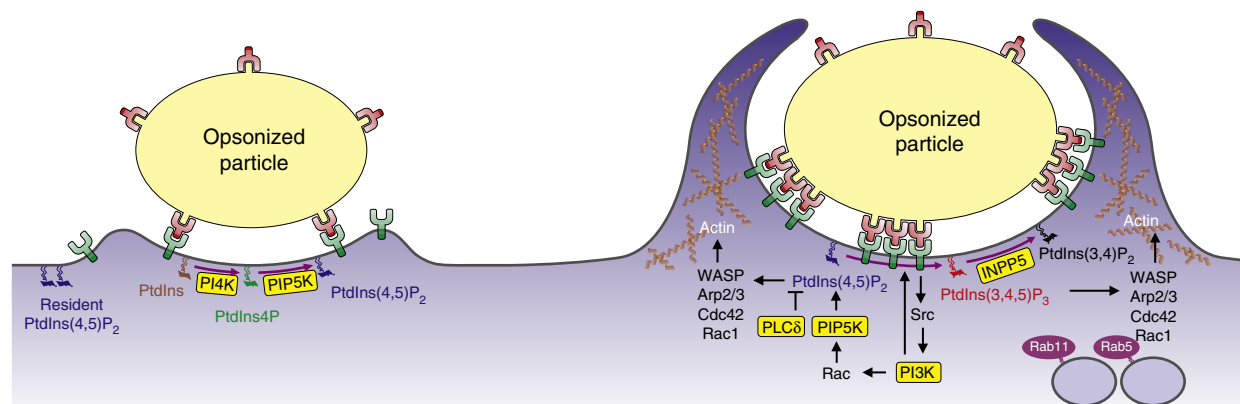


Figure 4 Phosphoinositide pathway in phagocytosis. Phagocytosis is an example of a specialized form of endocytosis in which PtdIns(3,4,5)P₃ has a prominent role. Shown is a schematic depicting Fc γ R-mediated phagocytosis. Antibody opsonised particles engage multiple Fc γ Rs via their Fc domains, clustering the receptors and initiating signaling via the Src tyrosine kinases which activates Fc γ Rs, to recruit Syk and Class I PI3-kinase. PtdIns(4,5)P₂ present in the plasma membrane is enriched at phagocytic cups, together with PIP5K α . Class I PI3-kinase generates PtdIns(3,4,5)P₃ from PtdIns(4,5)P₂ at the phagocytic cup, and these phosphoinositides contribute to the dramatic changes in actin polymerization necessary for the advancement of pseudopods, via interactions with actin nucleating factors such as WAVE, Cdc42, and Rac1. Turnover of PtdIns(4,5)P₂ occurs via phospholipase C δ (PLC δ) and Class I PI3-kinase (SHIP2) and is required for normal actin remodeling to complete phagocytosis. Additionally, PtdIns(3,4,5)P₃ is converted to PtdIns(3,4)P₂ at the phagocytic cup by the action of 5-phosphatases such as Inpp5e. Pseudopod membrane grows by the recruitment of intracellular membrane sources, such as Rab11- and Rab5-positive endosomes.

or for storage in the ERC. Alternatively, within a 5–10 min time point, early endosomes can develop a progressively lower pH, together with the acquisition of a hydrolytic environment in a process termed maturation (Cullen and Carlton, 2012). During maturation internalized cargo is transported in an endosome which converts from an early to a late endosome, a degradative organelle which fuses with lysosomes. Additionally, vesicular traffic can leave the early and late endosomal compartment via the retrograde pathway, for transport of membrane back to the TGN (Cullen and Carlton, 2012).

The early endosome is marked by the monophosphorylated phosphoinositide PtdIns(3)P, Rab5, and the endosomal fusion protein EEA1 (Zoncu *et al.*, 2009). Immediately following endocytosis, some cargos such as epidermal growth factor and nerve growth factor can enter a precursor to early endosomes, which is Rab5-positive, and is labeled by the presence of the membrane adaptor complex APPL, but the absence of EEA1 (Compagnon *et al.*, 2009; Zoncu *et al.*, 2009). The recruitment of PtdIns(3)P is required for the transition from the APPL-positive precursor endosome to APPL-negative, EEA1 positive early endosome (Zoncu *et al.*, 2009). Rab5 is itself involved in the formation of PtdIns(3)P, by recruiting the Class III PI3-kinase which generates PtdIns(3)P from PtdIns. The Class III PI3-kinase subunit p150/Vps15 interacts with GTP-bound Rab5 via HEAT and WD40 domains and localizes to early endosomes (Murray *et al.*, 2002). Early endosomal PtdIns(3)P may also be formed indirectly by the Class I PI3-kinase acting in conjunction with phosphoinositide phosphatases in a sequential cascade on the inner leaflet of the plasma membrane. Class I PI3-kinase catalyzes the formation of PtdIns(3,4,5)P₃ from PtdIns(4,5)P₂ at the plasma membrane, and the degradation of PtdIns(3,4,5)P₃ is in part accomplished by the recruitment of phosphoinositide 5-phosphatases to generate PtdIns(3,4)P₂, and then 4-phosphatases to generate PtdIns(3)P (Shin *et al.*, 2005). Interestingly, Rab5 interacts with the Class I PI3-kinase as well as the 5-phosphatase, Inpp5b, and 4-phosphatase, Inpp4a. It has been estimated that up to 30% of endosomal PtdIns(3)P may be formed in a Rab5-dependent cascade producing a phosphoinositide gradient from PtdIns(3,4,5)P₃ to PtdIns(3)P between the plasma membrane and the early endosome (Shin *et al.*, 2005). In some specialized circumstances, PtdIns(3)P is generated by Class II PI3-kinase on endosomes, including in *Drosophila* hemocytes and at pericentriolar endosomes at the base of primary cilia in mouse embryonic fibroblasts (Franco *et al.*, 2014; Velichkova *et al.*, 2010).

PtdIns(3)P on early endosomal membranes recruits effector proteins containing FYVE and PX domains (Balla, 2013). The FYVE domain is a zinc binding finger, and a large number of proteins have been identified containing this domain (Kutateladze, 2006). In most contexts a single FYVE domain is insufficient for high-affinity PtdIns(3)P binding alone. Many FYVE domain containing proteins have been identified, however, their regulation and function has only begun to be characterized. Of these, EEA1 is the best studied, but whether this is a model for the function of the other FYVE domain containing proteins remains to be shown. Early endosome recruitment of EEA1 is accomplished by FYVE domain interaction with PtdIns(3)P, together with EEA1 dimer formation between EEA1 coiled-coil domains (Kutateladze,

2006). Additionally, EEA1 also binds Rab5 directly by a separate domain adjacent to the FYVE domain, an example of the importance of cooperative binding of an effector to both a Rab GTPase and a phosphoinositide. In the case of other FYVE domain containing proteins, oligomerization (as for early endosomal Hrs and SARA) and tandem FYVE domains (as in the case of the ER and Golgi localized double FYVE containing protein 1 (DFCP1)), have been described (Kutateladze, 2006). The SNX proteins are an example of early endosomal recruitment via PX domains, with some SNXs additionally containing a curved membrane sensing BAR domain, that targets proteins to high-curvature subdomains of the early endosome. This facilitates the pinching off of tubular membrane extensions and fission (Cullen and Carlton, 2012). The SNXs direct cargo and receptors out of the early endosome into the recycling endosomal compartment or back to the secretory pathway via retrograde recycling (Cullen and Carlton, 2012). An example of this is SNX4, present on early and recycling endosomes, which facilitates the channeling of membrane-embedded transferrin and its receptor away from the degradative pathway and into the recycling pathway by virtue of regulating egress of tubular membranes with a higher surface area to volume ratio (Traer *et al.*, 2007).

Sequential transport from early to late endosomes in the degradative pathway involves the development of multi-vesicular bodies, with intraluminal vesicles that sequester ubiquitinated cargo for degradation away from the recycling process. PtdIns(3)P is present on intraluminal vesicles (Cullen and Carlton, 2012). Turnover of PtdIns(3)P can occur via conversion to PtdIns(3,5)P₂ by the PIKfyve complex, and inhibition of PIKfyve produces dramatic endosomal morphologic changes (De Lartigue *et al.*, 2009). PtdIns(3,5)P₂ is important for retrograde trafficking, epidermal growth factor receptor (EGFR) degradation and autophagy. The recent development of a PtdIns(3,5)P₂ probe has demonstrated that PtdIns(3,5)P₂ is predominantly present on Lamp1-positive late endosomes, where it is dynamically enriched immediately preceding late endosomal fusion events (Li *et al.*, 2013).

Intraluminal vesicle development is accomplished by four phosphoinositide interacting complexes, ESCRT 0-III (Cullen and Carlton, 2012). Hrs is a FYVE domain containing protein and a component of ESCRT-0, and ESCRT I-II may bind PtdIns(3)P by other mechanisms (Cullen and Carlton, 2012). The ESCRT-II subunit Vps36 has been reported to bind PtdIns(3)P, and its mammalian homologue binds PtdIns(3,4,5)P₃ (Slagsvold *et al.*, 2005; Teo *et al.*, 2006). ESCRT-III is responsible for fission and may bind PtdIns(3,5)P₂ (Cullen and Carlton, 2012). Additionally SNX3 may represent an ESCRT independent means for forming a multivesicular body (Cullen and Carlton, 2012). In *C. elegans*, SAND-1/Mon1 has been proposed as the switch in early to late endosome transition. SAND-1/Mon1 is recruited to early endosomes, possibly by binding to PtdIns(3)P, and also complexes with Rabex5, a GEF for Rab5, to displace this GEF from early endosomes (Poteryaev *et al.*, 2010). SAND-1/Mon1 and Rab5 both interact with the HOPS complex, which possesses GEF activity for Rab7. Early endosomes thus lose Rab5 and acquire Rab7 in the maturation process of the degradative pathway (Rink *et al.*, 2005).

The normal maturation of endosomes in the degradative pathway requires the turnover of PtdIns(3)P. The PIKfyve

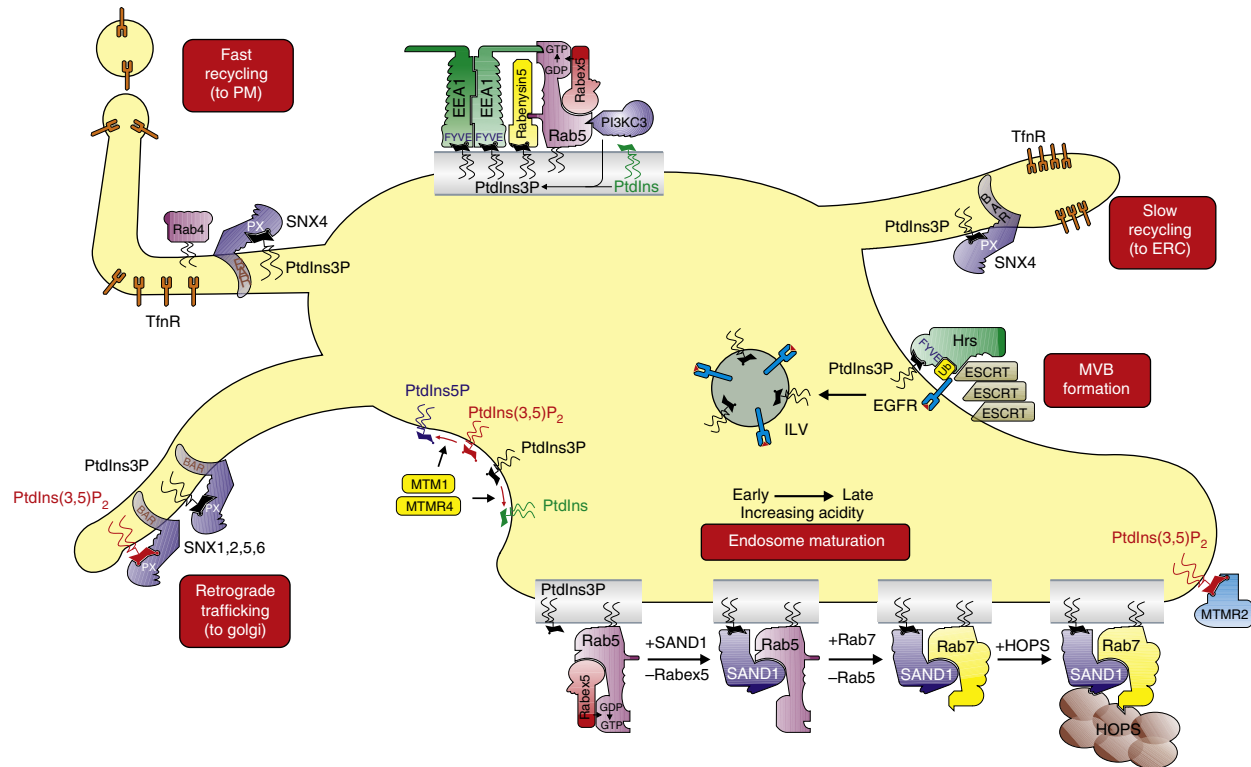


Figure 5 Endosomal phosphoinositide compartment maturation. The early endosome represents a way station for endocytosed vesicles, characterized by PtdIns(3)P. Rab5 recruits the Class I PI3-kinase, representing the source of the majority of early endosomal PtdIns(3)P. The PtdIns(3)P effector EEA1 is recruited and retained on early endosomes by coincidence detection of PtdIns(3)P interaction with its FYVE domain, and binding to Rab5. EEA1 and other FYVE domain containing effectors such as Phafin2 and rabenosyn 5 facilitate fusion between endosomes. Egress from the early endosome can occur via tubular branching for recycling and retrograde trafficking, mediated by the SNXs, which interact with PtdIns(3)P through their PX domains, in coincidence detection with BAR domain detection of membrane curvature. Receptors such as the transferrin receptor bound to transferrin can recycle directly to the plasma membrane, mediated by Rab4, or indirectly via a storage reservoir termed the ERC, mediated by Rab11. During retrograde trafficking, vesicles traffic from endosomes to the Golgi network, mediated by the retromer complex of SNXs which interact with PtdIns(3)P via PX domains. Turnover of PtdIns(3)P into PtdIns(3,5)P₂ is necessary for retrograde trafficking, and is impaired by pharmacologic inhibition of PIKfyve. Early endosomes can also mature into late endosomes for lysosomal fusion. Hrs is FYVE domain containing effector protein which is part of the ESCRT0 complex and directs the formation of internal vesicles. PtdIns(3)P also recruits SAND-1/Mon1 which permits the switch from Rab5 to Rab7 and the recruitment of the HOPS complex for the progression of endosomal maturation. The late endosome is positive for PtdIns(3,5)P₂. The myotubularins MTM1 and MTMR4 (contains a FYVE domain) have been identified at early endosomes and MTMR2 at late endosomes where they regulate PtdIns(3)P turnover.

complex can convert PtdIns(3)P to PtdIns(3,5)P₂. Additionally, myotubularins MTM1, MTMR2, and MTMR4 localize to early and late endosomes, and represent a means of degradation of both PtdIns(3)P and PtdIns(3,5)P₂, and mechanism for generating PtdIns(5)P (Davies *et al.*, 2012). The importance of a normal flux in endosomal PtdIns(3)P is highlighted by the observations that both increased as well as decreased or dominant negative PtdIns(3)P 3-phosphatase activity retards the degradation of EGFR (Lorenzo *et al.*, 2006; Pedersen *et al.*, 2012).

Phosphoinositides in Golgi Trafficking

The Golgi apparatus is a prominent network of flat membrane cisternae responsible for the modification and sorting of proteins and lipids. The *cis*-Golgi receives proteins and lipids from the ER, and after passage with modification through

the Golgi cisternae, vesicle carriers can exit from the TGN. Anterograde vesicular sorting occurs to the cell surface, and also to the endosomal network via clathrin-coated protein vesicles. Retrograde sorting back to the ER is via coated protein I (COPI) vesicles. The predominant signaling phosphoinositide at the Golgi is PtdIns(4)P, with small amounts of other phosphoinositides, including PtdIns(4,5)P₂ and PtdIns(3,4,5)P₃ also present (D'angelo *et al.*, 2012; Figure 6). PtdIns(4)P at the Golgi was initially thought to be a precursor for PtdIns(4,5)P₂ until subsequent studies demonstrated PtdIns(4)P to be essential for Golgi secretion in yeast. More recently, growing numbers of PtdIns(4)P-specific effectors have been shown to regulate Golgi sorting (D'angelo *et al.*, 2012). In yeast, one Class III PI4-kinase (PIK1) is responsible for Golgi PtdIns(4)P production. In mammals, there are four different phosphoinositide 4-kinases which localize to the Golgi, however the Class III (wortmannin-sensitive) PI4-kinase III β and the Class II (wortmannin-insensitive) PI4-kinase II α are the

predominant sources of PtdIns(4)P synthesis. Disruption of PtdIns(4)P signaling has dramatic effects on TGN function (D'angelo *et al.*, 2012).

PI4-kinase III β is present at the Golgi, endosomes, exocytic vesicles, lysosomes, and in the nucleus. Its most well-established role is to facilitate vesicle export to the plasma membrane (D'angelo *et al.*, 2012; Sridhar *et al.*, 2013). PI4-Kinase III β is recruited to the Golgi by the small GTPase Arf1, and is required for normal secretory function (Haynes *et al.*, 2005). The calcium sensor NCS-1 is also present at the TGN and recruits and activates PI4-Kinase III β , however, the complex of PI4-Kinase III β , Arf1 and NCS-1 is inhibitory to PI4-Kinase III β function (Haynes *et al.*, 2005). PI4-Kinase III β activity requires phosphorylation by the serine/threonine kinases PKD1 and 2 which are recruited to the Golgi by diacylglycerol, and increase transport from the TGN to the plasma membrane (Hausser *et al.*, 2005). PI4-Kinase III β kinase activity is stabilized by interaction of 14-3-3 protein dimers, which form scaffolds with CtBS/BARS and PI4-Kinase III β , essential for the budding of tubular domains from the TGN (Hausser *et al.*, 2005; Valente *et al.*, 2012). PtdIns(4)P and Arf1 can participate in coincidence detection of PH domain containing effectors. The PI4-Kinase III β pool of PtdIns(4)P mediates TGN to cell surface sorting of vesicles via FAPP1 and FAPP2, which are recruited to the TGN by interaction of their PH domains with PtdIns(4)P. The correct localization of FAPP1/FAPP2 requires Arf1 in addition to PtdIns(4)P, and inhibition of Arf1 activation prevents FAPP recruitment (Godi *et al.*, 2004). FAPP2 is also involved in glycosphingolipid biosynthesis, and initially associates with the *cis*-Golgi, where it binds glucosylceramide with subsequent increase in its affinity for PtdIns(4)P, resulting in relocation of FAPP2 to the TGN and delivery of glucosylceramide to this site (D'angelo *et al.*, 2013). During sphingomyelin synthesis, ceramide is transferred from the ER to the Golgi, and PtdIns(4)P is required at this early stage to recruit the PH domain containing CERT protein (Toth *et al.*, 2006). Oxysterol-binding protein 1 (OSBP) also regulates sphingomyelin synthesis, contains a PH domain, and binds PtdIns(4)P and Arf1 at the Golgi (Levine and Munro, 2002).

PI4-kinase II α is present on Golgi and endosomal membranes and regulates Golgi to endosome sorting (D'angelo *et al.*, 2012). Unlike PI4-Kinase III β , PI4-kinase II α recruitment to the Golgi is Arf1-independent. Cholesterol regulates PI4-kinase II α activity, possibly by altering its membrane mobility and therefore restricting PtdIns(4)P production to microdomains (Minogue *et al.*, 2010). Additionally, OSBP is involved in positive feedback to activate PI4-kinase II α , for recruitment of the PtdIns(4)P effector, CERT, to the TGN during sphingomyelin synthesis (Banerji *et al.*, 2010). Transfer from the TGN to endosomes/lysosomes occurs via clathrin-coated vesicles. PtdIns(4)P formed predominantly by PI4-kinase II α is required for the recruitment of the phosphoinositide effectors AP1, GGA, and Epsins to the TGN, a process that requires coincidence detection with Arf1 (Wang *et al.*, 2003, 2007).

PtdIns(4)P also regulates anterograde transport from the ER to the early Golgi, via COPII coated vesicles, where it acts at the SNARE-dependent fusion stage of vesicles with the Golgi (Lorente-Rodriguez and Barlowe, 2011). PtdIns(4)P

additionally regulates Golgi morphology, through interaction with positively charged domains on the PtdIns(4)P effectors Golgi phosphoprotein 3 (GOLPH3) and Golgi phosphoprotein 3-like (GOLPH3L) (Dippold *et al.*, 2009; Ng *et al.*, 2013). These proteins exert opposing effects on the shape of the Golgi (Ng *et al.*, 2013). GOLPH3 links PtdIns(4)P to the actin cytoskeleton via the myosin Myo18, maintaining a compressed Golgi ribbon structure and contributing to vesicle budding (Dippold *et al.*, 2009). The yeast homologue of GOLPH3 is Vps74p, and functions in retrograde traffic of COPI vesicles from the medial Golgi. Additionally, Vps74p contributes to PtdIns(4)P turnover. Vps74p binds the phosphoinositide phosphatase, SAC1, which degrades PtdIns(4)P to form PtdIns and terminate PtdIns(4)P signaling at the early Golgi apparatus to maintain cisternal identity (Cai *et al.*, 2014). In yeast, turnover of PtdIns(4)P has also been demonstrated in the anterograde pathway. Recruitment of the GEF Sec2p occurs via coincidence detection between Ypt32p (mammalian homologue Rab11b) and PtdIns(4)P, resulting in the activation of the Rab GTPase Sec4p (mammalian homologue Rab8). However, as vesicles reach secretory sites, PtdIns(4)P signals decline, and this switch dissociates Sec2p from Ypt32p, allowing Sec2p to interact with Sec15p, a component of the exocyst required for vesicle tethering for exocytosis (Mizuno-Yamasaki *et al.*, 2012).

Retrograde Trafficking from Endosomes to Golgi

Retrograde transport is important for the retrieval of proteins and lipids back to the Golgi apparatus. Retrograde trafficking of cargo such as the mannose-6-phosphate receptor, which delivers hydrolases to endosomes before being recycled, occurs via the multimeric complex retromer (Bonifacino and Hurley, 2008). Additionally, there are retromer independent retrograde pathways, such as the transport of the plant toxin ricin (Cullen and Carlton, 2012). The yeast multimeric retromer complex consists of the trimer Vps26p–Vps29p–Vps35p and also Vps5p and Vps17p (Cullen and Carlton, 2012). In the mammalian complex there are two Vps5p homologues, SNX1 and SNX2 and two Vps17p homologues, SNX5 and SNX6 (Cullen and Carlton, 2012). Retrograde transport is achieved through tubule based sorting, whereby cargo is enriched in forming tubules which can then undergo fission for trafficking (Van Weering *et al.*, 2012). A number of SNX proteins undergo coincidence detection via their phosphoinositide binding PX domain, and a membrane curvature-sensing BAR domain, targeted to the highly curved tubulovesicular endosomes where they are involved tubule formation (Van Weering *et al.*, 2012). The phosphoinositides involved in retrograde trafficking include PtdIns(3)P and PtdIns(3,5)P₂, however, this remains an emerging field. In yeast, Vps34, the Class III PI3-kinase has a direct role in the regulation of retromer function (Burda *et al.*, 2002). SNX1 and SNX2 localize to early endosomes, the site of PtdIns(3)P enrichment, and bind PtdIns(3)P and PtdIns(3,5)P₂ (Carlton *et al.*, 2005; Cozier *et al.*, 2002). SNX3 a non-BAR domain containing SNX binds PtdIns(3)P and interacts with retromer in an alternative pathway for the recycling of Wntless (Harterink *et al.*, 2011). The PIKfyve 5-kinase catalyzes the conversion of PtdIns(3)P to PtdIns(3,5)P₂,

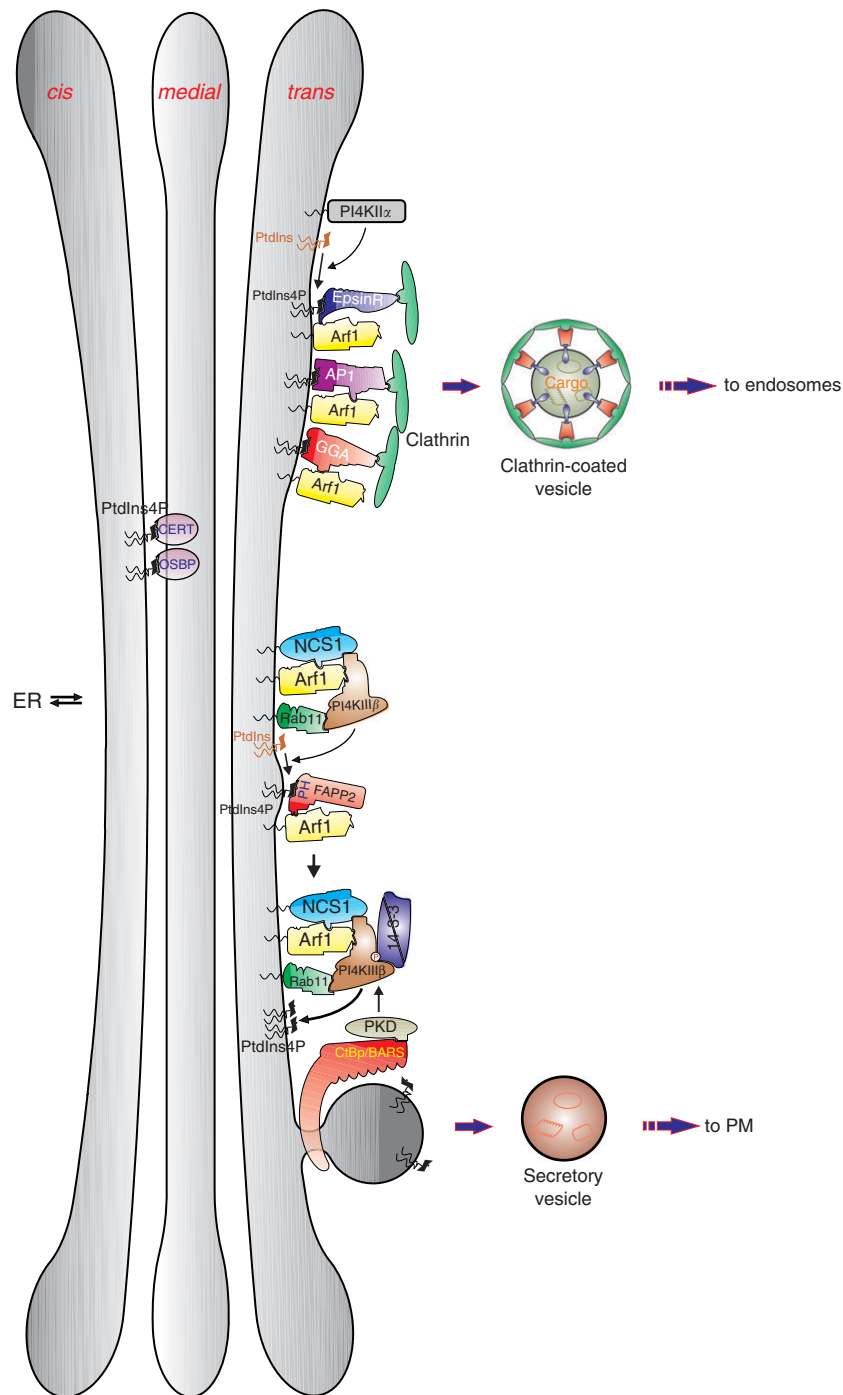


Figure 6 Golgi trafficking regulation by phosphoinositides. PtdIns(4)P is important in Golgi membrane trafficking. PI4KII α mediates the formation of PtdIns(4)P at the TGN, best described in the formation of coated vesicles for Golgi to endosome trafficking. Coincidence detection occurs between PtdIns(4)P and Arf1 for the recruitment of effector proteins such as AP1, GGAs, and EpsinR for the formation of clathrin-coated endosomes. PI4KIII β is also recruited to the TGN mediated via Arf1 in combination with NCS and Rab11, best described for the formation of vesicles destined for exocytosis at the plasma membrane. PtdIns(4)P and Arf1 recruit the PH domain containing effector FAPP2, involved in the initiation of vesicle budding. Positive feedback of PI4KIII β to make more PtdIns(4)P, stabilized by a scaffold of PKD, 14-3-3 proteins, Rab11, and CtBp/BARS contributes to the fission of vesicles from the TGN.

and is necessary for normal retromer function, suggesting that turnover of PtdIns(3)P into PtdIns(3,5)P₂ is an important requirement for retrograde trafficking (Rutherford *et al.*, 2006). Consistent with this, the myotubularin 3-phosphatases,

MTM-6 and MTM-9, which degrade PtdIns(3)P and PtdIns(3,5)P₂ are required for Wnt signaling via retromer in *C. elegans* (Silhankova *et al.*, 2010). The tubulation function of SNX1 is timed to Rab conversion in endosome maturation,

and is most prominently observed on endosomes transitioning from Rab5 (early)-positive to Rab7 (late)-positive (Van Weering *et al.*, 2012). Additionally, coincidence detection occurs between retromer, SNX3 and Rab7 on endosomes (Lorenowicz *et al.*, 2014).

Phosphoinositides in the Regulation of Autophagy

Macroautophagy (hereafter called autophagy) is a conserved membrane-regulated degradative mechanism which functions in cell homeostasis and under conditions of stress, such as starvation. Autophagy is implicated in disease processes such as neurodegeneration, cancer, and heart disease. During autophagy, cytosolic contents are sequestered within a membrane-bound compartment, and delivered to the lysosome (Figure 7). Autophagy permits the regeneration of damaged organelles, such as mitochondria and the ER, the regulation of lipid metabolism, the clearance of protein aggregates from the cell and the destruction of intracellular microorganisms (Choi *et al.*, 2013; Deretic, 2006). In yeast, 37 autophagy-related proteins (Atg proteins) have been described (Lamb *et al.*, 2013). The direction of autophagic signals to the membrane source is termed initiation, and autophagy can be stimulated by amino acid deprivation. Nucleation is the creation of the pre-autophagosomal structure, which subsequently leads to the formation of an isolation membrane which sequesters cytosolic content and fuses to surround that content within a double membrane autophagosome. The autophagosome then undergoes a maturation process to fuse with the vacuole, the yeast equivalent of the mammalian lysosomes (Lamb *et al.*, 2013; Mizushima, 2007).

Phosphoinositides are able to influence autophagy at different steps. Growth factor or insulin stimulated production of PtdIns(3,4,5)P₃ at the plasma membrane catalyzed by the Class I PI3-kinase negatively regulates autophagy (Devereaux *et al.*, 2013). Plasma membrane PtdIns(3,4,5)P₃ recruits and activates Akt, which contributes to a complex signaling cascade which results in mTORC1 activation (Devereaux *et al.*, 2013). mTORC1 inhibits autophagy at the initiation stage by forming a complex with ULK1/ULK2 (mammalian homolog of Atg1), Atg13, and FIP200 (postulated homolog of Atg17) under nutrient rich conditions. mTORC1 is excluded from the complex upon starvation (Hosokawa *et al.*, 2009). When not associated with mTORC1, ULK1/2 and Atg13 are dephosphorylated and in complex with Atg17 translocate to membranes to initiate autophagy (Hosokawa *et al.*, 2009). Intriguingly, the class I PI3-kinase subunit p110 β is a positive regulator of autophagy, independent of its PI3-kinase activity/mTORC1 influence, and forms a complex with the Class III PI3-kinase (Vps34)–Beclin 1–Vps15–Atg14L, involved in the initiation of autophagy (see below) (Dou *et al.*, 2010).

PtdIns(3)P is a critical signal to stimulate autophagy. The main source of PtdIns(3)P in autophagy is the Class III PI3-kinase (Vps34), with contribution from Class II PI3-kinase (Devereaux *et al.*, 2013). Initiation and nucleation require both the translocation of the ULK1 complex (ULK1–Atg13–FIP200–Atg101) to a membrane source, as well as the creation of PtdIns(3)P by the complex of Vps34–Beclin1 (homolog of Atg6)–Vps15–Atg14L (Lamb *et al.*, 2013). PtdIns(3)P

facilitates the recruitment of the ULK1 complex to autophagic puncta, and the ULK1 complex itself feeds back to increase the formation of PtdIns(3)P by activating the PI3-kinase complex (Karanasios *et al.*, 2013; Russell *et al.*, 2013). The early autophagosomal membrane in some settings is termed an omegasome, connected to the ER, marked by PtdIns(3)P and the FYVE domain containing PtdIns(3)P effector DFCEP1. However, the exact function of DFCEP1 in autophagy remains unclear (Lamb *et al.*, 2013). The WD-repeat domain containing phosphoinositide interacting proteins (WIPI) 1/2 (yeast homologue Atg18) and 3/4 proteins are additional PtdIns(3)P effectors which contain WD-repeat PROPPIN domains important in autophagosome formation. WIPI1 may have an inhibitory role in autophagy (Polson *et al.*, 2010). WIPI2b contributes to the maturation of the autophagosome, by directly binding Atg16L1 on omegasomes, which is responsible for the recruitment of the Atg12–Atg5–Atg16L complex to autophagosomes (Dooley *et al.*, 2014). This complex has E3 ubiquitin ligase-like function that promotes the lipidation of the autophagy Atg8 homologues microtubule associated protein light chain 3 (LC3), GABARAP, and GATE16 (Lamb *et al.*, 2013). The *C. elegans* protein EPG-6 is the homologue of WIPI4, and is important in directing autophagosomal membrane size, possibly via forming a complex with Atg2 and regulating the membrane delivery function of the transmembrane protein Atg9 (Lu *et al.*, 2011). Another PtdIns(3)P effector which coordinates membrane delivery to the autophagosome via Atg9 is UV radiation resistance-associated gene protein (UVRAG), which binds a pool of PtdIns(3)P at ER membrane regions (He *et al.*, 2013). Under basal conditions, UVRAG coordinates *cis*-Golgi to ER retrograde trafficking, but upon stimulation of autophagy UVRAG is part of an alternative PI3-kinase complex which activates Beclin1 and Vps34 (He *et al.*, 2013). Additionally, the PH domain containing effector tectonin β -propeller repeat containing 1 TECPR1 contributes to the recruitment of Atg5, and one report suggests it binds PtdIns(3)P and is required for autophagosome–lysosome fusion (Chen *et al.*, 2012).

Cytosolic LC3-I (Atg8 homologue) is recruited to autophagosomes and undergoes lipidation to become LC3-II via the Atg12–Atg5–Atg16L complex. During autophagosome maturation, most of the other Atg proteins dissociate from the membrane, but lipidated LC3-II remains bound and traverses to the autophagolysosome. LC3 is the proposed receptor for p62, which has a ubiquitin-binding domain and may mediate recruitment of ubiquitin containing proteins to the autophagosome membrane (Mizushima, 2007). FYVE and coiled-coil domain containing 1 (FYCO1) is a recently described FYVE domain containing protein which regulates autophagosome motility away from the perinuclear region by directing microtubule plus end motility through interactions with Rab7 and LC3 (Pankiv *et al.*, 2010). The autophagosome matures by fusion with endosomes and lysosomes (Mizushima, 2007). Early evidence suggests that PtdIns(3,5)P₂ may also regulate autophagy, although the site of PtdIns(3,5)P₂ action is unclear (De Lartigue *et al.*, 2009). Inhibition of the 5-kinase PIKfyve, responsible for turnover of PtdIns(3)P into PtdIns(3,5)P₂, dysregulates autophagy, possibly by preventing autophagosome consumption by lysosomes (De Lartigue *et al.*, 2009; Martin *et al.*, 2013). Neurons of mice deficient in Fig 4 or

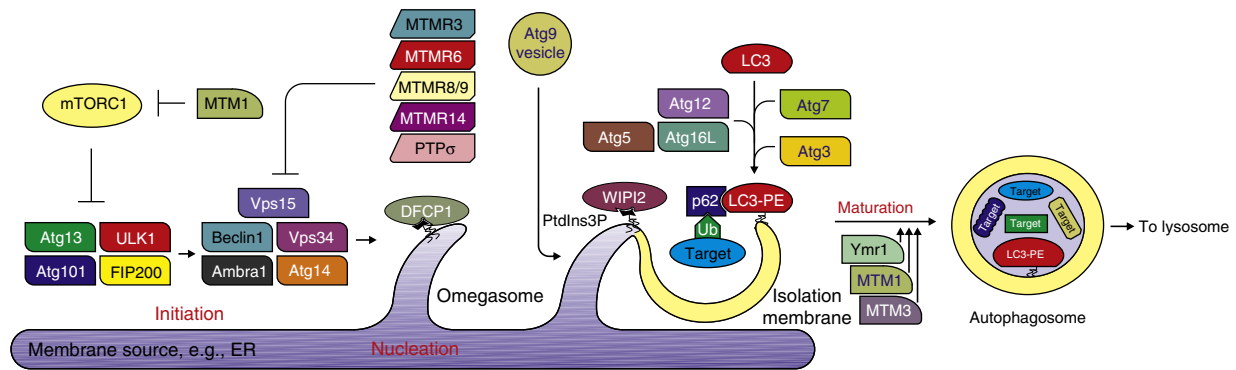


Figure 7 Role of phosphoinositides in autophagy stages. PtdIns(3)P is important in autophagic trafficking. Autophagy is initiated by the recruitment of two protein complexes to a membrane source, the ULK1 complex (ULK1-Atg13-Fip200-Atg101) and the PI3K complex (Vps34-Beclin1-Atg14-Vps15). Dissociation of mTORC1 from the ULK1 complex is necessary for autophagy initiation which is PtdIns(3)P-dependent. Recruitment of the PI3K complex creates a positive feedback loop of PtdIns(3)P production and further ULK1 complex recruitment, and nucleation occurs with the interaction of the FYVE domain containing effector protein DFCP1 with PtdIns(3)P on omegasomes. This serves as a platform for the development of the isolation membrane, which is mediated by WIPI2, a WD40 domain containing PtdIns(3)P effector. The myotubularins MTMR3, MTMR6, MTMR8/MTMR9 complex and MTMR14 act to decrease PtdIns(3)P and inhibit autophagy initiation. The transmembrane protein Atg9 recruits vesicles for the formation of the isolation membrane, and the Atg16L-Atg12-Atg5-Atg3 complex lipidates LC3 to LC3-PE for incorporation into the isolation membrane. LC3 and p62 mediate interactions with the species targeted for autophagy. The Atg proteins are lost upon closure of the autophagosome, which matures by fusion with lysosomes. The myotubularins Ymr1, MTM1 and MTM-3 are required for normal autophagy maturation. ER, endoplasmic reticulum.

Vac14, enzymes which are required for normal PIKfyve function exhibit abnormal autophagy (Ferguson *et al.*, 2009).

The turnover of PtdIns(3)P and possibly PtdIns(3,5)P₂ may be accomplished by the action of phosphoinositide 3-phosphatases. There have been reports that decreased 3-phosphatase function results in increased induction of autophagy, as well as blocking of the maturation of autophagosomes. This may be a result of individual 3-phosphatases acting to degrade PtdIns(3)P at unique points in time and space. The myotubularins MTMR3 and MTMR8 in complex with MTMR9, MTMR6, and MTMR14 as well as the protein phosphatase PTP σ all negatively regulate autophagy induction (Martin *et al.*, 2011; Mochizuki *et al.*, 2013; Taguchi-Atarashi *et al.*, 2010; Zou *et al.*, 2012). Two myotubularins ymr1 and MTM-3, in yeast and *C. elegans*, respectively, have been shown to act at later stages and are required for turning over PtdIns(3)P in order for autophagosomes to mature normally (Cebollero *et al.*, 2012; Wu *et al.*, 2014). Additionally, the muscle defects arising from loss of MTM1 function are thought to result from aberrant autophagy, produced as a consequence of a late-stage block in autophagy, as well as an increase in mTORC1 signaling, which negatively regulates autophagy initiation (Al-Qusairi *et al.*, 2013; Fetalvero *et al.*, 2013).

Recently, a role for PtdIns(4)P and PtdIns(4,5)P₂ in normal lysosomal function has been shown, particularly for lysosomal regeneration in response to starvation in the process of autophagic lysosome reformation (Kistakis and Tooze, 2013; Rong *et al.*, 2012; Sridhar *et al.*, 2013). PtdIns(4)P, formed by Class III PI4K β is important in restricting the lysosomal cargo from being sorted out of the lysosome during starvation when recycling of lysosomal membrane is required (Sridhar *et al.*, 2013). PtdIns(4,5)P₂ is needed for normal tubule formation for lysosomal egress, and fission of these tubules away from the lysosome, via the action of PI5K1B and PI5K1A respectively (Rong *et al.*, 2012).

Concluding Remarks

Refinements in biochemical and microscopic techniques continue to provide new opportunities to dissect the role of phosphoinositides in determining organelle identity and function. Many details of the contribution of phosphoinositide kinases and phosphatases in regulating the localization and abundance of individual phosphoinositide species at the membrane-cytosol interface remain to be characterized. Increasingly it is being understood that the orchestrated turnover of one phosphoinositide species into another is required not only to initiate and terminate individual organelle processes, but the transition of phosphoinositides ensures orderly flow of traffic between networked organelle compartments. The number of effector proteins which interact with phosphoinositides on membrane organelles continues to grow, however, much remains to be discovered about their regulation and function. Coordination of phosphoinositides and their effector proteins in regulating membrane organelle trafficking is essential for normal cellular function, and when aberrant leads to many human diseases, including infection, cancer, and degenerative diseases.

See also: Cell Division/Death: Apoptosis: Autophagy. Interorganellar Communication: Components: Adaptor Proteins: Inter-Organellar Traffic Controllers; SNAREs: Membrane Fusion and Beyond. Interorganellar Communication: Interplay and Processes: Clathrin and Clathrin-Dependent Endocytosis

References

- Abe, N., Inoue, T., Galvez, T., Klein, L., Meyer, T., 2008. Dissecting the role of PtdIns(4,5)P₂ in endocytosis and recycling of the transferrin receptor. *Journal of Cell Science* 121, 1488–1494.

- Al-Qusairi, L., Prokic, I., Amosii, L., *et al.*, 2013. Lack of myotubularin (MTM1) leads to muscle hypotrophy through unbalanced regulation of the autophagy and ubiquitin-proteasome pathways. *FASEB Journal* 27, 3384–3394.
- Alajlouni, R., Drahos, K.E., Finkielstein, C.V., Capelluto, D.G., 2011. Lipid-mediated membrane binding properties of Disabled-2. *Biochimica et Biophysica Acta* 1808, 2734–2744.
- Antonescu, C.N., Aguet, F., Danuser, G., Schmid, S.L., 2011. Phosphatidylinositol-(4,5)-bisphosphate regulates clathrin-coated pit initiation, stabilization, and size. *Molecular Biology of the Cell* 22, 2588–2600.
- Balla, T., 2013. Phosphoinositides: Tiny lipids with giant impact on cell regulation. *Physiological Reviews* 93, 1019–1137.
- Banerji, S., Ngo, M., Lane, C.F., *et al.*, 2010. Oxysterol binding protein-dependent activation of sphingomyelin synthesis in the golgi apparatus requires phosphatidylinositol 4-kinase IIalpha. *Molecular Biology of the Cell* 21, 4141–4150.
- Bonifacino, J.S., Hurley, J.H., 2008. Retromer. *Current Opinion in Cell Biology* 20, 427–436.
- Burda, P., Padilla, S.M., Sarkar, S., Emr, S.D., 2002. Retromer function in endosome-to-Golgi retrograde transport is regulated by the yeast Vps34 PtdIns 3-kinase. *Journal of Cell Science* 115, 3889–3900.
- Cai, Y., Deng, Y., Horenkamp, F., Reinisch, K.M., Burd, C.G., 2014. Sac1-Vps74 structure reveals a mechanism to terminate phosphoinositide signaling in the Golgi apparatus. *Journal of Cell Biology* 206, 485–491.
- Carlton, J.G., Bujny, M.V., Peter, B.J., *et al.*, 2005. Sorting nexin-2 is associated with tubular elements of the early endosome, but is not essential for retromer-mediated endosome-to-TGN transport. *Journal of Cell Science* 118, 4527–4539.
- Cebollero, E., Van Der Vaart, A., Zhao, M., *et al.*, 2012. Phosphatidylinositol-3-phosphate clearance plays a key role in autophagosome completion. *Current Biology* 22, 1545–1553.
- Chen, D., Fan, W., Lu, Y., *et al.*, 2012. A mammalian autophagosome maturation mechanism mediated by TECPR1 and the Atg12-Atg5 conjugate. *Molecular Cell* 45, 629–641.
- Choi, A.M., Ryter, S.W., Levine, B., 2013. Autophagy in human health and disease. *New England Journal of Medicine* 368, 1845–1846.
- Choudhury, P., Srivastava, S., Li, Z., *et al.*, 2006. Specificity of the myotubularin family of phosphatidylinositol-3-phosphatase is determined by the PH/GRAM domain. *Journal of Biological Chemistry* 281, 31762–31769.
- Compagnon, J., Gervais, L., Roman, M.S., Chamot-Boeuf, S., Guichet, A., 2009. Interplay between Rab5 and PtdIns(4,5)P2 controls early endocytosis in the *Drosophila* germline. *Journal of Cell Science* 122, 25–35.
- Corbalan-Garcia, S., Gómez-Fernández, J.C., 2014. Signaling through C2 domains: More than one lipid target. *Biochimica et Biophysica Acta* 1838, 1536–1547.
- Cozier, G.E., Carlton, J., McGregor, A.H., *et al.*, 2002. The phox homology (PX) domain-dependent, 3-phosphoinositide-mediated association of sorting nexin-1 with an early sorting endosomal compartment is required for its ability to regulate epidermal growth factor receptor degradation. *Journal of Biological Chemistry* 277, 48730–48736.
- Cullen, P.J., Carlton, J.G., 2012. Phosphoinositides in the mammalian endo-lysosomal network. *Subcellular Biochemistry* 59, 65–110.
- D'angelo, G., Uemura, T., Chuang, C.C., *et al.*, 2013. Vesicular and non-vesicular transport feed distinct glycosylation pathways in the Golgi. *Nature* 501, 116–120.
- D'angelo, G., Vicinanza, M., Wilson, C., De Matteis, M.A., 2012. Phosphoinositides in Golgi complex function. *Subcellular Biochemistry* 59, 255–270.
- Davies, E.M., Sheffield, D.A., Tibarewal, P., *et al.*, 2012. The PTEN and myotubularin phosphoinositide 3-phosphatases: Linking lipid signaling to human disease. *Subcellular Biochemistry* 58, 281–336.
- De Lartigue, J., Polson, H., Feldman, M., *et al.*, 2009. PIKfyve regulation of endosome-linked pathways. *Traffic* 10, 883–893.
- Deretic, V., 2006. Autophagy as an immune defense mechanism. *Current Opinion in Immunology* 18, 375–382.
- Devereaux, K., Dall'armi, C., Alcazar-Roman, A., *et al.*, 2013. Regulation of mammalian autophagy by class II and III PI 3-kinases through PI3P synthesis. *PLoS One* 8, e76405.
- Di Paolo, G., De Camilli, P., 2006. Phosphoinositides in cell regulation and membrane dynamics. *Nature* 443, 651.
- Dippold, H.C., Ng, M.M., Farber-Katz, S.E., *et al.*, 2009. GOLPH3 bridges phosphatidylinositol-4-phosphate and actomyosin to stretch and shape the Golgi to promote budding. *Cell* 139, 337–351.
- Dominguez, V., Raimondi, C., Somanath, S., *et al.*, 2011. Class II phosphoinositide 3-kinase regulates exocytosis of insulin granules in pancreatic beta cells. *Journal of Biological Chemistry* 286, 4216–4225.
- Dooley, H.C., Razi, M., Polson, H.E., *et al.*, 2014. WIPI2 links LC3 conjugation with PI3P, autophagosome formation, and pathogen clearance by recruiting Atg12-5-16L1. *Molecular Cell* 55, 238–252.
- Dou, Z., Chattopadhyay, M., Pan, J.A., *et al.*, 2010. The class IA phosphatidylinositol 3-kinase p110-beta subunit is a positive regulator of autophagy. *Journal of Cell Biology* 191, 827–843.
- Dove, S.K., Dong, K., Kobayashi, T., Williams, F.K., Michell, R.H., 2009. Phosphatidylinositol 3,5-bisphosphate and Fab1p/PIKfyve underPPIn endo-lysosome function. *Biochemical Journal* 419, 1–13.
- Falasca, M., Maffucci, T., 2007. Role of class II phosphoinositide 3-kinase in cell signaling. *Biochemical Society Transactions* 35, 211–214.
- Ferguson, C.J., Lenk, G.M., Meisler, M.H., 2009. Defective autophagy in neurons and astrocytes from mice deficient in PI(3,5)P2. *Human Molecular Genetics* 18, 4868–4878.
- Fetalvero, K.M., Yu, Y., Goetschkes, M., *et al.*, 2013. Defective autophagy and mTORC1 signaling in myotubularin null mice. *Molecular and Cellular Biology* 33, 98–110.
- Franco, I., Gulluni, F., Campa, C.C., *et al.*, 2014. PI3K class II alpha controls spatially restricted endosomal PtdIns3P and Rab11 activation to promote primary cilium function. *Developmental Cell* 28, 647–658.
- Fujita, A., Cheng, J., Tauchi-Sato, K., Takenawa, T., Fujimoto, T., 2009. A distinct pool of phosphatidylinositol 4,5-bisphosphate in caveolae revealed by a nanoscale labeling technique. *Proceedings of the National Academy of Sciences of the United States of America* 106, 9256–9261.
- Gaidarov, I., Keen, J.H., 1999. Phosphoinositide-AP-2 interactions required for targeting to plasma membrane clathrin-coated pits. *Journal of Cell Biology* 146, 755–764.
- Godi, A., Di Campli, A., Konstantakopoulos, A., *et al.*, 2004. FAPPs control Golgi-to-cell-surface membrane traffic by binding to ARF and PtdIns(4)P. *Nature Cell Biology* 6, 393–404.
- Godlee, C., Kaksonen, M., 2013. Review series: From uncertain beginnings: Initiation mechanisms of clathrin-mediated endocytosis. *Journal of Cell Biology* 203, 717–725.
- Harterink, M., Port, F., Lorenowicz, M.J., *et al.*, 2011. A SNX3-dependent retromer pathway mediates retrograde transport of the Wnt sorting receptor Wntless and is required for Wnt secretion. *Nature Cell Biology* 13, 914–923.
- Hausser, A., Storz, P., Martens, S., *et al.*, 2005. Protein kinase D regulates vesicular transport by phosphorylating and activating phosphatidylinositol-4 kinase IIb at the Golgi complex. *Nature Cell Biology* 7, 880–886.
- Hayashi, M., Raimondi, A., O'toole, E., *et al.*, 2008. Cell- and stimulus-dependent heterogeneity of synaptic vesicle endocytic recycling mechanisms revealed by studies of dynamin 1-null neurons. *Proceedings of the National Academy of Sciences of the United States of America* 105, 2175–2180.
- Haynes, L.P., Thomas, G.M., Burgoyne, R.D., 2005. Interaction of neuronal calcium sensor-1 and ADP-ribosylation factor 1 allows bidirectional control of phosphatidylinositol 4-kinase beta and trans-Golgi network-plasma membrane traffic. *Journal of Biological Chemistry* 280, 6047–6054.
- He, S., Ni, D., Ma, B., *et al.*, 2013. PtdIns(3)P-bound UVRAG coordinates Golgi-ER retrograde and Atg9 transport by differential interactions with the ER tether and the beclin 1 complex. *Nature Cell Biology* 15, 1206–1219.
- Hom, R.A., Vora, M., Regner, M., *et al.*, 2007. pH-dependent binding of the Epsin ANTH domain and the AP180 ANTH domain to PI(4,5)P2-containing bilayers. *Journal of Molecular Biology* 373, 412–423.
- Horan, K.A., Watanabe, K., Kong, A.M., *et al.*, 2007. Regulation of Fc gammaR-stimulated phagocytosis by the 72-kDa inositol polyphosphate 5-phosphatase: SHIP1, but not the 72-kDa 5-phosphatase, regulates complement receptor 3 mediated phagocytosis by differential recruitment of these 5-phosphatases to the phagocytic cup. *Blood* 110, 4480–4491.
- Hosokawa, N., Hara, T., Kaizuka, T., *et al.*, 2009. Nutrient-dependent mTORC1 association with the ULK1-Atg13-FIP200 complex required for autophagy. *Molecular Biology of the Cell* 20, 1981–1991.
- James, D.J., Khodthong, C., Kowalchuk, J.A., Martin, T.F., 2008. Phosphatidylinositol 4,5-bisphosphate regulates SNARE-dependent membrane fusion. *Journal of Cell Biology* 182, 355–366.
- Jovic, M., Kieken, F., Naslavsky, N., *et al.*, 2009. Eps15 homology domain 1-associated tubules contain phosphatidylinositol-4-phosphate and phosphatidylinositol-(4,5)-bisphosphate and are required for efficient recycling. *Molecular Biology of the Cell* 20, 2731–2743.
- Kabachinski, G., Yamaga, M., Kieler-Grevstad, D.M., Bruinsma, S., Martin, T.F., 2014. CAPS and Munc13 utilize distinct PIP2-linked mechanisms to promote vesicle exocytosis. *Molecular Biology of the Cell* 25, 508–521.
- Karanasios, E., Stapleton, E., Manifava, M., *et al.*, 2013. Dynamic association of the ULK1 complex with omegasomes during autophagy induction. *Journal of Cell Science* 126, 5224–5238.
- Ktistakis, N.T., Tooze, S.A., 2013. PIPing on lysosome tubes. *EMBO Journal* 32, 315–317.

- Kutateladze, T.G., 2006. Phosphatidylinositol 3-phosphate recognition and membrane docking by the FYVE domain. *Biochimica et Biophysica Acta (BBA) – Molecular and Cell Biology of Lipids* 1761, 868.
- Kutateladze, T.G., 2007. Mechanistic similarities in docking of the FYVE and PX domains to phosphatidylinositol 3-phosphate containing membranes. *Progress in Lipid Research* 46, 315–327.
- Lamb, C.A., Yoshimori, T., Tooze, S.A., 2013. The autophagosome: Origins unknown, biogenesis complex. *Nature Reviews Molecular Cell Biology* 14, 759–774.
- Lemmon, M.A., 2007. Pleckstrin homology (PH) domains and phosphoinositides. *Biochemical Society Symposium* 74, 81–93.
- Levine, T.P., Munro, S., 2002. Targeting of Golgi-specific pleckstrin homology domains involves both PtdIns 4-kinase-dependent and -independent components. *Current Biology* 12, 695–704.
- Li, X., Wang, X., Zhang, X., *et al.*, 2013. Genetically encoded fluorescent probe to visualize intracellular phosphatidylinositol 3,5-bisphosphate localization and dynamics. *Proceedings of the National Academy of Sciences of the United States of America* 110, 21165–21170.
- Lim, J.P., Teasdale, R.D., Gleeson, P.A., 2012. SNX5 is essential for efficient macropinocytosis and antigen processing in primary macrophages. *Biology Open* 1, 904–914.
- Lorenowicz, M.J., Macurkova, M., Harterink, M., *et al.*, 2014. Inhibition of late endosomal maturation restores Wnt secretion in *Caenorhabditis elegans* vps-29 retromer mutants. *Cell Signal* 26, 19–31.
- Lorente-Rodriguez, A., Barlowe, C., 2011. Requirement for Golgi-localized PI(4)P in fusion of COPII vesicles with Golgi compartments. *Molecular Biology of the Cell* 22, 216–229.
- Lorenzo, O., Urbe, S., Clague, M.J., 2006. Systematic analysis of myotubularins: Heteromeric interactions, subcellular localisation and endosomerelevant functions. *Journal of Cell Science* 119, 2953–2959.
- Lu, Q., Yang, P., Huang, X., *et al.*, 2011. The WD40 repeat PtdIns(3)P-binding protein EPG-6 regulates progression of omegasomes to autophagosomes. *Developmental Cell* 21, 343–357.
- Maekawa, M., Terasaka, S., Mochizuki, Y., *et al.*, 2014. Sequential breakdown of 3-phosphorylated phosphoinositides is essential for the completion of macropinocytosis. *Proceedings of the National Academy of Sciences of the United States of America* 111, E978–E987.
- Mani, M., Lee, S.Y., Lucast, L., *et al.*, 2007. The dual phosphatase activity of synaptotagmin1 is required for both efficient synaptic vesicle endocytosis and reavailability at nerve terminals. *Neuron* 56, 1004–1018.
- Maniti, O., Khalifat, N., Goggia, K., *et al.*, 2012. Binding of moesin and ezrin to membranes containing phosphatidylinositol (4,5) bisphosphate: A comparative study of the affinity constants and conformational changes. *Biochimica et Biophysica Acta* 1818, 2839–2849.
- Martin, K.R., Xu, Y., Looyenga, B.D., *et al.*, 2011. Identification of PTPsigma as an autophagic phosphatase. *Journal of Cell Science* 124, 812–819.
- Martin, S., Harper, C.B., May, L.M., *et al.*, 2013. Inhibition of PIKfyve by YM-201636 dysregulates autophagy and leads to apoptosis-independent neuronal cell death. *PLoS One* 8, e60152.
- Martin, T.F., 2012. Role of PI(4,5)P₂ in vesicle exocytosis and membrane fusion. *Subcellular Biochemistry* 59, 111–130.
- McLauchlan, H., Newell, J., Morrice, N., *et al.*, 1998. A novel role for Rab5-GDI in ligand sequestration into clathrin-coated pits. *Current Biology* 8, 34.
- Meunier, F.A., Osborne, S.L., Hammond, G.R., *et al.*, 2005. Phosphatidylinositol 3-kinase C2alpha is essential for ATP-dependent priming of neurosecretory granule exocytosis. *Molecular Biology of the Cell* 16, 4841–4851.
- Milosevic, I., Giovedi, S., Lou, X., *et al.*, 2011. Recruitment of endophilin to clathrin-coated pit necks is required for efficient vesicle uncoating after fission. *Neuron* 72, 587–601.
- Minogue, S., Chu, K.M., Westover, E.J., *et al.*, 2010. Relationship between phosphatidylinositol 4-phosphate synthesis, membrane organization, and lateral diffusion of PI4KIIalpha at the trans-Golgi network. *Journal of Lipid Research* 51, 2314–2324.
- Mizuno-Yamasaki, E., Rivera-Molina, F., Novick, P., 2012. GTPase networks in membrane traffic. *Annual Review of Biochemistry* 81, 637–659.
- Mizushima, N., 2007. Autophagy: Process and function. *Genes & Development* 21, 2861–2873.
- Mochizuki, Y., Ohashi, R., Kawamura, T., *et al.*, 2013. Phosphatidylinositol 3-phosphatase myotubularin-related protein 6 (MTMR6) is regulated by small GTPase Rab1B in the early secretory and autophagic pathways. *Journal of Biological Chemistry* 288, 1009–1021.
- Murray, J.T., Panaretou, C., Stenmark, H., Miaczynska, M., Backer, J.M., 2002. Role of Rab5 in the recruitment of hVps34/p150 to the early endosome. *Traffic* 3, 416–427.
- Naga Prasad, S.V., Laporte, S.A., Chamberlain, D., *et al.*, 2002. Phosphoinositide 3-kinase regulates beta2-adrenergic receptor endocytosis by AP-2 recruitment to the receptor/beta-arrestin complex. *Journal of Cell Biology* 158, 563–575.
- Nakatsu, F., Perera, R.M., Lucast, L., *et al.*, 2010. The inositol 5-phosphatase SHIP2 regulates endocytic clathrin-coated pit dynamics. *Journal of Cell Biology* 190, 307–315.
- Ng, M.M., Dippold, H.C., Buschman, M.D., Noakes, C.J., Field, S.J., 2013. GOLPH3L antagonizes GOLPH3 to determine Golgi morphology. *Molecular Biology of the Cell* 24, 796–808.
- Osborne, S.L., Wen, P.J., Boucheron, C., *et al.*, 2008. PIKfyve negatively regulates exocytosis in neurosecretory cells. *Journal of Biological Chemistry* 283, 2804–2813.
- Padrick, S.B., Rosen, M.K., 2010. Physical mechanisms of signal integration by WASP family proteins. *Annual Review of Biochemistry* 79, 707–735.
- Pankiv, S., Alemu, E.A., Brech, A., *et al.*, 2010. FYCO1 is a Rab7 effector that binds to LC3 and PI3P to mediate microtubule plus end-directed vesicle transport. *Journal of Cell Biology* 188, 253–269.
- Pedersen, N.M., Raiborg, C., Brech, A., *et al.*, 2012. The PtdIns3P-binding protein Phafin 2 mediates epidermal growth factor receptor degradation by promoting endosome fusion. *Traffic* 13, 1547–1563.
- Polson, H.E., De Lartigue, J., Rigden, D.J., *et al.*, 2010. Mammalian Atg18 (WIP1) localizes to omegasome-anchored phagophores and positively regulates LC3 lipidation. *Autophagy* 6, 506–522.
- Posor, Y., Eichhorn-Gruenig, M., Puchkov, D., *et al.*, 2013. Spatiotemporal control of endocytosis by phosphatidylinositol-3,4-bisphosphate. *Nature* 499, 233–237.
- Poteryaev, D., Datta, S., Ackema, K., Zerial, M., Spang, A., 2010. Identification of the switch in early-to-late endosome transition. *Cell* 141, 497–508.
- Rink, J., Ghigo, E., Kalaidzidis, Y., Zerial, M., 2005. Rab conversion as a mechanism of progression from early to late endosomes. *Cell* 122, 735–749.
- Rong, Y., Liu, M., Ma, L., *et al.*, 2012. Clathrin and phosphatidylinositol-4,5-bisphosphate regulate autophagic lysosome reformation. *Nature Cell Biology* 14, 924–934.
- Russell, R.C., Tian, Y., Yuan, H., *et al.*, 2013. ULK1 induces autophagy by phosphorylating Beclin-1 and activating VPS34 lipid kinase. *Nature Cell Biology* 15, 741–750.
- Rusten, T.E., Stenmark, H., 2006. Analyzing phosphoinositides and their interacting proteins. *Nature Methods* 3, 251–258.
- Rutherford, A.C., Traer, C., Wassmer, T., *et al.*, 2006. The mammalian phosphatidylinositol 3-phosphate 5-kinase (PIKfyve) regulates endosome-to-TGN retrograde transport. *Journal of Cell Science* 119, 3944–3957.
- Semerdjiova, S., Short, B., Maxwell, E., *et al.*, 2008. Coordinated regulation of AP2 uncoating from clathrin-coated vesicles by rab5 and hRME-6. *Journal of Cell Biology* 183, 499–511.
- Shin, H.W., Hayashi, M., Christoforidis, S., *et al.*, 2005. An enzymatic cascade of Rab5 effectors regulates phosphoinositide turnover in the endocytic pathway. *Journal of Cell Biology* 170, 607–618.
- Silhankova, M., Port, F., Harterink, M., Basler, K., Korswagen, H.C., 2010. Wnt signaling requires MTM-6 and MTM-9 myotubularin lipid-phosphatase function in Wnt-producing cells. *EMBO Journal* 29, 4094–4105.
- Slagsvold, T., Aasland, R., Hirano, S., *et al.*, 2005. Eap45 in mammalian ESCRT-II binds ubiquitin via a phosphoinositide-interacting GLUE domain. *Journal of Biological Chemistry* 280, 19600–19606.
- Sridhar, S., Patel, B., Aphkhasava, D., *et al.*, 2013. The lipid kinase PI4KIIbeta preserves lysosomal identity. *EMBO Journal* 32, 324–339.
- Sun, Y., Drubin, D.G., 2012. The functions of anionic phospholipids during clathrin-mediated endocytosis site initiation and vesicle formation. *Journal of Cell Science* 125, 6157–6165.
- Swanson, J.A., 2008. Shaping cups into phagosomes and macropinosomes. *Nature Reviews Molecular Cell Biology* 9, 639–649.
- Taguchi-Atarashi, N., Hamasaki, M., Matsunaga, K., *et al.*, 2010. Modulation of local PtdIns3P levels by the PI phosphatase MTMR3 regulates constitutive autophagy. *Traffic* 11, 468–478.
- Teo, H., Gill, D.J., Sun, J., *et al.*, 2006. ESCRT-I core and ESCRT-II GLUE domain structures reveal role for GLUE in linking to ESCRT-I and membranes. *Cell* 125, 99–111.
- Toth, B., Balla, A., Ma, H., *et al.*, 2006. Phosphatidylinositol 4-kinase IIbeta regulates the transport of ceramide between the endoplasmic reticulum and Golgi. *Journal of Biological Chemistry* 281, 36369–36377.
- Traer, C.J., Rutherford, A.C., Palmer, K.J., *et al.*, 2007. SNX4 coordinates endosomal sorting of TtnR with dynein-mediated transport into the endocytic recycling compartment. *Nature Cell Biology* 9, 1370–1380.
- Valente, C., Turacchio, G., Mariggio, S., *et al.*, 2012. A 14-3-3gamma dimer-based scaffold bridges CIBP1-S/BARS to PI(4)KIIb to regulate post-Golgi carrier formation. *Nature Cell Biology* 14, 343–354.

- Van Weering, J.R., Verkade, P., Cullen, P.J., 2012. SNX-BAR-mediated endosome tubulation is co-ordinated with endosome maturation. *Traffic* 13, 94–107.
- Velichkova, M., Juan, J., Kadandale, P., *et al.*, 2010. *Drosophila* Mtm and class II PI3K coregulate a PI(3)P pool with cortical and endolysosomal functions. *Journal of Cell Biology* 190, 407–425.
- Wang, J., Sun, H.Q., Macia, E., *et al.*, 2007. PI4P promotes the recruitment of the GGA adaptor proteins to the trans-Golgi network and regulates their recognition of the ubiquitin sorting signal. *Molecular Biology of the Cell* 18, 2646–2655.
- Wang, Y.J., Wang, J., Sun, H.Q., *et al.*, 2003. Phosphatidylinositol 4 phosphate regulates targeting of clathrin adaptor AP-1 complexes to the Golgi. *Cell* 114, 299–310.
- Wawrzyniak, A.M., Kashyap, R., Zimmermann, P., 2013. Phosphoinositides and PDZ domain scaffolds. *Advances in Experimental Medicine and Biology* 991, 41–57.
- Weber-Boyvat, M., Zhong, W., Yan, D., Olkkonen, V.M., 2013. Oxysterol-binding proteins: Functions in cell regulation beyond lipid metabolism. *Biochemical Pharmacology* 86, 89–95.
- Wen, P.J., Osborne, S.L., Zanin, M., *et al.*, 2011. Phosphatidylinositol(4,5) biphosphate coordinates actin-mediated mobilization and translocation of secretory vesicles to the plasma membrane of chromaffin cells. *Nature Communications* 2, 491.
- Wu, Y., Cheng, S., Zhao, H., *et al.*, 2014. PI3P phosphatase activity is required for autophagosome maturation and autolysosome formation. *EMBO Reports* 15, 973–981.
- Yoon, Y., Zhang, X., Cho, W., 2012. Phosphatidylinositol 4,5-bisphosphate (PtdIns (4,5)P₂) specifically induces membrane penetration and deformation by Bin/amphiphysin/Rvs (BAR) domains. *Journal of Biological Chemistry* 287, 34078–34090.
- Zoncu, R., Perera, R.M., Balkin, D.M., *et al.*, 2009. A phosphoinositide switch controls the maturation and signaling properties of APPL endosomes. *Cell* 136, 1110–1121.
- Zoncu, R., Perera, R.M., Sebastian, R., *et al.*, 2007. Loss of endocytic clathrin-coated pits upon acute depletion of phosphatidylinositol 4,5-bisphosphate. *Proceedings of the National Academy of Sciences of the United States of America* 104, 3793–3798.
- Zou, J., Zhang, C., Marjanovic, J., *et al.*, 2012. Myotubularin-related protein (MTMR) 9 determines the enzymatic activity, substrate specificity, and role in autophagy of MTMR8. *Proceedings of the National Academy of Sciences of the United States of America* 109, 9539–9544.

Unconventional Protein Secretion: Fibroblast Growth Factor 2 and Interleukin-1 β as Examples

W Nickel, JP Steringer, and H-M Müller, Heidelberg University Biochemistry Center, Heidelberg, Germany
D Brough, University of Manchester, Manchester, UK

© 2016 Elsevier Inc. All rights reserved.

Mechanisms of Eukaryotic Protein Secretion

The vast majority of soluble secretory proteins make use of the ER/Golgi-dependent pathway to get access to the extracellular space (Rothman and Wieland, 1996; Schekman and Orci, 1996; Figure 1). Entering this route of secretion requires the presence of N-terminal signal peptides that are recognized by SRP (Rapoport, 2007). Following a slow-down of translation, ribosomes carrying nascent chains with N-terminal signal peptides are directed to the cytoplasmic surface of the ER coupling protein synthesis to translocation into the ER lumen mediated by Sec61 (Osborne *et al.*, 2005). Translocation of newly synthesized chains proceeds in a largely unfolded form followed by folding of the translocation substrate into the native state within the ER lumen (Park and Rapoport, 2012). Once secretory proteins pass ER quality control (Hegde and Ploegh, 2010; Sitia and Braakman, 2003), they are delivered to the cell surface based on vesicular transport and, upon fusion of post-Golgi carriers with the plasma membrane, are released into the extracellular space (Rothman, 1994; Figure 1).

It was long believed that the presence of a signal peptide concomitant with a protein's ability to penetrate the lumen of the ER was an essential requirement for protein secretion from eukaryotic cells. This view was challenged by the identification of signal-peptide-lacking cytokines that mediate defined extracellular functions of fundamental physiological importance (Muesch *et al.*, 1990; Nickel, 2003) along with reports presenting direct evidence that they can be secreted in an ER/Golgi-independent manner (Engling *et al.*, 2002; Mignatti *et al.*, 1992; Rubartelli *et al.*, 1990; Trudel *et al.*, 2000). The most classical examples are fibroblast growth factor 2 (FGF2), a proangiogenic mitogen (Presta *et al.*, 2005), as well as interleukin-1 β (IL-1 β), a key pro-inflammatory cytokine (Garlanda *et al.*, 2013; Dinarello, 2009). For both molecules, the cell surface receptors on target cells have been identified and crystal structures of the corresponding ligand–receptor complexes are available (Pellegrini *et al.*, 2000; Plotnikov *et al.*, 2000; Schlessinger *et al.*, 2000; Vigers *et al.*, 1997). Based on this, alternative pathways of protein secretion that operate independently of the ER/Golgi system were proposed to exist in eukaryotic cells (Muesch *et al.*, 1990).

As illustrated in Figure 1, two principal types of unconventional secretion appear to exist that are exemplified by FGF2 and IL-1 β as cargo proteins. While FGF2 is not only excluded from the lumen of the ER, it does not enter any other type of vesicular intermediates inside cells. Instead, as detailed below, FGF2 is capable of translocating across plasma membranes to reach the extracellular space (Figure 1; Nickel, 2005; Schäfer *et al.*, 2004). This may also be the case for IL-1 β (Brough and Rothwell, 2007; Shirasaki *et al.*, 2014). However, a number of additional secretory routes for IL-1 β have also been described. IL-1 β has been described in membrane-bound intermediates of

the endocytic membrane system (Andrei *et al.*, 1999, 2004; Malhotra, 2013; Dupont *et al.*, 2011; Piccioli and Rubartelli, 2013). The identity of these intermediates has not unequivocally been clarified, however, a role for multi-vesicular bodies (MVBs) has been suggested (Qu *et al.*, 2009). In this case, IL-1 β localized in the cytoplasm would be internalized by endosomes resulting in its localization in intra-endosomal vesicles called exosomes. Alternatively, cytoplasmic IL-1 β might be initially captured by autophagosomes followed by delivery to endocytic compartments such as MVBs (Figure 1; Malhotra, 2013; Dupont *et al.*, 2011). This mechanism would be consistent with a proposed pathway of unconventional secretion in lower eukaryotes that has been described for acyl coenzyme A binding protein (AcbA/Acb1) (Malhotra, 2013; Bruns *et al.*, 2011; Duran *et al.*, 2010; Kinseth *et al.*, 2007). Other suggestions have been a role for secretory lysosomes (Rubartelli *et al.*, 1990; Andrei *et al.*, 1999; Piccioli and Rubartelli, 2013) or plasma membrane shedding, i.e., secretion of IL-1 β in microvesicles that form at the extracellular surface of cells and incorporate cytoplasmic material (Figure 1; Piccioli and Rubartelli, 2013; MacKenzie *et al.*, 2001). However, for these options of IL-1 β release, no unambiguous evidence has been presented and the exact mechanism remains elusive. Below, we comment on how these mechanisms of secretion may contribute to the cellular release of IL-1 β during an inflammatory episode.

Hallmarks of FGF2 Membrane Translocation during Unconventional Secretion

As illustrated in Figure 1, our current view on how FGF2 is secreted from cells is based on a mechanism characterized by direct translocation across plasma membranes, i.e., FGF2 is not incorporated into intracellular membrane-bound compartments as intermediates in this alternative pathway of secretion (Nickel, 2005; Schäfer *et al.*, 2004). Instead, FGF2 is capable of interacting with phosphoinositides at the inner leaflet of plasma membranes, a key step that initiates membrane translocation (Temmerman *et al.*, 2008; Temmerman and Nickel, 2009). This process, however, does not appear to rely on a protein-conducting channel but rather on the ability of membrane-recruited FGF2 to oligomerize concomitant with the formation of a membrane pore (Nickel, 2011; Steringer *et al.*, 2012). This structure has been proposed to have a toroidal architecture with the FGF2 oligomer in the center and the lipid binding sites being localized in the periphery. In a way, this structure might share structural similarities with toroidal membrane pores formed by components of the apoptotic machinery involved in the disintegration of the outer mitochondrial membrane (Terrones *et al.*, 2004; Qian *et al.*, 2008). Once FGF2 oligomers have formed membrane pores, membrane-proximal heparan sulphate chains of cell surface localized proteoglycans can act as

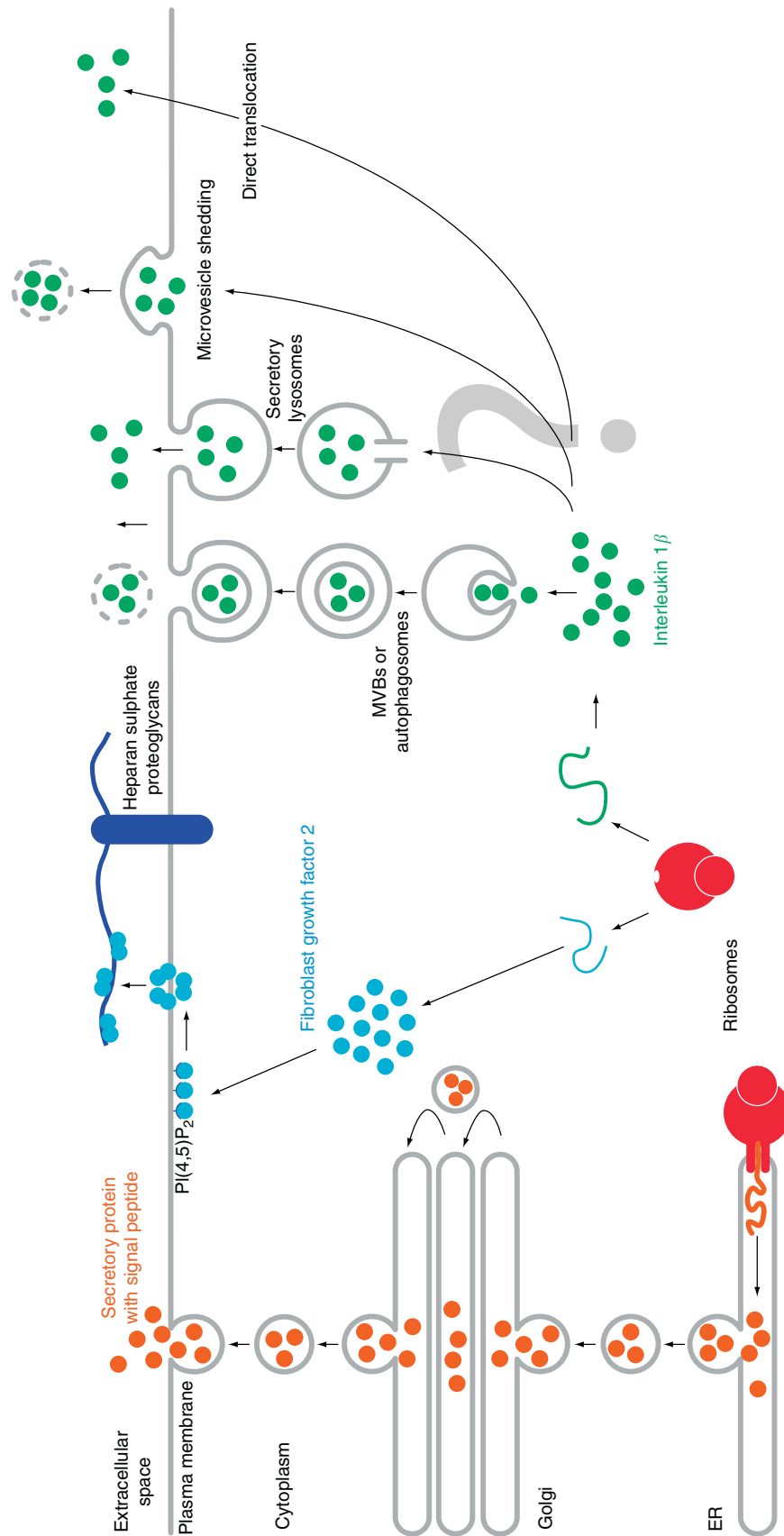


Figure 1 Principle pathways of protein secretion from eukaryotic cells.

an extracellular trap for FGF2 molecules resulting in their exposure on cell surfaces (Nickel, 2007; Zehe *et al.*, 2006). Thus, the biochemical and biophysical properties of FGF2 allow for a mechanism of membrane translocation with the cargo protein (= FGF2 monomer) forming its own translocation machinery (= membrane-inserted FGF2 oligomer) triggered by phosphoinositide-induced multimerization. This model is supported by evidence for membrane translocation of FGF2 in a fully folded state (Backhaus *et al.*, 2004; Torrado *et al.*, 2009) requiring the formation of defined oligomers during membrane pore formation. Further support for this model comes from the observation that Tec kinase, a non-receptor tyrosine kinase that has previously been described in the context of T cell development and activation (Yang *et al.*, 2000), is a regulatory factor of FGF2 secretion. The role of Tec kinase in FGF2 secretion was initially recognized through a genome-wide RNAi screen designed to identify factors involved in this process (Ebert *et al.*, 2010; Nickel, 2010). Tec kinase was further shown to phosphorylate FGF2, a process that facilitates the formation of membrane pores by FGF2 oligomers (Nickel, 2011; Steringer *et al.*, 2012). Since Tec kinase contains a PH domain that mediates both phosphoinositide-dependent membrane recruitment and enzymatic activation (Bradshaw, 2010), FGF2 phosphorylation is likely to occur at the inner leaflet of the plasma membrane (Nickel, 2011). Since FGF2 is a key signaling molecule in the context of many cancers, Tec kinase regulated secretion of FGF2 represents an interesting link to the up-regulation of PI3 kinases in many tumor cells (Liu *et al.*, 2009). PI3 kinases catalyze the formation of PI(3,4,5)P₃ and high cellular levels of this phosphoinositide are likely to support efficient secretion of FGF2 which, in turn, promotes tumor cell proliferation.

The Molecular Mechanism of FGF2 Membrane Translocation and Its Path of Discovery

Our current knowledge about the mechanism by which FGF2 physically traverses the plasma membrane to reach the extracellular space has previously been described explaining the sequence of events as they occur in a cell (Nickel, 2010, 2011; Nickel and Sedorf, 2008; Nickel and Rabouille, 2009; Rabouille *et al.*, 2012). Here, we would like to recount this path of discovery in chronological order.

In 2002, the setup of a cell-based assay to quantify FGF2 secretion was the starting point of these studies as this experimental system provided convincing evidence that FGF2 is secreted by a controlled mechanism (Engling *et al.*, 2002). It also confirmed observations by other laboratories demonstrating that FGF2 secretion does not result in the appearance of soluble FGF2 in cellular supernatants (Trudel *et al.*, 2000). Rather, following membrane translocation, FGF2 was found retained on cell surfaces bound to heparan sulphate proteoglycans. Further observations in the initial phase of these studies indicated that the mechanism of FGF2 secretion from cells is unusual in at least two ways. On the one hand we found that FGF2 has the ability to directly translocate across plasma membranes that had been affinity-purified as so-called inside-out vesicles (Schäfer *et al.*, 2004). This finding suggested that FGF2 is not only incapable of penetrating the lumen of the ER/Golgi system but also does not need to make use of

other membrane-bound intermediates such as endocytic compartments to reach the extracellular space. These results were initially interpreted as a possibility for a plasma membrane resident protein-conducting channel being responsible for membrane translocation of FGF2. This hypothesis appeared plausible also because ABC transporters were known in the plasma membrane mediating the externalization of peptides (McGrath and Varshavsky, 1989; Michaelis, 1993). However, a second key finding at the time was that FGF2 secretion from cells does not involve an unfolded form of FGF2 as an obligatory intermediate (Backhaus *et al.*, 2004). Since both peptide and protein-conducting channels from various membranes including the ER and mitochondria were known to form narrow channels through which largely unfolded translocation substrates are threaded, it appeared unlikely that a plasma membrane resident protein-conducting channel is behind the principle of FGF2 membrane translocation. This view was supported by additional findings demonstrating that membrane-proximal heparan sulphate proteoglycans on cell surfaces are required to complete FGF2 membrane translocation into the extracellular space (Zehe *et al.*, 2006). Since FGF2 binding to heparin depends on FGF2 being folded properly (Torrado *et al.*, 2009), folding of FGF2 is required for translocation to cell surfaces. The same is true for FGF2 binding to the phosphoinositide PI(4,5)P₂ that is both required for FGF2 recruitment at the inner leaflet and essential for FGF2 secretion from cells (Temmerman *et al.*, 2008; Temmerman and Nickel, 2009). Again, binding to PI(4,5)P₂ was found to require FGF2 in a folded form demonstrating that FGF2 needs to be folded to undergo membrane translocation (Torrado *et al.*, 2009). Finally, it was shown that FGF2 is capable of transporting other proteins out of cells based on engineered non-covalent interactions that depend on proper folding (Torrado *et al.*, 2009). These experiments confirmed earlier findings mentioned above and unequivocally demonstrated that unfolded forms are not obligatory intermediates in FGF2 membrane translocation and secretion from cells. In conclusion, as illustrated in Figure 1, it was clear that unconventional secretion of FGF2 occurs by direct translocation across plasma membranes, involves sequential interactions with the phosphoinositide PI(4,5)P₂ at the inner as well as heparan sulphates at the outer leaflet, and is based on a mechanism that is likely not only to be compatible but rather requires FGF2 in a properly folded form. The latter aspect was interpreted as a potential quality control mechanism ensuring that only folded and, therefore, biologically active forms of FGF2 reach the extracellular space (Torrado *et al.*, 2009). Thus, in 2009, the search for a novel type of protein translocation across membranes that is independent of protein-conducting channels began.

PI(4,5)P₂-Induced FGF2 Oligomerization and Membrane Pore Formation: A Transient Intermediate in FGF2 Membrane Translocation

As a starting point to obtain insight into the mechanism of FGF2 membrane translocation, we sought to conduct a global analysis of gene products involved in unconventional secretion of FGF2. In 2010, a genome-wide RNAi screen was completed identifying a large set of candidate proteins

potentially involved in FGF2 membrane translocation. The bioinformatical analysis, validation and functional analysis of these proteins is still ongoing, however, we were lucky to identify and validate one gene product already during pilot experiments that turned out to play a key role, Tec kinase (Ebert *et al.*, 2010). Both RNAi-mediated downregulation and pharmacological inhibition of Tec kinase impaired FGF2 secretion. Moreover, biochemical experiments revealed a direct interaction and phosphorylation of FGF2 by Tec kinase at tyrosine residue 81 (Ebert *et al.*, 2010). FGF2 variant forms lacking this tyrosine were found to be impaired in secretion and, most importantly, a phosphomimetic substitution of tyrosine 81 rendered FGF2 secretion from cells independent of Tec kinase providing conclusive evidence for its role in FGF2 secretion. The identification of Tec kinase was particularly intriguing as this protein contains a PH domain that mediates phosphoinositide-dependent recruitment to the inner leaflet of plasma membranes (Bradshaw, 2010). Tec kinase has a binding preference for PI(3,4,5)P₃ and this interaction causes activation of its enzymatic activity (Bradshaw, 2010). With this finding, principal components required for FGF2 membrane translocation were discovered: the phosphoinositides PI(4,5)P₂ and PI(3,4,5)P₃, the heparan sulphate chains of cell surface proteoglycans and Tec kinase. All of these factors are physically associated with the plasma membrane, the proposed site of FGF2 membrane translocation.

But what is the function of Tec kinase mediated tyrosine phosphorylation in FGF2 membrane translocation? To tackle this problem, we decided to generate a recombinant form of FGF2 in which tyrosine 81 was substituted by an unnatural amino acid, p-carboxymethylphenylalanine (pCMF). This residue is a close structural mimetic of a phosphotyrosine, so we analyzed the biochemical and biophysical properties of this variant form of FGF2 in biochemical reconstitution experiments. Though occurring independently of the phosphomimetic residue, a key observation in these experiments was that binding to PI(4,5)P₂-containing membranes causes FGF2 to oligomerize (Steringer *et al.*, 2012). Based on small fluorescent tracers and various kinds of artificial membranes including large and giant unilamellar vesicles, it became clear that FGF2 oligomers form membrane pores allowing for the passage of small molecules. In addition, transbilayer diffusion of membrane lipids was observed during FGF2-induced membrane pore formation suggesting a toroidal architecture (Steringer *et al.*, 2012). Most intriguingly, a phosphomimetic substitution in FGF2 (FGF2-Y81pCMF) strongly increased the rate of membrane pore formation providing direct evidence that Tec kinase mediated phosphorylation of FGF2 facilitates membrane insertion of FGF2 oligomers. This, in turn, suggests that inhibition of FGF2 secretion in the absence of Tec kinase activity (e.g., following RNAi-mediated down-regulation) is due to inefficient membrane insertion of FGF2 oligomers as membrane translocation intermediates.

But how are membrane-inserted FGF2 oligomers removed from the plasma membrane in order to appear on cell surfaces? The proposed mechanism of PI(4,5)P₂-induced membrane pore formation actually provides a simple explanation for our early findings establishing a role for cell surface heparan sulphates in the final step of FGF2 membrane translocation (Nickel, 2007; Zehe *et al.*, 2006). Since basic residues

in FGF2 that are required for heparin binding are also needed for binding to PI(4,5)P₂ (e.g., lysine 133; Temmerman *et al.*, 2008) it is possible that binding of FGF2 to heparan sulphates versus PI(4,5)P₂ is mutually exclusive. At the same time, the affinity of FGF2 for heparan sulphates is about 100 times higher ($K_D \approx 10$ nM; Faham *et al.*, 1996) compared to PI(4,5)P₂ ($K_D \approx 1$ μ M; Temmerman and Nickel, 2009). This suggests that, once FGF2 molecules become accessible to extracellular heparan sulphates as part of oligomers within a lipidic membrane pore, they can compete against PI(4,5)P₂ for binding to FGF2. We propose that this is the underlying mechanism by which cell surface heparan sulphates mediate the final step of FGF2 membrane translocation resulting in the exposure of FGF2 on cell surfaces. This model also provides a basis for directionality of FGF2 membrane translocation that is defined through sequential interactions of FGF2 with PI(4,5)P₂ at the inner leaflet and heparan sulphates at the outer leaflet. It also emphasizes that PI(4,5)P₂ is not only required for membrane recruitment of FGF2 but also to support oligomerization in the context of plasma membranes. Finally, due to its cone-shaped structure, PI(4,5)P₂ is likely to be important for the toroidal architecture of lipidic membrane pores induced by FGF2 oligomerization. Direct experimental evidence for this view comes from the fact that diacylglycerol, a membrane lipid with the opposite cone-shaped structure of PI(4,5)P₂ inhibits transbilayer diffusion of membrane lipids during membrane pore formation (Steringer *et al.*, 2012). Thus, PI(4,5)P₂ is likely to be required to stabilize the high membrane curvature that is generated during the formation of a toroidal membrane pore (Figure 2).

Understanding the Regulation of IL-1 β Processing

For almost 30 years, since the discoveries that the 'master' cytokine IL-1 β lacks a signal peptide (Auron *et al.*, 1984) and does not traffic through the ER and Golgi (Rubartelli *et al.*, 1990), immunologists have speculated how IL-1 β is secreted. Given the importance of IL-1 β to the host defense response, and its seemingly unique biology, understanding the mechanism of IL-1 β secretion is of outstanding fundamental interest, and holds considerable downstream therapeutic potential. As discussed above, despite a number of interesting ideas, all supported by biochemical data, no consensus for a mechanism of IL-1 β secretion exists (Lopez-Castejon and Brough, 2011).

IL-1 β is produced by many cells, most commonly those of macrophage lineage, as a pro-IL-1 β precursor. This precursor is expressed in response to pathogen associated molecular patterns (PAMPs) or damage associated molecular patterns (DAMPs) that bind to pattern recognition receptors (PRRs) on macrophages to upregulate pro-inflammatory gene expression (Schroder and Tschopp, 2010; Takeuchi and Akira, 2010). PAMPs are motifs carried by pathogens, such as bacterial endotoxin (or LPS) of Gram-negative bacteria, and DAMPs are commonly endogenous molecules released by necrosis. In macrophages this initial stimulus is commonly referred to as a 'priming' step and pro-IL-1 β is inactive and remains intracellular until a further PAMP or DAMP stimulation activates cytosolic PRRs, often of the NLR family, to form large multi-protein complexes called inflammasomes (Latz *et al.*, 2013).

needed for, i.e., why do cells need an alternative to the ER/Golgi-dependent secretory pathway? One way of looking into this was to analyze FGF2 variant forms that contain engineered signal peptides at their N-termini. Following expression in mammalian cells we asked whether they are secreted in an ER/Golgi-dependent manner and whether the secreted population is biologically active (Wegehingel *et al.*, 2008). The first question by no means was a trivial one since FGF2 is a lectin that, upon binding to heparan sulphate chains, tends to oligomerize. So premature binding to its ligands in the lumen of the *trans*-Golgi network might cause aggregation resulting in a block of secretion and potentially even disintegration of Golgi structure. However, FGF2 containing a signal peptide derived from FGF4 was secreted with high efficiency (Wegehingel *et al.*, 2008), yet a remarkable difference to authentic FGF2 was observed. FGF2 secreted by the ER/Golgi-dependent pathway was no longer retained on cell surfaces but rather accumulated in the supernatant of cells. When analyzed by biochemical means we realized that the signal-peptide-containing version of FGF2 was modified running as a rather undefined population with an apparent molecular weight of about 80 kDa. A biochemical analysis revealed that, when forced to travel through the lumen of the ER/Golgi system, FGF2 becomes posttranslational modified by O-linked chondroitin sulphates. As a consequence, this form of FGF2 failed to interact with heparan sulphates on cell surfaces resulting in its accumulation in cellular supernatants (Wegehingel *et al.*, 2008). Since binding of FGF2 to heparan sulphate proteoglycans is essential to form ternary signaling complexes containing FGF high affinity receptors, FGF2 secreted through the ER/Golgi-dependent pathway is impaired in its biological activity. Based on these findings, it is a plausible hypothesis that unconventional secretion permits export from cells of unmodified forms of FGF2 that retains their biological activity. Only limited evidence for such studies exists for IL-1 β . IL-1 β directed through the ER/Golgi is efficiently secreted, but is also N-glycosylated (Baldari *et al.*, 1987; Pecceu *et al.*, 1991). However, how glycosylation affects the biological function of IL-1 β is not clear with evidence suggesting that glycosylated IL-1 β is active (Baldari *et al.*, 1987), and evidence suggesting that glycosylation results in a 95% inhibition of biological activity (Fleer *et al.*, 1991). This concept may actually reflect a general aspect of unconventional secretory processes, the ability of mammalian cells to secrete proteins avoiding modifications that potentially occur in the lumen of the ER/Golgi system such as glycosylation.

Another interesting aspect emerges when one considers the evolution of the FGF protein family (Itoh and Ornitz, 2008). Four of the more than 20 members have been defined as so-called intracellular FGFs (FGF 11–14, termed iFGFs) that are not secreted from cells but are functioning solely inside cells. The iFGFs are believed to be the ancestors of the FGF family (Itoh and Ornitz, 2008). It appears possible that they were followed by FGF1 and FGF2, the two members that are secreted by unconventional means. Only later in evolution, signal-peptide-containing FGFs appeared complementing the FGF family as we know it today. This suggests that FGF1 and FGF2 were the first family members that managed to exit cells in order to execute new functions outside cells. But how did FGF2 find its way into the extracellular space? Based on the

mechanism proposed here (Figure 2), one may wonder whether the ability of FGF2 to form lipidic membrane pores is related to membrane pores generated by some bacterial toxins (Shai, 1999) or by components of the apoptotic machinery that promote the formation of toroidal membrane pores in the outer membrane of mitochondria to cause programmed cell death (Qian *et al.*, 2008; Garcia-Saez *et al.*, 2006). In case of FGF2, however, these pores are not formed to destruct the membrane but rather to build transient intermediates of membrane translocation. Unconventional secretion may allow for a double life of proteins with functions in different localizations based on an equilibrium that can be dynamically changed by cells to match physiological needs. This is also exemplified by the IL-1 family member IL-1 α . Also secreted by poorly defined nonconventional pathways, IL-1 α is also actively trafficked to the cell nucleus (Luheshi *et al.*, 2009b), where it is proposed to have functions leading to the 'dual function' hypothesis (Luheshi *et al.*, 2009a). By contrast, signal-peptide-containing proteins can be considered professional secretory proteins that are doomed for quantitative secretion with their sole function in the extracellular space.

Future Perspectives

To challenge the model of FGF2 membrane translocation discussed in this article new experimental strategies are required to verify its hallmarks and predictions. A major aim will be to reconstitute a full cycle of FGF2 membrane translocation with chemically defined components. A promising approach will be to use giant unilamellar vesicles with heparin trapped in their lumen to mimic membrane-proximal cell surface heparan sulphates. This will allow for testing the capability of fluorescently labeled species of phosphomimetic FGF2 to physically traverse the membrane under these conditions, a key prediction of the model shown in Figure 2. In this context, it is also attractive to address other unconventional secretory proteins such as HIV-Tat, a critical transactivator of transcription required for the life cycle of HIV (Jeang *et al.*, 1999). In addition to its classical role in the nucleus of infected T cells, HIV-Tat was shown to be secreted (Ensoli *et al.*, 1990, 1993) and is present in substantial amounts in the blood of HIV-infected individuals (Xiao *et al.*, 2000). In cell-based experiments, it has been shown that HIV-Tat is indeed secreted from infected T cells using an ER/Golgi-independent pathway (Rayne *et al.*, 2010a). These studies further suggested that HIV-Tat is capable of translocation across plasma membranes, a process that depends on the ability of HIV-Tat to bind to the phosphoinositide PI(4,5)P₂ (Rayne *et al.*, 2010a,b). Another example is FGF1, the second FGF family member secreted by unconventional means (Nickel, 2003). Its secretion has been suggested to occur at the level of plasma membranes depending on binding to acidic phospholipids and local destabilization of membrane integrity (Prudovsky *et al.*, 2008). As discussed above, it is not even clear whether the basic principle behind unconventional secretion of IL-1 β differs from that of FGF2. This is because the reported evidence for a role of endocytic sub-compartments is of indirect nature. A critical shortcoming of experiments supporting a role of, for example, MVBs and exosomes is that IL-1 β so far has not been

localized to transport vesicles whose nature as intermediates of this pathway has been demonstrated unambiguously (Zhang and Schekman, 2013). Therefore, one may wonder whether the substantial structural similarity of mature IL-1 β to FGF2 as well as the typical disruption of plasma membrane integrity observed during LPS/ATP triggered secretion of IL1 β (Shirasaki *et al.*, 2014) indicates a core mechanism that resembles the one we propose for FGF2 (Figure 2).

Other key aims in the analysis of the molecular mechanism, by which FGF2 is secreted from cells to analyze the structure–function relationship of the membrane-inserted FGF2 oligomers as translocation intermediates in unconventional secretion and the identification of further components of this pathway. In the long run, we aim to develop pharmacological inhibitors against each molecular component and step of this process. Furthermore, this knowledge will allow for systematically addressing the question whether indeed multiple, unrelated pathways of unconventional secretion exist or whether one common principle is behind this phenomenon with variations at the regulatory level depending on cell type and physiological conditions.

Acknowledgments

Work in the laboratory of the authors is supported by the German Research Council (DFG-SFB 638, DFG-TRR 83, and DFG-GRK1188), the AID-NET program of the Federal Ministry for Education and Research of Germany, and the DFG cluster of excellence CellNetworks (WN), as well as the Wellcome Trust and the BBSRC (DB).

See also: Molecular Principles, Components, Technology, and Concepts: Membranes: Composition, Physical Properties, and Curvature

References

- Andrei, C., Dazzi, C., Lotti, L., *et al.*, 1999. The secretory route of the leaderless protein interleukin 1 β involves exocytosis of endolysosome-related vesicles. *Molecular Biology of the Cell* 10, 1463–1475.
- Andrei, C., Margiocco, P., Poggi, A., *et al.*, 2004. Phospholipases C and A2 control lysosome-mediated IL-1 β secretion: Implications for inflammatory processes. *Proceedings of the National Academy of Sciences of the United States of America* 101, 9745–9750.
- Antonopoulos, C., El Sanadi, C., Kaiser, W.J., Mocarski, E.S., Dubyak, G.R., 2013. Proapoptotic chemotherapeutic drugs induce noncanonical processing and release of IL-1 β via caspase-8 in dendritic cells. *Journal of Immunology* 191, 4789–4803.
- Auron, P.E., Webb, A.C., Rosenwasser, L.J., *et al.*, 1984. Nucleotide sequence of human monocyte interleukin 1 precursor cDNA. *Proceedings of the National Academy of Sciences of the United States of America* 81, 7907–7911.
- Backhaus, R., Zehe, C., Wegehingel, S., *et al.*, 2004. Unconventional protein secretion: Membrane translocation of FGF-2 does not require protein unfolding. *Journal of Cell Science* 117, 1727–1736.
- Baldari, C., Murray, J.A., Ghiara, P., Cesareni, G., Galeotti, C.L., 1987. A novel leader peptide which allows efficient secretion of a fragment of human interleukin 1 β in *Saccharomyces cerevisiae*. *EMBO Journal* 6, 229–234.
- Bergsbaken, T., Fink, S.L., Cookson, B.T., 2009. Pyroptosis: Host cell death and inflammation. *Nature Reviews Microbiology* 7, 99–109.
- Bossaller, L., Chiang, P.I., Schmidt-Lauber, C., *et al.*, 2012. Cutting edge: FAS (CD95) mediates noncanonical IL-1 β and IL-18 maturation via caspase-8 in an RIP3-independent manner. *Journal of Immunology* 189, 5508–5512.
- Bradshaw, J.M., 2010. The Src, Syk, and Tec family kinases: Distinct types of molecular switches. *Cell Signaling* 22, 1175–1184.
- Brodsky, I.E., Medzhitov, R., 2011. Pyroptosis: Macrophage suicide exposes hidden invaders. *Current Biology* 21, R72–R75.
- Brough, D., Rothwell, N.J., 2007. Caspase-1-dependent processing of pro-interleukin-1 β is cytosolic and precedes cell death. *Journal of Cell Science* 120, 772–781.
- Bruns, C., McCaffery, J.M., Curwin, A.J., Duran, J.M., Malhotra, V., 2011. Biogenesis of a novel compartment for autophagosome-mediated unconventional protein secretion. *Journal of Cell Biology* 195, 979–992.
- Cassel, S.L., Sutterwala, F.S., 2010. Sterile inflammatory responses mediated by the NLRP3 inflammasome. *European Journal of Immunology* 40, 607–611.
- Dinarello, C.A., 2009. Immunological and inflammatory functions of the interleukin-1 family. *Annual Review of Immunology* 27, 519–550.
- DuPont, N., Jiang, S., Pilli, M., *et al.*, 2011. Autophagy-based unconventional secretory pathway for extracellular delivery of IL-1 β . *EMBO Journal* 30, 4701–4711.
- Duran, J.M., Anjard, C., Stefan, C., Loomis, W.F., Malhotra, V., 2010. Unconventional secretion of Acb1 is mediated by autophagosomes. *Journal of Cell Biology* 188, 527–536.
- Ebert, A.D., Lausmann, M., Wegehingel, S., *et al.*, 2010. Tec-kinase-mediated phosphorylation of fibroblast growth factor 2 is essential for unconventional secretion. *Traffic* 11, 813–826.
- England, H., Summersgill, H.R., Edye, M.E., Rothwell, N.J., Brough, D., 2014. Release of interleukin-1 α or interleukin-1 β depends on mechanism of cell death. *Journal of Biological Chemistry* 289, 15942–15950.
- Engling, A., Backhaus, R., Stegmayer, C., *et al.*, 2002. Biosynthetic FGF-2 is targeted to non-lipid raft microdomains following translocation to the extracellular surface of CHO cells. *Journal of Cell Science* 115, 3619–3631.
- Ensolli, B., Barillari, G., Salahuddin, S.Z., Gallo, R.C., Wong-Staal, F., 1990. Tat protein of HIV-1 stimulates growth of cells derived from Kaposi's sarcoma lesions of AIDS patients. *Nature* 345, 84–86.
- Ensolli, B., Buonaguro, L., Barillari, G., *et al.*, 1993. Release, uptake, and effects of extracellular human immunodeficiency virus type 1 Tat protein on cell growth and viral transactivation. *Journal of Virology* 67, 277–287.
- Faham, S., Hileman, R.E., Fromm, J.R., Linhardt, R.J., Rees, D.C., 1996. Heparin structure and interactions with basic fibroblast growth factor. *Science* 271, 1116–1120.
- Fleer, R., Chen, X.J., Amellal, N., *et al.*, 1991. High-level secretion of correctly processed recombinant human interleukin-1 β in *Kluyveromyces fragilis*. *Gene* 107, 285–295.
- Garcia-Saez, A.J., Coraiola, M., Serra, M.D., *et al.*, 2006. Peptides corresponding to helices 5 and 6 of Bax can independently form large lipid pores. *FEBS Journal* 273, 971–981.
- Garlanda, C., Dinarello, C.A., Mantovani, A., 2013. The interleukin-1 family: Back to the future. *Immunity* 39, 1003–1018.
- Hegde, R.S., Ploegh, H.L., 2010. Quality and quantity control at the endoplasmic reticulum. *Current Opinion in Cell Biology* 22, 437–446.
- Itoh, N., Ornitz, D.M., 2008. Functional evolutionary history of the mouse Fgf gene family. *Developmental Dynamics* 237, 18–27.
- Jeang, K.T., Xiao, H., Rich, E.A., 1999. Multifaceted activities of the HIV-1 transactivator of transcription, Tat. *Journal of Biological Chemistry* 274, 28837–28840.
- Kinseth, M.A., Anjard, C., Fuller, D., *et al.*, 2007. The Golgi-associated protein GRASP is required for unconventional protein secretion during development. *Cell* 130, 524–534.
- Latz, E., Xiao, T.S., Stutz, A., 2013. Activation and regulation of the inflammasomes. *Nature Reviews Immunology* 13, 397–411.
- Le Feuvre, R.A., Brough, D., Iwakura, Y., Takeda, K., Rothwell, N.J., 2002. Priming of macrophages with lipopolysaccharide potentiates P2 $\times$ 7-mediated cell death via a caspase-1-dependent mechanism, independently of cytokine production. *Journal of Biological Chemistry* 277, 3210–3218.
- Liu, P., Cheng, H., Roberts, T.M., Zhao, J.J., 2009. Targeting the phosphoinositide 3-kinase pathway in cancer. *Nature Reviews Drug Discovery* 8, 627–644.
- Lopez-Castejon, G., Brough, D., 2011. Understanding the mechanism of IL-1 β secretion. *Cytokine & Growth Factor Reviews* 22, 189–195.
- Luheshi, N.M., Rothwell, N.J., Brough, D., 2009a. Dual functionality of interleukin-1 family cytokines: Implications for anti-interleukin-1 therapy. *British Journal of Pharmacology* 157, 1318–1329.
- Luheshi, N.M., Rothwell, N.J., Brough, D., 2009b. The dynamics and mechanisms of interleukin-1 α and β nuclear import. *Traffic* 10, 16–25.
- MacKenzie, A., Wilson, H.L., Kiss-Toth, E., *et al.*, 2001. Rapid secretion of interleukin-1 β by microvesicle shedding. *Immunity* 15, 825–835.

- Malhotra, V., 2013. Unconventional protein secretion: An evolving mechanism. *EMBO Journal* 32, 1660–1664.
- McGrath, J.P., Varshavsky, A., 1989. The yeast STE6 gene encodes a homologue of the mammalian multidrug resistance P-glycoprotein. *Nature* 340, 400–404.
- Michaelis, S., 1993. STE6, the yeast a-factor transporter. *Seminars in Cell Biology* 4, 17–27.
- Mignatti, P., Morimoto, T., Rifkin, D.B., 1992. Basic fibroblast growth factor, a protein devoid of secretory signal sequence, is released by cells via a pathway independent of the endoplasmic reticulum-Golgi complex. *Journal of Cellular Physiology* 151, 81–93.
- Muesch, A., Hartmann, E., Rohde, K., *et al.*, 1990. A novel pathway for secretory proteins? *Trends in Biochemical Sciences* 15, 86–88.
- Netea, M.G., Nold-Petry, C.A., Nold, M.F., *et al.*, 2009. Differential requirement for the activation of the inflammasome for processing and release of IL-1 β in monocytes and macrophages. *Blood* 113, 2324–2335.
- Nickel, W., 2003. The mystery of nonclassical protein secretion. A current view on cargo proteins and potential export routes. *European Journal of Biochemistry* 270, 2109–2119.
- Nickel, W., 2005. Unconventional secretory routes: Direct protein export across the plasma membrane of mammalian cells. *Traffic* 6, 607–614.
- Nickel, W., 2007. Unconventional secretion: An extracellular trap for export of fibroblast growth factor 2. *Journal of Cell Science* 120, 2295–2299.
- Nickel, W., 2010. Pathways of unconventional protein secretion. *Current Opinion in Biotechnology* 21, 621–626.
- Nickel, W., 2011. The unconventional secretory machinery of fibroblast growth factor 2. *Traffic* 12, 799–805.
- Nickel, W., Rabouille, C., 2009. Mechanisms of regulated unconventional protein secretion. *Nature Reviews Molecular Cell Biology* 10, 148–155.
- Nickel, W., Seedorf, M., 2008. Unconventional mechanisms of protein transport to the cell surface of eukaryotic cells. *Annual Review of Cell and Developmental Biology* 24, 287–308.
- Osborne, A.R., Rapoport, T.A., van den Berg, B., 2005. Protein translocation by the Sec61/SecY channel. *Annual Review of Cell and Developmental Biology* 21, 529–550.
- Park, E., Rapoport, T.A., 2012. Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annual Review of Biophysics* 41, 21–40.
- Pecceu, F., Dousset, P., Shire, D., *et al.*, 1991. Human interleukin 1 beta fused to the human growth hormone signal peptide is N-glycosylated and secreted by Chinese hamster ovary cells. *Gene* 97, 253–258.
- Pelegrin, P., Barroso-Gutierrez, C., Surprenant, A., 2008. P2 $\times$ 7 receptor differentially couples to distinct release pathways for IL-1 β in mouse macrophage. *Journal of Immunology* 180, 7147–7157.
- Pellegrini, L., Burke, D.F., von Delft, F., Mulloy, B., Blundell, T.L., 2000. Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature* 407, 1029–1034.
- Piccoli, P., Rubartelli, A., 2013. The secretion of IL-1 β and options for release. *Seminars in Immunology* 25 (6), 425–429.
- Plotnikov, A.N., Hubbard, S.R., Schlessinger, J., Mohammadi, M., 2000. Crystal structures of two FGF-FGFR complexes reveal the determinants of ligand-receptor specificity. *Cell* 101, 413–424.
- Presta, M., Dell'Era, P., Mitola, S., *et al.*, 2005. Fibroblast growth factor/fibroblast growth factor receptor system in angiogenesis. *Cytokine & Growth Factor Reviews* 16, 159–178.
- Prudovsky, I., Tarantini, F., Landriscina, M., *et al.*, 2008. Secretion without Golgi. *Journal of Cellular Biochemistry* 103, 1327–1343.
- Qian, S., Wang, W., Yang, L., Huang, H.W., 2008. Structure of transmembrane pore induced by Bax-derived peptide: Evidence for lipidic pores. *Proceedings of the National Academy of Sciences of the United States of America* 105, 17379–17383.
- Qu, C., Gardner, P., Schrijver, I., 2009. The role of the cytoskeleton in the formation of gap junctions by Connexin 30. *Experimental Cell Research* 315, 1683–1692.
- Rabouille, C., Malhotra, V., Nickel, W., 2012. Diversity in unconventional protein secretion. *Journal of Cell Science* 125, 5251–5255.
- Rapoport, T.A., 2007. Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes. *Nature* 450, 663–669.
- Rayne, F., Debaisieux, S., Bonhoure, A., Beaumelle, B., 2010a. HIV-1 Tat is unconventionally secreted through the plasma membrane. *Cell Biology International* 34, 409–413.
- Rayne, F., Debaisieux, S., Yezid, H., *et al.*, 2010b. Phosphatidylinositol-(4,5)-bisphosphate enables efficient secretion of HIV-1 Tat by infected T-cells. *EMBO Journal* 29, 1348–1362.
- Rothman, J.E., 1994. Mechanisms of intracellular protein transport. *Nature* 372, 55–63.
- Rothman, J.E., Wieland, F.T., 1996. Protein sorting by transport vesicles. *Science* 272, 227–234.
- Rubartelli, A., Cozzolino, F., Talio, M., Sitia, R., 1990. A novel secretory pathway for interleukin-1 beta, a protein lacking a signal sequence. *EMBO Journal* 9, 1503–1510.
- Schäfer, T., Zentgraf, H., Zehe, C., *et al.*, 2004. Unconventional secretion of fibroblast growth factor 2 is mediated by direct translocation across the plasma membrane of mammalian cells. *Journal of Biological Chemistry* 279, 6244–6251.
- Schekman, R., Orci, L., 1996. Coat proteins and vesicle budding. *Science* 271, 1526–1533.
- Schlessinger, J., Plotnikov, A.N., Ibrahimi, O.A., *et al.*, 2000. Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Molecular Cell* 6, 743–750.
- Schroder, K., Tschopp, J., 2010. The inflammasomes. *Cell* 140, 821–832.
- Shai, Y., 1999. Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides. *Biochimica et Biophysica Acta* 1462, 55–70.
- Shimada, K., Crother, T.R., Karlin, J., *et al.*, 2012. Oxidized mitochondrial DNA activates the NLRP3 inflammasome during apoptosis. *Immunity* 36, 401–414.
- Shirasaki, Y., Yamagishi, M., Suzuki, N., *et al.*, 2014. Real-time single-cell imaging of protein secretion. *Scientific Reports* 4, 4736.
- Sitia, R., Braakman, I., 2003. Quality control in the endoplasmic reticulum protein factory. *Nature* 426, 891–894.
- Steringer, J.P., Bleicken, S., Andreas, H., *et al.*, 2012. Phosphatidylinositol 4,5-Bisphosphate (PI(4,5)P₂)-dependent oligomerization of fibroblast growth factor 2 (FGF2) triggers the formation of a lipidic membrane pore implicated in unconventional secretion. *Journal of Biological Chemistry* 287, 27659–27669.
- Takeuchi, O., Akira, S., 2010. Pattern recognition receptors and inflammation. *Cell* 140, 805–820.
- Temmerman, K., Ebert, A.D., Muller, H.M., *et al.*, 2008. A direct role for phosphatidylinositol-4,5-bisphosphate in unconventional secretion of fibroblast growth factor 2. *Traffic* 9, 1204–1217.
- Temmerman, K., Nickel, W., 2009. A novel flow cytometric assay to quantify interactions between proteins and membrane lipids. *Journal of Lipid Research* 50 (6), 1245–1254.
- Terrones, O., Antonsson, B., Yamaguchi, H., *et al.*, 2004. Lipid pore formation by the concerted action of proapoptotic BAX and tBID. *Journal of Biological Chemistry* 279, 30081–30091.
- Torrado, L.C., Temmerman, K., Muller, H.M., *et al.*, 2009. An intrinsic quality-control mechanism ensures unconventional secretion of fibroblast growth factor 2 in a folded conformation. *Journal of Cell Science* 122, 3322–3329.
- Trudel, C., Faure-Desire, V., Florkiewicz, R.Z., Baird, A., 2000. Translocation of FGF2 to the cell surface without release into conditioned media [In Process Citation]. *Journal of Cellular Physiology* 185, 260–268.
- Vigers, G.P.A., Anderson, L.J., Caffes, P., Brandhuber, B.J., 1997. Crystal structure of the type-I interleukin-1 receptor complexed with interleukin-1 beta. *Nature* 386, 190–194.
- Wegehingel, S., Zehe, C., Nickel, W., 2008. Rerouting of fibroblast growth factor 2 to the classical secretory pathway results in post-translational modifications that block binding to heparan sulfate proteoglycans. *FEBS Letters* 582, 2387–2392.
- Xiao, H., Neuveut, C., Tiffany, H.L., *et al.*, 2000. Selective CXCR4 antagonism by Tat: Implications for *in vivo* expansion of coreceptor use by HIV-1. *Proceedings of the National Academy of Sciences of the United States of America* 97, 11466–11471.
- Yang, W.C., Collette, Y., Nunes, J.A., Olive, D., 2000. Tec kinases: A family with multiple roles in immunity. *Immunity* 12, 373–382.
- Zehe, C., Engling, A., Wegehingel, S., Schafer, T., Nickel, W., 2006. Cell-surface heparan sulfate proteoglycans are essential components of the unconventional export machinery of FGF-2. *Proceedings of the National Academy of Sciences of the United States of America* 103, 15479–15484.
- Zhang, M., Schekman, R., 2013. Cell biology. Unconventional secretion, unconventional solutions. *Science* 340, 559–561.

Endoplasmic Reticulum Stress in Disease

RJ Kaufman, Sanford Burnham Medical Research Institute, La Jolla, CA, USA

L Popolo, Università degli Studi di Milano, Via Celoria, Milano, Italy

© 2016 Elsevier Inc. All rights reserved.

Glossary

Conditional deletion The genetic ablation of a specific gene in a cell-specific or temporal-restricted fashion in a multicellular organism by use of the site-specific recombinase technology. Frequently, this is achieved by flanking the target region for deletion by LoxP sequences and placing the Cre recombinase under the control of a cell type-specific or regulatable promoter that can be switched on only in the cells of interest. An example is the use of the rat insulin promoter to restrict the expression of Cre in β -cells and therefore achieve the specific ablation of the gene of interest only in β -cells. In some cases the Cre recombinase may be fused to a protein that may be activated by a ligand. An example would be a Cre recombinase-estrogen receptor fusion protein that upon treatment with the estrogen antagonist, tamoxifen, the fusion protein translocates to the nucleus to mediate deletion of the LoxP-flanked gene. The gene deletion can generate a heterozygous or a homozygous state.

Knock out The deletion of a gene in all the cells of an organism. It can be obtained in a heterozygous or homozygous state.

Polyubiquitination A posttranslational protein modification consisting of the formation of a chain of

molecules of ubiquitin, a 76 amino acid protein, linked one to another, and originated from the linkage of the first ubiquitin to a single lysine residue in the target protein. The presence of ubiquitin chain signals the disposal of the protein via the 26S proteasome or may signal trafficking, such as internalization from the cell membrane.

RNA interference (RNA_i) technology In the presence of a small interfering (si) RNAs for a gene, the activity of that gene is reduced, often giving a phenocopy of a hypomorphic allele in the gene. The technology is based on the annealing of a short antisense RNA to the corresponding mRNA and consequent destruction of the specific mRNA molecule or translational inhibition.

Signaling pathway A chain of molecular events that converts the presence of a stimulus into a physiological response. After the activation of the response, regulatory mechanisms reduce the sensitivity of the cell to the signal.

Synthetic lethality A lethal effect caused by the combination of the inactivation of two genes, or specific inhibition of their products. A functional interaction among the gene products or a reciprocal compensation often underlies a synthetic lethal interaction.

Introduction

The endoplasmic reticulum (ER) is a membranous system that expands from the outer membrane of the nuclear envelope and plays several fundamental functions. The ER coordinates the biosynthesis and input of proteins into the secretory pathway along which they are transported to their final destination (the Golgi apparatus, lysosomes, the endocytic compartment, plasma membrane, or the extracellular milieu). The importance of protein secretion in humans is underlined by the estimation that about one-third of proteins encoded by the human genome enter the secretory pathway and also many secreted proteins, such as hormones, growth factors, digestive enzymes, and blood proteins, are fundamental for human physiology. Cells specialized in secretion of large amounts of proteins, such as plasma cells, pancreatic endocrine or exocrine cells, and hepatocytes, exhibit a marked expansion of the ER indicating that this organelle has the striking ability to adapt its size to the needs of the cell. The ER is also the site of biosynthesis of lipid and sterols, essential membrane components that are required for compartmentalization of eukaryotic cells. The biophysical properties of membranes differ between different organelles to ensure proper localization and compartmentalization of proteins within the cell, to optimize

activities of membrane-localized enzymes and to direct protein trafficking mediated by vesicle transport. The ER is the principal site for calcium (Ca^{++}) storage and signaling and detoxification reactions. The ER orchestrates all these activities and consequently is susceptible to stress if the balance is perturbed.

One crucial requirement for cell survival is the ability to rapidly respond to environmental changes and adapt to it. Alterations in the extracellular environment, such as decreased oxygen or nutrient supply, as well as nutrient excess such as cholesterol or fatty acids, or changes in intracellular homeostasis, such as Ca^{++} flux, altered redox status, cell cycle progression, cellular responses to hormonal, or growth factor stimulation, can change the protein-folding load and/or efficiency in the ER leading to accumulation of misfolded proteins. Overall, this article focuses on the current knowledge on how the ER perceives alterations in cell homeostasis and communicates information to other organelles to elicit an appropriate biological response. This communication is mediated in part through a collection of signaling pathways termed the unfolded protein response (UPR) (Schroder and Kaufman, 2005b; Walter and Ron, 2011). The UPR is essential for cells to adapt to altered cellular environments, and if insufficient, initiates a cell death response (Tabas and Ron, 2011; Cao and Kaufman, 2014).

Protein Synthesis, Folding, and Quality Are Tightly Controlled in the ER

Proteins enter the ER during their synthesis. In this co-translational process, the polypeptide is translocated across the ER membrane through the Sec61 channel in an unfolded state. As nascent polypeptides enter the ER lumen during synthesis, they fold and assemble into their native conformations by the assistance of resident chaperones, which include the peptide-binding chaperones, such as binding immunoglobulin protein (BiP)/glucose-regulated protein of 78 kDa (GRP78) and GRP94, and their associated co-chaperones, including the DnaJ family members. As the ER is an oxidizing environment, folding involves the formation of disulfide bonds. In addition to peptide-binding chaperones, enzymes that catalyze protein folding include thiol-disulfide isomerases which catalyze disulfide bond formation (the protein disulfide isomerase (PDI) protein family, ERp57, ERp72) by oxidation of protein thiols (2 SH to S-S) at the expense of reduction of their own disulfide bonds (S-S to 2 SH), and peptidyl-prolyl cis-trans isomerase family members, such as cyclophilin B, that enhance the rate-limiting steps of protein folding. As the protein enters the ER it is subject to addition of core oligosaccharide structures to selective asparagine (N) residues that have the consensus Asn-X[P]-Ser/Thr where X may be any residue except proline. N-linked glycosylation increases protein solubility and trimming of the oligosaccharide core is coupled with interactions with lectins that assist protein folding, trafficking, and even degradation within the early secretory pathway (Ferris *et al.*, 2014). Properly folded proteins transit to the Golgi apparatus and are further modified and sorted to their final destination or stored in vesicles (granules) until a signal stimulates their release into the extracellular space.

The ER harbors an exquisite surveillance system that ensures only properly folded and assembled proteins traffic to the Golgi apparatus, a process termed ER protein quality control (ERQC). Proteins that do not pass ERQC are diverted to degradation by retro-translocation back into the cytosol, polyubiquitination, and degradation by the 26S proteasome in a process termed ER-associated protein degradation (ERAD) (Hampton and Sommer, 2012) or by a process of bulk auto-digestion called autophagy (Buchberger, 2014).

Misfolding of mutant proteins in the ER is the cause of many human genetic diseases. These pathological conditions and those associated with cytosolic misfolded proteins, collectively referred to as 'conformational diseases,' are outside the scope of this article.

ER protein-folding homeostasis depends on a strict coordination between the rates of protein synthesis, folding, and degradation. Perturbations, such as an increase in the rate of protein synthesis, reduced folding capacity, inefficient disulfide bond formation, energy depletion, changes in the ER redox status, or inhibition of degradation cause accumulation of misfolded protein in the ER, a condition termed ER stress.

The Discovery of the UPR as an Example of Interorganellar Communication

The UPR was discovered in the mid-1980s when it was found that misfolded proteins bind to the protein chaperone BiP/GRP78, a process that activates transcription of a set of genes,

originally characterized as genes induced upon glucose deprivation of cells, hence the name glucose-regulated proteins (Dorner *et al.*, 1987, 1989; Kozutsumi *et al.*, 1988). Subsequently, the first and most conserved sensor of misfolded proteins was discovered in yeast in the early 1990s through a genetic selection. Since the gene was previously identified as one required for growth of yeast in the absence of inositol, it was termed inositol requiring enzyme (IRE). Subsequently, mammalian homologs were identified, IRE1 α and IRE1 β . It was later discovered that IRE1 has dual catalytic activities of protein kinase and endoribonuclease (RNase) that cleaves an intron from its substrate Hac1p mRNA in yeast and X-box binding protein 1 (XBP1) mRNA in metazoans. Subsequently, PKR-related ER kinase (PERK) was identified as a protein kinase that phosphorylates the α -subunit of eIF2 to attenuate protein synthesis (Harding *et al.*, 1999) and activating transcription factor 6 (ATF6) was identified as a transcription factor that is cleaved from the ER membrane upon ER stress to generate a cytosolic fragment that migrates to the nucleus to activate gene transcription (Schroder and Kaufman, 2005b).

The UPR can be experimentally activated by inhibition of N-glycosylation (by the antibiotic tunicamycin), nutrient deprivation (amino acids or glucose), Ca⁺⁺ depletion (thapsigargin, an inhibitor of the sarcoplasmic reticulum Ca⁺⁺-dependent ATPase, SERCA), reductive stress (dithiothreitol or homocysteine), increases in protein expression (increased protein-folding load), synthesis of mutant proteins that misfold or inhibition of ERAD (inhibitors of the 26S proteasome). In mammalian cells the activation of the UPR provokes three biological responses to relieve cells from ER stress: (1) attenuation of protein synthesis to decrease protein influx into the ER, (2) stimulation of protein chaperone synthesis to promote protein folding, and (3) increased expression of factors required for ERAD or autophagy, i.e., to increase removal of terminally misfolded proteins (Schroder and Kaufman, 2005b). However, if these mechanisms fail and ER stress and protein misfolding persist, the UPR initiates an apoptotic pathway to cell death, which can contribute to the pathogenesis of many diseases (Wang and Kaufman, 2012).

ER Transmembrane Proteins Sense Misfolded Proteins to Activate UPR Signaling

In mammalian cells, ER stress activates the three sensors: IRE1 α/β , PERK, and ATF6 α . As depicted in the inset of Figure 1, IRE1 and PERK are type I transmembrane proteins that share a related luminal N-terminal ER sensing domain whereas ATF6 α is a type II transmembrane protein that has a C-terminal sensing domain. Each protein has a cytosolic portion that transduces the signal. Both IRE1 and PERK contain a cytosolic serine/threonine protein kinase domain but IRE1 has an additional RNase domain. The cytosolic N-terminal domain of ATF6 α is a basic leucine zipper (bZIP) transcriptional activator. Under normal conditions, the luminal domains of the three sensors bind BiP and are maintained in an inactive state. When misfolded proteins accumulate in the ER, BiP binding is diverted to the misfolded proteins, and this permits dimerization and activation of the protein kinases IRE1 and PERK and release of ATF6 α from the ER for

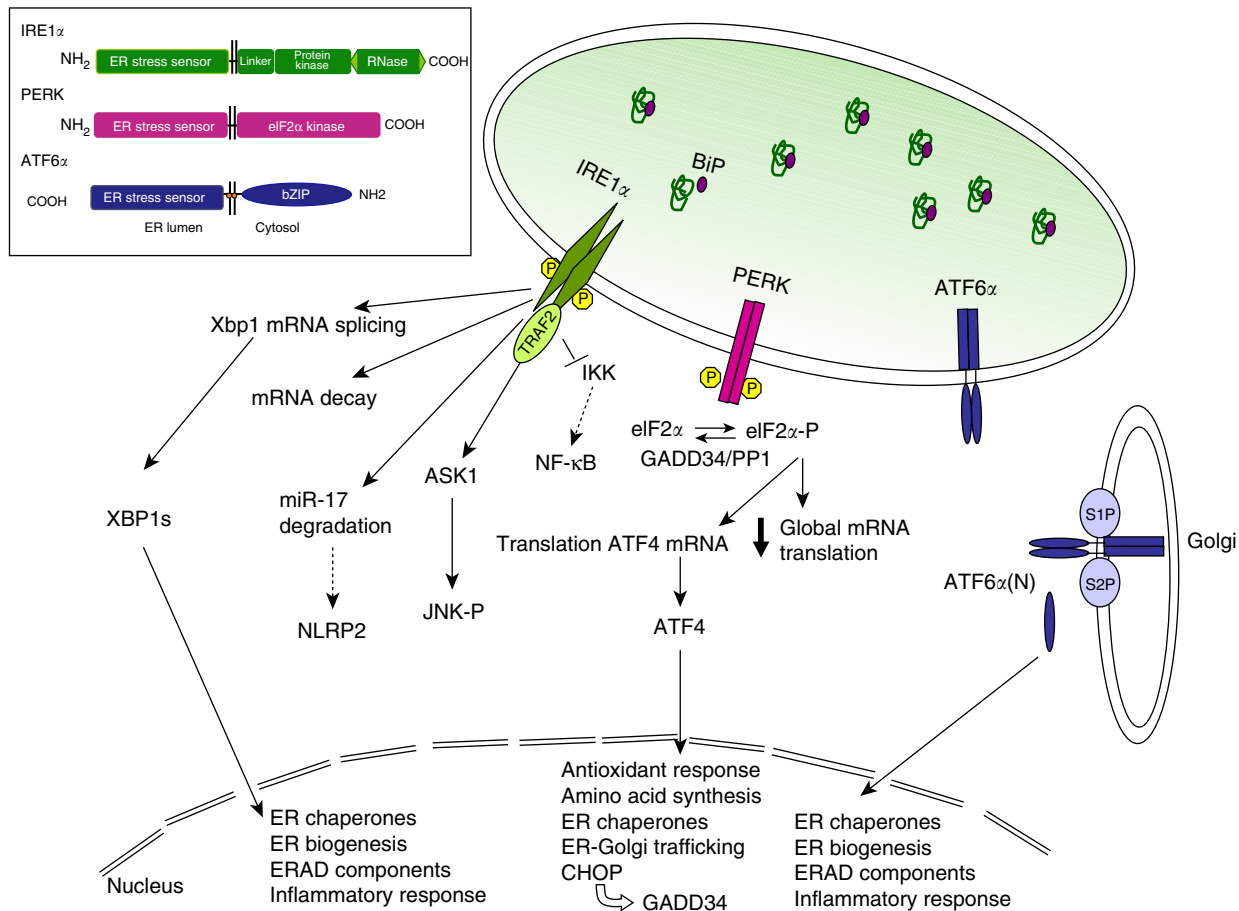


Figure 1 The intracellular signaling pathways induced during the UPR in metazoans. In the inset: a schematic representation of the domain organization of the three ER transmembrane protein sensors of the UPR: IRE1 α , PERK, and ATF6 α . The accumulation of unfolded/misfolded proteins in the ER lumen sequesters BiP and induces the dimerization and consequent *trans*-autophosphorylation of IRE1 α and PERK protein kinase. Phosphorylation of IRE1 α activates its endoribonuclease activity to initiate splicing of XBP1 mRNA. Upon release from BiP, ATF6 α transits to the Golgi compartment for processing by S1P and S2P cleavage to release the transcriptional active proteolytic fragment of 50 kDa, ATF6 α (N). The different arms of the UPR induce in concert a wide number of transcriptional and translational responses and regulation of selective RNAs degradation as explained in detail in the text.

trafficking to the Golgi compartment (Figure 1). The recent X-ray crystal structures of yeast and mammalian IRE1 and PERK provided insights in the molecular mechanism of ER stress sensing (Kimata and Kohno, 2011).

IRE1 Is an Evolutionarily Conserved UPR Sensor

IRE1 is the only UPR sensor conserved in all eukaryotes, from yeast to mammals, indicating that it plays fundamental functions for the cell. Mammalian IRE1 exists in two isoforms encoded by two homologous genes: IRE1 α is the best characterized and ubiquitously expressed while IRE1 β expression is restricted to the intestinal and lung epithelial cells where it may play a role in lipid adsorption. ER stress induces IRE1 α oligomerization and *trans*-autophosphorylation of juxtaposed kinase domains to activate the RNase domain (Figure 1). The RNase activity initiates unconventional splicing of the XBP1 mRNA that removes an internal 26-nucleotide intron to create

a translational frameshift. Translation of the spliced transcript produces XBP1s (s stands for spliced), a potent bZIP transcription factor that enters the nucleus and activates the transcription of UPR genes, notably chaperones or components of ERAD that have a UPR element (UPRE) in their promoter regions (Figure 1). IRE1 can also cleave mRNAs leading to their degradation, a process named regulated IRE1 α -dependent RNA decay or mediate the degradation of selected microRNAs (miRNAs) (Chitnis *et al.*, 2013; Coelho and Domingos, 2014). Activation of IRE1 α also induces a conformational change that allows the sensor to bind to the adaptor protein tumor-necrosis factor α receptor-associated factor 2 (TRAF2) and 6 (TRAF6) to contribute to an inflammatory response signaling through activation of apoptosis signal regulating-kinase 1 (ASK1) and the c-Jun N-terminal kinase (JNK) and activation of nuclear factor kappa B (NF- κ B) (Urano *et al.*, 2000; Wang and Kaufman, 2012). Overall, these discoveries highlight the multiplicity of molecular mechanisms that mediate the IRE1 α arm of the UPR.

PERK-Mediated Phosphorylation of α -Subunit of the Eukaryotic Translation Initiation Factor 2 Regulates mRNA Translation and Induces Death Signals

Upon ER stress, PERK undergoes dimerization/oligomerization and *trans*-autophosphorylation that activates the phosphorylation on Serine-51 of the α -subunit of eIF2 (eukaryotic translation initiation factor 2). This modification blocks eIF2 recycling and causes a shortage of the ternary complex (eIF2-GTP-Met-tRNAi) required for translation initiation. As shown in **Figure 1**, upon phosphorylation of the α -subunit of eIF2 (eIF2 α), protein synthesis is globally attenuated to reduce the protein flux into the ER. Paradoxically, activating transcription factor-4 (ATF4) mRNA, as well as a rapidly growing list of additional mRNAs, is preferentially translated under these conditions due to the presence of open reading frames within the 5'-untranslated region of its mRNA (Harding *et al.*, 2000a). ATF4 induces the expression of genes involved in ER-function antioxidant responses and amino acid biosynthesis and transport, as well as ER-stress-induced apoptosis (**Figure 1**; Harding *et al.*, 2003; Han *et al.*, 2013a). PERK was also reported to phosphorylate nuclear factor erythroid 2 (NF-E2) p45-related factor 2 (NRF2) to induce antioxidant response genes, including heme oxygenase1 and glutathione S-transferase. However, the gene expression changes induced by PERK activation are completely prevented by mutation at the Ser51 phosphorylation site to Ala in fibroblasts, suggesting all transcriptional regulation downstream of PERK activation is mediated through eIF2 α phosphorylation (Lu *et al.*, 2004).

ATF6 Is a Membrane-Bound Transcription Factor Regulated by Proteolysis

ATF6 α is the third arm of the mammalian UPR. The ATF6 β isoform is a distantly related homolog of ATF6 α and both are expressed in all tissues, although in mammals ATF6 β does not activate UPR gene expression. ATF6 α is an ER-resident transcription factor. As shown in **Figure 1**, during ER stress, ATF6 α is released from BiP and traffics to the Golgi apparatus, where it is sequentially cleaved by site-1 protease (S1P) and site-2 protease (S2P) at the transmembrane sites to release a cytosolic portion, p50 ATF6 α (N), that migrates to the nucleus to activate transcription (Shen and Prywes, 2005). In the nucleus ATF6 α (N) acts independently or synergistically with XBP1s to induce UPR target genes having the ER stress element (ERSE) in their promoter regions (Wu *et al.*, 2007; Yamamoto *et al.*, 2007). Interestingly, S1P and S2P are also responsible for cleavage of SREBPs (sterol-regulatory element binding proteins) that regulate lipid and cholesterol biosynthesis at the ER through activating transcription of lipid biosynthetic genes in response to reduced lipid or cholesterol levels.

The UPR Pathways Have Distinct and Overlapping Roles

Studies in genetically manipulated cells and organisms indicated that the three arms of the UPR provide different functions depending on the cell type and differentiation state. The IRE1 α /XBP1 pathway is particularly important for differentiation of cells that secrete large amounts of protein, such as plasma cells, pancreatic acinar, and β -cells (Wang and

Kaufman, 2012; Moore and Hollien, 2012). In an exocrine model of *in vitro* differentiation, inhibition of IRE1 α RNase activity did not alter the cellular response to ER stress but reduced membrane expansion (Cross *et al.*, 2012). These results suggest that in this model system, IRE1 α may play a more significant role in ER expansion associated with differentiation of secretory cell type than adaptation to ER stress (Kaufman *et al.*, 2002).

Studies in *Caenorhabditis elegans*, a nematode model of development, have provided a unique approach to study the UPR through application of RNA interference (RNAi) technology to inactivate genes, the use of mRNA microarray expression technology and the advantage of that each UPR sensor is encoded by a single gene. In *C. elegans*, the *Ire-1/Xbp1* pathway mediates the majority of UPR signaling during normal development and in response to acute ER stress. Deletion of either *atf6* or *pek-1* or their combination did not create a significant phenotype. In contrast, deletion of either *atf6* or *pek-1* was essential for the larval development in either *ire1* (RNAi) or *xbp-1* (RNAi) animals indicating that the two arms complement each other. Also upon acute ER stress, at least two coordinated transcriptional responses mediated by either *atf-6* or *pek-1* are required together with *ire-1/xbp1*. These findings indicate that *Ire-1/Xbp1* pathway is the most significant UPR pathway in *C. elegans* (Shen *et al.*, 2005).

Recent studies using *C. elegans* deleted in *ire-1* unveiled a link between ER homeostasis and nonsense-mediated RNA decay that degrades mRNAs carrying premature stop codons created by nonsense mutations (Sakaki *et al.*, 2012). Nonsense-mediated RNA decay prevents expression of C-terminal truncated proteins that would misfold. It was proposed this new connection protects cells from toxic effects of truncated proteins that are translocated into the ER and misfold and is an example of how the UPR intersects with other intracellular signaling pathways.

Chronic UPR Activation Causes Apoptosis in Mammalian Cells

Cells exposed to intense and/or chronic ER stress undergo apoptosis (Tabas and Ron, 2011; Logue *et al.*, 2013). Swelling and dilation of the ER is a hallmark of chronic ER stress. The PERK-eIF2 α -ATF4 arm of the UPR is a primary determinant for apoptosis since it culminates in the induction of the pro-apoptotic transcription factor C/EBP homologous protein (CHOP)/growth-arrest and DNA damage inducible protein (GADD153) (Rutkowski *et al.*, 2006). CHOP up-regulates apoptosis-related genes including the death receptor DR5 and the pro-apoptotic proteins BIM and PUMA, the pseudokinase TRB3 to promote cell death in response to ER stress (Lu *et al.*, 2014; Wang and Kaufman, 2012). CHOP also mediates the induction of GADD34, a regulatory subunit of protein phosphatase 1 (PP1) that dephosphorylates eIF2 α (**Figure 1**). This feed-back loop on eIF2 restores mRNA translation after the protein-folding defect is resolved. However, if protein misfolding is not resolved, resumption of protein synthesis causes oxidative stress leading to apoptosis (Han *et al.*, 2013a). This is suggested by the protection from ER-stress-induced tissue damage exhibited by CHOP or GADD34 knockout mice and

rescue from protein misfolding stress in cells treated with inhibitors of PP1 or GADD34, as well as antioxidants (Wang and Kaufman, 2012). Thus, when protein folding is disturbed, the sustained action of GADD34 increases protein synthesis leading to cell death. Moreover, CHOP induces oxidoreductin 1 (ERO1 α), an ER localized oxidoreductase that hyperoxidizes the ER environment and further commits cells to apoptosis (Li *et al.*, 2009). In addition, CHOP suppresses expression of anti-apoptotic B-cell lymphoma-2 (Bcl-2) protein (McCullough *et al.*, 2001) causing an increased production of reactive oxygen species and injury to the cell (Kim *et al.*, 2008). The sensitivity of the ER to oxidative stress revealed a link between ER and mitochondria, that are also physically associated at specific sites called mitochondrial-associated membranes that are likely sites of ER-stress-induced Ca⁺⁺ release from ER stores for uptake into mitochondria (Kaufman and Malhotra, 2014). The consequence of dumping ER Ca⁺⁺ into mitochondria initially activates oxidative phosphorylation to produce ATP, but prolonged or excessive Ca⁺⁺ loading leads to permeability transition of the inner mitochondrial membrane and activation of the intrinsic caspase cascade to cause apoptosis in a number of disease states (Cao and Kaufman, 2014).

IRE1 α is also implicated in cell death. Activated IRE1 α recruits the adaptor TRAF2 that signals through the stress-activated protein kinases ASK1 and JNK (Urano *et al.*, 2000; Nishitoh *et al.*, 2002).

Another link between IRE1 α and apoptosis is provided by the demonstration that hyperactivated IRE1 α causes relatively nonspecific RNA degradation that can lead to cell death (Ghosh *et al.*, 2014). In addition, thioredoxin-interacting protein (TXNIP) expression is induced by IRE1 α -mediated degradation of miR-17 that stabilizes TXNIP mRNA. Elevated levels of TXNIP activate the NLRP2 inflammasome and cause procaspase-1 cleavage and IL- β secretion. Accordingly, *Txnip* deletion reduced β -cell death during ER stress (Lemer *et al.*, 2012).

The Role of the UPR in Human Health and Disease

One of the remarkable findings regarding the UPR is that extracellular stimuli or fluctuations in intracellular homeostasis can disrupt the protein-folding environment of the ER (Figure 2). Therefore, through its unique ability to monitor the protein-folding status, UPR signaling senses perturbations

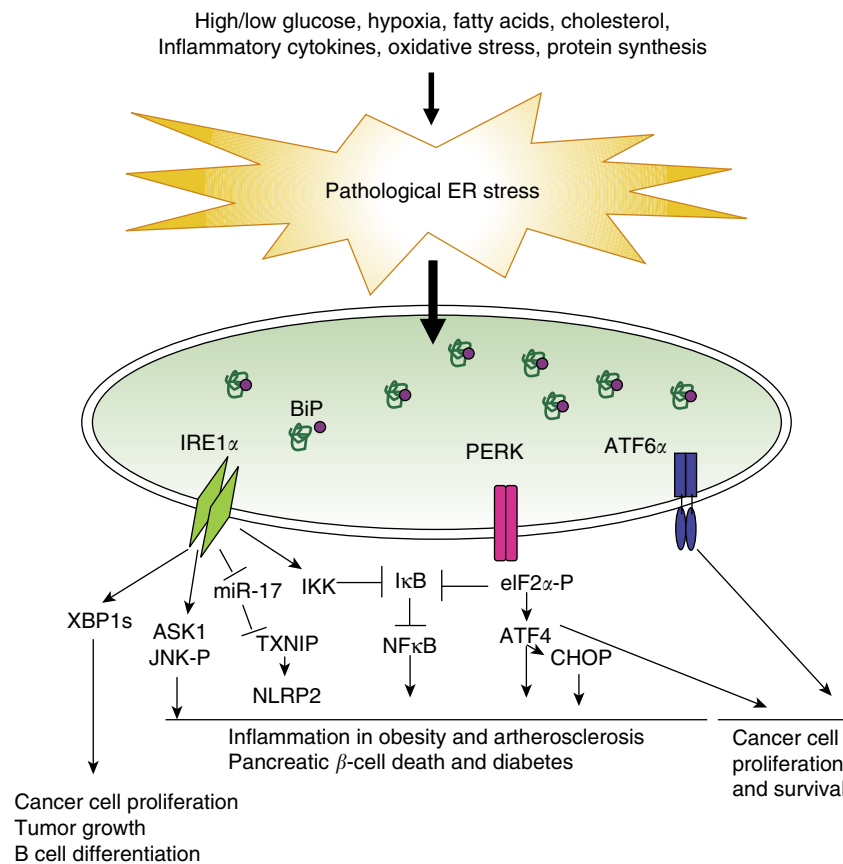


Figure 2 ER stress in disease. Pathophysiological conditions such as high/low glucose, hypoxia, high levels of fatty acids or cholesterol, inflammatory cytokines, and oxidative stress induce ER stress and activate the UPR chronically. The UPR activation is observed in a number of diseases including metabolic and inflammatory diseases, neurodegeneration, and cancer. The three UPR branches mediated by IRE1 α -XBP1, PERK-eIF2 α -phosphorylation-ATF4, and ATF6 α are important for tumor survival and growth under hypoxic conditions. IRE1 α and PERK also activate JNK and NF κ B to promote inflammation and apoptosis that contribute in inflammation in obesity and β -cell death in diabetes. PERK-mediated expression of CHOP contributes in aggravating diabetic states and atherosclerosis by exacerbating oxidative stress.

in cell homeostasis. Notably, pathological conditions, including metabolic disorders such as hyperlipidemia, hypercholesterolemia, hyperhomocysteinemia, hyperglycemia, inflammatory cytokines, and intracellular signaling such as the insulin anabolic response, pathogen infection, and tumorigenesis, all disrupt protein folding in the ER. Therefore, UPR activation is a hallmark of pathological conditions in human disease states, as well as in mouse models of human disease. Beside UPR activation, it was proposed that ER stress itself contributes to the pathology of many diseases (Schroder and Kaufman, 2005a; Kaufman, 2002; Wang and Kaufman, 2012).

The progress in the use of animal models of disease has greatly contributed to our knowledge regarding how the UPR impacts on normal physiology and disease pathogenesis. In addition, the identification of mutations in human genes often identifies or strengthens the casual relations between UPR defects and pathogenesis.

UPR Signaling May Contribute to Pancreatic β -Cell Degeneration and Diabetes

Pancreatic β -cell function is essential to maintain blood glucose homeostasis and its loss leads to diabetes, a chronic and degenerative disease. In normal individuals, high glucose stimulates insulin release from stored granules in β -cells and also increases proinsulin mRNA translation to replenish the pool of insulin granules. Circulating insulin stimulates glucose uptake in fat and muscle and reduces gluconeogenesis in the liver to maintain euglycemia. Therefore, proinsulin synthesis, folding and trafficking to secretory granules is coupled with insulin secretion in β -cells. Type 2 diabetes, i.e., adult onset, results from peripheral insulin resistance that increases levels of glucose, fatty acids, and inflammatory cytokines in the blood that can contribute to β -cell failure. Insulin resistance in fat and muscle causes the β -cell to increase insulin production. The increase in insulin production causes a load on the ER protein-folding capacity that is sufficient to initiate a series of events including proinsulin misfolding, insulin granule depletion, loss of glucose-stimulated insulin secretion, and oxidative stress leading to β -cell apoptosis (Laybutt et al., 2007; Huang et al., 2007).

The secretory function of β -cells depends on a functional UPR. The β -cell is unique because it is the only cell in the body that increases protein (mostly representing proinsulin) synthesis up to 10-fold upon glucose stimulation. As a consequence, several lines of evidence support the notion that the PERK-eIF2 arm is essential for β -cells to adapt to changes in protein synthesis (Harding et al., 2000b; Scheuner et al., 2001; Volchuk and Ron, 2010; Back et al., 2009). First, individuals afflicted with the autosomal recessive genetic disease called the Wolcott–Rallison syndrome require insulin at the age of 3 years. This syndrome is caused by loss-of-function mutations in PERK that cause β -cell failure. In addition, PERK deletion in mice recapitulates the β -cell failure observed in Wolcott–Rallison syndrome in humans (Harding et al., 2000b; Scheuner et al., 2001). Conditional deletion of PERK, i.e., deletion restricted to β -cells, further supports a homeostatic role for PERK in signaling β -cell survival (Cavener et al., 2010). In addition, mice with a Ser51Ala mutation at PERK phosphorylation

site in eIF2 α in a homozygous state, or in a heterozygous state combined with high fat diet, manifest β -cell loss due to proinsulin misfolding, ER stress, oxidative stress, and apoptosis (Back et al., 2009; Scheuner et al., 2005). Thus, PERK-mediated phosphorylation of eIF2 α is essential for normal β -cell physiology and loss of its function leads to β -cell failure and apoptosis and glucose intolerance.

β -cell failure in type 2 diabetes can also be exacerbated by pro-apoptotic components of the UPR, such as CHOP. Studies in mouse models of type 2 diabetes indicate that CHOP deletion not only protects β -cells from apoptosis but improves β -cell function by reducing oxidative stress and improving protein folding in the ER in diabetic *Akita* mice that express mutant Cys96Tyr misfolded proinsulin and in mice challenged with insulin resistance (Wang and Kaufman, 2012).

Other UPR components also contribute to the onset of type 2 diabetes. Inhibiting the IRE1 α -TXNIP axis by use of IRE1 α RNase inhibitors or deletion of *Txnip* reduces pancreatic β -cell death upon ER stress and suppresses diabetes in *Akita* mice (Lerner et al., 2012). Moreover, either deletion or increased expression of XBP1 in β -cells causes loss of β -cell homeostasis. β -cell specific XBP1 deletion in mice decreased the number of insulin granules, impaired proinsulin processing, blunted glucose-stimulated insulin secretion, and inhibited cell proliferation (Lee et al., 2011a). In addition, recent evidence suggests the defective IRE1 α and ATF6 α signaling accompany development of type 1 diabetes (Engin et al., 2014).

Additional recent findings point to a role for ATF6 α in protecting β -cells from ER stress. WFS1 (Wolfram syndrome 1 gene) encodes an ER transmembrane protein thought to regulate Ca⁺⁺ homeostasis. WFS1, a downstream target of XBP1s, is a negative regulator of ATF6 α and prevents β -cell death upon prolonged ATF6 α activation (Fonseca et al., 2010). Where mutations in WFS1 cause β -cell failure associated with Wolfram syndrome, polymorphisms in WFS1 are associated with impaired β -cell function and risk of type 2 diabetes (Sparso et al., 2008). In addition, some evidence suggests that ATF6 α variants are associated with type 2 diabetes (Wang and Kaufman, 2012). These findings provide a new link between ATF6 and protection from ER stress and β -cell death. However, ATF6 α deletion in the mouse did not dramatically interfere with β -cell function (Usui et al., 2012; Wu et al., 2007).

Metabolic Syndrome Activates UPR Signaling

Metabolic syndrome is associated with obesity, dyslipidemia (increase of circulating lipids), decrease of high-density lipoprotein (HDL) cholesterol, and hypertriglyceridemia that are often a cause of hypernutrition. The pathologies associated with metabolic syndrome are likely due to insulin resistance that develops when excessive lipids are deposited in insulin sensitive cells (hepatocytes and myocytes) instead than adipocytes, generating lipotoxicity. Accumulation of lipids in muscle generates insulin resistance, due to defective insulin signaling, with consequent inhibition of glucose uptake. In an effort to sustain hormone action, β -cells increase insulin production and this creates ER stress and β -cell death. The loss of β -cell function contributes to the onset of type 2 diabetes, that is, a complication in one-third of the patients afflicted with metabolic syndrome.

The increase in fatty acids or inability to oxidize excessive fatty acids, mainly a task of mitochondria, also causes lipid accumulation in the liver hepatocytes, called nonalcoholic fatty liver disease (Malhi and Kaufman, 2011). This pathology eventually gives rise to inflammation and fibrosis associated with nonalcoholic steatohepatitis and cirrhosis. Several reports provide evidence that the UPR is essential to maintain lipid homeostasis in the liver. Hepatic steatosis is alleviated by activation of UPR, as mimicked by overexpression of BiP, a downstream effector of the pathway. The UPR also inhibits activation of the cholesterol biosynthesis regulator SREBP-c and also improves glucose homeostasis (Kammoun *et al.*, 2009). Accordingly, selective inactivation of any single arm of UPR aggravated steatosis under pharmacologically induced ER stress, although it is not known whether this is due to increased hepatic lipogenesis or decreased lipoprotein secretion (Rutkowski *et al.*, 2008; Zhang *et al.*, 2011). Moreover, XBP1 liver-deleted mice manifest hypolipidemia, but neither hypoglycemia nor hyperglycemia suggesting that the regulatory role of XBP1 in hepatic metabolism is primarily to maintain lipid but not glucose homeostasis (So *et al.*, 2012). One study demonstrated that the IRE1 α /XBP1 pathway is essential to maintain functional levels of PDI, which is a cofactor for microsomal triglyceride transfer protein, to promote triglyceride assembly onto apolipoprotein B during assembly of very low-density lipoprotein (VLDL) (Wang *et al.*, 2012).

Finally, the discovery of cAMP responsive element-binding protein, hepatocyte specific (CREBH), a liver-specific transcription factor, for the first time provided a molecular link between ER stress and the innate systemic inflammatory response (Zhang *et al.*, 2006). CREBH expression is induced at a transcriptional level by inflammatory cytokines such as TNF, interleukin 1 β (IL-1 β), and IL-6. CREBH is similar to ATF6 α being an ER transmembrane transcription factor of the ATF family (Figure 1). In response to ER stress, CREBH traffics to the Golgi and is cleaved by S1P and S2P, as previously described for ATF6 and SREBP, to generate an N-terminal soluble bZIP transcription factor. The nuclear form of CREBH activates the expression of a subset of the acute phase response (APR) genes, including those encoding serum amyloid P component and C-reactive protein but has only minor effects on the induction of UPR genes. These studies highlight that ER stress can initiate a proinflammatory response. These findings are important because inflammation contributes to the atherosclerotic process and also to the pathogenesis of cardiovascular diseases, among which are complications associated with metabolic syndrome. The findings put ER function in a central position to coordinate metabolism with inflammation.

Further, CREBH is also a regulator of lipid homeostasis (Zhang *et al.*, 2011). CREBH was demonstrated to regulate hepatic lipogenesis, fatty acid oxidation, and lipolysis under conditions of metabolic stress and control the activity of a lipoprotein lipase regulating VLDL-triglyceride clearance in the plasma through up-regulation of its coactivator apolipoproteins C2, A4, and A5 and down-regulation of its inhibitor Apoc3. Consistently, nonsynonymous hypomorphic mutations in human CREBH are associated with severe hypertriglyceridemia in humans (Lee *et al.*, 2011b). Moreover, the UPR is activated during adipocyte differentiation, a crucial step in body weight gain. Both eIF2 α phosphorylation and splicing

of XBP1 mRNA were detected (Han *et al.*, 2013b). The IRE1 α -XBP1 axis is required for the expression of the key adipogenic factor C/EBP α . ER stress and UPR signaling appears to be a stimulus for adipogenesis.

The UPR Provides a Protective Role in Infectious and Inflammatory Disease

The UPR plays a role in pathogen infection and in the immune response. During viral infection IRE1 α and PERK are activated by the high-level expression of viral glycoproteins. Misfolded viral proteins induce transcription of BiP. Moreover, PERK and PKR-mediated eIF2 α phosphorylation is required to repress viral proteins synthesis, but viruses have evolved mechanisms to prevent the effect of eIF2 α phosphorylation to promote their replication (Cheng *et al.*, 2005). Every interferon resistant virus has evolved a unique and selective mechanism to bypass the inhibitory impact of PKR-mediated phosphorylation of eIF2 α , which is a major component of the interferon antiviral response.

Activated IRE1 α binds to TRAF2 and this complex recruits I κ B-kinase (IKK) that phosphorylates I κ B leading to its degradation and nuclear translocation of NF- κ B that induces the production of inflammatory cytokines. IRE1 α is important to integrate ER stress with inflammatory response signaling (Zhang and Kaufman, 2008). In addition, signaling through toll-like receptors can recruit TRAF6 to cause activation and subsequent degradation of IRE1 α , independent of ER stress (Qiu *et al.*, 2013).

The differentiation of B lymphocytes into plasma cells, that actively secrete high levels of antibodies during the humoral immune response, is coupled to expansion of ER and of the secretory pathway. In this process, XBP1s plays an essential role. In contrast, PERK-eIF2 α and ATF6 α arms are not required in differentiation of mature B cells into plasma cells (Zhang *et al.*, 2005).

UPR Signaling Provides a Double-Edged Sword in Cancer

Tumors result from a complex, multistep process involving progressive cellular degeneration and selection for cells that can propagate under unfavorable conditions. Recent studies suggest protein misfolding in the ER may be an initiating event in tumorigenesis (Nakagawa *et al.*, 2014). Due to poor vascularization, tumor cells proliferate under conditions of low oxygen and nutrients. As a consequence tumor cell proliferation requires active UPR signaling (Wang and Kaufman, 2014). Anaerobic metabolism prevails and pro-survival responses are activated in cancer cells.

Several studies support a role for the UPR in survival of cancer cells to extreme environmental stress (Clarke *et al.*, 2014; Wang and Kaufman, 2014). In particular, PERK-eIF2 α -ATF4 signaling pathway is critical under hypoxic conditions. The deletion of PERK, mutations that inactivate the kinase domain and phosphorylation-resistant forms of eIF2 α , all impair cell survival under hypoxic conditions. PERK also promotes cancer cell proliferation by limiting oxidative DNA damage through ATF4. Moreover, ATF4 is overexpressed in many solid tumors consistent with its role in regulating the expression of pro-survival genes.

The IRE1 α -XBP1 arm also contributes to growth and survival under hypoxia conditions. In a mouse glioma model, it has been shown that inhibition of IRE1 α reduces tumor growth and angiogenesis (Auf *et al.*, 2010). In addition, deletion of XBP1 increased sensitivity to hypoxia-induced cell death and reduced solid tumor formation (Fujimoto *et al.*, 2007). Moreover, since spliced XBP1 is essential for plasma cell differentiation, there has been interest to develop IRE1 α inhibitors for multiple myeloma, a cancer formed by malignant plasma cells. XBP1s is overexpressed in plasma cells derived from myeloma and this may be associated with the physiological role of XBP1s in promoting secretory functions during plasma cell differentiation. Thus, XBP1s may be a critical component in the induction of multiple myeloma. For this reason, studies have tested the therapeutic potential of IRE1 α RNase activity inhibitors (Papandreou *et al.*, 2011; Mimura *et al.*, 2012), although the results suggest a therapeutic benefit, there are concerns about general toxicity to other cell types. At present recommended treatment of multiple myeloma involves the use of proteasome inhibitors which induce death of multiple myeloma cells that accumulate misfolded immunoglobulin chains that cannot be degraded (Dimopoulos *et al.*, 2011). Unfortunately, resistance to proteasome inhibitors frequently occurs and this was associated with selection for cells that lose immunoglobulin expression and display preplasmablast characteristics (Leung-Hagesteijn *et al.*, 2013).

The studies of the role of the ATF6 arm of UPR in cancer are still limited but indicate that ATF6 α interacts with the other signaling pathways of UPR in promoting survival of tumor cells. Notably, ATF6 and XBP1 are responsible of up-regulation of BiP expression in hepatocarcinogenesis. Elevated expression of BiP is present in breast cancer, lung cancer, prostate cancer and other malignancies. The increase in BiP level plays a major role in pro-survival and cytoprotective response of cancer cells to environmental stress. The mechanism is through inhibition of apoptosis.

On the other hand, other studies indicate that alterations of UPR can also be involved in cellular apoptosis, inflammation leading to precancerous state especially in hepatocarcinogenesis. An indication is provided by single nucleotide polymorphisms in the ATF6 α gene that correlate with a decrease in ATF6 α mRNA and ATF6 α gene target contributing to hepatocellular carcinoma (Wu *et al.*, 2014). Moreover, *Chop*-null mutation promotes hepatic inflammatory gene expression and hepatic carcinogenesis. Thus, UPR can be involved also at the early steps of carcinogenesis (Dezwaan-McCabe *et al.*, 2013). Expression of BiP, a major ATF6 α target encoding an ER chaperone, not only correlates with cancer cell proliferation and histological grade, but also predicts response to therapies and prognosis (Pfaffenbach and Lee, 2011). Indeed, approaches to target BiP expression are under evaluation in the clinic (Rosenes *et al.*, 2013).

Targeting the UPR in cancer probably is a double-edged sword. On one hand, activation of the UPR is probably important in providing surveillance system to eliminate by apoptosis cells with mutations that might be oncogenic. On the other hand, tumor progression likely relies on UPR signaling to survive harsh conditions in the microenvironment. In conclusion, components of the UPR arms, including BiP and ERAD, are promising targets of new antitumor drugs. New

therapeutic strategies aim to exacerbate ER stress by inhibiting UPR and divert tumor cells from pro-survival to apoptotic pathways. Thus, the dualistic nature of the UPR may be exploited in developing new anticancer therapies (Wang and Kaufman, 2014).

UPR Activation Is Associated with Neurodegenerative Diseases

Whereas the role of the UPR in protein secretion and metabolic homeostasis is well characterized in nonneuronal cells, recent reports have documented a potential key role of the UPR in neuronal function (Wang and Kaufman, 2012). For example, XBP1 was recognized to regulate the induction of GABAergic markers to control the neurite growth. Translational control of ATF4 by GCN2-eIF2 α phosphorylation appears important for hippocampal synaptic plasticity and memory. eIF2 α phosphorylation appears required for learning and memory. Several studies document a requirement for protein synthesis for long-term potentiation (LTP) as well and mGluR1 long-term depression (LTD) (Costa-Mattioli *et al.*, 2007; Sidrauski *et al.*, 2013). Unlike LTD induced by stimulation of ionotropic glutamate receptors (N-methyl-D-aspartate receptor (NMDAR), LTD) induced by activation of metabotropic glutamate receptors (mGluRs) requires protein synthesis (Huber *et al.*, 2000). Alterations in mGluR-LTD have been linked to drug addiction, Alzheimer's disease, and intellectual disability (Luscher and Huber, 2010). Recent findings indicate that eIF2 α phosphorylation increases translation of selective mRNAs, one of which encodes OPHN1, a RhoGAP that stimulates synaptic vesicle endocytosis and is required for stabilization of dendritic spines (Di Prisco *et al.*, 2014). OPHN1 has two open reading frames within the 5' end of its mRNA that increase translation of OPHN1 upon phosphorylation of eIF2 α . Mutations in OPHN1 cause X-linked mental retardation in humans and its deletion in mice impairs spatial memory (Billuart *et al.*, 1998; Khelifaoui *et al.*, 2007). These studies suggest that inhibition of particular eIF2 α kinases may increase cognitive function.

All neurodegenerative diseases are associated with protein aggregation and UPR activation has been observed in a number of these diseases which include amyotrophic lateral sclerosis, Parkinson's disease, Huntington's disease, prion-related disorders, Alzheimer's disease, and demyelinating autoimmune diseases such as multiple sclerosis, retinal neuronal degeneration, and the list is not exhaustive. In particular, these studies have highlighted specific and even opposite effects of distinct UPR components, indicating that UPR importance in neurodegeneration depends on the disease context (see as a review, Hetz and Mollereau, 2014).

Where reduced eIF2 α phosphorylation may increase cognitive function, recent studies suggest that increased eIF2 α phosphorylation may reduce pathogenesis due to reduced synthesis of aggregation-prone proteins (Kim *et al.*, 2014; Ma *et al.*, 2013; Tribouillard-Tanvier *et al.*, 2008; Wang *et al.*, 2014). There is a need to identify under which conditions increased eIF2 α phosphorylation will increase cognitive function and those for which increased phosphorylation may reduce pathologies associated with expression of pathogenic

proteins. Importantly, inhibitors of GADD34 can protect from pathologies associated with protein aggregation associated with amyotrophic lateral sclerosis. Guanabenz is an antihypertensive agent that was shown to prevent pathogenicity associated with prion diseases and to inhibit GADD34 interaction with PP1 and prevent eIF2 α dephosphorylation, thereby reducing protein synthesis (Tsaytler *et al.*, 2011). Based on these findings, it seems that structural alterations to guanabenz to increase GADD34 interaction and reduce alpha adrenergic agonist activity may be encouraged.

Conclusions

The UPR is essential for the normal cell physiology and for adaptation to ER stress. The understanding of how the UPR intersects and cross-talks under normal conditions and in response to ER stress, and how ER function impacts other organelles is still the subject of intense research. Our increased understanding of the mechanisms that signal the UPR pathways has generated a wealth of knowledge, but there remain major unanswered questions. First and utmost, under what conditions is activation of any particular UPR signaling pathway beneficial or detrimental to any disease process? With an answer to this complex question, specific therapies may be envisioned for particular disease states, such as diabetes and cancer, and may provide in the future new strategies for other chronic and degenerative diseases. The dualistic nature of the UPR may be carefully explored to either inhibit the signaling to prevent cell death in response to ER stress or exacerbate it to prevent survival as in the case of tumor cells.

Acknowledgments

RJK is supported by NIH grants DK042394, DK088227, and HL052173.

See also: Nucleic Acid Synthesis/breakdown: Transcription: Eukaryotic Transcriptional Regulation. Protein synthesis/degradation: Protein Degradation – Intracellular: Endoplasmic Reticulum-Associated Degradation and Protein Quality Control. Protein Synthesis/Degradation: Protein Degradation – Pathological Aspects: Alpha-1-Antitrypsin Deficiency: A Misfolded Secretory Glycoprotein Damages the Liver by Proteotoxicity and Its Reduced Secretion Predisposes to Emphysematous Lung Disease Because of Protease-Inhibitor Imbalance

References

- Auf, G., Jabouille, A., Guert, S., *et al.*, 2010. Inositol-requiring enzyme 1 α is a key regulator of angiogenesis and invasion in malignant glioma. *Proceedings of the National Academy of Sciences of the United States of America* 107, 15553–15558.
- Back, S.H., Scheuner, D., Han, J., *et al.*, 2009. Translation attenuation through eIF2 α phosphorylation prevents oxidative stress and maintains the differentiated state in beta cells. *Cell Metabolism* 10, 13–26.
- Billuart, P., Bienvenu, T., Ronce, N., *et al.*, 1998. Oligophrenin-1 encodes a rhoGAP protein involved in X-linked mental retardation. *Nature* 392, 923–926.
- Buchberger, A., 2014. ERQC autophagy: Yet another way to die. *Molecular Cell* 54, 3–4.
- Cao, S.S., Kaufman, R.J., 2014. Endoplasmic reticulum stress and oxidative stress in cell fate decision and human disease. *Antioxidants & Redox Signaling* 21 (3), 396–413.
- Cavener, D.R., Gupta, S., McGrath, B.C., 2010. PERK in beta cell biology and insulin biogenesis. *Trends in Endocrinology & Metabolism* 21, 714–721.
- Cheng, G., Feng, Z., He, B., 2005. Herpes simplex virus 1 infection activates the endoplasmic reticulum resident kinase PERK and mediates eIF2 α dephosphorylation by the gamma(1)34.5 protein. *Journal of Virology* 79, 1379–1388.
- Chitnis, N., Pytel, D., Diehl, J.A., 2013. UPR-inducible miRNAs contribute to stressful situations. *Trends in Biochemical Sciences* 38, 447–452.
- Clarke, H.J., Chambers, J.E., Liniker, E., Marciniak, S.J., 2014. Endoplasmic reticulum stress in malignancy. *Cancer Cell* 25, 563–573.
- Coelho, D.S., Domingos, P.M., 2014. Physiological roles of regulated Ire1 dependent decay. *Frontiers in Genetics* 5, 76.
- Costa-Mattioli, M., Gobert, D., Stern, E., *et al.*, 2007. eIF2 α phosphorylation bidirectionally regulates the switch from short- to long-term synaptic plasticity and memory. *Cell* 129, 195–206.
- Cross, B.C., Bond, P.J., Sadowski, P.G., *et al.*, 2012. The molecular basis for selective inhibition of unconventional mRNA splicing by an IRE1-binding small molecule. *Proceedings of the National Academy of Sciences of the United States of America* 109 (15), E869–E878.
- Dezwaan-McCabe, D., Riordan, J.D., Arensdorf, A.M., *et al.*, 2013. The stress-regulated transcription factor CHOP promotes hepatic inflammatory gene expression, fibrosis, and oncogenesis. *PLoS Genetics* 9, e1003937.
- Di Prisco, G.V., Huang, W., Buffington, S.A., *et al.*, 2014. Translational control of mGluR-dependent long-term depression and object-place learning by eIF2 α . *Nature Neuroscience* 17 (8), 1073–1082.
- Dimopoulos, M.A., San-Miguel, J.F., Anderson, K.C., 2011. Emerging therapies for the treatment of relapsed or refractory multiple myeloma. *European Journal of Haematology* 86, 1–15.
- Dorner, A.J., Bole, D.G., Kaufman, R.J., 1987. The relationship of N-linked glycosylation and heavy chain-binding protein association with the secretion of glycoproteins. *Journal of cell biology* 105, 2665–2674.
- Dorner, A.J., Wasley, L.C., Kaufman, R.J., 1989. Increased synthesis of secreted proteins induces expression of glucose-regulated proteins in butyrate-treated Chinese hamster ovary cells. *Journal of Biological Chemistry* 264, 20602–20607.
- Engin, F., Nguyen, T., Yermalovich, A., Hotamisliligil, G.S., 2014. Aberrant islet unfolded protein response in type 2 diabetes. *Scientific Reports* 4, 4054.
- Ferris, S.P., Kodali, V.K., Kaufman, R.J., 2014. Glycoprotein folding and quality-control mechanisms in protein-folding diseases. *Disease Models & Mechanisms* 7, 331–341.
- Fonseca, S.G., Ishigaki, S., Oslowski, C.M., *et al.*, 2010. Wolfram syndrome 1 gene negatively regulates ER stress signaling in rodent and human cells. *Journal of Clinical Investigation* 120, 744–755.
- Fujimoto, T., Yoshimatsu, K., Watanabe, K., *et al.*, 2007. Overexpression of human X-box binding protein 1 (XBP-1) in colorectal adenomas and adenocarcinomas. *Anticancer Research* 27, 127–131.
- Ghosh, R., Wang, L., Wang, E.S., *et al.*, 2014. Allosteric inhibition of the IRE1 α RNase preserves cell viability and function during endoplasmic reticulum stress. *Cell* 158 (3), 534–548.
- Hampton, R.Y., Sommer, T., 2012. Finding the will and the way of ERAD substrate retrotranslocation. *Current Opinion in Cell Biology* 24, 460–466.
- Han, J., Back, S.H., Hur, J., *et al.*, 2013a. ER-stress-induced transcriptional regulation increases protein synthesis leading to cell death. *Nature Cell Biology* 15, 481–490.
- Han, J., Murthy, R., Wood, B., *et al.*, 2013b. ER stress signalling through eIF2 α and CHOP, but not IRE1 α , attenuates adipogenesis in mice. *Diabetologia* 56 (4), 911–924.
- Harding, H.P., Novoa, I., Zhang, Y., *et al.*, 2000a. Regulated translation initiation controls stress-induced gene expression in mammalian cells. *Molecular Cell* 6, 1099–1108.
- Harding, H.P., Zhang, Y., Bertolotti, A., Zeng, H., Ron, D., 2000b. Perk is essential for translational regulation and cell survival during the unfolded protein response. *Molecular Cell* 5, 897–904.
- Harding, H.P., Zhang, Y., Ron, D., 1999. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature* 397, 271–274.
- Harding, H.P., Zhang, Y., Zeng, H., *et al.*, 2003. An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Molecular Cell* 11, 619–633.
- Hetz, C., Mollereau, B., 2014. Disturbance of endoplasmic reticulum proteostasis in neurodegenerative diseases. *Nature Reviews Neuroscience* 15, 233–249.

- Huang, C.J., Lin, C.Y., Haataja, L., *et al.*, 2007. High expression rates of human islet amyloid polypeptide induce endoplasmic reticulum stress mediated beta-cell apoptosis, a characteristic of humans with type 2 but not type 1 diabetes. *Diabetes* 56, 2016–2027.
- Huber, K.M., Kayser, M.S., Bear, M.F., 2000. Role for rapid dendritic protein synthesis in hippocampal mGluR-dependent long-term depression. *Science* 288, 1254–1257.
- Kammoun, H.L., Chabanon, H., Hainault, I., *et al.*, 2009. GRP78 expression inhibits insulin and ER stress-induced SREBP-1c activation and reduces hepatic steatosis in mice. *Journal of Clinical Investigation* 119, 1201–1215.
- Kaufman, R.J., 2002. Orchestrating the unfolded protein response in health and disease. *Journal of Clinical Investigation* 110, 1389–1398.
- Kaufman, R.J., Malhotra, J.D., 2014. Calcium trafficking integrates endoplasmic reticulum function with mitochondrial bioenergetics. *Biochimica et Biophysica Acta* 1843 (10), 2233–2239.
- Kaufman, R.J., Scheuner, D., Schroder, M., *et al.*, 2002. The unfolded protein response in nutrient sensing and differentiation. *Nature Reviews Molecular Cell Biology* 3, 411–421.
- Khelifaoui, M., Denis, C., Van Galen, E., *et al.*, 2007. Loss of X-linked mental retardation gene oligophrenin1 in mice impairs spatial memory and leads to ventricular enlargement and dendritic spine immaturity. *Journal of Neuroscience* 27, 9439–9450.
- Kim, H.J., Raphael, A.R., Ladow, E.S., *et al.*, 2014. Therapeutic modulation of eIF2alpha phosphorylation rescues TDP-43 toxicity in amyotrophic lateral sclerosis disease models. *Nature Genetics* 46, 152–160.
- Kim, J., Choi, T.G., Ding, Y., *et al.*, 2008. Overexpressed cyclophilin B suppresses apoptosis associated with ROS and Ca²⁺ homeostasis after ER stress. *Journal of Cell Science* 121, 3636–3648.
- Kimata, Y., Kohno, K., 2011. Endoplasmic reticulum stress-sensing mechanisms in yeast and mammalian cells. *Current Opinion in Cell Biology* 23, 135–142.
- Kozutsumi, Y., Segal, M., Normington, K., Gething, M.J., Sambrook, J., 1988. The presence of misfolded proteins in the endoplasmic reticulum signals the induction of glucose-regulated proteins. *Nature* 332, 462–464.
- Laybutt, D.R., Preston, A.M., Akerfeldt, M.C., *et al.*, 2007. Endoplasmic reticulum stress contributes to beta cell apoptosis in type 2 diabetes. *Diabetologia* 50, 752–763.
- Lee, A.H., Heidman, K., Hotamisligil, G.S., Glimcher, L.H., 2011a. Dual and opposing roles of the unfolded protein response regulated by IRE1alpha and XBP1 in proinsulin processing and insulin secretion. *Proceedings of the National Academy of Sciences of the United States of America* 108, 8885–8890.
- Lee, J.H., Giannikopoulos, P., Duncan, S.A., *et al.*, 2011b. The transcription factor cyclic AMP-responsive element-binding protein H regulates triglyceride metabolism. *Nature Medicine* 17, 812–815.
- Lerner, A.G., Upton, J.P., Praveen, P.V., *et al.*, 2012. IRE1alpha induces thioredoxin-interacting protein to activate the NLRP3 inflammasome and promote programmed cell death under irremediable ER stress. *Cell Metabolism* 16, 250–264.
- Leung-Hagsteström, C., Erdmann, N., Cheung, G., *et al.*, 2013. Xbp1s-negative tumor B cells and pre-plasmablasts mediate therapeutic proteasome inhibitor resistance in multiple myeloma. *Cancer Cell* 24, 289–304.
- Li, G., Mongillo, M., Chin, K.T., *et al.*, 2009. Role of ERO1- α -mediated stimulation of inositol 1,4,5-triphosphate receptor activity in endoplasmic reticulum stress-induced apoptosis. *Journal of Cell Biology* 186, 783–792.
- Logue, S.E., Cleary, P., Saveljeva, S., Samali, A., 2013. New directions in ER stress-induced cell death. *Apoptosis* 18, 537–546.
- Lu, M., Lawrence, D.A., Marsters, S., *et al.*, 2014. Cell death. Opposing unfolded-protein-response signals converge on death receptor 5 to control apoptosis. *Science* 345, 98–101.
- Lu, P.D., Jousse, C., Marciniak, S.J., *et al.*, 2004. Cytoprotection by pre-emptive conditional phosphorylation of translation initiation factor 2. *EMBO Journal* 23, 169–179.
- Luscher, C., Huber, K.M., 2010. Group 1 mGluR-dependent synaptic long-term depression: Mechanisms and implications for circuitry and disease. *Neuron* 65, 445–459.
- Ma, T., Trinh, M.A., Wexler, A.J., *et al.*, 2013. Suppression of eIF2alpha kinases alleviates Alzheimer's disease-related plasticity and memory deficits. *Nature Neuroscience* 16, 1299–1305.
- Malhi, H., Kaufman, R.J., 2011. Endoplasmic reticulum stress in liver disease. *Journal of Hepatology* 54, 795–809.
- McCullough, K.D., Martindale, J.L., Klotz, L.O., AW, T.Y., Holbrook, N.J., 2001. Gadd153 sensitizes cells to endoplasmic reticulum stress by down-regulating Bcl2 and perturbing the cellular redox state. *Molecular and Cellular Biology* 21, 1249–1259.
- Mimura, N., Fulciniti, M., Gorgun, G., *et al.*, 2012. Blockade of XBP1 splicing by inhibition of IRE1alpha is a promising therapeutic option in multiple myeloma. *Blood* 119 (24), 5772–5781.
- Moore, K.A., Hollien, J., 2012. The unfolded protein response in secretory cell function. *Annual Review of Genetics* 46, 165–183.
- Nakagawa, H., Umemura, A., Taniguchi, K., *et al.*, 2014. ER stress cooperates with hypernutrition to trigger TNF-dependent spontaneous HCC development. *Cancer Cell* 26, 331–343.
- Nishitoh, H., Matsuzawa, A., Tobiume, K., *et al.*, 2002. ASK1 is essential for endoplasmic reticulum stress-induced neuronal cell death triggered by expanded polyglutamine repeats. *Genes & Development* 16, 1345–1355.
- Papandreou, I., Denko, N.C., Olson, M., *et al.*, 2011. Identification of an IRE1alpha endonuclease specific inhibitor with cytotoxic activity against human multiple myeloma. *Blood* 117, 1311–1314.
- Pfaffenbach, K.T., Lee, A.S., 2011. The critical role of GRP78 in physiologic and pathologic stress. *Current Opinion in Cell Biology* 23, 150–156.
- Qiu, Q., Zheng, Z., Chang, L., *et al.*, 2013. Toll-like receptor-mediated IRE1alpha activation as a therapeutic target for inflammatory arthritis. *EMBO Journal* 32, 2477–2490.
- Rosenes, Z., Mok, Y.F., Yang, S., *et al.*, 2013. Simultaneous binding of the anti-cancer IgM monoclonal antibody PAT-SM6 to low density lipoproteins and GRP78. *PLoS One* 8, e61239.
- Rutkowski, D.T., Arnold, S.M., Miller, C.N., *et al.*, 2006. Adaptation to ER stress is mediated by differential stabilities of pro-survival and pro-apoptotic mRNAs and proteins. *PLoS Biology* 4, e374.
- Rutkowski, D.T., Wu, J., Back, S.H., *et al.*, 2008. UPR pathways combine to prevent hepatic steatosis caused by ER stress-mediated suppression of transcriptional master regulators. *Developmental Cell* 15, 829–840.
- Sakaki, K., Yoshina, S., Shen, X., *et al.*, 2012. RNA surveillance is required for endoplasmic reticulum homeostasis. *Proceedings of the National Academy of Sciences of the United States of America* 109, 8079–8084.
- Scheuner, D., Song, B., McEwen, E., *et al.*, 2014. Translational control is required for the unfolded protein response and *in vivo* glucose homeostasis. *Molecular Cell* 7, 1165–1176.
- Scheuner, D., Vander Mierde, D., Song, B., *et al.*, 2005. Control of mRNA translation preserves endoplasmic reticulum function in beta cells and maintains glucose homeostasis. *Nature Medicine* 11, 757–764.
- Schroder, M., Kaufman, R.J., 2005a. ER stress and the unfolded protein response. *Mutation Research* 569, 29–63.
- Schroder, M., Kaufman, R.J., 2005b. The mammalian unfolded protein response. *Annual Review of Biochemistry* 74, 739–789.
- Shen, J., Prywes, R., 2005. ER stress signaling by regulated proteolysis of ATF6. *Methods* 35, 382–389.
- Shen, X., Ellis, R.E., Sakaki, K., Kaufman, R.J., 2005. Genetic interactions due to constitutive and inducible gene regulation mediated by the unfolded protein response in *C. elegans*. *PLoS Genetics* 1, e37.
- Sidrauski, C., Acosta-Alvear, D., Khoutorsky, A., *et al.*, 2013. Pharmacological brake-release of mRNA translation enhances cognitive memory. *Elife* 2, e00498.
- So, J.S., Hur, K.Y., Tarrio, M., *et al.*, 2012. Silencing of lipid metabolism genes through IRE1alpha-mediated mRNA decay lowers plasma lipids in mice. *Cell Metabolism* 16, 487–499.
- Sparso, T., Andersen, G., Nielsen, T., *et al.*, 2008. The GCKR rs780094 polymorphism is associated with elevated fasting serum triacylglycerol, reduced fasting and OGTT-related insulinaemia, and reduced risk of type 2 diabetes. *Diabetologia* 51, 70–75.
- Tabas, I., Ron, D., 2011. Integrating the mechanisms of apoptosis induced by endoplasmic reticulum stress. *Nature Cell Biology* 13, 184–190.
- Tribouillard-Tanvier, D., Beringue, V., Desban, N., *et al.*, 2008. Antihypertensive drug guanabenz is active *in vivo* against both yeast and mammalian prions. *PLoS One* 3, e1981.
- Tsaytler, P., Harding, H.P., Ron, D., Bertolotti, A., 2011. Selective inhibition of a regulatory subunit of protein phosphatase 1 restores proteostasis. *Science* 332, 91–94.
- Urano, F., Wang, X., Bertolotti, A., *et al.*, 2000. Coupling of stress in the ER to activation of JNK protein kinases by transmembrane protein kinase IRE1. *Science* 287, 664–666.
- Usui, M., Yamaguchi, S., Tanji, Y., *et al.*, 2012. Atf6alpha-null mice are glucose intolerant due to pancreatic beta-cell failure on a high-fat diet but partially resistant to diet-induced insulin resistance. *Metabolism* 61, 1118–1128.
- Volchuk, A., Ron, D., 2010. The endoplasmic reticulum stress response in the pancreatic beta-cell. *Diabetes, Obesity and Metabolism* 12 (Suppl. 2), 48–57.

- Walter, P., Ron, D., 2011. The unfolded protein response: From stress pathway to homeostatic regulation. *Science* 334, 1081–1086.
- Wang, L., Popko, B., Roos, R.P., 2014. An enhanced integrated stress response ameliorates mutant SOD1-induced ALS. *Human Molecular Genetics* 23, 2629–2638.
- Wang, M., Kaufman, R.J., 2014. The impact of the endoplasmic reticulum protein-folding environment on cancer development. *Nature Reviews Cancer* 14, 581–597.
- Wang, S., Chen, Z., Lam, V., *et al.*, 2012. IRE1alpha-XBP1s induces PDI expression to increase MTP activity for hepatic VLDL assembly and lipid homeostasis. *Cell Metabolism* 16, 473–486.
- Wang, S., Kaufman, R.J., 2012. The impact of the unfolded protein response on human disease. *Journal of Cell Biology* 197, 857–867.
- Wu, J., Rutkowski, D.T., Dubois, M., *et al.*, 2007. ATF6alpha optimizes long-term endoplasmic reticulum function to protect cells from chronic stress. *Developmental Cell* 13, 351–364.
- Wu, X., Xin, Z., Zhang, W., *et al.*, 2014. A missense polymorphism in ATF6 gene is associated with susceptibility to hepatocellular carcinoma probably by altering ATF6 level. *International Journal of Cancer* 135, 61–68.
- Yamamoto, K., Sato, T., Matsui, T., *et al.*, 2007. Transcriptional induction of mammalian ER quality control proteins is mediated by single or combined action of ATF6alpha and XBP1. *Developmental Cell* 13, 365–376.
- Zhang, K., Kaufman, R.J., 2008. From endoplasmic-reticulum stress to the inflammatory response. *Nature* 454, 455–462.
- Zhang, K., Shen, X., Wu, J., *et al.*, 2006. Endoplasmic reticulum stress activates cleavage of CREBH to induce a systemic inflammatory response. *Cell* 124, 587–599.
- Zhang, K., Wang, S., Malhotra, J., *et al.*, 2011. The unfolded protein response transducer IRE1alpha prevents ER stress-induced hepatic steatosis. *EMBO Journal* 30, 1357–1375.
- Zhang, K., Wong, H.N., Song, B., *et al.*, 2005. The unfolded protein response sensor IRE1alpha is required at 2 distinct steps in B cell lymphopoiesis. *Journal of Clinical Investigation* 115, 268–281.

Further Reading

- Bussey, H., Andrews, B., Boone, C., 2006. From worm genetic networks to complex human diseases. *Nature Genetics* 18, 862–863.
- Kanekura, K., Ishigaki, S., Merksamer, P., Papa, F.R., Urano, F., 2013. Establishment of a system of monitoring endoplasmic reticulum redox state in mammalian cells. *Laboratory Investigation* 93, 1254–1258.
- Valastyan, J.S., Lindquist, S., 2014. Mechanisms of protein-folding diseases at a glance. *Disease Models & Mechanisms* 7, 9–14.

CYTOSKELETON AND MOTORS: CYTOSKELETAL COMPONENTS

Contents

Microtubules and Microtubule-Associated Proteins (MAPs)

Actin Assembly Dynamics and Its Regulation in Motile and Morphogenetic Processes

Intermediate Filaments

Septins: Cytoskeletal Filaments with Structural and Regulatory Functions

Bacterial and Archaeal Cytoskeletons

Myosins

Kinesin Superfamily Proteins (KIFs) as a Fundamental Component of Life: Intracellular Transport and Beyond

Dyneins

Microtubules and Microtubule-Associated Proteins (MAPs)

A Straube, University of Warwick, Coventry, UK

© 2016 Elsevier Inc. All rights reserved.

Glossary

A-lattice Microtubule lattice arrangement in a 3-start helix where protofilaments are effectively staggered by 4.9 nm. Beta-tubulin makes lateral contacts to alpha-tubulin at every protofilament–protofilament interface.

B-lattice Microtubule lattice arrangement where protofilaments are staggered by 0.9 nm. In a 13-protofilament 3-start helix, this results in a discontinuity at the seam. Perfect helical B-lattice microtubules can be formed from 15 protofilaments. A 13-protofilament 3-start helix with B-lattice arrangement is the usual conformation of microtubules found in most cells.

Catastrophe The transition from assembly (growth) to disassembly (shrinkage) at a microtubule end.

Dynamic instability Stochastic switching between phases of assembly and disassembly at microtubule ends.

Microtubule organizing center (MTOC) MTOC, such as the centrosome in animal cells or the spindle pole body in yeast, is the main focus of microtubule minus ends in the cell and point of highest nucleation activity.

Minus end Less dynamic end of the microtubule, embedded in the centrosome/spindle pole in most cells,

exposes alpha-tubulin or bound to gamma-tubulin ring complex.

Nucleation Initiation of *de novo* formation of microtubules occurs at a low rate by spontaneous aggregation of tubulin subunits, facilitated in cells by the gamma-tubulin ring complex, which has a 13-fold symmetry and acts as a template for the addition of alpha/beta-tubulin dimers.

Pause Term used for a state of little or no observable length change of a microtubule.

Plus end Dynamic end of a microtubule that grows faster than the minus end and exposes beta-tubulin.

Rescue The transition from disassembly (shrinkage) to reassembly (growth) at a microtubule end.

Seam Contact between the 13th and 1st protofilament in a B-lattice microtubule where neighboring protofilaments are offset by 4.9 nm instead of 0.9 nm as for the other protofilament–protofilament interfaces. This results in beta-tubulin making lateral contacts to alpha-tubulin at the seam.

+TIP or tip tracker Microtubule-associated protein that localizes to the terminal region of growing microtubules.

Microtubule Structure

Microtubules are dynamic polymers of 8 nm long, heterodimeric tubulin subunits that are formed by the non-covalent association of alpha-tubulin and beta-tubulin. Within the microtubule, tubulin dimers associate head-to-tail to form protofilaments,

which associate laterally into a tube of about 25 nm diameter (Figure 1). Microtubules with 13 protofilaments are most commonly found in cells, but 11- or 15-protofilament microtubules have been observed in specialised cell types in mammals and invertebrates (Chalfie and Thomson, 1979; Cueva *et al.*, 2012; Saito and Hama, 1982). Protofilaments are staggered by 0.9 nm

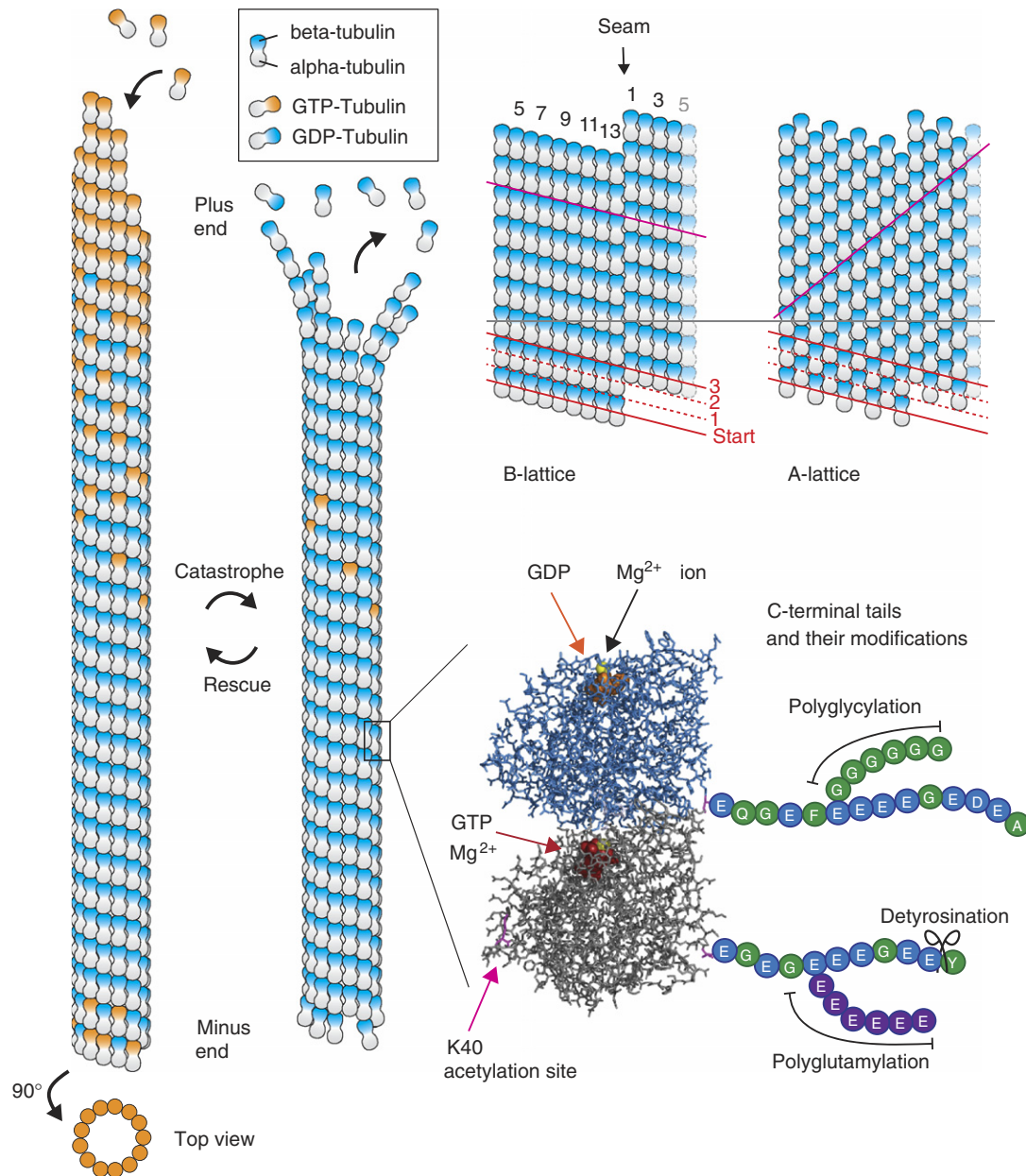


Figure 1 Microtubules are formed from alpha/beta-tubulin heterodimers. The most common arrangement is a B-lattice of 13 protofilaments that close into a tube. Microtubules assemble and disassemble at their ends. The plus end (exposing beta-tubulin) is more dynamic than the minus end. Microtubules show dynamic instability, switching from phases of growth to phases of shrinkage, an event called catastrophe. The reverse switch is a rescue. Dynamic instability is due to a cap of GTP-tubulin at the end of growing microtubules, which protects it from shrinking. Stick model of tubulin dimer (pdb accession 1SA0) shows non-exchangeable GTP (red) in alpha-tubulin and GDP bound to beta-tubulin (orange). Sites of posttranslational modifications at Lysine 40 in alpha-tubulin and a number of residues in the C-terminal tails of both alpha- and beta-tubulin are indicated.

or 4.9 nm, resulting either in beta-tubulin to laterally contact beta-tubulin (B-lattice) or beta-tubulin contacting alpha-tubulin (A-lattice). The most commonly found arrangement is 13-protofilament B-lattice microtubules with a seam between the 13th and 1st protofilament (Mandelkow *et al.*, 1986). All tubulin heterodimers are oriented with beta-tubulin pointing towards the plus end of the microtubule, therefore, the microtubule has intrinsic polarity and two different ends (Figure 1).

Tubulin is highly conserved in all eukaryotes and tubulin-related genes are expressed in bacteria. In vertebrates, the

tubulin gene family comprises 9 alpha-tubulin and 10 beta-tubulin genes (Khodiyar *et al.*, 2007; Huzil *et al.*, 2007). The highest variability between these tubulin isoforms is located within the C-terminal 15–20 amino acids. This negatively charged C-terminal tail of tubulin is further diversified by a number of posttranslational modifications (Figure 1). For example, the C-terminal tyrosine residue in alpha-tubulin is removed by a carboxypeptidase and added back by tubulin tyrosine ligase (Ersfeld *et al.*, 1993). Further modifications include the addition of polyglutamate and polyglycine side

chains to the C-terminal tail of alpha- and beta-tubulin and the acetylation of Lysine 40 on alpha-tubulin. Tubulin post-translational modifications occur on a subset of microtubules in cells and have been implicated in regulating the access of microtubule-associated proteins (MAPs) and motors to the microtubule. Specific modifications can either protect microtubules from depolymerizing enzymes, label them for destruction by severing enzymes or guide motor transport (Peris *et al.*, 2009; Lacroix *et al.*, 2010; Sirajuddin *et al.*, 2014). The details of this tubulin code and its implication in microtubule processes remain to be elucidated (See also Hammond *et al.*, 2008; Janke and Bulinski, 2011).

Microtubule Dynamics

Microtubules assemble and disassemble by addition and loss of subunits from their ends. The alpha-tubulin always contains a non-hydrolysable guanosine triphosphate (GTP), whereas the beta-tubulin contains a GTP molecule that can be hydrolyzed to guanosine diphosphate (GDP) during microtubule polymerization (Figure 1). It is the GTP hydrolysis in beta-tubulin that governs the dynamic switching between microtubule assembly and disassembly (Howard and Hyman, 2009; Mitchison and Kirschner, 1984). In the GTP cap model of dynamic instability, GTP-tubulin (i.e., tubulin with GTP bound at its exchangeable site in beta-tubulin) is incorporated at the growing microtubule end. Addition of the next layer of tubulin subunits accelerates GTP hydrolysis of beta-tubulin as the alpha-tubulin of the next subunit serves as GTPase activation protein (Nogales *et al.*, 1998). Therefore, the majority of the microtubule lattice is formed of GDP-tubulin with a small GTP cap at the growing end. This GTP cap stabilises the microtubule structure and keeps the microtubule in a polymerizing state. When the GTP cap is lost and GDP-tubulin is exposed at the microtubule end, a catastrophic depolymerization of the microtubule occurs. There are two models that explain the stabilising effect of GTP-tubulin. GTP hydrolysis generates mechanical strain in the microtubule lattice due to compaction at the nucleotide binding site (Alushin *et al.*, 2014; Hyman *et al.*, 1995). Loss of the GTP cap results in release of the strain by outward curling of the GDP-tubulin protofilament as seen for depolymerizing microtubules in electron micrographs (Mandelkow *et al.*, 1991). However, also GTP-tubulin has a curved conformation in solution and straightening occurs at incorporation into the microtubule lattice (Rice *et al.*, 2008; Nawrotek *et al.*, 2011; Buey *et al.*, 2006). Others have observed the nucleotide to affect the lateral interactions of tubulin subunits resulting in stronger, more difficult to break protofilament–protofilament interactions in GTP-tubulin lattices (Yajima *et al.*, 2012; Quiniou *et al.*, 2013).

While depolymerizing microtubules in cells and *in vitro* show predominantly frayed ends of the peeling protofilaments (Arnal *et al.*, 2000; Mandelkow *et al.*, 1991; Zovko *et al.*, 2008), the structure of growing microtubules is less clear. Electron microscopic observations suggest that microtubules grow as gently curved sheets that gradually close into the tube (Vitre *et al.*, 2008; Chretien *et al.*, 1995). However, many of these images are also compatible with tapered microtubule

ends, that is, incomplete microtubule lattice due to the different length of protofilaments. Tapered microtubule ends are in agreement with the variable decay of signal intensity at the ends of fluorescently labeled microtubules in cells and *in vitro* (Gardner *et al.*, 2011a; Coombes *et al.*, 2013).

At transitions between growth and shrinkage phases, the GTP cap is lost and regained. A microtubule will grow as long as the addition of new GTP-tubulin subunits is faster or equal to the rate of GTP hydrolysis. If assembly slows down, for example, if the microtubule is growing into a physical barrier such as the edge of the cell, GTP hydrolysis can catch up with the tip and the microtubule shrinks. This transition is called a catastrophe. The distribution of microtubule growth durations *in vitro* and in cells suggests that transitions through an intermediate state or the evolution of a more tapered tip structure occurs before catastrophe (Coombes *et al.*, 2013; Gardner *et al.*, 2011b; Odde *et al.*, 1995; Stepanova *et al.*, 2010). When a depolymerizing microtubule regrows, it is called a rescue. One idea is that GTP remnants in the microtubule lattice stimulate rescues (Dimitrov *et al.*, 2008).

Cellular Functions of Microtubules

Microtubules are the stiffest of the cytoskeletal filaments, 200-times stiffer than actin filaments when isolated (Gittes *et al.*, 1993). In crowded environments like the cytoplasm that restrict buckling, microtubules can bear very high compressive loads (Figure 2; Brangwynne *et al.*, 2006). Therefore, microtubules serve as structural elements in cells and can generate pulling and pushing forces. At the same time, microtubules provide the infrastructure for intracellular long-distance transport. Molecular motors of the kinesin and dynein family walk on the surface of microtubules to transport organelles, vesicles and molecular cargoes within the cell (Figure 2). In the case of a neuron, transport of neurotransmitters from the cell body to the presynaptic site requires covering distances in excess of 1 m. While single microtubules in the axon are on an average only 5 µm long, they are arranged in parallel bundles (Yu and Baas, 1994) to allow continuous motor-based transport.

In cells, the microtubule minus end is usually capped by gamma-tubulin and associated proteins, while the more dynamic plus end explores the intracellular space. The highest rate of microtubule catastrophe is observed at the cell edges (Komarova *et al.*, 2002; Drummond and Cross, 2000). This ensures that microtubule-based transport reaches the cell periphery. Furthermore, microtubule contacts with the cell cortex allow centering forces to be applied to the microtubule network (Figure 2). Such forces can be pushing of growing microtubule bundles against the cell tips as used for nuclear positioning in fission yeast (Tran *et al.*, 2001) or pulling forces mediated by cortically anchored dynein as used for spindle positioning in human cells (Samora *et al.*, 2011; Laan *et al.*, 2012). Coupling microtubule dynamics to subcellular structure allows harnessing forces mediated by polymerization and depolymerization (Figure 2). The segregation of chromosomes during

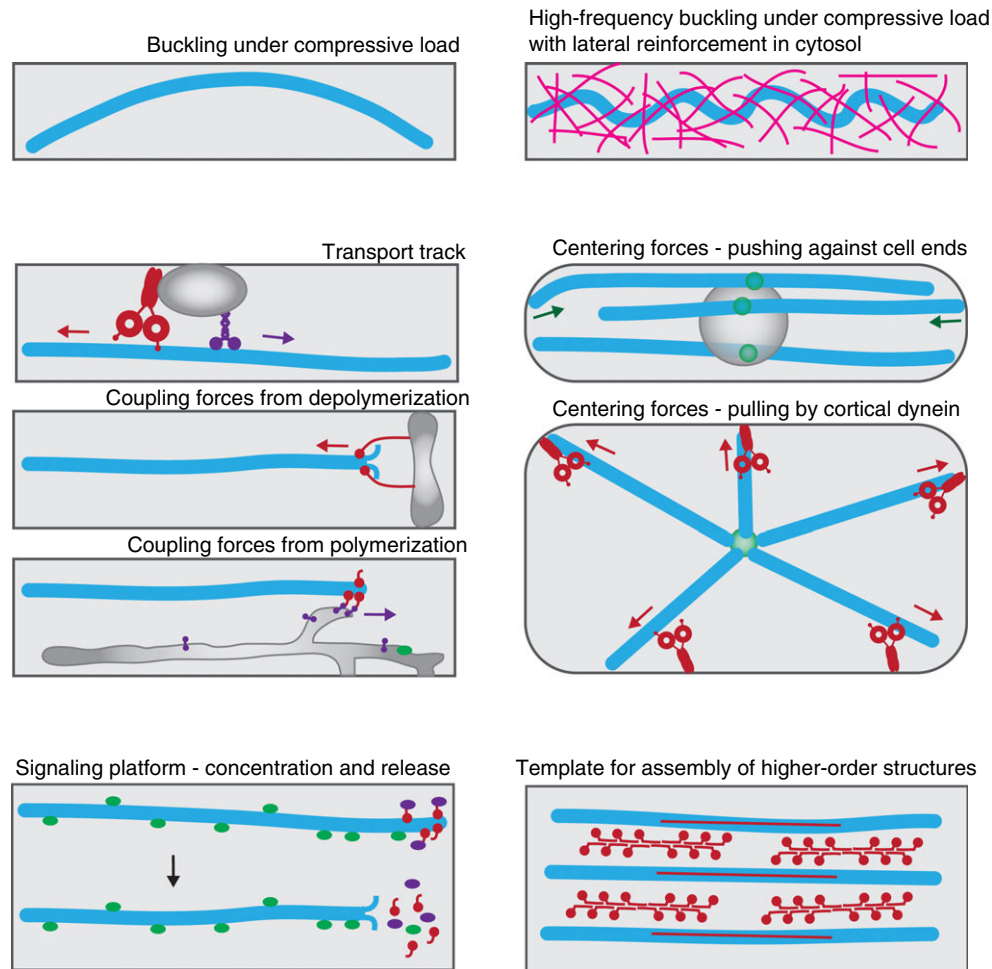


Figure 2 Microtubules and their associated proteins have a number of functions in cells. Being the stiffest cytoskeletal filaments, they carry high-compressive loads, especially when laterally reinforced. Microtubules serve as long-distance transport tracks, generate pushing and pulling forces, and act as signalling platforms and as templates for cellular self-organization.

mitosis is driven by coupling shrinking microtubules to kinetochores (Asbury *et al.*, 2006; Coue *et al.*, 1991), while polymerizing microtubule ends push membrane tubules of the endoplasmic reticulum (Waterman-Storer and Salmon, 1998; Grigoriev *et al.*, 2008).

The dynamic turnover of microtubules allows the rapid reorganization of the microtubule cytoskeleton into suitable arrays throughout the cell cycle and when cells change shape during cell locomotion or differentiation. Undifferentiated animal cells have a predominantly radial array of microtubules emanating from the centrosome. To form the mitotic spindle, microtubule nucleation is increased dramatically at mitotic entry, the two spindle poles separate and microtubules are stabilized upon contact with the kinetochores (Huitorel and Kirschner, 1988; Rieder, 2005; Snyder and McIntosh, 1975). Bundles of parallel microtubules fill the length of the axon. The minus ends of these microtubules are no longer embedded in the centrosome, but scattered throughout the cell (Yu and Baas, 1994). Acentrosomal microtubule arrangements are also commonly found in

muscle cells and polarized epithelia (Warren, 1974; Mogensen *et al.*, 2000). In plant cells, an antiparallel microtubule network lines the cell cortex (Hardham and Gunning, 1978). These specialized arrangements support cell polarity by allowing directional long-distance transport, serve as templates for the formation of higher-order structures such as the cell wall in plants and the sarcomers in skeletal muscle and act as structural scaffold to strengthen the cell axis (Figure 2; Paredez *et al.*, 2006; Pizon *et al.*, 2005; Franker and Hoogenraad, 2013).

Organelles are linked to microtubules to control their distribution in the cytoplasm. Examples for this are the distribution and dynamics of ER and mitochondria (Waterman-Storer and Salmon, 1998; Morris and Hollenbeck, 1995), ribosomes (Suprenant *et al.*, 1993) and the positioning of nuclei (Tran *et al.*, 2001; Metzger *et al.*, 2012). Microtubules probably also serve as signaling platforms as they bind and activate RhoGTPase regulators such as GEF-H1, STEF/Tiam and kinases such as Arg (Krendel *et al.*, 2002; Rooney *et al.*, 2010; Miller *et al.*, 2004). These molecules probably couple

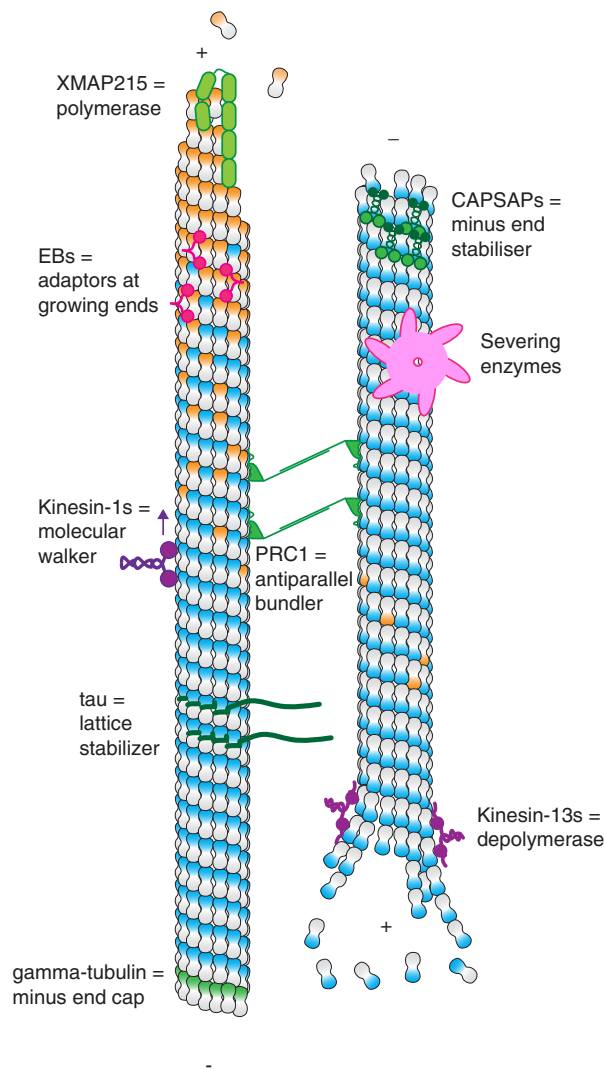


Figure 3 MAPs are a structurally diverse group of proteins that either bind along microtubules (e.g., tau) or preferentially bind substructures on the microtubule such as the microtubule ends (e.g., EBs, CAMSAPs, and XMAP215). Other MAPs accumulate in microtubule overlaps (PRC1). MAPs modulate microtubule dynamics. They either induce catastrophe (e.g., kinesin-13s), cut the microtubule lattice to make two microtubules (severing enzymes), nucleate new microtubules (gamma-tubulin), or facilitate microtubule growth (e.g., XMAP215).

microtubule dynamics to signaling pathways controlling cell migration and shape changes (Figure 2; Kaverina and Straube, 2011).

MAPs

Microtubule dynamics, rigidity and functions are regulated by MAPs. MAPs can be split into two categories based on whether they bind to specific sites on the microtubule or decorate them along their length. The classic structural MAPs from the MAP2/tau family fall into the latter category. These MAPs recognize

the lattice of the microtubule and through their binding to more than one tubulin dimer often act as glue to stabilize the microtubule. Further lattice-binding MAPs are MAP4, Enscin/Map7, EMAPs and EMAP-like proteins, and GEF-H1 (Suprenant *et al.*, 1993; Chapin and Bulinski, 1994; Masson and Kreis, 1993; Krendel *et al.*, 2002; Rooney *et al.*, 2010). The second category includes MAPs that preferentially recognize specific features of microtubules or microtubule arrangements such as

- the microtubule plus end: XMAP215 family (Brouhard *et al.*, 2008)
- the microtubule minus end: gamma-tubulin (Li and Joshi, 1995); CAMSAP/Nezha/Patronin family (Meng *et al.*, 2008; Goodwin and Vale, 2010); Fidgetin (Mukherjee *et al.*, 2012)
- the assembly status of the microtubule end: growing ends – EB family and interacting + TIPs (Maurer *et al.*, 2012; Jiang *et al.*, 2012); shrinking ends – MCAK/Kif2C (Hunter *et al.*, 2003)
- bundles of microtubules: PRC1/Map65/Ase1 family (Gaillard *et al.*, 2008; Subramanian *et al.*, 2010; Liodice *et al.*, 2005); HSET/KifC1 family (Fink *et al.*, 2009; Braun *et al.*, 2009); Eg5/Kif11 (Kapitein *et al.*, 2005)
- microtubule crossovers: Katanin (Zhang *et al.*, 2013; Lindeboom *et al.*, 2013)
- regions of microtubule curvature: Doublecortin (Bechstedt and Brouhard, 2012)
- microtubule defects: Katanin (Diaz-Valencia *et al.*, 2011).

The ability to recognize specific structural features of microtubules is usually paired with the ability to specifically stabilize or destabilize the recognized microtubule substructure, enabling a multitude of ways how MAPs regulate microtubule structure, dynamics, and organization (Figure 3).

Regulation of Microtubule Structure and Mechanics by MAPs

MAPs control microtubule structure and rigidity. A number of MAPs are implicated in controlling the number of protofilaments. For example, in the *Caenorhabditis elegans* touch receptor neurons, the formation of 15-protofilament microtubules depends on the presence of alpha-tubulin acetyl transferase, the enzyme that acetylates Lysine 40 on alpha-tubulin in the lumen of the microtubule (Cueva *et al.*, 2012). 13-Protofilament microtubules are promoted through the action of doublecortin and EB1, both proteins localize between protofilaments and thereby constrain lateral tubulin interactions. In addition, both proteins track growing microtubule ends ensuring their presence at microtubule assembly (Fourniol *et al.*, 2010; Moores *et al.*, 2004; Vitre *et al.*, 2008; Maurer *et al.*, 2012; Bechstedt and Brouhard, 2012). Tau stabilizes microtubules and significantly increases their rigidity if added during microtubule assembly (Hawkins *et al.*, 2013). While tau still decorates the microtubule lattice of preformed microtubules, only copolymerization allows it to bind to sites in the microtubule lumen (Makrides *et al.*, 2004; Kar *et al.*, 2003). The tau-related, non-neuronal protein MAP4 does not

change microtubule rigidity, while MAP65-1 and Ase1 decrease rigidity (Hawkins *et al.*, 2013; Portran *et al.*, 2013). Thus binding to the microtubule lattice in itself does not stiffen the microtubule, but MAPs instead seem to modulate intradimer and interdimer distances to tune microtubule stiffness. Different microtubule stiffness could be of functional importance: In an environment such as a neuron, where microtubules play a structural role to support long and thin cellular processes, stiff tau-stabilized microtubules will be an advantage. In the microtubule spindle, formation of antiparallel overlaps in the spindle midzone by PRC1/MAP65/Ase1 requires bending of microtubules that encounter each other at high incident angles.

Regulation of Microtubule Dynamics by MAPs

Microtubule assembly and disassembly is regulated by an arsenal of cellular factors (Van der Vaart *et al.*, 2009). While some of those can be classified as microtubule stabilizers or destabilizers, there are a number of factors that have more complicated effects. Microtubule dynamics is usually described by four or five parameters: the speed of growth and shrinkage, the rates of catastrophe and rescue and the fraction of time spent in pause (Straube, 2011). Microtubule stabilizers, such as the lattice-decorating MAPs tau and MAP2, promote growth and rescue, prevent catastrophe and slow shrinkage (Drechsel *et al.*, 1992; Kowalski and Williams, 1993). All these parameter changes are compatible with MAP2/tau acting as glue to stabilize the contacts between tubulin dimers in the microtubule lattice. Microtubule destabilisers, such as stathmin-family proteins and kinesin-13s, promote catastrophe and prevent rescue by sequestering free tubulin dimers and inducing a curled protofilament conformation (Newton *et al.*, 2004; Steinmetz, 2007). Proteins of the chTOG/XMAP215 family accelerate microtubule assembly and disassembly in a catalytic manner. Thus depending on whether the reaction conditions favor assembly or disassembly, XMAP215 speeds up the addition or loss of tubulin dimers at the microtubule plus end, probably through stabilization of an intermediate state of tubulin assembly/disassembly (Brouhard *et al.*, 2008). At the same time transitions between growth and shrinkage are suppressed. In contrast, a number of proteins promote fast transitions between assembly and shrinkage. Among those are CLASPs and ACF7, factors that mediate the capture of microtubules at the cell cortex or kinetochores (Kodama *et al.*, 2003; Mimori-Kiyosue *et al.*, 2005). The mechanism of their action is less clear. CLASPs contain a number of microtubule-interacting domains and preferentially bind to a region about 1 μm away from the microtubule tip. Stabilizing the interaction of several neighboring tubulin dimers at its binding site could delay microtubule shrinkage and favor regrowth.

Minus end-capping proteins such as gamma-tubulin, ninein and members of the CAMSAP/Nezha/Patronin family stabilize the microtubule minus end and protect it from depolymerization (Goodwin and Vale, 2010; Li and Joshi, 1995; Meng *et al.*, 2008; Mogensen *et al.*, 2000). Microtubule dynamics is restricted to the ends of microtubules. However, severing proteins Spastin, Katanin and Fidgetin cut the microtubule lattice (Sharp and Ross, 2012). This generates a

new minus and plus end, respectively, on the two resulting microtubules. In addition to severing, Katanin has been shown to accelerate depolymerization from both microtubule plus ends, while Fidgetin is a minus end depolymerase (Mukherjee *et al.*, 2012; Zhang *et al.*, 2011). Thus the combined activity of severing and depolymerization can result in the rapid disassembly of microtubules unless the new ends recruit minus end-capping proteins or conditions favor microtubule assembly.

Especially when looking at proteins that track the growing ends of microtubules, the so-called +TIPs, understanding the exact contribution of each protein can be difficult due to the dynamic interactions of a large number of proteins in the +TIP complex. Depleting one factor can result in freeing up space for another factor. Therefore, depletion experiments provide potentially misleading information about the function of these proteins. Reconstitution of microtubule dynamics and tip tracking from purified components has proven to be powerful to study those proteins individually and the synergistic effects of combinations of proteins in a controlled environment. One of such examples is the interaction of end-binding proteins (EBs) and mitotic centromere-associated kinesin (MCAK). MCAK is a potent depolymerase inducing catastrophe and accelerating shrinkage. In cells, MCAK tracks growing microtubule ends via EB1 and EB3. Reconstitution of this module reveals that EBs promote robust growth phases even in the presence of MCAK, while the targeting to microtubule ends via EBs increases MCAK-mediated catastrophes (Montenegro Gouveia *et al.*, 2010).

Arrangement of Microtubules into Higher-Order Structures by MAPs

Microtubules are organized into suitable arrays by the cooperation of mechanisms that control nucleation, selective stabilization and destabilization, microtubule transport and sliding, bundling and crossbridging. Microtubules nucleated from centrosomes will result in radial arrays, therefore, parallel microtubule arrays with either uniform directionality or antiparallel arrangements are often formed by acentrosomal nucleation. Acentrosomal nucleation can occur free in the cytoplasm or by gamma-tubulin containing complexes residing on existing microtubules (Musa *et al.*, 2003; Petry *et al.*, 2013; Straube *et al.*, 2003). Microtubules can also be amplified by severing. Cutting a microtubule produces two microtubules with the same orientation. Thus a cycle of severing and assembly can very quickly lead to the amplification of microtubules of a certain orientation, a process that has been implicated in the light-induced reorganization of the cortical microtubule array in plants and microtubule reorganization events in meiotic spindles (Lindeboom *et al.*, 2013; McNally *et al.*, 2006).

While nucleation and severing will determine the initial position of the microtubule minus end, microtubules can be transported along other microtubules or the cell cortex. Motors implicated in microtubule sliding are dynein, kinesin-1, Eg5, and kinesin-14s (Straube *et al.*, 2006; Jolly *et al.*, 2010; Fink, 2006; Rusan *et al.*, 2002; Kapitein *et al.*, 2005; Tanenbaum *et al.*, 2013; Fink *et al.*, 2009). Force generated by microtubule

sliding is crucial for the initiation of neurite growth and microtubule fragments are transported into neurites where they extend into parallel arrays (Lu *et al.*, 2013; Yu *et al.*, 1996). Microtubules are locked into position by crosslinking proteins such as proteins of the PRC1 family, which preferentially bind and stabilize antiparallel overlaps. While we start to understand from reconstitution experiments how individual MAPs and motors function, there are only few studies thus far looking at the cooperation of two or more components. These revealed functional modules with interesting new properties. For example, PRC1 and kinesin-4 cooperate to keep antiparallel microtubule overlaps at a specific length (Bieling *et al.*, 2010). Likewise PRC1-homolog Ase1 shows adaptive breaking to prevent antiparallel overlaps from being slid apart completely by kinesin-14 motors (Braun *et al.*, 2011).

The selective stabilization or destruction of microtubules is another important mechanism to microtubule array organization. Selective stabilization often occurs when microtubules engage with the cell cortex or other subcellular structures, such as the kinetochore in mitotic cells. Microtubule capture limits microtubule dynamics to short length oscillations and serves to maintain long-lived contacts that allow cargo delivery to sites of capture, transduction of forces from microtubule assembly/disassembly to drive chromosome movements and the maintenance of a polarized, front-biased microtubule cytoskeleton in migrating cells (Kaverina and Straube, 2011; Asbury *et al.*, 2006; Straube, 2011). Attraction of severing enzymes to microtubules containing posttranslational modifications such as poly-glutamylation allows the removal of a subset of microtubules, while deetyrosination protects microtubules from depolymerization by MCAK (Lacroix *et al.*, 2010; Peris *et al.*, 2009). However, we still have to understand how selective modification and de-modification is regulated in cells. Modifications accumulate on microtubules that are stable and long-lived, but microtubule age cannot be the only factor as microtubules in the same cell can acquire different modifications and these tend to extend over the entire length of a particular microtubule.

Recent years have seen important advances in the techniques to study how complex assemblies of microtubules, MAPs and motors are formed and maintained while constantly turning over. These will in the next decade provide exciting new insights into cellular self-organization principles.

Microtubules in Human Disease

Microtubules are of central importance for cell division, cell morphology, and intracellular transport. Consequently, a number of diseases are associated with impaired microtubule function, most notably neuronal development defects, neurodegenerative disorders and ciliary dysfunction (Field *et al.*, 2014; Reiner, 2013; Bisgrove and Yost, 2006). Microtubule-targeting agents cause mitotic arrest and associated cell death, block cell migration and have vascular-disrupting properties. They are routinely used for the treatment of cancer and chronic inflammatory diseases and are in clinical trials for treating neurodegenerative diseases (Field *et al.*, 2014).

See also: Cytoskeleton and Motors: Complex Cytoskeletal Structures: Centrioles and the Centrosome; Cilia and Flagella; The Mitotic Spindle. Cytoskeleton and Motors: Cytoskeletal Components: Kinesin Superfamily Proteins (KIFs) as a Fundamental Component of Life: Intracellular Transport and Beyond. Dynamic Integration: Dynamics: Dynamics of Microtubule Self-Assembly

References

- Alushin, G.M., Lander, G.C., Kellogg, E.H., *et al.*, 2014. High-resolution microtubule structures reveal the structural transitions in α -tubulin upon GTP hydrolysis. *Cell* 157, 1117–1129.
- Arnal, I., Karsenti, E., Hyman, A.A., 2000. Structural transitions at microtubule ends correlate with their dynamic properties in *Xenopus* egg extracts. *Journal of Cell Biology* 149, 767–774.
- Asbury, C.L., Gestaut, D.R., Powers, A.F., Franck, A.D., Davis, T.N., 2006. The Dam1 kinetochore complex harnesses microtubule dynamics to produce force and movement. *Proceedings of the National Academy of Sciences of the United States of America* 103, 9873–9878.
- Bechstedt, S., Brouhard, G.J., 2012. Doublecortin recognizes the 13-prot filament microtubule cooperatively and tracks microtubule ends. *Developmental Cell* 23, 181–192.
- Bieling, P., Telley, I.A., Surrey, T., 2010. A minimal midzone protein module controls formation and length of antiparallel microtubule overlaps. *Cell* 142, 420–432.
- Bisgrove, B.W., Yost, H.J., 2006. The roles of cilia in developmental disorders and disease. *Development* 133, 4131–4143.
- Brangwynne, C.P., Mackintosh, F.C., Kumar, S., *et al.*, 2006. Microtubules can bear enhanced compressive loads in living cells because of lateral reinforcement. *Journal of Cell Biology* 173, 733–741.
- Braun, M., Drummond, D.R., Cross, R.A., Mcainsh, A.D., 2009. The kinesin-14 Klp2 organizes microtubules into parallel bundles by an ATP-dependent sorting mechanism. *Nature Cell Biology* 11, 724–730.
- Braun, M., Lansky, Z., Fink, G., *et al.*, 2011. Adaptive braking by Ase1 prevents overlapping microtubules from sliding completely apart. *Nature Cell Biology* 13, 1259–1264.
- Brouhard, G.J., Stear, J.H., Noetzel, T.L., *et al.*, 2008. XMAP215 is a processive microtubule polymerase. *Cell* 132, 79–88.
- Buey, R.M., Diaz, J.F., Andreu, J.M., 2006. The nucleotide switch of tubulin and microtubule assembly: A polymerization-driven structural change. *Biochemistry* 45, 5933–5938.
- Chalfie, M., Thomson, J.N., 1979. Organization of neuronal microtubules in the nematode *Caenorhabditis elegans*. *Journal of Cell Biology* 82, 278–289.
- Chapin, S.J., Bulinski, J.C., 1994. Cellular microtubules heterogeneous in their content of microtubule-associated protein 4 (MAP4). *Cell Motility and the Cytoskeleton* 27, 133–149.
- Chretien, D., Fuller, S.D., Karsenti, E., 1995. Structure of growing microtubule ends: Two-dimensional sheets close into tubes at variable rates. *Journal of Cell Biology* 129, 1311–1328.
- Coomes, C.E., Yamamoto, A., Kenzie, M.R., Odde, D.J., Gardner, M.K., 2013. Evolving tip structures can explain age-dependent microtubule catastrophe. *Current Biology* 23, 1342–1348.
- Coue, M., Lombillo, V.A., McIntosh, J.R., 1991. Microtubule depolymerization promotes particle and chromosome movement *in vitro*. *Journal of Cell Biology* 112, 1165–1175.
- Cueva, J.G., Hsin, J., Huang, K.C., Goodman, M.B., 2012. Posttranslational acetylation of α -tubulin constrains protofilament number in native microtubules. *Current Biology* 22, 1066–1074.
- Diaz-Valencia, J.D., Morelli, M.M., Bailey, M., *et al.*, 2011. *Drosophila* katanin-60 depolymerizes and severs at microtubule defects. *Biophysical Journal* 100, 2440–2449.
- Dimitrov, A., Quesnoit, M., Moutel, S., *et al.*, 2008. Detection of GTP-tubulin conformation *in vivo* reveals a role for GTP remnants in microtubule rescues. *Science* 322, 1353–1356.
- Drechsel, D.N., Hyman, A.A., Cobb, M.H., Kirschner, M.W., 1992. Modulation of the dynamic instability of tubulin assembly by the microtubule-associated protein tau. *Molecular Biology of the Cell* 3, 1141–1154.

- Drummond, D.R., Cross, R.A., 2000. Dynamics of interphase microtubules in *Schizosaccharomyces pombe*. *Current Biology* 10, 766–775.
- Ersfeld, K., Wehland, J., Plessmann, U., *et al.*, 1993. Characterization of the tubulin-tyrosine ligase. *Journal of Cell Biology* 120, 725–732.
- Field, J.J., Kanakanthara, A., Miller, J.H., 2014. Microtubule-targeting agents are clinically successful due to both mitotic and interphase impairment of microtubule function. *Bioorganic & Medicinal Chemistry* 22, 5050–5059.
- Fink, G., 2006. Dynein-dependent motility of microtubules and nucleation sites supports polarization of the tubulin array in the fungus *Ustilago maydis*. *Molecular Biology of the Cell* 17, 3242–3253.
- Fink, G., Hajdu, L., Skowronek, K.J., *et al.*, 2009. The mitotic kinesin-14 Ncd drives directional microtubule-microtubule sliding. *Nature Cell Biology* 11, 717–723.
- Fourniol, F.J., Sindelar, C.V., Amigues, B., *et al.*, 2010. Template-free 13-prot filament microtubule-MAP assembly visualized at 8 Å resolution. *Journal of Cell Biology* 191, 463–470.
- Franker, M.A., Hoogenraad, C.C., 2013. Microtubule-based transport - basic mechanisms, traffic rules and role in neurological pathogenesis. *Journal of Cell Science* 126, 2319–2329.
- Gaillard, J., Neumann, E., Van Damme, D., *et al.*, 2008. Two microtubule-associated proteins of Arabidopsis MAP65s promote antiparallel microtubule bundling. *Molecular Biology of the Cell* 19, 4534–4544.
- Gardner, M.K., Charlebois, B.D., Janosi, I.M., *et al.*, 2011a. Rapid microtubule self-assembly kinetics. *Cell* 146, 582–592.
- Gardner, M.K., Zanic, M., Gell, C., Bormuth, V., Howard, J., 2011b. Depolymerizing kinesins Kip3 and MCAK shape cellular microtubule architecture by differential control of catastrophe. *Cell* 147, 1092–1103.
- Gittes, F., Mickey, B., Nettleton, J., Howard, J., 1993. Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape. *Journal of Cell Biology* 120, 923–934.
- Goodwin, S.S., Vale, R.D., 2010. Patronin regulates the microtubule network by protecting microtubule minus ends. *Cell* 143, 263–274.
- Grigoriev, I., Gouveia, S.M., Van Der Vaart, B., *et al.*, 2008. STIM1 is a MT-plus-end-tracking protein involved in remodeling of the ER. *Current Biology* 18, 177–182.
- Hammond, J.W., Cai, D., Verhey, K.J., 2008. Tubulin modifications and their cellular functions. *Current Opinion in Cell Biology* 20, 71–76.
- Hardham, A.R., Gunning, B.E., 1978. Structure of cortical microtubule arrays in plant cells. *Journal of Cell Biology* 77, 14–34.
- Hawkins, T.L., Sept, D., Mogessie, B., Straube, A., Ross, J.L., 2013. Mechanical properties of doubly stabilized microtubule filaments. *Biophysical Journal* 104, 1517–1528.
- Howard, J., Hyman, A.A., 2009. Growth, fluctuation and switching at microtubule plus ends. *Nature Reviews Molecular Cell Biology* 10, 569–574.
- Huitorrel, P., Kirschner, M.W., 1988. The polarity and stability of microtubule capture by the kinetochore. *Journal of Cell Biology* 106, 151–159.
- Hunter, A.W., Caplow, M., Coy, D.L., *et al.*, 2003. The kinesin-related protein MCAK is a microtubule depolymerase that forms an ATP-hydrolyzing complex at microtubule ends. *Molecular Cell* 11, 445–457.
- Huzli, J.T., Chen, K., Kurgan, L., Tuszyński, J.A., 2007. The roles of beta-tubulin mutations and isotype expression in acquired drug resistance. *Cancer Information* 3, 159–181.
- Hyman, A.A., Chretien, D., Arnal, I., Wade, R.H., 1995. Structural changes accompanying GTP hydrolysis in microtubules: Information from a slowly hydrolyzable analogue guanylyl-(alpha,beta)-methylene-diphosphonate. *Journal of Cell Biology* 128, 117–125.
- Janke, C., Bulinski, J.C., 2011. Post-translational regulation of the microtubule cytoskeleton: Mechanisms and functions. *Nature Reviews Molecular Cell Biology* 12, 773–786.
- Jiang, K., Toedt, G., Montenegro Gouveia, S., *et al.*, 2012. A proteome-wide screen for mammalian SxIP motif-containing microtubule plus-end tracking proteins. *Current Biology* 22, 1800–1807.
- Jolly, A.L., Kim, H., Srinivasan, D., *et al.*, 2010. Kinesin-1 heavy chain mediates microtubule sliding to drive changes in cell shape. *Proceedings of the National Academy of Sciences of the United States of America* 107, 12151–12156.
- Kapitein, L.C., Peterman, E.J., Kwok, B.H., *et al.*, 2005. The bipolar mitotic kinesin Eg5 moves on both microtubules that it crosslinks. *Nature* 435, 114–118.
- Kar, S., Fan, J., Smith, M.J., Goedert, M., Amos, L.A., 2003. Repeat motifs of tau bind to the insides of microtubules in the absence of taxol. *EMBO Journal* 22, 70–77.
- Kaverina, I., Straube, A., 2011. Regulation of cell migration by dynamic microtubules. *Seminars in Cell & Developmental Biology* 22, 968–974.
- Khodiyar, V.K., Maltais, L.J., Ruef, B.J., *et al.*, 2007. A revised nomenclature for the human and rodent alpha-tubulin gene family. *Genomics* 90, 285–289.
- Kodama, A., Karakesisoglou, I., Wong, E., Vaezi, A., Fuchs, E., 2003. ACF7: An essential integrator of microtubule dynamics. *Cell* 115, 343–354.
- Komarova, Y.A., Vorobjev, I.A., Borisy, G.G., 2002. Life cycle of MTs: Persistent growth in the cell interior, asymmetric transition frequencies and effects of the cell boundary. *Journal of Cell Science* 115, 3527–3539.
- Kowalski, R.J., Williams Jr., R.C., 1993. Microtubule-associated protein 2 alters the dynamic properties of microtubule assembly and disassembly. *Journal of Biological Chemistry* 268, 9847–9855.
- Krendel, M., Zenke, F.T., Bokoch, G.M., 2002. Nucleotide exchange factor GEF-H1 mediates cross-talk between microtubules and the actin cytoskeleton. *Nature Cell Biology* 4, 294–301.
- Laan, L., Pavin, N., Husson, J., *et al.*, 2012. Cortical dynein controls microtubule dynamics to generate pulling forces that position microtubule asters. *Cell* 148, 502–514.
- Lacroix, B., Van Dijk, J., Gold, N.D., *et al.*, 2010. Tubulin polyglutamylation stimulates spastin-mediated microtubule severing. *Journal of Cell Biology* 189, 945–954.
- Li, Q., Joshi, H.C., 1995. gamma-tubulin is a minus end-specific microtubule binding protein. *Journal of Cell Biology* 131, 207–214.
- Lindeboom, J.J., Nakamura, M., Hibbel, A., *et al.*, 2013. A mechanism for reorientation of cortical microtubule arrays driven by microtubule severing. *Science* 342, 1245533.
- Loiodice, I., Staub, J., Setty, T.G., *et al.*, 2005. Ase1p organizes antiparallel microtubule arrays during interphase and mitosis in fission yeast. *Molecular Biology of the Cell* 16, 1756–1768.
- Lu, W., Fox, P., Lakonishok, M., Davidson, M.W., Gelfand, V.I., 2013. Initial neurite outgrowth in Drosophila neurons is driven by kinesin-powered microtubule sliding. *Current Biology* 23, 1018–1023.
- Makrides, V., Massie, M.R., Feinstein, S.C., Lew, J., 2004. Evidence for two distinct binding sites for tau on microtubules. *Proceedings of the National Academy of Sciences of the United States of America* 101, 6746–6751.
- Mandelkow, E.M., Mandelkow, E., Milligan, R.A., 1991. Microtubule dynamics and microtubule caps: A time-resolved cryo-electron microscopy study. *Journal of Cell Biology* 114, 977–991.
- Mandelkow, E.M., Schultheiss, R., Rapp, R., Muller, M., Mandelkow, E., 1986. On the surface lattice of microtubules: Helix starts, protofilament number, seam, and handedness. *Journal of Cell Biology* 102, 1067–1073.
- Masson, D., Kreis, T.E., 1993. Identification and molecular characterization of E-MAP-115, a novel microtubule-associated protein predominantly expressed in epithelial cells. *Journal of Cell Biology* 123, 357–371.
- Maurer, S.P., Fourniol, F.J., Bohner, G., Moores, C.A., Surrey, T., 2012. EBs recognize a nucleotide-dependent structural cap at growing microtubule ends. *Cell* 149, 371–382.
- McNally, K., Audhya, A., Oegema, K., McNally, F.J., 2006. Katanin controls mitotic and meiotic spindle length. *Journal of Cell Biology* 175, 881–891.
- Meng, W., Mushika, Y., Ichii, T., Takeichi, M., 2008. Anchorage of microtubule minus ends to adherens junctions regulates epithelial cell–cell contacts. *Cell* 135, 948–959.
- Metzger, T., Gache, V., Xu, M., *et al.*, 2012. MAP and kinesin-dependent nuclear positioning is required for skeletal muscle function. *Nature* 484, 120–124.
- Miller, A.L., Wang, Y., Mooseker, M.S., Koleske, A.J., 2004. The Abl-related gene (Arg) requires its F-actin-microtubule cross-linking activity to regulate lamellipodial dynamics during fibroblast adhesion. *Journal of Cell Biology* 165, 407–419.
- Mimori-Kiyosue, Y., Grigoriev, I., Lansbergen, G., *et al.*, 2005. CLASP1 and CLASP2 bind to EB1 and regulate microtubule plus-end dynamics at the cell cortex. *Journal of Cell Biology* 168, 141–153.
- Mitchison, T., Kirschner, M., 1984. Dynamic instability of microtubule growth. *Nature* 312, 237–242.
- Mogensen, M.M., Malik, A., Piel, M., Bouckson-Castaing, V., Bornens, M., 2000. Microtubule minus-end anchorage at centrosomal and non-centrosomal sites: The role of ninein. *Journal of Cell Science* 113, 3013–3023.
- Montenegro Gouveia, S., Leslie, K., Kapitein, L.C., *et al.*, 2010. *In vitro* reconstitution of the functional interplay between MCAK and EB3 at microtubule plus ends. *Current Biology* 20, 1717–1722.
- Moores, C.A., Perderiset, M., Francis, F., *et al.*, 2004. Mechanism of microtubule stabilization by doublecortin. *Molecular Cell* 14, 833–839.
- Morris, R.L., Hollenbeck, P.J., 1995. Axonal transport of mitochondria along microtubules and F-actin in living vertebrate neurons. *Journal of Cell Biology* 131, 1315–1326.
- Mukherjee, S., Diaz Valencia, J.D., Stewman, S., *et al.*, 2012. Human fidgetin is a microtubule severing the enzyme and minus-end depolymerase that regulates mitosis. *Cell Cycle* 11, 2359–2366.

- Musa, H., Orton, C., Morrison, E.E., Peckham, M., 2003. Microtubule assembly in cultured myoblasts and myotubes following nocodazole induced microtubule depolymerisation. *Journal of Muscle Research and Cell Motility* 24, 301–308.
- Nawrotek, A., Knossow, M., Gigant, B., 2011. The determinants that govern microtubule assembly from the atomic structure of GTP-tubulin. *Journal of Molecular Biology* 412, 35–42.
- Newton, C.N., Wagenbach, M., Ovechkina, Y., Wordeman, L., Wilson, L., 2004. MCAK, a Kin I kinesin, increases the catastrophe frequency of steady-state HeLa cell microtubules in an ATP-dependent manner *in vitro*. *FEBS Letters* 572, 80–84.
- Nogales, E., Downing, K.H., Amos, L.A., Lowe, J., 1998. Tubulin and FtsZ form a distinct family of GTPases. *Nature Structural & Molecular Biology* 5, 451–458.
- Odde, D.J., Cassimeris, L., Buettner, H.M., 1995. Kinetics of microtubule catastrophe assessed by probabilistic analysis. *Biophysical Journal* 69, 796–802.
- Paredes, A.R., Somerville, C.R., Ehrhardt, D.W., 2006. Visualization of cellulose synthase demonstrates functional association with microtubules. *Science* 312, 1491–1495.
- Peris, L., Wagenbach, M., Lafanechere, L., *et al.*, 2009. Motor-dependent microtubule disassembly driven by tubulin tyrosination. *Journal of Cell Biology* 185, 1159–1166.
- Petry, S., Groen, A.C., Ishihara, K., Mitchison, T.J., Vale, R.D., 2013. Branching microtubule nucleation in *Xenopus* egg extracts mediated by augmin and TPX2. *Cell* 152, 768–777.
- Pizon, V., Gerbal, F., Diaz, C.C., Karsenti, E., 2005. Microtubule-dependent transport and organization of sarcomeric myosin during skeletal muscle differentiation. *EMBO Journal* 24, 3781–3792.
- Portran, D., Zoccoler, M., Gaillard, J., *et al.*, 2013. MAP65/Ase1 promote microtubule flexibility. *Molecular Biology of the Cell* 24, 1964–1973.
- Quiniou, E., Guichard, P., Perahia, D., Marco, S., Mouawad, L., 2013. An atomistic view of microtubule stabilization by GTP. *Structure* 21, 833–843.
- Reiner, O., 2013. LIS1 and DCX: Implications for brain development and human disease in relation to microtubules. *Scientifica (Cairo)* 2013, 393975.
- Rice, L.M., Montabana, E.A., Agard, D.A., 2008. The lattice as allosteric effector: Structural studies of alpha-beta- and gamma-tubulin clarify the role of GTP in microtubule assembly. *Proceedings of the National Academy of Sciences of the United States of America* 105, 5378–5383.
- Rieder, C.L., 2005. Kinetochore fiber formation in animal somatic cells: Dueling mechanisms come to a draw. *Chromosoma* 114, 310–318.
- Rooney, C., White, G., Nazgiewicz, A., *et al.*, 2010. The Rac activator STEF (Tiam2) regulates cell migration by microtubule-mediated focal adhesion disassembly. *EMBO Reports* 11, 292–298.
- Rusan, N.M., Tulu, U.S., Fagerstrom, C., Wadsworth, P., 2002. Reorganization of the microtubule array in prophase/prometaphase requires cytoplasmic dynein-dependent microtubule transport. *Journal of Cell Biology* 158, 997–1003.
- Saito, K., Hama, K., 1982. Structural diversity of microtubules in the supporting cells of the sensory epithelium of guinea pig organ of Corti. *Journal of Electron Microscopy (Tokyo)* 31, 278–281.
- Samora, C.P., Mogessie, B., Conway, L., *et al.*, 2011. MAP4 and CLASP1 operate as a safety mechanism to maintain a stable spindle position in mitosis. *Nature Cell Biology* 13, 1040–1050.
- Sharp, D.J., Ross, J.L., 2012. Microtubule-severing enzymes at the cutting edge. *Journal of Cell Science* 125, 2561–2569.
- Sirajuddin, M., Rice, L.M., Vale, R.D., 2014. Regulation of microtubule motors by tubulin isoforms and post-translational modifications. *Nature Cell Biology* 16, 335–344.
- Snyder, J.A., McIntosh, J.R., 1975. Initiation and growth of microtubules from mitotic centers in lysed mammalian cells. *Journal of Cell Biology* 67, 744–760.
- Steinmetz, M.O., 2007. Structure and thermodynamics of the tubulin-stathmin interaction. *Journal of Structural Biology* 158, 137–147.
- Stepanova, T., Smal, I., Van Haren, J., *et al.*, 2010. History-dependent catastrophes regulate axonal microtubule behavior. *Current Biology* 20, 1023–1028.
- Straube, A., 2011. How to measure microtubule dynamics? *Methods in Molecular Biology* 777, 1–14.
- Straube, A., Brill, M., Oakley, B.R., Horio, T., Steinberg, G., 2003. Microtubule organization requires cell cycle-dependent nucleation at dispersed cytoplasmic sites: Polar and perinuclear microtubule organizing centers in the plant pathogen *Ustilago maydis*. *Molecular Biology of the Cell* 14, 642–657.
- Straube, A., Hause, G., Fink, G., Steinberg, G., 2006. Conventional kinesin mediates microtubule-microtubule interactions *in vivo*. *Molecular Biology of the Cell* 17, 907–916.
- Subramanian, R., Wilson-Kubalek, E.M., Arthur, C.P., *et al.*, 2010. Insights into antiparallel microtubule crosslinking by PRC1, a conserved nonmotor microtubule binding protein. *Cell* 142, 433–443.
- Suprenant, K.A., Dean, K., McKee, J., Hake, S., 1993. EMAP, an echinoderm microtubule-associated protein found in microtubule-ribosome complexes. *Journal of Cell Science* 104, 445–450.
- Tanenbaum, M.E., Vale, R.D., McKenney, R.J., 2013. Cytoplasmic dynein crosslinks and slides anti-parallel microtubules using its two motor domains. *Elife* 2, e00943.
- Tran, P.T., Marsh, L., Doye, V., Inoue, S., Chang, F., 2001. A mechanism for nuclear positioning in fission yeast based on microtubule pushing. *Journal of Cell Biology* 153, 397–411.
- Van der Vaart, B., Akhmanova, A., Straube, A., 2009. Regulation of microtubule dynamic instability. *Biochemical Society Transactions* 37, 1007–1013.
- Vitre, B., Coquelle, F.M., Heichette, C., *et al.*, 2008. EB1 regulates microtubule dynamics and tubulin sheet closure *in vitro*. *Nature Cell Biology* 10, 415–421.
- Warren, R.H., 1974. Microtubular organization in elongating myogenic cells. *Journal of Cell Biology* 63, 550–566.
- Waterman-Storer, C.M., Salmon, E.D., 1998. Endoplasmic reticulum membrane tubules are distributed by microtubules in living cells using three distinct mechanisms. *Current Biology* 8, 798–806.
- Yajima, H., Ogura, T., Nitta, R., *et al.*, 2012. Conformational changes in tubulin in GMPCPP and GDP-taxol microtubules observed by cryoelectron microscopy. *Journal of Cell Biology* 198, 315–322.
- Yu, W., Baas, P.W., 1994. Changes in microtubule number and length during axon differentiation. *Journal of Neuroscience* 14, 2818–2829.
- Yu, W., Schwei, M.J., Baas, P.W., 1996. Microtubule transport and assembly during axon growth. *Journal of Cell Biology* 133, 151–157.
- Zhang, D., Grode, K.D., Stewman, S.F., *et al.*, 2011. *Drosophila* katanin is a microtubule depolymerase that regulates cortical-microtubule plus-end interactions and cell migration. *Nature Cell Biology* 13, 361–370.
- Zhang, Q., Fishel, E., Bertroche, T., Dixit, R., 2013. Microtubule severing at crossover sites by katanin generates ordered cortical microtubule arrays in *Arabidopsis*. *Current Biology* 23, 2191–2195.
- Zovko, S., Abrahams, J.P., Koster, A.J., Galjart, N., Mommaas, A.M., 2008. Microtubule plus-end conformations and dynamics in the periphery of interphase mouse fibroblasts. *Molecular Biology of the Cell* 19, 3138–3146.

Relevant Websites

- <http://celldynamics.org/celldynamics/research/cytokinesis/index.html>
Center for Cell Dynamics.
- <http://www.ibiology.org/ibioseminars.html>
iBiology.
- <http://mechanochemistry.org/>
Mechanochemistry.org.
- <http://mitchison.med.harvard.edu/protocols.html>
Mitchison Lab.
- <http://valelab.ucsf.edu/external/research/mt-mitosis/mt-mitosis-mt-regulators.html>
Vale Lab.

Actin Assembly Dynamics and Its Regulation in Motile and Morphogenetic Processes

J Pernier, P Montaville, and M-F Carlier, Cytoskeleton Dynamics and Motility Group, CNRS, Gif-sur-Yvette, France

© 2016 Elsevier Inc. All rights reserved.

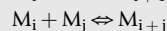
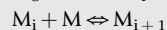
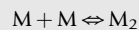
Glossary

Barbed end and pointed end of the actin filament The polarity of the actin filament is visualized in electron microscopy by decoration with the motor domain of myosin. Myosin heads bind the side of each subunit making an angle with the axis of the filament that generates arrowhead structures, defining the 'barbed' and the 'pointed' ends. It has been established that the barbed end is the most dynamic (rapidly associates with monomeric actin, and rapidly disassembles), while the pointed end has slow dynamics.

Capping proteins Bind the filament end in a way that prevents association of G-actin and disassembly of F-actin. Most capping proteins act on barbed ends.

Focal adhesions Complex structures by which links are established between the extracellular substrate and the internal actin cytoskeleton of the adhering live cell. Focal adhesions are mechanosensitive, their size and molecular composition vary under tensile strength.

Isodesmic self-assembly A polymerization process defined by one single equilibrium reaction for formation of dimers, trimers, ... from monomeric units M , as follows:



$$K = [M_i] \cdot [M] / [M_{i+1}] = [M_i] \cdot [M_j] / [M_{i+j}]$$

Lamellipodia and filopodia Cellular extensions of the plasma membrane that adopt lamellar sheet-like or thin spiky morphology respectively, and are generated and maintained by actin assembly.

Persistence length of a polymer The length at which the polymer loses memory of its original tangential direction. The persistence length L_p is correlated to the stiffness S of the polymer, $L_p = S/k_B T$. The persistence length of standard F-ADP-actin filaments is 7–8 μm . It can increase up to 20 μm in the presence of tropomyosin, or in the transient F-ADP-Pi state of ATP hydrolysis.

Processive assembly of filaments Defines the function of formins, which remain individually bound to the barbed end of the polymerizing filament through many cycles of self-assembly.

Sequestering proteins Proteins that bind monomeric actin in a complex that cannot assemble in filaments.

Severing proteins Proteins able to fragment filaments, making them shorter without changing the mass of assembled actin in this process. At the molecular level, severing implies transient binding of the protein (or drug) to the side of one filament followed by a structural change that enhances the affinity of the protein (or drug) for the final state, which consists of two filaments. In the final state the severing protein is often bound to one of the two ends created by the cut. For instance gelsolin when added to a solution of F-actin, severs rapidly to remain very tightly bound to the barbed end created by the cut. This reaction is actually faster than direct binding of gelsolin to the existing barbed ends which are present at 4 orders of magnitude lower concentration than the sites along the side of filaments. Some marine-organism derived macrolides also act as filament severers. Some severers alternately strip off an actin subunit from the severed filament.

Stochastic optical reconstruction microscopy

(STORM) A light microscopy super resolution imaging method that emerged, together with related PALM (photoactivated localization microscopy) and FPALM (fluorescence photoactivation localization microscopy) in 2006. These methods, derived from the pioneering work Werner Heisenberg (1930) and in depth mathematical developments, enable localization of individual molecules with high precision (in the range of the nanometer) due to the high number of photons that are captured following photoactivation by laser illumination of single molecules labeled with a probe, capable of photoswitching between a brightly fluorescent and a dark state. Spatial resolution is determined by the square root of the number of captured photons before inactivation of the probe, which is the critical point in the method. Probes must have a high contrast ratio (i.e., fluorescence ratio before and after photoactivation) to minimize the background.

WASP Protein responsible for the Wiskott–Aldrich syndrome – A syndrome of X chromosome-linked recessive immune deficiency characterized by thrombocytopenia and eczema. Immune deficiency is linked to mutations that are found mostly in the N-terminal moiety of the WAS protein. WAS proteins form a family that all share the ability to bind actin and Arp2/3 complex and catalyze filament branching by the same molecular mechanism.

Introduction

Actin is one of the most abundant and conserved protein in the eukaryotic kingdom. First identified as one of the main

component and partner of myosin in muscle (Straub, 1942, 1943), actin is also present in all non-muscle cells, where it stands, together with microtubules and intermediate filaments, as one of the major polymers of the cell cytoskeleton.

The essential property of actin is to self-assemble into non-covalent, dynamic, helical, and polar filaments. The dynamic nature of actin self-assembly is at the origin of its biological function in cell migration, cancer cell invasion, cytokinesis, endo- and phago-cytosis, wound healing, immune response, synaptic plasticity and a large number of morphogenetic and axis patterning processes (Lauffenburger and Horwitz, 1996; Pollard and Cooper, 2009; Bugyi and Carlier, 2010; Salbreux *et al.*, 2012; Heisenberg and Bellaiche, 2013).

In contrast to microtubules, which, in animal cells, are organized in one major radial array from the microtubule organizing center (MTOC), actin filaments are engaged in various structurally and functionally different ‘modules.’ These structures develop elementary protrusive, adhesive, and contractile activities (Small *et al.*, 1996; Le Clainche and Carlier, 2008; Lammermann and Sixt, 2009), which are coordinated in more complex biological functions. In response to signals in particular through the small GTPases Rho, Rac and Cdc42, the actin cytoskeleton is remodeled rapidly, via local disassembly and site-directed nucleation-growth reactions, generating shape changes (Ridley *et al.*, 2003; Ridley, 2011; Hall, 2012; Chaki and Rivera, 2013). The dynamic architecture of the actin cytoskeleton thus enables the cells to move, feed, divide, interact with neighbors, and organize intracellular transport of particles. In the special case of plant cells which have no centrosome and possess a rigid wall, the cytoskeleton (both actin and microtubules) plays a role in shape maintenance, traffic, and cell growth.

Outstanding questions about how actin drives cell motility have attracted approaches from various disciplines from structural biology to cell biology and physics in the past 20 years. This article focuses on the following issues: What intrinsic properties of actin are at the origin of its functions? How are the nucleation and the assembly–disassembly dynamics of actin controlled? How do the molecular reactions that control actin assembly convert into macroscopic motile processes at the cell scale?

Actin: Structural and Mechanical Properties

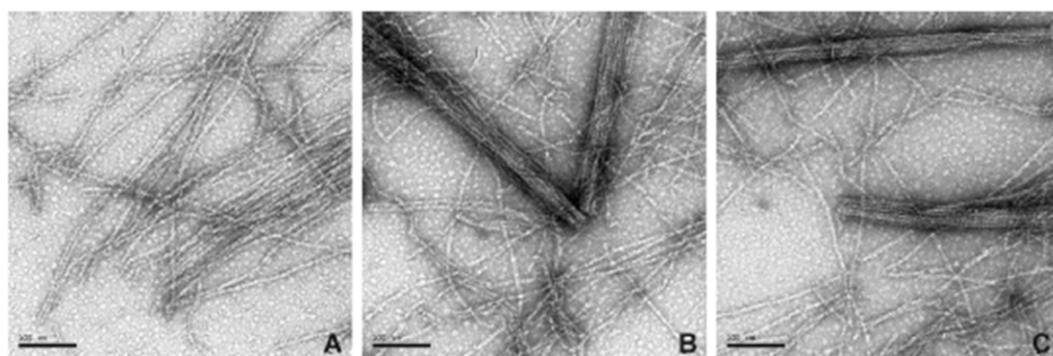
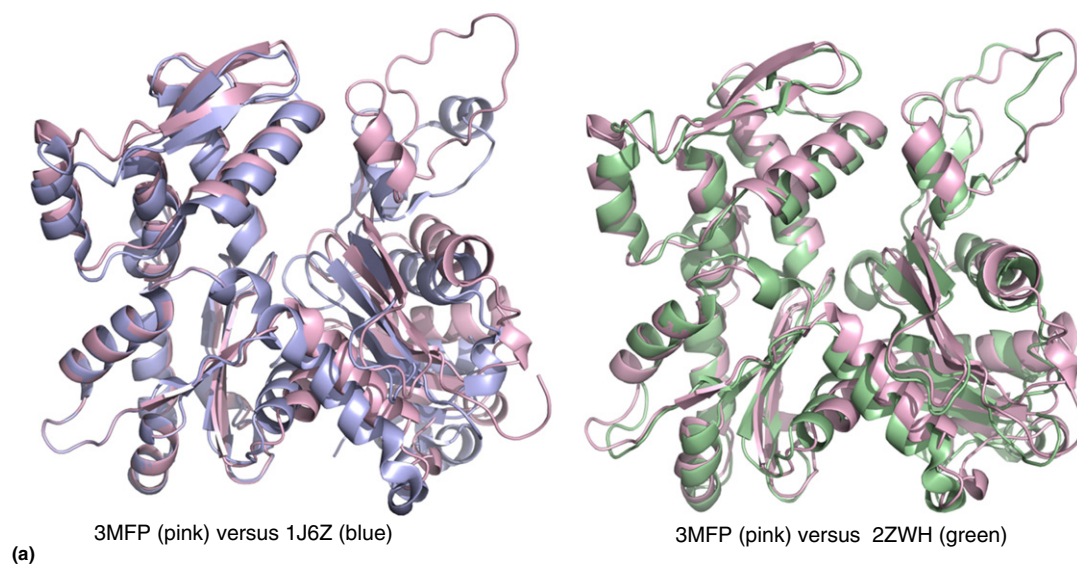
Actin is a 42 kDa globular acidic protein, its structure is known at atomic resolution (Kabsch *et al.*, 1990). Actin is organized in four subdomains (Figure 1). The four subdomains define a cleft with subdomains 1 and 3 at the bottom of the cleft (actin barbed face), subdomains 2 and 4 at the top of the cleft (pointed face). ATP is bound in the cleft, in complex with a divalent metal ion (Mg^{++} under physiological conditions; Valentin-Ranc and Carlier, 1989). In low-ionic strength (non-physiological) conditions, actin remains monomeric (G-actin, standing for globular actin) due to repulsive electronegative interactions. At physiological ionic strength (0.1–0.15 M KCl, 1 mM Mg^{++}), repulsive forces are abolished and actin self-assembles in filaments (F-actin, standing for filamentous actin) that may contain thousands of subunits (370 subunits = 1 μm). The structural change linked to the G to F transition involves a propeller-twist movement of the outer and inner domains of the actin subunit around an axis roughly perpendicular to the filament axis, which makes actin flatter in the F state (Oda *et al.*, 2009; Figure 1(a)).

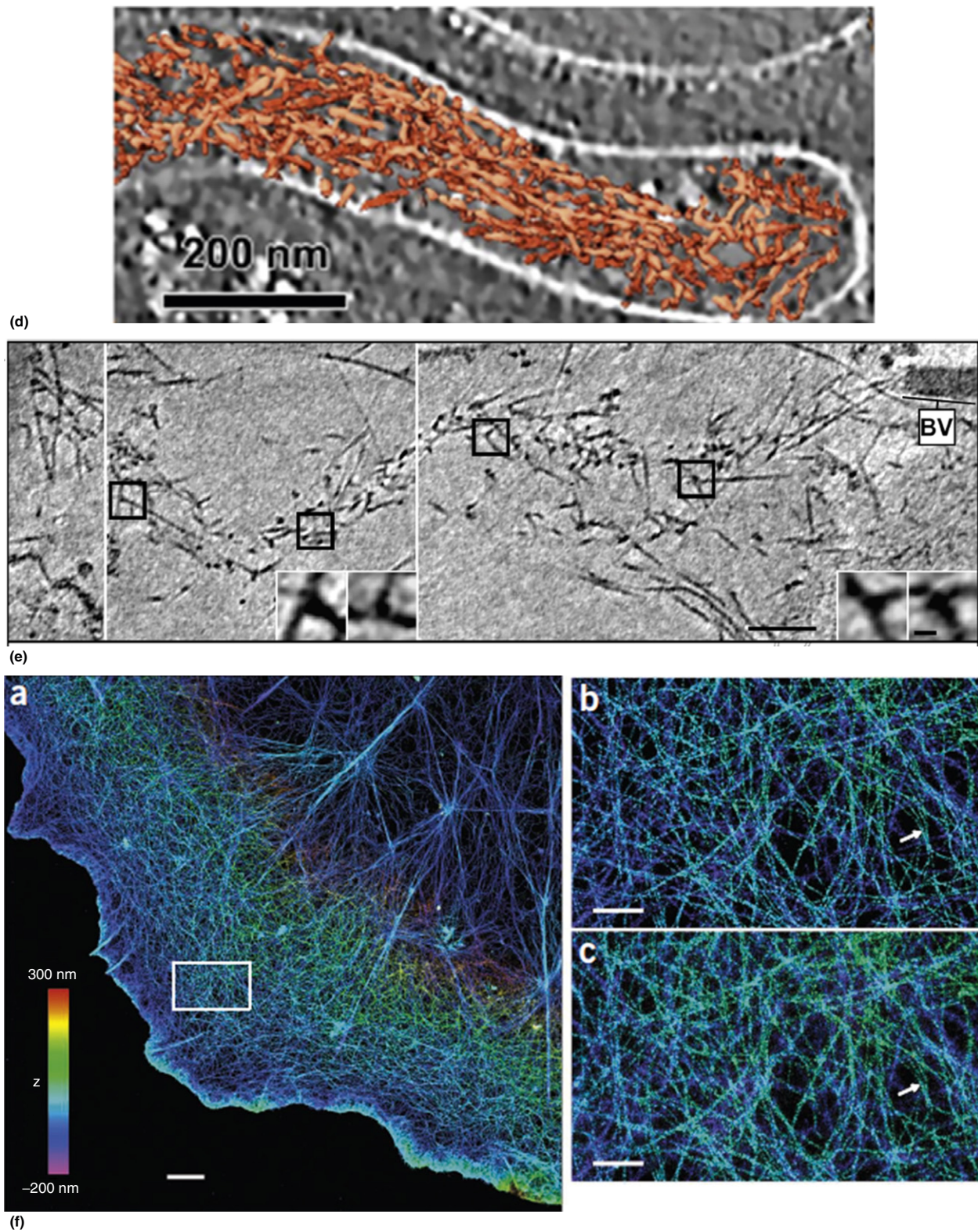
In agreement with the original atomic model of the actin filament (Holmes *et al.*, 1990) and the improved resolution structure derived from the fiber diffraction analysis (Oda *et al.*, 2009), the highest resolution structure of the actin filament in cryo-EM (Fujii *et al.*, 2010; Figure 1(b)) shows a polar helical structure in which assembled subunits are connected by a left-hand rotation of 167° and axial rise of 27.6 Å, describing the ‘genetic’ or short-pitch helix. The filament is also described as a two strand, right-handed, long-pitch helix that displays a crossover distance of 36 nm. The acidic N-terminal region of actin is exposed at the surface of the filament, accounting for its polyelectrolyte nature (Fujii *et al.*, 2010). Binding of positively charged ligands (cations, basic proteins, poly-L-lysine...) thus promotes association of actin filaments in rigid bundles (Tang and Janmey, 1996).

The highly ordered structure of actin filaments in very thin ice cryo-EM contrasts with the reported structural disorder in other electron microscopy observations, and the view that the structural disorder is used by regulators to generate multiple structural states and associated specified functions (Galkin *et al.*, 2003, 2010, 2011).

The structural organization of terminal subunits at the two ends of the filament may differ from the core subunits, since some of the filament actin–actin contacts are exposed to the solvent. How the potential structural plasticity of the terminal subunits is modulated by filament end-binding proteins is a timely issue. Only the structure of Capping Protein bound to barbed ends has been solved by combining crystallographic structure of free Capping Protein and cryo-electron microscopy images of capped filaments (Narita *et al.*, 2006). The cryo-EM structure of the pointed end (Narita *et al.*, 2011) reveals a tilted conformation of the terminal subunit toward the penultimate subunit, enhancing the strength of actin–actin contacts, thus accounting for the slow dynamics at this end. This structure has not been used so far in modeling pointed end binding proteins, but will be useful in understanding how proteins like tropomodulin associate with and regulate filament pointed ends. So far only crystal structures of isolated domains of tropomodulin in complex with monomeric actin are available (Rao *et al.*, 2014), from which a working model of tropomodulin bound to the pointed end is proposed. Similarly, the crystal structure of the FH2 domain of formin Bni1 in complex with actin is known, from which a working model is proposed for barbed end-bound formins (Otomo *et al.*, 2005), but no high-resolution cryo-EM structures are available for filament end-bound formins nor tropomodulin. However, often conclusions derived from crystal structures of complexes of monomeric actin with regulators do not match conclusions derived from functional or structural solution studies, due to drawbacks coming from crystallogenes conditions or crystal packing. The upcoming spectacular improvements in resolution brought by single detector camera in cryo-EM analysis of large macromolecular assemblies open a new age in relating structure and function.

Filaments can also be observed *in vitro* by negative staining electron microscopy (Figure 1(c)) and *in vivo* by cryo-electrotomography (Medalia *et al.*, 2007; Rigort *et al.*, 2012; Mueller *et al.*, 2014; Figures 1(d) and 1(e)), as well as using high-resolution STORM (Bates *et al.*, 2013; Xu *et al.*, 2012; Figure 1(f)).





Actin filaments are semiflexible polymers, i.e., their persistence length L_p is of the same order of magnitude ($8 \mu\text{m}$) as their average length (Isambert *et al.*, 1995). The flexural rigidity

of filaments is affected by bound ligands, reflecting the structural plasticity. Filaments are more rigid when tropomyosin or phalloidin is bound or when the filaments are made of

ATP-actin rather than ADP-actin subunits ($L_p = 18 \mu\text{m}$). In contrast, filaments are more flexible when ADF/cofilin is bound, ($L_p = 3\text{--}4 \mu\text{m}$), linked to a lower stability of actin-actin bonds in the filament (Carlier *et al.*, 1997, 1999; Pfaendtner *et al.*, 2010).

While filaments resist traction forces, they are easily fragmented by torsion (Yasuda *et al.*, 1996). Filaments are chiral polymers, not simple rods, hence it is possible that bending leading to severing, caused by either brownian motion or accessory proteins, affects the twist of the filament (Prochniewicz *et al.*, 2005; Matsushita *et al.*, 2012). In conclusion, the mechanosensitivity of actin, intimately linked to its structural plasticity, is at the heart of its multiple versatile functions in cell motility. It is used by several actin binding proteins. Additionally, protein machineries that couple actin filaments to membranes often are mechanosensitive and transduce mechanical force into biochemical regulation of actin dynamics. An interdisciplinary approach has actively developed in the recent years to understand the regulation of actin dynamics at a multiscale level.

Actin Self-Assembly: Nucleation-Growth, Thermodynamics, Polarity, ATP Hydrolysis

Bulk Solution Studies

The elementary helical unit of the filament is an actin trimer (Figure 1(b)). Consistently, spontaneous assembly triggered by a sudden increase in ionic strength is well described by a nucleation-growth process. The trimer is the smallest oligomer that shows a higher tendency to bind a 4th actin subunit than to dissociate (Wegner and Engel, 1975; Oosawa and Asakura, 1975; Korn, 1982). Nucleated self-assembly contrasts with the isodesmic self-assembly of linear polymers, described by a single type of interaction between subunits, leading to an equilibrium between monomers, dimers, trimers, and higher order oligomers.

The self-assembly of ADP-actin is well described by the laws of reversible polymerization (Oosawa and Asakura, 1975). Filaments grow faster at one end (the barbed end) than at the other (the pointed end). At the end of the polymerization process, actin exists in equilibrium between monomers

(ADP-G-actin) and filaments (ADP-F-actin) of several hundred subunits. The concentration of monomers at equilibrium with filaments is the equilibrium monomer-polymer dissociation constant $K_D = k_-/k_+$, called critical concentration for assembly, because actin remains monomeric below it. Filament fragmentation and re-annealing reactions also occur infrequently in F-actin solutions depending on filament number, length, and thermal brownian motion (Wegner, 1982; Schmolle *et al.*, 2011).

In nature, under physiological conditions, ATP is tightly bound to G-actin and is hydrolyzed irreversibly to ADP in a single turnover reaction following incorporation of an actin subunit in the filament, within two consecutive steps, cleavage of the gamma-phospho-ester bond followed by slower release of inorganic phosphate (Carlier and Pantaloni, 1986; Korn *et al.*, 1987; Murakami *et al.*, 2010). Although ATP hydrolysis is not required for polymerization, the assembly of actin filaments dissipates energy (as well as the assembly of microtubules from GTP-tubulin, which is associated with GTP hydrolysis). The fact that actin is an irreversible ATPase is pivotal in all its functions in cell motility (Carlier, 1989; Pantaloni *et al.*, 2001). Hydrolysis of ATP weakens actin-actin interactions in the filament. Due to the slow rate of Pi release, ATP or ADP-Pi is bound to terminal subunits at the rapidly growing end (barbed end) while ADP is bound to pointed end subunits, creating an energetic (stability) difference between the two ends of filaments. In ATP, filaments are thus maintained at a steady state – not at an equilibrium. A polarized flux of subunits through the filament, from the barbed to the pointed ends, called ‘treadmilling,’ describes the assembly dynamics of filaments both *in vitro* (Wegner, 1976) and *in vivo* (Lai *et al.*, 2008). A stationary concentration of ATP-G-actin allows the net flux of assembly at barbed ends to balance the net flux of filament disassembly at pointed ends (Wegner, 1976; Narita *et al.*, 2011). This process is highly regulated *in vivo* by proteins that bind differentially to the ATP- and ADP-bound forms of G and F-actin and by proteins that bind and affect the reactivity of either the barbed or pointed end, hereby affecting the rates of all elementary steps involved in the treadmilling cycle (Figure 2 and following sections).

Another mechanism differing from treadmilling, such as a sudden increase in G-actin availability, has been evoked to account for the observed growth of actin filaments in cellular

Figure 1 (a) Atomic structure of monomeric actin (G-actin-PDB code:1J6Z) and conformational change linked to its incorporation in a filament (F-actin PDB codes:3MFP) – local structural variability of actin monomer in the filament, adapted from Oda, T., Iwasa, M., Aihara, T., Maeda, Y., Narita, A., 2009. The nature of the globular- to fibrous-actin transition. *Nature* 457 (7228), 441–445 and Fujii, T., Iwane, A.H., Yanagida, T., Namba, K., 2010. Direct visualization of secondary structures of F-actin by electron cryomicroscopy. *Nature* 467 (7316), 724–728. (b) Structure of the actin filament (F-actin) derived from X-ray fiber diffraction – adapted from Oda, T., Iwasa, M., Aihara, T., Maeda, Y., Narita, A., 2009. The nature of the globular- to fibrous-actin transition. *Nature* 457 (7228), 441–445. An elementary actin trimer assembly is encircled with a black line. (c) Actin filaments assembled *in vitro* observed by electron microscopy (negative staining). Bundles (B,C) are induced by fesselin – adapted with permission from Schroeter, M.M., Orlova, A., Egelman, E.H., Beall, B., Chalovich, J.M., 2013. Organization of F-actin by Fesselin (avian smooth muscle synaptopodin 2). *Biochemistry* 52 (29), 4955–4961. Copyright 2013, American Chemical Society. Bar 100 nm. (d) Cryo-electron tomogram of actin filaments *in vivo* in filopodia – adapted from Medalia, O., Beck, M., Ecke, M., *et al.*, 2007. Organization of actin networks in intact filopodia. *Current Biology* 17 (1), 79–84. Actin filaments highlighted in red-orange. (e) Cryo-electrotomogram of branched filaments initiated at the surface of the baculovirus propelling in a B16 melanoma cell – adapted from Mueller, J., Pfanzelter, J., Winkler, C., *et al.*, 2014. Electron tomography and simulation of baculovirus actin comet tails support a tethered filament model of pathogen propulsion. *PLoS Biology* 12 (1), e1001765. (f) STORM image of actin arrays *in vivo* in lamellipodia – adapted with permission from Xu, K., Babcock, H.P., Zhuang, X., 2012. Dual-objective STORM reveals three-dimensional filament organization in the actin cytoskeleton. *Nature Methods* 9 (2), 185–188. Copyright 2012, Rights Managed by Nature Publishing Group. Z-positions are color coded purple (close to substrate) to red (far from substrate). (b) and (c) are enlarged views of the boxed area in (a), obtained using dual objective (b) or single objective (c).

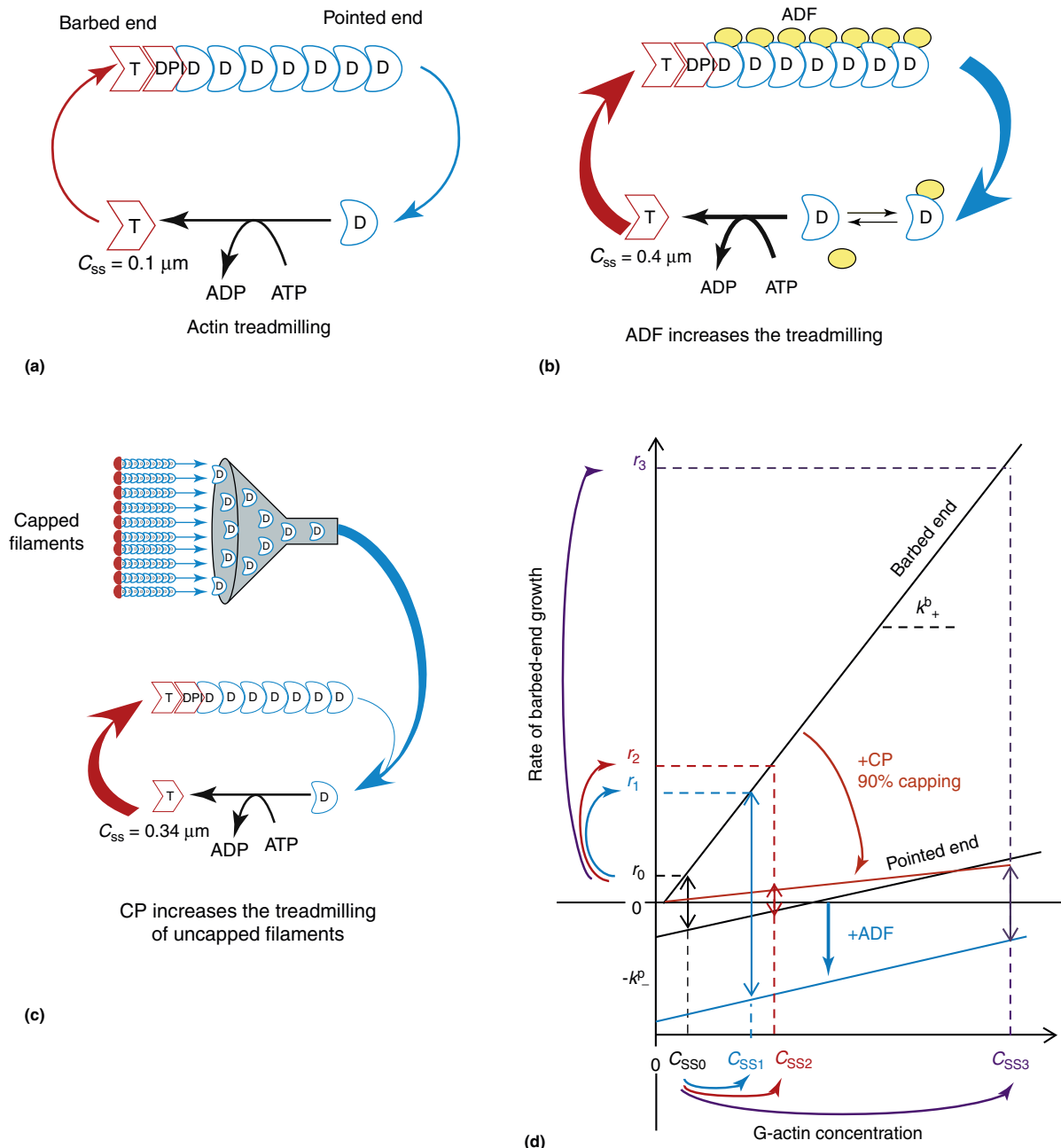


Figure 2 Regulation of treadmilling by ADF and capping proteins – adapted from Pantaloni, D., Le Clainche, C., Carlier, M.F., 2001. Mechanism of actin-based motility. *Science* 292 (5521), 1502–1506. (a) Pure actin, slow treadmilling, and low value of C_{ss} . (b) Effect of ADF, increase in C_{ss} via enhanced rate of disassembly. (c) Effect of capping proteins, enhanced rate of individual free barbed-end growth. (d) Graphic representation of the effects of ADF and capping proteins on treadmilling. Lines represent the dependence of the rates of filament elongation at barbed and pointed ends on the concentration of ATP-G-actin. The values of C_{ss} are C_{ss0} , which satisfy equal rates of barbed-end growth (r_1) and pointed-end disassembly (double headed arrow). Black lines: actin alone (characterized by C_{ss0} and r_0). Blue lines: the effect of 90% capping on the barbed-end kinetics (characterized by C_{ss1} and r_1). Red lines: effect of ADF on pointed-end kinetics (characterized by C_{ss2} and r_2). Purple lines: in the presence of both ADF and capping proteins (characterized by C_{ss3} and r_3).

processes. The monitored stationary rates of protrusion or pathogen propulsion in cytoplasm, for extended periods of time, argue against this mechanism, which would generate transient movements and exponentially slowing down of actin

assembly to arrest as observed in a test tube when G-actin is induced to polymerize at time zero. It is difficult to escape the view of a generalized regulated treadmilling mechanism to explain actin-based movements.

Single Filament Assembly Dynamics: Total Internal Reflection Fluorescence Microscopy

In the early 2000s, the use of evanescent wave microscopy allowed light microscopy observation of individual filaments growing from fluorescently labeled actin within a short distance (200 nm) from the coverslip, even in the presence of a large amount of labeled monomers in the chamber (Funatsu *et al.*, 1995; Kuhn and Pollard, 2005; Watanabe, 2012; Breitsprecher *et al.*, 2009). The method avoids the limitations due to the averaging of many filaments in bulk solution measurements. It is essential in monitoring processive assembly motors or the behavior of heterogeneous populations of filaments. Recent improvements of the standard total internal reflection fluorescence (TIRF) method have been brought by implementing microfluidics (Jegou *et al.*, 2011; Carlier *et al.*, 2014). Filaments are anchored at one end only at the surface of the coverslip, aligned with a microflow whose contents can be changed in 1 s, allowing kinetic studies. Molecular mechanisms by which proteins control filament assembly by binding barbed or pointed ends are addressed at the filament scale, and accurate statistical analysis of kymographs is greatly facilitated by alignment of filaments (Jegou *et al.*, 2011; Montaville *et al.*, 2014). The method thus represents the future of actin biochemistry at the single molecule level. Mechanical properties of actin filaments and of mechanosensitive protein machineries that control barbed-end growth such as formins is also addressed, using the flow to develop traction forces (Jegou *et al.*, 2013). Important future developments of this method are expected, as expressed in detail in perspective and review papers (Jegou *et al.*, 2011; Carlier *et al.*, 2014).

Regulation of Actin Assembly by Associated Proteins

A remarkable feature of the organization of the actin cytoskeleton is the specific set of actin regulators associated with each actin module within the cytoplasm. The mechanism underlying this functional and structural compartmentation is not fully understood. It is possible that site-directed nucleation of filaments by a given protein machinery directly recruits a set of associated regulators of the filament, or is linked to some signaling process that will impose a specific cascade of actin associated proteins binding in a cooperative fashion. The resulting exclusion of other sets of regulators may thus build a distinct composition/structure of a given actin meshwork. How different actin-binding proteins affect the structure and dynamic properties of actin individually and how they act in synergy or in antagonism provides insight into the emergence of biological complexity.

A general principle of the regulation of actin assembly by accessory proteins is based on their effect on the monomer-polymer steady state. Proteins that bind G-actin specifically displace the equilibrium toward the monomer and cause filament disassembly. Proteins that bind F-actin specifically displace the equilibrium toward the polymer. In addition to these simple cases, proteins bind the filament ends and affect their reactivity, i.e., modulate the rate parameters for actin assembly or disassembly at this end. Some proteins may also bind G-actin and not F-actin, but their complex with G-actin

may associate to filament ends with distinct rate parameters, hence they transiently interact with F-actin ends. Finally, some proteins may bind both G-actin and F-actin and change the dynamic monomer-polymer equilibrium.

The complexity of the regulation of actin assembly is enhanced when the regulators show some specificity for the ATP- or ADP-bound forms of G-actin and/or F-actin, or when they affect nucleotide exchange on G-actin. The relatively slow hydrolysis of ATP on F-actin acts as a timer that causes specific enrichment of old regions of the filaments in proteins that specifically bind ADP-F-actin, such as ADF/cofilin.

Some proteins are able to sever filaments. The severing mechanism implies that these proteins harbor two actin binding subsites, one on the side of filaments and one at the barbed face of actin, which enables them to strengthen their binding by cutting the filament and remaining bound to either the barbed end created by the cut (for instance gelsolin or Spire, 4 WH2 repeat protein), or to an actin monomer stripped off the filament (for instance Cordon-Bleu, a 3 WH2 repeat protein). Severing multiplies the number of filaments. Depending on the context, severing may either accelerate polymerization (as observed with Cordon-Bleu, which acts as a dynamizer of actin assembly in key developmental processes), or accelerate depolymerization and facilitate destruction of a complex filament meshwork.

Nucleation of actin filaments occurs mostly in a site-directed, signal-mediated fashion. Newly formed filaments often grow in a polarized fashion. Thus protein machineries which nucleate filaments often also bind their barbed ends to organize polarized growth. This is the case of formins and of proteins, which multiply rather than nucleate filaments by branching them (WASP family proteins, see following sections).

While a large number of actin regulators exhibit one of the above well-characterized individual activities, more recently discovered proteins display multifunctional regulation of actin assembly. These versatile regulators often combine several actin-binding modules of different functions (see Section 'The WASP Homology 2 (WH2) Domain: A Versatile Module').

Research on the regulation of actin assembly has been facilitated by the use of small molecules or drugs that mimic regulatory proteins, acting either by binding G-actin and preventing polymerization like thymosins, thus promoting depolymerization of filaments (Latrunculins: Spector *et al.*, 1983) or by binding G-actin as well as capping barbed ends (Cytochalasins) or stabilizing filaments abolishing dynamic turnover (phalloidin, jasplakinolide: Holzinger, 2001), or by slowing down dynamics (cyclic polyamines: Nedeva *et al.*, 2013). Chemicals also regulate actin dynamics by binding Arp2/3 complex (CK666: Nolen *et al.*, 2009; Hetrick *et al.*, 2013) or WASP proteins (Wiskostatin: Leung *et al.*, 2006) or formins (SMIFH2: Rizvi *et al.*, 2009).

The following sections illustrate the various modes of regulation of actin filament assembly dynamics by specific regulatory proteins.

G-Actin Binding Proteins

Proteins that bind specifically monomeric actin (Table 1(a)) may form a complex that either is non-polymerizable or displays polymerization properties that differ from G-actin.

Table 1 Main regulators of actin assembly

ABPs	Biochemical functions	Cellular localization	Regulators	Accession number
(a) G-Actin binding proteins				
b-Thymosins (Tß4,9,10,15) Ciboulot	Sequester G-actin	Blood, thymus, and hematopoietic cells brain	–	NP 066932.1 NP 066926.1 NP 068832.1 NP 919305.2
Profilins (three isoforms, profilin 1 most abundant and ubiquitous)	Promote actin monomer addition at barbed ends, enhance nucleotide exchange on G-actin, and enhance barbed-end depolymerization	Cytoplasm	PIP ₂	NP 005013.1 NP 444252.1 NP 001025057.1 NP 955378.1
WH2 domains	Sequester G-actin, may display profilin function, and also multifunctional regulators of barbed-end dynamics	Inserted in a large number of proteins involved in motile processes	–	NA (see Paunola et al., 2002)
ADF/cofilin Destrin Depactin Actophorin ADFH proteins: Drebrin Twintilin Abp1p coactosin	Bind ADP-actin specifically (both G- and F-actin)	Ubiquitous and essential	Lim kinase and Slingshot phosphatase	NP 005498.1 NP 006861.1 P20690.1 XP 003190377.1 NP 004386.2 NP 001229326.1 NP 010012.1 XP 009654332.1
Vitamin D-binding protein Gc globulin	Sequester G-actin	Plasma and cerebrospinal fluid	Vitamin D and C5A complement	NP 000574.2
(b) Barbed end-capping proteins				
Gelsolin	Severs actin filament	Blood, cytosol, and mitochondria	Calcium and PIP ₂	NP 000168.1
Brevin Severin Villin	Severs actin filament Severs actin filament Severs actin filament	Plasma <i>Dictyostelium discoideum</i> Epithelium	Calcium and tropomyosin Calcium PIP ₂	P06396.1 XP 001134515.1 NP 009058.2
Cap G	Caps barbed ends (calcium dependent)	Macrophages, nuclear and cytoplasmic	Calcium and PIP ₂	NP 001738.2
Capping protein (Cap Z)	Heterodimer caps barbed ends	Muscle	CP Interaction motif-containing proteins	NP 006126.1 NP 001193469.1
Vinculin	Weak barbed-end capper	Focal adhesion	Talin/actin	NP 003364.1
(c) Dimerized WH2 domains				
VASP	F-actin elongation, bundling, and uncapping	Focal adhesions and filopodia	AMPK, PKA, PKC	NP 003361.1
VopF	B-end uncapping and processive elongation	Intestinal epithelial cells	Profilin	YP 007301278.1
Pointed end binding proteins Tropomodulin DNase 1	Caps pointed ends Binds G-actin, caps pointed end	Muscle cells Zymogen granules	Tropomyosin PIP ₂	NP 003266.1 NP 005214.2
(d) Filament side binding proteins				
Tropomyosin	Binds filament	Muscle	–	NP 001018005.1 NP 003280.2 NP 003079.1
Fascin Fimbrin/plastin Espin Calponin (1–3)	Bundling activity Bundling activity Bundling activity Binds G and F-actin	Filopodia Focal adhesions Stereocilia Smooth muscle (1), Non-muscle cells (2), and Neurons, glial cells (3)	– Calcium and PIP ₂ Autoinhibition Calmodulin and Tropomyosin	NP 001138791.1 NP 113663.2 NP 001290.2 NP 004359.1 NP 001830.1
a-actinin (actinogelin)	Bundling activity	Muscle Z-disks, Non-muscle focal adhesions, and stress fibers	Vinculin, Zyxin, Integrin, PIP ₂ , and NMDA receptor and selectin	NP 001123476.1 NP 001094.1 NP 001095.2

(Continued)

Table 1 Continued

<i>ABPs</i>	<i>Biochemical functions</i>	<i>Cellular localization</i>	<i>Regulators</i>	<i>Accession number</i>
Scruin Ezrin/moesin/radixin	Bundling activity Binds filament	Acrosome bundle Microvilli and membrane ruffles in a wide variety of non-muscle cells	Calmodulin Calpain, PIP ₂ , PI3-K, PKA, RhoA, Cdc42, Dbl b2-integrin, and Integral membrane proteins	NP 004915.2 CAA86292 NP 001104547.1 NP 002435.1 NP 001247422.1
Spectrin (fodrin, calspectin)	Crosslinking and binding bundling activities	Submembranous cytoskeleton	Ca ²⁺ , PIP ₂ , lipids, ankyrin, band 4.1, and adducin	NP 003117.2 NP 001020029.1 NP 001123910.1 NP 003119.2
Talin	Binds filament	Heart, brain, and plasma membrane	PIP ₂ , PKC, Rap1A, and RIAM	NP 006280.3 NP 055874.2
(e) WASP family				
N-WASP	Filament branching	Filopodia	Cdc42 and PIP ₂	BAA20128
WASH	Activates arp2/3 and endosomal fission	Endosome	PIP3	NP 878908
WHAMM	Activates arp2/3	Golgi membrane and perinuclear region	–	NP 001073904
WAVE	Activates arp2/3	Filopodia	PIP3 and IRSp53	NP 001020107
JMY	Activates arp2/3	Brain tissue and neuronal cells	–	AAI30625
(f) Linkage BAR-domain containing proteins				
IRSp53	Formation of filopodia and lamellipodia	Neurons	Cdc42, Tiam1, and Eps8	Q9UQB8
CIP4/Toca-1	Scission and protrusion during cell migration	Filopodium	Cdc42	NP 001275891
Syndapin 1 or Paccin 1	Recruitment of dynamin, membrane constriction, and tubulation	Filopodium	PIP ₂	NP 001186512
MIM	Formation of plasma membrane protrusions	Liver, kidney, and Purkinje cells	Gli1 and Gli2 transcription factors	NP 001269900
ABBA-1	Lamellipodial dynamics	Central nervous system in certain glial cells	Rac1	Q765P7
FBP17	Endocytosis and induced membrane tubules <i>in vitro</i>	Podosome	Cdc42	NP 055848
Actin regulators in adhesion structures				
Talin	Integrin activation and cell adhesion	Heart, brain, and plasma membrane	PIP ₂ , PKC, Rap1A, and RIAM	AAF23322

The former class of so-called G-actin sequestering proteins generates a reservoir of unassembled subunits in equilibrium with free G-actin. The stationary concentration of free G-actin in the cell, buffered by monomer-polymer exchange, thus determines the size of this pool. The important consequence is that any regulator of actin filament assembly that causes a decrease in critical concentration lowers the pool of sequestered actin; conversely, capping proteins which cause an increase in critical concentration, elevate the amount of sequestered actin and lower the concentration of F-actin (Carlier and Pantaloni, 1994). β -Thymosins (5 kDa), most abundant (200–400 μ M) in blood and hematopoietic cells, are major representatives of the G-actin sequestering proteins in vertebrates (Safer and Nachmias, 1994; Sun *et al.*, 1995). β -Thymosins bind specifically ATP-G-actin, very poorly ADP-G-actin (Carlier *et al.*, 1993).

The drug Latrunculin A binds G-actin on top of the ATP binding cleft at the pointed face of actin and is a major depolymerizing drug that acts like a sequestering protein (Ayscough, 1998).

Profilin (15 kDa) is another major, abundant G-actin binding protein that also binds ATP-G-actin preferentially. The profilin-actin complex productively assembles in filaments at the barbed end exclusively, with kinetic parameters similar to G-actin (Pollard and Cooper, 1984). Following association of profilin-actin complex to the filament barbed end, barbed end-bound profilin dissociates leaving the barbed end free for further association of actin or profilin-actin. The same regulatory mode is often adopted by a protein module called WH2 (WASP-Homology 2 domain, see Section 'The WASP Homology 2 (WH2) Domain: A Versatile Module'). Profilin lowers the steady-state concentration of ATP-G-actin

(Pantaloni and Carlier, 1993; Yarmola and Bubb, 2006). Profilin is generally essential in formin function *in vivo* (Sagot *et al.*, 2002; Manseau *et al.*, 1996).

Proteins of the ADF/cofilin family bind preferentially ADP-actin both in the monomer and polymer form (see section below).

Barbed End-Capping Proteins

These represent one of the most functionally important proteins in the regulation of actin assembly (Table 1(b)). Strong cappers generally bind exclusively filament barbed-end subunits with very high-affinity (K_d in the 10^{-9} – 10^{-10} M range) and block-filament growth and depolymerization. Several classes of strong cappers exist, among which proteins of the gelsolin family (brevin, severin, villin, CapG: for recent review, Nag *et al.*, 2013) and of the CapZ family (Capping Protein beta2; for review Cooper and Sept, 2008). Proteins that contain WH2 repeats, like Spire, may also work as barbed-end cappers. Capping of a large fraction of barbed ends in a population of filaments at steady state in ATP causes the amount of assembled F-actin to be controlled by pointed end dynamics only. The resulting increase in steady-state concentration of G-actin is amplified by sequestering proteins (Carlier and Pantaloni, 1994). Profilin becomes a G-actin sequesterer when all barbed ends are capped. Hence, capping proteins control the ratio of assembled and unassembled actin in cells. Barbed-end capping has to be highly regulated in cells. Proteins of the CARMIL family bind capping protein via a specific module called CPI (capping protein interaction) found in CapZIP, CIN85, CKIP-1, and FAM21, a subunit of the WASH complex, which controls capping protein allosterically and can uncap filament barbed ends. Myotrophin, in contrast, only sequesters capping protein (Hernandez-Valladares *et al.*, 2010; Takeda *et al.*, 2010). Regulation of Capping Protein by sequestration and uncapping is important since its cellular concentration is high (in the micromolar range in blood cells) while its optimum concentration in motility is at least one order of magnitude lower (Edwards *et al.*, 2014 for review).

Some proteins, like vinculin, present in focal adhesions, work as weak cappers, meaning that they bind weakly to barbed ends and only slow down barbed-end dynamics (Ramarao *et al.*, 2007).

The protein Eps8, which links signaling to actin, binds actin both on the sides of filaments with low affinity, inducing bundling, and caps barbed ends with high affinity (Disanza *et al.*, 2004), which supports its function in filopodia formation, synaptic plasticity, and intestinal morphogenesis.

The drug Cytochalasin D binds the barbed face of G-actin and thus causes F-actin depolymerization, however, in contrast with Latrunculin A, it also binds filament barbed ends with an affinity of 10^9 M $^{-1}$, and thus is a good tool as a substitute to capping proteins.

Barbed-End Trackers

Some proteins bind barbed ends specifically, but in contrast with cappers, they are permissive for filament assembly,

and regulate the reactivity (growth rate) of filament barbed ends by modulating the association rate constant of G-actin (Tables 1(c), 2(a), 2(b)). Their linkage to barbed ends may be more or less processive, that is the same molecule may remain associated with terminal subunits at the barbed ends through a variable number of consecutive steps of assembly or disassembly. Formins are the most remarkable representatives of this class of proteins. Formins are major regulators of nucleation and growth of actin filaments from profilin-actin in a large variety of motile and morphogenetic processes like cytokinesis, filopodia extension, which they mediate by site-directed assembly of several filaments associated in bundles (Goode and Eck, 2007 for review). Most formins actually catalyze processive filament barbed-end assembly at variable rates, which can be up to 10-fold faster than the standard rate of assembly at free barbed ends, the same molecule remaining bound to barbed ends for up to thousands of assembly rounds. Many but not all formins are also nucleators of filament barbed-end assembly; some formins nucleate only if associated to a ligand, for instance Bud6 affects nucleation of Bni1 and Bnr1 (Tu *et al.*, 2012). Formins are distributed in several classes, depending on their mode of regulation. A large number of formins (DRFs, standing for diaphanous-related formins), are regulated by membrane-bound small G-proteins of the Rho/Cdc42 family (Table 2(a)). The regulation of non-DRFs formins is less well known (Table 2(b)). Formins generally act in a site-directed fashion and maintain actively growing filament barbed ends attached to the membrane. The active moiety of formins comprises conserved formin homology domains 1 and 2 (FH1 and FH2) which bind profilin and actin respectively. Profilin is required for the *in vivo* function of most formins, and is required *in vitro* for fast processive growth (Romero *et al.*, 2004; Kovar *et al.*, 2006). The FH1 domain contains a variable number of proline repeats, which are thought to bind several profilin-actin complexes and bring them close to the barbed end (Paul and Pollard, 2008). However complexes of formins with a defined stoichiometry of profilin-actin have not been isolated. The processive property relies on the dimeric nature of formins. The FH2 domain crystal structure reveals a head-to-tail ring structure thought to embrace two terminal subunits at the barbed end (Xu *et al.*, 2004), which is supported by the structure of the FH2-actin complex (Otomo *et al.*, 2005). Processive tracking of barbed ends is suggested to be coupled to a cycle of transitions between two states of actin-bound FH2 at barbed ends (Goode and Eck, 2007 for review). In maintaining active polarized growth at filament barbed ends, formins antagonize barbed-end cappers and control the level of F-actin assembly at steady state accordingly. Formin rotation around the filament is associated with processive growth (Mizuno *et al.*, 2011). Finally, formins are mechanosensitive protein machines, which increase the rate of filament assembly when submitted to tensile forces of a few picoNewtons, compatible with the mechanical effects of membrane tension *in vivo* and forces involved in cytokinetic ring contractility (Jegou *et al.*, 2013).

Other barbed-end trackers include multimeric WH2 domain proteins such as Ena/VASP family of proteins and VopF/VopL (Pernier *et al.*, 2013). The exact molecular and structural mechanism of filament end-tracking is not well known.

Table 2 Diaphanous-related formins (DRFs) and non-DRFs

<i>Mammalian formins</i>	<i>Cellular functions</i>	<i>Intracellular localization</i>	<i>Cell type expression</i>	<i>Regulators</i>	<i>Accession number</i>
(a) DRF					
Daam1	Planar cell polarity (PCP), dynamin dependent endocytosis, axonal morphogenesis, and filopodia formation	Focal adhesion, cell cortex, and filopodia	Wide	Dvl1-3, Src, Syndapin, Toca-1, CIP4, FBP1, and Cdc42	NP 055807.1
Daam2					NP 001188356.1
Dia1 (DRF1)	Stress fiber formation, endosome motility, mitochondrial motility, adherens junction stability, vesicle trafficking, microtubule stabilization, cell motility, cytokinesis, phagocytosis, filopodia information, and microtubule attachment to kinetochores	Filopodia, lamellipodia, phagocytotic cup, nucleus, endoplasmic reticulum, focal adhesion, microtubules, golgi complex, and plasma membrane	Wide	ABI1, APC, CLIP170, Ga12/13, IQGAP1, RhoA, RhoB, RhoC, RhoD, EB1, Rif, DIP, POPX2, Pax6, YWK-II, PKD2, PKCz, PKCe, Src, IRsp53, Syndapin, Toca-1, CIP4, FBP17, VASP, and Cdc42	NP 001073280.1
Dia2 (DRF2)					NP 006720.1
Dia3 (DRF3)					NP 001035982.1
FMNL1 (FRL1, FHOD4)	Actin filaments bundling, centrosome, polarity regulation, lamellipodia protrusion, and membrane blebbing regulation	Lamellipodia, microtubule-organizing center, and plasma membrane	Wide	Cdc42 and Rac1	NP 005883.2
FMNL2 (FRL3, FHOD2)					NP 443137.2
FMNL3 (FRL2)					NP 783863.4
FHOD1 (FHOS)	Lamellipodia and stress fiber formation, membrane blebbing, and microtubules and actin filament alignment	Plasma membrane and Focal adhesion	Wide	ROCK1, ERK, Raf-1, and Rac1	NP 037373.2
FHOD3 (FHOS2)					
(b) Non-DRF					
Delphilin (GRID2IP)	Synaptogenesis and synaptic plasticity	Parallel fibers Purkinje cell postsynapses	Neurons	Glutamate receptor d2	NP 001138590.1
FMN1	Interphase microtubule binding, leading edge protrusion regulation, and focal adhesion formation	Microtubules, focal adhesion, and lamellipodia	Nervous system, kidney, and oocytes	a-catenin, Fyn, and Src	NP 001264242.1
FMN2	Asymmetric spindle positioning in meiotic oocytes, and cell cycle regulation	Around premeiotic chromosomes	Neurons, germinal B-cells, oocytes, and retina	Spire	NP 064450.3
INF1 (FHDC1)	Microtubule stabilization	Microtubules	Wide	—	NP 203751.2
INF2 (WHIF1)	Required for efficient mitochondrial fission, and modulating glomerular podocyte phenotype and function	Endoplasmic reticulum, golgi complex, plasma membrane, and mitochondrial fission	Wide	—	NP 071934.3

Pointed End Binding Proteins

These proteins (Table 1(c)) are less prominent than barbed-end cappers. Tropomodulin is a muscle protein that caps filament pointed ends while interacting with tropomyosin, thus regulating the length of actin filaments in skeletal muscle cells as well as in a variety of cell types (Yamashiro *et al.*,

2012). DNase 1 binds G-actin tightly at its pointed face, and can also cap filament pointed ends. Arp2/3 has been reported to cap filament pointed ends (Mullins *et al.*, 1998; Volkman *et al.*, 2014), but also to maintain barbed end dynamics in competition with capping proteins (Pantaloni *et al.*, 2001), which leaves a puzzling issue open regarding the mechanism of filament branching.

Proteins that Bind to the Sides of Actin Filaments

These proteins fall in several classes depending on their effect on the structure and related thermodynamic stability of the filament (Table 1(d)).

- a. Tropomyosin was originally discovered as a homodimeric, coiled coil muscle protein, however the appearance of tropomyosins in eukaryotes long predates the emergence of the muscle-specific isoform. In striated muscle, tropomyosin together with troponin regulates the acto-myosin crossbridge in a Ca^{++} -dependent manner. About 40 isoforms of tropomyosin, more than half of which cytoplasmic, are known and expressed in a tissue-specific manner (Gunning *et al.*, 2008; Wang and Coluccio, 2010). Tropomyosin is generally thought to stabilize filaments. It binds in the groove of the filament long-pitch helix, spans 4–7 actin subunits per molecule and thus strengthens actin–actin bonds in the filament. It competes and functionally antagonizes with other regulatory proteins like ADF/cofilin, which instead destabilizes the filament. The detailed effects of each isoform on actin assembly dynamics, the mechanism of isoform sorting, and how each isoform might recruit other specific regulators like Caldesmon and build up a complex regulation of actin assembly are unsolved issues.
- b. Because actin filaments are negatively charged polyelectrolytes, basic proteins bind and neutralize the exposed acidic N-terminal residues, promoting association of filaments in bundles. A large number of bundling proteins are known, including fascin, fimbrin/plastin, espins, alpha-actinin, and villin. Again several isoforms of all these proteins specify different actin higher order structures.
- c. Crosslinking proteins organize the actin meshworks in nonparallel bundles in various processes, often mediating their association with membranes and signaling pathways. These scaffolding, mechanosensitive proteins play essential roles in most morphogenetic processes, via elementary mechanisms that are not fully understood. Filamins are one prominent example (Razinia *et al.*, 2012). Other important scaffolds include spectrins and spectraplakins (Baines *et al.*, 2014; Suozzi *et al.*, 2012), cadherins (Budnar and Yap, 2013), proteins of the ezrin-radixin-moesin family (Neisch and Fehon, 2011), whose roles in essential motile processes are gradually emerging.
- d. Actin depolymerizing factor (ADF)/cofilin represents a family of conserved small (13–19 kDa) actin binding proteins (also called destrin, drebrin, etc...) that bind both G-actin and F-actin in their ADP-bound state specifically. Filaments made of ADF-ADP-F-actin are less stable than ADP-F-actin filaments (the critical concentration for assembly of ADP-actin is increased by ADF), the helical structure of the ADF-filament is modified (Bamburg *et al.*, 1999) and filaments fragment easily. In addition, the rate of nucleotide exchange on G-actin is greatly slowed down by ADF binding. In the presence of ADF, actin assembles in filaments that bind ADF as soon as P_i has been released following hydrolysis of ATP. The ensuing destabilization of the ‘aged’ ADF-ADP-F-actin stretches promotes partial depolymerization into ADF-ADP-G-actin, which accumulates up to a steady-state amount that depends on the

renewal rate of ATP-G-actin, that is on the concentration of ADF. The destabilization of F-actin by ADF correlates with enhanced rate of disassembly. In a population of treadmill filaments in ATP, pointed end disassembly of ADP-actin is the rate limiting process in the ATPase cycle. ADF accelerates this step 30-fold (Carlier *et al.*, 1997). Consequently, the measured steady state concentration of G-actin appears set to a higher level, which results in faster barbed-end growth. In other words, individual filaments treadmill faster in ADF. It is by enhancing depolymerization that enhanced barbed-end assembly occurs. Another view has been expressed according to which severing of filaments by ADF is responsible for their disappearance (Michelot *et al.*, 2007). It is worth noting that severing in itself only multiplies the number of filaments and does not cause faster depolymerization of individual filaments. Additionally, the proposed notion that treadmill is enhanced by severing is not sound at the scale of individual filaments. Severing can promote faster turnover of the mass amount of F-actin in solution, but each small filament is renewed at an unchanged rate. Thus this mechanism fails to explain ADF-enhanced barbed-end growth. This major regulatory activity of ADF/cofilin is observed *in vitro* (Carlier *et al.*, 1997) and *in vivo* (Lappalainen and Drubin, 1997; Kiuchi *et al.*, 2011). ADF is down regulated by phosphorylation of a conserved N-terminally located serine by LIM-kinase (Arber *et al.*, 1998). The phosphatase slingshot activates ADF by dephosphorylation (Niwa *et al.*, 2002). ADF, like profilin, is an essential protein in all actin-based motility processes including cell migration, endocytosis, and cytokinesis (Carlier *et al.*, 1999; Bernstein and Bamburg, 2010; Mizuno, 2013; Pfaendtner *et al.*, 2010).

The ADF Homology module (ADFH) is present in a variety of proteins (Drebrin, Abp1, and coactosin, which bind essentially F-actin, not G-actin) and sometimes duplicated like in twinfilin (Poukkula *et al.*, 2011). With two ADFH, twinfilin is able to cap and sever actin filaments, in addition to binding ADP-G-actin (Helfer *et al.*, 2006; Moseley *et al.*, 2006).

Another ADF-related protein is GMF (glia maturation factor), which binds specifically the ADP-bound form of Arp2 in Arp2/3 complex, thus enhancing the debranching of Arp2/3-branched filaments, regulating the plasticity of the meshwork (Gandhi *et al.*, 2010; Boczkowska *et al.*, 2013; Luan and Nolen, 2013).

The WASP Homology 2 (WH2) Domain: A Versatile Module

The WH2 domain (first identified in WAS proteins) deserves a special section because its functional versatility places it in a bunch of classes (Table 1(e)). The WH2 domain is a short protein module of about 20–30 amino acids. Intrinsically disordered in aqueous solutions, it folds upon binding to actin, adopting a binding mode similar to β -thymosins (Paunola *et al.*, 2002). Although they share with β -thymosins a binding mode supported by a short consensus LKKT ‘actin binding’ motif and a N-terminal amphipathic helix, WH2 domains display a vast sequence variability and are not evolutionarily related to β -thymosins. They are present, either as single motif or in tandem-repeated units, in several dozens of proteins which all play a role in motility. The WH2 domain

appears functionally versatile and becomes multifunctional when expressed as tandem repeats. It sometimes sequesters G-actin like β -thymosin, thus preventing actin assembly at both ends, and sometimes acts as a functional homolog of profilin, supporting filament barbed-end assembly specifically (Hertzog *et al.*, 2004). The change in function was dependent on the dynamics of interaction of the C-terminal stretch of WH2 with actin (Hertzog *et al.*, 2004). WH2 repeats were identified as nucleators of actin filaments (Quinlan *et al.*, 2005; Ahuja *et al.*, 2007; Tam *et al.*, 2007; Liverman *et al.*, 2007), but appear multifunctional regulators of filament barbed-end dynamics, spanning a range of activities from barbed-end capping (Spire) to barbed-end tracking and uncapping CP from barbed ends (Bosch *et al.*, 2007; Pernier *et al.*, 2013; Breitsprecher *et al.*, 2011; Hansen and Mullins, 2010; Khanduja and Kuhn, 2014; Co *et al.*, 2007). Moreover, Cordon-Bleu or Spire are potent filament severers (Husson *et al.*, 2011). In the cell context, where profilin competes with WH2 domains for binding actin and profilin-actin is the major monomeric form of actin, the WH2 repeats fail to nucleate filaments. Their multifunctional activities may correlate with their function in morphogenetic processes, brain development and adult brain functions involved in learning and memory, in which rapid filament remodeling is involved.

Site-directed Branching of Actin Filaments by WASP Protein Family and Arp2/3 Complex

In vivo, actin filaments are organized in a polarized fashion, with their elongating barbed ends facing the membranes. They are also generally initiated at specific sites by signal-responsive 'nucleators' (Campellone and Welch, 2010). Often nucleators are also regulators of filament barbed-end growth, thus they enforce site directed polarized growth of actin filaments, at the origin of force-producing processes (Chesarone and Goode, 2009). This is the case in lamellipodium and filopodium extension, in endocytosis, in Golgi trafficking, in focal adhesions, podosomes, in pathogen propulsion and maybe in cytokinesis and oocyte cortical actin organization.

A major way to nucleate new filaments in a site-directed fashion is harbored by WASP family proteins (Table 1(e)). These proteins are actual enzymes that use Arp2/3 complex to multiply filament barbed ends by branching, making two actively growing filaments from one, in an autocatalytic fashion. WASP proteins are activated by Rho family small GTPases using various mechanisms. They comprise many members, WASP in hematopoietic cells, neural (ubiquitous) N-WASP, WASH, WHAMM, WAVE, JMY (see Table 1), which all mediate formation of dendritic actin arrays in various cellular contexts, via the same molecular mechanism (Pollard, 2007). The association of WASP protein with proteins that recognize or induce membrane curvature (for instance BAR domains, Table 1(f)) specifies the type of membrane deformation resulting from site-directed dendritic array formation. Activation of WASP protein at the membrane exposes a C-terminal active domain called 'VCA' comprising a WH2 domain ('V,' sometimes a tandem repeat of WH2) followed by a connecting 'C' region and an acidic 'A' C-terminal short

stretch. The V region binds a G-actin molecule, the CA region binds Arp2/3 complex. Thus VCA bridges one G-actin and Arp2/3 complex. *In vitro* polymerization assays and microscopy imaging in cells (Mullins and Pollard, 1999; Svitkina and Borisy, 1999; Machesky *et al.*, 1999; Rohatgi *et al.*, 1999; Pantaloni *et al.*, 2000) showed that N-WASP used Arp2/3 to promote formation of branched filaments in an autocatalytic fashion. The 'branching complex' WASP-actin-Arp2/3 actually associates with a 'mother filament' to make a 'branch junction' in which Arp2/3 is incorporated and from which a 'daughter filament' is initiated and grows at the same rate as the mother filament. The detailed molecular mechanism of filament branching via this complex reaction is not understood (Rotty *et al.*, 2013 for review). The Arp2/3 complex is a highly conserved stable complex of seven subunits, including two actin-related proteins Arp3 and Arp2, and five ArpCs subunits (C1 to C5). The crystal structure of Arp2/3 complex is known at 2.9 Å resolution (Robinson *et al.*, 2001). Reconstitution of the full and partial complexes after expression of the seven subunits in insect cells provides insight into its structural organization (Gournier *et al.*, 2001). The cryo-EM structure of the branched junction is known at about 10 Å resolution (Rouiller *et al.*, 2008). The solution structure of the VCA-actin-Arp2/3 complex has been studied by SAXS (Boczkowska *et al.*, 2008). All these structures do not provide direct nonambiguous evidence for a unique mechanism of filament branching due to the limited resolution. Postulated structural changes of Arp2/3 upon binding VCA and filaments are required to accommodate the data. The upcoming advances in high-resolution cryo-EM are anticipated to elucidate the molecular mechanism of filament branching. Repeated cycles of filament branching by WASP using Arp2/3, G-actin and a filament as substrates, interrupted by filament barbed-end growth, generate a dendritic meshwork of branched filaments that promotes stationary propulsion of N-WASP-coated beads or vesicles in reconstituted motility assays (Wiesner *et al.*, 2003; Delatour *et al.*, 2008). The WASP protein remains immobilized at the bead/membrane during these cycles. Therefore, transient attachment of filaments in the branching process alternates with their growth, and maintains barbed ends constantly in proximity to the membrane during actin-based movement. This biochemical evidence supports the model of 'tethered ratchet' proposed by Mogilner and Oster for actin-based movement (Mogilner and Oster, 2003), and further developed recently to account for quantitative measurements of actin-based propulsion of the baculovirus (Mueller *et al.*, 2014).

The Arp2/3 complex, which is the substrate of WASP proteins, is in itself inactive in actin assembly. It can be sequestered by a variety of proteins (gadkin, arpin) which mimic the CA domain of VCA and in absence of the WH2 domain cannot elicit branching.

The plasticity of branched actin arrays is regulated by debranching, a reaction which is consecutive to hydrolysis of ATP on Arp2, and is enhanced by GMF, a protein of the ADF family that binds specifically to ADP-Arp2 in Arp2/3 complex (Gandhi *et al.*, 2010; Luan and Nolen, 2013; Boczkowska *et al.*, 2013). Debranching is also facilitated by coronin (Chan *et al.*, 2011), while cortactin in contrast stabilizes the branched filaments (Helgeson and Nolen, 2013).

Integrated Protein Modules in Reconstituted Actin-Based Motility

The principles of actin assembly dynamics and its regulation by essential actin binding proteins have led to successful reconstitution of self-organized systems showing production of force and movement by a minimal number of purified components (Figure 3). This achievement (Loisel *et al.*, 1999) validates the principle of regulated treadmilling at the origin of actin's function in motility. The two known actin-based self-organized motile systems (WASP-Arp2/3 and formins) generate movement using these principles. Indeed, in a reconstituted motility assay, like *in vivo* in lamellipodia and filopodia (Figure 3(a)), filament barbed ends elongate at a sustained rate, which implies that a constant concentration of polymerizable monomers is maintained in the solution. A sustained flux of depolymerization maintains this pool of monomers (Pantaloni *et al.*, 2001; Pollard and Borisy, 2003). The motility assay (Loisel *et al.*, 1999; Wiesner *et al.*, 2003; Romero *et al.*, 2004) reveals how molecular chemical reactions convert into macroscopic mechanical effects (Figure 3(b)). Only ATP as a source of energy is required for movement. The actin-based velocity displays a bell shape dependence on the concentration of the main players (Arp2/3 complex, ADF/cofilin, and capping protein), fully consistent with their biochemical properties (Carlier *et al.*, 1997, 1999). Specifically, depletion of Arp2/3 (by overexpression of the CA domain of WASP proteins or by Arp2/3 sequestering proteins using this motif) is predicted to slow down actin-based motility processes; conversely overexpression of Arp2/3 or increase in the density of branching by overexpression of N-WASP slow down actin-based motility, extensive branching of filament barbed ends competing with barbed-end growth.

Capping proteins are required for formation of the dendritic array and are inhibitory for formin-induced processive assembly (Wiesner *et al.*, 2003; Romero *et al.*, 2004), which was corroborated by *in vivo* observations (Mejillano *et al.*, 2004). The different optimum biochemical conditions for dendritic (WASP/Arp2/3 based) and linear (formin-mediated bundles) actin meshworks raise the issue of the mechanism that allows the coordinated function of these coexisting machineries in motile processes like lamellipodia or invadopodia (Figure 3(a)). Reconstitution and analysis of more integrated systems combining WASP- and formin-functionalized particles will provide answers to this issue.

In the future, more complex actin-based modules, like elementary focal adhesions, vesicularisation of Golgi tubules, endocytosis (Figure 3(c)), membrane-bound cortical actin arrays and their conversion to a cytokinetic ring should as well be reconstituted from a minimal number of actin partners, once the functional analysis of their individual components is completed. In many of these modules, actin assembly dynamics and force production are coupled to the dynamic and mechanical properties of the membrane. Linkage BAR-domain containing proteins that either sense or induce positive or negative membrane curvature are often associated with actin regulators (Figure 3(c)). These scaffold proteins, like IRSp53, CIP4/Toca-1, syndapin 1, MIM, ABBA-1, FBP17, bridge actin cytoskeleton and membrane by appropriate domains, and play essential roles in coupling membrane/vesicle recycling to

actin remodeling (Scita *et al.*, 2008; Suetsugu and Gautreau, 2012 for reviews).

The physical mechanism of force production against the membrane by directed assembly of actin filaments has been debated and several models have been proposed. In some physical models, it was argued that the mechanical properties of the dense actin gel were at the origin of the movement. Relaxation of the constraints resulting from insertional polymerization at the surface of a functionalized microsphere would exert a pressure causing its propulsion. The same view however cannot account for the production of a protrusive force like in lamellipodia, by the same chemical reactions and the same molecules. In addition, no relaxation of such constraints was observed in the propulsion of a vesicle (Delatour *et al.*, 2008). Finally, the observation that the baculovirus is propelled by site-directed growth and branching of less than a dozen filaments, with their barbed ends oriented toward the surface of the virus, a few of them attached to the surface (Figure 1(e)), indicates that the pushing force is directly produced by the growth of individual filaments rather than by the macroscopic elastic properties of an actin gel formed by intricated filaments (Mueller *et al.*, 2014). Microscopic models developed in closer relation with the observed chemical reactions that control movement include the brownian ratchet (Peskin *et al.*, 1993), followed by the tethered ratchet model (Mogilner and Oster, 2003) and the 'actoclampin' model (Dickinson and Purich, 2002) that actually preceded the experimental evidence for formin-promoted processive assembly (Pruyne *et al.*, 2002) or WASP-mediated barbed-end tracking (Hansen and Mullins, 2010). Progress in understanding the mechanism of actin-based motility requires detailed structural and biochemical studies of the elementary complexes formed between WASP, Arp2/3, and actin filaments during the cycle of branching, as well as between formin and the barbed end during processive filament elongation.

Finally, some motile processes, like translocation of the spindle in oocyte meiosis (Figure 3(d)), require actin assembly, but the detailed mechanism that leads to movement is not understood in detail. Progress is expected from high-resolution imaging, which now allows quantitation of the amounts of G-actin and F-actin *in vivo* (Figure 3(e)).

Actin Regulators in Adhesion Structures

Cells move by coordinated protrusion, adhesion, and contraction steps. Adhesion of cells to the ECM is strengthened by actin assembly, through maturation steps in which the chemical composition of the adhesive structures varies (focal complexes becoming focal adhesions) in association with regulatory phosphorylation reactions (focal adhesion kinase, FAK). A large number of proteins link extracellular matrix components to the membrane and the actin cytoskeleton (integrins, talin, and vinculin) and establish mechanosensitive coupling between the membrane and the actin cytoskeleton (Figure 3(a); see recent reviews Ross *et al.*, 2013; Wehrle-Haller, 2012). Focal adhesions are the archetype of structures transducing force into biochemical signals, in large part via actin assembly (Sheetz *et al.*, 2006; Schwarz and Gardel, 2012). However the detailed molecular mechanisms by which

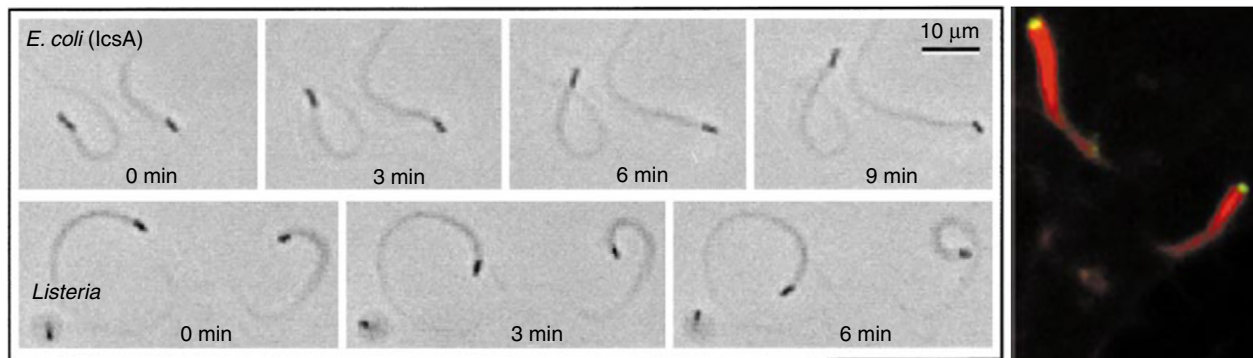
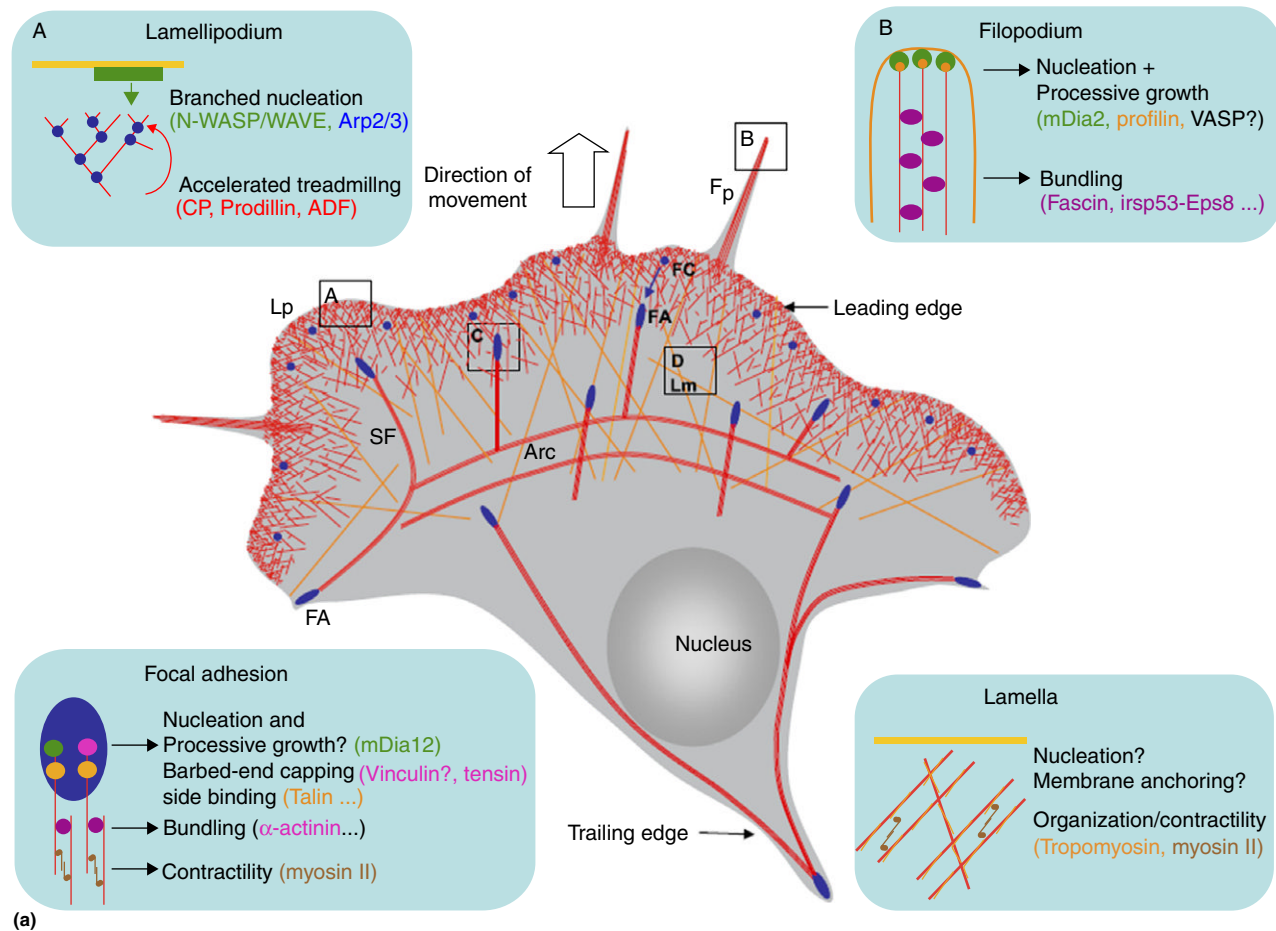


Figure 3 Cellular examples of motile processes in which actin is involved. (a) Protrusive force: lamellipodia and filopodia – adapted from Le Clainche, C., Carlier, M.F., 2008. Regulation of actin assembly associated with protrusion and adhesion in cell migration. *Physiological Reviews* 88 (2), 489–513. (b) Propulsive force: pathogens movement. Actin-based movement of *E. coli* (lcsA) and *L. monocytogenes* in reconstituted motility medium shown with time-lapse phase-contrast video-microscopy. Actin-based movement of vaccinia-infected cells shown by immunofluorescence microscopy – adapted from Loisel, T.P., Boujemaa, R., Pantaloni, D., Carlier, M.F., 1999. Reconstitution of actin-based motility of *Listeria* and *Shigella* using pure proteins. *Nature* 401 (6753), 613–616 and Moreau, V., Frischknecht, F., Reckmann, I., *et al.*, 2000. A complex of N-WASP and WIP integrates signalling cascades that lead to actin polymerization. *Nature Cell Biology* 2 (7), 441–448. Copyright 2000, Rights Managed by Nature Publishing Group. (c) Vesicle scission: time-lapse sketch of canonical clathrin-mediated endocytosis, for terminal (Ai)- and nonterminal (Aii) events, the latter generating repeated scission events – adapted from Taylor, M.J., Perrais, D., Merrifield, C.J., 2011. A high precision survey of the molecular dynamics of mammalian clathrin-mediated endocytosis. *PLoS Biology* 9 (3), e1000604. (d) Spindle translocation in asymmetric meiotic division – adapted from Pfender, S., Kuznetsov, V., Pleiser, S., Kerkhoff, E., Schuh, M., 2011. Spire-type actin nucleators cooperate with Formin-2 to drive asymmetric oocyte division. *Current Biology* 21 (11), 955–960. (e) Neural growth cone extension. Localization of G-actin (G panels) and F-actin (F panels). The green dashed region is expanded in middle row. G- and F-actins shown in lower row, in overlaid colors and in intensity ratio – adapted from Lee, C.W., Vitriol, E.A., Shim, S., *et al.*, 2013. Dynamic localization of G-actin during membrane protrusion in neuronal motility. *Current Biology* 23 (12), 1046–1056.

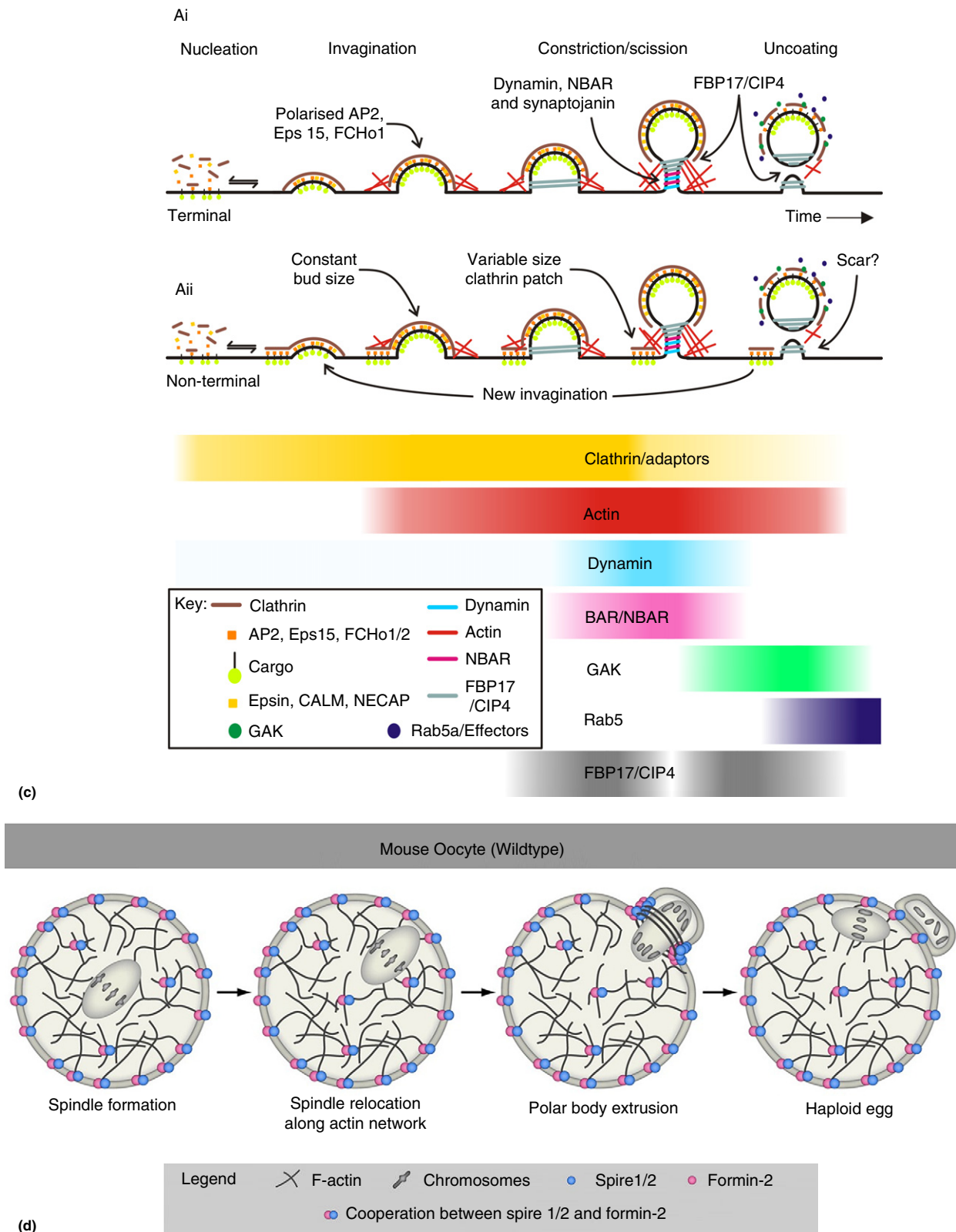


Figure 3 Continued.

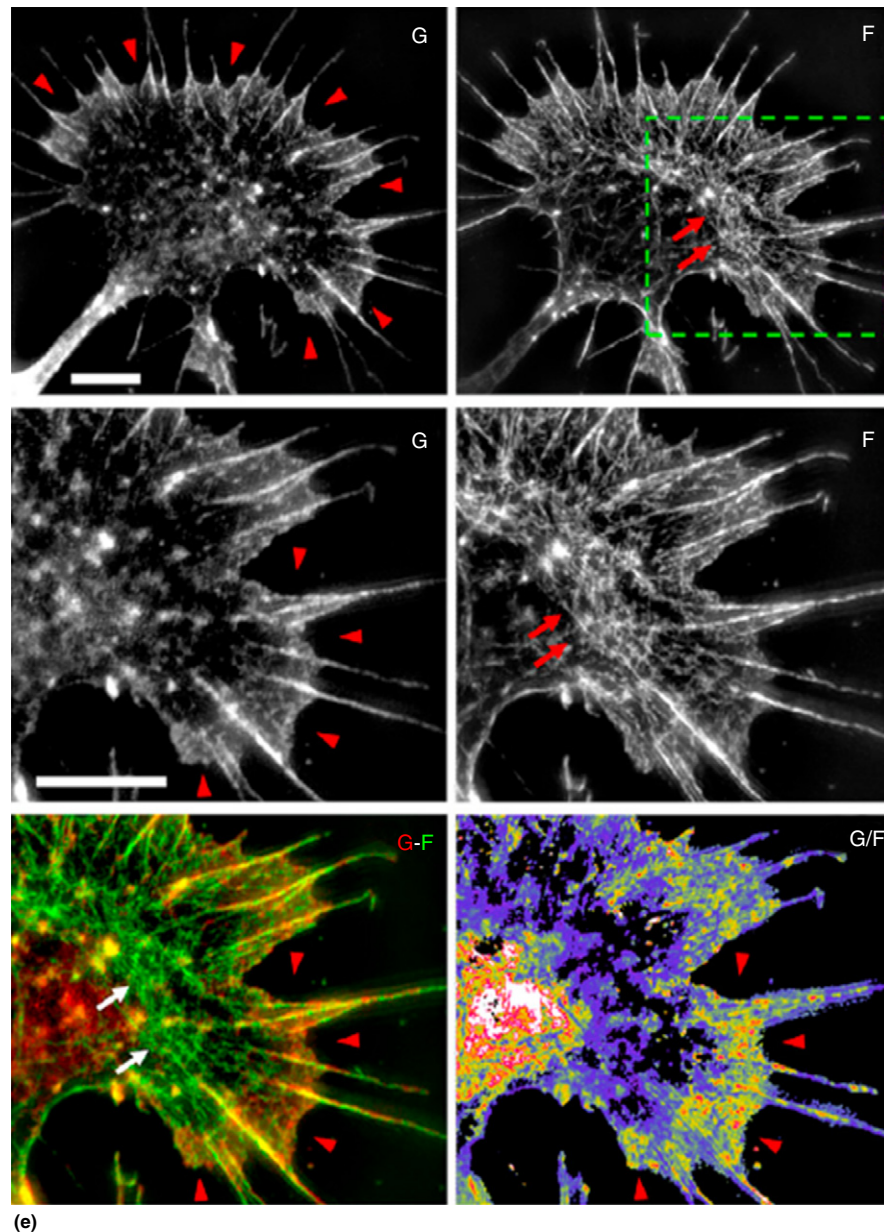


Figure 3 Continued.

nucleation and possibly processive assembly of actin filaments is regulated at focal adhesions is not yet understood in great detail and represents a challenge for the future. The biochemistry of cell adhesion will better be addressed in other articles.

Actin-Based Molecular Motors: The Family of Myosins

Myosins represent the archetype of molecular motors that associate with actin filaments in contractile structures in muscle but also in non-muscle cells (stress fibers). Myosins also organize intracellular traffic by translocating vesicles along actin filaments. The multiple functions of the many myosin families fall out of the scope of this article.

Actin in the Nucleus? Potential Role of Actin in the Regulation of Gene Expression

Actin is also an abundant component of the nucleus of amphibian oocytes, and plays a role in transcription of lampbrush chromosomes (Gounon and Karsenti, 1981; Scheer *et al.*, 1984). In somatic cells, most of actin is cytoplasmic, however the presence of actin in the nucleus – albeit in much lower amounts than in oocyte nucleus – has been detected (Egly *et al.*, 1984) and there too, actin in complex with actin-related proteins Arp4, Arp5, and Arp8 (Ino80) participates in remodeling of the chromatin (Visa and Percipalle, 2010; Grosse and Vartiainen, 2013). The molecular mechanisms involved in chromatin remodeling are still unknown. Whether

a monomer–polymer dynamic equilibrium of actin exists in the nucleus is elusive. Recent results suggest that nuclear F-actin assembly, revealed by the GFP-fused utrophin and LifeAct actin binding probes, occurs subsequent to signaling to transcription factors (Belin *et al.*, 2013). Since actin shuttles in and out from the nucleus, how the state of the large pool of cytoplasmic actin affects this equilibrium in the nucleus has to be clarified. Known actin regulators such as N-WASP, JMY, Arp2/3, formins mDia1 and mDia2, CapG and ADF/cofilin are present in the nucleus (Weston *et al.*, 2012) but in which actin-mediated processes they are involved is not clear. The MAL protein, an activator of the transcription factor SRF (serum response factor) is regulated by G-actin (Baarlink *et al.*, 2013; Treisman, 2013). MAL actually binds and sequesters G-actin via RPEL motifs, structurally and functionally similar to β -thymosins binding elements (Mouilleron *et al.*, 2008). Formation of the MAL–actin complex prevents nuclear import of MAL and activation of transcription. Consistently, an increase in F-actin leads to a decrease in sequestered nuclear G-actin, release of MAL, and activation of transcription (Baarlink *et al.*, 2013).

Bacterial Actin Ancestors

Actin-like proteins (Alps) have been identified in prokaryotes and archaea. Bacterial actins (MreB, AlfA, and ParM) show a high structural similarity with G-actin, but display highly divergent modes of helical assembly, as compared to the conserved eukaryotic F-actin. MreB assembles in single-stranded helical filaments, while ParM and AlfA assemble in a two-stranded polar helix like F-actin (Popp and Robinson, 2011 for review). In bacterial cells, the filaments arrange in bundles and cables, involved in bacterial shape building (MreB) or in plasmid segregation (ParM). Bacterial actins bind ATP or GTP in contrast with eukaryotic actin which selectively binds ATP. Nucleotide hydrolysis regulates the function of Alps.

See also: Cell Communication: Cellular Machineries: The Molecular Architecture of Cell–Cell Adhesions. Cell Communication: Cellular Outcomes: Regulation of Cell Migration. Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell Migration

References

- Ahuja, R., Pinyol, R., Reichenbach, N., *et al.*, 2007. Cordon-bleu is an actin nucleation factor and controls neuronal morphology. *Cell* 131 (2), 337–350.
- Arber, S., Barbayannis, F.A., Hanser, H., *et al.*, 1998. Regulation of actin dynamics through phosphorylation of cofilin by LIM-kinase. *Nature* 393 (6687), 805–809.
- Ayscough, K., 1998. Use of latrunculin-A, an actin monomer-binding drug. *Methods in Enzymology* 298, 18–25.
- Baarlink, C., Wang, H., Grosse, R., 2013. Nuclear actin network assembly by formins regulates the SRF coactivator MAL. *Science* 340 (6134), 864–867.
- Baines, A.J., Lu, H.C., Bennett, P.M., 2014. The Protein 4.1 family: Hub proteins in animals for organizing membrane proteins. *Biochimica et Biophysica Acta* 1838 (2), 605–619.
- Bamburg, J.R., McGough, A., Ono, S., 1999. Putting a new twist on actin: ADF/cofilins modulate actin dynamics. *Trends in Cell Biology* 9 (9), 364–370.
- Bates, M., Jones, S.A., Zhuang, X., 2013. Stochastic optical reconstruction microscopy (STORM): A method for superresolution fluorescence imaging. *Cold Spring Harbor Protocols* 2013 (6), 498–520.
- Belin, B.J., Cimini, B.A., Blackburn, E.H., Mullins, R.D., 2013. Visualization of actin filaments and monomers in somatic cell nuclei. *Molecular Biology of the Cell* 24 (7), 982–994.
- Bernstein, B.W., Bamburg, J.R., 2010. ADF/cofilin: A functional node in cell biology. *Trends in Cell Biology* 20 (4), 187–195.
- Boczkowska, M., Rebowski, G., Dominguez, R., 2013. Gila maturation factor (GMF) interacts with Arp2/3 complex in a nucleotide state-dependent manner. *Journal of Biological Chemistry* 288 (36), 25683–25688.
- Boczkowska, M., Rebowski, G., Petoukhov, M.V., *et al.*, 2008. X-ray scattering study of activated Arp2/3 complex with bound actin-WCA. *Structure* 16 (5), 695–704.
- Bosch, M., Le, K.H., Bugyi, B., *et al.*, 2007. Analysis of the function of Spire in actin assembly and its synergy with formin and profilin. *Molecular Cell* 28 (4), 555–568.
- Breitsprecher, D., Kieseewetter, A.K., Linkner, J., Faix, J., 2009. Analysis of actin assembly by *in vitro* TIRF microscopy. *Methods in Molecular Biology* 571, 401–415.
- Breitsprecher, D., Kieseewetter, A.K., Linkner, J., *et al.*, 2011. Molecular mechanism of Ena/VASP-mediated actin-filament elongation. *EMBO Journal* 30 (3), 456–467.
- Budnar, S., Yap, A.S., 2013. A mechanobiological perspective on cadherins and the actin-myosin cytoskeleton. *F1000Prime Reports* 5, 35.
- Bugyi, B., Carlier, M.F., 2010. Control of actin filament treadmilling in cell motility. *Annual Review of Biophysics* 39, 449–470.
- Campellone, K.G., Welch, M.D., 2010. A nucleator arms race: Cellular control of actin assembly. *Nature Reviews Molecular Cell Biology* 11 (4), 237–251.
- Carlier, M.F., 1989. Role of nucleotide hydrolysis in the dynamics of actin filaments and microtubules. *International Review of Cytology* 115, 139–170.
- Carlier, M.F., Jean, C., Rieger, K.J., Lenfant, M., Pantaloni, D., 1993. Modulation of the interaction between G-actin and thymosin beta 4 by the ATP/ADP ratio: Possible implication in the regulation of actin dynamics. *Proceedings of the National Academy of Sciences of the United States of America* 90 (11), 5034–5038.
- Carlier, M.F., Laurent, V., Santolini, J., *et al.*, 1997. Actin depolymerizing factor (ADF/cofilin) enhances the rate of filament turnover: Implication in actin-based motility. *Journal of Cell Biology* 136 (6), 1307–1322.
- Carlier, M.F., Pantaloni, D., 1986. Direct evidence for ADP-Pi-F-actin as the major intermediate in ATP-actin polymerization. Rate of dissociation of Pi from actin filaments. *Biochemistry* 25 (24), 7789–7792.
- Carlier, M.F., Pantaloni, D., 1994. Actin assembly in response to extracellular signals: Role of capping proteins, thymosin beta 4 and profilin. *Seminars in Cell Biology* 5 (3), 183–191.
- Carlier, M.F., Ressad, F., Pantaloni, D., 1999. Control of actin dynamics in cell motility. Role of ADF/cofilin. *Journal of Biological Chemistry* 274 (48), 33827–33830.
- Carlier, M.F., Romet-Lemonne, G., Jegou, A., 2014. Actin filament dynamics using microfluidics. *Methods in Enzymology* 540, 3–17.
- Chaki, S.P., Rivera, G.M., 2013. Integration of signaling and cytoskeletal remodeling by Nck in directional cell migration. *Bioarchitecture* 3 (3), 57–63.
- Chan, K.T., Creed, S.J., Bear, J.E., 2011. Unraveling the enigma: Progress towards understanding the coronin family of actin regulators. *Trends in Cell Biology* 21 (8), 481–488.
- Chesarone, M.A., Goode, B.L., 2009. Actin nucleation and elongation factors: Mechanisms and interplay. *Current Opinion in Cell Biology* 21 (1), 28–37.
- Co, C., Wong, D.T., Gierke, S., Chang, V., Taunton, J., 2007. Mechanism of actin network attachment to moving membranes: Barbed end capture by N-WASP WH2 domains. *Cell* 128 (5), 901–913.
- Cooper, J.A., Sept, D., 2008. New insights into mechanism and regulation of actin capping protein. *International Review of Cell and Molecular Biology* 267, 183–206.
- Delatour, V., Shekhar, S., Reymann, C., *et al.*, 2008. Actin-based propulsion of functionalized hard versus fluid objects. *New Journal of Physics* 10 (2), 20.
- Dickinson, R.B., Purich, D.L., 2002. Clamped-filament elongation model for actin-based motors. *Biophysical Journal* 82 (2), 605–617.
- Disanza, A., Carlier, M.F., Stradal, T.E., *et al.*, 2004. Eps8 controls actin-based motility by capping the barbed ends of actin filaments. *Nature Cell Biology* 6 (12), 1180–1188.
- Edwards, M., Zwolak, A., Schafer, D.A., *et al.*, 2014. Capping protein regulators fine-tune actin assembly dynamics. *Nature Reviews Molecular Cell Biology* 15 (10), 677–689.
- Egley, J.M., Miyamoto, N.G., Moncollin, V., Chambon, P., 1984. Is actin a transcription initiation factor for RNA polymerase B? *EMBO Journal* 3 (10), 2363–2371.

- Fujii, T., Iwane, A.H., Yanagida, T., Namba, K., 2010. Direct visualization of secondary structures of F-actin by electron cryomicroscopy. *Nature* 467 (7316), 724–728.
- Funatsu, T., Harada, Y., Tokunaga, M., Saito, K., Yanagida, T., 1995. Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature* 374 (6522), 555–559.
- Galkin, V.E., Orlova, A., Kudryashov, D.S., *et al.*, 2011. Remodeling of actin filaments by ADF/cofilin proteins. *Proceedings of the National Academy of Sciences of the United States of America* 108 (51), 20568–20572.
- Galkin, V.E., Orlova, A., Schroder, G.F., Egelman, E.H., 2010. Structural polymorphism in F-actin. *Nature Structural & Molecular Biology* 17 (11), 1318–1323.
- Galkin, V.E., Orlova, A., VanLoock, M.S., *et al.*, 2003. ADF/cofilin use an intrinsic mode of F-actin instability to disrupt actin filaments. *Journal of Cell Biology* 163 (5), 1057–1066.
- Gandhi, M., Smith, B.A., Bovellan, M., *et al.*, 2010. GMF is a cofilin homolog that binds Arp2/3 complex to stimulate filament debranching and inhibit actin nucleation. *Current Biology* 20 (9), 861–867.
- Goode, B.L., Eck, M.J., 2007. Mechanism and function of formins in the control of actin assembly. *Annual Review of Biochemistry* 76, 593–627.
- Gounon, P., Karsenti, E., 1981. Involvement of contractile proteins in the changes in consistency of oocyte nucleoplasm of the newt *Pleurodeles waltlii*. *Journal of Cell Biology* 88 (2), 410–421.
- Gournier, H., Goley, E.D., Niederstrasser, H., Trinh, T., Welch, M.D., 2001. Reconstitution of human Arp2/3 complex reveals critical roles of individual subunits in complex structure and activity. *Molecular Cell* 8 (5), 1041–1052.
- Grosse, R., Vartiainen, M.K., 2013. To be or not to be assembled: Progressing into nuclear actin filaments. *Nature Reviews Molecular Cell Biology* 14 (11), 693–697.
- Gunning, P., O'Neill, G., Hardeman, E., 2008. Tropomyosin-based regulation of the actin cytoskeleton in time and space. *Physiological Reviews* 88 (1), 1–35.
- Hall, A., 2012. Rho family GTPases. *Biochemical Society Transactions* 40 (6), 1378–1382.
- Hansen, S.D., Mullins, R.D., 2010. VASP is a processive actin polymerase that requires monomeric actin for barbed end association. *Journal of Cell Biology* 191 (3), 571–584.
- Heisenberg, C.P., Bellaiche, Y., 2013. Forces in tissue morphogenesis and patterning. *Cell* 153 (5), 948–962.
- Helfer, E., Nevalainen, E.M., Naumanen, P., *et al.*, 2006. Mammalian twinfilin sequesters ADP-G-actin and caps filament barbed ends: Implications in motility. *EMBO Journal* 25 (6), 1184–1195.
- Helgeson, L.A., Nolen, B.J., 2013. Mechanism of synergistic activation of Arp2/3 complex by cortactin and N-WASP. *Elife* 2, e00884.
- Hernandez-Valladares, M., Kim, T., Kannan, B., *et al.*, 2010. Structural characterization of a capping protein interaction motif defines a family of actin filament regulators. *Nature Structural & Molecular Biology* 17 (4), 497–503.
- Hertzog, M., van Heijenoort, C., Didry, D., *et al.*, 2004. The beta-thymosin/WH2 domain: structural basis for the switch from inhibition to promotion of actin assembly. *Cell* 117 (5), 611–623.
- Hetrick, B., Han, M.S., Helgeson, L.A., Nolen, B.J., 2013. Small molecules CK-666 and CK-869 inhibit actin-related protein 2/3 complex by blocking an activating conformational change. *Chemistry & Biology* 20 (5), 701–712.
- Holmes, K.C., Popp, D., Gebhard, W., Kabsch, W., 1990. Atomic model of the actin filament. *Nature* 347 (6288), 44–49.
- Holzinger, A., 2001. Jaspaklinolide. An actin-specific reagent that promotes actin polymerization. *Methods in Molecular Biology* 161, 109–120.
- Husson, C., Renaut, L., Didry, D., Pantaloni, D., Carlier, M.F., 2011. Cordon-Bleu uses WH2 domains as multifunctional dynamizers of actin filament assembly. *Molecular Cell* 43 (3), 464–477.
- Isambert, H., Venier, P., Maggs, A.C., *et al.*, 1995. Flexibility of actin filaments derived from thermal fluctuations. Effect of bound nucleotide, phalloidin, and muscle regulatory proteins. *Journal of Biological Chemistry* 270 (19), 11437–11444.
- Jegou, A., Carlier, M.F., Romet-Lemonne, G., 2013. Formin mDia1 senses and generates mechanical forces on actin filaments. *Nature Communications* 4, 1883.
- Jegou, A., Niedermayer, T., Orban, J., *et al.*, 2011. Individual actin filaments in a microfluidic flow reveal the mechanism of ATP hydrolysis and give insight into the properties of profilin. *PLoS Biology* 9 (9), e1001161.
- Kabsch, W., Mannherz, H.G., Suck, D., Pai, E.F., Holmes, K.C., 1990. Atomic structure of the actin: DNase I complex. *Nature* 347 (6288), 37–44.
- Khanduja, N., Kuhn, J.R., 2014. Processive acceleration of actin barbed-end assembly by N-WASP. *Molecular Biology of the Cell* 25 (1), 55–65.
- Kiuchi, T., Nagai, T., Ohashi, K., Watanabe, N., Mizuno, K., 2011. Live-cell imaging of G-actin dynamics using sequential FDAP. *Bioarchitecture* 1 (5), 240–244.
- Korn, E.D., 1982. Actin polymerization and its regulation by proteins from nonmuscle cells. *Physiological Reviews* 62 (2), 672–737.
- Korn, E.D., Carlier, M.F., Pantaloni, D., 1987. Actin polymerization and ATP hydrolysis. *Science* 238 (4827), 638–644.
- Kovar, D.R., Harris, E.S., Mahaffy, R., Higgs, H.N., Pollard, T.D., 2006. Control of the assembly of ATP- and ADP-actin by formins and profilin. *Cell* 124 (2), 423–435.
- Kuhn, J.R., Pollard, T.D., 2005. Real-time measurements of actin filament polymerization by total internal reflection fluorescence microscopy. *Biophysical Journal* 88 (2), 1387–1402.
- Lai, F.P., Szczodrak, M., Block, J., *et al.*, 2008. Arp2/3 complex interactions and actin network turnover in lamellipodia. *EMBO Journal* 27 (7), 982–992.
- Lammermann, T., Sixt, M., 2009. Mechanical modes of 'amoeboid' cell migration. *Current Opinion in Cell Biology* 21 (5), 636–644.
- Lappalainen, P., Drubin, D.G., 1997. Cofilin promotes rapid actin filament turnover *in vivo*. *Nature* 388 (6637), 78–82.
- Lauffenburger, D.A., Horwitz, A.F., 1996. Cell migration: A physically integrated molecular process. *Cell* 84 (3), 359–369.
- Le Clainche, C., Carlier, M.F., 2008. Regulation of actin assembly associated with protrusion and adhesion in cell migration. *Physiological Reviews* 88 (2), 489–513.
- Leung, D.W., Morgan, D.M., Rosen, M.K., 2006. Biochemical properties and inhibitors of (N)-WASP. *Methods in Enzymology* 406, 281–296.
- Liverman, A.D., Cheng, H.C., Trosky, J.E., *et al.*, 2007. Arp2/3-independent assembly of actin by *Vibrio* type III effector VopL. *Proceedings of the National Academy of Sciences of the United States of America* 104 (43), 17117–17122.
- Loisel, T.P., Boujemaa, R., Pantaloni, D., Carlier, M.F., 1999. Reconstitution of actin-based motility of *Listeria* and *Shigella* using pure proteins. *Nature* 401 (6753), 613–616.
- Luan, Q., Nolen, B.J., 2013. Structural basis for regulation of Arp2/3 complex by GMF. *Nature Structural & Molecular Biology* 20 (9), 1062–1068.
- Machesky, L.M., Mullins, R.D., Higgs, H.N., *et al.*, 1999. Scar, a WASp-related protein, activates nucleation of actin filaments by the Arp2/3 complex. *Proceedings of the National Academy of Sciences of the United States of America* 96 (7), 3739–3744.
- Manseau, L., Calley, J., Phan, H., 1996. Profilin is required for posterior patterning of the *Drosophila* oocyte. *Development* 122 (7), 2109–2116.
- Matsumita, S., Inoue, Y., Adachi, T., 2012. Quantitative analysis of extension-torsion coupling of actin filaments. *Biochemical and Biophysical Research Communications* 420 (4), 710–713.
- Medalia, O., Beck, M., Ecker, M., *et al.*, 2007. Organization of actin networks in intact filopodia. *Current Biology* 17 (1), 79–84.
- Mejillano, M.R., Kojima, S., Applewhite, D.A., *et al.*, 2004. Lamellipodial versus filopodial mode of the actin nanomachinery: Pivotal role of the filament barbed end. *Cell* 118 (3), 363–373.
- Michelot, A., Berro, J., Guerin, C., *et al.*, 2007. Actin-filament stochastic dynamics mediated by ADF/cofilin. *Current Biology* 17 (10), 825–833.
- Mizuno, H., Higashida, C., Yuan, Y., *et al.*, 2011. Rotational movement of the formin mDia1 along the double helical strand of an actin filament. *Science* 331 (6013), 80–83.
- Mizuno, K., 2013. Signaling mechanisms and functional roles of cofilin phosphorylation and dephosphorylation. *Cellular Signalling* 25 (2), 457–469.
- Mogilner, A., Oster, G., 2003. Force generation by actin polymerization II: The elastic ratchet and tethered filaments. *Biophysical Journal* 84 (3), 1591–1605.
- Montaville, P., Jegou, A., Pernier, J., *et al.*, 2014. Spire and Formin 2 synergize and antagonize in regulating actin assembly in meiosis by a ping-pong mechanism. *PLoS Biology* 12 (2), e1001795.
- Moseley, J.B., Okada, K., Balcer, H.I., *et al.*, 2006. Twinfilin is an actin-filament-severing protein and promotes rapid turnover of actin structures *in vivo*. *Journal of Cell Science* 119 (Pt 8), 1547–1557.
- Moulleron, S., Guettler, S., Langer, C.A., Treisman, R., McDonald, N.Q., 2008. Molecular basis for G-actin binding to RPEL motifs from the serum response factor coactivator MAL. *EMBO Journal* 27 (23), 3198–3208.
- Mueller, J., Pflanzler, J., Winkler, C., *et al.*, 2014. Electron tomography and simulation of baculovirus actin comet tails support a tethered filament model of pathogen propulsion. *PLoS Biology* 12 (1), e1001765.
- Mullins, R.D., Heuser, J.A., Pollard, T.D., 1998. The interaction of Arp2/3 complex with actin: Nucleation, high affinity pointed end capping, and formation of branching networks of filaments. *Proceedings of the National Academy of Sciences of the United States of America* 95 (11), 6181–6186.
- Mullins, R.D., Pollard, T.D., 1999. Rho-family GTPases require the Arp2/3 complex to stimulate actin polymerization in *Acanthamoeba* extracts. *Current Biology* 9 (8), 405–415.

- Murakami, K., Yasunaga, T., Noguchi, T.Q., *et al.*, 2010. Structural basis for actin assembly, activation of ATP hydrolysis, and delayed phosphate release. *Cell* 143 (2), 275–287.
- Nag, S., Larsson, M., Robinson, R.C., Burtneck, L.D., 2013. Gelsolin: The tail of a molecular gymnast. *Cytoskeleton (Hoboken)* 70 (7), 360–384.
- Narita, A., Oda, T., Maeda, Y., 2011. Structural basis for the slow dynamics of the actin filament pointed end. *EMBO Journal* 30 (7), 1230–1237.
- Narita, A., Takeda, S., Yamashita, A., Maeda, Y., 2006. Structural basis of actin filament capping at the barbed-end: A cryo-electron microscopy study. *EMBO Journal* 25 (23), 5626–5633.
- Nedeva, I., Koripelly, G., Caballero, D., *et al.*, 2013. Synthetic polyamines promote rapid lamellipodial growth by regulating actin dynamics. *Nature Communications* 4, 2165.
- Neisch, A.L., Fehon, R.G., 2011. Ezrin, Radixin and Moesin: Key regulators of membrane-cortex interactions and signaling. *Current Opinion in Cell Biology* 23 (4), 377–382.
- Niwa, R., Nagata-Ohashi, K., Takeichi, M., Mizuno, K., Uemura, T., 2002. Control of actin reorganization by Slingshot, a family of phosphatases that dephosphorylate ADF/cofilin. *Cell* 108 (2), 233–246.
- Nolen, B.J., Tomasevic, N., Russell, A., *et al.*, 2009. Characterization of two classes of small molecule inhibitors of Arp2/3 complex. *Nature* 460 (7258), 1031–1034.
- Oda, T., Iwasa, M., Aihara, T., Maeda, Y., Narita, A., 2009. The nature of the globular- to fibrous-actin transition. *Nature* 457 (7228), 441–445.
- Oosawa, F., Asakura, S., 1975. *Thermodynamics of the Polymerization of Protein*. London: Academic Press.
- Otomo, T., Tomchick, D.R., Otomo, C., *et al.*, 2005. Structural basis of actin filament nucleation and processive capping by a formin homology 2 domain. *Nature* 433 (7025), 488–494.
- Pantaloni, D., Boujemaa, R., Didry, D., Gounon, P., Carlier, M.F., 2000. The Arp2/3 complex branches filament barbed ends: Functional antagonism with capping proteins. *Nature Cell Biology* 2 (7), 385–391.
- Pantaloni, D., Carlier, M.F., 1993. How profilin promotes actin filament assembly in the presence of thymosin beta 4. *Cell* 75 (5), 1007–1014.
- Pantaloni, D., Le Clainche, C., Carlier, M.F., 2001. Mechanism of actin-based motility. *Science* 292 (5521), 1502–1506.
- Paul, A.S., Pollard, T.D., 2008. The role of the FH1 domain and profilin in formin-mediated actin-filament elongation and nucleation. *Current Biology* 18 (1), 9–19.
- Paunola, E., Mattila, P.K., Lappalainen, P., 2002. WH2 domain: A small, versatile adapter for actin monomers. *FEBS Letters* 513 (1), 92–97.
- Pernier, J., Orban, J., Avvaru, B.S., *et al.*, 2013. Dimeric WH2 domains in Vibrio VopF promote actin filament barbed-end uncapping and assisted elongation. *Nature Structural & Molecular Biology* 20 (9), 1069–1076.
- Peskin, C.S., Odell, G.M., Oster, G.F., 1993. Cellular motions and thermal fluctuations: The Brownian ratchet. *Biophysical Journal* 65 (1), 316–324.
- Ptaendner, J., De La Cruz, E.M., Voth, G.A., 2010. Actin filament remodeling by actin depolymerization factor/cofilin. *Proceedings of the National Academy of Sciences of the United States of America* 107 (16), 7299–7304.
- Pollard, T.D., 2007. Regulation of actin filament assembly by Arp2/3 complex and formins. *Annual Review of Biophysics and Biomolecular Structure* 36, 451–477.
- Pollard, T.D., Borisy, G.G., 2003. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* 112 (4), 453–465.
- Pollard, T.D., Cooper, J.A., 1984. Quantitative analysis of the effect of Acanthamoeba profilin on actin filament nucleation and elongation. *Biochemistry* 23 (26), 6631–6641.
- Pollard, T.D., Cooper, J.A., 2009. Actin, a central player in cell shape and movement. *Science* 326 (5957), 1208–1212.
- Popp, D., Robinson, R.C., 2011. Many ways to build an actin filament. *Molecular Microbiology* 80 (2), 300–308.
- Poukkula, M., Kremneva, E., Serlachius, M., Lappalainen, P., 2011. Actin-depolymerizing factor homology domain: A conserved fold performing diverse roles in cytoskeletal dynamics. *Cytoskeleton (Hoboken)* 68 (9), 471–490.
- Prochiewicz, E., Janson, N., Thomas, D.D., De la Cruz, E.M., 2005. Cofilin increases the torsional flexibility and dynamics of actin filaments. *Journal of Molecular Biology* 353 (5), 990–1000.
- Pruyne, D., Evangelista, M., Yang, C., *et al.*, 2002. Role of formins in actin assembly: Nucleation and barbed-end association. *Science* 297 (5581), 612–615.
- Quinlan, M.E., Heuser, J.E., Kerkhoff, E., Mullins, R.D., 2005. Drosophila Spire is an actin nucleation factor. *Nature* 433 (7024), 382–388.
- Ramarao, N., Le Clainche, C., Izard, T., *et al.*, 2007. Capping of actin filaments by vinculin activated by the Shigella IpaA carboxyl-terminal domain. *FEBS Letters* 581 (5), 853–857.
- Rao, J.N., Madasu, Y., Dominguez, R., 2014. Mechanism of actin filament pointed-end capping by tropomodulin. *Science* 345 (6195), 463–467.
- Razinia, Z., Makela, T., Ylanne, J., Calderwood, D.A., 2012. Filamins in mechanosensing and signaling. *Annual Review of Biophysics* 41, 227–246.
- Ridley, A.J., 2011. Life at the leading edge. *Cell* 145 (7), 1012–1022.
- Ridley, A.J., Schwartz, M.A., Burridge, K., *et al.*, 2003. Cell migration: Integrating signals from front to back. *Science* 302 (5651), 1704–1709.
- Rigort, A., Gunther, D., Hegerl, R., *et al.*, 2012. Automated segmentation of electron tomograms for a quantitative description of actin filament networks. *Journal of Structural Biology* 177 (1), 135–144.
- Rizvi, S.A., Neidt, E.M., Cui, J., *et al.*, 2009. Identification and characterization of a small molecule inhibitor of formin-mediated actin assembly. *Chemistry & Biology* 16 (11), 1158–1168.
- Robinson, R.C., Turbedsky, K., Kaiser, D.A., *et al.*, 2001. Crystal structure of Arp2/3 complex. *Science* 294 (5547), 1679–1684.
- Rohatgi, R., Ma, L., Miki, H., *et al.*, 1999. The interaction between N-WASP and the Arp2/3 complex links Cdc42-dependent signals to actin assembly. *Cell* 97 (2), 221–231.
- Romero, S., Le Clainche, C., Didry, D., *et al.*, 2004. Formin is a processive motor that requires profilin to accelerate actin assembly and associated ATP hydrolysis. *Cell* 119 (3), 419–429.
- Ross, T.D., Coon, B.G., Yun, S., *et al.*, 2013. Integrins in mechanotransduction. *Current Opinion in Cell Biology* 25 (5), 613–618.
- Rotty, J.D., Wu, C., Bear, J.E., 2013. New insights into the regulation and cellular functions of the ARP2/3 complex. *Nature Reviews Molecular Cell Biology* 14 (1), 7–12.
- Rouiller, I., Xu, X.P., Amann, K.J., *et al.*, 2008. The structural basis of actin filament branching by the Arp2/3 complex. *Journal of Cell Biology* 180 (5), 887–895.
- Safer, D., Nachmias, V.T., 1994. Beta thymosins as actin binding peptides. *Bioessays* 16 (8), 590.
- Sagot, I., Rodal, A.A., Moseley, J., Goode, B.L., Pellman, D., 2002. An actin nucleation mechanism mediated by Bni1 and profilin. *Nature Cell Biology* 4 (8), 626–631.
- Salbreux, G., Charras, G., Paluch, E., 2012. Actin cortex mechanics and cellular morphogenesis. *Trends in Cell Biology* 22 (10), 536–545.
- Scheer, U., Hinssen, H., Franke, W.W., Jockusch, B.M., 1984. Microinjection of actin-binding proteins and actin antibodies demonstrates involvement of nuclear actin in transcription of lampbrush chromosomes. *Cell* 39 (1), 111–122.
- Schmoller, K.M., Niedermayer, T., Zensen, C., Wurm, C., Bausch, A.R., 2011. Fragmentation is crucial for the steady-state dynamics of actin filaments. *Biophysical Journal* 101 (4), 803–808.
- Schwarz, U.S., Gardel, M.L., 2012. United we stand: Integrating the actin cytoskeleton and cell-matrix adhesions in cellular mechanotransduction. *Journal of Cell Science* 125 (Pt 13), 3051–3060.
- Scita, G., Confalonieri, S., Lappalainen, P., Suetsugu, S., 2008. IRSp53: Crossing the road of membrane and actin dynamics in the formation of membrane protrusions. *Trends in Cell Biology* 18 (2), 52–60.
- Sheetz, M.P., Sable, J.E., Dobereiner, H.G., 2006. Continuous membrane-cytoskeleton adhesion requires continuous accommodation to lipid and cytoskeleton dynamics. *Annual Review of Biophysics and Biomolecular Structure* 35, 417–434.
- Small, J.V., Anderson, K., Rottner, K., 1996. Actin and the coordination of protrusion, attachment and retraction in cell crawling. *Bioscience Reports* 16 (5), 351–368.
- Spector, I., Shochet, N.R., Kashman, Y., Groweiss, A., 1983. Latrunculin: Novel marine toxins that disrupt microfilament organization in cultured cells. *Science* 219 (4584), 493–495.
- Straub, F.B., 1942. Actin. Institute of Medical Chemistry, University of Szeged II. 3–15.
- Straub, F.B., 1943. Actin, II. Institute of Medical Chemistry, University of Szeged II. 23–37.
- Suetsugu, S., Gautreau, A., 2012. Synergistic BAR-NPF interactions in actin-driven membrane remodeling. *Trends in Cell Biology* 22 (3), 141–150.
- Sun, H.Q., Kwiatkowska, K., Yin, H.L., 1995. Actin monomer binding proteins. *Current Opinion in Cell Biology* 7 (1), 102–110.
- Suozi, K.C., Wu, X., Fuchs, E., 2012. Spectraplakins: Master orchestrators of cytoskeletal dynamics. *Journal of Cell Biology* 197 (4), 465–475.
- Svitkina, T.M., Borisy, G.G., 1999. Arp2/3 complex and actin depolymerizing factor/cofilin in dendritic organization and treadmilling of actin filament array in lamellipodia. *Journal of Cell Biology* 145 (5), 1009–1026.
- Takeda, S., Minakata, S., Koike, R., *et al.*, 2010. Two distinct mechanisms for actin capping protein regulation – steric and allosteric inhibition. *PLoS Biology* 8 (7), e1000416.

- Tam, V.C., Serruto, D., Dziejman, M., Briehner, W., Mekalanos, J.J., 2007. A type III secretion system in *Vibrio cholerae* translocates a formin/spire hybrid-like actin nucleator to promote intestinal colonization. *Cell Host & Microbe* 1 (2), 95–107.
- Tang, J.X., Janmey, P.A., 1996. The polyelectrolyte nature of F-actin and the mechanism of actin bundle formation. *Journal of Biological Chemistry* 271 (15), 8556–8563.
- Treisman, R., 2013. Shedding light on nuclear actin dynamics and function. *Trends in Biochemical Sciences* 38 (8), 376–377.
- Tu, D., Graziano, B.R., Park, E., *et al.*, 2012. Structure of the formin-interaction domain of the actin nucleation-promoting factor Bud6. *Proceedings of the National Academy of Sciences of the United States of America* 109 (50), E3424–E3433.
- Valentin-Ranc, C., Carlier, M.F., 1989. Evidence for the direct interaction between tightly bound divalent metal ion and ATP on actin. Binding of the lambda isomers of beta gamma-bidentate CrATP to actin. *Journal of Biological Chemistry* 264 (35), 20871–20880.
- Visa, N., Percipalle, P., 2010. Nuclear functions of actin. *Cold Spring Harbor Perspectives in Biology* 2 (4), a000620.
- Volkman, N., Page, C., Li, R., Hanein, D., 2014. Three-dimensional reconstructions of actin filaments capped by Arp2/3 complex. *European Journal of Cell Biology* 93 (5–6), 179–183.
- Wang, C.L., Coluccio, L.M., 2010. New insights into the regulation of the actin cytoskeleton by tropomyosin. *International Review of Cell and Molecular Biology* 281, 91–128.
- Watanabe, N., 2012. Fluorescence single-molecule imaging of actin turnover and regulatory mechanisms. *Methods in Enzymology* 505, 219–232.
- Wegner, A., 1976. Head to tail polymerization of actin. *Journal of Molecular Biology* 108 (1), 139–150.
- Wegner, A., 1982. Spontaneous fragmentation of actin filaments in physiological conditions. *Nature* 296 (5854), 266–267.
- Wegner, A., Engel, J., 1975. Kinetics of the cooperative association of actin to actin filaments. *Biophysical Chemistry* 3 (3), 215–225.
- Wehrle-Haller, B., 2012. Assembly and disassembly of cell matrix adhesions. *Current Opinion in Cell Biology* 24 (5), 569–581.
- Weston, L., Coutts, A.S., La Thangue, N.B., 2012. Actin nucleators in the nucleus: An emerging theme. *Journal of Cell Science* 125 (Pt 15), 3519–3527.
- Wiesner, S., Helfer, E., Didry, D., *et al.*, 2003. A biomimetic motility assay provides insight into the mechanism of actin-based motility. *Journal of Cell Biology* 160 (3), 387–398.
- Xu, K., Babcock, H.P., Zhuang, X., 2012. Dual-objective STORM reveals three-dimensional filament organization in the actin cytoskeleton. *Nature Methods* 9 (2), 185–188.
- Xu, Y., Moseley, J.B., Sagot, I., *et al.*, 2004. Crystal structures of a Formin Homology-2 domain reveal a tethered dimer architecture. *Cell* 116 (5), 711–723.
- Yamashiro, S., Gokhin, D.S., Kimura, S., Nowak, R.B., Fowler, V.M., 2012. Tropomodulins: Pointed-end capping proteins that regulate actin filament architecture in diverse cell types. *Cytoskeleton (Hoboken)* 69 (6), 337–370.
- Yarmola, E.G., Bubb, M.R., 2006. Profilin: Emerging concepts and lingering misconceptions. *Trends in Biochemical Sciences* 31 (4), 197–205.
- Yasuda, R., Miyata, H., Kinoshita Jr., K., 1996. Direct measurement of the torsional rigidity of single actin filaments. *Journal of Molecular Biology* 263 (2), 227–236.

Further Reading

- Lee, C.W., Vitriol, E.A., Shim, S., *et al.*, 2013. Dynamic localization of G-actin during membrane protrusion in neuronal motility. *Current Biology* 23 (12), 1046–1056.
- Moreau, V., Frischknecht, F., Reckmann, I., *et al.*, 2000. A complex of N-WASP and WIP integrates signalling cascades that lead to actin polymerization. *Nature Cell Biology* 2 (7), 441–448.
- Pfender, S., Kuznetsov, V., Pleiser, S., Kerkhoff, E., Schuh, M., 2011. Spire-type actin nucleators cooperate with Formin-2 to drive asymmetric oocyte division. *Current Biology* 21 (11), 955–960.
- Schroeter, M.M., Orlova, A., Egelman, E.H., Beall, B., Chalovich, J.M., 2013. Organization of F-actin by Fesselin (avian smooth muscle synaptopodin 2). *Biochemistry* 52 (29), 4955–4961.
- Taylor, M.J., Perrais, D., Merrifield, C.J., 2011. A high precision survey of the molecular dynamics of mammalian clathrin-mediated endocytosis. *PLoS Biology* 9 (3), e1000604.

Intermediate Filaments

RE Leube and N Schwarz, RWTH Aachen University, Aachen, Germany

© 2016 Elsevier Inc. All rights reserved.

Glossary

Astrocyte Star-shaped glial cell of the central nervous system.

Coiled coil Coiled coils are typically composed of two α -helices which consist of heptad repeats (abcdefg, where a and d are hydrophobic and e and f are charged amino acid residues) and are held together through strong hydrophobic interactions.

Cytolinker Protein connecting cytoskeletal filaments to each other and to cellular organelles.

Cytolysis Cell rupture.

Intermediate filament Cytoskeletal filament with a diameter of about 10 nm between that of actin filaments and microtubules.

Keratinocyte Predominant cell type in the epidermis.

The Intermediate Filament Protein Family is Subdivided into Six Types with Cell Type-Specific Expression Patterns

The members of the mammalian intermediate filament protein family are classified into six types based on sequence homology, gene organization, net charge, assembly mechanism, and expression pattern (Table 1; Coulombe and Wong, 2004; Herrmann *et al.*, 2009).

More than 50 type I and type II keratins comprise the 'soft' cytokeratins that are expressed in epithelial cells and the 'hard' keratins found in hair and nails. They are among the most abundant cellular proteins. Equal amounts of the more acidic type I and more basic type II keratins assemble into intermediate filaments in tissue- and differentiation-specific combinations. For example, the type I keratin 18 together with its type II partner keratin 8 is expressed in simple, one-layered epithelia of the intestinal mucosa, endometrium, or exocrine glands. In some of these epithelial cells, for example, hepatocytes, they are the only keratins expressed. In multilayered epithelia, such as the epidermis, other keratin pairs are expressed with specific keratin pairs in the basal cell layers (keratins 5 and 14) and suprabasal cell layers (keratins 1 and 10).

The type III intermediate filament proteins include vimentin, which is typically expressed in cells of mesenchymal origin such as fibroblasts, leukocytes, and endothelial cells, whereas desmin and syncoilin are found in muscle cells. Glial fibrillary acidic protein (GFAP) is produced in glial cells. Peripherin is mainly detected in peripheral neurons and in central nervous system neurons, which have projections toward the peripheral nervous system.

The type IV low-, middle-, and high-molecular weight neurofilament proteins NF-L, NF-M, and NF-H, whose C-termini differ dramatically in length, are highly expressed, together with α -internexin, in axonal processes of central nervous system neurons. Neurofilaments co-assemble *in vivo* as heteropolymers of NF-L and either NF-M or NF-H, respectively. Synemin, which is produced in glial astrocytes and muscle cells, and nestin, which is synthesized in developing and regenerating cells of the central nervous system and myogenic tissue, do not assemble into intermediate filaments on their

own but have regulatory functions for the assembly and disassembly of other intermediate filaments.

The type V lamin intermediate filament proteins are localized in the nucleus. At least one B-type lamin (B1, B2, and B2 splice variant B3) is expressed in every nucleated cell whereas A-type lamins (lamin splice variants A and C) have a more restricted distribution and are absent during early embryogenesis. Lamins are the major components of the nuclear lamina, a layer subjacent to the inner surface of the nuclear envelope.

Filensin and phakinin make unusual beaded filaments that are solely expressed in lens fiber cells.

Intermediate Filament Proteins Are Reliable Markers of Cellular Differentiation

The cell type-specific expression of intermediate filaments is exploited in tissue and tumor diagnosis (Moll *et al.*, 2008; Bragulla and Homberger, 2009; Karantzis, 2011). It is well

Table 1 The intermediate filament protein family is subdivided into six types with distinct distribution patterns

Type	Members	Distribution
I	'Acidic' keratins	Epithelial cells
II	'Basic' keratins	Epithelial cells
III	Vimentin	Mesenchymal cells
	Desmin and syncoilin	Muscle cells
	Glial fibrillary acidic protein (GFAP)	Glial cells
	Peripherin	Neurons
IV	Neurofilament proteins (NF-L, NF-M, and NF-H) and α -Internexin	Neurons
	Synemin	Astrocytes and muscle cells
	Nestin	Developing and regenerating cells
V	Lamins	Nuclear lamina of all nucleated cells
VI	Filensin and phakinin	Lens fiber cells

Source: Available at: <http://www.interfil.org/> (accessed 29.12.14).

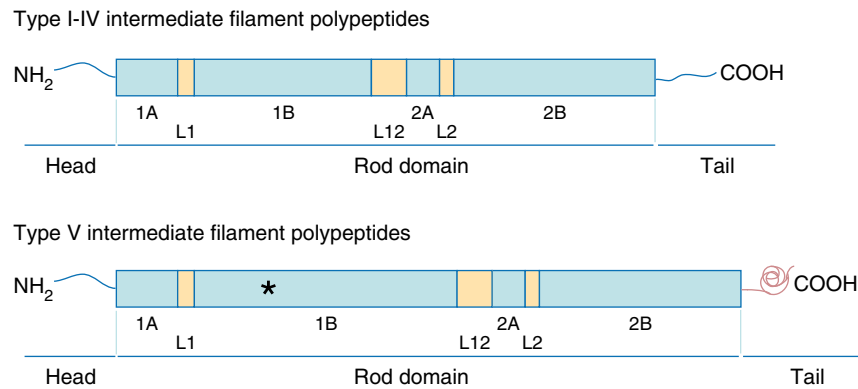


Figure 1 Model of intermediate filament protein structure. The central α -helical rod domain is subdivided into segments 1A, 1B, 2A, and 2B by non-helical linker domains L1, L12, and L2. The rod is flanked by the highly variable N-terminal head and the C-terminal tail domain. Type V intermediate filament polypeptides differ from Types I–IV by a 42 amino acid insertion in coil-domain 1B (star) and by a nuclear localization signal and immunoglobulin-like beta fold in the tail domain.

established that cellular differentiation is reflected by specific intermediate filament proteins. The main tissue types can thus be distinguished on the basis of their intermediate filament composition: epithelia are characterized by keratins, connective tissues by vimentin, muscles by desmin, and nervous tissues by neurofilaments in neurons and GFAP in glial cells. Intermediate differentiation states usually go along with defined admixtures of intermediate filament proteins. For example, myoepithelial cells contain desmin together with keratins or myofibroblasts vimentin together with desmin. In addition, further classification of tissue differentiation is possible by analysis of intermediate filament protein isotype combinations. This is especially true for the abundant epithelial keratins. Changes of differentiation, which occur during development or as a reaction to various insults, are also reflected by changes in intermediate filament expression. Taken together, the intermediate filament composition of a given cell can be taken as a reliable molecular marker of its differentiation and activity status. Remarkably, the same holds true for tumor cells. Epithelium-derived carcinomas are generally characterized by keratins, whereas mesenchymal tumors continue to produce vimentin. Furthermore, specific keratin pairs may provide clues for the origin of a given carcinoma and its differentiation status. Thus, simple epithelium-derived adenocarcinomas produce keratins 8 and 18, whereas multilayered epithelium-derived squamous cell carcinomas synthesize other keratin pairs, which are typical for multilayered epithelia. Changes in differentiation, however, which usually occur during malignant transformation and accompany dedifferentiation go along with alterations in intermediate filament protein expression. A prominent example is epithelial to mesenchymal transition, which is paralleled by dramatic shape changes and increased motility of invading cancer cells. These cells start to co-express vimentin together with reduced amounts of keratins.

Intermediate Filaments Share a Common Secondary Structure

Despite the sequence diversity among intermediate filament proteins they all share a common tripartite secondary structure

(Figure 1; Herrmann *et al.*, 2009): a ~ 45 nm-long α -helical rod domain is flanked by non-helical N-terminal head and C-terminal tail domains. The central rod domain is highly conserved and consists of four α -helices (coils 1A, 1B, 2A, and 2B) with interspersed non-helical linkers (L1, L12, and L2). Lamins carry an additional 42 amino acid insertion in the first central helical subdomain. The head and tail domains differ considerably in size, sequence, and secondary structure between different intermediate filament polypeptides. While the tail domain of human keratin 19 contains only 15 amino acids, nestin has more than 1300.

The rod domains interact with each other to form filaments, whereas the head and tail domains bind to various cytoplasmic elements including other cytoskeletal components. The N-terminal head domain regulates intermediate filament protein assembly, while the C-terminal tail domain is dispensable and only affects filament stability.

The amino acid sequences at the beginning and the end of the rod domain are evolutionarily most highly conserved. They are referred to as the helix initiation and helix termination motifs and are genetic hotspots for mutations in many hereditary intermediate filament disorders.

Intermediate Filaments Can Be Assigned to Three Different Assembly Groups

The initial building block of all intermediate filaments is the parallel, in-register, two-stranded coiled-coil dimer whose α -helices are strongly bound to each other by non-covalent hydrophobic forces (Figure 2). However, there are no crystallization data of polymerized full-length intermediate filament proteins available. Since the entire rod domains could not be crystallized, several rod fragments were analyzed. Coil 2B forms dimeric complexes while coil 1A remains monomeric in solution even at high protein concentrations although it assumes the helical twist that is necessary for coiled-coil formation (Strelkov *et al.*, 2002).

Depending on the subsequent mode of assembly intermediate filaments can be assigned to three different assembly groups (Figure 2; Herrmann *et al.*, 2009).

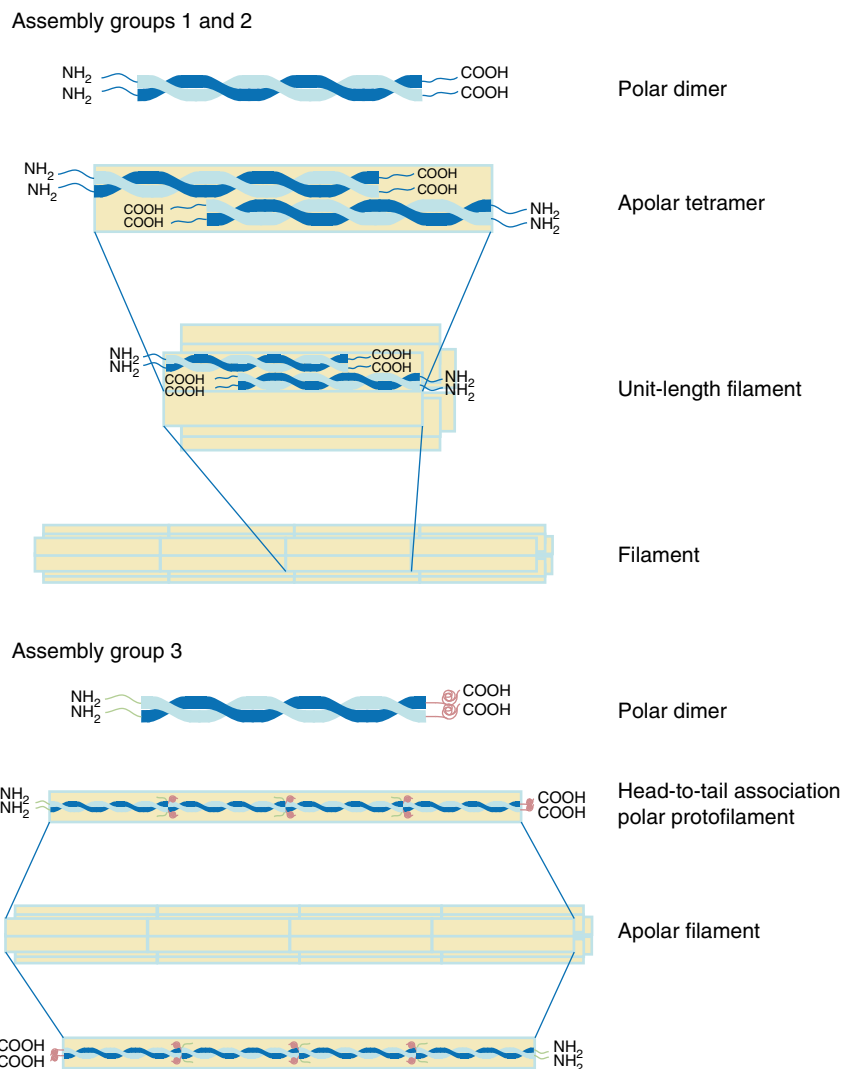


Figure 2 Schemes depicting major stages of intermediate filament assembly. Assembly groups 1 and 2 are characterized by parallel, non-staggered dimers consisting either of obligatory heteropolymers (assembly group 1) or facultative homopolymers (assembly group 2). Two antiparallel dimers associate in a staggered fashion to form a tetramer. Six to ten tetramers associate laterally into ~60 nm-long unit-length filaments, which elongate longitudinally and compact as 8–12 nm intermediate filaments. Nuclear lamins are members of assembly group 3. They also form polar, non-staggered dimers. However, dimers attach head-to-tail to generate protofilaments. Association of protofilaments occurs antiparallel to generate apolar intermediate filaments. Amino-termini (NH₂) and carboxy-termini (COOH) are labeled.

Type I and type II keratins belong to assembly group 1. Keratins are obligatory type I-type II heterodimers. Two antiparallel-oriented dimers associate laterally into staggered tetramers. The resulting lack of polarity of this and all further assembly intermediates is one of the distinguishing features of cytoplasmic intermediate filaments in comparison to the polar actin- and tubulin-based filament systems. Lateral association of 6–10 tetramers forms short ~60 nm unit-length filaments (ULFs) with a diameter of 8–16 nm. ULFs then anneal longitudinally giving rise to longer filaments that subsequently undergo radial compaction yielding 8–12 nm mature intermediate filaments. Keratin filament assembly is extremely fast and occurs spontaneously in physiologic salt solutions without any cofactors.

Assembly group 2 includes type III and type IV intermediate filaments, which typically assemble from homodimers. Further assembly intermediates are comparable to those found in assembly group 1 but may differ with respect to the number of subunits per filament cross section. In addition, the assembly rate is much slower than that observed in group 1.

The type V lamin intermediate filaments belong to assembly group 3. Lamins have a large C-terminal globular domain that contains an immunoglobulin-like beta-fold together with a nuclear localization sequence. Lamin dimers associate head-to-tail into polar protofilaments with protruding globular C-termini. Protofilaments then further associate laterally in an antiparallel fashion into extended ~10 nm mature apolar fibrils.

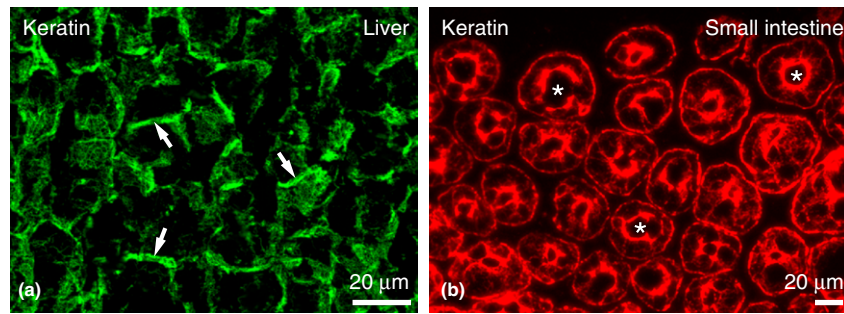


Figure 3 Intermediate filament network organization in murine tissues. The immunofluorescence micrographs depict the organization of keratins in hepatocytes of murine liver (a) and in a cross section of epithelial crypts of the intestinal mucosa (b). Note the enrichment of filaments around bile canaliculi (arrows) in liver and the subapical concentration of keratins in the epithelial cells surrounding crypt lumina (stars).

Intermediate Filament Networks Have Distinct Properties That Render Them Different from the Other Cytoskeletal Systems

Intermediate filaments differ in many aspects from the actin- and tubulin-based cytoskeletal filaments (Coulombe and Wong, 2004; Chou *et al.*, 2007; Herrmann *et al.*, 2009).

- Biochemically, intermediate filaments are difficult to solubilize. High concentrations of chaotropic reagents are needed to dissociate the extremely stable coiled-coil dimers.
- Intermediate filaments are apolar fibrous structures, which cannot guide motor protein-driven directional transport.
- Intermediate filament assembly occurs spontaneously in physiological salt solutions and does not require nucleotide triphosphates or proteinaceous cofactors.
- Intermediate filaments do not only elongate by addition of subunits to either end but also grow by incorporation of subunits along their entire length. Consistently, subunits can also dissociate from intact filaments. The process is referred to as ‘dynamic subunit exchange’ and results in considerable filament width inhomogeneity.
- Intermediate filament networks are very stable and retain their arrangement even when cells are treated with highly concentrated salt solutions and nonionic detergents.
- Intermediate filament proteins have a long half-life of up to several months. On the other hand, intermediate filaments have been shown to be dynamic. Based on high-resolution time-lapse imaging of cells transfected with fluorescently tagged keratin a model for the keratin filament turnover cycle was described (Windoffer *et al.*, 2011). Keratin filament precursors are formed in the periphery near the cell membrane where they assemble. They subsequently integrate into the existing filament network by moving inward. Near the nucleus filaments bundle and become either part of the perinuclear cage or dissolve and diffuse back toward the periphery where a new cycle is initiated starting with the formation of keratin filament precursors.
- The intermediate filament network provides compliance to small deformations on one hand and strengthens cells on the other hand when large deformations are applied. *In vitro* assays have shown that intermediate filaments are much more deformable than the other cytoskeletal

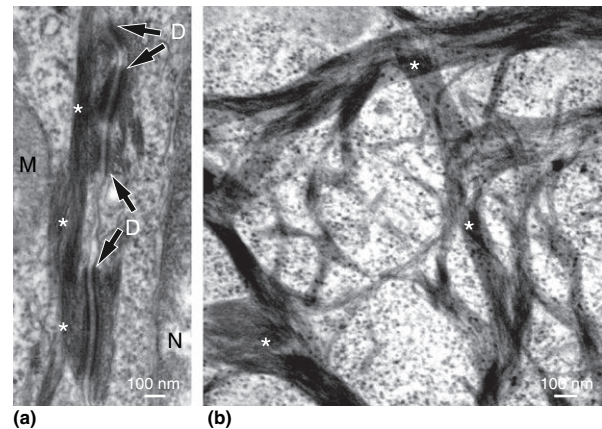


Figure 4 Electron microscopy of keratin intermediate filaments in epidermal cells. The micrograph in (a) depicts thick bundles of keratin (stars) connecting multiple desmosomes (D). M, mitochondrion; N, nucleus. The image in (b) shows a cytoplasmic region with thick intermediate filament bundles (stars) that are part of the highly interlaced three-dimensional cytoskeletal network. Images were kindly provided by Sabine Eisner.

filaments and can be stretched several times their initial length before they tear (Kreplak *et al.*, 2005). Upon application of force the filaments become more viscoelastic, a behavior that is referred to as strain stiffening (Janmey *et al.*, 1991).

Cytoplasmic Intermediate Filament Network Organization Is Determined by Interaction with Major Cellular Components through Associated Proteins

Intermediate filament networks present many different types of spatial arrangement. They form dense bundles of closely packed filaments in neuronal axons, they are part of a complex anchoring structure of Z-discs to the plasma membrane in striated muscle, they are concentrated in cortical compartments of hepatocytes, they are enriched subapically in enterocytes or make up a 3D-meshwork in epidermal keratinocytes (Figures 3 and 4). The molecular basis of the different organizational forms is still poorly understood. But the linkage of intermediate filaments to each other and to various other

determinants of cellular architecture including microtubules, actin filaments, and adhesive complexes at cell membranes, and to cellular organelles such as mitochondria, Golgi apparatus, endoplasmic reticulum, and the nucleus certainly contribute to it. These interactions are mediated by intermediate filament-associated proteins many of which contain a plakin domain (Sonnenberg and Liem, 2007; Bouameur *et al.*, 2014). A prominent plakin family member is the 500 kDa protein plectin with its multiple splice variants containing different binding modules (Wiche and Winter, 2011). This ubiquitously expressed and multifunctional cytolinker integrates the intermediate filament system into the cytoskeleton and connects it to other cellular components thereby affecting cell shape, tissue stability, and cell migration. The cytolinker function of plakins is particularly evident in epithelial cells, which are attached to each other by desmosomal cell–cell contacts and to the basement membrane by hemidesmosomes. Plectin has been localized to hemidesmosomes, whereas the plakin desmoplakin is present in desmosomes. In electron micrographs both structures are characterized by electron-dense plaques at the cell membrane with associated keratin intermediate filament bundles (Figure 4(a)). In this way forces are evenly distributed throughout the epithelial layer and propagated to the underlying connective tissue providing mechanical coherence and resilience.

Other intermediate filament-associated proteins have been identified with distinct functions for network organization. For example, filaggrins promote keratin filament bundling in epidermal cells. In muscles the intermediate filament type IV protein synemin, which does not form filaments on its own, integrates into existing intermediate filament networks through polymerization with vimentin and desmin type III intermediate filaments. Synemin recruits these mixed intermediate filament networks to extracellular matrix adhesion sites by interaction with focal adhesion components (Bellin *et al.*, 2001). In heart muscle, synemin thereby helps to distribute forces by attaching the contractile apparatus to the surrounding connective tissue.

Lamin Intermediate Filaments Contribute to Nuclear Structure and Function

The nuclear lamina of most mammalian somatic cells contains A- and B-type lamins, which form separate but interacting meshworks (Dechat *et al.*, 2008; Burke and Stewart, 2013; Zwerger and Medalia, 2013). These meshworks are attached to the inner nuclear membrane through lamina-associated proteins that are either integral or peripheral membrane proteins. The LINC (linker of nucleoskeleton and cytoskeleton)-complex, which traverses the nuclear envelope, further connects the nuclear lamina to the cytoplasmic cytoskeleton. The outer nuclear membrane components of the LINC-complex are nesprins (nuclear envelope spectrin repeat proteins). Specific nesprin isoforms bind directly to actin or indirectly to intermediate filaments via plectin to determine positioning of the nucleus (Wilhelmsen *et al.*, 2005).

Nuclear deformability depends on the level of lamin expression. High lamin levels correspond to stiffer nuclei (Pajeroski *et al.*, 2007). Consequently, lamin A/C-deficient

nuclei are more deformable, more fragile, and prone to spontaneous rupture, and lamin B-deficient cells show increased extrusions of nuclear material into the cytoplasm that are referred to as nuclear blebs (Lammerding *et al.*, 2006).

Lamins and their associated proteins control positioning of chromosomes. Chromatin is bound to lamin through specific DNA sequences, called matrix attachment regions. For example, the *Drosophila melanogaster* lamin Dm0 has been shown to interact with transcriptionally silent DNA sequences that are often localized to the nuclear periphery (Pickersgill *et al.*, 2006). In addition, lamin domains bind to core histones and through several lamin-binding proteins indirectly to chromatin. Further evidence for the involvement of lamins in the regulation of transcription was provided by studies that examined lamin mutants lacking the head domain and revealed inhibition of RNA polymerase II (Spann *et al.*, 2002). Furthermore, lamins A and C were shown to interact with a nuclear zinc finger transcription factor affecting transcription and posttranscriptional processing (Dreuillet *et al.*, 2002).

Lamins and their associated proteins have been shown to regulate the activity and accessibility of signaling molecules (Burke and Stewart, 2013). For example, lamin A is able to interfere with TGFβ signaling via segregation and modification of TGFβ-dependent intracellular mediators (Berlo *et al.*, 2005; Pan *et al.*, 2005).

Intermediate Filament Proteins Are Posttranslationally Modified

Intermediate filament proteins are subject to several types of posttranslational modification resulting in amino acid modification, protein aggregation and rearrangement, and degradation. In this way the intermediate filament system adapts to and reflects different functional states of a given cell (Snider and Omary, 2014).

During mitosis cells need to disrupt their cytoskeleton which is tightly coupled to phosphorylation of intermediate filament proteins. While actin filaments and microtubules are depolymerized and used for specific functions during cell division, intermediate filaments do not participate in essential mitotic processes. But they need to be redistributed to the daughter cells. Some cells disassemble their cytoplasmic intermediate filament network into aggregates (Figure 5) while others only show restricted disassembly at the cleavage furrow. After cell division, fully extended cytoplasmic intermediate filament networks are reestablished. During mitosis an increase in serine and threonine phosphorylation of the head and tail domains occurs, which leads to increased intermediate filament protein solubility and altered intermediate filament network morphology. Phosphorylation-deficient vimentin mutants have been shown to interfere with cytokinesis resulting in binucleated cells (Goto and Inagaki, 2014). Similarly, an increase in phosphorylation of the nuclear lamins occurs during nuclear lamina disassembly in mitosis. Ablation of the phosphorylated serine residues leads to the inhibition of nuclear lamina disassembly (Heald and McKeon, 1990).

In contrast, phosphorylation of neurofilaments has a stabilizing function. Neurofilaments are among the most

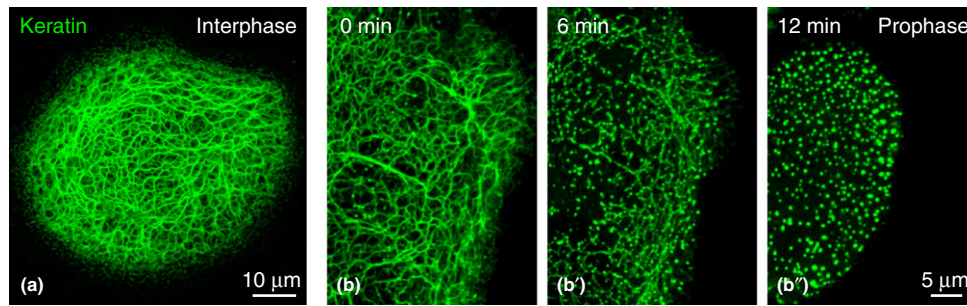


Figure 5 Intermediate filament network organization in a cultured cell in interphase and at the onset of mitosis. The fluorescence micrograph in (a) was taken from an interphase cell producing fluorescent protein-tagged keratin intermediate filament proteins. Note the dense network extending throughout the cytoplasm with smaller particles in the periphery. The time series in (b–b'') shows part of a similar cell entering mitosis. Within a few minutes the filament network is broken down into multiple small granules. Images were kindly provided by Marcin Moch (a) and Dr. Stefan Wöll (b–b'').

abundant and most highly phosphorylated proteins in neurons and form tightly clustered bundles providing mechanical support for axonal processes. Furthermore, phosphorylation of neurofilaments is linked to an increased axonal diameter resulting in increased conduction velocity (Perrot *et al.*, 2008).

Sumoylation, the covalent and reversible attachment of small ubiquitin-related modifier (SUMO) proteins, is a conserved posttranslational modification of intermediate filaments and regulates filament formation and solubility. Inhibition of sumoylation in *Caenorhabditis elegans* resulted in aggregate formation of the intermediate filament protein IFB-1 in the epidermis which was accompanied by reduced intermediate filament turnover and abnormal embryonic elongation (Kaminsky *et al.*, 2009). Furthermore, sumoylation has been shown to be important for the localization of lamin A (Zhang and Sarge, 2008).

O-linked glycosylation, the enzymatic addition of beta-D-N-acetylglucosamine to serine and threonine residues, is related to nutrient sensing and stress responses. In simple epithelia glycosylation of keratin 18 promotes the phosphorylation and activation of cell survival kinases in stress situations (Ku *et al.*, 2010).

Intermediate filaments are cleaved by caspases during apoptosis and are ubiquitinated to be subjected to proteasomal degradation during normal cell homeostasis (Snider and Omary, 2014).

A well worked-out example of successive posttranslational modifications is provided by lamins (Burke and Stewart, 2013). They are initially translated as pre-lamins whose cysteine residue in their C-terminal -CAAX motif becomes modified by addition of the 15-carbon isoprenoid farnesyl. Next, the -AAX part is removed. Subsequently, the cysteine residue becomes carboxymethylated generating the mature membrane-anchored B-type lamin. In case of A-type lamins, the C-terminal 15 amino acids including the modified cysteine residue are cleaved off producing mature lamin A, which lacks the farnesyl membrane anchor.

Intermediate Filaments Serve Many Functions

For a long time intermediate filaments were thought to exclusively support the structural integrity of cells. Recent

research efforts, however, uncovered many other functions in cell growth, migration, cellular architecture, and stress-related signaling (Toivola *et al.*, 2005, 2010; Pekny and Lane, 2007; Magin *et al.*, 2007; Kim and Coulombe, 2007).

Intermediate Filaments Provide Structural Support

The identification of intermediate filament mutations in diseases that are characterized by tissue fragility provides compelling evidence for the structural function of intermediate filaments. This is best exemplified in human skin disease, where single point mutations in keratin genes induce blistering in mechanically challenged areas. Furthermore, epithelial cells lacking cytoplasmic intermediate filaments are less resistant to mechanical stress evoked by magnetic tweezers or optical stretching (Ramms *et al.*, 2013; Seltmann *et al.*, 2013). Similarly, fibroblasts without vimentin display decreased stiffness and increased instability when subjected to mechanical challenges (Eckes *et al.*, 1998). Desmin deficiency results in myopathy with impaired force transmission in skeletal and heart muscle (Capetanaki *et al.*, 2007).

Evidence for mechanosensitive functions of intermediate filaments was provided in shear stress experiments of lung epithelial cells, which reacted by increasing keratin solubility, aggregation, and degradation (Ridge *et al.*, 2005). Vimentin expression in blood vessels is increased when high shear stress is applied strengthening the arterial wall structure and counteracting arterial diameter enlargement (Schnittler *et al.*, 1998).

Nuclear shape is at least partly determined by lamins. For example, spermatocytes which have a unique hook-shaped nucleus specifically express lamin B3. Ectopic expression of lamin B3 in somatic cells, which usually have a spherical nucleus, induced hook-shaped nuclei (Furukawa and Hotta, 1993). Lamin mutations furthermore result in elongated skeletal and myocardial nuclei and increased susceptibility to nuclear fragility (Burke and Stewart, 2013).

Intermediate Filaments Support Specialized Tissue Functions

In order to maintain the structure and barrier function of the epidermis, keratinocytes need to differentiate from proliferating cells in the basal layer toward postmitotic, tightly sealed cells in the upper cell layer which eventually die by apoptosis

and give rise to the stratum corneum that is made up by the cornified envelope and aggregated keratins. These differences in cellular differentiation are reflected by different keratin pairs with keratins 5 and 14 in the basal cell layer and increasing levels of keratins 1 and 10 in the suprabasal cell layers. Injury induces the expression of keratins 6 and 16 to support wound closure. Disruption of keratin genes therefore results in reduced mechanical stability, altered wound closure and compromised barrier function in the epidermis (Homberg and Magin, 2014).

Keratins have been linked to secretion in endocrine cells of the pancreas. Reduced insulin levels were observed in mice lacking keratin 8 (Alam *et al.*, 2013). In colonic mucosa, keratins are important for proper electrolyte transport (Toivola *et al.*, 2005).

During axonal growth, neurofilament subunits are incorporated throughout the axon in a dynamic process that involves the addition of subunits along the entire filament length as well as at filament ends. The amount and specific admixtures of neurofilaments along with their phosphorylation state determine axonal diameter and are therefore responsible for diameter-dependent axonal conduction velocity (Liem and Messing, 2009).

Intermediate Filaments Are Guardians of Cells in Stressful Situations

In stress situations intermediate filaments are usually hyperphosphorylated resulting in changed network morphology and intermediate filament protein solubility (Ridge *et al.*, 2005; Snider and Omary, 2014). Interestingly, heat shock proteins, which are increased during heat stress, bind some intermediate filaments (Toivola *et al.*, 2010).

Intermediate filament proteins protect cells from apoptosis. Their hyperphosphorylation acts as a phosphate sink preventing the phosphorylation and therefore activation of proapoptotic factors (Lai *et al.*, 1993; Ku and Omary, 2006). Consequently, keratin-deficient mice are generally more susceptible to apoptosis induced by toxins or other stressors.

Nestin, an intermediate filament protein that is usually expressed during development and downregulated in the adult organism, is reexpressed during pathological situations such as glial scar formation in the central nervous system and regeneration of injured muscle tissue. Relevant in this context may be the interaction of nestin and Cdk5 to promote cell survival (Sahlgren *et al.*, 2006).

Intermediate Filaments Are Reorganized During Migration, Wound Healing, and Tissue Remodeling

Cell migration and wound healing strongly depend on the fundamental reorganization of the intermediate filament cytoskeleton (DePianto and Coulombe, 2004; Ivaska *et al.*, 2007; Chung *et al.*, 2013). Thus, induction of usually absent keratin isoforms is observed during wound healing in epidermal keratinocytes to make them more malleable for repopulating the wounded area. Similarly, there is a rapid upregulation of vimentin and nestin during repair of damaged muscle. Peripheral blood mononuclear cells of mice lacking

vimentin have a reduced capacity to home to lymph nodes and spleen because surface receptors involved in homing are aberrantly expressed and distributed in both the blood mononuclear cells and endothelial cells. This indicates an important role of intermediate filaments in anchoring and organizing adhesion molecules.

In the brain, tissue reorganization following insults is also coupled to altered intermediate filament expression. Peripherin is induced in cerebral ischemia and GFAP is upregulated in glial scar formation.

Intermediate Filaments Contribute to Organelle Localization

Intermediate filaments are involved in the positioning, shaping, and function of cellular organelles (Styers *et al.*, 2005; Toivola *et al.*, 2005). Mitochondria in cardiac muscle of desmin knock-out mice are therefore mislocalized, fragmented, and functionally altered. Mitochondria are also aberrantly distributed in skin diseases caused by mutations in keratin genes. Disease-associated mutations of neurofilament proteins cause fragmentation of the Golgi apparatus.

Vimentin, peripherin, and alpha-internexin bind to an adaptor protein that carries vesicles between endosomal and lysosomal compartments and regulates sorting of different vesicles. In vimentin-null cells the uniform localization of lysosomes throughout the cytoplasm is lost and lysosomes are concentrated near the nucleus.

Cytoplasmic Intermediate Filament Dysfunctions Cause Tissue-Specific Diseases

Hereditary skin diseases and degenerative diseases of muscles and neurons are characterized by disruption of the intermediate filament network and prominent intermediate filament protein-containing cytoplasmic aggregates (Liem and Messing, 2009; Coulombe *et al.*, 2009; Clemen *et al.*, 2013). The highly conserved helix initiation and termination motifs of the central rod domain contain genetic hotspots for the most severe disease mutations. These mutations interfere with filament assembly and network stability.

The clinical importance of keratin intermediate filaments became first evident when single point mutations in the epidermal keratin 14 were linked to the skin disease epidermolysis bullosa simplex. A hallmark feature of this disease is the formation of blisters in the basal layer of the epidermis because of cytolysis, i.e., cell rupture. Cytoplasmic keratin aggregates are pathognomonic features in electron micrographs. Subsequently, skin disease-related mutations were identified in all other epidermal keratins. The localization of the respective pathologies reflected in each case the distribution of the affected keratin. In addition to blistering, increased proliferation and altered differentiation were observed to different degrees. Similar phenotypes are reproduced in mice lacking keratins. Surprisingly, multilayered epidermis is still established in the absence of all epidermal keratins. In addition to pronounced blistering, however, considerable deficiencies in epidermal differentiation and barrier formation occur in this situation (Figure 6).

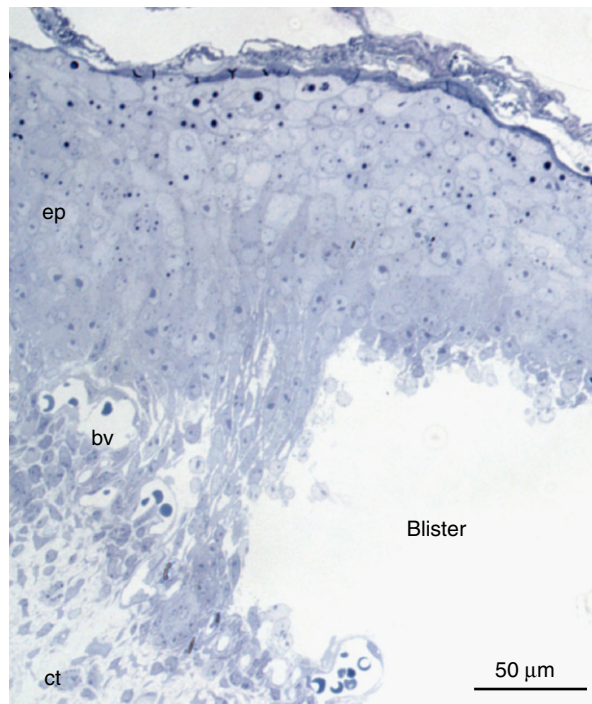


Figure 6 Blister formation in keratin-deficient mouse epidermis. The picture shows a toluidine blue-stained histological section of murine skin lacking keratin filaments. A multilayered epithelial tissue is still formed. But tissue resilience is considerably compromised resulting in blister formation because of cytolysis. bv, blood vessel; ct, connective tissue; ep, epidermis. The tissue sample was kindly provided by Dr. Thomas Magin, Leipzig University.

Mutations of hard keratins have been linked to human diseases affecting hair and nails (Langbein and Schweizer, 2005). In contrast to soft keratins, which possess glycine- and serine-rich head and tail domains, hard keratins possess cysteine- and proline-rich end domains. They form tight cross-links with each other and with keratin-associated proteins. The most common hard keratin-related disease is monilethrix. It is characterized by fragile hair and nail deformities. Affected individuals display periodic changes in the width of the hair shaft leading to beaded hair that easily breaks.

In simple keratins mutations of the helix entry and termination regions have not been identified as direct causes of human disease. But mutations in their head and tail domains predispose to chronic liver disease. This is apparently related to decreased stress protection. In addition, simple keratin-rich aggregates, Mallory bodies, are typical for cirrhotic liver disease (Strnad *et al.*, 2008).

Mutations in vimentin, phakinin, and filensin have been identified in human cataract patients (Song *et al.*, 2009). Similarly, deletion of filensin or phakinin or expression of mutant vimentin leads to opacification and cataract formation in mice.

Desmin mutations are associated with progressive skeletal myopathy that leads to skeletal muscle weakness in combination with cardiac arrhythmia and heart failure (Clemen *et al.*, 2013). A distinctive histopathological feature is

the accumulation of amorphous cytoplasmic deposits containing desmin and other proteins.

Mutations in the GFAP gene cause Alexander disease, a rare degenerative disease of the white matter in the brain which leads to loss of myelin (Liem and Messing, 2009). Again, cytoplasmic protein aggregates occur that contain several ubiquitinated stress proteins along with the mutant GFAP. Mutations of neurofilament proteins have been identified in neurodegenerative Charcot-Marie-Tooth disease. Affected individuals progressively lose muscle tissue and touch sensation.

Lamin Intermediate Filament Mutations Cause Complex Syndromes

Progeria infantilis or Hutchinson-Gilford syndrome is one of the most remarkable laminopathies (Worman and Bonne, 2007). It leads to premature aging and cardiac dysfunction. The rare disease has been linked to mutations in the lamin A gene, which cannot be processed into the mature form by cleavage of the farnesylated and methylated carboxy-terminus. Restructuring of the nuclear lamina is observed in several cell types. It is characterized by dissociation of heterochromatin from the nuclear periphery and by changes in histone methylation. The lamin A mutations also induce increased nuclear stiffness. In contrast, laminopathies, which are caused by other lamin A mutations and lamin B mutations and lead to muscle dystrophy syndromes, are associated with reduced nuclear stiffness and increased nuclear envelope fragility.

Highly Diverse Intermediate Filament Proteins Occur Throughout the Animal Kingdom

A prominent feature of intermediate filament proteins is their high degree of sequence diversity. Evolutionary pressure is apparently placed on the basic design of the rod domain but neither on its specific amino acid composition nor on the sequences in the highly variable end domains. This diversity enables intermediate filaments to perform isotype-specific functions in different cellular environments. It may explain why large intermediate filament multigene families have evolved and why the individual family members are still rapidly evolving. For example, silencing of a hair keratin gene in humans occurred as recently as 200 000–240 000 years ago (Winter *et al.*, 2001).

Although intermediate filaments are apparently absent in plants and fungi, they are present throughout the animal kingdom (Erber *et al.*, 1998; Peter and Stick, 2012). Comparing intron positions of intermediate filament genes and the amino acid sequences of their encoded polypeptides indicates that they evolved from a lamin-like ancestor. Simple metazoans such as *Hydra attenuata* express only nuclear intermediate filaments. In the invertebrate nematode *C. elegans*, however, 1 nuclear and 11 cytoplasmic intermediate filament genes have been identified. The cytoplasmic intermediate filament proteins contain an elongated coil 1B domain as is typical for lamins. They can be subdivided into two expression groups: Group 1 involves IFB-1 co-assembled with IFA variants in a cell-specific manner; group 2 consists of the 6 remaining intermediate filament proteins that are co-expressed in the

intestine. The nuclear intermediate filament encoded by the single *lmn-1* gene is expressed in all nuclei and localizes to the nuclear envelope. The insect *D. melanogaster* has two nuclear lamin-type intermediate filament proteins but lacks cytoplasmic intermediate filaments. The recent identification of the cytoplasmic intermediate filament-like protein isomin in the intestine of the basal hexapod *Isotomurus maculatus* provides an evolutionary link between the different metazoan phyla (Mencarelli *et al.*, 2011).

See also: Cytoskeleton and Motors: Cytoskeletal Components: Actin Assembly Dynamics and Its Regulation in Motile and Morphogenetic Processes; Bacterial and Archaeal Cytoskeletons

References

- Alam, C.M., Silvander, J.S.G., Daniel, E.N., *et al.*, 2013. Keratin 8 modulates β -cell stress responses and normoglycaemia. *Journal of Cell Science* 126, 5635–5644.
- Bellin, R.M., Huiatt, T.W., Critchley, D.R., Robson, R.M., 2001. Synemin may function to directly link muscle cell intermediate filaments to both myofibrillar Z-lines and costameres. *Journal of Biological Chemistry* 276, 32330–32337.
- Berlo, J.H.V., Voncken, J.W., Kubben, N., *et al.*, 2005. A-type lamins are essential for TGF- β 1 induced PP2A to dephosphorylate transcription factors. *Human Molecular Genetics* 14, 2839–2849.
- Bouameur, J.-E., Favre, B., Borradori, L., 2014. Plakins, a versatile family of cytolinkers: Roles in skin integrity and in human diseases. *Journal of Investigative Dermatology* 134, 885–894.
- Bragulla, H.H., Homberger, D.G., 2009. Structure and functions of keratin proteins in simple, stratified, keratinized and cornified epithelia. *Journal of Anatomy* 214, 516–559.
- Burke, B., Stewart, C.L., 2013. The nuclear lamins: Flexibility in function. *Nature Reviews Molecular Cell Biology* 14, 13–24.
- Capetanaki, Y., Bloch, R.J., Kouloumenta, A., Mavroidis, M., Psarras, S., 2007. Muscle intermediate filaments and their links to membranes and membranous organelles. *Experimental Cell Research* 313, 2063–2076. **Special Issue – Intermediate Filaments.**
- Chou, Y.-H., Flitney, F.W., Chang, L., *et al.*, 2007. The motility and dynamic properties of intermediate filaments and their constituent proteins. *Experimental Cell Research* 313, 2236–2243.
- Chung, B.-M., Rotty, J.D., Coulombe, P.A., 2013. Networking galore: Intermediate filaments and cell migration. *Current Opinion in Cell Biology* 25, 600–612.
- Clemen, C.S., Herrmann, H., Strelkov, S.V., Schröder, R., 2013. Desminopathies: Pathology and mechanisms. *Acta Neuropathologica* 125, 47–75.
- Coulombe, P.A., Kerns, M.L., Fuchs, E., 2009. Epidermolysis bullosa simplex: A paradigm for disorders of tissue fragility. *Journal of Clinical Investigation* 119, 1784–1793.
- Coulombe, P.A., Wong, P., 2004. Cytoplasmic intermediate filaments revealed as dynamic and multipurpose scaffolds. *Nature Cell Biology* 6, 699–706.
- Dechat, T., Pfeleghaar, K., Sengupta, K., *et al.*, 2008. Nuclear lamins: Major factors in the structural organization and function of the nucleus and chromatin. *Genes & Development* 22, 832–853.
- DePianto, D., Coulombe, P.A., 2004. Intermediate filaments and tissue repair. *Experimental Cell Research* 301, 68–76.
- Dreuillet, C., Tillit, J., Kress, M., Ernout-Lange, M., 2002. *In vivo* and *in vitro* interaction between human transcription factor MOK2 and nuclear lamin A/C. *Nucleic Acids Research* 30, 4634–4642.
- Eckes, B., Dogic, D., Colucci-Guyon, E., *et al.*, 1998. Impaired mechanical stability, migration and contractile capacity in vimentin-deficient fibroblasts. *Journal of Cell Science* 111, 1897–1907.
- Erber, A., Riemer, D., Bovenschulte, M., Weber, K., 1998. Molecular phylogeny of metazoan intermediate filament proteins. *Journal of Molecular Evolution* 47, 751–762.
- Furukawa, K., Hotta, Y., 1993. cDNA cloning of a germ cell specific lamin B3 from mouse spermatocytes and analysis of its function by ectopic expression in somatic cells. *EMBO Journal* 12, 97–106.
- Goto, H., Inagaki, M., 2014. New insights into roles of intermediate filament phosphorylation and progeria pathogenesis. *IUBMB Life* 66, 195–200.
- Heald, R., McKeon, F., 1990. Mutations of phosphorylation sites in lamin A that prevent nuclear lamina disassembly in mitosis. *Cell* 61, 579–589.
- Herrmann, H., Strelkov, S.V., Burkhard, P., Aebi, U., 2009. Intermediate filaments: Primary determinants of cell architecture and plasticity. *Journal of Clinical Investigation* 119, 1772–1783.
- Homberg, M., Magin, T.M., 2014. Beyond expectations: Novel insights into epidermal keratin function and regulation. *International Review of Cell and Molecular Biology* 311, 265–306.
- Ivaska, J., Pallari, H.-M., Nevo, J., Eriksson, J.E., 2007. Novel functions of vimentin in cell adhesion, migration, and signaling. *Experimental Cell Research* 313, 2050–2062.
- Jannmey, P.A., Euteneuer, U., Traub, P., Schliwa, M., 1991. Viscoelastic properties of vimentin compared with other filamentous biopolymer networks. *Journal of Cell Biology* 113, 155–160.
- Kaminsky, R., Denison, C., Bening-Abu-Shach, U., *et al.*, 2009. SUMO regulates the assembly and function of a cytoplasmic intermediate filament protein in *C. elegans*. *Developmental Cell* 17, 724–735.
- Karantza, V., 2011. Keratins in health and cancer: More than mere epithelial cell markers. *Oncogene* 30, 127–138.
- Kim, S., Coulombe, P.A., 2007. Intermediate filament scaffolds fulfill mechanical, organizational, and signaling functions in the cytoplasm. *Genes & Development* 21, 1581–1597.
- Kreplak, L., Bär, H., Leterrier, J.F., Herrmann, H., Aebi, U., 2005. Exploring the mechanical behavior of single intermediate filaments. *Journal of Molecular Biology* 354, 569–577.
- Ku, N.-O., Omary, M.B., 2006. A disease- and phosphorylation-related nonmechanical function for keratin 8. *Journal of Cell Biology* 174, 115–125.
- Ku, N.-O., Toivola, D.M., Strnad, P., Omary, M.B., 2010. Cytoskeletal keratin glycosylation protects epithelial tissue from injury. *Nature Cell Biology* 12, 876–885.
- Lai, Y.K., Lee, W.C., Chen, K.D., 1993. Vimentin serves as a phosphate sink during the apparent activation of protein kinases by okadaic acid in mammalian cells. *Journal of Cellular Biochemistry* 53, 161–168.
- Lammerding, J., Fong, L.G., Ji, J.Y., *et al.*, 2006. Lamins A and C but not lamin B1 regulate nuclear mechanics. *Journal of Biological Chemistry* 281, 25768–25780.
- Langbein, L., Schweizer, J., 2005. Keratins of the human hair follicle. *International Review of Cytology* 243, 1–78.
- Liem, R.K.H., Messing, A., 2009. Dysfunctions of neuronal and glial intermediate filaments in disease. *Journal of Clinical Investigation* 119, 1814–1824.
- Magin, T.M., Vijayaraj, P., Leube, R.E., 2007. Structural and regulatory functions of keratins. *Experimental Cell Research* 313, 2021–2032. **Special Issue – Intermediate Filaments.**
- Mencarelli, C., Cioffi, S., Caroti, D., Lupetti, P., Dallai, R., 2011. Isomin: A novel cytoplasmic intermediate filament protein from an arthropod species. *BMC Biology* 9, 17.
- Moll, R., Divo, M., Langbein, L., 2008. The human keratins: Biology and pathology. *Histochemistry and Cell Biology* 129, 705–733.
- Pajerowski, J.D., Dahl, K.N., Zhong, F.L., Sammak, P.J., Discher, D.E., 2007. Physical plasticity of the nucleus in stem cell differentiation. *Proceedings of the National Academy of Sciences of the United States of America* 104, 15619–15624.
- Pan, D., Estévez-Salmerón, L.D., Stroschein, S.L., *et al.*, 2005. The integral inner nuclear membrane protein MAN1 physically interacts with the R-Smad proteins to repress signaling by the transforming growth factor- β superfamily of cytokines. *Journal of Biological Chemistry* 280, 15992–16001.
- Pekny, M., Lane, E.B., 2007. Intermediate filaments and stress. *Experimental Cell Research* 313, 2244–2254. **Special Issue – Intermediate Filaments.**
- Perrot, R., Berges, R., Bocquet, A., Eyer, J., 2008. Review of the multiple aspects of neurofilament functions, and their possible contribution to neurodegeneration. *Molecular Neurobiology* 38, 27–65.
- Peter, A., Stick, R., 2012. Evolution of the lamin protein family: What introns can tell. *Nucleus* 3, 44–59.
- Pickersgill, H., Kalverda, B., de Wit, E., *et al.*, 2006. Characterization of the *Drosophila melanogaster* genome at the nuclear lamina. *Nature Genetics* 38, 1005–1014.
- Ramms, L., Fabris, G., Windoffer, R., *et al.*, 2013. Keratins as the main component for the mechanical integrity of keratinocytes. *Proceedings of the National Academy of Sciences of the United States of America* 110, 18513–18518.
- Ridge, K.M., Linz, L., Flitney, F.W., *et al.*, 2005. Keratin 8 phosphorylation by protein kinase C δ regulates shear stress-mediated disassembly of keratin

- intermediate filaments in alveolar epithelial cells. *Journal of Biological Chemistry* 280, 30400–30405.
- Sahlgren, C.M., Pallari, H.-M., He, T., *et al.*, 2006. A nestin scaffold links Cdk5/p35 signaling to oxidant-induced cell death. *EMBO Journal* 25, 4808–4819.
- Schnittler, H.J., Schmandra, T., Drenckhahn, D., 1998. Correlation of endothelial vimentin content with hemodynamic parameters. *Histochemistry and Cell Biology* 110, 161–167.
- Seltmann, K., Fritsch, A.W., Käs, J.A., Magin, T.M., 2013. Keratins significantly contribute to cell stiffness and impact invasive behavior. *Proceedings of the National Academy of Sciences of the United States of America* 110, 18507–18512.
- Snider, N.T., Omary, M.B., 2014. Post-translational modifications of intermediate filament proteins: Mechanisms and functions. *Nature Reviews Molecular Cell Biology* 15, 163–177.
- Song, S., Landsbury, A., Dahm, R., *et al.*, 2009. Functions of the intermediate filament cytoskeleton in the eye lens. *Journal of Clinical Investigation* 119, 1837–1848.
- Sonnenberg, A., Liem, R.K.H., 2007. Plakins in development and disease. *Experimental Cell Research* 313, 2189–2203. **Special Issue – Intermediate Filaments.**
- Spann, T.P., Goldman, A.E., Wang, C., Huang, S., Goldman, R.D., 2002. Alteration of nuclear lamin organization inhibits RNA polymerase II-dependent transcription. *Journal of Cell Biology* 156, 603–608.
- Strelkov, S.V., Herrmann, H., Geisler, N., *et al.*, 2002. Conserved segments 1A and 2B of the intermediate filament dimer: Their atomic structures and role in filament assembly. *EMBO Journal* 21, 1255–1266.
- Strnad, P., Zatloukal, K., Stumpfner, C., Kulaksiz, H., Denk, H., 2008. Mallory–Denk-bodies: Lessons from keratin-containing hepatic inclusion bodies. *Biochimica et Biophysica Acta (BBA) – Molecular Basis of Disease* 1782, 764–774.
- Styers, M.L., Kowalczyk, A.P., Faundez, V., 2005. Intermediate filaments and vesicular membrane traffic: The odd couple's first dance? *Traffic* 6, 359–365.
- Toivola, D.M., Strnad, P., Habtezion, A., Omary, M.B., 2010. Intermediate filaments take the heat as stress proteins. *Trends in Cell Biology* 20, 79–91.
- Toivola, D.M., Tao, G.-Z., Habtezion, A., Liao, J., Omary, M.B., 2005. Cellular integrity plus: Organelle-related and protein-targeting functions of intermediate filaments. *Trends in Cell Biology* 15, 608–617.
- Wiche, G., Winter, L., 2011. Plectin isoforms as organizers of intermediate filament cytoarchitecture. *Bioarchitecture* 1, 14–20.
- Wilhelmsen, K., Litjens, S.H.M., Kuikman, I., *et al.*, 2005. Nesprin-3, a novel outer nuclear membrane protein, associates with the cytoskeletal linker protein plectin. *Journal of Cell Biology* 171, 799–810.
- Windoffer, R., Beil, M., Magin, T.M., Leube, R.E., 2011. Cytoskeleton in motion: The dynamics of keratin intermediate filaments in epithelia. *Journal of Cell Biology* 194, 669–678.
- Winter, H., Langbein, L., Krawczak, M., *et al.*, 2001. Human type I hair keratin pseudogene *kihHaa* has functional orthologs in the chimpanzee and gorilla: Evidence for recent inactivation of the human gene after the Pan-Homo divergence. *Human Genetics* 108, 37–42.
- Worman, H.J., Bonne, G., 2007. "Laminopathies": A wide spectrum of human diseases. *Experimental Cell Research* 313, 2121–2133. **Special Issue – Intermediate Filaments.**
- Zhang, Y.-Q., Sarge, K.D., 2008. Sumoylation regulates lamin A function and is lost in lamin A mutants associated with familial cardiomyopathies. *Journal of Cell Biology* 182, 35–39.
- Zwarger, M., Medalia, O., 2013. From lamins to lamina: A structural perspective. *Histochemistry and Cell Biology* 140, 3–12.

Further Reading

- Burke, B., Stewart, C.L., 2014. Functional architecture of the cell's nucleus in development, aging, and disease. *Current Topics in Developmental Biology* 109, 1–52.
- Herrmann, H., Aebi, U., 2004. Intermediate filaments: Molecular structure, assembly mechanism, and integration into functionally distinct intracellular scaffolds. *Annual Review of Biochemistry* 73, 749–789.
- Omary, M.B., Coulombe, P.A., 2004. *Intermediate Filament Cytoskeleton*. Amsterdam; Boston: Elsevier/Academic Press.
- Swift, J., Ivanovska, I.L., Buxboim, A., *et al.*, 2013. Nuclear lamin-A scales with tissue stiffness and enhances matrix-directed differentiation. *Science* 341, 1240104.

Relevant Websites

- <http://www.interfil.org/>
Human Intermediate Filament Database.
- <http://labs.feinberg.northwestern.edu/goldman/intro.htm>
Northwestern University.
- <http://www.ukaachen.de/kliniken-institute/institut-fuer-molekulare-und-zellulaere-anatomie/movie-gallery.html>
RWTH Aachen University.

Septins: Cytoskeletal Filaments with Structural and Regulatory Functions

ET Spiliotis and L Dolat, Drexel University, Philadelphia, PA, USA

© 2016 Elsevier Inc. All rights reserved.

Similar to microtubules, actin microfilaments and intermediate filaments, septins constitute a network of filamentous polymers, which are essential for the development and physiological functions of eukaryotic organisms. Septin filaments associate with cell membranes and the actin and microtubule cytoskeleton, exhibiting a structural and functional interdependence that distinguishes them apart from the conventional cytoskeleton. Septin filaments function as diffusion barriers and scaffolds that affect the compartmentalization of cell membranes and the spatial organization and functions of the microtubules and actin cytoskeleton (Caudron and Barral, 2009; Spiliotis and Gladfelter, 2012).

Septins are guanosine-5'-triphosphate (GTP)-binding proteins that self assemble into nonpolar oligomers and polymers (Weirich *et al.*, 2008; Mostowy and Cossart, 2012). Like actin and microtubules, septin assembly is coupled to nucleotide-binding and hydrolysis (Gasper *et al.*, 2009). Resembling the hetero-polymeric nature of intermediate filaments, septin filaments consist of various subunits, which in mammalian organisms are encoded by 13–14 different genes with alternative splicing and translation initiation sites that give rise to multiple septin isoforms (Russell and Hall, 2011).

Septin Structure, Assembly, and Dynamics

Classified under the P-loop superfamily of nucleotide-binding and hydrolyzing proteins, which includes the Ras-like GTPases and the kinesin and myosin motors, septins possess a highly conserved GTP-binding domain (G domain), which is structurally characterized by α -helices and β -sheets that alternate between flexible loops (Leipe *et al.*, 2002). Septin G domains contain the G1 Walker A (GxxxGKS/T), G3 Walker B (DxxG), and G4 (NKxD) motifs, which interact directly with atoms of the guanine nucleotide. In addition to these signature motifs,

the N- and C-termini of the septin G domains are respectively characterized by a polybasic region, which binds phosphoinositides (Zhang *et al.*, 1999; Casamayor and Snyder, 2003), and a conserved sequence of unknown function termed the septin unique element (SUE) (Versele and Thorner, 2005). Beyond these homology domains, septins differ significantly in the length and sequence of their N- and C-terminal extensions, which flank their G domains and consist of proline-rich and α -helical domains, respectively (Pan *et al.*, 2007).

Based on amino acid similarity, mammalian septin paralogs are categorized into four groups: SEPT2, SEPT6, SEPT7 and SEPT9 (Pan *et al.*, 2007). The majority of septins fall under the SEPT2 group, which includes SEPT1, SEPT2, SEPT4, and SEPT5, and the SEPT6 group, which consists of SEPT6, SEPT8, SEPT10, SEPT11, and SEPT14 (Figure 1). Lacking a critical threonine residue from their G domains, septins of the SEPT6 group are posited to be hydrolytically inactive and always bound to GTP (Sirajuddin *et al.*, 2009; Zent and Wittinghofer, 2014). The remaining septins are classified under the SEPT9 group, which comprises SEPT3, SEPT9, and SEPT12, and the SEPT7 group, which contains SEPT7 isoforms (Figure 1). SEPT2, SEPT7, and SEPT9 are nearly ubiquitously expressed, while SEPT1, SEPT3, SEPT12, and SEPT14 are tissue-specific (Dolat *et al.*, 2014a). The remaining septins are expressed widely, but not present in every tissue (Dolat *et al.*, 2014a).

Septin polymerization is driven by the propensity of their G domains to homo- and hetero-dimerize (Sirajuddin *et al.*, 2007). In the presence of guanosine nucleotide, septin G domains dimerize by interacting directly through their GTP-binding pockets (G interface), which are oriented opposite to one another, or, alternatively, through their N- and C-terminal helices (NC interface). Using the G and NC dimerization interfaces in tandem, septin monomers assemble linearly into nonpolar oligomers (Sirajuddin *et al.*, 2007). In mammalian systems, the basic septin oligomer is a hetero-octamer

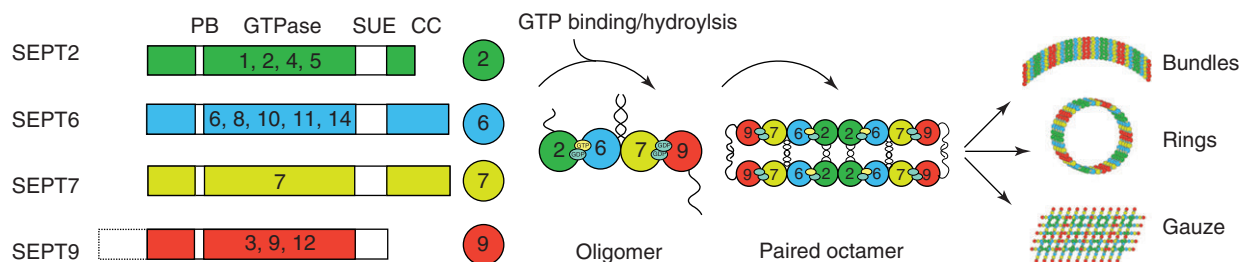


Figure 1 Septin filament assembly. Mammalian septins are classified in four groups (SEPT2, SEPT6, SEPT7, and SEPT9) and characterized by highly conserved domains such as a phosphoinositide-binding polybasic domain (PD), a GTP-binding domain, a septin unique element (SUE) and C-terminal α -helical coiled coil (CC) domains. Septins form nonpolar hetero-oligomeric complexes by dimerizing in tandem using their GTP-binding domains, which interact with one another via their N- and C-terminal helices (NC interface) or their GTP-binding pockets (G interface). GTP hydrolysis favors oligomerization by stabilizing the G interface, but septins of the SEPT6 group do not hydrolyze GTP. End-to-end assembly of septin hetero-oligomers results in the formation of nonpolar filaments, which interact laterally through the N- and C-terminal extensions of their septin subunits. Owing to the flexibility of the NC interface, paired septin filaments can bend into curved and ring-like bundles. In addition to septin bundles and rings, membrane-associated septins form gauze-like structures, which consist of orthogonally arranged filaments. Septin filaments and structures are not drawn to scale.

composed of septins from each of the SEPT2, SEPT6, SEPT7, and SEPT9 groups at a 2:2:2:2 stoichiometry (SEPT9-SEPT7-SEPT6-SEPT2-SEPT2-SEPT6-SEPT7-SEPT9) (Kim *et al.*, 2011; Sellin *et al.*, 2011b). This mode of hetero-octameric assembly is conserved between mammals and yeast (Bertin *et al.*, 2008; Farkasovsky *et al.*, 2005), but septin filaments could also consist of hetero-hexameric (SEPT7-SEPT6-SEPT2-SEPT2-SEPT6-SEPT7) and – tetrameric (SEPT6-SEPT7-SEPT7-SEPT6) units (Kinoshita, 2003; Sellin *et al.*, 2014). In addition, ‘non-canonical’ hetero-oligomers that consist of multiple subunits of the same septin group may also exist (Dolat *et al.*, 2014a). Overall, septin hetero-oligomers could vary in size and composition depending on cell and tissue type. Irrespective of subunit identity, presence of guanosine-5'-diphosphate (GDP) and thus GTP hydrolysis favors oligomerization by stabilizing the G interface, while GTP-binding appears to destabilize the NC interface (Zent and Wittinghofer, 2014). With the exception of SEPT6, oligomeric septins are GDP-bound and septin polymers contain an excess of GDP relative to GTP (Sirajuddin *et al.*, 2007; Farkasovsky *et al.*, 2005). Therefore, in contrast to microtubules and actin microfilaments whose polymerization is favored by a non-hydrolyzed nucleotide, septin polymerization is favored by GTP hydrolysis (Table 1).

Septin hetero-octamers polymerize into higher order structures such as linear and curved filaments, rings and gauze-like meshworks (Kinoshita *et al.*, 2002; Bertin *et al.*, 2008; Garcia *et al.*, 2011). End-to-end binding of septin hetero-octomers results in filaments, which are 5–10 nm wide and up to several microns long (Kinoshita *et al.*, 2002; Bertin *et al.*, 2008, 2010). Septin filaments pair with one another by interacting laterally through the C- (coiled coil domains) and possibly N-terminal extensions that stretch outwards from the plane of oligomerization (Bai *et al.*, 2013; John *et al.*, 2007; Bertin *et al.*, 2008; de Almeida Marques *et al.*, 2012). Thus, septin bundles form by lateral stacking of individual septin filaments and adapter proteins further enhance septin bundling (Sadian *et al.*, 2013). Owing to the flexibility of the SEPT2 NC interface, which behaves like a hinge, septin oligomers can bend within a 25–30° angle of one another generating a curvature, which allows for the formation ring-like structures with 0.5–0.7 µm diameter (Sirajuddin *et al.*, 2007). Septin subunits that enhance the flexibility of NC interface favor the formation of septin rings and posttranslational modifications can promote the assembly of a gauze-like meshwork of orthogonally arranged filaments (Garcia *et al.*, 2011). The

filamentous pattern of septins is further influenced by their association with membrane bilayers and the actin and microtubule cytoskeleton. Membranes enriched with phosphatidylinositol-4,5-bisphosphate promote the formation of tightly paired septin filaments (Bertin *et al.*, 2010; Tanaka-Takiguchi *et al.*, 2009). Conversely, actin-templated assembly of septins favors the formation of linear filaments as opposed to septin rings, which become more prevalent upon actin depolymerization (Kinoshita *et al.*, 2002; Schmidt and Nichols, 2004b). Similarly, the disk-like organization of septins in non-adherent cells depends on microtubules (Sellin *et al.*, 2011a; Martinez *et al.*, 2006).

Septin filaments are less dynamic than microtubules and F-actin. Half-times of subunit exchange and turnover could vary from 10 s of seconds for cytoplasmic septin filaments in mammalian cells to 3 min for membrane-associated septin polymers in yeast organisms (Schmidt and Nichols, 2004b; Hagiwara *et al.*, 2011; Hu *et al.*, 2008; Dobbelaere *et al.*, 2003; DeMay *et al.*, 2010). Recent studies show that membrane-bound septin filaments grow by end-to-end annealing (Bridges *et al.*, 2014). Rod-shaped septin oligomers diffuse, collide and anneal to the free ends of septin filaments, which do not exchange any subunits along their length (Bridges *et al.*, 2014). In contrast to membrane-bound septin filaments, which incorporate new subunits after fragmentation, cytoplasmic septin filaments are more dynamic, exchanging subunits along their length (Schmidt and Nichols, 2004b; Hu *et al.*, 2008; Bridges *et al.*, 2014). Septin filament dynamics are affected by posttranslational modifications such as phosphorylation and acetylation, and vary depending on the phase of the cell cycle (Hernandez-Rodriguez and Momany, 2012). For example, in the budding yeast *Saccharomyces cerevisiae*, septin filaments are dynamic in G1, but highly stable during the S, G2 and M phases of the cell cycle (Dobbelaere *et al.*, 2003). These changes in septin dynamics are mediated by septin-dependent kinases and phosphatases that are recruited to septin filaments in the beginning and end of the cell cycle, respectively (Dobbelaere *et al.*, 2003).

The molecular mechanisms that regulate septin assembly and dynamics are not fully understood. Recent advances indicate that regulation of GTP hydrolysis, septin synthesis and degradation as well as posttranslational modifications and protein–protein interactions influence the filamentous organization and dynamics of septins. Septin folding is mediated by the cytosolic chaperonin-containing TCP-1 (CCT), which also

Table 1 Septins in comparison to actin microfilaments, microtubules and intermediate filaments

	<i>Septins</i>	<i>F-actin</i>	<i>Microtubules</i>	<i>Intermediate filaments</i>
Subunit(s)	SEPT1-14	α -/ β -/ γ -actin	α -/ β -tubulin	Keratin, vimentin, lamin, desmin, peripherin, nestin, and neurofilament proteins
Nucleotide	GTP	ATP	GTP	None
Diameter	~5 nm	~7 nm	~25 nm	~10 nm
Length	25–32 nm to several microns	Several microns	Several microns	Several microns
Structure	Straight or curved paired filaments, rings and gauzes	Right-handed helix	Hollow tube	Twisted filaments
Polarity	No	Yes	Yes	No
Subunit turnover ($t_{1/2}$)	10–200 s	10 s of seconds	10 s of seconds	10 s of minutes

controls the folding of actin and tubulin, and septin degradation involves E3 ubiquitin ligases including parkin, the von Hippel-Lindau (VHL) tumor suppressor and the Ring Finger protein 8 (RNF8) (Craven *et al.*, 2006; Chahwan *et al.*, 2013; Dekker *et al.*, 2008; Zhang *et al.*, 2000). While these factors could regulate the availability of septin subunits for polymerization, proteins that promote the GTPase activity of septins (e.g., Orc6) and crosslink septin filaments (e.g., Gic1) can directly enhance septin assembly (Sadian *et al.*, 2013; Huijbregts *et al.*, 2009). In contrast, Cdc42, a signaling molecule of the Rho family of GTPases, induces the unbundling and dissociation of septin filaments (Sadian *et al.*, 2013). Post-translational modifications such as phosphorylation, acetylation and sumoylation further modulate septin assembly and dynamics, but many of the details remain unknown (Hernandez-Rodriguez and Momany, 2012).

Septins and the Membrane Skeleton

Septin filaments assemble on the cytoplasmic leaflet of membrane bilayers forming a meshwork that resembles the spectrin-actin membrane skeleton. In yeast cells, septins form filamentous rings and gauze-like patches in actin-free areas of the cell cortex (Oh and Bi, 2011; McMurray and Thorner, 2009), while in mammalian cells septin filaments associate with the actin membrane skeleton (Hagiwara *et al.*, 2011). These membrane-bound septin filaments restrict protein diffusion and affect the shape and rigidity of cell membranes (Spiliotis and Gladfelter, 2012; Caudron and Barral, 2009).

In the budding yeast *S. cerevisiae*, a membrane-bound septin ring forms at the base of the emerging bud early in the cell cycle and evolves to a collar that underlies the mother-bud neck (McMurray and Thorner, 2009; Oh and Bi, 2011). Septin filaments are positioned between the plasma membrane and the smooth endoplasmic reticulum (ER), impeding the lateral diffusion of cortical and ER membrane proteins during polarized bud growth (Barral *et al.*, 2000; Luedeke *et al.*, 2005). In mitosis, splitting of the septin collar into two rings results in the formation of a membrane compartment, which enables the accumulation of membrane-associated factors (e.g., vesicle tethering complexes) without diffusing into the mother and bud cell membranes (Dobbelaere and Barral, 2004). During mammalian mitosis, a septin diffusion barrier is posited to similarly limit the diffusion of inner leaflet membrane proteins across the equatorial plane of a dividing cell (Schmidt and Nichols, 2004a).

Septins form membrane-bound rings at the base of the primary cilium and the dendritic spines of neurons (Hu *et al.*, 2010; Tada *et al.*, 2007; Xie *et al.*, 2007). At the interface of the plasma and ciliary membrane, septins control the localization of ciliary membrane proteins and thus, they are required for the formation and maintenance of the primary cilium (Hu *et al.*, 2010). Septins could similarly impede the lateral diffusion of post-synaptic proteins from the dendritic spines to the shaft, aiding to the compartmentalization of synaptic receptors and channels (Caudron and Barral, 2009). At the annulus of spermatozoa, septins form a barrier between the principal and mid pieces of the sperm, and are required for the proper localization of a protein that is essential for

spermiogenesis (Kwitny *et al.*, 2010). The mechanism by which septins restrict lateral diffusion in cell membranes is unknown, but recent evidence indicates that septins influence the organization of lipid domains and interact with the cytoplasmic tails of membrane transporters (Sharma *et al.*, 2013; Hagiwara *et al.*, 2011). Protein-lipid compartmentalization and domain identity could be further reinforced by septin-mediated inhibition of vesicle delivery; septins interact with the exocyst, a vesicle tethering complex, and components of the soluble N-ethylmaleimide-sensitive factor activating protein receptor (SNARE) machinery of vesicle fusion, and may introduce a physical barrier that interferes with vesicle delivery to the plasma membrane (Beites *et al.*, 1999; Hsu *et al.*, 1998; Yang *et al.*, 2010). Hence, septins are an integral component of the structure of cell membranes, contributing to the organization of membrane domains by inhibiting selectively the lateral diffusion and delivery of membrane proteins.

In addition to their diffusion barrier properties, membrane-bound septins affect the shape and physical properties of lipid bilayers. Similar to the membrane-remodeling functions of the Bin-Amphiphysin-Rvs (BAR)-domain proteins, septins have been shown to induce the tubulation of spherical liposomes made of phosphatidylcholine and phosphoinositides (Tanaka-Takiguchi *et al.*, 2009). These membrane tubules have the diameter of septins rings ($\sim 0.5 \mu\text{m}$) and are circumferentially braced by septin filaments (Tanaka-Takiguchi *et al.*, 2009). It is unknown, however, whether septins affect membrane curvature and induce membrane tubules *in vivo*. Conversely, several studies have shown that septin depletion decreases the rigidity and stiffness of cell membranes (Tooley *et al.*, 2009; Mostowy *et al.*, 2011). In T lymphocytes, septins suppress membrane blebbing and protrusive activity is spatially restricted to septin-free areas of the cell cortex (Tooley *et al.*, 2009; Gilden *et al.*, 2012). Measurements of membrane elasticity and viscosity show a marked decrease in the rigidity of cell membranes after septin depletion, which alters cell shape and morphology (Mostowy *et al.*, 2011). Thus, like the spectrin-actin membrane skeleton, septin filaments contribute to the organization and mechanochemical properties of cell membranes.

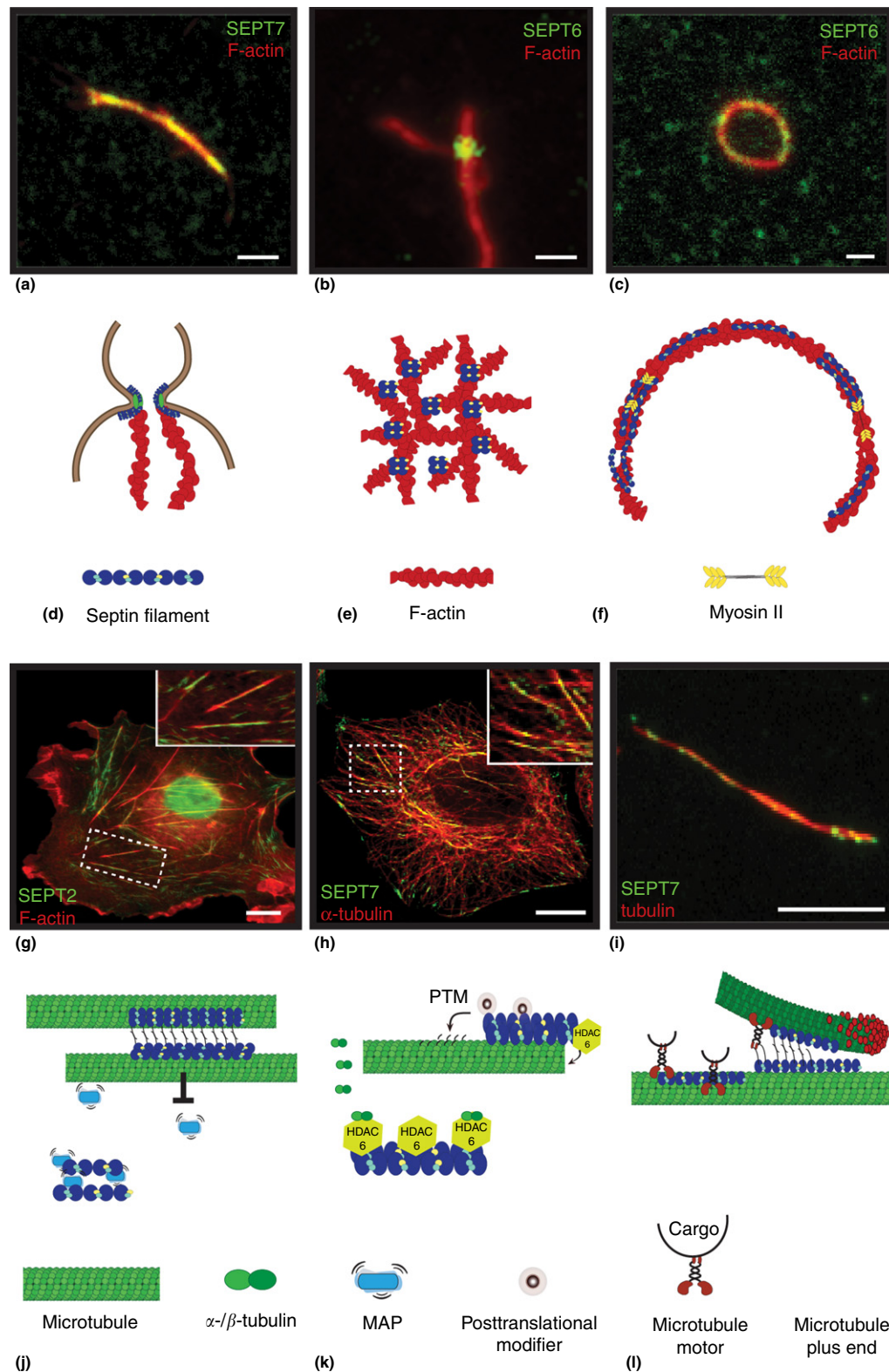
Septins and the Actin Cytoskeleton

From yeast to mammals, septins are implicated in the organization and functions of the actomyosin cytoskeleton. Septins crosslink actin filaments and interact with actin-binding proteins and signaling effectors that regulate the organization of the actin cytoskeleton (Mavrakakis *et al.*, 2014; Nagata and Inagaki, 2005; Kinoshita *et al.*, 2002). Notably, septins interact directly with myosin II, regulating its localization and activation in interphase and mitotic cells (Joo *et al.*, 2007).

In the budding yeast *S. cerevisiae*, septins are essential for the recruitment of proteins involved in actin cable assembly during bud growth. At the site of bud formation, linear actin cables, which enable the polarized transport of vesicles, are polymerized by the formins Bni1 and Bnr1 (Bretscher, 2013). In the absence of the nonessential septin Shs1, Bnr1 is mis-localized and actin cables fail to assemble (Gao *et al.*, 2010; Buttery *et al.*, 2012). The actin assembly activity of Bnr1 appears to depend on Shs1 (Buttery *et al.*, 2012) and the septin-

interacting partner Bni5 is required for the recruitment of myosin II, which drives actin cable retrograde flow (Nam *et al.*, 2007). Thus, septins scaffold and may activate the assembly of actin cables during bud growth. Interestingly, mammalian

septins scaffold the recruitment of cortactin, an activator of Arp2/3, to actin patches in the axons of sensory neurons (Hu *et al.*, 2012). Septin 6 (SEPT6) localizes to actin patches, which consist of branched actin filaments, and SEPT6 depletion



reduces the recruitment of cortactin to actin patches and their transition to filopodia (Hu *et al.*, 2012). Reconstitution of branched actin assembly *in vitro* and in the presence of Arp2/3 shows that SEPT6 localizes preferentially to actin branch points (Hu *et al.*, 2012). In agreement with this observation, overexpression of SEPT6 increases the recruitment of cortactin to the lamellipodium, a branched actin network at the leading edge of motile cells (Hu *et al.*, 2012). Additional septins (e.g., SEPT1 and SEPT4) are enriched in the lamellipodia of squamous carcinoma cells (Mizutani *et al.*, 2013). Septins, therefore, have an evolutionarily conserved role in scaffolding proteins involved in the formation of linear and branched actin filaments.

More recently, septins were found to crosslink and bundle actin filaments directly, contributing to the assembly of contractile actin rings in *Drosophila* embryos (Mavrakakis *et al.*, 2014). In developing *Drosophila* embryos, the blastoderm contains thousands of nuclei, which become individual cells. Cellularization requires the assembly of contractile rings of actomyosin filaments that furrow the plasma membrane resulting in membrane canals that envelope the nuclei of the syncytial embryo. Embryos expressing a septin mutant are characterized by a loss of circular actin filaments, which are less bundled and contract at slower rates (Mavrakakis *et al.*, 2014). *Drosophila* septins bundle actin filaments and copolymerization of actin with septins results in curved and circular actin bundles (Mavrakakis *et al.*, 2014). Mammalian septins also bundle actin filaments *in vitro* and septin depletion results in loss of actin stress fibers (Schmidt and Nichols, 2004b; Kremer *et al.*, 2007; Dolat *et al.*, 2014b; Kinoshita *et al.*, 2002). These findings indicate that septins are actin-binding and bundling proteins that regulate the organization of actin microfilaments (Figures 2(a)–(g)).

In addition to their direct roles in actin organization, septins influence actin organization and contractility by interacting with signaling factors and myosin II. Assembly of stress fibers is under the control of Rho GTPases, which activate actin nucleation factors and promote the contraction of myosin II (Narumiya *et al.*, 2009). Several studies have indicated that septins are involved in the Rho signaling pathways that control stress fiber assembly. Septin 9 (SEPT9), which is positioned at the terminal ends of septin hetero-octamers, interacts directly with a Rho guanine exchange factor (GEF), termed the septin-

associated Rho GEF (SA-Rho GEF), and Rhotekin, a secondary effector of Rho signaling (Nagata and Inagaki, 2005; Ito *et al.*, 2005). Both SA-Rho GEF and Rhotekin regulate septin assembly and localization to stress fibers. Interestingly, SEPT9 appears to negatively regulate Rho activation by inhibiting GTP loading by the SA-Rho GEF (Nagata and Inagaki, 2005). Septins also interact with Binders of Rho GTPases (Borgs), which are effectors of Cdc42 (Joberty *et al.*, 2001). New studies show that Gic1, a yeast homolog of the human Borg, binds and cross-links septin filaments *in vitro* in a Cdc42-dependent manner (Sadian *et al.*, 2013). Binding of the active Cdc42-GTP to Gic1 results in the dissociation of Gic1 from septin filaments. The inactive Cdc42-GDP, however, also inhibits the assembly of septin filaments by competing with Gic1 for binding to septins (Sadian *et al.*, 2013). In mammalian cells, Borgs have been shown to affect both septin and stress fiber assembly (Joberty *et al.*, 2001).

While more work is needed to elucidate how the interactions between septins and signaling molecules affect actin organization, structural and functional analysis of the septin-myosin II interaction has provided a novel mechanism for the activation of the non-muscle myosin II. The mammalian SEPT2 has been shown to interact directly with the heavy chain of the non-muscle myosin II and yeast two-hybrid data indicate that Myo1, the *S. cerevisiae* type II myosin, binds the sporulation-specific septin Spr3 (Joo *et al.*, 2007; Drees *et al.*, 2001). Disruption of the SEPT2-myosin II interaction results in loss of stress fibers in interphase cells and incomplete ingression of the cleavage furrow during mitosis (Joo *et al.*, 2007). These defects are accompanied by significant reduction in the phosphorylation of the myosin light chain, which stimulates myosin II contractility. Importantly, SEPT2 appears to scaffold the interaction of myosin II with the citron and Rho-activated myosin kinases, which in the absence of septins dissociate from the cleavage furrow (Joo *et al.*, 2007). Thus, septins may control spatially and temporally the localization and activation of myosin II. Importantly, this function might be evolutionarily conserved as myosin activation and actin organization is also disrupted in *Xenopus* embryos after Sept7 depletion, decreasing planar tension and protrusive activity during the collective cell migration that drives body axis elongation (convergent extension) (Shindo and Wallingford, 2014).

Figure 2 Septins interact with the actin and microtubule cytoskeleton. (a–e) Fluorescence images show purified recombinant septins tagged with the green fluorescent protein decorating linear (a), branched (b) and circular (c) actin filaments. Prepolymerized rhodamine-actin (3 μ M) decorated with 100 nM SEPT7-GFP (a), or 400 nM SEPT6-GFP in the presence (b) or absence (c) of unlabeled 15 nM Arp 2/3 and 100 nM VCA (Reproduced from Hu, J., Bai, X., Bowen, J.R., *et al.* 2012. Septin-driven coordination of actin and microtubule remodeling regulates the collateral branching of axons. *Current Biology* 22, 1109–1115.). Recent evidence indicates that septins crosslink actin filaments into curved and circular bundles (Mavrakakis *et al.*, 2014) and scaffold the recruitment of cortactin, an actin nucleation promoting factor, at Arp2/3-mediated branch points (Hu *et al.*, 2012). In the budding yeast *Saccharomyces cerevisiae*, septins are required for the localization and activity of formins, which stimulate the formation of actin cables (d). In the axons of sensory neurons, septins accumulate to actin patches (e), which consist of Arp2/3-nucleated branched filaments, and in *Drosophila* embryos, septins are essential for the assembly of the contractile rings of actomyosin (f), which drives membrane invagination and furrow ingression during cellularization. Scale bars, 1 μ m. (g–l) Fluorescent images show colocalization of septin filaments with actin stress fibers in Cos7 cells (g) and microtubules in Madin-Darby canine kidney cells (h), and *in vitro* decoration of 10 μ m prepolymerized rhodamine-microtubules with 20 nM purified recombinant SEPT7-GFP (Reproduced from Bai, X., Bowen, J.R., Knox, T.K., *et al.*, 2013. Novel septin 9 repeat motifs altered in neuralgic amyotrophy bind and bundle microtubules. *Journal of Cell Biology* 203, 895–905.). (i). Septins bundle microtubules and inhibit their interaction with other microtubule-associated proteins (j). Posttranslational modifications (PTM) of tubulin such as acetylation and polyglutamylation are influenced by septins, which can scaffold the interaction of tubulin with the deacetylase HDAC6 (k). In addition, septins have been shown to guide spatially microtubule-microtubule interactions and may influence the localization of microtubule motors (l). Scale bars, 10 μ m.

In summary, septins are integral components of the actomyosin cytoskeleton. Recent studies have revealed that septins interact directly with actin filaments and myosin II, but many of the molecular and functional details remain unknown. Future studies will address the mechanisms by which septins affect the formation of linear and branched actin filaments and myosin II-dependent contractility, and how septins interface with signaling pathways that affect actin organization and contractility.

Septins and the Microtubule Cytoskeleton

Septin filaments colocalize with microtubules in a diversity of eukaryotic organisms and cells. The extent of septin-microtubule colocalization varies between septin paralogs and isoforms, and depends on cell type, phase of the cell cycle and tubulin isoform expression and posttranslational modifications. Recent studies have shown that septins interact directly with microtubules, microtubule-associated proteins (MAPs) and motors as well as enzymes that modify tubulin's posttranslational modification. Thus, septins have emerged as novel regulators of microtubule organization and dynamics, and may play a key role in the spatial regulation of microtubule-dependent transport (**Figures 2(h)-(l)**).

Mammalian SEPT2/6/7/9 complexes interact directly with microtubules through their SEPT9, SEPT7, and SEPT6 subunits; SEPT2 does not bind microtubules (**Bai et al., 2013**). It is unknown how SEPT6 and SEPT7 associate with microtubules, but SEPT9 interacts with the acidic tails of β II-tubulin via its N-terminal extension, which contains multiple repeats of the K/R-R/x-x-D/E motif (**Bai et al., 2013**). Similar to the microtubule-binding sequence repeats of filamentous MAPs (e.g., tau, MAP1A), the SEPT9 repeat motifs consist of positively and negatively charged residues, which interact electrostatically with one another and the acidic C-terminal tails of β -tubulin (**Bai et al., 2013**). Thus, the N-termini of SEPT9 mediate septin-septin and septin-microtubule interactions that crosslink microtubules into bundles. By forming filamentous cross-bridges that bind microtubules in a staggered arrangement, septins could increase the thickness and length of microtubule bundles as observed *in vitro* (**Bai et al., 2013**). Consistent with this function, septins colocalize preferentially with elongated microtubule bundles in cultured kidney epithelia (**Bowen et al., 2011**). Moreover, expression of SEPT9 isoforms that lack N-terminal repeat motifs reduces both microtubule bundling and septin-microtubule colocalization (**Bai et al., 2013**). Hence, through the properties of SEPT9, mammalian septins function as *bona fide* MAPs that facilitate the bundling of microtubules.

Like most MAPs, septins affect not only microtubule bundling, but also microtubule dynamics. In kidney epithelia, SEPT2 knock-down increases microtubule catastrophe and reduces the rate of microtubule growth, indicating that septins are required for persistent microtubule growth (**Bowen et al., 2011**). Interestingly, loss of SEPT2 also affects microtubule-microtubule interactions and the overall directionality of microtubule growth (**Bowen et al., 2011**). In agreement with a septin requirement for microtubule growth and bundling, SEPT9 knock-down decreased the density of the microtubule

network in mammary epithelia (**Bai et al., 2013**). Depletion of SEPT7, however, has the opposite effect in HeLa cells, increasing microtubule density and acetylation, a marker of microtubule stability (**Kremer et al., 2005**). A direct interaction between SEPT2 and MAP4 indicates that cytoplasmic SEPT2/6/7 complexes inhibit MAP4 from binding and bundling microtubules, but SEPT2/6/7 and MAP4 also compete for binding to microtubules and a new study shows that SEPT7 mediates the interaction of α -tubulin with the deacetylase histone deacetylase 6 (HDAC6) (**Spiliotis et al., 2008**; **Ageta-Ishihara et al., 2013**). Surprisingly, SEPT7 is required for α -tubulin deacetylation and SEPT7-depleted neurons have increased levels of acetylated tubulin and reduced rates of microtubule growth (**Ageta-Ishihara et al., 2013**). While the cause-and-effect relationship between microtubule 'stability' and acetylation is unclear, septins appear to affect microtubule bundling, dynamics, and posttranslational modifications including acetylation and polyglutamylation. Given that septins interact with β -tubulin, MAP4 and HDAC6 in a paralog- and isoform-specific manner, it is possible that septins have distinctly different roles, targeting different aspects of microtubule organization and dynamics.

Microtubule posttranslational modifications and MAPs affect the binding and processive movement of motors, generating a code for the spatial regulation of intracellular transport (**Verhey and Gaertig, 2007**). Presence of septins on select microtubule tracks (e.g., perinuclear microtubule bundles) raises the possibility that septins regulate microtubule-dependent transport (**Spiliotis et al., 2008**). In kidney epithelial cells, tubular vesicular carriers egress from the trans-Golgi network on septin-coated microtubule tracks (**Spiliotis et al., 2008**). Microinjection of septin function-blocking antibodies or SEPT2 depletion results in stalling of vesicle transport *en route* to the plasma membrane (**Spiliotis et al., 2008**). Competition between SEPT2 and MAP4 for binding to microtubules suggests that septins could block the microtubule binding of MAP4, which inhibits kinesin-dependent transport (**Spiliotis et al., 2008**). It is unknown whether septins affect directly the binding or processive movement of microtubule motors, but SEPT7 interacts directly with the mitotic motor centromere associated protein E (CENP-E) (kinesin 7) (**Zhu et al., 2008**). Interestingly, septins are required for the proper localization of CENP-E and consequently, for the stable capture and biorientation of the kinetochores of congressing chromosomes (**Spiliotis et al., 2005**). Because SEPT7 interacts with the C-terminal tail of CENP-E, which binds and inhibits the motor domain of CENP-E, SEPT7 could promote the activation and motility of CENP-E by relieving its auto-inhibition. Alternatively, septins may scaffold the interaction of CENP-E with the Aurora A and B kinases or the protein phosphatase 1 (PP1), which regulate CENP-E binding to microtubules; Aurora B has been reported to interact directly with SEPT1 (**Qi et al., 2005**).

For many years now, the association of microtubules with septins has been an intriguing and contentious observation. Recent studies show that septins interact directly with microtubules, affecting their organization, dynamics and posttranslational modifications. However, more work is needed to determine if and how septins influence the binding and processivity of microtubule motors and whether septins could be

part of the molecular mechanisms that regulate spatially the transport of intracellular cargo.

Conclusions

Septins comprise a network of filaments that interface with cell membranes and the actin and microtubule cytoskeleton. Septins are integral components of the cell membrane's skeleton impeding lateral diffusion and vesicle delivery, and thereby, affecting the generation and maintenance of distinct membrane compartments and domains. In the cytoplasm, septins act as filamentous actin- and MAPs that regulate the spatial organization and functions of the cytoskeleton. Septins interact directly with actomyosin and actin-binding proteins, affecting the localization, formation, cross-linking, and contractility of actin microfilaments. Septins bundle microtubules and appear to regulate the organization, dynamics, and post-translational modifications of microtubules as well as their interactions with motors and MAPs. Thus, septins fine-tune the spatial organization and functions of the cell, regulating a diversity of cellular processes ranging from cell division and motility to cell death and autophagy.

Abnormalities in septin expression and septin mutations underlie a number of disease states (Dolat *et al.*, 2014a). Male infertility and Hereditary Neuralgic Amyotrophy (HNA), an autosomal dominant neuromuscular disorder, are linked to septin gene mutations (Kuo *et al.*, 2012; Kühlenbaumer *et al.*, 2005). Blood disorders such as the Bernard–Soulier Syndrome and acute myeloid and lymphoblastic leukemias are characterized by septin gene deletions and fusions with the mixed lineage leukemia (MLL) gene (Cerveira *et al.*, 2011; Bartsch *et al.*, 2011). Septins are abnormally expressed in many solid tumors of epithelial origin including breast, ovarian, renal and colorectal carcinomas, and changes in epigenetic marks of septin expression have been utilized as diagnostic biomarkers (Connolly *et al.*, 2011). Alterations in septin expression have also been reported in neurodevelopmental disorders (e.g., Down syndrome and schizophrenia) and septin aggregates have been found in the neurofibrillary tangles and Lewy bodies of patients with neurodegenerative disorders such as Alzheimer's and Parkinson's (Dolat *et al.*, 2014a). Recent studies indicate that septins are also important of host defense against pathogenic bacteria (Mostowy and Cossart, 2011). While more studies are required to elucidate how abnormalities in septin expression contribute to the pathogenesis of human diseases, septins are integral components of the cytoskeleton and its functions. Future studies will enhance our understanding of septins as regulators of the spatial organization of the cell.

See also: Cell Communication: Cellular Outcomes: Regulation of Cell Migration. Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell Migration; Filopodia and Lamellipodia; Skeletal Muscle. Cytoskeleton and Motors: Cytoskeletal Components: Actin Assembly Dynamics and Its Regulation in Motile and Morphogenetic Processes; Dyneins; Kinesin Superfamily Proteins (KIFs) as a Fundamental Component of Life: Intracellular Transport and Beyond. Dynamic Integration:

Dynamics: Dynamics of Microtubule Self-Assembly; Modeling Actin Dynamics. Interorganellar Communication: Interplay and Processes: Rabs and Other G Proteins

References

- Ageta-Ishihara, N., Miyata, T., Ohshima, C., *et al.*, 2013. Septins promote dendrite and axon development by negatively regulating microtubule stability via HDAC6-mediated deacetylation. *Nature Communications* 4, 2532.
- de Almeida Marques, I., Valadares, N.F., Garcia, W., *et al.*, 2012. Septin C-terminal domain interactions: Implications for filament stability and assembly. *Cell Biochemistry and Biophysics* 62, 317–328.
- Bai, X., Bowen, J.R., Knox, T.K., *et al.*, 2013. Novel septin 9 repeat motifs altered in neurogenic amyotrophy bind and bundle microtubules. *Journal of Cell Biology* 203, 895–905.
- Barral, Y., Mermall, V., Mooseker, M.S., Snyder, M., 2000. Compartmentalization of the cell cortex by septins is required for maintenance of cell polarity in yeast. *Molecular Cell* 5, 841–851.
- Bartsch, I., Sandrock, K., Lanza, F., *et al.*, 2011. Deletion of human GP1BB and SEPT5 is associated with Bernard-Soulier syndrome, platelet secretion defect, polymicrogyria, and developmental delay. *Thrombosis and Haemostasis* 106, 475–483.
- Beites, C.L., Xie, H., Bowser, R., Trimble, W.S., 1999. The septin CDCrel-1 binds syntaxin and inhibits exocytosis. *Nature Neuroscience* 2, 434–439.
- Bertin, A., McMurray, M.A., Grob, P., *et al.*, 2008. Saccharomyces cerevisiae septins: supramolecular organization of heterooligomers and the mechanism of filament assembly. *Proceedings of the National Academy of Sciences of the United States of America* 105, 8274–8279.
- Bertin, A., McMurray, M.A., Thai, L., *et al.*, 2010. Phosphatidylinositol-4,5-bisphosphate promotes budding yeast septin filament assembly and organization. *Journal of Molecular Biology* 404, 711–731.
- Bowen, J.R., Hwang, D., Bai, X., Roy, D., Spiliotis, E.T., 2011. Septin GTPases spatially guide microtubule organization and plus end dynamics in polarizing epithelia. *Journal of Cell Biology* 194, 187–197.
- Bretscher, A., 2013. Deconstructing formin-dependent actin cable assembly. *Proceedings of the National Academy of Sciences of the United States of America* 110, 18744–18745.
- Bridges, A.A., Zhang, H., Mehta, S.B., *et al.*, 2014. Septin assemblies form by diffusion-driven annealing on membranes. *Proceedings of the National Academy of Sciences of the United States of America* 111, 2146–2151.
- Buttery, S.M., Kono, K., Stokasimov, E., Pellman, D., 2012. Regulation of the formin Bnr1 by septins and MARK/Par1-family septin-associated kinase. *Molecular Biology of the Cell* 23, 4041–4053.
- Casamayor, A., Snyder, M., 2003. Molecular dissection of a yeast septin: Distinct domains are required for septin interaction, localization, and function. *Molecular and Cellular Biology* 23, 2762–2777.
- Caudron, F., Barral, Y., 2009. Septins and the lateral compartmentalization of eukaryotic membranes. *Developmental Cell* 16, 493–506.
- Cerveira, N., Bizarro, S., Teixeira, M.R., 2011. MLL-SEPTIN gene fusions in hematological malignancies. *Journal of Biological Chemistry* 392, 713–724.
- Chahwan, R., Gravel, S., Matsusaka, T., Jackson, S.P., 2013. Dma/RNF8 proteins are evolutionarily conserved E3 ubiquitin ligases that target septins. *Cell Cycle* 12, 1000–1008.
- Connolly, D., Abdesselam, I., Verdier-Pinard, P., Montagna, C., 2011. Septin roles in tumorigenesis. *Biological Chemistry* 392, 725–738.
- Craven, R.A., Hanrahan, S., Totty, N., *et al.*, 2006. Proteomic identification of a role for the von Hippel Lindau tumour suppressor in changes in the expression of mitochondrial proteins and septin 2 in renal cell carcinoma. *Proteomics* 6, 3880–3893.
- Dekker, C., Stirling, P.C., McCormack, E.A., *et al.*, 2008. The interaction network of the chaperonin CCT. *EMBO Journal* 27, 1827–1839.
- Demay, B.S., Meseroll, R.A., Occhipinti, P., Gladfelter, A.S., 2010. Cellular requirements for the small molecule forchlorfenuron to stabilize the septin cytoskeleton. *Cytoskeleton (Hoboken)* 67, 383–399.
- Dobbelaere, J., Barral, Y., 2004. Spatial coordination of cytokinetic events by compartmentalization of the cell cortex. *Science* 305, 393–396.
- Dobbelaere, J., Gentry, M.S., Hallberg, R.L., Barral, Y., 2003. Phosphorylation-dependent regulation of septin dynamics during the cell cycle. *Developmental Cell* 4, 345–357.

- Dolat, L., Hu, Q., Spiliotis, E.T., 2014a. Septin functions in organ system physiology and pathology. *Biological Chemistry* 395, 123–141.
- Dolat, L., Hunyara, J.L., Bowen, J.R., *et al.*, 2014b. Septins promote stress fiber-mediated maturation of focal adhesions and renal epithelial motility. *Journal of Cell Biology* 207, 225–235.
- Drees, B.L., Sundin, B., Brazeau, E., *et al.*, 2001. A protein interaction map for cell polarity development. *Journal of Cell Biology* 154, 549–571.
- Farkasovsky, M., Herter, P., Voss, B., Wittinghofer, A., 2005. Nucleotide binding and filament assembly of recombinant yeast septin complexes. *Biological Chemistry* 386, 643–656.
- Gao, L., Liu, W., Bretscher, A., 2010. The yeast formin Bnr1p has two localization regions that show spatially and temporally distinct association with septin structures. *Molecular Biology of the Cell* 21, 1253–1262.
- Garcia 3rd, G., Bertin, A., Li, Z., *et al.*, 2011. Subunit-dependent modulation of septin assembly: Budding yeast septin Shs1 promotes ring and gauze formation. *Journal of Cell Biology* 195, 993–1004.
- Gaspar, R., Meyer, S., Gotthardt, K., Sirajuddin, M., Wittinghofer, A., 2009. It takes two to tango: Regulation of G proteins by dimerization. *Nature Reviews Molecular Cell Biology* 10, 423–429.
- Gilden, J.K., Peck, S., Chen, Y.C., Krummel, M.F., 2012. The septin cytoskeleton facilitates membrane retraction during motility and blebbing. *Journal of Cell Biology* 196, 103–114.
- Hagiwara, A., Tanaka, Y., Hikawa, R., *et al.*, 2011. Submembranous septins as relatively stable components of actin-based membrane skeleton. *Cytoskeleton (Hoboken)* 68, 512–525.
- Hernandez-Rodriguez, Y., Momany, M., 2012. Posttranslational modifications and assembly of septin heteropolymers and higher-order structures. *Current Opinion in Microbiology* 15, 660–668.
- Hsu, S.C., Hazuka, C.D., Roth, R., *et al.*, 1998. Subunit composition, protein interactions, and structures of the mammalian brain sec6/8 complex and septin filaments. *Neuron* 20, 1111–1122.
- Hu, J., Bai, X., Bowen, J.R., *et al.*, 2012. Septin-driven coordination of actin and microtubule remodeling regulates the collateral branching of axons. *Current Biology* 22, 1109–1115.
- Hu, Q., Milenkovic, L., Jin, H., *et al.*, 2010. A septin diffusion barrier at the base of the primary cilium maintains ciliary membrane protein distribution. *Science* 329, 436–439.
- Hu, Q., Nelson, W.J., Spiliotis, E.T., 2008. Forchlorfenuron alters mammalian septin assembly, organization, and dynamics. *Journal of Biological Chemistry* 283, 29563–29571.
- Huijbregts, R.P., Svitin, A., Stinnett, M.W., Renfrow, M.B., Chesnokov, I., 2009. *Drosophila* Orc6 facilitates GTPase activity and filament formation of the septin complex. *Molecular Biology of the Cell* 20, 270–281.
- Ito, H., Iwamoto, I., Morishita, R., *et al.*, 2005. Possible role of Rho/Rhotekin signaling in mammalian septin organization. *Oncogene* 24, 7064–7072.
- Joberty, G., Perlungher, R.R., Sheffield, P.J., *et al.*, 2001. Borg proteins control septin organization and are negatively regulated by Cdc42. *Nature Cell Biology* 3, 861–866.
- John, C.M., Hite, R.K., Weirich, C.S., *et al.*, 2007. The *Caenorhabditis elegans* septin complex is nonpolar. *EMBO Journal* 26, 3296–3307.
- Joo, E., Surka, M.C., Trimble, W.S., 2007. Mammalian SEPT2 is required for scaffolding nonmuscle myosin II and its kinases. *Developmental Cell* 13, 677–690.
- Kim, M.S., Froese, C.D., Estey, M.P., Trimble, W.S., 2011. SEPT9 occupies the terminal positions in septin octamers and mediates polymerization-dependent functions in abscission. *Journal of Cell Biology* 195, 815–826.
- Kinoshita, M., 2003. Assembly of mammalian septins. *Journal of Biochemistry* 134, 491–496.
- Kinoshita, M., Field, C., Coughlin, M., Straight, A., Mitchison, T., 2002. Self- and actin-templated assembly of mammalian septins. *Developmental Cell* 3, 791–802.
- Kremer, B.E., Adang, L.A., Macara, I.G., 2007. Septins regulate actin organization and cell-cycle arrest through nuclear accumulation of NCK mediated by SOCS7. *Cell* 130, 837–850.
- Kremer, B.E., Haystead, T., Macara, I.G., 2005. Mammalian septins regulate microtubule stability through interaction with the microtubule-binding protein MAP4. *Molecular Biology of the Cell* 16, 4648–4659.
- Kuhlenbaumer, G., Hannibal, M.C., Nelis, E., *et al.*, 2005. Mutations in SEPT9 cause hereditary neuralgic amyotrophy. *Nature Genetics* 37, 1044–1046.
- Kuo, Y.C., Lin, Y.H., Chen, H.I., *et al.*, 2012. SEPT12 mutations cause male infertility with defective sperm annulus. *Human Mutation* 33, 710–719.
- Kwitny, S., Klaus, A.V., Hunnicutt, G.R., 2010. The annulus of the mouse sperm tail is required to establish a membrane diffusion barrier that is engaged during the late steps of spermiogenesis. *Biology of Reproduction* 82, 669–678.
- Leipe, D.D., Wolf, Y.I., Koonin, E.V., Aravind, L., 2002. Classification and evolution of P-loop GTPases and related ATPases. *Journal of Molecular Biology* 317, 41–72.
- Luedeke, C., Frei, S.B., Sbalzarini, I., *et al.*, 2005. Septin-dependent compartmentalization of the endoplasmic reticulum during yeast polarized growth. *Journal of Cell Biology* 169, 897–908.
- Martinez, C., Corral, J., Dent, J.A., *et al.*, 2006. Platelet septin complexes form rings and associate with the microtubular network. *Journal of Thrombosis and Haemostasis* 4, 1388–1395.
- Mavrikis, M., Azou-Gros, Y., Tsai, F.-C., *et al.*, 2014. Septins promote F-actin ring formation by crosslinking actin filaments into curved bundles. *Nature Cell Biology* 16, 322–334.
- McMurray, M.A., Thorner, J., 2009. Septins: Molecular partitioning and the generation of cellular asymmetry. *Cell Division* 4, 18.
- Mizutani, Y., Ito, H., Iwamoto, I., *et al.*, 2013. Possible role of a septin, SEPT1, in spreading in squamous cell carcinoma DJM-1 cells. *Biological Chemistry* 394, 281–290.
- Mostowy, S., Cossart, P., 2011. Septins as key regulators of actin based processes in bacterial infection. *Biological Chemistry* 392, 831–835.
- Mostowy, S., Cossart, P., 2012. Septins: The fourth component of the cytoskeleton. *Nature Reviews Molecular Cell Biology* 13, 183–194.
- Mostowy, S., Janel, S., Forestier, C., *et al.*, 2011. A role for septins in the interaction between the *Listeria monocytogenes* INVASION PROTEIN InlB and the Met receptor. *Biophysical Journal* 100, 1949–1959.
- Nagata, K., Inagaki, M., 2005. Cytoskeletal modification of Rho guanine nucleotide exchange factor activity: Identification of a Rho guanine nucleotide exchange factor as a binding partner for Sept9b, a mammalian septin. *Oncogene* 24, 65–76.
- Nam, S.C., Sun, H., Kang, S.H., *et al.*, 2007. Phosphorylation-dependent septin interaction of Bni5 is important for cytokinesis. *Journal of Microbiology* 45, 227–233.
- Narumiya, S., Tanji, M., Ishizaki, T., 2009. Rho signaling, ROCK and mDia1, in transformation, metastasis and invasion. *Cancer and Metastasis Reviews* 28, 65–76.
- Oh, Y., Bi, E., 2011. Septin structure and function in yeast and beyond. *Trends in Cell Biology* 21, 141–148.
- Pan, F., Malmberg, R.L., Momany, M., 2007. Analysis of septins across kingdoms reveals orthology and new motifs. *BMC Evolutionary Biology* 7, 103.
- Qi, M., Yu, W., Liu, S., *et al.*, 2005. Septin1, a new interaction partner for human serine/threonine kinase aurora-B. *Biochemical and Biophysical Research Communications* 336, 994–1000.
- Russell, S.E., Hall, P.A., 2011. Septin genomics: A road less travelled. *Biological Chemistry* 392, 763–767.
- Sadian, Y., Gatsogiannis, C., Patasi, C., *et al.*, 2013. The role of Cdc42 and Gic1 in the regulation of septin filament formation and dissociation. *Elife* 2, e01085.
- Schmidt, K., Nichols, B.J., 2004a. A barrier to lateral diffusion in the cleavage furrow of dividing mammalian cells. *Current Biology* 14, 1002–1006.
- Schmidt, K., Nichols, B.J., 2004b. Functional interdependence between septin and actin cytoskeleton. *BMC Cell Biology* 5, 43.
- Sellin, M.E., Holmfeldt, P., Stenmark, S., Gullberg, M., 2011a. Microtubules support a disk-like septin arrangement at the plasma membrane of mammalian cells. *Molecular Biology of the Cell* 22, 4588–4601.
- Sellin, M.E., Sandblad, L., Stenmark, S., Gullberg, M., 2011b. Deciphering the rules governing assembly order of mammalian septin complexes. *Molecular Biology of the Cell* 22, 3152–3164.
- Sellin, M.E., Stenmark, S., Gullberg, M., 2014. Cell type-specific expression of SEPT3-homology subgroup members controls the subunit number of heteromeric septin complexes. *Molecular Biology of the Cell* 25, 1594–1607.
- Sharma, S., Quintana, A., Findlay, G.M., *et al.*, 2013. An siRNA screen for NFAT activation identifies septins as coordinators of store-operated Ca entry. *Nature* 499, 238–242.
- Shindo, A., Wallingford, J.B., 2014. PCP and septins compartmentalize cortical actomyosin to direct collective cell movement. *Science* 343, 649–652.
- Sirajuddin, M., Farkasovsky, M., Hauer, F., *et al.*, 2007. Structural insight into filament formation by mammalian septins. *Nature* 449, 311–315.
- Sirajuddin, M., Farkasovsky, M., Zent, E., Wittinghofer, A., 2009. GTP-induced conformational changes in septins and implications for function. *Proceedings of the National Academy of Sciences of the United States of America* 106, 16592–16597.
- Spiliotis, E.T., Gladfelter, A.S., 2012. Spatial guidance of cell asymmetry: Septin GTPases show the way. *Traffic* 13, 195–203.

- Spiliotis, E.T., Hunt, S.J., Hu, Q., Kinoshita, M., Nelson, W.J., 2008. Epithelial polarity requires septin coupling of vesicle transport to polyglutamylated microtubules. *Journal of Cell Biology* 180, 295–303.
- Spiliotis, E.T., Kinoshita, M., Nelson, W.J., 2005. A mitotic septin scaffold required for Mammalian chromosome congression and segregation. *Science* 307, 1781–1785.
- Tada, T., Simonetta, A., Batterton, M., *et al.*, 2007. Role of septin cytoskeleton in spine morphogenesis and dendrite development in neurons. *Current Biology* 17, 1752–1758.
- Tanaka-Takiguchi, Y., Kinoshita, M., Takiguchi, K., 2009. Septin-mediated uniform bracing of phospholipid membranes. *Current Biology* 19, 140–145.
- Tooley, A.J., Gilden, J., Jacobelli, J., *et al.*, 2009. Amoeboid T lymphocytes require the septin cytoskeleton for cortical integrity and persistent motility. *Nature Cell Biology* 11, 17–26.
- Verhey, K.J., Gaertig, J., 2007. The tubulin code. *Cell Cycle* 6, 2152–2160.
- Versele, M., Thorner, J., 2005. Some assembly required: Yeast septins provide the instruction manual. *Trends in Cell Biology* 15, 414–424.
- Weirich, C.S., Erzberger, J.P., Barral, Y., 2008. The septin family of GTPases: Architecture and dynamics. *Nature Reviews Molecular Cell Biology* 9, 478–489.
- Xie, Y., Vessey, J.P., Konecna, A., *et al.*, 2007. The GTP-binding protein Septin 7 is critical for dendrite branching and dendritic-spine morphology. *Current Biology* 17, 1746–1751.
- Yang, Y.M., Fedchyshyn, M.J., Grande, G., *et al.*, 2010. Septins regulate developmental switching from microdomain to nanodomain coupling of Ca(2+) influx to neurotransmitter release at a central synapse. *Neuron* 67, 100–115.
- Zent, E., Wittinghofer, A., 2014. Human septin isoforms and the GDP-GTP cycle. *Biological Chemistry* 395, 169–180.
- Zhang, J., Kong, C., Xie, H., *et al.*, 1999. Phosphatidylinositol polyphosphate binding to the mammalian septin H5 is modulated by GTP. *Current Biology* 9, 1458–1467.
- Zhang, Y., Gao, J., Chung, K.K., *et al.*, 2000. Parkin functions as an E2-dependent ubiquitin- protein ligase and promotes the degradation of the synaptic vesicle-associated protein, CDCrel-1. *Proceedings of the National Academy of Sciences of the United States of America* 97, 13354–13359.
- Zhu, M., Wang, F., Yan, F., *et al.*, 2008. Septin 7 interacts with centromere-associated protein E and is required for its kinetochore localization. *Journal of Biological Chemistry* 283, 18916–18925.

Further Reading

- Caudron, F., Barral, Y., 2009. Septins and the lateral compartmentalization of eukaryotic membranes. *Developmental Cell* 16, 493–506.
- Connolly, D., Abdesselam, I., Verdier-Pinard, P., Montagna, C., 2011. Septin roles in tumorigenesis. *Biological Chemistry* 392, 725–738.
- Dolat, L., Hu, Q., Spiliotis, E.T., 2014. Septin functions in organ system physiology and pathology. *Biological Chemistry* 395, 123–141.
- Hall, P.A., Russell, H.P.A., Pringle, J.R. (Eds.), 2014. *The Septins*. Chichester: John Wiley & Sons, Ltd.
- McMurray, M.A., Thorner, J., 2009. Septins: Molecular partitioning and the generation of cellular asymmetry. *Cell Division* 4, 18.
- Mostowy, S., Cossart, P., 2012. Septins: The fourth component of the cytoskeleton. *Nature Reviews Molecular Cell Biology* 13, 183–194.
- Oh, Y., Bi, E., 2011. Septin structure and function in yeast and beyond. *Trends in Cell Biology* 21, 141–148.
- Russell, S.E., Hall, P.A., 2011. Septin genomics: A road less travelled. *Biological Chemistry* 392, 763–767.
- Weirich, C.S., Erzberger, J.P., Barral, Y., 2008. The septin family of GTPases: Architecture and dynamics. *Nature Reviews Molecular Cell Biology* 9, 478–489.

Bacterial and Archaeal Cytoskeletons

D Ghosal, J Löwe, and LA Amos, MRC Laboratory of Molecular Biology, Cambridge, UK

© 2016 Elsevier Inc. All rights reserved.

Introduction

Although it was long assumed that the contents of bacteria moved only by diffusion, the discovery of the bacterial cytoskeleton in the 1990s totally changed this view. Prokaryote cells (bacteria and archaea, which all lack a nucleus enclosed by membrane) are now known to possess a variety of filament-forming proteins used for cell shaping, polarization, and division, as outlined in this article (Figure 1). Many of these filament-forming proteins are clearly related to eukaryotic cytoskeletal filaments such as F-actin, microtubules, and intermediate filaments. The great variation in bacteria and archaea within each protein family reflects their long evolutionary history compared with the time that eukaryotes have been in existence. These proteins occur in both Gram-positive bacteria (such as *Bacillus subtilis*) and Gram-negative bacteria (such as *Escherichia coli*) and also in the three archaeal phyla, Euryarchaeota, Crenarchaeota, and Thaumarchaeota. From comparisons among bacteria, archaea, and eukarya of a variety of genetically encoded components, including cytoskeletal filaments, it seems most likely that archaea emerged from early Gram-positive bacteria, while eukaryotes seem to come from archaeal host cells encompassing endosymbiotic organelles (notably mitochondria and chloroplasts) derived from Gram-negative bacteria.

Many dynamic members of the nucleotide-dependent filamentous proteins work as linear motors, pushing or pulling objects through the cell and are thus known as cytomotive filaments. A confusing variety of these have been found. Some bacteria have acquired additional genes only distantly related to their own typical sequences, apparently through horizontal

gene transfer (HGT). Thus, *Verrucomicrobia* have eukaryotic tubulin-like BtubA/B heterodimers that assemble into narrow 5-protofilament microtubules (Pilhofer and Jensen, 2013), whose function in these bacteria is currently unknown. Other species contain plasmids and large phages that code for cytomotive filaments resembling archaeal proteins.

Accessory proteins that link the prokaryotic filaments to their cargoes or to the cell membrane, or act as catalysts of the cytomotive activity, are gradually being investigated; some appear to work in similar ways to known eukaryotic proteins but, unlike the conserved filaments, seem to result from convergent evolution rather than being directly related. So far, no motor proteins related to kinesin, myosin or dynein have been found, from which it can be assumed that the linear motor activity of the filaments is evolutionarily older than their use as static cytoskeletal tracks for motor proteins, as in the eukaryotic cytoskeletal network.

Cytokinesis

Almost all bacteria and many archaea divide by means of a constricting ring at mid-cell consisting of protofilaments of tubulin-related FtsZ (Figures 1(a) and 2). However, some groups of archaea divide using ESCRT filaments (Figure 3) instead of an FtsZ-ring. On either side of the membrane constriction, the 'divisome complex' exports components to the periplasm, where new cell wall is assembled to partition the daughter cells.

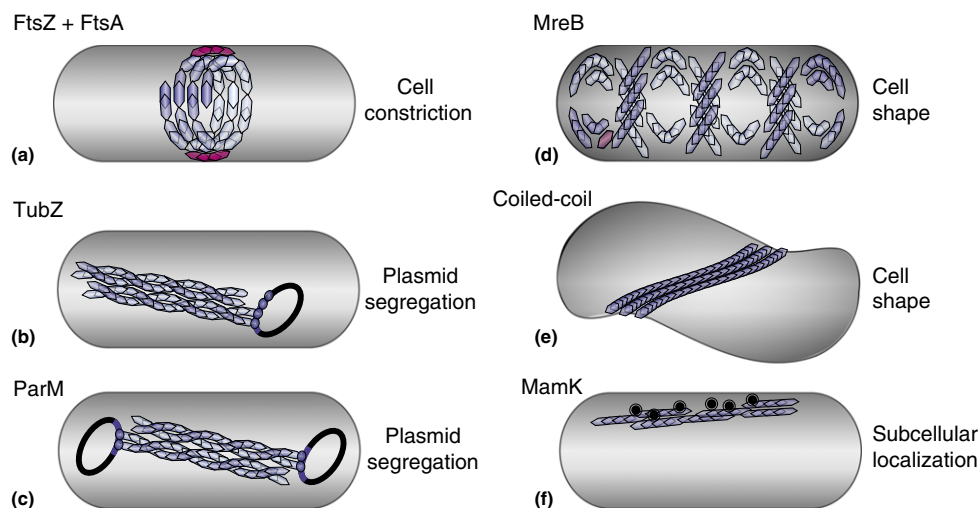


Figure 1 Diverse functions of cytoskeletal elements in bacteria. Bacterial cytoskeletal filaments perform cell division, DNA segregation, cell shape maintenance, and cellular compartment organization. Bacterial tubulin homologs shown: FtsZ performs cytokinesis (a), TubZ executes plasmid segregation (b). Bacterial actin homologs shown: FtsA is needed for cell division (a), ParM executes plasmid segregation (c), MamK organizes magnetosome chains (f), MreB maintains cell shape (d). Intermediate filament-like coiled-coil proteins (e.g., Crescentin and Bactofilins) also regulate cell shape (e).

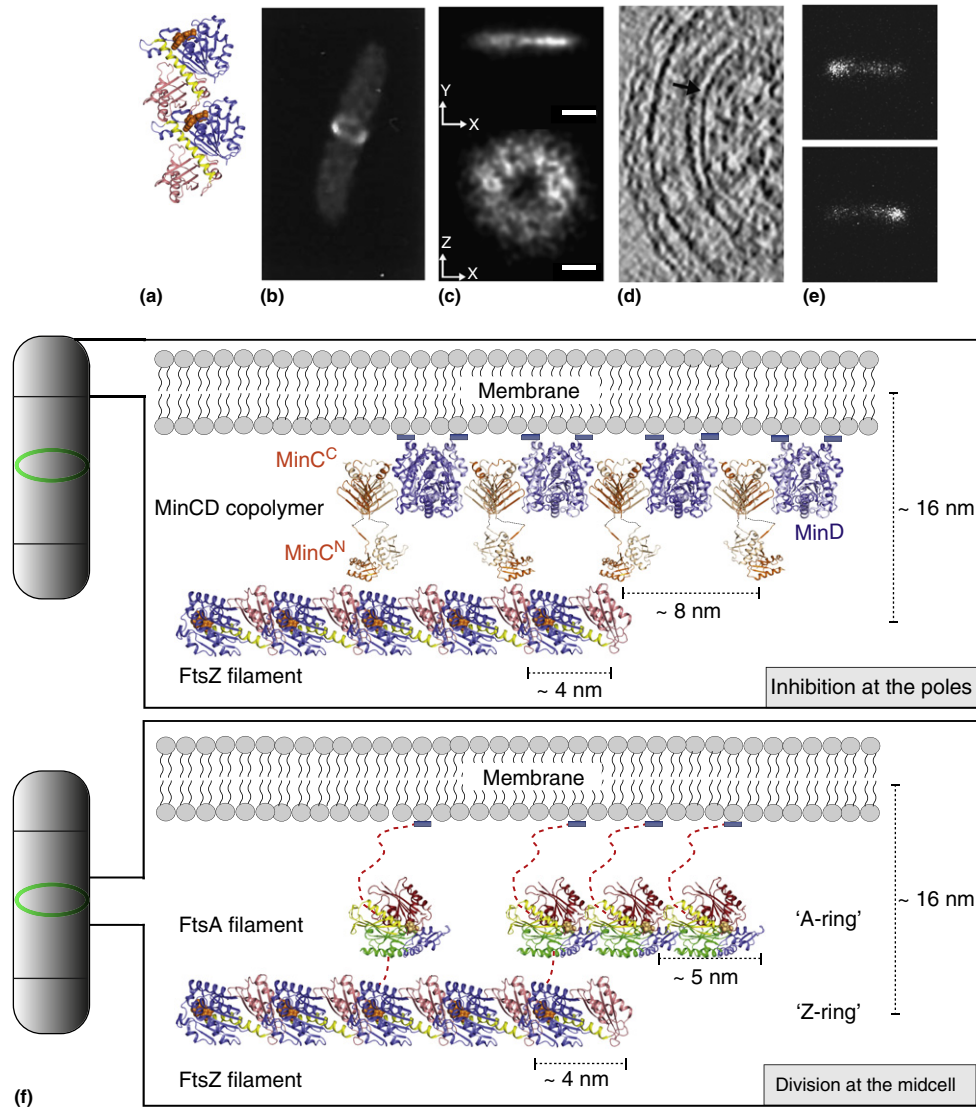


Figure 2 Bacterial filament structures involved in cell division. (a) FtsZ dimer from crystal structure of *Staphylococcus aureus* (PDB ID 3V08), representing two subunits of a straight protofilament (see f). The monomers are colored to distinguish the GTPase domain (blue), the central helix (yellow), and the activation domain (red). The longitudinal interaction between monomers is conserved throughout the tubulin family. Nucleotide (orange) is hydrolyzed by the activation domain of a new subunit adding to the protofilament during assembly. (b) Fluorescence image of an *Escherichia coli* (JM109) cell expressing FtsZ-GFP. FtsZ polymerizes at mid-cell to form the bright Z-ring shown here (Ma *et al.*, 1996). Copyright (1996) National Academy of Sciences, USA. (c) 3D super-resolution imaging of the Z-ring (FtsZ-Dendra2) in a *Caulobacter crescentus* cell. 2D projections of 3D images: side view and cross-sectional view; brighter segments occur where more filaments overlap. Scale bar 200 nm (Biteen *et al.*, 2012). (d) Section through an electron cryotomogram of a *C. crescentus* cell showing Z-ring. (e) In *E. coli*, the Min system oscillates between the cell poles and restricts the FtsZ-ring assembly to the mid-cell. Shown here is GFP-MinC oscillation in fluorescence images of the same cell at different times, 20 s apart. (f) Models showing how FtsZ-based cell division is aided by FtsA at the mid-cell, and how the Min system prevents Z-ring assembly near the cell poles. The FtsZ protofilament shown here is the same structure as (a); FtsA is from *Thermotoga maritima* (PDB ID 4A2B). The MinCD copolymer model was made by combining *Aquifex aeolicus* MinCD co-crystal structure (PDB ID 4V02) and *E. coli* MinD dimer structure (PDB ID 3Q9L). Top: at the cell poles, MinD dimers (blue) bind to the cell membrane, while MinC subunits (yellow) interact with FtsZ filaments via dimerized C-terminal domains and inhibit Z-ring assembly via N-terminal domains. The 8 nm spacing of MinCD copolymers is well-matched to the 4 nm spacing of FtsZ protofilaments (Ghosal *et al.*, 2010). Bottom: at the mid-cell, the MinCD copolymers are unstable due to a dynamic MinE ring (Zheng *et al.*, 2014), but FtsA bound to membrane via the end of its C-terminal extension (red) recruits FtsZ protofilaments which can bundle together to form a persistent Z-ring; the ring constricts when completed.

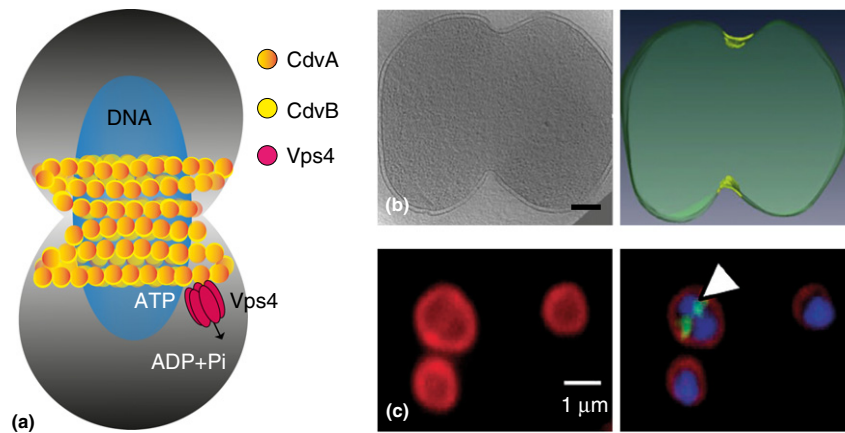


Figure 3 ESCRT-based cell division in archaea. (a) Schematic of ESCRT-mediated cell division in archaea. CdvA polymers (orange) recruit CdvB protein (ESCRT-III: yellow) to the cytoplasmic side of the cell membrane. CdvA and CdvB interaction occurs through the C-terminal winged Helix like (wH) domain of CdvB. These two proteins together generate force, leading to constriction. The mechanism involves sliding of filaments. Upon completion of cytokinesis, the AAA-type ATPase Vps4 (pink) releases the ESCRT-III subunits. DNA is in blue and cell membrane is shown in gray. (b) Left: central section through cryotomogram of a dividing *Sulfolobus acidocaldarius* cell dividing using the ESCRT machinery. Right: segmentations of the same cell. Cell membrane is shown in green and the ESCRT belt is shown in yellow. Scale bar 200 nm (Dobro *et al.*, 2013). Localization of the ESCRT-III in *S. acidocaldarius*, as seen using fluorescence microscopy. Membrane is stained by FM4-64X (red) and ESCRT-III localization is visualized by antibody staining (green). Scale bar 1 μm (Samson *et al.*, 2008).

Structure of the Z-ring and Proposed Constriction Mechanisms

FtsZ protofilaments (Figure 2(f)) observed *in vitro* can be straight filaments, single or bundled, or can be curved to form arcs and mini-rings, especially in the presence of ADP. *In vivo*, parallel FtsZ filaments form the so-called Z-ring around the waist of the cell (Figures 2(b)–2(d)), thus defining the site of division and recruiting other proteins of the divisome complex. They are held in place near the membrane by binding dynamically to FtsA, ZipA, and other membrane-anchored proteins. Filaments may be cross-bridged and stabilized by accessory proteins such as ZapA or SepF. A complete ring of overlapping, though individually dynamic, FtsZ filaments have the ability to exercise a constrictive force. As the diameter of the ring decreases, the divisome complex inserts new peptidoglycan into the ingrowing cell septum.

FtsA, which is similar to actin family member MreB (described below), apart from the replacement of one subdomain by one in a different position, is essential for assembling the Z-ring on the cell membrane; the C-terminus of FtsZ binds to FtsA and the C-terminus of FtsA binds to lipid. The interactions are sufficiently dynamic to allow FtsZ filaments to treadmill (Loose and Mitchison, 2014), at least before there is any constriction of the Z-ring. FtsA has been shown to be capable of assembly into protofilaments itself. During initial establishment of the ring at mid-cell, the mismatch between subunit lengths in the two polymers will inhibit FtsA assembly but pairs of FtsA monomers may exert a tension on FtsZ filaments to favor curvature and thus help to orient the ring. However, some components of the divisome can only bind to FtsA that is not self-associating so, as the curvature increases and the difference in radii compensates more for the subunit mismatch, longer FtsA polymers become possible and more divisome components will be displaced; this may be important for finishing off the septum neatly.

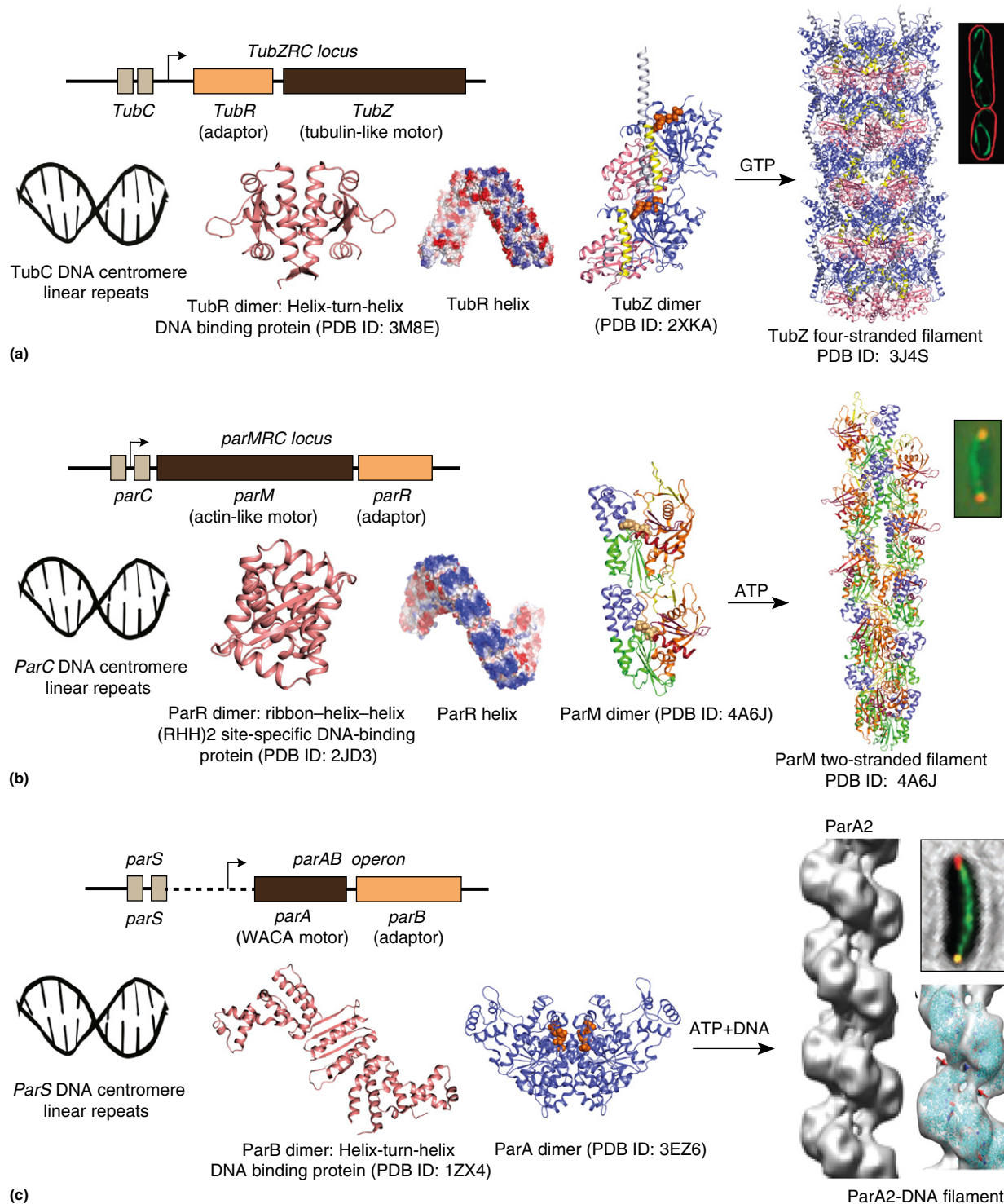
Until very recently, the most popular theory for membrane constriction was an ‘iterative pinching’ mechanism involving nucleotide-dependent bending of membrane-tethered FtsZ protofilaments. However, it is now known that even protofilaments bound to non-hydrolyzable nucleotide cause some curvature on binding to membrane (Osawa and Erickson, 2011). Also the curvature seen for GDP-bound FtsZ (Li *et al.*, 2013) is in the wrong direction for FtsZ, when tethered by its C-terminus to FtsA (Szwedziak *et al.*, 2012), to be able to constrict the membrane from inside. Moreover, it has been seen that the individual filaments, which overlap to form a complete Z-ring, undergo continuous treadmilling rather than cycles of assembly, bending and disassembly. Constriction could be driven in a ring of treadmilling filaments if growth at plus ends were blocked at some stage while minus ends continued to shrink; filament-sliding in the direction that maximized filament-overlap should be more favorable than random sliding. The stress-response protein SulA that binds to the minus end of FtsZ may inhibit ring formation either by preventing treadmilling or by sequestering subunits.

Placement of the Z-ring at Mid-Cell

MinC and MinD are part of the system that prevents assembly of FtsZ filaments in the wrong places (Figures 2(e) and 2(f)) and thus ensures proper positioning of the Z ring for cell division. Close to either side of the proper Z ring, nucleoid occlusion factors on the surfaces of the pair of nucleoids protects them from being bisected by extra rings and, nearer the poles, the Min system prevents the formation of empty mini-cells. MinD, like other proteins in the WACA (Walker A Cytoskeletal ATPases) family (see below), undergoes ATP-dependent dimerization. In the case of MinD, this change also enables the protein to bind to membrane. MinC uses its

C-terminal domain to bind to MinD dimers and its N-terminal domain to inhibit the assembly of FtsZ into a ring. Alternating MinC/D dimers assemble on the membrane surface as bipolar filaments (Ghosal *et al.*, 2010). They bear some structural resemblance to eukaryotic septins. In Gram-negative bacteria, the effector protein MinE (not present in Gram positives)

displaces MinC and releases MinD from the membrane as ADP-bound monomers. Free MinD and MinC can then diffuse to the other pole. A long-lasting ring of MinE at mid-cell prevents binding there. Thus the Min proteins in species such as *E. coli* are observed to oscillate back and forth between the two available binding regions near the two poles. *In vitro*,



MinC/D/E mixtures, supplied with ATP, have been observed to bind in dynamic fashion to a lipid-coated surface, producing a variety of waves characteristic of reversible reaction-diffusion systems (Loose *et al.*, 2011).

ESCRT-Based Cell Division in Archaea

ESCRT-III filaments play a critical role in membrane budding and abscission in the eukaryotic endosome system for transporting membrane proteins between the plasma membrane, the *trans*-Golgi network and the lysosome/vacuole; related structures are found in archaea but not bacteria. In Crenarchaea and Thaumarchaea they are responsible for cell division. Proteins called CdvB, found in *Sulfolobus acidocaldarius* and *Nitrosopumilus maritimus*, have been investigated in particular. A protein called CdvA provides a link between the cell membrane and the CdvB filaments. The filaments have been visualized as cone-shaped spiral assemblies that appear to constrict the membrane between 2 daughter cells (Figure 3). The constriction mechanism is unknown but must involve some sliding of filaments. ESCRT-III filaments self-assemble but can only disassemble with the help of an AAA (ATPase associated with various cellular activities) protein called Vps4 (vacuolar protein sorting 4), which catalyzes dissociation of the ESCRT machinery from endosomal membranes. Vps4-like proteins are also present in the archaea.

DNA Segregation

After, and to some extent during, duplication, copies of chromosomal or plasmid DNA are separated to opposite halves of the mother cell by a variety of co-operating mechanisms. We focus here on those employing cytomotive filaments (Figure 4).

Segregation of Plasmids by Actin-Like ParM

Actin-family proteins that assemble free in the bacterial cytoplasm are involved in plasmid partitioning and include ParM,

AlfA and Alps1-6. Of these, the most well studied is ParM (Figures 1(c) and 4(b)), whose sequence and crystal structure are closer to known archaeal actin-like proteins than to bacterial MreB.

ParM/ParR/*parC* (ParMRC) (a 'type II' partitioning system) is responsible for the equal distribution of R1 plasmids in the daughter cells during cell division. The system has many parallels with TubZRC (a 'type III' partitioning system; Figures 1(b) and 4(a)). This is a clear example of convergent evolution, suggesting that bundles of double helical filaments are very good for distributing objects within a cell. Currently, ParMRC is the best understood cytomotive system in prokaryotes. Two ParM protofilaments form a left-handed double helical filament. Dynamic instability of ParM filaments assembling alone involves rapid filament formation and breakup, approximately 200 times faster than actin, allowing ParM to 'search' for plasmids. ParM filaments attached to plasmids, either captured or initiated *in situ*, grow steadily and more slowly. ParR is a protein that binds to a centromere-like region of the DNA (*parC*). The C-terminus of ParR binds to the barbed end of a ParM filament. Two or more ParM filaments associate side-by-side with mixed polarity to form a mitotic-spindle-like bundle with ParRC bound at each end and ParM filament elongation pushes the plasmids to opposite ends of the cell (cf. eukaryotic anaphase B). A curved polymer of ParR connecting each plasmid to its end of the growing spindle is functionally similar to a eukaryotic formins complex.

Segregation of Plasmids by Tubulin-Like TubZ

The plasmid-encoded TubZ protein (Figure 4(b)) found in *Bacillus* appears to be more closely related to archaeal CetZs than to FtsZ; CetZs are more similar to tubulins than FtsZs are. However, although the tubulin-like longitudinal contacts are preserved in TubZ (Figure 4(a)), a slight kink within each monomer produces a twist in an individual protofilament and paired protofilaments are superficially like the structure of F-actin. A pair of double filaments or bundle of 4 single protofilaments seem to be needed for activity in the TubZ/TubR/

Figure 4 Bacterial filaments involved in DNA segregation. (a) Top: schematic illustration of a type III plasmid partitioning system. This includes centromeric DNA linear repeats, followed by genes for a DNA binding adaptor and a tubulin-like motor. Below: structure of the active components; both TubZ and TubR dimer structures shown are from *Bacillus thuringiensis*. TubR polymer forms a helix on *tubC* DNA (Aylett and Löwe, 2012). TubZ monomer domains are colored according to the same rule as FtsZ in Figure 2(a). Monomers assemble into slightly twisted GTP-dependent protofilaments, which associate as two-stranded and four-stranded filaments. The latter, shown on the right, are thought to be active *in vivo* (Montabana and Agard, 2014). They seem to move plasmids by treadmilling, pulling a TubRC complex that maintains attachment at the shrinking end. At far right: super-resolution imaging of TubZ-GFP filaments in a *Bacillus sphaericus* cell (Ge *et al.*, 2014). (b) Top: schematic illustration of a type II plasmid partitioning system, which includes centromeric DNA linear repeats, a DNA binding adaptor and an actin-like motor. Below: the ParR dimer structure is from plasmid PB171 and the ParM dimer structure from *E. coli* R1 plasmid. ParR polymer forms a helix on *parC* DNA (Møller-Jensen *et al.*, 2007). The four subdomains typical of actin-family proteins are colored red, yellow, green and blue. The longitudinal interaction seen here between ParM monomers is conserved throughout the actin family. ParM forms ATP-dependent two-stranded helical filaments (right) that move ParRC by inserting subunits at the filament tip. Antiparallel filaments slide against each other as they grow, to push plasmids apart (Gayathri *et al.*, 2012). At far right: immunofluorescence micrograph of an *E. coli* cell carrying R1 plasmid. ParM is visualized in green using ParM-specific antibodies. Plasmids, labeled using the lactose operon LacI-lacO marker system, are shown in red. (c) Top: Schematic illustration of a type I plasmid partitioning system, which includes centromeric DNA linear repeats, a DNA binding adaptor and a WACA (Walker A Cytoskeletal ATPases) motor. Below: the ParB dimer structure is from *Enterobacteria phage P1* (PDB ID 1ZX4) and the ParA dimer structure is from *E. coli* P1 plasmid (PDB ID 3EZ6). ParA forms an ATP-dependent nucleoprotein filament, whose structure is shown (Hui *et al.*, 2010). The filament may work by transporting a plasmid attached to the end that shrinks during treadmilling. At top right: super-resolution imaging of ParA-eYFP (green) and CFP-ParB (red) in *C. crescentus* cells, showing DNA segregation (Ptacin *et al.*, 2010).

tubC plasmid partitioning system, in which it seems probable that shrinking TubZ filaments pull the replicated plasmids into the daughter cells (cf. eukaryotic anaphase A). TubR is a protein that forms a short curved polymer and binds to a centromere-like region of the DNA (*tubC*). The C-termini of TubZ interact with the TubR filament.

Segregation of Chromosomal DNA or Plasmids by WACAs

A large class of proteins known as WACAs (Walker A Cytoskeletal ATPases, having sequences that are placed in the 'P loop' superfamily but have a deviant Walker A motif) also use nucleotide-dependent polymerization (mostly on DNA/membrane) to carry out complex tasks. Their roles include chromosome segregation (Soj), plasmid segregation 'type I' (ParA; [Figure 4\(c\)](#)) or controlling the position of the Z-ring (MinD; [Figures 2\(e\)](#) and [2\(f\)](#)). They differ from actin and tubulin in forming bipolar filaments, as do eukaryotic septins. Many have a partner protein that is needed in activating the ATPase cycle. Plasmid ParAs bind non-specifically to nucleoid DNA in a nucleotide-dependent reversible way. The activating partner protein, ParB, with cargo attached (the centromere site *parS*, located on the plasmid) stimulates the ATPase cycle. ParA2 has been shown to form filaments *in vitro*; other ParAs may form filaments only on the nucleoid surface. Replicated plasmids may then be segregated by being transported over the surface by filament depolymerization. Alternatively, ParA dimers may bind all over the surface and dynamic relocation of each plasmid may be explained by a 'burnt bridge' molecular ratchet mechanism: interaction of the ParB/*parS* complex with a ParA dimer bound to the nucleoid would release the ParA monomers and the ParB/*parS* complex would then be free to bind to a neighboring ParA dimer. Even a random walk would tend to separate plasmid copies.

ParB binds tightly only to *parS* on the plasmid but other ParB dimers may associate with the complex and bind weakly to adjacent stretches of DNA. Spreading of the complex in 2 or 3 dimensions is thought to help the DNA to condense into a compact form, with SMC (structural maintenance of chromosomes) proteins being recruited once folding has begun.

Cell Shape Determination

Bacteria mostly have simple shapes, spheres, rods, plates, helices, and crescents; nonspherical cells depend on filaments associated with the membrane to maintain their shapes ([Figures 1\(d\)](#), [1\(e\)](#), and [5](#)). Motile cells have elongated shapes, usually with one or more flagella attached to one pole. Curved or helical shapes seem to be particularly good, both for swimming in liquid and for gliding over surfaces.

Actin-Like MreB in Rod-Shaped Cells

The longitudinal protofilament contacts between MreB subunits ([Figure 5\(a\)](#)) are almost identical to those of

polymerized actin but pairs of protofilaments are antiparallel, rather than parallel, and straight, rather than twisting around each other. MreB is found in all bacteria with elongated shapes and its filaments lie close to the cell membrane, in a similar manner to that of actin in the eukaryotic submembrane cytoskeleton. Besides constricting the cell into an elongated shape, MreB also organizes the elongasome complex, responsible for cell wall synthesis and cell growth, that inserts new peptidoglycan into the cell wall in a helical pattern ([Kawai et al., 2009](#)). Initial reports suggesting that these roles are achieved by a band of continuous MreB filaments stretching along the cell have been disputed; it seems most likely that individual filaments are normally short and dynamic but overlap to form short-lived continuous structures ([Figures 1\(d\)](#) and [5\(b\)–5\(d\)](#)). Continuous MreB structures are relatively more stable in the gliding bacteria (*Myxococcus*) ([Mauriello et al., 2010](#)). Association of MreB with the cell membrane is both direct and via accessory proteins such as RodZ.

Filamentous Proteins that Promote Curvature

A variety of filamentous proteins with long coiled-coil segments have been identified in different bacteria. Most, like lamin filaments in eukaryotic nuclei, seem to associate closely with membrane, in supporting and shaping roles. Some, like Crescentin in *Caulobacter crescentus* and AglZ in *Myxococcus xanthus*, extend along the whole cell length; others are found only at cell poles. Filaments of Crescentin, the first intermediate filament-like protein to be identified, produce the typical banana shape of *Caulobacter crescentus* by forming a membrane-stiffening band of polymers on one side of the cell ([Figures 5\(e\)](#) and [5\(f\)](#)); the series of coiled-coil domains resembles that of eukaryotic nuclear lamins and IF proteins such as Vimentin but no Crescentin atomic structure is currently available to confirm the similarity. Bactofilins are also predicted to form IF-like structures and are involved in the regulation of murein synthesis in *C. crescentus*. Normally, except when they are overexpressed ([Figure 5\(g\)](#)), they are restricted to the more curved regions of plasma membrane, at the cell poles. The DivIVA peripheral membrane protein, found in Gram-positive bacteria, is similar to the bar-domain proteins of eukaryotes and also locates to curved regions of membrane ([Figures 5\(h\)](#) and [5\(i\)](#)). DivIVA plays a role in locating MinCD to polar regions as well as promoting membrane curvature.

Many other filaments have been noticed in the cytoplasm of individual species and their functions are usually unknown. However, the presence of the enzyme CTP-synthase (CtpS), which forms filaments in eukaryotes as well as bacteria, affects the curvature of *C. crescentus* ([Figure 5\(f\)](#)).

Organization of Intracellular Membrane Compartments

Although there are no separate membranous organelles in bacteria, large intracellular compartments are often found, especially in more specialized bacteria. Cyanobacteria ('blue-green algae'), for example, have their photosynthetic

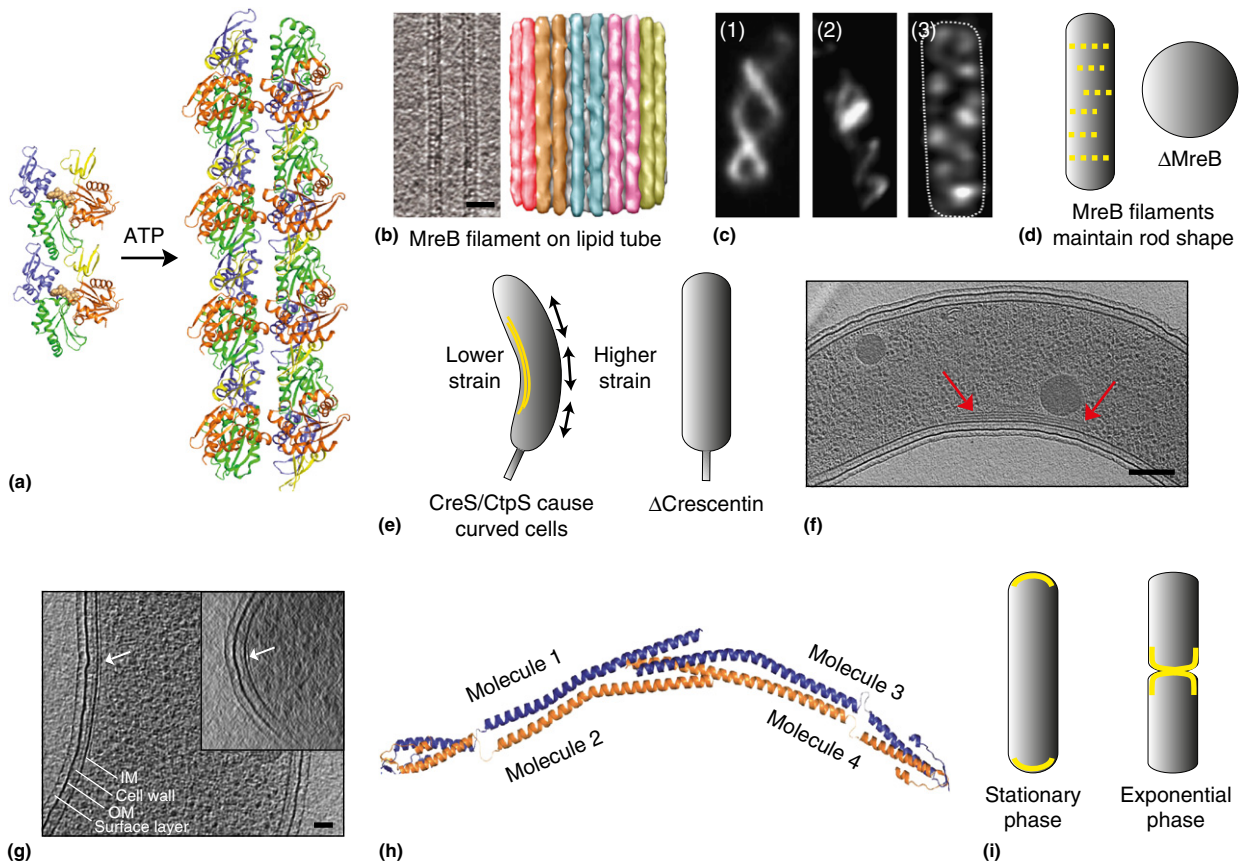


Figure 5 Cell shape determination by bacterial filamentous proteins. (a) Bacterial actin homologue MreB, with the four subdomains of a monomer colored red, yellow, green, and blue. The longitudinal interactions between monomers are as conserved throughout the actin family. MreB forms nucleotide-dependent flat antiparallel double filaments on membrane and maintains rod shape. Shown here are *Caulobacter crescentus* MreB dimer (left), and (AMPPNP) antiparallel double filaments (right) (PDB ID 4CZJ). (b) Central section through cryotomogram of a lipid tube decorated with *Thermotoga maritima* MreB and a corresponding subtomogram average, showing flat double filaments on the membrane (van den Ent *et al.*, 2014). (c) Organization of MreB filaments in *E. coli* cells. Fluorescence images of YFP-MreB showing (1) helical filaments and (2) band/ring-like structures (Vats and Rothfield, 2007). (3) GFP-MreB as small patches in *Bacillus subtilis* (Dominguez-Escobar *et al.*, 2011). (1,2) Copyright (2007) National Academy of Sciences, USA. (d) Summary: MreB filaments maintain a rod shape in many bacteria and deletion of the gene causes a distorted spherical cell shape. (e) Coiled-coil protein Crescentin (CreS) polymerizes along the length of the *C. crescentus* cell, causing higher strain on one side. This results in a bent cell shape. Cells lacking CreS grow with a straight rod shape. A combination of both Crescentin and CTP-synthase (CtpS) regulate cell shape in *Caulobacter*. (f) Section through tomographic reconstruction of *C. crescentus* showing CtpS filaments. Scale bar 100 nm (Ingerson-Mahar *et al.*, 2010). (g) Cryo-tomographic section of a *C. crescentus* cell overproducing Bactofilin, showing the extra filaments. Scale bar 50 nm (Kühn *et al.*, 2010). (h) Crystal structure of the *B. subtilis* peripheral membrane protein DivIVA (PDB ID 2WUJ). A composite model of the full-length protein has been made by combining N-terminal and C-terminal domain crystal structures (Oliva *et al.*, 2010). (i) Schematic of DivIVA localization in cells. During stationary phase DivIVA remains at the poles, while during exponential growth DivIVA localizes on the ingrowing septation (Bach *et al.*, 2014).

machinery embedded in extensive inward folds of the cell membrane. In other phyla, invaginations may be small and transient. The roles of bacterial dynamin-like proteins (BDLPs) are still unclear but it seems likely that they include the creation and maintenance of such membrane compartments. *In vitro*, BDLPs form polymers on lipid tubes (Figures 6(a) and 6(b)) like their eukaryotic counterparts.

Magnetotactic bacteria such as *Magnetospirillum*, have bag-like cell-membrane invaginations called magnetosomes that contain iron oxide or iron sulfide crystals (Figure 1(f)). MamK, a specialized actin-like protein, forms straight double filaments that organize a linear arrangement of magnetosomes to act as a magnetic field sensor (Figure 6(c)).

Conclusion

The cytoskeletons of prokaryotes are highly dynamic systems of proteins, of which the central components are cytomotive filaments, present in some form throughout all kingdoms of life. It is now clear that the cytoskeleton of the last common ancestor of all extant cells, including the 'simplest' bacteria, was already highly sophisticated. The properties of these filaments were so well adapted for intracellular movement that they have been conserved throughout subsequent evolution but, since the recruitment of different interacting proteins modulates the behavior of the conserved filaments, it is the explosion in the variety of regulatory factors and motor

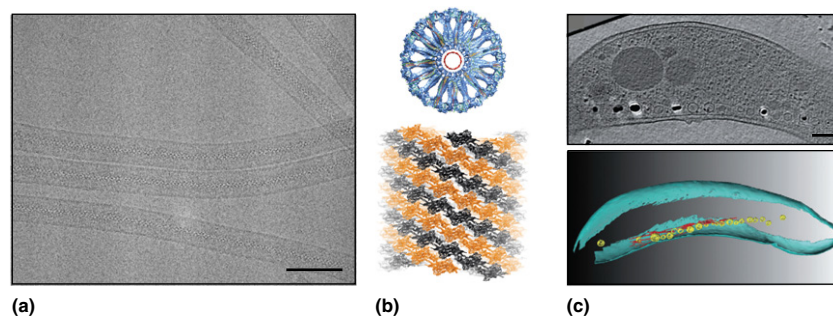


Figure 6 Filaments involved in organization of intracellular membrane compartments. (a) Bacterial dynamin-like protein (BDLP) decorating a lipid tube, viewed in thin ice by electron cryo-microscopy. Scale bar 100 nm (Low *et al.*, 2009). (b) Top: 3D reconstruction of the BDLP/lipid tube, shown in cross-section. The arrangement of spokes that connect the globular GTPase domains to the membrane is simpler than in equivalent eukaryotic complexes. Bottom: model showing how BDLP assembles into left-handed helical filaments on lipid. (c) Electron cryo-tomography (ECT) of a magnetotactic bacterium *Magnetospirillum magneticum* (AMB-1) showing how magnetosome-chain (yellow) assembly is supported by MamK filaments (red). Right: reconstructed sections (top) and segmentation analysis (bottom). Scale bar 200 nm (Draper *et al.*, 2011).

proteins that spawned the more extensive and more immediately obvious cytoskeletons of eukaryotes.

See also: Cytoskeleton and Motors: Cytoskeletal Components: Septins: Cytoskeletal Filaments with Structural and Regulatory Functions

References

- Aylett, C.H.S., Löwe, J., 2012. Superstructure of the centromeric complex of TubZRC plasmid partitioning systems. *Proceedings of the National Academy of Sciences of the United States of America* 109, 16522–16527.
- Bach, J.N., Albrecht, N., Bramkamp, M., 2014. Imaging DivIVA dynamics using photo-convertible and activatable fluorophores in *Bacillus subtilis*. *Frontiers in Microbiology* 5, 59.
- Biteen, J.S., Goley, E.D., Shapiro, L., Moerner, W.E., 2012. Three-dimensional super-resolution imaging of the midplane protein FtsZ in live *Caulobacter crescentus* cells using astigmatism. *Chemphyschem* 13, 1007–1012.
- Dobro, M.J., Samson, R.Y., Yu, Z., *et al.*, 2013. Electron cryotomography of ESCRT assemblies and dividing *Sulfolobus* cells suggests that spiraling filaments are involved in membrane scission. *Molecular Biology of the Cell* 24, 2319–2327.
- Dominguez-Escobar, J., Chastanet, A., Crevenna, A.H., *et al.*, 2011. Processive movement of MreB-associated cell wall biosynthetic complexes in bacteria. *Science* 333, 225–228.
- Draper, O., Byrne, M.E., Li, Z., *et al.*, 2011. MamK, a bacterial actin, forms dynamic filaments *in vivo* that are regulated by the acidic proteins MamJ and LimJ. *Molecular Microbiology* 82, 342–354.
- van den Ent, F., Izore, T., Bharat, T.A., Johnson, C.M., Löwe, J., 2014. Bacterial actin MreB forms antiparallel double filaments. *Elife (Cambridge)* 3, e02634.
- Gayathri, P., Møller-Jensen, J., van den Ent, F., Namba, K., Lowe, J., 2012. A bipolar spindle of antiparallel ParM filaments drives bacterial plasmid segregation. *Science* 338, 1334–1337.
- Ge, Y., Hu, X., Zhao, N., *et al.*, 2014. A new tubRZ operon involved in the maintenance of the *Bacillus sphaericus* mosquitoicidal plasmid pBspH. *Microbiology* 160, 1112–1124.
- Ghosal, D., Trambalolo, D., Amos, L.A., Löwe, J., 2010. MinCD cell division proteins form a new class of alternating co-polymeric cytomotive filaments. *Proceedings of the National Academy of Sciences of the United States of America* 107 (18), 8071–8078.
- Hui, M.P., Galkin, V.E., Yu, X., *et al.*, 2010. ParA2, a *Vibrio cholerae* chromosome partitioning protein, forms left-handed helical filaments on DNA. *Proceedings of the National Academy of Sciences of the United States of America* 107, 4590–4595.
- Ingerson-Mahar, M., Briegel, A., Werner, J.N., Jensen, G.J., Gitai, Z., 2010. The metabolic enzyme CTP synthase forms cytoskeletal filaments. *Nature Cell Biology* 12, 739–746.
- Kawai, Y., Asai, K., Errington, J., 2009. Partial functional redundancy of MreB isoforms, MreB, Mbl and MreBH, in cell morphogenesis of *Bacillus subtilis*. *Molecular Microbiology* 73, 719–731.
- Kühn, J., Briegel, A., Mörschel, E., *et al.*, 2010. Bactofilins, a ubiquitous class of cytoskeletal proteins mediating polar localization of a cell wall synthase in *Caulobacter crescentus*. *EMBO Journal* 29, 327–339.
- Li, Y., Hsin, J., Zhao, L., *et al.*, 2013. FtsZ protofilaments use a hinge-opening mechanism for constrictive force generation. *Science* 341, 392–395.
- Loose, M., Fischer-Friedrich, E., Herold, C., Kruse, K., Schwill, P., 2011. Min protein patterns emerge from rapid rebinding and membrane interaction of MinE. *Nature Structural & Molecular Biology* 18, 577–583.
- Loose, M., Mitchison, T.J., 2014. The bacterial cell division proteins FtsA and FtsZ self-organize into dynamic cytoskeletal patterns. *Nature Cell Biology* 16, 38–46.
- Low, H.H., Sachse, C., Amos, L.A., Löwe, J., 2009. Structure of a bacterial dynamin-like protein lipid tube provides a mechanism for assembly and membrane curving. *Cell* 139, 1342–1352.
- Ma, X., Ehrhardt, D.W., Margolin, W., 1996. Colocalization of cell division proteins FtsZ and FtsA to cytoskeletal structures in living *Escherichia coli* cells by using green fluorescent protein. *Proceedings of the National Academy of Sciences of the United States of America* 93, 12998–13003.
- Mauriello, E.M., Mouhamar, F., Nan, B., *et al.*, 2010. Bacterial motility complexes require the actin-like protein, MreB and the Ras homologue, MglA. *EMBO Journal* 29, 315–326.
- Møller-Jensen, J., Ringgaard, S., Mercogliano, C.P., Gerdes, K., Löwe, J., 2007. Structural analysis of the ParR/parC plasmid partition complex. *EMBO Journal* 26, 4413–4422.
- Montabana, E.A., Agard, D.A., 2014. Bacterial tubulin TubZ-Bt transitions between a two-stranded intermediate and a four-stranded filament upon GTP hydrolysis. *Proceedings of the National Academy of Sciences of the United States of America* 111, 3407–3412.
- Oliva, M.A., Halbedel, S., Freund, S.M., *et al.*, 2010. Features critical for membrane binding revealed by DivIVA crystal structure. *EMBO Journal* 29, 1988–2001.
- Osawa, M., Erickson, H.P., 2011. Inside-out Z rings — constriction with and without GTP hydrolysis. *Molecular Microbiology* 81, 571–579.
- Pilhofer, M., Jensen, G.J., 2013. The bacterial cytoskeleton: More than twisted filaments. *Current Opinion in Cell Biology* 25, 125–133.
- Placín, J.L., Lee, S.F., Garner, E.C., *et al.*, 2010. A spindle-like apparatus guides bacterial chromosome segregation. *Nature Cell Biology* 12, 791–798.
- Samson, R.Y., Obita, T., Freund, S.M., Williams, R.L., Bell, S.D., 2008. A role for the ESCRT system in cell division in archaea. *Science* 322, 1710–1713.
- Szwedziak, P., Wang, Q., Freund, S.M., Löwe, J., 2012. FtsA forms actin-like protofilaments. *EMBO Journal* 31, 2249–2260.
- Vats, P., Rothfield, L., 2007. Duplication and segregation of the actin (MreB) cytoskeleton during the prokaryotic cell cycle. *Proceedings of the National Academy of Sciences of the United States of America* 104, 17795–17800.

Zheng, M., Chiang, Y.L., Lee, H.L., *et al.*, 2014. Self-assembly of MinE on the membrane underlies formation of the MinE ring to sustain function of the *Escherichia coli* Min system. *Journal of Biological Chemistry* 289, 21252–21266.

Further Reading

Adams, D.W., Errington, J., 2009. Bacterial cell division: Assembly, maintenance and disassembly of the Z ring. *Nature Reviews Microbiology* 7, 642–653.

Aylett, C.H.S., Löwe, J., Amos, L.A., 2011. New insights into the mechanisms of cytomotive actin and tubulin filaments. *International Review of Cell and Molecular Biology* 292, 1–71.

Bagchi, S., Tomerius, H., Belova, L.M., Ausmees, N., 2008. Intermediate filament-like proteins in bacteria and a cytoskeletal function in *Streptomyces*. *Molecular Microbiology* 70, 1037–1050.

Bernander, R., Lindas, A.E., Ettema, T.J., 2011. An archaeal origin for the actin cytoskeleton: Implications for eukaryogenesis. *Communicative & Integrative Biology* 4, 664–667.

den Blaauwen, T., 2013. Prokaryotic cell division: Flexible and diverse. *Current Opinion in Microbiology* 16, 738–744.

Bramkamp, M., 2012. Structure and function of bacterial dynamin-like proteins. *Biological Chemistry* 393, 1203–1214.

Cabeen, M.T., Jacobs-Wagner, C., 2010. The bacterial cytoskeleton. *Annual Review of Genetics* 44, 365–392.

Campos, M., Jacobs-Wagner, C., 2013. Cellular organization of the transfer of genetic information. *Current Opinion in Microbiology* 16, 171–176.

Carballido-Lopez, R., 2006. The bacterial actin-like cytoskeleton. *Microbiology and Molecular Biology Reviews* 70, 888–909.

Celler, K., Koning, R.I., Koster, A.J., van Wezel, G.P., 2013. Multidimensional view of the bacterial cytoskeleton. *Journal of Bacteriology* 195, 1627–1636.

Chen, S., McDowell, A., Dobro, M.J., *et al.*, 2010. Electron cryotomography of bacterial cells. *Journal of Visualized Experiments* 6, 1943–1948.

Dajkovic, A., Lutkenhaus, J., 2006. Z ring as executor of bacterial cell division. *Journal of Molecular Microbiology and Biotechnology* 11, 140–151.

Dobro, M.J., Samson, R.Y., Yu, Z., *et al.*, 2013. Electron cryotomography of ESCRT assemblies and dividing *Sulfolobus* cells suggests that spiraling filaments are involved in membrane scission. *Molecular Biology of the Cell* 24, 2319–2327.

Dye, N.A., Shapiro, L., 2007. The push and pull of the bacterial cytoskeleton. *Trends in Cell Biology* 17, 239–245.

Erickson, H.P., Anderson, D.E., Osawa, M., 2010. FtsZ in bacterial cytokinesis: Cytoskeleton and force generator all in one. *Microbiology and Molecular Biology Reviews* 74, 504–528.

Ettema, T.J., Bernander, R., 2009. Cell division and the ESCRT complex: A surprise from the archaea. *Communicative & Integrative Biology* 2, 86–88.

Gerdes, K., Howard, M., Szardenings, F., 2010. Pushing and pulling in prokaryotic DNA segregation. *Cell* 141, 927–942.

Ghazi-Tabatabai, S., Obita, T., Pobbati, A.V., *et al.*, 2009. Evolution and assembly of ESCRTs. *Biochemical Society Transactions* 37, 151–155.

Graumann, P.L., 2009. Dynamics of bacterial cytoskeletal elements. *Cell Motility and the Cytoskeleton* 66, 909–914.

Haeussler, D.P., Margolin, W., 2012. Bacteriophage tubulins: carrying their own cytoskeleton key. *Current Biology* 22, R639–R641.

Harry, E., Monahan, L., Thompson, L., 2006. Bacterial cell division: The mechanism and its precision. *International Review of Cytology* 253, 27–94.

Ingerson-Mahar, M., Gitai, Z., 2012. A growing family: The expanding universe of the bacterial cytoskeleton. *FEMS Microbiology Reviews* 36, 256–266.

Lin, L., Thanbichler, M., 2013. Nucleotide-independent cytoskeletal scaffolds in bacteria. *Cytoskeleton (Hoboken)* 70, 409–423.

Low, H.H., Löwe, J., 2010. Dynamin architecture — from monomer to polymer. *Current Opinion in Structural Biology* 20, 791–798.

Löwe, J., Amos, L.A., 2009. Evolution of cytomotive filaments: The cytoskeleton from prokaryotes to eukaryotes. *International Journal of Biochemistry & Cell Biology* 41, 323–329.

Lutkenhaus, J., Pichoff, S., Du, S., 2012. Bacterial cytokinesis: From Z ring to divisome. *Cytoskeleton (Hoboken)* 69, 778–790.

Margolin, W., 2009. Sculpting the bacterial cell. *Current Biology* 19, R812–R822.

Mauriello, E.M., Mouhamar, F., Nan, B., *et al.*, 2010. Bacterial motility complexes require the actin-like protein, MreB and the Ras homologue, MglA. *EMBO Journal* 29, 315–326.

Michie, K.A., Löwe, J., 2006. Dynamic filaments of the bacterial cytoskeleton. *Annual Review of Biochemistry* 75, 467–492.

Nan, B., McBride, M.J., Chen, J., Zusman, D.R., Oster, G., 2014. Bacteria that glide with helical tracks. *Current Biology* 24, R169–R173.

Ozyamak, E., Kollman, J.M., Komeili, A., 2013. Bacterial actins and their diversity. *Biochemistry* 52, 6928–6939.

Pogliano, J., 2008. The bacterial cytoskeleton. *Current Opinion in Cell Biology* 20, 19–27.

Ptacin, J.L., Lee, S.F., Garner, E.C., *et al.*, 2010. A spindle-like apparatus guides bacterial chromosome segregation. *Nature Cell Biology* 12, 791–798.

Rico, A.I., Krupka, M., Vicente, M., 2013. In the beginning, *Escherichia coli* assembled the proto-ring: An initial phase of division. *Journal of Biological Chemistry* 288, 20830–20836.

Salje, J., Gayathri, P., Löwe, J., 2010. The ParMRC system: molecular mechanisms of plasmid segregation by actin-like filaments. *Nature Reviews Microbiology* 8, 683–692.

Samson, R.Y., Bell, S.D., 2009. Ancient ESCRTs and the evolution of binary fission. *Trends in Microbiology* 17, 507–513.

Schumacher, M.A., 2012. Bacterial plasmid partition machinery: A minimalist approach to survival. *Current Opinion in Structural Biology* 22, 72–79.

Shih, Y.L., Rothfield, L., 2006. The bacterial cytoskeleton. *Microbiology and Molecular Biology Reviews* 70, 729–754.

Szardenings, F., Guymer, D., Gerdes, K., 2011. ParA ATPases can move and position DNA and subcellular structures. *Current Opinion in Microbiology* 14, 712–718.

Szwedziak, P., Löwe, J., 2013. Do the divisome and elongasome share a common evolutionary past? *Current Opinion in Microbiology* 16, 745–751.

van Teeffelen, S., Wang, S., Furchtgott, L., *et al.*, 2011. The bacterial actin MreB rotates, and rotation depends on cell-wall assembly. *Proceedings of the National Academy of Sciences of the United States of America* 108, 15822–15827.

Typas, A., Banzhaf, M., Gross, C.A., Vollmer, W., 2012. From the regulation of peptidoglycan synthesis to bacterial growth and morphology. *Nature Reviews Microbiology* 10, 123–136.

Vats, P., Yu, J., Rothfield, L., 2009. The dynamic nature of the bacterial cytoskeleton. *Cellular and Molecular Life Sciences* 66, 3353–3362.

Wang, S., Shaevitz, J.W., 2013. The mechanics of shape in prokaryotes. *Frontiers in Bioscience (Scholar Edition)* 5, 564–574.

Wilson, M.Z., Gitai, Z., 2013. Beyond the cytoskeleton: Mesoscale assemblies and their function in spatial organization. *Current Opinion in Microbiology* 16, 177–183.

Relevant Websites

<http://labs.cellbio.duke.edu/Erickson/>
Erickson Lab.

<http://www.ibiology.org/ibioseminars/microbiology/lucy-shapiro-part-1.html>
iBiology Shapiro Seminar.

<http://www2.mrc-lmb.cam.ac.uk/groups/jyl/slideshows/encyclopedia2014/index.html>
Löwe Lab Slideshow.

<http://www.nature.com/milestones/milecyto/full/milecyto21.html>
Nature Milestone Series.

http://labs.biology.ucsd.edu/jpogliano/jpogliano/Bacterial_Cytoskeletal_Proteins.html
Pogliano Lab.

Myosins

SM Heissler and JR Sellers, National Institutes of Health, Bethesda, MD, USA

Published by Elsevier Inc.

Glossary

Coiled-coil A structural motif whereby α -helices from two different proteins associate to form a larger helix with a seven amino acid repeat.

Duty ratio The fraction of time during the ATPase cycle the motor spends attached on F-actin.

IQ motif An structural motif found in myosins and other proteins that binds calmodulin-like EF-hand family

member proteins. The consensus sequence is IQXXRGXXR.

Myosin A family of motor proteins which use the energy derived from ATP hydrolysis to interact with and move relative to actin.

Processivity The ability of a motor protein to take more than one step along F-actin per diffusional encounter.

Introduction

Myosins constitute a large multigene family of actin-based molecular motors, which are essential to eukaryotic homeostasis across the phylogenetic spectrum. Myosins are involved in growth and tissue formation, metabolism, reproduction, communication, reshaping, and movement of all 100 trillion cells in the human body. Further, myosins power the rapid entry of microbial pathogens such as parasites, viruses, and bacteria in eukaryotic host cells. In plants, essential processes like development, size control and cytoplasmic streaming are promoted by myosin motors.

These examples highlight the multifaceted functional activity of the myosin superfamily which is divided in 35 phylogenetic classes, including enzymatically inactive pseudoenzymes (Odronitz and Kollmar, 2007). In humans, the 39 myosin genes are categorized into 13 classes, as shown in Figure 1. Per definition, class-2 myosins which encompass the myosins involved in muscle contraction are referred to as 'conventional,' all other myosins as 'unconventional.' Despite myosin's functional diversity, all enzymatically active myosins couple the hydrolysis of ATP to force generation and unidirectional movement along filamentous actin (F-actin). This interaction, and consequently the mechanoenzymatic properties of the motor, is myosin specific and strictly fine-tuned by a plethora of up- and downstream effectors.

Myosin Structure and Organization

Myosins are heteromers composed of one or two myosin heavy chains and a variable number of calmodulin-like light chains. The myosin heavy chain is characterized by a prototypic motor domain, a neck region and a tail domain (Figures 2(a)–2(c)). Generally, the motor domain constitutes the N-terminus of the myosin heavy chain, but some myosins have N-terminal extensions of variable lengths preceding the motor domain (Odronitz and Kollmar, 2007).

The motor domain is highly allosteric and the catalytically active part of the molecule. It comprises of an ATP binding site and an F-actin binding region and links a repeated cycle of ATP hydrolysis to mechanical force generation on F-actin

(Figure 3), as outlined in greater detail below. The myosin motor domain core is highly conserved across phyla. Structural, kinetic, and functional diversifications are conferred by variable surface loops and the N-terminus. Surface loops are regulatory hotspots due to extensive phosphorylation, alternate splicing, and their implication in the establishment of the actomyosin interface (Lorenz and Holmes, 2010; Murphy and Spudich, 2000).

The α -helical neck domain comprises a variable number of light chain binding IQ motifs and can be lengthened by a monomeric stable single α -helix (Baboolal *et al.*, 2009). All characterized myosin light chains are calmodulin-like and comprise one to four EF-hand domains. As a central dogma, conventional myosins bind a light chain set composed of essential light chain (ELC) and regulatory light chain (RLC), whereas unconventional myosins bind the calcium-binding protein calmodulin. However, recent reports describe unconventional myosins as off-targets of the ELC and RLC (Guzik-Lendrum *et al.*, 2013; Lu *et al.*, 2014). Neck:light chain interactions increase the stiffness of the neck region, which allows the neck to function as a rigid lever arm during the force generating power stroke.

Myosin tail architecture is versatile due to a vast complement (>50) of domains and motifs such as coiled-coil domains, myosin tail homology 4/band 4.1, ezrin, radixin, and moesin (MyTH-FERM), postsynaptic density protein (PSD95), *Drosophila* disc large tumor suppressor (Dlg1), and zonula occludens-1 protein (zo-1) (PDZ) domain, src homology (SH) and pleckstrin homology (PH) domains (Figure 2(c)). Tail-domain variation is a crucial contributor to myosin-specific structural, regulatory, and functional properties as well as subcellular localization. For example, dimerization of myosins from classes-2, -5, and -18 is mediated by coiled-coil domains in the myosin tail. Solely the heavy chain of conventional class-2 myosins can further oligomerize and promote filament assembly. Additionally, tail domains establish the interaction with phospholipids and the complex number of myosin binding proteins (Sellers, 2000).

Bigger, faster, different – extreme diversification of the body plan of the myosin heavy chain is a prerequisite of the enormous functional activity of the myosome. The molecular weight of the heavy chain can vary from 94 to 495 kDa,

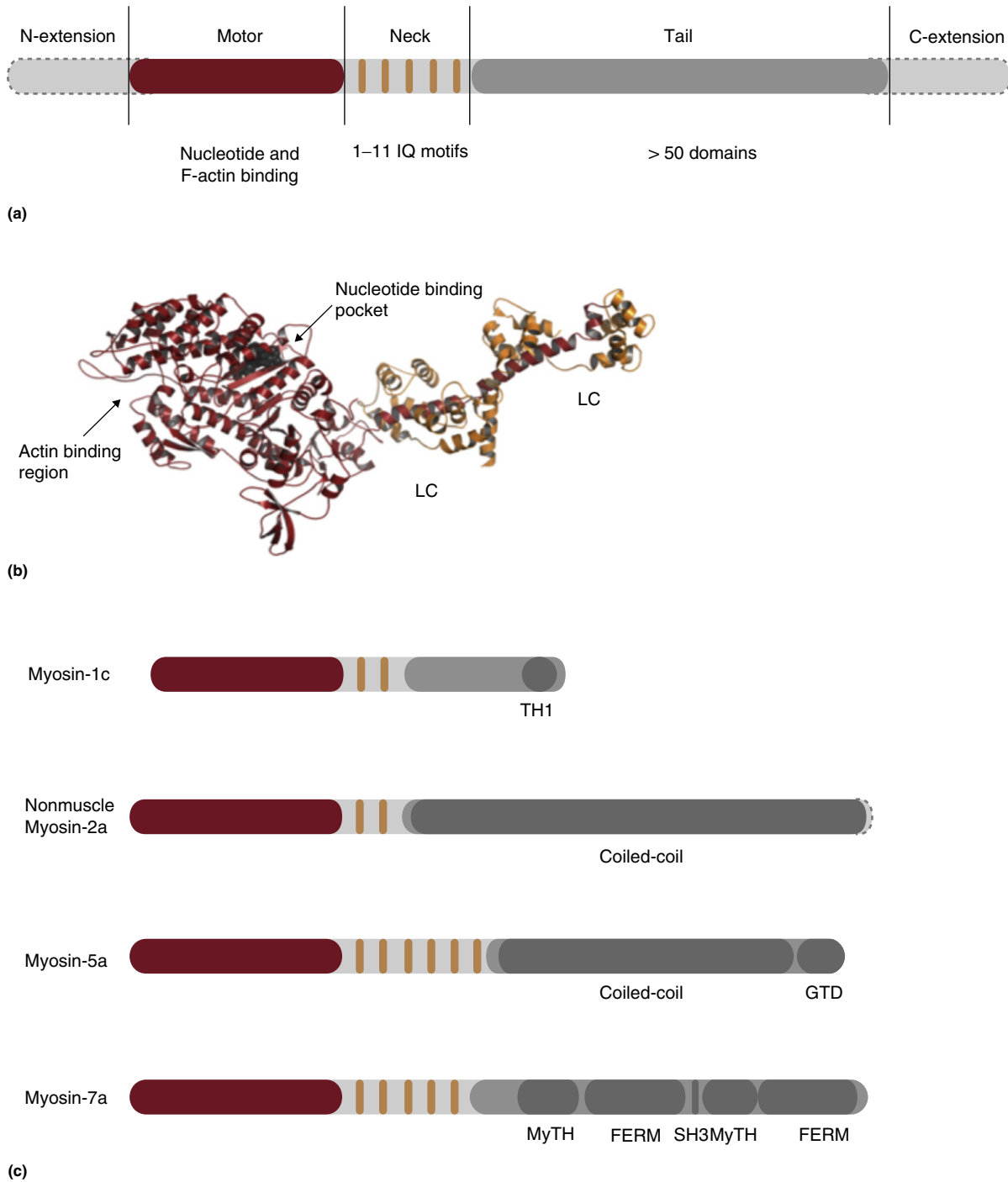


Figure 2 Structural domains found in typical myosins. (a) The myosin heavy chain has a motor domain (red) which harbors the nucleotide binding site and the F-actin binding region, followed by an IQ motif (orange) containing neck domain (gray) and a tail domain (dark gray). N- and C-terminal extensions (dashed line) for the myosin heavy chain are reported. (b) Three-dimensional model of the globular myosin motor domain (red). The light chains (orange, LC) bind to the IQ motifs in the α -helical neck region, distal from the motor domain. The nucleotide is represented in dark gray spheres. The F-actin binding region and the nucleotide binding pocket are indicated by an arrow. (c) Schematic diagrams of heavy chains of representative myosins from classes-1, -2- 5 and -7. Domains found in the myosin tails such as coiled-coiled and the myosin-5 globular, cargo binding tail domain (GTD) are colored in dark gray. The coiled-coil in the tail domain of nonmuscle myosin-2a and myosin-5a promotes the dimerization of the two heavy chains. Schemes are not drawn to scale.

faster than skeletal muscle myosin and more than 1000-fold faster than some myosins involved in the maintenance of the cytoplasmic architecture in animal cells (Higashi-Fujime *et al.*,

1995). In contrast, myosin-18 homologs and myosin-20 from the fruitfly *Drosophila* do not observably hydrolyze ATP and are immotile *in vitro* (Guzik-Lendrum *et al.*, 2013; Cao *et al.*,

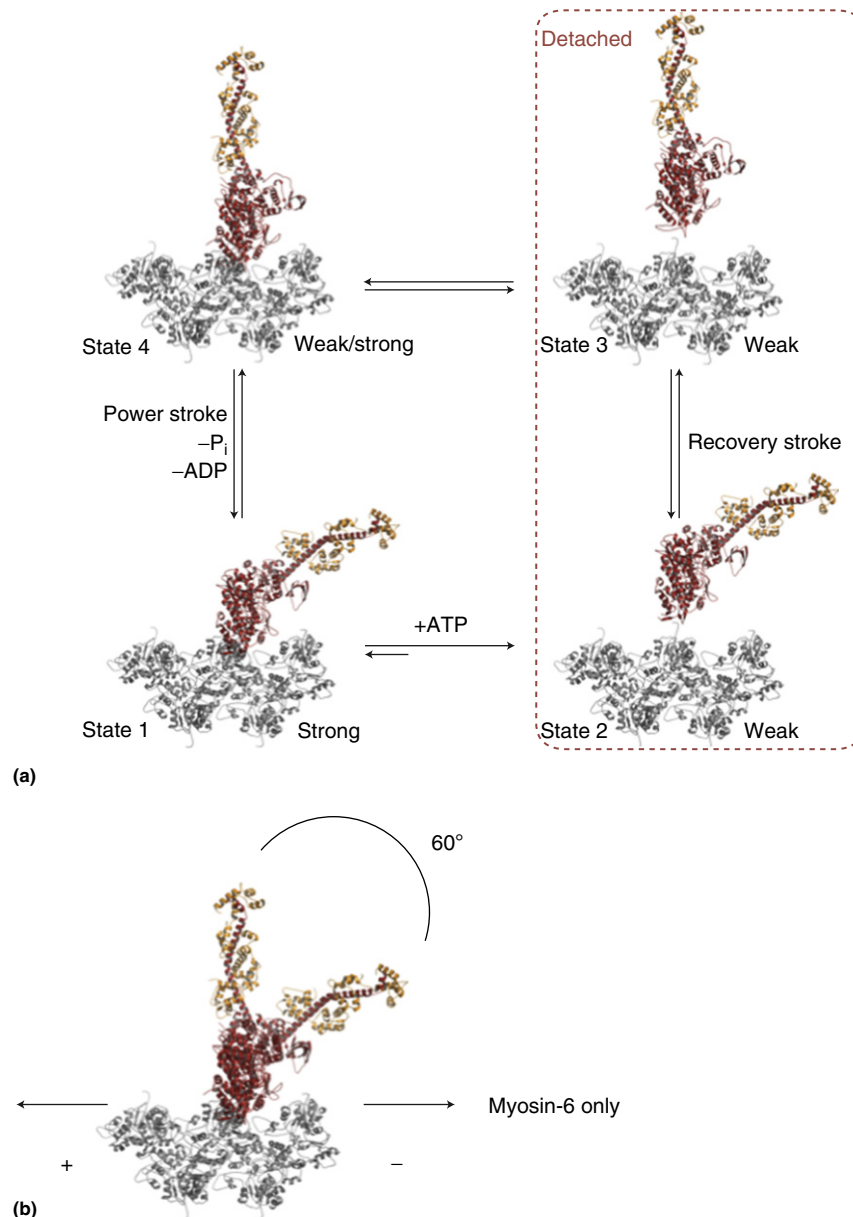


Figure 3 (a) Kinetic cycle of actomyosin interaction during ATP hydrolysis: State 1: In the absence of ATP, myosin has a high affinity for actin and the lever arm is in a post-powerstroke state. State 2: Upon binding of ATP, the affinity of myosin declines ~ 1000 fold and myosin dissociates. State 3: Upon hydrolysis of ATP the lever arm recovers to the pre-powerstroke state and the myosin remains in a state ($ADP \cdot P_i$) with weak affinity for actin. State 4: Upon release of P_i the myosin affinity for actin increases and myosin is strongly bound with the lever arm in the pre-powerstroke configuration. The powerstroke is associated with the events involved in ADP release. (b) Schematic comparison of the lever arm conformation in the pre- and the post-power stroke state. The lever arm rotates around 60° during the hydrolysis cycle and results in a lateral displacement of the myosin motor domain on F-actin. With the exception of unconventional myosin-6 which moves toward the (-) end of F-actin, all other myosins move to the (+) end of the filament.

2014). The kinetic tuning likely coincides with myosin function. Fast kinetics is required for contracting a skeletal muscle fiber or for transporters that move cargos over long distances in cells. In contrast, the kinetics of some myosins are more adapted to a role as an F-actin tether that is able to maintain force for long periods of time with less energy usage.

Apart from kinetic and structural signatures of the heavy chain, myosins operate under different regimes to increase the

functional diversity: As monomers, dimers, or multimers. Monomeric myosins are associated with tethering functions on F-actin but in some cases can be rendered into dimers by external or internal cues such as cargo-induced oligomerization or clustering. Dimeric motors such as class-5 and -11 myosins are efficient transporters, which move cargoes like organelles, vesicles, and RNA along the actin cytoskeleton. Uniquely, class-2 myosins assemble into higher order sidepolar or

bipolar filamentous arrays consisting of 14–300 motors (Billington *et al.*, 2013).

Mechanochemical ATPase Cycle

The mechanochemical actomyosin ATPase cycle is conserved in all enzymatically active myosins. The basic concept of the cycle are antagonistic affinities for ATP and F-actin, as shown in the Lymn–Taylor cycle (Figure 3; Lymn and Taylor, 1971). Myosin strongly binds to F-actin in the absence of nucleotide (state 1). This state is also referred to as rigor state. ATP binding causes the dissociation of the actomyosin complex (state 2). ATP hydrolysis occurs in the actin-detached state and is associated with a lever arm rotation, the recovery stroke. In the posthydrolytic ADP.Pi state, myosin weakly rebinds to F-actin (state 3). The subsequent P_i-release is associated with small conformational changes in the myosin motor domain, which are transmitted to the neck region, which functions as a rigid lever during its force producing swing. This power stroke produces effective displacement of the myosin.ADP complex on F-actin and shifts the myosin to a strong F-actin binding state. ADP-release transitions the motor back to the rigor state, in which the cycle repeats (state 4). Similar to Archimedes' law of the lever, the length of the myosin lever arm positively correlates with the displacement of the motor. With the exception of class-6 myosins, which move toward the pointed (–) end, all myosins characterized so far traffic toward the barbed (+) end of F-actin (Figure 3(b)).

The fraction of time the myosin spends in the strong actin-binding states (A.M and A.M.ADP) is the 'duty ratio' – a measure for the functional capacities and limitations of a motor (De La Cruz and Ostap, 2004; Bloemink and Geeves, 2011). Processive myosins which take multiple steps before detaching from F-actin are characterized by a high duty ratio (>0.50). A prototypic example for a high duty ratio motor is the double-headed unconventional myosin-5a, which functions as an effective cargo transporter because one head is in continuous interaction with the F-actin track. By contrast, a single-headed myosin, such as skeletal muscle myosin-2 with a low duty ratio of 0.02 must work in arrays composed of >300 heads to maintain continuous interaction with the F-actin. Additionally, single-headed myosins can work in small ensembles such as clusters on, for example, cargo vesicles to ensure bona fide transport function. Based on the mechanokinetic signatures, myosins are categorized in four different classes: fast movers, slow/efficient force holders, strain sensors, and processive motors (Bloemink and Geeves, 2011). Catalytically inactive pseudoenzymes such as class-18 and -20 myosins form a separate group.

Pathophysiology

Mutations, malfunction, and dysregulation of myosins are associated with heterogeneous pathophysiological conditions in mammals like cardiomyopathies, cancer, blindness, diabetes, deafness, and blood disorders. The mutational spectrum on the genes encoding for the myosin heavy chain and the associated light chains contain hundreds of mutations. Well-

studied examples include cardiac myosin-2 heavy chain gene mutations which result in cardiac disorders associated with aberrant cardiac muscle function and the impact of mutations in conventional and unconventional myosins in multi-spectrum disorders associated with congenital and progressive vision and hearing impairment (Petit and Richardson, 2009; Mansson, 2014; Spudich, 2014).

Due to the involvement of the myosome in processes such as the organization of the mitotic spindle, cytokinesis, and cell migration, myosin malfunction can cause various cancer types, including prostate cancer, where enhanced carcinogenesis and metastasis are linked to increased myosin-6 levels (Puri *et al.*, 2010). Malfunction or ablation of the tumor suppressor nonmuscle myosin-2a is associated with a predisposition to squamous cell carcinomas with poor survival (Schramek *et al.*, 2014). Nonmuscle myosin-2a is also linked to blood disorders and associated syndromic phenotypes (Ma and Adelstein, 2014). Other diseases are associated with myosin's function as cargo transporters and can result in cases where physiological cargoes cannot be delivered appropriately. A prototypic example is the pigment dilution of skin and hair due to impaired transport of melanosomes in melanocytes by myosin-5a (Wu *et al.*, 1997). Utilization of the actomyosin transport system by non-physiological cargoes such as viruses facilitates host cell entry, transport, and viral egress (Schudt *et al.*, 2013; Arii *et al.*, 2010; Roberts and Baines, 2010; van Leeuwen *et al.*, 2002). Indirectly, parasite myosins cause a broad-range of both zoonotic and non-zoonotic diseases with medical and veterinary importance along with high economic impact. Because of the plethora of myosin-dependent functions, pharmacotherapeutic bench to bedside approaches will be a substantial part of translational myosin research in the future.

Class-1 Myosins

Class-1 myosins are very abundant across phylogeny and the first class of unconventional myosins discovered 41 years ago (Pollard and Korn, 1973). Class-1 myosins are small, monomeric motors which control membrane tension and dynamics by acting as crosslinkers between the F-actin cytoskeleton and the latter. Crosslinking properties are mediated by phospholipid binding tail homology (TH) domains and F-actin binding properties of the motor domain. Based on the tail domain, class-1 myosins can be classified as short and long-tailed. Short-tailed class-1 myosins have a TH1 domain in their tails whereas long-tailed class-1 myosins harbor additional domains that can mediate protein–protein interactions or ATP-insensitive F-actin binding (McConnell and Tyska, 2010).

Class-1 myosins are involved in pathophysiological relevant tasks such as the transport of glucose transporter GLUT4 containing vesicles or organelles. Further, they mediate mechanosensory functions of sensory hair cell stereocilia and play pivotal roles in hematopoietic cells where they modulate immune functions (Bose *et al.*, 2002; Gillespie 2004; Santos-Argumedo *et al.*, 2013). Further, class-1 myosins might play a role in the regulation of RNA polymerase 1, transcription, and cell cycle progression in the nucleus (Ye *et al.*, 2008; Sarshad *et al.*, 2013; Percipalle *et al.*, 2006).

Class-2 Myosins

Citius, Altius, Fortius: Ubiquitous class-2 myosins constitute one of the largest subfamilies with up to 22 myosins per vertebrate organism (Odrionitz and Kollmar, 2007). Class-2 myosins are hexameric, consisting of a myosin heavy chain homodimer that non-covalently associates with a set of ELC and RLC. A tail coiled-coil domain mediates homodimerization. Uniquely, conventional myosins form filaments with characteristic signatures: Sarcomeric class-2 myosins from striated and cardiac muscle ensemble into bipolar filaments consisting of ~400 myosins and form the thick filaments that – together with actin thin filaments – assemble into parallel, highly stereotyped sarcomeres which cause the striped pattern of muscle. At the molecular level, sarcomere shortening and hence muscle contraction occurs when thick myosin filaments and thin actin filaments slide past each other.

Mammalian smooth muscle does not form precise sarcomeres and instead their myosins assemble into side-polar filaments consisting of hundreds of myosin molecules (Xu *et al.*, 1996). In contrast to sarcomeric myosins, these filaments continuously assemble and disassemble (Seow, 2005). Phosphorylation of the associated RLC by calcium/calmodulin-dependent myosin light chain kinase or dephosphorylation by myosin phosphatase acts as on/off switch for myosin's catalytic activity. In the absence of RLC phosphorylation, smooth muscle myosin adopts an autoinhibited, compact 10S conformation and adopts an extended 6S conformation upon RLC phosphorylation which readily assembles into filaments (Dulyaninova and Bresnick, 2013). Similar regulatory features are described for cytoplasmic nonmuscle myosin-2 paralogs. As the closely related myosin from smooth muscle, nonmuscle myosin-2 adapts a compact, autoinhibited conformation in the absence of RLC phosphorylation. Nonmuscle myosin-2 paralogs are regulators of spatiotemporal cytoskeletal dynamics during essential processes such as cell migration, adhesion, and cytokinesis and form short bipolar filaments with an average of 14–30 molecules (Billington *et al.*, 2013; Vicente-Manzanares *et al.*, 2009; Heissler and Manstein, 2013). In contrast to conventional myosins from smooth and skeletal muscle, the three nonmuscle myosin-2 paralogs found in mammalian cells have a vast tissue distribution and are expressed in a developmentally dependent manner. Though their kinetic features are related, functional activity is non-redundant but overlapping may occur (Vicente-Manzanares *et al.*, 2009).

Class-3 and -9 Myosins

Class-3 and -9 myosins have dichotomous functions as actin-based motor proteins and signaling molecules. Class-3 represents small monomeric myosins that harbor an N-terminal kinase domain prior to the motor domain. The kinase domain confers serine/threonine kinase activity and implicates a function for class-3 myosins in intracellular signaling. Autophosphorylation has been reported for class-3 myosins, though other protein targets for the kinase domain have not been identified (Dalal *et al.*, 2011). Class-3 myosins are implicated in the function of ciliary hair cells and retinal

photoreceptors. In mammals, mutations of *MYO3A* lead to deafness (Walsh *et al.*, 2011). In *Drosophila*, the myosin-3 analog NINAC plays a crucial role in phototransduction and rhabdome maintenance (Porter *et al.*, 1992). In mammals, class-3 myosins have been proposed move along parallel F-actin bundles. In myosin-3a, this motility is thought to be conferred by a second F-actin binding site in the tail domain, which allows the motor to be in continuous interaction with its track (Les Erickson *et al.*, 2003). Conversely, myosin-3b lacks the respective site and associates with binding partners such as espin-1 that harbor an actin binding site for facilitated motility (Merritt *et al.*, 2012).

Signaling activity is also reported for class-9 myosins, which are processive, single-headed molecular motors exclusively found in animals. Invertebrates contain one class-9 myosin, mammals two. Myosin-9a functions in the regulation of epithelial cell morphology, differentiation, and migration (Abouhamed *et al.*, 2009; Omelchenko and Hall, 2012). Distinctively, myosin-9b regulates shape and motility of immune cells, a crucial function during the immune response (Hanley *et al.*, 2010). The myosin-9 heavy chain has three structural signatures: A N-terminal extension with a pseudo Ras-binding domain, an unusual long surface loop in the myosin motor domain that contains a calmodulin binding site and a tail domain which contains a Rho GTPase-activating protein (Rho-GAP) domain (Bahler *et al.*, 2011). The latter increases the GTPase activity of members of the Rho family such as RhoA and Cdc42, which are regulators of cytoskeletal dynamics and cell cycle, thereby negatively regulating the Rho signaling pathways. The presence of a characteristic myosin motor domain and Rho-GAP domain enables myosin-9 to function dichotomously as molecular motor and signaling molecule, regulating and organizing the actin cytoskeleton at the same time (Bahler *et al.*, 2011). The long surface loop in the myosin motor domain has been shown to facilitate processive motion of the single-headed myosin by acting as an electrostatic tether, a unique feature of class-9 myosins (Elfrink *et al.*, 2014).

Class-5 Myosins

Class-5 is the best understood myosin class and expressed in most eukaryotes except plants – where class-8 and -11 myosins have a similar functional and structural features. Class-5 myosins are dimeric, short-range cargo transporters, and dynamic organelle tethers. Cargos range from mRNA to organelles such as the endoplasmic reticulum and pigment granules (Wu *et al.*, 1997; Wagner *et al.*, 2011; Yoshimura *et al.*, 2006). Dimerization via a coiled-coil forming sequence paralleled by mechanochemical characteristics like a high duty ratio support a processive stepping motion on actin filaments and bundles – a prerequisite for efficient cargo transport. Myosin-5 dimers walk in a hand-over-hand fashion on F-actin, taking multiple 36 nm steps which match the 36-nm F-actin pseudorepeat (Mehta *et al.*, 1999). During processive movements, the two heads of a myosin-5 bind 36 nm apart on an actin filament defining a leading and trailing head. It is likely that both heads are in an ADP-bound state during these dwells (Walker *et al.*, 2000). Intramolecular strain results in the dissociation of ADP from the trailing head, shifting it into the rigor state.

Subsequently, ATP binding to the trailing head results in its dissociation from F-actin, thereby triggering the leading head to undergo a power stroke which positions the new leading head (in the ADP.Pi state) to a forward F-actin binding site. The new leading head releases Pi and binds strongly to F-actin, thereby returning to the starting point of the mechanoenzymatic cycle, but translating 36 nm forward on F-actin (Sakamoto *et al.*, 2008). Processivity, the ability to make multiple steps before detaching from its tack, and run length of class-5 myosin is further determined by binding partners which function as regulatory switches for myosin-5: In the absence of cargo, myosin-5 adopts an autoinhibited conformation in which both heads fold back on the tail domain. The backfolded conformer is kinetically inert, preventing futile ATP consumption (Sellers and Knight, 2007). Cargo-loading is associated with an unfolding of myosin from an autoinhibited conformation and rendering the motor to an efficient transporter (Wu *et al.*, 1997; Sellers and Knight, 2007; Skolnick *et al.*, 2013).

Class-6 Myosins

Myosin-6 is the sole myosin identified so far that moves toward the pointed (–) end of F-actin (Figure 3(b)). Direction-reversal is caused by a unique insert located between the myosin motor domain and the neck region. The insert binds calmodulin and redirects the lever arm respective to the myosin motor domain (Menetrey *et al.*, 2005). Myosin-6 is involved in various cellular processes such as endo- and exocytosis, cell migration, the maintenance of Golgi morphology, and cancer cell metastasis (Puri *et al.*, 2010; Sahlender *et al.*, 2005; Tumbarello *et al.*, 2013; Buss and Kendrick-Jones, 2008). Recent work indicates a modulatory function for nuclear myosin-6 in RNA polymerase 2 dependent transcription regulation (Vreugde *et al.*, 2006). The interaction between the myosin tail domain and specific adapter and cargo proteins mediates the vast physiological tasks as F-actin tether or processive transporter. Further, the interaction of myosin-6 and its binding partners changes the mechanoenzymatic properties of the motor: Cargo-free myosin-6 is a nonprocessive monomer with actin tethering properties whereas the cargo-bound dimer acts as processive transporter (Yu *et al.*, 2009; Buss *et al.*, 2004; Lister *et al.*, 2004).

Class-7, -10, and -15 Myosins

MyTH-FERM myosins is a collective term for myosin classes-7, -10, and -15. Those myosins share the conserved dyad of a MyTH4 domain followed by a FERM domain in their tails, suggesting a close phylogenetic and functional relationship. The light chain content and the oligomeric state of MyTH-FERM myosins are still controversial. Among MyTH-FERM myosins, myosin-10 is ubiquitously expressed whereas class-7 and -15 myosins are often targeted to sensory organs and the endocrine system (Lloyd *et al.*, 2001; Hasson *et al.*, 1995; Berg *et al.*, 2000). A common signature of MyTH-FERM myosins is their localization to parallel actin-rich structures as they are found in stereocilia, microvilli, and filopodia (Kerber and

Cheney, 2011; El-Amraoui and Petit, 2005; Hartman *et al.*, 2011; Lu *et al.*, 2014).

Myosin-7a is implicated in vision and hearing in mammals and *Drosophila* and mutations in MYO7A associated with multisystem disorders characterized by deafness and retinal degeneration (Weil *et al.*, 1995). In the eye, myosin-7a localizes to the retinal pigmented epithelium where it is required for the visual retinoid cycle and the transport of opsin, phagosomes, melanosomes, and vesicles of the endoplasmic reticulum along with microtubule motors (Lopes *et al.*, 2011; Gibbs *et al.*, 2004). In the densely packed F-actin-rich structures in sensory hair cell stereocilia, myosin-7a plays a pivotal part in the mechanotransduction complex, serving in tip-link tensing and the maintenance of stereocilia structure (Grati and Kachar, 2011).

Myosin-15a is the largest heavy chain due to a long N-terminal proline-rich extension without recognizable motifs. The function of the extension is unknown, however, mutations in the domain are linked to congenital deafness in humans (Liang *et al.*, 1999). Myosin-15a is a critical determinant for stereocilia elongation, maturation, and the formation of the staircase-like architecture (Belyantseva *et al.*, 2003, 2005). Furthermore, myosin-15a is part of the stereocilia elongation complex and transports cargos such as the scaffolding protein whirlin to the tip of hair cell stereocilia (Belyantseva *et al.*, 2005).

Myosin-10 functions in pseudopod extension during phagocytosis in macrophages (Cox *et al.*, 2002). Myosin-10 localizes to the tips of filopodia and is significant for filopodial dynamics, formation, and intrafilopodial motility (Bohil *et al.*, 2006; Berg and Cheney, 2002). The myosin-10 MyTH4 domain was shown to bind microtubules, suggesting myosin-10 to be an integrator of the F-actin and microtubule cytoskeleton (Weber *et al.*, 2004).

Class-8 and 11 Myosins

Class-8 and -11 represent large multigene families and are the sole plant specific myosins. Class-8 myosins are involved in new cell wall formation, endocytosis, cell–cell transport and communication through plasmodesmata (Tominaga and Nakano, 2012). In moss, myosin-8 is required for bona fide protonemal patterning and development. Ablation of all class-8 myosins is associated with growth defects and cell expansion (Wu *et al.*, 2011). The action of the actomyosin cytoskeleton on plasmodesmata and intracellular movement is of particular interest since plant viruses such as tobacco mosaic virus and grapevine fanleaf virus hijack the actomyosin cytoskeleton for intracellular transport and cell-to-cell spreading of the viral infection (Kawakami *et al.*, 2004; Amari *et al.*, 2011; Harries *et al.*, 2009). Hijacking of the actomyosin cytoskeleton by plant viruses causes severe loss in crop yields worldwide and has a high socio-economic impact.

Structure follows function: Class-11 myosins are morphologically similar to mammalian class-5 myosins and involved in fast intracellular trafficking of organelles and vesicles such as chloroplasts, mitochondria, the endoplasmic reticulum, and peroxisomes. Organelle-associated myosin-11 is involved in cytoplasmic streaming, a form of directed, active actin-based

flow of the cytoplasm, and organelles in cells from algae to higher plants. In the alga *Chara*, myosin-11 propels F-actin with a velocity of $50\text{--}60\ \mu\text{m s}^{-1}$ – making it the fastest myosin motor characterized so far (Higashi-Fujime *et al.*, 1995). Studies with chimeric myosins suggest that the myosin-based cytoplasmic streaming is a determinant of plant size in *Arabidopsis*, implicating a potential future application in artificial size control in plants (Tominaga *et al.*, 2013). In *Arabidopsis*, simultaneous knockout of the three essential class-11 myosins that are responsible for organelle transport results in plant-development and fertility defects (Ojangu *et al.*, 2012). Ablation of both class-11 myosins in moss indicates their function the establishment of cell polarity and growth (Vidali *et al.*, 2010).

Class-14 Myosins and -21 Myosins

Class-14 myosins are specific to protozoan apicomplexan parasites such as *Plasmodium* and *Toxoplasma* which cause malaria and toxoplasmosis, respectively. Myosin-21 is found in *Leishmania*, a trypanosomatid which is responsible for leishmaniasis in humans. Structurally, class-14 myosins are atypical by lacking consensus IQ motifs and having a very short tail domain. The actomyosin cytoskeleton plays an integral part in locomotion, host cell invasion and infection of apicomplexan invasive stages (Kappe *et al.*, 2004). Apicomplexan parasites actively use a gliding motility for host cell entry, which relies on the parasite substrate-dependent actomyosin-mediated gliding motility, which pushes the parasite in the host cell – independent from the host cell cytoskeleton (Matuschewski and Schuler, 2008). Despite a highly degenerated IQ motif and the absence of a tail domain, class-14 myosins bind light chains, which contribute to both, gliding motility and invasion and uniquely anchor the myosin to the plasma membrane (Polonais *et al.*, 2011).

Myosin-21, which plays a role in the assembly of the flagellum and intracellular trafficking, has the distinguishing feature to undergo a light chain mediated dimerization (Batters *et al.*, 2014; Katta *et al.*, 2009, 2010). Calmodulin binding to the neck region prevents dimerization of the monomeric heavy chain, calmodulin dissociation its dimerization. Uniquely, the presence of calmodulin is paralleled by the ability of the motor to bind phospholipids and to power motility. Its absence renders myosin-21 an immobile actin-crosslinker that is unable to bind lipids (Batters *et al.*, 2014).

Class-16 and -18 Myosins

Class-16 myosins are single-headed motor proteins characterized by a large N-terminal ankyrin repeat domain which associates with the catalytic subunits PP1 α and PP1 γ of protein phosphatase 1 (Patel *et al.*, 2001). Myosin-16b localizes throughout the cytoplasm and the nucleoplasm, excepting the nucleolus. The nuclear expression of myosin-16b is cell cycle dependent and plays a pivotal regulatory role in its progression (Cameron *et al.*, 2013). Nucleolar import is mediated by the myosin tail domain (Cameron *et al.*, 2007). Apart from class-16 myosins, nuclear localization is solely described for

myosin -1c, -2, -5a, and -5b, myosin-6, -9b, -10, and 18b (de Lanerolle, 2012).

Class-18 myosins are exclusive to metazoan and are catalytically inactive pseudoenzymes. Myosin-18a dimerizes via an extended coiled-coil in the tail domain and assembles into dimeric structures, similar to conventional myosin-2. Additionally, unconventional class-18 myosins bind ELC and RLC as light chains (Guzik-Lendrum *et al.*, 2013). Even though there is a high degree of overall structural similarity with conventional class-2 myosins, filamentation is not reported for class-18 myosins. Myosin-18a isoforms have a long N-terminal extension which contains a PDZ domain and an ATP-insensitive F-actin binding site, though overall, the motor has a very low affinity for F-actin (Mori *et al.*, 2005). Sequence alignments highlight that conserved structural elements in the nucleotide binding pocket are altered, which might be the cause of the comprised catalytic activity (Guzik-Lendrum *et al.*, 2013). The same might be true for class-16 myosins suggesting that their physiological role is structural rather than catalytic. A physiological role of myosin-18a as binding partner of the PAK2/betaPIX/GIT1 complex suggests a potential function in modulating epithelial cell migration (Hsu *et al.*, 2010). Moreover, myosin-18a might be implicated in the maintenance of Golgi morphology and myogenesis (Bonn *et al.*, 2013; Dippold *et al.*, 2009).

Class-19 Myosins

Myosin-19 is a small monomeric, vertebrate specific myosin with a high duty ratio that traffics to the plus-end of F-actin. Interestingly, it binds the RLC as light chain and but its phosphorylation has no significant effect on the mechano-enzymatic motor properties (Lu *et al.*, 2014). Myosin-19 localizes to mitochondria and might be actively involved in mitochondrial positioning and transport in vertebrate cells (Quintero *et al.*, 2009).

Summary and Perspectives

The physiological tasks myosins are involved in are as versatile as the family members. Present in almost all eukaryotes, myosin motors are extensively required in all aspects of life. Actomyosin-driven contractility is indispensable during embryogenesis, development, and the maintenance of essential bodily functions in the adult organism including motile and sensory functions. At the single molecule level, myosin motors actively define the cytoplasmic architecture and transport cargos such as vesicles, mRNA, and organelles along the actin cytoskeleton to their destination in a spatiotemporal and developmentally dependent manner. Myosin malfunction – caused, for example, by mutations or misregulation – can cause severe disorders such as cardiomyopathies and cancer. Further, parasite- and virus-borne diseases often rely on the parasitic or host cell actomyosin cytoskeleton for invasion. As different as they are, all myosins share a common motor domain that interacts with ATP and F-actin – though the rates and affinities vary considerably. Structural diversity around the motor domain, particularly seen in the myosin tail domain

and both termini confer different regulatory and functional signatures. Those differences, paired with individual kinetic features, render individual myosin motor for its particular, physiological functions.

Despite an increasing number of characterized family members of different classes, the molecular mechanism underlying ATP hydrolysis and mechanotransduction as well as regulatory aspects of myosin motor function are only beginning to be understood. This is especially true for the recently identified and characterized mammalian and parasitic unconventional myosins, where information about the mechanoenzymology, oligomerization state, binding partners, and physiological functions in the molecular architecture of the cytoskeleton remain largely elusive. As research continues, the answers about the dynamics of a single myosin in time and space will be elucidated in *in vitro* and *in vivo* and substantially contribute to our understanding of physiological and aberrant myosin function in health and disease. Translational structural, pharmacological, and clinical approaches will lead to the development of myosin-specific drugs. The compound-mediated alteration of actomyosin-driven contractility or the myosin:binding partner interaction will be a milestone for basic research and the future treatment of myosin-related diseases.

Acknowledgment

This work was funded by the National Heart, Lung and Blood Institute Division of Intramural Research.

See also: Cell Communication: Cellular Outcomes: Regulation of Cell Migration. Cell Communication: Intracellular Pathways: Calcium and Calmodulin Signaling. Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell Adhesion to the Extracellular Matrix; Cell Migration. Cell Polarity; Cell–Cell Adhesion and the Cytoskeleton; Filopodia and Lamellipodia; Skeletal Muscle. Cytoskeleton and Motors: Cytoskeletal Components: Actin Assembly Dynamics and Its Regulation in Motile and Morphogenetic Processes; Dyneins; Kinesin Superfamily Proteins (KIFs) as a Fundamental Component of Life: Intracellular Transport and Beyond. Dynamic Integration: Dynamics: Dynamics of Microtubule Self-Assembly. Interorganellar Communication: Interplay and Processes: Rabs and Other G Proteins. Intracellular Infectiology: Cell Processes: Cellular Invasion by Bacterial Pathogens

References

- Abouhamed, M., Grobe, K., San, I.V., *et al.*, 2009. Myosin IXa regulates epithelial differentiation and its deficiency results in hydrocephalus. *Molecular Biology of the Cell* 20, 5074–5085.
- Amari, K., Lerich, A., Schmitt-Keichinger, C., Dolja, V.V., Ritzenthaler, C., 2011. Tubule-guided cell-to-cell movement of a plant virus requires class XI myosin motors. *PLoS Pathogens* 7, e1002327.
- Arij, J., Goto, H., Suenaga, T., *et al.*, 2010. Non-muscle myosin IIA is a functional entry receptor for herpes simplex virus-1. *Nature* 467, 859–862.
- Baboolal, T.G., Sakamoto, T., Forgacs, E., *et al.*, 2009. The SAH domain extends the functional length of the myosin lever. *Proceedings of the National Academy of Sciences of the United States of America* 106, 22193–22198.
- Bahler, M., Elfink, K., Hanley, P.J., Thelen, S., Xu, Y., 2011. Cellular functions of class IX myosins in epithelia and immune cells. *Biochemical Society Transactions* 39, 1166–1168.
- Batters, C., Ellrich, H., Helbig, C., *et al.*, 2014. Calmodulin regulates dimerization, motility, and lipid binding of Leishmania myosin XXI. *Proceedings of the National Academy of Sciences of the United States of America* 111, E227–E236.
- Belyantseva, I.A., Boger, E.T., Friedman, T.B., 2003. Myosin XVa localizes to the tips of inner ear sensory cell stereocilia and is essential for staircase formation of the hair bundle. *Proceedings of the National Academy of Sciences of the United States of America* 100, 13958–13963.
- Belyantseva, I.A., Boger, E.T., Naz, S., *et al.*, 2005. Myosin-XVa is required for tip localization of whirlin and differential elongation of hair-cell stereocilia. *Nature Cell Biology* 7, 148–156.
- Berg, J.S., Cheney, R.E., 2002. Myosin-X is an unconventional myosin that undergoes intrafilopodial motility. *Nature Cell Biology* 4, 246–250.
- Berg, J.S., Derfler, B.H., Pennisi, C.M., Corey, D.P., Cheney, R.E., 2000. Myosin-X, a novel myosin with pleckstrin homology domains, associates with regions of dynamic actin. *Journal of Cell Science* 113 (Pt 19), 3439–3451.
- Billington, N., Wang, A., Mao, J., Adelstein, R.S., Sellers, J.R., 2013. Characterization of three full-length human nonmuscle myosin II paralogs. *Journal of Biological Chemistry* 288, 33398–33410.
- Bloemink, M.J., Geeves, M.A., 2011. Shaking the myosin family tree: Biochemical kinetics defines four types of myosin motor. *Seminars in Cell & Developmental Biology* 22, 961–967.
- Bohil, A.B., Robertson, B.W., Cheney, R.E., 2006. Myosin-X is a molecular motor that functions in filopodia formation. *Proceedings of the National Academy of Sciences of the United States of America* 103, 12411–12416.
- Bonn, B.R., Rudolf, A., Hornbruch-Freitag, C., *et al.*, 2013. Myosin heavy chain-like localizes at cell contact sites during Drosophila myoblast fusion and interacts in vitro with Rolling pebbles 7. *Experimental Cell Research* 319, 402–416.
- Bose, A., Guilherme, A., Robida, S.I., *et al.*, 2002. Glucose transporter recycling in response to insulin is facilitated by myosin Myo1c. *Nature* 420, 821–824.
- Buss, F., Kendrick-Jones, J., 2008. How are the cellular functions of myosin VI regulated within the cell? *Biochemical and Biophysical Research Communications* 369, 165–175.
- Buss, F., Spudis, G., Kendrick-Jones, J., 2004. Myosin VI: Cellular functions and motor properties. *Annual Review of Cell and Developmental Biology* 20, 649–676.
- Cameron, R.S., Liu, C., Mixon, A.S., *et al.*, 2007. Myosin16b: The COOH-tail region directs localization to the nucleus and overexpression delays S-phase progression. *Cell Motility and the Cytoskeleton* 64, 19–48.
- Cameron, R.S., Liu, C., Pihkala, J.P., 2013. Myosin 16 levels fluctuate during the cell cycle and are downregulated in response to DNA replication stress. *Cytoskeleton (Hoboken)* 70, 328–348.
- Cao, Y., White, H.D., Li, X.D., 2014. Drosophila myosin-XX functions as an actin-binding protein to facilitate the interaction between Zyx102 and actin. *Biochemistry* 53, 350–360.
- Cox, D., Berg, J.S., Cammer, M., *et al.*, 2002. Myosin X is a downstream effector of PI(3)K during phagocytosis. *Nature Cell Biology* 4, 469–477.
- Dalal, J.S., Stevens Jr., S.M., Alvarez, S., *et al.*, 2011. Mouse class III myosins: Kinase activity and phosphorylation sites. *Journal of Neurochemistry* 119, 772–784.
- De La Cruz, E.M., Ostap, E.M., 2004. Relating biochemistry and function in the myosin superfamily. *Current Opinion in Cell Biology* 16, 61–67.
- Dippold, H.C., Ng, M.M., Farber-Katz, S.E., *et al.*, 2009. GOLPH3 bridges phosphatidylinositol-4-phosphate and actomyosin to stretch and shape the Golgi to promote budding. *Cell* 139, 337–351.
- Dulyaninova, N.G., Bresnick, A.R., 2013. The heavy chain has its day: Regulation of myosin-II assembly. *BioArchitecture* 3 (4), 77–85. doi:10.4161/bioa.26133.
- El-Amraoui, A., Petit, C., 2005. Usher I syndrome: Unravelling the mechanisms that underlie the cohesion of the growing hair bundle in inner ear sensory cells. *Journal of Cell Science* 118, 4593–4603.
- Elfink, K., Liao, W., Pieper, U., Oeding, S.J., Bahler, M., 2014. The loop2 insertion of type IX myosin acts as an electrostatic actin tether that permits processive movement. *PLoS One* 9, e84874.
- Gibbs, D., Azarian, S.M., Lillo, C., *et al.*, 2004. Role of myosin VIIa and Rab27a in the motility and localization of RPE melanosomes. *Journal of Cell Science* 117, 6473–6483.

- Gillespie, P.G., 2004. Myosin I and adaptation of mechanical transduction by the inner ear. *Philosophical Transactions of the Royal Society B: Biological Sciences* 359, 1945–1951.
- Grati, M., Kachar, B., 2011. Myosin VIIa and sans localization at stereocilia upper tip-link density implicates these Usher syndrome proteins in mechanotransduction. *Proceedings of the National Academy of Sciences of the United States of America* 108, 11476–11481.
- Guzik-Lendrum, S., Heissler, S.M., Billington, N., *et al.*, 2013. Mammalian myosin-18A, a highly divergent myosin. *Journal of Biological Chemistry* 288, 9532–9548.
- Hanley, P.J., Xu, Y., Kronlage, M., *et al.*, 2010. Motorized RhoGAP myosin IXb (Myo9b) controls cell shape and motility. *Proceedings of the National Academy of Sciences of the United States of America* 107, 12145–12150.
- Harries, P.A., Park, J.W., Sasaki, N., *et al.*, 2009. Differing requirements for actin and myosin by plant viruses for sustained intercellular movement. *Proceedings of the National Academy of Sciences of the United States of America* 106, 17594–17599.
- Hartman, M.A., Finan, D., Sivaramakrishnan, S., Spudich, J.A., 2011. Principles of unconventional myosin function and targeting. *Annual Review of Cell and Developmental Biology* 27, 133–155.
- Hasson, T., Heintzelman, M.B., Santos-Sacchi, J., Corey, D.P., Mooseker, M.S., 1995. Expression in cochlea and retina of myosin VIIa, the gene product defective in Usher syndrome type 1B. *Proceedings of the National Academy of Sciences of the United States of America* 92, 9815–9819.
- Heissler, S.M., Manstein, D.J., 2013. Nonmuscle myosin-2: Mix and match. *Cellular and Molecular Life Sciences* 70, 1–21.
- Higashi-Fujime, S., Ishikawa, R., Iwasawa, H., *et al.*, 1995. The fastest actin-based motor protein from the green algae, *Chara*, and its distinct mode of interaction with actin. *FEBS Letters* 375, 151–154.
- Hsu, R.M., Tsai, M.H., Hsieh, Y.J., Lyu, P.C., Yu, J.S., 2010. Identification of MYO18A as a novel interacting partner of the PAK2/betaPIX/GIT1 complex and its potential function in modulating epithelial cell migration. *Molecular Biology of the Cell* 21, 287–301.
- Kappe, S.H., Buscaglia, C.A., Bergman, L.W., Coppens, I., Nussenzweig, V., 2004. Apicomplexan gliding motility and host cell invasion: Overhauling the motor model. *Trends in Parasitology* 20, 13–16.
- Katta, S.S., Sahasrabudhe, A.A., Gupta, C.M., 2009. Flagellar localization of a novel isoform of myosin, myosin XXI, in *Leishmania*. *Molecular and Biochemical Parasitology* 164, 105–110.
- Katta, S.S., Tammana, T.V., Sahasrabudhe, A.A., Bajpai, V.K., Gupta, C.M., 2010. Trafficking activity of myosin XXI is required in assembly of *Leishmania* flagellum. *Journal of Cell Science* 123, 2035–2044.
- Kawakami, S., Watanabe, Y., Beachy, R.N., 2004. Tobacco mosaic virus infection spreads cell to cell as intact replication complexes. *Proceedings of the National Academy of Sciences of the United States of America* 101, 6291–6296.
- Kerber, M.L., Cheney, R.E., 2011. Myosin-X: A MyTH-FERM myosin at the tips of filopodia. *Journal of Cell Science* 124, 3733–3741.
- de Lanerolle, P., 2012. Nuclear actin and myosins at a glance. *Journal of Cell Science* 125, 4945–4949.
- van Leeuwen, H., Elliott, G., O'Hare, P., 2002. Evidence of a role for nonmuscle myosin II in herpes simplex virus type 1 egress. *Journal of Virology* 76, 3471–3481.
- Les Erickson, F., Corsa, A.C., Dose, A.C., Burnside, B., 2003. Localization of a class III myosin to filopodia tips in transfected HeLa cells requires an actin-binding site in its tail domain. *Molecular Biology of the Cell* 14, 4173–4180.
- Liang, Y., Wang, A., Belyantseva, I.A., *et al.*, 1999. Characterization of the human and mouse unconventional myosin XV genes responsible for hereditary deafness DFNB3 and shaker 2. *Genomics* 61, 243–258.
- Lister, I., Schmitz, S., Walker, M., *et al.*, 2004. A monomeric myosin VI with a large working stroke. *EMBO Journal* 23, 1729–1738.
- Lloyd, R.V., Vidal, S., Jin, L., *et al.*, 2001. Myosin XVA expression in the pituitary and in other neuroendocrine tissues and tumors. *American Journal of Pathology* 159, 1375–1382.
- Lopes, V.S., Gibbs, D., Libby, R.T., *et al.*, 2011. The Usher 1B protein, MYO7A, is required for normal localization and function of the visual retinoid cycle enzyme, RPE65. *Human Molecular Genetics* 20, 2560–2570.
- Lorenz, M., Holmes, K.C., 2010. The actin-myosin interface. *Proceedings of the National Academy of Sciences of the United States of America* 107, 12529–12534.
- Lu, Q., Li, J., Zhang, M., 2014a. Cargo recognition and cargo-mediated regulation of unconventional myosins. *Accounts of Chemical Research* 47 (10), 3061–3070. doi:10.1021/ar500216z.
- Lu, Z., Ma, X.-N., Zhang, H.-M., *et al.*, 2014b. Mouse myosin-19 is a plus-end-directed, high-duty ratio molecular motor. *Journal of Biological Chemistry* 289, 18535–18548. doi:10.1074/jbc.M114.569087.
- Lymn, R.W., Taylor, E.W., 1971. Mechanism of adenosine triphosphate hydrolysis by actomyosin. *Biochemistry* 10, 4617–4624.
- Ma, X., Adelstein, R.S., 2014. The role of vertebrate nonmuscle Myosin II in development and human disease. *BioArchitecture* 4, 88–102.
- Mansson, A., 2014. Hypothesis and theory: Mechanical instabilities and non-uniformities in hereditary sarcomere myopathies. *Frontiers in Physiology* 5, 350.
- Matuschewski, K., Schuler, H., 2008. Actin/myosin-based gliding motility in apicomplexan parasites. *Subcell Biochemistry* 47, 110–120.
- McConnell, R.E., Tyska, M.J., 2010. Leveraging the membrane — cytoskeleton interface with myosin-1. *Trends in Cell Biology* 20, 418–426.
- Mehta, A.D., Rock, R.S., Rief, M., *et al.*, 1999. Myosin-V is a processive actin-based motor. *Nature* 400, 590–593.
- Menetrey, J., Bahloul, A., Wells, A.L., *et al.*, 2005. The structure of the myosin VI motor reveals the mechanism of directionality reversal. *Nature* 435, 779–785.
- Merritt, R.C., Manor, U., Salles, F.T., *et al.*, 2012. Myosin IIIB uses an actin-binding motif in its espin-1 cargo to reach the tips of actin protrusions. *Current Biology: CB* 22, 320–325.
- Mori, K., Matsuda, K., Furusawa, T., *et al.*, 2005. Subcellular localization and dynamics of MysPDZ (Myo18A) in live mammalian cells. *Biochemical and Biophysical Research Communications* 326, 491–498.
- Murphy, C.T., Spudich, J.A., 2000. Variable surface loops and myosin activity: Accessories to a motor. *Journal of Muscle Research and Cell Motility* 21, 139–151.
- Odrionitz, F., Kollmar, M., 2007. Drawing the tree of eukaryotic life based on the analysis of 2,269 manually annotated myosins from 328 species. *Genome Biology* 8, R196.
- Ojangu, E.L., Tanner, K., Pata, P., *et al.*, 2012. Myosins XI-K, XI-1, and XI-2 are required for development of pavement cells, trichomes, and stigmatic papillae in *Arabidopsis*. *BMC Plant Biology* 12, 81.
- Omelchenko, T., Hall, A., 2012. Myosin-IXA regulates collective epithelial cell migration by targeting RhoGAP activity to cell-cell junctions. *Current Biology: CB* 22, 278–288.
- Patel, K.G., Liu, C., Cameron, P.L., Cameron, R.S., 2001. Myr 8, a novel unconventional myosin expressed during brain development associates with the protein phosphatase catalytic subunits 1alpha and 1gamma1. *Journal of Neuroscience* 21, 7954–7968.
- Percipalle, P., Fomproix, N., Cavellan, E., *et al.*, 2006. The chromatin remodelling complex WSTF-SNF2h interacts with nuclear myosin 1 and has a role in RNA polymerase I transcription. *EMBO Reports* 7, 525–530.
- Petit, C., Richardson, G.P., 2009. Linking genes underlying deafness to hair-bundle development and function. *Nature Neuroscience* 12, 703–710.
- Pollard, T.D., Korn, E.D., 1973. *Acanthamoeba* myosin. I. Isolation from *Acanthamoeba castellanii* of an enzyme similar to muscle myosin. *Journal of Biological Chemistry* 248, 4682–4690.
- Polonais, V., Javier Foth, B., Chinthalapudi, K., *et al.*, 2011. Unusual anchor of a motor complex (MyoD–MLC2) to the plasma membrane of *Toxoplasma gondii*. *Traffic* 12, 287–300.
- Porter, J.A., Hicks, J.L., Williams, D.S., Montell, C., 1992. Differential localizations of and requirements for the two *Drosophila* ninaC kinase/myosins in photoreceptor cells. *Journal of Cell Biology* 116, 683–693.
- Puri, C., Chibalina, M.V., Arden, S.D., *et al.*, 2010. Overexpression of myosin VI in prostate cancer cells enhances PSA and VEGF secretion, but has no effect on endocytosis. *Oncogene* 29, 188–200.
- Quintero, O.A., DiVito, M.M., Adikes, R.C., *et al.*, 2009. Human Myo19 is a novel myosin that associates with mitochondria. *Current Biology: CB* 19, 2008–2013.
- Roberts, K.L., Baines, J.D., 2010. Myosin Va enhances secretion of herpes simplex virus 1 virions and cell surface expression of viral glycoproteins. *Journal of Virology* 84, 9889–9896.
- Sahlender, D.A., Roberts, R.C., Arden, S.D., *et al.*, 2005. Optineurin links myosin VI to the Golgi complex and is involved in Golgi organization and exocytosis. *Journal of Cell Biology* 169, 285–295.
- Sakamoto, T., Webb, M.R., Forgacs, E., White, H.D., Sellers, J.R., 2008. Direct observation of the mechanochemical coupling in myosin Va during processive movement. *Nature* 455, 128–132.
- Santos-Argumedo, L., Maravillas-Montero, J.L., Lopez-Ortega, O., 2013. Class I myosins in B-cell physiology: Functions in spreading, immune synapses, motility, and vesicular traffic. *Immunological Reviews* 256, 190–202.
- Sarshad, A., Sadeghifar, F., Louvet, E., *et al.*, 2013. Nuclear myosin 1c facilitates the chromatin modifications required to activate rRNA gene transcription and cell cycle progression. *PLoS Genetics* 9, e1003397.
- Schramek, D., Sendoel, A., Segal, J.P., *et al.*, 2014. Direct in vivo RNAi screen unveils myosin IIa as a tumor suppressor of squamous cell carcinomas. *Science* 343, 309–313.

- Schudt, G., Kolesnikova, L., Dolnik, O., Sodeik, B., Becker, S., 2013. Live-cell imaging of Marburg virus-infected cells uncovers actin-dependent transport of nucleocapsids over long distances. *Proceedings of the National Academy of Sciences of the United States of America* 110, 14402–14407.
- Skolnick, M., Kremntsova, E.B., Warshaw, D.M., Trybus, K.M., 2013. More than just a cargo adapter, melanophilin prolongs and slows processive runs of myosin Va. *Journal of Biological Chemistry* 288, 29313–29322.
- Sellers, J.R., 2000. Myosins: A diverse superfamily. *Biochimica et Biophysica Acta* 1496, 3–22.
- Sellers, J.R., Knight, P.J., 2007. Folding and regulation in myosins II and V. *Journal of Muscle Research and Cell Motility* 28, 363–370.
- Seow, C.Y., 2005. Myosin filament assembly in an ever-changing myofilament lattice of smooth muscle. *American Journal of Physiology* 289, C1363–C1368.
- Spudich, J.A., 2014. Hypertrophic and dilated cardiomyopathy: Four decades of basic research on muscle lead to potential therapeutic approaches to these devastating genetic diseases. *Biophysical Journal* 106, 1236–1249.
- Tominaga, M., Kimura, A., Yokota, E., *et al.*, 2013. Cytoplasmic streaming velocity as a plant size determinant. *Developmental Cell* 27, 345–352.
- Tominaga, M., Nakano, A., 2012. Plant-specific myosin XI, a molecular perspective. *Frontiers in Plant Science* 3, 211.
- Tumbarello, D.A., Kendrick-Jones, J., Buss, F., 2013. Myosin VI and its cargo adaptors – linking endocytosis and autophagy. *Journal of Cell Science* 126, 2561–2570.
- Vicente-Manzanares, M., Ma, X., Adelstein, R.S., Horwitz, A.R., 2009. Non-muscle myosin II takes centre stage in cell adhesion and migration. *Nature Reviews Molecular Cell Biology* 10, 778–790.
- Vidali, L., Burkart, G.M., Augustine, R.C., *et al.*, 2010. Myosin XI is essential for tip growth in *Physcomitrella patens*. *Plant Cell* 22, 1868–1882.
- Vreugde, S., Ferriai, C., Miluzio, A., *et al.*, 2006. Nuclear myosin VI enhances RNA polymerase II-dependent transcription. *Molecular Cell* 23, 749–755.
- Wagner, W., Brenowitz, S.D., Hammer 3rd, J.A., 2011. Myosin-Va transports the endoplasmic reticulum into the dendritic spines of Purkinje neurons. *Nature Cell Biology* 13, 40–48.
- Walker, M.L., Burgess, S.A., Sellers, J.R., *et al.*, 2000. Two-headed binding of a processive myosin to F-actin. *Nature* 405, 804–807.
- Walsh, V.L., Raviv, D., Dror, A.A., *et al.*, 2011. A mouse model for human hearing loss DFNB30 due to loss of function of myosin IIIA. *Mammalian Genome* 22, 170–177.
- Weber, K.L., Sokac, A.M., Berg, J.S., Cheney, R.E., Bement, W.M., 2004. A microtubule-binding myosin required for nuclear anchoring and spindle assembly. *Nature* 431, 325–329.
- Weil, D., Blanchard, S., Kaplan, J., *et al.*, 1995. Defective myosin VIIA gene responsible for Usher syndrome type 1B. *Nature* 374, 60–61.
- Wu, S.Z., Ritchie, J.A., Pan, A.H., Quatrano, R.S., Bezanilla, M., 2011. Myosin VIII regulates protonemal patterning and developmental timing in the moss *Physcomitrella patens*. *Molecular Plant* 4, 909–921.
- Wu, X., Bowers, B., Wei, Q., Kocher, B., Hammer 3rd, J.A., 1997. Myosin V associates with melanosomes in mouse melanocytes: Evidence that myosin V is an organelle motor. *Journal of Cell Science* 110 (Pt 7), 847–859.
- Xu, J.Q., Harder, B.A., Uman, P., Craig, R., 1996. Myosin filament structure in vertebrate smooth muscle. *Journal of Cell Biology* 134, 53–66.
- Ye, J., Zhao, J., Hoffmann-Rohrer, U., Grummt, I., 2008. Nuclear myosin I acts in concert with polymeric actin to drive RNA polymerase I transcription. *Genes & Development* 22, 322–330.
- Yoshimura, A., Fujii, R., Watanabe, Y., *et al.*, 2006. Myosin-Va facilitates the accumulation of mRNA/protein complex in dendritic spines. *Current Biology: CB* 16, 2345–2351.
- Yu, C., Feng, W., Wei, Z., *et al.*, 2009. Myosin VI undergoes cargo-mediated dimerization. *Cell* 138, 537–548.

Further Reading

- Bahler, M., 2000. Are class III and class IX myosins motorized signaling molecules? *Biochimica et Biophysica Acta* 1496, 52–59.
- Chibalina, M.V., Puri, C., Kendrick-Jones, J., Buss, F., 2009. Potential roles of myosin VI in cell motility. *Biochemical Society Transactions* 37, 966–970.
- Coluccio, L.M., 2008. Myosins: A superfamily of molecular motors. *Proteins and Cell Regulation Series*, vol. 7. The Netherlands: Springer Verlag.
- Hammer 3rd, J.A., Sellers, J.R., 2012. Walking to work: roles for class V myosins as cargo transporters. *Nature Reviews Molecular Cell Biology* 13, 13–26.
- Heissler, S.M., Manstein, D.J., 2013. Nonmuscle myosin-2: Mix and match. *Cellular and Molecular Life Sciences: CMLS* 70, 1–21.
- Sellers, J.R., 1999. Myosins, *Protein Profile Series*, second ed. Oxford: Oxford University Press.
- Sousa, A.D., Cheney, R.E., 2005. Myosin-X: A molecular motor at the cell's fingertips. *Trends in Cell Biology* 15, 533–539.
- Sparkes, I.A., 2010. Motoring around the plant cell: Insights from plant myosins. *Biochemical Society Transactions* 38, 833–838.
- Sweeney, H.L., Houdusse, A., 2010. Structural and functional insights into the Myosin motor mechanism. *Annual Review of Biophysics* 39, 539–557.
- Williams, D.S., Lopes, V.S., 2011. The many different cellular functions of MYO7A in the retina. *Biochemical Society Transactions* 39, 1207–1210.

Relevant Websites

- www.cymobase.org
CyMoBase.
- www.mrc-lmb.cam.ac.uk/myosin/myosin.html
Myosin.
- <http://www1.imperial.ac.uk/nhli/molecular/musclebio/introduction/>
National Heart & Lung Institute.

Kinesin Superfamily Proteins (KIFs) as a Fundamental Component of Life: Intracellular Transport and Beyond

Y Tanaka, University of Tokyo, Tokyo, Japan

N Hirokawa, University of Tokyo, Tokyo, Japan and King Abdulaziz University, Jeddah, Saudi Arabia

© 2016 Elsevier Inc. All rights reserved.

Glossary

Cilia/flagella Microtubule-based cellular protrusions that function in the generation of fluid flow and signal transduction.

Glutamate receptors Major neurotransmitter receptors of excitatory neurons; involved in synaptic plasticity; classified into ionotropic receptors (including NMDA, kainite, and AMPA receptors) and metabotropic receptors.

Intraflagellar transport Bidirectional transport within cilia/flagella using the microtubule motors kinesin-2 (KIF3) and cytoplasmic dynein 2; involved in ciliogenesis and other ciliary functions.

Microtubules A major cytoskeletal component; 25-nm diameter tubular filaments consisting of alpha- and beta-tubulin heterodimers; the faster-growing plus ends generally point to the cell periphery.

Molecular motors Enzymes that generate motile force by utilizing the chemical energy from the hydrolysis of ATP into ADP and Pi.

Ventral node An embryonic endomesodermal ventral midline organ essential for left–right determination; the rapidly rotating cilia on its surface generate a leftward fluid flow that initially breaks the lateral asymmetry (nodal flow).

Introduction

Recently Emerging Concept of Intracellular Transport

Similar to railway cargoes connecting the center and the periphery of a country, intracellular transport connects the center and the periphery of the cell and plays a critical role in cellular function. Because the semantics of biological molecules vary considerably in different locations, the excess or deficiency of such molecules, including enzymes, receptors, signal transduction components, and structural molecules, at a specific site leads to serious abnormalities in cellular functions that may cause systemic diseases. Kinesins are molecular motors that are largely responsible for transporting these critical cargos along the microtubule tracks in order to properly localize these cargos within the cell. Some kinesins also serve for microtubule dynamics, mitosis and meiosis, intraflagellar transport, and signal transduction. This concept of intracellular transport has been gradually accepted by a wider audience, thanks especially to real-time visualization via the innovative technology of time-lapse microscopy, which can clearly resolve the directional and nondirectional movements of macromolecular complexes and organelles within cells.

Axonal transport in neurons is one of the best studied examples of intracellular transport. It plays a key physiological role in maintaining the structure and function of both the central and peripheral nervous systems. Astonishingly, neuronal axons, sometimes up to 1 m long, for the most part do not contain protein synthesis machinery. Thus, most of the essential proteins, macromolecular complexes, and lipids that are essential for neuronal axon function are synthesized at the cell body and are transported to axons by kinesins.

Discovery of Kinesin and Kinesin Superfamily Proteins

In a historic metabolic labeling experiment of eyeballs, protein transport toward optic nerve axons was classified into several categories, including fast and slow axonal transport ([Grafstein and Forman, 1980](#)). These transport mechanisms function at maximum speeds of 240 mm per day (group I), 34–68 mm per day (group II), 4–8 mm per day (group III), 2–4 mm per day (group IV, slow component b), and 0.7–1.1 mm per day (group V, slow component a). Structural proteins and metabolic enzymes are primarily transported by the slow mechanisms, while membrane organelles are primarily transported by the fast mechanism. High resolution electron microscopy described short filamentous structures, which are probable candidates for motor proteins that link microtubules to organelles ([Hirokawa, 1982](#)). In fact, conventional kinesin, now defined as kinesin-1 holoenzyme, was identified in 1985 via biochemical purification ([Brady, 1985](#); [Vale et al., 1985](#)) as one of the enzymes that is responsible for anterograde fast axonal transport along the microtubule cytoskeleton. One kinesin-1 holoenzyme consists of two kinesin heavy chains (KHCs; KIF5A, 5B, or 5C) that display microtubule-activated motor ATPase activity and two kinesin light chains (KLCs) that serve as cargo connectors. Later, MAP1C, also referred to as cytoplasmic dynein, was identified as a major retrograde motor ([Paschal et al., 1987](#)). We first visualized purified kinesin-1 molecules via electron microscopy, resulting in an appearance that was identical to the organelle-microtubule cross-linking structure in axons that were specifically decorated by anti-kinesin-1 antibodies ([Hirokawa et al., 1989](#); [Figure 1](#)).

The physiological relevance of kinesin-1 in intracellular transport was revealed by genetics studies of fungi, *Drosophila* ([Hurd et al., 1996](#)), and mice ([Tanaka et al., 1998](#)). For example, kinesin-1 depletion via gene targeting in mice revealed

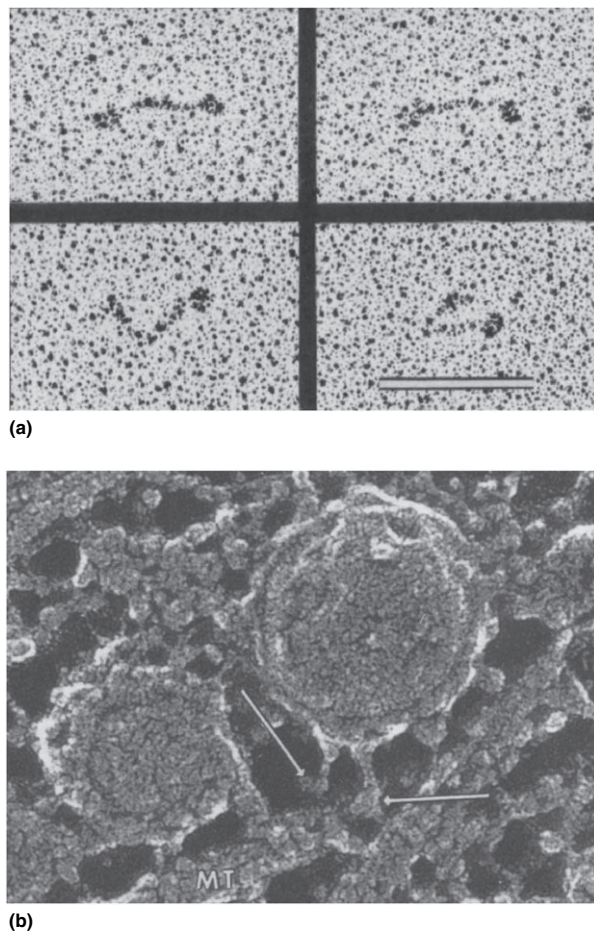


Figure 1 Structures of kinesin *in vitro* and *in vivo*. (a) High magnification image of rotary-shadowed kinesin. One or two globular domains are visible at the right end of each molecule, whereas less well-defined, but generally fan-shaped, domains can be detected at the opposite ends. (b) Quick-freeze, deep-etch electron micrograph of rat brain neurites. Many of the cross-bridges contain globular bulges where they appear to contact microtubules (see arrows). Bars, 100 nm. Reproduced with permission from Hirokawa, N., Pfister, K.K., Yorifuji, H., *et al.*, 1989. Submolecular domains of bovine brain kinesin identified by electron microscopy and monoclonal antibody decoration. *Cell* 56, 867–878.

striking perinuclear clustering of its mitochondrial cargo in extra-embryonic yolk sac cells (Tanaka *et al.*, 1998). This perinuclear clustering was proposed to be due to an imbalance of bidirectional transport in these cells, providing genetic evidence that kinesin can significantly alter the intracellular localization of organelles. Concurrently, positional cloning of yeast and fungal cell division mutants revealed the existence of kinesin-like proteins that shared several amino acid sequences with kinesin-1, such as KAR3 in yeast (Meluh and Rose, 1990). Based on RT-PCR cloning using degenerate primers corresponding to consensus sequences from a mouse brain cDNA library, our group first identified all 45 genes for the mammalian kinesin superfamily proteins (KIFs) (Miki *et al.*, 2001; Nakagawa *et al.*, 1997; Aizawa *et al.*, 1992).

Strategies in Kinesin Biology

Elucidating the structure and function of KIFs became the next focus of intracellular transport research (Goldstein, 1993; Hirokawa, 1998). Because KIFs are thought to bind to their specific cargos directly or via adaptor/scaffold molecules, the identification of specific cargos and their adaptors and scaffolds via molecular cell biology has been the most powerful approach to determine the specific intracellular roles of KIFs, together with analyses of genetic loss- or gain-of-function mutations. This approach has resulted in the occasional identification of human KIF genes that are responsible for systemic diseases and anomalies via positional cloning of hereditary disease pedigrees, for which exome sequencing represents a powerful tool. Another important question in KIF biology is how these motors are able to move forward (or backward or to perform atypical functions) using the chemical energy of ATP. Together with sophisticated biochemistry and biophysics, structural biology has been used to address this question, providing a rough outline of the activity of the engine of this remarkable mechanochemical ATPase (see Further Readings). From this perspective, in this article, we will briefly summarize the major findings and unresolved questions in kinesin biology.

Molecular Phylogeny and Nomenclature of Kinesins

A standardized kinesin nomenclature and classification system was developed after several rounds of discussions (Lawrence *et al.*, 2004). All mammalian kinesins, which are defined as those molecules which contain kinesin motor domains, follow the KIF numbering system. However, those proteins which contain a KIFC consensus sequence, which drives the motor toward microtubule minus ends, are assigned KIFC numbers. These kinesins are phylogenetically classified into subfamilies from kinesin-1 to kinesin-14, each of which often comprises a separate functional entity, as summarized in Figure 2. Kinesins 1–12 are classified as N-kinesins, which generally contain a microtubule-plus-end-directed motor domain in the N-terminus of the molecule. This group includes atypical non-motor kinesins, such as kinesin-11, which functions in signal transduction. Kinesin-13 (M-kinesin, or Kin I), whose motor domain lies in the middle of the molecule, depolymerizes microtubules. Kinesin 14 A/B is a C-kinesin, containing the minus-end-directed motor domain in the C-terminus. C-kinesins, together with cytoplasmic dynein, are involved in retrograde transport. For example, KIFC2 has been reported to retrogradely transport multivesicular body-like organelles toward the cell body (Saito *et al.*, 1997). In a more recent phylogenetic study, more types of kinesins were identified from previously uncharacterized eukaryotes, leading to the classification of 17 families in total (Wickstead *et al.*, 2010).

Detailed structural analyses of the kinesin motor domain at the atomic level have revealed that the kinesin motor domain has somewhat diverged and evolved in the course of molecular phylogeny, to give rise to new functions other than as a conventional plus-end-directed dimeric microtubule motor. For example, the monomeric kinesin-3 motor KIF1A has an extended Lysine-rich loop called 'K-loop' to gain the electrostatic interaction with microtubules that significantly augment the processivity of motility (Kikkawa *et al.*, 2001; Okada and

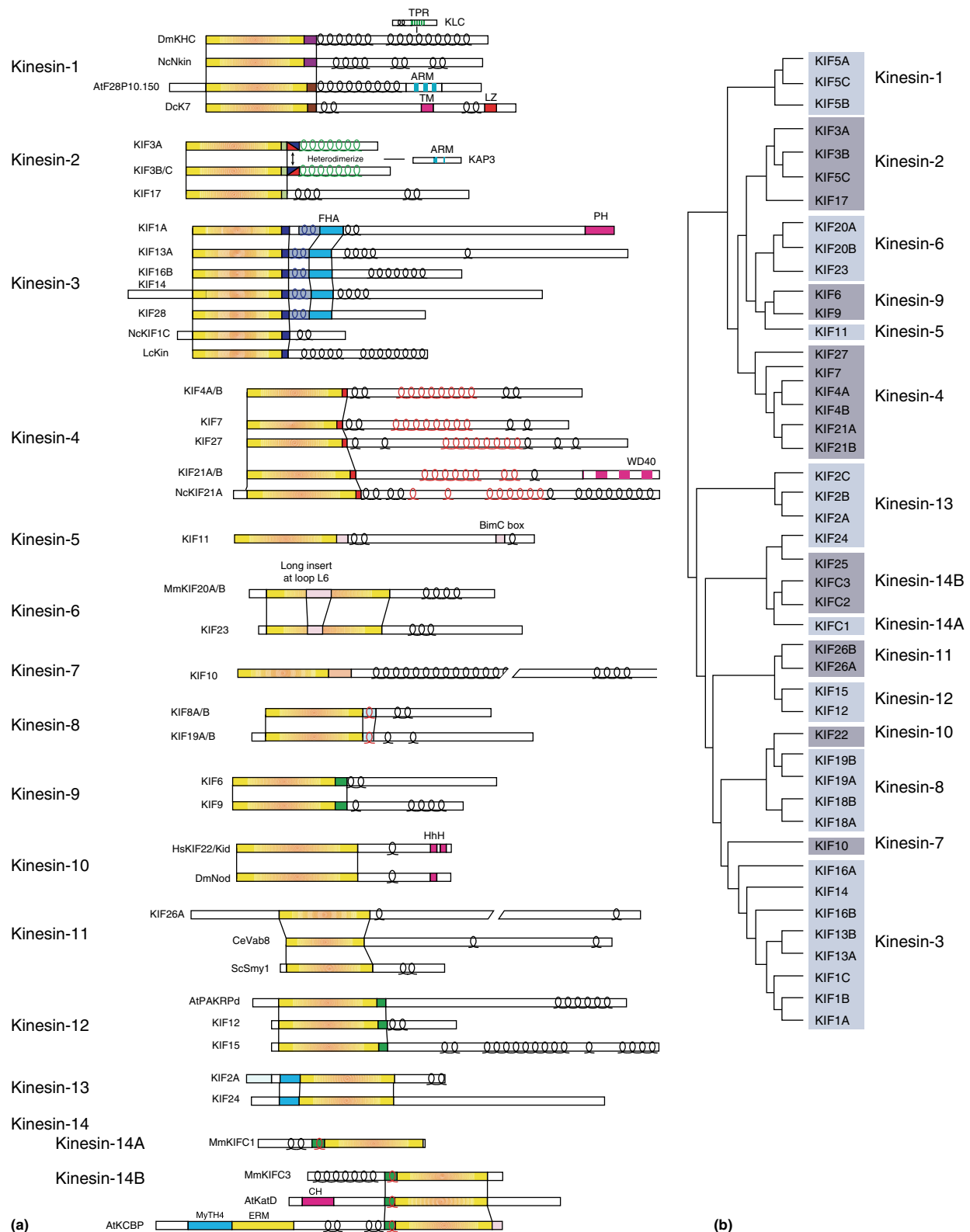


Figure 2 Diversity among KIFs. (a) Structure of KIFs. Orange boxes, kinesin motor domains. Reproduced from Miki, H., Okada, Y., Hirokawa, N., 2005. Analysis of the kinesin superfamily: Insights into structure and function. *Trends in Cell Biology* 15, 467–476. (b) Phylogenetic tree of KIFs. Reproduced with permission from Hirokawa, N., Niwa, S., Tanaka, Y., 2010. Molecular motors in neurons: Transport mechanisms and roles in brain function, development, and disease. *Neuron* 68, 610–638.

Hirokawa, 1999). In the case of the depolymerizing kinesin-13 motor KIF2A/B/C, the extended loop 2 is responsible for a depolymerizing activity (Ogawa *et al.*, 2004; Shipley *et al.*, 2004). The retrograde kinesin-14 motors (KIFCs) including Ncd have a specific C-terminal segment which appeared to be crucial for retrograde motility (Heuston *et al.*, 2010; Szczesna and Kasprzak, 2012). Thus, the vast differences in the functions of kinesins are partly due to the diversity of the pleiotropic domain structures in the motor domains.

Motor–Cargo Relationship and Pathophysiological Relevance

In this section, we describe some key examples of the roles of KIFs in important aspects of human diseases, development, and brain functions that have been discovered via molecular genetics and molecular and cellular biology studies.

Axonal Transport Abnormalities Responsible for Neurodegeneration and Common Neurological Diseases

Axonal transport is thought to be affected by several pathological signal transduction cascades, revealing overt symptoms of neuropathies (Holzbaur and Scherer, 2011). As summarized in Figure 3, the major kinesins responsible for axonal transport are kinesin-1 (KIF5A, 5B, and 5C) and kinesin-3 (KIF1A, 1Ba/β, 1C, 13A, 13B, 14, 16A, 16B, and 28). Kinesin-1 transports mitochondria, lysosomes, endosomes, and amyloid precursor protein (APP). APP is responsible for generating beta-amyloid in the brain of patients with Alzheimer disease (Selkoe, 2001), but the detailed mechanistic link between deficiencies in axonal transport and beta-amyloid accumulation and neurodegeneration in Alzheimer disease remains controversial (Lazarov *et al.*, 2005).

The kinesin-3 motors KIF1A and 1Bβ synergistically transport synaptic vesicle precursors (SVPs) containing synaptophysin, synaptotagmin, and Rab3 (Okada *et al.*, 1995; Zhao *et al.*, 2001), which are functional homologs of the *Caenorhabditis elegans* protein Unc-104 (Hall and Hedgecock, 1991). This cargo–motor association is precisely regulated by the Rab3 GTPase cycle, in which Rab3-GTP but not Rab3-GDP can bind to the KIF adaptor protein DENN/MADD. As a result, only SVPs carrying the GTP-bound form of Rab3 can move forward (Niwa *et al.*, 2008). Hippocampal KIF1A is transcriptionally upregulated by environmental stimuli downstream of brain-derived neurotrophic factor (BDNF) signaling, potentially facilitating learning and memory (Kondo *et al.*, 2012). Heterozygous KIF1B-knockout animals suffer from peripheral neuropathy, and one human pedigree of Charcot-Marie-Tooth disease type 2A1 was reported to carry a loss-of-function point mutation in the KIF1B motor domain (Zhao *et al.*, 2001). These phenotypes suggest that the gene dosage of these kinesin-3 motors could strongly affect the viability and performance of neurons.

Dendritic transport of essential molecules such as neurotransmitter receptors

Some kinesins are involved in the trafficking of neurotransmitter receptors in dendrites, thereby modulating the

sensitivity of neurons to neurotransmitters. The kinesin-2 motor KIF17 transports the glutamate receptor NR2B and contributes to learning and memory by generating a self-activation loop via cAMP response element-binding protein (CREB) phosphorylation (Yin *et al.*, 2011, 2012; Wong *et al.*, 2002; Setou *et al.*, 2000; Guillaud *et al.*, 2003, 2008). The kinesin-3 motor KIF13A transports the serotonin type 1A receptor and controls anxiety (Zhou *et al.*, 2013). The kinesin-1 motor KIF5A interacts with GABARAP to control the surface transport of the GABA_AR, which inhibits the occurrence of epilepsy (Nakajima *et al.*, 2012). Kinesin-3 motors KIF1A, KIF13A, and KIF13B are suggested to carry dendritic vesicles in a unique ‘split kinesin’ assay (Jenkins *et al.*, 2012). Kinesin-1 motor KIF5A/B/C interacts with a huge ribonucleoprotein (RNP) complex and carries mRNA into dendrites (Kanai *et al.*, 2004). These data strongly suggest that kinesins regulate brain function.

Ciliogenesis and Left–Right Determination

Cilia and flagella are unique cellular protrusions that principally function in fluid flow generation and signal transduction. The heterotrimeric kinesin-2 complex KIF3A/KIF3B/KAP3 performs an essential role in ciliogenesis by transporting structural components to the tips of cilia, a process defined as intraflagellar transport (IFT) (Rosenbaum and Witman, 2002; Nonaka *et al.*, 1998). Mouse mutants lacking kinesin-2 or IFT components display disorganized left–right determination of the internal organs (Nonaka *et al.*, 1998; Marszalek *et al.*, 1999; Takeda *et al.*, 1999), accompanied in part by other signs of ciliopathy, such as polycystic kidneys (Lin *et al.*, 2003). Ciliogenesis in the ventral node, an embryonic organ, is a key element of this phenomenon. The ventral node is a transient triangular concave structure on the ventral midline surface of the embryo at the early somite stage. Interestingly, monocilia on the ventral node rapidly rotate clockwise around a posteriorly-tilted axis, generating leftward surface fluid flow (nodal flow) (Nonaka *et al.*, 1998). Extracellular vesicles called nodal vesicular parcels (NVPs) were found to be secreted from cell surface protrusions, transferred to the left, and trapped at the left periphery of the node, where they deposit their contents (Tanaka *et al.*, 2005). This leftward flow may induce effects such as left-specific calcium elevation and expression of left-specific genes, which are the bases of the asymmetrical development of the internal organs (Figure 4).

Conversely, KIF19A, another ciliary kinesin-8 motor, depolymerizes microtubules and is essential for decreasing ciliary length (Niwa *et al.*, 2012). KIF19A-knockout mice suffered from hydrocephalus and female infertility caused by abnormally elongated cilia. Thus, the KIF3 complex and KIF19A may synergistically regulate the ciliary length within a proper range.

Depolymerizing Kinesins and Brain Wiring

The kinesin-13 motors KIF2A/2B/2C display microtubule depolymerizing activity. KIF2A-knockout mice display impaired brain wiring and layer formation defects in the developing brain. Primary cultured neurons exhibit abnormally elongated axon collaterals and defects in migration. This abnormality was proposed to be due to defects in the retraction of

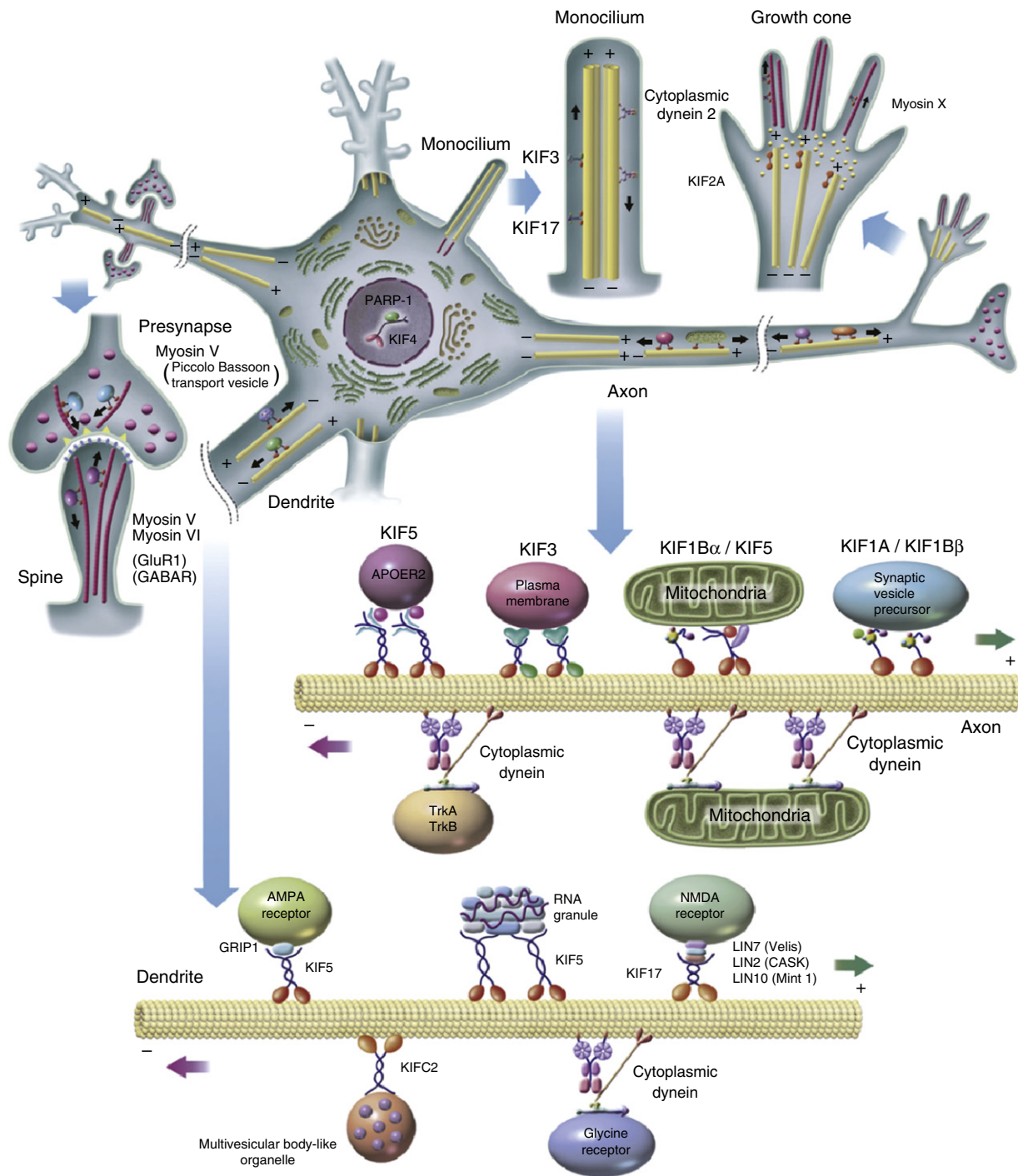


Figure 3 Kinesins in neurons. Various KIFs transport membranous organelles anterogradely in axons and dendrites, whereas cytoplasmic dynein 1 and KIFC2 transport cargos retrogradely. In neuronal cilia, anterograde transport is performed by KIF3 and KIF17, whereas retrograde transport is performed by cytoplasmic dynein 2. In short-range transport, such as transport within pre- and postsynapses and growth cone filopodia, the myosin family of proteins functions as the molecular motors. Reproduced with permission from Hirokawa, N., Niwa, S., Tanaka, Y., 2010. Molecular motors in neurons: Transport mechanisms and roles in brain function, development, and disease. *Neuron* 68, 610–638.

unnecessary collateral branches because of the loss of microtubule depolymerizing activity, which could be regulated by phosphoinositide signaling (Homma *et al.*, 2003; Noda *et al.*, 2012). KIF2C (mitotic centromere-associated protein, or

MCAK) is an Aurora kinase-regulated kinesin essential for the proper function of the mitotic apparatus. KIF2C is especially critical for kinetochore-microtubule attachments, cooperating with KIF18B in this process (Tanenbaum *et al.*, 2011).

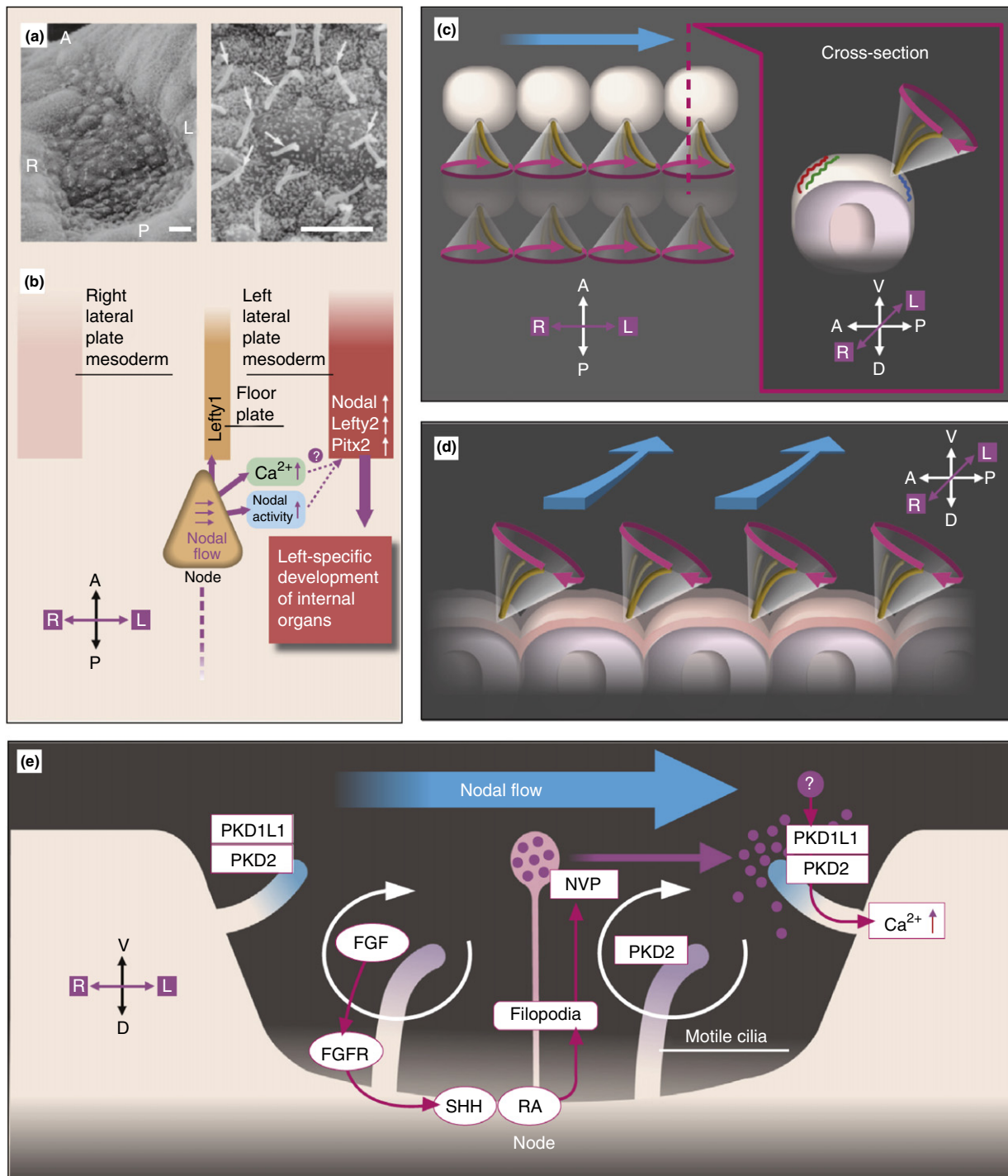


Figure 4 Kinesin-2-dependent ciliogenesis is essential for left–right body patterning: the initial symmetry-breaking process in vertebrates. (a) Scanning electron micrographs of ventral nodes of wild-type mouse embryos in the ventral view. Arrows, monocilia. Bars, 5 μm . Upper, anterior (A); lower, posterior (P); left of the figure, right (R); right of the figure, left (L) in (a)–(c). (b) Readouts of nodal flow. According to the direction of the nodal flow, the downstream side (left half) of the floor plate expresses *Lefty1*, and the left lateral plate mesoderm expresses *Nodal*, *Lefty2*, and *Pitx2*, possibly via Ca²⁺ elevation and nodal activation at the left periphery of the node. (c and d) Generation of leftward flow by tilted cilia, represented by ventral (c, left) and lateral (c, right; d) views of the node. Due to the expression of planar cell polarity (PCP) genes in the respective anterior and posterior poles of the cells (red, blue, and green lines in the inset of c), the cilia are positioned on the posterior ends of the node cells, resulting in posterior tilting of the clockwise-rotating axes. Accordingly, shear resistance on the surface influences the left-to-right cilia sweep more so than the right-to-left sweep, generating net leftward flow of the extracellular fluids. V, ventral; D, dorsal. (e) NVP hypothesis of signal propagation in the node. FGF signaling in cilia may activate sonic hedgehog (SHH) and retinoic acid (RA) signaling, inducing the release of NVPs from the cell surface with the help of filopodia-like cellular protrusions. NVPs are transported to the left via nodal flow and elevate intracellular Ca²⁺ on the left margin of the node. Reproduced with permission from Hirokawa, N., Tanaka, Y., Okada, Y., 2012. Cilia, KIF3 molecular motor and nodal flow. *Current Opinion in Cell Biology* 24, 31–39.

Kinesins and Cancer

There are some KIFs being largely involved in cell division rather than intracellular transport. For example, KIF11 and KIF15 play roles in generating and maintaining a bipolar spindle, CENPE in chromosome congression, kinesin-4 and kinesin-8 motors in microtubule regulation, and kinesin-6 motors in the central spindle and cytokinesis (Rath and Kozielski, 2012). These mitotic kinesins (kinesin spindle proteins; KSPs; Figure 5) have been considered as major targets for anticancer drugs because cancer cells possess infinite proliferative capacity. Monastrol has been developed as a prototype drug that specifically inhibits a kinesin-5 motor (KIF11, or Eg5) that is essential for spindle bipolarity (Mayer *et al.*, 1999).

Conversely, neuron-specific conditional knockout of the kinesin-2 component KAP3 leads to the formation of brain tumors (Teng *et al.*, 2005). This result was proposed to be due to the cytoplasmic role of kinesin-2 in transporting N-cadherin and β -catenin, which negatively regulates Wnt signaling and augments cell adhesion. Kinesin-2 is also responsible for cilogenesis, which serves for hedgehog signaling (Nonaka *et al.*, 1998; Han *et al.*, 2008). Thus, the kinesin-2 complex is regarded as a tumor suppressor motor via the downregulation of oncogenic signal transduction cascades.

Kinesins as Components of Signal Transduction

Other kinesins serve as components of signal transduction by simply transporting signaling molecules or modulating protein-protein interactions. As an example of the former, the kinesin-3 motor KIF16B acts in concert with the Rab14-GTP adaptor to transport fibroblast growth factor (FGF) receptor 2 through the Golgi-to-endosome pathway, and KIF16B is essential for early embryonic FGF signaling in the differentiation of the hypoblast and epiblast (Ueno *et al.*, 2011). This activity may partially oppose the activity of KIF13B in facilitating caveolin-mediated endocytosis, which is involved in cholesterol metabolism in the adult liver (Kanai *et al.*, 2014).

As an example of protein-protein interactions, the kinesin-11 protein KIF26A lacks motor activity, negatively regulates BDNF-Ret signaling by competitively inhibiting Grb2-SHC binding and is responsible for megacolon (Zhou *et al.*, 2009). Another example of such a nonmotor microtubule-associated adaptor in signal transduction is the *Drosophila* kinesin-11 protein Costal2 (Cos2), which is an essential component of the hedgehog (Hh) signaling cascade that interacts with cubitus interruptus (Ci) and suppressor of fused [Su(fu)] (Sisson *et al.*, 1997; Robbins *et al.*, 1997). However, in mammalian cells, this function in Hh signaling may be substituted by ciliary components, and thus, the precise mammalian homologue of Cos2 has yet to be identified.

Some kinesins function in nuclear signaling in addition to their roles in cytoplasmic transport. The kinesin-4 motor KIF4A binds to and suppress the activity of the enzyme poly(ADP-ribose) polymerase 1 (PARP-1) in the nucleus. This function of KIF4A is responsible for activity-dependent neuronal survival in the developing brain (Midorikawa *et al.*, 2006). KIF17b interacts with a transcriptional coactivator in male germ cells and participates in transcriptional regulation (Kotaja *et al.*, 2005).

These are only a few examples of the multimodal interactions between kinesins and their binding partners and cargos. Future studies will continue to reveal additional surprises regarding the functional relevance of KIFs.

Regulation and Synergy of Kinesins

How are kinesins regulated? This question is essential for incorporating kinesin function into broader aspects of biology, but our understanding of this topic remains limited. However, the following studies described are some examples of such studies that have examined this topic.

Regulation of the Expression Levels of Kinesins

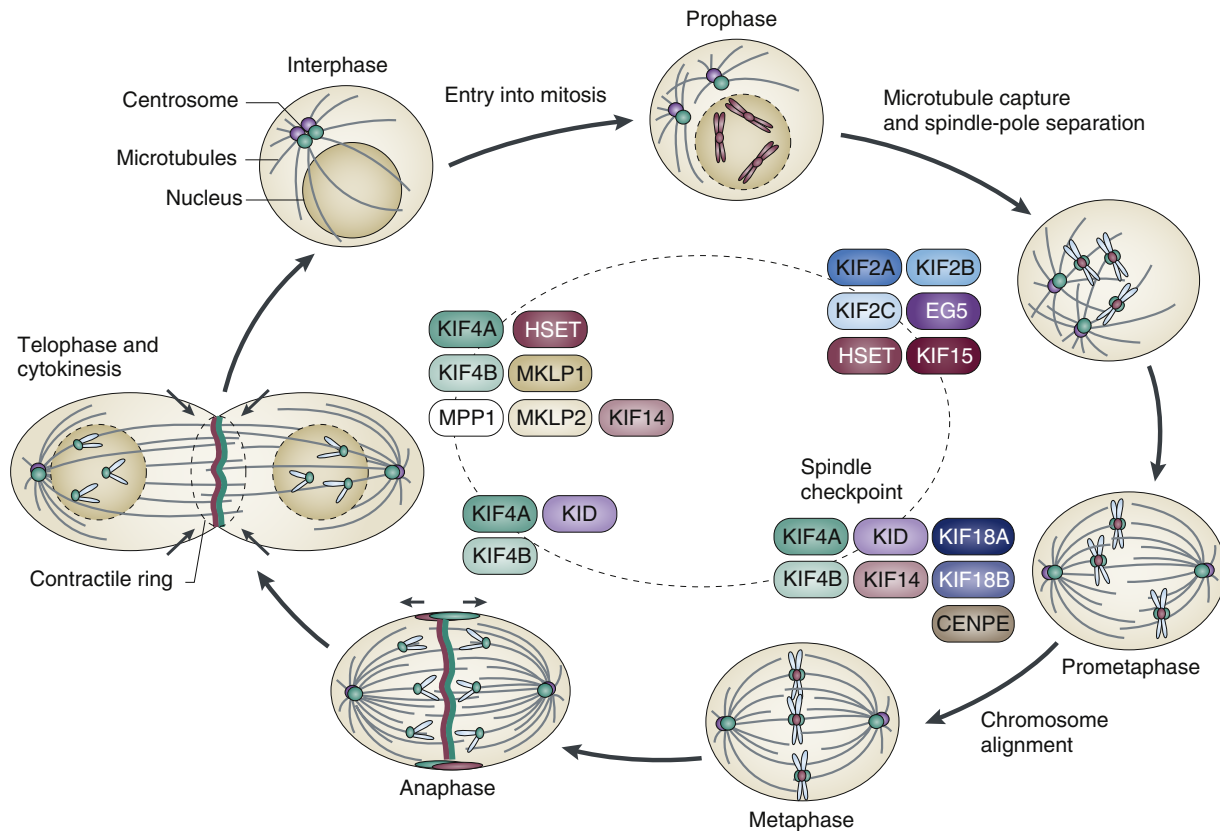
Haploinsufficiency of the *Kif1b* gene can lead to peripheral neuropathy due to a reduction in the levels of the KIF1B β protein (Zhao *et al.*, 2001), demonstrating that changes in the gene dosage of KIFs may be reflected by altered protein expression levels and rates of cargo trafficking. Point mutations in human *KIF* that cause haploinsufficiency are harmful to neurons (Hirokawa *et al.*, 2010; Tanaka and Hirokawa, 2002). Significant KIF mutations found in human neuropathies include the following examples. Knockout of murine *Kif5a* or mutation of the human homolog results in neuronal loss and SPG10-mediated spastic paraplegia (Reid *et al.*, 2002). Knockout of murine *Kif1b* or mutation of human *Kif1b* results in neuronal dysfunction and is responsible for Charcot-Marie-Tooth disease type 2A1 (Zhao *et al.*, 2001) and multiple sclerosis (Aulchenko *et al.*, 2008). *KIF21A* mutation in humans results in atrophy of extraocular muscles and CFEOM1 (Yamada *et al.*, 2003). Thus, pharmacologically augmenting the gene dosage or the residual motor activity of the healthy allele may represent a valuable therapy for these chronic neuropathies.

In contrast, some KIFs that are essential for higher brain functions are known to be upregulated by environmental stimuli. Transcription of the murine SVP transporter *Kif1a* is enhanced by BDNF signaling induced by living in enriched environments with toys in the mouse cage (Kondo *et al.*, 2012). The *Kif17* gene, which encodes a glutamate receptor transporter, is also transcriptionally activated by neuronal stimulation via phosphorylated CREB (Yin *et al.*, 2011; Wong *et al.*, 2002). It has been also known that some mitotic kinesins are transcriptionally regulated over the cell cycle (Verhey and Hammond, 2009). Examples of these types of transcriptional regulation of KIFs corresponding to physiological or pathological changes will be further elucidated by translational studies.

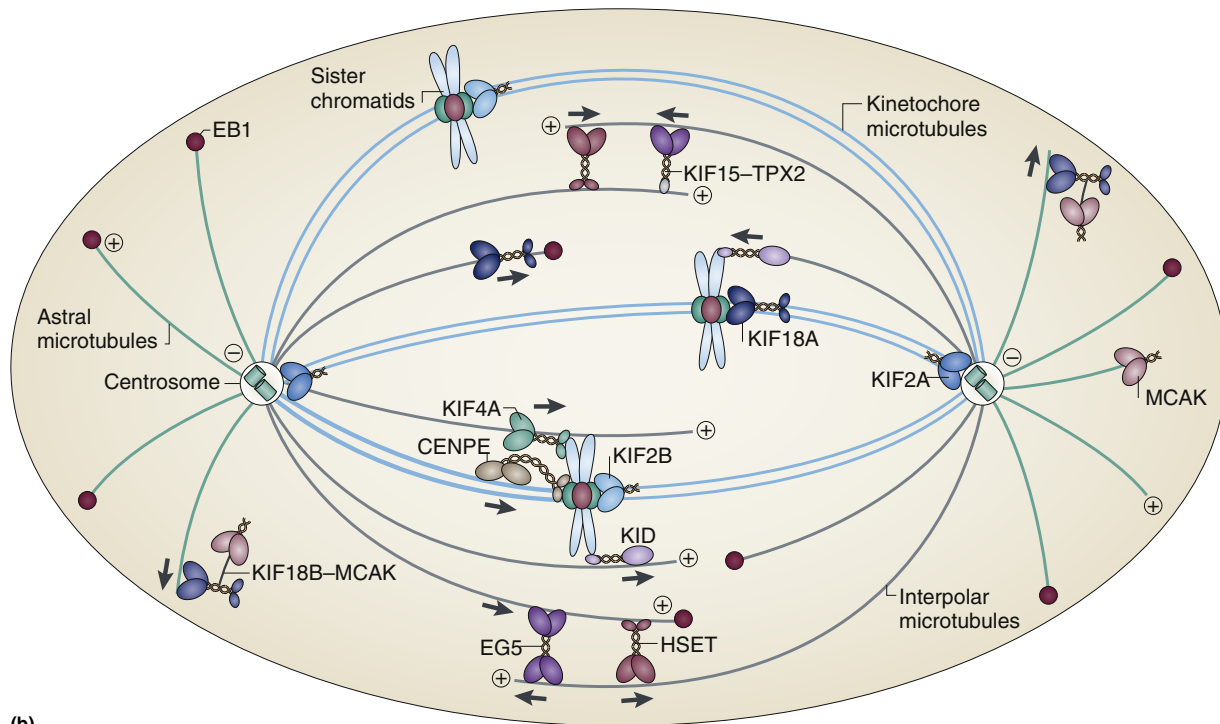
Regulation of Kinesin by Signal Transduction

The function of kinesins can be significantly regulated by global and local stimuli. In this section, we will describe several mechanisms of kinesin regulation (Figure 6).

Kinesins are phosphorylated and dephosphorylated by many kinases and phosphatases in signal transduction cascades, and these events alter their mechano-chemical ATPase activity or cargo-binding capacity. For example, kinesin-1



(a)



(b)

Figure 5 Kinesins in mitosis. (a) Kinesins in different phases of mitosis. CENPE, centromere-associated protein E (=KIF10); EG5, the bipolar mitotic kinesin KIF11; HSET, the minus-end-directed kinesin KIF1; KID, kinesin-like DNA-binding protein (=KIF22); MKLP1, mitotic kinesin-like protein 1 (KIF23); MKLP2, mitotic kinesin-like protein 2 (KIF20A, RAB6KIFL); MPP1, M-phase phosphoprotein-1 (KIF20B). (b) Kinesins in the mitotic apparatus. EB1, end binding 1; MCAK, mitotic centromere-associated protein (=KIF2C). Reproduced with permission from Rath, O., Kozielski, F., 2012. Kinesins and cancer. *Nature Reviews Cancer* 12, 527–539.

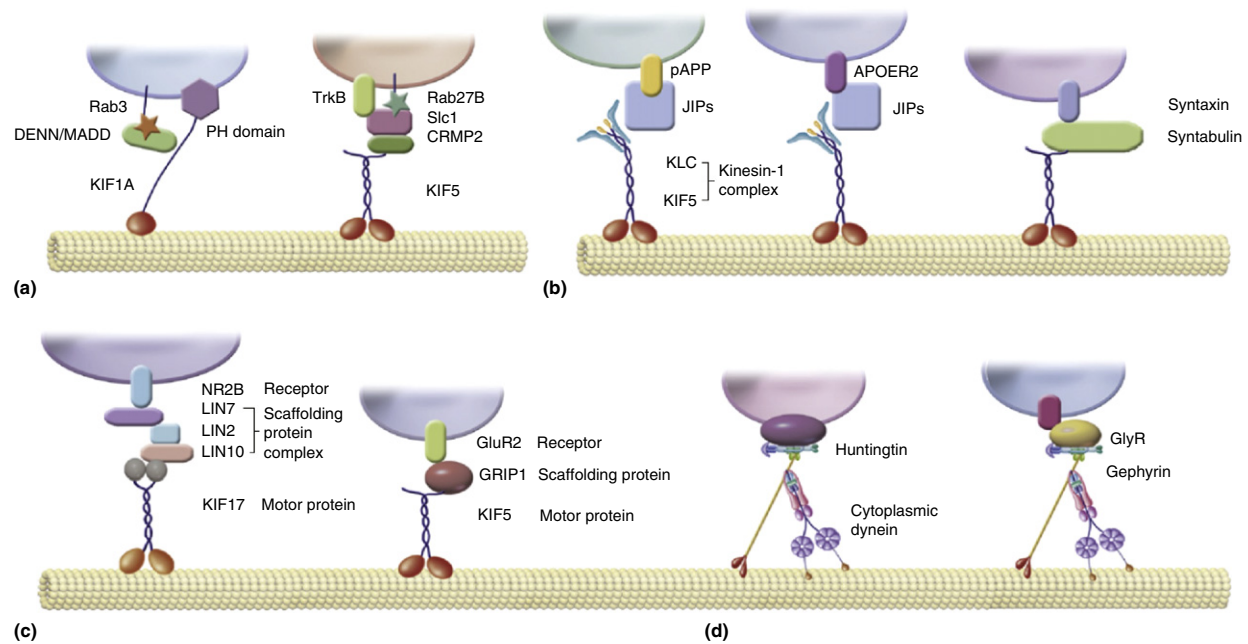


Figure 6 Regulation of kinesins. (a) KIF17 releases NMDA-receptor-2B-containing vesicles when phosphorylated by CaMKII, the activity of which is high near postsynapses. (b) KIF1A or KIF1B β recognizes GTP-Rab3 via DENN/MADD and transports Rab3-carrying synaptic vesicle precursors. Rab3 then cycles to the GDP-bound form, most likely via the action of Rab3 GAP at axon terminals, and is released from the motor. (c) A schematic illustrating high Ca^{2+} modifying the EF hand motifs of Miro to inhibit kinesin motor activity. In this model, KIF5 is constitutively localized to mitochondria. (d) A schematic illustrating the release of mitochondria from KIF5 via Ca^{2+} . Reproduced with permission from Hirokawa, N., Niwa, S., Tanaka, Y., 2010. Molecular motors in neurons: Transport mechanisms and roles in brain function, development, and disease. *Neuron* 68, 610–638.

activity is negatively modulated by JNK phosphorylation and this phosphorylation is possibly involved in the pathogenesis of Huntington's disease (Morfini *et al.*, 2009; DeBerg *et al.*, 2013). The cargo-binding domain of the kinesin-2 motor KIF17 is phosphorylated by CaM kinase II (CaMKII). This induces the release of NR2B-containing cargoes, which may be essential for dynamic changes in the cargo-motor-binding cycle, such as the efficient release of cargoes at nerve endings (Yin *et al.*, 2012; Guillaud *et al.*, 2008). The PARP1 binding capacity of KIF4A can also be regulated via phosphorylation by CaM kinase to function in cell survival signaling (Midorikawa *et al.*, 2006). Via these phosphorylation events, kinesins can respond to environmental stimuli and alter their activity to globally modulate cellular activities.

Rab GTPases have been found to be adaptors of kinesin-3 motors (Rab3 for KIF1A/1B β and Rab14 for KIF16B), and the nucleotide cycle of Rabs dramatically modulates the motor-cargo-binding capacity of kinesins (Niwa *et al.*, 2008; Ueno *et al.*, 2011). TrkB receptor is transported by kinesin-1 by the help of the Slp1/Rab27B/CRMP-2 complex (Arimura *et al.*, 2009). Because only the cargos carrying GTP-bound Rabs can associate with the motors, cellular signaling processes that activate Rab GTP exchange factors (GEFs) or GTPase-activating proteins (GAPs) could turn on or turn off axonal transport switches, respectively, both locally and globally.

The regulation of mitochondrial transport has been intensely studied. Calcium entry at synapses could halt the motility of nearby mitochondria, essentially recruiting these mitochondria to active synaptic regions that may require

additional energy. Multiple models have been proposed for the mechanism of calcium-mediated regulation of the interaction between the Milton (Trak1/2)/Miro/syntaxin adaptors, mitochondria, and kinesin-1 motors (Chen and Sheng, 2013; Macaskill *et al.*, 2009; Wang and Schwarz, 2009).

Autonomous Regulation of Kinesin Motility

In parallel to the signal-mediated mechanisms described above, autonomous regulation of motor function has been proposed at the single-molecule level (Verhey *et al.*, 2011). When GFP-tagged full-length kinesins are transfected into neurons, the GFP signal remains in the cell body. However, constructs containing only the motor domains, in most cases, translocate to the plus ends of microtubules and accumulate at the cell periphery or nerve endings. This phenomenon could be explained by the existence of an intramolecular interaction, in which the cargo-binding domain inhibits the motor domain when no cargo is present. This mechanism of self-inhibition is considered to be essential for the efficient recruitment of kinesin motors to microtubules upon cargo binding.

The relationship between kinesin motors with opposing directionalities is also an interesting issue. A 'tug-of-war' mechanism is the simplest model to explain the sudden directional changes in axonally transported organelles, which are likely simultaneously bound to multiple motors with different directionalities. Under this circumstance, what is the status of

the motor with the directionality opposite to the direction of movement? Is it inactivated, more weakly associated with the cargo, dragged along, or generating a normal force that is neutralized by the force of motors moving in the opposite direction? To date, no definitive answer to this question has been found, and multiple models could function in parallel. In the case of the transport of vesicles carrying BDNF, IGF-1/Akt-mediated phosphorylation of huntingtin augments the association of kinesin-1 with the cargo, such that anterograde transport is dominant (Humbert *et al.*, 2002). This finding suggests that receptor tyrosine kinase signaling could affect the directionality of motility by altering the balance in the accessibility of opposing motors. In the case of fungal hyphae, kinesin-3 puncta were detected on an early endosome before a switch in direction (Bielska *et al.*, 2014). However, in the case of *Drosophila* axons, eliminating the expression of a single motor (a kinesin or dynein) was sufficient to halt bi-directional transport (Martin *et al.*, 1999). This result has been largely explained by a model in which the existence of an opposing force increases motor processivity. Alternatively, the status of the local microtubule could be altered by motor binding, but changes in signal transduction or impairments of the transport of opposing motors downstream of motor loss could also be possible.

Regulation of Cargo Destinations

In some cases, the destination of intracellularly transported materials can be precisely regulated by cell polarity. This phenomenon, referred to as 'polarized sorting,' has been intensely studied, particularly in the case of kinesin-1 with respect to the axo-dendritic polarity of neurons.

First, a truncated motor containing only the KIF5 motor domain displayed a preference for axons based on their specific binding preference for GTP-tubulin, which is enriched in axons (Nakata *et al.*, 2011). In fact, a recent cryo-EM study revealed a structural difference between GTP-tubulin and GDP-tubulin and suggested structural cues for the preferential binding of KIF5 to axonal microtubules (Yajima *et al.*, 2012).

On other hand, when the kinesin attaches to dendritic cargos, such as the AMPA receptor subunit GluR2 and the EphB2 receptor, the adaptor GRIP1 directs the kinesin to the somatodendritic pathway (Misra and Ziff, 2005; Hoogenraad *et al.*, 2005; Setou *et al.*, 2002). Conversely, association of the stress protein Hsc70 with kinesin-1 is involved in determining the mode between fast and slow axonal transport (Terada *et al.*, 2010). As another factor regulating the kinesin-1 destination, two different isoforms of TRAK family adaptor proteins differentially regulates the mitochondrial transport into axons and dendrites (van Spronsen *et al.*, 2013). TRAK1 binds to both kinesin-1 and dynein/dynactin and required for axonal transport and axonal outgrowth. However, TRAK2 interferes kinesin-1 binding and predominantly localizes to dendrites to serve for dendritic development.

Posttranslational modifications of C-terminal of tubulin is also essential for kinesin activities (Janke and Kneussel, 2010). Kinesin-1 motility on neuronal β -tubulin (TUBB3) is increased by polyglutamylation and robust kinesin-2 motility requires

detyrosination of α -tubulin (Sirajuddin *et al.*, 2014). Microtubule depolymerization activity of kinesin-13 is regulated by the cytosolic carboxypeptidase CCP-6, which promotes $\Delta 2$ modification of α -tubulin and serves for axon regeneration (Ghosh-Roy *et al.*, 2012).

Furthermore, the microtubule-associated proteins (MAPs), such as MAP2 and tau, are important structural factors that stabilize and bundle the microtubules. They generally give competitive or steric inhibition on kinesin motility (Hagiwara *et al.*, 1994; Dixit *et al.*, 2008; Ebner *et al.*, 1998), which would be another essential factor regulating the axonal transport.

Thus, kinesins may find their destination using multimodal cues, including the identities of their tracks, associated proteins, and cargos.

Concluding Remarks

As has been briefly explored in this article, kinesin molecular motors have been classified into a new category of functional proteins in the past few decades based on investigations using multidisciplinary approaches, including biophysics, biochemistry, molecular and cellular biology, electrophysiology, molecular genetics, and behavioral sciences. Due to the vast mechanochemical properties of their motor domains and due to evolutionary gene shuffling, resulting in the generation of multiple protein-protein interaction domains, kinesins are now regarded as one of the most essential and exquisite motors within living cells. Not only are kinesins regulated by signal transduction cascades, but they also modulate multiple signal transduction pathways by transporting or associating with signaling molecules. Through these mechanisms, kinesins are responsible for multiple human diseases, including cancer, metabolic diseases, neuropathies, and epilepsies, as well as malformations during development and deficits in higher brain function in adults. Thus, kinesin biology provides enormous conceptual advances to various biological concepts and represents a growing field within molecular and cellular biology.

See also: Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cilia and Flagella. Cytoskeleton and Motors: Cytoskeletal Components: Microtubules and Microtubule-Associated Proteins (MAPs)

References

- Aizawa, H., Sekine, Y., Takemura, R., *et al.*, 1992. Kinesin family in murine central nervous system. *Journal of Cell Biology* 119, 1287–1296.
- Arimura, N., Kimura, T., Nakamura, S., *et al.*, 2009. Anterograde transport of TrkB in axons is mediated by direct interaction with Slp1 and Rab27. *Developmental Cell* 16, 675–686.
- Aulchenko, Y.S., Hoppenbrouwers, I.A., Ramagopalan, S.V., *et al.*, 2008. Genetic variation in the KIF1B locus influences susceptibility to multiple sclerosis. *Nature Genetics* 40, 1402–1403.
- Bielska, E., Schuster, M., Roger, Y., *et al.*, 2014. Hook is an adapter that coordinates kinesin-3 and dynein cargo attachment on early endosomes. *Journal of Cell Biology* 204, 989–1007.

- Brady, S.T., 1985. A novel brain ATPase with properties expected for the fast axonal transport motor. *Nature* 317, 73–75.
- Chen, Y., Sheng, Z.H., 2013. Kinesin-1-syntrophin coupling mediates activity-dependent regulation of axonal mitochondrial transport. *Journal of Cell Biology* 202, 351–364.
- Deberg, H.A., Blehm, B.H., Sheung, J., *et al.*, 2013. Motor domain phosphorylation modulates kinesin-1 transport. *Journal of Biological Chemistry* 288, 32612–32621.
- Dixit, R., Ross, J.L., Goldman, Y.E., Holzbaur, E.L., 2008. Differential regulation of dynein and kinesin motor proteins by tau. *Science* 319, 1086–1089.
- Ebneth, A., Godemann, R., Stamer, K., *et al.*, 1998. Overexpression of tau protein inhibits kinesin-dependent trafficking of vesicles, mitochondria, and endoplasmic reticulum: Implications for Alzheimer's disease. *Journal of Cell Biology* 143, 777–794.
- Ghosh-Roy, A., Goncharov, A., Jin, Y., Chisholm, A.D., 2012. Kinesin-13 and tubulin posttranslational modifications regulate microtubule growth in axon regeneration. *Developmental Cell* 23, 716–728.
- Goldstein, L.S., 1993. With apologies to Scheherazade: Tails of 1001 kinesin motors. *Annual Review of Genetics* 27, 319–351.
- Grafstein, B., Forman, D.S., 1980. Intracellular transport in neurons. *Physiological Reviews* 60, 1167–1283.
- Guillaud, L., Setou, M., Hirokawa, N., 2003. KIF17 dynamics and regulation of NR2B trafficking in hippocampal neurons. *Journal of Neuroscience* 23, 131–140.
- Guillaud, L., Wong, R., Hirokawa, N., 2008. Disruption of KIF17-Mint1 interaction by CaMKII-dependent phosphorylation: A molecular model of kinesin-cargo release. *Nature Cell Biology* 10, 19–29.
- Hagiwara, H., Yorifuji, H., Sato-Yoshitake, R., Hirokawa, N., 1994. Competition between motor molecules (kinesin and cytoplasmic dynein) and fibrous microtubule-associated proteins in binding to microtubules. *Journal of Biological Chemistry* 269, 3581–3589.
- Hall, D.H., Hedgecock, E.M., 1991. Kinesin-related gene *unc-104* is required for axonal transport of synaptic vesicles in *C. elegans*. *Cell* 65, 837–847.
- Han, Y.G., Spassky, N., Romaguera-Ros, M., *et al.*, 2008. Hedgehog signaling and primary cilia are required for the formation of adult neural stem cells. *Nature Neuroscience* 11, 277–284.
- Heuston, E., Bronner, C.E., Kull, F.J., Endow, S.A., 2010. A kinesin motor in a force-producing conformation. *BMC Structural Biology* 10, 19.
- Hirokawa, N., 1982. Cross-linker system between neurofilaments, microtubules, and membranous organelles in frog axons revealed by the quick-freeze, deep-etching method. *Journal of Cell Biology* 94, 129–142.
- Hirokawa, N., 1998. Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science* 279, 519–526.
- Hirokawa, N., Niwa, S., Tanaka, Y., 2010. Molecular motors in neurons: Transport mechanisms and roles in brain function, development, and disease. *Neuron* 68, 610–638.
- Hirokawa, N., Pfister, K.K., Yorifuji, H., *et al.*, 1989. Submolecular domains of bovine brain kinesin identified by electron microscopy and monoclonal antibody decoration. *Cell* 56, 867–878.
- Holzbaur, E.L., Scherer, S.S., 2011. Microtubules, axonal transport, and neuropathy. *New England Journal of Medicine* 365, 2330–2332.
- Homma, N., Takei, Y., Tanaka, Y., *et al.*, 2003. Kinesin superfamily protein 2A (KIF2A) functions in suppression of collateral branch extension. *Cell* 114, 229–239.
- Hoogenraad, C.C., Milstein, A.D., Ethell, I.M., Henkemeyer, M., Sheng, M., 2005. GRIP1 controls dendrite morphogenesis by regulating EphB receptor trafficking. *Nature Neuroscience* 8, 906–915.
- Humbert, S., Bryson, E.A., Cordelieres, F.P., *et al.*, 2002. The IGF-1/Akt pathway is neuroprotective in Huntington's disease and involves Huntingtin phosphorylation by Akt. *Developmental Cell* 2, 831–837.
- Hurd, D.D., Stern, M., Saxton, W.M., 1996. Mutation of the axonal transport motor kinesin enhances paralytic and suppresses Shaker in *Drosophila*. *Genetics* 142, 195–204.
- Janke, C., Kneussel, M., 2010. Tubulin post-translational modifications: Encoding functions on the neuronal microtubule cytoskeleton. *Trends in Neurosciences* 33, 362–372.
- Jenkins, B., Decker, H., Bentley, M., Luisi, J., Banker, G., 2012. A novel split kinesin assay identifies motor proteins that interact with distinct vesicle populations. *Journal of Cell Biology* 198, 749–761.
- Kanai, Y., Dohmae, N., Hirokawa, N., 2004. Kinesin transports RNA: Isolation and characterization of an RNA-transporting granule. *Neuron* 43, 513–525.
- Kanai, Y., Wang, D., Hirokawa, N., 2014. KIF13B enhances the endocytosis of LRP1 by recruiting LRP1 to caveolae. *Journal of Cell Biology* 204, 395–408.
- Kikkawa, M., Sablin, E.P., Okada, Y., *et al.*, 2001. Switch-based mechanism of kinesin motors. *Nature* 411, 439–445.
- Kondo, M., Takei, Y., Hirokawa, N., 2012. Motor protein KIF1A is essential for hippocampal synaptogenesis and learning enhancement in an enriched environment. *Neuron* 73, 743–757.
- Kotaja, N., Macho, B., Sassone-Corsi, P., 2005. Microtubule-independent and protein kinase A-mediated function of kinesin KIF17b controls the intracellular transport of activator of CREM in testis (ACT). *Journal of Biological Chemistry* 280, 31739–31745.
- Lawrence, C.J., Dawe, R.K., Christie, K.R., *et al.*, 2004. A standardized kinesin nomenclature. *Journal of Cell Biology* 167, 19–22.
- Lazarov, O., Morfini, G.A., Lee, E.B., *et al.*, 2005. Axonal transport, amyloid precursor protein, kinesin-1, and the processing apparatus: Revisited. *Journal of Neuroscience* 25, 2386–2395.
- Lin, F., Hiesberger, T., Cordes, K., *et al.*, 2003. Kidney-specific inactivation of the KIF3A subunit of kinesin-II inhibits renal ciliogenesis and produces polycystic kidney disease. *Proceedings of the National Academy of Sciences of the United States of America* 100, 5286–5291.
- Macaskill, A.F., Rinholm, J.E., Twelvetrees, A.E., *et al.*, 2009. Miro1 is a calcium sensor for glutamate receptor-dependent localization of mitochondria at synapses. *Neuron* 61, 541–555.
- Marszalek, J.R., Ruiz-Lozano, P., Roberts, E., Chien, K.R., Goldstein, L.S., 1999. Situs inversus and embryonic ciliary morphogenesis defects in mouse mutants lacking the KIF3A subunit of kinesin-II. *Proceedings of the National Academy of Sciences of the United States of America* 96, 5043–5048.
- Martin, M., Iyadurai, S.J., Gassman, A., *et al.*, 1999. Cytoplasmic dynein, the dynactin complex, and kinesin are interdependent and essential for fast axonal transport. *Molecular Biology of the Cell* 10, 3717–3728.
- Mayer, T.U., Kapoor, T.M., Haggarty, S.J., *et al.*, 1999. Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen. *Science* 286, 971–974.
- Meluh, P.B., Rose, M.D., 1990. KAR3, a kinesin-related gene required for yeast nuclear fusion. *Cell* 60, 1029–1041.
- Midorikawa, R., Takei, Y., Hirokawa, N., 2006. KIF4 motor regulates activity-dependent neuronal survival by suppressing PARP-1 enzymatic activity. *Cell* 125, 371–383.
- Miki, H., Setou, M., Kaneshiro, K., Hirokawa, N., 2001. All kinesin superfamily protein, KIF, genes in mouse and human. *Proceedings of the National Academy of Sciences of the United States of America* 98, 7004–7011.
- Misra, C., Ziff, E.B., 2005. EphB2 gets a GRIP on the dendritic arbor. *Nature Neuroscience* 8, 848–850.
- Morfini, G.A., You, Y.M., Pollema, S.L., *et al.*, 2009. Pathogenic huntingtin inhibits fast axonal transport by activating JNK3 and phosphorylating kinesin. *Nature Neuroscience* 12, 864–871.
- Nakagawa, T., Tanaka, Y., Matsuoka, E., *et al.*, 1997. Identification and classification of 16 new kinesin superfamily (KIF) proteins in mouse genome. *Proceedings of the National Academy of Sciences of the United States of America* 94, 9654–9659.
- Nakajima, K., Yin, X., Takei, Y., *et al.*, 2012. Molecular motor KIF5A is essential for GABA(A) receptor transport, and KIF5A deletion causes epilepsy. *Neuron* 76, 945–961.
- Nakata, T., Niwa, S., Okada, Y., Perez, F., Hirokawa, N., 2011. Preferential binding of a kinesin-1 motor to GTP-tubulin-rich microtubules underlies polarized vesicle transport. *Journal of Cell Biology* 194, 245–255.
- Niwa, S., Nakajima, K., Miki, H., *et al.*, 2012. KIF19A is a microtubule-depolymerizing kinesin for ciliary length control. *Developmental Cell* 23, 1167–1175.
- Niwa, S., Tanaka, Y., Hirokawa, N., 2008. KIF1Bbeta- and KIF1A-mediated axonal transport of presynaptic regulator Rab3 occurs in a GTP-dependent manner through DENN/MADD. *Nature Cell Biology* 10, 1269–1279.
- Noda, Y., Niwa, S., Homma, N., *et al.*, 2012. Phosphatidylinositol 4-phosphate 5-kinase alpha (PIP5Kalpha) regulates neuronal microtubule depolymerase kinesin, KIF2A and suppresses elongation of axon branches. *Proceedings of the National Academy of Sciences of the United States of America* 109, 1725–1730.
- Nonaka, S., Tanaka, Y., Okada, Y., *et al.*, 1998. Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. *Cell* 95, 829–837.
- Ogawa, T., Nitta, R., Okada, Y., Hirokawa, N., 2004. A common mechanism for microtubule destabilizers-M type kinesins stabilize curling of the protofilament using the class-specific neck and loops. *Cell* 116, 591–602.
- Okada, Y., Hirokawa, N., 1999. A processive single-headed motor: Kinesin superfamily protein KIF1A. *Science* 283, 1152–1157.
- Okada, Y., Yamazaki, H., Sekine-Aizawa, Y., Hirokawa, N., 1995. The neuron-specific kinesin superfamily protein KIF1A is a unique monomeric motor for anterograde axonal transport of synaptic vesicle precursors. *Cell* 81, 769–780.

- Paschal, B.M., Shpetner, H.S., Vallee, R.B., 1987. MAP 1C is a microtubule-activated ATPase which translocates microtubules *in vitro* and has dynein-like properties. *Journal of Cell Biology* 105, 1273–1282.
- Rath, O., Kozielski, F., 2012. Kinesins and cancer. *Nature Reviews Cancer* 12, 527–539.
- Reid, E., Kloos, M., Ashley-Koch, A., *et al.*, 2002. A kinesin heavy chain (KIF5A) mutation in hereditary spastic paraplegia (SPG10). *American Journal of Human Genetics* 71, 1189–1194.
- Robbins, D.J., Nybakken, K.E., Kobayashi, R., *et al.*, 1997. Hedgehog elicits signal transduction by means of a large complex containing the kinesin-related protein costal2. *Cell* 90, 225–234.
- Rosenbaum, J.L., Witman, G.B., 2002. Intraflagellar transport. *Nature Reviews Molecular Cell Biology* 3, 813–825.
- Saito, N., Okada, Y., Noda, Y., *et al.*, 1997. KIFC2 is a novel neuron-specific C-terminal type kinesin superfamily motor for dendritic transport of multivesicular body-like organelles. *Neuron* 18, 425–438.
- Selkoe, D.J., 2001. Alzheimer's disease: Genes, proteins, and therapy. *Physiological Reviews* 81, 741–766.
- Setou, M., Nakagawa, T., Seog, D.H., Hirokawa, N., 2000. Kinesin superfamily motor protein KIF17 and mLin-10 in NMDA receptor-containing vesicle transport. *Science* 288, 1796–1802.
- Setou, M., Seog, D.H., Tanaka, Y., *et al.*, 2002. Glutamate-receptor-interacting protein GRIP1 directly steers kinesin to dendrites. *Nature* 417, 83–87.
- Shipley, K., Hekmat-Nejad, M., Turner, J., *et al.*, 2004. Structure of a kinesin microtubule depolymerization machine. *EMBO Journal* 23, 1422–1432.
- Sirajuddin, M., Rice, L.M., Vale, R.D., 2014. Regulation of microtubule motors by tubulin isotypes and post-translational modifications. *Nature Cell Biology* 16, 335–344.
- Sisson, J.C., Ho, K.S., Suyama, K., Scott, M.P., 1997. Costal2, a novel kinesin-related protein in the Hedgehog signaling pathway. *Cell* 90, 235–245.
- van Spronsen, M., Mikhaylova, M., Lipka, J., *et al.*, 2013. TRAK/Milton motor-adaptor proteins steer mitochondrial trafficking to axons and dendrites. *Neuron* 77, 485–502.
- Sczesna, E., Kasprzak, A.A., 2012. The C-terminus of kinesin-14 Ncd is a crucial component of the force generating mechanism. *FEBS Letters* 586, 854–858.
- Takeda, S., Yonekawa, Y., Tanaka, Y., *et al.*, 1999. Left-right asymmetry and kinesin superfamily protein KIF3A: New insights in determination of laterality and mesoderm induction by kif3A^{-/-} mice analysis. *Journal of Cell Biology* 145, 825–836.
- Tanaka, Y., Hirokawa, N., 2002. Mouse models of Charcot-Marie-Tooth disease. *Trends in Genetics* 18, S39–S44.
- Tanaka, Y., Kanai, Y., Okada, Y., *et al.*, 1998. Targeted disruption of mouse conventional kinesin heavy chain, kif5B, results in abnormal perinuclear clustering of mitochondria. *Cell* 93, 1147–1158.
- Tanaka, Y., Okada, Y., Hirokawa, N., 2005. FGF-induced vesicular release of Sonic hedgehog and retinoic acid in leftward nodal flow is critical for left-right determination. *Nature* 435, 172–177.
- Tanenbaum, M.E., Macurek, L., van der Vaart, B., *et al.*, 2011. A complex of Kif18b and MCAK promotes microtubule depolymerization and is negatively regulated by Aurora kinases. *Current Biology* 21, 1356–1365.
- Teng, J., Rai, T., Tanaka, Y., *et al.*, 2005. The KIF3 motor transports N-cadherin and organizes the developing neuroepithelium. *Nature Cell biology* 7, 474–482.
- Terada, S., Kinjo, M., Aihara, M., Takei, Y., Hirokawa, N., 2010. Kinesin-1/Hsc70-dependent mechanism of slow axonal transport and its relation to fast axonal transport. *EMBO Journal* 29, 843–854.
- Ueno, H., Huang, X., Tanaka, Y., Hirokawa, N., 2011. KIF16B/Rab14 molecular motor complex is critical for early embryonic development by transporting FGF receptor. *Developmental Cell* 20, 60–71.
- Vale, R.D., Reese, T.S., Sheetz, M.P., 1985. Identification of a novel force-generating protein, kinesin, involved in microtubule-based motility. *Cell* 42, 39–50.
- Verhey, K.J., Hammond, J.W., 2009. Traffic control: Regulation of kinesin motors. *Nature Reviews Molecular Cell Biology* 10, 765–777.
- Verhey, K.J., Kaul, N., Soppina, V., 2011. Kinesin assembly and movement in cells. *Annual Review of Biophysics* 40, 267–288.
- Wang, X., Schwarz, T.L., 2009. The mechanism of Ca²⁺-dependent regulation of kinesin-mediated mitochondrial motility. *Cell* 136, 163–174.
- Wickstead, B., Gull, K., Richards, T.A., 2010. Patterns of kinesin evolution reveal a complex ancestral eukaryote with a multifunctional cytoskeleton. *BMC Evolutionary Biology* 10, 110.
- Wong, R.W., Setou, M., Teng, J., Takei, Y., Hirokawa, N., 2002. Overexpression of motor protein KIF17 enhances spatial and working memory in transgenic mice. *Proceedings of the National Academy of Sciences of the United States of America* 99, 14500–14505.
- Yajima, H., Ogura, T., Nitta, R., *et al.*, 2012. Conformational changes in tubulin in GMPCPP and GDP-taxol microtubules observed by cryoelectron microscopy. *Journal of Cell Biology* 198, 315–322.
- Yamada, K., Andrews, C., Chan, W.M., *et al.*, 2003. Heterozygous mutations of the kinesin KIF21A in congenital fibrosis of the extraocular muscles type 1 (CFEOM1). *Nature Genetics* 35, 318–321.
- Yin, X., Feng, X., Takei, Y., Hirokawa, N., 2012. Regulation of NMDA receptor transport: A KIF17-cargo binding/releasing underlies synaptic plasticity and memory *in vivo*. *Journal of Neuroscience* 32, 5486–5499.
- Yin, X., Takei, Y., Kido, M.A., Hirokawa, N., 2011. Molecular motor KIF17 is fundamental for memory and learning via differential support of synaptic NR2A/2B levels. *Neuron* 70, 310–325.
- Zhao, C., Takita, J., Tanaka, Y., *et al.*, 2001. Charcot-Marie-Tooth disease type 2A caused by mutation in a microtubule motor KIF1Bbeta. *Cell* 105, 587–597.
- Zhou, R., Niwa, S., Guillaud, L., Tong, Y., Hirokawa, N., 2013. A molecular motor, KIF13A, controls anxiety by transporting the serotonin type 1A receptor. *Cell Reports* 3, 509–519.
- Zhou, R., Niwa, S., Homma, N., Takei, Y., Hirokawa, N., 2009. KIF26A is an unconventional kinesin and regulates GDNF-Ret signaling in enteric neuronal development. *Cell* 139, 802–813.

Further Reading

- Hirokawa, N., Nitta, R., Okada, Y., 2009a. The mechanisms of kinesin motor motility: Lessons from the monomeric motor KIF1A. *Nature Reviews Molecular Cell Biology* 10, 877–884.
- Hirokawa, N., Noda, Y., Tanaka, Y., Niwa, S., 2009b. Kinesin superfamily motor proteins and intracellular transport. *Nature Reviews Molecular Cell Biology* 10, 682–696.
- Hirokawa, N., Tanaka, Y., Okada, Y., 2009c. Left-right determination: Involvement of molecular motor KIF3, cilia, and nodal flow. *Cold Spring Harbor Perspectives in Biology* 1, a000802.
- Hirokawa, N., Tanaka, Y., Okada, Y., Takeda, S., 2006. Nodal flow and the generation of left-right asymmetry. *Cell* 125, 33–45.
- Vernos, I. (Ed.), 2001. *Kinesin Protocols*, Vol. 164. Totowa, NJ: Humana Press.

Relevant Website

<http://cb.m.u-tokyo.ac.jp/KIF/>
KIF homepage at the University of Tokyo.

Dyneins

CA Cypranowska and A Yildiz, University of California, Berkeley, CA, USA
T Ishikawa, Paul Scherrer Institute, Villigen, Switzerland

© 2016 Elsevier Inc. All rights reserved.

Glossary

Axoneme A microtubule-based structure found in cilia and flagella. It is composed of nine outer doublets of microtubules (MT) arranged in a ring around two central MT singlets. Inner- and outer-arm dyneins drive bending of the axoneme for cell motility.

Cilia and flagella Cellular organelles composed of >400 proteins. Motile cilia generate bending motion driven by axonemal dyneins. Eukaryotic flagella are almost identical to motile cilia and these two terms are exchangeably referred.

Dynactin A protein complex known to associate with the dynein motor. Dynactin increases the processivity of the dynein motor and controls the localization of dynein in the cell.

Dynein A cytoskeletal motor protein that participates in minus-end-directed transport on MTs by coupling adenosine triphosphate (ATP) hydrolysis to mechanical

work. Dyneins are responsible for trafficking of vesicles and organelles, controlling mitotic spindle orientation, driving retrograde axonal transport, and retrograde intraflagellar transport.

Intraflagellar transport The bidirectional trafficking of proteins from the base of a cilium or flagellum to its distal end and back. This process is important in assembly and maintenance of cilia and flagella.

Mechanochemical cycle The cycle of events during which a molecular motor converts chemical energy, usually in the form of ATP, into mechanical work.

Microtubule A tubular, filamentous cytoskeletal structure assembled of 13 linear protofilaments made of polymerized tubulin. Microtubules provide the substrate for dynein motors.

Power stroke The force generating step in the mechanochemical cycle of a molecular motor.

Introduction

Eukaryotic cells are presented with the problems of intracellular organization, coordination, and locomotion, and use adenosine triphosphate (ATP)-driven molecular motors to overcome these challenges. Dyneins are processive, minus-end-directed microtubule (MT) motors that are involved in essential cellular processes, including axonal transport, mitotic spindle formation, and intraflagellar transport (IFT). Dyneins are evolutionarily distinct from the two other families of cytoskeletal motors, myosins and kinesins, and are organized into two classes: axonemal and cytoplasmic. Cytoplasmic dyneins are responsible for much of the minus-end-directed transport of cargoes both within the cell body and in cilia. Axonemal dyneins play a major role in the bending of axonemes that drive ciliary motility. In this article, we discuss the structure, function, and biophysical mechanics of dynein motors.

Architecture of Dyneins

The dynein holoenzyme is composed of 1–3 dynein heavy chains (DHCs), which contain the necessary components for ATP hydrolysis and MT binding, bound to intermediate (IC), light-intermediate (LIC), and light (LC) chains. ICs, LICs, and LCs are responsible for regulating the motility of the HC and recruiting cargo. Each dynein holoenzyme contains two HCs, two ICs, two LICs, and a variable number of LCs (Vallee *et al.*, 2004). Three types of LCs, known as TCTEX, LC8, and Roadblock, have been identified (Roberts *et al.*, 2013). The dynein holoenzyme can be thought of as a modular structure; a

relatively small selection of HCs is repurposed through the binding of a variety of ICs, LICs, and LCs to accomplish a broad array of cellular tasks. Herein we discuss the structure of the DHC and the insights it provides on dynein function.

DHC

Genes which encode DHCs can be sorted into three classes: cytoplasmic, outer-arm, and inner-arm dyneins.

Cytoplasmic DHCs

Cytoplasmic dynein 1 (also known as DYNC1) is the dynein complex that mediates nearly all minus-end-directed transport in eukaryotic cells. Cytoplasmic dynein 2 (DYNC2, referred to here as IFT dynein) has a strong sequence identity to DYNC1, and is responsible for mediating retrograde IFT. Both IFT dynein and cytoplasmic dynein complexes are composed of a homodimer of the respective HCs (Perrone *et al.*, 2003). The cytoplasmic dynein HC is composed of an N-terminal tail domain, a pseudohexameric AAA+ ATPase domain, and a stalk domain (Figure 1). The tail domain is involved in dimerization of the dynein HCs and recruitment of ICs, LICs, and LCs, and contacts the AAA+ domains through a linker region (Carter *et al.*, 2011; Kon *et al.*, 2011; Schmidt *et al.*, 2012; Tynan *et al.*, 2000). The AAA+ ATPase domains (numbered AAA1–AAA6) are responsible for binding and hydrolyzing ATP to power motility. The stalk region emerges from the motor domain at AAA4 as a 15 nm antiparallel coiled-coil with a globular MT-binding domain (MTBD) at its distal end (Kon *et al.*, 2011).

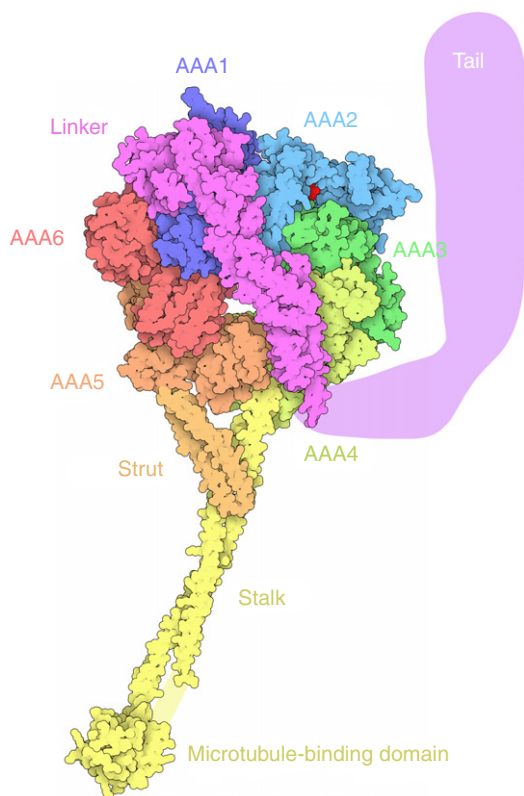


Figure 1 Architecture of the cytoplasmic DHC: the N-terminal tail domain (lavender) is involved in dimerization, the recruitment of ICs, LICs, and adaptor proteins. The linker region (magenta) is docked on the face of the AAA + ring domain at the AAA1 site (dark blue) and the AAA4 site (yellow). The AAA + modules are arranged from AAA1 to AAA6 in a clockwise fashion around the ring. Between AAA4 and AAA5, a ~15 nm coiled-coil stalk emerges with a globular microtubule-binding domain (MTBD) at its distal end. AAA5 contains a short coiled-coil protrusion known as a strut or buttress. The image is courtesy of David S. Goodsell and the RCSB PDB and is available under a CC-BY-3.0 license (<http://creativecommons.org/licenses/by/3.0/us/>).

The AAA + ring domains are part of the ASCE (additional strand catalytic E) superfamily, which contains a diverse array of proteins, including those involved in protein quality control and genome replication (Erzberger and Berger, 2006). The AAA + fold is derived from the core, ancestral ASCE α fold with a C-terminal α -helical domain (Erzberger and Berger, 2006). These two subdomains, often called the large and small subdomains, contain a flexible hinge region between them which allows a rigid-body motion between the large and small domains (Glynn *et al.*, 2009). Biochemical and structural studies suggest that this rigid-body motion communicates nucleotide-dependent conformational changes between AAA + domains (Carter *et al.*, 2011; Glynn *et al.*, 2012; Stinson *et al.*, 2013).

AAA + domains assemble as an asymmetric hexameric ring structure (Erzberger and Berger, 2006). However, dynein is one of two proteins whose AAA + domains are contained within a single polypeptide chain and the sequence and structure of each AAA + domain is unique (Neuwald *et al.*, 1999). The

canonical AAA + fold contains several motifs associated with nucleotide-triphosphate binding, including sensor-I, sensor-II, arginine finger, Walker A, and Walker B motifs (Figure 2). Only AAA1–4 have conserved Walker A and B motifs for catalytic activity. Of these, only AAA1, AAA3, and AAA4 have meaningful ATPase activities (Cho *et al.*, 2008; Kon *et al.*, 2004). Although AAA2 lacks the sensor-II motif and the catalytic glutamate residue of the Walker B motif, structural studies have consistently found electron density from nucleotide binding in the AAA2 pocket (Kon *et al.*, 2012; Schmidt *et al.*, 2012). This evidence suggests that the nucleotide-binding activity of AAA2 serves a structural role. AAA5 and AAA6 have lost their Walker A and B motifs and fail to binding nucleotide in crystallographic studies (Kon *et al.*, 2012; Neuwald *et al.*, 1999; Schmidt *et al.*, 2012).

The stalk domain is an antiparallel coiled-coil that emerges from and returns to the α -helical domain of AAA4 (Kon *et al.*, 2011). The stalk region is a unique feature of dynein; kinesins and myosins have globular head domains that both hydrolyze ATP and bind and release their cytoskeletal track (Kull *et al.*, 1996; Rayment *et al.*, 1993; Woehlke *et al.*, 1997). The coiled-coil stalk is kinked by two highly conserved proline residues near the MTBD (Carter *et al.*, 2008). These proline residues are thought to disturb the heptad-repeat pattern and packing of hydrophobic residues of the coiled-coil (Carter *et al.*, 2008; Nishikawa *et al.*, 2014). The MTBD is composed of a bundle of six α -helices (Carter *et al.*, 2008). Helices H1, H3, and H6 are enriched with positively charged residues and forms the MT-binding interface (Carter *et al.*, 2008; Koonce and Tikhonenko, 2000). Altering the registry of the coiled-coil stalk causes the MTBD to adopt a high-affinity (α) or low-affinity ($\pm\beta$) conformation (Gibbons *et al.*, 2005; Kon *et al.*, 2009). The α conformation of the MTBD is hallmarked by basic residues forming salt bridges with glutamate residues on the surface of the tubulin dimer (Redwine *et al.*, 2012). Moreover, these electrostatic interactions are satisfied by intramolecular salt-bridges within the MTBD in the low-affinity conformation (Redwine *et al.*, 2012).

Dynein Motility

Dynein Stepping Is Uncoordinated

Dynein stepping behavior differs significantly from other cytoskeletal motors. Myosin V and kinesin-1 walk hand-over-hand, in which the trailing head releases from the motor's substrate and advances in the forward direction past the leading head (Yildiz *et al.*, 2003, 2004). Unlike these motors, dynein's step size is highly variable, with most steps being 8–16 nm in size toward the MT minus-end, as well as ~20% steps in backward direction (DeWitt *et al.*, 2012; Qiu *et al.*, 2012; Reck-Peterson *et al.*, 2006). Dynein moves processively by uncoordinated stepping of its two heads (DeWitt *et al.*, 2012; Qiu *et al.*, 2012; Reck-Peterson *et al.*, 2006). The tethered excursion mechanism suggested that each head is able to release from MT and move forward while the partner head serves as a tether to prevent dissociation from the track. Although dynein's stepping is uncoordinated under most conditions, the probability of the trailing head taking a step is highly

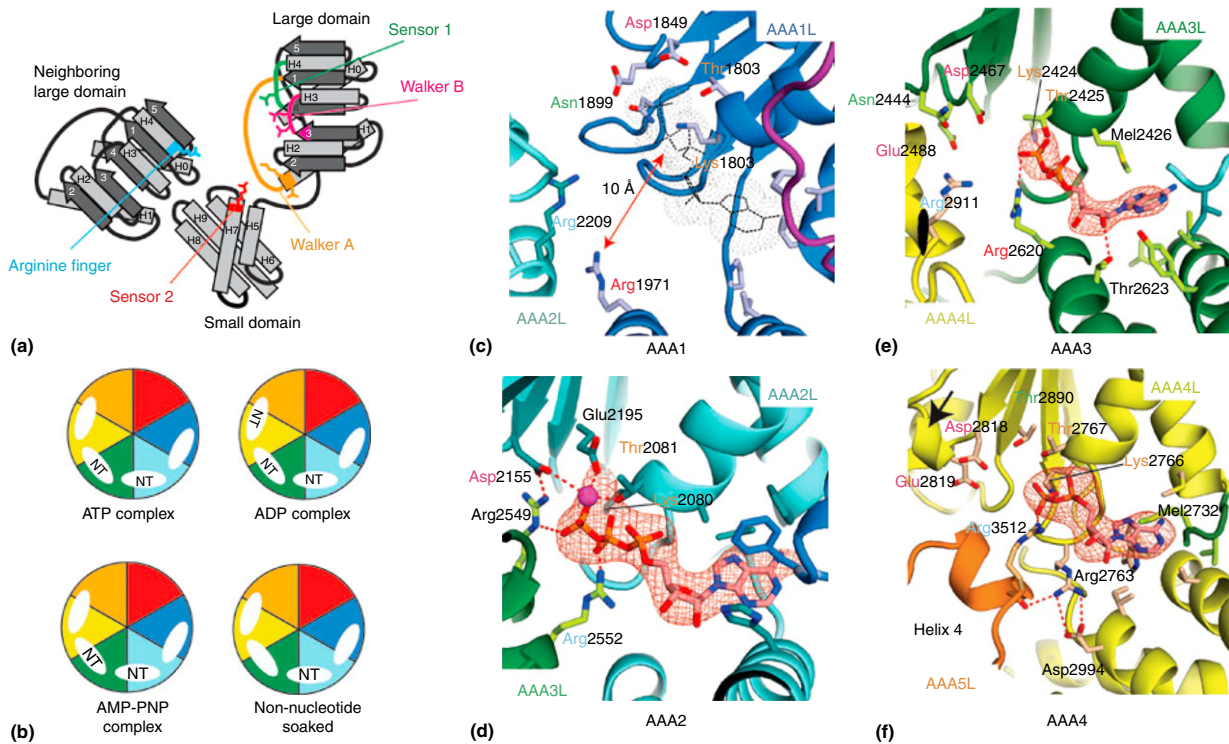


Figure 2 Details of the crystal structure of cytoplasmic dynein's motor domain. (a) Topology of the AAA+ domains in the cytoplasmic dynein motor domain. The Sensor 1 (green), Walker B (pink), Walker A (orange), and Sensor 2 (red) motifs are shown on the large and small domains of the complete AAA+ domain shown. The arginine finger (blue) of the neighboring large domain interacts with the small domain. (b) Diagram of nucleotide binding to the AAA+ ring of cytoplasmic dynein for non-soaked crystals and crystals soaked in ATP, adenosine diphosphate (ADP), and AMP-PNP. AAA1 (blue), AAA2 (light blue), AAA3 (green), AAA4 (yellow), AAA5 (orange), and AAA6 (red) are shown with nucleotide-binding pockets represented by white ovals. NT indicates the presence of nucleotide in the indicated binding pocket. (c–f) Details of the nucleotide-binding pocket. (c) The binding cleft of AAA1 from the ATP-soaked crystal has an open conformation, with the putative ATP-binding site indicated in gray. (d) The AAA2 subunit binds tightly to ATP (ball-and-stick representation with orange mesh). Magnesium ion density is shown in magenta. (e) The AAA3 subunit from the ADP-soaked crystal is shown bound to an exchangeable ADP molecule (ball-and-stick representation with orange mesh). (f) The AAA4 subunit is bound to ADP in the ADP-soaked crystal structure. The catalytic glutamate residue of the Walker B motif lies at the end of a short α -helical fragment (black arrow). Reprinted by Schmidt, H., Gleave, E.S., Carter, A.P., 2012. Insights into dynein motor domain function from a 3.3-Å crystal structure. *Nature Structural & Molecular Biology* 19, 492–497. © 2012, with permission from Macmillan Publishers Ltd.

dependent on the separation between the leading and trailing heads, suggesting that interhead tension plays an important role in dynein motility (Cleary *et al.*, 2014; DeWitt *et al.*, 2012). Dynein has also been shown to take sideways steps onto neighboring protofilaments of the MT, and walk in a helical path around the MT surface (Can *et al.*, 2014; Reck-Peterson *et al.*, 2006).

The AAA+ Domains Power Dynein Motility

The DHC contains the necessary components for motility, including the AAA+ ring, which contains conserved motifs for ATP binding and hydrolysis. Because ATPase and MTBDs are segregated from each other in dynein, this brings into question how ATP hydrolysis at ATPase domains is coupled to MT release. The ATPase activity of AAA1 is strictly required for motility and loss of ATP hydrolysis at AAA3 resulted in a greatly reduced velocity compared to wild-type dynein (Kon *et al.*, 2004). Abolishing nucleotide binding at AAA2 reduced the velocity, and AAA2 in the nucleotide-bound state can

propagate conformational changes from AAA1 to the other AAA+ domains via rigid-body motion (Roberts *et al.*, 2013). Loss of AAA3 function results in a >10-fold decrease in velocity and a ~2-fold reduction in stall force, whereas mutation of AAA4 doubles the motor's run length (Cho *et al.*, 2008). LIS1, a dynein adaptor protein that binds at the AAA3–AAA4 junction, increases the time that dynein spends strongly bound to the MT in between steps, indicating that these domains participate in communication between AAA1 and the MTBD (Huang *et al.*, 2012). The ATP hydrolysis cycles of AAA1 and AAA3 are uncoordinated and that the post-hydrolysis state of AAA3 site facilitates communication between AAA1 and the MTBD (Bhabha *et al.*, 2014; Dewitt *et al.*, 2015).

Communication of the nucleotide state of AAA1 is propagated by rigid-body motions of the other AAA+ domains (Schmidt *et al.*, 2014). When ATP is bound and hydrolyzed at AAA1, the gap between the AAA1 and AAA2 sites closes. Closure of this site triggers the buttress, a short coiled-coil projection from AAA5, to slide the stalk's coiled-coil helices with respect to each other (Schmidt *et al.*, 2014). This ultimately renders the MTBD in a low-affinity state, allowing the

motor to release from the MT during its stepping cycle (Schmidt *et al.*, 2014).

Force Generation

The force generation cycles of other cytoskeletal motors depend on ATP binding or hydrolysis activity to drive the forward movement by rotating a lever arm-like mechanical element (Vale and Milligan, 2000). Dynein, on the other hand, is thought to operate by a winch mechanism, where the power stroke is used to reduce the distance between dynein and its cargo (Carter *et al.*, 2008; Movassagh *et al.*, 2010; Ueno *et al.*, 2008). The first step of the dynein mechanochemical cycle involves binding of ATP at the AAA1 site, which triggers the release of the MTBD from the MT (Figure 3; Roberts *et al.*, 2013). During ATP hydrolysis at AAA1, the gap between AAA1 and AAA2 closes and forces the linker into a bent conformation (Roberts *et al.*, 2012; Schmidt *et al.*, 2014). After a diffusional search for a new MT-binding site in the minus-end direction, the linker undergoes a second conformational change, known as the power stroke (Figure 3; Roberts *et al.*, 2013). During the power stroke, the linker returns to its original conformation and docks near the AAA5 site (Roberts *et al.*, 2012; Schmidt *et al.*, 2012). Each motor head in the dynein dimer produces half the stall force of the dimer (Belyy *et al.*, 2014). Therefore, dynein motor heads share the load, which can produce motility against high-hindering forces, and

explains how multiple dyneins attached to a single cargo can team together more efficiently (Derr *et al.*, 2012).

Force production of dynein per mechanochemical cycle varies from organism to organism. Optical trap studies showed that the motility of mammalian dynein stalls at resistive forces of ~ 1.1 pN resistive forces (Hendricks *et al.*, 2012; Mallik *et al.*, 2004; McKenney *et al.*, 2010; Rai *et al.*, 2013; Schroeder *et al.*, 2010). In contrast, cytoplasmic dynein from porcine brain and *Saccharomyces cerevisiae* have 3–7 pN stall forces (Belyy *et al.*, 2014; Gennerich *et al.*, 2007; Toba *et al.*, 2006; Walter *et al.*, 2010). Estimates from *in vivo* studies in *Drosophila melanogaster* and *Reticulomyxa* indicate that dynein's stall force is between 2 and 5 pN (Ashkin *et al.*, 1990; Shubeita *et al.*, 2008). Resistive loads increase the frequency of backward stepping and decrease the frequency of forward steps larger than 12 nm (Gennerich *et al.*, 2007). Resisting forces greater than the stall force resulted in plus-end-directed stepping (Gennerich *et al.*, 2007). The rate of backward stepping was independent of ATP concentration, in contrast to the rate of forward stepping under assisting load (Gennerich *et al.*, 2007).

Intramolecular tension plays a role in coordination of the stepping motility of cytoskeletal motors (Guydosh and Block, 2006; Spudich, 2006; Yildiz *et al.*, 2008). A dynein dimer may experience tension when the two motor heads are separated by a large distance. Tension on the linker inhibits ATP-induced release of the motor from the MT, and dynein exhibits partially coordinated stepping at high interhead separations (Cleary *et al.*, 2014; DeWitt *et al.*, 2012). Tension on the linker may

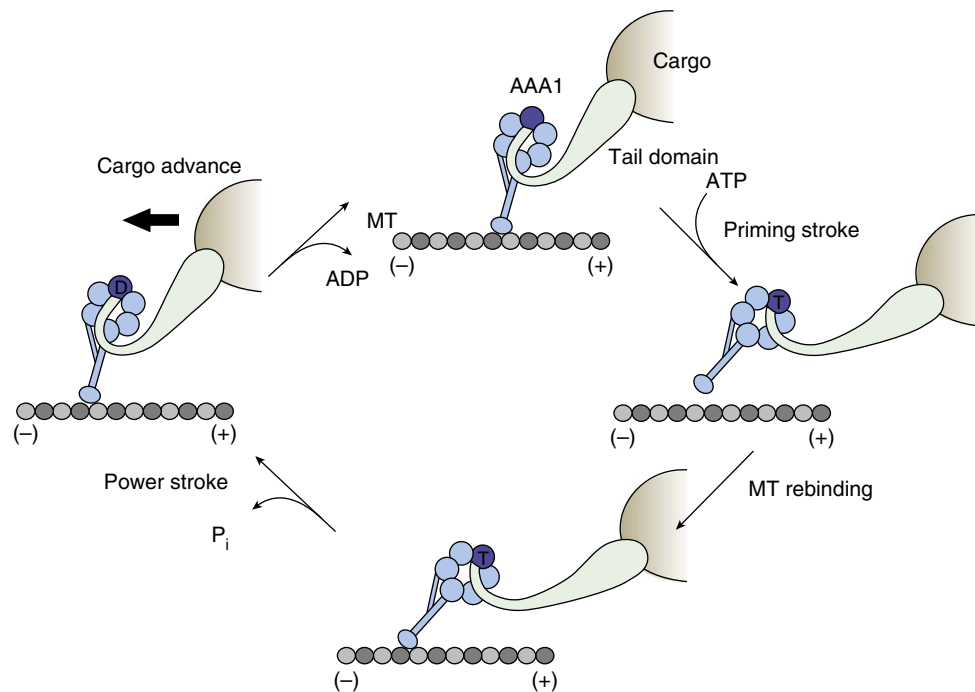


Figure 3 Model for the cytoplasmic dynein mechanochemical cycle and power stroke. The AAA1 ATP-binding pocket is in the apo state and the linker region is in the unprimed conformation at the beginning of the cycle. ATP binding at AAA1 triggers release of the MTBD from the MT. The linker undergoes the priming stroke and adopts a bent conformation, after which the MTBD advances toward the minus-end after a diffusional search for a new binding site. ATP is then hydrolyzed, inorganic phosphate is released, and the linker undergoes the power stroke to pull the bound cargo toward the MT minus-end. Cytoplasmic dynein returns to the beginning of its mechanochemical cycle after the linker returns to the unprimed position and ADP is released.

disrupt contacts between the linker and the AAA+ ring domain, as mutations to the linker-AAA+ ring interface reduce the ATPase activity and the velocity of the motor (Kon *et al.*, 2012; Schmidt *et al.*, 2012).

The Mechanism of Dynein Directionality

The original model for dynein's minus-end directionality came from a cryo-EM study of outer-arm dynein, which posited that movement of the N-terminal tail domain is responsible for minus-end-directed movement (Burgess *et al.*, 2003). It has been proposed that the priming stroke, during which the linker adopts a bent shape across the face of the AAA+ ring, allows the unbound head to push against the MT-bound head (Roberts *et al.*, 2009). However, the force generation occurs in the MT-bound state of the motor, and the priming stroke does not produce substantial amount of work (Belyy *et al.*, 2014). In addition, the reversal of dynein's power stroke by heptad insertions in the coiled-coil stalk did not reverse directionality (Carter *et al.*, 2008). These observations suggested that dynein directionality is determined by an element outside the AAA+ ring and the linker (Carter *et al.*, 2008).

Regulation of Dynein Motility

Because dynein is the only minus-end-directed motor, excluding kinesin-14, many adaptor proteins exist to repurpose the cytoplasmic DHC for different cellular processes. Dynactin, a 1.2 MDa protein complex composed of 11 different proteins, is required for all dynein-mediated processes in all cell types (Schroer, 2004). Most other proteins that regulate dynein motility are used for specific purposes, such as controlling mitotic spindle orientation, recruitment of dynein to the kinetochore, and trafficking organelles and mRNAs (Kardon and Vale, 2009).

Dynactin

Dynactin was initially identified as a factor that promoted trafficking of vesicles mediated by cytoplasmic dynein (Schroer, 2004). Dynactin is composed of the Arp1 rod, which contains Arp1, Arp11, actin, CapZ, p62, p27, and p25, along with p150^{glued}, dynamitin, and p24/p22 (Schroer, 2004). P150^{glued} is the largest component of the dynactin complex, and contains a cytoskeleton-associated protein, glycine-rich (CAP-Gly) domain (Schroer, 2004). The CAP-Gly domain of p150^{glued} has been shown to bind MTs *in vivo* and *in vitro* (Vaughan *et al.*, 2002; Waterman-Storer *et al.*, 1995). It has been proposed that the MTBD present in p150^{glued} acts as an extra tether to the MT (Waterman-Storer *et al.*, 1995).

Features of the dynein-dynactin complex are enhanced processivity and bidirectional motility (King and Schroer, 2000; Ross *et al.*, 2006). *In vitro* reconstitution has shown that dynactin, along with BICD2, is strictly required for robust, processive motion of human cytoplasmic dynein (McKenney *et al.*, 2014; Schlager *et al.*, 2014). Single-molecule fluorescence experiments have confirmed that dynactin increases yeast dynein processivity by ~2-fold. However, deletion of the CAP-

Gly domain of p150^{glued} did not alter the processivity *in vitro* or *in vivo* (Kardon *et al.*, 2009). In contrast, another group found that a removal or mutation of the CAP-Gly domain of p150^{glued} prevented the dynein-dependent translocation of the mitotic spindle and nucleus into the bud neck of *S. cerevisiae* yeast, and decreased the frequency of spindle oscillations (Moore *et al.*, 2009). The mechanistic details of dynactin's effect on dynein processivity remain unknown and highly controversial.

Dynactin also plays a role in recruiting cargo to the cytoplasmic dynein holoenzyme. Dynactin binds EB1 to target dynein to the plus-ends of MTs (Kardon and Vale, 2009). Dynactin's Arp1 also interacts with β III spectrin found on the surface of the Golgi (Kardon and Vale, 2009). P150^{glued} has also been shown to interact with SEC23, to bind vesicles emerging from the endoplasmic reticulum (ER) (Watson *et al.*, 2005). Thus, dynactin not only regulates the dynein HC's motility, but also adapts it for multiple functions within the cell.

Lissencephaly 1 and Nuclear Distribution E

The *LIS1* gene's namesake is a condition known as lissencephaly, which causes the gyri and sulci of the cerebral cortices to take on a smoothed appearance (Dobyns *et al.*, 1993). Lis1 is homologous to the *Aspergillus nidulans* protein NudF (Efimov and Morris, 2000). Loss of NudF function causes nuclear migration defects in this filamentous fungus species (Xiang *et al.*, 1994). Lis1/NudF interacts both physically and genetically with NudE (Efimov and Morris, 2000). Lis1-NudE, and its mammalian homolog LIS1-NudE-like (protein) (NUDEL), have been shown to be instrumental in recruiting dynein to the kinetochore, transport of mRNAs, and positioning of the mitotic spindle (Kardon and Vale, 2009). Moreover, both LIS1 and NUDEL are required for normal transport of organelles in the axons of mammalian neuronal cells (Pandey and Smith, 2011).

Lis1 binds to the AAA3 and AAA4 domains, both of which have been previously implicated in controlling speed and processivity of the motor (Huang *et al.*, 2012). Lis1/LIS1 binding renders dynein in a strongly bound MT state, while disengaging the ATPase activity at AAA1 from MT release (Huang *et al.*, 2012; McKenney *et al.*, 2010).

Functions of Cytoplasmic Dynein *In Vivo*

Transport of Cellular Cargoes

Cytoplasmic dynein can transport a large array of cargoes within a cell, including mitochondria, lysosomes, phagosomes, melanosomes, peroxisomes, vesicles emerging from the ER, lipid droplets, and endosomes (Roberts *et al.*, 2013). In recent years, dynein has been shown to transport more exotic cargo, such as transcription factors and viral particles (Dodding and Way, 2011; Harrell *et al.*, 2004). The cytoplasmic dynein holoenzyme even plays roles in specialized cells, including trafficking of MTs into the immunological synapse of T cells and growth cone of neuronal precursor cells (Combs *et al.*, 2006; Grabham *et al.*, 2007).

mRNA Transport

Bicaudal D (BicD) is a dynein adaptor protein identified in *D. melanogaster* embryos. Mutation of the BicD gene results in a lack of anterior structures in *D. melanogaster*, suggesting this gene's role in developmental patterning (Mohler and Wieschaus, 1986; Steward and Nüsslein-Volhard, 1986). BicD and Egalitarian (Egl) are required for dynein-dependent transport and localization of mRNAs during *D. melanogaster* oogenesis and embryogenesis (Delanoue and Davis, 2005; Mach and Lehmann, 1997; Navarro *et al.*, 2004). Egl binds to mRNA localization elements, in addition to interacting with the cargo-binding domain of BicD (Dienstbier *et al.*, 2009). The localization of an mRNA is dependent of the number of copies of Egl/BicD-bound dynein complexes recruited to the mRNA molecule (Bullock *et al.*, 2006). Outside of mRNA transport, BicD also promotes transport of clathrin-coated vesicles in *D. melanogaster*, and its metazoan homolog BICD-1 is involved in nuclear migration in *Caenorhabditis elegans* (Fridolfsson *et al.*, 2010; Li *et al.*, 2010).

Axonal Transport

Axonal transport describes the phenomenon in which organelles, vesicles, proteins, and other cargoes are transported down the axon of neuronal cells. Kinesin-1 and cytoplasmic dynein are the workhorses of axonal transport machinery, and dysfunction of these motors and their accessory proteins results in neurodevelopmental and neurodegenerative disease (Maday *et al.*, 2014). Anterograde transport describes the trafficking of cargoes from the neuron cell body out to the axon termini, while retrograde transport removes aging proteins and organelles from the axon termini for recycling at the cell body (Maday *et al.*, 2014). Dynein has also been discovered to be involved in the bidirectional transport of late endosomes, lysosomes, neurofilaments, and autophagosomes (Maday *et al.*, 2014). Dynein and dynactin have also been implicated in establishing the uniform polarity of MTs in the axon (Ahmad *et al.*, 1998). Dynein and dynactin also participate in the injury-induced signaling complex that transports transcription factors, such as pSTAT3, from the axon back to nucleus so that the axon can be regenerated (Rishal and Fainzilber, 2014).

Nuclear Migration and Cell Motility

Clues about cytoplasmic dynein's role in nuclear migration first came from the identification of NudE in the filamentous, multinucleated fungi, *A. nidulans*. The LIS1–NUDEL complex is known to couple the migration of the nucleus and centrosome (also known as the MT organization center) in nascent neurons during corticogenesis (Shu *et al.*, 2004; Tanaka *et al.*, 2004; Tsai *et al.*, 2007). LIS1–NUDEL plays a similar role in directed cell movement of fibroblasts (Dujardin *et al.*, 2003; Palazzo *et al.*, 2001). Cytoplasmic dynein maintains the position of the centrosome at the cell centroid during the re-orientation of motile fibroblasts, and inhibition of dynein–dynactin results in reduced cell motility (Dujardin *et al.*, 2003; Gomes *et al.*, 2005). Additionally, cytoplasmic dynein drives nuclear rotation in this cell type (Levy and Holzbaur, 2008).

Mitosis and Spindle Assembly

Cytoplasmic dynein is essential for the morphology and assembly of the mitotic spindle. Studies using *Xenopus laevis* egg extracts revealed that cytoplasmic dynein interacts with nuclear mitotic apparatus (NuMA) to focus the minus-ends of the mitotic spindle poles (Heald *et al.*, 1996; Merdes *et al.*, 1996). Cytoplasmic dynein, dynactin, and NuMA affect spindle morphology by regulating the length of the mitotic spindle and MT depolymerization at the minus-end (Gaetz and Kapoor, 2004). The mechanism of MT organization at the mitotic spindle is unknown, but there is strong evidence to suggest that cytoplasmic dynein plays an important role. Fusion of preassembled mitotic spindles depends on dynein-driven sliding of antiparallel MTs (Gatlin *et al.*, 2009). Dynein also counters the kinesin-5 driven outward sliding of antiparallel MTs to maintain the proper tension and length of the spindle, in addition to driving the MT severance response of the mitotic spindle (Elting *et al.*, 2014; Mitchison *et al.*, 2005; Tanenbaum *et al.*, 2008).

Cytoplasmic dynein is also required for the appropriate segregation of chromosomes during mitosis and spindle assembly checkpoint (SAC) inactivation. Kinetochore-associated dynein has been shown to maintain kinetochore orientation and stabilize kinetochore MTs (Figure 4; Varma *et al.*, 2008; Yang *et al.*, 2007). Cytoplasmic dynein is recruited to the kinetochore by the Rod–Zw10–Zwilch (RZZ) complex, which also serves to recruit mitotic arrest deficient 1 (MAD1), MAD2, and Spindly (Kardon and Vale, 2009). RZZ only interacts with a specific phosphorylated epitope of cytoplasmic dynein, which only occurs before chromosomes are aligned at the metaphase plate (Figure 4; Kardon and Vale, 2009). Once bioriented kinetochore attachment to the mitotic spindle is achieved, dynein transports MAD1, MAD2, and Spindly toward the spindle poles to inactivate the SAC (Figure 4; Howell *et al.*, 2001).

Dynein also helps orient astral MTs on the cell cortex. Ric-8A and Gi α recruit a ternary complex of dynein, LGN, and NuMA to the cell cortex (Woodard *et al.*, 2010), but the spindle position and orientation are regulated by pole-localized polo-like kinase 1 (Plk1) and RanGTP (Kiyomitsu and Cheeseman, 2012). Plk1 negatively regulates cortical dynein positioning by dissociating the dynein–dynactin complex from the LGN–NuMA complex, while the chromosome-derived RanGTP gradient restricts LGN–NuMA binding to lateral cell cortex (Kiyomitsu and Cheeseman, 2012). Dynein's role in asymmetric cell division and stem cell differentiation remains unknown, but Lis1 has been implicated in asymmetric cell division of *D. melanogaster* neuroblasts (Siller and Doe, 2008).

IFT Dynein

IFT is crucial for the maintenance of cilia and flagella and facilitates the turnover of ciliary and flagellar proteins (Roberts *et al.*, 2013). IFT machinery traffics cargo from the tip to the base of the cilia in repetitive multiprotein complexes, known as 'IFT trains' (Roberts *et al.*, 2013). Retrograde transport of IFT trains is driven by IFT dynein (Pazour *et al.*, 1998, 1999; Porter *et al.*, 1999), which has strong sequence similarity to that of cytoplasmic dynein (Gibbons *et al.*, 1994). IFT trains are

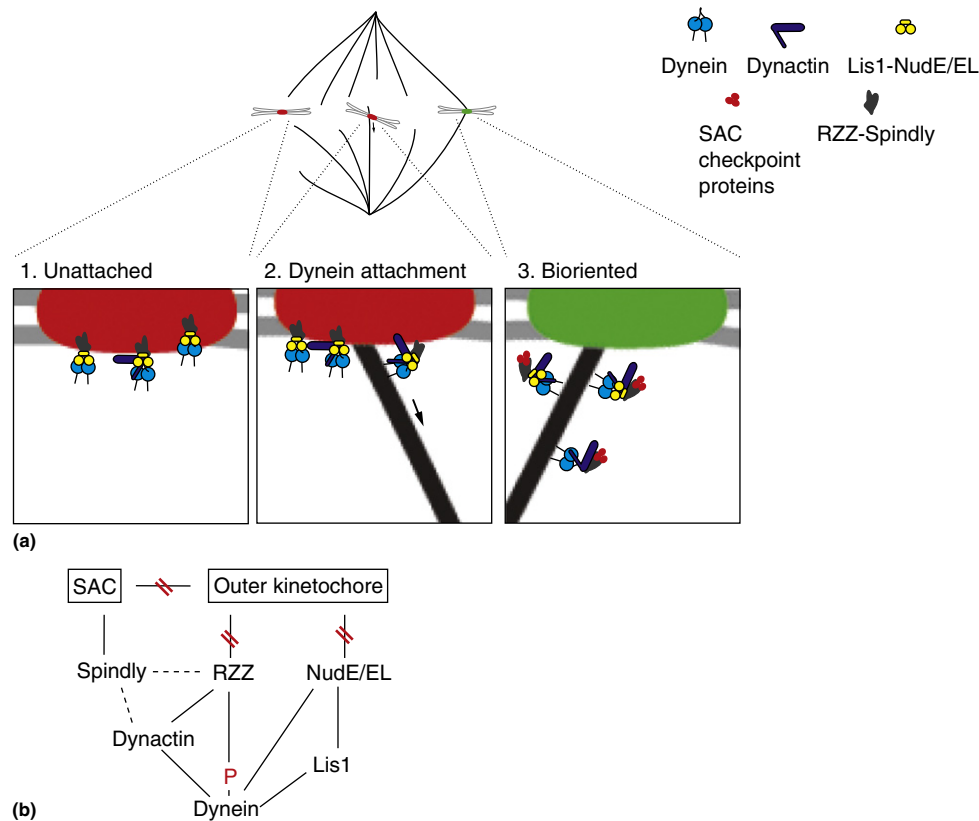


Figure 4 Function of cytoplasmic dynein and its adaptors at the mitotic kinetochore. (a) Geometry of the mitotic spindle showing how dynein interacts with the kinetochore when the chromosome is unattached to any MTs, attached to one MT in an end-on fashion, and attached to two bioriented MTs. In end-on MT attachment, dynein, bound to LIS1, NUDE/NUDEL, dynactin, and RZZ, exerts tension on a captured astral MT. Dynein traffics SAC proteins away from the kinetochore after bioriented MT attachment is achieved. (b) Schematic of the interactions between dynein, LIS1, NUDE/NUDEL, RZZ, Spindly, and dynactin at the kinetochore. NUDE/NUDEL, RZZ, and SAC proteins directly interact with the kinetochore. Dynein has direct interactions with dynactin, LIS1, RZZ, and NUDE/NUDEL. RZZ only directly interacts with dynein when it is phosphorylated immediately before the alignment of chromosomes at the metaphase plate.

transported by multiple dynein motors (Laib *et al.*, 2009; Shih *et al.*, 2013) at 1–3.5 $\mu\text{m s}^{-1}$ (Signor *et al.*, 1999). IFT dynein also drives whole cell gliding motility in *Chlamydomonas reinhardtii* (Shih *et al.*, 2013). IFT dynein has been shown to have roles outside of IFT, including trafficking signaling proteins involved in vertebrate development. Loss of IFT dynein function at the primary cilium has been associated with disruption of Smo localization and Gli2 trafficking and activation, establishing a link between proper IFT dynein function to Hedgehog-dependent developmental patterning (Kim *et al.*, 2009; May *et al.*, 2005).

(Lawrence *et al.*, 2001) (higher plants have no cytoplasmic dynein, either), *Cryptosporidium*, *Phaeodactylum*, *Caenorhabditis*, *Saccharomyces*, and *Schizosaccharomyces* (they have only cytoplasmic dynein). Interestingly, *Ostreococcus* has axonemal dynein (dynein f) (Wickstead and Gull, 2007), but no cilia. There are variety of motion in cilia from different species and under different conditions – symmetric, asymmetric, and, in multiciliated cells and tissues, metachronal (wavy and sequential). Characterization of axonemal dyneins and their interaction with other components are essential to reveal the mechanism of ciliary motion.

Axonemal Dyneins: Overall Arrangement and Structure

Axonemal dynein is a family of MT-based ATPase motor proteins, which drive the axoneme, a bending cellular appendage in motile cilia and flagella (we use the term cilia in this article) in eukaryotes, from protozoa to vertebrates. Axonemal dynein does not exist in primary cilia. Wide varieties of species have motile cilia in organs such as tracheae, brains, oviducts and sperm, and therefore axonemal dyneins. Examples of species which lack axonemal dynein are higher plant

Molecular Arrangement in the Axoneme

The arrangement of axonemal dyneins in cilia is highly conserved among species – from unicellular algae to human. Motile cilia from most species have a common 9 + 2 axoneme structure, in which nine microtubule doublets (MTD) surround two single MTs. There are also 9 + 0 motile cilia, such as eel sperm (Gibbons *et al.*, 1983, 1985), nodal cilia (Basu and Brueckner, 2008), and centric diatom (*Thalassiosira*) flagella (Armbrust *et al.*, 2004). Exceptionally there are species which have

axonemes with more than nine MTDs, for example, the cecidomyid dipteran *Monarthropalpus flavus* (Lupetti *et al.*, 2005).

Each MTD has 96 nm periodic units. The structure based on cryo-electron tomography of *Chlamydomonas* flagella is shown in Figure 5. Within a periodic unit, there are outer dynein arms (ODA), inner dynein arms (IDA), radial spokes (RS), which connect MTD and central pair singlet MTs (CP), dynein regulatory complexes (DRC), which connect adjacent MTDs, and other components. Both ODA and IDA consist of

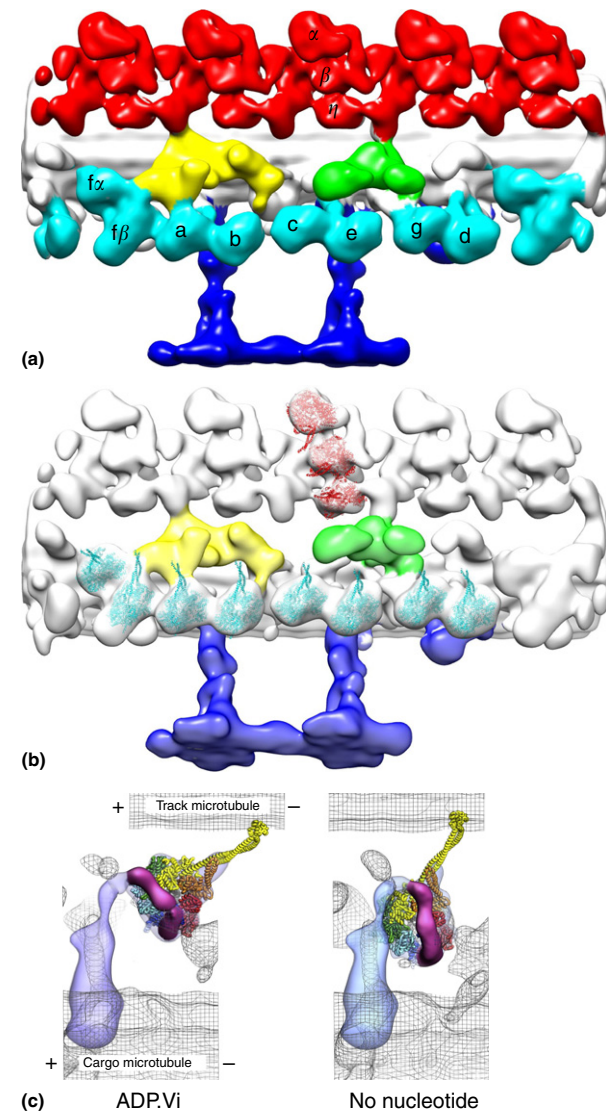


Figure 5 3D structure and dynein arrangement in the axoneme. (a) Periodic unit of MTD, out of nine, from *Chlamydomonas* flagella. Red, ODA; Cyan, IDA; Blue, RS; Yellow, IC/LC of dynein f; Green, DRC. The structure is available in the EM Databank as EMD2117 (Bui *et al.*, 2012). (b) Atomic models of cytoplasmic dynein heads (Kon *et al.*, 2012) are fitted to the cryo-ET structure. Cyan, IAD; Red, OAD. (c) Conformational change of axonemal dynein, modeled by fitting the atomic model to cryo-EM structure (Roberts *et al.*, 2012). Modified from Roberts, A.J., Malkova, B., Walker, M.L., *et al.*, 2012. ATP-driven remodeling of the linker domain in the dynein motor. Structure 20, 1670–1680.

components with various sizes: HCs (HC) (400–500 kDa), IC (45–140 kDa), and light chains (LC) (8–28 kDa).

An ODA consists of two or three outer-arm dyneins (OAD), forming bilobed or trilobed structure, and make an array with 24 nm repeat. Unikonts (single flagellum) have two HC in the ODA, while bikonts have three. These OAD are named α , β , and γ . In the two-headed ODA, innermost and outermost dyneins are called the α - and β -families, respectively. In the three-headed ODA, except *Chlamydomonas*, outermost is called γ -dynein. By historical reasons, only in *Chlamydomonas* the innermost dynein is named gamma and the outermost alpha (Figure 5(a)). No homologue of *Chlamydomonas* α -dynein has been found in human by sequence search (Altschul *et al.*, 1990).

There are normally eight inner-arm dyneins (IAD) within one 96 nm unit, longitudinally along MTD. In *Chlamydomonas* flagella, loci of eight major IAD are known: from the proximal end to the distal end, dynein $f\beta$, dynein $f\alpha$, dynein a, dynein b, dynein c, dynein e, dynein g, and dynein d (Bui *et al.*, 2012; Figure 5(a)). Although sequence homology gives us some hints, location of inner dyneins from other organisms is still to be analyzed. In both ODA and IDA, dyneins link two adjacent MTDs. The N-terminal tails are anchored permanently on the A-tubule, while the MTBD interacts with the B-tubule of the adjacent MTD. The N-terminal tail extends from the AAA-ring toward the distal end of the axoneme (plus end of the MTD) (Bui *et al.*, 2008; Figure 5(b)).

Nucleotide-Induced Conformational Change

By fitting atomic models (Kon *et al.*, 2012) to single-particle cryo-EM analysis of IAD dynein c, it was shown that rearrangement of the linker domain takes place during the power stroke of dynein (Roberts *et al.*, 2012). At the post-power stroke conformation solved by crystallography (Kon *et al.*, 2012; Schmidt *et al.*, 2012), the linker spans over AAA4. As proved by single-particle EM (Roberts *et al.*, 2012), the linker changes its conformation and extends over AAA2 and AAA3 (Figure 5(c)). This result was supported by tomography of the axoneme as well (Lin *et al.*, 2014; Ueno *et al.*, 2014). The angle of the stalk extending from the AAA-ring to the B-tubule is, regardless of nucleotide states, tilted toward the proximal end (Movassagh *et al.*, 2010; Ueno *et al.*, 2008), implying the absence of large swing for power stroke (Figure 5(c)). Instead, ~ 8 nm shift of the AAA-ring with respect to the A-tubule was found (Movassagh *et al.*, 2010), suggesting the shift of the head 'pulls' the adjacent B-tubule. However, according to cryo-ET, slight rotation was detected: the angle of the stalk with respect to the B-tubule is $\sim 5^\circ$ shallower in the pre-power-stroke state than in the post-powerstroke state (Lin *et al.*, 2014; Movassagh *et al.*, 2010; Ueno *et al.*, 2014), which may suggest active rotation of axonemal dynein to 'push' the B-tubule.

Diversity of DHC Genes

Although there are two or three OADs and eight IADs in one 96 nm unit, each species has more genes of DHCs. *Chlamydomonas* has 16 DHC genes, while *Tetrahymena* has 25. Based on sequence analysis, they are classified into nine major classes of HC (Morris *et al.*, 2006; Wickstead and Gull, 2007; Wilkes

et al., 2008), seven out of which are axonemal – called outer alpha, outer beta, IAD-1 α , IAD-1 β , IAD-3, IAD-4, and IAD-5 (Morris *et al.*, 2006; Wickstead and Gull, 2007). IAD-1 α and IAD-1 β form a dimer. IAD-3, IAD-4, and IAD-5 are single-headed dyneins.

Products of *Chlamydomonas* IAD genes (DHC1-12) are named, based on biochemical analysis, dyneins a–g. Dynein f (DHC1 and DHC10) is dimeric, corresponding to IAD-1 α and -1 β (Figure 5(a)). There are 10 monomeric IADs. While there are six loci for single-headed IADs in one 96 nm unit, occupied by dyneins a, b, c, e, g, and d (from proximal to distal) (Bui *et al.*, 2008, 2012), other four IADs are called minor dyneins (DHC3, DHC4, DHC11, and DHC12) (Yagi *et al.*, 2009). Dynein d belongs to IAD-4. DHC3 and dynein g (DHC7) belong to IAD-5. All other except DHC12 belong to IAD-3. In *Tetrahymena* DYH1 and 2 are cytoplasmic. DYH3–5 are OAD. DYH 6 and 7 are two-headed IAD. The other 18 DHCs are single headed. Dissimilarly to *Chlamydomonas* and *Tetrahymena*, human has duplication in OAD. β -HC orthologs of human are DNAH11, DNAH17, and DNAH9, while DNAH5 and DNAH8 are γ -HC homologs (Table 1).

The phylogenetic tree of dynein spreads from cytoplasmic dynein (Figure 6(a)). Judging from the dynein family from diverse lineages, a common ancestor of dyneins likely existed before the diversification of eukaryotes and then was lost in some species (Wickstead and Gull, 2012). In the two-headed ODA, α - and β -families are not necessarily monophyletic (Morris *et al.*, 2006; Wickstead and Gull, 2007; Wilkes *et al.*, 2008). In the three-headed ODA, the outermost also belongs to β -family (Sakakibara *et al.*, 1991) and is not classifiable from the central ODAs by phylogenetic analysis (Wickstead and Gull, 2007; Wilkes *et al.*, 2008). It is not clear why and how so many duplications took place, such as in the OAD- β family, IAD-3, and IAD-5 of *Tetrahymena* (Wickstead and Gull, 2007; Wilkes *et al.*, 2008).

Localization and Asymmetry in the Axoneme

One clue to that question is localization of dynein isoforms. In human distribution of OAD, isoforms depends on organs: DNAH5 in the lung, brain, and testis (Olbrich *et al.*, 2002); DNAH9 in respiratory, lung, and sperm (Reed *et al.*, 2000); DNAH8 testis specific (Samant *et al.*, 2002). Uneven distribution can be observed within one cilium. While DNAH5 exists in the entire length of cilia, DNAH9 is localized to the distal part of cilia (Fliegauf *et al.*, 2007; Figure 6(b)).

Localization of dynein species has been studied at the ultrastructural level in *Chlamydomonas*. Electron microscopy of sections of biflagellated *Chlamydomonas* proved that apposed MTDs (we name it MTD1) in two cilia lack ODA (Figure 6(c); Hoops and Witman, 1983). They also reported unusual linkers between this MTD and one adjacent MTD (MTD2) at the proximal area of cilia (Figure 6(c)). After circumferential asymmetry of IAD from bop2 mutant was reported (King *et al.*, 1994), detailed 3D reconstruction from cryo-ET depicted circumferential and longitudinal asymmetry of IAD in wild-type *Chlamydomonas*. Dynein b is missing in MTD1, MTD9, and in the proximal area of all the MTDs (Bui *et al.*, 2009, 2012). Dynein c and f are missing in the entire MTD1 and the proximal area of MTD1, respectively (Bui *et al.*, 2012). Minor dyneins were proved to be localized in the proximal area of *Chlamydomonas* flagella (Yagi *et al.*, 2009; Figure 6(b)), replacing other IADs (Bui *et al.*, 2012). The asymmetry of dynein distribution suggests the sliding force and direction is not homogeneous in the entire cilia and this heterogeneity leads specific waveforms of cilia.

Function of Axonemal Dynein HCs in Cilia

MT sliding caused by dynein is the source of bending (Brokaw, 1989; Satir, 1968; Shingyoji *et al.*, 1977; Summers and

Table 1 16 isoforms of dyneins in *Chlamydomonas*

Protein	Gene	Binding components	Oligomerization	<i>Tetrahymena</i> homologue	Human homologue
OAD α			Trimer	DYH5 (γ)	
OAD β			Trimer	DYH4 (β)	DNAH11, 17, 9
OAD γ			Trimer	DYH3 (α)	DNAH5, 8
Dynein f α	DHC1	IC140, IC138, IC97,	Dimer	DYH6	DYH10
Dynein f β	DHC10	FAP120, Tctex1, Tctex2b, LC7a, LC7b, LC8	Dimer	DYH7	DYH9, 17
Dynein a	DHC6	Actin, p28, RS1	Single headed		
Dynein b	DHC5	Actin, centrin, RS1	Single headed		
Dynein c	DHC9	Actin, p28, RS2	Single headed		
Dynein e	DHC8	Actin, centrin, RS2	Single headed		
Dynein g	DHC7	Actin, centrin, RS3	Single headed		
Dynein d	DHC2	Actin, p28, p38, p44, RS3	Single headed		
Minor IAD	DHC3	Actin, centrin	Single headed		
Minor IAD	DHC4	Actin, centrin	Single headed		
Minor IAD	DHC11	Actin, p28	Single headed		
Minor IAD	DHC12		Single headed		
Cytoplasmic dynein			Homodimer	DYH2	DYNC2H1

Source: Based on Alford, L.M., Wirschell, M., Yamamoto, R., Sale, W.S., 2012. Control of axonemal inner dynein arms. In: King, S.M. (Ed.), Dyneins – Structure, Biology and Disease. Academic Press, pp. 313–335. ISBN: 978-0-12-382004-4.

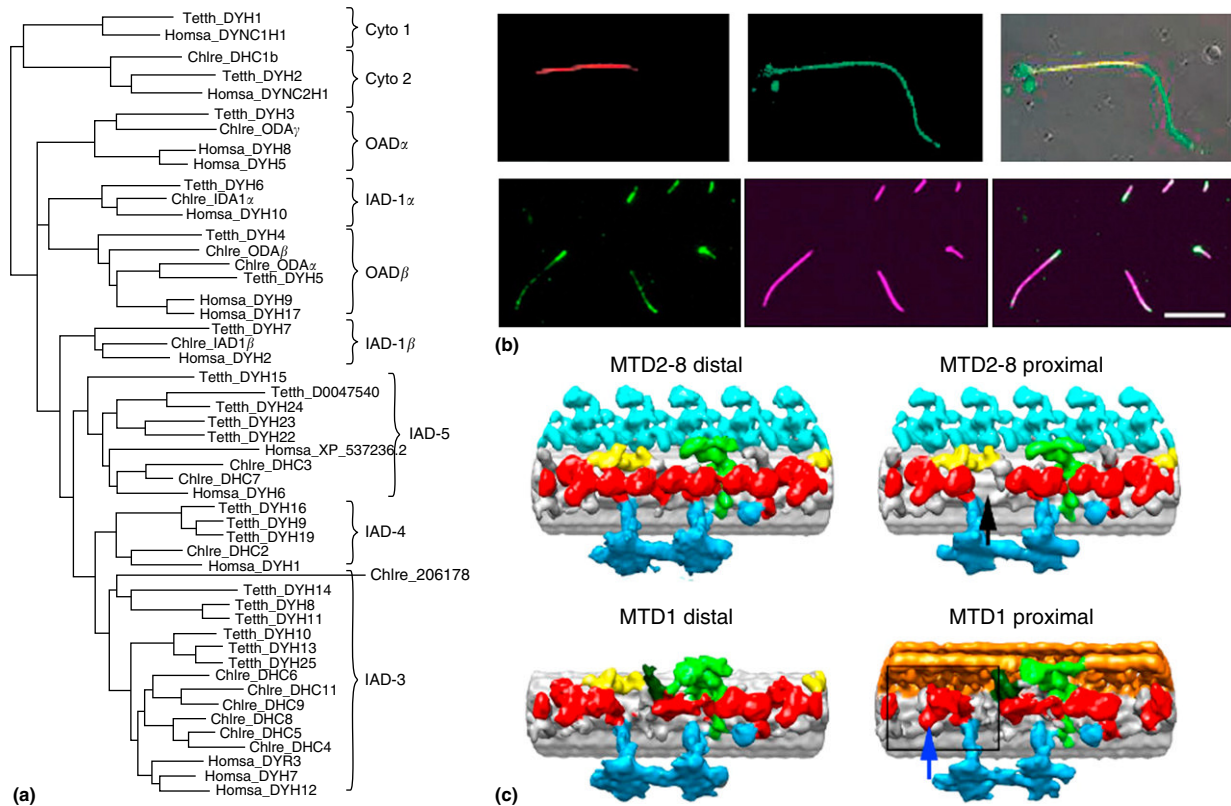


Figure 6 Diversity of axonemal dyneins. (a) Phylogenetic trees of dyneins from *Chlamydomonas*, *Tetrahymena*, and human. Modified from Wickstead, B., Gull, K., 2007. Dyneins across eukaryotes: A comparative genomic analysis. *Traffic* (Copenhagen, Denmark) 8, 1708–1721. (b) Localization of human outer (upper; DNAH5, one outer dynein species in red and acetylated tubulin in green) and *Chlamydomonas* inner (lower; minor inner dynein DHC11 in green and α -tubulin in purple) arm dyneins shown by immunofluorescence. Both demonstrate localization of these dynein isoforms in the proximal region of the axoneme. Modified from Fliegauf, M., Olbrich, H., Horvath, J., *et al.*, 2005. Mislocalization of DNAH5 and DNAH9 in respiratory cells from patients with primary ciliary dyskinesia. *American Journal of Respiratory and Critical Care Medicine* 171, 1343–1349; Yagi, T., Uematsu, K., Liu, Z., Kamiya, R., 2009. Identification of dyneins that localize exclusively to the proximal portion of *Chlamydomonas* flagella. *Journal of Cell Science* 122, 1306–1314, respectively. (c) Arrangement of inner and OAD on different MTDs in the proximal and distal regions. Modified from Bui, K.H., Yagi, T., Yamamoto, R., Kamiya, R., Ishikawa, T., 2012. Polarity and asymmetry in the arrangement of dynein and related structures in the *Chlamydomonas* axoneme. *Journal of Cell Biology* 198, 913–925.

Gibbons, 1971). But, the mechanism to integrate the sliding of individual DHCs into bending is unclear. The Switching model (Satir and Matsuoka, 1989; Wais-Steider and Satir, 1979) proposed that dyneins on one side of the axoneme is active and the other side is inactive, depending on the curvature. Recent cryo-ET of sperm flagella supports this model (Lin *et al.*, 2014). A number of theoretical studies were done to formulate force regulation as function of curvature or other geometric parameter of the axoneme to explain bending motion (Brokaw, 2009; Lindemann and Lesich, 2010; Riedel-Kruse *et al.*, 2007; Woolley, 2010). While switching is still a phenomenon to be detected, the mechanism of switching is unknown. It may be caused by regulation by IC/LC and RS/CP (Inoue and Shingyoji, 2007; Lehtreck and Witman, 2007; Lehtreck *et al.*, 2008; Nakano *et al.*, 2003; Porter and Sale, 2000; Smith, 2002), mechano-sensing of DHC or combination of them.

To address these questions, function of dynein is evaluated at multiple scales. Isolated DHC is measured by *in vitro* motility assay. Many dynein species with regulatory proteins are integrated into bending motion in the axoneme and therefore

its waveform, beating frequency, and the velocity of the whole cell are investigated by light microscopy. Between *in vitro* and *in vivo* levels, there is an experiment called sliding disintegration, in which MTDs slide out from enzymatically treated axonemes (Sale *et al.*, 1993; Summers and Gibbons, 1971).

At the level of single HC of axonemal dyneins, it is known that AAA2, 3, and 4 are non-catalytic (Kikushima *et al.*, 2004; Shiroguchi and Toyoshima, 2001). When ATP binds to AAA2–4, crossbridge is formed; when ADP binds, no crossbridge (Shingyoji, 2009; Yoshimura *et al.*, 2007).

Some studies indicate mechano-sensory function of axonemal dyneins. Mechanical stimulation can induce bending in paralyzed flagella (Hayashibe *et al.*, 1997; Inoue and Shingyoji, 2007). Mechanical induction occurs only when ADP binds. Already active dynein does not need ADP-binding (Hayashi and Shingyoji, 2008). Some ODA mutants of *Chlamydomonas* decrease beating under viscous condition (Minoura and Kamiya, 1995), which might be an evidence of mechano-sensing by dynein. Another phenomenon suggesting sensory function was found in cryo-electron tomography of the axoneme in the presence of ATP analog (Movassagh *et al.*,

2010). In the presence of non-hydrolysable nucleotide analog, OADs in the axoneme form nucleotide-bound or apo conformations, making clusters. This experiment indicates cooperativity between OADs on MTD – conformational changes of OADs are not independent of each other. This phenomenon could be explained by mechano-sensing by dyneins.

Distinct Roles of Outer and Inner Dynein Arm HCs

ODA mutants are slow and reduce beat frequency, but swim with the proper waveform (Kamiya and Okamoto, 1985; Mitchell and Rosenbaum, 1985). Also in sea urchin sperm defect of ODA causes decrease of frequency and power (Gibbons and Gibbons, 1973), while IDA-deficient mutant reduces the amplitude of bending motion or totally paralyzed.

Dynein f mutant shows alternation of waveform and less efficient beat (Kamiya *et al.*, 1991). Dynein f *in vitro* is processive, but slow ($1.5\text{--}3.0\ \mu\text{m sec}^{-1}$) (Kotani *et al.*, 2007), compared to other dyneins ($20\ \mu\text{m sec}^{-1}$), which lead to proposals of dynein f as a modulator of motility (Bower *et al.*, 2009; Kotani *et al.*, 2007; Wirschell *et al.*, 2007).

Dynein c proteins is a processive fastest motor (Sakakibara *et al.*, 1999). Dynein c mutant reduces velocity in viscous media. Mutant of single-headed IAD of *Tetrahymena*, DYH8 (IAD-3), DYH9 (IAD-4), and DYH12 (IAD-3) mutants swim $\sim 70\%$ velocity (Liu *et al.*, 2004). Only DYH8 mutant shows frequency reduction. All the three mutants exhibit a more sensitive reversal reaction (Liu *et al.*, 2005).

Regulatory Components

Localized dynein control by downstream signal will be the mechanism of bending motion (Elam *et al.*, 2009; Porter and Sale, 2000). Various biochemical reactions regulate dynein: phosphorylation utilizing such kinases as PKA (Salathe, 2007) and CK1 (Gokhale *et al.*, 2009), AKAP (Gaillard *et al.*, 2001, 2006), and phosphatases (PP1, PP2A) (Habermacher and Sale, 1996, 1997; Yang *et al.*, 2000). ODA α -HC is phosphorylated at the sites near the AAA1 and C-term (King and Witman, 1994). Calcium is another regulatory factor (King, 2010; Salathe, 2007). Calcium (and Cyclic adenosine monophosphate (cAMP)) induces alteration of waveform between symmetric and asymmetric waveform (Hasegawa *et al.*, 1987; Kamiya and Witman, 1984; Omoto and Brokaw, 1985). To elucidate the mechanism of waveform formation it is essential to identify and characterize the regulatory components and their signal transduction pathway, which will regulate axonemal dynein HCs in the end.

Intermediate Chains and Light Chains

OAD has WD-repeat ICs and LCs as well as other coiled-coil proteins (reviews in King, 2010, 2012). IC1 and IC2 are essential for ODA assembly (Fowkes and Mitchell, 1998). In *Chlamydomonas* mutation and knockdown of LC1, which has a leucine-rich repeat and binds to γ -dynein and tubulin (Benashski *et al.*, 1999; Wu *et al.*, 2000), results in reduced velocity and stalled beat (Rompolas *et al.*, 2010). ODA

contains a few redox-sensitive components (Wakabayashi, 2009), while LC4 (42% identical to calmodulin) and IC1 are Ca^{2+} sensitive (Sakato and King, 2003).

ODA is preassembled in the cytoplasm and transported by IFT with adaptors such as Oda16p, a WD-repeat protein (Ahmed *et al.*, 2008) and LC8 (Pazour *et al.*, 1998). Lis1 (see 'Lis1 and NudE') is found also in mammalian respiratory and oviduct cilia and known to be essential for ODA assembly (Pedersen *et al.*, 2007; Rompolas *et al.*, 2012). ODA7, a protein with a leucine-rich repeat and a coiled-coil, is also necessary for assembling α -HC and interacts with ODA and dynein f (Freshour *et al.*, 2007). To anchor ODA on MTD, docking complex, including coiled-coil proteins (DC1 and DC2) are necessary (Koutoulis *et al.*, 1997; Owa *et al.*, 2014; Takada *et al.*, 2002).

Dimeric IAD f has a large complex of ICs and LCs (Figure 7(a); Table 1). The following ICs and LCs were identified: IC138, IC97, FAP120, and LC7b form a phosphoregulatory subcomplex (Bower *et al.*, 2009; Ikeda *et al.*, 2009; Wirschell *et al.*, 2009). IC138 is phosphoprotein and its phosphorylation inactivates dynein f (Porter and Sale, 2000; Wirschell *et al.*, 2007; Yang and Sale, 2000). Hyperphosphorylation of IC138 is involved in control of flagella beating (Habermacher and Sale, 1997). IC97 is non-WD IC (Wirschell *et al.*, 2009). FAP120 is an ankyrin-repeat protein. IC140, LC8, Tctex1, Tctex2b, and LC7a form another subcomplex (Perrone *et al.*, 1998; Yang and Sale, 1998; Figure 7).

Monomeric IADs have centrin, actin (Piperno and Luck, 1979; Yanagisawa and Kamiya, 2001), and p28 (a LC) (LeDizet and Piperno, 1995) (summarized in Table 1). Until now no ICs have been found in single-headed IADs. Dynein d binds p38 (Yamamoto *et al.*, 2006) and p44 (Yamamoto *et al.*, 2008). These two proteins are homologs of LIC (30–60 kDa LC in cytoplasmic dynein) (King *et al.*, 1996; Yamamoto *et al.*, 2006, 2008).

RS and Central Pair

The RS is a complex of 23 proteins (it is not known how many copies of each component exist in one complex) (Yang *et al.*, 2006), protruding from MTD toward the CP (Goodenough and Heuser, 1985). Since RS includes calmodulin (Yang *et al.*, 2001, 2006), RS is supposed to be responsible for calcium-induced alteration of waveforms. There are either two (such as *Chlamydomonas*) or three (such as *Tetrahymena*) RSs within a 96 nm periodic unit. The N-terminal tail part of the singlet IAD is merged with RS (Pigino *et al.*, 2011; Figure 7(b)), suggesting interaction between them. Since RS is required for normal motility and contains AKAP, calmodulin, and PKA, it is likely the transducer and a dynein regulator (Porter and Sale, 2000; Smith and Yang, 2004; Wirschell *et al.*, 2007), although the detailed molecular mechanism of signal transduction is not known.

There are a few studies indicating multiple pathways to regulate dyneins to generate bending motion (reviewed in Kamiya and Yagi, 2014). Paralyzed RS/CP mutants are rescued by nucleotide concentration or mechanical stimulation, but only in the presence of ODA (Frey *et al.*, 1997; Hayashibe *et al.*, 1997; Kamiya, 2002; Yagi and Kamiya, 2000). Sliding velocity

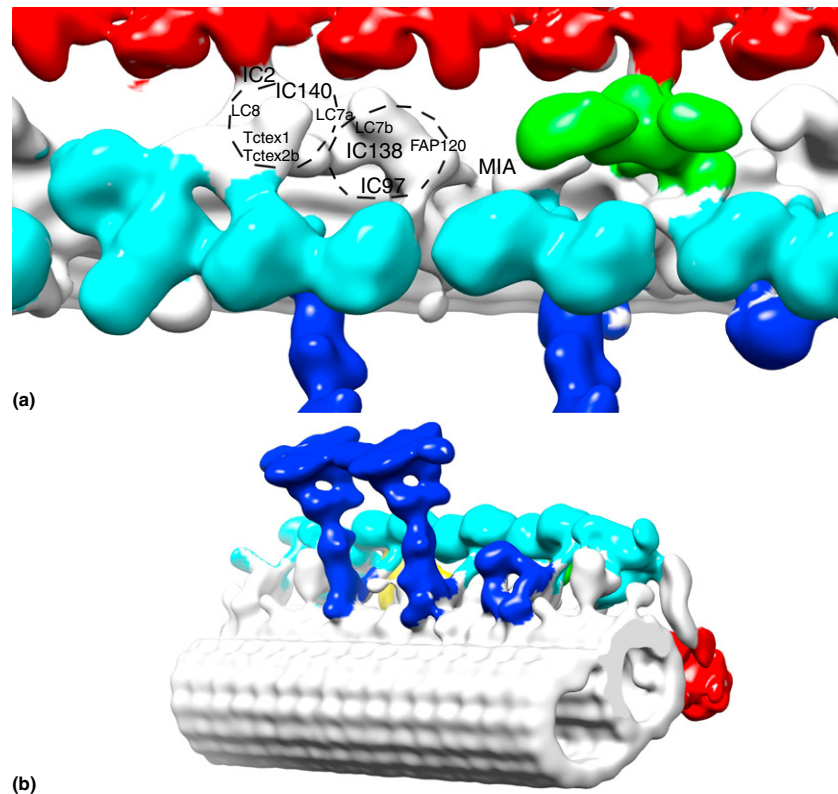


Figure 7 Connection between dyneins and other ciliary components. (a) Possible positions of IC/LC are plotted on the cryo-ET structure. Based on data from Bower, R., VanderWaal, K., O'Toole, E., *et al.*, 2009. IC138 defines a subdomain at the base of the I1 dynein that regulates microtubule sliding and flagellar motility. *Molecular Biology of the Cell* 20, 3055–3063; Oda, T., Kikkawa, M., 2013. Novel structural labeling method using cryo-electron tomography and biotin-streptavidin system. *Journal of Structural Biology* 183, 305–311; Perrone, C.A., Yang, P., O'Toole, E., Sale, W.S., Porter, M.E., 1998. The *Chlamydomonas* IDA7 locus encodes a 140-kDa dynein intermediate chain required to assemble the I1 inner arm complex. *Molecular Biology of the Cell* 9, 3351–3365; and Yamamoto, R., Song, K., Yanagisawa, H.-A., *et al.*, 2013. The MIA complex is a conserved and novel dynein regulator essential for normal ciliary motility. *Journal of Cell Biology* 201, 263–278. (b) RS (blue) and IAD (cyan) are connected at the base, suggesting the pathway of signal transduction for dynein regulation. Based on the data from Pigino, G., Bui, K.H., Maheshwari, A., *et al.*, 2011. Cryo-electron tomography of radial spokes in cilia and flagella. *Journal of Cell Biology* 195, 673–687.

is increased in RS/ODA double mutants (Smith and Sale, 1992). These studies indicate pathways to regulate dyneins by RS-IDA and ODA are independent or compensatory. IC138, phosphorylation of which inhibits dynein f, plays an important role of the RS-IDA pathway. When CP/RS is defective, IC138 phosphorylation is abnormal, resulting in reduced velocities (Habermacher and Sale, 1997; Hendrickson *et al.*, 2004). Kinase inhibitors rescue sliding (Gokhale *et al.*, 2009; Yang and Sale, 2000). This suggests dynein f and IC138 phosphorylation are targets for RS. Based on morphological changes of RS, steric collision was hypothesized as the cause of RS/CP interaction (Mitchell and Nakatsugawa, 2004; Warner and Satir, 1974). Genetic tagging to increase the volume of RS rescues paralyzed deletion mutant of CP proteins, supporting this hypothesis (Oda *et al.*, 2014).

Axonemal Dynein-Related Ciliopathies

Dynein-related ciliopathy such as primary ciliary dyskinesia (PCD) was identified from respiratory cilia lacking dynein

arms from patients with Kartagener's syndrome (Afzelius *et al.*, 1975; Pedersen and Rebbe, 1975).

Cilia in PCD defecting DNAH11 shows deficient in IC78/IC1, similar to *Chlamydomonas* oda9 mutant (Pennarun *et al.*, 1999), indicating assembly defect of ODA in the distal part (Fliegauf *et al.*, 2007). DNAH5 mutation causes random left/right asymmetry (Fliegauf *et al.*, 2005; Olbrich *et al.*, 2002) and hydrocephalus, situs inversus, respiratory infection in mouse (Ibañez-Tallon *et al.*, 2002; Olbrich *et al.*, 2002). Patients with mutation of DNAH11, an ortholog of *Chlamydomonas* β -HC and abundant in the node, suffer from cystic fibrosis (Bartoloni *et al.*, 2002). DNAH11 mutant shows defect in LR determination in mice (Supp *et al.*, 1997) and reduced flow in the Kupffer's vesicle and in the spinal cord and causes hydrocephals in zebrafish (Essner *et al.*, 2005; Kramer-Zucker *et al.*, 2005). Mutation of DNAH2, an ortholog of *Chlamydomonas* IC69 (IC2), defects assembly of ODA (Loges *et al.*, 2008, 2009).

Defect of essential genes for dynein assembly can be a cause of PCD. Cytoplasmic assembly factor, kintoun (ktu), is a member of PIH group and related to PCD (Omran *et al.*,

2008). Depletion of PF13, *Chlamydomonas* ortholog of ktu, causes failure of ODA and IDA assembly (Huang *et al.*, 1979; Omran *et al.*, 2008). Another PIH protein, MOT48, encoded at IDA10, is required for assembly of dyneins b, c, d, and e (Yamamoto *et al.*, 2010), suggesting that this type of PCD is caused by ODA defect. Lacking LRRC50 defects ODA and IDA (Fowkes and Mitchell, 1998; Freshour *et al.*, 2007; Sullivan-Brown *et al.*, 2008; Supp *et al.*, 1997), disrupting DNAH9 (distal), DNAH5, and DNAI2, causing PCD.

PCD (airway, left/right, and immotile sperm) is caused also from DNALI1 (IAD) (Zariwala *et al.*, 2013). Morphological change in DRC and IAD were found. Mouse DNAH1 knockout (dynein d homolog) (Neesen *et al.*, 2001) reduces beat frequency of tracheal cilia and incapability of sperm swimming in viscous media (Woolley *et al.*, 2005).

See also: Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cilia and Flagella

References

- Afzelius, B.A., Eliasson, R., Johnsen, O., Lindholmer, C., 1975. Lack of dynein arms in immotile human spermatozoa. *Journal of Cell Biology* 66, 225–232.
- Ahmad, F.J., Echeverri, C.J., Vallee, R.B., Baas, P.W., 1998. Cytoplasmic dynein and dynactin are required for the transport of microtubules into the axon. *Journal of Cell Biology* 140, 391–401.
- Ahmed, N.T., Gao, C., Luckner, B.F., Cole, D.G., Mitchell, D.R., 2008. ODA16 aids axonemal outer row dynein assembly through an interaction with the intraflagellar transport machinery. *Journal of Cell Biology* 183, 313–322.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215, 403–410.
- Armbrust, E.V., Berges, J.A., Bowler, C., *et al.*, 2004. The genome of the diatom *Thalassiosira pseudonana*. Ecology, evolution, and metabolism. *Science* 306, 79–86.
- Ashkin, A., Schütze, K., Dziedzic, J.M., Euteneuer, U., Schliwa, M., 1990. Force generation of organelle transport measured *in vivo* by an infrared laser trap. *Nature* 348, 346–348.
- Bartoloni, L., Blouin, J.-L., Pan, Y., *et al.*, 2002. Mutations in the DNAH11 (axonemal heavy chain dynein type 11) gene cause one form of situs inversus totalis and most likely primary ciliary dyskinesia. *Proceedings of the National Academy of Sciences of the United States of America* 99, 10282–10286.
- Basu, B., Brueckner, M., 2008. Cilia multifunctional organelles at the center of vertebrate left-right asymmetry. *Current Topics in Developmental Biology* 85, 151–174.
- Belyy, V., Hendel, N.L., Chien, A., Yildiz, A., 2014. Cytoplasmic dynein transports cargos via load-sharing between the heads. *Nature Communications* 5, 1–9.
- Benashski, S.E., Patel-King, R.S., King, S.M., 1999. Light chain 1 from the *Chlamydomonas* outer dynein arm is a leucine-rich repeat protein associated with the motor domain of the gamma heavy chain. *Biochemistry (Moscow)* 38, 7253–7264.
- Bhabha, G., Cheng, H., Zhang, N., *et al.*, 2014. Article allosteric communication in the dynein motor domain. *Cell* 159, 857–868.
- Bower, R., VanderWaal, K., O'Toole, E., *et al.*, 2009. IC138 defines a subdomain at the base of the I1 dynein that regulates microtubule sliding and flagellar motility. *Molecular Biology of the Cell* 20, 3055–3063.
- Brokaw, C.J., 1989. Direct measurements of sliding between outer doublet microtubules in swimming sperm flagella. *Science* 243, 1593–1596.
- Brokaw, C.J., 2009. Thinking about flagellar oscillation. *Cell Motility and the Cytoskeleton* 66, 425–436.
- Bui, K.H., Sakakibara, H., Movassagh, T., Oiwa, K., Ishikawa, T., 2008. Molecular architecture of inner dynein arms in situ in *Chlamydomonas reinhardtii* flagella. *Journal of Cell Biology* 183, 923–932.
- Bui, K.H., Sakakibara, H., Movassagh, T., Oiwa, K., Ishikawa, T., 2009. Asymmetry of inner dynein arms and inter-doublet links in *Chlamydomonas* flagella. *Journal of Cell Biology* 186, 437–446.
- Bui, K.H., Yagi, T., Yamamoto, R., Kamiya, R., Ishikawa, T., 2012. Polarity and asymmetry in the arrangement of dynein and related structures in the *Chlamydomonas* axoneme. *Journal of Cell Biology* 198, 913–925.
- Bullock, S.L., Nicol, A., Gross, S.P., Zicha, D., 2006. Guidance of bidirectional motor complexes by mRNA Cargoes through control of dynein number and activity. *Current Biology* 16, 1447–1452.
- Burgess, S.A., Walker, M.L., Sakakibara, H., Knight, P.J., Oiwa, K., 2003. Dynein structure and power stroke. *Nature* 421, 715–718.
- Can, S., Dewitt, M.A., Yildiz, A., 2014. Bidirectional helical motility of cytoplasmic dynein around microtubules. *eLife*. e03205.
- Carter, A., Garbarino, J., Wilson-Kubalek, E., *et al.*, 2008. Structure and functional role of dynein's microtubule-binding domain. *Science* 322, 1691–1695.
- Carter, A.P., Cho, C., Jin, L., Vale, R.D., 2011. Crystal structure of the dynein motor domain. *Science* 331, 1159–1165.
- Cho, C., Reck-Peterson, S.L., Vale, R.D., 2008. Regulatory ATPase sites of cytoplasmic dynein affect processivity and force generation. *Journal of Biological Chemistry* 283, 25839–25845.
- Cleary, F.B., Dewitt, M.A., Bilyard, T., *et al.*, 2014. Tension on the linker gates the ATP-dependent release of dynein from microtubules. *Nature Communications* 5, 4587.
- Combs, J., Kim, S.J., Tan, S., *et al.*, 2006. Recruitment of dynein to the Jurkat immunological synapse. *Proceedings of the National Academy of Sciences of the United States of America* 103, 14883–14888.
- Delanoue, R., Davis, I., 2005. Dynein anchors its mRNA cargo after apical transport in the *Drosophila* blastoderm embryo. *Cell* 122, 97–106.
- Derr, N.D., Goodman, B.S., Jungmann, R., *et al.*, 2012. Tug-of-war in motor protein ensembles revealed with a programmable DNA origami scaffold. *Science* 338, 662–665.
- DeWitt, M.A., Chang, A.Y., Combs, P.A., Yildiz, A., 2012. Cytoplasmic dynein moves through uncoordinated stepping of the AAA + ring domains. *Science* 335, 221–225.
- Dewitt, M.A., Cyranowska, C.A., Cleary, F.B., Belyy, V., Yildiz, A., 2015. The AAA3 domain of cytoplasmic dynein acts as a "switch" to facilitate microtubule release. *Nature Structural & Molecular Biology* 22, 73–80.
- Dienstbier, M., Boehl, F., Li, X., Bullock, S.L., 2009. Egalitarian is a selective RNA-binding protein linking mRNA localization signals to the dynein motor. *Genes & Development* 23, 1546–1558.
- Dobyns, W.B., Reiner, O., Carrozzo, R., Ledbetter, D.H., 1993. Lissencephaly. A human brain malformation associated with deletion of the LIS1 gene located at chromosome 17p13. *Journal of the American Medical Association* 270, 2838–2842.
- Dodding, M.P., Way, M., 2011. Coupling viruses to dynein and kinesin-1. *EMBO Journal* 30, 3527–3539.
- Dujardin, D.L., Barnhart, L.E., Stehman, S.A., *et al.*, 2003. A role for cytoplasmic dynein and LIS1 in directed cell movement. *Journal of Cell Biology* 163, 1205–1211.
- Efimov, V.P., Morris, N.R., 2000. Coiled-coil domain of the NUDE/RO11 protein. *Cell* 150, 681–688.
- Elam, C.A., Sale, W.S., Wirschell, M., 2009. The regulation of dynein-driven microtubule sliding in *Chlamydomonas* flagella by axonemal kinases and phosphatases. *Methods in Cell Biology* 92, 133–151.
- Elting, M.W., Hueschen, C.L., Udy, D.B., Dumont, S., 2014. Force on spindle microtubule minus ends moves chromosomes. *Journal of Cell Biology* 206, 245–256.
- Erzberger, J.P., Berger, J.M., 2006. Evolutionary relationships and structural mechanisms of AAA + proteins. *Annual Review of Biophysics and Biomolecular Structure* 35, 93–114.
- Essner, J.J., Amack, J.D., Nyholm, M.K., Harris, E.B., Yost, H.J., 2005. Kupffer's vesicle is a ciliated organ of asymmetry in the zebrafish embryo that initiates left-right development of the brain, heart and gut. *Development (Cambridge, England)* 132, 1247–1260.
- Fliegauf, M., Benzing, T., Omran, H., 2007. When cilia go bad: Cilia defects and ciliopathies. *Nature Reviews Molecular Cell Biology* 8, 880–893.
- Fliegauf, M., Olbrich, H., Horvath, J., *et al.*, 2005. Mislocalization of DNAH5 and DNAH9 in respiratory cells from patients with primary ciliary dyskinesia. *American Journal of Respiratory and Critical Care Medicine* 171, 1343–1349.
- Fowkes, M.E., Mitchell, D.R., 1998. The role of preassembled cytoplasmic complexes in assembly of flagellar dynein subunits. *Molecular Biology of the Cell* 9, 2337–2347.
- Freshour, J., Yokoyama, R., Mitchell, D.R., 2007. *Chlamydomonas* flagellar outer row dynein assembly protein ODA7 interacts with both outer row and I1 inner row dyneins. *Journal of Biological Chemistry* 282, 5404–5412.

- Frey, E., Brokaw, C.J., Omoto, C.K., 1997. Reactivation at low ATP distinguishes among classes of paralyzed flagella mutants. *Cell Motility and the Cytoskeleton* 38, 91–99.
- Fridolfsson, H.N., Ly, N., Meyerzon, M., Starr, D.A., 2010. UNC-83 coordinates kinesin-1 and dynein activities at the nuclear envelope during nuclear migration. *Developmental Biology* 338, 237–250.
- Gaetz, J., Kapoor, T.M., 2004. Dynein/dynactin regulate metaphase spindle length by targeting depolymerizing activities to spindle poles. *Journal of Cell Biology* 166, 465–471.
- Gaillard, A.R., Diener, D.R., Rosenbaum, J.L., Sale, W.S., 2001. Flagellar radial spoke protein 3 is an A-kinase anchoring protein (AKAP). *Journal of Cell Biology* 153, 443–448.
- Gaillard, A.R., Fox, L.A., Rhea, J.M., Craig, B., Sale, W.S., 2006. Disruption of the A-kinase anchoring domain in flagellar radial spoke protein 3 results in unregulated axonemal cAMP-dependent protein kinase activity and abnormal flagellar motility. *Molecular Biology of the Cell* 17, 2626–2635.
- Gatlin, J.C., Matov, A., Groen, A.C., et al., 2009. Spindle fusion requires dynein-mediated sliding of oppositely oriented microtubules. *Current Biology* 19, 287–296.
- Gennerich, A., Carter, A.P., Reck-Peterson, S.L., Vale, R.D., 2007. Force-induced bidirectional stepping of cytoplasmic dynein. *Cell* 131, 952–965.
- Gibbons, B.H., Asai, D.J., Tang, W.J., Hays, T.S., Gibbons, I.R., 1994. Phylogeny and expression of axonemal and cytoplasmic dynein genes in sea urchins. *Molecular Biology of the Cell* 5, 57–70.
- Gibbons, B.H., Baccetti, B., Gibbons, I.R., 1985. Live and reactivated motility in the 9 + 0 flagellum of *Anguilla* sperm. *Cell Motility* 5, 333–350.
- Gibbons, B.H., Gibbons, I.R., 1973. The effect of partial extraction of dynein arms on the movement of reactivated sea-urchin sperm. *Journal of Cell Science* 13, 337–357.
- Gibbons, B.H., Gibbons, I.R., Baccetti, B., 1983. Structure and motility of the 9 + 0 flagellum of eel spermatozoa. *Journal of Submicroscopic Cytology* 15, 15–20.
- Gibbons, I.R., Garbarino, J.E., Tan, C.E., et al., 2005. The affinity of the dynein microtubule-binding domain is modulated by the conformation of its coiled-coil stalk. *Journal of Biological Chemistry* 280, 23960–23965.
- Glynn, S.E., Martin, A., Nager, A.R., Baker, T., Sauer, R.T., 2009. Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA + protein-unfolding machine. *Cell* 139, 744–756.
- Glynn, S.E., Nager, A.R., Baker, T.A., Sauer, R.T., 2012. Dynamic and static components power unfolding in topologically closed rings of a AAA + proteolytic machine. *Nature Structural & Molecular Biology* 19, 616–622.
- Gokhale, A., Wirschell, M., Sale, W.S., 2009. Regulation of dynein-driven microtubule sliding by the axonemal protein kinase CK1 in *Chlamydomonas* flagella. *Journal of Cell Biology* 186, 817–824.
- Gomes, E.R., Jani, S., Gundersen, G.G., 2005. Nuclear movement regulated by Cdc42, MRCK, myosin, and actin flow establishes MTOC polarization in migrating cells. *Cell* 121, 451–463.
- Goodenough, U.W., Heuser, J.E., 1985. Substructure of inner dynein arms, radial spokes, and the central pair/projection complex of cilia and flagella. *Journal of Cell Biology* 100, 2008–2018.
- Graham, P.W., Seale, G.E., Bennecib, M., Goldberg, D.J., Vallee, R.B., 2007. Cytoplasmic dynein and LIS1 are required for microtubule advance during growth cone remodeling and fast axonal outgrowth. *Journal of Neuroscience Society for Neuroscience* 27, 5823–5834.
- Guydosh, N.R., Block, S.M., 2006. Backsteps induced by nucleotide analogs suggest the front head of kinesin is gated by strain. *Proceedings of the National Academy of Sciences of the United States of America* 103, 8054–8059.
- Habermacher, G., Sale, W.S., 1996. Regulation of flagellar dynein by an axonemal type-1 phosphatase in *Chlamydomonas*. *Journal of Cell Science* 109 (Pt 7), 1899–1907.
- Habermacher, G., Sale, W.S., 1997. Regulation of flagellar dynein by phosphorylation of a 138-kD inner arm dynein intermediate chain. *Journal of Cell Biology* 136, 167–176.
- Harrell, J.M., Murphy, P.J.M., Morishima, Y., et al., 2004. Evidence for glucocorticoid receptor transport on microtubules by dynein. *Journal of Biological Chemistry* 279, 54647–54654.
- Hasegawa, E., Hayashi, H., Asakura, S., Kamiya, R., 1987. Stimulation of *in vitro* motility of *Chlamydomonas* axonemes by inhibition of cAMP-dependent phosphorylation. *Cell Motility and the Cytoskeleton* 8, 302–311.
- Hayashi, S., Shingyoji, C., 2008. Mechanism of flagellar oscillation-bending-induced switching of dynein activity in elastase-treated axonemes of sea urchin sperm. *Journal of Cell Science* 121, 2833–2843.
- Hayashibe, K., Shingyoji, C., Kamiya, R., 1997. Induction of temporary beating in paralyzed flagella of *Chlamydomonas* mutants by application of external force. *Cell Motility and the Cytoskeleton* 37, 232–239.
- Heald, R., Tournebise, R., Blank, T., et al., 1996. Self-organization of microtubules into bipolar spindles around artificial chromosomes in *Xenopus* egg extracts. *Nature* 382, 420–425.
- Hendricks, A.G., Holzbaur, E.L.F., Goldman, Y.E., 2012. Force measurements on cargoes in living cells reveal collective dynamics of microtubule motors. *Proceedings of the National Academy of Sciences of the United States of America* 109, 18447–18452.
- Hendrickson, T.W., Perrone, C.A., Griffin, P., et al., 2004. IC138 is a WD-repeat dynein intermediate chain required for light chain assembly and regulation of flagellar bending. *Molecular Biology of the Cell* 15, 5431–5442.
- Hoops, H.J., Witman, G.B., 1983. Outer doublet heterogeneity reveals structural polarity related to beat direction in *Chlamydomonas* flagella. *Journal of Cell Biology* 97, 902–908.
- Howell, B.J., McEwen, B.F., Canman, J.C., et al., 2001. Cytoplasmic dynein/dynactin drives kinetochore protein transport to the spindle poles and has a role in mitotic spindle checkpoint inactivation. *Journal of Cell Biology* 155, 1159–1172.
- Huang, B., Piperno, G., Luck, D.J., 1979. Paralyzed flagella mutants of *Chlamydomonas reinhardtii*. Defective for axonemal doublet microtubule arms. *Journal of Biological Chemistry* 254, 3091–3099.
- Huang, J., Roberts, A.J., Leschziner, A.E., Reck-Peterson, S.L., 2012. Lis1 acts as a “clutch” between the ATPase and microtubule-binding domains of the dynein motor. *Cell* 150, 975–986.
- Ibañez-Tallon, I., Gorokhova, S., Heintz, N., 2002. Loss of function of axonemal dynein Mdnah5 causes primary ciliary dyskinesia and hydrocephalus. *Human Molecular Genetics* 11, 715–721.
- Ikeda, K., Yamamoto, R., Wirschell, M., et al., 2009. A novel ankyrin-repeat protein interacts with the regulatory proteins of inner arm dynein f (I1) of *Chlamydomonas reinhardtii*. *Cell Motility and the Cytoskeleton* 66, 448–456.
- Inoue, Y., Shingyoji, C., 2007. The roles of noncatalytic ATP binding and ADP binding in the regulation of dynein motile activity in flagella. *Cell Motility and the Cytoskeleton* 64, 690–704.
- Kamiya, R., 2002. Functional diversity of axonemal dyneins as studied in *Chlamydomonas* mutants. *International Review of Cytology* 219, 115–155.
- Kamiya, R., Kurimoto, E., Muto, E., 1991. Two types of *Chlamydomonas* flagellar mutants missing different components of inner-arm dynein. *Journal of Cell Biology* 112, 441–447.
- Kamiya, R., Okamoto, M., 1985. A mutant of *Chlamydomonas reinhardtii* that lacks the flagellar outer dynein arm but can swim. *Journal of Cell Science* 74, 181–191.
- Kamiya, R., Witman, G.B., 1984. Submicromolar levels of calcium control the balance of beating between the two flagella in demembrated models of *Chlamydomonas*. *Journal of Cell Biology* 98, 97–107.
- Kamiya, R., Yagi, T., 2014. Functional diversity of axonemal dyneins as assessed by *in vitro* and *in vivo* motility assays of *chlamydomonas* mutants. *Zoological Science* 31, 633–644.
- Kardon, J.R., Reck-Peterson, S.L., Vale, R.D., 2009. Regulation of the processivity and intracellular localization of *Saccharomyces cerevisiae* dynein by dynactin. *Proceedings of the National Academy of Sciences of the United States of America* 106, 5669–5674.
- Kardon, J.R., Vale, R.D., 2009. Regulators of the cytoplasmic dynein motor. *Nature Reviews Molecular Cell Biology* 10, 854–865.
- Kikushima, K., Yagi, T., Kamiya, R., 2004. Slow ADP-dependent acceleration of microtubule translocation produced by an axonemal dynein. *FEBS Letters* 563, 119–122.
- Kim, J., Kato, M., Beachy, P.A., 2009. Gli2 trafficking links Hedgehog-dependent activation of smoothened in the primary cilium to transcriptional activation in the nucleus. *Proceedings of the National Academy of Sciences of the United States of America* 106, 21666–21671.
- King, S.J., Inwood, W.B., O'Toole, E.T., Power, J., Dutcher, S.K., 1994. The bop2-1 mutation reveals radial asymmetry in the inner dynein arm region of *Chlamydomonas reinhardtii*. *Journal of Cell Biology* 126, 1255–1266.
- King, S.M., 2010. Sensing the mechanical state of the axoneme and integration of Ca²⁺ signaling by outer arm dynein. *Cytoskeleton* 67, 207–213.
- King, S.M., 2012. Integrated control of axonemal dynein AAA(+) motors. *Journal of Structural Biology* 179, 222–228.
- King, S.M., Barbaresi, E., Dillman, J.F., et al., 1996. Brain cytoplasmic and flagellar outer arm dyneins share a highly conserved Mr 8000 light chain. *Journal of Biological Chemistry* 271, 19358–19366.
- King, S.M., Witman, G.B., 1994. Multiple sites of phosphorylation within the alpha heavy chain of *Chlamydomonas* outer arm dynein. *Journal of Biological Chemistry* 269, 5452–5457.
- King, S.J., Schroer, T.A., 2000. Dynactin increases the processivity of the cytoplasmic dynein motor. *Nature Cell Biology* 2, 20–24.

- Kiyomitsu, T., Cheeseman, I.M., 2012. Chromosome- and spindle-pole-derived signals generate an intrinsic code for spindle position and orientation. *Nature Cell Biology* 14, 311–317.
- Kon, T., Imamura, K., Roberts, A.J., *et al.*, 2009. Helix sliding in the stalk coiled coil of dynein couples ATPase and microtubule binding. *Nature Structural & Molecular Biology* 16, 325–333.
- Kon, T., Nishiura, M., Ohkura, R., Toyoshima, Y.Y., Sutoh, K., 2004. Distinct functions of nucleotide-binding/hydrolysis sites in the four AAA modules of cytoplasmic dynein. *Biochemistry (Moscow)* 43, 11266–11274.
- Kon, T., Oyama, T., Shimo-Kon, R., *et al.*, 2012. The 2.8 Å crystal structure of the dynein motor domain. *Nature* 484, 345–350.
- Kon, T., Sutoh, K., Kurisu, G., 2011. X-ray structure of a functional full-length dynein motor domain. *Nature Structural & Molecular Biology* 18, 638–642.
- Koonce, M.P., Tikhonenko, I., 2000. Functional elements within the dynein microtubule-binding domain. *Molecular Biology of the Cell* 11, 523–529.
- Kotani, N., Sakakibara, H., Burgess, S.A., Kojima, H., Oiwa, K., 2007. Mechanical properties of inner-arm dynein-f (dynein I1) studied with *in vitro* motility assays. *Biophysical Journal* 93, 886–894.
- Koutoulis, A., Pazour, G.J., Wilkerson, C.G., *et al.*, 1997. The *Chlamydomonas reinhardtii* ODA3 gene encodes a protein of the outer dynein arm docking complex. *Journal of Cell Biology* 137, 1069–1080.
- Kramer-Zucker, A.G., Olale, F., Haycraft, C.J., *et al.*, 2005. Cilia-driven fluid flow in the zebrafish pronephros, brain and Kupffer's vesicle is required for normal organogenesis. *Development (Cambridge, England)* 132, 1907–1921.
- Kull, F., Sablin, E., Lau, R., Fletterick, R., Vale, R., 1996. Crystal structure of the kinesin motor domain reveals a structural similarity to myosin. *Nature* 380, 550–555.
- Laib, J.A., Marin, J.A., Bloodgood, R., Guilford, W.H., 2009. The reciprocal coordination and mechanics of molecular motors in living cells. *Proceedings of the National Academy of Sciences of the United States of America* 106, 3190–3195.
- Lawrence, C.J., Morris, N.R., Meagher, R.B., Dawe, R.K., 2001. Dyneins have run their course in plant lineage. *Traffic (Copenhagen, Denmark)* 2, 362–363.
- Lechtreck, K.-F., Delmotte, P., Robinson, M.L., Sanderson, M.J., Witman, G.B., 2008. Mutations in Hydin impair ciliary motility in mice. *Journal of Cell Biology* 180, 633–643.
- Lechtreck, K.-F., Witman, G.B., 2007. *Chlamydomonas reinhardtii* hydin is a central pair protein required for flagellar motility. *Journal of Cell Biology* 176, 473–482.
- LeDizet, M., Piperno, G., 1995. The light chain p28 associates with a subset of inner dynein arm heavy chains in *Chlamydomonas* axonemes. *Molecular Biology of the Cell* 6, 697–711.
- Levy, J.R., Holzbaur, E.L.F., 2008. Dynein drives nuclear rotation during forward progression of motile fibroblasts. *Journal of Cell Science* 121, 3187–3195.
- Li, X., Kuromi, H., Briggs, L., *et al.*, 2010. Bicaudal-D binds clathrin heavy chain to promote its transport and augments synaptic vesicle recycling. *EMBO Journal* 29, 992–1006.
- Lin, J., Okada, K., Raytchev, M., Smith, M.C., Nicastro, D., 2014. Structural mechanism of the dynein power stroke. *Nature Cell Biology* 16, 479–485.
- Lindemann, C.B., Lesich, K.A., 2010. Flagellar and ciliary beating: The proven and the possible. *Journal of Cell Science* 123, 519–528.
- Liu, S., Hard, R., Rankin, S., Hennessey, T., Pennock, D.G., 2004. Disruption of genes encoding predicted inner arm dynein heavy chains causes motility phenotypes in *Tetrahymena*. *Cell Motility and the Cytoskeleton* 59, 201–214.
- Liu, S., Hennessey, T., Rankin, S., Pennock, D.G., 2005. Mutations in genes encoding inner arm dynein heavy chains in *Tetrahymena thermophila* lead to axonemal hypersensitivity to Ca^{2+} . *Cell Motility and the Cytoskeleton* 62, 133–140.
- Loges, N.T., Olbrich, H., Becker-Heck, A., *et al.*, 2009. Deletions and point mutations of LRRC50 cause primary ciliary dyskinesia due to dynein arm defects. *American Journal of Human Genetics* 85, 883–889.
- Loges, N.T., Olbrich, H., Fenske, L., *et al.*, 2008. DNAI2 mutations cause primary ciliary dyskinesia with defects in the outer dynein arm. *American Journal of Human Genetics* 83, 547–558.
- Lupetti, P., Lanzavecchia, S., Mercati, D., *et al.*, 2005. Three-dimensional reconstruction of axonemal outer dynein arms in situ by electron tomography. *Cell Motility and the Cytoskeleton* 62, 69–83.
- Mach, J.M., Lehmann, R., 1997. An Egalitarian-BicaudalD complex is essential for oocyte specification and axis determination in *Drosophila*. *Genes & Development* 11, 423–435.
- Maday, S., Twelvetrees, A.E., Moughamian, A.J., Holzbaur, E.L.F., 2014. Review axonal transport: Cargo-specific mechanisms of motility and regulation. *Neuron* 84, 292–309.
- Mallik, R., Carter, B.C., Lex, S.A., King, S.J., Gross, S.P., 2004. Cytoplasmic dynein functions as a gear in response to load. *Nature* 430, 180–183.
- May, S.R., Ashique, A.M., Karlen, M., *et al.*, 2005. Loss of the retrograde motor for IFT disrupts localization of Smo to cilia and prevents the expression of both activator and repressor functions of Gli. *Developmental Biology* 287, 378–389.
- McKenney, R.J., Huynh, W., Tanenbaum, M.E., Bhabha, G., Vale, R.D., 2014. Activation of cytoplasmic dynein motility by dynactin-cargo adapter complexes. *Science* 345, 337–341.
- McKenney, R.J., Vershinin, M., Kunwar, A., Vallee, R.B., Gross, S.P., 2010. LIS1 and NudE induce a persistent dynein force-producing state. *Cell* 141, 304–314.
- Merdes, A., Ramyar, K., Vechio, J.D., Cleveland, D.W., 1996. A complex of NuMA and cytoplasmic dynein is essential for mitotic spindle assembly. *Cell* 87, 447–458.
- Minoura, I., Kamiya, R., 1995. Strikingly different propulsive forces generated by different dynein-deficient mutants in viscous media. *Cell Motility and the Cytoskeleton* 31, 130–139.
- Mitchell, D.R., Nakatsugawa, M., 2004. Bend propagation drives central pair rotation in *Chlamydomonas reinhardtii* flagella. *Journal of Cell Biology* 166, 709–715.
- Mitchell, D.R., Rosenbaum, J.L., 1985. A motile *Chlamydomonas* flagellar mutant that lacks outer dynein arms. *Journal of Cell Biology* 100, 1228–1234.
- Mitchison, T.J., Maddox, P., Gaetz, J., *et al.*, 2005. Roles of polymerization dynamics, opposed motors, and a tensile element in governing the length of *Xenopus* extract meiotic spindles. *Molecular Biology of the Cell* 16, 3064–3076.
- Mohler, J., Wieschaus, E.F., 1986. Dominant maternal-effect mutations of *Drosophila melanogaster* causing the production of double-abdomen embryos. *Genetics* 112, 803–822.
- Moore, J.K., Sept, D., Cooper, J.A., 2009. Neurodegeneration mutations in dynactin impair dynein-dependent nuclear migration. *Proceedings of the National Academy of Sciences of the United States of America* 106, 5147–5152.
- Morris, R.L., Hoffman, M.P., Obar, R.A., *et al.*, 2006. Analysis of cytoskeletal and motility proteins in the sea urchin genome assembly. *Developmental Biology* 300, 219–237.
- Movassagh, T., Bui, K.H., Sakakibara, H., Oiwa, K., Ishikawa, T., 2010. Nucleotide-induced global conformational changes of flagellar dynein arms revealed by *in situ* analysis. *Nature Structural & Molecular Biology* 17, 761–767.
- Nakano, I., Kobayashi, T., Yoshimura, M., Shingyoji, C., 2003. Central-pair-linked regulation of microtubule sliding by calcium in flagellar axonemes. *Journal of Cell Science* 116, 1627–1636.
- Navarro, C., Puthalakath, H., Adams, J.M., Strasser, A., Lehmann, R., 2004. Egalitarian binds dynein light chain to establish oocyte polarity and maintain oocyte fate. *Nature Cell Biology* 6, 427–435.
- Neesen, J., Kirschner, R., Ochs, M., *et al.*, 2001. Disruption of an inner arm dynein heavy chain gene results in asthenozoospermia and reduced ciliary beat frequency. *Human Molecular Genetics* 10, 1117–1128.
- Neuwald, A.F., Aravind, L., Spouge, J.L., Koonin, E.V., 1999. AAA + : A Class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Research* 9, 27–43.
- Nishikawa, Y., Oyama, T., Kamiya, R., *et al.*, 2014. Structure of the entire stalk region of the dynein motor domain. *Journal of Molecular Biology* 426, 3232–3245.
- Oda, T., Yanagisawa, H., Yagi, T., Kikkawa, M., 2014. Mechanosignaling between central apparatus and radial spokes controls axonemal dynein activity. *Journal of Cell Biology* 204, 807–819.
- Olbrich, H., Häfner, K., Kispert, A., *et al.*, 2002. Mutations in DNAH5 cause primary ciliary dyskinesia and randomization of left-right asymmetry. *Nature Genetics* 30, 143–144.
- Omoto, C.K., Brokaw, C.J., 1985. Bending patterns of *Chlamydomonas* flagella: II. Calcium effects on reactivated *Chlamydomonas* flagella. *Cell Motility* 5, 53–60.
- Omran, H., Kobayashi, D., Olbrich, H., *et al.*, 2008. Ktu/PF13 is required for cytoplasmic pre-assembly of axonemal dyneins. *Nature* 456, 611–616.
- Owa, M., Furuta, A., Usukura, J., *et al.*, 2014. Cooperative binding of the outer arm-docking complex underlies the regular arrangement of outer arm dynein in the axoneme. *Proceedings of the National Academy of Sciences of the United States of America* 111, 9461–9466.
- Palazzo, A.F., Joseph, H.L., Chen, Y.J., *et al.*, 2001. Cdc42, dynein, and dynactin regulate MTOC reorientation independent of Rho-regulated microtubule stabilization. *Current Biology* 11, 1536–1541.
- Pandey, J.P., Smith, D.S., 2011. A Cdk5-dependent switch regulates Lis1/Ndel1/dynein-driven organelle transport in adult axons. *Journal of Neuroscience* 31, 17207–17219.
- Pazour, G.J., Dickert, B.L., Witman, G.B., 1999. The DHC1b (DHC2) isoform of cytoplasmic dynein is required for flagellar assembly. *Journal of Cell Biology* 144, 473–481.
- Pazour, G.J., Wilkerson, C.G., Witman, G.B., 1998. A dynein light chain is essential for the retrograde particle movement of intraflagellar transport (IFT). *Journal of Cell Biology* 141, 979–992.

- Pedersen, H., Rebbe, H., 1975. Absence of arms in the axoneme of immobile human spermatozoa. *Biology of Reproduction* 12, 541–544.
- Pedersen, L.B., Rempel, P., Christensen, S.T., Rosenbaum, J.L., King, S.M., 2007. The lissencephaly protein Lis1 is present in motile mammalian cilia and requires outer arm dynein for targeting to *Chlamydomonas* flagella. *Journal of Cell Science* 120, 858–867.
- Pennarun, G., Escudier, E., Chapelin, C., *et al.*, 1999. Loss-of-function mutations in a human gene related to *Chlamydomonas reinhardtii* dynein IC78 result in primary ciliary dyskinesia. *American Journal of Human Genetics* 65, 1508–1519.
- Perrone, C.A., Tritschler, D., Taulman, P., *et al.*, 2003. A novel dynein light intermediate chain colocalizes with the retrograde motor for intraflagellar transport at sites of axoneme assembly in *chlamydomonas* and mammalian cells. *Molecular Biology of the Cell* 14, 2041–2056.
- Perrone, C.A., Yang, P., O'Toole, E., Sale, W.S., Porter, M.E., 1998. The *Chlamydomonas* IDA7 locus encodes a 140-kDa dynein intermediate chain required to assemble the 11 inner arm complex. *Molecular Biology of the Cell* 9, 3351–3365.
- Pigino, G., Bui, K.H., Maheshwari, A., *et al.*, 2011. Cryoelectron tomography of radial spokes in cilia and flagella. *Journal of Cell Biology* 195, 673–687.
- Piperno, G., Luck, D.J., 1979. An actin-like protein is a component of axonemes from *Chlamydomonas* flagella. *Journal of Biological Chemistry* 254, 2187–2190.
- Porter, M.E., Bower, R., Knott, J.A., Byrd, P., Dentler, W., 1999. Cytoplasmic dynein heavy chain 1b is required for flagellar assembly in *Chlamydomonas*. *Molecular Biology of the Cell* 10, 693–712.
- Porter, M.E., Sale, W.S., 2000. The 9 + 2 axoneme anchors multiple inner arm dyneins and a network of kinases and phosphatases that control motility. *Journal of Cell Biology* 151, F37–F42.
- Qiu, W., Derr, N.D., Goodman, B.S., *et al.*, 2012. Dynein achieves processive motion using both stochastic and coordinated stepping. *Nature Structural & Molecular Biology* 19, 193–200.
- Rai, A.K., Rai, A., Ramaiya, A.J., Jha, R., Mallik, R., 2013. Molecular adaptations allow dynein to generate large collective forces inside cells. *Cell* 152, 172–182.
- Rayment, I., Rypniewski, W., Schmidt-Base, K., 1993. Three-dimensional structure of myosin subfragment-1: A molecular motor. *Science* 261, 50–58.
- Reck-Petersen, S.L., Yildiz, A., Carter, A.P., *et al.*, 2006. Single-molecule analysis of dynein processivity and stepping behavior. *Cell* 126, 335–348.
- Redwine, W.B., Hernandez-Lopez, R., Zou, S., *et al.*, 2012. Structural basis for microtubule binding and release by dynein. *Science* 337, 1532–1536.
- Reed, W., Carson, J.L., Moats-Staats, B.M., *et al.*, 2000. Characterization of an axonemal dynein heavy chain expressed early in airway epithelial ciliogenesis. *American Journal of Respiratory Cell and Molecular Biology* 23, 734–741.
- Riedel-Kruse, I.H., Hilfinger, A., Howard, J., Jülicher, F., 2007. How molecular motors shape the flagellar beat. *HFSP Journal* 1, 192–208.
- Rishal, I., Fainzilber, M., 2014. Axon-soma communication in neuronal injury. *Nature Reviews Neuroscience* 15, 32–42.
- Roberts, A.J., Kon, T., Knight, P.J., Sutoh, K., Burgess, S.A., 2013. Functions and mechanics of dynein motor proteins. *Nature Reviews Molecular Cell Biology* 14, 713–726.
- Roberts, A.J., Malkova, B., Walker, M.L., *et al.*, 2012. ATP-driven remodeling of the linker domain in the dynein motor. *Structure* 20, 1670–1680.
- Roberts, A.J., Numata, N., Walker, M.L., *et al.*, 2009. AAA + Ring and linker swing mechanism in the dynein motor. *Cell* 136, 485–495.
- Rompolas, P., Patel-King, R.S., King, S.M., 2010. An outer arm dynein conformational switch is required for metachronal synchrony of motile cilia in planaria. *Molecular Biology of the Cell* 21, 3669–3679.
- Rompolas, P., Patel-King, R.S., King, S.M., 2012. Association of Lis1 with outer arm dynein is modulated in response to alterations in flagellar motility. *Molecular Biology of the Cell* 23, 3554–3565.
- Ross, J.L., Wallace, K., Shuman, H., Goldman, Y.E., Holzbaur, E.L.F., 2006. Processive bidirectional motion of dynein-dynactin complexes *in vitro*. *Nature Cell Biology* 8, 562–570.
- Sakakibara, H., Kojima, H., Sakai, Y., Katayama, E., Oiwa, K., 1999. Inner-arm dynein c of *Chlamydomonas* flagella is a single-headed processive motor. *Nature* 400, 586–590.
- Sakakibara, H., Mitchell, D.R., Kamiya, R., 1991. A *Chlamydomonas* outer arm dynein mutant missing the alpha heavy chain. *Journal of Cell Biology* 113, 615–622.
- Sakato, M., King, S.M., 2003. Calcium regulates ATP-sensitive microtubule binding by *Chlamydomonas* outer arm dynein. *Journal of Biological Chemistry* 278, 43571–43579.
- Salathe, M., 2007. Regulation of mammalian ciliary beating. *Annual Review of Physiology* 69, 401–422.
- Sale, W.S., Fox, L.A., Smith, E.F., 1993. Assays of axonemal dynein-driven motility. *Methods in Cell Biology* 39, 89–104.
- Samant, S.A., Ogunkua, O., Hui, L., Fossella, J., Pilder, S.H., 2002. The T complex distorter 2 candidate gene, Dnahc8, encodes at least two testis-specific axonemal dynein heavy chains that differ extensively at their amino and carboxyl termini. *Developmental Biology* 250, 24–43.
- Satir, P., 1968. Studies on cilia. 3. Further studies on the cilium tip and a "sliding filament" model of ciliary motility. *Journal of Cell Biology* 39, 77–94.
- Satir, P., Matsuoka, T., 1989. Splitting the ciliary axoneme: Implications for a "switch-point" model of dynein arm activity in ciliary motion. *Cell Motility and the Cytoskeleton* 14, 345–358.
- Schlager, M.A., Hoang, H.T., Urnavicius, L., Bullock, S.L., Carter, A.P., 2014. *In vitro* reconstitution of a highly processive recombinant human dynein complex. *EMBO Journal* 33, 1–14.
- Schmidt, H., Gleave, E.S., Carter, A.P., 2012. Insights into dynein motor domain function from a 3.3-Å crystal structure. *Nature Structural & Molecular Biology* 19, 492–497.
- Schmidt, H., Zalyte, R., Urnavicius, L., Carter, A.P., 2014. Structure of human cytoplasmic dynein-2 primed for its power stroke. *Nature* 518, 435–438.
- Schroeder, H.W., Mitchell, C., Shuman, H., Holzbaur, E.L.F., Goldman, Y.E., 2010. Motor number controls cargo switching at actin-microtubule intersections *in vitro*. *Current Biology* 20, 687–696.
- Schroer, T.A., 2004. Dynactin. *Annual Review of Cell and Developmental Biology* 20, 759–779.
- Shih, S.M., Engel, B.D., Kocabas, F., *et al.*, 2013. Intraflagellar transport drives flagellar surface motility. *eLife* 2013, 1–19.
- Shingyoji, C., 2009. Analysis of the role of nucleotides in axonemal dynein function. *Methods in Cell Biology* 92, 113–131.
- Shingyoji, C., Murakami, A., Takahashi, K., 1977. Local reactivation of triton-extracted flagella by iontophoretic application of ATP. *Nature* 265, 269–270.
- Shiroguchi, K., Toyoshima, Y.Y., 2001. Regulation of monomeric dynein activity by ATP and ADP concentrations. *Cell Motility and the Cytoskeleton* 49, 189–199.
- Shu, T., Ayala, R., Nguyen, M.D., *et al.*, 2004. Ndel1 operates in a common pathway with LIS1 and cytoplasmic dynein to regulate cortical neuronal positioning. *Neuron* 44, 263–277.
- Shubeita, G.T., Tran, S.L., Xu, J., *et al.*, 2008. Consequences of motor copy number on the intracellular transport of kinesin-1-driven lipid droplets. *Cell* 135, 1098–1107.
- Signor, D., Wedatman, K.P., Orozco, J.T., *et al.*, 1999. Role of a class DHC1b dynein in retrograde transport of IFT motors and IFT raft particles along cilia, but not dendrites, in chemosensory neurons of living *Caenorhabditis elegans*. *Journal of Cell Biology* 147, 519–530.
- Siller, K.H., Doe, C.Q., 2008. Lis1/dynactin regulates metaphase spindle orientation in *Drosophila* neuroblasts. *Developmental Biology* 319, 1–9.
- Smith, E.F., 2002. Regulation of flagellar dynein by calcium and a role for an axonemal calmodulin and calmodulin-dependent kinase. *Molecular Biology of the Cell* 13, 3303–3313.
- Smith, E.F., Sale, W.S., 1992. Regulation of dynein-driven microtubule sliding by the radial spokes in flagella. *Science* 257, 1557–1559.
- Smith, E.F., Yang, P., 2004. The radial spokes and central apparatus: Mechanochemical transducers that regulate flagellar motility. *Cell Motility and the Cytoskeleton* 57, 8–17.
- Spudich, J.A., 2006. Molecular motors take tension in stride. *Cell* 126, 242–244.
- Steward, R., Nüsslein-Volhard, C., 1986. The genetics of the dorsal-bicaudal-D region of *Drosophila melanogaster*. *Genetics* 113, 665–678.
- Stinson, B.M., Nager, A.R., Glynn, S.E., *et al.*, 2013. Nucleotide binding and conformational switching in the hexameric ring of a AAA + machine. *Cell* 153, 628–639.
- Sullivan-Brown, J., Schottenfeld, J., Okabe, N., *et al.*, 2008. Zebrafish mutations affecting cilia motility share similar cystic phenotypes and suggest a mechanism of cyst formation that differs from pkd2 morphants. *Developmental Biology* 314, 261–275.
- Summers, K.E., Gibbons, I.R., 1971. Adenosine triphosphate-induced sliding of tubules in trypsin-treated flagella of sea-urchin sperm. *Proceedings of the National Academy of Sciences of the United States of America* 68, 3092–3096.
- Supp, D.M., Witte, D.P., Potter, S.S., Brueckner, M., 1997. Mutation of an axonemal dynein affects left-right asymmetry in *inversus viscerum* mice. *Nature* 389, 963–966.
- Takada, S., Wilkerson, C.G., Wakabayashi, K., Kamiya, R., Witman, G.B., 2002. The outer dynein arm-docking complex: Composition and characterization of a subunit (oda1) necessary for outer arm assembly. *Molecular Biology of the Cell* 13, 1015–1029.
- Tanaka, T., Sereno, F.F., Higgins, C., *et al.*, 2004. Lis1 and doublecortin function with dynein to mediate coupling of the nucleus to the centrosome in neuronal migration. *Journal of Cell Biology* 165, 709–721.

- Tanenbaum, M.E., Macürek, L., Galjart, N., Medema, R.H., 2008. Dynein, Lis1 and CLIP-170 counteract Eg5-dependent centrosome separation during bipolar spindle assembly. *EMBO Journal* 27, 3235–3245.
- Toba, S., Watanabe, T.M., Yamaguchi-Okimoto, L., Toyoshima, Y.Y., Higuchi, H., 2006. Overlapping hand-over-hand mechanism of single molecular motility of cytoplasmic dynein. *Proceedings of the National Academy of Sciences of the United States of America* 103, 5741–5745.
- Tsai, J.-W., Bremner, K.H., Vallee, R.B., 2007. Dual subcellular roles for LIS1 and dynein in radial neuronal migration in live brain tissue. *Nature Neuroscience* 10, 970–979.
- Tynan, S.H., Gee, M.A., Vallee, R.B., 2000. Distinct but overlapping sites within the cytoplasmic dynein heavy chain for dimerization and for intermediate chain and light intermediate chain binding. *Journal of Biological Chemistry* 275, 32769–32774.
- Ueno, H., Bui, K.H., Ishikawa, T., *et al.*, 2014. Structure of dimeric axonemal dynein in cilia suggests an alternative mechanism of force generation. *Cytoskeleton* 71, 412–422.
- Ueno, H., Yasunaga, T., Shingyoji, C., Hirose, K., 2008. Dynein pulls microtubules without rotating its stalk. *Proceedings of the National Academy of Sciences of the United States of America* 105, 19702–19707.
- Vale, R.D., Milligan, R.A., 2000. The way things move: Looking under the hood of molecular motor proteins. *Science* 288, 88–95.
- Vallee, R.B., Williams, J.C., Varma, D., Barnhart, L.E., 2004. Dynein: An ancient motor protein involved in multiple modes of transport. *Journal of Neurobiology* 58, 189–200.
- Varma, D., Monzo, P., Stehman, S.A., Vallee, R.B., 2008. Direct role of dynein motor in stable kinetochore-microtubule attachment, orientation, and alignment. *Journal of Cell Biology* 182, 1045–1054.
- Vaughan, P.S., Miura, P., Henderson, M., Byrne, B., Vaughan, K.T., 2002. A role for regulated binding of p150(Glued) to microtubule plus ends in organelle transport. *Journal of Cell Biology* 158, 305–319.
- Wais-Steider, J., Satir, P., 1979. Effect of vanadate on gill cilia: Switching mechanism in ciliary beat. *Journal of Supramolecular Structure* 11, 339–347.
- Wakabayashi, K., 2009. Analysis of redox-sensitive dynein components. *Methods in Cell Biology* 92, 153–161.
- Walter, W.J., Brenner, B., Steffen, W., 2010. Cytoplasmic dynein is not a conventional processive motor. *Journal of Structural Biology* 170, 266–269.
- Warner, F.D., Satir, P., 1974. The structural basis of ciliary bend formation. Radial spoke positional changes accompanying microtubule sliding. *Journal of Cell Biology* 63, 35–63.
- Waterman-Storer, C.M., Karki, S., Holzbaur, E.L., 1995. The p150Glued component of the dynactin complex binds to both microtubules and the actin-related protein centractin (Arp-1). *Proceedings of the National Academy of Sciences of the United States of America* 92, 1634–1638.
- Watson, P., Forster, R., Palmer, K.J., Pepperkok, R., Stephens, D.J., 2005. Coupling of ER exit to microtubules through direct interaction of COPII with dynactin. *Nature Cell Biology* 7, 48–55.
- Wickstead, B., Gull, K., 2007. Dyneins across eukaryotes: A comparative genomic analysis. *Traffic (Copenhagen, Denmark)* 8, 1708–1721.
- Wickstead, B., Gull, K., 2012. Evolutionary biology of dyneins. In: King, S.M. (Ed.), *Dyneins — Structure, Biology and Disease*. Academic Press, pp. 89–121. ISBN: 978-0-12-382004-4.
- Wilkes, D.E., Watson, H.E., Mitchell, D.R., Asai, D.J., 2008. Twenty-five dyneins in *Tetrahymena*: A re-examination of the multidynein hypothesis. *Cell Motility and the Cytoskeleton* 65, 342–351.
- Wirschell, M., Hendrickson, T., Sale, W.S., 2007. Keeping an eye on I1: I1 dynein as a model for flagellar dynein assembly and regulation. *Cell Motility and the Cytoskeleton* 64, 569–579.
- Wirschell, M., Yang, C., Yang, P., *et al.*, 2009. IC97 is a novel intermediate chain of I1 dynein that interacts with tubulin and regulates interdoubt sliding. *Molecular Biology of the Cell* 20, 3044–3054.
- Woehlke, G., Ruby, A.K., Hart, C.L., *et al.*, 1997. Microtubule interaction site of the kinesin motor. *Cell* 90, 207–216.
- Woodard, G.E., Huang, N.-N., Cho, H., *et al.*, 2010. Ric-8A and Gi alpha recruit LGN, NuMA, and dynein to the cell cortex to help orient the mitotic spindle. *Molecular and Cellular Biology* 30, 3519–3530.
- Woolley, D.M., 2010. Flagellar oscillation: A commentary on proposed mechanisms. *Biological reviews of the Cambridge Philosophical Society* 85, 453–470.
- Woolley, D.M., Neesen, J., Vernon, G.G., 2005. Further studies on knockout mice lacking a functional dynein heavy chain (MDHC7). 2. A developmental explanation for the asthenozoospermia. *Cell Motility and the Cytoskeleton* 61, 74–82.
- Wu, H., Maciejewski, M.W., Marintchev, A., *et al.*, 2000. Solution structure of a dynein motor domain associated light chain. *Nature Structural & Molecular Biology* 7, 575–579.
- Xiang, X., Beckwith, S.M., Morris, N.R., 1994. Cytoplasmic dynein is involved in nuclear migration in *Aspergillus nidulans*. *Proceedings of the National Academy of Sciences of the United States of America* 91, 2100–2104.
- Yagi, T., Kamiya, R., 2000. Vigorous beating of *Chlamydomonas* axonemes lacking central pair/radial spoke structures in the presence of salts and organic compounds. *Cell Motility and the Cytoskeleton* 46, 190–199.
- Yagi, T., Uematsu, K., Liu, Z., Kamiya, R., 2009. Identification of dyneins that localize exclusively to the proximal portion of *Chlamydomonas* flagella. *Journal of Cell Science* 122, 1306–1314.
- Yamamoto, R., Hirono, M., Kamiya, R., 2010. Discrete PIH proteins function in the cytoplasmic preassembly of different subsets of axonemal dyneins. *Journal of Cell Biology* 190, 65–71.
- Yamamoto, R., Yanagisawa, H., Yagi, T., Kamiya, R., 2006. A novel subunit of axonemal dynein conserved among lower and higher eukaryotes. *FEBS Letters* 580, 6357–6360.
- Yamamoto, R., Yanagisawa, H.-A., Yagi, T., Kamiya, R., 2008. Novel 44-kilodalton subunit of axonemal dynein conserved from *chlamydomonas* to mammals. *Eukaryotic Cell* 7, 154–161.
- Yanagisawa, H.A., Kamiya, R., 2001. Association between actin and light chains in *Chlamydomonas* flagellar inner-arm dyneins. *Biochemical and Biophysical Research Communications* 288, 443–447.
- Yang, P., Diener, D.R., Rosenbaum, J.L., Sale, W.S., 2001. Localization of calmodulin and dynein light chain LC8 in flagellar radial spokes. *Journal of Cell Biology* 153, 1315–1326.
- Yang, P., Diener, D.R., Yang, C., *et al.*, 2006. Radial spoke proteins of *Chlamydomonas* flagella. *Journal of Cell Science* 119, 1165–1174.
- Yang, P., Fox, L., Colbran, R.J., Sale, W.S., 2000. Protein phosphatases PP1 and PP2A are located in distinct positions in the *Chlamydomonas* flagellar axoneme. *Journal of Cell Science* 113 (Pt 1), 91–102.
- Yang, P., Sale, W.S., 1998. The Mr 140,000 intermediate chain of *Chlamydomonas* flagellar inner arm dynein is a WD-repeat protein implicated in dynein arm anchoring. *Molecular Biology of the Cell* 9, 3335–3349.
- Yang, P., Sale, W.S., 2000. Casein kinase I is anchored on axonemal doublet microtubules and regulates flagellar dynein phosphorylation and activity. *Journal of Biological Chemistry* 275, 18905–18912.
- Yang, Z., Tulu, U.S., Wadsworth, P., Rieder, C.L., 2007. Kinetochore dynein is required for chromosome motion and congression independent of the spindle checkpoint. *Current Biology* 17, 973–980.
- Yildiz, A., Forkey, J.N., McKinney, S.A., *et al.*, 2003. Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization. *Science* 300, 2061–2065.
- Yildiz, A., Tomishige, M., Gennerich, A., Vale, R.D., 2008. Intramolecular strain coordinates kinesin stepping behavior along microtubules. *Cell* 134, 1030–1041.
- Yildiz, A., Tomishige, M., Vale, R.D., Selvin, P.R., 2004. Kinesin walks hand-over-hand. *Science* 303, 676–678.
- Yoshimura, A., Nakano, I., Shingyoji, C., 2007. Inhibition by ATP and activation by ADP in the regulation of flagellar movement in sea urchin sperm. *Cell Motility and the Cytoskeleton* 64, 777–793.
- Zariwala, M.A., Gee, H.Y., Kurkowiak, M., *et al.*, 2013. ZMYND10 is mutated in primary ciliary dyskinesia and interacts with LRRC6. *American Journal of Human Genetics* 93, 336–345.

Further Reading

- Hirose, K., Amos, L.A., 2012. *Handbook of Dynein*. Singapore: Pan Stanford. ISDN B008468FV1.
- King, S.M., 2011. *Dyneins: Structure, Biology and Disease*. Boston/Amsterdam: Academic Press. ISDN B007GVU3L2.

CYTOSKELETON AND MOTORS: COMPLEX CYTOSKELETAL STRUCTURES

Contents

- The Mitotic Spindle**
- Centrioles and the Centrosome**
- Cilia and Flagella**
- Skeletal Muscle**
- Filopodia and Lamellipodia**
- The Extracellular Matrix**
- Cell–Cell Adhesion and the Cytoskeleton**
- Cell Adhesion to the Extracellular Matrix**
- Cell Migration**
- Rho GTPases**
- Cell Polarity**

The Mitotic Spindle

M Barisic and H Maiato, Instituto de Biologia Molecular e Celular, Porto, Portugal

© 2016 Elsevier Inc. All rights reserved.

Architecture and Dynamics of the Mitotic Spindle

The mitotic spindle is a bipolar, spindle-shaped molecular ‘machine’ that is assembled at the onset of mitosis. Its main function is to accurately distribute the chromosomes into two daughter cells during cell division (for detailed recent reviews see [Helmke et al., 2013](#); [McIntosh et al., 2012](#)). Any failure in this critical role could lead to development of diseases, including cancer ([Noatynska et al., 2012](#); [Orticello et al., 2014](#)), emphasizing the robustness of the mitotic spindle apparatus as an absolute requirement for normal tissue development.

The first insight into the shape and structure of the mitotic spindle and its connection with chromosomes was provided in the early 1880s by the detailed drawings of Walther Flemming (reproduced in [Flemming, 1965](#)), making this research field over a century old. More than 50 years later, in the late 1930s, followed up in the 1950s, it became possible to confirm the existence of this fibrous structure in living cells owing to the invention and implementation of polarized light microscopy ([Schmidt, 1937](#); [Inoue, 1953](#)). In a series of experiments it was shown that spindle fibers, which we now know to be formed by microtubules, could disappear by lowering the temperature or by adding the disassembling agent colchicine (reviewed in [Inoue, 1981](#)). This process was reversible if the temperature was increased or if the colchicine was washed out, postulating the dynamic and polymeric nature of microtubules. By isolating and identifying the target of colchicine, it was confirmed that microtubules are indeed polymers, with a protein called tubulin being their main building unit ([Borisy and Taylor, 1967](#); [Mohri, 1968](#)).

Research over the following decades revealed that each microtubule is a hollow cylindrical structure normally composed of 13 laterally associated protofilaments that are assembled out of heterodimers of α - and β -tubulin molecules. The head-to-tail arrangement of α and β subunits causes the intrinsic polarity to microtubules, with α -tubulin exposed at the so-called minus end and β -tubulin at the plus end. Although tubulins are GTP-binding proteins, only the β subunit acts as a GTPase once a new heterodimer has been added. After GTP has been hydrolyzed into GDP, the heterodimers have a tendency to bend, which destabilizes the linear array and results in plus-end depolymerization. This, however, happens with a delay that is long enough to allow the accumulation of GTP-bound β -tubulin at the plus end and formation of a GTP-cap, which keeps the heterodimers straight and, consequently, stabilizes the protofilaments. Lateral interactions between the protofilaments prevent previously integrated heterodimers from bending even when their GTP has been hydrolyzed, resulting in energy storage along the microtubule lattice (reviewed in [Howard and Hyman, 2009](#)). Sudden loss of the GTP-cap causes curving of the tubulin heterodimers and rapid depolymerization of microtubule plus ends ([Chretien et al., 1995](#); [Mandelkow et al., 1991](#)), accompanied by the release of stored energy that can be harvested to facilitate cell division, e.g. by promoting the movement of chromosomes bound to depolymerizing microtubules. This transition from growth to shrinkage is termed ‘catastrophe,’ while regrowth of shrinking microtubules upon the retrieval of their GTP-cap is called ‘rescue.’ The ability of microtubules to rapidly switch between these two states is known as ‘dynamic instability’ ([Mitchison](#)

and Kirschner, 1984). However, the dynamic instability of microtubules is not uniform across the different stages of the cell cycle, becoming more dynamic during mitosis.

Different Paths of Spindle Assembly

In higher eukaryotes, mitotic spindle assembly starts with nuclear envelope breakdown in the prophase and relies largely on dynamic instability. Almost all animal cells contain centrosomes as the main microtubule organizing centers; these are duplicated in the interphase and separated at the onset of mitosis. Centrosomes consist of a pair of centrioles embedded in a pericentriolar material in which multiple copies of the γ -tubulin ring complex (γ -TuRC) are clustered and linked by coiled-coil proteins, such as pericentrin and ninein (Mennella *et al.*, 2014). γ -TuRC makes microtubule minus ends less dynamic by anchoring them at the centrosome, while providing the template for microtubule growth and defining their polarity (reviewed in Kollman *et al.*, 2011).

As soon as the nuclear envelope breaks, the barrier between the cytoplasm and the nucleus vanishes and microtubules can freely interact with chromosomes. At this stage, centrosomes mature and are able to nucleate 5–10 times more microtubules than in the interphase (Snyder and McIntosh, 1975). These dynamic microtubules nucleated from the centrosomes are known as astral microtubules. As a result of centrosome maturation and more frequent switches between growth and shrinkage due to dynamic instability, astral microtubules greatly increase their chances of interacting with chromosomes. This interaction is stabilized at the kinetochore, a structural component at the centromeric region built from more than a hundred different proteins (Cheeseman and Desai, 2008). Such random exploration of the cell space by dynamic astral microtubules in their search for chromosomes is known as ‘search-and-capture’ (Kirschner and Mitchison, 1986). However, computational studies have shown that spindle assembly could not rely exclusively on dynamic instability, since search-and-capture of all kinetochores would require several hours, in contrast to the observed spindle assembly time in living mitotic cells (Wollman *et al.*, 2005; Paul *et al.*, 2009). High-resolution time-lapse imaging, followed by 3D reconstructions, revealed that chromosome positioning during the prometaphase significantly facilitates search-and-capture and chromosome congression (Magidson *et al.*, 2011). This study showed that polar ejection forces (PEFs) generated by the plus-end directed motor proteins chromokinesins promote a toroid, ring-like arrangement of chromosomes that are laterally attached to the surface of the forming spindle. This promotes kinetochore exposure to a microtubule-dense area while preventing kinetochores from being shielded by other chromosomes.

The absence of centrosomes in most meiotic and plant spindles suggests, however, the existence of an alternative pathway of spindle assembly. The ‘self-organization’ of spindle assembly was clearly revealed using *Xenopus* egg extracts. In this system, spindles could be successfully formed, even in the absence of centrosomes, just by adding either whole chromosomes or even only chromatin-coated beads to the cell extracts (Heald *et al.*, 1996). Microtubule polymerization

around chromosomes is driven by the stabilizing effect of GTP-bound Ran molecules (Carazo-Salas *et al.*, 1999, 2001), which are brought to chromatin through their interaction with the guanine nucleotide exchange factor RCC1, establishing a concentration gradient throughout the cytoplasm (Kalab *et al.*, 2002). This Ran-GTP gradient is thought to release spindle assembly factors such as TPX2 from the inhibitory effect of Importin, favoring microtubule nucleation (Gruss and Vernos, 2004). However, in spite of a well-documented role of the Ran-GTP gradient in acentrosomal spindle assembly in *Xenopus* eggs, in other systems it does not seem to play an essential role (Dumont *et al.*, 2007; O’Connell *et al.*, 2009; Moutinho-Pereira *et al.*, 2013).

Although the ability of spindle assembly around chromatin-coated beads excluded an essential role for kinetochores in this process (Heald *et al.*, 1996), chromosome-based microtubule polymerization is strongest at the kinetochores (Pease, 1946; Snyder and McIntosh, 1975; Witt *et al.*, 1980; Maiato *et al.*, 2004a; Khodjakov *et al.*, 2003; O’Connell *et al.*, 2009). Careful analysis by serial section electron microscopy revealed that microtubule polymerization is initiated in close proximity to kinetochores (Witt *et al.*, 1980). Kinetochores then capture and stably attach those short-growing microtubule plus ends, for which kinetochores have higher affinity (Huitorel and Kirschner, 1988). This allows for the correct direction of their further growth (with the microtubule plus end situated at the kinetochore and the minus end oriented toward the pole) (Maiato *et al.*, 2004a).

Importantly, chromatin-based microtubule polymerization is complementary to search-and-capture in cells containing centrosomes. Remarkably, spindle formation was still occurring in cells with mutations that destroyed the functionality of centrosomes (Bonaccorsi *et al.*, 1998; Basto *et al.*, 2006; Megraw *et al.*, 2001; Debec *et al.*, 1982; Sir *et al.*, 2013), as well as in the cells in which centrosomes were physically removed or destroyed (Khodjakov *et al.*, 2000; Hinchcliffe *et al.*, 2001). However, the function of these spindles was somewhat compromised, showing that, even though they are not essential for spindle assembly, centrosomes are still important for the accuracy of cell division in some systems that normally have them (Khodjakov and Rieder, 2001; Sir *et al.*, 2013).

Computer-based simulations showed that a two-motor system would be sufficient to assemble the spindle in a chromatin-based manner. In this system, which the authors named ‘slide-and-cluster,’ a plus end-directed motor slides microtubules from the chromatin toward the pole, where they are clustered and focused by a minus end-directed motor (Burbank *et al.*, 2007). This model is in line with observations that some spindles are assembled as ‘tiled arrays’ from a larger number of short microtubules that are cross-linked by motor proteins and MAPs (Yang *et al.*, 2007a). Using laser microsurgery to ablate microtubules inside the large spindles of *Xenopus* egg extracts, it was recently shown that these spindles consist of microtubules of inhomogeneous lengths, with their plus and minus ends distributed throughout the spindle, indicating the importance of spatial variation in microtubule nucleation combined with the local transport of short microtubules (Brugues *et al.*, 2012).

An additional mechanism, which might significantly contribute to the efficiency of mitotic spindle assembly, depends

on non-centrosome- and non-chromatin-mediated microtubule nucleation initiated by Augmin. Augmin is an eight-subunit protein complex that recruits γ -TuRC to the outer lattice of preexisting microtubules, initiating their branching through the polymerization of new microtubules from within the spindle (Goshima *et al.*, 2008; Uehara *et al.*, 2009; Kamasaki *et al.*, 2013). Interestingly, RNAi-mediated knock-down of Augmin showed more severe spindle phenotypes in systems without functional centrosomes (Petry *et al.*, 2011; Goshima *et al.*, 2008; Uehara *et al.*, 2009), highlighting the importance of the interplay between different assembly pathways in the formation of an efficient spindle machinery.

The Role of Motors and MAPs in Mitotic Spindle Assembly

Microtubules in living cells exhibit more dynamic behavior compared to microtubules reconstituted *in vitro*, which suggests the existence of additional stabilizing and destabilizing factors in living cells (Desai and Mitchison, 1997). A large number of microtubule-associated proteins (MAPs) and motor proteins that affect the stability of microtubules were revealed through biochemical and proteomic approaches, as well as genetic and genome-wide RNAi-based screens (Bonner *et al.*, 2011; Gache *et al.*, 2010; Goshima *et al.*, 2007; Moutinho-Pereira *et al.*, 2013; Maiato *et al.*, 2004b). While MAPs directly influence microtubule stability through their interaction with the microtubule lattice and/or ends, motor proteins use ATP hydrolysis to provide the energy used either to destabilize the microtubule lattice or to transport different cargos along the microtubules, both of which have an impact on the efficiency of mitotic spindle assembly.

Microtubule Cross-Linkers

Another important property of several different MAPs and motors that contributes to spindle assembly and maintenance is their ability to cross-link microtubules. One of the most studied microtubule cross-linkers is Eg5, a member of the kinesin-5 protein family. Eg5 is a plus end-directed motor that assembles into homo-tetramers with two-motor domains located on each of the opposite sides of the complex (Cole *et al.*, 1994). This allows it to bind two different microtubules and, because of its preference for antiparallel microtubules (Van Den Wildenberg *et al.*, 2008), to slide their minus ends away from each other (Sharp *et al.*, 1999). Inhibition of Eg5 prevents assembly of a bipolar spindle, leading to formation of monopolar asters (Kapoor *et al.*, 2000; Blangy *et al.*, 1995). However, once spindle bipolarity has been achieved, Eg5 is not required to maintain it, due to the cooperative activity of another homo-tetramer microtubule cross-linker, Kif15, which was proposed to share the ability to slide antiparallel microtubules with Eg5 (Tanenbaum *et al.*, 2009; Vanneste *et al.*, 2009; Drechsler *et al.*, 2014).

An additional cross-linker of antiparallel microtubules is PRC1, which is involved in the assembly of the central spindle during late mitosis. During early mitosis, CDK1 and Plk1 phosphorylate PRC1, keeping it mostly inactive (Zhu *et al.*, 2006; Hu *et al.*, 2012). At anaphase onset it is activated

by dephosphorylation and then brought to the plus ends of antiparallel microtubules of the midzone by the plus end-directed motor protein Kif4A (Zhu and Jiang, 2005). It was recently shown that Aurora B kinase regulates relocation of Kif4A from the chromosome arms to the spindle midzone, as well as its interaction with PRC1 and the ability to suppress microtubule dynamics, which plays an important role in the regulation of spindle elongation and cytokinesis (Nunes Bastos *et al.*, 2013).

In addition to the described microtubule plus end-directed motors, two minus end-directed motor proteins also have the ability to cross-link microtubules and contribute to efficient mitotic spindle assembly. Dynein, a 1.6-MDa multisubunit protein complex (Allan, 2011), and its binding partner dynactin recruit NuMA to the spindle poles to cross-link and focus microtubules (Merdes *et al.*, 1996). Moreover, a recent *in vitro* study showed that Dynein alone is also able to cross-link and slide antiparallel microtubules (Tanenbaum *et al.*, 2013), resulting in the generation of an inward force within the spindle and further explaining the antagonizing effect on outward microtubule sliding forces generated by Eg5 (Ferenz *et al.*, 2009) and possibly Kif15. The other minus end-directed motor protein involved in spindle assembly is HSET (Kinesin 14). This kinesin has the ability to cross-link and slide antiparallel microtubules, generating an inward force within the spindle, similar to that proposed for Dynein (Fink *et al.*, 2009; Mountain *et al.*, 1999), and is required for kinesin-5-driven microtubule sliding (Hepperla *et al.*, 2014). Thus, without the assistance of kinesin-14, kinesin-5 fails to cross-link microtubules and cannot generate the outward force required for spindle elongation, resulting in short spindles.

Microtubule Polymerization-Promoting Proteins

Many MAPs contribute to spindle assembly and maintenance through their stabilizing effect on microtubules. A significant fraction of MAPs belong to the microtubule plus end tracking protein family or +TIPs, including members of the CLIP, CLASP, and TOG protein families (Ferreira *et al.*, 2014). CLIP170 stabilizes microtubules by promoting rescue events (Arnal *et al.*, 2004; Komarova *et al.*, 2002) and helps to maintain spindle bipolarity by counteracting outward microtubule sliding through interaction with Dynein (Tanenbaum *et al.*, 2008). Proteins of the CLASP family (CLASP1 and CLASP2 in mammals) also promote microtubule rescue events (Akhmanova *et al.*, 2001; Al-Bassam *et al.*, 2010). Depletion of CLASPs resulted in shorter, collapsed, and multipolar spindles (Maiato *et al.*, 2003, 2002; Logarinho *et al.*, 2012; Mimori-Kiyosue *et al.*, 2006; Lemos *et al.*, 2000; Pereira *et al.*, 2006), highlighting their role in bipolar spindle organization. The *Xenopus* member of the TOG protein family, XMAP215, promotes microtubule polymerization at the plus ends (Gard and Kirschner, 1987; Brouhard *et al.*, 2008), and its depletion prevents bipolar spindle assembly, resulting in short spindles (Tournéize *et al.*, 2000). RNAi-mediated knockdown of its mammalian homolog, ch-TOG, also affected mitotic spindle assembly, resulting in disorganized and multipolar spindles (Gergely *et al.*, 2003). In addition to its growth-promoting effect, ch-TOG also contributes to microtubule stabilization by

protecting kinetochore microtubules from depolymerization by MCAK (Barr and Gergely, 2008).

Another protein that specifically stabilizes k-fibers is HURP. HURP is localized to kinetochore microtubules in a Ran-GTP-dependent manner (Sillje *et al.*, 2006) and promotes chromatin-induced spindle assembly (Casanova *et al.*, 2008). HURP depletion impaired the stability of k-fibers, leading to chromosome congression problems (Wong and Fang, 2006; Sillje *et al.*, 2006), while its overexpression increased microtubule stability (Wong and Fang, 2006). Depletion of its *Drosophila* ortholog, Mars, resulted in mitotic spindle malformations due to the lack of kinetochore microtubules (Yang and Fan, 2008).

Besides the plus ends, MAPs are able to stabilize microtubule minus ends as well. Both Patronin (Goodwin and Vale, 2010; Hendershott and Vale, 2014; Jiang *et al.*, 2014) and MCRC1 (Meunier and Vernos, 2011; Petry and Vale, 2011) localize specifically to the microtubule minus ends, where they prevent their kinesin-13-mediated depolymerization.

Microtubule Depolymerization-Promoting Proteins

The best studied member of the kinesin-13 motor protein family is MCAK, which localizes to the spindle poles, centromeres, kinetochores, and microtubule plus ends, where it can use the energy from ATP hydrolysis to depolymerize microtubules from both ends (Ems-McClung and Walczak, 2010). Its ortholog in *Xenopus*, XKCM1, was shown to negatively regulate microtubule growth, as its depletion in egg extracts interfered with spindle assembly, creating abnormally long microtubules (Walczak *et al.*, 1996). In mammals, microtubule length was less affected in MCAK-depleted cells, but strongly increased by knocking down Kif2A, another member of the kinesin-13 protein family (reviewed in Ems-McClung and Walczak, 2010), emphasizing a conserved role of kinesin-13 in mitotic spindle assembly. *In vitro* reconstitution experiments with MCAK and Kif2A *Xenopus* orthologs demonstrated their direct microtubule-destabilizing activity, showing that these proteins were able to completely depolymerize chemically stabilized microtubules (Desai *et al.*, 1999).

Microtubule growth is also regulated by members of the kinesin-8 motor protein family (Kif18A in mammals), as their depletion from cells also results in elongated microtubules. Kinesin-8 proteins are highly processive motors, which accumulate at microtubule plus ends and, once the plus ends become saturated, they initiate microtubule disassembly in a length-dependent manner (reviewed in Wordeman, 2010).

Finally, the microtubule-severing proteins Katanin, Spastin, and Fidgetin disassemble microtubules in a completely different manner, using ATP hydrolysis to cut microtubules into shorter fragments. They bind to the tubulin C-terminal tails at the surface of the microtubule lattice and initiate microtubule depolymerization by removing tubulin subunits from it (Sharp and Ross, 2012). Loss of katanin in *Caenorhabditis elegans* resulted in disorganized spindles with fewer microtubules, but this phenotype is weaker in mammals, probably due to the presence of centrosomes that might compensate by increasing microtubule nucleation (reviewed in Sharp and Ross, 2012). In *Drosophila* S2 cells, microtubule-severing proteins also play a role in poleward chromosomal movements

during anaphase A by inducing depolymerization of kinetochore microtubules' plus and minus ends (Zhang *et al.*, 2007).

Mitotic Spindle Properties

The mitotic spindle is a highly dynamic molecular machine, in which forces generated by microtubule dynamics and motor proteins must balance out at steady state. Next we discuss some of the spindle properties that result from this balance.

Microtubule Turnover

Development of microscopy techniques such as photo-bleaching, photo-activation, and photo-conversion have revealed that spindle microtubules turn over rapidly, exhibiting typical half-lives of several seconds to several minutes. For instance, astral microtubules are less stable than kinetochore microtubules. Microtubule turnover does not differ only between different types of microtubules, but also varies within the same microtubule population depending on the mitotic stage. Accordingly, kinetochore microtubules become more stable toward the later stages of mitosis, with their turnover decreasing from prometaphase to metaphase and finally to anaphase (Zhai *et al.*, 1995; Gorbisky and Borisy, 1989; Kabeche and Compton, 2013). Kinetochore microtubule turnover depends on the tension established between sister kinetochores (Nicklas and Koch, 1969), and on CDK1 kinase activity regulated by Cyclins A and B (Kabeche and Compton, 2013).

Poleward Microtubule Flux

In metaphase, spindle size and shape is at steady state, but microtubules remain highly dynamic. This translates into another dynamic property, the so-called microtubule poleward flux, discovered by photo-activation of microtubules labeled with caged fluorescent tubulin (Mitchison, 1989). Although undetectable in yeast, this feature is highly conserved in higher eukaryotes. It is defined by the translocation of tubulin dimers added at the microtubule plus ends and their simultaneous removal at the minus ends (Mitchison *et al.*, 1986), generating a continuous poleward motion of tubulin subunits inside spindle microtubules. Further advances in the investigation of this cellular phenomenon were achieved using fluorescence speckle microscopy (FSM) (reviewed in Barisic *et al.*, 2012). FSM-based studies revealed that different classes of microtubules display different flux rates, from which the astral microtubules did not flux at all (Waterman-Storer *et al.*, 1998; Maddox *et al.*, 2003; LaFountain *et al.*, 2004).

Although it is widely accepted that the flux of the interpolar microtubules is driven by kinesin Eg5, which interconnects and slides antiparallel microtubules, the origin of kinetochore microtubule flux is still debatable. One view is that it is driven by kinesin-13-mediated depolymerization at the microtubule minus ends (reviewed in Kwok and Kapoor, 2007; Shimamoto *et al.*, 2011). Another view is that flux relies on the coupling of microtubules within the spindle (Shimamoto *et al.*, 2011; Itabashi *et al.*, 2009), likely through the action of microtubule

cross-linking proteins (see above). In this model, microtubule cross-linkers transduce the force generated by motor-driven sliding of interpolar microtubules (Brust-Mascher and Scholey, 2002; Matos *et al.*, 2009; Pereira and Maiato, 2012). Interestingly, RNAi-mediated depletion of chromokinesin Kif4A also decreases kinetochore microtubule flux, but it remains unknown whether this is a direct role or results from unbalanced spindle forces (Wandke *et al.*, 2012). Kif4A is localized exclusively at the chromosome arms during metaphase, where it interacts with non-kinetochore microtubules and actively participates in the generation of polar ejection forces (PEFs) (Barisic *et al.*, 2014). As chromosomes align to the metaphase plate, Kif4A cannot push chromosomes further away from the pole due to equivalent forces applied from the opposite pole. However, Kif4A remains processive on the interacting microtubules and might result in the generation of kinetochore microtubule flux by the action of cross-linking molecules. This view is also supported by the recent observation that the movements of neighboring sister kinetochore pairs in metaphase are correlated and could be explained by simulating elastic cross-linkers between kinetochore microtubules (Vladimirov *et al.*, 2013).

Although inhibition of poleward microtubule flux did not affect spindle assembly and chromosome alignment, it increased chromosome missegregation, highlighting its importance for accurate cell division (Ganem *et al.*, 2005). This negative consequence of microtubule flux abrogation could be explained through the role of flux-driven microtubule slippage at kinetochores, which facilitates error correction and equal distribution of spindle forces along the metaphase plate (Matos *et al.*, 2009).

Spindle Length Control

Significant progress in our understanding of the mechanisms behind spindle length control was obtained from the relationship between spindle size and cell size in meiotic and mitotic divisions during *Xenopus* development, as well as in meiotic and mitotic *Xenopus* cell extracts (reviewed in Goshima and Scholey, 2010). Accordingly, the cell wall and/or cytoplasmic material are limiting factors for spindle growth. However, cells reach a plateau in spindle size, after which the correlation is lost, indicating that in addition to cell size, some intrinsic mechanisms must control spindle length (Goshima and Scholey, 2010). These conclusions were elegantly confirmed in two recent studies in which droplets of different sizes containing mitotic *Xenopus* extracts were generated using microfluidics (Hazel *et al.*, 2013; Good *et al.*, 2013). These studies concluded that the amount of cytoplasmic material is a limiting factor that ultimately defines spindle length.

Some of the mechanisms that directly affect the regulation of spindle size involve polymerization- and depolymerization-promoting factors, cross-linking motor proteins involved in microtubule sliding, microtubule branching and severing proteins, poleward microtubule flux, and microtubule nucleation from chromosomes (reviewed in Goshima and Scholey, 2010). In addition, a strong correlation between centrosome size and spindle length has been reported (Greenan *et al.*, 2010), possibly due to enhanced microtubule nucleation from larger centrosomes, as well as the stronger pulling effect

of cortical dynein on a larger number of astral microtubules (Grill *et al.*, 2003).

Visco-Elastic Nature of the Mitotic Spindle

Deeper insight into the mechanical properties of the spindle was provided using micromechanical tools combined with high-resolution fluorescence imaging of metaphase spindles from *Xenopus* egg extracts. When spindles were compressed parallel to their axis by two cantilevers, one of them working as a force sensor and the other as a manipulator, they were able to completely recover their original shape if the applied force was not above a certain threshold, suggesting the existence of elastic properties. However, the recovery process after compression was slower than expected for strictly elastic elements, suggesting visco-elastic properties (Itabashi *et al.*, 2009). Application of forces above the threshold led to plastic irreversible deformations. Additionally, the observation of visco-elastic behavior suggested the existence of coupling elements between spindle microtubules, strengthening previous micromanipulation studies in grasshopper spermatocytes (Itabashi *et al.*, 2009; Nicklas *et al.*, 1982). Spindle micromanipulation by two force-calibrated microneedles further confirmed the visco-elastic behavior of spindles and showed that it indeed depends on microtubule cross-linking (Shimamoto *et al.*, 2011). Finally, spindle visco-elasticity was also confirmed in a recent study in which a stretching force was applied to spindles (Takagi *et al.*, 2014).

Mitotic Spindle Function

The main function of the mitotic spindle is to successfully distribute the genetic material packed on chromosomes to the two daughter cells during cell division. This essential function can be divided into two crucial steps: chromosome congression (Figure 1), which defines prometaphase and metaphase mitotic stages, and chromosome segregation (Figure 2), which defines anaphase and telophase.

Chromosome Congression

After nuclear envelope breakdown at the onset of mitosis, a large number of microtubules originate from the centrosomes, and, together with chromokinesins, generate PEFs that help chromosomes to arrange in a toroid conformation on the spindle surface. Chromosome arms become oriented away from the spindle, protecting their kinetochores from being shielded by other chromosomes and making them more reachable by growing microtubules from both spindle poles (Magidson *et al.*, 2011). Chromosomes that are initially situated much closer to one pole relative to the other have less chance to be reached from both spindle poles, and their congression is facilitated by the motor proteins Dynein and CENP-E (Barisic *et al.*, 2014; Figure 1(a)). Kinetochores of these chromosomes first become laterally attached by centrosomal microtubules and then are transported closer to the pole by the minus end-directed action of Dynein (Rieder and Alexander, 1990; Li *et al.*, 2007; Vorozhko *et al.*, 2008; Yang

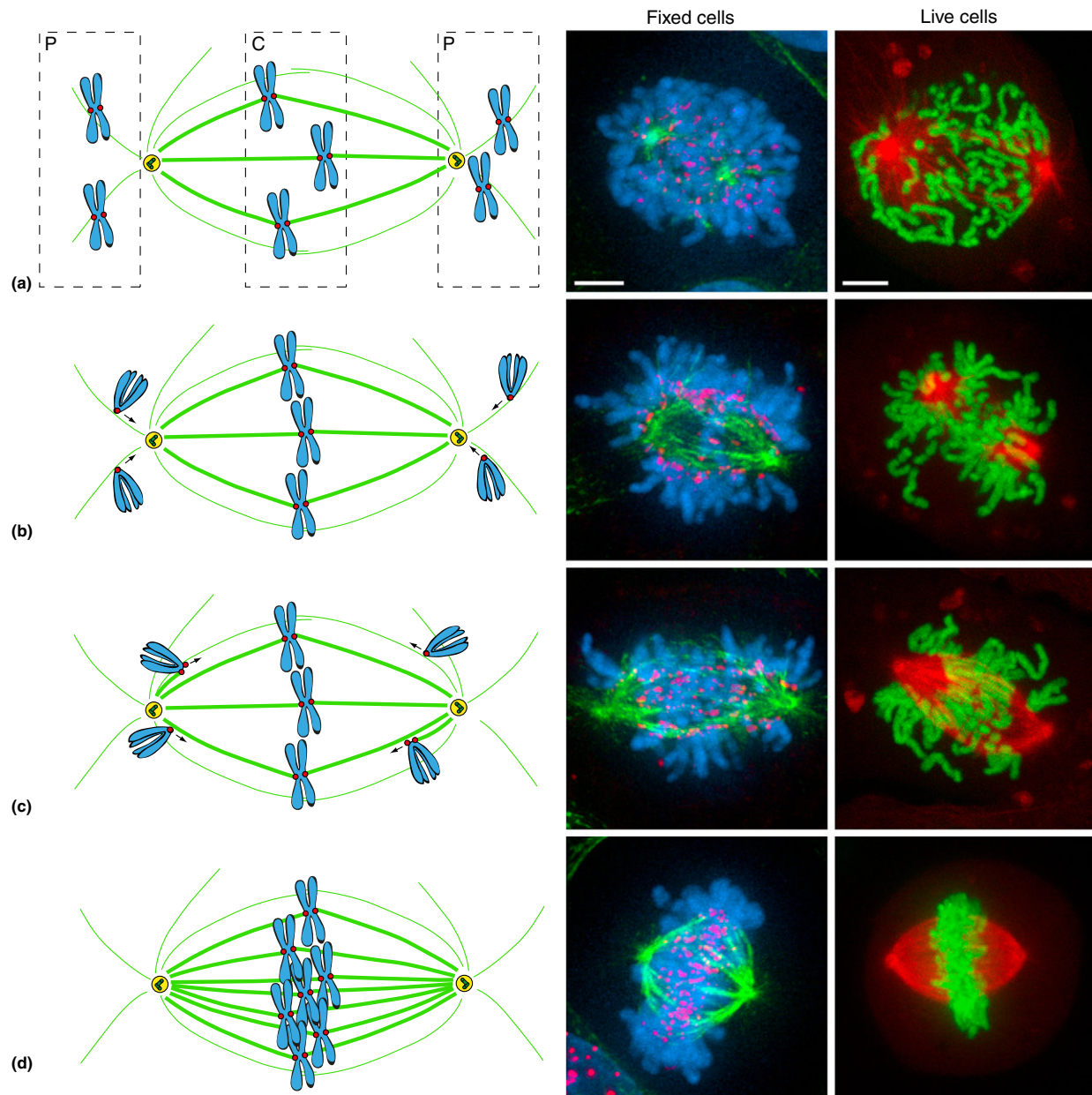


Figure 1 Mechanisms of chromosome congression. (a) Initial positioning of chromosomes after the nuclear envelope breakdown affects chromosome congression. Peripheral chromosomes that are initially closer to one pole than the other (P) depend more on motor proteins to congress to the metaphase plate than the centrally localized chromosomes (C), which can easily be reached from both poles. (b) Centrally localized chromosomes then become bi-oriented and align via changes in microtubule polymerization dynamics, while peripheral chromosomes are carried to the spindle poles by motor activity of Dynein. (c) After reaching the pole chromosomes continue to congress towards the metaphase plate mainly via activity of another motor protein – CENP-E. (d) Coordination of different motor protein activities and microtubule dynamics results in successful alignment of all chromosomes at the mitotic spindle equator in metaphase. Micrographs of fixed and live samples represent parental and H2B-GFP/mCherry- α -tubulin expressing U2OS cells, respectively. In the micrographs of fixed cells, chromosomes are labeled in blue (DAPI), kinetochores in red (ACA), and microtubules in green (mouse anti- α -tubulin antibody); the similar color code as used in the figure. Scale bar: 5 μ m.

et al., 2007b; Figure 1(b)), whereas the chromosomes that are initially positioned between the two poles quickly biorient, stabilizing end-on kinetochore microtubule attachment via the KMN network (Cheeseman and Desai, 2008). Thus, Dynein exerts the dominant force required for congression of peripheral chromosomes, overcoming the activities of CENP-E and chromokinesins and preventing the establishment of premature

end-on kinetochore microtubule attachments (Barisic *et al.*, 2014). Once at the pole, CENP-E becomes the dominant force, promoting chromosome congression to the metaphase plate (Kapoor *et al.*, 2006; Figures 1(c) and 1(d)). At the equator, chromokinesins are required to maintain chromosome alignment, possibly by stabilizing end-on kinetochore microtubule attachments (Wandke *et al.*, 2012; Cane *et al.*, 2013).

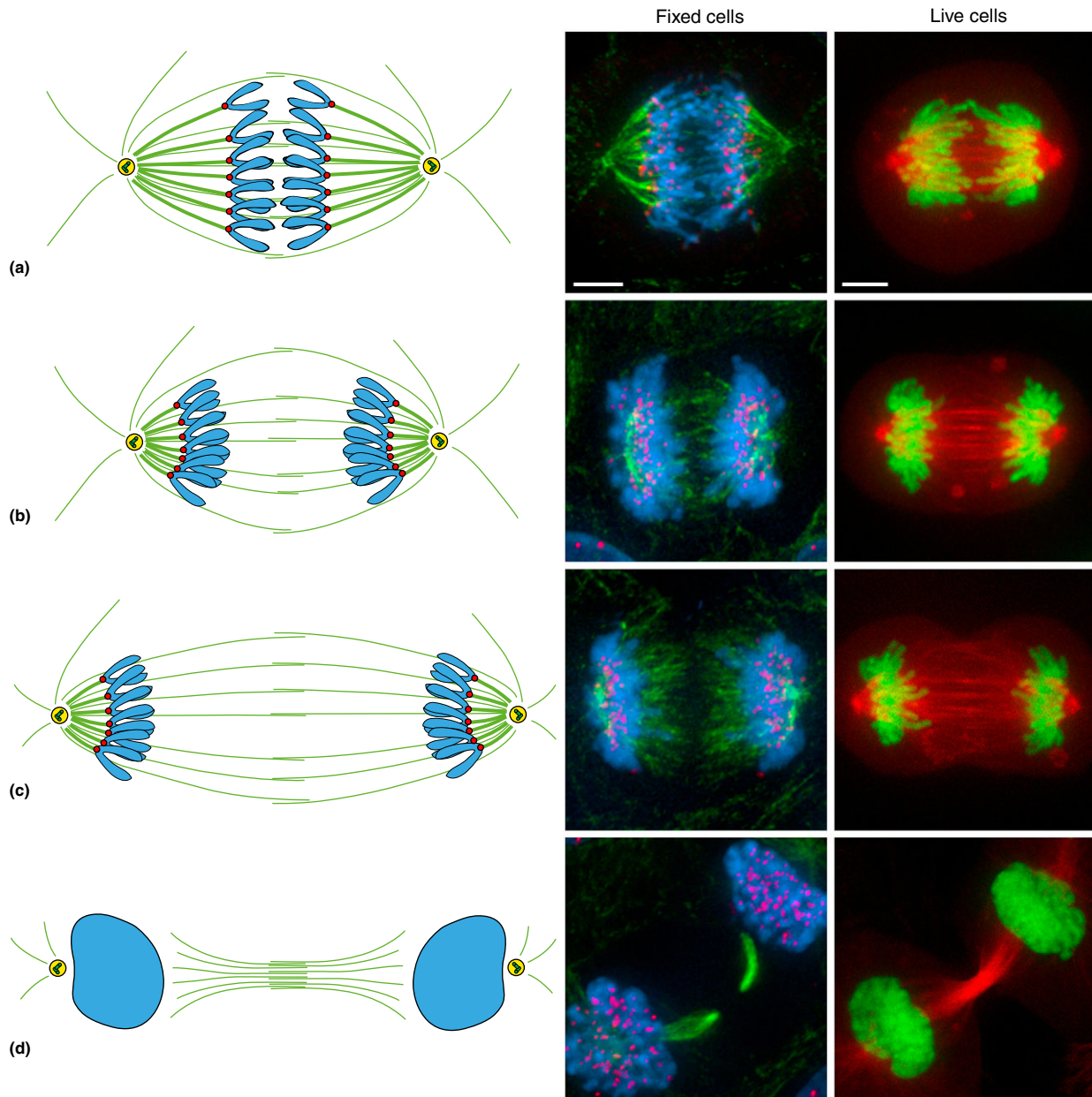


Figure 2 Mechanisms of chromosome segregation. (a) Once the last chromosome becomes bi-oriented and aligned, the cell is ready to enter anaphase. Spindle assembly checkpoint (SAC) is now satisfied, resulting in the activation of anaphase-promoting complex/cyclosome (APC/C) and consequent drop in CDK1 kinase activity and resolution of sister chromatid cohesion. (b) Chromosome separation in anaphase A starts with the shortening of kinetochore microtubules that is driven by microtubule depolymerases and poleward flux. (c) In anaphase B sister chromatids further separate due to spindle elongation caused by motor protein-driven microtubule sliding. (d) Finally, the chromatids decondense, and a cleavage furrow is formed by contraction of actin–myosin ring, surrounding the central spindle and preparing the cell for the abscission in cytokinesis. Micrographs of fixed and live samples represent parental and H2B-GFP/mCherry- α -tubulin expressing U2OS cells, respectively. In the micrographs of fixed cells, chromosomes are labeled in blue (DAPI), kinetochores in red (ACA), and microtubules in green (mouse anti- α -tubulin antibody); the similar color code as used in the figure. Scale bar: 5 μ m.

The transition from lateral to end-on attachments depends on several pathways. The localization of the dynein motor complex to the outer kinetochore layer is very dynamic and depends on Spindly, which is loaded on kinetochores through interaction with the RZZ complex (Rod-ZW10-Zwilch) (reviewed in Barisic and Geley, 2011). Apparently, the role of the RZZ complex is to support lateral contacts by decreasing the

binding affinity of the Ndc80 complex for microtubules, which is regulated by Spindly and Dynein. When Dynein removes this whole complex through its poleward motor activity, the inhibition of Ndc80 is relieved, allowing the formation of stable end-on attachments (Barisic *et al.*, 2010; Gassmann *et al.*, 2010; Cheerambathur *et al.*, 2013). A recent study also proposed roles for CENP-E and MCAK in lateral to end-on

attachment conversion (Shrestha and Draviam, 2013). Accordingly, while CENP-E is required to tether the kinetochore to the microtubule lattice, MCAK resolves these interactions in a subsequent conversion step. In addition to these two pathways, it was also shown that Bub1, a spindle assembly checkpoint kinase (Meraldi and Sorger, 2005), as well as the astrin/SKAP complex (Dunsch *et al.*, 2011), play an important role in this transition.

Kinetochore-driven formation of microtubules may additionally promote chromosome congression and biorientation, since these microtubules extend toward the poles and can become captured by growing astral microtubules and incorporated into the bipolar spindle in a Dynein-dependent manner (Maiato *et al.*, 2004a; Elting *et al.*, 2014; Khodjakov *et al.*, 2003).

Error Correction

Once stably attached, the kinetochore attachment site must be robust enough to resist the forces applied by spindle microtubules. However, before stabilization, kinetochore microtubule attachments must be rapidly reversible to allow for correction of errors (Cheeseman and Desai, 2008). Indeed, during early mitosis, erroneous attachments happen frequently, for instance, when both kinetochores are connected to the same spindle pole (syntelic attachment). Merotelic attachments may also arise when one kinetochore is attached to both spindle poles simultaneously. This type of aberrant attachment is particularly dangerous, because it cannot be detected by the spindle assembly checkpoint (discussed in more detail below). In contrast to merotelic attachments, syntelic attachments are highly unstable and are corrected before biorientation is achieved. Failure to correct these attachments may result in chromosome segregation errors, giving rise to aneuploidy, a hallmark of many cancers.

Error correction is carried out mainly through the activity of Aurora B kinase. When the activity of this kinase was disrupted by small-molecule inhibitors, the frequencies of both merotelic and syntelic errors increased significantly (Hauf *et al.*, 2003; Ditchfield *et al.*, 2003; Cimini *et al.*, 2006). Moreover, error correction requires active Aurora B kinase (Lampson *et al.*, 2004). Aurora B kinase phosphorylates several microtubule depolymerases, such as MCAK and Kif2B (Andrews *et al.*, 2004; Bakhoun *et al.*, 2009; Lan *et al.*, 2004), as well as proteins required for the stabilization of end-on kinetochore microtubule attachments, such as Ndc80 (DeLuca *et al.*, 2006; Cheeseman *et al.*, 2006; Ciferri *et al.*, 2008).

Aurora B localizes to the inner centromere between two sister kinetochores. Upon chromosome biorientation, tension is applied along the centromeric region, resulting in increased distances between the sister kinetochores, as well as alterations in kinetochore architecture (reviewed in Maresca and Salmon, 2010). This leads to physical separation of Aurora B from the Ndc80 complex, causing its dephosphorylation and increasing kinetochore affinity for microtubules (Fuller *et al.*, 2008; Liu *et al.*, 2009).

Finally, Aurora B activity is counteracted by protein phosphatase 1 (PP1), localized at the outer kinetochore plate, causing dephosphorylation of Aurora B substrates at the

kinetochore (Emanuele *et al.*, 2008; Francisco *et al.*, 1994; Liu *et al.*, 2010).

Chromosome Separation

After completing the biorientation of chromosomes, cells can proceed to the next step of cell division, anaphase. The initiation of this step is under the control of the spindle assembly checkpoint (SAC), which monitors kinetochore microtubule attachments (Khodjakov and Pines, 2010). The SAC inhibits the anaphase-promoting complex/cyclosome (APC/C), a ubiquitin ligase that targets proteins for degradation by the proteasome (reviewed in Musacchio, 2011). Two main APC/C targets are Cyclin B and Securin. Degradation of Cyclin B leads to inactivation of CDK1 kinase, whereas degradation of Securin activates Separase, the enzyme required for the cleavage of cohesin molecules that hold sister chromatids together (reviewed in Nasmyth and Haering, 2009).

To prevent the premature separation of sister chromatids, the checkpoint protein Mad2 is recruited to kinetochores and is released either alone or in complex with other proteins that interact with and inhibit the APC/C activator Cdc20. This mechanism ensures that a single unattached kinetochore is sufficient to inhibit APC/C activity, preventing anaphase onset until all kinetochores become attached (Rieder *et al.*, 1995). In addition to Mad2, BubR1 and Bub3 can also directly bind to the APC/C-Cdc20 and inhibit its ubiquitin ligase activity (reviewed in Musacchio, 2011). Once all kinetochores become stably attached the SAC is satisfied, Cyclin B and Securin become degraded and the cohesion between sister chromatids is resolved, allowing spindle forces to pull them apart (Figure 2(a)). These movements are initially generated by shortening of the kinetochore microtubules during anaphase A, which is driven by increased microtubule depolymerization at the kinetochore (known as the Pacman mechanism), as well as through microtubule depolymerization near the poles, both of which contribute to chromosome segregation in different species and cell types (Maiato and Lince-Faria, 2010; Figure 2(b)). This is normally followed by motor protein-driven microtubule sliding and spindle elongation, further contributing to chromosome segregation during anaphase B (Figure 2(c)).

Spindle Disassembly and Cytokinesis

Due to the decline in CDK1 activity, the mitotic spindle disassembles as chromosomes start to decondense and the nuclear envelope reforms around the two daughter nuclei. Nuclear envelope reformation and chromosome decondensation are additionally negatively regulated by a gradient of Aurora B activity spreading from the spindle midzone (Afonso *et al.*, 2014; Fuller *et al.*, 2008). The spindle midzone is also required for the assembly of the centralspindlin complex, which governs the positioning and activation of the cortical actin-myosin contractile ring that is assembled at the inner side of the plasma membrane to initiate cytokinesis (reviewed in Green *et al.*, 2012). As the actin-myosin ring contracts, a cleavage furrow forms, which pulls the cell membrane inward and surrounds the central spindle to

become compacted into the so-called midbody (Figure 2(d)). The midbody is then finally dismantled (Guizetti *et al.*, 2011), cell division is terminated, and two daughter cells are completely separated.

See also: Cytoskeleton and Motors: Cytoskeletal Components: Microtubules and Microtubule-Associated Proteins (MAPs). Dynamic Integration: Dynamics: Dynamics of Microtubule Self-Assembly

References

- Afonso, O., Matos, I., Pereira, A.J., *et al.*, 2014. Feedback control of chromosome separation by a midzone Aurora B gradient. *Science* 345, 332–336.
- Akhmanova, A., Hoogenraad, C.C., Drabek, K., *et al.*, 2001. Clasps are CLIP-115 and -170 associating proteins involved in the regional regulation of microtubule dynamics in motile fibroblasts. *Cell* 104, 923–935.
- Al-Bassam, J., Kim, H., Brouhard, G., *et al.*, 2010. CLASP promotes microtubule rescue by recruiting tubulin dimers to the microtubule. *Developmental Cell* 19, 245–258.
- Allan, V.J., 2011. Cytoplasmic dynein. *Biochemical Society Transactions* 39, 1169–1178.
- Andrews, P.D., Ovechkina, Y., Morrice, N., *et al.*, 2004. Aurora B regulates MCAK at the mitotic centromere. *Developmental Cell* 6, 253–268.
- Arnal, I., Heichette, C., Diamantopoulos, G.S., Chretien, D., 2004. CLIP-170/tubulin-curved oligomers coassemble at microtubule ends and promote rescues. *Current Biology* 14, 2086–2095.
- Bakhoum, S.F., Thompson, S.L., Manning, A.L., Compton, D.A., 2009. Genome stability is ensured by temporal control of kinetochore-microtubule dynamics. *Nature Cell Biology* 11, 27–35.
- Barisic, M., Aguiar, P., Geley, S., Maiato, H., 2014. Kinetochore motors drive congression of peripheral polar chromosomes by overcoming random arm-ejection forces. *Nature Cell Biology* 16, 1249–1256.
- Barisic, M., Geley, S., 2011. Spindly switch controls anaphase: Spindly and RZZ functions in chromosome attachment and mitotic checkpoint control. *Cell Cycle* 10, 449–456.
- Barisic, M., Pereira, A.J., Maiato, H., 2012. Fluorescent speckle microscopy in cultured cells. *Methods in Enzymology* 504, 147–161.
- Barisic, M., Sohm, B., Mikolcevic, P., *et al.*, 2010. Spindly/CCDC99 is required for efficient chromosome congression and mitotic checkpoint regulation. *Molecular Biology of the Cell* 21, 1968–1981.
- Barr, A.R., Gergely, F., 2008. MCAK-independent functions of ch-Tog/XMAP215 in microtubule plus-end dynamics. *Molecular and Cellular Biology* 28, 7199–7211.
- Basto, R., Lau, J., Vinogradova, T., *et al.*, 2006. Flies without centrioles. *Cell* 125, 1375–1386.
- Blangy, A., Lane, H.A., D'Herin, P., *et al.*, 1995. Phosphorylation by p34cdc2 regulates spindle association of human Eg5, a kinesin-related motor essential for bipolar spindle formation *in vivo*. *Cell* 83, 1159–1169.
- Bonaccorsi, S., Giansanti, M.G., Gatti, M., 1998. Spindle self-organization and cytokinesis during male meiosis in asterless mutants of *Drosophila melanogaster*. *Journal of Cell Biology* 142, 751–761.
- Bonner, M.K., Poole, D.S., Xu, T., *et al.*, 2011. Mitotic spindle proteomics in Chinese hamster ovary cells. *PLoS One* 6, e20489.
- Borisy, G.G., Taylor, E.W., 1967. The mechanism of action of colchicine. Binding of colchicine-3H to cellular protein. *Journal of Cell Biology* 34, 525–533.
- Brouhard, G.J., Stear, J.H., Noetzel, T.L., *et al.*, 2008. XMAP215 is a processive microtubule polymerase. *Cell* 132, 79–88.
- Brugues, J., Nuzzo, V., Mazur, E., Needleman, D.J., 2012. Nucleation and transport organize microtubules in metaphase spindles. *Cell* 149, 554–564.
- Brust-Mascher, I., Scholey, J.M., 2002. Microtubule flux and sliding in mitotic spindles of *Drosophila* embryos. *Molecular Biology of the Cell* 13, 3967–3975.
- Burbank, K.S., Mitchison, T.J., Fisher, D.S., 2007. Slide-and-cluster models for spindle assembly. *Current Biology* 17, 1373–1383.
- Cane, S., Ye, A.A., Luks-Morgan, S.J., Maresca, T.J., 2013. Elevated polar ejection forces stabilize kinetochore-microtubule attachments. *Journal of Cell Biology* 200, 203–218.
- Carazo-Salas, R.E., Gruss, O.J., Mattaj, I.W., Karsenti, E., 2001. Ran-GTP coordinates regulation of microtubule nucleation and dynamics during mitotic-spindle assembly. *Nature Cell Biology* 3, 228–234.
- Carazo-Salas, R.E., Guarguaglini, G., Gruss, O.J., *et al.*, 1999. Generation of GTP-bound Ran by RCC1 is required for chromatin-induced mitotic spindle formation. *Nature* 400, 178–181.
- Casanova, C.M., Rybina, S., Yokoyama, H., Karsenti, E., Mattaj, I.W., 2008. Hepatoma up-regulated protein is required for chromatin-induced microtubule assembly independently of TPX2. *Molecular Biology of the Cell* 19, 4900–4908.
- Cheerambathur, D.K., Gassmann, R., Cook, B., Oegema, K., Desai, A., 2013. Crosstalk between microtubule attachment complexes ensures accurate chromosome segregation. *Science* 342, 1239–1242.
- Cheeseman, I.M., Chappie, J.S., Wilson-Kubalek, E.M., Desai, A., 2006. The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. *Cell* 127, 983–997.
- Cheeseman, I.M., Desai, A., 2008. Molecular architecture of the kinetochore-microtubule interface. *Nature Reviews Molecular Cell Biology* 9, 33–46.
- Chretien, D., Fuller, S.D., Karsenti, E., 1995. Structure of growing microtubule ends: two-dimensional sheets close into tubes at variable rates. *Journal of Cell Biology* 129, 1311–1328.
- Ciferri, C., Pasqualato, S., Screpanti, E., *et al.*, 2008. Implications for kinetochore-microtubule attachment from the structure of an engineered Ndc80 complex. *Cell* 133, 427–439.
- Cimini, D., Wan, X., Hirel, C.B., Salmon, E.D., 2006. Aurora kinase promotes turnover of kinetochore microtubules to reduce chromosome segregation errors. *Current Biology* 16, 1711–1718.
- Cole, D.G., Saxton, W.M., Sheehan, K.B., Scholey, J.M., 1994. A “slow” homotetrameric kinesin-related motor protein purified from *Drosophila* embryos. *Journal of Biological Chemistry* 269, 22913–22916.
- Debec, A., Szollosi, A., Szollosi, D., 1982. A *Drosophila melanogaster* cell line lacking centriole. *Biology of the Cell* 44, 133–138.
- DeLuca, J.G., Gall, W.E., Ciferri, C., *et al.*, 2006. Kinetochore microtubule dynamics and attachment stability are regulated by Hec1. *Cell* 127, 969–982.
- Desai, A., Mitchison, T.J., 1997. Microtubule polymerization dynamics. *Annual Review of Cell and Developmental Biology* 13, 83–117.
- Desai, A., Verma, S., Mitchison, T.J., Walczak, C.E., 1999. Kin I kinesins are microtubule-destabilizing enzymes. *Cell* 96, 69–78.
- Ditchfield, C., Johnson, V.L., Tighe, A., *et al.*, 2003. Aurora B couples chromosome alignment with anaphase by targeting BubR1, Mad2, and CenP-E to kinetochores. *Journal of Cell Biology* 161, 267–280.
- Drechsler, H., McHugh, T., Singleton, M.R., Carter, N.J., McAnish, A.D., 2014. The Kinesin-12 Kif15 is a processive track-switching tetramer. *Elife* 3, e01724.
- Dumont, J., Petri, S., Pellegrin, F., *et al.*, 2007. A centriole- and RanGTP-independent spindle assembly pathway in meiosis I of vertebrate oocytes. *Journal of Cell Biology* 176, 295–305.
- Dunsch, A.K., Linnane, E., Barr, F.A., Gruneberg, U., 2011. The astrin-kinastrin/SKAP complex localizes to microtubule plus ends and facilitates chromosome alignment. *Journal of Cell Biology* 192, 959–968.
- Elting, M.W., Hueschen, C.L., Udy, D.B., Dumont, S., 2014. Force on spindle microtubule minus ends moves chromosomes. *Journal of Cell Biology* 206, 245–256.
- Emanuele, M.J., Lan, W., Jwa, M., *et al.*, 2008. Aurora B kinase and protein phosphatase 1 have opposing roles in modulating kinetochore assembly. *Journal of Cell Biology* 181, 241–254.
- Ems-McClung, S.C., Walczak, C.E., 2010. Kinesin-13s in mitosis: Key players in the spatial and temporal organization of spindle microtubules. *Seminars in Cell and Developmental Biology* 21, 276–282.
- Ferenz, N.P., Paul, R., Fagerstrom, C., Mogilner, A., Wadsworth, P., 2009. Dynein antagonizes eg5 by crosslinking and sliding antiparallel microtubules. *Current Biology* 19, 1833–1838.
- Ferreira, J.G., Pereira, A.L., Maiato, H., 2014. Microtubule plus-end tracking proteins and their roles in cell division. *International Review Of Cell and Molecular Biology* 309, 59–140.
- Fink, G., Hajdo, L., Skowronek, K.J., *et al.*, 2009. The mitotic kinesin-14 Ncd drives directional microtubule-microtubule sliding. *Nature Cell Biology* 11, 717–723.
- Flemming, W., 1965. Historical paper. Contributions to the knowledge of the cell and its vital processes. *Journal of Cell Biology* 25, Suppl. 1–69.
- Francisco, L., Wang, W., Chan, C.S., 1994. Type 1 protein phosphatase acts in opposition to Ipl1 protein kinase in regulating yeast chromosome segregation. *Molecular and Cellular Biology* 14, 4731–4740.
- Fuller, B.G., Lampson, M.A., Foley, E.A., *et al.*, 2008. Midzone activation of aurora B in anaphase produces an intracellular phosphorylation gradient. *Nature* 453, 1132–1136.

- Gache, V., Waridel, P., Winter, C., *et al.*, 2010. *Xenopus* meiotic microtubule-associated interactome. *PLoS One* 5, e9248.
- Ganem, N.J., Upton, K., Compton, D.A., 2005. Efficient mitosis in human cells lacking poleward microtubule flux. *Current Biology* 15, 1827–1832.
- Gard, D.L., Kirschner, M.W., 1987. A microtubule-associated protein from *Xenopus* eggs that specifically promotes assembly at the plus-end. *Journal of Cell Biology* 105, 2203–2215.
- Gassmann, R., Holland, A.J., Varma, D., *et al.*, 2010. Removal of Spindly from microtubule-attached kinetochores controls spindle checkpoint silencing in human cells. *Genes & Development* 24, 957–971.
- Gergely, F., Draviam, V.M., Raff, J.W., 2003. The ch-TOG/XMAP215 protein is essential for spindle pole organization in human somatic cells. *Genes & Development* 17, 336–341.
- Good, M.C., Vahey, M.D., Skandarajah, A., Fletcher, D.A., Heald, R., 2013. Cytoplasmic volume modulates spindle size during embryogenesis. *Science* 342, 856–860.
- Goodwin, S.S., Vale, R.D., 2010. Patronin regulates the microtubule network by protecting microtubule minus ends. *Cell* 143, 263–274.
- Gorbsky, G.J., Borisy, G.G., 1989. Microtubules of the kinetochore fiber turn over in metaphase but not in anaphase. *Journal of Cell Biology* 109, 653–662.
- Goshima, G., Mayer, M., Zhang, N., Stuurman, N., Vale, R.D., 2008. Augmin: A protein complex required for centrosome-independent microtubule generation within the spindle. *Journal of Cell Biology* 181, 421–429.
- Goshima, G., Scholey, J.M., 2010. Control of mitotic spindle length. *Annual Review of Cell and Developmental Biology* 26, 21–57.
- Goshima, G., Wollman, R., Goodwin, S.S., *et al.*, 2007. Genes required for mitotic spindle assembly in *Drosophila* S2 cells. *Science* 316, 417–421.
- Green, R.A., Paluch, E., Oegema, K., 2012. Cytokinesis in animal cells. *Annual Review of Cell and Developmental Biology* 28, 29–58.
- Greenan, G., Brangwynne, C.P., Jaensch, S., *et al.*, 2010. Centrosome size sets mitotic spindle length in *Caenorhabditis elegans* embryos. *Current Biology* 20, 353–358.
- Grill, S.W., Howard, J., Schaffer, E., Stelzer, E.H., Hyman, A.A., 2003. The distribution of active force generators controls mitotic spindle position. *Science* 301, 518–521.
- Gruss, O.J., Vernos, I., 2004. The mechanism of spindle assembly: Functions of Ran and its target TPX2. *Journal of Cell Biology* 166, 949–955.
- Guizetti, J., Schermelleh, L., Mantler, J., *et al.*, 2011. Cortical constriction during abscission involves helices of ESCRT-III-dependent filaments. *Science* 331, 1616–1620.
- Hauf, S., Cole, R.W., Laterra, S., *et al.*, 2003. The small molecule Hesperadin reveals a role for Aurora B in correcting kinetochore-microtubule attachment and in maintaining the spindle assembly checkpoint. *Journal of Cell Biology* 161, 281–294.
- Hazel, J., Krutkramelis, K., Mooney, P., *et al.*, 2013. Changes in cytoplasmic volume are sufficient to drive spindle scaling. *Science* 342, 853–856.
- Heald, R., Tournebise, R., Blank, T., *et al.*, 1996. Self-organization of microtubules into bipolar spindles around artificial chromosomes in *Xenopus* egg extracts. *Nature* 382, 420–425.
- Helmke, K.J., Heald, R., Wilbur, J.D., 2013. Interplay between spindle architecture and function. *International Review Of Cell and Molecular Biology* 306, 83–125.
- Hendershott, M.C., Vale, R.D., 2014. Regulation of microtubule minus-end dynamics by CAMSAPs and Patronin. *Proceedings of the National Academy of Sciences of the United States of America* 111, 5860–5865.
- Hepperla, A.J., Willey, P.T., Coombes, C.E., *et al.*, 2014. Minus-end-directed Kinesin-14 motors align antiparallel microtubules to control metaphase spindle length. *Developmental Cell* 31, 61–72.
- Hinchcliffe, E.H., Miller, F.J., Cham, M., Khodjakov, A., Sluder, G., 2001. Requirement of a centrosomal activity for cell cycle progression through G1 into S phase. *Science* 291, 1547–1550.
- Howard, J., Hyman, A.A., 2009. Growth, fluctuation and switching at microtubule plus ends. *Nature Reviews Molecular Cell Biology* 10, 569–574.
- Hu, C.K., Ozlu, N., Coughlin, M., Steen, J.J., Mitchison, T.J., 2012. Plk1 negatively regulates PRC1 to prevent premature midzone formation before cytokinesis. *Molecular Biology of the Cell* 23, 2702–2711.
- Huitorel, P., Kirschner, M.W., 1988. The polarity and stability of microtubule capture by the kinetochore. *Journal of Cell Biology* 106, 151–159.
- Inoue, S., 1953. Polarization optical studies of the mitotic spindle. I. The demonstration of spindle fibers in living cells. *Chromosoma* 5, 487–500.
- Inoue, S., 1981. Cell division and the mitotic spindle. *Journal of Cell Biology* 91, 131s–147s.
- Itabashi, T., Takagi, J., Shimamoto, Y., *et al.*, 2009. Probing the mechanical architecture of the vertebrate meiotic spindle. *Nature Methods* 6, 167–172.
- Jiang, K., Hua, S., Mohan, R., *et al.*, 2014. Microtubule minus-end stabilization by polymerization-driven CAMSAP deposition. *Developmental Cell* 28, 295–309.
- Kabeche, L., Compton, D.A., 2013. Cyclin A regulates kinetochore microtubules to promote faithful chromosome segregation. *Nature* 502, 110–113.
- Kalab, P., Weis, K., Heald, R., 2002. Visualization of a Ran-GTP gradient in interphase and mitotic *Xenopus* egg extracts. *Science* 295, 2452–2456.
- Kamasaki, T., O'Toole, E., Kita, S., *et al.*, 2013. Augmin-dependent microtubule nucleation at microtubule walls in the spindle. *Journal of Cell Biology* 202, 25–33.
- Kapoor, T.M., Lampson, M.A., Hergert, P., *et al.*, 2006. Chromosomes can congress to the metaphase plate before biorientation. *Science* 311, 388–391.
- Kapoor, T.M., Mayer, T.U., Coughlin, M.L., Mitchison, T.J., 2000. Probing spindle assembly mechanisms with monastrol, a small molecule inhibitor of the mitotic kinesin, Eg5. *Journal of Cell Biology* 150, 975–988.
- Khodjakov, A., Cole, R.W., Oakley, B.R., Rieder, C.L., 2000. Centrosome-independent mitotic spindle formation in vertebrates. *Current Biology* 10, 59–67.
- Khodjakov, A., Copenagle, L., Gordon, M.B., Compton, D.A., Kapoor, T.M., 2003. Minus-end capture of preformed kinetochore fibers contributes to spindle morphogenesis. *Journal of Cell Biology* 160, 671–683.
- Khodjakov, A., Pines, J., 2010. Centromere tension: A divisive issue. *Nature Cell Biology* 12, 919–923.
- Khodjakov, A., Rieder, C.L., 2001. Centrosomes enhance the fidelity of cytokinesis in vertebrates and are required for cell cycle progression. *Journal of Cell Biology* 153, 237–242.
- Kirschner, M., Mitchison, T., 1986. Beyond self-assembly: From microtubules to morphogenesis. *Cell* 45, 329–342.
- Kollman, J.M., Merdes, A., Mourey, L., Agard, D.A., 2011. Microtubule nucleation by gamma-tubulin complexes. *Nature Reviews Molecular Cell Biology* 12, 709–721.
- Komarova, Y.A., Akhmanova, A.S., Kojima, S., Galjart, N., Borisy, G.G., 2002. Cytoplasmic linker proteins promote microtubule rescue *in vivo*. *Journal of Cell Biology* 159, 589–599.
- Kwok, B.H., Kapoor, T.M., 2007. Microtubule flux: Drivers wanted. *Current Opinion in Cell Biology* 19, 36–42.
- LaFountain Jr., J.R., Cohan, C.S., Siegel, A.J., LaFountain, D.J., 2004. Direct visualization of microtubule flux during metaphase and anaphase in crane-fly spermatocytes. *Molecular Biology of the Cell* 15, 5724–5732.
- Lampson, M.A., Renduchitala, K., Khodjakov, A., Kapoor, T.M., 2004. Correcting improper chromosome-spindle attachments during cell division. *Nature Cell Biology* 6, 232–237.
- Lan, W., Zhang, X., Kline-Smith, S.L., *et al.*, 2004. Aurora B phosphorylates centromeric MCAK and regulates its localization and microtubule depolymerization activity. *Current Biology* 14, 273–286.
- Lemos, C.L., Sampaio, P., Maiato, H., *et al.*, 2000. Mast, a conserved microtubule-associated protein required for bipolar mitotic spindle organization. *EMBO Journal* 19, 3668–3682.
- Li, Y., Yu, W., Liang, Y., Zhu, X., 2007. Kinetochore dynein generates a poleward pulling force to facilitate congression and full chromosome alignment. *Cell Research* 17, 701–712.
- Liu, D., Vader, G., Vromans, M.J., Lampson, M.A., Lens, S.M., 2009. Sensing chromosome bi-orientation by spatial separation of aurora B kinase from kinetochore substrates. *Science* 323, 1350–1353.
- Liu, D., Vleugel, M., Backer, C.B., *et al.*, 2010. Regulated targeting of protein phosphatase 1 to the outer kinetochore by KNL1 opposes Aurora B kinase. *Journal of Cell Biology* 188, 809–820.
- Logarinho, E., Maffini, S., Barisic, M., *et al.*, 2012. CLASPs prevent irreversible multipolarity by ensuring spindle-pole resistance to traction forces during chromosome alignment. *Nature Cell Biology* 14, 295–303.
- Maddox, P., Straight, A., Coughlin, P., Mitchison, T.J., Salmon, E.D., 2003. Direct observation of microtubule dynamics at kinetochores in *Xenopus* extract spindles: Implications for spindle mechanics. *Journal of Cell Biology* 162, 377–382.
- Magidson, V., O'Connell, C.B., Loncarek, J., *et al.*, 2011. The spatial arrangement of chromosomes during prometaphase facilitates spindle assembly. *Cell* 146, 555–567.
- Maiato, H., Fairley, E.A., Rieder, C.L., *et al.*, 2003. Human CLASP1 is an outer kinetochore component that regulates spindle microtubule dynamics. *Cell* 113, 891–904.
- Maiato, H., Lince-Faria, M., 2010. The perpetual movements of anaphase. *Cellular and Molecular Life Sciences* 67, 2251–2269.
- Maiato, H., Rieder, C.L., Khodjakov, A., 2004a. Kinetochore-driven formation of kinetochore fibers contributes to spindle assembly during animal mitosis. *Journal of Cell Biology* 167, 831–840.
- Maiato, H., Sampaio, P., Lemos, C.L., *et al.*, 2002. MAST/Orbit has a role in microtubule-kinetochore attachment and is essential for chromosome alignment and maintenance of spindle bipolarity. *Journal of Cell Biology* 157, 749–760.

- Maiato, H., Sampaio, P., Sunkel, C.E., 2004b. Microtubule-associated proteins and their essential roles during mitosis. *International Review of Cytology* 241, 53–153.
- Mandelkow, E.M., Mandelkow, E., Milligan, R.A., 1991. Microtubule dynamics and microtubule caps: a time-resolved cryo-electron microscopy study. *Journal of Cell Biology* 114, 977–991.
- Maresca, T.J., Salmon, E.D., 2010. Welcome to a new kind of tension: Translating kinetochore mechanics into a wait-anaphase signal. *Journal of Cell Science* 123, 825–835.
- Matos, I., Pereira, A.J., Lince-Faria, M., *et al.*, 2009. Synchronizing chromosome segregation by flux-dependent force equalization at kinetochores. *Journal of Cell Biology* 186, 11–26.
- McIntosh, J.R., Molodtsov, M.I., Ataullakhanov, F.I., 2012. Biophysics of mitosis. *Quarterly Reviews of Biophysics* 45, 147–207.
- Megraw, T.L., Kao, L.R., Kaufman, T.C., 2001. Zygotic development without functional mitotic centrosomes. *Current Biology* 11, 116–120.
- Mennella, V., Agard, D.A., Huang, B., Pelletier, L., 2014. Amorphous no more: Subdiffraction view of the pericentriolar material architecture. *Trends in Cell Biology* 24, 188–197.
- Meraldi, P., Sorger, P.K., 2005. A dual role for Bub1 in the spindle checkpoint and chromosome congression. *EMBO Journal* 24, 1621–1633.
- Merdes, A., Ramyar, K., Vechio, J.D., Cleveland, D.W., 1996. A complex of NuMA and cytoplasmic dynein is essential for mitotic spindle assembly. *Cell* 87, 447–458.
- Meunier, S., Vernos, I., 2011. K-fibre minus ends are stabilized by a RanGTP-dependent mechanism essential for functional spindle assembly. *Nature Cell Biology* 13, 1406–1414.
- Mimori-Kiyosue, Y., Grigoriev, I., Sasaki, H., *et al.*, 2006. Mammalian CLASPs are required for mitotic spindle organization and kinetochore alignment. *Genes Cells* 11, 845–857.
- Mitchison, T., Evans, L., Schulze, E., Kirschner, M., 1986. Sites of microtubule assembly and disassembly in the mitotic spindle. *Cell* 45, 515–527.
- Mitchison, T., Kirschner, M., 1984. Dynamic instability of microtubule growth. *Nature* 312, 237–242.
- Mitchison, T.J., 1989. Polewards microtubule flux in the mitotic spindle: Evidence from photoactivation of fluorescence. *Journal of Cell Biology* 109, 637–652.
- Mohri, H., 1968. Amino-acid composition of “Tubulin” constituting microtubules of sperm flagella. *Nature* 217, 1053–1054.
- Mountain, V., Simerly, C., Howard, L., *et al.*, 1999. The kinesin-related protein, HSET, opposes the activity of Eg5 and cross-links microtubules in the mammalian mitotic spindle. *Journal of Cell Biology* 147, 351–366.
- Moutinho-Pereira, S., Stuurman, N., Afonso, O., *et al.*, 2013. Genes involved in centrosome-independent mitotic spindle assembly in *Drosophila* S2 cells. *Proceedings of the National Academy of Sciences of the United States of America* 110, 19808–19813.
- Musacchio, A., 2011. Spindle assembly checkpoint: The third decade. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 366, 3595–3604.
- Nasmyth, K., Haering, C.H., 2009. Cohesin: Its roles and mechanisms. *Annual Review of Genetics* 43, 525–558.
- Nicklas, R.B., Koch, C.A., 1969. Chromosome micromanipulation. 3. Spindle fiber tension and the reorientation of mal-oriented chromosomes. *Journal of Cell Biology* 43, 40–50.
- Nicklas, R.B., Kubai, D.F., Hays, T.S., 1982. Spindle microtubules and their mechanical associations after micromanipulation in anaphase. *Journal of Cell Biology* 95, 91–104.
- Noatynska, A., Gotta, M., Meraldi, P., 2012. Mitotic spindle (DIS) orientation and DiSease: Cause or consequence? *Journal of Cell Biology* 199, 1025–1035.
- Nunes Bastos, R., Gandhi, S.R., Baron, R.D., *et al.*, 2013. Aurora B suppresses microtubule dynamics and limits central spindle size by locally activating KIF4A. *Journal of Cell Biology* 202, 605–621.
- O’Connell, C.B., Loncarek, J., Kalab, P., Khodjakov, A., 2009. Relative contributions of chromatin and kinetochores to mitotic spindle assembly. *Journal of Cell Biology* 187, 43–51.
- Orticello, M., Fiore, M., Totta, P., *et al.*, 2014. N-terminus-modified Hec1 suppresses tumour growth by interfering with kinetochore-microtubule dynamics. *Oncogene*. doi:10.1038/onc.2014.265.
- Paul, R., Wollman, R., Silkworth, W.T., *et al.*, 2009. Computer simulations predict that chromosome movements and rotations accelerate mitotic spindle assembly without compromising accuracy. *Proceedings of the National Academy of Sciences of the United States of America* 106, 15708–15713.
- Pease, D.C., 1946. Hydrostatic pressure effects upon the spindle figure and chromosome movement; experiments on the meiotic divisions of Trade-scantia pollen mother cells. *Biological Bulletin* 91, 145–169.
- Pereira, A.J., Maiato, H., 2012. Maturation of the kinetochore-microtubule interface and the meaning of metaphase. *Chromosome Research* 20, 563–577.
- Pereira, A.L., Pereira, A.J., Maia, A.R., *et al.*, 2006. Mammalian CLASP1 and CLASP2 cooperate to ensure mitotic fidelity by regulating spindle and kinetochore function. *Molecular Biology of the Cell* 17, 4526–4542.
- Petry, S., Pugieux, C., Nedelec, F.J., Vale, R.D., 2011. Augmin promotes meiotic spindle formation and bipolarity in *Xenopus* egg extracts. *Proceedings of the National Academy of Sciences of the United States of America* 108, 14473–14478.
- Petry, S., Vale, R.D., 2011. A new cap for kinetochore fibre minus ends. *Nature Cell Biology* 13, 1389–1391.
- Rieder, C.L., Alexander, S.P., 1990. Kinetochores are transported poleward along a single astral microtubule during chromosome attachment to the spindle in newt lung cells. *Journal of Cell Biology* 110, 81–95.
- Rieder, C.L., Cole, R.W., Khodjakov, A., Sluder, G., 1995. The checkpoint delaying anaphase in response to chromosome monoorientation is mediated by an inhibitory signal produced by unattached kinetochores. *Journal of Cell Biology* 130, 941–948.
- Schmidt, W.J., 1937. Die Doppelbrechung von Karyoplasma, Metaplasma und Zytoplasma. Berlin: Gebruder Borntraeger.
- Sharp, D.J., McDonald, K.L., Brown, H.M., *et al.*, 1999. The bipolar kinesin, KLP61F, cross-links microtubules within interpolar microtubule bundles of *Drosophila* embryonic mitotic spindles. *Journal of Cell Biology* 144, 125–138.
- Sharp, D.J., Ross, J.L., 2012. Microtubule-severing enzymes at the cutting edge. *Journal of Cell Science* 125, 2561–2569.
- Shimamoto, Y., Maeda, Y.T., Ishiwata, S., Libchaber, A.J., Kapoor, T.M., 2011. Insights into the micromechanical properties of the metaphase spindle. *Cell* 145, 1062–1074.
- Shrestha, R.L., Draviam, V.M., 2013. Lateral to end-on conversion of chromosome-microtubule attachment requires kinesins CENP-E and MCAK. *Current Biology* 23, 1514–1526.
- Sillje, H.H., Nagel, S., Korner, R., Nigg, E.A., 2006. HURP is a Ran-importin beta-regulated protein that stabilizes kinetochore microtubules in the vicinity of chromosomes. *Current Biology* 16, 731–742.
- Sir, J.H., Putz, M., Daly, O., *et al.*, 2013. Loss of centrioles causes chromosomal instability in vertebrate somatic cells. *Journal of Cell Biology* 203, 747–756.
- Snyder, J.A., McIntosh, J.R., 1975. Initiation and growth of microtubules from mitotic centers in lysed mammalian cells. *Journal of Cell Biology* 67, 744–760.
- Takagi, J., Itabashi, T., Suzuki, K., *et al.*, 2014. Micromechanics of the vertebrate meiotic spindle examined by stretching along the pole-to-pole axis. *Biophysical Journal* 106, 735–740.
- Tanenbaum, M.E., Macurek, L., Galjart, N., Medema, R.H., 2008. Dynein, Lis1 and CLIP-170 counteract Eg5-dependent centrosome separation during bipolar spindle assembly. *EMBO Journal* 27, 3235–3245.
- Tanenbaum, M.E., Macurek, L., Janssen, A., *et al.*, 2009. Kif15 cooperates with eg5 to promote bipolar spindle assembly. *Current Biology* 19, 1703–1711.
- Tanenbaum, M.E., Vale, R.D., McKenney, R.J., 2013. Cytoplasmic dynein crosslinks and slides anti-parallel microtubules using its two motor domains. *Elife* 2, e00943.
- Tournebise, R., Popov, A., Kinoshita, K., *et al.*, 2000. Control of microtubule dynamics by the antagonistic activities of XMAP215 and XKCM1 in *Xenopus* egg extracts. *Nature Cell Biology* 2, 13–19.
- Uehara, R., Nozawa, R.S., Tomioka, A., *et al.*, 2009. The augmin complex plays a critical role in spindle microtubule generation for mitotic progression and cytokinesis in human cells. *Proceedings of the National Academy of Sciences of the United States of America* 106, 6998–7003.
- Van Den Wildenberg, S.M., Tao, L., Kapitein, L.C., *et al.*, 2008. The homotetrameric kinesin-5 KLP61F preferentially crosslinks microtubules into antiparallel orientations. *Current Biology* 18, 1860–1864.
- Vanneste, D., Takagi, M., Imamoto, N., Vernos, I., 2009. The role of Hkfp2 in the stabilization and maintenance of spindle bipolarity. *Current Biology* 19, 1712–1717.
- Vladimirov, E., Mchedlishvili, N., Gasic, I., *et al.*, 2013. Nonautonomous movement of chromosomes in mitosis. *Developmental Cell* 27, 60–71.
- Vorozhko, V.V., Emanuele, M.J., Kallio, M.J., Stukenberg, P.T., Gorbysky, G.J., 2008. Multiple mechanisms of chromosome movement in vertebrate cells mediated through the Ndc80 complex and dynein/dynactin. *Chromosoma* 117, 169–179.
- Walczak, C.E., Mitchison, T.J., Desai, A., 1996. XKCM1: A *Xenopus* kinesin-related protein that regulates microtubule dynamics during mitotic spindle assembly. *Cell* 84, 37–47.
- Wandke, C., Barisic, M., Sigl, R., *et al.*, 2012. Human chromokinesins promote chromosome congression and spindle microtubule dynamics during mitosis. *Journal of Cell Biology* 198, 847–863.

- Waterman-Storer, C.M., Desai, A., Bulinski, J.C., Salmon, E.D., 1998. Fluorescent speckle microscopy, a method to visualize the dynamics of protein assemblies in living cells. *Current Biology* 8, 1227–1230.
- Witt, P.L., Ris, H., Borisy, G.G., 1980. Origin of kinetochore microtubules in Chinese hamster ovary cells. *Chromosoma* 81, 483–505.
- Wollman, R., Cytrynbaum, E.N., Jones, J.T., *et al.*, 2005. Efficient chromosome capture requires a bias in the 'search-and-capture' process during mitotic-spindle assembly. *Current Biology* 15, 828–832.
- Wong, J., Fang, G., 2006. HURP controls spindle dynamics to promote proper interkinetochore tension and efficient kinetochore capture. *Journal of Cell Biology* 173, 879–891.
- Wordeman, L., 2010. How kinesin motor proteins drive mitotic spindle function: Lessons from molecular assays. *Seminars in Cell and Developmental Biology* 21, 260–268.
- Yang, C.P., Fan, S.S., 2008. *Drosophila* mars is required for organizing kinetochore microtubules during mitosis. *Experimental Cell Research*, 314, 3209–3220.
- Yang, G., Houghtaling, B.R., Gaetz, J., *et al.*, 2007a. Architectural dynamics of the meiotic spindle revealed by single-fluorophore imaging. *Nature Cell Biology* 9, 1233–1242.
- Yang, Z., Tulu, U.S., Wadsworth, P., Rieder, C.L., 2007b. Kinetochore dynein is required for chromosome motion and congression independent of the spindle checkpoint. *Current Biology* 17, 973–980.
- Zhai, Y., Kronebusch, P.J., Borisy, G.G., 1995. Kinetochore microtubule dynamics and the metaphase-anaphase transition. *Journal of Cell Biology* 131, 721–734.
- Zhang, D., Rogers, G.C., Buster, D.W., Sharp, D.J., 2007. Three microtubule severing enzymes contribute to the "Pacman-flux" machinery that moves chromosomes. *Journal of Cell Biology* 177, 231–242.
- Zhu, C., Jiang, W., 2005. Cell cycle-dependent translocation of PRC1 on the spindle by Kif4 is essential for midzone formation and cytokinesis. *Proceedings of the National Academy of Sciences of the United States of America* 102, 343–348.
- Zhu, C., Lau, E., Schwarzenbacher, R., Bossy-Wetzel, E., Jiang, W., 2006. Spatiotemporal control of spindle midzone formation by PRC1 in human cells. *Proceedings of the National Academy of Sciences of the United States of America* 103, 6196–6201.

Relevant Websites

- <http://www.ibiology.org/ibiomagazine/issue-6/shinya-inoue-the-dynamic-mitotic-spindle.html>
iBiology.
- <http://www.ibiology.org/ibiomagazine/issue-6/shinya-inoue-the-dynamic-mitotic-spindle.html>
iBiology.
- <http://www.ibiology.org/ibioseminars/cell-biology/dick-mcintosh-part-1.html>
iBiology.
- www.mitocheck.org
MitoCheck.
- <http://rmai.ucsf.edu>
University of California, San Francisco.
- <http://rmai.ucsf.edu/cnn>
University of California, San Francisco.

Centrioles and the Centrosome

J Sillibourne, Commissariat à l'Energie Atomique et aux Energies Alternatives (CEA), Grenoble, France
M Bornens, Institut Curie, Paris, France

© 2016 Elsevier Inc. All rights reserved.

Glossary

Cartwheel An early scaffolding structure assembled during centriole replication, consisting of a central hub with nine radial spokes that establishes the ninefold symmetry of the centriole.

Centriole A microtubule-based, barrel-like structure, approximately 200 nm in diameter and 450 nm long, forming the core of the centrosome.

Centrosome An organelle containing two similar but not identical centrioles embedded in a pericentriolar material.

Daughter centriole The younger centriole of the centriole pair.

Distal appendages and sub-distal appendages Stick-like structures that are attached to the distal end of the mother centriole.

Mother centriole An appendage-bearing centriole that is older than the daughter centriole.

Pericentriolar material A layered network of proteins surrounding the proximal ends of the centrioles.

Primary cilium A sensory organelle at the surface of the cell that is formed from the distal end of the mother centriole.

Procentriole A newly forming centriole assembled around a cartwheel.

Introduction

The centrosome is present in many organisms, including yeast where it is referred to as a spindle pole body, but is absent from plants. Two scientists, Van de Beneden and Boveri, almost simultaneously made the first observations of the centrosome in the late nineteenth century, although it was Boveri who gave the organelle the name 'centrosome' (central body). From their observations of the fertilized eggs of the transparent parasitic roundworm *Ascaris megalocephala*, they described the centrosome as an organelle with an attraction sphere and a central corpuscle that appeared to replicate itself and to be essential for the accurate segregation of chromosomes (Scheer, 2014). Boveri went on to show that manipulation of sea urchin embryos to produce blastomeres with too many centrosomes resulted in the formation of multipolar spindles and aberrant chromosome segregation. It is remarkable that these early observations of the centrosome are fundamentally true today and their proposed hypotheses are still the subject of investigation.

Structure of the Centrosome

At the centre of the centrosome there are two centrioles (Figure 1). These are barrel-like structures formed of microtubules that typically have a diameter of 200 nm and a length of 400–500 nm (although interspecies variation exists). In general, a centriole consists of nine sets of microtubules (microtubule blades; MB) that are arranged in a radially symmetrical manner to form the barrel of the centriole and these are tilted with respect to the barrel, giving the aspect of a ratchet wheel to transverse sections (Paintrand *et al.*, 1992). The centriole is a chiral structure as it is formed of microtubules with a defined polarity: the microtubule minus ends are at the proximal end of the centriole and the plus ends at the distal end of the centriole. The exact number of microtubules in each blade

can vary, with centrioles from *Caenorhabditis elegans* being very short and formed of singlet microtubules, those of *Drosophila melanogaster* being of intermediate size and formed of doublet microtubules, and in the case of *Homo sapiens* the proximal ends of centrioles are formed of triplet microtubules, while the distal ends are formed of doublet microtubules (Paintrand *et al.*, 1992). A triplet microtubule blade consists of a single complete microtubule, the A microtubule, of 13 protofilaments and two incomplete microtubules, the B and C microtubules, which contain 10 protofilaments each (Figure 1; Guichard *et al.*, 2013). Short links connect the A and C microtubules from neighboring MBs and help to stabilize the structure of the centriole. Tubulin posttranslational modifications, such as polyglutamylation and acetylation, also play a role in stabilizing the centriole structure and render it apparently insensitive to the action of microtubule destabilizing agents (Bobinnec *et al.*, 1998).

While the two centrioles in the centrosome share the same ninefold radial structure, they are not identical and differ both in age and composition. One of the two centrioles, the mother centriole, possesses two sets of appendages that extend from the distal end of the centriole (Paintrand *et al.*, 1992), while the other centriole, the daughter centriole, is shorter by 20% and lacks these appendages (Figure 1). Because the appendages are unique to the mother centriole, the composition of the two centrioles differs, although some proteins localizing to the sub-distal appendages, such as ninein (Delgehyr *et al.*, 2005) and Cep170 (Guarguaglini *et al.*, 2005) are also found at the proximal ends of both the mother and daughter centrioles. The daughter centriole acquires appendages after it transits through mitosis and enters the next cell cycle, when it will become a mother centriole.

Surrounding the proximal ends of each centriole is a network of proteins referred to as the pericentriolar material (PCM). The PCM used to be described as being amorphous, because of its appearance in electron micrographs, but recent high resolution microscopy studies have revealed that the

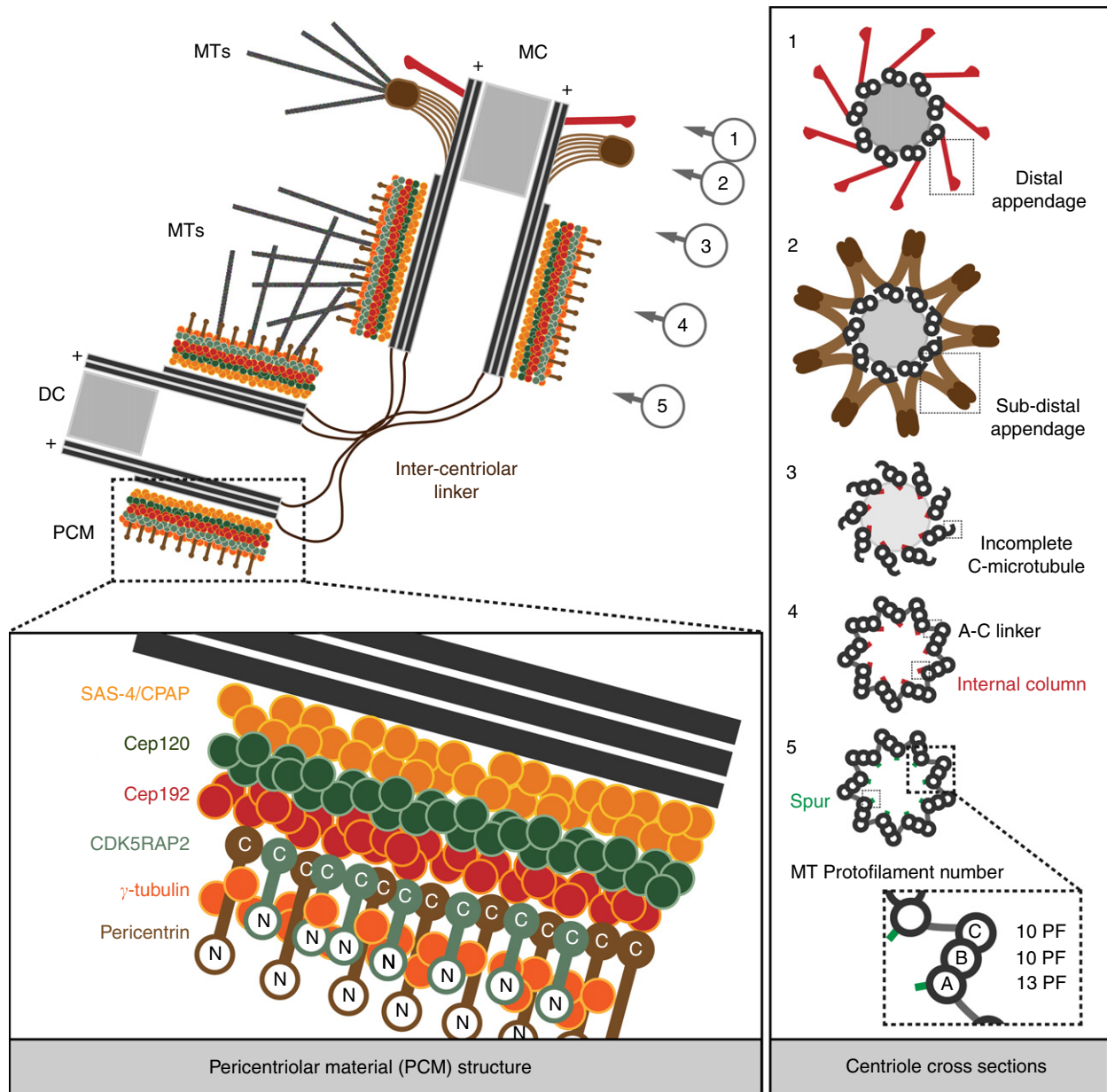


Figure 1 Structure of the centrosome. A diagram of a G1 centrosome consisting of a mother centriole (MC) and a daughter centriole (DC) attached to each other via an inter-centriolar linker. The MC possesses sub-distal appendages, which are involved in microtubule (MT) anchoring, and distal appendages, both of which the DC lacks. The right panel shows cross sections through the MC from the distal (section 1) to the proximal end (section 5) of the centriole. Note that the microtubule blades (MBs) consist of triplet MTs at the proximal end of the centriole, with one complete MT (A-MT) containing 13 protofilaments and two MTs (B- and C-MTs) that are formed of only 10 protofilaments each. At the distal end of the centrioles the MBs are formed of doublet MT (A- and B-MTs) and there is a progressive reduction in the tilt of the MBs from 45° at the proximal end to 20° at the distal end of the centriole. Adapted from Paintrand, M., Moudjou, M., Delacroix, H., Bornens, M., 1992. Centrosome organization and centriole architecture: Their sensitivity to divalent cations. *Journal of Structural Biology* 108, 107–128. The MBs are interconnected by A-C linkers and spurs also protrude from the C-MT into the centriolar lumen. Surrounding the proximal ends of the centrioles is the pericentriolar material (PCM), from which MTs are nucleated and in which procentrioles are assembled during centriole duplication. While the exact structure of the PCM has not been fully elucidated, recent work has shown that its components are organized into concentric layers surrounding the centrioles. A few PCM components are represented in the diagram, some of which, such as CDK5RAP2 and pericentrin, traverse the concentric layers and are oriented with their C-termini closest to the centriolar barrel and their N-termini facing outward.

PCM is highly ordered, consisting of concentric defined layers of proteins. Some of the proteins within these layers, such as pericentrin and CDK5RAP2, are oriented in a specific direction with their C-termini being located closer to the centriolar

barrel than their N-termini (Mennella *et al.*, 2014). The PCM is crucial for the function of the centrosome as it is the site from which microtubules are nucleated and is also the site where new centrioles, procentrioles, are assembled during replication

of the centrosome. PCM size is regulated throughout the cell cycle and there is a dramatic recruitment of PCM proteins to the centrosome during mitosis in a process that is referred to as centrosome maturation.

The mother and daughter centrioles are connected to one another by fibers emanating from their proximal ends, which are at the same time both flexible, to enable inter-centriolar movement (Piel *et al.*, 2000), and robust enough to maintain cohesion between the two centrioles. Cohesion between the mother and daughter centriole is maintained throughout most of the cell cycle, but it is disrupted in late G2, through the concerted action of kinases (Agircan *et al.*, 2014) to allow the mother and daughter centrioles, which will have replicated at this point (see next section), to separate and participate in bipolar spindle formation.

Centriole Duplication

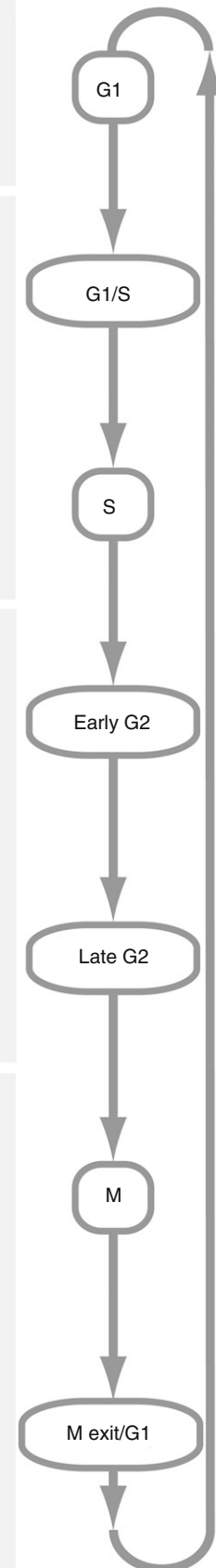
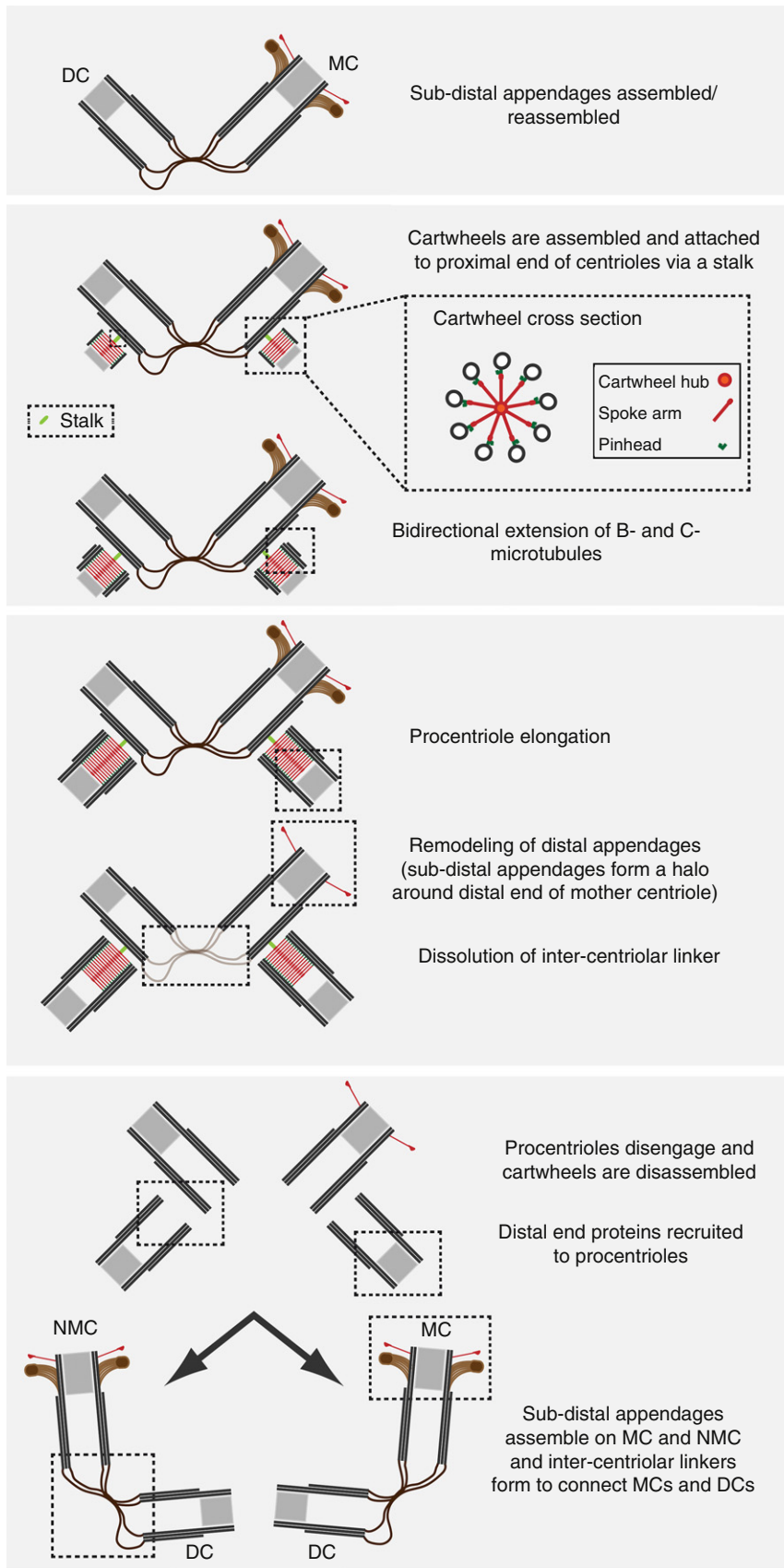
The centrosome is subject to strict copy number control and is duplicated only once per cell cycle. In G1 phase a cell usually possesses a single centrosome, containing a mother and a daughter centriole, but during the course of interphase the centrioles duplicate so that by the time a cell reaches G2 phase it has two centrosomes. Only one new centriole, a procentriole, is assembled at each existing centriole, meaning that at the end of G2 phase there are three generations of centrioles present in the cell, with one centrosome containing the mother centriole and a procentriole and the other containing the daughter centriole and a procentriole. How centriole duplication is controlled to ensure that only one procentriole forms at each existing centriole is not fully understood, but there is a licensing mechanism governing this process. It seems that once a procentriole begins to be assembled and is attached or engaged orthogonally to an existing centriole, another procentriole cannot form (Tsou and Stearns, 2006). Only when the procentrioles have disengaged from their parental centrioles, a process that takes place during mitosis, can centriole duplication occur again. In the absence of this licensing mechanism centriole amplification could occur and ultimately lead to chromosome instability (CIN) due to the improper segregation of chromosomes (Ganem *et al.*, 2009).

Many of the structural and regulatory components involved in centriole duplication were identified from RNAi and mutagenic screens carried out in *Drosophila* and *C. elegans* (Table 1; Dammermann *et al.*, 2004; Delattre *et al.*, 2004; Dobbelaere *et al.*, 2008; Leidel *et al.*, 2005). Procentriole assembly begins at the G1/S phase transition with the formation of a cartwheel structure that is attached to the proximal end of each existing centriole via a stalk (Guichard *et al.*, 2010). A cartwheel consists of a central hub with nine spoke arms (Figure 2). One of its key components is spindle assembly abnormal protein 6 (SAS-6) (Dammermann *et al.*, 2004; Leidel *et al.*, 2005), which is capable of oligomerizing via domains within its N-terminus, to form a ring-shaped structure with stalks emanating from it (Kitagawa *et al.*, 2011; van Breugel *et al.*, 2011). The radial periodicity of these stalks is 40° and corresponds well to the organization of the spoke arms, suggesting that SAS-6 alone might be responsible for conferring the ninefold symmetry to the centriole. It is to the spoke arms that the microtubule blades are attached, beginning with the A microtubule, which is capped at its minus end and elongates in a unidirectional manner (Guichard *et al.*, 2010). In contrast, the B and C microtubules are not capped and are extended from both their plus and minus ends (Figure 2). Capping of the plus ends of centriolar microtubules is important for regulating the overall length of the procentriole/centriole and CP110 has been identified as one of the factors controlling this process (Kleylein-Sohn *et al.*, 2007; Kohlmaier *et al.*, 2009; Schmidt *et al.*, 2009). Procentriole assembly also requires SAS-5/Ana2/SCL/TAL1 interrupting locus (STIL; Arquint *et al.*, 2012; Delattre *et al.*, 2004; Stevens *et al.*, 2010; Tang *et al.*, 2011; Vulprecht *et al.*, 2012), and Cep135 (Kleylein-Sohn *et al.*, 2007), which are probably cartwheel components, as well as Cep120 (Comartin *et al.*, 2013; Lin *et al.*, 2013) and SAS-4/CPAP (Kirkham *et al.*, 2003; Kohlmaier *et al.*, 2009; Leidel and Gonczy, 2003; Schmidt *et al.*, 2009; Tang *et al.*, 2009), which participate in the elongation of centriolar microtubules. Procentriole elongation is only finalized in mitosis and is dependent upon the recruitment of proteins such as hPOC5 (proteome of centriole 5) to the distal ends of procentrioles (Azimzadeh *et al.*, 2009).

There are many factors controlling the regulation of centriole duplication but there is one master regulator the kinase

Table 1 Key proteins required for centriole duplication

Caenorhabditis elegans	Drosophila melanogaster	Homo sapiens	Function
SAS-4	DSAS-4	CPAP	Involved in the assembly of the microtubule blades on to the cartwheel, as well as controlling centriole length
SAS-5	Ana2	STIL	Interacts with CPAP, probably a component of the cartwheel structure
SAS-6	DSAS-6	Hs SAS-6	Major cartwheel component dictating the ninefold symmetry of the procentriole/centriole
–	–	Cep120	Interacts with CPAP and regulates centriole length
–	Cep135	Cep135	Interacts with SAS-6 and CPAP suggesting that it is part of the pinhead
–	Asterless	Cep152	Scaffolding protein interacting with CPAP and PLK4
SPD-2	DSPD-2	Cep192	Controls PCM size and tethers PLK4 to centrioles
–	CP110	CP110	Caps centrioles and is involved in regulating centriole length
γ-Tubulin	γ-Tubulin	γ-Tubulin	Caps the minus end of the A-microtubule and localizes to a site within the lumen of the centriole
ZYG-1	SAK	PLK4	Master regulatory kinase controlling centriole assembly



Polo-like kinase 4 (PLK4) (Habedanck *et al.*, 2005), which is also known as Snk/Plk-akin kinase (SAK) in *Drosophila* (Bettencourt-Dias *et al.*, 2005) and its functional equivalent in *C. elegans* is ZYG-1 (O'Connell *et al.*, 2001). Overexpression of PLK4 can bypass the centriole licensing mechanisms and trigger centriole amplification, with multiple procentrioles forming around each parental centriole instead of just one. Due to its prominent role in centriole duplication, much attention has been focused on the regulation of PLK4 levels in the cell. PLK4 auto-regulates itself by trans-autophosphorylating residues within a degron motif located just outside its kinase domain, which targets the kinase for destruction via SCF ^{β -TrCP}-mediated ubiquitinylation (Cunha-Ferreira *et al.*, 2009; Guderian *et al.*, 2010; Rogers *et al.*, 2009; Sillibourne *et al.*, 2010). Counterbalancing this auto-destructive activity is the phosphatase PP2A that dephosphorylates the degron motif in PLK4 and relieves the kinase from ubiquitin-mediated degradation. The anchoring of PLK4 to centrioles has also been intensively studied and two PLK4-interacting proteins, Cep152 (Asterless in *Drosophila*) (Cizmecioglu *et al.*, 2010; Dzhindzhev *et al.*, 2010; Hatch *et al.*, 2010) and Cep192 (SPD-2 in *C. elegans*) (Kim *et al.*, 2013; Sonnen *et al.*, 2013) appear to act together in this process. Cep192 is involved in the recruitment of PCM to the centrosome (O'Connell *et al.*, 2000) and it is clear that there is interdependency between centriolar/centrosomal proteins for their localization to the centrosome, with the depletion of one protein affecting the centrosomal levels of another. It is important to mention that procentrioles are assembled within the PCM and in cases where the centriole duplication licensing mechanism is overridden, PCM size can influence the number of procentrioles formed. Artificial inflation of the PCM through overexpression of pericentrin in cells where procentrioles have been eliminated by laser ablation results in centriole amplification with the number formed being proportional to the amount of PCM (Loncarek *et al.*, 2008).

During centriole duplication there is the parallel process of centriole maturation, whereby the daughter centriole becomes a mother centriole. It has been observed that distal appendage proteins are recruited to the daughter centriole in late G2 (Graser *et al.*, 2007; Sillibourne *et al.*, 2013), but full maturation requires the transition through mitosis and it is in the subsequent cell cycle that the new mother centriole will fully mature by acquiring both sub-distal and distal appendages. This process is dependent on the mother centriole specific protein cenexin/ODF2 that is essential for the assembly of both the distal and sub-distal appendages (Ishikawa *et al.*,

2005). Other proteins are intimately associated to the duplication process, present in most and highly divergent species, but a comprehensive description of their roles is not yet known except in yeasts, which possess a simpler centrosome. This is the case for the centrin protein, a common marker of centrioles in most animal species, which is essential for the centrosome duplication in yeasts (Paoletti *et al.*, 2003).

Functions of the Centrosome

Microtubule Nucleation and Anchoring

The centrosome nucleates and anchors microtubules to organize a radial microtubule array during interphase and to assist in bipolar spindle formation during mitosis (Figures 3(a) and (b)). Microtubule nucleation is promoted by a number of factors but the most important is the γ -tubulin ring complex (γ -TuRC), which comprises six different subunits, γ -tubulin, GCP2, GCP3, GCP4, GCP5, and GCP6 (Teixido-Travesa *et al.*, 2012). In addition to these γ -TuRC components there are also a number of accessory factors including GCP-WD/Nedd1 and Mozart. Up to 80% of γ -tubulin is cytoplasmic (Moudjou *et al.*, 1996) and its centrosomal level varies during the course of the cell cycle, with it being at its lowest in early G1 and highest in M phase (Khodjakov and Rieder, 1999; Piehl *et al.*, 2004). The mitotic recruitment of microtubule nucleating factors to the centrosome is referred as centrosome maturation and is governed by the kinases Aurora A and PLK1 (Haren *et al.*, 2009). These phosphorylate the γ -TuRC accessory factors GCP-WD/Nedd1 and promote the recruitment of the γ -TuRC to the centrosome, leading to increased microtubule nucleation.

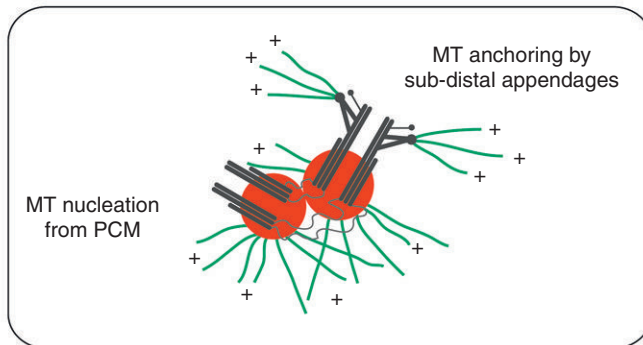
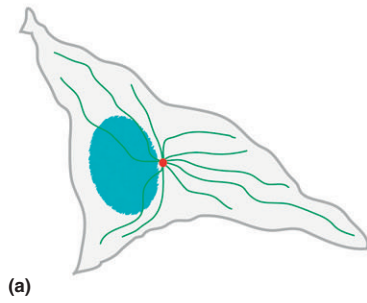
In addition to nucleating microtubules, the centrosome also anchors microtubules via the sub-distal appendages located on the mother centriole. This not only helps to generate a radial microtubule array, it also enables the mother centriole to position itself at the center of the cell (Piel *et al.*, 2000). Upon entry into mitosis, the sub-distal appendages are disassembled, resulting in decreased microtubule anchoring, coincident with increased nucleation activity.

Centrosomes in Cell Division

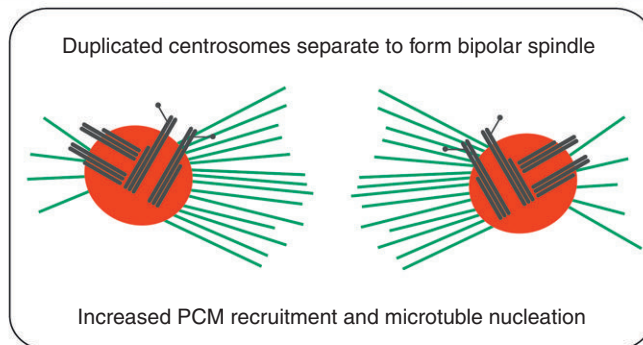
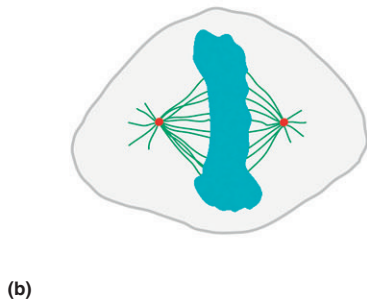
While there is a dramatic increase in the microtubule nucleation capacity of centrosomes during mitosis, they are not essential for bipolar spindle formation (Figure 3(b)), as this

Figure 2 Centriole duplication. Centriole duplication begins at the G1/S boundary with the assembly of a cartwheel at the proximal end of each existing centriole via a stalk (light green line). The cartwheel serves as a template for the assembly of a new centriole (procentriole) and consists of a hub from which nine spoke arms emanate. At the ends of the spoke arms there are pinhead structures to which the A-microtubules are connected and are elongated in a unidirectional manner. The B- and C-microtubules are elongated bidirectionally along the A-microtubules to form the centriolar barrel. Some distal luminal proteins, such as centrin and CP110, are recruited to the procentrioles after initiation of assembly (represented as grey rectangles). Procentriole elongation continues throughout interphase and is only completed in mitosis. Several structural changes occur at the mother centriole during late G2, notably the remodelling or disassembly of the sub-distal appendages to form a halo surrounding the mother centriole. In some cell lines, such as HeLa cells, the distal appendages are also disassembled although other cell lines maintain their distal appendages throughout the cell cycle. The inter-centriolar linker is dissolved at this point in the cell cycle, allowing the replicating centrosomes to separate (indicated by arrows). In mitosis (M) the cartwheel structures are disassembled from the procentrioles, which disengage from their parental centrioles. Upon exit from mitosis the sub-distal and distal appendages are assembled on the old and the new mother centriole (NMC). For simplicity the PCM is not shown in the figure although it should be noted that procentrioles do not possess their own PCM and only acquire it when they have matured to daughter centrioles.

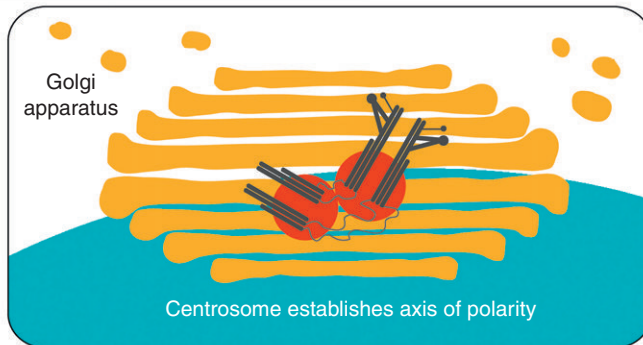
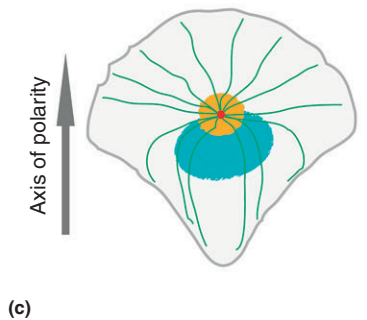
Microtubule organization



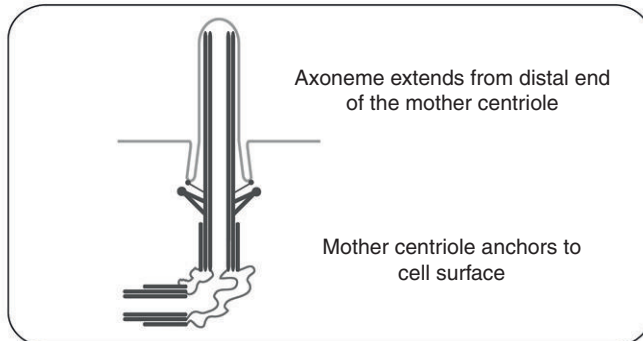
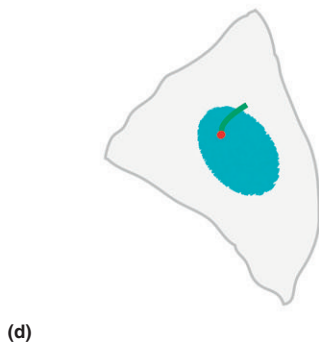
Bipolar spindle formation



Establishing polarity



Primary cilium formation



can occur through a chromatin-mediated pathway. However, centrosomes expedite bipolar spindle formation and in their absence the rate of CIN increases, indicating that they are required for maintaining genomic integrity (Sir *et al.*, 2013). Centrosomes also participate in the final stages of division, cytokinesis and abscission, where one of the two mother centrioles usually moves toward the cytokinetic bridge connecting the two daughter cells and apparently triggers cell abscission (Piel *et al.*, 2001).

Establishing Polarity

The idea that the centrosome is able to establish cell polarity was suggested by van Beneden in 1883, when he realised that a line drawn through the centrosome and the center of the nucleus defined an axis of polarity (Figure 3(c); Harris, 2000). In migrating adherent cells, such as fibroblasts, the centrosome is often positioned in front of the nucleus although this is not always the case, with the centrosome being located behind the nucleus in migrating non-adherent or weakly adherent cells, such as neutrophils. Regardless of the position of the centrosome relative to the nucleus, polarization of the microtubule network is still observed in both cases. The example shown in Figure 3(c) represents a cell confined on a crossbow-shaped adhesive micropattern, mimicking a migrating adherent cell, which has an asymmetrical microtubule (MT) array with more MTs orientated toward the leading edge than the trailing edge due to MT plus end stabilization at the leading edge (Théry *et al.*, 2006). This plus end MT stabilization reinforces cell polarity established by the centrosome. The close association between the centrosome and the Golgi apparatus (GA) establishes polarity by directing asymmetrical intracellular trafficking. One protein that is particularly important in mediating the connection between the centrosome and the GA is AKAP450 (A kinase-anchoring protein 450). This connection is important for coordinating directed trafficking, as the disruption of AKAP450 function through the overexpression of fragments responsible for targeting AKAP450 to the GA perturbs cell migration and the trafficking of vesicles from the GA (Hurtado *et al.*, 2011). These findings highlight the important role the centrosome plays in directing intracellular trafficking and establishing cell polarity.

Primary Cilium Formation

One of the key functions of the mother centriole, likely due to the evolutionary origin of the centrosome, is to template the assembly of a primary cilium, a sensory organelle that forms at

cell surface (Figure 3(d)). This is a multistep process that involves the recruitment of membrane to the distal end of the mother centriole, where it forms a vesicle capping the end of the centriole and into which the centriolar doublet microtubules are elongated to create an axoneme. During the process of primary ciliogenesis the mother centriole is anchored to the cell cortex, to enable extension of the primary cilium into the extracellular environment, and it is often referred to as a basal body in this context. Formation of a primary cilium is crucial for embryonic development and mutations in centrosomal proteins result in diseases known as ciliopathies that are characterized by mental retardation, polydactyl and retinal degeneration. Phylogeny of centrosome-associated proteins supports the notion that the centrosome evolved by direct filiation from the ancestral and highly conserved basal body/axoneme, i.e., of a cell compartment participating in two functions essential for cell polarity, sensation, and motion. Its reproduction based on a conservative mechanism (old/new) allows the transmission of polarity to daughter cells during cell division. Asymmetric orientation of activities in animal cells is critical for development and morphogenesis, for locomotion and chemotaxis, of multicellular organisms, and impaired cell polarity is the hallmark of malignant transformation. Only metazoa have evolved a centriole/basal body-containing centrosome and conserved all functions critical for cell polarity associated with the ancestral basal body/axoneme, sensation, motion and division. At the uni- to multicellular transition that lead to metazoa, selective pressure to maintain individual cell polarity could have triggered the transition from the plasma membrane-associated basal body/flagellum to the cytoplasmic centrosome, when growth requirements are no longer needed, to switch back to the former type of cell polarity by growing a primary cilium (Bornens, 2012). The transition from centrosome-to-primary cilium – a sensory organelle in quiescent cells – corresponds to a switch in cell polarity architecture.

The Centrosome in Development

There has been much debate as to whether the centrosome is necessary for development. The planarian *Schmidtea mediterranea* possesses ciliated epithelia and thus basal bodies/centrioles but lacks centrosomes, having lost centrosomal genes during evolution, and develops fully in the absence of centrosomes (Azimzadeh *et al.*, 2012). On the other hand, studies in *Drosophila* have shown that centrosomes are required for early embryonic development, but are dispensable for later stages of development. Loss-of-function mutations in 'asterless' (Cep152), probably the only PLK4-anchoring protein in

Figure 3 Functions of the centrosome. (a) The centrosome acts as a microtubule-organizing center by nucleating microtubules from the PCM and through the anchoring microtubules to the sub-distal appendages of the mother centriole. (b) At the end of G2 phase, the amount of PCM associated with the centrioles begins to increase, as does the microtubule nucleation capacity of the centrosome, thereby aiding in the formation of a bipolar spindle and ensuring the accurate segregation of sister chromatids. (c) The centrosome establishes polarity in the cell by organizing microtubules and through its connection with the Golgi apparatus that directly influences intracellular trafficking (image adapted from Théry, M., Racine, V., Piel, M., *et al.*, 2006. Anisotropy of cell adhesive microenvironment governs cell internal organization and orientation of polarity. Proceedings of the National Academy of Sciences 103, 19771–19776). (d) Upon exit from the cell cycle and entry into a quiescent state (G0), most cells form a primary cilium, a sensory organelle located at the surface of the cell. This is formed from the distal end of the mother centriole through the extension of the A- and B-microtubules to create an axoneme that is covered by a membrane sheath. At the base of the primary cilium, where the distal appendages are located, the ciliary membrane is invaginated to form a structure referred to as the ciliary pocket. In the left panel, showing cells, the DNA is labelled in blue, microtubules in green, the Golgi apparatus in yellow and centrosomes in red.

Drosophila, cause arrest in embryonic development soon after fertilization, whereas DSAS-4 mutants are able to develop into adult flies. However, these DSAS-4 mutant flies, while appearing morphologically normal, lack cilia in their sensory neurons and consequently exhibit an uncoordinated phenotype, being unable to walk or fly, and die soon after hatching. Male DSAS-4 flies are also sterile because their sperm lack flagella (Basto *et al.*, 2006; Varmark *et al.*, 2007). In the case of *C. elegans*, centrosomes are essential for embryonic development, where the first division is asymmetric and produces two founder cells of different fates. This is achieved by the segregation of different polarity factors, PAR proteins, to the anterior and posterior sides of the zygote and acentric positioning of the bipolar spindle that forms during the first division. Depletion of either SPD-2/Cep192 (O'Connell *et al.*, 2000) or SPD-5 (Hamill *et al.*, 2002), results in defective segregation of the PAR proteins and failed asymmetric division due to impaired PCM recruitment and microtubule nucleation, demonstrating the importance of the centrosome in this process.

Asymmetric cell division is also important for maintaining stem cell populations, so that one cell retains its self-renewing potential, while the other differentiates to a specialized cell type. In *Drosophila*, the male germline stem cells (GSCs) are maintained within a niche, where they are attached to hub cells and undergo asymmetric cell division to produce one cell that differentiates and another that maintains its stem cell fate. During this asymmetric division the older mother centriole is inherited by the stem cell, closest to the hub, while the differentiating cell receives the daughter centriole (Figure 4(a)). The asymmetric partitioning of the mother centriole is thought to be linked to its greater microtubule nucleating capacity that enables it to anchor to the hub/GSC interface (Yamashita *et al.*, 2007). It has been shown that if the mother centrosome is not positioned at the hub/GSC interface, cell cycle arrest occurs until it is repositioned (Cheng *et al.*, 2008). This delay serves as a mechanism to preserve the stem cell population, but it causes a reduction in both the size of the differentiating cell population and the rate of spermatogenesis. In ageing male *Drosophila*, centrosome misorientation occurs more frequently and this correlates with reduced spermatogenesis, demonstrating that asymmetric centrosome segregation is important for maintaining stem and differentiating cell populations. While the mother centriole is inherited in GSCs, a different scenario takes place in neuroblasts of the same species where the daughter centriole is inherited by the stem cell (Conduit and Raff, 2010; Januschke *et al.*, 2013). The relevance of this difference still remains to be determined.

Differences in mother versus daughter centriole inheritance have also been observed during asymmetric cell division in the mouse developing neocortex. Radial glial progenitors divide asymmetrically at the ventricular zone (VZ) to produce a self-renewing glial cell and a differentiating cell, usually a neuron, which migrates away from the VZ to form the cortical plate of the neocortex (Figure 4(b)). In this asymmetric cell division the self-renewing glial cell invariably inherits the older (mother) centrosome and this is important for maintaining the progenitor pool, as the ablation of ninein by RNAi results in depletion of cells from the VZ (Wang *et al.*, 2009). Asymmetric centrosome inheritance might have importance

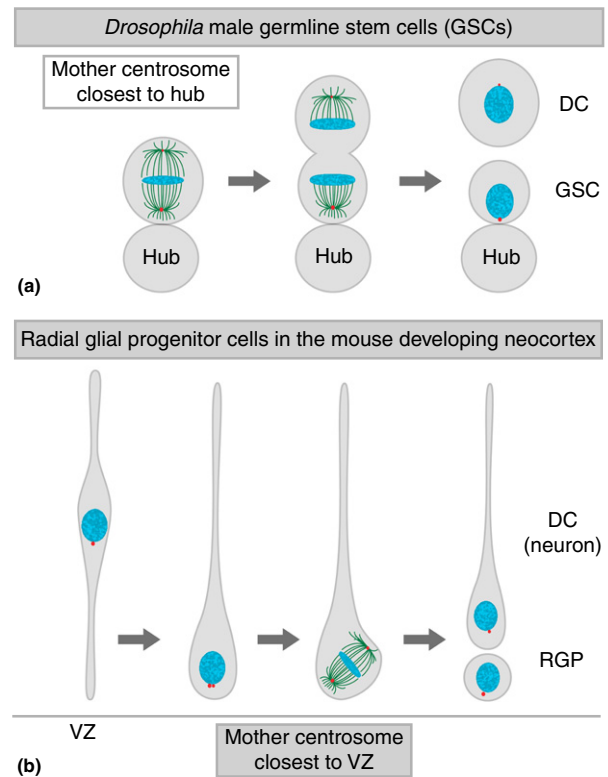


Figure 4 Asymmetric centrosome inheritance. (a) In *Drosophila*, male germline stem cells (GSCs) divide asymmetrical to produce a self-renewing stem cell and a differentiating cell (DC). The mother centrosome of the GSC nucleates more microtubules and is anchored to the side closest to the hub; during division the GSC invariably inherits the mother centrosome while the differentiating cell inherits the daughter centrosome. It should be noted that this situation is reversed in *Drosophila* neuroblasts where the stem cell inherits the daughter centrosome rather than the mother centrosome. (b) In the developing mouse neocortex, the mother centrosome is inherited by the radial glial progenitor cell (RGP), after dividing close to the ventricular zone, while the daughter centrosome is inherited by the differentiating cell (DC), which will ultimately become a neuron.

consequences for the fate of the cell, as it has been shown that the older mother centriole forms a primary cilium before the newly formed mother centriole and has the ability to respond to extracellular signals first (Paridaen *et al.*, 2013).

Finally, somatic centrosomes are able to induce parthenogenesis – development in the absence of fertilization – in amphibian eggs. The pricking of *Xenopus* oocytes has been found to trigger cell cycling resumption but it does not result in cell cleavage or development, whereas the injection of centrosomes, even when they are heterologous, induces parthenogenesis (Tournier *et al.*, 1989). Furthermore, only replication competent somatic centrosomes are capable of inducing parthenogenesis, indicating that duplication of the centrosome is required for development.

Centrosome in Disease

The centrosome has long been implicated in cancer. Examination of tumor biopsies revealed that cancerous cells

frequently possessed supernumerary centrosomes capable of nucleating microtubules, leading to the suggestion that these additional microtubule-organizing centers (MTOCs) could lead to multipolar spindle formation, unequal chromosome segregation, and aneuploidy (**Figure 5(a)**), an idea first proposed by Boveri in 1902 (**Boveri, 2008**). Testing of this hypothesis has proven difficult, because centrosome amplification can occur as a result of cell division defects, and has it required the use of model systems such as *D. melanogaster*, in which centrosome amplification can be induced in a controllable manner to address this problem. Induced overexpression of SAK/PLK4 in developing fly embryos after the early syncytial stage, during which centrosomes are essential (**Varmark et al., 2007**), results in a short developmental delay due to the presence of supernumerary centrioles, but the flies develop normally and exhibit only a low level of aneuploidy (**Basto et al., 2008**). This is due to an efficient centrosome clustering mechanism that is dependent on the non-essential kinesin *ncd*/HSET and enables a bipolar spindle to form that contains multiple centrosomes at each pole (**Basto et al., 2008; Kwon et al., 2008**). However, subsequent transplantation of larval brain cells, containing supernumerary centrosomes, into wild-type hosts resulted in metastatic tumor formation, demonstrating that centrosome amplification can promote tumorigenesis (**Basto et al., 2008**). Centrosome loss or centrosome dysregulation can also lead to tumor formation in

Drosophila larval neural stem cells, reiterating the importance of the centrosome in maintaining cell polarity during asymmetric stem cell division (**Castellanos et al., 2008**).

Inhibitors of HSET are being developed (**Watts et al., 2013**) because inhibition of its activity does not affect division in normal cells, but it disrupts centrosome clustering in cancer cells, causing them to form multipolar spindles and produce unviable progeny that undergo apoptosis (**Ganem et al., 2009**). Even though centrosome clustering mechanisms exist to prevent multipolar spindle formation, supernumerary centrosomes still contribute to CIN by promoting the formation of merotelic kinetochore-microtubule attachments (**Ganem et al., 2009**), where microtubules from the two opposing spindle poles attach to a single kinetochore (**Figure 5(b)**). This type of attachment satisfies the spindle assembly checkpoint (SAC), thereby allowing chromatid separation, and results in chromosome missegregation due to the presence of a lagging chromosome (**Ganem et al., 2009**).

Summary

- The centrosome consists of two microtubule-based centrioles (a mother and a daughter centriole) that differ in age and are structurally similar but not identical.
- The centrosome is the major microtubule-organizing center in the cell.

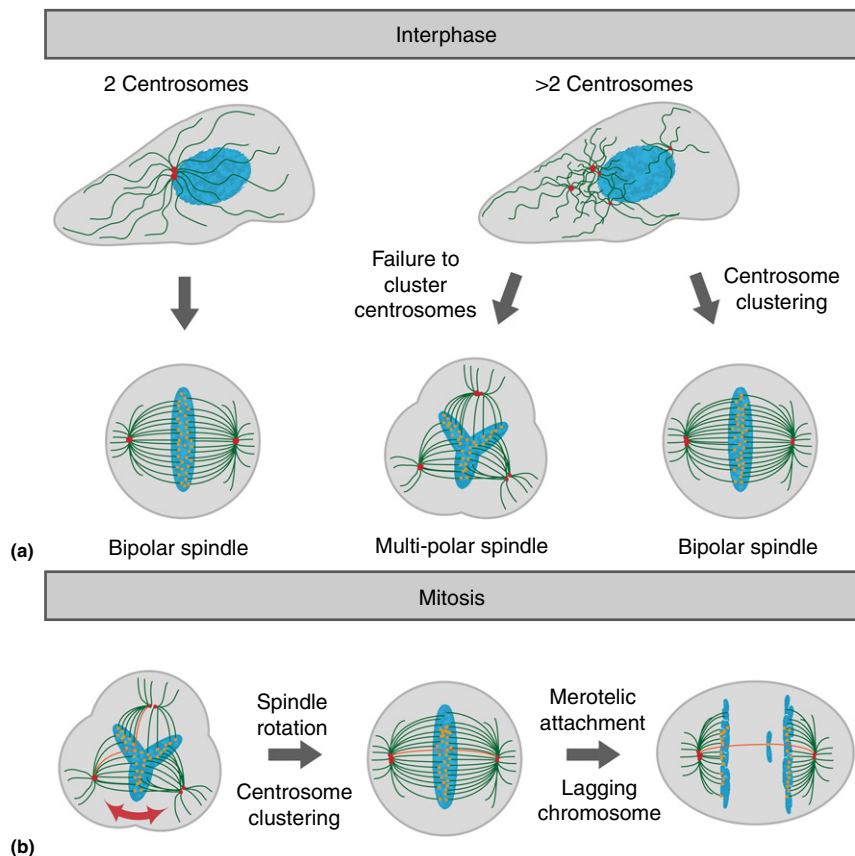


Figure 5 Centrosome amplification. (a) Upon entry into mitosis, cells possessing supernumerary centrosomes form bipolar spindles, unless centrosome clustering mechanisms fail, in which case multipolar spindles form. (b) Supernumerary centrosomes promote merotelic attachment where microtubules from opposing poles attach to a single kinetochore (orange microtubule), which can lead to aneuploidy due to chromosome missegregation.

- During interphase the centrosome is responsible for creating a microtubule array and during mitosis it assists in bipolar spindle assembly.
- Tissue development and homeostasis depend on the polarizing activity of the centrosome.
- The centrosome is involved in cell cycle progression and interconverts to a primary cilium when the cell enters a quiescent state (G0).
- Centrosome amplification is frequently observed in tumors where it promotes CIN.
- There are many outstanding questions in centrosome biology with the precise mechanisms of centriole duplication and the centrosome-to-primary cilium interconversion remaining to be fully elucidated, as well as the centrosome's role in development and tumorigenesis.

References

- Agircan, F.G., Schiebel, E., Mardin, B.R., 2014. Separate to operate: Control of centrosome positioning and separation. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 369.
- Arquint, C., Sonnen, K.F., Stierhof, Y.D., Nigg, E.A., 2012. Cell-cycle-regulated expression of STIL controls centriole number in human cells. *Journal of Cell Science* 125, 1342–1352.
- Azimzadeh, J., Hergert, P., Delouvee, A., *et al.*, 2009. hPOC5 is a centrin-binding protein required for assembly of full-length centrioles. *Journal of Cell Biology* 185, 101–114.
- Azimzadeh, J., Wong, M.L., Downhour, D.M., Sanchez Alvarado, A., Marshall, W.F., 2012. Centrosome loss in the evolution of planarians. *Science* 335, 461–463.
- Basto, R., Brunk, K., Vinadogrova, T., *et al.*, 2008. Centrosome amplification can initiate tumorigenesis in flies. *Cell* 133, 1032–1042.
- Basto, R., Lau, J., Vinogradova, T., *et al.*, 2006. Flies without centrioles. *Cell* 125, 1375–1386.
- Bettencourt-Dias, M., Rodrigues-Martins, A., Carpenter, L., *et al.*, 2005. SAK/PLK4 is required for centriole duplication and flagella development. *Current Biology* 15, 2199–2207.
- Bobinnec, Y., Khodjakov, A., Mir, L.M., *et al.*, 1998. Centriole disassembly *in vivo* and its effect on centrosome structure and function in vertebrate cells. *Journal of Cell Biology* 143, 1575–1589.
- Bornens, M., 2012. The centrosome in cells and organisms. *Science* 335, 422–426.
- Boveri, T., 2008. Concerning the origin of malignant tumours by Theodor Boveri. Translated and annotated by Henry Harris. *Journal of Cell Science* 121 (Suppl. 1), 1–84.
- van Breugel, M., Hirono, M., Andreeva, A., *et al.*, 2011. Structures of SAS-6 suggest its organization in centrioles. *Science* 331, 1196–1199.
- Castellanos, E., Dominguez, P., Gonzalez, C., 2008. Centrosome dysfunction in *Drosophila* neural stem cells causes tumors that are not due to genome instability. *Current Biology* 18, 1209–1214.
- Cheng, J., Turkel, N., Hemati, N., *et al.*, 2008. Centrosome misorientation reduces stem cell division during ageing. *Nature* 456, 599–604.
- Cizmecioglu, O., Arnold, M., Bahtz, R., *et al.*, 2010. Cep152 acts as a scaffold for recruitment of Plk4 and CPAP to the centrosome. *Journal of Cell Biology* 191, 731–739.
- Comartin, D., Gupta, G.D., Fussner, E., *et al.*, 2013. CEP120 and SPICE1 cooperate with CPAP in centriole elongation. *Current Biology* 23, 1360–1366.
- Conduit, P.T., Raff, J.W., 2010. Cnn dynamics drive centrosome size asymmetry to ensure daughter centriole retention in *Drosophila* neuroblasts. *Current Biology* 20, 2187–2192.
- Cunha-Ferreira, I., Rodrigues-Martins, A., Bento, I., *et al.*, 2009. The SCF/Slimb ubiquitin ligase limits centrosome amplification through degradation of SAK/PLK4. *Current Biology* 19, 43–49.
- Dammermann, A., Muller-Reichert, T., Pelletier, L., *et al.*, 2004. Centriole assembly requires both centriolar and pericentriolar material proteins. *Developmental Cell* 7, 815–829.
- Delattre, M., Leidel, S., Wani, K., *et al.*, 2004. Centriolar SAS-5 is required for centrosome duplication in *C. elegans*. *Nature Cell Biology* 6, 656–664.
- Delgehr, N., Sillibourne, J., Bornens, M., 2005. Microtubule nucleation and anchoring at the centrosome are independent processes linked by ninein function. *Journal of Cell Science* 118, 1565–1575.
- Dobbelaere, J., Josue, F., Suijkerbuijk, S., *et al.*, 2008. A genome-wide RNAi screen to dissect centriole duplication and centrosome maturation in *Drosophila*. *PLoS Biology* 6, e224.
- Dzhindzhev, N.S., Yu, Q.D., Weiskopf, K., *et al.*, 2010. Asterless is a scaffold for the onset of centriole assembly. *Nature* 467, 714–718.
- Ganem, N.J., Godinho, S.A., Pellman, D., 2009. A mechanism linking extra centrosomes to chromosomal instability. *Nature* 460, 278–282.
- Graser, S., Stierhof, Y.D., Lavoie, S.B., *et al.*, 2007. Cep164, a novel centriole appendage protein required for primary cilium formation. *Journal of Cell Biology* 179, 321–330.
- Guarguaglini, G., Duncan, P.I., Stierhof, Y.D., *et al.*, 2005. The forkhead-associated domain protein Cep170 interacts with Polo-like kinase 1 and serves as a marker for mature centrioles. *Molecular Biology of the Cell* 16, 1095–1107.
- Guderian, G., Westendorf, J., Uldschmid, A., Nigg, E.A., 2010. Plk4 trans-autophosphorylation regulates centriole number by controlling betaTrCP-mediated degradation. *Journal of Cell Science* 123, 2163–2169.
- Guichard, P., Chretien, D., Marco, S., Tassin, A.M., 2010. Procentriole assembly revealed by cryo-electron tomography. *EMBO Journal* 29, 1565–1572.
- Guichard, P., Hachet, V., Majubu, N., *et al.*, 2013. Native architecture of the centriole proximal region reveals features underlying its 9-fold radial symmetry. *Current Biology* 23, 1620–1628.
- Habedanck, R., Stierhof, Y.D., Wilkinson, C.J., Nigg, E.A., 2005. The Polo kinase Plk4 functions in centriole duplication. *Nature Cell Biology* 11, 1140–1146.
- Hamill, D.R., Severson, A.F., Carter, J.C., Bowerman, B., 2002. Centrosome maturation and mitotic spindle assembly in *C. elegans* require SPD-5, a protein with multiple coiled-coil domains. *Developmental Cell* 3, 673–684.
- Haren, L., Stearns, T., Luders, J., 2009. Plk1-dependent recruitment of gamma-tubulin complexes to mitotic centrosomes involves multiple PCM components. *PLoS One* 4, e5976.
- Harris, H., 2000. *The Birth of the Cell*. New Haven and London: Yale University Press.
- Hatch, E.M., Kulukian, A., Holland, A.J., Cleveland, D.W., Stearns, T., 2010. Cep152 interacts with Plk4 and is required for centriole duplication. *Journal of Cell Biology* 191, 721–729.
- Hurtado, L., Caballero, C., Gavilan, M.P., *et al.*, 2011. Disconnecting the Golgi ribbon from the centrosome prevents directional cell migration and ciliogenesis. *Journal of Cell Biology* 193, 917–933.
- Ishikawa, H., Kubo, A., Tsukita, S., 2005. Odf2-deficient mother centrioles lack distal/subdistal appendages and the ability to generate primary cilia. *Nature Cell Biology* 7, 517–524.
- Januschke, J., Reina, J., Llamazares, S., *et al.*, 2013. Centrobins controls mother-daughter centriole asymmetry in *Drosophila* neuroblasts. *Nature Cell Biology* 15, 241–248.
- Khodjakov, A., Rieder, C.L., 1999. The sudden recruitment of gamma-tubulin to the centrosome at the onset of mitosis and its dynamic exchange throughout the cell cycle, do not require microtubules. *Journal of Cell Biology* 146, 585–596.
- Kim, T.S., Park, J.E., Shukla, A., *et al.*, 2013. Hierarchical recruitment of Plk4 and regulation of centriole biogenesis by two centrosomal scaffolds, Cep192 and Cep152. *Proceedings of the National Academy of Sciences of the United States of America* 110 (50), E4849–E4857.
- Kirkham, M., Muller-Reichert, T., Oegema, K., Grill, S., Hyman, A.A., 2003. SAS-4 is a *C. elegans* centriolar protein that controls centrosome size. *Cell* 112, 575–587.
- Kitagawa, D., Vakonakis, I., Olieric, N., *et al.*, 2011. Structural basis of the 9-fold symmetry of centrioles. *Cell* 144, 364–375.
- Kleylein-Sohn, J., Westendorf, J., Le Clech, M., *et al.*, 2007. Plk4-induced centriole biogenesis in human cells. *Developmental Cell* 13, 190–202.
- Kohlmaier, G., Loncarek, J., Meng, X., *et al.*, 2009. Overly long centrioles and defective cell division upon excess of the SAS-4-related protein CPAP. *Current Biology* 19, 1012–1018.
- Kwon, M., Godinho, S.A., Chandhok, N.S., *et al.*, 2008. Mechanisms to suppress multipolar divisions in cancer cells with extra centrosomes. *Genes & Development* 22, 2189–2203.
- Leidel, S., Delattre, M., Cerutti, L., Baumer, K., Gonczy, P., 2005. SAS-6 defines a protein family required for centrosome duplication in *C. elegans* and in human cells. *Nature Cell Biology* 7, 115–125.
- Leidel, S., Gonczy, P., 2003. SAS-4 is essential for centrosome duplication in *C. elegans* and is recruited to daughter centrioles once per cell cycle. *Developmental Cell* 4, 431–439.
- Lin, Y.N., Wu, C.T., Lin, Y.C., *et al.*, 2013. CEP120 interacts with CPAP and positively regulates centriole elongation. *Journal of Cell Biology* 202, 211–219.

- Loncarek, J., Hergert, P., Magidson, V., Khodjakov, A., 2008. Control of daughter centriole formation by the pericentriolar material. *Nature Cell Biology* 10, 322–328.
- Mennella, V., Agard, D.A., Huang, B., Pelletier, L., 2014. Amorphous no more: Subdiffraction view of the pericentriolar material architecture. *Trends in Cell Biology* 24, 188–197.
- Moudjou, M., Bordes, N., Paintrand, M., Bornens, M., 1996. gamma-Tubulin in mammalian cells: The centrosomal and the cytosolic forms. *Journal of Cell Science* 109 (Pt 4), 875–887.
- O'Connell, K.F., Caron, C., Kopish, K.R., *et al.*, 2001. The *C. elegans* zyg-1 gene encodes a regulator of centrosome duplication with distinct maternal and paternal roles in the embryo. *Cell* 105, 547–558.
- O'Connell, K.F., Maxwell, K.N., White, J.G., 2000. The spd-2 gene is required for polarization of the anteroposterior axis and formation of the sperm asters in the *Caenorhabditis elegans* zygote. *Developmental Biology* 222, 55–70.
- Paintrand, M., Moudjou, M., Delacroix, H., Bornens, M., 1992. Centrosome organization and centriole architecture: Their sensitivity to divalent cations. *Journal of Structural Biology* 108, 107–128.
- Paoletti, A., Bordes, N., Haddad, R., *et al.*, 2003. Fission yeast cdc31p is a component of the half-bridge and controls SPB duplication. *Molecular Biology of the Cell* 14, 2793–2808.
- Paridaen, J.T., Wilsch-Brauninger, M., Huttner, W.B., 2013. Asymmetric inheritance of centrosome-associated primary cilium membrane directs ciliogenesis after cell division. *Cell* 155, 333–344.
- Piehl, M., Tulu, U.S., Wadsworth, P., Cassimeris, L., 2004. Centrosome maturation: Measurement of microtubule nucleation throughout the cell cycle by using GFP-tagged EB1. *Proceedings of the National Academy of Sciences of the United States of America* 101, 1584–1588.
- Piel, M., Meyer, P., Khodjakov, A., Rieder, C.L., Bornens, M., 2000. The respective contributions of the mother and daughter centrioles to centrosome activity and behavior in vertebrate cells. *Journal of Cell Biology* 149, 317–330.
- Piel, M., Nordberg, J., Euteneuer, U., Bornens, M., 2001. Centrosome-dependent exit of cytokinesis in animal cells. *Science* 291, 1550–1553.
- Rogers, G.C., Rusan, N.M., Roberts, D.M., Peifer, M., Rogers, S.L., 2009. The SCF Slimb ubiquitin ligase regulates Plk4/Sak levels to block centriole reduplication. *Journal of Cell Biology* 184, 225–239.
- Scheer, U., 2014. Historical roots of centrosome research: Discovery of Boveri's microscope slides in Würzburg. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 369.
- Schmidt, T.I., Kleylein-Sohn, J., Westendorf, J., *et al.*, 2009. Control of centriole length by CPAP and CP110. *Current Biology* 19, 1005–1011.
- Sillibourne, J.E., Hurbain, I., Grand-Perret, T., *et al.*, 2013. Primary ciliogenesis requires the distal appendage component Cep123. *Biology Open* 2, 535–545.
- Sillibourne, J.E., Tack, F., Vloemans, N., *et al.*, 2010. Autophosphorylation of polo-like kinase 4 and its role in centriole duplication. *Molecular Biology of the Cell* 21, 547–561.
- Sir, J.H., Putz, M., Daly, O., *et al.*, 2013. Loss of centrioles causes chromosomal instability in vertebrate somatic cells. *Journal of Cell Biology* 203, 747–756.
- Sonnen, K.F., Gabryjonczyk, A.M., Anselm, E., Stierhof, Y.D., Nigg, E.A., 2013. Human Cep192 and Cep152 cooperate in Plk4 recruitment and centriole duplication. *Journal of Cell Science* 126, 3223–3233.
- Stevens, N.R., Dobbelaere, J., Brunk, K., Franz, A., Raff, J.W., 2010. *Drosophila* Ana2 is a conserved centriole duplication factor. *Journal of Cell Biology* 188, 313–323.
- Tang, C.J., Fu, R.H., Wu, K.S., Hsu, W.B., Tang, T.K., 2009. CPAP is a cell-cycle regulated protein that controls centriole length. *Nature Cell Biology* 11, 825–831.
- Tang, C.J., Lin, S.Y., Hsu, W.B., *et al.*, 2011. The human microcephaly protein STIL interacts with CPAP and is required for procentriole formation. *EMBO Journal* 30, 4790–4804.
- Teixido-Travesa, N., Roig, J., Luders, J., 2012. The where, when and how of microtubule nucleation — One ring to rule them all. *Journal of Cell Science* 125, 4445–4456.
- Théry, M., Racine, V., Piel, M., *et al.*, 2006. Anisotropy of cell adhesive microenvironment governs cell internal organization and orientation of polarity. *Proceedings of the National Academy of Sciences* 103, 19771–19776.
- Tournier, F., Karsenti, E., Bornens, M., 1989. Parthenogenesis in *Xenopus* eggs injected with centrosomes from synchronized human lymphoid cells. *Developmental Biology* 136, 321–329.
- Tsou, M.F., Stearns, T., 2006. Mechanism limiting centrosome duplication to once per cell cycle. *Nature* 442, 947–951.
- Varmark, H., Llamazares, S., Rebollo, E., *et al.*, 2007. Asterless is a centriolar protein required for centrosome function and embryo development in *Drosophila*. *Current Biology* 17, 1735–1745.
- Vulprecht, J., David, A., Tibelius, A., *et al.*, 2012. STIL is required for centriole duplication in human cells. *Journal of Cell Science* 125, 1353–1362.
- Wang, X., Tsai, J.W., Imai, J.H., *et al.*, 2009. Asymmetric centrosome inheritance maintains neural progenitors in the neocortex. *Nature* 461, 947–955.
- Watts, C.A., Richards, F.M., Bender, A., *et al.*, 2013. Design, synthesis, and biological evaluation of an allosteric inhibitor of HSET that targets cancer cells with supernumerary centrosomes. *Chemistry & Biology* 20, 1399–1410.
- Yamashita, Y.M., Mahowald, A.P., Perlin, J.R., Fuller, M.T., 2007. Asymmetric inheritance of mother versus daughter centrosome in stem cell division. *Science* 315, 518–521.

Cilia and Flagella

F Schmid, ST Christensen, and LB Pedersen, University of Copenhagen, Universitetsparken, Copenhagen OE, Denmark

© 2016 Elsevier Inc. All rights reserved.

Glossary

Axoneme The microtubule-based skeleton of cilia and flagella, which usually consists of nine peripheral outer doublets of microtubules and 0 or 2 central microtubule singlets. In motile (9 + 2) cilia, the axoneme contains additional structures required for motility, including outer and inner dynein arms, radial spokes, nexin links, and central pair projections.

Basal body Modified centriole that provides the basis for the outgrowth of the ciliary and flagellar microtubule skeleton, the axoneme.

Centriole Tubulin-based barrel-shaped Encyclopedia of Cell Biology structure composed of 9 microtubule triplets.

Centrioles are the core components of the centrosome and also function as basal bodies in cells with cilia or flagella.

Centrosome Microtubule organizing center composed of two orthogonally arranged centrioles and pericentriolar material, which nucleates and anchors microtubules.

During mitosis, centrosomes function as mitotic spindle poles. In ciliogenesis, the older of the two centrioles (the mother centriole) functions as the basal body from which the ciliary axoneme emanates.

Ciliopathy Genetic disorder whose affected gene product (s) localizes to the cilium-centrosome axis.

Dyneins Motor proteins that consist of several subunits and move from the plus to the minus end of microtubules upon hydrolysis of ATP. In cilia and flagella, dyneins provide the force for motility and for retrograde intraflagellar transport.

Intraflagellar transport Process by which proteins are transported along the axoneme of the cilium from the base to the tip and back. Intraflagellar transport is essential for assembly and maintenance of most cilia and flagella, and is powered by kinesin and dynein motor proteins.

Kinesins Motor proteins that move along microtubules upon hydrolysis of ATP. Most kinesins move toward the plus end of microtubules, but some can move towards the minus end and promote microtubule disassembly.

Primary cilium A non-motile 9 + 0 cilium present in a single copy on the surface of many quiescent or differentiated vertebrate cells. Primary cilia are enriched for numerous signaling proteins and play important roles during development and in the adult as coordinators of many different signaling pathways.

Transition zone Region between the basal body and the cilium proper, which is thought to act as a barrier for selective transport of molecules into and out of the ciliary compartment.

Introduction

Cilia and flagella are microtubule-based organelles protruding from the surface of a wide variety of eukaryotic cells. There are two general types of cilia: motile cilia and non-motile cilia. Motile cilia are usually present in large numbers per cell, for example, in the epithelium of the human airway, oviduct and brain ventricles, or on the surface of protists such as *Paramecium* and *Tetrahymena* (Figure 1(a)). These cilia beat with a forward power stroke and a backward recovery stroke and function in cell locomotion or movement of fluid across epithelia. Flagella are also motile, but are usually present in a single copy per cell and motion is often planar and wave-like from base to the tip of the organelle producing a propellar-like motion (Figure 2(a); Ibanez-Tallon *et al.*, 2003; Satir and Christensen, 2006; Ginger *et al.*, 2008). Flagella and motile cilia are ultrastructurally similar (the terms cilia and flagella are used interchangeably), typically comprising a 9 + 2 microtubule axoneme with radial spokes (RS) and outer (ODAs) and inner (IDAs) dynein arms (Figures 1(b), and 1(c)). Non-motile cilia, such as the so-called primary cilia (Sorokin, 1968), which are present on many cell types of the human body, are solitary organelles that function as specialized signaling devices (Satir and Christensen, 2006; Satir *et al.*, 2010). These cilia typically have a 9 + 0 axoneme and lack motility-related structures such as dynein arms (Figures 1(b), and 1(c)).

Proteomics, transcriptomics, and comparative genomics studies have revealed that cilia contain hundreds of different proteins, many of which are highly conserved between different species (Arnaiz *et al.*, 2009). Some of these proteins assemble to form the elaborate structures of the microtubule axoneme or are incorporated into the ciliary membrane that surrounds it. Cilia are typically rod-shaped, ranging in length from a few (e.g., chondrocyte primary cilia) to several hundred microns in length (e.g., olfactory cilia). Some types of non-motile cilia have a highly specialized morphology. For example, the vertebrate rod and cone photoreceptors consist of an inner segment linked to the outer segment by a short connecting cilium, which ends in a highly differentiated outer segment composed of a cylinder-shaped stack of membrane disks (Insinna and Besharse, 2008). Certain neuronal sensory cilia of the nematode *Caenorhabditis elegans* also display atypical cilia with wing- or finger-like membranous appendages (Silverman and Leroux, 2009).

Although motile and non-motile cilia differ in their axoneme structure and function, all cilia have the ability to perceive signals from the environment and transmit these to the cell body (Christensen *et al.*, 2007). While it has been known for many years that motile cilia are important for many physiological processes, such as sperm locomotion and clearance of mucus from the airways (Ibanez-Tallon *et al.*, 2003), the functional importance of non-motile cilia has only recently

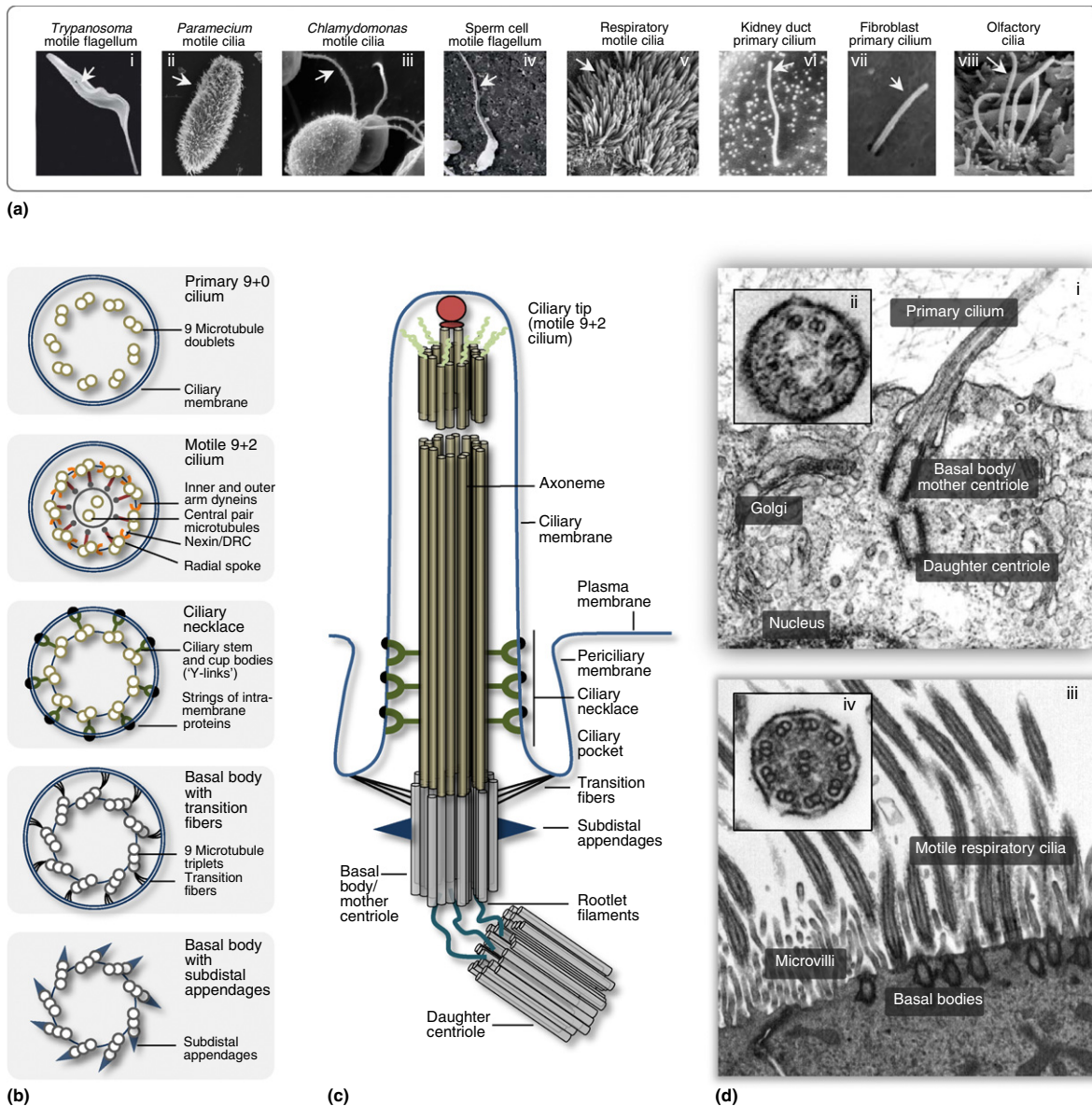


Figure 1 Overview of different types of cilia and flagella. (a) Examples of organisms and cell types with different types of motile or non-motile cilia. Images reproduced from the following references with permission: (i) Vaughan *et al.*, 2008, (ii) Vincensini *et al.*, 2011, (iii) and (v) Dartmouth College Electron Microscope Facility (<http://www.dartmouth.edu/~emlab/gallery/>), (iv) Sanchez and Mardomingo, 2007, (vi) Wheatley *et al.*, 1996, (vii) Christensen *et al.*, 2012, and (viii) Zeiske *et al.*, 2003. (b) Cartoon illustrating cross-section through a 9+0 and 9+2 cilium as well as the ciliary necklace and basal body region. (c) Cartoon showing longitudinal view of a 9+0 cilium and associated structures. Also included is a diagram illustrating typical tip structures of a 9+2 cilium. (d) Transmission electron micrographs of cross-sections (insets) or longitudinal sections of a primary 9+0 cilium (upper panel) or motile 9+2 cilia (lower panel). Reproduced with permission from (i) Jensen *et al.*, 2004, (ii) Kiprilov *et al.*, 2008, (iii) and (iv) Dartmouth College Electron Microscope Facility (<http://www.dartmouth.edu/~emlab/gallery/>).

been appreciated. For example, primary cilia were discovered more than a century ago (Zimmermann, 1898) but were, until recently, considered vestigial organelles. It is now well established that primary cilia are indispensable sensory organelles that coordinate a variety of important signaling pathways during development and in tissue homeostasis (Goetz and Anderson, 2010; Satir *et al.*, 2010). Consequently, mutations in genes that lead to dysfunction of motile or primary cilia can lead to diseases and developmental defects in humans. These

diseases are called ciliopathies (Ibanez-Tallon *et al.*, 2003; Hildebrandt *et al.*, 2011).

Structure of Cilia and Flagella

Motile cilia and flagella are widely used in protists and metazoans for locomotion, extracellular fluid flow and sensory perception (Ibanez-Tallon *et al.*, 2003; Ginger *et al.*, 2008).

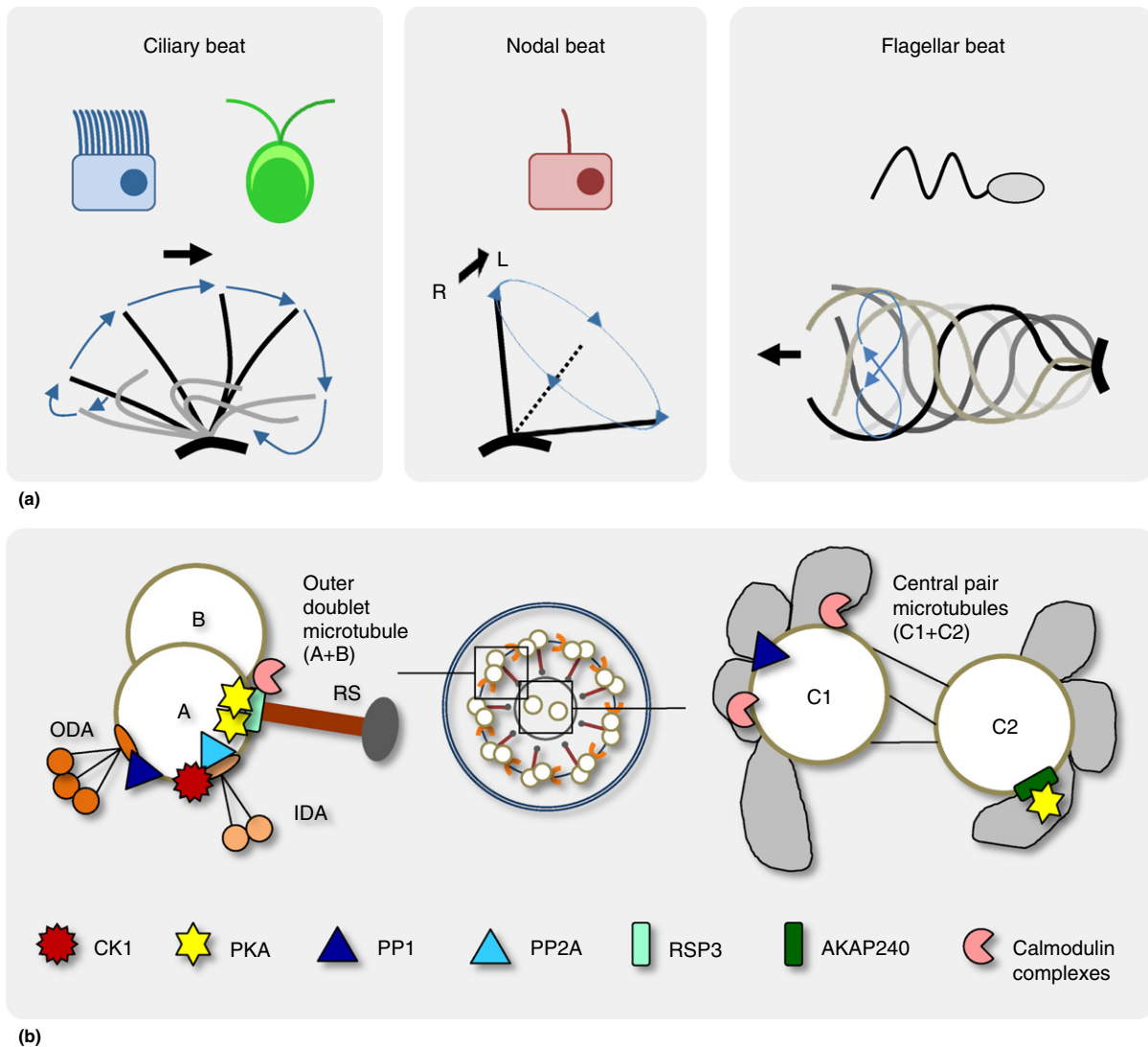


Figure 2 Structural and functional features of motile cilia and flagella. (a) Distinct beating patterns of motile cilia in different organisms and cell types. Cilia on multi-ciliated epithelial cells and *Chlamydomonas* use a power stroke (black lines) followed by a recovery stroke (gray lines) for movement in or of extracellular fluid. In contrast, single motile cilia on the embryonic node of vertebrates rotate around a tilted axis (dashed line) to establish a morphogen gradient and/or to induce Ca^{2+} influx in crown cells on the left-hand side of the node, via activation of the polycystins (PC2 and the PC1 homolog, PC1 like 1) in mechanosensory, immotile primary cilia, which is necessary for the establishment of left-right symmetry (Hirokawa *et al.*, 2006). The flagellum of sperm cells oscillates during locomotion. Black arrow represents net flow. (b) Cartoon illustrating cross-section of a motile 9+2 cilium showing associated structures on microtubule outer doublets (left; A and B) or central pair microtubules (right; C1 and C2), including the RS, ODA, IDA, kinases PKA and CK1, phosphatases PP2A and PP1, AKAP240 and RSP3, calmodulin complexes, and central pair appendages (gray). Figure panels redrawn from Okada *et al.* (2005), Wirschell *et al.* (2011), and Wirschell *et al.* (2013) and the website of Dr. Charles Lindemann, Oakland University (<http://www2.oakland.edu/biology/lindemann/cf.htm>) with permission.

They consist of a microtubule-based skeleton called axoneme, which extends from a specialized centriole called a basal body, and is surrounded by a bilayered ciliary membrane (Figure 1(c)). The axonemal microtubules are oriented with their growing (plus end) toward the cilium tip (Johnson and Rosenbaum, 1992) and comprise nine peripheral microtubule doublets, each containing an A-tubule fused with a B-tubule, which form a ring-shaped structure connected by nexin links. In addition to the nine peripheral microtubules, motile cilia usually contain a central pair (CP) of microtubules (9+2 configuration; Figures 1(b), 1(c), and 2(b)), although there

are exceptions to this 9+2 pattern. For example, motile cilia with 9+0 or 9+4 configuration were observed on the embryonic node of mouse and rabbit embryos (Nonaka *et al.*, 1998; Feistel and Blum, 2006) whereas non-motile cilia of the fish inner ear contain a CP (9+2 configuration) (Flock and Duvall, 1965). Thus, the presence of a CP does not always correlate with motility. Motility rather correlates with the presence of outer (ODAs) and inner (IDAs) dynein arms, which interconnect one outer axonemal microtubule doublet with the adjacent outer doublet (Figure 2(b)) and move along the doublets in a microtubule minus-end direction (Satir and

Christensen, 2006). ODAs have a homogenous structure and composition whereas IDAs are composed of at least seven different dynein isoforms (King and Kamiya, 2009). Another characteristic feature of motile cilia is the attachment of RS to the A-tubule of each microtubule outer doublet (Figure 2(b)). RS are T-shaped multiprotein complexes with a head pointing toward the CP microtubules (Yang and Smith, 2009). Appropriate attachment of the base of the RS to the A-tubule requires RS protein 3 (RSP3), an A-kinase anchoring protein (AKAP) involved in recruiting specific kinases to RS (Diener *et al.*, 1993; Wirschell *et al.*, 2008; Gaillard *et al.*, 2001, 2006).

The Basis of Motility

A common, but still incomplete, model describing the mechanical processes generating ciliary motion is based on directed movement of dynein motors from the ciliary tip to microtubule minus ends at the cilium base generating a molecular force which results in sliding of microtubule doublets along the ciliary axoneme (Sale and Satir, 1977; Fox and Sale, 1987). In this 'sliding model' sliding of microtubule doublets is temporally and spatially regulated. Controlled activation and inhibition of dynein activity across the axonemal axis and along the length leads to alternate bending of microtubule doublets in a oscillatory manner that confers cilium motility (Satir and Christensen, 2006; Lindemann and Lesich, 2010; Wirschell *et al.*, 2011). There is evidence in *Chlamydomonas* and mammalian cells that activity of ODAs regulates ciliary beat frequency whereas activity of IDAs is responsible for the waveform of the stroke (i.e., size and shape of the ciliary bend) (Brokaw, 1994; Brokaw and Kamiya, 1987; Chilvers *et al.*, 2003). Both ciliary beat frequency as well as the shape of the stroke are at least in part regulated by kinases and phosphatases that modify dynein activity and are physically attached to microtubule doublets, RS and CP usually through anchoring proteins such as AKAP240 and RSP3 (Salathe, 2007; Wirschell *et al.*, 2011). The protein network regulating dynein activity includes cAMP-dependent protein kinase (PKA), cGMP-dependent protein kinase (PKG), casein kinase 1 (CK1), protein phosphatase (PP)1 and 2A, and the dynein regulatory complex (DRC) (Figure 2(b); Wirschell *et al.*, 2009, 2011).

The processes underlying ciliary motion are very complex and still not completely understood. However, a common view is that spatial regulation of dynein activity is accomplished by direct interactions of RS heads with CP microtubules and further signal transmission via the DRC to dynein arms (Lindemann and Lesich, 2010; Wirschell *et al.*, 2011). Indeed, RS-CP interaction is regulating microtubule doublet sliding and can inhibit IDA activity (Smith and Sale, 1992; Witman *et al.*, 1978; Nakano *et al.*, 2003). The CP-associated protein Hydin is involved in the regulation of dynein activity possibly through its interaction with RS heads (Lechtreck and Witman, 2007). Notably, Hydin mutations cause Hydrocephalus in mice and humans due to motility defects in cilia in the brain causing insufficient transport of cerebrospinal fluid (Lechtreck *et al.*, 2008; Olbrich *et al.*, 2012). Moreover, regulation of dynein activity requires association of the DRC with ODAs and IDAs (Wirschell *et al.*, 2009, 2011). High-resolution cryoelectron tomography revealed that nexin is part of the

DRC, referred to as nexin-DRC, which interconnects outer microtubule doublets with ODAs and IDAs acting as a regulatory hub to control activity of dynein motors (Nicastro *et al.*, 2006; Heuser *et al.*, 2009). Indeed, nexin-DRC associates with inner dynein I1 and activity of I1 subunit IC138 is critical for dynein activity and microtubule sliding (Bower *et al.*, 2009; Heuser *et al.*, 2012). Importantly, mutations in the gene encoding the nexin-DRC subunit DRC1 cause primary ciliary dyskinesia (PCD) in humans (Wirschell *et al.*, 2013).

Ciliary Tip Structures

The distal tip of motile cilia typically features a central cap and filaments that link the microtubule ends to the ciliary membrane (Figure 1(c)). The organization and structure of these tip complexes vary among species and cell types, and their composition is not well understood (Fisch and Dupuis-Williams, 2011). Some of the proteins known to localize specifically to the ciliary tip include the EF-hand Ca^{2+} binding domain-containing protein Sentan (Kubo *et al.*, 2008), microtubule plus end-tracking EB proteins (Pedersen *et al.*, 2003; Schröder *et al.*, 2011), and centrosomal protein (CEP) 104 (Satish Tammana *et al.*, 2013). Both EB proteins and CEP104 are required for ciliogenesis (Satish Tammana *et al.*, 2013; Schröder *et al.*, 2011; Jiang *et al.*, 2012), and lack of CEP104 causes defects in the flagellar tip structures in *Chlamydomonas* (Satish Tammana *et al.*, 2013).

Axoneme Structure of Non-Motile Cilia

Like motile cilia the axonemes of non-motile cilia typically consist of nine peripheral microtubule pairs arranged in a ring (Figure 1), but lack the structures associated with motility, for example, dynein arms and radial spokes. In addition, the central pair of microtubules is usually lacking (i.e., 9+0 configuration) although there are exceptions to this rule (Gluenz *et al.*, 2010). The axoneme tip seems to lack distinct capping structures, ending in an electron-dense area close to the ciliary membrane (Fisch and Dupuis-Williams, 2011). Nevertheless, similar to motile cilia, non-motile cilia contain specific tip proteins such as EB proteins and CEP104 (Hao *et al.*, 2011; Larsen *et al.*, 2012; Satish Tammana *et al.*, 2013).

The Ciliary Membrane

The ciliary membrane is a bilayer lipid membrane that is continuous with but chemically distinct from the plasma membrane, and contains specific lipids and membrane proteins (Nachury *et al.*, 2010; Rohatgi and Snell, 2010). It is derived from vesicles that are transported from the Golgi or recycling endosomes toward the basal body, where vesicles fuse and become incorporated into the growing ciliary membrane (Sorokin, 1962, 1968; Bouck, 1971; Papermaster *et al.*, 1985; Hsiao *et al.*, 2012; Nachury *et al.*, 2010; Sung and Leroux, 2013). Lateral transport of proteins from the plasma membrane to the ciliary membrane has also been reported (Hunnicuttt *et al.*, 1990; Milenkovic *et al.*, 2009). Structural and functional barriers at the transition zone (TZ) are thought to prevent diffusion and mixing of ciliary and plasma membrane

proteins and lipids, ensuring maintenance of a unique ciliary membrane composition (see Section 'Passage of Proteins through the Transition Zone'). The ciliary membrane is dynamic, and incorporation of lipids and membrane proteins into the ciliary membrane is counterbalanced by endocytosis at the ciliary pocket (Rattner *et al.*, 2010; Kaplan *et al.*, 2012; Benmerah, 2013) and budding of ectosomes directly from the ciliary membrane (Wood *et al.*, 2013; Wang *et al.*, 2014).

The ciliary membrane is enriched in sterols (Montesano, 1979; Iomini *et al.*, 2006; Lehtreck *et al.*, 2013) and its lipid composition is regulated by enzymes such as phospholipase D, which is actively transported into and out of cilia (Lehtreck *et al.*, 2013). Appropriate regulation of ciliary membrane lipid content is important for membrane protein localization and signaling (Rohatgi *et al.*, 2007). Many different membrane receptors and ion channels are known to be enriched in cilia (Christensen *et al.*, 2007; Nachury *et al.*, 2010; Rohatgi and Snell, 2010) and this is fundamental for the ability of cilia to regulate various signaling pathways (see Section 'Signaling Pathways Coordinated by Cilia'). Some prominent examples of ciliary membrane proteins include the polycystic kidney disease gene products polycystin-1 and -2 (Barr and Sternberg, 1999; Pazour *et al.*, 2002; Yoder *et al.*, 2002), various receptor tyrosine kinases (RTKs) (Christensen *et al.*, 2012), and the transmembrane proteins Patched 1 and Smoothened that coordinate Hedgehog (Hh) signaling (Rohatgi *et al.*, 2007).

The Transition Zone

The TZ is the region between the basal body and cilium proper, which contains structural and functional barriers that regulate the selective passage of molecules into and out of the ciliary compartment. The most obvious structural part of the TZ comprises the ciliary necklace (Figures 1(b) and 1(c)). The necklace lies above the transition fibers and, for 9 + 2 cilia, it ends at the so-called basal plate where the axonemal CP microtubules begin (Czarnecki and Shah, 2012; Szymanska and Johnson, 2012). The necklace is composed of multiple rows of the so-called 'Y-links' each consisting of a 'Champagne-glass'-like structure with a pore stem attached to the doublet microtubules and a pore cup attached to the ciliary membrane via strings of intramembrane proteins (Figures 1(b) and 1(c); Gilula and Satir, 1972). Formation of these 'Champagne-glass'-like structures requires the protein CEP290 (Craig *et al.*, 2010). Mutations in the gene encoding CEP290 or other TZ components are responsible for a large number of ciliopathies, including nephronophthisis (NPHP) and Meckel-Gruber Syndrome (MKS) (Szymanska and Johnson, 2012; Czarnecki and Shah, 2012). In some cell types, a ring-like structure formed by the oligomeric GTPase, Septin, decorates the membrane between the basal body and the transition fibers, i.e., the junction between the ciliary and periciliary membranes (Hu *et al.*, 2010; Fliegauf *et al.*, 2013). The periciliary membrane, which is often embedded in the ciliary pocket where the plasma membrane transitions into the ciliary membrane (Benmerah, 2013; Figure 1(c)), has a specialized lipid composition that is thought to be critical for regulating transport of lipids and proteins into and out of the cilium (Vieira *et al.*, 2006; Cevik *et al.*, 2013).

Ciliogenesis

All types of eukaryotic cilia and flagella are assembled onto a basal body. In general, many aspects of ciliogenesis are highly conserved among different organisms and cell types, but specific differences exist (Dawe *et al.*, 2007; Ginger *et al.*, 2008). Most axonemes assemble within the confines of the ciliary membrane (compartmentalized ciliogenesis), but the axonemes of, for example, *Drosophila* and *Plasmodium* gamete flagella are assembled in the cytosol (Tokuyasu, 1975; Sinden *et al.*, 2010). In mammalian cells with primary cilia the basal body is derived from the mother centriole of the centrosome (Paintrand *et al.*, 1992; Kobayashi and Dynlacht, 2011), whereas in multi-ciliated cells basal body formation depends on production of many new centrioles (Dawe *et al.*, 2007). In some cell types (e.g., multi-ciliated epithelial cells) the centrioles migrate to the cell surface and dock at the plasma membrane prior to ciliogenesis, whereas in other cell types (e.g., fibroblasts and smooth muscle cells) ciliogenesis is initiated within the cell body by attachment of a vesicle to the distal end of the mother centriole (Sorokin, 1962, 1968; Molla-Herman *et al.*, 2010; Dawe *et al.*, 2007). In such cells, the basal body often remains localized deep in the cell body even after cilium formation is complete, and the proximal part of the mature cilium lies within a membrane invagination known as the ciliary pocket or ciliary pit (Figure 1(c); Benmerah, 2013).

Docking of centrioles to the membrane or a vesicle depends on the centriole distal appendages (Anderson, 1972; Schmidt *et al.*, 2012; Tateishi *et al.*, 2013; Joo *et al.*, 2013; Sillibourne *et al.*, 2013), which correspond to the basal body transition fibers (Paintrand *et al.*, 1992; Tateishi *et al.*, 2013). Depletion or loss of centriolar distal appendage proteins such as ODF2, CEP164, CEP83 [CCDC41], CEP89 [CCDC123], SCLT1, and FBF1 impairs ciliogenesis (Ishikawa *et al.*, 2005; Graser *et al.*, 2007; Schmidt *et al.*, 2012; Tanos *et al.*, 2013; Tateishi *et al.*, 2013; Joo *et al.*, 2013; Sillibourne *et al.*, 2013). Once centrioles have docked to the plasma membrane or vesicle, the axoneme emerges from the distal end of the basal body (Sorokin, 1962, 1968). The axoneme keeps elongating from its distal end (Johnson and Rosenbaum, 1992) in an intraflagellar transport (IFT)-dependent manner and the growing axoneme is enclosed by membrane along its entire length (Figure 3). This membrane, which becomes the ciliary membrane, is formed by continuous docking and fusion of vesicles at the ciliary base (Figure 3; Sorokin, 1962, 1968; Hill and Outka, 1974; Papermaster *et al.*, 1985; Hsiao *et al.*, 2012; Nachury *et al.*, 2010). Trafficking of such vesicles to the ciliary base is complex and involves different sorting routes and numerous GTPases, adapters and coat proteins (Nachury *et al.*, 2010; Barenz *et al.*, 2011; Hsiao *et al.*, 2012; Deretic, 2013; Sung and Leroux, 2013). Some prominent examples include the GTPases Rab8 and Rab11 (Moritz *et al.*, 2001; Nachury *et al.*, 2007; Knodler *et al.*, 2010) as well as IFT20 (Follit *et al.*, 2006). Proteins destined for the ciliary compartment may not only be associated with such vesicles (Hill and Outka, 1974; Wood and Rosenbaum, 2014), but can also be transported in the form of non-membraneous granules called centriolar satellites (Kubo and Tsukita, 2003; Barenz *et al.*, 2011). Transport of ciliary proteins and membrane from the cytosol toward the basal body is thought to depend on cytoplasmic dynein 1,

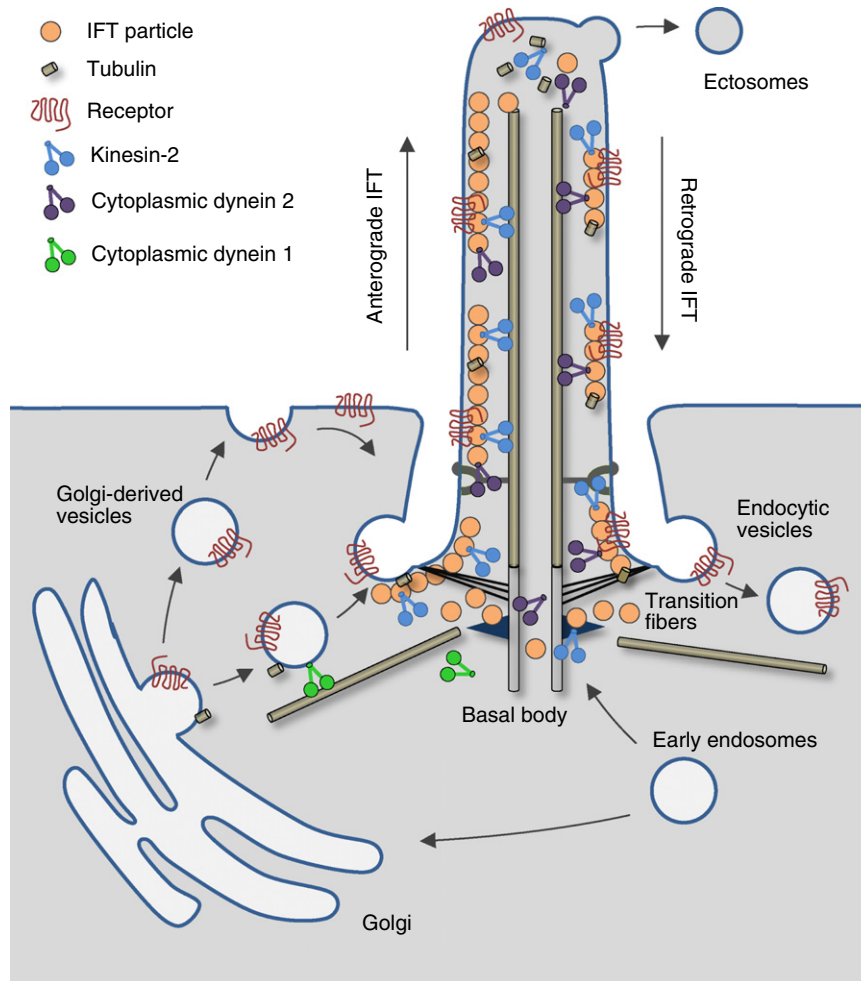


Figure 3 Diagram of IFT and vesicle transport pathways involved in cilium assembly, disassembly, and signaling. See text for details.

which travels along cytoplasmic microtubules toward the basal body (Kubo and Tsukita, 2003; Kong *et al.*, 2013), as well as the presence of specific ciliary targeting sequences in the proteins destined for the cilium (Pazour and Bloodgood, 2008; Berbari *et al.*, 2008). Some ciliary proteins, such as components of the RS, may be preassembled in large complexes in the cytoplasm prior to transport to the cilium (Qin *et al.*, 2004).

Intraflagellar Transport (IFT)

Compartmentalized ciliogenesis relies on IFT for delivery of ciliary building blocks, produced in the cell body, to their site of incorporation in the cilium (Pedersen and Rosenbaum, 2008; Ishikawa and Marshall, 2011). Consequently, mutations in genes coding for IFT components lead to defects in ciliary assembly and function and cause disease (see Section ‘Diseases caused by defects in cilia’). IFT is the process by which granular particles (IFT particles or trains) move along axonemal outer doublet microtubules, beneath the ciliary membrane, from the basal body to the flagellar tip and back (Figure 3). IFT was first discovered in 1993 by enhanced digital interference contrast

microscopy of immobilized flagella of *Chlamydomonas*, where the particles were seen to move in the anterograde (base to tip) and retrograde (tip to base) directions at approximately $2 \mu\text{s}^{-1}$ and $3.5 \mu\text{s}^{-1}$, respectively (Kozminski *et al.*, 1993). Subsequent studies using fluorescently tagged IFT particle or motor subunits revealed similar bidirectional movement in cilia from other organisms, such as neuronal sensory cilia in *C. elegans* (Orozco *et al.*, 1999), primary cilia of cultured mammalian cells (Follit *et al.*, 2006; Tran *et al.*, 2008), and motile flagella of *Trypanosoma brucei* (Absalon *et al.*, 2008). IFT is thought to occur in all eukaryotic cilia and flagella that display compartmentalized ciliogenesis, but the specific motors employed may vary depending on organism and cell type (Pedersen and Rosenbaum, 2008; Silverman and Leroux, 2009; Ishikawa and Marshall, 2011).

The anterograde IFT motor: Kinesin-2

Studies using a temperature sensitive (ts) *Chlamydomonas* mutant in the *FLA10* gene (Adams *et al.*, 1982; Huang *et al.*, 1977), which encodes one of the two motor subunits of heterotrimeric kinesin-2, also known as kinesin-II (Walther *et al.*, 1994; Table 1), revealed that this motor is responsible for anterograde IFT (Kozminski *et al.*, 1995). Subsequent studies

IFT particle polypeptides and cargo

IFT particles were first isolated from *Chlamydomonas* flagella using sucrose density gradient centrifugation (Piperno and Mead, 1997; Cole *et al.*, 1998). The IFT particles consist of two sub-complexes termed IFT-A and IFT-B (Cole *et al.*, 1998), which comprise at least 6 and 16 different polypeptides, respectively (Table 1) (Ishikawa and Marshall, 2011; Taschner *et al.*, 2012; van Dam *et al.*, 2013). Mutations in IFT-B genes generally lead to very short or absent cilia, whereas mutations in IFT-A genes typically lead to bulbous cilia with excessive amount of IFT material (Pedersen and Rosenbaum, 2008; Ishikawa and Marshall, 2011). Thus, it is commonly accepted that IFT-B proteins participate in anterograde IFT whereas IFT-A proteins are involved in retrograde IFT. However, some IFT-A proteins also participate in anterograde IFT, and some IFT-B proteins are important in retrograde IFT (Pedersen *et al.*, 2005; Qin *et al.*, 2005; Tsao and Gorovsky, 2008; Mukhopadhyay *et al.*, 2010; Behal *et al.*, 2012; Liem *et al.*, 2012; Wei *et al.*, 2012; Keady *et al.*, 2012).

Many IFT polypeptides harbor protein–protein interaction motifs such as tetratricopeptide repeats (TPRs), WD40 repeats, and coiled coils (Cole, 2003). Phylogenetic analyses indicated that several IFT-A and IFT-B proteins share common ancestry with the COPI vesicle coat complex (Jekely and Arendt, 2006; van Dam *et al.*, 2013), two IFT-B proteins (IFT22 and IFT27) are similar to Rab-like GTPases (Schafer *et al.*, 2006; Qin *et al.*, 2007), while four of the IFT-B proteins (IFT54, IFT57, IFT81, and CLUAP1) contain calponin homology domains (Taschner *et al.*, 2012; Bhogaraju *et al.*, 2013a; Schou *et al.*, 2013). The overall structure of the IFT particle trains was analyzed by electron tomography (Pigino *et al.*, 2009), whereas biochemical, bioinformatics, and crystallography studies have provided information regarding the organization of IFT polypeptides within the IFT-A and IFT-B complexes (Lucker *et al.*, 2005; Taschner *et al.*, 2011; Behal *et al.*, 2012) as well as the three dimensional structure of individual IFT particle polypeptides and sub-complexes (Bhogaraju *et al.*, 2011, 2013a). These studies are beginning to shed light on the molecular mechanisms by which IFT particle polypeptides interact with specific ciliary cargo proteins to promote axoneme assembly (Bhogaraju *et al.*, 2013b). For example, a recent study showed that IFT81 and IFT74 of the IFT-B complex form a dimer that directly binds tubulin (Bhogaraju *et al.*, 2013a). In addition to tubulin, several other axonemal proteins were found to interact with IFT particle polypeptides, including RSPs and ODA subunits (Qin *et al.*, 2004; Ahmed *et al.*, 2008; Gao *et al.*, 2010). A recent study in *Chlamydomonas* using single-particle imaging showed that IFT is required for transport of DRC4, a component of the DRC (see Section “The Basis of Motility”), from the cell body toward the tip of flagella where the majority of IFT-bound DRC4 is released (Wren *et al.*, 2013).

In addition to transport of structural components of the axoneme, IFT may also mediate transport of ciliary membrane and signaling proteins. For example, IFT-A proteins interact with the Tubby-like protein 3 (TULP3), which in turn interacts with distinct G protein coupled receptors to mediate their ciliary localization in an IFT-dependent manner (Mukhopadhyay *et al.*, 2010). IFT proteins also interact with Bardet Biedl Syndrome proteins (Blacque *et al.*, 2004; Lechtreck *et al.*, 2009; Wei *et al.*, 2012; Domire *et al.*, 2010), which form a protein complex

called the BBSome that is involved in ciliary membrane protein trafficking (Nachury *et al.*, 2007; Loktev *et al.*, 2008; Lechtreck *et al.*, 2009; Wei *et al.*, 2012). Similar to many IFT-A and IFT-B protein, the BBSome is thought to have evolved from a coat-omer-like progenitor (van Dam *et al.*, 2013).

Passage of Proteins through the Transition Zone

The selective passage of molecules into and out of the ciliary compartment is thought to be regulated by structural and functional barriers at the TZ. The presence of such barriers or ‘ciliary pore complexes’ was first proposed by Rosenbaum and Witman (2002), due to the observation that IFT particle proteins accumulate around the transition fibers in *Chlamydomonas* (Deane *et al.*, 2001). Several recent studies support the presence of such barriers at the TZ (Craigie *et al.*, 2010; Hu *et al.*, 2010; Kee *et al.*, 2012; Breslow *et al.*, 2013; Lin *et al.*, 2013). The barriers include the Septin ring complex that is thought to limit mobility of ciliary membrane proteins (Hu *et al.*, 2010) as well as the structures of the ciliary necklace (Craigie *et al.*, 2010). In the case of non-membrane proteins, a diffusion barrier with a pore size of approximately 8–9 nm in diameter that allows diffusion of molecules below about 50–100 kD into the cilium was reported (Kee *et al.*, 2012). Larger molecules such as the DRC are actively transported across the TZ by IFT (Wren *et al.*, 2013) and probably depend on additional factors for ciliary entry. Such factors may include proteins orthologous to those operating in nuclear trafficking, including importin and components in the RanGTP cycle (Fan *et al.*, 2007; Dishinger *et al.*, 2010; Hurd *et al.*, 2011; Maiuri *et al.*, 2013). Further, proteins similar to structural components of nuclear pore complexes, nucleoporins (Strambio-De-Castillia *et al.*, 2010), may function at the ciliary base region to regulate the import of ciliary proteins (Kee *et al.*, 2012).

Once proteins have passed the ciliary necklace they may become allocated to specific areas of the ciliary compartment. Transmembrane proteins, such as Smoothed and Serotonin receptors lose contact to IFT trains and are positioned along the ciliary membrane predominantly by diffusion (Ye *et al.*, 2013). In other cases, protein localization may be regulated by yet another set of barriers that define discrete intraciliary compartments. Prominent examples include the ‘Inversin/NPHP2 compartment’ that runs for 1–2 μm above the necklace region in primary cilia of vertebrate cells (Shiba *et al.*, 2009) and the ‘Arl-13/ARL13b compartment’ that corresponds to the middle segment of sensilla in *C. elegans* (Cevik *et al.*, 2010, 2013). These compartments likely play a critical role in the regulation of processes associated with ciliary formation, maintenance, and sensory function. Finally, proteins may be positioned uniquely to the ciliary tip, such as Gli transcription factors, which when activated are transported out of the cilium for nuclear localization (Haycraft *et al.*, 2005; see Section ‘Signaling Pathways Coordinated by Cilia’).

Cilium Disassembly and Length Maintenance

Cilia are dynamic organelles that can assemble, disassemble, or regulate their length in response to cellular and environmental cues. Cilia can disassemble either by severing of the

axoneme at its base or by depolymerization of axonemal microtubules from the tip (Quarmany, 2009; Seeley and Nachury, 2010; Ishikawa and Marshall, 2011). For example, in *Chlamydomonas*, the microtubule doublets between the basal body and flagellar TZ are severed prior to mitosis to liberate the centrioles for mitotic spindle pole formation (Parker *et al.*, 2010). Severing of the axoneme likely involves the microtubule severing enzyme katanin (Rasi *et al.*, 2009) whereas disassembly of the microtubule axonemes from the tip likely involves depolymerizing kinesins such as kinesin-8 or kinesin-13 family members (Blaineau *et al.*, 2007; Piao *et al.*, 2009; Dawson *et al.*, 2007; Niwa *et al.*, 2012). In addition, several kinases such as never in mitosis gene A (NIMA)-related kinases (NEKs), mitogen activated protein kinases (MAPKs), and Aurora A-like kinases were shown to regulate cilium length and disassembly (Berman *et al.*, 2003; Pan *et al.*, 2004; Quarmany and Mahjoub, 2005; Pugacheva *et al.*, 2007), but the exact mechanisms involved are not well understood (Plotnikova *et al.*, 2009; Ishikawa and Marshall, 2011). In steady state-length cilia, axonemal disassembly at the tip is balanced by IFT-dependent anterograde transport of tubulin (Marshall and Rosenbaum, 2001; Kozminski *et al.*, 1995; Bhogaraju *et al.*, 2013a). In addition, removal of ciliary membrane by endocytosis at the ciliary pocket (Rattner *et al.*, 2010; Kaplan *et al.*, 2012; Benmerah, 2013) or via budding of ectosomes directly from the ciliary membrane (Wood *et al.*, 2013; Wang *et al.*, 2014) may counterbalance the delivery of membrane vesicles at the ciliary base (see Section ‘The Ciliary Membrane’). These processes are also important for regulating signaling.

Signaling Pathways Coordinated by Cilia

All types of cilia likely function as sensory organelles, which register and convey extracellular signals to control a variety of cellular processes and behavioral responses (Christensen *et al.*, 2007). Specialized sensory cilia are found at dendritic endings on invertebrate neurons, neurons in the vertebrate olfactory epithelium of the nasal cavity, and on photoreceptor cells of vertebrates and certain invertebrates. These include sensilla in the nematode sensory organs that sense chemical, mechanical, osmotic, and thermal cues to control behavioral responses such as feeding, mating, and avoidance responses (Inglis *et al.*, 2007; Bae and Barr, 2008). In some cases, the sensory capacity is maintained by transient receptor potential (TRP) ion channels as well as G-Protein Coupled Receptors (GPCRs), which comprise a large family of 7-pass transmembrane (7TM) receptors that convey signals into a cellular response through the activation of G proteins on the cytoplasmic side of the cilium. A prominent example is that of Opsins, which in the outer segment of rod photoreceptors in the retina convert light signals to nerve impulses via cyclic nucleotide gated ion channels (Pearring *et al.*, 2013).

Sensory Capacity of Primary Cilia

Primary cilia play a unique role in coordinating signaling systems to control cellular processes during development and

in tissue homeostasis (Satir *et al.*, 2010; Goetz and Anderson, 2010). Many types of receptors and other signaling components are enriched in cilia and depend on cilium localization for appropriate function. Examples of signaling pathways coordinated by primary cilia include Hh (Huangfu *et al.*, 2003; Corbit *et al.*, 2005), Wntless/int (Wnt) (Wallingford and Mitchell, 2011; Lienkamp *et al.*, 2012), Notch (Ezratty *et al.*, 2011), RTK (Schneider *et al.*, 2005; Christensen *et al.*, 2012) and Transforming Growth Factor (TGF) β signaling (Clement *et al.*, 2013), as well as signaling through extracellular matrix (ECM) receptors (McGlashan *et al.*, 2006; Seeger-Nukpezah and Golemis, 2012), GPCRs (Berbari *et al.*, 2008; Green and Myktyyn, 2010; Loktev and Jackson, 2013), Purinergic receptors (Masyuk *et al.*, 2008; Wann *et al.*, 2012; Praetorius and Leipziger, 2013) and ion channels (Jones and Nauli, 2012; Figure 4). In this way, primary cilia function as chemo-, mechano-, and osmosensory organelles that regulate cell cycle control as well as polarization, migration, differentiation, size, and homeostasis of cells during organogenesis in embryonic and fetal stages of development as well as in tissue and organ function in the adult (see references above).

Primary cilia and Hedgehog signaling

The Hh signaling pathway is one of the key morphogenic regulators in tissue patterning and organogenesis during animal development (Ingham *et al.*, 2011). Hh signaling was originally discovered in *Drosophila*, where a mutation in this pathway produced a short and ‘spiked’ phenotype of the cuticle of the larvae (Nusslein-Volhard and Wieschaus, 1980). In mammals, Hh signaling is critically linked to the primary cilium, which functions as a cellular switch for the activation and deactivation of the Hh signaling cascade (Rohatgi *et al.*, 2007). In the absence of Hh ligands, the 12TM receptor, Patched 1 localizes to the ciliary membrane, where it blocks the activation of Gli transcription factors in the cilium. Hh signaling is initially turned on by binding of Hh ligands to ciliary Patched 1, which then leaves the ciliary compartment followed by entrance of the 7TM protein, Smoothened, which promotes formation of active versions Gli transcription factors in the cilium for targeted gene transcription in the nucleus. In other ciliary pathways, such as RTK and GPCR signaling, ligands may bind and activate receptors, which directly impinge on secondary messenger components in signal transduction. Finally, receptor-mediated signal transduction may rely on clathrin-dependent endocytosis at the ciliary pocket for activation of Smad transcription factors, such as in TGF β signaling (Clement *et al.*, 2013).

In summary, primary cilia orchestrate the regulation of multiple signaling pathways that control cellular processes during development and in tissue homeostasis. In this scenario, diverse pathways could act in concert in a spatiotemporal manner to regulate specified cellular responses in different cell types and tissues.

Diseases Caused by Defects in Cilia

Owing to the importance of cilia in motility and signal transduction, as well as their wide tissue distribution, mutations in genes required for cilium formation or function

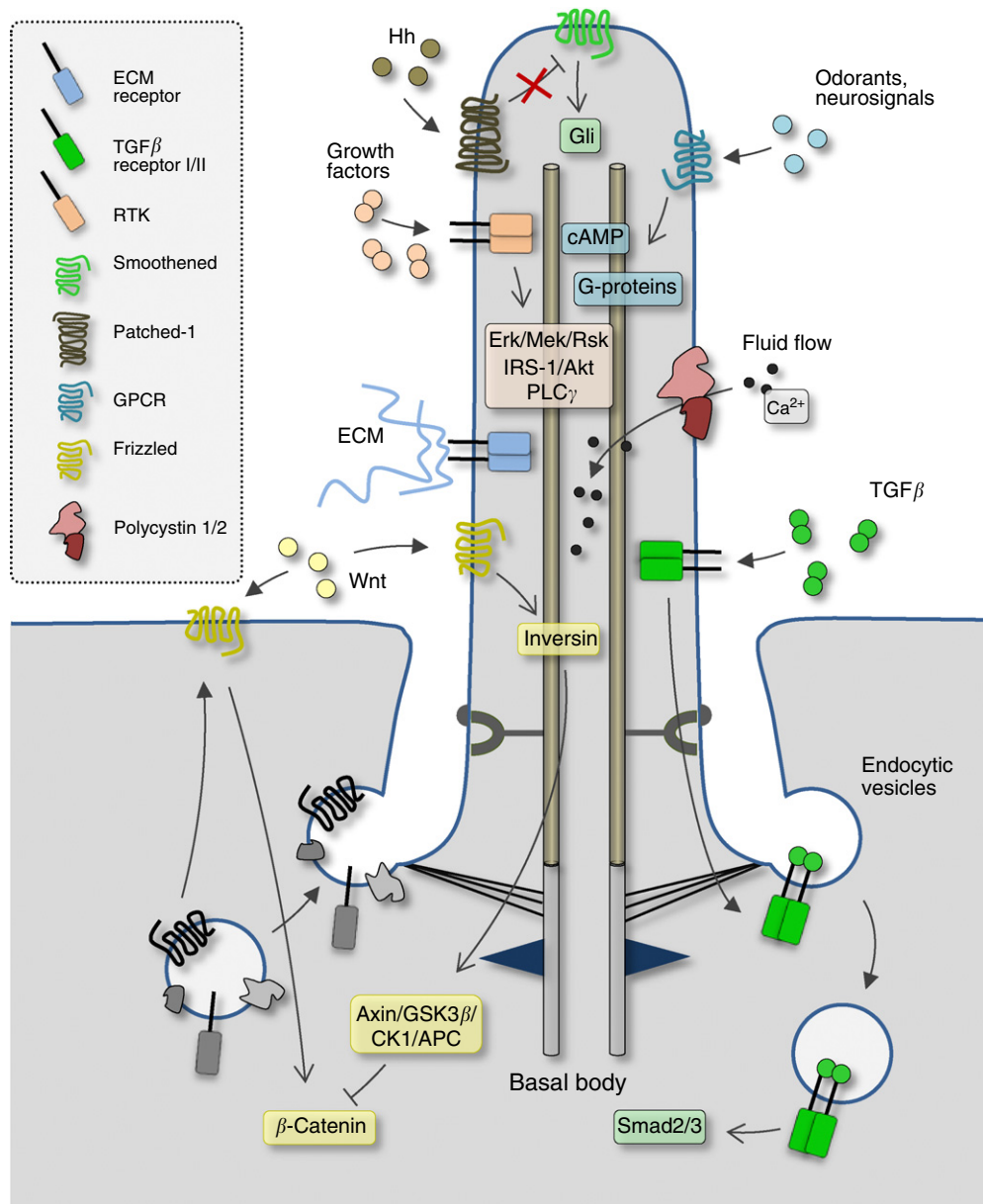


Figure 4 Simplified overview of signaling pathways that are coordinated by primary cilia. Various receptors, ion channels, and other signaling proteins are localized to cilia allowing coordination and integration of diverse signals in time and space to generate an appropriate cellular response. See text for details.

cause a broad range of human diseases called ciliopathies. Most prominently, defects in primary and sensory cilia inevitably lead to a series of developmental disorders and diseases, such as neurodevelopmental disorders, congenital heart disease, obesity, and cystic kidney diseases (Badano *et al.*, 2006; Fliegauf *et al.*, 2007; Hildebrandt *et al.*, 2011; Kenny and Beales, 2014; Tables 2 and 3). Ciliopathies are often rare and pleiotropic with many different tissues affected, and disease phenotypes are often complex with overlapping conditions making precise diagnosis challenging. Furthermore, complex genetic inheritance patterns and genetic modifiers of disease phenotypes have been reported (Lee and Gleeson, 2011;

Beales and Kenny, 2014). A few illustrative examples of ciliopathies will be described briefly below.

Diseases Caused by Defective Motile Cilia

It has been known for many years that defects in motile cilia can cause diseases in humans (Afzelius, 2004). The first human patients with defects in motile cilia were found among males in Denmark and Sweden in the seventies. These patients suffered from a variety of ailments, including infertility, chronic airway infections or situs inversus. Careful examination of immotile sperm cells or bronchial cells from these patients revealed lack

Table 2 List of known ciliopathies^a

Abbreviation	Full name	OMIM
Syndromic		
ACLS	Acrocallosal syndrome	#200990
ADPKD	Autosomal dominant polycystic kidney disease	N/A
ALS	Alström syndrome	N/A
ARPKD	Autosomal recessive polycystic kidney disease	N/A
BBS	Bardet–Biedl syndrome	#209900
CED	Cranial-ectodermal dysplasia	#218330
COACH	COACH syndrome	#216360
CORS	Cerebellar-ocular-renal syndrome	#608091
DKA	Dekaban-arima syndrome	#243910
EVC	Ellis-van Creveld syndrome	#225500
HLS	Hydrolethrus syndrome	N/A
JATD	Jeune asphyxiating thoracic dystrophy	#208500
JBTS	Joubert syndrome	N/A
MKS	Meckel–Gruber syndrome	N/A
MORM	Mental retardation, obesity, congenital retinal dystrophy and micropenis in males syndrome	#610156
MSS	Mainzer–Saldino syndrome	#266920
N/A	Usher syndrome	N/A
N/A	Cogan oculomotor apraxia	#257550
N/A	Weyers acrofacial dysostosis	#193539
NPHP	Nephronophthisis	N/A
OFDs	Oral-facial-digital syndromes (1 to 11)	#311200, #252100, #258850, #258860, #174300, #174300, #277170, #300484, #258865, #165590, #612913
PCD	Primary ciliary dyskinesia	N/A
SLS	Senior–Løken syndrome	N/A
SRPS	Short-rib-polydactyly syndromes	#263510 to #263530, #269860, #614091
Non-syndromic		
CD	Cone dystrophy	N/A
CRD	Cone rod dystrophy	N/A
LCA	Leber congenital amaurosis	N/A
MD	Macular dystrophy	N/A
RP	Retinitis pigmentosa	#268000

^aKenny, T.D., Beales, P.L., 2014. Ciliopathies: A Reference for Clinicians. Oxford: Oxford University Press.

of axonemal dynein arms (Afzelius *et al.*, 1975; Pedersen and Rebbe, 1975; Afzelius, 1976). The disorder, which was named ‘immotile-cilia syndrome’ (Eliasson *et al.*, 1977) showed phenotypical overlap with the previously described Kartagener’s syndrome (Kartagener and Stucki, 1962) and the more recently described syndrome, primary ciliary dyskinesia (PCD; OMIM 244400). PCD is a rare, genetically heterogeneous syndrome with symptoms including infections of the airways, laterality defects, and infertility (Bush *et al.*, 2007; Zariwala *et al.*, 2007). PCD is the most common disorder affecting motile cilia and is usually caused by mutations in ODA genes, although mutations in other genes affecting motility, for example, the gene encoding the nexin-DRC subunit DRC1, have also been reported (Wirschell *et al.*, 2013; Zariwala *et al.*, 2007; Hogg, 2014).

Diseases Caused by Defective Non-Motile Cilia

The first evidence that defects in primary cilia cause disease was provided by studies in *Chlamydomonas* and mouse. In 1994 an insertional mutation was identified in the mouse gene *Tg737* causing a disease phenotype similar to that found in human patients with autosomal recessive polycystic kidney

disease (ARPKD; OMIM 263200) (Moyer *et al.*, 1994). In 2000, Pazour *et al.* demonstrated that the *Chlamydomonas* homolog of Tg737, IFT88, which is part of the IFT-B complex (see Section ‘Intraflagellar Transport’), is required for cilium assembly and that cilia in polycystic kidneys of Tg737 mutant mice are shorter than normal (Pazour *et al.*, 2000). Concomitant studies from the Yoder lab confirmed that Tg737/IFT88 is a ciliary protein required for ciliogenesis (Haycraft *et al.*, 2001; Taulman *et al.*, 2001).

Barr and Sternberg had previously shown that products of the *C. elegans* homologs of the mammalian autosomal dominant polycystic kidney disease (ADPKD) genes, PKD1 and PKD2, are localized to non-motile cilia on male cephalic neurons (Barr and Sternberg, 1999). These studies indicated for the first time that correct assembly and maintenance of cilia are required for appropriate functioning of the kidney and that malfunctions in this organelle can lead to kidney cysts similar to those found in patients with ARPKD. These studies also spurred an intense interest in primary cilia, and lead a virtual boom in research in this organelle.

ARPKD is mainly caused by mutations in the *PKHD1* gene, whose protein product Fibrocystin localizes to cilia of

Table 3 List of symptoms commonly associated with ciliopathies

Symptom
Anosmia
Brain anomalies
Cancer
Cognitive disorders
Diabetes
Hearing loss
Heart defects
Hydrocephalus
Hypertension
Infertility
Kidney cysts
Lateral defects
Liver cysts
Obesity
Pancreas cysts
Respiratory disorders
Retinal degeneration
Skeletal malformations

kidney-derived cells (Ward *et al.*, 2002, 2003). ADPKD is an ARPKD-related disease in which the form of kidneys and cysts differ significantly from those of ARPKD patients; a well-established criterium for diagnosis (Bergmann, 2014). Mutations in *PKD1* and *PKD2* account for the majority of ADPKD cases (Rossetti *et al.*, 2007). The population prevalence varies between 1:400 and 1:1000 affected (Sandford, 2014).

Another prominent ciliopathy is BBS (OMIM 209900), a pleiotropic and genetically heterogeneous disorder with at least 16 causative genes found to date (Forsythe and Beales, 2013). Many of the encoded BBS proteins are involved in selective transport of membrane proteins to and from the ciliary compartment (Nachury *et al.*, 2007; Loktev *et al.*, 2008; Lechtreck *et al.*, 2009; Wei *et al.*, 2012). Mutations in *BBS* genes lead to defective transport of several membrane proteins to cilia affecting signaling and homeostasis of many different tissues (Zaghloul and Katsanis, 2009). Patients with BBS suffer from primary features including retinal degeneration, polydactyly, mental retardation, truncal obesity, and genital and renal anomalies as well as secondary features including speech disorders, developmental delay, hearing loss, and lateral defects (Tobin and Beales, 2007).

See also: Cell Communication: Cellular Outcomes: Regulation of Cell Polarity. Cell Communication: Growth Factor Mediated Cell Signaling: G Protein-Coupled Receptors; Hedgehog Signaling in Development and Disease; Receptor Tyrosine Kinases and Their Ligands. Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell Polarity; Centrioles and the Centrosome. Cytoskeleton and Motors: Cytoskeletal Components: Dyneins; Kinesin Superfamily Proteins (KIFs) as a Fundamental Component of Life: Intracellular Transport and Beyond; Microtubules and Microtubule-Associated Proteins (MAPs). Dynamic Integration: Dynamics: Dynamics of Microtubule Self-Assembly; Theory of Cargo and Membrane Trafficking

References

- Absalon, S., Blisnick, T., Kohl, L., *et al.*, 2008. Intraflagellar transport and functional analysis of genes required for flagellum formation in trypanosomes. *Molecular Biology of the Cell* 19, 929–944.
- Adams, G.M.W., Huang, B., Luck, D.J.L., 1982. Temperature-sensitive, assembly-defective flagella mutants of *Chlamydomonas reinhardtii*. *Genetics* 100, 579–586.
- Atzelius, B.A., 1976. A human syndrome caused by immotile cilia. *Science* 193, 317–319.
- Atzelius, B.A., 2004. Cilia-related diseases. *Journal of Pathology* 204, 470–477.
- Atzelius, B.A., Eliasson, R., Johnsen, O., Lindholmer, C., 1975. Lack of dynein arms in immotile human spermatozoa. *Journal of Cell Biology* 66, 225–232.
- Ahmed, N.T., Gao, C., Luckner, B.F., Cole, D.G., Mitchell, D.R., 2008. ODA16 aids axonemal outer row dynein assembly through an interaction with the intraflagellar transport machinery. *Journal of Cell Biology* 183, 313–322.
- Anderson, R.G., 1972. The three-dimensional structure of the basal body from the rhesus monkey oviduct. *Journal of Cell Biology* 54, 246–265.
- Arnaiz, O., Malinowska, A., Klotz, C., *et al.*, 2009. Cildb: A knowledgebase for centrosomes and cilia. Database: The Journal of Biological Databases and Curation 2009, bap022.
- Badano, J.L., Mitsuma, N., Beales, P.L., Katsanis, N., 2006. The ciliopathies: An emerging class of human genetic disorders. *Annual Review of Genomics and Human Genetics* 7, 125–148.
- Bae, Y.K., Barr, M.M., 2008. Sensory roles of neuronal cilia: Cilia development, morphogenesis, and function in *C. elegans*. *Frontiers in Bioscience* 13, 5959–5974.
- Barenz, F., Mayilo, D., Gruss, O.J., 2011. Centriolar satellites: Busy orbits around the centrosome. *European Journal of Cell Biology* 90, 983–989.
- Barr, M.M., Sternberg, P.W., 1999. A polycystic kidney-disease gene homologue required for male mating behaviour in *C. elegans*. *Nature* 401, 386–389.
- Beales, P.L., Kenny, T.D., 2014. Towards the diagnosis of a ciliopathy. In: Kenny, T.D., Beales, P.L. (Eds.), *Ciliopathies: A Reference for Clinicians*, first ed. Oxford: Oxford University Press, pp. 1–282.
- Behal, R.H., Miller, M.S., Qin, H., *et al.*, 2012. Subunit interactions and organization of the *Chlamydomonas reinhardtii* intraflagellar transport complex A proteins. *Journal of Biological Chemistry* 287, 11689–11703.
- Benmerah, A., 2013. The ciliary pocket. *Current Opinion in Cell Biology* 25, 78–84.
- Berbari, N.F., Johnson, A.D., Lewis, J.S., Askwith, C.C., Mykytyk, K., 2008. Identification of ciliary localization sequences within the third intracellular loop of G protein-coupled receptors. *Molecular Biology of the Cell* 19, 1540–1547.
- Bergmann, C., 2014. Autosomal recessive polycystic kidney disease. In: Kenny, T.D., Beales, P.L. (Eds.), *Ciliopathies: A Reference for Clinicians*, first ed. Oxford: Oxford University Press, pp. 194–217.
- Berman, S.A., Wilson, N.F., Haas, N.A., Lefebvre, P.A., 2003. A novel MAP kinase regulates flagellar length in *Chlamydomonas*. *Current Biology* 13, 1145–1149.
- Bhogaraju, S., Cajanek, L., Fort, C., *et al.*, 2013a. Molecular basis of tubulin transport within the cilium by IFT74 and IFT81. *Science* 341, 1009–1012.
- Bhogaraju, S., Engel, B.D., Lorentzen, E., 2013b. Intraflagellar transport complex structure and cargo interactions. *Cilia* 2, 10.
- Bhogaraju, S., Taschner, M., Morawetz, M., Basquin, C., Lorentzen, E., 2011. Crystal structure of the intraflagellar transport complex 25/27. *EMBO Journal* 30, 1907–1918.
- Blacque, O.E., Reardon, M.J., Li, C., *et al.*, 2004. Loss of *C. elegans* BBS-7 and BBS-8 protein function results in cilia defects and compromised intraflagellar transport. *Genes & Development* 18, 1630–1642.
- Blaineau, C., Tessier, M., Dubessay, P., *et al.*, 2007. A novel microtubule-depolymerizing kinesin involved in length control of a eukaryotic flagellum. *Current Biology* 17, 778–782.
- Bouck, G.B., 1971. The structure, origin, isolation, and composition of the tubular mastigonemes of the *Ochromonas* flagellum. *Journal of Cell Biology* 50, 362–384.
- Bower, R., Vanderwaal, K., O'Toole, E., *et al.*, 2009. IC138 defines a subdomain at the base of the I1 dynein that regulates microtubule sliding and flagellar motility. *Molecular Biology of the Cell* 20, 3055–3063.
- Breslow, D.K., Koslover, E.F., Seydel, F., Spakowitz, A.J., Nachury, M.V., 2013. An *in vitro* assay for entry into cilia reveals unique properties of the soluble diffusion barrier. *Journal of Cell Biology* 203, 129–147.
- Brokaw, C.J., 1994. Control of flagellar bending: A new agenda based on dynein diversity. *Cell Motility and the Cytoskeleton* 28, 199–204.
- Brokaw, C.J., Kamiya, R., 1987. Bending patterns of *Chlamydomonas* flagella: IV. Mutants with defects in inner and outer dynein arms indicate differences in dynein arm function. *Cell Motility and the Cytoskeleton* 8, 68–75.
- Brown, J.M., Marsala, C., Kosoy, R., Gaertig, J., 1999. Kinesin-II is preferentially targeted to assembling cilia and is required for ciliogenesis and

- normal cytokinesis in Tetrahymena. *Molecular Biology of the Cell* 10, 3081–3096.
- Bush, A., Chodhari, R., Collins, N., *et al.*, 2007. Primary ciliary dyskinesia: Current state of the art. *Archives of Disease in Childhood* 92, 1136–1140.
- Cevik, S., Hori, Y., Kaplan, O.I., *et al.*, 2010. Joubert syndrome ARL13b functions at ciliary membranes and stabilizes protein transport in *Caenorhabditis elegans*. *Journal of Cell Biology* 188, 953–969.
- Cevik, S., Sanders, A.A., Van Wijk, E., *et al.*, 2013. Active transport and diffusion barriers restrict Joubert Syndrome-associated ARL13B/ARL-13 to an Inv-like ciliary membrane subdomain. *PLoS Genetics* 9, e1003977.
- Chilvers, M.A., Rutman, A., O'Callaghan, C., 2003. Ciliary beat pattern is associated with specific ultrastructural defects in primary ciliary dyskinesia. *Journal of Allergy and Clinical Immunology* 112, 518–524.
- Christensen, S.T., Clement, C.A., Satir, P., Pedersen, L.B., 2012. Primary cilia and coordination of receptor tyrosine kinase (RTK) signalling. *Journal of Pathology* 226, 172–184.
- Christensen, S.T., Pedersen, L.B., Schneider, L., Satir, P., 2007. Sensory cilia and integration of signal transduction in human health and disease. *Traffic* 8, 97–109.
- Clement, C.A., Ajbro, K.D., Henriques De Jesus, M.P.R., *et al.*, 2013. Regulation of TGF β signaling by endocytosis at the pocket region of the primary cilium. *Cell Reports* 3, 1806–1814.
- Cole, D.G., 2003. The intraflagellar transport machinery of *Chlamydomonas reinhardtii*. *Traffic* 4, 1–8.
- Cole, D.G., Diener, D.R., Himelblau, A.L., *et al.*, 1998. *Chlamydomonas* kinesin-II-dependent intraflagellar transport (IFT): IFT particles contain proteins required for ciliary assembly in *Caenorhabditis elegans* sensory neurons. *Journal of Cell Biology* 141, 993–1008.
- Corbit, K.C., Aanstad, P., Singla, V., *et al.*, 2005. Vertebrate Smoothed functions at the primary cilium. *Nature* 437, 1018–1021.
- Craige, B., Tsao, C.C., Diener, D.R., *et al.*, 2010. CEP290 tethers flagellar transition zone microtubules to the membrane and regulates flagellar protein content. *Journal of Cell Biology* 190, 927–940.
- Czarniecki, P.G., Shah, J.V., 2012. The ciliary transition zone: From morphology and molecules to medicine. *Trends in Cell Biology* 22, 201–210.
- Dagoneau, N., Goulet, M., Genevieve, D., *et al.*, 2009. DYNC2H1 mutations cause asphyxiating thoracic dystrophy and short rib-polydactyly syndrome, type III. *American Journal of Human Genetics* 84, 706–711.
- van Dam, T.J., Townsend, M.J., Turk, M., *et al.*, 2013. Evolution of modular intraflagellar transport from a coatomer-like progenitor. *Proceedings of the National Academy of Sciences of the United States of America* 110, 6943–6948.
- Dawe, H.R., Farr, H., Gull, K., 2007. Centriole/basal body morphogenesis and migration during ciliogenesis in animal cells. *Journal of Cell Science* 120, 7–15.
- Dawson, S.C., Sagolla, M.S., Mancuso, J.J., *et al.*, 2007. Kinesin-13 regulates flagellar, interphase, and mitotic microtubule dynamics in *Giardia intestinalis*. *Eukaryotic Cell* 6, 2354–2364.
- Deane, J.A., Cole, D.G., Seeley, E.S., Diener, D.R., Rosenbaum, J.L., 2001. Localization of intraflagellar transport protein IFT52 identifies basal body transitional fibers as the docking site for IFT particles. *Current Biology* 11, 1586–1590.
- Demonchy, R., Blisnick, T., Deprez, C., *et al.*, 2009. Kinesin 9 family members perform separate functions in the trypanosome flagellum. *Journal of Cell Biology* 187, 615–622.
- Deretic, D., 2013. Crosstalk of Arf and Rab GTPases en route to cilia. *Small GTPases* 4, 70–77.
- Diener, D.R., Ang, L.H., Rosenbaum, J.L., 1993. Assembly of flagellar radial spoke proteins in *Chlamydomonas*: Identification of the axoneme binding domain of radial spoke protein 3. *Journal of Cell Biology* 123, 183–190.
- Dishinger, J.F., Kee, H.L., Jenkins, P.M., *et al.*, 2010. Ciliary entry of the kinesin-2 motor KIF17 is regulated by importin- β 2 and RanGTP. *Nature Cell Biology* 12, 703–710.
- Domire, J.S., Green, J.A., Lee, K.G., *et al.*, 2010. Dopamine receptor 1 localizes to neuronal cilia in a dynamic process that requires the Bardet-Biedl syndrome proteins. *Cellular and Molecular Life Sciences: CMLS* 68 (17), 2951–2961.
- Eliasson, R., Mossberg, B., Camner, P., Afzelius, B.A., 1977. The immotile-cilia syndrome. A congenital ciliary abnormality as an etiologic factor in chronic airway infections and male sterility. *New England Journal of Medicine* 297, 1–6.
- Endoh-Yamagami, S., Evangelista, M., Wilson, D., *et al.*, 2009. The mammalian Cos2 homolog Kif7 plays an essential role in modulating Hh signal transduction during development. *Current Biology: CB* 19, 1320–1326.
- Ezratty, E.J., Stokes, N., Chai, S., *et al.*, 2011. A role for primary cilia in Notch signaling and epidermal differentiation during skin development. *Cell* 145, 1129–1141.
- Fan, S., Fogg, V., Wang, Q., *et al.*, 2007. A novel Crumbs3 isoform regulates cell division and ciliogenesis via importin β interactions. *Journal of Cell Biology* 178, 387–398.
- Feistel, K., Blum, M., 2006. Three types of cilia including a novel 9 + 4 axoneme on the notochordal plate of the rabbit embryo. *Developmental Dynamics* 235, 3348–3358.
- Fisch, C., Dupuis-Williams, P., 2011. Ultrastructure of cilia and flagella – back to the future! *Biology of the Cell/Under the Auspices of the European Cell Biology Organization* 103, 249–270.
- Fliegauf, M., Benzing, T., Omran, H., 2007. When cilia go bad: Cilia defects and ciliopathies. *Nature Reviews Molecular Cell Biology* 8, 880–893.
- Fliegauf, M., Kahle, A., Hafner, K., Zieger, B., 2013. Distinct localization of septin proteins to ciliary sub-compartments in airway epithelial cells. *Biology Chemistry* 395 (2), 151–156.
- Flock, A., Duvall 3rd, A.J., 1965. The ultrastructure of the Kinocilium of the sensory cells in the inner ear and lateral line organs. *Journal of Cell Biology* 25, 1–8.
- Follit, J.A., Tuft, R.A., Fogarty, K.E., Pazour, G.J., 2006. The intraflagellar transport protein IFT20 is associated with the Golgi complex and is required for cilia assembly. *Molecular Biology of the Cell* 17, 3781–3792.
- Forsythe, E., Beales, P.L., 2013. Bardet-Biedl syndrome. *European Journal of Human Genetics* 21, 8–13.
- Fox, L.A., Sale, W.S., 1987. Direction of force generated by the inner row of dynein arms on flagellar microtubules. *Journal of Cell Biology* 105, 1781–1787.
- Gaillard, A.R., Diener, D.R., Rosenbaum, J.L., Sale, W.S., 2001. Flagellar radial spoke protein 3 is an A-kinase anchoring protein (AKAP). *Journal of Cell Biology* 153, 443–448.
- Gaillard, A.R., Fox, L.A., Rhea, J.M., Craige, B., Sale, W.S., 2006. Disruption of the A-kinase anchoring domain in flagellar radial spoke protein 3 results in unregulated axonemal cAMP-dependent protein kinase activity and abnormal flagellar motility. *Molecular Biology of the Cell* 17, 2626–2635.
- Gao, C., Wang, G., Amack, J.D., Mitchell, D.R., 2010. Oda16/Wdr69 is essential for axonemal dynein assembly and ciliary motility during zebrafish embryogenesis. *Developmental Dynamics* 239, 2190–2197.
- Gibbons, B.H., Asai, D.J., Tang, W.J., Hays, T.S., Gibbons, I.R., 1994. Phylogeny and expression of axonemal and cytoplasmic dynein genes in sea urchins. *Molecular Biology of the Cell* 5, 57–70.
- Gilula, N.B., Satir, P., 1972. The ciliary necklace. A ciliary membrane specialization. *Journal of Cell Biology* 53, 494–509.
- Ginger, M.L., Portman, N., McKean, P.G., 2008. Swimming with protists: Perception, motility and flagellum assembly. *Nature Reviews. Microbiology* 6, 838–850.
- Gluenz, E., Hoog, J.L., Smith, A.E., *et al.*, 2010. Beyond 9 + 0: Noncanonical axoneme structures characterize sensory cilia from protists to humans. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 24, 3117–3121.
- Goetz, S.C., Anderson, K.V., 2010. The primary cilium: A signalling centre during vertebrate development. *Nature Reviews. Genetics* 11, 331–344.
- Graser, S., Stierhof, Y.D., Lavoie, S.B., *et al.*, 2007. Cep164, a novel centriole appendage protein required for primary cilium formation. *Journal of Cell Biology* 179, 321–330.
- Green, J.A., Mykytyn, K., 2010. Neuronal ciliary signaling in homeostasis and disease. *Cellular and Molecular Life Sciences: CMLS* 67, 3287–3297.
- Hao, L., Thein, M., Brust-Mascher, I., *et al.*, 2011. Intraflagellar transport delivers tubulin isotypes to sensory cilium middle and distal segments. *Nature Cell Biology* 13, 790–798.
- Haycraft, C.J., Banizs, B., Aydin-Son, Y., *et al.*, 2005. Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function. *PLoS Genetics* 1, e53.
- Haycraft, C.J., Swoboda, P., Taulman, P.D., Thomas, J.H., Yoder, B.K., 2001. The *C. elegans* homolog of the murine cystic kidney disease gene *Tg737* functions in a ciliogenic pathway and is disrupted in *osm-5* mutant worms. *Development* 128, 1493–1505.
- He, M., Subramanian, R., Bangs, F., *et al.*, 2014. The kinesin-4 protein KIF7 regulates mammalian hedgehog signaling by organizing the cilia tip compartment. *Nature Cell Biology* 16, 663–672.
- Heuser, T., Barber, C.F., Lin, J., *et al.*, 2012. Cryoelectron tomography reveals doublet-specific structures and unique interactions in the 11 dynein. *Proceedings of the National Academy of Sciences of the United States of America* 109, E2067–E2076.

- Heuser, T., Raytchev, M., Krell, J., Porter, M.E., Nicastro, D., 2009. The dynein regulatory complex is the nexin link and a major regulatory node in cilia and flagella. *Journal of Cell Biology* 187, 921–933.
- Hildebrandt, F., Benzing, T., Katsanis, N., 2011. Ciliopathies. *New England Journal of Medicine* 364, 1533–1543.
- Hill, F.G., Outka, D.E., 1974. The structure and origin of mastigonemes in *Ochromonas minute* and *Monas* sp. *Journal of Protozoology* 21, 299–312.
- Hirokawa, N., Tanaka, Y., Okada, Y., Takeda, S., 2006. Nodal flow and the generation of left-right asymmetry. *Cell* 125, 33–45.
- Hogg, C., 2014. Primary ciliary dyskinesia. In: Kenny, T.D., Beales, P.L. (Eds.), *Ciliopathies: A Reference for Clinicians*, first ed. Oxford, UK: Oxford University Press, pp. 219–235.
- Hou, Y., Pazour, G.J., Witman, G.B., 2004. A dynein light intermediate chain, D1bLIC, is required for retrograde intraflagellar transport. *Molecular Biology of the Cell* 15, 4382–4394.
- Hsiao, Y.C., Tuz, K., Ferland, R.J., 2012. Trafficking in and to the primary cilium. *Cilia* 1, 4.
- Hu, Q., Milenkovic, L., Jin, H., et al., 2010. A septin diffusion barrier at the base of the primary cilium maintains ciliary membrane protein distribution. *Science* 329, 436–439.
- Huang, B., Rifkin, M.R., Luck, D.J.L., Kozler, V., 1977. Temperature-sensitive mutations affecting flagellar assembly and function in *Chlamydomonas reinhardtii*. *Journal of Cell Biology* 72, 67–85.
- Huangfu, D., Anderson, K.V., 2005. Cilia and Hedgehog responsiveness in the mouse. *Proceedings of the National Academy of Sciences of the United States of America* 102, 11325–11330.
- Huangfu, D., Liu, A., Rakeman, A.S., et al., 2003. Hedgehog signalling in the mouse requires intraflagellar transport proteins. *Nature* 426, 83–87.
- Huber, C., Wu, S., Kim, A.S., et al., 2013. WDR34 mutations that cause short-rib polydactyly syndrome type III/severe asphyxiating thoracic dysplasia reveal a role for the NF-kappaB pathway in cilia. *American Journal of Human Genetics* 93, 926–931.
- Hunnicutt, G.R., Kosfisz, M.G., Snell, W.J., 1990. Cell body and flagellar agglutinins in *Chlamydomonas reinhardtii*. The cell body plasma membrane is a reservoir for agglutinins whose migration to the flagella is regulated by a functional barrier. *Journal of Cell Biology* 111, 1605–1616.
- Hurd, T.W., Fan, S., Margolis, B.L., 2011. Localization of retinitis pigmentosa 2 to cilia is regulated by Importin beta2. *Journal of Cell Science* 124, 718–726.
- Ibanez-Tallon, I., Heintz, N., Omran, H., 2003. To beat or not to beat: Roles of cilia in development and disease. *Human Molecular Genetics* 12, R27–R35.
- Ingham, P.W., Nakano, Y., Seger, C., 2011. Mechanisms and functions of Hedgehog signalling across the metazoa. *Nature Reviews Genetics* 12, 393–406.
- Inglis, P.N., Ou, G., Leroux, M.R., Scholey, J.M., 2007. The sensory cilia of *Caenorhabditis elegans*. *WormBook: The Online Review of C. elegans Biology* 8, 1–22.
- Insinna, C., Besharse, J.C., 2008. Intraflagellar transport and the sensory outer segment of vertebrate photoreceptors. *Developmental Dynamics* 237, 1982–1992.
- Insinna, C., Pathak, N., Perkins, B., Drummond, I., Besharse, J.C., 2008. The homodimeric kinesin, Kif17, is essential for vertebrate photoreceptor sensory outer segment development. *Developmental Biology* 316, 160–170.
- Iomini, C., Li, L., Mo, W., Dutcher, S.K., Piperno, G., 2006. Two flagellar genes, AGG2 and AGG3, mediate orientation to light in *Chlamydomonas*. *Current Biology* 16, 1147–1153.
- Ishikawa, H., Kubo, A., Tsukita, S., 2005. Odf2-deficient mother centrioles lack distal/subdistal appendages and the ability to generate primary cilia. *Nature Cell Biology* 7, 517–524.
- Ishikawa, H., Marshall, W.F., 2011. Ciliogenesis: Building the cell's antenna. *Nature Reviews Molecular Cell Biology* 12, 222–234.
- Jekely, G., Arendt, D., 2006. Evolution of intraflagellar transport from coated vesicles and autogenous origin of the eukaryotic cilium. *Bioessays* 28, 191–198.
- Jenkins, P.M., Hurd, T.W., Zhang, L., et al., 2006. Ciliary targeting of olfactory CNG channels requires the CNGB1b subunit and the kinesin-2 motor protein, KIF17. *Current Biology* 16, 1211–1216.
- Jensen, C.G., Poole, C.A., McGlashan, S.R., et al., 2004. Ultrastructural, tomographic and confocal imaging of the chondrocyte primary cilium in situ. *Cell Biology International* 28, 101–110.
- Jiang, K., Toedt, G., Montenegro Gouveia, S., et al., 2012. A proteome-wide screen for mammalian SxIP motif-containing microtubule plus-end tracking proteins. *Current Biology* 22 (19), 1800–1807.
- Johnson, K.A., Rosenbaum, J.L., 1992. Polarity of flagellar assembly in *Chlamydomonas*. *Journal of Cell Biology* 119, 1605–1611.
- Jones, T.J., Nauli, S.M., 2012. Mechanosensory calcium signaling. *Advances in Experimental Medicine and Biology* 740, 1001–1015.
- Joo, K., Kim, C.G., Lee, M.S., et al., 2013. CCDC41 is required for ciliary vesicle docking to the mother centriole. *Proceedings of the National Academy of Sciences of the United States of America* 110, 5987–5992.
- Kaplan, O.I., Doroquez, D.B., Cevik, S., et al., 2012. Endocytosis genes facilitate protein and membrane transport in *C. elegans* sensory cilia. *Current Biology* 22, 451–460.
- Kartagener, M., Stucki, P., 1962. Bronchiectasis with situs inversus. *Archives of Pediatrics* 79, 193–207.
- Keady, B.T., Samtani, R., Tobita, K., et al., 2012. IFT25 links the signal-dependent movement of Hedgehog components to intraflagellar transport. *Developmental Cell* 22, 940–951.
- Kee, H.L., Dishinger, J.F., Blasius, T.L., et al., 2012. A size-exclusion permeability barrier and nucleoporins characterize a ciliary pore complex that regulates transport into cilia. *Nature Cell Biology* 14, 431–437.
- Kenny, T.D., Beales, P.L., 2014. *Ciliopathies: A Reference for Clinicians*. Oxford: Oxford University Press.
- King, S.M., Kamiya, R., 2009. Axonemal dyneins: Assembly, structure and force generation. In: Harris, E.H. (Ed.), *The Chlamydomonas Sourcebook*, second ed. London: Academic Press, pp. 131–208.
- Kiprilov, E.N., Awan, A., Desprat, R., et al., 2008. Human embryonic stem cells in culture possess primary cilia with hedgehog signaling machinery. *Journal of Cell Biology* 180, 897–904.
- Knodler, A., Feng, S., Zhang, J., et al., 2010. Coordination of Rab8 and Rab11 in primary ciliogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 107, 6346–6351.
- Kobayashi, T., Dynlacht, B.D., 2011. Regulating the transition from centriole to basal body. *Journal of Cell Biology* 193, 435–444.
- Kong, S., Du, X., Peng, C., et al., 2013. Dlic1 deficiency impairs ciliogenesis of photoreceptors by destabilizing dynein. *Cell Research* 23, 835–850.
- Kozminski, K.G., Beech, P.L., Rosenbaum, J.L., 1995. The *Chlamydomonas* kinesin-like protein FLA10 is involved in motility associated with the flagellar membrane. *Journal of Cell Biology* 131, 1517–1527.
- Kozminski, K.G., Johnson, K.A., Forscher, P., Rosenbaum, J.L., 1993. A motility in the eukaryotic flagellum unrelated to flagellar beating. *Proceedings of the National Academy of Sciences of the United States of America* 90, 5519–5523.
- Krock, B.L., Mills-Henry, I., Perkins, B.D., 2009. Retrograde intraflagellar transport by cytoplasmic dynein-2 is required for outer segment extension in vertebrate photoreceptors but not arrestin translocation. *Investigative Ophthalmology & Visual Science* 50, 5463–5471.
- Kubo, A., Tsukita, S., 2003. Non-membranous granular organelle consisting of PCM-1: Subcellular distribution and cell-cycle-dependent assembly/disassembly. *Journal of Cell Science* 116, 919–928.
- Kubo, A., Yuba-Kubo, A., Tsukita, S., Amagai, M., 2008. Sentan: A novel specific component of the apical structure of vertebrate motile cilia. *Molecular Biology of the Cell* 19, 5338–5346.
- Larsen, J., Grigoriev, I., Akhmanova, A., Pedersen, L.B., 2012. Analysis of microtubule plus end-tracking proteins in cilia. *Methods in Enzymology* 524, 105–122.
- Lechtreck, K.F., Brown, J.M., Sampaio, J.L., et al., 2013. Cycling of the signaling protein phospholipase D through cilia requires the BBSome only for the export phase. *Journal of Cell Biology* 201, 249–261.
- Lechtreck, K.F., Delmotte, P., Robinson, M.L., Sanderson, M.J., Witman, G.B., 2008. Mutations in Hydin impair ciliary motility in mice. *Journal of Cell Biology* 180, 633–643.
- Lechtreck, K.F., Johnson, E.C., Sakai, T., et al., 2009. The *Chlamydomonas reinhardtii* BBSome is an IFT cargo required for export of specific signaling proteins from flagella. *Journal of Cell Biology* 187, 1117–1132.
- Lechtreck, K.F., Witman, G.B., 2007. *Chlamydomonas reinhardtii* hydin is a central pair protein required for flagellar motility. *Journal of Cell Biology* 176, 473–482.
- Lee, J.E., Gleeson, J.G., 2011. A systems-biology approach to understanding the ciliopathy disorders. *Genome Medicine* 3, 59.
- Liem Jr., K.F., Ashe, A., He, M., et al., 2012. The IFT-A complex regulates Shh signaling through cilia structure and membrane protein trafficking. *Journal of Cell Biology* 197, 789–800.
- Liem Jr., K.F., He, M., Ocbina, P.J., Anderson, K.V., 2009. Mouse Kif7/Costal2 is a cilia-associated protein that regulates Sonic hedgehog signaling. *Proceedings of the National Academy of Sciences of the United States of America* 106, 13377–13382.
- Lienkamp, S., Ganner, A., Walz, G., 2012. Inversin, Wnt signaling and primary cilia. *Differentiation* 83, S49–S55.
- Lin, Y.C., Niewiadomski, P., Lin, B., et al., 2013. Chemically inducible diffusion trap at cilia reveals molecular sieve-like barrier. *Nature Chemical Biology* 9, 437–442.
- Lindemann, C.B., Lesich, K.A., 2010. Flagellar and ciliary beating: The proven and the possible. *Journal of Cell Science* 123, 519–528.

- Loktev, A.V., Jackson, P.K., 2013. Neuropeptide Y family receptors traffic via the Bardet-Biedl syndrome pathway to signal in neuronal primary cilia. *Cell Reports* 5, 1316–1329.
- Loktev, A.V., Zhang, Q., Beck, J.S., *et al.*, 2008. A BBSome subunit links ciliogenesis, microtubule stability, and acetylation. *Developmental Cell* 15, 854–865.
- Lucker, B.F., Behal, R.H., Qin, H., *et al.*, 2005. Characterization of the intraflagellar transport complex B core: Direct interaction of the IFT81 and IFT74/72 subunits. *Journal of Biological Chemistry* 280, 27688–27696.
- Maiuri, T., Woloshansky, T., Xia, J., Truant, R., 2013. The huntingtin N17 domain is a multifunctional CRM1 and Ran-dependent nuclear and ciliary export signal. *Human Molecular Genetics* 22, 1383–1394.
- Marshall, W.F., Rosenbaum, J.L., 2001. Intraflagellar transport balances continuous turnover of outer doublet microtubules: Implications for flagellar length control. *Journal of Cell Biology* 155, 405–414.
- Marszalek, J.R., Ruiz-Lozano, P., Roberts, E., Chien, K.R., Goldstein, L.S., 1999. Situs inversus and embryonic ciliary morphogenesis defects in mouse mutants lacking the KIF3A subunit of kinesin-II. *Proceedings of the National Academy of Sciences of the United States of America* 96, 5043–5048.
- Masyuk, A.I., Gradilone, S.A., Banales, J.M., *et al.*, 2008. Cholangiocyte primary cilia are chemosensory organelles that detect biliary nucleotides via P2Y12 purinergic receptors. *American Journal of Physiology. Gastrointestinal and Liver Physiology* 295, G725–G734.
- May, S.R., Ashique, A.M., Karlen, M., *et al.*, 2005. Loss of the retrograde motor for IFT disrupts localization of Smo to cilia and prevents the expression of both activator and repressor functions of Gli. *Developmental Biology* 287, 378–389.
- McGlashan, S.R., Jensen, C.G., Poole, C.A., 2006. Localization of extracellular matrix receptors on the chondrocyte primary cilium. *Journal of Histochemistry and Cytochemistry: Official Journal of the Histochemistry Society* 54, 1005–1014.
- Merrill, A.E., Merriman, B., Farrington-Rock, C., *et al.*, 2009. Ciliary abnormalities due to defects in the retrograde transport protein DYNC2H1 in short-rib polydactyly syndrome. *American Journal of Human Genetics* 84, 542–549.
- Milenkovic, L., Scott, M.P., Rohatgi, R., 2009. Lateral transport of Smoothened from the plasma membrane to the membrane of the cilium. *Journal of Cell Biology* 187, 365–374.
- Molla-Herman, A., Ghossoub, R., Blisnick, T., *et al.*, 2010. The ciliary pocket: An endocytic membrane domain at the base of primary and motile cilia. *Journal of Cell Science* 123, 1785–1795.
- Montesano, R., 1979. Inhomogeneous distribution of filipin-sterol complexes in the ciliary membrane of rat tracheal epithelium. *American Journal of Anatomy* 156, 139–145.
- Moritz, O.L., Tam, B.M., Hurd, L.L., *et al.*, 2001. Mutant rab8 impairs docking and fusion of rhodopsin-bearing post-Golgi membranes and causes cell death of transgenic *Xenopus* rods. *Molecular Biology of the Cell* 12, 2341–2351.
- Morris, R.L., Scholey, J.M., 1997. Heterotrimeric kinesin-II is required for the assembly of motile 9 + 2 ciliary axonemes on sea urchin embryos. *Journal of Cell Biology* 138, 1009–1022.
- Moyer, J.H., Lee-Tischler, M.J., Kwon, H.Y., *et al.*, 1994. Candidate gene associated with a mutation causing recessive polycystic kidney disease in mice. *Science* 264, 1329–1333.
- Mukhopadhyay, S., Wen, X., Chih, B., *et al.*, 2010. TULP3 bridges the IFT-A complex and membrane phosphoinositides to promote trafficking of G protein-coupled receptors into primary cilia. *Genes & Development* 24, 2180–2193.
- Nachury, M.V., Loktev, A.V., Zhang, Q., *et al.*, 2007. A core complex of BBS proteins cooperates with the GTPase Rab8 to promote ciliary membrane biogenesis. *Cell* 129, 1201–1213.
- Nachury, M.V., Seeley, E.S., Jin, H., 2010. Trafficking to the ciliary membrane: How to get across the periciliary diffusion barrier? *Annual Review of Cell and Developmental Biology* 26, 59–87.
- Nakano, I., Kobayashi, T., Yoshimura, M., Shingyoji, C., 2003. Central-pair-linked regulation of microtubule sliding by calcium in flagellar axonemes. *Journal of Cell Science* 116, 1627–1636.
- Nicastro, D., Schwartz, C., Pierson, J., *et al.*, 2006. The molecular architecture of axonemes revealed by cryoelectron tomography. *Science* 313, 944–948.
- Niwa, S., Nakajima, K., Miki, H., *et al.*, 2012. KIF19A is a microtubule-depolymerizing kinesin for ciliary length control. *Developmental Cell* 23, 1167–1175.
- Nonaka, S., Tanaka, Y., Okada, Y., *et al.*, 1998. Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. *Cell* 95, 829–837.
- Nusslein-Volhard, C., Wieschaus, E., 1980. Mutations affecting segment number and polarity in *Drosophila*. *Nature* 287, 795–801.
- Okada, Y., Takeda, S., Tanaka, Y., Izpisua Belmonte, J.C., Hirokawa, N., 2005. Mechanism of nodal flow: A conserved symmetry breaking event in left-right axis determination. *Cell* 121, 633–644.
- Olbrich, H., Schmidts, M., Werner, C., *et al.*, 2012. Recessive HYDIN mutations cause primary ciliary dyskinesia without randomization of left-right body asymmetry. *American Journal of Human Genetics* 91, 672–684.
- Orozco, J.T., Wedaman, K.P., Signor, D., *et al.*, 1999. Movement of motor and cargo along cilia. *Nature* 398, 674.
- Paintrand, M., Moudjou, M., Delacroix, H., Bornens, M., 1992. Centrosome organization and centriole architecture: Their sensitivity to divalent cations. *Journal of Structural Biology* 108, 107–128.
- Pan, J., Wang, Q., Snell, W.J., 2004. An aurora kinase is essential for flagellar disassembly in *Chlamydomonas*. *Developmental Cell* 6, 445–451.
- Papermaster, D.S., Schneider, B.G., Besharse, J.C., 1985. Vesicular transport of newly synthesized opsin from the Golgi apparatus toward the rod outer segment. Ultrastructural immunocytochemical and autoradiographic evidence in *Xenopus retinas*. *Investigative Ophthalmology & Visual Science* 26, 1386–1404.
- Parker, J.D., Hilton, L.K., Diener, D.R., *et al.*, 2010. Centrioles are freed from cilia by severing prior to mitosis. *Cytoskeleton* 67, 425–430.
- Patel-King, R.S., Gilberti, R.M., Hom, E.F., King, S.M., 2013. WD60/FAP163 is a dynein intermediate chain required for retrograde intraflagellar transport in cilia. *Molecular Biology of the Cell* 24, 2668–2677.
- Pazour, G.J., Bloodgood, R.A., 2008. Targeting proteins to the ciliary membrane. *Current Topics in Developmental Biology* 85, 115–149.
- Pazour, G.J., Dickert, B.L., Vucica, Y., *et al.*, 2000. *Chlamydomonas* IFT88 and its mouse homologue, polycystic kidney disease gene *tg737*, are required for assembly of cilia and flagella. *Journal of Cell Biology* 151, 709–718.
- Pazour, G.J., Dickert, B.L., Witman, G.B., 1999. The DHC1b (DHC2) isoform of cytoplasmic dynein is required for flagellar assembly. *Journal of Cell Biology* 144, 473–481.
- Pazour, G.J., San Agustin, J.T., Folliot, J.A., Rosenbaum, J.L., Witman, G.B., 2002. Polycystin-2 localizes to kidney cilia and the ciliary level is elevated in orpk mice with polycystic kidney disease. *Current Biology* 12, R378–R380.
- Pazour, G.J., Wilkerson, C.G., Witman, G.B., 1998. A dynein light chain is essential for retrograde particle movement in intraflagellar transport (IFT). *Journal of Cell Biology* 141, 979–992.
- Pearring, J.N., Salinas, R.Y., Baker, S.A., Arshavsky, V.Y., 2013. Protein sorting, targeting and trafficking in photoreceptor cells. *Progress in Retinal and Eye Research* 36, 24–51.
- Peden, E.M., Barr, M.M., 2005. The KLP-6 kinesin is required for male mating behaviors and polycystin localization in *Caenorhabditis elegans*. *Current Biology* 15, 394–404.
- Pedersen, H., Rebbe, H., 1975. Absence of arms in the axoneme of immobile human spermatozoa. *Biology of Reproduction* 12, 541–544.
- Pedersen, L.B., Geimer, S., Sloboda, R.D., Rosenbaum, J.L., 2003. The microtubule plus end-tracking protein EB1 is localized to the flagellar tip and basal bodies in *Chlamydomonas reinhardtii*. *Current Biology* 13, 1969–1974.
- Pedersen, L.B., Miller, M.S., Geimer, S., *et al.*, 2005. *Chlamydomonas* IFT172 is encoded by *FLA17*, interacts with CrEB1, and regulates IFT at the flagellar tip. *Current Biology* 15, 262–266.
- Pedersen, L.B., Rosenbaum, J.L., 2008. Intraflagellar transport (IFT) role in ciliary assembly, resorption and signalling. *Current Topics in Developmental Biology* 85, 23–61.
- Perrone, C.A., Tritschler, D., Taulman, P., *et al.*, 2003. A novel dynein light intermediate chain colocalizes with the retrograde motor for intraflagellar transport at sites of axoneme assembly in *Chlamydomonas* and mammalian cells. *Molecular Biology of the Cell* 14, 2041–2056.
- Piao, T., Luo, M., Wang, L., *et al.*, 2009. A microtubule depolymerizing kinesin functions during both flagellar disassembly and flagellar assembly in *Chlamydomonas*. *Proceedings of the National Academy of Sciences of the United States of America* 106, 4713–4718.
- Pigino, G., Geimer, S., Lanzavecchia, S., *et al.*, 2009. Electron-tomographic analysis of intraflagellar transport particle trains in situ. *Journal of Cell Biology* 187, 135–148.
- Piperno, G., Mead, K., 1997. Transport of a novel complex in the cytoplasmic matrix of *Chlamydomonas* flagella. *Proceedings of the National Academy of Sciences of the United States of America* 94, 4457–4462.
- Plotnikova, O.V., Pugacheva, E.N., Golemis, E.A., 2009. Primary cilia and the cell cycle. *Methods in Cell Biology* 94, 137–160.
- Porter, M.E., Bower, R., Knott, J.A., Byrd, P., Dentler, W., 1999. Cytoplasmic dynein heavy chain 1b is required for flagellar assembly in *Chlamydomonas*. *Molecular Biology of the Cell* 10, 693–712.
- Praetorius, H.A., Leipziger, J., 2013. Primary cilium-dependent sensing of urinary flow and paracrine purinergic signaling. *Seminars in Cell and Developmental Biology* 24, 3–10.

- Pugacheva, E.N., Jablonski, S.A., Hartman, T.R., Henske, E.P., Golemis, E.A., 2007. HEF1-dependent Aurora A activation induces disassembly of the primary cilium. *Cell* 129, 1351–1363.
- Qin, H., Burnette, D., Bae, Y.K., *et al.*, 2005. Intraflagellar transport is required for the vectorial movement of TRPV channels in the ciliary membrane. *Current Biology* 15, 1695–1699.
- Qin, H., Diener, D.R., Geimer, S., Cole, D.G., Rosenbaum, J.L., 2004. Intraflagellar transport (IFT) cargo: IFT transports flagellar precursors to the tip and turnover products to the cell body. *Journal of Cell Biology* 164, 255–266.
- Qin, H., Wang, Z., Diener, D., Rosenbaum, J., 2007. Intraflagellar transport protein 27 is a small G protein involved in cell-cycle control. *Current Biology* 17, 193–202.
- Quarby, L.M., 2009. Deflagellation. In: Harris, E.H., Stern, D.B., Witman, G.B. (Eds.), *The Chlamydomonas Sourcebook*, second ed. London: Academic Press, pp. 43–69. (Chapter 3).
- Quarby, L.M., Mahjoub, M.R., 2005. Caught Nek-ing: Cilia and centrioles. *Journal of Cell Science* 118, 5161–5169.
- Rana, A.A., Barbera, J.P., Rodriguez, T.A., *et al.*, 2004. Targeted deletion of the novel cytoplasmic dynein mD2LIC disrupts the embryonic organizer, formation of the body axes and specification of ventral cell fates. *Development* 131, 4999–5007.
- Rasi, M.Q., Parker, J.D., Feldman, J.L., Marshall, W.F., Quarby, L.M., 2009. Katanin knockdown supports a role for microtubule severing in release of basal bodies before mitosis in *Chlamydomonas*. *Molecular Biology of the Cell* 20, 379–388.
- Rattner, J.B., Sciore, P., Ou, Y., Van Der Hoorn, F.A., Lo, I.K., 2010. Primary cilia in fibroblast-like type B synoviocytes lie within a cilium pit: A site of endocytosis. *Histology and Histopathology* 25, 865–875.
- Rohatgi, R., Milenkovic, L., Scott, M.P., 2007. Patched1 regulates hedgehog signaling at the primary cilium. *Science* 317, 372–376.
- Rohatgi, R., Snell, W.J., 2010. The ciliary membrane. *Current Opinion in Cell Biology* 22, 541–546.
- Rompalas, P., Pedersen, L.B., Patel-King, R.S., King, S.M., 2007. *Chlamydomonas* FAP133 is a dynein intermediate chain associated with the retrograde intraflagellar transport motor. *Journal of Cell Science* 120, 3653–3665.
- Rosenbaum, J.L., Witman, G.B., 2002. Intraflagellar transport. *Nature Reviews Molecular Cell Biology* 3, 813–825.
- Rossetti, S., Sengul, M.B., Chapman, A.B., *et al.*, 2007. Comprehensive molecular diagnostics in autosomal dominant polycystic kidney disease. *Journal of the American Society of Nephrology* 18, 2143–2160.
- Salathe, M., 2007. Regulation of mammalian ciliary beating. *Annual Review of Physiology* 69, 401–422.
- Sale, W.S., Satir, P., 1977. Direction of active sliding of microtubules in *Tetrahymena* cilia. *Proceedings of the National Academy of Sciences of the United States of America* 74, 2045–2049.
- Sanchez, E.S., Mardomingo, C.T., 2007. Ultrastructural study of an atypical case of infertility. *ScientificWorldJournal* 7, 818–821.
- Sandford, R., 2014. Autosomal dominant polycystic kidney disease. In: Kenny, T.D., Beales, P.L. (Eds.), *Ciliopathies: A Reference for Clinicians*, first ed. Oxford: Oxford University Press, pp. 174–193.
- Sarpal, R., Todi, S.V., Sivan-Loukianova, E., *et al.*, 2003. *Drosophila* KAP interacts with the kinesin II motor subunit KLP64D to assemble chordotonal sensory cilia, but not sperm tails. *Current Biology* 13, 1687–1696.
- Satir, P., Christensen, S.T., 2006. Overview of structure and function of mammalian cilia. *Annual Review of Physiology* 69, 377–400.
- Satir, P., Pedersen, L.B., Christensen, S.T., 2010. The primary cilium at a glance. *Journal of Cell Science* 123, 499–503.
- Satish Tammana, T.V., Tammana, D., Diener, D.R., Rosenbaum, J., 2013. Centrosomal protein CEP104 (*Chlamydomonas* FAP256) moves to the ciliary tip during ciliary assembly. *Journal of Cell Science* 126, 5018–5029.
- Schafer, J.C., Haycraft, C.J., Thomas, J.H., Yoder, B.K., Swoboda, P., 2003. *XB1* encodes a dynein light intermediate chain required for retrograde intraflagellar transport and cilia assembly in *Caenorhabditis elegans*. *Molecular Biology of the Cell* 14, 2057–2070.
- Schafer, J.C., Winkelbauer, M.E., Williams, C.L., *et al.*, 2006. IFTA-2 is a conserved cilia protein involved in pathways regulating longevity and dauer formation in *Caenorhabditis elegans*. *Journal of Cell Science* 119, 4088–4100.
- Schmidt, K.N., Kuhns, S., Neuner, A., *et al.*, 2012. Cep164 mediates vesicular docking to the mother centriole during early steps of ciliogenesis. *Journal of Cell Biology* 199, 1083–1101.
- Schmidts, M., Arts, H.H., Bongers, E.M., *et al.*, 2013a. Exome sequencing identifies DYNC2H1 mutations as a common cause of asphyxiating thoracic dystrophy (Jeune syndrome) without major polydactyly, renal or retinal involvement. *Journal of Medical Genetics* 50, 309–323.
- Schmidts, M., Vodopivec, J., Christou-Savina, S., *et al.*, 2013b. Mutations in the gene encoding IFT dynein complex component WDR34 cause Jeune asphyxiating thoracic dystrophy. *American Journal of Human Genetics* 93, 932–944.
- Schneider, L., Clement, C.A., Teilmann, S.C., *et al.*, 2005. PDGFRα signaling is regulated through the primary cilium in fibroblasts. *Current Biology* 15, 1861–1866.
- Schou, K.B., Andersen, J.S., Pedersen, L.B., 2013. A divergent calponin homology (NN-CH) domain defines a novel family: Implications for evolution of ciliary IFT complex B proteins. *Bioinformatics* 30 (7), 899–902.
- Schrader, J.M., Larsen, J., Komarova, Y., *et al.*, 2011. EB1 and EB3 promote cilia biogenesis by several centrosome-related mechanisms. *Journal of Cell Science* 124, 2539–2551.
- Seeger-Nukpezah, T., Golemis, E.A., 2012. The extracellular matrix and ciliary signaling. *Current Opinion in Cell Biology* 24, 652–661.
- Seeley, E.S., Nachury, M.V., 2010. The perennial organelle: Assembly and disassembly of the primary cilium. *Journal of Cell Science* 123, 511–518.
- Shiba, D., Yamaoka, Y., Hagiwara, H., *et al.*, 2009. Localization of Inv in a distinctive intraciliary compartment requires the C-terminal ninein-homolog-containing region. *Journal of Cell Science* 122, 44–54.
- Signor, D., Wedaman, K.P., Orozco, J.T., *et al.*, 1999a. Role of a class DHC1b dynein in retrograde transport of IFT motors and IFT raft particles along cilia, but not dendrites, in chemosensory neurons of living *Caenorhabditis elegans*. *Journal of Cell Biology* 147, 519–530.
- Signor, D., Wedaman, K.P., Rose, L.S., Scholey, J.M., 1999b. Two heteromeric kinesin complexes in chemosensory neurons and sensory cilia of *Caenorhabditis elegans*. *Molecular Biology of the Cell* 10, 345–360.
- Sillibourne, J.E., Hurbain, I., Grand-Perret, T., *et al.*, 2013. Primary ciliogenesis requires the distal appendage component Cep123. *Biology Open* 2, 535–545.
- Silverman, M.A., Leroux, M.R., 2009. Intraflagellar transport and the generation of dynamic, structurally and functionally diverse cilia. *Trends in Cell Biology* 19, 306–316.
- Sinden, R.E., Talman, A., Marques, S.R., Wass, M.N., Sternberg, M.J., 2010. The flagellum in malarial parasites. *Current Opinion in Microbiology* 13, 491–500.
- Smith, E.F., Sale, W.S., 1992. Regulation of dynein-driven microtubule sliding by the radial spokes in flagella. *Science* 257, 1557–1559.
- Snow, J.J., Ou, G., Gunnarson, A.L., *et al.*, 2004. Two anterograde intraflagellar transport motors cooperate to build sensory cilia on *C. elegans* neurons. *Nature Cell Biology* 6, 1109–1113.
- Sorokin, S., 1962. Centrioles and the formation of rudimentary cilia by fibroblasts and smooth muscle cells. *Journal of Cell Biology* 15, 363–377.
- Sorokin, S.P., 1968. Reconstructions of centriole formation and ciliogenesis in mammalian lungs. *Journal of Cell Science* 3, 207–230.
- Strambio-De-Castillia, C., Niepel, M., Rout, M.P., 2010. The nuclear pore complex: Bridging nuclear transport and gene regulation. *Nature Reviews Molecular Cell Biology* 11, 490–501.
- Sung, C.H., Leroux, M.R., 2013. The roles of evolutionarily conserved functional modules in cilia-related trafficking. *Nature Cell Biology* 15, 1387–1397.
- Szymanska, K., Johnson, C.A., 2012. The transition zone: An essential functional compartment of cilia. *Cilia* 1, 10.
- Takeda, S., Yonekawa, Y., Tanaka, Y., *et al.*, 1999. Left-right asymmetry and kinesin superfamily protein KIF3A: New insights in determination of laterality and mesoderm induction by kif3A^{-/-} mice analysis. *Journal of Cell Biology* 145, 825–836.
- Tanos, B.E., Yang, H.J., Soni, R., *et al.*, 2013. Centriole distal appendages promote membrane docking, leading to cilia initiation. *Genes & Development* 27, 163–168.
- Taschner, M., Bhogaraju, S., Lorentzen, E., 2012. Architecture and function of IFT complex proteins in ciliogenesis. *Differentiation* 83, S12–S22.
- Taschner, M., Bhogaraju, S., Vetter, M., Morawetz, M., Lorentzen, E., 2011. Biochemical mapping of interactions within the intraflagellar transport (IFT) B core complex: IFT52 binds directly to four other IFT-B subunits. *Journal of Biological Chemistry* 286, 26344–26352.
- Tateishi, K., Yamazaki, Y., Nishida, T., *et al.*, 2013. Two appendages homologous between basal bodies and centrioles are formed using distinct Odf2 domains. *Journal of Cell Biology* 203, 417–425.
- Taulman, P.D., Haycraft, C.J., Balkovetz, D.F., Yoder, B.K., 2001. Polaris, a protein involved in left-right axis patterning, localizes to basal bodies and cilia. *Molecular Biology of the Cell* 12, 589–599.
- Tobin, J.L., Beales, P.L., 2007. Bardet-Biedl syndrome: Beyond the cilium. *Pediatric Nephrology* 22, 926–936.
- Tokuyasu, K.T., 1975. Dynamics of spermiogenesis in *Drosophila melanogaster*. VI. Significance of "onion" nebenkern formation. *Journal of Ultrastructure Research* 53, 93–112.

- Tran, P.V., Haycraft, C.J., Besschetnova, T.Y., *et al.*, 2008. THM1 negatively modulates mouse sonic hedgehog signal transduction and affects retrograde intraflagellar transport in cilia. *Nature Genetics* 40, 403–410.
- Tsao, C.C., Gorovsky, M.A., 2008. Different effects of *Tetrahymena* IFT172 domains on anterograde and retrograde intraflagellar transport. *Molecular Biology of the Cell* 19, 1450–1461.
- Vaughan, S., Kohl, L., Ngai, I., Wheeler, R.J., Gull, K., 2008. A repetitive protein essential for the flagellum attachment zone filament structure and function in *Trypanosoma brucei*. *Protist* 159, 127–136.
- Vieira, O.V., Gaus, K., Verkade, P., *et al.*, 2006. FAPP2, cilium formation, and compartmentalization of the apical membrane in polarized Madin-Darby canine kidney (MDCK) cells. *Proceedings of the National Academy of Sciences of the United States of America* 103, 18556–18561.
- Vincensini, L., Blisnick, T., Bastin, P., 2011. 1001 model organisms to study cilia and flagella. *Biology of the Cell* 103, 109–130.
- Wallingford, J.B., Mitchell, B., 2011. Strange as it may seem: The many links between Wnt signaling, planar cell polarity, and cilia. *Genes & Development* 25, 201–213.
- Walther, Z., Vashishtha, M., Hall, J.L., 1994. The *Chlamydomonas* FLA10 gene encodes a novel kinesin-homologous protein. *Journal of Cell Biology* 126, 175–188.
- Wang, J., Silva, M., Haas, L.A., *et al.*, 2014. *C. elegans* ciliated sensory neurons release extracellular vesicles that function in animal communication. *Current Biology* 24 (5), 519–525.
- Wann, A.K., Zuo, N., Haycraft, C.J., *et al.*, 2012. Primary cilia mediate mechanotransduction through control of ATP-induced Ca^{2+} signaling in compressed chondrocytes. *FASEB Journal* 26, 1663–1671.
- Ward, C.J., Hogan, M.C., Rossetti, S., *et al.*, 2002. The gene mutated in autosomal recessive polycystic kidney disease encodes a large, receptor-like protein. *Nature Genetics* 30, 259–269.
- Ward, C.J., Yuan, D., Masyuk, T.V., *et al.*, 2003. Cellular and subcellular localization of the ARPKD protein; fibrocystin is expressed on primary cilia. *Human Molecular Genetics* 12, 2703–2710.
- Wei, Q., Zhang, Y., Li, Y., *et al.*, 2012. The BBSome controls IFT assembly and turnaround in cilia. *Nature Cell Biology* 14, 950–957.
- Wheatley, D.N., Wang, A.M., Strugnell, G.E., 1996. Expression of primary cilia in mammalian cells 20, 73–81.
- Wicks, S.R., De Vries, C.J., Van Luenen, H.G., Plasterk, R.H., 2000. CHE-3, a cytosolic dynein heavy chain, is required for sensory cilium structure and function in *Caenorhabditis elegans*. *Developmental Biology* 221, 295–307.
- Wirschell, M., Nicastrò, D., Porter, M.E., Sale, W.S., 2009. The regulation of axonemal bending. In: Harris, E.H. (Ed.), *The Chlamydomonas Sourcebook*. London: Academic Press, pp. 253–282.
- Wirschell, M., Olbrich, H., Werner, C., *et al.*, 2013. The nexin-dynein regulatory complex subunit DRC1 is essential for motile cilium function in algae and humans. *Nature Genetics* 45, 262–268.
- Wirschell, M., Yamamoto, R., Alford, L., *et al.*, 2011. Regulation of ciliary motility: Conserved protein kinases and phosphatases are targeted and anchored in the ciliary axoneme. *Archives of Biochemistry and Biophysics* 510, 93–100.
- Wirschell, M., Zhao, F., Yang, C., *et al.*, 2008. Building a radial spoke: Flagellar radial spoke protein 3 (RSP3) is a dimer. *Cell Motility and the Cytoskeleton* 65, 238–248.
- Witman, G.B., Plummer, J., Sander, G., 1978. *Chlamydomonas* flagellar mutants lacking radial spokes and central tubules. Structure, composition, and function of specific axonemal components. *Journal of Cell Biology* 76, 729–747.
- Wood, C.R., Huang, K., Diener, D.R., Rosenbaum, J.L., 2013. The cilium secretes bioactive ectosomes. *Current Biology* 23, 906–911.
- Wood, C.R., Rosenbaum, J.L., 2014. Proteins of the ciliary axoneme are found on cytoplasmic membrane vesicles during growth of cilia. *Current Biology* 24, 1114–1120.
- Wren, K.N., Craft, J.M., Tritschler, D., *et al.*, 2013. A differential cargo-loading model of ciliary length regulation by IFT. *Current Biology* 23, 2463–2471.
- Yang, P., Smith, E.F., 2009. The flagellar radial spokes. In: Harris, E.H. (Ed.), *The Chlamydomonas Sourcebook*, second ed. Oxford: Academic Press, pp. 209–234.
- Ye, F., Breslow, D.K., Koslover, E.F., *et al.*, 2013. Single molecule imaging reveals a major role for diffusion in the exploration of ciliary space by signaling receptors. *Elife* 2, e00654.
- Yoder, B.K., Hou, X., Guay-Woodford, L.M., 2002. The polycystic kidney disease proteins, polycystin-1, polycystin-2, polaris, and cystin, are co-localized in renal cilia. *Journal of the American Society of Nephrology* 13, 2508–2516.
- Zaghloul, N.A., Katsanis, N., 2009. Mechanistic insights into Bardet-Biedl syndrome, a model ciliopathy. *Journal of Clinical Investigation* 119, 428–437.
- Zariwala, M.A., Knowles, M.R., Omran, H., 2007. Genetic defects in ciliary structure and function. *Annual Review of Physiology* 69, 423–450.
- Zeiske, E., Kasumyan, A., Bartsch, P., Hansen, A., 2003. Early development of the olfactory organ in sturgeons of the genus *Acipenser*: A comparative and electron microscopic study. *Anatomy and Embryology* 206, 357–372.
- Zimmermann, K.W., 1898. Beiträge zur Kenntniss einiger Drüsen und Epithelien. *Archiv für Mikroskopische Anatomie* 52, 552–706.

Further Reading

- Bhogaraju, S., Engel, B.D., Lorentzen, E., 2013b. Intraflagellar transport complex structure and cargo interactions. *Cilia* 2, 10.
- Christensen, S.T., Pedersen, L.B., Schneider, L., Satir, P., 2007. Sensory cilia and integration of signal transduction in human health and disease. *Traffic* 8, 97–109.
- Fisch, C., Dupuis-Williams, P., 2011. Ultrastructure of cilia and flagella — back to the future! *Biology of the Cell/Under the Auspices of the European Cell Biology Organization* 103, 249–270.
- Fliegauf, M., Benzing, T., Omran, H., 2007. When cilia go bad: Cilia defects and ciliopathies. *Nature Reviews Molecular Cell Biology* 8, 880–893.
- Ginger, M.L., Portman, N., McKean, P.G., 2008. Swimming with protists: Perception, motility and flagellum assembly. *Nature Reviews. Microbiology* 6, 838–850.
- Goetz, S.C., Anderson, K.V., 2010. The primary cilium: A signalling centre during vertebrate development. *Nature Reviews. Genetics* 11, 331–344.
- Hildebrandt, F., Benzing, T., Katsanis, N., 2011. Ciliopathies. *New England Journal of Medicine* 364, 1533–1543.
- Hsiao, Y.C., Tuz, K., Ferland, R.J., 2012. Trafficking in and to the primary cilium. *Cilia* 1, 4.
- Ishikawa, H., Marshall, W.F., 2011. Ciliogenesis: Building the cell's antenna. *Nature Reviews Molecular Cell Biology* 12, 222–234.
- Kenny, T.D., Beales, P.L., 2014. *Ciliopathies: A Reference for Clinicians*. Oxford: Oxford University Press.
- Nachury, M.V., Seeley, E.S., Jin, H., 2010. Trafficking to the ciliary membrane: How to get across the periciliary diffusion barrier? *Annual Review of Cell and Developmental Biology* 26, 59–87.
- Pedersen, L.B., Rosenbaum, J.L., 2008. Intraflagellar transport (IFT) role in ciliary assembly, resorption and signalling. *Current Topics in Developmental Biology* 85, 23–61.
- Satir, P., Christensen, S.T., 2006. Overview of structure and function of mammalian cilium. *Annual Review of Physiology* 69, 377–400.
- Satir, P., Pedersen, L.B., Christensen, S.T., 2010. The primary cilium at a glance. *Journal of Cell Science* 123, 499–503.
- Scholey, J.M., 2008. Intraflagellar transport motors in cilia: Moving along the cell's antenna. *Journal of Cell Biology* 180, 23–29.
- Silverman, M.A., Leroux, M.R., 2009. Intraflagellar transport and the generation of dynamic, structurally and functionally diverse cilia. *Trends in Cell Biology* 19, 306–316.
- Sung, C.H., Leroux, M.R., 2013. The roles of evolutionarily conserved functional modules in cilia-related trafficking. *Nature Cell Biology* 15, 1387–1397.
- Wirschell, M., Yamamoto, R., Alford, L., *et al.*, 2011. Regulation of ciliary motility: Conserved protein kinases and phosphatases are targeted and anchored in the ciliary axoneme. *Archives of Biochemistry and Biophysics* 510, 93–100.

Relevant Websites

- <http://www.bowserlab.org/primarycilium/ciliumpage2.htm>
Bowser Lab Primary Cilium Home Page.
- <http://cildb.cgm.cnrs-gif.fr>
Cildb: A Knowledgebase for Centrosomes and Cilia.
- <http://www.syscilia.org>
Homepage of the European Project SYSCILIA.
- <http://www.omim.org>
Online Mendelian Inheritance in Man: An Online Catalogue of Human Genetic Disorders.
- http://labs.umassmed.edu/chlamyftp/protector_login.php
The Chlamydomonas Flagellar Proteome Database.
- <http://www.ciliaproteome.org>
The Cilia Proteome Database from the Katsanis Lab.
- <http://www.ciliopathyalliance.org>
The Ciliopathy Alliance Homepage.

Skeletal Muscle

B Cadot, Sorbonne Universités UPMC, Paris Cedex, France

ER Gomes, Sorbonne Universités UPMC, Paris Cedex, France and Universidade de Lisboa, Lisbon, Portugal

© 2016 Elsevier Inc. All rights reserved.

The skeletal muscle is a highly differentiated tissue responsible for force generation involved in body movements. Coordination of different muscles is important for posture of the body, movement of arms and legs, and for fine movements of the fingers. Such activities are achieved by the connection of the muscle with the skeleton under the control of motor neurons. Muscle contraction is achieved by the contraction of individual muscle cells, known as myofibers, which are highly specialized long cells containing multiple nuclei that are distributed along the periphery of the cell, right next to the plasma membrane (sarcolemma).

Myofibers are the cells in the human body with the largest volume. Each myofiber is connected to a single neuron forming a neuromuscular synapse. Individual myofibers are surrounded by a basal lamina (endomysium). A small number of muscle stem cells (satellite cells) are found between the myofiber and basal lamina. Myofibers are grouped in muscle bundles surrounded by a connective tissue (perimysium) which together form a muscle. The extremities of each muscle are connected to the skeleton by tendons. All these structures lead to the translation of force generated by individual myofibers into the movement of the skeleton.

Myofibers are formed by the fusion of mononucleated myoblasts during embryogenesis in a process involving a complex change of the gene expression program of cells (Bentzinger *et al.*, 2012). In adult muscle, satellite cells can proliferate and give rise to mononucleated myoblasts that are able to repair damaged myofibers and form *de novo* myofibers. The proliferation of satellite cells is asymmetric so that a small number of cells retain their stem cell capability and give rise to satellite cells of the newly formed myofibers (Tajbakhsh, 2009).

Sarcomeres Are the Basic Unit of Contraction

Skeletal muscles exhibit regular striations along their length with about 2 μm spacing. These striations are formed by sarcomeres, the basic unit of contraction in the myofiber (Huxley, 1969). Sarcomeres are composed of multiple thick and thin filaments that overlap and each sarcomere is delimited by the Z disk (or Z line) (Cheng and Deatherage, 1989). Thick filaments are made of myosin and occupy a more central region of the sarcomere (the A band) whereas thin filaments are made of actin that is anchored to the Z disk. Different regions of the sarcomere have been named as bands due to their appearance when observed with polarized light (Figure 1). The region of thin filaments that does not overlap with the thick filaments is called the I band. The Z disk is in the middle of the I band.

Thick filaments of myosin are bipolar. Myosin tails form the core of the thick filament and point away from the extremities of the filament, whereas myosin heads stick out of the filament. The middle of the thick filament is called the M

line, which is also the axis of symmetry of the thick filament and the sarcomere (Agarkova and Perriard, 2005).

Thin filaments of actin are unipolarized; the plus (barbed) end of the filament is anchored to the Z disk and the minus (pointed) end is pointing to the middle of the sarcomere, the H band.

The region where myosin and actin filaments overlap (AI zone), one myosin filament is surrounded by six actin filaments and each actin filament is surrounded by two or three myosin filaments, depending on the species (Milligan and Flicker, 1987).

Multiple regulatory proteins bound to myosin and actin filaments are required for the structure and function of sarcomeres (Table 1). At least three gigantic proteins, titin, nebulin, and obscurin, play an important role on the function of sarcomeres. Each nebulin protein extends along the length of the actin filaments, from the Z disk toward the middle of the sarcomere and has a role stabilizing actin filaments. Titin (probably the biggest protein in humans) extends from the Z disk to the M line. Titin is an elastic protein and acts like a spring to keep the thick filaments in the center of the sarcomere. Obscurin is involved in the anchoring of sarcomeres to the sarcoplasmic reticulum (SR) and costameres (see below) (Gokhin and Fowler, 2013; Pappas *et al.*, 2011; Randazzo *et al.*, 2013; Tskhovrebova and Trinick, 2003).

Sarcomeric components are very dynamic, contrary to what we would expect from such a highly organized structure. For instance, actin, tropomyosin, and multiple Z disk proteins can incorporate into matured sarcomeres within minutes without any changes in sarcomere structure (Ono, 2010; Sanger *et al.*, 2010).

Sarcomeres are connected to the sarcolemma by the costamere, a protein complex. Costameres physically link the sarcolemma to the Z disk of the sarcomeres that are located right under the sarcolemma. In addition, costameres also link the sarcolemma to the extracellular matrix under the basal lamina (Ervasti, 2003). Costamere integrity is important for the transduction of force generated by the sarcomere to the sarcolemma and basal lamina, resulting in the contraction of the muscle (Rybakova *et al.*, 2000).

Sarcomere Contraction Is Driven by Sliding of Myosin and Actin Filaments

During muscle contraction, the length of myosin and actin filaments does not change. What changes is the distance between the extremities of the myosin filaments and the Z disks (the I band). Thus, during muscle contraction, the actin filaments move along the myosin filaments. This occurs because myosin heads bind to and move on actin filaments. Because both filaments are polarized, and myosin moves toward the barbed end of actin (at the Z disk), the movement of myosin

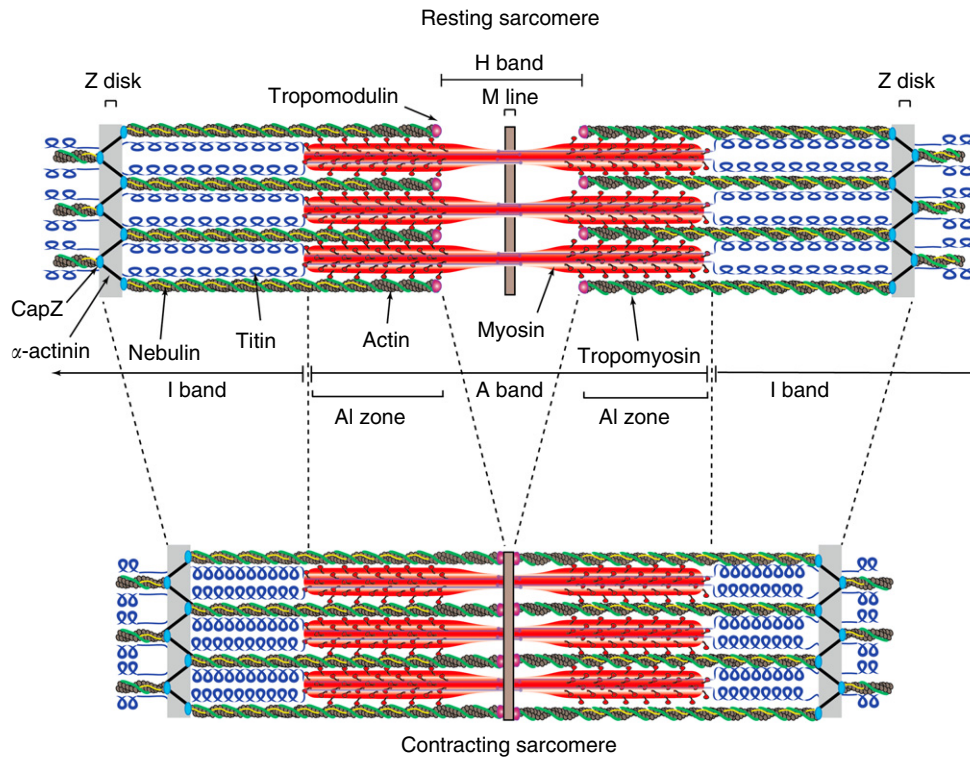


Figure 1 Sarcomere structure in resting and contracting states. Sarcomeres are mainly composed of actin (gray) and myosin (red) filaments. Nebulin (yellow) extends along the actin filament and titin (dark blue) stretches all the way from the Z disk to the M line. α -actinin (black) and CapZ (light blue) are Z disk components. Tropomyosin (green) extends along the actin filament and tropomodulin (purple) is at the extremity of the actin filament, the barbed end. In a contracted sarcomere, the length of the actin and myosin filaments does not change. Contraction occurs due to the sliding of actin filaments on myosin filaments. Therefore the AI zone increases whereas I band decreases, and the A band remains the same length. In contracted sarcomeres, titin is also shorter due to its function of spring to keep the integrity of the sarcomere.

Table 1 Sarcomere components

Name	Localization	Function	Disease
Actin	Actin filaments	Filament formation and sliding	Nemaline myopathy (Romero <i>et al.</i> , 2013)
Myosin	Myosin filaments	Filament formation and sliding	
C-protein	Myosin filaments	Stabilization	
M-protein	M line	Stabilization of myosin at M line	
α -actinin	Z disk	Connect actin filaments to Z disk	Desmin myopathy (Goldfarb <i>et al.</i> , 2004)
CapZ	Z disk	Caps actin filament barbed ends	
Desmin	Z disk	Stabilization	
Myotilin	Z disk	Binds to α -actinin	
Tropomyosin	Actin filaments	Prevents binding of myosin to actin	Nemaline myopathy (Vindin and Gunning, 2013)
Troponin I	Actin filaments	Inhibitory subunit of troponin	
Troponin T	Actin filaments	Tropomyosin-binding subunit of troponin	Nemaline myopathy (Kee and Hardeman, 2008)
Troponin C	Actin filaments	Calcium-binding subunit of troponin C	
Tropomodulin	Actin filaments	Binds to tropomyosin and caps actin filament pointed ends	Muscular dystrophy (Selcen and Carpén, 2008)
ZASP	Z disk	Stabilizes Z disk	
Nebulin	Actin filaments	Stabilizes actin filaments	
Titin	Actin and myosin filaments	Muscle elasticity	
Leiomodin3	Actin filaments	Actin nucleator	Nemaline myopathy (Chereau <i>et al.</i> , 2008; Yuen <i>et al.</i> , 2014)
Obscurin	M line and Z disk	Connect sarcomeres to sarcoplasmic reticulum and costameres	

heads connected to the actin filaments results in the sliding of these filaments and shortening of the sarcomere (distance between Z disks). The force produced by the sliding is proportional to the overlap between the myosin and actin filaments determined by the length of the sarcomere (Cooke, 1986; Gordon *et al.*, 1966).

Myosin movement is powered by the hydrolysis of ATP to ADP + Pi and each cycle of ATP binding and hydrolysis results in one stroke of the myosin head (Figure 2).

This process occurs in four steps. When ATP binds to myosin, myosin heads detach from actin filaments (step 1). Next, ATP is hydrolyzed to ADP + Pi (that remain bound to myosin) and results in the rotation of myosin heads to a position perpendicular to actin filaments (step 2). Next, in the presence of Ca^{++} , actin filament binds to the myosin head resulting in the release of Pi from myosin (step 3). Finally, the myosin head moves to about 45° of the actin filament, resulting in the movement of the actin filament and release of ADP (step 4). Step 4 is followed by step 1 with the binding of a new ATP molecule (Eisenberg and Hill, 1985; Geeves and Holmes, 1999; Irving, 1985).

A stroke of myosin that occurs during one cycle results in about 7 nm displacement of actin filament. However, each sarcomere contracts around 2000 nm. In order to achieve such a long range of movement, the detachment of myosin from actin and the existence of multiple myosin–actin binding sites that are in different stages of the ATP hydrolysis cycle is essential. This way, myosin filaments move on actin filaments like the legs of a walking centipede instead of the synchronized rows of a rowing galley.

When muscles are depleted of ATP, they become stiff and rigid, and this condition is named rigor. The best example of rigor is observed shortly after death. In rigor, the depletion of ATP and increase of cytosolic Ca^{++} concentration results in the tight binding of actin filaments to myosin heads preventing sliding of these filaments and causing muscle rigidity.

Calcium Release from the SR Activates Contraction

Muscle contraction is controlled by the motor neurons axons that form synapses with myofibers called the neuromuscular junction. Action potentials on the axons trigger the fusion of acetylcholine-containing synaptic vesicles with the plasma membrane, resulting in the release of acetylcholine into the cleft between the axon and the myofiber. Acetylcholine binds to the acetylcholine receptors, cation channels, leading to the opening of these channels and the generation of a new action potential in the myofiber plasma membrane (the sarcolemma).

Muscle contraction is triggered by the increase of cytosolic Ca^{++} concentration. Therefore, as long as the Ca^{++} concentration is high in the cytosol, and there is ATP, the myosin–actin connection and movement will continue to cycle and the muscle will remain contracted (Goldman *et al.*, 1982).

Sarcomeres form repetitive structures along the myofiber that are named myofibrils. Each myofibril is encircled by smooth endoplasmic reticulum, named SR that is the main Ca^{++} reservoir of the myofiber (Figure 3).

The SR network is interrupted by T-tubules, which are invaginations of the sarcolemma. T-tubules can be found near the Z disk or the A–I junction, depending on the organism and muscle type. The region of the SR that is next to each T-tubule is called terminal cisterna. A T-tubule is surrounded by two SR terminal cisternae forming a triad (Table 2). Triads play an important role on the control of muscle contraction. The T-tubule has voltage-sensitive Ca^{++} channels named dihydropyridine receptors (DHPR) that are physically connected to Ca^{++} channels in the SR named ryanodine receptor (Table 2; Franzini-Armstrong *et al.*, 1998). Depolarization of the T-tubule membrane due to the action potential generated at the neuromuscular synapse causes a conformation change of the DHPRs at the T-tubule. DHPRs directly induce the opening

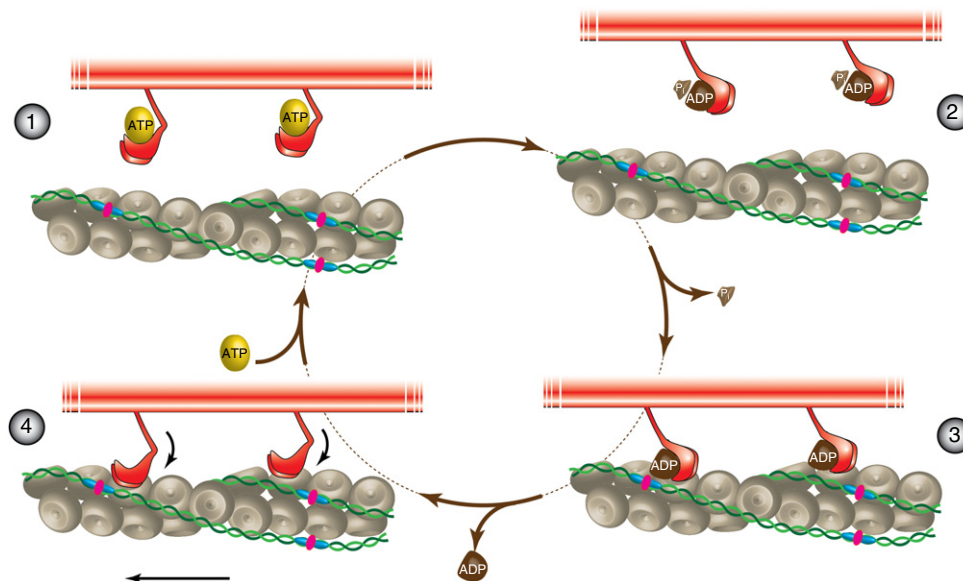


Figure 2 ATP hydrolysis and myosin stroke cycle. Step 1 – ATP binding to myosin and detachment from actin. Step 2 – ATP hydrolysis into ADP + Pi and myosin rotation. Step 3 – Binding of myosin to actin and Pi release. Step 4 – Myosin head stroke and ADP release.

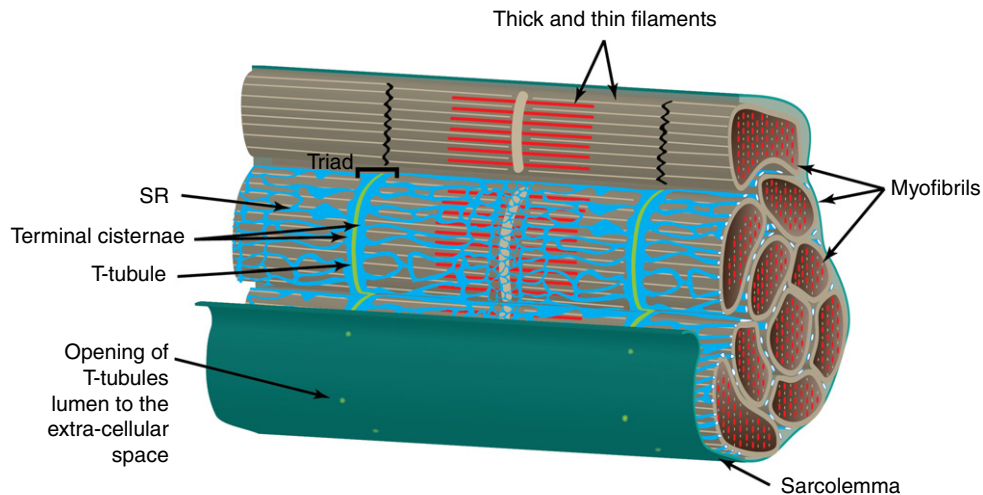


Figure 3 Organization of SR and T-tubules in skeletal muscle. Each myofiber is formed by multiple myofibrils (brown and red) that are surrounded by SR (light blue). The SR is interrupted by T-tubules (light green) that are invaginations of the sarcolemma, and therefore their lumen communicates with the extracellular space.

Table 2 Triad components

Name	Localization	Function	Disease
BIN1/ amphiphysin	T-tubule	Scaffolding	Centronuclear myopathy (Jungbluth <i>et al.</i> , 2008)
DHPR	T-tubule	Depolarization sensor	Malignant hyperthermia (Monnier <i>et al.</i> , 1997)
RyR	SR	Ca ⁺⁺ channel	Congenital muscular dystrophy (Bönnemann <i>et al.</i> , 2014)
Triadin	SR	Scaffolding	
Junctophilin	SR	Connect SR to T-tubule	
Ca ⁺⁺ ATPase	SR	Pumps Ca ⁺⁺ from cytosol into SR	Brody myopathy (Odermatt <i>et al.</i> , 1996)

of the Ca⁺⁺ channels of the ryanodine receptor at the SR which results in the release of Ca⁺⁺ from the SR into the cytosol. The increase of cytosolic Ca⁺⁺ triggers sarcomere contraction (Ridgway and Ashley, 1967). This process is called excitation-contraction coupling (Ríos *et al.*, 1992). Because the depolarization along the sarcolemma and T-tubules occurs in milliseconds, every sarcomere in a myofiber contracts simultaneously. When depolarization stops, Ca⁺⁺ ATPases in the SR pump Ca⁺⁺ back to the SR resulting in the decrease of cytosolic Ca⁺⁺ calcium concentration and muscle relaxation (Pellegrino and Franzini-Armstrong, 1969).

Increase of cytosolic Ca⁺⁺ concentration allows the binding of myosin heads to the actin filaments during step 3 of the cycle of ATP hydrolysis and myosin stroke. This process is mediated by tropomyosin and troponin, which are bound to actin filaments. When cytosolic Ca⁺⁺ concentration is low, tropomyosin and troponin prevent the binding of actin filaments to myosin heads. Tropomyosin binds laterally to actin filaments forming a long filament around the actin filament. Troponin is composed of three subunits: Troponin I, Troponin T, and Troponin C. When cytosolic Ca⁺⁺ concentration increases, it binds to Troponin C and results in a conformational change of Troponin I that leads to a shift of tropomyosin relative to actin filament (Figure 4).

The exposed actin filament surface is then more likely to bind to ADP-Pi myosin which is followed by Pi release and myosin movement. The binding of the first myosin head to actin shifts the tropomyosin filament beyond the original binding sites facilitating the binding of adjacent myosin heads to actin filaments (Phillips *et al.*, 1986).

Muscles can contract different lengths and at different speeds and forces. Such regulation does not occur at the level of each individual myofiber since each fiber is either contracted or relaxed. Instead this regulation occurs at the level of the motor neurons that are connected to the myofibers. Motor neurons are connected to a myofiber by the neuromuscular junction, also known as the neuromuscular synapse. This synapse works as a neuron-neuron synapse where the action potential of an axon is transferred to a downstream neurite. At the neuromuscular junction, the myofiber is downstream of the axon, and the action potential of the axon results in the depolarization of sarcolemma that quickly travels along the all sarcolemma and T-tubules leading to myofiber contraction. Each motor neuron is connected to one or more myofibers, forming a motor unit (English and Wolf, 1982). However, each myofiber has only one neuromuscular junction connected to a single motor neuron. Therefore, muscle length and speed is regulated by the number of motor units that are

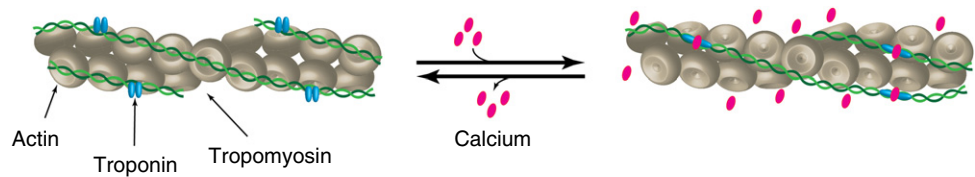


Figure 4 Conformation change of tropomyosin and troponin relative to actin filaments upon increase of cytosolic Ca^{++} . In low cytosolic Ca^{++} (pink) concentration, troponin (blue) and tropomyosin (green) occupies the myosin binding side of each actin subunit along the actin filaments. When cytosolic Ca^{++} concentration increases, the binding of Ca^{++} to troponin C leads to the shifting of tropomyosin relative to the actin filament, exposing myosin binding sites and allowing the binding of myosin heads to actin.

Table 3 Costamere components

Name	Localization	Function	Disease
Dystrophin	Cytosol	Connect actin to sarcolemma	Muscular dystrophy (Rahimov and Kunkel, 2013)
β -Dystroglycan	Sarcolemma	Connects sarcolemma to extracellular matrix	Limb-girdle muscular dystrophy (Rahimov and Kunkel, 2013)
α -, β -, and γ - Sarcoglycan	Sarcolemma	Connects sarcolemma to extracellular matrix	Limb-girdle muscular dystrophy (Rahimov and Kunkel, 2013)
Integrin $\alpha 7$	Extracellular	Extracellular matrix structure	Congenital myopathy (Hayashi <i>et al.</i> , 1998)
Collagen VI	Extracellular	Extracellular matrix structure	Bethlem myopathy (Briñas <i>et al.</i> , 2010)

activated. The more the motor units are activated, the higher and faster is the contraction. On the other hand, the rate of stimulation by the motor neuron controls the force produced by each myofiber in a motor unit.

Muscle Disorders Associated with Skeletal Muscle

Multiple genetic disorders have been identified to be caused by mutation in genes encoding proteins from the sarcomeres, costameres, extracellular matrix, triads, and nuclear envelope in the skeletal muscle (Tables 1–3). The first mutation to be identified was a protein from the costamere that was named dystrophin, since it was found in a patient with Duchenne’s muscular dystrophy. Mutations in dystrophin and other costameric proteins lead to the weakness of the connection between the sarcomere and the sarcolemma which results in myofiber damage during muscle activity (Rahimov and Kunkel, 2013). Due to this damage, the muscle of these patients is constantly undergoing regeneration. However, with age, the ability for the muscle to regenerate is lost, leading to muscle impairment. Patients with Duchenne muscular dystrophy start losing their locomotion capability and the premature cause of death is mostly caused by failure of respiratory muscles.

See also: Cytoskeleton and Motors: Cytoskeletal Components: Actin Assembly Dynamics and Its Regulation in Motile and Morphogenetic Processes; Myosins

References

Agarkova, I., Perriard, J.-C., 2005. The M-band: An elastic web that crosslinks thick filaments in the center of the sarcomere. *Trends in Cell Biology* 15, 477–485. doi:10.1016/j.tcb.2005.07.001.

Bentzinger, C.F., Wang, Y.X., Rudnicki, M.A., 2012. Building muscle: Molecular regulation of myogenesis. *Cold Spring Harbor Perspectives in Biology* 4, a008342. doi:10.1101/cshperspect.a008342.

Bönnemann, C.G., Wang, C.H., Quijano-Roy, S., *et al.*, 2014. Diagnostic approach to the congenital muscular dystrophies. *Neuromuscular Disorders* 24, 289–311. doi:10.1016/j.nmd.2013.12.011.

Briñas, L., Richard, P., Quijano-Roy, S., *et al.*, 2010. Early onset collagen VI myopathies: Genetic and clinical correlations. *Annals of Neurology* 68, 511–520. doi:10.1002/ana.22087.

Chauveau, C., Rowell, J., Ferreira, A., 2014. A rising titan: TTN review and mutation update. *Human Mutation* 35, 1046–1059. doi:10.1002/humu.22611.

Cheng, N.Q., Deatherage, J.F., 1989. Three-dimensional reconstruction of the Z disk of sectioned bee flight muscle. *Journal of Cell Biology* 108, 1761–1774.

Chereau, D., Boczkowska, M., Skwarek-Maruszewska, A., *et al.*, 2008. Leiomodin is an actin filament nucleator in muscle cells. *Science* 320, 239–243. doi:10.1126/science.1155313.

Cooke, R., 1986. The mechanism of muscle contraction. *CRC Critical Reviews in Biochemistry* 21, 53–118.

Eisenberg, E., Hill, T.L., 1985. Muscle contraction and free energy transduction in biological systems. *Science* 227, 999–1006.

English, A.W., Wolf, S.L., 1982. The motor unit. *Anatomy and physiology. Physical Therapy* 62, 1763–1772.

Ervasti, J.M., 2003. Costameres: The Achilles’ heel of Herculean muscle. *Journal of Biological Chemistry* 278, 13591–13594. doi:10.1074/jbc.R200021200.

Franzini-Armstrong, C., Protasi, F., Ramesh, V., 1998. Comparative ultrastructure of Ca^{2+} release units in skeletal and cardiac muscle. *Annals of the New York Academy of Sciences* 853, 20–30.

Geeves, M.A., Holmes, K.C., 1999. Structural mechanism of muscle contraction. *Annual Review of Biochemistry* 68, 687–728. doi:10.1146/annurev.biochem.68.1.687.

Gokhin, D.S., Fowler, V.M., 2013. A two-segment model for thin filament architecture in skeletal muscle. *Nature Reviews Molecular Cell Biology* 14, 113–119. doi:10.1038/nrm3510.

Goldfarb, L.G., Vicart, P., Goebel, H.H., Dalakas, M.C., 2004. Desmin myopathy. *Brain Journal of Neurology* 127, 723–734. doi:10.1093/brain/awh033.

Goldman, Y.E., Hibberd, M.G., McCray, J.A., Trentham, D.R., 1982. Relaxation of muscle fibres by photolysis of caged ATP. *Nature* 300, 701–705.

Gordon, A.M., Huxley, A.F., Julian, F.J., 1966. The variation in isometric tension with sarcomere length in vertebrate muscle fibres. *Journal of Physiology* 184, 170–192.

Hayashi, Y.K., Chou, F.-L., Engvall, E., *et al.*, 1998. Mutations in the integrin $\alpha 7$ gene cause congenital myopathy. *Nature Genetics* 19, 94–97. doi:10.1038/ng0598-94.

Huxley, H.E., 1969. The mechanism of muscular contraction. *Science* 164, 1356–1365.

- Irving, M., 1985. Muscle contraction: Weak and strong crossbridges. *Nature* 316, 292–293. doi:10.1038/316292a0.
- Jungbluth, H., Wallgren-Pettersson, C., Laporte, J., 2008. Centronuclear (myotubular) myopathy. *Orphanet Journal of Rare Diseases* 3, 26. doi:10.1186/1750-1172-3-26.
- Kee, A.J., Hardeman, E.C., 2008. Tropomyosins in skeletal muscle diseases. *Advances in Experimental Medicine and Biology* 644, 143–157.
- Labeit, S., Ottenheijm, C.A.C., Granzier, H., 2011. Nebulin, a major player in muscle health and disease. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 25, 822–829. doi:10.1096/fj.10-157412.
- Milligan, R.A., Flicker, P.F., 1987. Structural relationships of actin, myosin, and tropomyosin revealed by cryo-electron microscopy. *Journal of Cell Biology* 105, 29–39.
- Monnier, N., Procaccio, V., Stieglitz, P., Lunardi, J., 1997. Malignant-hyperthermia susceptibility is associated with a mutation of the alpha 1-subunit of the human dihydropyridine-sensitive L-type voltage-dependent calcium-channel receptor in skeletal muscle. *American Journal of Human Genetics* 60, 1316–1325.
- Odermatt, A., Taschner, P.E., Khanna, V.K., *et al.*, 1996. Mutations in the gene-encoding SERCA1, the fast-twitch skeletal muscle sarcoplasmic reticulum Ca²⁺-ATPase, are associated with Brody disease. *Nature Genetics* 14, 191–194. doi:10.1038/ng1096-191.
- Ono, S., 2010. Dynamic regulation of sarcomeric actin filaments in striated muscle. *Cytoskeleton* 67, 677–692. doi:10.1002/cm.20476.
- Pappas, C.T., Bliss, K.T., Zieseniss, A., Gregorio, C.C., 2011. The Nebulin family: An actin support group. *Trends in Cell Biology* 21, 29–37. doi:10.1016/j.tcb.2010.09.005.
- Pellegrino, C., Franzini-Armstrong, C., 1969. Recent contributions of electron microscopy to the study of normal and pathological muscle. *International Review of Experimental Pathology* 7, 139–226.
- Phillips, G.N., Fillers, J.P., Cohen, C., 1986. Tropomyosin crystal structure and muscle regulation. *Journal of Molecular Biology* 192, 111–131.
- Rahimov, F., Kunkel, L.M., 2013. Cellular and molecular mechanisms underlying muscular dystrophy. *Journal of Cell Biology* 201, 499–510. doi:10.1083/jcb.201212142.
- Randazzo, D., Giacomello, E., Lorenzini, S., *et al.*, 2013. Obscurin is required for ankyrinB-dependent dystrophin localization and sarcolemma integrity. *Journal of Cell Biology* 200, 523–536. doi:10.1083/jcb.201205118.
- Ridgway, E.B., Ashley, C.C., 1967. Calcium transients in single muscle fibers. *Biochemical and Biophysical Research Communications* 29, 229–234.
- Ríos, E., Pizarro, G., Stefani, E., 1992. Charge movement and the nature of signal transduction in skeletal muscle excitation-contraction coupling. *Annual Review of Physiology* 54, 109–133. doi:10.1146/annurev.ph.54.030192.000545.
- Romero, N.B., Sandaradura, S.A., Clarke, N.F., 2013. Recent advances in nemaline myopathy. *Current Opinion in Neurology* 26, 519–526. doi:10.1097/WCO.0b013e328364d681.
- Rybakova, I.N., Patel, J.R., Ervasti, J.M., 2000. The dystrophin complex forms a mechanically strong link between the sarcolemma and costameric actin. *Journal of Cell Biology* 150, 1209–1214. doi:10.1083/jcb.150.5.1209.
- Sanger, J.W., Wang, J., Fan, Y., White, J., Sanger, J.M., 2010. Assembly and dynamics of myofibrils. *Journal of Biomedicine and Biotechnology* 2010. doi:10.1155/2010/858606.
- Selcen, D., Carpén, O., 2008. The Z-disk diseases. *Advances in Experimental Medicine and Biology* 642, 116–130.
- Tajbakhsh, S., 2009. Skeletal muscle stem cells in developmental versus regenerative myogenesis. *Journal of Internal Medicine* 266, 372–389. doi:10.1111/j.1365-2796.2009.02158.x.
- Tskhovrebova, L., Trinick, J., 2003. Titin: Properties and family relationships. *Nature Reviews Molecular Cell Biology* 4, 679–689. doi:10.1038/nrm1198.
- Vindin, H., Gunning, P., 2013. Cytoskeletal tropomyosins: Choreographers of actin filament functional diversity. *Journal of Muscle Research and Cell Motility* 34, 261–274. doi:10.1007/s10974-013-9355-8.
- Yuen, M., Sandaradura, S.A., Dowling, J.J., *et al.*, 2014. Leiomodion-3 dysfunction results in thin filament disorganization and nemaline myopathy. *Journal of Muscle Research and Cell Motility*. doi:10.1172/JCI75199.

Further Reading

- Brown, R.H., 1997. Dystrophin-associated proteins and the muscular dystrophies. *Annual Review of Medicine* 48, 457–466. doi:10.1146/annurev.med.48.1.457.
- Clark, K.A., McElhinny, A.S., Beckerle, M.C., Gregorio, C.C., 2002. Striated muscle cytoarchitecture: An intricate web of form and function. *Annual Review of Cell and Developmental Biology* 18, 637–706. doi:10.1146/annurev.cellbio.18.012502.105840.
- Constantin, B., 2014. Dystrophin complex functions as a scaffold for signalling proteins. *Biochimica et Biophysica Acta* 1838, 635–642. doi:10.1016/j.bbame.2013.08.023.
- Davies, K.E., Nowak, K.J., 2006. Molecular mechanisms of muscular dystrophies: Old and new players. *Nature Reviews Molecular Cell Biology* 7, 762–773. doi:10.1038/nrm2024.
- Engel, A.G.M.D., Franzini-Armstrong, C., 1994. *Myology*, 2nd edn. New York, NY: McGraw-Hill Professional Publishing.
- Gordon, A.M., Homsher, E., Regnier, M., 2000. Regulation of contraction in striated muscle. *Physiological Reviews* 80, 853–924.
- Sugi, H., 2007. *Sliding Filament Mechanism in Muscle Contraction: Fifty Years of Research*. New York, NY: Springer Science & Business Media.
- Treves, S., Vukcevic, M., Maj, M., *et al.*, 2009. Minor sarcoplasmic reticulum membrane components that modulate excitation-contraction coupling in striated muscles. *Journal of Physiology* 587, 3071–3079. doi:10.1113/jphysiol.2009.171876.

Filopodia and Lamellipodia

TM Svitkina, University of Pennsylvania, Philadelphia, PA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Barbed end/pointed end Opposite ends of intrinsically polar actin filaments with the barbed end (plus end) exhibiting more rapid dynamics than the pointed end (minus end).

Filopodia Thin, filamentous protrusions extending from the cell surface and responsible for probing the surrounding environment and reaching distant targets.

Lamellipodia Broad and flat sheet-like cellular protrusions.

Nucleation Initial slow phase of polymerization that involves thermodynamically unfavorable clustering of first few monomers. This cluster, or nucleus, then undergoes elongation by incorporating additional subunits.

Ruffles Cellular protrusions formed by uplifting of lamellipodia.

Treadmilling Phenomenon of actin dynamics at which actin monomers incorporate at the barbed end of the actin filament and dissociate from its pointed end at the same ambient concentration of actin monomers.

Introduction

Individual cell migration consists of repeating cycles of protrusion and attachment of the cell front followed by detachment and retraction of the rear. Protrusion of the leading edge is an important component of cell locomotion. It typically occurs in the form of planar protrusions termed lamellipodia or slender protrusions termed filopodia (Figure 1). Lamellipodia can generate large pushing forces and carry the main load of propelling the cell front, whereas filopodia can extend far beyond the cell edge and reach or sense distant targets. A balance between filopodial and lamellipodial modes of protrusion varies broadly among cell types and may reflect relative needs of the cell to undergo fast locomotion versus precise navigation (Figure 2).

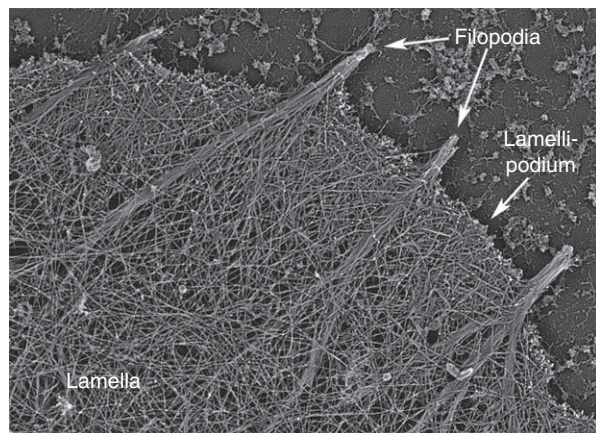


Figure 1 Organization of actin filaments at the leading edge of a fibroblast-like mouse melanoma cell visualized by transmission electron microscopy after detergent extraction, chemical fixation, critical point drying, and rotary shadowing of cultured cells with platinum and carbon. Modified from Yang, C., Czech, L., Gerboth, S., *et al.*, 2007. Novel roles of formin mDia2 in lamellipodia and filopodia formation in motile cells. *PLoS Biology* 5, e317.

The molecular machinery driving cell locomotion converges on the actin cytoskeleton, a complex system of actin filaments and numerous accessory proteins regulating actin filament dynamics and spatial organization. Actin cytoskeleton is the major force-generating machinery in the cell that can produce pushing, pulling, and resistance forces by forming various higher order ensembles, networks and bundles, adapted to specific tasks. Dynamic properties of the actin cytoskeleton allow a cell to quickly redesign actin structures according to its changing needs.

Both lamellipodia and filopodia generate pushing force using the same molecular strategy, coordinated polymerization of multiple interconnected actin filaments. Acting as a cohort, actin filaments can achieve necessary stiffness to produce large forces. Filopodia and lamellipodia coordinate polymerization of multiple filaments in different ways. Specifically, actin filaments form parallel bundles in filopodia, but branching networks in lamellipodia. Accordingly, actin polymerization in these protrusions is regulated by distinct, but overlapping, sets of accessory proteins.

General Principles of Actin Polymerization-Driven Protrusion

The concept that actin polymerization alone can produce protrusive force has emerged from a series of studies of the acrosomal reaction in invertebrate sperm by Tilney and coworkers (Tilney, 1975). Although this discovery was made in a specialized cellular system, the concept has general significance and applies to lamellipodia and filopodia. The directionality of pushing force originates from intrinsic polarity of the actin filament, which is further amplified by hydrolysis of actin-bound ATP accompanying actin polymerization. Due to structural and biochemical differences, one end of the actin filament ('barbed' or 'plus' end) can grow in the same conditions, in which the other ('pointed' or 'minus') end shrinks (Figure 3), the process being called treadmilling. In lamellipodia and filopodia, actin filaments are oriented with their barbed ends facing the plasma membrane (Small *et al.*, 1978).

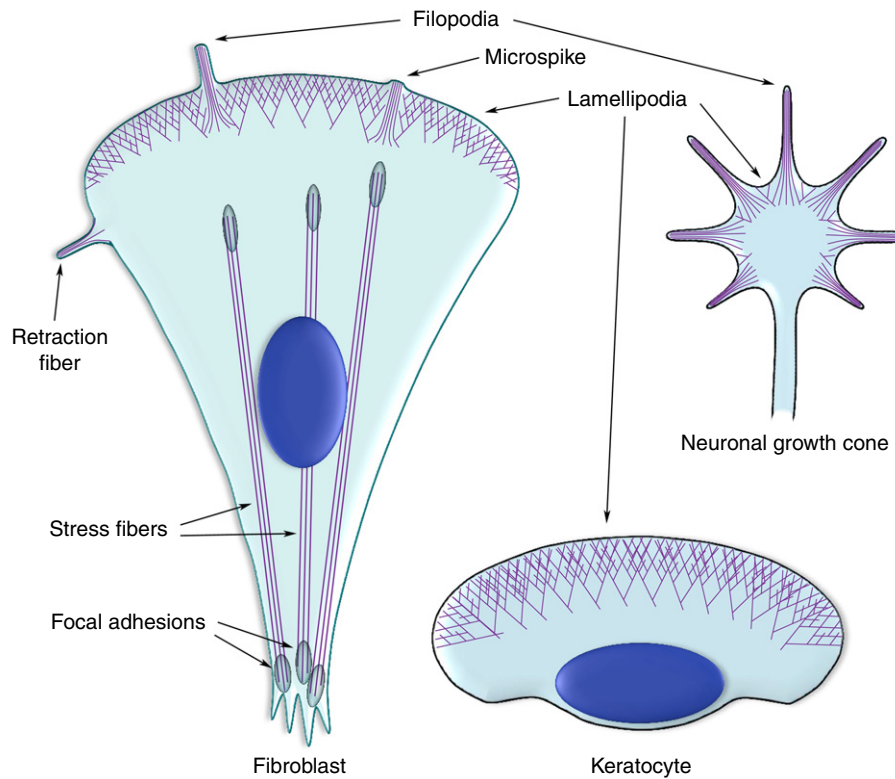


Figure 2 Schematics of actin filament organization at the leading edge of different cell types. Fibroblasts display a mixture of lamellipodia, filopodia, microspikes, and retraction fibers. They also contain contractile actin-myosin II bundles (stress fibers in the cell interior). Leading edge of neuronal growth cones is especially rich in filopodia with small lamellipodia located between filopodia. Fish or frog epidermal keratocytes typically express only lamellipodia. Modified from Svitkina, T., 2013. Actin organization. In: Lennarz, W.J., Daniel Lane, M. (Eds.), *Encyclopedia of Biological Chemistry*. Waltham, MA: Academic Press, pp. 27–35.

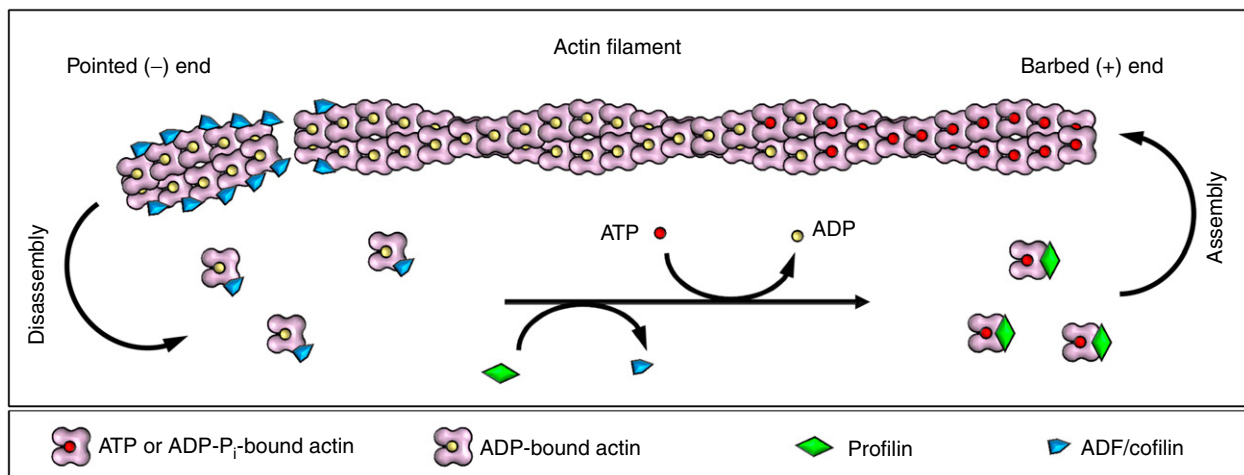


Figure 3 Actin filament treadmilling. Intrinsic structural polarity and polymerization-linked ATP hydrolysis allows an actin filament to undergo treadmilling by assembling subunits at the barbed end (right) and releasing subunits from the pointed end (left). This dynamic behavior is accelerated by profilin and ADF/cofilin. Modified from Svitkina, T., 2013. Actin organization. In: Lennarz, W.J., Daniel Lane, M. (Eds.), *Encyclopedia of Biological Chemistry*. Waltham, MA: Academic Press, pp. 27–35.

Polymerizing barbed ends push onto the membrane, whereas actin filament disassembly occurs closer to the pointed ends. Balanced rates of polymerization and depolymerization would allow a cell to maintain steady-state actin turnover. Intrinsic

treadmilling of actin filaments observed biochemically is too slow to account for fast rates of leading edge protrusion in living cells (Pollard and Mooseker, 1981), but can be accelerated by accessory proteins (Carlier *et al.*, 1997). Key

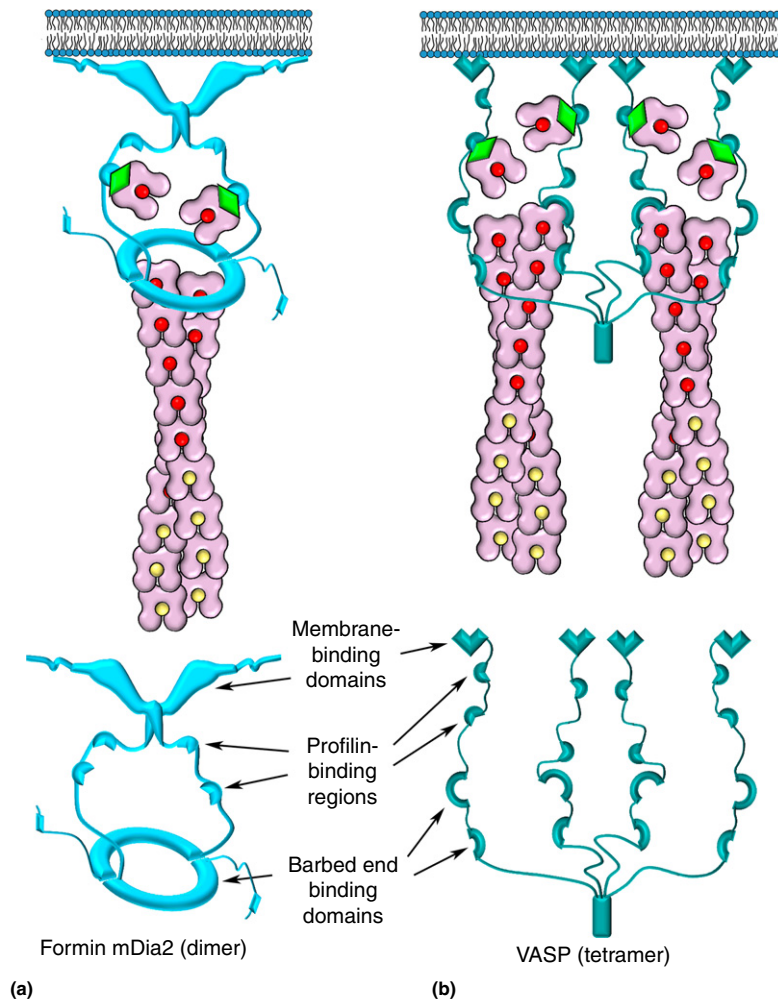


Figure 4 Formins (a) and Ena/VASPs (b) promote barbed end elongation. The mechanism of action of these unrelated proteins involves processive tracking and anti-capping protection of the growing end using actin-binding domains, recruitment of profilin-actin complexes using profilin-binding proline-rich regions, and anchorage to the plasma membrane using membrane-binding domains.

regulators of facilitated actin turnover are common for filopodia and lamellipodia (Figures 3 and 4).

Profilin is a small actin monomer binding protein, which functions to ensure that monomer addition occurs specifically at the barbed end of preexisting filaments (Tilney *et al.*, 1983). Profilin binds the barbed end surface of an actin monomer thereby precluding monomer interaction with the pointed end of other monomers or polymers (Figure 3). As a result, both spontaneous filament nucleation and elongation of filaments at their pointed ends are suppressed by profilin (Yarmola and Bubb, 2009). After addition of a profilin-actin complex into a filament barbed end, profilin dissociates leaving the barbed end free to assemble additional subunits. Profilin also functions as a nucleotide exchange factor by loading actin monomers with ATP.

Proteins of the actin depolymerizing factor (ADF)/cofilin family sever actin filaments and thus facilitate their depolymerization by increasing the number of disassembling ends (Bernstein and Bamburg, 2010; Figure 3). ADF/cofilins can bind both actin monomers and filaments with preference for ADP-containing actin subunits in either case (Carlier *et al.*,

1997). By cooperatively binding to the side of the actin filament, ADF/cofilins overtwist the actin filament helix and destabilize subunit-subunit interaction at the boundary between cofilin-loaded and cofilin-free actin filament segments. Higher affinity for ADP-containing actin subunits allows ADF/cofilin to preferentially accelerate depolymerization of an older part of the actin filament located closer to its pointed end.

Two unrelated protein families, Ena/VASP and formins, use analogous mechanisms to accelerate barbed end elongation and extend duration of elongation (Dominguez, 2010). Both formins and Ena/VASPs bind barbed ends and allow for addition of new actin subunits to the barbed end without dissociating from the end, a process called processive tracking (Figure 4). Using actin-binding or profilin-binding domains, Ena/VASPs and formins increase the local concentration of polymerizable monomers near the growing barbed end and thus accelerate its elongation. Membrane-binding domains within these proteins allow them to keep elongating ends near the membrane to increase efficiency of pushing. Formins and Ena/VASPs also prevent binding of

capping proteins, which could terminate barbed end elongation.

Filopodia

Definition

Filamentous pseudopodia had been known since the mid-1800s, when the morphologies of protrusive organelles characteristic of various Protozoa were classified. The term 'filopodia' was introduced to define these long and slender cellular protrusions. With an advent of cell culture (Harrison, 1907), filopodia were also documented in mammalian cells. At present, this term is variably used to designate a broad class of cellular protrusions characterized by a high length-to-width ratio, as introduced originally, or to refer to a best-studied class of filopodia at the leading edge of migrating animal cells.

Behavior and Functions

Filopodia are found in many different cell types and broadly vary in morphology. They often display not only protrusive, retractile, and sweeping motility, but also exist as stationary anchored cellular extensions. Because of their elongated geometry, filopodia are well designed to function as 'sensors,' which can reach, detect, and grasp distant targets. Accordingly, filopodia are believed to guide cell locomotion during normal tissue morphogenesis or cancer metastasis by recognizing adhesive surfaces and making initial contacts with the extracellular matrix. Filopodia of neuronal growth cones are thought to recognize attractive and repulsive cues and direct the axon to its destination (Drees and Gertler, 2008), whereas filopodia of endothelial tip cells appear to guide the growth of blood vessels during angiogenesis (De Smet *et al.*, 2009). Filopodia can also establish communication with other cells by finding a proper cellular partner and making initial contacts. For example, filopodial bridges formed between contacting endothelial (Hoelzle and Svitkina, 2012) or epithelial cells (Vasioukhin *et al.*, 2000) are important for subsequent formation of adherens junctions. In neurons, dendritic filopodia establish initial excitatory synapses with axons and then differentiate into dendritic spines (Bourne and Harris, 2008). Filopodia can also capture various particles, such as pathogens, for subsequent internalization (Romero *et al.*, 2011).

Structure and Dynamics

Filopodia that are formed at the leading edge of migrating cells, such as fibroblasts, or cells undergoing dramatic shape changes, such as neurons, are best characterized structurally and dynamically. These filopodia contain an actin filament bundle, in which individual filaments span the entire length of the filopodium and are uniformly oriented with barbed ends toward the filopodial tip (Small *et al.*, 1978; Figure 5). During protrusion, actin subunits are added at the filopodium tip, move away from the tip as part of the filament lattice, and are released at the rear of filopodium, thus undergoing treadmilling (Wang, 1985). Introduction of fiducial marks into the actin filament bundle in filopodia demonstrated that while the

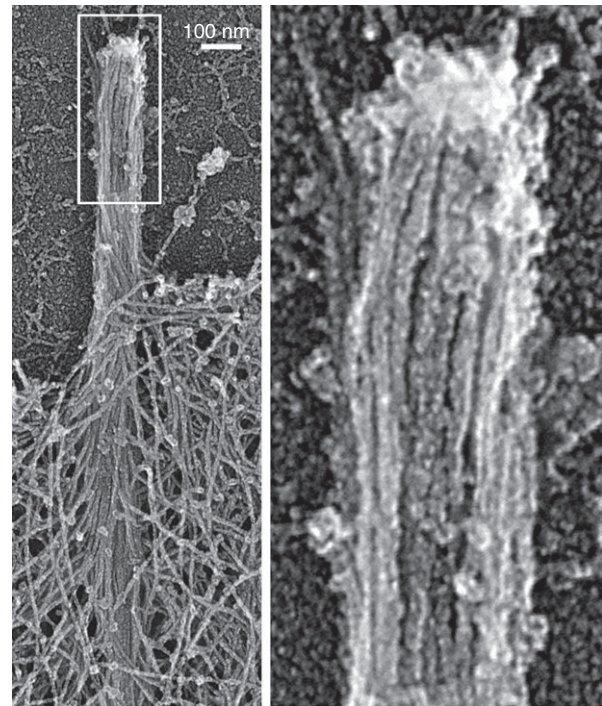


Figure 5 Actin filament organization in a filopodium at the leading edge of a mouse melanoma cell. Boxed region is enlarged in the right panel and shows long parallel actin filaments. Transmission electron microscopy of the cell after detergent extraction, chemical fixation, critical point drying, and rotary shadowing with platinum and carbon. Modified from Svitkina, T., 2013. Actin organization. In: Lennarz, W.J., Daniel Lane, M. (Eds.), *Encyclopedia of Biological Chemistry*. Waltham, MA: Academic Press, pp. 27–35.

filopodial tip grows forward the entire bundle slides backward (Mallavarapu and Mitchison, 1999). A balance between rates of protrusion and retrograde flow varies among individual filopodia and can be controlled, at least in part, by the strength of linkage(s) between the filopodial cytoskeleton and the substrate (Suter *et al.*, 1998).

Other cellular structures, such as microspikes and retraction fibers (Figure 1), are similar to leading edge filopodia in structural organization. Like filopodia, both microspikes and retraction fibers contain a parallel bundle of actin filaments. However, microspikes in contrast to filopodia do not significantly extend beyond the cell leading edge, but protrude at the same rate as the surrounding cell edge. Retraction fibers have an elongated narrow shape, like filopodia, but are formed in the course of cell edge retraction, rather than via their own protrusion. Yet, some retraction fibers can elongate at the tip, thus sharing dynamic features with filopodia. Similar structure and dynamics, as well as frequent interconversion between these organelles, indicate that filopodia, microspikes, and retraction fibers are related structures (Svitkina *et al.*, 2003).

Parallel actin filament bundles similar to those in filopodia are also present in other cylindrical cell surface extensions, such as brush border microvilli in intestinal epithelium (Mooseker and Tilney, 1975) and hair cell stereocilia in the inner ear (Tilney *et al.*, 1980). Unlike filopodia, microvilli and stereocilia maintain a constant length and exhibit very slow

(microvilli) (Tyska and Mooseker, 2002) or negligible (stereocilia) actin dynamics (Zhang *et al.*, 2012).

Notably, not all cylindrical membrane protrusions contain parallel actin bundles. It is becoming increasingly clear that filopodia defined purely by shape significantly vary in their structure, dynamics, and molecular composition. For example, dendritic filopodia of neuronal cells, as well as narrow elongated necks of dendritic spines, contain loosely cross-linked actin networks, in which a small fraction of actin filaments is oriented with their barbed end away from the tip of the protrusion (Korobova and Svitkina, 2010). Some experimentally induced filopodia also contain relatively sparse actin filaments having variable lengths and mutual positions (Yang *et al.*, 2009).

Molecular Machinery for Filopodial Protrusion

Filopodia protrusion involves coordinated changes in the membrane and the actin cytoskeleton. Actin polymerization and membrane deformation during filopodia protrusion are regulated by distinct, but interacting molecular machineries.

Actin cytoskeleton

Formation of tight bundles of uniformly oriented and synchronously polymerizing actin filaments is a key strategy used by leading edge filopodia to generate pushing force. The dynamics of the actin cytoskeleton during steady state filopodial protrusion is described by the treadmilling model. It states that actin filaments within the filopodial bundle continuously elongate at their distal barbed ends and depolymerize at the rear close to the pointed ends (Mallavarapu and Mitchison, 1999; Wang, 1985). Various aspects of actin dynamics in filopodia are regulated by proteins controlling actin filament elongation, cross-linking, and disassembly (Figure 6).

Elongation of filopodial filaments in the presence of abundant capping proteins in the cytoplasm is largely controlled by proteins of Ena/VASP or formin families, although relative contribution of these protein families varies among cell types. Ena/VASP proteins are particularly important for formation of filopodia in neuronal growth cones (Lebrand *et al.*, 2004), but share this function with formins in other cell types (Homem and Peifer, 2009). Consistent with their ability to processively track growing barbed ends, Ena/VASP proteins and specific formins are enriched at filopodial tips and promote actin filament elongation *in vivo* (Svitkina *et al.*, 2003; Yang *et al.*, 2007). Since formins are also actin nucleators (Pruyne *et al.*, 2002; Sagot *et al.*, 2002), in principle, they may contribute to filopodia formation in this capacity. However, formins poorly nucleate actin filaments from profilin-actin complexes, the main polymerizable actin species in the cytoplasm, and need to cooperate with other actin filament nucleators, such as the Arp2/3 complex (Svitkina *et al.*, 2003; Yang *et al.*, 2007), adenomatous polyposis coli (APC) (Okada *et al.*, 2010), or Spire (Bosch *et al.*, 2007; Quinlan *et al.*, 2007) to overcome this barrier. Formins and Ena/VASPs also function to processively anchor growing actin filament barbed ends to the plasma membrane using membrane-binding domains.

Filopodium geometry is unfavorable for the delivery of building components to the tip by simple diffusion (Mogilner

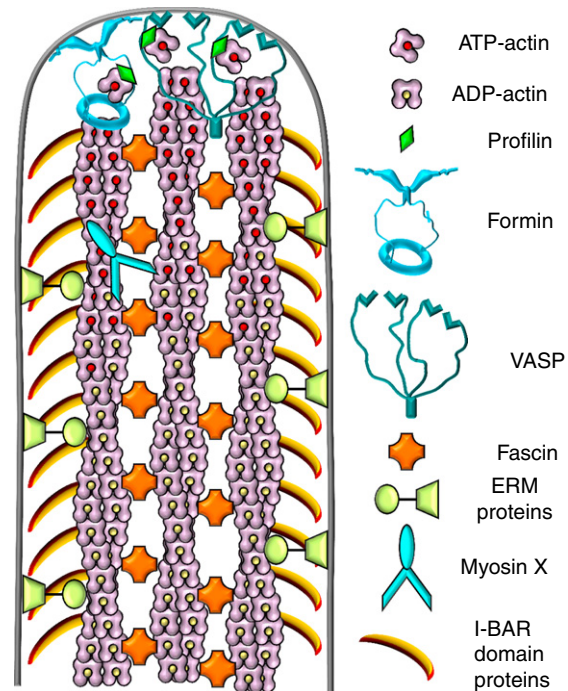


Figure 6 Schematics of the filopodial molecular machinery. Actin filaments are oriented with barbed ends toward the filopodial tip (top). They are bundled by fascin and laterally attached to the plasma membrane by ERM proteins. Formin and/or Ena/VASP proteins associate with the barbed ends, promote their elongation, and anchor them to the membrane at the filopodial tip. Myosin X travels along actin filaments to the filopodial tip. I-BAR domain-containing proteins, such as IRSp53, help maintaining the tubular shape of the plasma membrane. Modified from Svitkina, T., 2013. Actin organization. In: Lennarz, W.J., Daniel Lane, M. (Eds.), *Encyclopedia of Biological Chemistry*. Waltham, MA: Academic Press, pp. 27–35.

and Rubinstein, 2005). Unconventional myosin X, which plays an important role in filopodia elongation (Berg and Cheney, 2002) is believed to use its motor activity to deliver cytoplasmic components, such as Ena/VASP and surface receptors, to filopodial tips. Enrichment of myosin X at filopodial tips suggests that it may have additional roles in promoting filopodium protrusion.

Long actin filaments individually are not efficient at pushing, as they tend to buckle. To produce efficient pushing, filopodial filaments are immobilized by cross-linking along their length, which increases their collective stiffness. The major actin filament cross-linker in filopodia is fascin (Vignjevic *et al.*, 2006). It was first isolated from sea urchin eggs (Kane, 1975) and subsequently shown to bundle actin filaments in filopodia of sea urchin coelomocytes and mammalian cells. Fascin is a bivalent monomeric protein (Sedeh *et al.*, 2010) that *in vitro* induces needle-shaped bundles of actin filaments characterized by high rigidity. In cells, fascin allows actin filaments to overcome membrane resistance and avoid buckling during filopodial protrusion (Vignjevic *et al.*, 2006). Fascin undergoes fast turnover within filopodial bundles, which may help to release stresses and torques within the bundle as it assembles. Other monomeric actin cross-linkers,

such as fimbrin, espin, and villin, form parallel bundles in other cylindrical protrusions, such as microvilli and stereocilia.

Depolymerization of filopodial filaments is accelerated by their severing. Two actin filament severing proteins, ADF/cofilin (Breitsprecher *et al.*, 2011; Gehler *et al.*, 2004) and gelsolin (Lu *et al.*, 1997), play a role in this process. Although ADF/cofilin is absent in most filopodia (Svitkina and Borisy, 1999), it accumulates in stationary and retracting filopodia and promotes their disassembly (Breitsprecher *et al.*, 2011). Conventional myosin II, which is typically involved in contractile cellular activities, also contributes to generation of additional filament ends by breaking proximal regions of filopodial bundles (Medeiros *et al.*, 2006).

Many proteins besides Ena/VASPs, formins, and myosin X are associated with filopodial tips, where they form a morphologically recognizable tip complex (Svitkina *et al.*, 2003). Although functions of all these proteins are not yet fully established, talin (DePasquale and Izzard, 1991) and integrins (Wu *et al.*, 1996) may be needed to initiate adhesions with a suitable surface, whereas N-WASP, CR16 (Ho *et al.*, 2001), and Abi proteins (Stradal *et al.*, 2001), as well as some signaling molecules, like Vav (Kranewitter *et al.*, 2001), may function to trigger cellular responses when a filopodium finds an appropriate target.

Some proteins are found in proximal filopodial regions. A cross-linking protein α -actinin is found at the base of filopodia (Vignjevic *et al.*, 2006), where it may integrate filopodial bundles into the cytoplasmic actin cytoskeleton. Myosin II also frequently interacts with the cytoplasmic root of filopodial bundles (Bridgman and Dailey, 1989). By pulling on filopodial actin filaments, myosin II can contribute to their retrograde flow and reinforce adhesions for efficient protrusion.

Lateral binding of filopodial bundles to the plasma membrane helps to maintain the shape of filopodia. Proteins of ezrin–radixin–moesin (ERM) family, which can bind both actin filaments and the membrane, are likely candidates for this role (Bretscher *et al.*, 2002). ERM proteins localize to filopodia and promote their formation in some systems.

Membrane deformation

Changes in the membrane geometry accompanying filopodial protrusion are powered in part by the membrane-binding and deforming proteins, not only by cytoskeleton activity. Most of these proteins contain one of two related protein domains, I-BAR (Zhao *et al.*, 2011) or IF-BAR (Carlson *et al.*, 2011), which bind membrane lipids via their convex surface and cause negative membrane curvature (away from the cytoplasm). When expressed in cells, an isolated I-BAR domain from IRSp53 induces filopodia-like protrusions (Yamagishi *et al.*, 2004). These protrusions initially represent empty membrane tubes, but are subsequently filled with actin filaments (Yang *et al.*, 2009). In a physiological situation, IRSp53 upon activation is thought to deform the plasma membrane using its I-BAR domain and recruit actin polymerization machinery using its other protein–protein interaction domains, thus coordinating two molecular machineries in filopodia (Disanza *et al.*, 2013).

Lamellipodia

Definition

Flat undulating cellular protrusions that appeared to control the direction of cell motility were first observed by Abercrombie and colleagues, when they studied locomotion of cultured fibroblasts. By analogy with filopodia, they named these flat dynamic protrusions ‘lamellipodia,’ and reserved the term ‘lamella’ to refer to a broad flattened sheet of anterior cytoplasm that carries lamellipodia at its tip (Abercrombie *et al.*, 1970).

Behavior and Functions

Lamellipodia are found in virtually all migrating cells and serve as the major cellular engine to propel the leading edge forward. Lamellipodia can generate much greater protrusive forces than filopodia (Cojoc *et al.*, 2007). They also function as navigation devices for guiding a cell around obstacles, sense soluble guidance cues and probe chemical and mechanical properties of the substratum. Although lamellipodia have been mostly studied using cells cultured on a flat solid surface, they are also used for cell migration in 3D environment, but their morphology changes depending on the geometry of available adhesive surface and a balance of intracellular signaling pathways (Petrie and Yamada, 2012).

Lamellipodia typically display cyclic behavior with alternating phases of protrusion and retraction (Abercrombie *et al.*, 1970). The amplitude of these cycles broadly varies among cells. During protrusion phases, the lamellipodia typically lie flat on the substratum, whereas during retraction they often make ruffles, structures formed by the uplifting of flat protrusions. Protrusion–retraction cycles are not tightly coordinated along the leading edge of a cell and show no correlation among leading edge regions separated by more than 5–6 μm .

During migration, lamellipodia need traction to prevent slippage during protrusion. Traction for lamellipodial protrusion is provided by small and dynamic adhesion structures, typically called focal complexes or nascent adhesions (Beningo *et al.*, 2001). These structures mediate linkages between the actin cytoskeleton and the extracellular matrix through transmembrane adhesion receptors that typically belong to the integrin family. Nascent adhesions are formed at the interface between the lamellipodium and the adhesive substrate (Choi *et al.*, 2008) and can subsequently develop into large and stable mature focal adhesions that usually reside behind the lamellipodium in the cell lamella and cell body. Adhesion formation, in general, is a force-dependent process (Burrage and Wittchen, 2013). Formation of nascent adhesions requires fairly small forces, which can derive from myosin II-dependent contraction, actin filament cross-linking (Choi *et al.*, 2008), or a drag applied by the actin retrograde flow (Alexandrova *et al.*, 2008).

Lamellipodia are involved not only in cell migration, but also in formation of cell–cell junctions, such as adherens junctions in epithelial (McNeill *et al.*, 1993) and endothelial cells (Hoelzle and Svitkina, 2012). Flat protrusions homologous to lamellipodia and ruffles also function in phagocytosis (Groves *et al.*, 2008) and micropinocytosis (Kerr and Teasdale,

2009), where they protrude to enclose a solid particle or a portion of extracellular fluid.

Structure

Lamellipodia formed at the leading edge of cells migrating or spreading in 2D culture are best characterized structurally and dynamically. Early electron microscopic studies using thin sections of resin-embedded cells suggested that actin filaments in lamellipodia have distinct organization compared with contractile actin bundles in the same cell (Ludueno and Wesells, 1973). However, preservation of the cytoskeleton was not adequate at that time to define specific details. Subsequently, electron microscopy of negatively stained lamellipodia revealed a network of long diagonal actin filaments oriented with their barbed ends toward the leading edge (Small *et al.*, 1978). Further scrutiny of the actin network architecture by rotary shadowing electron microscopy revealed multiple branched actin filaments in lamellipodia (Svitkina *et al.*, 1997; Svitkina and Borisy, 1999). At branch points, a pointed end of one filament was linked to the side of another filament with the formation of approximately 70° angle between barbed ends of two filaments (Figure 7).

The specific parameters of branched (or dendritic) actin networks in lamellipodia, such as density, length, and orientation of actin filaments or frequency and distribution of branches, vary among cell types and a phase of the protrusion. For example, nascent lamellipodia appear to contain a greater density of branches and shorter individual filaments compared with established lamellipodia undergoing steady state

protrusion (Svitkina and Borisy, 1999; Vinzenz *et al.*, 2012). On the other hand, lamellipodia of neuronal growth cones contain relatively few branches and many long filaments (Korobova and Svitkina, 2008). These parameters of the actin network architecture correlate with the protrusive behavior of the lamellipodium. Thus, lamellipodia containing longer and less branched filaments protrude faster, but are more prone to ruffling and retraction, whereas lamellipodia with shorter and more frequently branched filaments are more persistent, although protrude slower (Bear *et al.*, 2002).

Dendritic actin networks are not restricted to lamellipodia and related membrane protrusions, but also provide pushing force for other cellular activities, such as formation of an endocytic vesicle during receptor-mediated endocytosis (Mooren *et al.*, 2012), or motility and biogenesis of intracellular membrane organelles (Burianek and Soderling, 2013). Certain intracellular pathogens, such as *Listeria monocytogenes* and *Shigella flexneri*, exploit the dendritic nucleation machinery of a host cell to spread from cell to cell while avoiding immune surveillance in the extracellular environment. These and other pathogens use a variety of ways to stimulate assembly of a branched network in the form of elongated 'comet tails,' which propel them through the cytoplasm and into a cellular protrusion, which then can be phagocytosed by another cell (Lambrechts *et al.*, 2008).

Actin Dynamics

Actin dynamics during lamellipodial protrusion displays typical features of treadmilling behavior, as revealed by tracking fiducial marks in the lamellipodial actin network

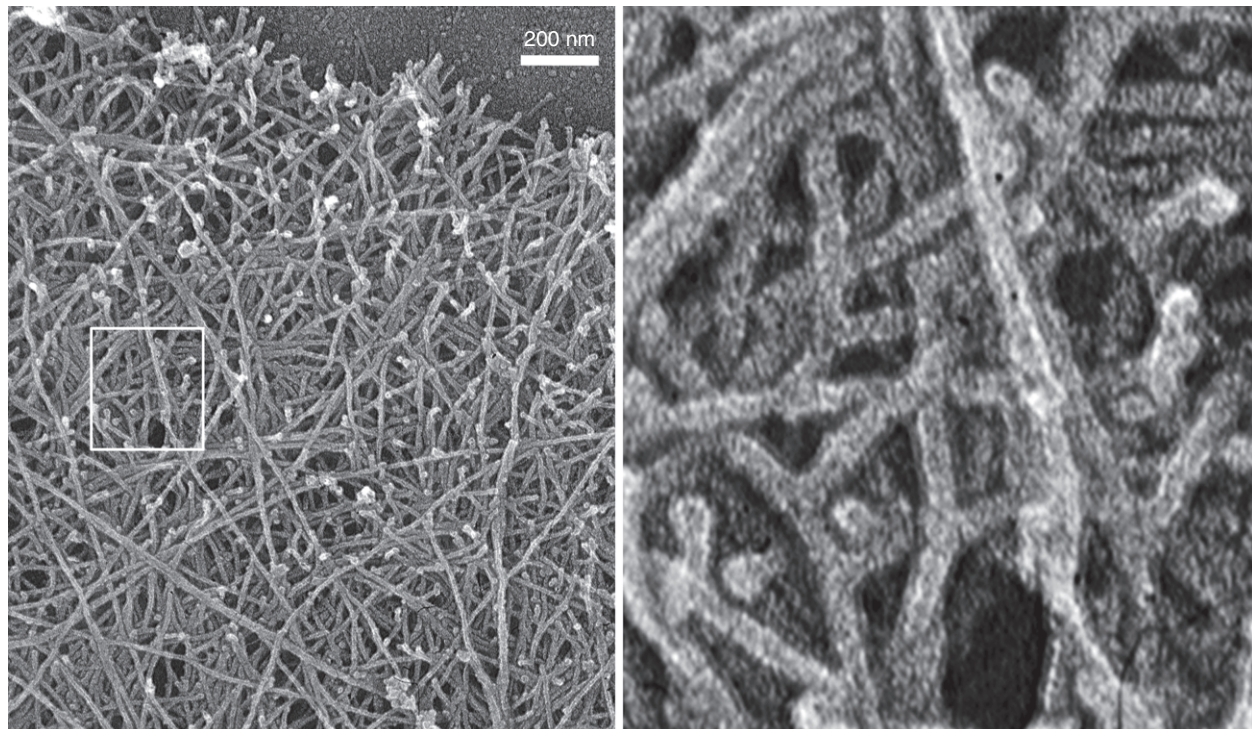


Figure 7 Actin filament organization in a lamellipodium at the leading edge of a fish epidermal keratocyte. Boxed region is enlarged in the right panel and shows branched actin filaments. Modified from Svitkina, T., 2013. Actin organization. In: Lennarz, W.J., Daniel Lane, M. (Eds.), *Encyclopedia of Biological Chemistry*. Waltham, MA: Academic Press, pp. 27–35.

(Wang, 1985). Thus, actin subunits are added at the leading edge near the plasma membrane, move away from the membrane, and dissipate from the network by the time they reach the lamellipodium base. While polymerizing at the front, the lamellipodial network also undergoes the retrograde flow. As in the case of filopodia, absolute and relative rates of protrusion and retrograde flow vary significantly. The balance between two processes can be controlled in part by the strength of the linkages between the substrate and the lamellipodial actin network (Theriot and Mitchison, 1992).

Despite clear signatures of treadmilling behavior, some parameters of actin dynamics in lamellipodia are not consistent with a simple actin filament treadmilling model, which states that individual actin filaments continuously elongate at their distal barbed ends and depolymerize from their proximal pointed ends. Quantitative analyses of the rate of actin depolymerization in lamellipodia using photoactivated actin marks revealed an extremely fast actin turnover, which was inconsistent with disassembly of a limited number of long filaments from their pointed ends. Also, actin depolymerization occurred evenly throughout the lamellipodium (Theriot and Mitchison, 1991), not only at the base, like in filopodia (Mallavarapu and Mitchison, 1999).

A combination of structural, biochemical, and kinetic data led to a revised model of lamellipodia protrusion, termed the

dendritic nucleation model (Mullins *et al.*, 1998; Figure 8). It posits that the actin filaments in lamellipodia are constantly nucleated as branches on preexisting 'mother' filaments. These 'daughter' filaments elongate and produce force. After a period of elongation, their growth is terminated by capping, while new filaments are nucleated to maintain protrusion. Disassembly of the network occurs through a combination of debranching and severing of actin filaments followed by depolymerization of filament fragments. Thus, the branched network in lamellipodia undergoes treadmilling as a whole by assembling at the front and disassembling throughout its body. In contrast, an individual filament in the network does not treadmill, but is 'born' at a branch point, grows at the barbed end, becomes capped and later 'dies' by depolymerization (Pollard and Borisy, 2003).

The branching network has many advantages for pushing force generation. Newly nucleated filaments are more efficient at pushing, because they are short and stiff. Furthermore, being anchored to the mother filament and through it to the entire lamellipodial network, daughter filaments can immediately do useful work, rather than slip backward in the absence of traction. Repetitive branching allows the network to easily expand or reorient itself by adjusting the rates and sites of nucleation and capping. An angled orientation of pushing actin filaments relative to the membrane is also useful, as it makes it easier to

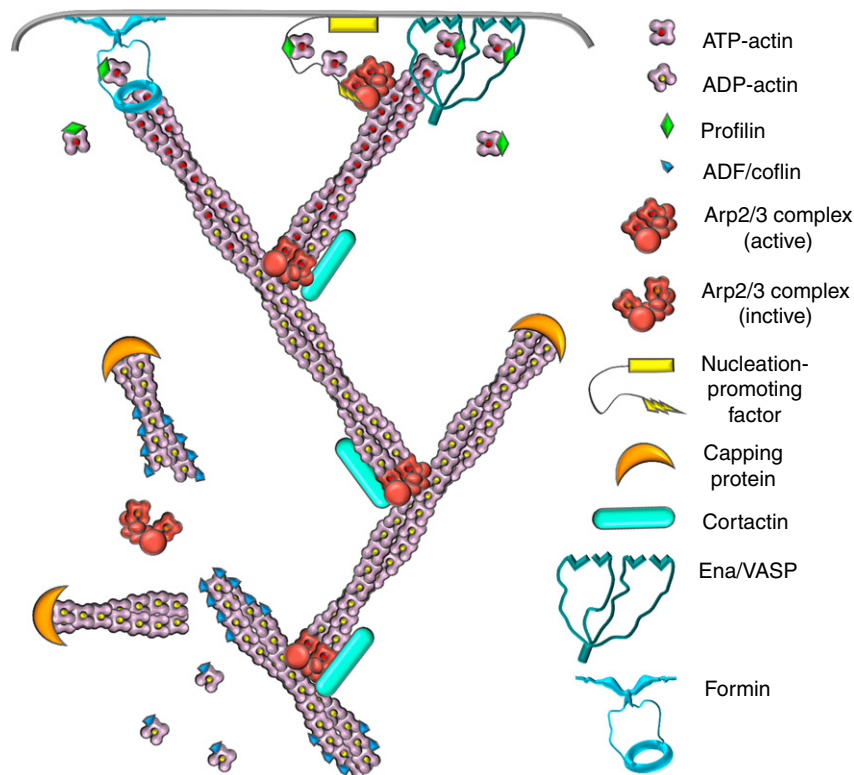


Figure 8 Schematics of the lamellipodial molecular machinery. The Arp2/3 complex is cooperatively activated by a membrane-targeted nucleation-promoting factor and a pre-existing actin filament. Upon activation, it nucleates a new actin filament at the side of the pre-existing filament and remains associated with the branch point. Branch points are further stabilized by cortactin. The nascent daughter filament elongates with its barbed end oriented toward the membrane. Its elongation is promoted by formins and/or Ena/VASP proteins. After a period of elongation, the barbed end is capped by capping protein and lags behind the advancing leading edge. Disassembly at the rear of the lamellipodial network occurs through dissociation of branches and ADF/cofilin-mediated severing. Modified from Svitkina, T., 2013. Actin organization. In: Lennarz, W.J., Daniel Lane, M. (Eds.), *Encyclopedia of Biological Chemistry*. Waltham, MA: Academic Press, pp. 27–35.

insert an actin monomer between the growing barbed end and the membrane (Mogilner and Oster, 1996).

Molecular Machinery

Assembly of branched networks of actin filaments is controlled by a large number of essential and accessory molecules (Figure 8). Functions of many of these proteins in motility have been revealed by analyzing molecular requirements for *Listeria* motility. A minimal set of key proteins involved in this process have been identified by reconstituting motility of *Listeria* or *Shigella* in a solution of mixed purified proteins (Loisel *et al.*, 1999). A set of essential proteins besides actin itself consisted of the Arp2/3 complex, capping protein, and ADF/cofilin.

The Arp2/3 complex (Machesky *et al.*, 1994; Welch *et al.*, 1997) is a key part of the dendritic nucleation machinery. This heteroheptameric complex nucleates a daughter filament from the side of a mother filament. After nucleation, Arp2/3 complex remains at the branch point, where it links the mother and daughter filaments at the defined angle of 70° and caps the pointed end of the daughter filament to prevent its premature depolymerization (Mullins *et al.*, 1998). Two actin-related proteins within the Arp2/3 complex, Arp2 and Arp3, mimic the barbed end of an actin filament, which allows them to promote nucleation and pointed end capping. Other subunits of the Arp2/3 complex mediate binding to the mother filament. The Arp2/3 complex is intrinsically inactive and has Arp2 and Arp3 separated from each other. To acquire nucleating capability, the Arp2/3 complex needs to be activated by a nucleation-promoting factor that acts in cooperation with the mother filament (Machesky *et al.*, 1999). Following activation, Arp2 and Arp3 within the complex acquire a mutual arrangement mimicking the barbed end. Most important nucleation-promoting factor belong to WASP family (Burianek and Soderling, 2013; Campellone and Welch, 2010). In lamellipodia, Arp2/3 complex is predominantly activated by WAVE proteins, which exist in the form of a pentameric complex (Eden *et al.*, 2002). Other nucleation-promoting factors function at distinct subcellular locations, such as various membrane organelles (Campellone and Welch, 2010).

Heterodimeric capping protein consisting of α and β subunits tightly binds barbed ends of actin filaments preventing their polymerization and depolymerization. Such capping activity is necessary to terminate excessive barbed ends constantly generated by the Arp2/3 complex and thus maintain the steady state actin dynamics in the branched network (Mejillano *et al.*, 2004). Capping protein is the predominant barbed end capper in lamellipodia (Schafer *et al.*, 1998), but the other barbed end capping proteins, such as Eps8, gelsolin, and Aip1, are also present in lamellipodia and may contribute to the dynamics of branched networks (Zigmond, 2004).

Under steady state conditions, actin polymerization should be balanced by depolymerization. Disassembly of branched actin networks is primarily carried out by ADF/cofilin proteins (Bamburg *et al.*, 1980; Bernstein and Bamburg, 2010), which sever actin filaments throughout the lamellipodium, except for the very front, where actin filaments are enriched in ATP-loaded subunits and therefore poorly bind ADF/cofilin (Svitkina and Borisy, 1999). ADF/cofilin typically functions together

with Aip1, which activates cofilin's severing activity and also caps the newly formed barbed end to prevent its growth or annealing (Okada *et al.*, 1999). However, in some cases, ADF/cofilin appears to sever actin filaments without capping, so that newly formed barbed ends promote polymerization, rather than depolymerization (Zebda *et al.*, 2000). Srv2/CAP proteins also enhance ADF/cofilin-mediated severing (Chaudhry *et al.*, 2013). ADF/cofilin can also debranch daughter filaments by changing conformation of the mother filament (Blanchoin *et al.*, 2000). Debranching can also be induced by cofilin-related GMF proteins (Gandhi *et al.*, 2010) and by coronins (Cai *et al.*, 2008); these proteins act through binding to the Arp2/3 complex.

Many other proteins improve the performance of the dendritic nucleation machinery in the cell. Actin elongation is assisted by profilin (Syriani *et al.*, 2008), formins (Yang *et al.*, 2007) and Ena/VASP proteins (Bear *et al.*, 2002). Actin-binding protein cortactin promotes binding of the Arp2/3 complex to actin filaments and thus stabilizes branches (Weaver *et al.*, 2001). Cross-linkers filamin A (Flanagan *et al.*, 2001) and α -actinin (Hamill *et al.*, 2013) likely consolidate the whole network. Numerous upstream regulators modulate functions of various components of the lamellipodial molecular machinery and thus control behavior of lamellipodia. For example, high activity of actin filament elongation factors, such as Ena/VASP (Bear *et al.*, 2002) and formin (Yang *et al.*, 2007), produces numerous long actin filaments within the branched actin network, which can be bundled with the formation of filopodia (Svitkina *et al.*, 2003), or buckled under the influence of membrane tension and produce ruffles (Bear *et al.*, 2002).

See also: Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell Migration. Cytoskeleton and Motors: Cytoskeletal Components: Actin Assembly Dynamics and Its Regulation in Motile and Morphogenetic Processes; Myosins. Interorganellar Communication: Components: BAR Domains and BAR Domain Superfamily Proteins. Intracellular Infectiology: Cell Processes: Macropinocytosis; Phagocytosis

References

- Abercrombie, M., Heaysman, J.E., Pegrum, S.M., 1970. The locomotion of fibroblasts in culture. II. Ruffling. *Experimental Cell Research* 60, 437–444.
- Alexandrova, A.Y., Arnold, K., Schaub, S., *et al.*, 2008. Comparative dynamics of retrograde actin flow and focal adhesions: Formation of nascent adhesions triggers transition from fast to slow flow. *PLoS One* 3, e3234.
- Bamburg, J.R., Harris, H.E., Weeds, A.G., 1980. Partial purification and characterization of an actin depolymerizing factor from brain. *FEBS Letters* 121, 178–182.
- Bear, J.E., Svitkina, T.M., Krause, M., *et al.*, 2002. Antagonism between Ena/VASP proteins and actin filament capping regulates fibroblast motility. *Cell* 109, 509–521.
- Beningo, K.A., Dembo, M., Kaverina, I., Small, J.V., Wang, Y.L., 2001. Nascent focal adhesions are responsible for the generation of strong propulsive forces in migrating fibroblasts. *Journal of Cell Biology* 153, 881–888.
- Berg, J.S., Cheney, R.E., 2002. Myosin-X is an unconventional myosin that undergoes intrafilopodial motility. *Nature Cell Biology* 18, 246–250.

- Bernstein, B.W., Bamberg, J.R., 2010. ADF/cofilin: A functional node in cell biology. *Trends in Cell Biology* 20, 187–195.
- Blanchoin, L., Pollard, T.D., Mullins, R.D., 2000. Interactions of ADF/cofilin, Arp2/3 complex, capping protein and profilin in remodeling of branched actin filament networks. *Current Biology* 10, 1273–1282.
- Bosch, M., Le, K.H., Bugyi, B., *et al.*, 2007. Analysis of the function of Spire in actin assembly and its synergy with formin and profilin. *Molecular Cell* 28, 555–568.
- Bourne, J.N., Harris, K.M., 2008. Balancing structure and function at hippocampal dendritic spines. *Annual Review of Neuroscience* 31, 47–67.
- Breitsprecher, D., Koestler, S.A., Chizhov, I., *et al.*, 2011. Cofilin cooperates with fascin to disassemble filopodial actin filaments. *Journal of Cell Science* 124, 3305–3318.
- Bretscher, A., Edwards, K., Fehon, R.G., 2002. ERM proteins and merlin: Integrators at the cell cortex. *Nature Reviews Molecular Cell Biology* 3, 586–599.
- Bridgman, P.C., Dailey, M.E., 1989. The organization of myosin and actin in rapid frozen nerve growth cones. *Journal of Cell Biology* 108, 95–109.
- Burianek, L.E., Soderling, S.H., 2013. Under lock and key: Spatiotemporal regulation of WASP family proteins coordinates separate dynamic cellular processes. *Seminars in Cell and Developmental Biology* 24, 258–266.
- Burridge, K., Wittchen, E.S., 2013. The tension mounts: Stress fibers as force-generating mechanotransducers. *Journal of Cell Biology* 200, 9–19.
- Cai, L., Makhov, A.M., Schafer, D.A., Bear, J.E., 2008. Coronin 1B antagonizes cortactin and remodels Arp2/3-containing actin branches in lamellipodia. *Cell* 134, 828–842.
- Campellone, K.G., Welch, M.D., 2010. A nucleator arms race: Cellular control of actin assembly. *Nature Reviews Molecular Cell Biology* 11, 237–251.
- Carlier, M.F., Laurent, V., Santolini, J., Melki, R., Didry, D., 1997. Actin depolymerizing factor (ADF/cofilin) enhances the rate of filament turnover: Implication in actin-based motility. *Journal of Cell Biology* 136, 1307–1322.
- Carlson, B.R., Lloyd, K.E., Kruszewski, A., *et al.*, 2011. WRP/srGAP3 facilitates the initiation of spine development by an inverse F-BAR domain, and its loss impairs long-term memory. *Journal of Neuroscience* 31, 2447–2460.
- Chaudhry, F., Breitsprecher, D., Little, K., *et al.*, 2013. Srv2/cyclase-associated protein forms hexameric shurikens that directly catalyze actin filament severing by cofilin. *Molecular Biology of the Cell* 24, 31–41.
- Choi, C.K., Vicente-Manzanares, M., Zareno, J., *et al.*, 2008. Actin and alpha-actinin orchestrate the assembly and maturation of nascent adhesions in a myosin II motor-independent manner. *Nature Cell Biology* 10, 1039–1050.
- Cojoc, D., Difato, F., Ferrari, E., *et al.*, 2007. Properties of the force exerted by filopodia and lamellipodia and the involvement of cytoskeletal components. *PLoS One* 2, e1072.
- De Smet, F., Segura, I., de Bock, K., Hohensinner, P.J., Carmeliet, P., 2009. Mechanisms of vessel branching: Filopodia on endothelial tip cells lead the way. *Arteriosclerosis, Thrombosis, and Vascular Biology* 29, 639–649.
- DePasquale, J.A., Izzard, C.S., 1991. Accumulation of talin in nodes at the edge of the lamellipodium and separate incorporation into adhesion plaques at focal contacts in fibroblasts. *Journal of Cell Biology* 113, 1351–1359.
- Disanza, A., Bisi, S., Winterhoff, M., *et al.*, 2013. CDC42 switches IRSp53 from inhibition of actin growth to elongation by clustering of VASP. *EMBO Journal* 32, 2735–2750.
- Dominguez, R., 2010. Structural insights into de novo actin polymerization. *Current Opinion in Structural Biology* 20, 217–225.
- Drees, F., Gertler, F.B., 2008. Ena/VASP: Proteins at the tip of the nervous system. *Current Opinion in Neurobiology* 18, 53–59.
- Eden, S., Rohatgi, R., Podtelejnikov, A.V., Mann, M., Kirschner, M.W., 2002. Mechanism of regulation of WAVE1-induced actin nucleation by Rac1 and Nck. *Nature* 418, 790–793.
- Flanagan, L.A., Chou, J., Falet, H., *et al.*, 2001. Filamin A, the Arp2/3 complex, and the morphology and function of cortical actin filaments in human melanoma cells. *Journal of Cell Biology* 155, 511–517.
- Gandhi, M., Smith, B.A., Bovellan, M., *et al.*, 2010. GMF is a cofilin homolog that binds Arp2/3 complex to stimulate filament debranching and inhibit actin nucleation. *Current Biology* 20, 861–867.
- Gehler, S., Shaw, A.E., Samiere, P.D., Bamberg, J.R., Letourneau, P.C., 2004. Brain-derived neurotrophic factor regulation of retinal growth cone filopodial dynamics is mediated through actin depolymerizing factor/cofilin. *Journal of Neuroscience* 24, 10741–10749.
- Groves, E., Dart, A.E., Covarelli, V., Caron, E., 2008. Molecular mechanisms of phagocytic uptake in mammalian cells. *Cellular and Molecular Life Sciences* 65, 1957–1976.
- Hamill, K.J., Hopkinson, S.B., Skalli, O., Jones, J.C., 2013. Actinin-4 in keratinocytes regulates motility via an effect on lamellipodia stability and matrix adhesions. *FASEB Journal* 27, 546–556.
- Harrison, R.G., 1907. Observations on the living developing nerve fiber. *Anatomical Record* 1, 116–118.
- Ho, H.Y., Rohatgi, R., MA, L., Kirschner, M.W., 2001. CR16 forms a complex with N-WASP in brain and is a novel member of a conserved proline-rich actin-binding protein family. *Proceedings of the National Academy of Sciences of the United States of America* 98, 11306–11311.
- Hoelzel, M.K., Svitkina, T., 2012. The cytoskeletal mechanisms of cell-cell junction formation in endothelial cells. *Molecular Biology of the Cell* 23, 310–323.
- Homem, C.C., Peifer, M., 2009. Exploring the roles of diaphanous and enabled activity in shaping the balance between filopodia and lamellipodia. *Molecular Biology of the Cell* 20, 5138–5155.
- Kane, R.E., 1975. Preparation and purification of polymerized actin from sea urchin egg extracts. *Journal of Cell Biology* 66, 305–315.
- Kerr, M.C., Teasdale, R.D., 2009. Defining macropinocytosis. *Traffic* 10, 364–371.
- Korobova, F., Svitkina, T., 2008. Arp2/3 complex is important for filopodia formation, growth cone motility, and neurogenesis in neuronal cells. *Molecular Biology of the Cell* 19, 1561–1574.
- Korobova, F., Svitkina, T., 2010. Molecular architecture of synaptic actin cytoskeleton in hippocampal neurons reveals a mechanism of dendritic spine morphogenesis. *Molecular Biology of the Cell* 21, 165–176.
- Kranewitter, W.J., Danninger, C., Gimona, M., 2001. GEF at work: Vav in protruding filopodia. *Cell Motility and the Cytoskeleton* 49, 154–160.
- Lambrechts, A., Gevaert, K., Cossart, P., Vandekerckhove, J., van Troys, M., 2008. Listeria comet tails: The actin-based motility machinery at work. *Trends in Cell Biology* 18, 220–227.
- Lebrand, C., Dent, E.W., Strasser, G.A., *et al.*, 2004. Critical role of Ena/VASP proteins for filopodia formation in neurons and in function downstream of netrin-1. *Neuron* 42, 37–49.
- Loisel, T.P., Boujemaa, R., Pantaloni, D., Carlier, M.F., 1999. Reconstitution of actin-based motility of Listeria and Shigella using pure proteins. *Nature* 401, 613–616.
- Lu, M., Witke, W., Kwiatkowski, D.J., Kosik, K.S., 1997. Delayed retraction of filopodia in gelsolin null mice. *Journal of Cell Biology* 138, 1279–1287.
- Ludena, M.A., Wessells, N.K., 1973. Cell locomotion, nerve elongation, and microfilaments. *Developmental Biology* 30, 427–440.
- Machesky, L.M., Atkinson, S.J., Ampe, C., Vandekerckhove, J., Pollard, T.D., 1994. Purification of a cortical complex containing two unconventional actins from Acanthamoeba by affinity chromatography on profilin-agarose. *Journal of Cell Biology* 127, 107–115.
- Machesky, L.M., Mullins, R.D., Higgs, H.N., *et al.*, 1999. Scar, a WASP-related protein, activates nucleation of actin filaments by the Arp2/3 complex. *Proceedings of the National Academy of Sciences of the United States of America* 96, 3739–3744.
- Mallavarapu, A., Mitchison, T., 1999. Regulated actin cytoskeleton assembly at filopodium tips controls their extension and retraction. *Journal of Cell Biology* 146, 1097–1106.
- McNeill, H., Ryan, T.A., Smith, S.J., Nelson, W.J., 1993. Spatial and temporal dissection of immediate and early events following cadherin-mediated epithelial cell adhesion. *Journal of Cell Biology* 120, 1217–1226.
- Medeiros, N.A., Burnette, D.T., Forscher, P., 2006. Myosin II functions in actin-bundle turnover in neuronal growth cones. *Nature Cell Biology* 8, 215–226.
- Mejillano, M.R., Kojima, S., Applewhite, D.A., *et al.*, 2004. Lamellipodial versus filopodial mode of the actin nanomachinery: pivotal role of the filament barbed end. *Cell* 118, 363–373.
- Mogilner, A., Oster, G., 1996. Cell motility driven by actin polymerization. *Biophysical Journal* 71, 3030–3045.
- Mogilner, A., Rubinstein, B., 2005. The physics of filopodial protrusion. *Biophysical Journal* 89, 782–795.
- Moore, O.L., Galletta, B.J., Cooper, J.A., 2012. Roles for actin assembly in endocytosis. *Annual Review of Biochemistry* 81, 661–686.
- Moosker, M.S., Tilney, L.G., 1975. Organization of an actin filament-membrane complex. Filament polarity and membrane attachment in the microvilli of intestinal epithelial cells. *The Journal of Cell Biology* 67, 725–743.
- Mullins, R.D., Heuser, J.A., Pollard, T.D., 1998. The interaction of Arp2/3 complex with actin: Nucleation, high affinity pointed end capping, and formation of branching networks of filaments. *Proceedings of the National Academy of Sciences of the United States of America* 95, 6181–6186.
- Okada, K., Bartolini, F., Deaconescu, A.M., *et al.*, 2010. Adenomatous polyposis coli protein nucleates actin assembly and synergizes with the formin mDia1. *Journal of Cell Biology* 189, 1087–1096.

- Okada, K., Obinata, T., Abe, H., 1999. XAIP1: A Xenopus homologue of yeast actin interacting protein 1 (AIP1), which induces disassembly of actin filaments cooperatively with ADF/cofilin family proteins. *Journal of Cell Science* 112, 1553–1565.
- Petrie, R.J., Yamada, K.M., 2012. At the leading edge of three-dimensional cell migration. *Journal of Cell Science* 125, 5917–5926.
- Pollard, T.D., Borisy, G.G., 2003. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* 112, 453–465.
- Pollard, T.D., Mooseker, M.S., 1981. Direct measurement of actin polymerization rate constants by electron microscopy of actin filaments nucleated by isolated microvillus cores. *Journal of Cell Biology* 88, 654–659.
- Pruyne, D., Evangelista, M., Yang, C., *et al.*, 2002. Role of formins in actin assembly: Nucleation and barbed-end association. *Science* 297, 612–615.
- Quinlan, M.E., Hilgert, S., Bedrossian, A., Mullins, R.D., Kerkhoff, E., 2007. Regulatory interactions between two actin nucleators, Spire and Capping protein. *Journal of Cell Biology* 179, 117–128.
- Romero, S., Grompone, G., Carayol, N., *et al.*, 2011. ATP-mediated Erk1/2 activation stimulates bacterial capture by filopodia, which precedes *Shigella* invasion of epithelial cells. *Cell Host & Microbe* 9, 508–519.
- Sagot, I., Rodal, A.A., Moseley, J., Goode, B.L., Pellman, D., 2002. An actin nucleation mechanism mediated by Bni1 and profilin. *Nature Cell Biology* 4, 626–631.
- Schafer, D.A., Welch, M.D., MaChesky, L.M., *et al.*, 1998. Visualization and molecular analysis of actin assembly in living cells. *Journal of Cell Biology* 143, 1919–1930.
- Sedeh, R.S., Fedorov, A.A., Fedorov, E.V., *et al.*, 2010. Structure, evolutionary conservation, and conformational dynamics of Homo sapiens fascin-1, an F-actin crosslinking protein. *Journal of Molecular Biology* 400, 589–604.
- Small, J.V., Isenberg, G., Celis, J.E., 1978. Polarity of actin at the leading edge of cultured cells. *Nature* 272, 638–639.
- Stradal, T., Courtney, K.D., Rottner, K., *et al.*, 2001. The Abl interactor proteins localize to sites of actin polymerization at the tips of lamellipodia and filopodia. *Current Biology* 11, 891–895.
- Suter, D.M., Errante, L.D., Belotserkovsky, V., Forscher, P., 1998. The Ig superfamily cell adhesion molecule, apCAM, mediates growth cone steering by substrate-cytoskeletal coupling. *Journal of Cell Biology* 141, 227–240.
- Svitkina, T.M., Borisy, G.G., 1999. Arp2/3 complex and actin depolymerizing factor/cofilin in dendritic organization and treadmilling of actin filament array in lamellipodia. *Journal of Cell Biology* 145, 1009–1026.
- Svitkina, T.M., Bulanova, E.A., Chaga, O.Y., *et al.*, 2003. Mechanism of filopodia initiation by reorganization of a dendritic network. *Journal of Cell Biology* 160, 409–421.
- Svitkina, T.M., Verkhovsky, A.B., McQuade, K.M., Borisy, G.G., 1997. Analysis of the actin-myosin II system in fish epidermal keratocytes: Mechanism of cell body translocation. *Journal of Cell Biology* 139, 397–415.
- Syriani, E., Gomez-Cabrero, A., Bosch, M., *et al.*, 2008. Profilin induces lamellipodia by growth factor-independent mechanism. *Faseb J* 22, 1581–1596.
- Theriot, J.A., Mitchison, T.J., 1991. Actin microfilament dynamics in locomoting cells. *Nature* 352, 126–131.
- Theriot, J.A., Mitchison, T.J., 1992. Comparison of actin and cell surface dynamics in motile fibroblasts. *Journal of Cell Biology* 119, 367–377.
- Tilney, L.G., 1975. The role of actin in nonmuscle cell motility. *Society of General Physiologists Series* 30, 339–388.
- Tilney, L.G., Bonder, E.M., Coluccio, L.M., Mooseker, M.S., 1983. Actin from Thyone sperm assembles on only one end of an actin filament: A behavior regulated by profilin. *Journal of Cell Biology* 97, 112–124.
- Tilney, L.G., Derosier, D.J., Mulroy, M.J., 1980. The organization of actin filaments in the stereocilia of cochlear hair cells. *Journal of Cell Biology* 86, 244–259.
- Tyska, M.J., Mooseker, M.S., 2002. MYO1A (brush border myosin I) dynamics in the brush border of LLC-PK1-CL4 cells. *Biophysical Journal* 82, 1869–1883.
- Vasioukhin, V., Bauer, C., Yin, M., Fuchs, E., 2000. Directed actin polymerization is the driving force for epithelial cell-cell adhesion. *Cell* 100, 209–219.
- Vignjevic, D., Kojima, S., Aratyn, Y., *et al.*, 2006. Role of fascin in filopodial protrusion. *Journal of Cell Biology* 174, 863–875.
- Vinzenz, M., Nemethova, M., Schur, F., *et al.*, 2012. Actin branching in the initiation and maintenance of lamellipodia. *Journal of Cell Science* 125, 2775–2785.
- Wang, Y.L., 1985. Exchange of actin subunits at the leading edge of living fibroblasts: Possible role of treadmilling. *Journal of Cell Biology* 101, 597–602.
- Weaver, A.M., Karginov, A.V., Kinley, A.W., *et al.*, 2001. Cortactin promotes and stabilizes Arp2/3-induced actin filament network formation. *Current Biology* 11, 370–374.
- Welch, M.D., Depace, A.H., Verma, S., Iwamatsu, A., Mitchison, T.J., 1997. The human Arp2/3 complex is composed of evolutionarily conserved subunits and is localized to cellular regions of dynamic actin filament assembly. *Journal of Cell Biology* 138, 375–384.
- Wu, D.Y., Wang, L.C., Mason, C.A., Goldberg, D.J., 1996. Association of beta 1 integrin with phosphotyrosine in growth cone filopodia. *Journal of Neuroscience* 16, 1470–1478.
- Yamagishi, A., Masuda, M., Ohki, T., Onishi, H., Mochizuki, N., 2004. A novel actin bundling/filopodium-forming domain conserved in Insulin receptor tyrosine kinase substrate p53 and Missing in metastasis protein. *Journal of Biological Chemistry* 279, 14929–14936.
- Yang, C., Czech, L., Gerboth, S., *et al.*, 2007. Novel roles of formin mDia2 in lamellipodia and filopodia formation in motile cells. *PLoS Biology* 5, e317.
- Yang, C., Hoelzle, M., Disanza, A., Scita, G., Svitkina, T., 2009. Coordination of membrane and actin cytoskeleton dynamics during filopodia protrusion. *PLoS One* 4, e5678.
- Yarmola, E.G., Bubb, M.R., 2009. How depolymerization can promote polymerization: The case of actin and profilin. *Bioessays* 31, 1150–1160.
- Zebda, N., Bernard, O., Bailly, M., *et al.*, 2000. Phosphorylation of ADF/cofilin abolishes EGF-induced actin nucleation at the leading edge and subsequent lamellipod extension. *Journal of Cell Biology* 151, 1119–1128.
- Zhang, D.S., Piazza, V., Perrin, B.J., *et al.*, 2012. Multi-isotope imaging mass spectrometry reveals slow protein turnover in hair-cell stereocilia. *Nature* 481, 520–524.
- Zhao, H., Pykalainen, A., Lappalainen, P., 2011. I-BAR domain proteins: Linking actin and plasma membrane dynamics. *Current Opinion in Cell Biology* 23, 14–21.
- Zigmond, S.H., 2004. Beginning and ending an actin filament: Control at the barbed end. *Current Topics in Developmental Biology* 63, 145–188.

The Extracellular Matrix

AD Bradshaw, Medical University of South Carolina, Charleston, SC, USA

© 2016 Elsevier Inc. All rights reserved.

Extracellular Matrix

Multicellular organisms require extracellular matrix (ECM) for viability and some form of ECM is even produced by single cell bacteria. The ECM of mammals can be divided into two broad categories: basement membranes/basal lamina and connective tissue. Whereas basement membranes are characterized by a thin, planar structure, connective tissue or interstitium is characterized as a space-filling type of ECM rich in fibrillar collagens, proteoglycans, and glycosaminoglycans (Figure 1). These two types of ECM will be discussed in general terms here and the reader should be informed that each organ (and in most cases, specific structures within each organ) contains tissue-specific ECM elements both in terms of basement membranes and connective tissue. As an example, shown in Figure 2 is a schematic representation of a large blood vessel that contains basement membrane, elastic lamellae, and connective tissue. For the sake of brevity with respect to references, in many cases pertinent reviews of each topic are provided so that the reader is encouraged to seek out for further detail and for direction to original references with experimental detail.

Basement Membranes

Early in development, during the peri-implantation period, deposition of the first basement membrane components are produced by the primitive endoderm of the inner cell mass

and assembled on trophoectodermal surfaces (Miner and Yurchenco, 2004). The formation of this first basement membrane is necessary for survival and further development of the embryo illustrating the critical role of this ECM in mammalian life. In fact, proteins of the basement membrane are represented in all forms of multicellular life where they serve vital functions in maintaining and instructing cellular processes. Thus the study of basement membrane proteins in model organisms such as *Drosophila melanogaster* and *Caenorhabditis elegans* has provided valuable insight into structure and function of this ECM (Kelley et al., 2014; Pastor-Pareja and Xu, 2011).

Developmental processes that rely on basement membrane assembly include, cell survival, migration, differentiation, and organization of tissues and organs. In the adult, polarized cell types such as epithelial and endothelial cells have a basement membrane associated with the basal side of each cell. In contrast, muscle cells, Schwann cells, and adipocytes are surrounded by a basement membrane on all sides. The basement membrane serves to provide structural support to cells that is needed for maintaining mechanical stability of tissues. However, the basement membrane also serves many other important functions including control of cell polarization, macromolecular diffusion/filtration, sequestration of cytokine and growth factors, and regulation of both immune and oncogenic cell intravasation/extravasation. Basement membranes serve to delineate cell compartments and also to facilitate linking of epithelial and endothelial layers to underlying connective tissues.

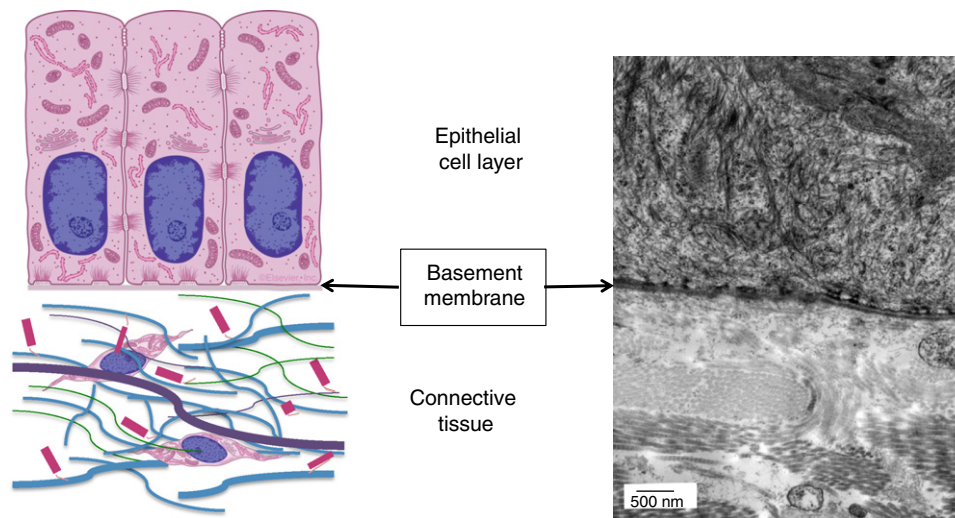


Figure 1 The left panel shows a schematic depiction of an epithelial cell layer with an underlying basement membrane that separates the cells from the connective tissue. Represented structures in the connective tissue are collagen fibers (blue), microfibrils (green), and elastic fibers (purple). Proteoglycans that bind to ECM components are shown by pink rectangles. Fibroblasts, cells that do not have an associated basement membrane, populate the connective tissue. The right panel is a transmission electron microscopy image of the dermal-epidermal border of murine skin. Note the bilayer structure of the basement membrane comprised of the basal lamina and the reticular lamina visible in this image. Collagen fibers in the dermal connective tissue are evident beneath the basement membrane.

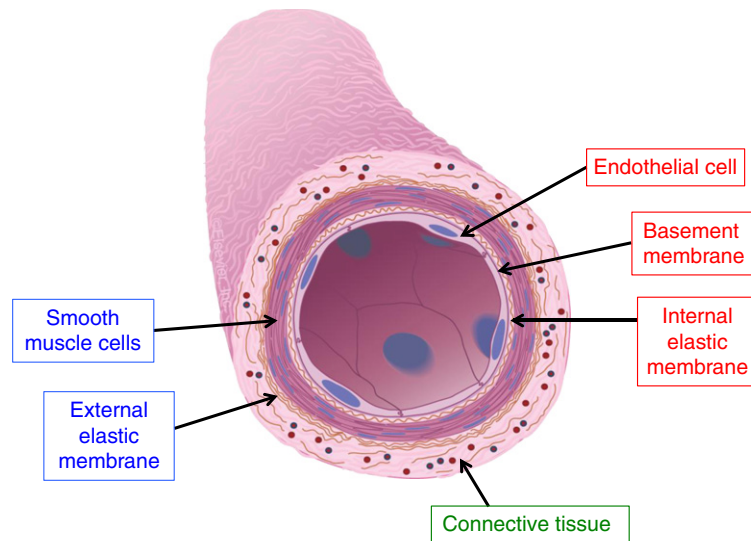


Figure 2 Schematic representation of a large blood vessel. The endothelial cells, the associated underlying basement membrane, and the internal elastic membrane comprise the intima (red lettering). The smooth muscle cells and external elastic membrane comprise the media (blue lettering). The outer layer is the adventitia (green lettering), a connective tissue rich in fibrillar collagens.

Of note, the terms basal lamina and basement membrane have been used somewhat interchangeably. Historically, basement membranes were defined as bilayer structures as viewed under the electron microscope composed of a basal lamina - an internal felt-like structure - and an external reticular lamina (see [Figure 1](#); [Sanes, 2003](#)). Hence, the term basement membrane was to include the basal lamina but also the external reticular lamina. However, recent data suggest that the bilayer appearance of basement membranes arises as an artifact of fixation and that, *in vivo*, these structures resemble a homogenous single layer ([Chan and Inoue, 1994](#); [Miosge, 2001](#)). Each basement membrane contains four basic components: laminin, collagen IV, nidogen, and heparan sulfate proteoglycans, perlecan and agrin ([Figure 3](#)), although many basement membranes in specific tissues have a more complex protein composition that includes additional components ([Yurchenco, 2011](#)).

Laminins

Laminin is a structural protein that forms a network-like structure. The three subunits of laminin - α , β , and γ - are translated into the endoplasmic reticulum where the subunits assemble to form a trimeric protein which then undergo glycosylation. The post-translational glycosylation modifications on laminin are further processed in the Golgi Apparatus prior to secretion into the extracellular space ([Hohenester and Yurchenco, 2013](#)). The laminin network undergoes self-assembly in close association with the cell surface. Each laminin protein forms either a cross-like structure, a Y-shaped structure, or a rod-shaped structure dependent upon the subunit composition that consists of one α , one β , and one γ subunit ([Suh and Miner, 2013](#)). Interaction of laminin with cell surface receptors serves to increase the effective concentration of laminin to facilitate assembly into networks ([Hohenester and Yurchenco, 2013](#)). Assembly of the laminin network

requires Ca^{2+} and is a reversible process in contrast to collagen IV networks, for example, that are stabilized by covalent cross-linking (see below). Laminins serve as a primary cell-binding ligand via cell surface receptors such as integrins and α -dystroglycan. Cytoplasmic domains of integrins and α -dystroglycan receptors form complexes with intracellular scaffolding proteins that interact with cytoskeletal elements. Thus, engagement of laminin by these cell surface receptors serves to link basement membrane ECM with the cytoskeleton to facilitate, for example, cell adhesion and cell movement ([Yurchenco, 2011](#)).

Although the laminin nomenclature has undergone several iterations, the current presiding nomenclature defines each laminin family member from its respective subunit composition. For example, laminin 111 is composed of $\alpha 1$, $\beta 1$, $\gamma 1$ subunits whereas laminin 511 is composed of $\alpha 5$, $\beta 1$, and $\gamma 1$ subunits ([Aumailley et al., 2005](#)). Laminin trimers are formed by the nonrandom assembly of one of each of the five α , four β , and three γ chains to produce at least 15 known different heterotrimers ([Aumailley et al., 2005](#)). The molecular weight of laminin 111 is ~ 900 kD.

Self-assembly of the laminin network occurs at the cell surface and is mediated by interaction of the three 'short arms' of the laminin cross-shaped molecule ([Hohenester and Yurchenco, 2013](#)). Notably, laminins that lack short arms, such as 411 and 421, are not able to independently polymerize to form a network, although can be incorporated into basement membranes via interaction with other forms of laminin. The 'long arm' of the laminin cross contains the LG domains that are the primary sites for laminin receptor binding including laminin-binding integrins and α -dystroglycan ([Yurchenco, 2011](#)). In addition, the proteoglycan agrin binds $\gamma 1$ at a site in the long arm of the molecule with high affinity, which provides additional linkages of laminin networks to cell surfaces ([Suh and Miner, 2013](#)).

Polymerization of the laminin network at the cell surface is required for basement membrane assembly. Abrogation of

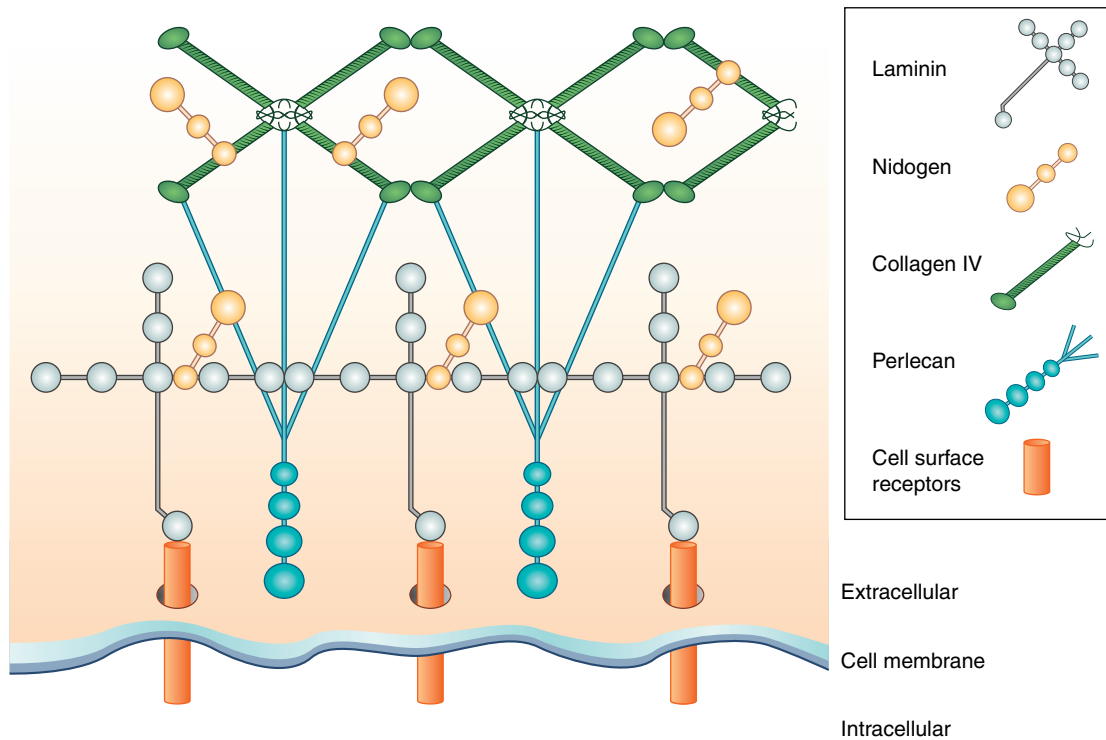


Figure 3 A representation of the current model of interaction of the four principle components of basement membranes in association with cell surfaces. The laminin network forms adjacent to the cell surface and engages cell surface receptors. The collagen IV network forms more distal from the cell surface. Heparan sulfate proteoglycans, for example, perlecan, link the laminin network to the collagen IV network, whereas nidogen interacts with each network independently (see text for specific references).

laminin expression leads to a complete absence of basement membranes in the embryo (Smyth *et al.*, 1999). In contrast, deletion of collagen IV, the other primary basement membrane protein network, does not lead to an absence of a laminin network and basement membranes form to some extent in collagen IV null mice (Poschl *et al.*, 2004). However, basement membranes produced in the absence of collagen IV eventually fail presumably due to an inability to maintain structural integrity that results in embryonic lethality (Poschl *et al.*, 2004).

Collagen IV

Collagen IV is the most abundant protein in basement membranes and also forms a polymerized network, similar to that of laminin (Khoshnoodi *et al.*, 2008). Collagen IV networks are assembled adjacent to, but independent of, the laminin network. In fact, laminin networks can be biochemically separated from collagen IV networks to reveal two distinct polymeric structures (Behrens *et al.*, 2012). Whereas laminin assembles in association with the cell surface, collagen IV networks are elaborated at a more distal location from the cell surface. In some cases, such as the dermal–epidermal junction and skeletal muscle, collagen IV has been shown to be produced by cells in the interstitial ECM and thus these cells are also thought to contribute to basement membrane construction (Kuhl *et al.*, 1984).

Like all members of the collagen family, collagen IV is a trimeric protein made up of three α subunits (Prockop and

Kivirikko, 1995). Collagen IV has a molecular weight of ~ 160 kD. Collagen IV proteins are encoded by six separate α chains that assemble to produce distinct collagen IV family members (Khoshnoodi *et al.*, 2008). The most abundant form of collagen IV is comprised of two collagen $\alpha 1(\text{IV})$ chains and one $\alpha 2(\text{IV})$ chain $[(\alpha 1(\text{IV}))_2 \alpha 2(\text{IV}); \alpha 112]$. The other two known forms of collagen IV have the composition $[(\alpha 3(\text{IV}) \alpha 4(\text{IV}) \alpha 5(\text{IV}); \alpha 345)]$ and $[(\alpha 5(\text{IV}))_2 \alpha 6(\text{IV}); \alpha 556]$. These different forms of collagen IV can assemble into either homotypic or heterotypic polymers or networks as described below (Cummings and Hudson, 2014).

Characteristically, collagen IV proteins have spans of Gly-X-Y repeats indicative of all collagen family members where glycines (Gly) are often followed by proline (X) and hydroxyproline (Y). However, in contrast to fibrillar collagens, the triple helical coiled coil segments formed from the repeated Gly-X-Y sequences (the collagenous domain) in collagen IV are interrupted such that the overall tertiary structure of collagen IV is more flexible than the rod-like fibrillar collagens (Prockop and Kivirikko, 1995). The C-termini of each α subunit, termed the non-collagenous (NC) 1 domain, designates chain recognition in the endoplasmic reticulum to initiate trimerization of the appropriate α subunits to form the three protomers of collagen IV: $\alpha 112$, $\alpha 345$, or $\alpha 556$. Once secreted into the extracellular space, a sulfilimine cross-link in the NC1 domains of one protomer forms a covalent bond with another protomer of collagen IV to initiate network assembly (Cummings and Hudson, 2014). The enzyme peroxidase catalyzes the formation of this sulfilimine bond and requires ionic bromide

(Br-) (McCall *et al.*, 2014). The N-terminal region of collagen IV is referred to as the 7S domain based on its sedimentation coefficient. The 7S domain mediates intermolecular association of four collagen IV molecules and is also a site of intermolecular cross-linking of the collagen IV network. Cross-linking of collagen IV molecules at the N and C termini stabilize the network which is further reinforced through lateral intermolecular associations of the collagenous domains (Khoshnoodi *et al.*, 2008).

The composition of the networks and the relative amounts of collagen IV has significant influence on functional parameters of basement membranes such as permeability, charge-selectivity, tensile strength, and cell interaction. For example, the $\alpha 345$ network is more interconnected and more highly cross-linked in comparison to the network generated from $\alpha 112$ and is therefore likely to provide increased tensile strength to the basement membrane (Gunwar *et al.*, 1998). In kidney development, for example, the glomerular basement membrane is composed primarily of $\alpha 112$ at early stages but is replaced by $\alpha 345$ in the mature glomerulus presumably to provide increased structural support as blood pressure through the glomerulus increases with age (Suh and Miner, 2013). Mutations in genes encoding $\alpha 345$ have been found in patients with kidney disease, including Alport's syndrome, demonstrating the functional importance of collagen IV in the glomerular basement membrane. Likewise mutations in genes encoding the $\alpha 1$ and $\alpha 2$ subunits (COL4A1 and COL4A2) lead to perinatal cerebral hemorrhage and porencephaly, for example (Suh and Miner, 2013; Gould *et al.*, 2005). Hence, the distribution of different polymers of collagen IV is both tissue-specific and dictated by functional requirements of specific basement membranes within a given tissue.

Nidogens

The proteoglycan nidogen (also known as entactin) is also a ubiquitous component of basement membranes with a molecular weight of ~ 150 kD (Yurchenco, 2011). There are two nidogen family members, Nidogen 1 and Nidogen 2. Expression of mRNA encoding nidogens is predominantly associated with cells of mesenchymal origin whereas nidogen proteins are found within basement membranes (Breitkreutz *et al.*, 2013). Therefore nidogens provide another example of the mesenchymal contribution to basement membranes that form adjacent to epithelial cells. Expression patterns and protein localization patterns of nidogen 1 and 2 are overlapping (Breitkreutz *et al.*, 2013). Transgenic mice that lack either nidogen 1 or lack nidogen 2 expression displayed mild phenotypes with no apparent alterations in basement membranes, presumably due to compensation from continued expression of the alternate nidogen. However, mice with abrogated expression of both nidogen 1 and 2 display perinatal lethality with defects in lung and heart basement membranes (Bose *et al.*, 2006). Interestingly, other basement membranes in the double nidogen null mice exhibited apparently normal morphology including the dermal-epidermal junction of the skin. Thus, different basement membranes in the body appear to have different dependencies upon the presence of nidogen for their integrity.

In vitro, nidogen binds to both collagen IV and laminin with fairly high affinity and has been proposed to serve as a molecular bridge or linking molecule that tethers the laminin and collagen IV networks together and thus stabilizes basement membranes. However, recent studies have challenged the strength of the molecular cross-bridge formed by nidogen. In studies by Behrens *et al.* (2012) linking of the laminin and collagen IV networks was dependent upon the heparan sulfate proteoglycan perlecan whereas nidogens were found to associate within both laminin and collagen IV networks but not as a cross-bridge between the two. Nidogens also function to sequester and distribute growth factors in the basement membrane. For example, the growth factor FGF-8 demonstrated altered localization in nidogen double null animals and this was associated with syndactyly and aberrant limb development in these mice (Bose *et al.*, 2006).

Heparan Sulfate Proteoglycans

Heparan sulfate proteoglycans are integral components of basement membranes. Like nidogen, heparan sulfate proteoglycans perlecan and agrin contain both laminin and collagen IV binding sites (Iozzo, 2005; Bezakova and Ruegg, 2003). As mentioned above, current evidence supports a stronger function of perlecan in tethering the two networks together than that of nidogen (Behrens *et al.*, 2012). Agrin is also known to provide additional indirect linkage of laminin trimers to the cell surface. Agrin was originally identified as a critical component of the basement membrane at the neuromuscular junction. Non-muscle agrin is now appreciated as having a broad expression profile in basement membranes throughout the body (Bezakova and Ruegg, 2003). Perlecan expression is not limited to basement membranes and is also found in cartilage and in some connective tissues (Iozzo, 2005).

Heparan sulfate proteoglycans bind to a number of cytokines and growth factors including fibroblast growth factors, platelet derived growth factor, vascular endothelial growth factor among many others (Iozzo, 2005). Hence these proteoglycans can sequester molecules in the ECM. These molecules can serve to establish gradients, inhibit activity, or increase activity of these factors by presenting them to cell surface receptors thereby enhancing receptor binding.

Other ECM Proteins Associated with Basement Membranes

In addition to the four primary constituents of the basement membrane outlined above, other proteins associated with this ECM include collagen XV and XVIII, also heparan sulfate proteoglycans, and collagen VI (Iozzo, 2005; Fitzgerald *et al.*, 2013). These collagen family members are thought to function at the interface between the outer basement membrane and the underlying connective tissue and serve to link the basement membrane to the stroma. In the case of the dermal-epidermal junction, collagen VII is another important linker molecule that attaches the epidermal basement membrane to the dermal interstitium through looping structures

(Nystrom *et al.*, 2013). Blistering disorders of the skin are associated with mutations in collagen VII due to weakening of these collagen VII anchoring fibrils (Nystrom *et al.*, 2013).

The proteins represented in this overview of basement membranes represent the basic components present in the greatest quantity. In fact, the number of ECM proteins associated with basement membranes is rapidly expanding. A recent proteomic analysis using human glomerulus revealed 144 extracellular proteins that were identified in glomerular extracts (Lennon *et al.*, 2014). Likely, other basement membranes isolated from other tissue sources will reveal similar complexity albeit with differences in the profile of extracellular proteins. The field of basement membrane biology is rapidly expanding as the importance of this fundamental aspect of multicellular life, with great significance to development, health, and disease progression, becomes increasingly more appreciated.

Fibronectin

Fibronectin is a critically important ECM protein that mediates cell:ECM interaction during fundamental events such as development, wound healing, fibrosis, and tumor progression (Schwarzbauer and DeSimone, 2011). Fibronectin is secreted as a soluble, covalently bound dimer of ~440 kD molecular weight. Fibronectin is encoded by a single gene however different forms of fibronectin arise through alternative splicing. Circulating fibronectin is produced primarily in the liver and this soluble form of fibronectin is referred to as plasma fibronectin. Cellular fibronectin is produced in a variety of different tissues by resident cells including connective tissues, bone, and periodontal tissues. Antibodies have been generated that distinguish between plasma and cellular fibronectin based on sequence specific epitopes that represent individual forms of fibronectin. The splice variants of fibronectin have been shown to exhibit differential solubility and sensitivity to protease digestion suggesting that expression of each fibronectin variant can influence fibronectin network formation and stability (White and Muro, 2011).

Like laminin, fibronectin assembles into a network on cell surfaces. However, in contrast to laminin where increasing effective concentrations in the pericellular space drives polymerization, in the current model of fibronectin assembly, force generated by the actin cytoskeleton relayed through cell surface integrin receptors drives fibronectin fibrillogenesis (Schwarzbauer and DeSimone, 2011). Fibronectin-binding integrins link to the cytoskeleton through scaffolding proteins that allow forces generated by the actin/myosin cytoskeletal elements to pull on fibronectin to expose an assembly site that is not available in the folded, soluble conformation. The tension generated by the actin cytoskeleton transferred through integrin engagement reveals the site that initiates fibronectin polymerization. The force generated through the actin cytoskeleton is required for fibronectin fibrillogenesis. Syndecans, transmembrane proteoglycans, enhance fibronectin-integrin engagement on cell surfaces and also influence fibronectin assembly (Schwarzbauer and DeSimone, 2011).

Once fibronectin fibrillogenesis is initiated a series of steps that are not completely understood convert the polymerized fibronectin into an insoluble ECM. Formation of insoluble

fibronectin likely involves protein:protein interactions and additional conformational changes that result in a stabilized ECM (Schwarzbauer and DeSimone, 2011). Notably, experiments that have contributed to our current model of fibronectin network formation were predominantly carried out in 2-dimensional cell culture systems. The process that occurs in developing and adult tissues is likely to be more complex and depend upon additional factors such as intrinsic levels of tissue-specific tension, for example. The cortical actin cytoskeleton and regulation by Wnt/PCP (planar cell polarity) pathways *in vivo* have also been implicated in fibronectin assembly in zebra fish models (Dohn *et al.*, 2013).

Formation of a stable fibronectin network is thought to precede and further enhance assembly of other ECM components including collagen fibrils and microfibril networks - a precursor of elastic fiber formation. Hence, fibronectin has been described as a 'master regulator' of ECM assembly. Accordingly, fibronectin has been shown to interact with a variety of ECM proteins including, fibrin, thrombospondin, transglutaminase, and collagens (Halper and Kjaer, 2014).

As mentioned previously, fibronectin plays critical roles in many important cellular processes both in development and adult tissues. Notably, fibronectin is required in the early embryo for mesodermal migration demonstrated in mice, chick, and zebra fish. Polymerization of fibronectin has also been shown to drive branching morphogenesis of developing glands in organ culture models (Larsen *et al.*, 2006). In adult tissues, fibronectin is an important component of blood clots, a critical factor in cell migration during wound healing events, and an important factor in bone maintenance and repair - to name just a few events that depend upon fibronectin expression and deposition.

Connective Tissue/Interstitium

In contrast to basement membranes that form a thin, planar structure, connective tissue ECMs are a space-filling type of matrix. Connective tissue is abundant in fibrillar collagens, proteoglycans and glycosaminoglycans and provides critical structural support in vertebrates (Prockop and Kivirikko, 1995). Mineralized tissues such as bone and teeth are also examples of tissues rich in fibrillar collagens and associated proteoglycans. Similar to basement membranes, the profile of ECM proteins in a given connective tissue is governed by functional requirements of the particular tissue and can vary widely throughout an organism. Skin, tendon, cornea, blood vessel walls, and ligaments are all examples of connective tissue and each has a tissue-specific pattern of ECM components that gives rise to each unique and specialized organ (Figure 4). Notably, the primary component of each of these tissues is collagen I, the prototypic fibrillar collagen, that serves as a remarkably versatile 'building block.' Manipulation and assembly of the same collagen I protein and associated tissue-specific ECM partners by resident cells gives rise to a transparent tissue in cornea and to the strength bearing fibers of tendon. The orientation and the diameter of collagen fibers assembled are two examples of critical parameters that influence tissue architecture and ultimately connective tissue function.

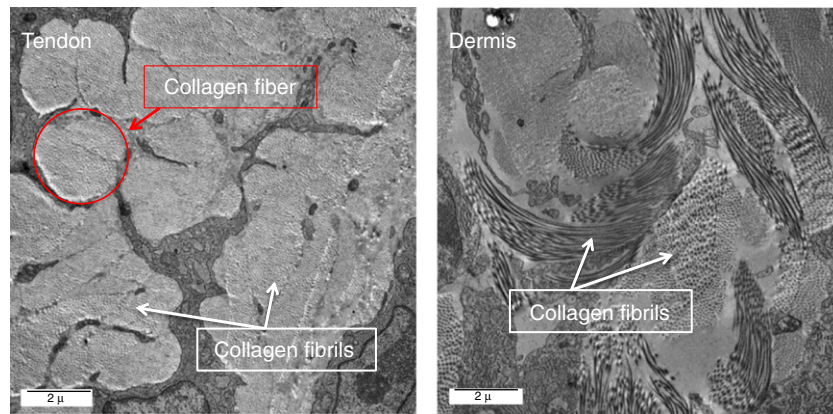


Figure 4 Transmission electron microscope images of the collagenous ECM of tendon and dermis. Individual collagen fibrils are visible by EM. Bundles of collagen fibrils make up collagen fibers. Note the differences in collagen fibril orientation between tendon (parallel alignment) and dermis (various orientations). The 67 nm *D*-periodicity of collagen fibrils are visible in dermal fibrils that run parallel to the plane of section.

Fibrillar Collagens

Collagens I, II, III, and V are structurally very similar (Kadler *et al.*, 2007). These fibril-forming collagen family members are triple helical, like all collagens, and have a long rod-like structure generated by a relatively long, uninterrupted span of Gly-X-Y amino acids (see Section Collagen IV). The three α chains of each are held together through interchain hydrogen bonds. Whereas collagen I is a heterotrimer composed of two $\alpha 1(I)$ and one $\alpha 2(I)$ chains, most fibril-forming collagens are homotrimers composed of a single type of α chain, for example, $\alpha 1(II)_3$ and $\alpha 1(III)_3$ for type II and type III collagen, respectively (Prockop and Hulmes, 1994). Collagen I has an approximate molecular weight of 300 kD. However due to the characteristic rod-like shape of fibrillar collagens, the migration of these proteins on, for example, sodium dodecyl sulfate polyacrylamide electrophoresis (SDS-PAGE) gels, is not relative to standard globular proteins.

Indicative of most secreted proteins, the α subunits of collagens are translated into the endoplasmic reticulum where the triple helices are assembled. A number of chaperones in the endoplasmic reticulum and Golgi Apparatus facilitate collagen subunit assembly and modifications including glycosylation and hydroxylation of prolines and lysines. Chaperones for collagen include heat shock protein 47, protein disulfide isomerase, hydroxylases, and glycosyltransferases (Myllyharju and Kivirikko, 2004). Chain recognition sequences in the C-terminal propeptide domain of each subunit allow for triple helical assembly of appropriate α subunits so that α subunits of collagen I do not form trimers with those of type III, for example. Initiation of triple helical folding occurs at the C-terminus and proceeds toward the N-terminus of the protein. Secretion of the rod-shaped procollagen molecules require specialized machinery as typical transport vesicles used for globular proteins cannot accommodate fibrillar procollagens (Canty and Kadler, 2005).

Procollagen molecules are soluble precursors that are processed to fibril-forming collagen monomers. The C-propeptide is removed by bone morphogenic protein-1/tolloid proteinases whereas the N-terminal propeptide is released by enzymes in the A Disintegrin and Metalloproteinase (ADAM-TS)

family. Enzymatic clipping of the prodomains of collagen convert soluble procollagen to collagen monomers that can be incorporated into insoluble ECM (Canty and Kadler, 2005). Therefore, regulatory factors that influence cleavage of propeptides of fibrillar collagens influence collagen deposition into fibrillar, insoluble ECM. In embryonic tendon, specialized cellular structures termed fibropositors have been described as sites of procollagen processing and nascent fibril assembly sites. These structures are characterized by extended pockets that associate both with intracellular vesicles and extend into the extracellular space (Canty *et al.*, 2004). These structures are abundant in embryonic tissues undergoing high rates of collagen fibril production but have also been observed in some adult tissues that undergo high rates of collagen turnover, such as the periodontal ligament. The exact mechanisms by which cells secrete collagen and influence collagen assembly, particularly in regards to the generation of tissue-specific architectures, continue to be an active area of research.

Collagen monomers are assembled into collagen fibrils in an ordered, staggered array that gives rise to fibrils with a characteristic 67-nm *D*-periodic repeat structure (Orgel *et al.*, 2011). Collagen fibrils undergo further aggregation in the extracellular space to form collagen fibers. Collagen fibrils are visible by electron microscopy and can range from ~ 12 to ~ 500 nm. Bundles of collagen fibrils make up collagen fibers that are visible by light microscopy (Figure 4). Insight into the macromolecular organization of collagen fibers is progressing although many aspects of fiber architecture remain to be determined (Orgel *et al.*, 2011). Importantly, collagen fibrils are not composed of single collagen types but are composite in nature. Evidence that collagen V initiates formation of collagen I containing fibrils, for example, was found in the analysis of collagen V null mice. Collagen fibrils containing types I and III are found in many connective tissues including: skin, tendon, bone, and ligament. Type II collagen fibers are found predominantly in cartilages, vitreous, and intervertebral disc and can also include type III collagen, as another example (Wu *et al.*, 2010).

Insoluble collagenous ECM is generated through the formation of inter and intramolecular covalent cross-links (Eyre *et al.*, 1984). Lysyl oxidase mediated collagen cross-links are the best characterized but other types of cross-links include

those mediated by transglutaminase and nonenzymatic advanced glycation end-products (Wang *et al.*, 2014; Goh and Cooper, 2008). Lysyl oxidase functions in the extracellular space to modify specific amino acids in the telopeptide domain of fibrillar collagens. Upon incorporation of lysyl oxidase-modified monomers into collagen fibrils, covalent cross-link formation occurs. Differences in collagen cross-linking can influence collagen solubility, fibril diameter, sensitivity to degradation, ECM composition, tensile strength, stiffness, and matrix permeability (Eyre *et al.*, 1984).

Collagen fibers can be organized into a number of different types of tissues. The orthogonal array of collagen fibers in cornea with fairly uniform diameters, gives rise to a transparent tissue. Collagen fibers in the skin are tightly arranged in various orientations with a range of fibril and fiber diameters that provide tensile strength in different directions as well as providing a barrier function for the dermis (Figure 4). Collagen fibers in tendons tend to be large bundles of collagen fibrils that align in one direction to transmit forces from skeletal muscles to bones (Figure 4). A number of different types of molecules can contribute to tissue-specific collagen organization.

FACIT Collagens and SLRPs

Fibril-associated Collagens with Interrupted Triple Helices (FACIT) collagens include collagens types IX, XII, XIV, XVI, XIX, XX, XXI, and XXII (Ricard-Blum, 2011). FACIT collagens are frequently found associated with the surface of collagen fibrils and are considered important mediators of collagen fiber organization (Kadler *et al.*, 2007). Small leucine rich proteoglycans (SLRPs) also contribute to collagen fibril and fiber architecture. Decorin, for example, was so named for its capacity to 'decorate' collagen fibers. Mice that do not express decorin exhibit differences in collagen fibril diameter that become more phenotypically distinct with the compound loss of another SLRP, biglycan (Chen and Birk, 2013). Other SLRPs implicated in collagen fiber morphology include: lumican, asporin, fibromodulin, and osteoglycin (for a complete list and review see reference Chen and Birk (2013)).

In addition to the role of proteoglycans in collagen fiber organization, these molecules are also hydrophilic and increase the water content of connective tissues. Glycosaminoglycans such as hyaluronic acid (HA) also facilitate water retention and play key roles in cell migration and ECM deposition (Wang *et al.*, 2011). Connective tissue ECMs can serve as reservoirs for secreted factors such as growth factors, cytokines, and inactive forms of digestive enzymes. Proteoglycans are considered a primary component facilitating the sequestration of signaling molecules in the ECM.

Elastic Fibers

Elastic fibers, the largest ECM structures, provide elasticity to tissues. This property is especially critical for organs such as lung and large blood vessels to allow for elastic recoil of these tissues (Figure 2). Elastic fibers are composed of two components, one made up primarily of elastin protein and the second being a complex of proteins, rich in fibrillins, that form

microfibrils (Wagenseil and Mecham, 2007). Elastin is encoded by a single gene in humans that, like fibronectin, gives rise to different protein forms as a result of tissue-specific alternative splicing of the mRNA. In the current model of elastic fiber assembly, elastin is secreted as tropoelastin and coacervates or aggregates through self-assembly to form an amorphous, hydrophobic core that is cross-linked by lysyl oxidase, the enzyme that also cross-links fibrillar collagens. Formation of elastin aggregates occurs at assembly sites on the cell surface and is facilitated by cell surface components such as heparan sulfate proteoglycans. The aggregates are then transferred to microfibrils. Component of microfibrils include fibrillins, fibulins, EMILIN-1, and microfibril associated glycoproteins (MAGPs) (Wagenseil and Mecham, 2007). Microfibrils are also present in tissues independent of elastic fibers and function as structural elements as well as in the sequestration of growth factors such as transforming growth factor (TGF) β . Examples of diseases associated with mutations in genes encoding fibrillin are Marfan's syndrome and congenital contractural arachnodactyly (Ramirez and Dietz, 2004). Mutations in elastin can give rise to the condition cutis laxa, a disorder characterized by sagging, inelastic skin among other symptoms (Halper and Kjaer, 2014).

Matricellular Proteins

Matricellular proteins were originally defined as proteins that associate with the ECM but do not contribute structural elements and were therefore distinguished from classical ECM components such as laminins and collagens (Bornstein, 1995). Primarily, matricellular proteins are viewed as molecules that mediate cell:ECM interaction (Murphy-Ullrich and Sage, 2014). Many matricellular proteins bind to cell surface receptors and to structural components of the ECM and thus influence a broad range of cellular events. Importantly, expression and activity of matricellular proteins in general tends to be highly contextual and can vary significantly depending on the cellular milieu. The matricellular family includes, SPARC and SPARC-like proteins, thrombospondins, tenascins, osteopontin, periostin, fibulins, and Cyr, CTGF, Nov (CCN) proteins.

SPARC is a prototypic matricellular protein characterized by a modular domain structure. In the absence of SPARC expression, differences in connective tissue ECM are detected that generally become more advanced with age (Bradshaw, 2009). Shown in the top panel of Figure 5 is a SPARC-null mouse that appears to have a grossly normal phenotype with the exception of a curled tail. At the level of the electron microscope, differences in collagen fibril organization are apparent in the tail tendon of these mice (Figure 5, lower panels). Similar to the phenotype of other mice with abrogated expression of specific matricellular proteins, SPARC-null mice have subtle phenotypic differences until challenged (Bradshaw, 2012; Murphy-Ullrich and Sage, 2014). However, SPARC-null mice, for example, are unable to mount a significant fibrotic response after injury in many tissues including liver, kidney, lung, heart, and skin (Bradshaw, 2009).

In most cases, matricellular proteins are expressed more highly in developing tissues and have reduced expression in adult tissues. In response to different types of injury or tumor

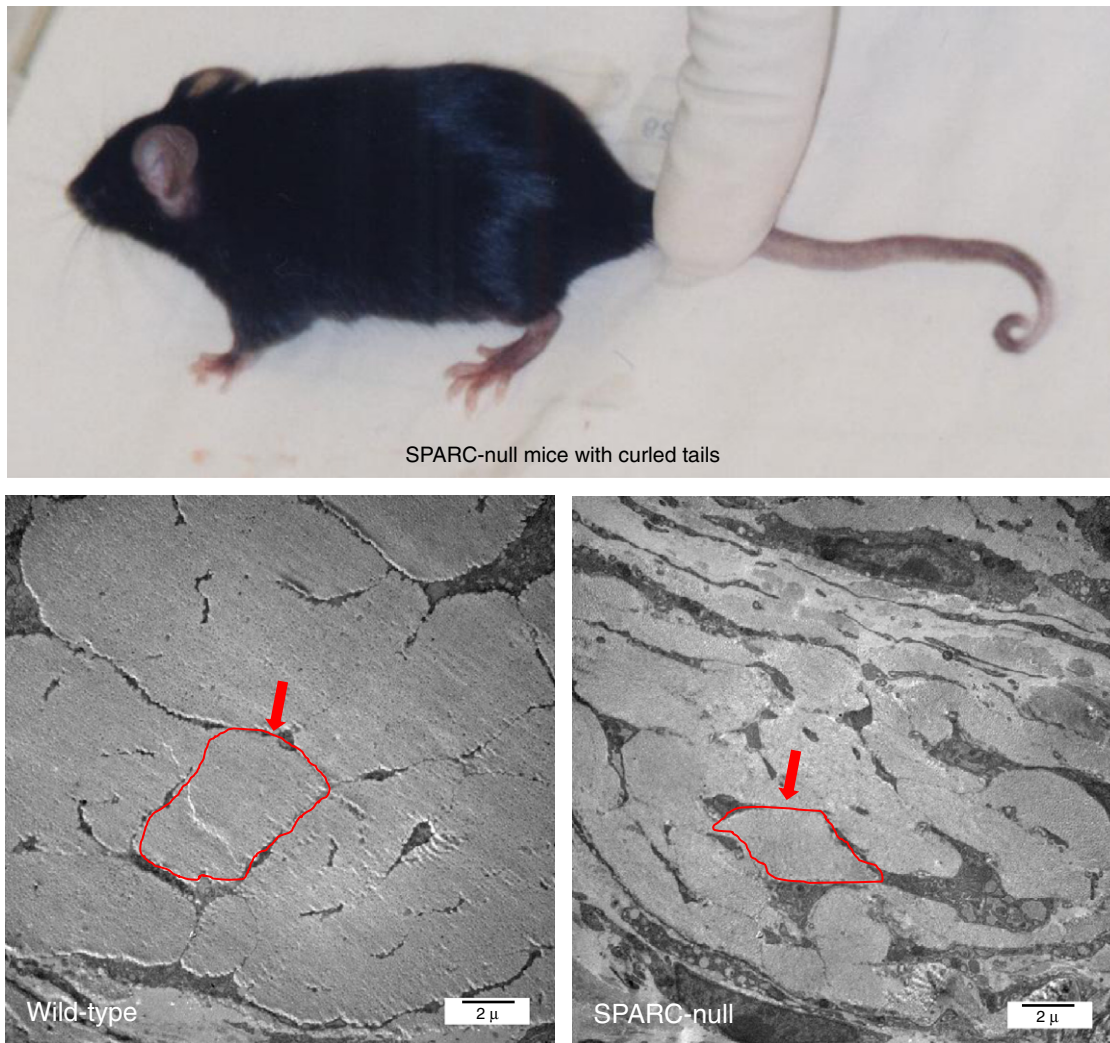


Figure 5 Top panel: SPARC-null mice exhibit curled tails. Bottom panel: Images of tail tendon from normal (wild-type) mice and SPARC-null mice at 1 month of age taken by transmission electron microscopy. Note the differences in collagen fiber organization in the absence of SPARC expression (highlighted in red).

development, re-expression of robust amounts of particular matricellular proteins are frequently observed. For example, significant increases in matricellular protein expression are often found in tissues undergoing wound healing, tumor progression, and fibrosis. The expression patterns and the general phenotype of the null mice suggest that matricellular proteins are important mediators of cell activity in response to injury and ECM remodeling. These proteins are recognized as critical regulators of a wide variety of cellular activities including: ECM assembly, inflammation, fibrosis, wound healing, migration/metastasis, growth factor activity, synapse formation, calcification, and apoptosis (Murphy-Ullrich and Sage, 2014).

Conclusions

- ECMs are required for multicellular life.
- Tissue-specific ECMs dictate form and function in tissues and organs.
- There are two primary types of ECM: basement membrane/basal lamina and connective tissue.
- There are four primary components of basement membranes: laminins, collagen IV, nidogens, and heparan sulfate proteoglycans.
- Laminin and collagen IV form independent networks and each type has the capacity to self-assemble.
- Collagen IV networks are stabilized by covalent cross-links whereas laminin network assembly is reversible.
- Nidogen and heparan sulfate proteoglycans confer stability to basement membranes as well as serving to sequester growth factors and to interact with cell surface receptors.
- Fibronectin is a critical ECM protein essential for development and serves as a 'master regulator' of many structural elements in the ECM.
- Connective tissue ECM is primarily composed of fibrillar collagens, proteoglycans, and glycosaminoglycans.
- Fibrillar collagens are synthesized as soluble procollagen molecules.

- Removal of the propeptides of fibrillar collagen precedes insoluble incorporation of collagen monomers to collagen fibrils.
- Collagen fibrils and fibers are composites of multiple types of collagens and associated proteoglycans such as those of the SLRP family.
- Collagen fibrils and fibers are stabilized by the formation of intra and intermolecular cross-links.
- Elastic fibers include an elastin component assembled onto microfibrils. Protein components of microfibrils include fibrillins, fibulins, EMILIN-1, and MAGP.
- Matricellular proteins mediate cellular interaction with structural ECM components and are essential factors in a variety of tissue processes

See also: Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell Adhesion to the Extracellular Matrix

References

- Aumailley, M., Bruckner-Tuderman, L., Carter, W.G., *et al.*, 2005. A simplified laminin nomenclature. *Matrix Biology* 24, 326–332.
- Behrens, D.T., Villone, D., Koch, M., *et al.*, 2012. The epidermal basement membrane is a composite of separate laminin- or collagen IV-containing networks connected by aggregated perlecan, but not by nidogens. *Journal of Biological Chemistry* 287, 18700–18709.
- Bezakova, G., Ruegg, M.A., 2003. New insights into the roles of agrin. *Nature Reviews Molecular Cell Biology* 4, 295–308.
- Bornstein, P., 1995. Diversity of function is inherent in matricellular proteins: An appraisal of thrombospondin 1. *Journal of Cell Biology* 130, 503–506.
- Bose, K., Nischt, R., Page, A., *et al.*, 2006. Loss of nidogen-1 and -2 results in syndactyly and changes in limb development. *Journal of Biological Chemistry* 281, 39620–39629.
- Bradshaw, A.D., 2009. The role of SPARC in extracellular matrix assembly. *Journal of Cell Communication and Signaling* 3, 239–246.
- Bradshaw, A.D., 2012. Diverse biological functions of the SPARC family of proteins. *International Journal of Biochemistry & Cell Biology* 44, 480–488.
- Breitkreutz, D., Koxholt, I., Thiemann, K., Nischt, R., 2013. Skin basement membrane: The foundation of epidermal integrity – BM functions and diverse roles of bridging molecules nidogen and perlecan. *BioMed Research International* 2013, 1–16.
- Canty, E.G., Kadler, K.E., 2005. Procollagen trafficking, processing and fibrillogenesis. *Journal of Cell Science* 118, 1341–1353.
- Canty, E.G., Lu, Y., Meadows, R.S., *et al.*, 2004. Coalignment of plasma membrane channels and protrusions (fibripositors) specifies the parallelism of tendon. *Journal of Cell Biology* 165, 553–563.
- Chan, F.L., Inoue, S., 1994. Lamina lucida of basement membrane: An artefact. *Microscopy Research and Technique* 28, 48–59.
- Chen, S., Birk, D.E., 2013. The regulatory roles of small leucine-rich proteoglycans in extracellular matrix assembly. *FEBS Journal* 280, 2120–2137.
- Cummings, C.F., Hudson, B.G., 2014. Lens capsule as a model to study type IV collagen. *Connective Tissue Research* 55, 8–12.
- Dohn, M.R., Mundell, N.A., Sawyer, L.M., Dunlap, J.A., Jessen, J.R., 2013. Planar cell polarity proteins differentially regulate extracellular matrix organization and assembly during zebrafish gastrulation. *Developmental Biology* 383, 39–51.
- Eyre, D.R., Paz, M.A., Gallop, P.M., 1984. Cross-linking in collagen and elastin. *Annual Review of Biochemistry* 53, 717–748.
- Fitzgerald, J., Holden, P., Hansen, U., 2013. The expanded collagen VI family: New chains and new questions. *Connective Tissue Research* 54, 345–350.
- Goh, S.Y., Cooper, M.E., 2008. Clinical review: The role of advanced glycation end products in progression and complications of diabetes. *Journal of Clinical Endocrinology and Metabolism* 93, 1143–1152.
- Gould, D.B., Phalan, F.C., Breedveld, G.J., *et al.*, 2005. Mutations in Col4a1 cause perinatal cerebral hemorrhage and porencephaly. *Science* 308, 1167–1171.
- Gunwar, S., Ballester, F., Noelken, M.E., *et al.*, 1998. Glomerular basement membrane. Identification of a novel disulfide-cross-linked network of alpha3, alpha4, and alpha5 chains of type IV collagen and its implications for the pathogenesis of Alport syndrome. *Journal of Biological Chemistry* 273, 8767–8775.
- Halper, J., Kjaer, M., 2014. Basic components of connective tissues and extracellular matrix: Elastin, fibrillin, fibulins, fibrinogen, fibronectin, laminin, tenascins and thrombospondins. *Advances in Experimental Medicine and Biology* 802, 31–47.
- Hohenester, E., Yurchenco, P.D., 2013. Laminins in basement membrane assembly. *Cell Adhesion & Migration* 7, 56–63.
- Iozzo, R.V., 2005. Basement membrane proteoglycans: From cellar to ceiling. *Nature Reviews Molecular Cell Biology* 6, 646–656.
- Kadler, K.E., Baldock, C., Bella, J., Boot-Handford, R.P., 2007. Collagens at a glance. *Journal of Cell Science* 120, 1955–1958.
- Kelley, L.C., Lohmer, L.L., Hagedorn, E.J., Sherwood, D.R., 2014. Traversing the basement membrane *in vivo*: A diversity of strategies. *Journal of Cell Biology* 204, 291–302.
- Khoshnoodi, J., Pedchenko, V., Hudson, B.G., 2008. Mammalian collagen IV. *Microscopy Research and Technique* 71, 357–370.
- Kuhl, U., Ocalan, M., Timpl, R., *et al.*, 1984. Role of muscle fibroblasts in the deposition of type-IV collagen in the basal lamina of myotubes. *Differentiation* 28, 164–172.
- Larsen, M., Wei, C., Yamada, K.M., 2006. Cell and fibronectin dynamics during branching morphogenesis. *Journal of Cell Science* 119, 3376–3384.
- Lennon, R., Byron, A., Humphries, J.D., *et al.*, 2014. Global analysis reveals the complexity of the human glomerular extracellular matrix. *Journal of the American Society of Nephrology* 25, 939–951.
- McCall, A.S., Cummings, C.F., Bhawe, G., *et al.*, 2014. Bromine is an essential trace element for assembly of collagen IV scaffolds in tissue development and architecture. *Cell* 157, 1380–1392.
- Miner, J.H., Yurchenco, P.D., 2004. Laminin functions in tissue morphogenesis. *Annual Review of Cell and Developmental Biology* 20, 255–284.
- Miosge, N., 2001. The ultrastructural composition of basement membranes *in vivo*. *Histology and Histopathology* 16, 1239–1248.
- Murphy-Ullrich, J.E., Sage, E.H., 2014. Revisiting the matricellular concept. *Matrix Biology* 37, 1–14.
- Mylyharju, J., Kivirikko, K.I., 2004. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends in Genetics* 20, 33–43.
- Nystrom, A., Velati, D., Mittapalli, V.R., *et al.*, 2013. Collagen VII plays a dual role in wound healing. *Journal of Clinical Investigation* 123, 3498–3509.
- Orgel, J.P., San Antonio, J.D., Antipova, O., 2011. Molecular and structural mapping of collagen fibril interactions. *Connective Tissue Research* 52, 2–17.
- Pastor-Pareja, J.C., Xu, T., 2011. Shaping cells and organs in *Drosophila* by opposing roles of fat body-secreted Collagen IV and perlecan. *Developmental Cell* 21, 245–256.
- Poschl, E., Schlotzer-Schrehardt, U., Brachvogel, B., *et al.*, 2004. Collagen IV is essential for basement membrane stability but dispensable for initiation of its assembly during early development. *Development* 131, 1619–1628.
- Prockop, D.J., Hulmes, D.J.S., 1994. Assembly of Collagen Fibrils De Novo from Soluble Precursors: Polymerization and Copolymerization of Procollagen, pN-Collagen, and Mutated Collagens. San Diego, CA: Academic Press.
- Prockop, D.J., Kivirikko, K.I., 1995. Collagens: molecular biology, diseases, and potentials for therapy. *Annual Review of Biochemistry* 64, 403–434.
- Ramirez, F., Dietz, H.C., 2004. Therapy insight: Aortic aneurysm and dissection in Marfan's syndrome. *Nature Clinical Practice Cardiovascular Medicine* 1, 31–36.
- Ricard-Blum, S., 2011. The collagen family. *Cold Spring Harbor Perspectives in Biology* 3, a004978.
- Sanes, J.R., 2003. The basement membrane/basal lamina of skeletal muscle. *Journal of Biological Chemistry* 278, 12601–12604.
- Schwarzbauer, J.E., DeSimone, D.W., 2011. Fibronectins, their fibrillogenesis, and *in vivo* functions. *Cold Spring Harbor Perspectives in Biology* 3, a005041.
- Smyth, N., Vatansever, H.S., Murray, P., *et al.*, 1999. Absence of basement membranes after targeting the LAMC1 gene results in embryonic lethality due to failure of endoderm differentiation. *Journal of Cell Biology* 144, 151–160.
- Suh, J.H., Miner, J.H., 2013. The glomerular basement membrane as a barrier to albumin. *Nature Reviews Nephrology* 9, 470–477.
- Wagenseil, J.E., Mecham, R.P., 2007. New insights into elastic fiber assembly. *Birth Defects Research Part C: Embryo Today* 81, 229–240.
- Wang, A., De La Motte, C., Lauer, M., Hascall, V., 2011. Hyaluronan matrices in pathobiological processes. *FEBS Journal* 278, 1412–1418.

- Wang, L., Uhlig, P.C., Eikenberry, E.F., *et al.*, 2014. Lateral growth limitation of corneal fibrils and their lamellar stacking depend on covalent collagen cross-linking by transglutaminase-2 and lysyl oxidases, respectively. *Journal of Biological Chemistry* 289, 921–929.
- White, E.S., Muro, A.F., 2011. Fibronectin splice variants: Understanding their multiple roles in health and disease using engineered mouse models. *IUBMB Life* 63, 538–546.
- Wu, J.J., Weis, M.A., Kim, L.S., Eyre, D.R., 2010. Type III collagen, a fibril network modifier in articular cartilage. *Journal of Biological Chemistry* 285, 18537–18544.
- Yurchenco, P.D., 2011. Basement membranes: Cell scaffoldings and signaling platforms. *Cold Spring Harbor Perspectives in Biology* 3 (2), a004911.

Relevant Websites

<http://www.asmb.net>
American Society for Matrix Biology.
<http://proteopedia.org/wiki/index.php/Collagen>
Collagen I Structure.
<http://ismb.org>
International Society for Matrix Biology.
<http://www.extracellularmatrixnews.com>
News Source.

Cell–Cell Adhesion and the Cytoskeleton

BR Acharya and AS Yap, The University of Queensland, Brisbane, QLD, Australia

© 2016 Elsevier Inc. All rights reserved.

Introduction

Cell–cell adhesion helps shape tissue organization during morphogenesis and preserves tissue cohesion in post-developmental life. In doing so, cell adhesion molecules influence diverse cellular processes, including cell–cell cohesion, recognition, signaling, and the regulation of cell proliferation. A variety of types of cell–cell adhesion molecules are now known, such as cadherins, nectins, and adhesion molecules of the immunoglobulin (Ig) superfamily. In this article, we focus on how cytoskeletal systems interact with cell adhesion receptors to influence their biological functions. Cell–cell adhesion is commonly supported by transmembrane adhesion receptors whose cytoplasmic tails can associate physically with cytoskeletal elements and, often, also regulate the dynamics and organization of those cytoskeletal partners. Indeed, all of the three major cytoskeletal components of eukaryotic cells – actin microfilaments (MFs), intermediate filaments (IFs), and microtubules (MTs) – can interact with cell–cell adhesion receptors. Here we will outline how these interactions occur and their functional implications. Increasingly, it is apparent that cell–cell adhesion cooperates with the cytoskeleton: adhesion systems can regulate the cytoskeleton with which they are associated and, in turn, those cytoskeletal elements contribute to the biology of adhesion. Accordingly, we have structured our discussion around the individual cytoskeletal systems and focused on solid tissues of the body, particularly epithelia and endothelia, where much of the research on adhesion–cytoskeletal cooperation has been conducted.

Cell–Cell Adhesion and the Actin Cytoskeleton

The F-actin cytoskeleton is necessary for the integrity of many types of cell–cell interactions; this substantially reflects its ability to cooperate with classical cadherin adhesion receptors (Figure 1). Classical cadherins are single-pass transmembrane proteins that engage in homophilic cell–cell adhesion, their cytoplasmic tails directly associate with β -catenin and p120-catenin. β -catenin serves in turn as an anchor for α -catenin, thereby forming a quaternary complex that is commonly called the cadherin–catenin complex (Niessen *et al.*, 2011; Takeichi, 2014). However, many other proteins can interact directly or indirectly with cadherins, although often in a sub-stoichiometric fashion, implying that their recruitment may be subject to cellular regulation.

Integrating Classical Cadherin Adhesion with the Actin Cytoskeleton

Different tissues express different members of the classical cadherin protein family, such as E-cadherin (epithelia), N-cadherin (neurons and mesenchymal cells), and VE-cadherin (endothelia) (Niessen *et al.*, 2011; Takeichi, 2014).

Although it is commonly believed that most, if not all, classical cadherins interact with F-actin, this process has been best studied in epithelia and endothelia. In those confluent tissues, a variety of F-actin networks interact with cadherins, depending on the tissue type being studied and its developmental or functional context. F-actin often concentrates at the cortex of the junctions where cadherins accumulate, as is seen in many simple and stratified epithelia and endothelia (Shewan *et al.*, 2005; Smutny *et al.*, 2010; Huveneers *et al.*, 2012). Similarly, N-cadherin accumulates at the tips of actin-rich dendritic spikes in neurons (Abe *et al.*, 2004; Chazneau *et al.*, 2015). As well, more distinctive patterns of F-actin organization are seen in specialized cellular locations. In simple epithelia, a subpopulation of E-cadherin is stabilized in a specialized junction, the zonula adherens (ZA), which is located at the apical region of cell–cell contacts (Meng *et al.*, 2008). At the ZA, E-cadherin localizes with cortical F-actin and also F-actin bundles that accumulate parallel to the junction itself (Kovacs *et al.*, 2011). In other circumstances, endothelial and epithelial cells can also display F-actin bundles that are

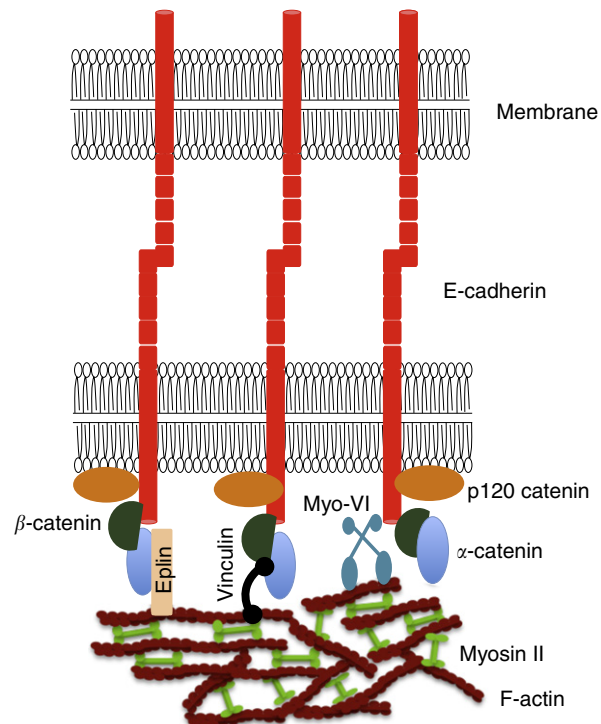


Figure 1 Interactions between classical cadherins and the actin cytoskeleton. Classical cadherins can associate with cortical F-actin through α -catenin, actin-binding proteins that are recruited by α -catenin (EPLIN, vinculin) and by other proteins that can potentially interact with the cadherin cytoplasmic tail (Myosin VI). Junctional actin networks can incorporate Myosin II to form contractile apparatuses.

oriented perpendicular to the cell–cell junctions and terminate in puncta of cadherin (Kwiatkowski *et al.*, 2010; Huveneers *et al.*, 2012; Taguchi *et al.*, 2011). Cortical junctional patterns of F-actin can be converted into perpendicular actin bundles (Taguchi *et al.*, 2011), but the precise mechanisms that control this organizational change are poorly understood. Finally, actin networks (often described as medial-apical networks) have been identified at the apical poles of embryonic epithelia and extend to associate with cadherin junctions (Roh-Johnson *et al.*, 2012; Rauzi *et al.*, 2010).

The F-actin networks that interact with cadherin junctions also associate with a range of other proteins (Figure 1). In particular, non-muscle myosin II interacts with these networks, decorating bundles and cortical F-actin networks both at junctions and the medial-apical poles of cells (Shewan *et al.*, 2005; Smutny *et al.*, 2010; Rauzi *et al.*, 2010). Thus, the actin networks that interact with cadherin junctions often represent actomyosin contractile apparatuses, which can exert force on the junctions. Contractile force is manifest as cellular-scale tension in the junctions (Ratheesh *et al.*, 2012; Wu *et al.*, 2014; Fernandez-Gonzalez *et al.*, 2009); as molecular-scale tension upon cadherin itself (Borghi *et al.*, 2012; Conway *et al.*, 2013); and as movements of the junctions that follow contractions in their associated medial-apical actomyosin networks (Rauzi *et al.*, 2010). It is increasingly apparent that mechanical force influences many aspects of cadherin adhesion, suggesting that cadherin junctions may be regarded as contractile structures that couple cells with one another (Gomez *et al.*, 2011; Pruijt *et al.*, 2014). The integration of actomyosin and cadherin junctions reflects: (1) physical connections between cadherin–catenin complexes and actin filaments and (2) the ability of cadherin adhesions to regulate the F-actin cytoskeleton.

Physical connections between cadherins and F-actin

Multiple molecular mechanisms can couple cadherin adhesion complexes to the F-actin cytoskeleton (Figure 1). One of the most intensively studied of those mechanisms involves α -catenin (Maiden and Hardin, 2011), which is a homologue of vinculin that can directly bind to F-actin (Rimm *et al.*, 1995). Taken with its ability to bind β -catenin, this suggested a model where a ternary complex of cadherin, β -catenin and α -catenin represented the minimal mechanism to couple cadherin adhesion and F-actin (Takeichi, 1991). Consistent with this, biological studies indicate that both the F-actin- and β -catenin-binding domains of α -catenin contribute to cell–cell integrity (Kwiatkowski *et al.*, 2010). However, this model proved controversial, because it was difficult to reconstitute the predicted biochemical complex *in vitro* (Yamada *et al.*, 2005; Drees *et al.*, 2005). However, those early studies were performed in solution and it has recently been revealed that force must be applied for the ternary cadherin–catenin complex to stably bind actin filaments (Buckley *et al.*, 2014). This is consistent with increasing evidence that the conformation of α -catenin and its binding interactions can be regulated by mechanical force (Yao *et al.*, 2014; Yonemura *et al.*, 2010). Therefore, α -catenin may provide a ubiquitous pathway for cadherins to associate with F-actin that becomes apparent when under force.

However, α -catenin is not the only mechanism for cadherins to associate with F-actin. α -Catenin can also anchor the association of other F-actin-binding proteins, such as vinculin

(Huveneers *et al.*, 2012; Yao *et al.*, 2014; Yonemura *et al.*, 2010) and epithelial protein lost in neoplasm (EPLIN) (Abe and Takeichi, 2008), thereby indirectly mediating an association with actin filaments. Alternatively, other actin-binding proteins, such as Myosin VI (Maddugoda *et al.*, 2007) and cortactin (Han *et al.*, 2014), can bind directly to the cadherin cytoplasmic tail. Unlike α -catenin, these other actin-binding proteins are not ubiquitous components of the cadherin molecular complex, which suggests that their contribution to cadherin–actin interactions may depend on the functional context of the junctions. Here, too, mechanical force can play an important influence, as illustrated by the observation that contractility promotes the junctional recruitment of both vinculin (Yonemura *et al.*, 2010; Yao *et al.*, 2014; Leerberg *et al.*, 2014) and EPLIN (Taguchi *et al.*, 2011).

Dynamic regulation of the junctional actin cytoskeleton

The F-actin networks that associate with cadherin junctions are dynamic. F-actin is assembled at junctions when they are assembled (Kovacs *et al.*, 2002b; Tang and Brieher, 2012) and even in mature epithelia junctional F-actin undergoes dynamic turnover (Drees *et al.*, 2005; Kovacs *et al.*, 2011). Cadherin adhesion itself influences the junctional cytoskeleton. Ligation of E-cadherin receptors promotes actin assembly at newly forming adhesions (Kovacs *et al.*, 2002b), while the F-actin content at cell–cell junctions is reduced when cadherins are depleted (Wu *et al.*, 2014). This reflects two broad pathways for cadherins to influence the dynamic actin cytoskeleton.

Firstly, cadherins can directly or indirectly recruit regulators of the actin cytoskeleton to the junctional cortex. These include Arp2/3 (actin-related proteins) (Kovacs *et al.*, 2002b; Tang and Brieher, 2012) and formin nucleators (Kobiela *et al.*, 2004; Carramusa *et al.*, 2007), which can initiate the process of actin assembly at junctions. In addition, cadherin junctions recruit modulators of actin assembly, such as α -actinin-4 (Tang and Brieher, 2012), which promotes Arp2/3 activity; N-WASP (Neuronal Wiskott–Aldrich Syndrome protein), which can stabilize nascent filaments (Kovacs *et al.*, 2011); and Ena/VASP (enabled/vasodilator-stimulated phosphoprotein) proteins that promote filament elongation (Scott *et al.*, 2006). Thus, cadherin adhesions provide cortical platforms that localize the many proteins responsible for regulating junctional actin dynamics and organization.

Secondly, cadherin adhesions promote cortical signals that influence actin dynamics. These signals include the Rac GTPase (Kovacs *et al.*, 2002a; Yamada and Nelson, 2007; Lambert *et al.*, 2002), which stimulates Arp2/3-based actin nucleation (Insall and Machesky, 2009), and also the Rho GTPase (Yamada and Nelson, 2007), which can stimulate formins (Chesarone and Goode, 2009). Rho also promotes the activation and recruitment of Myosin II to complete assembly of the actomyosin contractile apparatus (Jaffe and Hall, 2005; Bresnick, 1999).

Many biochemical interactions underlie the ability of cadherin adhesions to recruit actin regulators and their upstream signals. For example, E-cadherin recruits the Arp2/3 complex through the scaffolding protein, cortactin (Han *et al.*, 2014), while α -catenin can recruit formin-1 (Kobiela *et al.*, 2004), vinculin and Ena/VASP proteins (Leerberg *et al.*, 2014), and the Ect2 Rho activator (Ratheesh *et al.*, 2012). How these are

all coordinated has yet to be elucidated, but mechanical force likely contributes; it is increasingly apparent that cadherins play important roles in mechanosensing at cell–cell adhesions (Ladoux *et al.*, 2010). This can be appreciated when force is experimentally applied to cells bearing cadherins, a process that can lead to cell stiffening (le Duc *et al.*, 2010) or changes in cell migration (Weber *et al.*, 2012). This implies that cells possess signaling pathways that allow them to detect, and respond to, force that is applied to cadherins. At native cell–cell contacts, the source of those forces is likely to include contractility generated within neighboring cells themselves (Liu *et al.*, 2010; Maitre *et al.*, 2012). Such contractility will involve the interactions that classical cadherins make with the actomyosin cytoskeleton that we have already discussed. At the molecular level, force can modulate the ability of α -catenin to bind proteins, such as vinculin (Yao *et al.*, 2014), however the pathways responsible for cadherin mechanotransduction have yet to be characterized in detail.

Biological Functions of Cadherin–Actin Interactions

Classical cadherins participate in a wide range of biological processes, including the maintenance of tissue integrity, morphogenesis, cell–cell communication, and contact inhibition of locomotion (Takeichi, 2014). Integration with the actin cytoskeleton contributes to the action of cadherins at many levels, which encompass the strengthening of adhesion and stabilization of cadherin junctions, to cadherin-dependent morphogenetic movements.

Cadherin adhesion and adhesive strengthening

A core function of adhesion molecules is to resist the external forces that tend to break adhesive bonds and cause cell detachment. Adhesion by classical cadherins is mediated by adhesive ligation of their extracellular domains. However, resistance to detachment is also strengthened when the cytoplasmic domain interacts with the actin cytoskeleton. Adhesion is compromised when cadherin receptors are unable to interact with F-actin (Takeichi, 1991; Yap *et al.*, 1997) and also when either actin filament integrity or myosin II function are inhibited (Takeichi, 1991; Smutny *et al.*, 2010). Actomyosin may strengthen cadherin adhesion through cadherin receptor clustering and by modulating cortical tension.

Classical cadherins form lateral clusters of varying size (Iino *et al.*, 2001; Yap *et al.*, 1997), which can strengthen adhesion by increasing the avidity of adhesive binding. Recent superresolution studies suggest a minimal cluster size of ~5–6 individual cadherin molecules which can then be organized into larger-scale groups (Wu *et al.*, 2015) to form puncta that are readily identified by conventional light microscopy. It has been proposed that cadherin ectodomains alone can mediate clustering by establishing a lattice of adhesive (*trans*-) interactions between ectodomains on opposing surfaces and lateral (*cis*-) interactions between those ectodomains on the same cell surface (Harrison *et al.*, 2011; Brasch *et al.*, 2012). However, cytoplasmic interactions also contribute to clustering (Yap *et al.*, 1997), and, indeed, the formation of larger-scale clusters is perturbed when either F-actin or myosin are inhibited (Hong *et al.*, 2013; Wu *et al.*, 2015; Smutny *et al.*, 2010).

Cortical actin networks can promote clustering by acting as corrals that limit lateral diffusion of individual cadherin receptors (Kusumi *et al.*, 2005; Morone *et al.*, 2006; Wu *et al.*, 2015). Interestingly, elemental clusters can be seen even in the absence of adhesive ligation (Wu *et al.*, 2015), suggesting that a cortical actin mesh may serve to generate the minimal clusters of cadherins. In addition, myosin may promote further clustering by cross-linking and condensing actin networks (Backouche *et al.*, 2006; Soares e Silva *et al.*, 2011).

Junctional actomyosin can also influence cadherin adhesion by modulating cortical rigidity. Cells detach by peeling away from adhesive surfaces (Ward *et al.*, 1994) and mechanisms that retard peeling antagonize detachment, thereby promoting adhesion. Actomyosin contributes to rigidity of the cell cortex by cross-linking actin filaments and promoting cortical tension (Salbreux *et al.*, 2012). When coupled to cadherins, cortical actomyosin may then increase rigidity of the junctional membrane and strengthen adhesion.

Cadherin stabilization and junction formation

The actomyosin cytoskeleton also regulates the dynamic turnover of cadherins at cell–cell junctions. Measured by fluorescence recovery after photobleaching, the stability of E-cadherin at junctions requires myosin II and its upstream regulator, Rho (Priya *et al.*, 2013). This may explain why myosin is necessary in simple epithelial cells for E-cadherin to concentrate in the apical junction of the ZA (Smutny *et al.*, 2010). Such stabilization may reflect the ability of actomyosin to limit the lateral diffusion of cadherins in the plasma membrane, both through lateral corrals and also by myosin-dependent cross-linking and actin network condensation. In addition, association with the actin cytoskeleton can retard cadherin endocytosis (Izumi *et al.*, 2004). Association with actin may potentially antagonize cadherin internalization and thereby stabilize E-cadherin at the cell surface.

Junctional contractility and morphogenesis

Cadherin junctions participate in morphogenetic interactions where cells undergo active rearrangements that are driven by actomyosin-based contractility. Here, cadherins serve not only to allow junctions to resist detachment forces, but also to transmit those contractile forces between neighboring cells. We will outline three forms of morphogenetic movement that entail integration of cadherin adhesion with actomyosin contractility.

Firstly, cell sorting is a process in which cell populations within embryos physically segregate away from one another. It is classically illustrated by the ability of cells from different germ layers to associate with other cells from the same germ layer and dissociate from those of other germ layers (Fagotto, 2015). This was long attributed to expression of different classical cadherins, which allows cells to sort out when mixed in tissue culture (Takeichi, 1991). However, the differential binding capacity of cadherin ectodomains is not sufficient to explain the observed sorting behaviors (Niessen and Gumbiner, 2002). Recently, it was revealed that differences in cortical tension, determined by the actomyosin cytoskeleton, also contribute to the segregation of embryonic germ layers (Krieg *et al.*, 2008). Cadherin adhesion may then participate in cell

sorting by transmitting cortical tension between neighboring cells (Maitre *et al.*, 2012).

Secondly, rearrangements amongst morphogenetically active cells also require cadherin-based cell–cell junctions and actomyosin contractility (Guillot and Lecuit, 2013). These types of morphogenetic movements include convergent-extension (Bertet *et al.*, 2004; Blankenship *et al.*, 2006; Sawyer *et al.*, 2011) and cell invagination (Roh-Johnson *et al.*, 2012). These require cadherins to transmit the contractile forces that are the basis for cell-upon-cell locomotion. Another example of cell-upon-cell movement is the process of cell extrusion (Gu and Rosenblatt, 2012), where individual cells are expelled from epithelial monolayers. Extrusion is associated with diverse processes, including apoptosis (Rosenblatt *et al.*, 2001) and the expression of oncogenes (Hogan *et al.*, 2009). While there are likely to be important mechanistic differences between these different forms of cell extrusion, oncogenic extrusion requires cadherin-based cell–cell adhesion in the neighboring cells (Hogan *et al.*, 2009; Wu *et al.*, 2014), while cadherins at the interface between expelled and neighboring cells are cleaved during the extrusion process (Grieve and Rabouille, 2014). As well, contractile events in both the extruded cells and their neighbors drive extrusion (Slattum *et al.*, 2009; Wu *et al.*, 2014; Hogan *et al.*, 2009). This reflects both activation of myosin and also regulation of the actin cytoskeleton to increase the contractile stress that is generated at the cell–cell interface where extrusion occurs (Wu *et al.*, 2014).

Finally, coupling of cadherin adhesion to contractility also contributes to morphogenetic shaping of whole tissues. This is exemplified by the process of apical constriction (Sawyer *et al.*, 2010), where contractility at the apical poles of epithelial cells allows tissues to be bent and folded, leading to the formation of tissue furrows and invaginations. Physical coupling of cadherins to actomyosin is necessary to transmit the forces that drive invagination (Sawyer *et al.*, 2009; Martin *et al.*, 2009). Of note, coupling of actomyosin to adherens junctions can be a developmentally regulated process that initiates apical constriction (Roh-Johnson *et al.*, 2012); this illustrates the potential for the integration of cadherins and actomyosin to be rate-limiting during morphogenesis. Similarly, cadherin junctions integrate patterns of cellular contractility to yield anisotropic tissue-level patterns of tension during development (Martin *et al.*, 2010). To date, the contribution of cadherins to generating morphogenetic forces has principally focused on the capacity of cadherins to passively couple the contractile apparatuses of cells together. However, as we have discussed, cadherin adhesion can also promote the biogenesis of junctional actomyosin by promoting actin assembly and supporting Rho associated protein kinase (ROCK) signaling (Ratheesh *et al.*, 2012), while cortical stiffening elicited by cadherin mechanotransduction is likely to influence morphogenesis (le Duc *et al.*, 2010). It remains to be determined whether this active contribution of cadherins to cellular contractility participates in morphogenesis.

Cell–Cell Adhesion and the Intermediate Filament Cytoskeleton

IFs associate with desmosomes (Figure 2), specialized cell–cell junctions that are found in many solid tissues, including both

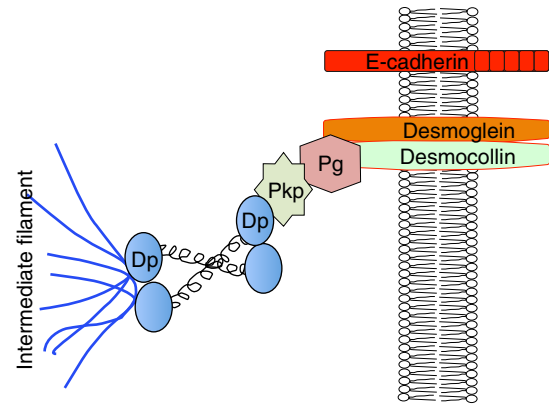


Figure 2 Association of desmosomes with IFs. Desmosomes comprise desmosomal cadherins (desmogleins and desmocollins) that associate with Pg and members of the Pkp family, which anchor Dp to bind IFs.

simple and stratified epithelia and cardiac muscle. An analogous junction that interacts with IF in endothelia is called the complexus adhaerentes and shares molecular similarities with desmosomes. IFs are believed to provide mechanical stability for cells (Gruenbaum and Aebi, 2014) and coupling of IF to desmosomes can then extend mechanical stabilization by IF to the tissue level. Indeed, as we shall see, the functional impact of linking desmosomes to IF is especially apparent in tissues that are subject to mechanical stress, such as the epidermis and cardiac muscle.

Integrating IFs and Desmosomes

Adhesion by desmosomes is mediated by distinct members of the cadherin superfamily. There are two classes of desmosomal cadherins, desmogleins and desmocollins, and each of these classes contains several isoforms (Delva *et al.*, 2009). The specific desmogleins and desmocollins that are expressed in cells vary with their tissue of origin. Nonetheless, mature, functionally competent desmosomes require at least one member from each of the desmosomal cadherin family, leading to the suggestion that they may involve heterodimers of desmocollins and desmogleins.

The cytoplasmic tails of the desmosomal cadherins associate with a range of cytoplasmic proteins to form plaques that link desmosomes to IF (Johnson *et al.*, 2014; Figure 2). This physical linkage is mediated by two sets of proteins (Nekrasova and Green, 2013). First, proteins of the armadillo family, plakophilins (Pkp) and plakoglobin (Pg), interact directly with the desmosomal cadherins. These armadillo family proteins then anchor desmoplakin (Dp), which can associate directly with IFs, thereby completing the physical association between desmosomal cadherin receptors and IF. Interestingly, Dp is common to all tissues where cell–cell adhesion is linked to IF (Johnson *et al.*, 2014). In contrast, the desmosomal cadherin isoforms and armadillo plaque proteins that are used for adhesion vary significantly between tissues.

The C-terminus of Dp mediates its interaction with IF. Of note, IF proteins show great diversity, being encoded by a large (70 member) family of genes with a common domain

organization, but highly divergent primary sequences whose complexity is further increased by alternative splicing (Herrmann *et al.*, 2009). IF proteins are differentially expressed during development and in different tissues. To date, desmosomes have been reported to interact with a number of the major IF proteins, including keratin, desmin, and vimentin (Green and Simpson, 2007). For example, in cardiac myocytes desmosomes associate with desmin, whereas complexus adherente in endothelia interact with vimentin (Getsios *et al.*, 2004). This degree of cell type specificity may reflect the major IF genes that are expressed, although additional levels of regulation are possible.

Physiological Functions of Desmosomal-IF Integration

The necessary role that desmosomes play in maintaining tissue integrity is well illustrated by diseases that target desmosomal proteins in stratified epithelia of the skin and mucosal membranes. For example, patients with the skin diseases pemphigus vulgaris and pemphigus foliaceus produce autoantibodies that target desmogleins 1 and 3 (Green and Simpson, 2007). These inhibit desmosomal adhesion causing loss of cell–cell adhesion (acantholysis) that manifests as blistering of the skin and mucosae. One mechanism for these autoantibodies to disrupt adhesion is by inducing the down-regulation of Dsg expression on the cell surface through endocytosis (Delva *et al.*, 2008). Human genetic diseases also provide compelling evidence that their interaction with IF is necessary for desmosomes to support tissue integrity. For example, severe acantholysis bullosa and extensive skin erosion were reported to be due to compound heterozygous mutations that truncated the IF-binding C-terminus of Dp (Delva *et al.*, 2008). Similarly, mutations in IF genes are associated with skin blistering and loss of adhesion. This implies that integration of desmosomes and IF is necessary to allow the epidermis to resist mechanical stresses.

Several distinctive biophysical properties of desmosomal adhesion and IF may cooperate to confer resistance to stress. Firstly, desmosomal cadherins display the phenomenon of 'hyper-adhesion' (Garrod and Tabernero, 2014). Whereas initial adhesion by desmosomes requires extracellular calcium, as desmosomes mature they display increased stability and adhesive strength and become resistant to the removal of calcium (Kimura *et al.*, 2007). In contrast, classical cadherins do not display 'hyper-adhesion' and remain dependent on extracellular calcium. Hyper-adhesion is a regulated process that responds to Protein Kinase C- α (PKC- α), as desmosomes remain hyper-adhesive when PKC- α is depleted (Johnson *et al.*, 2014).

IF also display distinctive biophysical properties that allow them to resist mechanical stress (Herrmann *et al.*, 2007). Firstly, individual IFs are soft and flexible, in contrast to actin filaments and MTs which readily break upon stress. However, high concentrations and cross-linking can significantly increase the stiffness of the IF network, into the MPa–GPa range, as found in skin and hair, respectively. Secondly, IFs are much more extensible than other cytoskeletal elements, single keratin filaments being able to be stretched up to 3.5-fold before breaking. Finally, IF undergo hardening when strain is applied, a property that is unique to IF and confers greater resistance to breakage than actin filaments or MTs. It seems likely that these physical

properties of IF cooperate with hyper-adhesion by desmosomes to confer a unique capacity to resist mechanical stresses.

Finally, in addition to conferring resistance to mechanical stress, desmosomes also influence cell proliferation and differentiation by regulating cell signaling pathways (Green and Simpson, 2007). These effects have principally been studied by manipulating the functional of desmosomal cadherins and their associated proteins. However, it is possible that association with IF also contributes by regulating desmosomal assembly and dynamics (Getsios *et al.*, 2004). For example, keratins modulate Dp phosphorylation by PKC- α to influence desmosomal stability (Johnson *et al.*, 2014).

Cell–Cell Adhesion and MTs

In contrast to the well-established contributions of F-actin and IF to cell–cell adhesion, the interrelationship between cell adhesion and the microtubule cytoskeleton is an emerging area. MTs are dynamic, polarized polymers of tubulin. The polarity of MTs is reflected in their free ends (typically called the plus and minus ends), which differ in dynamics and interacting proteins; polarity also encompasses the MT lattice itself to allow the directed movement of motor proteins with a preference to move to either the plus or minus end. Many MTs are anchored by their minus ends at the perinuclear centrosome (called centrosomal MTs) and extend with their plus-ends directed outwards. In addition, non-centrosomal MTs are also found, which are not anchored at the centrosomes.

Linkages between MTs and Cell–Cell Junctions

MTs can interact with both classical cadherins at adherens junctions and with desmosomes (Figure 3). Non-centrosomal MT have been found to concentrate at cell–cell junctions both in simple and stratified epithelia (Bacallao *et al.*, 1989; Lechler and Fuchs, 2007; Stehbens *et al.*, 2006; Bellett *et al.*, 2009). In simple polarized epithelia, these can run in an apical-to-basal direction, with their plus-ends oriented basally (Bacallao *et al.*, 1989). In contrast, MTs appear to align with the X–Y plane of the junctions in the flattened stratified squamous epithelial cells of the epidermis (Lechler and Fuchs, 2007). As well, populations of MTs, some of which may be anchored at centrosomes, have been identified which extend with their plus-ends directed toward the junctional cortex (Stehbens *et al.*, 2006). The precise pattern of MTs that interacts with junctions may depend on the developmental context of the cells. For example, as cultured epithelial cells first make contacts with one another, centrosomal arrays tend to predominate, extending their plus-ends toward the junctions, to be replaced by non-centrosomal junctional arrays as the monolayers mature and polarize (Bellett *et al.*, 2009). Similarly, centrosomal arrays predominate in the basal layer of the skin to be replaced by non-centrosomal junctional MTs in the suprabasal layers (Lechler and Fuchs, 2007). What controls these changes in organization of MTs relative to junctions is not yet understood, but may reflect the impact of cell polarization on cytoskeletal organization (Bryant and Mostov, 2008).

These different patterns of association between MTs and junctions are likely to reflect the different classes of MT-

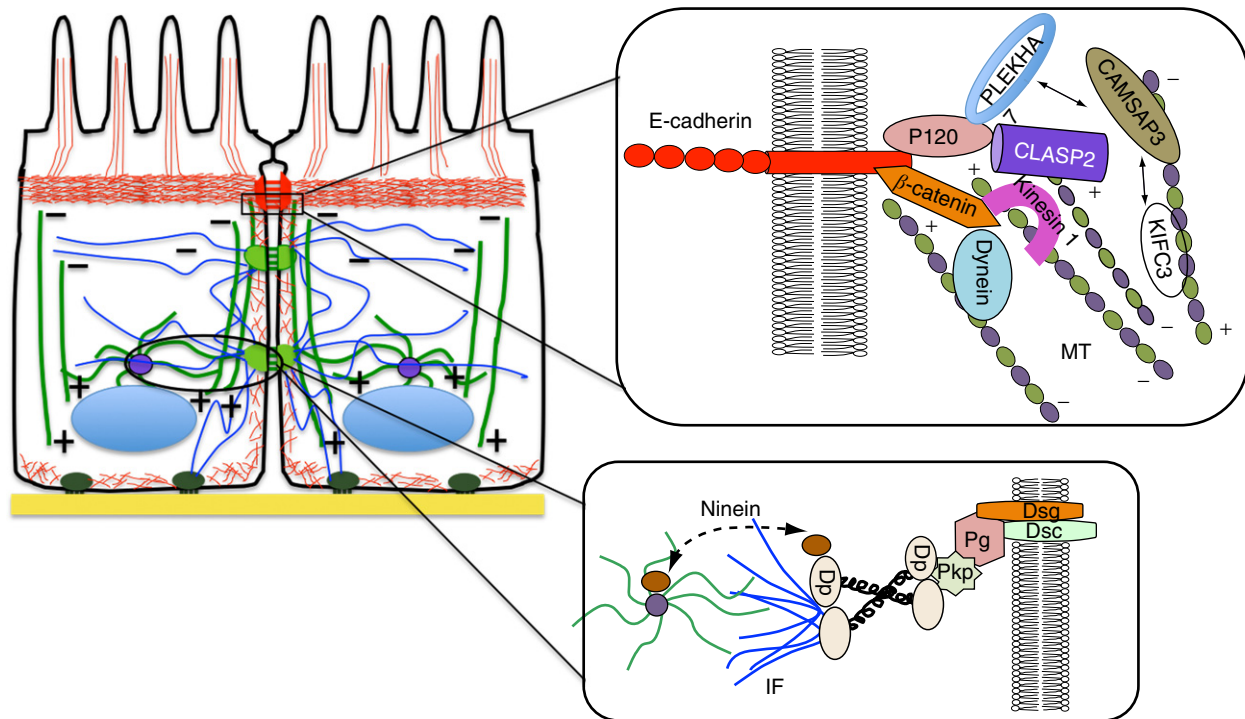


Figure 3 Interactions between MTs in epithelial cells. MTs are organized as centrosomal and non-centrosomal arrays which can interact with cell-cell junctions. At adherens junctions, classical cadherin/catenin complexes can interact with a range of MT-binding proteins, such as CAMSAP3/Nezha, CLASP2, kinesins, and dynein. Desmosomal cadherins can recruit MT-binding proteins such as CLIP190 and ninein.

binding proteins that can interact with either classical cadherin adhesions or desmosomes (Figure 3). Desmosomes can recruit the minus-end binding protein, ninein, through a biochemical interaction with desmoplakin (Lechler and Fuchs, 2007). Ninein is also found at centrosomes but relocalizes to accumulate at cell-cell junctions as epidermal cells differentiate. This coincides with the accumulation of non-centrosomal MTs at junctions (Lechler and Fuchs, 2007). Dynein, a minus-end directed motor, can be found at E-cadherin junctions, where it associates with β -catenin (Ligon *et al.*, 2001). Calmodulin-Regulated Spectrin-Associated Protein 3 (CAMSAP3, originally called Nezha) is a minus-end binding protein that is recruited to E-cadherin adhesions in simple epithelial cells by a complex that consists of the scaffolding protein, Pleckstrin homology domain-containing family A member 7 (PLEKHA7), which binds directly to p120-catenin (Meng *et al.*, 2008). In keratinocytes p120-catenin also binds directly to cytoplasmic linker associated protein 2 (CLASP2), a plus-end binding protein, providing a mechanism to capture MT plus-ends at adherens junctions (Shahbazi *et al.*, 2013).

MT-binding proteins can influence the interactions between MT and junction by mediating physical associations and/or influencing MT dynamics. For example, when anchored to E-cadherin, dynein can potentially bind to MTs and exert force to anchor them to the junctions (Ligon *et al.*, 2001). On the other hand, both CAMSAP3 and ninein can anchor and stabilize the minus ends of MTs (Meng *et al.*, 2008; Lechler and Fuchs, 2007). As MTs grow preferentially from their plus-ends, anchorage of their minus ends at junctions can serve to determine where they assemble and thereby influence their junctional accumulation.

Functional Contributions of MTs at Junctions

MT integrity and dynamics are necessary for the integrity of cell-cell adhesive junctions. The integrity of the ZA in simple epithelial cells is compromised when MTs are depolymerized with drugs, such as nocodazole; when the dynamic behavior of their plus-ends is inhibited, either using low concentrations of nocodazole or by inhibiting proteins that associate with dynamic plus-ends; and when minus-end anchorage via CAMSAP3 is perturbed (Stehbens *et al.*, 2006; Meng *et al.*, 2008). Similarly, the ability to form stable desmosomes was compromised in keratinocytes depleted of LIS1 (Sumigra *et al.*, 2011). In the latter case, the defect in desmosomal biogenesis was accompanied by a compromised epithelial barrier and terminal differentiation of keratinocytes. These functional effects may reflect several ways in which interactions with MTs can affect cell-cell junctions.

MT-dependent cellular transport

MTs are major mechanisms for the active transport of vesicles, organelles and molecular complexes. Such transport is mediated by the MT-dependent motors, dynein and kinesins. Therefore, the anchorage or assembly of MTs at junctions could provide a mechanism to facilitate the transport of cortical and membrane components to the junctions. Consistent with this, the minus-end directed motor, Kinesin superfamily protein 3 (KIFC3), accumulated at junctions in epithelial cells and this required anchorage of MTs at their minus ends via CAMSAP3 and PLEKHA7 (Meng *et al.*, 2008). Further, many specialized junctional components, including classical cadherins, desmosomal cadherins, and gap junction elements, are

transported to cell–cell junctions by MT-dependent motors (Bryant and Stow, 2004; Shaw *et al.*, 2007; Nekrasova and Green, 2013; Chen *et al.*, 2003). This can facilitate both the assembly of junctions and their maintenance, as junctional components undergo endocytosis and turnover in established epithelia (Bryant and Stow, 2004). Indeed, junctions are depleted when their transporting MT-based motors are inhibited. Thus, the delivery of junctional components represents one mechanism for MTs to support cell–cell junctions.

Alternatively, MT-based transport can indirectly affect junctional integrity by regulating the turnover of junctional components. For example, KIFC3 can transport Ubiquitin-specific protease (USP) 47 to junctions; this enzyme catalyzes the removal of ubiquitin from proteins and may protect E-cadherin from degradation (Sako-Kubota *et al.*, 2014).

Regulation of cortical signaling

MTs can also influence cortical signaling by diverse pathways. One example involves the Rho GTPase, although the impact of MTs differs depending on the biological context. Epithelial cell–cell junctions are sites where Rho signaling is activated to stabilize adhesion and promote contractility (Ratheesh *et al.*, 2012). In cultured mammalian epithelial cells, dynamic MTs can support junctional Rho signaling (Ratheesh *et al.*, 2012), whereas in embryonic *Drosophila* epithelia dynamic MTs appear to antagonize Rho signaling (Bulgakova *et al.*, 2013). This appears to reflect different impacts of MTs upon the molecular machinery that regulates Rho signaling. In mammalian cells, dynamic MTs facilitate the junctional recruitment of the centralspindlin complex, a cytokinetic regulator that is also found at junctions of interphase epithelial cells. Centralspindlin promotes Rho signaling by recruiting and activating the Rho Guanine nucleotide exchange factors (GEF), Epithelial cell transforming sequence 2 protein (Ect2) (Ratheesh *et al.*, 2012). In contrast, in the embryonic *Drosophila* epithelium, dynamic MTs appeared to antagonize another Rho activator, DRho-GEF2, thereby reducing Rho activity (Bulgakova *et al.*, 2013). By another mechanism, MTs can sequester the Rho GEF-H1 preventing it from stimulating Rho; however, MT depolymerization can stimulate Rho signaling when GEF-H1 is released (Krendel *et al.*, 2002). Understanding the basis for these context-dependent differences will therefore rely upon elucidating how dynamic MTs control these upstream regulators of Rho.

Concluding Comments

In summary, cytoskeletal connections play intimate roles in the biology of cell–cell adhesions. It is useful to consider cell–cell junctions as the product of integrating adhesion with the cytoskeleton. As we have illustrated in our current discussion, such integration involves the coordination of dynamic adhesive and cytoskeletal systems. While we have focused on the individual cytoskeletal systems and their associated adhesive junctions, it is important to emphasize that cross talk between these systems can occur at many levels. Firstly, cell signaling can link different junctional and cytoskeletal systems to one another. For example, regulation of cortical Rho signaling by dynamic MTs can affect actomyosin-based contractility at cell–cell junctions (Ratheesh *et al.*, 2012). Secondly, regulation

of junctional actin and its coupling to cadherins can potentially recruit other adhesion systems. As an example, classical cadherins can recruit nectins into maturing adherens junctions, through their capacity to interact with α -catenin (Trojanovsky *et al.*, 2015). Finally, different cytoskeletal networks may be physically and/or functionally linked together. For example, formins can regulate both actin and MT dynamics (Palazzo *et al.*, 2001); proteins like spectroplakins can link MTs, MFs, and IFs (Huelsmann and Brown, 2014); and IFs are implicated in force sensing by classical cadherins (Weber *et al.*, 2012). How these, and other modes, of cross talk influence the integration of cell–cell adhesion and the cytoskeleton will be an interesting topic of research for the future.

See also: Cytoskeleton and Motors: Cytoskeletal Components: Actin Assembly Dynamics and Its Regulation in Motile and Morphogenetic Processes. Cell Communication: Cellular Machineries: The Molecular Architecture of Cell–Cell Adhesions

References

- Abe, K., Chisaka, O., Vanroy, F., Takeichi, M., 2004. Stability of dendritic spines and synaptic contacts is controlled by a N-catenin. *Nature Neuroscience* 4, 357–363.
- Abe, K., Takeichi, M., 2008. EPLIN mediates linkage of the cadherin catenin complex to F-actin and stabilizes the circumferential actin belt. *Proceedings of the National Academy of Sciences of the United States of America* 105, 13–19.
- Bacallao, R., Antony, C., Karsenti, E., Stelzer, E., Simons, K., 1989. The subcellular organization of Madin-Darby canine kidney cells during the formation of a polarized epithelium. *Journal of Cell Biology* 109, 2817–2832.
- Backouche, F., Haviv, L., Groswasser, D., Bernheim-Groswasser, A., 2006. Active gels: Dynamics of patterning and self-organization. *Physical Biology* 3, 264–273.
- Bellett, G., Carter, J.M., Keynton, J., *et al.*, 2009. Microtubule plus-end and minus-end capture at adherens junctions is involved in the assembly of apico-basal arrays in polarised epithelial cells. *Cell Motility and the Cytoskeleton* 66, 893–908.
- Bertet, C., Sulak, L., Lecuit, T., 2004. Myosin-dependent junction remodelling controls planar cell intercalation and axis elongation. *Nature* 429, 667–671.
- Blankenship, J.T., Backovic, S.T., Sanny, J.S., Weitz, O., Zallen, J.A., 2006. Multicellular rosette formation links planar cell polarity to tissue morphogenesis. *Developmental Cell* 11, 459–470.
- Borghi, N., Sorokina, M., Shcherbakova, O.G., *et al.*, 2012. E-cadherin is under constitutive actomyosin-generated tension that is increased at cell–cell contacts upon externally applied stretch. *Proceedings of the National Academy of Sciences of the United States of America* 109, 12568–12573.
- Brasch, J., Harrison, O.J., Honig, B., Shapiro, L., 2012. Thinking outside the cell: How cadherins drive adhesion. *Trends in Cell Biology* 22, 299–310.
- Bresnick, A.R., 1999. Molecular mechanisms of nonmuscle myosin-II regulation. *Current Opinion in Cell Biology* 11, 26–33.
- Bryant, D.M., Mostov, K.E., 2008. From cells to organs: Building polarized tissue. *Nature Reviews Molecular Cell Biology* 9, 887–901.
- Bryant, D.M., Stow, J.L., 2004. The ins and outs of E-cadherin trafficking. *Trends in Cell Biology* 14, 427–434.
- Buckley, C.D., Tan, J.L., Anderson, K.L., *et al.*, 2014. The minimal cadherin-catenin complex binds to actin filaments under force. *Science* 346 (6209), 1254211.
- Bulgakova, N.A., Grigoriev, I., Yap, A.S., Akhmanova, A., Brown, N.H., 2013. Dynamic microtubules produce an asymmetric E-cadherin-Bazooka complex to maintain segment boundaries. *Journal of Cell Biology* 201, 887–901.
- Carramusa, L., Ballestrem, C., Zilberman, Y., Bershadsky, A.D., 2007. Mammalian diaphanous-related formin Dia1 controls the organization of E-cadherin-mediated cell–cell junctions. *Journal of Cell Science* 120, 3870–3882.
- Chazeau, A., Garcia, M., Czondor, K., *et al.*, 2015. A mechanical coupling between trans-synaptic N-cadherin adhesions and the actin flow stabilizes dendritic spines. *Molecular Biology of the Cell* 26 (5), 859–873.

- Chen, X., Kojima, S., Borisy, G.G., Green, K.J., 2003. p120 Catenin associates with kinesin and facilitates the transport of cadherin-catenin complexes to intercellular junctions. *Journal of Cell Biology* 163, 547–557.
- Chesarone, M.A., Goode, B.L., 2009. Actin nucleation and elongation factors: Mechanisms and interplay. *Current Opinion in Cell Biology* 21, 28–37.
- Conway, D.E., Breckenridge, M.T., Hinde, E., *et al.*, 2013. Fluid shear stress on endothelial cells modulates mechanical tension across VE-cadherin and PECAM-1. *Current Biology* 23, 1024–1030.
- Delva, E., Jennings, J.M., Calkins, C.C., *et al.*, 2008. Pemphigus vulgaris IgG-induced desmoglein-3 endocytosis and desmosomal disassembly are mediated by a clathrin- and dynamin-independent mechanism. *Journal of Biological Chemistry* 283, 18303–18313.
- Delva, E., Tucker, D.K., Kowalczyk, A.P., 2009. The desmosome. *Cold Spring Harbor Perspectives in Biology* 1, a002543.
- Drees, F., Pokutta, S., Yamada, S., Nelson, W.J., Weis, W.I., 2005. Alpha-catenin is a molecular switch that binds E-cadherin-beta-catenin and regulates actin-filament assembly. *Cell* 123, 903–915.
- le Duc, Q., Shi, Q., Blonk, I., *et al.*, 2010. Vinculin potentiates E-cadherin mechanosensing and is recruited to actin-anchored sites within adherens junctions in a myosin II-dependent manner. *Journal of Cell Biology* 189, 1107–1115.
- Fagotto, F., 2015. Regulation of cell adhesion and cell sorting at embryonic boundaries. *Current Topics in Developmental Biology* 112, 19–64.
- Fernandez-Gonzalez, R., Simoes Sde, M., Roper, J.C., Eaton, S., Zallen, J.A., 2009. Myosin II dynamics are regulated by tension in intercalating cells. *Developmental Cell* 17, 736–743.
- Garrod, D., Taberner, L., 2014. Hyper-adhesion: A unique property of desmosomes. *Cell Communication and Adhesion* 21, 249–256.
- Getsios, S., Huen, A.C., Green, K.J., 2004. Working out the strength and flexibility of desmosomes. *Nature Reviews Molecular Cell Biology* 5, 271–281.
- Gomez, G.A., Mclachlan, R.W., Yap, A.S., 2011. Productive tension: Force-sensing and homeostasis of cell-cell junctions. *Trends in Cell Biology* 21, 499–505.
- Green, K.J., Simpson, C.L., 2007. Desmosomes: New perspectives on a classic. *Journal of Investigative Dermatology* 127, 2499–2515.
- Grieve, A.G., Rabouille, C., 2014. Extracellular cleavage of E-cadherin promotes epithelial cell extrusion. *Journal of Cell Science* 127, 3331–3346.
- Gruenbaum, Y., Aebi, U., 2014. Intermediate filaments: A dynamic network that controls cell mechanics. *F1000Prime Reports* 6, 54.
- Gu, Y., Rosenblatt, J., 2012. New emerging roles for epithelial cell extrusion. *Current Opinion in Cell Biology* 24, 865–870.
- Guillot, C., Lecuit, T., 2013. Mechanics of epithelial tissue homeostasis and morphogenesis. *Science* 340, 1185–1189.
- Han, S.P., Gambin, Y., Gomez, G.A., *et al.*, 2014. Cortactin scaffolds Arp2/3 and WAVE2 at the epithelial zonula adherens. *Journal of Biological Chemistry* 289, 7764–7775.
- Harrison, O.J., Jin, X., Hong, S., *et al.*, 2011. The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins. *Structure* 19, 244–256.
- Herrmann, H., Bar, H., Kreplak, L., Strelkov, S.V., Aebi, U., 2007. Intermediate filaments: From cell architecture to nanomechanics. *Nature Reviews Molecular Cell Biology* 8, 562–573.
- Herrmann, H., Strelkov, S.V., Burkhard, P., Aebi, U., 2009. Intermediate filaments: Primary determinants of cell architecture and plasticity. *Journal of Clinical Investigation* 119, 1772–1783.
- Hogan, C., Dupre-Crochet, S., Norman, M., *et al.*, 2009. Characterization of the interface between normal and transformed epithelial cells. *Nature Cell Biology* 11, 460–467.
- Hong, S., Troyanovsky, R.B., Troyanovsky, S.M., 2013. Binding to F-actin guides cadherin cluster assembly, stability, and movement. *Journal of Cell Biology* 201, 131–143.
- Huelsmann, S., Brown, N.H., 2014. Spectraplakins. *Current Biology* 24, R307–R308.
- Huveneers, S., Oldenburg, J., Spanjaard, E., *et al.*, 2012. Vinculin associates with endothelial VE-cadherin junctions to control force-dependent remodeling. *Journal of Cell Biology* 196, 641–652.
- Iino, R., Koyama, I., Kusumi, A., 2001. Single molecule imaging of green fluorescent proteins in living cells: E-cadherin forms oligomers on the free cell surface. *Biophysical Journal* 80, 2667–2677.
- Insall, R.H., Machesky, L.M., 2009. Actin dynamics at the leading edge: From simple machinery to complex networks. *Developmental Cell* 17, 310–322.
- Izumi, G., Sakisaka, T., Baba, T., *et al.*, 2004. Endocytosis of E-cadherin regulated by Rac and Cdc42 small G proteins through IQGAP1 and actin filaments. *Journal of Cell Biology* 166, 237–248.
- Jaffe, A.B., Hall, A., 2005. Rho GTPases: Biochemistry and biology. *Annual Review of Cell and Developmental Biology* 21, 247–269.
- Johnson, J.L., Najor, N.A., Green, K.J., 2014. Desmosomes: Regulators of cellular signaling and adhesion in epidermal health and disease. *Cold Spring Harbor Perspectives in Medicine* 4, a015297.
- Kimura, T.E., Merritt, A.J., Garrod, D.R., 2007. Calcium-independent desmosomes of keratinocytes are hyper-adhesive. *Journal of Investigative Dermatology* 127, 775–781.
- Kobielak, A., Pasolli, H.A., Fuchs, E., 2004. Mammalian formin-1 participates in adherens junctions and polymerization of linear actin cables. *Nature Cell Biology* 6, 21–30.
- Kovacs, E.M., Ali, R.G., McCormack, A.J., Yap, A.S., 2002a. E-cadherin homophilic ligation directly signals through Rac and PI3-kinase to regulate adhesive contacts. *Journal of Biological Chemistry* 277, 6708–6718.
- Kovacs, E.M., Goodwin, M., Ali, R.G., Paterson, A.D., Yap, A.S., 2002b. Cadherin-directed actin assembly: E-cadherin physically associates with the Arp2/3 complex to direct actin assembly in nascent adhesive contacts. *Current Biology* 12, 379–382.
- Kovacs, E.M., Verma, S., Ali, R.G., *et al.*, 2011. N-WASP regulates the epithelial junctional actin cytoskeleton through a non-canonical post-nucleation pathway. *Nature Cell Biology* 13, 934–943.
- Krendel, M., Zenke, F.T., Bokoch, G.M., 2002. Nucleotide exchange factor GEF-H1 mediates cross-talk between microtubules and the actin cytoskeleton. *Nature Cell Biology* 4, 294–301.
- Krieg, M., Arboleda-Estudillo, Y., Puech, P.H., *et al.*, 2008. Tensile forces govern germ-layer organization in zebrafish. *Nature Cell Biology* 10, 429–436.
- Kusumi, A., Nakada, C., Ritchie, K., *et al.*, 2005. Paradigm shift of the plasma membrane concept from the two-dimensional continuum fluid to the partitioned fluid: High-speed single-molecule tracking of membrane molecules. *Annual Review of Biophysics and Biomolecular Structure* 34, 351–378.
- Kwiatkowski, A.V., Maiden, S.L., Pokutta, S., *et al.*, 2010. *In vitro* and *in vivo* reconstitution of the cadherin-catenin-actin complex from *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America* 107, 14591–14596.
- Ladoux, B., Anon, E., Lambert, M., *et al.*, 2010. Strength dependence of cadherin-mediated adhesions. *Biophysical Journal* 98, 534–542.
- Lambert, M., Choquet, D., Mege, R.M., 2002. Dynamics of ligand-induced, Rac1-dependent anchoring of cadherins to the actin cytoskeleton. *Journal of Cell Biology* 157, 469–479.
- Lechler, T., Fuchs, E., 2007. Desmoplakin: An unexpected regulator of microtubule organization in the epidermis. *Journal of Cell Biology* 176, 147–154.
- Leerberg, J.M., Gomez, G.A., Verma, S., *et al.*, 2014. Tension-sensitive actin assembly supports contractility at the epithelial zonula adherens. *Current Biology* 24, 1689–1699.
- Ligon, L.A., Karki, S., Tokito, M., Holzbaur, E.L., 2001. Dynein binds to beta-catenin and may tether microtubules at adherens junctions. *Nature Cell Biology* 3, 913–917.
- Liu, Z., Tan, J.L., Cohen, D.M., *et al.*, 2010. Mechanical tugging force regulates the size of cell-cell junctions. *Proceedings of the National Academy of Sciences of the United States of America* 107, 9944–9949.
- Maddugoda, M.P., Crampton, M.S., Shewan, A.M., Yap, A.S., 2007. Myosin VI and vinculin cooperate during the morphogenesis of cadherin cell-cell contacts in mammalian epithelial cells. *Journal of Cell Biology* 178, 529–540.
- Maiden, S.L., Hardin, J., 2011. The secret life of alpha-catenin: Moonlighting in morphogenesis. *Journal of Cell Biology* 195, 543–552.
- Maitre, J.L., Berthoumieux, H., Krens, S.F., *et al.*, 2012. Adhesion functions in cell sorting by mechanically coupling the cortices of adhering cells. *Science* 338, 253–256.
- Martin, A.C., Gelbart, M., Fernandez-Gonzalez, R., Kaschube, M., Wieschaus, E., 2010. Integration of contractile forces during tissue invagination. *Journal of Cell Biology* 188, 735–749.
- Martin, A.C., Kaschube, M., Wieschaus, E.F., 2009. Pulsed contractions of an actin-myosin network drive apical constriction. *Nature* 457, 495–499.
- Meng, W., Mushika, Y., Ichii, T., Takeichi, M., 2008. Anchorage of microtubule minus ends to adherens junctions regulates epithelial cell-cell contacts. *Cell* 135, 948–959.
- Morone, N., Fujiwara, T., Murase, K., *et al.*, 2006. Three-dimensional reconstruction of the membrane skeleton at the plasma membrane interface by electron tomography. *Journal of Cell Biology* 174, 851–862.
- Nekrasova, O., Green, K.J., 2013. Desmosome assembly and dynamics. *Trends in Cell Biology* 23, 537–546.
- Niessen, C., Leckband, D., Yap, A.S., 2011. Tissue organization by classical cadherin adhesion molecules: Dynamic molecular and cellular mechanisms of morphogenetic regulation. *Physiological Reviews* 91, 691–731.
- Niessen, C.M., Gumbiner, B.M., 2002. Cadherin-mediated cell sorting not determined by binding or adhesion specificity. *Journal of Cell Biology* 156, 389–400.

- Palazzo, A.F., Cook, T.A., Alberts, A.S., Gundersen, G.G., 2001. mDia mediates Rho-regulated formation and orientation of stable microtubules. *Nature Cell Biology* 3, 723–729.
- Priya, R., Yap, A.S., Gomez, G.A., 2013. E-cadherin supports steady-state Rho signaling at the epithelial zonula adherens. *Differentiation* 86, 133–140.
- Pruitt, B.L., Dunn, A.R., Weis, W.I., Nelson, W.J., 2014. Mechano-transduction: From molecules to tissues. *PLOS Biology* 12, e1001996.
- Ratheesh, A., Gomez, G.A., Priya, R., *et al.*, 2012. Centralspindlin and alpha-catenin regulate Rho signalling at the epithelial zonula adherens. *Nature Cell Biology* 14, 818–828.
- Rauzi, M., Lenne, P.F., Lecuit, T., 2010. Planar polarized actomyosin contractile flows control epithelial junction remodelling. *Nature* 468, 1110–1114.
- Rimm, D.L., Koslov, E.R., Kebriaei, P., Cianci, C.D., Morrow, J.S., 1995. α_1 (E)-Catenin is an actin-binding and -bundling protein mediating the attachment of F-actin to the membrane adhesion complex. *Proceedings of the National Academy of Sciences of the United States of America* 92, 8813–8817.
- Roh-Johnson, M., Shemer, G., Higgins, C.D., *et al.*, 2012. Triggering a cell shape change by exploiting preexisting actomyosin contractions. *Science* 335, 1232–1235.
- Rosenblatt, J., Raff, M.C., Cramer, L.P., 2001. An epithelial cell destined for apoptosis signals its neighbors to extrude it by an actin- and myosin-dependent mechanism. *Current Biology* 11, 1847–1857.
- Sako-Kubota, K., Tanaka, N., Nagae, S., Meng, W., Takeichi, M., 2014. Minus end-directed motor KIFC3 suppresses E-cadherin degradation by recruiting USP47 to adherens junctions. *Molecular Biology of the Cell* 25 (24), 3851–3860.
- Salbreux, G., Charras, G., Paluch, E., 2012. Actin cortex mechanics and cellular morphogenesis. *Trends in Cell Biology* 22, 536–545.
- Sawyer, J.K., Choi, W., Jung, K.C., *et al.*, 2011. A contractile actomyosin network linked to adherens junctions by Canoe/afadin helps drive convergent extension. *Molecular Biology of the Cell* 22 (14), 2491–2508.
- Sawyer, J.K., Harris, N.J., Slep, K.C., Gaul, U., Peifer, M., 2009. The *Drosophila* afadin homologue Canoe regulates linkage of the actin cytoskeleton to adherens junctions during apical constriction. *Journal of Cell Biology* 186, 57–73.
- Sawyer, J.M., Harrell, J.R., Shemer, G., *et al.*, 2010. Apical constriction: A cell shape change that can drive morphogenesis. *Developmental Biology* 341, 5–19.
- Scott, J.A., Shewan, A.M., Den Elzen, N.R., *et al.*, 2006. Ena/VASP Proteins Can Regulate Distinct Modes of Actin Organization at Cadherin-adhesive Contacts. *Molecular Biology of the Cell* 17, 1085–1095.
- Shahbazi, M.N., Megias, D., Epifano, C., *et al.*, 2013. CLASP2 interacts with p120-catenin and governs microtubule dynamics at adherens junctions. *Journal of Cell Biology* 203, 1043–1061.
- Shaw, R.M., Fay, A.J., Puthenveedu, M.A., *et al.*, 2007. Microtubule plus-end-tracking proteins target gap junctions directly from the cell interior to adherens junctions. *Cell* 128, 547–560.
- Shewan, A.M., Maddugoda, M., Kraemer, A., *et al.*, 2005. Myosin 2 is a key rho kinase target necessary for the local concentration of E-cadherin at cell-cell contacts. *Molecular Biology of the Cell* 16, 4531–4532.
- Slattum, G., McGee, K.M., Rosenblatt, J., 2009. P115 RhoGEF and microtubules decide the direction apoptotic cells extrude from an epithelium. *Journal of Cell Biology* 186, 693–702.
- Smutny, M., Cox, H.L., Leerberg, J.M., *et al.*, 2010. Myosin II isoforms identify distinct functional modules that support integrity of the epithelial zonula adherens. *Nature Cell Biology* 12, 696–702.
- Soares e Silva, M., Depken, M., Stuhmann, B., *et al.*, 2011. Active multistage coarsening of actin networks driven by myosin motors. *Proceedings of the National Academy of Sciences of the United States of America* 108, 9408–9413.
- Stebbens, S.J., Paterson, A.D., Crampton, M.S., *et al.*, 2006. Dynamic microtubules regulate the local concentration of E-cadherin at cell-cell contacts. *Journal of Cell Science* 119, 1801–1811.
- Sumigay, K.D., Chen, H., Lechler, T., 2011. Lis1 is essential for cortical microtubule organization and desmosome stability in the epidermis. *Journal of Cell Biology* 194, 631–642.
- Taguchi, K., Ishiuchi, T., Takeichi, M., 2011. Mechanosensitive EPLIN-dependent remodeling of adherens junctions regulates epithelial reshaping. *Journal of Cell Biology* 194, 643–656.
- Takeichi, M., 1991. Cadherin cell adhesion receptors as a morphogenetic regulator. *Science* 251, 1451–1455.
- Takeichi, M., 2014. Dynamic contacts: Rearranging adherens junctions to drive epithelial remodelling. *Nature Reviews Molecular Cell Biology* 15, 397–410.
- Tang, V.W., Brieher, W.M., 2012. Alpha-Actinin-4/FSGS1 is required for Arp2/3-dependent actin assembly at the adherens junction. *Journal of Cell Biology* 196, 115–130.
- Troyanovsky, R.B., Indra, I., Chen, C.S., Hong, S., Troyanovsky, S.M., 2015. Cadherin controls nectin recruitment into adherens junctions by remodeling the actin cytoskeleton. *Journal of Cell Science* 128, 140–149.
- Ward, M.D., Dembo, M., Hammer, D.A., 1994. Kinetics of cell detachment: Peeling of discrete receptor clusters. *Biophysical Journal* 67, 2522–2534.
- Weber, G.F., Bjerke, M.A., Desimone, D.W., 2012. A mechanoresponsive cadherin-keratin complex directs polarized protrusive behavior and collective cell migration. *Developmental Cell* 22, 104–115.
- Wu, S.K., Gomez, G.A., Michael, M., *et al.*, 2014. Cortical F-actin stabilization generates apical-lateral patterns of junctional contractility that integrate cells into epithelia. *Nature Cell Biology* 16, 167–178.
- Wu, Y., Kanchanawong, P., Zaidel-Bar, R., 2015. Actin-delimited adhesion-independent clustering of E-cadherin forms the nanoscale building blocks of adherens junctions. *Developmental Cell* 32, 139–154.
- Yamada, S., Nelson, W.J., 2007. Localized zones of Rho and Rac activities drive initiation and expansion of epithelial cell-cell adhesion. *Journal of Cell Biology* 178, 517–527.
- Yamada, S., Pokutta, S., Drees, F., Weis, W.I., Nelson, W.J., 2005. Deconstructing the cadherin-catenin-actin complex. *Cell* 123, 889–901.
- Yao, M., Qiu, W., Liu, R., *et al.*, 2014. Force-dependent conformational switch of alpha-catenin controls vinculin binding. *Nature Communications* 5, 4525.
- Yap, A.S., Brieher, W.M., Pruschy, M., Gumbiner, B.M., 1997. Lateral clustering of the adhesive ectodomain: A fundamental determinant of cadherin function. *Current Biology* 7, 308–315.
- Yonemura, S., Wada, Y., Watanabe, T., Nagafuchi, A., Shibata, M., 2010. Alpha-Catenin as a tension transducer that induces adherens junction development. *Nature Cell Biology* 12, 533–542.

Cell Adhesion to the Extracellular Matrix

B Geiger and A Livne, Weizmann Institute of Science, Rehovot, Israel

© 2016 Elsevier Inc. All rights reserved.

Glossary

Integrins A diverse family of adhesion receptors, mediating cell–matrix or cell–cell contacts. Integrins are heterodimeric transmembrane receptors consisting of α and β chains; in combination, they define the specificity of the adhesive interaction at the cell's exterior, and the local interaction with the cytoskeleton within the cells.

Invadopodia Protrusive cell adhesions associated with cancer cell invasion and metastasis. Invadopodia consist of an integrin-based adhesion domain through which they are attached to the matrix, a protrusive, actin-rich domain, and a proteolytic domain, responsible for local degradation of the matrix.

The Extracellular Matrix (ECM)

The ECM is composed of fibrillar networks that serve as scaffolds on which and within which tissue cells live (Hay, 1991; Alberts *et al.*, 2007). Interaction with the ECM affects the survival, proliferation, differentiation, and migration of individual, the so-called ‘anchorage dependent’ cells (Stoker *et al.*, 1968). In intact tissues it supports structural stability, sustains functional homeostasis and regulates the fate and behavior of the constituting cells (Davis and Camarillo, 1995; Menko and Boettiger, 1987). Adhesion-mediated signaling is based on the capacity of cells to sense and respond to changes in the chemical and physical properties of the ECM (Watt and Driskell, 2010) in a wide variety of processes, including embryonic development, tissue remodeling, and wound healing, as well as in disease states such as cancer invasion and metastasis (Egeblad *et al.*, 2010a,b). However, despite the strong evidence for physiological cell-ECM cross-talk *in vivo*, the molecular mechanisms underlying adhesion-dependent signaling, are poorly characterized, most likely due to the chemical and physical diversity and molecular complexity of living organisms (Hynes and Naba, 2012). Thus, many studies addressing adhesion-mediated processes are conducted *ex vivo*, enabling the design and fabrication of matrices with well-defined properties. In this chapter we will primarily address the mechanisms underlying cell adhesion to the ECM in such well-defined culture conditions.

Indeed, the ECM is highly diverse in structure, ranging from loose connective tissue to densely packed tendons, and sheets of basement membrane. Its variability can mainly be attributed to distinct molecular compositions and postdeposition remodeling. Connective tissue fascia and tendons, for example, contain high levels of collagen I, while basement membranes are enriched with collagen V, laminin, and different proteoglycans (Ricard-Blum, 2011; Yurchenco, 2011). Especially interesting is the fact that in addition to scaffolding networks, the ECM is loaded with growth factors (e.g., fibroblast-, transforming- and heparin-binding epidermal growth factors) that associate with specific components of the matrix (e.g., heparin sulfate proteoglycan) and stimulate the nearby adhering cells (Gospodarowicz *et al.*, 1980; Hay, 1991; Hynes, 2009; Munger and Sheppard, 2011; Sarrazin *et al.*, 2011). This process, whereby the ECM serves as a spatially defined signaling environment, is being applied as a powerful tool for culturing tissue cells under close-to-native conditions, *ex vivo*.

Another feature of the ECM that affects the attached cells is its dimensionality. Thus, cells can differentially respond through adhesion to one-dimensional (1D) matrices dominated by elongated fibrils (Doyle *et al.*, 2009); two-dimensional (2D) matrices such as basement membranes; and three-dimensional (3D) matrices, in which the cells are totally or partially embedded in the matrix (Cukierman *et al.*, 2001; Elsdale and Bard, 1972; Nelson and Bissell, 2006). Recent studies have demonstrated that cells can respond to the microtopography or even nano-topography of the surface to which they adhere (Baharloo *et al.*, 2005; Cukierman *et al.*, 2001; Curtis and Wilkinson, 1997; Geblinger *et al.*, 2010; Geiger *et al.*, 2001; Grossner-Schreiber *et al.*, 2006; Vogel *et al.*, 2006). Nano-patterning technologies have also been utilized to test the capacity of cells to recognize particular spacings between adhesive ligand epitopes, thus demonstrating that cells require certain threshold ligand proximities, usually within the range of a few tens of nanometers, for stimulating cell spreading and cytoskeletal reorganization (Cavalcanti-Adam *et al.*, 2007; Geiger *et al.*, 2009; Grinnell and Petroll, 2010; Massia and Hubbell, 1991; Yamada and Cukierman, 2007).

Diversity of ECM Adhesions

Clearly defined adhesions with the ECM are formed by essentially all adherent animal cells, in culture and *in vivo*; yet their molecular composition, subcellular distribution, size, and morphology can be quite heterogeneous. Particularly widespread and important are focal adhesions, usually located near the cell periphery, and associated with bundles of actin microfilaments (Geiger *et al.*, 2009; Figure 1). Focal adhesions share certain features in common: they are oval structures, several square microns in size, that are mediated by integrins, and interact with the actin cytoskeleton at the inner aspects of the membrane. Their development is stimulated by the small GTPase Rho, and requires actomyosin contractility. Characteristic focal adhesion plaque proteins include vinculin, paxillin, talin, and tyrosine-phosphorylated proteins. Interestingly, the molecular composition of the ECM (e.g., fibronectin vs. vitronectin or laminin) and the cellular interactions with these matrices via different integrins (e.g., $\alpha_5\beta_1$ vs. $\alpha_v\beta_3$) lead to the formation of focal adhesions of distinct morphology, size, and subcellular distribution (Figure 2). Focal adhesions are

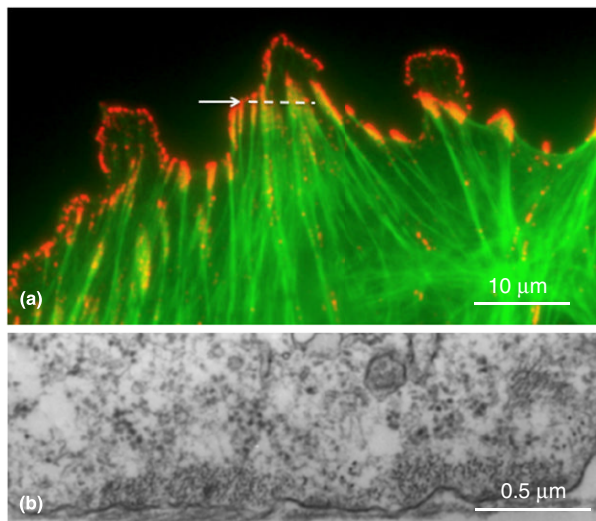


Figure 1 (a) Focal adhesions (red) are usually located at the cell periphery, anchoring contractile actin bundles (green), termed ‘stress fibers,’ to the ECM. (b) Electron microscope image showing a side view of the cell edge (e.g., dashed line in (a)). From bottom to top: ECM (dense patches), cell membrane (continuous line), actin fibers associated with focal adhesions (sparse patches next to cell membrane), microtubules, and cell organelles.

mainly prominent in cultured cells growing on solid surfaces; yet structures with similar molecular properties are also found *in vivo* (e.g., myotendinous junctions, adhesions to the basement membrane). One clear example for *in vivo* analogs of focal adhesions is the dense plaque of smooth muscle tissues, which is the anchorage site of the contractile actin machinery to the plasma membrane. This was demonstrated both by antibody labeling of smooth muscle sections, which displayed an abundance of adhesion components in these sites (North *et al.*, 1993; Winograd-Katz *et al.*, 2014), and by cryo-electron tomography (Bokstad *et al.*, 2012), which revealed, in these sites, specific focal adhesion-associated particles (Patla *et al.*, 2010).

Another form of integrin adhesion is focal complexes, small (around 100 nm in diameter), dot-like structures that are primarily located along the lamellipodium’s edge (Zamir and Geiger, 2001; Figure 3(a)). These sites can be associated with cell migration, or serve as precursors of focal adhesions. Their formation is induced by the Rho family GTPase Rac. In the more central locations of many cell types, ‘fibrillar adhesions’ (known also as ‘ECM contacts’), dot-like or elongated structures associated with ECM fibrils, are also found (Zamir *et al.*, 2000; Figure 3(b)). Fibrillar adhesions are typically associated with ‘soft’ extracellular fibronectin fibrils, mediated by the fibronectin receptor $\alpha_5\beta_1$ integrin, and enriched with the cytoplasmic protein tensin. Fibrillar adhesions emerge from focal adhesions in an actomyosin-dependent fashion. Another form of ECM adhesion is the invadosome, including podosomes and invadopodia, small (around 0.5 μm in diameter) cylindrical structures consisting of typical focal adhesion proteins, including vinculin and paxillin (Linder, 2009; Revach and Geiger, 2014). Podosomes are found in cells such as macrophages and osteoclasts (Figure 4(a)), whereas invadopodia are

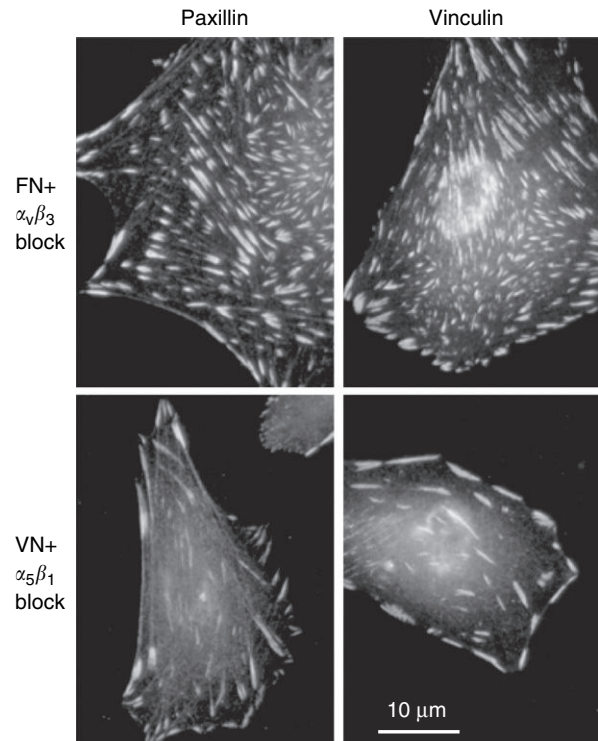


Figure 2 The molecular composition of the ECM (here, the glass surface is coated with fibronectin or vitronectin) and the cellular interaction with these matrices via different integrins ($\alpha_5\beta_1$ and $\alpha_v\beta_3$, respectively) lead to the formation of focal adhesions (stained here for vinculin and paxillin) with distinct morphology, size, and subcellular distribution. Images courtesy of B. Zimmerman.

predominantly found in invasive or metastatic cancer cells (Figure 4(b)). Characteristic proteins, indispensable for invadosome formation are gelsolin and dynamin, both of which localize to the actin cores of these structures.

It is noteworthy that most of the matrix adhesions described above, interact, through their cytoplasmic aspects, with the actin cytoskeleton. Nevertheless, the adhesion of epithelial cells to the underlying basement membrane is also mediated by another class of adhesions, known as hemidesmosomes, which typically associate, within the cell, with the keratin-based intermediate filament system. The primary integrin that mediates these interactions is $\alpha_6\beta_4$, which is accompanied, in some tissues, by additional transmembrane matrix receptors as well (Litjens *et al.*, 2006). The third cytoskeletal network, microtubules, is apparently not associated directly with matrix adhesions, yet it has been shown that microtubules negatively regulate focal adhesion stability (Bershadsky *et al.*, 1996; Kaverina *et al.*, 1999). In addition, podosomes in osteoclasts and other cell types depend on the presence of intact microtubules, and dissociate upon cell treatment with microtubule-disrupting drugs (Destaing *et al.*, 2003, 2005).

The Integrin Adhesome

The molecular components of the adhesion machinery, consisting of over 200 distinct proteins that localize to focal

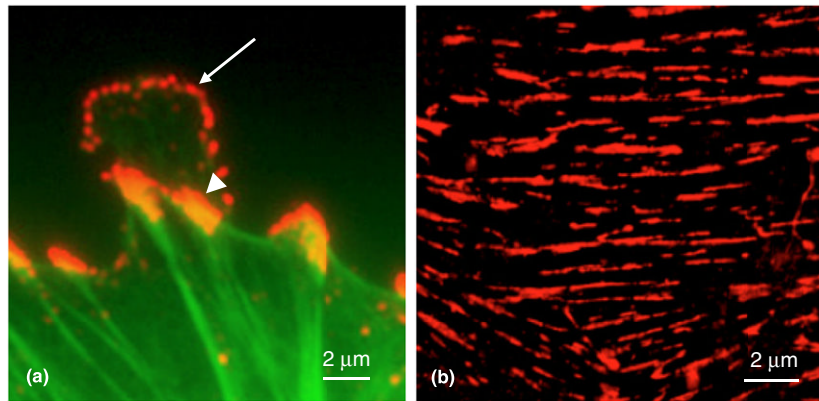


Figure 3 (a) Focal complexes (arrow) are small (~ 100 nm in diameter), dot-like adhesions primarily located near the lamellipodium edges. These sites can serve as precursors of focal adhesions (arrowhead), and may be involved in cell migration. (b) Fibrillar adhesions are dot-like or elongated structures, and associated with ECM fibrils located near the cell center in various cell types.

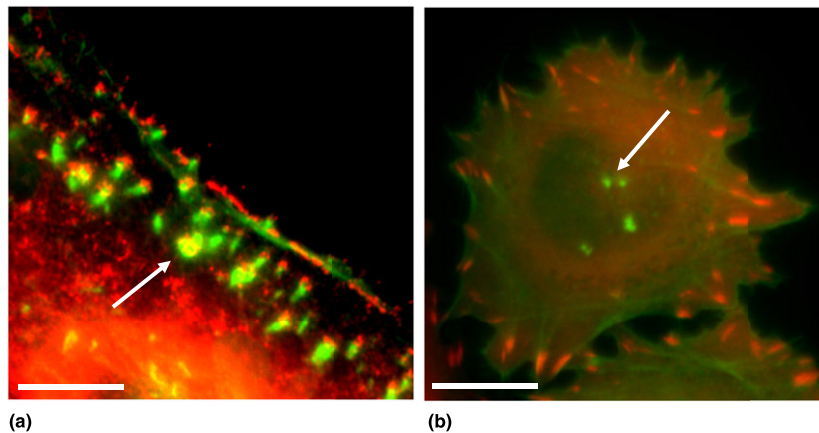


Figure 4 Invadosomes, including podosomes and invadopodia, are small (~ 0.5 μ m in diameter) cylindrical adhesion structures consisting of focal adhesion proteins such as vinculin (red) surrounded by an actin core (green). (a) Podosomes (arrow) at the periphery of an osteoclast. (b) Invadopodia (arrow) are prominently found in the vicinity of the cell nucleus in invasive or metastatic cancer cells. Images in (a) are courtesy of S. Batsir; images in (b) courtesy of O. Revach, Weizmann Institute of Science.

adhesions either constitutively or transiently, are known, collectively, as the ‘integrin adhesome’ (Zaidel-Bar and Geiger, 2010; Zaidel-Bar *et al.*, 2007). An updated list of components, available in (Winograd-Katz *et al.*, 2014) can be divided into two major domains: a ‘scaffolding domain,’ consisting of the adhesion receptors (mostly different integrins), a multitude of adapter proteins that connect the receptors to the cytoskeleton, and various actin-binding and regulatory proteins. In addition, a ‘regulatory domain’ consisting of a multitude of signaling proteins (e.g., kinases, phosphatases, G-proteins and their activators or inhibitors), generates and mediates adhesion-dependent signaling cascades, thereby affecting cell behavior, as well as the development and modulation of the adhesions themselves (Figure 5).

The transmembrane components of focal adhesions contain different pairs of α - and β integrin subunits, which bind to specific matrix components via their extracellular domains (Giancotti and Ruoslahti, 1999). The matrix specificity of cells is primarily attributed to the type of integrin they possess, and to their state of activation. The most common integrins found

in focal- and related ECM adhesions are the fibronectin receptor, $\alpha_5\beta_1$, and the vitronectin receptor, $\alpha_v\beta_3$. Additional membrane receptors, though less characterized at the structural and functional levels, include other integrins, as well as potentially adhesive molecules such as syndecan-4 and the hyaluronan-binding protein layilin (Lanza *et al.*, 2013).

Associated with the cytoplasmic aspect of the plasma membrane is the focal adhesion plaque, consisting of multi-protein complexes that interact with the cytoplasmic tails of the integrin chains and, eventually, link them to the actin cytoskeleton. Some of these plaque components, including α -actinin, talin, filamin, and tensin, contain binding sites for both integrins and actin, while other integrin-associated molecules do not directly bind to actin, but rather to different actin-associated proteins or additional ‘adapter’ proteins. Notably, these adapters, which include signaling proteins such as focal adhesion kinase (FAK), the LIM-domain protein DRAL, or adhesion mechano-regulators such as vinculin, can play regulatory roles. In fact, vinculin plays a pivotal role as a central linker interacting with many plaque proteins, including

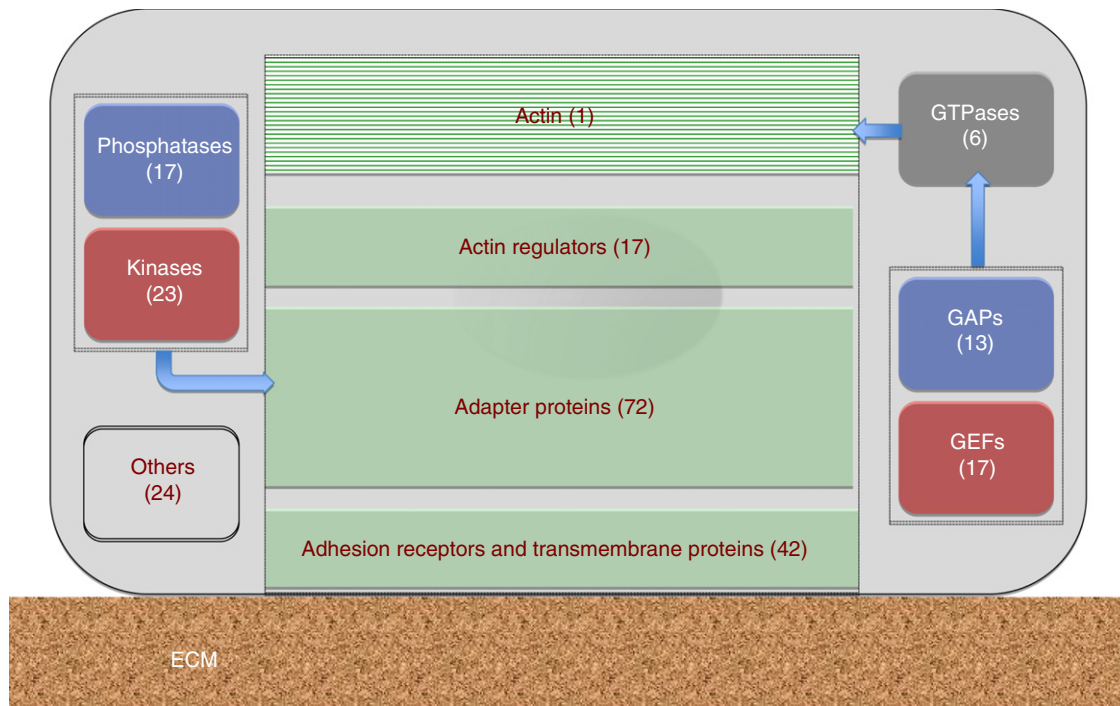


Figure 5 Diagram depicting the different functional and chemical groups comprising the integrin adhesome. Some 232 distinct proteins and other molecules are known to be contained in, or directly associated with, focal adhesions and other, similar adhesion sites. The adapter proteins, together with the actin regulators, form the scaffolding domain of the adhesion site, providing it with its unique mechanical properties. The stability and dynamics of the adhesion are controlled by its regulatory components, which include diverse elements of signaling cascades, mostly kinases and their corresponding phosphatases, as well as small Rho GTPases and their GEF and GAP regulators.

talin, α -actinin, VASP/Ena, ponsin and vinexin, as well as with acidic phospholipids and actin. Taken together, the adapter components and actin regulators of focal adhesions form the 'scaffolding domain' of the adhesion site, and generate the unique mechanical properties of these structures. Furthermore, many signaling molecules were shown to interact with these adhesions, and regulate their stability and dynamics. These 'regulatory components' include diverse components of signaling cascades, mostly kinases (tyrosine- and serine/threonine-specific) and the corresponding phosphatases, as well as small Rho GTPases and their GEF and GAP regulators. It is likely that dynamic, continuous cross-talk between the scaffolding and signaling domains is responsible for both the global 'adhesion-mediated signaling' believed to be generated in focal adhesions, as well as for the local dynamic reorganization of the adhesion sites in which they reside. Further details about the interplay between these two functional domains of adhesion sites are described in (Zaidel-Bar and Geiger, 2010).

Geometric and Mechanical Sensory (Mechanosensing) Functions of ECM Adhesions

The ECM presents cells with highly complex and diverse physical environments. Different tissue types may vary dramatically in rigidity (bone vs. cartilage), activity (beating heart vs. static brain), and geometry (2D basement membrane vs. 3D stroma), all of which influence cell organization and

function. Over the past two decades, a number of physical features of the cellular microenvironment have been observed, primarily *ex vivo*, to play key biological roles:

- **Spatial distribution of ECM ligand molecules:** The spacing of ECM ligands in physiological conditions may vary greatly. To mimic such conditions, cultured cells were plated on 2D substrates with controlled, nanoscale-positioning of the ECM ligands (Arnold *et al.*, 2008). Cells spread and migrated on the nanopatterned surfaces when ligand spacing was below a critical value (~ 70 nm), while at higher spacings, cell viability was severely hampered, eventually leading to apoptosis. Moreover, cells could detect and respond to minute gradients in ligand spacing (comparable to ligand positioning inhomogeneities), indicating that the ECM cues were averaged over a number of adhesion sites, to improve cellular sensitivity. Notably, an intrinsic focal adhesion length scale was observed in the same range as that of 'optimal' ligand spacing (Patla *et al.*, 2010). Three-dimensional structural reconstruction of focal adhesions using cryo-electron tomography revealed particles located at the cell membrane, and attached to actin fibers with a mean interspacing of approximately 45 nm.
- **ECM geometry:** The ECM provides cells with a 3D microenvironment. However, in some cases, the effective cell-matrix contacts are geometrically confined. For example, epithelial and endothelial cells adhere to the basal lamina, a practically 2D substrate (Stamenovic *et al.*, 2009). In addition, cells adhering to single ECM filaments effectively sense a 1D geometry (Sharma *et al.*, 2012). *In vitro* studies

- mimicking these different geometries have demonstrated particularly diverse cell shapes, polarities and migration speeds (e.g., Doyle *et al.*, 2009). In addition, the porosity level of the 3D ECM scaffold was shown to influence cell morphology and migration dynamics (e.g., ameboid vs. mesenchymal motion, and individual vs. collective migration) (Friedl and Gilmour, 2009).
- ECM topography: While the vast majority of cell culture studies are conducted, for imaging purposes, on flat substrates, cells *in vivo* adhere to rough surfaces as well. In such cases, the cell membrane can bend and deform to remain in contact with the ECM. In turn, the resulting curvature of the cell membrane and adhesion sites may affect cell behavior, by exposing cryptic protein binding sites or activating mechanosensitive ion channels (Hoffman *et al.*, 2011). One example of the correlation between surface roughness and cell response was observed in bone-resorbing osteoclasts, in which the resorption dynamics are highly dependent on ECM topography (Geblinger *et al.*, 2010).
 - ECM rigidity: *In vivo*, ECM rigidities span over 7 orders of magnitude, with different cell types typically associated with characteristic stiffness regimes (Discher *et al.*, 2005; Swift *et al.*, 2013). Loss of the cell's ability to sense changes in ECM rigidity are associated with cell dysfunction and disease (Janmey and Miller, 2011). For example, tumor stroma is typically stiffer than normal stroma, and its progression in 3D models correlates with ECM rigidity. In 2D cultured systems, substrate rigidity has been associated with cell morphology (Geiger *et al.*, 2009; Vogel and Sheetz, 2006), migration (Lo *et al.*, 2000; Saez *et al.*, 2007), proliferation (Janmey and Miller, 2011; Orr *et al.*, 2006), gene expression and differentiation (Discher *et al.*, 2005; Guilak *et al.*, 2009). The mechanism through which cells

probe the rigidity of their microenvironment involves application of actomyosin contractile forces via the cell-ECM adhesion sites, notably focal adhesions.

- External forces: The ECM mechanically couples cells anchored to it. In this manner, cells are exposed to external forces, which may be critical for tissue development, function, remodeling and healing. Remarkably, both constant (e.g., blood flow experienced by endothelial cells) and rhythmic force perturbations (e.g., originating in the respiratory or cardio-vascular systems) have been associated with cell realignment at well-defined and uniform angles (Wang and Grood, 2000).

The Cross-Talk between the Cell and the ECM

The ECM is a highly complex network of adhesive filaments with specialized chemical, mechanical and geometric properties (Geiger and Yamada, 2011). It serves as an effective tissue scaffold, providing cells with the precise environmental conditions required for proper development, survival and proliferation. Similar to the cytoplasmic aspects of focal adhesions described above, the ECM is characterized by both scaffolding and signaling properties. The scaffolding elements attached to the extracellular faces of focal adhesions consist of tightly packed arrays of ECM filaments, which are nearly colinear with the cytoplasmic actin filaments that attach to the focal adhesion at the cell interior (Singer, 1979). In addition, there is ample evidence that the ECM serves as a depot for growth factors that can stimulate the growth, differentiation and dynamics of the attached cells (Brizzi *et al.*, 2012; Gospodarowicz *et al.*, 1980; Hynes, 2009). This combination of external stimulation of adherent cells by chemical and physical

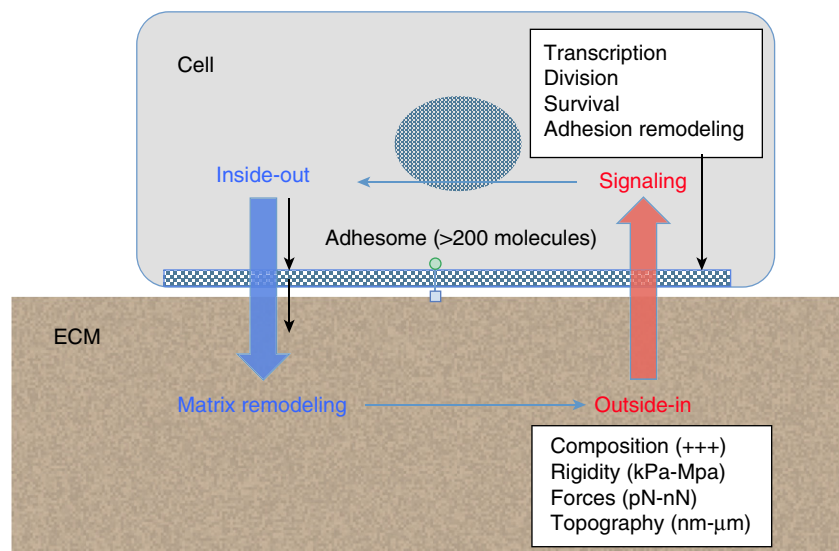


Figure 6 Diagram depicting the cross-talk between cells and the ECM. The properties of the ECM scaffold, such as composition, rigidity, and topography, and associated external forces, induce the formation of cellular adhesion structures (the 'outside-in' effect). The resulting signaling cascades that follow, through the integrin adhesome, may in turn drive an 'inside-out' process of ECM remodeling by the cell. Thus, bi-directional cross-talk between the cell, through its adhesion machinery, and the surrounding microenvironment is perpetually maintained, remodeling the ECM and, in turn, affecting cell behavior, structure, and fate.

ECM-mediated cues is, most likely, underlying the ‘anchorage dependence’ properties of normal cells, which are often lost upon malignant transformation (Stoker *et al.*, 1968).

ECM adhesion sites also serve as major environmental sensors, which are induced to assemble by means of chemical and physical cues generated at their points of contact with the ECM and activating, across the membrane, various signaling cascades that affect cell behavior and fate (Figure 6). Typical signaling events triggered by the adhesion include the activation of tyrosine phosphorylation caused by the accretion and activation of FAK and Src, the activation of Rho and Ras small GTPases, and the subsequent activation of the extracellular signal-regulated kinase (ERK) pathway (Geiger *et al.*, 2009). It is proposed that the tethering of FAK to clustered integrins leads to its autophosphorylation, followed by the recruitment of Src, which binds to phospho-FAK via its phosphotyrosine-binding domain (SH2), and followed by further phosphorylation of a series of downstream targets (Mitra and Schlaepfer, 2006). It is further believed that the local accumulation of a rich variety of kinases, phosphatases and their substrates in the sub-membrane focal adhesion plaque plays a key role in the adhesion-mediated signaling process (Geiger *et al.*, 2009). That said, definitive information concerning these signaling processes and their cellular effects remains scarce. It is interesting to note that, apart from the long-term and long-range adhesion signals that affect overall cell fate, there are also local signals, which act at, and affect focal adhesion structure and stability.

Clearly, however, the ECM scaffold and the sensory-signaling machinery of the cell conduct a complex molecular dialog that has far-reaching consequences for the cells. The focal adhesion-based sensory system is capable of sensing multiple molecular and physical cues such as ligand chemistry and spacing, surface nano-topography, geometry, rigidity and external forces. This rich information is integrated by focal adhesion-associated networks and triggers a cytoplasmic response; a process known as an ‘outside-in response.’ In turn, the same signaling cascades, triggered by environmental cues, can induce an ‘inside-out response,’ modifying the neighboring ECM (e.g., through dissolution with proteases, mechanical perturbation, or additional fiber deposition) which, in turn, can affect the original environmental signals. It is proposed that this bi-directional cross-talk between the cell, through its adhesion machinery, and the microenvironment is perpetually maintained, remodeling the ECM and, in turn, regulating cell behavior, structure and fate.

Acknowledgments

Research conducted in our laboratory on the molecular diversity of integrin adhesions was supported by the Israel Science Foundation and by the Henry Kreuter Foundation. Benjamin Geiger is the incumbent of the Erwin Neter Professorial Chair in Cell and Tumor Biology.

See also: Cytoskeleton and Motors: Complex Cytoskeletal Structures: The Extracellular Matrix

References

- Alberts, B., Johnson, A., Lewis, J., *et al.*, 2007. *Molecular Biology of the Cell*, fifth ed. New York: Garland Science.
- Arnold, M., Hirschfeld-Warneken, V.C., Lohmuller, T., *et al.*, 2008. Induction of cell polarization and migration by a gradient of nanoscale variations in adhesive ligand spacing. *Nano Letters* 8, 2063–2069.
- Baharloo, B., Textor, M., Brunette, D.M., 2005. Substratum roughness alters the growth, area, and focal adhesions of epithelial cells, and their proximity to titanium surfaces. *Journal of Biomedical Materials Research Part A* 74A, 12–22.
- Bershadsky, A., Chausovsky, A., Becker, E., Lyubimova, A., Geiger, B., 1996. Involvement of microtubules in the control of adhesion-dependent signal transduction. *Current Biology* 6, 1279–1289.
- Bokstad, M., Sabanay, H., Dahan, I., Geiger, B., Medalia, O., 2012. Reconstructing adhesion structures in tissues by cryo-electron tomography of vitrified frozen sections. *Journal of Structural Biology* 178, 76–83.
- Brizzi, M.F., Tarone, G., Defilippi, P., 2012. Extracellular matrix, integrins, and growth factors as tailors of the stem cell niche. *Current Opinion in Cell Biology* 24, 645–651.
- Cavalcanti-Adam, E.A., Volberg, T., Micoulet, A., *et al.*, 2007. Cell spreading and focal adhesion dynamics are regulated by spacing of integrin ligands. *Biophysical Journal* 92, 2964–2974.
- Cukierman, E., Pankov, R., Stevens, D.R., Yamada, K.M., 2001. Taking cell-matrix adhesions to the third dimension. *Science* 294, 1708–1712.
- Curtis, A., Wilkinson, C., 1997. Topographical control of cells. *Biomaterials* 18, 1573–1583.
- Davis, G.E., Camarillo, C.W., 1995. Regulation of endothelial-cell morphogenesis by integrins, mechanical forces, and matrix guidance pathways. *Experimental Cell Research* 216, 113–123.
- Destaing, O., Saltel, F., Geminard, J.C., Jurdic, P., Bard, F., 2003. Podosomes display actin turnover and dynamic self-organization in osteoclasts expressing actin-green fluorescent protein. *Molecular Biology of the Cell* 14, 407–416.
- Destaing, O., Saltel, F., Gilquin, B., *et al.*, 2005. A novel Rho-Mdia2-Hdac6 pathway controls podosome patterning through microtubule acetylation in osteoclasts. *Journal of Cell Science* 118, 2901–2911.
- Discher, D.E., Janmey, P., Wang, Y.L., 2005. Tissue cells feel and respond to the stiffness of their substrate. *Science* 310, 1139–1143.
- Doyle, A.D., Wang, F.W., Matsumoto, K., Yamada, K.M., 2009. One-dimensional topography underlies three-dimensional fibrillar cell migration. *Journal of Cell Biology* 184, 481–490.
- Egeblad, M., Nakasone, E.S., Werb, Z., 2010a. Tumors as organs: Complex tissues that interface with the entire organism. *Developmental Cell* 18, 884–901.
- Egeblad, M., Rasch, M.G., Weaver, V.M., 2010b. Dynamic interplay between the collagen scaffold and tumor evolution. *Current Opinion in Cell Biology* 22, 697–706.
- Elsdale, T., Bard, J., 1972. Collagen substrata for studies on cell behavior. *Journal of Cell Biology* 54, 626–637.
- Friedl, P., Gilmour, D., 2009. Collective cell migration in morphogenesis, regeneration and cancer. *Nature Reviews Molecular Cell Biology* 10, 445–457.
- Geblinger, D., Addadi, L., Geiger, B., 2010. Nano-topography sensing by osteoclasts. *Journal of Cell Science* 123, 1503–1510.
- Geiger, B., Bershadsky, A., Pankov, R., Yamada, K.M., 2001. Transmembrane extracellular matrix-cytoskeleton crosstalk. *Nature Reviews Molecular Cell Biology* 2, 793–805.
- Geiger, B., Spatz, J.P., Bershadsky, A.D., 2009. Environmental sensing through focal adhesions. *Nature Reviews Molecular Cell Biology* 10, 21–33.
- Geiger, B., Yamada, K.M., 2011. Molecular architecture and function of matrix adhesions. *Cold Spring Harbor Perspectives in Biology* 3, 21.
- Giancotti, F.G., Ruoslahti, E., 1999. Transduction – Integrin signaling. *Science* 285, 1028–1032.
- Gospodarowicz, D., Delgado, D., Vlodavsky, I., 1980. Permissive effect of the extracellular-matrix on cell-proliferation in vitro. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* 77, 4094–4098.
- Grinnell, F., Petroll, W.M., 2010. Cell motility and mechanics in three-dimensional collagen matrices. In: Schekman, R., Goldstein, L., Lehmann, R. (Eds.), *Annual Review of Cell and Developmental Biology*, vol. 26. Palo Alto, CA: Annual Reviews.
- Grossner-Schreiber, B., Herzog, M., Hedderich, J., *et al.*, 2006. Focal adhesion contact formation by fibroblasts cultured on surface-modified dental implants: An *in vitro* study. *Clinical Oral Implants Research* 17, 736–745.

- Guilak, F., Cohen, D.M., Estes, B.T., *et al.*, 2009. Control of stem cell fate by physical interactions with the extracellular matrix. *Cell Stem Cell* 5, 17–26.
- Hay, E.D., 1991. *Cell Biology of Extracellular Matrix*. New York, NY: Plenum Press.
- Hoffman, B.D., Grashoff, C., Schwartz, M.A., 2011. Dynamic molecular processes mediate cellular mechanotransduction. *Nature* 475, 316–323.
- Hynes, R.O., 2009. The extracellular matrix: Not just pretty fibrils. *Science* 326, 1233–1239.
- Hynes, R.O., Naba, A., 2012. Overview of the matrisome — An inventory of extracellular matrix constituents and functions. *Cold Spring Harbor Perspectives in Biology* 4, 16.
- Janmey, P.A., Miller, R.T., 2011. Mechanisms of mechanical signaling in development and disease. *Journal of Cell Science* 124, 9–18.
- Kaverina, I., Krylyshkina, O., Small, J.V., 1999. Microtubule targeting of substrate contacts promotes their relaxation and dissociation. *Journal of Cell Biology* 146, 1033–1043.
- Lanza, R., Langer, R., Vacanti, J.P., 2013. *Principles of Tissue Engineering*. New York, NY: Elsevier Science.
- Linder, S., 2009. Invadosomes at a glance. *Journal of Cell Science* 122, 3009–3013.
- Litjens, S.H.M., De Pereda, J.M., Sonnenberg, A., 2006. Current insights into the formation and breakdown of hemidesmosomes. *Trends in Cell Biology* 16, 376–383.
- Lo, C.M., Wang, H.B., Dembo, M., Wang, Y.L., 2000. Cell movement is guided by the rigidity of the substrate. *Biophysical Journal* 79, 144–152.
- Massia, S.P., Hubbell, J.A., 1991. An RGD spacing of 440 nm is sufficient for integrin alpha-V-beta-3-mediated fibroblast spreading and 140 nm for focal contact and stress fiber formation. *Journal of Cell Biology* 114, 1089–1100.
- Menko, A.S., Boettiger, D., 1987. Occupation of the extracellular-matrix receptor, integrin, is a control point for myogenic differentiation. *Cell* 51, 51–57.
- Mitra, S.K., Schlaepfer, D.D., 2006. Integrin-regulated FAK-Src signaling in normal and cancer cells. *Current Opinion in Cell Biology* 18, 516–523.
- Munger, J.S., Sheppard, D., 2011. Cross talk among TGF-beta signaling pathways, integrins, and the extracellular matrix. *Cold Spring Harbor Perspectives in Biology* 3, 17.
- Nelson, C.M., Bissell, M.J., 2006. Of extracellular matrix, scaffolds, and signaling: Tissue architecture regulates development, homeostasis, and cancer. *Annual Review of Cell and Developmental Biology*. Palo Alto, CA: Annual Reviews.
- North, A.J., Galazkiewicz, B., Byers, T.J., Glenney, J.R., Small, J.V., 1993. Complementary distributions of vinculin and dystrophin define 2 distinct sarcolemma domains in smooth-muscle. *Journal of Cell Biology* 120, 1159–1167.
- Orr, A.W., Helmke, B.P., Blackman, B.R., Schwartz, M.A., 2006. Mechanisms of mechanotransduction. *Developmental Cell* 10, 11–20.
- Patla, I., Volberg, T., Elad, N., *et al.*, 2010. Dissecting the molecular architecture of integrin adhesion sites by cryo-electron tomography. *Nature Cell Biology* 12, 909–915.
- Revach, O., Geiger, B., 2014. The interplay between the proteolytic, invasive, and adhesive domains of invadopodia and their roles in cancer invasion. *Cell Adhesion & Migration* 8, 215–225.
- Ricard-Blum, S., 2011. The collagen family. *Cold Spring Harbor Perspectives in Biology* 3, 19.
- Saez, A., Ghibaudo, M., Buguin, A., Silberzan, P., Ladoux, B., 2007. Rigidity-driven growth and migration of epithelial cells on microstructured anisotropic substrates. *Proceedings of the National Academy of Sciences of the United States of America* 104, 8281–8286.
- Sarrazin, S., Lamanna, W.C., Esko, J.D., 2011. Heparan sulfate proteoglycans. *Cold Spring Harbor Perspectives in Biology* 3, 33.
- Sharma, V.P., Beaty, B.T., Patsialou, A., *et al.*, 2012. Reconstitution of *in vivo* macrophage-tumor cell pairing and streaming motility on one-dimensional micro-patterned substrates. *IntraVital* 1, 77–85.
- Singer, I.I., 1979. Fibronexus — Transmembrane association of fibronectin-containing fibers and bundles of 5 nm microfilaments in hamster and human-fibroblasts. *Cell* 16, 675–685.
- Stamenovic, D., Lazopoulos, K.A., Pirentis, A., Suki, B., 2009. Mechanical stability determines stress fiber and focal adhesion orientation. *Cellular and Molecular Bioengineering* 2, 475–485.
- Stoker, M., Oneill, C., Berryman, S., Waxman, V., 1968. Anchorage and growth regulation in normal and virus-transformed cells. *International Journal of Cancer* 3, 683–693.
- Swift, J., Ivanovska, I.L., Buxboim, A., *et al.*, 2013. Nuclear lamin-a scales with tissue stiffness and enhances matrix-directed differentiation. *Science* 341, 975.
- Vogel, V., Sheetz, M., 2006. Local force and geometry sensing regulate cell functions. *Nature Reviews Molecular Cell Biology* 7, 265–275.
- Vogel, W.F., Abdulhussein, R., Ford, C.E., 2006. Sensing extracellular matrix: An update on discoidin domain receptor function. *Cellular Signalling* 18, 1108–1116.
- Wang, J.H.C., Grood, E.S., 2000. The strain magnitude and contact guidance determine orientation response of fibroblasts to cyclic substrate strains. *Connective Tissue Research* 41, 29–36.
- Watt, F.M., Driskell, R.R., 2010. The therapeutic potential of stem cells. *Philosophical Transactions of the Royal Society B-Biological Sciences* 365, 155–163.
- Winograd-Katz, S., Fassler, R., Geiger, B., Legate, K.R., 2014. The integrin adhesome: From genes and proteins to human disease. *Nature Reviews Molecular Cell Biology* 15, 273–288.
- Yamada, K.M., Cukierman, E., 2007. Modeling tissue morphogenesis and cancer in 3D. *Cell* 130, 601–610.
- Yurchenco, P.D., 2011. Basement membranes: Cell scaffoldings and signaling platforms. *Cold Spring Harbor Perspectives in Biology* 3, 27.
- Zaidel-Bar, R., Geiger, B., 2010. The switchable integrin adhesome. *Journal of Cell Science* 123, 1385–1388.
- Zaidel-Bar, R., Itzkovitz, S., Ma'ayan, A., Iyengar, R., Geiger, B., 2007. Functional atlas of the integrin adhesome. *Nature Cell Biology* 9, 858–868.
- Zamir, E., Geiger, B., 2001. Molecular complexity and dynamics of cell-matrix adhesions. *Journal of Cell Science* 114, 3583–3590.
- Zamir, E., Katz, M., Posen, Y., *et al.*, 2000. Dynamics and segregation of cell-matrix adhesions in cultured fibroblasts. *Nature Cell Biology* 2, 191–196.

Relevant Websites

<http://www.adhesome.org/>
Adhesome Network.

<http://www.weizmann.ac.il/>
Benny Geiger's Homepage.

<https://www.cellmigration.org/>
Cell Migration Gateway.

<http://www.mechanobio.info/>
Mechanics of Cellular and Molecular Function Wiki.

Cell Migration

A Ortega-Carrion, I Feo-Lucas, and M Vicente-Manzanares, Universidad Autónoma de Madrid School of Medicine, Madrid, Spain

© 2016 Elsevier Inc. All rights reserved.

Introduction

Cell migration is a fundamental process throughout evolution. Some prokaryotic cells are able to swim in fluid media using specialized structures, for example, flagellae, to propel the cell body directionally, or move using different strategies, for example, sliding, gliding, etc. These movements aim to establish new colonies away from the original or search for nutrients.

Unicellular eukaryotic cells, for example, amoebae, undergo chemotaxis while searching for food. In higher organisms, for example, mammals, different cell types display various modes of migration while responding to varied extracellular stimuli. Migration is crucial in tissue morphogenesis, regeneration, or inflammation. Alterations of the migratory patterns of specific cell types, for example, leukocytes, underlie disease, for example, autoimmune disease. Migration is also an important component of the pathogenesis of cancer during invasion and metastasis. Conversely, lack of migration is a major cause of failure of potential regenerative therapies, for example, using stem cells.

Cell migration requires the integration of complex cellular responses responding to multiple signals. It is now a fully formed field that encompasses almost 40 years of research. In this article, we provide an outline of eukaryotic cell migration; its fundamental modes and underlying subprocesses; molecular effectors and regulators; and roles in homeostasis and disease.

Alone or in the Company of Others: Modes of Eukaryotic Cell Migration

Eukaryotic cells display different migratory modes depending on their intrinsic nature or microenvironment. We classify migration into two major categories, single-cell and multicellular (Friedl and Wolf, 2010). Single-cell migration is typical of highly differentiated cells with specific functions, for example, leukocytes or neuronal progenitors; collective migration refers to the simultaneous and coordinated migration of multiple cells while in contact with others. This is typical of monolayers with structural/support roles, for example, fibroblasts or epithelial cells.

Single-cell migration is further subdivided according to the morphology of the cell: most constitutively adherent cells, for example, fibroblasts, neuroblasts, etc., display mesenchymal cell migration, characterized by low speed, high adherence, extension of large protrusions and reorganization, and/or degradation of the extracellular matrix (ECM) (Ridley *et al.*, 2003). Leukocytes display amoeboid migration, in which cells are faster, less adherent, display smaller protrusions, and are much less contractile (Lammermann and Sixt, 2009). Some cells display bleb-based motion, in which cells move slowly by extending membrane blebs and do not require receptor-mediated adhesion (Charras *et al.*, 2008; Figure 1(a)).

Multicellular migration comprises chain migration, which displays leader cells at the forward edge and trailing cells behind those. This is typical of neural crest cells (Theveneau and Mayor, 2012). Collective migration occurs when multiple cells form a coordinated front that moves in unison, with layers of trailing cells following them: this occurs during early embryogenesis and wound healing (Shaw and Martin, 2009; Figure 1(b)).

Most cells exhibit a degree of 'migratory plasticity' that enables them to switch back and forth between migratory modes. For example, mesenchymal cells may undergo amoeboid or bleb-based migration under specific adhesive conditions (Chen *et al.*, 2012; Bergert *et al.*, 2012), particularly during malignant transformation (Sanz-Moreno *et al.*, 2008). These changes reflect the adaptability to the microenvironment and dynamic nature of the migratory process.

Individual Cell Migration

Mesenchymal migration

This type of cellular movement is characterized by strong adhesion and front-back cell polarization driven by actin (Figure 2(a)). Mesenchymal cells, for example, fibroblasts, develop a leading edge in the direction of migration, which appears as a forward deformation of the plasma membrane, or protrusion. Protrusion is driven by actin polymerization (Pollard and Borisy, 2003) and is enabled by adhesion to the substrate (Choi *et al.*, 2008). This is a process that is dependent on integrin receptors (Hynes, 2002), which form extended interactive signaling networks that coordinately control actin polymerization and adhesion (Aguilar-Cuenca *et al.*, 2014). In most mesenchymal cells, protrusions contain two major regions based on the conformation of the actin cytoskeleton as well as the shape and dynamics of the adhesive structures that they contain. The lamellipodium is closer to the leading edge and is formed by newly polymerized, branched actin (Svitkina and Borisy, 1999). In most cells, it is $\approx 1 \mu\text{m}$ wide. Width is relatively stable due to actin treadmilling, in which the lamellipodium grows at the outer edge due to actin monomer incorporation to the actin barbed ends closer to the edge and disassembles at its rear due to actin depolymerization. Adhesions in the lamellipodium are short lived and small (Alexandrova *et al.*, 2008; Choi *et al.*, 2008). The $\approx 5\text{--}10 \mu\text{m}$ behind the lamellipodium are commonly referred to as the lamellum. Actin reorganizes into larger bundles of filaments due to the action of actin cross-linking proteins, for example, non-muscle myosin II (NMII). Adhesions become larger and more stable. The function of the lamellum is likely to isolate cellular structures from the rapid retrograde flow of the actin at the lamellipodium while connecting this subdomain to the rest of the cell body (Verkhovsky *et al.*, 1999).

Increased adhesion size in the lamellum results from the process of adhesion maturation, which occurs along actin

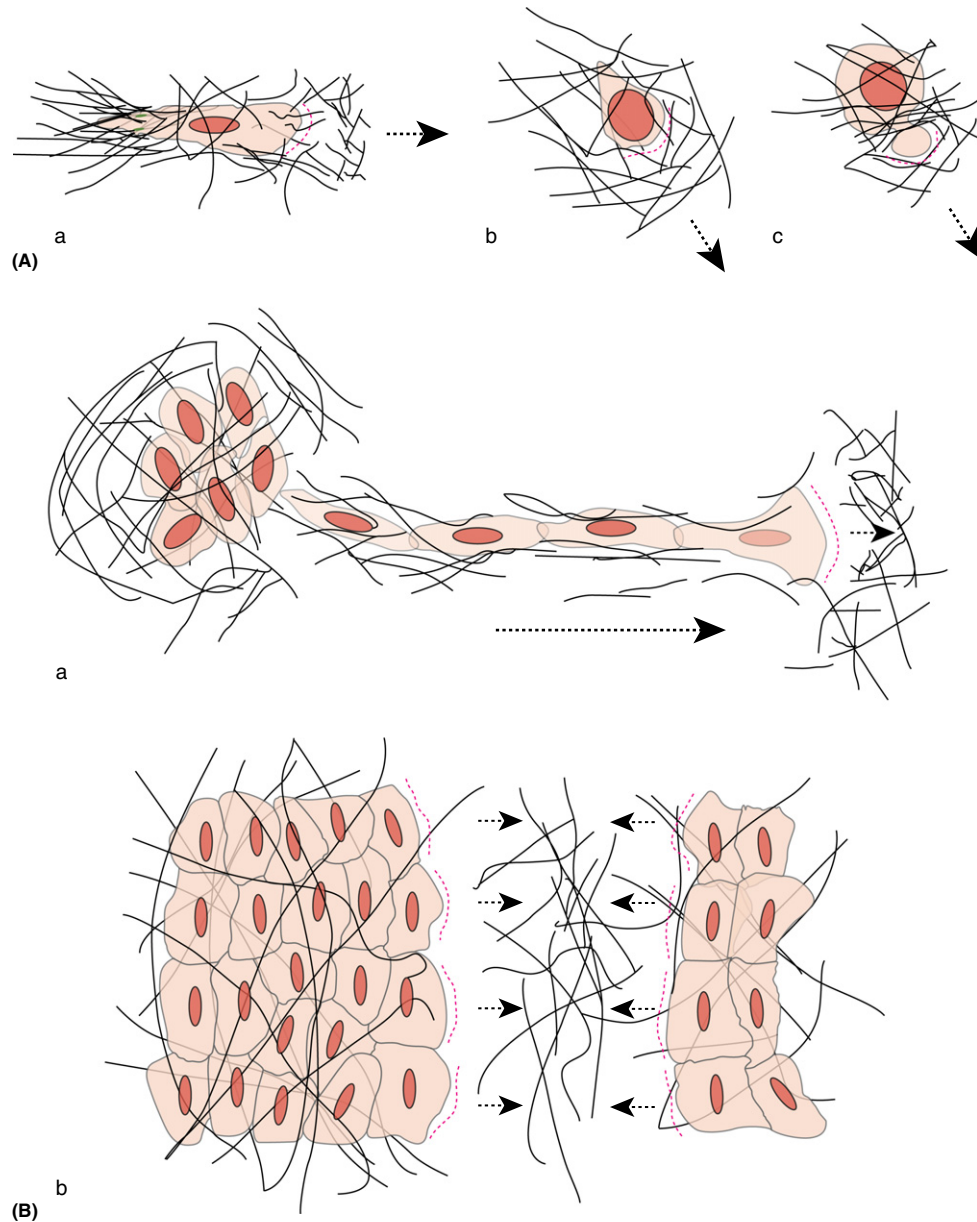


Figure 1 Modes of eukaryotic cell migration. (A) Types of individual cell migration. (a) Mesenchymal migrating cell, the cell is polarized and it rearranges the ECM (shown as densely packed black strands at one pole of the cell) as it migrates. (b) Amoeboid, the cell is also polarized but it is less contractile, reorganizes the ECM (depicted as loosely organized strands) to a much lower extent and shows less interaction with the ECM. The cell also is polarized, however are much less contractile. (c) Blebbing motility. Hydrostatic pressure generates a bleb on one side of the cell, but the overall morphology is not polarized. (B) Types of multicellular migration (a) Chain migration, a leader cell drives migration of a chain of follower cells bound to each other head to tail. The leading cell is polarized (b) Collective migration. A layer of cells invades a tissue gap led by a row of protruding cells (protrusion is represented as a discontinuous line along the leading edge of the first row of cells). Movement may occur on one side only or on both sides (as depicted). In all images, arrows point to the direction of migration.

templates (bundles) that form due to actin cross-linking (Choi *et al.*, 2008). The precursors of the lamellar adhesions are dynamic nascent adhesions generated in the lamellipodium that remain behind but do not disassemble as the lamellipodium moves forward. The frequency of adhesion maturation depends on the cell type and its microenvironment, and is controlled by the stability and length of the actin templates that guide their elongation (Choi *et al.*, 2008). These adhesions likely generate tension to drive forward the cell body.

The trailing edge of these cells undergoes constant adhesion disassembly driven by tension as well as proteolysis of adhesion proteins within focal adhesions.

Different lines of evidence indicate that single mesenchymal cells form the rear first, i.e., protrude away from a non-dynamic area that becomes the rear as the cell polarize. This region is dependent on early assembly of NMII, which acts as an anti-protrusive factor that shuts down the prospective trailing edge of the cell as it spreads (Vicente-Manzanares *et al.*, 2011).

A likely similar mechanism takes place in collective chain migration, in which the tail of the leader cell shuts down the leading edge of the followers, promoting cadherin-based cell–cell adhesion instead.

Amoeboid migration

Amoeboid motion is typical of some protozoa and leukocytes. **Figure 2(b)** shows that, in 2-D, leukocytes require integrins for migration, but adhesion is not as strong as in mesenchymal cells. Several authors have indicated that in 3-D, integrin-mediated adhesion may be dispensable (Friedl and Weigelin, 2008; Lammermann *et al.*, 2008). Weaker adhesion is on display in the nature of the adhesive structures, which very seldom reach the size and stability of those observed in fibroblasts (Chen *et al.*, 2012). Amoeboid cells sustain high migratory speeds, which can be 10-fold higher than that of mesenchymal cells ($> 10 \mu\text{m min}^{-1}$ compared to $0.2\text{--}2 \mu\text{m min}^{-1}$ in fibroblasts). Some principles of mesenchymal migration apply, i.e., formation of a protrusive leading edge that contains small, dynamic adhesions. The lamellipodium is similar to that observed in mesenchymal cells, but the lamellum is poorly defined and contains no large adhesions. The trailing edge of these cells is very contractile and pushes the nucleus forward, in some cases to squeeze it through confined spaces, for example, during leukocyte extravasation (Jacobelli *et al.*, 2013). Amoeboid cells are more agile than mesenchymal cells to migrate in response to microenvironment changes. An example is the ability of cells to turn in response to an orientation change of an extracellular chemoattractant. Mesenchymal cells make a new leading edge in the new direction of the gradient, retract the old leading edge, and repolarize. Conversely, amoeboid cells turn, repositioning their existing leading edge in the new direction (Kwon *et al.*, 2013). This may be important for the ability of leukocytes to chase pathogens or migrate rapidly in response to sudden microenvironment changes, for example, during inflammation.

Blebbing motion

Blebbing motility does not require actin polymerization for membrane extension. This is due to variations in hydrostatic pressure within the cell. Blebs are membrane ‘bubbles’ that emerge in the cellular periphery. Blebs form when the actomyosin cortex that underlies the plasma membrane separates from the phospholipid bilayer by contraction or disassembly. The newly formed space fills rapidly with cytoplasm, which increases the volume locally and expands the plasma membrane. In most cases, the underlying actin cortex is reformed by mechanisms unknown (the most common actin nucleators, i.e., actin-related protein 2/3 (Arp2/3) or formins, have not

been found in these regions) (Charras *et al.*, 2008). Finally, bleb retraction requires recruitment of NMII to the bleb to support inward contraction.

Membrane blebbing is mostly used for migration in 3-D environment. In this context, cells squeeze the blebs into narrow gaps and undergoes forward motion by using the bleb to cling to the sides of the gap, using them as footholds (Renkawitz and Sixt, 2010; **Figure 2(c)**).

Multicellular Migration

Chain migration

In this migratory mode, a leader cell drives forward migration in a similar manner to that described for mesenchymal cells. The trailing edge of this cell remains adhered to a follower cell using cadherin receptors; this cell may drag another cell and so forth, forming ‘Mexican trains.’ This is characteristic of neuroblast migration, which occurs along glial patterns (Solecki *et al.*, 2009). Sahai and coworkers also described how skin fibroblasts lead squamous cell carcinoma cells (SCC) as they invade the surrounding tissue in a radial manner (Gaggioli *et al.*, 2007).

Collective migration

This is typical of regenerative processes, for example, wound healing, and in development. A row of leader cells (which contact each other side to side) moves coordinately forward, followed by a cellular monolayer. Only cells from the leading row display protrusive leading edges. Follower cells have no leading edges; instead, they remain attached to the leading row and among themselves via cell–cell adhesion receptors. Posterior cells move using actomyosin contractility and are pushed forward by proliferation.

Building a Migrating Cell

A multitude of extracellular and intracellular molecules control the dominant migratory mode of a cell, or even whether it migrates or not. In the following sections, we outline most of these factors, from the outside to the inside of the cell, and their major roles in cell migration.

Extracellular Factors

ECM

ECM provides most of the adhesive backbone for cells to migrate. ECM proteins include collagens, fibronectin, laminin, vitronectin, and many others. Most of them support adhesion

Figure 2 Features of individually migrating cells. (a) Mesenchymal migration. Figure represents a top and side view of a migrating mesenchymal cell. Close to the leading edge, the lamellipodium displays branched actin (right insert) and small, nascent adhesions (in yellow). The lamellum connects the lamellipodium to the cell body. Actin switches from branched to bundled and adhesions become elongated along actin bundles (yellow ellipses on thick red lines). In the cell body and trailing edge, adhesions are big and actin bundles are thick and long (left insert). Adhesions act as focal points for ECM reorganization (black strands). The MTOC (green) localizes in front of the nucleus (blue). (b) Amoeboid migration. The cell depicted is polarized and displays a thick lamellipodium (red band on right) decorated with small adhesions (yellow dots). The lamellum is not apparent and the nucleus appears forward, leaving the MTOC behind (green). The trailing edge is very contractile and pushes the nucleus to leading edge. (c) Migration by blebs, an initial disassembly of the actin cortex (1) induces a local increase in hydrostatic pressure (represented as a color gradient toward the right), promoting membrane deformation and sprouting of a membrane bleb (2–3). Actin reforms into the bleb, stabilizing the structure (4) and enabling its use as an anchor point on the ECM (represented with black arrows).

via one, or several, integrin-type receptors. ECM proteins (and some cellular integrin ligands) display specific epitopes, for example, the RGD (Arg-Gly-Asp) sequence, which support integrin binding. ECM proteins are major constituents of the extracellular milieu in solid tissues and form elastic bundles (e.g., collagen) that cells use as migratory tracks. At the same time, migrating and contracting cells, particularly mesenchymal, constantly remodel the ECM fibers (e.g., during wound healing). Migrating amoeboid or bleb-based cells reorganize the ECM to a much lower extent.

Some cellular transmembrane proteins also display integrin-binding ability (they contain RGD or RGD-like motifs). Most belong to the Ig-like cell adhesion molecule family, for example, intercellular cell adhesion molecules (ICAMs), vascular cell adhesion molecule (VCAM), neural cell adhesion molecule (NCAM), etc.

Cell adhesion depends on ECM concentration following a saturable curve. Conversely, cell migration depends on the concentration of ECM in a biphasic manner. This means that, for each cell type, an optimal concentration of substrate best supports migration; under that threshold, cells do not grip sufficiently and cannot move; over the threshold, cells are too adhesive and remain attached (Polecek *et al.*, 1997).

Soluble chemoattractants

Several families of cytokines and growth factors are chemotactic, i.e., attract migrating cells. Chemotaxis triggers cellular protrusion in the direction of maximal concentration of a soluble gradient of the chemoattractant. Chemoattractants comprise several large groups. Chemokines are a large (>40) family of potent chemoattractants for leukocytes and other cell types that act through G protein-coupled receptors (GPCR). Some bacterial peptides, for example, formyl-Met-Leu-Pro (fMLP), and small subunits of complement, for example, C5a, also behave as chemoattractants through GPCR. Another family includes those GF that interact with Tyr kinase receptors, for example, insulin, vascular endothelial growth factor (VEGF), or ephrins. Most of these are chemotactic for cells that migrate in a mesenchymal manner, for example, fibroblasts, endothelial cells, or neuron progenitors.

Mechanical forces

Cells generate and respond to mechanical forces as they migrate. Through adhesions, cells generate traction forces on the ECM to promote tissue remodeling (Legant *et al.*, 2012). Likewise, forces imposed on migrating cells modulate their response. Cells not only react to vector forces and pressure (force per area), but they also modulate their migration by sensing the mechanical properties of their microenvironment. Durotaxis is the property by which cells migrate better on stiff than soft substrates and, given the choice, migrate toward stiffer substrates (Raab *et al.*, 2012).

Membrane Receptors

Integrins and focal adhesion proteins

Integrins are heterodimeric proteins that function as adhesion receptors (Hynes, 2002). To understand their role in cell migration, we will focus on two major properties: the

regulation of their activation and their ability to connect the ECM to the cellular cytoskeleton. According to their activation, there are two major classes of integrins: those that display constitutive activation, for example, fibroblast integrins; and those activated in response to stimulation, for example, those expressed in platelets and leukocytes. Their affinity increases in response to ligand exposure but also requires inside-out signals: activation of the T-cell receptor, GPCR, Tyr kinase GF receptors, etc., promotes integrin association to the actin cytoskeleton through an array of adaptors, referred to as adhesion proteins or the 'adhesome' (Zaidel-Bar *et al.*, 2007). These behave as a molecular clutch that modulates the extent of integrin activation and adhesion as well as migration (Aguilar-Cuenca *et al.*, 2014). Integrins also display regulated avidity, i.e., their ability to form plasma membrane aggregates (Carman and Springer, 2003). Adhesion proteins also regulate integrin avidity by forming interactive platforms through their association to actin.

A few notable adhesion proteins include talin, which links the cytoplasmic domain of the β chain of the integrin with actin. Talin is directly responsible for the conformational extension of the integrin in response to inside-out signaling by inducing an integrin conformational change that renders it ligand binding-competent. Vinculin is an adaptor protein that binds talin and actin and reinforces adhesion protein clustering (Bakolitsa *et al.*, 2004; Dumbauld *et al.*, 2013). Focal adhesion kinase (FAK) is a Tyr kinase that phosphorylates several substrates, most notably Src, in response to integrin activation and amplifies adhesive signals (Parsons, 2003). Paxillin is another scaffold protein that is heavily phosphorylated in response to adhesion and has multiple partners, being one of the best characterized adaptors in adhesive signaling (Deakin and Turner, 2008). A large interactive network for adhesion proteins underneath integrins has been published recently (Zaidel-Bar *et al.*, 2007; Figure 3).

Chemoattractant receptors

Most chemoattractant receptors belong to the GPCR family of seven transmembrane domains. These receptors are phosphorylated and internalized in response to ligand binding by a GRK- β -arrestin-dependent mechanism (Luttrell and Lefkowitz, 2002). G protein activation initiates a signaling cascade that includes cAMP and PIP₂ synthesis and activation of the Ras/PI 3-kinase pathway (McCudden *et al.*, 2005). They also promote actin polymerization dependent of Rac (see Section Actin polymerization (lamellipodium)) to generate protrusion in the direction of higher concentration of the gradient (Van Haastert and Devreotes, 2004). Localized protrusion occurs because cells are able to perceive small ($\leq 2\%$) differences in the concentration of the chemoattractant, and polarize the signaling intermediates toward the higher concentration of activated receptors (Iglesias and Devreotes, 2008; Figure 4).

Tyr kinase receptors usually have one transmembrane domain and contain a kinase domain in their C-terminus that usually undergoes autophosphorylation. Receptor autophosphorylation recruits other kinases, for example, Src, which trigger signaling cascades that also induce actin polymerization to form protrusions (Endres *et al.*, 2011).

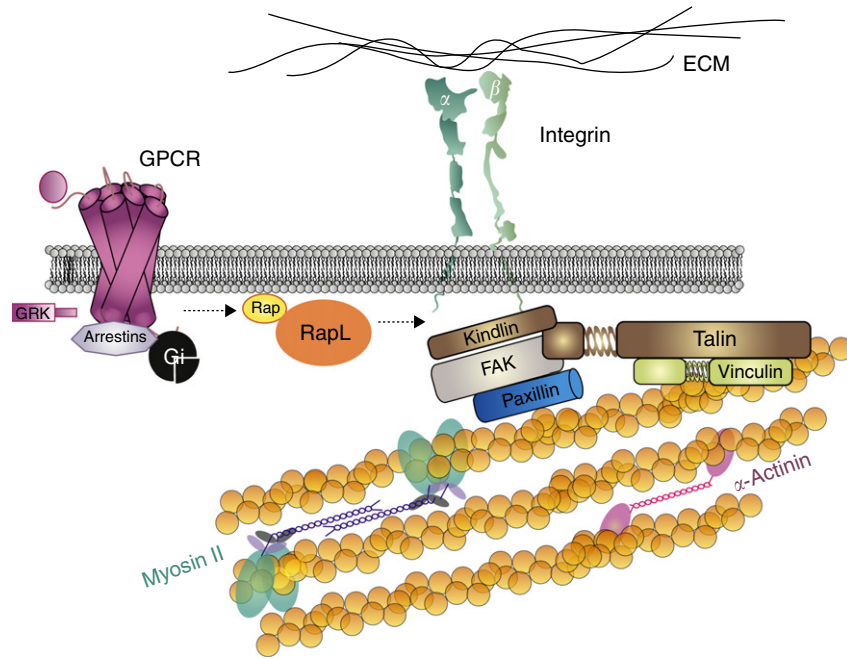


Figure 3 Integrin transactivation by GPCR. In addition to the canonical Gi-dependent and GRK-arrestin internalization pathways, chemoattractant binding to specific GPCR (left, shown in purple) activates the small GTPase Rap (yellow) and its effector RapL (orange), which promote integrin transactivation via kindlins. GPCR also transactivate integrins by promoting talin extension and association to the β chain of the integrin and inducing a conformational change to the integrin that triggers integrin binding to its ligands (ECM) formation of adhesive complexes to ligand-bound integrins. Some adhesion proteins are shown, including vinculin, FAK, and paxillin.

Intracellular Determinants

Actin filaments

Actin is the fundamental cytoskeletal molecule that regulates cell migration. Actin polymerization underlies protrusion formation, adhesion generation, and stabilization/turnover. Stable actin provides cortical integrity and the balance between dynamic, stable, and monomeric actin controls multiple phenomena, from migratory polarization to retraction of the rear, etc. In migrating cells, actin appears in two major forms: polymerized and monomeric. Many studies and reviews have covered the subject (Campellone and Welch, 2010; Pollard and Borisov, 2003).

Polymerized actin appears in different conformations, from densely packed branched thin filaments in lamellipodia (Svitkina and Borisy, 1999) to large and contractile actin bundles in the sides and rear of mesenchymal cells (Cramer *et al.*, 1997). Other migratory structures are podosomes and invadopodia, which are actin-rich rings that support cell migration and spatially confine metalloprotease secretion (Linder and Kopp, 2005). Actin also forms a subcortical network that controls the hydrostatic pressure of the cytoplasm as well as its integrity (see blebbing motility, see Section Blebbing motion). Each of these structures displays different architectures and assembly dynamics. This is due to the different actin-binding proteins in each structure as well as the mechanical properties of the subcellular region in which they each appear. In the lamellipodium, the actin mostly adopts a branched conformation due to the activity of the Arp2/3 complex. The Arp2/3 complex is an actin nucleator that binds to the side of an existing actin filament and generates a new

barbed end forming an angle of approximately 70° with the mother filament. At the tip of the lamellipodium, formins, for example, Diaphanous proteins, nucleate actin monomer incorporation to the barbed ends in a processive manner (Campellone and Welch, 2010). At the base of the lamellipodium, cofilin and normal filament aging disassemble actin filaments (Pollard and Borisy, 2003).

Actin dynamics direct adhesion formation and disassembly. Adhesions are linked to actin and disassemble in its absence. Likewise, adhesion elongation is driven by actin bundling in the lamellum. In areas of dendritic actin, actin cross-linkers are not abundant (Zaidel-Bar *et al.*, 2003). Conversely, NMII and α -actinin are both found in the lamellum and promote the actin assembly into larger bundles. Actin bundling also regulates cortical integrity and membrane coherence (Choi *et al.*, 2008; Cai *et al.*, 2010).

Microtubules

The role of microtubules in cell migration has been the focus of recent reviews (Etienne-Manneville, 2013; Kaverina and Straube, 2011). Migrating mesenchymal cells devoid of microtubules disconnect the front from the rear and lose cellular coherence (Verkhovsky *et al.*, 1999; Vicente-Manzanares *et al.*, 2011). However, amoeboid cells migrate almost normally (Ratner *et al.*, 1997). Microtubules and microtubule-associated proteins also control the relative positioning of the microtubule-organizing center (MTOC) and the nucleus during mesenchymal cell migration (Gomes *et al.*, 2005) and target adhesions to promote their disassembly (Kaverina *et al.*, 1999).

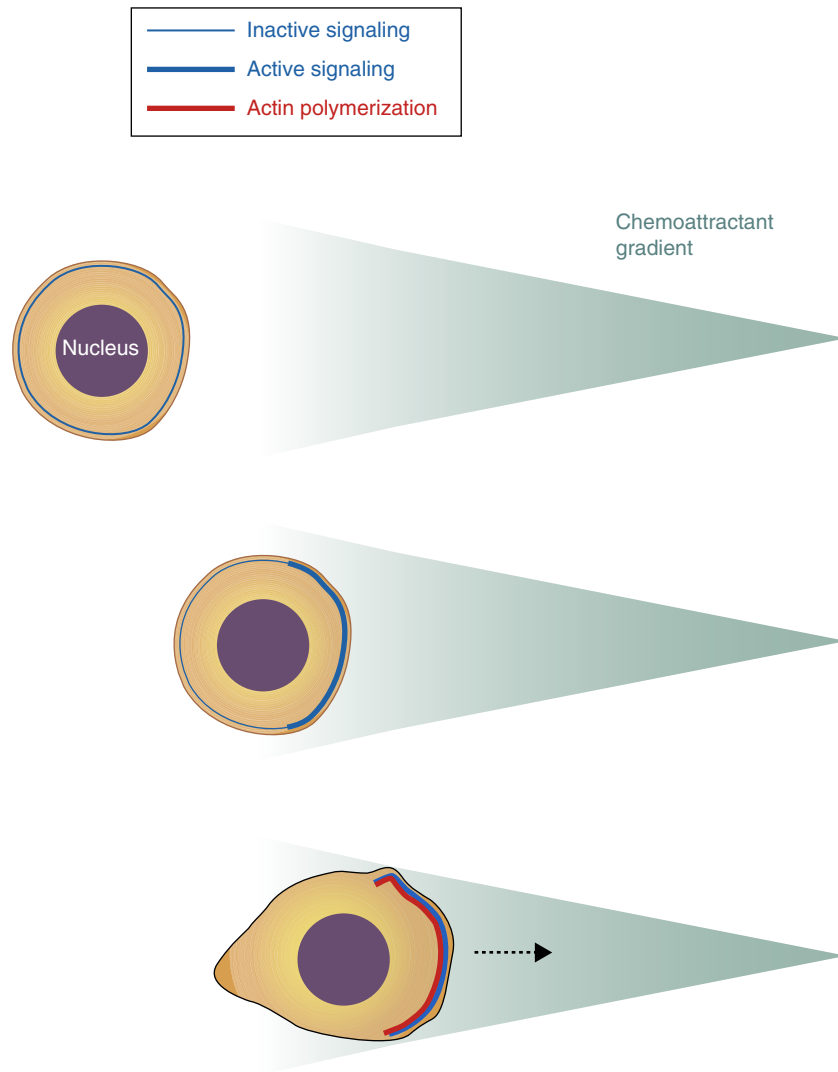


Figure 4 Cellular response to a uniform chemoattractant gradient. Top row, cells display initial signaling quiescence (shown as thin blue line) when it does not perceive the gradient. Middle row, polarized activation (depicted as thick blue line) of the signaling upon gradient sensing on one side of the cell (right side). Bottom row, polarized signaling promotes localized actin polymerization, making a leading edge in the direction of the gradient.

Regulation and integration

In this section, we overview several well-defined regulatory pathways involved in the generation of different migratory cellular structures. For more details, see [Vicente-Manzanares et al. \(2005\)](#).

Actin polymerization (lamellipodium)

Actin polymerization is controlled by small Rho GTPases (mainly RhoA and Rac1). Integrins cluster and activate Rho activators (GEFs) in nascent adhesions ([Kuo et al., 2011](#)), while chemoattractant receptors activate GEFs at the leading edge of the plasma membrane ([Iglesias and Devreotes, 2012](#)). Initially, RhoA activates formins at the very leading edge to drive actin polymerization at the barbed ends closer to the plasma membrane; immediately afterwards, Rac1 activation drive local nucleation of actin

filaments in the lamellipodium via Arp2/3 ([Machacek et al., 2009](#); [Figure 5](#)).

Actin reorganization (lamellum)

As adhesions grow, Rac1 signaling subsides and RhoA signaling becomes predominant ([Machacek et al., 2009](#)). In this area RhoA acts through the Ser/Thr kinase Rho-associated protein kinase (ROCK) to promote NMII recruitment and activation ([Totsukawa et al., 2004](#)). NMII and other nonmotor actin cross-linkers induce actin filament coalescence into bundles. In turn, actin bundling increases adhesion stabilization ([Vicente-Manzanares et al., 2011](#)).

Front-back polarity

Different molecular mechanisms, for example, microtubules and the Cdc42/aPKC/Par3/6 complex coordinately control

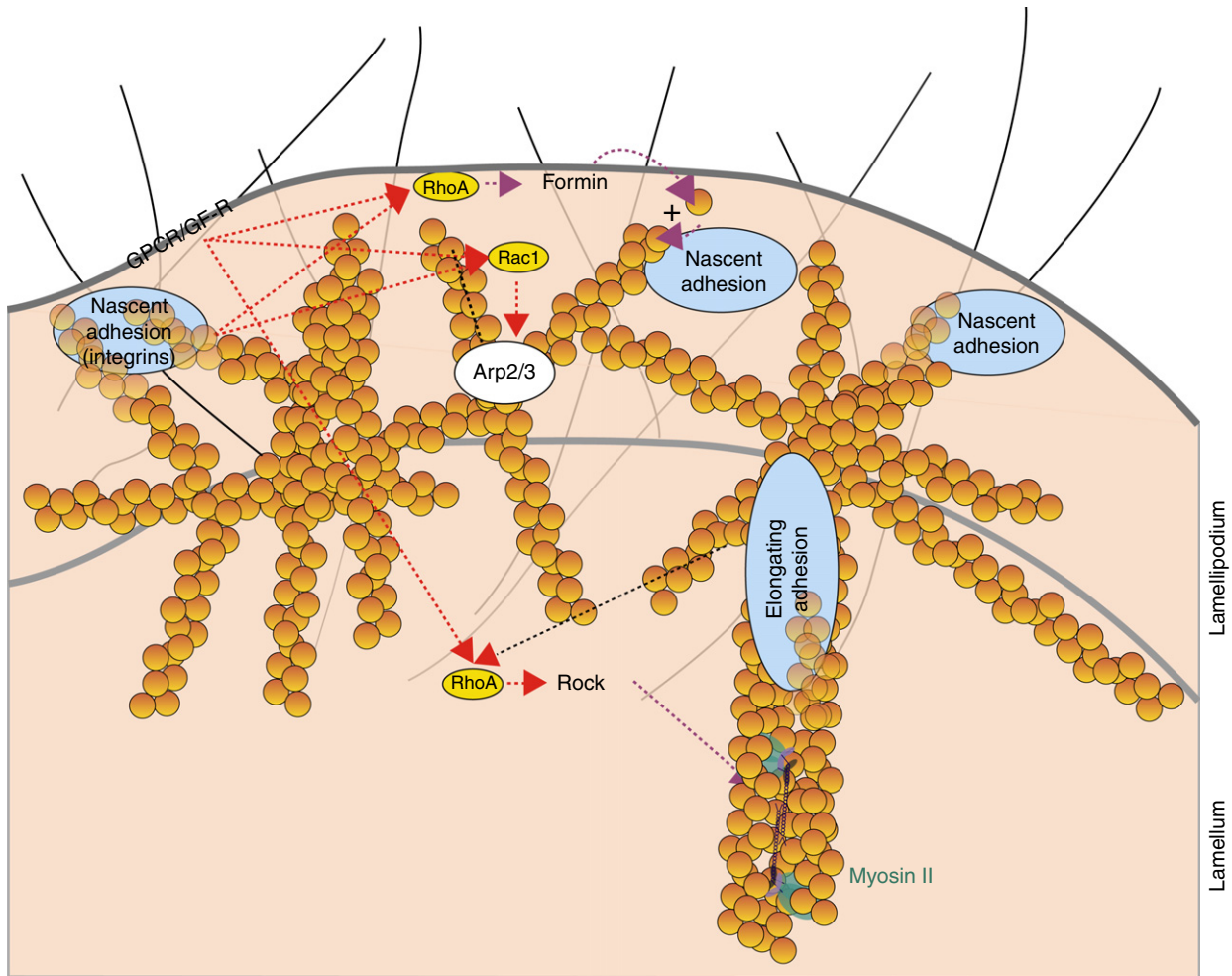


Figure 5 Actin organization at the leading edge. Image depicts a representation of the major pathways of actin organization downstream of chemoattractant and integrin receptors. GPCR or Tyr kinase-type receptors (GF-R) trigger activation of small GTPases and actin polymerization. RhoA (yellow) is activated in the lamellipodium and triggers actin polymerization at the barbed end (+) via formins. These receptors also activate Arp2/3 via Rac1 and promote actin branching at a fixed angle (dotted lines). During lamellipodial formation, actin polymerization guides nascent adhesion formation and integrins also activate the same GTPases. At the lamellum, signaling promotes myosin II activation via RhoA/ROCK, which in turn induces actin bundling and adhesion maturation.

nuclear localization and the relative position of the front and rear in response to migratory cues (Gundersen and Worman, 2013).

Rear detachment

Three mechanisms have been postulated: NMII retraction (Chrzanowska-Wodnicka and Burridge, 1996), proteolytic activation (Huttenlocher *et al.*, 1997), and integrin endocytosis, probably dependent of microtubules (Ezratty *et al.*, 2009; Kaverina *et al.*, 1999).

Migration in Development and Homeostasis

Migratory events underlie embryo and organ development. It also maintains homeostasis of the adult and participates in normal regeneration as well in the immune response.

Embryonic Development

During gastrulation, the cells that form the prospective mesoderm and endoderm invade the ectoderm to form the three primordial layers. This is a process of collective migration in chains, in which prospective mesoderm cells constitute the leading edge of the chain toward the animal pole of the embryo (Keller, 2005). Directional guidance of these chains involves haptotactic (immobilized/adherent) and chemotactic (soluble) cues, including platelet-derived growth factor (PDGF) and members of the Wnt family (Nagel *et al.*, 2004). Haptotactic gradients of morphogens evolve over time and space to segregate the different cell types into their proper locations to begin differentiation. Migratory events also shape the later stages of embryonic development. For example, during dorsal closure, border cells migrate to occlude the amnioserosa (Young *et al.*, 1993). Closure requires mechanical tension at the leading edges of the layer: the edges pull side to

side to close the gap (Franke *et al.*, 2005). In muscle development, progenitors migrate into the limbs driven by hepatocyte growth factor (HGF), which is expressed gradient-wise by non-progenitor mesoderm cells. HGF binds to the c-Met receptor expressed by the progenitors (Heymann *et al.*, 1996).

Migration in the Central Nervous System

Migration is most active during formation of the Central Nervous System (CNS). Although the initial signals that form the neural plate induce differentiation of the neural plate, the segregation of the brain into its definitive compartments requires neuroblast migration, which occurs following these models:

- Tangential migration: Neuroblasts navigate in a mesenchymal manner toward their differentiation niches. They are guided by an exquisite balance of chemoattractant, chemorepellent, and adhesive signals (Nakajima, 2007). An example is the migration of neuroblasts from the sub-ventricular zone (SVZ) to the olfactory bulb (OB) (Muras and Horwitz, 2004).
- Radial migration: Initially, progenitors migrate as individual cells and extend long processes as they differentiate into radial glial cells, protruding toward the pia and contracting the cell body toward the leading edge. These cells can differentiate into mature glial cells or neurons (Fricker-Gates, 2006). Their long processes act as haptotactic/durotactic cues used by subsequent waves of individual neuroblasts to migrate to the different layers of the brain. Confinement to the processes is optimized by chemorepellent signals on the sides of the cue (Marin and Rubenstein, 2003).
- Axophilic migration: Cells use their own axons as a guidance cue to move the soma (Wray, 2010).

Inflammation

Inflammation in response to a pathogen involves two major migratory steps: emigration of the responder leukocytes from the bloodstream into the target tissue and tissue migration of the leukocytes. The former is coordinately regulated by shear stress and transmembrane receptors of different families, e.g., selectins, integrins, and chemokine receptors expressed by leukocytes that interact with endothelial ligands (Vestweber, 2012).

In tissues, leukocytes migrate in an amoeboid manner. The role of integrin-based adhesion is diminished compared to mesenchymal cells in 2-D and 3-D and amoeboid cells in 2-D. This is due to the geometry of the ECM, which displays narrow openings through which leukocytes squeeze using the sides of the gaps as footholds (Renkawitz and Sixt, 2010). For guidance, cells use gradients of immobilized chemoattractants, for example, CCL21 in skin (Weber *et al.*, 2013).

Migration in Pathology

Abnormal migratory phenomena are also cause or part of the onset of several pathologies, most notably cancer and autoimmune disease.

Cancer

Tumor cells migrate to distal sites to form secondary tumors (metastasis). They undergo egress from the primary tumor and move toward blood vessels or lymphatics. Then, they migrate to distal sites and extravasate by mimicking the process used by leukocytes, or bursting through the capillary vessels. Once within the new tissue, tumor cells migrate to explore their new microenvironment, immobilize and proliferate to generate a new tumor.

At the edges of the tumor, migration is mainly multicellular, with chains or sheets of tumor cells emerging from the primary tumor in waves. As they migrate away from the primary tumor, some of these begin migrating individually (Friedl *et al.*, 2012). The predominant migratory mode depends on the origin of the tumor. Carcinomas, which are epithelial in origin, are thought to migrate mostly in a mesenchymal manner. Conversely, leukemic cells are hematopoietic and thus migrate in an amoeboid manner. However, some tumor cells can switch between modes rapidly to adapt to the mechano-chemical properties of their microenvironment (Bergert *et al.*, 2012).

The prevalent hypothesis regarding initiation of carcinoma migration involves the conversion of epithelial-like tumor into migration-competent cells through the epithelial-mesenchymal transition (EMT). This is a genetic switch that inactivates genes that confer epithelial traits to the cells and activate genes that enable migration in a mesenchymal manner. Once motile, cells are able to switch between migrating modes in response, for example, to variations in the composition of the ECM or structural or geometrical differences (Bergert *et al.*, 2012).

This is a very active area of research since metastasis is responsible for approximately 90% of cancer-related deaths. Therefore, the ability to halt cancer cell migration becomes the focal point of potential anti-metastasis therapy (Stock *et al.*, 2013).

Angiogenesis is an additional component of the onset of cancer in which migration plays a fundamental role. Tumors secrete chemotactic molecules for endothelial precursors, for example, VEGF. Precursors then invade the tumor and form new blood vessels connected to the general bloodstream, thereby generating a source of oxygen and nutrients that feed the tumors directly. Therefore, another therapeutic opportunity is blocking endothelial cell migration to the tumor to prevent *de novo* vascularization, thereby starving the tumor (Weis and Cheresh, 2011).

Autoimmune Disease

Part of the onset of autoimmune diseases is the migration of leukocytes to sites where they recognize self-antigens as foreign and mount an immune response against them. Some examples include rheumatoid arthritis (RA), diabetes, psoriasis, multiple sclerosis (MS), Crohn's disease (CD), and some others. The etiology of these diseases is complex and includes genetic and environment factors. However, the manifestation of the disease includes self-perpetuation of the inflammatory response and leukocyte infiltration (mainly neutrophils and autorreactive T cells) that causes inflammatory damage (Wright *et al.*, 2010). CD4+ T cell migration into the tissue

promotes and sustains inflammation (Toh and Miossec, 2007); whereas migrating CD8⁺ T cells destroy tissue cells that display autorreactive epitopes, for example, Langerhans cells in diabetes (Tsai *et al.*, 2008). Anti-migratory therapy works to treat some of these diseases. For example, natalizumab is a humanized anti-CD49d ($\alpha 4$) antibody that is used to control multiple sclerosis. Its mechanism of action prevents leukocytes from migrating to the CNS (Von Andrian and Engelhardt, 2003). Another example is efalizumab, which targets CD11a (αL) and has been used to treat psoriasis by preventing T cell and neutrophil migration to the injured skin (Frampton and Plosker, 2009).

Harnessing Cell Migration in Therapy

As described in the previous section, inadequate migration can cause severe disease. Conversely, there are instances in which the lack of migration prevents therapy. For example, in spinal cord injury, lack of neuroblast migration to the site of injury is a major reason for regenerative failure. A similar reason prevents myocardial regeneration upon congestive heart failure and necrosis. Most stem cell based approaches fail due to the inability of these cells to position correctly within the injured tissue prior to undergoing differentiation into target cells.

Several reasons concur to cause migratory failure. For example, damaged or scarred tissue does not provide an adequate microenvironment to stimulate and support stem cell migration. Additional reasons include physical resistance due to fibrosis and residual inflammation that can kill the regenerating cells. Future research will address these issues to be able to promote controlled cell migration in these hostile environments to support efficient regenerative therapies.

Conclusions

- Cell migration is a fundamental process in physiology and pathology that requires the interaction of many proteins and cell structures in response to extracellular and intracellular signals.
- Eukaryotic cells display different migratory modes: Single-cell migration is typical of highly differentiated cells and collective migration refers to the simultaneous and coordinated migration of multiple cells while in contact with others.
- Single-cell migration can be mesenchymal or amoeboid. Mesenchymal movement is slow and features strong adhesion to the ECM and complex cytoskeletal structures and produces tissue remodeling. Amoeboid migration is faster, with weak adhesion and modest perturbation to tissue morphology.
- Migratory processes are crucial in embryo and organ development. In later stages, it maintains tissue homeostasis and participates in normal regeneration as well as in the immune response.
- Abnormal migratory phenomena are cause or part of the onset of several pathologies, most notably cancer and autoimmune disease.

Acknowledgments

MV-M is supported by the Ramón y Cajal Research Program and grants from the Spanish Ministry of Science and Education (MINECO, SAF-2011-24953) the Fundación Ramón Areces (Spain) and a Marie Curie Career Reintegration Grant (CIG293719). AO is a member of the Immunology Specialist Program sponsored by the Spanish Ministry of Health.

See also: Cell Communication: Cellular Outcomes: Regulation of Cell Polarity. Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell Adhesion to the Extracellular Matrix; Cell–Cell Adhesion and the Cytoskeleton; Cell Polarity; Filopodia and Lamellipodia; Rho GTPases; The Extracellular Matrix. Cytoskeleton and Motors: Cytoskeletal Components: Actin Assembly Dynamics and Its Regulation in Motile and Morphogenetic Processes

References

- Aguilar-Cuenca, R., Juanes-Garcia, A., Vicente-Manzanares, M., 2014. Myosin II in mechanotransduction: Master and commander of cell migration, morphogenesis, and cancer. *Cellular and Molecular Life Sciences* 71, 479–492.
- Alexandrova, A.Y., Arnold, K., Schaub, S., *et al.*, 2008. Comparative dynamics of retrograde actin flow and focal adhesions: Formation of nascent adhesions triggers transition from fast to slow flow. *PLoS ONE* 3, e3234.
- Bakolitsa, C., Cohen, D.M., Bankston, L.A., *et al.*, 2004. Structural basis for vinculin activation at sites of cell adhesion. *Nature* 430, 583–586.
- Bergert, M., Chandross, S.D., Desai, R.A., Paluch, E., 2012. Cell mechanics control rapid transitions between blebs and lamellipodia during migration. *Proceedings of the National Academy of Sciences of the United States of America* 109, 14434–14439.
- Cai, Y., Rossier, O., Gauthier, N.C., *et al.*, 2010. Cytoskeletal coherence requires myosin-IIA contractility. *Journal of Cell Science* 123, 413–423.
- Campellone, K.G., Welch, M.D., 2010. A nucleator arms race: Cellular control of actin assembly. *Nature Reviews Molecular Cell Biology* 11, 237–251.
- Carman, C.V., Springer, T.A., 2003. Integrin avidity regulation: Are changes in affinity and conformation underemphasized? *Current Opinion in Cell Biology* 15, 547–556.
- Charras, G.T., Coughlin, M., Mitchison, T.J., Mahadevan, L., 2008. Life and times of a cellular bleb. *Biophysical Journal* 94, 1836–1853.
- Chen, L., Vicente-Manzanares, M., Potvin-Trottier, L., Wiseman, P.W., Horwitz, A.R., 2012. The integrin–ligand interaction regulates adhesion and migration through a molecular clutch. *PLoS ONE* 7, e40202.
- Choi, C.K., Vicente-Manzanares, M., Zareno, J., *et al.*, 2008. Actin and alpha-actinin orchestrate the assembly and maturation of nascent adhesions in a myosin II motor-independent manner. *Nature Cell Biology* 10, 1039–1050.
- Chrzanowska-Wodnicka, M., Burridge, K., 1996. Rho-stimulated contractility drives the formation of stress fibers and focal adhesions. *Journal of Cell Biology* 133, 1403–1415.
- Cramer, L.P., Siebert, M., Mitchison, T.J., 1997. Identification of novel graded polarity actin filament bundles in locomoting heart fibroblasts: Implications for the generation of motile force. *Journal of Cell Biology* 136, 1287–1305.
- Deakin, N.O., Turner, C.E., 2008. Paxillin comes of age. *Journal of Cell Science* 121, 2435–2444.
- Dumbauld, D.W., Lee, T.T., Singh, A., *et al.*, 2013. How vinculin regulates force transmission. *Proceedings of the National Academy of Sciences of the United States of America* 110, 9788–9793.
- Endres, N.F., Engel, K., Das, R., Kovacs, E., Kuriyan, J., 2011. Regulation of the catalytic activity of the EGF receptor. *Current Opinion in Structural Biology* 21, 777–784.
- Etienne-Manneville, S., 2013. Microtubules in cell migration. *Annual Review of Cell and Developmental Biology* 29, 471–499.
- Ezratty, E.J., Bertaux, C., Marcantonio, E.E., Gundersen, G.G., 2009. Clathrin mediates integrin endocytosis for focal adhesion disassembly in migrating cells. *Journal of Cell Biology* 187, 733–747.

- Frampton, J.E., Plosker, G.L., 2009. Efalizumab: A review of its use in the management of chronic moderate-to-severe plaque psoriasis. *American Journal of Clinical Dermatology* 10, 51–72.
- Franke, J.D., Montague, R.A., Kiehart, D.P., 2005. Nonmuscle myosin II generates forces that transmit tension and drive contraction in multiple tissues during dorsal closure. *Current Biology* 15, 2208–2221.
- Fricker-Gates, R.A., 2006. Radial glia: A changing role in the central nervous system. *NeuroReport* 17, 1081–1084.
- Friedl, P., Locker, J., Sahai, E., Segall, J.E., 2012. Classifying collective cancer cell invasion. *Nature Cell Biology* 14, 777–783.
- Friedl, P., Weigelin, B., 2008. Interstitial leukocyte migration and immune function. *Nature Immunology* 9, 960–969.
- Friedl, P., Wolf, K., 2010. Plasticity of cell migration: A multiscale tuning model. *Journal of Cell Biology* 188, 11–19.
- Gaggioli, C., Hooper, S., Hidalgo-Carcedo, C., *et al.*, 2007. Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells. *Nature Cell Biology* 9, 1392–1400.
- Gomes, E.R., Jani, S., Gundersen, G.G., 2005. Nuclear movement regulated by Cdc42, MRCK, myosin, and actin flow establishes MTOC polarization in migrating cells. *Cell* 121, 451–463.
- Gundersen, G.G., Worman, H.J., 2013. Nuclear positioning. *Cell* 152, 1376–1389.
- Heymann, S., Koudrova, M., Arnold, H., Koster, M., Braun, T., 1996. Regulation and function of SF/HGF during migration of limb muscle precursor cells in chicken. *Developmental Biology* 180, 566–578.
- Huttenlocher, A., Palecek, S.P., Lu, Q., *et al.*, 1997. Regulation of cell migration by the calcium-dependent protease calpain. *Journal of Biological Chemistry* 272, 32719–32722.
- Hynes, R.O., 2002. Integrins: Bidirectional, allosteric signaling machines. *Cell* 110, 673–687.
- Iglesias, P.A., Devreotes, P.N., 2008. Navigating through models of chemotaxis. *Current Opinion in Cell Biology* 20, 35–40.
- Iglesias, P.A., Devreotes, P.N., 2012. Biased excitable networks: How cells direct motion in response to gradients. *Current Opinion in Cell Biology* 24, 245–253.
- Jacobelli, J., Estlin Matthews, M., Chen, S., Krummel, M.F., 2013. Activated T cell trans-endothelial migration relies on myosin-IIA contractility for squeezing the cell nucleus through endothelial cell barriers. *PLoS ONE* 8, e75151.
- Kaverina, I., Krylyshkina, O., Small, J.V., 1999. Microtubule targeting of substrate contacts promotes their relaxation and dissociation. *Journal of Cell Biology* 146, 1033–1044.
- Kaverina, I., Straube, A., 2011. Regulation of cell migration by dynamic microtubules. *Seminars in Cell & Developmental Biology* 22, 968–974.
- Keller, R., 2005. Cell migration during gastrulation. *Current Opinion in Cell Biology* 17, 533–541.
- Kuo, J.C., Han, X., Hsiao, C.T., Yates 3rd, J.R., Waterman, C.M., 2011. Analysis of the myosin-II-responsive focal adhesion proteome reveals a role for beta-Pix in negative regulation of focal adhesion maturation. *Nature Cell Biology* 13, 383–393.
- Kwon, K.W., Park, H., Doh, J., 2013. Migration of T cells on surfaces containing complex nanotopography. *PLoS ONE* 8, e73960.
- Lammermann, T., Bader, B.L., Monkley, S.J., *et al.*, 2008. Rapid leukocyte migration by integrin-independent flowing and squeezing. *Nature* 453, 51–55.
- Lammermann, T., Sixt, M., 2009. Mechanical modes of 'amoeboid' cell migration. *Current Opinion in Cell Biology* 21, 636–644.
- Legant, W.R., Chen, C.S., Vogel, V., 2012. Force-induced fibronectin assembly and matrix remodeling in a 3D microtissue model of tissue morphogenesis. *Integrative Biology (Camb)* 4, 1164–1174.
- Linder, S., Kopp, P., 2005. Podosomes at a glance. *Journal of Cell Science* 118, 2079–2082.
- Luttrell, L.M., Lefkowitz, R.J., 2002. The role of beta-arrestins in the termination and transduction of G-protein-coupled receptor signals. *Journal of Cell Science* 115, 455–465.
- Machacek, M., Hodgson, L., Welch, C., *et al.*, 2009. Coordination of Rho GTPase activities during cell protrusion. *Nature* 461, 99–103.
- Marin, O., Rubenstein, J.L., 2003. Cell migration in the forebrain. *Annual Review of Neuroscience* 26, 441–483.
- McCudden, C.R., Hains, M.D., Kimple, R.J., Siderovski, D.P., Willard, F.S., 2005. G-protein signaling: Back to the future. *Cellular and Molecular Life Sciences* 62, 551–577.
- Murase, S., Horwitz, A.F., 2004. Directions in cell migration along the rostral migratory stream: The pathway for migration in the brain. *Current Topics in Developmental Biology* 61, 135–152.
- Nagel, M., Tahinci, E., Symes, K., Winklbauer, R., 2004. Guidance of mesoderm cell migration in the *Xenopus* gastrula requires PDGF signaling. *Development* 131, 2727–2736.
- Nakajima, K., 2007. Control of tangential/non-radial migration of neurons in the developing cerebral cortex. *Neurochemistry International* 51, 121–131.
- Palecek, S.P., Loftus, J.C., Ginsberg, M.H., Lauffenburger, D.A., Horwitz, A.F., 1997. Integrin-ligand binding properties govern cell migration speed through cell-substratum adhesiveness. *Nature* 385, 537–540.
- Parsons, J.T., 2003. Focal adhesion kinase: The first ten years. *Journal of Cell Science* 116, 1409–1416.
- Pollard, T.D., Borisy, G.G., 2003. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* 112, 453–465.
- Raab, M., Swift, J., Dingal, P.C., *et al.*, 2012. Crawling from soft to stiff matrix polarizes the cytoskeleton and phosphoregulates myosin-II heavy chain. *Journal of Cell Biology* 199, 669–683.
- Ratner, S., Sherrod, W.S., Lichlyter, D., 1997. Microtubule retraction into the uropod and its role in T cell polarization and motility. *Journal of Immunology* 159, 1063–1067.
- Renkawitz, J., Sixt, M., 2010. Mechanisms of force generation and force transmission during interstitial leukocyte migration. *EMBO Reports* 11, 744–750.
- Ridley, A.J., Schwartz, M.A., Burridge, K., *et al.*, 2003. Cell migration: Integrating signals from front to back. *Science* 302, 1704–1709.
- Sanz-Moreno, V., Gadea, G., Ahn, J., *et al.*, 2008. Rac activation and inactivation control plasticity of tumor cell movement. *Cell* 135, 510–523.
- Shaw, T.J., Martin, P., 2009. Wound repair at a glance. *Journal of Cell Science* 122, 3209–3213.
- Solekwa, J., Trivedi, N., Govek, E.E., *et al.*, 2009. Myosin II motors and F-actin dynamics drive the coordinated movement of the centrosome and soma during CNS glial-guided neuronal migration. *Neuron* 63, 63–80.
- Stock, A.M., Troost, G., Niggemann, B., Zanker, K.S., Entschladen, F., 2013. Targets for anti-metastatic drug development. *Current Pharmaceutical Design* 19, 5127–5134.
- Svitkina, T.M., Borisy, G.G., 1999. Arp2/3 complex and actin depolymerizing factor/cofilin in dendritic organization and treadmilling of actin filament array in lamellipodia. *Journal of Cell Biology* 145, 1009–1026.
- Theveneau, E., Mayor, R., 2012. Neural crest delamination and migration: From epithelium-to-mesenchyme transition to collective cell migration. *Developmental Biology* 366, 34–54.
- Toh, M.L., Miossec, P., 2007. The role of T cells in rheumatoid arthritis: New subsets and new targets. *Current Opinion in Rheumatology* 19, 284–288.
- Totsukawa, G., Wu, Y., Sasaki, Y., *et al.*, 2004. Distinct roles of MLCK and ROCK in the regulation of membrane protrusions and focal adhesion dynamics during cell migration of fibroblasts. *Journal of Cell Biology* 164, 427–439.
- Tsai, S., Shameli, A., Santamaria, P., 2008. CD8+ T cells in type 1 diabetes. *Advances in Immunology* 100, 79–124.
- Van Haastert, P.J., Devreotes, P.N., 2004. Chemotaxis: Signalling the way forward. *Nature Reviews Molecular Cell Biology* 5, 626–634.
- Verkhovsky, A.B., Svitkina, T.M., Borisy, G.G., 1999. Self-polarization and directional motility of cytoplasm. *Current Biology* 9, 11–20.
- Vestweber, D., 2012. Novel insights into leukocyte extravasation. *Current Opinion in Hematology* 19, 212–217.
- Vicente-Manzanares, M., Newell-Litwa, K., Bachir, A.I., Whitmore, L.A., Horwitz, A.R., 2011. Myosin IIA/IIB restrict adhesive and protrusive signaling to generate front-back polarity in migrating cells. *Journal of Cell Biology* 193, 381–396.
- Vicente-Manzanares, M., Webb, D.J., Horwitz, A.R., 2005. Cell migration at a glance. *Journal of Cell Science* 118, 4917–4919.
- Von Andrian, U.H., Engelhardt, B., 2003. Alpha 4 integrins as therapeutic targets in autoimmune disease. *New England Journal of Medicine* 348, 68–72.
- Weber, M., Hauschild, R., Schwarz, J., *et al.*, 2013. Interstitial dendritic cell guidance by haptotactic chemokine gradients. *Science* 339, 328–332.
- Weis, S.M., Cheresh, D.A., 2011. Tumor angiogenesis: Molecular pathways and therapeutic targets. *Nature Medicine* 17, 1359–1370.
- Wray, S., 2010. From nose to brain: Development of gonadotrophin-releasing hormone-1 neurones. *Journal of Neuroendocrinology* 22, 743–753.
- Wright, H.L., Moots, R.J., Bucknall, R.C., Edwards, S.W., 2010. Neutrophil function in inflammation and inflammatory diseases. *Rheumatology (Oxford)* 49, 1618–1631.
- Young, P.E., Richman, A.M., Ketchum, A.S., Kiehart, D.P., 1993. Morphogenesis in *Drosophila* requires nonmuscle myosin heavy chain function. *Genes & Development* 7, 29–41.
- Zaidel-Bar, R., Ballestrem, C., Kam, Z., Geiger, B., 2003. Early molecular events in the assembly of matrix adhesions at the leading edge of migrating cells. *Journal of Cell Science* 116, 4605–4613.
- Zaidel-Bar, R., Itzkovitz, S., Ma'ayan, A., Iyengar, R., Geiger, B., 2007. Functional atlas of the integrin adhesome. *Nature Cell Biology* 9, 858–867.

Rho GTPases

MJ Lukey and RA Cerione, Cornell University, Ithaca, NY, USA

© 2016 Elsevier Inc. All rights reserved.

General Features

There appear to be at least 20 members of the Rho GTPase family in mammals, based on genome sequence analyses (Jaffe and Hall, 2005; Boureux *et al.*, 2007). Rac, Rho, and Cdc42 each have more than one form (e.g., Rac1, Rac2, RhoA, RhoC etc.) and are found in virtually all animals. Cdc42 is rigorously conserved from yeast to humans, which suggested that it along with its close cousins were likely to play fundamentally important roles in cell biology. Indeed, their requirement for the maintenance of cell shape, polarity, and for the necessary actin cytoskeletal changes supporting these processes, is consistent with this assumption. Like the Ras GTPases, members of the Rho subfamily undergo a tightly regulated exchange of GDP for GTP in order to mediate their specific signaling activities. Moreover, similar to the Ras GTPases, the Rho proteins make use of conformationally sensitive regions, referred to as Switch 1 and 2, to trigger their signaling events, as well as a regulated GTP-hydrolytic activity to switch off their signal outputs.

Nonetheless, the Rho GTPases have a number of important features that distinguish them from their Ras GTPase counterparts. One of these involves the posttranslational modifications that occur within the carboxyl terminal CAAX box, and which are essential for membrane association. While the Ras GTPases contain a farnesyl moiety at their CAAX box, as well as in some cases (H- and N-Ras) a palmitoyl group at a cysteine located 4 residues upstream from the cysteine that starts the CAAX motif, the Rho GTPases typically contain a geranyl-geranyl moiety at their CAAX box cysteine residue (Michaelson *et al.*, 2001). One exception is the brain-specific isoform of Cdc42, which has a double cysteine motif within its CCAX carboxyl terminal and has been suggested to undergo a combination of geranyl-geranylation, palmitoylation, and farnesylation modifications (Kang *et al.*, 2008).

A second distinction between the Ras and Rho proteins is that a significant fraction of the Rho GTPases is typically found in the cytosol. This cytosolic localization is due to the actions of a family of regulators collectively called RhoGDI (for Rho GDP-dissociation inhibitor). Yoshimi Takai, who together with his colleagues was largely responsible for showing that the Ras proteins were the prototypes for a large superfamily of small GTPases (originally called Smgs for small molecular weight GTPases) (Takai *et al.*, 1988), was the first to identify two classes of regulatory proteins named GDI that slowed GDP dissociation from the Rab and Rho GTPases (thus called the Rab and RhoGDIs, respectively) (Sasaki *et al.*, 1990; Ueda *et al.*, 1990; Fukumoto *et al.*, 1990). The GDIs also were capable of facilitating the membrane-to-cytosolic transition of the Rab and Rho proteins (Leonard *et al.*, 1992). The RhoGDIs include RhoGDI α (or RhoGDI1), the hematopoietic-specific RhoGDI β (RhoGDI2), and RhoGDI γ (RhoGDI3). The X-ray crystal structure for GDP-bound Cdc42 complexed to RhoGDI demonstrated how the geranyl-geranyl moiety of Cdc42 fits

into a hydrophobic pocket within the GDI molecule (Hoffman *et al.*, 2000). More recently, new insights were obtained into the mechanism by which RhoGDI plays one of its most important roles, namely, in stabilizing the soluble, cytosolic state of Cdc42 (Johnson *et al.*, 2009). Contrary to what was originally assumed, RhoGDI does not appear to actively extract Cdc42 from membranes. Rather, the membrane association of Cdc42 is a dynamic process, i.e., it associates and dissociates from membranes with measurable rates. RhoGDI binds Cdc42 while it is associated with membranes. Surprisingly, it was shown that the RhoGDI-Cdc42 complex dissociates from membranes with a rate that matches (and does not greatly exceed) that for the dissociation of Cdc42 from membranes in the absence of RhoGDI. Once the RhoGDI-Cdc42 complex has dissociated from membranes, RhoGDI can sequester the geranyl-geranyl moiety of Cdc42 and help stabilize the GTPase in the cytosol until the appropriate time for Cdc42 to engage its upstream activators or downstream signaling targets.

Regulation of the GTP-Binding/GTP-Hydrolytic Cycles of Rho GTPases

As is the case for the Ras proteins, the GTP-binding and GTP-hydrolytic activities of the Rho GTPases are tightly regulated by the combined actions of guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs) (Figure 1). The founding member of an important superfamily of GEFs is the product of the *DBL* oncogene (Eva and Aaronson, 1985; Hart *et al.*, 1991). Proto-Dbl is an 115 kDa protein, whereas oncogenic Dbl was generated as an outcome of a recombination event when transfecting DNA from a diffuse B-cell lymphoma into fibroblasts. The amino-terminal half of proto-Dbl was removed and replaced by ~10 kDa of unrelated human sequence, yielding the 66 kDa oncogenic Dbl protein. Studies with the recombinant oncogenic Dbl protein showed that the limit region required for GEF activity was within a stretch of ~250 amino acids, subsequently termed the Dbl homology (DH) domain, which is conserved throughout a number of other proteins that would turn out to be Rho GEFs (Hart *et al.*, 1994). Immediately downstream from this region is another conserved stretch of sequence, now known as the Pleckstrin-homology (PH) domain. This tandem arrangement of DH and PH domains would become a characteristic feature of the Dbl-family of Rho GEFs, which includes the *Saccharomyces cerevisiae* Cdc24 protein, the Vav oncogene product which was first identified in hematopoietic tumors, the highly specific Rac-GEF Tiam-1 which was discovered when screening for gene products linked to metastasis, the Cool/Pix proteins, and p115RhoGEF which contains an RGS (regulator of G protein signaling) domain and couples the large G protein G₁₃ to the activation of RhoA (Figures 2(a) and 2(b); Cerione and Zheng, 1996; Rossman *et al.*, 2005). The fact that many of

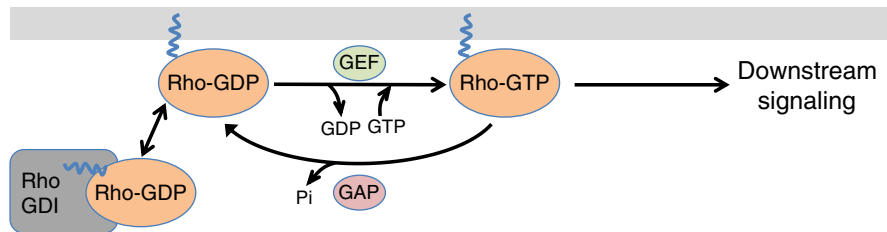


Figure 1 The Rho GTPase cycle. Rho GTPases are inactive when bound to GDP, and are sequestered in the cytoplasm by guanine nucleotide dissociation inhibitors (GDIs). Guanine nucleotide exchange factors (GEFs) stimulate exchange of GDP for GTP, activating the GTPase. When active, Rho GTPases interact with and activate numerous downstream effectors. GTPase-activating proteins (GAPs) stimulate the intrinsic GTPase activity, which generates GDP and inorganic phosphate (Pi), and restores the inactive state.

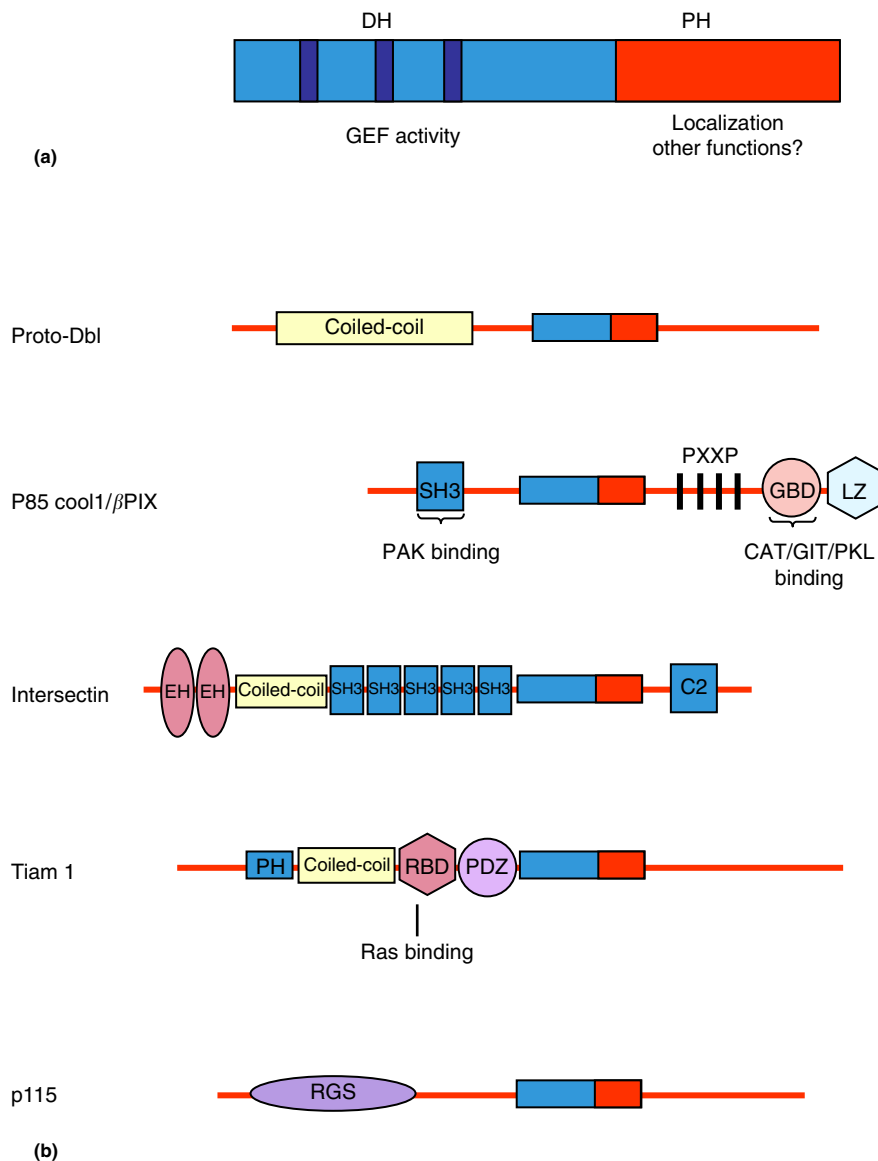


Figure 2 Diagram of DH domain-containing proteins. (a) Schematic diagram of the DH-PH module is shown. (b) The domain architectures of a number of Dbl-family proteins are shown.

these GEFs have been linked to cellular transformation or uncontrolled cell growth, starting with the initial discovery of oncogenic Dbl, represented the first significant clues that Rho

GTPases were involved in mitogenic signaling and cancer progression (see below). Insight into the mechanisms underlying the GEF activity of Dbl-family proteins was obtained

from the X-ray crystal structure of Rac1 in complex with the DH and PH domains of Tiam1 (Worthylake *et al.*, 2000).

A second class of GEFs for Cdc42 and Rac, that lack tandem DH and PH domains, has also been identified and characterized. This is called the DOCK180 (for Director of cytokinesis 180) family, named for its founding member (Cote and Vuori, 2002). The GEFs comprising this family are ~200 kDa in size and contain conserved domains called DHR1 and DHR2 (for DOCK homology regions 1 and 2). Many members of the DOCK180 family have been shown to activate Rac to regulate cell adhesion, cell migration, and phagocytosis. Some members of the family are specific GEFs for Cdc42, including DOCK9 (also known as Zizimin) and DOCK11 (also known as ACG for activated Cdc42-associated GEF), while others like DOCK7 are effective at activating both Rac and Cdc42 (Meller *et al.*, 2002).

The RhoGAPs also make up a large family of regulatory proteins. The founding member was an ~50 kDa protein, originally called p50RhoGAP, although it is also referred to as Cdc42GAP as it is much more effective at stimulating the GTP-hydrolytic activity of Cdc42 than of RhoA (Lancaster *et al.*, 1994). The members of this family share three regions of similarity spanning ~200 amino acids, referred to as the RhoGAP homology domain, and include the DLC1 and DLC2 proteins (for Deleted in liver cancer 1 and 2) which function as putative tumor suppressors, as well as the breakpoint cluster region (Bcr) protein which interestingly also contains tandem DH and PH domains, the 85 kDa regulatory subunit of the PI-3 kinase, and the RasGAP-binding protein p190 (Tcherkezian and Lamarche-Vane, 2007). Thus, like the Dbl-family of GEFs, a number of the RhoGAP family members have been implicated in different aspects of cell growth control and cancer, thereby providing additional links between the Rho GTPases and the transformed state. X-ray crystallographic analysis of Cdc42 bound to the Cdc42GAP shows that RhoGAPs catalyze GTP hydrolysis through a similar mechanism to that used by RasGAPs (Nassar *et al.*, 1998).

Rho GTPases and Their Roles in Regulating the Actin Cytoskeleton

Dynamic regulation of the actin cytoskeleton is required for many cellular processes, including changes in cell shape, cell migration, and cytokinesis. The results from microinjection experiments provided the first indication that the Rho GTPases are involved in controlling cytoskeletal organization. In a set of landmark publications, Ridley and Hall (1992) demonstrated that RhoA stimulated actin stress fiber and focal adhesion formation when microinjected into Swiss 3T3 fibroblasts. A number of different growth factors including PDGF, EGF, and bombesin stimulated actin reorganization and triggered a rapid formation of membrane ruffles, which was subsequently accompanied by the formation of stress fibers. Growth factor-stimulated focal adhesion assembly and actin stress fiber formation were blocked when endogenous RhoA function was inhibited by ADP ribosylation of the RhoA protein by the exoenzyme C3 transferase from *Clostridium botulinum*. Ridley, Hall and their colleagues also showed that the microinjection of Rac1 into confluent serum-starved Swiss

3T3 fibroblasts stimulated actin filament accumulation at the plasma membrane, resulting in the formation of lamellipodia (broad actin-based projections at the leading edge of motile cells) or membrane ruffles (Ridley *et al.*, 1992). Various growth factors and activated H-Ras were also shown to induce membrane ruffling with this response being blocked by the injection of a dominant-negative Rac mutant, which binds tightly to endogenous Rac-GEFs but cannot undergo guanine nucleotide exchange.

The likelihood that Cdc42 would similarly impact the cytoskeletal architecture seemed evident from its original discovery in *S. cerevisiae* by Johnson and Pringle (1990), where it was implicated in the establishment of cell polarity and the proper localization of budding sites. The human homolog was discovered based on its apparent involvement in EGF receptor signaling (Hart *et al.*, 1990). Remarkably, the cloning of the cDNAs encoding the yeast and human proteins, both reported in 1990 (Johnson and Pringle, 1990; Shinjo *et al.*, 1990; Munemitsu *et al.*, 1990), showed that they were more than 90% identical in amino acid sequence. This extraordinary conservation from yeast to humans foretold that Cdc42 would play fundamentally important roles in cell biology. Subsequently, it was shown that microinjection of activated Cdc42 mutants into Swiss 3T3 fibroblasts stimulated the formation of filopodia (slender actin-based projections that extend beyond the leading edge of lamellipodia) and microspikes (Nobes and Hall, 1995), and that at least in this cellular context, there is a GTPase cascade (Cdc42 to Rac to RhoA) that induces a specific sequence of cytoskeletal alterations (Figure 3).

The Rho GTPases regulate the actin cytoskeleton through several effector proteins (Figure 4), which together coordinate actin polymerization, branching, and actomyosin contractibility (Sit and Manser, 2011). Activated Cdc42 and Rac1 promote filamentous actin (F-actin) branching by indirectly activating the actin-related protein (ARP)2/3 complex (Jaffe and Hall, 2005). Once activated, the ARP2/3 complex binds to preexisting F-actin and nucleates the growth of branching daughter filaments, which can push forward the leading edge of motile cells (Pollard, 2007). Cdc42 signals to ARP2/3 through members of the Wiskott–Aldrich Syndrome protein (WASP) family, which, *in vitro*, are auto-inhibited by an intramolecular interaction between the C-terminal functional 'VCA' module and the N-terminal regulatory domain. Binding

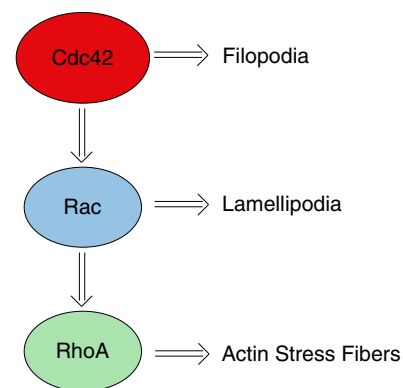


Figure 3 Depiction of a Rho GTPase cascade that regulates cell shape and actin cytoskeletal organization.

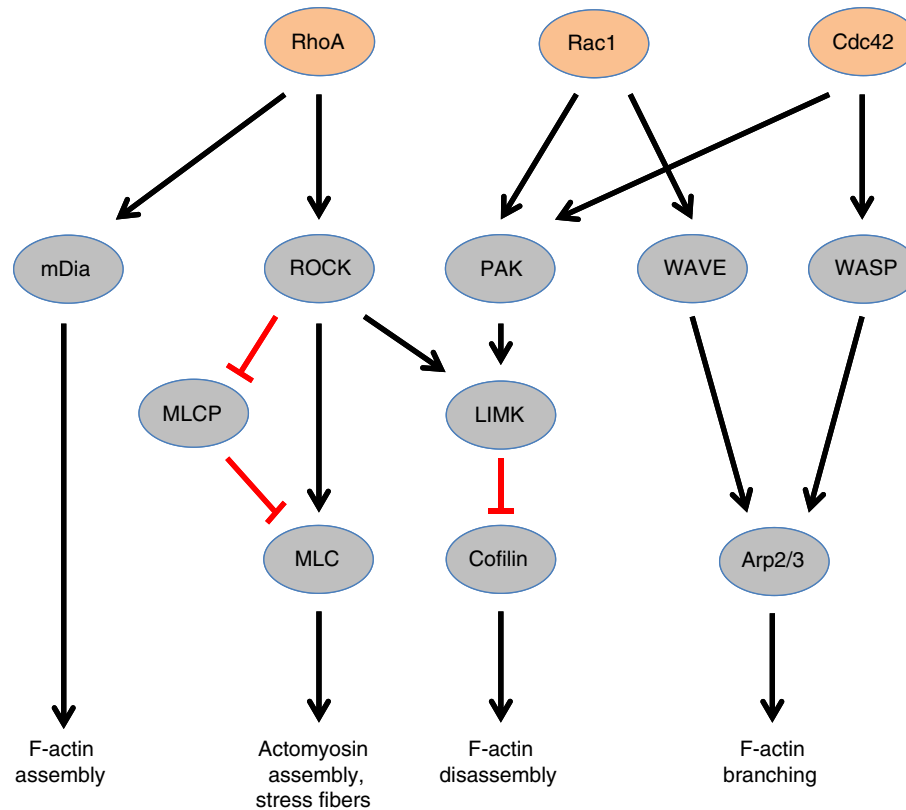


Figure 4 Regulation of the actin cytoskeleton by the Rho GTPases. Rho, Rac, and Cdc42 signal through a range of kinase and non-kinase effectors to regulate aspects of the cytoskeleton. Downstream of RhoA, the diaphanous-related formin mDia catalyzes actin polymerization, and Rho-associated coiled-coil-containing protein kinase (ROCK) promotes F-actin stabilization and actomyosin assembly to generate stress fibers. Downstream of Cdc42 and Rac, the Wiskott–Aldrich Syndrome protein (WASP) family and WASP family verprolin homologous (WAVE) proteins, respectively, promote F-actin branching by activating the actin-related protein 2/3 (Arp2/3) complex, and p21-activated kinase (PAK) promotes F-actin stabilization by activating LIM kinase (LIMK), which in turn inhibits cofilin.

of activated Cdc42 to the Cdc42-Rac-interactive binding (CRIB) motif within the regulatory domain relieves this inhibitory interaction, exposing the ARP2/3 binding and activation site of WASP (Hoffman and Cerione, 2000; Bompard and Caron, 2004). Rac signals to ARP2/3 through the WASP family verprolin homologous (WAVE) proteins, which lack a GTPase-binding domain and consequently do not bind Rac directly. The three WAVE proteins (1, 2, and 3) each exist in a pentameric heterocomplex consisting of WAVE, ABI, NAP1, SRA1, and HSPC300 (Takenawa and Suetsugu, 2007). The activity of WAVE toward ARP2/3 is inhibited by intra-complex sequestration of the VCA module, and binding of activated Rac1 to Sra1 causes an allosteric release of the VCA module to activate the complex (Chen *et al.*, 2010).

A study of Rho GTPase coordination in fibroblasts found that RhoA is activated at the advancing leading edge, and Rac1 and Cdc42 are activated slightly behind the leading edge after a short delay (Machacek *et al.*, 2009). This suggests that RhoA has a primary role in leading edge protrusion, while Cdc42 and Rac1 reinforce and stabilize newly formed protrusions. Whereas Cdc42 and Rac1 stimulate F-actin branching through the ARP2/3 complex, RhoA signals through a family of proteins called formins to induce formation of straight actin

filaments. Formins are characterized by the presence of one or more formin homology (FH) domains. The diaphanous-related formin mDia1 is a direct effector of RhoA, and contains an N-terminal Rho-binding domain (RBD) within a larger autoinhibitory domain. Binding of activated RhoA to the RBD relieves an inhibitory intramolecular interaction, exposing an FH2 domain that can then bind to the barbed end of an actin filament and nucleate actin polymerization (Lammers *et al.*, 2005). mDia1 also contains a proline-rich FH1 domain, which binds monomeric profilin-bound actin and delivers it to the filament end. As the actin filament grows, mDia1 remains bound to the barbed end, recruiting further actin monomers and preventing binding of capping proteins (Zigmond *et al.*, 2003).

Rho, Rac, and Cdc42 also signal to the actin cytoskeleton through kinases such as the Rho-associated coiled-coil-containing protein kinase (ROCK) and the p21-activated kinase (PAK). When bound to activated Rho, ROCK increases myosin light chain (MLC) phosphorylation by directly phosphorylating MLC and also by phosphorylating, and inactivating, myosin phosphatase (Riento and Ridley, 2003). The result of increased MLC phosphorylation downstream of ROCK is enhanced binding of myosin to F-actin. Because myosin is

effective at cross-linking actin filaments, this enhanced binding leads to the formation of contractile actomyosin bundles known as stress fibers. In addition to promoting actomyosin assembly, ROCK signals to stabilize F-actin by phosphorylating and activating LIM kinase (LIMK). LIMK in turn phosphorylates and inactivates cofilin, an actin depolymerizing and severing factor (Riento and Ridley, 2003). The Cdc42- and Rac-activated PAK family of kinases can also signal through LIMK to inactivate cofilin (Edwards *et al.*, 1999).

Other Targets and Multiple Roles in Biology

A variety of proteins have been identified as 'downstream' targets for the different Rho family members, demonstrating that these GTPases are involved in diverse aspects of cell biology far beyond the originally identified roles in membrane ruffling and filopodia formation. For example, it was shown that Cdc42 and Rac, by directly binding and activating PAK, triggered the activation of the stress-responsive nuclear MAP kinases, c-Jun N-terminal kinase (JNK) and p38 (Jaffe and Hall, 2005). This, together with the uncovering of a similar signaling pathway in *S. cerevisiae* (Leberer *et al.*, 1997), provided the first evidence that members of the Rho GTPase family can mediate the regulation of gene expression. Similarly, RhoA was shown to signal to the serum response factor and thereby activate transcription (Jaffe and Hall, 2005).

In an elegant set of biochemical studies, the Rac2 protein (highly similar in sequence to Rac1) was shown to play an essential role in the ability of phagocytes to destroy foreign pathogens by stimulating the oxidative burst *via* the activation of the NADPH oxidase complex (Abo *et al.*, 1992; Diebold and Bokoch, 2001). Indeed, this was the first conclusive evidence that a GTP-bound Rho protein could directly engage and stimulate an effector activity. However, there probably has not been a more clear cut demonstration of the signaling flexibility of the Rho GTPases than the discovery of how Cdc42, through its engagement of a diversity of target proteins (Figure 5), impacts a wide variety of cellular functions. One rather surprising discovery was the apparent roles played by Cdc42 in intracellular trafficking events. The initial suggestions for such

a role came from immunofluorescence studies showing that a predominant cellular localization of Cdc42 was in Golgi membranes (Erickson *et al.*, 1996). This was further reinforced by experiments in which the entire COP1 coatamer complex could be affinity precipitated with GTP γ S-bound GST-Cdc42 (Wu *et al.*, 2000). Several other studies would then implicate Cdc42 in intracellular trafficking events (Harris and Tepass, 2010). The interaction of Cdc42 with another kinase target, ACK (for activated Cdc42-associated tyrosine kinase), was suggested to be involved in the down-regulation of the EGF receptor (EGFR) (Lin *et al.*, 2006). Moreover, the role of Cdc42 in extending EGFR lifetime may have provided the most direct evidence for the involvement of this GTPase in cell growth control. A hyperactivated Cdc42 mutant (see below) was shown to influence EGFR interactions with the E3 ubiquitin ligase c-Cbl, which catalyzes receptor ubiquitination and degradation (Feng *et al.*, 2006; Wu *et al.*, 2003). This was the result of the interaction of Cdc42 with the Cool-1 (for Cloned-out-of library)/ β -Pix (for Pak interactive exchange factor) protein, which also binds to c-Cbl. When a ternary complex was formed between Cdc42, Cool-1, and Cbl, the latter was unable to bind to EGFR and catalyze its degradation and down-regulation (Figure 6).

Many of the processes regulated by Rho family GTPases – cell cycle progression, transcriptional activation, cell survival, cell migration and invasion, vesicle trafficking – are important during oncogenesis. However, in contrast to the Ras GTPases, which are constitutively activated by point mutations in ~30% of human tumors (Fernández-Medarde and Santos, 2011), mutations in Rho GTPases are rare. Instead, many tumors show increased expression of Rho proteins, along with altered expression of the GEFs, GAPs, and GDIs that control their activity. With 20 human Rho members regulated by more than 80 GEFs, ~70 GAPs, and 3 GDIs, and with >70 downstream effectors, establishing the contributions of Rho signaling to cancer development has been a challenging research undertaking. However, early *in vitro* studies of cellular transformation have now been complemented with several mouse models, and a picture of Rho GTPase contribution to tumor progression *in vivo* is beginning to emerge (Figure 7).

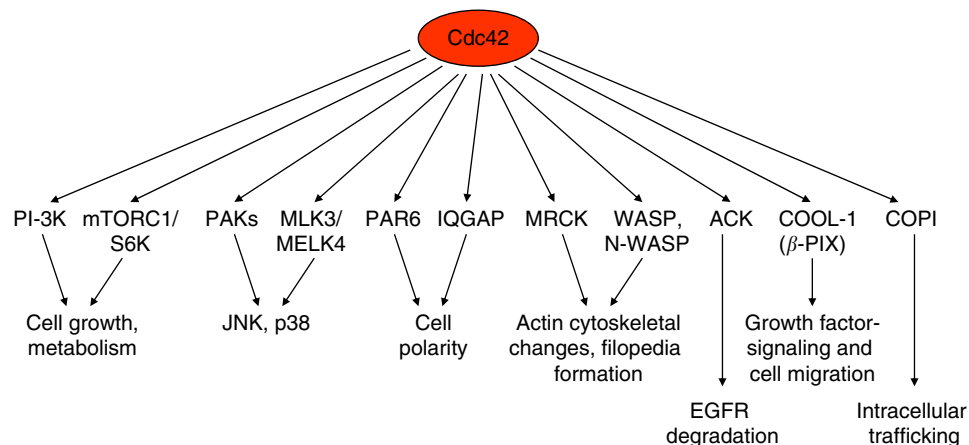


Figure 5 Schematic of the mammalian targets of Cdc42.

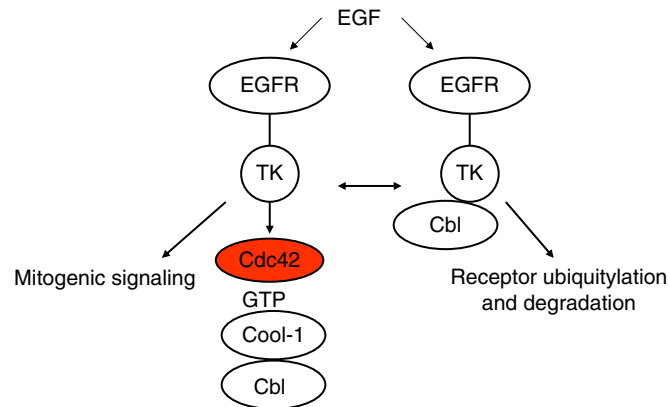


Figure 6 Model of the Cdc42-mediated regulation of EGF receptor interactions with c-Cbl. EGF receptor activation stimulates GDP-GTP exchange on both Ras and Cdc42, resulting in the formation of Cdc42•GTP-Cool-1-Cbl complexes and signaling through the Ras-ERK pathway. When the Cdc42-Cool-1-Cbl complex disassembles, Cbl can then bind to the EGF receptor and is phosphorylated, thus enabling Cbl to function as an ubiquitin E3 ligase that catalyzes receptor ubiquitination and degradation.

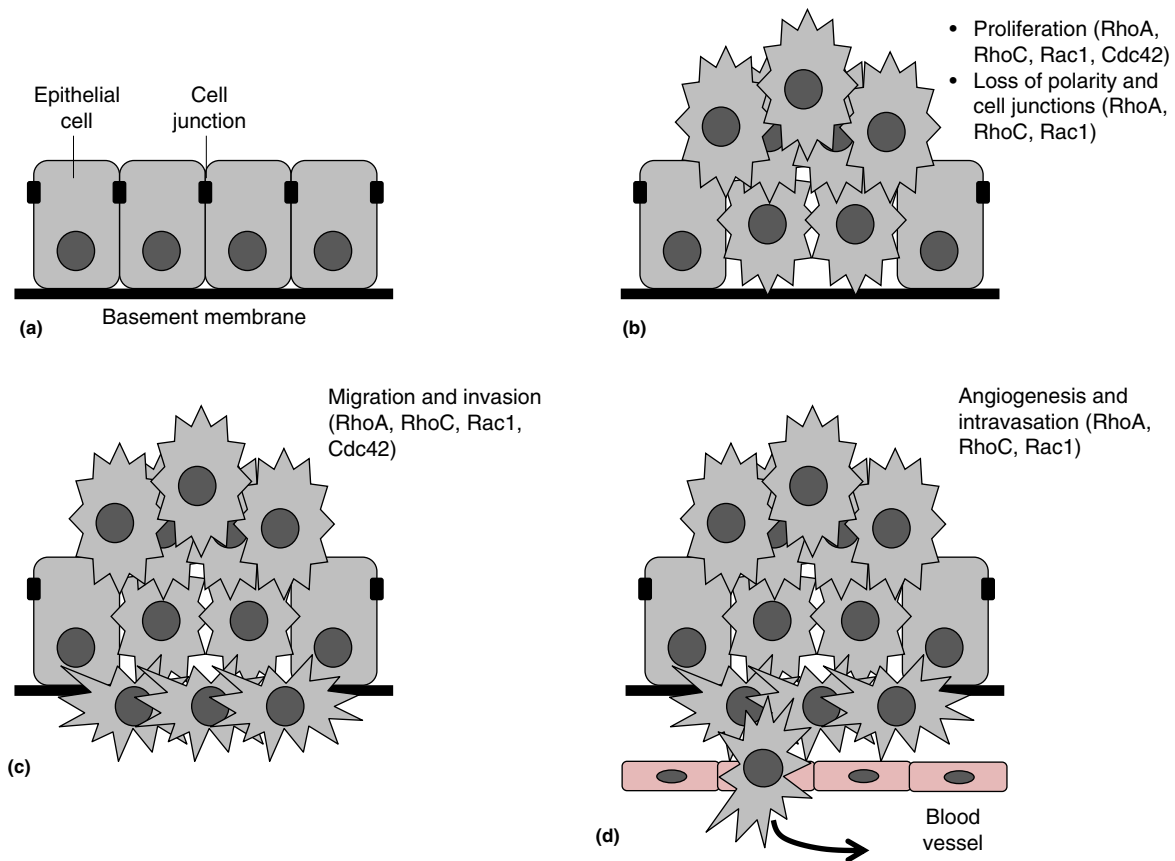


Figure 7 Contributions of Rho GTPases to different stages of tumorigenesis. RhoA, RhoC, Rac1, and Cdc42 can promote cell proliferation and, through the upregulation of NF κ B, protection from apoptosis. During the development of benign tumors, RhoA, RhoC, and Rac1 can also contribute to the loss of cell polarity and cell junctions. Increased cellular motility and invasion through the basement is driven by RhoA, RhoC, Rac1, and Cdc42. RhoC signals to increase expression of angiogenic factors, and RhoA, RhoC, and Rac1 play important roles in intravasation and metastasis.

Rho GTPases in Oncogenic Transformation

Members of the Rho family of GTPases were initially identified on the basis of their homology with the *Ras* oncogenes, raising

the question of whether the Rho proteins might also play important roles in oncogenesis. In fact, several of the GEFs that activate Rho GTPase signaling were found in screens for proteins that cause cellular transformation. The founding member

of the largest family of Rho GEFs, Dbl, was identified in 1985 as an oncogene in a screen of genomic DNA from a human diffuse B-cell lymphoma (Eva and Aaronson, 1985; Hart *et al.*, 1991). Similar screens for transforming proteins identified other GEFs of the Dbl-family, including Vav, Ect2, and Tiam1 (Cook *et al.*, 2013). Subsequently it was found that these Dbl-family members all contained an activating N-terminal truncation that had arisen as a cloning artifact (Cook *et al.*, 2013). Nevertheless, these early results provided strong evidence that aberrant Rho GTPase signaling could contribute to cancer development.

The first studies to probe the oncogenic potential of individual Rho GTPases made use of constitutively active or dominant-negative variants of these proteins. Substitutions analogous to those found in oncogenic Ras (e.g., G12V and Q61L) abolish the intrinsic GTP hydrolysis activity of Rho GTPases, resulting in constitutively GTP-bound, active proteins (Khosravi-Far *et al.*, 1995). A distinct type of activating mutation is based on the Ras F28L variant, which shows reduced binding affinity for guanine nucleotides and thus is referred to as 'fast-cycling' (Reinstein *et al.*, 1991). Because intracellular GTP is considerably more abundant than GDP, this substitution results in the GTPase being mostly GTP-bound but still able to function as a molecular switch. The final class of mutation is based on the S17N variant of Ras, which has a dominant-negative phenotype. This substitution promotes a nucleotide-free state of the GTPase, which binds GEFs with higher affinity than the native protein and consequently inhibits GEF function (Feig, 1999).

A series of studies in the mid-1990s demonstrated that Rho, Rac, and Cdc42 are all essential for Ras-mediated transformation of fibroblasts (Sahai and Marshall, 2002). Expression of dominant-negative variants of any one of these Rho proteins is sufficient to block Ras-driven oncogenesis. Rho GTPases are also required for the transforming activity of numerous other oncoproteins, including receptor tyrosine kinases, Dbl-family members, and G protein-coupled receptors (Boerner *et al.*, 2001; Sachdev *et al.*, 2001; Whitehead *et al.*, 2001). In addition to Rho signaling being required for transformation by multiple oncogenes, ectopic expression of constitutively active or fast-cycling variants of some Rho GTPases, including RhoA, Rac1, and Cdc42, is sufficient to transform fibroblasts (Vega and Ridley, 2008).

Rho GTPases in Mammalian Tumors

Although constitutively activated Rho GTPases can promote cellular transformation *in vitro*, these mutations have generally not been found in human tumors. Similarly, the N-terminal truncation of Dbl-family GEFs, which is potently transforming *in vitro*, has not been found to occur in human cancer. However, aberrant Rho GTPase function is common in cancer, and is typically driven by altered expression of Rho proteins, and their regulators (Vega and Ridley, 2008). Abnormal expression levels of several Rho GTPases, including RhoA, RhoB, RhoC, Rac1, Rac2, Rac3, and Cdc42, have been detected in human tumors (Karlsson *et al.*, 2009), and overexpression of Rho GEFs, including Ect2, P-Rex, Tiam1, and Vav1 is also common (Cook *et al.*, 2013). Some RhoGAPs appear to function as

tumor suppressors, and genomic deletion of the gene encoding DLC1 occurs in hepatocellular carcinoma and many other malignancies. Altered expression of downstream effectors also occurs. For example, the Rac1/Cdc42 effector PAK is upregulated in some breast cancers, and the Rho effector ROCK is upregulated in advanced testicular cancers (Ellenbroek and Collard, 2007).

RhoA overexpression is associated with progression of many cancers, and has been implicated in all stages of cancer development from cellular transformation to metastasis. It promotes G1-S cell cycle progression by suppressing expression of the p21 and p27 cell cycle inhibitors along with multiple INK4 family tumor suppressors (Zhang *et al.*, 2009), and also plays important roles in amoeboid cell migration through the effector ROCK (Vega and Ridley, 2008). Ectopic overexpression of RhoA induces preneoplastic transformation and immortalization of primary mammary epithelial cells (Zhao *et al.*, 2009), and forced overexpression in an ovarian carcinoma cell line significantly increases invasiveness *in vitro* and peritoneal dissemination of tumor cells in nude mice (Horiuchi *et al.*, 2008). Reciprocally, knockdown of RhoA expression inhibits cell cycle progression and proliferation of gastric cancer cells and breast cancer cell lines (Wang *et al.*, 2010; Zhang *et al.*, 2009). Like RhoA, RhoC can induce transformation of some cell types (Xie *et al.*, 2013), but its most important function in cancer development is probably as a driver of cellular invasion and metastasis, and it can additionally stimulate the production of angiogenic factors (Merajver and Usmani, 2005). *Rhoc* knockout mice are viable with no overt phenotype, and are born at a Mendelian ratio (Hakem *et al.*, 2005). The contribution of RhoC to oncogenesis was investigated by crossing the mice with Polyomavirus Middle T mice, which develop palpable mammary tumors at ~50 days of age and lung metastases at ~70 days of age. Onset, frequency, and size of the primary lesions were unaffected by RhoC loss, but the RhoC-null mice showed a dramatic decrease in the size and number of lung metastases at 110–140 days of age (Hakem *et al.*, 2005).

In contrast to RhoA and RhoC, several studies have indicated that RhoB functions as a tumor suppressor, and its expression is downregulated in certain cancers (Karlsson *et al.*, 2009). RhoB-null mice are viable and born at a Mendelian ratio (Liu *et al.*, 2001). In a classical carcinogenesis test, in which skin tumors are initiated by DMBA/TPA application, RhoB-null mice developed more tumors than heterozygotes. When RhoA, RhoB, or RhoC were ectopically overexpressed in gastric cancer lines, RhoA promoted proliferation, RhoC promoted migration and invasion, whereas RhoB inhibited proliferation, migration, and invasion (Zhou *et al.*, 2011). Given that RhoA, RhoB, and RhoC share many effectors, it is likely that their different contributions to oncogenesis are based at least in part on their intracellular localization. RhoA and RhoC are localized to the plasma membrane, whereas RhoB is found mainly at endosomal membranes, and it is possible that it functions as a tumor suppressor by regulating delivery of signaling proteins including growth factor receptors to specific intracellular compartments (Vega and Ridley, 2008).

Rac1 promotes cell cycle progression by stimulating expression of cyclin D1, and transmits an anti-apoptotic signal by activating NF κ B (Gastonguay *et al.*, 2012). The contribution

of Rac1 to cell invasion and migration varies with physiological setting: active Rac1 can inhibit invasion by strengthening epithelial cell–cell contacts, but can also promote a more migratory phenotype by mediating the loss of adherens junctions (Vega and Ridley, 2008). Rac1 knockout in mice is embryonic lethal (Sugihara *et al.*, 1998), and therefore conditional knockout models have been generated. When oncogenic K-Ras was expressed in lung epithelial cells, mice with a conditional Rac1 deletion showed reduced tumor burden and prolonged survival (Kissil *et al.*, 2007). Rac1 is unusual among the Rho GTPases, in that it is frequently mutated in human cancer (Hodis *et al.*, 2012; Krauthammer *et al.*, 2012). The activating P29S substitution is the third most common alteration in melanoma, after mutations in *BRAF* and *NRAS*, strongly implicating *RAC1* as an oncogene in this malignancy. Expression of the Rac1 GEF Tiam1 is also upregulated in melanoma and other cancers (Cook *et al.*, 2013), and Tiam1 deficiency inhibited tumor formation and growth in mice after DMBA/TPA application (Malliri *et al.*, 2002).

The role played by Cdc42 in oncogenesis is complex and likely varies with physiological context. Cdc42 promotes cell survival, cell cycle progression and mitosis, and like Rac1 can regulate NF κ B and other transcription factors including JUN, STAT3 and SRF (Olson *et al.*, 1995; Stengel and Zheng, 2011). Hyperactive Cdc42 is able to induce cellular transformation *in vitro*, and Cdc42 is overexpressed in several cancers (Stengel and Zheng, 2011). However, mouse models are yet to demonstrate a tumor-promoting role for Cdc42. In a hepatocyte-specific deletion model, tumor nodules develop in the liver of mice by 6 months of age (Van Hengel *et al.*, 2008). In all mice, these subsequently develop into hepatocellular carcinoma with metastases to the lung. It is likely that tumorigenesis in this case is a consequence of chronic liver damage, rather than being directly driven by Cdc42 loss; indeed, Cdc42 is upregulated in human hepatocellular carcinoma (Stengel and Zheng, 2011). The dual function of Cdc42 in establishing and maintaining cell polarity, but also in transmitting mitogenic signals, likely underlies the conflicting evidence for its roles in cancer development and progression.

Rho GTPases and Cancer Cell Metabolism

Rho GTPases have recently been implicated in the reprogramming of cellular metabolism that occurs during malignant transformation. A key metabolic adaptation exhibited by many cancer cells involves the pathways that utilize glutamine (Lukey *et al.*, 2013). During cellular proliferation, the TCA cycle intermediate citrate is exported from mitochondria in order to generate cytosolic acetyl-CoA for lipid biosynthesis. Under these conditions, many cells upregulate a mitochondrial metabolic pathway to replenish the carbon that has been lost from the TCA cycle (anaplerosis). Glutamine is converted by the enzyme glutaminase (GLS) to glutamate, which in turn is deaminated by glutamate dehydrogenase (GLUD1) to yield the TCA cycle intermediate α -ketoglutarate. The discovery that many cancer cells are addicted to this metabolic pathway has raised the possibility of targeting these enzymatic activities as a therapeutic strategy against cancer (Lukey *et al.*, 2013). Rho GTPases are able to support TCA cycle anaplerosis by

upregulating GLS activity in the mitochondrion (Wang *et al.*, 2010). Transformation of fibroblasts by oncogenic Dbl, or by fast-cycling variants of RhoC, Rac1, or Cdc42, results in a marked increase in mitochondrial GLS activity. Human breast cancer lines with deregulated Rho GTPase signaling also show elevated GLS activity, and knockdown or inhibition of GLS in these cells strongly inhibits proliferation. The pathway through which Rho GTPases signal to upregulate glutamine metabolism has not yet been fully elucidated, although the NF- κ B member p65 has been implicated as a critical component (Wang *et al.*, 2010).

Rho GTPases and Cancer Cell Microvesicles

Cancer cells can influence their environment through a variety of mechanisms, including the secretion of soluble factors and the release of two classes of extracellular vesicles, known as exosomes and microvesicles (Raposo and Stoorvogel, 2013). Exosomes are 40–100 nm vesicles that are released from a cell when a multivesicular endosome fuses with the plasma membrane, whereas microvesicles are typically larger (up to 3 μ m in diameter) and are shed directly from the plasma membrane. Both exosomes and microvesicles carry a diverse cargo from the parent cell, and when these structures fuse with a recipient cell the contents are released. The cargo of microvesicles includes cell surface receptors, downstream signaling proteins, and metabolic enzymes, and these components can promote transformed behavior in recipient cells (Raposo and Stoorvogel, 2013). Rho GTPase signaling drives the production of microvesicles, particularly downstream of RhoA and ROCK (Li *et al.*, 2012).

Concluding Remarks

The first established roles for the Rho GTPases involved regulation of the actin cytoskeleton. Over the last two decades, it has become clear that Rho family members signal to influence diverse processes, including cell cycle progression and gene expression. It has long been appreciated that many of these signaling pathways would have important consequences in cancer development and metastasis. This has been borne out by the construction of animal models, and by the discovery of recurring Rac1 mutations in human malignancies. Recent discoveries implicate Rho proteins in unexpected areas, including the regulation of glutamine metabolism and the generation of microvesicles. All of this points to the possibility of developing novel therapeutic strategies to interfere with these Rho GTPase-driven processes that promote and support the malignant state.

Key Points

- The 20 mammalian Rho GTPases act as molecular switches that regulate many important processes in cell biology.
- Early studies implicated the Rho GTPases as regulators of the actin cytoskeleton, through effectors such as mDia, WASP, WAVE, ROCK, and PAK.

- By regulating cytoskeletal dynamics, the Rho GTPases influence cell shape, cell migration, vesicle trafficking, and cytokinesis.
- Rho GTPases also signal to regulate gene expression, cell cycle progression, cell survival, cellular metabolism, and extracellular vesicle production.
- Rho GTPases are frequently deregulated in cancer. Many are overexpressed or hyperactivated, while some (e.g. RhoB) are typically downregulated. Rac1 has recently been implicated as an oncoprotein in melanoma.

See also: Cell Communication: Cellular Outcomes: Regulation of Cell Migration. Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell Migration

References

- Abo, A., Boyhan, A., West, I., Thrasher, A.J., Segal, A.W., 1992. Reconstitution of neutrophil NADPH oxidase activity in the cell-free system by four components: p67-phox, p47-phox, p21rac1, and cytochrome b-245. *Journal of Biological Chemistry* 267, 16767–16770.
- Boerner, J.L., Danielsen, A., McManus, M.J., Maihle, N.J., 2001. Activation of Rho is required for ligand-independent oncogenic signaling by a mutant epidermal growth factor receptor. *Journal of Biological Chemistry* 276, 3691–3695.
- Bompard, G., Caron, E., 2004. Regulation of WASP/WAVE proteins: Making a long story short. *Journal of Cell Biology* 166, 957–962.
- Boureux, A., Vignal, E., Faure, S., Fort, P., 2007. Evolution of the Rho family of ras-like GTPases in eukaryotes. *Molecular Biology and Evolution* 24, 203–216.
- Cerione, R.A., Zheng, Y., 1996. The Dbl family of oncogenes. *Current Opinion in Cell Biology* 8, 216–222.
- Chen, Z., Borek, D., Padrick, S.B., *et al.*, 2010. Structure and control of the actin regulatory WAVE complex. *Nature* 468, 533–538.
- Cook, D.R., Rossman, K.L., Der, C.J., 2013. Rho guanine nucleotide exchange factors: Regulators of Rho GTPase activity in development and disease. *Oncogene* 1–15. doi:10.1038/ncr.2013.362.
- Cote, J.F., Vuori, K., 2002. Identification of an evolutionarily conserved superfamily of DOCK180-related proteins with guanine nucleotide exchange activity. *Journal of Cell Science* 115, 4901–4913.
- Diebold, B.A., Bokoch, G.M., 2001. Molecular basis for Rac2 regulation of phagocyte NADPH oxidase. *Nature Immunology* 2, 211–215.
- Edwards, D.C., Sanders, L.C., Bokoch, G.M., Gill, G.N., 1999. Activation of LIM-kinase by Pak1 couples Rac/Cdc42 GTPase signalling to actin cytoskeletal dynamics. *Nature Cell Biology* 1, 253–259.
- Ellenbroek, S.I.J., Collard, J.G., 2007. Rho GTPases: Functions and associations with cancer. *Clinical and Experimental Metastasis* 24, 657–672.
- Erickson, J.W., Zhang, C.J., Kahn, R.A., Evans, T., Cerione, R.A., 1996. Mammalian Cdc42 is a brefeldin A-sensitive component of the Golgi apparatus. *Journal of Biological Chemistry* 271, 26850–26854.
- Eva, A., Aaronson, S.A., 1985. Isolation of a new human oncogene from a diffuse B-cell lymphoma. *Nature* 316, 273–275.
- Feig, L.A., 1999. Tools of the trade: Use of dominant-inhibitory mutants of Ras-family GTPases. *Nature Cell Biology* 1, E25–E27.
- Feng, Q., Baird, D., Peng, X., *et al.*, 2006. Cool-1 functions as an essential regulatory node for EGF receptor- and Src-mediated cell growth. *Nature Cell Biology* 8, 945–956.
- Fernández-Medarde, A., Santos, E., 2011. Ras in cancer and developmental diseases. *Genes & Cancer* 2, 344–358.
- Fukumoto, Y., Kaibuchi, K., Hori, Y., *et al.*, 1990. Molecular cloning and characterization of a novel type of regulatory protein (GDI) for the rho proteins, ras p21-like small GTP-binding proteins. *Oncogene* 5, 1321–1328.
- Gastonguay, A., Berg, T., Hauser, A.D., *et al.*, 2012. The role of Rac1 in the regulation of NF- κ B activity, cell proliferation, and cell migration in non-small cell lung carcinoma. *Cancer Biology & Therapy* 13, 647–656.
- Hakem, A., Sanchez-Sweetman, O., You-Ten, A., *et al.*, 2005. RhoC is dispensable for embryogenesis and tumor initiation but essential for metastasis. *Genes & Development* 19, 1974–1979.
- Harris, K.P., Tepass, U., 2010. Cdc42 and vesicle trafficking in polarized cells. *Traffic* 11, 1272–1279.
- Hart, M.J., Eva, A., Evans, T., Aaronson, S.A., Cerione, R.A., 1991. Catalysis of guanine nucleotide exchange on the CDC42Hs protein by the dbl oncogene product. *Nature* 354, 311–314.
- Hart, M.J., Eva, A., Zangrilli, D., *et al.*, 1994. Cellular transformation and guanine nucleotide exchange activity are catalyzed by a common domain on the dbl oncogene product. *Journal of Biological Chemistry* 269, 62–65.
- Hart, M.J., Polakis, P.G., Evans, T., Cerione, R.A., 1990. The identification and characterization of an epidermal growth factor-stimulated phosphorylation of a specific low molecular weight GTP-binding protein in a reconstituted phospholipid vesicle system. *Journal of Biological Chemistry* 265, 5990–6001.
- Hodis, E., Watson, I.R., Kryukov, G.V., *et al.*, 2012. A landscape of driver mutations in melanoma. *Cell* 150, 251–263.
- Hoffman, G.R., Cerione, R.A., 2000. Flipping the switch: The structural basis for signaling through the CRIB motif. *Cell* 102, 403–406.
- Hoffman, G.R., Nassar, N., Cerione, R.A., 2000. Structure of the Rho family GTP-binding protein Cdc42 in complex with the multifunctional regulator RhoGDI. *Cell* 100, 345–356.
- Horiuchi, A., Kikuchi, N., Osada, R., *et al.*, 2008. Overexpression of RhoA enhances peritoneal dissemination: RhoA suppression with Lovastatin may be useful for ovarian cancer. *Cancer Science* 99, 2532–2539.
- Jaffe, A.B., Hall, A., 2005. Rho GTPases: Biochemistry and biology. *Annual Review of Cell and Developmental Biology* 21, 247–269.
- Johnson, D.I., Pringle, J.R., 1990. Molecular characterization of CDC42, a *Saccharomyces cerevisiae* gene involved in the development of cell polarity. *Journal of Cell Biology* 111, 143–152.
- Johnson, J.L., Erickson, J.W., Cerione, R.A., 2009. New insights into how the Rho guanine nucleotide dissociation inhibitor regulates the interaction of Cdc42 with membranes. *Journal of Biological Chemistry* 284, 23860–23871.
- Kang, R., Wan, J., Arstikaitis, P., *et al.*, 2008. Neural palmitoyl-proteomics reveals dynamic synaptic palmitoylation. *Nature* 456, 904–909.
- Karlsson, R., Pedersen, E.D., Wang, Z., Brakebusch, C., 2009. Rho GTPase function in tumorigenesis. *Biochimica et Biophysica Acta — Reviews on Cancer* 1796, 91–98.
- Khosravi-Far, R., Solski, P.A., Clark, G.J., Kinch, M.S., Der, C.J., 1995. Activation of Rac1, RhoA, and mitogen-activated protein kinases is required for Ras transformation. *Molecular and Cellular Biology* 15, 6443–6453.
- Kissil, J.L., Walmsley, M.J., Hanlon, L., *et al.*, 2007. Requirement for Rac1 in a K-ras induced lung cancer in the mouse. *Cancer Research* 67, 8089–8094.
- Krauthammer, M., Kong, Y., Ha, B.H., *et al.*, 2012. Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. *Nature Genetics* 44, 1006–1014.
- Lammers, M., Rose, R., Scrima, A., Wittinghofer, A., 2005. The regulation of mDia1 by autoinhibition and its release by Rho.GTP. *EMBO Journal* 24, 4176–4187.
- Lancaster, C.A., Taylor-Harris, P.M., Self, A.J., *et al.*, 1994. Characterization of rhoGAP. A GTPase-activating protein for rho-related small GTPases. *Journal of Biological Chemistry* 269, 1137–1142.
- Leberer, E., Thomas, D.Y., Whiteway, M., 1997. Pheromone signalling and polarized morphogenesis in yeast. *Current Opinion in Genetics and Development* 7, 59–66.
- Leonard, D., Hart, M.J., Platko, J.V., *et al.*, 1992. The identification and characterization of a GDP-dissociation inhibitor (GDI) for the CDC42Hs protein. *Journal of Biological Chemistry* 267, 22860–22868.
- Li, B., Antonyak, M.A., Zhang, J., Cerione, R.A., 2012. RhoA triggers a specific signaling pathway that generates transforming microvesicles in cancer cells. *Oncogene* 31, 4740–4749.
- Lin, Q., Yang, W., Cerione, R.A., 2006. Measurement of epidermal growth factor receptor turnover and effects of Cdc42. *Methods in Enzymology* 406, 614–625.
- Liu, A.X., Rane, N., Liu, J.P., Prendergast, G.C., 2001. RhoB is dispensable for mouse development, but it modifies susceptibility to tumor formation as well as cell adhesion and growth factor signaling in transformed cells. *Molecular and Cellular Biology* 21, 6906–6912.
- Lukey, M.J., Wilson, K.F., Cerione, R.A., 2013. Therapeutic strategies impacting cancer cell glutamine metabolism. *Future Medicinal Chemistry* 5, 1685–1700.
- Machacek, M., Hodgson, L., Welch, C., *et al.*, 2009. Coordination of Rho GTPase activities during cell protrusion. *Nature* 461, 99–103.
- Malliri, A., van der Kammen, R.A., Clark, K., *et al.*, 2002. Mice deficient in the Rac activator Tiam1 are resistant to Ras-induced skin tumours. *Nature* 417, 867–871.
- Meller, N., Irani-Tehrani, M., Klosses, W.B., Del Pozo, M.A., Schwartz, M.A., 2002. Zizimin1, a novel Cdc42 activator, reveals a new GEF domain for Rho proteins. *Nature Cell Biology* 4, 639–647.

- Merajver, S.D., Usmani, S.Z., 2005. Multifaceted role of Rho proteins in angiogenesis. *Journal of Mammary Gland Biology and Neoplasia* 10, 291–298.
- Michaelson, D., Silletti, J., Murphy, G., *et al.*, 2001. Differential localization of Rho GTPases in live cells: Regulation by hypervariable regions and RhoGDI binding. *Journal of Cell Biology* 152, 111–126.
- Munemitsu, S., Innis, M.A., Clark, R., *et al.*, 1990. Molecular cloning and expression of a G25K cDNA, the human homolog of the yeast cell cycle gene CDC42. *Molecular and Cellular Biology* 10, 5977–5982.
- Nassar, N., Hoffman, G.R., Manor, D., Clardy, J.C., Cerione, R.A., 1998. Structures of Cdc42 bound to the active and catalytically compromised forms of Cdc42GAP. *Nature Structural Biology* 5, 1047–1052.
- Nobes, C.D., Hall, A., 1995. Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. *Cell* 81, 53–62.
- Olson, M.F., Ashworth, A., Hall, A., 1995. An essential role for Rho, Rac, and Cdc42 GTPases in cell cycle progression through G1. *Science* 269, 1270–1272.
- Pollard, T.D., 2007. Regulation of actin filament assembly by Arp2/3 complex and formins. *Annual Review of Biophysics and Biomolecular Structure* 36, 451–477.
- Raposo, G., Stoorvogel, W., 2013. Extracellular vesicles: Exosomes, microvesicles, and friends. *Journal of Cell Biology* 200, 373–383.
- Reinstein, J., Schlichting, I., Frech, M., Goody, R.S., Wittinghofer, A., 1991. p21 with a phenylalanine 28→leucine mutation reacts normally with the GTPase activating protein GAP but nevertheless has transforming properties. *Journal of Biological Chemistry* 266, 17700–17706.
- Ridley, A.J., Hall, A., 1992. The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. *Cell* 70, 389–399.
- Ridley, A.J., Paterson, H.F., Johnston, C.L., Dickmann, D., Hall, A., 1992. The small GTP-binding protein rac regulates growth factor-induced membrane ruffling. *Cell* 70, 401–410.
- Riento, K., Ridley, A.J., 2003. ROCKs: Multifunctional kinases in cell behaviour. *Nature Reviews Molecular Cell Biology* 4, 446–456.
- Rossman, K.L., Der, C.J., Sondek, J., 2005. GEF means go: Turning on RHO GTPases with guanine nucleotide-exchange factors. *Nature Reviews Molecular Cell Biology* 6, 167–180.
- Sachdev, P., Jiang, Y.X., Li, W., *et al.*, 2001. Differential requirement for Rho family GTPases in an oncogenic insulin-like growth factor-I receptor-induced cell transformation. *Journal of Biological Chemistry* 276, 26461–26471.
- Sahai, E., Marshall, C.J., 2002. Rho-GTPases and cancer. *Nature Reviews Cancer* 2, 133–142.
- Sasaki, T., Kikuchi, A., Araki, S., *et al.*, 1990. Purification and characterization from bovine brain cytosol of a protein that inhibits the dissociation of GDP from and the subsequent binding of GTP to smg p25A, a ras p21-like GTP-binding protein. *Journal of Biological Chemistry* 265, 2333–2337.
- Shinjo, K., Koland, J.G., Hart, M.J., *et al.*, 1990. Molecular cloning of the gene for the human placental GTP-binding protein Gp (G25K): Identification of this GTP-binding protein as the human homolog of the yeast cell-division-cycle protein CDC42. *Proceedings of the National Academy of Science of the United States of America* 87, 9853–9857.
- Sit, S.T., Manser, E., 2011. Rho Gtpases and their role in organizing the actin cytoskeleton. *Journal of Cell Science* 124, 679–683.
- Stengel, K., Zheng, Y., 2011. Cdc42 in oncogenic transformation, invasion, and tumorigenesis. *Cellular Signalling* 23, 1415–1423.
- Sugihara, K., Nakatsuji, N., Nakamura, K., *et al.*, 1998. Rac1 is required for the formation of three germ layers during gastrulation. *Oncogene* 17, 3427–3433.
- Takai, Y., Kikuchi, A., Yamashita, T., *et al.*, 1988. Small molecular weight GTP-binding protein from bovine brain membranes: Purification, characterization and possible functions. In: Imura, H., Shizume, K., Yoshida, S. (Eds.), *Progress in Endocrinology* 2. Amsterdam: Elsevier Science Publishers, pp. 995–1000.
- Takenawa, T., Suetsugu, S., 2007. The WASP-WAVE protein network: Connecting the membrane to the cytoskeleton. *Nature Reviews Molecular Cell Biology* 8, 37–48.
- Tcherkezian, J., Lamarche-Vane, N., 2007. Current knowledge of the large RhoGAP family of proteins. *Biology of the Cell* 99, 67–86.
- Ueda, T., Kikuchi, A., Ohga, N., Yamamoto, J., Takai, Y., 1990. Purification and characterization from bovine brain cytosol of a novel regulatory protein inhibiting the dissociation of GDP from and the subsequent binding of GTP to rhoB p20, a ras p21-like GTP-binding protein. *Journal of Biological Chemistry* 265, 9373–9380.
- Van Hengel, J., D'Hooge, P., Hooghe, B., *et al.*, 2008. Continuous cell injury promotes hepatic tumorigenesis in Cdc42-deficient mouse liver. *Gastroenterology* 134, 781–792.
- Vega, F.M., Ridley, A.J., 2008. Rho GTPases in cancer cell biology. *FEBS Letters* 582, 2093–2101.
- Wang, J.B., Erickson, J.W., Fuji, R., *et al.*, 2010. Targeting mitochondrial glutaminase activity inhibits oncogenic transformation. *Cancer Cell* 18, 207–219.
- Whitehead, I.P., Zohn, I.E., Der, C.J., 2001. Rho GTPase-dependent transformation by G protein-coupled receptors. *Oncogene* 20, 1547–1555.
- Worthylake, D.K., Rossman, K.L., Sondek, J., 2000. Crystal structure of Rac1 in complex with the guanine nucleotide exchange region of Tiam1. *Nature* 408, 682–688.
- Wu, W.J., Erickson, J.W., Lin, R., Cerione, R.A., 2000. The γ -subunit of the coatomer complex binds Cdc42 to mediate transformation. *Nature* 405, 800–804.
- Wu, W.J., Tu, S., Cerione, R.A., 2003. Activated Cdc42 sequesters c-Cbl and prevents EGF receptor degradation. *Cell* 114, 715–725.
- Xie, S., Zhu, M., Lv, G., *et al.*, 2013. Overexpression of Ras homologous C (RhoC) induces malignant transformation of hepatocytes *in vitro* and in nude mouse xenografts. *PLoS One* 8, e54493.
- Zhang, S., Tang, Q., Xu, F., *et al.*, 2009. RhoA regulates G1-S progression of gastric cancer cells by modulation of multiple INK4 family tumor suppressors. *Molecular Cancer Research* 7, 570–580.
- Zhao, X., Lu, L., Pokhriyal, N., *et al.*, 2009. Overexpression of RhoA induces preneoplastic transformation of primary mammary epithelial cells. *Cancer Research* 69, 483–491.
- Zhou, J., Zhu, Y., Zhang, G., *et al.*, 2011. A distinct role of RhoB in gastric cancer suppression. *International Journal of Cancer* 128, 1057–1068.
- Zigmond, S.H., Evangelista, M., Boone, C., *et al.*, 2003. Formin leaky cap allows elongation in the presence of tight capping proteins. *Current Biology* 13, 1820–1823.

Cell Polarity

F Martin-Belmonte, I Bernascone, and M Galvez-Santisteban, Centro de Biología Molecular Severo Ochoa (CSIC-UAM), Madrid, Spain

© 2016 Elsevier Inc. All rights reserved.

Glossary

Cell fate determinants Asymmetrically localized proteins that are required for cell fate specification. In *Drosophila* proteins such as Numb, Prospero, and Brat are localized at the basal side of neuroblasts ready for asymmetric division.

Epithelial–mesenchymal transition The epithelial–mesenchymal transition (EMT) is an orchestrated process by which polarized epithelial cells are converted to motile mesenchymal cells. During this process cell–cell and cell–extracellular matrix interactions are altered, the cytoskeleton is reorganized, and a new transcriptional program is induced to maintain the mesenchymal phenotype. EMT is essential for normal embryonic development, but it is also associated with metastasis and tumor progression.

Imaginal discs Larval or pupal epithelial precursors of most adult *Drosophila* organs. Imaginal discs are formed by two contiguous epithelia, a columnar layer and a squamous layer; together they form a bag-like folded epithelium. Each imaginal disc forms a separate structure, e.g., there is a

separate imaginal disc that forms wings while another imaginal disc forms legs.

MDCK (Madin-Darby canine kidney) cells MDCK cell line derived from dog kidney. Used as model for studying trafficking and establishment of cell polarity in epithelia.

Mitotic spindle The mitotic spindle is the macromolecular machine that segregates chromosomes to two daughter cells during mitosis. The spindle consists primarily of microtubules, polarized filaments composed of α/β -tubulin heterodimers arranged in a head-to-tail configuration within protofilaments.

Scaffold protein A multivalent protein that is known to interact and/or bind with multiple members of a signaling pathway. They help coordinate the location of proteins to specific areas of the cell and tether them into complexes, thus enhancing the efficiency and/or specificity of cellular signaling pathways.

Transmembrane protein Integral membrane protein that spans from one side of a membrane through to the other side of the membrane.

Basic Principles of Cell Polarity and Common Mechanisms of Control

Cell polarity can be defined as an asymmetry in molecular composition or structure between two sides, thus defining a polarity axis along which cellular processes will be differentially regulated. Evidence of polarity can be detected across evolutionary history, from the simplest unicellular organisms, such as the bacterium *Escherichia coli* and the budding yeast *Saccharomyces cerevisiae*, up to the most complex multicellular organisms, including flies, plants, and animals. Cell polarity is therefore a universal attribute of almost all cells. Whereas in unicellular organisms polarity apparently plays a role in control of lineage senescence by ensuring differential inheritance of damaged material at cell division (Macara and Mili, 2008), in multicellular organisms it is responsible for differential cell fate specification in asymmetric divisions as well as for a wider range of cellular functions. For example, neurons are highly polarized cells with segregated domains specialized for either receiving (dendrites) or transmitting (axons) cellular signals. Mesenchymal cells need to acquire a front-to-back polarity in order to migrate, while epithelial cells display both apical–basal and planar polarity.

A great effort has been made in the last 50 years to investigate cell polarity, first in simple model systems such as yeast, the worm *Caenorhabditis elegans*, and the fruitfly *Drosophila melanogaster*, and later in mammals, by the combination of different approaches, including traditional genetic and biochemistry techniques as well as modern live imaging and computational methods (Slaughter *et al.*, 2009; Sawa, 2012;

Tepass, 2012). Three important pieces of evidence have emerged from these studies. First, irrespective of the organism or cell type, the establishment of cell polarity relies on general and common steps that need to be integrated: (1) a polarity cue, both intrinsic and external; (2) dedicated polarity determinants that are localized to specific domains of the plasma membrane and act to establish polarity; (3) polarized trafficking of proteins and lipids that deliver cargo to specific membrane domains; and (4) transduction of polarity information. Second, even more fascinatingly, studies in recent years have shown that some polarity players have highly conserved roles across animal species. One example involves the Cdc42 small GTPase, which acts as a key molecule in orchestrating the establishment of cell polarity in many well-studied systems, including yeast cells, *Drosophila* neuroblasts, neurons, and epithelial cells. Cdc42 integrates the variety of extra- and intracellular signals received by the cell and, once activated, can potentially interact with several distinct target proteins in order to select a single appropriate polarity axis (Etienne-Manneville, 2004). Another example is that constituted by Par3, Par6, and the atypical protein kinase C aPKC, forming the Par complex (Goehring, 2014). They were first identified in the *C. elegans* zygote by pioneering genetic screens showing that mutants in ‘partitioning defective’ (*par*) genes fail to establish the correct anterior–posterior (AP) cortical polarity axis. Par mutants do not divide asymmetrically but result in two equal-sized anterior and posterior daughter cells (Kemphues *et al.*, 1988). As far as we know, the Par complex is conserved in the *Drosophila* oocyte and epithelia and in mammalian epithelial cells, where it functions to regulate cell

polarity and proliferation, but it does not exist in yeasts or plants. This divergence may be ascribed to the different mechanisms by which a plant or animal cell controls cell surface tension, with the former using cell-wall glucans or cellulose to modify surface tension, and the latter maintaining surface stability through a contractile actomyosin cortex (Li and Bowerman, 2010). A third principle hence arises: regulatory elements, such as the Par module, and cytoskeletal polymers cooperate and influence each other to generate and maintain asymmetry. For instance, in a *C. elegans* one-cell embryo and in mammalian polarized cells polarity molecules control the directional and polarized assembly of actin filaments. On the other hand, in the *Drosophila* oocyte the cytoskeleton is responsible for shaping the distribution and activity of signaling molecules (Goehring and Grill, 2013). The majority of polarity systems can therefore be considered mechanochemical, since cellular asymmetry is controlled by the coupling of mechanical processes and chemical elements, like diffusible species forming a gradient maintained by self-amplifying feedback (Howard *et al.*, 2011).

In the following sections we will dissect each of the general steps of polarity establishment outlined above, giving a molecular description of the key players and discussing their interplay in epithelial cells.

Epithelial Cells

Epithelial cells are the most common cell type in multicellular organisms. They form organized layers, either simple or stratified, that separate the body from the outside world (e.g., epidermis) and subdivide the body into morphologically and physiologically different compartments (e.g., intestinal and lung epithelium). The function of epithelia is to provide a barrier between the inside and outside of the organism, and to allow vectorial transport of fluid or directed secretion of specialized components from one side to the other. To achieve this function, epithelial cells must be highly polarized. They are indeed characterized by an apical plasma membrane domain, facing the external environment, and a basolateral domain, contacting neighboring cells and the extracellular matrix (ECM) with different protein and lipid composition. Epithelial cells differ from other polarized cell types, such as the *C. elegans* embryo and *Drosophila* oocytes having two opposite cortical domains as well, because they form specialized cell junctions with neighboring cells that mark the boundary between the apical and lateral domains. The differential positioning of components (membrane domains and organelles) along the apical-basal axis of an epithelial cell is referred to as apical-basal polarity.

Apical-Basal Polarity of Epithelial Cells

The mechanisms that promote the establishment and maintenance of epithelial cell polarity are not completely understood, and can vary in different epithelial cells in mammals and other model organisms. However, work over the last two decades has pointed to intercellular junctions and apical-basal polarity complexes as important mediators of polarity in simple epithelial cells in both vertebrates and invertebrates.

Mediators of Epithelial Cell Polarity

Intercellular adhesion complexes

Two types of intercellular junctions are established at the level of the lateral surface of epithelial cells: tight junctions (TJs) and adherens junctions (AJs). Although made up of different proteins, both junctional complexes associate with actin, and formation and maturation of cell-cell contacts involves reorganization of the actin cytoskeleton. Both TJs and AJs are composed of transmembrane proteins, responsible for the adhesion of adjacent cells, and of intracellular scaffolding proteins for anchorage to the cytoskeleton. In mammalian epithelia TJs localize apically to AJs at the border of the apical and basolateral domain (Figure 1). In *Drosophila* epithelia the septate junction (SJ), the structure analogous to TJs in invertebrates, sits basal to AJs, while *C. elegans* epithelial cells have a single apical junction that acts as both AJs and TJs. TJs form a belt-like structure and provide both a fence function, to limit the diffusion of proteins and lipids between apical and basolateral membrane domains, and a barrier function, which restricts the passage of molecules between cells. TJs consist of a number of integral proteins, such as claudin, occludin, and junctional adhesion molecules (JAMs), and associated cytoplasmic proteins such as zonula occludens-1 (ZO-1) and its homologs ZO-2 and ZO-3, which in turn interact with F-actin (Tsukita *et al.*, 2009; Figure 1). AJs provide the main mechanical link between neighboring epithelial cells. AJs consist of two basic adhesive units, the cadherin-catenin and nectin-afadin complexes (Niessen and Gottardi, 2008). Classical cadherins mediate strong cell-cell adhesion by a homophilic calcium-dependent binding of their extracellular domain. Intracellularly, they bind through β -catenin to α -catenin, which in turn interacts with actin-binding proteins (actin-BP) (Kobiela and Fuchs, 2004; Figure 1). Nectins link AJs to the actin cytoskeleton through the scaffolding protein afadin (AF6) (Takai and Nakanishi, 2003; Figure 1). In addition, nectins bind to ZO-1 and α -catenin, thus providing a link between AJs and TJs.

Polarity complexes

Two main apical and one basolateral groups of proteins have been found to be involved in the establishment of cell polarity (Figure 1). The first major apical polarity complex is the Par complex, identified in *C. elegans* embryonic lethal mutants and conserved in *Drosophila* oocytes and neuroblasts and mammalian epithelial cells. The Par complex, which establishes the apical-basal membrane border, consists of Par3 (Bazooka, in *Drosophila*) and Par6, which are the scaffolding proteins, and the kinase aPKC, core effector of the complex. Other proteins, including Cdc42 GTPase and Rac1, are considered part of the same group (Lin *et al.*, 2000). The second apical polarity complex is the Crb complex, which controls the extension of the apical surface (Roh *et al.*, 2003). Crb (Crumbs 1–3 in vertebrates) forms a complex with protein associated with Lin Seven 1 (PALS1; Stardust, in *Drosophila*) and PALS1-associated tight junction protein (PATJ). In both *Drosophila* and vertebrate epithelial cells, the Crb complex is enriched at the apical margin of the lateral domain. The Crb complex interacts with the Par complex through the interaction of PALS1 to Par6. In *Drosophila* and mammalian

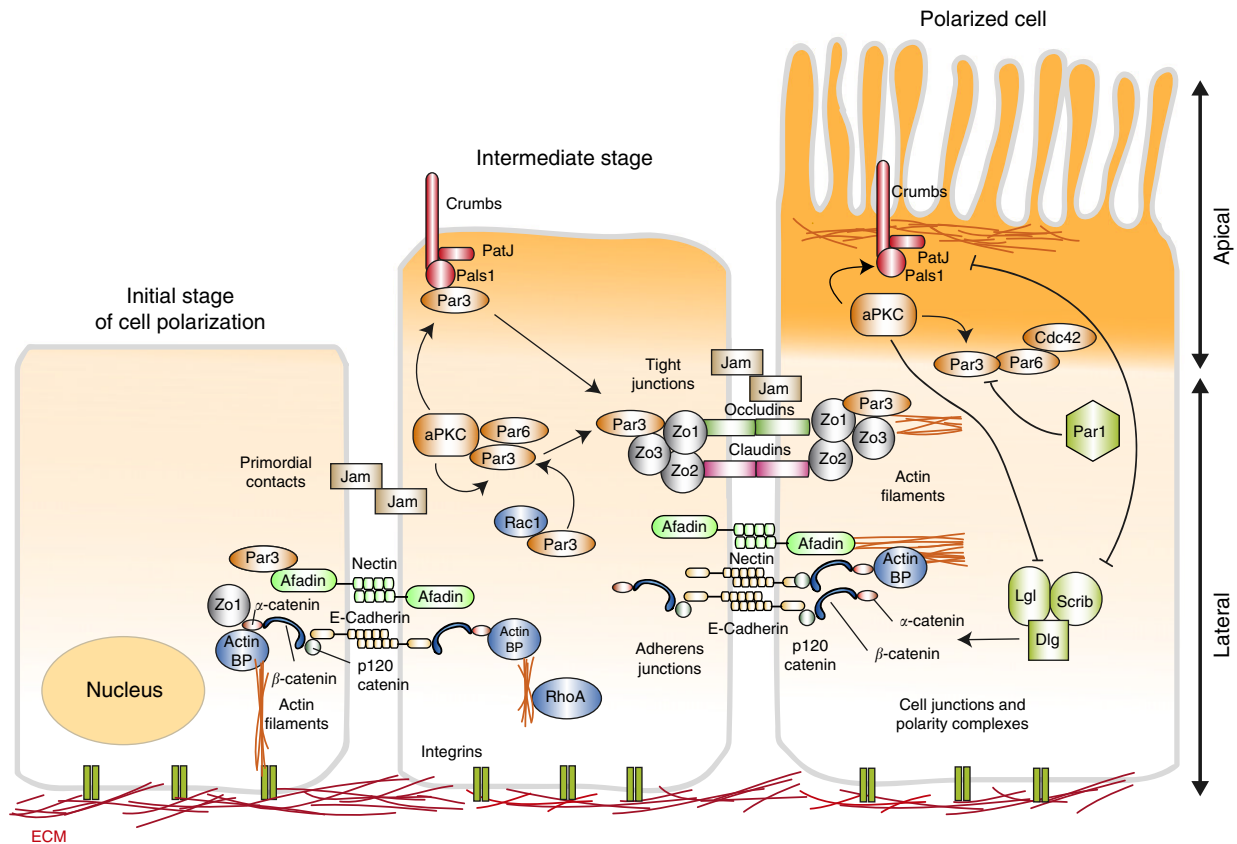


Figure 1 Establishment of cell polarity and junctional complexes in mammalian epithelial cell. From left to right, maturation of primordial junctions into distinct TJs and AJs during epithelial cell polarization. The process requires the exclusion of Par3 through the localized activation of Rac1 GTPase at intermediate stages of polarization. Upon phosphorylation by aPKC, Par3 accumulates at TJs. The Par and Crb complexes establish the apical and the apical/lateral membrane, whereas the Scrib complex defines the basolateral plasma membrane domain. On the right, overview of the antagonistic interactions between the Par complex (orange), the Crumbs complex (red), and the Scribble complex (green).

epithelia, Par6/aPKC colocalize with the Crb complex, while most Par3 localizes at TJs, slightly basally to the Crumbs complex. One exception is the primary epithelium of *Drosophila* (see below), where Bazooka is enriched at the level of the AJs.

The basolateral complex is the Scrib complex, composed of Scribble (Scrib), lethal giant larvae homolog (Lgl), and discs-large homolog (Dlg). The Scrib complex localizes to AJs in mammals (with SJs in *Drosophila*) and defines the basolateral membrane domain (Laprise et al., 2004). At least in *Drosophila*, another group of proteins, the Yurt/Coracle (Yrt/Coracle) group, has been identified as a novel additional player in polarity establishment. The Yrt/Coracle group promotes basolateral membrane stability by negatively interacting with the apical determinant Crumbs (Laprise et al., 2009). It is important to note that the Yrt/Coracle group is essential for epithelial polarity during organogenesis but not in late embryogenesis, when apical and basolateral markers can be normally segregated even in the absence of Yrt/Coracle. This implies the existence of yet unknown polarity proteins that act to define lateral membrane stability in late-stage embryos. Notably, the mammalian Yrt ortholog binds to Crb and contributes to epithelial organization (Laprise et al., 2006).

Initiating Cues of Polarization

Epithelial cells use different cues to establish apical-basal polarity. Most of our knowledge has come from studies on cultured epithelial cells, for example, Madin-Darby canine kidney (MDCK) cells grown in a two-dimensional (2D) system, i.e., as a monolayer, and in three-dimensional (3D) system, where they form spheroid structures termed cysts, consisting of a single layer of cells with an apical surface lining an inner lumen and a basolateral surface in contact with the ECM (Zegers et al., 2003; Martin-Belmonte and Mostov, 2008; Figure 2(d)). In these systems the initiating polarity cue is provided by contact with the basement membrane (BM) and with neighboring cells. BM collagen activates – through $\beta 1$ -integrin – Rac1, which mediates laminin organization and assembly of the nascent BM (Yu et al., 2005). In turn, assembly of laminin activates Rac1, thus generating a positive feedback loop that reinforces the orientation of this signaling pathway. A similar mechanism of matrix-dependent orientation of polarity acts in *Drosophila* follicular epithelium (Denef et al., 2008), and it has been also revealed in mammalian systems *in vivo*, such as in the mammary gland (Naylor et al., 2005). Remarkably, an outstanding article recently published unveils

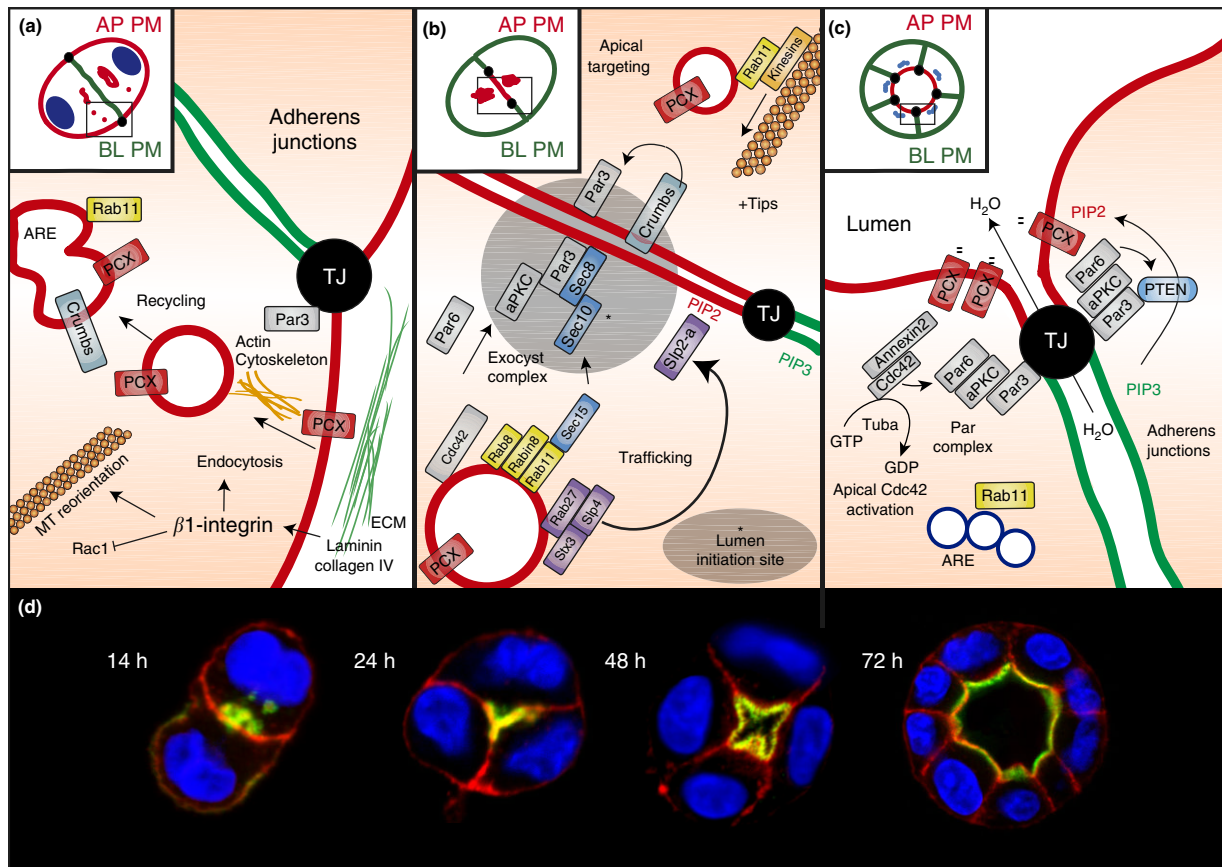


Figure 2 Polarized trafficking and lumen formation in 3D-MDCK cysts. (a–c) Scheme of vectorial trafficking. (a) In unpolarized cells apical membrane components (like PCX) face the ECM, and they need to be internalized and transported to their new destination, in a process of inversion of polarity. β 1-integrin mediates this inversion through its interaction with the ECM, thus triggering MT reorientation, Rac1 inhibition, and induction of endocytosis. Once endocytosed, the apical components are targeted to Rab11-positive AREs. Par3 is present at TJs. (b) Rab11-positive AREs travel to the lumen initiation site (*), probably by association with kinesins (upper right) along the microtubules, which have been previously reoriented. Slp2-a binds to PtdIns(4,5)P₂ enriched at the forming apical membrane and clusters apically destined vesicles in a Rab27a/b-dependent fashion. Concomitantly, Slp4-a bridges Rab27-Rab8 and the apical SNARE syntaxin 3 to promote vesicle tethering and fusion at the lumen initiation site (lower left). The exocyst complex carries out docking of vesicles to the lumen initiation site. Par3 is delivered to the lumen initiation site, a process reinforced by the Crb complex. (c) Once the vesicles reach the lumen initiation site, a tight junction–delimited lumen begins to appear. Cdc42 is activated by aPKC/Par6 in the apical membrane, in collaboration with Tuba. The Par complex recruits PTEN, which maintains the PIP2/PIP3 segregation between the AP and BL membranes. At this stage lumen filling takes place by paracellular transport of ions through the TJs. (d) Time-course analysis of lumen formation in 3D-MDCK cysts. AP and BL membranes are defined by green (Spl2) and red (actin) staining, respectively. Nuclei are depicted in blue (unpublished data). AP, apical; BL, basolateral; ECM, extracellular membrane; PCX, podocalyxin.

the importance of the BM niche surrounding epiblast (EPI) cells in guiding the polarization of these amorphous cells into rosette-like structures such as mouse blastocyst implants (Bedzhov and Zernicka-Goetz, 2014). The BM provides polarization cues that signal through β 1-integrin receptors to orient the establishment of basal–apical axis of the epiblast cells, which later change their shape and organize in a rosette with a small luminal space, precursor of the pre-amniotic cavity. The ECM proteins present in the matrigel provide cues similar to those of the BM in the embryo and are thus sufficient to induce polarization of 3D cultured ES cells, which organize in spheres strongly resembling the EPI rosettes of the peri-implantation embryos (Bedzhov and Zernicka-Goetz, 2014).

In cultured cells, an additional polarization cue comes from neighboring cells. Epithelial cells extend thin protrusions

filled by actin filaments in order to contact neighboring cells and build up cohesive layers of cells. The first cell–cell contact sites, known as ‘puncta,’ are positive for both typical AJ and TJ proteins (Tsukita *et al.*, 2009) and probably serve as a ‘landmark’ or ‘positional cue’ for membrane growth and recruitment of other integral and peripheral membrane proteins. These primordial adhesion sites later mature to form distinct AJs and TJs (Figure 1). Evidence in mammalian embryonic development and MDCK cells cultures suggests that AJs are required for cell polarization. However, it is noteworthy that some cell types can polarize in the absence of intercellular contacts, as is the case of *Drosophila* multinucleated embryonic cells which are forced into contact by cellularization, i.e., they undergo a process of plasma membrane invagination in order to compartmentalize the nuclei and generate single epithelial

cells (Lecuit, 2004). As we will see in the following section, the polarity cue acting in *Drosophila* primary epithelium seems to be provided by the cytoskeleton, and not by extrinsic cell–cell adhesion signals.

Proteins Interplay Organizing Apical–basal Polarity

In the *C. elegans* zygote the anterior (Par3/Par6/aPKC) and posterior (Par1/Par2/Lgl) polarity determinants act in a mutually antagonistic manner to exclude one another from the plasma membrane. Similarly, the generation and maintenance of distinct apical and basolateral domains in epithelial cells depend on mutually antagonistic interactions between the three polarity complexes described above and on their interaction with adhesion proteins. The three polarity complexes interact by a system of mutual phosphorylation-dependent exclusion that ensures their correct position in different membrane domains. The formation of the apical cellular domain relies on the activity of the kinase aPKC. Par3 is initially recruited to the plasma membrane of unpolarized cells through its binding to the nectin–afadin complexes (Ooshio *et al.*, 2007), but it will be later excluded from this region in order to allow the maturation of the primordial adhesion sites (Figure 1). E-cadherin enriched at mature AJs triggers phosphorylation of aPKC, which in turn phosphorylates Par3. Once phosphorylated, Par3 dissociates from aPKC and accumulates at TJs (Nagai-Tamai *et al.*, 2002; Figure 1). In addition, aPKC phosphorylates Crumbs at the apical cortex to control the extension of the apical surface and/or apical junction assembly. The formation of the basolateral domain depends on the activity of the Scribble complex and the kinase Par1. Lgl can physically interact with Par6/aPKC early during cell–cell contact formation, thus blocking the formation of the Par3/Par6/aPKC complex. Upon phosphorylation by aPKC, Lgl dissociates from the apical membrane cortex and distributes to the lateral membrane, where it may interact with Dlg and Scrib (Yamanaka *et al.*, 2003). Furthermore, Par1 phosphorylates Par3 at the basolateral domain, thus dissociating it from Par6/aPKC and excluding the Par6/aPKC complex from this domain (Suzuki *et al.*, 2004; Figure 1).

Apical–basal polarity is generated in quite a different way in the primary epithelium of *Drosophila*, where Bazooka acts upstream of the AJs. As the plasma membrane invaginates to form single cells, Bazooka accumulates at apical/lateral margins of each cell before and independently of cadherin, thus acting as a primary landmark to position the AJs. Bazooka then clusters E-cadherin and the nectin-like protein Echinoid to form the AJs. Bazooka localization depends on the apical actin network and on dynein-dependent transport of Bazooka along apical–basal microtubules (Harris and Peifer, 2005).

Polarized Trafficking Machinery

Membrane trafficking contributes to the generation of cortical epithelial cell polarity. Newly synthesized components of the plasma membrane are sorted in the trans-Golgi network (TGN) and use exocytotic pathways to be delivered to specific domains of the plasma membrane. After surface delivery, apical and basolateral cargoes can be internalized and sorted

into either late endosomes (LE)/lysosomes for degradation, or recycling endosomes (RE) for recycling back to either membrane domain. Endocytosis and recycling from both surface domains are thus essential to maintain the composition of the different membrane domains. Endocytosis ensures the appropriate levels of plasma membrane proteins that control apical–basal polarity, such as Crumbs, which is constantly endocytosed at the basolateral membrane to avoid apical expansion of *Drosophila* follicle epithelium. Recycling allows the proper repositioning of apical and basolateral proteins that require transcytosis to reach their correct membrane domain. As an example, Crumbs is recycled back to the plasma membrane through the action of Rab11 GTPase and contributes to AJs remodeling during embryogenesis. Indeed, disruption of the Rab11 function in the *Drosophila* embryo leads to fragmentation and loss of AJs, and is preceded by loss of Crumbs from the apical domain (Roeth *et al.*, 2009).

Membrane trafficking requires sorting determinants present in the protein sequence, which are decoded by the transport machinery that ensures correct membrane delivery. Apical sorting signals have been localized to all portions of apically destined proteins: the extracellular, transmembrane, and cytoplasmic domains. Apical sorting signals include N-glycans and O-glycans, the glycosyl phosphatidylinositol (GPI) anchor, and specialized transmembrane domains such as that of influenza hemagglutinin. Moreover, some apical proteins associate with lipid rafts, which are formed by cholesterol and glycosphingolipids and act as functional sorting platforms. By contrast, basolateral sorting signals are simple peptide motifs in the cytoplasmic tail of a protein, including tyrosine-based and dileucine-based motifs (Cao *et al.*, 2012). Vesicle traffic is dependent on the cytoskeleton and is driven by microtubule- and actin-based motors. Tethering of vesicles to both apical and basolateral plasma membrane domains has been shown to depend on different subunits of the exocyst, an eight-subunit complex first identified in yeast and conserved in *Drosophila* and mammalian cells (Oztan *et al.*, 2007; Zuo *et al.*, 2009). Finally, vesicle fusion is mediated by interaction between t-SNARE complexes on the target membrane, syntaxins, and v-SNARE complexes on the vesicle membranes.

Interestingly, an increasing body of evidence shows that polarity regulators act as important controllers of vesicle trafficking machinery. Genetic screens in *C. elegans* provided the first link between the Par complex, Cdc42, and regulation of endocytosis (Balklava *et al.*, 2007). In *Drosophila* neuroectoderm, loss of Cdc42 causes accumulation of apical proteins, including Crumbs, in sorting endosomes, probably owing to aPKC activity, a defect that in turn causes the disorganization of the AJs (Harris and Tepass, 2008). However, in other *Drosophila* tissues, such as imaginal discs, Cdc42 seems to play a role in AJs endocytosis through a Par-independent pathway involving actin polymerization. Recently, *par-5* has been identified in the *C. elegans* intestine as a controller of the subapical localization of Rab11-positive REs (Winter *et al.*, 2012).

Transcytosis and lumen formation

An instructive system to investigate the reciprocal interplay between polarity complexes and membrane trafficking is the formation of lumen, a structure found in many epithelial

organs of the body. 3D-MDCK cultures represent a good model to explore the process of *de novo* lumen generation, which depends on a transcytotic pathway (Rodriguez-Fraticelli *et al.*, 2011; Apodaca *et al.*, 2012). In unpolarized MDCK cells aggregates, apical membrane proteins, such as podocalyxin and the Crb complex, are distributed at the periphery. They are first endocytosed and transported in vesicles toward Rab11-positive apical recycling endosomes (ARE), a process regulated by the actin cytoskeleton (Madrid *et al.*, 2010; Figure 2(a)). Then apical membrane proteins travel to the lumen initiation site, probably by association with motor proteins (such as kinesins) that move through polarized microtubules. Their trafficking from ARE to the site of lumen initiation is regulated by the Rab family of GTPases, specifically by Rab11, Rab8, and its specific GEF Rabin8 (Bryant *et al.*, 2010; Figure 2(b)). Finally, the exocyst proteins Sec8, Sec10, and Sec15A facilitate the tethering of vesicles to the forming apical membrane. Synaptotagmin-like proteins 2a (Slp2-a) and 4a (Slp4-a) coordinate the spatiotemporal organization of vectorial apical transport to ensure that only a single apical surface occurs per cell (Galvez-Santisteban *et al.*, 2012). Exocytosis is also important for the initial recruitment of the Par3/aPKC complex to the lumen initiation site. Active Rab8 stimulates Cdc42 interaction with the exocytic carriers and its delivery to the forming apical membrane, where it activates the Par complex (Figure 2(b)). A similar type of transcytotic process has been observed *in vivo* in mouse mammary gland (Akhtar and Streuli, 2013). Furthermore, some aspects of the polarization process of EPI cells organizing into rosette-like structures attractively resemble lumen formation in 3D-MDCK cultures (Bedzhov and Zernicka-Goetz, 2014). EPI cells change their shape as a result of actomyosin constriction, transmitted through AJs, and aPKC and Par6 accumulate in the apical membrane in the center of the sphere. The small luminal space that appears in the center of the rosette is a result of charge repulsion of apical membrane proteins such as podocalyxin (Bedzhov and Zernicka-Goetz, 2014).

Phosphoinositides and polarized trafficking

Phosphoinositides and other lipids are essential for membrane trafficking and, ultimately, for generating epithelial polarity. As MDCK cells polarize in 3D cultures, *PtdIns(3,4,5)P₃* (PIP₃) concentrates at the basolateral membrane, whereas *PtdIns(4,5)P₂* (PIP₂) becomes restricted to the apical membrane (Cassama-Diagne *et al.*, 2006; Martin-Belmonte *et al.*, 2007). The lumen is formed at a previously basolateral cell–cell contact through the accumulation of PIP₂ via the lipid phosphatase PTEN, which converts *PtdIns(3,4,5)P₃* to *PtdIns(4,5)P₂* (Martin-Belmonte *et al.*, 2007), and the activity of the PI5K (Figure 2(c)). PIP₂ leads to the recruitment of active Cdc42 (Etienne-Manneville, 2004; Martin-Belmonte and Mostov, 2007), where it binds and activates the Par6/aPKC complex. PIP₂ is thus proposed to be the main determinant of apical identity. Moreover, PIP₂ directs the vectorial traffic of secretory vesicles to the lumen initiation site through Rab-GTPases and SLP2-a/SLP4-a (Galvez-Santisteban *et al.*, 2012; Yasuda *et al.*, 2012; Figure 2(b)). As in mammalian systems, apical membranes of *Drosophila* tubular epithelia are enriched in PIP₂ and mediate apical targeting through the recruitment of the formin-family protein Diaphanous (Dia) (Roussio *et al.*, 2013).

Cytoskeletal Rearrangements

Acquisition of apical–basal polarity entails an elaborate reorganization of microtubules and actin cytoskeleton. In mammalian epithelial cells, some microtubules are nucleated from microtubule organizing centers (MTOCs) such as the centrosome and are captured by the apical junction complex, thus generating a polarized network of cortical microtubules, with basally oriented plus-ends (Sugioka and Sawa, 2012). Rho-GTPases of the Par complex, such as Cdc42 and Rac, stimulate actin nucleators such as formin and Arp2/3 and suppress destabilizing factors. Additionally, Par3 present at TJs regulates the Rac exchange factor TIAM1, thus recruiting a ring of F-actin to the cell cortex adjacent to the AJs (Chen and Macara, 2005). Like mammalian Par3, Bazooka stabilizes the formation of an actin belt around the apical cell cortex in *Drosophila*. In the *C. elegans* embryo the asymmetric distribution of Par proteins along the anterior–posterior axis generates an asymmetry of force applied to the spindle, which results in the activation of the posterior pulling motors and the inhibition of the anterior motors (Colombo *et al.*, 2003).

Planar Cell Polarity

In addition to the apical–basal polarity, most epithelial tissues present a second type of polarity, which is in the plane of the tissue, perpendicular to the apical–basal axis. This type of polarity is known as planar cell polarity (PCP), and it can be found not only in epithelial tissues but also in mesenchymal sheets. PCP was first described in the *Drosophila* eye and wing imaginal disc (Adler, 2002). Although gene redundancy hampers the study of PCP generation in vertebrates, mutation and overexpression studies performed in recent years have revealed that many of the proteins linked to PCP in flies are evolutionarily conserved and have analogous roles in vertebrates. The most evident examples of mammalian epithelia with PCP features are ‘hairy’ organs such as the cochlear cells of the inner ear and skin/epidermis (Montcouquiol *et al.*, 2003; Guo *et al.*, 2004). In *Drosophila* and vertebrates, PCP controls the organization of cells within tissues and oriented cell division (see below). In addition, PCP shows an additional novel role, namely the regulation of polarized and directed cell movements. This process, known as convergent extension or cell intercalation, occurs during gastrulation (Wallingford *et al.*, 2002) and during tubular organ morphogenesis in vertebrates, such as in neural tube closure and renal tubule elongation (Mao *et al.*, 2011).

Molecular Basis

Genetic studies of *Drosophila* have revealed three major groups of PCP genes. The first group consists of the core PCP proteins, including *Frizzled* (Fz), *Dishevelled* (dsh), *Diego* (dg), *Van Gogh* (vang), *Flamingo* (fmi), and *Prickle* (pk). These proteins coordinate polarity between adjacent cells and amplify subcellular asymmetry, producing a local alignment of polarity. In the fly wing imaginal disc, vang and pk accumulate at the proximal side, while Fz, ds, and dg accumulate at the distal side (Zallen, 2007). Fmi is localized to both proximal and distal

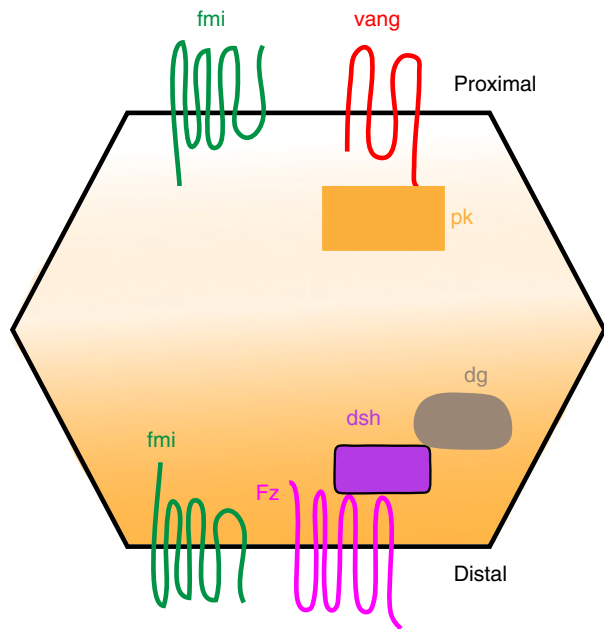


Figure 3 Asymmetric distribution of core PCP proteins in *Drosophila* imaginal disc.

sides (Figure 3). Homologs of these genes have analogous subcellular locations, roles, and interactions in vertebrates. A second group of PCP genes, known as the Fat/Ds PCP group, consists of *fat*, *dachsous* (*ds*), *atrophin* (*atro*), and *four-jointed* (*fj*). Opposing expression gradients of *ds* and *fj* are thought to provide global directional information (Axelrod, 2009).

Downstream of the core PCP genes and *fat/ds* are the PCP effectors that are tissue-specific and convert PCP signals into the corresponding cellular response, such as physical remodeling, adhesion and movement, and orientation of the mitotic spindle (Wang and Nathans, 2007). Recent evidence has shown an exciting link between PCP and apical-basal polarity. To start with, core PCP proteins colocalize with AJs (Lindqvist *et al.*, 2010). In addition, Lgl has been reported to be involved in PCP signaling in *Drosophila* (Kaplan and Tolwinski, 2010), and Scrib has been shown to be an effector of the core PCP protein Vang in mammalian epithelial cell lines (Courbard *et al.*, 2009).

Relationship between Cell Polarity and Spindle Orientation Control

The orientation of the cell division plane is one of the mechanisms that coordinate the rate of cell division with cell fate choice and cell position. This coordination depends on the orientation of the mitotic spindle relative to a defined polarity axis, and it is essential for animal development, as in the case of epidermis. Basal epidermal cells are symmetrically dividing cells, and therefore their spindle is oriented in the plane of the tissue, thus shaping, expanding, and maintaining tissue morphology. At later developmental stages, in order to generate cellular diversity and stratification of the skin, cells start to divide asymmetrically, i.e., they need to align the spindle

parallel to the polarity axis so that basal cell fate determinants segregate into only one daughter cell (Knoblich, 2010).

It is worth highlighting that many of the genes controlling epithelial cell polarity also regulate spindle orientation. In the *C. elegans* zygote, the Par complex, a fundamental player in establishing AP cortical polarity axis, also drives spindle orientation (Kemphues *et al.*, 1988). Upon fertilization, the *C. elegans* zygote divides asymmetrically to generate a larger anterior cell and a smaller posterior cell. The orientation and positioning of the mitotic spindle depend on microtubules that emanate from the spindle poles and make contact with the cortical Par complex and the cortical polarity proteins LIN-5 and GPR1/2. These polarity factors are unequally distributed in the cell cortex, so they generate an asymmetric distribution of net pulling forces acting on the two centrosomes of the mitotic spindle (Betschinger and Knoblich, 2004).

In asymmetrically dividing *Drosophila* neuroblasts the apical Par complex is connected via Inscuteable (Insc) to the apical machinery composed of Mud and dynein, which generates pulling forces on the mitotic spindle (Schober *et al.*, 1999; Figure 4(a)). A similar pathway is used for spindle orientation in mammalian epithelial cells (Lechler and Fuchs, 2005). The initial attachment of the astral microtubules to the basolateral membranes has recently been shown to occur through IQGAP1 and EGFR proteins (Banon-Rodriguez *et al.*, 2014). In *Drosophila* neuroblasts a second pathway, involving Dlg and the kinesin motor protein Khc-73, has been shown to provide a parallel mechanism of spindle orientation through microtubule-capture or attachment (Bergstrahl *et al.*, 2013; Figure 4(a)). Also, aPKC phosphorylates cell fate determinants to exclude them from the apical cortex and to ensure that they are inherited by the ganglion mother cell (GMC) only (Atwood and Prehoda, 2009).

In *Drosophila* sensory organ precursors (SOP), the Mud-dynein complex has been shown to directly bind to the Frizzled-Dishevelled complex, thus providing the molecular link between PCP proteins and spindle orientation machinery (Segalen *et al.*, 2010; Figure 4(b)). These results highlight a striking example of convergent evolution of a shared signaling pathway (Mud-dynein) downstream of divergent upstream activators (Pins/Dsh).

Loss of Cell Polarity and Disease

Loss of cell polarity is a characteristic hallmark of cancer. When polarity is disrupted, cells may become unresponsive to growth inhibitory signaling and thus circumvent differentiation, senescence, and apoptosis, promoting cancer initiation. In addition, disruption of epithelial cell polarity is associated with loss of key components of intercellular junctions, which precedes cells undergoing epithelial-mesenchymal transition (EMT), a crucial step in metastasis (Nieto, 2013). In humans, more than 80% of tumors originate from epithelial tissues.

Exciting studies in *Drosophila* have clearly demonstrated that any perturbation of correct asymmetric cell division results in abnormal accumulation of dividing cells and cancer. The polarity proteins Lgl, Dlg, and Scrib, which control asymmetric division (see above), are therefore potent tumor

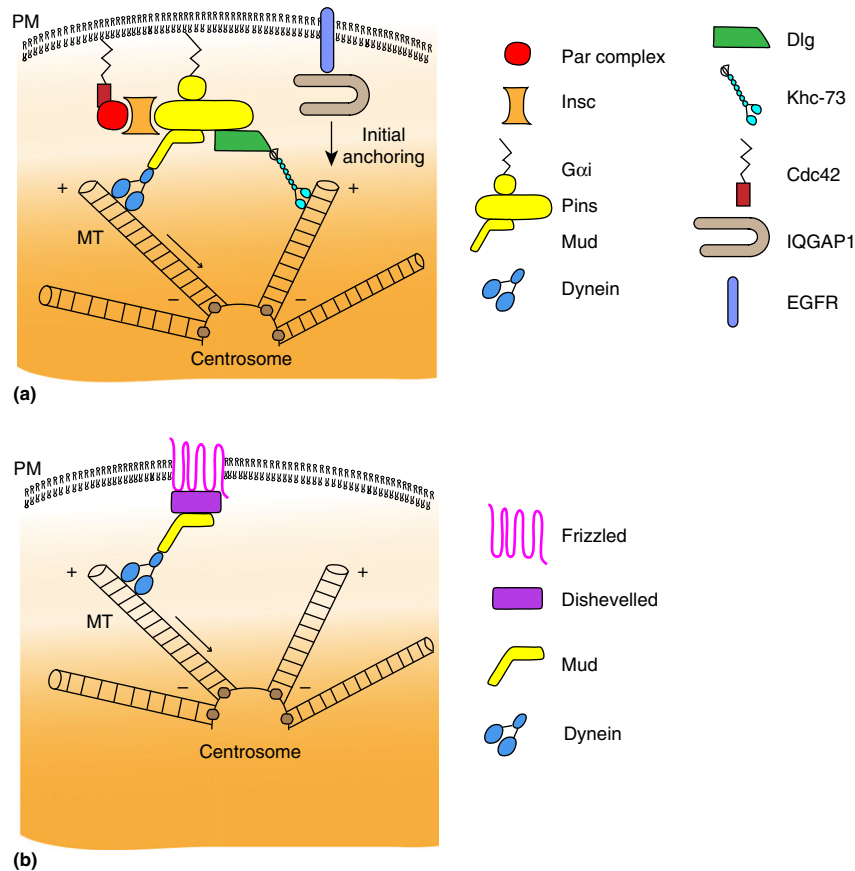


Figure 4 Molecular models for Pins- and Dsh-mediated spindle orientation. (a) The Par complex associates with Pins (Partner of Inscuteable) through the adapter protein Insc. Pins binds directly to Mud, which associates with the microtubule motor protein dynein, which moves toward microtubule minus-ends and generates pulling forces on the mitotic spindle. Upon phosphorylation by Aurora-A, Pins also binds to Dlg. Dlg in turn associates with Khc-73, a kinesin motor protein that moves toward and localizes at microtubule plus-ends, thus providing a microtubule-capture mechanism. (b) The Fz-Dsh complex regulates spindle orientation in *Drosophila* SOP through association with Mud, which binds to dynein.

suppressors. *Drosophila* neuroblasts with mutant *lgl*, *dlg*, and *scrib* fail to localize cell fate determinants correctly, and they generate daughter cells that divide and accumulate rapidly, leading to tumor development in larval imaginal discs and brains (Albertson and Doe, 2003). Currently, there is no definitive evidence that alteration of the mechanism of asymmetric division could be the cause of mammalian tumors. However, a strong correlation has been clearly described between changes (either downregulation or upregulation) in the normal expression pattern of polarity/junctional complexes and cancer. For example, mutations in the *CRB1* gene cause autosomal recessive retinitis pigmentosa (arRP) and autosomal Leber congenital amaurosis (arLCA). Several lines of evidence, obtained in both *Drosophila* and mice, show that loss of function of *Crb* or some of its intracellular interactors leads to morphological defects, loss of adhesion between photoreceptors, and light-induced degeneration of photoreceptor cells (Mehalow *et al.*, 2003), features comparable to those observed in patients lacking *CRB1* function (Den Hollander *et al.*, 2004). Furthermore, downregulation and loss of E-cadherin has been found in familial gastric cancers and breast cancers, respectively (Schrader *et al.*, 2008). Also, loss of α -catenin has been shown to increase cell proliferation,

decrease apoptosis, and suppress actin polymerization, which in turn decreases cell–cell adhesion (Benjamin and Nelson, 2008).

It should be added that cell proliferation and invasion have been correlated not only with loss of apical–basal polarity, but also with disruption of PCP signaling. Loss of the PCP protein Vangl 2 has been shown to promote migration and invasion of human cancer cells (Cantrell and Jessen, 2010). Polycystic kidney disease (PKD), a common genetic disorder in which abnormally growing fluid-filled cysts displace normal renal tubules, is a prime example of disease caused by defects in both planar and apical–basal polarity (Wilson, 2011).

Conclusions

- Cell polarity is responsible for orienting many cellular functions, such as cell shape, cell migration, and cell division.
- Conserved polarity complexes and common mechanisms generate and transduce cell polarity in both vertebrates and invertebrates.
- Epithelial cells display both apical–basal and planar polarity.

- Many of the genes controlling epithelial cell polarity regulate spindle orientation.
- Alterations of junctional and cell polarity complexes have been found to be associated with hyperproliferation and carcinogenesis.

See also: Cell Communication: Cellular Outcomes: Regulation of Cell Polarity

References

- Adler, P.N., 2002. Planar signaling and morphogenesis in *Drosophila*. *Developmental Cell* 2, 525–535.
- Akhtar, N., Streuli, C.H., 2013. An integrin-ILK-microtubule network orients cell polarity and lumen formation in glandular epithelium. *Nature Cell Biology* 15, 17–27.
- Albertson, R., Doe, C.Q., 2003. Dlg, Scrib and Lgl regulate neuroblast cell size and mitotic spindle asymmetry. *Nature Cell Biology* 5, 166–170.
- Apodaca, G., Gallo, L.I., Bryant, D.M., 2012. Role of membrane traffic in the generation of epithelial cell asymmetry. *Nature Cell Biology* 14, 1235–1243.
- Atwood, S.X., Prehoda, K.E., 2009. aPKC phosphorylates Miranda to polarize fate determinants during neuroblast asymmetric cell division. *Current Biology* 19, 723–729.
- Axelrod, J.D., 2009. Progress and challenges in understanding planar cell polarity signaling. *Seminars in Cell and Developmental Biology* 20, 964–971.
- Baklava, Z., Pant, S., Fares, H., Grant, B.D., 2007. Genome-wide analysis identifies a general requirement for polarity proteins in endocytic traffic. *Nature Cell Biology* 9, 1066–1073.
- Banon-Rodriguez, I., Galvez-Santisteban, M., Vergara-Jauregui, S., et al., 2014. EGFR controls IQGAP basolateral membrane localization and mitotic spindle orientation during epithelial morphogenesis. *EMBO Journal* 33, 129–145.
- Bedzhov, I., Zernicka-Goetz, M., 2014. Self-organizing properties of mouse pluripotent cells initiate morphogenesis upon implantation. *Cell* 156, 1032–1044.
- Benjamin, J.M., Nelson, W.J., 2008. Bench to bedside and back again: Molecular mechanisms of alpha-catenin function and roles in tumorigenesis. *Seminars in Cancer Biology* 18, 53–64.
- Bergstrahl, D.T., Lovegrove, H.E., St Johnston, D., 2013. Discs large links spindle orientation to apical-basal polarity in *Drosophila* epithelia. *Current Biology* 23, 1707–1712.
- Betschinger, J., Knoblich, J.A., 2004. Dare to be different: Asymmetric cell division in *Drosophila*, *C. elegans* and vertebrates. *Current Biology* 14, R674–R685.
- Bryant, D.M., Datta, A., Rodriguez-Fraticelli, A.E., et al., 2010. A molecular network for *de novo* generation of the apical surface and lumen. *Nature Cell Biology* 12, 1035–1045.
- Cantrell, V.A., Jessen, J.R., 2010. The planar cell polarity protein Van Gogh-Like 2 regulates tumor cell migration and matrix metalloproteinase-dependent invasion. *Cancer Letters* 287, 54–61.
- Cao, X., Surma, M.A., Simons, K., 2012. Polarized sorting and trafficking in epithelial cells. *Cell Research* 22, 793–805.
- Chen, X., Macara, I.G., 2005. Par-3 controls tight junction assembly through the Rac exchange factor Tiam1. *Nature Cell Biology* 7, 262–269.
- Colombo, K., Grill, S.W., Kimple, R.J., et al., 2003. Translation of polarity cues into asymmetric spindle positioning in *Caenorhabditis elegans* embryos. *Science* 300, 1957–1961.
- Courbard, J.R., Djiane, A., Wu, J., Mlodzik, M., 2009. The apical/basal-polarity determinant Scribble cooperates with the PCP core factor Stbm/Vang and functions as one of its effectors. *Developmental Biology* 333, 67–77.
- Den Hollander, A.I., Davis, J., Van Der Velde-Visser, S.D., et al., 2004. CRB1 mutation spectrum in inherited retinal dystrophies. *Human Mutation* 24, 355–369.
- Denef, N., Chen, Y., Weeks, S.D., Barcelo, G., Schupbach, T., 2008. Crag regulates epithelial architecture and polarized deposition of basement membrane proteins in *Drosophila*. *Developmental Cell* 14, 354–364.
- Etienne-Manneville, S., 2004. Cdc42 — The centre of polarity. *Journal of Cell Science* 117, 1291–1300.
- Galvez-Santisteban, M., Rodriguez-Fraticelli, A.E., Bryant, D.M., et al., 2012. Synaptotagmin-like proteins control the formation of a single apical membrane domain in epithelial cells. *Nature Cell Biology* 14, 838–849.
- Gassama-Diagne, A., Yu, W., Ter Beest, M., et al., 2006. Phosphatidylinositol-3,4,5-trisphosphate regulates the formation of the basolateral plasma membrane in epithelial cells. *Nature Cell Biology* 8, 963–970.
- Goehring, N.W., 2014. PAR polarity: From complexity to design principles. *Experimental Cell Research* 328 (2), 258–266.
- Goehring, N.W., Grill, S.W., 2013. Cell polarity: Mechanochemical patterning. *Trends in Cell Biology* 23, 72–80.
- Guo, N., Hawkins, C., Nathans, J., 2004. Frizzled6 controls hair patterning in mice. *Proceedings of the National Academy of Sciences of the United States of America* 101, 9277–9281.
- Harris, K.P., Tepass, U., 2008. Cdc42 and Par proteins stabilize dynamic adherens junctions in the *Drosophila* neuroectoderm through regulation of apical endocytosis. *Journal of Cell Biology* 183, 1129–1143.
- Harris, T.J., Peifer, M., 2005. The positioning and segregation of apical cues during epithelial polarity establishment in *Drosophila*. *Journal of Cell Biology* 170, 813–823.
- Howard, J., Grill, S.W., Bois, J.S., 2011. Turing's next steps: The mechanochemical basis of morphogenesis. *Nature Reviews Molecular Cell Biology* 12, 392–398.
- Kaplan, N.A., Tolwinski, N.S., 2010. Spatially defined Dsh-Lgl interaction contributes to directional tissue morphogenesis. *Journal of Cell Science* 123, 3157–3165.
- Kemphues, K.J., Priess, J.R., Morton, D.G., Cheng, N.S., 1988. Identification of genes required for cytoplasmic localization in early *C. elegans* embryos. *Cell* 52, 311–320.
- Knoblich, J.A., 2010. Asymmetric cell division: Recent developments and their implications for tumour biology. *Nature Reviews Molecular Cell Biology* 11, 849–860.
- Kobielak, A., Fuchs, E., 2004. Alpha-catenin: At the junction of intercellular adhesion and actin dynamics. *Nature Reviews Molecular Cell Biology* 5, 614–625.
- Laprise, P., Beronja, S., Silva-Gagliardi, N.F., et al., 2006. The FERM protein Yurt is a negative regulatory component of the Crumbs complex that controls epithelial polarity and apical membrane size. *Developmental Cell* 11, 363–374.
- Laprise, P., Lau, K.M., Harris, K.P., et al., 2009. Yurt, Coracle, Neurexin IV and the Na⁺(+),K⁺(+)-ATPase form a novel group of epithelial polarity proteins. *Nature* 459, 1141–1145.
- Laprise, P., Viel, A., Rivard, N., 2004. Human homolog of disc-large is required for adherens junction assembly and differentiation of human intestinal epithelial cells. *Journal of Biological Chemistry* 279, 10157–10166.
- Lechler, T., Fuchs, E., 2005. Asymmetric Cell divisions promote stratification and differentiation of mammalian skin. *Nature* 437, 275–280.
- Lecuit, T., 2004. Junctions and vesicular trafficking during *Drosophila* cellularization. *Journal of Cell Science* 117, 3427–3433.
- Li, R., Bowerman, B., 2010. Symmetry breaking in biology. *Cold Spring Harbor Perspectives in Biology* 2, a003475.
- Lin, D., Edwards, A.S., Fawcett, J.P., et al., 2000. A mammalian PAR-3-PAR-6 complex implicated in Cdc42/Rac1 and aPKC signalling and cell polarity. *Nature Cell Biology* 2, 540–547.
- Lindqvist, M., Horn, Z., Bryja, V., et al., 2010. Vang-like protein 2 and Rac1 interact to regulate adherens junctions. *Journal of Cell Science* 123, 472–483.
- Macara, I.G., Mili, S., 2008. Polarity and differential inheritance — Universal attributes of life? *Cell* 135, 801–812.
- Madrid, R., Aranda, J.F., Rodriguez-Fraticelli, A.E., et al., 2010. The formin INF2 regulates basolateral-to-apical transcytosis and lumen formation in association with Cdc42 and MAL2. *Developmental Cell* 18, 814–827.
- Mao, Y., Mulvaney, J., Zakaria, S., et al., 2011. Characterization of a Dchs1 mutant mouse reveals requirements for Dchs1-Fat4 signaling during mammalian development. *Development* 138, 947–957.
- Martin-Belmonte, F., Gassama, A., Datta, A., et al., 2007. PTEN-mediated apical segregation of phosphoinositides controls epithelial morphogenesis through Cdc42. *Cell* 128, 383–397.
- Martin-Belmonte, F., Mostov, K., 2007. Phosphoinositides control epithelial development. *Cell Cycle* 6, 1957–1961.
- Martin-Belmonte, F., Mostov, K., 2008. Regulation of cell polarity during epithelial morphogenesis. *Current Opinion in Cell Biology* 20, 227–234.
- Mehalow, A.K., Kameya, S., Smith, R.S., et al., 2003. CRB1 is essential for external limiting membrane integrity and photoreceptor morphogenesis in the mammalian retina. *Human Molecular Genetics* 12, 2179–2189.
- Montcouquiol, M., Rachel, R.A., Lanford, P.J., et al., 2003. Identification of Vangl2 and Scrb1 as planar polarity genes in mammals. *Nature* 423, 173–177.
- Nagai-Tamai, Y., Mizuno, K., Hirose, T., Suzuki, A., Ohno, S., 2002. Regulated protein-protein interaction between aPKC and PAR-3 plays an essential role in the polarization of epithelial cells. *Genes to Cells* 7, 1161–1171.
- Naylor, M.J., Li, N., Cheung, J., et al., 2005. Ablation of beta1 integrin in mammary epithelium reveals a key role for integrin in glandular morphogenesis and differentiation. *Journal of Cell Biology* 171, 717–728.

- Niessen, C.M., Gottardi, C.J., 2008. Molecular components of the adherens junction. *Biochimica et Biophysica Acta* 1778, 562–571.
- Nieto, M.A., 2013. Epithelial plasticity: A common theme in embryonic and cancer cells. *Science* 342, 1234850.
- Ooshio, T., Fujita, N., Yamada, A., *et al.*, 2007. Cooperative roles of Par-3 and afadin in the formation of adherens and tight junctions. *Journal of Cell Science* 120, 2352–2365.
- Oztan, A., Silvis, M., Weisz, O.A., *et al.*, 2007. Exocyst requirement for endocytic traffic directed toward the apical and basolateral poles of polarized MDCK cells. *Molecular Biology of the Cell* 18, 3978–3992.
- Rodriguez-Fraticelli, A.E., Galvez-Santisteban, M., Martin-Belmonte, F., 2011. Divide and polarize: Recent advances in the molecular mechanism regulating epithelial tubulogenesis. *Current Opinion in Cell Biology* 23, 638–646.
- Roeth, J.F., Sawyer, J.K., Wilner, D.A., Peifer, M., 2009. Rab11 helps maintain apical crumbs and adherens junctions in the *Drosophila* embryonic ectoderm. *PLoS ONE* 4, e7634.
- Roh, M.H., Fan, S., Liu, C.J., Margolis, B., 2003. The Crumbs3-Pals1 complex participates in the establishment of polarity in mammalian epithelial cells. *Journal of Cell Science* 116, 2895–2906.
- Roussos, T., Shewan, A.M., Mostov, K.E., Schejter, E.D., Shilo, B.Z., 2013. Apical targeting of the formin Diaphanous in *Drosophila* tubular epithelia. *Elife* 2, e00666.
- Sawa, H., 2012. Control of cell polarity and asymmetric division in *C. elegans*. *Current Topics in Developmental Biology* 101, 55–76.
- Schober, M., Schaefer, M., Knoblich, J.A., 1999. Bazooka recruits Inscuteable to orient asymmetric cell divisions in *Drosophila* neuroblasts. *Nature* 402, 548–551.
- Schrader, K.A., Masciari, S., Boyd, N., *et al.*, 2008. Hereditary diffuse gastric cancer: Association with lobular breast cancer. *Familial Cancer* 7, 73–82.
- Segalen, M., Johnston, C.A., Martin, C.A., *et al.*, 2010. The Fz-Dsh planar cell polarity pathway induces oriented cell division via Mud/NuMA in *Drosophila* and zebrafish. *Developmental Cell* 19, 740–752.
- Slaughter, B.D., Smith, S.E., Li, R., 2009. Symmetry breaking in the life cycle of the budding yeast. *Cold Spring Harbor Perspectives in Biology* 1, a003384.
- Sugioka, K., Sawa, H., 2012. Formation and functions of asymmetric microtubule organization in polarized cells. *Current Opinion in Cell Biology* 24, 517–525.
- Suzuki, A., Hirata, M., Kamimura, K., *et al.*, 2004. aPKC acts upstream of PAR-1b in both the establishment and maintenance of mammalian epithelial polarity. *Current Biology* 14, 1425–1435.
- Takai, Y., Nakanishi, H., 2003. Nectin and afadin: Novel organizers of intercellular junctions. *Journal of Cell Science* 116, 17–27.
- Tepass, U., 2012. The apical polarity protein network in *Drosophila* epithelial cells: Regulation of polarity, junctions, morphogenesis, cell growth, and survival. *Annual Review of Cell and Developmental Biology* 28, 655–685.
- Tsukita, S., Katsuno, T., Yamazaki, Y., Umeda, K., Tamura, A., 2009. Roles of ZO-1 and ZO-2 in establishment of the belt-like adherens and tight junctions with paracellular permselective barrier function. *Annals of the New York Academy of Sciences* 1165, 44–52.
- Wallingford, J.B., Fraser, S.E., Harland, R.M., 2002. Convergent extension: The molecular control of polarized cell movement during embryonic development. *Developmental Cell* 2, 695–706.
- Wang, Y., Nathans, J., 2007. Tissue/planar cell polarity in vertebrates: New insights and new questions. *Development* 134, 647–658.
- Wilson, P.D., 2011. Apico-basal polarity in polycystic kidney disease epithelia. *Biochimica et Biophysica Acta* 1812, 1239–1248.
- Winter, J.F., Hopfner, S., Korn, K., *et al.*, 2012. *Caenorhabditis elegans* screen reveals role of PAR-5 in RAB-11-recycling endosome positioning and apicobasal cell polarity. *Nature Cell Biology* 14, 666–676.
- Yamanaka, T., Horikoshi, Y., Sugiyama, Y., *et al.*, 2003. Mammalian Lgl forms a protein complex with PAR-6 and aPKC independently of PAR-3 to regulate epithelial cell polarity. *Current Biology* 13, 734–743.
- Yasuda, T., Saegusa, C., Kamakura, S., Sumimoto, H., Fukuda, M., 2012. Rab27 effector Slp2-a transports the apical signaling molecule podocalyxin to the apical surface of MDCK II cells and regulates claudin-2 expression. *Molecular Biology of the Cell* 23, 3229–3239.
- Yu, W., Datta, A., Leroy, P., *et al.*, 2005. Beta1-integrin orients epithelial polarity via Rac1 and laminin. *Molecular Biology of the Cell* 16, 433–445.
- Zallen, J., 2007. Planar polarity and tissue morphogenesis. *Cell* 129 (6), 1051–1063.
- Zegers, M.M., O'Brien, L.E., Yu, W., Datta, A., Mostov, K.E., 2003. Epithelial polarity and tubulogenesis *in vitro*. *Trends in Cell Biology* 13, 169–176.
- Zuo, X., Guo, W., Lipschutz, J.H., 2009. The exocyst protein Sec10 is necessary for primary ciliogenesis and cystogenesis *in vitro*. *Molecular Biology of the Cell* 20, 2522–2529.

Further Reading

- Apodaca, G., Gallo, L.I., Bryant, D.M., 2012. Role of membrane traffic in the generation of epithelial cell asymmetry. *Nature Cell Biology* 14, 1235–1243.
- Eaton, S., Martin-Belmonte, F., 2014. Cargo sorting in the endocytic pathway: A key regulator of cell polarity and tissue dynamics. *Cold Spring Harbor Perspectives in Biology* 6, a016899.
- Goldstein, B., Macara, I.G., 2007. The PAR proteins: Fundamental players in animal cell polarization. *Developmental Cell* 13, 609–622.
- Martin-Belmonte, F., Perez-Moreno, M., 2012. Epithelial cell polarity, stem cells and cancer. *Nature Reviews Cancer* 12, 23–38.
- Nieto, M.A., 2013. Epithelial plasticity: A common theme in embryonic and cancer cells. *Science* 342, 1234850.
- St Johnston, D., Ahringer, J., 2010. Cell polarity in eggs and epithelia: Parallels and diversity. *Cell* 141, 757–774.
- Thompson, B.J., 2013. Cell polarity: Models and mechanisms from yeast, worms and flies. *Development* 140, 13–21.

INTRACELLULAR INFECTIOLOGY: CELL PROCESSES

Contents

Phagocytosis

Macropinocytosis

Microbicidal Mechanisms

Bacterial Subversion of Phagocytic Killing

Cellular Invasion by Bacterial Pathogens

Phagocytosis

F Niedergang, Institut Cochin, Paris, France; Inserm, CNRS, Paris, France; and Université Paris Descartes, Paris, France

© 2016 Elsevier Inc. All rights reserved.

Glossary

Autophagy (from the Greek *auto-*, 'self,' and *phagein*, 'to eat') The mechanism that involves intracellular degradation of unnecessary or dysfunctional cellular components through the formation of an isolation membranous compartment (the autophagosome) and its fusion with lysosomes.

Cytoskeleton A scaffold or skeleton that determines the shape of all cells. It forms structures such as flagella and cilia and plays a critical role in intracellular traffic, cell division, and cell migration. It is composed of three types of filaments: actin microfilaments, intermediate filaments, and microtubules. Phagocytosis strictly relies on the actin cytoskeleton as a force to deform the plasma membrane and on microtubules for the maturation of phagosomes into phagolysosomes.

Efferocytosis The process of internalization of dead cells or debris. It is mediated by recognition of ligands on the surface of the dead cell or debris by special receptors on the surface of the cells that will mediate ingestion. From *efferre*, Latin for 'to take to the grave,' 'to bury.'

Lysosomes Intracellular membrane-bound organelles of animal cells, where degradation of biomolecules occurs. They contain lytic enzymes that are active at an acidic pH.

Macropinocytosis A form of fluid-phase endocytosis allowing internalization of large volumes. It relies on actin- and PI3-kinase-dependent membrane ruffle formation that leads to the formation of intracellular macropinosomes. In immature DCs, macropinocytosis is constitutive and is down-regulated when DCs are mature. In other cells, macropinocytosis can be induced by growth factors, viruses, or bacteria.

Opsonization The coating of a particle with molecules, called opsonins, that renders it capable of being bound and ingested by phagocytic cells. Opsonins are host serum factors such as immunoglobulins and components of the complement cascade. From Greek, meaning 'to prepare a meal.'

Professional and nonprofessional phagocytic cells Some cells of the immune system are called professional phagocytes, as they perform phagocytosis very efficiently. They include dendritic cells, monocytes/macrophages, and polymorphonuclear neutrophils. Other cells might perform phagocytosis of cell debris, either occasionally or routinely (such as epithelial cells from the retina). The phagocytosis of cell debris has been called efferocytosis.

Introduction

Phagocytosis is a universal cell function, which starts with the recognition and binding of a particle (over 0.5 μm in diameter), generally in a receptor-dependent manner, and leads to its internalization and degradation. Single-celled eukaryotes such as the mold *Dictyostelium discoideum* and amoebae use phagocytosis for feeding. In higher organisms, phagocytosis is fundamental for host defense against invading pathogens and contributes to the immune and inflammatory responses

(Flannagan *et al.*, 2012; Swanson, 2008; Stuart and Ezekowitz, 2008). In mammals, phagocytosis is the hallmark of specialized cells including macrophages, dendritic cells, and polymorphonuclear neutrophils that can engulf large particles, microorganisms, and cell debris. These cells are collectively referred to as professional phagocytes. Phagocytosis is also important during development and in adulthood for normal turnover and remodeling of tissues and disposal of dead cells (Poon *et al.*, 2014). Indeed, epithelial cells from the retina are specialized in clearing fragments shed by photoreceptor

cells, which is important to enable normal vision. In addition, thyroid and bladder epithelial cells or mesangial cells in the kidney are able to perform phagocytosis. In this article, we will describe phagocytosis but not a related mechanism of internalization called macropinocytosis, which relies on related molecular mechanisms, but does not depend on surface receptors (Swanson, 2008).

Phagocytic Receptors

Many receptors on the surface of phagocytes participate to phagosome formation and maturation, although not all receptors are phagocytic per se, i.e., capable of mediating phagocytosis when expressed in non-phagocytic cells. Receptors on the surface of the phagocytes can be classified into two main classes: receptors for host serum factors (opsonins) such as immunoglobulin (Ig)G and the complement fragment C3b that engage Fc γ receptor (Fc γ Rs) and complement receptors (CR3, α M β 2), respectively, and non-opsonic receptors. The former are the archetypes of phagocytic receptors and have been best characterized. The latter include the Toll-like receptors (TLRs) and the non-TLRs such as lectins and scavenger receptors. They recognize components on the particle surface such as mannose or fucose residues, phosphatidylserine, and lipopolysaccharides (Aderem and Underhill, 1999; Underhill and Goodridge, 2012; Underhill and Ozinsky, 2002; Flannagan *et al.*, 2012; Canton *et al.*, 2013). Scavenger receptors were initially defined to characterize the uptake activity of receptors that recognized specifically oxidized low-density lipoproteins (Brown *et al.*, 1979; Canton *et al.*, 2013). A new nomenclature and classification has recently been proposed for scavenger receptors, which recognize altered molecules of the organism but also pathogen determinants (Prabhudas *et al.*, 2014).

Although recognition of a pathogen or cell debris involves binding to and clustering of several types of phagocytic receptors, the signaling events leading to actin polymerization and membrane deformation were essentially analyzed in experimental models that allow to trigger a single phagocytic receptor (Figure 1). For this, the Fc γ Rs or the two chains of the CR3 integrin receptor can be expressed in non-phagocytic cells, which confers them phagocytic properties per se, and allows

the signaling pathways to be dissected (Groves *et al.*, 2008). Another experimental approach to target a specific receptor is to coat or opsonize inert particles such as beads or red blood cells with a single ligand of interest. Zymosan, which is a heat-inactivated yeast, has also been used as opsonized particle to target Fc γ Rs or CR3, although direct binding of intrinsic yeast surface ligands to other phagocytic receptors like Dectin or mannose receptors cannot be excluded in this case.

Signaling to Actin Polymerization

Phagocytosis is strictly dependent on actin polymerization that will provide the force driving membrane deformation (Figure 2). Actin polymerization is intense and transient, and corresponds to the formation of a specific F-actin ring-like structure, called the phagocytic cup. Actin filaments experience a high turnover, with depolymerization at the base of the phagocytic cup and intense polymerization in the tips of the pseudopods. The actin ring diameter progressively shrinks until the two membrane extensions eventually fuse. This step probably involves a constriction activity generated by Myosin II, and possibly other myosins (Deschamps *et al.*, 2013; Swanson, 2008).

The signaling cascades leading to actin polymerization have been extensively studied. The signaling pathway leading to F-actin polymerization downstream of the Fc γ Rs will be described in detail, because it is considered as an archetype of the phagocytic receptors. Signaling requires receptor clustering, presumably via the binding to adjacent ligands on the surface of the target, in a zipper-like mechanism. Despite earlier suggestions to this effect, lipid microdomains do not seem to be implicated in the local concentration of receptors. It was recently described that receptors have two distinct mobilities in the plasma membrane: some diffuse freely and some are confined in fenestrated cortical actin structures that depend on the activity of Src-family and SYK kinases (Jaumouille *et al.*, 2014).

Upon receptor clustering, the tyrosines in the immunoreceptor tyrosine-based activation motif (ITAM) of the FcR or the associated gamma chain are phosphorylated by Src-family kinases, and serve as docking sites for Src homology 2 (SH2)

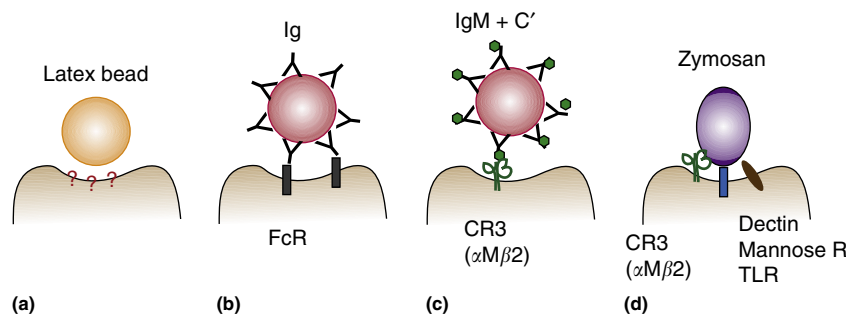


Figure 1 Experimental models of phagocytosis. (a) Beads, especially latex beads, have been extensively studied to induce phagocytosis, although the receptors triggered are not characterized; electrostatic interactions are thought to cluster unknown receptors. (b) and (c) Inert particles, such as beads, zymosan, or fresh red blood cells, can be opsonized with immunoglobulins to target the Fc receptors or with immunoglobulins and complement to target complement receptors such as CR3. (d) Zymosan is yeast inactivated by heat that exposes various pathogen molecular patterns recognized by receptors such as CR3, Dectin, Mannose receptors, and TLR.

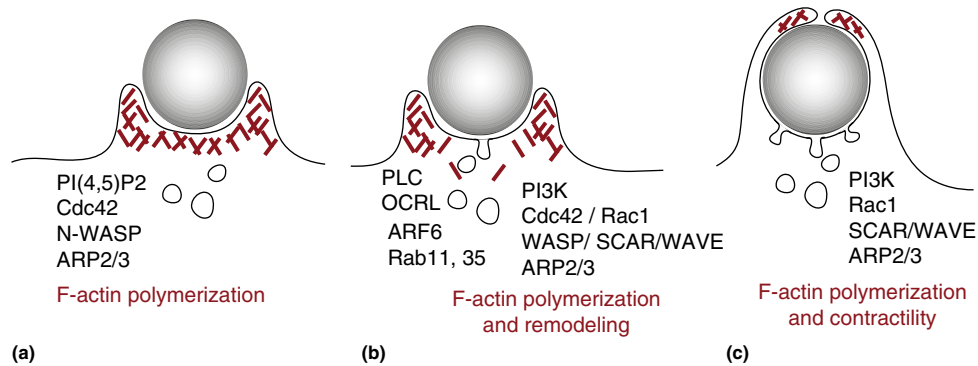


Figure 2 Steps of phagosome formation. (a) Receptor clustering induces initial signaling and cup formation, with a wave of tyrosine phosphorylation, recruitment of adaptor proteins and other protein, and lipid kinases. Small GTPases lead to cascades of activation of actin NPF s and actin polymerization. (b) Pseudopod extension is characterized by further actin polymerization and remodeling at the base of the phagocytic cup, where depolymerization occurs. This is coordinated by membrane-associated phosphatases and allows focal delivery of intracellular compartments, under the control of the small GTPases of the Rab and ARF families. (c) Phagosome extension and closure. Further actin polymerization and membrane secretion, together with contractile activities are required for phagosome completion.

domain-containing cytosolic proteins. Among these are adaptors and scaffolding proteins such as Gab2, SLP-76, and CrkII. In particular, the tyrosine kinase Syk is recruited and in turn phosphorylates several substrates. The p85 subunit of type I phosphatidylinositol 3'-kinase (PI3K) is also recruited to ITAM sites, which allows a local accumulation of phosphatidylinositol 3,4,5-trisphosphate (PI(3,4,5)P₃), necessary to engulf large particles. Indeed, there is a dramatic change from a phase of phosphatidylinositol 4,5-bisphosphate (PI(4,5)P₂) accumulation that is predominant at the plasma membrane and necessary early to form the phagosome, to a phase of quick disappearance, just before phagosome scission (Botelho *et al.*, 2000). The loss of PI(4,5)P₂ leads to a rapid detachment of polymerized actin. Several enzymatic reactions might account for this disappearance and they will be discussed below in the section on actin dynamics.

Small guanosine 5'-triphosphatases (GTPases) of the Rho family are crucial regulators of actin polymerization. During FcγR phagocytosis, Cdc42 and Rac were identified as the key players and RhoG was more recently described to play a predominant role in phagocytosis triggered by other receptors, although this role has not been fully characterized (Caron and Hall, 1998; Cox *et al.*, 1997; Tzircotis *et al.*, 2011). Dynamic studies by fluorescence resonance energy transfer (FRET) revealed different patterns of activation for Rac and Cdc42. Active GTP-Cdc42 is present at the tip of the advancing pseudopod where it colocalizes with polymerizing actin, and Rac1 activation is biphasic. GTP-Rac1 is induced at a low level early after particle binding, and peaks at the time of pseudopod fusion (Hoppe and Swanson, 2004). Cdc42 activation and PI(4,5)P₂ accumulation in the nascent phagocytic cup activates effectors like N-WASP that is an actin nucleation-promoting factor (NPF). It acts on the actin related protein 2/3 (Arp2/3) actin nucleation complex, responsible for pseudopod extension. Rac1 is then essential for F-actin polymerization to complete extension and closure, through activation of another NPF, the WAVE complex. There might be a positive feedback loop leading to a transition from PI(4,5)P₂ and Cdc42 to PI(3,4,5)P₃ and Rac1 in which PI3K plays a role in PI(4,5)P₂ consumption and PI(3)P formation. Indeed, PI(3,4,5)P₃ represents a platform to recruit

other adaptor proteins and guanine nucleotide exchange factors (GEFs) that activate the small GTPases. The accumulated 3' phosphatidylinositols provide a negative feedback response to deactivate Cdc42 that allows actin disassembly necessary for phagosome closure (Beemiller *et al.*, 2010), while Rac1 and PI3K have been reported to signal in a positive feedback loop to sustain PI3K activity (Welch *et al.*, 2003).

Other Receptors and Signaling Pathways to Actin Polymerization

Various signaling pathways are induced downstream of surface phagocytic receptors, but they share signaling modules like kinases, small GTPases, and nucleation complexes. The signaling pathway induced by the CR3 has been studied in model systems using opsonized particles. This approach initially led to the description of two types of phagocytosis, type II being CR3-mediated phagocytosis. The CR3 integrins need an inside-out signal to reach an active conformation, bind their ligands, and become competent for uptake (Dupuy and Caron, 2008). The small GTPases RhoG and RhoA are critical for the signaling to actin polymerization. RhoA in turn activates effectors like the Rho-Kinase (ROK), the formin mDia1, and myosin II that are implicated in polymerization and contraction of F-actin around the particles (Caron and Hall, 1998; Colucci-Guyon *et al.*, 2005; Cox *et al.*, 1997; Olazabal *et al.*, 2002). The microtubules are important for this pathway and CLIP1 is especially required for efficient recruitment of mDia1 downstream of CR3 (Binker *et al.*, 2007; Lewkowicz *et al.*, 2008). The Dectin-1 (CLEC7A, SR-E2) receptor for β-glucans of bacterial and yeast cell walls and the Dectin-2 (CLEC6A) receptors for α-mannans also recently received a lot of attention. They contain ITAM motifs either in their cytosolic domain or in association with the Fcγ chain, which are important for the recruitment of SYK. Of note, phagocytosis of zymosan initiated by Dectin-1 is SYK independent (Kerrigan and Brown, 2010).

Finally the receptors involved in the uptake of cell debris or apoptotic cells induce a process recently termed 'efferocytosis'

(taken from *effero*, Latin for taking to the grave or to bury). Multiple receptors have been described and their redundancy or requirement in various environmental situations (tissue renewal and inflammation) is not well known. The combination of multiple low-affinity receptors provides an avidity that leads to engulfment. The externalization of phosphatidylserine (PtdSer) or its oxidized forms on apoptotic cells is recognized by a number of candidate receptors as well as groups of bridging molecules that bind to PtdSer and then to secondary receptors on the phagocytic cell. Among the former are T-cell immunoglobulin- and mucin domain-containing molecules (TIM1, TIM3, and TIM4), Stabilin-2, and RAGE (Receptor for advanced glycation end products). PtdSer-binding bridging molecules that include milk fat globule EGF 8 (MFG-E8), Protein S, and growth arrest-specific 6 (Gas6) are in turn recognized by phagocytic receptors such as $\alpha_v\beta_3$ integrins and the Tyro-Axl-Mer family (Korns *et al.*, 2011; Poon *et al.*, 2014; Canton *et al.*, 2013). Similar to PtdSer, externally exposed calreticulin is recognized by CD91 on the surface of phagocytic cells. Members of the scavenger receptor family formerly known as CD36 or CD68 have also been implicated in the uptake of apoptotic cells. In addition to these 'eat me' signals, apoptotic cells may dampen some inhibitory signals referred to as 'don't eat me' signals that normally prevent uptake by phagocytic cells (Poon *et al.*, 2014).

Signaling downstream of all these receptors is still a matter of debate and awaits detailed studies. As for other types of receptors, actin polymerization is crucial for efficient uptake. Engagement of efferocytic receptors initiates signaling events modulated by two main complexes, CrkII/ELMO/Dock180 or PTB domain-containing engulfment adaptor protein 1 (GULP), both resulting in activation of Rac1, which initiates cytoskeletal rearrangement and subsequent engulfment (Niedergang and Chavrier, 2005; Poon *et al.*, 2014). Since there are five mammalian DOCK180 family members and three ELMO isoforms, there may be finely tuned interplays between all these molecules, depending on the context and the cell function. Beside the positive regulation of engulfment by Rac1, a negative regulation of this pathway by Rho has been recently described in mammalian cells (Niedergang and Chavrier, 2005; Poon *et al.*, 2014).

Membrane Remodeling and Focal Exocytosis

Vesicular traffic is required for efficient phagosome formation in the case of large ($> 3 \mu\text{m}$) particles. The concept of membrane remodeling and 'focal delivery' of intracellular compartment at the site of phagocytosis emerged in the late 1990s (Hackam *et al.*, 1996, 1998). This was consistent with the observation that, instead of diminishing, the cell surface area increased upon phagocytosis (Hackam *et al.*, 1998; Holeyvinsky and Nelson, 1998). Evidence that focal exocytosis is required for optimal phagocytosis came from studies interfering with the fusion machinery composed of vesicle-soluble *N*-ethylmaleimide-sensitive factor attachment protein receptors (v-SNAREs). Several intracellular compartments have been implicated in focal recruitment of fusion concomitant with phagosome formation, including recycling endosomes bearing VAMP3 on their surface (Bajno *et al.*, 2000; Niedergang *et al.*,

2003; Braun *et al.*, 2007) and late endocytic VAMP7⁺ compartments or lysosomes (Braun *et al.*, 2004; Czibener *et al.*, 2006). The events of compartment recruitment and fusion are regulated by small GTPases of the Rab and ARF family. Rab11, localized on the recycling compartments, is implicated in phagosome formation (Cox *et al.*, 2000; Damiani *et al.*, 2004; Patel and Harrison, 2008). ARF6 was shown to control the delivery of VAMP3 recycling endosomes, and Rab35 regulates actin-dependent phagosome formation by recruiting ACAP2, an ARF6 GTPase-activating protein (Egami *et al.*, 2011) or by regulating the localization of Rac1 and Cdc42 (Shim *et al.*, 2010). In addition, Rab11 and ARF6 activities are coordinated via common effectors like the Rab11-FIP3/4/RIP/RCP (Rab coupling proteins), also named arfophilins, or the exocyst complex. The exocyst complex is composed of eight subunits, among which Sec15, an effector of Rab11 and Sec10, an effector of ARF6, that is conserved from yeast to mammals and controls vesicle docking to the plasma membrane. The process of intracellular compartment recruitment and fusion plays a major role in phagosome formation by releasing tension for phagosome completion, but might also deliver specific cargoes or remodel the lipid composition. It has an additional role in the regulation of actin dynamics as explained in details below.

Linking Actin the Polymerization/Depolymerization Cycle with Vesicular Trafficking

Actin polymerization is strictly required for phagocytosis, but a high turnover of F-actin is also necessary for phagosome extension and closure. One way to generate actin turnover is supported by cofilin, which severs and depolymerizes F-actin filaments, and therefore provides G-actin subunits to support the rapid turnover-barbed end growth of filaments. Cofilin activity is negatively regulated by phosphorylation by the LIM kinase, which is activated by PAK, a direct effector of Rac1 and Cdc42. It was shown that Cdc42 inactivation and F-actin disappearance at the basis of the cup are directly correlated with and dependent on PI(4,5)P₂ hydrolysis (Scott *et al.*, 2005). During phagosome formation, several pathways contribute to the disappearance of PI(4,5)P₂. First, there is an arrest of its synthesis, due to the detachment of the phosphatidylinositol phosphate kinases (PIPKs) from the phagosome. Second, phospholipase C (PLC) γ hydrolyzes PI(4,5)P₂ into diacylglycerol and inositol-1,4,5-trisphosphate. Third, PI(4,5)P₂ is also phosphorylated and thereby consumed by class I PI3-kinases (PI3K) into PI(3,4,5)P₃. The PI(4,5)P₂ and PI(3,4,5)P₃ phosphatase oculocerebrorenal syndrome of Lowe (OCRL) has also been implicated in PI(4,5)P₂ hydrolysis and F-actin removal during phagocytosis in mammalian cells as in *Dicystostelium* (Marion *et al.*, 2012; Looovers *et al.*, 2007). OCRL is membrane-associated and recruited at the site of phagocytosis on AP1-positive compartments under the control of the NF- κ B signaling protein Bcl10 (Marion *et al.*, 2012; Deschamps *et al.*, 2013). As in other systems, the Rab5 and Rab35 GTPases could be also implicated in the recruitment of OCRL at sites of phagocytosis, and the Rab35/OCRL and the Rab11/FIP3 recycling pathways could act in parallel to restrict F-actin accumulation. The small GTP binding protein ARF6 is also a potent regulator of membrane trafficking and F-actin

remodeling (D'Souza-Schorey and Chavrier, 2006). ARF6 was shown to be activated during phagosome formation and to regulate efficient phagocytosis by controlling the focal delivery of intracellular components (Beemiller *et al.*, 2006; Niedergang *et al.*, 2003; Zhang *et al.*, 1998). Therefore, a complex interplay between intracellular compartments and signaling controls the actin dynamics and turnover that is required for vesicle exocytosis and further membrane extension and closure.

Maturation of Phagosomes into Phagolysosomes

Once sealed, the phagosome undergoes a series of fusion and fission events that lead to its progressive maturation into a phagolysosome compartment (Figure 3). The maturation process is accompanied by an acidification of the lumen and its migration on microtubules toward the cell center to reach a perinuclear localization. The early phagosome is similar to early endosomes, with the archetypal marker Rab5 and one of its effectors, the early endosome antigen 1 (EEA1). It has a pH of 6.1–6.5 and is poorly degradative. Effectors of Rab5A include EEA1, p150-hVPS34 complex, and SNARE proteins. The Ser and Thr kinase p150 allow the recruitment of hVPS34, a class III PI3K. The PI3P produced is recognized by proteins harboring a FYVE or PX domain, such as sorting nexins, Hrs, or p40phox. The products of PI3K are involved in the dissociation of Rab5 but are not essential for the recruitment of Rab7 on the phagosome (Vieira *et al.*, 2003). A critical step in phagosome maturation is the conversion from a Rab5

to a Rab7-positive compartment. The Rab5 effector identified as Mon1/SAND1, in a complex with Ccz1, acts as a GEF for Rab7 during endosome maturation and phagocytosis of apoptotic debris (Nordmann *et al.*, 2010; Kinchen and Ravichandran, 2010; Poteryaev *et al.*, 2010). The complex plays an important role in removing the Rabex5 GEF for Rab5 and attracting and activating Rab7 on phagosomes. The homotypic fusion and protein sorting (HOPS) complex, one of the many downstream regulators of Rab7 is implicated in phagosome maturation. The complex is composed of Vps11, Vps16, Vps18, Vps33, Vps39, and Vps41. The Vps41 protein, a key component of the HOPS complex, is required for the stabilization of the HOPS complex on the endosomal membrane prior to fusion with the vacuole in yeast. A key role and regulation by p38 MAPkinases of Vps41 has recently been highlighted to explain the differential traffic of virulent LPS of *Coxiella burnetii* (Barry *et al.*, 2012). Finally, emerging evidences suggest that maturation of phagosomes is also regulated by the recruitment of elements of the autophagy machinery, in particular the microtubule-associated protein 1 light chain 3 (LC3). Intersections of noncanonical autophagy and phagocytosis might have important roles in cell activation and antigen presentation (Mehta *et al.*, 2014; Romao and Munz, 2014).

The conversion from a Rab5- to a Rab7-positive compartment leads to a late phagosome with a more acidic pH of 5.5–6.0 due to acquisition of proton-pumping v-ATPases. The late phagosome is enriched in proteases and lysosomal-associated proteins (LAMPs). Various Rab7 effectors are then recruited to the late phagosome, including RILP and ORP1L, which act as adaptors for dynein (Harrison *et al.*, 2003; Johansson *et al.*,

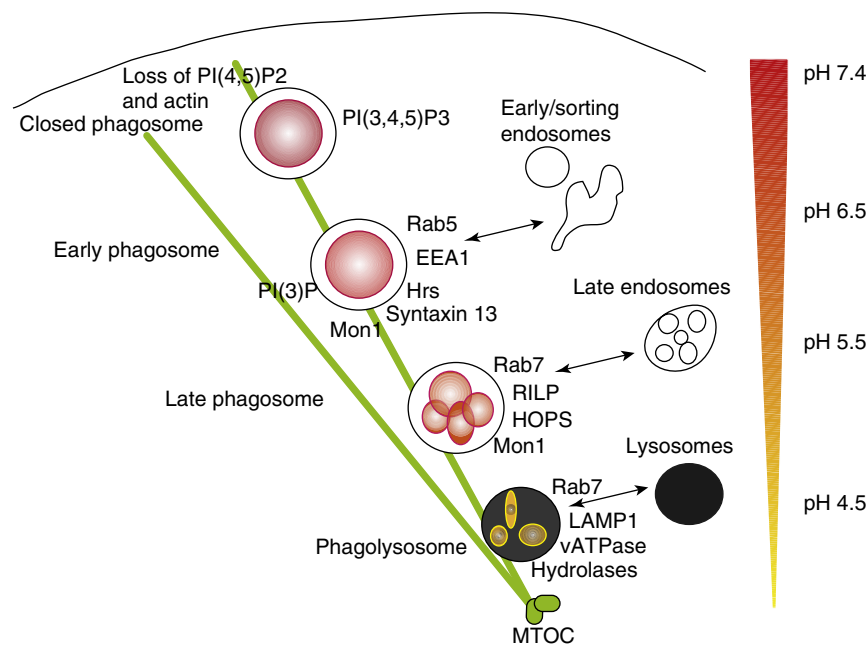


Figure 3 Phagosome maturation. Once closed, the newly formed phagosome quickly loses the actin coat and modifies the lipid composition of its membrane. The phagosome then acquires characteristics of early endosomes and a mildly acidic lumen. A series of fusion and fission events with the endocytic compartments lead to the formation of a late phagosome with a more acidic lumen. Concomitantly, the phagosome moves along microtubules in a centripetal manner. Phagolysosomes, with an acidic pH guaranteeing the efficiency of acid hydrolases, concentrate around the nucleus, where the degradation of the ingested material occurs.

2007). This microtubule motor promotes the movement toward the minus end of microtubules and allows the centripetal movement of the phagosomes, as described in pioneering studies (Blocker *et al.*, 1997). The coordination between the microtubule-based movements and the fusion machinery is still not well defined, as well as the fate of the phagolysosomal products. Indeed, many studies were performed with nondegradable targets such as latex beads, and 'regurgitation' of the ingested material could be detected in that particular case. The biogenesis and homeostasis of lysosomes, as well as the signaling pathways associated to this organelle, are currently the subject of intense research, indicating that it represents much more than a garbage disposal step. There is no doubt that the same holds true for phagolysosomes.

Conclusion

Different types of phagocytosis stimulated by different receptors induce signaling cascades leading to relatively similar morphologic changes and the engulfment of particulate material. Several pathogens have evolved fine-tuned strategies to induce internalization processes in non-phagocytic cells, mimicking the phagocytic properties of professional phagocytes. In addition, later steps of phagocytosis are also hijacked by the intracellular microorganisms that shape the compartment in which they live to their advantage. These aspects will be developed in dedicated articles.

The signaling cascades triggered downstream of the surface receptors, however, induce various activation profiles of the cells leading to antigen presentation and more or less production of inflammatory cytokines. This important aspect will be the focus of further studies in the future to better understand and potentially manipulate the response of phagocytes that is crucial during the inflammatory response. Phagocytosis is indeed crucial at the onset of the inflammatory response to engulf pathogens. The resulting vascular changes and inflammatory mediators attract other leukocytes to the site of inflammation, in particular the short-lived neutrophils. These changes are normally transient and inflammation undergoes a resolution phase in which monocytes/macrophages play again a crucial role to clear the cell debris resulting from tissue damage and massive neutrophil death. Phagocytosis is therefore critical to control the resolution phase of inflammation. Recent evidence indicates that failure to induce efficient phagocytosis and clearance can lead to pathogen or debris persistence and chronic inflammatory responses responsible for widespread and debilitating human diseases such as atherosclerosis, multiple sclerosis, or type 2 diabetes (Tabas and Glass, 2013). Inefficient triggering of proinflammatory transcriptional responses, on the other hand, can favor the appearance of opportunistic infections. Targeting pathophysiological inflammatory processes without affecting vital immune responses is therefore an immense therapeutic challenge.

See also: Cellular Immunology: Innate Immunity: Scavenger Receptors; Specialized Subsets of Tissue-Resident Macrophages in

Secondary Lymphoid Organs; The Phagocyte NADPH Oxidase: Structure and Assembly of the Key Multicomponent Enzyme of Innate Immunity. Interorganellar Communication: Components: Adaptor Proteins: Inter-Organelle Traffic Controllers; SNAREs: Membrane Fusion and Beyond. Interorganellar Communication: Interplay and Processes: Rabs of the Endosomal Recycling Pathway. Intracellular Infectiology: Cell Processes: Bacterial Subversion of Phagocytic Killing; Cellular Invasion by Bacterial Pathogens; Macropinocytosis; Microbicidal Mechanisms. Organelles: Structure and Function: Conventional and Secretory Lysosomes; Early Endosomal Compartments; The Late Endosome

References

- Aderem, A., Underhill, D.M., 1999. Mechanisms of phagocytosis in macrophages. *Annual Review of Immunology* 17, 593–623.
- Bajno, L., Peng, X.-R., Schreiber, A.D., *et al.*, 2000. Focal exocytosis of VAMP3-containing vesicles at sites of phagosome formation. *Journal of Cell Biology* 149, 697–705.
- Barry, A.O., Boucherit, N., Mottola, G., *et al.*, 2012. Impaired stimulation of p38alpha-MAPK/Vps41-HOPS by LPS from pathogenic *Coxiella burnetii* prevents trafficking to microbicidal phagolysosomes. *Cell Host & Microbe* 12, 751–763.
- Beemiller, P., Hoppe, A.D., Swanson, J.A., 2006. A phosphatidylinositol-3-kinase-dependent signal transition regulates ARF1 and ARF6 during Fc gamma receptor-mediated phagocytosis. *PLoS Biology* 4, e162.
- Beemiller, P., Zhang, Y., Mohan, S., *et al.*, 2010. A Cdc42 activation cycle coordinated by PI 3-kinase during Fc receptor-mediated phagocytosis. *Molecular Biology of the Cell* 21, 470–480.
- Binker, M.G., Zhao, D.Y., Pang, S.J., Harrison, R.E., 2007. Cytoplasmic linker protein-170 enhances spreading and phagocytosis in activated macrophages by stabilizing microtubules. *Journal of Immunology* 179, 3780–3791.
- Blocker, A., Severin, F.F., Burkhardt, J.K., *et al.*, 1997. Molecular requirements for bi-directional movement of phagosomes along microtubules. *Journal of Cell Biology* 137, 113–129.
- Botelho, R.J., Teruel, M., Dierckman, R., *et al.*, 2000. Localized biphasic changes in phosphatidylinositol-4,5-bisphosphate at sites of phagocytosis. *Journal of Cell Biology* 151, 1353–1368.
- Braun, V., Deschamps, C., Raposo, G., *et al.*, 2007. AP-1 and ARF1 control endosomal dynamics at sites of FcR mediated phagocytosis. *Molecular Biology of the Cell* 18, 4921–4931.
- Braun, V., Fraissier, V., Raposo, G., *et al.*, 2004. TI-VAMP/VAMP7 is required for optimal phagocytosis of opsonised particles in macrophages. *EMBO Journal* 23, 4166–4176.
- Brown, M.S., Goldstein, J.L., Krieger, M., Ho, Y.K., Anderson, R.G., 1979. Reversible accumulation of cholesteryl esters in macrophages incubated with acetylated lipoproteins. *Journal of Cell Biology* 82, 597–613.
- Canton, J., Neculai, D., Grinstein, S., 2013. Scavenger receptors in homeostasis and immunity. *Nature Reviews Immunology* 13, 621–634.
- Caron, E., Hall, A., 1998. Identification of two distinct mechanisms of phagocytosis controlled by different Rho GTPases. *Science* 282, 1717–1721.
- Colucci-Guyon, E., Niedergang, F., Wallar, B.J., *et al.*, 2005. A role for mammalian diaphanous-related formins in complement receptor (CR3)-mediated phagocytosis in macrophages. *Current Biology* 15, 2007–2012.
- Cox, D., Chang, P., Zhang, Q., *et al.*, 1997. Requirements for both Rac1 and Cdc42 in membrane ruffling and phagocytosis in leukocytes. *Journal of Experimental Medicine* 186, 1487–1494.
- Cox, D., Lee, D.J., Dale, B.M., Calafat, J., Greenberg, S., 2000. A Rab11-containing rapidly recycling compartment in macrophages that promotes phagocytosis. *Proceedings of the National Academy of Sciences of the United States of America* 97, 680–685.
- Czibener, C., Sherer, N.M., Becker, S.M., *et al.*, 2006. Ca²⁺ and synaptotagmin VII-dependent delivery of lysosomal membrane to nascent phagosomes. *Journal of Cell Biology* 174, 997–1007.
- D'Souza-Schorey, C., Chavrier, P., 2006. ARF proteins: Roles in membrane traffic and beyond. *Nature Reviews Molecular Cell Biology* 7, 347–358.

- Damiani, M.T., Pavarotti, M., Leiva, N., *et al.*, 2004. Rab coupling protein associates with phagosomes and regulates recycling from the phagosomal compartment. *Traffic* 5, 785–797.
- Deschamps, C., Echard, A., Niedergang, F., 2013. Phagocytosis and cytokinesis: Do cells use common tools to cut and to eat? Highlights on common themes and differences. *Traffic* 14, 355–364.
- Dupuy, A.G., Caron, E., 2008. Integrin-dependent phagocytosis: Spreading from microadhesion to new concepts. *Journal of Cell Science* 121, 1773–1783.
- Egami, Y., Fukuda, M., Araki, N., 2011. Rab35 regulates phagosome formation through recruitment of ACAP2 in macrophages during FcγR-mediated phagocytosis. *Journal of Cell Science* 124, 3557–3567.
- Flannagan, R.S., Jaumouille, V., Grinstein, S., 2012. The cell biology of phagocytosis. *Annual Review of Phytopathology* 7, 61–98.
- Groves, E., Dart, A.E., Covarelli, V., Caron, E., 2008. Molecular mechanisms of phagocytic uptake in mammalian cells. *Cellular and Molecular Life Sciences* 65, 1957–1976.
- Hackam, D.J., Rotstein, O.D., Bennett, M.K., *et al.*, 1996. Characterization and subcellular localization of target membrane soluble NSF attachment protein receptors (t-SNAREs) in macrophages. Syntaxis 2, 3, and 4 are present on phagosomal membranes. *Journal of Immunology* 156, 4377–4383.
- Hackam, D.J., Rotstein, O.D., Sjölin, C., *et al.*, 1998. v-SNARE-dependent secretion is required for phagocytosis. *Proceedings of the National Academy of Sciences of the United States of America* 95, 11691–11696.
- Harrison, R.E., Bucci, C., Vieira, O.V., Schroer, T.A., Grinstein, S., 2003. Phagosomes fuse with late endosomes and/or lysosomes by extension of membrane protrusions along microtubules: Role of Rab7 and RILP. *Molecular and Cellular Biology* 23, 6494–6506.
- Holevinsky, K.O., Nelson, D.J., 1998. Membrane capacitance changes associated with particle uptake during phagocytosis in macrophages. *Biophysical Journal* 75, 2577–2586.
- Hoppe, A.D., Swanson, J.A., 2004. Cdc42, Rac1, and Rac2 display distinct patterns of activation during phagocytosis. *Molecular Biology of the Cell* 15, 3509–3519.
- Jaumouille, V., Farkash, Y., Jaqaman, K., *et al.*, 2014. Actin cytoskeleton reorganization by syk regulates fcyγR receptor responsiveness by increasing its lateral mobility and clustering. *Developmental Cell* 29, 534–546.
- Johansson, M., Rocha, N., Zwart, W., *et al.*, 2007. Activation of endosomal dynein motors by stepwise assembly of Rab7-RILP-p150Glued, ORP1L, and the receptor betall spectrin. *Journal of Cell Biology* 176, 459–471.
- Kerrigan, A.M., Brown, G.D., 2010. Syk-coupled C-type lectin receptors that mediate cellular activation via single tyrosine based activation motifs. *Immunology Review* 234, 335–352.
- Kinchen, J.M., Ravichandran, K.S., 2010. Identification of two evolutionarily conserved genes regulating processing of engulfed apoptotic cells. *Nature* 464, 778–782.
- Korns, D., Frasch, S.C., Fernandez-Boyanapalli, R., Henson, P.M., Bratton, D.L., 2011. Modulation of macrophage efferocytosis in inflammation. *Frontiers in Immunology* 2, 57.
- Lewkowicz, E., Herit, F., Le Clairche, C., *et al.*, 2008. The microtubule-binding protein CLIP-170 coordinates mDia1 and actin reorganization during CR3-mediated phagocytosis. *Journal of Cell Biology* 183, 1287–1298.
- Loovers, H.M., Kortholt, A., De Groot, H., *et al.*, 2007. Regulation of phagocytosis in Dictyostelium by the inositol 5-phosphatase OCRL homolog Dd5P4. *Traffic* 8, 618–628.
- Marion, S., Mazzolini, J., Herit, F., *et al.*, 2012. The NF-κB signaling protein Bcl10 regulates actin dynamics by controlling AP1 and OCRL-bearing vesicles. *Developmental Cell* 23, 954–967.
- Mehta, P., Henault, J., Kolbeck, R., Sanjuan, M.A., 2014. Noncanonical autophagy: One small step for LC3, one giant leap for immunity. *Current Opinion in Immunology* 26, 69–75.
- Niedergang, F., Chavrier, P., 2005. Regulation of phagocytosis by Rho GTPases. *Current Topics in Microbiology and Immunology* 291, 43–60.
- Niedergang, F., Colucci-Guyon, E., Dubois, T., Raposo, G., Chavrier, P., 2003. ADP ribosylation factor 6 is activated and controls membrane delivery during phagocytosis in macrophages. *Journal of Cell Biology* 161, 1143–1150.
- Nordmann, M., Cabrera, M., Perz, A., *et al.*, 2010. The Mon1-Ccz1 complex is the GEF of the late endosomal Rab7 homolog Ypt7. *Current Biology* 20, 1654–1659.
- Olazabal, I.M., Caron, E., May, R.C., *et al.*, 2002. Rho-kinase and myosin-II control phagocytic cup formation during CR, but not FcγR, phagocytosis. *Current Biology* 12, 1413–1418.
- Patel, P.C., Harrison, R.E., 2008. Membrane ruffles capture C3bi-opsonized particles in activated macrophages. *Molecular Biology of the Cell* 19, 4628–4639.
- Poon, I.K., Lucas, C.D., Rossi, A.G., Ravichandran, K.S., 2014. Apoptotic cell clearance: Basic biology and therapeutic potential. *Nature Reviews Immunology* 14, 166–180.
- Poteryaev, D., Datta, S., Ackema, K., Zerial, M., Spang, A., 2010. Identification of the switch in early-to-late endosome transition. *Cell* 141, 497–508.
- Prabhudas, M., Bowditch, D., Drickamer, K., *et al.*, 2014. Standardizing scavenger receptor nomenclature. *Journal of Immunology* 192, 1997–2006.
- Romao, S., Munz, C., 2014. LC3-associated phagocytosis. *Autophagy* 10, 526–528.
- Scott, I.C., Dobson, W., Botelho, R.J., *et al.*, 2005. Phosphatidylinositol-4,5-bisphosphate hydrolysis directs actin remodeling during phagocytosis. *Journal of Cell Biology* 169, 139–149.
- Shim, J., Lee, S.M., Lee, M.S., *et al.*, 2010. Rab35 mediates transport of Cdc42 and Rac1 to the plasma membrane during phagocytosis. *Molecular and Cellular Biology* 30, 1421–1433.
- Stuart, L.M., Ezekowitz, R.A., 2008. Phagocytosis and comparative innate immunity: Learning on the fly. *Nature Reviews Immunology* 8, 131–141.
- Swanson, J.A., 2008. Shaping cups into phagosomes and macropinosomes. *Nature Reviews Molecular Cell Biology* 9, 639–649.
- Tabas, I., Glass, C.K., 2013. Anti-inflammatory therapy in chronic disease: Challenges and opportunities. *Science* 339, 166–172.
- Tzircotis, G., Braga, V.M., Caron, E., 2011. RhoG is required for both FcγR- and CR3-mediated phagocytosis. *Journal of Cell Science* 124, 2897–2902.
- Underhill, D.M., Goodridge, H.S., 2012. Information processing during phagocytosis. *Nature Reviews Immunology* 12, 492–502.
- Underhill, D.M., Ozinsky, A., 2002. Phagocytosis of microbes: Complexity in action. *Annual Review of Immunology* 20, 825–852.
- Vieira, O.V., Bucci, C., Harrison, R.E., *et al.*, 2003. Modulation of Rab5 and Rab7 recruitment to phagosomes by phosphatidylinositol 3-kinase. *Molecular and Cellular Biology* 23, 2501–2514.
- Welch, H.C., Coadwell, W.J., Stephens, L.R., Hawkins, P.T., 2003. Phosphoinositide 3-kinase-dependent activation of Rac. *FEBS Letters* 546, 93–97.
- Zhang, Q., Cox, D., Tseng, C.-C., Donaldson, J.G., Greenberg, S., 1998. A requirement for ARF6 in Fcγ receptor-mediated phagocytosis in macrophages. *Journal of Biological Chemistry* 273, 19977–19981.

Macropinocytosis

JA Swanson and S Yoshida, University of Michigan Medical School, Ann Arbor, MI, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Pinocytosis was discovered and named by Warren Lewis, whose time-lapse movies of macrophages and cultured rat sarcoma cells revealed actively moving cell surface protrusions which folded back to enclose extracellular fluids into intracellular vesicles (Lewis, 1931; Lewis, 1937; Video 1). Smaller pinocytic vesicles discovered later by electron microscopy, initially called micropinosomes, were resolved into many different kinds of endocytotic activity: clathrin-coated vesicles, caveolae and uncoated small vesicles (Kumari *et al.*, 2010). Currently, endocytosis refers to all kinds of cellular ingestion. The term pinocytosis refers to fluid-phase endocytosis regardless of endocytic vesicle size. It will be used here to describe activities with undetermined contributions from macro- and micropinocytosis. Macropinocytosis is the process originally imaged by Lewis, the cellular ingestion of extracellular fluid into 0.2 μm or larger pinosomes derived from the plasma membrane.

Macropinosomes appear in cells after apparently chaotic movements of the cell cortex, often in response to acute stimulation by growth factors or other agents. Typically, actin-rich, petal-shaped extensions of plasma membrane, called ruffles, fold back against the cell surface or curve into circular ruffles that close at their distal margins to form intracellular vacuoles (Figure 1(a); Video 2). In some cultured cells that have been deprived of growth factor, the re-addition of growth factor stimulates cell spreading and the formation of two kinds of ruffle: peripheral ruffles and circular dorsal ruffles (CDR; Figure 1(b); Video 3) (Hoon *et al.*, 2012; Itoh and Hasegawa, 2013). Peripheral ruffles and CDR exhibit different dynamics and regulation, but are both capable of forming macropinosomes. CDR and the circular ruffles that form macropinosomes are sometimes considered equivalent structures.

However, the relationship of CDR to mechanisms of macropinosome formation under steady state conditions is not known.

Functions of Macropinocytosis

Although many cells exhibit macropinocytosis constitutively, it is an inducible process in other cells. Macropinocytosis is a common cellular response to stimulation by growth factors and may be important for growth factor receptor (GFR) signaling (Porat-Shliom *et al.*, 2008). Increased pinocytosis was observed after stimulation of amoebae or other cells with cationic molecules or nutrients (Chapman-Andresen, 1971), or of serum-starved cells with their cognate growth factors, including epidermal growth factor (Brunk *et al.*, 1976), platelet-derived growth factor (Davies and Ross, 1978), hepatocyte growth factor (Dowrick *et al.*, 1993), and macrophage-colony-stimulating factor (M-CSF) (Racoosin and Swanson, 1989). Stimulation of macropinocytosis has also been observed in response to cell-penetrating peptides (Nakase *et al.*, 2007), the tumor promoter phorbol myristate acetate (PMA) (Swanson, 1989), and lipopolysaccharide (Knight *et al.*, 1992).

Macropinocytosis contributes to physiology. Cells transformed by oncogenic K-Ras exhibit constitutively elevated macropinocytosis (Bar-Sagi and Feramisco, 1986), which internalizes sufficient protein to provide a source of amino acids for growth (Commisso *et al.*, 2013). This indicates that macropinocytosis provides an efficient route for the ingestion and degradation of solute nutrients in normal cells. In adaptive immunity, macropinocytosis by immature dendritic cells provides a mechanism for indiscriminate sampling of extracellular proteins for antigen processing and presentation by

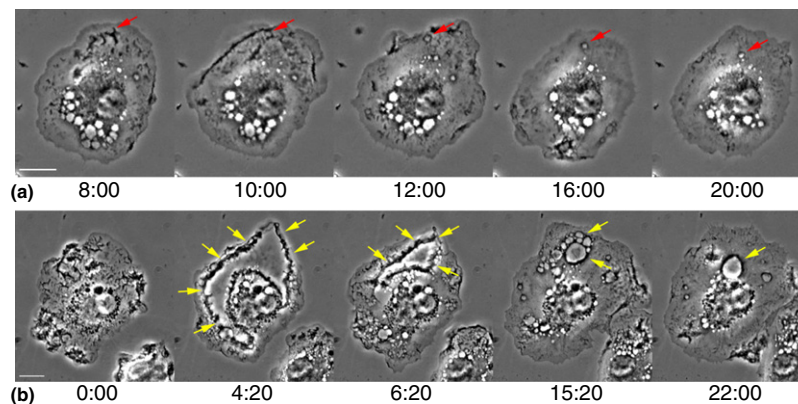


Figure 1 Macropinocytosis and CDR formation in macrophages after stimulation with M-CSF. (a) Phase-contrast microscopic images at the indicated times after addition of M-CSF (min:sec), taken from Video 2. Arrows indicate stages of macropinosome formation: curved ruffle (8:00), circular ruffle (10:00), closed macropinosome (12:00), motile macropinosome (16:00 and 20:00). (b) Phase-contrast images of CDR, taken from Video 3. Arrows indicate the large continuous region of ruffling, which constricts to a smaller region of the cell before closing into several macropinosomes (15:20), which then fuse together and migrate centripetally (22:00). Scale bars = 10 μm .

MHC class I and MHC class II proteins. Constitutive macropinocytosis by immature dendritic cells internalizes extracellular proteins, followed by the partial degradation of proteins by lysosomal enzymes (Sallusto *et al.*, 1995). Upon maturation of dendritic cells, macropinocytosis decreases, with concomitant increases in recycling and presentation of antigen on cell surface MHC class II complexes. Moreover, leakage of internalized proteins from macropinosomes into the cytosol contributes to cross-presentation of exogenous antigen on dendritic cell MHC class I (Norbury *et al.*, 1995).

Although the term macropinocytosis refers to the engulfment of fluid and solutes, the movements of macropinocytosis have been implicated in various kinds of particle phagocytosis, including the cellular ingestion of bacteria, viruses, and apoptotic cells (Francis *et al.*, 1993; Hoffmann *et al.*, 2001; Mercer and Helenius, 2009). Macropinocytosis of low density lipoprotein by macrophages leads to cholesterol accumulation and foam cell formation in atherosclerotic plaques (Kruth, 2013). In neuronal cells, macropinocytosis is an essential part of growth cone collapse in repulsive axon guidance (Kolpak *et al.*, 2009).

Solute Flow During Pinocytosis

Lewis first suggested that pinocytosis provides a mechanism for cells to internalize and degrade extracellular solutes (Lewis, 1931). It was later shown that macrophages ingest proteins and carbohydrates via pinocytosis and degrade them in lysosomes (Cohn and Ehrenreich, 1969; Ehrenreich and Cohn, 1967). Electron microscopic studies of pinocytosis determined that the equivalent of the entire cell surface area in membrane is internalized several times per hour in *Acanthamoeba castellanii* (Bowers and Olszewski, 1972) and every 30 min in macrophages (Steinman *et al.*, 1976). This prompted studies which demonstrated considerable recycling of membrane and fluid-phase solutes during constitutive pinocytosis (Besterman *et al.*, 1981; Swanson *et al.*, 1985) and increased retention of internalized solutes when macropinocytosis was stimulated (Swanson, 1989).

In many cell types, macropinosomes interact to varying degrees with other endocytic organelles, including early endosomes, late endosomes, lysosomes and other macropinosomes (Amyere *et al.*, 2002; Kerr *et al.*, 2006; Racoosin and Swanson, 1993). The dynamics of organelle traffic following pinocytosis can lead to size fractionation of internalized solute molecules into distinct compartments (Berthiaume *et al.*, 1995). In some cells, however, macropinosomes recycle to the plasma membrane after minimal interactions with other organelles (Hewlett *et al.*, 1994).

Mechanisms of Macropinosome Formation

Actin dynamics and myosin-based contractility are necessary for ruffling and macropinosome closure. Many proteins implicated in the functioning of the actin cytoskeleton are required for macropinocytosis, including actinin-4 (Araki *et al.*, 2000), coronin (Grogan *et al.*, 1997) and myosins I, IIB, and V (Edgar and Bennett, 1997; Jiang *et al.*, 2010). The regulation of

the cytoskeleton for CDR formation has been extensively studied (Hoon *et al.*, 2012; Itoh and Hasegawa, 2013). Much of what is known there likely also applies to the regulation of the cytoskeleton for macropinocytosis.

GFR Signals that Affect Macropinosome Formation

Binding of growth factors to GFR initiates signal cascades which lead to cell growth. Many of those signals are also necessary for GFR-induced macropinocytosis. Growth factor binding leads to GFR dimerization, followed by activation of c-Src (Kasahara *et al.*, 2007) and Src-family kinases (SFK) (Metten *et al.*, 2006; Veracini *et al.*, 2006). v-Src-transformed fibroblasts exhibit high rates of macropinocytosis (Veithen *et al.*, 1996). SFK facilitate the recruitment of phospholipase C- γ 1 (PLC γ 1) (Veracini *et al.*, 2006) which hydrolyzes phosphatidylinositol (4,5)-bisphosphate (PtdIns(4,5)P₂) to diacylglycerol (DAG) and inositol trisphosphate (InsP₃). DAG activates protein kinase C, which is required for macropinosome formation (Miyata *et al.*, 1989).

Small GTPases

The movements of macropinocytosis are organized by small GTPases of the Ras superfamily. These proteins alternate between inactive, GDP-bound conformations and active, GTP-bound conformations. GTPases activate various effector proteins, primarily enzymes which regulate the actin cytoskeleton, membrane fusion and fission reactions or other GTPases. Guanine nucleotide exchange factors (GEFs) activate GTPases by accelerating the exchange of GTP for bound GDP. GTPase-activating proteins (GAPs) stimulate GTP hydrolysis and the return to inactive, GDP-bound conformations. GEFs and GAPs are often regulated by membrane phosphoinositides (PIs), which restrict the subcellular domains where GTPases may function (Bohdanowicz and Grinstein, 2013).

Ras

Although Ras is clearly important for macropinocytosis, the exact roles for Ras in macropinosome formation and trafficking remain poorly defined. Ras triggers ruffling through activation of Rac, and stimulates pinocytosis via Rab5 (Li *et al.*, 1997). Ras activates multiple effector proteins, including the p110 α catalytic subunit of type I phosphatidylinositol 3-kinase (PI3K) (Tamura *et al.*, 2009) and the Rab5 GEF Rin1 (Tall *et al.*, 2001), both of which are required for macropinosome formation. Ras remains active on macropinosome membranes, which serve as signaling platforms for growth stimulation (Porat-Shliom *et al.*, 2008; Welliver and Swanson, 2012).

Rac

Rac is required for nearly all kinds of macropinocytosis. Cells expressing constitutively active mutants of Rac exhibit increased ruffling and macropinosome formation (Ridley *et al.*, 1992). Rac is transiently activated in unsealed macropinocytic cups (Yoshida *et al.*, 2009), and experimental activation of Rac in sub-regions of cells induces localized ruffling and macropinosome formation (Fujii *et al.*, 2013). Rac effectors

implicated in macropinocytosis include Alsln/ALS2, which is a GEF for Rab5 (Otomo *et al.*, 2011), WAVE2, which stimulates actin polymerization (Sun *et al.*, 2003) and p21-activated kinase-1 (Pak1) (Dharmawardhane *et al.*, 1997; Dharmawardhane *et al.*, 2000). Pak1 phosphorylates and activates Ctb1/BARS, which mediates closure of the macropinosome and the activation of phospholipase D1 (PLD1) (Haga *et al.*, 2009; Liberali *et al.*, 2008). Significantly, Rac1 must be deactivated for complete macropinosome closure (Fujii *et al.*, 2013).

Rab5

Rab5 regulates macropinosome closure and subsequent trafficking. It localizes to macropinocytic cups prior to closure and persists on the macropinosome until it matures to a Rab7-positive organelle (Feliciano *et al.*, 2011; Porat-Shliom *et al.*, 2008; Welliver and Swanson, 2012). The Ras effector and Rab5 GEF Rin1 stimulates macropinocytosis (Balaji *et al.*, 2012). Rab5 effectors with roles in macropinocytosis include raban-kyrin-5 (Schnatwinkel *et al.*, 2004), Vps34 and EEA1 (Hamasaki *et al.*, 2004); these effectors regulate macropinosome closure or maturation. Together with Ras, Rac, and PI3K, Rab5 is necessary for CDR formation, likely by activating the Rab5 effector RNTre (Lanzetti *et al.*, 2004). Rab5 inhibition by overexpression of Rab GAPs destabilizes macropinosomes and promotes their recycling to the plasma membrane, indicating that a normal function of Rab5 is to facilitate closure or to prevent macropinosome recycling to the plasma membrane (Feliciano *et al.*, 2011). Rab5 may function on membranes of unclosed macropinocytic cups, as Rab5 activates OCRL and Inpp5B, which facilitate phagosome closure (Bohdanowicz *et al.*, 2012).

Other GTPases

Cdc42 regulates actin polymerization. It is active during macropinocytosis in dendritic cells (Garrett *et al.*, 2000) and mediates the macropinocytosis-like phagocytic engulfment of

Salmonella (Chen *et al.*, 1996). Active Arf6 stimulates actin polymerization and macropinocytosis (Porat-Shliom *et al.*, 2008; Schafer *et al.*, 2000). Other small GTPases implicated in macropinocytosis include RhoG (Ellerbroek *et al.*, 2004), Rab20 (Egami and Araki, 2012), Rab21 (Egami and Araki, 2009), and Rab34/Rah (Sun *et al.*, 2003).

Lipids

Phosphoinositides

The phosphoinositide (PI) phosphatidylinositol (PtdIns) and its phosphorylated relatives PtdIns3P, PtdIns4P, PtdIns5P, PtdIns(4,5)-bisphosphate (PtdIns(4,5)P₂), PtdIns(3,4)P₂, PtdIns(3,5)P₂, and phosphatidylinositol (3,4,5)-trisphosphate (PtdIns(3,4,5)P₃) are present in plasma membrane and the membranes of endocytic compartments (Bohdanowicz and Grinstein, 2013). PtdIns3P, PtdIns(4,5)P₂, PtdIns(3,4)P₂, and PtdIns(3,4,5)P₃ have been localized to macropinosomes and several PI kinases, phosphatases and lipases have been implicated in macropinosome formation or maturation (Hasegawa *et al.*, 2011; Maekawa *et al.*, 2014). PtdIns(3,4,5)P₃ appears transiently in circular ruffles (Yoshida *et al.*, 2009). PtdIns3P appears late during macropinosome closure and is required for macropinosome maturation (Araki *et al.*, 2006; Yoshida *et al.*, 2009; Figure 2; Video 4). In macrophages stimulated with M-CSF, PI3K inhibition does not alter ruffling or circular ruffle formation (ruffle closure), but instead inhibits a contractile activity that closes the macropinosome (cup closure) (Araki *et al.*, 1996; Swanson *et al.*, 1999).

Diacylglycerol

During constitutive macropinosome formation in v-Src-transformed fibroblasts, PLC functions after PI3K (Amyere *et al.*, 2000). The PI3K product PtdIns(3,4,5)P₃ activates PLC (Falasca *et al.*, 1998). The PLC product DAG and the DAG analog PMA stimulate pinocytosis (Keller, 1990; Swanson, 1989). DAG appears in macropinocytic cup membranes

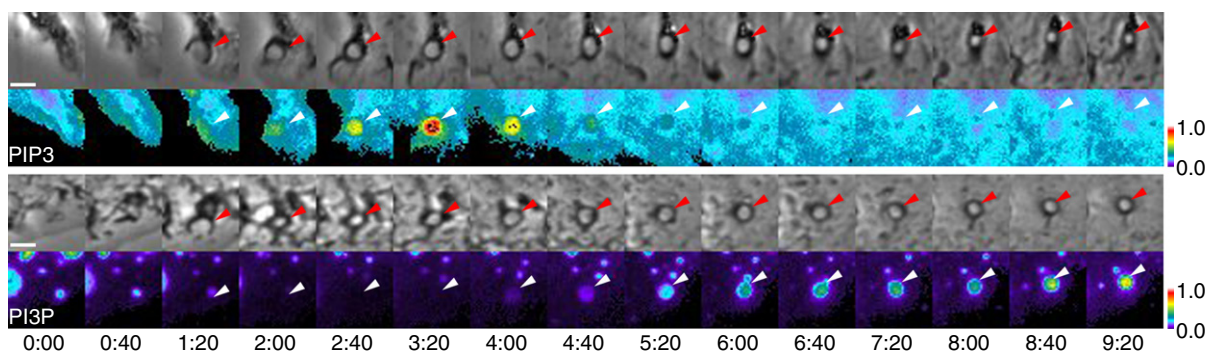


Figure 2 The sequence of 3'PIs during macropinosome formation. The top and bottom sequences are aligned relative to the stages of macropinosome formation; hence, time (min:sec) is indicated relative to the first frame (rather than the time of M-CSF addition). Top: A time series showing macropinosome formation and a spike of PtdIns(3,4,5)P₃/PtdIns(3,4)P₂ (PIP₃) in a macropinocytic cup (taken from Video 4). A mouse macrophage expressing cyan fluorescent protein (CFP) and a yellow fluorescent protein (YFP)-labeled PH domain from Akt (YFP-AktPH), which binds to PtdIns(3,4,5)P₃ and PtdIns(3,4)P₂, was imaged at 20-second intervals after stimulation with M-CSF. Corresponding phase-contrast and YFP/CFP ratio images are shown in a time series. A spike of PtdIns(3,4,5)P₃/PtdIns(3,4)P₂ (PIP₃) (2:40 to 4:00) appears after ruffle closure (2:00) and before cup closure (4:40). Bottom: A time series showing the generation of PtdIns3P during macropinosome formation. A macrophage expressing CFP and a YFP-2xFYVE chimera, which binds to PtdIns3P, was imaged at 20-second intervals after stimulation by M-CSF. PtdIns3P appears during or just after cup closure (4:40) and persists on the closed macropinosome. Color scale bars indicate relative YFP/CFP ratio values. Scale bars = 2 μm.

coincident with or shortly following the spike in PtdIns(3,4,5)P₃ (Welliver and Swanson, 2012).

Phosphatidic Acid

PLD, which synthesizes phosphatidic acid (PA) from phosphatidylcholine, is necessary for macropinosome formation in v-Src-transformed fibroblasts (Platek *et al.*, 2004). PA is required for macropinocytosis (Bohdanowicz *et al.*, 2013) and may function indirectly by activating the PtdIns(4,5)P₂-synthesizing enzyme PtdIns-4-phosphate-5-kinase (Honda *et al.*, 1999) or by recruiting the Rac1 GEF DOCK1 to plasma membrane (Sanematsu *et al.*, 2013).

Cholesterol

Although the role for cholesterol in macropinosome formation is not known, methods which deplete cellular cholesterol inhibit macropinocytosis (Grimmer *et al.*, 2002).

Other Necessary Molecules

Dynamin

Dynamin facilitates scission of clathrin-coated pits into coated vesicles (Schmid and Frolov, 2011). Its contribution to macropinocytosis is unsettled. Inhibition of dynamin 2 function inhibits macropinosome formation (Liu *et al.*, 2008). However, Cao *et al.*, found that dynamin was not required for macropinocytosis (Cao *et al.*, 2007). Dynamin 2 interacts with cortactin and Arp2/3 in the formation of CDR (Krueger *et al.*, 2003).

SNX proteins

The nascent macropinosome transforms from an irregular profile to a more rounded morphology with thin tubulovesicular extensions radiating along cytoplasmic microtubules. These extensions are formed by SNX-family proteins, which are regulated by PtdIns3P or PtdIns(3,4)P₂. The morphological changes may increase hydrostatic pressure inside the macropinosome (Kerr *et al.*, 2006). SNX1, 5, 9, 18, and 33 affect macropinosome formation or maturation (Lim *et al.*, 2012; Wang *et al.*, 2010).

Microtubules

Microtubule-depolymerizing drugs inhibit macropinocytosis by unknown mechanisms (Racoosin and Swanson, 1992; Swanson *et al.*, 1987). Conversely, microtubule destabilization promotes macropinosome formation. Histone deacetylase-6 (HDAC6) removes acetyl groups from tubulin in microtubules, leading to microtubule destabilization. HDAC6 localizes to ruffles and macropinosomes and stimulates macropinocytosis (Gao *et al.*, 2007). Moreover, a *Shigella* effector protein stimulates macropinocytosis for bacteria entry into host cells using an effector protein which destabilizes microtubules (Yoshida *et al.*, 2002). This suggests that dynamic microtubules are required for macropinosome formation.

The Organization of Cytoplasm for Macropinosome Formation

The movements of macropinosome formation and maturation are coordinated by interrelated networks of PIs, GTPases, GEFs, GAPs, and effector proteins which appear at distinct stages of macropinosome formation. A cascade of chemical reactions begins shortly after cell surface ruffles become circular ruffles (a.k.a. ruffle closure) (Maekawa *et al.*, 2014; Yoshida *et al.*, 2009). A spike of PtdIns(4,5)P₂ is followed by transient peaks of several signals, including PtdIns(3,4,5)P₃, Rac1 and DAG, which are terminated before the cup closes into a macropinosome. This is followed by a transient spike in PtdIns(3,4)P₂, which coincides with cup closure, then a slower peak comprised of PtdIns3P, Rab5a, Ras, and PKC α (Welliver and Swanson, 2012). This sequence of signals occurs during the formation of each macropinosome, and at different times after growth factor stimulation, which suggests that signaling is a self-organized cascade of chemical reactions guided by cell surface structure. Lateral diffusion of fluorescent proteins tethered to the inner leaflet of macrophage plasma membrane is restricted to macropinocytic cups following ruffle closure and prior to scission of the macropinosome from plasma membrane into a discrete intracellular vacuole (a.k.a. cup closure) (Welliver *et al.*, 2011). Restricted diffusion in subdomains of plasma membrane may facilitate signal amplification and the signal transitions necessary for macropinosome formation.

Experimental Diagnosis

Morphology

The clearest diagnostic method for detecting macropinocytosis is to observe macropinosomes directly. Time-lapse, light microscopy can show macropinosome formation in regions of cell surface ruffling or blebbing (Figure 1(a); Video 2). Alternatively, macropinosomes can be fluorescently labeled by brief, 3- to 15-minute incubation in medium containing a fluid-phase tracer, followed by brief rinsing in unlabeled medium, fixation and observation by light and fluorescence microscopy. Large, phase-bright, fluorescent macropinosomes can be counted directly (Commisso *et al.*, 2014). Fluorescein dextran (Fdx) and Lucifer yellow are nearly ideal fluorescent probes, as they do not adsorb to cell surfaces or induce endocytosis, and remain undegraded and membrane-impermeant for some time after uptake (Berthiaume *et al.*, 1995; Swanson *et al.*, 1985). Larger macromolecules enter macropinosomes more efficiently than smaller endocytic vesicles (Araki *et al.*, 1996; Berthiaume *et al.*, 1995); thus, high molecular weight Fdx (e.g., avg. mol. wt. 70 000) can be used to bias endocytic compartment labeling for macropinosomes.

Actin and Myosin

Agents which selectively inhibit actin polymerization, depolymerization, or the contractile activities of myosin, can inhibit macropinocytosis. Cytochalasin D and E, which inhibit actin filament turnover (Cooper, 1987), latrunculin A or B,

which depolymerize actin filaments (Coue *et al.*, 1987) and jasplakinolide, which inhibits actin depolymerization (Bubb *et al.*, 1994), all inhibit macropinocytosis. Macropinocytosis is also inhibited by the myosin light chain kinase inhibitor ML-7 (Araki *et al.*, 2003) and by the myosin II inhibitor blebbistatin (Jiang *et al.*, 2010; Shu *et al.*, 2005).

Amiloride

The sodium/proton antiport inhibitor amiloride and the related compounds ethyl isopropyl amiloride (EIPA) and HOE-694 inhibit macropinocytosis selectively (Koivusalo *et al.*, 2010; West *et al.*, 1989). They lower cytoplasmic pH near the plasma membrane; consequently inhibiting membrane-proximal cytoskeletal regulatory proteins (Koivusalo *et al.*, 2010). The inhibition of endocytosis by amiloride remains the most common diagnostic test for macropinocytosis.

Other Inhibitors

Additional inhibitors of macropinocytosis include wortmannin and LY294002 (Araki *et al.*, 1996), which inhibit PI 3-kinase, and rottlerin (Sarkar *et al.*, 2005), which works through unknown mechanisms. The PI 3-kinase inhibitor 3-methyladenine does not prevent macropinosome formation, but inhibits macropinosome maturation (Araki *et al.*, 2006).

Macropinocytosis and related engulfment activities have been indirectly linked with metabolic regulation. Uptake of Flock House virus by macropinocytosis requires AMP kinase, whose activity is regulated by cellular energy levels (Konradtowitz *et al.*, 2013). Macropinocytosis can be inhibited by rapamycin (Hackstein *et al.*, 2002), an inhibitor of the metabolic regulatory protein mTOR.

Dysregulation

Cancer

Oncogenic K-Ras, H-Ras, v-Src, and Abi1 all increase rates of macropinocytosis (Bar-Sagi and Feramisco, 1986; Dubielecka *et al.*, 2010; Veithen *et al.*, 1996). Unregulated macropinocytosis may contribute to unrestrained growth of tumor cells (Commisso *et al.*, 2013).

Bacterial Invasion

Various pathogenic bacteria are engulfed by host cells, including otherwise non-phagocytic cells, by triggering macropinocytosis. This was first described for *Salmonella Typhi* var. Typhimurium (Alpuche-Aranda *et al.*, 1994; Francis *et al.*, 1993) and later observed for *Legionella pneumophila* (Watarai *et al.*, 2001), *Brucella abortus* (Watarai *et al.*, 2002), *Neisseria gonorrhoeae* (Zenni *et al.*, 2000), and *Shigella flexneri* (Ogawa *et al.*, 2008).

Virus Infection

Eukaryotic cells internalize viruses by different routes, including clathrin-mediated endocytosis, caveolin-mediated endocytosis and macropinocytosis (Mercer *et al.*, 2010).

Vaccinia virus infects by triggering macropinocytosis (Mercer and Helenius, 2008). Other viruses which infect via macropinocytosis include echovirus 1 (Krieger *et al.*, 2013), adenovirus (Amstutz *et al.*, 2008), Flock House virus (Nakase *et al.*, 2009), Ebola virus (Saeed *et al.*, 2010), and HIV-1 (Liu *et al.*, 2002; Marechal *et al.*, 2001).

Therapeutics

Drug Delivery

Cell-penetrating peptides and cationic liposomes have been developed for therapeutic delivery of macromolecules into the cytoplasm. Many of these delivery vehicles enter cytoplasm from inside macropinosomes (Cardarelli *et al.*, 2012; Nakase *et al.*, 2007).

Cancer

Methods and reagents for manipulating macropinocytosis may find use as anti-cancer therapies. Stimulation of macropinocytosis in neuroblastoma cells by transfection with Ras or Rac causes an unusual form of cell death called methuosis, in which macropinosomes form, enlarge and rupture (Maltese and Overmeyer, 2014; Overmeyer *et al.*, 2008). Stimulation of methuosis may provide a therapeutic strategy for selectively killing these cells (Kitambi *et al.*, 2014).

See also: Intracellular Infectiology: Cell Processes: Phagocytosis

References

- Alpuche-Aranda, C.M., Racoosin, E.L., Swanson, J.A., Miller, S.I., 1994. *Salmonella* stimulate macrophage macropinocytosis and persist within spacious phagosomes. *Journal of Experimental Medicine* 179, 601–608.
- Amstutz, B., Gastaldelli, M., Kalin, S., *et al.*, 2008. Subversion of CIBP1-controlled macropinocytosis by human adenovirus serotype 3. *EMBO Journal* 27, 956–969.
- Amyere, M., Mettlen, M., Van Der Smissen, P., *et al.*, 2002. Origin, originality, functions, subversions and molecular signaling of macropinocytosis. *International Journal of Medical Microbiology: IJMM* 291, 487–494.
- Amyere, M., Payrastra, B., Krause, U., *et al.*, 2000. Constitutive macropinocytosis in oncogene-transformed fibroblasts depends on sequential permanent activation phosphoinositide 3-kinase and phospholipase C. *Molecular Biology of the Cell* 11, 3453–3467.
- Araki, N., Hamasaki, M., Egami, Y., Hatae, T., 2006. Effect of 3-methyladenine on the fusion process of macropinosomes in EGF-stimulated A431 cells. *Cell Structure and Function* 31, 145–157.
- Araki, N., Hatae, T., Furukawa, A., Swanson, J.A., 2003. Phosphoinositide-3-kinase-independent contractile activities associated with Fcγ-receptor-mediated phagocytosis and macropinocytosis in macrophages. *Journal of Cell Science* 116, 247–257.
- Araki, N., Hatae, T., Yamada, T., Hirohashi, S., 2000. Actinin-4 is preferentially involved in circular ruffling and macropinocytosis in mouse macrophages: Analysis by fluorescence ratio imaging. *Journal of Cell Science* 113, 3329–3340.
- Araki, N., Johnson, M.T., Swanson, J.A., 1996. A role for phosphoinositide 3-kinase in the completion of macropinocytosis and phagocytosis in macrophages. *Journal of Cell Biology* 135, 1249–1260.
- Balaji, K., Mooser, C., Janson, C.M., *et al.*, 2012. RIN1 orchestrates the activation of RAB5 GTPases and ABL tyrosine kinases to determine the fate of EGFR. *Journal of Cell Science* 125, 5887–5896.

- Bar-Sagi, D., Feramisco, J.R., 1986. Induction of membrane ruffling and fluid-phase pinocytosis in quiescent fibroblasts by *ras* proteins. *Science* 233, 1061–1066.
- Berthiaume, E.P., Medina, C., Swanson, J.A., 1995. Molecular size-fractionation during endocytosis in macrophages. *Journal of Cell Biology* 129, 989–998.
- Besterman, J.M., Airhart, J.A., Woodworth, R.C., Low, R.B., 1981. Exocytosis of pinocytosed fluid in cultured cells: Kinetic evidence for rapid turnover and compartmentation. *Journal of Cell Biology* 91, 716–727.
- Bohdanowicz, M., Balkin, D.M., De Camilli, P., Grinstein, S., 2012. Recruitment of OCR1 and Inpp5B to phagosomes by Rab5 and APPL1 depletes phosphoinositides and attenuates Akt signaling. *Molecular Biology of the Cell* 23, 176–187.
- Bohdanowicz, M., Grinstein, S., 2013. Role of phospholipids in endocytosis, phagocytosis, and macropinocytosis. *Physiological Reviews* 93, 69–106.
- Bohdanowicz, M., Schlamm, D., Hermansson, M., *et al.*, 2013. Phosphatidic acid is required for the constitutive ruffling and macropinocytosis of phagocytes. *Molecular Biology of the Cell* 24 (1700–1712), S1701–S1707.
- Bowers, B., Olszewski, T.E., 1972. Pinocytosis in *Acanthamoeba castellanii*. Kinetics and morphology. *Journal of Cell Biology* 53, 681–694.
- Brunk, U., Schellens, J., Westermarck, B., 1976. Influence of epidermal growth factor on ruffling activity, pinocytosis and proliferation of cultivated human glial cells. *Experimental Cell Research* 103, 295–302.
- Bubb, M.R., Senderowicz, A.M., Sausville, E.A., Duncan, K.L., Korn, E.D., 1994. Jaspaklinolide, a cytotoxic natural product, induces actin polymerization and competitively inhibits the binding of phalloidin to F-actin. *Journal of Biological Chemistry* 269, 14869–14871.
- Cao, H., Chen, J., Awoniyi, M., Henley, J.R., McNiven, M.A., 2007. Dynamin 2 mediates fluid-phase micropinocytosis in epithelial cells. *Journal of Cell Science* 120, 4167–4177.
- Cardarelli, F., Pozzi, D., Bifone, A., Marchini, C., Caracciolo, G., 2012. Cholesterol-dependent macropinocytosis and endosomal escape control the transfection efficiency of lipoplexes in CHO living cells. *Molecular Pharmaceutics* 9, 334–340.
- Chapman-Andresen, C., 1971. Biology of the large amoebae. *Annual Review of Microbiology* 25, 27–48.
- Chen, L.-M., Hobbie, S., Galán, J.E., 1996. Requirement of CDC42 for *Salmonella*-induced cytoskeletal and nuclear responses. *Science* 274, 2115–2118.
- Cohn, Z.A., Ehrenreich, B.A., 1969. The uptake, storage and intracellular hydrolysis of carbohydrates by macrophages. *Journal of Experimental Medicine* 129, 201–225.
- Commisso, C., Davidson, S.M., Soydaner-Azeloglu, R.G., *et al.*, 2013. Macropinocytosis of protein is an amino acid supply route in Ras-transformed cells. *Nature* 497, 633–637.
- Commisso, C., Flinn, R.J., Bar-Sagi, D., 2014. Determining the macropinocytic index of cells through a quantitative image-based assay. *Nature Protocols* 9, 182–192.
- Cooper, J.A., 1987. Effects of cytochalasin and phalloidin on actin. *Journal of Cell Biology* 105, 1473–1478.
- Coue, M., Brenner, S.L., Spector, I., Korn, E.D., 1987. Inhibition of actin polymerization by latrunculin A. *FEBS Letters* 213, 316–318.
- Davies, P.F., Ross, R., 1978. Mediation of pinocytosis in cultured arterial smooth muscle and endothelial cells by platelet-derived growth factor. *Journal of Cell Biology* 79, 663–671.
- Dharmawardhane, S., Sanders, L.C., Martin, S.S., Daniels, R.H., Bokoch, G.M., 1997. Localization of p21-activated kinase-1 (PAK-1) to pinocytic vesicles and cortical actin structures in stimulated cells. *Journal of Cell Biology* 138, 1265–1278.
- Dharmawardhane, S., Schurmann, A., Sells, M.A., *et al.*, 2000. Regulation of macropinocytosis by p21-activated kinase-1. *Molecular Biology of the Cell* 11, 3341–3352.
- Dowrick, P., Kenworthy, P., McCann, B., Warn, R., 1993. Circular ruffle formation and closure lead to macropinocytosis in hepatocyte growth factor/scatter factor-treated cells. *European Journal of Cell Biology* 61, 44–53.
- Dubielecka, P.M., Machida, K., Xiong, X., *et al.*, 2010. Abi1/Hsh3bp1 pY213 links Abi kinase signaling to p85 regulatory subunit of PI-3 kinase in regulation of macropinocytosis in LNCaP cells. *FEBS Letters* 584, 3279–3286.
- Edgar, A.J., Bennett, J.P., 1997. Circular ruffle formation in rat basophilic leukemia cells in response to antigen stimulation. *European Journal of Cell Biology* 73, 132–140.
- Egami, Y., Araki, N., 2009. Dynamic changes in the spatiotemporal localization of Rab21 in live RAW264 cells during macropinocytosis. *PLoS One* 4, e6689.
- Egami, Y., Araki, N., 2012. Spatiotemporal localization of Rab20 in live RAW264 macrophages during macropinocytosis. *Acta Histochemica et Cytochemica* 45, 317–323.
- Ehrenreich, B.A., Cohn, Z.A., 1967. The uptake and digestion of iodinated human serum albumin by macrophages *in vitro*. *Journal of Experimental Medicine* 126, 941–958.
- Ellerbroek, S.M., Wennerberg, K., Arthur, W.T., *et al.*, 2004. SGEF, a RhoG guanine nucleotide exchange factor that stimulates macropinocytosis. *Molecular Biology of the Cell* 15, 3309–3319.
- Falasca, M., Logan, S.K., Lehto, V.P., *et al.*, 1998. Activation of phospholipase C γ by PI 3-kinase-induced PH domain-mediated membrane targeting. *EMBO Journal* 17, 414–422.
- Feliciano, W.D., Yoshida, S., Straight, S.W., Swanson, J.A., 2011. Coordination of the Rab5 cycle on macropinosomes. *Traffic* 12, 1911–1922.
- Francis, C.L., Ryan, T.A., Jones, B.D., Smith, S.J., Falkow, S., 1993. Ruffles induced by *Salmonella* and other stimuli direct macropinocytosis of bacteria. *Nature* 364, 639–642.
- Fujii, M., Kawai, K., Egami, Y., Araki, N., 2013. Dissecting the roles of Rac1 activation and deactivation in macropinocytosis using microscopic photo-manipulation. *Scientific Reports* 3, 2385.
- Gao, Y.S., Hubbert, C.C., Lu, J., *et al.*, 2007. Histone deacetylase 6 regulates growth factor-induced actin remodeling and endocytosis. *Molecular and Cellular Biology* 27, 8637–8647.
- Garrett, W.S., Chen, L.-M., Kroschewski, R., *et al.*, 2000. Developmental control of endocytosis in dendritic cells by Cdc42. *Cell* 102, 325–334.
- Grimmer, S., van Deurs, B., Sandvig, K., 2002. Membrane ruffling and macropinocytosis in A431 cells require cholesterol. *Journal of Cell Science* 115, 2953–2962.
- Grogan, A., Reeves, E., Keep, N., *et al.*, 1997. Cytosolic p135 proteins interact with and regulate the assembly of coronin in neutrophils. *Journal of Cell Science* 110 (Pt 24), 3071–3081.
- Hackstein, H., Taner, T., Logar, A.J., Thomson, A.W., 2002. Rapamycin inhibits macropinocytosis and mannose receptor-mediated endocytosis by bone marrow-derived dendritic cells. *Blood* 100, 1084–1087.
- Haga, Y., Miwa, N., Jahangeer, S., Okada, T., Nakamura, S., 2009. CtBP1/BARS is an activator of phospholipase D1 necessary for agonist-induced macropinocytosis. *EMBO Journal* 28, 1197–1207.
- Hamasaki, M., Araki, N., Hatae, T., 2004. Association of early endosomal autoantigen 1 with macropinocytosis in EGF-stimulated A431 cells. The Anatomical Record Part A, Discoveries in Molecular, Cellular, and Evolutionary Biology 277, 298–306.
- Hasegawa, J., Tokuda, E., Tenno, T., *et al.*, 2011. SH3YL1 regulates dorsal ruffle formation by a novel phosphoinositide-binding domain. *Journal of Cell Biology* 193, 901–916.
- Hewlett, L.J., Prescott, A.R., Watts, C., 1994. The coated pit and macropinocytic pathways serve distinct endosome populations. *Journal of Cell Biology* 124, 689–703.
- Hoffmann, P.R., deCathelineau, J.M., Ogden, C.A., *et al.*, 2001. Phosphatidylserine (PS) induces PS receptor-mediated macropinocytosis and promotes clearance of apoptotic cells. *Journal of Cell Biology* 155, 649–659.
- Honda, A., Nogami, M., Yokozeki, T., *et al.*, 1999. Phosphatidylinositol 4-phosphate 5-kinase α is a downstream effector of the small G protein ARF6 in membrane ruffle formation. *Cell* 99, 521–532.
- Hoon, J.L., Wong, W.K., Koh, C.G., 2012. Functions and regulation of circular dorsal ruffles. *Molecular and Cellular Biology* 32, 4246–4257.
- Itoh, T., Hasegawa, J., 2013. Mechanistic insights into the regulation of circular dorsal ruffle formation. *Journal of Biochemistry* 153, 21–29.
- Jiang, J., Kolpak, A.L., Bao, Z.Z., 2010. Myosin IIB isoform plays an essential role in the formation of two distinct types of macropinosomes. *Cytoskeleton* 67, 32–42.
- Kasahara, K., Nakayama, Y., Sato, I., *et al.*, 2007. Role of Src-family kinases in formation and trafficking of macropinosomes. *Journal of Cellular Physiology* 211, 220–232.
- Keller, H.U., 1990. Diacylglycerols and PMA are particularly effective stimulators of fluid pinocytosis in human neutrophils. *Journal of Cellular Physiology* 145, 465–471.
- Kerr, M.C., Lindsay, M.R., Luettner, R., *et al.*, 2006. Visualisation of macropinosome maturation by the recruitment of sorting nexins. *Journal of Cell Science* 119, 3967–3980.
- Kitambi, S.S., Toledo, E.M., Usoskin, D., *et al.*, 2014. Vulnerability of glioblastoma cells to catastrophic vacuolization and death induced by a small molecule. *Cell* 157, 313–328.
- Knight, K.R., Vairo, G., Hamilton, J.A., 1992. Regulation of pinocytosis in murine macrophages by colony-stimulating factors and other agents. *Journal of Leukocyte Biology* 51, 350–359.

- Koivusalo, M., Welch, C., Hayashi, H., *et al.*, 2010. Amiloride inhibits macropinocytosis by lowering submembranous pH and preventing Rac1 and Cdc42 signaling. *Journal of Cell Biology* 188, 547–563.
- Kolpak, A.L., Jiang, J., Guo, D., *et al.*, 2009. Negative guidance factor-induced macropinocytosis in the growth cone plays a critical role in repulsive axon turning. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29, 10488–10498.
- Kondratowicz, A.S., Hunt, C.L., Davey, R.A., Cherry, S., Maury, W.J., 2013. AMP-activated protein kinase is required for the macropinocytic internalization of ebolavirus. *Journal of Virology* 87, 746–755.
- Krieger, S.E., Kim, C., Zhang, L., Marjomaki, V., Bergelson, J.M., 2013. Echovirus 1 entry into polarized Caco-2 cells depends on dynamin, cholesterol, and cellular factors associated with macropinocytosis. *Journal of Virology* 87, 8884–8895.
- Krueger, E.W., Orth, J.D., Cao, H., McNiven, M.A., 2003. A dynamin-cortactin-Arp2/3 complex mediates actin reorganization in growth factor-stimulated cells. *Molecular Biology of the Cell* 14, 1085–1096.
- Kruth, H.S., 2013. Fluid-phase pinocytosis of LDL by macrophages: A novel target to reduce macrophage cholesterol accumulation in atherosclerotic lesions. *Current Pharmaceutical Design* 19, 5865–5872.
- Kumari, S., Mg, S., Mayor, S., 2010. Endocytosis unplugged: Multiple ways to enter the cell. *Cell Research* 20, 256–275.
- Lanzetti, L., Palamidessi, A., Arecas, L., Scita, G., Di Fiore, P.P., 2004. Rab5 is a signaling GTPase involved in actin remodelling by receptor tyrosine kinases. *Nature* 429, 309–314.
- Lewis, W.H., 1931. Pinocytosis. *Bulletin of the Johns Hopkins Hospital* 49, 17–27.
- Lewis, W.H., 1937. Pinocytosis by malignant cells. *American Journal of Cancer* 29, 666.
- Li, G., D'Souza-Schorey, C., Barbieri, M.A., Cooper, J.A., Stahl, P.D., 1997. Uncoupling of membrane ruffling and pinocytosis during Ras signal transduction. *Journal of Biological Chemistry* 272, 10337–10340.
- Liberali, P., Kakkonen, E., Turacchio, G., *et al.*, 2008. The closure of Pak1-dependent macropinosomes requires the phosphorylation of CIBP1/BARS. *EMBO Journal* 27, 970–981.
- Lim, J.P., Teasdale, R.D., Gleeson, P.A., 2012. SNX5 is essential for efficient macropinocytosis and antigen processing in primary macrophages. *Biology Open* 1, 904–914.
- Liu, N.Q., Lossinsky, A.S., Popik, W., *et al.*, 2002. Human immunodeficiency virus type 1 enters brain microvascular endothelia by macropinocytosis dependent on lipid rafts and the mitogen-activated protein kinase signaling pathway. *Journal of Virology* 76, 6689–6700.
- Liu, Y.W., Surka, M.C., Schroeter, T., Lukiychuk, V., Schmid, S.L., 2008. Isoform and splice-variant specific functions of dynamin-2 revealed by analysis of conditional knock-out cells. *Molecular Biology of the Cell* 19, 5347–5359.
- Maekawa, M., Terasaka, S., Mochizuki, Y., *et al.*, 2014. Sequential breakdown of 3-phosphorylated phosphoinositides is essential for the completion of macropinocytosis. *Proceedings of the National Academy of Sciences of the United States of America* 111, E978–E987.
- Maltese, W.A., Overmeyer, J.H., 2014. Methuosis: Nonapoptotic cell death associated with vacuolization of macropinosome and endosome compartments. *American Journal of Pathology* 184, 1630–1642.
- Marechal, V., Prevost, M.C., Petit, C., *et al.*, 2001. Human immunodeficiency virus type 1 entry into macrophages mediated by macropinocytosis. *Journal of Virology* 75, 11166–11177.
- Mercer, J., Helenius, A., 2008. Vaccinia virus uses macropinocytosis and apoptotic mimicry to enter host cells. *Science* 320, 531–535.
- Mercer, J., Helenius, A., 2009. Virus entry by macropinocytosis. *Nature Cell Biology* 11, 510–520.
- Mercer, J., Schelhaas, M., Helenius, A., 2010. Virus entry by endocytosis. *Annual Review of Biochemistry* 79, 803–833.
- Mettlen, M., Platek, A., Van Der Smitten, P., *et al.*, 2006. Src triggers circular ruffling and macropinocytosis at the apical surface of polarized MDCK cells. *Traffic* 7, 589–603.
- Miyata, Y., Nishida, E., Koyasu, S., Yahara, I., Sakai, H., 1989. Protein kinase C-dependent and -independent pathways in the growth factor-induced cytoskeletal reorganization. *Journal of Biological Chemistry* 264, 15565–15568.
- Nakase, I., Hirose, H., Tanaka, G., *et al.*, 2009. Cell-surface accumulation of flock house virus-derived peptide leads to efficient internalization via macropinocytosis. *Molecular Therapy: The Journal of the American Society of Gene Therapy* 17, 1868–1876.
- Nakase, I., Tadokoro, A., Kawabata, N., *et al.*, 2007. Interaction of arginine-rich peptides with membrane-associated proteoglycans is crucial for induction of actin organization and macropinocytosis. *Biochemistry* 46, 492–501.
- Norbury, C.C., Hewlett, L.J., Prescott, A.R., Shastri, N., Watts, C., 1995. Class I MHC presentation of exogenous soluble antigen via macropinocytosis in bone marrow macrophages. *Immunity* 3, 783–791.
- Ogawa, M., Handa, Y., Ashida, H., Suzuki, M., Sasakawa, C., 2008. The versatility of *Shigella* effectors. *Nature Reviews Microbiology* 6, 11–16.
- Otomo, A., Kunita, R., Suzuki-Utsunomiya, K., Ikeda, J.E., Hadano, S., 2011. Defective relocalization of ALS2/alsin missense mutants to Rac1-induced macropinosomes accounts for loss of their cellular function and leads to disturbed amphisome formation. *FEBS Letters* 585, 730–736.
- Overmeyer, J.H., Kaul, A., Johnson, E.E., Maltese, W.A., 2008. Active ras triggers death in glioblastoma cells through hyperstimulation of macropinocytosis. *Molecular Cancer Research: MCR* 6, 965–977.
- Platek, A., Mettlen, M., Camby, I., *et al.*, 2004. v-Src accelerates spontaneous motility via phosphoinositide 3-kinase, phospholipase C and phospholipase D, but abrogates chemotaxis in Rat-1 and MDCK cells. *Journal of Cell Science* 117, 4849–4861.
- Porat-Shliom, N., Kloog, Y., Donaldson, J.G., 2008. A unique platform for H-Ras signaling involving clathrin-independent endocytosis. *Molecular Biology of the Cell* 19, 765–775.
- Racoosin, E.L., Swanson, J.A., 1989. Macrophage colony stimulating factor (rM-CSF) stimulates pinocytosis in bone marrow-derived macrophages. *Journal of Experimental Medicine* 170, 1635–1648.
- Racoosin, E.L., Swanson, J.A., 1992. M-CSF-induced macropinocytosis increases solute endocytosis but not receptor-mediated endocytosis in mouse macrophages. *Journal of Cell Science* 102, 867–880.
- Racoosin, E.L., Swanson, J.A., 1993. Macropinosome maturation and fusion with tubular lysosomes in macrophages. *Journal of Cell Biology* 121, 1011–1020.
- Ridley, A.J., Paterson, H.F., Johnston, C.L., Diekmann, D., Hall, A., 1992. The small GTP-binding protein rac regulates growth factor-induced membrane ruffling. *Cell* 70, 401–4020.
- Saeed, M.F., Kolokoltsov, A.A., Albrecht, T., Davey, R.A., 2010. Cellular entry of ebola virus involves uptake by a macropinocytosis-like mechanism and subsequent trafficking through early and late endosomes. *PLoS Pathogens* 6, e1001110.
- Sallusto, F., Cella, M., Danieli, C., Lanzavecchia, A., 1995. Dendritic cells use macropinocytosis and the mannose receptor to concentrate macromolecules in the major histocompatibility complex class II compartment: Downregulation by cytokines and bacterial products. *Journal of Experimental Medicine* 182, 389–400.
- Sanematsu, F., Nishikimi, A., Watanabe, M., *et al.*, 2013. Phosphatidic acid-dependent recruitment and function of the Rac activator DOCK1 during dorsal ruffle formation. *Journal of Biological Chemistry* 288, 8092–8100.
- Sarkar, K., Kruhlak, M.J., Erlandsen, S.L., Shaw, S., 2005. Selective inhibition by rottlerin of macropinocytosis in monocyte-derived dendritic cells. *Immunology* 116, 513–524.
- Schafer, D.A., D'Souza-Schorey, C., Cooper, J.A., 2000. Actin assembly at membranes controlled by ARF6. *Traffic* 1, 892–903.
- Schmid, S.L., Frolov, V.A., 2011. Dynamin: Functional design of a membrane fission catalyst. *Annual Review of Cell and Developmental Biology* 27, 79–105.
- Schnatwinkel, C., Christoforidis, S., Lindsay, M.R., *et al.*, 2004. The Rab5 effector Rabankyrin-5 regulates and coordinates different endocytic mechanisms. *PLoS Biology* 2, E261.
- Shu, S., Liu, X., Korn, E.D., 2005. Blebbistatin and blebbistatin-inactivated myosin II inhibit myosin II-independent processes in Dictyostelium. *Proceedings of the National Academy of Sciences of the United States of America* 102, 1472–1477.
- Steinman, R.M., Brodie, S.E., Cohn, Z.A., 1976. Membrane flow during pinocytosis. A stereological analysis. *Journal of Cell Biology* 68, 665–687.
- Sun, P., Yamamoto, H., Suetsugu, S., *et al.*, 2003. Small GTPase Rac/Rab34 is associated with membrane ruffles and macropinosomes and promotes macropinosome formation. *Journal of Biological Chemistry* 278, 4063–4071.
- Swanson, J.A., 1989. Phorbol esters stimulate macropinocytosis and solute flow through macrophages. *Journal of Cell Science* 94, 135–142.
- Swanson, J.A., Burke, E., Silverstein, S.C., 1987. Tubular lysosomes accompany stimulated pinocytosis in macrophages. *Journal of Cell Biology* 104, 1217–1222.
- Swanson, J.A., Johnson, M.T., Beningo, K., *et al.*, 1999. A contractile activity that closes phagosomes in macrophages. *Journal of Cell Science* 112, 307–316.
- Swanson, J.A., Yirinec, B.D., Silverstein, S.C., 1985. Phorbol esters and horseradish peroxidase stimulate pinocytosis and redirect flow of pinocytosed fluid through macrophages. *Journal of Cell Biology* 100, 851–859.
- Tall, G.G., Barbieri, M.A., Stahl, P.D., Horazdovsky, B.F., 2001. Ras-activated endocytosis is mediated by the Rab5 guanine nucleotide exchange activity of RIN1. *Developmental Cell* 1, 73–82.

- Tamura, N., Hazeki, K., Okazaki, N., *et al.*, 2009. Specific role of phosphoinositide 3-kinase p110 α in the regulation of phagocytosis and pinocytosis in macrophages. *Biochemical Journal* 423, 99–108.
- Veithen, A., Cupers, P., Baudhuin, P., Courtoy, P.J., 1996. v-Src induces constitutive macropinocytosis in rat fibroblasts. *Journal of Cell Science* 109, 2005–2012.
- Veracini, L., Franco, M., Boueux, A., *et al.*, 2006. Two distinct pools of Src family tyrosine kinases regulate PDGF-induced DNA synthesis and actin dorsal ruffles. *Journal of Cell Science* 119, 2921–2934.
- Wang, J.T., Kerr, M.C., Karunaratne, S., *et al.*, 2010. The SNX-PX-BAR family in macropinocytosis: The regulation of macropinosome formation by SNX-PX-BAR proteins. *PLoS One* 5, e13763.
- Watarai, M., Derre, I., Kirby, J., *et al.*, 2001. *Legionella pneumophila* is internalized by a macropinocytotic uptake pathway controlled by the Dot/Icm system and the mouse *Lgn1* locus. *Journal of Experimental Medicine* 194, 1081–1095.
- Watarai, M., Makino, S., Fujii, Y., Okamoto, K., Shirahata, T., 2002. Modulation of Brucella-induced macropinocytosis by lipid rafts mediates intracellular replication. *Cellular Microbiology* 4, 341–355.
- Welliver, T.P., Chang, S.L., Linderman, J.J., Swanson, J.A., 2011. Ruffles limit diffusion in the plasma membrane during macropinosome formation. *Journal of Cell Science* 124, 4106–4114.
- Welliver, T.P., Swanson, J.A., 2012. A growth factor signaling cascade confined to circular ruffles in macrophages. *Biology Open* 1, 754–760.
- West, M.A., Bretscher, M.S., Watts, C., 1989. Distinct endocytic pathways in epidermal growth factor-stimulated human carcinoma A431 cells. *Journal of Cell Biology* 109, 2731–2739.
- Yoshida, S., Hoppe, A.D., Araki, N., Swanson, J.A., 2009. Sequential signaling in plasma-membrane domains during macropinosome formation in macrophages. *Journal of Cell Science* 122, 3250–3261.
- Yoshida, S., Katayama, E., Kuwae, A., *et al.*, 2002. *Shigella* deliver an effector protein to trigger host microtubule destabilization, which promotes Rac1 activity and efficient bacterial internalization. *EMBO Journal* 21, 2923–2935.
- Zenni, M.K., Giardina, P.C., Harvey, H.A., *et al.*, 2000. Macropinocytosis as a mechanism of entry into primary human urethral epithelial cells by *Neisseria gonorrhoeae*. *Infection and Immunity* 68, 1696–1699.

Microbicidal Mechanisms

L-AH Allen, University of Iowa, Iowa City, IA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Apoptosis A noninflammatory form of programmed cell death.

Inflammasome One of several multiprotein complexes in the macrophage cytosol that contain caspase-1 and are required for secretion of proinflammatory cytokines such as IL-1 β .

Inflammation A rapid response to infection that is driven by macrophage cytokines and is characterized by tissue swelling, heat, redness, and pain.

NADPH oxidase An enzyme complex in phagocytes that is essential for host defense and catalyzes the conversion of oxygen into superoxide anions.

Neutrophil extracellular traps (NETs) NETs are networks of DNA, histones, and granule proteins released by activated neutrophils that can ensnare extracellular microbes.

Pathogen The subset of all microorganisms that have the capacity to cause disease.

Phagocytosis A receptor- and actin-driven endocytic process that allows macrophages, neutrophils, and dendritic cells to ingest whole bacteria and fungi.

Phagosome maturation The process of converting a nascent phagosome into a lysosome-like compartment for optimal microbe killing.

Pyroptosis A proinflammatory form of macrophage death that is triggered by inflammasome activation.

ROS Reactive oxygen species include superoxide anions, hydrogen peroxide, hypochlorous acid, and other toxic oxygen radicals and derivatives that are critical for host defense.

Phagocytes and Phagocytosis

Phagocytes are certain types of white blood cells (leukocytes) that are unique in their ability to bind and engulf large particles such as bacteria, fungi, and parasites. The term phagocyte means 'cells that can eat,' and the two main types of phagocytes are macrophages and neutrophils. Small numbers of macrophages are present in all body tissues. In contrast, neutrophils, which are also called polymorphonuclear leukocytes, reside in the bloodstream and are rapidly recruited from this locale to sites of infection or tissue damage. Phagocytes and the process of phagocytosis were originally described by Elie Metchnikoff in the late nineteenth century, and these early studies also revealed unique morphological features of neutrophils such as their lobed nuclei and abundant cytoplasmic granules (reprinted in [Cavaillon, 2011](#)).

Mechanisms of Phagocytosis

Phagocytosis is a specialized type of endocytosis that is unique to phagocytes. At the molecular level, phagocytosis is defined as a receptor-mediated and F-actin-dependent mechanism for internalization of particles that are at least 0.5 μm in diameter ([Figure 1](#)). There are many receptors on macrophages and neutrophils that mediate phagocytosis ([Freeman and Grinstein, 2014](#)). Some of these receptors bind directly to microbe surface proteins or carbohydrates, and examples of receptors in this class are the macrophage mannose receptor, dectin-1, and scavenger receptors. Other receptors that mediate phagocytosis bind indirectly to microbes that have been pre-coated with molecules termed opsonins that enhance their ability to be engulfed. The process of coating microbes with opsonins is called opsonization, which means 'to prepare for eating.' The two major types of opsonins are antibodies of the

IgG class, and complement proteins C3b and C3bi. Receptors that bind microbes coated with IgG are called Fc γ receptors and include CD16, CD32, and CD64. The receptor that binds C3b is called complement receptor 1 (CR1, also called CD35), whereas both CR3 (also called CD11b/CD18, $\alpha_M\beta_2$ or Mac-1) and CR4 (CD11c/CD18, $\alpha_X\beta_2$) bind particles that have been opsonized with C3bi. In many instances, multiple receptors can participate in uptake of a given microorganism. However, to study the properties of individual receptors in isolation, research studies often use latex beads or other particles that

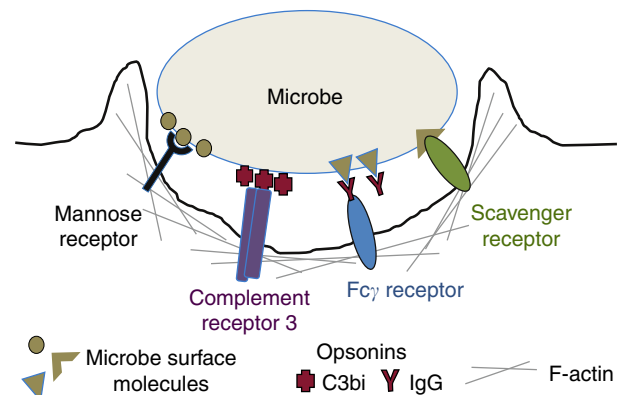


Figure 1 Receptors for phagocytosis. Phagocytosis of microbes is a receptor-dependent endocytic process that allows macrophages and neutrophils to ingest bacteria, fungi, and parasites. Some phagocytic receptors, such as the mannose receptor or scavenger receptors, bind directly to sugars, proteins, or other molecules on microbe surfaces. On the other hand, Fc γ receptors and complement receptor 3 (also called CR3, Mac-1, $\alpha_M\beta_2$) bind indirectly to microbes that have been opsonized (pre-coated) with IgG and C3bi, respectively. Regardless of the receptor engaged, extension of the plasma membrane around the microbe also requires local polymerization of actin filaments (F-actin).

are coated with ligands for individual receptor types, such as IgG.

Many of the fundamental principles of phagocytosis were defined by Silverstein and colleagues in the 1970s (Griffin *et al.*, 1975). These studies demonstrated that ligands for phagocytic receptors must be present over the entire surface of the microbe or particle being engulfed, since hemispherically opsonized targets were only partially ingested. They also showed that local actin polymerization provides the driving force for plasma membrane extension. Moreover, microbe binding and uptake could be distinguished, as binding was intact when phagocytes were incubated at 4 °C or in the presence of drugs that blocked actin polymerization, yet engulfment did not occur.

Subsequent studies have shown that many actin-binding proteins, guanosine triphosphatases (GTPases) and protein and lipid kinases are also recruited to forming phagosomes (Aderem and Underhill, 1999; Freeman and Grinstein, 2014). These molecules provide additional essential signals that regulate phagocytosis and also modulate the morphology of forming phagosomes. Thus, in some instances, phagocytes appear to reach out and grab their prey using long pseudopodia, whereas in other instances, the particles appear to sink into the cell surface (Allen and Aderem, 1996). Regardless of the specific receptor engaged, it is also noteworthy that macrophage surface area increases during phagocytosis, a finding which suggested a mechanism for plasma membrane expansion during particle engulfment that was confirmed by measurements of membrane capacitance (Holevinsky and Nelson, 1998). The results of additional studies indicate that multiple intracellular membrane compartments can be mobilized and contribute to this process, but as yet no one compartment or type of vesicle appears to be essential.

Phagocytic Killing

Central to effective immunity and host defense is the ability of phagocytes to rapidly and efficiently kill the microbes that they ingest. To achieve this, a multifaceted approach is utilized that includes production of toxic reactive oxygen species (ROS), sequestration of critical nutrients, and exposure to numerous hydrolytic enzymes and cationic peptides that directly damage and degrade microbes and microbe components. The efficiency of this process is enhanced by the phagosome itself, which allows the components of the killing arsenal to be concentrated in close proximity to the ingested microbe. The mechanisms that contribute to phagocytic killing are described here along with the associated process of phagosome maturation.

Oxygen-Dependent Killing

In general, the processes that contribute to phagocytic killing can be divided into two groups, those that require oxygen and those that do not. Three enzymes that are critical for oxidative host defense are the phagocyte nicotinamide adenine dinucleotide phosphate (NADPH) oxidase complex, myeloperoxidase (MPO), and inducible nitric oxide synthase (iNOS).

NADPH oxidase and MPO

The central player in oxidative host defense is the phagocyte NADPH oxidase (*phox*) complex which catalyzes the conversion of molecular oxygen into superoxide anions. Because toxic oxidants can damage host tissue as well as microbes, the activity of this enzyme is under tight spatial and temporal control. The enzyme itself consists of six subunits (Nauseef, 2004; DeLeo and Quinn, 1996). Two integral membrane components gp91^{phox} (also called Nox2) and p22^{phox} form a heterodimer called flavocytochrome b₅₅₈ that contains the catalytic core of the enzyme and localizes to the plasma membrane and early endosomes of macrophages, and the plasma membrane, specific granules and gelatinase granules of neutrophils (DeLeo and Quinn, 1996; Casbon *et al.*, 2009). The four other subunits are p47^{phox}, p67^{phox}, p40^{phox}, and the GTPase Rac, all of which are soluble proteins distributed throughout the cytosol.

In resting (unstimulated) phagocytes the NADPH oxidase complex is unassembled and inactive, with subunits segregated in different subcellular compartments (Figure 2). However, within 30 s of microbe binding flavocytochrome b₅₅₈ is recruited to forming phagosomes from the plasma membrane and intracellular stores, and becomes highly enriched in this locale (Allen *et al.*, 1999). Simultaneously, protein kinase C (PKC) phosphorylates multiple serine residues of p47^{phox} which triggers *en bloc* membrane translocation of p47^{phox}/p67^{phox}/p40^{phox} complexes, whereas active Rac translocates to the membrane independently (Allen *et al.*, 1999; DeLeo *et al.*, 1999; DeLeo and Quinn, 1996; Nauseef, 2004). The assembled, active NADPH oxidase complex catalyzes the conversion of molecular oxygen into superoxide anions, which accumulate in the phagosome lumen. Because large amounts of oxygen are consumed in this process, NADPH oxidase activation is often referred to as 'the respiratory burst.' Nonetheless, superoxide itself is only moderately toxic. Microbe killing is therefore enhanced by rapid dismutation of superoxide into hydrogen peroxide (H₂O₂). Moreover, neutrophils also contain MPO, an enzyme that catalyzes the conversion of H₂O₂ into highly toxic hypochlorous acid (HOCl), which kills microbes by chlorination of proteins, lipids, and DNA, and notably is also the active component of chlorine bleach (Klebanoff *et al.*, 2013; Schwartz *et al.*, 2009). Thus, in neutrophils the NADPH oxidase complex and MPO act in concert to generate microbicidal ROS during the first 60–90 min after phagocytosis (DeLeo *et al.*, 1999). Although the signals that terminate the respiratory burst in macrophages and neutrophils are not well defined, it is established that this process is accompanied by NADPH oxidase disassembly, as indicated by reverse translocation of Rac, p47^{phox}, p67^{phox}, and p40^{phox} from the phagosome membrane to the cytosol (DeLeo *et al.*, 1999; McCaffrey *et al.*, 2010).

iNOS

Although macrophages do not express MPO, they can be induced to express another enzyme that also has the capacity to increase the toxicity of superoxide anions, iNOS (also called NOS2). Specifically, iNOS expression is upregulated in macrophages exposed to certain bacterial products or proinflammatory cytokines such as lipopolysaccharide (LPS) or

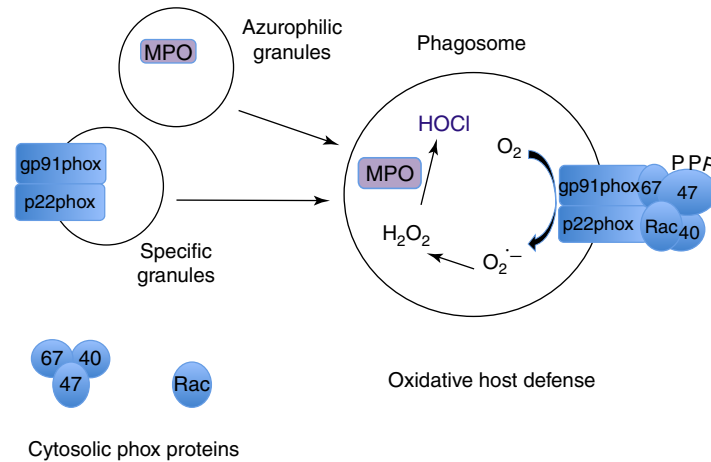


Figure 2 Oxidative host defense mechanisms. The phagocyte NADPH oxidase complex is comprised of six subunits: an integral membrane heterodimer (gp91^{phox}/p22^{phox}), and four soluble subunits (p47^{phox}, p67^{phox}, p40^{phox}, and Rac). This enzyme is disassembled in cells at rest, but is rapidly activated during phagocytosis. To this end, specific granules deliver gp91^{phox}/p22^{phox} to forming phagosomes. Simultaneously, phosphorylation of p47^{phox} and other signals trigger membrane translocation of the soluble subunits. The assembled, active enzyme catalyzes the conversion of molecular oxygen into superoxide anions, which then dismutate to form H₂O₂. Neutrophil phagosomes also acquire MPO from azurophilic granules. MPO catalyzes the conversion of H₂O₂ into HOCl, which is highly toxic to microbes and is also the active component of chlorine bleach.

interferon- γ , respectively (MacMicking *et al.*, 1997). iNOS catalyzes the conversion of L-arginine and oxygen into NO and citrulline, and highly toxic peroxynitrite (ONOO⁻) is generated when NO reacts with superoxide (MacMicking *et al.*, 1997). Although it is firmly established that iNOS is important for microbe killing by rodent macrophages, the contribution of this enzyme to the defense repertoire of human macrophages remains controversial, and the results of recent studies indicate that the iNOS promoter is silenced in alveolar macrophages isolated from human lungs (Gross *et al.*, 2014).

Chronic granulomatous disease

The importance of the respiratory burst and ROS in antimicrobial host defense is exemplified by the phenotype of persons with chronic granulomatous disease (CGD) (Matute *et al.*, 2009; Curnutte, 1993). Persons with CGD suffer from repeated, life-threatening bacterial and fungal infections due to inherited mutations in any one of the genes that encode NADPH oxidase subunits. Notably, CGD macrophages and neutrophils exhibit normal phagocytosis and oxygen-independent killing capacity (Allen *et al.*, 1999), and because of this persons with CGD are not uniformly susceptible to all infections. Rather, they are particularly susceptible to infection with bacteria such as *Staphylococcus aureus* and *Burkholderia cepacia*, and fungi such as *Aspergillus fumigatus* (Ben-Ari *et al.*, 2012). This illustrates the fact that each bacterial and fungal species exhibits differential sensitivity to the individual host defense mechanisms – including ROS – that are deployed by phagocytes.

Phagosome Maturation

Phagosome maturation is the process of converting a nascent phagosome into a degradative, lysosome-like compartment called a phagolysosome. Phagosome maturation has been extensively studied, and it is now clear that this process ensues

rapidly after microbe binding or uptake. However, both the kinetics of this process and the underlying molecular events differ significantly in macrophages and neutrophils, and this reflects the fact that neutrophils are granulocytes whereas macrophages are not. Thus, phagosome maturation in macrophages is a relatively slow process that involves incremental modification of phagosome composition over an hour or more via sequential interactions of this compartment with organelles of the endosomal pathway (Figure 3). In contrast, phagosome maturation in neutrophils is extremely rapid and is mediated by fusion of three types of cytoplasmic granules with forming and nascent phagosomes (Figure 4). Despite differences in the underlying molecular mechanisms, mature phagosomes in both phagocyte types are optimized for microbe killing and degradation.

Macrophage phagosome maturation

Until the mid-1980s it was generally believed that phagosome maturation was achieved by direct fusion of lysosomes with nascent phagosomes (Steinman *et al.*, 1983). This view began to change with the realization that remodeling of the phagosome membrane was rapid, and that most plasma membrane proteins disappeared from these structures within minutes of uptake (Muller *et al.*, 1980). Additional studies revealed that phagosomes become progressively acidified due to the action of an enzyme called the vacuolar proton adenosine triphosphatase (V-ATPase) that accumulates in phagosome membranes as they mature and pumps protons into the phagosome lumen (Mellman *et al.*, 1986).

Shortly thereafter, two research groups provided the first direct evidence to indicate that phagosome-endosome fusion occurs in macrophages and precedes phagosome fusion with lysosomes (Hart and Young, 1991; Mayorga *et al.*, 1991). Hart and Young examined phagosome composition using electron microscopy and central to their success was the use of macrophages that had been pretreated with ammonium

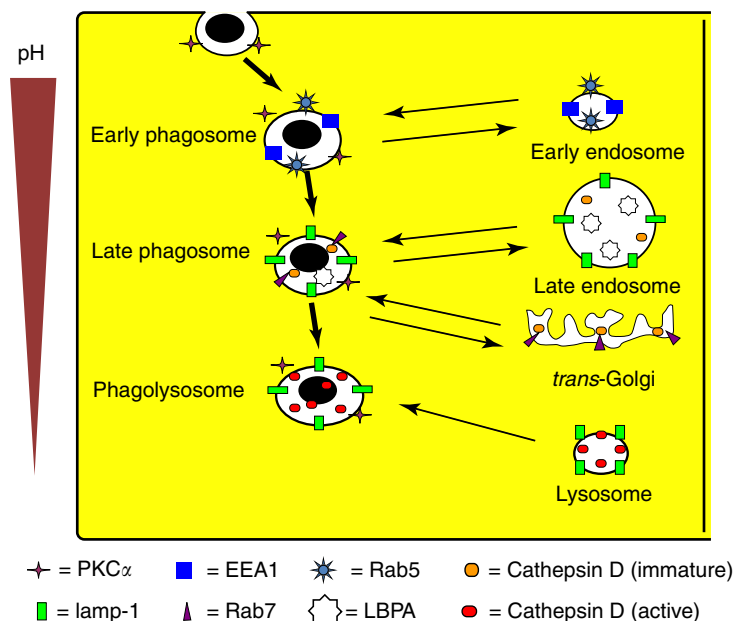


Figure 3 Macrophage phagosome maturation. The composition of nascent phagosomes in macrophages is modified incrementally by sequential interactions with early endosomes, late endosomes, *trans*-Golgi network-derived vesicles, and lysosomes. Prominent markers of early, late, and mature phagosomes include: Rab5 and EEA1; Rab7, LBPA, immature cathepsins and lamp-1; and mature cathepsins and lamp-1, respectively. Collectively, this process is called phagosome maturation, and it culminates in generation of an acidic compartment called a phagolysosome. The low pH and abundant hydrolases in mature phagolysosomes are critical components of the oxygen-independent killing arsenal of macrophages. EEA1, early endosome antigen-1; LBPA, lysobisphosphatidic acid; lamp-1, late endosome/lysosome-associated membrane protein-1.

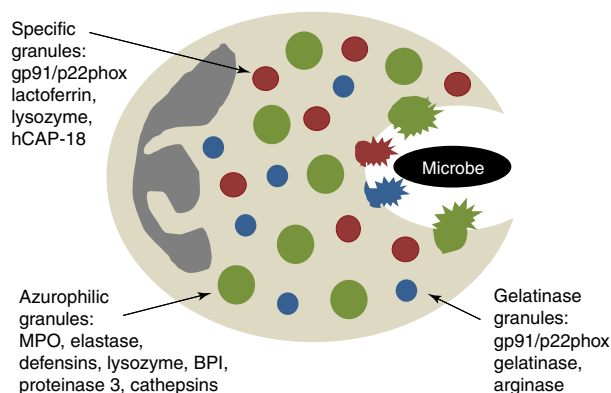


Figure 4 Neutrophil phagosome maturation. Maturation of neutrophil phagosomes is extremely rapid. Thus, fusion of all three types of cytoplasmic granules with forming phagosome begins within seconds of microbe binding and is apparent even before uptake and phagosome sealing is complete. Host defense components delivered into phagosomes by each granule population are listed. MPO, myeloperoxidase; BPI, bacteriocidal permeability-increasing protein.

chloride (NH₄Cl), which accumulates inside acidic membrane compartments (such as lysosomes) and neutralizes their pH. Under these conditions phagosome-lysosome fusion is blocked, and this allows phagosomes with endosome-like characteristics to accumulate (Hart and Young, 1991). The Stahl group also provided direct evidence of phagosome-endosome fusion in intact cells and were the first to reconstitute this process *in vitro*, thereby defining requirements for GTP and *N*-ethylmaleimide-sensitive factors in this process

(Mayorga *et al.*, 1991). Biochemical analyses of phagosomes isolated from macrophages at several time points during the first 20 min or so after particle uptake confirmed the rapid disappearance of phagocytic receptors and other plasma membrane proteins from these structures, as well as progressive accumulation of V-ATPase, and also demonstrated that phagosome acquisition of late endosome markers – such as the mannose-6-phosphate receptor (M6PR) – preceded acquisition of lysosomal enzymes such as cathepsin D (Pitt *et al.*, 1992). Collectively, the results of these pioneering studies provided a framework for in-depth studies of the underlying mechanisms.

Today, it is firmly established that maturation of macrophage phagosomes is intimately linked to membrane trafficking in the endosomal pathway, and as such phagosome composition is dynamic and changes rapidly during the first 5–60 min after microbe uptake via sequential interactions with early endosomes, late endosomes, vesicles derived from the *trans*-Golgi network, and lysosomes (Figure 3). Current understanding of this process has been informed by in depth analyses of phagosome protein and lipid composition, together with immunofluorescence and confocal microscopy and the development of probes for live cell imaging (Garin *et al.*, 2001; Russell *et al.*, 2009; Botelho *et al.*, 2000; Henry *et al.*, 2004; Vieira *et al.*, 2001).

Although our understanding of the molecular mechanisms is incomplete, several points of regulation have been identified, including a role for Rab family GTPases. Humans express more than 60 Rab GTPases, each of which has a distinct subcellular localization and regulates specific membrane trafficking and fusion events (Stenmark, 2009). Among these,

Rab5 and Rab7 are critical for phagosome maturation (Aderem and Underhill, 1999; Farin and Grinstein, 2012). Rab5 is a marker of early endosomes and mediates fusion of nascent phagosomes with these compartments. Consequently, in the first few minutes after sealing, early phagosomes are enriched for Rab5 and its effectors early endosome antigen-1 (EEA1) and the class 3 phosphatidylinositol 3-kinase Vps34, as well as other early endosome markers such as transferrin receptor. Subsequent disappearance of Rab5 and other early markers is accompanied by the appearance of late endosome markers on maturing phagosomes, including Rab7, late endosome/lysosome-associated membrane protein-1 (lamp-1), lamp-2, CD63, and lysobisphosphatidic acid (LBPA) (Henry *et al.*, 2004; Farin and Grinstein, 2012; Aderem and Underhill, 1999; Figure 3). Rab5-to-Rab7 conversion is essential for phagosome maturation, and this is due in part to the fact that Rab7 effectors including Rab-interacting lysosomal protein promote interactions of late phagosomes with microtubules and the dynein-driven movement of these compartments to the lysosome-rich perinuclear area (Farin and Grinstein, 2012; Harrison *et al.*, 2003). Late phagosomes also contain newly synthesized, immature hydrolases such as cathepsin D, which are delivered to phagosomes and late endosomes from the *trans*-Golgi network by M6PRs, which then recycle back to the Golgi via a Rab9-dependent mechanism (Vieira *et al.*, 2002; Stenmark, 2009). Thus, aberrant expression, targeting, or activity of Rab5, Rab7, Rab9, or Vps34 impairs phagosome maturation (Stenmark, 2009; Farin and Grinstein, 2012). Phagolysosomes are distinguished from late phagosomes by their greater density of lamp-1, lamp-2, and V-ATPase and lower pH, as well as their lack of LBPA and M6PRs and their enrichment for mature cathepsin D (Figure 3; Vieira *et al.*, 2002). Of note, more recent studies have identified roles for several additional Rab GTPases in regulation of phagosome acidification and maturation including Rab20, Rab22b, Rab32, Rab34, Rab38, Rab39, and Rab43 (Seto *et al.*, 2011).

Additional regulators of phagosome maturation include calcium/calmodulin and the alpha isoform of PKC (PKC α) (Allen and Aderem, 1995; Holm *et al.*, 2003; Vergne *et al.*, 2003). Several PKC isoforms are expressed in macrophages and contribute to phagocytosis, NADPH oxidase activation, and phagosome maturation in an isoform- and particle-dependent manner (Allen and Aderem, 1995; Larsen *et al.*, 2000; Allen and Allgood, 2002). PKC α differs from the other phagosome markers discussed above as it accumulates on forming phagocytic cups within seconds of microbe or particle binding, yet is retained throughout the maturation process and is essential for phagosome-lysosome fusion (Allen and Aderem, 1995; Holm *et al.*, 2003). Consistent with this, phagosomes also accumulate PKC substrates and molecules that modulate PKC α activation, including diacylglycerol and phosphatidylserine (Allen and Aderem, 1995; Levin *et al.*, 2015).

Neutrophil phagosome maturation

Phagosome maturation occurs by a different mechanism in neutrophils, and this reflects the distinct structure of this cell type as a granulocyte (Figure 4). Three types of granules are synthesized during neutrophil development in the bone

marrow: azurophilic (primary) granules (AG), specific (secondary) granules (SG), and gelatinase (tertiary) granules (GG). These structures fill the neutrophil cytoplasm, and each granule type contains molecules that contribute to host defense (Borregaard *et al.*, 2007; Pham, 2006), as shown in Figure 4. Conversely, the abundance of 'typical' organelles including mitochondria, endoplasmic reticulum, Golgi, early endosomes, and late endosomes is significantly diminished.

Speed and efficiency are defining features of phagosome maturation in neutrophils, and all three granule types are mobilized within seconds of microbe or particle binding (Figure 4). Thus, phagosome-granule fusion can be detected even before phagosome sealing and this process goes to completion within minutes (Tapper *et al.*, 2002; Bainton, 1973). Near-simultaneous delivery of SG and AG components to phagosomes is critical for cooperation among host defense mechanisms and is essential for optimal microbe killing. For example, the AG enzyme proteinase 3 processes the SG protein hCAP-18 to its active form, which is called LL-37 and kills bacteria via membrane disruption (Pham, 2006). As noted above, the catalytic core of the NADPH oxidase complex, flavocytochrome b₅₅₈, is stored in SG and to a lesser extent in GG, whereas MPO is stored in AG, and MPO allows neutrophils to convert NADPH oxidase-derived ROS into highly toxic HOCl in the phagosome lumen (Figure 2). The roles of other granule components in host defense are discussed below.

The mechanisms that regulate phagosome maturation in neutrophils are not well defined at the molecular level. It has been known for some time that degranulation is influenced by intracellular calcium, with progressively higher concentrations of this cation required for mobilization of GG, SG, and AG (Tapper *et al.*, 2002). Nonetheless, recent studies of Tapper and colleagues demonstrate that the role of calcium in phagosome-AG fusion is more complex than previously appreciated, as AG fusion with forming phagocytic cups is calcium-dependent whereas AG fusion with sealed phagosomes is not (Nordenfelt and Tapper, 2010, 2011; Nordenfelt *et al.*, 2009). Other molecules that have been implicated in AG mobilization include Rab27a, Munc13-4, Munc18-2, Rac, and the tyrosine kinase Hck (Monfregola *et al.*, 2012; Catz, 2013; Brochetta *et al.*, 2008; Abdel-Latif *et al.*, 2005; Welch *et al.*, 1996; N'Diaye *et al.*, 1998). In contrast, another Munc18 isoform, Munc18-3, preferentially associates with SG and GG (Brochetta *et al.*, 2008), and the efficiency of SG-phagosome fusion is influenced to some extent by opsonins and the receptors used for particle uptake (Joiner *et al.*, 1989).

It is also important to note that some proteins that localize to late endosomes and lysosomes in macrophages have a different subcellular localization in neutrophils. For example, CD63, lamp-1, and lamp-2 are all markers of macrophage late endosomes, yet in neutrophils CD63 is an AG marker, but lamp-1 and lamp-2 are not (Dahlgren *et al.*, 1995; Kallquist *et al.*, 2008). Rather, lamp-1 and lamp-2 associate with residual late endosome-like organelles in neutrophils along with Rab7, and Rab5 associates with neutrophil early endosomes. Rab5, Rab7, and lamp-1 have been detected on neutrophil phagosomes in some studies but not in others (Perskvist *et al.*, 2001; Minakanmi *et al.*, 2010; Nordenfelt and Tapper, 2011). The contribution of neutrophil endosomes to microbe killing remains uncertain because of their low abundance, and

because host defense molecules associated with endosomes in other cell types, such as lysozyme, cathepsins, and V-ATPase, localize instead to neutrophil AG. Indeed, AG and to a lesser extent SG are considered to be the lysosome-like compartments of neutrophils.

Oxygen-Independent Killing

Mechanisms for microbe killing that do not require oxygen are essential for host defense, as oxygen concentrations in many tissues and in abscesses are very low. Synergy between multiple defense mechanisms increases the efficacy of microbe killing, and in as much as certain microorganisms are preferentially controlled by ROS, others are predominantly controlled by nonoxidative mechanisms. Processes and molecules that comprise the oxygen-independent killing arsenal of phagocytes include phagosome acidification, iron sequestration, cationic antimicrobial peptides, and degradative enzymes. Some of these mechanisms are common to both macrophages and neutrophils, whereas others are cell type-specific.

V-ATPase and phagosome acidification

As noted above, macrophage phagosomes become progressively more acidic as they mature in parallel with accumulation of V-ATPase in the phagosome membrane. The pH of mature phagolysosomes can be as low as 4.5, and very few microbes are able to survive under these conditions. In addition to the direct toxicity of acidic pH, phagosome acidification is also required for phagosome-lysosome fusion in macrophages, and this process is blocked by V-ATPase inhibitors as well as agents that artificially elevate compartment pH, such as NH_4Cl (Hart and Young, 1991; Farin and Grinstein, 2012). Neutrophil phagosomes acquire V-ATPase from AG (Feuk-Lagerstedt *et al.*, 2007). Nonetheless, phagosomes in this cell type remain near neutral pH during the first hour after particle uptake, in parallel with robust NADPH oxidase activation (Nordenfelt and Tapper, 2011).

Lactoferrin, transferrin, and iron sequestration

Iron is an essential micronutrient that is required by microbes as well as host cells for their growth and survival. Limiting the availability of free iron for capture and use by microbes is therefore an important aspect of host defense and illustrates the concept of 'nutritional immunity' (Jurado, 1997). Iron sequestration is mediated primarily by transferrin and lactoferrin, which are delivered to macrophage and neutrophil phagosomes from early endosomes and SG, respectively (Figures 3 and 4; Pham, 2006; Borregaard *et al.*, 2007; Vieira *et al.*, 2002). Of note, lactoferrin is also secreted by epithelial cells and is present at high concentrations in tears, breast milk, respiratory and cervical secretions, seminal plasma and other extracellular fluids, and in this manner contributes to elimination of bacteria, viruses, and fungi in the extracellular space (Legrand *et al.*, 2005).

Defensins, LL-37, and other antimicrobial peptides

Antimicrobial proteins and peptides (AMPs) are ancient components of host defense that contribute to intraphagosomal killing by neutrophils and are also secreted into

extracellular fluids by epithelial cells at mucosal surfaces and in the skin (Wiesner and Vilcinskas, 2010). Neutrophil AMPs include the cathelicidin LL-37, the α -defensins HNP-1-HNP-4, and bacteriocidal permeability-increasing protein (BPI) (Borregaard *et al.*, 2007; Wiesner and Vilcinskas, 2010). All AMPs are cationic, and for this reason bind avidly to the negatively charged surfaces of microbes. Cathelicidins are peptides that disrupt microbe membranes. The α -defensins have similar effects on microbe membranes and may also inhibit DNA, RNA, and protein synthesis. Similar to α -defensins, BPI is stored in AGs and is delivered into neutrophil phagosomes. However, BPI is unique among AMPs as a full-length protein with two functional domains that facilitates elimination of gram-negative bacteria by multiple mechanisms (Schultz and Weiss, 2007). Specifically, the N-terminal domain of BPI binds LPS and neutralizes its proinflammatory capacity, and the C-terminal domain facilitates bacterial clearance. Additional AMPs, including those of the human β -defensin family, are synthesized and secreted by epithelial cells (Wiesner and Vilcinskas, 2010).

Degradative enzymes

Many enzymes contribute to microbe killing and degradation in phagosomes (Pham, 2006; Thorne *et al.*, 1976). Among these, elastase degrades proteins in the outer membranes of gram-negative bacteria, leading to membrane fragmentation and bacterial death, whereas enzymes of the cathepsin family can induce death of gram-positive bacteria. Lysozyme specifically degrades peptidoglycan, which is a structural component of the cell walls of both gram-positive and gram-negative bacteria. Cooperation between enzymes also increases killing efficiency. For example, elastase and cathepsins collaborate for intraphagosomal killing of fungi, whereas elastase-mediated degradation of the outer membranes of gram-negative bacteria increases lysozyme access to peptidoglycan in the intermembrane space. Finally, arginase is present in human neutrophils and murine macrophages, and catalyzes the conversion of arginine into urea and citrulline. As urea is a substrate for iNOS, arginase contributes indirectly to oxygen-dependent microbe killing.

Autophagy and Xenophagy

Autophagy (self-eating) is a conserved catabolic process that was first described as a mechanism to delay or prevent cell death in response to nutrient starvation (Levine *et al.*, 2011; Green and Levine, 2014). During starvation-induced autophagy crescents of membrane call phagophores (or isolation membranes) appear in the cytoplasm. These structures expand and then close to form double-membrane vesicles called autophagosomes. Under these conditions the content of each autophagosome is a random sample of cytoplasmic organelles and proteins that are subsequently delivered to lysosomes for degradation and amino acid reutilization. It was subsequently discovered that autophagy is also essential for degradation of damaged or aged mitochondria, peroxisomes, other organelles, as well as protein aggregates that are too large to fit inside the proteasome. In this case, selectivity is achieved by the use of cargo capture receptors (p62, NBD52, and NBR1). In brief, the organelles or proteins to be degraded are tagged with

ubiquitin, which allows them to bind one or more cargo capture receptors. These receptors can also bind LC3-II, a protein that is found exclusively in the phagophore membrane, thereby ensuring that the cargo of interest will be tethered to and trapped inside the forming autophagosome. The importance of selective autophagy to normal cell function is illustrated by the fact that defects in this process contribute to aging and to the pathology of many diseases, including Parkinson's disease, Huntington's disease, and Alzheimer's disease. Selective autophagy is also required for elimination of mitochondria and other organelles during red blood cell development.

Xenophagy

Xenophagy (foreign-eating) is essential for innate defense and refers to the use of selective autophagy by macrophages as a method to capture bacteria, fungi, parasites, and viruses inside autophagosomes, with subsequent fusion with lysosomes for elimination (Levine *et al.*, 2011; Figure 5). In practical terms, xenophagy gives macrophages a second chance to kill pathogens that inhibit or evade conventional phagosome maturation. For example, *Mycobacterium tuberculosis* inhibits phagosome maturation and replicates in compartments that resemble early endosomes, whereas *Listeria monocytogenes* escapes the phagosome and replicates in macrophage cytosol. Nonetheless, following induction of xenophagy, cytosolic *Listeria*, and intraphagosomal *Mycobacteria* can both be sequestered inside autophagosomes (Figure 5), and once formed these compartments mature by sequential interactions with late endosomes and lysosomes in a manner similar to the typical phagosome maturation pathway discussed above.

Cell Death Mechanisms and Host Defense

Several mechanisms of regulated (programmed) cell death have been described. Among these, pyroptosis and apoptosis

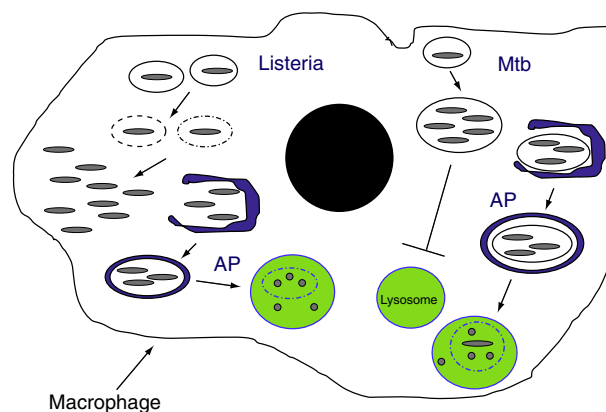


Figure 5 Xenophagy. The ability of autophagosomes to capture microbial pathogens and deliver them to lysosomes for degradation and killing is called xenophagy. Autophagosomes (AP) are double-membrane structures that can either capture bacteria replicating in macrophage cytosol, such as *L. monocytogenes*, or can ensnare whole phagosomes containing pathogens such as *M. tuberculosis* (Mtb), that otherwise blocks phagosome maturation and phagosome-lysosome fusion.

are critical for microbe killing and regulation of the inflammatory response by phagocytes. Another process, neutrophil extracellular trap formation, impairs bacterial spreading throughout an infected host and can also be coupled to cell death.

Pyroptosis

Pyroptosis is a proinflammatory form of macrophage death that is evoked by detection of pathogen components in the cytosol. 'Contamination' of the cytosol can occur by multiple mechanisms. As noted above, some pathogens escape from the phagosome to the cytosol, and bacterial replication releases small amounts of DNA, peptidoglycan, and proteins in this locale. Other pathogens act from the cell surface or from inside phagosomes by releasing toxins or using type III or type IV secretion systems to inject bacterial 'effector proteins' directly into the cytosol in order to manipulate the cytoskeleton, membrane trafficking, or other aspects of cell function as part of their virulence strategies. Proteins of the NOD-like receptor (NLR) family allow detection of microbe components in the cytosol (Broz and Monack, 2011; Elinav *et al.*, 2011). For example, AIM2 detects cytosolic DNA, NLRC4 detects secreted effector proteins, and NLRP1 detects secreted toxins. The NLRs then initiate formation of an inflammasome complex for activation of an enzyme called caspase-1. Active caspase-1 catalyzes the activation and secretion of two proinflammatory cytokines, IL-1 β and IL-18, and also induces plasma membrane disruption and cell death (Broz and Monack, 2011; Elinav *et al.*, 2011; Figure 6). This mechanism of cell death is proinflammatory because it is coupled to cytokine secretion, and because macrophage lysis also allows other proinflammatory molecules to be released into the extracellular space, including DAMPs. Inflammasome activation and

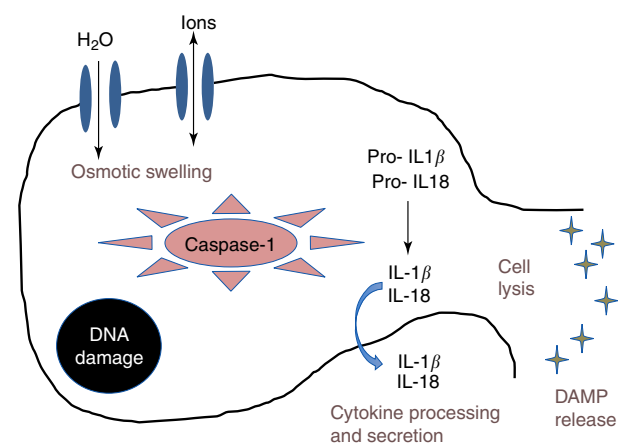


Figure 6 Macrophage pyroptosis. Pyroptosis is a proinflammatory form of cell death. Central to this process is caspase-1, which activates proinflammatory cytokines such as IL-1 β and IL-18. Other features of pyroptosis are DNA damage and loss of plasma membrane integrity. In particular, ion flux across the cell membrane and osmotic swelling leads to cell lysis and death. Concomitant spilling of cytotoxic cell contents, including DAMPs, is also proinflammatory. Pyroptosis contributes to host defense by inducing inflammation and by depriving intra-macrophage pathogens of their replicative niche. DAMPs, damage-associated molecular pattern molecules.

pyroptosis contribute to host defense by depriving intra-macrophage pathogens (such as *Listeria* and *Salmonella*) of their replicative niche. In addition, IL-1 β recruits neutrophils, which engulf and kill the released organisms (Elinav *et al.*, 2011). Thus, the contribution of pyroptosis to host defense requires cooperation between phagocyte types, and also illustrates the fact that most intra-macrophage pathogens are readily eliminated by neutrophils.

Apoptosis

Unlike macrophages which can survive for months, neutrophils are very short lived and are preprogrammed to die by constitutive apoptosis approximately 24 h after release into the bloodstream. For this reason, humans produce 10^{11} new neutrophils each day. Elegant studies of DeLeo and colleagues demonstrated that apoptosis is a final, regulated stage of the neutrophil lifecycle that is driven not only by intracellular signaling, but also by changes in gene expression that collectively comprise an 'apoptosis differentiation program' (Kobayashi *et al.*, 2003). Typically, neutrophil apoptosis is accelerated further by phagocytosis that is coupled to NADPH oxidase activation, and this phagocytosis-induced cell death response also requires changes in gene expression (Kobayashi *et al.*, 2002). During this process, neutrophil functional and proinflammatory capacity are downregulated, and the dying cells express surface markers that facilitate their internalization by macrophages via a process called efferocytosis. Rapid and efficient clearance of dying neutrophils by this mechanism is noninflammatory, and is therefore essential for daily neutrophil turnover and for restoration of tissue homeostasis following infection. Consistent with this, defects in neutrophil turnover and efferocytosis are indicative of a defective and aberrantly regulated inflammatory response that favors neutrophil progression to necrosis, exacerbates tissue destruction, and impairs microbe clearance. Inflammatory diseases associated with abnormal apoptosis include chronic obstructive pulmonary disease and rheumatoid arthritis (McCracken and Allen, 2014). In addition, bacterial pathogens that evade

elimination by neutrophils can affect the viability of these cells in two different ways. Pathogens that can grow inside neutrophils, such as *Anaplasma phagocytophilum* and *Francisella tularensis*, inhibit apoptosis and prolong the lifespan of the host cell to sustain the viability of their replicative niche, whereas extracellular bacteria, such as *Streptococcus pyogenes*, secrete toxins that induce neutrophil lysis in order to curtail intracellular killing and regain access to the extracellular milieu (McCracken and Allen, 2014).

Neutrophil Extracellular Traps and NETosis

Neutrophil extracellular traps (NETs) are complexes of chromosomal DNA, histones, and granule proteins that are released by neutrophils and ensnare extracellular microbes (Yipp and Kubes, 2013; Figure 7). As originally described, NET deployment is coupled to plasma membrane disruption and therefore culminates in a form of cell death called 'NETosis.' An initiating event in this process is robust NADPH oxidase activation at the plasma membrane, as occurs on response to soluble stimuli such as phorbol esters. Subsequent ROS-induced DNA damage leads to chromatin decondensation and loss of nuclear membrane integrity. Chromatin then mixes with elastase and other enzymes released into the cytosol from oxidant-damaged granules prior to cell lysis, DNA release, and NET formation (Figure 7). Additional *in vivo* and *in vitro* studies indicate that NET formation can also occur by mechanisms that do not require ROS, and by mechanisms that involve incremental DNA release and preserve cell viability and functional capacity (Yipp and Kubes, 2013). Thus, although our understanding of mechanisms that lead to NET formation is incomplete, it is established that trapping of extracellular bacteria and fungi in these structures can curtail their dissemination and limit the spread of infection, even if the NETs are not necessarily microbicidal.

Pathogen Subversion of Host Defense

Much of what is known about phagocyte function and mechanisms of host defense has been informed by studies of

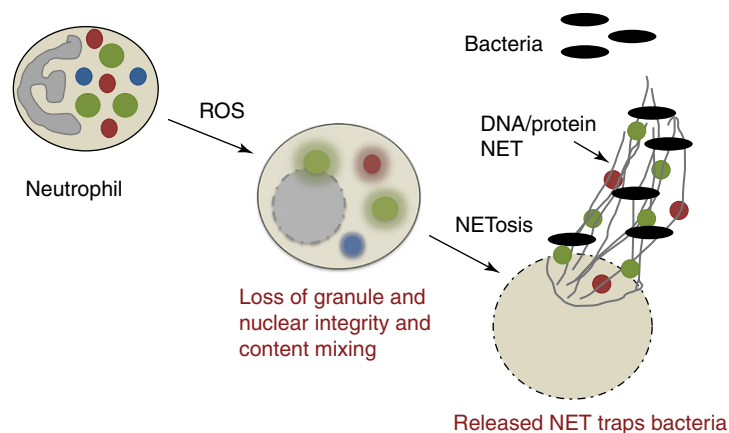


Figure 7 NETs and NETosis. Under certain circumstances ROS generated by activated neutrophils can disrupt both granule and nuclear membrane integrity. This allows granule proteins such as elastase and MPO to enter the nucleus and bind chromatin. Subsequent rupture of the plasma membrane allows deployment of this complex of DNA and granule proteins as a 'NET' or neutrophil extracellular trap that can ensnare extracellular bacteria and prevent their dissemination. Neutrophil death that is coupled to NET release is called NETosis.

microbial pathogens that perturb macrophage and neutrophil function as part of their virulence strategies (Smith and May, 2013; McCracken and Allen, 2014; Reddick and Alto, 2014). Although a detailed discussion of this topic is beyond the scope of this article, mechanisms of evasion include avoidance or inhibition of phagocytosis, disruption of NADPH oxidase assembly, activation or targeting, blockade of phagosome maturation at various stages of the pathway in either macrophages or neutrophils, neutralization of phagosome pH or adaptation for survival in acidic lysosomes, neutralization of AMPs, secretion of siderophores for iron capture that counteract the withholding effects of lactoferrin and transferrin, disruption of phagosome integrity and escape to the cytosol, interference with microbe detection systems including NLRs, and manipulation of phagocyte viability of mechanisms of cell death including apoptosis and pyroptosis.

Summary

Both macrophages and neutrophils are essential for microbe killing and host defense. These white blood cells accumulate at sites of infection and in this locale engulf bacteria and other microbes using a receptor and actin-dependent process called phagocytosis. The NADPH oxidase complex, AMPs, degradative enzymes, and iron sequestering proteins constitute the killing arsenal of macrophages and neutrophils, and allow creation of a highly lethal intraphagosomal environment. Concentration of these defense systems in close proximity to the ingested prey and collaboration between the different enzymes and molecules increases intracellular killing capacity and the efficiency of this process. Regulated mechanisms of cell death, including apoptosis and pyroptosis, also contribute to antimicrobial defense and regulation of the inflammatory response, and much of what we have learned about these processes has been informed by studies of bacterial and fungal pathogens that subvert certain aspects of macrophage or neutrophil defense in order to cause disease.

See also: Cellular Immunology: Innate Immunity: The Phagocyte NADPH Oxidase: Structure and Assembly of the Key Multicomponent Enzyme of Innate Immunity. Intracellular Infectiology: Cell Processes: Phagocytosis

References

- Abdel-Latif, D., Steward, M., Lacy, P., 2005. Neutrophil primary granule release and maximal superoxide generation depend on Rac2 in a common signalling pathway. *Canadian Journal of Physiology and Pharmacology* 83, 69–75.
- Aderem, A., Underhill, D.M., 1999. Mechanisms of phagocytosis in macrophages. *Annual Review of Immunology* 17, 593–623.
- Allen, L.-A.H., Aderem, A., 1995. A role for MARCKS, the alpha isozyme of protein kinase C and myosin I in zymosan phagocytosis by macrophages. *Journal of Experimental Medicine* 182, 829–840.
- Allen, L.-A.H., Aderem, A., 1996. Molecular definition of distinct cytoskeletal structures involved in complement- and Fc receptor-mediated phagocytosis in macrophages. *Journal of Experimental Medicine* 184, 627–637.
- Allen, L.-A.H., Allgood, J.A., 2002. Atypical protein kinase C- ζ is essential for delayed phagocytosis of *Helicobacter pylori*. *Current Biology* 13, 1762–1766.
- Allen, L.-A.H., DeLeo, F.R., Gallois, A., *et al.*, 1999. Transient association of the nicotinamide adenine dinucleotide phosphate oxidase subunits p47phox and p67phox with phagosomes in neutrophils from patients with X-linked chronic granulomatous disease. *Blood* 93, 3521–3530.
- Bainton, D.F., 1973. Sequential degranulation of the two types of polymorphonuclear leukocyte granules during phagocytosis of microorganisms. *Journal of Cell Biology* 58, 249–264.
- Ben-Ari, J., Wolbach, O., Gavirelli, R., Wolbach, B., 2012. Infections associated with chronic granulomatous disease: Linking genetics to phenotypic expression. *Expert Review of Anti-Infective Therapy* 10, 881–894.
- Borregaard, N., Sorensen, O.E., Theilgaard-Monch, K., 2007. Neutrophil granules: A library of innate immunity proteins. *Trends in Immunology* 28, 340–345.
- Botelho, R.J., Teruel, M., Dierckman, R., *et al.*, 2000. Localized biphasic changes in phosphatidylinositol-4,5-bisphosphate at sites of phagocytosis. *Journal of Cell Biology* 151, 1353–1367.
- Brochetta, C., Vita, F., Tiwari, N., *et al.*, 2008. Involvement of Munc18 isoforms in the regulation of granule exocytosis in neutrophils. *Biochimica et Biophysica Acta* 1783, 1781–1791.
- Broz, P., Monack, D.M., 2011. Molecular mechanisms of inflammasome activation during microbial infections. *Immunological Reviews* 243, 174–190.
- Casbon, A.-J., Allen, L.-A.H., Dunn, K.W., Dinan, M.C., 2009. Macrophage NADPH oxidase flavocytochrome b localizes to the plasma membrane and Rab11-positive recycling endosomes. *Journal of Immunology* 182, 2325–2339.
- Catz, S.D., 2013. Regulation of vesicular trafficking and leukocyte function by Rab27 GTPases and their effectors. *Journal of Leukocyte Biology* 94, 613–622.
- Cavaillon, J.-M., 2011. The historical milestones in the understanding of leukocyte biology initiated by Elie Metchnikoff. *Journal of Leukocyte Biology* 90, 413–424.
- Curnutte, J.T., 1993. Chronic granulomatous disease: The solving of a clinical riddle at the molecular level. *Clinical Immunology and Immunopathology* 67, S2–S15.
- Dahlgren, C., Carlsson, S.R., Karlsson, A., Lundqvist, H., Sjolin, C., 1995. The lysosomal membrane glycoproteins lamp-1 and lamp-2 are present in mobilizable organelles, but are absent from the azurophilic granules of human neutrophils. *Biochemical Journal* 311, 667–674.
- DeLeo, F.R., Allen, L.-A.H., Apicella, M., Nauseef, W.M., 1999. NADPH oxidase activation and assembly during phagocytosis. *Journal of Immunology* 163, 6732–6740.
- DeLeo, F.R., Quinn, M.T., 1996. Assembly of the phagocyte NADPH oxidase: Molecular interaction of oxidase proteins. *Journal of Leukocyte Biology* 60, 677–691.
- Elinav, E., Strowig, T., Henao-Mejia, J., Flavell, R.A., 2011. Regulation of the antimicrobial response by NLR proteins. *Immunity* 34, 665–679.
- Farin, G.D., Grinstein, S., 2012. How nascent phagosomes mature to become phagolysosomes. *Trends in Immunology* 33, 397–405.
- Feuk-Lagerstedt, E., Movitz, C., Pellme, S., Dahlgren, C., Karlsson, A., 2007. Lipid raft proteome of the human neutrophil azurophilic granule. *Proteomics* 7, 194–205.
- Freeman, S.A., Grinstein, S., 2014. Phagocytosis: Receptors, signal integration, and the cytoskeleton. *Immunological Reviews* 262, 193–215.
- Garin, J., Diez, R.D., Kieffer, S., *et al.*, 2001. The phagosome proteome: Insight into phagosome functions. *Journal of Cell Biology* 152, 165–180.
- Green, D.R., Levine, B., 2014. To be or not to be? How selective autophagy and cell death govern cell fate. *Cell* 157, 65–675.
- Griffin Jr., F.M., Griffin, J.A., Leider, J.E., Silverstein, S.C., 1975. Studies on the mechanism of phagocytosis. I. Requirements for circumferential attachment of particle-bound ligands to specific receptors on the macrophage plasma membrane. *Journal of Experimental Medicine* 142, 1263–1282.
- Gross, T.J., Kremens, K., Powers, L.S., *et al.*, 2014. Epigenetic silencing of the human NOS2 Gene: Rethinking the role of nitric oxide in human macrophage inflammatory responses. *Journal of Immunology* 192, 2326–2338.
- Harrison, R.E., Bucci, C., Viera, O.V., Schroer, T.A., Grinstein, S., 2003. Phagosomes fuse with late endosomes and/or lysosomes by extension of membrane protrusions along microtubules: Role of Rab7 and RILP. *Molecular and Cellular Biology* 23, 6494–6506.
- Hart, P.D., Young, M.R., 1991. Ammonium chloride, an inhibitor of phagosome-lysosome fusion in macrophages, concurrently induces phagosome-endosome fusion, and opens a novel pathway: Studies of pathogenic mycobacterium and nonpathogenic yeast. *Journal of Experimental Medicine* 174, 881–889.
- Henry, R.M., Hoppe, A.D., Joshi, N., Swanson, J.A., 2004. The uniformity of phagosome maturation in macrophages. *Journal of Cell Biology* 164, 185–194.
- Holevinsky, K.O., Nelson, D.J., 1998. Membrane capacitance changes associated with particle uptake during phagocytosis in macrophages. *Biophysical Journal* 75, 2577–2586.
- Holm, Å., Tejle, K., Gunnarsson, T., *et al.*, 2003. Role of protein kinase C α for uptake of unopsonized prey and phagosomal maturation in macrophages. *Biochemical and Biophysical Research Communications* 302, 653–658.

- Joiner, K.A., Ganz, T., Albert, J., Rotrosen, D., 1989. The opsonizing ligand on *Salmonella typhimurium* influences incorporation of specific, but not azurophilic, granule constituents into neutrophil phagosomes. *Journal of Cell Biology* 109, 2771–2782.
- Jurado, R.L., 1997. Iron, infections, and anemia of inflammation. *Clinical Infectious Diseases* 25, 888–895.
- Kallquist, L., Hansson, M., Persson, A.M., *et al.*, 2008. The tetraspanin CD63 is involved in granule targeting of neutrophil elastase. *Blood* 112, 3444–3454.
- Klebanoff, S.J., Kettle, A.J., Rosen, H., Winterbourn, C.C., Nauseef, W.M., 2013. Myeloperoxidase: A front-line defender against phagocytosed microorganisms. *Journal of Leukocyte Biology* 93, 185–198.
- Kobayashi, S.D., Voyich, J.M., Buhl, C.L., Stahl, R.M., DeLeo, F.R., 2002. Global changes in gene expression by human polymorphonuclear leukocytes during receptor-mediated phagocytosis: Cell fate is regulated at the level of gene expression. *Proceedings of the National Academy of Sciences of the United States of America* 99, 6901–6906.
- Kobayashi, S.D., Voyich, J.M., DeLeo, F.R., 2003. Regulation of the neutrophil-mediated inflammatory response to infection. *Microbes and Infection* 5, 1337–1344.
- Larsen, E.C., Digennaro, J.A., Saito, N., *et al.*, 2000. Differential requirement for classic and novel PKC isoforms in respiratory burst and phagocytosis in RAW 264.7 cells. *Journal of Immunology* 165, 2809–2817.
- Legrand, D., Elaiss, E., Carpentier, M., Mazurier, J., 2005. Lactoferrin: A modulator of immune and inflammatory responses. *Cellular and Molecular Life Sciences* 62, 2549–2559.
- Levin, R., Grinstein, S., Schlam, D., 2015. Phosphoinositides in phagocytosis and macropinocytosis. *Biochimica et Biophysica Acta* 1851 (6), 805–823. doi:10.1016/j.bbalip.2014.09.005.
- Levine, B., Mizushima, N., Virgin, H.W., 2009. Autophagy in immunity and inflammation. *Nature* 469, 323–335.
- MacMicking, J., Xie, Q.-W., Nathan, C., 1997. Nitric oxide and macrophage function. *Annual Review of Immunology* 15, 323–350.
- Matute, J.D., Aarias, A.A., Wright, N.A.M., *et al.*, 2009. A new genetic subgroup of chronic granulomatous disease with autosomal recessive mutations in p40phox and selective defects in neutrophil NADPH oxidase activity. *Blood* 114, 3309–3315.
- Mayorga, L.S., Bertini, F., Stahl, P.D., 1991. Fusion of newly formed phagosomes with endosomes in intact cells and in a cell-free system. *Journal of Biological Chemistry* 268, 6511–6517.
- McCaffrey, R.L., Schwartz, J.T., Lindemann, S.R., *et al.*, 2010. Multiple mechanisms of NADPH oxidase inhibition by type A and type B *Francisella tularensis*. *Journal of Leukocyte Biology* 88, 791–805.
- McCracken, J.M., Allen, L.-A.H., 2014. Regulation of human neutrophil apoptosis and lifespan in health and disease. *Journal of Cell Death* 7, 15–23.
- Mellman, I., Fuchs, R., Helenius, A., 1986. Acidification of the endocytic and exocytic pathways. *Annual Review of Biochemistry* 55, 663–700.
- Minakami, R., Maehara, Y., Kamakura, S., *et al.*, 2010. Membrane phospholipid metabolism during phagocytosis in human neutrophils. *Genes to Cells* 15, 409–424.
- Montfregola, J., Johnson, J.L., Meijler, M.M., Napolitano, G., Catz, S.D., 2012. MUNC13-4 protein regulates the oxidative response and is essential for phagosomal maturation and bacterial killing in neutrophils. *Journal of Biological Chemistry* 287, 44603–44618.
- Muller, W.A., Steinman, R.M., Cohn, Z.A., 1980. The membrane proteins of the vacuolar system II. Bidirectional flow between secondary lysosomes and the plasma membrane. *Journal of Cell Biology* 86, 304–314.
- N'Diaye, E.-N., Darzacq, X., Astarie-Degeker, C., *et al.*, 1998. Fusion of azurophilic granules with phagosomes and activation of the tyrosine kinase Hck are specifically inhibited during phagocytosis of Mycobacteria in human neutrophils. *Journal of Immunology* 161, 4983–4991.
- Nauseef, W.M., 2004. Assembly of the phagocyte NADPH oxidase. *Histochemistry and Cell Biology* 122, 277–291.
- Nordenfelt, P., Tapper, H., 2010. The role of calcium in neutrophil granule–phagosome fusion. *Communicative & Integrative Biology* 3, 224–226.
- Nordenfelt, P., Tapper, H., 2011. Phagosome dynamics during phagocytosis by neutrophils. *Journal of Leukocyte Biology* 90, 271–284.
- Nordenfelt, P., Winberg, M.E., Lonnbro, P., Rasmusson, B., Tapper, H., 2009. Different requirements for early and late phases of azurophilic granule–phagosome fusion. *Traffic* 10, 1881–1893.
- Perskvist, N., Roberg, K., Kulyte, A., Stendahl, O., 2001. Rab5a GTPase regulates fusion between pathogen-containing phagosomes and cytoplasmic organelles in human neutrophils. *Journal of Cell Science* 115, 1321–1330.
- Pham, C.T.N., 2006. Neutrophil serine proteases: Specific regulators of inflammation. *Nature Reviews Immunology* 6, 541–550.
- Pitt, A., Mayorga, L.S., Stahl, P.D., Schwartz, A.L., 1992. Alterations in the protein composition of maturing phagosomes. *Journal of Clinical Investigation* 90, 1978–1983.
- Reddick, L.E., Alto, N.M., 2014. Bacteria fighting back: How pathogens target and subvert the host innate immune system. *Molecular Cell* 54, 321–328.
- Russell, D.G., Vavderen, B.C., Glennie, S., Mwandumba, H., Heyderman, R.S., 2009. The macrophage marches on its phagosome: Dynamic assays of phagosome function. *Nature Reviews Immunology* 9, 594–600.
- Schultz, H., Weiss, J.P., 2007. The bactericidal/permeability-increasing protein (BPI) in infection and inflammatory disease. *Clinica Chimica Acta* 384, 12–23.
- Schwartz, J., Leidal, K.G., Fleming, J.K., Weiss, J.P., Nauseef, W.M., 2009. Neutrophil bleaching of GFP-expressing Staphylococci: Probing the intraphagosomal fate of individual bacteria. *Journal of Immunology* 183, 2632–2641.
- Seto, S., Tsujimura, K., Koide, Y., 2011. Rab GTPases regulating phagosome maturation are differentially recruited to mycobacterial phagosomes. *Traffic* 12, 407–420.
- Smith, L.M., May, R.C., 2013. Mechanisms of microbial escape from phagocyte killing. *Biochemical Society Transactions* 41, 475–490.
- Steinman, R.M., Mellman, I., Muller, W.A., Cohn, Z.A., 1983. Endocytosis and the recycling of plasma membrane. *Journal of Cell Biology* 96, 1–27.
- Stenmark, H., 2009. Rab GTPases as coordinators of vesicle traffic. *Nature Reviews Molecular Cell Biology* 10, 513–525.
- Tapper, H., Furuya, W., Grinstein, S., 2002. Localized exocytosis of primary (lysosomal) granules during phagocytosis: Role of Ca²⁺-dependent tyrosine phosphorylation and microtubules. *Journal of Immunology* 168, 5287–5296.
- Thorne, K.J.I., Oliver, R.C., Barrett, A.J., 1976. Lysis and killing of bacteria by lysosomal proteinases. *Infection and Immunity* 14, 555–563.
- Vergne, I., Chua, J., Deretic, V., 2003. Tuberculosis toxin blocking phagosome maturation inhibits a novel Ca²⁺/calmodulin-P13K hVps34 cascade. *Journal of Experimental Medicine* 198, 653–659.
- Vieira, O.V., Botelho, R.J., Grinstein, S., 2002. Phagosome maturation: Aging gracefully. *Biochemical Journal* 366, 689–704.
- Vieira, O.V., Botelho, R.J., Rameh, L., *et al.*, 2001. Distinct roles of class I and class III phosphatidylinositol 3-kinases in phagosome formation and maturation. *Journal of Cell Biology* 155, 19–25.
- Welch, H., Maura, C., Maridonneau-Parini, I., 1996. Nonreceptor protein tyrosine kinases in neutrophil activation. *Methods* 9, 607–618.
- Wiesner, J., Vilcinskas, A., 2010. Antimicrobial peptides: The ancient arm of the human immune system. *Virulence* 1, 440–464.
- Yipp, B.G., Kubes, P., 2013. NETosis: How vital is it? *Blood* 122, 2784–2794.

Further Reading

- Amulic, B., Cazalet, C., Metzler, K.D., Zychlinsky, A., 2012. Neutrophil function: From mechanisms to disease. *Annual Review of Immunology* 30, 459–489.
- Flannagan, R.S., Jaumouille, V., Grinstein, S., 2012. The cell biology of phagocytosis. *Annual Review of Pathology-Mechanisms of Disease* 7, 61–98.
- Kennedy, A., DeLeo, F., 2009. Neutrophil apoptosis and the resolution of infection. *Immunologic Research* 43, 25–61.
- Kobayashi, S.D., Braughton, K.R., Whitney, A.R., *et al.*, 2003. Bacterial pathogens modulate an apoptosis differentiation program in human neutrophils. *Proceedings of the National Academy of Sciences of the United States of America* 100, 10948–10953.
- Nauseef, W.M., 2004. Assembly of the phagocyte NADPH oxidase. *Histochemistry and Cell Biology* 122, 277–291.
- Nauseef, W.M., 2007. How human neutrophils kill and degrade microbes: An integrated view. *Immunological Reviews* 219, 88–102.
- Von Moltke, J., Ayres, J.S., Kofoed, E.M., Chavarria-Smith, J., Vance, R.E., 2013. Recognition of bacteria by inflammasomes. *Annual Review of Immunology* 31, 73–106.

Bacterial Subversion of Phagocytic Killing

RR Isberg, S Asrat, DA de Jesús, AD Hempstead, and V Ramabhadran, Tufts University School of Medicine, Boston, MA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The default pathway for uptake of bacteria by mammalian cells leads to the formation of a membrane-bound phagosome that matures and eventually fuses with lysosomal compartments. Movement via this pathway is usually thought to be antimicrobial because it immediately activates oxygen-dependent killing by the NADPH oxidase complex. Whether a microbe will survive this attack, is a further danger from acidic organelles such as lysosomes, which are sites of concentrated antimicrobial molecules including proteases and small antimicrobial peptides (Asrat *et al.*, 2014). As uptake by macrophages and neutrophils is a primary host innate defense against disease, many pathogenic bacteria have strategies to either prevent phagocytosis, or change the progression of the uptake pathway so that antimicrobial activities can be suppressed and microbial survival and growth can be supported. As a rule, extracellular pathogens attempt to prevent phagocytosis, while intracellular pathogens allow uptake, but support intracellular routing in a fashion that avoids degradation and attack by the host cell. In this article, we will review these strategies.

Interference of Uptake by Bacterial Pathogens

After microbial entry into the host, the first interaction is with either the mucosal layer or the host cell plasma membrane. In case of an encounter with phagocytic cells, contact with the plasma membrane can lead to uptake and phagocytic killing. An efficient mechanism to avoid targeting by phagocytic cells is to directly misregulate or destroy target cells via a focused pathogen attack on the plasma membrane.

Blocking Uptake via the Deposition of Bacterial Pore-Forming Toxins

A common strategy of targeting phagocytes is for extracellular pathogens to synthesize soluble oligomeric proteins that, on contact with host cells, form pores (Gonzalez *et al.*, 2008). The pore-forming toxins (PFTs) can have either α -helical or β -sheet structures that form a channel in the target plasma membrane. Cholesterol-dependent cytolysins are the best characterized of the β -sheet class, forming oligomers of approximately 35–44 monomers after targeting host cell cholesterol via a Thr-Leu motif on the toxins (Dunstone and Tweten, 2012). The toxins that form α -helical motifs in the target plasma membrane are made of monomers with a central hydrophobic helix that generates a pore via oligomerization after partial unfolding (Gonzalez *et al.*, 2008).

PFTs play a number of roles in the disease process, including killing phagocytes and targeting lymphocytes for destruction, and it is common for a single bacterial pathogen to encode multiple types of toxins. The membrane receptors

targeted by these toxins can also be diverse, which allows the targeting of specific cell types. For instance, *Staphylococcus aureus* has multiple PFTs with distinct cell tropisms based on their receptor binding profiles (Dumont and Torres, 2014). Receptors include ADAM10, which is located on a number of host cell types, and is targeted by Staphylococcal α -hemolysin, resulting in either lysis or altered regulation of host cell signaling pathways, such as enhanced chemokine secretion in neutrophils and cellular proliferation of keratinocytes. Another receptor targeted is the HIV co-receptor CCR5 by the LukED toxin. Binding via this receptor allows killing of macrophages and dendritic cells, while binding of the non-cholesterol-dependent LukED to CXCR1 and CXCR2 is linked to killing of neutrophils (Reyes-Robles *et al.*, 2013).

Inactivation of Phagocytosis After Intimate Contact with Pathogens

After contact with phagocytes, a number of pathogens are able to misregulate phagocytic signaling, interfering with uptake and locking the pathogen on the host cell surface. The most common strategy for accomplishing this task involves translocating bacterial-encoded misregulators into host cells using type III secretion systems (T3SS) (Burkinshaw and Strynadka, 2014). T3SS are multiprotein complexes that form needle-like structures, assembled in the envelope of many gram-negative bacterial pathogens. They are thought to inject proteins directly through the host cell plasma membrane. The best characterized of these translocated antiphagocytic proteins are Rho GTPase-activating proteins (RhoGAPs). These proteins target multiple Rho family members, known to be central regulators of actin polymerization events that support phagocytic cup formation. The RhoGAPs act by stimulating the rate of intrinsic GTPase hydrolysis, resulting in a Rho family protein that is charged with GDP and therefore inactive.

Each member of the *Yersinia* species, including the causative agent of bubonic plague *Yersinia pestis*, encodes a T3SS that translocates YopE, a RhoGAP that targets RhoA, RhoG, Rac1, and Cdc42, inhibiting phagocytosis (Black and Bliska, 2000). Similarly, the *Pseudomonas aeruginosa* ExoS protein has RhoGAP activity that interferes with phagocytosis (Goehring *et al.*, 1999). A different strategy of targeting Rho family members is taken by the extracellular *Vibrio parahaemolyticus* species. The *V. parahaemolyticus* T3SS substrate VopS AMPylates Rho, Rac, and Cdc42, resulting in steric hindrance that prevents interactions of Rho family members with downstream effectors of actin-mediated rearrangements (Yarbrough *et al.*, 2009).

After contact with pathogens, phagocytes can target the microbe for destruction via both oxygen-dependent and -independent pathways. The most important pathway for O₂-dependent killing is via the assembly and activation of the NADPH oxidase complex on the plasma membrane.

Pathogens can overcome this form of killing by either directly inactivating components of the complex, or by inducing a form of uptake into host cells that avoids activation of the complex. For the most part, the mechanistic bases of these strategies are poorly understood, although it is clear that the bacterial proteins that interfere with phagocytosis also directly target the NADPH oxidase complex. For instance, assembly of this complex requires Rac2 and/or Rac1, depending on the cell-type, to promote assembly (Hordijk, 2006). The *Yersinia* YopE RhoGAP protein inactivates both of these Rho family members and prevents the oxidative burst. Therefore, a single bacterial protein can interfere with two distinct steps in phagocytic killing, by preventing O₂-dependent killing and uptake into a compartment that could target the bacterium to be killed by O₂-independent strategies.

The Replication Vacuole 1: Manipulation of the Endocytic Pathway to Support Bacterial Replication

Phagosome acidification and targeting into a degradative lysosome is an O₂-independent strategy that leads to microbial demise, and arrest or bypass of this trafficking pathway is a common strategy used by bacterial pathogens that grow in membrane-bound vacuoles. Pathogens such as *Mycobacterium tuberculosis* (*Mtb*), *Legionella pneumophila*, *Salmonella enterica* serovar Typhimurium and *Chlamydia trachomatis* appear to interfere with acidification and, depending on the pathogen, can be internalized via a pathway that totally bypasses entry into endocytic compartments. In some cases, maturation and acidification of the vacuole seem to be encouraged, as exemplified by *Brucella abortus* and *Coxiella burnetii*. These pathogens use strategies to convert normally antimicrobial environments of phagolysosomes into ones that support microbial replication.

Blockade of Endocytic Maturation

Tuberculosis is caused by the growth of *M. tuberculosis* within lung alveolar macrophages. After uptake, the compartment harboring *Mtb* appears to mimic early endosomes, but there is a block in acidification and little evidence for further progress down the endocytic path. This block requires the bacterial ESX-1 secretion system (MacGurn and Cox, 2007). As expected for a compartment with impaired acidification, there is reduced acquisition of the v-ATPase, which may be due to a blockade in maturation from a Rab5-rich to a Rab7-rich compartment (Via et al., 1997; Gonzalez et al., 2008). *Mtb* also releases a number of bioactive lipids as well as manipulates host lipid signaling pathways to prevent maturation of the bacterium-containing compartment. The bacterial lipids include manose-capped lipoarabinomannan (ManLAM), which lowers phosphatidylinositol-3-phosphate (PI3P) levels on the vacuolar membrane, possibly by interfering with Ca²⁺ flux (Vergne et al., 2005). Similarly, the bacterial SapM phosphatase targets PI3P to generate PI, which further blocks maturation into a late phagosome (Vergne et al., 2005). Another strategy for interfering with maturation involves the secreted phosphatase PtpA, which targets the vacuolar protein VPS33B, a regulator of

fusion of compartments within the lysosomal network (Bach et al., 2008).

Survival After Entry into the Endocytic Pathway

An excellent example of a bacterial species that survives after entering into an endocytic pathway is *Salmonella*. Immediately after pathogen-induced uptake by T3SS-translocated proteins, the bacterium is found in a compartment that has some characteristics of both early and late endosomes. Host proteins Rab5 and the PI3-kinase hVPS34, diagnostic of an endocytic route, are observed on the *Salmonella*-containing vacuole (SCV), with consequent PI3P accumulation (Bakowski et al., 2010). Accumulation of this lipid species prevents docking of the recycling Rab35 GTPase (Heo et al., 2006).

Once SCV construction is initiated, there is a selective series of interactions that appear to mimic movement into a lysosomal niche, but clearly keep the microbe separate from antimicrobial effectors of the host. Although markers such as EEA1, Rab5, Rab7, and Lamp-1/2, indicative of some maturation along the endocytic pathway are localized to the SCV, clear steps are taken to prevent entry into a lysosomal compartment (Figure 1; Méresse et al., 1999). There is an absence of key markers of the endocytic path on the SCV, such as mannose-6-phosphate receptors, but there may be some limited exchange of material with lysosomes, making the SCV into a compartment of unique composition (Drecktrah et al., 2007).

Pathogens That Intimately Interact with Lysosomal Components

The gram-negative pathogens *Coxiella burnetii* and *B. abortus* appear to actively enter into a compartment that is lysosomal in nature. The Q-fever agent, *C. burnetii* constructs a vacuole that, similar to *Salmonella*, accumulates Rab5 followed by Rab7 (Romano et al., 2007). Initially, the *Coxiella*-containing vacuole (CCV) forms a compartment with the membrane in tight opposition to the bacterium. This tight vacuole enlarges as a result of fusion to both endocytic compartments as well as to other CCVs (Coleman et al., 2004; Figure 1). Eventually, the CCV occupies most of the host cell cytoplasm (Coleman et al., 2004). Unlike the SCV, the CCV has clear phagolysosomal characteristics, with presence of markers of the lysosome, such as cathepsin D (Romano et al., 2007).

Formation of the CCV requires that the bacterium assemble a dedicated protein translocation system called a type IV secretion system (T4SS). In addition, the function of specific v-SNAREs, such as VAMP7 and VAMP3, is required for full maturation into large CCVs (Campoy et al., 2013). One T4SS translocated substrate known to be critical for successful CCV formation is CvpA, which interacts with the clathrin complex adaptor protein AP2 (Larson et al., 2013). In addition, CCV formation probably involves interaction with the *trans*-Golgi network (TGN), as the retromer complex is necessary for maturation of the vacuole (McDonough et al., 2013).

In the case of *B. abortus*, construction of the *Brucella*-containing vacuole (BCV) involves early contact with lysosomal compartments (Atluri et al., 2011). Fusion with the lysosome

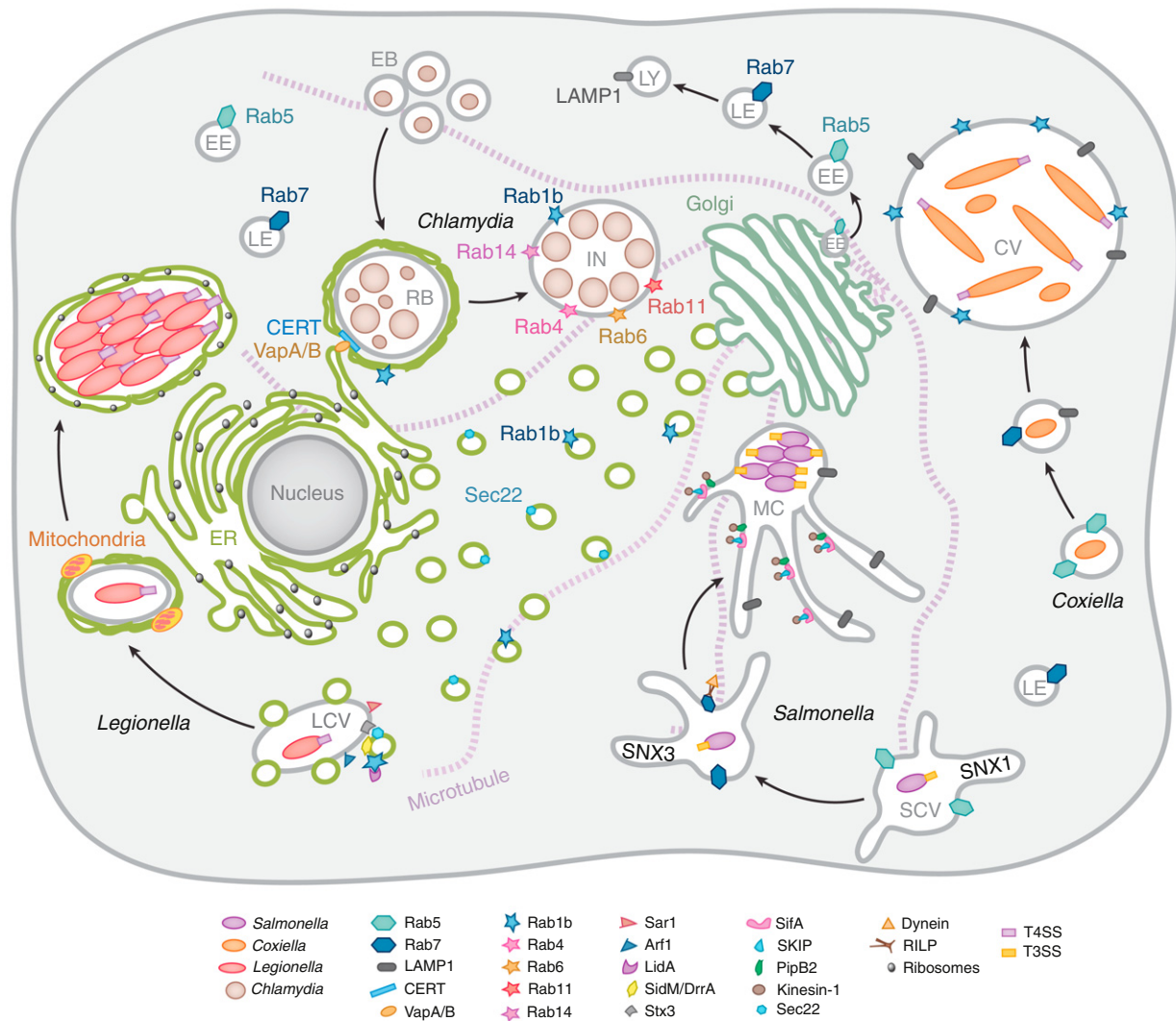


Figure 1 Targeting of host organelles by bacterial pathogens. Shown are a variety of bacterial pathogens that grow within vacuolar compartments. *Salmonella* and *Coxiella* are found in compartments that have proteins associated with the endocytic pathway, whereas other intracellular pathogens avoid this route. In the case of *Coxiella*, the bacterium-containing compartment moves deep into the endocytic pathway, fusing with the lysosome, while with *Salmonella*, movement is truncated. *Legionella* and *Chlamydia* avoid entry into this route. *Legionella* interacts intimately with ER-derived membrane components while the chlamydial inclusion recruits ceramide transfer protein (CERT) and associates with the ER and the Golgi. Reproduced from Asrat, S., De Jesus, D.A., Hempstead, A.D., Ramabhadran, V., Isberg, R.R., 2014. Bacterial pathogen manipulation of host membrane trafficking. *Annual Reviews of Cell and Developmental Biology* 30, 79–109.

appears transient, but is necessary for vacuole construction and initiating intracellular growth. Interestingly, this early event, which is associated with acidification, may activate the bacterial T4SS (Boschiroli *et al.*, 2002). After initial lysosomal interactions, T4SS translocated substrates block further maturation, and prevent the appearance of lysosomal markers such as cathepsin D (Atluri *et al.*, 2011). Once this transient interaction with the lysosome is complete, endoplasmic reticulum (ER)-derived vesicles are recruited to the BCV to mature it into a replication-competent vacuole. Bacterial components other than the T4SS, and its effectors, are also important for BCV construction. For example, the bacterial cyclic β -1,2 glucan interferes with long term BCV-lysosome

fusion, presumably by disrupting lipid rafts (Arellano-Reynoso *et al.*, 2005).

The Replication Vacuole 2: Manipulation of the Secretory Pathway to Support Bacterial Replication

A common strategy to avoid internalization into an endocytic antimicrobial pathway is for the pathogen to induce uptake in a fashion that promotes immediate interaction with the secretory pathway. This allows the bacterium to totally bypass many of the default killing mechanisms associated with phagocytosis. Numerous pathogens use this strategy by translocating

microbial molecules into the host cell that control the nature of the trafficking of the bacterium-containing compartment.

Intimate Interaction with the Reticulum

Targeting of ER dynamics as a strategy to allow pathogen replication is best characterized with the causative agent of Legionnaire's disease, *L. pneumophila*. A pathogen of amoebae in the environment, the organism grows within alveolar macrophages during human disease in the *Legionella*-containing vacuole (LCV) that is closely associated with rough endoplasmic reticulum (RER) (Isberg *et al.*, 2009). Interactions with the host cell early secretory pathway, involving budding of vesicles from the ER, occurs almost immediately after contact with host cells as a consequence of bacterial proteins translocated into the host cell via a T4SS.

Consistent with the microbe targeting membrane traffic that exits the ER, *L. pneumophila* translocates a large number of proteins that control steps associated with this process. The Rab1 GTPase, which modulates docking and fusion of ER-derived vesicles with downstream compartments such as the Golgi, is a common target of many of these bacterial proteins. The incredible diversity of strategies used to control this GTPase indicates that the bacterium takes over total control of its function throughout the replication cycle. Rab1 is recruited and activated by the SidM/DrrA guanine nucleotide exchange factor (GEF), and the Rab1GAP LepB can reverse its activation (Ham and Orth, 2011). Additionally, the Rab1 GTPase can be locked in its nucleotide-bound form in a number of ways. SidM/DrrA has a domain that activates Rab1 by adenylation (AMPylation), while SidD can reverse this activation by removing the modification. Finally, both LidA and AnkX can lock Rab1 in the nucleotide-bound form that the GTPase assumes at the time of contact (Tan and Luo, 2011). AnkX performs this function by a novel phosphorylcholine reaction, which can be reversed by the *L. pneumophila* Lem3 protein (Tan *et al.*, 2011).

Added onto this hijacking of Rab1 function are multiple other levels of attack on the host cell early secretory apparatus. The bacterial RalF protein was the first *Legionella* T4SS translocated substrate identified, and is responsible for activation of Arf1, a GTPase involved in vesicle egress from the ER (Isberg *et al.*, 2009). SNARE proteins are also targeted. Activation of Rab1 by SidM/DrrA promotes the formation of an unusual pairing of ER-derived and plasma membrane (PM)-derived SNARE proteins, potentially driving a membrane docking and fusion event between these two membrane sites. This event involves forming a Stx3/Stx4 (PM) and Sec22b (ER) complex, a process that requires active Sar1 and Arf1 (Figure 1; Arasaki *et al.*, 2012). Finally, *Legionella* inhibits retrograde traffic from the Golgi. The bacterial T4SS substrate RidL binds to the retromer protein VPS29, preventing accumulation of cargo receptors on the LCV (Finsel *et al.*, 2013).

In contrast to *Legionella*, *Brucella* hijacks retrograde Golgi-ER traffic. Although budding from the ER is necessary for construction of the *Brucella*-containing vacuole (BCV), the Rab1 GTPase that is involved in anterograde traffic is dispensable (Atluri *et al.*, 2011). Consistent with the model that retrograde transport is important for BCV formation, the retrograde GTPase Rab2 is critical for bacterial replication

(Fugier *et al.*, 2009). In fact, *Brucella melitensis* encodes a protein, RicA, which binds to the GDP-bound form Rab2, but there is no evidence that this protein modulates GTPase activity (de Barsy *et al.*, 2011). The details of establishment of the BCV are still being worked out, but similar to *Legionella*, the control of replication vacuole formation by T4SS translocated proteins appears to hold true. For instance, the T4SS substrate SepA was shown to be important during early stages of BCV establishment, and in its absence there is increased fusion of the BCV with lysosomes (Dohmer *et al.*, 2014).

In addition to pathogens that hijack retrograde or anterograde secretory traffic in host cells, some take advantage of bidirectional traffic to control formation of the vacuole that harbors the pathogen. In the case of *Salmonella* species, this results in the SCV assuming a tubular structure in epithelial cells (called a Sif) although the SCV in macrophages does not appear filamentous. Formation of the SCV involves positive interactions with retromer cargo proteins and sorting-nexin proteins-1 and 3 (SNX1 and SNX3) (Bujny *et al.*, 2008). In addition, the SCV must associate with the Golgi. This is apparently mediated by transient interactions with Rab7 and its effector RILP which allows positioning of the SCV in a perinuclear site (Guignot *et al.*, 2004).

Formation of Sifs appears to be tightly linked to the control of kinesin. To this end, a number of *Salmonella* T3SS substrates modulate kinesin function. For instance, kinesin-1 is recruited by the bacterial PipB2 protein (Henry *et al.*, 2006; Figure 1). By far, the most important bacterial protein for controlling this structure is the SifA protein, which controls Sif formation by modulating activities of a number of proteins associated with microtubule and membrane trafficking dynamics. Early in infection, RILP is recruited to the vacuole, and this is potentially terminated by SifA-Rab7 interactions (Harrison *et al.*, 2004). Secondly, SifA controls kinesin function by binding to the host protein kinesin-1-interacting protein SKIP, which promotes Sif formation by stimulating anterograde transport of vesicles from the replication compartment (Boucrot *et al.*, 2005). In the absence of SifA or SKIP, a proper replication compartment cannot form, and intracellular replication is defective (Boucrot *et al.*, 2005).

SifA may play a role in coordinating control of the microtubule network with Rab-regulated membrane dynamics, although there are some differences regarding mechanistic details of these interactions. It was first shown that SifA acts to interfere with the function of the late endosomal GTPase Rab9 by outcompeting its binding to SKIP, which is thought to control the distribution of LAMP1-containing compartments (Jackson *et al.*, 2008). Recent data suggests that dual control of Rab9 and SKIP enables *Salmonella* to manipulate mannose phosphate receptor (MPR) traffic and lysosomal degradation (McGourty *et al.*, 2012). By looking at either model, it is reasonable to speculate that the inability of SifA to control Rab9 function results in a defective niche that cannot protect the bacterium.

Salmonella is not the only pathogen that translocates proteins to support maintenance of replication vacuole integrity. The *Legionella* T4SS SdhA protein prevents the vacuole from falling apart, ensuring that microbial products do not contaminate the host cell cytosol and induce inflammatory signaling (Creasey and Isberg, 2014). There are a number of

functional similarities between SifA and SdhA, even though the proteins show no sequence similarity. For instance, *sdhA* and *sifA* mutations both cause replication vacuole rupture. Secondly, in each case, the instability of the compartment can be partially reversed by deleting a bacterial phospholipase gene (Creasey and Isberg, 2014; Ruiz-Albert *et al.*, 2002). Finally, both appear to control microtubule function.

The Replication Vacuole 3: Multi-Organelle Interactions

Chlamydia trachomatis is an obligate gram-negative bacterium that undergoes a developmental switch to promote disease in its hosts. The elementary body (EB) developmental form is infectious and is internalized into cells, where it differentiates into the reticulate body, which is the replicative form. After internalization, and simultaneously with this developmental switch, the compartment holding the bacterium forms a replication vacuole, called an inclusion, whose formation is promoted by a number of bacterial proteins called Incs that insert into the vacuole membrane (Mital *et al.*, 2010). The inclusion localizes to a perinuclear locale where it avoids fusion with lysosomal compartments. This locale in the cell appears perfectly situated to interface with anterograde traffic from the Golgi, and the bacterium directly manipulates Golgi structure and traffic. The bacterium promotes cleavage of Golgin-84, which results in formation of Golgi mini-stacks in cells harboring *C. trachomatis* (Heuer *et al.*, 2009). This fragmentation appears to allow movement of ceramide into the inclusion, which is then incorporated as sphingomyelin into the inclusion membrane (Heuer *et al.*, 2009). The importance of sphingolipid incorporation is emphasized by the fact that interference of trafficking of ceramide impairs *Chlamydia* development and causes loss of inclusion membrane integrity (Heuer *et al.*, 2009; Elwell *et al.*, 2011; Derre *et al.*, 2011).

Sphingomyelin biosynthesis on the inclusion membrane appears to be a consequence of nonvesicular trafficking of ceramide, and this is the first evidence that nonvesicular traffic contributes to intracellular survival of pathogens (Elwell *et al.*, 2011; Derre *et al.*, 2011). There is clear docking between the ER-localized ceramide transfer protein CERT and the inclusion that is mediated by the bacterial IncD protein, indicating that the bacterial replication compartment controls both ER and Golgi function (Elwell *et al.*, 2011; Derre *et al.*, 2011). CERT, in turn, is linked to ER-associated VapA/B (Figure 1) and sphingomyelin synthases SMS1/2, allowing this multiprotein complex to support sphingomyelin synthesis on the inclusion membrane (Elwell *et al.*, 2011; Derre *et al.*, 2011).

Breaking Free to Avoid Host Cell Killing

It would appear that the simplest strategy to avoid killing via lysosomal antimicrobial responses would be to simply break open a phagosome after uptake and set up a replication niche within the host cell cytosol. This strategy raises complex problems for the pathogen, because in many cell types the presence of cytosolic bacteria stimulates innate immune signaling, which can result in either host cell death, autophagic

clearing of the pathogen, or a robust cytokine response. Therefore, pathogens that follow this strategy must also have the capacity to bypass nonlysosomal antimicrobial responses.

Listeria is a gram-positive rod that enters into the host cytosol shortly after phagocytosis, before the bacterium-containing compartment is able to traffic to the lysosome. *Listeria* appears to interfere with the endocytic Rab5a protein, with the consequence that fusion of the compartment with early endosomes is blocked, potentially preventing maturation of the compartment (Alvarez-Dominguez *et al.*, 2008). The modulation of Rab5a activity may be a consequence of the ADP ribosyltransferase activity of the *Listeria* protein Lmo 2459, which interferes with GDP/GTP exchange by the Rab protein (Alvarez-Dominguez *et al.*, 2008).

The single bacterial protein that is most important for both the intracellular survival of *Listeria* as well as its ability to enter the host cell cytosol is the secreted PFT listeriolysin O (LLO). The activity of the protein is quite pH sensitive with optimal activity observed at approximately pH=5.5, appropriate for a protein that can be activated in an acidic compartment followed by down-modulation in the more neutral cytosolic pH. Other bacterial proteins are also involved in breakdown of the compartment, including the phosphatidylinositol-specific phospholipase C (PI-PLC), and a broad-range phosphatidylcholine phospholipase C (PC-PLC), which is believed to be involved in destruction of the vacuolar double-membranes after the bacterium spreads to neighboring cells (Smith *et al.*, 1995).

The details of vacuole escape by *Shigella* have been more difficult to decipher, and escape may be modulated by proteins whose primary function is to allow movement of proteins from the bacterial cytosol into the host cytosol via a T3SS. These proteins, often called translocon proteins, are inserted into the host cell membrane, forming pores that allow the final step in translocation via the T3SS apparatus. In the case of *Shigella* species, these proteins are IpaB and IpaC, which associate preferentially with cholesterol-rich regions of the membrane (Page *et al.*, 1999; Ferraris *et al.*, 2010). Insertion of these proteins into the host cell is facilitated by another T3SS-translocated protein IpaD (Picking *et al.*, 2005). As *Shigella* encodes a number of phospholipases that are translocated via this system, it is likely that escape from the vacuole is complex, and may involve coordination between the host and the bacterium.

In the case of *Rickettsia* species, the compartment harboring the bacterium appears to be degraded by the action of bacterial phospholipases. The most likely candidate is the secreted phospholipase D (Renesto *et al.*, 2003). In fact, *Salmonella* harboring the gene for this protein encoded by *Rickettsia prowazekii* were able to escape from the replication vacuole (Whitworth *et al.*, 2005). It is believed that lipid substrates of this phospholipase D are found in the compartment, and their depletion by the action of the protein leads to loss of membrane integrity and vacuolar disruption.

Bypass and Exploitation of Host Autophagic Clearing of Microorganisms

One of the primary strategies that host cells utilize to clear cytoplasmically localized pathogens is autophagy. Autophagy was traditionally associated with the ordered engulfment of

cytosolic proteins and organelles that need to be degraded to maintain cellular homeostasis, but it is now clear that this process can also be used to free the cell of unwanted intruders (Christen *et al.*, 2009). In particular, the ubiquitination of either intracellular microbes or the compartments encompassing them are important signals for targeting to the autophagic machinery. As expected, a strategy that allows microbes to survive uptake into host cells is for them to evade autophagic killing. A clever strategy used by *Salmonella* to avoid autophagy is to activate mTOR, a negative regulator of autophagy (Tattoli *et al.*, 2012). The action of mTOR is disrupted if it is removed from endomembranes. In the case of *Salmonella* entry into host cells, the protein gets detached leading to initial inactivation, but the bacterium is able to reverse this effect by recruiting mTOR to the surface of the bacterium-containing vacuole (Tattoli *et al.*, 2012). Although there is significant down-modulation of autophagy, the bacterium uses a second strategy to further bypass this host response. There is ubiquitination of the bacterial compartment, and the *Salmonella* protein SseL deubiquitinates proteins, preventing recognition by host autophagic receptors (Figure 2; Mesquita *et al.*, 2012).

Listeria interferes with autophagy at several levels. The *Listeria* protein ActA allows the cytosol-localized bacterium to move via actin-based motility, and this is important for avoiding recognition by autophagy because mutants lacking this protein get ubiquitinated and cleared by the autophagic machinery (Ogawa *et al.*, 2009). In parallel, the host

autophagic machinery is impaired by *Listeria* phospholipases PI-PLC and PC-PLC, possibly by reducing PI3P levels (Tattoli *et al.*, 2013). Finally, the bacterial surface protein InlK recruits the host's major vault protein (MVP) to the bacterial surface, which interferes with recognition of bacterial cell surface ubiquitination by the host cell autophagic machinery (Dortet *et al.*, 2011; Figure 2).

Shigella faces similar problems to those encountered by *Listeria*, as both organisms are localized in the host cell cytosol and use actin-based motility for cell-to-cell spread. An important host initiator of autophagy, Atg5, binds to the bacterial IcsA protein, which is needed to promote bacterial actin-based motility and spread to other cells. Thus the bacterium is faced with the conundrum of needing to spread via a protein that could target it into a degradative compartment. *Shigella* escapes autophagy by secreting IcsB, which binds IcsA and prevents Atg5 recognition (Ogawa *et al.*, 2011). In addition, the bacterium secretes VirA, a RabGAP protein that inactivates Rab1, a protein important in the formation of autophagosomes (Dong *et al.*, 2012; Zoppino *et al.*, 2010).

It was originally hypothesized that *L. pneumophila* delays autophagy (Amer and Swanson, 2005), and recent evidence agrees with this model. *Legionella pneumophila* translocates a protein, RavZ, which interferes with autophagy by inactivating a key component of this pathway. RavZ irreversibly deconjugates LC3, a critical protein that needs to be linked to the surface of pre-autophagosomal membranes to allow the process of

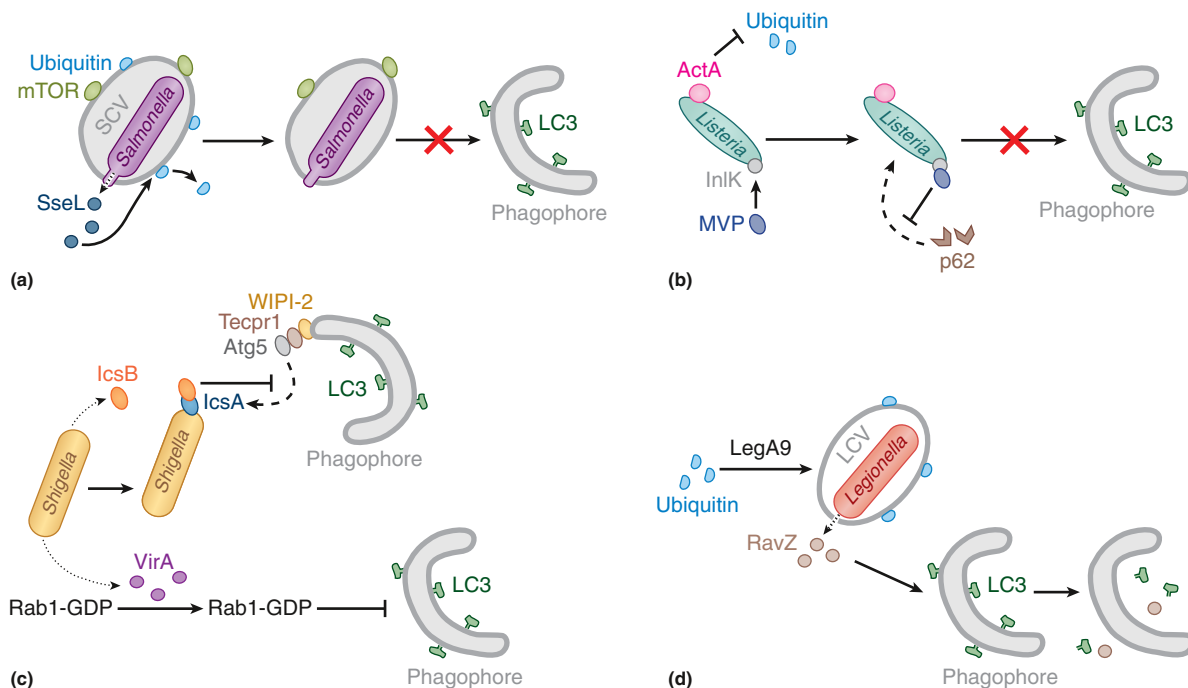


Figure 2 Modulation of host cell autophagy pathways. (a) Tug of war – *Legionella* both stimulates autophagy and blocks it. Bacterial proteins such as LegA9 cause ubiquitination of the bacterial compartment which could lead to autophagy. The bacterium also translocates RavZ into host cells, which blocks autophagic recognition by deconjugating LC3 from the phagophore. (b) Ubiquitination of *Listeria* is blocked by the bacterial protein ActA. The bacterial InlK protein recruits MVP which interferes with the targeting of the bacterium to the p62 linker protein that directs cargo to the autophagic machinery. (c) *Salmonella* interferes with autophagy by recruiting mTOR to the SCV. The bacterium also encodes a deubiquitinase, SseL, which inhibits autophagy. (d) *Shigella* prevents IcsA from interacting with the autophagy machinery by producing IcsB, which interferes with this interaction. The *Shigella* protein VirA also blocks autophagy by inactivating Rab1. See Asrat *et al.* (2014) for further details.

clearing of cytoplasmic structures to be completed (Choy *et al.*, 2013). A recent report indicated that another *L. pneumophila* protein, LegA9 stimulates recognition of the LCV by the host autophagy machinery, although the role of this recognition during the replication process is unclear (Khweek *et al.*, 2013). LegA9 mutants show less autophagic markers around the replication compartment, indicating that the protein may target the autophagy machinery to the site of bacterial replication (Khweek *et al.*, 2013). A summary of bacterial interactions with the autophagic machinery is displayed in Figure 2.

Bacterial Embrace of Autophagy

Surprisingly, some bacteria actively exploit the autophagic machinery to ensure survival and intracellular growth. The compartment encompassing *Coxiella* is composed of a number of autophagic markers and disruption of autophagy interferes with bacterial growth (Romano *et al.*, 2007). Furthermore, it appears that proteins translocated by the bacteria are required to form this autophagosome-like compartment (Carey *et al.*, 2011). Autophagy may support the growth of *Brucella*, as well (Pizarro-Cerda *et al.*, 1998). The *Brucella* replication compartment is encompassed by an autophagosome, and its formation is dependent on a subset of autophagy proteins, including Beclin-1 (ATG6) (Starr *et al.*, 2012). In addition to aiding growth, there is evidence that cell-to-cell spread of *Brucella* is supported by the formation of these autophagosomes (Starr *et al.*, 2012). Finally, although *Francisella* grows within the cell cytosol, it appears to enter a compartment that has autophagosome-like characteristics late in the infection cycle (Checroun *et al.*, 2006). These compartments stain positive for both late endosomal as well as autophagosomal markers, such as LC3 (Checroun *et al.*, 2006). Once inside this compartment, the bacteria cease to replicate, but they are not killed, thus it is possible that this autophagosomal compartment supports bacterial egress from the host cell.

Perspective

Phagocytic attack is the initial strategy that allows a host to battle against invading microorganisms. Therefore, being able to survive this attack is a primary determinant of whether a microbe is able to cause disease. The primary strategies that microbes use to combat the phagocyte are to kill the cell, prevent phagocytosis, or promote survival within an intracellular niche after phagocytosis. There are clearly a number of ways in which pathogenic bacteria can follow each of these strategies, and the approach taken by each is an important factor that defines the nature of the disease that occurs in the host. Studying how pathogens subvert and manipulate the host cell has contributed greatly to our understanding of a number of cell biological processes and has allowed detailed dissection of how pathogens cause disease at the molecular level.

See also: Intracellular Infectiology: Cell Processes: Phagocytosis

References

- Alvarez-Dominguez, C., Madrazo-Toca, F., Fernandez-Prieto, L., *et al.*, 2008. Characterization of a *Listeria monocytogenes* protein interfering with Rab5a. *Traffic* 9, 325–337.
- Amer, A.O., Swanson, M.S., 2005. Autophagy is an immediate macrophage response to *Legionella pneumophila*. *Cell Microbiology* 7, 765–778.
- Arasaki, K., Toomre, D., Roy, C., 2012. The *Legionella pneumophila* effector DrrA is sufficient to stimulate SNARE-dependent membrane fusion. *Cell Host & Microbe* 11, 46–57.
- Arellano-Reynoso, B., Lapaque, N., Salcedo, S., *et al.*, 2005. Cyclic beta-1,2-glucan is a *Brucella* virulence factor required for intracellular survival. *Nature Immunology* 6, 618–625.
- Asrat, S., De Jesus, D.A., Hempstead, A.D., Ramabhadran, V., Isberg, R.R., 2014. Bacterial pathogen manipulation of host membrane trafficking. *Annual Reviews of Cell and Developmental Biology* 30, 79–109.
- Atliuri, V., Xavier, M., De Jong, M., Den Hartigh, A., Tsois, R.E., 2011. Interactions of the human pathogenic *Brucella* species with their hosts. *Annual Review of Microbiology* 65, 523–541.
- Bach, H., Papavinasandaram, K., Wong, D., Hmama, Z., Av-Gay, Y., 2008. *Mycobacterium tuberculosis* virulence is mediated by PtpA dephosphorylation of human vacuolar protein sorting 33B. *Cell Host & Microbe* 3, 316–322.
- de Barse, M., Jamet, A., Filopon, D., *et al.*, 2011. Identification of a *Brucella* spp. secreted effector specifically interacting with human small GTPase Rab2. *Cellular Microbiology* 13, 1044–1058.
- Bakowski, M., Braun, V., Lam, G., *et al.*, 2010. The phosphoinositide phosphatase SopB manipulates membrane surface charge and trafficking of the *Salmonella*-containing vacuole. *Cell Host & Microbe* 7, 453–462.
- Black, D.S., Bliska, J.B., 2000. The RhoGAP activity of the *Yersinia pseudotuberculosis* cytotoxin YopE is required for antiphagocytic function and virulence. *Molecular Microbiology* 37, 515–527.
- Boschiroli, M., Ouahrani-Bettache, S., Foulongne, V., *et al.*, 2002. The *Brucella* suis virB operon is induced intracellularly in macrophages. *Proceedings of the National Academy of Sciences of the United States of America* 99, 1544–1549.
- Boucrot, E., Henry, T., Borg, J.-P., Gorvel, J.-P., Mèresse, S., 2005. The intracellular fate of *Salmonella* depends on the recruitment of kinesin. *Science* 308, 1174–1178.
- Bujny, M., Ewels, P., Humphrey, S., *et al.*, 2008. Sorting nexin-1 defines an early phase of *Salmonella*-containing vacuole-remodeling during *Salmonella* infection. *Journal of Cell Science* 121, 2027–2036.
- Burkinshaw, B.J., Strynadka, N.C., 2014. Assembly and structure of the T3SS. *Biochimica et Biophysica Acta* 1843, 1649–1663.
- Campoy, E., Mansilla, M., Colombo, M., 2013. Endocytic SNAREs are involved in optimal *Coxiella burnetii* vacuole development. *Cellular Microbiology* 15, 922–941.
- Carey, K., Newton, H.J., Lührmann, A., Roy, C.R., 2011. The *Coxiella burnetii* Dot/Icm system delivers a unique repertoire of type IV effectors into host cells and is required for intracellular replication. *PLoS Pathogens* 7, e1002056.
- Checroun, C., Wehrly, T., Fischer, E., Hayes, S., Celli, J., 2006. Autophagy-mediated reentry of *Francisella tularensis* into the endocytic compartment after cytoplasmic replication. *Proceedings of the National Academy of Sciences of the United States of America* 103, 14578–14583.
- Choy, A., Dancourt, J., Mugo, B., *et al.*, 2013. The *Legionella* effector RavZ inhibits host autophagy through irreversible Atg8 deconjugation. *Science* 338, 1072–1076.
- Christen, M., Coye, L.H., Hontz, J.S., *et al.*, 2009. Activation of a bacterial virulence protein by the GTPase RhoA. *Science Signaling* 2, ra71.
- Coleman, S., Fischer, E., Howe, D., Mead, D., Heinzen, R., 2004. Temporal analysis of *Coxiella burnetii* morphological differentiation. *Journal of Bacteriology* 186, 7344–7352.
- Creasey, E.A., Isberg, R., 2014. Maintenance of vacuole integrity by bacterial pathogens. *Current Opinion in Microbiology* 17, 42–46.
- Derre, I., Swiss, R., Agaisse, H., 2011. The lipid transfer protein CERT interacts with the *Chlamydia* inclusion protein IncD and participates to ER-*Chlamydia* inclusion membrane contact sites. *PLoS Pathogens* 7, e1002092.
- Dohmer, P.H., Valguarnera, E., Czibener, C., Ugalde, J.E., 2014. Identification of a type IV secretion substrate of *Brucella abortus* that participates in the early stages of intracellular survival. *Cellular Microbiology* 16, 396–410.
- Dong, N., Zhu, Y., Lu, Q., *et al.*, 2012. Structurally distinct bacterial TBC-like GAPs link Arf GTPase to Rab1 inactivation to counteract host defenses. *Cell* 150, 1029–1041.

- Dortet, L., Mostowy, S., Samba-Louaka, A., *et al.*, 2011. Recruitment of the major vault protein by IINIK: A *Listeria monocytogenes* strategy to avoid autophagy. *PLoS Pathogens* 7, e1002168.
- Drecktrah, D., Knodler, L., Howe, D., Steele-Mortimer, O., 2007. *Salmonella* trafficking is defined by continuous dynamic interactions with the endolysosomal system. *Traffic* 8, 212–225.
- Dumont, A.L., Torres, V.J., 2014. Cell targeting by the *Staphylococcus aureus* pore-forming toxins: It's not just about lipids. *Trends in Microbiology* 22, 21–27.
- Dunstone, M.A., Tweten, R.K., 2012. Packing a punch: The mechanism of pore formation by cholesterol dependent cytolysins and membrane attack complex/perforin-like proteins. *Current Opinion in Structural Biology* 22, 342–349.
- Elwell, C.A., Jiang, S., Kim, J.H., *et al.*, 2011. *Chlamydia trachomatis* co-opts GBF1 and CERT to acquire host sphingomyelin for distinct roles during intracellular development. *PLoS Pathogens* 7, e1002198.
- Ferraris, D.M., Gherardi, E., Di, Y., Heinz, D.W., Niemann, H.H., 2010. Ligand-mediated dimerization of the Met receptor tyrosine kinase by the bacterial invasion protein IINB. *Journal of Molecular Biology* 395, 522–532.
- Finsel, I., Ragaz, C., Hoffmann, C., *et al.*, 2013. The *Legionella* effector RidL inhibits retrograde trafficking to promote intracellular replication. *Cell Host & Microbe* 14, 38–50.
- Fugier, E., Salcedo, S.P., de Chastellier, C., *et al.*, 2009. The glyceraldehyde-3-phosphate dehydrogenase and the small GTPase Rab 2 are crucial for *Brucella* replication. *PLoS Pathogens* 5, e1000487.
- Goehring, U.M., Schmidt, G., Pederson, K.J., Aktories, K., Barbieri, J.T., 1999. The N-terminal domain of *Pseudomonas aeruginosa* exoenzyme S is a GTPase-activating protein for Rho GTPases. *Journal of Biological Chemistry* 274, 36369–36372.
- Gonzalez, M.R., Bischofberger, M., Pernot, L., Van Der Goot, F.G., Freche, B., 2008. Bacterial pore-forming toxins: The (w)hole story? *Cellular and Molecular Life Sciences* 65, 493–507.
- Guignot, J., Caron, E., Beuzón, C., *et al.*, 2004. Microtubule motors control membrane dynamics of *Salmonella*-containing vacuoles. *Journal of Cell Science* 117, 1033–1045.
- Ham, H., Orth, K., 2011. De-AMPylation unmasked: Modulation of host membrane trafficking. *Science Signaling* 4, pe42.
- Harrison, R., Brumell, J., Khandani, A., *et al.*, 2004. *Salmonella* impairs RILP recruitment to Rab7 during maturation of invasion vacuoles. *Molecular Biology of the Cell* 15, 3146–3154.
- Henry, T., Couillault, C., Rockenfeller, P., *et al.*, 2006. The *Salmonella* effector protein PipB2 is a linker for kinesin-1. *Proceedings of the National Academy of Sciences of the United States of America* 103, 13497–13502.
- Heo, W., Inoue, T., Park, W., *et al.*, 2006. PI(3,4,5)P3 and PI(4,5)P2 lipids target proteins with polybasic clusters to the plasma membrane. *Science* 314, 1458–1461.
- Heuer, D., Rejman Lipinski, A., Machuy, N., *et al.*, 2009. *Chlamydia* causes fragmentation of the Golgi compartment to ensure reproduction. *Nature* 457, 731–735.
- Hordijk, P.L., 2006. Regulation of NADPH oxidases: The role of Rac proteins. *Circulation Research* 98, 453–462.
- Isberg, R.R., O'Connor, T.J., Heidtman, M., 2009. The *Legionella pneumophila* replication vacuole: Making a cosy niche inside host cells. *Nature Reviews Microbiology* 7, 13–24.
- Jackson, L., Nawabi, P., Hentea, C., Roark, E., Haldar, K., 2008. The *Salmonella* virulence protein SifA is a G protein antagonist. *Proceedings of the National Academy of Sciences of the United States of America* 105, 14141–14146.
- Khweek, A.A., Caution, K., Akhter, A., *et al.*, 2013. A bacterial protein promotes the recognition of the *Legionella pneumophila* vacuole by autophagy. *European Journal of Immunology* 43, 1333–1344.
- Larson, C., Beare, P., Howe, D., Heinzen, R., 2013. *Coxiella burnetii* effector protein subverts clathrin-mediated vesicular trafficking for pathogen vacuole biogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 110, 9.
- MacGurn, J., Cox, J., 2007. A genetic screen for *Mycobacterium tuberculosis* mutants defective for phagosome maturation arrest identifies components of the ESX-1 secretion system. *Infection and Immunity* 75, 2668–2678.
- McDonough, J., Newton, H., Klum, S., *et al.*, 2013. Host pathways important for *Coxiella burnetii* infection revealed by genome-wide RNA interference screening. *mBio* 4, 12.
- McGourty, K., Thurston, T., Matthews, S., *et al.*, 2012. *Salmonella* inhibits retrograde trafficking of mannose-6-phosphate receptors and lysosome function. *Science* 338, 963–967.
- Méresse, S., Steele-Mortimer, O., Finlay, B., Gorvel, J., 1999. The rab7 GTPase controls the maturation of *Salmonella typhimurium*-containing vacuoles in HeLa cells. *EMBO Journal* 18, 4394–4403.
- Mesquita, F.S., Thomas, M., Sachse, M., *et al.*, 2012. The *Salmonella* deubiquitinase SseL inhibits selective autophagy of cytosolic aggregates. *PLoS Pathogens* 8, e1002743.
- Mital, J., Miller, N., Fischer, E., Hackstadt, T., 2010. Specific chlamydial inclusion membrane proteins associate with active Src family kinases in microdomains that interact with the host microtubule network. *Cellular Microbiology* 12, 1235–1249.
- Ogawa, M., Nakagawa, I., Yoshikawa, Y., *et al.*, 2009. *Streptococcus*-, *Shigella*-, and *Listeria*-induced autophagy. *Methods Enzymology* 452, 363–381.
- Ogawa, M., Yoshikawa, Y., Kobayashi, T., *et al.*, 2011. A Tecpr1-dependent selective autophagy pathway targets bacterial pathogens. *Cell Host & Microbe* 9, 376–389.
- Page, A.L., Ohayon, H., Sansonetti, P.J., Parsot, C., 1999. The secreted IpaB and IpaC invasins and their cytoplasmic chaperone IpgC are required for intercellular dissemination of *Shigella flexneri*. *Cellular Microbiology* 1, 183–193.
- Picking, W.L., Nishioka, H., Hearn, P.D., *et al.*, 2005. IpaD of *Shigella flexneri* is independently required for regulation of Ipa protein secretion and efficient insertion of IpaB and IpaC into host membranes. *Infection and Immunity* 73, 1432–1440.
- Pizarro-Cerda, J., Meresse, S., Parton, R., *et al.*, 1998. *Brucella abortus* transits through the autophagic pathway and replicates in the endoplasmic reticulum of nonprofessional phagocytes. *Infection and Immunity* 66, 5711–5724.
- Renesto, P., Dehoux, P., Gouin, E., *et al.*, 2003. Identification and characterization of a phospholipase D-superfamily gene in rickettsiae. *Journal of Infectious Diseases* 188, 1276–1283.
- Reyes-Robles, T., Alonzo 3rd, F., Kozhaya, L., *et al.*, 2013. *Staphylococcus aureus* leukotoxin ED targets the chemokine receptors CXCR1 and CXCR2 to kill leukocytes and promote infection. *Cell Host & Microbe* 14, 453–459.
- Romano, P., Gutierrez, M., Berón, W., Rabinovitch, M., Colombo, M., 2007. The autophagic pathway is actively modulated by phase II *Coxiella burnetii* to efficiently replicate in the host cell. *Cellular Microbiology* 9, 891–909.
- Ruiz-Albert, J., Yu, X.-J., Beuzón, C., *et al.*, 2002. Complementary activities of SseJ and SifA regulate dynamics of the *Salmonella typhimurium* vacuolar membrane. *Molecular Microbiology* 44, 645–661.
- Smith, G.A., Marquis, H., Jones, S., *et al.*, 1995. The two distinct phospholipases C of *Listeria monocytogenes* have overlapping roles in escape from a vacuole and cell-to-cell spread. *Infection and Immunity* 63, 4231–4237.
- Starr, T., Child, R., Wehrly, T., *et al.*, 2012. Selective subversion of autophagy complexes facilitates completion of the *Brucella* intracellular cycle. *Cell Host & Microbe* 11, 33–45.
- Tan, Y., Arnold, R.J., Luo, Z.Q., 2011. *Legionella pneumophila* regulates the small GTPase Rab1 activity by reversible phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* 108, 21212–21217.
- Tan, Y., Luo, Z.Q., 2011. Take it and release it: The use of the Rab1 small GTPase at a bacterium's will. *Cellular Logistics* 1, 125–127.
- Tattoli, I., Sorbara, M.T., Vuckovic, D., *et al.*, 2012. Amino acid starvation induced by invasive bacterial pathogens triggers an innate host defense program. *Cell Host & Microbe* 11, 563–575.
- Tattoli, I., Sorbara, M.T., Yang, C., *et al.*, 2013. *Listeria* phospholipases subvert host autophagic defenses by stalling pre-autophagosomal structures. *EMBO Journal* 32, 3066–3078.
- Vergne, I., Chua, J., Lee, H.-H., *et al.*, 2005. Mechanism of phagolysosome biogenesis block by viable *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America* 102, 4033–4038.
- Via, L., Deretic, D., Ulmer, R., *et al.*, 1997. Arrest of mycobacterial phagosome maturation is caused by a block in vesicle fusion between stages controlled by rab5 and rab7. *Journal of Biological Chemistry* 272, 13326–13331.
- Whitworth, T., Popov, V.L., Yu, X.J., Walker, D.H., Bouyer, D.H., 2005. Expression of the *Rickettsia prowazekii* pld or tlyC gene in *Salmonella enterica* serovar Typhimurium mediates phagosomal escape. *Infection and Immunity* 73, 6668–6673.
- Yarborough, M.L., Li, Y., Kinch, L.N., *et al.*, 2009. AMPylation of Rho GTPases by *Vibrio* VopS disrupts effector binding and downstream signaling. *Science* 323, 269–272.
- Zoppino, F.C., Militello, R.D., Slavin, I., Alvarez, C., Colombo, M.I., 2010. Autophagosome formation depends on the small GTPase Rab1 and functional ER exit sites. *Traffic* 11, 1246–1261.

Cellular Invasion by Bacterial Pathogens

ML Ferrari, Institut Pasteur, Paris Cédex, France

PJ Sansonetti, Institut Pasteur, Paris Cédex, France and Collège de France, Paris, France

© 2016 Elsevier Inc. All rights reserved.

Introduction

Bacterial pathogens evolved similar traits to manipulate host cell functions for their own benefit and to evade or manipulate the host immune system. The ability to invade cells and live intracellularly is one of these strategies developed in a process of coevolution between the bacteria and the host. Unlike the process of phagocytosis, in which a microbe is engulfed by professional phagocytic cells, invasion refers to the entry process whereby a microbe actively triggers its own uptake by non-phagocytic cells. Depending on their capacity to survive and replicate outside their host, invasive bacteria can be obligate intracellular organisms, like *Rickettsia* spp. or *Chlamydia* spp., or facultative intracellular organisms, like *Salmonella* spp., *Shigella* spp., or *Listeria monocytogenes*. For all these bacteria, however, invasion is a feature of their pathogenesis. On the other hand, an increasing number of bacteria which are generally considered exclusively extracellular can be found in some circumstances inside several types of cells, a feature which has been linked to their capacity for persistence and their pathogenesis (Schwarz-Linek *et al.*, 2006). Some examples of this last group are *Streptococcus pyogenes* and *Staphylococcus aureus*.

In a general way, the invasion process involves four main stages: entry, survival, replication, and exit or spreading. Pathogenic invasive bacteria enter cells by binding to host plasma membrane receptors or injecting bacterial proteins and promoting their uptake by actin remodeling. Once inside, they either transform the internalization vacuole into a replicative niche by manipulating its maturation, adapt to the intravacuolar environment, or they escape from the vacuole and adopt a cytoplasmic lifestyle. During the infectious process,

they develop mechanisms to avoid killing by cellular innate immunity and to manipulate inflammatory responses, cell survival pathways, autophagy, or intracellular trafficking.

Entry Mechanisms

The invasion mechanisms are historically classified based on two main morphological rearrangements of the host plasma membrane, as it was firstly proposed for the general phagocytic process. In the first mechanism, called the 'zipper mechanism,' bacterial surface molecules, generally called adhesins, bind to host plasma membrane receptors with high affinity to trigger the intracellular signaling needed for their uptake. The host cell receptors are typically transmembrane adhesion molecules involved in cell-to-cell or cell-extracellular matrix interactions. During this entry process, the cell membrane tightly surrounds the bacteria while undergoing subtle cytoskeleton rearrangements (Figure 1(a)). The second mechanism is called the 'trigger mechanism,' in which bacteria use a type III secretion system (TTSS) to inject virulence proteins, or effectors, into the host plasma membrane and cytoplasm, triggering massive actin polymerization, with membrane ruffling and filopodia forming the entry focus (Figure 1(b)).

Despite the differences between these two entry mechanisms, three common events can be distinguished: (1) contact with the host cell, (2) induction of actin remodeling and formation of an internalization vacuole, and (3) actin depolymerization and vacuole closure.

These mechanisms will be described in detail in the following sections.

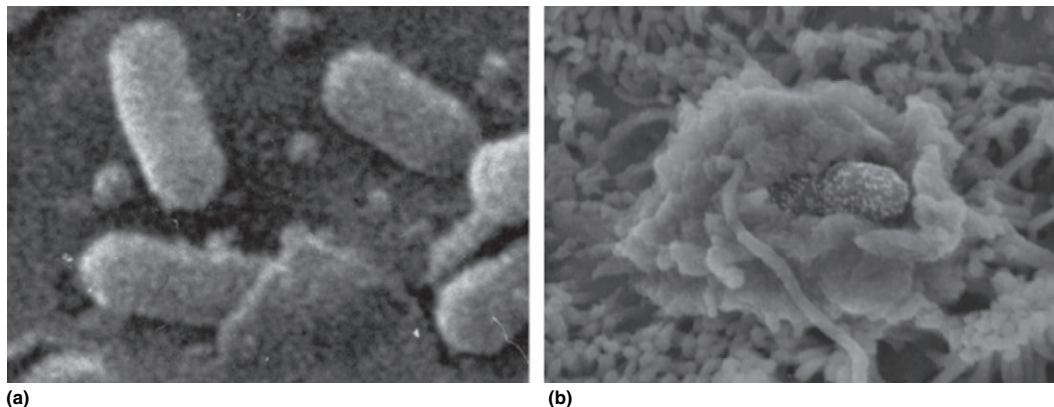


Figure 1 Mechanisms of bacterial invasion. (a) The 'zipper mechanism.' Scanning electron micrograph of *L. monocytogenes* entering into epithelial cells. Reprinted from Mengaud, J., Ohayon, H., Gounon, P., Mege, R.-M., Cossart, P., 1996. E-cadherin is the receptor for internalin, a surface protein required for entry of *L. monocytogenes* into epithelial cells. Cell 84 (6), 923–932. (b) The 'trigger mechanism.' Scanning electron micrographs of *Salmonella* entry focus in polarized epithelial cells. Courtesy of N. Lhocine.

Zipper

The best-known examples of bacteria using the zipper entry process are the enteropathogens *L. monocytogenes*, *Yersinia pseudotuberculosis*, and *Yersinia enterocolitica*. These bacteria evolved a similar mechanism to induce their uptake, yet with some variations in the molecular machineries involved in the process.

The *Y. pseudotuberculosis* adhesin, called invasins, was the first bacterial adhesion molecule identified to bind a host mammalian receptor, the $\beta 1$ chain of integrins (Isberg and Leong, 1990). Though *Yersinia* is able to impair its uptake by phagocytic cells through the injection of antiphagocytic effectors by its TTSS, it is able to invade epithelial cells using invasins during the first steps of infection whereby the bacteria traverse the intestinal epithelial lining. Invasins mimics fibronectin binding to the adhesion molecule $\beta 1$ -integrin but with higher affinity, inducing receptor clustering that triggers the downstream integrin signaling cascade. Thereby, *Yersinia* entry involves the repertoire of kinases and adaptor proteins participating in focal adhesion dynamics, such as focal adhesion kinase (FAK), Src, Pyk2, Cas, and Crk, with the consequent activation of the RhoGTPase Rac1, which directs Arp2/3-mediated cytoskeleton remodeling around the bacteria (Wong and Isberg, 2005).

Yersinia pseudotuberculosis can also enter cells through another adhesin, YadA, presumably after crossing the intestinal epithelium and in cases where invasins expression is repressed. This adhesin, unlike invasins, binds proteins from the extracellular matrix that in turn bind to host cell $\beta 1$ -integrins (Heise and Dersch, 2006). A similar strategy is used by the two facultative invasive pathogens, *S. pyogenes* and *S. aureus*, which express fibronectin-binding proteins on their cell wall that permit cellular invasion. Notably, the resulting intracellular signaling cascade is similar to the one induced by *Yersinia* invasins (Schwarz-Linek et al., 2006).

In the case of *L. monocytogenes*, two adhesins harboring leucine-rich repeats were described as mediators of invasion, each of them binding to different host receptors: internalin A (InlA) and internalin B (InlB). InlA is a bacterial membrane-attached protein that binds the extracellular domain of the host adherens junction adhesion molecule E-cadherin (Menga et al., 1996). InlB, on the other hand, is a loosely attached surface protein that induces the invasion of cells mainly by interaction with the extracellular domain of the Met receptor tyrosine kinase, the receptor for the hepatocyte growth factor (HGF) (Shen et al., 2000). *In vivo*, InlA-mediated invasion is restricted to cells expressing E-cadherin and is responsible for translocation across the intestinal epithelium barrier, whereas InlB is necessary, in conjunction with InlA, for translocation across the placenta (Disson et al., 2008).

Listeria monocytogenes entry through InlA mimics the signaling induced in adherens junctions formation. For example, both processes are dependent on binding of the cytoplasmic domain of E-cadherin to β -catenin and α -catenin, and involve myosin VII and the surface molecule vezatin, molecules bridging adherens junctions with the actin cytoskeleton and with motor proteins (Pizarro-Cerdá and Cossart, 2009). ARHGAP10, a Cdc42/Rho guanine-activating protein (GAP), is involved in α -catenin recruitment to adherens junctions and also plays a role during infection (Sousa et al., 2005). As ARHGAP10 binds also Arf6, which in turn interacts

with vezatin, it was proposed that Arf6 could mediate the coupling between the tracking force exerted by myosin VIIA and adherens junctions (Sousa et al., 2005; Figure 2(a)).

Conversely, InlB phenocopies the HGF-induced signaling cascade, inducing Met receptor dimerization, autophosphorylation of the cytoplasmic domain, and the recruitment and phosphorylation of several adaptors, such as Cbl, Gab1, CrkII, and Shc, with the ensuing activation of phosphatidylinositol (PI)-3-kinase (PI3K) (Dokainish et al., 2007; Pizarro-Cerdá and Cossart, 2009; Figure 2(b)).

In both entry pathways, an Arp2/3-dependent local actin polymerization takes place. In the case of InlA-mediated entry, the process is initiated by Src kinase activation, which in turn phosphorylates cortactin. Rac1 is also involved in this entry mechanism. In InlB-mediated entry, on the other hand, this event is triggered by PI3K activation. PI3K induces PI(3,4,5)P₃ production at the entry site and the local activation of Rac1/Cdc42, with the ensuing recruitment of N-WASP/WAVE (Sousa et al., 2007; Pizarro-Cerdá and Cossart, 2009). Src also phosphorylates the cytoplasmic domain of E-cadherin, inducing its subsequent ubiquitination by Hakai ubiquitin ligase (Bonazzi et al., 2008).

Interestingly, the ubiquitination of E-cadherin by Hakai and that of Met and $\beta 1$ -integrin by Cbl ubiquitin ligases, respectively, induces the recruitment of clathrin to the site of entry, with the participation of dynamin-2 and noncanonical clathrin adaptors. This recruitment happens upstream of actin polymerization, and also appears to be induced by *S. aureus* (Veiga et al., 2007). However, it is still not clear whether clathrin is involved directly in the engulfment of the bacteria or rather serves as a signaling platform for the recruitment of the actin machinery. Moreover, caveolin-1, the main component of plasmalemmal caveolae, is important for E-cadherin clustering and InlA-bacterial entry (Bonazzi et al., 2008). Thus, several types of endocytic machineries seem to be involved in *L. monocytogenes* entry.

Closure and scission of the internalization vacuole formed during *Yersinia* invasion involve several steps, including the activation of PI-4-phosphate-5-kinase (PIP5K) by Rac1 and Arf6, inducing localized PI(4,5)P₂ production, and subsequent PI(4,5)P₂ consumption by PI3K and the 5-phosphatases OCRL1 and Inpp5b, which are recruited in a Rab5-dependent manner (Sarantis et al., 2012). Similarly, the last step in *L. monocytogenes* invasion is mediated by the fine regulation of cofilin activity by LIM-kinase, and by OCRL1. Actin depolymerization by cofilin and the consumption of PI(4,5)P₂ by OCRL1 and PI3K help the scission process (Pizarro-Cerdá et al., 2015).

Another example of bacteria invading host cells through the zipper mechanism is provided by Rickettsiae, which attach to the host receptor Ku70 on endothelial cells via the surface protein OmpB, inducing an intracellular signaling cascade and bacterial uptake in a way similar to *L. monocytogenes* and *Yersinia* spp. (Martinez et al., 2005).

Trigger

Salmonella, *Shigella*, and *Chlamydia* use a TTSS to inject effectors into the host cell in order to trigger their uptake. The TTSS molecular machinery comprises a cytoplasmic domain, a basal

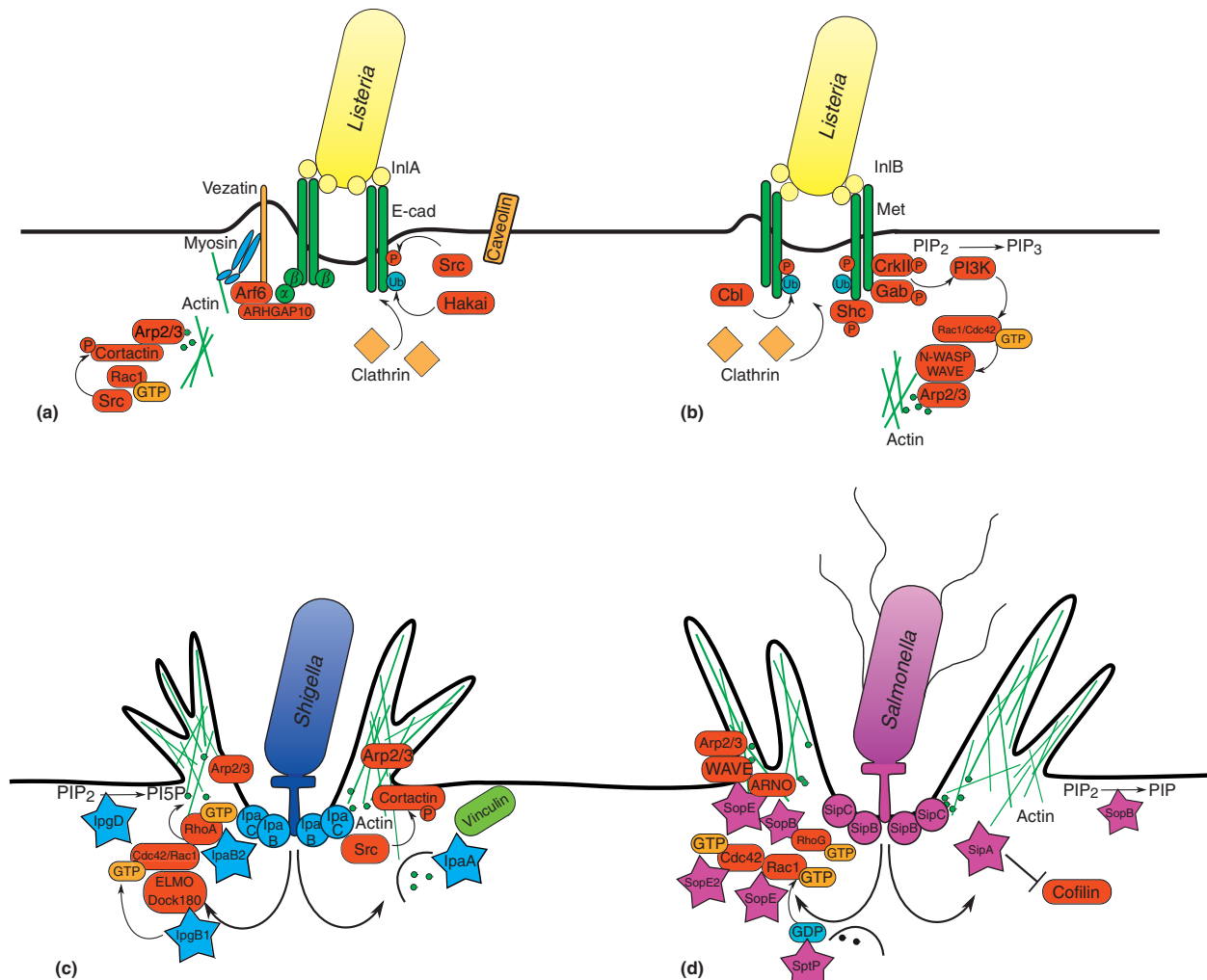


Figure 2 Signaling pathways triggered by bacterial invasion. (a) *Listeria monocytogenes* InlA-mediated entry. Inl A interacts with E-cadherin, mimicking the signaling induced in adherens junctions formation. E-cadherin cytoplasmic domain binds β -catenin and α -catenin, which recruits ARHGAP10, vezatin, and myosin VII. Arf6 interacts with ARHGAP10 and vezatin, bridging adherens junctions with actin cytoskeleton. Caveolin is involved in Src activation. Activated Src phosphorylates E-cadherin, inducing subsequent ubiquitination by Hakai and clathrin recruitment. Src also phosphorylates cortactin, inducing Arp2/3-dependent actin polymerization. Rac1 is also activated inducing actin dynamics. (b) *Listeria monocytogenes* InlB-mediated entry. InlB binds and induces activation of Met receptor. The intracellular signaling cascade involves phosphorylation and ubiquitination of Met, which recruits clathrin; recruitment and phosphorylation of adaptors such as Cbl, Shc, Gab, and Crkl; and activation of PI3K, which triggers Rac1/Cdc42 activation and Arp2/3-dependent actin polymerization. (c) *Shigella* invasion. *Shigella* injected effectors target host RhoGTPases and kinases triggering massive actin polymerization. Cortactin activation occurs downstream of Src activation by IpaC. IpgB1 and IpgB2 activate Cdc42, Rac1, and RhoA GTPases. IpgD converts PIP2 into PI5P, inducing cortical actin dissociation from plasma membrane. IpaA binds the focal adhesion protein vinculin, promoting actin remodeling and disassembly. (d) *Salmonella* 'trigger' mechanism. Similar to *Shigella*, *Salmonella* injects effectors inducing its uptake. SopE and SopE2 are RhoGEF, activating Cdc42 and Rac1. SipA inhibits cofilin, stabilizing actin filaments. SipC is an actin nucleator and crosslinker of actin filaments. SopB is a PI phosphatase that can activate RhoG, and induces ARNO/Arf GEF activation, WAVE activation, and Arp2/3-dependent actin polymerization. SptP is a RhoGAP that counteracts SopE/E2 RhoGEF activities, inducing actin disassembly.

body, a needle-like structure, and the translocator complex, the latter forming a pore in the plasma membrane of the target cells to allow effector injection (Cornelis, 2010). *Shigella* and *Chlamydia* possess one TTSS encoded on a large virulence plasmid and on the chromosome, respectively. *Salmonella*, on the other hand, has two TTSS encoded on two different genomic *Salmonella* pathogenicity islands (SPI): SPI-1 and SPI-2. TTSS assembly and activation are initiated by environmental

stimuli such as temperature, pH, oxygen tension, bile salts, and contact with the target cell.

Bacterial attachment

The first interaction with the cell can be accomplished by different means. *Chlamydia*, for example, has no identified adhesin. However, the PDGF-receptor β , heparan sulfate glycosaminoglycans, and protein disulfide isomerase on the

target cell surface are involved in the attachment of chlamydial elementary bodies (EBs) (Bastidas *et al.*, 2013). The EB is the non-replicative, infectious form of the bacteria, in contrast to the reticular body (RB), which is the intracellular and replicative form.

For *Shigella*, two distinct mechanisms of adhesion to epithelial cells have been described. Firstly, bile salts in the intestinal lumen can induce adhesion in a TTSS-dependent manner, promoting the attachment via polar IcsA (Zumsteg *et al.*, 2014). Second, the interaction can occur between the tip of the TTSS and a cell surface receptor present in the filopodia, followed by subsequent filopodia retraction, bringing bacteria close to the cell body (Romero *et al.*, 2011).

On the other hand, *Salmonella*, which is a flagellated bacterium, possesses several fimbriae and adhesins that interact with cells in addition to the SPI-1 TTSS, which induces an irreversible interaction with the host membrane (Lara-Tejero and Galán, 2009). A 'near surface swimming' model proposes that *Salmonella* swims by flagellar movement to scan the host surface topology, docking at sites of prominent topology by irreversible SPI-1 TTSS-mediated interaction, thereby explaining the site selection and the cooperative infection (Misselwitz *et al.*, 2012). Bacterial docking to the membrane allows the assembly of the translocation complex, activating the coordinated secretion of prestored cytosolic TTSS invasion effectors.

Activation of secretion

The proteins forming the translocon of the TTSS are IpaB/IpaC in *Shigella* and their homologs SipB/SipC in *Salmonella*. Plasma membrane cholesterol and lipid microdomains are important for translocon assembly and activation. IpaB and SipB, for example, interact directly with cholesterol. The TTSS activation may occur by induction of architectural changes in the needle that allow the assembly of the translocon and the initiation of injection, and/or through the local enrichment of receptors important for entry (Lafont and van der Goot, 2005). Moreover, *Salmonella* invasion is enhanced, in a SipB-dependent manner, in mitotic cells due to the cholesterol enrichment in the outer leaflet of the plasma membrane during mitosis (Santos *et al.*, 2013). However, there are contradictory reports regarding this topic (Gilk *et al.*, 2013) and its *in vivo* relevance remains unclear.

Actin and membrane rearrangement

Rearrangement of the actin cytoskeleton and formation of membrane protrusions are induced by TTSS-injected effectors that manipulate, either directly or indirectly, the main regulators of these processes: RhoGTPases, phosphoinositides, and tyrosine kinases. *Shigella* IpgB1 and IpgB2 effectors, for example, are guanine nucleotide exchange factors (GEFs) for Cdc42/Rac1 and RhoA, respectively (Huang *et al.*, 2009), but IpgB1 can also activate Rac1 through the recruitment of the ELMO/Dock180 machinery (Handa *et al.*, 2007). As a result, Rac1 activation induces downstream Arp2/3-dependent actin polymerization and ruffling, whereas RhoA induces stress fiber formation. IpgB2 activity, in conjunction with the other Rho-activating effector IpaA, might promote bacterial anchorage (Carayol and Tran Van Nhieu, 2013). The IpaC C-terminal domain, on the other hand, recruits and activates Src, which in turn phosphorylates cortactin, amplifying actin polymerization (Mounier

et al., 2009). IpaC also induces activation of Rac1/Cdc42, which is probably mediated by Src-Crk or through Abl kinase activity, another kinase activated during the invasion process (Carayol and Tran Van Nhieu, 2013; Schroeder and Hilbi, 2008). Actin remodeling is enhanced by the PI-phosphatase IpgD. This effector converts PI(4,5)P₂ into PI5P and leads to the dissociation of the plasma membrane from the actin cytoskeleton (Niebuhr *et al.*, 2002; Figure 2(c)).

For *Salmonella*, five effectors secreted by the SPI-1 TTSS are critical for entry. SipC and SipA interact directly with actin: SipC nucleates actin and crosslinks actin filaments, and SipA stabilizes the latter and inhibits cofilin. The bacteria also secrete two RhoGEFs, SopE and SopE2. They activate Rac1 and Cdc42, inducing the branching of actin filaments in an Arp2/3-dependent process (Patel and Galán, 2006; Ly and Casanova, 2007). SopB, which is a pleiotropic PI phosphatase, consumes PI(4,5)P₂ at the plasma membrane, leading to the same effect than its homologous IpgD. Additionally, SopB induces PI3P production and can indirectly activate RhoG through the SH3-containing GEF (Patel and Galán, 2006). Local production of PI3P also induces ARNO/Arf GEF recruitment, which in turn activates Arf. WAVE, which is recruited by SopE, is further activated by Arf and Rac1, finally triggering actin polymerization (Humphreys *et al.*, 2012; Figure 2(d)).

SipC recruits to the plasma membrane the exocyst complex component Exo70, promoting the docking and fusion of exocytic vesicles. Exocyst complex activation is mediated by RalA GTPase via SopE (Nichols and Casanova, 2010). This process stimulates membrane expansion at the bacterial entry site.

For *Chlamydia trachomatis* the injected effector TARP is the only one so far with known functions in bacterial entry: it is an actin nucleator (Jewett *et al.*, 2006) and, through phosphorylation by Src/Abl kinases, binds host proteins involved in Arp2/3-dependent actin remodeling such as Rac1, Rac1-GEFs, WAVE2, and PI3K (Lane *et al.*, 2008). Rac1 activation is critical for *C. trachomatis* entry; however, TARP phosphorylation is not essential, meaning that Rac1 can be activated by an additional and yet uncharacterized mechanism (Bastidas *et al.*, 2013; Dumoux *et al.*, 2014).

Actin depolymerization and vacuole closure

The disassembly of actin filaments and the end of the entry process is promoted in *Shigella* by the effector IpaA, which binds the focal adhesion protein vinculin (Carayol and Tran Van Nhieu, 2013; Figure 2(c)). The *Chlamydia* effector CT694 may perform a similar function (Bastidas *et al.*, 2013). The *Salmonella* PI-phosphatase SopB induces PI(4,5)P₂ consumption, which is critical for membrane fission during vacuole closure (Terebiznik *et al.*, 2002). Furthermore, the effector SptP is a GAP for RhoGTPases, counteracting SopE/E2 RhoGEF activities (Figure 2(d)). The sequential activity of SopE and SptP has been suggested to be regulated by differential half-life or secretion rates (Van Engelenburg and Palmer, 2008; Kubori and Galán, 2003).

It should be noted that even if we distinguish two general entry mechanisms, there is increasing evidence suggesting more than one entry mechanism for some bacteria. *Salmonella*, for example, can also enter cells by an Arp2/3-independent mechanism whereby SopB activates RhoA and the downstream effector Rho kinase to induce a myosin II, contractility-dependent

pathway (Hänisch *et al.*, 2011). In addition, the *Salmonella* proteins Rck and PagN are involved in TTSS-independent, zipper-like entry mechanisms (Velge *et al.*, 2012). Another example is *Chlamydia* (Dumoux *et al.*, 2014), which injects protein effectors that remodel the actin cytoskeleton. At the same time, it was proposed that the clathrin machinery is important for *Chlamydia* entry (Hybiske and Stephens, 2007). Furthermore, *Chlamydia* binds and activates PDGF-receptor β (Bastidas *et al.*, 2013). Both clathrin-mediated entry and receptor binding constitute features generally proposed to be part of the zipper mechanism (Veiga *et al.*, 2007).

Intracellular Lifestyle and Manipulation of Host Intracellular Trafficking

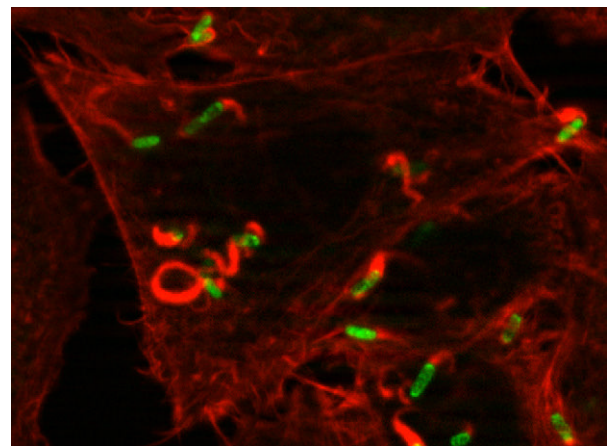
Invasive pathogenic bacteria have evolved mechanisms to survive inside cells after internalization into a primary vacuole. The great majority are able to live and replicate inside a pathogen-containing vacuole, either by blocking vacuole maturation or by adapting to harsh intravacuolar conditions (Kumar and Valdivia, 2009). However, some intracellular bacteria escape from the vacuole to the host cell cytoplasm, which is thought to be a richer source of nutrients than vacuoles and devoid of the antimicrobial mechanisms of phagosomes (Ray *et al.*, 2009). Despite the apparently safer conditions, the cytoplasm contains a range of antimicrobial, autophagic, and inflammatory factors that cytoplasmic bacteria need to avoid in order to survive. Indeed, the cytoplasm has been shown to be nonpermissive for replication of a variety of intravacuolar bacteria (Ray *et al.*, 2009).

Interestingly, whether invasive bacteria are vacuole-dwelling or live in the cytosol is independent from the mechanism of entry and from the property of being obligate or facultative intracellular bacteria (Casadevall, 2008). For example, both *Shigella* and *Salmonella* use the trigger mechanism to invade target cells, but *Shigella* lyses the vacuole and escapes to the cytoplasm, while *Salmonella* stays inside a *Salmonella*-containing vacuole (SCV). *Listeria monocytogenes*, on the other hand, uses the zipper mechanism for its entry but, just as *Shigella*, lyses the vacuole and moves into the cytosol using actin-based motility (ABM). In addition, obligate intracellular bacteria such as *Rickettsia* escape into the cytosol, while other obligate intracellular pathogens like *Coxiella* or *Chlamydia* remain in a vacuolar compartment (Figure 3).

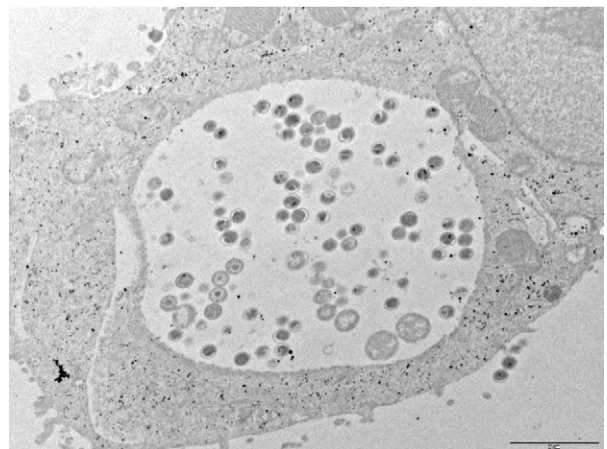
Intravacuolar Bacteria

Bacteria inhabiting a vacuolar compartment hijack host membrane trafficking pathways not only to avoid the cellular immune response, but also to acquire nutrients for their replication, and membrane to enlarge their vacuole. Vacuole maturation and integrity are regulated by the interplay between several factors, such as the cytoskeletal network of microtubules, actin and associated motor proteins, the lipid composition of the vacuole, host GTPases as well as bacterial effectors, which tightly regulate these processes for their own benefit (Figure 4).

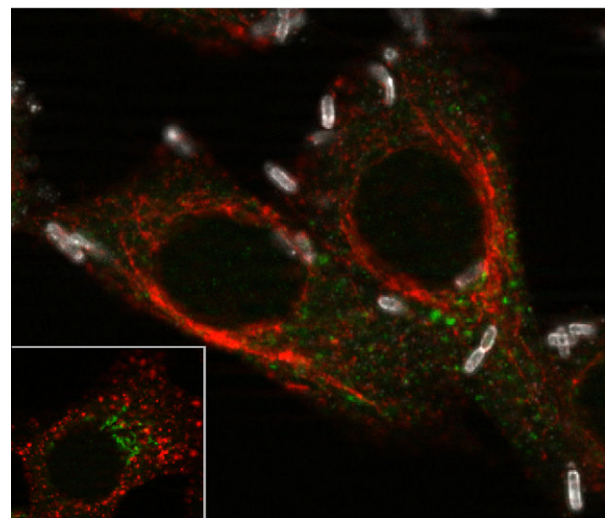
After internalization, *Salmonella* manipulates the host endolysosomal pathway to establish the SCV as a replicative



(a)



(b)



(c)

Figure 3 Bacterial intracellular lifestyles. (a) Fluorescence microscopy of *Shigella* ABM (red: actin; green: bacteria). (b) Electron micrograph of *Chlamydia* inclusion with replicating bacteria (courtesy of A. Subtil). (c) Fluorescence microscopy of Golgi fragmentation and recycling endosomes tubulation induced by *Shigella* invasion. Inset, noninfected cell (gray: bacteria; green: GM130, Golgi marker; red: transferrin, recycling endosomes).

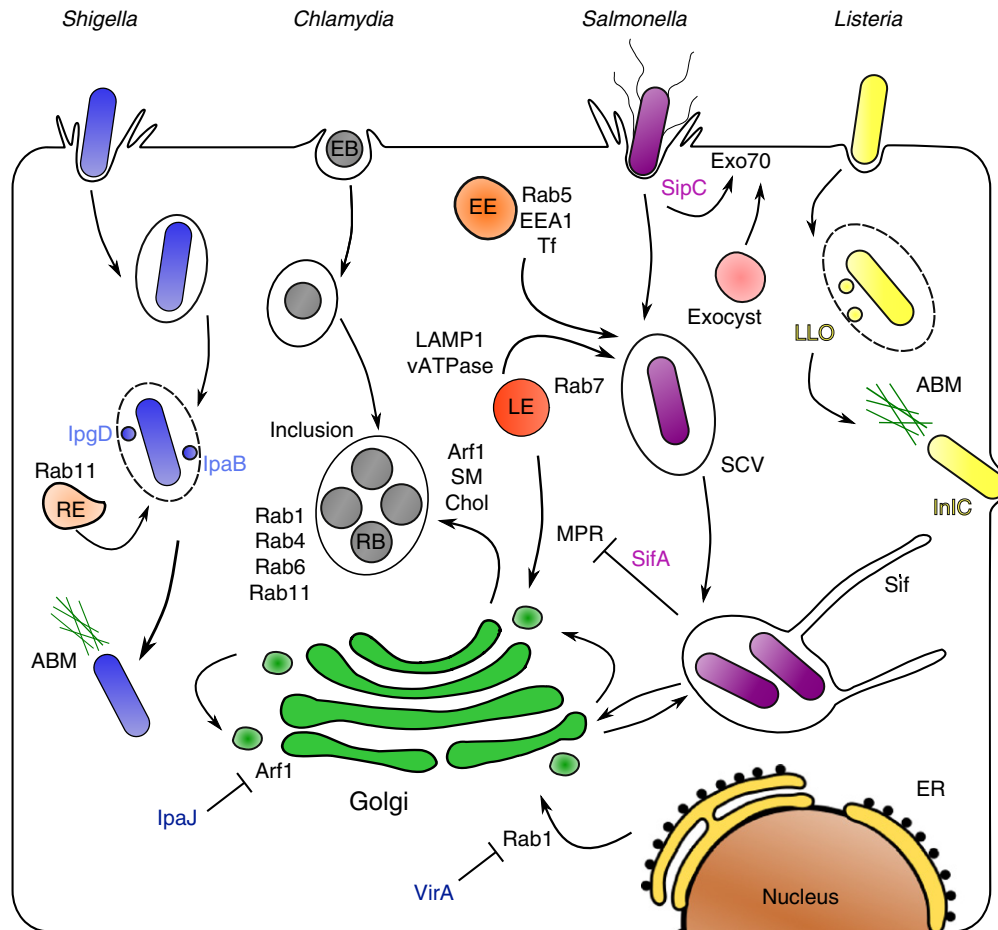


Figure 4 Schematic representation of bacterial intracellular lifestyles and manipulation of host intracellular trafficking. *Shigella* and *L. monocytogenes* lyse the vacuole after entry and move into the cytosol by ABM. *Shigella* vacuolar escape is mediated by the translocon protein IpaB, and is enhanced by the IpgD-dependent recruitment of Rab11 endosomes. *Shigella* effector IpaJ targets Arf1 inducing Golgi fragmentation, whereas VirA inactivates Rab1 blocking ER–Golgi transport. *Listeria monocytogenes* induces vacuolar disruption by secretion of the pore-forming toxin listeriolysin O (LLO). *Listeria monocytogenes* InlC protein promotes bacteria spreading to adjacent cells. *Chlamydia* and *Salmonella* remain in a vacuolar compartment, the *Chlamydia* inclusion and the SCV, respectively, which acquire different protein and lipid markers. *Salmonella* SipC recruits Exo70 to the entry site, promoting the docking and fusion of exocytic vesicles. SifA inhibits retrograde trafficking of MPR, attenuating lysosome function. Chol, cholesterol; EB, elementary body; EE, early endosomes; ER, endoplasmic reticulum; LE, late endosomes; RB, reticular body; RE, recycling endosomes; SM, sphingomyelin.

niche. First, the SCV acquires early endosomal markers such as early endosomal antigen 1 (EEA1), transferrin receptor (TfR), and Rab5. Then, the SCV matures to a vATPase- and LAMP1-positive compartment, in a Rab7-dependent mechanism (Malik-Kale *et al.*, 2011). The manipulation of RabGTPase dynamics during SCV maturation is dependent on the PI-phosphatase activity of SopB, which regulates the PI surface content on the SCV and consequently the negative charge of the SCV membrane, thereby regulating the recruitment and detachment of RabGTPases (Pizarro-Cerdá *et al.*, 2015). In addition, SifA blocks Rab9-dependent mannose-6-phosphate receptor (MPR) retrograde traffic, thereby affecting lysosomal maturation and avoiding the acquisition of lysosomal hydrolases by the SCV (McGourty *et al.*, 2012). During the maturation process, the SCV moves toward the microtubule organizing center (MTOC) and the peri-Golgi region (Salcedo and Holden, 2003), and develops *Salmonella*-induced filaments (Sifs), which expand to the cell periphery. These

processes are dependent on interactions with the cytoskeleton and dynein–kinesin motors and on the activity of several bacterial effectors, including SifA. A network of F-actin surrounds the SCV and provides structural stability; thus, its depolymerization induces loss of SCV integrity (Malik-Kale *et al.*, 2011). Sifs interact also with the Golgi apparatus owing to SPI-2 TTSS effectors, probably to acquire nutrients and lipids from the secretory pathway that are important for bacterial replication and vacuole enlargement or as a strategy to manipulate host immunity (Mota *et al.*, 2009). Interestingly, cholesterol is excluded from the SCV by the enzymatic activity of the SseJ effector, modifying the SCV membrane properties and allowing Sif formation (Lossi *et al.*, 2008).

Similarly, the *Chlamydia*-containing vacuole, called inclusion (Figure 3(b)), moves to the MTOC using dynein (Grieshaber *et al.*, 2003), but without acquiring endolysosomal markers. This large and rigid organelle interacts with components from the Golgi apparatus, lipid droplets,

multivesicular bodies, and mitochondria to efficiently acquire a large variety of nutrients and polar lipids, cholesterol, and sphingomyelin (Bastidas *et al.*, 2013). In addition, F-actin forms a structural scaffold with intermediate filaments around the *Chlamydia* inclusion (Kumar and Valdivia, 2008). Several Rabs interact with the inclusion, including Rab 4, 11, and 14 that are involved in receptor recycling, and Rab 1 and 6, which are implicated in endoplasmic reticulum (ER)–Golgi transport. Remarkably, *Chlamydia* infection induces Golgi fragmentation into mini-stacks that surround the inclusion (Heuer *et al.*, 2008) in a process regulated by Rab6 and Rab11 (Rejman Lipinski *et al.*, 2009), but independent of *Chlamydia* CFAP protease activity (Snaveley *et al.*, 2014), as previously suggested (Heuer *et al.*, 2008). Other important recruited factors are proteins involved in PI4P production, such as OCRL1, Arf1, and its downstream interactor PI4KII α , all of which are critical for inclusion growth (Bastidas *et al.*, 2013). Furthermore, *Chlamydia* secretes TTSS effectors, particularly the Inc proteins, which localize extensively to the inclusion membrane. These effectors are involved in diverse functions including Rab recruitment, inclusion biogenesis, homotypic fusion of the inclusion, and sphingomyelin acquisition (Bastidas *et al.*, 2013).

Occasionally, intravacuolar bacteria escape into the cytosol where they can replicate. In the case of *Salmonella*, for example, this event is accompanied with epithelial cell extrusion and may be a strategy for bacterial dissemination (Knodler *et al.*, 2010).

Cytosolic Bacteria

Bacteria from this group are able to escape from the internalization vacuole, replicate in the cytosol, and avoid cytosolic innate immunity for their further dissemination (Figure 4).

The escape from the vacuole is a very fast process driven by secreted bacterial factors. *Listeria monocytogenes* uses the pore-forming toxin listeriolysin O (LLO) and type C phospholipases (PLC). LLO is a thiol-activated and cholesterol-dependent toxin that, after activation by the acidic conditions of the maturing vacuole, inserts into the membrane to form large pores that inhibit vacuole maturation (Pizarro-Cerdá and Cossart, 2009). Vacuolar disruption is further promoted by PLCs, which destabilize the enclosing membrane. *Rickettsia* spp. also express phospholipases and haemolysin C that are probably involved in bacterial escape (Ray *et al.*, 2009).

On the other hand, the *Shigella* TTSS effector IpaB is required for vacuole escape in macrophages (High *et al.*, 1992), likely due to the formation of the bacterial translocon together with the effector IpaC. However, as IpaB/IpaC are essential to invasion, it is experimentally challenging to confirm this function in epithelial cells. It has been proposed that other effectors encoded in the bacterial virulence plasmid but not essential for the entry may promote vacuole rupture (Mellouk *et al.*, 2014; Ray *et al.*, 2009; Schroeder and Hilbi, 2008).

Once in the cytoplasm, bacteria replicate and move in the cytosol by actin-based motility. They express proteins at one pole that induce the polymerization of actin to form actin comet tails (Figure 3(a)). This propels bacteria toward the cell periphery and promotes cell-to-cell spread by induction of protrusions into adjacent cells. *Listeria monocytogenes* ActA mimics N-WASP, directly activating Arp2/3 complex; *Shigella*

IcsA, instead, binds N-WASP to recruit the Arp2/3 complex thereafter. In both cases bacteria promote the polymerization of branched-actin filaments (Iretton, 2013). Interestingly, *Rickettsia* have a slow motility at early times of infection that is driven by RickA and Arp2/3, while at later times it uses Sca2 to mimic formins, thereby inducing long, unbranched actin filaments (Reed *et al.*, 2014).

Membrane protrusions penetrate neighboring cells forming double membrane vacuoles that are lysed through the activity of bacterial and also host factors. In *Shigella*, this process is dependent on the reactivation of the TTSS (Campbell-Valois *et al.*, 2014) and on host motor proteins, adhesion molecules, and a noncanonical clathrin endocytic pathway exploited at tricellular junctions in the neighboring cell (Fukumatsu *et al.*, 2012). *Listeria monocytogenes* secretes InlC after invasion, which promotes formation of protrusions by disrupting complexes between the host proteins TUBA and N-WASP, which generate tension at tight junctions. Hence *L. monocytogenes* may use InlC to decrease membrane tension and promote spreading (Rajabian *et al.*, 2009).

It is interesting to note that cytosolic bacteria, just as vacuole-dwelling pathogens, are able to alter the intracellular trafficking pathways of the host cell (Figure 4). Since cytosolic bacteria already have access to nutrients, targeting trafficking pathways may provide them another kind of benefit. *Shigella*, for instance, recruits Rab11 endosomes to the entry site just after actin focus formation. This recruitment is dependent on the TTSS effector IpgD and enhances bacterial vacuolar escape (Mellouk *et al.*, 2014). Moreover, IpgD activity subverts endocytosis and signaling downstream of the EGF receptor, promoting host cell survival (Ramel *et al.*, 2011). After invasion, *Shigella* induces fragmentation of the Golgi apparatus *in vitro* and *in vivo*, and causes recycling endosome tubulation, thereby blocking TfR recycling and cell secretion (Mounier *et al.*, 2012; Figure 3(c)). IpaJ, a cysteine protease, cleaves the Arf1 N-myristoyl modification inducing Arf1 detachment from the Golgi and thereby its fragmentation (Burnaevskiy *et al.*, 2013). In addition, VirA, blocks ER–Golgi trafficking through its RabGAP activity, leading to Rab1 inactivation. It was also proposed that the VirA GAP activity mediates bacterial escape from autophagy (Dong *et al.*, 2012). On the other hand, soon after entry of *L. monocytogenes* into macrophages, the protein Lmo2459 inhibits Rab5a activation, which indirectly activates the bactericidal NADPH (nicotinamide adenine dinucleotide phosphate) oxidase at the vacuolar surface (Pizarro-Cerdá and Cossart, 2009).

Cellular Responses to Bacterial Invasion

Cellular innate immune defenses are triggered by extracellular or intracellular receptors in response to pathogen-associated molecular patterns (PAMPs) or pathogen-induced host cell danger-associated molecular patterns (DAMPs). Signaling induced by these different molecular patterns can lead to the activation of pathways involved in inflammatory responses, autophagic degradation, or cell death. However, invasive bacteria have evolved strategies to avoid or manipulate these pathways. A few examples of these mechanisms are mentioned below.

Autophagy

The *Shigella* IcsA protein is not only the bacterial factor responsible for ABM, but is also recognized by the Atg5 autophagy protein. It was proposed, however, that the *Shigella* TTSS effector IcsB binds IcsA and blocks its binding to Atg5, thus impairing autophagy (Ogawa *et al.*, 2005). This mechanism may function in addition to the VirA-mediated escape from autophagy through Rab1 inactivation mentioned before (Dong *et al.*, 2012). *Listeria monocytogenes* ActA, on the other hand, directly masks autophagic recognition through Arp2/3 recruitment (Yoshikawa *et al.*, 2009) and prevents ubiquitination of bacteria. *Listeria monocytogenes* InlK binds the major vault protein, which also shelters bacteria from recognition by the autophagic machinery (Dortet *et al.*, 2011).

Coxiella burnetii and *Brucella abortus*, on the other hand, are examples of intravacuolar bacteria manipulating autophagy for their own benefit (Kumar and Valdivia, 2009). *Coxiella burnetii* is an extreme case of a pathogen adapted to live and replicate in a parasitophorous vacuole, which is an autophagolysosome-like acidic compartment. The *Brucella*-containing vacuole matures into a vacuole with ER markers and some autophagy-initiation markers. In both cases, stimulation of autophagy increases bacterial replication.

Inflammatory Response

Shigella provides some excellent examples of inflammation modulation. For example, *Shigella flexneri* regulates, both *in vitro* and *in vivo*, the epithelial release of ATP, which is a proinflammatory danger signal upstream of cytokine production. ATP is released during infection with *Shigella*, *Salmonella*, and enteropathogenic *Escherichia coli* (EPEC) through opening of connexin hemichannels. However, the *Shigella* IpgD induces hemichannel closure through its PI phosphatase activity, thereby blocking ATP release and dampening tissue inflammation (Puhar *et al.*, 2013). Another interesting strategy is the ability developed by *Shigella* to remodel its lipopolysaccharide (LPS), the prototypical bacterial PAMP. Intracellular replicating bacteria have a hypoacylated version of the lipid A of the LPS, lowering immune recognition and inflammasome activation (Paciello *et al.*, 2013). Finally, *Shigella* secretes several TTSS effectors, in general with enzymatic activities, that directly target MAP-kinase and NF- κ B signaling, thereby counteracting the inflammatory response stimulated by Nod1 receptor recognition of bacterial peptidoglycan (PGN). These effectors are OspF, a phosphothreonine lyase inactivating p38 and Erk2 MAPKs; IpaH9.8, an E3 ligase targeting the NEMO complex and inducing its proteasomal degradation; OspG, which interferes with I κ B α ubiquitination; OspZ, which blocks nuclear translocation of NF- κ B; and finally OspI that deamidates the E2 enzyme required for TRAF6-NF- κ B signaling (Ashida *et al.*, 2011).

Cell Survival

Bacteria can manipulate cell apoptosis as a strategy for promoting not only host cell survival but also its own. The *Shigella* VirA effector, for example, induces p53 degradation by calpain activation at early times of infection, thus blocking the NF- κ B-

p53 proapoptotic signaling pathway (Bergounioux *et al.*, 2012). IpgD-mediated production of PI5P, on the other hand, subverts the endocytic machinery and growth factor receptor activation to induce activation of the PI3-kinase/Akt survival pathway (Pizarro-Cerdá *et al.*, 2015).

Finally, *Shigella* OspE binds to integrin-linked kinase, stabilizing focal adhesion and reinforcing epithelial cell adhesion to the extracellular matrix, as a strategy to prevent rapid cell turnover and protect its replicative niche (Kim *et al.*, 2009).

Concluding Remarks

Invasive bacteria are capable of exploiting the host cell cytoskeleton and many signaling pathways to enter non-phagocytic cells. Whether induced by receptor binding and activation or by injection of bacterial effectors, bacterial uptake by the target cell requires remodeling of the cytoskeleton and the plasma membrane at the entry site, especially in the case of *Salmonella* and *Shigella*. This is achieved by biochemical changes in the lipid composition of membranes and by protein recruitment that will induce changes in fluidity as well as membrane tension to favor bacterial entry. The downstream signaling events or their manipulation help the establishment of a vacuolar niche and the cross talk between factors inducing vacuole stability, as the vacuolar membrane and host cytoskeleton are constantly remodeled to maintain the biochemical and biophysical properties needed for bacterial proliferation.

Vacuolar remodeling occurs also during infection with cytosolic bacteria, but, in contrast to what happens with vacuolar pathogens, leads to vacuolar rupture and bacterial escape. Host kinases and small GTPases from the Rho and Rab families are key host factors that pathogens manipulate to orchestrate each of these steps. Remarkably, invasive bacteria evolved sophisticated strategies to avoid immune responses. In conclusion, pathogen–host coevolution resulted in several traits acquired by pathogenic bacteria to manipulate each of the important factors that are critical for a successful invasion.

See also: Intracellular Infectiology: Infectious Agents: Bacterial Protein Toxins as Tools in Cell Biology and Physiology

References

- Ashida, H., Ogawa, M., Mimuro, H., *et al.*, 2011. *Shigella* are versatile mucosal pathogens that circumvent the host innate immune system. *Current Opinion in Immunology* 23 (4), 448–455.
- Bastidas, R.J., Elwell, C.A., Engel, J.N., Valdivia, R.H., 2013. Chlamydial intracellular survival strategies. *Cold Spring Harbor Perspectives in Medicine* 3 (5), a010256. doi: 10.1101/cshperspect.a010256.
- Bergounioux, J., Elisei, R., Prunier, A.-L., *et al.*, 2012. Calpain activation by the *Shigella flexneri* effector VirA regulates key steps in the formation and life of the bacterium's epithelial niche. *Cell Host and Microbe* 11 (3), 240–252.
- Bonazzi, M., Veiga, E., Pizarro-Cerdá, J., Cossart, P., 2008. Successive post-translational modifications of E-cadherin are required for InlA-mediated internalization of *Listeria monocytogenes*. *Cellular Microbiology* 10 (11), 2208–2222.

- Burnaevskiy, N., Fox, T.G., Plymire, D.A., *et al.*, 2013. Proteolytic elimination of N-myristoyl modifications by the *Shigella* virulence factor IpaJ. *Nature* 496 (7443), 106–109.
- Campbell-Valois, F.-X., Schnupf, P., Nigro, G., *et al.*, 2014. A fluorescent reporter reveals on/off regulation of the *Shigella* type III secretion apparatus during entry and cell-to-cell spread. *Cell Host and Microbe* 15 (2), 177–189.
- Carayol, N., Tran Van Nhieu, G., 2013. Tips and tricks about *Shigella* invasion of epithelial cells. *Current Opinion in Microbiology* 16 (1), 32–37.
- Casadevall, A., 2008. Evolution of intracellular pathogens. *Annual Review of Microbiology* 62 (1), 19–33.
- Cornelis, G.R., 2010. The type III secretion injectisome, a complex nanomachine for intracellular 'toxin' delivery. *Biological Chemistry* 391 (7), 745–751.
- Disson, O., Grayo, S., Huillet, E., *et al.*, 2008. Conjugated action of two species-specific invasion proteins for fetoplacental listeriosis. *Nature* 455 (7216), 1114–1118.
- Dokainish, H., Gavicherla, B., Shen, Y., Ireton, K., 2007. The carboxyl-terminal SH3 domain of the mammalian adaptor CrkII promotes internalization of *Listeria monocytogenes* through activation of host phosphoinositide 3-kinase. *Cellular Microbiology* 9 (10), 2497–2516.
- Dong, N., Zhu, Y., Lu, Q., *et al.*, 2012. Structurally distinct bacterial TBC-like GAPs link Arf GTPase to Rab1 inactivation to counteract host defenses. *Cell* 150 (5), 1029–1041.
- Dortet, L., Mostowy, S., Louka, A.S., *et al.*, 2011. Recruitment of the major vault protein by InlK: A *Listeria monocytogenes* strategy to avoid autophagy. *PLoS Pathogens* 7 (8), e1002168.
- Dumoux, M., Nans, A., Saibil, H.R., Hayward, R.D., 2014. Making connections: Snapshots of chlamydial type III secretion systems in contact with host membranes. *Current Opinion in Microbiology* 23C, 1–7.
- Fukumatsu, M., Ogawa, M., Arakawa, S., *et al.*, 2012. *Shigella* targets epithelial tricellular junctions and uses a noncanonical clathrin-dependent endocytic pathway to spread between cells. *Cell Host and Microbe* 11 (4), 325–336.
- Gilk, S.D., Cockrell, D.C., Luterbach, C., *et al.*, 2013. Bacterial colonization of host cells in the absence of cholesterol. *PLoS Pathogens* 9 (1), e1003107.
- Grieshaber, S.S., Grieshaber, N.A., Hackstadt, T., 2003. *Chlamydia trachomatis* uses host cell dynein to traffic to the microtubule-organizing center in a p50 dynamitin-independent process. *Journal of Cell Science* 116 (Pt 18), 3793–3802.
- Handa, Y., Suzuki, M., Ohya, K., *et al.*, 2007. *Shigella* IpgB1 promotes bacterial entry through the ELMO-Dock180 machinery. *Nature Cell Biology* 9 (1), 121–128.
- Hänisch, J., Kölm, R., Wozniczka, M., *et al.*, 2011. Activation of a RhoA/myosin II-dependent but Arp2/3 complex-independent pathway facilitates *Salmonella* invasion. *Cell Host and Microbe* 9 (4), 273–285.
- Heise, T., Dersch, P., 2006. Identification of a domain in *Yersinia* virulence factor YadA that is crucial for extracellular matrix-specific cell adhesion and uptake. *Proceedings of the National Academy of Sciences of the United States of America* 103 (9), 3375–3380.
- Heuer, D., Lipinski, A.R., Machuy, N., *et al.*, 2008. *Chlamydia* causes fragmentation of the Golgi compartment to ensure reproduction. *Nature* 457 (7230), 731–735.
- High, N., Mounier, J., Prévost, M.C., Sansonetti, P.J., 1992. IpaB of *Shigella flexneri* causes entry into epithelial cells and escape from the phagocytic vacuole. *EMBO Journal* 11 (5), 1991–1999.
- Huang, Z., Sutton, S.E., Wallentang, A.J., *et al.*, 2009. Structural insights into host GTPase isoform selection by a family of bacterial GEF mimics. *Nature Structural & Molecular Biology* 16 (8), 853–860.
- Humphreys, D., Davidson, A., Hume, P.J., Koronakis, V., 2012. *Salmonella* virulence effector SopE and host GEF ARNO cooperate to recruit and activate WAVE to trigger bacterial invasion. *Cell Host and Microbe* 11 (2), 129–139.
- Hybiske, K., Stephens, R.S., 2007. Mechanisms of *Chlamydia trachomatis* entry into nonphagocytic cells. *Infection and Immunity* 75 (8), 3925–3934.
- Ireton, K., 2013. Molecular mechanisms of cell-cell spread of intracellular bacterial pathogens. *Open Biology* 3 (7), 130079. doi: 10.1098/rsob.130079.
- Isberg, R.R., Leong, J.M., 1990. Multiple beta 1 chain integrins are receptors for invasins, a protein that promotes bacterial penetration into mammalian cells. *Cell* 60 (5), 861–871.
- Jewett, T.J., Fischer, E.R., Mead, D.J., Hackstadt, T., 2006. Chlamydial TARP is a bacterial nucleator of actin. *Proceedings of the National Academy of Sciences of the United States of America* 103 (42), 15599–15604.
- Kim, M., Ogawa, M., Fujita, Y., *et al.*, 2009. Bacteria hijack integrin-linked kinase to stabilize focal adhesions and block cell detachment. *Nature* 459 (7246), 578–582.
- Knodler, L.A., Vallance, B.A., Celli, J., *et al.*, 2010. Dissemination of invasive *Salmonella* via bacterial-induced extrusion of mucosal epithelia. *Proceedings of the National Academy of Sciences of the United States of America* 107 (41), 17733–17738.
- Kubori, T., Galán, J.E., 2003. Temporal regulation of salmonella virulence effector function by proteasome-dependent protein degradation. *Cell* 115 (3), 333–342.
- Kumar, Y., Valdivia, R.H., 2008. Actin and intermediate filaments stabilize the *Chlamydia trachomatis* vacuole by forming dynamic structural scaffolds. *Cell Host and Microbe* 4 (2), 159–169.
- Kumar, Y., Valdivia, R.H., 2009. Leading a sheltered life: Intracellular pathogens and maintenance of vacuolar compartments. *Cell Host and Microbe* 5 (6), 593–601.
- Lafont, F., van der Goot, F.G., 2005. Bacterial invasion via lipid rafts. *Cellular Microbiology* 7 (5), 613–620.
- Lane, B.J., Mutchler, C., Khodor, A.I.S., Grieshaber, S.S., Carabeo, R.A., 2008. Chlamydial entry involves TARP binding of guanine nucleotide exchange factors. *PLoS Pathogens* 4 (3), e1000014.
- Lara-Tejero, M., Galán, J.E., 2009. *Salmonella enterica* serovar typhimurium pathogenicity island 1-encoded type III secretion system translocases mediate intimate attachment to nonphagocytic cells. *Infection and Immunity* 77 (7), 2635–2642.
- Lossi, N.S., Rolhion, N., Magee, A.I., Boyle, C., Holden, D.W., 2008. The *Salmonella* SPI-2 effector SseJ exhibits eukaryotic activator-dependent phospholipase A and glycerophospholipid: Cholesterol acyltransferase activity. *Microbiology (Reading, England)* 154 (Pt 9), 2680–2688.
- Ly, K.T., Casanova, J.E., 2007. Mechanisms of *Salmonella* entry into host cells. *Cellular Microbiology* 9 (9), 2103–2111.
- Malik-Kale, P., Jolly, C.E., Lathrop, S., *et al.*, 2011. *Salmonella* — at home in the host cell. *Frontiers in Microbiology* 2, 125. doi: 10.3389/fmicb.2011.00125.
- Martinez, J.J., Seveau, S., Veiga, E., Matsuyama, S., Cossart, P., 2005. Ku70, a component of DNA-dependent protein kinase, is a mammalian receptor for *Rickettsia conorii*. *Cell* 123 (6), 1013–1023.
- McGourty, K., Thurston, T.L., Matthews, S.A., *et al.*, 2012. *Salmonella* inhibits retrograde trafficking of mannose-6-phosphate receptors and lysosome function. *Science* 338 (6109), 963–967.
- Mellouk, N., Weiner, A., Aulner, N., *et al.*, 2014. *Shigella* subverts the host recycling compartment to rupture its vacuole. *Cell Host and Microbe* 16 (4), 517–530.
- Mengaud, J., Ohayon, H., Gounon, P., Mege, R.-M., Cossart, P., 1996. E-cadherin is the receptor for internalin, a surface protein required for entry of *L. monocytogenes* into epithelial cells. *Cell* 84 (6), 923–932.
- Misselwitz, B., Barrett, N., Kreibich, S., *et al.*, 2012. Near surface swimming of *Salmonella Typhimurium* explains target-site selection and cooperative invasion. *PLoS Pathogens* 8 (7), e1002810.
- Mota, L.J., Ramsden, A.E., Liu, M., Castle, J.D., Holden, D.W., 2009. SCAMP3 is a component of the *Salmonella*-induced tubular network and reveals an interaction between bacterial effectors and post-Golgi trafficking. *Cellular Microbiology* 11 (8), 1236–1253.
- Mounier, J., Boncompain, G., Senerovic, L., *et al.*, 2012. *Shigella* effector IpaB-induced cholesterol relocation disrupts the Golgi complex and recycling network to inhibit host cell secretion. *Cell Host and Microbe* 12 (3), 381–389.
- Mounier, J., Popoff, M.R., Enninga, J., *et al.*, 2009. The IpaC carboxyterminal effector domain mediates Src-dependent actin polymerization during *Shigella* invasion of epithelial cells. *PLoS Pathogens* 5 (1), e1000271.
- Nichols, C.D., Casanova, J.E., 2010. *Salmonella*-directed recruitment of new membrane to invasion foci via the host exocyst complex. *Current Biology* 20 (14), 1316–1320.
- Niebuhr, K., Giuriato, S., Pedron, T., *et al.*, 2002. Conversion of PtdIns(4,5)P(2) into PtdIns(5)P by the *Shigella flexneri* effector IpgD reorganizes host cell morphology. *EMBO Journal* 21 (19), 5069–5078.
- Ogawa, M., Yoshimori, T., Suzuki, T., *et al.*, 2005. Escape of intracellular *Shigella* from autophagy. *Science* 307 (5710), 727–731.
- Paciello, I., Silipo, A., Lembo-Fazio, L., *et al.*, 2013. Intracellular *Shigella* remodels its LPS to dampen the innate immune recognition and evade inflammasome activation. *Proceedings of the National Academy of Sciences of the United States of America* 110 (46), E4345–E4354.
- Patel, J.C., Galán, J.E., 2006. Differential activation and function of Rho GTPases during *Salmonella*-host cell interactions. *Journal of Cell Biology* 175 (3), 453–463.
- Pizarro-Cerdá, J., Cossart, P., 2009. *Listeria monocytogenes* membrane trafficking and lifestyle: The exception or the rule? *Annual Review of Cell and Developmental Biology* 25 (1), 649–670.
- Pizarro-Cerdá, J., Kühbacher, A., Cossart, P., 2015. Phosphoinositides and host-pathogen interactions. *Biochimica et Biophysica Acta — Molecular and Cell Biology of Lipids* 1851 (6), 911–918.
- Puhar, A., Tronchère, H., Payrastra, B., Tran Van Nhieu, G., Sansonetti, P.J., 2013. A *Shigella* effector dampens inflammation by regulating epithelial release of

- danger signal ATP through production of the lipid mediator PtdIns5P. *Immunity* 39 (6), 1121–1131.
- Rajabian, T., Gavicherla, B., Heisig, M., *et al.*, 2009. The bacterial virulence factor InlC perturbs apical cell junctions and promotes cell-to-cell spread of *Listeria*. *Nature Cell Biology* 11 (10), 1212–1218.
- Ramel, D., Lagarrigue, F., Pons, V., *et al.*, 2011. *Shigella flexneri* infection generates the lipid PI5P to alter endocytosis and prevent termination of EGFR signaling. *Science Signaling* 4 (191), ra61. doi: 10.1126/scisignal.2001619.
- Ray, K., Marteyn, B., Sansonetti, P.J., Tang, C.M., 2009. Life on the inside: The intracellular lifestyle of cytosolic bacteria. *Nature Reviews Microbiology* 7 (5), 333–340.
- Reed, S.C.O., Lamason, R.L., Risca, V.I., Abernathy, E., Welch, M.D., 2014. *Rickettsia* actin-based motility occurs in distinct phases mediated by different actin nucleators. *Current Biology* 24 (1), 98–103.
- Rejman Lipinski, A., Heymann, J., Meissner, C., *et al.*, 2009. Rab6 and Rab11 regulate *Chlamydia trachomatis* development and golgin-84-dependent Golgi fragmentation. *PLoS Pathogens* 5 (10), e1000615.
- Romero, S., Grompone, G., Carayol, N., *et al.*, 2011. ATP-mediated Erk1/2 activation stimulates bacterial capture by filopodia, which precedes *Shigella* invasion of epithelial cells. *Cell Host and Microbe* 9 (6), 508–519.
- Salcedo, S.P., Holden, D.W., 2003. SseG, a virulence protein that targets *Salmonella* to the Golgi network. *EMBO Journal* 22 (19), 5003–5014.
- Santos, A.J.M., Meinecke, M., Fessler, M.B., Holden, D.W., Boucrot, E., 2013. Preferential invasion of mitotic cells by *Salmonella* reveals that cell surface cholesterol is maximal during metaphase. *Journal of Cell Science* 126 (Pt 14), 2990–2996.
- Sarantis, H., Balkin, D.M., De Camilli, P., *et al.*, 2012. *Yersinia* entry into host cells requires rab5-dependent dephosphorylation of PI(4,5)P. *Cell Host and Microbe* 11 (2), 117–128.
- Schroeder, G.N., Hilbi, H., 2008. Molecular pathogenesis of *Shigella* spp.: Controlling host cell signaling, invasion, and death by type III secretion. *Clinical Microbiology Reviews* 21 (1), 134–156.
- Schwarz-Linek, U., Höök, M., Potts, J.R., 2006. Fibronectin-binding proteins of gram-positive cocci. *Microbes and Infection* 8 (8), 2291–2298.
- Shen, Y., Naujokas, M., Park, M., Ireton, K., 2000. InlB-dependent internalization of *Listeria* is mediated by the Met receptor tyrosine kinase. *Cell* 103 (3), 501–510.
- Snaveley, E.A., Kokes, M., Dunn, J.D., *et al.*, 2014. Reassessing the role of the secreted protease CPAF in *Chlamydia trachomatis* infection through genetic approaches. *Pathogens and Disease* 71 (3), 336–351.
- Sousa, S., Cabanes, D., Archambaud, C., *et al.*, 2005. ARHGAP10 is necessary for alpha-catenin recruitment at adherens junctions and for *Listeria* invasion. *Nature Cell Biology* 7 (10), 954–960.
- Sousa, S., Cabanes, D., Bougnères, L., *et al.*, 2007. Src, cortactin and Arp2/3 complex are required for E-cadherin-mediated internalization of *Listeria* into cells. *Cellular Microbiology* 9 (11), 2629–2643.
- Terebiznik, M.R., Vieira, O.V., Marcus, S.L., *et al.*, 2002. Elimination of host cell PtdIns(4,5)P(2) by bacterial SigD promotes membrane fission during invasion by *Salmonella*. *Nature Cell Biology* 4 (10), 766–773.
- Van Engelenburg, S.B., Palmer, A.E., 2008. Quantification of real-time *Salmonella* effector type III secretion kinetics reveals differential secretion rates for SopE2 and SptP. *Chemistry & Biology* 15 (6), 619–628.
- Veiga, E., Guttman, J.A., Bonazzi, M., *et al.*, 2007. Invasive and adherent bacterial pathogens co-Opt host clathrin for infection. *Cell Host and Microbe* 2 (5), 340–351.
- Velge, P., Wiedemann, A., Rosselin, M., *et al.*, 2012. Multiplicity of *Salmonella* entry mechanisms, a new paradigm for *Salmonella* pathogenesis. *Microbiologyopen* 1 (3), 243–258.
- Wong, K.-W., Isberg, R.R., 2005. Emerging views on integrin signaling via Rac1 during invasin-promoted bacterial uptake. *Current Opinion in Microbiology* 8 (1), 4–9.
- Yoshikawa, Y., Ogawa, M., Hain, T., *et al.*, 2009. *Listeria monocytogenes* ActA-mediated escape from autophagic recognition. *Nature Cell Biology* 11 (10), 1233–1240.
- Zumsteg, A.B., Goosmann, C., Brinkmann, V., Morona, R., Zychlinsky, A., 2014. IcsA is a *Shigella flexneri* adhesin regulated by the type III secretion system and required for pathogenesis. *Cell Host and Microbe* 15 (4), 435–445.

INTRACELLULAR INFECTIOLOGY: INFECTIOUS AGENTS

Contents

Bacterial Protein Toxins as Tools in Cell Biology and Physiology

The Gut Microbiome

Cell Biology of Virus Infection

Virus Factories and Mini-Organelles Generated for Virus Replication

HIV – The Cell Biology of Virus Infection and Replication

Prions

Bacterial Protein Toxins as Tools in Cell Biology and Physiology

M Rigoni and C Montecucco, CNR Institute of Neuroscience, Padua, Italy

© 2016 Elsevier Inc. All rights reserved.

Introduction

Pathogenic bacteria produce a large variety of protein toxins to increase their chance of multiplication within their hosts and/or to promote their diffusion. These toxins are the result of the coevolution of microorganisms and hosts, which select microbial products capable of subverting the host's physiology to the benefit of the producing bacteria. The 'preferred' targets of bacterial toxins are components of the host immune system, including immune cells and complement proteins, and the epithelial barrier, but several other host functions and cells are also targeted.

The understanding of the structure and function of bacterial toxins has allowed the development of vaccines, antibodies, and other therapeutics aimed at preventing or curing the diseases caused by the toxins. In addition, toxins have proven to be valuable tools for the dissection of several physiological processes of the cell. This review aims at describing the activity and use of bacterial toxins in cell biology.

Bacterial toxins can act: (1) on the cell surface (pore-forming toxins (PFT), sphingomyelinases, phospholipases, superantigens, etc.) or (2) in the cell cytosol. Bacterial toxins can reach their cytosolic target(s) either upon injection via complex molecular syringes made by the bacterium (secretion systems) or by exploiting internalization inside intracellular acidic compartments. This latter group of toxins will be presented here as they can be used by cell biologists to alter or knock out specific cellular processes in a very specific mode, whilst injected toxins have to be transfected, making the control of their intracellular action somehow problematic.

Exotoxins That Alter Signal Transduction

A large proportion of bacterial toxins appears to have been generated through evolution to avoid host immune

surveillance. A major bacterial strategy to escape immune response is to release protein toxins that alter crucial intracellular signaling pathways. The largest group of bacterial toxins is the PFT, which form pores in the plasma membrane and allow the permeation of a variety of ions, thus interfering with intracellular calcium signaling (Bischofberger *et al.*, 2012). PFTs bind to cell-surface receptors and oligomerize, forming a pore that spans the membrane. In this way, they subvert the osmotic equilibrium of the cell, causing calcium overload and water influx into leukocytes. They may boost cytosolic calcium so much as to cause calcium toxicity. Their application in the study of this major signaling pathway is less useful with respect to pharmacological tools such as calcium ionophores.

Other PFTs such as aerolysin, a toxin produced by the Gram-negative bacterium *Aeromonas hydrophila* and related species, form anion-selective transmembrane channels in host cells; as a consequence, the cell readily loses ions, water, and small molecules (i.e., sugars, amino acids, and nucleic acids) (Fivaz *et al.*, 2001). Loss of K^+ facilitates the formation and activation of the inflammasome, a host inflammatory signaling complex involved in the innate immune response to pathogens via the release of IL-1. Also *Staphylococcus aureus* α -toxin is a PFT that allows the permeation of monovalent ions across the membrane (Menestrina, 1986) and, like aerolysin, activates the inflammasome (Craven *et al.*, 2009). This fundamental intracellular component of the innate immune response is activated by a variety of bacterial toxins acting by various mechanisms, including the TcdB toxin of *Clostridium difficile*, responsible for most cases of nosocomial diarrhea. TcdB cytotoxin inactivates Rho GTPases by glucosylation, thereby affecting cytoskeletal regulation (see below) (Xu *et al.*, 2014).

A large group of protein toxins that deserve special mention are capable of binding to selected microdomains of the cell surface containing sphingomyelin and cholesterol or other

lipids. Their derivatives are very useful to study the structure and dynamics of the plasma membrane as well as the association of specific kinases and other signaling proteins to such microdomains (Skočaj *et al.*, 2013; Abe and Kobayashi, 2014).

Another major intracellular signaling pathway targeted by bacterial toxins is that triggered by the second messenger cyclic adenosine monophosphate (cAMP). The edema factor (EF), produced by *Bacillus anthracis*, combined with the protective antigen (PA) forms the edema toxin (PA + EF). These proteins can be easily produced in recombinant form as separate entities (Dal Molin *et al.*, 2006). PA binds to two different protein receptors which are widely expressed; therefore many different cell types can be affected. PA is activated by specific proteolysis on the cell surface, oligomerizes and binds EF. The PA + EF complex is internalized into endosomes and the acidic pH induces the formation of a mushroom-shaped PA channel that translocates EF into the cytosol (Collier and Young, 2003). Here, EF displays its calmodulin-dependent adenylate cyclase (CyA) activity which persists for many hours, and raises the cytosolic cAMP level with a gradient of concentrations higher around the nucleus and lower below the plasma membrane (Dal Molin *et al.*, 2006; Puhar *et al.*, 2008).

cAMP signaling in host cells is also affected by cholera toxin (CT) from *Vibrio cholerae*, a binary toxin (A–B): the A part is the catalytic subunit and the B part allows cell binding. The B part of CT is a pentamer of identical subunits (AB₅), each one harboring a binding site for the ganglioside GM1 which is present in many cells, including the apical portion of epithelial cells. Fluorescent derivatives of this pentamer are used to map ganglioside microdomains in cells (Conte *et al.*, 2009). CT is internalized inside endosomes, then reaches sequentially the trans-Golgi network (TGN), the Golgi cisternae, finally entering the endoplasmic reticulum (ER) lumen. As several other bacterial protein toxins do (see below), subunit A traverses the ER membrane through the reverse translocon, thereby reaching the cytosol. Subunit A then reaches the CyA complex located on the cytosolic surface of the basolateral membrane, and adenosine diphosphate (ADP)-ribosylates the α subunit of heterotrimeric G proteins of the CyA complex. As a consequence, this enzyme becomes permanently activated, resulting in an increased cAMP level, leading to stimulation of protein kinase A (PKA) and of chloride ion efflux through the cystic fibrosis transmembrane conductance regulator (CFTR) channel. The final result is a large efflux of ions and water across the intestinal epithelium, causing diarrhea (Simon *et al.*, 2014).

Another useful tool for the G heterotrimeric proteins is pertussis toxin (PTX), which is similarly formed by an oligomeric B subunit that binds to oligosaccharide receptors on various eukaryotic cells and by an A catalytic component (Locht *et al.*, 2011). Once internalized, the A protomer is delivered to the ER and from there to the cytosol, where it ADP-ribosylates the α subunit of a G_i protein, leading to a loss of CyA regulation and a rise in cAMP levels. PTX also acts on a number of G_o proteins which are implicated in controlling several intracellular processes (Carbonetti, 2010). *Bordetella pertussis* secretes another virulence factor that interferes with the cAMP signaling of the cell, the CyA. This calmodulin-activated enzyme is capable of crossing the plasma membrane of cells by itself and causes a rapid increase of cAMP below the

plasma membrane (Ladant and Ullmann, 1999; Vojtova *et al.*, 2006; Dal Molin *et al.*, 2006).

The MAPK pathway can also be targeted and inactivated by bacterial toxins. The anthrax lethal factor (LF) toxin combines with PA and enters the cytosol in the same way as EF. Once in the cytosol, LF cleaves the N-terminus of several members of the mitogen-activated protein kinase kinase (MAPKK) family (with the exception of MAPKK5) by virtue of its zinc-dependent metalloprotease activity (Vitale *et al.*, 2000). This results in the inhibition of the MAPK signal transduction pathway, which is key to cellular proliferation and signal transduction processes in the cell (Tonello and Montecucco, 2009). EF and LF synergize in their inhibitory action against immune cells, indicating that the cAMP and the MAPK pathway cross-talk at several points (Baldari *et al.*, 2006).

Photobacterium damsela subsp. *piscicida*, a Gram-negative pathogen responsible for an important fish septicemia, secretes the exotoxin AIP56 (apoptosis-inducing protein of 56 kDa), a major virulence factor. AIP56 is the first known A–B toxin that causes apoptosis by acting on the NF κ B p65 subunit, which is cleaved at a specific site by the metalloprotease subunit A. Other bacterial toxins act similarly, but they are injected via type III secretion systems, and are therefore less simple tools to use in cell biology. On the other hand, AIP56 is specific for fish immune cells (Silva *et al.*, 2013).

The complement system is also a target of bacterial toxins: *Pseudomonas aeruginosa* alkaline protease blocks C3b deposition and complement activation both via the classical and the lectin pathways, thus inhibiting the phagocytosis of the bacterium by human neutrophils (Laarman *et al.*, 2012). The *Neisseria meningitidis* protease NalP cleaves the C3 component of central importance for the activation of both complement pathways (Del Tordello *et al.*, 2014).

Exotoxins That Alter the Cytoskeleton

The cytoskeleton is an essential part of cell architecture and function: it controls cell structure and contacts, cell migration, phagocytosis, signaling, intracellular vesicle trafficking, and other processes. Therefore, it is not surprising that it is among the preferred targets of many virulence factors. Actin is often a direct substrate of ADP-ribosylating toxins. *Clostridium botulinum* C2 toxin, *Clostridium perfringens* iota toxin, *C. difficile* toxin (CDT), *Clostridium spiroforme* toxin (CST) and *Bacillus cereus* vegetative insecticidal protein (VIP), collectively called C2-like toxins, are A–B toxins that ADP-ribosylate actin, leading to its degradation (Simon *et al.*, 2014). Arg-177 of monomeric G-actin is the specific residue that is ADP-ribosylated. As this arginine is at the contact site between actin monomers, this bulky chemical modification prevents actin polymerization. It was observed that the ADP-ribosylation of actin and the consequent decrease of the cortical actin cytoskeleton cause changes in the microtubule system, with formation of microtubule protrusions which extend the cell surface. The superficial network of protrusions seems to enhance adhesion of bacteria and facilitates colonization (Aktories *et al.*, 2012). Intoxication leads to diminished cell–cell contacts and massive edema of intestinal endothelium with increase of fluid secretion.

However, the majority of bacterial toxins affecting the cell cytoskeleton do so indirectly, by interfering with its key

regulators: Rho, Rac, and Cdc42. By switching between a guanosine diphosphate (GDP)-bound inactive and a guanosine triphosphate (GTP)-bound active state, a transition controlled by activating (guanine nucleotide exchange factor (GEF) proteins) and deactivating (GTPase-activating proteins (GAP) proteins) components, these small GTPases control actin dynamics. The activity of these proteins can be modified by bacterial toxins via specific ADP-ribosylation, glycosylation, or deamidation, producing different cellular outcomes (Aktories, 2011).

Recently, the use of a variant of *C. difficile* toxin B (TcdB), a large (~270 kDa) glucosyltransferase that targets the Ras superfamily of small GTPases causing its irreversible glycosylation, has helped to elucidate the role of the Rho GTPases in the control of centrosome recruitment and the activation of the mitotic kinases. Treatment of synchronized hela cells with TcdB in the G2 phase prevented the centrosomal activation of PAK, Aurora A, and Cdk1, and delayed mitotic entry (May *et al.*, 2014).

Insect pathogens can produce toxin complexes that target the cytoskeleton of insect cells. Such tripartite complexes, known as Tc or ABC toxins, are produced, among others, by entomopathogenic *Photobacterium luminescens*, which lives in symbiosis with nematodes. These nematodes invade insect larvae and release the bacteria in the intestine, causing the death of insects by the action of toxin complexes. Tc toxin consists of TcA, TcB, and TcC components: TcA is the binding and membrane translocation component, and TcB is a functional linker between TcC and TcA, and is also involved in the translocation of the toxin. TccC3 ADP-ribosylates actin at threonine 148, resulting in actin polymerization, whereas TccC5 modifies Rho proteins at glutamine 61/63, inducing their activation. The concerted action of both toxins inhibits phagocytosis of target insect cells and induces extensive intracellular polymerization and clustering of actin. Since Tc toxin targets insect cells, it could be exploited in studies that employ *Drosophila melanogaster* as model system. Several human pathogenic bacteria produce related toxins (Lang *et al.*, 2010; Simon *et al.*, 2014).

In the nervous system, the Rho GTPases play a key role in several processes, controlling neurotransmitter release, the morphogenesis of dendritic spines, and synaptic plasticity; mutations in proteins involved in Rho GTPase signaling may be causative of some forms of intellectual disabilities. Rho GTPases can be blocked in their activated, GTP-bound state by a bacterial protein toxin from *Escherichia coli* named cytotoxic necrotizing factor 1 (CNF1). CNF1 catalyzes the deamidation of a single glutamine residue of the Rho molecules, thus inhibiting GTP hydrolysis and leading to their persistent activation. CNF1 is able to increase brain plasticity in adulthood (Ceri *et al.*, 2011). The ability of CNF1 to act on synaptic plasticity by modulating Rho GTPases activity may be exploited to promote brain repair (Travaglione *et al.*, 2014), but its action is not limited to neurons, and CNF1 is very useful for the study of the role of Rho in different cellular processes and in cancer growth (Vannini *et al.*, 2014).

In addition to C2 toxin and BoNT/C and BoNT/D neurotoxins, type C and D strains of *C. botulinum* produce a Rho-modifying ADP-ribosyltransferase named C3 toxin. C3 transferases are not restricted to Clostridia, but are released by many different bacterial species. C3 toxins modify an asparagine

residue of RhoA, B, and C, at variance from many bacterial ADP-ribosyltransferases that usually target arginines or cysteines, and have been widely used to dissect the cellular functions of Rho GTPases (Vogelsgesang *et al.*, 2007). Their cellular uptake mechanism has not been clarified yet, since they lack a transporter B subunit. Several mechanisms of entry into host cells have been proposed so far: (1) unspecific internalization, (2) internalization mediated by pore-forming cytotoxins produced by the same bacteria, and (3) internalization of the bacterium that, once inside the host cell, can secrete the toxin directly into the host cytosol. Due to their poor cell accessibility C3 toxins must be transfected, microinjected, or fused with the binding and translocation subunit of other toxins to be delivered into cells (Vogelsgesang *et al.*, 2007).

Exotoxins That Interfere with Intracellular Trafficking

The various organelles of eukaryotic cells are interconnected by a continuous and active trafficking of vesicles, which allows for the exchange, uptake, and delivery of molecules. Vesicle trafficking is governed by several proteins, among which the three SNARE (SNAP (Soluble NSF Attachment Protein) Receptor) proteins VAMP, SNAP-25, and syntaxin play a major role, together with members of the Rab family of small GTPases. There are many different isoforms of SNAREs and Rabs, which are involved in different events of vesicle traffic (Lang and Jahn, 2008). Several human bacterial pathogens (*Legionella pneumophila*, *Listeria monocytogenes*, *Salmonella enterica*, etc.) create specialized niches in macrophages and epithelia by altering Rab GTPase functions. This strategy allows them to escape degradation and obtain nutrients essential for growth and survival (Stein *et al.*, 2012).

Tetanus neurotoxin and botulinum neurotoxins (whose number is rapidly growing) cleave some isoforms of SNAREs at single peptide bonds (Rossetto *et al.*, 2014). However, they only enter neurons that express both their polysialoganglioside and protein receptors and their capability of cleaving the specific SNARE isoform under study has to be controlled before use. Cleavage by a botulinum neurotoxin of a particular SNARE isoform of a given animal species depends on the sequence of the cleavage site as well as on the sequence of several exosites which are located at a distance from the cleavage site (Pantano and Montecucco, 2014). Thus, botulinum neurotoxins are very useful in studying the role of neurotransmitters or other biological molecules released by fusion of intracellular vesicles with the plasma membrane and in intercellular signaling, because they provide a cheap knockout of the process mediated by a given SNARE protein.

Tetanus neurotoxin, together with the pentameric binding protomer of cholera toxin, are the two best characterized markers of retroaxonal transport inside motoneurons from the periphery to the spinal cord (Bercsenyi *et al.*, 2013). In addition, one can use bacterial toxins to trace vesicular transport inside cells. For example, Shiga toxin (STx) from *Shigella dysenteriae*, cholera toxin, and *P. aeruginosa* exotoxin A (ExoA) bind to their cell receptors via the B subunits, and are then endocytosed inside early endosomes which undergo progressive lumen acidification. Instead of being directed to late endosomes and then to lysosomes for degradation, these toxins transit to the TGN and from there to the ER. Once in the ER, they exploit a cellular

mechanism devoted to the degradation of newly synthesized misfolded proteins, called endoplasmic reticulum-associated protein degradation (ERAD). Usually misfolded proteins are driven through a reverse translocon pore to be discarded into the cytosol where they become substrates for ubiquitination and are degraded inside the proteasome (Johannes and Wunder, 2011). These bacterial toxins are ubiquitinated poorly or not at all in the cytosol; they therefore escape such degradation route and can exert their enzymatic activities within the host cell.

Exotoxins That Block Protein Synthesis or Protein Folding

Many bacterial toxins block protein synthesis and/or folding, with deleterious consequences particularly in organs that strictly rely on protein synthesis for their function. The most studied member of this family is diphtheria toxin (DT), which represents the paradigm of the bacterial toxins made of two parts (A for active and B for binding), held together by a single interchain disulfide bond. DT is endocytosed by most cells and when exposed to the acid pH of the endosomal lumen it penetrates the membrane, causing translocation of the A protomer into the cytosol. DT-A is an ADP-ribosyl-transferase which specifically transfers the ADP-ribose from NAD⁺ to the elongation factor 2 (EF-2), thereby causing a complete inhibition of protein synthesis and cell death (Collier, 2001). It was actually demonstrated that a single molecule of the enzymatically active fragment A of DT is by itself able to kill an eukaryotic cell (Yamaizumi *et al.*, 1978). Other bacterial toxins that block protein synthesis are ExoA, which acts in the same way as DT, as well as STx and Shiga-like toxins from some enterohemorrhagic strains of *E. coli*. STx is an AB₅ exotoxin which binds to glycolipids of the cell surface via the B5 protomer, and is endocytosed; the A subunit is then delivered to the cytosol, where it displays N-glycosidase activity specific for a single adenine of ribosomal RNA, thus blocking protein synthesis. These toxins, as well as the plant toxin ricin, could be used to study the kinetics of protein synthesis.

Cell life is dependent on the correct folding of proteins to be exported, an event that takes place in the ER. A central player in this process is the chaperone Binding Immunoglobulin Protein (BiP). BiP is cleaved by a bacterial toxin produced by *E. coli* associated to an uremic syndrome, leading to inactivation of protein folding (Paton *et al.*, 2006; Montecucco and Molinari, 2006).

Exotoxins That Alter Deoxyribonucleic Acid (DNA)

The cytolethal distending toxins (CDTs) are bacterial genotoxins that cause DNA damage in the target cells. They are produced by a variety of Gram-negative pathogenic bacteria, among them *E. coli*, *Shigella dysenteriae*, *Campylobacter sp.*, *Helicobacter sp.*, and *S. enterica*. The mechanism of cytotoxicity of CDTs is unique: they enter into eukaryotic cells, reach the nucleus, and break double-stranded DNA, resulting in the arrest of the cell cycle and, eventually, in the activation of the apoptotic program of cell death. CDT is the product of an operon encoding three proteins: CdtA, CdtB, and CdtC. The catalytic subunit CdtB has deoxyribonuclease I (DNase I)-like activity, whereas CdtA and CdtC are binding proteins for delivering CdtB into target cells. All three subunits are essential to confer full activity to the

holotoxin. Binding of CDT depends on the presence of intact lipid rafts, and the toxin is internalized via dynamin-dependent endocytosis into early and late endosomes. The CdtB subunit further transits to the Golgi complex, and is then retrogradely transported to the ER. Translocation from the ER does not require the ERAD pathway, and protein unfolding. How the nuclear translocation of the CdtB subunit takes place is still an open question. In response to DNA damage, cells activate a series of complex mechanisms that aim at repairing the genome and limiting the probability of lethal or permanent genetic damage (Guerra *et al.*, 2011). Recently, a novel toxin consisting of a B binding protomer and of two catalytic A subunits has been isolated from *S. enterica*. One A subunit is similar to PTX, whilst the other is a DNase (Song *et al.*, 2013).

Final Considerations

The field of bacterial toxins acting upon host cells is continuously expanding, thanks to the use of genomic, transcriptomic, and proteomic methods, which are revealing the large variety of tools that bacteria have devised in their continuous struggle to increase their presence in the world. In addition, the same evolutionary processes have taken place in the animal and plant kingdoms, and other toxins not considered here may be found useful in cell biology studies.

See also: Interorganellar Communication: Interplay and Processes: Retrograde Transport

References

- Abe, M., Kobayashi, T., 2014. Imaging local sphingomyelin-rich domains in the plasma membrane using specific probes and advanced microscopy. *Biochimica et Biophysica Acta* 1841, 720–726.
- Aktories, K., 2011. Bacterial protein toxins that modify host regulatory GTPases. *Nature Reviews Microbiology* 9, 487–498.
- Aktories, K., Schwan, C., Papatheodorou, P., Lang, A.E., 2012. Bidirectional attack on the actin cytoskeleton. Bacterial protein toxins causing polymerization or depolymerization of actin. *Toxicon* 60, 572–581.
- Baldari, C.T., Tonello, F., Paccani, S.R., Montecucco, C., 2006. Anthrax toxins: A paradigm of bacterial immune suppression. *Trends in Immunology* 27, 434–440.
- Bercsenyi, K., Giribaldi, F., Schiavo, G., 2013. The elusive compass of clostridial neurotoxins: Deciding when and where to go? *Current Topics in Microbiology and Immunology* 364, 91–113.
- Bischofberger, M., Iacovache, I., Van der Goot, F.G., 2012. Pathogenic pore-forming proteins: Function and host response. *Cell Host & Microbe* 12, 266–275.
- Carbonetti, N.H., 2010. Pertussis toxin and adenylate cyclase toxin: Key virulence factors of *Bordetella pertussis* and cell biology tools. *Future Microbiology* 5, 455–469.
- Cerri, C., Fabbri, A., Vannini, E., *et al.*, 2011. Activation of Rho GTPases triggers structural remodeling and functional plasticity in the adult rat visual cortex. *Journal of Neuroscience* 31, 15163–15172.
- Collier, R.J., 2001. Understanding the mode of action of diphtheria toxin: A perspective on progress during the 20th century. *Toxicon* 39, 1793–1803.
- Collier, R.J., Young, J.A., 2003. Anthrax toxin. *Annual Review of Cell and Developmental Biology* 19, 45–70.
- Conte, W.L., Kamishina, H., Reep, R.L., 2009. Multiple neuroanatomical tract-tracing using fluorescent Alexa Fluor conjugates of cholera toxin subunit B in rats. *Nature Protocols* 4, 1158–1166.
- Craven, R.R., Gao, X., Allen, I.C., *et al.*, 2009. *Staphylococcus aureus* alpha-hemolysin activates the NLRP3-inflammasome in human and mouse monocytic cells. *PLoS One* 4, e7446.

- Dal Molin, F., Tonello, F., Ladant, D., *et al.*, 2006. Cell entry and cAMP imaging of anthrax edema toxin. *EMBO Journal* 25, 5405–5413.
- Del Tordello, E., Vacca, I., Ram, S., Rappuoli, R., Serruto, D., 2014. *Neisseria meningitidis* NalP cleaves human complement C3, facilitating degradation of C3b and survival in human serum. *Proceedings of the National Academy of Sciences of the United States of America* 111, 427–432.
- Fivaz, M., Abrami, L., Tsitrin, Y., van der Goot, F.G., 2001. Not as simple as just punching a hole. *Toxicon* 39, 1637–1645.
- Guerra, L., Cortes-Bratti, X., Guidi, R., Frisan, T., 2011. The biology of the cytolethal distending toxins. *Toxins (Basel)* 3, 172–190.
- Johannes, L., Wunder, C., 2011. Retrograde transport: Two (or more) roads diverged in an endosomal tree? *Traffic* 12, 956–962.
- Laarman, A.J., Bardoel, B.W., Ruyken, M., *et al.*, 2012. *Pseudomonas aeruginosa* alkaline protease blocks complement activation via the classical and lectin pathways. *Journal of Immunology* 188, 386–393.
- Ladant, D., Ullmann, A., 1999. *Bordetella pertussis* adenylate cyclase: A toxin with multiple talents. *Trends in Microbiology* 7, 172–176.
- Lang, A.E., Schmidt, G., Schlosser, A., *et al.*, 2010. Photorhabdus luminescens toxins ADP-ribosylate actin and RhoA to force actin clustering. *Science* 327, 1139–1142.
- Lang, T., Jahn, R., 2008. Core proteins of the secretory machinery. *Handbook of Experimental Pharmacology* 184, 107–127.
- Locht, C., Coutte, L., Mielcarek, N., 2011. The ins and outs of pertussis toxin. *FEBS Journal* 278, 4668–4682.
- May, M., Schelle, I., Brakebusch, C., Rottner, K., Genth, H., 2014. Rac1-dependent recruitment of PAK2 to G 2 phase centrosomes and their roles in the regulation of mitotic entry. *Cell Cycle* 13, 2210–2220.
- Menestrina, G., 1986. Ionic channels formed by *Staphylococcus aureus* alpha-toxin: Voltage-dependent inhibition by divalent and trivalent cations. *Journal of Membrane Biology* 90, 177–190.
- Montecucco, C., Molinari, M., 2006. Microbiology: Death of a chaperone. *Nature* 443, 511–512.
- Pantano, S., Montecucco, C., 2014. The blockade of the neurotransmitter release apparatus by botulinum neurotoxins. *Cellular and Molecular Life Science* 71, 793–811.
- Paton, A.W., Beddoe, T., Thorpe, C.M., *et al.*, 2006. AB5 subtilase cytotoxin inactivates the endoplasmic reticulum chaperone BiP. *Nature* 443, 548–552.
- Puhar, A., Dal Molin, F., Horvath, S., Ladant, D., Montecucco, C., 2008. Anthrax edema toxin modulates PKA- and CREB-dependent signaling in two phases. *PLoS One* 3, e3564.
- Rossetto, O., Pirazzini, M., Montecucco, C., 2014. Botulinum neurotoxins: Genetic, structural and mechanistic insights. *Nature Reviews Microbiology* 12, 535–549.
- Silva, D.S., Pereira, L.M., Moreira, A.R., *et al.*, 2013. The apoptogenic toxin Alp56 is a metalloprotease A-B toxin that cleaves NF- κ B P65. *PLoS Pathogens* 9, e1003128.
- Simon, N.C., Aktories, K., Barbieri, J.T., 2014. Novel bacterial ADP-ribosylating toxins: Structure and function. *Nature Reviews Microbiology* 12, 599–611.
- Skočaj, M., Bakrač, B., Križaj, I., *et al.*, 2013. The sensing of membrane microdomains based on pore-forming toxins. *Current Medicinal Chemistry* 20, 491–501.
- Song, J., Gao, X., Galan, J.E., 2013. Structure and function of the *Salmonella typhi* chimaeric A(2)B(5) typhoid toxin. *Nature* 499, 350–354.
- Stein, M.P., Muller, M.P., Wandinger-Ness, A., 2012. Bacterial pathogens commandeer Rab GTPases to establish intracellular niches. *Traffic* 13, 1565–1588.
- Tonello, F., Montecucco, C., 2009. The anthrax lethal factor and its MAPK kinase-specific metalloprotease activity. *Molecular Aspects of Medicine* 30, 431–438.
- Travaglione, S., Loizzo, S., Ballan, G., Fiorentini, C., Fabbri, A., 2014. The *E. coli* CNF1 as a pioneering therapy for the central nervous system diseases. *Toxins (Basel)* 6, 270–282.
- Vannini, E., Panighini, A., Cerri, C., *et al.*, 2014. The bacterial protein toxin, cytotoxic necrotizing factor 1 (CNF1) provides long-term survival in a murine glioma model. *BMC Cancer* 14, 449.
- Vitale, G., Bernardi, L., Napolitani, G., Mock, M., Montecucco, C., 2000. Susceptibility of mitogen-activated protein kinase family members to proteolysis by anthrax lethal factor. *Biochemical Journal* 352, 739–745.
- Vogelsgesang, M., Pautsch, A., Aktories, K., 2007. C3 exoenzymes, novel insights into structure and action of Rho-ADP-ribosylating toxins. *Naunyn-schmiedeberg's Archives of Pharmacology* 374, 347–360.
- Vojtova, J., Kamanova, J., Sebo, P., 2006. Bordetella adenylate cyclase toxin: A swift saboteur of host defense. *Current Opinion in Microbiology* 29, 69–75.
- Xu, H., Yang, J., Gao, W., *et al.*, 2014. Innate immune sensing of bacterial modifications of Rho GTPases by the Pyrin inflammasome. *Nature* 513, 237–241.
- Yamaizumi, M., Mekada, E., Uchida, T., Okada, Y., 1978. One molecule of diphtheria toxin fragment A introduced into a cell can kill the cell. *Cell* 15, 245–250.

The Gut Microbiome

H Neuman and O Koren, Bar-Ilan University, Safed, Israel

© 2016 Elsevier Inc. All rights reserved.

The Gut Microbiome

The emerging field of microbiome research refers to the collection of microorganisms found within human, animal, or plant hosts, and their genetic information. Amazingly, the number of microbial cells found within the human body is approximately 10 times greater than the sum of all human somatic and germ cells. Accordingly, the microbial genetic information within our bodies is 150 times greater than our own. Our microbial partners can be found throughout the body, including the gastrointestinal tract, skin, nose, mouth, and urogenital system, with different typical compositions in different organs (Human Microbiome Project, C., 2012). The human microbiome is composed of bacteria, archaea, viruses, and fungi. For the purpose of this article, we will refer mainly to the bacterial component of the microbiome, as this has been the most extensively studied. The largest bacterial community in our body resides in the gastrointestinal tract. In the colon, for example, bacterial densities can reach up to 10^{11} cells/g luminal contents (Simon and Gorbach, 1984). Together, gut microbes form a community referred to as the microbiota. The human gut microbiota is made up of more than 1000 different microbial species (Human Microbiome Project, C., 2012). Most of these microbes are anaerobic and cannot be grown using routine laboratory procedures. Our gut microbiota has been shown to have broad effects on our health and well-being, impacting our immune system and metabolism, detoxifying various ingested components, and performing other functions. Thus, the majority of gut microbes are generally viewed as beneficial symbionts, having reciprocal interactions with their hosts. Changes in gut microbiota composition are correlated with a wide range of disease states.

The human microbiome is highly variable both within a single subject over time and different environmental conditions, and between different individuals. While the microbiome is greatly affected by environmental factors, the genetic contribution of the host seems to be polygenic and mild, as there can be large differences in microbial populations between identical twins (Turnbaugh *et al.*, 2009b).

Microbiome Research Methodology (Approaches and Tools)

Microbiome research aims to decipher which are the components that constitute the overall microbial populations, how they interact and affect one another, how they are affected by the host environment, and how they affect the host at various levels.

The ability to thoroughly characterize microbiota components came only with the development of next-generation sequencing (NGS) techniques, including Roche 454, ion torrent and Illumina, as many of the microbes in the human body are not readily cultured. Even with culture under

anaerobic conditions to enlarge the number of potentially culturable microbes, only about 50% of the microbial populations are represented (Goodman *et al.*, 2011). NGS also enables high-throughput experiments, sequencing millions of reads simultaneously and analyzing them with bioinformatic methods.

16S rRNA Gene Surveys

The most commonly used method to date for characterizing bacterial communities is sequencing of the 16S rRNA bacterial gene. While the 16S rRNA gene includes highly conserved regions, it also includes highly variable regions, enabling the distinction between thousands of bacterial sequences, representing different microbial species. These 16S surveys can be used to answer the question: 'What are the microbial components in the sample?' The sequences from a sample can be assigned taxonomy using databases of known 16S gene sequences. The number of sequences in a sample that are classified as a specific microbe can give an estimate of the relative abundance of that microbe within the sample. Microbial diversity can also be calculated both within a single sample (alpha-diversity), and between different samples (beta-diversity).

Metagenomics (Whole-Genome Shotgun Sequencing)

Another way to characterize the microbiome is with whole-genome sequencing (shotgun sequencing). In this technique all genetic material in the sample is sequenced together, and only later segregated as host versus microbial. This approach offers information regarding the bacterial genes present, thereby hinting at the roles and abilities of the microbial populations. Such functional metagenomic data answers the question: 'What could the microbial components potentially do?'

Metatranscriptomics

Additionally, reverse-transcribed RNA sequencing from a sample can reveal metatranscriptomic data, highlighting the microbial genes actually expressed, thereby answering the question: 'What are these microbes really doing?'

Metabolomics

Metabolic profiling, usually by chromatography and spectrometry, reveals the composition of small-molecule metabolites from both host and microbial metabolism in a sample. By combining these data with the genetic sequences of a sample, it is possible to find correlations between levels of specific metabolites and microbial populations (Ursell *et al.*, 2014). This approach can answer the question: 'How do these microbes affect host metabolites?'

Germ-Free Animals

Germ-free (GF) organisms, grown under sterile conditions, are a useful tool for understanding the roles of the microbiome. Observations of GF animals have revealed alterations in immune system components, hormonal levels, metabolism, and even behavior. GF models allow the introduction of specific bacteria, creating a gnotobiotic model, and studying their effects on the host. Microbiota transplants, in which fecal bacteria from a donor are transferred into a GF mouse, can also assess the effects of the microbiome. Such experiments have shown that exposure to different types of bacterial compositions can lead to a variety of health outcomes, proving a causal relationship between microbiota and health states. Another method used to alter microbiome composition in rodents is cohousing of animals with different microbial compositions.

Constituents of the Microbiome

Bacteria

The bacterial components of the microbiome are the most dominant and extensively studied to date. The two predominant bacterial phyla represented in the gut microbiome are Firmicutes and Bacteroidetes. Additional phyla include Actinobacteria, Proteobacteria, Verrucomicrobia, Fusobacteria, and Cyanobacteria (Eckburg *et al.*, 2005).

Archaea

The archaeal gut community is composed mainly of methanogens. While methanogens are not found in all fecal samples, the dominant methanogen in the human gut is *Methanobrevibacter smithii*. Methanogen abundance is higher following a high-carbohydrate diet, and in obese individuals. In fact, methanogens may be a risk factor for obesity, since in the process of producing methane they oxidize hydrogen and complete fermentation of carbohydrate substrates, leading to higher production of short-chain fatty acids (Conway de Macario and Macario, 2009; Hoffmann *et al.*, 2013). Methanogens may also play an immunological role, capable of inducing an inflammatory cytokine response (Bang *et al.*, 2014). Interestingly, their abundance is higher in inflammatory bowel diseases (IBD) patients, who also display an antigen-specific IgG response to the methanogens (Blais Lecours *et al.*, 2014).

Viruses

The human virome is unique to each individual and includes many viruses, including previously uncharacterized bacteriophages, which may play a defensive role against pathogens (Minot *et al.*, 2013). As part of the Human Microbiome Project (HMP), the dsDNA viruses of 102 individuals were analyzed, and on average included 5.5 viral genera per individual. These viruses included herpesviruses, papillomaviruses, polyomaviruses, adenoviruses, anelloviruses, parvoviruses, and circoviruses. Most of these viruses were stable over time (Wylie *et al.*, 2014). Nonetheless, due to their rapid evolution, horizontal gene transfers, and intimate interactions with host

nucleic acids, we are only beginning to decipher the human viral communities on a large scale (Abeles and Pride, 2014).

Eukaryotes

Our bodies are colonized with populations of fungi, also termed the mycobiome, as well as several protists. There are at least 66 genera of fungi found in the gut microbiome (Hoffmann *et al.*, 2013). Similarly to the other components of the microbiome, the fungal components are likely to impact host health and disease (Gouba *et al.*, 2014). Fungi have important roles interacting with innate immune cells, via receptors such as Dectin-1 (Underhill and Iliev, 2014). It has been found that disruption of the Dectin-1 receptor function can lead to colitis, perhaps due to loss of interactions with host fungi (Iliev *et al.*, 2012). Additionally, there have been studies on other eukaryotic members of the gut microbiome such as the single-celled protozoa, *Blastocystis*, which was present in over 50% of adult samples studied (Scanlan *et al.*, 2014).

The Roles of the Microbiome

The most obvious role of the gut microbiome is in food digestion, performing numerous biochemical reactions including breaking down starch, producing amines and phenols from proteins, and metabolizing xenobiotics into multiple metabolites. The commensal microbiota can also produce vitamins.

In addition, the gut microbiota affects the host immune system. Beyond influencing the gut barrier function and repressing pathogenic microbes, the gut microbiota plays a role in modulating the host immune response, both locally and systemically (Kamada *et al.*, 2013). In the absence of commensal bacteria, GF mice have impaired development of the innate and adaptive immune system (Hill and Artis, 2010; Honda and Littman, 2012; Hooper *et al.*, 2012; Littman and Pamer, 2011), reduced numbers of IgA-producing plasma cells (Crabbe *et al.*, 1969), impaired development of gut-associated lymphoid tissues (Kelly and Mulder, 2012) and a decreased percentage of CD4⁺ T cells (Ostman *et al.*, 2006). Additionally, T helper 17 (TH17) cells, which produce pro-inflammatory cytokines, are regulated by gut bacteria and are promoted specifically by segmented filamentous bacteria (Tanabe, 2013). Finally, autoimmune diseases have also been correlated with alterations of the microbiome (dysbiosis).

Another effect that the microbiota has on their hosts includes altering hormone levels. There have been numerous reports of microbiota both producing and regulating host hormones such as serotonin, epinephrine, norepinephrine, dopamine, etc. These, in turn, affect host behavior, metabolism, and immunity.

The effects of the microbiome on host brain functions and behavior are particularly interesting. Microbiota in mice has been shown to affect depression, anxiety, and exploratory behavior. GF mice display a variety of behavioral alterations compared to specific pathogen-free (SPF) mice, including changes in cognitive function, memory, stress response,

anxiety, and social behavior (Diaz Heijtz *et al.*, 2011; Gareau *et al.*, 2011; Neufeld *et al.*, 2011). These behavioral alterations are concurrent with alterations found in serotonin levels and lower levels of brain-derived neurotrophic factor, in GF vs. SPF mice (Bercik *et al.*, 2011; Grenham *et al.*, 2011). In a different study, *Lactobacillus* was found to regulate emotion and behavior, as well as central GABA receptor expression, via the vagus nerve (Bravo *et al.*, 2011). In *Drosophila melanogaster*, it was shown that microbial symbionts drive sexual preference, and that these change upon use of antibiotics or addition of specific bacteria (Sharon *et al.*, 2010). Although much remains to be learned, there are clues that bacteria may affect behavior in humans, as well. Minocycline, an antibiotic with a broad-spectrum activity, has been recently suggested as a candidate treatment for depression and schizophrenia (Jhamani *et al.*, 2013; Soczynska *et al.*, 2012). Furthermore, digestive diseases such as IBD have been reported to lead to depression and anxiety (Walker *et al.*, 2008), although precise mechanisms have not yet been described. Additionally, in a human study testing effects of probiotics on brain function, results showed altered brain activity in regions associated with central processing of emotion and sensation (Tillisch *et al.*, 2013). Thus, while further research is needed to understand precise relationships, mechanisms and interactions, there are significant clues for broad microbial effects on their hosts.

Changes in the Gut Microbiome Throughout Life

Shifts in the composition of the microbiome occur naturally at different stages in life, such as infancy, childhood, puberty, old age, pregnancy, and others (Spor *et al.*, 2011). Below, we describe the typical microbiome composition for the different age groups (also described in Figure 1).

Pregnancy

During pregnancy, there is an increase of Proteobacteria and Actinobacteria, and less richness (number of different types of microbial species) is observed. These changes correlate with increased weight gain, hormonal and immunological changes, and insulin desensitization. Interestingly, when GF mice receive fecal transplants from third trimester pregnant women, they develop induced adiposity and insulin desensitization, compared with mice receiving fecal transplants from women in their first trimester. This suggests that the bacterial populations may play a role in these metabolic processes. These aberrant microbiotas are then passed to the neonates at birth (Koren *et al.*, 2012).

Birth and Infancy

While it has long been assumed that we are born sterile, new evidence suggests that the placenta harbors bacteria (Aagaard

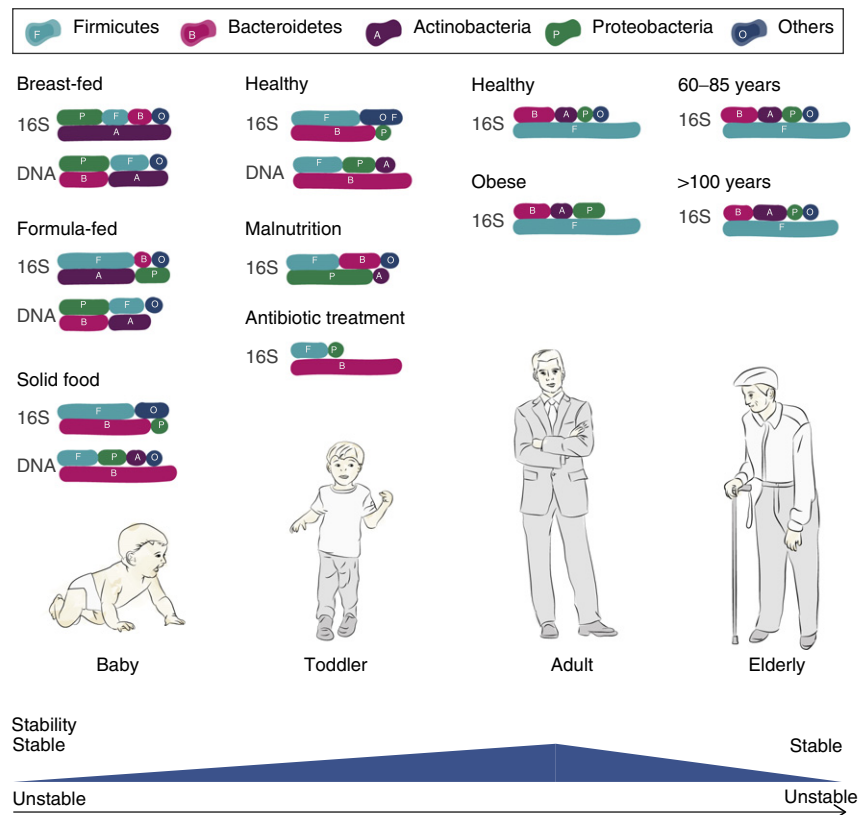


Figure 1 Changes in gut microbiota composition throughout life. Distinct age-related alterations in microbial composition as seen by the colored bars representing typical distribution of the major bacterial phyla following 16S analysis (upper part of figure, adapted from Ottman, N., *et al.*, 2012). The function of our microbiota: Who is out there and what do they do? *Frontiers in Cellular and Infection Microbiology* 2, 104), and by microbial stability (relative compositional changes throughout time) (bottom part of figure, adapted from Spor, A., Koren, O., Ley, R., 2011). Unraveling the effects of the environment and host genotype on the gut microbiome. *Nature Reviews. Microbiology* 9 (4), 279–290.

et al., 2014). A newborn's microbiome is greatly affected by the maternal microbiome (Spor *et al.*, 2011) and by the mode of delivery, so that babies born by vaginal birth obtain an initial microbial inoculum that resembles their mother's vaginal microbiome, as opposed to the microbiome of babies born by cesarean section, which is more similar to their mother's skin microbiota (Spor *et al.*, 2011). The microbiome composition changes dramatically during the first 2–3 years of life (Spor *et al.*, 2011), and then stabilizes. The infant's microbiota is characterized by less diversity compared to that of an adult, harboring a high abundance of bacteria that function in lactate utilization. Transition to solid foods has been shown to increase the abundance of Bacteroidetes as well as bacteria that utilize carbohydrates, and are involved in vitamin biosynthesis and xenobiotic degradation (Koenig *et al.*, 2011). It has been further shown that the gestational age greatly affects the development of the gut microbiota, so that premature infants take longer to assemble bacterial communities (La Rosa *et al.*, 2014).

Adulthood

The largest microbiome studies to date that describe the human adult microbiome have been performed by consortiums named the Metagenomics for the Human Intestinal Tract (MetaHIT), and the HMP. MetaHIT was the first to perform metagenomics on fecal samples from 124 human subjects, and found that the gut microbiota includes ~1000 species. Interestingly, they found great diversity of microbial species between subjects, so that only 18 core species were found in all subjects. However, when analyzing the metagenomic data, they found much greater similarity, suggesting that while the precise microbial species may differ between subjects, their genes and functions may be the same (Qin *et al.*, 2010). The HMP was initiated by the US National Institutes of Health in an attempt to identify and characterize the microorganisms found in association with health and disease in humans. So far, the first phase of the project characterized the composition and diversity of microbial communities inhabiting a variety of body sites, and evaluated the genetic metabolic potential of these communities. All together, almost 5000 samples were collected and sequenced from 242 healthy volunteers, representing 15 (in males) to 18 (in females) body sites, including nasal passages, oral cavities, skin, gastrointestinal tract, and the urogenital tract. Their reference database is the largest one available to date. It appears that the microbial populations at different body sites are unique, and that there is higher diversity between body sites rather than between different subjects. Even throughout the gastrointestinal tract there are significant population variations according to location. This may not be surprising as the microenvironments are different at various locations, including variations in pH levels. The current phase of studies by the HMP focuses on microbiome-associated diseases (Human Microbiome Project, C., 2012).

Aging

In the elderly, the microbiota once again becomes less stable (Spor *et al.*, 2011), and is dominated by members of Bacteroidetes. In a study comparing the gut microbiota

composition of 161 elders (over age 65) and 9 younger controls, a greater proportion of *Bacteroides* were found in the elders as well as distinct abundance patterns of *Clostridium* groups (Claesson *et al.*, 2011). These alterations may be due to changes in diet, a more sedentary lifestyle, frailty, background diseases, and/or states of inflammation (Candela *et al.*, 2014). On a functional level, there are over 100 bacterial genes correlated with host aging. The microbiome in the elderly is characterized by loss of SCFA-producing genes and an overall decrease in the saccharolytic potential, while proteolytic functions are more abundant than in younger adults. Additionally, in samples from the elderly, there were more opportunistic pro-inflammatory bacteria (Rampelli *et al.*, 2013). In a study that compared the microbiota of elders living in the community versus microbiota of elders in long-stay care, the latter was significantly less diverse and correlated with frailty (Claesson *et al.*, 2012).

Factors Influencing the Microbiome

Changes in environmental factors such as diet, geographic location, and use of antibiotics, which may occur throughout life, are associated with changes in the gut microbial communities.

Diet and Geographic Location

The factor that most strongly influences the composition of the microbiome appears to be diet. The effects of a dietary change on the microbiota composition can be extremely rapid. In a study of short-term dietary change (5 days long) in which subjects received an animal-based diet, there were observed changes in microbial communities compared to baseline after as little as 2 days, including changes in microbial gene expression (David *et al.*, 2014b). Certain dietary compositions have been associated with higher abundance of a specific genus, such as significant *Bacteroides* in a protein-rich animal fat diet and significant *Prevotella* in a carbohydrate-rich diet (Wu *et al.*, 2011).

In a unique study in which the gut microbiota of two subjects was followed daily over the course of two years, it was possible to examine the overall stability of the microbial components. Among the short-term alterations, changes in fiber intake correlated with abundance changes among 15% of gut microbiota members within 1 day. Additionally, travel to the developing world had a significant effect on the microbial diversity of one of the subjects by nearly doubling the Bacteroidetes to Firmicutes ratio, which reverted upon return (David *et al.*, 2014a).

Several studies showed significant differences in gut microbiota between populations from around the world, mostly due to dietary diversity. In a large study comparing the microbiome of 531 individuals from Venezuela, rural Malawi and US metropolitan areas, there were distinct differences between groups. These differences were most likely affected by the different diets; while a typical US diet is rich in protein, the diets in Malawi and Venezuela are dominated by corn and cassava. Thus, the US samples appeared to more strongly differ from the two other groups (Yatsunenko *et al.*, 2012). In another study, large

differences in microbial populations were found between children from the rural village of Boulpon in Burkina Faso (BF) and European children (De Filippo *et al.*, 2010). While the African children's diet was high in fiber and polysaccharides, the European children followed a modern western diet. Accordingly, the African (BF) children showed significantly higher levels of Bacteroidetes and lower levels of Firmicutes, while the European children presented an opposite profile with high Firmicutes and low Bacteroidetes (see Figure 2). Additionally, the African children exhibited unique bacteria from the genus *Prevotella* and *Xylanibacter*, which contain bacterial genes enabling digestion of cellulose and xylan hydrolysis. On the other hand, *Enterobacteriaceae* levels (*Shigella* and *Escherichia*) were significantly higher in the European children. In a study of the Hadza, a community of Tanzanian hunter-gatherers, unique microbial components were observed that were consistent with their foraging lifestyle and differences in activity levels between sexes. In addition to higher levels of microbial richness and biodiversity compared to Italian urban controls, they did not harbor any *Bifidobacterium* and they exhibited differences in microbial composition between sexes. Furthermore, enrichment in *Prevotella*, *Treponema*, and unclassified *Bacteroidetes*, as well as a peculiar arrangement of *Clostridiales* taxa were observed, which may enhance the Hadza's ability to digest fibrous plant foods (Schnorr *et al.*, 2014).

Exercise

Exercise has been shown to increase bacterial diversity in humans. In a study comparing the gut microbial composition of rugby athletes to healthy male controls, the number of phyla, families and genera were all doubled in the athletes (Clarke *et al.*, 2014). Changes in microbial composition due to exercise have also been shown in a rat model in which exercised rats had a different microbiota leading to greater levels of n-butyrate compared to sedentary rats (Matsumoto *et al.*, 2008). Two additional rodent studies have shown a higher abundance of *Lactobacillus* following exercise (Choi *et al.*, 2013; Petriz *et al.*, 2014).

Antibiotics

As expected, use of antibiotics affects not only the targeted pathogen, but also the overall microbial composition. While most of these changes are short-term, and recovery to baseline microbial composition occurs within a few days or weeks, long-term effects have also been reported. The typical short-term alterations of the microbiota due to antibiotics include decreased taxonomic richness, evenness, and diversity, but the precise effects vary between patients (Dethlefsen *et al.*, 2008). One study reported that a week-long treatment of clarithromycin and metronidazole led to an immediate decrease in Actinobacteria abundance (Jakobsson *et al.*, 2010). Another study showed that levels of Firmicutes were most affected after treatment with eight different types of antibiotics (Maurice *et al.*, 2013). However, long-term effects of antibiotics on the microbiome including depletion of certain taxa and altered microbial communities have been shown to last more than six months, and even up to four years after short-term exposure (Jakobsson *et al.*, 2010). The effects of antibiotics on the

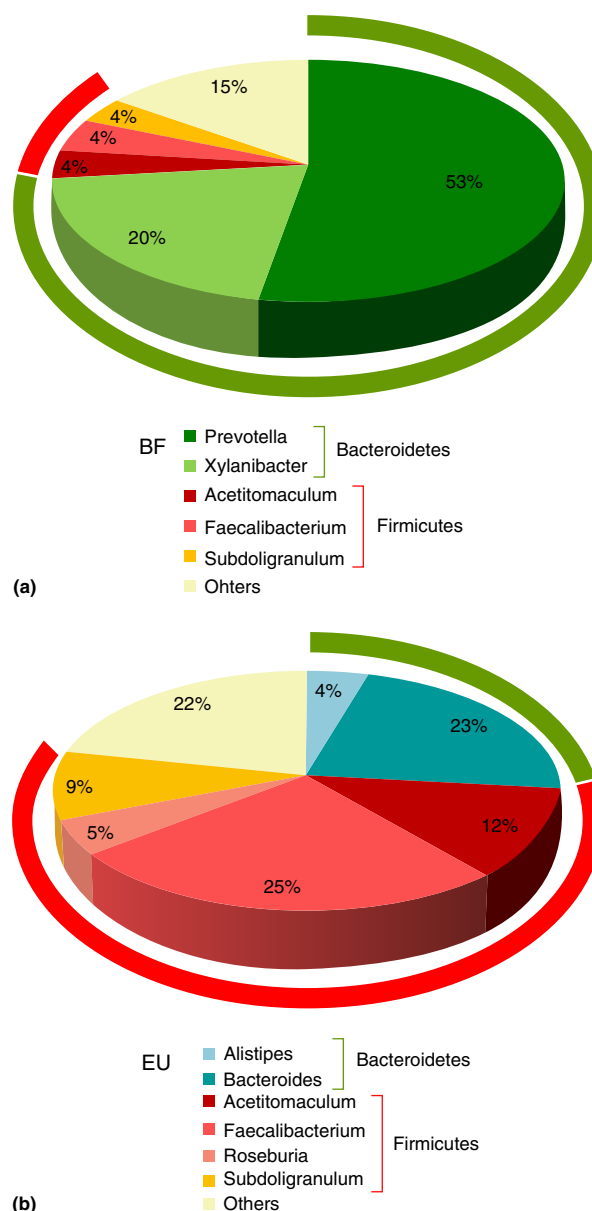


Figure 2 16S rRNA gene surveys reveal a clear separation of two populations of infants. (a and b) Pie charts of median values of bacterial genera present in fecal samples of rural Burkina Faso (a) versus European children (b), as identified by RDP classifier v. 2.1. Rings represent corresponding phyla (Bacteroidetes in green and Firmicutes in red) for each of the most frequently represented genera. (Adapted from De Filippo, C., *et al.*, 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences of the United States of America* 107 (33), 14691–14696).

microbiome can be especially crucial in infancy. In a mice study in which low-dose penicillin (LDP) was delivered from birth, perturbations in microbial communities were observed, as well as sustained alterations in host metabolism and adiposity. Transplantations of the microbiota resulting from LDP to germ-free mice proved that the microbiota had a causative effect on changes in metabolism and adiposity (Cox *et al.*,

2014). Furthermore, as one of the roles of the healthy microbiota is to enhance resistance to microbial pathogens, long-term use of antibiotics can weaken this defense mechanism, leading to susceptibility and subsequent disease (Keeney *et al.*, 2014). In conclusion, antibiotics have a significant effect on microbiota composition, a factor that should be taken into clinical consideration.

The Microbiome and Disease

Changes in the composition of the microbiome (dysbiosis) are associated with a growing list of diseases, including obesity, IBD, diabetes, metabolic syndrome, and others ((Spor *et al.*, 2011); see Figure 3). Such associations raise the question of whether the dysbiosis contributes to, or is a result of the disease. Microbiome transplant experiments have demonstrated a role of the microbiome in several diseases: several phenotypes could be transferred via the microbiota, such as enhanced fat gain (Ridaura *et al.*, 2013; Turnbaugh *et al.*, 2009a), metabolic syndrome (Vijay-Kumar *et al.*, 2010), colitis (Garrett *et al.*, 2007) and increased insulin sensitivity (Vrieze *et al.*, 2012). There have also been reports of successful fecal transplantation from healthy individuals to treat IBD patients (Bennet and Brinkman, 1989; Borody *et al.*, 2003). Together, these transplantation-based studies imply that the specific composition of the microbiota can be an important factor in the onset, and possibly the progression of chronic diseases.

Microbiota have also been implicated in a growing number of additional disease states including: autoimmune diseases, autism, cancer, *Clostridium difficile* infection (CDI), and necrotizing enterocolitis (NEC).

C. difficile Infection

This condition includes *C. difficile* overgrowth in the digestive tract, which can cause mucosal inflammation and damage, leading to colitis. The cause in many cases is prolonged use of broad-spectrum antibiotics which lead to dysbiosis and elimination of many microbial components, while enabling *C. difficile* overgrowth. Fecal microbial transplants have proven to be effective for treatment of severe CDI (Koenigsnecht and Young, 2013). While CDI is associated with reduced microbial diversity and richness, these are regained following fecal transplants (Shahinas *et al.*, 2012). Furthermore, while the microbial populations in CDI are dominated by Proteobacteria and Bacilli, following treatment, there are more species from the Bacteroidetes and Firmicutes phyla, including butyrate-producing bacteria (Fuentes *et al.*, 2014). Potentially, a combination of stable probiotics may lead to efficient treatment of CDI, and is therefore a major goal.

Obesity

The obesity-related changes to the microbiota are probably the most extensively studied so far. Changes in the abundances of bacterial phyla were first linked to obesity in mice. It was observed that the microbiota of obese mice was less diverse and that the ratio of Firmicutes to Bacteroidetes increased as

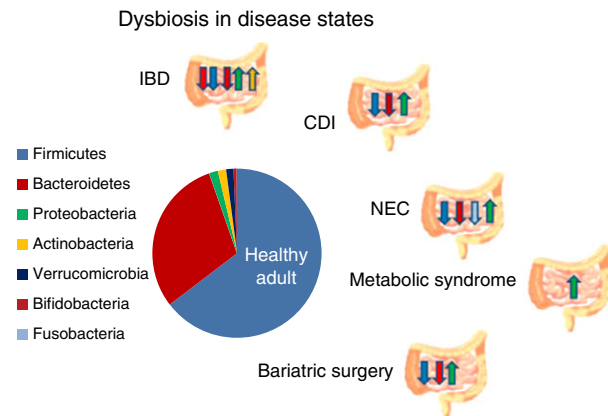


Figure 3 Dysbiosis in disease states. The gut microbial composition changes throughout various disease states, with typical bacterial phyla becoming overrepresented (up arrows), or under represented (down arrows) compared to the healthy control distribution. IBD- irritable bowel disease; CDI, *Clostridium difficile* infection; NEC, necrotizing enterocolitis. Data combined from sources Vijay-Kumar, M., *et al.*, 2010. Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. *Science* 328 (5975), 228–231; Fuentes, S., *et al.*, 2014. Reset of a critically disturbed microbial ecosystem: Fecal transplant in recurrent *Clostridium difficile* infection. *ISME Journal* 8 (8), 1621–1633; Zhang, H., *et al.*, 2009. Human gut microbiota in obesity and after gastric bypass. *Proceedings of the National Academy of Sciences of the United States of America* 106 (7), 2365–2370; Furet, J.P., *et al.*, 2010. Differential adaptation of human gut microbiota to bariatric surgery-induced weight loss: Links with metabolic and low-grade inflammation markers. *Diabetes* 59(12), 3049–3057; Morgan, X.C., *et al.*, 2012. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biology* 13 (9), R79; Shanahan, F., 2012. The colonic microbiota and colonic disease. *Current Gastroenterology Reports* 14 (5), 446–452; Lepage, P., *et al.*, 2011. Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. *Gastroenterology* 141 (1), 227–236; Rowan, F., *et al.*, 2010. *Desulfovibrio* bacterial species are increased in ulcerative colitis. *Diseases of the Colon and Rectum* 53 (11), 1530–1536; Fujimoto, T., *et al.*, 2012. Decreased abundance of *Faecalibacterium prausnitzii* in the gut microbiota of Crohn's disease. *Journal of Gastroenterology and Hepatology*; and Wang, Y., *et al.*, 2009. 16S rRNA gene-based analysis of fecal microbiota from preterm infants with and without necrotizing enterocolitis. *ISME Journal* 3 (8), 944–954.

mice became fatter (Ley *et al.*, 2005). This change was also observed in 12 humans who were on a 52-week weight-loss diet; as the subjects lost weight, the ratio of Firmicutes to Bacteroidetes decreased (Ley *et al.*, 2006). A different study tested energy extraction as an attribute of microbiota composition and weight. It was found that the more overweight people were, the more efficiently their microbiota extracted energy from the same amount of food (Turnbaugh *et al.*, 2006). Controversially, there have been other studies that exhibit different results. For example, a few human studies described a greater Bacteroidetes abundance in overweight (Schwiertz *et al.*, 2010) or a greater Firmicutes abundance in lean people (Zhang *et al.*, 2009). Other studies did not observe any changes at the phylum level of Firmicutes and Bacteroidetes (Armougom *et al.*, 2009), or found changes in different phyla such as Actinobacteria (Turnbaugh *et al.*, 2009a). Some

of the recent human studies investigated changes at other taxonomic levels, e.g. family, genus and species etc. (Ley *et al.*, 2005). Despite these different observations, the consensus is that the microbiota composition changes in obesity, and that these changes are associated with inflammation, insulin resistance and adiposity (Ley *et al.*, 2005). Microbiota transplantation studies in GF mice have shown that transplantation of microbiota from obese individuals results in mice with increased body mass and adiposity. These effects can be reversed when mice are co-housed with mice receiving microbiota transplants from lean individuals, or when mice receive a diet characterized by low levels of saturated fats and high consumption of fruits and vegetables (Ridaura *et al.*, 2013).

Bariatric Surgery

Another direction for studying the role of bacteria in obesity is by looking at the microbial effects of a Roux-en-Y-gastric bypass (RYGB), a treatment for obesity, which results in rapid weight loss. Several studies have looked at the effects of bariatric surgery on the diversity of the microbiota in human, rat and mouse fecal samples, and reported fairly congruent trends (Zhang *et al.*, 2009; Furet *et al.*, 2010; Li *et al.*, 2011; Liou *et al.*, 2013). A 16S rRNA microbial composition survey (Zhang *et al.*, 2009) reported decreased levels of Firmicutes and a proportional increase in the γ -Proteobacteria in patients who underwent surgery compared to normal-weight and obese humans. A different study using a quantitative PCR approach to quantify levels of specific groups of bacteria in fecal samples derived from humans before and after RYGB, reported an increase in the *Bacteroides/Prevotella* group and *Escherichia coli* three months after surgery and a decrease in the abundance of Bifidobacteria as well as subtypes of Firmicutes, notably the *Lactobacillus/Leuconostoc/Pediococcus* group (Furet *et al.*, 2010). A study looking at the effect of RYGB on the fecal microbiota of non-obese rats also reported a decrease in the abundance of Firmicutes and a substantial increase in the abundance of Proteobacteria, especially of *Enterobacter hormaechei* (Li *et al.*, 2011). An additional study investigating the effect of RYGB in mice reported that there was an increase in Proteobacteria and Verrucomicrobia (Liou *et al.*, 2013). This study also reported that transferring the microbiota from operated mice to GF mice resulted in weight loss and decreased fat mass in the GF mice. This study demonstrates that the microbiota plays an active role in the success of the surgery.

Metabolic Syndrome

Metabolic syndrome describes a group of disorders that when concurrent, increase the risk for type 2 diabetes and cardiovascular disease. These disorders also include excess weight; however, unlike obesity in which the microbiota is suggested to extract energy more efficiently, in metabolic syndrome the hypothesized mechanism is different. Mice with metabolic syndrome have higher levels of Proteobacteria and lower diversity than their healthy counterparts. Recently it was shown that changes in the microbiota composition correlated with metabolic changes and that the metabolic syndrome phenotype could be transferred to GF mice simply through transfer of the

microbiota (Vijay-Kumar *et al.*, 2010). The mechanism proposed for the involvement of the microbiota is their induction of low-grade inflammatory signaling, which in turn, could lead to metabolic syndrome (Vijay-Kumar *et al.*, 2010).

Malnutrition

Since diet plays an important factor in shaping the microbiota composition, it is not surprising that malnutrition affects the microbiota. In a study of Bangladeshi children with severe acute malnutrition (SAM) it was found that their microbiota composition was distinct from healthy children, and only partially changed after receiving caloric replenishment (Subramanian *et al.*, 2014). Comparing both healthy and SAM children at different ages enabled definition of the normal microbial composition by age, and indicated that the SAM children retained an immature microbial composition, even after nutritional treatment.

IBD (Ulcerative Colitis, Crohn's Disease)

Ulcerative colitis (UC) and Crohn's disease (CD) are commonly known as IBD. The microbiota of IBD patients has been found to differ from the microbiota of healthy subjects, including a decrease in diversity and evenness (relative abundance of species) of the microbiota, a decrease in Firmicutes species, and an increased abundance of members of the Enterobacteriaceae (Proteobacteria) (Morgan *et al.*, 2012). Members of the Bifidobacteria and lactobacilli have also been shown to decrease in abundance in both CD and UC (Shanahan, 2012). In UC there is also an increase in Actinobacteria, Proteobacteria and *Desulfovibrio* species (Lepage *et al.*, 2011; Rowan *et al.*, 2010). In CD patients, there was a significant decrease compared to controls in the abundance of the Clostridia class members, *Faecalibacterium prausnitzii* in particular (Fujimoto *et al.*, 2012). On the other hand there was an increase in the mucosa-associated bacterial numbers and an increase in the abundance of members of the Enterobacteriaceae, and of *Ruminococcus gnavus* (Shanahan, 2012).

Cancer

There have been several studies showing effects of the microbiota on host tumor promotion. This has been shown in various organs, including the skin, colon, liver, breast, and lungs (Schwabe and Jobin, 2013). Carcinogenesis driven by microbiota is thought to result from the presence of potentially harmful microbes or by lack of protective ones (Tlaskalova-Hogenova *et al.*, 2014). Among the proposed mechanisms for tumor promotion by microbiota are induction of inflammation and barrier failure. These can be achieved by a combination of microbial alteration encouraging growth of inflammation-inducing bacteria, mucin-degrading bacteria, genotoxic bacteria, and bacteria producing cancer-promoting metabolites that discourage growth of health-promoting symbionts such as butyrate-producing bacteria (Baxter *et al.*, 2014). Additionally, the microbiome composition can affect the abundance of specific bacterial pathogens such as *Helicobacter pylori*, which was found to be carcinogenic. Many of the studies have focused on the

potential roles of the microbiota in colon carcinogenesis. It has been shown that colon cancer patients have a distinct microbiome compared to healthy controls. In a mouse model of colon cancer, microbial composition showed increased abundance of *Bacteroides*, *Odoribacter*, and *Akkermansia* genera and decreased abundance of *Prevotellaceae* and *Porphyromonadaceae* families. GF mice receiving microbial transplants from tumor-bearing mice showed increased tumorigenesis in the colon compared to GF mice receiving transplants from healthy animals (Zackular *et al.*, 2013).

Future Microbiome Research

Microbiome research is still in its infancy. While many studies have shown correlations between disease states and alterations in the microbial populations, there are still very few studies describing the effects of specific microbes and mechanisms of action. As the microbiome includes populations of many different microbes, their interactions and environmental effects will likely be complex. It is presumed that future research will reveal precise microbial-host interaction networks and mechanisms. As part of the second stage of the HMP project, current ongoing research is focused on large-scale studies of the microbiome in human health and disease in four systems: (1) pregnancy, including those resulting in preterm birth; (2) gut disease onset, using inflammatory bowel disease (IBD) as a model; (3) respiratory viral infection; and (4) onset of type 2 diabetes (T2D). These models will likely shed more light on microbial-host interactions by integrating genomic, phylogenetic, transcriptomic, and metabolic data as well as measures of the host inflammatory and clinical state (Integrative HMP (iHMP) Research Network Consortium, 2014). Such findings will hopefully lead to effective probiotics and personalized medical treatments for a variety of conditions. To date, the only condition medically treated with fecal transplants is severe contamination by *C. difficile*. However, once the microbiome components and roles will be better understood, microbiome-based treatments including specific probiotics, prebiotics, and transplants may be developed for a variety of diseases and conditions.

See also: Intracellular Infectiology: Cell Processes: Cellular Invasion by Bacterial Pathogens

References

- Aagaard, K., Ma, J., Antony, K.M., *et al.*, 2014. The placenta harbors a unique microbiome. *Science Translational Medicine* 6 (237), 237ra65.
- Abeles, S.R., Pride, D.T., 2014. Molecular bases and role of viruses in the human microbiome. *Journal of Molecular Biology* 426 (23), 3892–3906.
- Armougom, F., Henry, M., Valette, B., Raccach, D., Raoult, D., 2009. Monitoring bacterial community of human gut microbiota reveals an increase in *Lactobacillus* in obese patients and *Methanogens* in anorexic patients. *PLoS One* 4 (9), e7125.
- Bang, C., Weidenbach, K., Gutschmann, T., Heine, H., Schmitz, R.A., 2014. The intestinal archaea *Methanosphaera stadtmanae* and *Methanobrevibacter smithii* activate human dendritic cells. *PLoS One* 9 (6), e99411.
- Baxter, N.T., Zackular, J.P., Chen, G.Y., Schloss, P.D., 2014. Structure of the gut microbiome following colonization with human feces determines colonic tumor burden. *Microbiome* 2, 20.
- Bennet, J.D., Brinkman, M., 1989. Treatment of ulcerative colitis by implantation of normal colonic flora. *Lancet* 1 (8630), 164.
- Bercik, P., Denou, E., Collins, J., *et al.*, 2011. The intestinal microbiota affect central levels of brain-derived neurotrophic factor and behavior in mice. *Gastroenterology* 141 (2), 599–609.e3.
- Blais Lecours, P., Marsolais, D., Cormier, Y., *et al.*, 2014. Increased prevalence of *Methanosphaera stadtmanae* in inflammatory bowel diseases. *PLoS One* 9 (2), e87734.
- Borody, T.J., Warren, E.F., Leis, S., Surace, R., Ashman, O., 2003. Treatment of ulcerative colitis using fecal bacteriotherapy. *Journal of Clinical Gastroenterology* 37 (1), 42–47.
- Bravo, J.A., Forsythe, P., Chew, M.V., *et al.*, 2011. Ingestion of *Lactobacillus* strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proceedings of the National Academy of Sciences of the United States of America* 108 (38), 16050–16055.
- Candela, M., Biagi, E., Brigidi, P., O'Toole, P.W., De Vos, W.M., 2014. Maintenance of a healthy trajectory of the intestinal microbiome during aging: A dietary approach. *Mechanisms of Ageing and Development* 136–137, 70–75.
- Choi, J.J., Eum, S.Y., Rampersaud, E., *et al.*, 2013. Exercise attenuates PCB-induced changes in the mouse gut microbiome. *Environmental Health Perspectives* 121 (6), 725–730.
- Claesson, M.J., Cusack, S., O'Sullivan, O., *et al.*, 2011. Composition, variability, and temporal stability of the intestinal microbiota of the elderly. *Proceedings of the National Academy of Sciences of the United States of America* 108 (Suppl. 1), 4586–4591.
- Claesson, M.J., Jeffery, I.B., Conde, S., *et al.*, 2012. Gut microbiota composition correlates with diet and health in the elderly. *Nature* 488 (7410), 178–184.
- Clarke, S.F., Murphy, E.F., O'Sullivan, O., *et al.*, 2014. Exercise and associated dietary extremes impact on gut microbial diversity. *Gut* 63 (12), 1913–1920.
- Conway de Macario, E., Macario, A.J., 2009. Methanogenic archaea in health and disease: A novel paradigm of microbial pathogenesis. *International Journal of Medical Microbiology* 299 (2), 99–108.
- Cox, L.M., Yamanishi, S., Sohn, J., *et al.*, 2014. Altering the intestinal microbiota during a critical developmental window has lasting metabolic consequences. *Cell* 158 (4), 705–721.
- Crabbe, P.A., Nash, D.R., Bazin, H., Eyssen, D.V., Heremans, J.F., 1969. Antibodies of the IgA type in intestinal plasma cells of germfree mice after oral or parenteral immunization with ferritin. *Journal of Experimental Medicine* 130 (4), 723–744.
- David, L.A., Materna, A.C., Friedman, J., *et al.*, 2014a. Host lifestyle affects human microbiota on daily timescales. *Genome Biology* 15, R89.
- David, L.A., Maurice, C.F., Carmody, R.N., *et al.*, 2014b. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505 (7484), 559–563.
- De Filippo, C., Cavalieri, D., Di Paola, M., *et al.*, 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences of the United States of America* 107 (33), 14691–14696.
- Dethlefsen, L., Huse, S., Sogin, M.L., Relman, D.A., 2008. The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biology* 6 (11), e280.
- Diaz Heijtz, R., Wang, S., Anuar, F., *et al.*, 2011. Normal gut microbiota modulates brain development and behavior. *Proceedings of the National Academy of Sciences of the United States of America* 108 (7), 3047–3052.
- Eckburg, P.B., Bik, E.M., Bernstein, C.N., *et al.*, 2005. Diversity of the human intestinal microbial flora. *Science* 308 (5728), 1635–1638.
- Fuentes, S., van Nood, E., Tims, S., *et al.*, 2014. Reset of a critically disturbed microbial ecosystem: Fecal transplant in recurrent *Clostridium difficile* infection. *ISME Journal* 8 (8), 1621–1633.
- Fujimoto, T., Imaeda, H., Takahashi, K., *et al.*, 2012. Decreased abundance of *Faecalibacterium prausnitzii* in the gut microbiota of Crohn's disease. *Journal of Gastroenterology and Hepatology* 28 (4), 613–619.
- Furet, J.P., Kong, L.C., Tap, J., *et al.*, 2010. Differential adaptation of human gut microbiota to bariatric surgery-induced weight loss: Links with metabolic and low-grade inflammation markers. *Diabetes* 59 (12), 3049–3057.
- Gareau, M.G., Wine, E., Rodrigues, D.M., *et al.*, 2011. Bacterial infection causes stress-induced memory dysfunction in mice. *Gut* 60 (3), 307–317.
- Garrett, W.S., Lord, G.M., Punit, S., *et al.*, 2007. Communicable ulcerative colitis induced by T-bet deficiency in the innate immune system. *Cell* 131 (1), 33–45.
- Goodman, A.L., Kallstrom, G., Faith, J.J., *et al.*, 2011. Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. *Proceedings of the National Academy of Sciences of the United States of America* 108 (15), 6252–6257.

- Gouba, N., Raoult, D., Drancourt, M., 2014. Eukaryote culturomics of the gut reveals new species. *PLoS One* 9 (9), e106994.
- Grenham, S., Clarke, G., Cryan, J.F., Dinan, T.G., 2011. Brain-gut-microbe communication in health and disease. *Frontiers in Physiology* 2, 94.
- Hill, D.A., Artis, D., 2010. Intestinal bacteria and the regulation of immune cell homeostasis. *Annual Review of Immunology* 28, 623–667.
- Hoffmann, C., Dollive, S., Grunberg, S., *et al.*, 2013. Archaea and fungi of the human gut microbiome: Correlations with diet and bacterial residents. *PLoS One* 8 (6), e66019.
- Honda, K., Littman, D.R., 2012. The microbiome in infectious disease and inflammation. *Annual Review of Immunology* 30, 759–795.
- Hooper, L.V., Littman, D.R., Macpherson, A.J., 2012. Interactions between the microbiota and the immune system. *Science* 336 (6086), 1268–1273.
- Human Microbiome Project, C., 2012. Structure, function and diversity of the healthy human microbiome. *Nature* 486 (7402), 207–214.
- Iliev, I.D., Funari, V.A., Taylor, K.D., *et al.*, 2012. Interactions between commensal fungi and the C-type lectin receptor Dectin-1 influence colitis. *Science* 336 (6086), 1314–1317.
- Integrative HMP (iHMP) Research Network Consortium, 2014. The Integrative Human Microbiome Project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease. *Cell Host & Microbe* 16 (3), 276–289.
- Jakobsson, H.E., Jernberg, C., Andersson, A.F., *et al.*, 2010. Short-term antibiotic treatment has differing long-term impacts on the human throat and gut microbiome. *PLoS One* 5 (3), e9836.
- Jhannani, K., Shivakumar, V., Kalmady, S., Rao, N.P., Venkatasubramanian, G., 2013. Successful use of add-on minocycline for treatment of persistent negative symptoms in schizophrenia. *Journal of Neuropsychiatry and Clinical Neurosciences* 25 (1), E06–E07.
- Kamada, N., Seo, S.U., Chen, G.Y., Núñez, G., 2013. Role of the gut microbiota in immunity and inflammatory disease. *Nature Reviews Immunology* 13 (5), 321–335.
- Keeney, K.M., Yurist-Doutsch, S., Arrieta, M.C., Finlay, B.B., 2014. Effect of antibiotics on human microbiota and subsequent disease. *Annual Review of Microbiology* 68, 217–235.
- Kelly, D., Mulder, I.E., 2012. Microbiome and immunological interactions. *Nutrition Reviews* 70 (Suppl. 1), S18–S30.
- Koenig, J.E., Spor, A., Scalfone, N., *et al.*, 2011. Succession of microbial consortia in the developing infant gut microbiome. *Proceedings of the National Academy of Sciences of the United States of America* 108 (Suppl. 1), 4578–4585.
- Koenigsnecht, M.J., Young, V.B., 2013. Fecal microbiota transplantation for the treatment of recurrent *Clostridium difficile* infection: Current promise and future needs. *Current Opinion in Gastroenterology* 29 (6), 628–632.
- Koren, O., Goodrich, J.K., Cullender, T.C., *et al.*, 2012. Host remodeling of the gut microbiome and metabolic changes during pregnancy. *Cell* 150 (3), 470–480.
- La Rosa, P.S., Warner, B.B., Zhou, Y., *et al.*, 2014. Patterned progression of bacterial populations in the premature infant gut. *Proceedings of the National Academy of Sciences of the United States of America* 111 (34), 12522–12527.
- Lepage, P., Häsler, R., Spehlmann, M.E., *et al.*, 2011. Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. *Gastroenterology* 141 (1), 227–236.
- Ley, R.E., Bäckhed, F., Turnbaugh, P., *et al.*, 2005. Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences of the United States of America* 102, 11070–11075.
- Ley, R.E., Turnbaugh, P.J., Klein, S., Gordon, J.I., 2006. Microbial ecology: Human gut microbes associated with obesity. *Nature* 444 (7122), 1022–1023.
- Li, J.V., Ashrafian, H., Bueter, M., *et al.*, 2011. Metabolic surgery profoundly influences gut microbial-host metabolic cross-talk. *Gut* 60 (9), 1214–1223.
- Liou, A.P., Paziuk, M., Luevano Jr., J.-M., *et al.*, 2013. Conserved shifts in the gut microbiota due to gastric bypass reduce host weight and adiposity. *Science Translational Medicine* 5 (178), 178ra41.
- Littman, D.R., Pamer, E.G., 2011. Role of the commensal microbiota in normal and pathogenic host immune responses. *Cell Host & Microbe* 10 (4), 311–323.
- Matsumoto, M., Inoue, R., Tsukahara, T., *et al.*, 2008. Voluntary running exercise alters microbiota composition and increases n-butyrate concentration in the rat cecum. *Bioscience, Biotechnology, and Biochemistry* 72 (2), 572–576.
- Maurice, C.F., Haiser, H.J., Turnbaugh, P.J., 2013. Xenobiotics shape the physiology and gene expression of the active human gut microbiome. *Cell* 152 (1–2), 39–50.
- Minot, S., Bryson, A., Chehoud, C., *et al.*, 2013. Rapid evolution of the human gut virome. *Proceedings of the National Academy of Sciences of the United States of America* 110 (30), 12450–12455.
- Morgan, X.C., Tickle, T.L., Sokol, H., *et al.*, 2012. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biology* 13 (9), R79.
- Neufeld, K.M., Kang, N., Bienenstock, J., Foster, J.A., 2011. Reduced anxiety-like behavior and central neurochemical change in germ-free mice. *Neurogastroenterology and Motility: The Official Journal of the European Gastrointestinal Motility Society* 23 (3), 255–264. e119.
- Ostman, S., Rask, C., Wold, A.E., Hultkrantz, S., Telmo, E., 2006. Impaired regulatory T cell function in germ-free mice. *European Journal of Immunology* 36 (9), 2336–2346.
- Petriz, B.A., Castro, A.P., Almeida, J.A., *et al.*, 2014. Exercise induction of gut microbiota modifications in obese, non-obese and hypertensive rats. *BMC Genomics* 15, 511.
- Qin, J., Li, R., Raes, J., *et al.*, 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464 (7285), 59–65.
- Rampelli, S., Candela, M., Turroni, S., *et al.*, 2013. Functional metagenomic profiling of intestinal microbiome in extreme ageing. *Aging (Albany, NY)* 5 (12), 902–912.
- Ridaura, V.K., Faith, J.J., Rey, F.E., *et al.*, 2013. Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* 341 (6150), 1241214.
- Rowan, F., Docherty, N.G., Murphy, M., *et al.*, 2010. *Desulfovibrio* bacterial species are increased in ulcerative colitis. *Diseases of the Colon and Rectum* 53 (11), 1530–1536.
- Scanlan, P.D., Stensvold, C.R., Rajilic-Stojanovic, M., *et al.*, 2014. The microbial eukaryote *Blastocystis* is a prevalent and diverse member of the healthy human gut microbiota. *FEMS Microbiology Ecology* 90 (1), 326–330.
- Schnorr, S.L., Candela, M., Rampelli, S., *et al.*, 2014. Gut microbiome of the Hadza hunter-gatherers. *Nature Communications* 5, 3654.
- Schwabe, R.F., Jobin, C., 2013. The microbiome and cancer. *Nature Reviews Cancer* 13 (11), 800–812.
- Schwiertz, A., Taras, D., Schäfer, K., *et al.*, 2010. Microbiota and SCFA in lean and overweight healthy subjects. *Obesity* 18 (1), 190–195.
- Shahinas, D., Silverman, M., Sittler, T., *et al.*, 2012. Toward an understanding of changes in diversity associated with fecal microbiome transplantation based on 16S rRNA gene deep sequencing. *mBio* 3 (5), e00338–12.
- Shanahan, F., 2012. The colonic microbiota and colonic disease. *Current Gastroenterology Reports* 14 (5), 446–452.
- Sharon, G., Segal, D., Ringo, J.M., *et al.*, 2010. Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America* 107 (46), 20051–20056.
- Simon, G.L., Gorbach, S.L., 1984. Intestinal flora in health and disease. *Gastroenterology* 86 (1), 174–193.
- Soczynska, J.K., Mansur, R.B., Brietzke, E., *et al.*, 2012. Novel therapeutic targets in depression: Minocycline as a candidate treatment. *Behavioural Brain Research* 235 (2), 302–317.
- Spor, A., Koren, O., Ley, R., 2011. Unraveling the effects of the environment and host genotype on the gut microbiome. *Nature Reviews. Microbiology* 9 (4), 279–290.
- Subramanian, S., Huq, S., Yatsunenko, T., *et al.*, 2014. Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature* 510 (7505), 417–421.
- Tanabe, S., 2013. The effect of probiotics and gut microbiota on Th17 cells. *International Reviews of Immunology* 32 (5–6), 511–525.
- Tillisch, K., Labus, J., Kilpatrick, L., *et al.*, 2013. Consumption of fermented milk product with probiotic modulates brain activity. *Gastroenterology* 144 (7), 1394–1401.e4.
- Titkaloova-Hogenova, H., Vannucci, L., Klimesova, K., *et al.*, 2014. Microbiome and colorectal carcinoma: Insights from germ-free and conventional animal models. *Cancer Journal* 20 (3), 217–224.
- Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., *et al.*, 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444 (7122), 1027–1031.
- Turnbaugh, P.J., Hamady, M., Yatsunenko, T., *et al.*, 2009b. A core gut microbiome in obese and lean twins. *Nature* 457 (7228), 480–484.
- Turnbaugh, P.J., Ridaura, V.K., Faith, J.J., *et al.*, 2009a. The effect of diet on the human gut microbiome: A metagenomic analysis in humanized gnotobiotic mice. *Science Translational Medicine* 1 (6), 6ra14.
- Underhill, D.M., Iliev, I.D., 2014. The mycobiota: Interactions between commensal fungi and the host immune system. *Nature Reviews Immunology* 14 (6), 405–416.
- Ursell, L.K., Haiser, H.J., Van Treuren, W., *et al.*, 2014. The intestinal metabolome: An intersection between microbiota and host. *Gastroenterology* 146 (6), 1470–1476.

- Vijay-Kumar, M., Aitken, J.D., Carvalho, F.A., *et al.*, 2010. Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. *Science* 328 (5975), 228–231.
- Vrieze, A., Van Nood, E., Holleman, F., *et al.*, 2012. Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology* 143 (4), 913–916. **e7**.
- Walker, J.R., Ediger, J.P., Graff, L.A., *et al.*, 2008. The Manitoba IBD cohort study: A population-based study of the prevalence of lifetime and 12-month anxiety and mood disorders. *American Journal of Gastroenterology* 103 (8), 1989–1997.
- Wu, G.D., Chen, J., Hoffmann, C., *et al.*, 2011. Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334 (6052), 105–108.
- Wylie, K.M., Mihindukulasuriya, K.A., Zhou, Y., *et al.*, 2014. Metagenomic analysis of double-stranded DNA viruses in healthy adults. *BMC Biology* 12 (1), 71.
- Yatsunenko, T., Rey, F.E., Manary, M.J., *et al.*, 2012. Human gut microbiome viewed across age and geography. *Nature* 486 (7402), 222–227.
- Zackular, J.P., Baxter, N.T., Iverson, K.D., *et al.*, 2013. The gut microbiome modulates colon tumorigenesis. *mBio* 4 (6), e00692–13.
- Zhang, H., DiBaise, J.K., Zuccolo, A., *et al.*, 2009. Human gut microbiota in obesity and after gastric bypass. *Proceedings of the National Academy of Sciences of the United States of America* 106 (7), 2365–2370.

Cell Biology of Virus Infection

A Helenius and Y Yamauchi, Institute for Biochemistry, Zürich, Switzerland

© 2016 Elsevier Inc. All rights reserved.

Viruses are best known as infectious agents that cause diseases and epidemics in humans and livestock. Unlike other infectious agents such as bacteria, they are, however, extremely small and simple in structure. The simplest among them contain one nucleic acid molecule (the viral genome) and a single protein in multiple copies forming the protective shell. Such simplicity is only possible because viruses have no metabolism, no means of locomotion, no source of energy, and no biosynthetic machinery. Their replication strategy differs completely from that of bacteria and other cells. Instead of multiplying by division, they replicate by taking over live cells and forcing them to synthesize new virus particles. The 'blueprint' is provided by the viral DNA or RNA genome, the building blocks and the biosynthetic machinery by the host cell. New virus particles are often produced in huge numbers, and serve as the vehicle for transmission of infection to further cells and organisms. The host cell generally ends up dying in the process. To understand viruses and their replication cycle, one must understand the biology of their host cells and host organisms. This is why cell biology is an important aspect of virological research and infectious disease medicine.

Our world is filled with thousands of different types of viruses. They are omnipresent, and capable of infecting all forms of life: bacteria, slime molds, plants, insects, fish, mammals, etc. Some are specific to a single host species, and may have a long common evolutionary history with their host organism. Polio-, small pox, measles, and Herpes Simplex are examples of viruses specific for humans. Other viruses can move between different organisms such as influenza A virus (IAV), yellow fever virus, and Western Equine Encephalitis virus, to name a few. The causative agent of AIDS, (human immunodeficiency virus-1) HIV-1, was originally a monkey virus.

The Virus Particle

Viruses can be classified according to several criteria. **Figure 1** shows examples of animal viruses from different virus families. Enveloped viruses have a lipid bilayer membrane; non-enveloped viruses are composed of nucleic acid and protein only. Their size varies from 25 to about 400 nm in diameter. While many viruses are spherical in shape, some are filamentous, like the Ebola virus, or bullet-shaped, like the rabies virus. A classification based on the genome, which can be DNA or RNA, and its properties (single-stranded or double-stranded, positive or negative sense, etc.), is often useful when considering the cell biology of infection.

Although different in size, shape, and overall appearance, viruses have in common that they contain a genome in the form of one or multiple molecules of DNA or RNA. The genome is present in a condensed and protected state, which together with proteins form a structure called the capsid ([Harrison, 2013](#)). One common capsid design involves hexameric and pentameric assemblies of capsid proteins. These are joined together according to rules of icosahedral symmetry to form a protein shell. This soccer ball-like architecture generates a protected space for the genome using a minimum number of different building elements.

In many non-enveloped viruses such as polio-, adeno-, and parvoviruses, the classical icosahedron-like capsid plus the genome constitute the entire virus particle. More complex viruses can, however, combine an icosahedral capsid with additional layers of protein, as well as a lipid bilayer envelope (herpes-, alpha-, and retroviruses) (**Figure 1**).

An alternative capsid structure is observed in many RNA viruses. Here numerous capsid proteins interact with the RNA over its full length forming a helical RNA/protein complex. Tobacco mosaic virus is the classical example of a

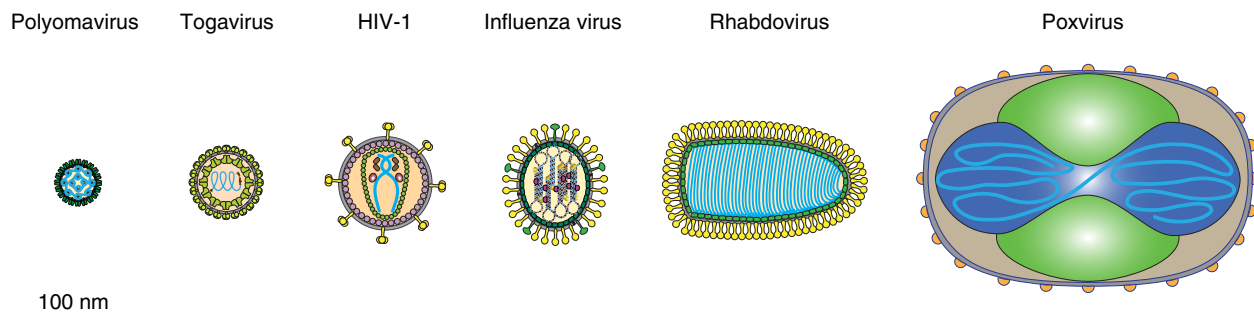


Figure 1 Animal viruses have many forms and sizes. Polyomaviruses are non-enveloped and have a naked icosahedral capsid (45 nm) with a circular dsDNA genome packaged as a minichromosome around histones. Togaviruses (alphaviruses, 70 nm) are enveloped with an icosahedral capsid with a positive-sense ssRNA that is capped at the 5' end and poly-adenylated at the 3' end. HIV-1 is enveloped (100 nm) with a conical capsid and a positive-sense ssRNA genome. Like other retroviruses, the capsid contains two copies of the genome associated with tRNA from the cell. Influenza viruses are enveloped, and can be round (80–120 nm) or filamentous with a negative-sense ssRNA wrapped around nucleoproteins and separated in 6–8 segments. Rhabdoviruses are bullet-shaped, rod-like, and enveloped (180 × 75 nm), with a genome composed of a single negative-sense ssRNA. Poxviruses are brick-shaped (300 × 250 × 150 nm) with an internal core (blue) and lateral bodies (green). The scheme shows a mature virus (MV) with one membrane. The genome is composed of dsDNA. Bar: 100 nm.

non-enveloped helical virus. Many enveloped RNA viruses that infect humans have helical capsids trapped within a lipid bilayer envelope. These include IAV, measles virus, and Rift Valley fever virus.

A single virus particle has everything needed to infect a cell. The proteins present in the particle in one or more copies are called structural proteins. Depending on the virus, the structural proteins can include – in addition to capsid proteins – enzymes such as RNA polymerases and host proteins such as histones. Nonstructural proteins are virally encoded and expressed within the infected cells, but they are not part of the mature virus particle. They are important because they have a variety of functions in nucleic acid replication, virus assembly, and in the suppression of host cell defense mechanisms. In a virus population only a minority of particles actually succeed in infecting cells.

Many of the enveloped viruses are surprisingly heterogeneous in size, shape, and composition and display rather poor infectivity if analyzed at the level of individual particles. However, it has been shown that one infectious virus particle is sufficient to infect a cell.

Transmission

The spread of virus infection occurs at several levels; from cell to cell, tissue to tissue, organism to organism, and population to population. Transmission to humans can occur via airborne droplets, direct contact, insect bites, injection needles, food items, open wounds, congenital transmission, etc. Most viruses enter through epithelial barriers (the skin, the airways, the digestive tract, etc.). Once initiated, infection may be cell type-specific and strictly localized, or the virus may spread widely and cause infection of many cell types and tissues. Many virus infections are asymptomatic and held in check by the immune system.

Introduction to the Replication Cycle

The mission of the virus particle as such is to mediate the transport of the viral genome and sometimes accessory proteins from the infected cell to a new host cell. The virus must be able to undergo proper assembly and release from the producing cells. It must protect the viral genome in the extracellular space during transit. Finally, its job is to deliver the capsid into the cytosol or the nucleus of a new host cell and release it. These are not trivial duties because they involve, for example, the transfer of large polar macromolecular assemblies through cellular membranes. In addition, they require orderly assembly and disassembly of the particle.

Either in the cytosol (for most RNA viruses) or the nucleus (for most DNA viruses), the uncoated genome is used for the synthesis of viral mRNAs and replication of new viral genomes, leading to amplification of viral components in the cell (Whelan, 2013). Enough viral components to assemble thousands of new virus particles (nucleic acids, proteins, glycoproteins, and lipids) have to be synthesized. The components are often generated in different locations inside the cell. The logistics of bringing them together for genome packaging, particle assembly, intracellular trafficking, and release is

complicated. The process often involves many cellular compartments and numerous viral and cellular factors. The high degree of dependence on the host cell in all stages of infection from entry to release is one of the reasons why cell biology is so important for understanding the life-style of viruses.

Virus Entry

Like most of the steps in the replication cycle, virus entry is critically dependent on assistance by the host cell. The virus needs receptor expression, membrane traffic, signal transduction, cytoskeleton motility, nuclear import, etc. As illustrated in Figure 2, the process of entry occurs in several steps. The majority of viruses make use of the endocytic machinery to enter cells (Mercer and Helenius, 2009; Mercer *et al.*, 2010; Yamauchi and Helenius, 2013). This strategy offers advantages including a free ride on the cytoskeleton from the surface to the perinuclear region inside endosomal vesicles. Passage through the crowded, obstacle-rich cytoplasm is thus made easy. By having the penetration step occur from intracellular vacuoles, the exposure of tell-tale viral components on the plasma membrane is minimized, which helps to delay detection by antibodies and immune cells. There are, however, some enveloped viruses that can fuse directly with the plasma membrane. In addition to HIV-1 and some other retroviruses, these include members of the herpes- and paramyxovirus families (Jardetzky and Lamb, 2014; Eisenberg *et al.*, 2012; Klasse, 2012).

After internalization by endocytosis, viruses are sorted into a network of endocytic vacuoles that comprise a dynamic, interacting collection of organelles that the cell uses to deal with the incoming vesicle traffic of different cargo molecules and hundreds of signaling receptors. Penetration of viruses into the cytosol generally occurs through the limiting membranes of endosomes or macropinosomes often triggered by low pH and/or proteolytic activation of viral proteins. By providing such biochemical cues, the endocytic organelles inadvertently ‘inform’ the incoming particles when and where it is time to enter the cytosol. Viruses that fail to respond correctly to these signals are delivered to lysosomes and eventually degraded.

The entry program is completed only when the viruses or viral capsids have reached a final destination in the cytosol or nucleus, and the genome has been uncoated so that it can be used as template for mRNA production or genome replication. The uncoating process often occurs in several steps at different stages during the entry process (Suomalainen and Greber, 2013).

The details of virus entry are highly variable depending on virus and cell type. However, the overall profile of events is similar for most viruses. The major steps will be discussed below without going into great detail.

Receptor Binding

Attachment of virus particles to the cell surface occurs by the interaction of viral proteins with cell surface proteins or carbohydrates. In enveloped viruses, binding to surface molecules occurs through specific envelope glycoproteins. Non-

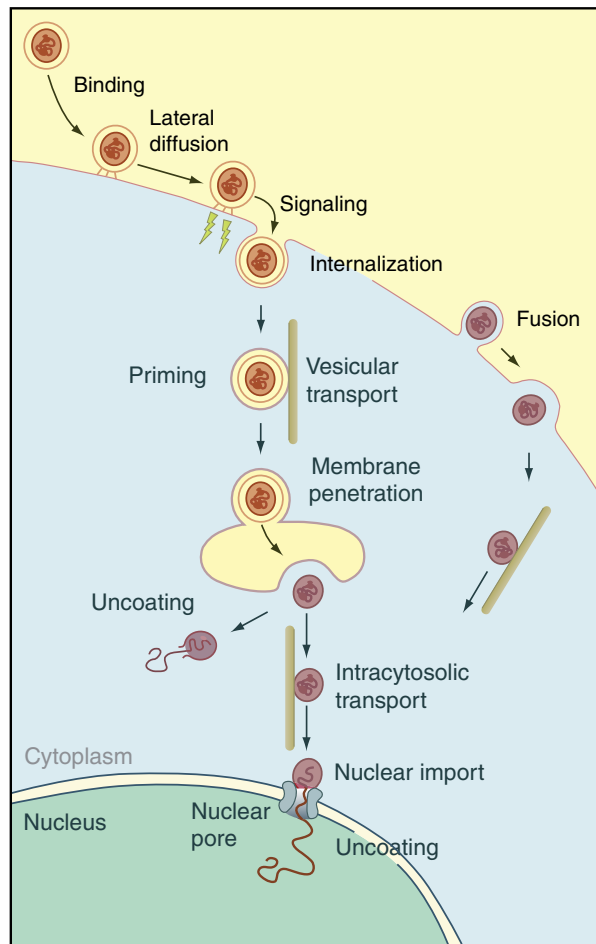


Figure 2 Stepwise host cell entry of animal viruses. When viruses enter, they follow a stepwise program that allows them to travel from the extracellular space to the site of replication in the cytoplasm or the nucleus of the host cell. Entry begins by binding to receptors on the cell surface and by lateral diffusion of viruses along the membrane. Receptor association and clustering activate signaling cascades that trigger endocytic internalization of virus particles. The viruses move into the cytoplasm as cargo particles inside endocytic vesicles and vacuoles. Cellular cues, such as low pH in these vacuoles, triggers changes in the virus particles or in viral proteins that activate membrane fusion or other mechanism of penetration into the cytosol. After penetration, some viral capsids stay in the cytosol where they undergo uncoating and replication. Others use the microtubule transport and nuclear import machineries to reach and enter the nucleus. Some enveloped viruses can enter cells directly by fusing with the plasma membrane. The details in the entry program vary for different viruses and cell types, but many of the key steps shown here are general. Modified from Marsh and Helenius (2006).

enveloped viruses usually bind through sites in the virus shell unless – as in the case of adenoviruses – specialized antennae are present (Grove and Marsh, 2011).

The binding partners on the cell surface are called ‘viral receptors,’ although their reason for being there has to do with normal cellular functions. Hundreds of different receptors for viruses have been identified. The list includes receptors for physiological ligands such as the transferrin receptor,

components of the extracellular matrix, surface antigens of the immunoglobulin family, ion channels, receptor tyrosine kinases, glycoproteins and proteoglycans. Viruses of the polyoma virus family have evolved to bind specifically to the glycan moieties of specific gangliosides, thus generating raft-like domains enriched in cholesterol and sphingolipids, activating an unusual endocytic trafficking pathway to the ER (Tsai and Qian, 2010; Qian *et al.*, 2009). Receptor-binding is often specific and therefore important in defining the tropism of viruses as well as eventually determining the type of disease.

By binding to CD4 and a cytokine receptor (Figure 3), HIV-1 is an example of viruses that need more than one receptor type to enter cells. In this case, the two receptors trigger a conformation change in the viral glycoprotein that activates its membrane fusion activity (Klasse, 2012). To be able to infect a variety of cell types and use different entry pathways, herpes simplex virus type 1 (HSV-1) carries multiple receptor-binding glycoproteins with different specificity (Eisenberg *et al.*, 2012).

Compared to most cellular ligands, viruses are large and multivalent. By binding simultaneously to many receptors, they can generate a patch or domain in the plasma membrane with unique composition and properties (Szklarczyk *et al.*, 2013). This can in turn result in trans-bilayer signaling, coat assembly on the cytosolic side, and changes in lateral mobility.

Live cell imaging of fluorescent virus particles has shown that immediately after attachment most virus particles are motile on the surface of cells. Some surf on filopodia propelled in the direction of the cell body by dynamic actin flow; others undergo random or directed movement on the cell body itself (Lehmann *et al.*, 2005; Burckhardt and Greber, 2009; Schelhaas *et al.*, 2008). Typically, the movement stops before the viruses are internalized by endocytosis. Several viruses have been shown to induce formation of clathrin coats and vesicles at the site of binding (Cureton *et al.*, 2010). Thus, they are capable of triggering their own endocytic internalization.

Signaling

Viruses are known to activate cellular signaling pathways and thus trigger a variety of changes in the cell, such as ruffling of the plasma membrane and endocytosis through macropinocytosis and other mechanisms. Some do this by interacting directly with signaling receptors, others by indirectly carrying phosphatidylserine (PS) in their envelope membrane. When exposed, this negatively charged phospholipid activates receptors, signaling pathways, and endocytic mechanisms that are normally reserved for uptake of residual bodies left behind after apoptotic cell death. In this way, pox- and filoviruses use ‘apoptotic mimicry’ as their strategy to enter cells (Mercer and Helenius, 2008, 2012; Moller-Tank *et al.*, 2013). Other viruses activate other signaling pathways.

Endocytosis and the Endosome Network

Many of the mechanisms of endocytosis known to cell biologists are exploited by viruses depending on particle size, receptor choice, and cell type (Figures 4 and 5). The clathrin-mediated pathway is commonly used by small and medium

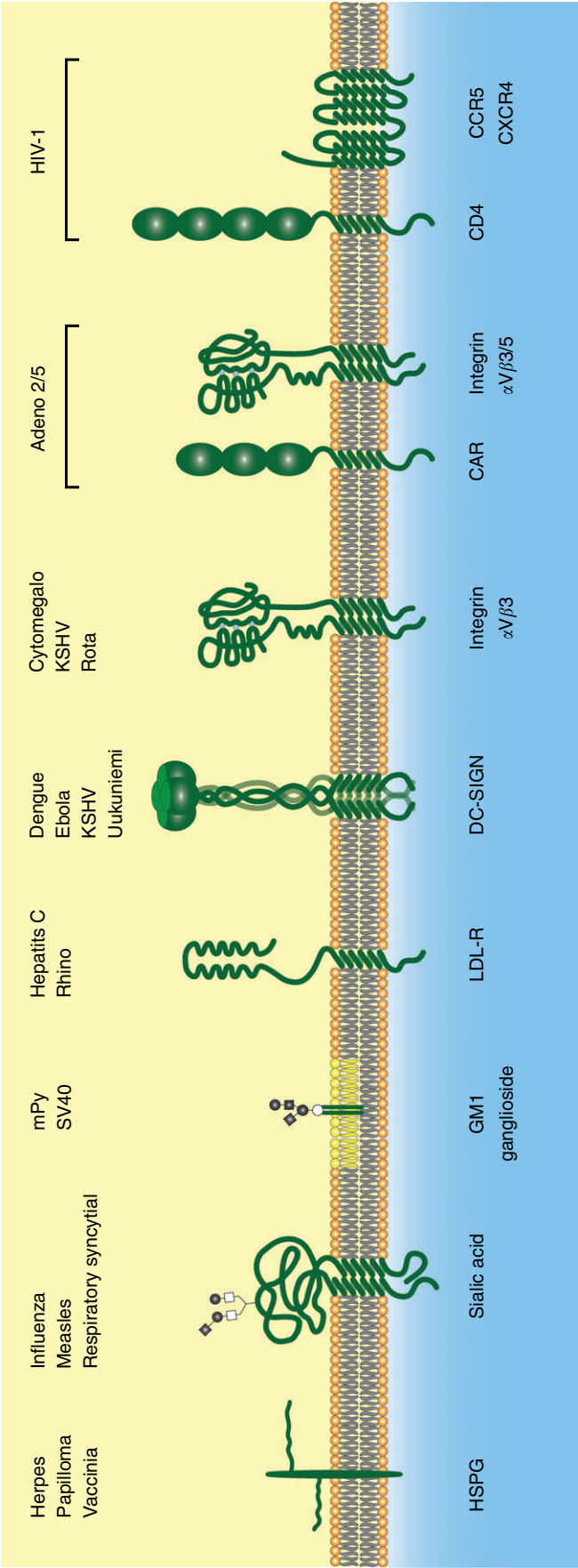


Figure 3 Attachment factors and receptors. Cell surface proteins, lipids, and carbohydrates can serve as virus attachment factors and receptors. This schematic shows examples of virus receptors and viruses that can bind to them. Some viruses such as adenoviruses 2 and 5 as well as HIV-1 require more than one receptor. mPy, mouse polyoma virus; KSHV, Kaposi's sarcoma-associated herpesvirus; HSPG, Heparan sulphate proteoglycan; LDL-R, low-density lipoprotein receptor; DC-SIGN, dendritic cell-specific intercellular adhesion molecule-3-grabbing nonintegrin; CAR, coxsackievirus and adenovirus receptor; CCR and CXCR, chemokine receptor.

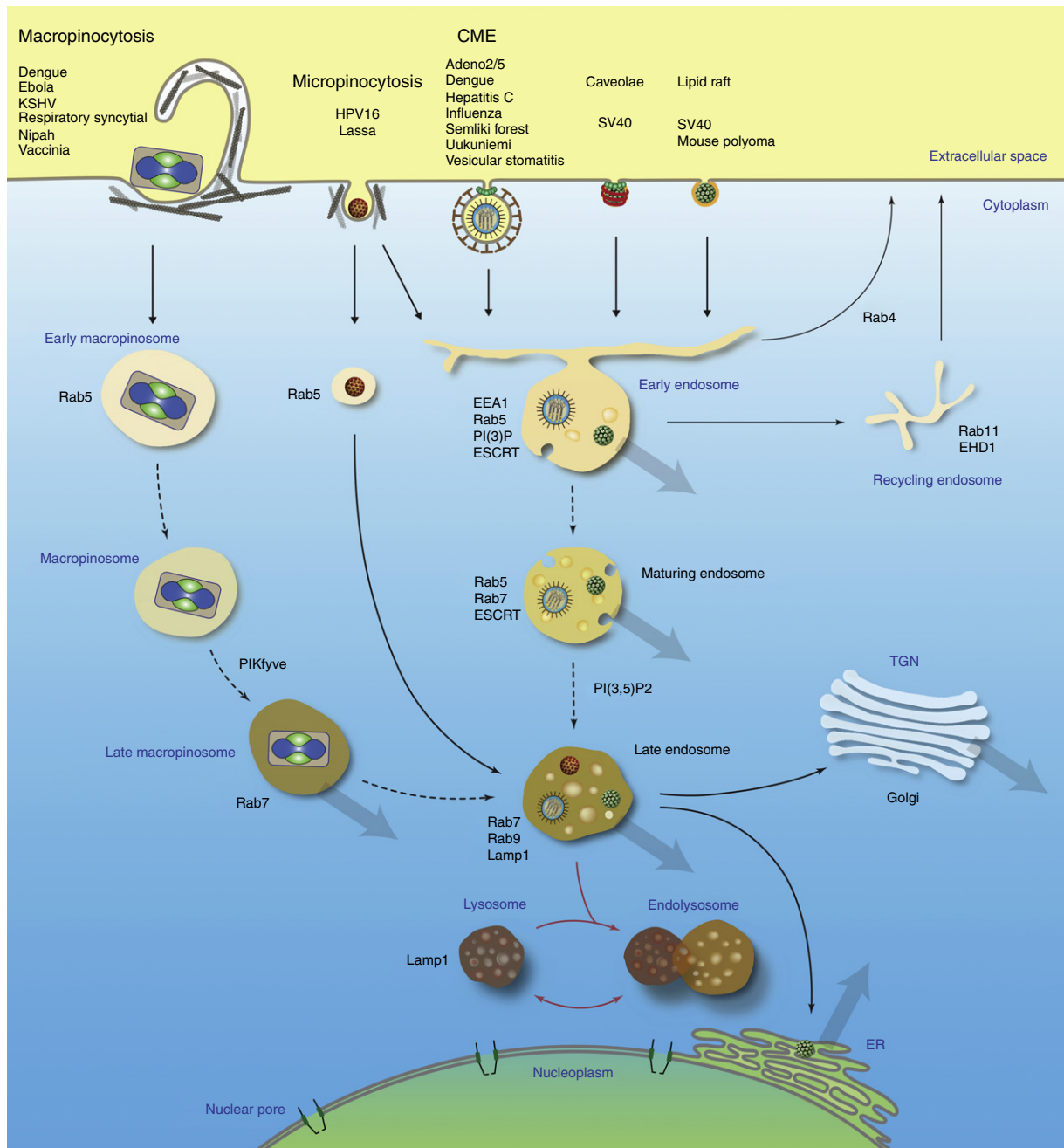


Figure 4 Endocytic entry pathways for animal viruses. The majority of viruses use endocytosis for entry. For this, they take advantage of different endocytic mechanisms. Many viruses of small and intermediate size (40–200 nm) use clathrin-mediated endocytosis (CME), or caveolar and lipid raft-mediated mechanisms involving tight-fitting vesicles. CME and caveolar uptake are dynamin-dependent (green beads). The virus virus-carrying vesicles and vacuoles often move along microtubules. The endocytic network that the viruses enter after internalization is composed of a large variety of interacting membrane compartments many of which are acidic. The low pH and other cues trigger the penetration in different organelles depending on the virus and its properties. As shown by the thick gray arrows, penetration can happen in early, maturing, and late endosomes as well as in the ER and possibly the Golgi complex. Viruses that use macropinocytosis – an actin-dependent transient process that involves membrane ruffling and fluid uptake into large vacuoles – penetrate from macropinosomes. Micropinocytosis represents an actin-dependent, dynamin-independent endocytic mechanism, which is still incompletely characterized. The names of some of the viruses using different pathways are indicated. Some viruses such as influenza A virus can exploit multiple pathways. Dotted arrows indicate vesicle maturation.

sized viruses (up to about 150 nM in diameter) macro-endocytosis by larger ones (Chigo, 2010; Mercer *et al.*, 2010). Some of the pathways discussed in the endocytosis chapters such as the CLIC/GEEC pathway do not seem to be used by

viruses. Most endocytic events are activated by viruses themselves either by triggering signals or by inducing plasma membrane indentation. After binding to its receptor, the GM1 ganglioside, SV40 has been shown to induce the formation of

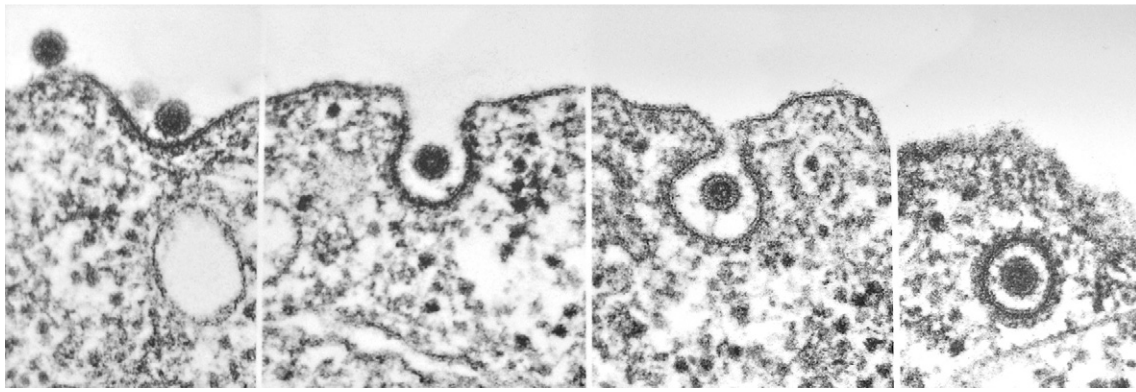


Figure 5 Alphaviruses are internalized by clathrin-mediated endocytosis. Semliki Forest virus, a member of the alphavirus family, is a small, enveloped, arthropod-borne RNA virus that enters host cells by clathrin-mediated endocytosis. After binding to the plasma membrane, it induces the assembly of a clathrin coat on the cytosolic surface of the plasma membrane. After about 1 min, the clathrin-coated pit is completed. It undergoes membrane fission, and the virus is internalized in a coated vesicle. Different stages of this process are shown in electron micrographs taken by Jürgen Kartenbeck and Ari Helenius.

a deep, tight-fitting pit in the plasma membrane, as if budding into the cell (Ewers *et al.*, 2010). Cellular factors are responsible for membrane fission resulting in a vesicle for transport to endosomes. This process is interesting because the cargo virus is here inducing membrane curvature by a zipper-like process when it binds to multiple GM1 gangliosides.

After endocytosis, many viruses are first delivered to early endosomes (EE) (Helenius, 2013). If penetration does not happen in the EEs, they generally move within maturing endosomes and late endosomes (LE) toward the perinuclear region to be delivered to endolysosomes (Figure 4). Viruses that enter via macropinocytosis or related pathways bypass the EEs and join the classical endosome pathway at the level of LE or endolysosomes. It has been proposed that some viruses enter recycling endosomes, autophagosomes, and the Golgi complex before penetration.

Late-penetrating viruses, such as IAV and bunyaviruses, are dependent on endosome maturation for infection. This means that their entry can be inhibited by perturbations that affect microtubule-mediated vesicle traffic, Rab conversion, formation of intraluminal vesicles, etc. Polyomaviruses travel from LEs to the smooth ER where penetration through the ER membrane into the cytosol occurs with the help of ER chaperones and components of the quality control and ERAD pathways (Inoue and Tsai, 2013).

Penetration

Where exactly the critical penetration steps occur often depends on pH, because exposure to low pH triggers changes in many viruses that make them penetration-competent (White *et al.*, 1980, 1982). Depending on the pH-threshold for such activation, the penetration reaction may occur in EEs (pH 6.5–6.0), in LEs or endolysosomes (pH 6.0–5.0), or in macropinosomes (Figures 4 and 6). However, there are other intracellular triggers such as proteolytic processing of viral proteins, exposure to specific vacuolar proteins, assistance by cytosolic factors, etc. For enveloped viruses, large conformational changes in the fusion glycoproteins of the viral envelope lead to hydrophobic attachment to the endosomal target

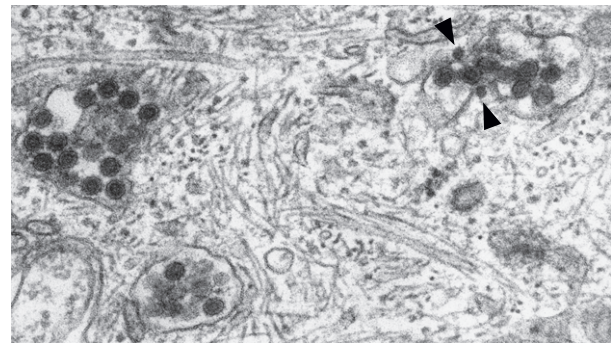


Figure 6 Incoming alphaviruses in endosomes. The low pH in early endosomes induces a dramatic, irreversible conformational change in the spike glycoproteins triggering their membrane fusion activity. The viruses fuse from the inside with the limiting membrane of the endosome, and the icosahedral capsids (see arrow heads) are released into the cytosol. Moments later the capsids are uncoated and the RNA begins to be translated. The image is by Jürgen Kartenbeck and Ari Helenius.

membrane, followed by fusion of the viral envelope (Kielian and Rey, 2006; Harrison, 2013). The fusion proteins have been analyzed in great detail. They belong to three different general classes. As a result of membrane fusion, the capsids are transferred into the cytosol.

While the penetration of non-enveloped viruses is less well understood, it is evident that there are several options. The viruses can escape by lysing endosomal vacuoles, they can induce the formation of pores through which the virus itself or just the genome can pass into the cytosol, or – as in the case of SV40 – they can make use of the ERAD pathway in the ER (Suomalainen and Greber, 2013; Cerqueira and Schelhaas, 2012).

Intracellular Transport and Nuclear Import

Replication of viruses occurs either in the nucleus or in the cytoplasm. In the cytoplasm, it is usually associated with membranes, granules, and complex virus-induced structures

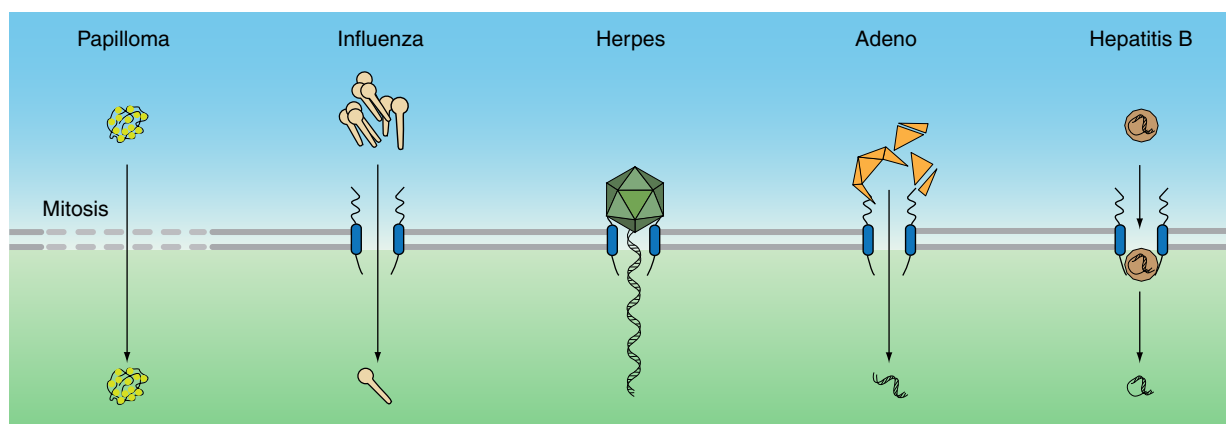


Figure 7 Entry into the nucleus. Viruses that replicate in the nucleus must transport their genome and sometimes accessory proteins across the nuclear membrane. Some of the mechanism and viruses that use them are shown. Subviral complexes of papillomaviruses wait in the cytosol until cell division breaks down the nuclear membrane barrier. The ribonucleoprotein complexes (vRNP) of influenza virus are imported through nuclear pores. Herpesvirus capsids dock at the cytosolic side of the nuclear pore and inject their genome into the nucleus. Adenovirus capsids are disassembled on the cytosolic side of the nuclear pore, and the genome is released into the nucleus. Hepatitis B capsids are small enough to be imported through to the nuclear pore complex. Uncoating of the viral DNA occurs on the inside in the nuclear basket.

serving as virus factories. To reach the sites of replication, intracellular transport of incoming viruses and capsids is required. Viruses often make use of microtubules and microtubule-dependent motors (Figure 2). For example, the transport of herpesvirus capsids from their site of penetration in the axons to the nuclear pores of neurons involves such minus end-directed movement. Specific viral proteins attach to dynactin, which mediates attachment to the dynein motor (Radtke *et al.*, 2010; Smith, 2012). Adenoviruses use a similar strategy (Kelkar *et al.*, 2004). Viruses also bind to kinesins that serve as plus end-directed motors.

Viruses that enter the nucleus must be able to overcome another barrier, the nuclear envelope (Figure 7). Most of the viruses have nuclear import signals and make use of the cell's nuclear import machinery. If small enough in diameter (< 40 nm), they can use this machinery to enter through the nuclear pore complex and uncoat inside the nucleus (polyomaviruses and hepatitis B virus) (Rabe *et al.*, 2003). They can also release the genome through the NPC without entering the nucleus (herpes- and adenoviruses) (Sodeik *et al.*, 1997; Greber *et al.*, 1997). In influenza virus, the RNA genome is present in short pieces individually packaged in particles small enough to enter through the nuclear pore complex (Cros and Palese, 2003).

Viruses can also cause local lysis of the nuclear envelope (parvoviruses) (Cohen *et al.*, 2011), or they can wait in the cytosol until mitosis occurs and the nuclear envelope is temporarily disassembled (papilloma- and retroviruses) (Aydin *et al.*, 2014; Suzuki and Craigie, 2007).

Replication and Synthesis of Viral mRNAs

The molecular biology of virus replication and mRNA production depends on the type of genome (DNA or RNA, single-stranded or double-stranded, etc.). A lot of information is available for many different viruses and virus families. Here, a

short review of some cell biologically relevant aspects of these important processes will have to suffice.

DNA and RNA viruses have a different set of options. DNA viruses can make use of the cell's DNA synthesis and transcription machinery as well as accessory functions such as mRNA splicing and processing in the nucleus. Therefore, with few exceptions, the replication of DNA viruses occurs in the nucleus.

Although profiting from cellular services in the nucleus, these viruses have to deal with the added problem of import/export across the nuclear envelope. This applies not only to incoming virus particles but also to viral protein components needed for capsid assembly. These are synthesized in the cytosol, then transported into the nucleus for virus assembly. The proteins in question therefore have nuclear localization signals and make use of the nuclear import machinery of the cell. To exit from the nucleus, non-enveloped viruses usually wait for the cell to die and the nuclear envelope and the plasma membrane to lyse.

The newly assembled capsids of herpes viruses escape into the cytosol by a complicated two-step membrane fission/fusion strategy that allows the icosahedral capsids assembled in the nucleus to cross the two membranes of the nuclear envelope (Figure 8). By undergoing a second budding event into the Golgi complex, they acquire their final membrane envelope before being secreted into the extracellular space (Mettenleiter, 2004).

Influenza virus belongs to the few RNA viruses that replicate in the nucleus. Individual viral ribonucleoproteins assembled in the nucleus are escorted through nuclear pore complexes by viral and cellular proteins to the cytosol and transported to the plasma membrane where they join up with viral glycoproteins and bud out (Cros and Palese, 2003).

Since cells do not have RNA-dependent RNA polymerases, RNA viruses must provide their own enzymes for genome replication and mRNA synthesis. With a few exceptions, they have evolved to replicate in the cytoplasm. If the RNA genome is single-stranded and has a positive sense, it can be used

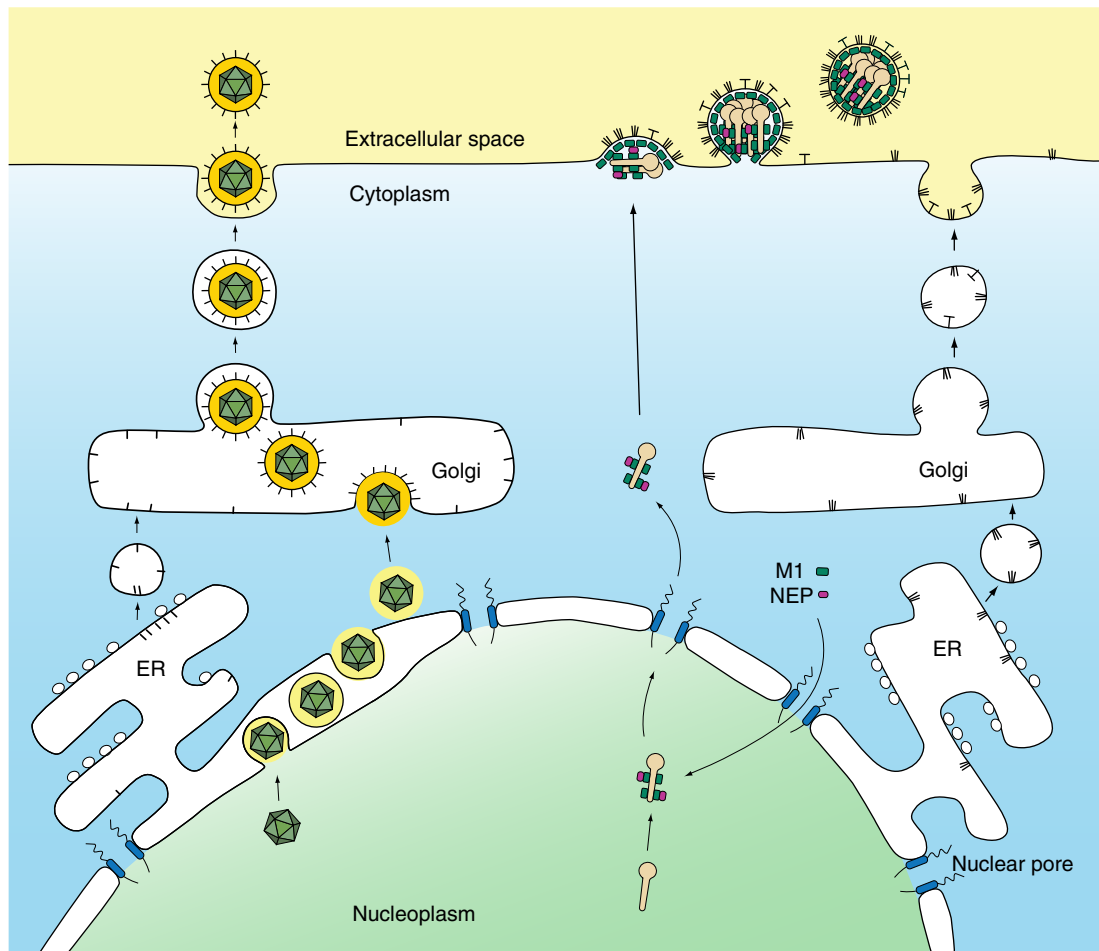


Figure 8 Nuclear exit and maturation of herpes simplex virus 1 and influenza A virus. Left: After capsid assembly and DNA encapsidation in the nucleus, the icosahedral HSV nucleocapsids (green) together with some tegument proteins (yellow) bud into the inner nuclear membrane in a process called primary envelopment, followed by fusion with the outer nuclear membrane. Next the nucleocapsids associate with additional tegument proteins (orange) and bud into the Golgi-endosome compartment (secondary envelopment) acquiring a new envelope membrane and mature glycoproteins derived from the endoplasmic reticulum (ER). The virion is released from the cell by exocytosis. Right: Newly assembled vRNPs of influenza virus bind to two soluble proteins, M1 and NEP, and exit to the cytosol through the nuclear pore complex. In the cytosol, the vRNP traffic along microtubules to the plasma membrane where they are encapsidated in a shell of M1 protein. A budding reaction follows because the assembled cores interact with the ER-derived viral glycoproteins that have been transported to the cell surface.

directly as a messenger RNA to translate proteins. The required RNA polymerase needed to replicate the viral RNA and generate additional messenger RNAs for viral proteins can now be synthesized with the incoming RNA as messenger. If the RNA genome has negative sense – as often is the case among animal RNA viruses – the required enzymes must be brought along as a structural component of the incoming virus.

Replication and mRNA synthesis usually occur in close contact with cytoplasmic membranes where the process is harder to detect by cellular defense systems that detect double-stranded RNA. The cell biological consequence is that virus infection causes major rearrangements in membranes of infected cells, leading to so-called cytopathic effects. Some RNA viruses induce a shut-off of host protein synthesis, others enhance the formation of stress granules and P-bodies, while many induce apoptosis and cell death (Onomoto *et al.*, 2014; Clemens, 2005). The cell biological consequences of infection are often dramatic. Virus infection does not, however, always have to

result in cell death, because there are many viruses that induce latent and persistent infections (Speck and Ganem, 2010).

Virus Assembly and Release

The components needed for assembly of new particles include, in addition to the viral genome, the protein components of the capsid and envelope, enzymes such as polymerases, and the lipid bilayer in the case of enveloped viruses (Perlmutter and Hagan, 2014). Host cell proteins are also sometimes included such as histones, ubiquitin, molecular chaperones, or even ribosomes. Through the glycans present in envelope glycoproteins and the lipids, the virus particles possess an ‘imprint’ of the host that they are derived from. This has consequences, for example, when dendritic cells in the skin recognize insect-derived virus particles as ‘foreign’ by the N-linked glycans that the glycoproteins carry (Freer and Matteucci, 2009).

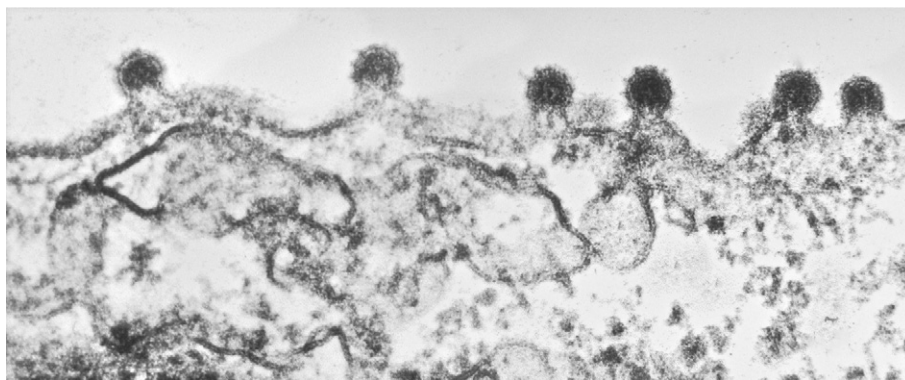


Figure 9 Virus budding at the plasma membrane. Like other alphaviruses and many other enveloped viruses, Semliki Forest virus buds at the plasma membrane of infected cells. The icosahedral capsid that contains the viral RNA genome is assembled in the cytosol. It is then transported to the plasma membrane where it interacts with the cytosolic domains of the glycoproteins. These interactions drive the progressive generation of membrane curvature around the capsid and eventually the membrane fission event that releases the mature particle. The image is by Jürgen Kartenbeck and Ari Helenius.

As already mentioned, the components of the virus are usually synthesized and modified in different compartments of the cell: lipids in the ER and Golgi, envelope proteins in the ER, nucleic acids in the nucleus or cytosol, and capsid proteins in the cytosol. How the components are brought together locally and how they assemble to generate virus particles is a complex process that varies between virus families (Perlmutter and Hagan, 2014). Like entry and uncoating, the assembly process relies on both cellular and viral factors. The formation of icosahedral capsids depends mainly on the properties of the capsid proteins, and how they can assemble into pentons and hexons.

The selective inclusion of viral RNA and DNA in the particle is typically determined by a sequence element (a so-called packaging sequence) in the viral genome recognized by one or more of the capsid proteins. The viral genome may be introduced into an almost completed capsid through one of the vertices, or it may have to be present before capsid assembly can even start. In the case of viruses with a negative sense genome, capsid proteins bind to the RNA already when it is being synthesized by the viral polymerase in the cytosol or the nucleus. The genomic RNA of these viruses therefore never exists in a completely uncoated state.

The viral glycoproteins are synthesized, modified, and folded in the ER as integral membrane proteins. Most of them are type I integral transmembrane proteins with large oligomeric ectodomains. Like cellular proteins, they undergo glycosylation, glycan trimming, disulfide oxidation, oligomerization, chaperone interaction, and quality control in this compartment. They then either stay in the ER (flaviviruses) or move along the secretory pathway to the Golgi complex (bunya- and herpesviruses) (Figure 8), the plasma membrane (most other enveloped viruses), or possibly the endosome membrane. For transport and targeting they possess the appropriate sequence signals and undergo similar posttranslational modifications as cellular membrane proteins. Influenza hemagglutinin and HIV-1 glycoprotein are often used as model proteins for analyzing the biogenesis, folding, quality control, and vesicular transport of glycoproteins.

With the exception of pox viruses, the final assembly of capsids and envelopes involves a budding reaction. The envelope proteins, capsids, and in some cases matrix proteins

with adaptor functions establish a network of interactions with the exclusion of most host proteins, and form an outward oriented bud in the membrane (Figures 8 and 9). When the membrane in the bud closes and undergoes fission, a fully membrane-enclosed particle is formed; this is the enveloped virus. For some enveloped viruses such as HIV-1, the fission reaction depends on components of the ESCRT complex, which localize to the bud neck (Martin-Serrano *et al.*, 2001). After budding, the virus particle may undergo some further maturation steps. A network of interactions between viral components makes the virus resistant enough to withstand stresses in the extracellular environment. However, the design of the particle is such that it provides for metastability, and thus the ability to uncoat and release the genome when the virus enters a new host cell. The envelope serves as an extracellular transport vesicle poised to fuse with a membrane in the new host cells, as already discussed.

While the lipid composition of the viral envelope membrane reflects the cellular membrane where budding occurred, it may be enriched in certain lipids such as those present in lipid rafts. The matrix proteins located between the envelope and the capsid often serves as important adaptors in the budding process connecting the membrane proteins with the capsid. If budding occurs at the membranes of cytoplasmic organelles (herpes-, flavi-, retro-, and bunyaviruses) (Garoff *et al.*, 1998), the particles formed are secreted into the extracellular space via the secretory pathway (Figure 8).

See also: Organizational Cell Biology: An Overview

References

- Aydin, I., Weber, S., Snijder, B., *et al.*, 2014. Large scale RNAi reveals the requirement of nuclear envelope breakdown for nuclear import of human papillomaviruses. *PLoS Pathogens* 10, e1004162.
- Burckhardt, C.J., Greber, U.F., 2009. Virus movements on the plasma membrane support infection and transmission between cells. *PLoS Pathogens* 5, e1000621.

- Cerrequeira, C., Schelhaas, M., 2012. Principles of polyoma- and papillomavirus uncoating. *Medical Microbiology and Immunology* 201, 427–436.
- Clemens, M.J., 2005. Translational control in virus-infected cells: Models for cellular stress responses. *Seminars in Cell & Developmental Biology* 16, 13–20.
- Cohen, S., Marr, A.K., Garcin, P., Pante, N., 2011. Nuclear envelope disruption involving host caspases plays a role in the parvovirus replication cycle. *Journal of Virology* 85, 4863–4874.
- Cros, J.F., Palese, P., 2003. Trafficking of viral genomic RNA into and out of the nucleus: influenza, Thogoto and Borna disease viruses. *Virus Research* 95, 3–12.
- Cureton, D.K., Massol, R.H., Whelan, S.P., Kirchhausen, T., 2010. The length of vesicular stomatitis virus particles dictates a need for actin assembly during clathrin-dependent endocytosis. *PLoS Pathogens* 6, e1001127.
- Eisenberg, R.J., Atanasiu, D., Cairns, T.M., *et al.*, 2012. Herpes virus fusion and entry: A story with many characters. *Viruses* 4, 800–832.
- Ewers, H., Romer, W., Smith, A.E., *et al.*, 2010. GM1 structure determines SV40-induced membrane invagination and infection. *Nature Cell Biology* 12, 11–18; Suppl. pp. 1–12.
- Freer, G., Matteucci, D., 2009. Influence of dendritic cells on viral pathogenicity. *PLoS Pathogens* 5, e1000384.
- Garoff, H., Hewson, R., Opstelten, D.J., 1998. Virus maturation by budding. *Microbiology and Molecular Biology Reviews* 62, 1171–1190.
- Ghigo, E., 2010. A dilemma for viruses and giant viruses: Which endocytic pathway to use to enter cells? *Intervirology* 53, 274–283.
- Greber, U., Suomalainen, M., Stidwill, R., *et al.*, 1997. The role of the nuclear pore complex in adenovirus DNA entry. *EMBO Journal* 16, 5998–6007.
- Grove, J., Marsh, M., 2011. The cell biology of receptor-mediated virus entry. *Journal of Cell Biology* 195, 1071–1082.
- Harrison, S.C., 2013. Principles of virus structure. In: Knipe, D.M., Howley, P.M. (Eds.), *Fields Virology*, sixth ed. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins, pp. 52–86.
- Helenius, A., 2013. Virus entry and uncoating. In: Knipe, D.M., Howley, P.M. (Eds.), *Fields Virology*, sixth ed. Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins, pp. 87–104.
- Inoue, T., Tsai, B., 2013. How viruses use the endoplasmic reticulum for entry, replication, and assembly. *Cold Spring Harbor Perspectives in Biology* 5, a013250.
- Jardetzky, T.S., Lamb, R.A., 2014. Activation of paramyxovirus membrane fusion and virus entry. *Current Opinion in Virology* 5, 24–33.
- Kelkar, S.A., Pfister, K.K., Crystal, R.G., Leopold, P.L., 2004. Cytoplasmic dynein mediates adenovirus binding to microtubules. *Journal of Virology* 78, 10122–10132.
- Kielian, M., Rey, F.A., 2006. Virus membrane-fusion proteins: More than one way to make a hairpin. *Nature Reviews Microbiology* 4, 67–76.
- Klasse, P.J., 2012. The molecular basis of HIV entry. *Cellular Microbiology* 14, 1183–1192.
- Lehmann, M.J., Sherer, N.M., Marks, C.B., Pypaert, M., Mothes, W., 2005. Actin- and myosin-driven movement of viruses along filopodia precedes their entry into cells. *Journal of Cell Biology* 170, 317–325.
- Martin-Serrano, J., Zang, T., Bieniasz, P.D., 2001. HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to sites of particle assembly to facilitate egress. *Nature Medicine* 7, 1313–1319.
- Mercer, J., Helenius, A., 2008. Vaccinia virus uses macropinocytosis and apoptotic mimicry to enter host cells. *Science* 320, 531–535.
- Mercer, J., Helenius, A., 2009. Virus entry by macropinocytosis. *Nature Cell Biology* 11, 510–520.
- Mercer, J., Helenius, A., 2012. Gulping rather than sipping: Macropinocytosis as a way of virus entry. *Current Opinion in Microbiology* 15, 490–499.
- Mercer, J., Schelhaas, M., Helenius, A., 2010. Virus entry by endocytosis. *Annual Review of Biochemistry* 79, 803–833.
- Mettenleiter, T.C., 2004. Budding events in herpesvirus morphogenesis. *Virus Research* 106, 167–180.
- Moller-Tank, S., Kondratowicz, A.S., Davey, R.A., Rennert, P.D., Maury, W., 2013. Role of the phosphatidylserine receptor TIM-1 in enveloped-virus entry. *Journal of Virology* 87, 8327–8341.
- Onomoto, K., Yoneyama, M., Fung, G., Kato, H., Fujita, T., 2014. Antiviral innate immunity and stress granule responses. *Trends in Immunology* 35, 420–428.
- Perlmutter, J.D., Hagan, M.F., 2014. Mechanisms of virus assembly. *Annual Review of Physical Chemistry* 66, 217–239.
- Qian, M., Cai, D., Verhey, K.J., Tsai, B., 2009. A lipid receptor sorts polyomavirus from the endolysosome to the endoplasmic reticulum to cause infection. *PLoS Pathogens* 5, e1000465.
- Rabe, B., Vlachou, A., Pante, N., Helenius, A., Kann, M., 2003. Nuclear import of hepatitis B virus capsids and release of the viral genome. *Proceedings of the National Academy of Sciences of the United States of America* 100, 9849–9854.
- Radtko, K., Kienke, D., Wolfstein, A., *et al.*, 2010. Plus- and minus-end directed microtubule motors bind simultaneously to herpes simplex virus capsids using different inner tegument structures. *PLoS Pathogens* 6, e1000991.
- Schelhaas, M., Ewers, H., Rajamaki, M.L., *et al.*, 2008. Human papillomavirus type 16 entry: Retrograde cell surface transport along actin-rich protrusions. *PLoS Pathogens* 4, e1000148.
- Smith, G., 2012. Herpesvirus transport to the nervous system and back again. *Annual Review of Microbiology* 66, 153–176.
- Sodeik, B., Ebersold, M.W., Helenius, A., 1997. Microtubule-mediated transport of incoming herpes simplex virus 1 capsids to the nucleus. *Journal of Cell Biology* 136, 1007–1021.
- Speck, S.H., Ganem, D., 2010. Viral latency and its regulation: Lessons from the gamma-herpesviruses. *Cell Host & Microbe* 8, 100–115.
- Suomalainen, M., Greber, U.F., 2013. Uncoating of non-enveloped viruses. *Current Opinion in Virology* 3, 27–33.
- Suzuki, Y., Craigie, R., 2007. The road to chromatin – nuclear entry of retroviruses. *Nature Reviews Microbiology* 5, 187–196.
- Szklarczyk, O.M., Gonzalez-Segredo, N., Kukura, P., *et al.*, 2013. Receptor concentration and diffusivity control multivalent binding of Sv40 to membrane bilayers. *PLOS Computational Biology* 9, e1003310.
- Tsai, B., Qian, M., 2010. Cellular entry of polyomaviruses. *Current Topics in Microbiology and Immunology* 343, 177–194.
- Whelan, S., 2013. Viral replication strategies. In: Knipe, D.M., Howley, P.M. (Eds.), *Fields Virology*, sixth ed. Philadelphia, PA: Wolters Kluwer, pp. 105–126.
- White, J., Kartenbeck, J., Helenius, A., 1980. Fusion of Semliki forest virus with the plasma membrane can be induced by low pH. *Journal of Cell Biology* 87, 264–272.
- White, J., Kartenbeck, J., Helenius, A., 1982. Membrane fusion activity of influenza virus. *EMBO Journal* 1, 217–222.
- Yamauchi, Y., Helenius, A., 2013. Virus entry at a glance. *Journal of Cell Science* 126, 1289–1295.

Relevant Websites

- https://www.google.ch/search?q=virus+entry&ie=utf-8&oe=utf-8&aq=t&rls=org.mozilla:en-US:official&client=firefox-a&gfe_rd=cr&ei=94NWU5_HCeHl8gfgpYDIDg
- <http://www.youtube.com/watch?v=Y8h2wzwdzZs>
HCV Life Cycle.
- <http://www.youtube.com/watch?v=odRyv7V8LAE>
HIV-1 Life Cycle.
- <http://www.ibiology.org/ibioseminars/microbiology/ari-helenius-part-1.html>
iBiology.
- http://www.youtube.com/watch?v=7_E8CXVeOUo
Racaniello Virus Assembly.
- <http://www.youtube.com/watch?v=UblJs2vT7oA>
Racaniello Virus Entry.
- <http://visualscience.ru/en/projects/influenza/illustration/>
Visual Science.

Virus Factories and Mini-Organelles Generated for Virus Replication

T Wileman, University of East Anglia, Norwich, UK

CL Netherton, Institute for Animal Health, Pirbright, Woking, UK

PP Powell, University of East Anglia, Norwich, UK

© 2016 Elsevier Inc. All rights reserved.

Introduction

Viruses are relatively simple life forms which use a protein capsid to carry short stretches of genomic material from one cell to another. All viruses are obligate intracellular parasites and need to enter cells to replicate. Nucleic acid genomes delivered to the cytoplasm provide a template to generate new genomes which in turn initiate the synthesis of mRNA that encode viral proteins. These processes require the assembly of complex replication machines that ensure that genome replication and packaging are carefully coordinated. For many viruses this is achieved by concentrating virus genomes, replicase proteins, and the host proteins required for replication in a 'mini-organelle' called a virus factory or 'viroplasm' (Netherton *et al.*, 2007; Fernandez de Castro *et al.*, 2012).

All viruses synthesize mRNA and use host ribosomes to make viral proteins. Viruses are grouped into the 'RNA' or 'DNA' viruses according to the nature of their genome. The (+) strand genomes make an mRNA that is immediately translatable, and the complementary RNA or DNA sequence is called negative (−) strand. This means that the genomes of (+) strand RNA viruses can function as mRNA while the (−) strand RNA viruses need to be delivered into the cell with an RNA-dependent RNA polymerase (RdRp) that uses the negative template to make the (+) strand. In both cases virus replication occurs in association with cytoplasmic organelles and (+) RNA is translated by ribosomes into protein. The genomes of most DNA viruses are made from double-stranded DNA and most DNA viruses replicate in the nucleus because they need to recruit host proteins for DNA replication. The DNA viruses also show temporal gene expression with genes being expressed at early, intermediate, and late times during infection. Early genes are expressed before the onset of replication while late genes are normally used for replication and assembly. The small DNA viruses, for example, polyoma and papillomaviruses have limited coding capacity and do not make their own DNA-dependent DNA polymerase (DdDp). The large DNA viruses, for example herpes viruses and poxviruses encode their own DdDp, but the herpes viruses still rely on nuclear factors for genome replication. Interestingly, the poxviruses encode all the enzymes needed for DNA and mRNA synthesis, and replicate in the cytoplasm.

Evolution has not been blind to viral infection and cells make several proteins that recognize viral genomes as danger signals because they are different from host DNA or RNA. The RNA viruses are a good example because replication of RNA from an RNA (rather than DNA) template generates a double-stranded RNA (dsRNA) intermediate that is not generated by the host cell. Recognition of dsRNA in plants, insects, and other primitive animals results in gene silencing while in vertebrates this leads to activation of interferon (Randall and Goodbourn, 2008). Cells also contain sensors for viral DNA that activate

proinflammatory responses (Rathinam and Fitzgerald, 2011). The concentration of replicase proteins within virus factories would be expected to provide a potent signal for activation of these defenses. Viruses have overcome this problem by using virus factories to shield replication intermediates from recognition by nucleic acid sensors, but at the same time the factories retain connections to the cytoplasm to allow export of nucleic acid to sites of genome packaging.

RNA Viruses

(+) Strand RNA Viruses Generate Mini-Organelles from Membranes Derived from the Secretory Pathway

The genomes of the (+) strand RNA viruses encode a single polyprotein synthesized by cellular ribosomes. The capsid proteins and the nonstructural proteins required for replication and are generated from the polyprotein by viral, and sometimes host, proteases. In each case nonstructural proteins with hydrophobic domains act as a scaffold to target the RdRp to the cytoplasmic face of membrane-bound organelles where assembly of the replicase complex induces membrane curvature (reviewed Cottam *et al.*, 2009; den Boon *et al.*, 2010; den Boon and Ahlquist, 2010). These remodeled membranes vary greatly between different families of (+) strand RNA viruses but are thought to be formed by similar mechanisms where their dimensions are determined by the assembly of the replicase proteins during invagination. This process is analogous to the way that virus capsids vary in size and shape depending on the assembly of capsid subunits (Ahlquist, 2006).

Spherules produced by the *Togaviridae* and *nodaviruses*

The mammalian alphaviruses, for example, Semliki Forest virus (SFV), Sindbis virus (SINV), and the rubivirus rubella virus are members of the 'Togaviridae' and encode a single polyprotein called nsP1234 that contains the helicase, methyltransferase, and RdRp. Proteolytic processing releases the RdRp (P4) which binds to P1,2,3 and is targeted to membranes by a hydrophobic domain in P1 (Figure 1(a)). The structural proteins are generated as a second polyprotein encoded by a sub-genomic message which is formed when the RdRp transcribes RNA from an internal initiation site. A signal sequence exposed after removal of the capsid protein directs synthesis of envelope proteins by ribosomes attached to the endoplasmic reticulum (ER). The 'Togaviridae' generate large 'cytopathic vacuoles' with 50 nm diameter invaginations called spherules aligned along the inside face of the vacuole. The vacuoles contain replicase proteins and are derived from endosomes and lysosomes. Spherules can be generated from many different membrane-bound organelles. Flock house virus, an insect virus, generates spherules in the outer

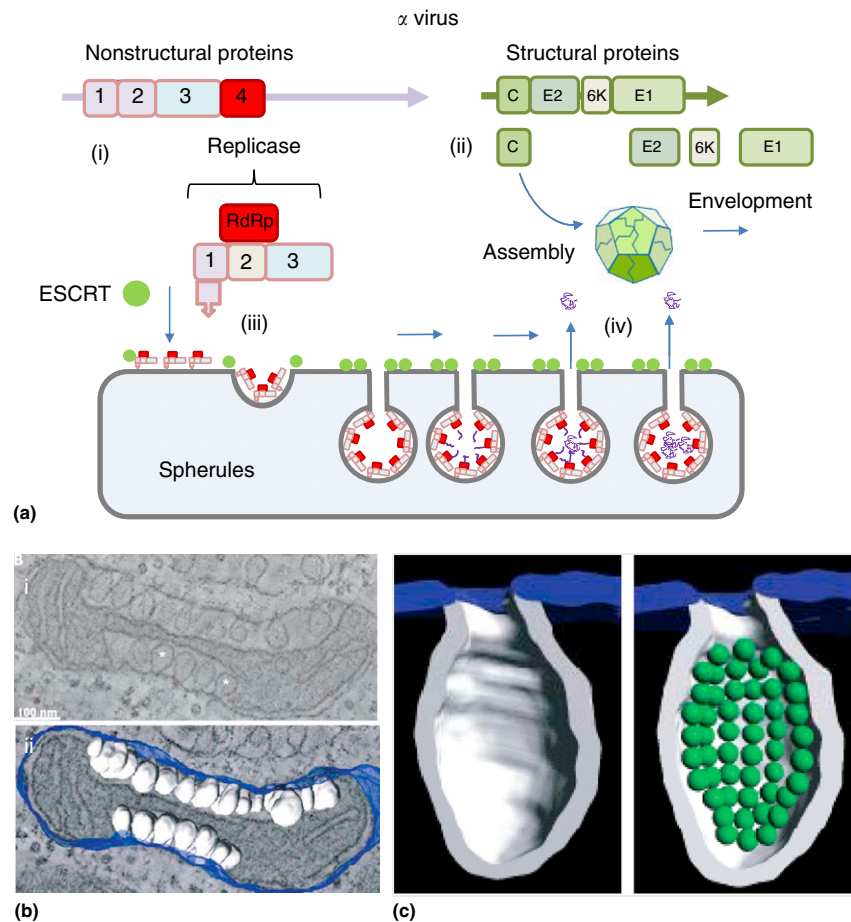


Figure 1 Formation of spherules during replication of alphaviruses. (a) The alphaviruses encode a single polyprotein called nsp1234 that contains the helicase, methyltransferase, and RdRp. Proteolytic processing releases the RdRp (P4) which is targeted to membranes by a hydrophobic domain in P1 (i). The structural proteins (ii) synthesized as a second polyprotein (C,E2,6K,E1) are directed by ribosomes to the endoplasmic reticulum. Assembly of replicase proteins induces membrane curvature and invagination forming a spherule (iii). The interior of each spherule is connected to the cytosol by a thin membrane neck that surrounds a channel wide enough to allow passage of (+) RNA into the cytosol allowing assembly with capsid proteins (iv). In the case of Bromovirus and tombusvirus, ESCRT proteins (green) recruited to the spherule by replicase proteins stabilize the membrane neck and prevent closure of the spherule. (b) Spherules generated in mitochondria by flock house virus. Electron micrographs (i) generated from a tilt series were processed digitally to generate the tomogram in (ii). (c) A 3D reconstruction of spherules shows the likely organization of replicase proteins of flock house virus modeled as 7 nm diameter spheres (green) lining the inner surface of the spherule. Reproduced from Kopek, B.G., Perkins, G., Miller, D.J., Ellisman, M.H., Ahlquist, P., 2007. Three-dimensional analysis of a viral RNA replication complex reveals a virus-induced mini-organelle. *PLoS Biology* 5, e220.

membrane of mitochondria (Figures 1(b) and 1(c); Miller *et al.*, 2001) while Brome mosaic virus, a plant virus, generates spherules in the ER. In each case the spherules are of similar size and contain approximately 100 copies of the replicase protein packed along the inner membrane surface (Kopek *et al.*, 2007). The interior of each spherule is connected to the cytosol by a thin membrane neck that surrounds a channel wide enough to allow passage of (+) RNA into the cytosol allowing assembly with capsid proteins. Spherule formation involves the assembly of replicase proteins on the cytoplasmic face of membrane-bound organelles where they induce negative curvature forcing membrane to invaginate into the organelle. Multivesicular body formation is topologically similar where ESCRT (endosomal sorting complex required for transport) proteins assemble on the cytoplasmic face of endosomes generating vesicles that bud away from the

cytoplasm into the endosome. Recent work has shown that proteins of the ESCRT-III complex are required for proper formation of spherules where they may drive spherule closure and at the same time maintain the channel that allow genomes to pass into the cytosol (Barajas *et al.*, 2014; Diaz *et al.*, 2015).

Convolved membranes and membrane webs generated by coronaviruses and flaviviruses

The flaviviruses, for example, yellow fever virus, West Nile virus, and hepatitis C (HCV) also generate a single polyprotein that, in common with picornaviruses, is processed to generate structural and nonstructural proteins (Figure 2(a)). The flaviviruses are enveloped viruses and the initial processing of the polyprotein exposes signal sequences that direct synthesis and processing of the polyprotein by the ER. The NS5B protein at

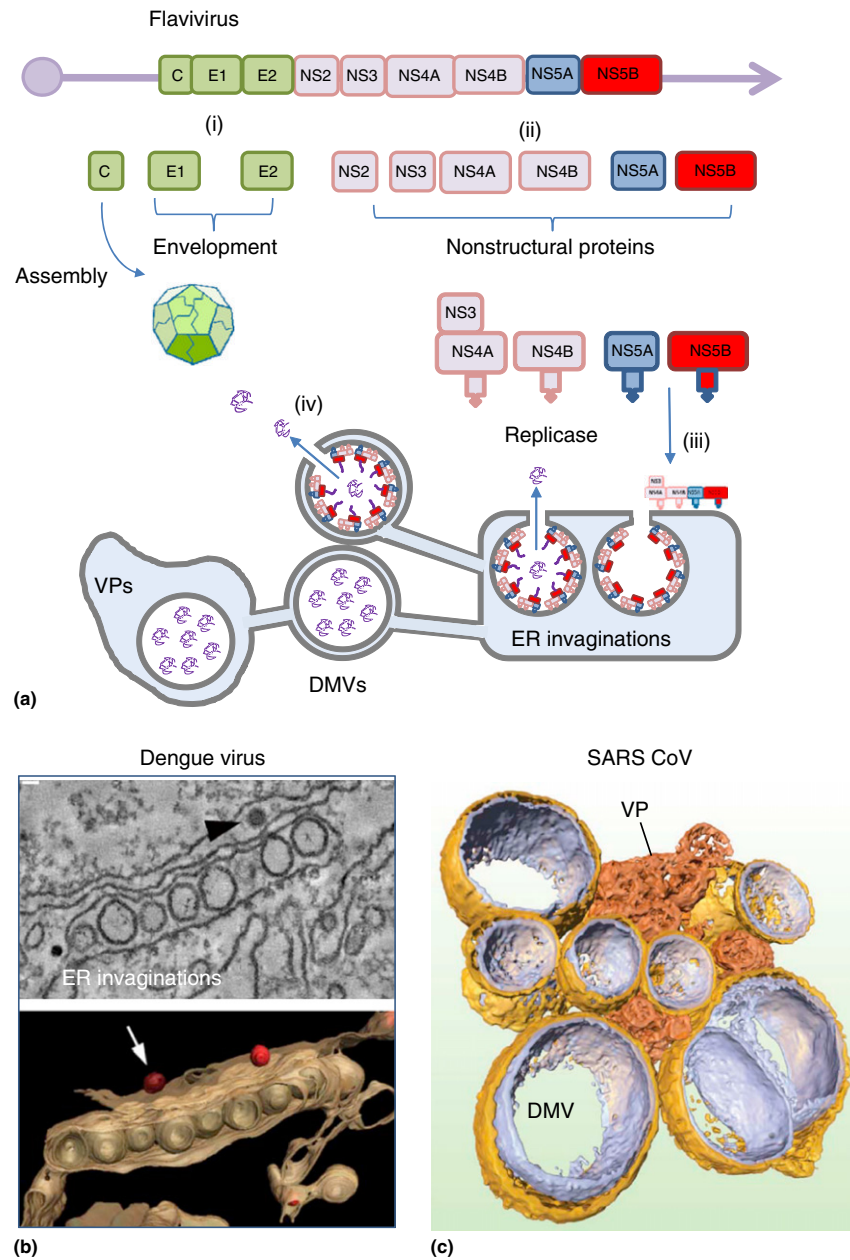


Figure 2 Formation of double-membraned vesicles by flaviviruses and coronaviruses. (a) The flaviviruses polyprotein encodes the capsid protein (C), envelope proteins (E1, E2) and the nonstructural proteins (NS2–NS5) required for replication. (i) Removal of the capsid protein by the NS2 protease exposes a signal peptide that directs synthesis of envelope proteins by the ER. (ii) Hydrophobic domains in NS4A, NS4B, NS5A and NS5B direct posttranslational insertion into the cytoplasmic face of the ER. (iii) NS5B is the RdRp. Assembly of replicase proteins induces ER invaginations and double membrane vesicles (DMV) and membraneous webs, also known as vesicular packets (VP). (b) Endoplasmic reticulum (ER) invaginations generated during dengue virus infection. Electron micrographs generated from a tilt series (top) were processed digitally to generate the tomogram (bottom). The 3D reconstruction shows the likely organization of ER invaginations. The arrow indicates viruses that may receive genomes from ER invaginations during assembly and envelopment. Reproduced from Welsch, S., Miller, S., Romero-Brey, I., *et al.*, 2009. Composition and three-dimensional architecture of the dengue virus replication and assembly sites. *Cell Host & Microbe* 5, 365–375. (c) The coronavirus polyprotein has the same overall organization as the flaviviruses but encodes as many as 15 nonstructural proteins. The figure shows double membrane vesicles and vesicular packets generated from the ER during SARS coronavirus infection. Electron micrographs generated from a tilt series were processed digitally to generate the tomogram. The 3D reconstruction shows the likely organization of interconnected DMVs (gold and silver) and vesicular packets (orange). Reproduced from Knoops, K., Kikkert, M., Worm, S.H., *et al.*, 2008. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biology* 6, e226.

the C-terminus of the polyprotein is targeted to the cytoplasmic face of the ER by hydrophobic domains in NS4A, NS5A, and NS5B. The flaviviruses generate 80–100 nm diameter invaginations into an ER membrane connected to spherical vesicles and convoluted membranes (Figure 2(b)). These invaginations are less well organized than spherules, even so the spherical vesicles contain the viral polymerase and dsRNA and each has a pore opening to the cytosol that could allow exit of viral RNA for packaging with capsid proteins and envelopment (Welsch *et al.*, 2009; Romero-Brey *et al.*, 2012). The genome of the coronaviruses (CoV) is similar in organization to the flaviviruses, but these viruses are the most complex of the (+) strand RNA viruses and encode as many as 15 non-structural proteins. Severe acute respiratory syndrome coronavirus (SARS-CoV) induces double membrane vesicles (DMVs) that are between 150 and 300 nm diameter connected to the ER (Figure 2(c); Knoops *et al.*, 2008). The infected cells also contain convoluted tubular and reticular membranes, and later during infection ‘vesicular packets’ appear where vesicles are surrounded by a common outer ER membrane (Knoops *et al.*, 2008). It is thought that replication occurs in the convoluted membranes and that genomes are transported to vesicular packets for envelopment and budding.

Membrane tubes and DMVs generated by picornaviruses

DMVs formed by membrane rearrangement

Picornaviruses, for example poliovirus, coxsackievirus, foot and mouth disease virus (FMDV), and the rhinoviruses that cause the common cold generate a single polyprotein that encodes the structural and nonstructural proteins (Figure 3(a)). This is cleaved by the viral protease 3C^{Pro} to generate capsid proteins from VP1–4 and replicase proteins from 2ABC and 3ABCD. The RdRp is encoded by 3D^{Pol} which is targeted to membranes by hydrophobic domains in 3A, 2B, and 2C. Picornaviruses generate large numbers of membrane vesicles which vary between 200 nm and 400 nm diameter and some have double membranes (DMVs) while others have single membranes. These DMVs were observed in the first electron micrographs generated from cells infected with viruses in the 1960s (Dales *et al.*, 1965) and their precise origins are still in dispute. The sizes and relative numbers of vesicles vary greatly between picornaviruses families and it is not clear whether single or DMVs house the RNA polymerase. For enteroviruses, such as poliovirus and coxsackievirus, replication may take place on tubular membranes generated during the log phase of virus production, and these may be seen as single membrane vesicles in cross section in electron micrographs (Limpens *et al.*, 2011; Belov *et al.*, 2012). Single membrane tubes may gain curvature following progressive assembly of replicase proteins or through recruitment of reticulon proteins (see below) leading to formation of DMVs (Figure 3(a)). The DMVs generated during picornavirus replication lack an obvious opening to the cytosol making it possible that they seal and are a by-product of replicase assembly and used to store viral RNA.

Double membrane autophagosomes

The DMVs seen in cells infected with enteroviruses may be formed from autophagosomes. Autophagosomes are generated by a membrane trafficking pathway called autophagy that

delivers portions of the cytosol to lysosomes for degradation. Autophagy is activated during many viral infections and capture by autophagosomes can lead to destruction of viruses in lysosomes. It was therefore a surprise when it was discovered that poliovirus benefits from autophagy, with virus yields increasing some 10-fold when autophagy is activated during infection (Jackson *et al.*, 2005; Wileman, 2006). The poliovirus replicase proteins associate with autophagosomes during infection and co-expression of 2BC and 3A in cells generates autophagosomes and DMVs (Schlegel *et al.*, 1996; Suhy *et al.*, 2000). It is currently believed that poliovirus 2BC and 3A activate autophagy during infection allowing replication to take place on the surface of autophagosomes. The enteroviruses infect via the gastrointestinal tract and are resistant to acid pH and proteases. It is likely that they can survive in lysosomes and there is evidence that low pH promotes maturation of the poliovirus capsid before release from cells by exocytosis (Richards and Jackson, 2012). Autophagosomes are used as sites of assembly for the replicases of other picornaviruses such as coxsackieviruses, human enterovirus 71, and encephalomyocarditis, but the role played by lysosomes during virus release is less clear. Unlike the enteroviruses, FMDV is very susceptible to low pH and this facilitates disassembly of the virus during endocytosis and cell entry. This sensitivity to low pH makes it unlikely that the FMDV would survive in autophagosomes or lysosomes. Autophagy is, however, activated during FMDV infection but this takes place early during infection before the nonstructural proteins are synthesized, and can be activated by empty capsids and UV inactivated viruses that cannot replicate (Berryman *et al.*, 2012). It is thought that FMDV capsids activate autophagy during cell entry and this is consistent with the co-localization of capsid proteins with autophagosomes.

Host cell contribution to membrane rearrangements

The assembly of replicase proteins during spherule formation induces negative curvature resulting in a membrane invagination that moves away from the cytosol into the organelle. The formation of DMVs may require additional proteins because the generation of the outer membrane requires induction of positive curvature, allowing the curved face of the DMV to extend into the cytoplasm. The positive curvature of ER tubes is maintained by a family of proteins called reticulons which have hydrophobic domains that insert into the outer leaflet of the phospholipid bilayer (Diaz and Ahlquist, 2012; Tang *et al.*, 2007). The picornavirus 2C proteins that provide a scaffold for the 3D^{Pol} polymerase and target the replicase complex to membranes bind reticulon proteins and replication is reduced in the absence of reticulons. Most DMVs are derived, at least in part, from the ER, making it possible that reticulons play a key role in the formation of DMVs.

The lipid composition of membranes is also important for the replication of many different (+) strand RNA viruses and they are inhibited by drugs that block fatty acid and sterol synthesis (Castorena *et al.*, 2010; Kapadia and Chisari, 2005; Perez *et al.*, 1991; Rassmann *et al.*, 2007). Lipids can also provide sites for assembly of replicase complexes. The RdRp (3D^{Pol}) of poliovirus and coxsackievirus, for example, binds phosphatidylinositol-4-phosphate (PI4-P). This means that recruitment of the 3D^{Pol} polymerase can be increased if

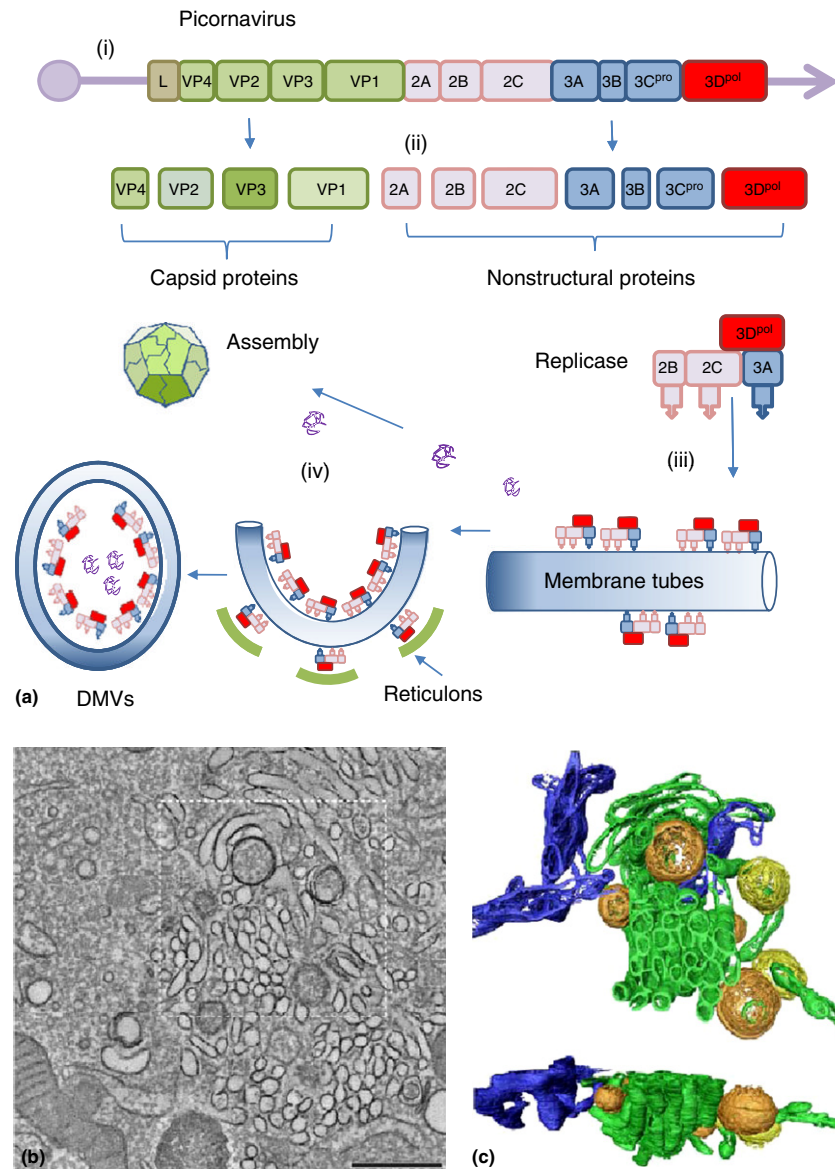


Figure 3 Formation of membrane tubes and double membrane vesicles by picornaviruses. (a) The picornaviruses polyprotein (i) encodes structural proteins (VP1–VP4) that generate the capsid and nonstructural proteins (2ABC and 3ABCD), required for replication. (ii) Processing of the polyprotein by 3C^{pro} generates the replicase from 2B, 2C, 2BC, 3A, and 3D^{pol} that are targeted to the cytoplasmic face of the endoplasmic reticulum by hydrophobic domains in 2B, 2C and 3A (iii). Replicase assembly leads to the generation of membrane tubes, the origins of the tubes are unknown but are likely to be derived from ER or cis Golgi. Later during infection tubes gain curvature, possibly powered by reticulon proteins (iv) and form double membrane vesicles (DMVs). (b) Electron micrograph of vesicles and tubes generated during coxsackievirus infection. The image is one of several tomographic slices used to generate rendered tomograms in panel C. Reproduced from Limpens, R.W., van der Scharr, H. M., Kumar, D., *et al.*, 2011. The transformation of enterovirus replication structures: A three dimensional study of single and double membraned compartments. *mBio* 2, e00166-11. (c) Side and top views of surfaced rendered models of membrane vesicles showing single membrane tubes (green), ER (blue) and open (orange) and closed (yellow) DMVs. Reproduced from Knoops, K., Kikkert, M., Worm, S.H., *et al.*, 2008. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biology* 6, e226.

membranes contain increased levels of PI4-P. For some viruses this is achieved by manipulation of host proteins that control membrane traffic through the secretory pathway. The 3A protein of poliovirus and coxsackievirus increases the activity of the Arf1 GTPase and its guanine nucleotide exchange factor GBF1. Under normal circumstances activation of Arf1 by GBF1 results in formation of the COP1-coated vesicles that carry proteins

from the Golgi to the ER. Activation of GBF1 by the viral 3A protein, however, results in recruitment of an alternative Arf1 effector protein, phosphatidylinositol-4-kinase-III β (PI4-III β) to membranes (Hsu *et al.*, 2010). The enzyme phosphorylates PI4 to generate PIP-4 providing membranes enriched in PIP-4 which recruit 3D^{pol}. A similar mechanism may operate for flaviviruses (Reiss *et al.*, 2011; Berger *et al.*, 2009) and

coronaviruses (Verheije *et al.*, 2008) where replication and membrane remodeling are reduced following inhibition of the Arf1 GTPases or phosphatidylinositol-4-kinase-III.

(–) Strand RNA Viruses Replicate in Factories Assembled in the Golgi

The 'Bunyaviridae' are good examples of (–) strand RNA viruses and include Hantavirus, Rift Valley fever virus, and Crimean Congo hemorrhagic fever virus which cause severe diseases in man. The Bunyamwera virus serves as a good model for this family of viruses. The genome is segmented into three (–) strand RNA strands named according to their size. The long (L) segment encodes the RdRp, the medium segment encodes a polyprotein that is processed by the ER to make the envelope proteins Gn and Gc which travel to the Golgi apparatus, while the small segment encodes the nuclear protein (NP). One copy of each segment is packaged within the virus and each segment binds one copy of the RdRp and the rest of the RNA strand is protected by several NP proteins. On entering the cell the RdRp copies each (–) strand to make (+) strands that can be read as mRNA. Infection of cells with Bunyamwera virus results in the formation of tubules within the Golgi apparatus which form the virus factory (Fontana *et al.*, 2008). These tubules contain the RdRp, NP, and viral RNA. It is likely that they function as specialized spherules for generating new genome segments which are assembled with the RdRp and NP. The formation of the factory in the Golgi apparatus would facilitate delivery of RNA segments budding into late Golgi compartments during virus assembly and envelopment.

DNA Viruses

The Nucleo-Cytoplasmic Large DNA Viruses Use Aggresomes to Generate Pericentriolar Factories

The nucleo-cytoplasmic large DNA viruses (NCLDV) are a family of large DNA viruses with genomes ranging from 150 kb to 1.2 MB. The family includes the poxviruses iridoviruses, African swine fever virus (ASFV), phycodnaviruses, and mimiviruses that share several conserved core genes. The NCLDV infect a wide range of hosts including mammals, arthropods, fish, and marine plankton, but all replicate and assemble in large factories located close to the nucleus at the microtubule organizing center (MTOC) (Figure 4). The factories contain viral DNA and structural proteins but lack cellular membranes and are very similar to inclusions called aggresomes that form during protein misfolding diseases (Heath *et al.*, 2001; Wileman, 2007). Protein misfolding leads to the formation of small protein aggregates in the cytosol which are transported along microtubules by dynein motors to the MTOC where they encounter the major cellular pools of chaperones and proteasomes. If aggregates fail to fold, or exceed the proteolytic capacity of the proteasomes, they accumulate within an aggresome surrounded by mitochondria and a cage made from the vimentin filaments (Figure 4(a)). The inclusions also contain ubiquitin and a linker protein called p62/SQSTM1 that binds ubiquitin and LC3. LC3 is the major membrane protein of the autophagosome and p62 is used to remove ubiquitinated

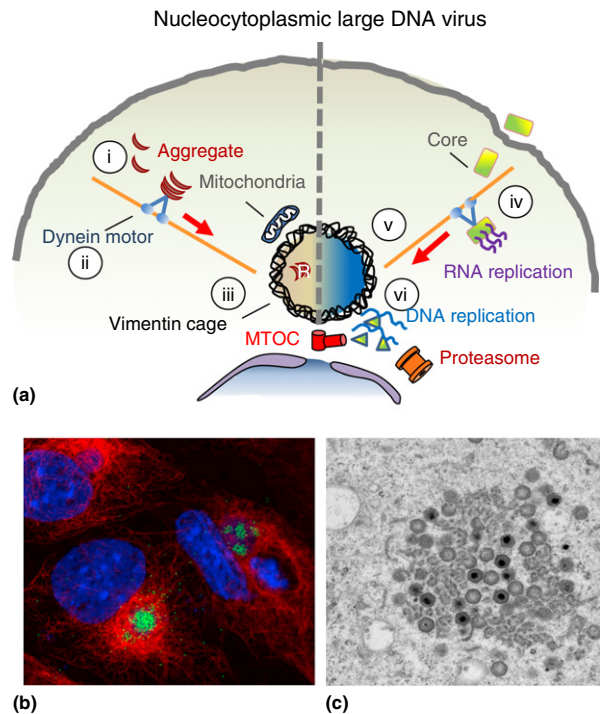


Figure 4 Aggresomes and perinuclear sites generated by nucleocytoplasmic large DNA viruses. (a) (i) Misfolded proteins aggregate in the cytoplasm and bind to cytoplasmic dynein. (ii) Dynein motors move aggregated proteins by retrograde transport to the microtubule organizing center (MTOC). (iii) An aggresome forms and induces vimentin rearrangement and recruits of mitochondria and chaperones. (iv) Poxvirus cores enter the cytoplasm and bind dynein and microtubules and begin RNA replication. (v) Poxvirus cores are disassembled at the MTOC by proteasomes dynein and microtubules and begin DNA replication. (vi) DNA replication and virus assembly take place in factories located at the MTOC. (b) Virus factory generated by African swine fever virus. The image shows a vimentin cage (red) surrounding a virus factory containing newly assembled viruses (green). The nucleus is blue. (c) Electron micrograph African swine fever virus factory showing viruses at different stages of assembly. Fully assembled viruses are seen as hexagons in cross section. Those with dark centers have packaged DNA genomes.

proteins from aggresomes into autophagosomes which deliver them to lysosomes for degradation. Virus factories and aggresomes share many features in common including recruitment of chaperones, proteasomes, and mitochondria and confinement within a vimentin cage. Many viruses are recognized by dynein after entering cells and delivered to the MTOC. It is possible that incoming virus capsids stimulate an aggresome response because they appear foreign or misfolded to cells (Wileman, 2007). For the NCLDV this may be beneficial and provide transport to a site for replication, for other viruses it may represent an innate immune response leading to confinement at the MTOC and degradation.

The factories formed by poxviruses and ASFV are the best characterized. During cell entry the poxviruses release a viral core into the cytoplasm which binds dynein motor proteins (Figure 4(a)). The core immediately starts the synthesis of at least 100 early genes using a viral DNA-dependent RNA

polymerase. The mRNA transcribed from early genes by the polymerase are delivered into the cytosol as the core travels to the MTOC. The early proteins encoded by the mRNA are used to break open the core to release viral DNA through a process requiring proteasome-dependent degradation of ubiquitinated capsid proteins. Poxviruses are unusual DNA viruses because they encode all the enzymes needed for DNA and mRNA synthesis and do not need to travel to the nucleus. The viral genomes remain at the MTOC where they create a factory by providing a template for DNA replication and generation of viral mRNA encoding intermediate and late genes. Interestingly, for poxviruses it appears that intermediate and late proteins are synthesized within the factory rather than in the cytoplasm. This means that viral mRNA is retained within the factory and that the factory recruits cellular factors needed for translation. Detailed analysis of poxvirus factories shows that cavities within the viral DNA recruit viral mRNA, intermediate and late proteins, and host cell translation initiation factors such as eIF4g and eIF4e. The results suggest that translation of viral mRNA takes place on ribosomes recruited into the factory, and that sites of translation are segregated from sites of genome replication (Katsafanas and Moss, 2007). Viral structural proteins are retained in the factory until they are assembled into new viruses which are carried into the cytosol along microtubules.

Similar to poxviruses, infection of cells by ASFV also results in delivery of a nucleoprotein core particle into the cytosol. The cores deliver the genome to the nucleus where it provides a template for the formation of short precursor fragments that are exported to the cytosol for replication. The formation of the ASFV factory begins with the removal of cellular material from the MTOC where it is replaced by a high concentration of vimentin filaments arranged into an 'aster' aligned along microtubules. This site also recruits chaperones and mitochondria. Following the onset of DNA replication the vimentin filaments are rearranged into a cage surrounding virus DNA and delineate the edges of the factory (Figure 4(b)). The factory is kept at the MTOC by microtubules and becomes dispersed through the cytosol if microtubules are depolymerized. The rearrangement of vimentin into a cage requires phosphorylation of the N-terminal domain of vimentin that is important for filament assembly by host cell calcium calmodulin-dependent kinase II (CamKinase II). This phosphorylation is required for replication because inhibition of CamKinase II blocks late gene expression (Stefanovic *et al.*, 2005). In common with poxviruses, ASFV structural proteins do not leave the factory until they are assembled into new viruses (Figure 4(c)). At late stages in the lifetime of the factory the MTOC is lost and microtubules form bundles in the cytosol. It is thought that this may weaken the factory to facilitate release of viruses into the cytosol.

The DNA Viruses Generate Replication Compartments in the Nucleus

Most DNA viruses replicate in the nucleus because they need nuclear factors to generate new genomes (reviewed in Schmid *et al.*, 2014; Jiang and Imperiale, 2012). DNA viruses entering the cell are transported by dynein motors along microtubules to the MTOC where they are transferred to the nuclear pore complex for disassembly and release of the DNA genome into the nucleus (Figure 5). The viral genome must recruit host

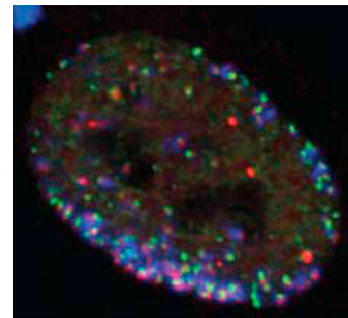
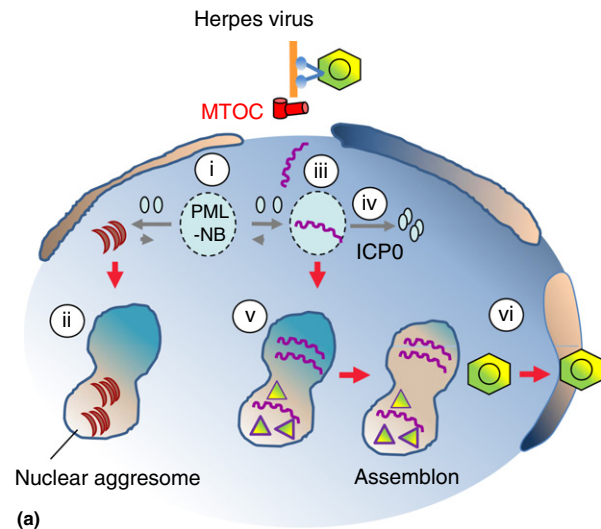


Figure 5 Replication compartments generated by DNA viruses in the nucleus. (a) Herpes virus replication sites in the nucleus. DNA viruses are transferred from the MTOC to the nuclear pore complex for disassembly and release of the DNA genome into the nucleus. In many cases the RCs form close to the nuclear envelope next to a nuclear substructures called promyelocytic leukemia nuclear bodies (PML-NB) which vary between 100 and 1000 μm in diameter and contain PML Daxx, Sp100, and ATRX (i). PML-NBs have also been called nuclear aggresomes (ii) because they accumulate misfolded proteins, cellular chaperones, proteasome subunits, and ubiquitinated proteins. Viral DNA entering the nucleus lacks chromatin and exposes a damage signal that recruits PML-NB proteins to add chromatin to genomes and repress transcription (iii). For HSV-1 this cellular defense is countered by the regulatory protein ICP0 which targets PML for degradation (iv). The residue of the PML body (v) may form the 'assemblon' that is seen late during infection which contains tegument proteins and aggregated capsids. The assemblon may serve as a site of virus assembly early during infection and become the VICE domain/nuclear aggresome at later times when structural proteins are made in excess. Fully assembled capsids (vi) pass across the nuclear envelope and enter the cytoplasm and envelop in the trans-Golgi network before release from the cell. (b) Recruitment of PML-NB proteins to sites of HSV-1 replication. Cells were infected with a strain of virus (ΔICP0) lacking ICP0 to slow dispersal of PML-NB proteins. The fluorescence micrograph shows the virus genome located through detection of transcriptional activator ICP4 (green), PML protein (red) and DNA damage response protein γH2AX (blue). The PML protein has redistributed to sites containing viral DNA. Reproduced from Everett, R.D., 2013. The spatial organisation of DNA virus genomes in the nucleus. PLoS Pathogens 9, e1003386.

factors required for replication and establish sites for DNA replication and transcription in the nucleus, and at the same time evade cellular defenses against infection. The delivery of viral DNA to the nucleus initiates large-scale rearrangement of nuclear structures and the formation of replication compartments that can expand to fill the nucleus late during infection. In many cases the RCs form next to nuclear substructures called promyelocytic leukemia nuclear bodies (PML-NB), ND10s or PODs which store proteins involved in apoptosis, chromatin remodeling, and DNA repair (reviewed in [Everett, 2013](#); [Boutell and Everett, 2013](#)). Many of the components of the PML-NBs are upregulated by interferon suggesting they play a role in defense against infection. The RCs initially form close to the nuclear envelope, possibly where the viral DNA first enters the nucleus. A key question has centered on determining whether sites of DNA virus replication are indeed PML-NBs, and understanding how a virus would survive in a compartment designed to inhibit replication. The interaction of DNA viruses with PML-NBs is best studied for herpes simplex virus-1 (HSV-1), but is likely to be similar during the replication of other DNA viruses.

Use of promyelocytic leukemia nuclear bodies (PML-NB) by herpes virus

PML-NBs vary between 100 µm and 1000 µm in diameter and contain PML, the major structural protein, and Daxx, Sp100, and ATRX which regulate chromatin assembly and inhibit transcription. The PML-NBs do not move through the nucleus but PML proteins exchange rapidly with the surrounding nucleoplasm. Recruitment into the PML-NB is dependent on addition of ubiquitin-like SUMO proteins to PML, and the presence of SUMO interaction motifs in Daxx and Sp100. Viral DNA entering the nucleus lacks chromatin and this may expose a damage signal that recruits PML-NB proteins which would add chromatin to genomes and repress transcription ([Figure 5](#)). For HSV-1 this cellular defense is countered by the regulatory protein ICP0 which is expressed very early during infection. The ubiquitin E3 ligase activity of ICP0 ubiquitinates PML leading to degradation of PML by proteasomes and loss of PML-NBs from the nucleus of infected cells. Evidence that PML-NBs suppress virus replication before they are dispersed by ICP0 comes from the observation that HSV-1 strains expressing defective ICP0s fail to degrade PML, have reduced replication and can fail to proceed to lytic infection. Herpes viruses also use structural proteins to combat PML-NBs. The herpes virus nucleoprotein core is surrounded by a layer of tegument proteins (i.e., viral matrix) that lie beneath the envelope. The human cytomegalovirus tegument phosphoprotein pp71 locates to PML-NBs soon after cell entry and induces degradation of PML, similarly the major tegument protein of Epstein-Barr virus BNRF1 disassembles the Daxx-ATRX complex. Daxx-ATRX complexes are also displaced from PML-NBs by specific capsid proteins of adenoviruses and papillomaviruses that travel to the nucleus following cell entry.

PML-NBs have also been called nuclear aggresomes because they are closely associated with nuclear inclusions that accumulate misfolded proteins, cellular chaperones, proteasome subunits, and ubiquitinated proteins ([Fu et al., 2005](#)). Aggregated proteins, for example, the polyQ repeat proteins associated with Huntington's disease or spinocerebellar ataxia accumulate

close to, or within PML-NBs which fuse together to form larger structures. Nuclear aggresomes also appear during herpes virus infection, and in some reports they are called virus-induced-chaperone-enriched (VICE) domains. These domains appear adjacent to sites of replication soon after the dispersal of PML-NBs by ICP0. Proteomic analysis has shown that 20 or more host proteins may be recruited to replication sites. These include proteins involved in DNA replication and transcription, chromatin remodeling and DNA repair, and RNA polymerase II which is used to transcribe the viral genome. As seen for the poxviruses the sub-compartmentalization of the virus factory through the formation of a VICE domain may facilitate replication by sequestering structural proteins away from sites of genome replication and transcription. VICE domains may also store viral proteins, such as the UL6 viral portal protein, and pp65 tegument protein, that are prone to aggregation when they are made in excess over levels needed for the assembly of new viruses. Evidence that aggregation of viral proteins is detrimental to replication comes from studies of HCMV strains expressing a defective tegument protein called UL97. UL97 is a kinase that can reduce protein aggregation ([Prichard et al., 2005](#)). Drugs that inhibit the kinase activity of UL97, or mutations that abolish kinase activity, reduce replication and induce the formation of aggresomes containing tegument proteins.

Successful subversion of the cellular defenses provided by the PML-NB allows the virus to initiate DNA replication and transcription. Precisely where this occurs in the nucleus is not known because the PML-NB markers are dispersed by ICP0. The genomes of HSV-1 expressing a defective ICP0 accumulate in large PML-NBs making it likely that replication takes place within the residue of the PML-NB or VICE domain. Once genomes have been generated in the factory they have to be assembled with capsids. Capsid proteins are synthesized on ribosomes in the cytosol and imported into the nucleus where they assemble with scaffold proteins and a portal protein to generate empty immature capsids. Import of the genome through the portal results in proteolysis and loss of the scaffold and formation of the mature capsid. These events take place in the nucleus and early studies identified an 'assemblon' as a site of virus assembly, seen late during infection which contains tegument proteins and aggregated capsids. The assemblon may serve as a site of virus assembly early during infection and becomes the VICE domain/nuclear aggresome at later times when structural proteins are made in excess over levels that can be packaged into viruses. Fully assembled capsids pass across the nuclear envelope and enter the cytoplasm where they are enveloped in the trans-Golgi network before release from the cell.

See also: intracellular infectiology: infectious agents: Cell Biology of Virus Infection

References

- Ahlquist, P., 2006. Parallels among positive-strand RNA viruses, reverse transcribing viruses and double stranded RNA viruses. *Nature Reviews Microbiology* 4, 371–382.

- Barajas, D., Fernandez de Castro Martin, I., Pogany, J., Risco, C., Nagy, P.D., 2014. Noncanonical role for the host Vps4 AAA + ATPase ESCRT protein in the formation of Tomato Bushy Stunt Virus replicase. *PLoS Pathogens* 10, e1004087.
- Belov, G.A., Nair, V., Hansen, B.T., *et al.*, 2012. Complex dynamic development of poliovirus membranous replication complexes. *Journal of Virology* 86, 302–312.
- Berger, K.L., Cooper, J.D., Heaton, N.S., *et al.*, 2009. Roles for endocytic trafficking and phosphatidylinositol 4-kinase III alpha in hepatitis C virus replication. *Proceedings of the National Academy of Sciences of the United States of America* 106, 7577–7582.
- Berryman, S., Brooks, E., Burman, A., *et al.*, 2012. Foot-and-mouth disease virus induces autophagosomes during cell entry via a Class III Phosphatidylinositol 3-Kinase-Independent pathway. *Journal of Virology* 86, 12940–12953.
- den Boon, J.A., Ahlquist, P., 2010. Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annual Review of Microbiology* 64, 241–256.
- den Boon, J.A., Diaz, A., Ahlquist, P., 2010. Cytoplasmic viral replication complexes. *Cell Host & Microbe* 8, 77–85.
- Boutell, C., Everett, R.D., 2013. Regulation of alphaherpesvirus infections by the ICP0 family of proteins. *Journal of General Virology* 94, 465–481.
- Castorena, K.M., Stapleford, K.A., Miller, D.J., 2010. Complementary transcriptomic, lipidomic and targeted functional genetic analyses in cultured *Drosophila* cells highlight the role of glycerophospholipid metabolism in Flock House virus RNA replication. *BMC Genomics* 11, 183.
- Cottam, E., Pierini, R., Roberts, R., Wileman, T., 2009. Origins of membrane vesicles generated during replication of positive-strand RNA viruses. *Future Virology* 4, 473–485.
- Dales, S., Eggers, H.J., Tamm, I., Palade, G.E., 1965. Electron microscopic study of the formation of poliovirus. *Virology* 26, 379–389.
- Diaz, A., Ahlquist, P., 2012. Role of host reticulon proteins in rearranging membranes for positive strand RNA virus replication. *Current Opinion in Microbiology* 15, 519–524.
- Diaz, A., Zhang, J., Ollwerther, A., Wang, X., Ahlquist, P., 2015. Host ESCRT proteins are required for bromovirus RNA replication compartment assembly and function. *PLoS Pathogens* 11 (3), e1004742.
- Everett, R.D., 2013. The spatial organisation of DNA virus genomes in the nucleus. *PLoS Pathogens* 9, e1003386.
- Fernandez de Castro, I., Volonte, L., Risco, C., 2012. Virus factories: Biogenesis and structural design. *Cellular Microbiology* 15, 24–34.
- Fontana, J., Lopez-Montero, N., Elliott, R.M., Fernandez, J.J., Risco, C., 2008. The unique architecture of Bunyamwera virus factories around the Golgi complex. *Cellular Microbiology* 10, 2012–2028.
- Fu, L., Gao, Y.-S., Tousson, A., *et al.*, 2005. Nuclear aggresomes form by fusion of PML-associated aggregates. *Molecular Biology of the Cell* 16, 4905–4917.
- Heath, C.M., Windsor, M., Wileman, T., 2001. Aggresomes resemble sites specialized for virus assembly. *Journal of Cell Biology* 153, 449–456.
- Hsu, N.Y., Ilnytska, O., Belov, G., *et al.*, 2010. Viral reorganization of the secretory pathway generates distinct organelles for RNA replication. *Cell* 141, 799–811.
- Jackson, W.T., Giddings Jr., T.H., Taylor, M.P., *et al.*, 2005. Subversion of cellular autophagosomal machinery by RNA viruses. *PLoS Biology* 3, e156.
- Jiang, M., Imperiale, M.J., 2012. Design stars: How small DNA viruses remodel the host nucleus. *Future Virology* 7, 445–459.
- Kapadia, S.B., Chisari, F.V., 2005. Hepatitis C virus RNA replication is regulated by host geranylgeranylation and fatty acids. *Proceedings of the National Academy of Sciences of the United States of America* 102, 2561–2566.
- Katsafanas, G.C., Moss, B., 2007. Colocalization of transcription and translation within cytoplasmic poxvirus factories coordinates viral expression and subjugates host functions. *Cell Host & Microbe* 2, 221–228.
- Knoops, K., Kikkert, M., Worm, S.H., *et al.*, 2008. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biology* 6, e226.
- Kopek, B.G., Perkins, G., Miller, D.J., Ellisman, M.H., Ahlquist, P., 2007. Three-dimensional analysis of a viral RNA replication complex reveals a virus-induced mini-organelle. *PLoS Biology* 5, e220.
- Limpens, R.W., van der Schar, H.M., Kumar, D., *et al.*, 2011. The transformation of enterovirus replication structures: A three dimensional study of single and double membraned compartments. *mBio* 2, e00166–11.
- Miller, D.J., Schwartz, M.D., Ahlquist, P., 2001. Flock house virus RNA replicates on outer mitochondrial membranes in *Drosophila* cells. *Journal of Virology* 75, 11664–11676.
- Netherton, C., Moffat, K., Brooks, E., Wileman, T., 2007. A guide to viral inclusions, membrane rearrangements, factories, and viroplasm produced during virus replication. *Advances in Virus Research* 70, 101–182.
- Perez, L., Guinea, R., Carrasco, L., 1991. Synthesis of Semliki Forest virus RNA requires continuous lipid synthesis. *Virology* 183, 74–82.
- Prichard, M.N., Britt, W.J., Daily, S.L., Hartline, C.B., Kern, E.R., 2005. Human cytomegalovirus UL97 kinase is required for the normal intranuclear distribution of pp65 and virion morphogenesis. *Journal of Virology* 79, 15494–15502.
- Randall, R.E., Goodbourn, S., 2008. Interferons and viruses: An interplay between induction, signalling, antiviral responses and virus countermeasures. *Journal of General Virology* 98, 1–47.
- Rassmann, A., Henke, A., Jarasch, N., *et al.*, 2007. The human fatty acid synthase: A new therapeutic target for coxsackievirus B3-induced diseases? *Antiviral Research* 76, 150–158.
- Rathinam, V.A.K., Fitzgerald, K.A., 2011. Innate immune sensing of DNA viruses. *Virology* 411, 153–162.
- Reiss, S., Rebhan, I., Backes, P., *et al.*, 2011. Recruitment and activation of a lipid kinase by hepatitis C virus NS5A is essential for integrity of the membranous replication compartment. *Cell Host & Microbe* 9, 32–45.
- Richards, A.L., Jackson, W.T., 2012. Intracellular vesicle acidification promotes maturation of infectious poliovirus particles. *PLoS Pathogens* 8, e1003046.
- Romero-Brey, I., Merz, A., Chiramel, A., *et al.*, 2012. Three-Dimensional Architecture and Biogenesis of Membrane Structures Associated with Hepatitis C Virus Replication. *PLoS Pathogens* 8 (12), e1003056.
- Schlegel, A., Giddings Jr., T.H., Ladinsky, M.S., Kirkegaard, K., 1996. Cellular origin and ultrastructure of membranes induced during poliovirus infection. *Journal of Virology* 70, 6576–6588.
- Schmid, M., Speiseder, T., Dobner, T., Gonzalez, R.A., 2014. DNA virus replication compartments. *Journal of Virology* 88, 1404–1420.
- Stefanovic, S., Windsor, M., Nagata, K.-I., Inagaki, M., Wileman, T., 2005. Vimentin rearrangement during African swine fever virus infection involves retrograde transport along microtubules and phosphorylation of vimentin by calcium calmodulin kinase II. *Journal of Virology* 79, 11766–11775.
- Suhy, D.A., Giddings Jr., T.H., Kirkegaard, K., 2000. Remodeling the endoplasmic reticulum by poliovirus infection and by individual viral proteins: An autophagy-like origin for virus-induced vesicles. *Journal of Virology* 74, 8953–8965.
- Tang, W.F., Yang, S.Y., Wu, B.W., 2007. Reticulon 3 binds the 2C protein of enterovirus 71 and is required for replication. *Journal of Biological Chemistry* 282, 5888–5898.
- Verheije, M.H., Raaben, M., Mari, M., *et al.*, 2008. Mouse hepatitis coronavirus RNA replication depends on GBF1-mediated ARF1 activation. *PLoS Pathogens* 4, e1000088.
- Welsch, S., Miller, S., Romero-Brey, I., *et al.*, 2009. Composition and three-dimensional architecture of the dengue virus replication and assembly sites. *Cell Host & Microbe* 5, 365–375.
- Wileman, T., 2006. Aggresomes and autophagy generate sites for viral infection. *Science* 312, 875–878.
- Wileman, T., 2007. Aggresomes and pericentriolar sites of virus assembly: Cellular defense or viral design? *Annual Review of Microbiology* 61, 149–167.

HIV – The Cell Biology of Virus Infection and Replication

S Giese, A Pelchen-Matthews, and M Marsh, University College London, London, UK

© 2016 Elsevier Inc. All rights reserved.

Glossary

Acquired immunodeficiency syndrome Immunodeficiency disease caused by human immunodeficiency virus (HIV). Infection with HIV gradually damages the immune system, leading to increased susceptibility to infection by multiple opportunistic pathogens.

Lentiviruses Subgroup of retroviruses including HIV-1, HIV-2, and closely related simian immunodeficiency viruses.

Restriction factors Cellular proteins that inhibit viral replication.

Retroviruses Group of RNA viruses in which the RNA genome is reverse transcribed to DNA through the action of a virally encoded reverse transcriptase that is packaged into virus particles.

Reverse transcriptase Viral enzyme that converts RNA to DNA. Hallmark of retroviruses.

Virological synapse Intercellular contact across which viruses may be efficiently transferred from an infected to an uninfected cell.

Introduction

The human immunodeficiency virus type 1 (HIV-1) was discovered in the early 1980s in patients with so-called acquired immunodeficiency syndrome (AIDS). Subsequently, a compelling body of evidence has emerged that supports the hypothesis that HIV-1 infection in man is the primary cause of AIDS. Currently, about 35 million people worldwide are infected with HIV and a similar number have died as a consequence of the infection. Despite detailed knowledge of the mode of person-to-person transfer, new HIV infections continue at a rate of 2–3 million per year, mostly in sub-Saharan Africa.

HIV-1 is a lentivirus – a genus of the retrovirus family, i.e., it contains an RNA genome that is reverse transcribed into DNA following infection of new target cells. The virus is believed to have originated in great apes (chimpanzees and gorillas) by recombination of two coinfecting simian immunodeficiency viruses (SIV). SIV cpz/gor transferred to man on more than one occasion, but the main or M group HIV-1, which is the predominant group of viruses pathogenic in man, became established in the human population in the early twentieth century (Faria *et al.*, 2014). A second HIV, HIV-2, originating in West Africa, has also been described. This virus, derived from an SIV infecting sooty mangabeys (SIV smm), is genetically distinct from HIV-1, and less pathogenic in man. Here we will restrict our discussion to HIV-1.

HIV is mainly transmitted through sexual intercourse, but can also be transferred through transfusion of infected blood or blood products, through the use of contaminated needles, during child birth, and through breast-feeding. Following transfer, HIV targets the CD4-expressing (CD4⁺) subset of T lymphocytes and CD4⁺ myeloid cells (primarily macrophages and dendritic cells). By infecting cells of the immune system, HIV compromises the host's ability to mount effective immune responses, resulting in the pathology of AIDS.

Despite intense research, no HIV vaccine is currently available. However, progress has been made in understanding the cellular and molecular mechanisms involved in HIV

replication, and this work has led to the development of drugs that can reduce viral replication and limit the pathological consequences of infection. As with many viruses, HIV is genetically simple – its genome contains just nine open reading frames encoding 15 proteins (Figure 1(c)). Thus, virtually all aspects of the cell biology of HIV replication require that the virus interacts with and exploits host cell proteins and synthetic processes.

The HIV Replication Cycle

Overview

In its cell-free form HIV is a 130–140 nm diameter enveloped virus, i.e., virus particles are enwrapped by a bilayered lipid membrane, derived from the plasma membrane (PM) of infected cells, that protects the viral genome during the extracellular phase of the replication cycle (Figures 1(a) and 1(b), Figure 2). Virus entry into host cells is initiated by interaction of the viral envelope glycoprotein (Env) with specific cellular receptors (CD4 and a chemokine receptor, usually CCR5 and/or CXCR4), which determine the viral tropism for CD4⁺ cells (see Figure 2). Following Env-mediated membrane fusion, the viral RNA genome is reverse transcribed through the action of reverse transcriptase (RT), a virally encoded enzyme that is incorporated into virions during assembly. The reverse transcribed DNA is imported into the nucleus and integrated into specific sites of the host DNA. In some cases the integrated viral DNA (or provirus) can remain inactive or latent, but mostly it functions as a template for the synthesis of viral proteins and viral genomic RNA. New viral proteins either modify the functions of the infected cell, or are directed to the inner surface of the PM where, together with two copies of the viral genomic RNA, they assemble to form new virus particles.

HIV virions are formed from about 2500 copies of p55Gag, a 55 kDa polyprotein comprising the matrix (MA), capsid (CA), and nucleocapsid (NC) proteins (Figures 1 and 3(a)). About 150 copies of a second polyprotein – GagPol, the fused

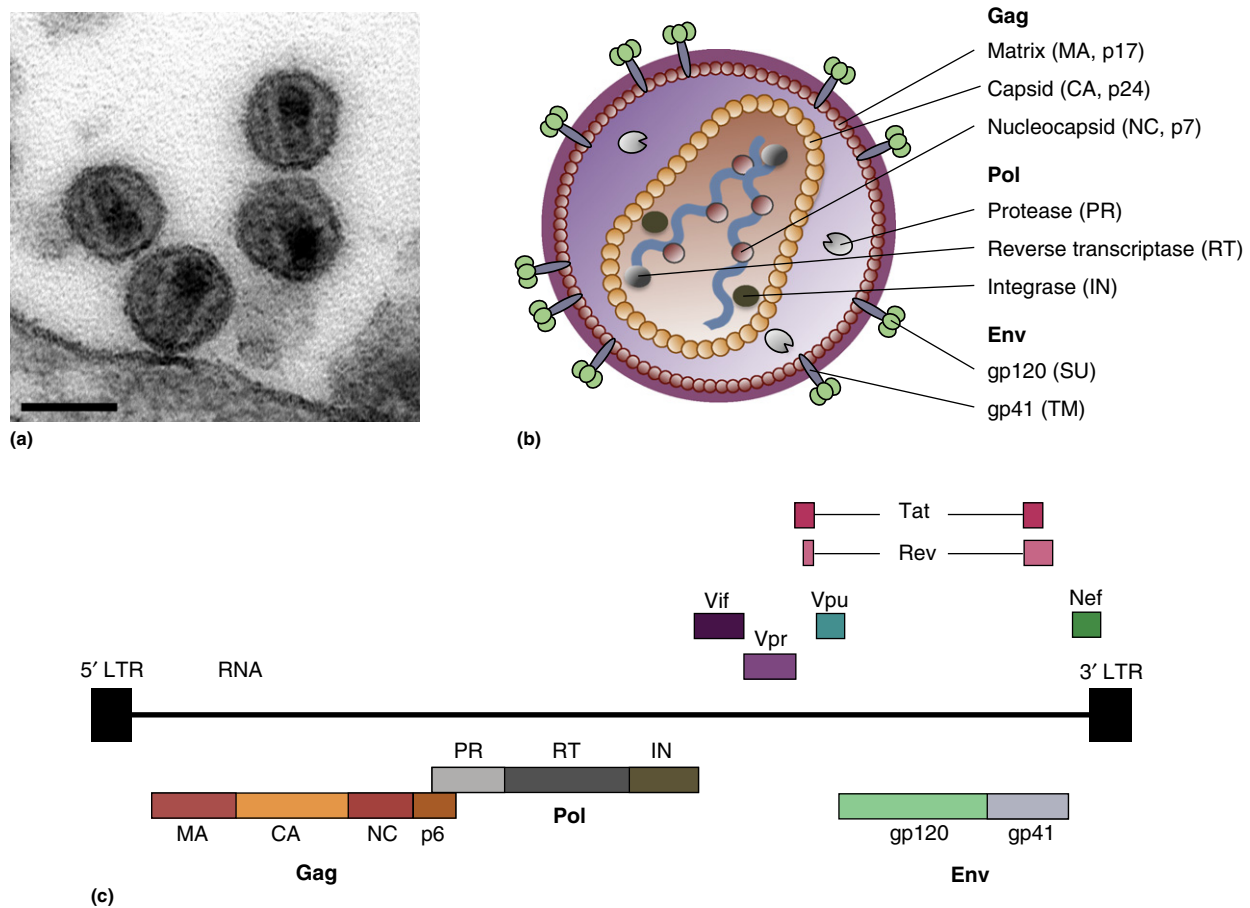


Figure 1 Structure and genomic organization of HIV-1. (a) Transmission EM image of lentivirus particles (SIV) at the surface of a lymphocyte. Scale bar, 100 nm. (b) Diagram of the structure of a mature HIV particle, showing the location of the major viral proteins and the host cell-derived lipid envelope. (c) Organization of the HIV-1 genomic RNA (in black), indicating the 5' and 3' long terminal repeats (LTR), as well as the nine open reading frames that encode the viral polypeptides Gag, GagPol and Env, and the accessory proteins Vif (virion infectivity factor), Vpr, Vpu (viral proteins R and U, respectively), Tat (trans-activator of transcription), Rev (regulator of expression of virion proteins), and Nef (negative factor; note that these protein names are historical and do not always reflect a functional activity).

product of the Gag and Pol genes, bring the essential virally encoded enzymes protease (PR), RT, and integrase (IN) into forming virus particles. Additionally, multiple copies of the trimeric viral envelope protein (Env) are incorporated into each particle, as well as variable amounts of other viral proteins, including Vpr and Nef, a tRNA^{Lys3}, and other host cytosolic and membrane proteins. During or shortly after budding is completed, GagPol and p55Gag are cleaved by the viral PR to generate mature infectious HIV particles. This basic outline of the HIV replication cycle is illustrated in [Figure 2](#). The interactions of the virus with host cells will be described in more detail below.

Virus Assembly

Gag Coordinates HIV Assembly

HIV is assembled at the PM of infected cells. Although in T cells the virus buds from the cell surface, in monocyte-derived macrophages deeply invaginated intracellular domains

of the PM are the preferred sites for assembly and release. HIV assembly is driven by the main structural protein Gag ([Figure 3](#)) and, when expressed in cells, Gag alone can form noninfectious virus-like particles. Gag is a polyprotein synthesized from full-length transcripts of integrated proviral DNA (these transcripts also form the viral genomic RNA). A second *gag* product translated from the same mRNA is GagPol (or GagProPol). Gag and GagPol are both synthesized from the same initiation codon, but in about 5% of the translation reactions ribosomal frameshifting (the ribosome slips back one nucleotide and shifts from the *gag* reading frame to the *pol* reading frame) results in the synthesis of a fusion protein that contains the bulk of Gag as well as the viral enzymes encoded by the *pol* gene. This ensures that Gag is made in adequate amounts, but smaller amounts of the viral enzymes (PR, RT, and IN) are produced for incorporation into virus particles. Both Gag and GagPol are synthesized on cytoplasmic ribosomes, co-translationally myristoylated on the glycine residue at position 2 (after removal of the N-terminal methionine), and are assumed to be directed to the PM through similar mechanisms (see [Figure 3\(b\)](#)). Gag also

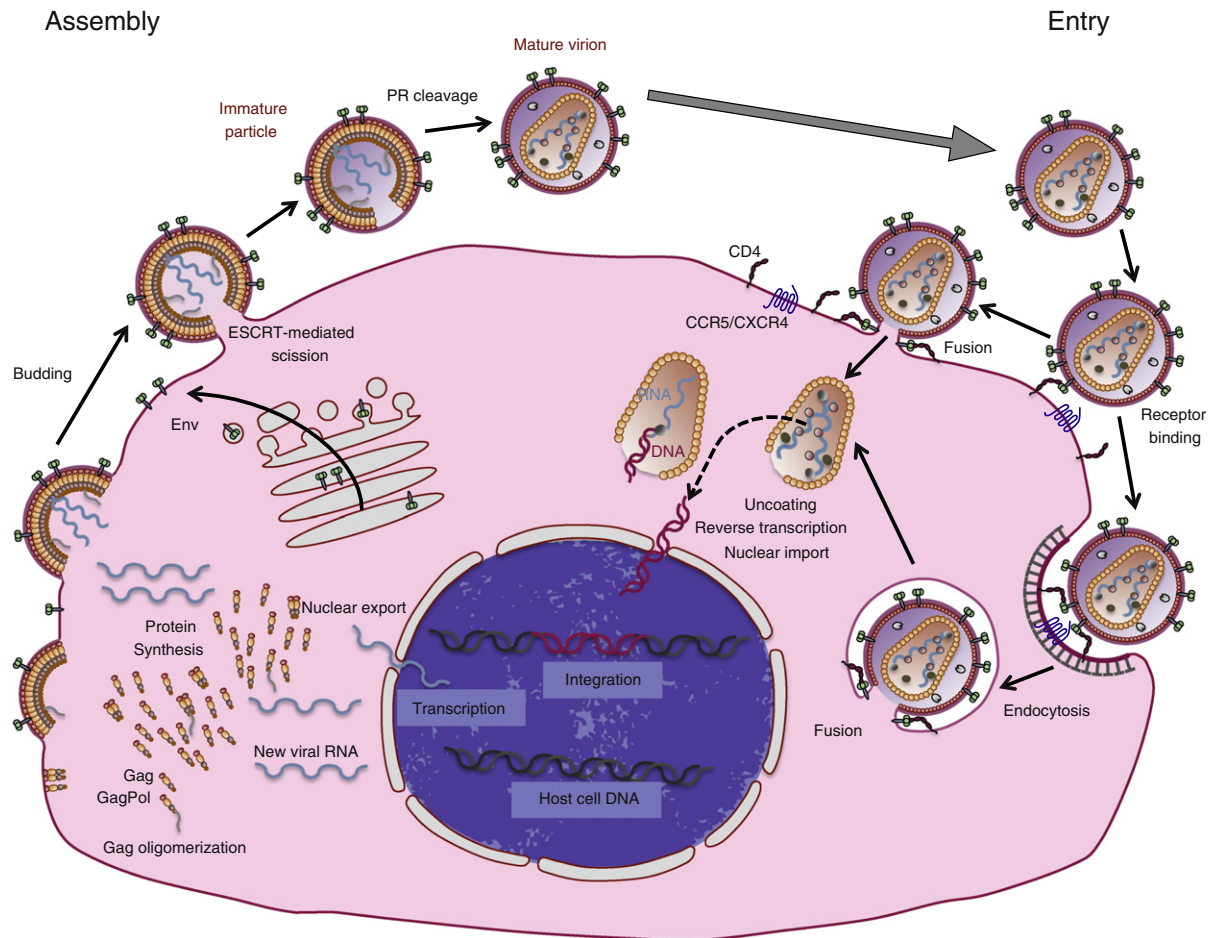


Figure 2 Cartoon illustrating the replication cycle of HIV. Left: HIV assembly. Viral genomic RNA and spliced mRNAs are produced by transcription from the integrated provirus and exported to the cytoplasm. Gag and GagPol polyproteins, as well as the HIV accessory proteins are synthesized on free ribosomes. Gag is targeted to the plasma membrane, where it oligomerizes, bending the membrane into a bud. The viral enzymes are incorporated into the bud by co-assembly with the GagPol polyprotein, and two copies of the viral genomic RNA are selected by the nucleocapsid portion of Gag. The Env protein is synthesized on the rough endoplasmic reticulum and transported to the plasma membrane via the Golgi/TGN. Immature virus particles are released from the plasma membrane by ESCRT-mediated scission of the bud stalk. Concomitant with or shortly after scission, the viral protease, PR, cleaves Gag and GagPol polyproteins, allowing the rearrangements that result in the formation of the characteristic cone-shaped cores of mature infectious viruses. Right: HIV entry. HIV binds its target cells via CD4 and a chemokine co-receptor, CCR5 or CXCR4, and fuses with the target cell plasma membrane, or from within an endosome after endocytosis of the virus. Following fusion, the RNA genome is reverse transcribed into DNA complex and transported into the nucleus – though the exact details of these steps are still poorly understood. In the nucleus, the viral integrase incorporates the DNA provirus into the host genome, where it may remain inactive (in a latently infected cell), or serve as a template for new viral RNAs.

recruits viral genomic RNA to the PM. The NC domain of Gag has two zinc fingers that select two copies of HIV RNA for packaging via the conserved ψ motif (Figure 3(d)).

Gag oligomerization is initiated in the cytoplasm and involves interaction with a number of cellular proteins, but the significance of most of these interactions for targeting Gag to the PM, and for virus assembly, remains unclear (Giese and Marsh, 2013). The ABCe1 ATPase complex, for example, has been proposed to function as a chaperone for newly synthesized Gag (Zimmerman *et al.*, 2002), and several reports have suggested that microtubule motor proteins play roles in Gag transport (Brass *et al.*, 2008; Azevedo *et al.*, 2009). In addition, the cellular lipid phosphatidylinositol-4,5-bisphosphate [PI(4,5)P₂] is particularly important for HIV assembly. PI(4,5)P₂ is located mainly in the inner leaflet of the PM, where it functions

in targeting many cytosolic proteins to the PM. When cellular PI(4,5)P₂ is depleted, by over-expression of PI 5-phosphatase IV, Gag is redirected to late endosomes, and virus production is impaired. Similarly, over-expression of a constitutively active Arf6 leads to the formation of intracellular vesicles rich in PI(4,5)P₂, and redirects HIV assembly to these internal vesicles (Ono *et al.*, 2004).

Gag is targeted to the cytoplasmic leaflet of the PM via its N-terminal MA domain, which is myristoylated; membrane association is further stabilized by a basic region on the surface of MA (residues 16–31) (Figure 3(b)). Gag molecules containing mutations that prevent myristoylation, or that disrupt the basic patch, are mis-targeted and fail to support budding (Sundquist and Kräusslich, 2012). The myristoylation, as well as binding of MA to PI(4,5)P₂, direct Gag to so-called lipid

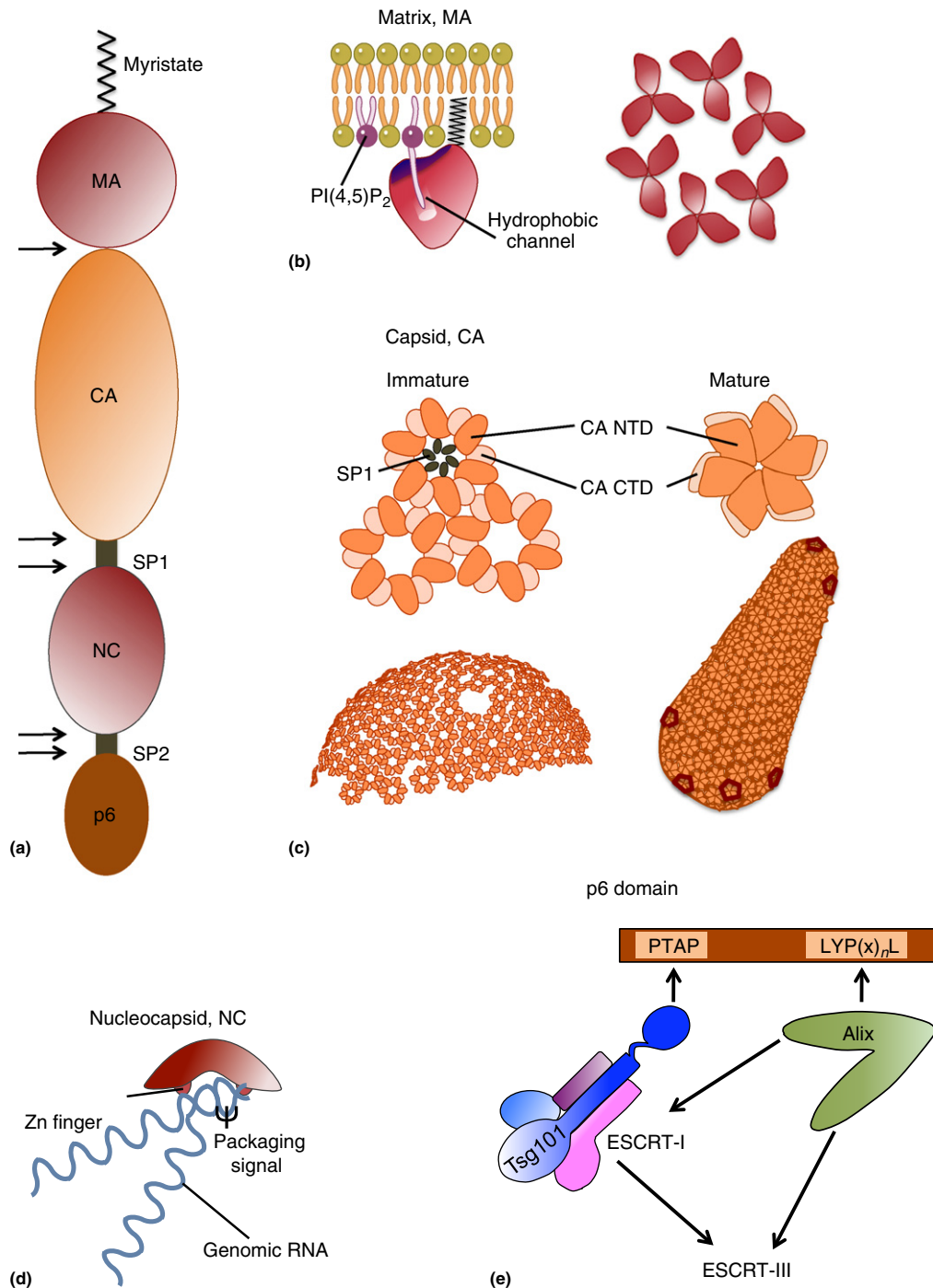


Figure 3 Cartoon illustrating the multiple functions of Gag in coordinating virus assembly. (a) Diagram of the HIV-1 Gag polyprotein, showing the N-terminal myristic acid moiety and the protein components Matrix (MA or p17), Capsid (CA or p24), Nucleocapsid (NC or p7) and the p6 domain, as well as the spacer peptides SP1 and SP2. Arrows indicate cleavage sites for the HIV-1 protease, PR. (b) The 'Matrix (MA)' domain targets the Gag polyprotein to the plasma membrane. MA binds to the plasma membrane via the N-terminal myristate and a highly basic region (indicated in blue). Membrane association is further stabilized by binding to the phospholipid PI(4,5)P₂ when the 2' acyl chain of PI(4,5)P₂ is extracted from the membrane and binds into a hydrophobic groove or channel in MA. The membrane-associated MA is organized into trimers (shown in the diagram on the right, orthogonal view to the membrane), which are loosely arranged into a hexameric lattice. (c) Role of the 'Capsid (CA)' in Gag oligomerization. CA can associate into at least two distinct hexameric configurations. As part of the Gag polyprotein in immature particles (left hand side), CA forms a lattice of hexamers of about 8 nm periodicity with occasional faults and gaps, allowing the formation of a spherical bud. After cleavage by PR and maturation of the viral core, CA reassembles into a distinct hexameric configuration (periodicity of about 9.5 nm, right hand side). The HIV core consists of about 250 CA hexagons and 12 pentagons to allow curvature (5 at the narrow and 7 at the blunt end). The CA N-terminal domain (NTD) and C-terminal domain (CTD) are indicated by dark and lighter orange shading, respectively. (d) The 'Nucleocapsid (NC)' of Gag contains two zinc finger domains that bind to the HIV genomic RNA via the ~150 nucleotide packaging sequence, Ψ. Two copies of the genomic RNA are incorporated into each virus particle. (e) The C-terminal 'p6' domain of Gag contains consensus sequences for binding of the ESCRT-I component Tsg101. In addition, p6 is able to bind the ESCRT-associated protein AIP1/ALIX that can recruit ESCRT-III or ESCRT-I to the assembling virion. p6 is also required for incorporation of the viral protein r (Vpr).

rafts, and HIV budding has been suggested to occur mainly through these liquid-ordered domains. Disruption of raft domains has been shown to reduce HIV release (Ono and Freed, 2001), and virus budding through raft domains may also explain why the membranes of HIV particles appear enriched for sphingolipids and cholesterol.

Following recruitment to the cytoplasmic leaflet of the PM, Gag assembles into a hexameric lattice that drives the formation of a membrane bud (Figure 3(c)). High-resolution cryo-electron tomography of immature particles shows the interactions between the CA domains of Gag that are crucial for lattice formation and virus particle assembly (Sundquist and Kräusslich, 2012).

Envelope Glycoprotein (Env) Trafficking

HIV Env is a type I integral membrane protein that must also be transported to the sites of virus assembly to be incorporated into nascent particles. Env is synthesized on the endoplasmic reticulum (ER) as a 160 kDa precursor, and in infected cells the bulk of the protein is located in the ER lumen. Gp160 trimerizes in the ER and is extensively glycosylated at up to 30 N-linked glycosylation motifs. Subsequently, the trimeric glycoprotein is transported to the Golgi apparatus, and eventually the PM. During transit through the secretory pathway some, but not all, of the N-linked oligosaccharides are processed to complex forms, and gp160 is cleaved by cellular furin or furin-like proteases to form the mature, non-covalently linked subunits, the surface unit (SU, or gp120) and the transmembrane protein (TM, gp41) (Figure 4). This proteolytic cleavage unmasks the hydrophobic fusion peptide at the new N-terminus of gp41, rendering Env competent for membrane fusion (Figure 5; Merk and Subramaniam, 2013). Currently, it remains unclear how exactly Env trimers are recruited into budding virions. However, EM tomography indicates that the number of Env trimers incorporated is low, with estimates ranging from 4 to 30 per virion (Zhu *et al.*, 2006).

The gp41 C-terminal cytoplasmic domain consists of about 150 amino acids and contains several conserved short sequence motifs homologous to trafficking signals of cellular proteins. The best characterized of these are a membrane-proximal GYxxØ motif and a C-terminal di-leucine motif (Figure 4). Both of these signals mediate AP-2 binding and clathrin-dependent endocytosis, and therefore regulate the distribution of Env in infected cells. Internalized Env is delivered to early endosomes, and can subsequently return to the PM via recycling endosomes. However, some Env is also sorted to late endosomes and lysosomes, where it is degraded. As a result, the steady state levels of Env on the surface of infected cells are low, which is likely to contribute to the low levels of Env incorporated into virions (Giese and Marsh, 2013).

HIV Release and Maturation

Live cell fluorescence imaging experiments indicate that it takes 5–10 min from the first appearance of Gag at the PM to complete recruitment of HIV components, though another 16–17 min may pass before the particles detach from the cells (Jouvenet *et al.*, 2008). Virus release requires that the narrow stalks, which connect budded particles to the PM, are cut

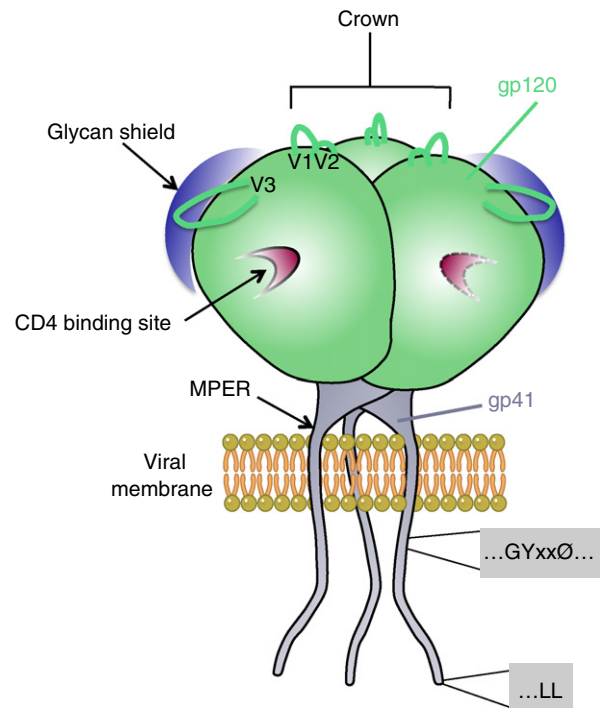


Figure 4 Cartoon showing the organization of the HIV-1 envelope glycoprotein (Env). HIV Env is a trimer of hetero-dimers comprising the surface glycoprotein SU (gp120, in light green) and the transmembrane protein TM (gp41, in gray). About half of the mass of gp120 consists of N-linked glycan residues, both complex as well as high mannose glycans, which shield the protein from antibody binding. The location of the V1V2 and V3 variable loops and the CD4 binding pocket are indicated. Other important neutralizing antibody epitopes are found at the crown of gp120 and in gp41 (the membrane-proximal external region, MPER). The locations of the conserved endocytosis signals in gp41 (GYxxØ and di-leucine) are also indicated.

by the endosomal sorting complex required for transport (ESCRT) machinery. A number of enveloped viruses use the cellular ESCRT complexes to allow their release from cells, and each has evolved one or more short amino acid sequences to mediate ESCRT recruitment. For HIV and related viruses, the C-terminal p6 domain of Gag (also termed the 'late domain' due to its requirement late in the assembly pathway) contains a PTAP motif that mimics the PSAP sequence in the ESCRT-0 protein Hrs that is responsible for recruiting the Tsg101 component of ESCRT-I to endosomes (Figure 3(e)). This PTAP-Tsg101 interaction is the main mechanism of ESCRT recruitment during HIV assembly, and mutations in the PTAP motif arrest virus budding at a late stage. The HIV Gag p6 domain also contains a second late domain motif, LYPx(n)L, which binds the ESCRT-associated protein AIP1/ALIX that can recruit both ESCRT-I and ESCRT-III. In the absence of a functional PTAP motif, over-expression of ALIX can partially rescue HIV release, though the affinity of ALIX for the LYPx(n)L motif is low, and thus it may not be able to rescue the release of PTAP mutants at physiological concentrations (Fisher *et al.*, 2007). The mechanism of ESCRT-mediated membrane scission is currently a major focus for research, with complementary views on how the machinery is assembled to

mediate membrane fission in the context of a budding virus (Sundquist and Kräusslich, 2012).

During, or shortly after virus release, autocatalytic activation of the viral PR leads to processing of both Gag and GagPol to generate the proteins of the mature virus. CA is reassembled around the viral RNA, NC, RT, and IN to form the conical capsid structure seen in mature viruses. As in immature particles, CA forms a hexagonal lattice, but with different contacts between the CA subunits (see Figure 3(c)). The inclusion of 12 pentameric units within the lattice allows for the formation of a closed capsid structure (Sundquist and Kräusslich, 2012). These distinct arrangements of the Gag CA domain in immature and mature virions drive two complementary processes in the HIV replication cycle, namely initial particle assembly and the formation of the mature capsid, which is essential for the subsequent activities of the virus core post fusion. The transition between the conformations is dependent on PR-mediated cleavage of Gag, thereby ensuring that these steps take place in the correct temporal sequence.

Virus Transmission

Following release from infected cells, viruses can spread by passive diffusion through extracellular fluids to infect receptor-positive cells. Most, if not all, viruses can use this cell-free mode of transmission. However, in many cases viruses may also spread by cell–cell transmission, during which an infected and a target cell come into direct contact. HIV has been suggested to exploit various structures for its direct cell–cell spread, of which the virological synapse (VS) is currently the best characterized, and is likely to be the most efficient. VS are named for their structural similarities to immune synapses, and are characterized by an accumulation of HIV proteins and the cellular receptors CD4 and CCR5/CXCR4 at intercellular junctions between infected and target cells. The high local concentrations of both infectious particles and their receptors are thought to promote efficient virus transfer (Sattentau, 2010).

Since there is no need for virus particles to diffuse over long distances, cell–cell transmission is likely to be faster than cell-free spread. As a consequence, viruses may be less likely to be detected by the immune system, and in some cases the intercellular contacts exploited for transmission may shield viruses from neutralizing antibodies (Sattentau, 2010).

In tissue culture, HIV exploits both cell-free and cell–cell transmission for its propagation, but the relative contributions of the two processes to virus spread *in vivo* remain unclear. In HIV-infected individuals, replication is observed in areas that are densely populated by migratory cells, e.g., gut-associated lymphoid tissue and peripheral lymph nodes, which would provide an ideal environment for direct cell–cell spread. Moreover, the observation of local ‘hotspots,’ or foci, of HIV replication in the tissues of HIV patients suggests that direct cell–cell propagation does occur *in vivo*. Recently, intravital multiphoton microscopy has revealed that HIV-infected T cells form stable contacts with uninfected CD4⁺ cells in the lymph nodes of humanized mice, but whether these contacts lead to the transfer of virus and infection of target cells remains unclear (Murooka *et al.*, 2012).

Virus Entry

Receptor Dependence and Specificity – Site of Fusion

HIV enters cells by fusion of its envelope with a target cell (Figure 5). Membrane fusion requires binding of the gp120 subunit of Env to its cellular receptor CD4 and to a chemokine co-receptor, mainly CCR5 and/or CXCR4. Most viruses use CCR5, but CXCR4-tropic viruses emerge in ~50% of patients at late stages of infection. Initial binding to CD4 triggers conformational changes in gp120, which reveal the co-receptor binding site, a bridging sheet at the base of the so-called gp120 V3 loop (see Figures 5(b) and 5(c)). Co-receptor binding triggers further conformational changes that result in major changes in the conformation of gp41. As a result, the gp41 N-terminal fusion peptide inserts into the target cell membrane (Figure 5(d)), and the subsequent refolding of the ectodomain of the trimeric gp41 to form a stable six-helix bundle (Figure 5(e)) is thought to bring the viral and cellular membranes into close apposition (Figure 5(f)) and trigger fusion of the viral and target membranes (Figure 5(g)). HIV is generally assumed to fuse with the PM of target cells, but some reports indicate that fusion may occur in endosomal compartments, particularly following the direct cell–cell transfer of viruses (Melikyan, 2014). Although HIV fusion is triggered by receptor engagement and not pH changes, the mechanism is likely to be similar to that of other viruses with type I fusion proteins, in particular influenza viruses.

Reverse Transcription

Regardless of where fusion occurs, it releases the HIV core into the cytoplasm of the target cell where reverse transcription and uncoating take place (see Figure 2). The temporal order and subcellular localization of reverse transcription and uncoating remain controversial (Hilditch and Towers, 2014). Recent evidence suggests that reverse transcription is initiated before significant uncoating occurs, though both processes may proceed in parallel. The resulting nucleic acid/protein complexes are referred to as reverse transcription complex (RTC), and subsequently the pre-integration complex (PIC), but the precise composition of the RTC and PIC remain elusive. For example, some studies suggest that most CA is lost during uncoating, but other researchers have reported functions for CA in nuclear import and even in the chromosomal integration of HIV DNA, suggesting that at least some CA remains associated with the RTC/PIC (Fassati, 2012).

Reverse transcription of the viral RNA genome takes place in the RTC. It is primed by a single molecule of tRNA^{Lys3} bound to the primer binding site of the viral RNA. RT mediates RNA-dependent DNA transcription of the positive-sense RNA genome, digestion of the transcribed RNA, and DNA-dependent DNA polymerization to produce double-stranded HIV DNA. HIV RT lacks proof-reading capacity, resulting in an error rate of 3×10^{-5} /nucleotide/replication cycle. Thus, reverse transcription is responsible for the high mutation rate seen in HIV, which in part leads to its genetic diversity (the nucleotide sequence of HIV Env differs by 20–30% between isolates and by up to 10% between viruses from one individual). In addition, recombination between different HIV

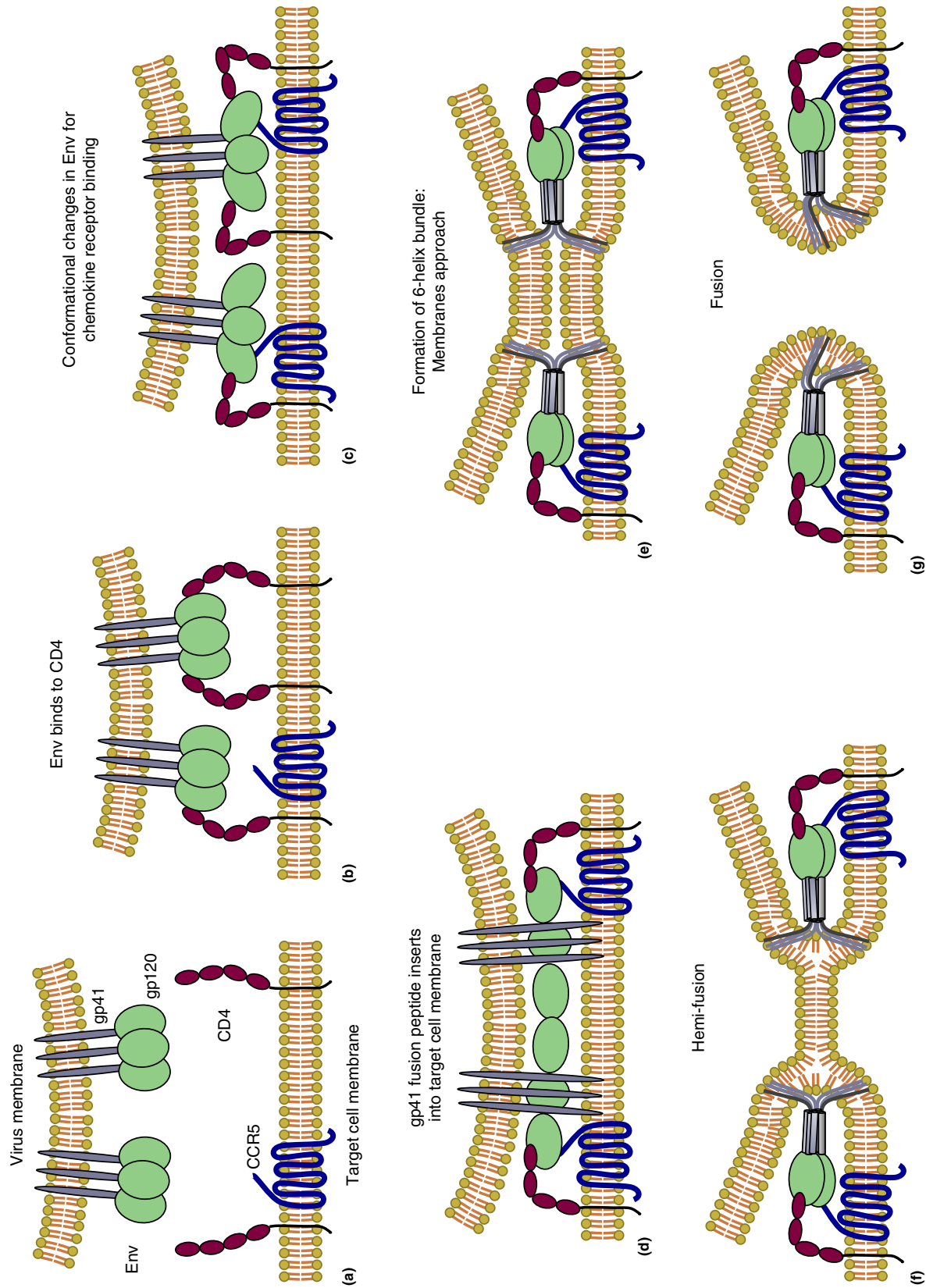


Figure 5 Cartoon showing the HIV envelope-induced membrane fusion process. The components of the fusion reaction shown in (a) comprise the Env trimers in the viral membrane, and the HIV receptor CD4 and a chemokine receptor (usually CCR5 or CXCR4) in the target cell membrane. Binding to CD4 (b) leads to conformational changes in Env that then allow binding to the co-receptor (c). This leads to further conformational rearrangements, with gp120 opening out and allowing the fusion peptide at the N-terminus of gp41 to insert into the target cell membrane (d). The extended conformation of gp41 is inherently unstable and the protein folds back on itself into a hairpin structure (the 6-helix bundle), pulling the viral and target cell membranes into close apposition (e) and leading to the formation of the metastable hemi-fusion intermediate (f) and full fusion of the viral and cell membranes (g).

genomes can occur when cells are simultaneously infected with several virions. The plethora of antigenically distinct HIV quasi-species found within HIV-infected individuals presents a challenge to the immune system, which consequently is unable to mount an effective response to the infection.

Nuclear Entry and Integration

Following reverse transcription, double-stranded genomic HIV DNA, associated with the PIC, is delivered into the nucleus of cells. In contrast to other retroviruses, HIV can infect post-mitotic as well as dividing cells. Therefore, the virus must be able to deliver its genome across the intact nuclear envelope, presumably through nuclear pore complexes. Nuclear localization signals in several PIC-associated viral proteins have been suggested to mediate nuclear import, but the extent to which any of these proteins is involved is unclear. The ability of HIV to infect nondividing cells makes HIV-1-based lentiviral vectors particularly useful for gene transfer.

Once inside the nucleus, HIV DNA is integrated into the host cell genome. The viral IN clips several nucleotides at the 3' ends of the double-stranded HIV DNA, cleaves the host cell chromosome, and joins the 3' ends of the viral and cellular DNA. Remaining gaps are filled by cellular DNA repair enzymes.

Integrated HIV DNA can serve as a template for the production of mRNA and the synthesis of viral proteins, but in some cells it is not transcribed, and the infected cells remain in a latent state. Latency represents a major obstacle to the cure of HIV, as cells that do not synthesize viral proteins are resistant to drugs and host immune defenses. The mechanisms by which latent HIV is reactivated currently remain unknown.

Synthesis of New Viral Components

In productively infected cells, integrated proviral HIV DNA is transcribed by the cellular RNA polymerase II. Transcription is initiated at the 5' long terminal repeat (LTR), and promoted by a TATA box located approximately 25 base pairs (bp) upstream of the transcription start site. Subsequently, the viral accessory protein Tat binds to the trans-activation response element (TAR) of the nascent mRNA molecule, recruits the human positive-transcriptional elongation factor b (P-TEF-b), and thereby further enhances transcription. Gag and GagPol are translated from unspliced HIV mRNA, whereas Env, Vif, Vpr, and Vpr are synthesized from partially spliced RNAs. Although unspliced and partially spliced mRNAs are normally retained in the nucleus, HIV uses the accessory protein Rev to overcome this barrier to replication. Rev, as well as the accessory proteins Tat and Nef, is translated from fully spliced mRNAs early in infection. Multiple copies of Rev bind to a Rev responsive element (RRE), which is located in the Env coding sequence of all unspliced or partially spliced mRNAs, and interact with the nuclear export factor CRM1 to transport unspliced and partially spliced HIV RNAs into the cytoplasm. In addition to its function as a template for the translation of Gag and GagPol, unspliced viral mRNA also forms the genomic RNA that is incorporated into infectious particles.

Cell Killing

HIV infection culminates in the depletion of CD4⁺ T cells, one of the defining characteristics of AIDS. However, numerous studies have shown that the death of infected cells alone cannot account for the rapid progression and extent of T cell depletion in patients. Moreover, HIV-infection also leads to a loss of CD4-negative (CD4⁻) cells, including B cells and neurons. These observations indicate that HIV kills both infected, as well as uninfected 'bystander' cells, though many of the underlying mechanisms remain controversial.

HIV-infected cells can die from direct consequences of virus replication, or from cytotoxicity mediated by other immune cells. A direct consequence of HIV replication is the integration of viral DNA into the host cell genome. In activated CD4⁺ T cells, this step of virus replication may trigger a cellular DNA damage response and culminate in cell death by apoptosis (Cooper *et al.*, 2013). The accumulation of HIV RNA and proteins in infected cells may also cause cytotoxicity.

Cytotoxic T lymphocytes (CTL) induce apoptotic cell death in HIV-infected cells that present MHC-I-associated viral peptides at their surface. To avoid CTL killing, many viruses down-regulate MHC-I molecules. However, cells that do not present MHC-I at their surface are in turn targeted by NK cells. To circumvent both the CTL and NK cell immune defense mechanisms, HIV only down-modulates subsets of MHC-I molecules (Cohen *et al.*, 1999).

NK cells, macrophages and neutrophils express Fc receptors at their surface, which enables them to recognize and kill antibody-coated HIV-infected cells in a process called antibody-dependent cellular cytotoxicity (ADCC). ADCC depends on the presence of HIV-specific antibodies, which bind to cell surface Env on infected cells, or to Env proteins incorporated into cell-associated virus particles, and recruit the cytotoxic immune cells via their constant antibody regions (Fc regions).

The depletion of uninfected bystander cells in HIV-infected patients is mainly attributed to the induction of apoptosis. HIV induces a state of chronic immune activation, which is characterized by an upregulation of genes of the TNF family, including the pro-apoptotic receptors Fas and TRAIL (Poonia *et al.*, 2009). Moreover, abortive infection of quiescent CD4⁺ T cells has been suggested to result in intracellular accumulation of unintegrated viral DNA, which may be detected by a cellular DNA sensor and lead to cell death by pyroptosis (Monroe *et al.*, 2014). The interplay of these multiple cell killing mechanisms can explain the rapid loss of CD4⁺ as well as CD4⁻ cells during HIV infection, and can account for diverse symptoms associated with AIDS.

Host Defences against HIV

Humoral and Cellular Immunity

The evolutionary pressure imposed by viruses and other pathogens has given rise to immune defences that counteract viral replication. *In vivo*, HIV infection triggers both innate and adaptive immune responses, which are further subdivided into cellular and humoral immunity.

Macrophages and dendritic cells are mediators of the innate cellular defense. As professional antigen presenting cells (APC), these cells phagocytose HIV-infected cells, present foreign antigen to other immune cells, and secrete pro-inflammatory cytokines that attract cells of the adaptive immune response. Following their recruitment, B and T cells are activated by APCs. B cells mature and secrete specific antibodies, which are the main mediators of the humoral immune response. However, due to the viral immune evasion strategies described above, antibodies are unable to control HIV infection *in vivo*. CTL recognize and eliminate cells that present MHC-I-associated HIV antigens on their surface. These cells appear to more efficiently counteract HIV, and are assumed to mediate the temporary drop in viral load following the acute phase of infection but ultimately also fail to clear infected individuals of HIV.

Restriction Factors

In addition to cellular and humoral immunity, protective responses to viral pathogens can involve cellular proteins that directly counteract virus replication within infected cells. These so-called restriction factors have a number of common characteristics: (1) Their expression is usually enhanced, or induced, by type I interferons (IFN), which are upregulated following infection with intracellular pathogens such as viruses. (2) They directly inhibit viral replication, rather than mediating their effects through activation of a broad and nonspecific immune response, though some restriction factors function as pattern recognition receptors and activate inflammatory cellular signaling pathways in addition to their direct effects on viral replication. (3) Since restriction factors can present major barriers to viral replication, successful viruses often encode antagonists that operate via direct protein-protein interactions. (4) Mutations in the restriction factors that abolish their interactions with the viral antagonists offer an advantage to the host, and are thus selected for during evolution. Viruses in turn adapt their antagonists to maintain activity against the modified restriction factors. As a result, restriction factors show signatures of rapid evolution. HIV replication is inhibited by a number of restriction factors, including TRIM5 α , SAMHD1, APOBEC3G, and Tetherin (Figure 6; Malim and Bieniasz, 2012).

TRIM5 α

TRIM5 α is an E3 ubiquitin ligase that acts on HIV-1 early in the replication cycle. It assembles into a lattice surrounding the incoming capsid, and is thought to disassemble the viral core before reverse transcription can occur. Ubiquitin ligase activity is necessary for its antiviral activity, and during restriction TRIM5 α is auto-ubiquitinated and ultimately degraded via proteasomes. TRIM5 α may also inhibit replication after reverse transcription, and has recently been shown to activate the NF κ B pathway to promote the expression of pro-inflammatory cytokines.

SAMHD1

SAMHD1 is an IFN-inducible restriction factor that has been suggested to interfere with HIV reverse transcription by

depleting cellular dNTP pools, though other mechanisms may exist. Although HIV-2 and SIV from macaques (SIV mac) can antagonize SAMHD1 through their Vpx protein, HIV-1 does not encode a Vpx analog and is unable to antagonize SAMHD1.

APOBEC3G

APOBEC3G also targets HIV reverse transcription, but it does so through a mechanism that is distinct to that employed by SAMHD1. APOBEC3G is a cytidine deaminase that catalyzes cytidine-to-uridine hypermutations during reverse transcription. Cytidine-to-uridine mutations in the negative-strand lead to guanine-to-adenine mutations in the proviral DNA. Thus, HIV DNA isolated from AIDS patients shows high levels of guanine-to-adenine changes. These mutations frequently lead to premature stop codons or amino acid changes that prevent subsequent steps of HIV replication. APOBEC3G has also been suggested to inhibit reverse transcription and integration by mechanisms independent of its deaminase activity. In virus-producing cells, APOBEC3G is packaged into HIV particles via RNA-dependent interactions with the NC domain of Gag, but it exerts its effects in newly infected cells. The HIV accessory protein Vif, together with the cellular co-factor CBF β , antagonizes APOBEC3G by recruiting the E3 ubiquitin ligase complex elongin B/elongin C/RBX/Cullin5, which mediates APOBEC3G ubiquitination and proteasomal degradation.

Tetherin

Tetherin (or Bst-2, HM1.24, CD317) impedes HIV replication late in the virus replication cycle. The 20 kDa glycoprotein has an unusual topology, in that it has both an N-terminal transmembrane domain and a C-terminal glycosylphosphatidylinositol (GPI) anchor. During virus assembly, one end of the dimeric protein (predominantly the C-terminus) is incorporated into budding particles, while the other remains inserted in the host cell PM. Consequently, fully formed HIV particles remain tethered to infected cells, and are subsequently internalized and targeted for lysosomal degradation. Tetherin may also inhibit HIV replication by activating the pro-inflammatory NF κ B signaling pathway. Since Tetherin targets the viral membrane, an immutable structure inherent to all enveloped viruses, many viruses that assemble at the PM are susceptible to Tetherin restriction. Indeed, Tetherin has been shown to restrict all tested retroviruses, the filoviruses Ebola and Marburg, the arenaviruses Lassa and Machupo, the paramyxoviruses Sendai and Nipah, the rhabdovirus vesicular stomatitis virus, as well as the herpes viruses HSV-1 and Kaposi's sarcoma-associated herpesvirus, which gain their membranes on intracellular compartments. Most, if not all, inhibited viruses encode proteins that antagonise Tetherin. In HIV-1 this is done by the accessory protein Vpu, which interacts with Tetherin and promotes its lysosomal degradation (Neil, 2013).

Drugs Interfering with HIV Replication

HIV relies on cellular processes throughout its replication cycle. However, some steps of replication are dependent on virally encoded proteins, and interfering with these can limit

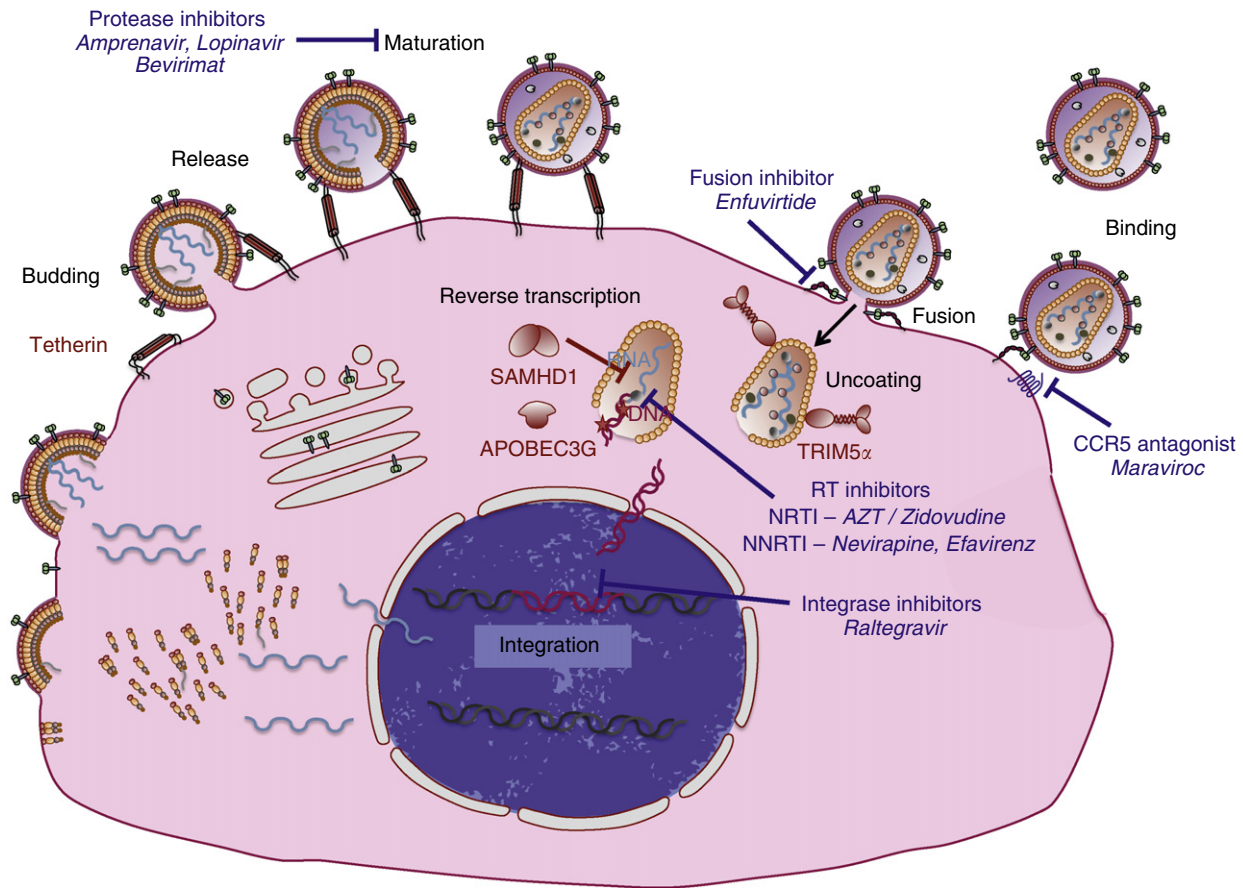


Figure 6 Cartoon illustrating points in the HIV replication cycle where restriction factors or antiretroviral drugs exert their effects. Diagram of the HIV replication cycle (see [Figure 1](#)). The site of action of four restriction factors, TRIM5 α , SAMHD1, APOBEC3G and Tetherin, are indicated in dark red. TRIM5 α acts on the incoming HIV core and interferes with the uncoating steps, SAMHD1 has been suggested to deplete the cellular dNTP pools required for reverse transcription, and APOBEC3G induces mutations in the viral DNA (indicated as stars). Tetherin retains budded virus particles on the surface of infected cells and leads to their internalization and degradation. The sites of action of various classes of antiretroviral drugs are indicated in purple, with selected examples. NNRTI non-nucleoside reverse transcriptase inhibitor; NRTI, nucleoside reverse transcriptase inhibitor.

replication without disrupting essential cellular processes. A number of antiretroviral drugs are now available that target these virus-specific steps. These include viral entry, reverse transcription, DNA integration, and maturation (protease) inhibitors ([Figure 6](#)).

RT inhibitors were the first anti-HIV drugs to receive approval by the US Food and Drug Administration (FDA). They are subdivided into two classes: nucleoside reverse transcriptase inhibitors (NRTI) and non-nucleoside reverse transcriptase inhibitors (NNRTI). NRTI consist of synthetic nucleoside analogs, which lack the 3'-OH group necessary for transcriptional elongation. When incorporated into the growing viral DNA, they terminate reverse transcription. NNRTI bind to RT outside the nucleotide binding site, induce conformational changes, and thereby inhibit RT activity. With currently 13 and six FDA-approved NRTI and NNRTI, respectively, RT inhibitors are the most abundant anti-HIV drugs.

Any of these drugs potentially inhibit HIV replication *in vitro*. However, when administered singly to patients, drug-resistant strains rapidly emerge due to the high mutation rate of the

virus. To limit the emergence of drug resistance, highly active antiretroviral therapy (HAART) nowadays consists of a combination of up to four antiretroviral drugs. Though this therapy can decrease the viral load to undetectable levels and prevent progression to AIDS, adverse effects are frequent and may negatively impact on patient compliance. Significantly, HIV usually reappears rapidly if HAART is stopped due to its ability to exist in latent form within subpopulations of CD4⁺ T cells and in other reservoirs such as macrophages.

Conclusion

Since its discovery in the early 1980s, HIV has become one of the best characterized viruses. Many aspects of its evolution, structure, mode of replication, interaction with its human host and mechanisms of pathogenesis are understood in detail. Nevertheless, and despite its relative genetic simplicity, there is much that remains to be understood. In particular, full understanding of how HIV evades cellular defenses, how the virus is assembled and transmitted within infected individuals,

and whether it is possible to develop effective immunological antiviral protection remain major challenges. Increasing our understanding of how HIV and related viruses engage with target cells, and elucidating the spectrum of molecular interactions involved in these engagements, will continue to be areas of active research for the foreseeable future.

Acknowledgments

The authors thank S. Kilcher, S. Lawrence, and M. Mazzon for assistance in the preparation of this article, and the UK Medical Research Council for core funding to the MRC-UCL University Unit (Grant ref. MC_U122665002). SG was supported by a Boehringer Ingelheim Fonds fellowship. Some sections of this article are based on elements of SG's PhD thesis "Cellular analysis of HIV transmission from primary macrophages", University College London, 2014.

See also: Cellular Immunology: Cell–Cell Interactions and the Immune System: Lymphocyte–Endothelial Interactions. Cellular Immunology: Transcriptional Basis of B and T cell Lineages and Memory Cells: NF-kappaB and the Immune System. Intracellular Infectiology: Cell Processes: Cellular Invasion by Bacterial Pathogens. Intracellular Infectiology: Infectious Agents: Cell Biology of Virus Infection; Virus Factories and Mini-Organelles Generated for Virus Replication. Nucleic Acid Synthesis/Breakdown: Nucleic Acid Technology: Viral Nucleic Acids. Protein Synthesis/Degradation: Protein Degradation – Pathological Aspects: Inhibitors of HIV Protease

References

- Azevedo, C., Burton, A., Ruiz-Mateos, E., Marsh, M., Saiardi, A., 2009. Inositol pyrophosphate mediated pyrophosphorylation of AP3B1 regulates HIV-1 Gag release. *Proceedings of the National Academy of Sciences of the United States of America* 106, 21161–21166.
- Brass, A.L., Dykxhoorn, D.M., Benita, Y., *et al.*, 2008. Identification of host proteins required for HIV infection through a functional genomic screen. *Science* 319, 921–926.
- Cohen, G.B., Gandhi, R.T., Davis, D.M., *et al.*, 1999. The selective downregulation of class I major histocompatibility complex proteins by HIV-1 protects HIV-infected cells from NK cells. *Immunity* 10, 661–671.
- Cooper, A., García, M., Petrovas, C., *et al.*, 2013. HIV-1 causes CD4 cell death through DNA-dependent protein kinase during viral integration. *Nature* 498, 376–379.
- Faria, N.R., Rambaut, A., Suchard, M.A., *et al.*, 2014. HIV epidemiology. The early spread and epidemic ignition of HIV-1 in human populations. *Science* 346, 56–61.
- Fassati, A., 2012. Multiple roles of the capsid protein in the early steps of HIV-1 infection. *Virus Research* 170, 15–24.
- Fisher, R.D., Chung, H.-Y., Zhai, Q., *et al.*, 2007. Structural and biochemical studies of ALIX/AIP1 and its role in retrovirus budding. *Cell* 128, 841–852.
- Giese, S., Marsh, M., 2013. Cellular trafficking mechanisms in the assembly and release of HIV. *Advances in HIV-1 Assembly and Release*. New York, NY: Springer.
- Hilditch, L., Towers, G.J., 2014. A model for cofactor use during HIV-1 reverse transcription and nuclear entry. *Current Opinion in Virology* 4, 32–36.
- Jouvenet, N., Bieniasz, P.D., Simon, S.M., 2008. Imaging the biogenesis of individual HIV-1 virions in live cells. *Nature* 454, 236–240.
- Malim, M.H., Bieniasz, P.D., 2012. HIV restriction factors and mechanisms of evasion. *Cold Spring Harbor Perspectives in Medicine* 2, a006940.
- Melikyan, G.B., 2014. HIV entry: A game of hide-and-fuse? *Current Opinion in Virology* 4, 1–7.
- Merk, A., Subramaniam, S., 2013. HIV-1 envelope glycoprotein structure. *Current Opinion in Structural Biology* 23, 268–276.
- Monroe, K.M., Yang, Z., Johnson, J.R., *et al.*, 2014. IFI16 DNA sensor is required for death of lymphoid CD4 T cells abortively infected with HIV. *Science* 343, 428–432.
- Murooka, T.T., Deruaz, M., Marangoni, F., *et al.*, 2012. HIV-infected T cells are migratory vehicles for viral dissemination. *Nature* 490, 283–287.
- Neil, S.J.D., 2013. The antiviral activities of tetherin. *Current Topics in Microbiology and Immunology* 371, 67–104.
- Ono, A., Ablan, S.D., Lockett, S.J., Nagashima, K., Freed, E.O., 2004. Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. *Proceedings of the National Academy of Sciences of the United States of America* 101, 14889–14894.
- Ono, A., Freed, E.O., 2001. Plasma membrane rafts play a critical role in HIV-1 assembly and release. *Proceedings of the National Academy of Sciences of the United States of America* 98, 13925–13930.
- Poonia, B., Pauza, C.D., Salvato, M.S., 2009. Role of the Fas/FasL pathway in HIV or SIV disease. *Retrovirology* 6, 91.
- Sattentau, Q.J., 2010. Cell-to-Cell spread of retroviruses. *Viruses* 2, 1306–1321.
- Sundquist, W.I., Kräusslich, H.-G., 2012. HIV-1 assembly, budding, and maturation. *Cold Spring Harbor Perspectives in Medicine* 2, a006924.
- Zhu, P., Liu, J., Bess, J., *et al.*, 2006. Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* 441, 847–852.
- Zimmerman, C., Klein, K.C., Kiser, P.K., *et al.*, 2002. Identification of a host protein essential for assembly of immature HIV-1 capsids. *Nature* 415, 88–92.

Further Reading

- Coffin, J.M., Hughes, S.H., Varmus, H.E (Eds.), 1997. *Retroviruses*. Cold Spring Harbor Laboratory Press. ISBN-10: 0–87969–571–4. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK19376/> (accessed 01.01.14).
- Shilts, R., 1987. *And the Band Played On – Politics, People and the AIDS Epidemic*. New York, NY: St Martin's. Available at: <http://us.macmillan.com/book.aspxisbn=9780312374631> (accessed 12.12.14).

Prions

G Legname and KE Pischke, Scuola Internazionale Superiore di Studi Avanzati, Trieste, Italy

© 2016 Elsevier Inc. All rights reserved.

Introduction

Prion diseases, or transmissible spongiform encephalopathies (TSEs), are a group of closely related, unfailingly fatal neurodegenerative disorders characterized by motor and cognitive impairments, extensive brain damage, and neuronal dysfunction. These diseases include Creutzfeldt–Jakob disease (CJD), Gerstmann–Sträussler–Scheinker syndrome (GSS), fatal familial insomnia (FFI), and kuru in humans; scrapie in sheep and goats; bovine spongiform encephalopathy (BSE) in cattle; and chronic wasting disease (CWD) in cervids.

Stanley B. Prusiner first proposed the then heretic idea that the infectious pathogen responsible for these diseases could be transmitted and propagated without the need for nucleic acids but rather with protein forming the essential pathogenic component. He referred to this idea as the ‘protein-only’ hypothesis. He also proposed the term ‘prion’ to define this unusual proteinaceous infectious agent that causes prion diseases. The prion, denoted PrP^{Sc} for scrapie prion protein (PrP), is the abnormal, disease-causing form of the physiologically expressed cellular prion protein, PrP^C.

History of the Prion Discovery

Scrapie, a naturally occurring prion disease affecting sheep and goats has been recognized in Europe for more than 250 years (McGowan, 1922). The name scrapie originates from a clinical sign of the disease, as afflicted animals will compulsively scrape off their fleeces against objects. Scrapie was shown to be transmissible by inoculation between sheep (Cuille and Chelle, 1936). Much later, the scrapie agent of sheep was successfully transmitted to mice and rats (Chandler, 1961; Chandler and Fisher, 1963).

In the 1950s, there arose an epidemic of a neurodegenerative disease among the indigenous Fore population of Papua New Guinea. This neurodegenerative disease, known as kuru, was characterized by progressive ataxia and was thought to be linked to symbolic cannibalistic activities practiced among the people. It was recognized that kuru (and later CJD) histopathologically resembled scrapie. This led to the idea that these diseases were somehow related and they too may be transmitted by intracerebral inoculation (Klatzo *et al.*, 1959). This was shown with two seminal experiments wherein brain homogenates from diseased humans were inoculated into chimpanzees, resulting in disease in the chimpanzees which mimicked the human diseases (Gajdusek *et al.*, 1966).

This work showed that the neurodegenerative diseases kuru and CJD (and later work showed that FFI and GSS also) were transmissible. Still unclear was the nature of the transmissible agent. Although many initially thought a ‘slow virus’ was responsible, this was challenged by the failure to convincingly demonstrate such a virus, lack of an immunological response by the body to infection and especially by evidence presented

by Alper *et al.* (1967) suggesting that the transmissible agent was resistant to treatments known to inactivate nucleic acid.

Building on Alper’s work, Griffith was the first to propose a model illustrating that the scrapie agent could be a protein devoid of nucleic acids (Griffith, 1967). This idea was based on the strong experimental evidence that no conventional pathogen could resist such high doses of ionizing or UV radiation, but a polypeptide could. The model challenged the central dogma of biology by suggesting that proteins, and not exclusively DNA or RNA, could store and propagate biological information.

The protein-only model was quite ahead of its time, and it was not until 1982 that Stanley B. Prusiner worked to forward Griffith’s model. Prusiner collected all the evidence amassed thus far and argued that since several procedures that inactivate nucleic acids would not prevent scrapie infectivity, the agent must be a protein. He coined the term ‘prion’ defined as ‘small proteinaceous infectious particles that resist inactivation by procedures which modify nucleic acids’ and thereby clearly distinguishing the infectious pathogen from viruses or viroids (Prusiner, 1982). Prusiner and colleagues found that a novel protein in the brains of animals inoculated with the scrapie agent was the major component of infective materials (Bolton *et al.*, 1982).

Soon after, this novel infectious protein was found to be encoded within the host’s own genome and named PrP (Oesch *et al.*, 1985). The discovery that the infectious agent of scrapie was a host-encoded protein helped to explain why the body had no immunological response to the disease and how nucleic acids were not needed for the disease’s propagation, but still unclear was how PrP could cause disease. Evidence from biophysical studies indicated that PrP associated with disease possessed a different folding pattern, or tertiary structure, than the physiological PrP present in most mammalian cells. This has led to the concept that the change of PrP structure can impart the characteristics of an infectious agent. Researchers today are still struggling to fully understand how a misfolded protein is able to cause disease.

Despite all of the evidence put forward, there still lingered doubts. To dispel such doubts the Prusiner group developed infectious prions outside of mammalian cells via recombinant proteins produced in bacteria. These synthetic mammalian prions caused disease in animal models, providing the final evidence of the prion puzzle (Legname *et al.*, 2004).

In 1997 Prusiner was awarded the Nobel Prize in physiology or medicine for his discovery of prions as a new biological principle of infection.

Prion Diseases

Prion diseases, or TSEs, are rare, infectious, lethal neurodegenerative diseases affecting humans as well as agricultural, captive and free-ranging animals. TSEs can be divided into

three etiological categories: sporadic, inherited, and acquired. Importantly, even the sporadic and genetic forms of the disease are experimentally transmissible.

One of several challenging aspects of prion diseases is the heterogeneity of clinical symptoms and histopathological phenotypes. This can make diagnosing these diseases a very difficult task. Currently, there are no proven treatments for prion diseases.

Human Prion Diseases

Human prion diseases are classified into CJD, GSS, FFI, and kuru. Most human prion diseases, 85%, are sporadic, meaning they have an unknown origin, which is often attributed to a spontaneous PrP^C to PrP^{Sc} conversion or to somatic mutation. Approximately 10–15% of human TSEs are inherited through mutations in the prion gene (PRNP) resulting in familial prion disease (familial CJD, GSS, and FFI). More than 30 disease-causing mutations in the open reading frame of PRNP have been characterized. Finally, less than 1% of human prion diseases are acquired by means of an infectious route. These rare instances, also referred to as acquired prion diseases, manifest in the form of variant CJD (vCJD) through the consumption of BSE contaminated beef; iatrogenic CJD (iCJD) by infected surgical procedures; or kuru through ritualistic cannibalism of the diseased (Collins *et al.*, 2004; Glatzel *et al.*, 2003; Mastrianni, 2010).

CJD

The neuropathological characteristics of CJD include neuronal loss, spongiform change, gliosis, and in some cases amyloid plaques.

Sporadic CJD

The commonest form, sporadic CJD (sCJD), results from a yet undefined, possibly stochastic event, in the central nervous system that induces PrP to misfold into a disease-causing form. It occurs with a worldwide incidence of between one and two deaths per one million people annually (Budka *et al.*, 2011). Clinical presentations of the disease vary, but in roughly 40% of patients the first symptom is cognitive dysfunction (often memory problems, language impairment and/or executive dysfunction) (Takada and Geschwind, 2013).

Variant CJD

The first case of variant CJD (vCJD) was recognized in the United Kingdom in 1995 (Will *et al.*, 1996). Since then, 228 cases of vCJD have been reported worldwide in 11 different countries, with the majority of cases found in the United Kingdom (Diack *et al.*, 2014). vCJD has been linked to the ingestion of BSE-infected meat, as it has been established that BSE and vCJD are the same prion strain (Bruce *et al.*, 1997). One prominent feature to distinguish between vCJD and sCJD is by analyzing PrP extracted from lymphoid organs such as tonsils. Likely due to its oral route of infection, vCJD shows PrP^{Sc} in the lymphoid organs, while sCJD does not (Frosch *et al.*, 2004).

The clinical presentation of vCJD differs from other forms of CJD. Patients with vCJD are generally younger, with a

median age of death of 28 years (as opposed to 68 years with sCJD) and duration of the disease is longer with a median of 14 months (compared to 7 months in sCJD) (Heath *et al.*, 2010; Pocchiari *et al.*, 2004). Patients usually suffer psychiatric symptoms early on in the disease course, commonly in the form of depression, anxiety, or apathy. Neurological symptoms arise as the disease progresses, including difficulty walking, unsteadiness, and involuntary movements; culminating at the complete loss of movement and death (Diack *et al.*, 2014).

Iatrogenic CJD

Although extremely rare, iatrogenic CJD (iCJD) results from the disease being accidentally transmitted through surgical procedures. More than 450 cases of iCJD have been recorded. The main sources have been contaminated growth hormone and dura mater grafts from cadavers with undiagnosed CJD; but a few cases have derived from neurosurgical instrument contamination, corneal grafts, gonadotrophic hormone, and secondary infection with vCJD transmitted by transfusion of blood products.

The infectious prion agent survives decontamination procedures that are standard procedure to destroy viruses and bacteria. It is hoped that the implementation of new sterilization techniques, exclusive use of synthetic growth hormone, combined with improved recognition of possibly infected individuals will end the outbreak of iCJD (Brown *et al.*, 2012).

It is interesting to note that iCJD, depending on the route of infection has markedly different incubation times. When iCJD occurs through direct transmission of CJD to the central nervous system (via corneal transplants, dura mater grafts, and contaminated surgical instruments during brain surgery), referred to as intracerebral transmission, the incubation period is 1.5–3.8 years. On the other hand, with iCJD transmitted through contaminated human growth hormone, a type of peripheral transmission, the incubation time is much longer, approximately 15 years.

Familial CJD

A small fraction of all CJD cases is due to point mutations on PRNP, resulting in familial CJD (fCJD). Experiments with fCJD showed for the first time that familial prion diseases are transmissible. This gives familial prion diseases the unique traits of being both inherited and transmissible (Gambetti *et al.*, 2003).

GSS

Gerstmann–Sträussler–Scheinker (GSS) is a slowly progressive hereditary autosomal dominant disease. It typically presents as a progressive ataxic disorder with late onset cognitive impairment. Onset of disease ranges from the fourth to the sixth decades with disease duration averaging 5 years. There is, however, substantial phenotypic variability among families (Arata *et al.*, 2006). GSS was the first human prion disease in which a mutation in PrP was described (Liberski, 2012).

FFI

Fatal familial insomnia (FFI) is a rare autosomal dominant inherited prion disease. It usually presents with progressive, severe insomnia, dysautonomia, and endocrine disturbances.

FFI is almost always caused by PrP^C mutations, but rare, sporadic cases have also been described, termed sporadic fatal insomnia (sFI). Disease onset usually occurs in the fifth and sixth decade of life and the mean duration of the disease is 18 months. Spongiform degeneration is rarely found in FFI patients, but neuronal loss and gliosis, specifically within the thalamus, is common (Capellari *et al.*, 2011; Pocchiari *et al.*, 2004).

Nearly 100 cases of FFI in almost 40 families have been reported from Italy, Germany, Austria, Spain, United Kingdom, France, Finland, the United States, Australia, Japan, China, and Morocco (Baldin *et al.*, 2009).

Animal Prion Diseases

TSEs in animals include scrapie of sheep and goats; BSE of cattle; CWD of cervids; transmissible mink encephalopathy (TME), feline spongiform encephalopathy, and ungulate spongiform encephalopathy (Kirkwood *et al.*, 1990; Marsh and Hadlow, 1992; Wells *et al.*, 1987; Williams and Young, 1980; Wyatt *et al.*, 1991). In addition to scrapie, the two most prevalent animal TSEs are BSE and CWD.

BSE

BSE also known as 'Mad Cow disease' is a prion disease that affects cattle. Clinical signs of BSE may include gait abnormalities (ataxia), tremors, aggressive behavior, apprehension, and hyperreactivity to stimuli.

BSE first appeared in the United Kingdom in the 1980s and quickly grew to epidemic proportions among cattle. The outbreak was fed by the practice of using meat and bone meal, which was then inadvertently contaminated with pathogenic prions, as a dietary supplement for the cattle. BSE has affected more than 180 000 cattle worldwide. A ban on the use of meat and bone meal in cattle feed has ultimately resulted in a decline of the epidemic (Ducrot *et al.*, 2008). The worst was realized when it became apparent that BSE had been transmitted to humans in the form of vCJD (Bruce *et al.*, 1997).

CWD

This TSE affects captive and free-ranging deer, elk, and moose in North America and South Korea. The origin of CWD is unknown. The potential for CWD transmission to humans is low and has yet to be documented (Beringue *et al.*, 2008; Sigurdson, 2008). Moreover, transgenic mice expressing the human, cattle, or sheep PrP^C did not develop TSE when inoculated with CWD prions (Tamguney *et al.*, 2006). CWD does, however, seem to spread readily among cervids and long-term effects of CWD on cervid populations and ecosystems could be very harmful, especially if the disease continues to spread.

The Prion Protein

The cellular form of the prion protein, PrP^C, has about 90% sequence similarity across mammals. It is expressed in almost all tissues, but the highest levels are found in the central

nervous system. Here, PrP^C is especially associated with synaptic membranes.

In humans, the prion gene (*PRNP*) encodes a 253 amino acid long precursor protein in a single, uninterrupted open reading frame. After translocation across the endoplasmic reticulum, the 22-amino acid N-terminal and 23-amino acid C-terminal signal peptide sequences are removed. This leaves a 209 amino acid long functional protein that is attached to the outer membrane surface of the cell by a glycosylphosphatidylinositol (GPI) anchor.

The precise function of PrP^C within the cellular context eludes researchers to this day. Its considerable conservation among mammals and high expression throughout the body suggest an important physiological role for PrP^C, however, PrP^C-deficient mice at first glance develop normally. Several possible functions for PrP^C have been proposed including signaling cascades, neuronal survival, apoptosis, oxidative stress, cell adhesion, cell differentiation, etc. (Grassmann *et al.*, 2013). It is known, however, that PrP^C has a high affinity for metals such as copper, zinc, and manganese (Garnett and Viles, 2003; Hornshaw *et al.*, 1995).

Structure

Much evidence now reinforces the idea that prions are principally or completely composed of an isoform of the normal host-encoded protein PrP^C. This abnormal isoform is designated PrP^{Sc} (Prusiner, 1991). PrP^{Sc} originates from PrP^C by a posttranslational mechanism (Caughey and Raymond, 1991). No consistent differences in amino acid sequence or posttranslational modifications have been found between PrP^C and PrP^{Sc} (Stahl *et al.*, 1993). It is thought that the differences between PrP^C and PrP^{Sc} only involve a conformational change and that PrP^{Sc} acts as a template for the conversion of PrP^C to PrP^{Sc}. The mechanism of conversion between the two is yet unknown, but discerning any conformational differences may help to better elucidate this mechanism.

The structure of the natively folded PrP^C was first available in 1996 and has been solved many times by X-ray crystallography and NMR since (Riek *et al.*, 1996; Surawicz and Apostol, 2011). These studies show PrP^C to have a mostly α -helical structure. This has been confirmed to be structurally equivalent to PrP^C found *in vivo* (Hornemann *et al.*, 2004).

PrP^C is attached to the outer cell membrane leaflet by a GPI anchor posttranslationally added. Additional posttranslational modifications include two variably occupied N-linked glycosylation sites at Asn181 and Asn197 as well as a stabilizing disulfide bridge between Cys 179 and Cys214.

Although the molecular structure of PrP^C is well characterized, that of PrP^{Sc} is still not well understood. Its tendency to aggregate and general insolubility have hindered attempts to elucidate its structure at high resolution. Generally, PrP^{Sc} is described as oligomeric in nature and abundant in β -sheet structures. Resolving the structure of PrP^{Sc} is of utmost importance in prion research today as the structural change from PrP^C to PrP^{Sc} is at the core of transmission, infection, and pathogenesis of prions (Requena and Wille, 2014).

Conversion and Replication

Structure plays a critical role in the prion saga, where PrP^C undergoes a structural change in conformation to convert it from the normal cellular protein found throughout the body to a misfolded form linked to disease, its infectious counterpart PrP^{Sc}. According to the 'protein-only' model, PrP^{Sc} itself is the infectious prion agent responsible for the transmissibility of prion disorders (Prusiner, 1998).

Two mechanisms have been proposed to explain how PrP^C is converted to PrP^{Sc}. In the heterodimer refolding model, (also known as the template assistance model), PrP^{Sc} is a monomer and thermodynamically favored, but kinetically inaccessible, conformer of PrP^C (Prusiner, 1991). In the rare event that a PrP^{Sc} molecule is spontaneously formed or delivered from an outside source, PrP^{Sc} would serve as a template for PrP^C by the formation of a heterodimer, misfolding PrP^C into the thermodynamically more stable PrP^{Sc}. This mechanism is possible, but there is no evidence for a stable PrP^{Sc} monomer, in fact, PrP^{Sc} is thought to be highly oligomeric (Silveira *et al.*, 2005).

The nucleation-polymerization model argues that PrP^{Sc} is less stable than PrP^C and that the limiting step to conversion is the formation of a PrP^{Sc} nucleus. Though the formation of this nucleus is not favored, once formed it is a stabilized aggregate. These aggregates of PrP^{Sc} then propagate by incorporating PrP^C monomers, which can easily add to the structure by adopting the PrP^{Sc} conformation. The rate-limiting nucleation step could be responsible for the lag phase in spontaneous conversion (Jarrett and Lansbury, 1993).

Transmission

A central characteristic of prion diseases is their transmissibility to other animals. However, the transmissibility of prions is not a straightforward process, in fact there seem to be some factors that hinder the ability of PrP^{Sc} to convert PrP^C, known as transmission barrier.

When transmitting an infectious prion, PrP^{Sc}, from one host to another identical host, where the amino acid sequences between PrP^{Sc} and PrP^C are identical, prion diseases are highly infectious. In this case the transmission barrier is very low or nonexistent. However, transmission of prions between species can be incomplete or with long incubation times, indicating some type of transmission barrier. This barrier is frequently linked to differences in the primary sequence of the PrP (Bruce *et al.*, 2002). Interestingly, a similar transmission barrier as is seen between species can be seen within the same species, often to a smaller extent, if the individuals express different PrP genotypes (Bossers *et al.*, 2000).

The transmission barrier has implications to human health. The barrier between humans and cattle with BSE is traversable, as evidenced by the many cases of vCJD. Fortunately, so far the transmission barrier seems to hold between humans and sheep/cervids as there are no known cases of humans developing a prion disease from scrapie or CWD.

The transmission barrier, however, is not a completely fixed characteristic. While initially there may exist a transmission barrier, several passages of PrP^{Sc} to identical hosts allows for the prions to stabilize. This process is known as prion

adaptation and it leads to characteristic, reproducible disease features of the prion disease (Thackray *et al.*, 2011). The process of adaptation leads to the generation of new prion strains.

Strains

Prion strains are molecules that share the same primary sequence of amino acids, but with differing incubation times or neuropathological/biochemical properties. Strains are believed to result from conformational differences in PrP^{Sc} (Telling *et al.*, 1996).

The existence of prion strains was first shown by the transmission of two isolates from scrapie-infected sheep into goats. It was observed that the goats showed two marked behavioral pathologies, a drowsy syndrome and a scratching syndrome, and that these behaviors were conserved upon subsequent passages in goats (Pattison and Millson, 1961). Later, five prion scrapie strains were isolated following inoculation into mice. These strains could be reliably recognized by histological parameters (degree of vacuolation and relative distribution of the damage within the brain) (Fraser and Dickinson, 1973). Additional studies have identified many prion strains with persisting characteristics after serial transmission in rodents (Solforosi *et al.*, 2013).

It was hard to imagine at the time that all of this diversity could be encoded simply within the protein structure of PrP^{Sc}. It was widely thought that the existence of strains was convincing evidence that some nucleic acid, which could mutate, must be involved. Even several genes within the host genome were suspected of influencing prion strains; but this was disproven when prion strains were propagated *in vitro* via non-genetic mechanisms (Bessen *et al.*, 1995).

Direct evidence for the molecular basis of prion strains was long wanting and many looked at prion strains as a key obstacle to the protein-only hypothesis. Ultimately, it was shown that hamsters infected with two strains of TME resulted in PrP^{Sc} molecules with distinct electrophoretic mobility (revealed as a specific pattern of bands on a Western Blot) and degree of resistance to protease digestion (Bessen and Marsh, 1994). The observed diversity is thought to arise from varying three-dimensional structures resulting in exposure of different combinations of proteinase-K cleavage sites found on the PrP (Parchi *et al.*, 2000). Unique banding patterns have also been described for protease-resistant PrP originating from vCJD and FFI (Collinge *et al.*, 1996; Telling *et al.*, 1996).

Prion strains are defined by their distinct prion disease phenotype (incubation time to disease onset, histopathological lesion profiles, and brain areas targeted) that persists upon serial transmission. Molecularly, strains have been described by their band pattern on a Western Blot after protease digestion and degree of protease digestion. In addition, pathological features are instrumental for fine characterization of closely related prion disorders.

Neurotoxicity

Perhaps the greatest mystery so far in the prion field is the mechanism by which PrP^{Sc} causes disease. TSEs are fatal

because they result in neuronal loss, extensive spongiform degeneration, and widespread synaptic damage, but the role of PrP^{Sc} in these processes is uncertain.

It is generally thought that PrP^{Sc} is the toxic species of prion diseases, responsible for the death of neurons. This is at first glance a reasonable assumption as PrP^{Sc} is the infectious agent able to transmit TSEs and its accumulation in the brain is generally seen as a hallmark of the diseases. However, much evidence has surfaced in favor of the 'dissociation' between the infectious agent and the toxic species (Halliday *et al.*, 2014).

In 1993, it was shown that PrP^{Sc} was harmless when inoculated into brains of mice devoid of PrP^C (PrP knockout (KO) mice) (Bueler *et al.*, 1993). This not only showed that PrP^C is essential for the propagation of PrP^{Sc}, but also that PrP^{Sc} itself is not neurotoxic. Another group used these resistant PrP KO mice and surgically grafted onto their brains neural tissue overexpressing PrP^C. When these chimeric mice were intracerebrally inoculated with PrP^{Sc} the grafted area accumulated high levels of PrP^{Sc} and showed grave histopathological characteristics typical of the disease. It was seen that a considerable amount of graft-derived PrP^{Sc} migrated to the host brain, but surprisingly no pathological changes were detected in the PrP KO tissue. This showed that brain tissue lacking PrP^C is not damaged by exogenous PrP^{Sc} (Brandner *et al.*, 1996).

It was also discovered that by genetically 'turning off' the production of PrP^C by neurons in mice already suffering from a prion infection is neuroprotective. By depleting endogenous neuronal PrP^C in already diseased mice the scientists were able to reverse early pathological changes, prevent neuronal loss and stop clinical progression of the disease. Interestingly, as PrP^{Sc} was still present in non-neuronal cells of the brain, PrP^{Sc} continued to accumulate outside neurons (to levels normally associated with terminally ill animals), but without the neurotoxic effects on the PrP^C-free neurons (Mallucci *et al.*, 2003). Similarly, by preventing PrP from attaching to the neuronal cell surface by removing the GPI anchor inhibits neurotoxicity, though once again PrP^{Sc} accumulates extraneuronally (Chesebro *et al.*, 2005).

These findings support the idea that PrP^{Sc} is not directly toxic to neurons, but that there is another mechanism involved. Currently, much work in the prion field is focusing on trying to understand the underlying cause(s) of neurotoxicity in prion diseases.

Conclusion

Prion diseases, although very rare, represent a new biological principle of infection and in fact, much recent research is indicating that the prion mechanism of disease may not be limited to prion diseases. Many neurodegenerative diseases, such as Alzheimer's disease, Parkinsons disease, and amyotrophic lateral sclerosis, may share with prion diseases an analogous mechanism of conversion, replication, and propagation (Prusiner, 2012). Consequently, self-propagating proteins such as prions will likely continue to be a source of scientific interest for some time.

See also: Organelles: Structure and Function: Extracellular Vesicles

References

- Alper, T., Cramp, W.A., Haig, D.A., Clarke, M.C., 1967. Does the agent of scrapie replicate without nucleic acid? *Nature* 214, 764–766.
- Arata, H., Takashima, H., Hirano, R., *et al.*, 2006. Early clinical signs and imaging findings in Gerstmann-Straussler-Scheinker syndrome (Pro102Leu). *Neurology* 66, 1672–1678.
- Baldin, E., Capellari, S., Provini, F., *et al.*, 2009. A case of fatal familial insomnia in Africa. *Journal of Neurology* 256, 1778–1779.
- Beringue, V., Vilotte, J.L., Laude, H., 2008. Prion agent diversity and species barrier. *Veterinary Research* 39, 47.
- Bessen, R.A., Kocisko, D.A., Raymond, G.J., *et al.*, 1995. Non-genetic propagation of strain-specific properties of scrapie prion protein. *Nature* 375, 698–700.
- Bessen, R.A., Marsh, R.F., 1994. Distinct PrP properties suggest the molecular basis of strain variation in transmissible mink encephalopathy. *Journal of Virology* 68, 7859–7868.
- Bolton, D.C., Mckinley, M.P., Prusiner, S.B., 1982. Identification of a protein that purifies with the scrapie prion. *Science* 218, 1309–1311.
- Bossers, A., De Vries, R., Smits, M.A., 2000. Susceptibility of sheep for scrapie as assessed by *in vitro* conversion of nine naturally occurring variants of PrP. *Journal of Virology* 74, 1407–1414.
- Brandner, S., Isenmann, S., Raeber, A., *et al.*, 1996. Normal host prion protein necessary for scrapie-induced neurotoxicity. *Nature* 379, 339–343.
- Brown, P., Brandel, J.P., Sato, T., *et al.*, 2012. Iatrogenic Creutzfeldt-Jakob disease, final assessment. *Emerging Infectious Diseases* journal 18, 901–907.
- Bruce, M.E., Boyle, A., Cousens, S., *et al.*, 2002. Strain characterization of natural sheep scrapie and comparison with BSE. *Journal of General Virology* 83, 695–704.
- Bruce, M.E., Will, R.G., Ironside, J.W., *et al.*, 1997. Transmissions to mice indicate that 'new variant' CJD is caused by the BSE agent. *Nature* 389, 498–501.
- Budka, H., Head, M.W., Ironside, J.W., *et al.*, 2011. Sporadic Creutzfeldt-Jakob Disease. *Neurodegeneration: The Molecular Pathology of Dementia and Movement Disorders*. Oxford: Wiley-Blackwell.
- Bueler, H., Aguzzi, A., Sailer, A., *et al.*, 1993. Mice devoid of PrP are resistant to scrapie. *Cell* 73, 1339–1347.
- Capellari, S., Strammiello, R., Saverioni, D., Kretzschmar, H., Parchi, P., 2011. Genetic Creutzfeldt-Jakob disease and fatal familial insomnia: Insights into phenotypic variability and disease pathogenesis. *Acta Neuropathologica* 121, 21–37.
- Caughey, B., Raymond, G.J., 1991. The scrapie-associated form of PrP is made from a cell surface precursor that is both protease- and phospholipase-sensitive. *Journal of Biological Chemistry* 266, 18217–18223.
- Chandler, R.L., 1961. Encephalopathy in mice produced by inoculation with scrapie brain material. *Lancet* 1, 1378–1379.
- Chandler, R.L., Fisher, J., 1963. Experimental transmission of scrapie to rats. *Lancet* 2, 1165.
- Chesebro, B., Trifilo, M., Race, R., *et al.*, 2005. Anchorless prion protein results in infectious amyloid disease without clinical scrapie. *Science* 308, 1435–1439.
- Collinge, J., Sidle, K.C., Meads, J., Ironside, J., Hill, A.F., 1996. Molecular analysis of prion strain variation and the aetiology of 'new variant' CJD. *Nature* 383, 685–690.
- Collins, S.J., Lawson, V.A., Masters, C.L., 2004. Transmissible spongiform encephalopathies. *Lancet* 363, 51–61.
- Cuille, J., Chelle, P.-L., 1936. Pathologie animale: La maladie dite tremblante du mouton est-elle inoculable? *Comptes Rendus de l'Académie des Sciences* 203, 1552–1554.
- Diack, A.B., Head, M.W., McCutcheon, S., *et al.*, 2014. Variant CJD: 18 years of research and surveillance. *Prion* 8, 286–295.
- Ducrot, C., Arnold, M., De Koeijer, A., Heim, D., Calavas, D., 2008. Review on the epidemiology and dynamics of BSE epidemics. *Veterinary Research* 39, 15.
- Fraser, H., Dickinson, A.G., 1973. Scrapie in mice. Agent-strain differences in the distribution and intensity of grey matter vacuolation. *Journal of Comparative Pathology* 83, 29–40.
- Frosh, A., Smith, L.C., Jackson, C.J., *et al.*, 2004. Analysis of 2000 consecutive UK tonsillectomy specimens for disease-related prion protein. *Lancet* 364, 1260–1262.
- Gajdusek, D.C., Gibbs, C.J., Alpers, M., 1966. Experimental transmission of a Kuru-like syndrome to chimpanzees. *Nature* 209, 794–796.

- Gambetti, P., Kong, Q., Zou, W., Parchi, P., Chen, S.G., 2003. Sporadic and familial CJD: Classification and characterisation. *British Medical Bulletin* 66, 213–239.
- Garnett, A.P., Viles, J.H., 2003. Copper binding to the octarepeats of the prion protein. Affinity, specificity, folding, and cooperativity: Insights from circular dichroism. *Journal of Biological Chemistry* 278, 6795–6802.
- Glatzel, M., Ott, P.M., Linder, T., *et al.*, 2003. Human prion diseases: Epidemiology and integrated risk assessment. *Lancet Neurology* 2, 757–763.
- Grassmann, A., Wolf, H., Hofmann, J., Graham, J., Vorberg, I., 2013. Cellular aspects of prion replication *in vitro*. *Viruses* 5, 374–405.
- Griffith, J.S., 1967. Self-replication and scrapie. *Nature* 215, 1043–1044.
- Halliday, M., Radford, H., Mallucci, G.R., 2014. Prions: Generation and spread versus neurotoxicity. *Journal of Biological Chemistry* 289, 19862–19868.
- Heath, C.A., Cooper, S.A., Murray, K., *et al.*, 2010. Validation of diagnostic criteria for variant Creutzfeldt-Jakob disease. *Annals of Neurology* 67, 761–770.
- Hornemann, S., Schorn, C., Wuthrich, K., 2004. NMR structure of the bovine prion protein isolated from healthy calf brains. *EMBO Reports* 5, 1159–1164.
- Hornshaw, M.P., McDermott, J.R., Candy, J.M., Lakey, J.H., 1995. Copper binding to the N-terminal tandem repeat region of mammalian and avian prion protein: Structural studies using synthetic peptides. *Biochemical and Biophysical Research Communications* 214, 993–999.
- Jarrett, J.T., Lansbury Jr., P.T., 1993. Seeding "one-dimensional crystallization" of amyloid: A pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* 73, 1055–1058.
- Kirkwood, J.K., Wells, G.A., Wilesmith, J.W., Cunningham, A.A., Jackson, S.I., 1990. Spongiform encephalopathy in an arabian oryx (*Oryx leucoryx*) and a greater kudu (*Tragelaphus strepsiceros*). *Veterinary Record* 127, 418–420.
- Klatzo, I., Gajdusek, D.C., Zigas, V., 1959. Pathology of Kuru. *Laboratory Investigation* 8, 799–847.
- Legname, G., Baskakov, I.V., Nguyen, H.O., *et al.*, 2004. Synthetic mammalian prions. *Science* 305, 673–676.
- Liberski, P.P., 2012. Gerstmann-Straussler-Scheinker disease. *Advances in Experimental Medicine and Biology* 724, 128–137.
- Mallucci, G., Dickinson, A., Linehan, J., *et al.*, 2003. Depleting neuronal PrP in prion infection prevents disease and reverses spongiosis. *Science* 302, 871–874.
- Marsh, R.F., Hadlow, W.J., 1992. Transmissible mink encephalopathy. *Revue scientifique et technique* 11, 539–550.
- Mastrianni, J.A., 2010. The genetics of prion diseases. *Genetics in Medicine* 12, 187–195.
- McGowan, J.P., 1922. Scrapie in sheep. *Scottish Journal of Agriculture* 5, 365–375.
- Oesch, B., Westaway, D., Walchli, M., *et al.*, 1985. A cellular gene encodes scrapie PrP 27-30 protein. *Cell* 40, 735–746.
- Parchi, P., Zou, W., Wang, W., *et al.*, 2000. Genetic influence on the structural variations of the abnormal prion protein. *Proceedings of the National Academy of Sciences of the United States of America* 97, 10168–10172.
- Pattison, I.H., Millson, G.C., 1961. Scrapie produced experimentally in goats with special reference to the clinical syndrome. *Journal of Comparative Pathology* 71, 101–109.
- Pocchiari, M., Puopolo, M., Croes, E.A., *et al.*, 2004. Predictors of survival in sporadic Creutzfeldt-Jakob disease and other human transmissible spongiform encephalopathies. *Brain* 127, 2348–2359.
- Prusiner, S.B., 1982. Novel proteinaceous infectious particles cause scrapie. *Science* 216, 136–144.
- Prusiner, S.B., 1991. Molecular biology of prion diseases. *Science* 252, 1515–1522.
- Prusiner, S.B., 1998. Prions. *Proceedings of the National Academy of Sciences of the United States of America* 95, 13363–13383.
- Prusiner, S.B., 2012. Cell biology. A unifying role for prions in neurodegenerative diseases. *Science* 336, 1511–1513.
- Requena, J.R., Wille, H., 2014. The structure of the infectious prion protein: Experimental data and molecular models. *Prion* 8, 60–66.
- Riek, R., Hornemann, S., Wider, G., *et al.*, 1996. NMR structure of the mouse prion protein domain PrP(121–231). *Nature* 382, 180–182.
- Sigurdson, C.J., 2008. A prion disease of cervids: Chronic wasting disease. *Veterinary Research* 39, 41.
- Silveira, J.R., Raymond, G.J., Hughson, A.G., *et al.*, 2005. The most infectious prion protein particles. *Nature* 437, 257–261.
- Solforsini, L., Milani, M., Mancini, N., Clementi, M., Burioni, R., 2013. A closer look at prion strains: Characterization and important implications. *Prion* 7, 99–108.
- Stahl, N., Baldwin, M.A., Teplow, D.B., *et al.*, 1993. Structural studies of the scrapie prion protein using mass spectrometry and amino acid sequencing. *Biochemistry* 32, 1991–2002.
- Surewicz, W.K., Apostol, M.I., 2011. Prion protein and its conformational conversion: A structural perspective. *Topics in Current Chemistry* 305, 135–167.
- Takada, L.T., Geschwind, M.D., 2013. Prion diseases. *Seminars Neurology* 33, 348–356.
- Tamguney, G., Giles, K., Bouzamondo-Bernstein, E., *et al.*, 2006. Transmission of elk and deer prions to transgenic mice. *Journal of Virology* 80, 9104–9114.
- Telling, G.C., Parchi, P., Dearmond, S.J., *et al.*, 1996. Evidence for the conformation of the pathologic isoform of the prion protein enciphering and propagating prion diversity. *Science* 274, 2079–2082.
- Thackray, A.M., Hopkins, L., Lockey, R., Spiropoulos, J., Bujdoso, R., 2011. Emergence of multiple prion strains from single isolates of ovine scrapie. *Journal of General Virology* 92, 1482–1491.
- Wells, G.A., Scott, A.C., Johnson, C.T., *et al.*, 1987. A novel progressive spongiform encephalopathy in cattle. *Veterinary Record* 121, 419–420.
- Will, R.G., Ironside, J.W., Zeidler, M., *et al.*, 1996. A new variant of Creutzfeldt-Jakob disease in the UK. *Lancet* 6, 921–925.
- Williams, E.S., Young, S., 1980. Chronic wasting disease of captive mule deer: A spongiform encephalopathy. *Journal of Wildlife Diseases* 16, 89–98.
- Wyatt, J.M., Pearson, G.R., Smerdon, T.N., *et al.*, 1991. Naturally occurring scrapie-like spongiform encephalopathy in five domestic cats. *Veterinary Record* 129, 233–236.

INTRACELLULAR INFECTIOLOGY: CELL PROCESSES

Cellular Responses to Infections in *Caenorhabditis elegans*

J Sun, Washington State University, Spokane, WA, USA

A Aballay, Duke University, Durham, NC, USA

V Singh, Indian Institute of Science, Bangalore, India

© 2016 Elsevier Inc. All rights reserved.

Caenorhabditis elegans as a Model System to Study Responses to Pathogens

Caenorhabditis elegans is a powerful experimental organism for a number of traits that facilitate genetic and genomic analysis, including the hermaphroditic lifestyle, short 2–3 week lifespan, and small genome, which offers an ideal compromise between complexity and tractability. Unlike bacteria, yeast, or other single-celled organisms, *C. elegans* has muscles, nerves, sexual organs, and guts, in spite of its one millimeter long body. It reacts to touch, can recognize odors, and exhibits different behaviors. In nature, *C. elegans* is found in the soil and decaying organic matter. To date, more than 40 microbes have been shown to be pathogenic to *C. elegans*, including bacteria, fungi, and viruses. The major routes of infection are through the intestine and the epidermis. Because the nematode lacks professional immune cells, intestinal and epidermal epithelial cells serve as the primary defense against pathogens, as they are in direct contact with different microbes. Despite lacking adaptive immunity, *C. elegans* can mount protective responses to pathogen infection by avoiding certain pathogens or by triggering evolutionarily conserved signaling pathways. Activation of these cellular signaling pathways can lead to production of immune effectors such as lectins, lysozymes, lipases, and

antimicrobial peptides, which act directly to fight off the invading pathogens. The defensive response that results in activated signaling pathways and induced immune effectors is pathogen-specific, indicating that the nematode is able to distinguish different infecting microbes.

Overview of *C. elegans* Pathogens

Caenorhabditis elegans can be infected in its natural environment and in the laboratory setting by a large number of pathogens, including Gram negative bacteria, Gram positive bacteria, fungi, and viruses. These pathogens may vary in their mode of infection, site of replication and host responses they induce. **Figure 1** illustrates site and mode of infection utilized by a number of the most highly studied pathogens in *C. elegans*.

The majority of bacteria and fungi remain extracellular. Gram negative pathogens, such as *Salmonella enterica* and *Pseudomonas aeruginosa*, colonize the digestive system of *C. elegans* causing lethal infections (Aballay *et al.*, 2000; Tan *et al.*, 1999). *Pseudomonas aeruginosa* can kill *C. elegans* using toxins in a matter of hours by a process termed ‘fast killing’ (Mahajan-Miklos *et al.*, 1999). It can also kill nematodes in 2–4 days by an infectious-like process termed ‘slow-killing,’ which requires bacterial

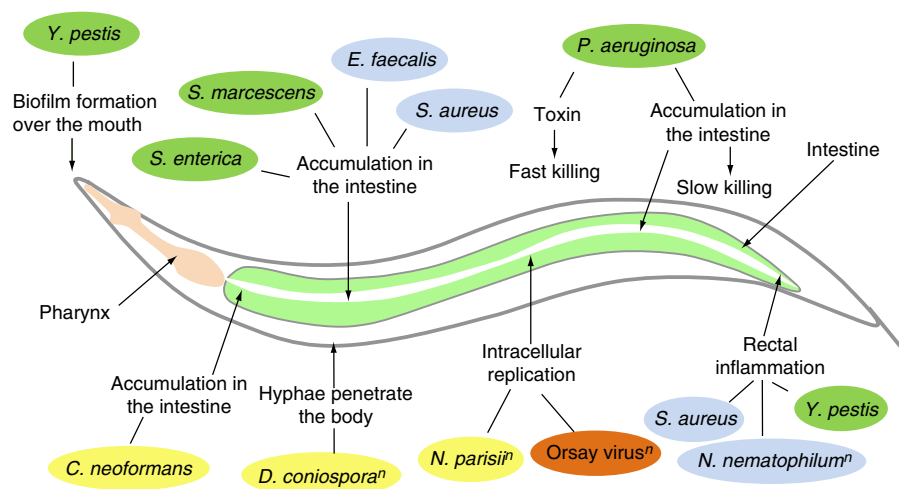


Figure 1 *Caenorhabditis elegans* pathogens colonize different tissues of the nematode. Gram negative bacteria, green; Gram positive bacteria, blue; fungi, yellow; viruses, orange. Natural pathogens of *C. elegans* are marked with superscript *n*.

replication in the intestinal lumen of the host (Tan *et al.*, 1999). Other Gram negative bacteria, such as *Microbacterium nematophilum*, and *Yersinia pestis* strains incapable of forming biofilms, also colonize the digestive tract, but cause rectal inflammation in the digestive tract (Hodgkin *et al.*, 2000; Styer *et al.*, 2005). *Yersinia pestis* strains capable of forming biofilm cover the mouth of *C. elegans* and prevent feeding, reminiscent of *Y. pestis* infection of plague carrier fleas (Darby *et al.*, 2002). Gram positive bacteria such as *Enterococcus faecalis* and *Staphylococcus aureus* also colonize the intestine (Irazoqui *et al.*, 2010; Garsin *et al.*, 2001). An interesting aspect of *S. aureus* infection is that it causes extensive loss of intestinal microvilli of the host and rectal inflammation (Irazoqui *et al.*, 2010).

Fungal pathogens can set up infection in the intestine of the nematode or the epidermis. The yeast form of the human opportunistic pathogen *Cryptococcus neoformans* accumulates in the nematode intestine and causes a lethal infection (Mylonakis *et al.*, 2002). *Drechmeria coniospora* is a natural fungal pathogen of *C. elegans* that enters through the cuticle and penetrates all body tissues (Couillault *et al.*, 2012; Jansson, 1994). *Candida albicans*, a fungus that causes nosocomial infections in humans, is pathogenic to *C. elegans* in both its hyphal and yeast forms. In turn, nematodes elicit distinct responses depending upon the site of infection (Pukkila-Worley *et al.*, 2009; Breger *et al.*, 2007).

A small number of natural pathogens can cause intracellular infections in *C. elegans*. Examples include *Nematocida parisii*, which is a microsporidian parasite related to fungi, and a natural pathogen of *C. elegans* (Troemel *et al.*, 2008; Felix and Duveau, 2012). Orsay virus and other viruses of the nodoviridae family cause intracellular invasion in the intestine of *C. elegans* and related species (Felix *et al.*, 2011; Franz *et al.*, 2014).

Pathogen Recognition by *C. elegans*

The mechanisms by which *C. elegans* identifies pathogens remain elusive. In plants and higher animals, pathogen recognition is mainly mediated by a set of PRRs that detect conserved pathogen-derived molecules, known as pathogen-associated molecular patterns (PAMPs) or MAMPs. The best characterized PRRs are a family of conserved transmembrane Toll-like receptors (TLRs) that recognize various microbial products, such as bacterial cell-wall components and flagellin, and trigger multiple defense response pathways. The genome of *C. elegans* does not encode the majority of known PRRs and lacks genes of some key components of the TLR pathways such as NF- κ B homologs and the TLR adaptor MyD88. Nematode orthologs of a single Toll-like receptor (TOL-1), TNF receptor associated factor-1 (TRF-1), Pelle and IL-1R-associated kinases (PIK-1) and inhibitor of NF- κ B (IKB-1) do not seem to play key roles in immunity. However, it is worth noting that TOL-1 has been implicated in immune signaling against *S. enterica* (Tenor and Aballay, 2008) and that TIR-1, a scaffold protein containing Toll/IL-1R (TIR) protein-protein interaction domain, is involved in antimicrobial defense against a variety of *C. elegans* pathogens (Liberati *et al.*, 2004; Couillault *et al.*, 2004). *Caenorhabditis elegans* is also capable of activating immune pathways when exposed to heat-killed *S. aureus* and *C. albicans*, suggesting that the nematode can detect pathogenic

microbes via MAMP recognition (Irazoqui *et al.*, 2010; Pukkila-Worley and Ausubel, 2012).

While MAMP/PRR interactions capable of eliciting immune responses in *C. elegans* remain elusive, emerging evidence supports the notion that the nematode can recognize pathogen attack through the detection of disturbances of cellular homeostasis. *Caenorhabditis elegans* appears to have evolved cellular surveillance systems to monitor its core cellular activities and interpret disruption of these activities as evidence of pathogen attack. The innate immunity elicited by this kind of indirect recognition is similar to effector-triggered immunity (ETI) that was originally demonstrated in plants (Jones and Dangl, 2006) and later found conserved across invertebrate and vertebrate animals (Boyer *et al.*, 2011; Kleino and Silverman, 2012). An example of ETI in *C. elegans* is its response to a virulence factor of *P. aeruginosa*, Exotoxin A (ToxA). Upon infection with *P. aeruginosa*, ToxA enters the intestinal epithelial cells, likely via endocytosis, and inhibits protein translation by modifying elongation factor 2 via ADP-ribosylation, which triggers the host to up-regulate immune defense genes (Dunbar *et al.*, 2012; McEwan *et al.*, 2012). Besides protein translation, disruption of many other cellular functions such as mitochondrial respiration, ubiquitin-proteasome system activity, or actin cytoskeleton and microtubule dynamics, results in activation of detoxification and immune responses in *C. elegans* (Melo and Ruvkun, 2012; Bakowski *et al.*, 2014). Pore-forming toxins (PFTs), the single largest class of proteinaceous bacterial virulence factors that perforate host cell membranes, can also trigger ETI in the nematode, although the precise molecular mechanisms remain unknown (Engelmann and Pujol, 2010). It has been suggested that the host cells can sense disruption of cell membrane, potassium efflux or calcium influx and respond with antimicrobial signaling and membrane repair pathways (Los *et al.*, 2013; Los *et al.*, 2011). ETI offers *C. elegans* an effective and versatile strategy to detect and distinguish pathogens from surrounding microbes. On the one hand, this allows the nematode to combat a wide variety of microbes without evolving the full repertoire of PRRs; on the other hand, it remains to be addressed how the nematode can distinguish infecting microbes if different pathogens disrupt the same cellular processes, as immune responses are not only pathogen-shared but also pathogen-specific (Wong *et al.*, 2007; Shivers *et al.*, 2008; Alper *et al.*, 2007; Irazoqui *et al.*, 2010; Schulenburg *et al.*, 2008).

G protein-coupled receptors (GPCRs) may play a direct or indirect role in pathogen recognition in *C. elegans*. The nematode uses neuronal GPCR-mediated signaling to sense bacterial compounds and to respond by avoiding pathogens and activate immune pathways (Sun *et al.*, 2011; Styer *et al.*, 2008; Pradel *et al.*, 2007; Reddy *et al.*, 2009). Furthermore, a GPCR that functions in the epidermis activates innate immunity by recognizing an endogenous ligand. This ligand potentially acts as a danger molecule to alert the organism of fungal infections (Zugasti *et al.*, 2014).

Immune Effectors

Upon infection, a number of immune effectors are produced in *C. elegans* and confer the nematode protective immunity

Table 1 Common immune effectors in *Caenorhabditis elegans*

Immune effector	Description	Gene family	References
Antibacterial factor (ABF)	Small cationic peptides similar to human defensins ● Antimicrobial activity	<i>abf-1</i> to <i>abf-6</i>	Kato <i>et al.</i> (2002)
Caenacin	Putative antimicrobial peptides, structurally related to neuropeptide-like proteins antimicrobial activity ● Antimicrobial activity	<i>cnc-1</i> to <i>cnc-11</i>	Zugasti and Ewbank (2009)
Caenopore	A family of proteins containing saposin-domain, similar to protozoan amoebapores, mammalian NK-lysin, and granulysin ● Antibacterial activity ● Membrane permeability ● Pore-forming activity	28 spp genes encoding 33 different caenopores	Roeder <i>et al.</i> (2010); Banyai and Pathy (1998); Hoeckendorf and Leippe (2012); Hoeckendorf <i>et al.</i> (2012)
C-type lectin	Proteins that contain one or more C-type lectin-like domains (CTLD) and recognize complex carbohydrates on cells and tissues ● Antimicrobial activity ● Contribute pathogen specificity of immunity	278 genes encoding CTLD proteins	Drickamer and Dodd (1999); Takeuchi <i>et al.</i> (2008); Schulenburg <i>et al.</i> (2008); Simonsen <i>et al.</i> (2012)
CUB domain protein	The CUB domain is a structural motif of approximately 110 residues. They resemble immunoglobulin and occur in large gene clusters ● Antimicrobial activity	More than 50 genes encoding CUB domain proteins	Bork and Beckmann (1993); Bolz <i>et al.</i> (2010)
Lysozyme	Lysozymes are glycosyl hydrolase enzymes present in microbes, insects, plants, and animals ● Degrade bacterial cell wall ● Antimicrobial activity	10 protist type lysozyme encoding genes (<i>lys-1</i> to <i>lys-10</i>) and 5 invertebrate type lysozyme encoding genes (<i>ilys-1</i> to <i>ilys-5</i>)	Schulenburg and Boehnisch (2008); Mallo <i>et al.</i> (2002); Jensen <i>et al.</i> (2010); Marsh <i>et al.</i> (2011)
Neuropeptide-like protein (NLP)	The NLPs are a diverse group of neuropeptides that have little similarity among them ● Acts as neurotransmitters ● Antimicrobial activity	42 genes encoding NLPs	Nathoo <i>et al.</i> (2001)

against microbial pathogens (Table 1). These effectors either directly inhibit the growth of invading pathogens or counteract their effects on host cells.

Antibacterial Factors (ABFs)

ABF-1 to ABF-6 in *C. elegans* are small cationic peptides similar to human defensins. ABF-2 has antimicrobial activity against Gram negative bacteria, Gram positive bacteria, and yeast (Kato *et al.*, 2002). ABF-1 and ABF-2 are expressed in the pharynx of *C. elegans*, which is one of the first organs that gets in direct contact with pathogens, and their expression requires TOL-1, and cell death receptor CED-1.

Caenacins

These are putative antimicrobial peptides, structurally related to neuropeptide-like proteins. Some of these genes are up-regulated in the epidermis following infection with the fungal pathogen *D. coniospora*. Up-regulation of caenacins in the epidermis requires neurally produced TGF β homolog DBL-1 (Zugasti and Ewbank, 2009). In addition, overexpression of caenacins enhances *C. elegans* immunity against *D. coniospora*.

Caenopores

Caenopores are a family of proteins containing a saposin domain in *C. elegans*, similar to protozoan amoebapores,

mammalian NK-lysin and granulysin. *Caenorhabditis elegans* has a group of 28 *spp* genes that encode 33 different caenopores (Roeder *et al.*, 2010). Antibacterial, membrane-permeabilizing, and pore-forming activities have been demonstrated for SPP-1, SPP-3, SPP-5, and SPP-12 (Banyai and Pathy, 1998; Roeder *et al.*, 2010; Hoeckendorf *et al.*, 2012; Hoeckendorf and Leippe, 2012). SPP-1 protects *C. elegans* against *S. enterica* and *P. aeruginosa*. Down-regulation of SPP-1 by *P. aeruginosa* is under the control of the FOXO transcription factor DAF-16.

C-type Lectins

C-type lectins are a group of proteins that contain one or more C-type lectin-like domains (CTLD) and that recognize complex carbohydrates on cells and tissues (Drickamer and Dodd, 1999; Takeuchi *et al.*, 2008). 278 genes encoding CTLD proteins are present in the *C. elegans* genome (Schulenburg *et al.*, 2008). More than 60 CTLD proteins are induced during pathogen infection. Distinct expression profiles of CTLD genes during different infections indicate that these proteins might contribute to pathogen specificity of *C. elegans* immune responses. For instance, RNAi knockdown of *clec-17*, *clec-60* and *clec-86* caused enhanced susceptibility to *M. nematophilum* (O'Rourke *et al.*, 2006). Also, knockdown of *clec-65* or *clec-70* resulted in enhanced susceptibility to *Escherichia coli* LF82 or *S. aureus*, respectively (Simonsen *et al.*, 2011; Irazoqui *et al.*, 2010). Simultaneous overexpression of *clec-60* and *clec-61*, or *clec-70* and *clec-71* led to increased resistance to *S. aureus* but decreased resistance to *P. aeruginosa* (Irazoqui *et al.*, 2010). C-type lectins often bind carbohydrates in a Ca^{2+} -dependent manner. However, a recent study revealed that CLEC-39 and CLEC-49 directly recognize live *Serratia marcescens* in a Ca^{2+} -independent manner, suggesting a nonclassical recognition of *S. marcescens* by both CTLD proteins (Miltch *et al.*, 2014).

CUB Domain Proteins

More than 50 genes in the *C. elegans* genome encode proteins containing a CUB (C1r/C1s, Uegf, and Bmp1) domain, which is a structural motif of about 100 residues (Bork and Beckmann, 1993). RNAi inhibition of three of the CUB genes (F08G5.6, F20G2.5, *dct-17*) enhances *C. elegans* susceptibility to *P. aeruginosa* and RNAi inhibition of two other CUB genes (C17H12.8 and C32H11.12) enhances *C. elegans* susceptibility to *Y. pestis* infection. Interestingly, CUB genes that are usually upregulated in response to bacterial infections, are down-regulated in response to fungi such as *C. albicans* and *D. coniospora* (Simonsen *et al.*, 2012).

Lysozymes

Lysozymes are glycosyl hydrolase enzymes that degrade bacterial cell walls and that are present in both plants and animals. The *C. elegans* genome has 10 protist type lysozyme encoding genes (*lys-1* to *lys-10*) and 5 invertebrate type lysozyme encoding genes (*ihys-1* to *ihys-5*) (Schulenburg and Boehnisch, 2008). Many lysozyme genes are expressed in the intestine and upregulated in response to microbial infections.

RNAi inhibition of *lys-1* causes enhanced susceptibility to *S. aureus*, while overexpression of *lys-4* and *lys-5* increases resistance to *S. aureus* (Mallo *et al.*, 2002; Jensen *et al.*, 2010). Overexpression of *lys-4* enhances resistance to *S. marcescens* (Mallo *et al.*, 2002). *lys-7* mutants are also more susceptible to *C. neoformans* infection but resistant to *S. typhimurium* infection (Marsh *et al.*, 2011).

Neuropeptide-Like Proteins

Neuropeptides are bioactive peptides that play a role in synaptic signaling. *Caenorhabditis elegans* has 32 genes encoding neuropeptide-like proteins (NLPs) (Nathoo *et al.*, 2001). Some of the NLPs encoding genes are upregulated during microbial infection. NLP-31 has antimicrobial activity *in vitro* and NLP-29 is expressed in the intestine and hypodermis (Couillault *et al.*, 2004). Up-regulation of nlp-29 cluster (6 genes) in the hypodermis during *D. coniospora* infection requires signaling molecules such as TIR-1, PMK-1/p38 MAPK and DCAR-1/GPCR.

Immune Signaling Pathways

Detection of pathogen attack leads to activation of cellular signaling pathways that induce expression of defensive genes to limit microbial growth, destroy invading microbes, detoxify xeno/endobiotics, and repair damage. These pathways include innate immune and stress response pathways activated in various tissues during infection with bacterial or fungal pathogens (Figures 2 and 3). Recent studies demonstrated the complexity and specificity of these responses: different signaling pathways regulate expression of distinct but overlapping sets of effector genes, individual pathogens induce genes controlled by multiple pathways, and the set of genes regulated by each pathway is also pathogen-specific.

The p38 MAPK Pathway and Related Pathways

The conserved p38 MAPK pathway plays a major role in the antimicrobial response of *C. elegans*. Forward and reverse genetic studies have demonstrated that the p38 MAPK pathway is required for immunity against a variety of pathogens and xenobiotics, including Gram negative and positive bacteria, fungi, PFTs, and many types of stress such as oxidative stress and heavy metal stress (Ermolaeva and Schumacher, 2014; Pukkila-Worley and Ausubel, 2012). This pathway includes the TIR domain protein TIR-1, neuronal symmetry family member 1 (*nsy-1*), SAPK/ERK kinase 1 (*sek-1*) and p38 MAPK family member 1 (*pmk-1*) (Kim *et al.*, 2002, 2004). Signal is transduced in the form of protein phosphorylation, where NSY-1 phosphorylates SEK-1, which in turn phosphorylates PMK-1. TIR-1 and two protein kinases PKC δ and PKD act upstream of NSY-1 (Liberati *et al.*, 2004; Ziegler *et al.*, 2009; Ren *et al.*, 2009). It has also been demonstrated that *C. elegans* ATF-7, a transcription factor orthologous to mammalian ATF2, functions downstream of PMK-1 as a repressor of PMK-1-regulated genes, and undergoes a switch to an activator upon phosphorylation by PMK-1 (Shivers *et al.*, 2010).

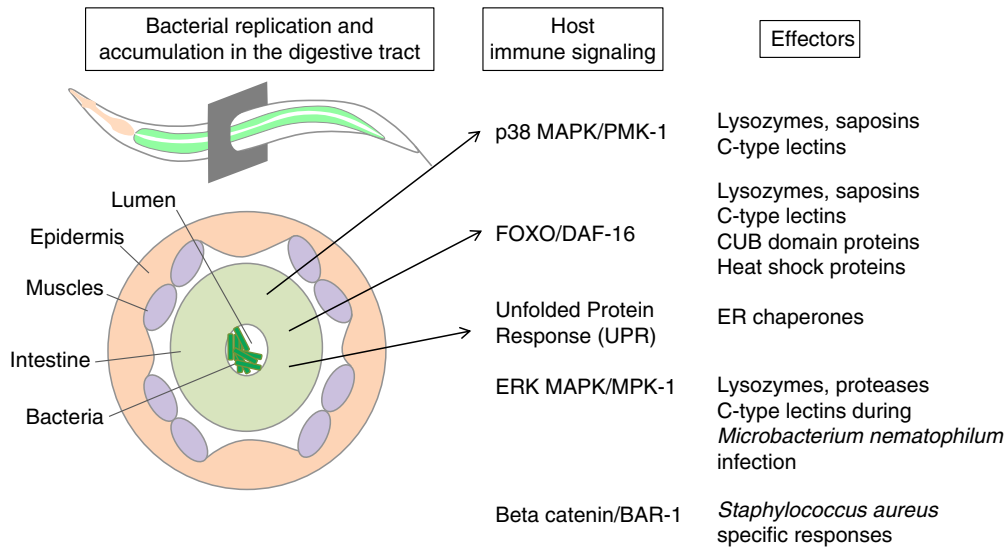


Figure 2 Cellular responses to bacterial infection in *C. elegans*. Replication and accumulation of Gram negative bacteria in the intestine leads to activation of many signaling pathways. Some of the better studied pathways include p38 MAPK, insulin signaling involving FOXO transcription factor Daf-16 as well as UPR in the intestinal cells. These produce effector molecules including antimicrobial molecules, chaperones etc. The site of activation of beta catenin or the ERK MAPK pathway is not known.

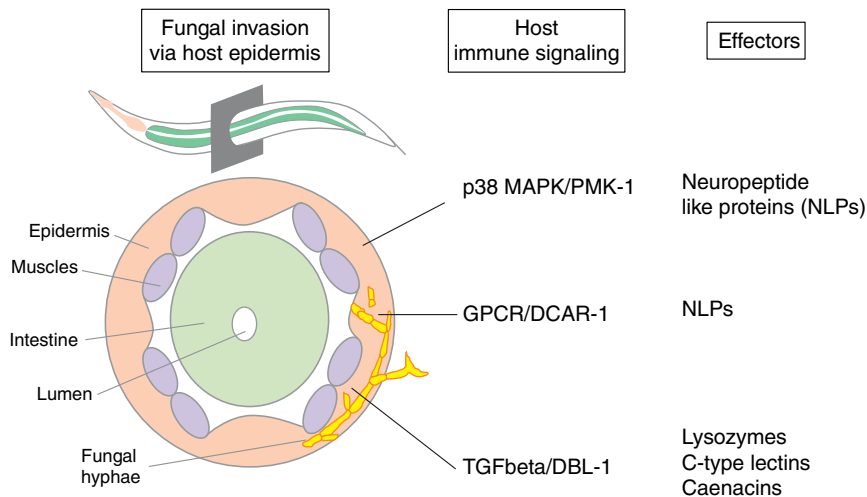


Figure 3 Cellular responses to fungal infection in *C. elegans*. Invasion of epidermis by hyphal form of fungus *Drechmeria coniospora* leads to induction of defenses in that tissue. The p38 MAPK pathway regulates antimicrobial peptides such as NLP genes in the epidermis while the TGFbeta pathway regulates anti-fungal products such as caenacins, C-type lectins and lysozymes. DCAR-1, a G protein-coupled receptor (GPCR) in the epidermis, regulates NLP gene expression partly through the PMK-1 pathway.

Activation of the p38 MAPK pathway induces expression of a set of secreted immune response genes. Several microarray studies showed that genes upregulated by PMK-1 include those encoding proteins containing CUB-like domains, C-type lectins, antibacterial peptides, lysozymes, NLPs, and caenacins. Besides p38, two other MAPKs, extracellular signal-regulated kinase (ERK) and c-Jun N-terminal kinase (JNK), are also involved in the nematode's defense to infections. Upon infection with *M. nematophilum*, the ERK pathway is activated and mediates a protective response by inducing production of C-type lectins, lysozymes, proteases, and other defense-related proteins (Nicholas and Hodgkin, 2004; O'Rourke et al., 2006).

The JNK pathway is a key regulator of the transcriptional and functional responses to bacterial PFTs (Kao et al., 2011).

The DAF-2/DAF-16 Pathway

The DAF-2/DAF-16 pathway is well known for its regulation on the lifespan of *C. elegans*, and also plays important roles in the immune response to infection by several bacteria. Activation of the DAF-2 receptor by an agonist ligand, such as the insulin-like peptide DAF-28, activates a phosphorylation cascade that results in the phosphorylation of the FOXO family

transcription factor DAF-16. Phosphorylated DAF-16 is retained in the cytoplasm. In *daf-2* mutants or in the presence of an antagonist ligand such as INS-1, un-phosphorylated DAF-16 translocates into the nucleus, where it regulates expression of a wide variety of genes (Lin *et al.*, 2001). It has been suggested that the PMK-1 and DAF-16 pathways act in parallel to promote immunity. While PMK-1 controls both basal and infection-induced expression of pathogen response genes, DAF-16 regulates a constitutively expressed response or a general stress response to microbes (Troemel *et al.*, 2006; Shivers *et al.*, 2008). This is in agreement with the role of DAF-16 in regulating the expression of many genes involved in resistance to various forms of stress, including heat, oxidative stress, hypoxia, osmotic stress, heavy metal toxicity, ultraviolet radiation, proteotoxicity, and microbial infection (Murphy and Hu, 2013). A number of antimicrobial effectors are regulated by DAF-16, including lysozymes LYS-7 and LYS-8, the saposins SPP-1, SPP-9 and SPP-12, DUF-23, and C-type lectins (Shivers *et al.*, 2008; Murphy and Hu, 2013). Most of these effectors appear to be secreted proteins that function to disrupt microbial cell membranes.

The DBL-1 Pathway

The DBL-1 pathway in *C. elegans* is homologous to the mammalian TGF β cascade, and was first described as involved in resistance to infection by *P. aeruginosa* and *S. marcescens* (Tan, 2001; Mochii *et al.*, 1999; Mallo *et al.*, 2002). The TGF β family ligand DBL-1 binds membrane-anchored heterodimeric DAF-4/SMA-6 receptor, leading to the phosphorylation and activation of the cytoplasmic signal transducer SMAD complex (SMA-2, -3 and -4); the transducer then translocates into the nucleus, where it activates gene expression. *dbl-1* mutants exhibit increased susceptibility to infection with *S. marcescens*, and the DBL-1 pathway up-regulates the expression of many defense genes, including lectins and lysozymes (Mallo *et al.*, 2002). Upon infection with fungus *D. coniospora*, *dbl-1* gene is also required for induction of *cnc-2* expression (Zugasti and Ewbank, 2009).

The Unfolded Protein Response Pathway

In the endoplasmic reticulum (ER), proteostasis surveillance is mediated by the unfolded protein response (UPR). ER is an important intracellular organelle in which newly synthesized transmembrane and secretory proteins are folded, assembled and matured. ER homeostasis is critical for ensuring proper folding and release of modified proteins to the Golgi. In response to pathogen infection, increased demands on the protein folding machinery causes stress in the ER, which activates a series of UPR pathways to restore ER homeostasis. Activation of UPR in *C. elegans* is required for defense against a variety of pathogens including *B. thuringiensis*, *S. enterica* and *P. aeruginosa* (Bischof *et al.*, 2008; Haskins *et al.*, 2008; Richardson *et al.*, 2010). UPR was found to be regulated by the p38, JNK and MAPK pathways in response to PFTs (Bischof *et al.*, 2008; Kao *et al.*, 2011). Interestingly, activation of UPR is also required by p38 MAP kinase PMK-1 mediated defense against *P. aeruginosa* (Richardson *et al.*, 2010). Upon infection

with *S. typhimurium*, the apoptotic receptor CED-1 activates the expression of *pqn/abu* genes in the noncanonical UPR pathway to promote immune response (Haskins *et al.*, 2008). In light of recent studies indicating that there is a strong developmental regulation of certain *abu/pqn* genes (George-Raizen *et al.*, 2014), further investigation will be required to fully address whether such genes are part of the UPR and whether their role in resistance to pathogen infection is related to cuticle and pharynx integrity rather than UPR. A recent study by Glover-Cutter *et al.* demonstrated that SKN-1, the *C. elegans* homolog of Nrf1/2/3 that functions in oxidative stress resistance and longevity, transcriptionally regulates core UPR transcription factors and downstream effectors in the ER-stress response, and the UPR plays a role in activation of the antioxidant/detoxification response (Glover-Cutter *et al.*, 2013).

Response Pathways to Intracellular Pathogen Infections

Unlike in the case of bacterial intestinal infections, canonical defense pathways, such as p38 MAPK and insulin/insulin-like growth factor signaling pathways, do not appear to play roles in resistance to the intracellular pathogen *N. parisii* (Troemel *et al.*, 2008). Ubiquitin-mediated pathways, the proteasome, and xenophagy components are involved in host response and defense against *N. parisii* infection (Bakowski *et al.*, 2014). The ubiquitin-mediated response is also important for defense against the Orsay virus, another natural intracellular pathogen of *C. elegans* (Bakowski *et al.*, 2014).

Both natural and nonnatural viral infection models have been established in *C. elegans*. Studies using laboratory introduced viruses such as Flock House virus (FHV), vesicular stomatitis virus (VSV), and vaccinia virus (VV) have identified a prominent role for the RNA interference (RNAi) machinery in viral restriction, a conserved defense mechanism that was first characterized in plants (Ermolaeva and Schumacher, 2014). These studies confirmed that the RNAi pathway is important for protection from viral infection. *Caenorhabditis elegans* mutants with defective RNAi had higher viral load, while mutants with hyperactive RNAi responses had lower viral load. More recently, three novel RNA viruses, Orsay, Santeuil, and Le Blanc viruses, were found to be able to naturally infect nematodes (Felix *et al.*, 2011; Franz *et al.*, 2012). All three viruses are distantly related to nodaviruses and cause morphological abnormalities in the nematode intestine (Franz *et al.*, 2014). The RNAi pathway has also been implicated in the defense against these natural viruses. Nematode mutants lacking RNAi factors (RDE-1, RDE-4, RNaseD MUT-7, or dicer-related helicase DRH-1) had an elevated infection load.

Conclusions and Perspectives

Caenorhabditis elegans has been proven to be an excellent model system for studying host-pathogen interactions. Although it lacks typical MAMP/PRR-mediated pathogen recognition mechanisms, it can distinguish pathogens from innocuous microbes through surveillance-mediated recognition mechanisms that are capable of detecting disturbances in cellular homeostasis. This ancient mechanism has been

evolutionarily conserved across species. For example, inhibition of host protein synthesis by secreted *Legionella pneumophila* virulence factors activates an innate immune response in mouse macrophages (Fontana *et al.*, 2011). *Escherichia coli*-encoded toxin CNF1 catalyzes deamidation of a host RhoGTPase and triggers NF- κ B immune responses in both insects and mammalian cells (Boyer *et al.*, 2011). Given that there are numerous microbial virulence factors and toxins in the environment aiming at crippling host functions, monitoring their own cellular activities offers the hosts a very effective and economic strategy to detect threats. The only caveat with this hypothesis is that it cannot explain how hosts distinguish between infecting pathogens and mount a pathogen-specific response if different pathogens disrupt the same cellular functions.

There is a complex signaling network in *C. elegans* that is activated in response to microbial infection and stress. This network includes several highly conserved signaling pathways, such as the MAPK pathways, the DAF-2/DAF-16 pathway, and the TGF- β /DBL-1 pathway. There are extensive interplays among immune signaling pathways, between immunity and stress response pathways, and between immunity and longevity pathways. For example, genetic analysis showed that the p38 MAPK pathway contributes to the enhanced longevity regulated by the DAF-2/DAF-16 pathway (Troemel *et al.*, 2006). p38 MAPK signaling also mediates response to oxidative stress generated by the dual oxidase Ce-Duox1/BLI-3 during infection in *C. elegans* (Hoeven *et al.*, 2011). The complexity and specificity of these pathways and their interactions make elucidating molecular details of innate immunity a challenging task, especially in humans. The simplicity of *C. elegans* anatomy and the wealth of genetic and genomic tools for *C. elegans* research make the nematode an outstanding system to dissect host responses to microbial infections at the cellular and molecular levels in a multicellular organism.

See also: Intracellular Infectiology: Cell Processes: Bacterial Subversion of Phagocytic Killing; Cellular Invasion by Bacterial Pathogens

References

- Aballay, A., Yorgey, P., Ausubel, F.M., 2000. *Salmonella typhimurium* proliferates and establishes a persistent infection in the intestine of *Caenorhabditis elegans*. *Current Biology* 10, 1539–1542.
- Alper, S., McBride, S.J., Lackford, B., Freedman, J.H., Schwartz, D.A., 2007. Specificity and complexity of the *Caenorhabditis elegans* innate immune response. *Molecular and Cellular Biology* 27, 5544–5553.
- Bakowski, M.A., Desjardins, C.A., Smelkinson, M.G., *et al.*, 2014. Ubiquitin-mediated response to microsporidia and virus infection in *C. elegans*. *PLoS Pathogens* 10, e1004200.
- Banyai, L., Patthy, L., 1998. Amoebapore homologs of *Caenorhabditis elegans*. *Biochimica et Biophysica Acta* 1429, 259–264.
- Bischof, L.J., Kao, C.Y., Los, F.C., *et al.*, 2008. Activation of the unfolded protein response is required for defenses against bacterial pore-forming toxin *in vivo*. *PLoS Pathogens* 4, e1000176.
- Bolz, D.D., Tenor, J.L., Aballay, A., 2010. A conserved PMK-1/p38 MAPK is required in *Caenorhabditis elegans* tissue-specific immune response to *Yersinia pestis* infection. *Journal of Biological Chemistry* 285 (14), 10832–10840.
- Bork, P., Beckmann, G., 1993. The CUB domain. A widespread module in developmentally regulated proteins. *Journal of Molecular Biology* 231, 539–545.
- Boyer, L., Magoc, L., Dejardin, S., *et al.*, 2011. Pathogen-derived effectors trigger protective immunity via activation of the Rac2 enzyme and the IMD or Rip kinase signaling pathway. *Immunity* 35, 536–549.
- Breger, J., Fuchs, B.B., Aperis, G., *et al.*, 2007. Antifungal chemical compounds identified using a *C. elegans* pathogenicity assay. *PLoS Pathogens* 3, e18.
- Couillault, C., Fourquet, P., Pophillat, M., Ewbank, J.J., 2012. A UPR-independent infection-specific role for a BiP/GRP78 protein in the control of antimicrobial peptide expression in *C. elegans* epidermis. *Virulence* 3, 299–308.
- Couillault, C., Pujol, N., Reboul, J., *et al.*, 2004. TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. *Nature Immunology* 5, 488–494.
- Darby, C., Hsu, J.W., Ghori, N., Falkow, S., 2002. *Caenorhabditis elegans*: Plague bacteria biofilm blocks food intake. *Nature* 417, 243–244.
- Drickamer, K., Dodd, R.B., 1999. C-Type lectin-like domains in *Caenorhabditis elegans*: predictions from the complete genome sequence. *Glycobiology* 9, 1357–1369.
- Dunbar, T.L., Yan, Z., Balla, K.M., Smelkinson, M.G., Troemel, E.R., 2012. *C. elegans* detects pathogen-induced translational inhibition to activate immune signaling. *Cell Host & Microbe* 11, 375–386.
- Engelmann, I., Pujol, N., 2010. Innate immunity in *C. elegans*. *Advances in Experimental Medicine and Biology* 708, 105–121.
- Ermolaeva, M.A., Schumacher, B., 2014. Insights from the worm: The *C. elegans* model for innate immunity. *Seminars in Immunology* 26, 303–309.
- Felix, M.A., Ashe, A., Piffaretti, J., *et al.*, 2011. Natural and experimental infection of *Caenorhabditis nematodes* by novel viruses related to nodaviruses. *PLoS Biology* 9, e1000586.
- Felix, M.A., Duveau, F., 2012. Population dynamics and habitat sharing of natural populations of *Caenorhabditis elegans* and *C. briggsae*. *BMC Biology* 10, 59.
- Fontana, M.F., Banga, S., Barry, K.C., *et al.*, 2011. Secreted bacterial effectors that inhibit host protein synthesis are critical for induction of the innate immune response to virulent *Legionella pneumophila*. *PLoS Pathogens* 7, e1001289.
- Franz, C.J., Renshaw, H., Frezal, L., *et al.*, 2014. Orsay, Santeuil and Le Blanc viruses primarily infect intestinal cells in *Caenorhabditis nematodes*. *Virology* 448, 255–264.
- Franz, C.J., Zhao, G., Felix, M.A., Wang, D., 2012. Complete genome sequence of Le Blanc virus, a third *Caenorhabditis nematode*-infecting virus. *Journal of Virology* 86, 11940.
- Garsin, D.A., Sifri, C.D., Mylonakis, E., *et al.*, 2001. A simple model host for identifying Gram-positive virulence factors. *Proceedings of the National Academy of Sciences of the United States of America* 98, 10892–10897.
- George-Raizen, J.B., Shockley, K.R., Trojanowski, N.F., Lamb, A.L., Raizen, D.M., 2014. Dynamically-expressed prion-like proteins form a cuticle in the pharynx of *Caenorhabditis elegans*. *Biology Open* 3, 1139–1149.
- Glover-Cutter, K.M., Lin, S., Blackwell, T.K., 2013. Integration of the unfolded protein and oxidative stress responses through SKN-1/Nrf. *PLoS Genetics* 9, e1003701.
- Haskins, K.A., Russell, J.F., Gaddis, N., Dressman, H.K., Aballay, A., 2008. Unfolded protein response genes regulated by CED-1 are required for *Caenorhabditis elegans* innate immunity. *Developmental Cell* 15, 87–97.
- Hodgkin, J., Kuwabara, P.E., Corneliussen, B., 2000. A novel bacterial pathogen, *Microbacterium nematophilum*, induces morphological change in the nematode *C. elegans*. *Current Biology* 10, 1615–1618.
- Hoeckendorf, A., Leippe, M., 2012. SPP-3, a saposin-like protein of *Caenorhabditis elegans*, displays antimicrobial and pore-forming activity and is located in the intestine and in one head neuron. *Developmental & Comparative Immunology* 38, 181–186.
- Hoeckendorf, A., Stanisak, M., Leippe, M., 2012. The saposin-like protein SPP-12 is an antimicrobial polypeptide in the pharyngeal neurons of *Caenorhabditis elegans* and participates in defence against a natural bacterial pathogen. *Biochemical Journal* 445, 205–212.
- Hoeven, R., McCallum, K.C., Cruz, M.R., Garsin, D.A., 2011. Ce-Duox1/BLI-3 generated reactive oxygen species trigger protective SKN-1 activity via p38 MAPK signaling during infection in *C. elegans*. *PLoS Pathogens* 7, e1002453.
- Irazoqui, J.E., Troemel, E.R., Feinbaum, R.L., *et al.*, 2010. Distinct pathogenesis and host responses during infection of *C. elegans* by *P. aeruginosa* and *S. aureus*. *PLoS Pathogens* 6, e1000982.
- Jansson, H.B., 1994. Adhesion of Conidia of *Drechmeria coniospora* to *Caenorhabditis elegans* Wild Type and Mutants. *Journal of Nematology* 26, 430–435.

- Jensen, V.L., Simonsen, K.T., Lee, Y.H., Park, D., Riddle, D.L., 2010. RNAi screen of DAF-16/FOXO target genes in *C. elegans* links pathogenesis and dauer formation. *PLoS One* 5, e15902.
- Jones, J.D., Dangl, J.L., 2006. The plant immune system. *Nature* 444, 323–329.
- Kao, C.Y., Los, F.C., Huffman, D.L., et al., 2011. Global functional analyses of cellular responses to pore-forming toxins. *PLoS Pathogens* 7, e1001314.
- Kato, Y., Aizawa, T., Hoshino, H., et al., 2002. *abf-1* and *abf-2*, ASABF-type antimicrobial peptide genes in *Caenorhabditis elegans*. *Biochemical Journal* 361, 221–230.
- Kim, D.H., Feinbaum, R., Alloing, G., et al., 2002. A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. *Science* 297, 623–626.
- Kim, D.H., Liberati, N.T., Mizuno, T., et al., 2004. Integration of *Caenorhabditis elegans* MAPK pathways mediating immunity and stress resistance by MEK-1 MAPK kinase and VHP-1 MAPK phosphatase. *Proceedings of the National Academy of Sciences of the United States of America* 101, 10990–10994.
- Kleino, A., Silverman, N., 2012. UnZIPping mechanisms of effector-triggered immunity in animals. *Cell Host & Microbe* 11, 320–322.
- Liberati, N.T., Fitzgerald, K.A., Kim, D.H., et al., 2004. Requirement for a conserved Toll/interleukin-1 resistance domain protein in the *Caenorhabditis elegans* immune response. *Proceedings of the National Academy of Sciences of the United States of America* 101, 6593–6598.
- Lin, K., Hsin, H., Libina, N., Kenyon, C., 2001. Regulation of the *Caenorhabditis elegans* longevity protein DAF-16 by insulin/IGF-1 and germline signaling. *Nature Genetics* 28, 139–145.
- Los, F.C., Kao, C.Y., Smitham, J., et al., 2011. RAB-5- and RAB-11-dependent vesicle-trafficking pathways are required for plasma membrane repair after attack by bacterial pore-forming toxin. *Cell Host & Microbe* 9, 147–157.
- Los, F.C., Randis, T.M., Aroian, R.V., Ratner, A.J., 2013. Role of pore-forming toxins in bacterial infectious diseases. *Microbiology and Molecular Biology Reviews* 77, 173–207.
- Mahajan-Miklos, S., Tan, M.W., Rahme, L.G., Ausubel, F.M., 1999. Molecular mechanisms of bacterial virulence elucidated using a *Pseudomonas aeruginosa*-*Caenorhabditis elegans* pathogenesis model. *Cell* 96, 47–56.
- Mallo, G.V., Kurz, C.L., Couillault, C., et al., 2002. Inducible antibacterial defense system in *C. elegans*. *Current Biology* 12, 1209–1214.
- Marsh, E.K., Van Den Berg, M.C., May, R.C., 2011. A two-gene balance regulates *Salmonella typhimurium* tolerance in the nematode *Caenorhabditis elegans*. *PLoS One* 6, e16839.
- McEwan, D.L., Kirienko, N.V., Ausubel, F.M., 2012. Host translational inhibition by *Pseudomonas aeruginosa* exotoxin A triggers an immune response in *Caenorhabditis elegans*. *Cell Host & Microbe* 11, 364–374.
- Melo, J.A., Ruvkun, G., 2012. Inactivation of conserved *C. elegans* genes engages pathogen- and xenobiotic-associated defenses. *Cell* 149, 452–466.
- Miltsch, S.M., Seeberger, P.H., Lepenies, B., 2014. The C-type lectin-like domain containing proteins Clec-39 and Clec-49 are crucial for *Caenorhabditis elegans* immunity against *Serratia marcescens* infection. *Developmental & Comparative Immunology* 45, 67–73.
- Mochii, M., Yoshida, S., Morita, K., Kohara, Y., Ueno, N., 1999. Identification of transforming growth factor-beta-regulated genes in *caenorhabditis elegans* by differential hybridization of arrayed cDNAs. *Proceedings of the National Academy of Sciences of the United States of America* 96, 15020–15025.
- Murphy, C.T., Hu, P.J., 2013. Insulin/insulin-like growth factor signaling in *C. elegans*. *WormBook* 26, 1–43.
- Mylonakis, E., Ausubel, F.M., Perfect, J.R., Heitman, J., Calderwood, S.B., 2002. Killing of *Caenorhabditis elegans* by *Cryptococcus neoformans* as a model of yeast pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 99, 15675–15680.
- Nathoo, A.N., Moeller, R.A., Westlund, B.A., Hart, A.C., 2001. Identification of neuropeptide-like protein gene families in *Caenorhabditis elegans* and other species. *Proceedings of the National Academy of Sciences of the United States of America* 98, 14000–14005.
- Nicholas, H.R., Hodgkin, J., 2004. The ERK MAP kinase cascade mediates tail swelling and a protective response to rectal infection in *C. elegans*. *Current Biology* 14, 1256–1261.
- O'Rourke, D., Baban, D., Demidova, M., Mott, R., Hodgkin, J., 2006. Genomic clusters, putative pathogen recognition molecules, and antimicrobial genes are induced by infection of *C. elegans* with *M. nematophilum*. *Genome Research* 16, 1005–1016.
- Pradel, E., Zhang, Y., Pujol, N., et al., 2007. Detection and avoidance of a natural product from the pathogenic bacterium *Serratia marcescens* by *Caenorhabditis elegans*. *Proceedings of the National Academy of Sciences of the United States of America* 104, 2295–2300.
- Pukkila-Worley, R., Ausubel, F.M., 2012. Immune defense mechanisms in the *Caenorhabditis elegans* intestinal epithelium. *Current Opinion in Immunology* 24, 3–9.
- Pukkila-Worley, R., Peleg, A.Y., Tampakakis, E., Mylonakis, E., 2009. Candida albicans hyphal formation and virulence assessed using a *Caenorhabditis elegans* infection model. *Eukaryotic Cell* 8, 1750–1758.
- Reddy, K.C., Andersen, E.C., Kruglyak, L., Kim, D.H., 2009. A polymorphism in *npr-1* is a behavioral determinant of pathogen susceptibility in *C. elegans*. *Science* 323, 382–384.
- Ren, M., Feng, H., Fu, Y., Land, M., Rubin, C.S., 2009. Protein kinase D is an essential regulator of *C. elegans* innate immunity. *Immunity* 30, 521–532.
- Richardson, C.E., Kooistra, T., Kim, D.H., 2010. An essential role for XBP-1 in host protection against immune activation in *C. elegans*. *Nature* 463, 1092–1095.
- Roeder, T., Stanisak, M., Gelhaus, C., et al., 2010. Caenopores are antimicrobial peptides in the nematode *Caenorhabditis elegans* instrumental in nutrition and immunity. *Developmental & Comparative Immunology* 34, 203–209.
- Schulenburg, H., Boehnisch, C., 2008. Diversification and adaptive sequence evolution of *Caenorhabditis* lysozymes (Nematoda: Rhabditidae). *BMC Evolutionary Biology* 8, 114.
- Schulenburg, H., Hoepfner, M.P., Weiner 3rd, J., Bornberg-Bauer, E., 2008. Specificity of the innate immune system and diversity of C-type lectin domain (CTL) proteins in the nematode *Caenorhabditis elegans*. *Immunobiology* 213, 237–250.
- Shivers, R.P., Pagano, D.J., Kooistra, T., et al., 2010. Phosphorylation of the conserved transcription factor ATF-7 by PMK-1 p38 MAPK regulates innate immunity in *Caenorhabditis elegans*. *PLoS Genetics* 6, e1000892.
- Shivers, R.P., Youngman, M.J., Kim, D.H., 2008. Transcriptional responses to pathogens in *Caenorhabditis elegans*. *Current Opinion in Microbiology* 11, 251–256.
- Simonsen, K.T., Gallego, S.F., Faergeman, N.J., Kallipolitis, B.H., 2012. Strength in numbers: “Omics” studies of *C. elegans* innate immunity. *Virulence* 3, 477–484.
- Simonsen, K.T., Moller-Jensen, J., Kristensen, A.R., et al., 2011. Quantitative proteomics identifies ferritin in the innate immune response of *C. elegans*. *Virulence* 2, 120–130.
- Stryer, K.L., Hopkins, G.W., Bartra, S.S., et al., 2005. *Yersinia pestis* kills *Caenorhabditis elegans* by a biofilm-independent process that involves novel virulence factors. *EMBO Reports* 6, 992–997.
- Stryer, K.L., Singh, V., Macosko, E., et al., 2008. Innate immunity in *Caenorhabditis elegans* is regulated by neurons expressing NPR-1/GPCR. *Science* 322, 460–464.
- Sun, J., Singh, V., Kajino-Sakamoto, R., Aballay, A., 2011. Neuronal GPCR controls innate immunity by regulating noncanonical unfolded protein response genes. *Science* 332, 729–732.
- Takeuchi, T., Sennari, R., Sugiura, K., et al., 2008. A C-type lectin of *Caenorhabditis elegans*: its sugar-binding property revealed by glycoconjugate microarray analysis. *Biochemical and Biophysical Research Communications* 377, 303–306.
- Tan, M.W., 2001. Genetic and genomic dissection of host–pathogen interactions using a *P. aeruginosa*–*C. elegans* pathogenesis model. *Pediatric Pulmonology* 32, 96–97.
- Tan, M.W., Mahajan-Miklos, S., Ausubel, F.M., 1999. Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 96, 715–720.
- Tenor, J.L., Aballay, A., 2008. A conserved Toll-like receptor is required for *Caenorhabditis elegans* innate immunity. *EMBO Reports* 9, 103–109.
- Troemel, E.R., Chu, S.W., Reinke, V., et al., 2006. p38 MAPK regulates expression of immune response genes and contributes to longevity in *C. elegans*. *PLoS Genetics* 2, e183.
- Troemel, E.R., Felix, M.A., Whiteman, N.K., Barriere, A., Ausubel, F.M., 2008. Microsporidia are natural intracellular parasites of the nematode *Caenorhabditis elegans*. *PLoS Biology* 6, 2736–2752.
- Wong, D., Bazopoulou, D., Pujol, N., Tavernarakis, N., Ewbank, J.J., 2007. Genome-wide investigation reveals pathogen-specific and shared signatures in the response of *Caenorhabditis elegans* to infection. *Genome Biology* 8, R194.
- Ziegler, K., Kurz, C.L., Cypowyj, S., et al., 2009. Antifungal innate immunity in *C. elegans*: PKCdelta links G protein signaling and a conserved p38 MAPK cascade. *Cell Host & Microbe* 5, 341–352.
- Zugasti, O., Bose, N., Squiban, B., et al., 2014. Activation of a G protein-coupled receptor by its endogenous ligand triggers the innate immune response of *Caenorhabditis elegans*. *Nature Immunology* 15, 833–838.
- Zugasti, O., Ewbank, J.J., 2009. Neuroimmune regulation of antimicrobial peptide expression by a noncanonical TGF-beta signaling pathway in *Caenorhabditis elegans* epidermis. *Nature Immunology* 10, 249–256.

ENCYCLOPEDIA OF CELL BIOLOGY

ENCYCLOPEDIA OF CELL BIOLOGY

EDITORS-IN-CHIEF

RALPH A BRADSHAW

University of California, Irvine, CA, USA

PHILIP D STAHL

Washington University School of Medicine, St. Louis, MO, USA

VOLUME 3

FUNCTIONAL CELL BIOLOGY

VOLUME EDITOR

ALAN EZEKOWITZ

Abide Therapeutics, San Diego, CA, USA

SECTION EDITORS

JASON D WEBER

Washington University, St. Louis, MO, USA

YOSEF YARDEN

Weizmann Institute of Science, Rehovot, Israel

MICHAEL DUSTIN

University of Oxford, Headington, Oxford, UK

DAVID A SASSOON

Université de Pierre et Marie Curie-Sorbonne Universités, Paris, France



AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
225 Wyman Street, Waltham, MA 02451, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

Copyright © 2016 Elsevier Inc. All rights reserved

Cover background image created by M W Goldberg and T D Allen.

Foreground images © Clifford Harding, John Heuser and Philip D Stahl, 1983, originally published in *The Journal of Cell Biology*, Volume 97, 329–339.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-394447-4

For information on all publications visit our
website at <http://store.elsevier.com>

Printed and bound in the United States of America

15 16 17 18 19 20 10 9 8 7 6 5 4 3 2 1



Working together
to grow libraries in
developing countries

www.elsevier.com • www.bookaid.org

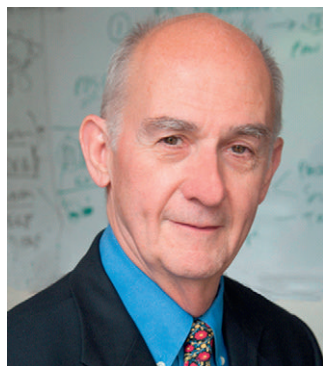
Publisher: Adam Gilbert
Acquisition Editor: Ginny Mills
Content Project Manager: Blerina Osmanaj
Associate Content Project Manager: Victoria Dyer
Cover Designer: Maria Inês Cruz

EDITORS-IN-CHIEF



Ralph A. Bradshaw received his BA in Chemistry from Colby College in 1962 and his PhD from Duke University in Biochemistry in 1966. He performed postdoctoral research at Indiana University and the University of Washington and in 1969 joined the faculty of Washington University School of Medicine in the Department of Biological Chemistry.

In 1982, he became Professor and Chair of the Department of Biological Chemistry at University of California, Irvine. He remained on the faculty of UCI for 25 years, the last 10 in the Department of Physiology and Biophysics. In 2006, he retired as Professor Emeritus at UCI and became Professor of Pharmaceutical Chemistry at the University of California San Francisco. His research has focused upon protein chemistry and the structure and function of growth factors and their receptors. He and his colleagues have determined the partial or complete sequence of over 75 proteins.



Philip D. Stahl is E. Mallinckrodt Jr. Professor Emeritus at Washington University School of Medicine in St. Louis, MO. He was educated at West Virginia University with post-doctoral work at Vanderbilt University. He served as Head of the Department of Cell Biology and Physiology and Director of the Division of Biology and Biomedical Sciences at Washington University. He has been the recipient of many awards including a MERIT award from the NIH and the WICB Senior Recognition award given by the American Society for Cell Biology in recognition of his work supporting the advancement of women in science. Among StahlLab contributions are the discovery of the lysosomal enzyme clearance pathway now implemented in the treatment of lysosomal storage disease, discovery of the innate immune receptor, the mannose receptor, discovery of the exosome secretion pathway, and the role of Rab5 and Arf6 in endocytosis.

Currently his research focuses on endocytosis, signal transduction, and exosome biogenesis and secretion.

VOLUME EDITORS



Gerald Hart is the Director and DeLamar Professor of Biological Chemistry at Johns Hopkins Medical School. He began his research on glycoconjugates about 40 years ago as a graduate student, and he has been active in the field of Glycobiology ever since. Among his research accomplishments, he determined the minimal sequence requirements for N-linked glycosylation while a postdoc in William Lennarz's lab, and he collaborated with Paul Englund's group at Johns Hopkins to elucidate the biosynthetic pathway for GPI-anchors. In the early 1980s, while probing cells with glycosyltransferases, Hart's laboratory discovered cytoplasmic and nuclear protein glycosylation by O-linked N-acetylglucosamine (O-GlcNAc) (e.g., *Journal of Biological Chemistry* 259: 3308; *Journal of Biological Chemistry* 261: 8049). Since that time, the Hart laboratory has published nearly 200 papers on O-GlcNAcylation, identifying and cloning the enzymes controlling cycling, characterizing O-GlcNAcylation and its interplay with phosphorylation on hundreds of proteins, and they have developed many of the tools and methods in use today to study this modification. In

1989, Hart founded the leading journal in the field, *Glycobiology*, serving as Editor-in-Chief until 2001. Hart received the first International Glycoconjugate Organization (IGO) Award in 1997, the Karl Meyer Award from the Society for Glycobiology in 2006, and served as the 2009–2011 president of the IGO. Hart is currently an Associate Editor for *The Journal of Biological Chemistry* and an Associate Editor for *Molecular and Cellular Proteomics*. To date, Hart has published about 263 papers, all in the area of glycosciences. H factor=84.



Bruno Goud studied biochemistry and immunology at the Ecole Normale Supérieure de Cachan and the University of Paris. He received his PhD in 1981 under the mentorship of Jean-Claude Antoine and Stratis Avrameas at Institut Pasteur and did postdoctoral work under the mentorship of Peter Novick at Yale University. In 1995, he established an independent research group in the Department of Cell Biology at the Institut Curie in Paris. The main focus of his studies is the regulation of intracellular transport and membrane trafficking in eukaryotic cells. In particular, his group is working on the functions of Rab GTPases, the mechanisms that sustain the global organization of intracellular compartments and the functions of myosins in membrane traffic. Since 2000, he has also developed original *in vitro* approaches to unravel physical parameters such as membrane tension or membrane curvature that underlie transport processes.

Bruno Goud is a recipient of an ERC (European Research Council) Advanced grant (2013). He is a member of the European Molecular Biology Organization (EMBO). He has been the Head of the Department of Cell Biology of Institut Curie since 2003.



Graça Raposo received her PhD in 1989 in Membrane Biology and Immunology at the University of Paris VII where she specialized in electron microscopy and membrane biology. From 1990 to 1995 she was a postdoctoral fellow at the Immunology Center in Marseille and then in the Department of Cell Biology, Utrecht University, The Netherlands. She is the deputy Director of the Department of Cell Biology at Institut Curie and Director of the Training unit. Her major research interests focus on the biogenesis and functions of exosomes and lysosome related organelles with implications in neurodegenerative disorders, lysosomal diseases, and cancer. Her group have started to unravel the cellular and molecular mechanisms regulating the biogenesis of melanosomes, the lysosome related organelles of epidermal melanocytes, studies that open a new avenue to modulate pigmentation in health and disease. In 2012 she was awarded the CNRS Silver Medal and in 2013 the Descartes Huygens Price from the Royal Dutch Academy of Sciences. She is a member of the European Molecular Biology Organization (EMBO).



Alan Ezekowitz, MBChB, DPhil, FAAP, is Co-Founder, President, and CEO of Abide Therapeutics, a company cofounded by Professors Dale Boger and Ben Cravatt from the Scripps Research Institute. He was previously Senior Vice President and Franchise Head, Bone, Respiratory, Immunology, Endocrine, Dermatology and Urology at Merck Research Laboratories. At Merck, he was responsible for the overall scientific direction of the drug discovery and development process. Prior to Merck, Dr. Ezekowitz was Chief of Pediatric Services at the Massachusetts General Hospital for Children and the Partners Healthcare System and the Charles Wilder Professor of Pediatrics at the Harvard Medical School. He chaired the committee that led to the establishment of the Academy at the Harvard Medical School where he served as a Scholar and Founding member. He served on the Board of Directors of the Partners Healthcare System and Massachusetts General Physicians Organization. In 2008 he was honored with the establishment R. Alan Ezekowitz Professorship in Pediatrics at the Harvard Medical School.

He is a member of the American Society of Clinical Investigation, the Association of American Physicians, the American Pediatric Society, and a Fellow of AAAS. He served on NIH Subcommittees on Biodefense and Vaccine Adjuvants; NAS panels on antibiotic resistance and academic pharmaceutical partnerships; and the ARISE 2 task force of American Academy of Arts & Science that is evaluating the impact of federal and industrial funding of science, engineering, and medicine on American universities.

Dr. Ezekowitz received his MBChB and DPhil from University of Cape Town and Oxford University, respectively. At Children's Hospital in Boston, he was a postdoctoral fellow in the Division of Hematology and Oncology with clinical training in pediatrics. He was on the staff of the Children's Hospital of Boston prior to moving to the Massachusetts General Hospital.

He is a pioneer in the field of innate immunity and has over 150 publications. In particular his group played a major role in defining the structure function of the mannose binding lectin and the macrophage mannose receptor.



Douglas A. Lauffenburger is Ford Professor of Bioengineering, and cofounder and Head of the Department of Biological Engineering at MIT, with affiliate appointment in the Department of Biology; he was a founding codirector of the MIT Computational and Systems Biology Initiative in 2002. His major research interests are in cell engineering, with central focus on cell-cell communication and intracellular signal transduction, emphasizing predictive computational models derived from quantitative experiment. Lauffenburger has coauthored a monograph entitled *Receptors: Models for Binding, Trafficking & Signaling* (Oxford University Press, 1993), has coedited the book entitled *Systems Biomedicine: Concepts and Perspectives* (Elsevier, 2010), and has supervised more than 100 doctoral and post-doctoral students. He is a member of the National Academy of Engineering and the American Academy of Arts and Sciences, and has served as President of the Biomedical Engineering Society, Chair of the AIMBE College of Fellows, and on the NIH NIGMS Advisory Council, and coauthored the NRC report on *A New Biology for the 21st Century*.

SECTION EDITORS



Kenneth E Neet received his PhD in Biochemistry in 1965 (with Dr. Frank W. Putnam) from the University of Florida. He was a postdoctoral fellow at the University of California, Berkeley (with Dr. Daniel E. Koshland, Jr.), and joined the faculty of Case Western Reserve University (CWRU) in 1967 as an Assistant Professor of Biochemistry.

Dr. Neet received a Faculty Research Award of the American Cancer Society from 1968 to 1973 and was on sabbatical leave at the National Institute for Medical Research, Mill Hill, England (with Dr. N. Michael Green) 1973–1974. From 1978 to 1990, he was Professor of Biochemistry at CWRU. During 1980–1981, he took a sabbatical year as a Josiah Macy Faculty Scholar in the Department of Neurobiology, Stanford University (with Dr. Eric M. Shooter).

Dr. Neet moved to Rosalind Franklin University of Medicine and Science/The Chicago Medical School in 1990 and was Professor and Chair of Biochemistry & Molecular Biology until 2005. Dr. Neet became Associate Dean for Research of the Chicago Medical School, RFUMS in 2004.

Dr. Neet has served on the Editorial Boards of the *Journal of Biological Chemistry*, *Protein Science*, and *Molecular & Cellular Proteomics* and as a member of study sections for the National Institutes of Health and the National Science Foundation. He was an Associate Editor of the *Journal of Biological Chemistry* (1996–2013).

Dr. Neet's research interests have been in the general areas of protein–protein interactions, allosteric interactions, protein conformational changes, and cell signaling. In his earlier years he studied theoretical and experimental aspects of slow transitions in enzymes, particularly the cooperativity of the monomeric enzyme glucokinase. Later, he studied ligand–receptor interactions of nerve growth factor, establishing a mutational system to study the interactions with its receptors (TrkA and p75), and ultimately the complex signaling emanating from these receptors within neuronal systems.

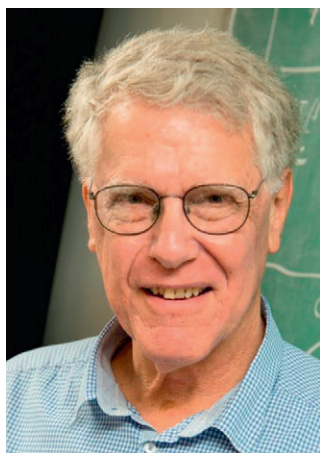


Sarah A Woodson is a biophysical chemist interested in how RNA molecules fold into specific three-dimensional structures and how they interact with proteins to turn genes on and off in the cell. She was born in Michigan, USA, and attended Kalamazoo College before receiving her PhD in Biophysical Chemistry from Yale University in 1987. After postdoctoral research in the laboratory of Tom Cech at the University of Colorado Boulder from 1987 to 1990, she joined the faculty of the University of Maryland in 1990 and the faculty of Johns Hopkins University in 1999, where she is currently the T.C. Jenkins Professor of Biophysics. Together with Mark Chance and Michael Brenowitz, she pioneered methods for visualizing how RNA molecules change shape in real time. She served on the board of the RNA Society and was elected an AAAS Fellow in 2011 and a Pew Scholar in the Biomedical Sciences in 1993. She is currently a reviewing editor of *Biopolymers* and *Journal of Molecular Biology* and serves on the editorial boards of *Nucleic Acids Research* and *RNA*.



Judith S Bond, Evan Pugh Emeritus Professor of Biochemistry and Molecular Biology at the Pennsylvania State University College of Medicine, was President of the Federation of American Societies for Experimental Biology (2012–13) and an Associate Editor of the *Journal of Biological Chemistry* (1999–2013). She received her BS degree from Bennington College in Vermont, her PhD from Rutgers University, and did postdoctoral research at Vanderbilt University. She rose through the academic ranks at the Medical College of Virginia, Virginia Commonwealth University, became Professor and Head of Biochemistry and Nutrition at Virginia Polytechnic Institute and State University, and served at Penn State University College of Medicine as Professor and Chair of Biochemistry and Molecular Biology from 1992 to 2012. At Penn State, she also served as assistant dean for graduate studies, codirector of the All-University Interdisciplinary Biological Sciences Program and founding director of the Medical Scientist Training Program. She recently directed NIH-funded summer research programs for high school students and teachers and for undergraduate students (particularly underrepresented minority groups). Her work on metalloproteases has been funded continuously by the NIH for over 35 years. She has

trained 5 Master's, 21 PhD students, and 19 postdoctoral trainees. Her professional service included member and chair National Institutes of Health Study Sections, member of the National Institute of Diabetes, Digestive and Kidney Diseases Advisory Council, member of the American Association of Medical Colleges – Howard Hughes Medical Institute Committee on the Scientific Foundations for Future Physicians, and President of the American Society for Biochemistry and Molecular Biology.



Elliot L. Elson received his AB at Harvard, his PhD at Stanford under the mentorship of R.L. Baldwin and did postdoctoral work at University of California, San Diego, under Bruno Zimm. His first independent faculty position was in the Department of Chemistry, Cornell University, where he stayed for 11 years. He then moved to the Department of Biological Chemistry (now the Department of Biochemistry and Molecular Biophysics) at Washington University in St. Louis, School of Medicine. He has pursued research in several biophysical areas. One of these is the development of Fluorescence Correlation Spectroscopy and Fluorescence Photobleaching Recovery and their application to studies of diffusion in cell membranes and of phase separation in model membrane bilayers as well as of other cellular and noncellular phenomena. He has also worked in the areas of cell mechanics, cell motility, and the mechanical properties of engineered tissues. His laboratory has developed approaches for measuring mechanical properties of cells in monolayer culture and for determining the contributions of cells and extracellular matrix to the mechanics of engineered heart and connective tissue constructs.



Tamotsu Yoshimori received his PhD degree in 1989 at Osaka University. After working at several places including European Molecular Biology Laboratory (Prof. Kai Simons' lab) and National Institute of Basic Biology (Prof. Yoshinori Ohsumi's lab), he is now a distinguished professor of Osaka University (Graduate School of Medicine and of Frontier Biosciences). His research interests are focused on intracellular membrane trafficking, and especially for the last 18 years, on autophagy. He identified LC3 as an autophagosome-binding protein, which has been widely used as the gold standard in autophagy assays. The paper has been cited over 3000 times. He also provided new insights into membrane biogenesis in autophagy and the role of autophagy in pathogen defense and suppression of various diseases. He authored or coauthored over 220 journal articles and book chapters. He is an editor of *Journal of Cell Science*, and on the editorial board of *Journal of Cell Biology*, *Molecular Biology of the Cell*, and so on. He is a member of the Faculty of 1000 and a vice president of Japan Society for Cell Biology. He was awarded Osaka University Presidential Award for 2012 and 2013, Prize for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology in 2013, Kakiuchi Saburo Memorial Prize by the Japanese Biochemical Society in 2014, and selected as Highly Cited Researchers 2014 by Thomson Reuters.



Paul A. Gleeson is currently Head of the Department of Biochemistry and Molecular Biology at the University of Melbourne, located at Bio21 Institute. His research interests include the molecular mechanisms of intracellular membrane transport and the molecular basis of organ-specific autoimmune diseases. Paul Gleeson obtained his PhD in 1980 from the University of Melbourne and did postdoctoral research in the biosynthesis and function of glycoproteins at the Hospital for Sick Children, Toronto, National Institute for Medical research, Mill Hill London and Department of Biochemistry, La Trobe University, Melbourne. He established an independent laboratory at Monash University in 1986 where he defined the targeting signals of Golgi glycosyltransferases, identified golgins of the *trans*-Golgi network and along with his colleagues developed mouse models of autoimmune gastritis. In 2001 he moved to Department of Biochemistry and Molecular Biology at the University of Melbourne and has been Head of the Department since 2006. He has a number of international research collaborations and has been a visiting scientist at the EMBL, Heidelberg, and the Institut Curie, Paris.

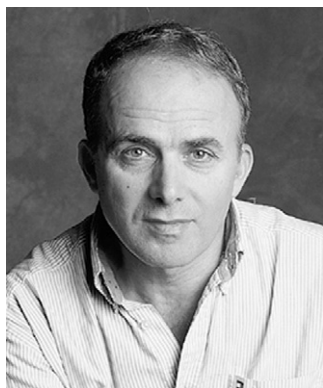


Anna Akhmanova studied biochemistry and molecular biology at the Moscow State University. She received her PhD in 1997 at the University of Nijmegen, the Netherlands. She worked as a postdoc at the Department of Microbiology and Evolutionary Biology at the University of Nijmegen and at the Department of Cell Biology at the Erasmus Medical Center in Rotterdam, the Netherlands. In 2001, she started her own research group at the Erasmus Medical Center. Since 2011, Anna Akhmanova is professor of Cell Biology at Utrecht University, the Netherlands.

Akhmanova studies cytoskeletal organization and trafficking processes, which contribute to cell polarization, differentiation, vertebrate development, and human disease. The main focus of her studies is the microtubule cytoskeleton. Her group has described key interactions between proteins associated with microtubule ends and determined their role in the regulation of microtubule organization and dynamics. She has also characterized mechanisms of bidirectional microtubule-based motility of membrane organelles such as cell nuclei and exocytotic vesicles. Her research relies on combining high-resolution live cell imaging and quantitative analysis of cytoskeletal remodeling, measurement of protein dy-

namics using advanced microscopic assays, *in vitro* reconstitution of dynamic cytoskeleton-based processes and different methods of identification of protein–protein interactions.

Anna Akhmanova is a recipient of the ALW Vernieuwingsimpuls VIDI (2001) and VICI awards (2007) (Netherlands Organization for Scientific Research), and an ERC Synergy grant (2013). She is the Chair of the board of the Netherlands Microscopy Society and a member of the European Molecular Biology Organization (EMBO).



Yosef (Yossi) Yarden is a Professor in the Department of Biological Regulation at Weizmann Institute of Science (Rehovot, Israel). Dr. Yarden obtained his PhD from Weizmann Institute (under the mentoring of Dr. Joseph Schlessinger) and later trained at both Genentech Inc., with Dr. Axel Ullrich, and at the Whitehead Institute (MIT), with Dr. Robert A. Weinberg. His group studies the roles played by growth factors and their receptors in cancer progression. They also explore strategies able to intercept growth factor signals in tumors.



Jason D Weber is an Associate Professor in the Departments of Internal Medicine and Cell Biology and Physiology at Washington University and is the Coleader of the Breast Cancer Research Program in the Siteman Comprehensive Cancer Center. Dr. Weber obtained his PhD from Saint Louis University and received postdoctoral training under the mentoring of Dr. Charles J. Sherr at St. Jude Children's Research Hospital in Memphis, TN. His group studies the molecular interplay between oncogenes and tumor suppressors, particularly focused on their role in regulating cellular growth processes.



Michael Dustin received his PhD in 1990 in Cell and Developmental Biology from Harvard University, where he worked in the laboratory of Timothy A. Springer, PhD. During his graduate work he was involved in the identification and cloning of ICAM-1 and ICAM-2, key ligands for the integrin LFA-1. He further demonstrated that LFA-1 dependent adhesion was stimulated by T cell receptor signaling. He did his postdoctoral training with Stuart Kornfeld at Washington University School of Medicine in St. Louis, focusing on mechanisms of lysosome biogenesis. He started his independent lab also at Washington University School of Medicine, and used supported lipid bilayers to define the first protein to modulate organization of the immunological synapse- CD2 associated protein- and the dynamics of immunological synapse formation. He moved to NYU School of Medicine in 2001 where he applied *in vivo* microscopy to tolerance induction and immune responses in effector sites including the liver, brain, and spleen. Continued work with the immunological synapse model resulted in discovery of signaling microclusters, the molecular basis of immunological synapse stability, and the direct budding of extracellular microvesicle enriched in T cell receptors in the immunological synapse. Professor Dustin recently took a position at the

University of Oxford with support of a Wellcome Trust Principal Research Fellowship. The focus of his new post is the translation of the immunological synapse. He received a Presidential Early Career Award in Science and Engineering and the DART-NYU Biotechnology Achievement Award.



David A Sassoon directs a research team at Paris Sorbonne. He received his PhD from Columbia University (NYC, USA) in 1986 in the biological sciences and was a professor at Boston University Medical School and Mt. Sinai Medical School before relocating to Paris in 2006. His long-standing interests are in developmental and stem cell biology in a number of tissues including skeletal muscle, skin, CNS, and heart. Dr. Sassoon recently concluded the coordination of a EC-funded 15 partner international consortium (Endostem) designed to identify novel therapeutics for mobilizing endogenous stem/progenitor cells (<http://www.endostem.eu/>) and is a recent recipient as coordinator of a multipartner Transatlantic Network of Excellence grant from the Fondation Leducq focused on cardiac stem cell biology (<http://www.fondationleducq.org/nivel2.aspx?idsec=1360>).

A major focus of his research is on adult progenitor/stem cell biology and how these cells respond to stress. His team has identified a parentally imprinted gene, PW1/Peg3, which is involved in both p53 and inflammatory responses. PW1 is expressed in adult stem cells in all tissues identified to date. Loss of PW1 function in stem cells results in a loss of stem cell competence including a reduced capacity to undergo self-renewal and respond to hypoxic stress. We are currently evaluating a role for PW1 in postnatal and adult heart with a particular focus on its role in vessel precursors using both myocardial infarction and stress-induced myocardial hypertrophy.



Jason M Haugh is a Professor of Chemical and Biomolecular Engineering at North Carolina State University in Raleigh, NC, USA, where he has been on the faculty since 2000. His laboratory has been among those to pioneer the synthesis of quantitative experiments and modeling to study signal transduction in mammalian cells. Since the lab's inception, their approach has combined biochemical measurements, live-cell fluorescence microscopy, and computational modeling to elucidate signaling mechanisms by analyzing their kinetics and spatial organization in cells. The systems studied by the Haugh Laboratory include: regulation of the phosphoinositide 3-kinase, Ras/extracellular signal-regulated kinase, and phospholipase C pathways mediated by receptor tyrosine kinases, and cross talk between these pathways; dynamic organization of multimolecular complexes at cell membranes; signaling mediated by cytokine and chemokine receptors in immune cells; and integration of adhesion, signaling, and cytoskeletal dynamics that direct cell migration.



After graduating with an MA in Mathematics from the University of Cambridge, Helen Mary Byrne received her Master of Science and DPhil from the University of Oxford. She worked as a postdoctoral researcher at the Universities of Oxford and Bath, before taking up a lectureship at the UMIST in Manchester. She moved to the University of Nottingham in 1998 and was awarded a prestigious Advanced Research Fellowship from the United Kingdom's Engineering and Physical Sciences Research Council (2000–2006). She was promoted to Reader in 2002 and Professor in 2003. While in Nottingham, she established and then led the Centre for Mathematical Medicine and Biology until she returned to Oxford in 2011 where she is now based in the Mathematical Institute at the University of Oxford. She has over 20 years' experience of developing, analyzing, and simulating continuum and multiscale models of biomedical systems, and a particular interest in studying the growth and response to treatment of solid tumors.



Rune Linding completed his PhD at the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany, followed by postdoctoral training at EMBL. He then jointly trained with professors Tony Pawson and Mike Yaffe at the Lunenfeld at Mount Sinai Hospital in Toronto, Canada, and the Massachusetts Institute of Technology (MIT) in Cambridge, US, respectively. Dr. Linding then established his own laboratory of Cellular and Molecular Logic at the Institute of Cancer Research (ICR) in London, UK, before returning to Denmark to take a position as professor of cellular signal integration at the Technical University of Denmark. In 2014, Dr. Linding moved his laboratory to the Biotech Research and Innovation Centre (BRIC) at University of Copenhagen where he is currently professor of cellular signaling. His research group focuses on big data network biology, exploring biological systems by developing and deploying algorithms aimed to predict cell behavior, in particular looking at cellular signal processing and decision making. A strategic focus is to continue to develop computational tools (such as KinomeExplorer, NetworKIN, and NetPhorest) and to deploy these on genome-scale quantitative data obtained by, for example, mass spectrometry, genomic, and phenotypic screens to understand the principles of how spatio and temporal assembly of mammalian signaling networks transmit and process information at a systems level in order to alter cell behavior. His overarching aim is to advance network

medicine by identifying and targeting signaling networks associated with complex diseases. To this end Dr. Linding is currently leading high-level, strategic, multidisciplinary studies of signaling network dynamics driving cancer metastasis in collaboration with other labs at Harvard, Yale, The Jackson Laboratory, Memorial Sloan Kettering Cancer Center, MIT, and BRIC.

CONTRIBUTORS TO VOLUME 3

AS Aing

Thomas Jefferson University, Philadelphia, PA, USA

B Alarcón

Centro Biología Molecular Severo Ochoa, CSIC/UAM, Madrid, Spain

C Albayrak

ETH Zürich, Basel, Switzerland

DW Andrews

Sunnybrook Research Institute, Toronto, ON, Canada

M Antsiferova

Institute of Molecular Health Sciences, Zurich, Switzerland

TI Arnon

University of Oxford, Oxford, UK

JD Axelrod

Stanford University School of Medicine, Stanford, CA, USA

JJ Babon

Walter and Eliza Hall Institute of Medical Research, Melbourne, VIC, Australia and The University of Melbourne, Melbourne, VIC, Australia

JJ Baldassare

Saint Louis University School of Medicine, St. Louis, MO, USA

A Baransk

University Hospital of Erlangen, Erlangen, Germany and Aix Marseille Université, Marseille-Luminy, France

ME Bekier II

University of Toledo, Toledo, OH, USA and University of Michigan, Ann Arbor, MI, USA

A Besser

University of Miami Miller School of Medicine, Miami, FL, USA

G Blandino

Regina Elena National Cancer Institute, Rome, Italy

ES Blomain

Thomas Jefferson University, Philadelphia, PA, USA

H Brahmbhatt

Sunnybrook Research Institute, Toronto, ON, Canada

CL Brooks

Stemline Therapeutics, Inc., New York, NY, USA

SE Broughton

St. Vincent's Institute of Medical Research, Fitzroy, VIC, Australia

HE Broxmeyer

Indiana University School of Medicine, Indianapolis, IN, USA

S Buczacki

University of Cambridge, Cambridge, UK

K Cadwell

New York University School of Medicine, New York, NY, USA

ML Capitano

Indiana University School of Medicine, Indianapolis, IN, USA

E Carafoli

University of Padova, Padova, Italy

CV Carman

Harvard Medical School, Boston, MA, USA

C Carmona-Rivera

National Institutes of Health, Bethesda, MD, USA

G Cesareni

University of Rome Tor Vergata, Rome, Italy and Santa Lucia Foundation, Rome, Italy

A Chauveau

University of Oxford, Oxford, UK

M Chircop

Children's Medical Research Institute, Westmead, NSW, Australia and University of Sydney, Sydney, NSW, Australia

M Ciganda

University at Buffalo, Buffalo, NY, USA

M Collado

IDIS, Clinical University Hospital (CHUS), SERGAS, Santiago de Compostela, Spain

RS Cook

Vanderbilt University School of Medicine, Nashville, TN, USA

R Correra

Université de Pierre et Marie Curie-Sorbonne Universités, Paris, France

EJ Coulson

The University of Queensland, Brisbane, QLD, Australia

M Cristina de Almagro
Genentech, South San Francisco, CA, USA

AM Cuervo
Institute for Aging Studies, Albert Einstein College of Medicine, NY, USA

S Da Silva-Álvarez
IDIS, Clinical University Hospital (CHUS), SERGAS, Santiago de Compostela, Spain

BW Darbro
University of Iowa, Iowa City, IA, USA and University of Iowa Hospitals and Clinics, Iowa City, IA, USA

JO DaSilva
Regeneron Pharmaceuticals Inc., Tarrytown, NY, USA

JA DeCaprio
Dana-Farber Cancer Institute, Boston, MA, USA

A-L Denizot
Université de Pierre et Marie Curie-Sorbonne Universités, Paris, France

AE Denton
University of Cambridge, Cambridge, UK

L Deriano
Institut Pasteur, Paris, France

U Dhagat
St. Vincent's Institute of Medical Research, Fitzroy, VIC, Australia

S Djuranovic
Washington University, St. Louis, MO, USA

D Dudziak
University Hospital of Erlangen, Erlangen, Germany

ML Dustin
New York University School of Medicine, New York, NY, USA and The University of Oxford, Headington, UK

M Dustin
University of Oxford, Headington, Oxford, UK

J El-Benna
INSERM U1149, Centre de Recherche sur l'Inflammation CRI, Paris, France and Université Paris Diderot, Paris, France

A Esteves
National Institutes of Health, Bethesda, MD, USA

C Evangelisti
National Research Council, Bologna, Italy and Rizzoli Orthopedic Institute, Bologna, Italy

KL Everett
F-star Biotechnology Ltd., Cambridge, UK

A Ezekowitz
Abide Therapeutics, San Diego, CA, USA

DT Fearon
Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA and Weill Cornell Medical College, New York, NY, USA

JR Fergusson
University of Oxford, Oxford, UK

M Ferraiuolo
Regina Elena National Cancer Institute, Rome, Italy

T Finkel
National Institutes of Health, Bethesda, MD, USA

L Formicola
Université de Pierre et Marie Curie-Sorbonne Universités, Paris, France

DA Franklin
The University of North Carolina, Chapel Hill, NC, USA

SM Frisch
West Virginia University, Morgantown, WV, USA

L Galluzzi
INSERM U1138, Paris, France; Gustave Roussy, Villejuif, France; Centre de Recherche des Cordeliers, Paris, France; and Université Paris Descartes/Paris V – Sorbonne Paris Cité, Paris, France

L García-Prat
Pompeu Fabra University (UPF), Barcelona, Spain

B Geiger
Weizmann Institute of Science, Rehovot, Israel

S Gerondakis
Monash University Central Clinical School, Melbourne, VIC, Australia

TD Gilmore
Boston University, Boston, MA, USA

HS Goodridge
Cedars-Sinai Medical Center, Los Angeles, CA, USA

S Gordon
Sir William Dunn School of Pathology, Oxford, UK and Stellenbosch Institute for Advanced Study, Stellenbosch, South Africa

L Gu
Beckman Research Institute at City of Hope, Duarte, CA, USA

MA Guthridge
Australian Centre for Blood Cancers, Monash University, Melbourne, VIC, Australia

MG Hall
MD Anderson Cancer Center, Houston, TX, USA

N Hassanzadeh-Kiabi
Cedars-Sinai Medical Center, Los Angeles, CA, USA

X He
Harvard Medical School, Boston, MA, USA

L Heger
University Hospital of Erlangen, Erlangen, Germany

GF Heidkamp
University Hospital of Erlangen, Erlangen, Germany

C-H Heldin
Uppsala University, Uppsala, Sweden

TR Hercus
Centre for Cancer Biology, SA Pathology and University of South Australia, Adelaide, SA, Australia

PJ Hertzog
MIMR-PHI Institute of Medical Research, Monash University, Clayton, VIC, Australia

RJ Hickey
Beckman Research Institute at City of Hope, Duarte, CA, USA

CS Hill
Cancer Research UK London Research Institute, London, UK

T Hines
University of South Carolina, Columbia, SC, USA

A Hoffmann
University Hospital of Erlangen, Erlangen, Germany

JB Huppa
Medical University of Vienna, Vienna, Austria

C Hutcheon
Institute for Systems Biology, Seattle, WA, USA

S Hutten
University of Dundee, Dundee, UK and Ludwig-Maximilians-University, Munich, Germany

K Inoki
University of Michigan, Ann Arbor, MI, USA

PA Janmey
University of Pennsylvania, Philadelphia, PA, USA

D Jenkins
UCL Institute of Child Health, London, UK

BJ Jenkins
MIMR-PHI Institute of Medical Research, Monash University, Clayton, VIC, Australia

AL Kalinski
Drexel University, Philadelphia, PA, USA

WL Kan
Centre for Cancer Biology, SA Pathology and University of South Australia, Adelaide, SA, Australia

MJ Kaplan
National Institutes of Health, Bethesda, MD, USA

M Katan
University College London, London, UK

P Klenerman
University of Oxford, Oxford, UK

JRP Knight
University of Leicester, Leicester, UK

J Krebs
MPI for Biophysical Chemistry, Goettingen, Germany

G Kroemer
INSERM U1138, Paris, France; Hôpital Européen Georges Pompidou, Paris, France; Gustave Roussy, Villejuif, France; Centre de Recherche des Cordeliers, Paris, France; and Université Paris Descartes/Paris V, Sorbonne Paris Cité, Paris, France

S Kumar
University of South Australia, Adelaide, SA, Australia

S Kyryachenko
Université de Pierre et Marie Curie-Sorbonne Universités, Paris, France

K Kyrylkova
Université de Pierre et Marie Curie-Sorbonne Universités, Paris, France

AI Lamond
University of Dundee, Dundee, UK

B Leber
Hamilton, ON, Canada

CHK Lehmann
University Hospital of Erlangen, Erlangen, Germany

J Lennartsson
Uppsala University, Uppsala, Sweden

NH Lents
The City University of New York, New York, NY, USA

C Lescale
Institut Pasteur, Paris, France

CM Li
Beckman Research Institute at City of Hope, Duarte, CA, USA

ATY Lian

*Children's Medical Research Institute, Westmead,
NSW, Australia and University of Sydney, Sydney,
NSW, Australia*

Y Lin

*University of Kentucky College of Medicine, Lexington,
KY, USA*

RG Lingeman

*Beckman Research Institute at City of Hope, Duarte,
CA, USA*

MA Linterman

Babraham Institute, Cambridge, UK

C Liu

University of Kentucky, Lexington, KY, USA

K Liu

*Columbia University Medical Center, New York, NY,
USA*

AF Lopez

*Centre for Cancer Biology, SA Pathology and University
of South Australia, Adelaide, SA, Australia*

J Lühr

University Hospital of Erlangen, Erlangen, Germany

T Ly

University of Dundee, Dundee, UK

F Macian

Albert Einstein College of Medicine, Bronx, NY, USA

LH Malkas

*Beckman Research Institute at City of Hope, Duarte,
CA, USA*

G Manic

Regina Elena National Cancer Institute, Rome, Italy

A Mansell

*MIMR-PHI Institute of Medical Research, Monash
University, Clayton, VIC, Australia*

G Marazzi

*Université de Pierre et Marie Curie-Sorbonne
Universités, Paris, France*

AM Marchiando

*New York University School of Medicine, New York,
NY, USA*

CL Marriott

University of Birmingham, Birmingham, UK

AM Martelli

University of Bologna, Bologna, Italy

R Martinelli

Harvard Medical School, Boston, MA, USA

CA McCulloch

University of Toronto, Toronto, ON, Canada

MA McNiven

Mayo Clinic, Rochester, MN, USA

D Merlino

Thomas Jefferson University, Philadelphia, PA, USA

DSJ Miller

*Cancer Research UK London Research Institute,
London, UK*

RT Miller

*University of Texas Southwestern Medical School,
Dallas, TX, USA and Dallas VA Medical Center,
Dallas, TX, USA*

M Mingueneau

Biogen Idec, Cambridge, MA, USA

P Muñoz-Cánoves

*Pompeu Fabra University (UPF), Barcelona, Spain and
Institutió Catalana de Recerca i Estudis Avançats
(ICREA), Barcelona, Spain*

JM Murphy

*Walter and Eliza Hall Institute of Medical Research,
Melbourne, VIC, Australia and The University of
Melbourne, Melbourne, VIC, Australia*

P My-Chan Dang

*INSERM U1149, Centre de Recherche sur
l'Inflammation CRI, Paris, France and Université Paris
Diderot, Paris, France*

MF Navarro

Sanford-Burnham Medical Research Institute, CA, USA

TL Nero

*St. Vincent's Institute of Medical Research, Fitzroy, VIC,
Australia*

C Neyen

*École Polytechnique Fédérale de Lausanne, Lausanne,
Switzerland*

SE Nicholson

*Walter and Eliza Hall Institute of Medical Research,
Melbourne, VIC, Australia and The University of
Melbourne, Melbourne, VIC, Australia*

NA Nicola

*Walter and Eliza Hall Institute of Medical Research,
Melbourne, VIC, Australia and The University of
Melbourne, Melbourne, VIC, Australia*

A Nicolas

University of Dundee, Dundee, UK

J Nteeba

University of Iowa, Iowa City, IA, USA

D Ollitrault
*Université de Pierre et Marie Curie-Sorbonne
Universités, Paris, France*

T Palmer
University of California, San Diego, CA, USA

VG Panse
ETH Zürich, Switzerland

MW Parker
*St. Vincent's Institute of Medical Research, Fitzroy, VIC,
Australia and University of Melbourne, Parkville, VIC,
Australia*

N Parnell
University of Michigan, Ann Arbor, MI, USA

SJ Parsons
*University of Virginia Health System, Charlottesville,
VA, USA*

P Paulvannan
Institute for Systems Biology, Seattle, WA, USA

D Peluso
*University of Rome Tor Vergata, Rome, Italy and Santa
Lucia Foundation, Rome, Italy*

E Perdiguero
Pompeu Fabra University (UPF), Barcelona, Spain

RYC Poon
*Hong Kong University of Science and Technology, Clear
Water Bay, Hong Kong*

J Puccini
University of South Australia, Adelaide, SA, Australia

PL Puri
*Sanford-Burnham Medical Research Institute, La Jolla,
CA, USA and IRCCS Fondazione Santa Lucia, Rome,
Italy*

DE Quelle
*University of Iowa, Iowa City, IA, USA and University
of Iowa Hospitals and Clinics, Iowa City, IA, USA*

A Rajput
*Sanford Burnham Medical Research Institute, La Jolla,
CA, USA*

MM Rasenick
*University of Illinois Chicago College of Medicine,
Chicago, IL, USA and Jesse Brown VA Medical Center,
Chicago, IL, USA*

GL Razidlo
Mayo Clinic, Rochester, MN, USA

Y Regimbald-Dumas
Harvard Medical School, Boston, MA, USA

EW Roberts
University of California, San Francisco, CA, USA

KT Roybal
University of California, San Francisco, CA, USA

A Sacco
*Sanford-Burnham Medical Research Institute, La Jolla,
CA, USA*

W Saengsawang
Mahidol University, Bangkok, Thailand

G Salvesen
Sanford-Burnham Medical Research Institute, CA, USA

DM Sansom
*UCL Institute of Immunity and Transplantation,
London, UK*

HM van Santen
*Centro Biología Molecular Severo Ochoa, CSIC/UAM,
Madrid, Spain*

M Santiago-Medina
University of California, San Diego, CA, USA

DA Sassoon
*Université de Pierre et Marie Curie-Sorbonne
Universités, Paris, France*

T Sbarrato
University of Leicester, Leicester, UK

GJ Schütz
Vienna University of Technology, Vienna, Austria

R Seger
The Weizmann Institute of Science, Rehovot, Israel

MM Shannack
Babraham Institute, Cambridge, UK

N Siddiqui
McGill University, Montreal, QC, Canada

TW Siggers
Boston University, Boston, MA, USA

S Skeldal
Aarhus University Hospital, Denmark

J Slingerland
*University of Miami Miller School of Medicine, Miami,
FL, USA*

DS Smith
University of South Carolina, Columbia, SC, USA

L Smith
University of Minnesota, Minneapolis, MN, USA

SJ Smith
Beckman Research Institute at City of Hope, Duarte, CA, USA

AE Snook
Thomas Jefferson University, Philadelphia, PA, USA

N Sonenberg
McGill University, Montreal, QC, Canada

B Soskic
UCL Institute of Immunity and Transplantation, London, UK

P Sousa-Victor
Pompeu Fabra University (UPF), Barcelona, Spain and Buck Institute for Research on Aging, Novato, CA, USA

JC Stanford
Vanderbilt University School of Medicine, Nashville, TN, USA

M Stoneley
University of Leicester, Leicester, UK

S Strano
Regina Elena National Cancer Institute, Rome, Italy

N Subramanian
Institute for Systems Biology, Seattle, WA, USA

Y Svitkin
McGill University, Montreal, QC, Canada

AM Tartakoff
Case Western Reserve University School of Medicine, Cleveland, OH, USA

S Tay
ETH Zürich, Basel, Switzerland

WR Taylor
University of Toledo, Toledo, OH, USA

J Thomas Parsons
University of Virginia Health System, Charlottesville, VA, USA

TD Tlsty
University of California, San Francisco, CA, USA

JL Twiss
University of South Carolina, Columbia, SC, USA

K Umaer
University at Buffalo, Buffalo, NY, USA

M Vaman Rao
National University of Singapore, Singapore

DB Vaught
Vanderbilt University School of Medicine, Nashville, TN, USA

I Vitale
Regina Elena National Cancer Institute, Rome, Italy

D Vucic
Genentech, South San Francisco, CA, USA

SA Waldman
Thomas Jefferson University, Philadelphia, PA, USA

CF Ware
Sanford Burnham Medical Research Institute, La Jolla, CA, USA

JD Weber
Washington University, St. Louis, MO, USA

NA de Weerd
MIMR-PHI Institute of Medical Research, Monash University, Clayton, VIC, Australia

CS Weirich
ETH Zürich, Switzerland

S Werner
Institute of Molecular Health Sciences, Zurich, Switzerland

S Wexler
The Weizmann Institute of Science, Rehovot, Israel

N Williams
University at Buffalo, Buffalo, NY, USA

AE Willis
University of Leicester, Leicester, UK

DR Withers
University of Birmingham, Birmingham, UK

C Wülfing
University of Bristol, Bristol, UK

MB Yaffe
Broad Institute Massachusetts Institute of Technology, Cambridge, MA, USA

J Yang
University of California, San Diego, CA, USA

Y Yao
University of Michigan, Ann Arbor, MI, USA

Y Yarden
Weizmann Institute of Science, Rehovot, Israel

G Zachos
University of Crete, Heraklion, Greece

HS Zaher
Washington University, St. Louis, MO, USA

R Zaidel-Bar
National University of Singapore, Singapore

X Zhang

Harvard Medical School, Boston, MA, USA

Y Zhang

*The University of North Carolina, Chapel Hill, NC,
USA*

BP Zhou

*University of Kentucky College of Medicine, Lexington,
KY, USA*

P Zweidler-McKay

MD Anderson Cancer Center, Houston, TX, USA

CONTENTS OF VOLUME 3

Preface xxvii

Cell Biology: An Overview *RA Bradshaw and PD Stahl* xxix

Functional Cell Biology

Functional Cell Biology: An Overview *JD Weber, Y Yarden, M Dustin, DA Sassoon, and A Ezekowitz* 1

Cell Communication

Growth Factor Mediated Cell Signaling

Receptor Tyrosine Kinases and Their Ligands *C-H Heldin and J Lennartsson* 8

Cytokine Receptors and Their Ligands *U Dhagat, TL Nero, WL Kan, TR Hercus, SE Broughton, NA de Weerd, BJ Jenkins, A Mansell, MA Guthridge, PJ Hertzog, AF Lopez, and MW Parker* 22

TGF- β Superfamily Signaling *DSJ Miller and CS Hill* 37

G Protein-Coupled Receptors *W Saengsawang and MM Rasenick* 51

Guanylyl Cyclase Receptors *AE Snook, ES Blomain, AS Aing, D Merlino, and SA Waldman* 56

Tumor Necrosis Factor Receptors: A Brief Digestion *Y Lin and BP Zhou* 61

Signaling and Function of Death Receptors of the Tumor Necrosis Factor Receptor Superfamily *S Skeldal and EJ Coulson* 67

Hedgehog Signaling in Development and Disease *D Jenkins* 76

The Wnt/ β -Catenin Pathway *C Liu, Y Regimbald-Dumas, X Zhang, and X He* 86

The Hippo Pathway *M Ferraiuolo, S Strano, and G Blandino* 99

The Notch Pathway *MG Hall, L Smith, and P Zweidler-McKay* 107

Intracellular Pathways

SH3 and SH2: Prototypic Domains to Mediate Regulatory Mechanisms in the Cell *G Cesareni and D Peluso* 112

The MAPK Signaling Cascades *R Seger and S Wexler* 122

The PI3K/Akt/mTOR Pathway *C Evangelisti and AM Martelli* 128

The JAK-STAT-SOCS Signaling Cascade *JM Murphy, JJ Babon, NA Nicola, and SE Nicholson* 136

The PLC Pathway *KL Everett and M Katan* 153

Calcium and Calmodulin Signaling *E Carafoli and J Krebs* 161

Cellular Machineries

Macromolecular Communication between Nucleus and Cytoplasm	<i>AM Tartakoff</i>	170
The Molecular Architecture of Cell–Cell Adhesions	<i>B Geiger, R Zaidel-Bar, and M Vaman Rao</i>	181
Extracellular Regulation of Cell-to-Matrix Adhesion	<i>PA Janmey, CA McCulloch, and RT Miller</i>	192

Cellular Outcomes

Regulation of Cell Polarity	<i>JD Axelrod</i>	199
Regulation of Cell Migration	<i>MA McNiven and GL Razidlo</i>	208

Prototypic Integrative Processes

Wound Healing: An Orchestrated Process of Cell Cycle, Adhesion, and Signaling	<i>S Werner and M Antsiferova</i>	216
Cancer Cell Invasion through Tissue Barriers	<i>M Santiago-Medina, T Palmer, and J Yang</i>	223
Neuronal Transport and Spatial Signaling Mechanisms in Neural Repair	<i>AL Kalinski, T Hines, DS Smith, and JL Twiss</i>	230

Cell Division/Death**Regulation of Cell Growth**

mTORC1: Upstream and Downstream	<i>Y Yao, N Parnell, and K Inoki</i>	243
The Nucleolus	<i>AI Lamond, T Ly, S Hutten, and A Nicolas</i>	254
Ribosomes and Stress – Linked from Birth to Death	<i>JRP Knight, T Sbarrato, M Stoneley, and AE Willis</i>	270
Extra-Ribosome Functions of Ribosomal Proteins	<i>DA Franklin and Y Zhang</i>	281
Ribosomal RNA Processing	<i>K Umaer, M Ciganda, and N Williams</i>	288
Eukaryotic Ribosome Assembly and Export	<i>VG Panse and CS Weirich</i>	297
Internal Ribosome Entry Site-Mediated Translation	<i>N Sonenberg, Y Svitkin, and N Siddiqui</i>	307
Targeted mRNA Degradation	<i>S Djuranovic and HS Zaher</i>	317

Apoptosis

Apoptosis	<i>MF Navarro and G Salvesen</i>	327
Anoikis	<i>SM Frisch</i>	335
Mitochondria in Cell Death Regulation	<i>I Vitale, G Manic, G Kroemer, and L Galluzzi</i>	341
Tumor Necrosis Factor Signaling Pathways	<i>A Rajput and CF Ware</i>	354
Caspases	<i>J Puccini and S Kumar</i>	364
Efferocytosis in the Tumor Microenvironment	<i>DB Vaught, JC Stanford, and RS Cook</i>	374
The Bcl-2 Family Proteins: Insights into Their Mechanism of Action and Therapeutic Potential	<i>DW Andrews, H Brahmabhatt, and B Leber</i>	379
Inhibitor of Apoptosis Proteins, the Sentinels of Cell Death and Signaling	<i>M Cristina de Almagro and D Vucic</i>	390

Mitotic Catastrophe	<i>RYC Poon</i>	399
Autophagy	<i>AM Cuervo</i>	404

Cell Cycle

Interplay between Oncogenes and Tumor Suppressor Genes in Human Disease	<i>SJ Parsons, JO DaSilva, and J Thomas Parsons</i>	411
Cyclins and Cyclin-Dependent Kinases	<i>NH Lents and JJ Baldassare</i>	423
The Restriction Point	<i>JA DeCaprio</i>	432
CDK Inhibitors in Normal and Malignant Cells	<i>A Besser and J Slingerland</i>	437
The <i>INK4a</i> /ARF Locus	<i>DE Quelle, J Nteeba, and BW Darbro</i>	447
S Phase	<i>SJ Smith, RG Lingeman, CM Li, L Gu, RJ Hickey, and LH Malkas</i>	458
Traveling through Mitosis with the Chromosomal Passenger Complex	<i>ME Bekier II and WR Taylor</i>	469
Mitosis in Animal Cells	<i>ATY Lian and M Chircop</i>	478
Regulating Cytokinesis	<i>G Zachos</i>	494
Regulation of the p53 Pathway	<i>CL Brooks</i>	504
Cellular Senescence	<i>S Da Silva-Álvarez and M Collado</i>	511

Cellular Immunology

Overview

Introduction to Functional Cell Biology of Immunity	<i>K Cadwell and ML Dustin</i>	518
---	--------------------------------	-----

Transcriptional Basis of B and T Cell Lineages and Memory Cells

The T-Cell Receptor Signalosome	<i>M Mingueneau</i>	528
T Follicular Helper Cells	<i>MM Shannack and MA Linterman</i>	541
V(D)J Recombination: Orchestrating Diversity without Damage	<i>C Lescale and L Deriano</i>	550
T Cell Memory to Viral Infections	<i>P Klenerman and JR Fergusson</i>	567
Nuclear Factor of Activated T Cells and Tolerance	<i>F Macian</i>	573
NF-kappaB and the Immune System	<i>TD Gilmore, TW Siggers, and S Gerondakis</i>	580

Cell–Cell Interactions and the Immune System

Functional Specialization of Dendritic Cell Subsets	<i>GF Heidkamp, CHK Lehmann, L Heger, A Baransk, A Hoffmann, J Lühr, and D Dudziak</i>	588
CD28 Costimulation and Regulatory T Cells	<i>B Soskic and DM Sansom</i>	605
Roles of Stromal Cells in the Immune System	<i>EW Roberts, AE Denton, and DT Fearon</i>	616
CXCL12/SDF-1 and Hematopoiesis	<i>ML Capitano and HE Broxmeyer</i>	624
Lymphocyte–Endothelial Interactions	<i>CV Carman and R Martinelli</i>	632
T Cell Receptor Triggering	<i>B Alarcón and HM van Santen</i>	650
Single Molecule Methods to Measure Receptor–Ligand Interaction in Immunological Synapses	<i>JB Huppa and GJ Schütz</i>	660

Single-Cell Interrogation of the Immune System Using Microtools	<i>C Albayrak and S Tay</i>	667
Cellular Structures Controlling T Cell Signaling in Time and Space	<i>C Wülfing and KT Roybal</i>	677

Innate Immunity

Function of Epithelial Barriers	<i>K Cadwell and AM Marchiando</i>	687
Phagocytic Synapses	<i>HS Goodridge and N Hassanzadeh-Kiabi</i>	695
The Phagocyte NADPH Oxidase: Structure and Assembly of the Key Multicomponent Enzyme of Innate Immunity	<i>J El-Benna, P My-Chan Dang, and MB Yaffe</i>	702
Cytoplasmic Sensing in Innate Immunity	<i>C Hutcheon, P Paulvannan, and N Subramanian</i>	710
Scavenger Receptors	<i>S Gordon and C Neyen</i>	727
Dendritic Cells	<i>K Liu</i>	741
Neutrophil Biology	<i>C Carmona-Rivera and MJ Kaplan</i>	750
Specialized Subsets of Tissue-Resident Macrophages in Secondary Lymphoid Organs	<i>A Chauveau and TI Arnon</i>	759
CD4 T Cell Memory and Role of TNF Receptor Family	<i>DR Withers and CL Marriott</i>	770

Stem Cell Biology

Structure and Function

Stem Cells and Aging	<i>E Perdiguerro, L García-Prat, P Sousa-Victor, and P Muñoz-Cánoves</i>	776
Epigenetic Regulation of Stem Cells	<i>A Sacco and PL Puri</i>	785
The Adult Stem Cell Niche: Multiple Cellular Players in Tissue Homeostasis and Regeneration	<i>S Kyryachenko, L Formicola, D Ollitrault, R Correra, A-L Denizot, K Kyrylkova, G Marazzi, and DA Sassoon</i>	794
Cancer Stem Cells	<i>S Buczacki</i>	807
Metabolic and Energetic Regulation of Stem Cells	<i>A Esteves and T Finkel</i>	813
Pluripotent Cells: New Tools for Disease Research and Therapies	<i>TD Tlsty</i>	819

PREFACE

Cell biology is the study of cells, the integral unit that is the basis of all living tissues. In its earliest phases, this area of investigation was largely defined by microscopic-based observations of structure and organization that were, by their nature, largely descriptive. This also served to delineate cell biology from its nearest neighbor, biochemistry, which, dating from its post WWII renaissance, was driven in large part by reductionist approaches, where tissues, cells, and organelles were broken down in order to isolate and then characterize individual components. The melding force that bridged these two areas was molecular biology that allowed the cell biologist to transform pictures into functions and allowed the biochemist to reassemble their components into ensembles and subcellular entities. Not surprisingly the boundary between these two key biological sciences has become blurred beyond recognition. Today a considerable part of what passes as cell biology is by any other name biochemistry (or molecular biology – the distinction between these two areas being equally indistinct) and vice versa. Indeed, one of the first challenges we faced in organizing this Encyclopedia was how much molecular detail to include. It was our conclusion that for the sake of completeness, and to appropriately introduce the volumes Organizational Cell Biology and Functional Cell Biology, we needed a description of the molecular components of the cell, particularly with respect to the synthesis and degradation of proteins and nucleic acids, along with outlines of important chemical principles and key methodology. We were guided to some degree in these decisions by input from undergraduate and graduate teachers and the contents of their introductory courses that they shared with us and by discussions with our volume and section editors. The result was the first volume which we recognize is very 'biochemical,' but certainly lacking the more detailed coverage of the *Encyclopedia of Biological Chemistry* (Lennarz and Lane, 2004) or the many excellent biochemical texts that dot the educational landscape. Readers who find 'gaps' or incomplete treatments in this section will likely be able to find the additional detail they seek from these sources.

The second major decision that we had to confront was whether we would attempt to cover the full range of living organisms or whether we would place some limits on the scope of the material we included. In the final result, it was decided to place the emphasis on higher eukaryotes with only very prescribed treatments of lower eukaryotes and prokaryotes. Volumes 2 and 3 thus deal with material that has perhaps been more traditionally thought of as cell biology. Again we were guided in our decisions about what to include by teaching considerations, brainstorming sessions with our editors, and a perceived need to separate cellular organization from cellular function. However, as even a quick perusal will reveal, this is not a sharp delineation either. Volume 2 begins with a methodological treatment of imaging, which is basically divided between light and electron microscopic applications, and is followed by sections on organelles, interorganellar communication, and the cytoskeleton (including motors). The

last part of volume 2 is devoted to intracellular infection and covers a variety of external agents and pathogens that impact cell function as well as human pathology. Volume 3, in contrast, largely deals with major cell functions: signaling, cell-cycle control, and apoptosis. The last two parts are devoted to cells of the immune system and their responses and stem cells – two highly specialized niches characterized by unique cellular involvement.

The last part of the Encyclopedia deals with systems biology. This is really the newest area of cell biology and considers cellular involvement as part of a phenotypic response. As such, it strives to integrate all of the levels of the first three volumes in a consistent manner to produce a seamless description of cellular function. This is clearly a newly evolving area and the most rapidly changing part of cell biology; indeed it is where we expect to see the greatest change in the next few years.

Assembling a treatment of a topic as large as cell biology had multiple challenges. Coverage was a significant problem: on the one hand it was almost impossible to avoid some redundancy in places while, on the other, there were inevitable gaps. Some of these arose from late cancellations; others from oversights on our part. We can only promise to fill these in future editions. We also note that as can be expected for a large multiauthor compilation the individual articles do differ in detail and treatment. We felt it was more important to allow our experts substantial latitude in deciding how to present their topics than to apply rigid guidelines. We did try to limit the number of references (to make it easier for people not familiar with a topic to make the transition to more extensive articles) but there is admittedly some considerable variation in the bibliographic material as well.

When we were well into the project, we were approached by Graham Nisbet (Elsevier) about including the Encyclopedia in a larger collection entitled the Reference Module in Biomedical Sciences. The Reference Module plan was grouped in single integrated work information from several other compendia and it was our view that this would greatly leverage the value of our efforts (and those of our contributors). While we were the 'new kid on the block' in that we were still compiling the Encyclopedia and the other collections to be included were already extant, we felt that it was too good an opportunity to pass up. The Reference Module has since been launched and although it does not yet contain the complete content from the Encyclopedia, this will be achieved in the not too distant future. We consider that this will be a particularly valuable resource for researchers that are just entering (or changing into) this field.

No work of this size could be completed without the help and advice of many people. In this regard we would like to thank a number of people who were instrumental in bringing this project to fruition. First and foremost we would like to recognize our volume editors: Jerry Hart, Graça Raposo, Bruno Goud, Alan Ezekowitz, and Doug Lauffenburger; and our section editors: Ken Neet, Sarah Woodson, Judy Bond,

Elliot Elson, Tamotsu Yoshimori, Paul Gleeson, Anna Akhmanova, Yosef Yarden, Jason Weber, Michael Dustin, David Sassoon, Jason Haugh, Helen Byrne, and Rune Linding; and thank them for their invaluable input, enthusiasm, and hard work. They are truly a global group and have been a great team to work with. We also wish to thank the staff at Elsevier including Janice Audet, Peter Labella, Nicky Carter, Karen Maloney, Erin Hill-Parks, Rashmi Phadnis, Blerina Osmanaj, and Ginny Mills for their many efforts, perseverance (particularly with our idiosyncrasies), and skillful management of every aspect of this project.

In the course of planning this work, we gathered information from a number of sources. At the outset the publishers polled some 167 faculty from both undergraduate and graduate schools to ascertain their perceived interest in such a project, which also garnered considerable input about potential topics. The response was highly favorable. We also talked with many colleagues about experiences at their institutions and would like to particularly thank Catherine Bevier (Colby College), Robert Simoni (Stanford University), Larry Marsh (UC Irvine), and John Cooper (Washington University) for providing syllabi of germane courses taught in their institutions. Special thanks also go to Ruedi Abersold (ETH, Zurich) and Sergio Grinstein (University of Toronto) for their expertise that they readily shared with us, which was invaluable in planning certain sections.

Finally we would like to recognize all of our authors. We have been extraordinarily fortunate in attracting individuals from all around the globe to take time from their busy schedules to prepare this splendid set of contributions. Without these efforts there would be no Encyclopedia. Under ordinary circumstances it would be appropriate to dedicate the work to them and all the colleagues whose works they wrote about. However, during the preparation of this work, we sadly lost one of the most important contributors to cell biology, Tony Pawson. Tony was a true pioneer in elucidating basic mechanisms of how cells transmit information intracellularly, a fundamental knowledge that permeates all of cell biology. In an effort that was initiated by one of the section editors, Rune Linding, it was decided to dedicate the entire work to his memory as recognition of his contributions and a lasting tribute to a man who left us too soon.

We hope that the Encyclopedia will be of value to students and researchers alike and will become a useful tool in the training of future generations of scientists.

Ralph A Bradshaw and Philip D Stahl

References

Lennarz, W.J., Lane, M.D. (Eds.), 2004. Encyclopedia of Biological Chemistry. Oxford: Elsevier.

Cell Biology: An Overview

RA Bradshaw, University of California, Irvine, CA, USA

PD Stahl, Washington University School of Medicine, St. Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

All living organisms can be divided into three principal categories: archaeobacteria, eubacteria, and eukaryotes; although they differ in structure and organization, they are all composed of cells as the fundamental life unit. At the molecular level, there is also a great deal of similarity in the basic materials that make up these entities because they use the same kinds of molecules to store and reproduce information, to run the cellular metabolism and machinery and to provide the structural framework. Thus nucleic acids, proteins, lipid membranes, and carbohydrates – alone and in various combinations – are universally present, albeit in distinguishable forms, along with innumerable metabolites and ions. There are components that are apparently essential for life and are found in one form or another in all species and there are many unique moieties and associated activities that are highly specialized and are found in relatively few organisms. Indeed, the similarities have underpinned the development of our understanding of cellular function at a rudimentary level and the differences, basically engendered by evolution, have illustrated and delineated the complexity that speciation has introduced. Perhaps the largest of these differences is that which separates single cell organisms from multicellular organisms. The latter are exclusively eukaryotes while the former are composed of both eukaryotes and prokaryotes. The cellular organization and architecture that distinguishes these two major life forms is striking; although cell biology correctly embraces both, traditionally prokaryotic organisms have been the province of the microbiologists and the majority of cell biology research has been devoted to the eukaryotic world. In practical terms this translates for the most part into the study of human cells and those of easily maintained laboratory animals and selected paradigms, for example, fruit flies, worms, and zebra fish.

Human and animal cell biology is not a tightly proscribed science with well-defined borders. Basically it serves at the interface between biochemistry, molecular biology, and genetics, on the one hand, and anatomy and physiology, on the other. The continuum of these disciplines forms the core of the biomedical sciences, which also include the related but separate fields of pharmacology, microbiology, immunology, and pathology that provide the connections to disease and health. Cell biology has strong connections to all of these. There are also specialized areas, for example, neuroscience, that are of such importance that they warrant their own category and the cell biology associated with them is also highly specialized. Thus, cell biology is as complex as the enormous variety of cells that exist and achieving an accurate description of all of them in terms of their components and functions has long been a major part of the research in this field.

Imaging and Organelle Organization

Among the developments that propelled cell biology into the modern era are the introduction of the ultracentrifuge and

the rapid advances in microscopy both electron and light microscopy – the former allowing investigators to fractionate and characterize the components of the cell and the latter to literally see them – either *in situ* or in isolated form. This progress is illustrated by a series of Nobel Prizes starting in the early 1970s that chronicle the grand confluence of classical biochemistry and general physiology that created modern cell biology (Claude *et al.*, 1974) or the discovery and characterization of intracellular organelles – lysosomes and endoplasmic reticulum (ER) among others (Mitchell, 1978), for the chemiosmotic hypothesis (Brown and Goldstein, 1985), for endocytosis (Ciechanover *et al.*, 2004), for ubiquitination and protein degradation, and just recently (Rothman *et al.*, 2013), for vesicle trafficking, and (Betzig *et al.*, 2014) for advances in light microscopy. The early cell biologists insisted on quantitative application of these new techniques and laid the groundwork for modern cell biology. As with biochemistry, reductionism has been the keystone for the amazing success over these past several decades. Each organelle has its own research history – the nucleus, mitochondria, peroxisomes, the proteasome, the ER, and cytoskeleton. The Golgi apparatus/secretory pathway and the endosome–lysosome network have all been examined in detail both in isolation and in their relationships with each other. A new member, the exosome (aka extracellular vesicle) has emerged more recently and stimulated the imagination of aspiring young investigators (Harding *et al.*, 2013). The rise of genomics, the development of spectacular imaging modalities, and evolving biochemical techniques (proteomics) have led to the discovery of molecular motors, the identification of macromolecular complexes such as the exocyst, ESCRT, GARP, SNAREs among others that choreograph complex intracellular pathways, including cell motility and cytokinesis, have brought cell biology into this new era where the whole is now seen as larger than the individual components. This creates the segue from Parts 1 and 2 of the *Encyclopedia (Molecular Cell Biology and Organizational Cell Biology)* to Part 3 (*Functional Cell Biology*) where signaling modalities will be seen as an integrative force for regulation and control.

Signaling

While all organisms can sense their environment and respond to cues from it, multicellular organisms must in addition coordinate their responses, which require intercellular communication at a sophisticated level (Bradshaw and Dennis, 2009). The higher the development, the more complex these communication systems become. Thus the cell biologist must focus not only on how molecular function is translated into cell organization and how these functions are coordinated from organelle to organelle but also on the external interactions and signals that control the larger functional responses of organs and ultimately organisms.

Intercellular communication is afforded by stimuli that can be transmitted across the cell membrane, either directly or via membrane bound molecules that in turn pass signals to the intracellular compartment. The origins of these stimuli may be quite diverse. Among other things, they can include cell–cell contacts, soluble factors/messengers, or foreign agents. Lipophilic molecules can cross the membrane by diffusion or by facilitated transport to recognize and bind to intracellular entities but most substances bind to membrane bound proteins and induce their signal by ‘activating’ them instead. These so-called receptors are capable of producing a variety of responses that invariably involve the generation of posttranslational modifications of preexisting proteins. Protein phosphorylation, which in eukaryotes mostly occurs on tyrosine, serine, and threonine residues, is highly prevalent and appears to be directly or indirectly involved in almost all transmembrane signaling (Gnad *et al.*, 2011). However, it is by no means the only vehicle for transmitting information as acetylation, methylation, monoglycosylation, and ubiquitination, to name a few, are both important and widespread. There are over three hundred different covalent modifications (Khoury *et al.*, 2011) that are introduced into proteins (and many more that have not yet been chemically defined) of both a transient and stable nature and most are likely to be involved to some extent in signal transduction processes. In addition, limited proteolysis can also be an important part of a pathway and this activity is a major player in cellular responses (Turk *et al.*, 2012).

The induction of an intracellular signal is basically perpetrated by the formation of new interactions between the modified proteins and other entities, which can range from small molecules to macromolecular protein complexes. One of the most important advances in understanding how cells process information induced from external stimuli was the appreciation that the newly formed sites produced by the protein modifications were recognized by other proteins with modules specifically designed for this purpose (Pawson, 1995). Thus, for example, phosphotyrosine residues could be recognized by other proteins containing a domain, termed SH2 for its relatedness to a similar domain found in the Src protein kinase, which could then be further modified allowing for additional interactions to occur. By these ‘docking’ events, signaling complexes could be assembled, often in multiple steps, that ultimately lead to the activation of key effectors. The end point of many of the pathways is the activation of transcription factors that lead to the modulation of gene expression of that cell by ultimately changing its protein expression profile (Bradshaw and Dennis, 2009). However, many molecules are usually modified and activated ‘along the way’ and these ‘new’ activities also add to the overall response to the original stimuli. Short term responses indeed require that the protein effectors necessary for it are already present; long term responses per force require new protein synthesis.

Signal transduction is thus dependent on two phenomena: posttranslational modifications, particularly of the readily reversible type, and protein–protein interactions. The extent to which these two activities take place, even in resting, unstimulated cells was greatly underestimated before the advent of proteomics and the introduction of mass spectrometric methodology into cell biological research (Bradshaw and

Burlingame, 2005) These high-throughput unbiased analyses of very complex mixtures, basically derived directly from cell lysates, revealed that essentially every protein had multiple interacting partners and that posttranslational modifications affected a very significant proportion of the proteins present. These were profound differences from what had been the prevailing wisdom and amounted to a paradigm shift in thinking about cellular organization and structure. As noted above, these same tools have gone on to cast new light on organelle organization in terms of resident proteins and have helped to disclose the structure of important cellular machines, such as the proteasome (Voges *et al.*, 1999), the nuclear pore (Routa *et al.*, 2000), and transcription complexes (Kornberg, 2007). They have also begun to elucidate the complexity of epigenetic regulation of gene transcription at the histone level (Cosgrove, 2012), which will also be essential in completing our understanding of signaling processes.

Concluding Remarks

Two of the most fundamental aspects of cells are their ability to reproduce themselves and, in higher organisms, to undergo changes that lead to new cell types and functions. These developmental processes leading to differentiation are what allow a single fertilized egg to form a complex adult organism and require the synthesis of all aspects of cell biology to understand. Knowing how cells go from one state to another in a timed and regulated fashion forms the core of developmental biology, which can be viewed as a sub discipline of cell biology. Also of major importance is the maintenance of viability and the associated turnover processes. The control of cell death is not only an essential part of development it is also key to managing situations that have gone awry and underlies many serious disease conditions.

It is clear that science is still a long way from understanding how even simple cells work. There is a long list of functions and structures that remain to be elucidated and integrating the various experimental approaches and the data they produce still lags far behind. This is the ‘Era of Big Data’ and vast amounts of new information that impact on our understanding of cells and their processes are collected every day. Genomic information is a good example – despite the rapidity that this sort of information can now be obtained, it still remains to be effectively mined in terms of what it can tell us about basic principles of how cells work. There has been an understandable pressure to apply this information to managing disease (Collins and Varmus, 2015), particularly those that are life threatening, but there certainly is information that is yet to be formulated that would apply to fundamental problems as well. The same can be said for transcriptomics, proteomics, and metabolomics. In addition to these powerful new technologies, singular advances in analyzing single cells promises to augment these molecular approaches significantly (Di Palma and Bodenmiller, 2015).

One aspect of cell biology that is looking to address the challenges of integrating these relatively newly minted studies is systems biology (Part 4). It is far too early to assess the value of these efforts in coordinating this flood of information but it certainly looks promising. It will be an elusive goal for some

time to come but also a stimulus for new generations of cell biologists that will be forthcoming.

References

- Betzig, E., Hell, S.W., Moerner, W.E., 2014. The Nobel Prize in chemistry — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2014/press.html (accessed 03.04.15).
- Bradshaw, R.A., Burlingame, A.L., 2005. From proteins to proteomics. *IUBMB Life* 57, 267–272.
- Bradshaw, R.A., Dennis, E.A. (Eds.), 2009. *Handbook of Cell Signaling*, second ed., vol. 1–3. San Diego, CA: Elsevier Academic Press, pp. 1–3047.
- Brown, M.S., Goldstein, J.L., 1985. Physiology or medicine — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1985/press.html (accessed 03.04.15).
- Ciechanover, A., Hershko, A., Rose, I., 2004. The Nobel Prize in chemistry — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2004/press.html (accessed 03.04.15).
- Claude, A., de Duve, C., Palade, G.E., 1974. Physiology or medicine — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1974/press.html (accessed 03.04.15).
- Collins, F.S., Varmus, H.A., 2015. New initiative on precision medicine. *New England Journal of Medicine* 372, 793–795.
- Cosgrove, M.S., 2012. Writers and readers: Deconvoluting the harmonic complexity of the histone code. *Nature Structural & Molecular Biology* 19, 739–740.
- Di Palma, S., Bodenmiller, B., 2015. Unraveling cell populations in tumors by single-cell mass cytometry. *Current Opinion in Biotechnology* 31, 122–129.
- Gnad, F., Gunawardena, J., Mann, M., 2011. PHOSIDA 2011: The posttranslational modification database. *Nucleic Acids Research* 39 (suppl. 1), D253–D260.
- Harding, C.V., Heuser, J.E., Stahl, P., 2013. Exosomes: Looking back three decades and into the future. *Journal of Cell Biology* 200, 367–371.
- Khoury, G.A., Baliban, R.C., Floudas, C.A., 2011. Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database. *Scientific Reports* 1, 90. doi:10.1038/srep00090.
- Kornberg, R.D., 2007. The molecular basis of eukaryotic transcription. *Proceedings of the National Academy of Sciences of the United States of America* 104, 12955–12961.
- Mitchell, P., 1978. The Nobel Prize in chemistry — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1978/press.html (accessed 03.04.15).
- Pawson, T., 1995. Protein modules and signaling networks. *Nature* 373, 573–580.
- Rothman, J.E., Schekman, R.W., Südhof, T.C., 2013. The 2013 Nobel Prize in physiology or medicine — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/2013/press.html (accessed 03.04.15).
- Rout, M.P., Aitchison, J.D., Supratka, A., *et al.*, 2000. The yeast nuclear pore complex — Composition, architecture, and transport mechanism. *Journal of Cell Biology* 148, 635–652.
- Türk, B., Türk, D., Türk, V., 2012. Protease signalling: The cutting edge. *EMBO Journal* 31, 1630–1643.
- Voges, D., Zwickl, P., Baumeister, W., 1999. The 26S proteasome: A molecular machine designed for controlled proteolysis. *Annual Review of Biochemistry* 68, 1015–1068.

Functional Cell Biology: An Overview

JD Weber, Washington University, St. Louis, MO, USA

Y Yarden, Weizmann Institute of Science, Rehovot, Israel

M Dustin, University of Oxford, Headington, Oxford, UK

DA Sassoon, Université de Pierre et Marie Curie-Sorbonne Universités, Paris, France

A Ezekowitz, Abide Therapeutics, San Diego, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

This section of the Encyclopedia addresses fundamental questions as they pertain to the physiology of cellular homeostasis. Basic questions about the regulation of cell growth and cell content are discussed. How transformation of cells is a result of misdirected homeostasis that affects selected cellular functions and diverts others. This intrinsic threat is contrasted to external threats posed by infectious agents that rouse the innate immune system, which in turn alerts the adaptive immune system, if required. Host defense might disrupt tissue continuity and function. How and when repair can evoke regeneration of tissue is discussed in the section on stem cell biology. The narratives underscore an elegant complexity that underlies the preservation of health and provides insights into the mechanism of disease.

Cell Signaling and Intercellular Communication

When unicellular creatures first assembled to generate the cooperatives we now call multicellular organisms, they inevitably handed over the lion share of their autonomy. The new biochemical 'language' that had concomitantly evolved confers to metazoans an amazingly complex and busy intercellular communication. This communication is crucial for benefitting the synergy offered by many collective cellular effects, which constantly establish sophisticated organs and whole-body systems, such as the immune system. Understanding the linguistic nuances of cell-to-cell interactions is not only interesting and informative but also very useful, for many diseases are rooted in miscommunication and unauthorized autonomy. This part of the *Encyclopedia of Cell Biology* teaches and exemplifies the vocabulary of intercellular communication. The 'receptors' receive special attention; several families of these molecular devices exist in mammals and they all come with two faces: the extracellular face receives a cue, in most cases a molecule carrying a physiological message, such as a growth factor. The receptor's part facing the cytoplasm instigates a biochemical pathway (a.k.a., signal transduction), variably mediated by posttranslational protein modifications, elevation of second messenger molecules (e.g., cyclic adenosine monophosphate or calcium ions), or formation of new protein complexes. Unlike the multiple extracellular cues and their many receptors, only a few 'signal transduction pathways' exist in eukaryotic cells. And while the principal pathways are reviewed in this part, it is becoming increasingly clearer that several pathways are simultaneously activated by each receptor. The identity and relative strength of such pathway combinations are decoded intracellularly by the transcriptional

machinery, adhesive properties of the cell surface, the endocytic pathway, and other 'cellular machineries,' as exemplified by several articles. To demonstrate 'cellular outcomes' of the enormous complexity of intercellular communication and the associated clinical implications, this part culminates by highlighting wound healing, colonization of distant organs by cancer cells (metastasis), and neural repair mechanisms.

Receptors

Intercellular communication may be likened to a large funnel: multiple polypeptide and other messengers, in the form of cytokines and growth factors, mediate cellular interactions by binding to several receptors, each able to bind with a few ligands and couple their biochemical messages to still fewer cytoplasmic signaling pathways. Accordingly, the cell surface receptors, which fall into a few families and subfamilies, carry out a critical function as translators of extracellular cues into intracellular signals. One important family, receptor tyrosine kinases (RTKs; see [Receptor Tyrosine Kinases and Their Ligands](#)), comprises 58 members and 20 subfamilies. Ligand binding promotes dimerization of RTKs, which causes activation of their kinases and autophosphorylation, thereby providing docking sites for downstream signaling molecules with phosphotyrosine binding specificity (e.g., SH2 domains). Several hormones, interleukins, interferons, and colony stimulating factors bind with other transmembrane receptors, which lack an intrinsic tyrosine kinase function and belong to the cytokine receptor family (see [Cytokine Receptors and Their Ligands](#)). Like the RTKs, receptors for the transforming growth factor beta (TGF- β) family of ligands harbor protein kinase activity, albeit with specificity to residues other than tyrosine, and they transmit signals through a dedicated group of adaptors (see [TGF- \$\beta\$ Superfamily Signaling](#)). Unlike the above-listed single-pass receptors, the largest family of receptors, G protein-coupled receptors (GPCRs), includes molecules that transverse the plasma membrane seven times (see [G Protein-Coupled Receptors](#)). GPCRs generate intracellular second messengers and also engage intracellular signaling pathways (through arrestins). Two additional, smaller groups of receptors, guanylyl cyclase receptors (see [Guanylyl Cyclase Receptors](#)) and receptors for tumor necrosis factors (TNFs; see [Tumor Necrosis Factor Receptors: A Brief Digestion](#)), are also described.

The Signal Transduction Pathways

Although it is already clear that cascades of proteins engaged by each type of ligand-occupied receptors often bifurcate and channel their signal laterally and even backward, these

pathways are commonly described as linear and unidirectional. Perhaps the best understood pathways are those recruited, by means of adaptor proteins containing SH2 or SH3 domains (see [SH3 and SH2: Prototypic Domains to Mediate Regulatory Mechanisms in the Cell](#)), to ligand-activated RTKs and cytokine receptors. These are the mitogen-activated protein kinase (MAPK) route (see [The MAPK Signaling Cascades](#)), the PI3K-AKT-mTOR pathway (see [The PI3K/Akt/mTOR Pathway](#)), the JAK-STAT (see [The JAK-STAT-SOCS Signaling Cascade](#)), the phospholipase C (PLC) to protein kinase C (PKC) (see [The PLC Pathway](#)) cascade and the calcium-calmodulin pathway (see [Calcium and Calmodulin Signaling](#)). Because several signaling pathways regulate bidirectional transport of proteins between the cytoplasm and nucleus, which occurs through nuclear pores, this rather common aspect is also reviewed (see [Macromolecular Communication between Nucleus and Cytoplasm](#)). Additional, more restricted pathways, include sonic hedgehog (SHH), which enables morphogens to pattern tissues such as the neural tube and limb bud by regulating processing of master transcription factors (see [Hedgehog Signaling in Development and Disease](#)). The Hippo signaling pathway similarly regulates organ size and tissue regeneration by means of a kinase cascade and retention of a complex of two transcription repressors in the cytoplasm (see [The Hippo Pathway](#)). As befits its many roles in tissue homeostasis, the Hippo pathway interacts with other signaling players, such as the Notch and WNT pathways. In mammals the Notch pathway consists of four receptors and five ligands (see [The Notch Pathway](#)). These ligands, expressed on the surface of a signal-sending cell, bind to and activate Notch receptors on neighboring cells. Ligand binding instigates receptor cleavage followed by translocation of the intracellular domain of Notch to the nucleus and formation of an activating complex that up-regulates target gene expression. In analogy, a major effector of the canonical WNT signaling pathway is the transcription factor β -catenin (see [The Wnt \$\beta\$ -Catenin Pathway](#)). In the absence of ligand, β -catenin interacts with two scaffold proteins and is a substrate for two kinases. Phosphorylated β -catenin is then ubiquitinated and destroyed. When WNT ligand binds to a Frizzled family receptor, the multiple component destruction complex is inhibited, leading to the stabilization of β -catenin and its translocation to the nucleus, where it interacts with TCF/LEF family transcription factors, so target gene expression is activated. In contrast to the WNT, Hippo, and Notch pathways, which culminate in transcription regulation, signaling downstream of death receptors, like TNF-receptors, initiates programmed cell death by means of activating multiple cytoplasmic proteases of the caspase family (see [Signaling and Function of Death Receptors of the Tumor Necrosis Factor Receptor Superfamily](#)).

Cellular Machineries

One cellular machinery that both regulates signal transduction and is profoundly regulated by intracellular signals is endocytosis of both surface proteins and extracellular, soluble molecules (see [Endocytosis of Cargo Proteins: LDL in Part 2](#)). Although the selection of cargo is wide, the principles

underlying endocytosis are quite common. The article included herein describes how cholesterol circulates in the form of low-density lipoprotein (LDL) particles, which must be internalized to enable entry into cellular membrane pools. On its own, endocytosis regulates cell polarity, which is vital to the function of specialized cells of the epithelium, immune, and nervous systems (see [Regulation of Cell Polarity](#)). Endocytosis regulates yet another machinery, which ensures cell-cell adhesion and employs diverse types of intercellular junctions, including cadherin-based junctions (see [The Molecular Architecture of Cell-Cell Adhesions](#)). The related machinery, namely cell adhesion to extracellular matrices, involves specific bonds between membrane proteins and their extracellular ligands (see [Extracellular Regulation of Cell-to-Matrix Adhesion](#)). This type of cell adhesion is altered by changes in the chemical or physical properties of the matrix, which critically affect normal development and pathology. How cellular machineries, such as adhesion, endocytosis, and cytoskeletal reorganization, are functionally integrated is well exemplified by cellular motility (see [Regulation of Cell Migration](#)). This process is heavily regulated at multiple steps of the migratory process, to enable control of development, wound healing, and immunity.

Cellular Outcomes

The impact of cell-to-cell communication on pathology is herein exemplified by the process that heals skin wounds (see [Wound Healing: An Orchestrated Process of Cell Cycle, Adhesion, and Signaling](#)). Stepwise blood clotting and inflammation, new tissue formation and, eventually, a long tissue remodeling process, engage many signaling pathways and some of the above-described cellular machineries. Even more life threatening and widely spread in the body is the metastatic process, in which cancer cells disseminate from their primary sites to form secondary tumors throughout the body (see [Cancer Cell Invasion through Tissue Barriers](#)). The initial step of tumor cell dissemination is followed by invasion into the surrounding tissue, subsequent intravasation into the circulatory system and, finally, extravasation at a distant site to instigate organ colonization. Intercellular communication between highly specialized cells, such as neurons, requires very long axonal processes connecting neurons to their targets (see [Neuronal Transport and Spatial Signaling Mechanisms in Neural Repair](#)). Injury of axons disrupts this intercellular communication, such that neural function can only be restored by regeneration of axons and their target reinnervation. Regeneration requires growth factors, receptors, intracellular trafficking of molecules, and organelles, as well as components of the neuronal cytoskeleton, all supporting growth of the injured axon.

Cell Growth

One of the outstanding fundamental questions in cell biology concerns how cells coordinate cellular growth (or macromolecular synthesis) with cell cycle progression and mitosis. Intuitively, rapidly dividing cells must have some control over these processes; if not, cell would continue to shrink in volume

with every passing cycle, similar to the cytoreductive divisions seen in the very early stages of embryogenesis. The problem is more easily solved in unicellular organisms, such as yeast, as their growth rates are entirely dependent on nutrient availability. Multicellular organisms must have acquired additional levels of control, as nutrient availability is seldom an issue and the organism has a prodigious capacity to store necessary metabolites in the form of glycogen, lipids, and protein. While myriad mechanisms of preventing unrestricted cellular proliferation exist via initiating programmed cell death (e.g., apoptosis, autophagy, necrosis), these mechanisms all depend on some external cue like those emanating from hypoxia, death signaling ligands, and lack of nutrients or survival signals. In the following articles, we will introduce the principle processes involved in cellular growth and macromolecular synthesis. Subsequent articles will then clarify the contrasting conditions of cell death and cell cycle progression and how cells are poised to choose between these two states in order to grow or die.

Ribosome Biogenesis Pathways

An essential feature of cell growth regulation is the tight coupling between ribosome synthesis and extracellular conditions. Although macromolecular synthesis is distinct from cell cycle progression and division, the two processes are coordinated. Cell mass and DNA content both double during each cell division, thus ensuring that cells remain a constant size under proliferative conditions. In yeast and mammalian systems target of rapamycin (TOR) is a key regulator of these two processes, providing the bridge between growth and proliferation. TOR was first identified by genetic screens in yeast aimed at identifying the protein TOR, a potent immunosuppressive macrolide produced by *Streptomyces hygroscopicus*. The mechanism of action for rapamycin is conserved in eukaryotes, with single TOR orthologs identified in *Caenorhabditis elegans*, *Drosophila melanogaster*, and mammals, and will be discussed in the first article of this section (see [mTORC1: Upstream and Downstream](#)).

In recent years, it has become apparent that a major function of TOR is to connect incoming environmental signals to ribosome production. It is estimated that approximately 50% of the cell's energy expenditure is directed toward ribosomal biogenesis. Given these costly energy projections, the cell has adopted stringent criteria that must be met before engaging the ribosome biogenesis machinery. As such, TOR is able to link these criteria to the cellular home of ribosome biogenesis, the nucleolus (see [The Nucleolus](#)). The nucleolus, long recognized as a marker for active cell growth, was first described as the center of ribosomal biogenesis and rDNA transcription in the early 1960s. This membrane-free organelle is composed of three regions on the basis of morphology at the ultrastructural level as well as the function of each distinct region. The nucleolus is sensitive to both positive and negative regulators of growth: ribosomal proteins themselves can act as signaling proteins often operating as communicators of stress to alter cell growth patterns in mammals (see [Ribosomes and Stress – Linked from Birth to Death](#)). Intriguingly, ribosomal proteins themselves may function as bona fide tumor suppressors. It is

well appreciated that ribosomal proteins possess extra-ribosomal functions in DNA damage responses and cell cycle regulation. Recently, several ribosomal proteins were found to antagonize the MDM2 oncogene, causing an induction of a p53-dependent cell cycle arrest. Thus, perturbations of ribosomal biogenesis or nucleolar architecture may lead to release of ribosomal proteins into the nucleoplasm where they might act in non-ribosome capacities (see [Extra-Ribosome Functions of Ribosomal Proteins](#)).

Ribosomal biogenesis begins with the Pol I-dependent transcription of the tandemly repeated rDNA loci on the five acrocentric chromosomes in humans. The product of Pol I transcription is the 47S pre-rRNA which, much like its mRNA counterparts, must be properly processed to yield functional rRNA and subsequent ribosomes (see [Ribosomal RNA Processing](#)). This consists of a series of base modifications (pseudouridylation and 2'-O-methylation) followed by an ordered cleavage by processing RNases to yield the 28S, 18S, and 5.8S rRNA species. Mature eukaryotic ribosomes have a small 40S and large 60S subunit, both of which are assembled within the nucleolus and exported into the cytoplasm. The large subunits are composed of the 28S, 5.8S, and 5S (made in the nucleoplasm) rRNAs plus large associated ribosomal proteins (e.g., L5, L7, etc.) while the small subunits contain the 18S rRNA and small ribosomal proteins (e.g., S3, S6, etc.). In all, these four rRNAs must assemble with more than 70 ribosomal proteins to undergo a series of maturation steps before gaining full functionality (see [Eukaryotic Ribosome Assembly and Export](#)).

Translation

Regulation of translation functions as an important biological mode to modulate gene and protein expression. The translational process can be separated into three phases: (1) initiation, (2) elongation, and (3) termination. While a limited number of factors are involved in elongation and termination, numerous protein-mRNA and protein-protein interactions are necessary for proper initiation. Indeed, more than 25 proteins have been implicated in eukaryotic initiation and at least 11 initiation factors are necessary for proper translation initiation. Thus, it is not surprising that initiation appears to be the rate-limiting step in translation for nearly all eukaryotic mRNAs. As such, translation initiation is the most frequent target of translational regulation. The majority of eukaryotic mRNAs are capped at their 5' ends by a 7-methyl guanosine moiety. A heterotrimeric protein complex known as eIF4F, which is the core of cap-dependent translation initiation, binds this 5' cap. However, a second independent initiation mechanism exists that utilizes unique internal ribosome entry sites (IRES) elements to stimulate initiation (see [Internal Ribosome Entry Site-Mediated Translation](#)). IRES elements are highly structured RNA sequences that were originally discovered in viral RNA genomes and have since been identified in a subset of mammalian mRNAs. IRES elements are defined exclusively by functional criteria rather than primary sequence, making their identification difficult. Unlike cap-dependent translation, the mechanisms of IRES-induced ribosome recruitment are only beginning to be understood. Differential

requirement of eIFs by distinct ORES elements provides support for the likelihood of mechanistic variability. In normal cells, global translation is thought to be cap-dependent while certain circumstances can lead to an increase in IRES-dependent translation. This variability has led many to hypothesize that selective translation plays an important role in how cells respond to external cues or stresses (see [Targeted mRNA Degradation](#)).

Cell Death Regulation

Programmed cell death is an organismal and cellular necessity, effectively removing damaged or unwanted cells in a process that minimizes damage to other nearby cells. Apoptosis involves both an extrinsic and intrinsic pathway that sets in motion programmed cell death (see [Apoptosis](#)). Notably, cross talk between the two pathways is possible since both share some common elements. Each cascading pathway contains both pro-apoptotic and anti-apoptotic components whose opposing activities ultimately decide the cell's fate. Specifically, caspases are thought of as the executioners of apoptosis. They consist of a family of cysteine proteases that deliver and execute apoptotic stimuli through their limited proteolysis activity (see [Caspases](#)). Conversely, the BCL family of proteins consists of three subfamilies associated with controlling mitochondrial apoptosis (see [The Bcl-2 Family Proteins: Insights into Their Mechanism of Action and Therapeutic Potential](#)). These are readily divided into anti-apoptotic, pro-apoptotic, and BH3-only proteins. This complicated family division results in a multifaceted network of antagonizing outcomes that either protect or promote apoptosis in response to upstream signals. Inhibitor of apoptosis (IAPs) family of proteins adds an additional layer of apoptotic regulation by inhibiting the function of caspases (see [Inhibitor of Apoptosis Proteins, the Sentinels of Cell Death and Signaling](#)). The eight IAP family members, not surprisingly, appear to play a central role in cancer progression.

The cell mitochondria have always played a central role in producing the majority of intracellular ATP and in participating in many anabolic processes that are required for cell survival. However, additional functions have been ascribed to mitochondria based on its role in various signaling pathways that actively provoke cell death (see [Mitochondria in Cell Death Regulation](#)). As stated above, both intrinsic and extrinsic signals are capable of eliciting a mitochondrial apoptotic response. In this manner, the mitochondria serve as a central location for both pro- and anti-apoptotic components.

Extrinsic cell death signals can come in the form of extracellular ligands or other receptor-mediated interactions. This implies that neighboring cells have the capacity to influence the programmed cell death of other cells either during development or to prevent local tumorigenic invasion. In the latter case, the elimination of tumor cells through induced apoptosis by neighboring cells or through changes in its own micro-environment interactions could serve as a potent antitumor mechanism. One such mechanism is termed anoikis and results when a normal epithelial cell detaches from the extracellular matrix or attaches to an inappropriate type of matrix

(see [Anoikis](#)). The Anoikis response provides protection from tumor cells that might detach from their local environment in order to become metastatic. TNF is a far more complicated ligand that is capable of eliciting both proliferative and apoptosis pathways upon binding to the TNF receptor (see [Tumor Necrosis Factor Signaling Pathways](#)). However, its ability to activate caspases under certain conditions makes it a potent inducer of cell death, particular during viral infection.

When cells are poised to die, their decaying contents must be dealt with by surrounding cells. Recent evidence has pointed toward an equally effective form of cellular recycling termed autophagy. In this setting, intracellular components are degraded in lysosomes to be recycled. This turnover allows for proteins and organelles to be renewed and also serves as a mechanism for quality control of newly synthesized cellular structures (see [Autophagy](#)). Additionally, organ homeostasis is often maintained by the rapid removal of newly apoptotic cells through efferocytosis. In order to prevent the dispersal of internal self-antigens to the extracellular environment, dying cells must be removed in an efficient manner and to promote healing of the surrounding tissue. Efferocytosis is the organism's natural response to both remove dying cells and to promote wound healing (see [Efferocytosis in the Tumor Microenvironment](#)). Not surprisingly, the local environment of the dying cell plays a central role in stimulating its engulfment and clearance.

Cell Cycle Progression

The ability of a cell to achieve an appropriate balance between cell division and arrest requires the coordinated function of primarily two classes of proteins: proto-oncogenes and tumor suppressors. The heightened state of dependence of malignant cells on oncogenes and on the inactivation of tumor suppressor genes underscores the importance of both classes in cancer development. Thus, an understanding of the functional interplay between proto-oncogenes and tumor suppressors under normal conditions as well as how perturbations that offset this balance trigger tumorigenesis is the focus of the first article on this topic (see [Interplay between Oncogenes and Tumor Suppressor Genes in Human Disease](#)).

Both proto-oncogenes and tumor suppressor proteins largely function to drive and stop the cell cycle, respectively. The mammalian cell cycle can be temporally partitioned into four segments: G1, S, G2, and M phases. Alternatively a fifth G0 phase is often termed to describe a quiescent, non-proliferative state into which the cell withdraws upon cessation of growth factor signaling or in response to growth inhibitory cues. During the initial G1 gap phase, the cell is primed by mitogenic stimuli in preparation for the synthesis of an exact replicate of its genetic material during S phase that will be distributed to its resultant daughter cell upon mitotic division in M phase. The second G2 gap phase serves a similar preparatory function as G1, checking the fidelity of replicated DNA and positioning the cell for its proper transition into M phase. Early in G1, external mitogenic stimuli activate several kinase cascades that promote the synthesis and stability of D-type cyclins which partner with cyclin-dependent kinases (CDKs) to drive the cell's progression through G1 (see [Cyclins](#)

and Cyclin-Dependent Kinases). As cells progress through the cell cycle other combinations of cyclins and CDKs (cyclin E-CDK2, cyclin A-CDK2, cyclin B-CDK1) act to further drive events associated with each cell cycle phase. Arguably, one of the most crucial events in the cell cycle is the hyperphosphorylation and inactivation of the retinoblastoma protein (Rb) at the restriction point (see [The Restriction Point](#)). Hyperphosphorylation of Rb derepresses E2F target genes whose transcription is required for the efficient transition from G1 into S phase and subsequent DNA synthesis during S phase (see [S Phase](#)). By maintaining the hyperphosphorylated state of Rb throughout M phase, cyclin-CDK complexes ensure the cell's committed passage through the restriction point to the completion of mitosis.

Two classes of CDK inhibitors (CKIs), the INK4 (p15, p16, p18, and p19) and the CIP/KIP (p21, p27, and p57) families provide an additional layer of regulation to the cell cycle (see [CDK Inhibitors in Normal and Malignant Cells](#)). CKIs are activated in response to either DNA damage or hyperproliferative signals and inhibit the CDK4/6 (INK4 family) or CDK2 (CIP/KIP family) holoenzymes. Whereas the INK4 family members block kinase activity by impairing the binding of CDK4/6 to D-type cyclins, the CIP/KIP proteins form heterotrimeric complexes with partnered cyclins and CDKs to impair their kinase activity. As all of the CKI genes display tumor suppressive functions *in vitro* and *in vivo*, clearly the cell's commitment to replicate and divide is a pivotal decision that is targeted for deregulation in the process of tumorigenesis.

A functionally independent tumor suppressor was identified in 1995 residing between the *INK4A* and *INK4B* tumor suppressor genes on human chromosome 9p21. This gene shares a common second and third exon with *INK4A* but is encoded by a distinct exon 1. The resulting mRNA splices in an alternative reading frame yielding an ARF tumor suppressor protein. Since *INK4A* and ARF are translated in alternative reading frames, they share no resemblance in their protein structure or their cellular function. The ARF tumor suppressor protein largely resides in the nucleolus and binds to the MDM2 proto-oncogene product to indirectly activate the p53 tumor suppressor (see [The INK4a/ARF Locus](#)). Over the last decade, ARF has been shown to also maintain cellular homeostasis through its ability to interact with numerous other proteins independent of p53, many of which have known roles in ribosome biogenesis. This provides a more than coincidental link between cell proliferation and cell growth pathways, squarely placing ARF at the nexus of the two processes. Likewise, p53 has long been labeled the guardian of the genome because of its role in preserving the stability of the genome. The p53 tumor suppressor is a key sensory molecule that regulates a plethora of downstream targets capable of triggering cell cycle arrest, apoptosis, senescence, DNA repair, and autophagy among many others in response to robust cellular stress signals (see [Regulation of the p53 Pathway](#)). While we have learned much about these tumor suppressors in the past decade, their additive list of cellular functions appears nearly limitless.

As proliferating cells travel into mitosis they must ensure the proper segregation of chromosomes prior to completion of the cell cycle. One of the main components of this checkpoint involves the function of the chromosomal passenger complex

(CPC) whose main role is to ensure proper chromosome inheritance in daughter cells (see [Traveling through Mitosis with the Chromosomal Passenger Complex](#)). As cells enter mitosis, this complex associates with centromeres where it monitors the attachment of newly formed spindles to kinetochores, fully capable of destabilizing and removing those attachments that could cause chromosome mis-segregation. Additionally the CPC must properly oversee chromosome positioning to ensure that no chromosomes are trapped in the cleavage furrow and in doing so, the CPC ultimately acts as a guardian of chromosome segregation. Any failure to properly regulate the six stages of mitosis, prophase, prometaphase, metaphase, anaphase, telophase, and cytokinesis could affect the tumorigenic potential of the cell (see [Mitosis in Animal Cells](#)). Moreover, the final step of mitosis is termed cytokinesis in which the cytoplasm of the mother cell is properly partitioned into two daughter cells. Cytoskeletal, membrane trafficking, and cell cycle pathways tightly coordinate the steps leading to the final cleavage furrow to ensure proper cleavage (see [Regulating Cytokinesis](#)). A failure to properly regulate cytokinesis leads to a genetically unstable cell ploidy that can lead to tumorigenesis. Not surprisingly, cancer cells often harbor an altered mitotic program and this has led to an increased focus on identifying mitotic components that could be used to treat cancer. The normal cell is programmed to undergo mitotic catastrophe in order to eliminate cells that have been irreversibly damaged (see [Mitotic Catastrophe](#)). While numerous chromosomal conditions can trigger mitotic catastrophe, cells that evade this process contribute to genome instability and tumor progression. However, using this strategy, many anticancer agents invoke mitotic catastrophe by damaging DNA beyond repair, demonstrating the utility of inducing this mitotic event to treat cancer.

We have seen cells that enter the cell cycle often encounter and interpret a multitude of signals in order to either properly transit the cell cycle or be removed from it. Cellular senescence is a state of permanent cell cycle arrest that was initially observed in cells grown in cell culture (see [Cellular Senescence](#)). Future experiments rapidly redefined the state, arguing that cellular senescence could be seen in tissues, particularly those in the aged. These observations led many to relate cell senescence with the process of aging and while that is certainly true, current viewpoints have placed an emphasis on the role of senescence in response to oncogenic stimulation. In this manner, cellular senescence is seen as a direct response to oncogenic insult, an opposing response to oncogenic transformation.

Functional Cell Biology of Immunity

The battle for homeostasis in the face of pathogens generates intense evolutionary pressure on immune defenses, pushing the boundaries of cell biology. The defensive strategies have been characterized as innate for recognition of evolutionarily conserved signatures of injury and infection, and adaptive when unique receptor repertoires are developed to deal with genetic adaptation of pathogens. A first level of innate immunity is the tissue barriers composed of epithelial coverings (see [Function of Epithelial Barriers](#)) and mesenchymal

connective tissue (see [Roles of Stromal Cells in the Immune System](#)) that together protect the integrity of the organism and provide pathways through tissues for defensive cells. Another type of innate immunity is based on chemical defences, such as the NADPH oxidase of neutrophils and phagocytes that specialize in generation of toxic reactive oxygen species in a tightly controlled manner (see [The Phagocyte NADPH Oxidase: Structure and Assembly of the Key Multicomponent Enzyme of Innate Immunity](#)). Both innate and adaptive immune receptors link to signal transduction and transcriptional pathways that support selection of effector functions and determine proliferation and survival of adaptive clones. These include the nuclear factor- κ B (NF κ B) and nuclear factor of activated T cells (NFAT) transcription factor families (see [NF- \$\kappa\$ B and the Immune System](#) and see [Nuclear Factor of Activated T Cells and Tolerance](#)). Phagocytes include neutrophils (see [Neutrophil Biology](#)), macrophages (see [Specialized Subsets of Tissue-Resident Macrophages in Secondary Lymphoid Organs](#)), and dendritic cells (see [Dendritic Cells](#) and see [Functional Specialization of Dendritic Cell Subsets](#)), which have the capacity to engulf and digest relatively large particles on the order of 1–10 μ m and even larger, which can involve phagocytic synapses (see [Phagocytic Synapses](#)). Phagocytes express many type of ‘scavenger’ receptors that allow them to participate in clearance of aged host proteins and cells or microbe-derived materials (see [Scavenger Receptors](#)). Adaptive immune cells generate unique antigen receptors during their development from progenitors in primary lymphoid tissues: the bone marrow (B cells make B cell receptors, BCR) and thymus (T cells make T cell receptors, TCR) (see [V\(D\)J Recombination: Orchestrating Diversity without Damage](#)). Once the BCR and TCR are made these cells undergo selection processes to eliminate auto-reactive cells (negative selection) and in the case of T cells a process of positive selection that ensures that T cells are predisposed to recognize appropriately processed antigens. Antigen processing is discussed in the context of dendritic cells, which are the primary antigen-presenting cells for initiation of immune responses (see [Dendritic Cells](#) and see [Functional Specialization of Dendritic Cell Subsets](#)). Antigen presentation requires the partial proteolysis of proteins to generate peptides that bind the major histocompatibility complex (MHC) proteins, which are expressed on the surface of the antigen-presenting cell. The TCR and peptide–MHC will span about 15 nm when engaged in the interface between T cells and antigen-presenting cells and this close proximity is facilitated by adhesion molecules in close coordination with the actin cytoskeleton (see [Cellular Structures Controlling T Cell Signaling in Time and Space](#)). The TCR interaction with the peptide–MHC complex is weak in solution, but is perfectly adapted to operate in a synaptic junction. The basic chemistry and physics of this environment is an area of active study (see [Single Molecule Methods to Measure Receptor–Ligand Interaction in Immunological Synapses](#); see [T Cell Receptor Triggering](#)). The downstream signal transduction process is well defined, although there are many nuances to this including how distinct co-stimulatory signals are integrated with TCR signals (see [The T-Cell Receptor Signalingosome](#)). We would define the immunological synapse as a (1) stationary cell–cell junction, (2) built from bona fide adhesion molecules, (3) with contiguous membranes, and (4)

vectoral communication by secretion (see [Introduction to Functional Cell Biology of Immunity](#)). One of the most important functions of the immunological synapse is in cytotoxicity mediated by CD8 T cells (see [T Cell Memory to Viral Infections](#)). Another setting for immunological synapse function is T cell help which is based on recognition of peptides presented by MHC class II and is critical for generation of high affinity antibodies (see [T Follicular Helper Cells](#)). A second major class of T cells that recognize peptides presented by MHC class II are referred to as regulatory T cells and they have adapted the immunological synapse to regulate the function of antigen-presenting cells through *trans*-endocytosis (see [CD28 Costimulation and Regulatory T Cells](#)). Cell populations within the bone marrow are maintained by chemokines (see [CXCL12/SDF-1 and Hematopoiesis](#)). Secondary lymphoid tissues serve as an important barrier to systemic pathogen spread and a hub for adaptive immunity (see [Roles of Stromal Cells in the Immune System](#)). Specialized endothelial cells control access to secondary lymphoid tissues and serve as inducible entry points into inflamed tissues (see [Lymphocyte–Endothelial Interactions](#)). These articles highlight the unique cell biology of the immune system and point to open questions accessible to tools of the cell biologist.

Stem Cell Biology

Stem cell biology has become a very active field in the last decade, due not only to the fascinating fundamental biology that has been uncovered, but also due to the application of discoveries in the field to therapeutic design for regenerative medicine. In these articles, a temporal snapshot of the field is presented that seeks to provide both experts and the students alike with a range of current lines of investigation into stem cell biology. These articles focus primarily upon adult stem cells, also referred to as progenitors that are responsible for maintaining tissue homeostasis and driving regeneration as opposed to embryonic stem cells (ES cells), which are formed during early embryogenesis and give rise to all tissues of the body. First we consider the adult stem cell in its cellular context in a number of tissues (see [The Adult Stem Cell Niche: Multiple Cellular Players in Tissue Homeostasis and Regeneration](#)). We examine tissues that are capable of robust regeneration such as the skin or skeletal muscle and the central role stem cells play in this process as well as their reliance on locally delivered cues from neighboring cells or their niche. We also examine tissues that are relatively incapable of regeneration such as the heart or brain and examine why this may be the case and how the niche may be different. We then turn our attention to the unique metabolic signature of stem cells (see [Metabolic and Energetic Regulation of Stem Cells](#)). Here we explore how the metabolic state of stem cells is different than other somatic cell types in the body and how this unique metabolic state may underlay their stem cell capacity and ability to respond to cues from the environment to expand and differentiate. Linked tightly with both the stem cell niche and the metabolic state of these stem cells is the age of the organism in which they are found. In the following article (see [Stem Cells and Aging](#)), we examine the molecular changes that occur in the ageing stem cell. Even in quiescent stem cells such

as muscle satellite cells, changes occur that render them less responsive to their environment and less capable of maintaining tissue integrity or properly executing a complete regenerative response. It has been found that stem cells lose their capacity to expand (proliferate) or that they fail to replenish a pool of stem cells (self-renew) as a reserve that is essential to the long-term health of the tissue. While the 'normal' course of ageing leads to a decrease in stem cell mitotic capacity, it has also been proposed that when stem cells lose the ability to either remain in quiescence or to differentiate, they enter of phase of unchecked proliferation leading to cancer. In the following article, (see [Cancer Stem Cells](#)), we examine the cancer stem cell hypothesis and examine a specific setting in which stem cells can be clearly demonstrated to be the cancer source and that once formed, the cancer maintains a 'stem-like' cell that gives rise to partially differentiated tumor cells. It is perhaps the focus of this article that also outlines an optimal cellular target for cancer treatment but also the potential risks in therapies aimed at increasing the proliferative potential of stem cells in chronic disease or ageing. A survey of stem cells would be incomplete if it did not take into account one of the most important paradigm shifts in the field which is the discovery of a means to derive pluripotent stem cells from ordinary somatic cells such as fibroblasts (see [Pluripotent Cells: New Tools for Disease Research and Therapies](#)). These cells, referred to as iPS cells (induced pluripotent stem cells) behave

quite similarly to ES cells and can be triggered to differentiate into most all of the tissues found in the body. From a fundamental viewpoint, this is remarkable as it involves erasing much of the epigenetic signature in the adult-derived fibroblast and reprogramming the cell completely. These cells have several potential clinical uses including the use as donor cells for engraftment into damaged tissue as well as for therapy screens to design drugs or the best combination of drugs for a specific patient suffering from a tissue-specific disease. Drugs that correct heart arrhythmias in one patient do not always work for another who presents seemingly identical symptoms. The patient's cells can be used to prescreen for a tailored or 'personalized medicine' approach. Lastly, we examine how the epigenome, which refers to modifications made to the DNA of the genome without changing the DNA itself is an important player in governing stem cell competence and identity (see [Epigenetic Regulation of Stem Cells](#)). This can be brought about by DNA methylation or modification of the histones via acetylation. Drugs that target the enzymes that drive histone acetylation are already in use for cancer treatment, and it is suggested that this same family of drugs may have a major potential potentiating regeneration of diseased tissue. While this collection of topics is not meant to be exhaustive, it covers many of the key topics of investigation in adult stem cell research today that will be the basis of future stem cell research and responsible for new cures in regenerative medicine.

CELL COMMUNICATION: GROWTH FACTOR MEDIATED CELL SIGNALING

Contents

Receptor Tyrosine Kinases and Their Ligands

Cytokine Receptors and Their Ligands

TGF- β Superfamily Signaling

G Protein-Coupled Receptors

Guanylyl Cyclase Receptors

Tumor Necrosis Factor Receptors: A Brief Digestion

Signaling and Function of Death Receptors of the Tumor Necrosis Factor Receptor Superfamily

Hedgehog Signaling in Development and Disease

The Wnt/ β -Catenin Pathway

The Hippo Pathway

The Notch Pathway

Receptor Tyrosine Kinases and Their Ligands

C-H Heldin and J Lennartsson, Uppsala University, Uppsala, Sweden

© 2016 Elsevier Inc. All rights reserved.

Glossary

Adaptor molecules Molecules without intrinsic enzymatic activity which serve to connect other signaling molecules with each other.

Autophosphorylation A kinase phosphorylates itself or a companion molecule in a dimer.

Endocytosis Internalization of membrane-bound molecules into intracellular vesicles (endosomes).

Growth factor A ligand, normally a protein, which binds to a receptor at the cell membrane and promotes cell proliferation.

Shedding Proteolytic cleavage of membrane proteins releasing their extracellular domains.

Introduction

Cells within a multicellular organism communicate with each other, by direct cell–cell or cell–matrix interactions or by release of molecules that bind to and activate receptors on the surface of, or inside, the receiving cell. Several growth factors and cytokines bind to plasma membrane-bound receptors associated with kinase activities. Certain receptors have intrinsic kinases, as are the cases for members of the large family of receptor tyrosine kinases (RTKs) and for members of the smaller family of receptor serine/threonine kinases. Other receptors, particularly those binding cytokines that have important functions to regulate immune cells, lack intrinsic kinase activity, but associate with intracellular tyrosine kinases that are activated upon ligand binding (Heldin *et al.*, 2014).

This review will focus on RTKs. First, we will describe some general features of RTK relating to their activation, turnover,

and signaling mechanisms, then we will comment on the structural and functional properties of each RTK family, and finally we will discuss their involvement in disease.

Families of RTKs and Their Ligands

RTKs have important functions in the regulation of the embryonal development, and in the regulation of wound healing and tissue homeostasis in the adult. RTKs have an extracellular, ligand-binding part, which is composed of different combinations of domains, for example, immunoglobulin (Ig)-like, fibronectin type III, cysteine-rich, epidermal growth factor (EGF)-like, cadherin-like, and discoidin-like domains (Blume-Jensen and Hunter, 2001; Lemmon and Schlessinger, 2010). A single transmembrane segment connects the extracellular part with the intracellular part, which harbors the kinase domain together



Figure 1 RTK families. The 20 RTK subfamilies and their characteristic structural domains are shown. Reproduced with permission from Lemmon, M.A., Schlessinger, J., 2010. Cell signaling by receptor tyrosine kinases. *Cell* 141 (7), 1117–1134.

with juxtamembrane and C-terminal sequences. Certain RTKs, in addition, have a characteristic insert in their kinase domain which protrudes from the kinase domain after it has been folded (Locascio and Donoghue, 2013; Figure 1).

Based on their structural properties, the 58 RTKs encoded in the human genome can be divided into 20 subfamilies (Blume-Jensen and Hunter, 2001; Lemmon and Schlessinger, 2010; Figure 1). A prototypic example is the EGF receptor family which consists of the EGF receptor, ErbB2, ErbB3, and ErbB4. Each RTK family corresponds to a family of structurally related ligands; the EGF family consists of 11 members (Citri and Yarden, 2006). Normally, each receptor binds more than one of the ligands in the family, and each ligand binds more than one receptor in the family (Table 1). Binding of ligands

outside the family has been observed, but is rare; examples include the Ryk and Ror RTKs which bind certain members of the Wnt family that have Frizzled receptors as their canonical receptors (van Amerongen *et al.*, 2008).

Some ligands are secreted as soluble molecules which may or may not need to undergo proteolytic activation. Other ligands are synthesized as transmembrane proteins; in some cases the transmembrane molecule is cleaved to release the ligand, in other cases the ligand presented on one cell may activate a receptor in another cell. In the case of ephrins and RTKs of the Eph receptor family, there are also examples of retrograde signaling, where the Eph 'receptor' by interacting with the ephrine 'ligand' induces signaling events in the ligand-bearing cell (Pasquale, 2008).

Table 1 Receptor tyrosine kinase family members and their ligands

RTK family	Receptors	Ligands ^a	References
EGF receptor	EGFR	Epidermal growth factor (EGF), transforming growth factor α (TGF α), amphiregulin, epiregulin, heparin-binding EGF-like growth factor (HB-EGF), β -cellulin, epigen	Fuller <i>et al.</i> , 2008
	ErbB2/Her2	Unknown	
	ErbB3/Her3	Neuregulin (NRG) 1, 2	
	ErbB4/Her4	HB-EGF, epiregulin, β -cellulin, NRG 3, 4	
Insulin receptor	InsR	Insulin, (insulin-like growth factor (IGF)-1, IGF-2)	Pollak, 2012
	IGF-1R	IGF-1, IGF-2, (insulin)	
	IRR	Unknown	
PDGF receptor	PDGFR α	Platelet-derived growth factor (PDGF)-AA, -BB, -AB, -CC	Dai <i>et al.</i> , 2002; Heldin and Lennartsson, 2013; Lin <i>et al.</i> , 2008; Lyman and Jacobsen, 1998
	PDGFR β	PDGF-BB, -DD	
	CSF-1R	Colony stimulating factor (CSF)-1, interleukin-34	
	SCFR/c-Kit	Stem cell factor (SCF)	
	Flt3/Flk2	Fms-related tyrosine kinase 3 ligand (FLT3LG)	
VEGF receptor	VEGFR1	Vascular endothelial growth factor (VEGF)-A, -B, placental growth factor (PlGF)	Olsson <i>et al.</i> , 2006
	VEGFR2	VEGF-A, -C, -D	
	VEGFR3	VEGF-C, -D	
FGF receptor	FGFR1 ^b	Fibroblast growth factor (FGF)-1, -2, -3, -4, -5, -6, -7, -8, -9, -10, -16, -17, -18, -19, -20, -21, -22, -23	Zhang <i>et al.</i> , 2006
	FGFR2 ^b	FGF-1, -2, -3, -4, -5, -6, -7, -8, -9, -10, -16, -17, -18, -19, -20, -21, -22, -23	
	FGFR3 ^b	FGF-1, -2, -4, -5, -6, -8, -9, -16, -17, -18, -19, -20, -21, -23	
	FGFR4 ^b	FGF-1, -2, -4, -5, -6, -8, -17, -18, -19, -21, -23	
KLK/CCK	CCK4	Unknown	
NGF receptor	TrkA	Nerve growth factor (NGF), (neurotrophin (NT)-3)	Skaper, 2012
	TrkB	Brain-derived neurotrophic factor (BDNF), NT-4, NT-5, (NT-3)	
	TrkC	NT-3	
HGF receptor	Met	Hepatocyte growth factor (HGF)	Graveel <i>et al.</i> , 2013; Wang <i>et al.</i> , 2013
	Ron	HGF-like (HGFL)/macrophage stimulating protein (MSP)	
Eph receptor	EphA1-8	Ephrin-A1, 2, 3, 4, 5	Lisabeth <i>et al.</i> , 2013
	EphB1-6	Ephrin-B1, 2, 3	
Axl receptor	Axl	Growth arrest-specific (GAS) 6, tubby-related protein (Tulp) 1	Cabero <i>et al.</i> , 2010; Lemke, 2013; Pacciez <i>et al.</i> , 2014
	Mer	GAS6, protein S (ProtS), Tulp1, tubby	
	Tyro3/Sky	GAS6, ProtS, Tulp1	
Tie receptor	Tie1	Unknown	Maisonpierre <i>et al.</i> , 1997
	Tie2	Angiopoietin (Angpt)-1, -2, -4	
Ryk receptor	Ryk	Wnt-1, -3a, -5a	van Amerongen <i>et al.</i> , 2008
DDR	DDR1	Collagens	Vogel <i>et al.</i> , 2006
	DDR2	Collagens	
Ret receptor	Ret	GDNF, neurturin, persephin, artemin	Mulligan, 2014
Ros receptor	Ros	Unknown	
ALK receptor	LTK	Unknown	Hallberg and Palmer, 2013
	ALK	Pleiotrophin, midkine	
Ror receptor	Ror1	Wnt1, 3, 5a	van Amerongen <i>et al.</i> , 2008
	Ror2	Wnt1, 3, 5a	
MuSK receptor	MuSK	Aggrin ^c	Hubbard and Gnanasambandan, 2013
LMR receptor	AATYK	Unknown	
	AATYK2	Unknown	
	AATYK3	Unknown	

^aLigand-receptor interactions of lower affinities are indicated by brackets.^bIncludes all FGFR splice forms and data represent *in vitro* activity which varies considerable between different FGF-FGFR combinations.^cInteracts with Lrp4 on the target cell and this is required for MuSK receptor dimerization and activation.

RTKs Undergo Ligand-Induced Dimerization

Many, maybe all, RTKs are activated by ligand-induced dimerization (Heldin, 1995; Lemmon and Schlessinger, 2010). The juxtaposition of the intracellular parts of the receptors makes it possible for one receptor molecule to phosphorylate the other *in trans*, a process that is referred to as autophosphorylation.

Dimerization of RTKs can occur by different mechanisms. Certain ligands are dimeric molecules that bind their receptors in a symmetric manner; in the case of nerve growth factor (NGF), the ligand acts as a bridge between the receptors and the receptors do not touch each other (Wehrman *et al.*, 2007), whereas in the cases of platelet-derived growth factor (PDGF) (Yang *et al.*, 2008), colony stimulating factor 1 (CSF1) (Chen *et al.*, 2008) and stem cell factor (SCF) (Yuzawa *et al.*, 2007) receptors the dimeric configuration is stabilized by interactions both between the ligands and the receptors, and by direct receptor–receptor interactions. Direct interactions between Eph receptors also appear to be important for their oligomerization and activation (Seiradake *et al.*, 2010).

The situation is very different for the EGF receptor family. These receptors have four extracellular domains; the monomeric ligand binds to domains I and III, which causes a dramatic change in conformation unmasking a dimerization arm in domain II which promotes direct receptor–receptor interaction, but the ligand itself does not take part in bridging between the two receptors (Burgess *et al.*, 2003). Moreover, the dimerization of receptors in the EGF family is dependent on an interaction between the N-terminal parts of the transmembrane helices, which promotes an antiparallel interaction between the juxtamembrane sequences releasing the inhibition by the membrane (Arkhipov *et al.*, 2013; Endres *et al.*, 2013).

Yet another variation on the dimerization theme is exemplified by fibroblast growth factor (FGF) receptors; in this case the binding of the monomeric ligand is stabilized by heparin/heparan sulfate polysaccharide chains (Schlessinger *et al.*, 2000). A 2:2:2 complex is formed in which each FGF molecule contacts both receptor molecules in the dimer, heparin/heparan sulfate contacts both ligand and receptor, and the two receptors contact each other.

There are indications that certain RTKs form dimers already before ligand binding, and members of the insulin receptor family are even disulfide-bonded dimers. In these cases, ligand binding is likely to induce a conformational change that activates the receptor kinases. There are indications that the relative orientation of the receptors in the dimer is important for the activation (Jiang and Hunter, 1999). Moreover, an initial ligand-induced dimerization may be followed by oligomerization into large clusters, which may be important for activation of, for example, the Tie2 (Barton *et al.*, 2006) and Eph (Himanen and Nikolov, 2003) RTKs.

Dimerization Activates the Kinase Activity of RTKs

The autophosphorylation that follows ligand-induced RTK dimerization has two important consequences, i.e., it changes the conformation of the receptor promoting activation of its

kinase, and it provides docking sites for SH2 domain containing signal transduction molecules.

The kinases of RTKs are kept inactive by one or more mechanisms involving the activation loop of the kinase domain, the juxtamembrane segment, or the C-terminal tail of the receptor. Autophosphorylations in these regions are connected with release of the kinase activity. Thus, in the resting state, the activation loop of most RTKs is folded over the active site of the kinase, thereby preventing ATP to access its pocket; the phosphorylation of a conserved tyrosine residue in the activation loop causes a conformational change moving it away from the active site. This mechanism for control of the kinase action appears to prevail for most RTKs, exceptions include members of the EGF receptor family (Zhang *et al.*, 2006) and Ret (Knowles *et al.*, 2006). The juxtamembrane segment is in the resting state of certain RTKs, including MuSK (Till *et al.*, 2002), Flt3 (Griffith *et al.*, 2004), the SCF receptor (Mol *et al.*, 2004) and Eph family members (Wybenga-Groot *et al.*, 2001), also folded over the active site; phosphorylation in this segment changes its conformation and moves it away from the active site (Hubbard, 2004). Finally, the C-terminal tail of some RTKs, including Tie2 (Shewchuk *et al.*, 2000), Ron (Yokoyama *et al.*, 2005) and probably also the PDGF β -receptor (Chiara *et al.*, 2004), can be folded over the kinase domain to inhibit the kinase activity; phosphorylation of certain tyrosine residues in the tail makes it swing open and allows activation of the kinase.

It is not obvious how the juxtaposition of two inactive kinase domains can lead to autophosphorylation *in trans* between the receptors. It is possible that the kinases are not completely inactive in the resting state and their juxtaposition may start the activation process in a reciprocal manner. Moreover, in the case of the EGF receptor family, a phosphorylation-independent mechanism for activation of the receptor kinase in the dimer has been elucidated. Thus, the C-terminal lobe of one kinase domain makes contact with the N-terminal lobe of the other, thereby activating it (Zhang *et al.*, 2006). Interestingly, the juxtamembrane segment, which in other RTKs has an inhibitory role, is needed to stabilize the interaction (Jura *et al.*, 2009, 2011; Red Brewer *et al.*, 2009).

Ligand binding induces homodimerization of RTKs, however, heterodimeric complexes can also be formed, for example, in the case of the EGF (Citri and Yarden, 2006), PDGF (Heldin and Lennartsson, 2013) or vascular endothelial growth factor (VEGF) (Olsson *et al.*, 2006) receptor families. Since downstream signaling depends on autophosphorylation of specific tyrosine residues in the RTKs, differences in the autophosphorylation patterns between homo- and heterodimeric receptors, will result in different signaling properties.

Signaling Downstream of RTKs

In addition to activating the kinases of RTKs, ligand-induced autophosphorylation of the receptors provide docking sites for SH2 domain-containing signaling molecules. The SH2 domains recognize phosphorylated tyrosine residues in a certain environment; in particular, the 3–6 amino acid residues on C-terminal of the phosphorylated tyrosine determine the specificity of SH2 domain binding (Songyang and

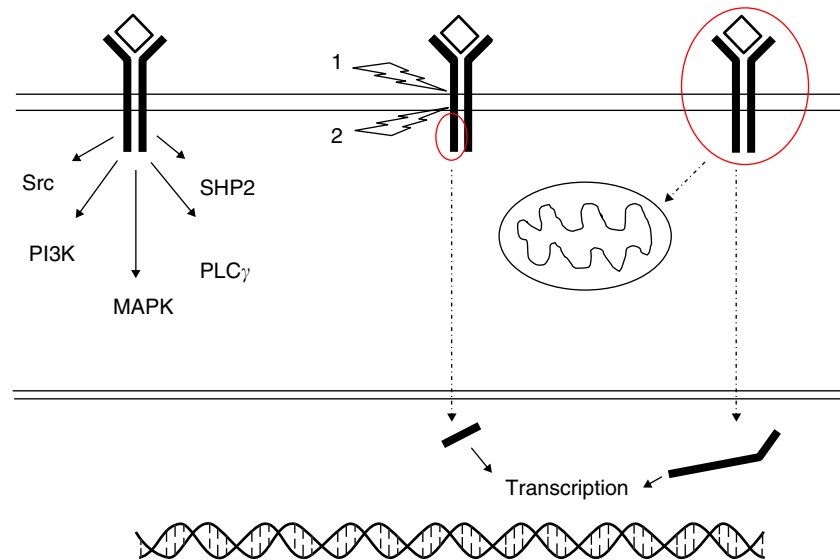


Figure 2 Schematic representation of signaling induced by receptor tyrosine kinases. Phosphorylated tyrosine residues in the ligand-activated receptor serve as docking sites for signaling proteins containing SH2 or PTB domains, which subsequently activates intracellular signaling pathways, for example Src, PI3K, MAPK, PLC γ , and SHP2. Another signaling mechanism includes proteolytic release of the intracellular tyrosine kinase domain, which is initiated by a protease cleaving the receptor extracellularly (1) followed by a subsequent proteolysis in the transmembrane segment (2). The released intracellular receptor fragment translocates into the nucleus where it can influence gene expression. In addition, the entire receptor can be translocated to the nucleus or, in certain cases, to the mitochondria.

Cantley, 2004). Autophosphorylation sites binding SH2 domains are often located in the juxtamembrane, kinase insert and C-terminal tail regions of the receptors, but more rarely in the kinase domains.

In addition to autophosphorylation sites in the receptors themselves, additional phosphorylation sites are found on scaffolding proteins that bind to the receptors in a selective manner. Thus, RTKs of the insulin receptor family bind the 4 members of the insulin receptor substrate (IRS) family, FGF and NGF receptors bind FRS2 α and β , and several RTKs bind Gab1 or Gab2. Formation of such multi-protein tyrosine-kinase-dependent complexes is important in RTK signaling (Pawson, 2004).

Of the SH2 domain-containing molecules that bind to phosphorylated tyrosines in RTK complexes, certain have intrinsic enzymatic activities, such as members of the Src family of tyrosine kinases, members of the SHP family of tyrosine phosphatases, GTPase-activating protein (GAP) for Ras, phospholipase C γ (PLC γ), and members of the Cbl family of ubiquitin ligases (Figure 2). Other receptor-interacting SH2 domain-containing molecules form stable complexes with enzymes, for example, Grb2 which forms a complex with SOS1, a guanine nucleotide exchange factor (GEF) for Ras, and the phosphatidylinositol-3'-kinase (PI3-kinase) regulatory subunit p85 which forms a complex with the catalytic subunit p110. Some RTKs also activate members of the signal transducer and activator of transcription (STAT) family, which after phosphorylation and activation are translocated to the nucleus where they regulate transcription by binding to the promoters of certain genes. Finally, RTKs bind certain adaptor molecules, such as Shc, Crk, and Nck, that contain, in addition to SH2 domains, other domains mediating interactions with certain proteins or phospholipids of membranes. Examples of such

domains are SH3- and WW-domains that recognize Pro-rich regions, PTB-domains that bind tyrosine-phosphorylated regions although the dependence on phosphorylation is not absolute, PDZ-domains that recognize a C-terminal Val residue or certain other bulky hydrophobic residues, and PH-, PX-, C1-, C2-, and FYVE-domains that recognize phospholipids. The common occurrence of the latter domains in signaling molecules reflects the fact that much signal transduction activities occur at the inner leaflet of the plasma membrane or close to membranes of intracellular vesicles.

Certain of the members of the RTK family have no, or low, kinase activity, for example, ErbB3, CCK4, EphB6, and Ryk, since they lack critical catalytic amino acid residues. Still, these receptors have important roles in signal transduction by acting as phosphorylatable scaffold molecules in heterodimers with other kinase-active RTKs.

Given that there is a limited number of signaling pathways activated downstream of RTKs, an interesting question is how specificity can be achieved. The binding specificities of SH2 domains provide a basis for specificity in signal transduction, where a given RTK may activate certain signaling pathways, but not others. However, the formation of large scaffolds with many phosphorylated tyrosines means that activation of most of the RTKs leads to activation of most of the signaling pathways induced by the binding of SH2 domain-containing molecules. Adding to the complexity is also the fact that there is an extensive cross-talk between the signaling pathways, creating an intracellular signaling network rather than linear signaling pathways (Lemmon and Schlessinger, 2010).

Another mechanism of signaling via certain RTKs, involves ligand-induced cleavage of the receptor just outside the cell membrane by metalloproteases, for example, ADAM17, followed by another cleavage in the transmembrane region by

γ -secretase, whereby the cytoplasmic part of the receptor is translocated to the nucleus where it may regulate transcription. Such a mechanism was originally described for ErbB4 (Ni *et al.*, 2001), but has subsequently been observed also for several other RTKs, including the CSF-1 receptor, EphB2, PTK7, Ret, Met, Ron, Ryk, Tie1, FGF receptor 3, and VEGF receptor 1 and 2 (Carpenter and Liao, 2013), as well as the receptors for insulin (Kasuga *et al.*, 2007) and IGF-I (McElroy *et al.*, 2007). In the case of ErbB4 the intracellular part forms a complex with TAB2 and the corepressor N-CoR which represses the transcription of several glial genes and thereby inhibits the formation of astrocytes (Sardi *et al.*, 2006). In addition, the ErbB4 intracellular domain has been found to associate with mitochondria and promote apoptosis of breast cancer cells (Naresh *et al.* 2006).

A consequence of the shedding of the extracellular domains of RTKs by ADAM17, or by other proteases, is that the released extracellular domain may scaffold the corresponding ligand and thus cause decreased signaling (López-Otín and Hunter, 2010).

In addition, certain RTKs, including members of the EGF receptor family, the IGF-I receptor, Ron, Met, FGF receptor 1 and 2, and VEGF receptor 1 and 2 can be translocated in uncleaved forms to the nucleus (Carpenter and Liao, 2013; Massie and Mills, 2006; Song *et al.*, 2013). It appears that the RTKs are mainly present in the nucleoplasm rather than in the nuclear envelope. Trafficking of the EGF receptor from the cell surface to the Golgi occurs in a microtubule-dependent manner and fusion with the Golgi involves syntaxin 6-mediated membrane fusion (Du *et al.*, 2014). ErbB2 (Wang *et al.*, 2004) and the EGF receptor (Lin *et al.*, 2001; Lo *et al.*, 2005) have been shown to regulate transcription in cooperation with other nuclear factors. The physiological implication of the translocation full-length RTKs into the nucleus, however, remains to be elucidated.

In addition, the intact ErbB2 molecule has been found to be translocated to the mitochondria, via association with mtHSP70, where it negatively regulates respiratory functions in a kinase-dependent manner (Ding *et al.*, 2012).

RTK Coreceptors

Signaling via RTKs is modulated by interactions between RTKs and other molecules expressed on the plasma membrane. Thus, certain RTKs interact with receptors for extracellular matrix molecules, such as CD44 that binds hyaluronan (Orian-Rousseau, 2010), and integrins that bind collagen, fibronectin, or laminins (Ivaska and Heino, 2011). Low-density lipoprotein (LDL) receptor-related proteins (LRPs) also interact with certain RTKs. Thus, the heparin sulfate proteoglycan agrin binds LRP4 and thereby activate the RTK MuSK (Kim *et al.*, 2008). On the other hand, LRP interacts with PDGF receptors and suppresses their signaling (Boucher *et al.*, 2002; Loukinova *et al.*, 2002). Other examples of RTK coreceptors are glial-derived neurotrophic factor (GDNF) receptor α , a glycosylphosphatidylinositol (GPI)-anchored molecule that is needed for GDNF to bind to and activate the RTK Ret (Schlee *et al.*, 2006), and neuropilin 1 and 2 that cooperate with VEGF receptors (Olsson *et al.*, 2006).

Feedback and Amplification of RTK Signaling

In the early phase after stimulation of RTK signaling, amplification mechanisms assure that a high level of signaling is rapidly achieved. For example, reactive oxygen species (ROS) are produced in a PI3 kinase-dependent manner (Bae *et al.*, 1997), which inhibit PTPs by oxidizing a cysteine residue in their active site (Tonks, 2006) and thereby enhance tyrosine phosphorylation. Another example is activation of the ADAM family of proteases downstream of EGF receptors, which cleaves and activates membrane-bound heparin-binding (HB) EGF further activating the EGF receptors through autocrine stimulation (Hynes and Schlang, 2006).

At later stages various feedback mechanisms are engaged to control RTK signaling. Examples include the docking of PTPs of the SHP family to the receptors, which leads to dephosphorylation of RTKs, as well as the docking of RasGAP which counteracts the activation of Ras. In addition, protein kinase C (PKC), which is activated downstream of PLC γ , is involved in feedback control by phosphorylating and inactivating certain RTKs, such as the EGF receptor (Hunter *et al.*, 1984), the insulin receptor (Bollag *et al.*, 1986), Met (Gandino *et al.*, 1990) and the SCF receptor (Blume-Jensen *et al.*, 1993).

Activation of RTKs rapidly leads to the induction of immediate-early genes, which encode different effector molecules involved in stimulation of cell proliferation, survival, and migration. Thereafter, delayed-early genes are induced, several of which encode proteins that suppress signaling. Examples of inhibitory mechanisms downstream of several RTKs include FOSL1, which binds to and inhibits AP1 transcription factors activated by RTKs; the transcription factor Id2, which inhibits the TCF transcription complex; dual specificity phosphatases (DUSPs), which dephosphorylate and inactivate MAP kinases; and ZFP36, which recognizes AU-motifs in the 3' end of mRNA molecules and promotes their degradation (Amit *et al.*, 2007). With regard to the EGF receptor, a number of additional feedback mechanisms occur, including induction of NAB2, LRIG-1, Kerkon-1, and Mig6 that inhibit the receptor itself, induction of Argos that sequesters ligand, and induction of Sprouty that has several inhibitory effects on downstream signaling (Citri and Yarden, 2006; Shilo, 2005).

Internalization of RTKs

Ligand binding induces clustering of RTKs in coated pits at the cell surface, whereafter the ligand-receptor complex is internalized via a clathrin-dependent pathway (Zwang and Yarden, 2009), but other internalization routes, for example, caveolin-dependent endocytosis, macropinocytosis, or phagocytosis, also occur (Parachoniak and Park, 2012; Schmees *et al.*, 2012; Sorkin and Goh, 2009). Signaling via RTKs continue after internalization until the pH of the endosomes drops to such a low value that the ligand dissociates from the receptor and the receptor is dephosphorylated by PTPs (von Zastrow and Sorkin, 2007). Interestingly, signaling via RTKs at the plasma membrane differ qualitatively from RTK signaling at endosomes, maybe because of differences in the location of downstream signaling molecules (Miaczynska *et al.*, 2004). In the case of VEGF receptor 2 and 3, the Eph receptor ligand

ephrin-B2 has been found to promote endocytosis, and absence of ephrin-B2 leads to diminished VEGF receptor signaling through Rac1, Akt, and Erk1/2 (Sawamiphak *et al.*, 2010; Wang *et al.*, 2010). Ephrin-B2 has also been found to control the internalization and signaling of the PDGF β -receptor in vascular smooth muscle cells (Nakayama *et al.*, 2013). Moreover, the adaptor protein Dab2, the polarity factor PAR3 and atypical PKC isoforms have been shown to regulate VEGF receptor endocytosis; properly controlled endocytosis has been found to be necessary for angiogenesis in the retina (Nakayama *et al.*, 2013).

Ubiquitination of RTKs by ubiquitin ligases of the Cbl family, and possibly also other E3 ligases, promotes their recognition by components of the endosomal sorting complex required for transport (ESCRT) machinery, which facilitates translocation to multivesicular bodies and degradation in lysosomes (Raiborg and Stenmark, 2009). In the case of the EGF receptor, Cbl initially modifies the receptor by monomers of the ubiquitin-like molecule Nedd8 (Oved *et al.*, 2006). Thereafter, Cbl mono- and poly-ubiquitinates the receptor (Haglund *et al.*, 2003; Huang *et al.*, 2006; Mosesson *et al.*, 2003). The receptor is then recognized by ubiquitin binding proteins of the clathrin coat, such as Eps15, Epsin1-3, and Hrs, to promote the transport to the ESCRT machinery (Avraham and Yarden, 2011). Alternatively, RTKs can be sorted to recycling via fast recycling vesicles characterized by expression of Rab4a, or via slow recycling vesicles characterized by expression of Rab11. After recycling and return to the plasma membrane, the RTK can serve again. Thus, recycling enhances RTK signaling. The balance between sorting to degradation vs. sorting to recycling can be affected by posttranslational modifications of the RTK and can significantly affect the amplitude of signaling (Hellberg *et al.*, 2009; Mosesson *et al.*, 2008; Parachoniak and Park, 2012).

Structural and Functional Properties of RTK Families

The EGF Receptor Family

The EGF receptor family consists of four members, i.e., the EGF receptor, ErbB2, ErbB3, and ErbB4 (Citri and Yarden, 2006). Their extracellular parts consist of four domains; ligand binding occurs to domains I and III, whereby a dramatic relocation of the cysteine-rich domains II and IV occur. Interestingly, ErbB2 does not bind any ligand; its extracellular domain is constitutively in the activated configuration and it takes part in signaling by forming heterodimers with other members in the family. Also notable is the fact that the intracellular kinase of ErbB3 lacks critical amino acid residues involved in catalysis, and therefore has no, or very low, kinase activity; it functions by forming dimers with other members of the family acting as a phosphorylatable scaffold. In particular, ErbB3 mediates a powerful activation of PI3-kinase.

There are 11 ligands in the EGF family; 7 of these bind to the EGF receptor, 2 to ErbB3 and 5 to ErbB4 (Table 1). Several of the receptors thus bind to more than one ligand, and some of the ligands bind to more than one receptor. Certain of the ligands are synthesized as transmembrane proteins, for

example, EGF itself or HB-EGF, that need to be cleaved off before they can bind to receptors.

Deletion of the EGF receptor gene in mice has revealed important roles in the development of epithelial cells in several organs, including skin, lung, intestine and placenta (Miettinen *et al.*, 1995; Sibilio and Wagner, 1995; Threadgill *et al.*, 1995). Knockout mice that survive after birth develop neurodegeneration (Sibilio *et al.*, 1998). Mice in which ErbB2 (Lee *et al.*, 1995) or ErbB4 (Gassmann *et al.*, 1995) has been deleted display defect cardiac development; trabeculae do not develop and the heart beats are irregular. In ErbB3 knockout mice trabeculation is delayed but normal, however, atrioventricular valves are dilated and thinned (Riethmacher *et al.*, 1997).

The Insulin Receptor Family

The insulin receptor family consists of three members, i.e., the insulin receptor and the IGF-I receptor, binding insulin and IGF-I, respectively, and the insulin receptor-related receptor for which no ligand is known (Ward *et al.*, 2007). The insulin receptor occurs as two splice forms, A and B; the A isoforms binds, in addition to insulin, also IGF-II. The members of the insulin receptor family are synthesized as precursor molecules that undergo proteolytic cleavage in the extracellular part, creating α - and β -chains that are held together by disulfide bonds. In addition, two α -chains are bound together with disulfide bonds, giving the receptors a heterotetrameric configuration. The extracellular domains of the insulin receptor consists of two leucine-rich domains separated by a cysteine-rich domain, and three FNIII domains; its three-dimensional configuration resembles a reverted V (Ward *et al.*, 2007).

The insulin receptor regulates lipid, protein, and carbohydrate metabolism, in particular glucose homeostasis (Kitamura *et al.*, 2003), whereas the IGF-I receptor regulates normal growth and development (Adams *et al.*, 2000). Insulin and IGF-I receptors also regulate longevity, as has been shown in different animal models from nematodes to the mouse (Giannakou and Partridge, 2007).

The PDGF Receptor Family

The PDGF receptor family consists of PDGF α - and β -receptors, SCF receptor (Kit), CSF-1 receptor (Fms) and the Flt3 receptor. These receptors are characterized by extracellular parts consisting of five Ig-like domains and an intracellular kinase containing a rather long kinase insert region.

The ligands for the PDGF receptor family of RTKs are all dimeric molecules. The PDGF α - and β -receptors bind homodimers of PDGF-A, -B, -C and -D polypeptide chains and the PDGF-AB heterodimer, the CSF-I receptor binds CSF-1 and interleukin-34, and the SCF receptor and Flt3 bind one ligand each. The ligands of this family bind to Ig domains 2 and 3 of their receptors; after ligand-induced dimerization, Ig-domains 4 and 5 stabilize the dimer by direct receptor-receptor interactions (Verstraete and Savvides, 2012; Yuzawa *et al.*, 2007). Binding of the distantly related VEGF-A to the PDGF α - and β -receptors has been reported (Ball *et al.*, 2007), but the physiological significance of this finding remains to be elucidated.

PDGF receptors have important functions during the embryonal development, as elucidated by knocking out the genes for PDGF isoforms or PDGF receptors in mice (Andrae *et al.*, 2008). Often the PDGF isoforms are produced by epithelial or endothelial cells and act in a paracrine manner on neighboring mesenchymal cells, such as fibroblasts, smooth muscle cells, and pericytes. Knockout of the PDGF α -receptor led to defects in neural crest mesenchymal derivatives, such as the cardiac outflow tract, the thymus, and skeletal components in the facial and other regions, and in the development of the palate and teeth (Soriano, 1997). PDGF α -receptor also mediates signals that are needed for the development and differentiation of oligodendrocyte progenitor cells (Calver *et al.*, 1998) and retinal astrocytes (Fruttiger *et al.*, 2000).

Knockout of the PDGF β -receptor led to abnormal development of the glomeruli in the kidney due to defect development of mesangial cells (Soriano, 1994), capillary microaneurysm (Lindahl *et al.*, 1997), cardiac defects (Van den Akker *et al.*, 2008), and placental defects (Ohlsson *et al.*, 1999).

In the adult, PDGF β -receptor controls the interstitial fluid pressure in tissues and thereby prevents edema (Rodt *et al.*, 1996). The mechanism involves integrin-mediated contacts between fibroblasts and myofibroblasts and extracellular matrix molecules; activation of the PDGF β -receptor causes contraction of the cells and thereby also of the tissue (Lidén *et al.*, 2006). Activation of PDGF receptors on cells involved in wound healing, such as fibroblasts, smooth muscle cells, neutrophils and macrophages, promotes healing of the wound (Robson *et al.*, 1992).

The SCF receptor is expressed in primitive hematopoietic cells and promotes survival and proliferation of these cells in synergy with other cytokines. The SCF receptor is also expressed on melanocytes and has been shown to be important for the migration of these cells from the neuronal crest to the dermis during development (Wehrle-Haller, 2003); defects in SCF receptor signaling was found to lead to pigmentation defects found in patients with piebald disease. SCF receptor knockout mice had reduced fertility (Blume-Jensen *et al.*, 2000; Kissel *et al.*, 2000) and were also constipated because of loss of interstitial cells of Cajal, which functions as pacemaker cells (Huizinga *et al.*, 1995; Ward *et al.*, 1995); they also had defects in spatial learning (Kataguchi *et al.*, 2000; Motro *et al.*, 1996).

The CSF-1 receptor is critical for the proliferation and differentiation of mononuclear phagocytic cells, such as monocytes, tissue macrophages, microglia, osteoclasts, and antigen presenting dendritic cells (Hamilton, 2008). Signaling via Flt3 is important for early haematopoiesis, particularly the proliferation of haematopoietic stem cells and B cell progenitors (Stirewalt and Radich, 2003).

The VEGF Receptor Family

The three members of the VEGF receptor family are characterized by 7 Ig-like domains extracellularly, except that Ig domain 5 in VEGF receptors is replaced by a disulfide bond (Olsson *et al.*, 2006). The kinase domains of VEGF receptors have characteristic insert regions. The VEGF family of ligands is

dimeric molecules and consists of 5 members, i.e., VEGF-A, -B, -C, and -D, and PlGF; they bind to Ig domain 2 and 3.

VEGF receptors have important roles in the establishment and maintenance of the blood and lymphatic vascular systems. Mice with the VEGFR1 gene deleted die at embryonic day 8.5–9.0 by vascular disorganization due to endothelial cell overgrowth (Fong *et al.*, 1995) and increased hemangioblast commitment (Fong *et al.*, 1999), suggesting that VEGFR1 has an inhibitory role in angiogenesis. Knockout of the VEGFR2 gene leads to embryonic lethality day 8.5–9.5 due to defective blood island formation and vasculogenesis (Shalaby *et al.*, 1995).

VEGFR3 has a specific role in the formation of the lymphatic system (Tammela and Alitalo, 2010). However, VEGFR3 knockout mice show embryonic lethality before the formation of the lymphatic system, due to cardiovascular failure (Dumont *et al.*, 1998), suggesting that VEGFR3 also is important for the development of the blood vascular system.

The FGF Receptor Family

The FGF receptor family has 4 members that are characterized by three Ig domains extracellularly; Ig-domain 3 is subject to differential splicing events which alters the ligand-binding specificities of the receptors (Goetz and Mohammadi, 2013). There are 18 ligands in the FGF family; whereas 15 members mainly act locally in a paracrine manner as is common for most RTK ligands, 3 of the FGF family members, FGF-19, -21, and -23, act in an endocrine, hormone-like manner.

FGF receptors are expressed on many cell types and have important roles during the embryonal development, including mesoderm patterning and development of several organ systems, and in the regulation of angiogenesis and wound healing in the adult (Turner and Grose, 2010). Loss of function mutations in the FGFR1 gene lead to Kallmann syndrome 2, which is characterized by hearing loss, cleft palate and tooth agenesis (Dode *et al.*, 2003), and loss of function mutations in the genes of FGFR2 or the FGF-10 ligand give rise to lacrimo-auriculo-dento-digital syndrome (Rohmann *et al.*, 2006).

Met and Ron

Met and Ron RTKs each consists of a Sema domain, a PSI domain and 4 Ig domains extracellularly; the receptors are cleaved to generate α - and β -chain heterodimers. The intracellular part with the kinase domain has a rather short C-terminal tail (Trusolino *et al.*, 2010). Their ligands, HGF/SF and MST1, respectively, are rather large molecules consisting of four N-terminal kringle domains and a C-terminal serine protease homology domain that lacks proteolytic activity.

During embryonal development, Met is essential for the growth, survival and migration of hepatocytes, placental trophoblasts and muscle progenitor cells. Met-null mouse embryos have considerably diminished livers, the placental labyrinth is severely hypomorphic and there is an absence of myogenic precursor cells (Birchmeier and Gherardi, 1998). Met has also an important role in the regeneration of the liver and the kidney after acute or chronic insults (Trusolino *et al.*, 2010). Ron is involved in the development of epithelial, bone

and neuroendocrine tissues (Gaudino *et al.*, 1995) and complete inactivation of Ron in mice leads to early embryonic lethality (Muraoka *et al.*, 1999).

The Eph Family

The Eph family is the largest of the RTK families. It consists of 9 EphA family members that bind to 5 glycosylphosphatidylinositol (GPI)-linked ephrin-A ligands, and 5 EphB members that bind three transmembrane ephrin-B ligands (Pasquale, 2010). Both ephrin-A:s and ephrin-B:s can mediate reversed signaling after binding to Ephs. Structurally, the Eph RTKs are composed of an ephrin binding domain and two leucine-rich domains extracellularly, and a kinase domain and a SAM domain intracellularly.

Eph RTKs have widespread functions in normal physiology (Pasquale, 2008). Eph-ephrin bidirectional signaling is important for communication between neurons, as well as between neurons and glial cells. Eph RTKs and ephrins also have important functions in cell contact-dependent communication in the immune system, for example, in the development of thymocytes into mature T cells in the thymus. Moreover, EphA-ephrin-A interactions are important in the control of insulin secretion in the β -cells, and EphB-ephrin-B interactions in bone development and in intestinal homeostasis (Pasquale, 2008).

The Tie Family

The extracellular parts of each of the two Tie RTKs are characterized by three Ig domains, three EGF domains and three leucine-rich domains (Huang *et al.*, 2010). Tie2 binds the agonistic ligands angiopoietin-1 (Angpt1) and -4; Angpt2, normally acts as a Tie2 antagonist (Maisonpierre *et al.*, 1997), but can under certain conditions have agonistic properties (Yuan *et al.*, 2009). Tie1 has no known ligand, but interacts with Tie2 and negatively regulates its activity (Kim *et al.*, 2006; Milner *et al.*, 2009; Saharinen *et al.*, 2005).

Activation of Tie2 promotes endothelial cell survival and the maintenance of the endothelial barrier (Huang *et al.*, 2010). Activation of Tie2 is also required for the formation of lymph vessels; in this case both Angpt1 and Angpt2 act as Tie2 agonists (Gale *et al.*, 2002).

Tie2-deficient mice die at embryonic day 10.5-12.5 since their vasculature does not develop beyond the primary capillary plexus. In contrast, Tie1-deficient mice die later, after embryonic day 13.5, when their vessels lose their integrity leading to edema and death (Augustin *et al.*, 2009).

The Trk Family

The Trk family consists of three members, i.e., TrkA, B, and C (Skaper, 2012). This RTK family is characterized by an extracellular ligand-binding part consisting of two cysteine-rich motifs, three leucine-rich repeats, and two Ig-like domains. The ligands for Trk receptors are denoted neurotrophins (NT), including NGF, BDNF, NT-3, NT-4, and NT-5. These factors are crucial for neuronal cell survival, growth, and plasticity during development. The different Trk receptors display selective ligand binding; TrkA binds NGF, TrkB binds BDNF, NT-4 as well

as NT-5, and TrkC binds NT-3. In addition, NT-3 can bind TrkA and B, but with lower affinity (Skaper, 2012).

In addition to binding to the Trk family of RTKs, the neurotrophins bind to p75NTR which belongs to the tumor neurosis factor (TNF) receptor family. It has been found that p75NTR assists in the binding of NGF to TrkA. In addition, p75NTR may suppress Trk receptor activation by the 'wrong' neurotrophin (Benedetti *et al.*, 1993). Thus, there is a close functional connection between the Trk and p75NTR receptors. It has been proposed that in the absence of Trk receptor expression, p75NTR can induce apoptosis (Brodeur *et al.*, 2009).

Trk RTKs have important roles during the development of the peripheral nervous system, as illustrated by the phenotypes of Trk knockout mice (Klein *et al.*, 1993, 1994; Smeyne *et al.*, 1994). The observed defects of the different Trk knockout mice display unique as well as overlapping functions in the nervous system of the different receptors. The TrkB knockout is the most severe.

Other RTKs

Besides the receptor families discussed above, there are also other less investigated RTKs. For example, discoidin domain receptors (DDR1 and 2) are widely expressed and are activated in a sustained manner by binding to certain types of collagens (Vogel *et al.*, 2006). Hence DDR1 and 2 are classified as non-integrin collagen receptors. Knockout of DDR genes in mice leads to skeletal and reproductive abnormalities (Fu *et al.*, 2013).

Another RTK that has been found to be relevant in the context of cancer, while its normal physiological role is unclear, is anaplastic lymphoma kinase (ALK) (Hallberg and Palmer, 2013). Ligands for mammalian ALK are heparin-binding growth factors called pleiotrophin and midkine. The kinase domains of ALK, and the related leukocyte tyrosine kinase (LTK), show similarities to the insulin receptor.

Ror1 and Ror2 were long thought to be orphan RTKs. However, their extracellular domains contain frizzled domains in addition to Ig-like, Cys-rich, and kringle domains, and Ror2 has recently been shown to bind Wnt1, 3, and 5 (Green *et al.*, 2008). Ror1 and 2 are expressed during embryonal development and are important for skeletal, respiratory and cardiac development. Moreover, Ror1 is expressed on chronic lymphatic leukemia cells, but not on normal lymphocytes (Daneshmanesh *et al.*, 2008), and on some other tumor types (Zhang *et al.*, 2012).

The Ryk RTK contains a Wnt-inhibitory factor-1 (WIF-1) domain and has, like Ror family members, been shown to bind Wnts, and to be important during neurogenesis, axon guidance, and synaptogenesis (Fradkin *et al.*, 2010).

The Ret tyrosine kinase has an extracellular domain that is related to the cell adhesion molecule cadherin; like cadherins, Ret may require Ca^{2+} binding to be functional (Nozaki *et al.*, 1998). The ligands for Ret are glial cell line-derived neurotrophic factor (GDNF), neurturin, persephin, and artemin (Mulligan, 2014). Ligand binding to Ret requires the expression of a coreceptor family denoted GDNF receptor α . Ret is required for the development of neural and genitourinary tissues.

Muscle-specific kinase (MuSK) is a receptor tyrosine kinase that is expressed on skeletal muscle cells and is important for the connection between muscle cells and motor neurons (Hubbard and Gnanasambandan, 2013). The activation of MuSK requires two proteins; agrin, produced by the motor neuron, interacts with LRP4 expressed on the muscle cells, which enables LRP4 to bind and activate MuSK.

Members of the TAM family of RTKs, which consists of Axl, Tyro3/Sky, and Mer, have important roles in immunoregulation and neurogenesis (Ji *et al.*, 2013; Paccez *et al.*, 2014). Moreover, Axl, the ligand of which is the vitamin K-dependent protein GAS6, has been found to be upregulated in several cancers and its expression correlates to poor prognosis. Axl promotes epithelial-mesenchymal transition and increased invasiveness; in addition, Axl may contribute to chemoresistance of cancers (Gjerdum *et al.*, 2010; Paccez *et al.*, 2014).

RTKs in Disease

RTKs have central roles in processes regulating cell proliferation, survival, and migration; they are frequently deregulated in various diseases, key examples being diabetes and cancer.

Diabetes

Diabetes is a disease that involves defects in insulin signaling. Type-1 diabetes is caused by loss of insulin production by the β -cells in the pancreas, whereas in type-2 diabetes the insulin receptor function is reduced, either by mutations in the receptor or in downstream signaling molecules making the cells resistant to insulin stimulation (Guo, 2014).

Malignancies

In malignancies, RTKs can be deregulated through different mechanisms, including amplification of the genes, activating point mutations or deletions, changed expression of RTKs or their ligands, and translocations creating chimeric proteins containing the RTK kinase domain.

For example, amplification of the ErbB2 gene occurs in 20% of breast carcinomas, overexpression of EGFR occurs in a wide array of carcinomas and a truncated form of EGFR called EGFRvIII has been observed in brain and prostate tumors (Yarden and Pines, 2012), and activating mutations in c-Kit or PDGFR α occur in gastrointestinal stromal tumors (Heinrich *et al.*, 2003). Fusion proteins of RTK kinase domains of the PDGF and FGF receptor families, with different components that have in common that they can oligomerize, occur in hypereosinophilic syndrome and leukemias (Matsumura *et al.*, 2008); fusion proteins with the ALK, Ros, Ret, Trks, FGFRs, Axl, and PDGFR α RTK kinase domains have been observed in lung cancer, colorectal cancer, breast cancer, thyroid cancer, and other cancers (Shaw *et al.*, 2013). In addition, autocrine stimulation of RTKs, as a consequence of production by tumor cells of the corresponding ligand, has been observed in a large number of tumors. All these mechanisms act directly on the cancer cells enhancing their survival and proliferation.

In addition, RTKs also indirectly contribute to tumor growth by regulating angiogenesis. The angiogenesis process is complex and involves many different signaling molecules; VEGF produced by tumor cells has the ability to promote nearby blood vessel endothelial cells to proliferate and migrate toward the tumor (Olsson *et al.*, 2006). Pericytes is another cell type required for normal blood vessel function and a key factor regulating them is PDGF (Lindahl *et al.*, 1997).

RTK Inhibitors in Treatment

The recurrent involvement of overactive RTKs in cancer has stimulated the development of therapeutic RTK antagonists. The clinically used RTK antagonists fall into two major groups; low molecular weight compounds that penetrate the plasma membrane and inhibit the kinase activities of RTKs, and antibodies recognizing the extracellular parts of RTKs, the binding of which interfere with RTK function or cause recruitment of immune cells.

The first tyrosine kinase inhibitor that was introduced into the clinic was imatinib that inhibits the kinase of the Bcr-Abl fusion protein which drives chronic myeloid leukemia. Imatinib also inhibits SCF and PDGF receptor kinases and are therefore also used to treat gastrointestinal stromal tumor (GIST) and other tumors driven by these receptors (Heinrich *et al.*, 2003). A number of other inhibitors have subsequently been developed, including lapatinib, erlotinib and gefitinib against EGF receptor family kinases, which are used to treat, for example, lung cancer (Shawver, 2002), sunitinib which inhibits several kinases including SCFR, PDGFRs, VEGFRs, Flt3, and Ret which is used to treat GIST and renal cell carcinoma (Chow and Eckhardt, 2007), and several other inhibitors.

Monoclonal antibodies that are used clinically include trastuzumab against ErbB2 which is used to treat breast cancer with ErbB2 amplification, cetuximab, and panitumumab against the EGF receptor which are used to treat colorectal cancer and head and neck cancer (Reichert and Valge-Archer, 2007). In addition, a monoclonal antibody against VEGF, bevacizumab, is used for anti-angiogenic therapy.

Unfortunately, drug resistance very often develops after some time of anti-RTK treatment. Mechanisms involve the appearance of RTK mutants that are insensitive to the inhibitor, activation of parallel signaling pathways that take over, or activation of membrane transporters that pump out the drug from the tumor cell (Holohan *et al.*, 2013). Anti-RTK treatment of malignancies therefore most likely need to be combined with other types of treatment in order to achieve complete tumor remission.

See also: Cell Communication: Intracellular Pathways: SH3 and SH2: Prototypic Domains to Mediate Regulatory Mechanisms in the Cell; The MAPK Signaling Cascades; The PLC Pathway. Horizontal Integration: Omics: Phosphoproteomics: Approaches, Developments, and Challenges; The Advent of Mass Spectrometry-Based Proteomics in Systems Biology Research. Organelles: Structure and Function: Early Endosomal Compartments

References

- Adams, T.E., Epa, V.C., Garrett, T.P., Ward, C.W., 2000. Structure and function of the type 1 insulin-like growth factor receptor. *Cellular and Molecular Life Sciences* 57 (7), 1050–1093.
- van Amerongen, R., Mikels, A., Nusse, R., 2008. Alternative wnt signaling is initiated by distinct receptors. *Science Signaling* 1 (35), re9.
- Amit, I., Citri, A., Shay, T., *et al.*, 2007. A module of negative feedback regulators defines growth factor signaling. *Nature Genetics* 39 (4), 503–512.
- Andrae, J., Gallini, R., Betsholtz, C., 2008. Role of platelet-derived growth factors in physiology and medicine. *Genes & Development* 22 (10), 1276–1312.
- Arkhipov, A., Shan, Y., Das, R., *et al.*, 2013. Architecture and membrane interactions of the EGF receptor. *Cell* 152 (3), 557–569.
- Augustin, H.G., Koh, G.Y., Thurston, G., Alitalo, K., 2009. Control of vascular morphogenesis and homeostasis through the angiopoietin-Tie system. *Nature Reviews Molecular Cell Biology* 10 (3), 165–177.
- Avraham, R., Yarden, Y., 2011. Feedback regulation of EGFR signalling: Decision making by early and delayed loops. *Nature Reviews Molecular Cell Biology* 12 (2), 104–117.
- Bae, Y.S., Kang, S.W., Seo, M.S., *et al.*, 1997. Epidermal growth factor (EGF)-induced generation of hydrogen peroxide. Role in EGF receptor-mediated tyrosine phosphorylation. *Journal of Biological Chemistry* 272, 217–221.
- Ball, S.G., Shuttleworth, C.A., Kiely, C.M., 2007. Vascular endothelial growth factor can signal through platelet-derived growth factor receptors. *Journal of Cell Biology* 177 (3), 489–500.
- Barton, W.A., Tzvetkova-Robev, D., Miranda, E.P., *et al.*, 2006. Crystal structures of the Tie2 receptor ectodomain and the angiopoietin-2-Tie2 complex. *Nature Structural & Molecular Biology* 13 (6), 524–532.
- Benedetti, M., Levi, A., Chao, M.V., 1993. Differential expression of nerve growth factor receptors leads to altered binding affinity and neurotrophin responsiveness. *Proceedings of the National Academy of Sciences of the United States of America* 90 (16), 7859–7863.
- Birchmeier, C., Gherardi, E., 1998. Developmental roles of HGF/SF and its receptor, the c-Met tyrosine kinase. *Trends in Cell Biology* 8 (10), 404–410.
- Blume-Jensen, P., Hunter, T., 2001. Oncogenic kinase signalling. *Nature* 411 (6835), 355–365.
- Blume-Jensen, P., Jiang, G., Hyman, R., *et al.*, 2000. Kit/stem cell factor receptor-induced activation of phosphatidylinositol 3'-kinase is essential for male fertility. *Nature Genetics* 24 (2), 157–162.
- Blume-Jensen, P., Siegbahn, A., Stabel, S., Heldin, C.-H., Rönstrand, L., 1993. Increased Kit/SCF receptor induced mitogenicity but abolished cell motility after inhibition of protein kinase C. *EMBO Journal* 12 (11), 4199–4209.
- Bollag, G.E., Roth, R.A., Beaudoin, J., Mochly-Rosen, D., Koshland Jr., D.E., 1986. Protein kinase C directly phosphorylates the insulin receptor *in vitro* and reduces its protein-tyrosine kinase activity. *Proceedings of the National Academy of Sciences of the United States of America* 83 (16), 5822–5824.
- Boucher, P., Liu, P., Gotthardt, M., *et al.*, 2002. Platelet-derived growth factor mediates tyrosine phosphorylation of the cytoplasmic domain of the low density lipoprotein receptor-related protein in caveolae. *Journal of Biological Chemistry* 277 (18), 15507–15513.
- Brodeur, G.M., Minturn, J.E., Ho, R., *et al.*, 2009. Trk receptor expression and inhibition in neuroblastomas. *Clinical Cancer Research* 15 (10), 3244–3250.
- Burgess, A.W., Cho, H.S., Eigenbrot, C., *et al.*, 2003. An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. *Molecular Cell* 12 (3), 541–552.
- Cabero, N.B., Zhou, Y., Li, W., 2010. Tubby and tubby-like protein 1 are new MerTK ligands for phagocytosis. *EMBO Journal* 29 (23), 3898–3910.
- Calver, A.R., Hall, A.C., Yu, W.P., *et al.*, 1998. Oligodendrocyte population dynamics and the role of PDGF *in vivo*. *Neuron* 20 (5), 869–882.
- Carpenter, G., Liao, H.-J., 2013. Receptor tyrosine kinases in the nucleus. *Cold Spring Harbor Perspectives in Biology*. a008979.
- Chen, X., Liu, H., Focia, P.J., Shim, A.H., He, X., 2008. Structure of macrophage colony stimulating factor bound to FMS: Diverse signaling assemblies of class III receptor tyrosine kinases. *Proceedings of the National Academy of Sciences of the United States of America* 105 (47), 18267–18272.
- Chiara, F., Bishayee, S., Heldin, C.-H., Demoulin, J.-B., 2004. Autoinhibition of the platelet-derived growth factor β receptor tyrosine kinase by its C-terminal tail. *Journal of Biological Chemistry* 279 (19), 19732–19738.
- Chow, L.Q., Eckhardt, S.G., 2007. Sunitinib: From rational design to clinical efficacy. *Journal of Clinical Oncology* 25 (7), 884–896.
- Citri, A., Yarden, Y., 2006. EGF-ERBB signalling: Towards the systems level. *Nature Reviews Molecular Cell Biology* 7 (7), 505–516.
- Dai, X.M., Ryan, G.R., Hapel, A.J., *et al.*, 2002. Targeted disruption of the mouse colony-stimulating factor 1 receptor gene results in osteopetrosis, mononuclear phagocyte deficiency, increased primitive progenitor cell frequencies, and reproductive defects. *Blood* 99 (1), 111–120.
- Daneshmandi, A.H., Mikaelsson, E., Jeddi-Tehrani, M., *et al.*, 2008. Ror1, a cell surface receptor tyrosine kinase is expressed in chronic lymphocytic leukemia and may serve as a putative target for therapy. *International Journal of Cancer* 123 (5), 1190–1195.
- Ding, Y., Liu, Z., Desai, S., *et al.*, 2012. Receptor tyrosine kinase ErbB2 translocates into mitochondria and regulates cellular metabolism. *Nature Communications* 3, 1271.
- Dode, C., Levilliers, J., Dupont, J.M., *et al.*, 2003. Loss-of-function mutations in FGFR1 cause autosomal dominant Kallmann syndrome. *Nature Genetics* 33 (4), 463–465.
- Du, Y., Shen, J., Hsu, J.L., *et al.*, 2014. Syntaxin 6-mediated Golgi translocation plays an important role in nuclear functions of EGFR through microtubule-dependent trafficking. *Oncogene* 33 (6), 756–770.
- Dumont, D.J., Jussila, L., Taipale, J., *et al.*, 1998. Cardiovascular failure in mouse embryos deficient in VEGF receptor-3'. *Science* 282 (5390), 946–949.
- Endres, N.F., Das, R., Smith, A.W., *et al.*, 2013. Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* 152 (3), 543–556.
- Fong, G.-H., Rossant, J., Gertsenstein, M., Breitman, M.L., 1995. Role of the Flt-1 receptor tyrosine kinase in regulating the assembly of vascular endothelium. *Nature* 376, 66–70.
- Fong, G.-H., Zhang, L., Bryce, D.-M., Peng, J., 1999. Increased hemangioblast commitment, not vascular disorganization, is the primary defect in flt-1 knock-out mice. *Development* 126 (13), 3015–3025.
- Fradkin, L.G., Dura, J.M., Noordermeer, J.N., 2010. Ryks: New partners for Wnts in the developing and regenerating nervous system. *Trends in Neurosciences* 33 (2), 84–92.
- Fruttiger, M., Calver, A.R., Richardson, W.D., 2000. Platelet-derived growth factor is constitutively secreted from neuronal cell bodies but not from axons. *Current Biology* 10 (20), 1283–1286.
- Fu, H.L., Valiathan, R.R., Arkwright, R., *et al.*, 2013. Discoidin domain receptors: Unique receptor tyrosine kinases in collagen-mediated signaling. *Journal of Biological Chemistry* 288 (11), 7430–7437.
- Fuller, S.J., Sivarajah, K., Sugden, P.H., 2008. ErbB receptors, their ligands, and the consequences of their activation and inhibition in the myocardium. *Journal of Molecular and Cellular Cardiology* 44 (5), 831–854.
- Gale, N.W., Thurston, G., Hackett, S.F., *et al.*, 2002. Angiopoietin-2 is required for postnatal angiogenesis and lymphatic patterning, and only the latter role is rescued by Angiopoietin-1'. *Developmental Cell* 3 (3), 411–423.
- Gandino, L., Di Renzo, M.F., Giordano, S., Bussolino, F., Comoglio, P.M., 1990. Protein kinase-c activation inhibits tyrosine phosphorylation of the c-met protein. *Oncogene* 5 (5), 721–725.
- Gassmann, M., Casagrande, F., Orioli, D., *et al.*, 1995. Aberrant neural and cardiac development in mice lacking the ErbB4 neuregulin receptor. *Nature* 378 (6555), 390–394.
- Gaudino, G., Avantaggiato, V., Follenzi, A., *et al.*, 1995. The proto-oncogene RON is involved in development of epithelial, bone and neuro-endocrine tissues. *Oncogene* 11 (12), 2627–2637.
- Giannakou, M.E., Partridge, L., 2007. Role of insulin-like signalling in *Drosophila* lifespan. *Trends in Biochemical Sciences* 32 (4), 180–188.
- Gjerdum, C., Tiron, C., Hoiby, T., *et al.*, 2010. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013. MET: A critical player in tumorigenesis and therapeutic target. *Cold Spring Harbor Perspectives in Biology* 5 (7), a009209.
- Green, J.L., Kuntz, S.G., Sternberg, P.W., 2008. Ror receptor tyrosine kinases: Orphans no more. *Trends in Cell Biology* 18 (11), 536–544.
- Griffith, J., Black, J., Faerman, C., *et al.*, 2004. Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proceedings of the National Academy of Sciences of the United States of America* 107 (3), 1124–1129.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14 (3), 166–180.
- Graveel, C.R., Tolbert, D., Vande Woude, G.F., 2013.

- Hallberg, B., Palmer, R.H., 2013. Mechanistic insight into ALK receptor tyrosine kinase in human cancer biology. *Nature Reviews Cancer* 13 (10), 685–700.
- Hamilton, J.A., 2008. Colony-stimulating factors in inflammation and autoimmunity. *Nature Reviews Immunology* 8 (7), 533–544.
- Heinrich, M.C., Corless, C.L., Demetri, G.D., *et al.*, 2003. Kinase mutations and imatinib response in patients with metastatic gastrointestinal stromal tumor. *Journal of Clinical Oncology* 21 (23), 4342–4349.
- Heldin, C.-H., 1995. Dimerization of cell surface receptors in signal transduction. *Cell* 80 (2), 213–223.
- Heldin, C.H., Lennartsson, J., 2013. Structural and functional properties of platelet-derived growth factor and stem cell factor receptors. *Cold Spring Harbor Perspectives in Biology*. a009100.
- Heldin, C.-H., Lu, B., Evans, R., Silvio Gutkind, J., 2014. Signals and receptors. In: Cantley, L.C., Hunter, T., Sever, R., Thorner, J. (Eds.), *Signal Transduction: Principles, Pathways, and Processes*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 3–30. Available at: http://www.cshlpress.com/default.tpl?cart=14199889722625091&fromlink=T&linkaction=full&linksortby=oap_title&eqSKUdata=991 (accessed 01.01.15).
- Hellberg, C., Schmees, C., Karlsson, S., Åhgren, A., Heldin, C.-H., 2009. Activation of protein kinase C α is necessary for sorting the PDGF β -receptor to Rab4a-dependent recycling. *Molecular Biology of the Cell* 20 (12), 2856–2863.
- Himanen, J.P., Nikolov, D.B., 2003. Eph signaling: A structural view. *Trends in Neurosciences* 26 (1), 46–51.
- Holohan, C., Van Schaeybroeck, S., Longley, D.B., Johnston, P.G., 2013. Cancer drug resistance: An evolving paradigm. *Nature Reviews Cancer* 13 (10), 714–726.
- Huang, F., Kirkpatrick, D., Jiang, X., Gygi, S., Sorkin, A., 2006. Differential regulation of EGF receptor internalization and degradation by multiubiquitination within the kinase domain. *Molecular Cell* 21 (6), 737–748.
- Huang, H., Bhat, A., Woodnutt, G., Lappe, R., 2010. Targeting the ANGPT-TIE2 pathway in malignancy. *Nature Reviews Cancer* 10 (8), 575–585.
- Hubbard, S.R., 2004. Juxtamembrane autoinhibition in receptor tyrosine kinases. *Nature Reviews Molecular Cell Biology* 5 (6), 464–471.
- Hubbard, S.R., Gnanasambandan, K., 2013. Structure and activation of MuSK, a receptor tyrosine kinase central to neuromuscular junction formation. *Biochimica et Biophysica Acta* 1834 (10), 2166–2169.
- Huizinga, J.D., Thuneberg, L., Kluppel, M., *et al.*, 1995. W/klt gene required for interstitial cells of Cajal and for intestinal pacemaker activity. *Nature* 373 (6512), 347–349.
- Hunter, T., Ling, N., Cooper, J.A., 1984. Protein kinase C phosphorylation of the EGF receptor at a threonine residue close to the cytoplasmic face of the plasma membrane. *Nature* 311 (5985), 480–483.
- Hynes, N.E., Schlange, T., 2006. Targeting ADAMS and ERBBs in lung cancer. *Cancer Cell* 10 (1), 7–11.
- Ivaska, J., Heino, J., 2011. Cooperation between integrins and growth factor receptors in signaling and endocytosis. *Annual Review of Cell and Developmental Biology* 27, 291–320.
- Ji, R., Tian, S., Lu, H.J., *et al.*, 2013. TAM receptors affect adult brain neurogenesis by negative regulation of microglial cell activation. *Journal of Immunology* 191 (12), 6165–6177.
- Jiang, G., Hunter, T., 1999. Receptor signaling: When dimerization is not enough. *Current Biology* 9 (15), R568–R571.
- Jura, N., Endres, N.F., Engel, K., *et al.*, 2009. Mechanism for activation of the EGF receptor catalytic domain by the juxtamembrane segment. *Cell* 137 (7), 1293–1307.
- Jura, N., Zhang, X., Endres, N.F., *et al.*, 2011. Catalytic control in the EGF receptor and its connection to general kinase regulatory mechanisms. *Molecular Cell* 42 (1), 9–22.
- Kasuga, K., Kaneko, H., Nishizawa, M., Onodera, O., Ikeuchi, T., 2007. Generation of intracellular domain of insulin receptor tyrosine kinase by γ -secretase. *Biochemical and Biophysical Research Communications* 360 (1), 90–96.
- Katafuchi, T., Li, A.J., Hirota, S., Kitamura, Y., Hori, T., 2000. Impairment of spatial learning and hippocampal synaptic potentiation in c-kit mutant rats. *Learning & Memory* 7 (6), 383–392.
- Kim, K.L., Shin, I.S., Kim, J.M., *et al.*, 2006. Interaction between Tie receptors modulates angiogenic activity of angiopoietin2 in endothelial progenitor cells. *Cardiovascular Research* 72 (3), 394–402.
- Kim, N., Stiegler, A.L., Cameron, T.O., *et al.*, 2008. Lrp4 is a receptor for Agrin and forms a complex with MuSK. *Cell* 135 (2), 334–342.
- Kissel, H., Timokhina, I., Hardy, M.P., *et al.*, 2000. Point mutation in kit receptor tyrosine kinase reveals essential roles for kit signaling in spermatogenesis and oogenesis without affecting other kit responses. *EMBO Journal* 19 (6), 1312–1326.
- Kitamura, T., Kahn, C.R., Accili, D., 2003. Insulin receptor knockout mice. *Annual Review of Physiology* 65, 313–332.
- Klein, R., Silos-Santiago, I., Smeyne, R.J., *et al.*, 1994. Disruption of the neurotrophin-3 receptor gene *trkC* eliminates la muscle afferents and results in abnormal movements. *Nature* 368 (6468), 249–251.
- Klein, R., Smeyne, R.J., Wurst, W., *et al.*, 1993. Targeted disruption of the *trkB* neurotrophin receptor gene results in nervous system lesions and neonatal death. *Cell* 75 (1), 113–122.
- Knowles, P.P., Murray-Rust, J., Kjaer, S., *et al.*, 2006. Structure and chemical inhibition of the RET tyrosine kinase domain. *Journal of Biological Chemistry* 281 (44), 33577–33587.
- Lee, K.F., Simon, H., Chen, H., *et al.*, 1995. Requirement for neuregulin receptor *erbB2* in neural and cardiac development. *Nature* 378 (6555), 394–398.
- Lemke, G., 2013. Biology of the TAM receptors. *Cold Spring Harbor Perspectives in Biology* 5 (11), a009076.
- Lemmon, M.A., Schlessinger, J., 2010. Cell signaling by receptor tyrosine kinases. *Cell* 141 (7), 1117–1134.
- Lidén, A., Berg, A., Nedrebø, T., Reed, R.K., Rubin, K., 2006. Platelet-derived growth factor BB-mediated normalization of dermal interstitial fluid pressure after mast cell degranulation depends on β 3 but not β 1 integrins. *Circulation Research* 98 (5), 635–641.
- Lin, H., Lee, E., Hestir, K., *et al.*, 2008. Discovery of a cytokine and its receptor by functional screening of the extracellular proteome. *Science* 320 (5877), 807–811.
- Lin, S.-Y., Makino, K., Xia, W., *et al.*, 2001. Nuclear localization of EGF receptor and its potential new role as a transcription factor. *Nature Cell Biology* 3 (9), 802–808.
- Lindahl, P., Johansson, B.R., Levéen, P., Betsholtz, C., 1997. Pericyte loss and microaneurysm formation in PDGF-B-deficient mice. *Science* 277 (5323), 242–245.
- Lisabeth, E.M., Falivelli, G., Pasquale, E.B., 2013. Eph receptor signaling and ephrins. *Cold Spring Harbor Perspectives in Biology* 5 (9), a009159.
- Lo, H.W., Hsu, S.C., Ali-Seyed, M., *et al.*, 2005. Nuclear interaction of EGFR and STAT3 in the activation of the iNOS/NO pathway. *Cancer Cell* 7 (6), 575–589.
- Locascio, L.E., Donoghue, D.J., 2013. KIDs rule: Regulatory phosphorylation of RTKs. *Trends in Biochemical Sciences* 38 (2), 75–84.
- López-Otín, C., Hunter, T., 2010. The regulatory crosstalk between kinases and proteases in cancer. *Nature Reviews Cancer* 10 (4), 278–292.
- Loukinova, E., Ranganathan, S., Kuznetsov, S., *et al.*, 2002. Platelet-derived growth factor (PDGF)-induced tyrosine phosphorylation of the low density lipoprotein receptor-related protein (LRP). Evidence for integrated co-receptor function between LRP and the PDGF. *Journal of Biological Chemistry* 277 (18), 15499–15506.
- Lyman, S.D., Jacobsen, S.E., 1998. c-kit ligand and Flt3 ligand: Stem/progenitor cell factors with overlapping yet distinct activities. *Blood* 91 (4), 1101–1134.
- Maisonpierre, P.C., Suri, C., Jones, P.F., *et al.*, 1997. Angiopoietin-2, a natural antagonist for Tie2 that disrupts *in vivo* angiogenesis. *Science* 277 (5322), 55–60.
- Massie, C., Mills, I.G., 2006. The developing role of receptors and adaptors. *Nature Reviews Cancer* 6 (5), 403–409.
- Matsumura, I., Mizuki, M., Kanakura, Y., 2008. Roles for deregulated receptor tyrosine kinases and their downstream signaling molecules in hematologic malignancies. *Cancer Science* 99 (3), 479–485.
- McElroy, B., Powell, J.C., McCarthy, J.V., 2007. The insulin-like growth factor 1 (IGF-1) receptor is a substrate for γ -secretase-mediated intramembrane proteolysis. *Biochemical and Biophysical Research Communications* 358 (4), 1136–1141.
- Miaczynska, M., Pelkmans, L., Zerial, M., 2004. Not just a sink: Endosomes in control of signal transduction. *Current Opinion in Cell Biology* 16 (4), 400–406.
- Miettinen, P.J., Berger, J.E., Meneses, J., *et al.*, 1995. Epithelial immaturity and multiorgan failure in mice lacking epidermal growth factor receptor. *Nature* 376 (6538), 337–341.
- Milner, C.S., Hansen, T.M., Singh, H., Brindle, N.P., 2009. Roles of the receptor tyrosine kinases Tie1 and Tie2 in mediating the effects of angiopoietin-1 on endothelial permeability and apoptosis. *Microvascular Research* 77 (2), 187–191.
- Mol, C.D., Dougan, D.R., Schneider, T.R., *et al.*, 2004. Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. *Journal of Biological Chemistry* 279 (30), 31655–31663.
- Mosesson, Y., Mills, G.B., Yarden, Y., 2008. Derailed endocytosis: An emerging feature of cancer. *Nature Reviews Cancer* 8 (11), 835–850.
- Mosesson, Y., Shtiegman, K., Katz, M., *et al.*, 2003. Endocytosis of receptor tyrosine kinases is driven by monoubiquitylation, not polyubiquitylation. *Journal of Biological Chemistry* 278 (24), 21323–21326.
- Motro, B., Wojtowicz, J.M., Bernstein, A., van der Kooy, D., 1996. Steel mutant mice are deficient in hippocampal learning but not long-term potentiation. *Proceedings of the National Academy of Sciences of the United States of America* 93 (5), 1808–1813.

- Mulligan, L.M., 2014. RET revisited: Expanding the oncogenic portfolio. *Nature Reviews Cancer* 14 (3), 173–186.
- Muraoka, R.S., Sun, W.Y., Colbert, M.C., *et al.*, 1999. The Ron/STK receptor tyrosine kinase is essential for peri-implantation development in the mouse. *Journal of Clinical Investigation* 103 (9), 1277–1285.
- Nakayama, A., Nakayama, M., Turner, C.J., *et al.*, 2013. Ephrin-B2 controls PDGFR β internalization and signaling. *Genes & Development* 27 (23), 2576–2589.
- Nakayama, M., Nakayama, A., van Lessen, M., *et al.*, 2013. Spatial regulation of VEGF receptor endocytosis in angiogenesis. *Nature Cell Biology* 15 (3), 249–260.
- Narresh, A., Long, W., Vidal, G.A., *et al.*, 2006. The ERBB4/HER4 intracellular domain 4ICD is a BH3-only protein promoting apoptosis of breast cancer cells. *Cancer Research* 66 (12), 6412–6420.
- Ni, C.-Y., Murphy, M.P., Golde, T.E., Carpenter, G., 2001. γ -Secretase cleavage and nuclear localization of ErbB-4 receptor tyrosine kinase. *Science* 294, 2179–2181.
- Nozaki, C., Asai, N., Murakami, H., *et al.*, 1998. Calcium-dependent Ret activation by GDNF and neurturin. *Oncogene* 16 (3), 293–299.
- Ohlsson, R., Falck, P., Hellström, M., *et al.*, 1999. PDGFB regulates the development of the labyrinthine layer of the mouse fetal placenta. *Developmental Biology* 212 (1), 124–136.
- Olsson, A.-K., Dimberg, A., Kreuger, J., Claesson-Welsh, L., 2006. VEGF receptor signalling – in control of vascular function. *Nature Reviews Molecular Cell Biology* 7 (5), 359–371.
- Orian-Rousseau, V., 2010. CD44, a therapeutic target for metastasising tumours. *European Journal of Cancer* 46 (7), 1271–1277.
- Oved, S., Mosesson, Y., Zwang, Y., *et al.*, 2006. Conjugation to Nedd8 instigates ubiquitylation and down-regulation of activated receptor tyrosine kinases. *Journal of Biological Chemistry* 281 (31), 21640–21651.
- Pacecz, J.D., Vogelsang, M., Parker, M.I., Zerbini, L.F., 2014. The receptor tyrosine kinase Axl in cancer: Biological functions and therapeutic implications. *International Journal of Cancer* 134 (5), 1024–1033.
- Parachoniak, C.A., Park, M., 2012. Dynamics of receptor trafficking in tumorigenicity. *Trends in Cell Biology* 22 (5), 231–240.
- Pasquale, E.B., 2008. Eph-ephrin bidirectional signaling in physiology and disease. *Cell* 133 (1), 38–52.
- Pasquale, E.B., 2010. Eph receptors and ephrins in cancer: Bidirectional signalling and beyond. *Nature Reviews Cancer* 10 (3), 165–180.
- Pawson, T., 2004. Specificity in signal transduction: From phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell* 116 (2), 191–203.
- Pollak, M., 2012. The insulin and insulin-like growth factor receptor family in neoplasia: An update. *Nature Reviews Cancer* 12 (3), 159–169.
- Raiborg, C., Stenmark, H., 2009. The ESCRT machinery in endosomal sorting of ubiquitylated membrane proteins. *Nature* 458 (7237), 445–452.
- Red Brewer, M., Choi, S.H., Alvarado, D., *et al.*, 2009. The juxtamembrane region of the EGF receptor functions as an activation domain. *Molecular Cell* 34 (6), 641–651.
- Reichert, J.M., Valge-Archer, V.E., 2007. Development trends for monoclonal antibody cancer therapeutics. *Nature Reviews Drug Discovery* 6 (5), 349–356.
- Riethmacher, D., Sonnenberg-Riethmacher, E., Brinkmann, V., *et al.*, 1997. Severe neuropathies in mice with targeted mutations in the ErbB3 receptor. *Nature* 389 (6652), 725–730.
- Robson, M.C., Phillips, L.G., Thomason, A., Robson, L.E., Pierce, G.F., 1992. Platelet-derived growth factor BB for the treatment of chronic pressure ulcers. *Lancet* 339, 23–25.
- Rodt, S.Å., Åhlén, K., Berg, A., Rubin, K., Reed, R.K., 1996. A novel physiological function for platelet-derived growth factor-BB in rat dermis. *Journal of Physiology* 495 (Pt 1), 193–200.
- Rohmann, E., Brunner, H.G., Kayserili, H., *et al.*, 2006. Mutations in different components of FGF signaling in LADD syndrome. *Nature Genetics* 38 (4), 414–417.
- Saharinen, P., Kerkela, K., Ekman, N., *et al.*, 2005. Multiple angiopoietin recombinant proteins activate the Tie1 receptor tyrosine kinase and promote its interaction with Tie2. *Journal of Cell Biology* 169 (2), 239–243.
- Sardi, S.P., Murtie, J., Koirala, S., Patten, B.A., Corfas, G., 2006. Presenilin-dependent ErbB4 nuclear signaling regulates the timing of astrogenesis in the developing brain. *Cell* 127 (1), 185–197.
- Sawamiphak, S., Seidel, S., Essmann, C.L., *et al.*, 2010. Ephrin-B2 regulates VEGFR2 function in developmental and tumour angiogenesis. *Nature* 465 (7297), 487–491.
- Schlee, S., Carmillo, P., Whitty, A., 2006. Quantitative analysis of the activation mechanism of the multicomponent growth-factor receptor Ret. *Nature Chemical Biology* 2 (11), 636–644.
- Schlessinger, J., Plotnikov, A.N., Ibrahimi, O.A., *et al.*, 2000. Crystal structure of a ternary FGF-FGFR-heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Molecular Cell* 6 (3), 743–750.
- Schmees, C., Villaseñor, R., Zheng, W., *et al.*, 2012. Macropinocytosis of the PDGF β -receptor promotes fibroblast transformation by H-RasG12V. *Molecular Biology of the Cell* 23 (13), 2571–2582.
- Seiradake, E., Harlos, K., Sutton, G., Aricescu, A.R., Jones, E.Y., 2010. An extracellular steric seeding mechanism for Eph-ephrin signaling platform assembly. *Nature Structural & Molecular Biology* 17 (4), 398–402.
- Shalaby, F., Rossant, J., Yamaguchi, T.P., *et al.*, 1995. Failure of blood-island formation and vasculogenesis in Flk-1-deficient mice. *Nature* 376, 62–66.
- Shaw, A.T., Hsu, P.P., Awad, M.M., Engelman, J.A., 2013. Tyrosine kinase gene rearrangements in epithelial malignancies. *Nature Reviews Cancer* 13 (11), 772–787.
- Shawver, L.K., Slamon, D., Ullrich, A., 2002. Smart drugs: Tyrosine kinase inhibitors in cancer therapy. *Cancer Cell* 1, 117–123.
- Shewchuk, L.M., Hassell, A.M., Ellis, B., *et al.*, 2000. Structure of the Tie2 RTK domain: Self-inhibition by the nucleotide binding loop, activation loop, and C-terminal tail. *Structure* 8 (11), 1105–1113.
- Shilo, B.Z., 2005. Regulating the dynamics of EGF receptor signaling in space and time. *Development* 132 (18), 4017–4027.
- Sibilia, M., Steinbach, J.P., Stingl, L., Aguzzi, A., Wagner, E.F., 1998. A strain-independent postnatal neurodegeneration in mice lacking the EGF receptor. *EMBO Journal* 17 (3), 719–731.
- Sibilia, M., Wagner, E.F., 1995. Strain-dependent epithelial defects in mice lacking the EGF receptor. *Science* 269 (5221), 234–238.
- Skaper, S.D., 2012. The neurotrophin family of neurotrophic factors: An overview. *Methods in Molecular Biology* 846, 1–12.
- Smeyne, R.J., Klein, R., Schnapp, A., *et al.*, 1994. Severe sensory and sympathetic neuropathies in mice carrying a disrupted Trk/NGF receptor gene. *Nature* 368 (6468), 246–249.
- Song, S., Rosen, K.M., Corfas, G., 2013. Biological function of nuclear receptor tyrosine kinase action. *Cold Spring Harbor Perspectives in Biology*. a009001.
- Songyang, Z., Cantley, L.C., 2004. ZIP codes for delivering SH2 domains. *Cell* 116 (2 suppl), S41–S43. 2 p following S48.
- Soriano, P., 1994. Abnormal kidney development and hematological disorders in PDGF β -receptor mutant mice. *Genes & Development* 8 (16), 1888–1896.
- Soriano, P., 1997. The PDGF α receptor is required for neural crest cell development and for normal patterning of the somites. *Development* 124 (14), 2691–2700.
- Sorkin, A., Goh, L.K., 2009. Endocytosis and intracellular trafficking of ErbBs. *Experimental Cell Research*, 315 (4), 683–696.
- Stirewalt, D.L., Radich, J.P., 2003. The role of FLT3 in haematopoietic malignancies. *Nature Reviews Cancer* 3 (9), 650–665.
- Tammela, T., Alitalo, K., 2010. Lymphangiogenesis: Molecular mechanisms and future promise. *Cell* 140 (4), 460–476.
- Threadgill, D.W., Dlugosz, A.A., Hansen, L.A., *et al.*, 1995. Targeted disruption of mouse EGF receptor: Effect of genetic background on mutant phenotype. *Science* 269 (5221), 230–234.
- Till, J.H., Becerra, M., Watty, A., *et al.*, 2002. Crystal structure of the MuSK tyrosine kinase: Insights into receptor autoregulation. *Structure* 10 (9), 1187–1196.
- Tonks, N.K., 2006. Protein tyrosine phosphatases: From genes, to function, to disease. *Nature Reviews Molecular Cell Biology* 7 (11), 833–846.
- Trusolino, L., Bertotti, A., Comoglio, P.M., 2010. MET signalling: Principles and functions in development, organ regeneration and cancer. *Nature Reviews Molecular Cell Biology* 11 (12), 834–848.
- Turner, N., Grose, R., 2010. Fibroblast growth factor signalling: From development to cancer. *Nature Reviews Cancer* 10 (2), 116–129.
- Van den Akker, N.M., Winkel, L.C., Nisancioglu, M.H., *et al.*, 2008. PDGF-B signaling is important for murine cardiac development: Its role in developing atrioventricular valves, coronaries, and cardiac innervation. *Developmental Dynamics* 237 (2), 494–503.
- Verstraete, K., Savvides, S.N., 2012. Extracellular assembly and activation principles of oncogenic class III receptor tyrosine kinases. *Nature Reviews Cancer* 12 (11), 753–766.
- Vogel, W.F., Abdulhussein, R., Ford, C.E., 2006. Sensing extracellular matrix: An update on discoidin domain receptor function. *Cellular Signalling* 18 (8), 1108–1116.
- Wang, M.H., Zhang, R., Zhou, Y.Q., Yao, H.P., 2013. Pathogenesis of RON receptor tyrosine kinase in cancer cells: Activation mechanism, functional crosstalk, and signaling addiction. *Journal of Biomedical Research* 27 (5), 345–356.
- Wang, S.C., Lien, H.C., Xia, W., *et al.*, 2004. Binding at and transactivation of the COX-2 promoter by nuclear tyrosine kinase receptor ErbB-2. *Cancer Cell* 6 (3), 251–261.
- Wang, Y., Nakayama, M., Pitulescu, M.E., *et al.*, 2010. Ephrin-B2 controls VEGF-induced angiogenesis and lymphangiogenesis. *Nature* 465 (7297), 483–486.

- Ward, C.W., Lawrence, M.C., Streltsov, V.A., Adams, T.E., McKern, N.M., 2007. The insulin and EGF receptor structures: New insights into ligand-induced receptor activation. *Trends in Biochemical Sciences* 32 (3), 129–137.
- Ward, S.M., Burns, A.J., Torihashi, S., Harney, S.C., Sanders, K.M., 1995. Impaired development of interstitial cells and intestinal electrical rhythmicity in steel mutants. *American Journal of Physiology* 269 (6 Pt 1), C1577–C1585.
- Wehrle-Haller, B., 2003. The role of Kit-ligand in melanocyte development and epidermal homeostasis. *Pigment Cell Research* 16 (3), 287–296.
- Wehrman, T., He, X., Raab, B., *et al.*, 2007. Structural and mechanistic insights into nerve growth factor interactions with the TrkA and p75 receptors. *Neuron* 53 (1), 25–38.
- Wybenga-Groot, L.E., Baskin, B., Ong, S.H., *et al.*, 2001. Structural basis for autoinhibition of the EphB2 receptor tyrosine kinase by the unphosphorylated juxtamembrane region. *Cell* 106 (6), 745–757.
- Yang, Y., Yuzawa, S., Schlessinger, J., 2008. Contacts between membrane proximal regions of the PDGF receptor ectodomains are required for receptor activation but not for receptor dimerization. *Proceedings of the National Academy of Sciences of the United States of America* 105 (22), 7681–7686.
- Yarden, Y., Pines, G., 2012. The ERBB network: At last, cancer therapy meets systems biology. *Nature Reviews Cancer* 12 (8), 553–563.
- Yokoyama, N., Ischenko, I., Hayman, M.J., Miller, W.T., 2005. The C terminus of RON tyrosine kinase plays an autoinhibitory role. *Journal of Biological Chemistry* 280 (10), 8893–8900.
- Yuan, H.T., Khankin, E.V., Karumanchi, S.A., Parikh, S.M., 2009. Angiopoietin 2 is a partial agonist/antagonist of Tie2 signaling in the endothelium. *Molecular and Cellular Biology* 29 (8), 2011–2022.
- Yuzawa, S., Opatowsky, Y., Zhang, Z., *et al.*, 2007. Structural basis for activation of the receptor tyrosine kinase KIT by stem cell factor. *Cell* 130 (2), 323–334.
- von Zastrow, M., Sorkin, A., 2007. Signaling on the endocytic pathway. *Current Opinion in Cell Biology* 19 (4), 436–445.
- Zhang, S., Chen, L., Wang-Rodriguez, J., *et al.*, 2012. The onco-embryonic antigen ROR1 is expressed by a variety of human cancers. *American Journal of Pathology* 181 (6), 1903–1910.
- Zhang, X., Gureasko, J., Shen, K., Cole, P.A., Kuriyan, J., 2006. An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell* 125 (6), 1137–1149.
- Zhang, X., Ibrahimi, O.A., Olsen, S.K., *et al.*, 2006. Receptor specificity of the fibroblast growth factor family. The complete mammalian FGF family. *Journal of Biological Chemistry* 281 (23), 15694–15700.
- Zwang, Y., Yarden, Y., 2009. Systems biology of growth factor-induced receptor endocytosis. *Traffic* 10 (4), 349–363.

Cytokine Receptors and Their Ligands

U Dhagat and TL Nero, St. Vincent's Institute of Medical Research, Fitzroy, VIC, Australia

WL Kan and TR Hercus, Centre for Cancer Biology, SA Pathology and University of South Australia, Adelaide, SA, Australia

SE Broughton, St. Vincent's Institute of Medical Research, Fitzroy, VIC, Australia

NA de Weerd, BJ Jenkins, and A Mansell, MIMR-PHI Institute of Medical Research, Monash University, Clayton, VIC, Australia

MA Guthridge, Australian Centre for Blood Cancers, Monash University, Melbourne, VIC, Australia

PJ Hertzog, MIMR-PHI Institute of Medical Research, Monash University, Clayton, VIC, Australia

AF Lopez, Centre for Cancer Biology, SA Pathology and University of South Australia, Adelaide, SA, Australia

MW Parker, St. Vincent's Institute of Medical Research, Fitzroy, VIC, Australia and University of Melbourne, Parkville, VIC, Australia

© 2016 Elsevier Inc. All rights reserved.

Glossary

Cysteine knot fold A protein structural motif containing three disulfide bridges. Two of the disulfide bridges create a polypeptide loop through which a third disulfide bond passes, forming a structurally stable mechanically interlocked molecular substructure known as a rotaxane.

The cysteine knot fold occurs in many proteins across a wide range of species and include the growth factor cysteine knot, cyclotides, and the inhibitor cysteine knot.

Elbow The hinge-like region that exists between two adjacent FnIII domains in the same protein chain. The shape is reminiscent of an elbow, hence the name.

Fibronectin Type III (FnIII) domains FnIII domains belong to the immunoglobulin (Ig)-like protein family and adopt one of the most common extracellular protein folds. These domains typically consist of two β -sheets, each containing at least seven antiparallel β -strands, packed against each other in an arrangement known as a β -sandwich.

Immunoglobulin (Ig)-like domains Immunoglobulin (Ig) and Ig-like proteins are modular and their domains have been classified into four types; V-set, C1-set, C2-set, and I-set. Each domain consists of a β -sandwich structure, with differing numbers of β -strands in each of the β -sheets making up the β -sandwich.

Leucine-rich repeats As the name suggests, the leucine-rich repeat (LRR) is a repetitive sequence pattern rich in leucines and approximately 20–30 amino acids in length. Two or more consecutive LRRs can form curved linear

solenoid structures that are able to participate in protein–protein interactions.

Pseudokinase domains Traditionally defined as a kinase domain that is incapable of phosphoryl group transfer due to an incomplete catalytic motif. Pseudokinases are now known to act as signal transducers and are able to allosterically activate protein kinases.

SEFIR domain The name stands for 'similar expression to fibroblast-growth factor genes and IL-17R.' These signaling domains consist of six α -helices wrapped around a five-stranded parallel β -sheet core.

Sushi domain Sushi domains are based on a β -sandwich structure, one face contains three β -strands hydrogen-bonded to form a triple-stranded central region and the opposite face is formed by two separate β -strands. These exist in a wide variety of adhesion and complement proteins; are also known as short consensus repeats (SCR) and complement control protein (CCP) modules.

The Toll/IL-1 receptor signaling domain The Toll/IL-1 receptor (TIR) signaling domain was previously known as the Toll/IL-1 resistance domain and consists of three conserved sequence motifs designated box 1 (FDFAFISY), box 2 (GYKLC-RD-PG), and box 3 (a conserved W surrounded by basic residues). Boxes 1 and 2 are thought to be involved in the binding of signaling proteins, whereas it has been suggested that the main function of box 3 is directing receptor localization. The TIR domains of human Toll receptor 1 (TLR1) and TLR2 contain a five-stranded parallel β -sheet core surrounded by five α -helices.

Introduction

Cytokines are responsible for a wide range of biological processes, particularly in the survival, proliferation, differentiation, and functional activity of many types of cells. Abnormal cytokine levels or aberrations in their signaling pathways can lead to a variety of diseases, including cancers and inflammatory conditions reflecting their importance in normal hematopoiesis and immunity.

On the basis of common structural features, cytokine receptors are grouped into six major families: type I cytokine receptors, type II cytokine receptors, TNF receptors, IL-1 receptors, tyrosine kinase receptors, and chemokine receptors. The type I

(hematopoietic) cytokine receptors are characterized by having cytokine-binding modules (CBM) consisting of two adjacent β -sheet sandwich fibronectin type III (FnIII) domains. Two distinctive features of the CBM are, firstly, four highly conserved cysteine residues in the N-terminal FnIII domain and, secondly, a Trp-Ser-X-Trp-Ser motif (also known as the 'WSXWS' box, where X is a nonconserved residue) in the C-terminal FnIII domain. Both features are believed to be important for maintaining the structural integrity of the extracellular region to facilitate ligand-receptor interactions (Bazan, 1990a; Kernebeck *et al.*, 1999). The type I cytokine receptors also have sequence motifs in their cytoplasmic domains, called box 1 and box 2, that are responsible for receptor interaction with protein

partners. The type II (non-hematopoietic) cytokine receptors are similar to type I cytokine receptors except they do not possess the signature WSXWS motif.

Activation of the human growth hormone (hGH) and cytokine receptors was originally proposed to occur via homo- or heterodimerization (Weiss and Schlessinger, 1998). However, subsequent molecular and structural studies have extended this model to include the formation of higher order receptor oligomers as well as ligand-induced conformational changes in preformed receptor dimers and oligomers (Brown *et al.*, 2005; de Vos *et al.*, 1992; Livnah *et al.*, 1999; Figure 1). Receptor oligomerization or ligand-induced preformed dimer conformational change is thought to initiate activation of kinases (such as the Janus kinase JAK2 in the case of archetypical human growth hormone receptor (hGHR)), that are normally bound to the cytoplasmic tails of the resting receptor chains, by bringing the kinase molecules close enough to each other thus allowing transphosphorylation of the receptor cytoplasmic tails on specific tyrosine residues (Figure 1). This in turn creates docking sites for signaling and adaptor molecules such as STAT (signal transducer and activator of transcription) and the Src homology-2 containing protein Shc respectively, leading to distinct functions and the further building and assembling of an active signalosome.

Recent years have seen a steady stream of new cytokine receptor crystal structures including those that are activated by GM-CSF, Type I interferon, and a variety of interleukins. Here we give a brief overview of the cytokine family of receptors with a particular focus on how structural biology has shed light on their mechanism of signaling with implications for the development of new drugs to treat a variety of inflammatory diseases and cancer.

hGHR and Related Receptors

hGHR is the paradigm for understanding cytokine recognition and signaling of much of the cytokine superfamily (Figures 1 and 2). The hGHR extracellular region consists of the minimal structural elements of a type I cytokine receptor comprising a tandem set of FnIII domains (a single CBM), a single transmembrane (TM) domain, and a cytoplasmic tail (Figure 1). Two possible cytokine-binding pathways leading to receptor activation and downstream signaling have been put forward. The first involves binding of the four α -helical bundle hormone to the elbow between the two FnIII domains of one receptor chain, called Site 1, to form a binary complex (Figure 1, far left). The hormone in the binary complex is then recognized by a second receptor chain at its elbow, via a Site 2 interface. The two receptors come together at the membrane-proximal region resulting in a Site 3 interaction. The second pathway involves the hormone binding directly to a preformed receptor dimer (Figure 1, far right). A striking feature of this recognition process is how a single ligand can be recognized in two different ways by the same receptor chain.

For hGHR and other related cytokine receptors such as the erythropoietin (EPO), prolactin (PRL), and thrombopoietin (TPO) receptors, the hormone-induced model has recently been superseded with the demonstration that these receptors exist largely as inactive dimers in the absence of ligand (Brown *et al.*, 2012; Hercus *et al.*, 2009; Hiwase *et al.*, 2010; Lopez *et al.*, 2010; Wang *et al.*, 2009; Wells and de Vos, 1996). It was found that the TM domains of these single-pass cytokine receptors play an important role in their constitutive dimerization and the existence of inactive dimers implies that a specific ligand-induced conformational change is required for signal transmission to the associated cytoplasmic JAK2

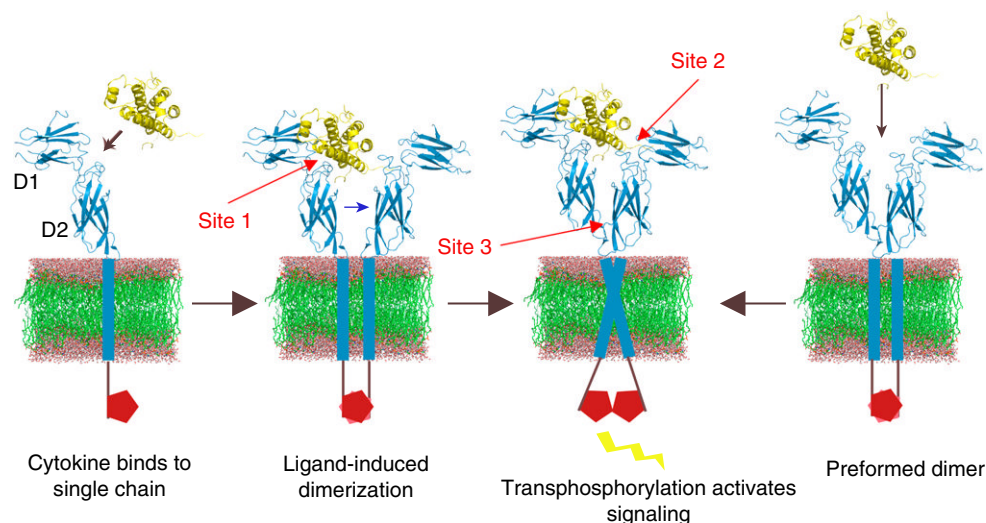
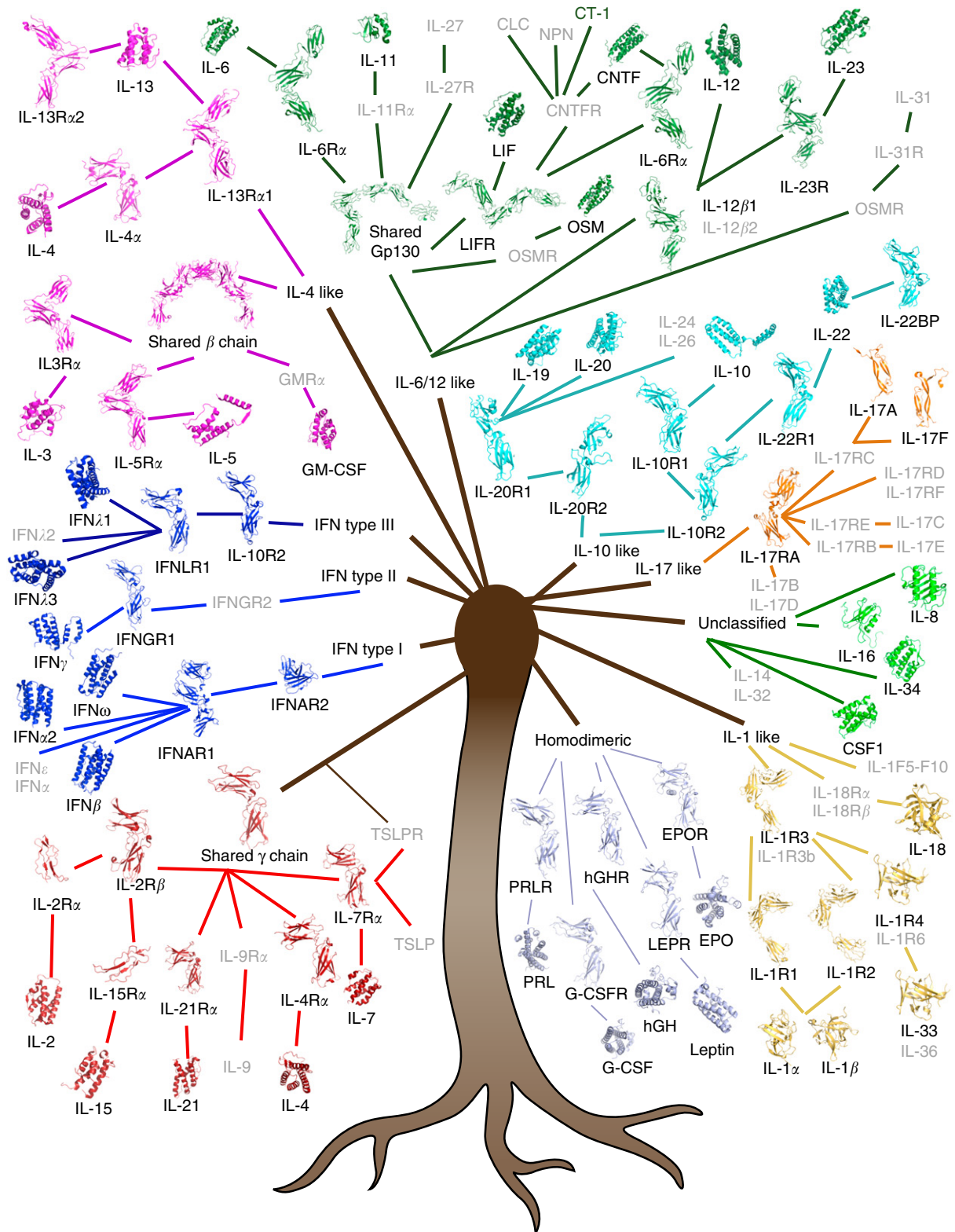


Figure 1 Human growth hormone (hGH) signaling as a paradigm for the cytokine receptor superfamily. A cartoon representation of hGH receptors (hGHR, in blue) shown bound to a lipid cell membrane. Domains D1 and D2 of hGHR define the cytokine-binding motif (CBM). Transmembrane (TM) domains are represented as blue rectangles. Red arrows indicate binding sites defined in the text. On the left hand side of the panel hGH (yellow colored cartoon) is shown binding first to a single receptor subunit (Site 1), which then induces receptor dimerization (Sites 2 and 3). The signaling complex is then activated via intracellular transphosphorylation of associated JAK molecules (represented as red hexagons) and cytoplasmic tails (brown sticks). From the right hand side of the panel moving left, a preformed symmetrical hGHR dimer is shown, which then undergoes conformational change upon ligand binding to activate the asymmetric complex. (PDB ID: 3HHR).

2014). This mechanism leads to removal of the pseudokinase domain from the kinase domain of the partner JAK2 and pairing of the two kinase domains, facilitating trans-activation (Figure 1).



The simple homodimeric receptor system seen in hGHR has evolved to much more complicated heteromeric systems, often involving shared receptors, but clear structural analogies in terms of conserved architectures and rules of receptor recognition can still be traced back to hGHR in spite of low (<20%) sequence identities between cytokines and between receptor chains in the same cytokine family (Figure 2). Such observations highlight the critical roles that crystallographic structural studies have played in unmasking unexpected similarities between proteins with highly divergent primary structures and disparate biological functions.

Beta Common Receptors

The β common (βc) family of cytokines consists of granulocyte-macrophage colony-stimulating factor (GM-CSF) and interleukins IL-3 and IL-5 (Figure 2). They are produced by activated T cells (and in the case of GM-CSF also by many other cell types) to regulate the survival, proliferation, differentiation, and functional activation of hematopoietic cells after infection or blood cell loss (Codarri *et al.*, 2011). The targets of GM-CSF are myeloid hemopoietic progenitor cells and mature cells such as neutrophils and macrophages (Burmester *et al.*, 2014; Hercus *et al.*, 2013). IL-3 is important for the production and function of hemopoietic stem cells, mast cells, and basophils (Dahl *et al.*, 2004; Lantz *et al.*, 1998), and IL-5 selectively stimulates the production and activation of eosinophils (Lopez *et al.*, 1988).

Recent evidence suggests that this family of cytokines plays a key role in cancer. Both IL-3 and GM-CSF contribute to the development and metastasis of solid tumors (Korpelainen *et al.*, 1993; Pylayeva-Gupta *et al.*, 2012; Su *et al.*, 2014). Intriguingly, the α -subunit of the IL-3 receptor (IL-3R α) is overexpressed on the surface of leukemic stem cells and primary blasts of patients with acute myelogenous leukemia (AML) and this correlates with poor prognosis (Jin *et al.*, 2009; Jordan *et al.*, 2000; Testa *et al.*, 2002). This observation has led to several strategies that target the IL-3R α such as antibodies that block IL-3 function and mediate killing of AML cells overexpressing IL-3R α (Busfield *et al.*, 2014; Gill *et al.*, 2014; Jin *et al.*, 2009; Nievergall *et al.*, 2014) by immune mechanisms.

GM-CSF, IL-3 and IL-5 are glycoproteins with architectures similar to that of hGH (Figures 1 and 2) as they are largely composed of four short antiparallel α -helices (Hercus *et al.*, 2013). While there are minor structural differences between GM-CSF and IL-3 (Feng *et al.*, 1996; Rozwarski *et al.*, 1996), IL-5 is distinct as it exists as a homodimer (Milburn *et al.*, 1993).

GM-CSF, IL-3, and IL-5 bind heterodimeric receptors composed of a type I cytokine-specific α -subunit and a common β -subunit (βc), which is shared by all three receptors and is the major signaling subunit. Unlike hGHR, the receptor α -subunits contain three FnIII domains: the N-terminal domain (D1), D2, and the membrane-proximal domain D3; D2-D3 comprise the CBM. The three FnIII domains wrap around the cytokine, as revealed by the crystal structure of the IL-5R α binary complex (Kusano *et al.*, 2012; Patino *et al.*, 2011). Apart from the IL-5R α binary complex structure, there is a partial crystal structure of the GMR α ternary complex (Hansen *et al.*, 2008), and a complete structure of IL-3R α bound to a blocking antibody (Broughton *et al.*, 2014). The crystal structure of βc (Figure 2) showed that it is an intertwined dimer that forms an arch spanning from the TM domain of one βc subunit to the TM domain of the second βc subunit (Carr *et al.*, 2001). Each of the βc monomers consists of four FnIII domains, designated D1 (N-terminal domain) through to D4 (membrane-proximal domain).

The GM-CSF ternary complex crystal structure revealed how the receptor is activated upon cytokine binding (Hansen *et al.*, 2008). Initially, GM-CSF binds to the elbow region between D2 and D3 of GMR α (analogous to Site 1 in Figure 1) to form a low-affinity binary complex ($K_d \sim 10$ nM) (Hercus *et al.*, 1994). Then, βc is recruited to the binary complex through interactions with GM-CSF and D3 of GMR α , analogous to Sites 2 and 3 respectively in Figure 1. Site 2 in βc uses a chimeric interaction surface composed of residues from D1 of one monomer and from D4 of the second βc monomer (Hansen *et al.*, 2008). This arrangement is unique in the type I cytokine receptor superfamily. The resulting ternary complex includes a βc homodimer, two GM-CSF molecules, and two α -subunits and converts the binding of cytokine to high affinity ($K_d = 100$ pM) (Hansen *et al.*, 2008).

Interestingly, the crystal lattice of the GM-CSF ternary receptor complex revealed a unique higher order dodecameric complex composed of two hexamers associating through the membrane-proximal D3 of GMR α and D4 of βc of each hexamer (Site 4). Formation of the dodecamer brings the intracellular βc domains into close proximity (~ 10 Å), which allows for JAK2 transphosphorylation and receptor activation (Broxmeyer *et al.*, 2012; Hansen *et al.*, 2008). Activated JAK2 phosphorylates tyrosine residues on βc (Guthridge *et al.*, 2004), which then act as docking sites for proteins that lead to the activation of JAK2/STAT5, Ras/Raf/MAPK, and PI3-kinase/Akt pathways (Ramshaw *et al.*, 2007). The IL-3 and IL-5 receptors are also predicted to form higher order complexes similar to that identified for the GM-CSF receptors (Dey *et al.*, 2009; Patino *et al.*, 2011; Zaks-Zilberman *et al.*, 2008).

Figure 2 Family tree of human cytokines and their receptors. Cytokines and their receptors are represented as cartoons and colored according to common features and structural motifs. The associated activating cytokines with known three-dimensional structures are shown with their respective receptor. For clarity, entire assembled receptors or partial structures are not shown. Some cytokines may be commonly assigned to more than one group. IFNLR1, IFN λ 1, IFN λ 2, and IFN λ 3 are also known as IL-28R α , IL-28 α , IL-28 β , and IL-29, respectively. The cytokines IL-36 α , IL-36 β , and IL-36 γ have not been included under the IL-1-like branch due to space limits. Similarly, thrombopoietin (TPO) and its receptor have not been included under the homodimeric branch due to space restrictions. There are currently no three-dimensional structures for the IL-36 α , IL-36 β , IL-36 γ and TPO cytokines or the TPO receptor. CT-1, cardiotrophin-1; CLC, cardiotrophin-like cytokine; CNTF, ciliary neurotrophic factor; EPO, erythropoietin; gp130, glycoprotein 130; G-CSF, granulocyte colony-stimulating factor; GM-CSF, granulocyte-macrophage colony-stimulating factor; GH, growth hormone; IFN, interferon; IL, interleukin; PRL, prolactin; LIF, leukemia inhibitory factor; OSM, oncostatin-M; TLSP, thymic stromal lymphopoietin (PDB IDs are: 2B51, 4GS7, 3BPO, 3BPL, 3DI2, 3TGX, 3SE4, 1FG9, 1AU1, 1RH2, 3S9D, 2GYS, 3CXE, 3QT2, 4JZJ, 1JLI, 3LB6, 3L5H, 1ALU, 111R, 1N26, 2KUM, 1EVS, 1EMR, 3EOG, 1F45, 3D87, 4DOH, 1N1F, 1J7V, 3LQM, 3DLQ, 3OG4, 4HSA, 4JPY, 1ICW, 1I16, 4DKC, 3UF2, 4DEP, 2L5X, 2VXT, 4KC3, 3G9V, 1EER, 3D48, 1CD9, 3HHR, 3HHH, 2Z3Q, 1P9M, 1X6D, 1AX8, 3O4O, 3V6O, and 1CNT).

Gamma Common Receptors

The gamma common (γ c) family of cytokines includes interleukins IL-2, IL-4, IL-7, IL-9, IL-15, IL-21, and thymic stromal lymphopoietin (TSLP) (Wang *et al.*, 2009). These cytokines regulate the development, proliferation, and homeostasis of B, T, and natural killer (NK) cells of the immune system (Leonard *et al.*, 1994a; Rochman *et al.*, 2009). The biological importance of the γ c receptor is evidenced by the fact that mutations in this receptor or its associated JAKs abolish functions of all γ c-dependent cytokines and causes X-linked severe combined immunodeficiency diseases (SCID) (Leonard *et al.*, 1994b). The γ c-dependent cytokines belong to the type I cytokine superfamily and they are functional monomers that adopt the classical left-handed four-helix bundle structure (Figures 1 and 2). Crystal structures of many γ c-dependent cytokine:receptor complexes have been solved and they are discussed below.

IL-2/IL-15 Receptors

IL-2 and IL-15 are key regulators of lymphocyte homeostasis and function. IL-2 is an autocrine/paracrine factor produced primarily by antigen-activated T cells and it was initially pursued as a therapeutic agent for patients with metastatic renal cell carcinoma and metastatic melanoma; however, severe toxicities associated with high-dose IL-2-based therapies have limited its clinical use (Levin *et al.*, 2012; Schwartzentruber *et al.*, 2011). IL-15 is produced by activated monocytes and dendritic cells and has been implicated in inflammatory diseases such as rheumatoid arthritis (Mattei *et al.*, 2001; Nelson and Willerford, 1998; Votavova *et al.*, 2014; Waldmann and Tagaya, 1999).

Both IL-2 and IL-15 signal through heterotrimeric 'high affinity' complexes consisting of a selective IL-2R α or IL-15R α receptor and two common receptors, IL-2R β and γ c (Chirifu *et al.*, 2007; Kono *et al.*, 1993; Nakamura *et al.*, 1994; Olsen *et al.*, 2007; Rickert *et al.*, 2005; Stauber *et al.*, 2006). The IL-2R α /15R α receptors are comprised of two sushi domains and upon binding to their respective cytokines form a stable binary complex that lacks signaling activity. This binary complex is then recruited by the signaling subunits IL-2R β and γ c to form an active signaling complex. Crystal structures of the quaternary complexes of both IL-2 and IL-15 have been determined to reveal a striking similarity in the membrane-proximal or 'signaling ends' of the cytokine receptors (Ring *et al.*, 2012). The shared signaling subunits, IL-2R β and γ c, have the same domain architecture as hGHR and the cytokine binds in a similar fashion as hGH (Figure 1). Since the assembly of the two cytokine heterotrimeric receptors is almost identical, differences in signaling is attributed to: distinct expression patterns of the IL-2R α /15R α receptors; differences in receptor binding kinetics; the presentation mode of the cytokine, i.e., *cis* versus *trans*, by the IL-2R α /15R α receptors to the common receptor subunits and the presence of additional signals such as those from the T-cell antigen receptor (Rickert *et al.*, 2005; Ring *et al.*, 2012; Stauber *et al.*, 2006).

IL-4/IL-13 Receptors

While IL-13 is not a γ c-dependent cytokine, it does share the IL-4 specific receptor, IL-4R α , and hence it will also be discussed here.

Like the γ c-dependent cytokines, IL-13 adopts the classical left-handed four-helix bundle structure observed for hGH (Figures 1 and 2) (Lupardus *et al.*, 2010). The cytokines IL-4 and IL-13 are produced by activated T helper type 2 (TH2) cells, mast cells and basophils (Chomarat and Banchereau, 1998; Nelms *et al.*, 1999; Wynn, 2003). Deregulation of the IL-4 and IL-13-mediated responses has been implicated in the pathology of asthma and allergic diseases (Barnes, 2002; Wills-Karp, 2004).

IL-4 binds to the IL-4R α receptor with high affinity (forming a binary complex), which then leads to receptor dimerization with either the γ c chain to form a type I receptor complex (in hematopoietic cells) or with the IL-13R α 1 receptor to form a type II receptor complex (in non-hematopoietic cells) (Arima *et al.*, 2005; Kelly-Welch *et al.*, 2003; Mueller *et al.*, 2002). The IL-4R α receptor shares the same domain architecture as hGHR (Figure 1), whereas IL-13R α 1 contains an additional membrane-distal FnIII domain (D1) set at $\sim 100^\circ$ to the middle FnIII domain (D2) so that the extracellular region adopts an overall zig-zag conformation, analogous to the domain architecture of IL-5R α (Figure 2; LaPorte *et al.*, 2008). IL-4 interacts with IL-4R α and γ c in an analogous fashion to hGH (Figure 1); whereas in the IL-4: IL-4R α :IL-13R α 1 ternary complex the N-terminal IL-13R α 1 FnIII domain (D1) folds down over the cytokine in a manner that has been likened to a wrench (IL-13R α 1) fitting around a nut (cytokine) (Patino *et al.*, 2011). IL-13 can form an analogous type II receptor complex by binding to both IL-13R α 1 and IL-4R α . Differential responses to IL-4 and IL-13 cytokines are achieved by 'tuning' the respective affinities of the cytokines for the IL-13R α 1 receptor (Andrews *et al.*, 2009; Ito *et al.*, 2009).

IL-13R α 2 is a high affinity receptor of IL-13 which serves as a decoy binding protein due to its short intracellular domains that lack JAK interaction motifs (Andrews *et al.*, 2009; Wood *et al.*, 2003). The crystal structure of the IL-13:IL-13R α 2 binary complex revealed that the receptor adopts a very similar fold to IL-13R α 1, but utilizes peripheral residues unused in the IL-13: IL-13R α 1 complex to generate a larger complementary interface for IL-13 binding (Lupardus *et al.*, 2010). Interestingly, IL-13R α 2 has been reported to be significantly up-regulated in a number of human cancers (Kioi *et al.*, 2006; Mintz *et al.*, 2002) making it a potential target for tumor-specific therapeutics (Kahlon *et al.*, 2004).

IL-7/TSLP Receptors

IL-7 and its receptors play an important role in the development and maintenance of B and T cells and extracellular matrix remodeling. Aberrant IL-7 signaling has been shown to lead to diseases such as cancer, heart disease, and allergic rhinitis upon overstimulation and SCID when understimulated (Kovanen and Leonard, 2004; Lundström *et al.*, 2012; Rochman *et al.*, 2009). IL-7 receptors signal via the JAK/STAT, PI3-kinase/Akt, and Src pathways (Mazzucchelli and Durum, 2007).

Like the IL-4R α , IL-2R β , and γ c receptors, IL-7R α has the same domain architecture as hGHR (Figure 1; McElroy *et al.*, 2012). In the crystal structure of the IL-7:IL-7R α binary complex, the interactions between the two proteins are generally more hydrophobic and nonspecific than those observed in

other γ c-dependent cytokine receptor complexes (McElroy *et al.*, 2012). Surface plasmon resonance studies of the IL-7:IL-7R α binary complex revealed that N-glycosylation of the receptor chain is critically important for high affinity binding, in stark contrast to the other members of the γ c receptor family (McElroy *et al.*, 2009; Walsh, 2012; Wickham and Walsh, 2007). It has been speculated that glycosylation may influence the engagement of IL-7 with its receptor by affecting the overall electrostatic potential of the receptor chain or altering the conformational states of the receptor (McElroy *et al.*, 2009). IL-7R α can form homodimers at the cell surface, with the C-termini of the extracellular domains separated by ~ 110 Å, but they do not activate downstream signaling (McElroy *et al.*, 2012). In addition to interacting with IL-7, IL-7R α also binds to TSLP and the TSLP receptor (TSLPR) (Wang *et al.*, 2013; Ziegler and Artis, 2010). TSLP is expressed by epithelial cells in the skin, lung, thymus, and intestine and plays a key role in intestinal immune homeostasis and in mediating allergic inflammatory processes, including those related to asthma and atopic dermatitis (Roan *et al.*, 2012; Zhang and Zhou, 2012).

Shared Glycoprotein 130 Receptors

The IL-6 cytokine family comprises the archetypal family member IL-6, along with IL-11, IL-27, IL-31, leukemia inhibitory factor (LIF), oncostatin-M (OSM), ciliary neurotrophic factor (CNTF), cardiotrophin-1 (CT-1), cardiotrophin-like cytokine (CLC), and neuropoietin (NPN). A unifying feature of these cytokines is their similar tertiary structure, which despite a lack of sequence homology, is made up of a conserved four α -helical bundle analogous to hGH (Figures 1 and 2; Bazan, 1990b). Another hallmark of this family of cytokines, with the exception of IL-31 (Dillon *et al.*, 2004), is their usage of the ubiquitously expressed shared signal-transducing receptor β -subunit, glycoprotein 130 (gp130) (Heinrich *et al.*, 2003), to drive their collective biological activities (e.g. cell proliferation, survival, functional maturation, and differentiation).

The gp130 signaling subunit belongs to the type I 'tall' cytokine receptor family; the 'tall' cytokine receptor family contains additional IL-6 family cytokine-specific β -receptors for LIF (LIFR), OSM (OSMR), IL-27 (IL-27R/WSX-1), and IL-31 (IL-31RA, also known as gp130-like protein (GPL)), as well as the β -receptors for IL-12 (IL-12R β 2), GM-CSF (β c), and leptin (LEPR) (Skiniotis *et al.*, 2005). IL-6 and IL-11 signal via gp130 homodimers whereas all other family members employ heterodimers of gp130 with their respective β -receptor, for example, LIFR, OSMR, and IL-27R (Figure 2; Heinrich *et al.*, 2003). The exception to this rule is IL-31, which utilizes a heterodimer of IL-31RA and OSMR (Dillon *et al.*, 2004). The extracellular region of these receptors contains numerous immunoglobulin (Ig)-like and FnIII-like domains; for example, gp130 consists of six structural domains (D1 through to D6) classified as N-terminal Ig-like domain (D1), two adjacent FnIII domains within the CBM (D2-D3) and three membrane-proximal FnIII domains (D4-D6) (Bazan, 1990a). Despite the lack of intrinsic tyrosine kinase activity, the membrane-proximal region of the cytoplasmic domain of gp130 and its related

cytokine-specific β -subunits (e.g. LIFR, OSMR) also contain a conserved proline-rich motif ('box 1' motif), which via mutagenesis has been demonstrated to be critical for the constitutive association and ligand-induced activation of members of the JAK family of non-receptor tyrosine kinases (JAK1, JAK2, and TYK2) (Lupardus *et al.*, 2011; Murakami *et al.*, 1991).

Among the IL-6 cytokine family, assembly of gp130-containing receptor complexes is best characterized for IL-6. IL-6, and in an analogous manner IL-11, first bind with high affinity to its cognate α -subunit (IL-6R α , IL-11R α), itself comprising an extracellular Ig-like N-terminal domain (D1) adjacent to a CBM (D2-D3), a TM domain, and a short cytoplasmic domain that does not have any signaling potential (Yawata *et al.*, 1993). In addition to this conventional mode of receptor assembly, it is worth noting two other mechanisms employed by IL-6 family cytokines, both of which involve IL-6R α , to promote their vast pleiotropic and/or redundant biological properties. Firstly, in a process called 'trans-signaling,' a naturally occurring soluble form of IL-6R α , which upon its production via either protease-induced shedding or alternative mRNA splicing, can bind IL-6 to facilitate association with gp130, thus providing IL-6 with an avenue to activate gp130 signaling in a vast array of cell types that do not normally express the membrane-bound IL-6R α (Chalaris *et al.*, 2011). Secondly, IL-6R α is also used by CNTF to induce LIFR: gp130 heterodimers (Schuster *et al.*, 2003), which suggests that in certain situations cells can become responsive to CNTF even in the absence of the CNTF ligand-binding α -subunit CNTFR. Despite these variations, a common outcome during the ligand-induced assembly of gp130-containing receptor complexes is the activation of the receptor-associated JAKs, which tyrosine phosphorylate gp130 (and the associated cytokine-specific β -subunits) to facilitate the subsequent activation of downstream signaling molecules, such as STAT, extracellular signal regulated kinase (ERK), MAPK, and PI3-kinase (Heinrich *et al.*, 2003).

Insights into the mechanisms by which ligand-bound, gp130-containing receptor complexes are formed have evolved greatly over the last couple of decades, from earlier molecular modeling and site-directed mutagenesis studies to more recent and advanced experimental approaches combining NMR spectroscopy, low resolution electron microscopy and crystallography to generate high resolution crystal structures (Boulanger *et al.*, 2003a; Chow *et al.*, 2001; Horsten *et al.*, 1997; Kalai *et al.*, 1997; Kernebeck *et al.*, 1999; Lupardus *et al.*, 2011; Schwantner *et al.*, 2004; Xu *et al.*, 2010; Yawata *et al.*, 1993). Importantly, the crystal structure for the IL-6-bound receptor complex revealed a hexameric IL-6:IL-6R α :gp130 complex with 2:2:2 stoichiometry, thus providing formal evidence for the presence of gp130 dimers (Boulanger *et al.*, 2003b). Furthermore, comparative analyses of the structure of the extracellular region of gp130 in a ligand-bound versus ligand-free state suggested that ligand binding induces only subtle structural changes in gp130 that are required to activate intracellular signaling cascades (Xu *et al.*, 2010). Collectively, the detailed structural information obtained on individual receptor subunits has further defined critical amino acids for subunit-subunit and ligand-subunit interactions during receptor assembly, which is important for the design of small peptide and chemical inhibitory molecules for potential use as therapeutics to combat the myriad of inflammatory and autoimmune diseases, as well as cancers, in

which the IL-6 family cytokines are implicated (Mansell and Jenkins, 2013).

IL-1 Receptors

IL-1 was the first cytokine identified (Dinarello, 2013; Dinarello *et al.*, 1977) and is closely linked to innate inflammatory and immune responses. The family of IL-1 ligands has now expanded to include IL-1 α , IL-1 β , IL-18, IL-33, IL-36 α , IL-36 β , and IL-36 γ , which are related through receptor recognition and structure (Figure 2), and the signal transduction pathways emanating from their respective receptors (Boraschi and Tagliabue, 2013; Garlanda *et al.*, 2013). IL-1 family members play a crucial role in maintaining tissue integrity, maturing the adaptive immune response, and maintaining homeostasis (Arend *et al.*, 2008), largely due to receptor-mediated activation of the key transcription factor, NF- κ B. Here we will focus on the prototypic pyrogenic cytokines IL-1 α and IL-1 β and their receptor complexes.

A characteristic feature of the IL-1 receptor (IL-1R) family is the cytosolic Toll/IL-1 receptor domain (TIR, originally termed Toll/IL-1 resistance) (Gay and Keith, 1991), common to both IL-1 family receptors and Toll-like receptors (TLR). A key structural difference, however, is the extracellular ligand-binding domain; IL-1Rs contain Ig domains, while TLRs consist of leucine-rich repeats (LRRs) that define ligand specificity. Ligand-induced receptor dimerization provokes homotypic dimerization of the cytosolic TIR domains, causing recruitment of the protein encoded by myeloid differentiation primary response gene 88 (MyD88) to the receptor complex. Recruitment of MyD88 initiates signal transduction via activation of IL-1R-associated kinase (IRAK)-1, IRAK-4, and tumor necrosis factor receptor-associated factor (TRAF)-6, leading to activation of the I κ B kinase signalosome, which instigates NF- κ B nuclear translocation and gene transcription. Activation of NF- κ B induces inflammation and amplification of the immune response (Pahl, 1999), which requires tight regulation, as excessive activation of NF- κ B by TIR-mediated signaling is now known to be associated with tissue and organ damage and is an underlying pathology in multiple diseases (Dinarello *et al.*, 2012; O'Neill *et al.*, 2009). As such, there is considerable interest in understanding the IL-1R complex with an aim to therapeutically targeting it to alleviate inflammation.

IL-1R1

The IL-1 receptor type 1 (IL-1R1) is the high affinity membrane-bound receptor that binds IL-1 α and IL-1 β . This was the first identified IL-1 receptor and it comprises an extracellular domain of three Ig-like domains and a cytosolic TIR domain. Resolving the IL-1R1 crystal structure to 2.5 Å resolution demonstrated that the two N-terminal Ig domains (D1 and D2) of IL-1R1 are configured in a rigid structure due to an interdomain disulfide bond, while the D3 membrane-proximal Ig domain is flexible (Figure 2; Vigers *et al.*, 1997). Binding of IL-1 β demonstrates that IL-1R1 wraps around the cytokine, contacting via two surface patches; a predominate patch in the groove between D1 and D2 and a smaller region on the side of D3. It is presumed, though not defined as yet,

that IL-1 α binds in a similar fashion due to homology and binding affinity (approximately 0.1–1.0 nM for both interleukins) (Symons *et al.*, 1995). While IL-1R1 binds its antagonistic ligand IL-1 α with similar high affinity as IL-1 α / β , the binding interface is different. While IL-1 α binds the large patch on IL-1R1 D1 and D2, it fails to bind the smaller patch on D3, resulting in only a partial wrapping of the receptor around IL-1 α , and the receptor structure being more open and extended than the IL-1 β :IL-1R1 complex (Schreuder *et al.*, 1997). The arrangement thus fails to recruit IL-1R3, whose involvement is described below, and the complex cannot induce signal transduction.

IL-1R1 can also undergo metalloproteinase-mediated shedding of the extracellular Ig domains (Elzinga *et al.*, 2009; Twomey *et al.*, 2009), producing a soluble receptor (sIL-1R1) that can act as a decoy receptor to prevent excessive IL-1-induced membrane-proximal signaling. However, sIL-1R1 can also sequester IL-1 α , potentially offsetting the inhibitory capacity of IL-1 α (Burger *et al.*, 1995; Symons *et al.*, 1995). Genetic deletion or ablation of IL-1 α causes severe systemic and local inflammation, further highlighting its role in balancing IL-1R-mediated inflammation (Masters *et al.*, 2009).

IL-1R2

Similar to IL-1R1, IL-1R2 consists of three extracellular Ig-like domains and is able to bind IL-1; however, because it lacks a cytoplasmic TIR domain, it acts as a decoy receptor (Colotta *et al.*, 1994; McMahan *et al.*, 1991). Following IL-1 binding, IL-1R2 can recruit IL-1R3; however, as IL-1R2 does not contain a TIR domain, a nonfunctional receptor complex is formed that cannot recruit MyD88.

Interestingly, while IL-1R2 can bind IL-1 β at similar affinity as IL-1R1 (2–5 nM) and has a five- to ten-fold lower affinity for IL-1 α than IL-1R1, it recognizes IL-1 α poorly with ~1000-fold lower affinity (Symons *et al.*, 1995). IL-1R2 can also be cleaved proteolytically from the membrane surface generating either a 60 kDa or 45 kDa soluble fragment (sIL-1R2) that can bind IL-1 (Lorenzen *et al.*, 2012; Orlando *et al.*, 1997). sIL-1R2 is able to bind pro-IL-1 β (Symons *et al.*, 1995), the inactive IL-1 β precursor, indicating that sIL-1R2 appears to recognize the presence of pro-IL-1 β from necrotic cells. sIL-1R2 is also able to engage soluble IL-1R3 (see below), which increases its binding affinity for IL-1 about 100-fold, a mechanism that appears to represent a further way of regulating IL-1 availability and its downstream effects on the inflammatory response (Smith *et al.*, 2003). An intracellular form of IL-1R2 has also been described to bind cytosolic IL-1 α , which protects it from maturation by proteolytic processing.

IL-1R3

IL-1R3 is the most promiscuous of the IL-1R family, not only forming a high affinity complex with IL-1R1 and IL-1R2, but also acting as an accessory signaling receptor with IL-1R4 and IL-1R6 to form agonist receptor complexes for IL-33 and IL-36 respectively (reviewed in Boraschi and Tagliabue, 2013; Garlanda *et al.*, 2013). Consistent with IL-1R1, IL-1R3 consists of three Ig-like extracellular domains linked to the cytosolic

TIR domain via a single-pass TM domain. While IL-1R3 does not bind IL-1 directly, it is necessary to form an active signaling complex with IL-1R1 (hence its previous name IL-1R Accessory Protein). An IL-1R1 interaction with IL-1R3 via the extracellular Ig-like domains causes cytosolic TIR:TIR association to initiate active downstream signaling to NF- κ B.

In silico and crystal structure elucidation of IL-1R3 in complex with IL-1 β and IL-1R1 (and IL-1R2) has demonstrated that IL-1R3 engages IL-1R1 on the reverse side of ligand binding, interacting with D2 and D3 of IL-1R1, but having little to no contact with IL-1 itself (Boraschi *et al.*, 1995; Casadio *et al.*, 2001). It was also demonstrated that IL-1R3 engagement with IL-1 α :IL-1R1 complexes only facilitated interaction with IL-1R1 D2, resulting in the formation of an unstable and hence inactive, receptor complex, which supports earlier experimental findings (Wang *et al.*, 2010). As with other IL-1R family members, alternative splicing produces a soluble IL-1R3 (sIL-1R3), which has been hypothesized to act as a further decoy receptor to negatively regulate IL-1 function (Greenfeder *et al.*, 1995). There are four other alternatively spliced variants of IL-1R3 which are produced under a variety of circumstances and expressed in specific tissues; these lack TM and intracellular domains (reviewed in Boraschi and Tagliabue, 2013).

IL-1R3b

An alternate transcript of IL-1R3, termed IL-1R3b, produces a variant of IL-1R3 that is 687 amino acid longer at its C-terminal domain (Lu *et al.*, 2008). IL-1R3b is exclusively expressed in the central nervous system (Smith *et al.*, 2009), in particular neurons, but absent in astrocytes (Huang *et al.*, 2011; Nguyen *et al.*, 2011; Smith *et al.*, 2009). Gene deficiency of IL-1R3b appears to confer a local neuroprotective role during cerebral inflammation, but no role in peripheral responses, as opposed to IL-1R3 (Gosselin *et al.*, 2013; Smith *et al.*, 2009). These findings seem to suggest that the same receptor may be used in different tissues or organelles to achieve specific functions that may further highlight the pleiotropic effects of IL-1 family members.

Therapeutic Targeting of IL-1R

Clinical treatment of patients suffering IL-1 driven disease cryopyrin-associated periodic syndrome (CAPS) with IL-1R directed therapies has highlighted the potential of targeting the IL-1R complex (Dinarello *et al.*, 2012). A recombinant form of IL-1 α called Anakinra, approved in 2001, has proved effective in treating a broad range of diseases such as autoimmune hearing loss, hidradenitis suppurativa, stroke, and arthritis. Canakinumab is a human monoclonal antibody that targets IL-1 β , disrupting its ability to bind IL-1R1. It has been successful in the treatment of CAPS patients along with refractory gout and systemic onset juvenile idiopathic arthritis. Canakinumab is in several clinical trials currently, including the CANTOS study (clinical trial number NCT01327846), to examine if reducing IL-1R activation decreases rates of recurrent myocardial infarction, stroke, and cardiovascular death among stable patients with coronary artery disease.

IL-10/IL-22 Receptors

The nine known IL-10 cytokine family members (interleukins IL-10, IL-19, IL-20, IL-22, IL-24, and IL-26; and the interferon (IFN)- λ members IFN λ 1 (IL-28 α), IFN λ 2 (IL-28 β), and IFN λ 3 (IL-29)) belong to the type II cytokines, with IL-10 described as the master anti-inflammatory cytokine. IL-10 is expressed by numerous immune cells including dendritic cells, macrophages, T cells, and eosinophils; whereas IL-22 is produced by T cells, innate lymphoid cells, and epithelial cells (Rutz *et al.*, 2013; Sabat, 2010; Sabat *et al.*, 2014; Striz *et al.*, 2014; Tsai *et al.*, 2013). Apart from IL-22, the IL-10 cytokine family directly regulates the function of immune cells and maintains the balance between overexuberant and beneficial innate immunity responses. IL-22 targets cells of the pancreas, liver, kidney, and joints; as well as cells at outer-body barriers like the skin (Kong *et al.*, 2013; Mizoguchi, 2012; Pan *et al.*, 2013; Rutz *et al.*, 2013; Sabat *et al.*, 2014). Viruses such as human immunodeficiency virus, dengue virus and Epstein-Barr virus can evade immune responses by increasing host IL-10 production or by expressing their own virally encoded IL-10 homologs (Ouyang *et al.*, 2014; Slobedman *et al.*, 2009). Recent evidence suggests that IL-10 cytokines may have potential clinical relevance in stroke, cardio-protection, pancreatitis, chronic kidney disease, inflammatory bowel disease, graft-versus-host disease, cancer, wound healing and immune protection against bacterial, nematode, fungal, viral, and protozoa infections (Bodhankar *et al.*, 2013; Hamidullah *et al.*, 2012; King *et al.*, 2014; Kong *et al.*, 2013; Leung and Loke, 2013; Mizoguchi, 2012; Pan *et al.*, 2013; Rutz *et al.*, 2013; Sabat, 2010; Sabat *et al.*, 2014; Wang *et al.*, 2013; Zdanov, 2010).

The IL-10 cytokines consist of five to seven α -helices with the signature left-handed four-helix bundle typical of helical cytokines (Figures 1 and 2); their pairwise sequence identity ranges from 9% to 40% (Bleicher *et al.*, 2008; de Moura *et al.*, 2009; Jones *et al.*, 2008; Josephson *et al.*, 2001; Nagem *et al.*, 2002; Sabat, 2010; Yoon *et al.*, 2010). IL-10 exists as a domain-swapped dimer; IL-19, IL-20, IL-22, IL-24, IFN λ 1, IFN λ 2, and IFN λ 3 are functional monomers and IL-26 is active either as a dimer or monomer (Josephson *et al.*, 2001; Nagem *et al.*, 2002; Zdanov, 2010). Generally the cytokine first binds with high affinity to one receptor (for IL-10 and IL-22 it is the IL-10R1 and IL-22R1 receptor, respectively) and then the binary complex binds with lower affinity to a second common receptor, IL-10R2, leading to JAK1/TYK2/STAT3 activation (Jones *et al.*, 2008; Logsdon *et al.*, 2004; Sabat *et al.*, 2014; Yoon *et al.*, 2006; Zdanov, 2010).

Although no ternary receptor complexes have yet been determined, the binary complexes of IL-10:IL-10R1, IL-22:IL-22R1 and IL-22:IL-22BP are structurally very similar to each other and the isolated IL-10R2 structure is also available (Figure 2; Bleicher *et al.*, 2008; de Moura *et al.*, 2009; Jones *et al.*, 2008; Josephson *et al.*, 2001; Yoon *et al.*, 2006, 2010). The receptors have the same domain architecture as hGHR, with two FnIII domains, and the cytokine interacts in a similar fashion to hGH (Figure 1). IL-10R1 and IL-22R1 share 22% sequence identity. IL-22 binding protein (IL-22BP) is a naturally occurring soluble receptor containing only an extracellular region and it shares ~30% sequence identity with the

extracellular regions of both IL-22R1 and IL-10R2 (de Moura *et al.*, 2009). Homology models of the IL-10:IL-10R1:IL-10R2 and IL-22:IL-22R1:IL-10R2 ternary complexes (Bleicher *et al.*, 2008; Pletnev *et al.*, 2005; Yoon *et al.*, 2010), together with surface plasmon resonance measurements (Logsdon *et al.*, 2004; Yoon *et al.*, 2006, 2010) and alanine scanning mutagenesis studies (Logsdon *et al.*, 2004), suggest that IL-10R1/22R1 and IL-10R2 interact via their membrane-proximal D2 domains (Figure 1, Site 3) and that IL-10/22 interacts with both receptor chains through distinct binding sites (Figure 1, Sites 1 and 2).

IL-17 Receptors

There are six members in the pro-inflammatory IL-17 family of cytokines, designated IL-17A through to IL-17F (Figure 2), and they are important regulators of both innate and adaptive immune responses (Baeten and Kuchroo, 2013; Ely *et al.*, 2009; Kirkham *et al.*, 2013; Striz *et al.*, 2014; Zhang *et al.*, 2013). IL-17 cytokines are produced by numerous cell types including TH17 helper T cells, macrophages, dendritic cells, NK cells, innate lymphoid cells and epithelial cells (Baeten and Kuchroo, 2013; Kirkham *et al.*, 2013; Striz *et al.*, 2014; Zhang *et al.*, 2013). These cytokines are involved in host defense against bacteria and fungi and dysregulation of their production may contribute to inflammatory and autoimmune disorders such as multiple sclerosis, rheumatoid arthritis, psoriasis, uveitis, and ankylosing spondylitis (Baeten and Kuchroo, 2013; Blalock *et al.*, 2013; Kirkham *et al.*, 2013; Korn *et al.*, 2009; Striz *et al.*, 2014). The crystal structures of IL-17A and IL-17F reveal that they adopt a typical cysteine knot fold and assemble as disulfide-linked homodimers, in addition IL-17A and IL-17F form functional heterodimers (Ely *et al.*, 2009; Hymowitz *et al.*, 2001; Kirkham *et al.*, 2013; Liu *et al.*, 2013). Each IL-17A/F monomer is composed of two pairs of anti-parallel β -sheets. The other IL-17 cytokines have varying homology with IL-17A ranging from 16% to 50% (Kirkham *et al.*, 2013).

The five IL-17 receptor subtypes, IL-17RA through to IL-17RE, are single TM proteins with their extracellular region containing two atypical FnIII domains and their cytoplasmic tail containing a unique receptor signaling motif known as the 'similar expression to fibroblast-growth factor genes and IL-17R' (SEFIR) domain (Ely *et al.*, 2009; Gu *et al.*, 2013; Liu *et al.*, 2013; Zhang *et al.*, 2013). IL-17 receptor signaling requires the NF κ B activator 1 (Act1), which also contains a SEFIR domain. Upon cytokine stimulation Act1 is recruited to IL-17R complexes in a SEFIR-dependent manner (Ely *et al.*, 2009; Kirkham *et al.*, 2013; Zhang *et al.*, 2013). The crystal structure of the isolated IL-17RB SEFIR domain consists of six α -helices wrapped around a five-stranded parallel β -sheet core (Zhang *et al.*, 2013).

In the IL-17A dimer:IL-17RA and IL-17F dimer:IL-17RA complexes the two FnIII domains of IL-17RA (Figure 2) are arranged in a linear fashion, rather than set at right angles to each other as observed in the classical receptors (Figure 1), and the manner in which the cytokine interacts is quite distinct from other cytokine receptor systems (Ely *et al.*, 2009; Liu *et al.*, 2013). The cytokine binds edge on to the two FnIII

domains of IL-17RA in a groove between the IL-17 dimer interface, burying $\sim 2200 \text{ \AA}^2$ of the receptor surface area. The majority of the interaction ($\sim 70\%$) involves the membrane-distal D1 FnIII domain of IL-17RA. The cytokine undergoes conformational changes in peripheral strands and loops to facilitate receptor binding (Ely *et al.*, 2009; Hymowitz *et al.*, 2001; Liu *et al.*, 2013). IL-17RA appears to act as a common receptor for the IL-17 cytokine family analogous to those in type I cytokine receptor complexes (Ely *et al.*, 2009; Liu *et al.*, 2013). IL-17A, IL-17F homodimers and the heterodimeric IL-17A/IL-17F require both IL-17RA and IL-17RC for signaling (Gu *et al.*, 2013; Kirkham *et al.*, 2013; Wright *et al.*, 2008). IL-17RA and IL-17RE are thought to interact with IL-17C, while IL-17RA and IL-17RB interact with IL-17E (also known as IL-25) (Gu *et al.*, 2013; Kirkham *et al.*, 2013). Binding to the first receptor with high affinity (low nM) modulates the affinity and specificity of the second receptor binding event (typically $\sim 150\text{--}300 \text{ nM}$), promoting heterodimeric rather than homodimeric complex formation (Ely *et al.*, 2009; Hymowitz *et al.*, 2001; Liu *et al.*, 2013).

IFN Receptors

The IFNs are broadly distinguished from other cytokines by their ability to directly interfere with viral replication (Isaacs and Lindenmann, 1957). Interferons are classified into three types based on their amino acid sequence, cognate receptor system and to a lesser extent producing cell and stimulus. Type I IFNs (incl. α , β and ϵ) signal through IFN alpha receptor (IFNAR)-1 and IFNAR2, Type II IFN (γ) signals through IFN gamma receptor (IFNGR)-1 and IFNGR2; Type III IFNs (λ) signal through IFN lambda receptor (IFNLR)-1 and IL10R2 (de Weerd *et al.*, 2007). (Unfortunately, the 'Type' nomenclature used to distinguish interferons can lead to confusion when discussing type I/II cytokine receptor families. In this article we distinguish the two by using all lower case for the latter). All identified IFN receptors are type II helical cytokine receptors, thereby possessing certain structural traits (Langer *et al.*, 2004) in common with their type I counterparts, such as tandem FnIII domains in the extracellular domain, conserved patterns of disulfide bonding cysteines and a single TM region (Bazan, 1990a).

Type I IFN Receptors

The Type I IFNs are a large family of proteins consisting of ~ 20 different subtypes that engage a shared receptor complex (Figure 2), comprised of high affinity (IFNAR2) and low affinity (IFNAR1) receptor components to activate JAK/STAT, CRKL, MAPK and PI3K signaling pathways. While IFNAR2-extracellular domain (ECD) is comprised of tandem FnIII domains, typical of all other known type II cytokine receptors, IFNAR1-ECD possesses four domains (D1 through to D4) in the extracellular region. Structures for IFNAR2-ECD have been resolved by both NMR (in liganded and unliganded states) (Chill *et al.*, 2002) and crystallography (liganded state) (Thomas *et al.*, 2011), predictably revealing a receptor reminiscent of other type II cytokine receptors. Although the

structure for unliganded, full-length IFNAR1-ECD has not been determined, a structure for the three membrane-distal subdomains (D1-D3) of IFNAR1-ECD has been resolved, showing a relatively planar arrangement for these domains (Thomas *et al.*, 2011). This planar architecture appears to be maintained upon ligand binding since the crystal structure of the mIFN β :IFNAR1-ECD binary complex revealed a similar planar arrangement for the three membrane-distal subdomains of the receptor (de Weerd *et al.*, 2013). However, this structure revealed for the first time the location and presentation of the membrane-proximal subdomain (D4) of IFNAR1-ECD showing it to be transplanted away from the plane of the three membrane-distal domains (D1-D3) giving the receptor a squat appearance (de Weerd *et al.*, 2013). While NMR of IFNAR2-ECD revealed a rigid structure with only limited backbone conformational changes upon ligand engagement (Chill *et al.*, 2002), ligand binding to IFNAR1 induces a conformational change that must be propagated through the membrane-proximal D4 subdomain in order to induce signal transduction (Strunk *et al.*, 2008). The structures of the IFN α 2 variant and IFN ω ternary complexes that were recently elucidated confirmed the location of the IFNAR1 and IFNAR2 binding sites on opposing faces of the IFNs, and revealed considerable conservation in the modes of docking of the two IFN subtypes and residues involved in the interfaces with IFNAR2-ECD (Thomas *et al.*, 2011).

Type II IFN Receptors

IFN γ is the lone Type II IFN, and signals through high affinity (IFNGR1) and low affinity (IFNGR2) receptors (de Weerd and Nguyen, 2012). The structure of the functional IFN γ dimer in complex with its high affinity receptor IFNGR1 has been determined revealing that the cytokine homodimer binds to two molecules of the high affinity IFNGR1 on identical related surfaces of the cytokine homodimer (Walter *et al.*, 1995). The interaction between IFN γ and IFNGR1 resembles other type II cytokine receptor interactions, including those of tissue factor (Banner *et al.*, 1996) and IL-10R1 (Josephson *et al.*, 2001) with their ligands, but differs from those of the type I cytokine receptors in the relative presentation of the subdomains and the orientation of ligands (Walter *et al.*, 1995). Crystal structures for the low affinity receptor, IFNGR2 and the ternary IFN γ :IFNGR1:IFNGR2 complex have not yet been determined.

Type III IFN Receptors

Initially termed IL-28/29 (Sheppard *et al.*, 2003), the Type III IFNs (IFN λ 1–3) function as monomeric cytokines engaging a unique high affinity receptor, IFNLR1, and also utilizing the IL-10R2 chain as their low affinity receptor as discussed above (Kotenko *et al.*, 2003). Activities of the Type III IFNs are limited to epithelial cells and some dendritic cells due to the restricted cell surface expression of the IFNLR1 receptor (de Weerd and Nguyen, 2012). The crystal structure of IFN λ 3 demonstrated the topological similarity between the Type III IFNs and IL-10 family cytokines, in particular IL-22 (Gad *et al.*, 2009); it is therefore not surprising that the Type III IFNs signal via the IL-10R2 receptor (Langer *et al.*, 2004). The crystal

structure of ligand-bound IFNLR1 further revealed the structural similarities between this receptor and IL-10R1 and IL-10R2 (Miknis *et al.*, 2010). However, the manner in which the ligand docks onto its high affinity receptor is distinct from that of the IL-10:IL-10R1 interaction, with IFN λ 1 and IL-10 binding at different angles around the elbow between the two FnIII domains of their respective receptors (Miknis *et al.*, 2010).

Intracellular Signaling Complexes

The binding of a cytokine to the extracellular domain of its cognate receptor triggers signaling events within the receptor cytoplasmic domain that ultimately lead to the regulation of specific cellular outcomes. The ability of receptor cytoplasmic tails to function as scaffolds upon which signaling complexes are assembled allows the activated receptor to physically couple to downstream pathways. In particular, phosphorylation of the cytoplasmic tails of receptors in response to cytokine allows the recruitment and binding of a diverse array of signaling proteins that include enzymes (e.g., kinases, phosphatases, GTPases, and ubiquitin ligases), adapter proteins, and transcription factors.

Type I and type II cytokine receptors lack intrinsic tyrosine kinase activity but are able to activate the JAKs which consist of four members (JAK1, JAK2, JAK3, and TYK2) and which constitutively associate with weakly conserved 'box 1' and 'box 2' motifs in the membrane-proximal regions of receptors (Kiu and Nicholson, 2012; Wilks, 1989). High-affinity ligand binding induces either receptor oligomerization (e.g., β c family of receptors, gp130 family of receptors) and/or conformational changes (e.g. EPOR, hGHR), which serve to juxtapose the receptor cytoplasmic domains and their associated JAKs leading to transphosphorylation of receptor cytoplasmic tails. The tyrosine phosphorylation of receptor cytoplasmic tails provides docking sites for the binding of proteins containing Src Homology-2 (SH2)- or phosphotyrosine binding (PTB)-domains (Pawson, 2004). Further tyrosine phosphorylation of recruited proteins then allows the formation of higher order complexes on the receptor cytoplasmic tail mediated by a network of SH2 and PTB interactions. For example, JAKs have been shown to phosphorylate the STAT proteins, which consist of a family of seven latent transcription factors (Kiu and Nicholson, 2012). The JAK/STAT pathway, initially identified as being activated by IFN receptors (Darnell *et al.*, 1994), has been shown to be a critical regulator of type I and type II cytokine receptor activities. For example, both the single chain cytokine receptors (EPOR, hGHR, PRLR, Figure 2), and the β c family of receptors (IL-3, GM-CSF, IL-5, Figure 2) utilize JAK2 that signal via STAT3 and STAT5 (Guthridge *et al.*, 1998; Hennighausen and Robinson, 2008). The IL-2/IL-4 receptor family activates JAK1 and JAK3 that signal via STAT1, STAT3, STAT5, and STAT6 (Imada and Leonard, 2000; Kiu and Nicholson, 2012). The gp130 cytokine receptor family activates JAK1 to signal via STAT1, STAT3, and STAT5 (Kiu and Nicholson, 2012). Furthermore, Type I and Type II IFN receptors utilize JAK1, JAK2, and TYK2 that signal via STAT1, STAT2, STAT5, and STAT6 (Darnell *et al.*, 1994; Kiu and Nicholson, 2012).

In many cases, selective recognition of phosphotyrosine motifs by different SH2- and PTB-domain proteins allows the

recruitment of unique combinations of signaling proteins to the cytoplasmic tails of different cytokine receptors providing at least one mechanism by which specific signals can be generated (Pawson, 2004). For example, a specific SH2 docking site for SH2-containing protein tyrosine phosphatase 2 (SHP2) has been demonstrated in gp130 while a specific PTB binding site for Shc has been identified in the IL-3 receptor, each of which is important for activating Ras/ERK signaling (Jenkins *et al.*, 2005; Okuda *et al.*, 1997). In contrast, the ability to recruit negative regulators is also important for ensuring the termination of cytokine signaling. For example, the recruitment of the suppressor of cytokine signaling (SOCS) proteins, consisting of eight family members, provides an important negative feedback mechanism for down-regulating cytokine signaling via several mechanisms which include JAK inhibition and the proteasomal degradation of SOCS binding partners (Yoshimura *et al.*, 2007).

In addition to tyrosine phosphorylation, cytokine receptors are also phosphorylated on serine residues, which play important roles in initiating downstream signaling events. In an analogous manner to the binding of PTB-domains to phosphotyrosine residues, signaling complexes can also be assembled by the binding of phosphoserine binding proteins/domains to phosphoserine residues in the cytoplasmic tails of cytokine receptors. For example, the 14-3-3 family of phosphoserine/threonine binding proteins are able to interact with a number of cytokine receptors including the IL-9, PRL, GM-CSF and IL-3 receptors allowing the recruitment of signaling proteins such as PI3-kinase (Guthridge *et al.*, 2000; Olayioye *et al.*, 2003; Sliva *et al.*, 2000). Furthermore, the tyrosine phosphorylation of phosphoserine/threonine binding proteins such as 14-3-3 as well as the serine phosphorylation of SH2- and PTB-domain proteins such as insulin receptor substrate-1 (IRS-1) and STATs not only provide mechanisms for the integration of both phosphotyrosine and phosphoserine pathways but also for mediating specificity in signaling downstream of pleiotropic cytokine receptors (Barry *et al.*, 2009; Guthridge *et al.*, 2000; Pesu *et al.*, 2000; Weigert *et al.*, 2006).

Concluding Remarks

Cytokine receptors are capable of mediating an extraordinary array of biological responses. Yet, following the cloning and sequencing of many cytokine receptors in the late 1980s and into the 1990s, it became evident that, despite the functional pleiotropy of type I and type II cytokine receptors, there was strong evolutionary conservation in terms of their domain organization and, to some extent, specific sequence motifs. It was not until the application of X-ray crystallography, first to cytokines, then to receptors and more recently to ligand-bound receptors that the structure–function relationships of cytokine receptors could be understood at the molecular level. This increasing knowledge in cytokine structural biology is driving new opportunities for developing novel therapies to modulate cytokine function in a diverse range of diseases including malignancies, chronic inflammation and immunological pathologies. Whilst approaches to interfere with cytokine receptor activation are currently largely based on antibodies to the cytokines themselves or to their surface receptors (Broughton

et al., 2014; Sun *et al.*, 1996) new strategies are emerging. For example, following the successes in the structure-guided rational design of therapeutics that function as antagonists of cytoplasmic protein–protein interactions (Oltsdorf *et al.*, 2005), similar approaches are now being applied to identify compounds that antagonize cytokine–receptor interactions. In contrast, our understanding of the molecular structure of cytokine receptor cytoplasmic tails has lagged considerably (Brooks *et al.*, 2014). Although cytokine receptor cytoplasmic tails are widely depicted as disordered linear sequences, this is unlikely to be the case given the frequency with which disease-related mutations in cytoplasmic domains (particularly proline substitutions which disrupt secondary structure) lie outside known phosphotyrosine and phosphoserine docking sites. Thus, an important future challenge will be to understand the structure–function relationships for cytokine receptor cytoplasmic domains and apply that understanding to the development of new therapies for the treatment of human diseases.

Acknowledgments

This work was supported by the National Health and Medical Research Council of Australia (NHMRC) Program Grant (565217) to A.F.L. and M.W.P. and by Cancer Australia (AL). Infrastructure support from the NHMRC Independent Research Institutes Infrastructure Support Scheme and the Victorian State Government Operational Infrastructure Support Program to St. Vincent's Institute and MIMR-PHI are gratefully acknowledged. U.D. is an NHMRC Postdoctoral Fellow, S.E.B. is a Postdoctoral Fellow supported by the Leukaemia Foundation of Australia. M.W.P. and M.A.G. are NHMRC Research Fellows.

See also: Cell Communication: Intracellular Pathways: The JAK–STAT–SOCS Signaling Cascade; The MAPK Signaling Cascades

References

- Andrews, A.L., Nordgren, I.K., Kirby, I., *et al.*, 2009. Cytoplasmic tail of IL-13R α 2 regulates IL-4 signal transduction. *Biochemical Society Transactions* 37, 873–876.
- Arend, W.P., Palmer, G., Gabay, C., 2008. IL-1, IL-18, and IL-33 families of cytokines. *Immunological Reviews* 223, 20–38.
- Arima, K., Sato, K., Tanaka, G., *et al.*, 2005. Characterization of the interaction between interleukin-13 and interleukin-13 receptors. *Journal of Biological Chemistry* 280, 24915–24922.
- Baeten, D.L.P., Kuchroo, V.K., 2013. Interleukin-17 and a tale of two autoimmune diseases. *Nature Medicine* 19, 824–825.
- Banner, D.W., D'Arcy, A., Chene, C., *et al.*, 1996. The crystal structure of the complex of blood coagulation factor VIIa with soluble tissue factor. *Nature* 380, 41–46.
- Barnes, P.J., 2002. Cytokine modulators as novel therapies for asthma. *Annual Review of Pharmacology and Toxicology* 42, 81–98.
- Barry, E.F., Felquer, F.A., Powell, J.A., *et al.*, 2009. 14-3-3:Shc scaffolds integrate phosphoserine and phosphotyrosine signaling to regulate phosphatidylinositol 3-kinase activation and cell survival. *Journal of Biological Chemistry* 284, 12080–12090.
- Bazan, J.F., 1990a. Structural design and molecular evolution of a cytokine receptor superfamily. *Proceedings of the National Academy of Sciences of the United States of America* 87, 6934–6938.

- Bazan, J.F., 1990b. Haemopoietic receptors and helical cytokines. *Immunology Today* 11, 350–354.
- Bialock, E.L., Chien, H., Dix, R.D., 2013. Murine cytomegalovirus downregulates interleukin-17 in mice with retrovirus-induced immunosuppression that are susceptible to experimental cytomegalovirus retinitis. *Cytokine* 61, 862–875.
- Bleicher, L., de Moura, P.R., Watanabe, L., *et al.*, 2008. Crystal structure of the IL-22/IL-22R1 complex and its implications for the IL-22 signaling mechanism. *FEBS Letters* 582, 2985–2992.
- Bodhankar, S., Chen, Y., Vandenbark, A.A., Murphy, S.J., Offner, H., 2013. IL-10-producing B-cells limit CNS inflammation and infarct volume in experimental stroke. *Metabolic Brain Disease* 28, 375–386.
- Boraschi, D., Bossu, P., Ruggiero, P., *et al.*, 1995. Mapping of receptor binding sites on IL-1 beta by reconstruction of IL-1ra-like domains. *Journal of Immunology* 155, 4719–4725.
- Boraschi, D., Tagliabue, A., 2013. The interleukin-1 receptor family. *Seminars in Immunology* 25, 394–407.
- Boulanger, M.J., Bankovich, A.J., Kortemmer, T., Baker, D., Garcia, K.C., 2003a. Convergent mechanisms for recognition of divergent cytokines by the shared signaling receptor gp130. *Molecular Cell* 12, 577–589.
- Boulanger, M.J., Chow, D.C., Brevnova, E.E., Garcia, K.C., 2003b. Hexameric structure and assembly of the interleukin-6/IL-6 alpha-receptor/gp130 complex. *Science* 300, 2101–2104.
- Brooks, A.J., Dai, W., O'Mara, M.L., *et al.*, 2014. Mechanism of activation of protein kinase JAK2 by the growth hormone receptor. *Science* 344, 1249783.
- Broughton, S.E., Hercus, T.R., Hardy, M.P., *et al.*, 2014. Dual mechanism of IL-3 receptor blockade by an anti-cancer antibody. *Cell Reports* 8 (2), 410–419.
- Brown, A.L., Salerno, D.G., Sadras, T., *et al.*, 2012. The GM-CSF receptor utilizes β -catenin and Tcf4 to specify macrophage lineage differentiation. *Differentiation* 83, 47–59.
- Brown, R.J., Adams, J.J., Pelekanos, R.A., *et al.*, 2005. Model for growth hormone receptor activation based on subunit rotation within a receptor dimer. *Nature Structural and Molecular Biology* 12, 814–821.
- Broxmeyer, H.E., Hoggatt, J., O'Leary, H.A., *et al.*, 2012. Dipeptidylpeptidase 4 negatively regulates colony-stimulating factor activity and stress hematopoiesis. *Nature Medicine* 18, 1786–1796.
- Burger, D., Chicheportiche, R., Giri, J.G., Dayer, J.M., 1995. The inhibitory activity of human interleukin-1 receptor antagonist is enhanced by type II interleukin-1 soluble receptor and hindered by type I interleukin-1 soluble receptor. *Journal of Clinical Investigation* 96, 38–41.
- Burmester, G.R., Feist, E., Dorner, T., 2014. Emerging cell and cytokine targets in rheumatoid arthritis. *Nature Reviews Rheumatology* 10, 77–88.
- Busfield, S.J., Biondo, M., Wong, M., *et al.*, 2014. Targeting of acute myeloid leukaemia *in vitro* and *in vivo* with an anti-CD123 mAb engineered for optimal ADCC. *Leukemia* 28, 2213–2221.
- Carr, P.D., Gustin, S.E., Church, A.P., *et al.*, 2001. Structure of the complete extracellular domain of the common β subunit of the human GM-CSF, IL-3 and IL-5 receptors reveals a novel dimer configuration. *Cell* 104, 291–300.
- Casadio, R., Frigimelica, E., Bossu, P., *et al.*, 2001. Model of interaction of the IL-1 receptor accessory protein IL-1RAcP with the IL-1 β /IL-1R $_I$ complex. *FEBS Letters* 499, 65–68.
- Chalaris, A., Garbers, C., Rabe, B., Rose-John, S., Scheller, J., 2011. The soluble interleukin 6 receptor: Generation and role in inflammation and cancer. *European Journal of Cell Biology* 90, 484–494.
- Chill, J.H., Nivasch, R., Levy, R., *et al.*, 2002. The human interferon receptor: NMR-based modeling, mapping of the IFN- α 2 binding site, and observed ligand-induced tightening. *Biochemistry* 41, 3575–3585.
- Chirifu, M., Hayashi, C., Nakamura, T., *et al.*, 2007. Crystal structure of the IL-15-IL-15R α complex, a cytokine-receptor unit presented in trans. *Nature Immunology* 8, 1001–1007.
- Chomarat, P., Banchereau, J., 1998. Interleukin-4 and interleukin-13: Their similarities and discrepancies. *International Reviews of Immunology* 17, 1–52.
- Chow, D., He, X., Snow, A.L., Rose-John, S., Garcia, K.C., 2001. Structure of an extracellular gp130 cytokine receptor signaling complex. *Science* 291, 2150–2155.
- Codarra, L., Gyölvési, G., Tosevski, V., *et al.*, 2011. ROR γ t drives production of the cytokine GM-CSF in helper T cells, which is essential for the effector phase of autoimmune neuroinflammation. *Nature Immunology* 12, 560–567.
- Colotta, F., Dower, S.K., Sims, J.E., Mantovani, A., 1994. The type II 'decoy' receptor: A novel regulatory pathway for interleukin 1. *Immunology Today* 15, 562–566.
- Dahl, C., Hoffmann, H.J., Saito, H., Schütz, P.O., 2004. Human mast cells express receptor for IL-3, IL-5, and GM-CSF; a partial map of receptors on human mast cells cultured *in vitro*. *Allergy* 59, 1087–1096.
- Darnell Jr., J.E., Kerr, I.M., Stark, G.R., 1994. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* 264, 1415–1421.
- Dey, R., Ji, K., Liu, Z., Chen, L., 2009. A cytokine-cytokine interaction in the assembly of higher-order structure and activation of the interleukin-3:receptor complex. *PLoS One* 4, e5188.
- Dillon, S.R., Sprecher, C., Hammond, A., *et al.*, 2004. Interleukin 31, a cytokine produced by activated T cells, induces dermatitis in mice. *Nature Immunology* 5, 752–760.
- Dinarello, C.A., 2013. Overview of the interleukin-1 family of ligands and receptors. *Seminars in Immunology* 25, 389–393.
- Dinarello, C.A., Renfer, L., Wolff, S.M., 1977. Human leukocytic pyrogen: Purification and development of a radioimmunoassay. *Proceedings of the National Academy of Sciences of the United States of America* 74, 4624–4627.
- Dinarello, C.A., Simon, A., van der Meer, J.W., 2012. Treating inflammation by blocking interleukin-1 in a broad spectrum of diseases. *Nature Reviews Drug Discovery* 11, 633–652.
- Ely, L.K., Fischer, S., Garcia, K.C., 2009. Structural basis of receptor sharing by interleukin 17 cytokines. *Nature Immunology* 10, 1245–1251.
- Elzinga, B.M., Twomey, C., Powell, J.C., Harte, F., McCarthy, J.V., 2009. Interleukin-1 receptor type 1 is a substrate for gamma-secretase-dependent regulated intramembrane proteolysis. *Journal of Biological Chemistry* 284, 1394–1409.
- Feng, Y., Klein, B.K., McWherter, C.A., 1996. Three-dimensional solution structure and backbone dynamics of a variant of human interleukin-3. *Journal of Molecular Biology* 259, 524–541.
- Gad, H.H., Dellgren, C., Hamming, O.J., *et al.*, 2009. Interferon- λ is functionally an interferon but structurally related to the interleukin-10 family. *Journal of Biological Chemistry* 284, 20869–20875.
- Garlanda, C., Dinarello, C.A., Mantovani, A., 2013. The interleukin-1 family: Back to the future. *Immunity* 39, 1003–1018.
- Gay, N.J., Keith, F.J., 1991. Drosophila Toll and IL-1 receptor. *Nature* 351, 355–356.
- Gill, S., Tasian, S.K., Ruella, M., *et al.*, 2014. Preclinical targeting of human acute myeloid leukemia and myeloblasts using chimeric antigen receptor-modified cells. *Blood* 123, 2343–2354.
- Gosselin, D., Bellavance, M.A., Rivest, S., 2013. IL-1RAcPb signaling regulates adaptive mechanisms in neurons that promote their long-term survival following excitotoxic insults. *Frontiers in Cellular Neuroscience* 7, 9. doi:10.3389/fncel.2013.00009.
- Greenfeder, S.A., Nunes, P., Kwee, L., *et al.*, 1995. Molecular cloning and characterization of a second subunit of the interleukin 1 receptor complex. *Journal of Biological Chemistry* 270, 13757–13765.
- Gu, C., Wu, L., Li, X., 2013. IL-17 family: Cytokines, receptors and signaling. *Cytokine* 64, 477–485.
- Guthridge, M.A., Barry, E.F., Felquer, F.A., *et al.*, 2004. The phosphoserine-585-dependent pathway of the GM-CSF/IL-3/IL-5 receptors mediates hematopoietic cell survival through activation of NF- κ B and induction of bcl-2. *Blood* 103, 820–827.
- Guthridge, M.A., Stomski, F.C., Barry, E.F., *et al.*, 2000. Site-specific serine phosphorylation of the IL-3 receptor is required for hemopoietic cell survival. *Molecular Cell* 6, 99–108.
- Guthridge, M.A., Stomski, F.C., Thomas, D., *et al.*, 1998. Mechanism of activation of the GM-CSF, IL-3, and IL-5 family of receptors. *Stem Cells* 16, 301–313.
- Hamidullah, Changkija, B., Konwar, R., 2012. Role of interleukin-10 in breast cancer. *Breast Cancer Research and Treatment* 133, 11–21.
- Hansen, G., Hercus, T.R., McClure, B.J., *et al.*, 2008. The structure of the GM-CSF receptor complex reveals a distinct mode of cytokine receptor activation. *Cell* 134, 496–507.
- Heinrich, P.C., Behrmann, I., Haan, S., *et al.*, 2003. Principles of interleukin (IL)-6-type cytokine signalling and its regulation. *Biochemical Journal* 374, 1–20.
- Hennighausen, L., Robinson, G.W., 2008. Interpretation of cytokine signaling through the transcription factors STAT5A and STAT5B. *Genes and Development* 22, 711–721.
- Hercus, T.R., Cambareri, B., Dottore, M., *et al.*, 1994. Identification of residues in the first and fourth helices of human granulocyte-macrophage colony-stimulating factor involved in biologic activity and in binding to the α - and β -chains of its receptor. *Blood* 83, 3500–3508.
- Hercus, T.R., Dhagat, U., Kan, W.L., *et al.*, 2013. Signalling by the β c family of cytokines. *Cytokine and Growth Factor Reviews* 24, 189–201.
- Hercus, T.R., Thomas, D., Guthridge, M.A., *et al.*, 2009. The granulocyte-macrophage colony-stimulating factor receptor: Linking its structure to cell signaling and its role in disease. *Blood* 114, 1289–1298.
- Hiwase, D.K., White, D.L., Powell, J.A., *et al.*, 2010. Blocking cytokine signalling along with intense Bcr-Abl kinase inhibition induces apoptosis in primary CML progenitors. *Leukemia* 24, 771–778.

- Horsten, U., Müller-Newen, G., Gerhartz, C., *et al.*, 1997. Molecular modeling-guided mutagenesis of the extracellular part of gp130 leads to the identification of contact sites in the interleukin-6 (IL-6)/IL-6 receptor/gp130 complex. *Journal of Biological Chemistry* 272, 23748–23757.
- Huang, Y., Smith, D.E., Ibanez-Sandoval, O., Sims, J.E., Friedman, W.J., 2011. Neuron-specific effects of interleukin-1beta are mediated by a novel isoform of the IL-1 receptor accessory protein. *Journal of Neuroscience* 31, 18048–18059.
- Hymowitz, S.G., Filvaroff, E.H., Yin, J.P., *et al.*, 2001. IL-17s adopt a cystine knot fold: Structure and activity of a novel cytokine, IL-17F, and implications for receptor binding. *EMBO Journal* 20, 5332–5341.
- Imada, K., Leonard, W.J., 2000. The Jak-STAT pathway. *Molecular Immunology* 37, 1–11.
- Isaacs, A., Lindenmann, J., 1957. Virus interference. I. The interferon. *Proceedings of the Royal Society B: Biological Sciences* 147, 258–267.
- Ito, T., Suzuki, S., Kanaji, S., *et al.*, 2009. Distinct structural requirements for interleukin-4 (IL-4) and IL-13 binding to the shared IL-13 receptor facilitate cellular tuning of cytokine responsiveness. *Journal of Biological Chemistry* 284, 24289–24296.
- Jenkins, B.J., Grail, D., Nheu, T., *et al.*, 2005. Hyperactivation of Stat3 in gp130 mutant mice promotes gastric hyperproliferation and desensitizes TGF- β signaling. *Nature Medicine* 11, 845–852.
- Jin, L., Lee, E.M., Ramshaw, H.S., *et al.*, 2009. Monoclonal antibody-mediated targeting of CD123, IL-3 receptor α chain, eliminates human acute myeloid leukemic stem cells. *Cell Stem Cell* 5, 31–42.
- Jones, B.C., Logsdon, N.J., Walter, M.R., 2008. Structure of IL-22 bound to its high-affinity IL-22R1 chain. *Structure* 16, 1333–1344.
- Jordan, C.T., Upchurch, D., Szilvassy, S.J., *et al.*, 2000. The interleukin-3 receptor alpha chain is a unique marker for human acute myelogenous leukemic stem cells. *Leukemia* 14, 1777–1784.
- Josephson, K., Logsdon, N.J., Walter, M.R., 2001. Crystal structure of the IL-10/IL-10R1 complex reveals a shared receptor binding site. *Immunity* 15, 35–46.
- Kahlon, K.S., Brown, C., Cooper, L.J., *et al.*, 2004. Specific recognition and killing of glioblastoma multiforme by interleukin 13-zetakine redirected cytolytic T cells. *Cancer Research* 64, 9160–9166.
- Kalai, M., Montero-Julian, F.A., Grötzinger, J., *et al.*, 1997. Analysis of the human interleukin-6/human interleukin-6 receptor binding interface at the amino acid level: Proposed mechanism of interaction. *Blood* 89, 1319–1333.
- Kelly-Welch, A.E., Hanson, E.M., Boothby, M.R., Keegan, A.D., 2003. Interleukin-4 and interleukin-13 signaling connections maps. *Science* 300, 1527–1528.
- Kernebeck, T., Pflanz, S., Müller-Newen, G., *et al.*, 1999. The signal transducer gp130: Solution structure of the carboxy-terminal domain of the cytokine receptor homology region. *Protein Science* 8, 5–12.
- King, A., Balaji, S., Le, L.D., Crombleholme, T.M., Keswani, S.G., 2014. Regenerative wound healing: The role of interleukin-10. *Advances in Wound Care (New Rochelle)* 3, 315–323.
- Kioi, M., Kawakami, M., Shimamura, T., Husain, S.R., Puri, R.K., 2006. Interleukin-13 receptor $\alpha 2$ chain: A potential biomarker and molecular target for ovarian cancer therapy. *Cancer* 107, 1407–1418.
- Kirkham, B.W., Kavanaugh, A., Reich, K., 2013. Interleukin-17A: A unique pathway in immune-mediated diseases: Psoriasis, psoriatic arthritis and rheumatoid arthritis. *Immunology* 141, 133–142.
- Kiu, H., Nicholson, S.E., 2012. Biology and significance of the JAK/STAT signalling pathways. *Growth Factors* 30, 88–106.
- Kong, X., Feng, D., Mathews, S., Gao, B., 2013. Hepatoprotective and anti-fibrotic functions of interleukin-22: Therapeutic potential for the treatment of alcoholic liver disease. *Journal of Gastroenterology and Hepatology* 28 (Suppl 1), 56–60.
- Kono, T., Minami, Y., Taniguchi, T., 1993. The interleukin-2 receptor complex and signal transduction: Role of the β -chain. *Seminars in Immunology* 5, 299–307.
- Korn, T., Bettelli, E., Oukka, M., Kuchroo, V.K., 2009. IL-17 and Th17 Cells. *Annual Review of Immunology* 27, 485–517.
- Korpelainen, E.I., Gamble, J.R., Smith, W.B., *et al.*, 1993. The receptor for interleukin-3 is selectively induced in human endothelial cells by tumor necrosis factor α and potentiates interleukin 8 secretion and neutrophil transmigration. *Proceedings of the National Academy of Sciences of the United States of America* 90, 11137–11141.
- Kotenko, S.V., Gallagher, G., Baurin, V.V., *et al.*, 2003. IFN- λ s mediate antiviral protection through a distinct class II cytokine receptor complex. *Nature Immunology* 4, 69–77.
- Kovanen, P.E., Leonard, W.J., 2004. Cytokines and immunodeficiency diseases: Critical roles of the γ c-dependent cytokines interleukins 2, 4, 7, 9, 15, and 21, and their signaling pathways. *Immunological Reviews* 202, 67–83.
- Kusano, S., Kukimoto-Niino, M., Hino, N., *et al.*, 2012. Structural basis of interleukin-5 dimer recognition by its α receptor. *Protein Science* 21, 850–864.
- Langer, J.A., Cutrone, E.C., Kotenko, S., 2004. The Class II cytokine receptor (CRF2) family: Overview and patterns of receptor-ligand interactions. *Cytokine Growth Factor Reviews* 15, 33–48.
- Lantz, C.S., Boesinger, J., Song, C.H., *et al.*, 1998. Role for interleukin-3 in mast-cell and basophil development and in immunity to parasites. *Nature* 392, 90–93.
- LaPorte, S.L., Juo, Z.S., Vacklavikova, J., *et al.*, 2008. Molecular and structural basis of cytokine receptor pleiotropy in the interleukin-4/13 system. *Cell* 132, 259–272.
- Leonard, W.J., Noguchi, M., Russell, S.M., 1994b. Sharing of a common γ chain, γ c, by the IL-2, IL-4, and IL-7 receptors: Implications for X-linked severe combined immunodeficiency (XSCID). *Advances in Experimental Medicine and Biology* 365, 225–232.
- Leonard, W.J., Noguchi, M., Russell, S.M., McBride, O.W., 1994a. The molecular basis of X-linked severe combined immunodeficiency: The role of the interleukin-2 receptor γ chain as a common γ chain, γ c. *Immunological Reviews* 138, 61–86.
- Leung, J.M., Loke, P., 2013. A role for IL-22 in the relationship between intestinal helminths, gut microbiota and mucosal immunity. *International Journal for Parasitology* 43, 253–257.
- Levin, A.M., Bates, D.L., Ring, A.M., *et al.*, 2012. Exploiting a natural conformational switch to engineer an interleukin-2 'superkine'. *Nature* 484, 529–533.
- Liu, S., Song, X., Chrnyk, B.A., *et al.*, 2013. Crystal structures of interleukin 17A and its complex with IL-17 receptor A. *Nature Communications* 4, 1888.
- Livnah, O., Stura, E.A., Middleton, S.A., *et al.*, 1999. Crystallographic evidence for preformed dimers of erythropoietin receptor before ligand activation. *Science* 283, 987–990.
- Logsdon, N.J., Jones, B.C., Allman, J.C., *et al.*, 2004. The IL-10R2 binding hot spot on IL-22 is located on the N-terminal helix and is dependent on N-linked glycosylation. *Journal of Molecular Biology* 342, 503–514.
- Lopez, A.F., Hercus, T.R., Ekert, P., *et al.*, 2010. Molecular basis of cytokine receptor activation. *International Union of Biochemistry and Molecular Biology Life* 62, 509–518.
- Lopez, A.F., Sanderson, C.J., Gamble, J.R., *et al.*, 1988. Recombinant human interleukin 5 is a selective activator of human eosinophil function. *Journal of Experimental Medicine* 167, 219–224.
- Lorenzen, I., Lokau, J., Dusterhöft, S., *et al.*, 2012. The membrane-proximal domain of A Disintegrin and Metalloprotease 17 (ADAM17) is responsible for recognition of the interleukin-6 receptor and interleukin-1 receptor II. *FEBS Letters* 586, 1093–1100.
- Lu, H.L., Yang, C.Y., Chen, H.C., *et al.*, 2008. A novel alternatively spliced interleukin-1 receptor accessory protein mIL-1RAcP687. *Molecular Immunology* 45, 1374–1384.
- Lundström, W., Fewkes, N.M., Mackall, C.L., 2012. IL-7 in human health and disease. *Seminars in Immunology* 24, 218–224.
- Lupardus, P.J., Birnbaum, M.E., Garcia, K.C., 2010. Molecular basis for shared cytokine recognition revealed in the structure of an unusually high affinity complex between IL-13 and IL-13R $\alpha 2$. *Structure* 18, 332–342.
- Lupardus, P.J., Skiniotis, G., Rice, A.J., *et al.*, 2011. Structural snapshots of full-length Jak1, a transmembrane gp130/IL-6/IL-6R α cytokine receptor complex, and the receptor-Jak1 holocomplex. *Structure* 19, 45–55.
- Mansell, A., Jenkins, B.J., 2013. Dangerous liaisons between interleukin-6 cytokine and toll-like receptor families: A potent combination in inflammation and cancer. *Cytokine and Growth Factor Reviews* 24, 249–256.
- Masters, S.L., Simon, A., Aksentjevich, I., Kastner, D.L., 2009. Horror autoinflammatus: The molecular pathophysiology of autoinflammatory disease. *Annual Review of Immunology* 27, 621–668.
- Mattei, F., Schiavoni, G., Belardelli, F., Tough, D.F., 2001. IL-15 is expressed by dendritic cells in response to type I IFN, double-stranded RNA, or lipopolysaccharide and promotes dendritic cell activation. *Journal of Immunology* 167, 1179–1187.
- Mazzuchelli, R., Durum, S.K., 2007. Interleukin-7 receptor expression: Intelligent design. *Nature Reviews Immunology* 7, 144–154.
- McElroy, C.A., Dohm, J.A., Walsh, S.T.R., 2009. Structural and biophysical studies of the human IL-7/IL-7R α complex. *Structure* 17, 54–65.
- McElroy, C.A., Holland, P.J., Zhao, P., *et al.*, 2012. Structural reorganization of the interleukin-7 signaling complex. *Proceedings of the National Academy of Sciences of the United States of America* 109, 2503–2508.
- McMahan, C.J., Slack, J.L., Mosley, B., *et al.*, 1991. A novel IL-1 receptor, cloned from B cells by mammalian expression, is expressed in many cell types. *EMBO Journal* 10, 2821–2832.
- Miknis, Z.J., Magracheva, E., Li, W., *et al.*, 2010. Crystal structure of human interferon- $\lambda 1$ in complex with its high-affinity receptor interferon- λ R1. *Journal of Molecular Biology* 404, 650–664.

- Milburn, M.V., Hassell, A.M., Lambert, M.H., *et al.*, 1993. A novel dimer configuration revealed by the crystal structure at 2.4 Å resolution of human interleukin-5. *Nature* 363, 172–176.
- Mintz, A., Gibo, D.M., Slagle-Webb, B., Christensen, N.D., Debinski, W., 2002. IL-13R α 2 is a glioma-restricted receptor for interleukin-13. *Neoplasia* 4, 388–399.
- Mizoguchi, A., 2012. Healing of intestinal inflammation by IL-22. *Inflammatory Bowel Diseases* 18, 1777–1784.
- de Moura, P.R., Watanabe, L., Bleicher, L., *et al.*, 2009. Crystal structure of a soluble decoy receptor IL-22BP bound to interleukin-22. *FEBS Letters* 583, 1072–1077.
- Mueller, T.D., Zhang, J.L., Sebald, W., Duschl, A., 2002. Structure, binding, and antagonists in the IL-4/IL-13 receptor system. *Biochimica et Biophysica Acta* 1592, 237–250.
- Murakami, M., Narazaki, M., Hibi, M., *et al.*, 1991. Critical cytoplasmic region of the interleukin 6 signal transducer gp130 is conserved in the cytokine receptor family. *Proceedings of the National Academy of Sciences of the United States of America* 88, 11349–11353.
- Nagem, R.A., Colau, D., Dumoutier, L., *et al.*, 2002. Crystal structure of recombinant human interleukin-22. *Structure* 10, 1051–1062.
- Nakamura, Y., Russell, S.M., Mess, S.A., *et al.*, 1994. Heterodimerization of the IL-2 receptor β - and γ -chain cytoplasmic domains is required for signalling. *Nature* 369, 330–333.
- Nelms, K., Keegan, A.D., Zamorano, J., Ryan, J.J., Paul, W.E., 1999. The IL-4 receptor: Signaling mechanisms and biologic functions. *Annual Review of Immunology* 17, 701–738.
- Nelson, B.H., Willerford, D.M., 1998. Biology of the interleukin-2 receptor. *Advances in Immunology* 70, 1–81.
- Nguyen, L., Rothwell, N.J., Pinteaux, E., Boutin, H., 2011. Contribution of interleukin-1 receptor accessory protein B to interleukin-1 actions in neuronal cells. *Neurosignals* 19, 222–230.
- Nievergall, E., Ramshaw, H.S., Yong, A.S., *et al.*, 2014. Monoclonal antibody targeting of IL-3 receptor α with CSL362 effectively depletes CML progenitor and stem cells. *Blood* 123, 1218–1228.
- Okuda, K., Smith, L., Griffin, J.D., Foster, R., 1997. Signaling functions of the tyrosine residues in the β c chain of the granulocyte-macrophage colony-stimulating factor receptor. *Blood* 90, 4759–4766.
- Olajoye, M.A., Guthridge, M.A., Stomski, F.C., *et al.*, 2003. Threonine 391 phosphorylation of the human prolactin receptor mediates a novel interaction with 14-3-3 proteins. *Journal of Biological Chemistry* 278, 32929–32935.
- Olsen, S.K., Ota, N., Kishishita, S., *et al.*, 2007. Crystal structure of the interleukin-15:interleukin-15 receptor α complex: Insights into trans and cis presentation. *Journal of Biological Chemistry* 282, 37191–37204.
- Oltersdorf, T., Elmore, S.W., Shoemaker, A.R., *et al.*, 2005. An inhibitor of Bcl-2 family proteins induces regression of solid tumours. *Nature* 435, 677–681.
- O'Neill, L.A., Bryant, C.E., Doyle, S.L., 2009. Therapeutic targeting of Toll-like receptors for infectious and inflammatory diseases and cancer. *Pharmacological Reviews* 61, 177–197.
- Orlando, S., Sironi, M., Bianchi, G., *et al.*, 1997. Role of metalloproteases in the release of the IL-1 type II decoy receptor. *Journal of Biological Chemistry* 272, 31764–31769.
- Ouyang, P., Rakus, K., van Beurden, S.J., *et al.*, 2014. IL-10 encoded by viruses: A remarkable example of independent acquisition of a cellular gene by viruses and its subsequent evolution in the viral genome. *Journal of General Virology* 95, 245–262.
- Pahl, H.L., 1999. Activators and target genes of Rel/NF- κ B transcription factors. *Oncogene* 18, 6853–6866.
- Pan, H.F., Li, X.P., Zheng, S.G., Ye, D.Q., 2013. Emerging role of interleukin-22 in autoimmune diseases. *Cytokine and Growth Factor Reviews* 24, 51–57.
- Patino, E., Kotzsch, A., Saremba, S., *et al.*, 2011. Structure analysis of the IL-5 ligand-receptor complex reveals a wrench-like architecture for IL-5R α . *Structure* 19, 1864–1875.
- Pawson, T., 2004. Specificity in signal transduction: From phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell* 116, 191–203.
- Pesu, M., Takaluoma, K., Aittomäki, S., *et al.*, 2000. Interleukin-4-induced transcriptional activation by stat6 involves multiple serine/threonine kinase pathways and serine phosphorylation of stat6. *Blood* 95, 494–502.
- Pletnev, S., Magracheva, E., Wlodawer, A., Zdanov, A., 2005. A model of the ternary complex of interleukin-10 with its soluble receptors. *BMC Structural Biology* 5, 10.
- Pylayeva-Gupta, Y., Lee, K.E., Hajdu, C.H., Miller, G., Bar-Sagi, D., 2012. Oncogenic Kras-induced GM-CSF production promotes the development of pancreatic neoplasia. *Cancer Cell* 21, 836–847.
- Ramshaw, H.S., Guthridge, M.A., Stomski, F.C., *et al.*, 2007. The Shc-binding site of the β c subunit of the GM-CSF/IL-3/IL-5 receptors is a negative regulator of hematopoiesis. *Blood* 110, 3582–3590.
- Rickert, M., Wang, X., Boulanger, M.J., Goriatheva, N., Garcia, K.C., 2005. The structure of interleukin-2 complexed with its alpha receptor. *Science* 308, 1477–1480.
- Ring, A.M., Lin, J.X., Feng, D., *et al.*, 2012. Mechanistic and structural insight into the functional dichotomy between IL-2 and IL-15. *Nature Immunology* 13, 1187–1195.
- Roan, F., Bell, B.D., Stoklasek, T.A., *et al.*, 2012. The multiple facets of thymic stromal lymphopoietin (TSLP) during allergic inflammation and beyond. *Journal of Leukocyte Biology* 91, 877–886.
- Rochman, Y., Spolski, R., Leonard, W.J., 2009. New insights into the regulation of T cells by γ c family cytokines. *Nature Reviews Immunology* 9, 480–490.
- Rozwarski, D.A., Diederichs, K., Hecht, R., Boone, T., Karplus, P.A., 1996. Refined crystal structure and mutagenesis of human granulocyte-macrophage colony-stimulating factor. *Proteins* 26, 304–313.
- Rutz, S., Eidschinken, C., Ouyang, W., 2013. IL-22, not simply a Th17 cytokine. *Immunological Reviews* 252, 116–132.
- Sabat, R., 2010. IL-10 family of cytokines. *Cytokine and Growth Factor Reviews* 21, 315–324.
- Sabat, R., Ouyang, W., Wolk, K., 2014. Therapeutic opportunities of the IL-22-IL-22R1 system. *Nature Reviews Drug Discovery* 13, 21–38.
- Schreuder, H., Tardif, C., Trump-Kallmeyer, S., *et al.*, 1997. A new cytokine-receptor binding mode revealed by the crystal structure of the IL-1 receptor with an antagonist. *Nature* 386, 194–200.
- Schuster, B., Kovaleva, M., Sun, Y., *et al.*, 2003. Signaling of human ciliary neurotrophic factor (CNTF) revisited. The interleukin-6 receptor can serve as an α -receptor for CNTF. *Journal of Biological Chemistry* 278, 9528–9535.
- Schwantner, A., Dingley, A.J., Ozbek, S., Rose-John, S., Grötzinger, J., 2004. Direct determination of the interleukin-6 binding epitope of the interleukin-6 receptor by NMR spectroscopy. *Journal of Biological Chemistry* 279, 571–576.
- Schwartzentruber, D.J., Lawson, D.H., Richards, J.M., *et al.*, 2011. gp100 peptide vaccine and interleukin-2 in patients with advanced melanoma. *New England Journal of Medicine* 364, 2119–2127.
- Sheppard, P., Kindsvogel, W., Xu, W., *et al.*, 2003. IL-28, IL-29 and their class II cytokine receptor IL-28R. *Nature Immunology* 4, 63–68.
- Skiniotis, G., Boulanger, M.J., Garcia, K.C., Walz, T., 2005. Signaling conformations of the tall cytokine receptor gp130 when in complex with IL-6 and IL-6 receptor. *Nature Structural and Molecular Biology* 12, 545–551.
- Sliva, D., Gu, M., Zhu, Y.X., *et al.*, 2000. 14–3-3 ζ interacts with the α -chain of human interleukin 9 receptor. *Biochemical Journal* 345, 741–747.
- Slobedman, B., Barry, P.A., Spencer, J.V., Avdic, S., Abendroth, A., 2009. Virus-encoded homologs of cellular interleukin-10 and their control of host immune function. *Journal of Virology* 83, 9618–9629.
- Smith, D.E., Hanna, R., Friend, D., *et al.*, 2003. The soluble form of IL-1 receptor accessory protein enhances the ability of soluble type II IL-1 receptor to inhibit IL-1 action. *Immunity* 18, 87–96.
- Smith, D.E., Lipsky, B.P., Russell, C., *et al.*, 2009. A central nervous system-restricted isoform of the interleukin-1 receptor accessory protein modulates neuronal responses to interleukin-1. *Immunity* 30, 817–831.
- Stauber, D.J., Debler, E.W., Horton, P.A., Smith, K.A., Wilson, I.A., 2006. Crystal structure of the IL-2 signaling complex: Paradigm for a heterotrimeric cytokine receptor. *Proceedings of the National Academy of Sciences of the United States of America* 103, 2788–2793.
- Striz, I., Brabcova, E., Kolesar, L., Sekerkova, A., 2014. Cytokine networking of innate immunity cells: A potential target of therapy. *Clinical Science* 126, 593–612.
- Strunk, J.J., Gregor, I., Becker, Y., *et al.*, 2008. Ligand binding induces a conformational change in ifnar1 that is propagated to its membrane-proximal domain. *Journal of Molecular Biology* 377, 725–739.
- Su, S., Liu, Q., Chen, J., *et al.*, 2014. A positive feedback loop between mesenchymal-like cancer cells and macrophages is essential to breast cancer metastasis. *Cancer Cell* 25, 605–620.
- Sun, Q., Woodcock, J.M., Rapoport, A., *et al.*, 1996. Monoclonal antibody 7G3 recognizes the N-terminal domain of the human interleukin-3 (IL-3) receptor α -chain and functions as a specific IL-3 receptor antagonist. *Blood* 87, 83–92.
- Symons, J.A., Young, P.R., Duff, G.W., 1995. Soluble type II interleukin 1 (IL-1) receptor binds and blocks processing of IL-1 β precursor and loses affinity for IL-1 receptor antagonist. *Proceedings of the National Academy of Sciences of the United States of America* 92, 1714–1718.
- Testa, U., Riccioni, R., Miliati, S., *et al.*, 2002. Elevated expression of IL-3R α in acute myelogenous leukaemia is associated with enhanced blast proliferation, increased cellularity, and poor prognosis. *Blood* 100, 2980–2988.

- Thomas, C., Moraga, I., Levin, D., *et al.*, 2011. Structural linkage between ligand discrimination and receptor activation by type I interferons. *Cell* 146, 621–632.
- Tsai, T.T., Chuang, Y.J., Lin, Y.S., *et al.*, 2013. An emerging role for the anti-inflammatory cytokine interleukin-10 in dengue virus infection. *Journal of Biomedical Science* 20, 40–48.
- Twomey, C., Qian, S., McCarthy, J.V., 2009. TRAF6 promotes ubiquitination and regulated intramembrane proteolysis of IL-1R1. *Biochemical Biophysical Research Communications* 381, 418–423.
- Vigers, G.P., Anderson, L.J., Caffes, P., Brandhuber, B.J., 1997. Crystal structure of the type-I interleukin-1 receptor complexed with interleukin-1 β . *Nature* 386, 190–194.
- de Vos, A.M., Ultsch, M., Kossiakoff, A.A., 1992. Human growth hormone and extracellular domain of its receptor: Crystal structure of the complex. *Science* 255, 306–312.
- Votavova, P., Tomala, J., Kovar, M., 2014. Increasing the biological activity of IL-2 and IL-15 through complexing with anti-IL-2 mAbs and IL-15R α -Fc chimera. *Immunology Letters* 159, 1–10.
- Waldmann, T.A., Tagaya, Y., 1999. The multifaceted regulation of interleukin-15 expression and the role of this cytokine in NK cell differentiation and host response to intracellular pathogens. *Annual Review of Immunology* 17, 19–49.
- Walsh, S.T.R., 2012. Structural insights into the common γ -chain family of cytokines and receptors from the interleukin-7 pathway. *Immunological Reviews* 250, 303–316.
- Walter, M.R., Windsor, W.T., Nagabhushan, T.L., *et al.*, 1995. Crystal structure of a complex between interferon- γ and its soluble high-affinity receptor. *Nature* 376, 230–235.
- Wang, D., Zhang, S., Li, L., *et al.*, 2010. Structural insights into the assembly and activation of IL-1 β with its receptors. *Nature Immunology* 11, 905–911.
- Wang, W.L., Li, H.Y., Zhang, M.S., *et al.*, 2013. Thymic stromal lymphopoietin: A promising therapeutic target for allergic diseases. *International Archives of Allergy and Immunology* 160, 18–26.
- Wang, X., Lupardus, P., LaPorte, S.L., Garcia, K.C., 2009. Structural biology of shared cytokine receptors. *Annual Review of Immunology* 27, 29–60.
- Wang, Y., Liu, X.H., Li, Y.H., Li, O., 2013. The paradox of IL-10-mediated modulation in cervical cancer. *Biomedical Reports* 1, 347–351.
- de Weerd, N.A., Nguyen, T., 2012. The interferons and their receptors — distribution and regulation. *Immunology and Cell Biology* 90, 483–491.
- de Weerd, N.A., Samarajiva, S.A., Hertzog, P.J., 2007. Type I interferon receptors: Biochemistry and biological functions. *Journal of Biological Chemistry* 282, 20053–20057.
- de Weerd, N.A., Vivian, J.P., Nguyen, T.K., *et al.*, 2013. Structural basis of a unique interferon- β signaling axis mediated via the receptor IFNAR1. *Nature Immunology* 14, 901–907.
- Weigert, C., Hennige, A.M., Lehmann, R., *et al.*, 2006. Direct cross-talk of interleukin-6 and insulin signal transduction via insulin receptor substrate-1 in skeletal muscle cells. *Journal of Biological Chemistry* 281, 7060–7067.
- Weiss, A., Schlessinger, J., 1998. Switching signals on or off by receptor dimerization. *Cell* 94, 277–280.
- Wells, J.A., de Vos, A.M., 1996. Hematopoietic receptor complexes. *Annual Review of Biochemistry* 65, 609–634.
- Wickham, J., Walsh, S.T., 2007. Crystallization and preliminary X-ray diffraction of human interleukin-7 bound to unglycosylated and glycosylated forms of its α -receptor. *Acta Crystallographica Section F: Structural Biology Crystallization Communications* 63, 865–869.
- Wilks, A.F., 1989. Two putative protein-tyrosine kinases identified by application of the polymerase chain reaction. *Proceedings of the National Academy of Sciences of the United States of America* 86, 1603–1607.
- Wills-Karp, M., 2004. Interleukin-13 in asthma pathogenesis. *Immunological Reviews* 202, 175–190.
- Wood, N., Whitters, M.J., Jacobson, B.A., *et al.*, 2003. Enhanced interleukin (IL)-13 responses in mice lacking IL-13 receptor α 2. *Journal of Experimental Medicine* 197, 703–709.
- Wright, J.F., Bennett, F., Li, B., *et al.*, 2008. The human IL-17F/IL-17A heterodimeric cytokine signals through the IL-17RA/IL-17RC receptor complex. *Journal of Immunology* 181, 2799–2805.
- Wynn, T.A., 2003. IL-13 effector functions. *Annual Review of Immunology* 21, 425–456.
- Xu, Y., Kershaw, N.J., Luo, C.S., *et al.*, 2010. Crystal structure of the entire ectodomain of gp130: Insights into the molecular assembly of the tall cytokine receptor complexes. *Journal of Biological Chemistry* 285, 21214–21218.
- Yawata, H., Yasukawa, K., Natsuka, S., *et al.*, 1993. Structure-function analysis of human IL-6 receptor: Dissociation of amino acid residues required for IL-6-binding and for IL-6 signal transduction through gp130. *EMBO Journal* 12, 1705–1712.
- Yoon, S.I., Jones, B.C., Logsdon, N.J., *et al.*, 2010. Structure and mechanism of receptor sharing by the IL-10R2 common chain. *Structure* 18, 638–648.
- Yoon, S.I., Logsdon, N.J., Sheikh, F., Donnelly, R.P., Walter, M.R., 2006. Conformational changes mediate interleukin-10 receptor 2 (IL-10R2) binding to IL-10 and assembly of the signaling complex. *Journal of Biological Chemistry* 281, 35088–35096.
- Yoshimura, A., Naka, T., Kubo, M., 2007. SOCS proteins, cytokine signalling and immune regulation. *Nature Reviews Immunology* 7, 454–465.
- Zaks-Zilberman, M., Harrington, A.E., Ishino, T., Chaiken, I.M., 2008. Interleukin-5 receptor subunit oligomerization and rearrangement revealed by fluorescence resonance energy transfer imaging. *Journal of Biological Chemistry* 283, 13398–13406.
- Zdanov, A., 2010. Structural analysis of cytokines comprising the IL-10 family. *Cytokine and Growth Factor Reviews* 21, 325–330.
- Zhang, B., Liu, C., Qian, W., *et al.*, 2013. Crystal structure of IL-17 receptor B SEFIR domain. *Journal of Immunology* 190, 2320–2326.
- Zhang, Y., Zhou, B., 2012. Functions of thymic stromal lymphopoietin in immunity and disease. *Immunologic Research* 52, 211–223.
- Ziegler, S.F., Artis, D., 2010. Sensing the outside world: TSLP regulates barrier immunity. *Nature Immunology* 11, 289–293.

Further Reading

- Banchereau, J., Pascual, V., O'Garra, A., 2012. From IL-2 to IL-37: The expanding spectrum of anti-inflammatory cytokines. *Nature Immunology* 13, 925–931.
- Brocker, C., Thompson, D., Matsumoto, A., Nebert, D.W., Vasilou, V., 2010. Evolutionary divergence and functions of the human interleukin (IL) gene family. *Human Genomics* 5, 30–55.
- Broughton, S.E., Hercus, T.R., Lopez, A.F., Parker, W.M., 2012. Cytokine activation at the cell surface. *Current Opinion in Structural Biology* 22, 350–359.
- Liongue, C., Ward, A.C., 2013. Evolution of the JAK-STAT pathway. *JAKSTAT* 2, e22756.

Relevant Websites

- <http://www.guidetopharmacology.org/IUPHAR/BPS>.
- <http://www.rcsb.org>
- RCSB Protein Data Bank.

TGF- β Superfamily Signaling

DSJ Miller and CS Hill, Cancer Research UK London Research Institute, London, UK

© 2016 Elsevier Inc. All rights reserved.

Introduction

The transforming growth factor β (TGF- β) superfamily of secreted ligands are found in all metazoans and emerged with multicellularity. They were originally thought to have evolved with dorsal-ventral (DV) patterning and thus only to be found in bilateria. However, it is now clear that TGF- β superfamily ligands are also found in sponges (Porifera) and in *Trichoplax adhaerens* (Placozoa) (Adamska *et al.*, 2007; Huminiecki *et al.*, 2009), indicating the ancient origin of these ligands. In humans the TGF- β superfamily comprises the TGF- β s themselves, activins, NODAL, bone morphogenetic proteins

(BMPs), growth and differentiation factors (GDFs), anti-Müllerian hormone (AMH) and also the two LEFTY proteins and INHA that act as ligand antagonists (Figure 1). In this article we will begin by outlining some of the key functions for which the TGF- β superfamily members are required. We will go on to describe how the ligands are produced and regulated, their cell surface receptors and how they signal to the nucleus through the intracellular transducers, the SMADs. After a general discussion about pathway regulation, we will end by highlighting a number of human diseases that are caused by mutation or loss of key components of TGF- β superfamily pathways.

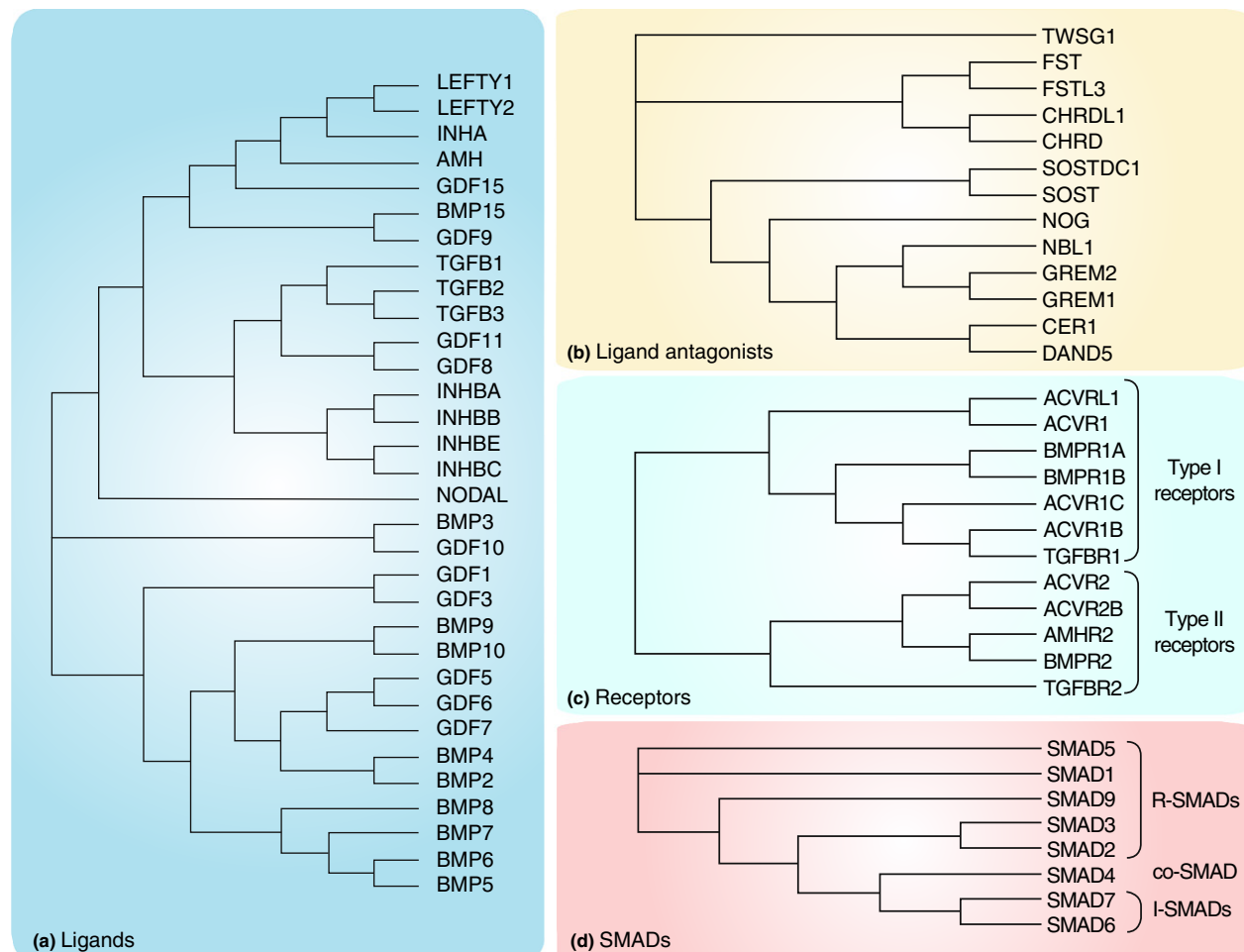


Figure 1 Components of TGF- β superfamily signaling. (a) Phylogenetic tree of all mammalian TGF- β superfamily ligands. (b) Phylogenetic tree of ligand antagonists. Antagonists are often described in the literature by their alternative names: TWSG1, twisted gastrulation 1; FST, follistatin; FSTL3, follistatin-like 3; CHRDL1, chordin-like 1; CHRDL, chordin; SOSTDC1, USAG1; SOST, sclerostin; NOG, noggin; NBL1, DAN; GREM2, PRDC; GREM1, gremlin 1; CER1, cerberus 1; DAND5, COCO. (c) Phylogenetic tree of TGF- β superfamily receptors. Type I and type II receptors are indicated. (d) Phylogenetic tree of SMAD transcription factors. R-SMADs, co-SMAD and I-SMADs are indicated. Note that SMAD9 was originally named SMAD8. For more details about the TGF- β superfamily components please see entries in the GeneCards website <http://www.genecards.org/> and, in the case of the SMADs, also the following website <http://smad.maciasnmr.net/>.

Function of TGF- β Superfamily Signaling

The TGF- β superfamily ligands primarily function in a paracrine fashion. They bind and activate cell surface receptors on neighboring cells, which signal to the nucleus to induce new programs of gene expression. Signaling by these ligands can

regulate many different cellular behaviors, among which are cell proliferation, differentiation, migration, invasion, adhesion, and apoptosis (Massague, 2012). As a result of their ability to regulate such a diversity of cellular activities, they play critical functional roles in many aspects of early embryonic development and tissue homeostasis in adults (Figure 2).

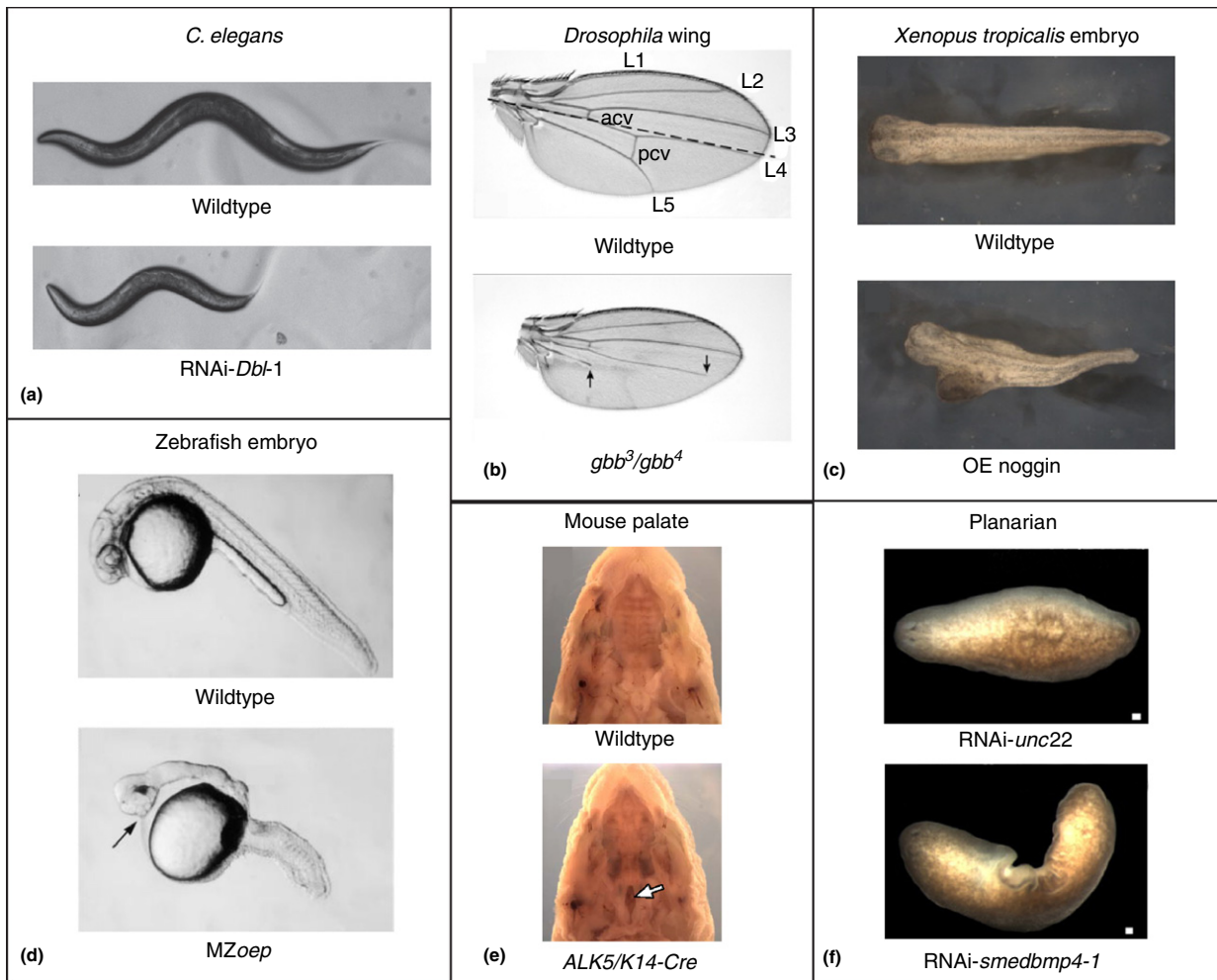


Figure 2 Perturbation of multiple components of TGF- β superfamily signaling lead to developmental defects in a wide range of organisms.

(a) *Caenorhabditis elegans* null for the TGF- β ligand homolog *Dbl-1* displays the small body size typical of all TGF- β pathway deletions in this organism. Reproduced with permission from Gumienny, T.L., Savage-Dunn, C., 2013. TGF- β signaling in *C. elegans*. WormBook 1–34.

(b) *Drosophila melanogaster* with a hypomorphic allele of the BMP homolog *gbb* have severely altered posterior vein development in the adult wing. The 5 longitudinal veins L1–L5 are labeled in the wildtype, as are the anterior crossvein (acv) and posterior crossvein (pcv). The dotted line depicts the anterior/posterior boundary. The arrows indicate vein truncations in the mutant wing. Reproduced with permission Ray, R.P., Wharton, K.A., 2001. Context-dependent relationships between the BMPs *gbb* and *dpp* during development of the *Drosophila* wing imaginal disk. Development 128, 3913–3925.

(c) *Xenopus tropicalis* embryos over-expressing the BMP antagonist NOG develop a secondary body axis. Reproduced with permission from Young, J.J., Cherone, J.M., Doyon, Y., et al., 2011. Efficient targeted gene disruption in the soma and germ line of the frog *Xenopus tropicalis* using engineered zinc-finger nucleases. Proceedings of the National Academy of Sciences of the United States of America 108, 7052–7057.

(d) Zebrafish embryos carrying a maternal zygotic (MZ) deletion of the co-receptor *tdgf1* (*oeop*) lack anterior and trunk mesoderm and all endoderm. Arrow indicates cyclopia. Reproduced with permission from Gritsman, K., Zhang, J., Cheng, S., et al., 1999. The EGF-CFC protein one-eyed pinhead is essential for nodal signaling. Cell 97, 121–132.

(e) Mice carrying a deletion of *Tgfb1* (*Alk5*) in the embryonic ectodermal and neural crest cell lineages display cleft palate (arrow). Reproduced with permission from Dudas, M., Kim, J., Li, W.Y., et al., 2006. Epithelial and ectomesenchymal role of the type I TGF- β receptor ALK5 during facial morphogenesis and palatal fusion. Developmental Biology 296, 298–314.

(f) A planarian with RNAi directed against the BMP2/4/DPP-like gene *smedbmp4-1* fails in lateral regeneration after amputation, while those treated with the control RNAi regenerated normally. Reproduced with permission from Reddien, P.W., Bermange, A.L., Kicza, A.M., Sanchez Alvarado, A., 2007. BMP signaling regulates the dorsal planarian midline and is needed for asymmetric regeneration. Development 134, 4043–4051.

Their roles are too numerous to outline in detail, so we will focus on some illustrative examples.

In very early embryos, the main roles of TGF- β superfamily members are in specifying the germ layers, patterning tissues across the different axes and maintaining pluripotency. For example, in zebrafish and *Xenopus*, graded BMP signaling determines the DV axis and is responsible for specifying and patterning tissues across this axis (Ramel and Hill, 2012). This role in DV patterning is also evident in arthropods such as *Drosophila*, although the DV axis is inverted in this case (De Robertis and Kuroda, 2004). The earliest role of NODAL in the mouse embryo is in the maintenance of the determinants of pluripotency, such as POU5F1 (Oct4) and NANOG at the blastocyst stage (Mesnard *et al.*, 2006). After implantation, graded NODAL signaling is important for defining the proximal–distal axis, which establishes the anterior–posterior axis of the embryo (Arnold and Robertson, 2009). NODAL signaling also functions in conjunction with WNT and BMP signaling to induce mesoderm and endoderm. At later stages, NODAL is required for left–right axis formation and patterning. The role for NODAL in mesendoderm differentiation and patterning and left–right asymmetry is also evident in zebrafish (Figure 2(d)), *Xenopus* and sea urchins (Duboc *et al.*, 2010; Schier, 2003). In many of these contexts, graded NODAL signaling is antagonized by a reverse gradient of BMP signaling, and in fact opposing gradients of BMP and NODAL signaling are sufficient to organize uncommitted cells of the zebrafish blastula animal pole into an embryo (Xu *et al.*, 2014). Later in development, BMPs and GDFs play critical roles in organogenesis, including development of bone and cartilage, heart, lung, kidney, eyes, reproductive systems, nervous system, placenta and placental connections, vasculogenesis, and angiogenesis (Miyazono *et al.*, 2010). Postnatally, BMPs play important roles in the induction of bone and cartilage formation, vascular homeostasis, and in iron metabolism (Corradini *et al.*, 2009).

The major developmental roles of the TGF- β s, as revealed by the phenotypes of mice deleted for TGF- β 1, 2 and 3, are in vasculogenesis, haematopoiesis, and development of a variety of organs (Goumans and Mummery, 2000). In adults, TGF- β s are important for controlling immunity, inflammation, wound healing and bone, muscle, adipose, vascular, and haematopoietic homeostasis (Massague, 2012). Activins were originally identified by virtue of their ability to stimulate release of follicle stimulating hormone from the pituitary and, in addition, control hormone production in the gonads and placenta, and regulate gametogenesis (Antsiferova and Werner, 2012). AMH plays a much more restricted role than many other TGF- β family members, being responsible for inducing regression of the Müllerian ducts during male sex differentiation, and it has an important role in ovarian folliculogenesis in females (Durlinger *et al.*, 2002).

TGF- β Superfamily Ligands

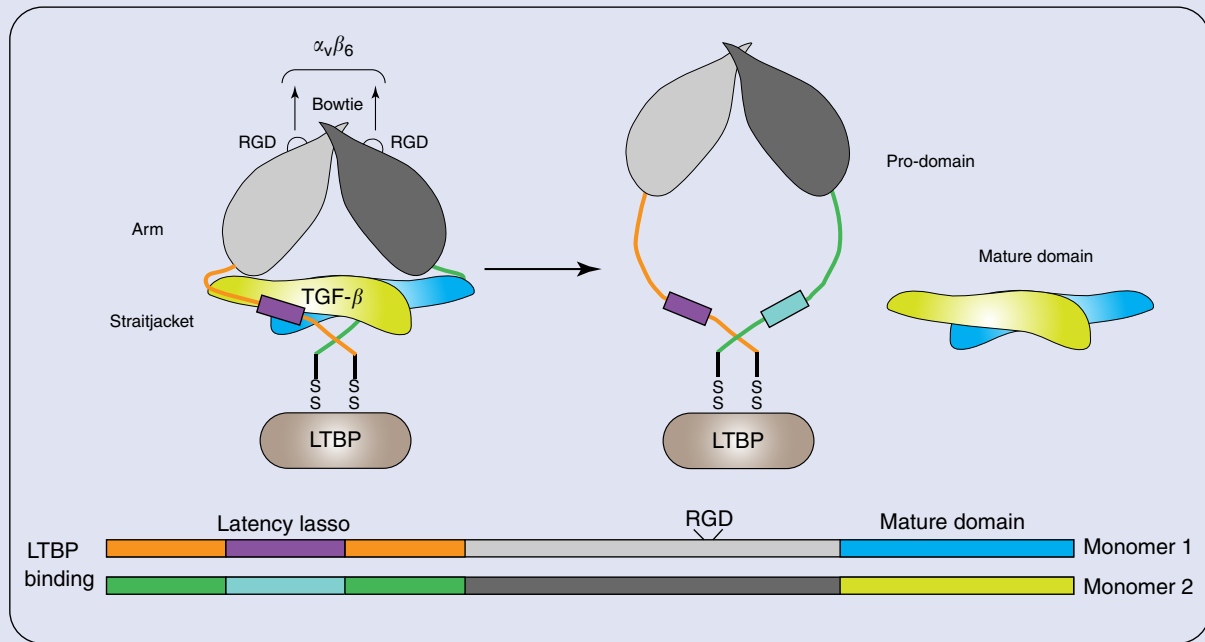
The TGF- β superfamily ligands function as disulphide-linked dimers, and the mature domain of the ligands forms a cystine knot, that is stabilized by three intramolecular disulphide bonds between six strictly conserved cysteine residues.

They are thus members of the so-called cystine knot family that also contains other hormones and growth factors such as nerve growth factor, platelet-derived growth factor and human chorionic gonadotrophin (Murray-Rust *et al.*, 1993). All the TGF- β superfamily ligands except GDF3, GDF9, and BMP15 use an additional conserved cysteine to form an inter-molecular disulphide bond that stabilizes the dimer. Although most ligands appear to act as homodimers, ligand heterodimerization has also been demonstrated. Prominent examples are activin β A– β B heterodimers (Antsiferova and Werner, 2012), and heterodimers formed between BMP15 and GDF9 (Peng *et al.*, 2013) and between NODAL and GDF1 (Tanaka *et al.*, 2007). Furthermore, in zebrafish, BMP2 and BMP7 form obligate heterodimers which are required for signaling in the early embryo (Little and Mullins, 2009).

TGF- β superfamily ligands are all synthesized as precursors, with a large prodomain and a C-terminal mature domain. Dimerization is thought to occur intracellularly and depends on the presence of the prodomain (Gray and Mason, 1990). Cleavage of the mature domain from the prodomain is subsequently mediated by proteases of the subtilisin-like proprotein convertase (SPC) family (Constam and Robertson, 1999, 2000). Some of the ligands have multiple cleavage sites and recent data suggest that for BMP4 these cleavages must be sequential rather than simultaneous in order to preserve ligand function (Tilak *et al.*, 2014). The TGF- β s themselves, as well as GDF8 and the closely related GDF11, are secreted as latent forms still non-covalently attached to their propeptides (Ge *et al.*, 2005; Shi *et al.*, 2011). In the case of the TGF- β s the latent ligand forms a disulphide-linked complex with LTBP (latent TGF- β binding proteins) (Annes *et al.*, 2003), which associate with the extracellular matrix through interactions with fibrillins (Ramirez and Sakai, 2010). This regulates diffusion, storage, presentation, and release of the active TGF- β s. A further activation step is required to release the active ligand (Box 1). In the case of TGF- β 1 and TGF- β 3, this involves the binding of $\alpha_v\beta_6$ or $\alpha_v\beta_8$ integrins to RGD motifs in the prodomains of the ligands (Shi *et al.*, 2011), but other mechanisms involving proteases, thrombospondin 1, reactive oxygen species and low pH, can also activate TGF- β (Annes *et al.*, 2003). In the case of GDF11, ligand activation requires metalloproteinases of the BMP1/tolloid family (Ge *et al.*, 2005).

TGF- β Receptors

TGF- β superfamily ligands signal through two types of serine/threonine kinase receptors, a type I (TGFBR1, ACVRL1, ACVR1, ACVR1B, ACVR1C, BMPR1A, or BMPR1B) and a type II (ACVR2A, ACVR2B, TGFBR2, BMPR2, or AMHR2) (Shi and Massague, 2003; Table 1; Figure 1). Type II receptors consist of an extracellular ligand-binding domain, a membrane-spanning region and an internal, constitutively active, kinase domain. Type I receptors are closely related to the type II receptors and have a similar structural organization (Figure 1). They additionally contain a highly conserved intracellular glycine- and serine-rich (GS) domain, close to the membrane-spanning region. Receptors signal through a conserved mechanism. Ligand binding permits the type II receptor to

Box 1 Release of mature TGF- β from the latency complex

Mature TGF- β is released from the latency complex after the exertion of force through the binding of integrins. The mature TGF- β dimer is anchored in place by the 'straitjacket' region formed by a dimer of its prodomains, with the 'latency lasso' preventing the binding of ligand to the receptors and thus keeping ligand inactive. The prodomains themselves are held together in the arm region by 4 disulphide bonds forming a bowtie motif. The prodomains attach to $\alpha_v\beta_6$ or $\alpha_v\beta_8$ integrins at an RGD site in the arm region. Latent TGF- β interacts with the extracellular matrix by forming a disulphide-linked complex with LTBP, which itself interacts with fibrillins. The binding of integrins exerts a force on the complex that unfastens the straitjacket and releases mature ligand. Shown below is a linearized representation of the two full-length TGF- β monomers, indicating their structural organization. The elements are color-coded, to indicate which monomer is associated with which prodomain (Shi, M., Zhu, J., Wang, R., *et al.*, 2011. Latent TGF- β structure and activation. *Nature* 474, 343–349).

phosphorylate the type I receptor at multiple serines and threonines in its GS domain. This phosphorylation drives conformational changes that activate the type I receptor kinase and provide a binding site for its downstream substrates, the receptor-regulated SMADs (R-SMADs, see below) (Figure 1), which it phosphorylates and activates (Figure 3). Traditionally, the pathway has been divided into two main branches, with the activins, TGF- β s and NODAL leading to the phosphorylation of SMAD2 and 3, and the BMPs and GDFs leading to the phosphorylation of SMAD1, 5 and 9 (Shi and Massague, 2003). Although each type I receptor appears to activate certain SMADs preferentially, there is also evidence for the formation of mixed receptor complexes that result in a single ligand being able to activate both classes of R-SMAD (Daly *et al.*, 2008; Goumans *et al.*, 2002; Holtzhausen *et al.*, 2014).

For the canonical TGF- β superfamily ligands, activated receptor complexes consist of two type I and two type II receptors (Ehrlich *et al.*, 2012), which signal as a heterotetrameric complex (Huang *et al.*, 2011; Weis-Garcia and Massague, 1996). Although it was initially believed that complex formation was sequential, with ligand binding to the type II receptor bringing the type I receptor into the complex (Wrana *et al.*, 1994), it has also been suggested that complexes are preformed on the plasma membrane to some extent (Rechtman *et al.*, 2009; Wells *et al.*, 1999). In the case of the

TGF- β s themselves, the ligand must first bind the type II receptors to bring the type I receptors into the complex. In contrast, the BMPs show little preference for one receptor type over the other, suggesting a less sequential mechanism for receptor complex formation (Groppe *et al.*, 2008). In the inactive state, FKBP1A (FKBP12) is bound at the GS domain of type I receptors and inhibits signaling, with phosphorylation of the type I receptor releasing FKBP1A and relieving this inhibition (Wang and Donahoe, 2004).

The trafficking of receptors has been demonstrated to be critical for pathway regulation. Receptors are constitutively internalized by endocytosis in the absence and presence of ligand and can signal from endosomes (Di Guglielmo *et al.*, 2003). They have been described to either internalize uniquely into clathrin-coated pits (Mitchell *et al.*, 2004) or to additionally internalize into caveolae (Di Guglielmo *et al.*, 2003), but conflicting evidence exists over whether internalization is necessary for signaling, with different cell lines, ligands and methods for blocking internalization leading to varying conclusions (Di Guglielmo *et al.*, 2003; Hagemann *et al.*, 2009; Jullien and Gurdon, 2005; Lu *et al.*, 2002; Figure 4). ZFYVE9 (also known as Smad anchor for receptor activation (SARA)) has been proposed to associate with endosomes and to bring Smads into contact with activated receptors, which could enhance the ability of internalized receptors to signal

Table 1 Ligand–receptor–SMAD usage in TGF- β superfamily signaling

Ligand ^a	Type I receptor	Type II receptor	Co-receptors	Smads ^d
INHA ^b	No type I receptor	ACVR2A		
AMH ^c	ACVR1 BMPR1A	AMHR2		SMAD1 SMAD5 SMAD9
GDF15 ^d	Unknown	Unknown		
BMP15 ^e	BMPR1B	BMPR2		SMAD1 SMAD5 SMAD9
GDF9 ^e	ACVR1B	BMPR2		SMAD2 SMAD3
TGFB1 ^{f,g}	TGFB1 ACVRL1	TGFB2	TGFB3 ENG	SMAD2 SMAD3
TGFB2 ^{f,g}	TGFB1 ACVRL1	TGFB2	TGFB3 ENG	SMAD2 SMAD3
TGFB3 ^{f,g}	TGFB1 ACVRL1	TGFB2	TGFB3 ENG	SMAD2 SMAD3
GDF11 ^h	ACVR1B TGFB1	ACVR2A ACVR2B		SMAD2 SMAD3
GDF8 ^h	ACVR1B TGFB1	ACVR2A ACVR2B		SMAD2 SMAD3
INHBA ⁱ	ACVR1B	ACVR2A ACVR2B		SMAD2 SMAD3
INHBB ⁱ	ACVR1B ACVR1C	ACVR2A ACVR2B		SMAD2 SMAD3
INHBE ^j	No receptor	No receptor		
INHBC ^j	No receptor	No receptor		
NODAL ^{k,l}	ACVR1B ACVR1C	ACVR2A ACVR2B	TDGF1 CFC1	SMAD2 SMAD3
BMP3 ^m	No type I receptor	ACVR2B		
GDF10 ⁿ	ACVR1B	ACVR2A		SMAD2 SMAD3
GDF1 ^k	ACVR1B	ACVR2A ACVR2B	TDGF1 CFC1	SMAD2 SMAD3
GDF3 ^k	ACVR1B	ACVR2A ACVR2B	TDGF1 CFC1	SMAD2 SMAD3
BMP9 ^o	ACVRL1	ACVR2A ACVR2B BMPR2		SMAD1 SMAD5 SMAD9
BMP10 ^o	ACVRL1	ACVR2A ACVR2B BMPR2		SMAD1 SMAD5 SMAD9
GDF5 ^p	BMPR1A BMPR1B	ACVR2A ACVR2B BMPR2		SMAD1 SMAD5 SMAD9
GDF6 ^p	BMPR1A BMPR1B	ACVR2A ACVR2B BMPR2		SMAD1 SMAD5 SMAD9
GDF7 ^p	BMPR1A BMPR1B	ACVR2A ACVR2B BMPR2		SMAD1 SMAD5 SMAD9
BMP4 ^p	BMPR1A BMPR1B	ACVR2A ACVR2B BMPR2	RGMA RGMB HFE2 ^q	SMAD1 SMAD5 SMAD9
BMP2 ^p	BMPR1A BMPR1B	ACVR2A ACVR2B BMPR2	RGMA RGMB HFE2 ^q	SMAD1 SMAD5 SMAD9
BMP8 ^p	BMPR1A BMPR1B ACVR1	ACVR2A ACVR2B BMPR2		SMAD1 SMAD5 SMAD9
BMP7 ^p	BMPR1A BMPR1B ACVR1	ACVR2A ACVR2B BMPR2		SMAD1 SMAD5 SMAD9
BMP6 ^p	BMPR1A BMPR1B ACVR1	ACVR2A ACVR2B BMPR2		SMAD1 SMAD5 SMAD9
BMP5 ^p	BMPR1A BMPR1B ACVR1	ACVR2A ACVR2B BMPR2		SMAD1 SMAD5 SMAD9

^aThe order of the ligands in the Table is the same as in the phylogenetic tree in Figure 1. Shown are the best understood cognate receptors for each ligand, their requirement for co-receptor and the SMADs that have traditionally been thought to be activated by each receptor type. In some cases, alternative SMAD complexes can also form, such as the mixed SMAD complexes that can form in response to TGF- β 1 and 3 (SMAD1/5–SMAD3) and the recently demonstrated SMAD2/3 activation by BMP2 and 4. The following references were used to construct the Table:

^bAntsiferova, M., Werner, S., 2012. The bright and the dark sides of activin in wound healing and cancer. *Journal of Cell Science* 125, 3929–3937.

^cOrvis, G.D., Jamin, S.P., Kwan, K.M., *et al.*, 2008. Functional redundancy of TGF- β family type I receptors and receptor-Smads in mediating anti-Mullerian hormone-induced Mullerian duct regression in the mouse. *Biology of Reproduction* 78, 994–1001.

^dUnsicker, K., Spittau, B., Kriegstein, K., *et al.*, 2013. The multiple facets of the TGF- β family cytokine growth/differentiation factor-15/macrophage inhibitory cytokine-1. *Cytokine & Growth Factor Reviews* 24, 373–384.

^ePeng, J., Li, Q., Wigglesworth, K., *et al.*, 2013. Growth differentiation factor 9:bone morphogenetic protein 15 heterodimers are potent regulators of ovarian functions. *Proceedings of the National Academy of Sciences of the United States of America* 110, E776–E785.

^fLebrin, F., Goumans, M.J., Jonker, L., *et al.*, 2004. Endoglin promotes endothelial cell proliferation and TGF- β /ALK1 signal transduction. *EMBO Journal* 23, 4018–4028.

^gShi, Y., Massague, J., 2003. Mechanisms of TGF- β signaling from cell membrane to the nucleus. *Cell* 113, 685–700.

^hTsuchida, K., Nakatani, M., Uezumi, A., *et al.*, 2008. Signal transduction pathway through activin receptors as a therapeutic target of musculoskeletal diseases and cancer. *Endocrine Journal* 55, 11–21.

ⁱTsuchida, K., Nakatani, M., Yamakawa, N., *et al.*, 2004. Activin isoforms signal through type I receptor serine/threonine kinase ALK7. *Molecular and Cellular Endocrinology* 220, 59–65.

^jGold, E., Risbridger, G., 2012. Activins and activin antagonists in the prostate and prostate cancer. *Molecular and Cellular Endocrinology* 359, 107–112.

^kShen, M.M., 2007. Nodal signaling: Developmental roles and regulation. *Development* 134, 1023–1034.

^lNadeem, L., Munir, S., Fu, G., *et al.*, 2011. Nodal signals through activin receptor-like kinase 7 to inhibit trophoblast migration and invasion: Implication in the pathogenesis of preeclampsia. *American Journal of Pathology* 178, 1177–1189.

^mKokabu, S., Garner, L., Cox, K., *et al.*, 2012. BMP3 suppresses osteoblast differentiation of bone marrow stromal cells via interaction with Acvr2b. *Molecular Endocrinology* 26, 87–94.

ⁿMatsumoto, Y., Otsuka, F., Hino, J., *et al.*, 2012. Bone morphogenetic protein-3b (BMP-3b) inhibits osteoblast differentiation via Smad2/3 pathway by counteracting Smad1/5/8 signaling. *Molecular and Cellular Endocrinology* 350, 78–86.

^oTownson, S.A., Martinez-Hackert, E., Greppi, C., *et al.*, 2012. Specificity and structure of a high affinity activin receptor-like kinase 1 (ALK1) signaling complex. *Journal of Biological Chemistry* 287, 27313–27325.

^pMueller, T.D., Nickel, J., 2012. Promiscuity and specificity in BMP receptor activation. *FEBS Letters* 586, 1846–1859.

^qCorradini, E., Babitt, J.L., Lin, H.Y., 2009. The RGM/DRAGON family of BMP co-receptors. *Cytokine & Growth Factor Reviews* 20, 389–398.

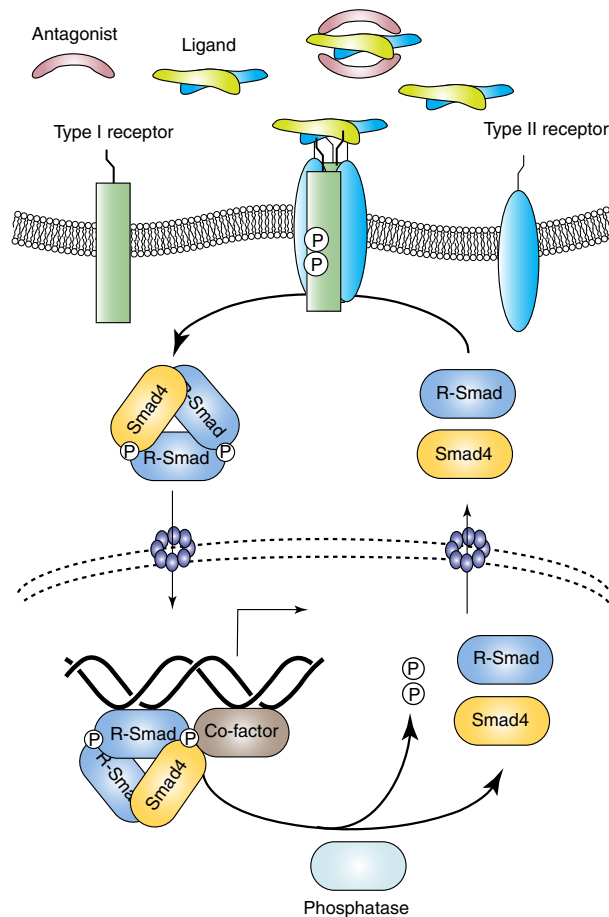


Figure 3 Schematic representation of the generic TGF- β /SMAD pathway. Binding of ligand dimers leads to the phosphorylation of type I receptors by type II receptors and formation of a heterotetrameric receptor complex. This allows the type I receptor to phosphorylate R-SMAD transcription factors, which form complexes with SMAD4 and accumulate in the nucleus, where they regulate gene transcription with transcriptional cofactors. The action of the receptor kinase is opposed by a putative nuclear phosphatase. Ligand activity is also regulated by binding of ligand antagonists.

(Di Guglielmo *et al.*, 2003; Tsukazaki *et al.*, 1998). ZFYVE16 (endofin), a ZFYVE9 homolog, has been suggested to play an analogous role in BMP signaling (Chen *et al.*, 2007). Once internalized, receptors are either targeted for degradation in the proteasome or lysosome or are recycled back to the plasma membrane in a RAB11 dependent manner (Mitchell *et al.*, 2004; Figure 4). Resolving the exact trafficking routes of receptors has proved difficult due to a lack of reagents to localize endogenous receptors. To date, there are no antibodies that can reliably be used for immunofluorescence studies of any TGF- β superfamily receptor and it is not certain that over-expression systems fully recapitulate the behavior of the endogenous signaling pathway.

TGF- β superfamily receptors can be posttranslationally modified to regulate their localization and activity. These modifications can modulate receptor trafficking, endocytosis, recycling and degradation, and thus determine the

duration of the phospho-SMAD response. Both receptor types are glycosylated during trafficking to the membrane, which may regulate their subsequent turnover (Kim *et al.*, 2012; Wells *et al.*, 1997). In addition, ubiquitination of receptors plays a key role in regulating their stability and abundance (Figure 4). The canonical TGF- β type I receptor TGFBR1 is degraded following ubiquitination by an E3 ubiquitin ligase. Several E3 ligases have been proposed to be involved in this process, including SMURF1 (Kavak *et al.*, 2000), SMURF2 (Ogunjimi *et al.*, 2005), WWP1 (Komuro *et al.*, 2004), and NEDD4L (Kuratori *et al.*, 2005). Deubiquitinating enzymes have also been identified that reverse this process and increase receptor stability such as USP15 (Eichhorn *et al.*, 2012), UCHL5 (UCH37) (Wicks *et al.*, 2005), and USP4 (Zhang *et al.*, 2012b). Recently, TRAF4 has been identified as antagonizing the inhibitory effects of SMURF2 by recruiting USP15 to activated TGF- β type I receptors (Zhang *et al.*, 2013). In addition, TGFBR1 is sumoylated in response to TGF- β , a modification that stabilizes SMAD3 binding to the receptor and promotes signaling (Kang *et al.*, 2008). Receptor regulation is also substantially influenced by the inhibitory SMADs (I-SMADs, see below), SMAD6 and SMAD7. SMAD7 is recruited to activated receptors and can in turn recruit negative regulators of receptor signaling, including SMURF2 (Kavak *et al.*, 2000) and the GADD34-PP1c phosphatase complex (Shi *et al.*, 2004).

Several ligand–receptor interactions also require membrane-bound co-receptors for full SMAD activation. The first of these to be identified was betaglycan, which was initially considered to be a third core receptor type (hence its official name, TGFBR3). This co-receptor can enhance canonical TGF- β signaling by presenting ligand to the type II receptor (Lopez-Casillas *et al.*, 1993), and is particularly important for TGF- β 2 signaling (ten Dijke and Arthur, 2007). It can also inhibit signaling, possibly by shedding its ectodomain, releasing a soluble form of betaglycan that is capable of sequestering ligand (Elderbrook *et al.*, 2014). Shortly after the identification of betaglycan, ENG (endoglin) was isolated as a TGF- β co-receptor in endothelial cells. In these cells ENG enhances TGF- β binding to ACVRL1 and promotes phosphorylation of SMAD1/5, while inhibiting signaling through TGFBR1 and SMAD2/3, promoting cell proliferation and migration (Lebrin *et al.*, 2004).

TGF- β superfamily co-receptors also play critical roles in embryonic development. TDGF1 (also known as cripto) is an EGF-CFC (epidermal growth factor-like cripto frl-1 cryptic) protein that was first identified as having a role in TGF- β superfamily signaling when a mutant in the zebrafish homolog, one-eyed pinhead (*oep*), was shown to have the same phenotype as loss of NODAL activity (Schier, 2003). It has since been shown to act as an obligate co-receptor for NODAL/GDF1/GDF3 signaling through ACVR1B and SMAD2/3, although its mechanism of action is still in debate. It has been proposed to act either by binding ligand and presenting it to receptor complexes, by directing trafficking of ligand or activated receptors or by directing the cleavage and activation of NODAL (Constam, 2009). A second family member, CFC1 (also known as cryptic) also exists, which is important for NODAL function in left–right patterning (Schier, 2003).

Co-receptors for BMP signaling have only been identified much more recently, and fall into a single family, the repulsive

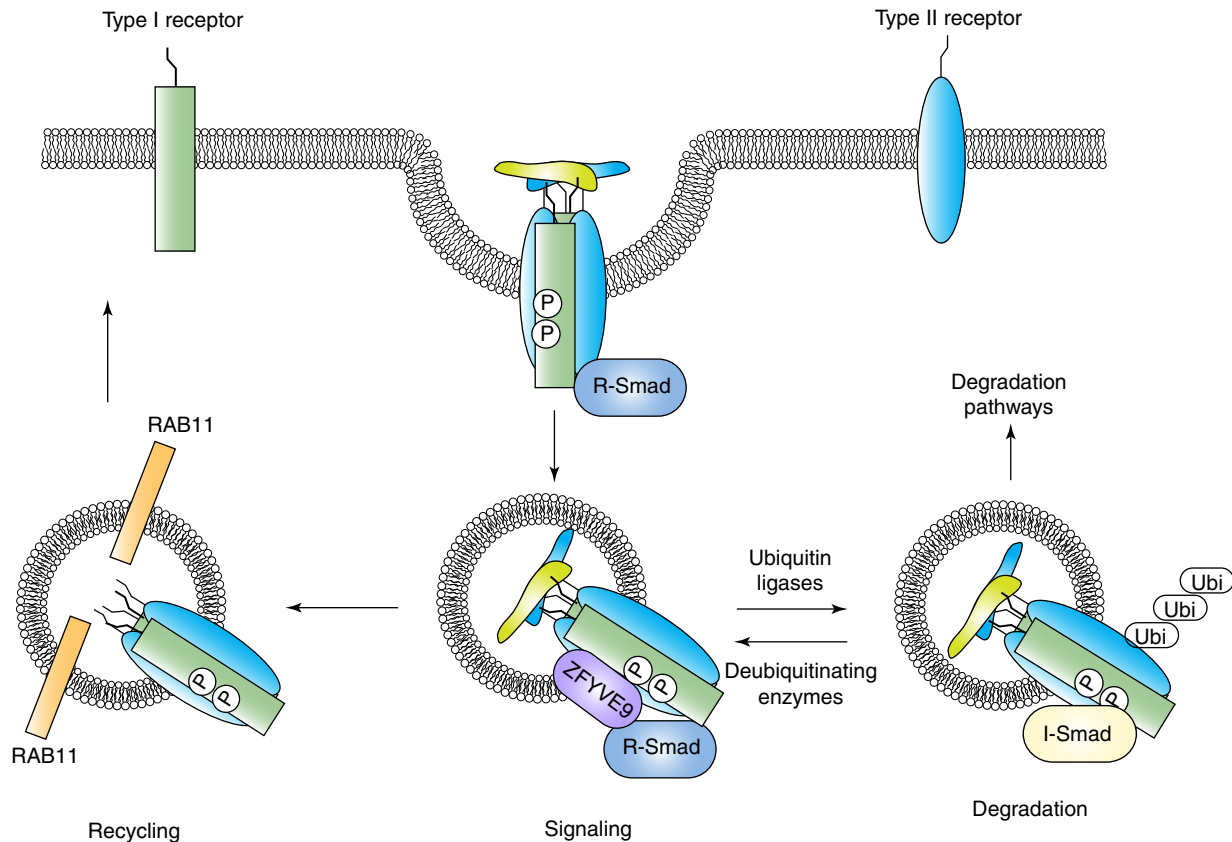


Figure 4 Receptor endocytosis following ligand stimulation. After binding ligand, receptors undergo endocytosis and continue to signal from endosomes. Signaling from endosomes may be enhanced by the presence of factors such as ZFYVE9 (SARA). Receptors could be subsequently either recycled back to the plasma membrane in RAB11-marked vesicles or poly-ubiquitinated and targeted for degradation. I-SMADs recruit inhibitory factors to activated receptors.

guidance molecule (RGM) or DRAGON family (Corradini *et al.*, 2009). These molecules may enhance signaling by allowing BMP2 and BMP4 to signal through ACVR2A in addition to BMPR2, though little is known about the precise molecular mechanism.

Co-receptors thus increase the specificity of ligand–receptor–SMAD interactions, and allow greater regulation of signaling in different cellular contexts. The co-receptors have also been shown to have important roles in the development of cancer and other diseases (ten Dijke and Arthur, 2007) (see also below).

The SMADs

SMAD transcription factors are the principal downstream effectors of TGF- β superfamily signaling. They fall into three subclasses, the receptor-regulated SMADs (R-SMADs), SMAD1, 2, 3, 5, and 9, the common mediator SMAD (co-SMAD) SMAD4, and the inhibitory SMADs (I-SMADs), SMAD6 and 7 (Figures 1 and 3; Table 1).

The R-SMADs are highly conserved across all metazoans. They contain three key structural features: an N-terminal Mad Homology 1 (MH1) domain, a C-terminal Mad Homology 2 (MH2) domain, and a less conserved proline-rich region

linking these two domains. The MH2 domain is the critical determinant of the interactions of the SMADs with other proteins, including receptors, other SMADs, co-activators and transcriptional cofactors, whereas the MH1 domain regulates nuclear import and in the case of all R-SMADs except SMAD2, can contact DNA. In addition, the MH1 domain can inhibit the activity of the MH2 domain when the SMADs are in an inactive state (Shi and Massague, 2003). Activated SMADs are phosphorylated by TGF- β superfamily type I receptors at two serine residues at an S-M/V-S motif at their extreme C-terminus. Once phosphorylated, activated R-SMADs form complexes with SMAD4, which accumulate in the nucleus where they activate or repress transcription. The exact stoichiometry of these complexes is still unclear, but they are likely to be dimeric or trimeric (with two R-SMADs and SMAD4). In addition, stimulation with TGF- β ligand can drive the formation of mixed SMAD complexes comprising phosphorylated SMAD1/5-SMAD3 that inhibit BMP responses (Gronroos *et al.*, 2012).

Pathway activation through the receptors is counteracted by the activity of nuclear phosphatases, which terminate signaling by either dissociating SMAD complexes or increasing their nuclear export rate. Several phosphatases have been identified that can dephosphorylate SMAD2/3 or SMAD1/5, including PPM1A, PP2A, MTMR4, and Pdp (Bruce and Sapkota, 2012).

It is unclear if any one of these is single-handedly responsible for terminating SMAD signaling or whether they act in combination. A proportion of activated SMADs are targeted for degradation in the proteasome. In one example of this, NEDD4L specifically recognizes TGF- β -induced phosphorylation of a threonine residue preceding a so-called PY motif in the SMAD2/3 linker, leading to SMAD polyubiquitination and degradation, thus limiting the full extent of TGF- β signaling (Gao *et al.*, 2009).

Once in the nucleus, SMADs can regulate transcription both positively and negatively. SMAD4 and all R-SMADs except SMAD2 bind DNA directly, with SMAD3 and 4 binding at SMAD binding elements (SBEs), comprising a repeated CAGAC (or its reverse complement, GTCTG) motif, while SMADs 1 and 5 bind at GC-rich GCCGNC or GRCGNC sequences (Ross and Hill, 2008). The SMADs have a low affinity for DNA, so transcriptional regulation relies heavily on the binding of co-factors that increase the specificity and affinity of SMADs for certain gene promoters and enhancers. The first of these to be identified was FOXH1, which was shown to be required for activin-mediated induction of the *Xenopus Mix.2* gene (Chen *et al.*, 1996), but has since been shown to be a general transcriptional co-activator of SMAD2 target genes during early embryonic development. Other transcriptional co-factors have subsequently been identified which can be either activatory or repressive, and cooperate with different SMADs (Ross and Hill, 2008). The requirement for these co-factors depends on both the gene under regulation and the cellular context.

SMAD binding leads to chromatin remodeling via changes in the acetylation and methylation state of histones surrounding SMAD binding sites (Ross *et al.*, 2006). The SMADs recruit histone acetyl transferases, such as p300 and CBP (Feng *et al.*, 1998) that acetylate key lysines in histone tails and promote transcriptional activation. Additionally, in response to TGF- β /activin signaling, SMAD2/3 can recruit the SWI/SNF nucleosome remodeling component SMARCA4 (BRG1) to create a favorable chromatin environment for other transcription factors to bind (Xi *et al.*, 2008). In different contexts, SMAD binding partners can also recruit histone deacetylases (HDACs) to inhibit gene expression (Kang *et al.*, 2005).

The I-SMADs, SMAD6 and 7, negatively regulate TGF- β signaling. As well as their role in recruiting negative regulators to receptors (see above), SMAD6 and 7 have been described to interfere with SMAD complex formation (Hata *et al.*, 1998; Moren *et al.*, 2005), the ability of SMADs to bind DNA and the activation of transcription by nuclear SMADs (Lin *et al.*, 2003).

SMAD activity is also regulated by several posttranslational modifications in addition to phosphorylation of the canonical S-M/V-S motif. MAP kinases (MAPKs) and cyclin-dependent kinases (CDKs) can phosphorylate R-SMADs at Ser/Thr-Pro motifs in the linker region (Massague, 2012). Linker phosphorylation has been shown to lead to the degradation of activated SMADs by recruiting E3 ubiquitin ligases which polyubiquitinate SMADs. SMAD1/5 linker phosphorylation by MAPKs and GSK3 recruits the E3 ligase SMURF1 (Sapkota *et al.*, 2007), while SMAD2/3 is ubiquitinated in an analogous fashion by NEDD4L (Gao *et al.*, 2009). The deubiquitinase USP15 was recently identified as being able to counteract these

processes and stabilize activated SMADs (Inui *et al.*, 2011). Chromatin-bound SMAD4 is regulated through mono-ubiquitination by TRIM33 (TIF1 γ), which acts to reduce the residence time of SMAD complexes on DNA (Agricola *et al.*, 2011). In addition, R-SMADs have been shown to be regulated by acetylation (Simonsson *et al.*, 2006), PARylation (Lonn *et al.*, 2010), and sumoylation (Lee *et al.*, 2003), highlighting the critical importance of controlling SMAD levels and localization within the cell.

The I-SMADs are also subject to regulation through post-translational modifications. The RING domain-containing E3 ligase RNF12 binds to SMAD7, inducing its polyubiquitination and degradation (Zhang *et al.*, 2012a). SMAD6 has been shown to be regulated by arginine methylation via the methyltransferase PRMT1 (Xu *et al.*, 2013). This methylation leads to relocalization of SMAD6 away from receptors, facilitating R-SMAD recruitment and activation.

Pathway Modulation

TGF- β signaling is subject to complex regulation, as might be expected from a pathway whose activation and inactivation at the right time and place is critical for the normal development of metazoans. Regulation occurs at all levels of the pathway, from ligands to SMADs, but the mechanisms that elicit this regulation vary in different cellular contexts and between TGF- β superfamily pathways. Although many positive and negative regulators of TGF- β signaling have been identified, common principles underlie their mechanism of action, which will be described below. These regulators do not necessarily lead to simple activation or suppression of the pathway as a whole, but instead act like a rheostat, where multiple positive and negative inputs favor more or less pathway activity and thus modulate signaling.

Ligand activity is regulated both by the release of ligand from inactive conformations and by secreted inhibitors. Gradients of ligand activity have long been understood to be involved in the patterning of embryos during development, and are often regulated by opposing gradients of antagonists. These antagonists generally function by binding and sequestering ligands, as is the case for BMP antagonists such as NOG (noggin) and CHRD (chordin) or the activin antagonist FST (follistatin) (Figure 1). One exception to this is the inhibition of NODAL signaling by LEFTY1/2. LEFTY1/2 are secreted antagonists necessary for normal mesendoderm induction and left-right patterning in vertebrates (Shiratori and Hamada, 2014). As members of the TGF- β superfamily, their sequence is similar to the ligands themselves, so they are likely to mediate ligand antagonism by a different mechanism, such as sequestration of the co-receptor TDGF1 (Chen and Shen, 2004; Cheng *et al.*, 2004). Because LEFTY1/2 are targets of NODAL signaling, they can generate a negative feedback loop in developmental contexts, where pathway activity is shut off when signaling is sustained (Schier, 2003). In addition to these negative regulators of TGF- β signaling, several ligand agonists have been identified that enhance the ability of ligand to activate receptors. One example of this is BMPER (crossveinless 2), which acts in developmental contexts to enhance BMP activity in specific spatial domains (Reichert *et al.*, 2013; Serpe *et al.*, 2008).

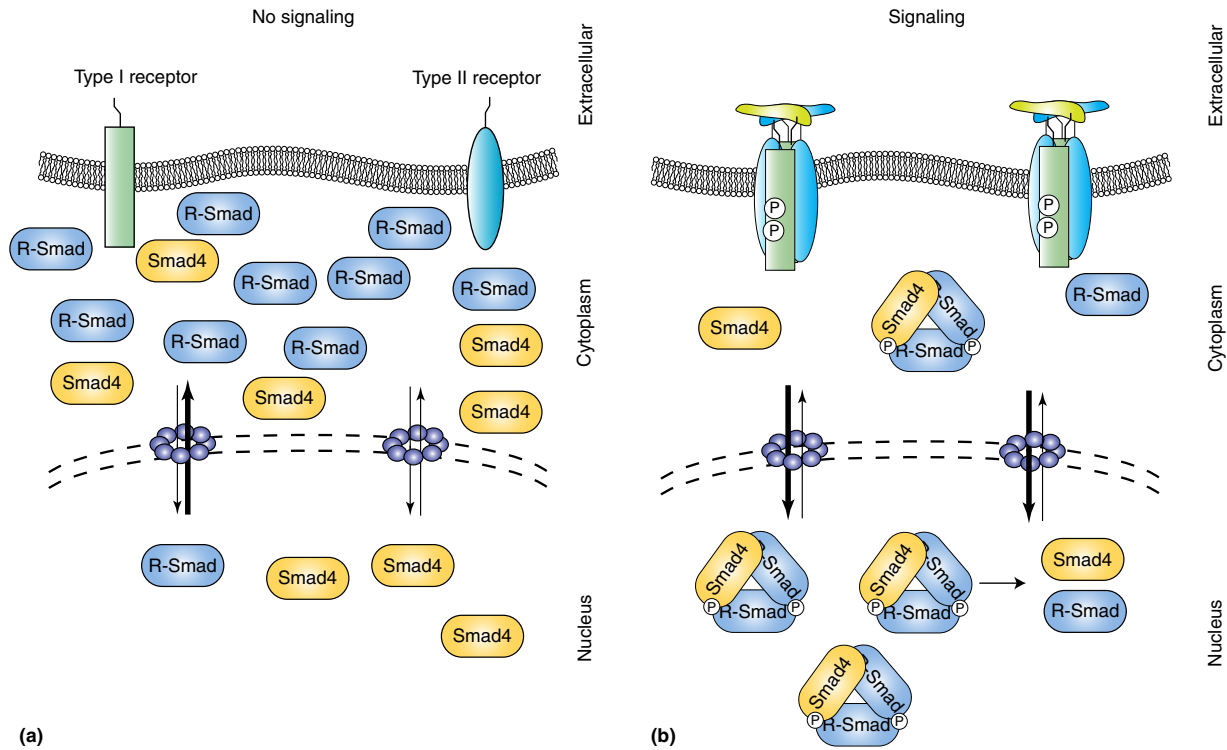


Figure 5 SMAD nucleocytoplasmic shuttling allows continuous monitoring of receptor activity. (a) In the absence of signal, SMADs shuttle into and out of the nucleus at a continuous rate. Monomeric R-SMADs have a higher nuclear export rate than import rate so appear predominantly located in the cytoplasm, while SMAD4 is distributed evenly between the cytoplasm and nucleus. (b) When receptors are signaling, the SMADs form complexes that have a higher nuclear import rate than the monomeric SMADs and cannot be exported until dissociated. Thus, the overall nuclear import rate exceeds the export rate and this drives SMAD accumulation in the nucleus.

Recent work has indicated that the level of activated receptors does not necessarily correlate with levels of ligand in the extracellular milieu. After acute TGF- β stimulation, receptors are rapidly depleted from the surface of cells, although signaling is sustained for several hours after this point (Vizan *et al.*, 2013). This renders the cells refractory to further acute stimulation, until the ligand is depleted and surface receptor levels recover. Sustained signaling from internalized receptors in the absence of extracellular ligand has also been observed in *Xenopus* embryos responding to activin (Jullien and Gurdon, 2005). In embryonic contexts, cells have to interpret a transient presence of ligands and translate that into a sustained response. *Xenopus* embryos appear to have a long-term cellular 'memory' for ligand, with activated receptors signaling for several hours after ligand has disappeared.

Although pathway activation is not a direct corollary of ligand levels, continuous monitoring of receptor activity does occur. The R-SMADs and SMAD4 continuously shuttle into and out of the nucleus, even in the absence of pathway activation (Inman *et al.*, 2002; Xu *et al.*, 2002). Receptor activity drives the formation of complexes, which are prevented from being exported from the nucleus, and have a higher nuclear import rate than monomeric SMADs and so accumulate in the nucleus upon ligand induction (Schmierer *et al.*, 2008; Figure 5; Movies 1 and 2). When receptors are turned off, the SMADs gradually reaccumulate in the cytoplasm. At any time, the level of activated SMAD complexes in the nucleus will be a

direct result of the relative activity of the receptor kinase and the nuclear phosphatase. This dynamic nucleocytoplasmic shuttling allows cells to continuously monitor pathway activation, with levels of nuclear SMADs directly reflecting receptor activity.

Pathway activation can lead to the initiation of both positive and negative feedback cycles in different contexts. These can act at the transcriptional level or feed back onto the core TGF- β pathway components. An example of a transcriptional feedforward loop involves the homeodomain transcription factor Mixer, which regulates the expression of key NODAL target genes in development, such as *gooseoid*, by acting as transcriptional co-factor for SMAD2-SMAD4 complexes (Germain *et al.*, 2000; Kunwar *et al.*, 2003). Mixer is itself a transcriptional target of NODAL signaling, so forms part of a positive feedback loop to upregulate developmental genes, and allows amplification of an initial NODAL signal. Another transcriptional feedback loop with a similar mechanism, but the opposite outcome, was termed the self-enabling response (Kang *et al.*, 2003). Here, activated SMAD3 directly induces the transcription of *ATF3*, which encodes a transcriptional repressor. This then enables a secondary response where SMAD3 co-operates with ATF3 to down-regulate the expression of a gene encoding a key inhibitor of differentiation, *ID1*.

Negative feedback loops can act on pathway components themselves to limit the overall level of signaling.

Negative regulators of the pathway, including the previously described I-SMADs and SMURF1/2, as well as transcriptional co-repressors such as c-SKI and SKIL (also known as SnoN) can be direct targets of TGF- β superfamily signaling. Feedback loops can also act together to generate transient responses to pathway activation. A well-understood example of this involves the E3 ubiquitin ligase RNF111 (also known as Arkadia) and SKIL (Levy *et al.*, 2007). SKIL binds with SMAD4 in the inactive state to recruit transcriptional co-repressors to DNA at SBEs. Upon pathway activation, RNF111 bound to phosphorylated SMAD2/3 can ubiquitinate and promote degradation of SKIL, allowing transcription of the target genes via activated SMAD complexes. One of these target genes is *SKIL* itself, which accumulates in response to signaling and represses further transcriptional activation by the pathway. Recently, several microRNAs (miRNAs) have been described that act to regulate TGF- β superfamily signaling in mammalian cells, by downregulating receptors (Martello *et al.*, 2007) or ligands (Choi *et al.*, 2007). The production of these miRNAs can also be regulated in response to pathway activation, so can contribute to feedback loops (Kang *et al.*, 2012).

TGF- β Superfamily Signaling and Disease

Consistent with the importance of TGF- β superfamily signaling in embryonic development and tissue homeostasis, mutation or loss of core components of the pathways results in serious human diseases that affect many different organs of the body (Figure 6). One of the earliest to be identified was Hereditary Hemorrhagic Telangiectasia (HHT), which is characterized by bleeding from small vascular lesions in the mucocutaneous tissues, and arteriovenous malformations in the lung, liver and/or cerebral vasculature (ten Dijke and Arthur, 2007). HHT results from heterozygous mutations either in the TGF- β co-receptor *ENG* that lead to haploinsufficiency, or loss-of-function or dominant negative mutations in the type I receptor *ACVRL1* (Johnson *et al.*, 1996; McAllister *et al.*, 1994). This can readily be rationalized as *ACVRL1* and *ENG*, along with *TGFBR1* and *TGFBR2* are known to transduce TGF- β responses in endothelial cells (ten Dijke and Arthur, 2007). Mutation of other key components of TGF- β superfamily pathways also affect the cardiovascular system. One of the main features of the connective tissue disorders Marfan syndrome (MFS) and Loeys-Dietz syndrome (LDS), is a strong predisposition for

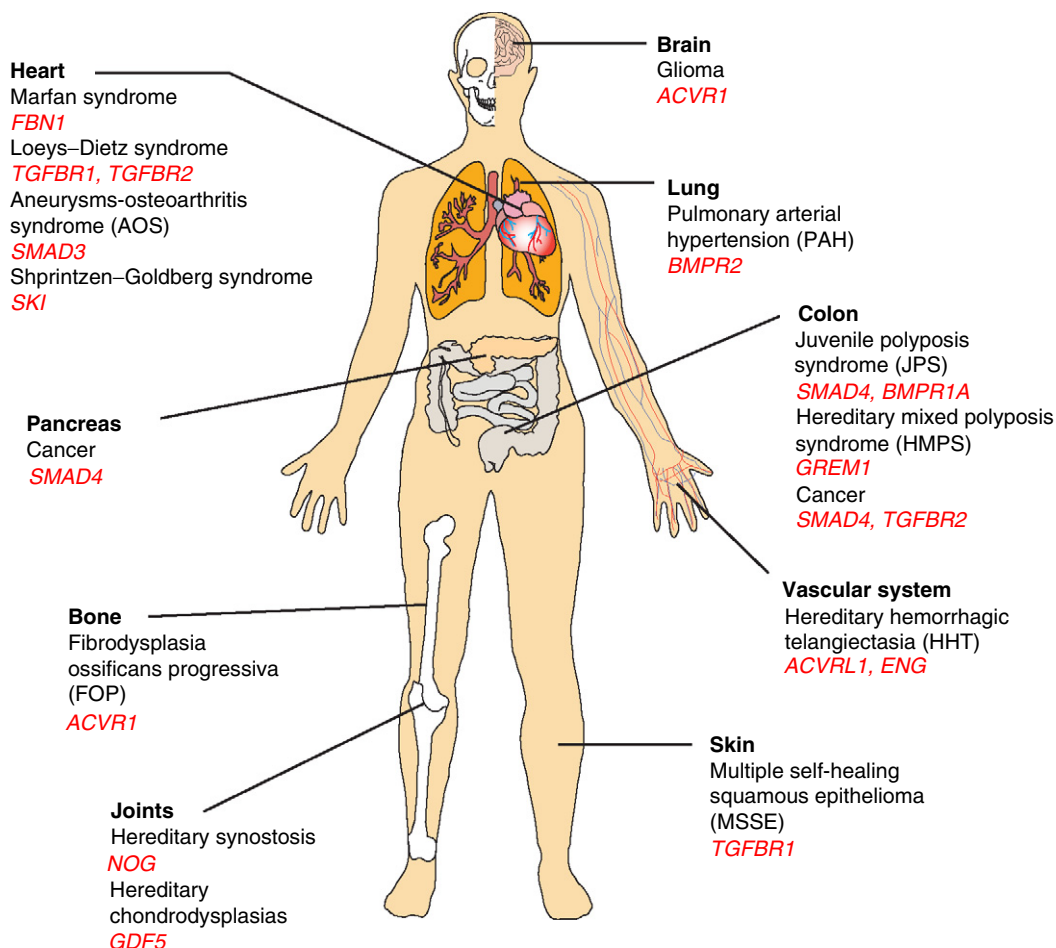


Figure 6 Mutations in TGF- β superfamily signaling components underlie a diverse range of human diseases. Indicated are the diseases that are best understood to commonly be driven by mutations in TGF- β signaling components, and the main organ or system they affect. The most common components to be mutated for each condition are highlighted in red.

aortic root aneurysms, along with long bone overgrowth, spine curvature and ocular lens dislocation, among other manifestations (Doyle *et al.*, 2012b). MFS is associated with heterozygous mutations in *FBN1* (fibrillin-1), whereas the related LDS results from heterozygous missense mutations in either *TGFBR1* or *TGFBR2*. Two other syndromes with significant overlap with MFS and LDS are Shprintzen–Goldberg syndrome (Doyle *et al.*, 2012a) and Aneurysms-osteoarthritis syndrome (van de Laar *et al.*, 2012), which are caused by heterozygous mutations in *SKI* and *SMAD3* respectively. Deregulated TGF- β signaling appears to be the underlying cause of all of these syndromes, but exactly how this results in altered extracellular matrix homeostasis and subsequent disease is not fully understood (Doyle *et al.*, 2012b).

Familial pulmonary hypertension is a lung disease characterized by vascular cell hyperproliferation resulting in obliteration of small pulmonary arteries, which leads to severe pulmonary hypertension and right-sided heart failure. It is associated with heterozygous loss-of-function mutations in the BMP type II receptor gene *BMPR2* (International PPH Consortium *et al.*, 2000). This is thought to lead to reduced BMP signaling in pulmonary artery smooth muscle cells and endothelial cells, causing vascular smooth muscle cell hyperproliferation. Interestingly, many of the disease-causing mutations lead to a truncation of the long cytoplasmic tail of BMPR2 protein, which is required for proper regulation of the cytoskeleton (ten Dijke and Arthur, 2007).

Germline mutations in TGF- β superfamily signaling components can also lead to cancer predisposition. The first example of this was the finding that heterozygous mutations of *SMAD4* that cause C-terminal truncations and loss of *SMAD4* function led to juvenile polyposis syndrome (JPS), which is characterized by the appearance of multiple juvenile polyps in the gastrointestinal tract, and increased risk of adenocarcinoma (Howe *et al.*, 1998). It is likely that it is BMP signaling that is affected, as other families with JPS have germline loss-of-function mutations in the BMP type I receptor *BMPR1A*. Recently, patients with a variant of JPS called hereditary mixed polyposis syndrome (HMPS) were found to have a duplication of the regulatory regions of the *GREM1* gene which encodes a BMP antagonist. The duplication results in increased expression of *GREM1*, which would be predicted to cause reduced BMP signaling (Jaeger *et al.*, 2012). Monoallelic mutations in the TGF- β type I receptor *TGFBR1* that result in truncation of the kinase domain or mutation of the ligand binding domain lead to multiple self-healing squamous epithelioma (MSSE) which is an autosomal-dominant skin cancer condition (Goudie *et al.*, 2011). These mutations are clearly loss-of-function and are entirely distinct from the missense mutations in *TGFBR1* that lead to LDS. Very recently, activating mutations in the gene encoding *ACVR1* have been associated with diffuse intrinsic pontine glioma (DIPG), which is an incurable pediatric high grade glioma (Buczkowicz *et al.*, 2014; Taylor *et al.*, 2014; Wu *et al.*, 2014). There are also numerous examples of somatic mutations in TGF- β superfamily components that result in cancer, highlighting the tumor suppressive activity of TGF- β superfamily pathways; the most prevalent being loss-of-function mutations in *SMAD4* in a high proportion of pancreatic and colon cancers, and loss-of-function mutations in *TGFBR2* in colon cancer (Levy and Hill,

2006; Wakefield and Hill, 2013). TGF- β superfamily signaling also has potent tumor promoting effects at later stages of cancer.

Intriguingly, the somatic mutations in *ACVR1* that result in DIPG give rise to a very different disease when present in the germline. This disease is the congenital childhood developmental disorder fibrodysplasia ossificans progressiva (FOP), where fibrous tissue such as skeletal muscle is transformed to bone spontaneously or when damaged (Kaplan *et al.*, 2011). Mutation of two other components of BMP/GDF signaling cause heritable developmental disorders affecting the joints. Heterozygous missense mutations in the gene encoding the BMP antagonist *NOG* causes the autosomal-dominant disorders proximal symphalangism and multiple synostoses syndrome (Gong *et al.*, 1999). In the former, osseous union results in fusions of joints in the hands and feet. In the latter, joints of the hip and cervical spine are affected. Mutations in *GDF5* result in human hereditary chondrodysplasias, which are characterized by shortening of limbs and loss of one or more joints (Luyten, 1997).

The range of diseases caused by misregulated TGF- β superfamily signaling reflects well the pleiotropic effects of these pathways in human development and homeostasis. However, the connection between the deregulated signaling and the disease manifestation in many cases is not yet clear, and will require considerable further work. The next challenge is to use our knowledge of the misregulated pathway concerned to inform the development of specific therapeutics.

Acknowledgments

We thank all the members of the Hill lab for discussions and useful comments on the manuscript. We thank Dr Bernhard Schmierer for the movies of *SMAD* nucleocytoplasmic shuttling. Work in the Hill lab is funded by Cancer Research UK, the Wates Foundation and the European Commission Network of Excellence EpiGeneSys [HEALTH-F4-2010-257082].

See also: Dynamic Integration: Dynamics: Dynamics of Protein Kinase Cascades. Horizontal Integration: Omics: Phosphoproteomics: Approaches, Developments, and Challenges; The Advent of Mass Spectrometry-Based Proteomics in Systems Biology Research

References

- Adamska, M., Degnan, S.M., Green, K.M., *et al.*, 2007. Wnt and TGF- β expression in the sponge *Amphimedon queenslandica* and the origin of metazoan embryonic patterning. *PLoS One* 2, e1031.
- Agricola, E., Randall, R.A., Gaarenstroom, T., Dupont, S., Hill, C.S., 2011. Recruitment of TIF1 γ to chromatin via its PHD finger-bromodomain activates its ubiquitin ligase and transcriptional repressor activities. *Molecular Cell* 43, 85–96.
- Annes, J.P., Munger, J.S., Rifkin, D.B., 2003. Making sense of latent TGF β activation. *Journal of Cell Science* 116, 217–224.
- Antsiferova, M., Werner, S., 2012. The bright and the dark sides of actinin in wound healing and cancer. *Journal of Cell Science* 125, 3929–3937.

- Arnold, S.J., Robertson, E.J., 2009. Making a commitment: Cell lineage allocation and axis patterning in the early mouse embryo. *Nature Reviews Molecular Cell Biology* 10, 91–103.
- Bruce, D.L., Sapkota, G.P., 2012. Phosphatases in SMAD regulation. *FEBS Letters* 586, 1897–1905.
- Buczkowicz, P., Hoeman, C., Rakopoulos, P., *et al.*, 2014. Genomic analysis of diffuse intrinsic pontine gliomas identifies three molecular subgroups and recurrent activating ACVR1 mutations. *Nature Genetics* 46, 451–456.
- Chen, C., Shen, M.M., 2004. Two modes by which Lefty proteins inhibit nodal signaling. *Current Biology* 14, 618–624.
- Chen, X., Rubock, M.J., Whitman, M., 1996. A transcriptional partner for MAD proteins in TGF- β signalling. *Nature* 383, 691–696.
- Chen, Y.G., Wang, Z., Ma, J., Zhang, L., Lu, Z., 2007. Endofin, a FYVE domain protein, interacts with Smad4 and facilitates transforming growth factor- β signaling. *Journal of Biological Chemistry* 282, 9688–9695.
- Cheng, S.K., Olale, F., Brivanlou, A.H., Schier, A.F., 2004. Lefty blocks a subset of TGF β signals by antagonizing EGF-CFC coreceptors. *PLoS Biology* 2, E30.
- Choi, W.Y., Giraldez, A.J., Schier, A.F., 2007. Target protectors reveal dampening and balancing of Nodal agonist and antagonist by miR-430. *Science* 318, 271–274.
- Constam, D.B., 2009. Riding shotgun: A dual role for the epidermal growth factor-Cripto/FRL-1/Cryptic protein Cripto in Nodal trafficking. *Traffic* 10, 783–791.
- Constam, D.B., Robertson, E.J., 1999. Regulation of bone morphogenetic protein activity by pro domains and proprotein convertases. *Journal of Cell Biology* 144, 139–149.
- Constam, D.B., Robertson, E.J., 2000. SPC4/PACE4 regulates a TGF β signaling network during axis formation. *Genes & Development* 14, 1146–1155.
- Corradini, E., Babitt, J.L., Lin, H.Y., 2009. The RGM/DRAGON family of BMP co-receptors. *Cytokine & Growth Factor Reviews* 20, 389–398.
- Daly, A.C., Randall, R.A., Hill, C.S., 2008. Transforming growth factor β -induced Smad1/5 phosphorylation in epithelial cells is mediated by novel receptor complexes and is essential for anchorage-independent growth. *Molecular and Cellular Biology* 28, 6889–6902.
- De Robertis, E.M., Kuroda, H., 2004. Dorsal-ventral patterning and neural induction in *Xenopus* embryos. *Annual Review of Cell and Developmental Biology* 20, 285–308.
- Di Guglielmo, G.M., Le Roy, C., Goodfellow, A.F., Wrana, J.L., 2003. Distinct endocytic pathways regulate TGF- β receptor signalling and turnover. *Nature Cell Biology* 5, 410–421.
- ten Dijke, P., Arthur, H.M., 2007. Extracellular control of TGF β signalling in vascular development and disease. *Nature Reviews Molecular Cell Biology* 8, 857–869.
- Doyle, A.J., Doyle, J.J., Bessling, S.L., *et al.*, 2012a. Mutations in the TGF- β repressor SKI cause Shprintzen-Goldberg syndrome with aortic aneurysm. *Nature Genetics* 44, 1249–1254.
- Doyle, J.J., Gerber, E.E., Dietz, H.C., 2012b. Matrix-dependent perturbation of TGF β signaling and disease. *FEBS Letters* 586, 2003–2015.
- Duboc, V., Lapraz, F., Saudemont, A., *et al.*, 2010. Nodal and BMP2/4 pattern the mesoderm and endoderm during development of the sea urchin embryo. *Development* 137, 223–235.
- Durlinger, A.L., Visser, J.A., Themmen, A.P., 2002. Regulation of ovarian function: The role of anti-Müllerian hormone. *Reproduction* 124, 601–609.
- Ehrlich, M., Gutman, O., Knaus, P., Henis, Y.I., 2012. Oligomeric interactions of TGF- β and BMP receptors. *FEBS Letters* 586, 1885–1896.
- Eichhorn, P.J., Rodon, L., Gonzalez-Junca, A., *et al.*, 2012. USP15 stabilizes TGF- β receptor I and promotes oncogenesis through the activation of TGF- β signaling in glioblastoma. *Nature Medicine* 18, 429–435.
- Elderbrook, J.L., Huang, J.J., Gatz, C.E., *et al.*, 2014. Ectodomain shedding of TBR1 is required for TBR1-mediated suppression of TGF- β signaling and breast cancer migration and invasion. *Molecular Biology of the Cell* 25, 2320–2332.
- Feng, X.H., Zhang, Y., Wu, R.Y., Derynck, R., 1998. The tumor suppressor Smad4/DPC4 and transcriptional adaptor CBP/p300 are coactivators for smad3 in TGF- β -induced transcriptional activation. *Genes & Development* 12, 2153–2163.
- Gao, S., Alarcón, C., Sapkota, G., *et al.*, 2009. Ubiquitin ligase Nedd4L targets activated Smad2/3 to limit TGF- β signaling. *Molecular Cell* 36, 457–468.
- Ge, G., Hopkins, D.R., Ho, W.B., Greenspan, D.S., 2005. GDF11 forms a bone morphogenetic protein 1-activated latent complex that can modulate nerve growth factor-induced differentiation of PC12 cells. *Molecular and Cellular Biology* 25, 5846–5858.
- Germain, S., Howell, M., Esslemont, G.M., Hill, C.S., 2000. Homeodomain and winged-helix transcription factors recruit activated Smads to distinct promoter elements via a common Smad interaction motif. *Genes & Development* 14, 435–451.
- Gong, Y., Krakow, D., Marcelino, J., *et al.*, 1999. Heterozygous mutations in the gene encoding noggin affect human joint morphogenesis. *Nature Genetics* 21, 302–304.
- Goudie, D.R., D'Alessandro, M., Merriman, B., *et al.*, 2011. Multiple self-healing squamous epithelioma is caused by a disease-specific spectrum of mutations in TGFBR1. *Nature Genetics* 43, 365–369.
- Goumans, M.J., Mummery, C., 2000. Functional analysis of the TGF β receptor/Smad pathway through gene ablation in mice. *International Journal of Developmental Biology* 44, 253–265.
- Goumans, M.J., Valdimarsdottir, G., Itoh, S., *et al.*, 2002. Balancing the activation state of the endothelium via two distinct TGF- β type I receptors. *EMBO Journal* 21, 1743–1753.
- Gray, A.M., Mason, A.J., 1990. Requirement for activin A and transforming growth factor- β 1 pro-regions in homodimer assembly. *Science* 247, 1328–1330.
- Gronroos, E., Kingston, I.J., Ramachandran, A., *et al.*, 2012. Transforming growth factor β inhibits bone morphogenetic protein-induced transcription through novel phosphorylated Smad1/5-Smad3 complexes. *Molecular and Cellular Biology* 32, 2904–2916.
- Groppe, J., Hinck, C.S., Samavarchi-Tehrani, P., *et al.*, 2008. Cooperative assembly of TGF- β superfamily signaling complexes is mediated by two disparate mechanisms and distinct modes of receptor binding. *Molecular Cell* 29, 157–168.
- Hagemann, A.I., Xu, X., Nentwich, O., Hyvonen, M., Smith, J.C., 2009. Rab5-mediated endocytosis of activin is not required for gene activation or long-range signalling in *Xenopus*. *Development* 136, 2803–2813.
- Hata, A., Lagna, G., Massague, J., Hemmati-Brivanlou, A., 1998. Smad6 inhibits BMP/Smad1 signaling by specifically competing with the Smad4 tumor suppressor. *Genes & Development* 12, 186–197.
- Holtzhausen, A., Golzio, C., How, T., *et al.*, 2014. Novel bone morphogenetic protein signaling through Smad2 and Smad3 to regulate cancer progression and development. *FASEB Journal* 28, 1248–1267.
- Howe, J.R., Roth, S., Ringold, J.C., *et al.*, 1998. Mutations in the SMAD4/DPC4 gene in juvenile polyposis. *Science* 280, 1086–1088.
- Huang, T., David, L., Mendoza, V., *et al.*, 2011. TGF- β signaling is mediated by two autonomously functioning T β R1:T β R2 pairs. *EMBO Journal* 30, 1263–1276.
- Huminiecki, L., Goldovsky, L., Freilich, S., *et al.*, 2009. Emergence, development and diversification of the TGF- β signalling pathway within the animal kingdom. *BMC Evolutionary Biology* 9, 28.
- Inman, G.J., Nicolas, F.J., Hill, C.S., 2002. Nucleocytoplasmic shuttling of Smads 2, 3, and 4 permits sensing of TGF- β receptor activity. *Molecular Cell* 10, 283–294.
- International PPH Consortium, Lane, K.B., Machado, R.D., Pauciulo, M.W., *et al.*, 2000. Heterozygous germline mutations in BMPR2, encoding a TGF- β receptor, cause familial primary pulmonary hypertension. *Nature Genetics* 26, 81–84.
- Inui, M., Manfrin, A., Mamidi, A., *et al.*, 2011. USP15 is a deubiquitylating enzyme for receptor-activated SMADs. *Nature Cell Biology* 13, 1368–1375.
- Jaeger, E., Leedham, S., Lewis, A., *et al.*, 2012. Hereditary mixed polyposis syndrome is caused by a 40-kb upstream duplication that leads to increased and ectopic expression of the BMP antagonist GREM1. *Nature Genetics* 44, 699–703.
- Johnson, D.W., Berg, J.N., Baldwin, M.A., *et al.*, 1996. Mutations in the activin receptor-like kinase 1 gene in hereditary haemorrhagic telangiectasia type 2. *Nature Genetics* 13, 189–195.
- Julien, J., Gurdon, J., 2005. Morphogen gradient interpretation by a regulated trafficking step during ligand-receptor transduction. *Genes & Development* 19, 2682–2694.
- Kang, H., Louie, J., Weisman, A., *et al.*, 2012. Inhibition of microRNA-302 (miR-302) by bone morphogenetic protein 4 (BMP4) facilitates the BMP signaling pathway. *Journal of Biological Chemistry* 287, 38656–38664.
- Kang, J.S., Alliston, T., Delston, R., Derynck, R., 2005. Repression of Runx2 function by TGF- β through recruitment of class II histone deacetylases by Smad3. *EMBO Journal* 24, 2543–2555.
- Kang, J.S., Saunier, E.F., Akhurst, R.J., Derynck, R., 2008. The type I TGF- β receptor is covalently modified and regulated by sumoylation. *Nature Cell Biology* 10, 654–664.
- Kang, Y., Chen, C.R., Massague, J., 2003. A self-enabling TGF β response coupled to stress signaling: Smad engages stress response factor ATF3 for Id1 repression in epithelial cells. *Molecular Cell* 11, 915–926.
- Kaplan, F.S., Lounev, V.Y., Wang, H., Pignolo, R.J., Shore, E.M., 2011. Fibrodysplasia ossificans progressiva: A blueprint for metamorphosis. *Annals of the New York Academy of Sciences* 1237, 5–10.
- Kavak, P., Rasmussen, R.K., Causing, C.G., *et al.*, 2000. Smad7 binds to Smurf2 to form an E3 ubiquitin ligase that targets the TGF β receptor for degradation. *Molecular Cell* 6, 1365–1375.

- Kim, Y.W., Park, J., Lee, H.J., Lee, S.Y., Kim, S.J., 2012. TGF- β sensitivity is determined by N-linked glycosylation of the type II TGF- β receptor. *Biochemical Journal* 445, 403–411.
- Komuro, A., Imamura, T., Saitoh, M., *et al.*, 2004. Negative regulation of transforming growth factor- β (TGF- β) signaling by WW domain-containing protein 1 (WWP1). *Oncogene* 23, 6914–6923.
- Kunwar, P.S., Zimmerman, S., Bennett, J.T., *et al.*, 2003. Mixer/Bon and FoxH1/Sur have overlapping and divergent roles in Nodal signaling and mesendoderm induction. *Development* 130, 5589–5599.
- Kuratomi, G., Komuro, A., Goto, K., *et al.*, 2005. NEDD4-2 (neural precursor cell expressed, developmentally down-regulated 4-2) negatively regulates TGF- β (transforming growth factor- β) signalling by inducing ubiquitin-mediated degradation of Smad2 and TGF- β type I receptor. *Biochemical Journal* 386, 461–470.
- van de Laar, I.M., van der Linde, D., Oei, E.H., *et al.*, 2012. Phenotypic spectrum of the SMAD3-related aneurysms-osteoarthritis syndrome. *Journal of Medical Genetics* 49, 47–57.
- Lebrin, F., Goumans, M.J., Jonker, L., *et al.*, 2004. Endoglin promotes endothelial cell proliferation and TGF- β /ALK1 signal transduction. *EMBO Journal* 23, 4018–4028.
- Lee, P.S., Chang, C., Liu, D., Derynck, R., 2003. Sumoylation of Smad4, the common Smad mediator of transforming growth factor- β family signaling. *Journal of Biological Chemistry* 278, 27853–27863.
- Levy, L., Hill, C.S., 2006. Alterations in components of the TGF- β superfamily signaling pathways in human cancer. *Cytokine & Growth Factor Reviews* 17, 41–58.
- Levy, L., Howell, M., Das, D., *et al.*, 2007. Arkadia activates Smad3/Smad4-dependent transcription by triggering signal-induced SnoN degradation. *Molecular and Cellular Biology* 27, 6068–6083.
- Lin, X., Liang, Y.Y., Sun, B., *et al.*, 2003. Smad6 recruits transcription corepressor CtBP to repress bone morphogenetic protein-induced transcription. *Molecular and Cellular Biology* 23, 9081–9093.
- Little, S.C., Mullins, M.C., 2009. Bone morphogenetic protein heterodimers assemble heteromeric type I receptor complexes to pattern the dorsoventral axis. *Nature Cell Biology* 11, 637–643.
- Lonn, P., van der Heide, L.P., Dahl, M., *et al.*, 2010. PARP-1 attenuates Smad-mediated transcription. *Molecular Cell* 40, 521–532.
- Lopez-Casillas, F., Wrana, J.L., Massague, J., 1993. Betaglycan presents ligand to the TGF β signaling receptor. *Cell* 73, 1435–1444.
- Lu, Z., Murray, J.T., Luo, W., *et al.*, 2002. Transforming growth factor β activates Smad2 in the absence of receptor endocytosis. *Journal of Biological Chemistry* 277, 29363–29368.
- Luyten, F.P., 1997. Cartilage-derived morphogenetic protein-1. *International Journal of Biochemistry & Cell Biology* 29, 1241–1244.
- Martello, G., Zaccagna, L., Inui, M., *et al.*, 2007. MicroRNA control of Nodal signalling. *Nature* 449, 183–188.
- Massague, J., 2012. TGF β signalling in context. *Nature Reviews Molecular Cell Biology* 13, 616–630.
- McAllister, K.A., Grogg, K.M., Johnson, D.W., *et al.*, 1994. Endoglin, a TGF- β binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. *Nature Genetics* 8, 345–351.
- Mesnard, D., Guzman-Ayala, M., Constam, D.B., 2006. Nodal specifies embryonic visceral endoderm and sustains pluripotent cells in the epiblast before overt axial patterning. *Development* 133, 2497–2505.
- Mitchell, H., Choudhury, A., Pagano, R.E., Leof, E.B., 2004. Ligand-dependent and -independent transforming growth factor- β receptor recycling regulated by clathrin-mediated endocytosis and Rab11. *Molecular Biology of the Cell* 15, 4166–4178.
- Miyazono, K., Kamiya, Y., Morikawa, M., 2010. Bone morphogenetic protein receptors and signal transduction. *Journal of Biochemistry* 147, 35–51.
- Moren, A., Imamura, T., Miyazono, K., Heldin, C.H., Moustakas, A., 2005. Degradation of the tumor suppressor Smad4 by WW and HECT domain ubiquitin ligases. *Journal of Biological Chemistry* 280, 22115–22123.
- Murray-Rust, J., McDonald, N.Q., Blundell, T.L., *et al.*, 1993. Topological similarities in TGF- β 2, PDGF-BB and NGF define a superfamily of polypeptide growth factors. *Structure* 1, 153–159.
- Ogunjimi, A.A., Briant, D.J., Pece-Barbara, N., *et al.*, 2005. Regulation of Smurf2 ubiquitin ligase activity by anchoring the E2 to the HECT domain. *Molecular Cell* 19, 297–308.
- Peng, J., Li, Q., Wigglesworth, K., *et al.*, 2013. Growth differentiation factor 9: bone morphogenetic protein 15 heterodimers are potent regulators of ovarian functions. *Proceedings of the National Academy of Sciences of the United States of America* 110, E776–E785.
- Ramel, M.C., Hill, C.S., 2012. Spatial regulation of BMP activity. *FEBS Letters* 586, 1929–1941.
- Ramirez, F., Sakai, L.Y., 2010. Biogenesis and function of fibrillin assemblies. *Cell and Tissue Research* 339, 71–82.
- Rechtman, M.M., Nakaryakov, A., Shapira, K.E., Ehrlich, M., Henis, Y.I., 2009. Different domains regulate homomeric and heteromeric complex formation among type I and type II transforming growth factor- β receptors. *Journal of Biological Chemistry* 284, 7843–7852.
- Reichert, S., Randall, R.A., Hill, C.S., 2013. A BMP regulatory network controls ectodermal cell fate decisions at the neural plate border. *Development* 140, 4435–4444.
- Ross, S., Cheung, E., Petrakis, T.G., *et al.*, 2006. Smads orchestrate specific histone modifications and chromatin remodeling to activate transcription. *EMBO Journal* 25, 4490–4502.
- Ross, S., Hill, C.S., 2008. How the Smads regulate transcription. *International Journal of Biochemistry & Cell Biology* 40, 383–408.
- Sapkota, G., Alarcon, C., Spagnoli, F.M., Brivanlou, A.H., Massague, J., 2007. Balancing BMP signaling through integrated inputs into the Smad1 linker. *Molecular Cell* 25, 441–454.
- Schier, A.F., 2003. Nodal signaling in vertebrate development. *Annual Review of Cell and Developmental Biology* 19, 589–621.
- Schmieder, B., Tournier, A.L., Bates, P.A., Hill, C.S., 2008. Mathematical modeling identifies Smad nucleocytoplasmic shuttling as a dynamic signal-interpreting system. *Proceedings of the National Academy of Sciences of the United States of America* 105, 6608–6613.
- Serpe, M., Umlulis, D., Ralston, A., *et al.*, 2008. The BMP-binding protein Crossveinless 2 is a short-range, concentration-dependent, biphasic modulator of BMP signaling in *Drosophila*. *Developmental Cell* 14, 940–953.
- Shi, M., Zhu, J., Wang, R., *et al.*, 2011. Latent TGF- β structure and activation. *Nature* 474, 343–349.
- Shi, W., Sun, C., He, B., *et al.*, 2004. GADD34-PP1c recruited by Smad7 dephosphorylates TGF β type I receptor. *Journal of Cell Biology* 164, 291–300.
- Shi, Y., Massague, J., 2003. Mechanisms of TGF- β signaling from cell membrane to the nucleus. *Cell* 113, 685–700.
- Shiratori, H., Hamada, H., 2014. TGF β signaling in establishing left-right asymmetry. *Seminars in Cell & Developmental Biology* 32, 80–84.
- Simonsson, M., Kanduri, M., Gronroos, E., Heldin, C.H., Ericsson, J., 2006. The DNA binding activities of Smad2 and Smad3 are regulated by coactivator-mediated acetylation. *Journal of Biological Chemistry* 281, 39870–39880.
- Tanaka, C., Sakuma, R., Nakamura, T., Hamada, H., Saijoh, Y., 2007. Long-range action of Nodal requires interaction with GDF1. *Genes & Development* 21, 3272–3282.
- Taylor, K.R., Mackay, A., Truffaux, N., *et al.*, 2014. Recurrent activating ACVR1 mutations in diffuse intrinsic pontine glioma. *Nature Genetics* 46, 457–461.
- Tilak, A., Nelsen, S.M., Kim, H.S., *et al.*, 2014. Simultaneous rather than ordered cleavage of two sites within the BMP4 prodomain leads to loss of ligand in mice. *Development* 141, 3062–3071.
- Tsukazaki, T., Chiang, T.A., Davison, A.F., Attisano, L., Wrana, J.L., 1998. SARA, a FYVE domain protein that recruits Smad2 to the TGF β receptor. *Cell* 95, 779–791.
- Vizan, P., Miller, D.S., Gori, I., *et al.*, 2013. Controlling long-term signaling: Receptor dynamics determine attenuation and refractory behavior of the TGF- β pathway. *Science Signaling* 6, ra106.
- Wakefield, L.M., Hill, C.S., 2013. Beyond TGF β : roles of other TGF β superfamily members in cancer. *Nature Reviews Cancer* 13, 328–341.
- Wang, T., Donahoe, P.K., 2004. The immunophilin FKBP12: A molecular guardian of the TGF- β family type I receptors. *Frontiers in Bioscience* 9, 619–631.
- Weis-Garcia, F., Massague, J., 1996. Complementation between kinase-defective and activation-defective TGF- β receptors reveals a novel form of receptor cooperativity essential for signaling. *EMBO Journal* 15, 276–289.
- Wells, R.G., Gilboa, L., Sun, Y., *et al.*, 1999. Transforming growth factor- β induces formation of a dithiothreitol-resistant type I/type II receptor complex in live cells. *Journal of Biological Chemistry* 274, 5716–5722.
- Wells, R.G., Yankeliev, H., Lin, H.Y., Lodish, H.F., 1997. Biosynthesis of the type I and type II TGF- β receptors. Implications for complex formation. *Journal of Biological Chemistry* 272, 11444–11451.
- Wicks, S.J., Haros, K., Maillard, M., *et al.*, 2005. The deubiquitinating enzyme UCH37 interacts with Smads and regulates TGF- β signaling. *Oncogene* 24, 8080–8084.
- Wrana, J.L., Attisano, L., Wieser, R., Ventura, F., Massague, J., 1994. Mechanism of activation of the TGF- β receptor. *Nature* 370, 341–347.
- Wu, G., Diaz, A.K., Paugh, B.S., *et al.*, 2014. The genomic landscape of diffuse intrinsic pontine glioma and pediatric non-brainstem high-grade glioma. *Nature Genetics* 46, 444–450.

- Xi, Q., He, W., Zhang, X.H., Le, H.V., Massague, J., 2008. Genome-wide impact of the BRG1 SWI/SNF chromatin remodeler on the transforming growth factor β transcriptional program. *Journal of Biological Chemistry* 283, 1146–1155.
- Xu, J., Wang, A.H., Oses-Prieto, J., *et al.*, 2013. Arginine methylation initiates BMP-induced Smad signaling. *Molecular Cell* 51, 5–19.
- Xu, L., Kang, Y., Col, S., Massague, J., 2002. Smad2 nucleocytoplasmic shuttling by nucleoporins CAN/Nup214 and Nup153 feeds TGF β signaling complexes in the cytoplasm and nucleus. *Molecular Cell* 10, 271–282.
- Xu, P.F., Houssin, N., Ferri-Lagneau, K.F., Thisse, B., Thisse, C., 2014. Construction of a vertebrate embryo from two opposing morphogen gradients. *Science* 344, 87–89.
- Zhang, L., Huang, H., Zhou, F., *et al.*, 2012a. RNF12 controls embryonic stem cell fate and morphogenesis in zebrafish embryos by targeting Smad7 for degradation. *Molecular Cell* 46, 650–661.
- Zhang, L., Zhou, F., Drabsch, Y., *et al.*, 2012b. USP4 is regulated by AKT phosphorylation and directly deubiquitylates TGF- β type I receptor. *Nature Cell Biology* 14, 717–726.
- Zhang, L., Zhou, F., García de Vinuesa, A., *et al.*, 2013. TRAF4 promotes TGF- β receptor signaling and drives breast cancer metastasis. *Molecular Cell* 51, 559–572.

Relevant Websites

<http://www.genecards.org/>
GeneCards.
<http://smad.maciasnmr.net>
Macías Lab.

G Protein-Coupled Receptors

W Saengsawang, Mahidol University, Bangkok, Thailand

MM Rasenick, University of Illinois Chicago College of Medicine, Chicago, IL, USA and Jesse Brown VA Medical Center, Chicago, IL, USA

© 2016 Elsevier Inc. All rights reserved.

G Protein-Coupled Receptors

G Protein-Coupled Receptors (GPCRs), which convey the messages borne by several ligands such as hormones, neurotransmitters, odorants, or phospholipids, are ubiquitous through the phylogenetic spectrum. GPCRs are coded by a gene family representing 5% of human genome. Activation of these receptors, in turn, regulates diverse downstream signaling pathways that are crucial for normal physiological responses. Therefore, GPCRs are the targets for about half of the drugs in current use (Morris and Malbon, 1999; Ji *et al.*, 1998; Flower, 1999). GPCRs form complexes with heterotrimeric G proteins ($G\alpha$ and $G\beta\gamma$ subunits). They are activated by an external signal (ligand) that induces a conformational change in the receptor, causing activation of a G protein. The diversity of GPCRs is determined by the variety of stimuli and their effects depend on different types of G proteins with which they couple.

We often associate GPCRs with conventional signaling pathways, the generation of intracellular second messengers such as cAMP, inositol trisphosphate, or calcium. More recently, it has become evident that these receptors also engage intracellular signaling pathways through integration of several networks. For example, MAPKs (mitogen-activated protein kinases) and JNK (c-Jun N-terminal kinases) network has been shown to link GPCRs to the nucleus where they mediate cell proliferation (Marinissen and Gutkind, 2001). In addition, several effectors for GPCR that are independent of G proteins have also been identified. β -arrestins internalize in response to GPCR activation and engage intracellular signaling cascades described in this section.

In this article we will diverge from either of these G protein signaling pathways and address the information connecting GFs and GPCR, emphasizing on cross talk between growth factor and GPCRs that results in modification of microtubule-based structures and, ultimately, changes in cellular morphology as those observed in neurite outgrowth.

Growth Factor Receptors and GPCRs

Several GFs mediate cell signaling through GPCRs. The convergences of GPCRs and various receptor tyrosine kinase (RTK) signaling pathways have been reported. This cross talk between GPCRs and RTK pathways could be at the receptor level or at some downstream signaling pathway (Figure 1; Malarkey *et al.*, 1995; Luttrell *et al.*, 1999; Gavi *et al.*, 2006; Natarajan and Berk, 2006).

Activation of epidermal growth factor receptor (EGFR) is linked to regulation of several G protein signaling cascades including $G\alpha$, $G_{i\alpha}$, and G_{12} by direct interaction between EGFR and either GPCR or G proteins (Poppleton *et al.*, 1996; Zhang *et al.*, 2001; Piiper *et al.*, 1997; Yang *et al.*, 1991; Profrock *et al.*, 1991). GPCR-EGFR cross talk plays important

roles in endocrine systems, including the neuroendocrine regulation of reproduction and ovulation (Hsieh and Conti, 2005). In addition, transactivation of EGFR by $\alpha 1$ -adrenoceptor ($\alpha 1$ -AR) is implicated in contraction and hypertrophy of vascular smooth muscle. This process may require participation of PI3 Kinase (Ulu *et al.*, 2013).

Several studies also indicate that GPCR and G proteins are required in some PDGF signaling pathways, such as PDGF-induced ROS production, endocytosis, and cell migration (Hobson *et al.*, 2001; Freedman *et al.*, 2002; Alderton *et al.*, 2001). S1P1 (sphingosine1-phosphate) receptor, a GPCR that plays a role in regulation of smooth muscle cell migration, proliferation, and vascular maturation, has also been shown to be required for PDGF-stimulated cell motility in several cell types (Pyne *et al.*, 2003; Hobson *et al.*, 2001).

Cross talk between GPCR and insulin or IGF have been reported in many cell types (Yang *et al.*, 1991; Rozengurt *et al.*, 2010). For example, G_i has been shown to interact with insulin and IGF receptors (Profrock *et al.*, 1991; Hallak *et al.*, 2000). Both $G_{i\alpha}$ and $G_{\beta\gamma}$ were shown to be involved in IGF-I-mediated mitogenic signal pathways (Dalle *et al.*, 2001). $G_{q\alpha}$ was also shown to play a role in insulin-induced glucose transport in adipocytes (Imamura *et al.*, 1999). Basic fibroblast growth factor

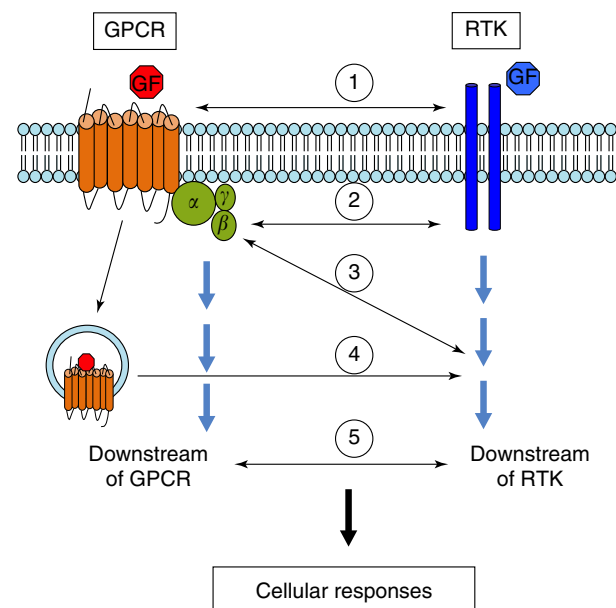


Figure 1 Model showing cross talk between GPCRs and RTK pathways. Ligand binding to either GPCR or RTK induces (1) direct interaction of the receptors, (2) direct interaction of G proteins to the RTK, or (3) to the downstream signaling molecules. (4) Endocytosis of GPCR may also be required to link GPCR to RTK signaling pathways. (5) Cross talk might also occur at downstream signaling pathways. GF, growth factors.

(bFGF) was shown to stimulate adenylyl cyclase by coupling to G_{α} but inhibit NADPH oxidase via $G_{\beta\gamma}$ (Krieger-Brauer *et al.*, 2000). In addition, $G_{\beta\gamma}$ was shown to be linked to FGF inducing skeletal muscle differentiation through ERK1/2 activation (Fedorov *et al.*, 1998). Likewise, $G_{\alpha q}$ and $G_{\beta\gamma}$ have been shown to be involved in cell migration and proliferation regulated by vascular endothelial growth factor (VEGF) (Zeng *et al.*, 2002a,b, 2003). In neurons, nerve growth factor (NGF) was shown to induce neuronal differentiation and neurite outgrowth through G protein cross talk, which will be discussed in more detail in the next section (Sarma *et al.*, 2003).

In addition to RTK, the WNT pathway is also known to involve activation of a GPCR, (frizzled) and the subsequent engagement of cellular growth pathways. For the most part, these involve translocation of β -catenin to the nucleus and subsequent gene activation. Curiously, despite exhaustive search, the nature of the G protein involved has not been identified. WNT may represent 'biased agonism,' or the engagement of an arrestin-dependent, G protein independent, pathway following activation of a given GPCR. On the other hand, activation of pertussis-toxin sensitive G proteins (G_i / G_o) has been reported to accompany activation of frizzled by WNT (Kilander *et al.*, 2014).

NGF Induces Neurite Outgrowth through G Protein

NGF promotes G_{α} translocation and interaction of tubulin and G_{α} without changing the expression of G_{α} in PC12 pheochromocytoma cells. In the presence of NGF the association of G_{α} (Kobilka and Schertler, 2008) with microtubules in the cellular processes is increased (Sarma *et al.*, 2003), suggesting an interplay between GPCR and RTK, which might be another possible mechanism of NGF inducing neuronal differentiation and neurite outgrowth.

In addition, expression of the dominant-negative G_{α} in PC12 cells greatly inhibited neurite outgrowth, whether evoked by expression of mutationally active G_{α} or by NGF (unpublished data). While the degree of inhibition was greater for the activated G_{α} than for NGF, neurite outgrowth evoked by the latter was inhibited, suggesting that there is some cross talk between G_{α} - and NGF-mediated events.

GPCRs, G_{α} , and Neurite Outgrowth

One recent study showed that activation of G_{α} and subsequent translocation from the plasma membrane resulted in G_{α} decoration of microtubules (Yu *et al.*, 2009). This occurred subsequent to either GPCR activation or direct activation of G_{α} with cholera toxin. Mutationally activated G_{α} behaved in a similar fashion and all elicited neurite outgrowth in PC12 pheochromocytoma cells. Only the activated (GTP-bound) conformation of G_{α} bound to microtubules, but the binding did not discriminate between GDP- and GTP-conformations of tubulin. Curiously, forskolin, which activates adenylyl cyclase and increases cAMP did not promote translocation of G_{α} , but did increase neurite outgrowth. When this was further examined using PC12 cells deficient in PKA, it was established that there was both a cAMP dependent and a cAMP independent component to neurite outgrowth elicited

by activation of GPCR and subsequently G_{α} (Yu *et al.*, 2009). In order to effect neurite outgrowth directly, it appeared that G_{α} bound directly to microtubules subsequent to translocation from the membrane. So, G_{α} acted directly as a 'second messenger' for the relevant GPCR.

Mechanisms for G_{α} -Induced Neurite Outgrowth

Microtubules are one of the important components of the cytoskeleton and are involved in many cellular processes including neurite formation and outgrowth. Structurally, microtubules are linear polymers of α - and β -tubulin. Structures of α and β tubulin are similar and they both bind GTP. However, GTP on α subunit is buried between the two subunits (α and β), while the GTP on the β subunit is exposed to solution. Therefore, only the β tubulin is able to hydrolyze GTP or exchange GDP for GTP (exchangeable). This hydrolysis occurs as a result of intrinsic GTPase activity of tubulin, which is activated during the process of polymerization upon addition of another tubulin dimer (Carlier and Pantaloni, 1983). Microtubules are highly dynamic and this dynamicity is based on an intrinsic property of tubulin capable binding, exchange, and hydrolysis of guanine nucleotides. Therefore, tubulin GTPase activity is a major determinant of microtubule polymerization and dynamics (Carlier, 1982).

After G_{α} is activated, it internalizes and binds directly to tubulin at the plus end of microtubules. Upon binding to tubulin, G_{α} stimulates hydrolysis of tubulin GTP (Wang *et al.*, 1990). Activated G_{α} stimulates tubulin GTPase with an EC_{50} of 1.2 μ M (Yu *et al.*, 2009). This suggests noncooperative activation of tubulin GTPase by G_{α} , and is consistent with a 1:1 stoichiometry between the two proteins (Dave *et al.*, 2011). This stimulation of tubulin-GTP hydrolysis results in an increase in microtubule dynamics (Yu *et al.*, 2009). Activated G_{α} increases microtubule dynamics by increasing the growing rate, the shortening rate and the catastrophe frequency, the frequency of transition from growth to shortening (Roychowdhury *et al.*, 1999; Yu *et al.*, 2009; Dave *et al.*, 2011).

Association between G Proteins and Tubulin

Several studies both *in vitro* and *in vivo* suggest direct interaction between G proteins and tubulin (Rudolph *et al.*, 1977; Kennedy and Insel, 1979; Rasenick *et al.*, 1981; Yan *et al.*, 1996). G_{α} , $G_{\alpha 1}$, and $G_{\alpha q}$ all show association with tubulin but other G protein alpha subunits, most notably $G_{\alpha t}$ (transducin) do not bind (Wang *et al.*, 1990). This allowed comparison of sequences in peptide and creation of a model for the binding interface for the tubulin- G_{α} complex (Layden *et al.*, 2008; Dave *et al.*, 2011). However, the difference between G_{α} and other G proteins is that the active conformation of G_{α} is required for tubulin interaction. GPCR stimulation by agonists or direct activation of G_{α} by cholera toxin causes translocation of G_{α} from plasma membrane to cytoplasm and microtubule networks (Yu and Rasenick, 2002). Note that since the only G_{α} that internalizes after activation is G_{α} (Allen *et al.*, 2005; Wedegaertner, 2012), this is likely the most relevant G protein to GPCR-mediated neurite outgrowth. Curiously, G_{α} , like most G protein alpha subunits has an intrinsic GTPase that would be

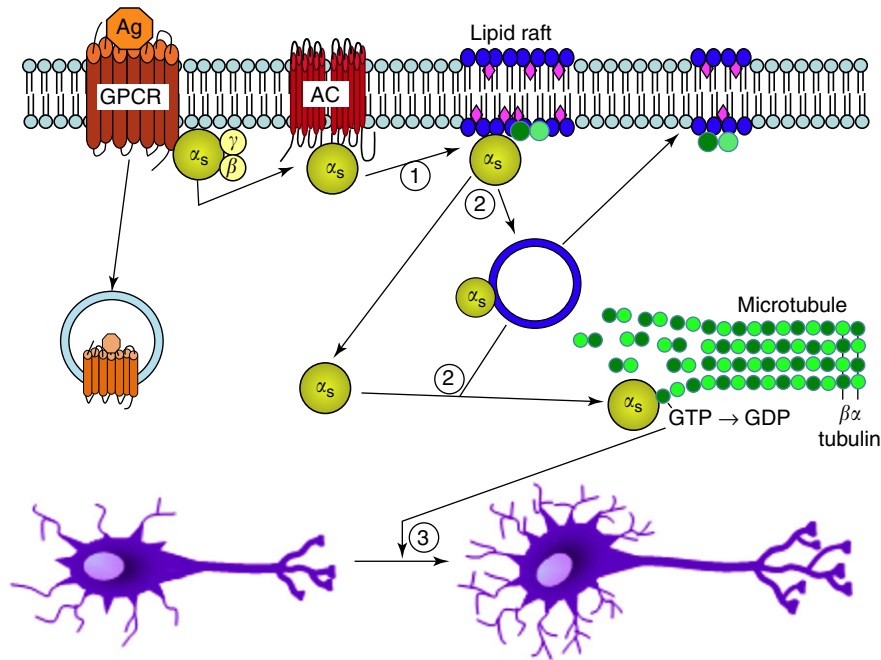


Figure 2 Translocation of $G_s\alpha$ from membrane to cytosol and association with microtubule plus end. Activated $G_s\alpha$ induces microtubule dynamics and increases neurite outgrowth. (1) Agonist binds to GPCR and promotes exchange of GDP for GTP on $G_s\alpha$, conferring the activated conformation to that G protein. (2) $G_s\alpha$ translocates to lipid rafts, where it is internalized, some of it associating with microtubule plus ends and activating the intrinsic tubulin GTPase. This hydrolyzes the GTP 'cap,' making, decreasing microtubule stability and increasing microtubule dynamics. (3) Greater microtubule dynamics increases neurite branching and neurite outgrowth. GPCR, G protein-coupled receptor; Ag-agonist; AC, adenylyl cyclase; α_s , alpha subunit of Gs; $\beta\gamma$, G protein beta/gamma subunits.

expected to hydrolyze rapidly GTP and convert $G_s\alpha$ to the GDP conformation, which does not bind to tubulin. Von Zastrow and colleagues have demonstrated, using conformation-sensitive nanobodies, that internalized $G_s\alpha$ retains the active, GTP-bound conformation (Irannejad *et al.*, 2013). This is consistent with GPCR/ $G_s\alpha$ -driven increase in microtubule dynamics and subsequent neurite outgrowth.

$G\beta\gamma$ subunits have also been shown to translocate to the cytosol following activation of certain $G\alpha$ (Jin *et al.*, 2000; Popova and Rasenick, 2003) and it appears that $G\beta\gamma$ binds to microtubules along a regular pattern, not unlike that seen with microtubule-associated proteins (Roychowdhury and Rasenick, 1997). Just as the transducin α subunit does not bind tubulin, the $G\gamma$ subunit most closely associated with transducin, $G\gamma 1$, does not bind tubulin. This appears due to the requirement that the $G\gamma$ be geranylgeranylated rather than farnesylated (as is the case with $G\gamma 1$) (Roychowdhury and Rasenick, 1997). The translocation to the cytosol of $G\beta\gamma$ follows activation of Gq-coupled receptors (m1, m3) that also evoke microtubule depolymerization due to increased Ca^{2+} . It appears that there is a continuum between microtubule depolymerization by Ca^{2+} and stabilization by $G\beta\gamma$ – both responding to the same external signal (acetylcholine) (Popova and Rasenick, 2003, 2004).

Conclusion

This article suggests interplay between GFs, GPCRs, and G proteins, which is crucial for several cellular responses and physiological functions. It also indicates a direct role for the G

protein, $G_s\alpha$ in regulating both microtubule dynamics and in regulating microtubule-dependent cellular outgrowth. Further, it demonstrates a direct role for G proteins and the GPCRs that activate them in the regulation of both microtubule dynamics and the modulation of cellular and neuronal morphology (Figure 2). Much remains unknown about both the interactions between GPCRs and growth factor receptors and the interface between their downstream signaling pathways. Ideally, this will be elucidated soon.

Acknowledgment

This work was supported, in part, by grant BX 0011049 from the US Veterans Administration.

See also: Cell Communication: Intracellular Pathways: The PLC Pathway

References

- Alderton, F., Rakhit, S., Kong, K.C., *et al.*, 2001. Tethering of the platelet-derived growth factor beta receptor to G-protein-coupled receptors. A novel platform for integrative signaling by these receptor classes in mammalian cells. *Journal of Biological Chemistry* 276, 28578–28585.
- Allen, J.A., Yu, J.Z., Donati, R.J., Rasenick, M.M., 2005. Beta-adrenergic receptor stimulation promotes G alpha s internalization through lipid rafts: A study in living cells. *Molecular Pharmacology* 67, 1493–1504.

- Carlier, M.F., 1982. Guanosine-5'-triphosphate hydrolysis and tubulin polymerization. Review article. *Molecular and Cellular Biochemistry* 47, 97–113.
- Carlier, M.F., Pantaloni, D., 1983. Taxol effect on tubulin polymerization and associated guanosine 5'-triphosphate hydrolysis. *Biochemistry* 22, 4814–4822.
- Dalle, S., Ricketts, W., Imamura, T., Vollenweider, P., Olefsky, J.M., 2001. Insulin and insulin-like growth factor I receptors utilize different G protein signaling components. *Journal of Biological Chemistry* 276, 15688–15695.
- Dave, R.H., Saengsawang, W., Lopus, M., *et al.*, 2011. A molecular and structural mechanism for G protein-mediated microtubule destabilization. *Journal of Biological Chemistry* 286, 4319–4328.
- Fedorov, Y.V., Jones, N.C., Olwin, B.B., 1998. Regulation of myogenesis by fibroblast growth factors requires beta-gamma subunits of pertussis toxin-sensitive G proteins. *Molecular and Cellular Biology* 18, 5780–5787.
- Flower, D.R., 1999. Modelling G-protein-coupled receptors for drug design. *Biochimica et Biophysica Acta* 1422, 207–234.
- Freedman, N.J., Kim, L.K., Murray, J.P., *et al.*, 2002. Phosphorylation of the platelet-derived growth factor receptor-beta and epidermal growth factor receptor by G protein-coupled receptor kinase-2. Mechanisms for selectivity of desensitization. *Journal of Biological Chemistry* 277, 48261–48269.
- Gavi, S., Shumay, E., Wang, H.Y., Malbon, C.C., 2006. G-protein-coupled receptors and tyrosine kinases: Crossroads in cell signaling and regulation. *Trends in Endocrinology and Metabolism* 17, 48–54.
- Hallak, H., Seiler, A.E., Green, J.S., Ross, B.N., Rubin, R., 2000. Association of heterotrimeric G(i) with the insulin-like growth factor-I receptor. Release of G (betagamma) subunits upon receptor activation. *Journal of Biological Chemistry* 275, 2255–2258.
- Hobson, J.P., Rosenfeldt, H.M., Barak, L.S., *et al.*, 2001. Role of the sphingosine-1-phosphate receptor EDG-1 in PDGF-induced cell motility. *Science* 291, 1800–1803.
- Hsieh, M., Conti, M., 2005. G-protein-coupled receptor signaling and the EGF network in endocrine systems. *Trends in Endocrinology and Metabolism* 16, 320–326.
- Imamura, T., Vollenweider, P., Egawa, K., *et al.*, 1999. G alpha-q/11 protein plays a key role in insulin-induced glucose transport in 3T3-L1 adipocytes. *Molecular and Cellular Biology* 19, 6765–6774.
- Irannejad, R., Tomshine, J.C., Tomshine, J.R., *et al.*, 2013. Conformational biosensors reveal GPCR signalling from endosomes. *Nature* 495, 534–538.
- Ji, T.H., Grossmann, M., Ji, I., 1998. G protein-coupled receptors. I. Diversity of receptor-ligand interactions. *Journal of Biological Chemistry* 273, 17299–17302.
- Jin, T., Zhang, N., Long, Y., Parent, C.A., Devreotes, P., 2000. Localization of the G protein betagamma complex in living cells during chemotaxis. *Science* 287, 1034–1036.
- Kennedy, M.S., Insel, P.A., 1979. Inhibitors of microtubule assembly enhance beta-adrenergic and prostaglandin E1-stimulated cyclic AMP accumulation in S49 lymphoma cells. *Molecular Pharmacology* 16, 215–223.
- Kilander, M.B., Petersen, J., Andressen, K.W., *et al.*, 2014. Dishevelled regulates precoupling of heterotrimeric G proteins to Frizzled 6. *FASEB Journal* 28, 2293–2305.
- Kobilka, B., Schertler, G.F., 2008. New G-protein-coupled receptor crystal structures: Insights and limitations. *Trends in Pharmacological Sciences* 29, 79–83.
- Krieger-Brauer, H.I., Medda, P., Kather, H., 2000. Basic fibroblast growth factor utilizes both types of component subunits of Gs for dual signaling in human adipocytes. Stimulation of adenylyl cyclase via G α ph(s) and inhibition of NADPH oxidase by G β gamma(s). *Journal of Biological Chemistry* 275, 35920–35925.
- Layden, B.T., Saengsawang, W., Donati, R.J., *et al.*, 2008. Structural model of a complex between the heterotrimeric G protein, G α alpha, and tubulin. *Biochimica et Biophysica Acta* 1783, 964–973.
- Luttrell, L.M., Daaka, Y., Lefkowitz, R.J., 1999. Regulation of tyrosine kinase cascades by G-protein-coupled receptors. *Current Opinion in Cell Biology* 11, 177–183.
- Malarkey, K., Belham, C.M., Paul, A., *et al.*, 1995. The regulation of tyrosine kinase signalling pathways by growth factor and G-protein-coupled receptors. *Biochemical Journal* 309 (Pt 2), 361–375.
- Marinissen, M.J., Gutkind, J.S., 2001. G-protein-coupled receptors and signaling networks: Emerging paradigms. *Trends in Pharmacological Sciences* 22, 368–376.
- Morris, A.J., Malbon, C.C., 1999. Physiological regulation of G protein-linked signaling. *Physiological Reviews* 79, 1373–1430.
- Natarajan, K., Berk, B.C., 2006. Crosstalk coregulation mechanisms of G protein-coupled receptors and receptor tyrosine kinases. *Methods in Molecular Biology* 332, 51–77.
- Piiper, A., Stryjek-Kaminska, D., Zeuzem, S., 1997. Epidermal growth factor activates phospholipase C-gamma 1 via G(i)1-2 proteins in isolated pancreatic acinar membranes. *American Journal of Physiology* 272, G1276–G1284.
- Popova, J.S., Rasenick, M.M., 2003. G beta gamma mediates the interplay between tubulin dimers and microtubules in the modulation of Gq signaling. *Journal of Biological Chemistry* 278, 34299–34308.
- Popova, J.S., Rasenick, M.M., 2004. Clathrin-mediated endocytosis of m3 muscarinic receptors. Roles for Gbetagamma and tubulin. *Journal of Biological Chemistry* 279, 30410–30418.
- Poppleton, H., Sun, H., Fulgham, D., Bertics, P., Patel, T.B., 1996. Activation of G α alpha by the epidermal growth factor receptor involves phosphorylation. *Journal of Biological Chemistry* 271, 6947–6951.
- Profrock, A., Schnefel, S., Schulz, I., 1991. Receptors for insulin interact with Gi-proteins and for epidermal growth factor with Gi- and Gs-proteins in rat pancreatic acinar cells. *Biochemical and Biophysical Research Communications* 175, 380–386.
- Pyne, N.J., Waters, C., Moughal, N.A., Sami, B.S., Pyne, S., 2003. Receptor tyrosine kinase-GPCR signal complexes. *Biochemical Society Transactions* 31, 1220–1225.
- Rasenick, M.M., Stein, P.J., Bitensky, M.W., 1981. The regulatory subunit of adenylate cyclase interacts with cytoskeletal components. *Nature* 294, 560–562.
- Roychowdhury, S., Panda, D., Wilson, L., Rasenick, M.M., 1999. G protein alpha subunits activate tubulin GTPase and modulate microtubule polymerization dynamics. *Journal of Biological Chemistry* 274, 13485–13490.
- Roychowdhury, S., Rasenick, M.M., 1997. G protein beta1gamma2 subunits promote microtubule assembly. *Journal of Biological Chemistry* 272, 31576–31581.
- Rozenfurt, E., Sinnott-Smith, J., Kisfalvi, K., 2010. Crosstalk between insulin/insulin-like growth factor-1 receptors and G protein-coupled receptor signaling systems: A novel target for the antidiabetic drug metformin in pancreatic cancer. *Clinical Cancer Research* 16, 2505–2511.
- Rudolph, S.A., Greengard, P., Malawista, S.E., 1977. Effects of colchicine on cyclic AMP levels in human leukocytes. *Proceedings of the National Academy of Sciences of the United States of America* 74, 3404–3408.
- Sarma, T., Voyno-Yasenetskaya, T., Hope, T.J., Rasenick, M.M., 2003. Heterotrimeric G-proteins associate with microtubules during differentiation in PC12 pheochromocytoma cells. *FASEB Journal* 17, 848–859.
- Ulu, N., Henning, R.H., Guner, S., *et al.*, 2013. Intracellular transactivation of epidermal growth factor receptor by alpha1A-adrenoceptor is mediated by phosphatidylinositol 3-kinase independently of activation of extracellular signal regulated kinases 1/2 and serine-threonine kinases in Chinese hamster ovary cells. *Journal of Pharmacology and Experimental Therapeutics* 347, 47–56.
- Wang, N., Yan, K., Rasenick, M.M., 1990. Tubulin binds specifically to the signal-transducing proteins, Gs alpha and Gi alpha 1. *Journal of Biological Chemistry* 265, 1239–1242.
- Wedegaertner, P.B., 2012. G protein trafficking. *Sub-Cellular Biochemistry* 63, 193–223.
- Yan, K., Greene, E., Belga, F., Rasenick, M.M., 1996. Synaptic membrane G proteins are complexed with tubulin *in situ*. *Journal of Neurochemistry* 66, 1489–1495.
- Yang, L.J., Baffy, G., Rhee, S.G., *et al.*, 1991. Pertussis toxin-sensitive Gi protein involvement in epidermal growth factor-induced activation of phospholipase C-gamma in rat hepatocytes. *Journal of Biological Chemistry* 266, 22451–22458.
- Yu, J.Z., Dave, R.H., Allen, J.A., Sarma, T., Rasenick, M.M., 2009. Cytosolic G α alpha s acts as an intracellular messenger to increase microtubule dynamics and promote neurite outgrowth. *Journal of Biological Chemistry* 284 (16), 10462–10472.
- Yu, J.Z., Rasenick, M.M., 2002. Real-time visualization of a fluorescent G(alpha)(s): Dissociation of the activated G protein from plasma membrane. *Molecular Pharmacology* 61, 352–359.
- Zeng, H., Zhao, D., Mukhopadhyay, D., 2002a. Flt-1-mediated down-regulation of endothelial cell proliferation through pertussis toxin-sensitive G proteins, beta gamma subunits, small GTPase CDC42, and partly by Rac-1. *Journal of Biological Chemistry* 277, 4003–4009.
- Zeng, H., Zhao, D., Mukhopadhyay, D., 2002b. KDR stimulates endothelial cell migration through heterotrimeric G protein Gq/11-mediated activation of a small GTPase RhoA. *Journal of Biological Chemistry* 277, 46791–46798.
- Zeng, H., Zhao, D., Yang, S., Datta, K., Mukhopadhyay, D., 2003. Heterotrimeric G alpha q/G alpha 11 proteins function upstream of vascular endothelial growth factor (VEGF) receptor-2 (KDR) phosphorylation in vascular permeability factor/VEGF signaling. *Journal of Biological Chemistry* 278, 20738–20745.
- Zhang, B.H., Ho, V., Farrell, G.C., 2001. Specific involvement of G(alpha)2 with epidermal growth factor receptor signaling in rat hepatocytes, and the inhibitory effect of chronic ethanol. *Biochemical Pharmacology* 61, 1021–1027.

Further Reading

- Allen, J.A., Halverson-Tamboli, R.A., Rasenick, M.M., 2007. Lipid raft microdomains and neurotransmitter signalling. *Nature Reviews Neuroscience* 8, 128–140.
- Dave, R.H., Saengsawang, W., Yu, J.Z., Donati, R., Rasenick, M.M., 2009. Heterotrimeric G-proteins interact directly with cytoskeletal components to modify microtubule-dependent cellular processes. *Neurosignals* 17, 100–108.
- Kalil, K., Dent, E.W., 2014. Branch management: Mechanisms of axon branching in the developing vertebrate CNS. *Nature Reviews Neuroscience* 15, 7–18.
- Pyne, N., Waters, C., Moughal, N., Sambhi, B., Pyne, S., 2003. Receptor tyrosine kinase–GPCR signal complexes. *Biochemical Society Transactions* 31, 1220–1225.
- Safinya, C.R., Raviv, U., Needleman, D.J., *et al.*, 2011. Nanoscale assembly in biological systems: From neuronal cytoskeletal proteins to curvature stabilizing lipids. *Advanced Materials* 23, 2260–2270.
- Schappi, J.M., Krbanjevic, A., Rasenick, M.M., 2014. Tubulin, actin and heterotrimeric G proteins: Coordination of signaling and structure. *Biochimica et Biophysica Acta* 1838, 674–681.

Guanylyl Cyclase Receptors

AE Snook, ES Blomain, AS Aing, D Merlino, and SA Waldman, Thomas Jefferson University, Philadelphia, PA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Cellular survival depends on an organism's ability to respond and adapt to extracellular signals in its environment. Signal transduction is, therefore, an important biologic process, allowing cells to react to a variety of these extracellular stimuli and respond adaptively. Guanylyl cyclases (GCs) comprise one ubiquitous class of signaling enzymes that sense and transduce extracellular signals to intracellular signaling cascades through production of the cyclic nucleotide second messenger known as cyclic guanosine monophosphate (cyclic GMP or cGMP). In turn, cGMP activates cGMP-dependent protein kinases (PKGs), cGMP-gated ion channels, and other downstream targets, which potentiate cellular responses. Importantly, these cGMP signaling cascades are involved in a wide array of diverse cellular processes including neurotransmission, vasodilation, intestinal homeostasis, and retinal phototransduction. These myriad functions underscore the near universal distribution of cGMP signaling in mammalian biology.

There are several members of the GC family of receptors, permitting significant structural and functional diversity. Importantly, these proteins all generally share a conserved dimeric catalytic domain, but are otherwise structurally diverse (Figure 1). Soluble GCs (sGCs) are cytoplasmic proteins that are responsive to intracellular nitric oxide (NO), a radical, which readily crosses the cell membrane to interface with this receptor. Particulate GCs (pGCs) are transmembrane receptors, primarily for extracellular ligands, that produce intracellular cGMP. These membrane-spanning proteins contain α helical transmembrane domains similar to those of the growth factor superfamily of receptors.

Although much of the structure and function of GCs was elucidated decades ago in a series of seminal studies, more recent work has uncovered novel physiological and pathophysiological roles for GCs in a variety of previously uncharacterized tissues and disease processes. Here, we summarize GC structure and function and provide an overview of the translation of this biology into understanding human physiology and disease and development of new therapeutics.

Soluble GCs

In the 1970s, a series of paradigm-shifting reports led to the identification of a novel signaling pathway consisting of NO-dependent induction of cGMP signaling through sGCs (Arnold *et al.*, 1977; Murad *et al.*, 1978). This original work was awarded the 1998 Nobel Prize in Physiology or Medicine and was focused primarily on the sGC-mediated signaling cascade as a mechanism of smooth muscle relaxation. This work resulted in clinical strategies for, as well as mechanistic insights into, the use of organic nitrates and nitrites to induce vasodilation for angina pectoris.

Six years after GCs were first described in 1969 (Hardman and Sutherland, 1969; White and Aurbach, 1969), efforts to characterize their subcellular location demonstrated enzymatic activity both in cytoplasmic as well as membrane-associated tissue homogenate fractions. The cytoplasmic GCs were named sGCs, whereas the membrane-bound GCs were termed pGCs. The catalytic activities of both particulate and soluble GCs are dependent on a dimerized structure (Figure 1). sGCs exist as heterodimers of homologous subunits designated as α and β . The most commonly studied isoform of sGC is the $\alpha1\beta1$

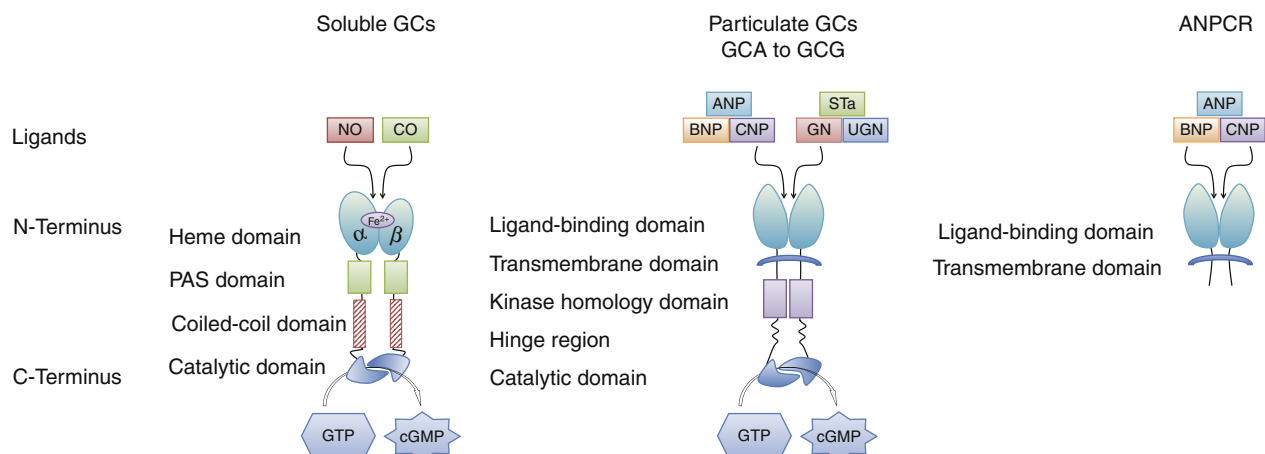


Figure 1 Mammalian GC domain structure. GCs can be subdivided into two broad groups: soluble and particulate. Soluble GCs are cytoplasmic proteins, which bind the intracellular ligands NO and CO and produce intracellular cGMP. In contrast, particulate GCs are transmembrane proteins, which bind extracellular ligands and transmit signals internally via intracellular cGMP production. ANPCR possess an extracellular natriuretic peptide-binding, but lacks intracellular catalytic domains, and acts primarily as a natriuretic peptide-clearance receptor.

protein, which is nearly ubiquitously expressed and possesses physiologically relevant ligand-binding and catalytic activity. However, the alternative isoforms $\alpha 2$ and $\beta 2$ also exist. The $\alpha 2$ subunit expression is limited to select tissues, including brain, lung, colon, heart, spleen, uterus, and placenta. The $\alpha 2\beta 1$ heterodimer resulting from this subunit has been shown to possess ligand-binding and catalytic functions similar to $\alpha 1\beta 1$. Interestingly, there is a splice variant of $\alpha 2$ known as $\alpha 2i$, which contains an in-frame insertion of 31 amino acids in the catalytic domain. This subunit is not only functionally inactive, but acts as a dominant-negative protein and is putatively involved in negative regulation of sGCs. The $\beta 2$ subunits have very limited expression, with messenger RNA (mRNA) studies localizing it mostly to the kidney. While the functions of $\beta 2$ subunits remain undefined, studies have shown that overexpression of $\beta 2$ results in catalytic activity when expressed in the absence of $\alpha 1$ and $\beta 1$ units, suggesting the possibility of homodimers of $\beta 2$ subunits.

Architecturally, each sGC subunit consists of four domains (Figure 1): an amino-terminal NO-binding heme-containing domain, a Per/Arnt/Sim (PAS) domain, a coiled-coil domain, and a carboxy-terminal catalytic domain. The different sGC α and β subunits isoforms possess more variability at the amino-terminus, whereas the carboxy-terminal catalytic domain is highly conserved, suggesting distinct regulatory functions amid conserved catalytic function.

The amino-terminal ligand-binding domain contains a heme prosthetic group that is indispensable to its function. Structurally, there is a histamine residue at position 105 that serves as a heme ligand in the $\beta 1$ subunit. This domain belongs to a larger family of proteins known as the heme-nitric oxide and oxygen binding proteins (H-NOX). In the context of sGCs, the heme domain can bind NO as well as carbon monoxide (CO). However, while other members of the H-NOX family exhibit O₂ binding, the heme domains in sGCs do not.

The carboxy-terminal domain of sGC is catalytic, and this domain in α and β subunits of sGC share significant homology, not only with one another, but also with pGCs. These catalytic domains have now been localized in sGC to residues 467–690 in the $\alpha 1$ subunit and 414–619 in the $\beta 1$ subunit. Several residues are critical to catalysis, reflecting their binding of Mg²⁺ ions, which act catalytically in several ways, activating both the 3'-hydroxyl and the α -phosphate in the guanine nucleotide, as well as stabilizing the β and γ phosphates on both substrate and product. Additional residues are important in orienting the ribose ring ($\beta 1$ N548), base recognition ($\beta 1$ E473 and C541), and interaction with nucleotide triphosphate ($\alpha 1$ R573 and $\beta 1$ R552).

The central region of sGC contains the PAS and coiled-coil domains. The exact function of these domains remains undefined, though they are generally thought to facilitate heterodimerization of sGC subunits. Indeed, PAS domains are regularly involved in protein-protein interactions, a function consistent with that hypothesis. Interestingly, the coiled-coil domain shares no homology with any known protein in the National Center for Biotechnology Information protein database. Despite the uncertainty regarding the central region (PAS and coiled-coil domains) of sGC, multiple reports have provided indirect evidence for the putative dimerization function of this region, an increasingly accepted model of its function.

NO binding to sGC is a well-studied interaction given its indispensability to sGC function. This interaction is fast (occurring in milliseconds) and potent with a 50% effective concentration (EC₅₀) of 45 nM. The heme prosthetic group of sGC associates with an axial histidine residue, H105, to form a five-member ring in the amino-terminal region, which binds NO. Upon NO binding, this H105 is displaced, and a new five-member ring is formed with the NO in the fifth position. CO gas also can bind and activate sGC, but with 100-fold lower affinity, and so the physiological role for this phenomenon is unclear.

Importantly, the source of sGC-activating NO is another class of enzyme known as nitric oxide synthase (NOS). Endothelial NOS (eNOS) is produced by endothelial cells and regulates cardiovascular homeostasis, including blood pressure. Indeed, eNOS-deficient mice exhibit a phenotype of hypertension. Another isoform, neuronal NOS (nNOS), also is involved in vasodilation in the nervous system. Interestingly, nNOS-deficient mice are characterized by increased aggression and abnormal pyloric sphincters, but have normal blood pressure. These findings suggest an involvement of the nNOS-sGC cascade in complex processes beyond vascular control in the nervous system.

The best-characterized function of sGC is the vasodilation of vascular smooth muscle cells (SMCs). In this process, eNOS is activated in endothelial cells, producing NO that diffuses to neighboring vascular SMCs. In SMCs, NO induces sGC production of cGMP, activating PKG, reducing levels of pro-contractile Ca²⁺ ions, as well as desensitizing cells to Ca²⁺. Modulation of Ca²⁺ levels is accomplished by increasing Ca²⁺ efflux, sequestering Ca²⁺ in the sarcoplasmic reticulum, and decreasing the mobilization of Ca²⁺. Reduced intracellular Ca²⁺ levels induce SMC relaxation leading to vasodilation.

Prior to elucidating the NO-sGC molecular mechanism of cross-talk between endothelium and vascular smooth muscle, organic nitrates which are enzymatically converted to NO in cells were exploited therapeutically in the clinical setting for more than 150 years to treat cardiovascular disease including angina pectoris. These clinical strategies have historically included nitroglycerine, but as knowledge of sGC and its associated vasodilating agents increased, so too did inquiry into additional clinical strategies for this pathway. Inhalational NO is now approved for treatment of pulmonary hypertension in infants, and other applications are being investigated. The list of potential pharmacologic agents influencing the sGC signaling cascade also has been expanded to include phosphodiesterases (PDEs) as targets, described below.

pGCs

Unlike sGC, which is comprised of a heterodimer of α and β subunits, pGCs exist and function as homodimers. Similarly, while sGC is activated upon binding NO, pGCs produce cGMP in response to extracellular peptide ligands (GCA, B and C) or intracellular calcium flux (GCE and F). Also, unlike sGC, seven different isoforms of pGCs exist in mammals (named GCA-GCG in order of discovery) and all but GCD and GCG are expressed in humans. GCA, GCB, and GCC are activated by the binding of specific peptide ligands to their amino-terminal extracellular domain (Figure 1). GCG and GCD genes have

been inactivated to become pseudogenes in humans and all primates, respectively. In rodents, GCD is expressed in olfactory neurons and may participate in sensing cues for hunger and thirst, while the tissue distribution, ligand, and function of GCG remains undefined. Like GCA-C, GCE and GCF also possess putative extracellular domains, but no extracellular ligands have been identified for these proteins. Instead, their activity is controlled by intracellular GC activating proteins (GCAPs) that sense light-induced changes in intracellular Ca^{2+} levels.

General Structure of pGCs

In 1989, the first pGC cloned was from sea urchin testis complementary DNA (cDNA), enabling subsequent screening for homologous genes in mammalian tissues. This resulted in the discovery of seven unique mammalian pGC genes, five of which are expressed in humans. At the structural level, these proteins are largely similar: all are composed of a single polypeptide chain of approximately 1050 amino acids, which encode the protein's ligand-binding domain, transmembrane domain, regulatory sites, and the catalytic domain (Figure 1). Slightly less than half of this sequence (about 500 residues) is devoted to the extracellular amino-terminal ligand-binding domain. This portion of the molecules exhibits the lowest degree of homology between the different pGCs, consistent with the observed differences in ligand-binding characteristics of the different isoforms. A domain of 21–25 hydrophobic amino acids spanning the plasma membrane is followed by an intracellular kinase homology domain (KHD), which is composed of approximately 250 amino acids. KHD does not possess kinase activity due to changes in residues necessary for kinase function, and instead plays a role in the regulation of pGC activity. pGC mutants lacking the KHD are constitutively active, and phosphorylation sites within this domain play an important role in ligand-induced desensitization of GCA and GCC.

Connecting the KHD to the catalytic domain is a sequence of about 70 amino acids called the hinge or dimerization domain. Mutations in this domain in GCE result in autosomal dominant cone-rod dystrophy, illustrating the importance of this domain in proper pGC function. The catalytic domain of pGC is responsible for mediating the synthesis of cGMP from guanosine triphosphate (GTP). This domain contains two active sites, each with three amino acids required for catalytic activity. Each active site utilizes an Asp from one monomer and an Arg/Asn pair from the other, underlying the dimer-dependence of GCs.

GCA

GCA, encoded on the first chromosome in humans, is expressed throughout the body, with particularly high expression in the vasculature, brain, liver, kidneys, adipose tissue, and lungs. Post-translational modifications include three intramolecular disulfide bonds, extracellular Asn glycosylation, and basal phosphorylation of six residues within the KHD. GCA is activated upon binding by any of the three natriuretic peptides (atrial natriuretic peptide (ANP), B-type natriuretic peptide (BNP), and C-type natriuretic peptide (CNP)). ANP and BNP are released by cardiac myocytes of the atria or ventricles,

respectively, in response to increased cardiac stretch, activating GCA to reduce cardiac load. This is achieved by inducing vasorelaxation, diuresis, and natriuresis, as well as by inhibiting the fluid retention-promoting renin-angiotensin-aldosterone system (RAAS). Genetic elimination of either GCA or ANP in mice produces cardiac hypertrophy and salt-resistant hypertension. Numerous efforts have been made to target the GCA signaling pathway with ANP peptide analogs for the treatment of hypertension, congestive heart failure, and renal disease. In 1995, a full-length analog of ANP, carperitide, was approved in Japan for the treatment of acute decompensated heart failure. However, the necessity to inject the drug limits its appeal, especially considering the numerous alternative oral antihypertensive agents. The development of a small-molecule drug targeting GCA could overcome this obstacle.

GCB

GCB, also known as natriuretic peptide receptor B, is encoded on the ninth chromosome in humans. Like GCA, it is heavily glycosylated and possesses three disulfide bonds, in addition to five phosphorylation sites within the KHD. It is expressed at high levels in the brain, heart, bones, ovaries, and lungs. The primary ligand for GCB is CNP, which, despite its name, does not induce natriuresis. Instead, CNP induces endochondral ossification and is required for proper long-bone development. Indeed, the impaired bone growth resulting from homozygous GCB loss causes a rare form of dwarfism called acromesomelic dysplasia, type Maroteaux. Furthermore, overstimulation of GCB results in bone overgrowth. Together, these data suggest that GCB may serve as a therapeutic target in diseases affecting skeletal growth.

GCC

Encoded on the 12th chromosome, GCC is expressed predominantly on the apical membrane of intestinal enterocytes, where its stimulation induces PKG-dependent phosphorylation and opening of the chloride channel cystic fibrosis transmembrane conductance regulator (CFTR). Open CFTR permits Cl^- diffusion down its concentration gradient out of the cell, into the intestinal lumen. Other ion channels and transporters maintain the electroneutrality of the cell by secreting Na^+ and water into the lumen. This system manifests as diarrhea during exposure to the enterotoxigenic *Escherichia coli* (ETEC)-produced GCC superagonist heat-stable enterotoxin (ST) during traveler's diarrhea. The endogenous ligands for GCC, guanylin and uroguanylin, also are expressed in the intestine. Uroguanylin is also produced by the kidney and is present in the circulation in its prohormone form. In 2012, an ST analog Linzess (linaclotide, Ironwood Pharmaceuticals) was approved by the FDA for the treatment of irritable bowel syndrome with constipation and chronic idiopathic constipation. In addition to maintaining fluid balance, GCC plays other homeostatic roles in the intestine, suppressing tumorigenesis by regulating cell cycle progression and reducing DNA damage, in addition to maintaining intestinal barrier integrity. As a result, linaclotide is also currently undergoing clinical trials for the prevention of colorectal

cancer. Interestingly, GCC also is expressed at low levels in the hypothalamus, where it mediates a gut–brain satiety axis. Indeed, GCC-deficient mice display increased susceptibility to weight gain and metabolic dysregulation in response to chronic high-fat diet, suggesting GCC ligands as a potential therapeutic for obesity, endemic in Western diets.

Retinal GCs

GCE and GCF are expressed primarily in photoreceptor cells of the retina, playing critical roles in phototransduction. These GCs possess no known extracellular peptide ligands. Rather, their activity is regulated by intracellular GCAPs during light/dark cycles. Exposure to light activates PDE 6, hydrolyzing cGMP to GMP. Reduced cGMP levels lead to closure of the cyclic nucleotide-gated channel, CNG1 inducing hyperpolarization of the photoreceptor cell, a key step in phototransduction. CNG1 closure also reduces Ca^{2+} influx, reducing intracellular Ca^{2+} levels, resulting in GCAP activation of retinal GCs to replenish intracellular cGMP levels, resetting the system.

The importance of retinal GCs in phototransduction is revealed by retinal diseases reflecting mutations in genes encoding GCE or GCF (*GUCY2D* and *GUCY2F* in humans, respectively). Leber congenital amaurosis type 1 (LCA-1) is a rare autosomal recessive, early onset retinal dystrophy caused by null mutations in the GCE gene (*GUCY2D*). Mutations in the *GUCY2D* locus also underlie dominant cone-rod dystrophy (CORD6), however, all mutations in CORD6 are restricted to the dimerization domain of GCE. CORD6 is a progressive disease, initiating with early loss of visual acuity and color vision following degeneration of cone photoreceptor cells. Later, rod photoreceptor cell degeneration

causes progressive night blindness and peripheral visual field loss.

ANP Clearance Receptor

ANP clearance receptor (ANPCR) is a truncated isoform of GCA encoded by the human *NPR3* gene that possesses an extracellular ligand-binding domain, but lacks all intracellular domains, including the cGMP catalytic domain (Figure 1). Thus, ANPCR binds ANP, BNP, and CNP, but produces no intracellular cGMP and acts primarily as a clearance receptor for the natriuretic peptides by constitutive endocytosis. However, ANPCR regulates a variety of physiological processes in human and rodent cells through a mechanism beyond natriuretic peptide clearance. ANPCR interacts with the heterotrimeric G proteins (G proteins) G_o and G_i and activates downstream processes through G_i . Importantly, this interaction is mediated by a short (about 25 amino acid) juxta-membrane region on the carboxy-terminus of ANPCR. This domain is conserved among pGCs and contains a consensus sequence that exists in other G protein-coupled single trans-membrane spanning receptors, suggesting that other pGCs may regulate G protein-mediated signaling cascades.

PDE

Intracellular levels of cGMP are regulated not only by its GC-mediated synthesis, but also by its removal, primarily by PDE-mediated hydrolysis (Figure 2), and to a lesser degree by export into the extracellular milieu or transit to an adjoining cell. The large PDE family of enzymes catalyzes hydrolysis of

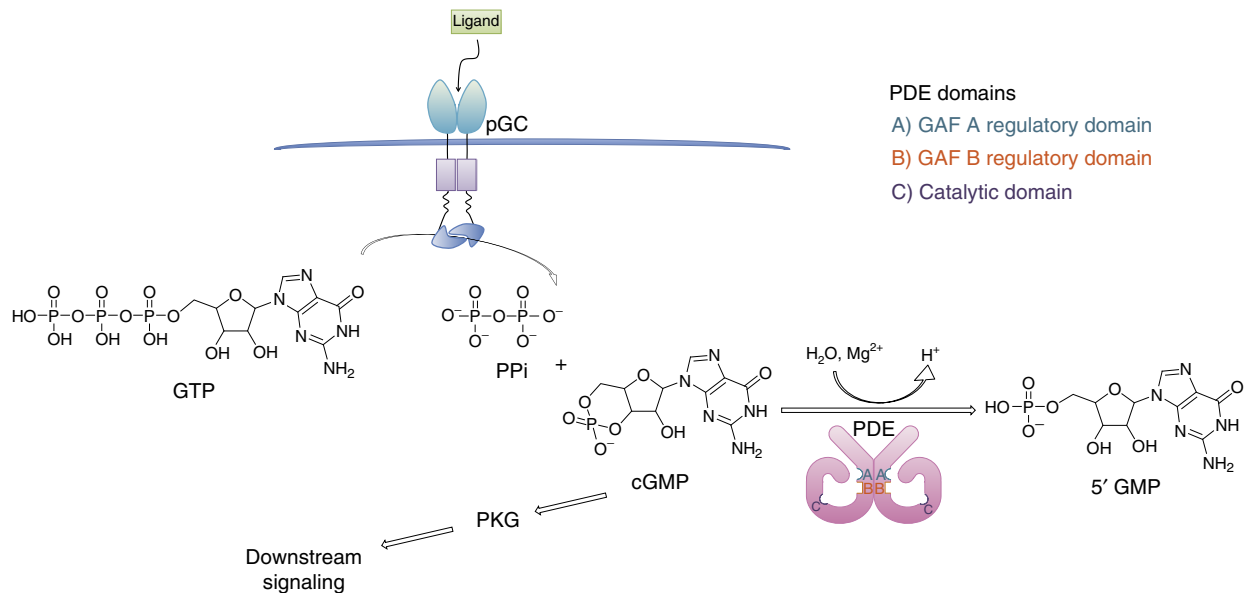


Figure 2 Intracellular cGMP regulation by GCs and PDEs. pGC receptor–ligand interaction catalyzes the conversion of GTP to cGMP, activating downstream signaling cascades via PKG. In contrast, PDEs catalyze the hydrolysis of the cGMP 3' cyclic phosphate bond, producing inactive 5' GMP from cGMP, returning cGMP levels to baseline. PDE structure is homologous to the α and β subunits of sGCs. Most cGMP-hydrolyzing PDEs are homodimeric enzymes with two catalytic centers (C) and tandem GAF-A and GAF-B regulatory domains (A and B, respectively). The C domain catalyzes the hydrolysis of cN, whereas the GAF domains play important roles in regulating PDE activity. GAF-A and GAF-B functions vary between the different PDE family members and some PDE family members lack GAF domains, but possess other regulatory elements, such as Ca^{2+} /CaM regulated domains.

the 3' cyclic phosphate bond of the cyclic nucleotides (cN) cyclic adenosine monophosphate (cAMP) and cGMP, resulting in its conversion into an inactive 5' monophosphate (Russell *et al.*, 1973; Cheung, 1970). The two classes of PDEs, class I and class II, differ in substrate selectivity. There are 11 families of class I PDEs, and each family has its own set of isomers and splice variants originating from 21 different gene products and more than 100 mRNA products. These PDEs form homodimers with unique structures, active sites, and functions. The commonality between all the isoenzymes is a 270 amino acid catalytic domain and heterogeneous regulatory domains.

The original classification of each PDE family was to organize them according to three characteristics: substrate specificity, mode of regulation, and elution order from a diethylaminoethyl (DEAE) ion-exchange column (Butcher and Sutherland, 1962). PDEs 1, 2, 3, 10, and 11 exhibit dual specificity and hydrolyze cAMP and cGMP, whereas PDEs 4, 7, and 8 selectively hydrolyze cAMP, and PDEs 5, 6, and 9 selectively hydrolyze cGMP. Each PDE possess a regulatory domain between the amino-terminus and catalytic domain, and these vary between the different isoforms. PDE1 has a Ca^{2+} /calmodulin (CaM) regulatory domain and is activated when four Ca^{2+} atoms cooperatively bind to CaM. PDEs 2, 5, 6, 10, and 11 have GAF (mammalian cGMP binding PDEs, *Anabaena* adenylyl cyclase, plant Fh1A transcription factors) domains, which regulate various aspects of PDE dimerization and activity. The name GAF is derived from the first three domains in which these were identified. PDE4 has two upstream conserved region (UCR1 and UCR2) domains. Specific isoforms of PDE4 can be phosphorylated by cAMP-dependent protein kinase (PKA), and UCR1 phosphorylation leads to a conformational change in the structure, increasing PDE4 activity. PDE8 has a PAS domain, the function of which is poorly defined, except that it is a site for ligand binding and protein interactions. All other PDEs have complements of various subdomains, such as UCR, neutralized homology repeat (NHR), PAS, and membrane association domains.

PDE isoenzymes are localized to various intracellular regions and in various tissues, allowing compartmentalized regulation of cGMP and cAMP. Like the structure, cellular location, and tissue specificity of PDEs, the downstream physiological effects of PDEs vary greatly. PDEs regulate cardiac functions, adrenal steroidogenesis, the male erectile response, and phototransduction, and are clinically associated with asthma, erectile dysfunction, chronic obstructive pulmonary disease (COPD), autoimmune diseases, hypertension, intermittent claudication, heart failure, schizophrenia, stroke, and depression. The goal of current clinical investigations inhibiting PDEs is to enhance the signaling cascade of cN to treat these associated maladies.

Two FDA-approved inhibitors of PDE3 are Primacor IV (milrinone), for the short-term treatment of acute chronic heart failure, and Pletal (cilostazol), for reduction of symptoms of intermittent claudication. In 2011, the PDE4 inhibitor Daliresp (roflumilast) was approved for the treatment of severe COPD associated with bronchitis, but not for patients with emphysema. Most recently (2014), Otezla (apremilast) was approved for the treatment of active psoriatic arthritis. The best known PDE inhibitors target PDE5 for erectile dysfunction, made popular with the approval of Viagra (sildenafil) in 1998. Since then, Cialis (tadalafil) and Levitra

(vardenafil), and more recently, a second-generation PDE5 inhibitor, Stendra (avanafil) in 2012, have been approved. In 2011, Cialis also was approved for the treatment of signs and symptoms of benign prostatic hyperplasia.

Importantly, the number of PDE-targeted pharmacological agents under development is increasing. PDE1 inhibitors have been developed, primarily due to its prevalence in specific regions of the brain, suggesting potential activity in targeting cognitive disorders. Two current PDE2 inhibitors are being investigated for activity in endothelial permeability and learning and memory. PDE3 has been linked to insulin and leptin signaling pathways, but no specific compounds have been identified. PDE4 inhibitors have been linked to improved long-term memory and PDE10 inhibitors have been indicated for the treatment of neurological diseases, with much focus on schizophrenia.

Conclusion

The second messenger cGMP plays critical roles in various physiologic and pathophysiologic processes from single-celled to multicellular organisms, including bacteria and humans. Intracellular levels of cGMP are tightly controlled by its synthesis and elimination, mediated by GCs and PDEs, respectively. Reflecting the broad cellular and tissue distribution of these enzymes and the myriad of processes they regulate, these enzymes are key targets for established and experimental therapeutics. However, many questions about these molecules and cGMP-dependent signaling remain unanswered. Indeed, regulation of satiety and obesity by an intestinal uroguanylin to hypothalamic GCC signaling axis was identified just 3 years ago, suggesting that many GC-mediated processes are yet to be discovered.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: G Protein-Coupled Receptors

References

- Arnold, W.P., Mittal, C.K., Katsuki, S., Murad, F., 1977. Nitric oxide activates guanylate cyclase and increases guanosine 3':5'-cyclic monophosphate levels in various tissue preparations. *Proceedings of the National Academy of Sciences of the United States of America* 74, 3203–3207.
- Butcher, R.W., Sutherland, E.W., 1962. Adenosine 3',5'-phosphate in biological materials. I. Purification and properties of cyclic 3',5'-nucleotide phosphodiesterase and use of this enzyme to characterize adenosine 3',5'-phosphate in human urine. *Journal of Biological Chemistry* 237, 1244–1250.
- Cheung, W.Y., 1970. Cyclic 3',5'-nucleotide phosphodiesterase. Demonstration of an activator. *Biochemical and Biophysical Research Communications* 38, 533–538.
- Hardman, J.G., Sutherland, E.W., 1969. Guanyl cyclase, an enzyme catalyzing the formation of guanosine 3',5'-monophosphate from guanosine triphosphate. *Journal of Biological Chemistry* 244, 6363–6370.
- Murad, F., Mittal, C.K., Arnold, W.P., *et al.*, 1978. Guanylate cyclase: Activation by azide, nitro compounds, nitric oxide, and hydroxyl radical and inhibition by hemoglobin and myoglobin. *Advances in Cyclic Nucleotide Research* 9, 145–158.
- Russell, T.R., Terasaki, W.L., Appleman, M.M., 1973. Separate phosphodiesterases for the hydrolysis of cyclic adenosine 3',5'-monophosphate and cyclic guanosine 3',5'-monophosphate in rat liver. *Journal of Biological Chemistry* 248, 1334–1340.
- White, A.A., Aurbach, G.D., 1969. Detection of guanyl cyclase in mammalian tissues. *Biochimica et Biophysica Acta — Bioenergetics* 191, 686–697.

Tumor Necrosis Factor Receptors: A Brief Digestion

Y Lin and BP Zhou, University of Kentucky College of Medicine, Lexington, KY, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

As early as the eighteenth century, it had been noticed that occasional tumor regression occurred in patients exposed to severe bacterial infection (Coley, 1991). It is, however, not until a few decades ago when we began to unveil the mechanism underlying this occurrence and really appreciate the function of tumor necrosis factors (TNF) and their receptors (TNFR). As indicated by their names, TNF/TNFR proteins are involved in the necrotic regression of tumor mass. TNF (formerly known as TNF α) and lymphotoxin (also referred to as TNF β) were successfully isolated from immune cells in 1984 (Gray *et al.*, 1984; Pennica *et al.*, 1984), with the protein coding sequences sharing 30% homology. Their specific receptors, TNFR1 and TNFR2, were cloned and expressed in 1990 (Dembic *et al.*, 1990; Loetscher *et al.*, 1990; Schall *et al.*, 1990; Smith *et al.*, 1990). In fact, TNF and lymphotoxin are two prototypic members of the TNF superfamily (TNFSF) consisting of 19 structurally related proteins; TNFR1 and TNFR2, in the meantime, belong to the TNF receptor superfamily (TNFRSF) harboring 29 members (Bodmer *et al.*, 2002; Locksley *et al.*, 2001). Since the crystal structure of TNF β -TNFR1 was revealed in 1993, there have been more than 35 specific ligand–receptor pairs identified (Bossen *et al.*, 2006). It eventually became evident that the functions of these proteins are far beyond the scope of tumor necrosis induction. Upon activation through specific ligand binding, TNFRs initiate distinct signaling pathways mediating cell survival, proliferation, differentiation, and apoptosis.

As an integral part of the immune system, TNFR signaling plays essential roles in numerous physiological events to regulate tissue development and homeostasis. Importantly, members of TNFSF and TNFRSF participate in organogenesis and maintenance of lymphoid microarchitecture, as well as induction of humoral immunoglobulin responses (Hehlgans and Pfeffer, 2005). During intracellular bacterial infections, specific TNFR pathways become activated to promote a series of inflammatory immune responses, including differentiation and cytokine secretion of T cells, activation and inducible nitric oxidase synthase (iNOS) production of macrophages, and recruitment of inflammatory cells (Hehlgans and Pfeffer, 2005). Beyond host defense, TNFSF and TNFRSF members are involved in bone and mammary gland homeostasis, hair follicle and sweat gland development, as well as neural development (Locksley *et al.*, 2001). Not surprisingly, deregulation of TNFR pathways is associated with autoimmune diseases, inflammatory diseases, and cancer. Studies have been focused on mechanistic clarification of the complicated TNFR signaling network, as well as evaluation of members of TNFSF and TNFRSF as therapeutic targets.

Classification of TNFRSF

TNFRSF members have originally been grouped according to sequence comparisons and genomic locations of their genes.

In humans, TNFRSF genes are dispersed across 14 chromosomes, with genes of several members involved in immune function residing in chr12p13 and its paralogous region chr1p36 (Locksley *et al.*, 2001). Functionally relevant receptors with their genes locating at the same chromosome clusters are speculated to derive from a common ancestor. Recently, a structure-based classification of TNFRSF was proposed. By using the T-Coffee structure-based MSAs (multiple sequence alignments) in combination with the T-RMSD (tree based on root mean square deviation) classification algorithm, TNFRSF members were divided into three groups, according to the patterns of their cysteine-rich domains (CRDs) (Magis *et al.*, 2012). Among the three groups, type S (for small) consists of small receptors with less conserved CRDs; type L-1 members are large receptors containing N-terminal CRDs, with highly conserved loop between cysteines 4 and 5; the third group is composed of large receptors with core CRDs away from the N-terminal and can be further divided into three subgroups L-2, L-3, and L-4 based on position of the core CRD within the receptor ectodomain architecture. Remarkably, this classification is in much stronger agreement with structural and functional features of TNFRSF members and may be applied to predict the convergent evolutionary process as well as interaction specificity of members of TNFSF and TNFRSF.

Structural Basis of Ligand–Receptor Interaction

Most of TNFSF members are type II transmembrane proteins with extracellular C-terminus and intracellular N-terminus. The last 150 residues of the C-terminus constitute the TNF homology domain (THD), which is responsible for receptor binding (Bodmer *et al.*, 2002). Upon proteolytic cleavage by metalloproteinases, the extracellular domain is released from the producer cell to generate mature cytokine protein in the form of soluble homotrimer. As to TNFRSF members, they largely consist of type I transmembrane proteins with extracellular N-terminus, while the remaining members are either classified as type III transmembrane proteins with extracellular N-terminus but without signal sequence (TNFR13B, TNFR13C, TNFR17, and XEDAR), attached to the membrane through a glycosylphosphatidylinositol (GPI) linker (TNFR10C), or exist in soluble forms due to proteolytic cleavage or alternative splicing (TNFR11B and TNFR6B) (Bodmer *et al.*, 2002).

Members of TNFRSF contain up to six CRDs, with three or four copies being the most frequent configurations (Locksley *et al.*, 2001). These hallmark pseudorepeats are characterized by highly conserved cysteine residues that form one to four intrachain disulfide bridges (Smith *et al.*, 1994). This unique structure enables receptor chains to fit into the trimeric ligand furrow. In fact, the ligand–receptor interaction presents a 3:3 scaffold, that is, the ligand trimer binds symmetrically to the extracellular region of three TNFR molecules (Banner *et al.*, 1993). It was originally inferred that the ligand ‘cross-linked’

three receptor monomers to form trimer. This view had been challenged by the findings that several TNFRs could oligomerize without presence of ligand (Chan *et al.*, 2000; Siegel *et al.*, 2000). Indeed, ligand-induced receptor trimerization would make possible the random assembly of 'mixed' receptor molecules, as the ligand–receptor recognition is not on a one-on-one basis, and this mess-up could interfere with downstream signaling (Chan, 2007). Of note, later findings indicated that TNFRs including TNFR1, TNFR2, TNFR5, TNFR6, TNFR10A, TNFR10B, and TNFR10D each contains a PLAD (pre-ligand binding assembly domain) within the first CRD, which favors ligand-independent, homotypic receptor assembly (Chan, 2000, 2007; Clancy *et al.*, 2005). This is, however, not a universal case, as TNFRs with only one or two CRDs do not require PLAD-mediated pre-assembling for ligand interaction (Hymowitz *et al.*, 2005). Pre-ligand assembly of TNFRs may hold advantage in binding ligand with higher affinity, circumventing potential signal interference, as well as preventing spontaneous receptor autoactivation.

Signal Transduction through TNFR Activation

Although the extracellular domains of TNFRs share limited homology, their intracellular regions are almost completely different, indicating their distinct biological functions. Indeed, signal transduction through TNFRs is initiated upon recruitment of various intracellular adaptor proteins. In most cases, the cytosolic tails of TNFRs contain either a 'death domain' (DD) or a TNF-receptor-associated factor (TRAF) interacting motif (Fesik, 2000; Inoue *et al.*, 2000). The DD is commonly present in a number of proteins regulating cell death; receptors containing DDs are defined as 'death receptors' (DR). These DRs are able to recruit DD-containing adaptor proteins through homotypic DD dimerization. DD-containing TNFRSF members include: TNFR1 (CD120a), FAS (CD95), TNF-related apoptosis inducing ligand receptor 1 (TRAIL-R1) (CD26, DR4), TRAIL-R2 (CD262, DR5), TRAIL-R4 (CD264, DCR2), TNF-like receptor apoptosis mediating protein (TRAMP) (DR3), and ectodysplasin (EDA) receptor (EDAR) (Hehlhans and Pfeffer, 2005). As the archetypal DR, TNFR1 dissociates from the inhibitory protein silencer of death domains (SODD) upon ligand binding and recruits TNF receptor-associated DD protein (TRADD), which subsequently recruits cellular inhibitor of apoptosis proteins 1 and 2 (c-IAP1 and 2), TRAF2, and receptor-interacting protein kinase 1 (RIP1). The E3 ligases TRAF2/cIAP can ubiquitinate RIP1, which then engages downstream signaling cascades to activate the pro-survival NF- κ B pathway (Li *et al.*, 2013). In addition, TRAF2 can activate ASK1, a mitogen-activated protein kinase kinase kinase (MAPKKK) and initiate a kinase cascade to eventually activate c-Jun N-terminal kinases (JNK) pathway mediating cell cycle progression and anti-apoptosis (Ichijo *et al.*, 1997). In a different cellular scenario when ubiquitination of RIP1 is removed, it tends to form complex with TRADD, Fas-associated DD protein (FADD) and pro-caspase 8 (Feoktistova *et al.*, 2011; Tenev *et al.*, 2011). It is believed that TNFR1 endocytosis and this secondary complex formation are inseparable events (Schneider-Brachert *et al.*, 2004). Within the new complex, pro-caspase 8 becomes activated to cleave and inactivate RIP1

and its family member RIP3, eventually leading to cell apoptosis characterized by chromatin condensation and fragmentation, membrane blebbing, and generation of apoptotic bodies (Micheau and Tschopp, 2003; Wang *et al.*, 2008; Bertrand *et al.*, 2008). Under certain physiological conditions such as viral infection, the functions of caspases are inhibited, and RIP1/RIP3 form the necrosome to initiate caspase-independent programmed necrosis, characterized by cell clumping and swelling, membrane rupture, and organelle degeneration (Cho *et al.*, 2009; He *et al.*, 2009; Zhang *et al.*, 2009). Taken together, TNFR1 is a pleiotropic receptor that may either lead to pro-survival NF- κ B and JNK activation or induce programmed cell death, determined by the assembly of specific higher-order signaling complexes under different conditions.

For the second group of TNFRSF members without DD, they transduce signal through the TRAF-interacting motif (TIM). These TNFRs include: TNFR2 (CD120b), LT β R, CD27, CD30, CD40, OX40 (CD134), 4-1BB (CD137), receptor activator of NF- κ B (RANK) (CD265), Fn14 (CD266, TWEAKR), transmembrane activator and calcium-signal modulating cyclophilin ligand (CAML) interactor (TACI) (CD267), BAFFR (CD268, BR3), B-cell maturation antigen (BCMA) (CD269), herpes virus entry mediator (HVEM) (CD270), activation induced TNF-receptor (AITR) (CD357), X-linked EDA-A2 receptor (XEDAR), and the member of the TNFR family TROY (Hehlhans and Pfeffer, 2005). Typically, these TNFRs use the cytoplasmic TIMs to interact with the highly conserved C-terminal TRAF domain of TRAF family members; TRAFs then recruit other signaling molecules to induce various downstream pathways. For example, activated TNFR2 associates with TRAF2, which then recruits TRAF1, c-IAP1 and c-IAP2, with subsequent signaling ending in NF- κ B activation (Rauert *et al.*, 2010; Marchetti *et al.*, 2004). Most frequently, TNFR2 signaling leads to cell survival and proliferation. However, it could induce different cellular events including differentiation and even apoptosis, depending on which genes NF- κ B transcribes.

Although different TNFRs may mediate signal transduction independently, functional overlap and even crosstalk between two certain pathways turned out to be not unusual, which is the case for TNFR1 and TNFR2. Both receptors are able to mediate cell survival/proliferation or apoptosis, although with different preferences. When these two receptors are co-expressed in the cells, there appears to be intracellular crosstalk between the two pathways. One tentative mechanism lies in that TNFR2 can modulate TNFR1 signaling to the apoptosis route through proteasomal degradation of TRAF2 and cIAPs (Fotin-Mlecsek *et al.*, 2002). However, overexpression of anti-apoptotic proteins such as NF- κ B, FLIP, Bcl-2 and Bcl-xL or downregulation of pro-apoptotic factors such as caspase 8 may render cells refractory to programmed cell death (Jaattela *et al.*, 1995; Irmeler *et al.*, 1997; Fulda *et al.*, 2001; Orłowski and Baldwin, 2002). Unlike TNFR1 that can be activated by both membrane-bound TNF (mTNF) and soluble TNF (sTNF), TNFR2 only responds to mTNF stimulation (Grell, 1995; Krippner-Heidenreich *et al.*, 2002). More intriguingly, normal TNF binding would inhibit rather than activate TNFR2 signaling, due to the presence of a TRAF2-binding site, T2bs-C, which locates at the C-terminal of TNFR2 and negatively

regulates TNFR2-TRAF2 binding and downstream signaling (Grech *et al.*, 2005). These unique characteristics of TNFR2 add additional layers of modulation and crosstalk with TNFR1. Overall, the functional consequences would depend on cell type and activation state, conformation and expression levels of TNFR1 and TNFR2, and availability of TNF and other critical signaling molecules.

The third group of TNFRSF members contain neither DD nor TIM, therefore cannot transduce signal within the cells. However, they can act as 'decoy' receptors to compete with the other two groups of TNFRSF for ligand binding. These TNFRs include TRAIL-R3 (CD263, DCR1), DCR3, and osteoprotegerin (OPG).

The Double-Edged Sword in Cancer: Signaling through TNFR1/2 as an Example

The role of TNFR signaling in cancer biology has been at the center of a long-term debate. As two of the most studied TNFRSF members, TNFR1 and TNFR2 can initiate signaling pathways leading to either tumor regression or its malignant progression. Apparently, there are several mechanisms underlying their anticancer activities (Bertazza and Mocellin, 2010). (1) Tumor vascular damage. TNF may promote the adhesion of tumor endothelial cells to leukocytes, and their subsequent migration into local areas of inflammation for clearance. In addition, TNF may suppress the activation of Integrin- α -V- β -3 selectively expressed by proliferating endothelium cells to induce apoptosis (Ruegg *et al.*, 1998). Furthermore, the high level of nitric oxide synthase (NOS) in tumor endothelial cells may sensitize them to TNF signals (Mocellin *et al.*, 2004). (2) Stimulation of innate immune system. As the specific ligand of TNFR1 and TNFR2, TNF can induce the production of other cytokines and cytotoxic factors by macrophages and natural killer (NK) cells (Vujanovic, 2001; Mocellin *et al.*, 2003). (3) Promoting adaptive immunity. TNF may mediate tumor cell rejection by T-lymphocytes, as well as promote the maturation of CD34⁺ myeloid cells into antigen-presenting dendritic cells (Prevost-Blondel *et al.*, 2000; Poehlein *et al.*, 2003; Baxeianis *et al.*, 2000; Figdor *et al.*, 2004).

In spite of a wealth of evidence supporting the anticancer activities, signaling through TNFR1 and/or TNFR2 may promote the development and progression of cancer in certain circumstances (Bertazza and Mocellin, 2010). Mechanistically, TNF may (1) function as an autocrine tumor growth factor to activate NF κ B and MAPKs, as mentioned in the earlier section; (2) upregulate cell cycle regulators such as Ras and c-Myc and downregulate Cdk inhibitors (Tselepis *et al.*, 2002); (3) cause DNA damage and inhibit DNA repair through enhancing the production of genotoxic molecules (Jaiswal *et al.*, 2000); (4) promote angiogenesis through upregulating factors such as VEGF, VEGFR2, and IL-8 (Yoshida *et al.*, 1997); (5) enhance tumor invasion through downregulation of c-Src (Kawai *et al.*, 2002); and (6) promote tumor metastasis possibly through downregulation of VCAM-1 (Kitakata *et al.*, 2002). Additionally, polymorphisms of TNF might affect its biological function and therefore cancer risk, although studies yet to obtain conclusive results.

Defining the roles of TNFR1/2 signaling in cancer remains a challenging task, due to the pleiotropic nature of the receptors,

the variety of survival/apoptotic pathway derangement in cancer cells, and the profound influence of tumor micro-environment. Further basic mechanistic studies and clinical trials are warranted to elucidate the function and regulation of this sophisticated signaling network.

Clinical Targeting of TNFSF and TNFRSF Members: Focusing on Cancer Immunotherapy

In 1992, the first TNF inhibitor entered clinical trial. Since then, a number of biologics targeting TNFSF and TNFRSF members have been developed (Croft *et al.*, 2013). For example, chimeric monoclonal antibody infliximab (Remicade; Centocor Ortho Biotech), human monoclonal antibodies adalimumab (Humira; Abbott), golimumab (Simponi; Centocor Ortho Biotech), denosumab (Prolia/Xgeva; Amgen), and belimumab (Benlysta; Human Genome Sciences/GlaxoSmithKline), as well as TNFR2-Fc fusion protein etanercept (Enbrel; Amgen/Pfizer) are clinically approved antagonists for treatment of autoimmune and inflammatory diseases. Biologics can also be specific antibodies lacking Fc region, such as Certolizumab pegol (Cimzia; UCB), a pegylated Fab fragment against TNF; and ESBA105, a humanized scFv against TNF. Other categories of biologics include naturally selected or engineered antibodies with enhanced cytotoxicity to eliminate pathogenic cells expressing specific TNFRSF members, RNA aptamers to TNFRSF molecules, and TNFSF ligand recombinant constructs.

The diverse biological functions of TNFR signaling ranging from selective induction of cell death to stimulation of effective immune responses make it a promising target for cancer immunotherapy. Earlier studies based on *in vitro* assay and mouse models showed that TNF has potent antitumor activity through the apoptotic TNFR1 pathway (Sugerman *et al.*, 1985; Old, 1988). However, since TNFR1 is universally expressed, systemic administration of TNF would cause severe toxicity due to its lack of tumor cell-binding selectivity. In the current clinical setting, TNF is administered only through locoregional delivery approaches to minimize systemic toxicity. This makes it impossible to assess the effectiveness of TNF on the overall patient survival, which depends on metastasis of tumor cells across the body that cannot be reached by locoregional treatments. In addition, signaling through TNFR1 may lead to either cell survival or apoptosis, adding another layer of uncertainty to use TNF as a single cancer therapy. Compared to TNFR1, TNFR2 is tightly regulated, with limited expression in certain immune cells, endothelial cells, and hematopoietic cells. Considering that TNFR2 activation favors the switch to the TNFR1 apoptosis pathway, and that soluble TNF can only trigger TNFR1 rather than TNFR2 signaling, TNFR2-selective TNF derivatives are developed as anticancer drugs. For example, TNC-scTNFR2 is a TNFR2 agonist designed by fusing the trimerization domain of tenascin C (TNC) to a TNFR2-specific single-chain TNF molecule to generate mTNF-like activity (Fischer *et al.*, 2011). In addition, scFv:TNF-based fusion proteins contain a TNFR2-selective TNF molecule and an antibody fragment that can specifically bind target antigen, usually associated with or overexpressed in tumor endothelial cells. Once the fusion protein becomes membrane-bound

upon antibody–antigen recognition, it will specifically activate TNFR2, which then sensitizes tumor endothelial cells to TNFR1 apoptotic signaling and results in destruction of tumor vasculature. This antibody-mediated delivery approach has been applied through targeting stromal (fibroblast activation protein (FAP)), endothelial (integrins) or cellular (EGFR and Her2) antigens (Wajant *et al.*, 2005). Similarly, derivatives of other TNFSF members, such as Fas-ligand (FasL) and TNF-related apoptosis inducing ligand (TRAIL), have also been developed (Bremer, 2013).

Furthermore, research has been focused on improving the anticancer activity of TNF. For example, various TNF analogues are designed to enhance receptor-binding selectivity; drug delivery systems are evaluated and optimized to minimize systemic toxicity without much compromising efficacy; TNF is administered in combination with NF κ B inhibitors to increase sensitivity; and TNF sensitizers such as nitric oxide and non-steroidal anti-inflammatory drugs (NSAID) are tested to increase anticancer activity (Bertazza and Mocellin, 2010). Intriguingly, evidence suggested that endogenous low-dose TNF chronically produced within the tumor microenvironment could promote cancer development and progression, in contrast to the anticancer activity of high-dose, single-shot TNF-based regimens. Therefore, upon administration of sTNF derivatives, its intracellular concentration should be carefully monitored.

On the basis of the tumor-promoting effects of TNF, the second category of therapy comprising TNF antagonists have also been developed for treatment of hematological malignancies, such as multiple myeloma, myelofibrosis, and acute myelogenous leukemia, as well as solid tumors, such as melanoma, renal cell carcinoma (RCC), and breast cancer (Du Bois *et al.*, 1997; Tsimberidou *et al.*, 2003; Szlosarek and Balkwill, 2003). Among them, humanized antibody RO5458640 against TWEAK is currently in Phase I clinical trial for treating solid tumors. Preliminary studies of etanercept on breast cancer patients have also been reported. Of the two categories of antagonists, ligand-specific antibodies bind to the antigenic epitope in the ectodomain of TNFs, whereas receptor-based antibodies bind directly to the receptor-binding region of TNFs, both of which being able to inhibit TNF–TNFR interaction.

The rationale for developing the third category of anticancer therapy is that many members of TNFRSF are costimulatory proteins that either promote B and T cell development and survival or potentiate immune responses. In this regard, agonists of these TNFRs may hold a promise as antitumor drugs. Following are a few examples. (1) CD137 (4–1BB): Monoclonal antibody urelumab (BMS-663513) has entered Phase I/II trial for the treatment of melanoma, RCC, and ovarian cancer. (2) CD134 (OX40): A mouse monoclonal antibody is in Phase I clinical trial; humanized antibodies and a fully human OX40L:IgG fusion protein have also been developed. Furthermore, OX40 agonists in combination with stereotactic radiation and/or cyclophosphamide are in Phase II clinical trial. (3) CD357 (GITR): Phase I trial of the humanized antibody TRX518 against melanoma is ongoing. (4) CD27: Human antibody SGN-75 is in Phase I trial for treating RCC and non-Hodgkin's lymphoma; CD27-expressing dendritic cells are evaluated in Phase I/II trial for treatment of

melanoma. (5) CD40: Human antibody CP-870893 is in Phase I trial treating melanoma and pancreatic cancer; chimeric antibody Chi Lob 7/4 is in Phase I trial targeting non-Hodgkin's lymphoma. (6) CD30: Chimeric antibody SGN-30 is in Phase II trial treating Hodgkin's lymphoma and Anaplastic large-cell lymphoma (ALCL). (7) CD26 (TRAILR1, DR4): Human antibody Mapatumumab/HGS-ETR1 is in Phase I trial targeting various cancers.

For the fourth category of anticancer therapy, antibodies are selected or engineered for enhanced cytotoxicity against tumor cells. Examples of these tumor cell-depleting biologics are: brentuximab vedotin (Adcetris; Seattle Genetics), a clinically approved CD30 specific chimeric antibody for treating ALCL and Hodgkin's lymphoma; SGN-75, MDX-1411 and MDX-1203, three humanized CD70 specific antibodies all in Phase I trials for treating RCC and non-Hodgkin's lymphoma; and lucatumumab/HCD122, a CD40 specific antibody in Phase I/II trial targeting various kinds of lymphoma, among others.

Collectively, a host of clinical and preclinical studies are underway to evaluate the targeting of specific TNFSF and TNFRSF members as immunotherapy of cancer, many of which having promising results. On the other side, activation of TNFR signaling may result in completely different cell fates, depending on cell types, metabolic status, and the surrounding microenvironment. Therefore, there is possibility that certain TNF/TNFR-based drug used to treat one type of cancer can increase the risk of developing another type of malignancy. Furthermore, a large part of TNFSF and TNFRSF members are not constitutively expressed, making both timing of treatment and target selection critical and challenging factors. Finally, challenges remain in inducing anticancer activities while avoiding increased susceptibility to infectious diseases or undesired inflammatory and autoimmune responses. It is advised that ubiquitous administration can have severe side effects, whereas tailored combination therapy may represent a safer and more efficient treatment strategy. All these factors must be taken into careful consideration when selecting TNF/TNFR targets as well as optimizing delivery systems.

Concluding Remarks

Challenging as it is, we are in a promising territory for the thorough appreciation and digestion of TNFR-mediated signaling with the help of contemporary molecular and cellular techniques in combination with bioinformatic approaches. With the expansion of our knowledge on the biology of TNFRSF and their ligands, the focus has widened from basic mechanistic research to therapeutic interventions at the clinical forefront. We envision the use of drugs targeting TNFSF/TNFRSF members as an important part of routine regimens for treatment of related diseases.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: Signaling and Function of Death Receptors of the Tumor Necrosis Factor Receptor Superfamily. Cell Division/Death: Apoptosis: Tumor Necrosis Factor Signaling Pathways

References

- Banner, D.W., D'Arcy, A., Janes, W., *et al.*, 1993. Crystal structure of the soluble human 55 kd TNF receptor-human TNF beta complex: Implications for TNF receptor activation. *Cell* 73, 431–445.
- Baxevas, C.N., Voutsas, I.F., Tsitsilonis, O.E., *et al.*, 2000. Compromised anti-tumor responses in tumor necrosis factor- α knockout mice. *European Journal of Immunology* 30, 1957–1966.
- Bertazza, L., Mocellin, S., 2010. The dual role of tumor necrosis factor (TNF) in cancer biology. *Current Medicinal Chemistry* 17, 3337–3352.
- Bertrand, M.J., Milutinovic, S., Dickson, K.M., *et al.*, 2008. cIAP1 and cIAP2 facilitate cancer cell survival by functioning as E3 ligases that promote RIP1 ubiquitination. *Molecular Cell* 30, 689–700.
- Bodmer, J.L., Schneider, P., Tschopp, J., 2002. The molecular architecture of the TNF superfamily. *Trends in Biochemical Sciences* 27, 19–26.
- Bossen, C., Ingold, K., Tardivel, A., *et al.*, 2006. Interactions of tumor necrosis factor (TNF) and TNF receptor family members in the mouse and human. *Journal of Biological Chemistry* 281, 13964–13971.
- Bremer, E., 2013. Targeting of the tumor necrosis factor receptor superfamily for cancer immunotherapy. *ISRN Oncology* 2013, 371854.
- Chan, F.K., 2000. The pre-ligand binding assembly domain: A potential target of inhibition of tumour necrosis factor receptor function. *Annals of the Rheumatic Diseases* 59 (Suppl. 1), i50–i53.
- Chan, F.K., 2007. Three is better than one: Pre-ligand receptor assembly in the regulation of TNF receptor signaling. *Cytokine* 37, 101–107.
- Chan, F.K., Chun, H.J., Zheng, L., *et al.*, 2000. A domain in TNF receptors that mediates ligand-independent receptor assembly and signaling. *Science* 288, 2351–2354.
- Cho, Y.S., Challa, S., Moquin, D., *et al.*, 2009. Phosphorylation-driven assembly of the RIP1-RIP3 complex regulates programmed necrosis and virus-induced inflammation. *Cell* 137, 1112–1123.
- Clancy, L., Mruk, K., Archer, K., *et al.*, 2005. Pre-ligand assembly domain-mediated ligand-independent association between TRAIL receptor 4 (TR4) and TR2 regulates TRAIL-induced apoptosis. *Proceedings of the National Academy of Sciences of the United States of America* 102, 18099–18104.
- Coley, W.B., 1991. The treatment of malignant tumors by repeated inoculations of erysipelas. With a report of ten original cases. 1893. *Clinical Orthopaedics and Related Research* 3–11.
- Croft, M., Benedict, C.A., Ware, C.F., 2013. Clinical targeting of the TNF and TNFR superfamilies. *Nature Reviews Drug Discovery* 12, 147–168.
- Dembic, Z., Loetscher, H., Gubler, U., *et al.*, 1990. Two human TNF receptors have similar extracellular, but distinct intracellular, domain sequences. *Cytokine* 2, 231–237.
- Du Bois, J.S., Trehu, E.G., Mier, J.W., *et al.*, 1997. Randomized placebo-controlled clinical trial of high-dose interleukin-2 in combination with a soluble p75 tumor necrosis factor receptor immunoglobulin G chimera in patients with advanced melanoma and renal cell carcinoma. *Journal of Clinical Oncology* 15, 1052–1062.
- Feoktistova, M., Geserick, P., Kellert, B., *et al.*, 2011. cIAPs block Ripoptosome formation, a RIP1/caspase-8 containing intracellular cell death complex differentially regulated by cFLIP isoforms. *Molecular Cell* 43, 449–463.
- Fesik, S.W., 2000. Insights into programmed cell death through structural biology. *Cell* 103, 273–282.
- Figdor, C.G., de Vries, I.J.M., Lesterhuis, W.J., Melief, C.J.M., 2004. Dendritic cell immunotherapy: Mapping the way. *Nature Medicine* 10, 475–480.
- Fischer, R., Maier, O., Siegemund, M., *et al.*, 2011. A TNF receptor 2 selective agonist rescues human neurons from oxidative stress-induced cell death. *PLoS One* 6, e27621.
- Fotin-Mlecsek, M., Henkler, F., Samel, D., *et al.*, 2002. Apoptotic crosstalk of TNF receptors: TNF-R2 induces depletion of TRAF2 and IAP proteins and accelerates TNF-R1-dependent activation of caspase-8. *Journal of Cell Science* 115, 2757–2770.
- Fulda, S., Kufer, M.U., Meyer, E., *et al.*, 2001. Sensitization for death receptor- or drug-induced apoptosis by re-expression of caspase-8 through demethylation or gene transfer. *Oncogene* 20, 5865–5877.
- Gray, P.W., Aggarwal, B.B., Benton, C.V., *et al.*, 1984. Cloning and expression of cDNA for human lymphotoxin, a lymphokine with tumour necrosis activity. *Nature* 312, 721–724.
- Grech, A.P., Gardam, S., Chan, T., *et al.*, 2005. Tumor necrosis factor receptor 2 (TNFR2) signaling is negatively regulated by a novel, carboxyl-terminal TNFR-associated factor 2 (TRAF2)-binding site. *Journal of Biological Chemistry* 280, 31572–31581.
- Grell, M., 1995. Tumor necrosis factor (TNF) receptors in cellular signaling of soluble and membrane-expressed TNF. *Journal of Inflammation* 47, 8–17.
- He, S., Wang, L., Miao, L., *et al.*, 2009. Receptor interacting protein kinase-3 determines cellular necrotic response to TNF- α . *Cell* 137, 1100–1111.
- Hehlgans, T., Pfeffer, K., 2005. The intriguing biology of the tumour necrosis factor/tumour necrosis factor receptor superfamily: Players, rules and the games. *Immunology* 115, 1–20.
- Hymowitz, S.G., Patel, D.R., Wallweber, H.J., *et al.*, 2005. Structures of APRIL-receptor complexes: Like BCMA, TACI employs only a single cysteine-rich domain for high affinity ligand binding. *Journal of Biological Chemistry* 280, 7218–7227.
- Ichijo, H., Nishida, E., Irie, K., *et al.*, 1997. Induction of apoptosis by ASK1, a mammalian MAPKKK that activates SAPK/JNK and p38 signaling pathways. *Science* 275, 90–94.
- Inoue, J., Ishida, T., Tsukamoto, N., *et al.*, 2000. Tumor necrosis factor receptor-associated factor (TRAF) family: Adapter proteins that mediate cytokine signaling. *Experimental Cell Research* 254, 14–24.
- Irmeler, M., Thome, M., Hahne, M., *et al.*, 1997. Inhibition of death receptor signals by cellular FLIP. *Nature* 388, 190–195.
- Jaattela, M., Benedict, M., Tewari, M., Shayman, J.A., Dixit, V.M., 1995. Bcl-x and Bcl-2 inhibit TNF and Fas-induced apoptosis and activation of phospholipase A2 in breast carcinoma cells. *Oncogene* 10, 2297–2305.
- Jaiswal, M., Larusso, N.F., Burgart, L.J., Gores, G.J., 2000. Inflammatory cytokines induce DNA damage and inhibit DNA repair in cholangiocarcinoma cells by a nitric oxide-dependent mechanism. *Cancer Research* 60, 184–190.
- Kawai, N., Tsuji, S., Tsujii, M., *et al.*, 2002. Tumor necrosis factor α stimulates invasion of Src-activated intestinal cells. *Gastroenterology* 122, 331–339.
- Kitakata, H., Nemoto-Sasaki, Y., Takahashi, Y., *et al.*, 2002. Essential roles of tumor necrosis factor receptor p55 in liver metastasis of intrasplenic administration of colon 26 cells. *Cancer Research* 62, 6682–6687.
- Krippner-Heidenreich, A., Tubing, F., Bryde, S., *et al.*, 2002. Control of receptor-induced signaling complex formation by the kinetics of ligand/receptor interaction. *Journal of Biological Chemistry* 277, 44155–44163.
- Li, J., Yin, Q., Wu, H., 2013. Structural basis of signal transduction in the TNF receptor superfamily. *Advances in Immunology* 119, 135–153.
- Locksley, R.M., Killeen, N., Lenardo, M.J., 2001. The TNF and TNF receptor superfamilies: Integrating mammalian biology. *Cell* 104, 487–501.
- Loetscher, H., Pan, Y.C., Lahm, H.W., *et al.*, 1990. Molecular cloning and expression of the human 55 kd tumor necrosis factor receptor. *Cell* 61, 351–359.
- Magis, C., van der Sloot, A.M., Serrano, L., Notredame, C., 2012. An improved understanding of TNF/TNFR interactions using structure-based classifications. *Trends in Biochemical Sciences* 37, 353–363.
- Marchetti, L., Klein, M., Schlett, K., Pfizenmaier, K., Eisel, U.L., 2004. Tumor necrosis factor (TNF)-mediated neuroprotection against glutamate-induced excitotoxicity is enhanced by N-methyl-D-aspartate receptor activation. Essential role of a TNF receptor 2-mediated phosphatidylinositol 3-kinase-dependent NF- κ B pathway. *Journal of Biological Chemistry* 279, 32869–32881.
- Micheau, O., Tschopp, J., 2003. Induction of TNF receptor I-mediated apoptosis via two sequential signaling complexes. *Cell* 114, 181–190.
- Mocellin, S., Provenzano, M., Lise, A., Nitti, D., Rossi, C.R., 2003. Increased TIA-1 gene expression in the tumor microenvironment after locoregional administration of tumor necrosis factor- α to patients with soft tissue limb sarcoma. *International Journal of Cancer* 107, 317–322.
- Mocellin, S., Provenzano, M., Rossi, C.R., *et al.*, 2004. Induction of endothelial nitric oxide synthase expression by melanoma sensitizes endothelial cells to tumor necrosis factor-driven cytotoxicity. *Clinical Cancer Research* 10, 6879–6886.
- Old, L.J., 1988. Tumor necrosis factor. *Scientific American* 258 (59–60), 69–75.
- Orlowski, R.Z., Baldwin Jr., A.S., 2002. NF- κ B as a therapeutic target in cancer. *Trends in Molecular Medicine* 8, 385–389.
- Pennica, D., Nedwin, G.E., Hayflick, J.S., *et al.*, 1984. Human tumour necrosis factor: Precursor structure, expression and homology to lymphotoxin. *Nature* 312, 724–729.
- Poelein, C.H., Hu, H.M., Yamada, J., *et al.*, 2003. TNF plays an essential role in tumor regression after adoptive transfer of perforin/IFN- γ double knockout effector T cells. *Journal of Immunology* 170, 2004–2013.
- Prevost-Blondel, A., Roth, e., Rosenthal, F.M., Pircher, H., 2000. Crucial role of TNF- α in CD8 T cell-mediated elimination of 3LL-A9 Lewis lung carcinoma cells *in vivo*. *Journal of Immunology* 164, 3645–3651.
- Rauert, H., Wicovsky, A., Muller, N., *et al.*, 2010. Membrane tumor necrosis factor (TNF) induces p100 processing via TNF receptor-2 (TNFR2). *Journal of Biological Chemistry* 285, 7394–7404.

- Ruegg, C., Yilmaz, A., Bieler, G., *et al.*, 1998. Evidence for the involvement of endothelial cell integrin $\alpha V\beta 3$ in the disruption of the tumor vasculature induced by TNF and IFN- γ . *Nature Medicine* 4, 408–414.
- Schall, T.J., Lewis, M., Koller, K.J., *et al.*, 1990. Molecular cloning and expression of a receptor for human tumor necrosis factor. *Cell* 61, 361–370.
- Schneider-Brachert, W., Tchikov, V., Neumeyer, J., *et al.*, 2004. Compartmentalization of TNF receptor 1 signaling: Internalized TNF receptosomes as death signaling vesicles. *Immunity* 21, 415–428.
- Siegel, R.M., Chan, F.K., Chun, H.J., Lenardo, M.J., 2000. The multifaceted role of Fas signaling in immune cell homeostasis and autoimmunity. *Nature Immunology* 1, 469–474.
- Smith, C.A., Davis, T., Anderson, D., *et al.*, 1990. A receptor for tumor necrosis factor defines an unusual family of cellular and viral proteins. *Science* 248, 1019–1023.
- Smith, C.A., Farrah, T., Goodwin, R.G., 1994. The TNF receptor superfamily of cellular and viral proteins: Activation, costimulation, and death. *Cell* 76, 959–962.
- Sugarman, B.J., Aggarwal, B.B., Hass, P.E., *et al.*, 1985. Recombinant human tumor necrosis factor- α : Effects on proliferation of normal and transformed cells *in vitro*. *Science* 230, 943–945.
- Szlosarek, P.W., Balkwill, F.R., 2003. Tumour necrosis factor α : A potential target for the therapy of solid tumours. *Lancet Oncology* 4, 565–573.
- Tenev, T., Bianchi, K., Darding, M., *et al.*, 2011. The Ripoptosome, a signaling platform that assembles in response to genotoxic stress and loss of IAPs. *Molecular Cell* 43, 432–448.
- Tselepis, C., Perry, I., Dawson, C., *et al.*, 2002. Tumour necrosis factor- α in Barrett's oesophagus: A potential novel mechanism of action. *Oncogene* 21, 6071–6081.
- Tsimberidou, A.M., Waddelow, T., Kantarjian, H.M., Albitar, M., Giles, F.J., 2003. Pilot study of recombinant human soluble tumor necrosis factor (TNF) receptor (p75) fusion protein (TNFR:Fc; Enbrel) in patients with refractory multiple myeloma: Increase in plasma TNF α levels during treatment. *Leukemia Research* 27, 375–380.
- Vujanovic, N.L., 2001. Role of TNF family ligands in antitumor activity of natural killer cells. *International Reviews of Immunology* 20, 415–437.
- Wajant, H., Gerspach, J., Pfizenmaier, K., 2005. Tumor therapeutics by design: Targeting and activation of death receptors. *Cytokine & Growth Factor Reviews* 16, 55–76.
- Wang, L., Du, F., Wang, X., 2008. TNF- α induces two distinct caspase-8 activation pathways. *Cell* 133, 693–703.
- Yoshida, S., Ono, M., Shono, T., *et al.*, 1997. Involvement of interleukin-8, vascular endothelial growth factor, and basic fibroblast growth factor in tumor necrosis factor α -dependent angiogenesis. *Molecular Cell Biol* 17, 4015–4023.
- Zhang, D.W., Shao, J., Lin, J., *et al.*, 2009. RIP3, an energy metabolism regulator that switches TNF-induced cell death from apoptosis to necrosis. *Science* 325 (5938), 332–336.

Signaling and Function of Death Receptors of the Tumor Necrosis Factor Receptor Superfamily

S Skeldal, Aarhus University Hospital, Denmark

EJ Coulson, The University of Queensland, Brisbane, QLD, Australia

© 2016 Elsevier Inc. All rights reserved.

Glossary

Apoptosome The apoptosome is a heptameric protein complex of the intrinsic apoptotic pathway, which drives the activation of pro-caspase-9, which in turn activates pro-caspase-3. The apoptosome, consisting of Apaf1, pro-caspase-9, and dATP, is assembled following cytochrome c release from mitochondria in response to low intracellular potassium concentrations, and Bak- and Bax-induced formation of mega-channels within the outer mitochondrial membrane.

Autoimmune disease An autoimmune disease develops as a result of an immune response toward healthy tissue and substances that are regarded as foreign. The condition may affect one or more types of tissue and can affect organ function and development.

Death domain A death domain is an ~80-amino-acid protein interaction domain consisting of a bundle of six antiparallel α -helices also known as a 'death fold.' Death domains have the capacity to oligomerize with each other, thereby promoting interactions between proteins such as TNF death receptors and FADD/TRADD adaptor proteins. A death domain is found in proteins involved in apoptosis and inflammation.

Death effector domain Similar to the death domain, the death effector domain consists of a bundle of six α -helices

that fold into a 'death fold.' Death effector domains interact with each other and are found in proteins such as FADD, pro-caspase-8 and pro-caspase-10, and cFLIP.

DISC Upon ligand binding and oligomerization of death receptors, their cytosolic death domains are brought into close proximity, creating a docking site for adaptor proteins such as FADD and pro-caspase-8. This multimeric protein death-inducing signaling complex (DISC) assembles at the cell surface and relays the extracellular apoptotic signal to the cell interior, thereby inducing downstream signaling cascades.

TRAF proteins TRAF proteins are cytosolic adaptor proteins capable of interacting with and transmitting signals from members of the TNF receptor superfamily. The TRAF proteins interact with TNF receptors via a conserved stretch of amino acids within the C-terminal domain of TRAF. Seven mammalian TRAF proteins have been described to date.

TRAPS A disease caused by a mutation within the TNFR1 extracellular domain causing aggregation and retention of TNFR1 within the cell, which in turn activates NF- κ B. TRAPS is characterized by periods of high fever as well as abdominal and joint pain, rash, and inflammation in various areas of the body.

Introduction to Cell Death

In metazoans, programmed cell death, or apoptosis, is crucial for eliminating excess cells. This process occurs during neural development, the acquisition of immunity, and the maintenance of tissue homeostasis (Strasser *et al.*, 2011). Apoptosis, or the lack of it, is also a feature of degenerative diseases and cancers, respectively. In vertebrates, apoptosis mainly occurs by one of two pathways named the intrinsic (mitochondrial) and extrinsic (receptor) pathways. The intrinsic pathway, which was first identified in *Caenorhabditis elegans*, is typically engaged in vertebrates following intracellular damage or growth factor deprivation and is executed following changes to the mitochondria. The extrinsic pathway is triggered by external factors and relies on death receptors (DRs) of the tumor necrosis factor (TNF) receptor superfamily (Croft *et al.*, 2012).

DRs, a Subset of the TNF Receptor Family

Receptors of the TNF receptor (TNFR) superfamily are characterized by the presence of one to four extracellular cysteine-rich domains and comprise a family of proteins with over 25 members, many of which are involved in immunological

functions. Within this superfamily is a subgroup of eight DRs thought to have evolved 350–450 million years ago, before the divergence of bony fish and tetrapods, with a single homologous receptor found in *Drosophila* (Croft *et al.*, 2012; Igaki and Miura, 2014). The eight DRs are Fas, tumor necrosis factor receptor 1 (TNFR1), p75 neurotrophin receptor (p75^{NTR}), DR 3, DR 4, DR 5, DR 6, and the ectodysplasin-A receptor (EDAR) (Figure 1), many of which are also known by a variety of other names (see Table 1) and fall into three homologous groups (Figure 1; Sessler *et al.*, 2013). In addition to having extracellular domain homology, each DR also contains a cytosolic ~80 amino acid so-called death domain, a protein–protein interaction module that is found in over 30 cell death-inducing proteins. However only four, including the tumor necrosis factor receptor type 1-associated death domain (TRADD), Fas-associated death domain protein (FADD), EDAR-associated death domain-containing adaptor protein (EDARADD), and receptor-interacting protein 1 (RIP1), bind directly to the death domain of TNFRs (Sessler *et al.*, 2013). Highlighting their functional importance, at least one DR, and often more, is expressed in every cell of the human body. However, the function of this receptor class is not solely to mediate cell death.

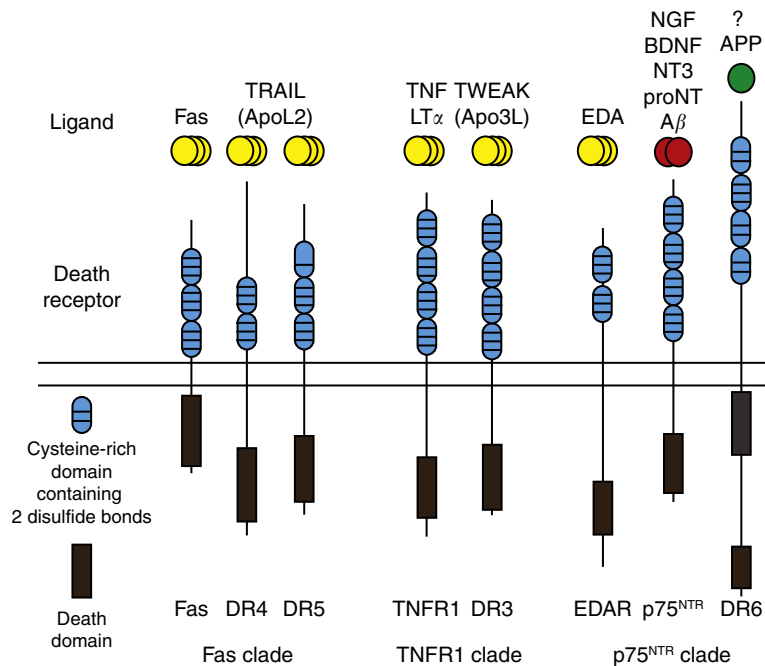


Figure 1 Members of the DR family and their ligand interactions. Not depicted are the decoy receptors DcR1 and DcR2, which both interact with TRAIL. p75^{NTR} also interacts with the pro-form of neurotrophins by forming a complex with the co-receptor sortilin or SorCS2. Domains are derived from NCBI protein sequence annotations (see also Bodmer, J.-L., Schneider, P., Tschopp, J., 2002. The molecular architecture of the TNF superfamily. *Trends in Biochemical Sciences* 27, 19–26).

Table 1 Commonly used and alternative names for DRs

Death receptor	Pseudonyms
Fas	Apo1, CD95, TNFRSF6, APT
DR4	TRAIL-R1, TNFRSF10A, CD261, APO2
DR5	TRAIL-R2, TNFRSF10B, CD262, KILLER, TRICK2, ZTNFR9
TNFR1	p55, CD120a
DR3	DDR3, Apo-3, TNFRSF25, TRAMP, LARD, WSL-1
EDAR	Anhidrotic ectodysplasin receptor 1, EDA-A1 receptor
p75 ^{NTR}	NGFR, LNGFR, TNFRSF16, p75(NTR), CD271, Gp80-LNGFR
DR7	TR-7

Ligand Activation of DRs Occurs by Receptor Oligomerization

The DRs are typically activated upon binding of ligands to their cysteine-rich repeats. The ligands are not always soluble molecules; they can also be membrane-bound and expressed on the surface of other cells, which activates the TNFRs *in trans*. Classically, upon binding ligands, the DRs oligomerize, bringing together the death domains, and recruiting other death domain-containing adaptor proteins, thereby facilitating signal transduction. Based on the trimeric crystal structure of TNFR1 and one of its ligands TNF β /lymphotoxin (LT) α (Banner *et al.*, 1993), the prevailing view is that activated TNFRs are trimers. However, this view is being challenged by more recent data. Unlike receptor kinases where dimerization

of two independent receptors is induced by ligands to activate the receptor–ligand complex, preformed inactive complexes exist at the cell surface for many of the TNFR family members (including TNFR1, Fas, DR4 and 5, and p75^{NTR}) (Chan *et al.*, 2000; Siegel *et al.*, 2000; Papoff *et al.*, 1999; Clancy *et al.*, 2005; Sykes *et al.*, 2012). Furthermore recent work has highlighted the fact that TNFRs may signal as symmetric dimers (p75^{NTR}), or as asymmetric trimers (DR5), with preformed receptor cluster rearrangement from trimer to dimer, or visa versa, being induced by ligand binding (Sessler *et al.*, 2013). Moreover, ligands may drive large-scale clustering of receptor complexes located within lipid microdomains into macrodomains to mediate signaling outcomes (Sessler *et al.*, 2013). There is currently no consensus on this issue, and more work needs to be performed to clarify the mechanism by which ligands result in receptor activation.

DR Signaling is Mediated by Protein Interactions

DR activation typically results in signal transduction pathways that are mediated by protein–protein interactions rather than phosphorylation cascades. Indeed phosphorylation of TNFR members has been relatively poorly studied, with receptor phosphorylation appearing to be more regulatory than required for function (Sessler *et al.*, 2013). Once receptors are clustered and activated, signal transduction occurs via the association of intracellular partners with binding sites within both the death domains and the other cytoplasmic domains of the DRs, thereby allowing additional signaling proteins to interact. Indeed, the variety of cytoplasmic interacting partners

underpins the capacity of DRs to activate multiple alternative signaling pathways, sometimes of opposing function (e.g., survival or proliferation versus cell death), depending on the cellular circumstance. Although each of the eight DRs can activate apoptosis, promoting cell death is not necessarily a receptor's main function. The pleiotropic biological outcomes of the DRs are further regulated by receptor internalization, ubiquitination, shedding/regulated intramembrane proteolysis, and other posttranslational modifications, including palmitoylation.

DR Signaling Activates Multiple Signal Pathways

This section will highlight the three signal transduction pathways frequently activated by DRs. The signaling and functional outcomes for each receptor are provided in more detail in the subsequent sections.

Extrinsic Apoptotic Pathway

The canonical apoptotic pathway (Figure 2(a)) is mediated by the binding of the death domain-containing protein FADD to the death domain of the DR Fas, DR4 or 5. In turn, pro-caspase-8 is recruited to the complex via the FADD death effector domain (DED), thereby forming the death-inducing signaling complex (DISC). DISC formation can autocatalyze the cleavage and thereby activation of caspase-8 (as well as pro-caspase-10 in humans) which activates downstream caspases such as caspase-3 and caspase-6. Caspases are endoproteases that cleave vital cellular proteins, and their caspase-mediated degradation triggers apoptotic processes, including membrane blebbing, nuclear condensation, and DNA fragmentation.

Caspase-8 can be inhibited by the inhibitor of apoptosis protein (IAP) and cellular inhibitor of apoptosis protein (cIAP).

Intrinsic Apoptotic Pathway

The second apoptotic pathway is the mitochondrial death pathway (Figure 2(b)), which is typically activated by environmental stressors. However, it can also be used by DRs as an amplification loop, and appears to be the main death signaling mechanism for p75^{NTR} (Murray *et al.*, 2003). This pathway is genetically conserved between mammals and worms (Metzstein *et al.*, 1998). Major regulators of the mitochondrial death pathway are the B-cell lymphoma 2 (Bcl-2) family of pro-apoptotic and anti-apoptotic proteins. In particular, proteolytic cleavage of the pro-apoptotic Bcl-2 protein Bid generates the truncated variant tBid, which translocates from the cytosol to the outer mitochondrial membrane, thereby promoting oligomerization of pro-apoptotic Bak and Bcl-2 associated X-protein (Bax). Oligomerized (Bcl-2 homologous antagonist/killer) Bak and Bax can subsequently form so-called mega-channels within the outer mitochondrial membrane, allowing the release of cytochrome c and second mitochondria-derived activator of caspases (SMAC)/DIABLO into the cytosol. Here, cytochrome c complexes with apoptotic protease activating factor 1 (Apaf1), dATP, and pro-caspase-9 forming the apoptosome. Within the apoptosome, pro-caspase-9 is activated, which then cleaves caspase-3, -6, -7. Released SMAC/DIABLO further facilitates apoptosis by suppressing the IAP X-linked inhibitor of apoptosis protein (XIAP), which otherwise inhibits the activity of caspase-9. Anti-apoptotic B-cell lymphoma-extra large (BclxL) is another inhibitor of the apoptosome. Bid can be cleaved by caspase-8

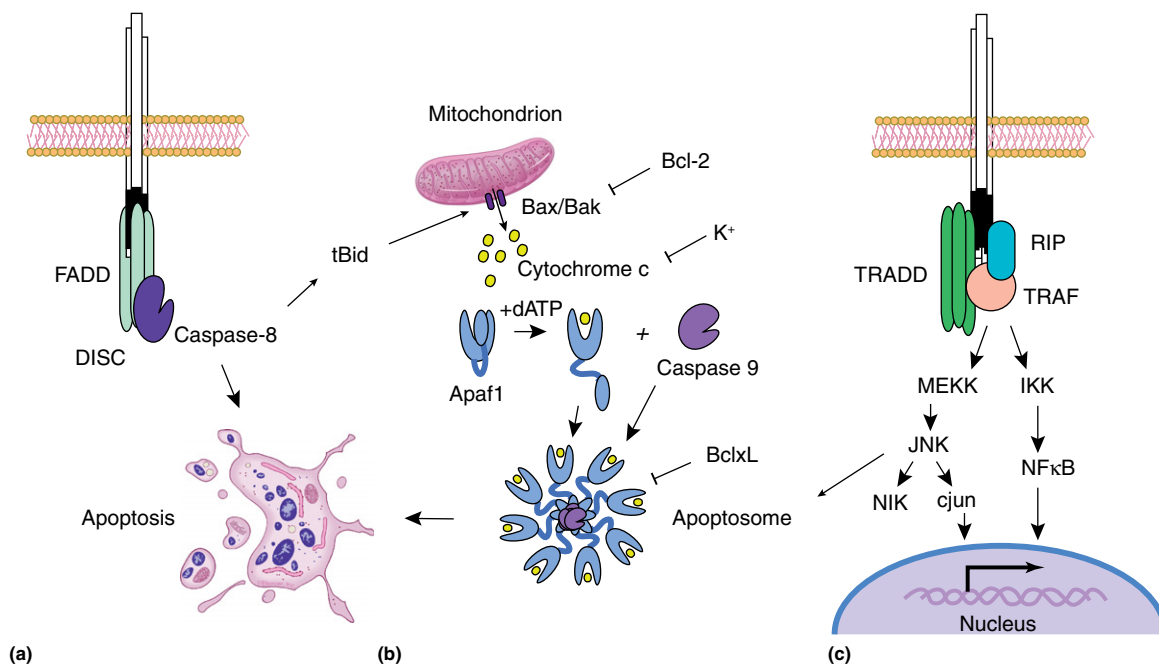


Figure 2 DR signaling pathways. (a) The canonical extrinsic apoptotic signaling pathway typically used by Fas. (b) The intrinsic mitochondrial death pathway. (c) Signaling via TRADD TNF receptor-associated factors (TRAFs) and RIP.

to trigger pathway activation, thereby initiating parallel apoptotic cascades.

Non-Apoptotic Pathways

The other major pathway activated by the DR family members promotes survival, differentiation, or inflammation by activating the mitogen-activated protein kinase (MAPK)/c-Jun N-terminal kinase (JNK) and/or nuclear factor κ B (NF- κ B) (Schneider-Brachert *et al.*, 2013). These pathways (Figure 2(c)) are also evolutionarily conserved from insects to humans (Igaki *et al.*, 2002). In mammals, the pathway can be activated by TNFR-associated factor (TRAF) family members, particularly TRAF 2, 4, and 6, which bind either directly to the death domain of the TNFRs or to death domain adaptor proteins. In turn, TRAF proteins bind to RIP kinases, which also contain a death domain (e.g., RIP1) or a caspase recruitment domain (RIP2). Association occurs between TRAF and/or RIP kinases with the inhibitor of kappa B ($\text{I}\kappa\text{B}$) kinase (IKK) via transforming growth factor β -activated kinase 1 (TAK1) (Wilson *et al.*, 2009). IKK activity regulates $\text{I}\kappa\text{B}$ ubiquitination and its subsequent degradation (Chung *et al.*, 2002; Zhang *et al.*, 2010), thereby allowing nuclear translocation of nuclear factor kappa-light-chain-enhancer of activated B cells (NF- κ B) to promote survival or inflammation. By a direct interaction, TRAF proteins can also regulate MEK kinases and thereby activate JNK, which promotes other cellular functions, and other regulatory kinases (e.g., NIK) (Chung *et al.*, 2002). TRAF protein signaling can be inhibited by cIAPs.

As a general rule, Fas, DR4, and 5 predominantly promote apoptosis of immune cells using FADD to bridge the activated DR with the cytosolic apoptotic machinery (Figure 2). TNFR1 and DR3 mainly control inflammatory responses by apoptotic and non-apoptotic pathways, which are mediated in the same way except that they use TRADD rather than FADD as the bridge (sometimes referred to as Complex I). DR6, EDAR, and p75^{NTR} are predominantly engaged in developmental processes, including apoptosis, and signal via the other conserved as well as receptor-specific pathways.

Function and Signaling Pathway of Each DR

The Fas Clade

Fas

Fas is widely expressed and induces apoptosis in a number of cell types, particularly within the immune system. It is important for killing pathogen-infected cells or immune cells remaining after an infection to protect against autoimmunity and tumor development. Mice homozygous for the naturally occurring lymphoproliferative (lpr) or generalized lymphoproliferative disease (gld) mutation suffer from a rearrangement of the Fas gene (Watanabe-Fukunaga *et al.*, 1992) or a point mutation in the Fas ligand gene (Takahashi *et al.*, 1994), respectively. Both conditions result in dysfunction of the Fas system, causing impaired negative selection of excess or self-reactive lymphocytes (Ramaswamy *et al.*, 2009). Accordingly, lpr and gld mice develop lpr autoimmune syndrome which manifests as lymphadenopathy, splenopathy,

and accumulation of CD4⁺, CD8⁺ T cells (Strasser *et al.*, 2009). lpr autoimmune disorders have also been found in human patients with a disruption of the Fas gene (Rieux-Laucat *et al.*, 1995) and Fas knockout mice suffer from liver hyperplasia (Adachi *et al.*, 1995).

Fas also promotes apoptosis in invading cells in immune privileged sites such as the brain, eye, placenta, and fetus. Here, Fas ligand-expressing cells induce apoptosis in Fas-expressing infiltrating cells and thereby avoid an inflammatory immunologic response (Strasser *et al.*, 2009).

As the Fas signaling system is involved in immune surveillance, Fas ligand-expressing T cells and natural killer cells suppress tumor development in Fas-bearing tumor cells. In line with this, older lpr and gld mice exhibit increased susceptibility to the development of plasmacytoid tumors (Ehrenschrwender and Wajant, 2009). However, the Fas ligand has also been identified in a variety of human tumors and is thought to play a role in the protection of cancer cells from Fas-bearing immune cells in a similar manner to that observed at immune privileged sites (Whiteside, 2007). Thus, the Fas system plays a dual role in tumor development and metastasis.

The apoptotic signal of Fas is executed upon *trans* binding of homotrimeric Fas ligand. Trimerization of Fas receptors and their intracellular death domains creates a docking site for FADD, which is required for induction of the extrinsic caspase-8-dependent apoptotic pathways (Strasser *et al.*, 2009). Fas DISC formation relies not only on Fas trimerization but also on further clustering of Fas receptor trimers (Scott *et al.*, 2009; Wang *et al.*, 2010) to induce a concomitant structural change of the Fas death domain, which allows FADD to engage DISC (Scott *et al.*, 2009). Furthermore, this pathway is only induced by membrane-bound Fas ligand (La *et al.*, 2009), possibly mediated by receptor capping into focal points within the cell membrane. This could occur by a two-step process (Grassme *et al.*, 2003). Low levels of caspase-8 activation result in enrichment of ceramide in lipid microdomains due to lysosomal acidic sphingomyelinase activity, which could lead to receptor capping within macrodomains due to the merging of the ceramide-rich microdomains. In support of this model, Fas palmitoylation (the posttranslational addition of a lipid to the receptor) is required for Fas-induced cell death (Chakrabandhu *et al.*, 2007).

Whereas the canonical death pathway is activated by Fas in so-called type-I cells (e.g., T cells) to mediate apoptosis, for apoptosis to occur in type-II cells (e.g., hepatocytes) the pathway requires amplification by the intrinsic mitochondrial pathway (Peter and Krammer, 2003; Strasser *et al.*, 2009). Thus the extrinsic and intrinsic apoptotic pathways are interconnected in type-II cells, with partial activation of caspase-8 being sufficient to cleave Bid and trigger the amplification loop. However, Fas is also capable of promoting a non-apoptotic response following Fas ligand binding in situations where the apoptotic pathway is blocked. In this paradigm, Fas has been shown to activate NF- κ B via FADD, TRAF2, and RIP1 (Kreuz *et al.*, 2004).

DR4 and DR5

The two DRs DR4 and DR5 belong to the same homology clade as Fas and promote apoptosis via a similar pathway to Fas (Figure 3(a)). In humans, the TNF-related apoptosis-inducing

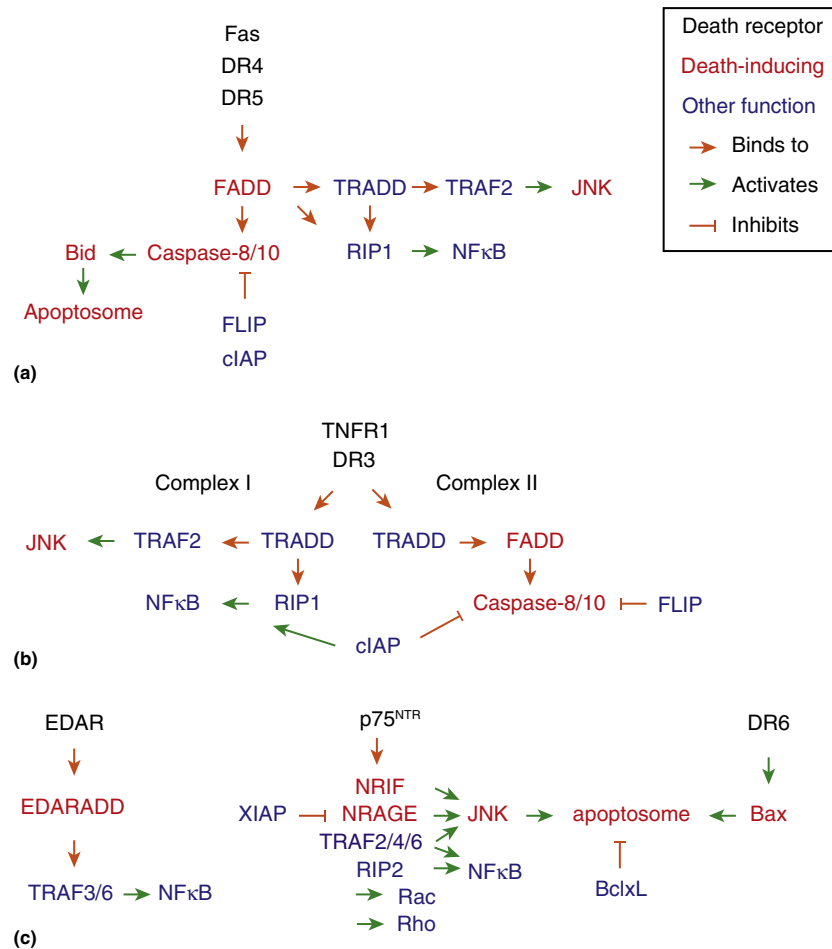


Figure 3 Overview of the binding partners and downstream signaling pathways initiated by the Fas clade (a), the TNFR1 clade (b), and the p75NTR clade (c).

ligand (TRAIL) interacts with DR4 and DR5, which in turn directly interact with FADD and cause DISC assembly. Hence, activation of DR4 or DR5 can mediate both the extrinsic and intrinsic apoptotic signaling pathways (Sessler *et al.*, 2013). Signaling through DR4 and DR5 has been suggested to modulate immune responses to avoid reaction to self-reactive antigens and kill virally infected cells (Gonzalez and Ashkenazi, 2010; Collison *et al.*, 2009). Furthermore, mDR5 (the mouse ortholog to human DR4 and DR5) and TRAIL are known to suppress tumor growth and metastasis by specifically promoting apoptosis among cancer cells while sparing non-transformed cells (Grosse-Wilde and Kemp, 2008; Wu, 2009). However, recent work has indicated that DR4 and DR5 also have the capacity to activate non-apoptotic signaling pathways via the MAPK, JNK, and NF-κB pathways (Collison *et al.*, 2009; Gonzalez and Ashkenazi, 2010).

TRAIL also interacts with the two so-called decoy-receptors DcR1 (TRAIL-R3) and DcR2 (TRAIL-R4), which are both incapable of transmitting an apoptotic signal. Whereas DcR1 does not possess a cytoplasmic or transmembrane domain and is instead attached to the membrane via a glycosylphosphatidylinositol-anchor, DcR2 has a truncated and non-functional death domain and has been reported to promote activation of NF-κB (Schneider-Brachert *et al.*, 2013). However,

as the decoy receptors (DcRs) are capable of binding TRAIL they reduce DR4 and DR5 apoptotic signaling by competing for ligand binding.

TNFR1 Clade

TNFR1

Due to the widespread expression of TNFR1 and its pleiotropic nature of signaling (see below; Figure 3(b)), the physiological roles of the TNF signaling system are numerous. TNFR1 plays a major role in maintaining immune homeostasis by promoting apoptosis, cell survival, differentiation, and inflammation. Early studies with mice deficient in TNFR1 revealed a prominent role of the receptor in fighting bacterial pathogens; mice devoid of TNFR1 exhibit increased susceptibility to infection by *Listeria monocytogenes* and *Mycobacterium tuberculosis* (Rothe *et al.*, 1993; Flynn *et al.*, 1995). In line with a crucial role of TNFR1 in modulating immune responses, dysregulation of TNFR1 signaling is thought to cause several autoimmune diseases. For example, a number of mutations within TNFR1 are associated with apparent autoimmune diseases collectively known as TNF receptor-associated periodic syndrome (TRAPS), which are characterized by inflammatory attacks of variable duration (Lobito *et al.*, 2011). Dysregulation of TNF

also causes autoimmune diseases such as psoriasis, autoimmune arthritis, multiple sclerosis, and diabetes mellitus. The majority of these diseases are thought to relate to the ability of TNF to promote expression of proinflammatory genes via the NF- κ B signaling pathway (Sethi *et al.*, 2009). The proinflammatory nature of TNF signaling is also thought to contribute to cancer. Although some studies support a protective role of TNFR1 and TNF against cancer, the majority of studies have shown that TNF has the ability to promote tumor development, proliferation, invasion, angiogenesis, and metastasis (Sethi *et al.*, 2009). TNFR1 has also been shown to be involved in neuronal activity in the brain, beyond mediating inflammation (Haase *et al.*, 2008). For example, a recent study has suggested a developmental role of TNFR1 during the development of sensory neurons known as nociceptors, allowing proper perception of pain (Wheeler *et al.*, 2014).

TNFR1 is activated by binding of homotrimers of TNF or lymphotoxin- α (LT α). At the cell surface, TNFR1 recruits TRADD, which also harbors a death domain similar to that found in FADD. TRADD then recruits RIP1 and TRAF2, collectively comprising what is known as Complex I (Figure 3(b); Micheau and Tschopp, 2003). Complex I induces survival and inflammatory responses by activating the MAPK and JNK pathways. Complex I further promotes survival and proliferation by activating NF- κ B via the pro-survival kinase TAK1 (Bertrand *et al.*, 2008). Activation of NF- κ B also requires ubiquitination of RIP1 by the E3 ubiquitin ligases cIAP1 and 2 (O'donnell *et al.*, 2007; Bertrand *et al.*, 2008).

However, upon receptor internalization, FADD-mediated death signaling can also be triggered. As TRADD and FADD both interact with TNFR1 at the same site, TRADD (TRAF2 and RIP1) must dissociate from TNFR1, thereby allowing access to FADD (now known as Complex II), at which point TNFR1 promotes apoptosis rather than survival (Wilson *et al.*, 2009; Lee *et al.*, 2012). Furthermore, dissociation of TRADD can be regulated by a conformational change in the receptor (allowing formation of Complex IIA) or by ubiquitination of RIP1 (Complex IIB). The activity of complex IIA is regulated by the caspase-8 inhibitor cellular FLICE inhibitory protein (cFLIP) whereas that of Complex IIB is regulated by the levels of cIAP1 and 2 (Wang *et al.*, 2008). One hypothesis as to why TNFR1/DR3 mediate a non-apoptotic response when DISC formation occurs, whereas apoptosis mediated by DISC formation is stimulated by Fas or DR4/5, is because the latter receptors trigger a global cellular response, whereas DISC formation following TNFR1 internalization allows for compartmentalization, restricting caspase activity. Depending on the context, TNFR1-mediated DISC assembly can also alternatively promote cell survival, proliferation or differentiation as described above. The cellular outcome critically depends on the constituents of DISC and on the subcellular localization of activated TNFR1 (Schneider-Brachert *et al.*, 2013).

DR3

As the receptor with highest homology to TNFR1, DR3 signaling pathways are similarly relayed through TRADD, RIP, and TRAF2 (Meylan *et al.*, 2011). Although DR3 is capable of executing an apoptotic response, promoting apoptosis among thymocytes (Wang *et al.*, 2001a), it plays an important modulatory role in immune responses through the formation of

Complex I, by activating the MAPK, JNK, and NF- κ B signaling pathways following binding of its ligand TL1A. Specifically, DR3 co-stimulates T cells to produce a number of cytokines and promotes expansion of activated and regulatory T cells (Meylan *et al.*, 2011). It is also involved in T-cell-dependent immune responses that fight viruses and bacteria (Buchan *et al.*, 2012; Twohig *et al.*, 2012).

p75^{NTR} Clade

EDAR

The EDAR gene was identified in association with the condition ectodermal dysplasia, which is also caused by a mutation in the EDARADD (Headon *et al.*, 2001; Monreal *et al.*, 1999). The EDAR gene, which contains a polymorphism common in Han Chinese populations, is required for development of the ectoderm that gives rise to skin, hair, nails, teeth, and mammary and sweat glands, and has been most widely studied in this developmental context (Sadier *et al.*, 2014). EDAR does not interact with TRADD or FADD but, via EDARADD, signals through the TRAF family members 3 and 6 to activate JNK and NF- κ B (Figure 3(c)), thereby regulating the gene transcription required for the development of ectoderm appendages (Kumar *et al.*, 2001; Morlon *et al.*, 2005).

p75^{NTR}

p75^{NTR} is a receptor for the neurotrophin family of growth factors and mediates a wide range of cellular outcomes, predominantly in the nervous system. The best-known function of p75^{NTR} is to promote apoptosis in the developing nervous system as neurons undergo selection during the period of synaptogenesis (Miller and Kaplan, 2001). It also causes neurodegeneration after injury or in neurological conditions, including motor neuron disease and Alzheimer's disease (Haase *et al.*, 2008).

p75^{NTR} mediates cell death by the mitochondrial death pathway, as it is inhibited in cells from Apaf1- and Bax-deficient animals, and is blocked by BclxL and intracellular potassium (which inhibits cytochrome c activation of the apoptosome (Figure 2(c); Roux and Barker, 2002; Coulson *et al.*, 2004; Coulson *et al.*, 2008). Like Fas, p75^{NTR} can generate ceramide, and its cell death signaling requires lipid-rich regions of the membrane and the palmitoylation of the receptor (Dobrowsky *et al.*, 1994; Underwood *et al.*, 2008).

However, like EDAR, p75^{NTR} does not interact with FADD or TRADD (Figure 3(c)), although a dominant-negative TRADD mutant was found to inhibit p75^{NTR}-mediated NF- κ B activation (Wang *et al.*, 2001b; Sessler *et al.*, 2013). Unlike the death domains in other DRs, the p75^{NTR} death domain does not self-associate and cannot substitute for the Fas death domain, having a different tertiary structure to that of the other family members (Liepinsh *et al.*, 1997; Kong *et al.*, 1999).

Nonetheless, p75^{NTR} signals via protein-protein interactions, including binding to RIP2 and TRAF2, 4, and 6 (Roux and Barker, 2002). In addition, it interacts with NRIF and NAGE, adaptor proteins that activate JNK to promote cell death (Roux and Barker, 2002; Charalampopoulos *et al.*, 2012). Alternatively, p75^{NTR} can bind to RIP2 to activate NF- κ B, and it can also activate RhoA, to produce actin remodeling

and neurite retraction, as well as Rac1, to mediate a range of cellular outcomes. The binding sites within the p75^{NTR} death domain that underpin these various signaling pathways have recently been mapped (Charalampopoulos *et al.*, 2012).

However, the signaling pathways of p75^{NTR} are determined not only by the expressed cellular components but also by its receptor complex and ligands. Firstly, it forms a co-receptor complex with tropomyosin-related kinase (Trk) receptors, increasing their affinity for the mature neurotrophin ligands and promoting survival signaling (Skeldal *et al.*, 2011). Secondly p75^{NTR} forms a receptor–pro-neurotrophin complex together with sortilin family members (sortilin and SorCS2) to signal atrophic functions (Glerup *et al.*, 2014), and, thirdly, it combines with Nogo and the Nogo receptor to mediate neurite retraction (Kraemer *et al.*, 2014). p75^{NTR} also binds the amyloid precursor protein-derived amyloid beta peptide which induces cell death (Yaar *et al.*, 1997; Coulson *et al.*, 2009).

p75^{NTR} is further distinguished from the wider TNFR members as it uses regulated intramembrane proteolysis to regulate its signaling (Skeldal *et al.*, 2011). While regulation of cell death appears to be at least partially dependent on receptor cleavage, the ability of p75^{NTR} to regulate neurogenesis (Young *et al.*, 2007) and neural progenitor cell migration (e.g., neural crest migration) also appears to be particularly dependent on this modification (Skeldal *et al.*, 2011). In peripheral tissues, p75^{NTR} has been demonstrated to be tumorigenic (e.g., breast cancer), and to promote metastasis, specifically following intramembrane cleavage (e.g., glioma) (Johnston *et al.*, 2007; Forsyth *et al.*, 2014). However in other cases, p75^{NTR} acts as a tumor suppressor by mediating cell death in cancer cells (Molloy *et al.*, 2011).

DR6

As for other DRs, DR6 is involved in modulation of immune responses. The DR6 knockout mouse exhibits increased proliferation of CD4⁺ T cells, suggesting that DR6 is a negative regulator of CD4⁺ T cell proliferation during immune responses (Liu *et al.*, 2001). However, recent research has revealed that it also acts as a negative regulator of cell survival of neural cells such as oligodendrocytes and motor neurons during development (Nikolaev *et al.*, 2009; Mi *et al.*, 2011) and in motor neuron disease (Huang *et al.*, 2013). DR6 further regulates axonal pruning by promoting axonal degeneration (Nikolaev *et al.*, 2009; Marik *et al.*, 2013). One ligand of DR6 appears to be the amyloid precursor protein itself (Nikolaev *et al.*, 2009), which is expressed both in the brain and in peripheral tissues including blood. The signaling pathway emanating from DR6 is not well defined (Figure 3(c)), but like others in this clade it does not appear to involve caspase-8 as its central apoptotic messenger; rather DR6 triggers a caspase-dependent pathway that includes translocation of Bax to the mitochondria (Zeng *et al.*, 2012).

Summary

The DRs use three major conserved signaling pathways to mediate their functions. These are the activation of a canonical extrinsic cell death cascade, activation of the intrinsic cell death cascade, and activation via intracellular binding proteins including TRAF, RIP proteins, and transcription factors such as

NF-κB. The structural basis of activation of individual signaling proteins by these protein–protein interactions is still far from resolved. Furthermore, despite conservation of signaling mechanisms, the DRs mediate a wide array of cellular functions, often within the same cell types, which depends on the cellular environment in which the receptor finds itself. Thus regulation of DR signals is mediated by a range of post-translational changes including palmitoylation and higher order clustering of the receptors, as well as the expression of downstream signaling proteins coupled with activation by proteolysis, and/or inhibition by ubiquitination. This is an emerging area of DR research, and we encourage further exploration of this topic on a receptor-by-receptor basis.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: Tumor Necrosis Factor Receptors: A Brief Digestion. Cell Division/Death: Apoptosis: Tumor Necrosis Factor Signaling Pathways

References

- Adachi, M., Suematsu, S., Kondo, T., *et al.*, 1995. Targeted mutation in the Fas gene causes hyperplasia in peripheral lymphoid organs and liver. *Nature Genetics* 11, 294–300.
- Banner, D.W., D'Arcy, A., Janes, W., *et al.*, 1993. Crystal structure of the soluble human 55 kd TNF receptor–human TNF beta complex: Implications for TNF receptor activation. *Cell* 73, 431–445.
- Bertrand, M.J., Milutinovic, S., Dickson, K.M., *et al.*, 2008. cIAP1 and cIAP2 facilitate cancer cell survival by functioning as E3 ligases that promote RIP1 ubiquitination. *Molecular Cell* 30, 689–700.
- Buchan, S.L., Taraban, V.Y., Slebiada, T.J., *et al.*, 2012. Death receptor 3 is essential for generating optimal protective CD4(+) T-cell immunity against *Salmonella*. *European Journal of Immunology* 42 (3), 580–588.
- Chakrabandhu, K., Herincs, Z., Huault, S., *et al.*, 2007. Palmitoylation is required for efficient Fas cell death signaling. *EMBO Journal* 26, 209–220.
- Chan, F.K., Chun, H.J., Zheng, L., *et al.*, 2000. A domain in TNF receptors that mediates ligand-independent receptor assembly and signaling. *Science* 288, 2351–2354.
- Charalampopoulos, I., Vicario, A., Padiaditakis, I., *et al.*, 2012. Genetic dissection of neurotrophin signaling through the p75 neurotrophin receptor. *Cell Reports* 2, 1563–1570.
- Chung, J.Y., Park, Y.C., Ye, H., Wu, H., 2002. All TRAFs are not created equal: Common and distinct molecular mechanisms of TRAF-mediated signal transduction. *Journal of Cell Science* 115, 679–688.
- Clancy, L., Mruk, K., Archer, K., *et al.*, 2005. Pre-ligand assembly domain-mediated ligand-independent association between TRAIL receptor 4 (TR4) and TR2 regulates TRAIL-induced apoptosis. *Proceedings of the National Academy of Sciences of the United States of America* 102, 18099–18104.
- Collison, A., Foster, P.S., Mattes, J., 2009. Emerging role of tumour necrosis factor-related apoptosis-inducing ligand (TRAIL) as a key regulator of inflammatory responses. *Clinical and Experimental Pharmacology and Physiology* 36, 1049–1053.
- Coulson, E.J., May, L.M., Osborne, S.L., *et al.*, 2008. p75 Neurotrophin receptor mediates neuronal cell death by activating GIRK channels through phosphatidylinositol 4,5-bisphosphate. *Journal of Neuroscience* 28, 315–324.
- Coulson, E.J., May, L.M., Sykes, A.M., Hamlin, A.S., 2009. The role of the p75 neurotrophin receptor in cholinergic dysfunction in Alzheimer's disease. *Neuroscientist* 15, 317–323.
- Coulson, E.J., Reid, K., Shipham, K.M., *et al.*, 2004. The role of neurotransmission and the chopper domain in p75 neurotrophin receptor death signaling. *Progress in Brain Research* 146, 41–62.
- Croft, M., Duan, W., Choi, H., *et al.*, 2012. TNF superfamily in inflammatory disease: Translating basic insights. *Trends in Immunology* 33, 144–152.

- Dobrowsky, R.T., Werner, M.H., Castellino, A.M., *et al.*, 1994. Activation of the sphingomyelin cycle through the low-affinity neurotrophin receptor. *Science* 265, 1596–1599.
- Ehrenschrwender, M., Wajant, H., 2009. The role of FasL and Fas in health and disease. *Advances in Experimental Medicine and Biology* 647, 64–93.
- Flynn, J.L., Goldstein, M.M., Chan, J., *et al.*, 1995. Tumor necrosis factor- α is required in the protective immune response against *Mycobacterium tuberculosis* in mice. *Immunity* 2, 561–572.
- Forsyth, P.A., Krishna, N., Lawn, S., *et al.*, 2014. p75 neurotrophin receptor cleavage by α - and γ -secretases is required for neurotrophin-mediated proliferation of brain tumor-initiating cells. *Journal of Biological Chemistry* 289, 8067–8085.
- Glerup, S., Nykjaer, A., Vaegter, C.B., 2014. Sortilins in neurotrophic factor signaling. *Handbook of Experimental Pharmacology* 220, 165–189.
- Gonzalez, F., Ashkenazi, A., 2010. New insights into apoptosis signaling by Apo2L/TRAIL. *Oncogene* 29, 4752–4765.
- Grassme, H., Cremesti, A., Kolesnick, R., Gulbins, E., 2003. Ceramide-mediated clustering is required for CD95-DISC formation. *Oncogene* 22, 5457–5470.
- Grosse-Wilde, A., Kemp, C.J., 2008. Metastasis suppressor function of tumor necrosis factor-related apoptosis-inducing ligand-R in mice: Implications for TRAIL-based therapy in humans? *Cancer Research* 68, 6035–6037.
- Haase, G., Pettmann, B., Raoul, C., Henderson, C.E., 2008. Signaling by death receptors in the nervous system. *Current Opinion in Neurobiology* 18, 284–291.
- Headon, D.J., Emmal, S.A., Ferguson, B.M., *et al.*, 2001. Gene defect in ectodermal dysplasia implicates a death domain adaptor in development. *Nature* 414, 913–916.
- Huang, G., Lee, X., Bian, Y., *et al.*, 2013. Death receptor 6 (DR6) antagonist antibody is neuroprotective in the mouse SOD1G93A model of amyotrophic lateral sclerosis. *Cell Death & Disease* 4, e841.
- Igaki, T., Kanda, H., Yamamoto-Goto, Y., *et al.*, 2002. Eiger, a TNF superfamily ligand that triggers the *Drosophila* JNK pathway. *EMBO Journal* 21, 3009–3018.
- Igaki, T., Miura, M., 2014. The *Drosophila* TNF ortholog Eiger: Emerging physiological roles and evolution of the TNF system. *Seminars in Immunology* 26, 267–274.
- Johnston, A.L., Lun, X., Rahn, J.J., *et al.*, 2007. The p75 neurotrophin receptor is a central regulator of glioma invasion. *PLOS Biology* 5, e212.
- Kong, H., Kim, A.H., Orlicki, J.R., Chao, M.V., 1999. A comparison of the cytoplasmic domains of the Fas receptor and the p75 neurotrophin receptor. *Cell Death & Disease* 6, 1133–1142.
- Kraemer, B.R., Yoon, S.O., Carter, B.D., 2014. The biological functions and signaling mechanisms of the p75 neurotrophin receptor. *Handbook of Experimental Pharmacology* 220, 121–164.
- Kreuz, S., Siegmund, D., Rumpf, J.J., *et al.*, 2004. NF κ B activation by Fas is mediated through FADD, caspase-8, and RIP and is inhibited by FLIP. *Journal of Cell Biology* 166, 369–380.
- Kumar, A., Eby, M.T., Sinha, S., *et al.*, 2001. The ectodermal dysplasia receptor activates the nuclear factor- κ B, JNK, and cell death pathways and binds to ectodysplasin A. *Journal of Biological Chemistry* 276, 2668–2677.
- La, O.R., Tai, L., Lee, L., *et al.*, 2009. Membrane-bound Fas ligand only is essential for Fas-induced apoptosis. *Nature* 461, 659–663.
- Lee, E.W., Seo, J., Jeong, M., *et al.*, 2012. The roles of FADD in extrinsic apoptosis and necroptosis. *BMB Reports* 45, 496–508.
- Liepinsh, E., Ilag, L.L., Otting, G., Ibanez, C.F., 1997. NMR structure of the death domain of the p75 neurotrophin receptor. *EMBO Journal* 16, 4999–5005.
- Liu, J., Na, S., Glasebrook, A., *et al.*, 2001. Enhanced CD4+ T cell proliferation and Th2 cytokine production in DR6-deficient mice. *Immunity* 15, 23–34.
- Lobito, A.A., Gabriel, T.L., Medema, J.P., Kimberley, F.C., 2011. Disease causing mutations in the TNF and TNFR superfamilies: Focus on molecular mechanisms driving disease. *Trends in Molecular Medicine* 17, 494–505.
- Marik, S.A., Olsen, O., Tessier-Lavigne, M., Gilbert, C.D., 2013. Death receptor 6 regulates adult experience-dependent cortical plasticity. *Journal of Neuroscience* 33, 14998–15003.
- Metzstein, M.M., Stanfield, G.M., Horvitz, H.R., 1998. Genetics of programmed cell death in *C. elegans*: Past, present and future. *Trends in Genetics* 14, 410–416.
- Meylan, F., Richard, A.C., Siegel, R.M., 2011. TL1A and DR3, a TNF family ligand-receptor pair that promotes lymphocyte costimulation, mucosal hyperplasia, and autoimmune inflammation. *Immunological Reviews* 244, 188–196.
- Mi, S., Lee, X., Hu, Y., *et al.*, 2011. Death receptor 6 negatively regulates oligodendrocyte survival, maturation and myelination. *Nature Medicine* 17, 816–821.
- Micheau, O., Tschopp, J., 2003. Induction of TNF receptor I-mediated apoptosis via two sequential signaling complexes. *Cell* 114, 181–190.
- Miller, F.D., Kaplan, D.R., 2001. Neurotrophin signalling pathways regulating neuronal apoptosis. *Cellular and Molecular Life Sciences* 58, 1045–1053.
- Molloy, N.H., Read, D.E., Gorman, A.M., 2011. Nerve growth factor in cancer cell death and survival. *Cancers (Basel)* 3, 510–530.
- Monreal, A.W., Ferguson, B.M., Headon, D.J., *et al.*, 1999. Mutations in the human homologue of mouse dl cause autosomal recessive and dominant hypohidrotic ectodermal dysplasia. *Nature Genetics* 22, 366–369.
- Morlon, A., Munnich, A., Smahi, A., 2005. TAB2, TRAF6 and TAK1 are involved in NF- κ B activation induced by the TNF-receptor, Edar and its adaptor Edaradd. *Human Molecular Genetics* 14, 3751–3757.
- Murray, S.S., Bartlett, P.F., Lopes, E.C., *et al.*, 2003. Low-affinity neurotrophin receptor with targeted mutation of exon 3 is capable of mediating the death of axotomized neurons. *Clinical and Experimental Pharmacology and Physiology* 30, 217–222.
- Nikolaev, A., McLaughlin, T., O'leary, D.D., Tessier-Lavigne, M., 2009. APP binds DR6 to trigger axon pruning and neuron death via distinct caspases. *Nature* 457, 981–989.
- O'donnell, M.A., Legarda-Addison, D., Skountzou, P., *et al.*, 2007. Ubiquitination of RIP1 regulates an NF- κ B-independent cell-death switch in TNF signaling. *Current Biology* 17, 418–424.
- Papoff, G., Hausler, P., Eramo, A., *et al.*, 1999. Identification and characterization of a ligand-independent oligomerization domain in the extracellular region of the CD95 death receptor. *Journal of Biological Chemistry* 274, 38241–38250.
- Peter, M.E., Krammer, P.H., 2003. The CD95(APO-1/Fas) DISC and beyond. *Cell Death & Differentiation* 10, 26–35.
- Ramaswamy, M., Cleland, S.Y., Cruz, A.C., Siegel, R.M., 2009. Many checkpoints on the road to cell death: Regulation of Fas-FasL interactions and Fas signaling in peripheral immune responses. *Results and Problems in Cell Differentiation* 49, 17–47.
- Rieux-Laucat, F., Le Deist, F., Hivroz, C., *et al.*, 1995. Mutations in Fas associated with human lymphoproliferative syndrome and autoimmunity. *Science* 268, 1347–1349.
- Rothe, J., Lesslauer, W., Lotscher, H., *et al.*, 1993. Mice lacking the tumour necrosis factor receptor 1 are resistant to TNF-mediated toxicity but highly susceptible to infection by *Listeria monocytogenes*. *Nature* 364, 798–802.
- Roux, P.P., Barker, P.A., 2002. Neurotrophin signaling through the p75 neurotrophin receptor. *Progress in Neurobiology* 67, 203–233.
- Sadier, A., Viriot, L., Pantalacci, S., Laudet, V., 2014. The ectodysplasin pathway: From diseases to adaptations. *Trends in Genetics* 30, 24–31.
- Schneider-Brachert, W., Heigl, U., Ehrenschrwender, M., 2013. Membrane trafficking of death receptors: Implications on signalling. *International Journal of Molecular Sciences* 14, 14475–14503.
- Scott, F.L., Stec, B., Pop, C., *et al.*, 2009. The Fas-FADD death domain complex structure unravels signalling by receptor clustering. *Nature* 457, 1019–1022.
- Smessler, T., Healy, S., Samali, A., Szegezdi, E., 2013. Structural determinants of DISC function: New insights into death receptor-mediated apoptosis signalling. *Pharmacology & Therapeutics* 140, 186–199.
- Sethi, G., Sung, B., Kunnumakkara, A.B., Aggarwal, B.B., 2009. Targeting TNF for Treatment of Cancer and Autoimmunity. *Advances in Experimental Medicine and Biology* 647, 37–51.
- Siegel, R.M., Frederiksen, J.K., Zacharias, D.A., *et al.*, 2000. Fas preassociation required for apoptosis signaling and dominant inhibition by pathogenic mutations. *Science* 288, 2354–2357.
- Skeldal, S., Matusica, D., Nykjaer, A., Coulson, E.J., 2011. Proteolytic processing of the p75 neurotrophin receptor: A prerequisite for signalling? Neuronal life, growth and death signalling are crucially regulated by intra-membrane proteolysis and trafficking of p75(NTR). *Bioessays* 33, 614–625.
- Strasser, A., Cory, S., Adams, J.M., 2011. Deciphering the rules of programmed cell death to improve therapy of cancer and other diseases. *EMBO Journal* 30, 3667–3683.
- Strasser, A., Jost, P.J., Nagata, S., 2009. The many roles of FAS receptor signaling in the immune system. *Immunity* 30, 180–192.
- Sykes, A.M., Palstra, N., Abankwa, D., *et al.*, 2012. The effects of transmembrane sequence and dimerization on cleavage of the p75 neurotrophin receptor by γ -secretase. *Journal of Biological Chemistry* 287, 43810–43824.
- Takahashi, T., Tanaka, M., Brannan, C.I., *et al.*, 1994. Generalized lymphoproliferative disease in mice, caused by a point mutation in the Fas ligand. *Cell* 76, 969–976.
- Twohig, J.P., Marsden, M., Cuff, S.M., *et al.*, 2012. The death receptor 3/TL1A pathway is essential for efficient development of antiviral CD4(+) and CD8(+) T-cell immunity. *FASEB Journal* 26, 3575–3586.

- Underwood, C.K., Reid, K., May, L.M., *et al.*, 2008. Palmitoylation of the C-terminal fragment of p75(NTR) regulates death signaling and is required for subsequent cleavage by gamma-secretase. *Molecular and Cellular Neuroscience* 37, 346–358.
- Wang, E.C., Thern, A., Denzel, A., *et al.*, 2001a. DR3 regulates negative selection during thymocyte development. *Molecular and Cellular Biology* 21, 3451–3461.
- Wang, L., Du, F., Wang, X., 2008. TNF-alpha induces two distinct caspase-8 activation pathways. *Cell* 133, 693–703.
- Wang, L., Yang, J.K., Kabaleeswaran, V., *et al.*, 2010. The Fas-FADD death domain complex structure reveals the basis of DISC assembly and disease mutations. *Nature Structural & Molecular Biology* 17, 1324–1329.
- Wang, X., Bauer, J.H., Li, Y., *et al.*, 2001b. Characterization of a p75(NTR) apoptotic signaling pathway using a novel cellular model. *Journal of Biological Chemistry* 276, 33812–33820.
- Watanabe-Fukunaga, R., Brannan, C.I., Copeland, N.G., *et al.*, 1992. Lymphoproliferation disorder in mice explained by defects in Fas antigen that mediates apoptosis. *Nature* 356, 314–317.
- Wheeler, M.A., Heffner, D.L., Kim, S., *et al.*, 2014. TNF-alpha/TNFR1 signaling is required for the development and function of primary nociceptors. *Neuron* 82, 587–602.
- Whiteside, T.L., 2007. The role of death receptor ligands in shaping tumor microenvironment. *Immunological Investigations* 36, 25–46.
- Wilson, N.S., Dixit, V., Ashkenazi, A., 2009. Death receptor signal transducers: Nodes of coordination in immune signaling networks. *Nature Immunology* 10, 348–355.
- Wu, G.S., 2009. TRAIL as a target in anti-cancer therapy. *Cancer Letters* 285, 1–5.
- Yaar, M., Zhai, S., Pilch, P.F., *et al.*, 1997. Binding of beta-amyloid to the p75 neurotrophin receptor induces apoptosis. A possible mechanism for Alzheimer's disease. *Journal of Clinical Investigation* 100, 2333–2340.
- Young, K.M., Merson, T.D., Sotthibundhu, A., *et al.*, 2007. p75 neurotrophin receptor expression defines a population of BDNF-responsive neurogenic precursor cells. *Journal of Neuroscience* 27, 5146–5155.
- Zeng, L., Li, T., Xu, D.C., *et al.*, 2012. Death receptor 6 induces apoptosis not through type I or type II pathways, but via a unique mitochondria-dependent pathway by interacting with Bax protein. *Journal of Biological Chemistry* 287, 29125–29133.
- Zhang, D., Lin, J., Han, J., 2010. Receptor-interacting protein (RIP) kinase family. *Cellular & Molecular Immunology* 7, 243–249.

Further Reading

- Boatright, K.M., Salvesen, G.S., 2003. Mechanisms of caspase activation. *Current Opinion in Cell Biology* 15, 725–731.
- Bodmer, J.-L., Schneider, P., Tschoop, J., 2002. The molecular architecture of the TNF superfamily. *Trends in Biochemical Sciences* 27, 19–26.
- Cremer, A.E., Goni, F.M., Kolesnick, R., 2002. Role of sphingomyelinase and ceramide in modulating rafts: Do biophysical properties determine biologic outcome? *FEBS Letters* 531, 47–53.
- Cremer, A., Paris, F., Grassme, H., *et al.*, 2001. Ceramide enables fas to cap and kill. *Journal of Biological Chemistry* 276, 23954–23961.
- Daniel, N.N., Korsmeyer, S.J., 2004. Cell death: Critical control points. *Cell* 116, 205–219.
- Dechant, G., Barde, Y.A., 2002. The neurotrophin receptor p75^{NTR}: Novel functions and implications for diseases of the nervous system. *Nature Neuroscience* 5, 1131–1136.
- Elmore, S., 2007. Apoptosis: A review of programmed cell death. *Toxicologic Pathology* 35, 495–516.
- Karin, M., Lin, A., 2002. NF-kappaB at the crossroads of life and death. *Nature Immunology* 3, 221–227.
- Kaufmann, S.H., Hengartner, M.O., 2001. Programmed cell death: Alive and well in the new millennium. *Trends in Cell Biology* 11, 526–534.
- Li, Q., Verma, I.M., 2002. NF-kappaB regulation in the immune system. *Nature Reviews Immunology* 2, 725–734.
- Merry, D.E., Korsmeyer, S.J., 1997. Bcl-2 gene family in the nervous system. *Annual Review of Neuroscience* 20, 245–267.
- Skaug, B., Jiang, X., Chen, Z.J., 2009. The role of ubiquitin in NF-kappaB regulatory pathways. *Annual Review of Biochemistry* 78, 769–796.
- Youle, R.J., Strasser, A., 2008. The BCL-2 protein family: Opposing activities that mediate cell death. *Nature Reviews Molecular Cell Biology* 9, 47–59.

Relevant Website

www.deathdomain.org
DeathDomain.

Hedgehog Signaling in Development and Disease

D Jenkins, UCL Institute of Child Health, London, UK

© 2016 Elsevier Inc. All rights reserved.

Introduction

Key components of the hedgehog (HH) signaling pathway were first isolated as mutants affecting larval cuticle patterning in the fruit fly, *Drosophila melanogaster* (Nusslein-Volhard and Wieschaus, 1980). The HH pathway has served as a paradigm of signal transduction for over three decades, leading in our understanding of many of the fundamental mechanisms involved in intercellular communication. The generally accepted canonical model of HH signaling is largely conserved from fruit flies to humans, although differences in some of the biochemical details are apparent between species. One recent development is the finding that primary cilia are particularly important for regulation of HH signaling in higher vertebrates, and this has spawned a whole new field of investigation. The HH pathway also provides text-book examples of: (1) how a long-range morphogen is produced and secreted; (2) how transmembrane receptors are regulated by a morphogen gradient; and (3) how this can feed into transcriptional networks. Other lessons that have emerged from investigation of the HH pathway include: (4) how long-range regulatory elements determine the expression of a morphogen; and (5) the way in which a whole spectrum of anticancer and teratogenic compounds can influence various aspects of signaling mechanisms. It should also be noted that the HH pathway is a paradigm for human genetic disease. It is the only signaling pathway for which most of its constituent components are mutated in human birth defects, while somatic mutations within the pathway also give rise to specific pediatric and adult cancers. Importantly, a battery of therapeutic compounds targeting HH signaling components has been identified. Finally, very recent evidence also suggests added complexity in the HH pathway, in the form of so-called noncanonical pathways.

The Basic Mechanisms of HH Signaling

Production and Secretion of a Morphogen – The HH Proteins

By definition, a morphogen is a secreted molecule that acts at long-range. While a single HH ligand exists in fruit flies, in mammals, three homologs of *Drosophila* HH exist, which are known as Sonic HH (SHH), Indian HH (IHH), and Desert HH (DHH). Each HH morphogen is expressed as a 52 kDa precursor protein that undergoes autoproteolytic cleavage to generate an N-terminal 19 kDa protein that fulfils all patterning activities both at short- and long-range, and a C-terminal 25 kDa species (Lee *et al.*, 1994; Porter *et al.*, 1995). The C-terminal region of the unprocessed precursor serves as a catalytic domain, which cleaves the precursor protein at a highly conserved cysteine residue by an intramolecular mechanism, and this reaction exhibits concentration-independent kinetics (Porter *et al.*, 1995). Processing of HH proteins to generate an active N-terminal morphogen requires

normal cholesterol metabolism, and chemical inhibition of sterol metabolism following treatment of cells with inhibitors of cholesterol biosynthesis impedes HH precursor processing into the two smaller fragments (Guy, 2000). It is now known that HH morphogens are modified by the addition of cholesterol following autoproteolysis of the precursor protein; the newly made N-terminus of the 19 kDa morphogen becomes covalently modified by the addition of cholesterol while its C-terminus becomes palmitoylated at the same time (Lee *et al.*, 1994; Porter *et al.*, 1995).

SHH and IHH have been elegantly shown to travel at a distance of many cell diameters from their source in mice using immunohistochemistry on tissues fixed in such a way that preserves proteoglycans/glycosaminoglycans and proteins in an insoluble form (Gritli-Linde *et al.*, 2001). It is thought that addition of cholesterol is important in determining the range over which HH proteins signal. By harboring a cholesterol moiety, HH morphogens have the tendency to remain at the surface of cells that have synthesized them, owing to hydrophobic interaction between the cholesterol group and cell surface lipids. In some tissues, however, the HH-producing cell also expresses Dispatched (DISP), a sterol-sensing domain protein dedicated to the release of cholesterol modified HH ligand (Burke *et al.*, 1999). Thus, HH signaling is long-range when expressed with DISP but juxtacrine/autocrine when DISP is absent. It is thought that DISP serves to mask the cholesterol motif of HH, reducing its affinity for the cell membrane and allowing its diffusion over many cell diameters, perhaps also promoting HH oligomerization.

Given that three mammalian HH morphogens exist, each of which function in an overlapping yet distinct range of tissues, the choice of morphogen therefore represents another mechanism by which HH signaling is modulated. While SHH, IHH, and DHH bind to their receptor with equal affinity, in certain assays the concentration of each morphogen required to elicit a response can be very different (up to 60-fold); SHH has the greatest potency in most scenarios, with IHH and then DHH inducing progressively reduced effects. These relationships may also be tissue-dependent, however, because in some contexts, such as the inhibition of chondrocyte differentiation, DHH can be more potent than IHH (Pathi *et al.*, 2001). Taken together, this combined evidence shows how processing to generate an active morphogen is also involved in the tight regulation of HH signaling.

Intracellular Effectors of Signal Transduction – Transmembrane Receptors

Cells that are potentially responsive to HH morphogens express the 12-transmembrane domain receptor, Patched (PTCH), which normally serves to repress a second transmembrane protein, Smoothened (SMO) (Figure 1). Therefore, in the absence of HH morphogen, HH signaling is silenced. In contrast, binding of secreted HH to PTCH relieves its normal

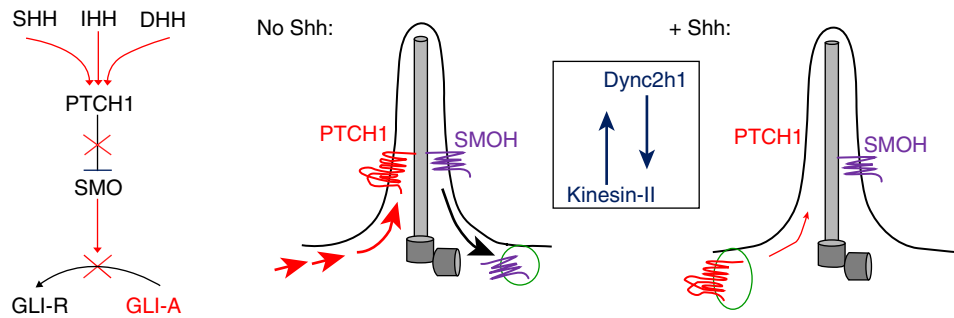


Figure 1 HH signal transduction. Left – Scheme showing the logic of HH signaling. Black lines indicate activity in the absence of SHH, IHH, or DHH ligand. Red lines indicate actions following binding of ligand to the PTCH1 receptor. Middle – Scheme showing a primary cilium with central microtubular doublets anchored to the basal body and encased by membrane. In the absence of HH ligand, the PTCH1 receptor is transported into the cilium, leading to the export of SMO and transport away from the cilium in endocytic vesicles. Right – In the presence of ligand, PTCH1 no longer enters the cilium, allowing SMO to accumulate in this organelle. The box indicates two key molecules, Kinesin-II and Dync2h1 (a key component of dynein-2), which regulate anterograde and retrograde intraflagellar transport, respectively.

repression of SMO by inducing a conformational change (Ingham *et al.*, 2000; Zhao *et al.*, 2007). Extensive genetic and biochemical analyses have demonstrated the importance of protein kinases in driving changes in the subcellular localization of SMO (Jia *et al.*, 2004; Zhang *et al.*, 2005; Liu *et al.*, 2007; Zhu *et al.*, 2003; Denef *et al.*, 2000). Inhibition of PTCH repressor function by HH leads to the cell surface accumulation of SMO, a process which requires phosphorylation at a number of protein kinase A (PKA) and casein kinase I (CKI) sites in the C-terminal tail of SMO (Jia *et al.*, 2004). This intracellular domain of SMO interacts with a protein complex that includes PKA, CKI, fused (FU; another kinase), suppressor of fused (SUFI), and the kinesin-like molecule Costal-2 (COS2), which facilitates interaction of the complex with microtubules (Zhang *et al.*, 2005).

Cholesterol biosynthesis may also be important for regulation of signaling downstream of SMO. Whereas cells over-expressing SMO alone displayed low but significant levels of pathway activity, as assessed by a responsive reporter assay, treatment of these cells with cyclodextrin (an oligosaccharide that complexes with sterols) inhibited this response in a dose-dependent manner. In contrast, the HH response was unaffected by cyclodextrin treatment in cells transfected with a constitutively active oncogenic form of SMO (SMOA1; see Section Mutations in HH Pathway Components in Congenital Disease and Cancer). These results suggest that proper cholesterol biosynthesis is required for normal SMO activity as well as morphogen autoprocessing (Cooper *et al.*, 2003). These results are also relevant for human disease, because mutations in 7-dehydrocholesterol reductase (*DHCR7*), which encodes the ultimate enzyme in the production of cholesterol, cause Smith–Lemli–Opitz syndrome (SLOS). SLOS was the first inborn error of metabolism found to manifest itself as a malformation syndrome, with the range of malformations resembling those of patients with mutations in other HH pathway components (Kelley and Henekam, 2000; see Section Mutations in HH Pathway Components in Congenital Disease and Cancer).

Several novel receptors for HH morphogens, HIP, BOC, CDO, and GAS1, have also been identified more recently. These molecules were identified through combined efforts that

screened cDNA libraries for genes whose protein products facilitate binding of SHH morphogen to the cell surface (Chuang and McMahon, 1999), and that used microarrays to identify genes that are up- or down-regulated in mice with mutations in HH pathway components, consistent with their being targets of HH signaling (Tenzen *et al.*, 2006) (see also the following section). These receptors bind HH morphogens, inhibiting their diffusion and possibly internalizing and degrading them, thus regulating the steepness of the morphogen gradient. In the case of BOC, CDO, and GAS1, they also promote the response of target cells to HH morphogen (Tenzen *et al.*, 2006; Martinelli and Fan, 2007; Allen *et al.*, 2007; Zhang *et al.*, 2006).

Gene Regulatory Networks Controlled by the HH Pathway – The Glioma-Associated Transcription Factors

The ultimate output of HH signaling is the regulation of gene expression, thereby translating concentration-dependent signals into geometric patterns of transcription within a tissue. In the fruit fly, the main transcription factor that mediates the HH response is known as cubitus interruptus (CI), whereas there are three orthologous proteins in vertebrates, known as glioma-associated oncogenes, GLI1-3. In the absence of a morphogen, GLI2 and GLI3 undergo proteolytic cleavage to remove the C-terminus that contains a transcriptional activator domain. The resultant truncated 89 kDa protein still contains its zinc-finger domain, and thus retains its ability to bind DNA. In this context GLI2/GLI3 act as transcriptional repressors owing to N-terminal transcriptional repressor domains. Binding of HH to its receptor and derepression of SMO leads to activation of a series of cellular processes that prevent the cleavage of GLI2/GLI3. In their full-length 190 kDa forms, GLI2/GLI3 possess C-terminal transcriptional activator domains in combination with their zinc-finger domains and thereby actively transcribe target genes. Both the mouse and human GLI2/GLI3 proteins possess an N-terminal repressor and C-terminal activator domain (Roessler *et al.*, 2005) and so are responsible for mediating the HH response. In contrast, the GLI1 protein does not possess an N-terminal transcriptional

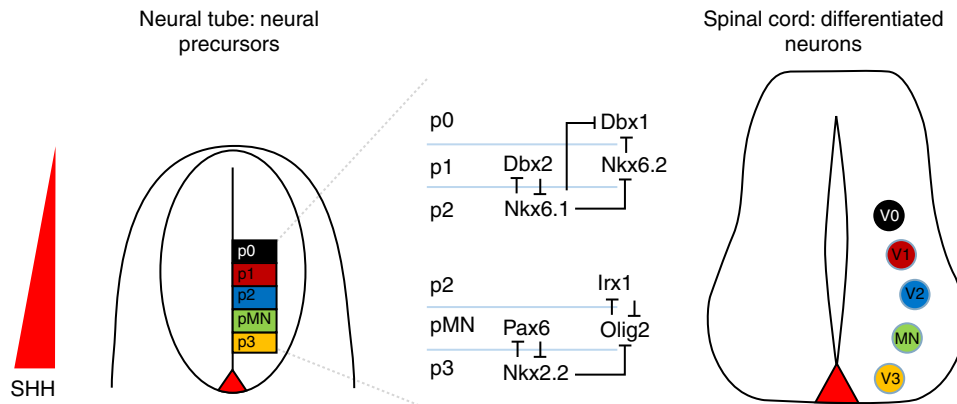


Figure 2 Neural tube patterning by long-range SHH morphogen. Left – Scheme of the neural tube. SHH is expressed by the floor plate (red) which is secreted, generating a ventral-to-dorsal gradient. This leads to the patterning of the neural tube into distinct neural precursor domains, p3–p0. Middle – Scheme showing that this SHH gradient establishes different domains expressing specific combinations of transcription factors. Some examples are given. These transcription factors have many mutually repressive activities, thereby converting the continuous SHH gradient into distinct boundaries of gene expression and neural identity. Right – Scheme of the spinal cord showing how these precursors give rise to specific neurons following this early patterning by SHH.

repressor domain and thus is not responsive to HH signaling. While GLI2/GLI3 mediate the response to HH in most developmental contexts, the GLI1 gene is dispensable for development although it does compensate for a loss of transcriptional activation in mice lacking GLI2/GLI3 (Bai and Joyner, 2001; Bai *et al.*, 2002; Ding *et al.*, 1998; Motoyama *et al.*, 1998; Park *et al.*, 2000). As such, GLI1-mediated transcriptional activation may simply serve to amplify the response to positive HH signaling mediated by GLI2/GLI3.

The cellular localization of CI/GLI proteins is also regulated according to the concentration of morphogen through modification of nuclear localization and nuclear export signals within these proteins. At low concentrations of morphogen, a scaffold of proteins will tether any uncleaved, potentially transcriptionally active GLI protein to the cytoskeleton, preventing its entry into the nucleus. One component of this scaffold is Costal-2, and two vertebrate homologs, Kif7 and Kif27, have been identified as negative regulators of HH signaling in vertebrates. Regulation of HH signaling by these proteins in vertebrates also relates to cilia, which involves tethering of cargo to, and its transport along, the ciliary axoneme (see Section Mutations in HH Pathway Components in Congenital Disease and Cancer). Taken together, the data suggest a ‘three-response’ model for HH signaling (Hooper and Scott, 2005). In the absence of HH morphogen, GLI-repressor is formed through phosphorylation and proteolytic cleavage, thereby inhibiting the expression of HH target genes. In the presence of low concentrations of morphogen, the processing of GLI-repressor is blocked, and some target genes are derepressed, however, a fully functional GLI-activator is either not synthesized, or it is sequestered in the cell cytoplasm. Finally, if the levels of morphogen are sufficiently high, GLI-activator protein isoforms will additionally be released from the protein complex and will enter the nucleus to activate the expression of target genes. Thus a cell can initiate graded responses depending on the concentration of morphogen sensed by the cell.

Work on patterning of the embryonic neural tube has provided an exquisite example of the way in which a gradient

of GLI-transcription factor activity can feed into a gene regulatory network that determines cellular identity. The neural tube is the anlagen of the vertebrate central nervous system. Differentiation of cell types along the dorsoventral (D–V) axis of the neural tube becomes evident as distinct classes of neurons appear at different D–V positions (Figure 2). Initial work implicating HH proteins as morphogens showed that the notochord and floor plate specifically expressed SHH and served as heterologous inducers of D–V neuronal identity. The finding that SHH could substitute for these structures to induce D–V cell fates, showed that SHH acts as a long-range morphogen within the developing neural tube (Yamada *et al.*, 1993; Roelink *et al.*, 1994).

Subsequent work showed that the GLI-transcription factors elicit the response to SHH in this tissue. *Gli2*^{−/−} mice do not correctly specify the ventral-most cells in the neural tube (Ding *et al.*, 1998; Park *et al.*, 2000), while this structure is dorsalized in *Gli3*^{−/−} mutants (Persson *et al.*, 2002; Bai *et al.*, 2002; Mo *et al.*, 1997; Park *et al.*, 2000). These findings are consistent with the notion that GLI2 acts primarily as a transcriptional-activator, while GLI3 acts primarily as a repressor. In the neural tube, this is likely to be the result of differential spatial expression of *Gli2* and *Gli3* in the dorsal–ventral axis rather than intrinsically different biochemical properties. Different neuronal subtypes within the neural tube are derived from different precursor populations that express a number of homeodomain and basic helix–loop–helix transcription factors in response to SHH (Figure 2). Many of these factors act as transcriptional repressors, and a number of mutually cross-repressive interactions have been identified (Dessaud, 2008, for review). As such, following the initial expression of these factors at defined locations along the neural tube, a gene regulatory network is established. This network serves to reinforce the boundaries between different domains of transcription factor expression. As such, a continuous gradient of SHH morphogen is translated into a segmented pattern of neuronal fates via the GLI-transcription factors.

Newer Aspects of HH Signal Transduction

The Role of Primary Cilia in Regulating Mammalian HH Signaling

Analysis of mouse mutants affecting neural tube patterning has shown that primary cilia play roles in regulating HH signal transduction. Cilia are membrane-encased, microtubular extensions present on the apical surface of most cells. They consist of a ciliary axoneme, composed of nine outer microtubule doublets (Jenkins and Beales, 2013). Some cilia also have a central pair of microtubules, and are therefore known as '9 + 2' cilia, whereas others do not have a central pair and are referred to as '9 + 0' cilia. The ciliary axoneme connects at its base to a cytoplasmically located structure known as the basal body that consists of a mother and a daughter centriole. Recent evidence has now shown that there is regulated movement of both proteins, such as transmembrane receptors, and lipids into the cilium which means that the ciliary membrane composition is separated from the cell surface membrane (Rohatgi and Snell, 2010; Hu *et al.*, 2010; Kim *et al.*, 2010). Recent evidence has also identified vesicle trafficking proteins with highly specific roles in trafficking various components to cilia.

Several HH pathway components also localize to cilia in a manner that correlates with HH pathway activity. Ciliary motor proteins are separated into complexes B or A, involved in anterograde and retrograde transport, respectively. GLI3 processing is abnormal in mice with mutations disrupting these motor proteins, with an imbalance in the levels of activator and repressor forms. Furthermore, disruption of different ciliary components can affect GLI3 processing differently. It is therefore clear that different classes of ciliary trafficking proteins regulate distinct aspects of GLI processing. The mechanisms that control this are only just beginning to be understood.

Although HH signaling is highly conserved, both the mechanisms by which Ptch1 represses Smo, and those leading to the activation of the Gli-transcription factors have remained unclear, and recent evidence suggests that this may involve primary cilia. Rohatgi *et al.* (2007) have shown that Ptch1 and Smo exhibit mutually exclusive patterns of ciliary localization – in unstimulated fibroblasts Ptch1 is present within cilia but Smo is not, whereas Smo becomes localized to cilia following application of Shh protein to cells, and Ptch1 concomitantly becomes excluded from cilia. Shh was shown to bind directly to Ptch1 at the ciliary membrane, and Smo was constitutively localized to cilia in *Ptch1*^{-/-} cells. Further work has suggested that Smo is trafficked directly from the plasma membrane to the ciliary membrane via a process of lateral transport, independent of dynamin-mediated endocytosis (Milenkovic *et al.*, 2009).

β -Arrestins are also essential for trafficking of Smo into cilia following Shh treatment (Kovacs *et al.*, 2008). β -Arrestins are key regulators of receptor internalization in clathrin-coated vesicles, and this work has suggested a model whereby Smo is internalized from the plasma membrane and recycled into the cilium through interaction with the type-II kinesin motor, Kif3a. Smo is constitutively localized to cilia in cells deficient for the dynein-heavy chain protein, *Dyn2ch1* (orthologous to

the gene mutated in asphyxiating thoracic dystrophy and short rib polydactyly) (Dagoneau *et al.*, 2009; Merrill *et al.*, 2009), although these cells fail to elicit a HH response (Ocbina and Anderson, 2008; Ocbina *et al.*, 2011). Therefore, Smo translocation into cilia is necessary, but not sufficient, for HH signaling. Instead it seems that the transport of Smo, and probably also GLI proteins, through the cilium is essential to regulate HH signaling appropriately.

Noncanonical HH Signaling

Early work into HH signaling was based largely on the analysis of gene expression in response to HH signaling in body segments and wing imaginal discs in transgenic flies. It is perhaps not surprising, therefore, that the large body of work that has followed has been strongly biased toward an analysis of the transcriptional outputs of signal transduction. In recent years, biochemical analyses of the PTCH1 protein and studies of cell migration/axon guidance are collectively revealing new, transcription-independent modes of response to HH morphogen.

Based on this evidence, Jenkins (2009) defined three modes of 'noncanonical' HH signaling:

Type 1 – Signaling that involves HH pathway components but which is independent of GLI-mediated transcription. An example where transcription-independent signaling has been formally demonstrated in a fibroblast cell line was described by Bijlsma *et al.* (2007). Soluble SHH protein was shown to act as a chemoattractant when applied to cultured fibroblasts, in a manner that was dependent on SMO. However, both chemoattraction and cytoskeletal changes were observed after SHH treatment for only 10 min, a matter of hours before GLI-mediated transcriptional responses to morphogen could first be observed in these cells. More recently, similar experiments in fibroblasts and endothelial cells have indicated that this transcription-independent regulation of cell migration is mediated by direct activation of Rac1 by SMO, in turn leading to modulation of the cytoskeleton (Polizio *et al.*, 2011a,b).

Further support for this mode of signaling initially came from studies of chick neural progenitor cells cultured on Fibronectin or Laminin-1 coated plates. These cells are initially adherent, and subsequently dissociate to become migratory. When cultured on plates onto which SHH had been adsorbed, these cells largely failed to adhere. In contrast, treatment of these explants with soluble SHH protein had no effect on adherence or migration (Jarov *et al.*, 2003; Testaz *et al.*, 2001). This suggests that SHH directly interacts with the extracellular matrix of neuroepithelial cells to stabilize them within an epithelium.

Type 2 – Direct interaction of HH signaling components with other molecular pathways, as opposed to the usual indirect regulation of cellular processes by GLI-mediated transcription. One example of this mode of signaling comes from work showing that SHH and PTCH1 participate in a novel G₂/M phase checkpoint independent of other downstream pathway components. Cyclin B1 is a critical regulator of mitotic cell division, entering the cell nucleus during late G₂ phase. Through a direct protein–protein interaction involving its large intracellular loop, PTCH1 has been shown to sequester the active, phosphorylated form of Cyclin B1 at the cell

surface, thereby inhibiting cell cycle progression (Barnes *et al.*, 2001). Binding of SHH to PTCH1 leads to release of Cyclin B1, thereby allowing it to enter the cell nucleus. As such, these HH pathway components can regulate cell cycle progression in certain situations, independently of SMO or downstream pathway components.

PTCH1 has also been shown to act as a so-called 'dependence-receptor,' inducing caspase-dependent cell death when transfected into cell lines in which it is not normally expressed (Thibert *et al.*, 2003). This effect could be inhibited by treatment with SHH morphogen in a dose-dependent manner, thereby generating a state of cellular dependency on SHH ligand for survival. PTCH1 has now been shown to function as a dependence receptor through its intracellular C-terminal tail, by forming a multi-protein complex consisting of the adaptor protein, DRAL, and Caspase-9 (Mille *et al.*, 2009). Formation of this complex in the absence of SHH morphogen leads to the ubiquitination and activation of Caspase-9, in turn triggering apoptosis. Once again, this pathway is thought to be independent of SMO or the transcriptional response to morphogen.

Since the initial definition of noncanonical HH signaling, many more examples of noncanonical HH signaling have been reported and their roles *in vivo* have been studied further. However, an in-depth evaluation of each of these pathways is beyond the scope of this article.

Long-Range Transcriptional Regulation of the Mammalian *SHH* Gene – The 'Zone of Polarizing Activity Regulatory Sequence'

As shown in Figure 3, the limb bud is patterned by a gradient of SHH morphogen which is normally expressed in a very specific location posteriorly, known as the zone of polarizing activity (ZPA). This gradient controls the number and identity of the digits. The mechanisms that regulate this expression

were illuminated by the identification of noncoding mutations that cause triphalangeal thumb in humans and polydactyly in mice and cats (Hill and Lettice, 2013, for review). These mutations lie within a highly conserved 780 bp *cis*-regulator, known as the ZPA regulatory sequence (ZRS), which is located within an intron of the ubiquitously expressed *Lmbr1* gene. This element lies approximately 1 Mb away from the *SHH* promoter, making it one of the longest-range gene regulatory elements reported to date. Using mouse transgenic reporter assays, Lettice *et al.* (2014) showed that the ZRS was able to drive the expression of a *LacZ* reporter gene within the ZPA. Interestingly, by testing the ability of truncated forms of this element to drive this pattern of expression, it was found that the 5' half was essential for driving expression, while the 3' half was dispensable in these transgenic assays. However, by creating mutations within the endogenous locus, it was found that the 3' element was required for expression of *Shh* within the ZPA. Therefore, the entire ZRS is required for proper regulation of *Shh* expression. Using 3D FISH, these authors and others have shown that the 3' sequence of the ZRS is required for the ZRS to localize with the *SHH* locus, which loops out of its chromosomal territory specifically within ZPA cells (Lettice *et al.*, 2014; Amano *et al.*, 2009).

Over 20 point mutations have been identified in the ZRS in patients with triphalangeal thumb. When these mutations have been tested in the same transgenic reporter assays, several have been shown to cause ectopic expression of *LacZ* within the anterior limb bud in addition to the ZPA, which would explain why the thumb takes on the identity of a finger. This result suggests that there are repressor elements within the ZRS that somehow inhibit expression in the anterior limb bud, perhaps through binding of transcriptional repressors. Indeed, one family has been described with a duplication of the ZRS. This has been interpreted as leading to the 'dilution' of bound repressive factors, i.e., these repressors bind to both the wild-type and duplicated ZRS, thereby halving binding at the ZRS

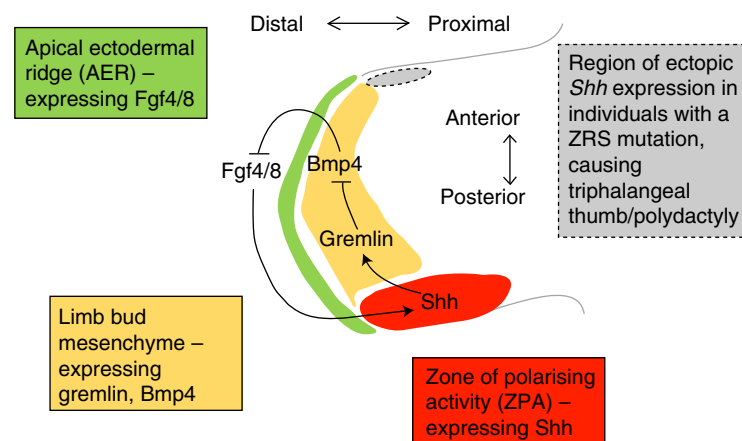


Figure 3 HH signaling in limb developing, and the effect of mutations in a long-range gene regulatory element. Scheme showing the early limb bud (~4 weeks in human fetuses), with proximo-distal (what will become shoulder-to-finger tip) and antero-posterior (little finger-to-thumb) axis indicated. SHH is secreted from the zone of polarizing activity (ZPA) where it signals to the limb bud mesenchyme and sets-up a posterior-to-anterior gradient of positional information. Antagonism between SHH and BMP signaling (by regulating expression of the BMP antagonist Gremlin) modulates expression of Fgf ligand in the apical ectodermal ridge (AER) which in turn signals back to the ZPA and mesenchyme. Mutations in the long-range ZRS leads to ectopic expression in the anterior domain of the limb bud in addition to the ZPA, which disrupts these signaling events and altered digit number and identity.

that interacts with the *Shh* promoter and allowing some ectopic expression. Initial support for the transcription factor binding hypothesis comes from analysis of a point mutation in a polydactylous mouse strain. This point mutant form of the ZRS is bound by HnRNP U, whereas the wild-type version is not. HnRNP U also binds the 5' UTR of the *Shh* gene, and this dual interaction leads to ectopic association between the ZRS and the *Shh* promoter (Zhao *et al.*, 2009). Taken together, the current evidence suggests that the ZRS acts as a likely platform for the binding of many transcriptional repressors and activators that interact with the *Shh* promoter through long-range chromatin looping.

HH Signaling in Human Disease

Mutations in HH Pathway Components in Congenital Disease and Cancer

Most of the molecular components that regulate the HH pathway, which have been discussed in previous sections, are constitutionally mutated in human birth defects, and a number of these components are also somatically mutated in human cancers. The HH pathway is unique, being the only signal transduction pathway for which all of its core components are mutated in human disease, as summarized in Tables 1 and 2. These tables list the phenotypes associated with each of these mutations and a summary of the clinical features of each disease. As discussed above, core HH signaling components have been shown to localize to primary cilia, and their transport through cilia is essential for the proper regulation of HH signaling. Diseases associated with mutations affecting cilia are known as 'ciliopathies.' More than 100 genes have currently been found to be mutated in ciliopathies, making this perhaps the largest category of disease in terms of genetic etiology, and this number is increasing rapidly with the advent of next-generation sequencing technologies. Consistent with the central role of cilia in HH signaling, patients with ciliopathies also exhibit defects that are characteristic of abnormal HH signaling, as discussed below, in particular limb (polydactyly and brachydactyly type-A1) and heart malformations. This category of disease will not be discussed

further as it has recently been reviewed in detail (Jenkins and Beales, 2013).

As can be seen in Tables 1 and 2, different birth defects are readily attributable to the distinct roles of SHH, IHH, or DHH as the active morphogen in different tissues. For example, patients with heterozygous mutations in the *SHH* gene exhibit craniofacial defects including cleft lip/palate and a spectrum of midline defects ranging from solitary median central incisor to cyclopia, which reflects a failure of signaling from the notochord (as discussed in Section Gene Regulatory Networks Controlled by the HH Pathway – The Glioma-Associated Transcription Factors) to instruct the midline structures (including the eye field) to expand during fetal development. Patients also exhibit neurological defects, such as agenesis of the corpus callosum, which reflects the important role of SHH during development of the brain and nervous system. IHH is known to play prominent roles during endochondral ossification of the skeleton, as shown by studies in mice, and patients with *IHH* mutations exhibit defects that reflect this function. Whereas patients with heterozygous mutations in *IHH* exhibit a reduction in the middle phalanges of their hands (known as brachydactyly type-A1, BDA1), patients with homozygous mutations have BDA1 together with a specific form of dwarfism, known as acrocapitofemoral dysplasia (ACFD), which stems from abnormal development of the appendicular skeleton. Finally, DHH is known to have much more restricted functions during embryonic development, with a restricted pattern of expression that is limited to the testis. Consistent with this, patients with *DHH* mutations only display gonadal dysgenesis.

Given that SHH, IHH, and DHH ligands all signal through the PTCH1 receptor, it comes as no surprise that patients with mutations in *PTCH1*, who have basal cell nevus syndrome (BCNS), exhibit a combination of defects that are attributable to the roles of both SHH and IHH ligands during development. These patients have skeletal defects, affecting the ribs, vertebrae and phalanges of the hands, which likely stem from defective IHH signaling. They also have polydactyly, which reflects abnormal patterning of the early limb bud by SHH (Figure 3), and abnormalities of the brain. Interestingly, whereas patients with *SHH* mutations, which render HH signaling in craniofacial structures less active, have a reduction in

Table 1 Summary of human mutations within the hedgehog pathway

Syndrome	Mutated gene	Type of mutation	Mode of inheritance	OMIM ^a accession
HPE3/SMMCI	<i>SHH</i>	LOF	AD	#142945
BDA1	<i>IHH</i>	LOF	AD	#112500
ACFD	<i>IHH</i>	LOF	AR	#607778
Gondal dysgenesis	<i>DHH</i>	ND	AR	#607080
BCNS	<i>PTCH1</i>	LOF	AD	#109400
Carpenter	<i>RAB23</i>	LOF	AR	#201000
Pallister–Hall	<i>GLI3</i>	DN	AD	#146510
GCPS/ACS/PAP	<i>GLI3</i>	LOF	AD	#175700, #200910
P/HPE-L	<i>GLI2</i>	DN	AD	#610829
P/HPE-L	<i>GLI2</i>	LOF	AD	#610829

^aOnline Mendelian Inheritance in Man – <http://www.ncbi.nlm.nih.gov/omim>

Abbreviations: ACFD, acrocapitofemoral dysplasia; ACS, acrocallosal syndrome; BDA1, brachydactyly type-A1; GCPS, Greig cephalopolysyndactyly; HPE3, holoprosencephaly-3; PAP, pre-axial polydactyly; P/HPE-L, pituitary involvement/holoprosencephaly-like; SMMCI, solitary median maxillary central incisor.

Table 2 Summary of birth defects commonly associated with hedgehog pathway mutations

Syndrome	Phenotypes					
	Anogenital	Heart	Craniofacial	Limb	Skeletal	Neurologic
HPH3/SMMCI	Hypoplastic genitalia	Patent ductus arteriosus	Hypertelorism, cyclopia, solitary central incisor, and cleft palate	–	Scoliosis	Dysplastic pituitary, corpus callosum and cerebellum; holoprosencephaly
BDA1	–	–	–	Short/absent middle phalanges	Short stature; musculoskeletal defects	–
ACFD	–	–	Macrocephaly	Short/absent middle phalanges, short/long bones, and cone-shaped epiphyses	Short stature; dysplastic pelvis, vertebrae and hips	–
Gonadal dysgenesis BCNS	–	–	–	–	–	–
Carpenter	Cryptorchidism	Complex	Macrocephaly, hypertelorism; Craniosynostosis; abnormal dentition	Polydactyly	Bifid ribs; scoliosis; dysplastic vertebrae	High birthweight
Pallister–Hall	–	–	Abnormal dentition	Polydactyly; cutaneous syndactyly	–	–
GCPS/ACS/PAP	–	–	Macrocephaly, hypertelorism	Polydactyly; cutaneous syndactyly	–	Hypoplastic corpus callosum
P/HPE-L	–	–	–	Polydactyly	–	–

midline structures (e.g., cyclopia), patients with mutations in *PTCH1*, which render HH signaling more active, have expanded midline structures (e.g., hypertelorism). Patients with mutations in *GLI2* or *GLI3* also have a similar combination of defects.

The HH pathway has also been implicated in driving several human cancers. As well as exhibiting a variety of birth defects, as described above, patients with BCNS who are heterozygous for inactivating mutations in *PTCH1* also have an increased risk for basal cell carcinomas (BCCs) and medulloblastomas (Hahn *et al.*, 1996). Given the role of *PTCH1* as a negative-regulator of the HH pathway, this finding links activated HH signaling to formation of these specific tumors which typically affect the adult skin and the brain of children, respectively. Consistent with this function of *PTCH1* as a classical tumor suppressor acting early in tumor formation, loss-of-heterozygosity at the *PTCH1* locus was reported in both familial and sporadic tumors biopsies (Gailani *et al.*, 1992). Interestingly, while germ-line *SMOH* mutations have not been reported in human birth defects (presumably because they would produce nonviable fetuses), specific somatic activating mutations in *SMOH* are found in sporadic BCCs (Lam *et al.*, 1999). Mutations in the negative-regulator *SUFU* have also been identified in these tumors type (Taylor *et al.*, 2002). Interestingly, genetic ablation of cilia in a mouse model of basal carcinoma which carries an activating mutation in *SMOH*, prevents tumors formation, thereby demonstrating the essential role of cilia in the activation of HH signaling (Wong *et al.*, 2009). Several drivers of cell proliferation and G1/S phase progression, including *C-MYC* and *CYCND1*, are positive transcriptional targets of HH signaling. Therefore, it seems logical that mutations that activate HH signaling have been described in cancer, perhaps leading to enhanced cell survival and proliferation.

Targeting the HH Pathway with Drugs

Given the importance of the HH pathway in normal development, and its deregulation in many forms of human disease, it is therefore of great clinical importance that many small molecules targeting different aspects of the pathway are available for potential treatment of disease. This area of study began over 40 years ago with the identification of cyclopamine. Epidemiologists noticed that clusters of sheep in certain mountainous areas of the United States were producing offspring with cyclopia, similar to patients with *SHH* mutations. Further investigation revealed that the source of these teratogenic effects was the Californian corn lily, which pregnant females were feeding on. These plants produce several steroidal alkaloids, including jervine and cyclopamine, which are teratogenic. A body of work subsequently showed that this compound acts as a small molecule inhibitor of SMO and thus a negative regulator of the HH pathway (Chen *et al.*, 2002). Subsequent chemical screening also identified other small molecules that target the HH pathway. Two such compounds, SAG and purmorphamine, activated HH signaling by agonizing SMO (Chen *et al.*, 2002; Sinha and Chen, 2006), suggesting that SMO is a very 'drug-gable' molecule. This is of great importance for the potential treatment of a variety of cancers caused by activation of the HH

pathway. Indeed, a next-generation orally bioavailable SMO inhibitor, GDC-0449, was used to successfully clear extensive metastatic medulloblastoma nodules in a patient with *PTCH1* loss-of-heterozygosity after only 2 months of treatment, although the metastases did relapse after 3 months, presumably owing to the acquisition of drug-resistant mutations (Rudin *et al.*, 2009). Promising results were also found in the treatment of advanced and metastatic BCC (Von Hoff *et al.*, 2009).

Interestingly, different SMO antagonists have different effects on SMO accumulation within primary cilia. For example, cyclopamine inhibits the downstream HH response, but does not prevent SMO from entering the cilium (Rohatgi *et al.*, 2009). This is analogous to the result described in *Dync2h1*^{-/-} mice, whereby SMO is constitutively localized to cilia but no HH response is elicited, showing that transport through the cilium necessary for HH signaling (Ocbina and Anderson, 2008; Ocbina *et al.*, 2011). Furthermore, constitutive localization of SMO to cilia in *Ptch1*^{-/-} cells is reversed by two other SMO antagonists, known as SANT1&2, but not by cyclopamine. These pharmacological experiments have therefore led to a model whereby SMO is regulated by trafficking to different locations within primary cilia, within which it adopts different active or inactive conformations. Clearly, the identification of pharmacological modulators of HH pathway activity will be important both for the treatment of patients with mutations in the pathway, and for dissection of the basic molecular mechanisms that regulate HH signal transduction.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: Tumor Necrosis Factor Receptors: A Brief Digestion

References

- Allen, B.L., Tenzen, T., McMahon, A.P., 2007. The Hedgehog-binding proteins Gas1 and Cdo cooperate to positively regulate Shh signaling during mouse development. *Genes & Development* 21, 1244–1257.
- Amano, T., Sagai, T., Tanabe, H., *et al.*, 2009. Chromosomal dynamics at the Shh locus: Limb bud-specific differential regulation of competence and active transcription. *Developmental Cell* 16, 47–57.
- Bai, C.B., Auerbach, W., Lee, J.S., Stephen, D., Joyner, A.L., 2002. Gli2, but not Gli1, is required for initial Shh signaling and ectopic activation of the Shh pathway. *Development* 129, 4753–4761.
- Bai, C.B., Joyner, A.L., 2001. Gli1 can rescue the *in vivo* function of Gli2. *Development* 128, 5161–5172.
- Barnes, E.A., Kong, M., Ollendorff, V., Donoghue, D.J., 2001. Patched1 interacts with cyclin B1 to regulate cell cycle progression. *EMBO Journal* 20, 2214–2223.
- Bijlsma, M.F., Borensztajn, K.S., Roelink, H., Peppelenbosch, M.P., Spek, C.A., 2007. Sonic hedgehog induces transcription-independent cytoskeletal rearrangement and migration regulated by arachidonate metabolites. *Cellular Signalling* 19, 2596–2604.
- Burke, R., Nellen, D., Bellotto, M., *et al.*, 1999. Dispatched, a novel sterol-sensing domain protein dedicated to the release of cholesterol-modified hedgehog from signaling cells. *Cell* 99, 803–815.
- Chen, J.K., Taipale, J., Young, K.E., Maiti, T., Beachy, P.A., 2002. Small molecule modulation of smoothened activity. *Proceedings of the National Academy of Sciences of the United States of America* 99, 14071–14076.
- Chuang, P.T., McMahon, A.P., 1999. Vertebrate Hedgehog signalling modulated by induction of a Hedgehog-binding protein. *Nature* 397 (6720), 617–621.
- Cooper, M.K., Wassif, C.A., Krakowiak, P.A., *et al.*, 2003. A defective response to Hedgehog signaling in disorders of cholesterol biosynthesis. *Nature Genetics* 33, 508–513.

- Dagoneau, N., Goulet, M., Geneviève, D., *et al.*, 2009. DYNC2H1 mutations cause asphyxiating thoracic dystrophy and short rib-polydactyly syndrome, type III. *American Journal of Human Genetics* 84, 706–711.
- Denef, N., Neubüser, D., Perez, L., Cohen, S.M., 2000. Hedgehog induces opposite changes in turnover and subcellular localization of patched and smoothened. *Cell* 102, 521–531.
- Dessaud, E., 2008. Pattern formation in the vertebrate neural tube: A sonic hedgehog morphogen-regulated transcriptional network. *Development* 135, 2489–2503.
- Ding, Q., Motoyama, J., Gasca, S., *et al.*, 1998. Diminished Sonic hedgehog signaling and lack of floor plate differentiation in Gli2 mutant mice. *Development* 125, 2533–2543.
- Gailani, M.R., Bale, S.J., Leffell, D.J., *et al.*, 1992. Developmental defects in Gorlin syndrome related to a putative tumor suppressor gene on chromosome 9. *Cell* 69, 111–117.
- Gritti-Linde, A., Lewis, P., McMahon, A.P., Linde, A., 2001. The whereabouts of a morphogen: Direct evidence for short- and graded long-range activity of hedgehog signaling peptides. *Developmental Biology* 236, 364–386.
- Guy, R.K., 2000. Inhibition of sonic hedgehog autoprocessing in cultured mammalian cells by sterol deprivation. *Proceedings of the National Academy of Sciences of the United States of America* 97, 7303–7312.
- Hahn, H., Wicking, C., Zaphiropoulos, P.G., *et al.*, 1996. Mutations of the human homolog of *Drosophila* patched in the nevoid basal cell carcinoma syndrome. *Cell* 85, 841–851.
- Hill, R.E., Lettice, L.A., 2013. Alterations to the remote control of Shh gene expression cause congenital abnormalities. *Philosophical Transactions of the Royal Society B: Biological Sciences* 368, 20120357.
- Hooper, J.E., Scott, M.P., 2005. Communicating with Hedgehogs. *Nature Reviews Molecular Cell Biology* 6, 306–317.
- Hu, Q., Milenkovic, L., Jin, H., *et al.*, 2010. A septin diffusion barrier at the base of the primary cilium maintains ciliary membrane protein distribution. *Science* 329, 436–439.
- Ingham, P.W., Nystedt, S., Nakano, Y., *et al.*, 2000. Patched represses the Hedgehog signalling pathway by promoting modification of the Smoothened protein. *Current Biology* 10, 1315–1318.
- Jarov, A., Williams, K.P., Ling, L.E., *et al.*, 2003. A dual role for Sonic hedgehog in regulating adhesion and differentiation of neuroepithelial cells. *Developmental Biology* 261 (2), 520–536.
- Jenkins, D., 2009. Hedgehog signalling: Emerging evidence for non-canonical pathways. *Cell Signal* 21, 1023–1034.
- Jenkins, D., Beales, P.L., 2013. Chapter 165 — Genes and mechanisms in human ciliopathies. In: Rimoin, D., Pyeritz, R., Korf, B. (Eds.), *Emery and Rimoin's Principles and Practice of Medical Genetics*, sixth ed. Oxford: Academic Press, pp. 1–36. ISBN 9780123838346. Available at: <http://dx.doi.org/10.1016/B978-0-12-383834-6.00174-9> (accessed 17.03.15).
- Jia, J., Tong, C., Wang, B., Luo, L., Jiang, J., 2004. Hedgehog signalling activity of Smoothened requires phosphorylation by protein kinase A and casein kinase I. *Nature* 432, 1045–1050.
- Kelley, R.I., Henekam, R.C., 2000. The Smith-Lemli-Opitz syndrome. *Journal of Medical Genetics* 37, 321–335.
- Kim, S.K., Shindo, A., Park, T.J., *et al.*, 2010. Planar cell polarity acts through septins to control collective cell movement and ciliogenesis. *Science* 10, 1337–1340.
- Kovacs, J.J., Whalen, E.J., Liu, R., *et al.*, 2008. Beta-arrestin-mediated localization of smoothened to the primary cilium. *Science* 320, 1777–1781.
- Lam, C.W., Xie, J., To, K.F., *et al.*, 1999. A frequent activated smoothened mutation in sporadic basal cell carcinomas. *Oncogene* 18, 833–836.
- Lee, J.J., Ekker, S.C., von Kessler, D.P., *et al.*, 1994. Autoproteolysis in hedgehog protein biogenesis. *Science* 266, 1528–1537.
- Lettice, L.A., Williamson, I., Devenney, P.S., *et al.*, 2014. Development of five digits is controlled by a bipartite long-range cis-regulator. *Development* 141, 1715–1725.
- Liu, Y., Cao, X., Jiang, J., Jia, J., 2007. Fused-costal2 protein complex regulates Hedgehog-induced Smo phosphorylation and cell-surface accumulation. *Genes & Development* 21, 1949–1963.
- Martinelli, D.C., Fan, C.M., 2007. Gas1 extends the range of Hedgehog action by facilitating its signaling. *Genes & Development* 21, 1231–1243.
- Merrill, A.E., Merriman, B., Farrington-Rock, C., *et al.*, 2009. Ciliary abnormalities due to defects in the retrograde transport protein DYNC2H1 in short-rib polydactyly syndrome. *American Journal of Human Genetics* 84, 542–549.
- Milenkovic, L., Scott, M.P., Rohatgi, R., 2009. Lateral transport of Smoothened from the plasma membrane to the membrane of the cilium. *Journal of Cell Biology* 187, 365–374.
- Mille, F., Thibert, C., Fombonne, J., *et al.*, 2009. The Patched dependence receptor triggers apoptosis through a DRAL-caspase-9 complex. *Nature Cell Biology* 11, 739–746.
- Mo, R., Freer, A.M., Zinyk, D.L., *et al.*, 1997. Specific and redundant functions of Gli2 and Gli3 zinc finger genes in skeletal patterning and development. *Development* 124, 113–123.
- Motoyama, J., Liu, J., Mo, R., *et al.*, 1998. Essential function of Gli2 and Gli3 in the formation of lung, trachea and oesophagus. *Nature Genetics* 20, 54–57.
- Nusslein-Volhard, C., Wieschaus, E., 1980. Mutations affecting segment number and polarity in *Drosophila*. *Nature* 287, 795–801.
- Ocbina, P.J., Anderson, K.V., 2008. Intraflagellar transport, cilia, and mammalian Hedgehog signaling: Analysis in mouse embryonic fibroblasts. *Developmental Dynamics* 237, 2030–2038.
- Ocbina, P.J., Eggenschwiler, J.T., Moskowitz, I., Anderson, K.V., 2011. Complex interactions between genes controlling trafficking in primary cilia. *Nature Genetics* 43, 547–553.
- Park, H.L., Bai, C., Platt, K.A., *et al.*, 2000. Mouse Gli1 mutants are viable but have defects in SHH signaling in combination with a Gli2 mutation. *Development* 127, 1593–1605.
- Pathi, S., Pagan-Westphal, S., Baker, D.P., *et al.*, 2001. Comparative biological responses to human Sonic, Indian and Desert Hedgehog. *Mechanisms of Development* 106, 107–117.
- Persson, M., Stamatakis, D., te Welscher, P., *et al.*, 2002. Dorsal–ventral patterning of the spinal cord requires Gli3 transcriptional repressor activity. *Genes and Development* 16 (22), 2865–2878.
- Polizio, A.H., Chinchilla, P., Chen, X., *et al.*, 2011a. Heterotrimeric Gi proteins link Hedgehog signaling to activation of Rho small GTPases to promote fibroblast migration. *Journal of Biological Chemistry* 286 (22), 19589–19596. doi:10.1074/jbc.M110.197111.
- Polizio, A.H., Chinchilla, P., Chen, X., Manning, D.R., Riobo, N.A., 2011b. Sonic Hedgehog activates the GTPases Rac1 and RhoA in a Gli-independent manner through coupling of smoothened to Gi proteins. *Science Signaling* 4 (200), pt7. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22114142> (accessed 17.03.15).
- Porter, J.A., von Kessler, D.P., Ekker, S.C., *et al.*, 1995. The product of hedgehog autoproteolytic cleavage active in local and long-range signaling. *Nature* 374, 363–366.
- Roelink, H., Augsburger, A., Heemskerk, J., *et al.*, 1994. Floor plate and motor neuron induction by vhh-1, a vertebrate homolog of hedgehog expressed by the notochord. *Cell* 76 (4), 761–775.
- Roessler, E., Ermilov, A.N., Grange, D.K., *et al.*, 2005. A previously unidentified amino-terminal domain regulates transcriptional activity of wild-type and disease-associated human GLI2. *Human Molecular Genetics* 14, 2181–2188.
- Rohatgi, R., Milenkovic, L., Corcoran, R.B., Scott, M.P., 2009. Hedgehog signal transduction by smoothened: Pharmacologic evidence for a 2-step activation process. *Proceedings of the National Academy of Sciences of the United States of America* 106, 3196–3201.
- Rohatgi, R., Milenkovic, L., Scott, M.P., 2007. Patched1 regulates hedgehog signaling at the primary cilium. *Science* 317, 372–376.
- Rohatgi, R., Snell, W.J., 2010. The ciliary membrane. *Current Opinion in Cell Biology* 22, 541–546.
- Rudin, C.M., Hann, C.L., Laterra, J., *et al.*, 2009. Treatment of medulloblastoma with hedgehog pathway inhibitor GDC-0449. *New England Journal of Medicine* 361, 1173–1178.
- Sinha, S., Chen, J.K., 2006. Purmorphamine activates the Hedgehog pathway by targeting Smoothened. *Nature Chemical Biology* 2, 29–30.
- Taylor, M.D., Liu, L., Raffel, C., *et al.*, 2002. Mutations in SUFU predispose to medulloblastoma. *Nature Genetics* 31, 306–310.
- Tenzen, T., Allen, B.L., Cole, F., *et al.*, 2006. The cell-surface membrane proteins Cdo and Boc are components and targets of the Hedgehog signaling pathway and feedback network in mice. *Developmental Cell* 10, 647–656.
- Testaz, S., Jarov, A., Williams, K.P., *et al.*, 2001. Sonic hedgehog restricts adhesion and migration of neural crest cells independently of the Patched-Smoothened-Gli signaling pathway. *Proceedings of the National Academy of Sciences of the United States of America* 98, 12521–12526.
- Thibert, C., Teillet, M.A., Lapointe, F., *et al.*, 2003. Inhibition of neuroepithelial patched-induced apoptosis by sonic hedgehog. *Science* 301, 843–846.
- Von Hoff, D.D., LoRusso, P.M., Rudin, C.M., *et al.*, 2009. Inhibition of the hedgehog pathway in advanced basal-cell carcinoma. *New England Journal of Medicine* 361, 1164–1172.
- Wong, S.Y., Seol, A.D., So, P.L., *et al.*, 2009. Primary cilia can both mediate and suppress Hedgehog pathway-dependent tumorigenesis. *Nature Medicine* 15, 1055–1061.

- Yamada, T., Pfaff, S.L., Edlund, T., Jessell, T.M., 1993. Control of cell pattern in the neural tube: Motor neuron induction by diffusible factors from notochord and floor plate. *Cell* 73 (4), 673–686.
- Zhang, W., Kang, J.-S., Cole, F., Yi, M.-J., Krauss, R.S., 2006. Cdo functions at multiple points in the Sonic Hedgehog pathway, and Cdo-deficient mice accurately model human holoprosencephaly. *Developmental Cell* 10, 657–665.
- Zhang, W., Zhao, Y., Tong, C., *et al.*, 2005. Hedgehog-regulated costal2-kinase complexes control phosphorylation and proteolytic processing of cubitus interruptus. *Developmental Cell* 8, 267–278.
- Zhao, J., Ding, J., Li, Y., *et al.*, 2009. HnRNP U mediates the long-range regulation of Shh expression during limb development. *Human Molecular Genetics* 18 (16), 3090–3097. doi:10.1093/hmg/ddp250.
- Zhao, Y., Tong, C., Jiang, J., 2007. Hedgehog regulates smoothened activity by inducing a conformational switch. *Nature* 450 (7167), 252–258.
- Zhu, A.J., Zheng, L., Suyama, K., Scott, M.P., 2003. Altered localization of *Drosophila* Smoothened protein activates Hedgehog signal transduction. *Genes & Development* 17, 1240–1252.

The Wnt/ β -Catenin Pathway

C Liu, University of Kentucky, Lexington, KY, USA

Y Regimbald-Dumas, X Zhang, and X He, Harvard Medical School, Boston, MA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The Wnt family of secreted lipoglycoproteins constitutes one of the most fundamental signaling systems in animals, from primitive sponges to humans. Wnt proteins guide most, if not all, aspects of embryogenesis, from the early body plan to generations of various tissues and organs. In adult animals, Wnt proteins govern tissue homeostasis, including regulation of stem cell self-renewal and differentiation, and repair and regeneration upon damages and injuries (Clevers and Nusse, 2012). These various vital functions of Wnt proteins in development and homeostasis are reflected by the fact that birth defects, congenital disorders, cancers, and other diseases are frequently associated with mutations or deregulation of Wnt proteins and their downstream signaling components (MacDonald et al., 2009). Wnt proteins exert their functions through many signaling pathways, which, as a result of our incomplete understanding, are roughly classified as canonical (β -catenin-dependent) and noncanonical (β -catenin-independent) pathways. This article will focus on canonical Wnt/ β -catenin signaling, which acts through the transcriptional co-activator β -catenin to govern gene expression, and has been, relatively speaking, better defined. For noncanonical Wnt signaling pathways, which are often involved in regulation of cell polarity, movements, and possibly gene expression through various but less well characterized signaling routes, we refer readers to other excellent reviews (Wallingford, 2012; Singh and Mlodzik, 2012).

Wnt Proteins

The first Wnt gene, *Wnt1*, was initially called *int1* and discovered as an oncogene causing mouse mammary tumorigenesis (Nusse and Varmus, 1982). The famous *Drosophila* *Wingless* locus encodes a protein, Wingless (Wg), that is homologous to the Int1 protein, leading to the name Wnt (Wingless and Int) (Nusse et al., 1991). Wnt genes are found in all sequenced animal genomes. The primitive *Hydra vulgaris* has 13 Wnt genes, whereas the genome of mice or humans has 19 Wnt genes (Willert and Nusse, 2012). Wnt proteins are secreted, with an N-terminal signal peptide followed by about 350–400 amino acid residues including 22 or 24 invariable cysteines across the polypeptide forming 11 or 12 disulfide bonds (Willert and Nusse, 2012). Wnt proteins are glycosylated, but the most unusual and functionally important feature is that Wnt proteins are lipid modified (Takada et al., 2006; Willert et al., 2003). This lipid modification is referred to as O-palmitoleylation, which conjugates a monounsaturated palmitoleic acid (C16:1) onto the hydroxyl group of a conserved serine residue of the Wnt protein (Takada et al., 2006; Willert et al., 2003). Thus, Wnt proteins are hydrophobic and insoluble in water, likely accounting for the notorious

difficulty purifying Wnt proteins. Palmitoleylation likely constrains the ability of Wnt proteins to diffuse in an aqueous extracellular environment, and indeed Wnt proteins act locally in most contexts (Clevers et al., 2014). However, Wnt proteins such as Wg also act over longer distances as morphogens, which regulate cell behavior in a concentration gradient (Hausmann et al., 2007). In these instances, binding of Wnt proteins to carriers such as lipoprotein particles (Neumann et al., 2009; Panakova et al., 2005), specific partners such as Swim (Mulligan et al., 2012), and various cell surface proteoglycans such as glypicans (Filmus et al., 2008; Yan and Lin, 2009) appears to be important.

While Wg has been extensively characterized genetically, the mouse Wnt3a was the first purified, and is the best biochemically characterized, Wnt protein, in particular, in regard to palmitoleylation (Willert and Nusse, 2012; Takada et al., 2006; Willert et al., 2003). On the other hand, *Xenopus* Wnt8 was the first to be crystallized (Janda et al., 2012; Figure 1). Wnt8 structure resembles a hand with a large palm region and two finger-like (the 'thumb' and 'index finger') protrusions, which consists of looping β strands stabilized by disulfide bonds, with the most prominent feature being the palmitoleate adduct at the tip of the Wnt8 thumb pointing outwards (Figure 1). Based on the Wnt structure and bioinformatics, it has been speculated that the primordial Wnt protein is evolved from an ancient fusion of a saposin-like and a cytokine-like domains (Bazan et al., 2012). Thus Wnt lipid

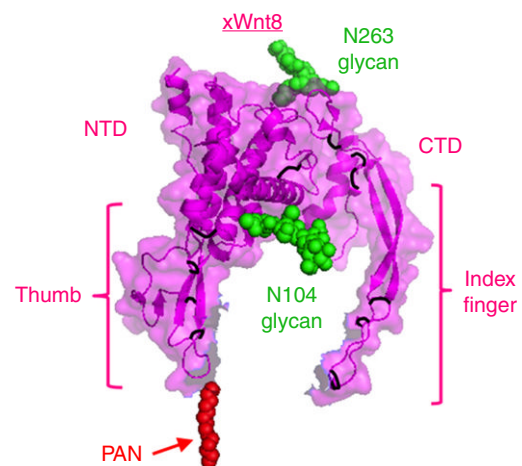


Figure 1 Structure of xWnt8. Surface representation and secondary structure of xWnt8 (magenta) as viewed 'face on.' The two N-Linked glycans of xWnt8 (N104 and N263) are drawn in green shapes. The palmitoleic acid (PAM) group is shown in red extending from the Wnt thumb. Disulfide bonds are shown in black. NTD and CTD, N- and C-terminal domains. This structure was generated using PyMOL software and the crystal structure of xWnt8 published by Janda et al. (2012).

modification is a cardinal feature that has profound effects on Wnt proteins, from secretion to signaling (see below).

Wnt Protein Biogenesis

The unusual properties of Wnt proteins, particularly palmitoleoylation, imply the existence of unique machineries for their biogenesis and secretion. Indeed, genetic screens in *Drosophila* and nematodes have identified a number of genes that have conserved functions in Wnt biogenesis and secretion from invertebrates to mammals (Willert and Nusse, 2012; Hausmann *et al.*, 2007; Figure 2). Porcupine (Porc) is a multi-span transmembrane protein residing in the endoplasmic reticulum (ER) and is likely a Wnt-specific O-acyltransferase responsible for Wnt palmitoleoylation. The p24 family of transmembrane cargo receptors, which are known to be involved in ER exit of glycosylphosphatidylinositol (GPI)-anchored proteins, are required for Wnt exit from the ER to Golgi. Wntless (Wls), also known as Evenness interrupted or Sprinter or Gpr177, is a multi-span transmembrane protein that binds to palmitoleoylated Wnt proteins in the Golgi, chaperoning them through the exocytic pathway to the plasma membrane (PM) for release. The critical importance of Wls in Wnt secretion is reflected in findings that Wls recycling from the PM to Golgi is critical. Thus the retromer complex, sorting nexin-3 (SNX3), and the MTM-6/MTM-9 myotubularin proteins that are lipid phosphatases for phosphatidylinositol-3-phosphate (PI3P), are each required for retrieval of Wls from the PM back to Golgi for Wnt secretion. Studies in mammalian cells have implicated stearoyl desaturase (SCD) and the V-ATPase in Wnt biogenesis: SCD converts abundant palmitic acid into palmitoleic acid for adduction into Wnt by Porc (Rios-Esteves and

Resh, 2013), whereas the V-ATPase, a proton pump for vesicular acidification, appears to be required for release of Wnt from Wls at the PM. In this sophisticated Wnt production cascade (Figure 2), Porc and Wls appear to be dedicated to Wnt biogenesis while the other proteins/complexes likely have additional functions in vesicular trafficking or lipid metabolism, although the loss of function of these genes commonly produces loss of Wnt phenotypes (a consequence of prominent influence of Wnt signaling on embryogenesis).

Wnt Receptors: Frizzled and LRP5/6

Several families of Wnt receptors have been described. For canonical Wnt/ β -catenin signaling, the Frizzled (Fz) family of serpentine receptors and LDL receptor-related proteins 5 and 6 (LRP5/6) are both essential and will be reviewed in detail here. Other Wnt receptors such as Ryk, Ror, and Ptk7, which belong to the atypical receptor tyrosine kinase (RTK) family, have primary roles, either alone or together with Fz, in non-canonical Wnt signaling, on which we refer readers to some excellent reviews (Green *et al.*, 2014; Peradziriyi *et al.*, 2012).

Fz receptors harbor an extracellular cysteine-rich domain (CRD) that binds to Wnt proteins with high affinities (1–10 nM), and participate in perhaps most, if not all, Wnt pathways (Figure 3; MacDonald and He, 2012). How a Wnt ligand engages Fz is illustrated by the xWnt8–Fz8CRD co-crystal structure (Janda *et al.*, 2012; Figure 3). In the complex, the xWnt8 thumb interacts with Fz8CRD through hydrophobic contacts and inserts the palmitoleoylate adduct into a hydrophobic cleft of Fz8CRD in a ‘lipid in groove’ fashion (site 1), while the xWnt8 index finger ends in a vicinal disulfide bond that wraps around the other edge of Fz8CRD through

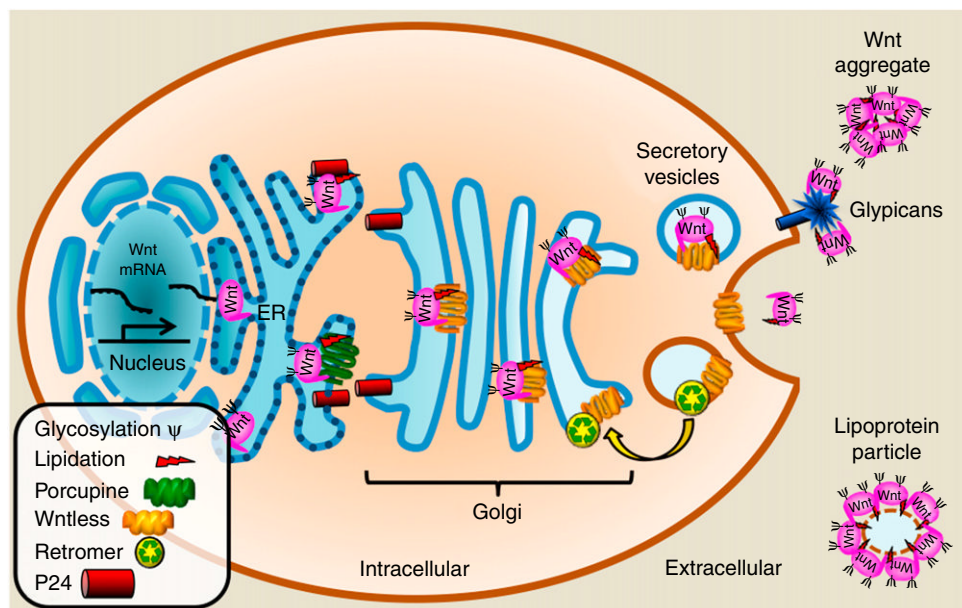


Figure 2 Wnt protein biogenesis. Glycosylation and lipidation of Wnt occur in the endoplasmic reticulum (ER) and require porcupine. P24 is required for Wnt export from the ER. Wnt proteins are transported by Wntless from the Golgi to the PM. Wntless is recycled from endocytic vesicles back to the golgi by the retromer complex formed by SNX3 and the myotubularins proteins. After secretion, mature Wnt can bind to glypicans, travel on lipoprotein particles or form aggregates.

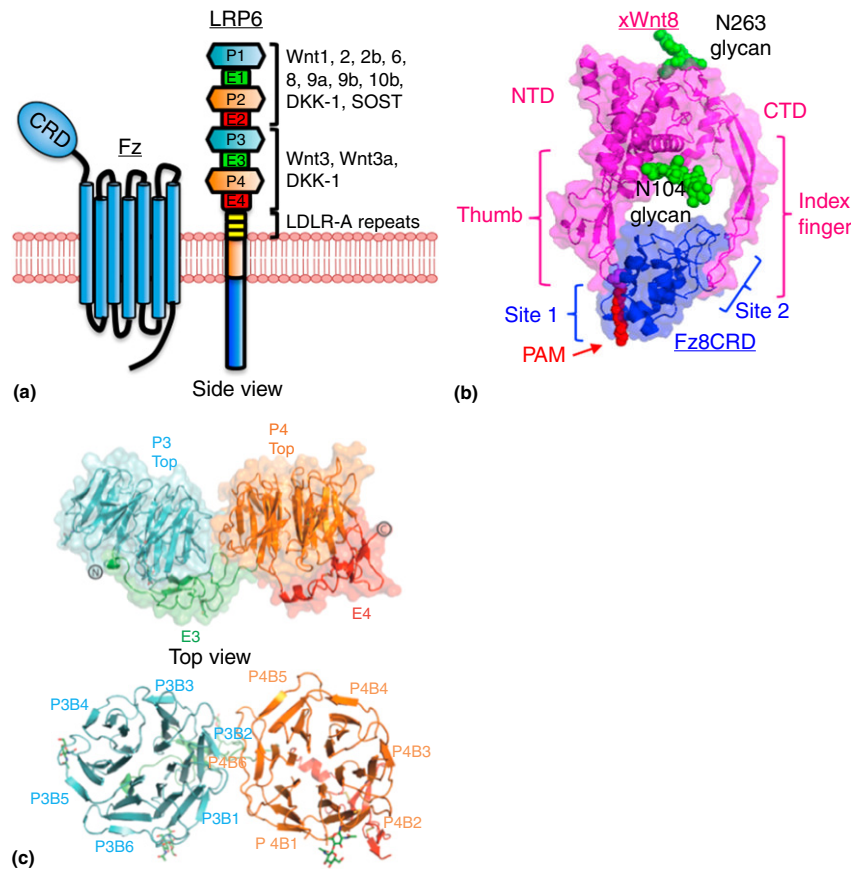


Figure 3 Wnt receptors: Frizzled and LRP6 (a) representation of Frizzled (Fz) and LRP6. LRP6 binding to different Wnt proteins and antagonists SOST and DKK1 are shown. CRD (cysteine-rich domain), E1–E4 (EGF-repeats), P1–P4 (β -propeller domains) (b) xWnt8–Fz8CRD co-crystal structure, shown as a ribbon diagram superimposed on surface representation (Janda *et al.*, 2012). (c) Side and top views of the secondary structure of LRP6 P3E3–P4E4 (Chen *et al.*, 2011).

hydrophobic contacts, ‘knob-in-hole’ fashion (site 2) (Figure 3; Janda *et al.*, 2012). Thus, xWnt8 resembles a human hand with a central palm that extends the thumb and the index finger to grab/pinch Fz8CRD (Figure 3). Because contact residues of Wnt and Fz proteins at site 1 and site 2 are conserved but with variations, the relative broad specificity of Wnt–Fz binding, i.e., one Wnt binding to multiple Fz proteins and vice versa, can be rationalized by this unusual ligand–receptor structure, in which Wnt palmitoleoylation has a critical role.

LRP5 and LRP6 (and their *Drosophila* homolog Arrow) are single-span transmembrane proteins and act as essential coreceptors for Fz in Wnt/ β -catenin signaling, but unlike Fz, they are not required for noncanonical Wnt pathways (MacDonald and He, 2012; He *et al.*, 2004). LRP5 and LRP6 share partially redundant functions, with LRP6 being a more dominant of the two, during mouse embryogenesis (He *et al.*, 2004). However, LRP5 is critical for human bone mass accrual and is thus a key player in osteoporosis pathogenesis (Regard *et al.*, 2012). The ectodomain of LRP5 and LRP6 are composed of four repeating units, P1E1 to P4E4, each of which is a so-called YWTD- β -propeller domain (P) juxtaposed by an EGF-repeat (E), followed by three LDL receptor type A (LDLR-A) repeats before the transmembrane domain (Figure 3; MacDonald and He, 2012; He *et al.*, 2004). Wnt proteins promote Fz–LRP6

complex formation *in vitro*, and therefore bind to Fz and LRP6 simultaneously (Tamai *et al.*, 2000; Bourhis *et al.*, 2010). The PE domains are involved in Wnt-binding, and interestingly, P1E1 and P3E3 appear to bind different sets of Wnt ligands, with Kd values at about 10 nM (Ahn *et al.*, 2011; Chen *et al.*, 2011). Structural studies suggest that the top surface of the YWTD- β -propeller domains, P1 and P3, are likely the binding sites for Wnt proteins (Ahn *et al.*, 2011; Chen *et al.*, 2011), although a lack of a Wnt–LRP5/6 co-crystal structure prohibits a definitive assignment. P2E2 and P4E4 are essential for the structural integrity of the PE units (Bourhis *et al.*, 2010; Liu *et al.*, 2009), but it is unknown whether they bind Wnt or other ligands (see below). The role of LDLR-A repeats is less clear, although their involvement in LRP5/6 dimerization has been suggested (Ahn *et al.*, 2011). The intracellular domain of LRP5/6 (and Arrow) is conserved with multiple phosphorylation motifs that are essential for Wnt/ β -catenin signaling activation (see below).

Wnt Antagonists

Many families of secreted or membrane-bound antagonists of Wnt signaling have been identified, and they expand strategies

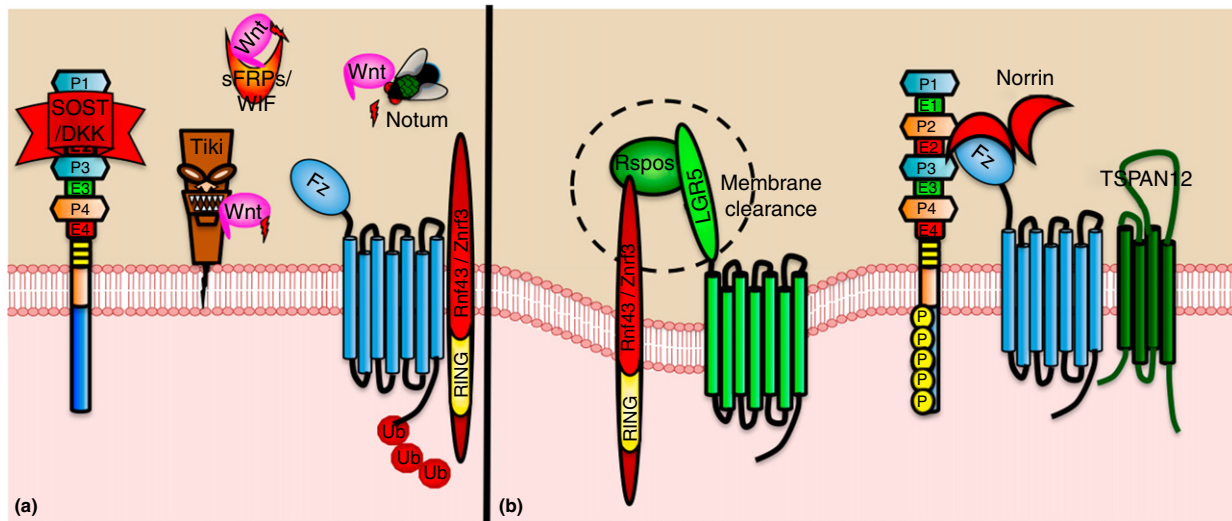


Figure 4 Wnt/ β -catenin signaling modulators. (a) Secreted inhibitors such as DKK1 and SOST prevent canonical signaling through negative actions on LRP5/6. SFRPs and WIF inhibit Wnt signaling by direct binding to Wnt ligands. Membrane tethered Tiki and Notum (depicted as a fly) inactivate Wnt by cleaving the N-terminal or deacylating Wnt respectively. Ubiquitin ligases Rnf43 and Znf3 promote LRP5/6 and Frizzled receptors turnover. (b) R-spondin proteins act as potent enhancers of Wnt signaling by clearing Rnf43 and Znf3 inhibitors from the plasma membrane. Norrin and TSPAN12 enhance Wnt signaling by promoting Fz4 and LRP5/6 clustering.

of Wnt signaling regulation far beyond an already complex Wnt-receptor network (Cruciat and Niehrs, 2013; Malinauskas and Jones 2014). These factors can be classified based on their mechanisms of actions (Figure 4(a)). Secreted Frizzled-related proteins (sFRPs) share homology with the Fz CRD, and bind and titrate Wnt from Fz (and may also heterodimerize with Fz); sFRPs may have biphasic functions and can enhance the range of Wnt proteins presumably by shielding the lipid moiety to facilitate Wnt diffusion (Mii and Taira, 2009). Other Wnt-binding antagonists include Wnt inhibitory factor (WIF) and *Xenopus* Cerberus (Cruciat and Niehrs, 2013). Dickkopf-1 (Dkk1) on the other hand binds to LRP5/6 with high affinity and competes with Wnt for the coreceptors; Wise and SOST, which is a human skeletal disease gene product, act in a similar manner. However, Dkk1 binds to P1E1 and P3E3 of LRP6 whereas SOST binds to P1E1, thereby exhibiting inhibition toward different Wnt proteins (Ahn et al., 2011; Chen et al., 2011; Bourhis et al., 2011). Fz is also a key regulatory target. Shisa is an ER-resident protein that traps Fz (and also the fibroblast growth factor receptor, FGFR) in the ER to prevent Wnt signaling, whereas RNF43 and ZNRF3 are homologous transmembrane proteins with an intracellular E3 ligase domain that ubiquitinate Fz, thereby triggering degradation of Fz and associated coreceptors including LRP6 (Furushima et al., 2007; Yamamoto et al., 2005; Koo et al., 2012; Hao et al., 2012) (also see below). More recently Wnt modifying and inactivating enzymes have been identified. Tiki1 and Tiki2 are membrane-anchored metalloproteases that cleave the N-terminus of some Wnt proteins, whereas Notum is a Wnt deacylase that removes the palmitoleoylate adduct, with both types of enzymes fully inactivating Wnt ligands (Zhang et al., 2012, 2015; Kakugawa et al., 2015). There are several additional Wnt antagonists that share mechanisms similar to those mentioned (Cruciat and Niehrs, 2013; Malinauskas and Jones, 2014).

Wnt Agonists

Norrin (Norrie Disease gene product) and R-spondin (Rspo) proteins are two families of secreted Wnt agonists (Figure 4 (b)). Norrin binds to Fz4 (but not any other Fz) and LRP5/6 (Wang et al., 2012; Ke et al., 2013) and thus acts in a manner analogous to Wnt proteins but with a strict Fz4 choice. Norrin (but not Wnt) shows dependence on a tetraspannin protein, TSPAN12, which appears to promote Fz4 clustering (Junge et al., 2009). On the other hand, Rspo proteins act primarily as enhancers of Wnt signaling in a sophisticated fashion. Rspo binds to two types of cell surface receptors: leucine-rich repeat containing G-protein-coupled receptors 4/5/6, LGR4/5/6, which like Fz are serpentine receptors, and the aforementioned RNF43/ZNRF3 ubiquitin ligases (Malinauskas and Jones, 2014; de Lau et al., 2014). The Rspo-LGR-RNF43/ZNRF3 trimeric complex promotes the removal of RNF43/ZNRF3 from the PM, preventing RNF43/ZNRF3 from degrading Fz and LRP6 (Malinauskas and Jones, 2014; de Lau et al., 2014). Thus, Rspo stabilizes Fz and LRP6 on the cell surface for Wnt signaling. LGR5 is an intestinal stem cell marker, and Rspo and Wnt proteins are key stem cell growth/self-renewal factors (de Lau et al., 2014).

The ‘Wnt-Off’ State: Keeping β -Catenin Levels Low

β -Catenin has two different roles in animal cells: it is an essential transcriptional co-activator in the Wnt pathway (as covered in this article), and it is also a critical component of the E-cadherin- β -catenin adherens junction at the PM (Valenta et al., 2012). Although only the cytoplasmic pool of β -catenin is under Wnt control, the junctional pool of β -catenin may crosstalk with the Wnt pathway in manners that remain to be understood.

Wnt regulation of β -catenin can be best described in a 'two-state' model (Clevers and Nusse, 2012; MacDonald *et al.*, 2009). In the 'Wnt-off' state, cytoplasmic β -catenin levels are kept low thanks to the so-called 'destruction complex,' whose main mission is to destruct β -catenin (Stamos and Weis, 2013). This multi-protein complex is assembled by the scaffolding protein Axin, and contains two kinases, casein kinase 1 α (CK1 α) and glycogen synthase kinase 3 (GSK3 α or GSK3 β), in addition to the APC (adenomatous polyposis coli) tumor suppressor/scaffolding protein and β -catenin (Figure 5). Axin is the ringleader and binds directly to each of the above core components, with help of additional interactions such as be-

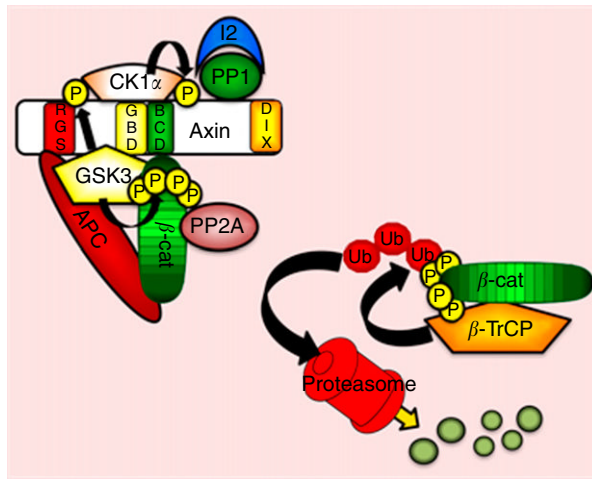


Figure 5 β -Catenin destruction complex. In absence of Wnt, Axin forms a complex with CK1 α , GSK3, APC, and β -catenin. CK1 α and GSK3 sequentially phosphorylate β -catenin, which creates a recognition motif for β -TrCP, thereby coupling β -catenin phosphorylation to its ubiquitination and degradation by the proteasome. In addition, phosphorylation of Axin by CK1 α , GSK3 promote Axin to adopt an 'open' conformation critical for the destruction complex assembly. Protein phosphatases 1 and 2A act to counter CK1 α and GSK3 actions.

tween APC and β -catenin (Figure 6). CK1 α and GSK3 are the executioners of the complex by sequentially phosphorylating β -catenin. Thus phosphorylation of S45 by CK1 α primes that of T41, S37, and S33 by GSK3; phosphorylation of S37 and S33 of β -catenin creates a recognition motif for binding by β -TrCP, a subunit of an E3 ubiquitin ligase complex, thereby coupling β -catenin phosphorylation to its ubiquitination and degradation by the proteasome. The destruction complex deconstructs β -catenin through phosphorylation-ubiquitination in a cyclic fashion.

Axin has a 'DIX' domain (Figure 6), which is also found in Disheveled (Dvl) (see below) and has the ability to polymerize (Schwarz-Romond *et al.*, 2007). Such DIX-mediated oligomerization of Axin is suggested to be a key feature of the destruction complex assembly, which has been referred to as the 'degradosome.' Although β -catenin phosphorylation by CK1 α and GSK3 is the central task of the destruction complex, Axin and APC themselves are substrates of CK1 α and GSK3, and their phosphorylation appear to be critical for the destruction complex assembly. Phosphorylation of Axin and APC enhances their binding to, and thus capture of, β -catenin (Willert *et al.*, 1999a; Spink *et al.*, 2000; Xing *et al.*, 2003), thereby facilitating destruction of β -catenin. In the case of Axin, this seems to be a result of phosphorylation-regulated conformational changes: phosphorylation by GSK3 (and likely CK1 α) promotes Axin to adopt an 'open' conformation that is conducive to its scaffolding function (Kim *et al.*, 2013b; Figure 5). Indeed protein phosphatases (PPs), which are also components of the destruction complex, act to counter CK1 α and GSK3 in multiple ways. PP1 dephosphorylates Axin to promote a 'closed' conformation through Axin intramolecular interaction/autoinhibition (Kim *et al.*, 2013b), whereas PP2A dephosphorylates β -catenin and prevents it from recognition- and ubiquitination by β -TrCP (Ikeda *et al.*, 2000; Seeling *et al.*, 1999; Figure 5). Therefore the destruction complex is an oiled machinery with accelerators (kinases) and brakes (phosphatases).

APC is a tumor suppressor in the gastrointestinal tract and its loss of function mutations are found in about 80% of

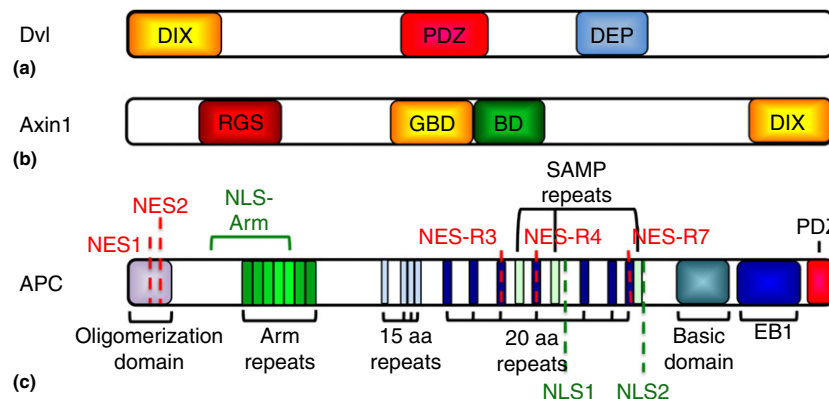


Figure 6 Hallmarks of Dvl, Axin1, and APC proteins. (A) Shows Dvl N-terminal DIX domain involved in oligomerization, PDZ and DEP domains. (B) Axin features a RGS domain, GSK3-binding domain (GBD), β -catenin-binding domain (BCD), and also a C-terminal DIX domain. (C) *Wt* APC has an oligomerization domain, followed by the Arm repeats, 15 and 20 a.a. repeats mediating β -catenin binding, SAMP motifs mediating Axin binding, a basic domain, EB1 and PDZ domains. APC nuclear localization sequences (NLS) and nuclear export sequences (NES) are also shown.

colorectal cancer cases (Polakis, 2012). But the biochemical nature of the APC protein in the destruction complex remains debated, with several (but not mutually exclusive) functions proposed in various models. These include assembly of the destruction complex, dynamic assembly and disassembly of the complex, inhibition of dephosphorylation of β -catenin, a requirement in both phosphorylation and ubiquitination of β -catenin, oligomerization, and cytoplasmic retention of β -catenin (Stamos and Weis, 2013). Another prominent but enigmatic component of the destruction complex is WTX (Wilms tumor gene on X) or Amer1, another tumor suppressor protein that is vertebrate-specific and required for β -catenin degradation (Regimbald-Dumas and He, 2011; Tanneberger et al., 2011).

Given the central role of Axin in the assembly of the destruction complex, regulation of Axin protein abundance/stability has attracted significant attention. Tankyrases (TNKS1 and TNKS2) promote Axin ubiquitination and proteasomal degradation through PARsylation of Axin (Zhang et al., 2011b; Huang et al., 2009). As such, small molecule inhibitors of TNKS have been identified that elevate Axin protein levels and thus decrease β -catenin protein amounts (Huang et al., 2009; Chen et al., 2009), highlighting a potential therapeutic strategy for inhibition of β -catenin signaling in cancer. However, TNKS does not appear to be under Wnt regulation.

The 'Wnt-On' State: Raising β -Catenin Levels High

Formation of the Wnt-receptor signaling complex/signalosome

Wnt binds to Fz and LRP6 and promotes Fz–LRP6 complex formation, triggering stabilization and accumulation of β -catenin (MacDonald and He, 2012; He et al., 2004). A hallmark of the receptor activation is Wnt-induced LRP6 phosphorylation on its five conserved PPPSPxS motifs and a few additional S/T residues outside these motifs (MacDonald and He, 2012). Dually phosphorylated PPPSPxS motifs are key signaling modules of LRP6 by serving as docking sites for Axin, thereby promoting the recruitment of the Axin destruction complex by the activated Wnt receptor complex. How Axin binds to phosphorylated LRP6 has not been resolved, with data and models supporting a direct (MacDonald et al., 2011) or indirect (Piao et al., 2008; Stamos et al., 2014) interaction.

A phosphorylated PPPSPxS motif under experimental conditions is sufficient to activate β -catenin signaling (Tamai et al., 2004), suggesting that LRP6 phosphorylation is the critical output of the Wnt receptor complex. Paradoxically, the kinases responsible for phosphorylation of the PPPSPxS motifs are GSK3 (on PPPSP) and CK1 (on xS) (MacDonald and He, 2012). In fact, GSK3 in complex with Axin carries out LRP6 phosphorylation (Zeng et al., 2008; Figure 7). Thus GSK3 and Axin, both of which have negative roles in Wnt signaling through promoting β -catenin phosphorylation and degradation, also have positive roles in Wnt signaling via promoting LRP6 phosphorylation, and these opposing roles of GSK3 in Wnt signaling have been experimentally demonstrated (Zeng et al., 2008, 2005). Multiple CK1 isoforms, including CK1 α , CK1 ϵ , and CK1 γ , may have roles in LRP6

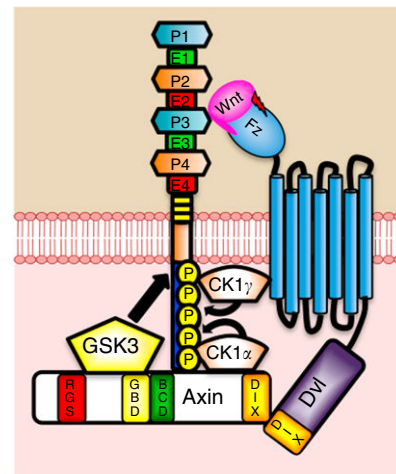


Figure 7 Wnt receptor activation. In the presence of Wnt ligand, a receptor complex forms between Fz and LRP6. Dvl interaction with Fz leads to Axin-GSK3 recruitment and phosphorylation of the PPPSP motif of LRP6 which primes xS phosphorylation by CK1, probably via CK1 γ (membrane-associated) and/or CK1 α and CK1 ϵ associated with Axin and Dvl. Partially phosphorylated LRP6 may be able to recruit and more efficiently bind Axin-GSK3 and promote more PPPSPxS phosphorylation.

phosphorylation (MacDonald and He, 2012). Additional kinases have been suggested to participate in LRP6 phosphorylation (MacDonald and He, 2012). Given its importance, most receptor proximal components are involved in controlling LRP6 phosphorylation, including Fz, the Fz cytoplasmic partner Dvl, and a number of Dvl-interacting proteins (MacDonald and He, 2012; Figure 7).

Dvl is an enigmatic scaffolding protein (Figure 6) and acts downstream of Fz, likely through direct Fz–Dvl binding (Tauriello et al., 2012; Wong et al., 2003). Unlike LRP5/6, which are specifically required for the Wnt/ β -catenin pathway, the Fz/Dvl pair is required for most, if not all, Wnt pathways (van Amerongen, 2012; Gao and Chen, 2010). Both Fz and Dvl functions are essential for Wnt-induced LRP6 phosphorylation, which depends on Fz–LRP6 proximity and Fz–Dvl interaction (Zeng et al., 2008). Like Axin, Dvl has a DIX domain (Figure 6) that mediates polymerization, as either homotypic Dvl assemblies or heterotypic Dvl–Axin assemblies (Schwarz-Romond et al., 2007). Dvl–Axin interaction appears to be critical for recruiting the Axin–GSK3 complex by Fz–Dvl into the Wnt receptor complex for phosphorylation of the PPPSP motif of LRP6 (Figure 7), which primes xS phosphorylation by CK1 (Zeng et al., 2008, 2005). Note that the phosphorylated PPPSPxS motif recruits Axin, which promotes PPPSPxS phosphorylation, creating a positive feed forward loop for Axin recruitment and LRP6 phosphorylation and supporting a Fz–LRP6 signaling ‘initiation–amplification’ model (MacDonald et al., 2008; Baig-Lewis et al., 2007). A related ‘LRP6 signalosome’ model, which posits LRP6 aggregation that depends on the property of Dvl polymerization and recruitment of Axin, has also been proposed (Schwarz-Romond et al., 2007; Bilic et al., 2007).

Several additional notable players have also been suggested to participate in LRP6 phosphorylation. Although a contentious subject, evidence exists to argue for a role of trimeric G-protein acting downstream of the Fz serpentine receptor (Katanaev *et al.*, 2005). A study suggests that the $G\beta\gamma$ subunits are associated with the LRP6 signaling complex including Dvl and Axin and recruit GSK3 to phosphorylate LRP6 (Jernigan *et al.*, 2010). Dvl is also associated with PI4K and PIP5K, two lipid kinases that stimulate PIP2 production to promote LRP6 phosphorylation (Pan *et al.*, 2008). PIP2 appears to enhance LRP6 aggregation or signalosome formation (Kim *et al.*, 2013a) and/or recruit WTX/Amer1, which in turn recruits Axin–GSK3 for LRP6 phosphorylation (Tanneberger *et al.*, 2011). These studies point to a multivalent assembly of the Fz–Dvl–LRP6–Axin signaling complex/signalosome, promoted by, and promoting, LRP6 phosphorylation.

β -Catenin Stabilization

How the Wnt receptor signaling complex/signalosome formation leads to β -catenin stabilization has been a contentious subject. A number of models have been proposed and reviewed (MacDonald and He, 2012), and we discuss four main and more recent models. An important consideration is β -catenin stabilization kinetics, which shows a rapid accumulation phase and a subsequent plateau phase upon Wnt treatment in mammalian cells (Kim *et al.*, 2013b; Hernandez *et al.*, 2012). This kinetics can be best explained in the following way: without Wnt, the rate of β -catenin protein synthesis equals that of its degradation (by the destruction complex), representing a steady state of no β -catenin accumulation; with Wnt, the rate of β -catenin degradation is slowed down and becomes less than that of β -catenin protein synthesis, leading to β -catenin accumulation until the level plateaus; this plateau represents a new steady state where the rate of β -catenin protein synthesis equals that of its degradation again, suggesting that the β -catenin destruction complex is fully operational at this steady state (in the presence of Wnt). A mechanistic model of β -catenin stabilization needs to account for this simple (but often underappreciated) kinetics.

A model of ‘Direct inhibition of GSK3 by LRP6’ (Figure 8(a)) was proposed based primarily on reconstitution *in vitro* (Piao *et al.*, 2008; Wu *et al.*, 2009; Cselenyi *et al.*, 2008). This model posits that LRP6 directly binds to GSK3 through phosphorylated PPPSPxS motifs and inhibits its phosphorylation of β -catenin, thereby stabilizing β -catenin. Within the LRP6 signaling complex/signalosome LRP6 and GSK3 are in proximity, and there is some evidence that a phosphorylated PPPSPxS motif binds to the catalytic cleft of GSK3 (Stamos *et al.*, 2014). However, it is unclear whether this binding represents GSK3–substrate or GSK3–inhibitor interaction. A related uncertainty remains whether LRP6 binds directly to Axin or GSK3 or both simultaneously. This static model has stringent requirements for the relative abundance of activated/phosphorylated LRP6 versus Axin and/or GSK3 in the cytoplasm.

A model of ‘MVB (multivesicular body) sequestration of GSK3’ suggests that after LRP6 signaling complex/signalosome formation, endocytosis of this complex leads to wrapping and thus sequestration of associated GSK3 into MVBs, thereby

separating GSK3 from β -catenin by two layers of membranes (Figure 8(b); Taelman *et al.*, 2010). Because GSK3 protein levels do not change during Wnt signaling, this model does not explain how GSK3 escapes MVBs, which normally destine cargos for lysosomal degradation. This model also requires that LRP6 recruits most or all GSK3 for sequestration, a challenging task if GSK3 is more abundant than LRP6. An important feature of this model is the ‘spillover’ effect in which Wnt signaling would inhibit most or all GSK3-mediated phosphorylation events.

While the prevailing view is that inhibition of β -catenin phosphorylation prevents β -catenin binding, and thus ubiquitination, by β -TrCP, a model of ‘Axin complex saturation’ argues that Wnt inhibits β -catenin binding and ubiquitination by β -TrCP without affecting β -catenin phosphorylation (Li *et al.*, 2012). In this scenario, phosphorylated but un-ubiquitinated and un-degraded β -catenin would remain in and saturate the Axin complex (that has been recruited to phosphorylated LRP6), thereby preventing the newly synthesized β -catenin from entering the destruction complex (Figure 8(c)). This model does not explain why phosphorylated β -catenin ‘prevents’ un-phosphorylated β -catenin from binding to the Axin complex based on protein thermodynamics and why phosphorylated β -catenin cannot be recognized by β -TrCP.

A more recent model is ‘Axin inactivation,’ which is based on the idea that phosphorylation regulates Axin intramolecular interaction and therefore its scaffolding function and destruction complex assembly (Kim *et al.*, 2013b; Figure 8(d)). Axin phosphorylation by GSK3 inhibits the intramolecular interaction, thereby keeping Axin in an active and ‘open’ conformation for β -catenin binding and poised for phosphorylated LRP6 binding; Axin dephosphorylation by PP1 promotes the intramolecular interaction, thereby rendering Axin in an inactive and ‘closed’ conformation that binds to neither β -catenin nor phosphorylated LRP6. Wnt signaling promotes Axin dephosphorylation and inactivation (Willert *et al.*, 1999a,b; Kim *et al.*, 2013b; Luo *et al.*, 2007). This model suggests that phosphorylated LRP6 binds to and recruits only active and phosphorylated Axin for inactivation in a cyclic manner. Further, this model posits that rising β -catenin levels upon Wnt signaling compete with and inhibit Axin intramolecular interaction, thereby promoting reassembly of the Axin destruction complex and potentially accounting for the steady state/plateau of β -catenin accumulation. Of the four models discussed here, the Axin inactivation model permits β -catenin stabilization regardless of relative abundance of LRP6 versus Axin, and is sufficient to explain β -catenin stabilization kinetics.

Wnt Signaling in the Nucleus

Nucleus Import

β -Catenin can shuttle between cytoplasm and nucleus. A classic nuclear protein contains a nuclear localization sequence (NLS), which strongly binds nuclear pore protein, importin/karyopherin. With the assistance of Ran GTPase, the cargo is moved through the nuclear envelop and released in the nucleus. However, β -catenin does not have a NLS. It has been found that β -catenin can enter nucleus in Ran-independent manner

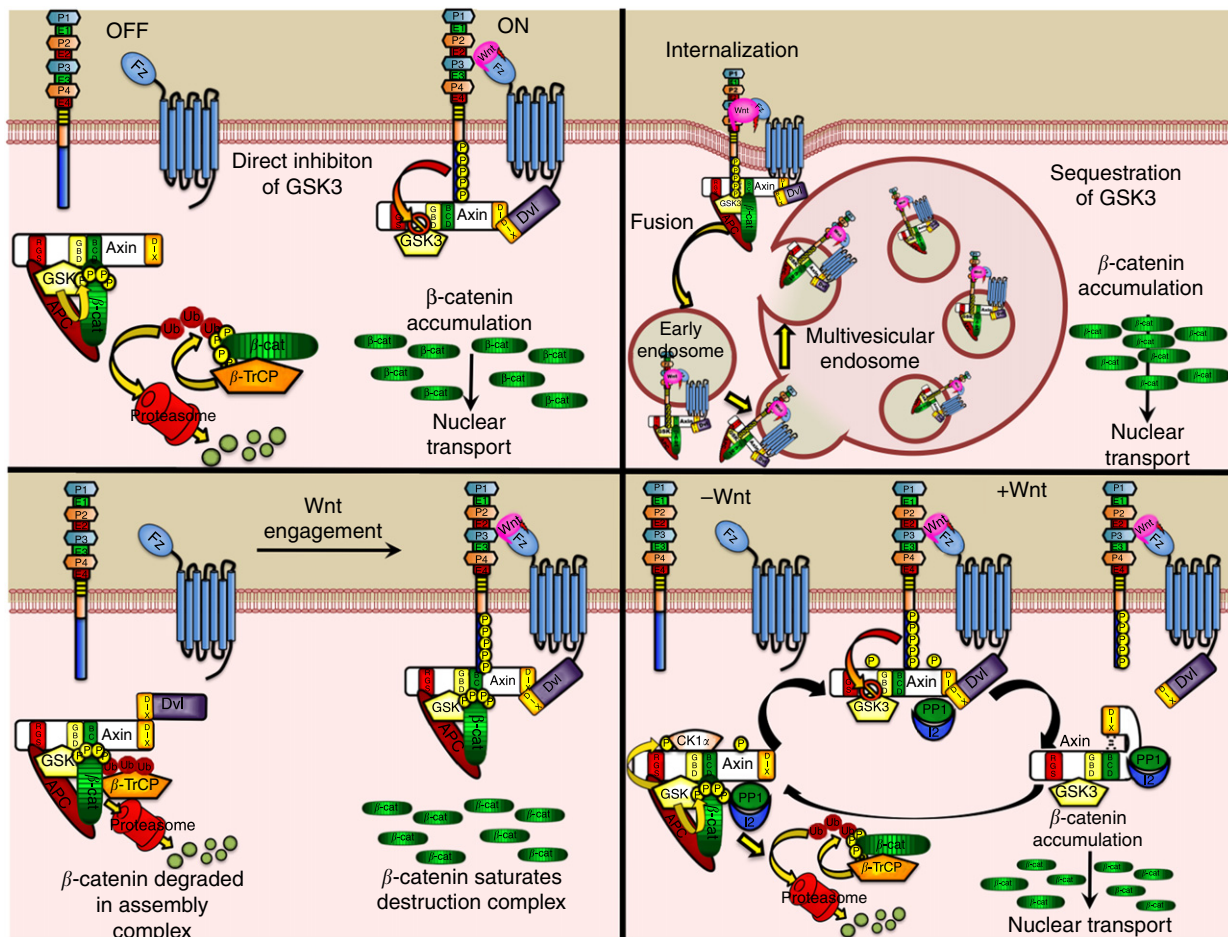


Figure 8 β -Catenin stabilization models. (a) In this model LRP6 directly binds to GSK3 through phosphorylated PPPSPxS motifs and inhibits phosphorylation of β -catenin, thereby stabilizing β -catenin. (b) Another model suggests that after LRP6 signaling complex/signalsome formation, endocytosis of this complex leads to wrapping and sequestration of associated GSK3 into multivesicular bodies. (c) In this view, Wnt inhibits β -catenin binding and ubiquitination by β -TrCP without affecting β -catenin phosphorylation. (d) This model suggests that Axin phosphorylation by GSK3 inhibits the intramolecular interaction, thereby keeping Axin in an active and 'open' conformation for β -catenin binding and poised for phosphorylated LRP6 binding; Axin dephosphorylation by PP1 promotes the intramolecular interaction, thereby rendering Axin in an inactive and 'closed' conformation that binds to neither β -catenin nor phosphorylated LRP6.

(Yokoya *et al.*, 1999). Since β -catenin and importin share similar armadillo-repeat domains, β -catenin might directly interact with nuclear pore components, bypassing the importin/karyopherin proteins (Yokoya *et al.*, 1999; Fagotto *et al.*, 1998). β -Catenin nuclear localization is regulated by phosphorylation. Protein kinase AKT can phosphorylate β -catenin at Ser552 and promote β -catenin nuclear localization (Fang *et al.*, 2007; He *et al.*, 2007). In response to Rac1 activation, JNK2 also phosphorylates β -catenin at Ser191 and Ser605 and regulate nuclear translocation of β -catenin (Wu *et al.*, 2008). In glioma cells, FoxM1, a forkhead box transcription factor, interacts with β -catenin and promotes its nuclear translocation (Zhang *et al.*, 2011a). Whether FoxM1 controls β -catenin nuclear localization in normal cells is not known.

Nuclear Export

β -Catenin can be shuttled in and out of the nucleus in response to Wnt signaling. β -Catenin does not contain a classical nuclear

export sequence (NES). β -Catenin nuclear export is controlled by other proteins in the Wnt pathway, such as Axin (Cong and Varmus, 2004) and APC (Rosin-Arbesfeld *et al.*, 2000; Henderson, 2000; Neufeld *et al.*, 2000). These components in the β -catenin destruction complex also participate in nuclear export of β -catenin. APC contains several NES sequences, which bind exportin, another member of the karyoporin family (f. Binding of Ran-GTP to exportin enhances its affinity to NES and promotes the export of exportin and its cargo through nuclear pore. Ran GTPase releases the cargo and recycles exportin to the nucleus by converting Ran-GTP to Ran-GDP. APC may act as a shuttle for β -catenin nuclear export. Some factors may not directly participate in shuttling, but rather as an 'anchor' to retain β -catenin within their respective compartments (Krieghoff *et al.*, 2006). The nuclear proteins described below, such as TCF-4, Pygopus, and BCL9 may function as nuclear 'anchors' (Townesley *et al.*, 2004). E-Cadherin function as a membrane 'anchor' and Axin may function as a cytoplasmic 'anchor' of β -catenin (Tolwinski and Wieschaus, 2001).

β -Catenin/TCF Complex

β -Catenin is a transcription co-activator that interacts with TCF/LEF (T-cell factor/lymphoid enhancer-binding factor) family of transcription factors in the nucleus (MacDonald *et al.*, 2009). The TCF/LEF factors are high-mobility group (HMG) DNA-binding proteins. There is only one TCF gene in *Drosophila* and worm. In mammalian cells, the TCF/LEF family includes TCF1 (TCF7 by HUGO), LEF-1, TCF-3 (TCF7L1), and TCF-4 (TCF7L2). Alternative promoters and splicing further expand the variety of TCF/LEF isoforms generating context dependent patterns of activity. The N-terminus of TCF/LEF binds β -catenin and the HMG domain binds a specific DNA sequence called Wnt responsive elements (WRE). The consensus WRE sequence is: CCTTTGNN (N represents either T or A). The TCF in invertebrates and some spliced form of TCF/LEF proteins in vertebrates have a second highly conserved DNA-binding domain referred to as 'C Clamp,' which binds to GC-rich sequences referred to as 'helper sites' that are found in proximity to some WREs (Figure 9(a); Bhambhani *et al.*, 2014; Hoverter *et al.*, 2014). Some alternative splicing products of TCF/LEF lack the C Clamp or other regions in N-terminal of the HMG domain (Van de Wetering *et al.*, 1996; Prokunina-Olsson *et al.*, 2009). Many Wnt target genes contain

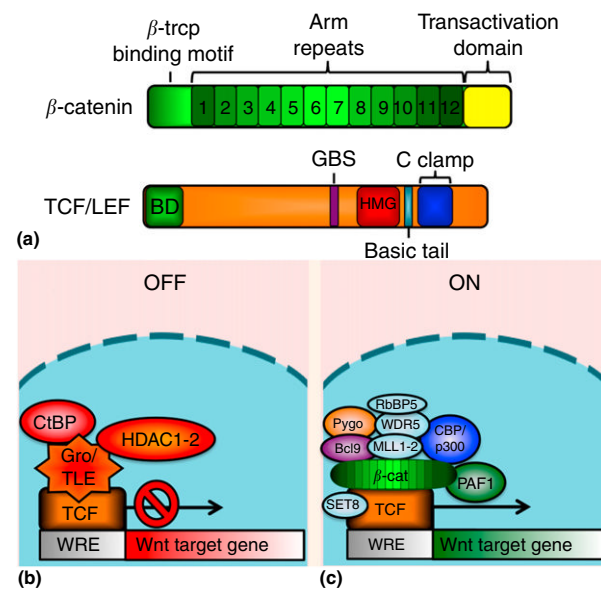


Figure 9 β -Catenin/TCF complex. (a) Hallmarks of β -catenin and TCF/LEF proteins. β -Catenin has a β -TrCP-binding motif in N-terminal, 12 armadillo repeats and a transactivation domain. TCF/LEF proteins have a β -catenin-binding domain (BD), a groucho-binding site (GBS) followed by the high-mobility group (HMG), a basic tail or NLS and a C clamp domain. (b) In absence of Wnt, target genes are constitutively inhibited by nuclear TCF in complex with corepressors such as CtBP, HDAC1, HDAC2, and groucho/TLE on Wnt response elements (WREs). (c) In presence of Wnt, β -catenin translocates to the nucleus, where it displaces TCF-bound corepressors or co-imports additional TCF to occupy WREs. Once bound, β -catenin functions as a scaffold to recruit co-activators involved in chromatin remodeling such as histone acetyltransferases CBP and P300 and histone methyltransferases MLL1-2, WDR5, RbBP5 and SET8, and components of the PAF1 complex itself interacting with RNA polymerase II to induce Wnt target gene expression.

multiple WRE sequences in the promoter region or distant sites (Hatzis *et al.*, 2008). Different sets of target genes are regulated by WREs with or without helper sites, suggesting an additional layer of regulation in *Drosophila* embryos and mammalian cells (Bhambhani *et al.*, 2014; Hoverter *et al.*, 2014).

TCF/LEF factors act as both transcription activators and repressors depending on the status of Wnt signaling. Without Wnt stimulation, TCF/LEF factors recruit corepressors such as CtBP, HDAC1, MTGR1, Coop, HIC5, and the most studied Groucho/TLE to inhibit transcription (Figure 9(b); MacDonald *et al.*, 2009; Cadigan and Waterman, 2012). In response to Wnt signaling, β -catenin is stabilized in the cytoplasm and enters nucleus, where it binds TCF/LEF, displaces Groucho/TLE from TCF/LEF and activates transcription by recruiting other co-activators (Figure 9(c)). Recently, it was found that the N-terminal domain of the Groucho/TLE1 forms a tetramer that specifically binds to K20 methylated Histone H4 tail (Chodaparambil *et al.*, 2014). Some factors regulate this competition between β -catenin and Groucho/TLE1 for TCF/LEF binding. In zebrafish the transcription factor Ladybird homeobox 2 (Lbx2) enhances Wnt/ β -catenin signaling pathway in the posterior lateral and ventral mesoderm during gastrulation, by directly interfering with Groucho/TLE binding to TCF (Lu *et al.*, 2014). This TCF/LEF transcriptional switch model is supported by observations in *Drosophila* and *Caenorhabditis elegans*. However, in vertebrates individual TCF/LEF have distinct ability to activate or repress Wnt target gene expression with TCF-3 generally known as a repressor and TCF1 and LEF-1 more often associated with gene activation (MacDonald *et al.*, 2009; Cadigan and Waterman, 2012). In addition, the switch between active and recessive forms of TCF/LEF could be mediated by factors other than the direct competition between β -catenin and Groucho/TLE1 for TCF/LEF binding. For instance, in *Xenopus* Wnt promotes the phosphorylation of the repressor TCF-3 by HIPK2, which lead to its dissociation, and replacement by TCF1 from *Vent2* gene WRE (Hikasa *et al.*, 2010).

Co-Activators

β -Catenin has multiple and extensive protein-protein interaction domains. The N-terminus contains the destructing box that upon phosphorylation binds β -TrCP. The C-terminus contains a transactivation domain, which binds histone acetyltransferase (HAT) CBP/p300 (MacDonald *et al.*, 2009; Cadigan and Waterman, 2012). The armadillo-repeat domain in the middle of β -catenin binds E-cadherin on the PM and Axin and APC in the cytoplasm and nucleus. In addition, the N-terminal armadillo repeats interact with transcription co-activators, BCL9/Legless and Pygopus (Figure 9(c); MacDonald *et al.*, 2009; Cadigan and Waterman, 2012).

HATs

Histone modification is required for Wnt-mediated transcription. CBP and p300 are HATs that acetylate lysines on histones. This modification is essential to reduce the overall positive charge and to loosen chromatin at Wnt target genes, in order to facilitate binding of transcription factors and activate transcription (Ogryzko *et al.*, 1996; Goldman *et al.*, 1997).

HATs also acetylate β -catenin and TCF and regulate the activity of the TCF complex (Levy *et al.*, 2004). Recently, Walker *et al.* (2015) showed that FOXP1 promotes acetylation of β -catenin by CBP and potentiation of β -catenin-dependent transcription. Histone deacetylases (HDACs) have opposite roles as HATs. It has been found that HDAC1 and HDAC2 disrupt the β -catenin/TCF complex and inhibit Wnt signaling (Figure 9(b); Ye *et al.*, 2009). Corepressor Groucho/TLE may maintain a repressive chromatin environment by interacting with hypoacetylated histone H3 (Palaparti *et al.*, 1997). However, there are four classes of HDACs. HDAC1 and HDAC2 belong to class I HDACs (Barneda-Zahonero and Parra, 2012). Whether other HDACs have roles in Wnt signaling is unclear.

Histone Methyltransferases

Histone methyltransferases (HMTs) are also involved in Wnt target gene activation. It has been reported that β -catenin interacts with Ash2L, a member of the Set1/MLL HMT complexes (Sierra *et al.*, 2006). There are six members of Set1/MLL HMTs; Ash2L, along with WDR5 and RbBP5 are shared subunits of these Set1/MLL complexes (Krivtsov and Armstrong, 2007). The Set1/MLL complexes catalyze Histone 3 lysine 4 trimethylation (H3K4Me3), which is generally required for gene expression including Wnt target gene expression. Pygopus also interacts with MLL2 and Ash2L (Chen *et al.*, 2010). Pygopus is recruited to the β -catenin/TCF complex by BCL9/Legless (Figure 9(c); Kramps *et al.*, 2002). Pygopus contains a plant homeodomain (PHD domain), which interacts with methylated histone H3 (Soliman and Riabowol, 2007). The PHD domain of Pygopus preferably binds dimethylated H3K4 upon interacting with BCL9 (Mosimann *et al.*, 2009), suggesting that Pygopus reads the 'histone code' and acts as an additional anchor for the β -catenin/TCF complex to chromatin. Parafibromin, a component of the PAF1 complex, is essential for the initiation and elongation steps of transcription through its interaction with RNA Polymerase II. The PAF1 complex interacts with the C-terminal transactivation domain of β -catenin and is required for transactivation (Figure 9(c)). Overexpression of Parafibromin can compensate for loss of BCL9/Legless *in vivo* (Mosimann *et al.*, 2006).

Other HMTs may also regulate Wnt-mediated transcription. For example, HMT SET8 directly interacts with TCF/LEF and regulates H4K20 monomethylation (H3K4Me1). Thus, Wnt stimulation increases H3K4Me1 levels at Wnt target gene promoters, indicating H3K4Me1 is a marker for gene activation in Wnt signaling (Li *et al.*, 2011). Dot1 (disruptor of telomeric silencing) is a lysine methyltransferase in *Drosophila*. Dot1 and its associated factors have been linked to H3K79 trimethylation and activation of Wnt/Wingless target genes in *Drosophila*, zebrafish, and human colorectal cancer cells (Mahmoudi *et al.*, 2010; Mohan *et al.*, 2010). However, DOT1L-mediated H3K79 methylation appears to be dispensable for Wnt pathway functions in the mouse intestine (Ho *et al.*, 2013).

Other Co-activators and Corepressors

β -Catenin interacts with many other co-activators and corepressors, such as the ATP-dependent chromatin remodeling

factors Brg-1/Brahma (Barker *et al.*, 2001) and the ATP-dependent helicase TIP49a/Pontin52 (Bauer *et al.*, 1998, 2000). Nuclear β -catenin antagonists ICAT and Chibby bind the C-terminus of β -catenin and disrupt the β -catenin/TCF complex (Tago *et al.*, 2000; Takemaru and Moon, 2000). Many transcription factors also interact and regulate the β -catenin/TCF complex. For example, a zinc finger protein, KLF4, interacts with both β -catenin and TCF and represses Wnt signaling in the intestine by inhibiting the binding of co-activator, p300/CBP (Evans *et al.*, 2010). However, the closely related transcription factor KLF5 highly expressed in the intestinal crypt compartment promotes cell growth and tumorigenesis by facilitating β -catenin-TCF-4 complex formation (Nakaya *et al.*, 2014; Guo *et al.*, 2015). In hepatocellular carcinoma SOX1 acts as a tumor suppressor. GST-pull-down and co-immunoprecipitation demonstrated that SOX1 interacts with β -catenin but not with the β -catenin/TCF complex (Tsao *et al.*, 2012). In mouse muscle progenitor cells, Barx2 and Pax7 have competing effects on Wnt effectors β -catenin and TCF-4. Barx2 promotes recruitment of β -catenin to TCF/lymphoid enhancer factor sites, while Pax7 represses Wnt/ β -catenin reporter and Barx2 effect (Zhuang *et al.*, 2014). In *Xenopus laevis*, co-injection of mRNA from β -catenin along with any member of the zinc finger family Zic (Zic1 to Zic5) leads to suppression of β -catenin transcriptional activity. Zic2 directly binds to TCF-4 HMG box while Zic3 interacts with overexpressed TCF1 (Pourelbrahim *et al.*, 2011; Fujimi *et al.*, 2012). Kindlin2, better known as a focal adhesion protein, was found to interact directly with active β -catenin and form a tripartite complex with β -catenin and TCF-4. Kindlin2 strengthens Wnt target gene expression in MCF7 breast cancer cells (Yu *et al.*, 2012). Some transcription factors not only enhance Wnt target gene expression but also shape genetic programs by binding to DNA elements found in a subset of Wnt target gene. Genome-wide analysis of *cis*-regulatory regions from colonic cells showed co-occupancy of CDX2 and TCF-4 across short distances. Another interesting example is the stem cell marker and transcription factor Ascl2 that binds DNA elements nearby WRE and drives the intestinal stem cell program (Schuijers *et al.*, 2015). Other signaling pathways also regulate Wnt signaling in the nucleus. TCF/LEF can be phosphorylated by Nemo-like kinase (NLK), which is activated by TAK1 (TGF β -activated kinase 1). The TAK1-NLK1 pathway inhibits the interaction of β -catenin/TCF complex with DNA (Ishitani *et al.*, 1999). Another kinase called TNIK is recruited by β -catenin on TCFs for phosphorylation and enhances Wnt target gene activation (Mahmoudi *et al.*, 2009).

Conclusions

Since the discovery of Wnt1 in 1982, our understanding of Wnt proteins, Wnt/ β -catenin and other Wnt signaling pathways, and their involvement in broad biological processes has exploded beyond anyone's imagination. Wnt signaling and its studies have permeated almost all realms of modern biological and biomedical research, and exciting discoveries are continuing to be made in and around Wnt signaling in development, stem cell, and cancer biology. As many fundamental questions of Wnt signaling remain to be understood, and the

potential of therapeutic targeting of Wnt signaling for cancer and disease treatment is yet to be fully realized, these unanswered questions and unmet needs present great challenges and also opportunities for generations of Wnt researchers.

Acknowledgments

Chunming Liu acknowledges grant support from the National Institutes of Health (NIH) (RO1 DK071976). Yannik Regimbald-Dumas is in part supported by a postdoctoral fellowship from FRSQ (QC, Canada). Xi He acknowledges grant support from NIH (RO1 GM057603, RO1 GM074241, and RO1 AR060359) and Boston Children's Hospital Intellectual and Developmental Disabilities Research Center (P30 HD-18655).

References

- Ahn, V.E., Chu, M.L., Choi, H.J., *et al.*, 2011. Structural basis of Wnt signaling inhibition by Dickkopf binding to LRP5/6. *Developmental Cell* 21 (5), 862–873.
- van Amerongen, R., 2012. Alternative Wnt pathways and receptors. *Cold Spring Harbor Perspectives in Biology* 4, 10.
- Baig-Lewis, S., Peterson-Nedry, W., Wehrli, M., 2007. Wingless/Wnt signal transduction requires distinct initiation and amplification steps that both depend on Arrow/LRP. *Developmental Biology* 306 (1), 94–111.
- Barker, N., Hurlstone, A., Musisi, H., *et al.*, 2001. The chromatin remodelling factor Brg-1 interacts with beta-catenin to promote target gene activation. *EMBO Journal* 20 (17), 4935–4943.
- Barneda-Zahonero, B., Parra, M., 2012. Histone deacetylases and cancer. *Molecular Oncology* 6 (6), 579–589.
- Bauer, A., Chauvet, S., Huber, O., *et al.*, 2000. Pontin52 and reptin52 function as antagonistic regulators of beta-catenin signalling activity. *EMBO Journal* 19 (22), 6121–6130.
- Bauer, A., Huber, O., Kemler, R., 1998. Pontin52, an interaction partner of beta-catenin, binds to the TATA box binding protein. *Proceedings of the National Academy of Sciences of the United States of America* 95 (25), 14787–14792.
- Bazan, J.F., Janda, C.Y., Garcia, K.C., 2012. Structural architecture and functional evolution of Wnts. *Developmental Cell* 23 (2), 227–232.
- Bhambhani, C., Ravindranath, A.J., Mentink, R.A., *et al.*, 2014. Distinct DNA binding sites contribute to the TCF transcriptional switch in *C. elegans* and *Drosophila*. *PLoS Genetics* 10 (2), e1004133.
- Bilic, J., Huang, Y.L., Davidson, G., *et al.*, 2007. Wnt induces LRP6 signalosomes and promotes dishevelled-dependent LRP6 phosphorylation. *Science* 316 (5831), 1619–1622.
- Bourhis, E., Tam, C., Franke, Y., *et al.*, 2010. Reconstitution of a frizzled8.Wnt3a. LRP6 signaling complex reveals multiple Wnt and Dkk1 binding sites on LRP6. *Journal of Biological Chemistry* 285 (12), 9172–9179.
- Bourhis, E., Wang, W., Tam, C., *et al.*, 2011. Wnt antagonists bind through a short peptide to the first beta-propeller domain of LRP5/6. *Structure* 19 (10), 1433–1442.
- Cadigan, K.M., Waterman, M.L., 2012. TCF/LEFs and Wnt signaling in the nucleus. *Cold Spring Harbor Perspectives in Biology* 4, 11.
- Chen, B., Dodge, M.E., Tang, W., *et al.*, 2009. Small molecule-mediated disruption of Wnt-dependent signaling in tissue regeneration and cancer. *Nature Chemical Biology* 5 (2), 100–107.
- Chen, J., Luo, Q., Yuan, Y., *et al.*, 2010. Pygo2 associates with MLL2 histone methyltransferase and GCN5 histone acetyltransferase complexes to augment Wnt target gene expression and breast cancer stem-like cell expansion. *Molecular and Cellular Biology* 30 (24), 5621–5635.
- Chen, S., Bubeck, D., MacDonald, B.T., *et al.*, 2011. Structural and functional studies of LRP6 ectodomain reveal a platform for Wnt signaling. *Developmental Cell* 21 (5), 848–861.
- Chodaparambil, J.V., Pate, K.T., Hepler, M.R., *et al.*, 2014. Molecular functions of the TLE tetramerization domain in Wnt target gene repression. *EMBO Journal* 33 (7), 719–731.
- Clevers, H., Loh, K.M., Nusse, R., 2014. Stem cell signaling. An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control. *Science* 346 (6205), 1248012.
- Clevers, H., Nusse, R., 2012. Wnt/beta-catenin signaling and disease. *Cell* 149 (6), 1192–1205.
- Cong, F., Varmus, H., 2004. Nuclear-cytoplasmic shuttling of Axin regulates subcellular localization of beta-catenin. *Proceedings of the National Academy of Sciences of the United States of America* 101 (9), 2882–2887.
- Cruciat, C.M., Niehrs, C., 2013. Secreted and transmembrane wnt inhibitors and activators. *Cold Spring Harbor Perspectives in Biology* 5 (3), a015081.
- Cselenyi, C.S., Jernigan, K.K., Tahinci, E., *et al.*, 2008. LRP6 transduces a canonical Wnt signal independently of Axin degradation by inhibiting GSK3's phosphorylation of beta-catenin. *Proceedings of the National Academy of Sciences of the United States of America* 105 (23), 8032–8037.
- Evans, P.M., Chen, X., Zhang, W., Liu, C., 2010. KLF4 interacts with beta-catenin/TCF4 and blocks p300/CBP recruitment by beta-catenin. *Molecular and Cellular Biology* 30 (2), 372–381.
- Fagotto, F., Gluck, U., Gumbiner, B.M., 1998. Nuclear localization signal-independent and importin/karyopherin-independent nuclear import of beta-catenin. *Current Biology* 8 (4), 181–190.
- Fang, D., Hawke, D., Zheng, Y., *et al.*, 2007. Phosphorylation of beta-catenin by AKT promotes beta-catenin transcriptional activity. *Journal of Biological Chemistry* 282 (15), 11221–11229.
- Filmus, J., Capurro, M., Rast, J., 2008. Glypicans. *Genome Biology* 9 (5), 224.
- Fujimi, T.J., Hatayama, M., Aruga, J., 2012. Xenopus Zic3 controls notochord and organizer development through suppression of the Wnt/beta-catenin signaling pathway. *Developmental Biology* 361 (2), 220–231.
- Furushima, K., Yamamoto, A., Nagano, T., *et al.*, 2007. Mouse homologues of Shisa antagonistic to Wnt and Fgf signalings. *Developmental Biology* 306 (2), 480–492.
- Gao, C., Chen, Y.G., 2010. Dishevelled: The hub of Wnt signaling. *Cellular Signaling* 22 (5), 717–727.
- Goldman, P.S., Tran, V.K., Goodman, R.H., 1997. The multifunctional role of the co-activator CBP in transcriptional regulation. *Recent Progress in Hormone Research* 52, 103–119. Discussion 119–120.
- Green, J., Nusse, R., van Amerongen, R., 2014. The role of Ryk and Ror receptor tyrosine kinases in Wnt signal transduction. *Cold Spring Harbor Perspectives in Biology* 6 (2), a009175.
- Guo, L., He, P., No, Y.R., Yun, C.C., 2015. Kruppel-like factor 5 incorporates into the beta-catenin/TCF complex in response to LPA in colon cancer cells. *Cellular Signaling* 27 (5), 961–968.
- Hao, H.X., Xie, Y., Zhang, Y., *et al.*, 2012. ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner. *Nature* 485 (7397), 195–200.
- Hatzis, P., van der Flier, L.G., van Driel, M.A., *et al.*, 2008. Genome-wide pattern of TCF7L2/TCF4 chromatin occupancy in colorectal cancer cells. *Molecular and Cellular Biology* 28 (8), 2732–2744.
- Hausmann, G., Banziger, C., Basler, K., 2007. Helping Wingless take flight: How WNT proteins are secreted. *Nature Reviews. Molecular Cell Biology* 8 (4), 331–336.
- He, X., Semenov, M., Tamai, K., Zeng, X., 2004. LDL receptor-related proteins 5 and 6 in Wnt/beta-catenin signaling: Arrows point the way. *Development* 131 (8), 1663–1677.
- He, X.C., Yin, T., Grindley, J.C., *et al.*, 2007. PTEN-deficient intestinal stem cells initiate intestinal polyposis. *Nature Genetics* 39 (2), 189–198.
- Henderson, B.R., 2000. Nuclear-cytoplasmic shuttling of APC regulates beta-catenin subcellular localization and turnover. *Nature Cell Biology* 2 (9), 653–660.
- Hernandez, A.R., Klein, A.M., Kirschner, M.W., 2012. Kinetic responses of beta-catenin specify the sites of Wnt control. *Science* 338 (6112), 1337–1340.
- Hikasa, H., Ezan, J., Itoh, K., *et al.*, 2010. Regulation of TCF3 by Wnt-dependent phosphorylation during vertebrate axis specification. *Developmental Cell* 19 (4), 521–532.
- Ho, L.L., Sinha, A., Verzi, M., *et al.*, 2013. DOT1L-mediated H3K79 methylation in chromatin is dispensable for Wnt pathway-specific and other intestinal epithelial functions. *Molecular and Cellular Biology* 33 (9), 1735–1745.
- Hoverter, N.P., Zeller, M.D., McQuade, M.M., *et al.*, 2014. The TCF C-clamp DNA binding domain expands the Wnt transcriptome via alternative target recognition. *Nucleic Acids Research* 42 (22), 13615–13632.
- Huang, S.M., Mishina, Y.M., Liu, S., *et al.*, 2009. Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* 461 (7264), 614–620.
- Ikeda, S., Kishida, M., Matsuura, Y., Usui, H., Kikuchi, A., 2000. GSK-3beta-dependent phosphorylation of adenomatous polyposis coli gene product can be modulated by beta-catenin and protein phosphatase 2A complexed with Axin. *Oncogene* 19 (4), 537–545.

- Ishitani, T., Ninomiya-Tsuji, J., Nagai, S., *et al.*, 1999. The TAK1-NLK-MAPK-related pathway antagonizes signalling between beta-catenin and transcription factor TCF. *Nature* 399 (6738), 798–802.
- Janda, C.Y., Waghay, D., Levin, A.M., Thomas, C., Garcia, K.C., 2012. Structural basis of Wnt recognition by Frizzled. *Science* 337 (6090), 59–64.
- Jernigan, K.K., Cselenyi, C.S., Thorne, C.A., *et al.*, 2010. Gbetagamma activates GSK3 to promote LRP6-mediated beta-catenin transcriptional activity. *Science Signaling* 3 (121), ra37.
- Junge, H.J., Yang, S., Burton, J.B., *et al.*, 2009. TSPAN12 regulates retinal vascular development by promoting Norrin- but not Wnt-induced FZD4/beta-catenin signaling. *Cell* 139 (2), 299–311.
- Kakugawa, S., Langton, P.F., Zebisch, M., *et al.*, 2015. Notum deacylates Wnt proteins to suppress signalling activity. *Nature* 519, 187–192.
- Katanaev, V.L., Ponzelli, R., Semeriva, M., Tomlinson, A., 2005. Trimeric G protein-dependent frizzled signaling in *Drosophila*. *Cell* 120 (1), 111–122.
- Ke, J., Harikumar, K.G., Erice, C., *et al.*, 2013. Structure and function of Norrin in assembly and activation of a Frizzled 4-Lrp5/6 complex. *Genes & Development* 27 (21), 2305–2319.
- Kim, I., Pan, W., Jones, S.A., *et al.*, 2013a. Clathrin and AP2 are required for Ptdlns (4,5)P2-mediated formation of LRP6 signalosomes. *Journal of Cell Biology* 200 (4), 419–428.
- Kim, S.E., Huang, H., Zhao, M., *et al.*, 2013b. Wnt stabilization of beta-catenin reveals principles for morphogen receptor-scaffold assemblies. *Science* 340 (6134), 867–870.
- Koo, B.K., Spit, M., Jordens, I., *et al.*, 2012. Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. *Nature* 488 (7413), 665–669.
- Kramps, T., Peter, O., Brunner, E., *et al.*, 2002. Wnt/wingless signaling requires BCL9/legless-mediated recruitment of pygopus to the nuclear beta-catenin-TCF complex. *Cell* 109 (1), 47–60.
- Krieghoff, E., Behrens, J., Mayr, B., 2006. Nucleo-cytoplasmic distribution of beta-catenin is regulated by retention. *Journal of Cell Science* 119 (Pt 7), 1453–1463.
- Krivtsov, A.V., Armstrong, S.A., 2007. MLL translocations, histone modifications and leukaemia stem-cell development. *Nature Reviews Cancer* 7 (11), 823–833.
- de Lau, W., Peng, W.C., Gros, P., Clevers, H., 2014. The R-spondin/Lgr5/Rnf43 module: Regulator of Wnt signal strength. *Genes & Development* 28 (4), 305–316.
- Levy, L., Wei, Y., Labalette, C., *et al.*, 2004. Acetylation of beta-catenin by p300 regulates beta-catenin–Tcf4 interaction. *Molecular and Cellular Biology* 24 (8), 3404–3414.
- Li, V.S., Ng, S.S., Boersema, P.J., *et al.*, 2012. Wnt signaling through inhibition of beta-catenin degradation in an intact Axin1 complex. *Cell* 149 (6), 1245–1256.
- Li, Z., Nie, F., Wang, S., Li, L., 2011. Histone H4 Lys 20 monomethylation by histone methylase SET8 mediates Wnt target gene activation. *Proceedings of the National Academy of Sciences of the United States of America* 108 (8), 3116–3123.
- Liu, C.C., Pearson, C., Bu, G., 2009. Cooperative folding and ligand-binding properties of LRP6 beta-propeller domains. *Journal of Biological Chemistry* 284 (22), 15299–15307.
- Lu, F.I., Sun, Y.H., Wei, C.Y., Thisse, C., Thisse, B., 2014. Tissue-specific derepression of TCF/LEF controls the activity of the Wnt/beta-catenin pathway. *Nature Communications* 5, 5368.
- Luo, W., Peterson, A., Garcia, B.A., *et al.*, 2007. Protein phosphatase 1 regulates assembly and function of the beta-catenin degradation complex. *EMBO Journal* 26 (6), 1511–1521.
- MacDonald, B.T., Yokota, C., Tamai, K., Zeng, X., He, X., 2008. Wnt signal amplification via activity, cooperativity, and regulation of multiple intracellular PPPSP motifs in the Wnt co-receptor LRP6. *Journal of Biological Chemistry* 283 (23), 16115–16123.
- MacDonald, B.T., Semenov, M.V., Huang, H., He, X., 2011. Dissecting molecular differences between Wnt coreceptors LRP5 and LRP6. *PLoS One* 6 (8), e23537.
- MacDonald, B.T., He, X., 2012. Frizzled and LRP5/6 receptors for Wnt/beta-catenin signaling. *Cold Spring Harbor Perspectives in Biology* 4, 12.
- MacDonald, B.T., Tamai, K., He, X., 2009. Wnt/beta-catenin signaling: Components, mechanisms, and diseases. *Developmental Cell* 17 (1), 9–26.
- Mahmoudi, T., Li, V.S., Ng, S.S., *et al.*, 2009. The kinase TNIK is an essential activator of Wnt target genes. *EMBO Journal* 28 (21), 3329–3340.
- Mahmoudi, T., Boj, S.F., Hatzis, P., *et al.*, 2010. The leukemia-associated Mllt10/Af10-Dot11 are Tcf4/beta-catenin coactivators essential for intestinal homeostasis. *PLoS Biology* 8 (11), e1000539.
- Malinauskas, T., Jones, E.Y., 2014. Extracellular modulators of Wnt signalling. *Current Opinion in Structural Biology* 29, 77–84.
- Mii, Y., Taira, M., 2009. Secreted Frizzled-related proteins enhance the diffusion of Wnt ligands and expand their signalling range. *Development* 136 (24), 4083–4088.
- Mohan, M., Herz, H.M., Takahashi, Y.H., *et al.*, 2010. Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom). *Genes & Development* 24 (6), 574–589.
- Mosimann, C., Hausmann, G., Basler, K., 2006. Parafibromin/Hyrax activates Wnt/Wg target gene transcription by direct association with beta-catenin/Armadillo. *Cell* 125 (2), 327–341.
- Mosimann, C., Hausmann, G., Basler, K., 2009. Beta-catenin hits chromatin: Regulation of Wnt target gene activation. *Nature Reviews Molecular Cell Biology* 10 (4), 276–286.
- Mulligan, K.A., Fuerer, C., Ching, W., *et al.*, 2012. Secreted Wingless-interacting molecule (Swim) promotes long-range signaling by maintaining Wingless solubility. *Proceedings of the National Academy of Sciences of the United States of America* 109 (2), 370–377.
- Nakaya, T., Ogawa, S., Manabe, I., *et al.*, 2014. KLF5 regulates the integrity and oncogenicity of intestinal stem cells. *Cancer Research* 74 (10), 2882–2891.
- Neufeld, K.L., Zhang, F., Cullen, B.R., White, R.L., 2000. APC-mediated downregulation of beta-catenin activity involves nuclear sequestration and nuclear export. *EMBO Reports* 1 (6), 519–523.
- Neumann, S., Coudreuse, D.Y., van der Westhuyzen, D.R., *et al.*, 2009. Mammalian Wnt3a is released on lipoprotein particles. *Traffic* 10 (3), 334–343.
- Nusse, R., Brown, A., Papkoff, J., *et al.*, 1991. A new nomenclature for int-1 and related genes: The Wnt gene family. *Cell* 64 (2), 231.
- Nusse, R., Varmus, H.E., 1982. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* 31 (1), 99–109.
- Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H., Nakatani, Y., 1996. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* 87 (5), 953–959.
- Palaparti, A., Baratz, A., Stifani, S., 1997. The Groucho/transducin-like enhancer of split transcriptional repressors interact with the genetically defined amino-terminal silencing domain of histone H3. *Journal of Biological Chemistry* 272 (42), 26604–26610.
- Pan, W., Choi, S.C., Wang, H., *et al.*, 2008. Wnt3a-mediated formation of phosphatidylinositol 4,5-bisphosphate regulates LRP6 phosphorylation. *Science* 321 (5894), 1350–1353.
- Panakova, D., Sprong, H., Marois, E., Thiele, C., Eaton, S., 2005. Lipoprotein particles are required for Hedgehog and Wingless signalling. *Nature* 435 (7038), 58–65.
- Peradziry, H., Tolwinski, N.S., Borchers, A., 2012. The many roles of PTK7: A versatile regulator of cell–cell communication. *Archives of Biochemistry and Biophysics* 524 (1), 71–76.
- Piao, S., Lee, S.H., Kim, H., *et al.*, 2008. Direct inhibition of GSK3beta by the phosphorylated cytoplasmic domain of LRP6 in Wnt/beta-catenin signaling. *PLoS ONE* 3 (12), e4046.
- Polakis, P., 2012. Wnt signaling in cancer. *Cold Spring Harbor Perspectives in Biology* 4, 5.
- Pourebahram, R., Houtmeyers, R., Ghogomu, S., *et al.*, 2011. Transcription factor Zic2 inhibits Wnt/beta-catenin protein signaling. *Journal of Biological Chemistry* 286 (43), 37732–37740.
- Prokunina-Olsson, L., Welch, C., Hansson, O., *et al.*, 2009. Tissue-specific alternative splicing of TCF7L2. *Human Molecular Genetics* 18 (20), 3795–3804.
- Regard, J.B., Zhong, Z., Williams, B.O., Yang, Y., 2012. Wnt signaling in bone development and disease: Making stronger bone with Wnts. *Cold Spring Harbor Perspectives in Biology* 4, 12.
- Regimbald-Dumas, Y., He, X., 2011. Wnt signalling: What The X@# is WTX? *EMBO Journal* 30 (8), 1415–1417.
- Rios-Esteves, J., Resh, M.D., 2013. Stearoyl CoA desaturase is required to produce active, lipid-modified Wnt proteins. *Cell Reports* 4 (6), 1072–1081.
- Rosin-Arbesfeld, R., Townsley, F., Bienz, M., 2000. The APC tumour suppressor has a nuclear export function. *Nature* 406 (6799), 1009–1012.
- Schuijers, J., Junker, J.P., Mokry, M., *et al.*, 2015. Ascl2 acts as an R-spondin/Wnt-responsive switch to control stemness in intestinal crypts. *Cell Stem Cell* 16 (2), 158–170.
- Schwarz-Romond, T., Fiedler, M., Shibata, N., *et al.*, 2007. The DIX domain of Dishevelled confers Wnt signaling by dynamic polymerization. *Nature Structural & Molecular Biology* 14 (6), 484–492.
- Seeling, J.M., Miller, J.R., Gil, R., *et al.*, 1999. Regulation of beta-catenin signaling by the B56 subunit of protein phosphatase 2A. *Science* 283 (5410), 2089–2091.

- Sierra, J., Yoshida, T., Joazeiro, C.A., Jones, K.A., 2006. The APC tumor suppressor counteracts beta-catenin activation and H3K4 methylation at Wnt target genes. *Genes & Development* 20 (5), 586–600.
- Singh, J., Mlodzik, M., 2012. Planar cell polarity signaling: Coordination of cellular orientation across tissues. *Wiley Interdisciplinary Reviews. Developmental Biology Society for Developmental Biology* 1 (4), 479–499.
- Soliman, M.A., Riabowol, K., 2007. After a decade of study-ING, a PHD for a versatile family of proteins. *Trends in Biochemical Science* 32 (11), 509–519.
- Spink, K.E., Polakis, P., Weis, W.I., 2000. Structural basis of the Axin-adenomatous polyposis coli interaction. *EMBO Journal* 19 (10), 2270–2279.
- Stamos, J.L., Chu, M.L., Enos, M.D., Shah, N., Weis, W.I., 2014. Structural basis of GSK-3 inhibition by N-terminal phosphorylation and by the Wnt receptor LRP6. *eLife* 3, e01998.
- Stamos, J.L., Weis, W.I., 2013. The beta-catenin destruction complex. *Cold Spring Harbor Perspectives in Biology* 5 (1), a007898.
- Taelman, V.F., Dobrowolski, R., Plouhinec, J.L., et al., 2010. Wnt signaling requires sequestration of glycogen synthase kinase 3 inside multivesicular endosomes. *Cell* 143 (7), 1136–1148.
- Tago, K., Nakamura, T., Nishita, M., et al., 2000. Inhibition of Wnt signaling by ICAT, a novel beta-catenin-interacting protein. *Genes & Development* 14 (14), 1741–1749.
- Takada, R., Satomi, Y., Kurata, T., et al., 2006. Monounsaturated fatty acid modification of Wnt protein: Its role in Wnt secretion. *Developmental Cell* 11 (6), 791–801.
- Takemaru, K.I., Moon, R.T., 2000. The transcriptional coactivator CBP interacts with beta-catenin to activate gene expression. *Journal of Cell Biology* 149 (2), 249–254.
- Tamai, K., Semenov, M., Kato, Y., et al., 2000. LDL-receptor-related proteins in Wnt signal transduction. *Nature* 407 (6803), 530–535.
- Tamai, K., Zeng, X., Liu, C., et al., 2004. A mechanism for Wnt coreceptor activation. *Molecular Cell* 13 (1), 149–156.
- Tanneberger, K., Pfister, A.S., Brauburger, K., et al., 2011. Amer1/WTX couples Wnt-induced formation of PtdIns(4,5)P2 to LRP6 phosphorylation. *EMBO Journal* 30 (8), 1433–1443.
- Tauriello, D.V., Jordens, I., Kirchner, K., et al., 2012. Wnt/beta-catenin signaling requires interaction of the dishevelled DEP domain and C terminus with a discontinuous motif in Frizzled. *Proceedings of the National Academy of Sciences of the United States of America* 109 (14), E812–E820.
- Tolwinski, N.S., Wieschaus, E., 2001. Armadillo nuclear import is regulated by cytoplasmic anchor Axin and nuclear anchor dTCF/Pan. *Development* 128 (11), 2107–2117.
- Townsend, F.M., Cliffe, A., Bienz, M., 2004. Pygopus and Legless target Armadillo/beta-catenin to the nucleus to enable its transcriptional co-activator function. *Nature Cell Biology* 6 (7), 626–633.
- Tsao, C.M., Yan, M.D., Shih, Y.L., et al., 2012. SOX1 functions as a tumor suppressor by antagonizing the WNT/beta-catenin signaling pathway in hepatocellular carcinoma. *Hepatology* 56 (6), 2277–2287.
- Valenta, T., Hausmann, G., Basler, K., 2012. The many faces and functions of beta-catenin. *EMBO Journal* 31 (12), 2714–2736.
- Van de Wetering, M., Castrop, J., Korinek, V., Clevers, H., 1996. Extensive alternative splicing and dual promoter usage generate Tcf-1 protein isoforms with differential transcription control properties. *Molecular and Cellular Biology* 16 (3), 745–752.
- Walker, M.P., Stopford, C.M., Cederlund, M., et al., 2015. FOXP1 potentiates Wnt/beta-catenin signaling in diffuse large B cell lymphoma. *Science Signaling* 8 (362), ra12.
- Wallingford, J.B., 2012. Planar cell polarity and the developmental control of cell behavior in vertebrate embryos. *Annual Review of Cell and Developmental Biology* 28, 627–653.
- Wang, Y., Rattner, A., Zhou, Y., et al., 2012. Norrin/Frizzled4 signaling in retinal vascular development and blood brain barrier plasticity. *Cell* 151 (6), 1332–1344.
- Willert, K., Logan, C.Y., Arora, A., Fish, M., Nusse, R., 1999. A *Drosophila* Axin homolog, Daxin, inhibits Wnt signaling. *Development* 126 (18), 4165–4173.
- Willert, K., Brown, J.D., Danenberg, E., et al., 2003. Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature* 423 (6938), 448–452.
- Willert, K., Nusse, R., 2012. Wnt proteins. *Cold Spring Harbor Perspectives in Biology* 4 (9), a007864.
- Willert, K., Shibamoto, S., Nusse, R., 1999. Wnt-induced dephosphorylation of axin releases beta-catenin from the axin complex. *Genes & Development* 13 (14), 1768–1773.
- Wong, H.C., Bourdelas, A., Krauss, A., et al., 2003. Direct binding of the PDZ domain of dishevelled to a conserved internal sequence in the C-terminal region of Frizzled. *Molecular Cell* 12 (5), 1251–1260.
- Wu, G., Huang, H., Garcia Abreu, J., He, X., 2009. Inhibition of GSK3 phosphorylation of beta-catenin via phosphorylated PPPSPXS motifs of Wnt coreceptor LRP6. *PLoS ONE* 4 (3), e4926.
- Wu, X., Tu, X., Joeng, K.S., et al., 2008. Rac1 activation controls nuclear localization of β -catenin during canonical Wnt signaling. *Cell* 133 (2), 340–353.
- Xing, Y., Clements, W.K., Kimelman, D., Xu, W., 2003. Crystal structure of a beta-catenin/axin complex suggests a mechanism for the beta-catenin destruction complex. *Genes & Development* 17 (22), 2753–2764.
- Yamamoto, A., Nagano, T., Takehara, S., Hibi, M., Aizawa, S., 2005. Shisa promotes head formation through the inhibition of receptor protein maturation for the caudalizing factors, Wnt and FGF. *Cell* 120 (2), 223–235.
- Yan, D., Lin, X., 2009. Shaping morphogen gradients by proteoglycans. *Cold Spring Harbor Perspectives in Biology* 1 (3), a002493.
- Ye, F., Chen, Y., Hoang, T., et al., 2009. HDAC1 and HDAC2 regulate oligodendrocyte differentiation by disrupting the beta-catenin-TCF interaction. *Nature Neuroscience* 12 (7), 829–838.
- Yokoya, F., Imamoto, N., Tachibana, T., Yoneda, Y., 1999. beta-catenin can be transported into the nucleus in a Ran-unassisted manner. *Molecular Biology of the Cell* 10 (4), 1119–1131.
- Yu, Y., Wu, J., Wang, Y., et al., 2012. Kindlin 2 forms a transcriptional complex with beta-catenin and TCF4 to enhance Wnt signalling. *EMBO Reports* 13 (8), 750–758.
- Zeng, X., Tamai, K., Doble, B., et al., 2005. A dual-kinase mechanism for Wnt co-receptor phosphorylation and activation. *Nature* 438 (7069), 873–877.
- Zeng, X., Huang, H., Tamai, K., et al., 2008. Initiation of Wnt signaling: Control of Wnt coreceptor Lrp6 phosphorylation/activation via frizzled, dishevelled and axin functions. *Development* 135 (2), 367–375.
- Zhang, N., Wei, P., Gong, A., et al., 2011a. FoxM1 promotes beta-catenin nuclear localization and controls Wnt target-gene expression and glioma tumorigenesis. *Cancer Cell* 20 (4), 427–442.
- Zhang, X., Abreu, J.G., Yokota, C., et al., 2012. Tiki1 is required for head formation via Wnt cleavage-oxidation and inactivation. *Cell* 149 (7), 1565–1577.
- Zhang, X., Cheong, S.M., Amado, N.G., et al., 2015. Notum is required for neural and head induction via Wnt deacylation, oxidation, and inactivation. *Developmental Cell* 32 (6), 719–730.
- Zhang, Y., Liu, S., Mickanin, C., et al., 2011b. RNF146 is a poly(ADP-ribose)-directed E3 ligase that regulates axin degradation and Wnt signalling. *Nature Cell Biology* 13 (5), 623–629.
- Zhuang, L., Hulin, J.A., Gromova, A., et al., 2014. Barx2 and Pax7 have antagonistic functions in regulation of wnt signaling and satellite cell differentiation. *Stem Cells* 32 (6), 1661–1673.

Relevant Website

<http://web.stanford.edu/group/nusselab/cgi-bin/wnt/>
The Wnt Homepage.

The Hippo Pathway

M Ferraiuolo, S Strano, and G Blandino, Regina Elena National Cancer Institute, Rome, Italy

© 2016 Elsevier Inc. All rights reserved.

Introduction

The formation of complex structures during the development of multicellular organisms is a very intriguing theme that raises many questions. The first answers came from a series of studies made in *Drosophila melanogaster* thanks to the development of genetic mosaics. Four tumor suppressor genes were identified, and a loss-of-function mutation in these genes resulted in overgrowth and reduced cell death (Tapon *et al.*, 2002; Lai *et al.*, 2005; Staley and Irvine, 2012). Products of these genes are the core component (the core kinase cassette) of the Hippo pathway (also known as the Salvador–Warts–Hippo pathway), and orthologs were discovered in mammals as well. The Hippo pathway is an evolutionarily conserved regulator of cell proliferation, organ size and shape during development, cell survival, cell competition, stem cell maintenance, metastasis, tissue regeneration, apoptosis, senescence, and differentiation (Figure 1). It is regulated by many factors such as planar cell polarity (PCP), apical-basal cell polarity (ABCP), cell–cell adhesion, contact inhibition, metabolism, mechano-transduction, senescence, and DNA damage (Fausti *et al.*, 2013; Harvey *et al.*, 2013; Strano *et al.*, 2013; Piccolo *et al.*, 2014; Sorrentino *et al.*, 2014; Figure 1). This pathway also cross-talks with other signaling players such as Notch, Wnt, Insulin-like growth factor/Epidermal growth factor receptor (IGF/EGF-R), transforming growth factor- β /bone morphogenetic protein (TGF- β /BMP), and Sonic hedgehog (Cho *et al.*, 2006; Baena-Lopez *et al.*, 2008; Karpowicz *et al.*, 2010). Dysregulation of the Hippo pathway is involved in tumorigenesis (Zhao *et al.*, 2008; Pan, 2010; Harvey *et al.*, 2013; Park and Guan, 2013).

The Hippo Pathway in *D. melanogaster*

The Hippo kinase cassette is composed of four tumor suppressor proteins: a member of the Sterile-20 family Serine-Threonine kinase, Hippo (Hpo) (Harvey *et al.*, 2003); a member of the Nuclear Dbf-2-related (NDR) family Serine-Threonine kinase, Warts (Wts) (Xu *et al.*, 1995); the scaffold protein interacting with Hpo, Salvador (Sav; also known as Shar-Pei) (Tapon *et al.*, 2002); and the Wts cofactor Mob as

tumor suppressor (Mats) (Lai *et al.*, 2005). Activation of Hpo and Wts involves intermolecular autophosphorylation. Activated Hpo phosphorylates Wts, Mats, and Sav bind and link together Hpo and Wts. Phosphorylated Sav activates Hpo kinase activity (Tapon *et al.*, 2002). Wts activity is crucial for the phosphorylation-dependent regulation of the transcriptional coactivator Yorkie (Yki) (Huang *et al.*, 2005; Figure 2). Yki promotes cell proliferation and survival and inhibits apoptosis (Buttitta and Edgar, 2007). Activated Wts/Mats phosphorylate Yki in three different sites (Ser111, Ser168, and Ser250), inhibiting its function (Huang *et al.*, 2005; Dong *et al.*, 2007). The phosphorylation of Ser168 creates a binding site for 14-3-3 proteins (these proteins recognize phosphorylated Serine or Threonine residues) that retain Yki in the cytoplasm, thus inhibiting its nuclear transcriptional coactivator function (Figure 2). Yki has two WW domains that enable interaction with short proline-rich or proline-containing motifs terminated with Tyrosine (PY motifs, i.e., PPXY) (Sudol and Hunter, 2000). Yki lacks its own DNA-binding domain and therefore interacts in the nucleus with different DNA-binding transcription factors such as Scalloped (Sd), Mothers against Dpp (Mad) (Oh and Irvine, 2011), Homothorax (Hth)-Teashirt (Tsh) complex (Peng *et al.*, 2009), and the interacting protein WW domain-binding protein 2 (Wbp2) (Oh and Irvine, 2010; Figure 2). This leads to increased cell proliferation and tissue growth (Staley and Irvine, 2012). In fact, some downstream targets of Yki are the *D. melanogaster* inhibitor of apoptosis protein 1 (Diap1), the growth-promoting microRNA gene *bantam* (*ban*) (Nolo *et al.*, 2006; Peng *et al.*, 2009; Oh and Irvine, 2011), dMyc, E2F1, and cyclins A, B, and E (Tapon *et al.*, 2002; Goulev *et al.*, 2008).

Upstream Regulators of Hippo Signaling in *D. melanogaster*

There are many upstream regulators of the Hpo kinase cassette. This strongly suggests that Hippo signaling functions as a network rather than a linear pathway. The two FERM (Four-point-one, Ezrin, Radixin, Moesin) domain-containing proteins Merlin (Mer) and Expanded (Ex) are tumor suppressors

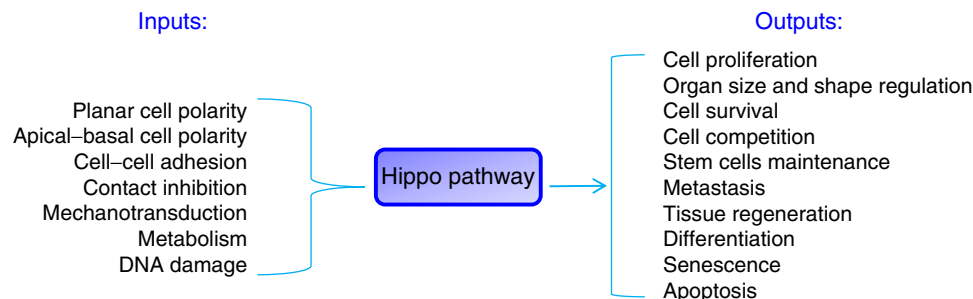


Figure 1 The Hippo pathway is regulated by many factors and is involved in the regulation of several processes.

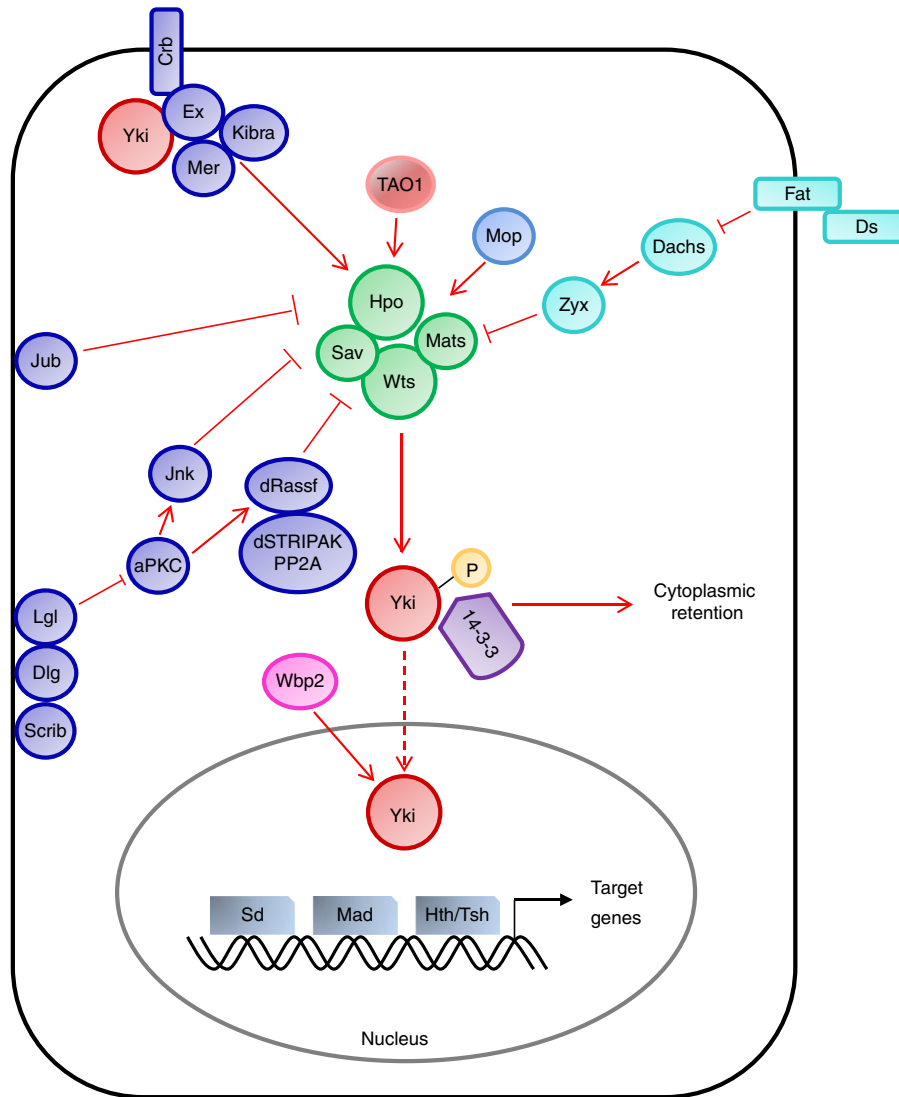


Figure 2 The core kinase cassette of the Hippo pathway in *D. melanogaster* is composed of Hpo, Wts, Sav, and Mats (in green color) and is regulated by many proteins, the majority of which are involved in cell polarity instauration (in blue and cyan colors). The Yki protein (in red color) is a transcriptional coactivator that promotes cell proliferation and survival; it is inhibited by the core kinase cassette.

whose mutation leads to Yki activation and causes tissue overgrowth (Hamaratoglu *et al.*, 2006). Mer and Ex bind to each other and co-localize at the apical membrane of polarized epithelial cells. A third partner in this complex is Kibra (Baumgartner *et al.*, 2010), another tumor suppressor protein (Yu and Guan, 2013) that, together with Mer and Ex, can phosphorylate and activate Hpo and Wts (Hamaratoglu *et al.*, 2006; Figure 2). Furthermore, Ex directly binds Yki and sequesters it in the cytoplasm, preventing its nuclear transcriptional function (Badouel *et al.*, 2009). Ex localization and function are regulated by the transmembrane protein Crumbs (Crb), which favors the acquisition of ABCP (Tepass *et al.*, 1990) and regulation of the Hippo pathway (Figure 2).

The ABCP is also regulated by a complex composed of Lethal giant larvae (Lgl), Disc large (Dlg), and Scribble (Scrib) that localizes to the basolateral membrane of polarized epithelial cells. Lgl inhibits the atypical protein kinase C (aPKC),

which activates the Ras association family member (dRassf) and the Jun Kinase (Jnk); both dRassf and Jnk inhibit activation of the core kinase cassette, increasing Yki activity. Rassf functions as a scaffold protein interacting with the Protein phosphatase 2A (PP2A) serine-threonine phosphatase complex, dSTRIPAK (Striatin-interacting phosphatase and kinase), which inhibits Hpo autophosphorylation and activation (Ribeiro *et al.*, 2010; Figure 2).

The atypical cadherins Fat and Dachsous (Ds) are involved in the acquisition of PCP and function as tumor suppressors, antagonizing the unconventional myosin Dachs, which activates the LIM-domain protein Zyx120 (Zyx). Zyx inhibits the activation of Wts in the core kinase cassette (Cho and Irvine, 2004; Rauskolb *et al.*, 2011); thus Fat, and Ds promote the activation of the Hippo pathway inhibiting Yki (Figure 2).

Other regulators of the Hippo pathway are the Sterile-20-like kinase TAO1, which phosphorylates and activates Hpo

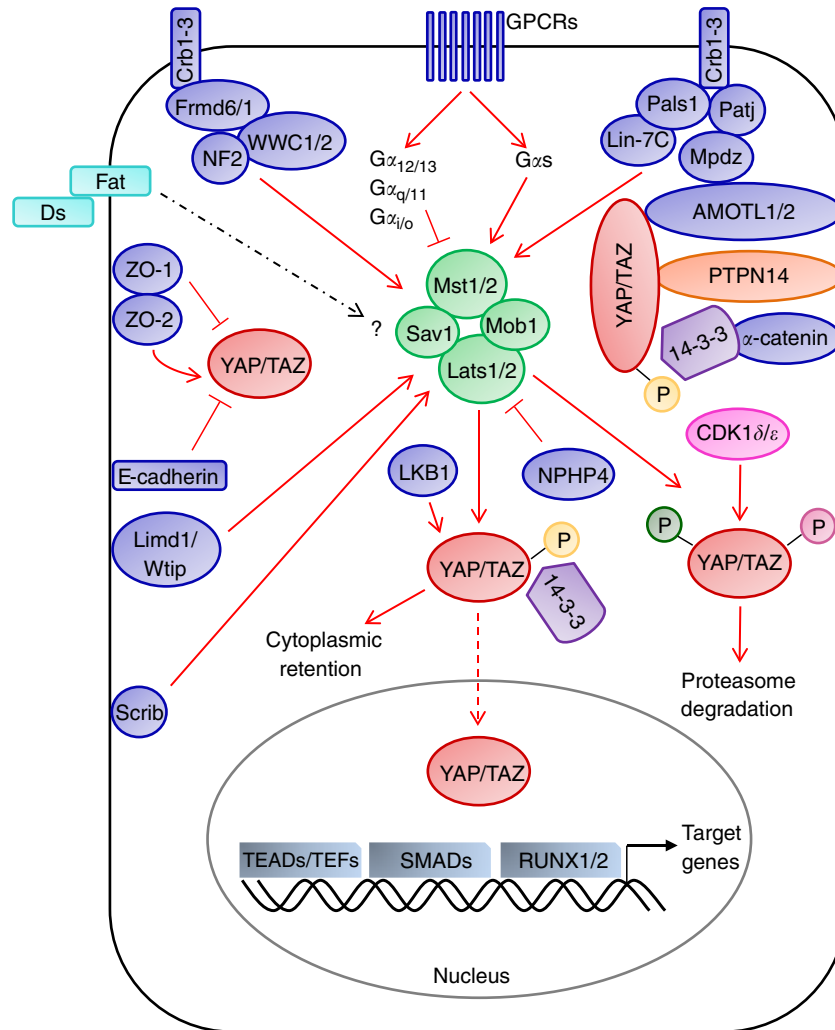


Figure 3 The Hippo pathway in mammals shown conservation in the following key components: Mst1/2, Lats1/2, Sav1, Mob1, and YAP/TAZ (Hpo, Wts, Sav, Mats, and Yki orthologs, respectively). As in *D. melanogaster*, the core kinase cassette (in green color) inhibits YAP/TAZ (in red color) transcriptional coactivators. Fat and Ds roles in the Hippo pathway activation are not yet understood.

(Poon *et al.*, 2011) and the component of the adherent junction Ajuba LIM protein (Jub), which binds and inhibits both Sav and Wts, thereby increasing Yki activity (Das Thakur *et al.*, 2010; Figure 2).

The Hippo Pathway in Mammals

Many components of the *D. melanogaster* Hippo pathway have ortholog human counterparts that exhibit the same or opposite functions. The Hippo kinase cassette is composed by the Mammalian Sterile-20 family Serine-Threonine kinases 1 and 2 (Mst1/2) (Hpo orthologs) that phosphorylate the adaptor protein Salvador 1 (Sav1) (Salvador ortholog) (Tapon *et al.*, 2002), also termed WW45, the large tumor suppressor 1 and 2 (Lats1/2) AGC family kinases (Wts orthologs) (Yabuta *et al.*, 2000) and the Mps one binder kinase activator-like A and B (MOBK1A and MOBK1B) (Mats orthologs), also known as Mob1. Mst1/2 activate Lats1/2 by phosphorylation; the

interaction between Lats1/2 and Mob1 is necessary for Lats1/2 activation. This interaction is enhanced by Mst1/2 phosphorylation of Mob1. Mst1/2 kinase activity is enhanced through interaction with Sav1 (Yabuta *et al.*, 2000). Similar to the core cassette of *D. melanogaster*, Mst1/2 and Lats1/2 trigger a kinase cascade regulated by Sav1 and Mob1 and phosphorylate the co-transcription activators yes-associated protein (YAP) and its paralog Tafazzin (TAZ, also known as WWTR1) (Yki orthologs) (Sudol, 1994; Dong *et al.*, 2007; Figure 3). In response to cell density, YAP/TAZ phosphorylation at Ser127 and Ser89, respectively (like Yki Ser168), creates a 14-3-3 binding site and consequent YAP/TAZ cytoplasmic retention and inhibition. Thus, cell density-dependent activation of the Hippo pathway is involved in cell contact inhibition (Zhao *et al.*, 2012). Moreover, Lats1/2 can phosphorylate YAP at Ser381, triggering a consequent phosphorylation by the casein kinase 1 delta/epsilon (CK1δ/ε) at an additional residue. This event generates a 'phospho-degron' signal (conserved also in TAZ which is phosphorylated at Ser311, but not in Yki)

inducing recruitment of the E3 ubiquitin ligase SCF ^{β -TRCP} and consequentially YAP ubiquitin-mediated proteasome degradation (Zhao *et al.*, 2012; Figure 3).

YAP is a 65-KDa protein existing in two splicing variants (YAP1 and YAP2) and displaying a different number of WW domains (one WW domain in YAP1 and two tandem WW domains in YAP2), one binding domain for the interaction with transcription factors, highly conserved PDZ-binding motifs, and an activation domain like its paralog TAZ (Sudol *et al.*, 1994, 1995; Yagi *et al.*, 1999). YAP is ubiquitously expressed in many tissues as well as from blastocyst to perinatal developmental stages. In the nucleus, YAP interacts through the WW domain with the PPXY motifs of diverse transcription factors such as Runt-related transcription factor-1 (RUNX1, also known as acute myeloid leukemia 1 protein or AML1) and Runt-related transcription factor 2 (RUNX2), which regulates osteoblast differentiation of mesenchymal stem cells (MSCs), chondrocyte hypertrophy, and endothelial cell migration and vascular invasion in bone development (Yagi *et al.*, 1999). YAP also interacts with the TEA domain proteins/transcription enhancer factors (TEADs/TEFs) family (Zhao *et al.*, 2008; Figure 3), regulating cell proliferation, anchorage-independent growth, epithelial-mesenchymal transition (EMT), and embryonic stem cell (ESC) pluripotency. YAP induces the transcription of connective tissue growth factor (CTGF), cysteine-rich angiogenic inducer 61 (CYR61), or CCN family member 1 (CCN1), of Baculoviral Inhibitor of the apoptosis repeat-containing protein 1 (Birc2/cIAP1) Birc5 (also known as surviving), the epidermal growth factor family member amphiregulin (AREG), and Myc (Dong *et al.*, 2007; Lian *et al.*, 2010). YAP's interaction with Smad proteins regulates ESC and iPSC (induced pluripotent stem cells) pluripotency; YAP enhances ErbB-4 transcriptional activity (Varelas *et al.*, 2010). YAP also interacts with the p53 family member p73 (Strano *et al.*, 2001). In response to DNA damage, YAP imparts transcriptional target specificity to p73, favoring its recruitment onto apoptotic, instead of growth suppression targets. The tumor suppressor promyelocytic leukemia protein (PML) gene is a target of the transcriptional competent complex p73/YAP; this leads to an auto-regulatory feedback loop that stabilizes YAP and maximizes p73's pro-apoptotic activity (Strano *et al.*, 2001, 2005; Levy *et al.*, 2007; Lapi *et al.*, 2008). YAP also interacts with β -catenin, inducing the expression of canonical Wnt target genes (SOXS2 and SNAI2) (Heallen *et al.*, 2011). All these studies underline the pleiotropic YAP functions as oncoprotein or transcriptional coactivators of tumor suppressor proteins in response to different stimuli, resulting in differential spatial/temporal regulation of YAP activity and its downstream targets.

Similar to YAP, TAZ is inhibited by the Hippo pathway and functions as a transcriptional coactivator (Figure 3); it interacts with RUNX, TEADs/TEFs, peroxisome proliferator-activated receptor γ (PPAR- γ), thyroid transcription factor-1 (TTF-1), paired box 3 (Pax3), T-box transcription factor 5 (Tbx5), and Smads transcription factors, inducing cell proliferation, cell migration and invasion, ESCs pluripotency, and EMT (Hong *et al.*, 2005).

Upstream Regulators of Hippo Signaling in Mammals

Many regulators of the *D. melanogaster* Hippo pathway are also present in mammals. The neurofibromin 2 (NF2)/FERM

domain-containing protein 6 and 1 (Fmrd6/1)/WW domain-containing protein 1 and 2 (WWC1/2) tumor suppressor complex (Mer/Ex/Kibra orthologs) is involved in the mammalian ABCP; its localization at the apical membrane of polarized epithelial cells is mediated by Crumbs1-3 (Crb1-3) (Crb orthologs). This complex is involved in Lats1/2 activation and consequent YAP/TAZ phosphorylation and inhibition (Zhao *et al.*, 2012; Figure 3). Moreover, depletion of extracellular calcium or knockdown of Crb3 results in YAP/TAZ nuclear localization and transcription coactivation of their target genes (Varelas *et al.*, 2010). Mammalian Scrib activates the Hippo signaling inhibiting YAP/TAZ function (Chen *et al.*, 2012; Figure 3). Cell junction proteins such as Lin-7 homolog C (Lin-7C), membrane-associated palmitoylated protein 5 (MMP5, also known as Pals1), Pals1-associated TJ protein (Patj), and multiple PDZ domain protein (Mpdz) interact with the Hippo core kinase cassette (Varelas *et al.*, 2010; Figure 3). The Angiomotin-like proteins 1 and 2 (AMOTL1/2) maintain epithelial cell polarity and interact with YAP/TAZ in a phosphorylation-independent manner. This leads to sequestering of YAP/TAZ at the apical membrane and inhibition of their nuclear localization (Zhao *et al.*, 2011; Figure 3). AMOT also acts as a scaffold protein interacting with Mst1/2 and Last1/2, promoting YAP/TAZ phosphorylation (Zhao *et al.*, 2011). The α -catenin protein, a component of cell junctions, binds YAP-phosphorylated per 14-3-3 protein interaction, and this trimeric complex sequesters YAP at the apical membrane, preventing its activation (Schlegelmilch *et al.*, 2011; Figure 3). The Tyrosine-protein phosphatase non-receptor type 14 (PTPN14) is another factor involved in cell-cell junctions that directly binds YAP and sequesters it (Liu-Chittenden *et al.*, 2012; Figure 3). Several other proteins involved in ABCP integrity modulate the Hippo pathway: E-cadherin inactivates YAP; Ajuba/LIM domains containing 1 (Limd1)/Wilms tumor protein 1-interacting protein (Wtip) complex (Jub orthologs) activate Sav1 and Lats1/2, inhibiting YAP; Liver kinase B1 (LKB1) induces YAP phosphorylation; Nephronophthisis 4 (NPHP4) protein inhibits Lats1; and zona occludens-2 (ZO-2) induces YAP activation and nuclear localization, whereas ZO-1 represses TAZ activation (Yu and Guan, 2013; Figure 3).

The Hippo pathway responds to mechanical stretch or compression sensing the tensile state of the actomyosin cytoskeleton (dynamic of actin/myosin filaments assembly): in mechanical stretch conditions YAP/TAZ activity is increased and is repressed when cells are compressed (Dupont *et al.*, 2011; Piccolo *et al.*, 2014). YAP/TAZ activation is also regulated by cell-cell contact abundance: high cell density induces YAP/TAZ phosphorylation and cytoplasmic retention; disruption of cell-cell junctions results in YAP/TAZ activation and nuclear localization (Varelas *et al.*, 2010).

Lysophosphatidic acid (LPA), Sphingosine-1-phosphate (S1P), and thrombin inhibit Lats1/2, inducing YAP/TAZ activation in cultured cell binding transmembrane G protein-coupled receptors (GPCRs), and consequently activating the Rho GTPases Rho, Rac, and Cdc42 (Mo *et al.*, 2012; Yu *et al.*, 2012; Zhao *et al.*, 2012). It has been shown that GPCRs activate different heterotrimeric G proteins: $G_{\alpha_{12/13}}$, $G_{\alpha_{q/11}}$, and $G_{\alpha_{i/o}}$ induce YAP/TAZ activation, whereas G_{α_s} represses YAP/TAZ activity (Figure 3); thus glucagon, adrenaline, and dopamine induce Lats1/2-mediated phosphorylation of YAP/TAZ

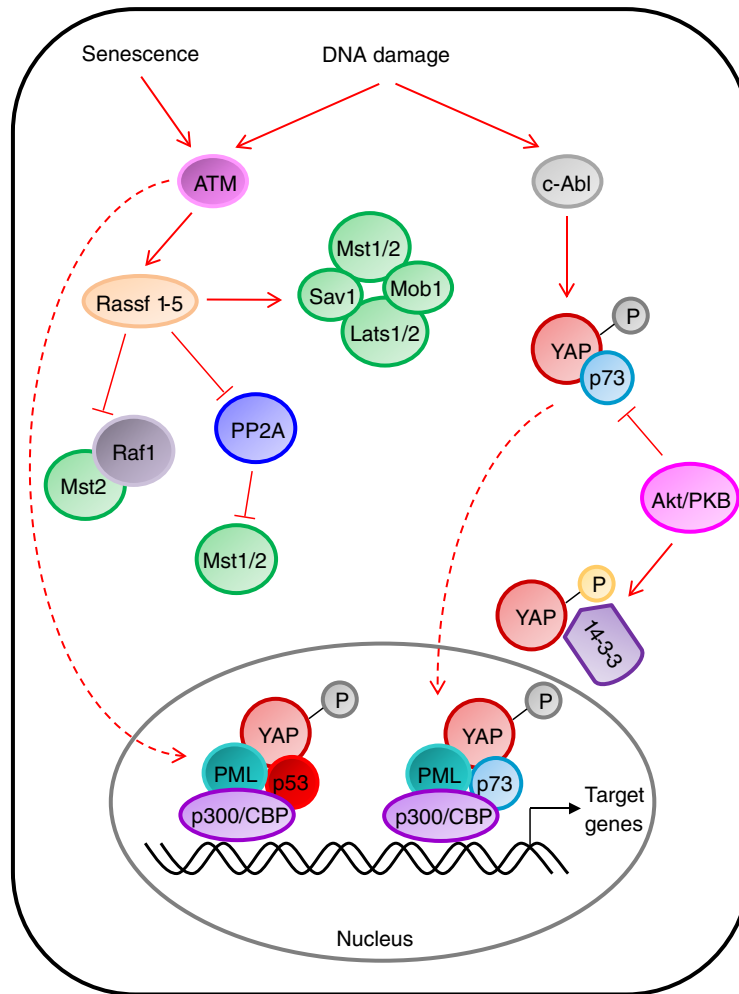


Figure 4 In mammals, YAP is phosphorylated and activated by c-Abl in response to DNA damage. The YAP/PML/p73 complex translocates into the nucleus, promoting pro-apoptotic gene expression. ATM-phosphorylated YAP interacts with PML and signals p53 to induce senescence.

through increasing cyclic adenosine monophosphate (cAMP) amounts, resulting in Lats1/2 activation via Protein kinase A (PKA). Therefore, YAP/TAZ can be up- or down-regulated depending on different G α -activated proteins (Yu *et al.*, 2012). YAP/TAZ activities are also regulated by metabolism; in particular, activation of the mevalonate pathway, involved in cholesterol synthesis, causes YAP/TAZ activation and nuclear localization. Statins (drugs that reduce cholesterol levels in plasma) impair the mevalonate pathway and cause YAP/TAZ inactivation and cytoplasmic retention (Sorrentino *et al.*, 2014).

In response to stress-induced apoptosis, Mst1/2 are activated by autophosphorylation and caspase-dependent cleavage. This induces YAP phosphorylation and inactivation via a Lats1/2-independent mechanism (Zhou *et al.*, 2009). In response to DNA damage, Ataxia telangiectasia mutated (ATM) protein kinase activates the Rassf1-5 mammalian proteins, which have opposite functions compared to dRassf in *D. melanogaster*; they act as tumor suppressors that bind and activate Mst1/2, contrasting with Mst1/2 dephosphorylation by PP2A, and can also disrupt the inhibitory complex between

Raf1 and Mst2, promoting Mst1/2 and Lats1/2 interaction (Figure 4). In this way, YAP is phosphorylated at Tyr357 and relocalizes into the nucleus, where it binds to p73 and promotes its transcriptional pro-apoptotic activity (Matallanas *et al.*, 2007; Figure 4). By contrast, Rassf6 binds and inhibits Mst2, blocking Hippo signaling (Ikeda *et al.*, 2009). Moreover, in response to DNA damage, c-Abl kinase directly phosphorylates YAP at Tyr357, increasing protein stability and affinity for p73 (Levy *et al.*, 2008), whereas Akt/PKB kinase signaling mediates YAP phosphorylation at Ser127 and consequent cytoplasmic retention by preventing its binding to p73 and transcriptional activation of pro-apoptotic genes (Figure 4). YAP is also involved in the establishment of senescence through the activation of PML and p53 tumor suppressor proteins; in detail, loss of Werner-induced senescence activates ATM, which phosphorylates YAP and triggers the instauration of a transcriptionally active YAP/PML/p53 axis (Fausti *et al.*, 2013; Strano *et al.*, 2013; Figure 4). Together these findings highlight the fact that several intrinsic and extrinsic signals regulate the Hippo mammalian pathway.

Regulation of Organ Size, Stem Cell Self-Renewal, and Tissue Regeneration

In *D. melanogaster*, loss-of-function mutations of Hpo or Wts kinases or overexpression of Yki during development result in overgrown imaginal discs, leading to overgrowth of the corresponding adult structures (Staley and Irvine, 2012).

Organ size is finely regulated during development, and the Hippo signaling pathway plays a key role in this process (Camargo *et al.*, 2007). YAP/TAZ are expressed during mammalian early embryogenesis, binding the transcription factor TEAD4 and inducing Caudal type homeobox 2 (Cdx2) expression. Cdx2 is a transcription factor involved in trophoblast (TE; outer layer of the blastocyst stage embryo) formation (Camargo *et al.*, 2007). Loss of YAP suppresses Cdx2 expression, indicating that Hippo signaling is crucial for inner cell mass (ICM) segregation from the TE in blastocyst development. Moreover, YAP/TAZ localization is spatially regulated in the blastocyst stage embryo: YAP/TAZ are nuclear in TE and cytoplasmic in ICM, and this differential distribution is determinant for lineage specification in the preimplantation embryo (Camargo *et al.*, 2007). YAP/TAZ double knockout embryos die at the morula stage, before TE formation. It appears that the actin cytoskeleton structure is determinant for YAP/TAZ localization: TE cells are polarized, and thus have more F-actin and tension than ICM cells, which by contrast display more cell density and cell-cell contacts. Moreover, TE cells are exposed, and the GPCRs expressed on the surface might be regulated by external diffusible signals (Yu and Guan, 2013). Alteration in the expression or activity of upstream regulators results in deregulation of the Hippo pathway, which causes severe developmental defects.

The Hippo pathway defines either differentiation or reprogramming to a pluripotent state that is determined by the precise balance of growth factor amounts and cytoskeletal-associated cues. YAP is found in the nucleus of MSCs and ESCs; its nuclear localization and protein levels decline during MSC and ESC differentiation. Concomitantly, Hippo pathway activity is enhanced and YAP accumulates in the cytoplasm of differentiated cells (Lian *et al.*, 2010). YAP is also nuclear, and its expression is increased in iPSCs in conjunction with Octamer-binding protein 4 (Oct4), SRY (Sex-determining region Y), SRY-box 2 (Sox2), and Kruppel-like factor 4 (Klf4) transcription factor activation (Lian *et al.*, 2010); the pluripotent state of ESCs is also mediated by the YAP/TAZ/Smad2/3 complex in response to TGF- β /BMP signaling. It was observed that YAP/TEAD2 bind and activate OCT3/4 and NANOG promoters, and Lats1/2 inhibition of TAZ impairs reprogramming. Therefore, nuclear YAP/TAZ promote ESC self-renewal and iPSC reprogramming (Lian *et al.*, 2010). Notably, YAP/TAZ localization may be critical for cell fate, despite the fact that the impact of YAP/TAZ cellular compartmentalization during development is not well understood. Increasing nuclear YAP/TAZ promotes tissue regeneration, rising cell proliferation in the liver, pancreas, salivary glands, kidney, lung, heart, intestine, skin, and nervous system (Varelas, 2014). This requires that the Hippo pathway cross-talk with the Notch, Wnt, IGF/EGF-R, TGF- β /BMP, and Sonic hedgehog signaling pathways (Pan, 2010; Yu and Guan, 2013; Varelas, 2014).

Hippo Pathway Deregulation in Cancer

Loss-of-function mutation or epigenetic silencing of core kinase cassette components or activators of the Hippo pathway result in uncontrolled tissue overgrowth; moreover, YAP/TAZ overexpression induces transformation in mammalian cells and promotes tumorigenesis (Pan, 2010; Zhao *et al.*, 2012). Lats1/2 inhibit cell division cycle 2 (Cdc2 or Cdk1), favoring G2/M arrest; Lats2 also induces G1/S arrest. Thus, loss-of-function mutation of these genes leads to formation of tumors such as tissue sarcoma and leukemia (Yabuta *et al.*, 2000). Inactivation of NF2 promotes metastatic osteosarcomas, brain tumor development (meningiomas, schwannomas, and acoustic neuromas), hepatocellular carcinoma (HCC), bile duct tumors, and malignant mesothelioma, as do Mst1/2, Lats2, and Sav1 mutations (Johnson and Halder, 2014). Patients with NF2 mutation exhibit F-actin alterations and morphological changes, resulting in increased YAP activation. Fat4 and Dchs1/2 are frequently mutated in colorectal cancer; mutations in Sav1 and Mob1 are present in renal cancer cell lines and human melanoma, respectively (Johnson and Halder, 2014); Mst1/2, Lats1/2, and Rassf family genes are often silenced through epigenetic mechanisms, and YAP/TAZ genes are amplified in the lung, pancreas, esophagus, skin, colon, prostate, liver, ovarian and mammary gland carcinomas, medulloblastomas, and oral squamous-cell carcinomas; moreover, chromosomal translocation involving TAZ and calmodulin-binding transcription activator 1 (CAMTA1) results in a fusion gene TAZ-CAMTA1 being highly expressed in epithelioid hemangioendothelioma (Pan, 2010; Harvey *et al.*, 2013).

Elevated nuclear levels of YAP/TAZ are present in many solid tumors. This leads to increased cell proliferation, acquisition of cancer stem cell features, EMT (forming metastases), inhibition of senescence, reduced apoptosis, and drug resistance (Harvey *et al.*, 2013). Invasion and metastasis are common in advanced solid tumors and are responsible for the majority of deaths from cancer. Loss of cell polarity promotes YAP/TAZ activation in a positive feed-forward loop, inducing EMT and metastasis formation (Cordenonsi *et al.*, 2011). Increasing cell survival and reduced cell death are mediated by YAP/TAZ transcriptional up-regulation of anti-apoptotic B cell lymphoma 2 (Bcl-2) family members, cIAP1/2 anti-apoptotic proteins, MYC, CTGF, and CRY61 (Pan, 2010). Moreover, YAP overexpression can suppress anoikis cell death induced by anchorage-dependent cells detaching, promoting dissemination of invading cells and inducing metastasis formation; loss of E-cadherin (linked to EMT) function may also induce YAP/TAZ activation in metastatic cells (Yu and Guan, 2013).

Finally, the Hippo pathway interacts with many other signaling pathways, creating a complex network in which YAP/TAZ are emerging as a critical node integrating and decoding both oncogenic and tumor suppressor inputs (Harvey *et al.*, 2013).

Conclusions

The pleiotropic activities of the Hippo pathway have attracted unprecedented scientific interest. This has led to the

uncovering of many yet unexplored mechanistic insights into organ size regulation, stem cell self-renewal, tissue regeneration, DNA damage response, senescence, cell metabolism, and cancer development. It can be reasonably hypothesized that regenerative medicine and cancer treatment will strongly benefit from such knowledge.

See also: Cell Communication: Cellular Outcomes: Regulation of Cell Polarity

References

- Badouel, C., Gardano, L., Amin, N., *et al.*, 2009. The FERM-domain protein expanded regulates Hippo pathway activity via direct interactions with the transcriptional activator Yorkie. *Developmental Cell* 16 (3), 411–420.
- Baena-Lopez, L.A., Rodríguez, I., Baonza, A., *et al.*, 2008. The tumor suppressor genes dachsous and fat modulate different signalling pathways by regulating dally and dally-like. *Proceedings of the National Academy of Sciences of the United States of America* 105 (28), 9645–9650.
- Baumgartner, R., Poernbacher, I., Buser, N., Hafen, E., Stocker, H., 2010. The WW domain protein Kibra acts upstream of Hippo in *Drosophila*. *Developmental Cell* 18 (2), 309–316.
- Buttitta, L.A., Edgar, B.A., 2007. How size is controlled: From Hippos to Yorkies. *Nature Cell Biology* 9 (11), 1225–1227.
- Camargo, F.D., Gokhale, S., Johnnidis, J.B., *et al.*, 2007. YAP1 increases organ size and expands undifferentiated progenitor cells. *Current Biology* 17 (23), 2054–2060.
- Chen, D., Sun, Y., Wei, Y., *et al.*, 2012. LIFR is a breast cancer metastasis suppressor upstream of the Hippo-YAP pathway and a prognostic marker. *Nature Medicine* 18 (10), 1511–1517.
- Cho, E., Feng, Y., Rauskolb, C., *et al.*, 2006. Delineation of a Fat tumor suppressor pathway. *Nature Genetics* 38, 1142–1150.
- Cho, E., Irvine, K.D., 2004. Action of fat, four-jointed, dachsous and dachs in distal-to-proximal wing signaling. *Development* 131 (18), 4489–4500.
- Cordenonsi, M., Zanconato, F., Azzolin, L., *et al.*, 2011. The Hippo transducer TAZ confers cancer stem cell-related traits on breast cancer cells. *Cell* 147 (4), 759–772.
- Das Thakur, M., Feng, Y., Jagannathan, R., *et al.*, 2010. Ajuba LIM proteins are negative regulators of the Hippo signaling pathway. *Current Biology* 20, 657–662.
- Dong, J., Feldmann, G., Huang, J., *et al.*, 2007. Elucidation of a universal size-control mechanism in *Drosophila* and mammals. *Cell* 130 (6), 1120–1133.
- Dupont, S., Morsut, L., Aragona, M., *et al.*, 2011. Role of YAP/TAZ in mechanotransduction. *Nature* 474 (7350), 179–183.
- Fausti, F., Di Agostino, S., Cioce, M., *et al.*, 2013. ATM kinase enables the functional axis of YAP, PML and p53 to ameliorate loss of Werner protein-mediated oncogenic senescence. *Cell Death & Differentiation* 20 (11), 1498–1509.
- Goulev, Y., Fauny, J.D., Gonzalez-Marti, B., *et al.*, 2008. SCALLOPED interacts with YORKIE, the nuclear effector of the hippo tumor-suppressor pathway in *Drosophila*. *Current Biology* 18 (6), 435–441.
- Hamaratoglu, F., Willecke, M., Kango-Singh, M., *et al.*, 2006. The tumour-suppressor genes NF2/Merlin and Expanded act through Hippo signalling to regulate cell proliferation and apoptosis. *Nature Cell Biology* 8 (1), 27–36.
- Harvey, K.F., Pfeleger, C.M., Hariharan, I.K., 2003. The *Drosophila* Mst ortholog, hippo, restricts growth and cell proliferation and promotes apoptosis. *Cell* 114 (4), 457–467.
- Harvey, K.F., Zhang, X., Thomas, D.M., 2013. The Hippo pathway and human cancer. *Nature Reviews Cancer* 13 (4), 246–257.
- Heallen, T., Zhang, M., Wang, J., *et al.*, 2011. Hippo pathway inhibits Wnt signaling to restrain cardiomyocyte proliferation and heart size. *Science* 332 (6028), 458–461.
- Hong, J.H., Hwang, E.S., McManus, M.T., *et al.*, 2005. TAZ, a transcriptional modulator of mesenchymal stem cell differentiation. *Science* 309 (5737), 1074–1078.
- Huang, J., Wu, S., Barrera, J., Matthews, K., Pan, D., 2005. The Hippo signaling pathway coordinately regulates cell proliferation and apoptosis by inactivating Yorkie, the *Drosophila* Homolog of YAP. *Cell* 122 (3), 421–434.
- Ikeda, M., Kawata, A., Nishikawa, M., *et al.*, 2009. Hippo pathway-dependent and -independent roles of RASSF6. *Science Signaling* 2 (90), ra59.
- Johnson, R., Halder, G., 2014. The two faces of Hippo: Targeting the Hippo pathway for regenerative medicine and cancer treatment. *Nature Reviews Drug Discovery* 13 (1), 63–79.
- Karpowicz, P., Perez, J., Perrimon, N., 2010. The Hippo tumor suppressor pathway regulates intestinal stem cell regeneration. *Development* 137 (24), 4135–4145.
- Lai, Z.C., Wei, X., Shimizu, T., *et al.*, 2005. Control of cell proliferation and apoptosis by mob as tumor suppressor, mats. *Cell* 120 (5), 675–685.
- Lapi, E., Di Agostino, S., Donzelli, S., *et al.*, 2008. PML, YAP, and p73 are components of a proapoptotic autoregulatory feedback loop. *Molecular Cell* 32 (6), 803–814.
- Levy, D., Adamovich, Y., Reuven, N., Shaul, Y., 2007. The Yes-associated protein 1 stabilizes p73 by preventing Itch-mediated ubiquitination of p73. *Cell Death & Differentiation* 14 (4), 743–751.
- Levy, D., Adamovich, Y., Reuven, N., Shaul, Y., 2008. Yap1 phosphorylation by c-Abl is a critical step in selective activation of proapoptotic genes in response to DNA damage. *Molecular Cell* 29 (3), 350–361.
- Lian, I., Kim, J., Okazawa, H., 2010. The role of YAP transcription coactivator in regulating stem cell self-renewal and differentiation. *Genes & Development* 24 (11), 1106–1118.
- Liu-Chittenden, Y., Huang, B., Shim, J., *et al.*, 2012. Genetic and pharmacological disruption of the TEAD-YAP complex suppresses the oncogenic activity of YAP. *Genes & Development* 26 (12), 1300–1305.
- Matallanas, D., Romano, D., Yee, K., *et al.*, 2007. RASSF1A elicits apoptosis through an MST2 pathway directing proapoptotic transcription by the p73 tumor suppressor protein. *Molecular Cell* 27 (6), 962–975.
- Mo, J.S., Yu, F.X., Gong, R., Brown, J.H., Guan, K.L., 2012. Regulation of the Hippo-YAP pathway by protease-activated receptors (PARs). *Genes & Development* 26 (19), 2138–2143.
- Nolo, R., Morrison, C.M., Tao, C., *et al.*, 2006. The bantam microRNA is a target of the hippo tumor-suppressor pathway. *Current Biology* 16 (19), 1895–1904.
- Oh, H., Irvine, K.D., 2010. Yorkie: The final destination of Hippo signaling. *Trends in Cell Biology* 20 (7), 410–417.
- Oh, H., Irvine, K.D., 2011. Cooperative regulation of growth by Yorkie and Mad through bantam. *Developmental Cell* 20 (1), 109–122.
- Pan, D., 2010. The hippo signaling pathway in development and cancer. *Developmental Cell* 19 (4), 491–505.
- Park, H.W., Guan, K.L., 2013. Regulation of the Hippo pathway and implications for anticancer drug development. *Trends in Pharmacological Sciences* 34 (10), 581–589.
- Peng, H.W., Slattery, M., Mann, R.S., 2009. Transcription factor choice in the Hippo signaling pathway: Homothorax and yorkie regulation of the microRNA bantam in the progenitor domain of the *Drosophila* eye imaginal disc. *Genes & Development* 23 (19), 2307–2319.
- Piccolo, S., Dupont, S., Cordenonsi, M., 2014. The Biology of YAP/TAZ: Hippo Signaling and Beyond. *Physiological Reviews* 94 (4), 1287–1312.
- Poon, C.L.C., Lin, J.I., Zhang, X., Harvey, K.F., 2011. The sterile 20-like kinase Tao-1 controls tissue growth by regulating the Salvador-Warts-Hippo pathway. *Developmental Cell* 21 (5), 896–906.
- Rauskolb, C., Pan, G., Reddy, B.V., Oh, H., Irvine, K.D., 2011. Zyxin links fat signaling to the hippo pathway. *PLOS Biology* 9 (6), e1000624.
- Ribeiro, P.S., Josué, F., Wepf, A., *et al.*, 2010. Combined functional genomic and proteomic approaches identify a PP2A complex as a negative regulator of Hippo signaling. *Molecular Cell* 39 (4), 521–534.
- Schlegelmilch, K., Mohseni, M., Kirak, O., *et al.*, 2011. Yap1 acts downstream of alpha-catenin to control epidermal proliferation. *Cell* 144 (5), 782–795.
- Sorrentino, G., Ruggeri, N., Specchia, V., *et al.*, 2014. Metabolic control of YAP and TAZ by the mevalonate pathway. *Nature Cell Biology* 16 (4), 357–366.
- Staley, B.K., Irvine, K.D., 2012. Hippo signaling in *Drosophila*: Recent advances and insights. *Developmental Dynamics* 241 (1), 3–15.
- Strano, S., Fausti, F., Di Agostino, S., *et al.*, 2013. PML Surfs into HIPPO Tumor Suppressor Pathway. *Frontiers in Oncology* 3, 36.
- Strano, S., Monti, O., Pediconi, N., *et al.*, 2005. The transcriptional coactivator Yes-associated protein drives p73 gene-target specificity in response to DNA Damage. *Molecular Cell* 18 (4), 447–459.
- Strano, S., Munarriz, E., Rossi, M., *et al.*, 2001. Physical interaction with Yes-associated protein enhances p73 transcriptional activity. *Journal of Biological Chemistry* 276 (18), 15164–15173.
- Sudol, M., 1994. Yes-associated protein (YAP65) is a proline-rich phosphoprotein that binds to the SH3 domain of the Yes proto-oncogene product. *Oncogene* 9 (8), 2145–2152.
- Sudol, M., Bork, P., Einbond, A., *et al.*, 1995. Characterization of the mammalian YAP (Yes-associated protein) gene and its role in defining a novel protein module, the WW domain. *Journal of Biological Chemistry* 270 (24), 14733–14741.

- Sudol, M., Hunter, T., 2000. NeW wrinkles for an old domain. *Cell* 103 (7), 1001–1004.
- Tapon, N., Harvey, K.F., Bell, D.W., *et al.*, 2002. salvador Promotes both cell cycle exit and apoptosis in *Drosophila* and is mutated in human cancer cell lines. *Cell* 110 (4), 467–478.
- Tepass, U., Theres, C., Knust, E., 1990. crumbs encodes an EGF-like protein expressed on apical membranes of *Drosophila* epithelial cells and required for organization of epithelia. *Cell* 61 (5), 787–799.
- Varelas, X., 2014. The Hippo pathway effectors TAZ and YAP in development, homeostasis and disease. *Development* 141 (8), 1614–1626.
- Varelas, X., Samavarchi-Tehrani, P., Narimatsu, M., *et al.*, 2010. The Crumbs complex couples cell density sensing to Hippo-dependent control of the TGF-beta-SMAD pathway. *Developmental Cell* 19 (6), 831–844.
- Xu, T., Wang, W., Zhang, S., Stewart, R.A., Yu, W., 1995. Identifying tumor suppressors in genetic mosaics: The *Drosophila* lats gene encodes a putative protein kinase. *Development* 121 (4), 1053–1063.
- Yabuta, N., Fujii, T., Copeland, N.G., *et al.*, 2000. Structure, expression, and chromosome mapping of LATS2, a mammalian homologue of the *Drosophila* tumor suppressor gene lats/warts. *Genomics* 63 (2), 263–270.
- Yagi, R., Chen, L.F., Shigesada, K., Murakami, Y., Ito, Y., 1999. A WW domain-containing yes-associated protein (YAP) is a novel transcriptional co-activator. *EMBO Journal* 18 (9), 2551–2562.
- Yu, F.X., Guan, K.L., 2013. The Hippo pathway: Regulators and regulations. *Genes & Development* 27 (4), 355–371.
- Yu, F.-X., Zhao, B., Panupinthu, N., *et al.*, 2012. Regulation of the Hippo-YAP pathway by G-protein-coupled receptor signaling. *Cell* 150 (4), 780–791.
- Zhao, B., Li, L., Lu, Q., *et al.*, 2011. Angiomin is a novel Hippo pathway component that inhibits YAP oncoprotein. *Genes & Development* 25 (1), 51–63.
- Zhao, B., Li, L., Wang, L., *et al.*, 2012. Cell detachment activates the Hippo pathway via cytoskeleton reorganization to induce anoikis. *Genes & Development* 26 (1), 54–68.
- Zhao, B., Ye, X., Yu, J., *et al.*, 2008. TEAD mediates YAP-dependent gene induction and growth control. *Genes & Development* 22 (14), 1962–1971.
- Zhou, D., Conrad, C., Xia, F., *et al.*, 2009. Mst1 and Mst2 maintain hepatocyte quiescence and suppress hepatocellular carcinoma development through inactivation of the Yap1 oncogene. *Cancer Cell* 16 (5), 425–438.

Further Reading

Oren, M., Aylon, Y. (Eds.), 2013. *The Hippo Signaling Pathway and Cancer*. New York; Heidelberg; Dordrecht; London: Springer.

The Notch Pathway

MG Hall, MD Anderson Cancer Center, Houston, TX, USA

L Smith, University of Minnesota, Minneapolis, MN, USA

P Zweidler-McKay, MD Anderson Cancer Center, Houston, TX, USA

© 2016 Elsevier Inc. All rights reserved.

The Notch Signaling Pathway

The Notch pathway was first described as a 'notch' seen in a *Drosophila* wing that resulted from haploinsufficiency of the Notch receptor. Subsequently, the Notch receptor, two ligands, Jagged and Serrate, and the downstream target gene *Hairy* were identified. In mammals this pathway has become more complex with four Notch receptors (Notch1, Notch2, Notch3, and Notch4), five ligands, and at least 10 *Hairy*-related genes. However, the basic mechanics of the pathway have been highly conserved across species. Briefly, the Notch receptors are each synthesized in a single polypeptide and then directed to the Golgi where they undergo glycosylation via fringe glycosyltransferases (Manic, Lunatic, and Radical) and proteolytic cleavage via a furin-like convertase. Glycosylation of the Notch receptors increases their affinity for specific Notch ligands. The two protein fragments resulting from proteolytic cleavage form a heterodimer in which the two domains are connected by a non-covalent calcium bond (Guruharsha *et al.*, 2012). Each receptor is thus composed of an extracellular domain and a transmembrane-intracellular domain (Figure 1). The extracellular domain is comprised of up to 36 epidermal growth factor (EGF) repeats, which bind Notch ligands, as well as a negative regulatory region (NRR) which binds to the transmembrane-intracellular domain, thereby preventing receptor activation. The NRR can be subdivided further into three Lin12/Notch repeats (LNR) and a heterodimerization domain (HD). The transmembrane-intracellular domain is non-covalently bound to the extracellular domain. The intracellular domain of Notch (ICN or NICD) is comprised of an RBPJk-associated module (RAM), ankyrin repeats (ANK), transactivating domain (TAD), and a proline, glutamic acid, serine, threonine (PEST) domain. The RAM and ANK domains allow ICN to interact with other proteins while the PEST domain is ubiquitinated, leading to degradation of the activated receptor (Li *et al.*, 2006). The five Notch ligands (Jagged1 (JAG1), JAG2, Delta-like ligand1 (DLL1), DLL3, and DLL4) are highly homologous to the two ligands first identified in *Drosophila*, Serrate and Delta, respectively. These ligands expressed on the cell surface of a signal-sending cell bind to and activate Notch receptors on neighboring cells. The ligands have a series of EGF repeats and a Delta-Serrate LAG-2 (DSL) domain. Additionally, the Jagged ligands also have a cysteine-rich domain. The DSL and EGF repeats cooperate to bind Notch receptors. Differences between the ligand structures result in a variety of different functions. Most of the ligands lead to receptor activation via a conformational change in the heterodimerization domain of the receptor with subsequent cleavage of the transmembrane domain. However, DLL3 can bind the Notch receptors but is unable to activate Notch, thereby inhibiting the level of signaling that occurs when other ligands bind to a receptor (Ladi *et al.*, 2005). The activating ligands also have preferences for different

receptors. This is partly based on the fringe glycosyltransferase modifications on the receptor EGF repeats from processing of the receptors in the Golgi. Glycosylated receptors have higher affinity for the Delta-like ligands compared to the Jagged ligands (Rampal *et al.*, 2005). Upon ligand binding and cleavage of the receptor, the signal-sending cells endocytose the ligand and the extracellular domain of the receptor. The remaining receptor is embedded in the membrane and is the target of enzymatic cleavage by ADAM17 (S2 cleavage) and gamma secretase (S3 cleavage). These sequential proteolytic cleavages from the transmembrane domain release ICN into the cytoplasm.

Upon release of the ICN from the transmembrane domain, the ICN translocates to the nucleus via two nuclear localization signals, and binds to a complex of ubiquitous transcription factor CBF-1, Suppressor of Hairless, Lag-2 (CSL or RBPJk). CSL normally forms a corepressor complex including SMRT, MINT, and SPEN, transcriptionally repressing CSL target genes. However, upon ICN binding, the corepressors are displaced and co-activators (i.e., mastermind-like) and histone acetyltransferases (i.e., CBP/p300) form an activating complex that upregulates target gene expression. The most common Notch/CSL target genes are in the *Hairy/Enhancer of Split* (HES) related family (HES1-HES7, HEY1, HEY2, and HEYL). These genes are basic-helix-loop-helix (bHLH) transcription factors which generally recruit corepressors of the groucho/TLE family to repress their target genes. Many other direct targets of Notch/CSL have been described in many cell type-specific patterns including MYC, Cyclin D1/3, CDK5, p21, Snail, PDGFRb, and many others. This downstream complexity provides for the diverse consequences of Notch signaling in different tissues.

Notch signaling is tightly controlled as low, moderate, and high levels and can have distinct consequences in a single cell type. Consequently, regulation of the Notch pathway occurs at multiple levels. Ligands on the signal-sending cell can be downregulated via ubiquitination by E3 ligases, Mindbomb and Neuralized, and undergo subsequent proteasomal degradation. Similarly, receptors can be degraded before the ICN is able to translocate to the nucleus. E3 ubiquitin ligases such as Deltex1 and Numb ubiquitinate the intracellular domain of the receptors, which leads to proteasomal degradation. Additionally, the Notch target gene NRARP can bind to ICN impairing the recruitment of a co-activation complex, forming another negative feedback loop. Finally, the HES family of Notch target genes can repress their own transcription, leading to oscillations of Notch downstream signaling.

Developmental Biology

Mechanisms

The Notch signaling pathway plays a key role in the developmental biology of all animals. It is involved in the growth

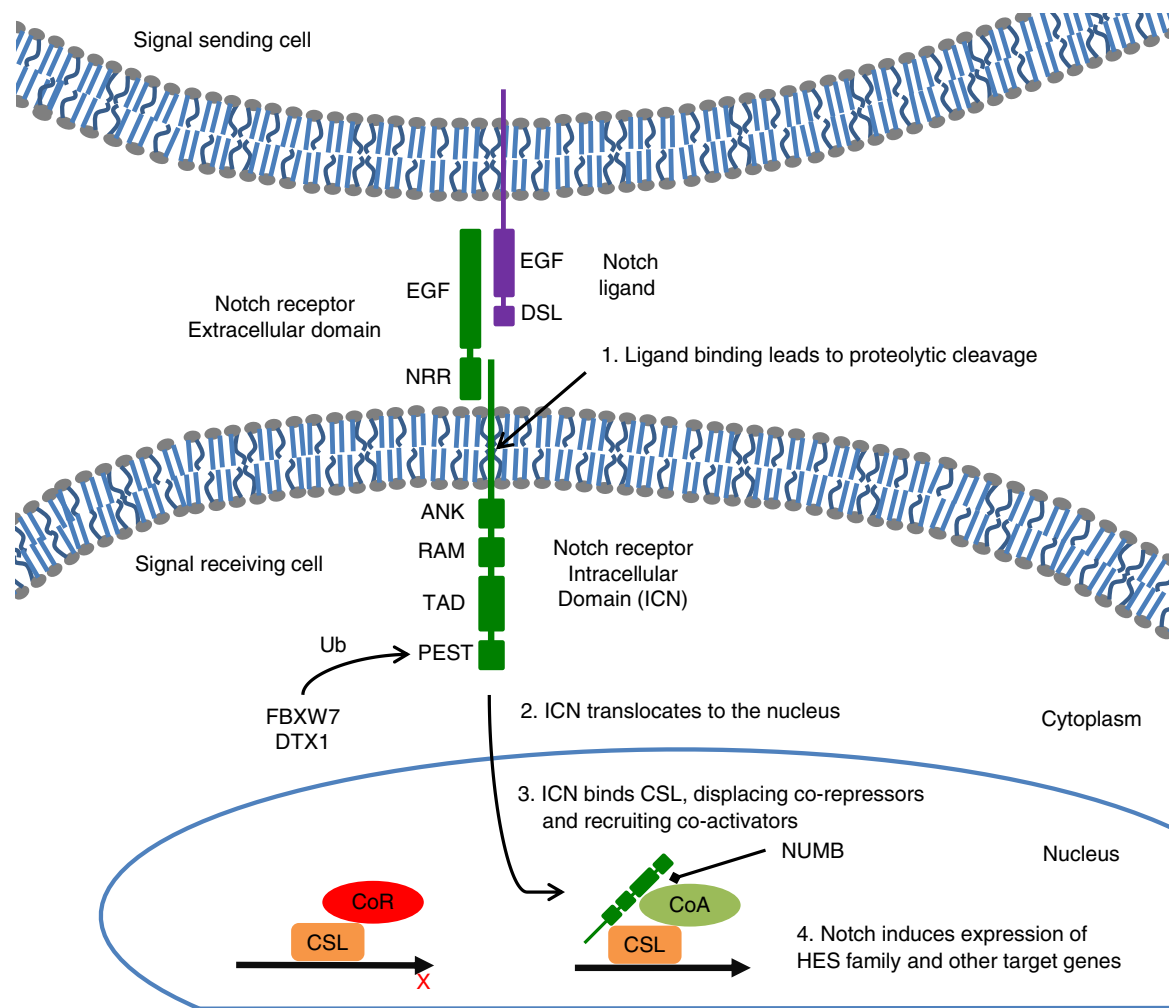


Figure 1 The Notch pathway. Notch ligand expressed on a signal-sending cell engages a Notch receptor on a signal-receiving cell and induces a conformational change in the Notch receptor, thereby allowing proteolytic cleavage of the receptor in the transmembrane domain. This releases the intracellular Notch receptor, which translocates to the nucleus, and binds to the CSL transcription factor, displacing corepressors and recruiting co-activators to drive the expression of HES family genes, and other target genes. See text for domain names.

and differentiation of many cell types, and has the ability to control cell fate decisions. Notch signaling can control the patterns of gene expression within a cell by either upregulating or downregulating various genes. These alterations of gene expression can either induce or impair differentiation. Thus, activating Notch in one cell type may impair differentiation, retaining stem-like properties in that cell type, while its activation may allow for differentiation down an alternate fate in another cell type. The distribution of Notch ligands and receptors can influence the cell fate decisions by three basic mechanisms (Figure 2; Lewis, 1998). During cell division the Notch pathway components can be asymmetrically distributed to two daughter cells (e.g., Numb), leading to different levels of Notch signaling and thus different cell fates. An additional mechanism of Notch influencing cell fate occurs during cell-cell interaction. Competitive Notch signaling can inhibit neighboring cells from adopting the same cell fate through lateral inhibition. This occurs in some cell populations when Notch signaling is induced in neighboring cells by presentation of Notch ligands in the signal-sending cell.

Signal-receiving cells downregulate expression of their own Notch ligand and thus decrease Notch signaling in the signal-sending cell. Ultimately, strong Notch signaling occurs in one cell and a lack of Notch signaling occurs in the other cell, thereby promoting different cell fates in neighboring cells. Another mechanism of how Notch influences cell fate decisions is boundary formation or inductive signaling. When a population of cells express Notch ligands they induce Notch signaling in neighboring cells, and this induces an alternate cell fate. Thus a group of cells can induce the differentiation of another group of cells, forming a boundary between different cell types. Whether these mechanisms occur in a particular cell type depends on expression of Notch pathway members and the downstream consequences of Notch signaling in a given cell type.

Using the cell fate mechanisms above, and its ability to regulate growth and apoptosis in a cell type-specific manner, Notch signaling plays numerous roles during embryogenesis, including critical roles in somitogenesis, neurogenesis, myogenesis, hematopoiesis, angiogenesis, among many others.

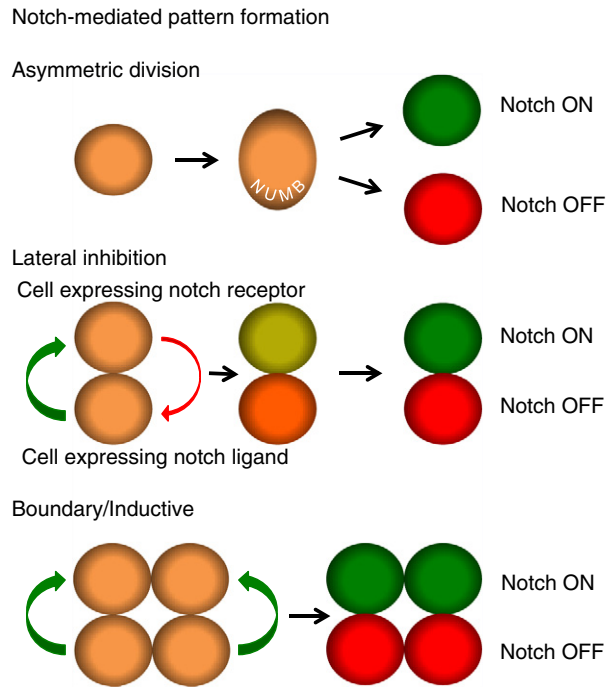


Figure 2 Mechanisms of cell fate determination or patterning. Through asymmetric segregation of Notch pathway genes (e.g., Numb), daughter cells can adopt distinct cell fates. Through cell–cell interaction, two similar cells can compete with Notch signaling, whereby Notch signaling inhibits expression of Notch ligands on the neighboring cell, resulting in high Notch signaling in one cell and low Notch signaling in the neighboring cell, and thus distinct cell fates. Finally, cells may express Notch ligands and induce Notch signaling in another group of cells, thereby inducing an alternate cell fate and creating a boundary of two cell types.

During these developmental processes, Notch signaling induces differentiation in some lineages while impairing it in others (Liu *et al.*, 2010).

Examples of Notch Roles in Development

During embryogenesis, Notch is important for anterior–posterior orientation, left–right orientation, and much of the anatomical layout of the organs and tissues. For example, during somitogenesis Notch signaling controls the formation of somite boundaries, which are the basis for the musculo-skeletal layout of the body. This pattern of distinct somites is achieved through the oscillatory negative feedback loop of HES1 creating waves of Notch signaling that influence the differentiation of multiple cell types along the embryo, for example, tendons, skeletal muscle, and cartilage (Rida *et al.*, 2004). In part, Notch activation allows myoblasts to differentiate at a slower rate, allowing the musculature to develop gradually while preventing the differentiation of muscle progenitor cells into mature muscle cells. Over the course of time, Notch signaling maintains progenitor cells that remain dormant until regeneration or growth of muscle is required (Bentzinger *et al.*, 2012).

Notch also plays a critical role in the development of the nervous system. Importantly, neural stem cells are maintained

by activation of the Notch signaling pathway, while inhibition of the pathway leads to differentiation into neurons (Hitoshi *et al.*, 2002). Notch controls the differentiation of multiple other cell types in the central and peripheral nervous system, such as Schwann cells and other glial cells.

Another system where Notch signaling plays multiple distinct roles is the hematopoietic system. Both activation and inactivation of Notch at specific times is imperative for the development of multiple lineages of hematopoietic cells (Kushwah *et al.*, 2014). During embryogenesis, Notch signaling directs development of definitive hematopoiesis via enabling the first hematopoietic stem cells to develop in the peri-aortic region. Later in development, Notch signaling appears to maintain the self-renewal capacity of hematopoietic stem cells. Stronger Notch signaling promotes differentiation toward the T cell lineage at the expense of B cell development. Interestingly, Notch signaling is then required in late B cell development to generate marginal zone B cells. Notch also plays multiple roles in myeloid lineages, though these are less well understood (Bigas *et al.*, 2012).

Many aspects of blood vessel development are dependent on Notch signaling. For example, Notch receptors 1 and 4 along with ligands DLL4 and JAG1 direct the specification of arteries versus veins, and contribute to new vessel sprouting and differentiation between endothelial and vascular smooth muscle cell fates (Fouillade *et al.*, 2012).

Notch also plays a role in the development of the liver, pancreas, skin and intestine, among many others. For a comprehensive review of Notch signaling in developmental processes, see Andersson *et al.* (2011).

Pathology

Given the broad effects of Notch signaling on normal development, it is not surprising that alterations in the Notch pathway will lead to abnormal development and disease. Indeed several disorders have been linked to mutations in Notch pathway members (Figure 3; Penton *et al.*, 2012).

For example, over 90% of patients with Alagille syndrome carry a micro-deletion or inactivating mutation in the Notch ligand JAG1 (Suskind and Murray, 2007), and a subset of the remaining patients carry mutations in the receptor Notch2 (Kamath *et al.*, 2012). This syndrome includes abnormal development of biliary ducts (in the liver), abnormal cardiac development leading to aortic valve dysfunction or Tetralogy of Fallot, abnormal vertebral development, as well as ocular and renal abnormalities. Cardiac abnormalities can also occur with Notch1 mutations, which have been found in patients with familial aortic valve disorders. Similarly, Notch3 receptor mutation or deletion leads to vascular abnormalities and is the major cause of CADASIL (cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy), a disorder characterized by migraine headaches, stroke, and dementia. These effects are due to loss of vascular smooth muscle leading to constriction of cerebral arteries, recurrent strokes, and specifically loss of white matter (Joutel *et al.*, 1996). Another example of musculoskeletal effects is revealed through mutations in the Notch ligand DLL3 in patients with spondylocostal dysostosis (SCDO). Aberrant Notch signaling leads to bone fusion and alignment issues, particularly

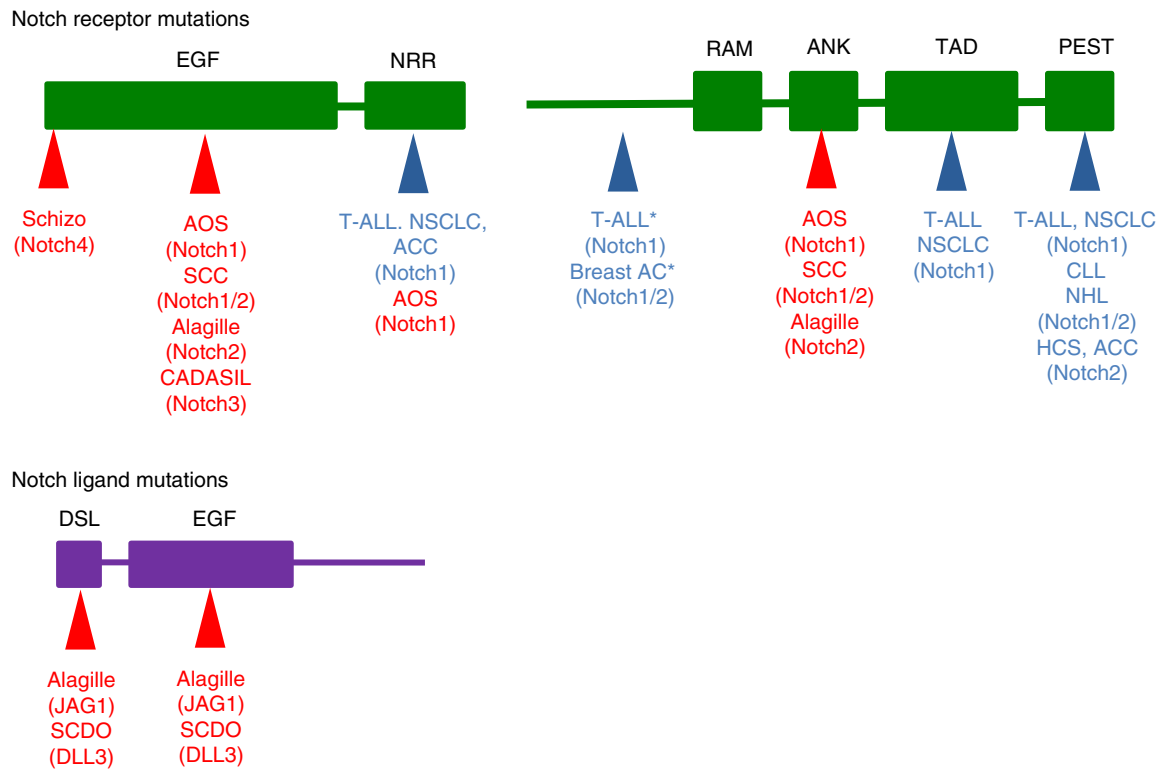


Figure 3 Notch pathway mutations in syndromes and cancer. Notch receptor ligand mutations can either activate (blue) or inhibit (red) Notch signaling. Specific receptor or ligand in parentheses. Schizophrenia (Schizo), Adams-Oliver syndrome (AOS), squamous cell carcinoma (SCC), cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL), T cell acute lymphoblastic leukemia (T-ALL), non-small-cell lung cancer (NSCLC), adenoid cystic carcinoma (ACC), chronic lymphocytic leukemia (CLL), non-Hodgkin lymphoma (NHL), Hajdu-Cheney syndrome (HCS), spondylocostal dysostosis (SCDO). See text for domain names.

in the ribs and spine (Tumpenny *et al.*, 2003). Interestingly, DLL3 is an inhibitory Notch ligand, which when expressed in a cell can bind Notch receptors in the Golgi and lead to their degradation. Thus, these DLL3 mutations lead to a loss of repression of Notch signaling, disrupting somitogenesis, and thus skeletal development (Chapman *et al.*, 2011). Rarely, mutations in HES7 or Lunatic Fringe are also found in patients with SCDO. Aberrant Notch signaling has a variety of consequences on bone development and disorders (Zanotti and Canalis, 2013). Similarly, five of six patients with the osteoporotic skeletal disorder Hajdu-Cheney syndrome (HCS) had activating mutations in Notch2 (Isidor *et al.*, 2011), suggesting Notch2 activation is involved in development of HCS. Finally, combined effects on vascular and musculoskeletal development can also be seen in patients with Adams-Oliver syndrome (AOS), a defect in skin and bone development of the head and distal extremities. Mutations in Notch1 or CSL were found in several families (Hassed *et al.*, 2012). Like Alagille syndrome, cardiac and vascular defects were also found in some patients with these mutations (Stittrich *et al.*, 2014).

Notch pathway dysregulation is also involved in schizophrenia, polycystic kidney syndrome, and Mayer-Rokitansky-Küster-Hauser syndrome (genital aplasia), further revealing the broad consequences of aberrant Notch signaling.

Importantly, Notch signaling has been found to be a major feature of cancer in many different tumor types. Indeed, Notch can have oncogenic or tumor-suppressor roles depending on

tissue or cell type (Lobry *et al.*, 2014). The first oncogenic role for aberrant Notch was described in T cell acute lymphoblastic leukemia (T-ALL) and lymphoma, where the discovery of activating Notch1 mutations occur in >50% of patients with T-ALL (Weng *et al.*, 2004) as well as other hematologic malignancies. Indeed, constitutive Notch activation leads to differentiation arrest in T cell progenitors and directly contributes to T cell malignancies in mice and humans. In addition, activating mutations in Notch1 or Notch2 have been found in 10–15% of chronic lymphocytic leukemia (CLL) patients (Puente *et al.*, 2011), and in several types of non-Hodgkin lymphoma (NHL) (Martinez *et al.*, 2014).

In other tumor types activating mutations in Notch receptors occur in smaller subsets. Notch1 mutations have been identified in 10% of non-small-cell lung cancer (Westhoff *et al.*, 2009) and Notch1/2 mutations in 10% of adenoid cystic carcinoma (ACC) (Stephens *et al.*, 2013). Additionally, rare translocations of Notch1 or Notch2 and amplification of Notch4 have been seen in breast cancers (Andre *et al.*, 2009).

In contrast, inactivating mutations of Notch1 or Notch2 occur in 75% of cutaneous squamous cell carcinomas (SCC) (Wang *et al.*, 2011), and less frequently in SCC arising in other tissues, such as the head and neck (Agrawal *et al.*, 2011; Stransky *et al.*, 2011). Such mutations either impair the function of the Notch receptors or lead to incomplete or unstable forms of the receptors, which are rapidly degraded. In these cancers, Notch signal appears to play a tumor-suppressor role.

These are a few examples of the role of aberrant Notch signaling in cancer, with additional roles for Notch activation in melanoma, colorectal adenocarcinoma, pancreatic carcinoma, medulloblastoma, among others. Additional tumors where Notch plays a tumor-suppressor role include B cell acute lymphoblastic leukemia, acute myeloblastic leukemia, small-cell lung cancer, hepatocellular carcinoma, among others. It is further reviewed in Ntziachristos *et al.* (2014).

Summary

In summary, the Notch pathway is a fundamental regulator of development in many tissue types. Not surprisingly, aberrant Notch signaling leads to multiple developmental abnormalities and Notch can play both oncogenic and tumor-suppressor roles in a wide variety of cancer types. The diversity of ligands, receptors, and downstream consequences make Notch a complex pathway requiring a detailed analysis of Notch signaling in each biologic system.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: The Hippo Pathway

References

- Agrawal, N., Frederick, M.J., Pickering, C.R., *et al.*, 2011. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. *Science* 333 (6046), 1154–1157.
- Andersson, E.R., Sandberg, R., Lendahl, U., 2011. Notch signaling: Simplicity in design, versatility in function. *Development* 138 (17), 3593–3612.
- Andre, F., Job, B., Dessen, P., *et al.*, 2009. Molecular characterization of breast cancer with high-resolution oligonucleotide comparative genomic hybridization array. *Clinical Cancer Research* 15 (2), 441–451.
- Bentzinger, C.F., Wang, Y.X., Rudnicki, M.A., 2012. Building muscle: Molecular regulation of myogenesis. *Cold Spring Harbor Perspectives in Biology* 4 (2), a008342.
- Bigas, A., D'Altri, T., Espinosa, L., 2012. The Notch pathway in hematopoietic stem cells. *Current Topics in Microbiology and Immunology* 360, 1–18.
- Chapman, G., Sparrow, D.B., Kremmer, E., Dunwoodie, S.L., 2011. Notch inhibition by the ligand DELTA-LIKE 3 defines the mechanism of abnormal vertebral segmentation in spondylocostal dysostosis. *Human Molecular Genetics* 20 (5), 905–916.
- Fouillade, C., Monet-Lepretre, M., Baron-Menguy, C., Joutel, A., 2012. Notch signalling in smooth muscle cells during development and disease. *Cardiovascular Research* 95 (2), 138–146.
- Guruharsha, K.G., Kankel, M.W., Artavanis-Tsakonas, S., 2012. The Notch signalling system: Recent insights into the complexity of a conserved pathway. *Nature Reviews Genetics* 13 (9), 654–666.
- Hassed, S.J., Wiley, G.B., Wang, S., *et al.*, 2012. RBPJ mutations identified in two families affected by Adams-Oliver syndrome. *American Journal of Human Genetics* 91 (2), 391–395.
- Hitoshi, S., Alexson, T., Tropepe, V., *et al.*, 2002. Notch pathway molecules are essential for the maintenance, but not the generation, of mammalian neural stem cells. *Genes & Development* 16 (7), 846–858.
- Isidor, B., Lindenbaum, P., Pichon, O., *et al.*, 2011. Truncating mutations in the last exon of NOTCH2 cause a rare skeletal disorder with osteoporosis. *Nature Genetics* 43 (4), 306–308.
- Joutel, A., Corpechot, C., Ducros, A., *et al.*, 1996. Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature* 383 (6602), 707–710.
- Kamath, B.M., Bauer, R.C., Loomes, K.M., *et al.*, 2012. NOTCH2 mutations in Alagille syndrome. *Journal of Medical Genetics* 49 (2), 138–144.
- Kushwah, R., Guezguez, B., Lee, J.B., Hopkins, C.I., Bhatia, M., 2014. Pleiotropic roles of Notch signaling in normal, malignant, and developmental hematopoiesis in the human. *EMBO Reports* 15 (11), 1128–1138.
- Ladi, E., Nichols, J.T., Ge, W., *et al.*, 2005. The divergent DSL ligand Dll3 does not activate Notch signaling but cell autonomously attenuates signaling induced by other DSL ligands. *J Cell Biol* 170 (6), 983–992.
- Lewis, J., 1998. Notch signalling and the control of cell fate choices in vertebrates. *Seminars in Cell and Developmental Biology* 9 (6), 583–589.
- Li, J., Mahajan, A., Tsai, M.D., 2006. Ankyrin repeat: A unique motif mediating protein-protein interactions. *Biochemistry* 45 (51), 15168–15178.
- Liu, J., Sato, C., Cerletti, M., Wagers, A., 2010. Notch signaling in the regulation of stem cell self-renewal and differentiation. *Current Topics in Developmental Biology* 92, 367–409.
- Lobry, C., Oh, P., Mansour, M.R., Look, A.T., Aifantis, I., 2014. Notch signaling: Switching an oncogene to a tumor suppressor. *Blood* 123 (16), 2451–2459.
- Martinez, N., Almaraz, C., Vaque, J.P., *et al.*, 2014. Whole-exome sequencing in splenic marginal zone lymphoma reveals mutations in genes involved in marginal zone differentiation. *Leukemia* 28 (6), 1334–1340.
- Ntziachristos, P., Lim, J.S., Sage, J., Aifantis, I., 2014. From fly wings to targeted cancer therapies: A centennial for notch signaling. *Cancer Cell* 25 (3), 318–334.
- Penton, A.L., Leonard, L.D., Spinner, N.B., 2012. Notch signaling in human development and disease. *Seminars in Cell and Developmental Biology* 23 (4), 450–457.
- Puente, X.S., Pinyol, M., Quesada, V., *et al.*, 2011. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* 475 (7354), 101–105.
- Rampal, R., Li, A.S.Y., Moloney, D.J., *et al.*, 2005. Lunatic fringe, manic fringe, and radical fringe recognize similar specificity determinants in O-fucosylated epidermal growth factor-like repeats. *Journal of Biological Chemistry* 280 (51), 42454–42463.
- Rida, P.C.G., Le Minh, N., Jiang, Y.-J., 2004. A Notch feeling of somite segmentation and beyond. *Developmental Biology* 265 (1), 2–22.
- Stephens, P.J., Davies, H.R., Mitani, Y., *et al.*, 2013. Whole exome sequencing of adenoid cystic carcinoma. *Journal of Clinical Investigation* 123 (7), 2965–2968.
- Stittrich, A.B., Lehman, A., Bodian, D.L., *et al.*, 2014. Mutations in NOTCH1 cause Adams-Oliver syndrome. *American Journal of Human Genetics* 95 (3), 275–284.
- Stransky, N., Egloff, A.M., Tward, A.D., *et al.*, 2011. The mutational landscape of head and neck squamous cell carcinoma. *Science* 333 (6046), 1157–1160.
- Suskind, D.L., Murray, K.F., 2007. Increasing the mutation rate for Jagged1 mutations in patients with Alagille syndrome. *Hepatology* 46 (2), 598–599.
- Turnpenny, P.D., Whitlock, N., Duncan, J., *et al.*, 2003. Novel mutations in DLL3, a somitogenesis gene encoding a ligand for the Notch signalling pathway, cause a consistent pattern of abnormal vertebral segmentation in spondylocostal dysostosis. *Journal of Medical Genetics* 40 (5), 333–339.
- Wang, N.J., Sanborn, Z., Arnett, K.L., *et al.*, 2011. Loss-of-function mutations in Notch receptors in cutaneous and lung squamous cell carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* 108 (43), 17761–17766.
- Weng, A.P., Ferrando, A.A., Lee, W., *et al.*, 2004. Activating mutations of NOTCH1 in human T cell acute lymphoblastic leukemia. *Science* 306 (5694), 269–271.
- Westhoff, B., Colaluca, I.N., D'Ario, G., *et al.*, 2009. Alterations of the Notch pathway in lung cancer. *Proceedings of the National Academy of Sciences of the United States of America* 106 (52), 22293–22298.
- Zanotti, S., Canalis, E., 2013. Notch signaling in skeletal health and disease. *European Journal of Endocrinology* 168 (6), R95–103.

CELL COMMUNICATION: INTRACELLULAR PATHWAYS

Contents

SH3 and SH2: Prototypic Domains to Mediate Regulatory Mechanisms in the Cell
The MAPK Signaling Cascades
The PI3K/Akt/mTOR Pathway
The JAK-STAT-SOCS Signaling Cascade
The PLC Pathway
Calcium and Calmodulin Signaling

SH3 and SH2: Prototypic Domains to Mediate Regulatory Mechanisms in the Cell

G Cesareni and D Peluso, University of Rome Tor Vergata, Rome, Italy and Santa Lucia Foundation, Rome, Italy

© 2016 Elsevier Inc. All rights reserved.

Glossary

3₁₀-Helix It is a type of protein secondary structure characterized by amino acids arranged in a right-handed helical structure.

The Domain Perspective of Protein Architecture and Function

Protein structure and function is dictated by the protein amino acid sequence. Proteins similar in amino acid sequence fold into structures that are similar and often have similar function. However, the overwhelming complexity of protein sequences and structures makes it difficult to identify some pattern in the apparent disorder. Soon it became clear, however, that proteins can be looked at from a higher, less-detailed perspective as a sequence of domains. Domains are peptide units, 50 to several hundreds of amino acid long, that are found repeated in several instances in proteomes and can be recognized on the basis of sequence homology. Domains have the property of being able to fold independently from the rest of the protein they are embedded in and are modular. To put it simply, each domain has a function, either enzymatic or structural – most often the ability to interact with other peptides – and the combinatorial assembly of a subset of domains in different protein instances contributes to build a versatile diversity in the proteome. The Simple Modular Architecture Research Tool (SMART) database (see Relevant Website section) annotated 1191 domain families in July 2014. For historical reasons, the discovery and the characterization of the so-called SRC homology (SH) domains contributed to the establishment of a signaling model based on a modular view of protein structure and function and will be described in some

detail in this article. The SH2 and SH3 domains are prototype modules that can promote the formation of signaling complexes in response to changes in the environment sensed by a cell.

A Brief History of the Discovery of SH Domains

In the mid-1980s, scientists studying oncogenic viruses realized that the SRC and ABL families of non-receptor tyrosine kinases share three regions of sequence homology. Although the first region SH1 (for SRC homology region 1) was soon identified as the kinase domain, the functions of the other two (SH2 and SH3) were not immediately clear. The observation that mutations either in the SH2 or the SH3 domains could have either positive or negative effects on the ability of the SRC protein to transform cells pointed to the prominent role played by these domains in the regulation of the enzymatic function of the protein containing them. Soon after, protein database searches revealed that these two domains are not restricted to non-receptor tyrosine kinases, but are present in a variety of protein contexts with apparently unrelated function. This observation implied that SH2 and SH3 domains are modular and confer some independent function that is not necessarily related to the protein kinase activity. The finding that many proteins containing SH3 or SH2 domains are involved in signal transduction immediately

alerted the researchers to investigate their prominent role in transferring regulatory signals from the cell membrane to the nucleus. The breakthrough came from the discovery that these domains can bind to peptides that share a common structure and/or sequence features. A variety of experimental approaches showed that SH3 domains have preference for peptides characterized by several prolines while SH2 bind to peptides containing a phosphorylated tyrosine. These seminal observations were instrumental in stimulating a large number of studies that eventually lead to the description of a complex network of interactions that underlies most of the cellular processes.

SH3

SH3 domains are about 60–70 residues long and are found in a large number of proteins involved in different cellular processes such as intracellular signaling, protein trafficking, cytoskeletal organization, and immune response. The SH3 domain family is one of the most populated family in many proteomes, from yeast to man. The number of SH3 domains roughly correlates with genome complexity, the human, drosophila, and yeast genome encoding about 300, 90, and 28 members, respectively. Although most SH3-containing proteins have a single SH3 domain, it is not uncommon to find

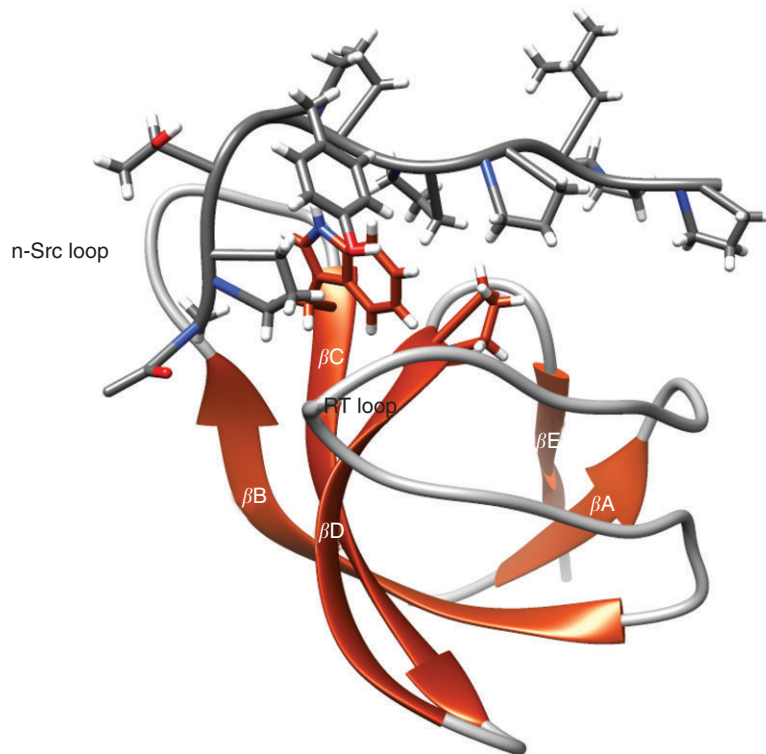
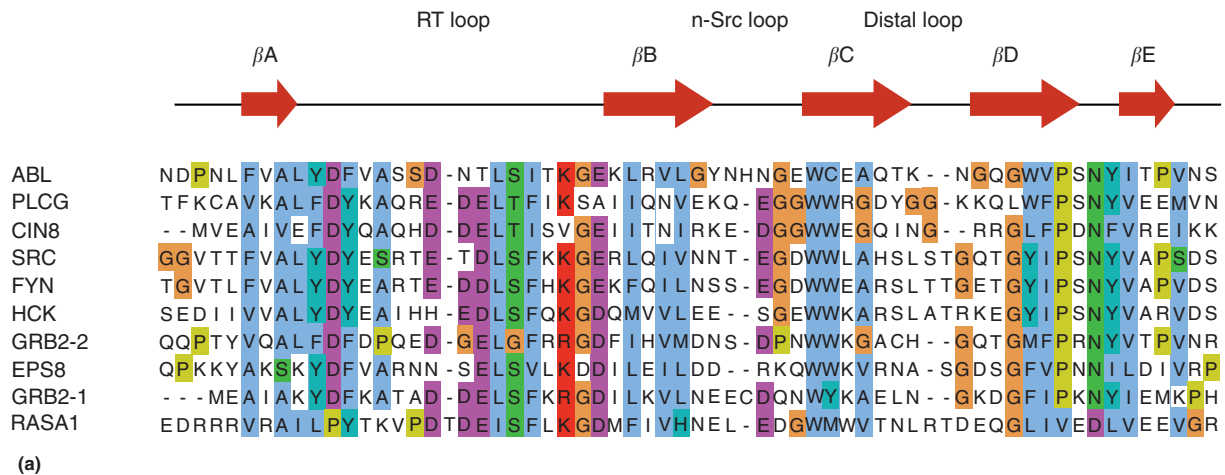


Figure 1 Structure of the SH3 domain. (a) Alignment of 10 representative SH3 domains to illustrate conserved positions. Conserved residues are colored according to amino acid type. Arrows above the alignment map the positions of β -strands. (b) Three-dimensional ribbon model of the SH3 domain of the kinase Abl complexed with the high-affinity peptide ligand 3BP-1 (APTYPPPLPP). β -Strands are depicted as red arrows. Some of the residues described in the text are shown with a ball and stick representation.

proteins with multiple copies and as many as five SH3 domains can be found in the dynamin binding protein, DNMBP, or in intersectin. The human proteome encodes 207 proteins containing a total of about 300 SH3 domains.

Structure

Even though the primary sequences of SH3 domains show little conservation, typically 25% between any two SH3 domains (Figure 1(a)), their three-dimensional fold is highly similar (Figure 1(b)).

The domain structure consists of five (or six) β -strands arranged in two perpendicular β -sheets to form a very conserved β -barrel. Each β -sheet is built from three antiparallel β -strands linked by three variable loops that are normally referred to as RT-loop, n-Src loop, and distal loop; the fourth and fifth strands are separated by a short 3_{10} -helix (Figure 1(b)).

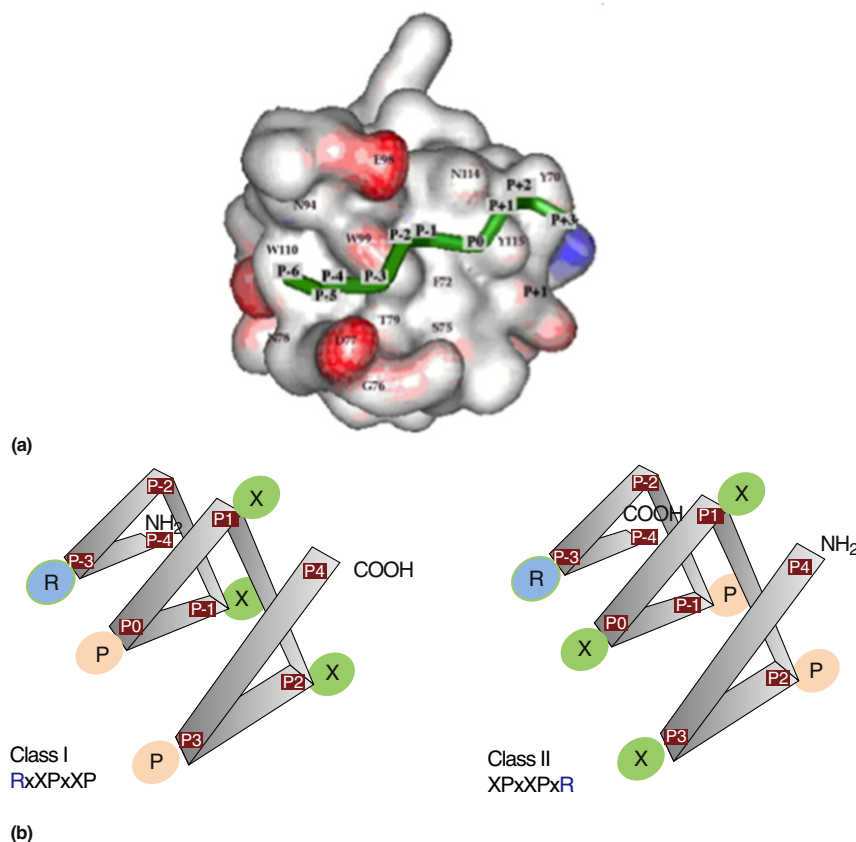
The surface utilized by most SH3 domains to bind their targets can be schematically split into three discrete patches consisting of two hydrophobic grooves, lined mainly by highly conserved hydrophobic residues and a specificity pocket (Musacchio *et al.*, 1992; Yu *et al.*, 1992). The shape and chemical characteristics of the specificity pocket is molded by the variable and flexible RT and n-Src loops that play an

important role in modulating recognition specificity within the domain family (Feng *et al.*, 1995).

Target Recognition Motifs

The observation that many SH3 ligands contain proline-rich motifs (Ren *et al.*, 1993) combined with the analysis of a growing number of structures of SH3 domains complexed with their ligands led to the formulation of a general SH3 binding model (Feng *et al.*, 1994; Lim *et al.*, 1994; Figure 2). Most SH3 ligands contain an (XPxXP) motif where P is a proline and X a hydrophobic residue. The two XP moieties in the core motif, when complexed with the SH3 surface, occupy the two hydrophobic grooves in the peptide-binding surface (Figures 1 and 2). The specificity pocket often contains negative residues and can host a positively charged side chain flanking the peptide ligand core motif. Some domains, such as the ABL SH3 domain in Figure 1, do not have the negative charge in the specificity pocket and preferentially bind peptides containing a large hydrophobic residue instead of the typical positively charged residue.

SH3 ligands bind to their receptors in a left-handed polyproline type II (PPII) helical conformation in either of two opposite orientations depending on the position of a positive



residue in the peptide sequence (Figure 2(b)). This ability to bind in the two orientations is a consequence of the pseudo symmetry of the left-handed PPII helix and it is also observed in other domains that bind poly-proline peptides. Positively charged residues, such as arginine and lysine, in the ligand peptide play an important role in the binding to SH3 domains, not only by providing additional binding energy, but also by orienting the ligand with respect to the SH3 domain binding surface (Feng *et al.*, 1994). In fact, depending on the position of the positive residue, either at the N- or C-terminus of the core motif, peptide ligands adopt either an N-to-C or C-to-N orientation. Peptides that conform to the consensus R/Kx#PX#P (where # is normally a hydrophobic residue) bind in a type I orientation, while peptides that are characterized by Px#PxR/K (type II motif) bind in the opposite orientation (Figure 2(b)). The SH3 domain of the protein kinase ABL, one of the most studied SH3 domains, is atypical since it binds to ligands in a class I conformation that have a tyrosine (or a large hydrophobic residue) at position P-3 (for residue nomenclature see Figure 2) in place of the positively charged side chain.

Atypical Binding Motifs

Two *in vitro* technologies, phage display and screening of peptide arrays, have made possible large surveys of SH3 domain binding preference (Tong *et al.*, 2002; Tonikian *et al.*, 2009). These studies have confirmed that most SH3 domains show preference for peptides that conform to classical type I or type II binding motifs. Over recent years, however, a growing number of reports have revealed several interactions that involve binding of the core PXXP-motif docking surface of the SH3 to a peptide that do not contain a classical #PX#P consensus. These findings have challenged the generality of the classical model (Saksela and Permi, 2012). In addition, it is possible that the perceived dominance of classical Class I or Class II motifs rather reflects a historical focus on non-receptor tyrosine kinases and that the relative prevalence of non-consensus ligands might eventually turn out to be significantly higher than currently appreciated.

For instance, the SH3 domain of amphiphysin 1 was shown to prefer class II peptides containing an arginine instead of an aliphatic residue at position P0 (Cestra *et al.*, 1999; Grabs *et al.*, 1997). A second atypical subfamily of SH3 domains includes the three CIN85 SH3 domains. This subfamily specifically selects peptides containing the consensus sequence PX(P/A)XXR (Musacchio *et al.*, 1992; Kurakin *et al.*, 2003; Yu *et al.*, 1992). Regions containing peptides conforming to this consensus were confirmed as the CIN85 binding sites in its established physiological partners Cbl-b, c-Cbl, BLNK, AIP1, SB1, and CD2. A similar consensus (PPPVIAPRP) was also observed in a region of PAK2 that is required for binding the SH3 domain of α PIX indicating that CIN85 and α PIX may have evolved similar strategies in peptide recognition (Feng *et al.*, 1995; Manser *et al.*, 1998).

More strikingly, some SH3 were shown to be able to accommodate peptides that are not proline rich. For instance, the SH3 domain of EPS8 binds to ligands that contain a PxxDY motif (Ren *et al.*, 1993; Mongiovi *et al.*, 1999). A second striking example of an atypical target peptide that is capable of binding to SH3 domains is the tyrosine-based peptide (RKXXYXXY)

found in the immune cell adaptor protein SKAP55. Peptides containing this motif can bind to a subset of SH3 domains, such as those of FYN and FYB, that prefer consensus Class I PXXP ligands in random peptide screening experiments (Feng *et al.*, 1994; Kang *et al.*, 2000; Lim *et al.*, 1994).

All the atypical binding modes described so far require the peptide to contact the same binding surface that was described for the classical type I and type II interactions (Figures 1 and 2), irrespective of whether the binding peptide contains a PxxP motif or not. However, a few SH3 domains have acquired in evolution the capacity to exploit different patches of their surface to promote the formation of functional complexes.

One classical example of an SH3-mediated interaction that does not involve a peptide in PPII conformation is the binding of the SH3 domain, a protein involved in peroxisomal assembly, PEX13p, to PEX5p. PEX5p does not contain a poly-proline motif and binds in α -helical conformation to an SH3 region that is different from the PPII peptide-binding pocket. Consistently, PEX14p, another peroxin which recognizes PEX13p SH3 via a canonical PXXP interaction, does not compete with binding of the PEX5p α -helical peptide (Feng *et al.*, 1994; Pires *et al.*, 2003). Interestingly, GRB2 and the guanine nucleotide exchange factor VAV were shown to form a complex via hetero-dimerization of their SH3 domains.

In summary, most of the 300 human SH3 domains *in vitro* bind to peptides that conform to classical class 1 and class 2 consensi. However, the list of atypical binders is growing, showing that the SH3 fold is very versatile and offers the possibility for exploring alternative binding modes of peptide binding and that this versatility has been extensively used in evolution.

SH2

The propagation of cell signals is often mediated by the reversible phosphorylation of specific tyrosine, threonine, or serine residues, which promote the formation of dynamic protein complexes. It was therefore exciting when it was proposed that SH2 domains preferentially bind peptides containing phosphorylated tyrosines (Tong *et al.*, 2002; Anderson *et al.*, 1990; Tonikian *et al.*, 2009; Mayer and Hanafusa, 1990). Together with tyrosine kinases and phosphatases, SH2 domains play a prominent role in many signaling processes by taking the 'reader' role in monitoring the modification events that add or remove phosphates from tyrosines in response to changes in the cell environment.

The SH2 domain is not the only module that is able to recognize phosphorylated tyrosine residues (Saksela and Permi, 2012; Yaffe, 2002). For instance, the phospho-tyrosine binding (PTB) domain, first identified in the adaptor proteins SHC and in the insulin receptor substrate 1 (IRS1), has the capability to bind phospho-tyrosines. The SH2 family though is by far the largest family displaying this property and, as such, has a central role in the signaling events triggered by receptor tyrosine kinases.

The development of a pTyr cell communication system, based on the interplay between tyrosine kinases, phosphatases, and proteins containing SH2 domains is likely to have played a considerable role in facilitating the evolutionary shift from unicellular organism to multi-cellularity (Cestra *et al.*, 1999;

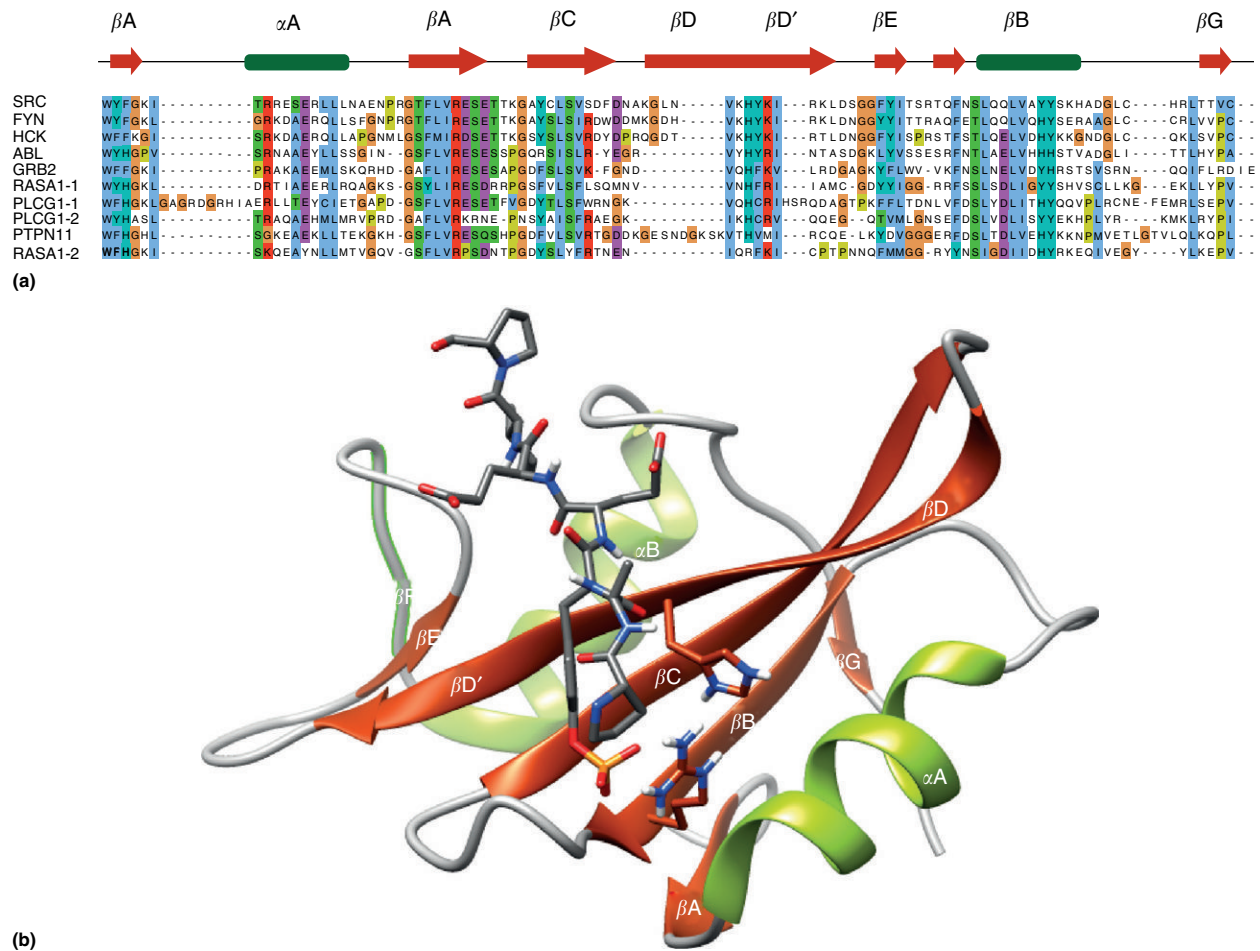


Figure 3 Structure of the SH2 domain. (a) Alignment of 10 representative SH2 domains to illustrate conserved positions. Conserved residues are colored according to amino acid type. Green boxes and red arrows above the alignment map the positions of α -helices and β -strands respectively. (b) Three-dimensional ribbon model of the SH2 domain of the kinase SRC in complex with the high-affinity phospho-tyrosine peptide ligand PQPYEEIP. Helices and strands are depicted as green ribbons or red arrows respectively. Some of the residues described in the text are shown with a ball and stick representation.

Lim and Pawson, 2010; Grabs *et al.*, 1997). It is therefore not surprising that functional SH2 domains are only found in the proteomes of metazoan. Homology searches in the human proteome identify 120 SH2 domains in 110 proteins.

Structure

The SH2 domain is approximately 100 residues long and folds into a conserved structure formed by an N-terminal α -helix (αA) and a 'core' four-stranded antiparallel β -sheet (βA , βB , βC , βD) followed by second β -sheet ($\beta D'$, βE , βF) that precedes an additional α -helix (αB) and a short C-terminal β -strand (βG) (Figure 3).

The structures of SH2 domains complexed with peptide ligands show the phospho-peptide hosted in two binding clefts that are separated by the core β -sheet (Pascal *et al.*, 1994; Lee *et al.*, 1994; Eck *et al.*, 1993; Figure 3(a)). The first cleft is positively charged and makes numerous important contacts with the phospho-tyrosine while the second cleft is much

more variable and is responsible for the selective recognition of residues that are C-terminal to the phospho-tyrosine.

The negatively charged phosphate inserts into a cleft in the core β -sheet, where its oxygen atoms are coordinated by two conserved arginines at position $\beta B5$ (β -strand B, position 5), and $\alpha A2$ and a histidine at $\beta D4$ (Figure 3(b)). These pTyr interactions are conserved in most SH2 domains and provide roughly half of the ligand binding energy. Consistently, mutations in either Arg $\beta B5$ or His $\beta D4$ abolish pTyr-specific binding (Marengere and Pawson, 1992).

Target Recognition Motifs

SH2 domains bind pTyr-containing peptide ligands with an affinity that is 1000-fold higher than that of the corresponding non-phosphorylated peptide suggesting that most of the binding energy depends on the interactions between the phosphate and the positively charged residues in the pTyr binding pocket (Piccione *et al.*, 1993). However, most SH2 domains display specificity and discriminate phospho-tyrosine

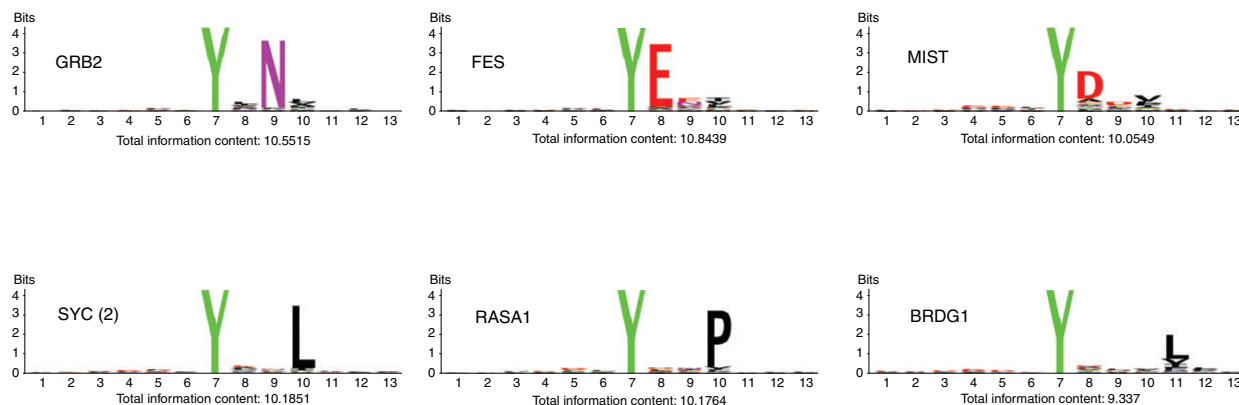


Figure 4 SH2 domain recognition motifs. The figure illustrates, by a sequence logo representation, the preferred motifs bound by 6 representative SH2 domains. The size of the font matches the level of conservation of the specific residue in a large collection of peptides that bind the domain in *in vitro* experiments. The green Y represents the phosphorylated tyrosine.

peptides on the basis of the phospho-tyrosine sequence context. Specificity is determined by the contacts with residues in the loop regions (CD, DE, BG) and strands β D and β E which define the specificity pocket in the SH2 domain. These loops and strands are more variable in sequence and are responsible for making contact with residues C-terminal to the pTyr (**Figure 1(b)**).

For instance, the GRB2 SH2 has a very strong preference for the presence of an asparagine two residues at the carboxy-terminus of the phosphorylated tyrosine (pYxNxx). Recent proteome wide peptide array experiments have allowed a nearly complete characterization of the recognition specificity of the entire family of the human SH2 domains (**Tinti et al., 2013**). These studies have pointed out that SH2 domains discriminate their targets by making important contacts with residues from position +1 up to +4 in the ligand peptide. For example, the SH2 domain of the FES, MST, SYK, RASA1, and BRDG1 proteins prefers the pYExxx, pYDxxx, pYxxLx, pYxxPx, or pYxxxP respectively (**Figure 4**). In addition to residues that in specific positions favor binding and confer specificity non permissive amino acid – residues that oppose binding – could play a role in shaping SH2 domain recognition specificity (**Liu et al., 2012**).

The preference of SH2 domains for specific phospho-peptide motifs has important biological consequences as demonstrated by the effects of manipulating their binding selectivity. Replacing a Thr at position EF1 in the SRC SH2 domain with a Trp, as in GRB2, determines a dramatic change in specificity with a strong preference for peptides having an Asn at the +2 position, virtually identical to that of GRB2 (**Marengere et al., 1994**). This demonstrates that domain selectivity can be attributed to specific residues. In addition, it was possible to show that binding selectivity correlates with the biological activity since SH2 domains are functionally interchangeable only if they share the same binding specificity.

Specificity and Promiscuity

The results of proteome wide studies of SH3/2 domain binding specificity, as outlined in the previous sections, have

shown that little divergence in domain sequence homology can account for relatively large changes in binding specificity. This is consistent with the observation that a few amino acid changes are sufficient to induce a specificity shift in peptide recognition modules (**Marengere et al., 1994; Panni et al., 2002**) and has implications for the interpretation of the observed rapid evolution of protein interaction networks (**Kiemer and Cesareni, 2007**).

Although individual SH3 and SH2 domains in general favor binding to peptides containing particular amino acid sequence motifs, these differences are often subtle and a considerable level of interaction promiscuity can be observed in experiments carried out *in vitro*. Any given domain can bind to a variety of short peptides with similar moderate affinity and many peptides in the proteome can bind to a number of different domains. Indeed there isn't really much room for specificity in the core peptide-binding grooves. Only a few core ligand residues directly contact the SH domains. Moderate affinities low (μ M to high nM) also imply that the interactions mediated by these domains are highly dynamic. In fact, the relatively rapid off rates imply that SH3/2-mediated interactions have the potential to reorganize, depending on available binding partners.

It has been argued that a certain level of promiscuity and redundancy can confer some robustness to the system (**Castagnoli et al., 2004**). For instance, a protein that exploits interactions mediated by its SH3 domain to localize to the cytoskeleton may find it advantageous to have several possibilities to dock to different cytoskeletal proteins.

However, a high degree of selectivity is achievable *in vivo*. Experiments in yeast have shown that at least one SH3 ligand peptide is absolutely selective for its known biologically relevant binding partner and does not detectably bind to any of the other 26 SH3 domains of the yeast *Saccharomyces cerevisiae* (**Zarrinpar et al., 2003**). In addition, the ability to link to several partners may contribute to the crosstalk between different pathways.

The limited specificity detected *in vitro*, which in several cases contrasts with the selectivity observed *in vivo*, implies that SH3/2-mediated interactions can be highly dependent on the molecular and cellular context. Several examples

have been reported, where additional surfaces on the domain and its peptide target within the physiological protein context can confer much greater affinity and/or specificity to an interaction. In addition other domains on the two potential binding partners can also contribute to overall specificity.

The studies described so far establish that the amino acid sequence of an SH3/2 domain can be used to predict its target peptide preference. This in turn has some predictive value for identifying physiologically relevant partner proteins for that domain. Although it is reasonable to assume that domain peptide-ligand pairs that have evolved the highest binding affinities and selectivities are likely to be physiologically relevant, the intrinsic specificity built into the fold of individual SH2 and SH3 domains are only one of the many factors that influence partner selection. To mention a few, the protein localization in cell compartments and the protein expression profiles of different cell types confer compartment and cell specificity to the functional complexes that are formed.

Function

The SH3 and SH2 folds are optimized for binding poly-proline type 2 helices or phospho-tyrosine peptides respectively, and

this is what they only do in their physiological context, if we neglect a few exceptions.

However, nature has exploited the ability to splice these simple functions to a variety of additional modular functions embedded in other domains, to develop, by a sort of 'molecular Lego,' a complex pattern of versatile regulatory networks. For instance, the combination of SH2 domains with either a kinase or a phosphatase domain may generate a positive or a negative feedback. The phosphorylation of a receptor can recruit via an SH2 domain a kinase or a phosphatase that either adds an extra phosphate (positive feedback) or removes an existing one from a different residue (negative feedback) (Pawson, 2004; Tonks and Neel, 2001). The combinatorial nature of the gene splicing processes that may occur in evolution offers a practically unlimited number of functional possibilities. We will describe here a few paradigmatic cases selected from a growing list of fascinating examples.

It needs to be pointed out that there is a fundamental difference between the interactions mediated by SH3 and by SH2 domains. The first ones are somewhat static and can only be affected by conformational changes of the domain or its target, which may result in affinity modulation. Alternatively, variations of the local concentrations of the partners or of competitor ligands may change the relative abundance of the complexes that are formed by the law of mass action.

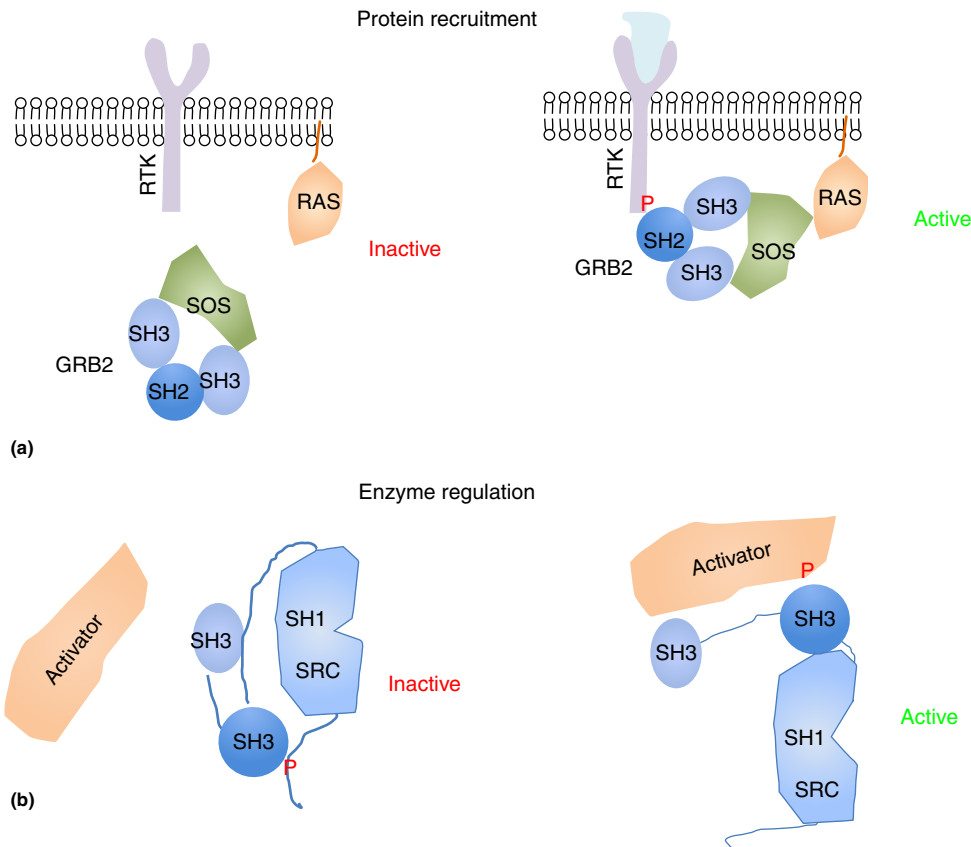


Figure 5 Mechanisms that regulate pathway activation. The two cartoons illustrate two examples by which intramolecular interactions mediated by SH domains modulate enzyme activity. (a) Regulations of RAS activation by re-localization of the SOS protein. (b) Regulation of activity of non-receptor protein kinases.

Complexes formed by the interaction between SH2 domains and phosphoproteins on the other hand are more dynamic since they can be switched on and off by changing the phosphorylation status of the target peptides either in *cis* (within the same protein) or in *trans* (in another protein). These types of interactions are ideally suited to modulate the cell response to growth signals that activate receptor tyrosine kinases as in the cases briefly described below.

Recruitment of the Son of Sevenless Protein to the Membrane

One of the most common function of SH domains is that of facilitating the colocation of proteins, for instance enzymes and substrates, thus promoting specific signaling cascades and pathway activation. A prototypic example of such a regulatory mechanism is offered by the re-localization of the son of sevenless (SOS) protein to activate the MAP kinase pathway (Figure 5(a)).

The stimulation of the MAP kinase growth pathway is triggered by the activation of the RAS small GTPase, which, in the active state, turns on the MAP kinase cascade (Katz *et al.*, 2007). RAS is attached to the membrane via a prenyl group and in the absence of growth factors it is in the inactive guanosine diphosphate (GDP) bound conformation. To be activated it needs to encounter the SOS protein that can promote the exchange of the GDP with the guanosine triphosphate (GTP). SOS, however, is normally located in the cytoplasmic compartment where it forms a complex with the growth factor receptor-bound protein 2 (GRB2) adaptor by docking two proline-rich regions to the two SH3 domains in GRB2. After stimulation by growth factors the growth factor receptor dimerizes and promotes the phosphorylation of a number of tyrosines in its cytoplasmic tail. Some of these phosphopeptides have affinity for the GRB2 SH2 domain thus promoting the prompt re-localization to the membrane of the GRB2-SOS complex. The re-localization of SOS to the membrane promotes the activation of RAS triggering the RAF/MEK/ERK phosphorylation cascade.

Modulation of the Activity of SRC Kinase and the SHP-2 Phosphatase

A second interesting and well-characterized mechanism controls the activity of non-receptor tyrosine kinases (Figure 5(b)). Many soluble tyrosine kinases such as, for instance, SRC and ABL are found tethered to SH3 and SH2 domains. Aside from guiding the enzymatic activity to specific substrates, these domains have evolved the ability to regulate the activity of the kinase via specific intramolecular interactions. In the inactive state the kinase is in a compact conformation stabilized by the interaction between the SH3 domain and the linker between the SH2 and the kinase domains. The compact structure is further stabilized by the interaction between the SH2 domain and a phospho-tyrosine peptide in the short carboxy-terminal tail (Tyr527 in SRC). Both interactions, however, are sub-optimal and they only occur because the two interacting partners are covalently linked.

Kinase activation can be promoted in several ways. For instance, de-phosphorylation of Tyr527 leads to dissociation of the SH2 domain and, as a consequence, also of the SH3 domain. Once relieved from interaction with the peptide targets in *cis* the two domains can bind proteins containing appropriate motifs leading to kinase activation and target phosphorylation. In addition, since the regulatory interactions are designed to be weak they can be readily reversed when the SH modules encounter better ligands.

Mechanisms that are similar to the ones that we have described for soluble tyrosine kinases are also involved in controlling the activity of the SHP-2 phosphatase. Despite being a phosphatase, a class of proteins that are normally associated with the dampening of growth signals, SHP2 is a positive regulator of the RAS-MAPK pathway because it dephosphorylates and inactivates negative regulators of the pathway such as RASGAP and SPROUTY (Jarvis *et al.*, 2006; Montagner *et al.*, 2005). SHP2 contains two SH2 domains linked on the amino-terminal side to the tyrosine phosphatase domain. The phosphatase activity is normally kept in check because the SH2 domains fold back onto enzymatic domain and inactivate it. However, when adaptor proteins such as GAB1 are phosphorylated on tyrosines by growth signals, the SH2 domains dock onto the phosphorylated peptides thereby releasing the inhibitory interaction with the phosphatase domain. These mechanisms have been extensively characterized because, as described below, mutations that alter it cause severe human pathologies. A large number of activating or inactivating mutations in the SH2 or phosphatase domains have been characterized thus shedding light on this sophisticated regulatory mechanism (Tartaglia *et al.*, 2006).

SH Domains and Diseases

Given the important role played by SH domains in signal transduction, it is not surprising that mutations within these domains or their targets have been implicated in human diseases. For instance, loss-of-function mutation in the SH2 domain of Bruton tyrosine kinase causes some cases of X-linked agammaglobulinemia, which is associated with a deficiency in mature B cells and susceptibility to bacterial infection (Saffran *et al.*, 1994). A distinct inherited immune defect, X-linked lymphoproliferative syndrome, results from mutations in the SAP gene encoding a small protein that contains only one SH2 domain and a short carboxy-terminal tail (Sayos *et al.*, 1998). Noonan's syndrome, a developmental disorder associated with cardiac defects, facial dysmorphism, and skeletal abnormalities, can be caused by mutations in the amino-terminal SH2 domain of the SHP-2 phosphatase. The mutations presumably remove the inhibitory effect of the SH2 domain on the phosphatase domain (Tartaglia *et al.*, 2001). Somatic mutations that activate Shp-2 are also common in juvenile myelomonocytic leukemia (JMML) (Tartaglia *et al.*, 2003).

Although many proteins containing SH3 domains are 'disease proteins,' clear examples of disease mutations affecting an SH3 domain or its targets are not numerous. A homozygous mutation (A335V) in the MED25 gene has been linked to Charcot-Marie-Tooth neuropathy. MED25 is a

transcription cofactor but its physiological function in transcriptional regulation remains obscure. The A335V missense mutation is located in a proline-rich region which has high affinity for the ABL SH3 domains. The mutation causes a decrease in binding specificity leading to the promiscuous recognition of a broader range of SH3 domain proteins.

It should also be considered that naturally occurring polymorphisms could affect predisposition to disease by altering SH domain recognition specificity. For instance, a Trp-to-Cys substitution in the C-SH2 domain of ZAP-70 causes in the SKG strain of mice the development of chronic arthritis. This is caused by the inability of the polymorphic ZAP-70 protein to associate with the T cell receptor thus leading to impaired T-cell signaling and survival of auto-reactive T cells.

Another interesting example is that of a common polymorphism in the human p53 gene. A single nucleotide variant in the p53 gene is reflected in the presence of either an Arg or a Pro at position 72. This polymorphism has been linked to susceptibility to several types of cancers probably as a consequence of a markedly different apoptotic potential of the two p53 protein variants (Dumont *et al.*, 2003). The polymorphism affects the sequence of a proline-rich peptide that has the potential to bind to SH3 domains. It was recently shown that some SH3 domains show a differential preference for either of the polymorphic proteins. It was proposed that the differential capability to form these complexes is at the basis of the observed differential susceptibility to cancer (Panni *et al.*, 2015).

Conclusions

The versatility of SH3 and SH2 domains has been exploited in evolution to develop a dynamic network of protein interactions that supports cell organization and signal transduction. The impact of this network of interactions in cell physiology is not confined to specific functions but pervades most cellular pathways. By providing the molecular basis for the formation of a sensitive and dynamic network, SH3 and SH2 domains can exert a broad effect on cellular behavior. Perhaps most importantly, their characterization has built a framework for the discovery and analysis of additional families of interaction domains and has crystallized a new way of thinking about cellular organization.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: Cytokine Receptors and Their Ligands; Receptor Tyrosine Kinases and Their Ligands

References

- Anderson, D., Koch, C.A., Grey, L., *et al.*, 1990. Binding of SH2 domains of phospholipase C gamma 1, GAP, and Src to activated growth factor receptors. *Science* (New York, NY) 250 (4983), 979–982.
- Castagnoli, L., Costantini, A., Dall'Armi, C., *et al.*, 2004. Selectivity and promiscuity in the interaction network mediated by protein recognition modules. *FEBS letters* 567 (1), 74–79.
- Cestra, G., Castagnoli, L., Dente, L., *et al.*, 1999. The SH3 domains of endophilin and amphiphysin bind to the proline-rich region of synaptotagmin 1 at distinct sites that display an unconventional binding specificity. *Journal of Biological Chemistry* 274 (45), 32001–32007.
- Dumont, P., Leu, J.I., Della Pietra 3rd, A.C., George, D.L., Murphy, M., 2003. The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. *Nature Genetics* 33 (3), 357–365.
- Eck, M.J., Shoelson, S.E., Harrison, S.C., 1993. Recognition of a high-affinity phosphotyrosyl peptide by the Src homology-2 domain of p56lck. *Nature* 362 (6415), 87–91.
- Feng, S., Chen, J.K., Yu, H., Simon, J.A., Schreiber, S.L., 1994. Two binding orientations for peptides to the Src SH3 domain: Development of a general model for SH3-ligand interactions. *Science* (New York, NY) 266 (5188), 1241–1247.
- Feng, S., Kasahara, C., Rickles, R.J., Schreiber, S.L., *et al.*, 1995. Specific interactions outside the proline-rich core of two classes of Src homology 3 ligands. *Proceedings of the National Academy of Sciences of the United States of America* 92 (26), 12408–12415.
- Grabs, D., Slepnev, V.I., Songyang, Z., *et al.*, 1997. The SH3 domain of amphiphysin binds the proline-rich domain of dynamin at a single site that defines a new SH3 binding consensus sequence. *Journal of Biological Chemistry* 272 (20), 13419–13425.
- Jarvis, L.A., Toering, S.J., Simon, M.A., Krasnow, M.A., Smith-Bolton, R.K., 2006. Sprouty proteins are *in vivo* targets of Corkscrew/SHP-2 tyrosine phosphatases. *Development* 133 (6), 1133–1142.
- Kang, H., Freund, C., Duke-Cohan, J.S., *et al.*, 2000. SH3 domain recognition of a proline-independent tyrosine-based RKxxYxxY motif in immune cell adaptor SKAP55. *EMBO Journal* 19 (12), 2889–2899.
- Katz, M., Amit, I., Yarden, Y., 2007. Regulation of MAPKs by growth factors and receptor tyrosine kinases. *Biochimica et Biophysica Acta* 1773 (8), 1161–1176.
- Kiemer, L., Cesareni, G., 2007. Comparative interactomics: Comparing apples and pears? *Trends in Biotechnology* 25 (10), 448–454.
- Kurakin, A.V., Wu, S., Bredesen, D.E., 2003. Atypical recognition consensus of CIN85/SETA/Ruk SH3 domains revealed by target-assisted iterative screening. *Journal of Biological Chemistry* 278 (36), 34102–34109.
- Lee, C.H., Kominos, D., Jacques, S., *et al.*, 1994. Crystal structures of peptide complexes of the amino-terminal SH2 domain of the Syp tyrosine phosphatase. *Structure* 2 (5), 423–438.
- Lim, W.A., Pawson, T., 2010. Phosphotyrosine signaling: Evolving a new cellular communication system. *Cell* 142 (5), 661–667.
- Lim, W.A., Richards, F.M., Fox, R.O., 1994. Structural determinants of peptide-binding orientation and of sequence specificity in SH3 domains. *Nature* 372 (6504), 375–379.
- Liu, B.A., Engelmann, B.W., Nash, P.D., 2012. The language of SH2 domain interactions defines phosphotyrosine-mediated signal transduction. *FEBS letters* 586 (17), 2597–2605.
- Manser, E., Loo, T.H., Koh, C.G., *et al.*, 1998. PAK kinases are directly coupled to the PIX family of nucleotide exchange factors. *Molecular Cell* 1 (2), 183–192.
- Marengere, L.E., Pawson, T., 1992. Identification of residues in GTPase-activating protein Src homology 2 domains that control binding to tyrosine phosphorylated growth factor receptors and p62. *Journal of Biological Chemistry* 267 (32), 22779–22786.
- Marengere, L.E., Songyang, Z., Gish, G.D., *et al.*, 1994. SH2 domain specificity and activity modified by a single residue. *Nature* 369 (6480), 502–505.
- Mayer, B.J., Hanafusa, H., 1990. Association of the v-crk oncogene product with phosphotyrosine-containing proteins and protein kinase activity. *Proceedings of the National Academy of Sciences* 87 (7), 2638–2642.
- Mongiovi, A.M., Romano, P.R., Panni, S., *et al.*, 1999. A novel peptide-SH3 interaction. *EMBO Journal* 18 (19), 5300–5309.
- Montagner, A., Yart, A., Dance, M., *et al.*, 2005. A novel role for Gab1 and SHP2 in epidermal growth factor-induced Ras activation. *Journal of Biological Chemistry* 280 (7), 5350–5360.
- Musacchio, A., Noble, M., Pauptit, R., Wierenga, R., Saraste, M., 1992. Crystal structure of a Src-homology 3 (SH3) domain. *Nature* 359 (6398), 851–855.
- Panni, S., Dente, L., Cesareni, G., 2002. *In vitro* evolution of recognition specificity mediated by SH3 domains reveals target recognition rules. *Journal of Biological Chemistry* 277 (24), 21666–21674.
- Panni, S., Salvio, S., Santonico, E., *et al.*, 2015. The adapter protein CD2AP binds to p53 protein in the cytoplasm and can discriminate its polymorphic variants P72R. *Journal of Biochemistry* 157 (2), 101–111.
- Pascal, S.M., Singer, A.U., Gish, G., *et al.*, 1994. Nuclear magnetic resonance structure of an SH2 domain of phospholipase C-gamma 1 complexed with a high affinity binding peptide. *Cell* 77 (3), 461–472.

- Pawson, T., 2004. Specificity in signal transduction: From phosphotyrosine-SH2 domain interactions to complex cellular systems. *Cell* 116 (2), 191–203.
- Piccione, E., Case, R.D., Domchek, S.M., *et al.*, 1993. Phosphatidylinositol 3-kinase p85 SH2 domain specificity defined by direct phosphopeptide/SH2 domain binding. *Biochemistry* 32 (13), 3197–3202.
- Pires, J.R., Hong, X., Brockmann, C., *et al.*, 2003. The ScPex13p SH3 domain exposes two distinct binding sites for Pex5p and Pex14p. *Journal of Molecular Biology* 326 (5), 1427–1435.
- Ren, R., Mayer, B.J., Cicchetti, P., Baltimore, D., 1993. Identification of a ten-amino acid proline-rich SH3 binding site. *Science* 259 (5098), 1157–1161.
- Saffran, D.C., Parolini, O., Fitch-Hilgenberg, M.E., *et al.*, 1994. Brief report: A point mutation in the SH2 domain of Bruton's tyrosine kinase in atypical X-linked agammaglobulinemia. *New England Journal of Medicine* 330 (21), 1488–1491.
- Saksela, K., Permi, P., 2012. SH3 domain ligand binding: What's the consensus and where's the specificity? *FEBS letters* 586 (17), 2609–2614.
- Sayos, J., Wu, C., Morra, M., *et al.*, 1998. The X-linked lymphoproliferative-disease gene product SAP regulates signals induced through the co-receptor SLAM. *Nature* 395 (6701), 462–469.
- Tartaglia, M., Martinelli, S., Stella, L., *et al.*, 2006. Diversity and functional consequences of germline and somatic PTPN11 mutations in human disease. *American Journal of Human Genetics* 78 (2), 279–290.
- Tartaglia, M., Mehler, E.L., Goldberg, R., *et al.*, 2001. Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. *Nature Genetics* 29 (4), 465–468.
- Tartaglia, M., Niemeyer, C.M., Fragale, A., *et al.*, 2003. Somatic mutations in PTPN11 in juvenile myelomonocytic leukemia, myelodysplastic syndromes and acute myeloid leukemia. *Nature Genetics* 34 (2), 148–150.
- Tinti, M., Kiemer, L., Costa, S., *et al.*, 2013. The SH2 domain interaction landscape. *Cell Reports* 3 (4), 1293–1305.
- Tong, A.H.Y., Drees, B., Nardelli, G., *et al.*, 2002. A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. *Science (New York, NY)* 295 (5553), 321–324.
- Tonikian, R., Xin, X., Toret, C.P., *et al.*, 2009. Bayesian modeling of the yeast SH3 domain interactome predicts spatiotemporal dynamics of endocytosis proteins. *PLoS biology* 7 (10), e1000218.
- Tonks, N.K., Neel, B.G., 2001. Combinatorial control of the specificity of protein tyrosine phosphatases. *Current Opinion in Cell Biology* 13 (2), 182–195. Available at: <http://www.sciencedirect.com/science/article/pii/S0955067400001964> (accessed 24.02.15).
- Yaffe, M.B., 2002. Phosphotyrosine-binding domains in signal Transduction. *Nature Reviews Molecular Cell Biology* 3 (3), 177–186.
- Yu, H., Rosen, M.K., Shin, T.B., *et al.*, 1992. Solution structure of the SH3 domain of Src and identification of its ligand-binding site. *Science (New York, NY)* 258 (5088), 1665–1668.
- Zarrinpar, A., Park, S.-H., Lim, W.A., 2003. Optimization of specificity in a cellular protein interaction network by negative selection. *Nature* 426 (6967), 676–680.

Further Reading

- Mayer, B.J., Saksela, K., 2004. SH3 domains. In: Cesareni, G., Sudol, M., Yaffe, M. B. (Eds.), *Modular Protein Domains*. Weinheim: WILEY-VCH, pp. 37–58.
- Pawson, T., Gish, G.D., Nash, P., 2004. The Sh2 domain: A prototype for protein interaction modules. In: Cesareni, G., Gimona, M., Sudol, M., Yaffe, M. (Eds.), *Modular Protein Domains*. Weinheim: WILEY-VCH, pp. 5–36.

Relevant Website

<http://smart.embl-heidelberg.de/>
EMBL-heidelberg.

The MAPK Signaling Cascades

R Seger and S Wexler, The Weizmann Institute of Science, Rehovot, Israel

© 2016 Elsevier Inc. All rights reserved.

Glossary

The mitogen-activated protein kinase (MAPK) Cascades are a central group of signaling pathways that respond to a variety of extracellular stimuli to induce and regulate cellular processes such as differentiation and proliferation.

They transmit their signals by sequential phosphorylation and activation of protein Ser/Thr kinases. Dysregulation of the cascade often results in pathologies such as cancer, diabetes, and other disorders.

General Overview

Cells respond to a large variety of extracellular signals through activation of transcription factors, which eventually induce the necessary cellular processes. These extracellular signals include signaling ligands such as hormones, growth factors, and cytokines, as well as environmental changes such as stress, changes in temperature, and osmotic pressure. Since most extracellular agents cannot cross the plasma membrane, in order for them to activate their corresponding genes, they use membranal receptors and intracellular signaling pathways to transmit the extracellular signals to target molecules either in the cytoplasm or in the nucleus. Such pathways are found in all organisms including bacteria, invertebrate, and ancient species and in many cases, they operate by sequential phosphorylations within protein kinase cascades. Most notable among them are the mitogen-activated protein kinase (MAPK) cascades that control fundamental cellular processes including proliferation, differentiation, motility, stress response, survival, and apoptosis (for recent reviews see: [Cargnello et al., 2012](#); [Keshet and Seger, 2010](#); [Plotnikov et al., 2011](#)).

The MAPK cascades are responsive to a wide variety of cues that transmit their signals via membranal receptors and other signaling components. These cascades funnel the information into either small G protein or adaptor proteins that are the immediate upstream activators of the cascades. The signals are then further transmitted through the cascade by three to five tiers of protein kinases. The kinases in each tier phosphorylate and activate the kinases in the downstream tier, to allow a rapid transmission of the signals to the cascade's targets. Three of the tiers; MAP3K, MAPKK, MAPK, are considered the core ones ([Figure 1](#)), while the upstream (MAP4K) or the downstream (MAPK-activated protein kinase; MAPKAPK) tiers are not always present. Kinases downstream of MAPKAPKs exist as well, but are not considered as part of the cascades. Each of the tiers is composed of several components, which are either distinct gene products, or alternatively spliced isoforms that often present distinct functions. The signals transmitted via the cascades are usually transient, with variable durations between minutes (transient) to hours (sustained), as determined by phosphatases and other regulatory proteins, as described below.

At present, four mammalian MAPK cascades are known and named according to their MAPK components: extracellular

signal-regulated kinase 1 and 2 (ERK1/2), c-Jun N-terminal kinase 1–3 (JNK), p38 α – δ (p38), and ERK5. Additional names for each of the cascades and their components are found in the literatures as well, and the full terminology can be found in ([Keshet and Seger, 2010](#)). Other kinases that share sequence similarities with the MAPKs have been identified as well (e.g., ERK3/4 and ERK7/8). However, the distinct mode of activation of these kinases suggests that they are not components of genuine MAPK cascades. Thus, about 70 genes are known today to encode for close to 200 distinct protein kinases that transmit signals within the cascades. Notably, phosphatases and other regulators bring the number of components of the entire MAPK system to well over 300. From the above, the MAPK and MAPKAPK components phosphorylate and modulate the activities of hundreds of substrates in various subcellular localizations and organelles ([Yoon and Seger, 2006](#)). Each of the cascades regulates several distinct, and sometime overlapping cellular processes. Generally, the ERK1/2 cascade plays a major role in proliferation and differentiation, JNK and p38 cascades are mainly stress-related, and the ERK5 cascade responds equally to both stress and mitogenic signals. Dependent on the cell and stimulation types, the cascades may regulate noncanonical and even opposing functions.

The ability of these seemingly simple, linear cascades to integrate a large number of extracellular stimuli, and regulate a variety of cellular processes has attracted much attention over the past two decades. Some of the mechanisms that are involved in the determination of signaling specificity are the intrinsic properties of the cascade, including: (1) the number of components in each tier of the cascade, (2) the unique selectivity of the MAP4K, MAP3K, and MAPKK components, (3) the progressive signal amplification, (4) the switch-like nature of the cascade, and (5) the rapid and transient nature of the transmitted signals. In addition, fine-tuning of the cascades' specificity is often achieved by: (1) interaction with other signaling cascades, (2) presence of many scaffolding proteins that bring components close together, (3) substrate competition, and (4) spatio-temporal regulation of the various components ([Shaul and Seger, 2007](#)). In this article, we cover the nature and functions of the four cascades, as well as their regulatory components. In addition we describe diseases that are induced and maintained by dysregulation of the cascades.

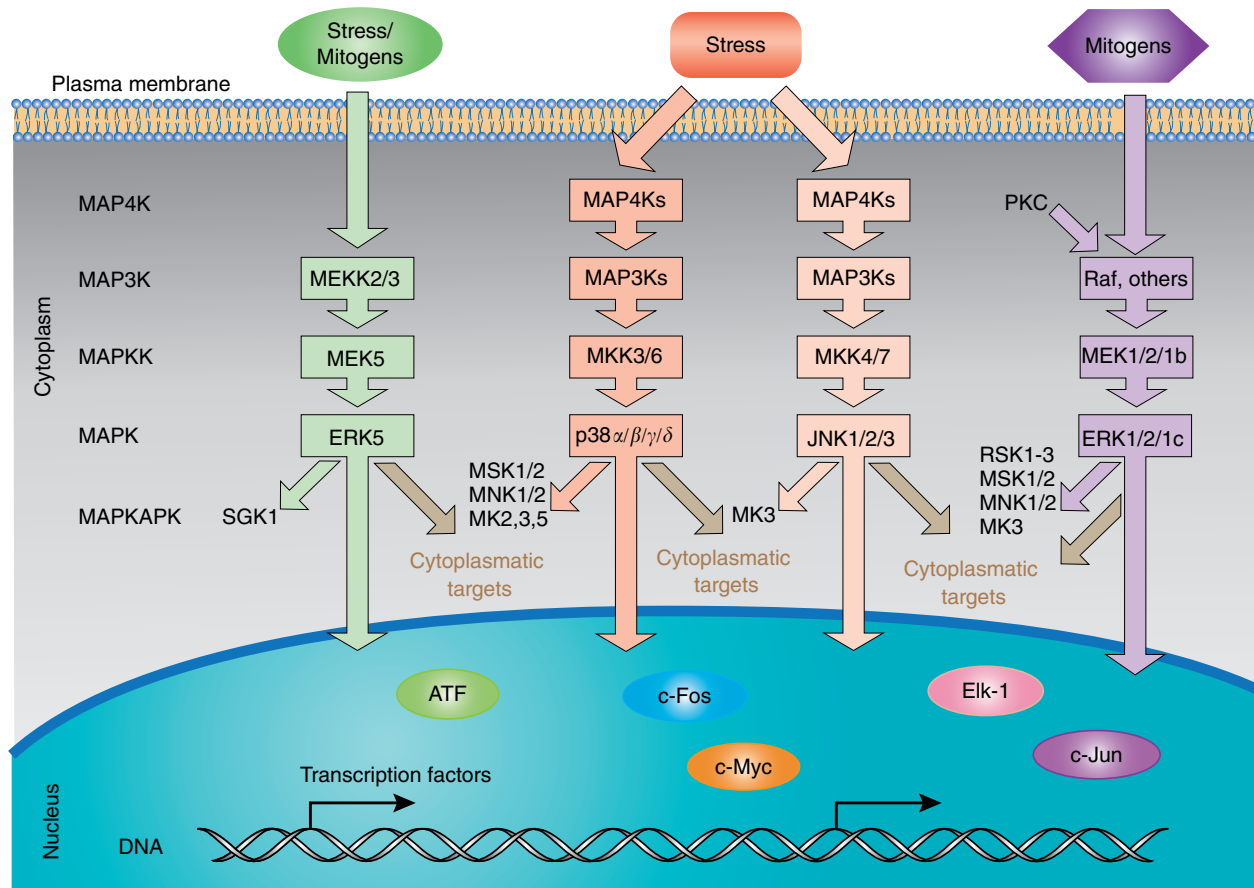


Figure 1 Schematic representation of the 4 MAPK signaling cascades.

Components of MAPK Cascades

The ERK1/2 Cascade

As all MAPK cascades, the ERK1/2 one is activated by many extracellular stimuli, including mainly growth factors, cytokines, and other mitogens (recent reviews are: [Matallanas et al., 2011](#); [Plotnikov et al., 2011](#); [Roskoski, 2012](#)). These stimuli usually transmit their signals by various membranal receptors that further activate a nucleotide exchange factor (SOS) and a small GTPase (Ras) at the plasma membrane. Other ways of Ras or Raf activation have been suggested as well, but those are much less frequent. No components are found in the MAP4K level of the cascade, and the main components at the MAP3K level are B-Raf and C-Raf, that are activated mainly by Ras-induced dimerization. A-Raf, KSR1/2 as well as MEKK1, Mos, and Cot1 are additional MAP3Ks, but those are activated by distinct mechanisms in restricted cell type- and stimulus-specific manners. The MAP3Ks then phosphorylate the MAPKKs MEK1 and MEK2 on the two serine residues within the activation motif Ser-Met-Ala-Asn-Ser in the kinases' activation loop. In turn, these MEKs phosphorylate ERK1 and ERK2 on the Tyr and Thr residues in the Tyr-Glu-Thr motif within their activation loop, inducing their full activation. Alternatively spliced forms of MEK1 (MEK1b) and ERK1 (ERK1c) have been described as well, and those seem to generate a unique

signaling route that regulates distinct processes than that of the MEK1/2-ERK1/2 routes ([Wortzel and Seger, 2011](#)). RSK1-3, MSK1/2, MNK1/2, and in some cases MAPKAPK3 and MK5, are the components at the MAPKAPK level, which are all phosphorylated and activated by ERK1/2 but can also be phosphorylated by other signaling kinases. The down-regulation of the cascade, which is mediated by several phosphatases, and the effect of the subcellular localization of the components on their activity are described below. Interestingly, while the activity of Rafs and MEK1/2 is restricted mainly to their downstream components, ERK1/2 and MAPKAPKs phosphorylate a large number of substrates, localized in various cellular compartments. The phosphorylation of all these substrates allows the cascade to induce and regulate a variety of cellular processes. Dysregulation of the cascade results in several diseases, including mainly cancer, as well as developmental and neurological disorders, as will be described below.

The p38 Cascade

The activity of this cascade is induced by various stress factors and ligands that operate via different receptors, including apoptosis-related receptors, GPCRs, and even receptor Tyr kinases (RTKs, reviewed in [Shiryaev and Moens, 2010](#)). In

addition, some of the physical stresses (e.g., heat and osmotic shock) are thought to operate in many cases via receptor-independent machinery that requires changes in membrane fluidity or other specialized signaling systems. Then, the signals are transmitted either by the small GTPases Rac and CDC42, or occasionally, activatory adaptor proteins. These two modes of upstream activators then stimulate protein kinases at the MAP3K tiers of the cascade. About 20 genes encode protein kinases at the MAP3K tier, and those are very similar to the MAP3Ks of the JNK cascade. They are activated mainly by small GTPases, but under some conditions, also by a few MAP4Ks, or even directly by adaptor proteins. This large number of MAP3Ks transmits the signals to a much smaller number of MAPKKs (mainly MKK3, MKK6), phosphorylating them on Ser and Thr residues at the typical Ser-Xaa-Ala-Xaa-Ser/Thr motif, in their activation loop. The next tier of the cascade is composed of products of four MAPK genes, including p38 α (SAPK2 α) and p38 β (SAPK2b) as well as p38 γ and δ . Importantly, p38 genes also express several alternatively spliced forms, bringing the number of isoforms of this group to ten, with molecular weights that vary between 32 and 54 kDa. These isoforms are all activated by phosphorylation of the Tyr and Thr residues in the Thr-Gly-Tyr motif in their activation loop. Unlike ERK1/2, p38 members can be localized both in the nucleus and/or in the cytosol, and their translocation upon stimulation seems to be bi-directional. The mechanism of nuclear translocation is similar to that of JNK (described below), while the export seems to be mediated by interaction with MAPKAPK2, which contains a nuclear export signal. Once these p38 members are activated, they either transmit the signal to the MAPKAPK level components (MAPKAPK2, MAPKAPK3, MNK1/2, MSK1/2, and MK5/PARK), or phosphorylate a large number of cytoplasmic and nuclear targets. The phosphorylation of all the substrates allows the cascade to induce and regulate a variety of cellular processes that include stress response, inflammation, immunological effects, and more. Dysregulation of the cascade results in several diseases, including inflammation-related diseases, neurological disorders, and diabetes, and even cancer.

The JNK Cascade

Similar the p38 cascade, the JNK cascade is responsive to stress/apoptosis-related receptors, receptor-independent physical stresses, GPCRs, and even RTKs (reviewed in [Bogoyevitch et al., 2010](#); [Haeusgen et al., 2011](#)). Those receptors, or receptor-independent stress-induced signals, further transmit the signals to adaptor proteins or small GTPases. The GTPases then are able to activate kinases in the MAP4K, and in some cases, MAP3K tiers of the JNK cascade. The kinases at the MAP4K tiers, include Ste20-like kinases, that can be shared with the p38 cascade, and can all phosphorylate and activate the kinases at the MAP3K tier. Again, most of the MAP3Ks are shared with those of the p38 cascade, although some of them have been reported only in the JNK cascade. Next, the activated MAP3Ks transmit the signals to kinases at the MAPKK level, which are mainly MKK4 and MKK7. Interestingly, MKK7 seems to express a particularly high number (~ 6) of alternatively spliced isoforms with distinct activities and those, together with alternative spliced isoforms of the MKK4, were

suggested to extend the specificity of the cascade. As the other MAPKKs, MKK4, and MKK7 are activated by phosphorylation of the typical Ser-Xaa-Ala-Xaa-Ser/Thr motif and are then able to transmit the signal further to the JNK level. Three genes (JNK1-3, SAPKs) encode the JNK proteins, which are translated from a total of ten JNK alternatively spliced transcripts. Interestingly, the spliced isoforms of JNK1 and JNK2 are either 46 kDa or ~ 54 kDa proteins, while JNK3 is mostly translated to a 52 kDa protein. The activation loop of JNK contains a Pro in the Xaa position (Thr-Pro-Tyr) and, as with the other MAPKs, both Thr and Tyr need to be phosphorylated to achieve activation. This cascade appears to be a major regulator of nuclear processes, specifically transcription, and the number of its cytoplasmic substrates is limited. Only one MAPKAPK, MAPKAPK3, is currently known to be a component of this MAPK cascade, indicating that most of the signaling is mediated by the JNK. These kinases usually associate with their targets and activate them to induce processes such as apoptosis, immunological effects, stress response and more. Again, dysregulation of the cascade was shown to lead to pathologies such as neurological disorders, diabetes, inflammation, and cancer.

The ERK5 Cascade

This cascade is activated by mitogens as well as stress stimuli, but is the least studied out of the four cascades (reviewed in [Roberts et al., 2009](#)). The mechanism of upstream activation of the cascade has not been fully elucidated, but may include activation of RTK/PTKs that either transmit their signals to the adaptor protein Lad1, or to WNK1, which acts as a MAP4K. These components activate the kinases in the MAP3K level, which are MEKK2/3 and possibly also TPL2 and MLTK. The MAP3K tier kinases then phosphorylate and activate the two alternative spliced MAPKK isoforms MEK5a (50 kDa) and MEK5b (40 kDa), which are both active kinases. As other MAPKKs, this phosphorylation occurs on the Ser-Xaa-Ala-Xaa-Thr activation motif of MEK5. The MEK5s then activate the MAPK, ERK5 by phosphorylating it on both Thr and Tyr residues within the sequence Thr-Glu-Tyr, which is very similar to that of ERK1/2. However, in spite of the similarity, ERK5 cannot be phosphorylated by MEK1/2, and MEK5 cannot phosphorylate ERK1/2. There is one gene that encodes ERK5, which is translated to one main protein of ~ 100 kDa, a size that gave this kinase its other name – big MAPK1 (BMK1). Upon stimulation, ERK5 phosphorylates the serum and glucocorticoid-activated kinase (SGK), which may serve as a MAPKAPK of this cascade. However, this does not seem to be the main substrate of the cascade, as several transcription factors have been identified as ERK5 substrates as well. Interestingly, unlike the other MAPKs, ERK5 was found to influence transcription by additional mechanisms, including direct protein–protein interactions with transcription factors, or intrinsic transcriptional activity that is mediated by its C-terminal noncatalytic half. These activities, allow ERK5 to mediate physiological processes including proliferation, angiogenesis, immunological processes, and stress responses. The cascade has also been implicated in cancer, and is likely to participate in other pathologies as well.

Regulation of the MAPK Cascades

The MAPK cascades are central signaling moieties, capable of communicating between many extracellular signals to various distinct and even opposing cellular processes. To avoid erroneous signaling, the activities of the MAPK cascades need to be well regulated. This is mediated by a set of regulatory proteins and processes that keep the level, duration, and localization of the signal under control, to ensure proper specificity and functioning, and avoid pathological outcomes. Some of the important proteins that directly regulate the cascades and their roles are described below.

Phosphatases

As mentioned above, all components of the MAPKK, MAPK, MAPKAPK and some of the MAP4K, and MAP3K tiers of the cascades are activated by phosphorylation. Inhibition of some of the components, as well as modulation of their activity often occurs by phosphorylation as well. Thus, phosphoprotein phosphatases are positioned to impact all aspects of MAPK cascade function (mainly negatively but in also positively; reviewed in Nunes-Xavier *et al.*, 2011; Yao and Seger, 2004). As of today, not much is known on the phosphatases that regulate the components at the upstream levels MAP4K, MAP3K, and MAPKK, but it is assumed that they are mostly regulated by the general Ser/Thr phosphatase PP2A. Another phosphatase that regulates MAP3K level components is PP5, which dephosphorylate activatory phosphorylations of ASK1 and Raf1, in order to inhibit their activity. On the other hand, the mechanism of MAPKs' dephosphorylation and inactivation is much better studied. Thus, dephosphorylation of either the pThr, pTyr or both in the activation loop of MAPKs causes a substantial decrease in activity. This is controlled by pTyr, pSer/Thr, or specific dual specificity phosphatases (DUSPs) known as MAPK phosphatases (MKPs). Interestingly, MKPs are usually not expressed in resting cells, and the initial MAPK regulation after stimulation is mediated by nonspecific phosphatases that are either pTyr or pSer/Thr phosphatases. The MKPs are then transcribed, and take over the dephosphorylation 45–75 min after stimulation.

This group of MKPs contains 11 proteins whose expression is different in various cell lines. In addition, each MKP member has different specificity toward the various MAPK. Some (e.g., MKP3) are specific to ERK1/2, while others are specific to JNK and p38, and MKP6 has similar specificity to ERK1/2, JNK, and p38. The specificity is mainly determined by the variant amino acids in the binding D-domain of MKPs. Not much is known about the MKPs that participate in the regulation of ERK5. Importantly, different MKPs are localized in different compartments, as some of them are exclusive to the nucleus, and some are found only in the cytoplasm. In addition, the phosphatases may strongly interact with their target MAPKs, to regulate not only their activity but also their subcellular localization. Thus, the distinct features of the MKPs enable differential regulation of MAPKs and their downstream processes.

Scaffold Proteins

Alongside mechanisms of cascade inhibition by phosphatases, scaffold proteins are thought to play a role in conferring

signaling specificity as well (reviewed in Pullikuth and Catling, 2007; Witzel *et al.*, 2012). These proteins allow formation of multicomponent complexes with two or more components of each of the cascades. Proteins that bind just one component are not considered as scaffolds, but may often serve as anchoring proteins to their interacting components, directing them to their right localization. The scaffold proteins have several functions, including: (1) facilitation of signaling rate by bringing distinct cascade components in close proximity to each other, (2) directing MAPK components to their proper subcellular localization, (3) upstream activators, (4) downstream targets, (5) various regulators, or even (6) stabilizing MAPK components due to protection from phosphatases and proteinases. These properties often determine signaling thresholds and allow regulation of the proper downstream cellular processes. Currently, the number of scaffold proteins known is limited, not exceeding ten for any cascade. However, it is likely this number is much higher.

One of the best-studied scaffolds of the ERK1/2 cascade is the KSR1, initially identified as a key component in the ERK signaling in *Drosophila melanogaster* and *Caenorhabditis elegans*. KSR1 has high homology to Raf1, and in some extreme cases can act as a MAP3K component by phosphorylating MEK1/2 (McKay *et al.*, 2011). However, it seems that the main function of this protein is the scaffolding of components of the cascade. Thus, in quiescent cells, KSR1 is associated with MEK1/2, but not with ERK1/2 or Rafs, and its activity is inhibited by the associated protein IMP1. Upon stimulation, the IMP1 is polyubiquitinated and degraded, which consequently induces regulatory changes in the protein. Consequently the KSR1 complex translocates to and interacts with the plasma membrane, recruiting an activated Raf1, which facilitates activation of MEK1/2. Simultaneously, ERK1/2 are recruited to the activated complex, activated, released from the complex and translocate to their sites of action. The mechanisms of action of most other scaffold proteins of the MAPK cascades seems to be less regulated than that of KSR1, as they undergo only limited numbers of regulatory phosphorylations.

Crosstalk with Other Signaling Cascades

Crosstalks and interplays between the distinct MAPK cascades and with other signaling components is an important regulatory mechanism in the determination of signaling specificity (reviewed in Fey *et al.*, 2012; Mendoza *et al.*, 2011). Such crosstalks can either affect the activity of components within the distinct MAPK cascades or impose combinatorial effects on targets downstream of the distinct cascades/pathways. For example, some of the cascades participate in negative upstream regulations of upstream components, such as the phosphorylation of EGFR by p38 that inhibits its activity toward the ERK1/2 cascade. The effects can occur within the various cascades, as exemplified by the phosphorylation of Ser298 of MEK1 by the PI3K-AKT-activated component PAK1. This phosphorylation induces further activation of the ERK1/2 cascade, demonstrating a positive influence of the PI3K/AKT pathway on the ERK1/2 cascade. An example of downstream crosstalk is the activation of IL2 gene promoter. This promoter requires binding of several transcription factors for its activation. Interestingly, some of the

factors lie downstream of distinct signaling cascades, including AP1 downstream of the JNK cascade, CREB downstream of PKA, or ERK1/2 cascades, and NFkB that is a representative of as yet another signaling pathway. Activation of several of these transcription factors upon stimulation is required for full promoter activity, while activity of just one of them causes a limited activation and may induce distinct physiological function. These mechanisms determine the signaling specificity by changing the output of the cascades at upstream components, various tiers, or downstream the MAPK signaling, to strongly affect the outcome of the MAPK signals.

Subcellular Localization of MAPK Components

Most components of the four MAPK cascades are localized in the cytoplasm of resting cells, and are translocated to the nucleus and other cellular compartments after stimulation (reviewed in Flores and Seger, 2013). In the ERK1/2 cascade the cytoplasmic localization is rapidly changing upon stimulation in various ways. The Rafs are recruited to the plasma membrane by the activated Ras, while MEK1/2 are rapidly shuttling in and out of the nucleus. Simultaneously, ERK1/2, as well as most of the MAPKAP, detach from their cytoplasmic anchors and translocate to the nucleus where they can stay for minutes to hours. The mechanism that allows these rapid changes of ERK1/2 and MEK1/2 translocation to the nucleus is phosphorylation-dependent binding to importin7. JNK and p38 also translocate to the nucleus in most cell types and upon many stimulations. Shortly after stimulation, these MAPKs detach from their cytoplasmic anchors, and interact with either importin7 or importin9 that further associate with importin3 to form a heterotrimer of importin7/9-importin3-MAPK that induces their translocation. ERK5 can be localized in the cytoplasm of resting cells and translocates to the nucleus upon stimulation as well. This is mainly mediated by a nuclear export signal, which is exposed in the inactive protein. However, unlike the other MAPKs, stimulation causes changes in conformation that hinder the export signal and expose canonical nuclear localization signals in the protein that interacts with importin- α/β to translocate the ERK5 to the nucleus. Interestingly, the various MAPK components can be localized in various cellular organelles and compartments as well. These localizations are mostly induced by stimulation-induced translocation, such as that of ERK1/2 to the mitochondria, or ERK1c to the Golgi. The mechanism of these translocations and their exact function is still obscured.

MAPK Signaling in Diseases

Since the MAPK cascades are such a central regulator of most if not all stimulated cellular processes, their dysregulation induces or is involved in many diseases (Kim and Choi, 2010). Below we discuss the involvement in cancer, developmental diseases and inflammation. However, the involvement of the cascade was noticed also in neurological disorders, diabetes, polycystic kidney disease, and others. Therefore, understanding the role of the MAPK cascades in conjunction with developing agents to circumvent these signaling-related diseases

may lead to novel strategies in the combat of a large number of diseases.

MAPK and Cancer

Dysregulation of the MAPK cascades is profoundly linked to cancer initiation and progression (reviewed in Lawrence *et al.*, 2008). Indeed, many known oncogenes encode proteins that participate in the activation or regulation of the MAPK cascades, and in particular the ERK1/2 one that is found activated in more than 85% of all tumors. Interestingly, this activation is not only mediated by upstream components of growth factor-related signaling, but often seen in cancer cells transformed by activated components of other signaling pathways. The most abundant oncogene upstream the ERK1/2 cascade is Ras, the various activating mutations of which account for about 33% of all cancer types. In addition, activating mutations of B/C-Raf, particularly B-Raf, act as driving oncogenes in about 5% of all cancers, but are particularly known as a cause for metastatic melanoma (~60%). Activating MEK1/2 mutations are the cause of about 1% of cancers, while no activating ERK1/2 mutations have been reported in cancer as yet. Being such a central cancer regulator, the ERK1/2 cascade is considered as a prominent therapeutic target. Indeed, several efficient Rafs and MEK1/2 inhibitors have been developed, and found to be efficient in combatting ERK1/2 cascade-addicted cancers, in particular B-Raf mutated metastatic melanoma. However, after several months of treatment, all patients develop resistance, indicating that more efficient inhibitors of the ERK1/2 cascade are required to better target the cascade in this cancer.

Other MAPK cascades have been reported to play a role in cancer, albeit their involvement does not seem to be as pronounced as the ERK1/2 cascade. Thus, activated JNK has been reported to phosphorylate and activate the transcription factor c-Jun, which participates in the induction of some cancer types. However, the deletion of *jnk* genes in various mouse cancer models have shown that in some cases JNK acts as a tumor suppressor. Similarly to JNK, dysregulation of the p38 cascade may have a dual role in cancer, as it can either mediate cell survival or cell death. Interestingly, its activation levels in patients were found to be mainly associated with advanced stages of the disease, but its role in improving prognosis has been reported as well. Finally, the ERK5 cascade has been reported to play an inducing role in cancer, and several inhibitors of ERK5 are currently used in an attempt to combat various cancer types.

MAPKs in Developmental Diseases

Germline mutations of components of the MAPK cascades, in particular, the ERK1/2 one induces various developmental diseases. The best-known group of such syndromes is the RASopathies (Rauen, 2013). This is a group of genetic diseases that include, among others, the Noonan, Costello and cardio-facio-cutaneous (CFC) syndromes. These congenital syndromes are characterized by a variety of defects that include craniofacial dysmorphism, cardiac malformations, cutaneous and skeletal defects, and more. Although the prevalence of each disease may vary significantly, it was recently estimated that the rate of all the syndromes is about 1 in 1000

individuals. This demarcates RASopathies as one of the largest group of neurodevelopmental and malformation syndromes, making them attractive targets for research and drug development. Mutations linked to RASopathies have been identified in many signaling components upstream the ERK1/2 cascade, mainly Ras, as well as several mutations in B/C-Raf and MEK1/2. However, the basic mechanisms underlying their onset, phenotypic consequences and chronic nature are not fully understood as yet.

Germline mutations have been shown to affect the JNK cascade as well. One such mutation is in the CYLD gene that has been described in the Brooke–Spiegler syndrome (BSS), which is the autosomal dominant predisposition to skin appendageal neoplasms. CYLD encodes a deubiquitinating (DUB) enzyme that negatively regulates the nuclear factor (NF)- κ B pathway and the JNK cascade. The mutants of CYLD stabilize and prolong the signaling through this cascade, most likely leading to the syndrome. Although no germline mutation has been reported for the p38 cascade, the activity of the latter was reported to be important in inducing a variety of syndromes including the skull abnormalities of Apert syndrome, Ataxia-telangiectasia and Rad3 (ATR)-related Seckel syndrome, which is associated with growth retardation and the Nakajo-Nishimura syndrome. ERK5 has not been reported to be involved in any genetic syndrome as yet.

MAPKs in Inflammation-Related Diseases

The various MAPK cascades, particularly the p38 one, play a role in the induction and regulation of various inflammation-related diseases (Kyriakis and Avruch, 2012). Thus, the p38 cascade has been implicated in bowel inflammation, including ulcerative colitis and Crohn's disease. Most of the effects of p38 in these diseases are mediated by MAPKAPK2 and activation of transcription factors. In addition, activation of p38 is involved in both asthma and COPD, which have chronic inflammation. In these diseases, the inflammatory responses are induced by cigarette smoke or others cues and results in the induction and release of proinflammatory cytokines/chemokines, and activation of inflammatory cell migration. These effects are all mediated by p38 in epithelial cells and macrophages, which also may underlie the corticosteroid insensitivity of severe asthma and COPD. Other inflammation-related diseases associated with the activation of p38 are degenerative diseases of the central nervous system, inflammation-related cancer, rheumatoid arthritis, multiple sclerosis, and more. Therefore, p38 inhibitors present an attractive treatment for these conditions. Currently p38 inhibitors have been disappointing in the treatment of patients, mainly due to their severe side effects such as liver toxicity, indicating that the use of these inhibitors might not be possible for most diseases. Aside from the p38, the JNK cascade is involved in a multitude of inflammation-related diseases, such as cancer, bowel diseases, neurological, and immunological conditions. However, the involvement of the JNK cascade in these diseases seems to be only secondary to that of p38. The ERK1/2 and ERK5 cascades seem to have a much smaller involvement in these diseases. Therefore, unlike the p38 inhibitors, the specific inhibitors of JNK, as well as ERK1/2 and ERK5 are currently not considered as beneficial drugs for inflammation-related diseases.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: Receptor Tyrosine Kinases and Their Ligands. Cell Communication: Intracellular Pathways: The PI3K/Akt/mTOR Pathway; The PLC Pathway. Dynamic Integration: Dynamics: Dynamics of Protein Kinase Cascades

References

- Bogoyevitch, M.A., Ngoei, K.R., Zhao, T.T., Yeap, Y.Y., Ng, D.C., 2010. c-Jun N-terminal kinase (JNK) signaling: Recent advances and challenges. *Biochimica et Biophysica Acta* 1804, 463–475.
- Cargnello, M., Tcherkezian, J., Dorn, J.F., *et al.*, 2012. Phosphorylation of the eukaryotic translation initiation factor 4E-transporter (4E-T) by c-Jun N-terminal kinase promotes stress-dependent P-body assembly. *Molecular and Cellular Biology* 32, 4572–4584.
- Fey, D., Croucher, D.R., Kolch, W., Kholodenko, B.N., 2012. Crosstalk and signaling switches in mitogen-activated protein kinase cascades. *Frontiers in Physiology* 3, 355.
- Flores, K., Seger, R., 2013. Stimulated nuclear import by beta-like importins. *F1000Prime Reports* 5, 41.
- Haeusgen, W., Herdegen, T., Waetzig, V., 2011. The bottleneck of JNK signaling: Molecular and functional characteristics of MKK4 and MKK7. *European Journal of Cell Biology* 90, 536–544.
- Keshet, Y., Seger, R., 2010. The MAP kinase signaling cascades: A system of hundreds of components regulates a diverse array of physiological functions. *Methods in Molecular Biology* 661, 3–38.
- Kim, E.K., Choi, E.J., 2010. Pathological roles of MAPK signaling pathways in human diseases. *Biochimica et Biophysica Acta* 1802, 396–405.
- Kyriakis, J.M., Avruch, J., 2012. Mammalian MAPK signal transduction pathways activated by stress and inflammation: A 10-year update. *Physiological Reviews* 92, 689–737.
- Lawrence, M.C., Jivan, A., Shao, C., *et al.*, 2008. The roles of MAPKs in disease. *Cell Research* 18, 436–442.
- Matalanas, D., Birtwistle, M., Romano, D., *et al.*, 2011. Raf family kinases: Old dogs have learned new tricks. *Genes Cancer* 2, 232–260.
- McKay, M.M., Freeman, A.K., Morrison, D.K., 2011. Complexity in KSR function revealed by Raf inhibitor and KSR structure studies. *Small GTPases* 2, 276–281.
- Mendoza, M.C., Er, E.E., Blenis, J., 2011. The Ras-ERK and PI3K-mTOR pathways: Cross-talk and compensation. *Trends in Biochemical Sciences* 36, 320–328.
- Nunes-Xavier, C., Roma-Mateo, C., Rios, P., *et al.*, 2011. Dual-specificity MAP kinase phosphatases as targets of cancer treatment. *Anti-cancer Agents in Medicinal Chemistry* 11, 109–132.
- Plotnikov, A., Zehorai, E., Procaccia, S., Seger, R., 2011. The MAPK cascades: Signaling components, nuclear roles and mechanisms of nuclear translocation. *Biochimica et Biophysica Acta* 1813, 1619–1633.
- Pullikuth, A.K., Catling, A.D., 2007. Scaffold mediated regulation of MAPK signaling and cytoskeletal dynamics: A perspective. *Cellular Signalling* 19, 1621–1632.
- Rauen, K.A., 2013. The RASopathies. *Annual Review of Genomics and Human Genetics* 14, 355–369.
- Roberts, O.L., Holmes, K., Muller, J., Cross, D.A., Cross, M.J., 2009. ERK5 and the regulation of endothelial cell function. *Biochemical Society Transactions* 37, 1254–1259.
- Roskoski Jr., R., 2012. ERK1/2 MAP kinases: Structure, function, and regulation. *Pharmacological Research* 66, 105–143.
- Shaul, Y.D., Seger, R., 2007. The MEK/ERK cascade: From signaling specificity to diverse functions. *Biochimica et Biophysica Acta* 1773, 1213–1226.
- Shiryaev, A., Moens, U., 2010. Mitogen-activated protein kinase p38 and MK2, MK3 and MK5: Menage a trois or menage a quatre? *Cellular Signalling* 22, 1185–1192.
- Witzel, F., Maddison, L., Bluthgen, N., 2012. How scaffolds shape MAPK signaling: What we know and opportunities for systems approaches. *Frontiers in Physiology* 3, 475.
- Wortzel, I., Seger, R., 2011. The ERK cascade: Distinct functions within various subcellular organelles. *Genes Cancer* 2, 195–209.
- Yao, Z., Seger, R., 2004. The molecular mechanism of MAPK/ERK inactivation. *Current Genomics* 5, 385–393.
- Yoon, S., Seger, R., 2006. The extracellular signal-regulated kinase: Multiple substrates regulate diverse cellular functions. *Growth Factors* 24, 21–44.

The PI3K/Akt/mTOR Pathway

C Evangelisti, National Research Council, Bologna, Italy and Rizzoli Orthopedic Institute, Bologna, Italy
AM Martelli, University of Bologna, Bologna, Italy

© 2016 Elsevier Inc. All rights reserved.

Brief History

Identification of the phosphatidylinositol 3'-kinase (PI3K)/Akt/mammalian target of rapamycin (mTOR) pathway began in the mid-1980s. Since its discovery, interest among the scientific community and the quantity of research done in this pathway have been remarkable. Over the last 20 years, an enormous amount of evidence has demonstrated that PI3K/Akt/mTOR is a key signaling pathway involved in many physiological and pathophysiological processes.

In 1988, Cantley and colleagues identified a phosphatidylinositol (PtdIns) 3-kinase able to phosphorylate the 3-OH group of the inositol ring of PtdIns *in vitro*, yielding the PtdIns-3,4,5-trisphosphate (PtdIns3,4,5P3) (Whitman *et al.*, 1988).

Scientific attention to this pathway increased in 1991 through the development of a natural product with antifungal and immunosuppressive activity derived from the bacterium *Streptomyces hygroscopicus*. This compound, called rapamycin (or sirolimus), showed inhibitory activity against mTOR (Heitman *et al.*, 1991). Following studies discovered several functions of mTOR in higher eukaryotes, making this protein kinase to gain its central position in the signaling network that regulates cell growth and survival.

In 1991, three independent groups discovered a gene that encoded a serine/threonine (Ser/Thr) kinase, corresponding to Akt (Bellacosa *et al.*, 1991; Jones *et al.*, 1991; Coffey and Woodgett, 1991). Four years later, two independent groups highlighted that Akt was activated by growth factors in a PI3K-dependent manner (Burgering and Coffey, 1995; Franke *et al.*, 1995). The exact mechanism of this phospho-regulation was demonstrated in 1996 by Alessi and colleagues, who showed that activation of Akt required a double phosphorylation event on residues Thr 308 and Ser 473 (Alessi *et al.*, 1996). A year later, the same group identified a specific kinase that phosphorylated the Thr 308 residue, which they referred to as phosphoinositide-dependent kinase 1 (PDK1) (Alessi *et al.*, 1997). Then, in 2005, an mTOR complex 2 (mTORC2) was identified as the kinase complex that phosphorylates the other amino acidic residue necessary for full Akt activation (Sarbasov *et al.*, 2005).

PI3K

PI3Ks catalyze the phosphorylation of the 3-OH group of the inositol ring in one or more PtdIns lipid substrates. In turn, the PI3K products activate a number of cytoplasmic effectors, mostly by recruiting them to the plasma or other organelle membrane.

There are eight closely related mammalian PI3Ks, classified into three groups (I–III), based on their substrate preference and structural homology. All isoforms display the so-called PI3K core, consisting of C2, helical, and catalytic domains (Figure 1).

Class I PI3K

Class I PI3Ks preferentially phosphorylate PtdIns-4,5-bisphosphate (PtdIns4,5P2) to yield PtdIns3,4,5P3. PtdIns3,4,5P3 enables the recruitment to the plasma membrane and activation of downstream proteins that contain lipid-binding domains, such as the pleckstrin homolog (PH) domain and the phox homolog (PX) domain. Among others, both PDK1 and Akt are recruited to the plasma membrane by PtdIns3,4,5P3. Class I PI3Ks consist of two subgroups, depending on the receptor to which they match. Class IA PI3Ks are heterodimeric lipid enzymes comprising a p110 catalytic subunit (α , β , and δ , encoded by the *PIK3CA*, *PIK3CB*, and *PIK3CD* genes, respectively) and an adaptor/regulatory subunit (p85 α , p85 β , and p55 γ , encoded by *PIK3R1*, *PIK3R2*, and *PIK3R3* genes, respectively, or p55 α and p50 α , which are splice variants of *PIK3R1* gene). The single class IB PI3K comprises a p110 γ catalytic subunit which binds one of two related regulatory subunits, p101 and p84/p87 (Cantley, 2002).

Class IA PI3Ks can be activated by direct interaction with RTKs or GPCRs, whereas single class IB enzymes are exclusively activated by GPCRs. The regulatory subunits mediate membrane localization, receptor binding, and activation, whereas the catalytic subunit transfers phosphate groups to the PtdIns3,4,5P3. The different catalytic subunits display different patterns of expression and function. p110 α and p110 β PI3Ks are expressed in every mammalian cell type, whereas p110 γ and p110 δ PI3Ks are mainly expressed in leukocytes (So and Fruman, 2012).

Class II PI3K

Class II PI3Ks are monomers of high molecular mass that yield PtdIns-3-phosphate (PtdIns3P) and PtdIns3,4P2 from PtdIns and PtdIns-4-phosphate (PtdIns4P), respectively (Engelman *et al.*, 2006).

Mammal class II PI3K comprises three isoforms: PI3KC2 α , PI3KC2 β , and PI3KC2 γ . In humans, PI3KC2 α is expressed ubiquitously, and it is highly enriched in the heart, placenta, and ovary; PI3KC2 β is expressed in the placenta and thymus, whereas PI3KC2 γ is found in the liver, prostate, breast, and salivary glands (Falasca and Maffucci, 2012). All class II PI3K isoforms have a long C-terminus containing a PX and additional C2 domains (Vanhaesebroeck *et al.*, 2010). Class II PI3Ks do not interact with regulatory subunits, but the C- and N-terminus can mediate the regulation of their activation and enzymatic activity (Falasca and Maffucci, 2007).

Recently, it has been reported that class II PI3Ks are downstream targets of both activated RTKs and GPCRs (Falasca and Maffucci, 2012). PI3KC2 α and PI3KC2 β can be activated by a large number of growth factors, chemokines, and hormones such as insulin (Falasca and Maffucci, 2012). Moreover, class II PI3Ks generate distinct lipid products from

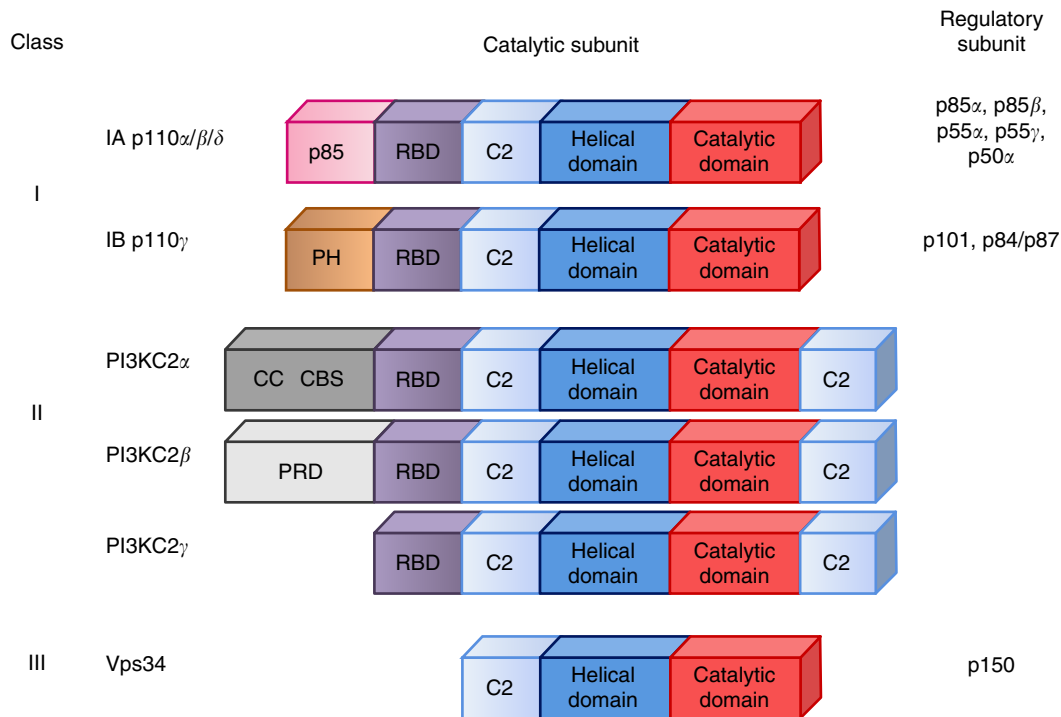


Figure 1 PI3K class domain structure: PI3Ks are divided into three classes based on their structural and biochemical features. All the PI3K isoforms have a catalytic domain PI3K core structure composed of a C2 domain, a helical domain, and a catalytic domain. C2, PKC homology 2 (-like) domain; CBS, clathrin-binding site; CC, coiled-coil domain; p85, p85-binding domain; PH, pleckstrin homolog domain; PRD, proline-rich domain; PtdIns, phosphatidylinositol; PtdIns4P, PtdIns-4-phosphate; PtdIns4,5P2, PtdIns-4,5-bisphosphate; RBD, ras-binding domain; Vps34, vacuolar protein sorting 34.

class I PI3Ks, and, by doing so, they activate different downstream effectors. For example, class II PI3Ks seem not to be involved in Akt activation.

Class III PI3K

The single class III isoform, termed vacuolar protein sorting 34 (Vps34), exclusively utilizes PtdIns as substrate to yield PtdIns3P and plays an important role in controlling vesicular protein sorting and autophagy. Vps34 is an endosomal kinase that can acquire multiple functions, forming a core complex with two other protein subunits: VPS15 and Beclin 1. The protein kinase VPS15 has been described as a regulatory protein, but its precise role in Vps34 regulation is still not completely defined (Backer, 2008). Beclin 1 has a central role in autophagy and participates in processes such as development, tumorigenesis, and neurodegeneration (Levine and Kroemer, 2008).

Akt

Akt, also termed protein kinase B (PKB), is the main downstream target of class I PI3Ks. Akt is a 57-kDa member of the AGC protein kinase family, which is the mammalian homolog of the *v-akt* oncogene. In mammals, there are three genes, located on different chromosomes, encoding for Akt1/α, Akt2/β, and Akt3/γ (Figure 2). Akt1 and Akt2 are present in every tissue (Hanada et al., 2004), whereas Akt3 is expressed in the

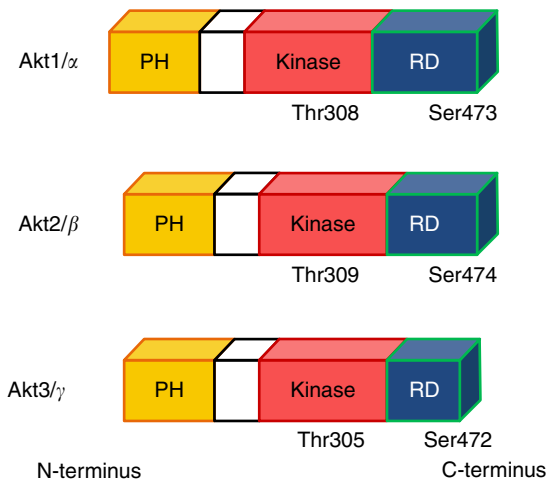


Figure 2 Akt isoform domain structure. All the Akt isoforms have a catalytic domain (kinase) in the central region of the molecule. The PH domain is located at the N-terminus and binds PtdIns3,4,5P3 t-membrane surfaces. The RD is at the C-terminus. The two phosphorylation sites necessary for Akt isoform activation are shown.

brain, kidney, and heart (Masure et al., 1999). Akt isoforms share a high degree of sequence homology in their catalytic domains, but diverge in other regions of the protein, i.e., the C-terminus regulatory domain (RD) and the N-terminus PH domain (Hanada et al., 2004). All three Akt isoforms bind to PtdIns3,4,5P3 through the PH domain, leading to

phosphorylation and activation of Akt itself (Scheid and Woodgett, 2003). The three isoforms contain similar phosphorylation sites: Thr 308 (Akt1)/Thr 309 (Akt2)/Thr 305 (Akt3) and Ser 473 (Akt1)/Ser 474 (Akt2)/Ser 472 (Akt3) (Hanada *et al.*, 2004). Threonine residues are within the catalytic domain, and their phosphorylation is mediated by PDK1 (Bayascas, 2010), whereas the serine residues are located in the RD and are phosphorylated by the mTORC2 or by other kinases, including integrin-linked kinase (ILK) or DNA-dependent protein kinase (DNA-PK) (Persad *et al.*, 2000; Figure 2). Full activation of Akt requires the phosphorylation of both residues; nevertheless, in the absence of Ser 473 phosphorylation, Thr 308 phospho-Akt can still phosphorylate some of its substrates (Guertin *et al.*, 2006). Once activated, Akt isoforms translocate to different subcellular compartments, including the endoplasmic reticulum, the mitochondria, the Golgi, and the nucleus, where they phosphorylate myriad downstream effectors regulating a variety of essential processes such as cell cycle progression, apoptosis, translation, glycolysis, and angiogenesis (Martelli *et al.*, 2012). Akt substrates often contain the phospho-Akt-substrate consensus motif (R-X-R-X-X-S/T-B), where X represents any amino acid and B represents bulky hydrophobic residues. Akt may either positively or negatively regulate the functions of its substrates, alter their subcellular localization, or modify their stability.

Akt enhances cell survival by inhibiting pro-apoptotic proteins and processes. Indeed, Akt negatively regulates the transcriptional activity of FoxO3a (Forkhead box O3), resulting in a decrease of expression of different pro-apoptotic genes such as caspase 9, caspase 3, Bim, Bad, Noxa, or Puma (You *et al.*, 2006). Moreover, Akt controls the activity of other transcription factors such as cyclic AMP response element-binding protein (CREB), the phosphorylation of which activates several anti-apoptotic genes such as Bcl-2 or Mcl-1 (Du and Montminy, 1998).

Other targets regulated by Akt and involved in the promotion of cell survival include the inhibitor of kappa B kinase- α (IKK α), which phosphorylates and activates the degradation of the inhibitory cofactor of NF- κ B (nuclear factor- κ B), I- κ B, inducing its degradation and the release of NF- κ B, which translocates to the nucleus, where it activates the transcription of pro-survival genes (Ozes *et al.*, 1999).

Another target of Akt is the murine double minute 2 (MDM2), an E3 ubiquitin ligase that negatively regulates p53 expression, the major cell death effector (Oren, 2003). Phosphorylation of MDM2 at Ser 166 and Ser 186 by Akt decreases MDM2 self-ubiquitination and renders the protein more stable. In this way, p53-mediated apoptosis is inhibited (Feng *et al.*, 2004).

Concerning the modulation of cell proliferation, Akt can phosphorylate and inactivate p21Waf1/Cip1 and p27Kip1, two members of the Cip/Kip family of cyclin-dependent kinase (CDK) inhibitors, which are required for cell cycle arrest in G₁.

Moreover, Akt can phosphorylate and then inhibit glycogen synthase kinase (GSK) 3 β . GSK3 β phosphorylates cyclin D1 at Thr 286 (Diehl *et al.*, 1998) and Myc at Thr 58 (Gregory *et al.*, 2003), leading to their nuclear export and degradation through a ubiquitin-mediated pathway. Consequently,

Akt-mediated GSK3 β inhibition results in cell cycle progression. Two other Akt substrates that play an important role in the regulation of cell cycle are Wee1 (Katayama *et al.*, 2005) and Myt1 (Okumura *et al.*, 2002). Both Wee1 and Myt1 are phosphorylated and inactivated by Akt, and these phosphorylative events facilitate cell cycle progression at the G₂/M phase transition.

As mentioned before, Akt also plays a key role in promoting cell growth, mainly through the activation of the mTOR complex 1 (mTORC1). Akt is therefore both an upstream activator of mTORC1 and a downstream effector of mTORC2. Akt phosphorylates 200-kDa tuberous sclerosis 2 (TSC2 or hamartin), leading to disruption of its interaction with tuberous sclerosis 1 (TSC1 or tuberlin) and to its degradation by the proteasome. TSC2 is a GTP-ase-activating protein (GAP) that associates with TSC1 to inactivate the small G protein Ras homolog enriched in the brain (Rheb), which accumulates in a GTP-bound state, which activates mTORC1. Phosphorylation by Akt negatively regulates the TSC1–TSC2 complex, leading to activation of mTOR (Inoki *et al.*, 2002). Moreover, the proline-rich Akt substrate of 40 kDa (PRAS40) inhibits mTORC1 activity. Phosphorylation by Akt represses the inhibitory function of PRAS40 on mTORC1 (Wang *et al.*, 2012).

mTOR

mTOR is a 289 kDa Ser/Thr kinase which belongs to the phosphatidylinositol kinase-related (PIKK) family (Memmott and Dennis, 2009) and is conserved throughout evolution. mTOR controls various cellular processes including cell cycle progression, survival, translation, metabolism, motility, autophagy, and aging. It collects input from upstream growth factor receptors, such as the PI3K/Akt, Ras/Raf/mitogen-activated protein kinase (MEK)/extracellular signal-regulated kinase (ERK), and liver kinase 1 (LKB1)/AMP-activated protein kinase (AMPK) pathways, in order to regulate several physiological events (Figure 3). In addition, it has recently been reported that mTORC1 also responds to inputs via the Hippo, WNT, and Notch signaling pathways (Shimobayashi and Hall, 2014).

mTOR is the catalytic subunit of two distinct complexes known as mTORC1 and mTORC2, which differ in partner proteins and substrate specificity. The two complexes have different functions: while mTORC1 positively controls cell survival, cell growth, and protein synthesis by promoting many anabolic processes and by limiting catabolic processes such as autophagy, mTORC2 regulates cytoskeletal rearrangement, cell survival, cell-cycle progression, and metabolism (Figure 3).

mTORC1 is composed of mTOR/regulatory-associated protein of mTOR (Raptor, a scaffolding protein)/mammalian Lethal-with-Sec-Thirteen 8 (mLST8)/PRAS40/FK-506-binding protein 38 (FKBP38)/DEP-domain-containing mTOR interacting protein (Deptor), and is sensitive to rapamycin and its derivatives (rapalogs). Different exogenous stimuli regulate mTORC1 activity, including growth factor signaling such as insulin and type I insulin-like growth factor (IGF-1), the energy state of the cell (that is, AMP levels), nutrient (e.g., amino

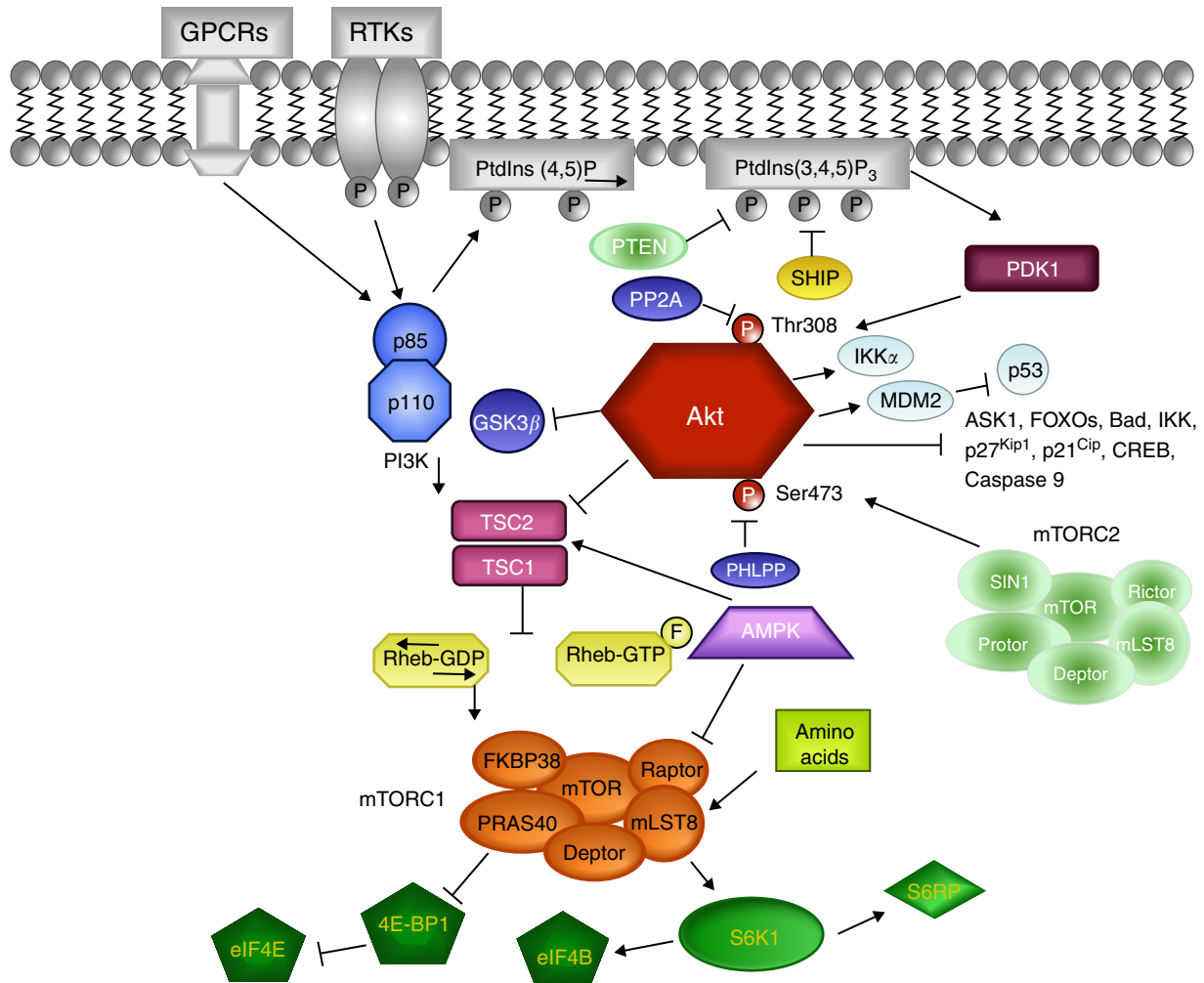


Figure 3 The PI3K/Akt/mTOR pathway activation. RTKs or GPCRs stimulate PI3K activity. Activated PI3K phosphorylates PtdIns4,5P2 to yield PtdIns3,4,5P3. PtdIns3,4,5P3 activity can be negatively regulated by PTEN and SHIP phosphatases. PtdIns3,4,5P3 recruits to the plasma membrane PDK1, which phosphorylates Akt at Thr 308. Full Akt activation requires Ser 473 phosphorylation by mTORC2. Akt activity can be negatively regulated by PP2A and PHLPP phosphatases. Active Akt inhibits TSC2 activity through direct phosphorylation. TSC2 is a GAP that functions in association with TSC1 to inactivate the small G protein Rheb. Akt-driven TSC1–TSC2 complex inactivation allows Rheb to accumulate in a GTP-bound state. Rheb-GTP then activates the protein kinase activity of mTORC1. mTORC1 targets S6K1, 4E-BP1, and S6RP. The phosphorylation of 4E-BP1 leads to the release of eIF4E, which interacts with eIF4G to activate translation initiation. However, other signaling cascades impinge on mTORC1, including GSK3 β , and the AMPK network which is sensitive to the ADP/ATP ratio. Arrows indicate activating events, while perpendicular lines indicate inhibitory events.

acids) and O₂ availability, and stress signals (Laplanche and Sabatini, 2009).

Regarding mTORC1 functions, protein synthesis is the best-characterized process controlled by this complex (Sengupta et al., 2010). When active, mTORC1 phosphorylates the translational regulator eukaryotic translation initiation factor 4E (eIF4E)-binding protein 1 (4E-BP1) and S6 kinase 1 (S6K1) (Thoreen et al., 2012). The phosphorylation of 4E-BP1 leads to the release of eIF4E, which interacts with eIF4G to activate translation initiation (Martelli et al., 2011).

Then, S6K1 phosphorylates the 40S ribosomal protein S6 (S6RP), leading to active translation, while 4E-BP1 interacts with the eukaryotic initiation factor 4B (eIF4B), which critically regulates cap-dependent translation (Browne and Proud, 2004). S6K1 activation also leads to increased mRNA

biogenesis, as well as translational initiation and elongation (Chauvin et al., 2014). S6K1 has several substrates, including insulin receptor substrate 1 (IRS1), which is upstream of mTORC1. Importantly, phosphorylation of IRS1 by S6K1 targets IRS1 for proteosomal degradation and thereby hampers the ability of growth factors (insulin, IGF-1) to signal downstream of RTKs (Laplanche and Sabatini, 2012). This results in inhibition of PI3K/Akt/mTOR activation, creating a negative feedback loop, which has an important role in the regulation of mTORC1.

Through the phosphorylation of various other targets, mTORC1 can trigger metabolic changes such as mitochondrial biogenesis, promotion of glycolysis, lipid biogenesis, and suppression of autophagy (Laplanche and Sabatini, 2009).

It has been reported that mTORC1 plays an important role in the maintenance of the mitochondrial oxidative function,

by regulating the interactions between the transcription factor yin-yang 1 (YY1) and the peroxisomal proliferator-activated receptor γ (PPAR γ) coactivator 1 (PGC-1), thereby preventing the coactivation of YY1 (Cunningham *et al.*, 2007).

Recently, an involvement of mTORC1 in lipid synthesis has been described. Lipid synthesis is required for proliferating cells to generate membranes, acting through the sterol regulatory element-binding protein 1/2 (SREBP1/SREBP2), which are transcription factors of lipogenic genes, as well as through PPAR γ , which are necessary and sufficient for the differentiation of preadipocytes and lipid accumulation (Laplane and Sabatini, 2012).

Furthermore, mTORC1 suppresses autophagy, a complex catabolic process that sustains cellular metabolism through recycling of cellular components during growth-unfavorable conditions. mTORC1, under stress conditions, transduces anti-autophagic signals by binding to the UNC51-like kinase (ULK) multi-protein complex, which is essential for the initial formation of the phagophore. mTORC1 can negatively regulate autophagy directly through phosphorylation of both ULK1, at Ser 638 and Ser 757, and ATG13 (Ganley *et al.*, 2009), or indirectly via TSC2 (Kim *et al.*, 2011).

mTORC2 is composed of mTOR/rapamycin-insensitive companion of mTOR (Rictor)/mLST8/stress-activated protein kinase-interacting protein 1 (SIN1)/protein observed with Rictor (Protor)/Deptor. mTORC2 is mainly activated by growth factors and insulin and seems to be insensitive to amino acids. mTORC2 regulates different downstream AGC kinases such as Akt, serum-, and glucocorticoid-stimulated kinase (SGK) and protein kinases C (PKCs). Moreover, mTORC2 controls cellular survival and spatial control of cell growth via cytoskeleton regulation, by stimulating of actin fibers, paxillin, RhoA, Rac1, and PKC α (Oh and Jacinto, 2011).

Regulation of the PI3K/Akt/mTOR Pathway

The regulation of the PI3K/Akt/mTOR axis is extremely complex, due to the existence of a very intricate network of negative feedback loops.

mTORC1 can regulate the PI3K/Akt/mTOR pathway at different levels. First, mTORC1 elicits a negative feedback that antagonizes the formation of mTORC2, thus lowering Akt activity (Sarbasov *et al.*, 2005). Therefore, inhibition of mTORC1 may result in Akt hyperactivation. Furthermore, when active, mTORC1 phosphorylates S6K1, which can induce the proteosomal degradation of IRS1, by phosphorylating it at multiple sites, resulting in a decrease of PI3K/Akt/mTORC1 activation (Bhaskar and Hay, 2007). mTORC1 can also down-regulate IRS2 expression by enhancing its proteosomal degradation (Bhaskar and Hay, 2007).

Furthermore, the existence of a rapamycin-sensitive, mTORC1/S6K1-mediated phosphorylation of Rictor on Thr 1135 has been documented. This phosphorylative event exerted a negative regulatory effect on the mTORC2-dependent phosphorylation of Akt *in vivo* (Boulbes *et al.*, 2010).

Moreover, it has been demonstrated that FoxO1, in a TSC2-dependent way, can activate AMPK, which in turn directly inhibits mTORC1 and activates Akt, suppressing the S6K1 feedback (Chen *et al.*, 2010). TSC1/2, in addition to its role in

the inhibition of mTORC1, can also regulate mTORC2 functioning. Indeed, the TSC1–TSC2 complex interacts with mTORC2, leading to its activation (Huang *et al.*, 2008).

Direct association between mTORC2 and ribosomes can activate mTORC2, and this event is regulated by PI3K activity (Zinzalla *et al.*, 2011). This mechanism suggests a possible link between both mTOR complexes in which mTORC2 promotes mTORC1 activity through the PI3K/Akt pathway and, in return, mTORC1 controls mTORC2 activity through the regulation of ribosomal biogenesis.

The PI3K/Akt/mTOR signaling cascade also cross talks with the p53 pathway. Indeed, p53 can regulate both TSC2 and AMPK (Feng *et al.*, 2007). Furthermore, AMPK may phosphorylate p53 at Ser 15, which leads to stabilization of p53 (Jones *et al.*, 2005). Hence, p53 and AMPK positively regulate each other. Moreover, phosphatase and tensin homolog (PTEN) may interact with p53, leading to its stabilization (Freeman *et al.*, 2003). As mentioned above, Akt is a negative regulator of p53 through MDM2. These cross regulations of AMPK, PTEN, and p53 oppose Akt activity and constitute a network coordinating cell division and growth after sensing nutrient availability, DNA integrity, and oncogenic activity.

Negative Regulation of PI3K/Akt/mTOR Pathway

The activity of several phosphatases counterbalances PI3K/Akt/mTOR activation, thus switching off this signaling cascade (Figure 3).

PI3Ks inactivation

The main phosphatases that can block the PI3K activity are PTEN, Src-homology 2 (SH2)-containing inositol 50-phosphatase (SHIP), and inositol polyphosphate 4-phosphatase type II (INPP4B), which hydrolyzes the 4-position of PtdIns(3,4)P₂.

PTEN is a ubiquitous phosphatase that can dephosphorylate the 3-position of PtdIns(3,4,5)P₃ to yield PtdIns(4,5)P₂ (Maehama and Dixon, 1998). Loss of PTEN activity due to inactivating mutations, deletions, or silencing has been reported in a wide range of both familial and sporadic human cancers (Sansal and Sellers, 2004).

The family of SHIP phosphatases includes two members, SHIP-1 and SHIP-2. These phosphatases dephosphorylate the 5-position of the inositol ring of PtdIns(3,4,5)P₃ to yield PtdIns3,4P₂ (Kalesnikoff *et al.*, 2003). While the SHIP-1 isoform is expressed in endothelial cells and in the hematopoietic lineage, SHIP-2 is ubiquitously expressed.

Akt inactivation

Akt function is also inactivated through the action of different phosphatases.

The phosphorylated Thr residues of Akt are dephosphorylated directly by protein phosphatase 2A (PP2A) (Ruvolo *et al.*, 2011). The core enzyme of PP2A is a dimer consisting of a catalytic subunit (PP2A/C α or β) and a regulatory/structural A subunit (PP2A/A α or β). A third regulatory B subunit (PP2A/B), which determines substrate specificity, can be associated with this core structure (Eichhorn *et al.*, 2009).

The phosphorylated Ser residues of Akt are targeted by the PH domain leucine-rich repeat protein phosphatase (PHLPP) family of isozymes. Three PHLPP isoforms have been identified so far: the alternatively spliced PHLPP1 α and PHLPP1 β , which target Akt2 and Akt3, and PHLPP2, which dephosphorylates Akt1 and Akt3 (Brognard *et al.*, 2007).

PI3K/Akt/mTOR Signaling in Human Disorders

Considering the crucial functions played by the PI3K/Akt/mTOR signaling pathway in several physiological processes, it appears evident that dysregulation of this axis could have dramatic consequences. Indeed, alterations in the PI3K/Akt/mTOR pathway are found in neurodegenerative diseases, cardiac hypertrophy, and metabolic disorders. Moreover, aberrantly activated PI3K/Akt/mTOR cascades are commonly observed in an increasing number of both solid tumors and hematological malignancies (Beauchamp and Platanias, 2013).

Regarding neurodegenerative diseases, it has been demonstrated that this pathway is involved in neuroprotection (Timmons *et al.*, 2009), enhancing cell survival via stimulation of cell proliferation and inhibition of apoptosis and autophagy (Heras-Sandoval *et al.*, 2014).

Recent evidence suggests that defective PI3K/Akt/mTOR signaling is associated with loss of dopaminergic neurons in Parkinson's disease (PD); in addition, a significant decrease in phospho-Akt at the Ser 473 residue has been reported in tyrosine hydroxylase TH immunoreactive dopaminergic neurons in the brains of PD patients (Timmons *et al.*, 2009).

Furthermore, this pathway appears to promote hyperphosphorylation of Tau, the main cause of formation of senile plaques and neurofibrillary tangles that characterize Alzheimer's disease (AD). In particular, GSK3 β , a downstream target of Akt, is probably the most documented kinase implicated in the abnormal hyper-phosphorylation of Tau protein (Hoover *et al.*, 2010). Moreover, alterations in the PI3K/Akt/mTOR pathway are linked to disruption of autophagy, and this event seems to be involved in the pathogenesis of AD and PD. Indeed, in both AD and PD, deficient elimination of abnormal and toxic protein aggregates promotes cellular stress and cell death. Therefore, induction of autophagy has been proposed as a reasonable strategy to help neurons in clearing abnormal protein aggregates and then to survive (Heras-Sandoval *et al.*, 2014).

Given the fundamental role of mTORC1 in the regulation of the autophagic process, beneficial outcomes can be achieved by modulating the PI3K/Akt/mTOR pathway: on the one hand, the induction of the above-mentioned cascade promotes survival and stops excessive autophagy that leads to cellular death; on the other hand, the inhibition of mTORC1 may stimulate autophagy of damaged or toxic proteins and then up-regulate cellular survival through mTORC2 activity on Akt.

Akt may also play a role in Huntington's disease (HD), as it phosphorylates huntingtin at Ser 421. Huntingtin phosphorylation is reduced at Ser 421 in HD patients (Warby *et al.*, 2005).

Another pathological condition linked to dysregulation of the PI3K/Akt/mTOR signaling axis is cardiac hypertrophy, one

of the main risk factors for heart failure. Cardiac hypertrophy is induced by increased cardiomyocyte volume, which is characterized by increased transcription and translation, as well as reduced protein turnover. The PI3K/Akt/mTOR pathway has been shown to directly regulate cardiomyocyte size, survival, angiogenesis, and inflammation in both physiological and pathological cardiac hypertrophy, through translational regulation by mTOR. Indeed, inhibition of mTORC1 by rapamycin treatment attenuated the hypertrophic response to pressure overload (McMullen *et al.*, 2004).

Insulin activates the PI3K/Akt signaling pathway in cardiomyocytes and substantially preserves and even improves regional and overall cardiac function. Insulin can protect cardiomyocytes from apoptosis by regulating a number of PI3K/Akt downstream signaling molecules, such as endothelial nitric oxide synthase (eNOS), GSK3 β , and mTORC1 (Yao and Han, 2014).

Type 2 diabetes is characterized by diminished pancreatic β -cell function, including decreased insulin secretion, as well as by peripheral insulin resistance in insulin-targeted tissues and organs. Insulin action elicits a series of signaling cascades initiated by insulin binding to its receptor, which causes receptor auto-phosphorylation and up-regulation of its RTK activity, resulting in tyrosine phosphorylation of IRSs. Phosphorylation of IRSs leads to activation of PI3K and, subsequently, of Akt and its downstream mediator, AS160, all of which are important steps in stimulating glucose transport induced by insulin. Studies on the skeletal muscle of type 2 diabetic patients have disclosed impaired insulin activation of the IRS1/PI3K/Akt signaling axis (Schultze *et al.*, 2012).

The involvement of PI3K/Akt/mTOR in tumorigenesis of multiple cancer types is well established (Katso *et al.*, 2001).

Dysregulation of the PI3K/Akt/mTOR pathway is more frequent than with any other pathway in human cancer (30–50% of all tumor types), with the possible exception of the p53 and retinoblastoma (Rb) pathways.

Moreover, PI3K/Akt/mTOR aberrant activity is not only found in many types of sporadic tumors but is also associated with familial cancer syndromes, such as Cowden's syndrome (CS), Peutz–Jegher's syndrome, tuberous sclerosis, and neurofibromatosis type-1 and type-2 (Carnero, 2010).

Pathologic activation of the PI3K/Akt/mTOR axis, leading to cancer onset, can occur at multiple levels and by various mechanisms, including PI3K/Akt point mutations, Akt amplification, and deletion/inactivation of tumor suppressor genes, including PTEN (Hennessy *et al.*, 2005).

Point mutations in PIK3CA, the gene encoding for the p110 α subunit of PI3K, are among the most commonly documented mutations in cancer (Samuels *et al.*, 2004). Indeed, these mutations have been described in nearly every major tumor type, including breast cancer, ovarian cancer, head and neck cancer, bowel cancer, gastric cancer, lung cancer, anaplastic oligodendrogliomas, anaplastic astrocytomas, glioblastoma multiforme, and medulloblastoma.

Mutations in PTEN have been identified in sporadic cancer; in particular, they have been found in endometrial, central nervous system, skin, and prostate cancers (Tamguney and Stokoe, 2007).

While somatic mutations are more common, germline loss of PTEN results in CS. CS is characterized by macrocephaly,

trichilemmomas, and papillomatous papules, as well as by a high risk of developing both benign and malignant tumors, especially of the thyroid, breast, and endometrium. CS is caused by PTEN monoallelic mutations that are sufficient to deregulate the PI3K/Akt/mTOR pathway, thus triggering carcinogenesis (Eng, 2003).

Furthermore, loss of PTEN activity could be due to deletions, transcriptional silencing, or protein instability due to posttranslational modification (Ali *et al.*, 1999).

In addition to inherent aberrations in members of the PI3K/Akt/mTOR pathway, pathologic signaling through this pathway can also occur by other means, including deregulation of RTKs, GPCRs, and oncogene, such as RAS (Chappell *et al.*, 2011).

Activation of the PI3K/Akt/mTOR pathway contributes to competitive growth advantage of cancer cells, metastatic competence, tumor angiogenesis, and, frequently, therapy resistance.

For all these reasons, PI3K/Akt/mTOR signaling is considered a valuable target for the development of innovative therapeutic strategies in cancer, and many small molecular inhibitors of this pathway are currently being tested in clinical trials.

Nevertheless, modulation of this signaling pathway could be beneficial to patients suffering from a wide variety of disorders other than cancer.

Highlights

- The PI3K/Akt/mTOR pathway plays a key role in many cellular processes, such as cell cycle progression, cellular survival, metabolism, and apoptosis.
- The PI3K/Akt/mTOR axis is extremely complex and tightly regulated.
- Dysregulation of PI3K/Akt/mTOR axis activity is associated with a wide range of pathological conditions.
- Targeting the PI3K/Akt/mTOR signaling cascade is considered a valuable and innovative therapeutic strategy for several disorders, including cancer.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: Receptor Tyrosine Kinases and Their Ligands

References

- Alessi, D.R., Andjelkovic, M., Caudwell, B., *et al.*, 1996. Mechanism of activation of protein kinase B by insulin and IGF-1. *EMBO Journal* 15, 6541–6551.
- Alessi, D.R., James, S.R., Downes, C.P., *et al.*, 1997. Characterization of a 3-phosphoinositide-dependent protein kinase which phosphorylates and activates protein kinase B α . *Current Biology* 7, 261–269.
- Ali, I.U., Schriml, L.M., Dean, M., 1999. Mutational spectra of PTEN/MMAC1 gene: A tumor suppressor with lipid phosphatase activity. *Journal of the National Cancer Institute* 91, 1922–1932.
- Backer, J.M., 2008. The regulation and function of class III PI3Ks: Novel roles for Vps34. *Biochemical Journal* 410, 1–17.
- Bayascas, J.R., 2010. PDK1: The major transducer of PI 3-kinase actions. *Current Topics in Microbiology and Immunology* 346, 9–29.
- Beauchamp, E.M., Platanias, L.C., 2013. The evolution of the TOR pathway and its role in cancer. *Oncogene* 32, 3923–3932.
- Bellacosa, A., Testa, J.R., Staal, S.P., Tsichlis, P.N., 1991. A retroviral oncogene, akt, encoding a serine-threonine kinase containing an SH2-like region. *Science* 254, 274–277.
- Bhaskar, P.T., Hay, N., 2007. The two TORCs and Akt. *Developmental Cell* 12, 487–502.
- Boulbes, D., Chen, C.H., Shaikenov, T., *et al.*, 2010. Rictor phosphorylation on the Thr-1135 site does not require mammalian target of rapamycin complex 2. *Molecular Cancer Research* 8, 896–906.
- Brogna, J., Sieracki, E., Gao, T., Newton, A.C., 2007. PHLPP and a second isoform, PHLPP2, differentially attenuate the amplitude of Akt signaling by regulating distinct Akt isoforms. *Molecular Cell* 25, 917–931.
- Browne, G.J., Proud, C.G., 2004. A novel mTOR-regulated phosphorylation site in elongation factor 2 kinase modulates the activity of the kinase and its binding to calmodulin. *Molecular and Cellular Biology* 24, 2986–2997.
- Burgering, B.M., Coffey, P.J., 1995. Protein kinase B (c-Akt) in phosphatidylinositol-3-OH kinase signal transduction. *Nature* 376, 599–602.
- Cantley, L.C., 2002. The phosphoinositide 3-kinase pathway. *Science* 296, 1655–1657.
- Carnero, A., 2010. The PKB/AKT pathway in cancer. *Current Pharmaceutical Design* 16, 34–44.
- Chappell, W.H., Steelman, L.S., Long, J.M., *et al.*, 2011. Ras/Raf/MEK/ERK and PI3K/PTEN/Akt/mTOR inhibitors: Rationale and importance to inhibiting these pathways in human health. *Oncotarget* 2, 135–164.
- Chauvin, C., Koka, V., Nusch, A., *et al.*, 2014. Ribosomal protein S6 kinase activity controls the ribosome biogenesis transcriptional program. *Oncogene* 33, 474–483.
- Chen, C.C., Jeon, S.M., Bhaskar, P.T., *et al.*, 2010. FoxOs inhibit mTORC1 and activate Akt by inducing the expression of Sestrin3 and Rictor. *Developmental Cell* 18, 592–604.
- Coffer, P.J., Woodgett, J.R., 1991. Molecular cloning and characterisation of a novel putative protein-serine kinase related to the cAMP-dependent and protein kinase C families. *European Journal of Biochemistry* 201, 475–481.
- Cunningham, J.T., Rodgers, J.T., Arlow, D.H., *et al.*, 2007. mTOR controls mitochondrial oxidative function through a YY1-PGC-1 α transcriptional complex. *Nature* 450, 736–740.
- Diehl, J.A., Cheng, M., Roussel, M.F., Sherr, C.J., 1998. Glycogen synthase kinase-3 β regulates cyclin D1 proteolysis and subcellular localization. *Genes & Development* 12, 3499–3511.
- Du, K., Montminy, M., 1998. CREB is a regulatory target for the protein kinase Akt/PKB. *Journal of Biological Chemistry* 273, 32377–32379.
- Eichhorn, P.J., Creighton, M.P., Bernards, R., 2009. Protein phosphatase 2A regulatory subunits and cancer. *Biochimica et Biophysica Acta* 1795, 1–15.
- Eng, C., 2003. PTEN: One gene, many syndromes. *Human Mutation* 22, 183–198.
- Engelman, J.A., Luo, J., Cantley, L.C., 2006. The evolution of phosphatidylinositol 3-kinases as regulators of growth and metabolism. *Nature Reviews Genetics* 7, 606–619.
- Falasca, M., Maffucci, T., 2007. Role of class II phosphoinositide 3-kinase in cell signalling. *Biochemical Society Transactions* 35, 211–214.
- Falasca, M., Maffucci, T., 2012. Regulation and cellular functions of class II phosphoinositide 3-kinases. *Biochemical Journal* 443, 587–601.
- Feng, J., Tamaskovic, R., Yang, Z., *et al.*, 2004. Stabilization of Mdm2 via decreased ubiquitination is mediated by protein kinase B/Akt-dependent phosphorylation. *Journal of Biological Chemistry* 279, 35510–35517.
- Feng, Z., Hu, W., de Stanchina, E., *et al.*, 2007. The regulation of AMPK β 1, TSC2, and PTEN expression by p53: Stress, cell and tissue specificity, and the role of these gene products in modulating the IGF-1-AKT-mTOR pathways. *Cancer Research* 67, 3043–3053.
- Franke, T.F., Yang, S.I., Chan, T.O., *et al.*, 1995. The protein kinase encoded by the Akt proto-oncogene is a target of the PDGF-activated phosphatidylinositol 3-kinase. *Cell* 81, 727–736.
- Freeman, D.J., Li, A.G., Wei, G., *et al.*, 2003. PTEN tumor suppressor regulates p53 protein levels and activity through phosphatase-dependent and -independent mechanisms. *Cancer Cell* 3, 117–130.
- Ganley, I.G., Lam du, H., Wang, J., *et al.*, 2009. ULK1/ATG13/FIP200 complex mediates mTOR signaling and is essential for autophagy. *Journal of Biological Chemistry* 284, 12297–12305.
- Gregory, M.A., Qi, Y., Hann, S.R., 2003. Phosphorylation by glycogen synthase kinase-3 controls c-myc proteolysis and subnuclear localization. *Journal of Biological Chemistry* 278, 51606–51612.
- Guertin, D.A., Stevens, D.M., Thoreen, C.C., *et al.*, 2006. Ablation in mice of the mTORC components raptor, rictor, or mLST8 reveals that mTORC2 is required for signaling to Akt-FOXO and PKC α , but not S6K1. *Developmental Cell* 11, 859–871.
- Hanada, M., Feng, J., Hemmings, B.A., 2004. Structure, regulation and function of PKB/AKT – A major therapeutic target. *Biochimica et Biophysica Acta* 1697, 3–16.

- Heitman, J., Movva, N.R., Hiestand, P.C., Hall, M.N., 1991. FK 506-binding protein proline rotamase is a target for the immunosuppressive agent FK 506 in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* 88, 1948–1952.
- Hennessy, B.T., Smith, D.L., Ram, P.T., Lu, Y., Mills, G.B., 2005. Exploiting the PI3K/AKT pathway for cancer drug discovery. *Nature Reviews Drug Discovery* 4, 988–1004.
- Heras-Sandoval, D., Perez-Rojas, J.M., Hernandez-Damian, J., Pedraza-Chaverri, J., 2014. The role of PI3K/AKT/mTOR pathway in the modulation of autophagy and the clearance of protein aggregates in neurodegeneration. *Cellular Signalling* 26, 2694–2701.
- Hoover, B.R., Reed, M.N., Su, J., et al., 2010. Tau mislocalization to dendritic spines mediates synaptic dysfunction independently of neurodegeneration. *Neuron* 68, 1067–1081.
- Huang, J., Dibble, C.C., Matsuzaki, M., Manning, B.D., 2008. The TSC1-TSC2 complex is required for proper activation of mTOR complex 2. *Molecular and Cellular Biology* 28, 4104–4115.
- Inoki, K., Li, Y., Zhu, T., Wu, J., Guan, K.L., 2002. TSC2 is phosphorylated and inhibited by Akt and suppresses mTOR signalling. *Nature Cell Biology* 4, 648–657.
- Jones, P.F., Jakubowicz, T., Pitossi, F.J., Maurer, F., Hemmings, B.A., 1991. Molecular cloning and identification of a serine/threonine protein kinase of the second-messenger subfamily. *Proceedings of the National Academy of Sciences of the United States of America* 88, 4171–4175.
- Jones, R.G., Plas, D.R., Kubek, S., et al., 2005. AMP-activated protein kinase induces a p53-dependent metabolic checkpoint. *Molecular Cell* 18, 283–293.
- Kalesnikoff, J., Sly, L.M., Hughes, M.R., et al., 2003. The role of SHIP in cytokine-induced signaling. *Reviews of Physiology, Biochemistry and Pharmacology* 149, 87–103.
- Katayama, K., Fujita, N., Tsuruo, T., 2005. Akt/protein kinase B-dependent phosphorylation and inactivation of WEE1Hu promote cell cycle progression at G2/M transition. *Molecular and Cellular Biology* 25, 5725–5737.
- Katso, R., Okkenhaug, K., Ahmadi, K., et al., 2001. Cellular function of phosphoinositide 3-kinases: Implications for development, homeostasis, and cancer. *Annual Review of Cell and Developmental Biology* 17, 615–675.
- Kim, J., Kundu, M., Viollet, B., Guan, K.L., 2011. AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1. *Nature Cell Biology* 13, 132–141.
- Laplane, M., Sabatini, D.M., 2009. mTOR signaling at a glance. *Journal of Cell Science* 122, 3589–3594.
- Laplane, M., Sabatini, D.M., 2012. mTOR signaling in growth control and disease. *Cell* 149, 274–293.
- Levine, B., Kroemer, G., 2008. Autophagy in the pathogenesis of disease. *Cell* 132, 27–42.
- Maehama, T., Dixon, J.E., 1998. The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. *Journal of Biological Chemistry* 273, 13375–13378.
- Martelli, A.M., Evangelisti, C., Chappell, W., et al., 2011. Targeting the translational apparatus to improve leukemia therapy: Roles of the PI3K/PTEN/Akt/mTOR pathway. *Leukemia* 25, 1064–1079.
- Martelli, A.M., Tabellini, G., Bressanin, D., et al., 2012. The emerging multiple roles of nuclear Akt. *Biochimica et Biophysica Acta* 1823, 2168–2178.
- Masure, S., Haefner, B., Wesselink, J.J., et al., 1999. Molecular cloning, expression and characterization of the human serine/threonine kinase Akt-3. *European Journal of Biochemistry* 265, 353–360.
- McMullen, J.R., Sherwood, M.C., Tarnavski, O., et al., 2004. Inhibition of mTOR signaling with rapamycin regresses established cardiac hypertrophy induced by pressure overload. *Circulation* 109, 3050–3055.
- Memmott, R.M., Dennis, P.A., 2009. Akt-dependent and -independent mechanisms of mTOR regulation in cancer. *Cellular Signalling* 21, 656–664.
- Oh, W.J., Jacinto, E., 2011. mTOR complex 2 signaling and functions. *Cell Cycle* 10, 2305–2316.
- Okumura, E., Fukuhara, T., Yoshida, H., et al., 2002. Akt inhibits Myt1 in the signalling pathway that leads to meiotic G2/M-phase transition. *Nature Cell Biology* 4, 111–116.
- Oren, M., 2003. Decision making by p53: Life, death and cancer. *Cell Death & Differentiation* 10, 431–442.
- Ozes, O.N., Mayo, L.D., Gustin, J.A., et al., 1999. NF-kappaB activation by tumour necrosis factor requires the Akt serine-threonine kinase. *Nature* 401, 82–85.
- Persad, S., Attwell, S., Gray, V., et al., 2000. Inhibition of integrin-linked kinase (ILK) suppresses activation of protein kinase B/Akt and induces cell cycle arrest and apoptosis of PTEN-mutant prostate cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* 97, 3207–3212.
- Ruvolo, P.P., Qui, Y.H., Coombes, K.R., et al., 2011. Low expression of PP2A regulatory subunit B55alpha is associated with T308 phosphorylation of AKT and shorter complete remission duration in acute myeloid leukemia patients. *Leukemia* 25, 1711–1717.
- Samuels, Y., Wang, Z., Bardelli, A., et al., 2004. High frequency of mutations of the PIK3CA gene in human cancers. *Science* 304, 554.
- Sansal, I., Sellers, W.R., 2004. The biology and clinical relevance of the PTEN tumor suppressor pathway. *Journal of Clinical Oncology* 22, 2954–2963.
- Sarbassov, D.D., Guertin, D.A., Ali, S.M., Sabatini, D.M., 2005. Phosphorylation and regulation of Akt/PKB by the rictor-mTOR complex. *Science* 307, 1098–1101.
- Scheid, M.P., Woodgett, J.R., 2003. Unravelling the activation mechanisms of protein kinase B/Akt. *FEBS Letters* 546, 108–112.
- Schultz, S.M., Hemmings, B.A., Niessen, M., Tschopp, O., 2012. PI3K/AKT, MAPK and AMPK signalling: Protein kinases in glucose homeostasis. *Expert Reviews in Molecular Medicine* 14, e1.
- Sengupta, S., Peterson, T.R., Sabatini, D.M., 2010. Regulation of the mTOR complex 1 pathway by nutrients, growth factors, and stress. *Molecular Cell* 40, 310–322.
- Shimobayashi, M., Hall, M.N., 2014. Making new contacts: The mTOR network in metabolism and signalling crosstalk. *Nature Reviews Molecular Cell Biology* 15, 155–162.
- So, L., Fruman, D.A., 2012. PI3K signalling in B- and T-lymphocytes: New developments and therapeutic advances. *Biochemical Journal* 442, 465–481.
- Tamguney, T., Stokoe, D., 2007. New insights into PTEN. *Journal of Cell Science* 120, 4071–4079.
- Thoreen, C.C., Chantranupong, L., Keys, H.R., et al., 2012. A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* 485, 109–113.
- Timmons, S., Coakley, M.F., Moloney, A.M., C, O.N., 2009. Akt signal transduction dysfunction in Parkinson's disease. *Neuroscience Letters* 467, 30–35.
- Vanhaesebroeck, B., Guillermet-Guibert, J., Graupera, M., Bilanges, B., 2010. The emerging mechanisms of isoform-specific PI3K signalling. *Nature Reviews Molecular Cell Biology* 11, 329–341.
- Wang, H., Zhang, Q., Wen, Q., et al., 2012. Proline-rich Akt substrate of 40 kDa (PRAS40): A novel downstream target of PI3K/Akt signaling pathway. *Cellular Signalling* 24, 17–24.
- Warby, S.C., Chan, E.Y., Metzler, M., et al., 2005. Huntingtin phosphorylation on serine 421 is significantly reduced in the striatum and by polyglutamine expansion in vivo. *Human Molecular Genetics* 14, 1569–1577.
- Whitman, M., Downes, C.P., Keeler, M., Keller, T., Cantley, L., 1988. Type I phosphatidylinositol kinase makes a novel inositol phospholipid, phosphatidylinositol-3-phosphate. *Nature* 332, 644–646.
- Yao, H., Han, X., 2014. The cardioprotection of the insulin-mediated PI3K/Akt/mTOR signaling pathway. *American Journal of Cardiovascular Drugs* 14, 433–442.
- You, H., Pellegrini, M., Tsuchihara, K., et al., 2006. FOXO3a-dependent regulation of Puma in response to cytokine/growth factor withdrawal. *Journal of Experimental Medicine* 203, 1657–1663.
- Zinzalla, V., Stracka, D., Oppliger, W., Hall, M.N., 2011. Activation of mTORC2 by association with the ribosome. *Cell* 144, 757–768.

The JAK–STAT–SOCS Signaling Cascade

JM Murphy, JJ Babon, NA Nicola, and SE Nicholson, Walter and Eliza Hall Institute of Medical Research, Melbourne, VIC, Australia and The University of Melbourne, Melbourne, VIC, Australia

© 2016 Elsevier Inc. All rights reserved.

Introduction

The JAK–STAT–SOCS signaling cascade is pivotal to the intracellular signal transduction initiated by many cytokine and growth factor receptors and as such drives numerous biological processes including hematopoiesis, innate and adaptive immunity, and growth and cellular differentiation. It has been the subject of intense research over the past 20 years and our understanding has grown exponentially. In essence, ligand binding to cognate cell surface receptors results in activation of the Janus kinases (JAK), which are constitutively associated with the receptor cytoplasmic domains. The JAKs, in turn, phosphorylate tyrosine residues within the receptor cytoplasmic domains to recruit signaling intermediates, such as the latent signal transducer and activator of transcription (STAT) proteins, which themselves become phosphorylated and move to the nucleus inducing a wave of gene transcription. The suppressors of cytokine signaling (SOCS) proteins act as a brake to limit the extent of JAK-mediated signaling, and are often induced as part of the STAT transcriptional response in a classic negative feedback loop.

The STATs were originally identified as components of an interferon (IFN)-regulated transcription complex (Levy *et al.*, 1988, 1989; Schindler *et al.*, 1992). Shortly thereafter, JAK1 and JAK2 were discovered in a degenerate polymerase chain reaction (PCR) screen (Wilks *et al.*, 1989; Wilks, 1991; Harpur *et al.*, 1992), whilst Tyk2 and JAK3 were independently discovered (Firmbach-Kraft *et al.*, 1990; Witthuhn *et al.*, 1993, 1994; Rane and Reddy, 1994; Takahashi and Shirasawa, 1994). The JAK family initially suffered from the misnomer of being ‘Just Another Kinase.’ However, it soon became clear that this unique family of protein tyrosine kinases were critical for the initiation of cytokine signaling (Johnston *et al.*, 1994; Witthuhn *et al.*, 1993, 1994) and they were renamed ‘Janus kinase’ for the two-headed Roman god of gates and in appreciation of the twin kinase domains (JH1 and JH2; see below). The positioning of the JAK and STAT proteins in the IFN cytokine cascades was delineated by an elegant series of somatic cell complementation studies performed in the Kerr and Stark laboratories in the early 1990s (Darnell *et al.*, 1994; Watling *et al.*, 1993; Shuai *et al.*, 1993; Velazquez *et al.*, 1992; Muller *et al.*, 1993). Whilst the full complement of mammalian STAT proteins (seven in total) were discovered in a burst of activity from 1992 to 1995 (Schindler *et al.*, 1992; Wakao *et al.*, 1994; Mui *et al.*, 1995; Liu *et al.*, 1995; Hou *et al.*, 1994; Zhong *et al.*, 1994; Akira *et al.*, 1994; Yamamoto *et al.*, 1994).

The field burgeoned and the analysis of genetically modified mice (Table 1) and the identification of various human mutations (Russell *et al.*, 1995; Macchi *et al.*, 1995) revealed the complex and highly specific nature of JAK–STAT recruitment by various receptor complexes.

Excessive cytokine signaling causes severe pathology in many inflammatory, allergic, and autoimmune disease states.

The SOCS proteins are small intracellular proteins that have a key role in controlling the immune response. SOCS1 was simultaneously discovered by three groups; in a screen for negative regulators of IL-6 signaling (Starr *et al.*, 1997) as a JAK-binding protein (Endo *et al.*, 1997) and based on antigenic similarity to the STAT3-SH2 domain (Naka *et al.*, 1997). Cytokine inducible SH2-containing protein (CIS) had been identified 2 years earlier and as its name suggests was identified as an interleukin (IL)-2, IL-3, granulocyte macrophage-colony stimulating factor (GM-CSF), and erythropoietin (EPO)-inducible protein (Yoshimura *et al.*, 1995). SOCS1 sequence similarity to CIS and database mining subsequently revealed a family of eight related mammalian SOCS proteins (Hilton *et al.*, 1998) that primarily function as adaptors to recruit an E3 ubiquitin ligase complex and regulate cytokine signaling.

A number of insights have been gained from the analysis of JAK–STAT signaling in *Drosophila*, which has an evolutionarily conserved but less redundant complement of JAKs, STATs, and SOCS. These are: Unpaired (ligand), Domeless (receptor), Hopscotch (JAK1/JAK2 homolog), STAT92E (STAT3/5 homolog), SOCS16D, SOCS44A (SOCS6/7 homolog), and SOCS36E (SOCS5 homolog) and are reviewed in Stec and Zeidler (2011). As an example, mutations in the pseudokinase domain of Hopscotch cause a leukemia-like disease (Luo *et al.*, 1995), not unlike the myeloproliferative disorders that result from mutation of valine 617 in the pseudokinase domain of human JAK2 (Cazzola and Kralovics, 2014).

It is important to acknowledge that in addition to the STATs, other pathways are activated following JAK activation and receptor phosphorylation (for instance, PI3K, PLC- γ , Ras-MAPK); similarly there are other ‘non-SOCS’ mechanisms for negative regulation of these pathways. The following summarizes our current understanding of the JAK–STAT–SOCS signaling cascade.

The JAK Family

The Evolution of the JAK Family

Cytokine receptor signal transduction relies on the association of the receptor intracellular portions with members of the JAK family of non-receptor tyrosine kinases. Cytokine receptors possess no intrinsic catalytic activity and as a result rely on recruitment of effector kinases, principally the JAKs, to convey signals. In vertebrates, there are four JAK family members, JAK1, JAK2, JAK3, and TYK2, each with common domain architecture (Figure 1(d)). The JAKs are composed of seven JAK-homology (JH) domains: at the N-terminus, JH7-4 constitute a FERM domain; JH3, an atypical SH2 domain; JH2, a pseudokinase domain; and at the C-terminus, JH1, the tyrosine kinase domain (Harpur *et al.*, 1992; Wilks *et al.*, 1991).

Table 1 The *in vivo* phenotypes caused by JAK/STAT/SOCS deficiency

Gene	Model	Viability	Phenotype	Cytokine signaling	References
Jak1	Mouse	Postnatal day 1 lethality	Death attributed to neurological defects from defective gp130 signaling. Small thymic and defective thymocyte production	IL-3; IL-2, IL-4, IL-7; IL-6, LIF; IFN α ; IFN γ ; IL-10; GM-CSF; IL-5; G-CSF	Rodrig <i>et al.</i> , 1998
Jak2	Mouse	Embryonic lethality at E12.5	Failure of definitive erythropoiesis	IL-3, IL-5, GM-CSF; IFN γ ; EPO; TPO; GH; prolactin (PRL); IL-6, LIF; IFN α / β ; G-CSF	Neubauer <i>et al.</i> , 1998; Parganas <i>et al.</i> , 1998; Wagner <i>et al.</i> , 2004
Jak2	Mouse	Lethal	PN4 lethality due to defective erythropoiesis. Adult mice succumbed to severe anemia, thrombocytopenia, stem cell loss, and bone marrow failure		Park <i>et al.</i> , 2013; Akada <i>et al.</i> , 2014; Grisouard <i>et al.</i> , 2014; Kremler <i>et al.</i> , 2004
		Viable	Impaired mammary gland development and maintenance of functionally differentiated alveolar cells		Wagner <i>et al.</i> , 2004
Jak3	Mouse	Viable	Develop severe combined immunodeficiency syndrome (SCID). Block in B-cell maturation in the bone marrow; smaller thymic and spleen size; and defective T cell function	IL-2, IL-4, IL-7, IL-9, IL-13, IL-15, IL-21	Nosaka <i>et al.</i> , 1995; Thomis <i>et al.</i> , 1995; Park <i>et al.</i> , 1995
JAK3	Human	Viable	Diminished T and NK cell numbers; normal B cell numbers		Macchi <i>et al.</i> , 1995; Russell <i>et al.</i> , 1995
Tyk2	Mouse	Viable	Macrophages unresponsive to lipopolysaccharide (LPS) stimulation. Reduced IL-12 mediated T-cell response following viral challenge	IL-12; IFN α / β and IFN γ (at low doses); IL-23; IL-3; IL-6, LIF; IL-10; G-CSF, TPO	Shimoda <i>et al.</i> , 2000; Karaghiosoff <i>et al.</i> , 2000, 2003; Shaw <i>et al.</i> , 2003; Ishizaki <i>et al.</i> , 2014
Tyk2	Mouse	Viable	Myeloid deletion reduced murine cytomegalovirus (MCMV) defense. Ubiquitous, but not lineage specific deletions, impaired tumor immunosurveillance		Vielhascher <i>et al.</i> , 2014
TYK2	Human	Viable	Susceptibility to viral, fungal, and mycobacteria infection. Atopic dermatitis		Minegishi <i>et al.</i> , 2006
STAT1	Mouse	Viable	Susceptibility to bacterial and viral infection	Type I IFN; IFN γ	Durbin <i>et al.</i> , 1996; Meraz <i>et al.</i> , 1996
STAT2	Mouse	Viable	Susceptibility to viral infection	Type I IFN	Park <i>et al.</i> , 2000

(Continued)

Table 1 Continued

Gene	Model	Model	Viability	Phenotype	Cytokine signaling	References
STAT3	Mouse	Germline knockout	Embryonic lethality at E6.5-7.5	Failure to form visceral endoderm	gp130 family cytokines; IL-10; G-CSF; leptin; IL-21, IL-23	Takeda <i>et al.</i> , 1997
		Conditional: Lung, bone, colon, heart, nervous system, and skin	Viable			Hokuto <i>et al.</i> , 2004; Zhang <i>et al.</i> , 2005; Welte <i>et al.</i> , 2003; Hliffiker-Kleiner <i>et al.</i> , 2004; Jacoby <i>et al.</i> , 2003; Okada <i>et al.</i> , 2006; Sano <i>et al.</i> , 1999
		'Knock-in' leptin receptor STAT3 mutant	Viable	Obesity		Bates <i>et al.</i> , 2003
STAT4	Mouse	Germline knockout	Viable	Reduced Th1 differentiation and natural killer cell function	IL-12; IL-23	Kaplan <i>et al.</i> , 1996; Thierfelder <i>et al.</i> , 1996
STAT5a	Mouse	Germline knockout	Viable	Defect in mammary gland development, failure to lactate	IL-3; GM-CSF; Prolactin; growth hormone; EPO; γ_c cytokines	Liu <i>et al.</i> , 1997
STAT5b STAT5a/b	Mouse Mouse	Germline knockout Compound germline knockout	Viable > 99% perinatal lethality	Impaired growth Severe anemia		Udy <i>et al.</i> , 1997 Socolovsky <i>et al.</i> , 1999, 2001; Yao <i>et al.</i> , 2006
STAT6	Mouse	Germline knockout	Viable	Impaired lymphoid development and differentiation Defective Th2 polarization, and IgG and IgE class switching; increased susceptibility to parasite infection Protection from allergic inflammation	IL-4; IL-13.	Shimoda <i>et al.</i> , 1996; Takeda <i>et al.</i> , 1996a,b; Urban <i>et al.</i> , 1998; Akimoto <i>et al.</i> , 1998; Kuperman <i>et al.</i> , 1998
CIS	Mouse	Germline knockout Conditional: T cells	Viable	Airway inflammation and increased susceptibility to experimental asthma. Enhanced generation of Th2 and Th9 T cell subsets	IL-2; IL-4.	Yang <i>et al.</i> , 2013
SOCS1	Mouse	Germline knockout	Perinatal lethality. Can be rescued by crossing KOs with mice deficient for IFN γ , IFN α receptor, Stat1 or Stat6	Multi-organ inflammatory infiltrate	Type I IFN; IFN γ ; IL-4; γ_c cytokines	Alexander <i>et al.</i> , 1999; Davey <i>et al.</i> , 2005; Cornish <i>et al.</i> , 2003; Starr <i>et al.</i> , 1998; Naka <i>et al.</i> , 2001; Fenner <i>et al.</i> , 2006

SOCS2	Mouse	Conditional: T cells Germline knockout	Viable	Gigantism	Growth hormone; prolactin	Chong <i>et al.</i> , 2003 Metcalf <i>et al.</i> , 2000
SOCS3	Mouse	Germline knockout	Embryonic lethality at E12-16	Placental deficiency as a result of excessive LIF signaling	gp130 cytokines; G- CSF; leptin; IL-12	Robb <i>et al.</i> , 2005; Takahashi <i>et al.</i> , 2003
		Conditional: hematopoietic and endothelial cells	Lethal as young adults	Inflammation in lung and peritoneal cavities		Croker <i>et al.</i> , 2003, 2004; Lang <i>et al.</i> , 2003
		Macrophages and neutrophils	Viable			
		'Knock-in' gp130Y757F mice (no SOCS3 recruitment)	Viable	Inflammatory joint disease		Yasukawa <i>et al.</i> , 2003a
SOCS4	Mouse	ENU-mutant	Viable	Increased susceptibility to influenza infection		Kedzierski <i>et al.</i> , 2014
SOCS5	Mouse	Germline knockout	Viable	No described phenotype		Brender <i>et al.</i> , 2004; Khaled <i>et al.</i> , 2007
SOCS6	Mouse	Germline knockout	Viable	Mild mammary gland hyperplasia		Krebs <i>et al.</i> , 2002
SOCS7	Mouse	Germline knockout	Perinatal lethality (C57BL/6)	Mild growth retardation		Krebs <i>et al.</i> , 2004; Banks <i>et al.</i> , 2005
			Viable (129/SvJ)	Hydrocephaly		
				Enhanced insulin signaling and improved glucose tolerance		

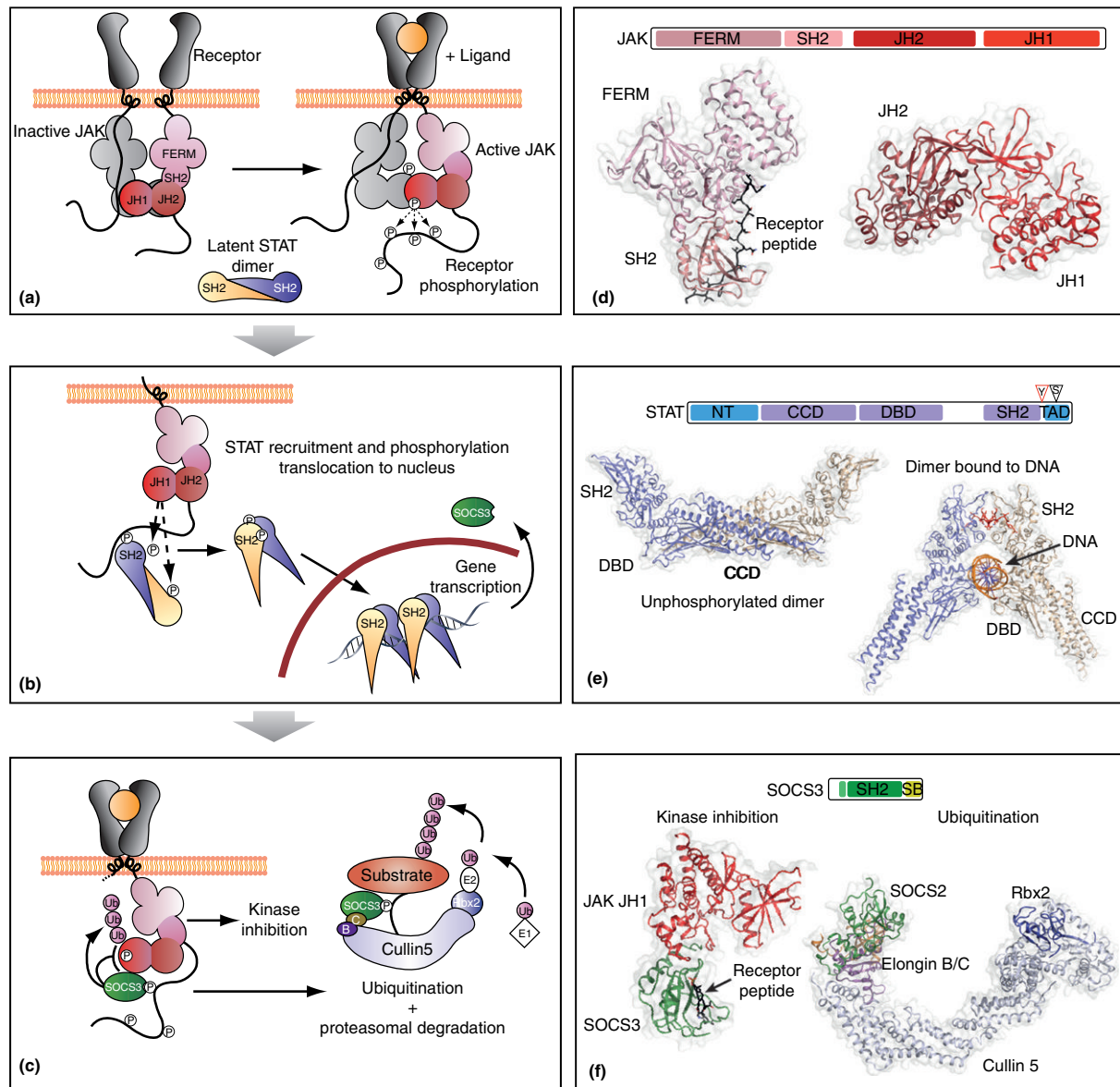


Figure 1 The JAK–STAT–SOCS signaling cascade. Cartoon representation of (a) JAK activation, (b) STAT activation, and (c) SOCS3 inhibition of JAK activity. (d–f) Schematics showing the generic domain architecture for JAK, STAT, and SOCS3 together with the three-dimensional crystal structures for (d) TYK2 FERM and SH2 (residues 23–583) in complex with a Box 2 peptide from IFNAR1 (residues 465–512) (Wallweber *et al.*, 2014) (pdb4po6), Tyk2 JH2 and JH1 (residues 566–1187) (Lupardus *et al.*, 2014) (pdb4OLI) domains; (e) unphosphorylated (Mao *et al.*, 2005) (pdb1yvl) and phosphorylated Stat1 dimer core (residues 132–713) bound to DNA (Chen *et al.*, 1998) (pdb1bf5); (f) SOCS3–SH2 domain (residues 22–185) bound to the gp130 phosphotyrosine peptide (residues 750–764) and JAK2 JH1 domain (Kershaw *et al.*, 2013) (pdb4GL9). The SOCS3–SOCS box–Elongin BC complex with Cullin5 and Rbx2 is a model created by juxtaposition of the following structures: SOCS2/BC/Cullin5 (pdb4JGH), Cullin5 N-terminal domain (pdb2WZK), Cullin5 C-terminal domain/Rbx2 (pdb3DPL). CCD, Coiled-coil domain; DBD, DNA binding domain; NT, N-terminal domain; SB, SOCS box.

Intriguingly, studies of each of the JAK tyrosine kinase (JH1) domains *in vitro* have not revealed any clear substrate specificities beyond a preference for acidic residues adjacent to the substrate tyrosine (Sanz *et al.*, 2011), a finding typical of tyrosine kinase domains when divorced from their regulatory domains. This then begs the question: if not solely via the substrate preferences of the catalytic, tyrosine kinase domain, how do JAKs achieve specificity in signaling? And secondly, if the tyrosine kinase domains of the four JAKs have overlapping

substrate specificities (Blouin *et al.*, 2011), why have four JAKs, rather than one evolved?

Our understanding of the capacity of JAKs to impart specificity in signaling has been greatly advanced by the development of knockout mouse models for each of the JAK family members (Table 1). Of note, JAK1, JAK2, and TYK2 are ubiquitously expressed, and only JAK3 expression is restricted to hematopoietic tissues, including the spleen, bone marrow, fetal liver, bone marrow and thymus, lymphoid and myeloid

cell lineages (Johnston *et al.*, 1994; Musso *et al.*, 1995; Witthuhn *et al.*, 1994). As a result, the ubiquity of JAK expression offers no solution to the origin of specificity. However, the phenotypes of mice genetically deleted for each JAK provide important insights into specificity. Firstly, the lethality of *Jak2*^{-/-} mice around embryonic day 12.5 demonstrates that JAK2 serves nonredundant functions in signaling that are not compensated for by JAK1, JAK3, and TYK2 (Neubauer *et al.*, 1998; Parganas *et al.*, 1998). Studies of fetal liver cells derived from *Jak2*^{-/-} mice revealed deficits in signaling downstream of several cytokine receptors (Table 1) and, relatedly the observed embryonic lethality was attributed to an absence of definitive erythropoiesis (Neubauer *et al.*, 1998; Parganas *et al.*, 1998). An essential role of JAK2 beyond embryonic development was underscored by conditional deletion of JAK2 in postnatal day 4 and adult mice (Akada *et al.*, 2014; Park *et al.*, 2013), where systemic deletion of JAK2 led to lethality due to defects in erythropoiesis, thrombopoiesis, hematopoietic stem cell maintenance and consequent bone marrow failure. While some cytokine receptors have evolved obligate relationships with effector JAKs (Table 1), others appear to have evolved promiscuous relationships and can equally utilize a subset or all JAK family members for signaling. This is evidenced by the granulocyte-colony stimulating factor receptor (G-CSFR), which exhibits no deficits in signaling upon deletion of *Jak1*, *Jak2*, or *Tyk2*, suggesting a redundancy where each JAK family member can signal equivalently downstream of G-CSFR activation.

These findings illustrate that, in many cases, cytokine receptors have developed obligate relationships with the JAK effectors they utilize, which cannot simply be accounted for by tissue-specific expression of the paralogs. Instead, there appears to be a biochemical, rather than biological, basis for the preference of receptors for JAKs. Two sequence motifs, termed Box 1 and Box 2, have been identified within the cytoplasmic portions of cytokine receptors and shown to function variably in JAK recruitment and activation. Box 1 is characterized by a consensus $\Pi\text{ITPV}[I/V]\text{PxP}[E/K]$ sequence (where Π is any hydrophobic residue), which is typically located in the intracellular juxtamembrane region, 6–10 residues C-terminal to the transmembrane domain (Tanner *et al.*, 1995; Usacheva *et al.*, 2002). Box 2, in contrast, is less well conserved and is characterized by a hydrophobic-charged-hydrophobic-hydrophobic motif, such as LEVL or VEVI. Relative to Box 1, the position of the Box 2 motif is highly variable, as exemplified by a Box 2 sequence positioned 7 residues C-terminal of Box 1 in IFNAR2, but 31 residues C-terminal to Box 1 in gp130 (Usacheva *et al.*, 2002). Broadly, these motifs can be thought to serve as recruitment sites for the JAKs via their FERM (JH7-JH4) domains. Typically, the function of JAK recruitment has been attributed to Box 1 in class I cytokine receptors, such as growth hormone receptor (GHR), EPOR, gp130, and beta common (β_c), with Box 2 believed to function in JAK activation rather than binding. Among the class II cytokine receptors, however, there is wide variability in the conservation and function of Box 1 and Box 2 motifs. Indeed, IFNGR1 and the IL-10R α have no apparent motif that conforms to the above Box 1 and Box 2 consensus sequences (Usacheva *et al.*, 2002), yet still bind and signal via JAK1 (Table 1). The absence of strict rules governing JAK recruitment by the receptors is

hardly surprising if one considers that each receptor system has coevolved with JAKs to permit combinatorial use of different JAK family members to enable signal pleiotropy. For example, both JAK1 and JAK3 are utilized as effectors downstream of γ_c receptor activation (Table 1), and as such the receptor has evolved motifs that enable binding of both JAKs, whereas GHR signals strictly via JAK2 and has thus evolved the capacity to strictly bind JAK2. In 2014, the first detailed insights into receptor binding by the JAKs were provided by the crystal structure of the TYK2 FERM-SH2 (JH7-3) domains in complex with the IFNAR1 Box 2 motif (Wallweber *et al.*, 2014; Figure 1(d)). This work revealed that the F1 lobe of the FERM domain intimately interacts with the SH2 domain, thereby positioning the SH2 domain adjacent to the F2 lobe of the FERM domain to form a composite Box 2 binding motif. Interestingly, the SH2 domain contains a His (H474) in place of the canonical phosphotyrosine binding Arg, consistent with a non-phosphotyrosine binding function. Instead, the adjacent residues, S476 and T477, were shown to form two hydrogen bonds with the carboxylate of the IFNAR1 Box 2 residue, E497: an interaction essential for complexation. Whether similar interactions are likely to occur between other JAKs and the receptors, especially in cases where Box 1 is the primary recruitment motif, awaits further structural characterization.

Mechanism of Activation

These receptors assemble into oligomeric structures and thereby act as scaffolds to recruit two (or more) JAK family proteins. Prior to receptor activation, the receptor-bound JAK kinases are maintained in an inactive state (Figure 1(a)). Upon ligation of the receptor ectodomains by a cognate cytokine ligand, the receptor subunits are reoriented, with this motion conveyed from the ectodomain, across the transmembrane domain, into the intracellular region. The resulting repositioning of the intracellular regions of the receptor subunits is believed to simultaneously reposition the receptor-associated JAKs to enable phosphorylation of one another and consequent activation (Figure 1(a)). It has long been established from biochemical studies that maximal catalytic activity by the JAK family tyrosine kinase domains depends on the trans-phosphorylation of the activation loop tyrosine residues (Chatti *et al.*, 2004). Inhibition of JAK catalytic activity prior to receptor activation can be attributed to the pseudokinase (JH2) domain, which recent data suggest acts *in trans* to suppress tyrosine kinase domain activation (Figure 1(a); Brooks *et al.*, 2014; Koppikar *et al.*, 2012; Varghese *et al.*, 2014). The functions of pseudokinase domains are still emerging (Murphy *et al.*, 2014) and there appears to be subtle differences between the functions of those within the JAK family. For example, the pseudokinase domain of JAK2 has been shown to exhibit a weak catalytic function (Bandaranayake *et al.*, 2012; Ungureanu *et al.*, 2011) that enables phosphorylation of negative regulatory sites: S523 in the SH2-pseudokinase domain linker and Y570 within the pseudokinase domain (Bandaranayake *et al.*, 2012; Ishida-Takahashi *et al.*, 2006; Mazurkiewicz-Munoz *et al.*, 2006; Ungureanu *et al.*, 2011). In contrast, the counterpart domain in JAK1 exhibited no detectable catalytic activity (Toms *et al.*, 2013) and the TYK2

pseudokinase domain did not detectably bind ATP (Murphy *et al.*, 2014). These data argue for the JAK pseudokinase domains principally serving catalysis-independent functions to inhibit the catalytic activities of the tyrosine kinase domains. The function of the JAK2 pseudokinase domain as a suppressor of the tyrosine kinase domain is well established (Saharinen and Silvennoinen, 2002; Saharinen *et al.*, 2000, 2003), and human myeloproliferative neoplasms have been attributed to mutations within the pseudokinase domain, exemplified by the V617F mutation first reported in 2005 (Zhao *et al.*, 2005; Levine *et al.*, 2005; James *et al.*, 2005; Kralovics *et al.*, 2005; Baxter *et al.*, 2005).

Recent work has laid a platform for understanding how the JAK pseudokinase domain can suppress the catalytic activity of the JAK tyrosine kinase domain. Many models have been generated in the absence of experimental support in which the pseudokinase domain was proposed to bind and suppress the adjacent tyrosine kinase domain *in cis*. In the case of JAK2, however, recent experimental data supports the existence of the pseudokinase–kinase tandem domains in an elongated conformation. Firstly, recombinant pseudokinase–kinase tandem domains expressed and purified from insect cells were phosphorylated at Y1007/Y1008, a hallmark of JAK2 activation, illustrating a limited capacity for the pseudokinase domain to bind and inhibit the catalytic activity of the adjacent kinase domain *in cis* (Varghese *et al.*, 2014). Further, rather than a condensed conformation in which an intimate association between the pseudokinase and tyrosine kinase domains mediates inhibition of catalytic activity, small-angle X-ray scattering studies supported the idea that, in solution, the JAK2 pseudokinase–kinase tandem domains exist in an extended conformation (Figure 1(d); Varghese *et al.*, 2014). JAK1 was similarly observed to adopt an elongated conformation in cryo-electron microscopy studies (Lupardus *et al.*, 2011).

These data are consistent with an attractive model that has recently been proposed based on experimental studies of JAK2 activation downstream of the GHR (Brooks *et al.*, 2014). This model is based on the observation that GHR ligation by GH results in a relative rotation of receptor subunits, which, based on mutational and crosslinking studies, was modeled to be transmitted via the helical transmembrane domains to the intracellular, helical juxtamembrane region and Box 1 motif. Unexpectedly, by fusing FRET probes either C-terminal to the GHR Box 1 motif or C-terminal to the JAK2 pseudokinase or kinase domain and monitoring their positions following GHR activation (Brooks *et al.*, 2014), it was proposed that the JAK2 molecules associated with each of the two receptor subunits are drawn apart, rather than into proximity as previously believed. Based on modeling, crosslinking and biochemical studies, other homodimeric class I cytokine receptors, such as EPOR and c-Mpl, were similarly proposed to contain helical regions adjacent to Box 1 in their juxtamembrane regions (Staerk *et al.*, 2011; Seubert *et al.*, 2003). Taken together these studies suggest a concerted model for JAK2 activation following homodimeric receptor activation, although the extent to which the same rules apply to the more complicated, heteromeric class I and class II cytokine receptors remains to be established.

Furthermore, it is unclear whether the proposed mechanism of JAK2 activation is likely to be conserved among all JAK family members. While JAK1 was found to exist in an

elongated conformation with no fixed pseudokinase:kinase orientation (Lupardus *et al.*, 2011), an observation consistent with a JAK2-like activation mechanism, the opposite has been proposed for TYK2 (Lupardus *et al.*, 2014). The recent crystal structure of TYK2 pseudokinase–kinase tandem domains suggests the two domains intimately interact via their N-lobes in a back-to-back arrangement (Figure 1(d)). This idea is supported by the observation that within the context of the purified recombinant TYK2 pseudokinase–kinase, the kinase domain activation is suppressed, and mutations that are predicted to perturb the interaction of the pseudokinase and kinase domains relieve this suppression (Lupardus *et al.*, 2014). These data, and the observation that the TYK2 pseudokinase–kinase domain linker is 14 residues shorter than the equivalent sequence in JAK2, raise the possibility that an *in cis* mode of pseudokinase-mediated kinase domain inhibition could predominate in TYK2, while the longer interdomain linker favours *in trans* inhibition in JAK2. The precise details of these mechanisms await further detailed structural analyses.

Signaling Downstream of JAK Activation

Following ligation of receptors by their ligands, receptor-associated JAK molecules are opposed to facilitate their activation by *trans*-phosphorylation. Principally, the activation is thought to comprise mutual phosphorylation of their kinase domain activation loops on a Tyr-Tyr motif, where phosphorylation of the first of these Tyr residues (Y1007 in JAK2) is known to be essential for elevated catalytic activity (Chatti *et al.*, 2004; Feng *et al.*, 1997). Following JAK activation, however, the precise details of how JAKs activate further signaling pathways remain unclear. Many tyrosines within receptor intracellular regions have been identified as JAK substrates and their phosphorylation is known to enable recruitment of phosphotyrosine binding-domain containing effectors to the receptor, including the STATs, as detailed in the next section. The outstanding questions remain: why have JAKs evolved as non-receptor tyrosine kinases rather than RTKs and, relatedly, do JAKs serve functions in addition to their roles as cytokine receptor effectors? One answer to these questions may come from studies attributing a nuclear function to JAKs, where divorced from cytokine receptors, JAK2 has been reported to phosphorylate histone H3 in the nucleus (Dawson *et al.*, 2009). On the other hand, convincing evidence for non-cytokine receptor based cytoplasmic functions of JAKs has yet to emerge. In fact, JAK1 phosphorylation following LIF receptor activation induced dissociation from the receptor, a step that was found to promote its proteasomal degradation via the negative regulator, SOCS3 (Boyle *et al.*, 2009). These findings suggest that rather than activated JAKs serving non-receptor, cytoplasmic signaling functions, mechanisms exist within the cell to eliminate activated JAKs once dissociated from their receptors. These and other negative regulatory mechanisms are outlined below.

STAT Proteins

The STAT proteins are characterized by a unique ability to transduce a signal from the receptor complex at the membrane

to the nucleus, where they function as transcription factors, inducing an appropriate program of gene expression. Conceptually, this makes the JAK-STAT pathway one of the simpler, and perhaps more efficient, signaling cascades.

The STAT proteins exist in the cytoplasm as preformed dimers and are recruited to the receptor complex following ligand binding, JAK activation, and phosphorylation of key tyrosine motifs within the receptor cytoplasmic domains. They bind to the receptor phosphotyrosine residues via their SH2 domains and become 'second tier' substrates for phosphorylation by the JAK kinases. Phosphorylation occurs on a key conserved tyrosine in the C-terminal STAT tail and induces a major change in the configuration of the dimer (antiparallel to parallel); stabilized by SH2 binding to the reciprocal STAT tyrosine residue. As a consequence, the STATs disengage from the receptor complex and translocate to the nucleus, traversing the nuclear pore complex with the assistance of various transport receptors, commonly referred to as importins (Liu *et al.*, 2005). Once in the nucleus they bind to promoter elements to induce gene transcription.

There are seven STAT family members: STAT1, 2, 3, 4, 5A and 5B, and STAT6. All contain amino-terminal (NT), coiled-coil (CCD), DNA binding (DBD) and linker regions, in addition to an SH2 domain and a C-terminal conserved tyrosine residue and transactivation domain (TAD) (Figure 1(e)). A number of differentially spliced isoforms exist for STAT1, STAT3, STAT4, and STAT5 that result in truncation of the C-terminal tail (Darnell *et al.*, 1994; Hoey *et al.*, 2003; Wang *et al.*, 1996), while STAT5A and 5B are the products of a gene duplication event (Liu *et al.*, 1995).

SH2 domains are approximately 100 amino acids in length and have the ability to bind phosphotyrosine residues that are contained within specific sequence motifs. In general, different SH2 domains will bind different motifs. STAT recruitment to different receptor complexes is determined by STAT-SH2 domain recognition of the residues located adjacent to the receptor phosphotyrosines. The result is ligand-dependent activation of the various STAT proteins (Heim *et al.*, 1995) and is consistent with their distinct biological roles, which for the most part, have been elucidated by the generation and analysis of STAT-deficient mice. For example, STAT1 is recruited to both the IFN γ receptor (IFNGR1) (Y440) (Greenlund *et al.*, 1995) and the IFN α/β receptor (IFNAR1) (Y510) (Zhao *et al.*, 2008) receptors. Accordingly, *Stat1*-deficient mice are susceptible to infection by both bacterial and viral pathogens due to an inability to signal via type I and type II IFNs (Meraz *et al.*, 1996; Durbin *et al.*, 1996). Type I IFNs also signal through STAT2 and *Stat2*-deficient mice are similarly highly susceptible to viral infection. G-CSF, leptin, and gp130 cytokines signal through STAT3, IL-4 and IL-13 through STAT6, IL-12 and IL-23 via STAT4 and the IL-2 family cytokines, IL-3, GM-CSF, EPO, prolactin, and GH signal through STAT5 (Table 1).

The STAT proteins can form homo- or heterodimers, each complex initiating a unique pattern of gene transcription. STAT1, STAT3, STAT4, STAT5, and STAT6 all function as homodimers. The STAT1 homodimer was originally described as a gamma-activated factor (GAF). STAT1 and STAT3 can additionally form a heterodimer, while STAT2, together with STAT1 and IRF9, forms a heterotrimeric transcription complex, which is often referred to as ISGF3 (IFN-stimulated gene 3;

reviewed in Fink and Grandvaux, 2013). STAT2 is also thought to have some STAT1-independent functions (Hahm *et al.*, 2005; Perry *et al.*, 2011).

STAT signaling within the cell is in constant flux; both phosphorylated and unphosphorylated STAT dimers can enter the nucleus and both can induce gene transcription. However, the import and export of STAT proteins from the nucleus appears to be differentially regulated within the STAT family (Reich, 2013). At least in the context of STAT1, the phosphorylated 'parallel' dimer is actively trafficked into the nucleus, is able to bind to specific DNA sequences and is slower to exit the nucleus, resulting in cytokine-driven signaling responses. That is not to say that 'unphosphorylated' STAT-driven transcription is not without biological consequence, as is evident from the transcriptional profiles observed in various cancers where STAT3 is elevated (Yang *et al.*, 2005) (discussed below).

In general, the STAT dimers bind to a palindromic core sequence: TTC(N)₂₋₄GAA (Xu *et al.*, 1996; Schindler *et al.*, 1995) in the promoter regions of target genes, but have preferences for the spacing between the two halves of the palindrome, with STAT1-5 preferring N₃, while STAT6 prefers N₄ (Ehret *et al.*, 2001). The two STAT molecules form a clamp-like structure which is held together by the reciprocal SH2:phosphotyrosine interactions, while the two DNA binding domains sandwich the DNA core (Chen *et al.*, 1998; Figure 1(e)). Interactions between the NT regions of adjacent STAT dimers stabilize binding to neighboring sites (Vinkemeier *et al.*, 1996; Xu *et al.*, 1996) and can also facilitate the recruitment of additional STAT1 dimers in the absence of multiple GAF sites. This latter results in a cooperative mode of binding that is required for STAT1 transcriptional responses to type II IFN but strikingly, not for type I IFN responses mediated by the ISGF3 complex (Begitt *et al.*, 2014), and provides a molecular explanation for the different transcriptional profiles generated by STAT1-containing complexes.

The key serine in the transactivator domain of STAT1, 3, 4, and 5 can be phosphorylated by various serine kinases and phosphorylation is required for full transcriptional activity (Wen *et al.*, 1995; Visconti *et al.*, 2000). Consequently, the truncated STAT isoforms appear to function as dominant negatives (Wang *et al.*, 1996). De-phosphorylation of STAT proteins is a prerequisite for STAT exit from the nucleus. This is best illustrated by the role of the nuclear phosphatase TC45 in STAT1 de-phosphorylation (ten Hoeve *et al.*, 2002).

The unphosphorylated dimer is stabilized by two interfaces, one between the NT regions and a reciprocal interaction between the CCD and DBD (Figure 1(e); Mao *et al.*, 2005). The contribution of the unphosphorylated STAT dimers to signaling has steadily emerged over the past few years. STAT genes are often the transcriptional targets of active STAT dimers, with the net result being an increase in the pool of unphosphorylated STAT proteins. At least for STAT1 and STAT3, it is apparent that this can drive a secondary wave of transcription that is distinct from the canonical STAT response. This has a biological context with, for instance, the prolonged transcription of antiviral genes sustaining immunity to a chronic infection (Cheon and Stark, 2009).

There are also implications for cancers, where in addition to the activation of STAT3 by cytokines and oncogenic tyrosine

signaling cascades, reviewed in (Ernst and Putoczki, 2012; Yu *et al.*, 2009), increased expression of the unphosphorylated STAT3 pool can enhance the transcription of tumor-promoting genes (Yang *et al.*, 2005). The unphosphorylated dimers are proposed to activate gene transcription via binding to NF- κ B (Yang *et al.*, 2007), although there is now some evidence that they can bind directly to DNA (Nkansah *et al.*, 2013). The ability of unphosphorylated STAT proteins to stabilize heterochromatin and regulate three-dimensional chromatin structure, is also an area of interest (Shi *et al.*, 2008; Zhao *et al.*, 2013; Hu *et al.*, 2013).

Finally, independent of their classic transcriptional activity in the nucleus, there is an emerging role for STAT proteins in mitochondria, reviewed in (Meier and Larner, 2014). STAT3 is found associated with mitochondrial electron transport chain complexes where it promotes cellular respiration (Wegrzyn *et al.*, 2009) and can contribute to Ras-mediated oncogenesis (Gough *et al.*, 2009). STAT5 also appears to have a non-canonical role in maintaining the structural integrity of the Golgi apparatus and endoplasmic reticulum (Lee *et al.*, 2012).

Among the key genes induced by phosphorylated STAT dimers is an important family of small cytoplasmic inhibitors, the SOCS proteins.

The SOCS

Once activated, it is important that the JAK/STAT signaling cascade is switched off after an appropriate duration. This is both to prevent excessive proliferation in response to cytokine as well as to allow the cell to return to its basal state, ready for subsequent rounds of stimulation. The primary effectors responsible for this temporal control are the SOCS family. SOCS expression is induced by the JAK/STAT cascade and they act as components of a negative feedback loop to shut down signaling (Endo *et al.*, 1997; Naka *et al.*, 1997; Starr *et al.*, 1997).

The SOCS family consists of eight proteins in humans, SOCS1-7 and CIS (Hilton *et al.*, 1998). These eight family members are the descendants of three ancestral SOCS proteins (SOCS1/2/3/CIS; SOCS4/5, SOCS6/7), also found in extant organisms such as sea anenomes (Liongue and Ward, 2013). Of these, the first subgroup in particular (containing SOCS1, SOCS2, SOCS3, and CIS) have all been shown to strongly inhibit JAK/STAT signaling; their expression is induced by activated STATs and they then inhibit the cytokine-induced JAK/STAT signaling cascade. SOCS4 and SOCS5 also play a role in inhibition of JAK/STAT signaling (Linossi *et al.*, 2013) whilst, in contrast, SOCS6 and SOCS7 appear to regulate receptor tyrosine kinase (RTK) signaling pathways (Krebs *et al.*, 2002; Banks *et al.*, 2005; Bayle *et al.*, 2004; Zadjali *et al.*, 2011).

SOCS proteins all share a similar domain architecture consisting of an N-terminal domain, a central SH2 domain and a C-terminal SOCS box domain (Figure 1(f)). The SOCS box domain is a small ~40 amino acid domain consisting of 2–3 alpha helices and contains two separable binding motifs. The first, contained in the N-terminal half of the domain, binds the adapter complex elonginBC (Zhang *et al.*, 1999; Kamura *et al.*, 1998) whilst the second, an L-P-L-P motif found in the C-terminal half of the domain binds the E3 ligase scaffold protein Cullin5 (Kamura *et al.*, 2004; Muniz *et al.*,

2013; Babon *et al.*, 2008, 2009). E3 ubiquitin ligases catalyze the ubiquitination of substrate proteins and the eight SOCS proteins can be considered the substrate recruitment modules for Cullin5-based E3 ligases, conveying substrates to the multi-subunit ligase in order for them to be ubiquitinated. The majority of these substrates are believed to be cytokine receptors (Irlandoust *et al.*, 2007; Metcalf *et al.*, 2000; Kershaw *et al.*, 2014; Qing *et al.*, 2005; Landsman and Waxman, 2005; Lavens *et al.*, 2006) and their ubiquitination leads either to proteosomal degradation (Kamizono *et al.*, 2001) or lysosomal routing (Irlandoust *et al.*, 2007) thus inhibiting JAK/STAT signaling by removing the activated cytokine receptor, and potentially their associated JAKs. The SOCS box is the defining feature of the eight SOCS family members but is also found in a much larger family of proteins, all of which act as substrate recruitment modules for E3 ubiquitin ligases (Hilton *et al.*, 1998; Kile *et al.*, 2002; Linossi and Nicholson, 2012).

SOCS1 and SOCS3

SOCS1 and SOCS3 are arguably the most potent inhibitors of cytokine signaling within the SOCS family. Their expression is highly upregulated within 15–30 minutes (Wormald *et al.*, 2006) of exposure to cytokine and they then rapidly inhibit signaling, approximately halving the duration of the signaling cascade that would occur in their absence (Wormald *et al.*, 2006). Although induced by a multitude of different cytokines and capable of inhibiting an almost equal number (when artificially overexpressed), knockout studies have shown that SOCS1 and SOCS3 are highly specific inhibitors of just a handful of different cytokines under normal, physiological conditions (Table 1).

Genetic deletion of either SOCS1 or SOCS3 is lethal (Roberts *et al.*, 2001; Alexander *et al.*, 1999; Table 1). *Socs1*^{−/−} mice die a few weeks after birth from an inflammatory disease due to dysregulated IFN γ signaling (Alexander *et al.*, 1999). This disease can be rescued by genetic deletion of IFN γ and partially alleviated by deletion of STAT1 or STAT6, although the resultant animals still develop a fatal inflammatory disease and die earlier than wild-type littermates (Naka *et al.*, 2001; Alexander *et al.*, 1999). In addition to IFN γ signaling, genetic deletion studies have shown that SOCS1 is also a potent inhibitor of IFN α/β (Gingras *et al.*, 2004; Fenner *et al.*, 2006), IL-12/23 (Eyles *et al.*, 2002), IL-4/IL-13 (Naka *et al.*, 2001) and cytokines that signal through the γ_c chain (IL-2, 15, 21) (Davey *et al.*, 2005).

Socs3^{−/−} embryos are non-viable due to defects in placenta (Roberts *et al.*, 2001), a phenotype that can be rescued by simultaneous genetic deletion of LIF (Robb *et al.*, 2005), however the resultant mice still die as young adults from a fatal inflammatory disease. Conditional knockout of SOCS3 in restricted cell types has shown it to be specific for IL-6 family cytokines (Yasukawa *et al.*, 2003b), G-CSF (Crocker *et al.*, 2004), and leptin (Mori *et al.*, 2004).

Direct Kinase Inhibition by SOCS1 and SOCS3

In addition to acting as E3 ubiquitin ligases (Kershaw *et al.*, 2014), SOCS1 and SOCS3 directly inhibit the catalytic activity

of JAKs, via the action of their so-called 'kinase inhibitory regions' (KIRs) (Figure 1(c)). The KIR was first identified by the groups of Yoshimura (Sasaki *et al.*, 1999; Yasukawa *et al.*, 1999) and Kishimoto (Narazaki *et al.*, 1998) as a short region in the N-terminal domain of SOCS1 and SOCS3, immediately upstream of their SH2 domains. Recent structural studies have shown that the KIR is unstructured in the absence of JAK (Babon *et al.*, 2005, 2006) yet becomes structured in its presence, where it binds to the substrate-binding groove of the JAK catalytic domains and prevents substrate binding (Kershaw *et al.*, 2013). This inhibits further downstream signaling. Inhibition of signaling by this direct kinase inhibitory mechanism is the dominant mode-of-action of both SOCS1 and SOCS3, as genetic deletion of their SOCS boxes (and hence abrogation of their ability to induce ubiquitination of any substrates) causes a much milder phenotype than the full knockout (Zhang *et al.*, 2001; Boyle *et al.*, 2007). SOCS3 is capable of inhibiting JAK1, JAK2 and TYK2 via this mechanism but not JAK3, due to the absence of a conserved sequence element in the JAK3 kinase domain (Babon *et al.*, 2012). The JAK targets of SOCS1 are currently unknown.

Although capable *in vitro* of inhibiting three of the four JAK family members, SOCS3 is a specific inhibitor of only a small subset of cytokines *in vivo*. This is because SOCS3 can bind JAK and cytokine receptor simultaneously, via two different molecular surfaces (Kershaw *et al.*, 2013). The phosphotyrosine-binding groove of its SH2 domain binds to receptor whilst the opposing surface of its SH2 domain, along with the KIR, binds JAK (Figure 1(f)). Only receptors containing the correct phosphotyrosine-based motifs are targeted with high affinity. The true target of SOCS is therefore a JAK-cytokine receptor complex, rather than a JAK or receptor alone. This allows for specific, and potent, inhibition of cytokines that signal via receptors that (A) bind JAK1, JAK2, or TYK2 and (B) contain the correct phosphotyrosine-based binding sites. Cytokines that fulfill these two criteria are leptin (De Souza *et al.*, 2002), G-CSF (Hortner *et al.*, 2002), and IL-6 family cytokines (Nicholson *et al.*, 2000) (reviewed in Babon and Nicola, 2012). The JAK- or receptor-binding specificities of SOCS1 are currently unclear but it may act via a similar mechanism.

SOCS2 and CIS

SOCS2 and CIS, like SOCS1 and SOCS3, consist primarily of only SH2 and SOCS box domains, having only very short N-terminal regions. However, unlike SOCS1 and SOCS3, these two proteins do not contain a kinase inhibitory region and therefore rely entirely upon their ability to ubiquitinate substrates in order to inhibit JAK/STAT signaling. Consistent with this, *in vitro* studies have shown that both SOCS2^{ΔSB} and CIS^{ΔSB} (which lack the SOCS Box) act as dominant negatives (Greenhalgh *et al.*, 2005; Ram and Waxman, 2000). SOCS2 binds to phosphorylated growth hormone receptor (at pY595) (Bullock *et al.*, 2006, 2007) and induces its ubiquitination. Consistent with this, *Socs2*^{-/-} mice display gigantism due to dysregulated growth hormone signaling (Metcalf *et al.*, 2000). *Cis*^{-/-} mice are born healthy but have recently been shown to suffer a spontaneous lung disease, similar to that of allergic

asthma (Yang *et al.*, 2013) due to dysregulated IL-4 signaling. The molecular determinants of CIS action are yet to be determined.

SOCS4 and SOCS5

The involvement of SOCS4 and SOCS5 in regulation of JAK/STAT signaling is less well understood than for SOCS1, SOCS2, SOCS3 and CIS. In addition to a role in JAK/STAT signaling both SOCS4 and SOCS5, as well as their *Drosophila* homolog (SOCS36E) have been shown to regulate EGF signaling (Callus and Mathey-Prevot, 2002; Rawlings *et al.*, 2004; Nicholson *et al.*, 2005; Kario *et al.*, 2005). In mammals, SOCS5 has been implicated in the regulation of IL-4 signaling (Seki *et al.*, 2002) whilst little is known of the actions of SOCS4. Both SOCS4 and SOCS5 are highly similar at the amino acid level, especially across their SH2 domains and a structured segment within their large N-terminal domains (Feng *et al.*, 2012) which suggests they may have similar molecular targets. Despite this, recent studies have shown that SOCS5 can inhibit JAK1 and JAK2 autophosphorylation whereas SOCS4 cannot (Linossi *et al.*, 2013; Figure 1).

JAK-STAT-SOCS and Human Disease

Involvement of JAK Proteins in Human Disease

Inherited mutations in JAK3 and TYK2 can result in human immune deficiency syndromes. JAK3 is exclusively associated with the common γ -chain receptor that is utilized by cytokines required for T- and NK cell production and function including IL-2, IL-4, IL-7, IL-9, IL-15, and IL-21. Bi-allelic inherited mutations in either JAK3 (7–14% of patients) or the common gamma chain are associated with severe combined immunodeficiency syndrome (SCID) with the latter the most common in X-linked SCID. The JAK3 mutations occur throughout the coding sequence and generally result in a complete loss of protein production and absence of T- and NK but not B-cells (O'Shea *et al.*, 2013). TYK2 associates with IFN α/β and IL-6 receptors and mutations have been described in two patients. Both showed severe herpes virus infection and mycobacterial infection but differed in other clinical features, suggesting a common deficiency in IFN signaling, but also unique features such as hyper IgE that may be due to different genetic backgrounds (Kilic *et al.*, 2012; Minegishi *et al.*, 2006).

Several activating mutations in JAK2 have been associated with myeloproliferative neoplasms (MPN), chronic and acute myeloid and lymphoid leukemias. In several chromosomal translocations the oligomerization domain of one partner is fused to the kinase domain (JH1) or both kinase and pseudokinase (JH2) domains of N-terminally truncated JAK2 (James *et al.*, 2005; Jatiani *et al.*, 2010). Constitutive dimerization of such fusion proteins is thought to allow cross-phosphorylation and activation of the JAK2 kinase domain. The fusion partners include the ETS domain of the transcription factor TEL, the centrosome-associating protein PCM1 and the GTPase-activating protein BCR and result in T-lymphocytic and myeloid leukemias; myeloid and lymphoid neoplasms;

and atypical chronic myeloid leukemia, respectively (Vainchenker and Constantinescu, 2013).

The MPNs, polycythemia vera (PCV), essential thrombocythemia (ET), and idiopathic myelofibrosis (MF) display a high frequency of somatically acquired activating mutations in JAK2. The most common mutations result in the V617F point mutation in the pseudokinase domain or are exon 12 mutations in the N-terminally adjacent domain of JAK2 (~100% in PCV, ~60% in ET, and ~60% in MF). These mutations result in a constitutively active kinase, although *in vitro* studies suggest that an appropriate homodimeric cytokine receptor (EPO, thrombopoietin, or G-CSF receptor) is required to activate mutant JAK2 and perhaps to scaffold other essential signaling molecules like STAT5. Additional mutations in negative regulators of JAK2 such as the lymphocyte-specific adaptor protein LNK or calreticulin mutations (in ~20% of ET and 20% of MF) are present in MPNs and result in constitutive JAK2 activation although the mechanisms are not clear (Cazzola and Kralovics, 2014). There is some evidence that particular mutations have selective effects on the disease phenotype, progression or outcome. For example, the JAK2 exon 12 mutations result in almost exclusive erythrocytosis (Scott *et al.*, 2007) and progression from ET to PV or PV to MF is associated with an increase in mutant allele burden (progression to homozygosity) (Cazzola and Kralovics, 2014). More rare activating mutations in the JH2 domain, such as point mutations of R683, have been associated with about 20% of acute lymphoblastic leukemias seen in patients with Down's syndrome (Mullighan and Downing, 2009).

Somatic activating mutations in the JAK1 JH2 domain have been detected in up to 20% of adult T-cell acute lymphoblastic leukemia (T-cell ALL) and less commonly in B-cell ALL (Flex *et al.*, 2008; Jeong *et al.*, 2008). Mutations in JAK1 have also been detected in AML and non-hemopoietic tumors but at a lower frequency. Finally, activating mutations located in the FERM or pseudokinase domains of JAK3 have also been detected in acute megakaryocytic leukemia (Walters *et al.*, 2006).

Involvement of STAT Proteins in Human Disease

A heterozygous dominant negative mutation in STAT1 was first described in a child susceptible to mycobacterial but not viral infections (Dupuis *et al.*, 2001) while homozygous STAT1 deficiency in other infants resulted in susceptibility to both infections (Dupuis *et al.*, 2003). This was due to failed IFN- γ signaling in the former case but failed IFN γ and IFN α signaling in the latter. Perhaps surprisingly, activating mutations in STAT1 have been associated with fungal infections possibly due to excessive IFN γ signaling but diminished IL-17 responses (Liu *et al.*, 2011; van de Veerdonk *et al.*, 2011).

Dominant negative mutations have also been found for STAT3 and result in hyper IgE syndrome with increased infections and some bone defects (Minegishi *et al.*, 2007; Holland *et al.*, 2007). Since STAT3 is essential for Th17 responses some of these defects appear to be due to diminished IL-17 responses (Milner *et al.*, 2008). Homozygous inactivating mutation of STAT5b in a child demonstrated growth retardation due to growth hormone insensitivity and autoimmune defects, apparently due to increased Th17 responses

resulting from lack of inhibition by IL-2 and reduced regulatory T cells (Kofoed *et al.*, 2003).

Constitutive activation of STAT3 has been associated with numerous solid (epithelial, liver, breast, prostate and melanoma) and hemopoietic tumors (CML, AML, CLL, ALL, multiple myeloma) (Bromberg, 2002). This probably reflects the common activation of STAT3 by growth-promoting tyrosine kinase receptors and by the tumor-promoting TH17-type inflammatory response in tumors as well as angiogenic signaling by STAT3. In contrast STAT1 activation is thought to promote antitumor immunity through a TH1-type response and is less frequently activated in tumors. Moreover, only STAT3 has been shown to be required for cancer cell line proliferation for both solid tumors and leukemias (Bromberg, 2002). On the other hand STAT5 seems to be the major STAT required for proliferation of myeloproliferative neoplasms induced by activating JAK2 mutations (Funakoshi-Tago *et al.*, 2010).

Activating STAT3 mutations (usually in the SH2 domain required for STAT dimerization) are relatively common (30–40%) in T cell large granular lymphocytic leukemia and NK cell leukemias (Ohgami *et al.*, 2013; Koskela *et al.*, 2012; Jerez *et al.*, 2012) and at lower frequency in a subset of diffuse B-cell lymphomas (Ding *et al.*, 2008).

Involvement of SOCS Proteins in Human Disease

Although there have been reports of increased expression of SOCS proteins associated with several diseases and in several tumor types (Trengeve and Ward, 2013) these studies are difficult to interpret because the expression of SOCS proteins is induced by cytokine signaling and in particular by STAT activation. Consequently, increased expression of SOCS may simply reflect activation of STAT pathways by alterations in the cytokine milieu or primary mutations in receptors or oncogenes that activate the JAK/STAT pathway with no causative relationship to disease. Similarly, although enforced expression of a SOCS protein may ameliorate disease this may represent an ability to dampen JAK/STAT signaling rather than reflecting loss of SOCS expression as a cause of disease.

It is more convincing that several studies have revealed a loss of SOCS1 or SOCS3 expression in a range of tumors and have correlated this with specific epigenetic regulation at the locus by bi-allelic methylation of the respective promoters or enhancers. These include high-frequency methylation of the SOCS1 gene in hepatocellular carcinoma (HCC 69%) (Yang *et al.*, 2003; Yoshikawa *et al.*, 2001), gastric carcinoma (GC 44%) (Oshimo *et al.*, 2004), myelodysplastic syndrome (MDS 31%) (Brakensiek *et al.*, 2005), ovarian cancer (23%) (Sutherland *et al.*, 2004), and pancreatic ductal adenocarcinoma (22%) (Fukushima *et al.*, 2003). Similarly, methylation of the SOCS3 gene has been reported in head and neck squamous cell carcinoma (90%) (Weber *et al.*, 2005), non-small cell lung cancer (80%) (He *et al.*, 2004), and HCC (30%) (Niwa *et al.*, 2005).

Inactivating somatic mutations in SOCS genes are rare but a significant proportion of classical Hodgkin Lymphoma patients (8/19) displayed a variety of inactivating mutations in SOCS1 resulting in enhanced nuclear accumulation of STAT5

(Weniger *et al.*, 2006) as did 9/20 primary mediastinal B-cell lymphomas (Melzner *et al.*, 2005).

Concluding Remarks

It has been quite a long journey over the last 20 years, as the intricacies of the JAK–STAT–SOCS network and the exquisite cellular control of this fascinating pathway have unfolded. Clearly there are many layers of regulation, extending from suppression of JAK activation by its pseudokinase domain, phosphorylation, and de-phosphorylation of the STATs and induction of the signaling suppressors, SOCS1 and SOCS3, which then act to directly inhibit JAK catalytic activity. The complexity of STAT action is only just beginning to be appreciated and includes their transit into and out of the nucleus and distinct roles in mitochondria. We are quite sure that there is still more to be discovered and eagerly await the next chapter in the story. Together, the multilayered regulation of the JAK–STAT pathway offers many entry points for therapeutic intervention in hematological diseases and solid tumors.

Acknowledgments

This work was supported in part by the National Health and Medical Research Council (NHMRC), Australia (Program grant #1016647), as well as an NHMRC IRISS grant 361646 and a Victorian State Government Operational Infrastructure Scheme grant. N.A.N. acknowledges fellowship support from the NHMRC, while J.M.M. and J.J.B. are supported by fellowships from the Australian Research Council. This work was also supported in part by the National Institutes of Health (RO1 CA22556-26).

See also: Cell Communication: Cellular Machineries: Macromolecular Communication between Nucleus and Cytoplasm. Cell Communication: Growth Factor Mediated Cell Signaling: Cytokine Receptors and Their Ligands. Cell Communication: Intracellular Pathways: SH3 and SH2: Prototypic Domains to Mediate Regulatory Mechanisms in the Cell. Protein Synthesis/Degradation: Protein Degradation – Intracellular: Ubiquitin, Ubiquitin-Like Proteins, and Proteasome-Mediated Degradation

References

- Akada, H., Akada, S., Hutchison, R.E., *et al.*, 2014. Critical role of Jak2 in the maintenance and function of adult hematopoietic stem cells. *Stem Cells* 32, 1878–1889.
- Akimoto, T., Numata, F., Tamura, M., *et al.*, 1998. Abrogation of bronchial eosinophilic inflammation and airway hyperactivity in signal transducers and activators of transcription (STAT)6-deficient mice. *Journal of Experimental Medicine* 187, 1537–1542.
- Akira, S., Nishio, Y., Inoue, M., *et al.*, 1994. Molecular cloning of APRF, a novel IFN-stimulated gene factor 3 p91-related transcription factor involved in the gp130-mediated signaling pathway. *Cell* 77, 63–71.
- Alexander, W.S., Starr, R., Fenner, J.E., *et al.*, 1999. SOCS1 is a critical inhibitor of interferon gamma signaling and prevents the potentially fatal neonatal actions of this cytokine. *Cell* 98, 597–608.
- Babon, J.J., Kershaw, N.J., Murphy, J.M., *et al.*, 2012. Suppression of cytokine signaling by SOCS3: Characterization of the mode of inhibition and the basis of its specificity. *Immunity* 36, 239–250.
- Babon, J.J., McManus, E.J., Yao, S., *et al.*, 2006. The structure of SOCS3 reveals the basis of the extended SH2 domain function and identifies an unstructured insertion that regulates stability. *Molecular Cell* 22, 205–216.
- Babon, J.J., Nicola, N.A., 2012. The biology and mechanism of action of suppressor of cytokine signaling 3. *Growth Factors* 30, 207–219.
- Babon, J.J., Sabo, J.K., Soetopo, A., *et al.*, 2008. The SOCS box domain of SOCS3: Structure and interaction with the elonginBC-cullin5 ubiquitin ligase. *Journal of Molecular Biology* 381, 928–940.
- Babon, J.J., Sabo, J.K., Zhang, J.G., Nicola, N.A., Norton, R.S., 2009. The SOCS box encodes a hierarchy of affinities for Cullin5: Implications for ubiquitin ligase formation and cytokine signalling suppression. *Journal of Molecular Biology* 387, 162–174.
- Babon, J.J., Yao, S., DeSouza, D.P., *et al.*, 2005. Secondary structure assignment of mouse SOCS3 by NMR defines the domain boundaries and identifies an unstructured insertion in the SH2 domain. *FEBS Journal* 272, 6120–6130.
- Bandaranayake, R.M., Ungureanu, D., Shan, Y., *et al.*, 2012. Crystal structures of the JAK2 pseudokinase domain and the pathogenic mutant V617F. *Nature Structural & Molecular Biology* 19 (8), 754–759.
- Banks, A.S., Li, J., McKeag, L., *et al.*, 2005. Deletion of SOCS7 leads to enhanced insulin action and enlarged islets of Langerhans. *Journal of Clinical Investigation* 115, 2462–2471.
- Bates, S.H., Stearns, W.H., Dundon, T.A., *et al.*, 2003. STAT3 signalling is required for leptin regulation of energy balance but not reproduction. *Nature* 421, 856–859.
- Baxter, E.J., Scott, L.M., Campbell, P.J., *et al.*, 2005. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. *Lancet* 365, 1054–1061.
- Bayle, J., Letard, S., Frank, R., Dubreuil, P., De Sepulveda, P., 2004. Suppressor of cytokine signaling 6 associates with KIT and regulates KIT receptor signaling. *Journal of Biological Chemistry* 279, 12249–12259.
- Begitt, A., Droscher, M., Meyer, T., *et al.*, 2014. STAT1-cooperative DNA binding distinguishes type 1 from type 2 interferon signaling. *Nature Immunology* 15, 168–176.
- Blouin, J., Roby, P., Arcand, M., Beaudet, L., Lipari, F., 2011. Catalytic specificity of human protein tyrosine kinases revealed by Peptide substrate profiling. *Current Chemical Genomics* 5, 115–121.
- Boyle, K., Egan, P., Rakar, S., *et al.*, 2007. The SOCS box of suppressor of cytokine signaling-3 contributes to the control of G-CSF responsiveness *in vivo*. *Blood* 110, 1466–1474.
- Boyle, K., Zhang, J.G., Nicholson, S.E., *et al.*, 2009. Deletion of the SOCS box of suppressor of cytokine signaling 3 (SOCS3) in embryonic stem cells reveals SOCS box-dependent regulation of JAK but not STAT phosphorylation. *Cell Signal* 21, 394–404.
- Brakensiek, K., Langer, F., Schlegelberger, B., Kreipe, H., Lehmann, U., 2005. Hypermethylation of the suppressor of cytokine signalling-1 (SOCS-1) in myelodysplastic syndrome. *British Journal of Haematology* 130, 209–217.
- Brender, C., Columbus, R., Metcalf, D., *et al.*, 2004. SOCS5 is expressed in primary B and T lymphoid cells but is dispensable for lymphocyte production and function. *Molecular and Cellular Biology* 24, 6094–6103.
- Bromberg, J., 2002. Stat proteins and oncogenesis. *Journal of Clinical Investigation* 109, 1139–1142.
- Brooks, A.J., Dai, W., O'Mara, M.L., *et al.*, 2014. Mechanism of activation of protein kinase JAK2 by the growth hormone receptor. *Science* 344, 1249783.
- Bullock, A.N., Debreczeni, J.E., Edwards, A.M., Sundstrom, M., Knapp, S., 2006. Crystal structure of the SOCS2-elongin C-elongin B complex defines a prototypical SOCS box ubiquitin ligase. *Proceedings of the National Academy of Sciences of the United States of America* 103, 7637–7642.
- Bullock, A.N., Rodriguez, M.C., Debreczeni, J.E., Songyang, Z., Knapp, S., 2007. Structure of the SOCS4-ElonginB/C complex reveals a distinct SOCS box interface and the molecular basis for SOCS-dependent EGFR degradation. *Structure* 15, 1493–1504.
- Callus, B.A., Mathey-Prevot, B., 2002. SOCS36E, a novel Drosophila SOCS protein, suppresses JAK/STAT and EGF-R signalling in the imaginal wing disc. *Oncogene* 21, 4812–4821.
- Cazzola, M., Kralovics, R., 2014. From Janus kinase 2 to calreticulin: The clinically relevant genomic landscape of myeloproliferative neoplasms. *Blood* 123, 3714–3719.

- Chatti, K., Farrar, W.L., Duhé, R.J., 2004. Tyrosine phosphorylation of the Janus kinase 2 activation loop is essential for a high-activity catalytic state but dispensable for a basal catalytic state. *Biochemistry* 43, 4272–4283.
- Chen, X., Vinkemeier, U., Zhao, Y., *et al.*, 1998. Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA. *Cell* 93, 827–839.
- Cheon, H., Stark, G.R., 2009. Unphosphorylated STAT1 prolongs the expression of interferon-induced immune regulatory genes. *Proceedings of the National Academy of Sciences of the United States of America* 106, 9373–9378.
- Chong, M.M., Cornish, A.L., Darwiche, R., *et al.*, 2003. Suppressor of cytokine signaling-1 is a critical regulator of interleukin-7-dependent CD8+ T cell differentiation. *Immunity* 18, 475–487.
- Cornish, A.L., Chong, M.M., Davey, G.M., *et al.*, 2003. Suppressor of cytokine signaling-1 regulates signaling in response to interleukin-2 and other gamma c-dependent cytokines in peripheral T cells. *Journal of Biological Chemistry* 278, 22755–22761.
- Crocker, B.A., Krebs, D.L., Zhang, J.G., *et al.*, 2003. SOCS3 negatively regulates IL-6 signaling *in vivo*. *Nature Immunology* 4, 540–545.
- Crocker, B.A., Metcalf, D., Robb, L., *et al.*, 2004. SOCS3 is a critical physiological negative regulator of G-CSF signaling and emergency granulopoiesis. *Immunity* 20, 153–165.
- Darnell Jr., J.E., Kerr, I.M., Stark, G.R., 1994. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* 264, 1415–1421.
- Davey, G.M., Starr, R., Cornish, A.L., *et al.*, 2005. SOCS-1 regulates IL-15-driven homeostatic proliferation of antigen-naïve CD8 T cells, limiting their autoimmune potential. *Journal of Experimental Medicine* 202, 1099–1108.
- Dawson, M.A., Bannister, A.J., Gottgens, B., *et al.*, 2009. JAK2 phosphorylates histone H3Y41 and excludes HP1alpha from chromatin. *Nature* 461, 819–822.
- De Souza, D., Fabri, L.J., Nash, A., *et al.*, 2002. SH2 domains from suppressor of cytokine signaling-3 and protein tyrosine phosphatase SHP-2 have similar binding specificities. *Biochemistry* 41, 9229–9236.
- Ding, B.B., Yu, J.J., Yu, R.Y., *et al.*, 2008. Constitutively activated STAT3 promotes cell proliferation and survival in the activated B-cell subtype of diffuse large B-cell lymphomas. *Blood* 111, 1515–1523.
- Dupuis, S., Dargemont, C., Fieschi, C., *et al.*, 2001. Impairment of mycobacterial but not viral immunity by a germline human STAT1 mutation. *Science* 293, 300–303.
- Dupuis, S., Jouanguy, E., Al-Hajjar, S., *et al.*, 2003. Impaired response to interferon-alpha/beta and lethal viral disease in human STAT1 deficiency. *Nature Genetics* 33, 388–391.
- Durbin, J.E., Hackenmiller, R., Simon, M.C., Levy, D.E., 1996. Targeted disruption of the mouse Stat1 gene results in compromised innate immunity to viral disease. *Cell* 84, 443–450.
- Ehret, G.B., Reichenbach, P., Schindler, U., *et al.*, 2001. DNA binding specificity of different STAT proteins. Comparison of *in vitro* specificity with natural target sites. *Journal of Biological Chemistry* 276, 6675–6688.
- Endo, T.A., Masuhara, M., Yokouchi, M., *et al.*, 1997. A new protein containing an SH2 domain that inhibits JAK kinases. *Nature* 387, 921–924.
- Ernst, M., Putoczki, T.L., 2012. Stat3: Linking inflammation to (gastrointestinal) tumorigenesis. *Clinical and Experimental Pharmacology and Physiology* 39, 711–718.
- Eyles, J.L., Metcalf, D., Grusby, M.J., Hilton, D.J., Starr, R., 2002. Negative regulation of interleukin-12 signaling by suppressor of cytokine signaling-1. *Journal of Biological Chemistry* 277, 43735–43740.
- Feng, J., Witthuhn, B.A., Matsuda, T., *et al.*, 1997. Activation of Jak2 catalytic activity requires phosphorylation of Y1007 in the kinase activation loop. *Molecular and Cellular Biology* 17, 2497–2501.
- Feng, Z.P., Chandrashekar, I.R., Low, A., *et al.*, 2012. The N-terminal domains of SOCS proteins: A conserved region in the disordered N-termini of SOCS4 and 5. *Proteins* 80, 946–957.
- Fenner, J.E., Starr, R., Cornish, A.L., *et al.*, 2006. Suppressor of cytokine signaling 1 regulates the immune response to infection by a unique inhibition of type I interferon activity. *Nature Immunology* 7, 33–39.
- Fink, K., Grandvaux, N., 2013. STAT2 and IRF9: Beyond ISGF3. *JAKSTAT* 2, e27521.
- Firmbach-Kraft, I., Byers, M., Shows, T., Dalla-Favera, R., Krolewski, J.J., 1990. tyk2, prototype of a novel class of non-receptor tyrosine kinase genes. *Oncogene* 5, 1329–1336.
- Flex, E., Petrangeli, V., Stella, L., *et al.*, 2008. Somatic acquired JAK1 mutations in adult acute lymphoblastic leukemia. *Journal of Experimental Medicine* 205, 751–758.
- Fukushima, N., Sato, N., Sahin, F., *et al.*, 2003. Aberrant methylation of suppressor of cytokine signalling-1 (SOCS-1) gene in pancreatic ductal neoplasms. *British Journal of Cancer* 89, 338–343.
- Funakoshi-Tago, M., Tago, K., Abe, M., Sonoda, Y., Kasahara, T., 2010. STAT5 activation is critical for the transformation mediated by myeloproliferative disorder-associated JAK2 V617F mutant. *Journal of Biological Chemistry* 285, 5296–5307.
- Gingras, S., Parganas, E., de Pauw, A., Ihle, J.N., Murray, P.J., 2004. Re-examination of the role of suppressor of cytokine signaling 1 (SOCS1) in the regulation of toll-like receptor signaling. *Journal of Biological Chemistry* 279, 54702–54707.
- Gough, D.J., Corlett, A., Schlessinger, K., *et al.*, 2009. Mitochondrial STAT3 supports Ras-dependent oncogenic transformation. *Science* 324, 1713–1716.
- Greenhalgh, C.J., Rico-Bautista, E., Lorentzon, M., *et al.*, 2005. SOCS2 negatively regulates growth hormone action *in vitro* and *in vivo*. *Journal of Clinical Investigation* 115, 397–406.
- Greenlund, A.C., Morales, M.O., Viviano, B.L., *et al.*, 1995. Stat recruitment by tyrosine-phosphorylated cytokine receptors: An ordered reversible affinity-driven process. *Immunity* 2, 677–687.
- Grisouard, J., Hao-Shen, H., Dirnhofer, S., Wagner, K.U., Skoda, R.C., 2014. Selective deletion of Jak2 in adult mouse hematopoietic cells leads to lethal anemia and thrombocytopenia. *Haematologica* 99, e52–e54.
- Hahn, B., Trifilo, M.J., Zuniga, E.I., Oldstone, M.B., 2005. Viruses evade the immune system through type I interferon-mediated STAT2-dependent, but STAT1-independent, signaling. *Immunity* 22, 247–257.
- Harpur, A.G., Andres, A.C., Ziemiecki, A., Aston, R.R., Wilks, A.F., 1992. JAK2, a third member of the JAK family of protein tyrosine kinases. *Oncogene* 7, 1347–1353.
- He, B., You, L., Xu, Z., *et al.*, 2004. Activity of the suppressor of cytokine signaling-3 promoter in human non-small-cell lung cancer. *Clinical Lung Cancer* 5, 366–370.
- Heim, M.H., Kerr, I.M., Stark, G.R., Darnell Jr., J.E., 1995. Contribution of STAT SH2 groups to specific interferon signaling by the Jak-STAT pathway. *Science* 267, 1347–1349.
- Hilfiker-Kleiner, D., Hilfiker, A., Fuchs, M., *et al.*, 2004. Signal transducer and activator of transcription 3 is required for myocardial capillary growth, control of interstitial matrix deposition, and heart protection from ischemic injury. *Circulation Research* 95, 187–195.
- Hilton, D.J., Richardson, R.T., Alexander, W.S., *et al.*, 1998. Twenty proteins containing a C-terminal SOCS box form five structural classes. *Proceedings of the National Academy of Sciences of the United States of America* 95, 114–119.
- ten Hoeve, J., de Jesus Ibarra-Sanchez, M., Fu, Y., *et al.*, 2002. Identification of a nuclear Stat1 protein tyrosine phosphatase. *Molecular and Cellular Biology* 22, 5662–5668.
- Hoey, T., Zhang, S., Schmidt, N., *et al.*, 2003. Distinct requirements for the naturally occurring splice forms Stat4alpha and Stat4beta in IL-12 responses. *EMBO Journal* 22, 4237–4248.
- Hokuto, I., Ikegami, M., Yoshida, M., *et al.*, 2004. Stat-3 is required for pulmonary homeostasis during hyperoxia. *Journal of Clinical Investigation* 113, 28–37.
- Holland, S.M., DeLeo, F.R., Elloumi, H.Z., *et al.*, 2007. STAT3 mutations in the hyper-IgE syndrome. *New England Journal of Medicine* 357, 1608–1619.
- Hortner, M., Nielsch, U., Mayr, L.M., *et al.*, 2002. Suppressor of cytokine signaling-3 is recruited to the activated granulocyte-colony stimulating factor receptor and modulates its signal transduction. *Journal of Immunology* 169, 1219–1227.
- Hou, J., Schindler, U., Henzel, W.J., *et al.*, 1994. An interleukin-4-induced transcription factor: IL-4 Stat. *Science* 265, 1701–1706.
- Hu, X., Dutta, P., Tsurumi, A., *et al.*, 2013. Unphosphorylated STAT5A stabilizes heterochromatin and suppresses tumor growth. *Proceedings of the National Academy of Sciences of the United States of America* 110, 10213–10218.
- Irlandoust, M.I., Aarts, L.H., Roovers, O., *et al.*, 2007. Suppressor of cytokine signaling 3 controls lysosomal routing of G-CSF receptor. *EMBO Journal* 26, 1782–1793.
- Ishida-Takahashi, R., Rosario, F., Gong, Y., *et al.*, 2006. Phosphorylation of Jak2 on Ser(523) inhibits Jak2-dependent leptin receptor signaling. *Molecular and Cellular Biology* 26, 4063–4073.
- Ishizaki, M., Muromoto, R., Akimoto, T., *et al.*, 2014. Tyk2 is a therapeutic target for psoriasis-like skin inflammation. *International Immunology* 26, 257–267.
- Jacoby, J.J., Kalinowski, A., Liu, M.G., *et al.*, 2003. Cardiomyocyte-restricted knockout of STAT3 results in higher sensitivity to inflammation, cardiac fibrosis, and heart failure with advanced age. *Proceedings of the National Academy of Sciences of the United States of America* 100, 12929–12934.

- James, C., Ugo, V., Le Couedic, J.P., *et al.*, 2005. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature* 434, 1144–1148.
- Jatiani, S.S., Baker, S.J., Silverman, L.R., Reddy, E.P., 2010. Jak/STAT pathways in cytokine signaling and myeloproliferative disorders: Approaches for targeted therapies. *Genes & Cancer* 1, 979–993.
- Jeong, E.G., Kim, M.S., Nam, H.K., *et al.*, 2008. Somatic mutations of JAK1 and JAK3 in acute leukemias and solid cancers. *Clinical Cancer Research* 14, 3716–3721.
- Jerez, A., Clemente, M.J., Makishima, H., *et al.*, 2012. STAT3 mutations unify the pathogenesis of chronic lymphoproliferative disorders of NK cells and T-cell large granular lymphocyte leukemia. *Blood* 120, 3048–3057.
- Johnston, J.A., Kawamura, M., Kirken, R.A., *et al.*, 1994. Phosphorylation and activation of the Jak-3 Janus kinase in response to interleukin-2. *Nature* 370, 151–153.
- Kamizono, S., Hanada, T., Yasukawa, H., *et al.*, 2001. The SOCS box of SOCS-1 accelerates ubiquitin-dependent proteolysis of TEL-JAK2. *Journal of Biological Chemistry* 276, 12530–12538.
- Kamura, T., Maenaka, K., Kotoshiba, S., *et al.*, 2004. VHL-box and SOCS-box domains determine binding specificity for Cul2-Rbx1 and Cul5-Rbx2 modules of ubiquitin ligases. *Genes & Development* 18, 3055–3065.
- Kamura, T., Sato, S., Haque, D., *et al.*, 1998. The Elongin BC complex interacts with the conserved SOCS-box motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families. *Genes & Development* 12, 3872–3881.
- Kaplan, M.H., Sun, Y.L., Hoey, T., Grusby, M.J., 1996. Impaired IL-12 responses and enhanced development of Th2 cells in Stat4-deficient mice. *Nature* 382, 174–177.
- Karaghiossioff, M., Neubauer, H., Lassnig, C., *et al.*, 2000. Partial impairment of cytokine responses in Tyk2-deficient mice. *Immunity* 13, 549–560.
- Karaghiossioff, M., Steinborn, R., Kovarik, P., *et al.*, 2003. Central role for type I interferons and Tyk2 in lipopolysaccharide-induced endotoxin shock. *Nature Immunology* 4, 471–477.
- Kario, E., Marmor, M.D., Adamsky, K., *et al.*, 2005. Suppressors of cytokine signaling 4 and 5 regulate epidermal growth factor receptor signaling. *Journal of Biological Chemistry* 280, 7038–7048.
- Kedzierski, L., Linossi, E.M., Kolesnik, T.B., *et al.*, 2014. Suppressor of cytokine signaling 4 (SOCS4) protects against severe cytokine storm and enhances viral clearance during influenza infection. *PLoS Pathogens* 10, e1004134.
- Kershaw, N.J., Laktuyshin, A., Nicola, N.A., Babon, J.J., 2014. Reconstruction of an active SOCS3-based E3 ubiquitin ligase complex *in vitro*: Identification of the active components and JAK2 and gp130 as substrates. *Growth Factors* 32, 1–10.
- Kershaw, N.J., Murphy, J.M., Liao, N.P., *et al.*, 2013. SOCS3 binds specific receptor-JAK complexes to control cytokine signaling by direct kinase inhibition. *Nature Structural & Molecular Biology* 20, 469–476.
- Khaled, W.T., Read, E.K., Nicholson, S.E., *et al.*, 2007. The IL-4/IL-13/Stat6 signalling pathway promotes luminal mammary epithelial cell development. *Development* 134, 2739–2750.
- Kile, B.T., Schulman, B.A., Alexander, W.S., *et al.*, 2002. The SOCS box: A tale of destruction and degradation. *Trends in Biochemical Sciences* 27, 235–241.
- Kilic, S.S., Hacimustafaoğlu, M., Boisson-Dupuis, S., *et al.*, 2012. A patient with tyrosine kinase 2 deficiency without hyper-IgE syndrome. *Journal of Pediatrics* 160, 1055–1057.
- Kofoed, E.M., Hwa, V., Little, B., *et al.*, 2003. Growth hormone insensitivity associated with a STAT5b mutation. *New England Journal of Medicine* 349, 1139–1147.
- Koppikar, P., Bhagwat, N., Kilpivaara, O., *et al.*, 2012. Heterodimeric JAK-STAT activation as a mechanism of persistence to JAK2 inhibitor therapy. *Nature* 489, 155–159.
- Koskela, H.L., Eldfors, S., Ellonen, P., *et al.*, 2012. Somatic STAT3 mutations in large granular lymphocytic leukemia. *New England Journal of Medicine* 366, 1905–1913.
- Kralovics, R., Passamonti, F., Buser, A.S., *et al.*, 2005. A gain-of-function mutation of JAK2 in myeloproliferative disorders. *New England Journal of Medicine* 352, 1779–1790.
- Krebs, D.L., Metcalf, D., Merson, T.D., *et al.*, 2004. Development of hydrocephalus in mice lacking SOCS7. *Proceedings of the National Academy of Sciences of the United States of America* 101, 15446–15451.
- Krebs, D.L., Uren, R.T., Metcalf, D., *et al.*, 2002. SOCS-6 binds to insulin receptor substrate 4, and mice lacking the SOCS-6 gene exhibit mild growth retardation. *Molecular and Cellular Biology* 22, 4567–4578.
- Krempler, A., Qi, Y., Triplett, A.A., *et al.*, 2004. Generation of a conditional knockout allele for the Janus kinase 2 (Jak2) gene in mice. *Genesis* 40, 52–57.
- Kuperman, D., Schofield, B., Wills-Karp, M., Grusby, M.J., 1998. Signal transducer and activator of transcription factor 6 (Stat6)-deficient mice are protected from antigen-induced airway hyperresponsiveness and mucus production. *Journal of Experimental Medicine* 187, 939–948.
- Landsman, T., Waxman, D.J., 2005. Role of the cytokine-induced SH2 domain-containing protein CIS in growth hormone receptor internalization. *Journal of Biological Chemistry* 280, 37471–37480.
- Lang, R., Pauleau, A.L., Parganas, E., *et al.*, 2003. SOCS3 regulates the plasticity of gp130 signaling. *Nature Immunology* 4, 546–550.
- Lavens, D., Montoye, T., Piessevaux, J., *et al.*, 2006. A complex interaction pattern of CIS and SOCS2 with the leptin receptor. *Journal of Cell Science* 119, 2214–2224.
- Lee, J.E., Yang, Y.M., Liang, F.X., *et al.*, 2012. Nongenomic STAT5-dependent effects on Golgi apparatus and endoplasmic reticulum structure and function. *American Journal of Physiology – Cell Physiology* 302, C804–C820.
- Levine, R.L., Wadleigh, M., Coombs, J., *et al.*, 2005. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. *Cancer Cell* 7, 387–397.
- Levy, D.E., Kessler, D.S., Pine, R., Darnell Jr., J.E., 1989. Cytoplasmic activation of ISGF3, the positive regulator of interferon-alpha-stimulated transcription, reconstituted *in vitro*. *Genes & Development* 3, 1362–1371.
- Levy, D.E., Kessler, D.S., Pine, R., Reich, N., Darnell Jr., J.E., 1988. Interferon-induced nuclear factors that bind a shared promoter element correlate with positive and negative transcriptional control. *Genes & Development* 2, 383–393.
- Linossi, E.M., Chandrashekar, I.R., Kolesnik, T.B., *et al.*, 2013. Suppressor of Cytokine Signaling (SOCS) 5 utilizes distinct domains for regulation of JAK1 and interaction with the adaptor protein Shc-1. *PLoS One* 8, e70536.
- Linossi, E.M., Nicholson, S.E., 2012. The SOCS box-adapting proteins for ubiquitination and proteasomal degradation. *IUBMB Life* 64, 316–323.
- Liongue, C., Ward, A.C., 2013. Evolution of the JAK-STAT pathway. *JAKSTAT* 2, e22756.
- Liu, L., McBride, K.M., Reich, N.C., 2005. STAT3 nuclear import is independent of tyrosine phosphorylation and mediated by importin-alpha3. *Proceedings of the National Academy of Sciences of the United States of America* 102, 8150–8155.
- Liu, L., Okada, S., Kong, X.F., *et al.*, 2011. Gain-of-function human STAT1 mutations impair IL-17 immunity and underlie chronic mucocutaneous candidiasis. *Journal of Experimental Medicine* 208, 1635–1648.
- Liu, X., Robinson, G.W., Gouilleux, F., Groner, B., Hennighausen, L., 1995. Cloning and expression of Stat5 and an additional homologue (Stat5b) involved in prolactin signal transduction in mouse mammary tissue. *Proceedings of the National Academy of Sciences of the United States of America* 92, 8831–8835.
- Liu, X., Robinson, G.W., Wagner, K.U., *et al.*, 1997. Stat5a is mandatory for adult mammary gland development and lactogenesis. *Genes & Development* 11, 179–186.
- Luo, H., Hanratty, W.P., Dearolf, C.R., 1995. An amino acid substitution in the Drosophila hopTum-I Jak kinase causes leukemia-like hematopoietic defects. *EMBO Journal* 14, 1412–1420.
- Lupardus, P.J., Skiniotis, G., Rice, A.J., *et al.*, 2011. Structural snapshots of full-length Jak1, a transmembrane gp130/IL-6/IL-6Ralpha cytokine receptor complex, and the receptor-Jak1 holocomplex. *Structure* 19, 45–55.
- Lupardus, P.J., Ultsch, M., Wallweber, H., *et al.*, 2014. Structure of the pseudokinase-kinase domains from protein kinase TYK2 reveals a mechanism for Janus kinase (JAK) autoinhibition. *Proceedings of the National Academy of Sciences of the United States of America* 111, 8025.
- Macchi, P., Villa, A., Giliani, S., *et al.*, 1995. Mutations of Jak-3 gene in patients with autosomal severe combined immune deficiency (SCID). *Nature* 377, 65–68.
- Mao, X., Ren, Z., Parker, G.N., *et al.*, 2005. Structural bases of unphosphorylated STAT1 association and receptor binding. *Molecular Cell* 17, 761–771.
- Mazurkiewicz-Munoz, A.M., Argetsinger, L.S., Kouadio, J.L., *et al.*, 2006. Phosphorylation of JAK2 at serine 523: A negative regulator of JAK2 that is stimulated by growth hormone and epidermal growth factor. *Molecular and Cellular Biology* 26, 4052–4062.
- Meier, J.A., Larner, A.C., 2014. Toward a new STATe: The role of STATs in mitochondrial function. *Seminars in Immunology* 26, 20–28.
- Melzner, I., Bucur, A.J., Bruderlein, S., *et al.*, 2005. Biallelic mutation of SOCS-1 impairs JAK2 degradation and sustains phospho-JAK2 action in the MedB-1 mediastinal lymphoma line. *Blood* 105, 2535–2542.
- Meraz, M.A., White, J.M., Sheehan, K.C., *et al.*, 1996. Targeted disruption of the Stat1 gene in mice reveals unexpected physiologic specificity in the JAK-STAT signaling pathway. *Cell* 84, 431–442.

- Metcalfe, D., Greenhalgh, C.J., Viney, E., *et al.*, 2000. Gigantism in mice lacking suppressor of cytokine signalling-2. *Nature* 405, 1069–1073.
- Milner, J.D., Brenchley, J.M., Laurence, A., *et al.*, 2008. Impaired T(H)17 cell differentiation in subjects with autosomal dominant hyper-IgE syndrome. *Nature* 452, 773–776.
- Minegishi, Y., Saito, M., Morio, T., *et al.*, 2006. Human tyrosine kinase 2 deficiency reveals its requisite roles in multiple cytokine signals involved in innate and acquired immunity. *Immunity* 25, 745–755.
- Minegishi, Y., Saito, M., Tsuchiya, S., *et al.*, 2007. Dominant-negative mutations in the DNA-binding domain of STAT3 cause hyper-IgE syndrome. *Nature* 448, 1058–1062.
- Mori, H., Hanada, R., Hanada, T., *et al.*, 2004. Socs3 deficiency in the brain elevates leptin sensitivity and confers resistance to diet-induced obesity. *Nature Medicine* 10, 739–743.
- Mui, A.L., Wakao, H., O'Farrell, A.M., Harada, N., Miyajima, A., 1995. Interleukin-3, granulocyte-macrophage colony stimulating factor and interleukin-5 transduce signals through two STAT5 homologs. *EMBO Journal* 14, 1166–1175.
- Muller, M., Briscoe, J., Laxton, C., *et al.*, 1993. The protein tyrosine kinase JAK1 complements defects in interferon- α / β and - γ signal transduction. *Nature* 366, 129–135.
- Mullighan, C.G., Downing, J.R., 2009. Genome-wide profiling of genetic alterations in acute lymphoblastic leukemia: Recent insights and future directions. *Leukemia* 23, 1209–1218.
- Muniz, J.R., Guo, K., Kershaw, N.J., *et al.*, 2013. Molecular architecture of the ankyrin SOCS box family of Cul5-dependent E3 ubiquitin ligases. *Journal of Molecular Biology* 425, 3166–3177.
- Murphy, J.M., Zhang, Q., Young, S.N., *et al.*, 2014. A robust methodology to subclassify pseudokinases based on their nucleotide-binding properties. *Biochemical Journal* 457, 323–334.
- Musso, T., Johnston, J.A., Linnekin, D., *et al.*, 1995. Regulation of JAK3 expression in human monocytes: Phosphorylation in response to interleukins 2, 4, and 7. *Journal of Experimental Medicine* 181, 1425–1431.
- Naka, T., Narazaki, M., Hirata, M., *et al.*, 1997. Structure and function of a new STAT-induced STAT inhibitor. *Nature* 387, 924–929.
- Naka, T., Tsutsui, H., Fujimoto, M., *et al.*, 2001. SOCS-1/SSI-1-deficient NKT cells participate in severe hepatitis through dysregulated cross-talk inhibition of IFN- γ and IL-4 signaling *in vivo*. *Immunity* 14, 535–545.
- Narazaki, M., Fujimoto, M., Matsumoto, T., *et al.*, 1998. Three distinct domains of SSI-1/SOCS-1/JAB protein are required for its suppression of interleukin 6 signaling. *Proceedings of the National Academy of Sciences of the United States of America* 95, 13130–13134.
- Neubauer, H., Cumano, A., Muller, M., *et al.*, 1998. Jak2 deficiency defines an essential developmental checkpoint in definitive hematopoiesis. *Cell* 93, 397–409.
- Nicholson, S.E., De Souza, D., Fabri, L.J., *et al.*, 2000. Suppressor of cytokine signalling-3 preferentially binds to the SHP-2-binding site on the shared cytokine receptor subunit gp130. *Proceedings of the National Academy of Sciences of the United States of America* 97, 6493–6498.
- Nicholson, S.E., Metcalfe, D., Sprigg, N.S., *et al.*, 2005. Suppressor of cytokine signalling (SOCS)-5 is a potential negative regulator of epidermal growth factor signaling. *Proceedings of the National Academy of Sciences of the United States of America* 102, 2328–2333.
- Niwa, Y., Kanda, H., Shikauchi, Y., *et al.*, 2005. Methylation silencing of SOCS-3 promotes cell growth and migration by enhancing JAK/STAT and FAK signalings in human hepatocellular carcinoma. *Oncogene* 24, 6406–6417.
- Nkansah, E., Shah, R., Collie, G.W., *et al.*, 2013. Observation of unphosphorylated STAT3 core protein binding to target dsDNA by PEMS and X-ray crystallography. *FEBS Letters* 587, 833–839.
- Nosaka, T., van Deursen, J.M., Tripp, R.A., *et al.*, 1995. Defective lymphoid development in mice lacking Jak3. *Science* 270, 800–802.
- O'Shea, J.J., Holland, S.M., Staudt, L.M., 2013. JAKs and STATs in immunity, immunodeficiency, and cancer. *New England Journal of Medicine* 368, 161–170.
- Ohgami, R.S., Ma, L., Merker, J.D., *et al.*, 2013. STAT3 mutations are frequent in CD30+ T-cell lymphomas and T-cell large granular lymphocytic leukemia. *Leukemia* 27, 2244–2247.
- Okada, S., Nakamura, M., Katoh, H., *et al.*, 2006. Conditional ablation of Stat3 or Socs3 discloses a dual role for reactive astrocytes after spinal cord injury. *Nature Medicine* 12, 829–834.
- Oshimo, Y., Kuraoka, K., Nakayama, H., *et al.*, 2004. Epigenetic inactivation of SOCS-1 by CpG island hypermethylation in human gastric carcinoma. *International Journal of Cancer* 112, 1003–1009.
- Parganas, E., Wang, D., Stravopodis, D., *et al.*, 1998. Jak2 is essential for signaling through a variety of cytokine receptors. *Cell* 93, 385–395.
- Park, C., Li, S., Cha, E., Schindler, C., 2000. Immune response in Stat2 knockout mice. *Immunity* 13, 795–804.
- Park, S.O., Wamsley, H.L., Bae, K., *et al.*, 2013. Conditional deletion of Jak2 reveals an essential role in hematopoiesis throughout mouse ontogeny: Implications for Jak2 inhibition in humans. *PLoS One* 8, e59675.
- Park, S.Y., Saijo, K., Takahashi, T., *et al.*, 1995. Developmental defects of lymphoid cells in Jak3 kinase-deficient mice. *Immunity* 3, 771–782.
- Perry, S.T., Buck, M.D., Lada, S.M., Schindler, C., Shresta, S., 2011. STAT2 mediates innate immunity to Dengue virus in the absence of STAT1 via the type I interferon receptor. *PLoS Pathogens* 7, e1001297.
- Qing, Y., Costa-Pereira, A.P., Watling, D., Stark, G., 2005. Role of tyrosine 441 of interferon- γ receptor subunit 1 in SOCS-1-mediated attenuation of STAT1 activation. *Journal of Biological Chemistry* 280, 1849–1853.
- Ram, P.A., Waxman, D.J., 2000. Role of the cytokine-inducible SH2 protein CIS in desensitization of STAT5b signaling by continuous growth hormone. *Journal of Biological Chemistry* 275, 39487–39496.
- Rane, S.G., Reddy, E.P., 1994. JAK3: A novel JAK kinase associated with terminal differentiation of hematopoietic cells. *Oncogene* 9, 2415–2423.
- Rawlings, J.S., Rennebeck, G., Harrison, S.M., Xi, R., Harrison, D.A., 2004. Two Drosophila suppressors of cytokine signaling (SOCS) differentially regulate JAK and EGFR pathway activities. *BMC Cell Biology* 5, 38.
- Reich, N.C., 2013. STATs get their move on. *JAKSTAT* 2, e27080.
- Robb, L., Boyle, K., Rakar, S., *et al.*, 2005. Genetic reduction of embryonic leukemia-inhibitory factor production rescues placentalization in SOCS3-null embryos but does not prevent inflammatory disease. *Proceedings of the National Academy of Sciences of the United States of America* 102, 16333–16338.
- Roberts, A.W., Robb, L., Rakar, S., *et al.*, 2001. Placental defects and embryonic lethality in mice lacking suppressor of cytokine signalling (SOCS)-3. *Proceedings of the National Academy of Sciences* 98, 9324–9329.
- Rodig, S.J., Meraz, M.A., White, J.M., *et al.*, 1998. Disruption of the Jak1 gene demonstrates obligatory and nonredundant roles of the Jaks in cytokine-induced biologic responses. *Cell* 93, 373–383.
- Russell, S.M., Tayebi, N., Nakajima, H., *et al.*, 1995. Mutation of Jak3 in a patient with SCID: Essential role of Jak3 in lymphoid development. *Science* 270, 797–800.
- Saharinen, P., Silvennoinen, O., 2002. The pseudokinase domain is required for suppression of basal activity of Jak2 and Jak3 tyrosine kinases and for cytokine-inducible activation of signal transduction. *Journal of Biological Chemistry* 277, 47954–47963.
- Saharinen, P., Takaluoma, K., Silvennoinen, O., 2000. Regulation of the Jak2 tyrosine kinase by its pseudokinase domain. *Molecular and Cellular Biology* 20, 3387–3395.
- Saharinen, P., Vihinen, M., Silvennoinen, O., 2003. Autoinhibition of Jak2 tyrosine kinase is dependent on specific regions in its pseudokinase domain. *Molecular Biology of the Cell* 14, 1448–1459.
- Sano, S., Itami, S., Takeda, K., *et al.*, 1999. Keratinocyte-specific ablation of Stat3 exhibits impaired skin remodeling, but does not affect skin morphogenesis. *EMBO Journal* 18, 4657–4668.
- Sanz, A., Ungureanu, D., Pekala, T., *et al.*, 2011. Analysis of Jak2 catalytic function by peptide microarrays: The role of the JH2 domain and V617F mutation. *PLoS One* 6, e18522.
- Sasaki, A., Yasukawa, H., Suzuki, A., *et al.*, 1999. Cytokine-inducible SH2 protein-3 (CIS3/SOCS3) inhibits Janus tyrosine kinase by binding through the N-terminal kinase inhibitory region as well as SH2 domain. *Genes to Cells* 4, 339–351.
- Schindler, C., Fu, X.Y., Imbrota, T., Aebersold, R., Darnell Jr., J.E., 1992. Proteins of transcription factor ISGF-3: One gene encodes the 91- and 84-kDa ISGF-3 proteins that are activated by interferon α . *Proceedings of the National Academy of Sciences of the United States of America* 89, 7836–7839.
- Schindler, C., Shuai, K., Prezioso, V.R., Darnell Jr., J.E., 1992. Interferon-dependent tyrosine phosphorylation of a latent cytoplasmic transcription factor. *Science* 257, 809–813.
- Schindler, U., Wu, P., Rothe, M., Brasseur, M., McKnight, S.L., 1995. Components of a Stat recognition code: Evidence for two layers of molecular selectivity. *Immunity* 2, 689–697.
- Scott, L.M., Tong, W., Levine, R.L., *et al.*, 2007. JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. *New England Journal of Medicine* 356, 459–468.
- Seki, Y., Hayashi, K., Matsumoto, A., *et al.*, 2002. Expression of the suppressor of cytokine signaling-5 (SOCS5) negatively regulates IL-4-dependent STAT6

- activation and Th2 differentiation. *Proceedings of the National Academy of Sciences of the United States of America* 99, 13003–13008.
- Seubert, N., Royer, Y., Staerk, J., *et al.*, 2003. Active and inactive orientations of the transmembrane and cytosolic domains of the erythropoietin receptor dimer. *Molecular Cell* 12, 1239–1250.
- Shaw, M.H., Boyartchuk, V., Wong, S., *et al.*, 2003. A natural mutation in the Tyk2 pseudokinase domain underlies altered susceptibility of B10.Q/J mice to infection and autoimmunity. *Proceedings of the National Academy of Sciences of the United States of America* 100, 11594–11599.
- Shi, S., Larson, K., Guo, D., *et al.*, 2008. Drosophila STAT is required for directly maintaining HP1 localization and heterochromatin stability. *Nature Cell Biology* 10, 489–496.
- Shimoda, K., Kato, K., Aoki, K., *et al.*, 2000. Tyk2 plays a restricted role in IFN alpha signaling, although it is required for IL-12-mediated T cell function. *Immunity* 13, 561–571.
- Shimoda, K., van Deursen, J., Sangster, M.Y., *et al.*, 1996. Lack of IL-4-induced Th2 response and IgE class switching in mice with disrupted Stat6 gene. *Nature* 380, 630–633.
- Shuai, K., Stark, G.R., Kerr, I.M., Darnell Jr., J.E., 1993. A single phosphotyrosine residue of Stat91 required for gene activation by interferon-gamma. *Science* 261, 1744–1746.
- Socolovsky, M., Fallon, A.E., Wang, S., Brugnara, C., Lodish, H.F., 1999. Fetal anemia and apoptosis of red cell progenitors in Stat5a/-5b/- mice: A direct role for Stat5 in Bcl-X(L) induction. *Cell* 98, 181–191.
- Socolovsky, M., Nam, H., Fleming, M.D., *et al.*, 2001. Ineffective erythropoiesis in Stat5a(-/-)5b(-/-) mice due to decreased survival of early erythroblasts. *Blood* 98, 3261–3273.
- Staerk, J., Defour, J.P., Pecquet, C., *et al.*, 2011. Orientation-specific signalling by thrombopoietin receptor dimers. *EMBO Journal* 30, 4398–4413.
- Starr, R., Metcalf, D., Elefanti, A.G., *et al.*, 1998. Liver degeneration and lymphoid deficiencies in mice lacking suppressor of cytokine signaling-1. *Proceedings of the National Academy of Sciences of the United States of America* 95, 14395–14399.
- Starr, R., Willson, T.A., Viney, E.M., *et al.*, 1997. A family of cytokine-inducible inhibitors of signalling. *Nature* 387, 917–921.
- Stec, W.J., Zeidler, M.P., 2011. Drosophila SOCS proteins. *Journal of Signal Transduction* 2011, 894510.
- Sutherland, K.D., Lindeman, G.J., Choong, D.Y., *et al.*, 2004. Differential hypermethylation of SOCS genes in ovarian and breast carcinomas. *Oncogene* 23, 7726–7733.
- Takahashi, T., Shirasawa, T., 1994. Molecular cloning of rat JAK3, a novel member of the JAK family of protein tyrosine kinases. *FEBS Letters* 342, 124–128.
- Takahashi, Y., Carpino, N., Cross, J.C., *et al.*, 2003. SOCS3: An essential regulator of LIF receptor signaling in trophoblast giant cell differentiation. *EMBO Journal* 22, 372–384.
- Takeda, K., Kamanaka, M., Tanaka, T., Kishimoto, T., Akira, S., 1996a. Impaired IL-13-mediated functions of macrophages in STAT6-deficient mice. *Journal of Immunology* 157, 3220–3222.
- Takeda, K., Noguchi, K., Shi, W., *et al.*, 1997. Targeted disruption of the mouse Stat3 gene leads to early embryonic lethality. *Proceedings of the National Academy of Sciences of the United States of America* 94, 3801–3804.
- Takeda, K., Tanaka, T., Shi, W., *et al.*, 1996b. Essential role of Stat6 in IL-4 signalling. *Nature* 380, 627–630.
- Tanner, J.W., Chen, W., Young, R.L., Longmore, G.D., Shaw, A.S., 1995. The conserved box 1 motif of cytokine receptors is required for association with JAK kinases. *Journal of Biological Chemistry* 270, 6523–6530.
- Thierfelder, W.E., van Deursen, J.M., Yamamoto, K., *et al.*, 1996. Requirement for Stat4 in interleukin-12-mediated responses of natural killer and T cells. *Nature* 382, 171–174.
- Thomis, D.C., Gurniak, C.B., Tivol, E., Sharpe, A.H., Berg, L.J., 1995. Defects in B lymphocyte maturation and T lymphocyte activation in mice lacking Jak3. *Science* 270, 794–797.
- Toms, A.V., Deshpande, A., McNally, R., *et al.*, 2013. Structure of a pseudokinase-domain switch that controls oncogenic activation of Jak kinases. *Nature Structural & Molecular Biology* 20, 1221–1223.
- Trengove, M.C., Ward, A.C., 2013. SOCS proteins in development and disease. *American Journal of Clinical and Experimental Immunology* 2, 1–29.
- Udy, G.B., Towers, R.P., Snell, R.G., *et al.*, 1997. Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression. *Proceedings of the National Academy of Sciences of the United States of America* 94, 7239–7244.
- Ungureanu, D., Wu, J., Pekkala, T., *et al.*, 2011. The pseudokinase domain of JAK2 is a dual-specificity protein kinase that negatively regulates cytokine signaling. *Nature Structural & Molecular Biology* 18 (9), 971–976.
- Urban Jr., J.F., Noben-Trauth, N., Donaldson, D.D., *et al.*, 1998. IL-13, IL-4Ralpha, and Stat6 are required for the expulsion of the gastrointestinal nematode parasite *Nippostrongylus brasiliensis*. *Immunity* 8, 255–264.
- Usacheva, A., Sandoval, R., Domanski, P., *et al.*, 2002. Contribution of the Box 1 and Box 2 motifs of cytokine receptors to Jak1 association and activation. *Journal of Biological Chemistry* 277, 48220–48226.
- Vainchenker, W., Constantinescu, S.N., 2013. JAK/STAT signaling in hematological malignancies. *Oncogene* 32, 2601–2613.
- Varghese, L.N., Ungureanu, D., Liao, N.P., *et al.*, 2014. Mechanistic insights into activation and SOCS3-mediated inhibition of myeloproliferative neoplasm-associated JAK2 mutants from biochemical and structural analyses. *Biochemical Journal* 458, 395–405.
- van de Veerdonk, F.L., Plantinga, T.S., Hoischen, A., *et al.*, 2011. STAT1 mutations in autosomal dominant chronic mucocutaneous candidiasis. *New England Journal of Medicine* 365, 54–61.
- Velazquez, L., Fellous, M., Stark, G.R., Pellegrini, S., 1992. A protein tyrosine kinase in the interferon alpha/beta signaling pathway. *Cell* 70, 313–322.
- Vielnasscher, R.M., Hainzl, E., Leitner, N.R., *et al.*, 2014. Conditional ablation of TYK2 in immunity to viral infection and tumor surveillance. *Transgenic Research* 23, 519–529.
- Vinkemeier, U., Cohen, S.L., Moarefi, I., *et al.*, 1996. DNA binding of *in vitro* activated Stat1 alpha, Stat1 beta and truncated Stat1: Interaction between NH2-terminal domains stabilizes binding of two dimers to tandem DNA sites. *EMBO Journal* 15, 5616–5626.
- Visconti, R., Gadina, M., Chiariello, M., *et al.*, 2000. Importance of the MKK6/p38 pathway for interleukin-12-induced STAT4 serine phosphorylation and transcriptional activity. *Blood* 96, 1844–1852.
- Wagner, K.U., Krempler, A., Triplett, A.A., *et al.*, 2004. Impaired alveologenesis and maintenance of secretory mammary epithelial cells in Jak2 conditional knockout mice. *Molecular and Cellular Biology* 24, 5510–5520.
- Wakao, H., Gouilleux, F., Groner, B., 1994. Mammary gland factor (MGF) is a novel member of the cytokine regulated transcription factor gene family and confers the prolactin response. *EMBO Journal* 13, 2182–2191.
- Wallweber, H.J., Tam, C., Franke, Y., Starovasnik, M.A., Lupardus, P.J., 2014. Structural basis of recognition of interferon-alpha receptor by tyrosine kinase 2. *Nature Structural & Molecular Biology* 21, 443–448.
- Walters, D.K., Mercher, T., Gu, T.L., *et al.*, 2006. Activating alleles of JAK3 in acute megakaryoblastic leukemia. *Cancer Cell* 10, 65–75.
- Wang, D., Stravopodis, D., Teglund, S., Kitazawa, J., Ihle, J.N., 1996. Naturally occurring dominant negative variants of Stat5. *Molecular and Cellular Biology* 16, 6141–6148.
- Watling, D., Guschin, D., Muller, M., *et al.*, 1993. Complementation by the protein tyrosine kinase JAK2 of a mutant cell line defective in the interferon-gamma signal transduction pathway. *Nature* 366, 166–170.
- Weber, A., Hengge, U.R., Bardenheuer, W., *et al.*, 2005. SOCS-3 is frequently methylated in head and neck squamous cell carcinoma and its precursor lesions and causes growth inhibition. *Oncogene* 24, 6699–6708.
- Wegrzyn, J., Potla, R., Chwae, Y.J., *et al.*, 2009. Function of mitochondrial Stat3 in cellular respiration. *Science* 323, 793–797.
- Welte, T., Zhang, S.S., Wang, T., *et al.*, 2003. STAT3 deletion during hematopoiesis causes Crohn's disease-like pathogenesis and lethality: A critical role of STAT3 in innate immunity. *Proceedings of the National Academy of Sciences of the United States of America* 100, 1879–1884.
- Wen, Z., Zhong, Z., Darnell Jr., J.E., 1995. Maximal activation of transcription by Stat1 and Stat3 requires both tyrosine and serine phosphorylation. *Cell* 82, 241–250.
- Weniger, M.A., Melzner, I., Menz, C.K., *et al.*, 2006. Mutations of the tumor suppressor gene SOCS-1 in classical Hodgkin lymphoma are frequent and associated with nuclear phospho-STAT5 accumulation. *Verhandlungen der Deutschen Gesellschaft für Pathologie* 90, 210–215.
- Wilks, A.F., 1991. Cloning members of protein-tyrosine kinase family using polymerase chain reaction. *Methods in Enzymology* 200, 533–546.
- Wilks, A.F., Harpur, A.G., Kurban, R.R., *et al.*, 1991. Two novel protein-tyrosine kinases, each with a second phosphotransferase-related catalytic domain, define a new class of protein kinase. *Molecular and Cellular Biology* 11, 2057–2065.
- Wilks, A.F., Kurban, R.R., Hovens, C.M., Ralph, S.J., 1989. The application of the polymerase chain reaction to cloning members of the protein tyrosine kinase family. *Gene* 85, 67–74.
- Witthuhn, B.A., Quelle, F.W., Silvennoinen, O., *et al.*, 1993. JAK2 associates with the erythropoietin receptor and is tyrosine phosphorylated and activated following stimulation with erythropoietin. *Cell* 74, 227–236.

- Witthuhn, B.A., Silvennoinen, O., Miura, O., *et al.*, 1994. Involvement of the Jak-3 Janus kinase in signalling by interleukins 2 and 4 in lymphoid and myeloid cells. *Nature* 370, 153–157.
- Wormald, S., Zhang, J.G., Krebs, D.L., *et al.*, 2006. The comparative roles of suppressor of cytokine signaling-1 and -3 in the inhibition and desensitization of cytokine signaling. *Journal of Biological Chemistry* 281, 11135–11143.
- Xu, X., Sun, Y.L., Hoey, T., 1996. Cooperative DNA binding and sequence-selective recognition conferred by the STAT amino-terminal domain. *Science* 273, 794–797.
- Yamamoto, K., Quelle, F.W., Thierfelder, W.E., *et al.*, 1994. Stat4, a novel gamma interferon activation site-binding protein expressed in early myeloid differentiation. *Molecular and Cellular Biology* 14, 4342–4349.
- Yang, B., Guo, M., Herman, J.G., Clark, D.P., 2003. Aberrant promoter methylation profiles of tumor suppressor genes in hepatocellular carcinoma. *American Journal of Pathology* 163, 1101–1107.
- Yang, J., Chatterjee-Kishore, M., Staugaitis, S.M., *et al.*, 2005. Novel roles of unphosphorylated STAT3 in oncogenesis and transcriptional regulation. *Cancer Research* 65, 939–947.
- Yang, J., Liao, X., Agarwal, M.K., *et al.*, 2007. Unphosphorylated STAT3 accumulates in response to IL-6 and activates transcription by binding to NFkappaB. *Genes & Development* 21, 1396–1408.
- Yang, X.O., Zhang, H., Kim, B.S., *et al.*, 2013. The signaling suppressor CIS controls proallergic T cell development and allergic airway inflammation. *Nature Immunology* 14, 732–740.
- Yao, Z., Cui, Y., Watford, W.T., *et al.*, 2006. Stat5a/b are essential for normal lymphoid development and differentiation. *Proceedings of the National Academy of Sciences of the United States of America* 103, 1000–1005.
- Yasukawa, H., Misawa, H., Sakamoto, H., *et al.*, 1999. The JAK-binding protein JAB inhibits Janus tyrosine kinase activity through binding in the activation loop. *EMBO Journal* 18, 1309–1320.
- Yasukawa, H., Ohishi, M., Mori, H., *et al.*, 2003a. IL-6 induces an anti-inflammatory response in the absence of SOCS3 in macrophages. *Nature Immunology* 4, 551–556.
- Yasukawa, H., Ohishi, M., Mori, H., *et al.*, 2003b. IL-6 induces an anti-inflammatory response in the absence of SOCS3 in macrophages. *Nature Immunology* 4, 551–556.
- Yoshikawa, H., Matsubara, K., Qian, G.S., *et al.*, 2001. SOCS-1, a negative regulator of the JAK/STAT pathway, is silenced by methylation in human hepatocellular carcinoma and shows growth-suppression activity. *Nature Genetics* 28, 29–35.
- Yoshimura, A., Ohkubo, T., Kiguchi, T., *et al.*, 1995. A novel cytokine-inducible gene CIS encodes an SH2-containing protein that binds to tyrosine-phosphorylated interleukin 3 and erythropoietin receptors. *EMBO Journal* 14, 2816–2826.
- Yu, H., Pardoll, D., Jove, R., 2009. STATs in cancer inflammation and immunity: A leading role for STAT3. *Nature Reviews Cancer* 9, 798–809.
- Zadjali, F., Pike, A.C., Vesterlund, M., *et al.*, 2011. Structural basis for c-KIT inhibition by the suppressor of cytokine signaling 6 (SOCS6) ubiquitin ligase. *Journal of Biological Chemistry* 286, 480–490.
- Zhang, J.G., Farley, A., Nicholson, S.E., *et al.*, 1999. The conserved SOCS box motif in suppressors of cytokine signaling binds to elongins B and C and may couple bound proteins to proteasomal degradation. *Proceedings of the National Academy of Sciences of the United States of America* 96, 2071–2076.
- Zhang, J.G., Metcalf, D., Rakar, S., *et al.*, 2001. The SOCS box of suppressor of cytokine signaling-1 is important for inhibition of cytokine action *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America* 98, 13261–13265.
- Zhang, Z., Welte, T., Troiano, N., *et al.*, 2005. Osteoporosis with increased osteoclastogenesis in hematopoietic cell-specific STAT3-deficient mice. *Biochemical and Biophysical Research Communications* 328, 800–807.
- Zhao, R., Xing, S., Li, Z., *et al.*, 2005. Identification of an acquired JAK2 mutation in polycythemia vera. *Journal of Biological Chemistry* 280, 22788–22792.
- Zhao, W., Lee, C., Piganis, R., *et al.*, 2008. A conserved IFN-alpha receptor tyrosine motif directs the biological response to type I IFNs. *Journal of Immunology* 180, 5483–5489.
- Zhao, Y., Zeng, C., Tarasova, N.I., *et al.*, 2013. A new role for STAT3 as a regulator of chromatin topology. *Transcription* 4, 227–231.
- Zhong, Z., Wen, Z., Darnell Jr., J.E., 1994. Stat3: A STAT family member activated by tyrosine phosphorylation in response to epidermal growth factor and interleukin-6. *Science* 264, 95–98.

Further Reading

- Reich, N.C., 2013. STATs get their move on. *JAK-STAT* 2 (4), e27080.
- Stark, G.R., Darnell, J.E., 2012. The JAK-STAT pathway at twenty. *Immunity* 36 (4), 503–514.

The PLC Pathway

KL Everett, F-star Biotechnology Ltd., Cambridge, UK
M Katan, University College London, London, UK

© 2016 Elsevier Inc. All rights reserved.

Families of Phosphoinoside-Specific Phospholipase C Enzymes

Phosphoinoside-specific phospholipase C (PLC) enzymes catalyze the hydrolysis of inositol lipids, mainly phosphatidylinositol 4,5-bisphosphate [PI(4,5)P₂] to inositol 1,4,5-trisphosphate (IP₃) and diacylglycerol (DAG) (Suh *et al.*, 2008; Gresset *et al.*, 2012; Kadamur and Ross, 2013; Figure 1(a)). The released IP₃ binds to IP₃ receptors on the endoplasmic reticulum resulting in Ca²⁺ release into the cytoplasm. DAG remains in the membrane and activates some isoforms of protein kinase C (PKC), other kinases, and GTPase regulating proteins (Brose and Rosenmund, 2002). Other functions of PI(4,5)P₂ are related to its interactions with a number of proteins, particularly ion channels and proteins involved in the actin cytoskeleton (Suh and Hille, 2008; Lemmon, 2008). PLC enzymes could impact on these interactions by changing local concentrations of PI(4,5)P₂.

There are six families of PLC enzymes (PLC β , δ , γ , ϵ , η , and ζ) and 13 human isoforms in total (β 1–4, δ 1, 3 and 4, γ 1 and 2, ϵ 1, η 1 and 2 and ζ 1); some isoforms also have splice variants (Suh *et al.*, 2008). With respect to regulatory mechanisms, the most extensive insights have been obtained for activation of PLC β family by G-protein coupled receptors (GPCRs) and PLC γ enzymes activated via growth factor receptors and immune cell receptors (Suh *et al.*, 2008; Gresset *et al.*, 2012; Kadamur and Ross, 2013). More recently, the role of small GTPases from Ras and Rho families in regulation of PLC β , PLC γ , and PLC ϵ enzymes has been also documented, extending the complexity of their regulatory networks (Harden and Sondek, 2006; Bunney and Katan, 2006).

Domain Organization and Structure of PLCs

As shown in Figure 1(b), PLC enzymes contain a pleckstrin homology (PH) domain (except for PLC ζ 1), four EF hands, a triose isomerase (TIM) barrel catalytic domain, and a C2 domain. The TIM barrel contains two regions of high sequence similarity, originally described as X and Y, corresponding to two halves of the TIM barrel structure. Between these two regions is a linker, the X–Y linker, which varies between PLC families. Each PLC family, except PLC δ and PLC ζ , has additional regulatory domains. PLC β and PLC η isoforms have C-terminal tails and PLC ϵ contains a Cdc25 domain and two Ras association (RA) domains. PLC γ isoforms contain a split PH domain (spPH), two Src homology 2 domains (nSH2 and cSH2), and a Src homology 3 (SH3) domain. The region containing these domains is known as the γ -specific array (γ SA) and is inserted between the two halves of the TIM barrel sequence, representing the most complex X–Y linker.

The crystal structure of PLC δ 1 (minus the PH domain) was the first structure of the catalytic region of a mammalian PLC enzyme to be solved (Essen *et al.*, 1996); this provided

structural insights into the PLC common core including inter-domain interactions and the mechanism of PIP₂ hydrolysis (Essen *et al.*, 1997). More recently, crystal structures of PLC β enzymes confirmed that inter-domain interactions between the TIM barrel and the EF hands and the TIM barrel and the C2 domain are conserved (Lyon *et al.*, 2011; Jezyk *et al.*, 2006). The only major difference is expected to be in the positioning of the PH domain. The PH domain of PLC δ 1 binds strongly to PI(4,5)P₂ and is attached to the rest of the enzyme by a flexible linker (Essen *et al.*, 1996) whereas the PH domain of some PLC β enzymes binds to small GTPases and is more integral to the enzyme structure (Jezyk *et al.*, 2006).

The progress has also been made in understanding the structural basis for activation of PLC enzymes, in particular for PLC β and PLC γ isoforms; the information covers complexes of PLC β with G α q and small GTPase Rac and of PLC γ with FGFR and Rac (Lyon *et al.*, 2011; Jezyk *et al.*, 2006; Waldo *et al.*, 2010; Bunney *et al.*, 2009; Bae *et al.*, 2009). Importantly, structure-based functional analysis revealed a general, auto-inhibitory mechanism that involves elements from the X–Y linker that in most PLC families contributes to occlude the active site in the TIM barrel; different regulatory interactions act to release this intra molecular inhibitory constraint as a step in a more complex activation process (Bunney and Katan, 2011).

Signaling Connectivity, Function, and Dysfunction

A number of different experimental approaches provided insights into physiologically important signaling connectivity for each PLC family. In addition, experimental evidence related to their dysfunction and links to disease highlighted their key regulatory roles. This information is summarized in Table 1.

PLC δ family

PLC δ isoforms 1, 3, and 4 comprise the PLC δ family. There is some evidence for signaling networks that involve PLC δ but they are not well defined. PLC δ 1 is the most studied of the PLC δ enzymes. Its activity is more sensitive to Ca²⁺ levels than most other PLCs and its PLC activity can be stimulated by Ca²⁺ levels within the physiological range (Allen *et al.*, 1997). It has been suggested, therefore, that PLC δ 1 could be stimulated downstream of the Ca²⁺ mobilized by activation of other PLCs or in the vicinity of calcium channels (Allen *et al.*, 1997; Thompson and Shuttleworth, 2011). The binding of the PLC δ 1 PH domain to PI(4,5)P₂ is also important for PLC δ 1 activity and can be competed out by IP₃ (Yagisawa *et al.*, 1998). Thus, changes in Ca²⁺, IP₃ or PI(4,5)P₂ could influence activity of PLC δ 1. Protein–protein interactions may also regulate PLC δ 1 activity. PLC δ 1 activation has been linked to two GTP binding proteins, G_h (also known as transglutaminase II), a high molecular weight GTP binding protein, and Ral, a member of the Ras family of small GTPases

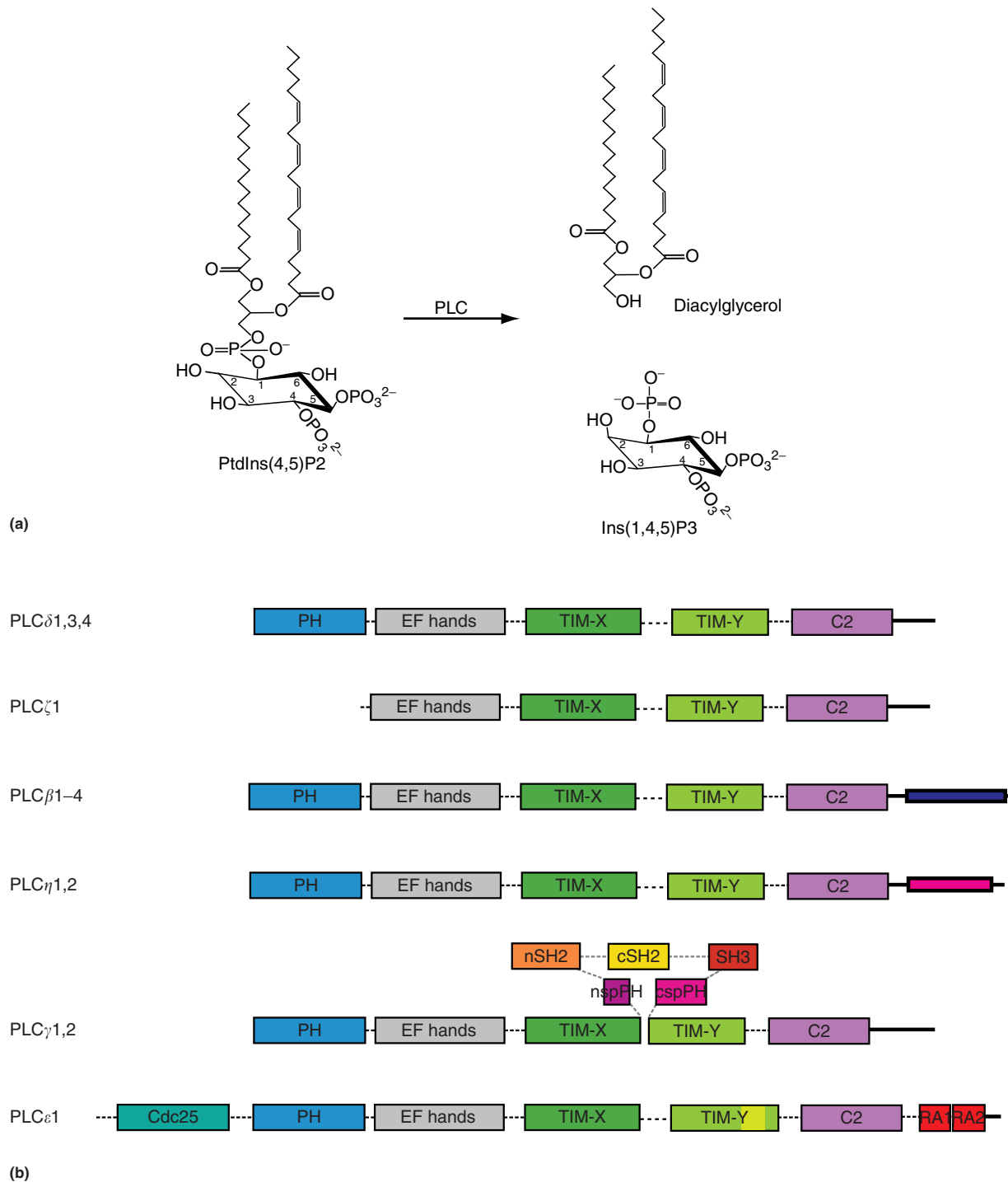


Figure 1 Domain organization of the PLC enzymes. (a) PLC enzymes catalyzed hydrolysis $PI(4,5)P_2$ to IP_3 and DAG. (b) The basic structure of PLC enzymes consists of a PH domain (missing in $PLC\zeta 1$), four EF hands, a split TIM barrel catalytic domain and a C2 domain. The TIM barrel is formed from the X region and the Y region and the linker between the two is referred to as the X-Y linker. The PLC enzymes, except $PLC\delta$ and $PLC\zeta 1$, have additional regulatory regions. $PLC\beta$ and $PLC\eta$ have extended C-terminal regions. $PLC\gamma$ contains a spPH domain, two SH2 domains, and an SH3 domain within its X-Y linker. $PLC\epsilon 1$ has an N-terminal Cdc25 domain and two C-terminal RA domains. $PLC\epsilon 1$ also contains an insert in the Y region of the TIM barrel that is otherwise well conserved between PLC enzymes.

(Feng *et al.*, 1996; Sidhu *et al.*, 2005). However, none of these proposed interactions have been evaluated in biological contexts.

Some physiological roles of $PLC\delta$ enzymes can be suggested based on the phenotypes of knockout mice (Fukami *et al.*, 2008). For example, *Plcd1* knockout suggests that $PLC\delta 1$

Table 1 Expression, function and dysfunction of PLC enzymes

Human gene/protein	Expression pattern ^a	Phenotype of knockout mice	Dysfunction/disease
PLCD1/PLC δ 1	Most tissues	Hair loss (Nakamura <i>et al.</i> , 2005) Inflammation in the skin (Ichinohe <i>et al.</i> , 2007)	Possible tumor suppressor, down regulated in several cancers (Fu <i>et al.</i> , 2007; Hu <i>et al.</i> , 2009; Danielsen <i>et al.</i> , 2011)
PLCD3/PLC δ 3	Most tissues	No abnormality (Nakamura <i>et al.</i> , 2005)	Down regulated in some cancers (Danielsen <i>et al.</i> , 2011)
PLCD4/PLC δ 4	Higher in neuronal tissue	Male infertility (Fukami <i>et al.</i> , 2001)	
PLCB1/PLC β 1	Higher in neuronal tissue	Retarded growth, low viability, and epilepsy (Kim <i>et al.</i> , 1997)	
PLCB2/PLC β 2	High in hematopoietic cells	Defective leukocyte response to chemoattractants (Jiang <i>et al.</i> , 1997)	
PLCB3/PLC β 3	Most tissues	Macrophage hypersensitivity and reduced atherosclerosis (Wang <i>et al.</i> , 2008) Premature death, myeloproliferative disease, and lymphoma (Xiao <i>et al.</i> , 2009)	Possible tumor suppressor, reduced expression in some cancers, no identified mutations (Xiao <i>et al.</i> , 2009)
PLCB4/PLC β 4	Higher in neuronal tissue	Ataxia (Kim <i>et al.</i> , 1997)	
PLCG1/PLC γ 1	Most tissues High in T cells	Embryonic lethal, day E9 (Ji <i>et al.</i> , 1997)	Cancer promoting Increased expression in several cancers (Arteaga <i>et al.</i> , 1991; Noh <i>et al.</i> , 1994; Thomas <i>et al.</i> , 2003; Sala <i>et al.</i> , 2008) Gain-of-function mutations in some cancers (Behjati <i>et al.</i> , 2014; Vaque <i>et al.</i> , 2014) Gain-of-function function/hypomorphic mutations in immune disorders (Ombrello <i>et al.</i> , 2012; Zhou <i>et al.</i> , 2012)
PLCG2/PLC γ 2	High in hematopoietic cells	Decreased mature B cells, loss of function in other hematopoietic cells (Wang <i>et al.</i> , 2000a,b)	Mutations in nephrotic syndrome (Hinkes <i>et al.</i> , 2006)
PLCE1/PLC ϵ 1	Generally low expression	Heart defects (Tadano <i>et al.</i> , 2005; Wang <i>et al.</i> , 2005; Zhang <i>et al.</i> , 2013) Decreased tumor development due to decreased inflammation (Ikuta <i>et al.</i> , 2008; Li <i>et al.</i> , 2009) Increased tumor development due to changes in cancer cells (Martins <i>et al.</i> , 2014)	Possible tumor suppressor. Reduced expression in some cancers (Martins <i>et al.</i> , 2014; Sorli <i>et al.</i> , 2005; Danielsen <i>et al.</i> , 2011; Wang <i>et al.</i> , 2012). Single nucleotide polymorphisms associated with some cancers (Abnet <i>et al.</i> , 2010; Wang <i>et al.</i> , 2010)
PLCH1/PLC η 1	Neuronal		
PLCH2/PLC η 2	Neuronal	No global effects. Impact on development and maturation of retina (Kanemaru <i>et al.</i> , 2010)	
PLCZ1/PLC ζ 1	Testis/sperm		Reduced expression associated with infertility (Yoon <i>et al.</i> , 2008; Heytens <i>et al.</i> , 2009).

^aExpression patterns were determined from USCS Genome Bioinformatics (<http://genome.ucsc.edu>), Human Protein Atlas (<http://www.proteinatlas.org>), and Suh *et al.* (2008).

may influence inflammation in the skin (Ichinohe *et al.*, 2007) while PLC δ 4 may play a role in male fertility (Fukami *et al.*, 2001). *Plcd3* knockout mice did not have an abnormal phenotype but combined *Plcd1/Plcd3* knockouts die between days E11.5 and E13.5 due to defects in the placenta (Nakamura *et al.*, 2005).

PLC δ enzymes may play a role in cancer. It has been suggested that PLC δ 1 acts as a tumor suppressor in esophageal squamous cell carcinoma, gastric cancer, and colorectal cancer (Fu *et al.*, 2007; Hu *et al.*, 2009; Danielsen *et al.*, 2011). The chromosomal region that contains the *PLCD1* gene and several other candidate tumor suppressor genes is deleted in many cancers (Fu *et al.*, 2007; Hu *et al.*, 2009). PLC δ 1 was also found to be down regulated and its promoter methylated in a number of gastric cancer cell lines. Sixty-two percent of gastric cancer tumors showed *PLCD1* methylation, whereas no methylation was seen in any of the normal tissue controls (Hu *et al.*, 2009). PLC δ 1 protein expression was also decreased in

tumors. Expression of PLC δ 1 in gastric cancer cell lines decreased growth and migration and altered cell morphology. This effect is proposed to be mediated via PLC δ 1's effect on proteins that rearrange the cytoskeleton. In esophageal squamous cell carcinoma the tumor suppressive effects of PLC δ 1 were attributed to an effect on cell cycle progression through the G1-S checkpoint (Fu *et al.*, 2007). In colorectal cancer, reduced expression of PLC δ 1 was validated in biopsies and cell lines by quantitative analyses and found associated with *KRAS* mutations (Danielsen *et al.*, 2011). Promoter methylation analyses revealed that also in this cancer type, methylation of *PLCD1* can contribute to reduced expression.

PLC β family

PLC β isozymes (β 1- β 4) are well-documented targets downstream of GPCRs, activated by $G\alpha$ and $G\beta\gamma$ subunits of heterotrimeric G-proteins. Members of the Rho family GTPases, such as the Rac isoforms, also directly bind and activate PLC β ,

linking PLC β activity to GPCR-independent signaling cascades (Suh *et al.*, 2008; Gresset *et al.*, 2012; Kadamur and Ross, 2013).

Activation of PLC β by members of the Gq subfamily have prominent roles in cardiovascular function, chemotaxis, and neuronal signaling. Recent studies provided structural and mechanistic insights into PLC β activation by G α_q (Harden *et al.*, 2011; Lyon *et al.*, 2014). The main binding site of G α_q is the proximal site (near the C2 domain) of the C-terminal tail of PLC β (designated as proximal-CTD). A recently suggested mechanistic model implies that membrane recruitment on its own is not sufficient for the activation (Lyon *et al.*, 2014). Instead, it has been proposed that the binding of G α_q displaces the auto-inhibitory helix from the catalytic core, allosterically activating PLC β . Simultaneously, the distal CTD interacts with the membrane and with G α_q in a manner that helps the G α_q -catalytic core complex adopt a specific orientation with respect to the membrane surface. The resulting electrostatic repulsion between the membrane and acidic regions in the X-Y linker helps to open the active site, allowing binding of the PI(4,5)P₂ substrate. Thus, the distal CTD and the membrane could be acting as scaffolds that help to not only optimize the orientation of the G α_q -catalytic core complex, but also allosterically activate it. Interestingly, another region within PLC β subsequently accelerates GTP hydrolysis by of G α_q , leading to its inactivation and signal termination (Harden *et al.*, 2011).

As originally observed for PLC β 2, PLC β 1-3 are also activated by G $\beta\gamma$ dimers which is thought to be mediated by activation of Gi-coupled GPCR (Camps *et al.*, 1992). G $\beta\gamma$ dimers interact with PLC β 2 via its PH domain and in a chimeric protein the PLC β 2 PH domain is sufficient to make PLC δ 1 sensitive to stimulation by G $\beta\gamma$ (Wang *et al.*, 2000b).

Experiments *in vitro* and in COS-7 cells have shown that PLC β 2 is activated by the Rac and Cdc42 members of the Rho family of GTPases in a GTP dependent manner. PLC β 3 is less sensitive to activation in this way and PLC β 1 is Rac insensitive. Experiments using chimeras of PLC β 1 and PLC β 2 identified the N-terminal PH domain of PLC β 2 as a potential binding site for Rac (Illenberger *et al.*, 2003). The crystal structure of PLC β 2 and Rac1 further demonstrated that the PH domain is the only part of PLC β 2 that binds to the GTPase (Jezzyk *et al.*, 2006). It has been suggested that a major role of Rac binding could be translocation of PLC β 2 to the membrane.

Knockout mice have been generated for each of the four PLC β isozymes and suggest diverse physiological roles. *Plcb1*^{-/-} mice display retarded growth and low viability after birth (Kim *et al.*, 1997). The mice have recurrent epileptic seizures indicative of a role for PLC β 1 in the normal functioning of inhibitory neuronal circuitry. *Plcb4*^{-/-} mice also have a neuronal defect. They suffer from ataxia caused by retarded development of the cerebellum (Kim *et al.*, 1997). Further pharmacological experiments using the mice suggested that PLC β 1 is active downstream of muscarinic acetylcholine receptors in the cerebral cortex, hippocampus and cerebellum and PLC β 4 is downstream of both muscarinic acetylcholine receptors and metabotropic glutamate receptors in the cerebellum (Kim *et al.*, 1997). *Plcb2*^{-/-} mice show no effects at the systemic or cellular level. Neutrophils from the mice show decreased production of inositol phosphates in response to chemo-attractants (Jiang *et al.*, 1997). However, chemotaxis toward

these signals is increased. This suggests that chemotaxis is regulated by inputs from a number of pathways. *Plcb2*^{-/-} mice also show an enhanced response to bacterial or viral infection. *Plcb3* knockout results in mice with macrophages that are more sensitive to apoptosis (Wang *et al.*, 2008). When *Plcb3*^{-/-} mice are crossed with a model of atherosclerosis the mice lacking PLC β 3 have less atherosclerotic plaques.

More recently *Plcb3*^{-/-} mice have been reported to die prematurely, display symptoms of myeloproliferative disease and have various tumors (Xiao *et al.*, 2009). Based on this phenotype, PLC β 3 has been identified as a tumor suppressor (Xiao *et al.*, 2009). The phenotype is mediated by bone marrow cells and *Plcb3*^{-/-} cell populations enhanced for hematopoietic stem cells exhibit increased proliferation and decreased apoptosis. Interestingly, PLC β 3's tumor suppressive effect is mediated by its C-terminal tail and is independent of lipase activity. As yet no PLC β 3 point mutations have been identified in cancers but a small number of acute lymphocytic leukemia patients have a deletion of the chromosomal region that includes *PLCB3* and 11% of chronic lymphocytic leukemia samples studied had low PLC β 3 expression (Xiao *et al.*, 2009).

Recent studies of uveal melanomas suggest that PLC β could have an important role in this type of cancer without itself harboring mutations; the predominance of activating, somatic mutations in G α_q or G α_{11} in uveal melanomas results in constitutive activation of their downstream signaling responses (Van Raamsdonk *et al.*, 2009). In this context, PLC β could promote cancer formation rather than act as a tumor suppressor.

PLC γ family

PLC γ isoforms (PLC γ 1 and PLC γ 2) are activated downstream of receptor and non-receptor tyrosine kinases (Suh *et al.*, 2008; Gresset *et al.*, 2012; Kadamur and Ross, 2013). Action of non-receptor tyrosine kinases (from Src, Syk, and Tec families) has been best defined in the context of activation of immune cell (T cell, B cell, and Fc) receptors associated with multi-protein complexes. Ubiquitously expressed PLC γ 1 is activated downstream of growth factor stimulation, including stimulation by epidermal growth factor, platelet derived growth factor receptor, vascular endothelial growth factor and fibroblast growth factor (FGF). These growth factors regulate a number of cellular processes including proliferation, motility and cell survival; although PLC γ 1 could be involved in regulation of several of these processes, the importance of this PLC in cell motility has been best documented (Bunney and Katan, 2010). PLC γ 2, with highest expression in hematopoietic cells, is activated by immune cell receptors such as B cell and Fc receptors, however, T-cell receptor activation is linked to PLC γ 1. More recently it has been discovered that PLC γ 2 is activated by the small GTPase Rac, but as yet no activation of PLC γ 1 by GTPases has been identified (Piechulek *et al.*, 2005).

Phosphorylation of PLC γ by tyrosine kinases is important for its activation by these kinases. In both PLC γ 1 and PLC γ 2, the critical tyrosine residues (such as PLC γ 1 Y783 and PLC γ 2 Y759) are within a linker between cSH2 and SH3 domains. Recent structural and mechanistic studies focused on PLC γ 1 activation by tyrosine kinase receptors (Bae *et al.*, 2009; Bunney *et al.*, 2012). In a basal state, PLC γ is auto-inhibited by intramolecular interactions between the cSH2 domain and the

catalytic core. This auto-inhibitory interaction is out competed following phosphorylation within the linker region and by the subsequent binding of the phosphotyrosine to the cSH2 domain. In contrast to the cSH2 domain that appears to be mainly involved in intramolecular interactions, the nSH2 domain is the main determinant of recognition of specific binding sites on at least some growth factor receptors; this is best documented for FGF receptor (Bae *et al.*, 2009; Bunney *et al.*, 2012). Interestingly, activation of PLC γ 2 by Rac, that involves binding to the spPH domain within the γ SA, is independent of tyrosine phosphorylation and the mechanism of activation in this case remains unclear (Piechulek *et al.*, 2005; Bunney *et al.*, 2009).

Plcg1 knockout is embryonic lethal in mice at around embryonic day 9 (Ji *et al.*, 1997). At embryonic day 8.5–9.5 the *Plcg1*^{-/-} embryos are smaller than wild-type but no gross abnormalities are visible. Further analysis demonstrated that both erythropoiesis and vasculogenesis were significantly disrupted in these mice (Liao *et al.*, 2002). They lack expression of an adhesion molecule marker for endothelial cells; however the lack of blood vessels could be due to lack of angiogenesis or a defect in differentiation. In contrast, *Plcg2* knockout mice are viable but have significant defects in the immune system and platelets (Wang *et al.*, 2000a). FACS analysis shows that they have a shift toward immature B cell populations and a decrease in the mature B cell population. B cells from *Plcg2*^{-/-} mice do not mobilize Ca²⁺ in response to stimulation with anti-IgM but do respond normally to phorbol esters and ionomycin. Mast cells and natural killer cells also show decreased function following stimulation via the Fc ϵ R and Fc γ RII/II respectively. *Plcg2*^{-/-} mice experience increased hemorrhaging and their platelets do not aggregate or release TXA₂ or ATP in response to collagen. T cell function and signaling is normal in these mice. This shows that PLC γ 2 is nonredundant and mediates the effects of several immunoglobulin superfamily receptors (Wang *et al.*, 2000a).

The first evidence that PLC γ enzymes could be linked to complex disorders by incorporating gain-of-function or hypermorphic mutations came from a large-scale *N*-ethyl-*N*-nitrosourea mutagenesis screen. Among mouse strains with abnormal limbs due to severe inflammation two strains, abnormal limb 5 and 14 (Ali5 and Ali14), contained single point mutations in PLC γ 2 that resulted in gain-of-function (Yu *et al.*, 2005; Everett *et al.*, 2009; Abe *et al.*, 2011). Subsequently, genetic analyses of various dominantly inherited immune disorders in humans have revealed that mutations in the *PLCG2* gene are indeed causatively linked to two clinical syndromes designated as PLAID and APLAID syndrome (Ombrello *et al.*, 2012; Zhou *et al.*, 2012). PLAID patients have familial cold-induced urticaria that is caused by mast cell degranulation. They also have variable B cell immunodeficiency and a predisposition to developing skin granulomas and autoimmunity. Through linkage analysis and candidate gene sequencing of PLAID patients, it was found that they have large (4–8 kb) deletions in *PLCG2* that resulted in a cold sensitive activation of PLC γ 2 enzyme (Ombrello *et al.*, 2012). Another study identified a family with an autosomal dominant autoinflammatory disease of the eyes, skin and lungs linked to a point mutation in *PLCG2* (Zhou *et al.*, 2012). The disease that was observed in this family was designated

autoinflammatory PLAID or APLAID. The B cells of APLAID patients differed from those of patients with PLAID in that surface ligand-mediated cellular signaling was actually enhanced in these cells and there was no cold sensitivity.

Because PLC γ 1 is a common component downstream of many growth factor receptors mutated in cancer, it was assumed that it could contribute to tumor generation and progression as the key link to deregulation of processes such as cell motility. Initial studies of tumor samples also indicated that PLC γ 1 is overexpressed in human cancers including breast, colon and head and neck squamous cell carcinomas (Arteaga *et al.*, 1991; Noh *et al.*, 1994; Thomas *et al.*, 2003). An increase in PLC γ 1 phosphorylation in tumor samples was also identified (Thomas *et al.*, 2003). Furthermore, increased levels of PLC γ 1 expression have been detected in lymph node metastases compared to primary tumors from breast cancer patients (Sala *et al.*, 2008). However, it was only recently that gain-of-function mutations in PLC γ 1 have been identified in tumor samples, providing evidence for its role as an oncogene. Two independent studies identified recurrent mutations in *PLCG1* in about 10% of angiosarcomas, characterized by aberrant angiogenesis (Behjati *et al.*, 2014), and in about 20% of cutaneous T-cell lymphomas (Vaque *et al.*, 2014); in both types of these relatively rare cancers, aggressive forms lack effective treatment.

It is also possible to interpret various gain-of-function mutations in PLC γ enzymes within the structural framework for PLC γ activation and auto-inhibition (Bunney *et al.*, 2012). It appears that deletions and point mutations, found in immune disorders and cancer, affect regions within the auto-inhibitory surface either within the γ SA (in particular cSH2 domain) or in the vicinity of the active site, resulting in the constitutive or enhanced activation.

PLC ϵ 1

PLC ϵ 1, the largest PLC enzyme, is downstream of both GPCRs and receptor tyrosine kinases by virtue of its ability to be regulated by small GTPases from Ras and Rho families (Bunney and Katan, 2006; Harden and Sodek, 2006; Smrcka *et al.*, 2012). However, PLC ϵ 1 appears to mediate more-sustained phosphoinositide hydrolysis compared to the more receptor-proximal PLC β and PLC γ enzymes; PLC ϵ 1 does so through its regulation by a distinct set of GPCRs and at least some receptor tyrosine kinases (Smrcka *et al.*, 2012).

PLC ϵ 1 activation by classical members of the Ras family (H-Ras, K-Ras, and N-Ras) is through binding to the RA2 domain (Bunney *et al.*, 2006). The affinity of Ras for the PLC ϵ 1 RA2 domain is similar to its affinity for PI3-kinases and RalGDS, however it is lower than its affinity for Raf-1. One of the roles of Ras binding appears to be translocation of PLC ϵ 1 to the membrane. Adding a membrane-targeting tag to PLC ϵ 1 removes the requirement for the RA2 domain in activation. However, Ras binding may also contribute to conformational change. PLC ϵ 1 is also activated by Rap1, Rap2, and R-Ras. *In vitro* the binding of these GTPases to the RA2 domain has 10-fold less affinity than the binding of classical Ras proteins but in cells other factors may contribute to activation (Bunney *et al.*, 2006). Activation by Rho GTPases is independent of the RA2 domain and may involve a unique insert in the Y-box.

PLC ϵ 1 also contains a region in its N-terminus that has homology with the Cdc25 domain found in Ras family guanine exchange factors (GEFs) including SOS and Epac. The Cdc25 domain is the catalytic domain of these GEFs and catalyzes the conversion of GTPases from their GDP to GTP bound form. In their GTP bound form they are able to bind to and influence their effectors. The presence of a Cdc25 domain in PLC ϵ 1 suggests that PLC ϵ 1 could regulate the activity of Ras family GTPases as well as being activated by them. One study has suggested that this domain has GEF activity for H-Ras as H-Ras-GTP levels increased when PLC ϵ 1b was expressed and this effect was independent of lipase activity (Lopez *et al.*, 2001). A separate study of the PLC ϵ 1a splice variant looked for GEF activity toward H-Ras, M-Ras, Rap1, Rap2, RalA, Rit, Rin, and Rheb. This study demonstrated that the GEF activity was selective for Rap1 (Jin *et al.*, 2001). The role of the Cdc25 domain is, therefore, still unclear and it remains possible that the domain has no direct GEF activity. If the domain is active, the regulation of its activity may involve the other domains of PLC ϵ 1.

PLC ϵ 1 transgenic mice were first generated by in-frame disruption of the catalytic domain (PLC ϵ 1 Δ x/ Δ x). These mice were found to have heart valve defects, although the mice were viable (Tadano *et al.*, 2005). Another study of PLC ϵ 1 mice disrupted the *Plce1* gene in a way that resulted in a near complete knockout (PLC ϵ 1 $^{-/-}$) (Wang *et al.*, 2005). This study detected different heart defects, pointing to a role for PLC ϵ 1 in GPCR-mediated hypertrophic growth of cardiomyocytes. Cardiac myocytes from these mice showed decreased Ca $^{2+}$ signaling in response to β -adrenergic stimulation. Subsequently, it was shown that the hypertrophy of neonatal rat ventricular myocytes driven by ET-1, norepinephrine, or isoproterenol was inhibited by small interfering RNA-mediated knockdown of PLC ϵ 1 (Zhang *et al.*, 2011). Furthermore, it was shown that PLC ϵ scaffolds to muscle-specific A kinase-anchoring protein at the nuclear envelope, and disruption of this scaffolding interaction prevents development of agonist-induced hypertrophy (Zhang *et al.*, 2011). The important insight into the mechanistic role of PLC ϵ 1 in cardiac function and failure was recently obtained from targeted deletion of *Plce1* specifically in cardiac myocytes in mice after development; these mice are protected from pressure-overload-induced hypertrophy, clearly identifying the *Plce1* gene as critical for the development of heart failure (Zhang *et al.*, 2013).

Findings that PLC ϵ 1 can directly bind and be activated by Ras GTPases that are frequently mutated in several cancer types, prompted search for possible links between the Ras oncogene and PLC ϵ 1 in the context of cancer. Notably, PLC ϵ 1 transgenic mice were tested in different cancer models. Initial studies using PLC ϵ 1 Δ x/ Δ x mice indicated that PLC ϵ 1 may play a role in tumorigenesis as they show decreased skin tumor formation in a two-stage model of skin chemical carcinogenesis (Bai *et al.*, 2004). Later studies suggested that PLC ϵ 1 was not involved in proliferation of keratinocytes driven by the oncogenic Ras but was instead involved in TPA induced inflammation (Ikuta *et al.*, 2008). In cultures of dermal fibroblasts TPA induces PLC ϵ 1 activity leading to the production of interleukin-1 α , a proinflammatory molecule. This suggests that PLC ϵ 1's role in cancer may be in the cells surrounding the tumor rather than the tumor itself. When

PLC ϵ 1 Δ x/ Δ x background was combined with a genetic mouse model for human colorectal tumorigenesis, also dependent on immune responses, the mice had resistance to spontaneous tumorigenesis (Li *et al.*, 2009). In contrast to these observations, PLC ϵ 1 $^{-/-}$ mice exhibited increased susceptibility to tumor formation in the two-stage skin carcinogenesis protocol revealing a tumor suppressor function for this PLC in cancer cells. Although significant differences were not seen in the *LSL-KrasG12D* non small cell lung carcinoma model, down-regulation of PLC ϵ was found in animal tumors and in cellular systems following expression of the oncogenic Ras (Martins *et al.*, 2014). Furthermore, tumor suppressor role of PLC ϵ 1 observed in these mouse models is consistent with observations of marked down-regulation in several types of human tumors, including lung and colorectal cancer, and the inhibitory effects of PLC ϵ 1 expression on cell growth (Martins *et al.*, 2014; Sorli *et al.*, 2005; Danielsen *et al.*, 2011; Wang *et al.*, 2012). Thus, it appears that PLC ϵ 1 could have multiple and context dependent roles in tumor environment and cancer cells themselves.

A link between PLC ϵ 1 and a development of human tumors was also suggested by genome-wide association studies; the studies identified *PLCE1* single nucleotide polymorphisms (SNPs) associated with esophageal squamous cell carcinoma (ESCC) (Abnet *et al.*, 2010; Wang *et al.*, 2010). One of the studies also found that the same *PLCE1* SNPs associated with ESCC were also associated with gastric adenocarcinoma (Abnet *et al.*, 2010). Whether these alterations associated with the *PLCE1* SNPs alter *PLCE1* function or expression is unknown and it remains unclear whether such alterations are causal or merely correlated with the ESCC or gastric adenocarcinoma.

Mutations in PLC ϵ 1 have been identified in patients with nephrotic syndrome, a kidney disease (Hinkes *et al.*, 2006). Seven homozygous mutations have been identified, six truncation mutations and one missense mutation. Patients with PLC ϵ mutations have impaired glomerular development, thus it has been suggested that PLC ϵ signaling may be involved in this process.

PLC η family

PLC η 1 and PLC η 2 were the most recently identified PLC isoforms (Katan, 2005). There are three splice variants of PLC η 1 and five splice variants of PLC η 2. These variants differ in their C-terminal tails (Hwang *et al.*, 2005; Zhou *et al.*, 2005). The PLC η isoforms are expressed in neuron enriched regions and are abundant in the brain (Stewart *et al.*, 2007). In COS-7 cells, co-expression of PLC η 2 with G β $_1\gamma$ $_2$ resulted in an increase in PLC activity (Zhou *et al.*, 2005). No other G $\beta\gamma$ dimer produced activation and PLC η 1 was not activated in this way. The activation did not require the C-terminal tail region of the enzyme. Activation was also seen in reconstitution assays using almost pure protein suggesting that the activation was due to a direct interaction.

PLC η 2 knockout mice are viable and further analysis suggested the role of PLC η 2 in development and maturation of retina where this PLC is highly expressed (Kanemaru *et al.*, 2010).

PLC ζ 1

PLC ζ 1 is the smallest mammalian PLC enzyme and is the only one not to contain a PH domain. It is around 100 times more

sensitive to Ca^{2+} than $\text{PLC}\delta 1$ and is activated at levels of Ca^{2+} found in resting cells (Nomikos *et al.*, 2005). This testis-specific PLC is expressed in sperm and is believed to be the sperm factor required to induce Ca^{2+} oscillations in fertilized eggs (Amdani *et al.*, 2013, Kashir *et al.*, 2014). These oscillations are required for egg activation and completion of meiosis. $\text{PLC}\zeta 1$ may be inactivated in some way in sperm, possibly by its localization, and then activated by resting Ca^{2+} levels in the egg following fertilization. At fertilization, the sperm initiates development by causing a series of oscillations in concentrations of calcium Ca^{2+} within the egg. $\text{PLC}\zeta 1$ mimics the sperm at fertilization, causing the same pattern of Ca^{2+} release as seen at fertilization.

A number of reports from the clinic support the link between defects in human $\text{PLC}\zeta 1$ and the inability of sperm to cause egg activation. Sperm of some infertile men are unable to induce Ca^{2+} oscillations upon microinjection into mouse oocytes and such patients exhibited reduced/absent levels of $\text{PLC}\zeta 1$ (Yoon *et al.*, 2008, Heytens *et al.*, 2009). The failure to activate egg can be in some instances treated by using artificial egg activation methods, such as applying Ca^{2+} ionophores. Indeed, high fertilization rates were achieved with patients whose sperm were deficient in $\text{PLC}\zeta 1$ when a Ca^{2+} ionophore was used to artificially activate eggs (Taylor *et al.*, 2010). It is hoped that the use of purified $\text{PLC}\zeta 1$ can provide a better option for this type of artificial fertilization, bypassing possible side effects of Ca^{2+} ionophores.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: G Protein-Coupled Receptors; Receptor Tyrosine Kinases and Their Ligands. Cell Communication: Intracellular Pathways: Calcium and Calmodulin Signaling; SH3 and SH2: Prototypic Domains to Mediate Regulatory Mechanisms in the Cell. Molecular Principles, Components, Technology, and Concepts: Lipids: Lipid Signaling

References

- Abe, K., Fuchs, H., Boersma, A., *et al.*, 2011. A novel N-ethyl-N-nitrosourea-induced mutation in phospholipase Cgamma2 causes inflammatory arthritis, metabolic defects, and male infertility *in vitro* in a murine model. *Arthritis and Rheumatism* 63, 1301–1311.
- Abnet, C.C., Freedman, N.D., Hu, N., *et al.*, 2010. A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. *Nature Genetics* 42, 764–767.
- Allen, V., Swigart, P., Cheung, R., Cockcroft, S., Katan, M., 1997. Regulation of inositol lipid-specific phospholipase Cdelta by changes in Ca^{2+} ion concentrations. *Biochemical Journal* 327 (Pt 2), 545–552.
- Amdani, S.N., Jones, C., Coward, K., 2013. Phospholipase C zeta (PLCzeta): Oocyte activation and clinical links to male factor infertility. *Advances in Biological Regulation* 53, 292–308.
- Arteaga, C.L., Johnson, M.D., Todderud, G., *et al.*, 1991. Elevated content of the tyrosine kinase substrate phospholipase C-gamma 1 in primary human breast carcinomas. *Proceedings of the National Academy of Sciences of the United States of America* 88, 10435–10439.
- Bae, J.H., Lew, E.D., Yuzawa, S., *et al.*, 2009. The selectivity of receptor tyrosine kinase signaling is controlled by a secondary SH2 domain binding site. *Cell* 138, 514–524.
- Bai, Y., Edamatsu, H., Maeda, S., *et al.*, 2004. Crucial role of phospholipase Cepsilon in chemical carcinogen-induced skin tumor development. *Cancer Research* 64, 8808–8810.
- Behjati, S., Tarpey, P.S., Sheldon, H., *et al.*, 2014. Recurrent PTPRB and PLCG1 mutations in angiosarcoma. *Nature Genetics* 46, 376–379.
- Brose, N., Rosenmund, C., 2002. Move over protein kinase C, you've got company: Alternative cellular effectors of diacylglycerol and phorbol esters. *Journal of Cell Science* 115, 4399–4411.
- Bunney, T.D., Esposito, D., Mas-Droux, C., *et al.*, 2012. Structural and functional integration of the PLCgamma interaction domains critical for regulatory mechanisms and signaling deregulation. *Structure* 20, 2062–2075.
- Bunney, T.D., Harris, R., Gandarillas, N.L., *et al.*, 2006. Structural and mechanistic insights into ras association domains of phospholipase C epsilon. *Molecular Cell* 21, 495–507.
- Bunney, T.D., Katan, M., 2006. Phospholipase C epsilon: Linking second messengers and small GTPases. *Trends in Cell Biology* 16, 640–648.
- Bunney, T.D., Katan, M., 2010. Phosphoinositide signaling in cancer: Beyond PI3K and PTEN. *Nature Reviews Cancer* 10, 342–352.
- Bunney, T.D., Katan, M., 2011. PLC regulation: Emerging pictures for molecular mechanisms. *Trends in Biochemical Sciences* 36, 88–96.
- Bunney, T.D., Opaleye, O., Roe, S.M., *et al.*, 2009. Structural insights into formation of an active signaling complex between Rac and phospholipase C gamma 2. *Molecular Cell* 34, 223–233.
- Camps, M., Carozzi, A., Schnabel, P., *et al.*, 1992. Isozyme-selective stimulation of phospholipase C-beta 2 by G protein beta gamma-subunits. *Nature* 360, 684–686.
- Danielsen, S.A., Cekaite, L., Agesen, T.H., *et al.*, 2011. Phospholipase C isozymes are deregulated in colorectal cancer – Insights gained from gene set enrichment analysis of the transcriptome. *PLoS One* 6, e24419.
- Essen, L.O., Perisic, O., Cheung, R., Katan, M., Williams, R.L., 1996. Crystal structure of a mammalian phosphoinositide-specific phospholipase C delta. *Nature* 380, 595–602.
- Essen, L.O., Perisic, O., Katan, M., *et al.*, 1997. Structural mapping of the catalytic mechanism for a mammalian phosphoinositide-specific phospholipase C. *Biochemistry* 36, 1704–1718.
- Everett, K.L., Bunney, T.D., Yoon, Y., *et al.*, 2009. Characterization of phospholipase C gamma enzymes with gain-of-function mutations. *Journal of Biological Chemistry* 284, 23083–23093.
- Feng, J.F., Rhee, S.G., Im, M.J., 1996. Evidence that phospholipase delta1 is the effector in the Gh (transglutaminase II)-mediated signaling. *Journal of Biological Chemistry* 271, 16451–16454.
- Fu, L., Qin, Y.R., Xie, D., *et al.*, 2007. Characterization of a novel tumor-suppressor gene PLC delta 1 at 3p22 in esophageal squamous cell carcinoma. *Cancer Research* 67, 10720–10726.
- Fukami, K., Ichinohe, M., Hirata, M., Nakamura, Y., 2008. Physiological functions of phospholipase C delta-type. *Advances in Enzyme Regulation* 48, 261–273.
- Fukami, K., Nakao, K., Inoue, T., *et al.*, 2001. Requirement of phospholipase Cdelta4 for the zona pellucida-induced acrosome reaction. *Science* 292, 920–923.
- Gresset, A., Sondek, J., Harden, T.K., 2012. The phospholipase C isozymes and their regulation. *Sub-Cellular Biochemistry* 58, 61–94.
- Harden, T.K., Sondek, J., 2006. Regulation of phospholipase C isozymes by ras superfamily GTPases. *Annual Review of Pharmacology and Toxicology* 46, 355–379.
- Harden, T.K., Waldo, G.L., Hicks, S.N., Sondek, J., 2011. Mechanism of activation and inactivation of Gq/phospholipase C-beta signaling nodes. *Chemical Reviews* 111, 6120–6129.
- Heytens, E., Parrington, J., Coward, K., *et al.*, 2009. Reduced amounts and abnormal forms of phospholipase C zeta (PLCzeta) in spermatozoa from infertile men. *Human Reproduction* 24, 2417–2428.
- Hinkes, B., Wiggins, R.C., Gbadegesin, R., *et al.*, 2006. Positional cloning uncovers mutations in PLCE1 responsible for a nephrotic syndrome variant that may be reversible. *Nature Genetics* 38, 1397–1405.
- Hu, X.T., Zhang, F.B., Fan, Y.C., *et al.*, 2009. Phospholipase C delta 1 is a novel 3p22.3 tumor suppressor involved in cytoskeleton organization, with its epigenetic silencing correlated with high-stage gastric cancer. *Oncogene* 28, 2466–2475.
- Hwang, J.I., Oh, Y.S., Shin, K.J., *et al.*, 2005. Molecular cloning and characterization of a novel phospholipase C, PLC-eta. *Biochemical Journal* 389, 181–186.
- Ichinohe, M., Nakamura, Y., Sai, K., *et al.*, 2007. Lack of phospholipase C-delta1 induces skin inflammation. *Biochemical and Biophysical Research Communications* 356, 912–918.
- Ikuta, S., Edamatsu, H., Li, M., Hu, L., Kataoka, T., 2008. Crucial role of phospholipase C epsilon in skin inflammation induced by tumor-promoting phorbol ester. *Cancer Research* 68, 64–72.
- Illenberger, D., Walliser, C., Nurnberg, B., Diaz Lorente, M., Gierschik, P., 2003. Specificity and structural requirements of phospholipase C-beta stimulation

- by Rho GTPases versus G protein beta gamma dimers. *Journal of Biological Chemistry* 278, 3006–3014.
- Jezyk, M.R., Snyder, J.T., Gershberg, S., *et al.*, 2006. Crystal structure of Rac1 bound to its effector phospholipase C-beta2. *Nature Structural & Molecular Biology* 13, 1135–1140.
- Ji, Q.S., Winnier, G.E., Niswender, K.D., *et al.*, 1997. Essential role of the tyrosine kinase substrate phospholipase C-gamma1 in mammalian growth and development. *Proceedings of the National Academy of Sciences of the United States of America* 94, 2999–3003.
- Jiang, H., Kuang, Y., Wu, Y., *et al.*, 1997. Roles of phospholipase C beta2 in chemoattractant-elicited responses. *Proceedings of the National Academy of Sciences of the United States of America* 94, 7971–7975.
- Jin, T.G., Satoh, T., Liao, Y., *et al.*, 2001. Role of the CDC25 homology domain of phospholipase Cepsilon in amplification of Rap1-dependent signaling. *Journal of Biological Chemistry* 276, 30301–30307.
- Kadamur, G., Ross, E.M., 2013. Mammalian phospholipase C. *Annual Review of Physiology* 75, 127–154.
- Kanemaru, K., Nakahara, M., Nakamura, Y., *et al.*, 2010. Phospholipase C-eta2 is highly expressed in the habenula and retina. *Gene Expression Patterns* 10, 119–126.
- Kashir, J., Normikos, M., Lai, F.A., Swann, K., 2014. Sperm-induced Ca release during egg activation in mammals. *Biochemical and Biophysical Research Communications* 450, 1204–1211.
- Katan, M., 2005. New insights into the families of PLC enzymes: Looking back and going forward. *Biochemical Journal* 391, e7–e9.
- Kim, D., Jun, K.S., Lee, S.B., *et al.*, 1997. Phospholipase C isozymes selectively couple to specific neurotransmitter receptors. *Nature* 389, 290–293.
- Lemmon, M.A., 2008. Membrane recognition by phospholipid-binding domains. *Nature Reviews Molecular Cell Biology* 9, 99–111.
- Li, M., Edamatsu, H., Kitazawa, R., Kitazawa, S., Kataoka, T., 2009. Phospholipase Cepsilon promotes intestinal tumorigenesis of Apc(Min/+) mice through augmentation of inflammation and angiogenesis. *Carcinogenesis* 30, 1424–1432.
- Liao, H.J., Kume, T., McKay, C., *et al.*, 2002. Absence of erythropoiesis and vasculogenesis in Plcg1-deficient mice. *Journal of Biological Chemistry* 277, 9335–9341.
- Lopez, I., Mak, E.C., Ding, J., Hamm, H.E., Lomasney, J.W., 2001. A novel bifunctional phospholipase c that is regulated by Galpha 12 and stimulates the Ras/mitogen-activated protein kinase pathway. *Journal of Biological Chemistry* 276, 2758–2765.
- Lyon, A.M., Taylor, V.G., Tesmer, J.J., 2014. Strike a pose: Galphaq complexes at the membrane. *Trends in Pharmacological Sciences* 35, 23–30.
- Lyon, A.M., Tesmer, V.M., Dhamania, V.D., *et al.*, 2011. An autoinhibitory helix in the C-terminal region of phospholipase C-beta mediates Galphaq activation. *Nature Structural & Molecular Biology* 18, 999–1005.
- Martins, M., McCarthy, A., Baxendale, R., *et al.*, 2014. Tumor suppressor role of phospholipase C epsilon in Ras-triggered cancers. *Proceedings of the National Academy of Sciences of the United States of America* 111, 4239–4244.
- Nakamura, Y., Hamada, Y., Fujiwara, T., *et al.*, 2005. Phospholipase C-delta1 and -delta3 are essential in the trophoblast for placental development. *Molecular and Cellular Biology* 25, 10979–10988.
- Noh, D.Y., Lee, Y.H., Kim, S.S., *et al.*, 1994. Elevated content of phospholipase C-gamma 1 in colorectal cancer tissues. *Cancer* 73, 36–41.
- Normikos, M., Blayney, L.M., Larman, M.G., *et al.*, 2005. Role of phospholipase C-zeta domains in Ca2+ -dependent phosphatidylinositol 4,5-bisphosphate hydrolysis and cytoplasmic Ca2+ oscillations. *Journal of Biological Chemistry* 280, 31011–31018.
- Ombrello, M.J., Remmers, E.F., Sun, G., *et al.*, 2012. Cold urticaria, immunodeficiency, and autoimmunity related to PLCG2 deletions. *New England Journal of Medicine* 366, 330–338.
- Piechulek, T., Rehlen, T., Walliser, C., *et al.*, 2005. Isozyme-specific stimulation of phospholipase C-gamma2 by Rac GTPases. *Journal of Biological Chemistry* 280, 38923–38931.
- Sala, G., Dituri, F., Raimondi, C., *et al.*, 2008. Phospholipase Cgamma1 is required for metastasis development and progression. *Cancer Research* 68, 10187–10196.
- Sidhu, R.S., Clough, R.R., Bhullar, R.P., 2005. Regulation of phospholipase C-delta1 through direct interactions with the small GTPase Ral and calmodulin. *Journal of Biological Chemistry* 280, 21933–21941.
- Smrcka, A.V., Brown, J.H., Holz, G.G., 2012. Role of phospholipase Cepsilon in physiological phosphoinositide signaling networks. *Cell Signal* 24, 1333–1343.
- Sorli, S.C., Bunney, T.D., Sugden, P.H., Paterson, H.F., Katan, M., 2005. Signaling properties and expression in normal and tumor tissues of two phospholipase C epsilon splice variants. *Oncogene* 24, 90–100.
- Stewart, A.J., Morgan, K., Farquharson, C., Millar, R.P., 2007. Phospholipase C-eta enzymes as putative protein kinase C and Ca2+ signaling components in neuronal and neuroendocrine tissues. *Neuroendocrinology* 86, 243–248.
- Suh, B.C., Hille, B., 2008. PIP2 is a necessary cofactor for ion channel function: How and why? *Annual Review of Biophysics* 37, 175–195.
- Suh, P.G., Park, J.I., Manzoli, L., *et al.*, 2008. Multiple roles of phosphoinositide-specific phospholipase C isozymes. *BMB Reports* 41, 415–434.
- Tadano, M., Edamatsu, H., Minamisawa, S., *et al.*, 2005. Congenital semilunar valvulogenesis defect in mice deficient in phospholipase C epsilon. *Molecular and Cellular Biology* 25, 2191–2199.
- Taylor, S.L., Yoon, S.Y., Morshedi, M.S., *et al.*, 2010. Complete globozoospermia associated with PLCzeta deficiency treated with calcium ionophore and ICSI results in pregnancy. *Reproductive BioMedicine Online* 20, 559–564.
- Thomas, S.M., Coppelli, F.M., Wells, A., *et al.*, 2003. Epidermal growth factor receptor-stimulated activation of phospholipase Cgamma-1 promotes invasion of head and neck squamous cell carcinoma. *Cancer Research* 63, 5629–5635.
- Thompson, J.L., Shuttleworth, T.J., 2011. Orai channel-dependent activation of phospholipase C-delta: A novel mechanism for the effects of calcium entry on calcium oscillations. *Journal of Physiology* 589, 5057–5069.
- Van Raamsdonk, C.D., Bezrukov, V., Green, G., *et al.*, 2009. Frequent somatic mutations of GNAQ in uveal melanoma and blue naevi. *Nature* 457, 599–602.
- Vaque, J.P., Gomez-Lopez, G., Monsalve, V., *et al.*, 2014. PLCG1 mutations in cutaneous T-cell lymphomas. *Blood* 123, 2034–2043.
- Waldo, G.L., Ricks, T.K., Hicks, S.N., *et al.*, 2010. Kinetic scaffolding mediated by a phospholipase C-beta and Gq signaling complex. *Science* 330, 974–980.
- Wang, D., Feng, J., Wen, R., *et al.*, 2000a. Phospholipase Cgamma2 is essential in the functions of B cell and several Fc receptors. *Immunity* 13, 25–35.
- Wang, H., Oestreich, E.A., Maekawa, N., *et al.*, 2005. Phospholipase C epsilon modulates beta-adrenergic receptor-dependent cardiac contraction and inhibits cardiac hypertrophy. *Circulation Research* 97, 1305–1313.
- Wang, L.D., Zhou, F.Y., Li, X.M., *et al.*, 2010. Genome-wide association study of esophageal squamous cell carcinoma in Chinese subjects identifies susceptibility loci at PLCE1 and C20orf54. *Nature Genetics* 42, 759–763.
- Wang, T., Dowal, L., El-Maghrabi, M.R., Rebecchi, M., Scarlata, S., 2000b. The pleckstrin homology domain of phospholipase C-beta(2) links the binding of gbetagamma to activation of the catalytic core. *Journal of Biological Chemistry* 275, 7466–7469.
- Wang, X., Zhou, C., Qiu, G., *et al.*, 2012. Phospholipase C epsilon plays a suppressive role in incidence of colorectal cancer. *Medical Oncology* 29, 1051–1058.
- Wang, Z., Liu, B., Wang, P., *et al.*, 2008. Phospholipase C beta3 deficiency leads to macrophage hypersensitivity to apoptotic induction and reduction of atherosclerosis in mice. *Journal of Clinical Investigation* 118, 195–204.
- Xiao, W., Hong, H., Kawakami, Y., *et al.*, 2009. Tumor suppression by phospholipase C-beta3 via SHP-1-mediated dephosphorylation of Stat5. *Cancer Cell* 16, 161–171.
- Yagisawa, H., Sakuma, K., Paterson, H.F., *et al.*, 1998. Replacements of single basic amino acids in the pleckstrin homology domain of phospholipase C-delta1 alter the ligand binding, phospholipase activity, and interaction with the plasma membrane. *Journal of Biological Chemistry* 273, 417–424.
- Yoon, S.Y., Jellerette, T., Salicioni, A.M., *et al.*, 2008. Human sperm devoid of PLC, zeta 1 fail to induce Ca(2+) release and are unable to initiate the first step of embryo development. *Journal of Clinical Investigation* 118, 3671–3681.
- Yu, P., Constien, R., Dear, N., *et al.*, 2005. Autoimmunity and inflammation due to a gain-of-function mutation in phospholipase C gamma 2 that specifically increases external Ca2+ entry. *Immunity* 22, 451–465.
- Zhang, L., Malik, S., Kelley, G.G., Kapiloff, M.S., Smrcka, A.V., 2011. Phospholipase C epsilon scaffolds to muscle-specific A kinase anchoring protein (mAKAPbeta) and integrates multiple hypertrophic stimuli in cardiac myocytes. *Journal of Biological Chemistry* 286, 23012–23021.
- Zhang, L., Malik, S., Pang, J., *et al.*, 2013. Phospholipase Cepsilon hydrolyzes perinuclear phosphatidylinositol 4-phosphate to regulate cardiac hypertrophy. *Cell* 153, 216–227.
- Zhou, Q., Lee, G.S., Brady, J., *et al.*, 2012. A hypermorphic missense mutation in PLCG2, encoding phospholipase Cgamma2, causes a dominantly inherited autoinflammatory disease with immunodeficiency. *American Journal of Human Genetics* 91, 713–720.
- Zhou, Y., Wing, M.R., Sondek, J., Harden, T.K., 2005. Molecular cloning and characterization of PLC-eta2. *Biochemical Journal* 391, 667–676.

Calcium and Calmodulin Signaling

E Carafoli, University of Padova, Padova, Italy

J Krebs, MPI for Biophysical Chemistry, Goettingen, Germany

© 2016 Elsevier Inc. All rights reserved.

Glossary

Apoptosis/programmed cell death Death of cells by mechanisms that are carefully controlled: a positive way to use the calcium signal.

Calcium pathologies Uncontrolled activation of hydrolytic enzymes (proteases, phospholipases, and nucleases) by excess calcium.

Calcium signaling Transfer of information by calcium to intracellular protein targets.

Calmodulin A protein that binds calcium to decode its message and to transfer it to protein targets.

Protein kinases Enzymes that hydrolyze ATP to phosphorylate proteins.

Calcium, the third most abundant metal in nature, is present in living organisms in salts of various compositions. In animals, calcium phosphate is the major salt of bones and teeth, which contain more than 99% of the total Ca^{2+} of the body. In man, the concentration of Ca^{2+} in the body fluids and the extracellular spaces oscillates between 2.1 and 2.6 mM (Woodhead, 1982). The concentration of Ca^{2+} in the blood and the extracellular spaces, and the equilibrium between its free (ionized), and bound and/or salt-precipitated forms, is regulated by three hormones, parathyroid hormone, calcitonin, and Vitamin D. They act on the three systems that preside over the dynamic equilibrium of Ca^{2+} in the body: the intestine, the kidney, and the bone compartment. About half of the blood Ca^{2+} is free, the other half is complexed to organic or inorganic compounds. Total Ca^{2+} is in the millimolar concentration range also inside cells, however, its free concentration in the cytosol, where most of the targets of its signaling function reside, is but a negligible fraction of the total: normally, it is maintained in the 100–300 nM level, which is the level in which the regulation of the signaling function of Ca^{2+} occurs. The balance between free and bound Ca^{2+} in the cytosol is a complex process in which the binding of the metal to inorganic ligands and small molecular weight organic molecules has a role (Brini *et al.*, 2013b). However, the complexation by these ligands cannot lower the concentration of free Ca^{2+} to the nanomolar level. This low concentration is achieved thanks to the intervention of specific proteins that coordinate Ca^{2+} to complex sites that have the appropriate affinity and specificity for it. These proteins exploit the peculiar coordination chemistry of Ca^{2+} , which determines the ease with which it can be ligated to the sites of irregular geometry their structures offer. Ca^{2+} -ligation usually occurs via carboxylates (mono- or bidentate) or neutral oxygen donors since the calcium ion prefers oxygen as donor groups of ligands. In proteins, calcium is generally bound by seven oxygen atoms in a conformation which is best represented by a pentagonal bipyramid.

The advantage of Ca^{2+} in the ligation process becomes clear on comparing it, for instance, to the other Group 2 metal, Mg^{2+} . The range of distances Mg^{2+} can accept from the coordinating oxygens of the binding site is much more restricted than in the case of Ca^{2+} : Mg^{2+} would thus demand perfectly

octahedral-binding sites, which are obviously unlikely to be found in proteins, whereas Ca^{2+} easily accepts the sites of more irregular geometry that one may expect in their structures (Carafoli and Penniston, 1985). Proteins endowed with the ability to specifically bind Ca^{2+} may be divided in two large families: those that transport it across membranes and those that reversibly complex it in the cytosol and in the lumen of the organelles.

The EF-Hand Calcium-Binding Principle

The intracellular level of free Ca^{2+} in eukaryotic cells can transiently increase above the 100–300 nM concentration level in response to extracellular signals. To function as a carrier of messages the Ca^{2+} -signal must be processed by Ca^{2+} -binding proteins. At variance with the extracellular proteins that bind Ca^{2+} with low affinity, intracellular proteins bind it with high affinity to a number of structural motifs. The most common (and most important) is a sequential arrangement of about 30 amino acids that ligate Ca^{2+} in a loop flanked by two helical segments. This helix-loop-helix Ca^{2+} -binding motif, which is known as the 'EF-hand' was introduced by Kretsinger (Kretsinger and Nockolds, 1973) from the 3D structure of parvalbumin. It revealed six α -helices, designated A through F: the spatial arrangement of the fifth (E) and sixth (F) α -helix enclosing the calcium-binding loop resemble a human hand, the 'EF-Hand,' as shown in Figure 1. After the original parvalbumin work, the 'EF-hand' motif was discovered in numerous other proteins as a highly conserved structural Ca^{2+} -binding arrangement in intracellular proteins that mediate cellular responses. The family of proteins which contain EF-Hand motifs, presently numbers more than 800 members, and is steadily increasing, (Kawasaki *et al.*, 1998). Calmodulin (CaM), troponin C, recoverin, and S100 are some of the important proteins of the superfamily.

All EF-hand domains show a pentagonal-bipyramidal coordination pattern of Ca^{2+} in the loop flanked by two helices approximately perpendicular to each other. The residues forming the ligands to Ca^{2+} are highly conserved within a contiguous sequence of 12 residues spanning the loop and the beginning of the second helix.

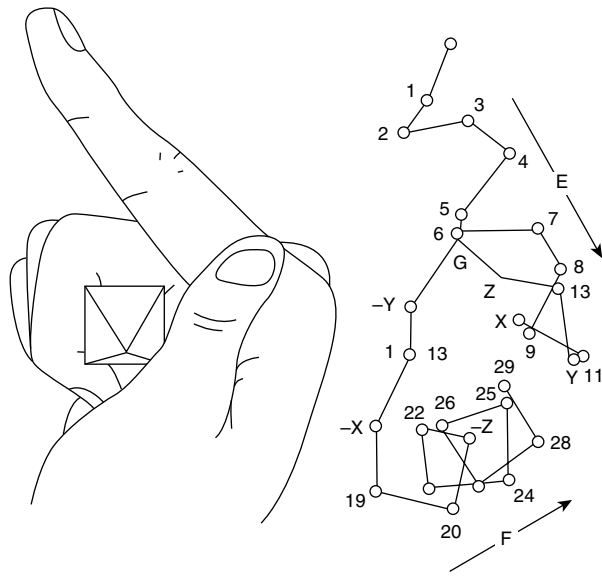


Figure 1 The EF-hand as proposed by Kretsinger based on the crystal structure of parvalbumin can be represented by the forefinger (helix E) and the thumb (helix F), both enclosing the Ca^{2+} -binding loop, represented by the bent middle finger. Adapted from Kretsinger, R.H., 1975. Hypothesis: Calcium modulated proteins contain EF hands. In: Carafoli, E., Clementi, F., Drabikowsky, W., Margreth, A. (Eds.), *Calcium Transport in Contraction and Secretion*. Amsterdam: Elsevier, pp. 469–478.

The Ca^{2+} -binding domains usually occur in pairs stabilized by hydrogen-bond bridges between the central residues of adjacent loops, which form a minimal antiparallel beta-sheet. The pair-forming Ca^{2+} -binding domains enhance the Ca^{2+} -affinity and the cooperativity of binding. CaM and troponin C, the Ca^{2+} -binding protein of muscles, contain four such Ca^{2+} -binding sites. Both proteins display Ca^{2+} -binding characteristics compatible with a 'pair of pairs'-model of EF-hands, i.e., the two globular at the N- and C-terminal domains of the protein each contain a pair of EF-hands, the two globular domains being connected by a long helix that confers a dumbbell-shaped appearance to the protein. However, the central portion of the helix connecting the terminal lobes acts as a flexible hinge, which plays an important role in the interaction of the proteins with targets as it permits a broad range of spacing between the hydrophobic anchor residues in the globular lobes. Upon binding of Ca^{2+} , and subsequent interaction with targets, a conformational change of the protein occurs that transforms its extended, dumbbell-shaped structure into a more compact hairpin-like form that wraps itself around the target (Yap *et al.*, 1999).

A number of cellular Ca^{2+} -binding proteins are not structurally organized on the EF-hand principle. Among them are the Ca^{2+} -buffering proteins of the endoplasmic reticulum calreticulin and calsequestrin, phospholipase A2, protein kinase C, membrane-intrinsic proteins such as the annexins, and proteins containing the structural motif defined as the C_2 domain. The large family of annexins is of particular interest. Its members bind to membranes containing negatively charged phospholipids in a Ca^{2+} -dependent manner and

have a role in a number of important cellular functions, such as membrane fusion, signal transduction, and phospholipase modulation (Gerke *et al.*, 2005). An important member of the C_2 domain containing proteins is synaptotagmin, the Ca^{2+} -sensor required for membrane fusion during endo- or exocytosis.

CaM

CaM is the most conserved protein of the EF-hand superfamily, and is ubiquitously expressed in eukaryotic organisms (Klee and Vanaman, 1982). The sequences of CaMs of all vertebrates are 100% identical. In fact, CaM is ranked among the most conserved proteins known so far, together with proteins like histones, actin, or ubiquitin (Copley *et al.*, 1999). Human CaM is encoded by three nonallelic genes (Fischer *et al.*, 1988). The three genes encode an identical protein, but the coding sequences differ substantially.

CaM is the modulator that mediates the Ca^{2+} signal to most targets in eukaryotic cells. It regulates numerous fundamental cellular activities, among them glycogen metabolism, intracellular motility, Ca^{2+} -transport across membranes, cyclic nucleotide metabolism, protein phosphorylation and dephosphorylation, cell cycle progression, and gene expression. It controls these processes by interacting directly with a number of key enzymes such as the Ca^{2+} ATPase (PMCA pump) of the plasma membrane, different kinases, and the protein phosphatase calcineurin (CaN). It can regulate proteins with opposite functions, like adenylate cyclases and cyclic nucleotide phosphodiesterases. In addition to interacting reversibly with targets, in some cases CaM acts as a permanent subunit of enzymes, for example, phosphorylase kinase (Shenolikar *et al.*, 1979) and CaM-dependent kinases of plants (Roberts and Harmon, 1992). CaM is a monomer of approximate molecular weight 17 000, which is mainly located in the cytosol. However, it can be translocated to the nucleus (Pruschy *et al.*, 1994) where it plays a role in gene transcription and regulation of gene expression.

The structure of CaM has been solved for both the Ca^{2+} -saturated (Babu *et al.*, 1985) and the Ca^{2+} -free form (Kuboniva *et al.*, 1995; Zhang *et al.*, 1995). As mentioned, upon binding of Ca^{2+} , CaM alters its conformation exposing hydrophobic patches on its surface (LaPorte *et al.*, 1980; Krebs *et al.*, 1984; Figure 2). This confers to the protein the ability to interact with targets. A defined consensus sequence for CaM-binding sites in target proteins does not exist, but a number of characteristic features have been described such as high helix propensity, net positive charge, and two hydrophobic anchor residues spaced by a certain number of amino acids (James *et al.*, 1995). A number of structures of Ca^{2+} -loaded CaM complexed with binding peptides corresponding to the binding domains of protein targets have been solved (e.g., see Elshorst *et al.*, 1999; Juranic *et al.*, 2010; for a detailed review see Villarroel *et al.*, 2014). These studies have provided evidence for the general conclusion briefly mentioned above that the binding of target peptides induces a large conformational change of Ca^{2+} saturated CaM, which bends from the extended dumbbell shape into a more globular structure in which the two halves of CaM wrap around the target peptide

(Ikura *et al.*, 1992; Meador *et al.*, 1992; Figure 3). High-resolution structures of CaM complexed to entire protein targets exist only for two calmodulin-dependent kinases, CaM kinase II and the death-associated protein kinase (DAPK) (Rellos *et al.*, 2010; De Diego *et al.*, 2010, see below). The general molecular pattern of the molecular interaction of CaM with target domains that has emerged from these studies is supported by experiments using small-angle X-ray diffraction and neutron scattering. In these experiments, CaM complexed to intact myosin light chain kinase (MLCK) CaM undergoes a conformational collapse identical to that observed with the peptide corresponding to the CaM-binding domain of MLCK (Krueger *et al.*, 1997).

In addition to the CaM-binding motif in which hydrophobic residues spaced by a certain number of amino acids bind to hydrophobic pockets in CaM, a different motif for

CaM binding has been characterized more recently, the so-called IQ-motif, to which CaM can bind even in the absence of Ca^{2+} . This motif contains a conserved isoleucine next to a conserved glutamine and, C-terminal to it, at least one basic residue, normally arginine, with a number of residues in between (usually 3 or 4). The consensus sequence is thus $\text{IQX}_{(3\text{or}4)}\text{RGX}_{(3\text{or}4)}\text{R}$. A prominent example of this motif is found in the voltage-dependent calcium channel, which contains an IQ-motif in the C-terminal portion of its α subunit (see below). X-ray work has shown that CaM binds to a binding peptide that includes the IQ-motif (Fallon *et al.*, 2005).

Regulation of Ca^{2+} by Proteins that Mediate Its Membrane Transport

Ca^{2+} Channels

Ca^{2+} enters the cytosol from the external spaces, and the lumen of the intracellular organelles, through carefully controlled proteinaceous channels (see Brini *et al.*, 2013a for a recent review). Those in the plasma membrane have been historically divided in three major groups, depending on the mechanism by which their operation is regulated: (1) the voltage gated channels (VOCs), which transduce the changes in membrane potential into intracellular Ca^{2+} transients, (2) the receptor-operated channels (ROCs), which are activated by extracellular ligands, and (3) the store-operated channels (SOCs), which are activated by the emptying of the cellular Ca^{2+} stores (Figure 4). The VOCs are complexes of five subunits: the α 1 subunit, which contains the channel proper, is organized in four repeated modules of six transmembrane domains, the fourth of which contains the voltage sensor that gates the channel. The most important ROCs are those activated by the brain excitatory transmitter L-glutamate, which also interacts with a canonical seven-transmembrane receptor

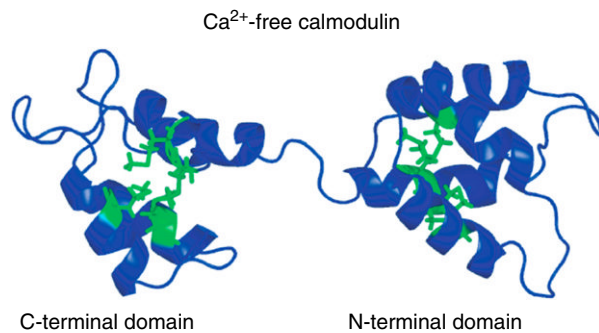


Figure 2 Ca^{2+} -free (apo) calmodulin (1DMO.pdb). The blue ribbons show secondary structural elements of CaM, the green sticks show different methionine residues important for the interaction with targets. The overall structure shows the dumbbell shape typical for target-free CaM. Reproduced from Stathopoulos, P.B., Ames, J.B., Ikura, M., 2007. Structural aspects of calcium binding proteins and their interactions with targets. In: Krebs, J., Michalak, M. (Eds.), Calcium: A Matter of Life or Death. Amsterdam: Elsevier, pp. 95–123.

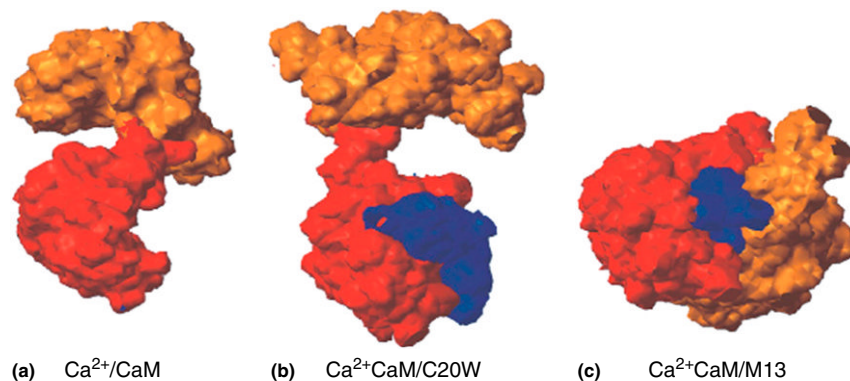


Figure 3 Surface representation of Ca^{2+} -bound CaM, and of its different binding modes to peptides. The N-terminal halves of CaM are in orange, the C-terminal halves are in red, the different peptides bound to CaM are in blue. The orientation of the C-terminal half of CaM is the same for all three structures. (a) Crystal structure of Ca^{2+} -bound calmodulin. (b) Solution structure of Ca^{2+} -CaM bound to a peptide (C20W) representing the N-terminal portion of the CaM-binding domain of the plasma membrane Ca^{2+} pump (lacking the C-terminal hydrophobic anchor residue): the peptide only binds to the C-terminal half of CaM. (c) Solution structure of the complex of Ca^{2+} -CaM with the CaM-binding domain of MLCK (M13) demonstrating that CaM changes to a compact globular conformation in which its N- and C-terminal halves wrap around M13. Reproduced from Elshorst, B., Hennig, M., Försterling, H., *et al.*, 1999. NMR solution structure of a complex of calmodulin with a binding peptide of the Ca^{2+} pump. Biochemistry 38, 12320–12332.

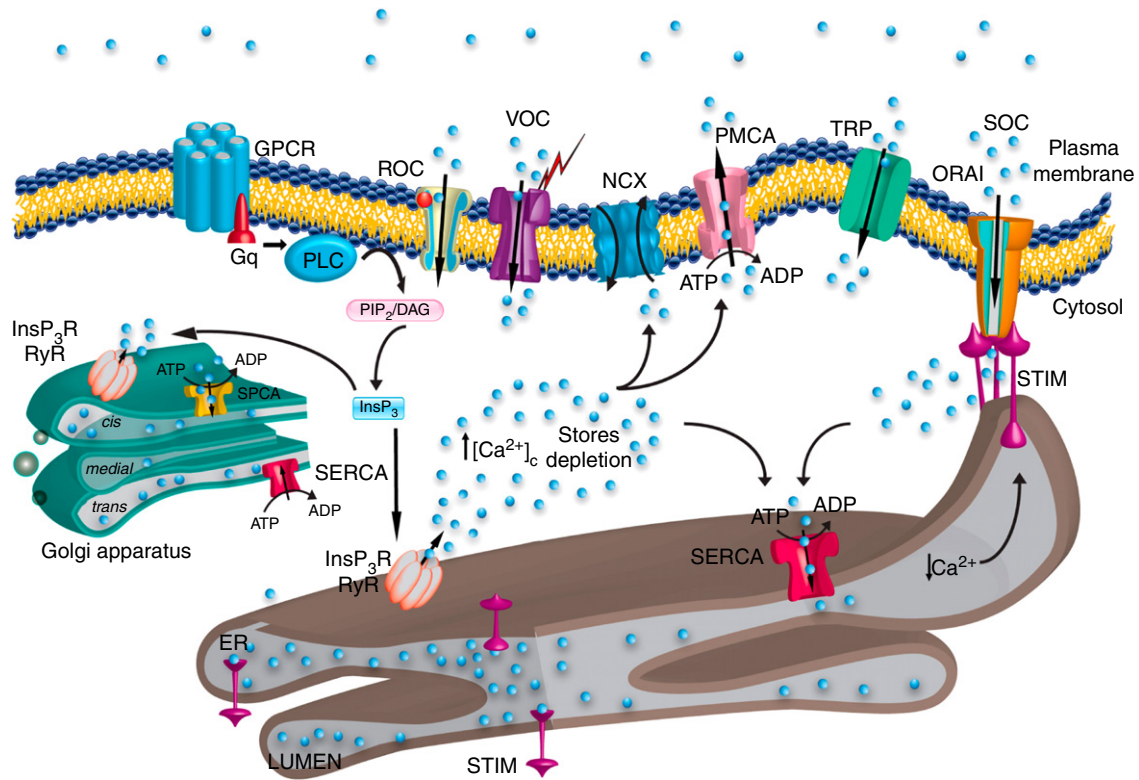


Figure 4 Schematic representation of the Ca^{2+} transporters and channels in the plasma membrane and the intracellular organelles. Most of the systems shown in the figure are mentioned in the text. G-protein-coupled receptor (GPCR) indicates the seven transmembrane domains receptors linked to Gq proteins that liberate InsP_3 in response to specific agonists by acting on the phospholipase C that hydrolyzes phosphatidyl-inositol-bis-phosphate (PIP₂). DAG is diacyl-glycerol. Transient receptor potentials (TRPs) are channels that can also mediate Ca^{2+} intake. The cartoon does not show the acidic organelles that have also been claimed to be a Ca^{2+} store. Modified from Brini, M., Ottolini, D., Cali, T., 2013a. Intracellular calcium homeostasis and signaling, In: Banci, L. (Ed.), *Metallomics and the Cell, Metal Ions in Life Sciences*, vol. 12. Dordrecht: Springer Science + Business Media, pp. 119–168.

coupled to G-proteins. The L-glutamate activated channels are divided into three groups based on the activity of specific agonists. They are multisubunit, relatively nonselective cation channels, in which the subunits contribute four transmembrane domains that form the cationic channel (however, transmembrane domain 2 does not cross the full width of the membrane). The SOCs are channels that are operated by the interplay of transmembrane Ca^{2+} -binding proteins in the endoplasmic reticulum that sense the level of Ca^{2+} filling within its lumen (STIM proteins). The STIM proteins communicate physically with a plasma membrane counterpart, which is the Ca^{2+} forming channel (the ORAI proteins). The ORAI protein family is composed of three subunits: ORAI 1 contains four transmembrane that form the pore (Hou *et al.*, 2012). When depletion of the Ca^{2+} store occurs, STIM proteins undergo a conformational change that redistributes them to discrete plasma membrane districts where ORAI proteins also redistribute (see Figure 4). Their interaction allows the opening of the ORAI1 channel.

Two channel types mediate the efflux of Ca^{2+} from the endo(sarco)plasmic reticulum and the Golgi system, which are the main intracellular Ca^{2+} stores: the inositol 1,4,5-trisphosphate (InsP_3) operated receptor/channel, which is expressed ubiquitously in tissues, and the ryanodine receptor/channel

(RyR), which has tissue restricted expression. Both are very large tetrameric molecules, which span the membrane with six hydrophobic helices. The very large portion of the molecules that protrudes into the cytosol contains binding sites for numerous regulators of the channel activity. One of them is Ca^{2+} , which acts either as a positive or a negative agonist of the InsP_3 R depending on its concentration. It activates the opening of the RyR in heart (triggering the Ca^{2+} -induced Ca^{2+} release mechanism, CICR), but not in skeletal muscles, where the activation of the channel requires the physical interaction between the plasma membrane VOCs and the RyR (the electromechanical coupling). Both the InsP_3 R and the RyRs are products of multigene families: three variants of each have been described, having tissue specific expression. Recent X-ray work has clarified many structural aspects of the function and regulation of the InsP_3 R, and the full 3D structure of the RyR has been solved by cryo-EM work (see for instance Zalk *et al.*, 2015).

Ca^{2+} Transporters

The ejection of Ca^{2+} from the cytosol to the extracellular space and from the lumen of intracellular organelles is performed by

ATPases (pumps) and $\text{Na}^+/\text{Ca}^{2+}$ -exchangers (Figure 4). Animal cells express Ca^{2+} pumps in the plasma membrane (PMCA pumps), in the endo(sarco)plasmic reticulum (SERCA pumps), and in the Golgi vesicles (SPCA pumps) (other Ca^{2+} pumps are expressed in plant cells and in lower eukaryotes). All these pumps belong to the superfamily of P-type ATPases, which are characterized by the temporary conservation of the energy of ATP in the form of an aspartyl-phosphate during the reaction cycle. They are organized in the membrane with 10 transmembrane domains, the larger portion of the structure protruding into the cytosol. A distinctive property of the PMCA pump is the presence of a long cytosolic C-terminal tail that houses important regulatory sites. The molecular details of the path of Ca^{2+} across the molecule of P-type pumps have been clarified by a number of X-ray studies on the SERCA pump (Toyoshima *et al.*, 2000). They have then been used as templates to model the molecular path for Ca^{2+} in the other two pumps. All three Ca^{2+} pumps are the products of multigene families, the isoforms having different tissue distribution and functional properties (mainly pertaining to their regulation) (Brini and Carafoli, 2009). Quantitatively, the SERCA pump has the most important role in the regulation of the homeostasis of Ca^{2+} in the cytosol, and thus in the overall control of the activity of the targets of Ca^{2+} signaling, whereas the SPCA pump may be predominant in special cell types, for example, the keratinocytes. The Ca^{2+} ejection activity of the PMCA pump probably does not have a significant role in the regulation of the homeostasis of Ca^{2+} in the bulk cytosol, considering the much greater amounts of the SERCA pump existing in most cells and, especially, the presence of the much more powerful $\text{Na}^+/\text{Ca}^{2+}$ -exchanger in the plasma membrane (see below). The concept has thus recently emerged that the main role of the PMCA pump would be that of regulating the concentration of Ca^{2+} in selected microdomains around its cytosolic protruding portion, which may contain important Ca^{2+} sensitive enzymes (Mohamed *et al.*, 2011).

The PMCA pump is of particular interest from another angle as well, as it is a target of the regulation by CaM, and is thus a striking example of the autoregulatory property of the Ca^{2+} signal. The activation of the PMCA pump by CaM has an additional point of interest: as the pump becomes activated, the concentration of Ca^{2+} in the microambient surrounding its cytosolic portion decreases, causing the detachment of CaM from its binding site in the cytosolic C-terminal tail of the pump. The detachment may be incomplete, as CaM may still remain partially bound to the pump with its C-terminal moiety, but the activity of the pump would necessarily decrease to a fraction of maximal. The lower activity would increase the concentration of Ca^{2+} in the pump microambient, promoting the full re-association of CaM with the pump, effectively reinitiating the activation/deactivation cycle. In essence, then, the activation of the PMCA pump by CaM cannot be permanent, but is bound to have oscillatory character. One last aspect of the function of the PMCA pump in the local regulation of the cytosolic Ca^{2+} is of interest: its isoforms have recently been found to regulate the Ca^{2+} transients produced by the triggering of the SOC channels (see above) in ways that differ depending on the peculiarities of their functional activity (Pászty *et al.*, 2015).

The plasma membrane of most animal cells contains a system that ejects Ca^{2+} in exchange for Na^+ (NCX). The

exchanger has lower affinity for Ca^{2+} than the Ca^{2+} pumps, and is particularly active in cells of excitable tissues like the heart and the nervous system. It has a much larger transport capacity than its companion plasma membrane Ca^{2+} transporter, the PMCA pump (Carafoli *et al.*, 2001), and thus efficiently restores the cytosolic Ca^{2+} concentration to resting values after the large increases physiologically occurring in cells like the neurons or the cardiomyocytes. It operates electrogenically, exchanging three Na^+ per one Ca^{2+} using the energy of the trans-plasma membrane electrochemical Na^+ gradient to eject Ca^{2+} ; however, the direction of the exchanger can be reversed by changes in the concentrations of the transported ions; if the concentration of Na^+ in the cytosol increases abnormally, as for example, in cardiomyocytes exposed to inhibitors of the plasma membrane Na^+/K^+ -ATPase like digitalis, the NCX will promote the influx of Ca^{2+} . The NCX is the product of a multigene family, the three basic isoforms having different tissue distribution and regulatory properties. All NCXs are organized in the membrane with nine transmembrane domains separated in two parts by a 500 amino acid cytosolic loop (Nicoll *et al.*, 2002). Partial 3D structures of the mammalian NCX have been published (Nicoll *et al.*, 2006; Hilge *et al.*, 2006), and the full structure of a prokaryotic exchanger has also been published (Liao *et al.*, 2012).

A phylogenetically ancestral branch of the $\text{Na}^+/\text{Ca}^{2+}$ exchanger superfamily also exchanges Li^+ , which is not accepted by the canonical NCXs. It has recently been identified (Palty *et al.*, 2010) as the variant of the $\text{Na}^+/\text{Ca}^{2+}$ exchanger that mediates the release of Ca^{2+} from the mitochondria (Carafoli *et al.*, 1974), which are now, after a long period of oblivion, have now become very important in the regulation of the cytosolic homeostasis of Ca^{2+} . Mitochondria had been shown to take up Ca^{2+} in the 1960s (De Luca *et al.*, 1961; Vasington and Murphy, 1962). However, their low Ca^{2+} affinity had negated their role in the regulation of cytosolic Ca^{2+} , relegating their Ca^{2+} transporting ability to a long period of obscurity (Carafoli, 2010). In the 1990s, it was discovered that the close proximity of endoplasmic reticulum to mitochondria generated local Ca^{2+} pools of sufficient Ca^{2+} concentration to activate their low affinity uptake system (Rizzuto *et al.*, 1992). Thus, mitochondria are thus now considered very important in the regulation of the cytosolic homeostasis of Ca^{2+} . Their uptake system is a channel uniporter (MCU) that operates in conjunction with one or more superficial components (MICU1) that may 'sense' Ca^{2+} to pass it out to the transmembrane MCU (De Stefani *et al.*, 2011; Baughman *et al.*, 2011; Figure 5).

CaM Targets

Calmodulin-Dependent Protein Kinases

Protein kinases regulate a great variety of cellular processes. More than 600 have been identified within the human genome. About 100 of them are calmodulin-dependent; they are involved in the regulation of numerous central cellular functions (see Swulius and Waxham, 2008 for a comprehensive review).

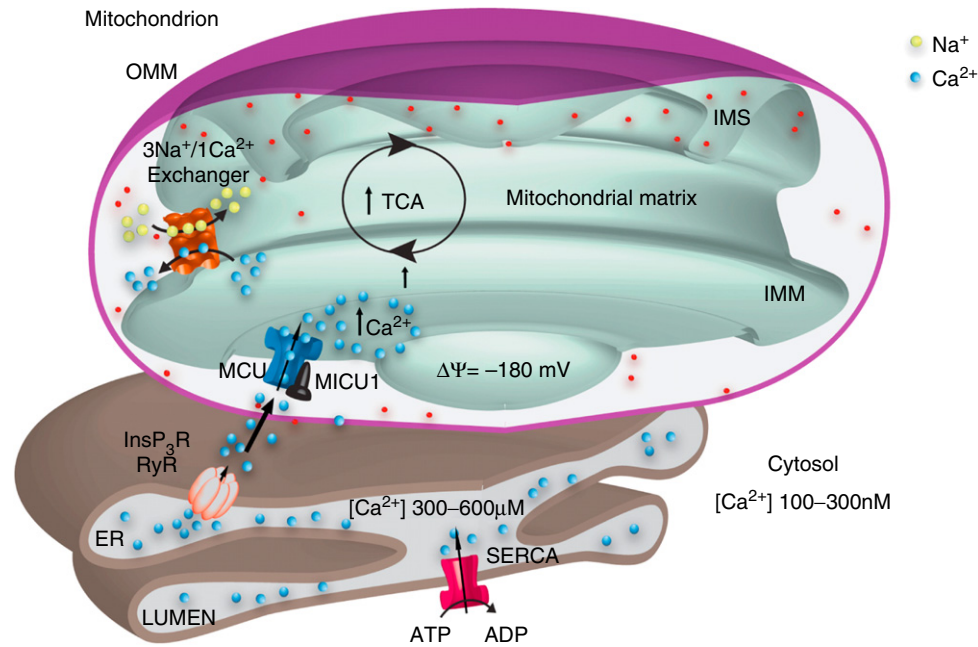


Figure 5 Schematic representation of a mitochondrion and its Ca^{2+} transporting systems. Details are found in the text. OMM is the outer mitochondrial membrane, IMM is the inner mitochondrial membrane, IMS is the inter membrane space. TCA is the citric acid cycle, which is also regulated by Ca^{2+} in the matrix. The cartoon also shows the endoplasmic reticulum closely associated with the mitochondrion and its two receptor/channels (InsP₃ R and RyR) that liberate Ca^{2+} to create micropools in which its concentration can temporarily reach the values well above micrometer necessary for the activation of the low affinity MCU uniporter. Other nonspecific channels (VDAC, mPTP) not described in the text also mediate the passage of Ca^{2+} under special conditions. Modified from Brini, M., Ottolini, D., Cali, T., 2013a. Intracellular calcium homeostasis and signaling. In: Banci, L. (Ed.), *Metallomics and the Cell, Metal Ions in Life Sciences*, vol. 12. Dordrecht: Springer Science + Business Media, pp. 119–168.

All members of the CaM kinase family are Ser/Thr kinases. Their general substrate consensus sequence is R-X-X-S/T (Kennelly and Krebs, 1991). According to their substrates they can be divided into two categories: monosubstrate and multisubstrate. MLCK, phosphorylase kinase, and CaM kinase III (elongation factor 2 kinase) are examples of monosubstrate kinases, whereas CaM kinase I, II, IV, and the CaM kinase-kinase (CaMKK) are multisubstrate kinases. The general structure of the CaM-kinases consists of a catalytic domain followed by a regulatory domain which contains both an auto-inhibitory and a CaM-binding domain.

The best understood CaM-dependent kinase at the molecular level is MLCK. It catalyzes the phosphorylation of a specific serine residue of the regulatory light chains of myosin II. Two MLCKs are encoded by different genes, and are specific for skeletal or smooth muscles (skMLCK and smMLCK). Both are important for the regulation of muscle contraction, however, the latter is also expressed in non-muscle cells.

Phosphorylase kinase is one of the largest and structurally most complex CaM-kinases. Its function is the phosphorylation of glycogen phosphorylase, the enzyme that modulates the energy source for muscle contraction. Phosphorylase kinase is composed of four subunits (α , β , γ , δ) that assemble in a heterotetramer of which four heterotetramers form the holoenzyme (molecular weight about 1.3 MDa). The γ subunit contains the catalytic core, whereas the α , β , and δ subunits are involved in the regulation of the enzyme. The δ subunit is CaM stably bound to the holoenzyme (Picton *et al.*, 1980).

A very special CaM kinase is CaMKIII, which phosphorylates elongation factor 2 as its only substrate (Nairn and Palfrey, 1987). Since its domain organization differs from that of CaMKI, II, or IV, CaMKIII has been renamed as elongation factor 2 kinase (EF2K). Phosphorylation of EF2 by EF2K leads to suppression of protein translation by the dissociation of EF2 from the ribosome.

CaMKI, II, and IV are multisubstrate kinases. CaMKI and IV are part of a phosphorylation cascade, since they are substrates of CaMKK β (Lee and Edelman, 1994; Tokumitsu *et al.*, 1994). They become fully active upon phosphorylation by CaMKK (CaMKKs are themselves substrates of other kinases, e.g., cyclin-dependent kinase 5, glycogen synthase kinase 3, and also DAPK). One of the best-studied substrates of CaMKI and IV is the transcription factor CREB, originally identified as a substrate of the cAMP-dependent protein kinase, but later shown to be an equally important substrate of CaMKI and IV in the control of Ca^{2+} -dependent gene expression (Matthews *et al.*, 1994).

CaMKI is ubiquitously expressed (the 3D structure of the autoinhibited form has been solved, Goldberg *et al.*, 1996). In contrast, the expression of CaMKIV is restricted to the nervous, immune, and fertilization systems (Wang *et al.*, 2001). Interestingly, recent observations have shown that CaMKIV is also involved in the regulation of alternative splicing (Xie and Black, 2001) of one of the isoforms of the PMCA pump, PMCA1a (Krebs, 2014), which is preferentially expressed during brain development.

CaMKII is very important for the regulation of synaptic plasticity and spatial memory (Giese *et al.*, 1998). In mammals four isoforms are encoded by different genes. In contrast to CamKI and IV, which are monomeric, the holoenzyme of CaMKII is a 12 subunit homo- or heteromeric complex. Autophosphorylation is important to CaMKII's activation, which is initiated by the Ca^{2+} -promoted binding of CaM, which releases the autoinhibited state (Hudmon and Schulman, 2002). Trans-phosphorylation between two adjacent Ca^{2+} /CaM-bound subunits generates a Ca^{2+} -independent activity, which enables the enzyme to decode the Ca^{2+} signals (Hanson *et al.*, 1994) necessary for synaptic plasticity, learning, and memory (Giese *et al.*, 1998). CaMKII's ability to detect changes in the frequency of Ca^{2+} oscillations has led to the suggestion that CaMKII may act as a trigger for synaptic plasticity (De Koninck and Schulman, 1998). Structural studies have added information on the regulation of neuronal (and cardiac) Ca^{2+} signaling by CaMKII (Stratton *et al.*, 2013).

Kimchi and coworkers have described a novel CaM Ser/Thr kinase (Deiss *et al.*, 1995), which is involved in the cell-death signaling network (DAPK). It was later demonstrated that DAPK belongs to a larger family of kinases, which are linked to apoptotic pathways. DAPK contains a conserved CaM autoregulatory domain with an inhibitory autophosphorylation site (Temmermann *et al.*, 2013). All members of this family of kinases can function as tumor suppressors, and possess regulatory domains important for function and localization within the cell (Shiloh *et al.*, 2014).

Calmodulin-Dependent Phosphatase

CaN is a 2B phosphatase regulated by CaM (Klee and Krinks, 1978), which is highly conserved from yeast to mammals. In mammals CaN is mainly found in the nervous tissues, hence its name (Klee *et al.*, 1979). The enzyme is composed of a catalytic domain CaN A that binds CaM and a regulatory domain CaN B, which itself is a CaM-like Ca^{2+} -binding subunit containing four EF-hands. The enzyme is rather substrate-selective, a major substrate being the nuclear factor of activated T-cells (NFAT). This highly phosphorylated transcription factor is dephosphorylated by CaN, which promotes its translocation to the nucleus (Jain *et al.*, 1993). An important finding was the demonstration that the immunosuppressants cyclosporin A and FK506 bind CaN with high affinity, inhibiting its phosphatase activity in T lymphocytes (Fruman *et al.*, 1992). The solution of the crystal structure of CaN bound to FK506 associated to its immunophylin FKBP12 has underlined its importance for the immune response (Griffith *et al.*, 1995; Kissinger *et al.*, 1995). In addition to Ca^{2+} /CaM, CaN is also regulated by an endogenous modulator, the regulator of calcineurin RCAN, a highly conserved protein whose gene is located in the critical Down Syndrome region of chromosome 21 (Fuentes *et al.*, 2000).

The Dark Side of Ca^{2+} Signaling: Ca^{2+} -Related Pathologies

One distinctive property of the Ca^{2+} signal is ambivalence. As should have become evident from the discussions above,

the Ca^{2+} message must be spatially and temporally controlled with great precision to permit the correct functioning of the most important processes of cell life. Unfortunately, the control may sometimes become defective, effectively transforming Ca^{2+} into a negative messenger: the uncontrolled activation of potentially harmful Ca^{2+} controlled activities, from proteases, to phospholipases, to nucleases, may create situations of cell discomfort that can in extreme cases culminate in cell death. In principle, death is not necessarily a negative phenomenon, as it is for instance essential for processes as organ remodeling and cell renewal, and is thus but one of the processes in which the Ca^{2+} signal operates under strict control. Conditions in which the control of the Ca^{2+} signal becomes grossly defective instead lead to the massive and global increase of cell Ca^{2+} and rapidly terminate cell life. These are the relatively uninteresting cases of toxic cell death. However, the dysregulation of the Ca^{2+} signal could also selectively affect individual components of the systems for its control; cell life would in these cases be permitted to continue, although with variable degrees of difficulty. These conditions are frequently genetic, and their study has now progressed to impressive molecular detail. A full discussion of Ca^{2+} related pathologies would be outside of the scope of this article, as the area has rapidly grown to be very large, and is extensively covered in a number of comprehensive contributions (see the additional readings suggested). These pathologies could affect practically every one of the systems that control Ca^{2+} by mediating its transport across membranes or by binding it for buffering purposes and/or to decode its message. Their study has offered interesting contributions to the understanding of the mechanisms for the control of cell Ca^{2+} and of the molecular processing of its message.

See also: Molecular Principles, Components, Technology, and Concepts: Membranes: Membrane Potential: Concepts

References

- Babu, Y.S., Sack, J.S., Greenhough, T.J., *et al.*, 1985. Three-dimensional structure of calmodulin. *Nature* 315, 37–40.
- Baughman, J.M., Perocchi, F., Girgis, H.S., *et al.*, 2011. Integrative genomics identifies MCU as an essential component of the mitochondrial calcium uniporter. *Nature* 476, 341–345.
- Brini, M., Carafoli, E., 2009. Calcium pumps in health and disease. *Physiological Reviews* 89, 1341–1378.
- Brini, M., Ottoloni, D., Cali, T., 2013a. Intracellular calcium homeostasis and signaling. In: Banci, L. (Ed.), *Metallomics and the Cell, Metal Ions in Life Sciences* 12. Dordrecht: Springer Science + Business Media, pp. 119–168.
- Brini, M., Ottoloni, D., Cali, T., Carafoli, E., 2013b. Calcium in health and disease. In: Sigel, A., Sigel, H., Sigel, R.K.O. (Eds.), *Interrelations between Metal Ions and Human Disease, Metal Ions in Life Sciences* 13. Dordrecht: Springer Science + Business Media, pp. 81–137.
- Carafoli, E., 2010. The fateful encounter of mitochondria with calcium: How did it happen? *Biochimica et Biophysica Acta* 1797, 595–606.
- Carafoli, E., Penniston, J.T., 1985. The calcium signal. *Scientific American* 253, 70–78.
- Carafoli, E., Santella, L., Branca, D., Brini, M., 2001. Generation, processing, and control of cellular calcium signals. *Critical Reviews in Biochemistry and Molecular Biology* 36, 107–260.
- Carafoli, E., Tiozzo, R., Lugli, G., Crovetto, F., Kratzing, C., 1974. The release of calcium from heart mitochondria by sodium. *Journal of Molecular and Cellular Cardiology* 6, 361–371.

- Copley, R.R., Schultz, J., Ponting, C.P., Bork, P., 1999. Protein families in multicellular organisms. *Current Opinion in Structural Biology* 9, 408–415.
- De Diego, I., Kuper, J., Bakalova, N., Kursula, P., Wilmanns, M., 2010. Molecular basis of the death-associated protein kinase-calcium/calmodulin regulator complex. *Science Signaling* 3, ra6. doi:10.1126/scisignal.2000552.
- De Koninck, P., Schulman, H., 1998. Sensitivity of CaM kinase II to the frequency of Ca^{2+} oscillations. *Science* 279, 227–230.
- De Luca, H.F., Engstrom, G.W., Rasmussen, H., 1961. The action of vitamin D and parathyroid hormone *in vitro* on calcium uptake and release by kidney mitochondria. *Proceedings of the National Academy of Sciences of the United States of America* 48, 1604–1609.
- De Stefani, D., Raffaello, A., Teardo, E., Szabo, I., Rizzuto, R., 2011. A forty kilodalton protein of the inner membrane is the mitochondrial calcium uniporter. *Nature* 476, 336–340.
- Deiss, L.P., Feinstein, E., Berissi, H., Cohen, O., Kimchi, A., 1995. Identification of a novel serine/threonine kinase and a novel 15-kD protein as potential mediators of the gamma interferon-induced cell death. *Genes & Development* 9, 15–30.
- Elshorst, B., Hennig, M., Försterling, H., *et al.*, 1999. NMR solution structure of a complex of calmodulin with a binding peptide of the Ca^{2+} pump. *Biochemistry* 38, 12320–12332.
- Fallon, J.L., Hailing, D.B., Hamilton, S.L., Quijcho, F.A., 2005. Structure of calmodulin bound to the hydrophobic IQ domain of the cardiac Ca(v)1.2 calcium channel. *Structure* 13, 1881–1886.
- Fischer, R., Koller, M., Flura, M., *et al.*, 1988. Multiple divergent mRNAs code for a single human calmodulin. *Journal of Biological Chemistry* 263, 17055–17062.
- Fruman, D.A., Klee, C.B., Bieder, B.E., Burakoff, S.J., 1992. Calcineurin phosphatase activity in T lymphocytes is inhibited by FK506 and cyclosporin A. *Proceedings of the National Academy of Sciences of the United States of America* 89, 3686–3690.
- Fuentes, J.J., Genescà, L., Kingsbury, T.J., *et al.*, 2000. DSCR1, overexpressed in Down syndrome, is an inhibitor of calcineurin-mediated signaling pathways. *Human Molecular Genetics* 9, 1681–1690.
- Gerke, V., Creutz, C.E., Moss, S.E., 2005. Annexins: linking Ca^{2+} signalling to membrane dynamics. *Nature Reviews Molecular Cell Biology* 6, 449–461.
- Giese, K.P., Fedorov, N.B., Filipkowski, R.K., Silva, A.J., 1998. Autophosphorylation at Thr286 of the alpha calcium-calmodulin kinase II in LTP and learning. *Science* 279, 870–873.
- Goldberg, J., Nairn, A.C., Kuriyan, J., 1996. Structural basis for the autoinhibition of calcium/calmodulin-dependent protein kinase I. *Cell* 84, 875–887.
- Griffith, J.P., Griffith, J.P., Kim, J.L., *et al.*, 1995. X-ray structure of calcineurin inhibited by the immunophilin-immunosuppressant FKBP12-FK506 complex. *Cell* 82, 507–522.
- Hanson, P.J., Meyer, T., Stryer, L., Schulman, H., 1994. Dual role of calmodulin in autophosphorylation of multi-functional CaM kinase may underlie decoding of calcium signals. *Neuron* 12, 943–956.
- Hilge, M., Aelen, J., Vuister, G.W., 2006. Ca^{2+} regulation in the $\text{Na}^{+}/\text{Ca}^{2+}$ exchanger involves two markedly different Ca^{2+} sensors. *Molecular Cell* 22, 15–25.
- Hou, X., Pedi, L., Diver, M.M., Long, S.B., 2012. Crystal structure of the calcium release-activated calcium channel Orai. *Science* 338, 1308–1313.
- Hudmon, A., Schulman, H., 2002. Structure-function of the multifunctional Ca^{2+} /calmodulin-dependent protein kinase II. *Biochemical Journal* 364, 593–611.
- Ikura, M., Clore, G.M., Gronenborn, A.M., *et al.*, 1992. Solution structure of a calmodulin-target peptide complex by multidimensional NMR. *Science* 256, 632–638.
- Jain, J., McCaffrey, P.G., Miner, Z., *et al.*, 1993. The T-cell transcription factor NFATp is a substrate for calcineurin and interacts with Fos and Jun. *Nature* 365, 352–355.
- James, P., Vorherr, T., Carafoli, E., 1995. Calmodulin-binding domains: Just two faced or multi-faceted? *Trends in Biochemical Sciences* 20, 38–42.
- Juranic, N., Atanasova, E., Filoteo, A.G., *et al.*, 2010. Calmodulin wraps around its binding domain in the plasma membrane Ca^{2+} pump anchored by a novel 18-1 motif. *Journal of Biological Chemistry* 285, 4015–4024.
- Kawasaki, H., Nakayama, S., Kretsinger, R.H., 1998. Classification and evolution of EF-hand proteins. *Biomaterials* 11, 277–295.
- Kennelly, P.J., Krebs, E.G., 1991. Consensus sequences as substrate specificity determinants for protein kinases and protein phosphatases. *Journal of Biological Chemistry* 266, 15555–15558.
- Kissinger, C.R., Parge, H.E., Knighton, D.R., *et al.*, 1995. Crystal structures of human calcineurin and the human FKBP12-FK506-calcineurin complex. *Nature* 378, 641–644.
- Klee, C.B., Crouch, T.H., Krinks, M.H., 1979. Calcineurin: A calcium- and calmodulin-binding protein of the nervous system. *Proceedings of the National Academy of Sciences of the United States of America* 76, 6270–6273.
- Klee, C.B., Krinks, M.H., 1978. Purification of cyclic 3',5'-nucleotide phosphodiesterase inhibitory protein by affinity chromatography on activator protein coupled to Sepharose. *Biochemistry* 17, 120–126.
- Klee, C.B., Vanaman, T.C., 1982. Calmodulin. *Advances in Protein Chemistry* 35, 213–321.
- Krebs, J., 2014. The plethora of PMCA isoforms: Alternative splicing and differential expression. *Biochimica et Biophysica Acta*. Available at: <http://dx.doi.org/10.1016/j.bbamcr.2014.12.020> (accessed 27.03.15).
- Krebs, J., Buerkler, J., Guerini, D., Brunner, J., Carafoli, E., 1984. 3-(Trifluoromethyl)-3-(m-[125]iodophenyl)diazirine, a hydrophobic, photoreactive probe, labels calmodulin and calmodulin fragments in a Ca^{2+} -dependent way. *Biochemistry* 23, 400–403.
- Kretsinger, R.H., Nockolds, C.E., 1973. Carp-muscle calcium-binding protein. II. Structure determination and general description. *Journal of Biological Chemistry* 248, 3313–3326.
- Krueger, J.K., Olah, G.A., Rokop, S.E., *et al.*, 1997. Structures of calmodulin and a functional scattering study. *Biochemistry* 36, 6017–6023.
- Kuboniva, H., Tjandra, N., Grzesiek, Z., *et al.*, 1995. Solution structure of calcium-free calmodulin. *Nature Structural & Molecular Biology* 2, 768–776.
- LaPorte, D.C., Wierman, B.M., Storm, D.R., 1980. Calcium-induced exposure of a hydrophobic surface on calmodulin. *Biochemistry* 19, 3814–3819.
- Lee, J.C., Edelman, A.M., 1994. A protein activator of Ca^{2+} -calmodulin-dependent protein kinase Ia. *Journal of Biological Chemistry* 269, 2158–2164.
- Liao, J., Li, H., Zeng, W., *et al.*, 2012. Structural insight into the ion-exchange mechanism of the sodium/calcium exchanger. *Science* 335, 686–690.
- Matthews, R.P., Guthrie, C.R., Wailes, L.M., *et al.*, 1994. Calcium/calmodulin-dependent protein kinase types II and IV differentially regulate CREB-dependent gene expression. *Molecular and Cellular Biology* 14, 6107–6116.
- Meador, W.E., Means, A.R., Quijcho, F.A., 1992. Target enzyme recognition by calmodulin: 2.4 Å structure of a calmodulin-peptide complex. *Science* 257, 1251–1255.
- Mohamed, T.M., Oeandy, D., Zi, M., *et al.*, 2011. Plasma membrane calcium pump (PMCA4)-neuronal nitric oxide synthase complex regulates cardiac contractility through modulation of a compartmentalized cyclic nucleotide domain. *Journal of Biological Chemistry* 286, 41520–41529.
- Nairn, A.C., Palfrey, H.C., 1987. Identification of the major Mr 100 000 substrate for calmodulin-dependent protein kinase III in mammalian cells as elongation factor-2. *Journal of Biological Chemistry* 262, 17299–17303.
- Nicoll, D.A., Ottolia, M., Philipson, K.D., 2002. Toward a topological model of NCX1exchanger. *Annals of the New York Academy of Sciences* 976, 11–18.
- Nicoll, D.A., Sawaya, M.R., Kwon, S., *et al.*, 2006. The crystal structure of the primary Ca^{2+} sensor of the $\text{Na}^{+}/\text{Ca}^{2+}$ exchanger reveals a novel Ca^{2+} binding motif. *Journal of Biological Chemistry* 281, 21577–21581.
- Palty, R., Silverman, W.F., Hershfinkel, M., *et al.*, 2010. NCLX is an essential component of mitochondrial $\text{Na}^{+}/\text{Ca}^{2+}$ exchange. *Proceedings of the National Academy of Sciences of the United States of America* 107, 436–441.
- Pászty, K., Caride, A.J., Bajzer, Z., *et al.*, 2015. Plasma membrane Ca^{2+} ATPases can shape the pattern of Ca^{2+} transients induced by store-operated Ca^{2+} entry. *Science Signaling* 8, 1–11.
- Picton, C., Klee, C.B., Cohen, P., 1980. Phosphorylase kinase from rabbit skeletal muscle: Identification of the calmodulin-binding subunits. *European Journal of Biochemistry* 111, 553–561.
- Pruschy, M., Ju, Y., Spitz, L., Carafoli, E., Goldfarb, D.S., 1994. Facilitated nuclear transport of calmodulin in tissue culture cells. *Journal of Cell Biology* 127, 1527–1536.
- Rellos, P., Pike, A.C., Niesen, F.H., *et al.*, 2010. Structure of the CaMKII δ /calmodulin complex reveals the molecular mechanism of CaMKII kinase activation. *PLOS Biology* 8, e1000426.
- Rizzuto, R., Simpson, A.W., Brini, M., Pozzan, T., 1992. Rapid changes of mitochondrial Ca^{2+} revealed by specifically targeted recombinant aequorin. *Nature* 358, 325–327.
- Roberts, D.M., Harmon, A.C., 1992. Calcium-modulated proteins: Targets of intracellular calcium signals in higher plants. *Annual Review of Plant Physiology and Plant Molecular Biology* 43, 375–414.
- Shenolikar, S., Cohen, P.T., Cohen, P., Nairn, A.C., Perry, S.V., 1979. The role of calmodulin in the structure and regulation of phosphorylase kinase from rabbit skeletal muscle. *European Journal of Biochemistry* 100, 329–337.

- Shiloh, R., Bialik, S., Kimchi, A., 2014. The DAPK family: A structure-function analysis. *Apoptosis* 19, 286–297.
- Stratton, M.M., Chao, L.H., Schulman, H., Kuriyan, J., 2013. Structural studies on the regulation of Ca^{2+} /calmodulin dependent protein kinase II. *Current Opinion in Structural Biology* 23, 292–301.
- Swulius, M.T., Waxham, M.N., 2008. Ca^{2+} /calmodulin-dependent protein kinases. *Cellular and Molecular Life Sciences* 65, 2637–2657.
- Temmermann, K., Simon, B., Wilmanns, M., 2013. Structural and functional diversity in the activity and regulation of DAPK-related protein kinases. *FEBS Journal* 280, 5533–5550.
- Tokumitsu, H., Brickey, D.A., Glod, J., *et al.*, 1994. Activation mechanisms for Ca^{2+} /calmodulin-dependent protein kinase IV. Identification of a brain CaM-kinase IV kinase. *Journal of Biological Chemistry* 269, 28640–28647.
- Toyoshima, C., Nakasako, M., Nomura, H., Ogawa, H., 2000. Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. *Nature* 405, 647–655.
- Vasington, F.D., Murphy, J.V., 1962. Ca ion uptake by kidney mitochondria and its dependence on respiration and phosphorylation. *Journal of Biological Chemistry* 237, 2670–2677.
- Villarroel, A., Tagliatalata, M., Bernardo-Seisdedos, G., *et al.*, 2014. The ever changing moods of calmodulin: How structural plasticity entails transductional adaptability. *Journal of Molecular Biology* 426, 2717–2735.
- Wang, S.L., Ribar, T.J., Means, A.R., 2001. Expression of Ca^{2+} /calmodulin-dependent protein kinase IV (caMKIV) messenger RNA during murine embryogenesis. *Cell Growth & Differentiation* 12, 351–361.
- Woodhead, J.S., 1982. Regulation of whole body calcium in man. In: Ekeles, R.S. (Ed.), *Biochemical Aspects of Human Disease*. Oxford: Blackwell.
- Xie, J., Black, D.L., 2001. A CaMKIV responsive RNA element mediates depolarization-induced alternative splicing of ion channels. *Nature* 410, 936–939.
- Yap, K.L., Ames, J.B., Swindells, M.B., Ikura, M., 1999. Diversity of conformational states and changes within the EF-Hand protein superfamily. *Proteins: Structure, Function and Genetics* 37, 499–507.
- Zalk, R., Clarke, O.B., des Georges, A., *et al.*, 2015. Structure of a mammalian ryanodine receptor. *Nature* 517, 44–49.
- Zhang, M., Tanaka, T., Ikura, M., 1995. Calcium-induced conformational transition revealed by the solution structure of apo calmodulin. *Nature Structural & Molecular Biology* 2, 758–767.

Further Reading

- Carafoli, E., Brini, M. (Eds.), 2007. *Calcium Signalling and Disease: Molecular Pathology of Calcium*. New York, NY: Springer.
- Carafoli, E., Klee, C.B. (Eds.), 1999. *Calcium as a Cellular Regulator*. New York, NY: Oxford University Press.
- Krebs, J., Michalak, M. (Eds.), 2007. *Calcium: A Matter of Life or Death*. Amsterdam: Elsevier.

CELL COMMUNICATION: CELLULAR MACHINERIES

Contents

Macromolecular Communication between Nucleus and Cytoplasm

The Molecular Architecture of Cell–Cell Adhesions

Extracellular Regulation of Cell-to-Matrix Adhesion

Macromolecular Communication between Nucleus and Cytoplasm

AM Tartakoff, Case Western Reserve University School of Medicine, Cleveland, OH, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Exportins and Importins (also known as karyopherins) They are the principal cargo receptors that recognize transport motifs in one compartment (nucleus or cytoplasm), interact with the nuclear pores, and then translocate across the pore. Their ability to bind cargo is sensitive to the GTP/GDP status of the Ran GTPase.

Nuclear envelope The double membrane that surrounds the nucleoplasm. The outer membrane is compositionally and functionally very similar to the endoplasmic reticulum.

Nuclear pores These are the eightfold symmetric mega-structures that are embedded in the nuclear envelope. All rapid transport into and out of the nucleus occurs through these pores.

Nucleocytoplasmic transport The process by which macromolecules translocate across the nuclear envelope.

Ran cycle The Ran GTPase belongs to the Ras superfamily. Its GTP/GDP state is critical for nuclear import and export.

Translocon The site in a membrane through which macromolecular transport occurs. Translocons of the nuclear envelope are nuclear pores.

Introduction

Because membranes form barriers between the interior and exterior of cells and between compartments within cells, the transport of macromolecules across membranes has been a classic object of study. Their elucidation has drawn from the fine application of genetic, biochemical, and cytological approaches. The understanding of corresponding mechanisms and their relation to transmembrane asymmetry lie at the heart of explaining the integrity and origin of highly differentiated multi-organelle eukaryotic cells.

History – Overview

Since transcription occurs in the nucleus, while most RNA resides in the cytoplasm where protein synthesis occurs, it was evident for many years that mechanisms must exist to export RNA and to import proteins. The first glimmer of understanding depended on the development of the electron microscope and satisfactory procedures for sample preparation, which made it possible to visualize nuclear pores (Hoelz *et al.*, 2011; Floch *et al.*, 2014; Figure 1). In fact, light microscopy can often predict the location of pores since

heterochromatin is absent from the regions of the nuclear perimeter where pores are inserted. Table 1 provides a chronology of relevant landmark observations that preceded investigations of nucleocytoplasmic transport per se.

Following upon the functional studies that defined some of the basic features of import, pore proteins were identified by purifying nuclei, recovering an insoluble residue after treatment with nucleases, detergent, concentrated salt and chelators, and using the residue to raise corresponding antibodies. When recovered from animal cells, the residue generally included the nuclear lamina as well as the pores themselves. The structure came to be known as the ‘nuclear pore complex’ (NPC) (Davis and Blobel, 1986; Dingwall and Laskey, 1986).

The distribution of ~100 (yeast) to thousands (animal cells) of pores along the surface of the nuclear envelope (NE) appears largely random; however, each pore often seems to define a territory that is free from other pores. Moreover, at least in *Saccharomyces cerevisiae*, NPCs are absent from the point at which the nucleus and vacuole contact each other (Pan *et al.*, 2000) and from multilayered regions of the NE (Wright *et al.*, 1988). Especially high-packing densities are characteristic of the NE of amphibian oocytes and portions of the NE of mature spermatids (Goldberg *et al.*, 1997;

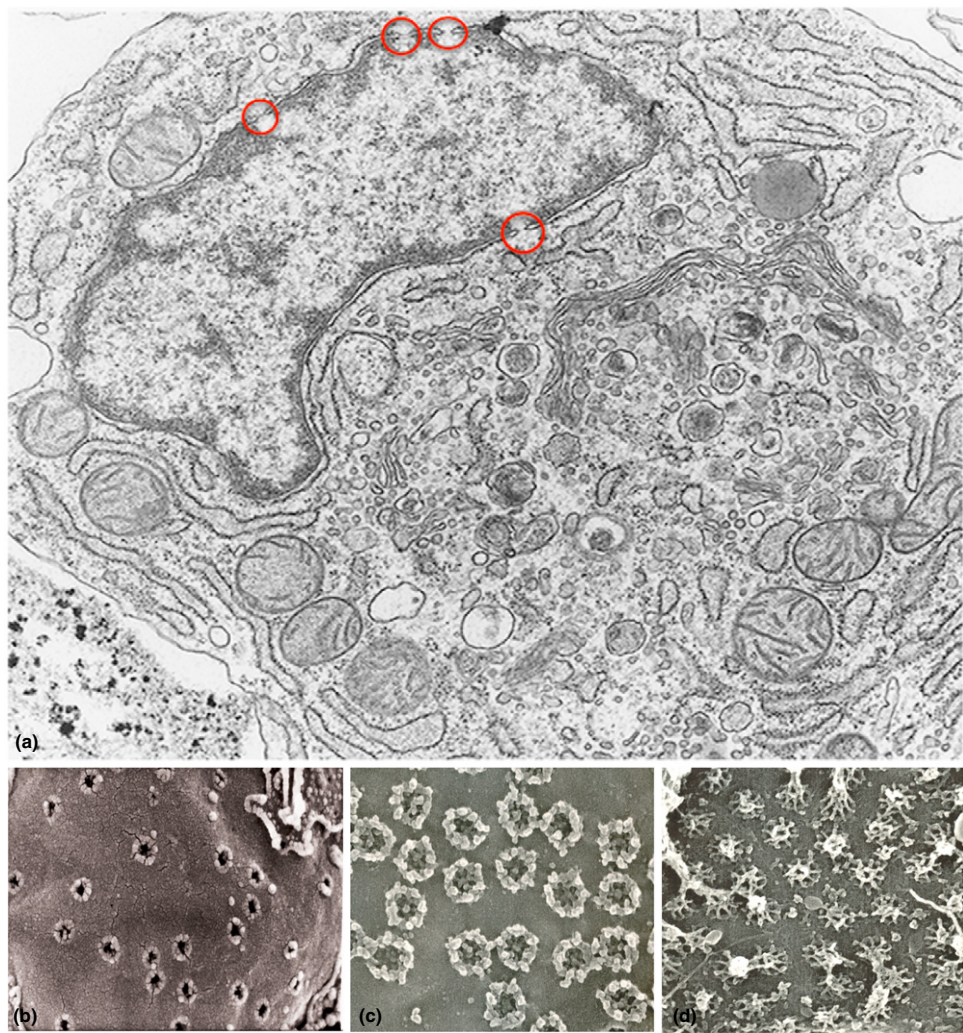


Figure 1 Electron micrographs of animal cell nuclei. (a) A transmission EM of a thin section of a myeloma cell. Pores are circled in red. (b) A scanning EM image of the surface of the NE of a serous cell from the submandibular gland, prepared by osmium maceration (Riva *et al.*, 1998). (c) and (d) Scanning EM images of the cytoplasmic and nucleoplasmic faces of an isolated nuclear envelope from the oocyte of *Xenopus*. Each pore complex measures ~100 nm in diameter. Images courtesy of Dr. A. Riva and Dr. B. Tandler (a); Dr. T. Allen (c–d).

Table 1 Setting the stage for discovery of nuclear pores

Nucleated fragments of cells can divide	Waller, Brandt, Nussbaum, and Gruber	1852–85
DNA is in the nucleus and RNA in the cytoplasm	Brachet and Caspersson	1933
First commercial transmission and scanning electron microscopes	Siemens, IG Farben, and Cambridge Instruments	1939–65
The nucleus programs cell phenotype	Hämmerling	1943
The transforming substance is DNA	Avery, MacLeod, and McCarty	1944
Phage transfer DNA into bacteria	Hershey and Chase	1952
Discovery of mRNA in bacteria	Brenner, Jacob, Meselson, and Monod	Late 1950s
EM visualization of pores in whole mounts and in thin sections	Callan and Tomlim; Scheer and Franke; Gall, Watson, Afzelius, and Rhodin	1950–66
Demonstration of nuclear synthesis of RNA, RNA export, and protein import and export	Prescott and Goldstein and Plaut	1955–67
Visualization of pores by freeze fracture	Friend and Fawcett	1974
Visualization of the nuclear basket by scanning EM	Ris	1989

Ho, 2010). In higher eukaryotes, pores are linked to the nuclear lamina and are not mobile in the plane of the membrane. By contrast, lateral translocation is evident in *S. cerevisiae* (Bucci and Wente, 1997; Belgareh and Doye, 1997) in which the inner face of the NE lacks lamin homologs (Hattier *et al.*, 2007; Gruenbaum *et al.*, 2005; Worman, 2012).

Abundant bidirectional translocation occurs through NPCs. Electron microscopic observations of colloidal gold tracers decorated with classical import or export cargos show that NPCs can transport cargo and can simultaneously participate in both import and export (Dingwall and Laskey, 1986; Feldherr and Dworetzky, 1988). The larger issue of the valency of translocation, i.e., whether more than two cargos can simultaneously pass in one or both directions, has not been addressed.

NPCs allow facile exchange of proteins <20–40 kDa, while larger protein cargos up to ~1000 kDa can enter if they carry nuclear localization signals (NLSs) that are accessible on their surface (Shulga *et al.*, 2000; Lanford *et al.*, 1986). By contrast to the cargos that access the endoplasmic reticulum (ER) or mitochondria, there is no need for the signals to be in the linear sequence of the cargo. Viruses that dock at the outer surface of NPCs as they import their genetic material may exceed the upper size limit. The entire NPC structure appears to be ‘belted’: No circumstances are known to change its external diameter. The striking observation that the NPC can accommodate huge cargos without being grossly leaky for monomeric proteins suggests that cargo in transit is in some way sequestered or intimately enclosed in a malleable milieu (see below).

Since NLSs and nuclear export signals (NESs) are not removed upon translocation, individual cargos can be translocated repeatedly, as for transcription factors that shuttle across the NE (Kopito and Elbaum, 2007). For cells that undergo NE breakdown during mitosis, the permanence of signals enormously facilitates the reestablishment of the nucleoplasmic milieu in telophase.

In addition to the translocation of soluble cargos, as discussed below, the NPC mediates transport of membrane proteins. Thus, proteins delivered to the ER membrane access the inner (and outer) membrane of the NE, and proteins of the inner membrane can access the ER (Powell and Burke, 1990). These events may result from traversal of a lateral channel at the margin of the NPC. Import of inner membrane proteins is facilitated by NLSs that are also functional for soluble cargos (Lusk *et al.*, 2007; Schirmer *et al.*, 2003).

Transport Signals and the Importin β Superfamily: A Code

Many NLS motifs are enriched in positively charged amino acids. These ‘classical’ signals can be monopartite, as in SV40 large T antigen, or bipartite, as in nucleoplasmin. Their identification – and the demonstration of ongoing import of proteins – largely depended on studies in which candidate cargos were delivered into the cytoplasm of *Xenopus* oocytes and then were followed to evaluate their import into the nucleus (Dingwall and Laskey, 1986; Goldfarb *et al.*, 1986; Stevens and Swift, 1966; Speer and Zimmerman, 1968; Bonner, 1975; Kalderon *et al.*, 1984). Initially it was unclear

whether corresponding receptors were in fact permanent components of the NPCs or rather might be soluble in the cytoplasm. It was also unclear as to whether filamentous docking sites might extend throughout the cytoplasm (and nucleoplasm) to facilitate cargo delivery to the NPC.

After attempts to use NLSs as probes for binding factors and for affinity retrieval, proper identification of receptors required development of a functional assay for import. This depended on the all-important realization that treatment of adherent cells in culture with the sterol-selective detergent, digitonin, permeabilized the plasma membrane, releasing soluble proteins, and exposing the surface of the nucleus. Since import of labeled cargos could be beautifully reconstituted by addition of cytosolic proteins, this assay was used to identify the essential equipment for import: importin β /karyopherins and the GTPase, Ran (Moore and Blobel, 1992; Fornerod *et al.*, 1997b; Gorlich *et al.*, 1997; Adam *et al.*, 1990).

Importin β superfamily members (KAPs, IPOs, and XPOs – 95–145 kDa) have a ‘Ran-GTP binding domain’ at their N-terminus, that is followed by many short α -helical ‘HEAT repeat’ motifs that bind both translocation cargos and the NPC (Strom and Weis, 2001; Kimura and Imamoto, 2014). Crystallographic studies have defined cargo-binding sites for several importins and exportins in great detail (Suel *et al.*, 2006; Cook *et al.*, 2007).

In the widely accepted model for import, cargos with classical NLSs bind complementary sites in adaptor proteins (importin α ’s – that are not members of the importin β family; Goldfarb *et al.*, 2004) that in turn bind importin β 1 (KPNB1), thus making it possible to interact with and traverse the NPC. Proteins with nonclassical NLSs bind directly to and are translocated by other members of the importin β family (Table 2; Figure 2). The recognition code for transport is in many cases degenerate: multiple cargos bind the same importin and some cargos bind to more than one importin. In fact, the varieties of signals that are recognized by many importins have not yet been exhaustively characterized and it remains in many cases quite unclear why individual cargos use one or another signal/importin pair. It is not evident, for example, that groups of proteins that are functionally related are generally imported by a single importin.

Several approaches have been used to identify the cargo that is transported by a given importin β . If an importin can be inactivated (using ts alleles or promoter shut off) one can ask whether the nucleocytoplasmic distribution of a given potential cargo will change upon inactivation. Such evidence is indirect, but strongly suggestive of function. More direct approaches have involved expressing a given importin β in living cells, recovering it from lysates, and then inquiring whether any putative cargo can be released from it upon denaturation. Since the GTPase, Ran, cycles between GTP- and GDP-bound forms (see below), this test becomes definitive if the release of cargo is achieved by the addition of Ran-GTP or a Ran mutant that is locked in the GTP-bound conformation. For exportins, a reciprocal strategy can involve binding in the presence of Ran-GTP and then inquiring whether a cargo is released upon addition of Ran-GAP or replacement of Ran-GTP with Ran-GDP.

Many export cargos are correspondingly translocated by ‘exportins,’ that belong to the very same importin β

Table 2 Importin β family members and typical cargos

	Importin/exportin		Known cargo	
	Homo sapiens	Saccharomyces cerevisiae	H. sapiens	S. cerevisiae
Import	<i>Homologs are readily identified^a</i>			
	KPNB1/ β 1	KAP95	Importin α and associated proteins with classical monopartite NLSs, snurportin	
	IPO9	KAP114	Actin, histones, ribosomal proteins	H2A, H2B, Nap1, Spt15, Sua7
	IPO11	KAP120	L12, UbcM2	Rfp1
	IPO4	KAP123	RPS3A, UL84	H3, H4
Export	XPO5	MSN5/KAP142 (bidirectional)	dsRNA export, microRNAs	Import of Rfa1; export of Far1, Msn2, Pho4, Swi5, Swi6, large subunits, tRNAs
	CSE1L	CSE1/KAP109	Importin α	
	XPO1	LOS1	tRNAs	
Import	<i>Diversification^a</i>			
	RANBP5/RANBP6	PSE1/KAP121	Histones, ribosomal proteins, influenza RNA polymerase	Aft1, Pdr1, Ste12, Yap1
	IPO7/IPO8	NMD5/KAP119	Ago, Smad, ERK, GR, ribosomal proteins	Crz1, Hog1, TfiIS
	IPO3/TNPO3	MTR10/KAP111	SR proteins, HuR	Npl3, tRNAs
	TNPO1/TNPO2	KAP104	hnRNP A1 (M9, PY motifs)	Hrp1, Nab2
Import	<i>Further diversification^a</i>			
		PDR6/KAP122		Toa1, Toa2
		SXM1/KAP108		SRP19, Smad
	XPO13 (bidirectional)		Lhp1, Pab1	UBC9, RBM8 (Y14) for import; eIF1A for export
	XPO4 (bidirectional)		Arx	
	XPO7		Imports SOX2, SRY, exports Smad3, eIF5A	
	XPO6		p50-RhoGAP	
	XPO7			

^aGroupings are according to whether the structure of each importin is widely conserved (Quan *et al.*, 2008).

Note: Cargos remain to be identified for many family members.

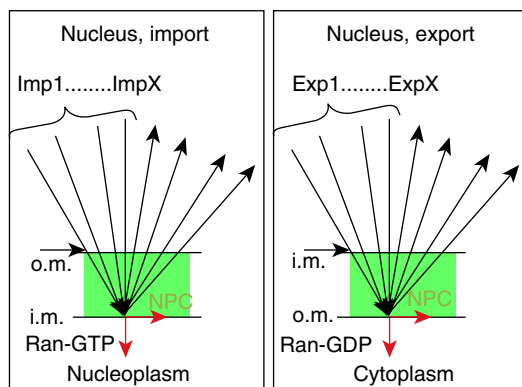


Figure 2 Overview of nuclear import and export. Membrane surfaces are designated by horizontal lines. The paired sets of black arrows designate the arrival of multiple importins (or exportins) which is followed by their return to their compartment of origin. The red arrows indicate the location of cargo after release from the translocon. o.m., outer membrane; i.m., inner membrane.

superfamily (Guttler and Gorlich, 2011). This elegant structural convergence between importins and exportins is remarkable. 'Leucine-rich' NESs bind the exportin, Crm1/XPO1. This cargo motif-exportin pair was identified as a result of interest in localization of incompletely spliced HIV mRNAs and a subunit of cAMP-dependent protein kinase, as well as identification of the target of the anti-HIV drug, leptomycin B (Fischer *et al.*, 1995; Kutay and Guttinger, 2005; Wen *et al.*, 1995; Fornerod *et al.*, 1997a; Fukuda *et al.*, 1997). Crm1 is responsible for the export of many proteins of diverse biological significance, as well as several types of RNPs (Table 3). Its name belies its initial identification as a protein that is important for 'chromosomal region maintenance' in *Schizosaccharomyces pombe* (Adachi and Yanagida, 1989).

Exportin-t/Los1 is one of several factors that contributes to export of tRNAs (Arts *et al.*, 1998; Lipowsky *et al.*, 1999; Hopper *et al.*, 2010). Perhaps because of their large size, several exportins and at least one 'mRNA export factor,' Mtr2/Mex67, contribute to ribosomal subunit export (Lo and Johnson, 2009; Woolford and Baserga, 2013; Wild *et al.*, 2010).

Table 3 Mediators of RNA export^a

Class of RNA	Export factors/binding factors
<i>Importin/exportin-related</i>	
rRNA, large subunit	Crm1, XPO5, Mtr2/Mex67, Arx1, Nmd3
rRNA, small subunit	Crm1
5s RNA	La, LS, TFIIA
tRNAs	XPO5, Los1, PAUSED
snRNAs	Crm1, PHAX, CBP
microRNA precursors	XPO5
<i>Independent of importins/exportins</i>	
mRNAs	Mex67/Mtr2 (TAP/p15) and special cases

^aKohler and Hurt, 2007; Taniguchi *et al.*, 2014

In generating these subunits, the nucleolus provides an example of massive transcription, RNP processing, intranuclear transport, and export. Localization of individual components that accomplish these events has defined the contiguous sub-compartments that are involved. Interestingly, the nucleolus is often associated with the inner aspect of the NE (Oakes *et al.*, 1998; Bourgeois *et al.*, 1979).

The repeated use of importins and exportins for successive translocation events mandates that there must be efficient mechanisms to normalize their distribution after release of their cargos (see below for a discussion of protein shuttling per se). One well-studied example of reexport of an import carrier is that of importin α , whose reexport is mediated by the exportin, Cse1. It is not known whether single carriers normalize the distribution of multiple importin family members.

There are also striking examples of nucleocytoplasmic transport that do not directly involve importin family members (Kumeta *et al.*, 2012). These include the import of Ran-GDP itself, that is mediated by Ntf2 (Ribbeck and Gorlich, 2002), import of β -catenin, and mRNA export (see below).

Directionality of Translocation by Importin β Superfamily Members

Cargo translocation by importin β 's occurs in conjunction with the Ras superfamily GTPase, Ran, that cycles between its GDP-bound form (in the cytoplasm) and its GTP-bound form (in the nucleus), due to the polarized distribution of Ran-GAP (GTPase activating protein) and the corresponding guanine nucleotide exchange factor (RanGEF). Ran-GTP confers directionality upon translocation by controlling the conformation of importin β 's, thereby regulating their affinity for cargos. Thus, once inside the nucleus, the binding of Ran-GTP to import complexes (cargo bound to an importin β) causes cargo release – sometimes with the participation of additional cofactors. By contrast, Ran-GTP binding to export complexes (cargo bound to an exportin) in the nucleus stabilizes the complexes until they arrive at the cytoplasmic surface of the NPC, where Ran-binding proteins cause release of Ran-GTP that is rapidly hydrolyzed by local Ran-GAP. Transit of protein cargos through the NPC is not obligatorily coupled to ATP or

GTP hydrolysis. Studies of several cargos show that the only energy requirement for importin-mediated nucleocytoplasmic translocation is due to the need to replenish Ran-GTP (Englmeier *et al.*, 1999; Ribbeck *et al.*, 1999; Schwoebel *et al.*, 1998). Net accumulation of many cargos on one side of the NPC thus requires their binding to targets that are selectively concentrated on the same side. It nevertheless remains possible that especially large cargos present further energy requirements.

Although Ran/TC4 was already known to affect RNA metabolism, DNA synthesis, and cell cycle control (Bischoff *et al.*, 2002), its relation to nucleocytoplasmic transport depended on discovery of its GEF in *S. cerevisiae* (Mtr1/Prp20) as a protein whose mutation causes nuclear accumulation of poly (A) + RNA (Kadowaki *et al.*, 1992; Amberg *et al.*, 1993). Mutations in what proved to be the corresponding Ran-GAP, Rna1, had a comparable effect. Moreover, hamster cells that carry a ts mutation in a homolog of Mtr1/Prp20 – first characterized because they cause premature chromosome condensation (Nishimoto *et al.*, 1978) – also accumulate poly (A) + RNA in the nucleus at the restrictive temperature. In-depth investigation of the Ran cycle was inspired by studies of Ras. This precedent provided quantitative assays for measuring the regulation of GTPase activity and made it evident that specific Ran mutants would be able to mimic the normal GTP- or GDP-bound conformations of Ran.

At least two members of the importin β superfamily (importin 13 (in man), Msn5 (in yeast)) can mediate cargo translocation in both directions. There has been no investigation of the structural correlates of this bidirectionality.

Apart from the issue of whether a given carrier normally translocates distinct cargos in both directions, there is the question of whether a carrier can be caused to translocate the same cargo both into and out of the nucleus. Since translocation requires docking at the NPC, and since the nucleocytoplasmic distribution of nucleoporins is in part asymmetric, one might expect such net reversal to be impossible. Nevertheless, experimental inversion of the Ran-GTP/Ran-GDP gradient can cause Crm1 to function as an importin (Nachury and Weis, 1999). This apparently symmetric performance of the NPC could reflect the relative nonspecificity of importin interactions with the FG-rich internal milieu and/or the observation that many nucleoporins reversibly dissociate from and reassociate with the NPC (Rabut *et al.*, 2004).

mRNA Export

The structural diversity of mRNAs poses a major challenge for their recognition (Wickramasinghe *et al.*, 2014; Ohno *et al.*, 2002). In the best-studied cases, rather than depending on exportins, export is accomplished by a complex set of proteins that bind transcripts during their transcription, while others are added during their covalent maturation (Rodriguez-Navarro and Hurt, 2011; Kyllberg *et al.*, 2008; Natalizio and Wente, 2013). Moreover, interestingly, covalently mature transcripts can be exported upon microinjection into the nucleus. Export of many mRNPs after splicing is mediated by Aly/REF and shuttling SR proteins, followed by intervention of Mex67p/TAP in conjunction with Mtr2p/p15. Release of the

mRNA cargo to the cytoplasm is then due to the activity of the RNA helicase, Dbp5, that is concentrated at the cytoplasmic surface of NPCs and could serve as a ratchet, roughly analogous to the *trans* factors that account for protein import into mitochondria. Inositol phosphates and Gle1 play a central role in these events (Ledoux and Guthrie, 2011; Tseng *et al.*, 1998; Natalizio and Wente, 2013; Schmitt *et al.*, 1999). After cargo release, reimport of mRNA export factors is mediated by the importin, Mtr10, that also functions in reimport of tRNAs (Pemberton *et al.*, 1997; Senger *et al.*, 1998; Hopper *et al.*, 2010).

Remarkably, it has been possible to visualize selected large mRNPs as they are exported through the nuclear pores in insect salivary glands (Kylberg *et al.*, 2008).

Identification of mRNA export factors initially was made possible by generation of corresponding temperature-sensitive mutants of yeast. The key requirements for this undertaking were development of a reliable *in situ* hybridization procedure to detect accumulation of poly(A)⁺ RNA in the nucleus and the devising of appropriate genetic selections. Although some of the resulting lesions were in nucleoporins, others were in the dedicated equipment that manages mRNAs (Farny *et al.*, 2008; Katahira *et al.*, 1999; Kadowaki *et al.*, 1994).

It has not been possible to develop corresponding export assays using permeabilized animal cells, probably because of the need to ensure splicing (Lichtenstein *et al.*, 2001). Oocyte microinjection studies have been highly useful for investigation of mRNA export (Jarmolowski *et al.*, 1994; Ohno *et al.*, 2002; Taniguchi *et al.*, 2014).

Like mRNAs, small structured snRNAs undergo cap methylation; however, their export is mediated by specific associated proteins (PHAX, CBP) that interact with Crm1. Subsequent to export, the cap become trimethylated, the set of Sm proteins is added, and reimport occurs upon recognition by the β family member, snurportin in conjunction with KPNB1 (Matera and Wang, 2014).

Nucleoporins

The NPC consists of about 30 different nucleoporins (Nups). It has eightfold symmetry in the plane of the membrane – as was first recognized by Gall – and each of its proteins is therefore likely to be present in multiples of eight (Antonin *et al.*, 2008; Beck *et al.*, 2007; Lim *et al.*, 2008; Suntharalingam and Wente, 2003).

There are three classes of Nups: ‘structural Nups’ that contribute to overall NPC architecture; ‘pore membrane proteins’ (Poms), that include a transmembrane domain and could contribute to anchoring the NPC in the NE; and ‘FG-Nups,’ that include multiple clusters of phenylalanine-glycine (FG), GLFG, or FxFG repeat motifs that are interspersed among polar sequences. The FG-rich domains of Nups are unstructured and are essential for maintaining the NPC permeability barrier (Denning *et al.*, 2003). Many FG-Nups line the central NPC channel, while others are found in the asymmetric filaments and ‘nuclear basket’ that extend from the cytoplasmic and nucleoplasmic faces of the NPC. Interestingly, the basket anchors components of the SUMO modification cycle, mRNP surveillance factors and spindle assembly checkpoint proteins.

Although the filaments are not essential for survival, they can serve as docking sites for cargo–importin complexes (Richardson *et al.*, 1988; Walther *et al.*, 2002).

During translocation, the surfaces of importins and exportins are thought to engage in multiple low-affinity interactions with FG repeats. Whether they follow a specific sequence of stepping-stones or rather associate indiscriminately with hydrophobic elements in the NPC channel, has been a matter of dispute (Ben-Efraim and Gerace, 2001; Frey and Gorlich, 2007; Peters, 2005; Weis, 2007). Although the latter model has many strengths – and has become amenable to further scrutiny now that macroscopic FG-repeat-containing ‘hydrogel’ preparations are available – the surfaces of importins are by no means uniformly hydrophobic. Moreover, it seems unlikely that importins cover or conceal most of the surface of some cargos that they transport. Translocation through the FG-rich interior of the NPC has been described by a ‘virtual gate model,’ an ‘oily spaghetti model,’ a ‘selective phase model,’ and a ‘reduction of dimensionality model’ (Peters, 2005; Hulsmann *et al.*, 2012).

A more complete understanding of translocation will depend on both the bottoms-up approach of studying small cargos and a top-down emphasis on the huge cargos that presumably were the *raison d’être* of the evolution of the NPC. Any satisfactory model will have to take into consideration the dynamic nature of the NPC itself since many nucleoporins frequently detach from the NPC, possibly as a result of transport of especially large ribonucleoprotein cargos (Rabut *et al.*, 2004). Crystallography of subcomplexes of nucleoporins, as well as methods to tag single cargo molecules, importins and nucleoporins, will surely refine present models of the NPC and understanding of translocation. A major present goal is to obtain a full 3D structure of the NPC, making use of preparations from vertebrates and fungi, including thermophiles. This work is discussed at length in (Aitchison and Rout, 2012) and in the article by Imamoto.

Regulation of Transport

The specificity of cargo–importin interactions raises a fundamental quandary: since many proteins are imported even though they do not share identical motifs, the overall scope of substrate recognition by importins must be broad. Moreover, many proteins of modest size equilibrate across the NE. Thus, at least selected cytoplasmic proteins could wreck havoc if they entered the nucleus. There therefore must be mechanisms that regulate transport in both directions. It is perhaps for this reason that multiple transport signals are often found on many single cargos. Moreover, the activity of transport signals is often regulated by their phosphorylation, acetylation, ubiquitylation, or methylation. Their activity can also be regulated by signal concealment, as for the transcription factor, NF- κ B, that cannot be recognized by importins until its inhibitory subunits have been removed, or for p53, that conceals its NESs when induced to form a tetramer and activate transcription. Interestingly, some translocation signals overlap with domains that have a distinct functional importance. A good example is that of homeobox proteins for which NLSs are situated within the DNA-binding homeodomain itself. Transport can also be

regulated by tethering to cytoplasmic structures, as for the transcription factor, SREBP, that is a transmembrane protein of the ER until released by proteolysis.

The issue of signal concealment and modulation is especially conspicuous for RNPs whose proteins are imported into the nucleus where they are assembled with RNA prior to export (SRP, ribosomal subunits, etc.). Were the multiple NLSs of their proteins to remain exposed after export, these RNPs could well sequester import factors or undergo repeated import.

With regard to global regulation of the transport equipment, phosphorylation of yet unidentified components of the nuclei of digitonin-permeabilized cells affects import (Kehlenbach and Gerace, 2000). Other studies indicate that the diffusional diameter of the NPC and its composition are affected by transformation and during the cell cycle (Feldherr *et al.*, 1998; Makhnevych *et al.*, 2003). Additionally, the cell cycle regulatory protein, Mad1 (that associates with NPCs), regulates both Kap121 and the spindle assembly checkpoint (Cairo *et al.*, 2013; Rodriguez-Bravo *et al.*, 2014). Furthermore, several viruses cause selective destruction of Nup's or interfere with the Ran cycle.

Individual nucleoporins are unevenly expressed among distinct cells types. Moreover, the relative titer of importin β proteins and transcripts varies among cell types and changes during development (Kimura and Imamoto, 2014; Quan *et al.*, 2008). As an extreme example of differential function, the macro- and micronuclei of protozoa express distinct sets of nucleoporins and use different subsets of importins (Malone *et al.*, 2008).

Shuttling

Although most of the transport options described above are unidirectional, many proteins – and even tRNAs and several snRNAs (Hopper *et al.*, 2010) – can be both imported and exported. There is no evidence for such recurrent import and export for the ER or mitochondria. For peroxisomes, at least some carriers do appear to enter and then return to the cytoplasm.

Shuttling between nucleus and cytoplasm can result from successive interactions with distinct importin family members, in which case the cargo has both NLS and NES motifs (Stade *et al.*, 1997). Alternatively, transport in one direction can be passive, as for proteins that are small enough to exit by diffusion through the NPC (Shulga *et al.*, 2000).

For proteins that are highly concentrated in the nucleus, a classical strategy to document shuttling has involved the formation of dikaryons (or multikaryons) as a result of cell–cell fusion. One then inquires whether a macromolecule that is initially present in the nucleus of only one of the two cell types will become detectable also in the second type of nucleus. In such protocols it is critical to stop new synthesis of the cargo in question. Yeast provides a novel way to generate dikaryons by studying zygotes in crosses of *kar1* mutants with wild-type cells. The *kar1* mutants of interest oppose most congression of nuclei after cell–cell fusion (Liang *et al.*, 1996; Flach *et al.*, 1994).

For proteins that are concentrated in the cytoplasm, one could use photobleaching to document repeated entry. Thus, if these cargos are fluorescent, iterated bleaching of the nuclear content would bleach the intranuclear pool and ultimately also deplete the cytoplasmic signal.

Related Issues/Unanticipated Roles of the Nucleus

Related aspects of transport have also been intensely investigated. These include

- The disassembly and ordered reassembly of NPCs during each cell cycle as the NE and lamina break down and reform in higher eukaryotic cells. Moreover, in growing cells, newly synthesized nucleoporins are continually added to existing NPCs (Antonin *et al.*, 2008; Wandke and Kutay, 2013; Rothballer and Kutay, 2013; Doucet and Hetzer, 2010). Interestingly, NPC assembly can regulate abscission at the end of the cell cycle (Mackay and Ullman, 2011).
- The importance of Ran-GTP in associating with chromatin and promoting assembly of the spindle, the NE and NPCs (Hetzer *et al.*, 2002; Walther *et al.*, 2003; Zheng, 2004).
- The reversible sequestration of many proteins in the nucleolus. One example is that of the key cell cycle regulatory phosphatase, Cdc14, that resides in inactive form in the nucleolus until it is progressively released to the nucleoplasm and cytoplasm in anaphase (Manzoni *et al.*, 2010; Audas *et al.*, 2012).

Other less-investigated areas concern

- Signaling, both across the inner nuclear membrane, which may involve SUN/KASH proteins, and also between the nucleoplasm and the vacuole at their site of mutual contact – the nucleus–vacuole junction (Pan *et al.*, 2000).
- Issues of stoichiometry and scale that define the surface area of the NE (Arnone *et al.*, 2013; Hattier *et al.*, 2007; Kessel, 1992; Ralle *et al.*, 2004), the stoichiometry of nucleoporins (some of which are remarkably long-lived; Savas *et al.*, 2012), and the number and turnover of NPCs.
- Turnover of cytoplasmic proteins inside the nucleus (Park *et al.*, 2013) and turnover of parts of the nucleus by 'piecemeal' autophagy (Krick *et al.*, 2008) as well as full nucleation (Keerthivasan *et al.*, 2011).
- Apparent perforation of the NE in laminopathies and upon aging (De Vos *et al.*, 2011; D'Angelo *et al.*, 2009), and as a means of releasing large particles into the cytoplasm (Speese *et al.*, 2012; Buser *et al.*, 2007; Afzelius, 1963).
- The mechanisms of nuclear fusion, in which the outer membranes fuse before the inner membranes, ultimately allowing encounter of parental genomes (Melloy *et al.*, 2007; Tartakoff and Jaiswal, 2009).
- Regulation of gene expression by NPCs and the NE (Kohler and Hurt, 2010; Bermejo *et al.*, 2012; Ishii *et al.*, 2002).

Evolutionary Considerations

Nuclear Envelope

In an attempt to explain the origin of the NE, many authors envisage the envelopment of genomic DNA nucleoids by elements of an endomembrane system related to the ER, causing the genome to be bounded by a double membrane. This model is directly reminiscent of both the wrapping of chromatin that occurs during anaphase of higher eukaryotes, and the

experimental generation of 'artificial nuclei' (with pores) upon incubation of chromatin (or DNA) in extracts of *Xenopus* oocyte cytoplasm (Richardson *et al.*, 1988; Bernis and Forbes, 2014; Newmeyer *et al.*, 1986). A number of prokaryotes have membrane-bounded internal compartments that surround nucleoids; however, they are not fully closed (Fuerst, 2005).

One way to account for the presence of pores is to postulate that the endomembrane in question already included structures related to annulate lamellae or was fenestrated, as are distal cisternae of the Golgi and *trans*-Golgi network. Alternatively, the extremities of the ER-related endomembrane, having already almost encircled the chromatin, could have been decorated with COPII components, as are the transitional elements that normally bud from the ER in the direction of the Golgi. If these coated elements could not fuse with each other, the result would be focal pores limited by COPII orthologs. This scenario could account for the presence of Sec13p both in the NPC and at transitional elements, and also for the presence of Nup's that share a domain with the COPII protein, Sec31p (Devos *et al.*, 2004; Brohawn *et al.*, 2008).

A related and critical aspect of this discussion concerns the origin of the inner membrane of the NE, that is distinct from the ER (Lusk *et al.*, 2007; Schirmer *et al.*, 2003). Among the few proteins of the NE that share motifs with prokaryotes are the proteins that include a HEH motif that is also found in prokaryotic DNA-binding proteins (Mans *et al.*, 2004; Ulbert *et al.*, 2006). The first appearance of membrane proteins that bind chromatin (or the nuclear lamina) could have been crucial for the development of an enclosed genome.

The presence of the nucleus could have made possible sequestration of genes that include introns, e.g. α -proteobacterial/promitochondrial genes that have type II self-splicing introns. Since intron-containing transcripts would code for nonfunctional truncated products, it is necessary to remove the introns before allowing translation. This condition is assured so long as the genome and splicing are segregated from translation, for example, by housing them in distinct compartments, while allowing selective nucleocytoplasmic transport to occur (Cavalier-Smith, 2005; Koonin, 2006; Lopez-Garcia and Moreira, 2006; Mans *et al.*, 2004).

Ran, Nucleoporins, and Importins

The appearance of Ran in evolutionary time may predate the rest of the equipment needed for nucleocytoplasmic transport, considering that the small GTPase superfamily of molecular switches is well-developed in prokaryotes (Jekely, 2003) and since Ran is also central to the organization of chromatin and the spindle. Ran lacks the C-terminal lipid moiety that is characteristic of Ras superfamily members that function in the cytoplasm. Considering the dynamic import and export of Ran, its lack of membrane anchoring is crucial.

Prokaryotes lack proteins closely related to nucleoporins and importins (Mans *et al.*, 2004). Their functional interdependence makes it highly likely that they coevolved, along with Ran (Stuwe *et al.*, 2014). Ongoing investigations, perhaps further advancing the pace of evolution of *S. cerevisiae*, have simplified the NPC by eliminating selected Nup's and removing some FG repeats from those that remain, to generate a 'minimal' NPC (Strawn *et al.*, 2004).

Comparative studies indicate that the most primitive eukaryote is likely to have had on the order of 13 nucleoporins, of which at least eight are present throughout eukaryotes (Mans *et al.*, 2004). Structural Nup's that lack repeats seem the most conserved, and a subset of them including Nup107, Nup160, and ELYS/Mel28 is also implicated in the initial steps of NPC assembly in vertebrates (Antonin *et al.*, 2008; Walther *et al.*, 2003). Orthologs of the Pom, Ndc1, are found in both yeast and vertebrates (but not in the primitive protest, *Giardia lamblia*). Ndc1 also is implicated in NPC formation (Mansfeld *et al.*, 2006). For FG-repeat Nup's, most of the sequences outside of the repeats appear especially variable, apart from those that mediate characteristic interactions (among Nup's, with importins, etc.) (Degrasse *et al.*, 2009; Denning *et al.*, 2003).

Importin β 's can be recognized from yeast to man (Strom and Weis, 2001) and in most species there are more importins than exportins. For example, in man there are 12 importins and 7 exportins. Among the 14 importin β s of *S. cerevisiae*, 10 can be matched with their human orthologs directly (Quan *et al.*, 2008). The C-terminal cargo-binding domain of one importin adapter, snurportin 1 (that imports snRNAs that have a tri-methyl cap, in conjunction with KPNB1), clearly is orthologous to the guanylyl transferase domain of the capping enzyme itself. The observation that no extant species has a small repertoire of importins implies that their multiplicity provides significant advantage.

As a trace of the possible prokaryotic origin of the machinery of export, the GTPases, Nug1p/Nug2p and Nog1p, that function in 60S ribosomal subunit export in eukaryotes, are orthologous to proteins that function in ribosome assembly in prokaryotes (Mans *et al.*, 2004).

The most primitive export equipment for mRNAs could have included proteins related to TAP/Mex67 and p15/Mtr2, considering their widespread distribution. Alternatively, there could have been 'urexportins' that bound Ran-GTP and mature mRNA transcripts. In this event, given the high prevailing GTP/GDP ratio of cells, the key element of directionality could have been provided by the emergence of a Ran equivalent and a cytoplasmic protein with Ran-GAP activity. Curiously, there are bacterial sterol isomerases that share a 'Ntf2 domain' with the Ran-GDP import factor, Ntf2. This motif is also characteristic of the export factors, Mtr2p and TAP/Mex67p (Mans *et al.*, 2004).

Translocons for Oversize Cargos

The need to accommodate large cargos could have first arisen to export RNPs from the nucleus. Other conspicuous examples of transport of oversize cargos are the entry of multisubunit complexes (and gold particles) into the peroxisome and the export of subviral particles from the ER. The fundamental characteristics of the corresponding transport channels could be quite distinct from each other. By contrast to the multi-spanning translocons of the ER, and the β -barrel translocon of the mitochondrion, such extreme distensibility is likely to be possible only when the channel is composed of multiple subunits whose positions relative to each other can be adjusted without dissipating the integrity of the mega-complex, perhaps because they are belted. In this regard, present models of the interior of the NPC envisage elements that self-associate, move relative to each other, and also bind importins, thereby being able to substitute one interaction

for the other. This type of organization would ensure a tight fit regardless of whether cargo is in transit.

Bacteria also translocate hetero-oligomeric and folded proteins. These pathways include (1) Type II secretion, in which cargos are first exported to the periplasmic space by SecY, where they fold and then are secondarily exported through the outer membrane channel formed by multimers of GspD and (2) TAT-dependent transport that is clearly related to the system that transports protein complexes into the thylakoids of chloroplasts. In this system, oligomerization of TatA is thought to generate the corresponding channel in the inner membrane. If this precedent is an accurate guide, the distensible channels of animal cells will also prove to be multimeric. When belted, such units begin to resemble the NPC.

Concluding Remarks

Shuttling between the nucleus and cytoplasm is integral to the organization of eukaryotic cells. It is gratifying to see that the fundamentals of these events have been extensively characterized in a few decades since discovery of nuclear pores. Exhilarating meetings, multiple reviews, and new journals have helped coordinate these efforts.

There are a vast number of specific applications of these fundamental observations and concepts for individual proteins, RNAs, and distinct cell types and organisms. As in much of cell biology, present understanding can be attributed to the application of biochemical, genetic, and cytological approaches. In many cases, this information makes it possible to understand malfunction that underlies disease. Reciprocally, investigations of the molecular basis of disease have often helped elucidate the basic mechanisms of nucleocytoplasmic transport.

Acknowledgments

This work was supported by NIH grant R01GM089872 and by the Visconsi family.

See also: Organelles: Structure and Function: Nuclear Organization (Nuclear Structure and Dynamics)

References

- Adachi, Y., Yanagida, M., 1989. Higher order chromosome structure is affected by cold-sensitive mutations in a *Schizosaccharomyces pombe* gene *crn1+* which encodes a 115-kD protein preferentially localized in the nucleus and its periphery. *Journal of Cell Biology* 108, 1195–1207.
- Adam, S.A., Marr, R.S., Gerace, L., 1990. Nuclear protein import in permeabilized mammalian cells requires soluble cytoplasmic factors. *Journal of Cell Biology* 111, 807–816.
- Atzelius, B.A., 1963. The nucleus of *Noctiluca scintillans*: Aspects of nucleocytoplasmic exchanges and the formation of nuclear membrane. *Journal of Cell Biology* 19, 229–238.
- Aitchison, J.D., Rout, M.P., 2012. The yeast nuclear pore complex and transport through it. *Genetics* 190, 855–883.
- Amberg, D.C., Fleischmann, M., Stagljar, I., Cole, C.N., Aebi, M., 1993. Nuclear PRP20 protein is required for mRNA export. *EMBO Journal* 12, 233–241.
- Antonin, W., Ellenberg, J., Dultz, E., 2008. Nuclear pore complex assembly through the cell cycle: Regulation and membrane organization. *FEBS Letters* 582, 2004–2016.
- Arnone, J.T., Walters, A.D., Cohen-Fix, O., 2013. The dynamic nature of the nuclear envelope: Lessons from closed mitosis. *Nucleus* 4, 261–266.
- Arts, G.J., Kuersten, S., Romby, P., Ehresmann, B., Mattaj, I.W., 1998. The role of exportin-t in selective nuclear export of mature tRNAs. *EMBO Journal* 17, 7430–7441.
- Audas, T.E., Jacob, M.D., Lee, S., 2012. Immobilization of proteins in the nucleolus by ribosomal intergenic spacer noncoding RNA. *Molecular Cell* 45, 147–157.
- Beck, M., Lucic, V., Forster, F., Baumeister, W., Medalia, O., 2007. Snapshots of nuclear pore complexes in action captured by cryo-electron tomography. *Nature* 449, 611–615.
- Belgareh, N., Doye, V., 1997. Dynamics of nuclear pore distribution in nucleoporin mutant yeast cells. *Journal of Cell Biology* 136, 747–759.
- Ben-Efraim, I., Gerace, L., 2001. Gradient of increasing affinity of importin beta for nucleoporins along the pathway of nuclear import. *Journal of Cell Biology* 152, 411–417.
- Bermejo, R., Kumar, A., Foiani, M., 2012. Preserving the genome by regulating chromatin association with the nuclear envelope. *Trends in Cell Biology* 22, 465–473.
- Bernis, C., Forbes, D.J., 2014. Analysis of nuclear reconstitution, nuclear envelope assembly, and nuclear pore assembly using *Xenopus in vitro* assays. *Methods in Cell Biology* 122, 165–191.
- Bischoff, F.R., Scheffzek, K., Ponstingl, H., 2002. How Ran is regulated. *Results and Problems in Cell Differentiation* 35, 49–66.
- Bonner, W.M., 1975. Protein migration into nuclei. I. Frog oocyte nuclei *in vivo* accumulate microinjected histones, allow entry to small proteins, and exclude large proteins. *Journal of Cell Biology* 64, 421–430.
- Bourgeois, C.A., Hemon, D., Bouteille, M., 1979. Structural relationship between the nucleolus and the nuclear envelope. *Journal of Ultrastructure Research* 68, 328–340.
- Brohawn, S.G., Leksa, N.C., Spear, E.D., Rajashankar, K.R., Schwartz, T.U., 2008. Structural evidence for common ancestry of the nuclear pore complex and vesicle coats. *Science* 322, 1369–1373.
- Bucci, M., Wente, S.R., 1997. *In vivo* dynamics of nuclear pore complexes in yeast. *Journal of Cell Biology* 136, 1185–1199.
- Buser, C., Walther, P., Mertens, T., Michel, D., 2007. Cytomegalovirus primary envelopment occurs at large infoldings of the inner nuclear membrane. *Journal of Virology* 81, 3042–3048.
- Cairo, L.V., Ptak, C., Wozniak, R.W., 2013. Mitosis-specific regulation of nuclear transport by the spindle assembly checkpoint protein Mad1p. *Molecular Cell* 49, 109–120.
- Cavalier-Smith, T., 2005. Economy, speed and size matter: Evolutionary forces driving nuclear genome miniaturization and expansion. *Annals of Botany (London)* 95, 147–175.
- Cook, A., Bono, F., Jinek, M., Conti, E., 2007. Structural biology of nucleocytoplasmic transport. *Annual Review of Biochemistry* 76, 647–671.
- D'Angelo, M.A., Raices, M., Panowski, S.H., Hetzer, M.W., 2009. Age-dependent deterioration of nuclear pore complexes causes a loss of nuclear integrity in postmitotic cells. *Cell* 136, 284–295.
- Davis, L.I., Blobel, G., 1986. Identification and characterization of a nuclear pore complex protein. *Cell* 45, 699–709.
- De Vos, W.H., Houben, F., Kamps, M., *et al.*, 2011. Repetitive disruptions of the nuclear envelope invoke temporary loss of cellular compartmentalization in laminopathies. *Human Molecular Genetics* 20, 4175–4186.
- Degrassat, J.A., Dubois, K.N., Devos, D., *et al.*, 2009. Evidence for a shared nuclear pore complex architecture that is conserved from the last common eukaryotic ancestor. *Molecular & Cellular Proteomics* 8 (9), 2119–2130.
- Denning, D.P., Patel, S.S., Uversky, V., Fink, A.L., Rexach, M., 2003. Disorder in the nuclear pore complex: The FG repeat regions of nucleoporins are natively unfolded. *Proceedings of the National Academy of Sciences of the United States of America* 100, 2450–2455.
- Devos, D., Dokudovskaya, S., Alber, F., *et al.*, 2004. Components of coated vesicles and nuclear pore complexes share a common molecular architecture. *PLoS Biology* 2, e380.
- Dingwall, C., Laskey, R.A., 1986. Protein import into the cell nucleus. *Annual Review of Cell and Developmental Biology* 2, 367–390.
- Doucet, C.M., Hetzer, M.W., 2010. Nuclear pore biogenesis into an intact nuclear envelope. *Chromosoma* 119, 469–477.
- Englmeier, L., Olivo, J.C., Mattaj, I.W., 1999. Receptor-mediated substrate translocation through the nuclear pore complex without nucleotide triphosphate hydrolysis. *Current Biology* 9, 30–41.

- Farny, N.G., Hurt, J.A., Silver, P.A., 2008. Definition of global and transcript-specific mRNA export pathways in metazoans. *Genes & Development* 22, 66–78.
- Feldherr, C., Akin, D., Moore, M.S., 1998. The nuclear import factor p10 regulates the functional size of the nuclear pore complex during oogenesis. *Journal of Cell Science* 111 (Pt 13), 1889–1896.
- Feldherr, C.M., Dworetzky, S.I., 1988. The pore complex in nucleocytoplasmic exchange. *Cell Biology International Reports* 12, 791–808.
- Fischer, U., Huber, J., Boelens, W.C., Mattaj, I.W., Luhrmann, R., 1995. The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* 82, 475–483.
- Flach, J., Bossie, M., Vogel, J., *et al.*, 1994. A yeast RNA-binding protein shuttles between the nucleus and the cytoplasm. *Molecular Cell Biology* 14, 8399–8407.
- Floch, A.G., Palancade, B., Doye, V., 2014. Fifty years of nuclear pores and nucleocytoplasmic transport studies: Multiple tools revealing complex rules. *Methods in Cell Biology* 122, 1–40.
- Fornerod, M., Ohno, M., Yoshida, M., Mattaj, I.W., 1997a. CRM1 is an export receptor for leucine-rich nuclear export signals. *Cell* 90, 1051–1060.
- Fornerod, M., Van Deursen, J., Van Baal, S., *et al.*, 1997b. The human homologue of yeast CRM1 is in a dynamic subcomplex with CAN/Nup214 and a novel nuclear pore component Nup88. *EMBO Journal* 16, 807–816.
- Frey, S., Gorlich, D., 2007. A saturated FG-repeat hydrogel can reproduce the permeability properties of nuclear pore complexes. *Cell* 130, 512–523.
- Fuerst, J.A., 2005. Intracellular compartmentation in planctomycetes. *Annual Review of Microbiology* 59, 299–328.
- Fukuda, M., Asano, S., Nakamura, T., *et al.*, 1997. CRM1 is responsible for intracellular transport mediated by the nuclear export signal. *Nature* 390, 308–311.
- Goldberg, M.W., Wiese, C., Allen, T.D., Wilson, K.L., 1997. Dimples, pores, star-rings, and thin rings on growing nuclear envelopes: Evidence for structural intermediates in nuclear pore complex assembly. *Journal of Cell Science* 110, 409–420.
- Goldfarb, D.S., Corbett, A.H., Mason, D.A., Harreman, M.T., Adam, S.A., 2004. Importin alpha: A multipurpose nuclear-transport receptor. *Trends in Cell Biology* 14, 505–514.
- Goldfarb, D.S., Garipey, J., Schoolnik, G., Kornberg, R.D., 1986. Synthetic peptides as nuclear localization signals. *Nature* 322, 641–644.
- Gorlich, D., Dabrowski, M., Bischoff, F.R., *et al.*, 1997. A novel class of RanGTP binding proteins. *Journal of Cell Biology* 138, 65–80.
- Gruenbaum, Y., Margalit, A., Goldman, R.D., Shumaker, D.K., Wilson, K.L., 2005. The nuclear lamina comes of age. *Nature Reviews Molecular Cell Biology* 6, 21–31.
- Guttler, T., Gorlich, D., 2011. Ran-dependent nuclear export mediators: A structural perspective. *EMBO Journal* 30, 3457–3474.
- Hattier, T., Andriulis, E.D., Tartakoff, A.M., 2007. Immobility, inheritance and plasticity of shape of the yeast nucleus. *BMC Cell Biology* 8, 47.
- Hetzer, M., Gruss, O.J., Mattaj, I.W., 2002. The Ran GTPase as a marker of chromosome position in spindle formation and nuclear envelope assembly. *Nature Cell Biology* 4, E177–E184.
- Ho, H.C., 2010. Redistribution of nuclear pores during formation of the redundant nuclear envelope in mouse spermatids. *Journal of Anatomy* 216, 525–532.
- Hoelz, A., Debler, E.W., Blobel, G., 2011. The structure of the nuclear pore complex. *Annual Review of Biochemistry* 80, 613–643.
- Hopper, A.K., Pai, D.A., Engelke, D.R., 2010. Cellular dynamics of tRNAs and their genes. *FEBS Letters* 584, 310–317.
- Hulsmann, B.B., Labokha, A.A., Gorlich, D., 2012. The permeability of reconstituted nuclear pores provides direct evidence for the selective phase model. *Cell* 150, 738–751.
- Ishii, K., Arib, G., Lin, C., Van Houwe, G., Laemmli, U.K., 2002. Chromatin boundaries in budding yeast: The nuclear pore connection. *Cell* 109, 551–562.
- Jarmolowski, A., Boelens, W.C., Izauralde, E., Mattaj, I.W., 1994. Nuclear export of different classes of RNA is mediated by specific factors. *Journal of Cell Biology* 124, 627–635.
- Jekely, G., 2003. Small GTPases and the evolution of the eukaryotic cell. *BioEssays* 25, 1129–1138.
- Kadowaki, T., Chen, S., Hitomi, M., *et al.*, 1994. Isolation and characterization of *Saccharomyces cerevisiae* mRNA transport-defective (mtr) mutants. *Journal of Cell Biology* 126, 649–659.
- Kadowaki, T., Zhao, Y., Tartakoff, A.M., 1992. A conditional yeast mutant deficient in mRNA transport from nucleus to cytoplasm. *Proceedings of the National Academy of Sciences of the United States of America* 89, 2312–2316.
- Kalderon, D., Roberts, B.L., Richardson, W.D., Smith, A.E., 1984. A short amino acid sequence able to specify nuclear location. *Cell* 39, 499–509.
- Katahira, J., Strasser, K., Podtelejnikov, A., *et al.*, 1999. The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *EMBO Journal* 18, 2593–2609.
- Keerthivasan, G., Wickrema, A., Crispino, J.D., 2011. Erythroblast enucleation. *Stem Cells International* 2011, 139851.
- Kehlenbach, R.H., Gerace, L., 2000. Phosphorylation of the nuclear transport machinery down-regulates nuclear protein import *in vitro*. *Journal of Biological Chemistry* 275, 17848–17856.
- Kessel, R.G., 1992. Annulate lamellae: A last frontier in cellular organelles. *International Review of Cytology* 133, 43–120.
- Kimura, M., Imamoto, N., 2014. Biological significance of the importin-beta family-dependent nucleocytoplasmic transport pathways. *Traffic* 15, 727–748.
- Kohler, A., Hurt, E., 2007. Exporting RNA from the nucleus to the cytoplasm. *Nature Reviews Molecular Cell Biology* 8, 761–773.
- Kohler, A., Hurt, E., 2010. Gene regulation by nucleoporins and links to cancer. *Molecular Cell* 38, 6–15.
- Koonin, E.V., 2006. The origin of introns and their role in eukaryogenesis: A compromise solution to the introns-early versus introns-late debate? *Biology Direct* 1, 22.
- Kopito, R.B., Elbaum, M., 2007. Reversibility in nucleocytoplasmic transport. *Proceedings of the National Academy of Sciences of the United States of America* 104, 12743–12748.
- Krick, R., Muehe, Y., Prick, T., *et al.*, 2008. Piecemeal microautophagy of the nucleus requires the core macroautophagy genes. *Molecular Biology of the Cell* 19, 4492–4505.
- Kumeta, M., Yamaguchi, H., Yoshimura, S.H., Takeyasu, K., 2012. Karyopherin-independent spontaneous transport of amphiphilic proteins through the nuclear pore. *Journal of Cell Science* 125, 4979–4984.
- Kutay, U., Guttinger, S., 2005. Leucine-rich nuclear-export signals: Born to be weak. *Trends in Cell Biology* 15, 121–124.
- Kylberg, K., Bjorkroth, B., Ivarsson, B., Fomproix, N., Daneholt, B., 2008. Close coupling between transcription and exit of mRNP from the cell nucleus. *Experimental Cell Research* 314, 1708–1720.
- Langford, R.E., Kanda, P., Kennedy, R.C., 1986. Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal. *Cell* 46, 575–582.
- Ledoux, S., Guthrie, C., 2011. Regulation of the Dbp5 ATPase cycle in mRNP remodeling at the nuclear pore: A lively new paradigm for DEAD-box proteins. *Genes & Development* 25, 1109–1114.
- Liang, S., Hitomi, M., Hu, Y.H., Liu, Y., Tartakoff, A.M., 1996. A DEAD-box-family protein is required for nucleocytoplasmic transport of yeast mRNA. *Molecular Cell Biology* 16, 5139–5146.
- Lichtenstein, M., Guo, W., Tartakoff, A.M., 2001. Control of nuclear export of hnRNP A1. *Traffic* 2, 261–267.
- Lim, R.Y., Ullman, K.S., Fahrenkrog, B., 2008. Biology and biophysics of the nuclear pore complex and its components. *International Review of Cell and Molecular Biology* 267, 299–342.
- Lipowsky, G., Bischoff, F.R., Izauralde, E., *et al.*, 1999. Coordination of tRNA nuclear export with processing of tRNA. *RNA* 5, 539–549.
- Lo, K.Y., Johnson, A.W., 2009. Reengineering ribosome export. *Molecular Biology of the Cell* 20, 1545–1554.
- Lopez-Garcia, P., Moreira, D., 2006. Selective forces for the origin of the eukaryotic nucleus. *BioEssays* 28, 525–533.
- Lusk, C.P., Blobel, G., King, M.C., 2007. Highway to the inner nuclear membrane: Rules for the road. *Nature Reviews Molecular Cell Biology* 8, 414–420.
- Mackay, D.R., Ullman, K.S., 2011. Coordinating postmitotic nuclear pore complex assembly with abscission timing. *Nucleus* 2, 283–288.
- Makhnevych, T., Lusk, C.P., Anderson, A.M., Aitchison, J.D., Wozniak, R.W., 2003. Cell cycle regulated transport controlled by alterations in the nuclear pore complex. *Cell* 115, 813–823.
- Malone, C.D., Falkowska, K.A., Li, A.Y., *et al.*, 2008. Nucleus-specific importin alpha proteins and nucleoporins regulate protein import and nuclear division in the binucleate *Tetrahymena thermophila*. *Eukaryotic Cell* 7, 1487–1499.
- Mans, B.J., Anantharaman, V., Aravind, L., Koonin, E.V., 2004. Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. *Cell Cycle* 3, 1612–1637.
- Mansfeld, J., Guttinger, S., Hawryluk-Gara, L.A., *et al.*, 2006. The conserved transmembrane nucleoporin NDC1 is required for nuclear pore complex assembly in vertebrate cells. *Molecular Cell* 22, 93–103.
- Manzoni, R., Montani, F., Visintin, C., *et al.*, 2010. Oscillations in Cdc14 release and sequestration reveal a circuit underlying mitotic exit. *Journal of Cell Biology* 190, 209–222.
- Matera, A.G., Wang, Z., 2014. A day in the life of the spliceosome. *Nature Reviews Molecular Cell Biology* 15, 108–121.

- Melloy, P., Shen, S., White, E., McIntosh, J.R., Rose, M.D., 2007. Nuclear fusion during yeast mating occurs by a three-step pathway. *Journal of Cell Biology* 179, 659–670.
- Moore, M.S., Blobel, G., 1992. The two steps of nuclear import, targeting to the nuclear envelope and translocation through the nuclear pore, require different cytosolic factors. *Cell* 69, 939–950.
- Nachury, M.V., Weis, K., 1999. The direction of transport through the nuclear pore can be inverted. *Proceedings of the National Academy of Sciences of the United States of America* 96, 9622–9627.
- Natalizio, B.J., Wente, S.R., 2013. Postage for the messenger: Designating routes for nuclear mRNA export. *Trends in Cell Biology* 23, 365–373.
- Newmeyer, D.D., Finlay, D.R., Forbes, D.J., 1986. *In vitro* transport of a fluorescent nuclear protein and exclusion of non-nuclear proteins. *Journal of Cell Biology* 103, 2091–2102.
- Nishimoto, T., Eilen, E., Basilico, C., 1978. Premature of chromosome condensation in a ts DNA- mutant of BHK cells. *Cell* 15, 475–483.
- Oakes, M., Aris, J.P., Brockenbrough, J.S., *et al.*, 1998. Mutational analysis of the structure and localization of the nucleolus in the yeast *Saccharomyces cerevisiae*. *Journal of Cell Biology* 143, 23–34.
- Ohno, M., Segref, A., Kuersten, S., Mattaj, I.W., 2002. Identity elements used in export of mRNAs. *Molecular Cell* 9, 659–671.
- Pan, X., Roberts, P., Chen, Y., *et al.*, 2000. Nucleus-vacuole junctions in *Saccharomyces cerevisiae* are formed through the direct interaction of Vac8p with Nvj1p. *Molecular Biology of the Cell* 11, 2445–2457.
- Park, S.H., Kukushkin, Y., Gupta, R., *et al.*, 2013. PolyQ proteins interfere with nuclear degradation of cytosolic proteins by sequestering the Sis1p chaperone. *Cell* 154, 134–145.
- Pemberton, L.F., Rosenblum, J.S., Blobel, G., 1997. A distinct and parallel pathway for the nuclear import of an mRNA-binding protein. *Journal of Cell Biology* 139, 1645–1653.
- Peters, R., 2005. Translocation through the nuclear pore complex: Selectivity and speed by reduction-of-dimensionality. *Traffic* 6, 421–427.
- Powell, L., Burke, B., 1990. Internuclear exchange of an inner nuclear membrane protein (p55) in heterokaryons: *In vivo* evidence for the interaction of p55 with the nuclear lamina. *Journal of Cell Biology* 111, 2225–2234.
- Quan, Y., Ji, Z.L., Wang, X., Tartakoff, A.M., Tao, T., 2008. Evolutionary and transcriptional analysis of karyopherin beta superfamily proteins. *Molecular & Cellular Proteomics* 7, 1254–1269.
- Rabut, G., Doye, V., Ellenberg, J., 2004. Mapping the dynamic organization of the nuclear pore complex inside single living cells. *Nature Cell Biology* 6, 1114–1121.
- Ralle, T., Grund, C., Franke, W.W., Stick, R., 2004. Intranuclear membrane structure formations by CaaX-containing nuclear proteins. *Journal of Cell Science* 117, 6095–6104.
- Ribbeck, K., Gorlich, D., 2002. The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO Journal* 21, 2664–2671.
- Ribbeck, K., Kutay, U., Paraskeva, E., Gorlich, D., 1999. The translocation of transport-cargo complexes through nuclear pores is independent of both Ran and energy. *Current Biology* 9, 47–50.
- Richardson, W.D., Mills, A.D., Dilworth, S.M., Laskey, R.A., Dingwall, C., 1988. Nuclear protein migration involves two steps: Rapid binding at the nuclear envelope followed by slower translocation through nuclear pores. *Cell* 52, 655–664.
- Riva, A., Loffredo, F., Segawa, A., Congiu, T., Testa Riva, F., 1998. Human salivary gland parenchymal cells seen by SEM from the cytoplasmic side using a new osmium maceration method. *European Journal of Morphology* 36 (Suppl.), 35–40.
- Rodriguez-Bravo, V., Maciejowski, J., Corona, J., *et al.*, 2014. Nuclear pores protect genome integrity by assembling a premitotic and Mad1-dependent anaphase inhibitor. *Cell* 156, 1017–1031.
- Rodriguez-Navarro, S., Hurt, E., 2011. Linking gene regulation to mRNA production and export. *Current Opinion in Cell Biology* 23, 302–309.
- Rothballer, A., Kutay, U., 2013. Poring over pores: Nuclear pore complex insertion into the nuclear envelope. *Trends in Biochemical Sciences* 38, 292–301.
- Savas, J.N., Toyama, B.H., Xu, T., Yates 3rd, J.R., Hetzer, M.W., 2012. Extremely long-lived nuclear pore proteins in the rat brain. *Science* 335, 942.
- Schirmer, E.C., Florens, L., Guan, T., Yates 3rd, J.R., Gerace, L., 2003. Nuclear membrane proteins with potential disease links found by subtractive proteomics. *Science* 301, 1380–1382.
- Schmitt, C., Von Kobbe, C., Bachi, A., *et al.*, 1999. Dbp5, a DEAD-box protein required for mRNA export, is recruited to the cytoplasmic fibrils of nuclear pore complex via a conserved interaction with CAN/Nup159p. *EMBO Journal* 18, 4332–4347.
- Schwoebel, E.D., Talcott, B., Cushman, I., Moore, M.S., 1998. Ran-dependent signal-mediated nuclear import does not require GTP hydrolysis by Ran. *Journal of Biological Chemistry* 273, 35170–35175.
- Senger, B., Simos, G., Bischoff, F.R., *et al.*, 1998. Mtr10p functions as a nuclear import receptor for the mRNA-binding protein Npl3p. *EMBO Journal* 17, 2196–2207.
- Shulga, N., Mosammaparast, N., Wozniak, R., Goldfarb, D.S., 2000. Yeast nucleoporins involved in passive nuclear envelope permeability. *Journal of Cell Biology* 149, 1027–1038.
- Speer, H.L., Zimmerman, E.F., 1968. The transfer of proteins from cytoplasm to nucleus in HeLa cells. *Biochemical and Biophysical Research Communications* 32, 60–65.
- Speese, S.D., Ashley, J., Jokhi, V., *et al.*, 2012. Nuclear envelope budding enables large ribonucleoprotein particle export during synaptic Wnt signaling. *Cell* 149, 832–846.
- Stade, K., Ford, C.S., Guthrie, C., Weis, K., 1997. Exportin 1 (Crm1p) is an essential nuclear export factor. *Cell* 90, 1041–1050.
- Stevens, B.J., Swift, H., 1966. RNA transport from nucleus to cytoplasm in *Chironomus* salivary glands. *Journal of Cell Biology* 31, 55–77.
- Strawn, L.A., Shen, T., Shulga, N., Goldfarb, D.S., Wente, S.R., 2004. Minimal nuclear pore complexes define FG repeat domains essential for transport. *Nature Cell Biology* 6, 197–206.
- Strom, A.C., Weis, K., 2001. Importin-beta-like nuclear transport receptors. *Genome Biology* 2, Reviews3008.
- Stuwe, T., Lin, D.H., Collins, L.N., Hurt, E., Hoelz, A., 2014. Evidence for an evolutionary relationship between the large adaptor nucleoporin Nup192 and karyopherins. *Proceedings of the National Academy of Sciences of the United States of America* 111, 2530–2535.
- Suel, K.E., Cansizoglu, A.E., Chook, Y.M., 2006. Atomic resolution structures in nuclear transport. *Methods* 39, 342–355.
- Suntharalingam, M., Wente, S.R., 2003. Peering through the pore: Nuclear pore complex structure, assembly, and function. *Developmental Cell* 4, 775–789.
- Taniguchi, I., McCloskey, A., Ohno, M., 2014. Analysis of RNA transport in *Xenopus* oocytes and mammalian cells. *Methods in Cell Biology* 122, 395–413.
- Tartakoff, A.M., Jaiswal, P., 2009. Nuclear fusion and genome encounter during yeast zygote formation. *Molecular Biology of the Cell* 20, 2932–2942.
- Tseng, S.S., Weaver, P.L., Liu, Y., *et al.*, 1998. Dbp5p, a cytosolic RNA helicase, is required for poly(A)⁺ RNA export. *EMBO Journal* 17, 2651–2662.
- Ulbert, S., Platani, M., Boue, S., Mattaj, I.W., 2006. Direct membrane protein-DNA interactions required early in nuclear envelope assembly. *Journal of Cell Biology* 173, 469–476.
- Walther, T.C., Askjaer, P., Gentzel, M., *et al.*, 2003. RanGTP mediates nuclear pore complex assembly. *Nature* 424, 689–694.
- Walther, T.C., Pickersgill, H.S., Cordes, V.C., *et al.*, 2002. The cytoplasmic filaments of the nuclear pore complex are dispensable for selective nuclear protein import. *Journal of Cell Biology* 158, 63–77.
- Wandke, C., Kutay, U., 2013. Enclosing chromatin: Reassembly of the nucleus after open mitosis. *Cell* 152, 1222–1225.
- Weis, K., 2007. The nuclear pore complex: Oily spaghetti or gummy bear? *Cell* 130, 405–407.
- Wen, W., Meinkoth, J.L., Tsien, R.Y., Taylor, S.S., 1995. Identification of a signal for rapid export of proteins from the nucleus. *Cell* 82, 463–473.
- Wickramasinghe, V.O., Andrews, R., Ellis, P., *et al.*, 2014. Selective nuclear export of specific classes of mRNA from mammalian nuclei is promoted by GANP. *Nucleic Acids Research* 42, 5059–5071.
- Wild, T., Horvath, P., Wyler, E., *et al.*, 2010. A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. *PLoS Biology* 8, e1000522.
- Woolford Jr., J.L., Baserga, S.J., 2013. Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*. *Genetics* 195, 643–681.
- Worman, H.J., 2012. Nuclear lamins and laminopathies. *Journal of Pathology* 226, 316–325.
- Wright, R., Basson, M., D'Ari, L., Rine, J., 1988. Increased amounts of HMG-CoA reductase induce "karmellae": A proliferation of stacked membrane pairs surrounding the yeast nucleus. *Journal of Cell Biology* 107, 101–114.
- Zheng, Y., 2004. G protein control of microtubule assembly. *Annual Review of Cell and Developmental Biology* 20, 867–894.

The Molecular Architecture of Cell–Cell Adhesions

B Geiger, Weizmann Institute of Science, Rehovot, Israel

R Zaidel-Bar and M Vaman Rao, National University of Singapore, Singapore

© 2016 Elsevier Inc. All rights reserved.

Cell–Cell Adhesion: The Essence of Multicellularity

All animals begin their lives as single cells. Cell growth and division lead to multicellularity and expansion of the organism; concomitantly, cell differentiation gives rise to a variety of cell types with specialized functions. Nonetheless, it is cell adhesion that ultimately organizes these cells into tissues and organs, and unites them into a single organism. Depending on their type and location, cells in our body engage in interactions with other cells, and/or with the extracellular matrix surrounding them. For example, endothelial cells lining the interior of blood vessels interact with neighboring endothelial cells and with the basal lamina, whereas chondrocytes producing the cartilage between our bones interact only with the matrix they secrete, and are embedded within.

Interactions between a cell and its environment serve multiple purposes: they provide the cell with positional cues (e.g., apical/basal and front/back); they facilitate biochemical signaling (e.g., growth factors); and they enable mechanical coupling, important for regulating cell shape (e.g., spreading and polarity), for integrating cells into tissues, and for transforming cell dynamics into tissue dynamics (e.g., gastrulation and migration). Both cell–matrix and cell–cell adhesions are critical for the development and homeostasis of multicellular organisms. While they share many functions as well as molecular pathways, they also display distinct features. In this

article, we will focus our attention on cell–cell adhesions. Cell–matrix adhesions are discussed in a separate article.

The adhesion of one cell to another is not a spontaneous process. The plasma membrane itself is not ‘sticky,’ and the intrinsic contractility of cells acts to round their surfaces. Cell–cell adhesion is primarily achieved through the action of transmembrane adhesion receptors that counteract the cell’s surface tension by generating adhesion tension, and through regulation of the cortical cytoskeleton (Maitre and Heisenberg, 2013).

As will be elaborated on below, mammalian cells possess several different types of adhesion receptors. Each type clusters with receptors of the same kind to form distinct cell–cell adhesion sites, also referred to as cell–cell junctions. The formation of such junctions is nucleated by adhesion-independent nanoclusters of receptors that are delimited by the actin cortex (Wu *et al.*, 2015). The cytoplasmic tails of the receptors recruit adaptor proteins, which in turn bind a variety of structural, scaffolding, and signaling proteins that ultimately integrate the nascent adhesion site with the cytoskeleton (Garrod and Chidgey, 2008; Zaidel-Bar, 2013; Guo *et al.*, 2014). Signaling components play a dual role: they respond to extracellular signals to affect cell behavior, and they regulate cell adhesion in response to internal signaling.

As demonstrated in Figure 1, interactions between adhesion receptors of the same kind (homophilic) or of a

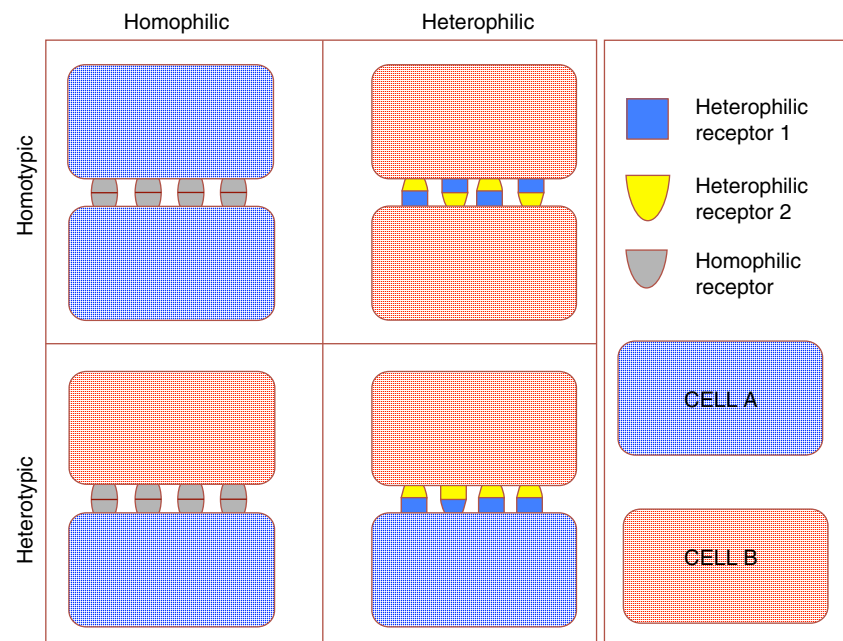


Figure 1 Different forms of cell–cell adhesions. Illustration depicting several possible modes of interaction between similar or different cells, mediated via similar or different adhesion receptors.

different kind (heterophilic) have been shown to mediate both interactions between cells of the same type (homotypic), and interactions between cells of different origins (heterotypic). It is important to note that while [Figure 1](#) highlights a specific example for each type of interaction, any two cells are likely to be interacting through multiple types of adhesion receptors.

The aim of this article is to provide molecular insights into the mechanisms by which the structural and signaling modules of cell–cell junctions function in tandem, in order to provide the cell with a responsive, dynamic junction that is fully integrated with the cell's activities and fate.

Structural and Functional Diversity of Cell–Cell Junctions

Cell–cell adhesions (=junctions) differ structurally, molecularly, and functionally, according to characteristic subcellular positions and functions. Based on their physiological role, cell–cell junctions can be broadly divided into three types: occluding, anchoring, and signaling ([Figure 2](#)).

Occluding junctions are only found in epithelial sheets, such as the lining of the gut and kidney, where formation of a barrier between the two sides of the monolayer is essential.

Proteins of the claudin and occludin family mediate occluding adhesions in invertebrate tight junctions ([Furuse et al., 1998](#)), and closely related proteins function similarly in invertebrate septate junctions. Tight junctions in invertebrates are located at the most apical position along the lateral membrane of neighboring cells. In addition to sealing the space between cells, tight junctions also create a ‘fence’ within the cell's membrane that helps to separate the apical and baso-lateral membrane domains ([Farquhar and Palade, 1963](#)).

The most ancient anchoring junction found in all animal cells is the adherens junction (AJ), which is based on cadherin adhesion receptors linked to the actin cytoskeleton ([Farquhar and Palade, 1963](#)). AJs provide a physical connection between cells of the body by reinforcing their cell–cell adhesions with the actomyosin cytoskeleton ([Yonemura, 2011](#)). In the following sections, we will focus on AJs, their receptors, their connection with the actin cytoskeleton, and regulation of their function by an ensemble of proteins we refer to as the ‘cadherin adhesome’ ([Zaidel-Bar, 2013](#)).

Unique to vertebrates is a second type of anchoring junction: the desmosome. Cell–cell adhesion in desmosomes is based on specialized cadherins known as desmogleins and desmocollins ([Chitaev and Troyanovsky, 1997](#)). Unlike AJs, these adhesion receptors are connected within the cell to

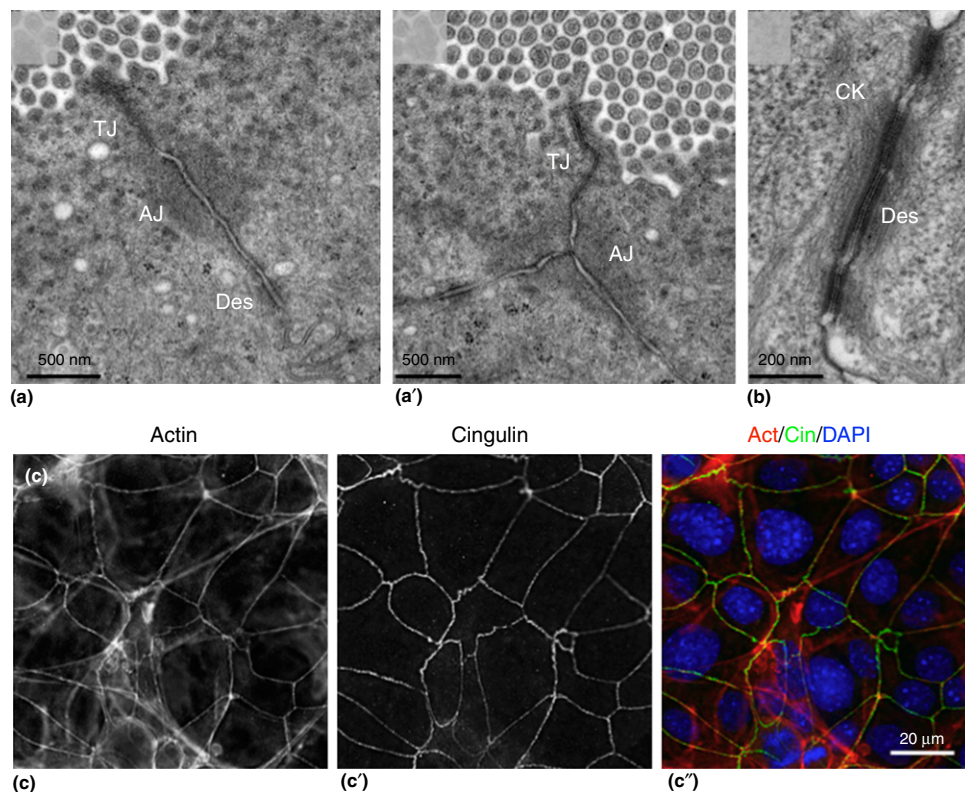


Figure 2 Diversity of cell–cell junctions in mouse epithelial cells. Transmission electron micrographs of mouse small intestine (a and a', presented at two orientations) and tongue (b), showing different cell–cell junctions, including tight junctions (TJ), adherens junctions (AJ), and desmosomes (Des). (b) Highly developed desmosomes in stratified epithelium of the mouse tongue, displaying the laminated structure of these junctions and the associated cytokeratins (CK). (c–c''): Cultured YAMC epithelial cells derived from mouse intestine, triple-labeled for actin (c), showing the adherens junctions and (out of focus) the microfilament bundles associated with the basal matrix adhesions of the cells; cingulin-labeled tight junctions (c') and the two labels together (actin – red; cingulin – green), plus DAPI staining of the nuclei (blue), in the triple-colored image (c''). Scale bar: 20 μ m.

intermediate filaments. Desmosomes provide additional mechanical resilience to vertebrate epithelial tissues, such as skin (Garrod and Chidgey, 2008).

Signaling or communication junctions include the channel-forming GAP junctions and neuronal synapses, as well as the more transient stimulatory immune synapse, involved in T-cell activation (Rozental *et al.*, 2000). In contrast to tight junctions that block the passage of molecules in the space between neighboring cells, GAP junctions bridge this gap, by forming channels through which small molecules such as ions, amino acids, and secondary messengers can freely pass from one cell to another (Mese *et al.*, 2007). GAP junctions, formed by four-pass transmembrane proteins from the connexin family, are found in many animal tissues, where they serve to coordinate the activity of cells, such as those in the beating heart muscle (Mese *et al.*, 2007).

Neuronal and immune synapses are specialized junctions that combine adhesion and signaling proteins for the localized transfer of information between particular pairs of cells. Though they will not be discussed here, the interested reader is directed to the following reviews (Rozental *et al.*, 2000; Shimizu and Stopfer, 2013).

The Cadherin Family and Diversity of AJs

Interactions between cadherin molecules are mediated through specific cadherin extracellular domains (EC). Structurally, these domains are about 110 amino acids long, and fold into a distinctive immunoglobulin-like shape (Parisini *et al.*, 2007). There are over a hundred proteins in the human genome that contain EC domains, which together comprise the cadherin protein superfamily (Sotomayor *et al.*, 2014). Various cadherins have been shown to be essential for basic processes occurring in cells and tissues, such as neurodevelopment, cell polarization, and migration.

In AJs, however, only a small subset of this diverse family, termed ‘classical cadherins,’ function as adhesion receptors. ‘Classical’ cadherins in humans contain five EC domains, but are best defined by their cytoplasmic tail, which contains a

juxtamembrane binding site for p120-catenin, and a C-terminal binding site for beta-catenin (van Roy and Berx, 2008). Based on conservation of the cytoplasmic tail and overall functional similarity, classical cadherins have also been identified in other organisms such as *Drosophila* and *Caenorhabditis elegans*, despite considerable differences in their extracellular domains (Hill *et al.*, 2001; Oda *et al.*, 1994). While cadherins were found in the pre-metazoan protist *Monosiga brevicollis*, the catenin-binding cytoplasmic tail is believed to be a metazoan development, which may have constituted an important evolutionary turning point in facilitating stable cell–cell adhesions in the earliest metazoans (Abedin and King, 2008; Murray and Zaidel-Bar, 2014).

The human genome encodes 18 classical cadherins (henceforth referred to as cadherins), but only a few of them have been extensively studied. The first cadherin (CDH1) was discovered using antibodies that were capable of inhibiting calcium-dependent cell–cell adhesions, and was thus named cadherin, for ‘Ca²⁺-dependent adhesion’ (Yoshida-Noro *et al.*, 1984). Subsequently, other cadherins were discovered and named after the tissues in which they were found to be prominently expressed. Thus, CDH1 is also known as E (epithelial)-cadherin, CDH2 is N (neuronal)-cadherin, CDH5 is VE (vascular endothelial)-cadherin, CDH3 is P (placental)-cadherin, and CDH15 is M (muscle)-cadherin (Hatta *et al.*, 1985).

As discussed below, the cytoplasmic tails of cadherins associate with a large number of structural and regulatory proteins. Whether the composition of this network of interacting proteins differs, if at all, among the different cadherin subtypes, is currently unknown. However, it is known that different cadherins organize into AJs with distinct structures, as observed with light microscopy. Thus, while E-cadherin forms a continuous adhesion belt along the circumference of the cell in the apical-most region of the lateral membranes (Figure 3(a)), N-cadherin and VE-cadherin commonly form intermittent, zipper-like adhesion structures perpendicular to the line of contact between neighboring cells (Takeichi, 2014).

Importantly, AJ morphology may differ within the same cell type, depending on parameters such as cell density and tension.

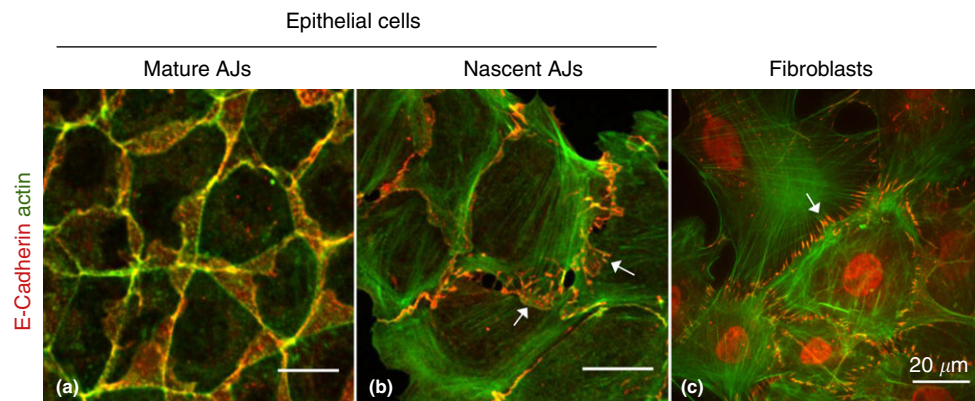


Figure 3 Morphological variations of adherens junctions. Illustration of different morphologies of AJs identified in epithelial monolayers and cultures of fibroblasts. (a) E-cadherin (red) and F-actin (green) co-localize to form a continuous adhesion belt along the apical circumference of epithelial cells in a confluent epithelial monolayer. (b) Epithelial cells at the edge of a monolayer forming punctate AJs (white arrows). (c) Streak-like AJs (white arrows) of fibroblasts, with associated actin cables perpendicular to the AJ membrane. Scale bars: 20 μm .

For example, in endothelial cells, stable junctions have been shown to be transformed into focal AJs after treatment with the hormone thrombin (Rabiet *et al.*, 1996); moreover, epithelial cells at the edge of a monolayer exhibit punctate AJs, in contrast to the linear junctions displayed by cells within the sheet (Figure 3(b); Yonemura *et al.*, 1995). Even within the same cell, more than one type of junction can be observed. For example, in both epithelial and endothelial cells, in addition to the more prominent apical junctions, there exists a pool of lateral, spot-like junctions (Takeichi, 2014), which have been observed to translocate along actin filaments in a basal to apical direction, and join the apical junctions (Hong *et al.*, 2010).

Little is known about the mechanism controlling the morphology of a given junction, but it appears to be tightly connected to the F-actin structures associated with it. The lateral adhesion belt in polarized epithelial cells is associated with a thick bundle of F-actin that runs parallel to the membrane, at a short distance from it (Yonemura *et al.*, 1995). Streak-like AJs in fibroblasts and endothelial cells, on the other hand, are associated with stress fibers that approach the membrane at an angle close to perpendicular, and appear to connect to the AJ (Figure 3(c); Noria *et al.*, 1999). Mechanical continuity has also been noted between stress fibers in neighboring cells, mediated by the AJ (Millan *et al.*, 2010).

The Molecular Basis for Cadherin–Cadherin Binding

The extracellular domains of several classical cadherins have been crystallized (Harrison *et al.*, 2011; Shapiro *et al.*, 1995), and based on their solved structures, a model for cadherin–cadherin interactions has emerged (Brasch *et al.*, 2012; Figure 4(a)). Calcium ions bind in between consecutive cadherin domains, and act to rigidify the crescent-shaped extracellular domain and facilitate *trans* interactions. The primary adhesive interaction involves strand swapping between the first cadherin domain (EC1) (closest to the N-terminus) of apposing cadherins. Depending on the particular cadherin subtype, this involves one or two tryptophan residues that insert themselves into a hydrophobic pocket of the apposing cadherin (Brasch *et al.*, 2011; Harrison *et al.*, 2011).

Another type of *trans* interaction, observed in the crystal structure when the tryptophan is mutated, is known as an X-dimer (Harrison *et al.*, 2010). It has been suggested that the X-dimer, which is weaker than the strand-swapped dimer, might represent an intermediate interaction that occurs as the strand-swapped dimer is being assembled or disassembled (Harrison *et al.*, 2010; Hong *et al.*, 2011). The crystal packing also reveals the basis of the *cis* interaction (between cadherins on the same cell), which is mediated by interfaces on EC1 and EC2 (Harrison *et al.*, 2011). Importantly, the *cis* and *trans* interactions occupy different surfaces of EC1, such that a given cadherin can engage in both *cis* and *trans* interactions, and thus create a lattice of tightly packed dimers (Harrison *et al.*, 2011). Molecular dynamics modeling suggests that *trans* dimerization stabilizes cadherin in such a way that *cis* interactions become more probable, and *vice versa*, creating a positive feedback mechanism promoting the growth of a polarized lattice (Wu *et al.*, 2011). Whether cadherin in cells organizes in a manner

similar to that observed in the crystal structure remains unproven. However, super-resolution microscopy has revealed molecular densities at the core of cadherin clusters that are equivalent to the densities expected for the crystal structure packing (Wu *et al.*, 2015).

The Cytoplasmic Plaque

Cadherin's 150 amino acid-long cytoplasmic tail is unstructured on its own, but is stabilized and protected from proteolysis by binding two armadillo repeat proteins from the catenin family of adaptors. P120-catenin (or one of its paralogs: delta-catenin, ARVCF, p0071, and plakophilin) binds at the juxtamembrane region, and beta-catenin (or its close paralog, plakoglobin) binds at the C-terminus of the cytoplasmic tail (Pokutta and Weis, 2007). The binding of p120-catenin not only protects the tail from proteolysis, but also serves to conceal an endocytic signal (Hartsock and Nelson, 2012). Beta-catenin, which was shown to bind to the cadherin tail already in the endoplasmic reticulum (Chen *et al.*, 1999), has many important binding partners, most notably vinculin (Hazan *et al.*, 1997) and its close relative, alpha-catenin (Drees *et al.*, 2005).

In addition to these two armadillo repeat catenins, cadherin cytoplasmic tails have been shown to bind directly with more than a dozen other proteins, some of which (e.g., ankyrin, IQGAP1, and NUMB) are structural proteins and adaptors, whereas others (e.g., RapGEF1, CSK, and Trio) primarily display regulatory functions (Zaidel-Bar, 2013). The binding of these many proteins to the relatively short cadherin tail is mostly mutually exclusive, and dependent on their availability and relative binding affinities.

The binding affinities of the various binding partners can be modulated by post-translational modifications, primarily phosphorylation of specific tyrosine, serine, and threonine residues along the cadherin tail. For example, tyrosine phosphorylation of cadherin creates docking sites for the SH2 domain of the adaptor SHC1 (Xu and Carpenter, 1999) and the PTB domain of the cell polarity protein Numb (Wang *et al.*, 2009). Conversely, phosphorylation of serine residues by casein kinase inhibits the binding of p120 and beta-catenin (Dupre-Crochet *et al.*, 2007). Similar switch-like regulation by post-translational modification governs the interactions between many, if not all, of the plaque proteins at the cytoplasmic face of AJs. In this manner, signaling pathways can rapidly change AJ structure and dynamics.

Dynamic Connections between Cadherin and the Actin Cytoskeleton

Studies in developmental models, as well as cells in cultured epithelial monolayers, provide convincing evidence of a mechanical link between cadherin receptors and the actin cytoskeleton (Borghi *et al.*, 2012; Martin *et al.*, 2009). Indeed, it has been established that lack of anchorage to the actin cytoskeleton renders cadherin clusters unstable, whereas coupling the latter to the cytoskeleton confers stability and directed motility (Hong *et al.*, 2013). Though the molecular

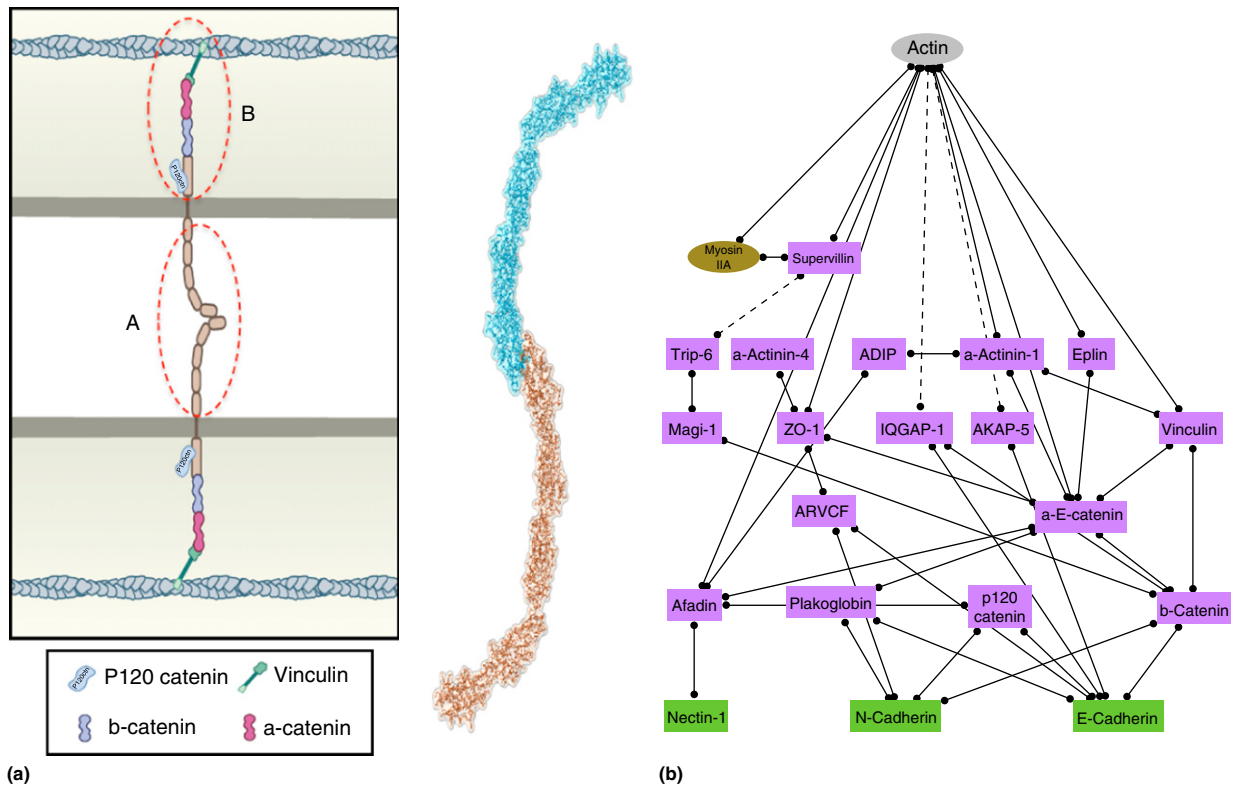


Figure 4 Schematic representation of E-cadherin receptors, interacting with each other via their EC1 domains, and linked to the actin cytoskeleton via α -catenin and vinculin. The encircled cadherin regions (A and B in the box on the left) are illustrated in (a), showing a ball-and-stick model for trans interactions between cadherin extracellular domains; and (b), presenting a network of interactions whereby adhesion receptors (E- or N-Cadherin and Nectin-1) are linked to the actin cytoskeleton, via adaptor proteins.

nature of this connection is still not fully understood, several key proteins have been identified. Foremost amongst the linkers is α -catenin, which contains binding sites for F-actin and for β -catenin (Kobielak and Fuchs, 2004). In *in vitro* biochemical assays, the binding of α -catenin to β -catenin precluded its binding to F-actin (Drees *et al.*, 2005). This finding highlights the fact that conventional biochemical assays fail to simulate the cellular conditions *in vivo*, such as tension applied by actomyosin to the proteins involved. Indeed, a binding assay incorporating forces was able to recapitulate the β -catenin/ α -catenin/F-actin complex (Buckley *et al.*, 2014). Moreover there is substantial evidence that α -catenin at AJs exists in a stretched conformation (Yonemura *et al.*, 2010). The stretching of α -catenin in response to force leads to the unfurling of the protein, and exposes an N-terminal modulatory 1 (M1) domain that binds to vinculin (Ishiyama *et al.*, 2013; Yonemura *et al.*, 2010).

Vinculin at AJs could also conceivably form a link between the cadherin/catenin complex and F-actin. Moreover, α -catenin- and vinculin-dependent recruitment of the vasodilator-stimulated phosphoprotein facilitates actin polymerization at AJs (Vasioukhin *et al.*, 2000). Several other actin-binding proteins, namely, eplin, filamin-A, α -actinin-4, and supervillin, can bind to α -catenin, and thus form a three-component link between cadherin and F-actin (Zaidel-Bar, 2013). Vinculin can also bind to β -catenin directly, and independently of

α -catenin (Peng *et al.*, 2010; Figure 4(b)). Furthermore, cadherin is by no means the only trans-membrane protein at AJs: several other receptors have known connections with F-actin. For example, afadin connects the nectin receptor (Takai and Nakanishi, 2003), and myosin VIIA connects vezatin with F-actin (Kussel-Andermann *et al.*, 2000).

Although the connection with F-actin is by far the most prominent, AJs have also been observed to form connections with the microtubule cytoskeleton and, in some cases, also link with intermediate filaments (Meng and Takeichi, 2009). Anchoring of the minus ends of microtubules to AJs has been described, mediated by dynein or a nezha-plekha7 complex that, in turn, binds to β -catenin or p120-catenin, respectively (Ligon *et al.*, 2001; Meng *et al.*, 2008).

Regulatory Elements of the Cadherin Adhesome

Epithelial cells not only build a complex, structural AJ composed of dozens of proteins, but also engage numerous regulatory networks to enable precise control of AJ assembly, turnover, maintenance and disassembly (Figure 5). The most important regulators of the adhesion plaque include GTPases, activating guanine-nucleotide exchange factors (GEFs) and inhibitory GTPase-activating proteins (GAPs) (Braga and Yap, 2005), kinases, and phosphatases (Bertocchi *et al.*, 2012).

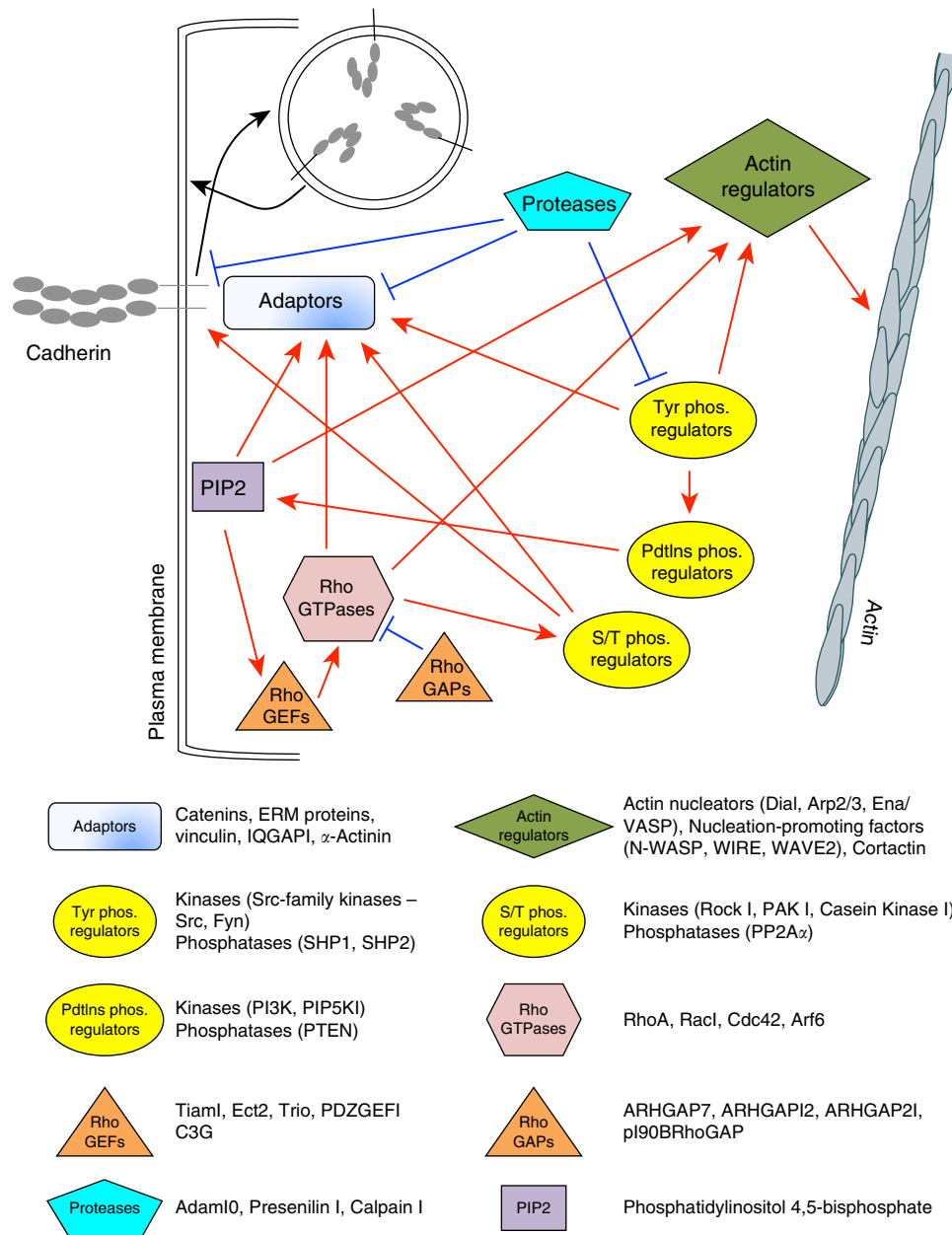


Figure 5 Interplay between regulators at Adherens Junctions. Regulation of cadherin receptors, adaptor proteins and actin regulators is achieved through the action of Rho GTPases, kinases, phosphatases, phospholipids and proteases. Red and blue arrows represent activating and inhibitory interactions, respectively, whereas black arrows represent exocytosis and endocytosis. Representative members from each group of proteins are listed in the box (GEFs, Guanine-nucleotide exchange factors; GAPs, GTPase-activating proteins; Phos, Phosphorylation).

Rho-family GTPases are important regulators of AJs, primarily due to their role in controlling actin polymerization and myosin-driven contractility in cells. The earliest stage of establishment of cell–cell adhesions is driven by high Rac1 activity that stimulates Arp2/3 complex-dependent formation of lamellipodial protrusions, and establishment of contact between cells (Yamada and Nelson, 2007). Following establishment of *de novo* contacts, lamellipodia experience a transient reduction in Rac1 activity, as E-cadherin accumulates and forms nascent AJs. A second wave of elevated RhoA activity

then promotes myosin-based contractility, necessary to expand and complete the formation of a mature junction (Yamada and Nelson, 2007).

Although this type of switch in GTPase signaling may appear conceptually simple, reality is often more complex, to ensure tight spatiotemporal control of GTPase signaling via regulated GEF and GAP activity. For instance, E-cadherin ligation is known to activate phosphatidylinositol 3-kinase (PI 3-kinase) that, in turn, regulates a Rac-GEF called Tiam1 (Nakagawa *et al.*, 2001). Such a signaling cascade may serve to

locally activate Rac1 at sites of cell–cell contact initiation. More recent studies demonstrate novel roles for proteins such as central spindlein, a known cytokinesis regulator, in locally recruiting RhoGEF (ect2) while inhibiting RhoGAP (p190B RhoGAP) localization at AJs, to fine-tune the levels of active RhoA at AJs (Ratheesh *et al.*, 2012).

Besides being regulated by GTPase activity, AJs serve as sites for kinase and phosphatase activity, with at least 21 serine/threonine and tyrosine kinases and 13 phosphatases known to function at AJs (Bertocchi *et al.*, 2012). Phosphorylation or dephosphorylation events can exert one of several possible effects on AJ resident proteins – including stabilizing, destabilizing or altering the binding affinities of protein pairs and complexes (Bertocchi *et al.*, 2012). Although several kinases and phosphatases are known to interact with E-cadherin or the catenins, as determined by immunofluorescence or co-immunoprecipitation, limited knowledge exists concerning the mechanism of localization, and activation of the majority of them.

Elucidating the effects of phosphorylation on AJs has proven to be challenging, due to both positive and negative effects of kinases on AJ assembly, maintenance, and function. For example, increased Src kinase activity in epithelial cells leads to elevated levels of tyrosine phosphorylation, resulting in the subsequent disassembly of AJs (Owens *et al.*, 2000). Although this would suggest negative regulation of AJs by Src kinase, evidence from other studies points to a positive role for Src at cell–cell junctions. Clustering of cadherin molecules during AJ formation triggers Src kinase activity, which is required for further activation of the PI 3-Kinase signaling cascade (Pang *et al.*, 2005). Another member of the Src kinase family, Fyn, has also been positively implicated in the formation of AJs in keratinocytes (Calautti *et al.*, 1998).

Phosphorylation events mediated by tyrosine kinases also contribute to AJ remodeling by regulating E-cadherin endocytosis. Activation of the tyrosine kinase Src in epithelial cells leads to a marked increase in E-cadherin and β -catenin phosphorylation, which in turn stimulates ubiquitination and endocytosis of E-cadherin from AJs (Fujita *et al.*, 2002). Ubiquitination entails the attachment of multiple ubiquitin molecules to a protein, which then induces degradation of the protein by the cell's proteolytic machinery. The E3 ubiquitin ligase Hakai has been found to bind E-cadherin in a tyrosine phosphorylation-dependent manner, and induce ubiquitination and endocytosis as a consequence of binding (Fujita *et al.*, 2002). More recent studies suggest competition in the binding of p120-catenin and Hakai to the juxtamembrane domain of E-cadherin (Hartsock and Nelson, 2012). When bound to the juxtamembrane region of E-cadherin, p120-catenin masks the site for ubiquitination; conversely, Hakai-mediated ubiquitination of the cadherin cytoplasmic tail prevents p120-catenin from binding to and stabilizing E-cadherin at AJs (Hartsock and Nelson, 2012).

Dynamics of Actin at AJs

The dynamics and remodeling of AJs are closely intertwined with those of the actomyosin cytoskeleton; therefore, it is essential to understand the nature of actin dynamics that facilitate AJ remodeling. AJs constitute sites of actin assembly and disassembly, with several nucleation factors known to play

prominent roles in supporting actin polymerization at junctions. Members of the formin family, most notably mammalian Diaphanous 1 (mDia1), have established roles in strengthening cadherin-mediated junctions (Carramusa *et al.*, 2007). The actin-related protein 2/3 (Arp2/3) complex is also known to localize at AJs; inhibiting Arp2/3 complex activity reduces actin nucleation and F-actin content at cadherin adhesions (Kovacs *et al.*, 2002). The Arp2/3 complex itself displays poor intrinsic actin polymerization activity, and is activated by several nucleation promoting factors (NPFs), including cortactin, and members of the WASP (Wiskott-Aldrich Syndrome Protein) and WAVE (WASP-family verprolin homologous protein) families. While several of these NPFs have been found at AJs (Han *et al.*, 2014; Verma *et al.*, 2012), the specific role of each in supporting junctional actin assembly remains unclear. Several other proteins may function as co-factors, to regulate Arp2/3-mediated actin polymerization at AJs. For instance, evidence from *in vitro* studies points to a role for alpha-catenin in inhibiting Arp2/3 activity (Drees *et al.*, 2005), while a more recent study demonstrates a role for alpha-actinin-4 in supporting Arp2/3-dependent actin assembly at cell–cell junctions (Tang and Brieher, 2012).

In addition to direct regulation of actin nucleation at AJs, several proteins contribute to homeostasis of actin filaments at AJs at a post-nucleation stage (Chesarone and Goode, 2009). Depletion of the pointed-end capping protein Tropomodulin-3 (TMOD3) is known to destabilize cell–cell adhesion, due to loss of F-actin along the lateral membranes of polarized epithelial cells (Weber *et al.*, 2007). Evidence from *in vivo* studies also demonstrates a role for tropomodulins in protecting actin filaments that are recruited by alpha-catenin during epithelial morphogenesis in the *C. elegans* embryo (Cox-Paulson *et al.*, 2012). Tropomyosins – integral components of actin filaments – also contribute to stabilization of F-actin at cell–cell junctions, and are essential for maintenance of proper cell shape *in vitro* and *in vivo* (Temm-Grove *et al.*, 1998). Conversely, proteins such as gelsolin that possess potent actin severing properties also modulate actin dynamics at AJs. Phosphoinositide signaling regulates activity of gelsolin that, in turn, contributes to regulation of adhesion strength at N-cadherin-mediated cell–cell junctions (El Sayegh *et al.*, 2007).

Cadherin and Embryonic Development

Cell–cell adhesion plays a critical role in embryonic development, beginning as early as compaction at the eight-cell stage in the developmental timeline of vertebrate embryos. Compaction marks the first differentiation event during development, and requires E-cadherin-mediated cell adhesion to coordinate blastocyst formation from an embryo composed of eight spherical blastomeres. Deletion of the *CDH1* gene in mouse embryos, or treatment with function-blocking E-cadherin antibodies, leads to compaction defects and nonviable embryos (Larue *et al.*, 1994). More recently, elegant experiments using live-embryo imaging demonstrated the extension of E-cadherin-dependent filopodia by blastomeres onto their neighboring cells (Fierro-Gonzalez *et al.*, 2013). Formation of such filopodia is tightly coordinated with cell shape changes, and facilitates

tension generation that helps brings cells closer together, to undergo compaction (Fierro-Gonzalez *et al.*, 2013).

Besides the expression of E-cadherin, several other types of cadherins are expressed in a tissue-specific manner during the course of development. Tissue-specific expression patterns of cadherins facilitate ‘cell sorting,’ a phenomenon that has been identified both *in vivo* and in tissue culture systems, and is based on their differential affinity to other cadherins (Duguay *et al.*, 2003). N-cadherin, for example, forms homodimers with higher affinity than those formed with E-cadherin, so a mixture of cells expressing either will tend to segregate into an inner cluster of N-cadherin expressing cells surrounded by E-cadherin expressing cells (Foty and Steinberg, 2005). Cell sorting mediated by expression and homophilic ligation of compatible cadherin subtypes enables the demarcation of cell types into distinct tissues during embryogenesis. For instance, during oogenesis in *Drosophila*, oocytes localize at the posterior end of the follicle, due to higher levels of expression of DE-cadherin in the posterior follicle cell that favors oocyte-follicle cell interactions (Godt and Tepass, 1998).

However, there also occurs during development the process of ‘cadherin switching,’ which leads to changes in expression from one cadherin type to another. Primary neurulation in chick embryos is associated with a switch in expression from E-cadherin to N-cadherin, to enable formation of a neural tube from a neural plate, and distinguish the neural tube from epidermal cells (Hatta and Takeichi, 1986). A shift towards N-cadherin expression provides neural plate cells with the ability to delaminate and migrate, a phenomenon known as epithelial-to-mesenchymal transition (EMT). Gastrulation during *Drosophila* development also involves a transition from DE-cadherin to DN-cadherin expression in developing mesodermal tissue, accompanied by EMT (Oda *et al.*, 1998).

AJs are intimately associated with the actomyosin cytoskeleton via several intermediate proteins, and the importance of this connection is evident in a variety of processes throughout the course of morphogenesis. Germband extension in *Drosophila* results in narrowing of the embryo along the dorsal-ventral axis, concomitant with elongation along the anterior-posterior axis (Bertet *et al.*, 2004). This process is driven by the intercalation of ectodermal cells that occurs as a result of AJ remodeling and actomyosin contractility (Rauzi *et al.*, 2010). The distributions of DE-cadherin and Myosin-II are planar polarized and complementary to one another, an arrangement that is essential for proper intercalation. The planar-polarized distribution of DE-cadherin orients the flow of apical actomyosin networks toward junctions with low DE-cadherin and higher myosin-II levels, thus causing these junctions to contract preferentially, to enable cell intercalation (Rauzi *et al.*, 2010). Several other processes such as apical constriction during gastrulation also require coordination between junction remodeling and contractions of the actomyosin network, thus illustrating the prominent role played by AJs during morphogenesis (Mason *et al.*, 2013).

AJs in Disease

Cancer metastasis is a complex process, often involving loss of cell–cell adhesion, increased migration, and invasion of cancer

cells into distant tissues, followed by colonization and formation of secondary tumors. By virtue of facilitating cell–cell adhesion via formation of AJs, loss of E-cadherin function appears to promote metastasis and invasiveness in human epithelial tumors, due to weaker contacts between neighboring cells. Down-regulation of E-cadherin expression is often correlated with highly metastatic cancers, and associated with adverse patient outcomes (Asgeirsson *et al.*, 2000). However, it is important to note that loss of E-cadherin, in and of itself, is not sufficient to drive cancer or metastasis, but rather accompanies cellular transformation that involves many other changes (Shamir *et al.*, 2014).

In the vast majority of primary cancers affecting multiple organ types, transcriptional down-regulation of E-cadherin expression via hyper-methylation of the *CDH1* promoter contributes to loss of cell–cell adhesion (Bex and van Roy, 2009). Excessive methylation of CpG islands in the promoter region of E-cadherin triggers histone 3 (H3) deacetylation, which shuts down transcription (Chen *et al.*, 2003). In addition to promoter hyper-methylation, the binding of transcription factors such as Snail, Slug and Twist to E-box sequences within the *CDH1* promoter are also known to suppress transcription and promote tumor invasion (Batlle *et al.*, 2000). Expression of these transcription factors is associated with EMT and increased migratory potential that, in turn, are commonly associated with cancer metastasis. Next-generation sequencing techniques have also demonstrated a role for inactivating mutations in the *CDH1* gene as a trigger for promotion of weaker cell adhesion and metastasis (Bex *et al.*, 1996). However, the effects of such mutations are limited to breast and gastric cancers (Soares *et al.*, 1997).

It is interesting to note that reduced expression of E-cadherin in tumor cells is frequently associated with higher levels of N-cadherin expression, an adhesion receptor predominantly found in neurons and mesenchymal cells. While this process of ‘cadherin switching’ occurs during developmental processes such as gastrulation, it is likely that increased N-cadherin expression might contribute to acquisition of migratory phenotypes among cancer cells, and hence promote the spread of tumors.

The function of E-cadherin in facilitating cell adhesion is intimately associated with its ability to bind to members of the catenin family via its cytoplasmic tail region. It is thus not surprising that changes in catenin expression, in particular alpha-catenin, closely mirror those of E-cadherin in a wide range of epithelial cancers (Zhou *et al.*, 2005). Although alpha-catenin levels can be reduced to as much as 40–80% in a variety of cancers (Zhou *et al.*, 2005), little is known about the mechanisms involved in down-regulation of its expression. While loss-of-function mutations in the alpha-catenin gene have been identified in *in vitro* culture systems for tumor cells, limited evidence exists that would clarify their roles *in vivo* in primary and secondary tumor formation.

Although E-cadherin and the catenins have been the focus of intensive research for their causal roles in cancer metastasis, it bears remembering that other human diseases also arise out of defects in function of nonclassical or desmosomal cadherins. Desmosomal cadherins have been predominantly implicated in diseases affecting skin and hair follicles, while defects in nonclassical cadherins often result in neurosensory complications that manifest as retinal or inner ear diseases.

For instance, the nonclassical cadherins, Cadherin-23 and Protocadherin-15, which form the tip link in hair cells in the inner ear, are mutated in Usher syndrome (USH), resulting in blindness and deafness (El-Amraoui and Petit, 2005). We refer readers to several reviews that further highlight the involvement of desmosomal and nonclassical cadherins in disease development and progression in humans (Al-Jassar *et al.*, 2013; El-Amraoui and Petit, 2005).

Concluding Comments

In this article, we have discussed, with some historical perspective, the structure, functions, and diversity of intercellular junctions, and their involvement in major processes in metazoan organisms, in health and disease. We have addressed the critical roles of these adhesions in the evolution of multicellularity, and focused on both their cardinal mechanical role in supporting tissue scaffolding and morphogenesis, and their involvement in intercellular communication, critical for body homeostasis. Zooming in on the nanostructure of cell–cell adhesions, we discussed the multimolecular networks that, in concert, perform the scaffolding, sensing, and signaling tasks of the different adhesion sites. In this context, much attention is focused on the roles of the cytoskeleton in physically supporting cell adhesions and in participating in local mechanosensing, and the involvement of a wide variety of adhesion-associated enzymes (e.g., kinases, phosphatases, and their regulators) in the signaling process. It is noteworthy that recent studies, based on advanced and powerful molecular genetic and imaging technologies, combined with *in vivo* and *ex vivo* approaches, provide a wealth of information on the molecular components of the various adhesion systems, and their individual contributions to the functions of these sites. These studies also make it increasingly clear that the next major challenge is to address junction biology via integrative, ‘systems level’ approaches, whereby the diverse functions of these cellular structures are evaluated in the context of networking of their multiple constituents.

Acknowledgments

BG is the incumbent of the Erwin Neter Professorial Chair in Cell and Tumor Biology. RZB was supported by the National Research Foundation Singapore under its NRF Fellowship (NRF-RF2009-RF001-074). The authors are grateful to Barbara Morgenstern for her expert editorial assistance.

See also: Cell Communication: Cellular Machineries: Extracellular Regulation of Cell-to-Matrix Adhesion

References

- Abedin, M., King, N., 2008. The premetazoan ancestry of cadherins. *Science* 319, 946–948.
- Al-Jassar, C., Bikker, H., Overduin, M., Chidgey, M., 2013. Mechanistic basis of desmosome-targeted diseases. *Journal of Molecular Biology* 425, 4006–4022.
- Asgeirsson, K.S., Jonasson, J.G., Tryggvadottir, L., *et al.*, 2000. Altered expression of E-cadherin in breast cancer patterns, mechanisms and clinical significance. *European Journal of Cancer* 36, 1098–1106.
- Battle, E., Sancho, E., Franci, C., *et al.*, 2000. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nature Cell Biology* 2, 84–89.
- Bertet, C., Sulak, L., Lecuit, T., 2004. Myosin-dependent junction remodelling controls planar cell intercalation and axis elongation. *Nature* 429, 667–671.
- Bertocchi, C., Vaman Rao, M., Zaidel-Bar, R., 2012. Regulation of adherens junction dynamics by phosphorylation switches. *Journal of Signal Transduction* 2012, 125295.
- Berx, G., Cleton-Jansen, A.M., Strumane, K., *et al.*, 1996. E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain. *Oncogene* 13, 1919–1925.
- Berx, G., van Roy, F., 2009. Involvement of members of the cadherin superfamily in cancer. *Cold Spring Harbor Perspectives in Biology* 1, a003129.
- Borghi, N., Sorokina, M., Shcherbakova, O.G., *et al.*, 2012. E-cadherin is under constitutive actomyosin-generated tension that is increased at cell–cell contacts upon externally applied stretch. *Proceedings of the National Academy of Sciences of the United States of America* 109, 12568–12573.
- Braga, V.M., Yap, A.S., 2005. The challenges of abundance: Epithelial junctions and small GTPase signalling. *Current Opinion in Cell Biology* 17, 466–474.
- Brasch, J., Harrison, O.J., Ahlsen, G., *et al.*, 2011. Structure and binding mechanism of vascular endothelial cadherin: A divergent classical cadherin. *Journal of Molecular Biology* 408, 57–73.
- Brasch, J., Harrison, O.J., Honig, B., Shapiro, L., 2012. Thinking outside the cell: How cadherins drive adhesion. *Trends in Cell Biology* 22, 299–310.
- Buckley, C.D., Tan, J., Anderson, K.L., *et al.*, 2014. Cell adhesion. The minimal cadherin–catenin complex binds to actin filaments under force. *Science* 346 (6209), 1254211.
- Calautti, E., Cabodi, S., Stein, P.L., *et al.*, 1998. Tyrosine phosphorylation and src family kinases control keratinocyte cell–cell adhesion. *Journal of Cell Biology* 141, 1449–1465.
- Carramusa, L., Ballestrem, C., Zilberman, Y., Bershadsky, A.D., 2007. Mammalian diaphanous-related formin Dia1 controls the organization of E-cadherin-mediated cell–cell junctions. *Journal of Cell Science* 120, 3870–3882.
- Chen, C.L., Liu, S.S., Ip, S.M., *et al.*, 2003. E-cadherin expression is silenced by DNA methylation in cervical cancer cell lines and tumours. *European Journal of Cancer* 39, 517–523.
- Chen, Y.T., Stewart, D.B., Nelson, W.J., 1999. Coupling assembly of the E-cadherin/beta-catenin complex to efficient endoplasmic reticulum exit and basal-lateral membrane targeting of E-cadherin in polarized MDCK cells. *Journal of Cell Biology* 144, 687–699.
- Chesarone, M.A., Goode, B.L., 2009. Actin nucleation and elongation factors: Mechanisms and interplay. *Current Opinion in Cell Biology* 21, 28–37.
- Chitaev, N.A., Troyanovsky, S.M., 1997. Direct Ca²⁺-dependent heterophilic interaction between desmosomal cadherins, desmoglein and desmocollin, contributes to cell–cell adhesion. *Journal of Cell Biology* 138, 193–201.
- Cox-Paulson, E.A., Walck-Shannon, E., Lynch, A.M., *et al.*, 2012. Tropomodulin protects alpha-catenin-dependent junctional-actin networks under stress during epithelial morphogenesis. *Current Biology*: CB 22, 1500–1505.
- Drees, F., Pokutta, S., Yamada, S., Nelson, W.J., Weis, W.I., 2005. Alpha-catenin is a molecular switch that binds E-cadherin–beta-catenin and regulates actin–filament assembly. *Cell* 123, 903–915.
- Duguay, D., Foty, R.A., Steinberg, M.S., 2003. Cadherin-mediated cell adhesion and tissue segregation: Qualitative and quantitative determinants. *Developmental Biology* 253, 309–323.
- Dupre-Crochet, S., Figueroa, A., Hogan, C., *et al.*, 2007. Casein kinase 1 is a novel negative regulator of E-cadherin-based cell–cell contacts. *Molecular and Cellular Biology* 27, 3804–3816.
- El Sayegh, T.Y., Arora, P.D., Ling, K., *et al.*, 2007. Phosphatidylinositol-4,5 biphosphate produced by PIP5Kgamma regulates gelsolin, actin assembly, and adhesion strength of N-cadherin junctions. *Molecular Biology of the Cell* 18, 3026–3038.
- El-Amraoui, A., Petit, C., 2005. Usher I syndrome: Unravelling the mechanisms that underlie the cohesion of the growing hair bundle in inner ear sensory cells. *Journal of Cell Science* 118, 4593–4603.
- Farquhar, M.G., Palade, G.E., 1963. Junctional complexes in various epithelia. *Journal of Cell Biology* 17, 375–412.
- Fierro-Gonzalez, J.C., White, M.D., Silva, J.C., Plachta, N., 2013. Cadherin-dependent filopodia control preimplantation embryo compaction. *Nature Cell Biology* 15, 1424–1433.

- Foty, R.A., Steinberg, M.S., 2005. The differential adhesion hypothesis: A direct evaluation. *Developmental Biology* 278, 255–263.
- Fujita, Y., Krause, G., Scheffner, M., *et al.*, 2002. Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex. *Nature Cell Biology* 4, 222–231.
- Furuse, M., Fujita, K., Hiiiragi, T., Fujimoto, K., Tsukita, S., 1998. Claudin-1 and -2: Novel integral membrane proteins localizing at tight junctions with no sequence similarity to occludin. *Journal of Cell Biology* 141, 1539–1550.
- Garrod, D., Chidgey, M., 2008. Desmosome structure, composition and function. *Biochimica et Biophysica Acta* 1778, 572–587.
- Godt, D., Tepass, U., 1998. *Drosophila* oocyte localization is mediated by differential cadherin-based adhesion. *Nature* 395, 387–391.
- Guo, Z., Neilson, L.J., Hang, Z., *et al.*, 2014. E-cadherin adhesome complexity and robustness resolved by quantitative proteomics. *Science Signaling* 7 (354), rs7.
- Han, S.P., Gambin, Y., Gomez, G.A., *et al.*, 2014. Cortactin scaffolds Arp2/3 and WAVE2 at the epithelial zonula adherens. *Journal of Biological Chemistry* 289, 7764–7775.
- Harrison, O.J., Bahna, F., Katsamba, P.S., *et al.*, 2010. Two-step adhesive binding by classical cadherins. *Nature Structural & Molecular Biology* 17, 348–357.
- Harrison, O.J., Jin, X., Hong, S., *et al.*, 2011. The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins. *Structure* 19, 244–256.
- Hartssock, A., Nelson, W.J., 2012. Competitive regulation of E-cadherin juxtamembrane domain degradation by p120-catenin binding and Hakai-mediated ubiquitination. *PLOS ONE* 7, e37476.
- Hatta, K., Okada, T.S., Takeichi, M., 1985. A monoclonal antibody disrupting calcium-dependent cell-cell adhesion of brain tissues: Possible role of its target antigen in animal pattern formation. *Proceedings of the National Academy of Sciences of the United States of America* 82, 2789–2793.
- Hatta, K., Takeichi, M., 1986. Expression of N-cadherin adhesion molecules associated with early morphogenetic events in chick development. *Nature* 320, 447–449.
- Hazan, R.B., Kang, L., Roe, S., Borgen, P.I., Rimm, D.L., 1997. Vinculin is associated with the E-cadherin adhesion complex. *Journal of Biological Chemistry* 272, 32448–32453.
- Hill, E., Broadbent, I.D., Chothia, C., Pettitt, J., 2001. Cadherin superfamily proteins in *Caenorhabditis elegans* and *Drosophila melanogaster*. *Journal of Molecular Biology* 305, 1011–1024.
- Hong, S., Troyanovsky, R.B., Troyanovsky, S.M., 2010. Spontaneous assembly and active disassembly balance adherens junction homeostasis. *Proceedings of the National Academy of Sciences of the United States of America* 107, 3528–3533.
- Hong, S., Troyanovsky, R.B., Troyanovsky, S.M., 2011. Cadherin exits the junction by switching its adhesive bond. *Journal of Cell Biology* 192, 1073–1083.
- Hong, S., Troyanovsky, R.B., Troyanovsky, S.M., 2013. Binding to F-actin guides cadherin cluster assembly, stability, and movement. *Journal of Cell Biology* 201, 131–143.
- Ishiyama, N., Tanaka, N., Abe, K., *et al.*, 2013. An autoinhibited structure of alpha-catenin and its implications for vinculin recruitment to adherens junctions. *Journal of Biological Chemistry* 288, 15913–15925.
- Kobieliak, A., Fuchs, E., 2004. Alpha-catenin: At the junction of intercellular adhesion and actin dynamics. *Nature Reviews Molecular Cell Biology* 5, 614–625.
- Kovacs, E.M., Goodwin, M., Ali, R.G., Paterson, A.D., Yap, A.S., 2002. Cadherin-directed actin assembly: E-cadherin physically associates with the Arp2/3 complex to direct actin assembly in nascent adhesive contacts. *Current Biology: CB* 12, 379–382.
- Kussel-Andermann, P., El-Amraoui, A., Safieddine, S., *et al.*, 2000. Vezatin, a novel transmembrane protein, bridges myosin VIIA to the cadherin-catenins complex. *EMBO Journal* 19, 6020–6029.
- Larue, L., Ohsugi, M., Hirchenhain, J., Kemler, R., 1994. E-cadherin null mutant embryos fail to form a trophectoderm epithelium. *Proceedings of the National Academy of Sciences of the United States of America* 91, 8263–8267.
- Ligon, L.A., Karki, S., Tokito, M., Holzbaur, E.L., 2001. Dynein binds to beta-catenin and may tether microtubules at adherens junctions. *Nature Cell Biology* 3, 913–917.
- Maitre, J.L., Heisenberg, C.P., 2013. Three functions of cadherins in cell adhesion. *Current Biology: CB* 23, R626–R633.
- Martin, A.C., Kaschube, M., Wieschaus, E.F., 2009. Pulsed contractions of an actin-myosin network drive apical constriction. *Nature* 457, 495–499.
- Mason, F.M., Tworoger, M., Martin, A.C., 2013. Apical domain polarization localizes actin-myosin activity to drive ratchet-like apical constriction. *Nature Cell Biology* 15, 926–936.
- Meng, W., Mushika, Y., Ichii, T., Takeichi, M., 2008. Anchorage of microtubule minus ends to adherens junctions regulates epithelial cell-cell contacts. *Cell* 135, 948–959.
- Meng, W., Takeichi, M., 2009. Adherens junction: Molecular architecture and regulation. *Cold Spring Harbor Perspectives in Biology* 1, a002899.
- Mese, G., Richard, G., White, T.W., 2007. Gap junctions: Basic structure and function. *Journal of Investigative Dermatology* 127, 2516–2524.
- Millan, J., Cain, R.J., Reglero-Real, N., *et al.*, 2010. Adherens junctions connect stress fibres between adjacent endothelial cells. *BMC Biology* 8, 11.
- Murray, P., Zaidel-Bar, R., 2014. Premetazoan origins and evolution of the cadherin adhesome. *Biology Open* 3, 1183–1195.
- Nakagawa, M., Fukata, M., Yamaga, M., Itoh, N., Kaibuchi, K., 2001. Recruitment and activation of Rac1 by the formation of E-cadherin-mediated cell-cell adhesion sites. *Journal of Cell Science* 114, 1829–1838.
- Noria, S., Cowan, D.B., Gotlieb, A.I., Langille, B.L., 1999. Transient and steady-state effects of shear stress on endothelial cell adherens junctions. *Circulation Research* 85, 504–514.
- Oda, H., Tsukita, S., Takeichi, M., 1998. Dynamic behavior of the cadherin-based cell-cell adhesion system during *Drosophila* gastrulation. *Developmental Biology* 203, 435–450.
- Oda, H., Uemura, T., Harada, Y., Iwai, Y., Takeichi, M., 1994. A *Drosophila* homolog of cadherin associated with armadillo and essential for embryonic cell-cell adhesion. *Developmental Biology* 165, 716–726.
- Owens, D.W., McLean, G.W., Wyke, A.W., *et al.*, 2000. The catalytic activity of the Src family kinases is required to disrupt cadherin-dependent cell-cell contacts. *Molecular Biology of the Cell* 11, 51–64.
- Pang, J.H., Kraemer, A., Stehens, S.J., Frame, M.C., Yap, A.S., 2005. Recruitment of phosphoinositide 3-kinase defines a positive contribution of tyrosine kinase signaling to E-cadherin function. *Journal of Biological Chemistry* 280, 3043–3050.
- Parisini, E., Higgins, J.M., Liu, J.H., Brenner, M.B., Wang, J.H., 2007. The crystal structure of human E-cadherin domains 1 and 2, and comparison with other cadherins in the context of adhesion mechanism. *Journal of Molecular Biology* 373, 401–411.
- Peng, X., Cuff, L.E., Lawton, C.D., DeMali, K.A., 2010. Vinculin regulates cell-surface E-cadherin expression by binding to beta-catenin. *Journal of Cell Science* 123, 567–577.
- Pokutta, S., Weis, W.I., 2007. Structure and mechanism of cadherins and catenins in cell-cell contacts. *Annual Review of Cell and Developmental Biology* 23, 237–261.
- Rabiet, M.J., Plantier, J.L., Rival, Y., *et al.*, 1996. Thrombin-induced increase in endothelial permeability is associated with changes in cell-to-cell junction organization. *Arteriosclerosis, Thrombosis, and Vascular Biology* 16, 488–496.
- Ratheesh, A., Gomez, G.A., Priya, R., *et al.*, 2012. Centralspindlin and alpha-catenin regulate Rho signalling at the epithelial zonula adherens. *Nature Cell Biology* 14, 818–828.
- Rauzi, M., Lenne, P.F., Lecuit, T., 2010. Planar polarized actomyosin contractile flows control epithelial junction remodelling. *Nature* 468, 1110–1114.
- Rozental, R., Giaume, C., Spray, D.C., 2000. Gap junctions in the nervous system. *Brain Research. Brain Research Reviews* 32, 11–15.
- Shamir, E.R., Pappalardo, E., Jorgens, D.M., *et al.*, 2014. Twist1-induced dissemination preserves epithelial identity and requires E-cadherin. *Journal of Cell Biology* 204 (5), 839–856.
- Shapiro, L., Fannon, A.M., Kwong, P.D., *et al.*, 1995. Structural basis of cell-cell adhesion by cadherins. *Nature* 374, 327–337.
- Shimizu, K., Stopfer, M., 2013. Gap junctions. *Current Biology: CB* 23, R1026–R1031.
- Soares, P., Bex, G., van Roy, F., Sobrinho-Simoes, M., 1997. E-cadherin gene alterations are rare events in thyroid tumors. *International Journal of Cancer. Journal International du Cancer* 70, 32–38.
- Sotomayor, M., Gaudet, R., Corey, D.P., 2014. Sorting out a promiscuous superfamily: Towards cadherin connectomics. *Trends in Cell Biology* 24, 524–536.
- Takai, Y., Nakanishi, H., 2003. Nectin and afadin: Novel organizers of intercellular junctions. *Journal of Cell Science* 116, 17–27.
- Takeichi, M., 2014. Dynamic contacts: Rearranging adherens junctions to drive epithelial remodelling. *Nature Reviews. Molecular Cell Biology* 15, 397–410.
- Tang, V.W., Brieher, W.M., 2012. alpha-Actinin-4/FGS1 is required for Arp2/3-dependent actin assembly at the adherens junction. *Journal of Cell Biology* 196, 115–130.
- Temmm-Grove, C.J., Jockusch, B.M., Weinberger, R.P., Schvezov, G., Helfman, D.M., 1998. Distinct localizations of tropomyosin isoforms in LLC-PK1 epithelial cells suggests specialized function at cell-cell adhesions. *Cell Motility and the Cytoskeleton* 40, 393–407.

- van Roy, F., Berx, G., 2008. The cell-cell adhesion molecule E-cadherin. *Cellular and Molecular Life Sciences: CMLS* 65, 3756–3788.
- Vasioukhin, V., Bauer, C., Yin, M., Fuchs, E., 2000. Directed actin polymerization is the driving force for epithelial cell-cell adhesion. *Cell* 100, 209–219.
- Verma, S., Han, S.P., Michael, M., *et al.*, 2012. A WAVE2-Arp2/3 actin nucleator apparatus supports junctional tension at the epithelial zonula adherens. *Molecular Biology of the Cell* 23, 4601–4610.
- Wang, Z., Sandiford, S., Wu, C., Li, S.S., 2009. Numb regulates cell-cell adhesion and polarity in response to tyrosine kinase signalling. *EMBO Journal* 28, 2360–2373.
- Weber, K.L., Fischer, R.S., Fowler, V.M., 2007. Tmod3 regulates polarized epithelial cell morphology. *Journal of Cell Science* 120, 3625–3632.
- Wu, Y., Kanchanawong, P., Zaidel-Bar, R., 2015. Actin-delimited adhesion-independent clustering of E-cadherin forms the nanoscale building blocks of adherens junctions. *Developmental Cell* 32 (2), 139–154.
- Wu, Y., Vendome, J., Shapiro, L., Ben-Shaul, A., Honig, B., 2011. Transforming binding affinities from three dimensions to two with application to cadherin clustering. *Nature* 475, 510–513.
- Xu, Y., Carpenter, G., 1999. Identification of cadherin tyrosine residues that are phosphorylated and mediate Shc association. *Journal of Cellular Biochemistry* 75, 264–271.
- Yamada, S., Nelson, W.J., 2007. Localized zones of Rho and Rac activities drive initiation and expansion of epithelial cell-cell adhesion. *Journal of Cell Biology* 178, 517–527.
- Yonemura, S., 2011. Cadherin-actin interactions at adherens junctions. *Current Opinion in Cell Biology* 23, 515–522.
- Yonemura, S., Itoh, M., Nagafuchi, A., Tsukita, S., 1995. Cell-to-cell adherens junction formation and actin filament organization: Similarities and differences between non-polarized fibroblasts and polarized epithelial cells. *Journal of Cell Science* 108 (Pt 1), 127–142.
- Yonemura, S., Wada, Y., Watanabe, T., Nagafuchi, A., Shibata, M., 2010. alpha-Catenin as a tension transducer that induces adherens junction development. *Nature Cell Biology* 12, 533–542.
- Yoshida-Noro, C., Suzuki, N., Takeichi, M., 1984. Molecular nature of the calcium-dependent cell-cell adhesion system in mouse teratocarcinoma and embryonic cells studied with a monoclonal antibody. *Developmental Biology* 101, 19–27.
- Zaidel-Bar, R., 2013. Cadherin adhesome at a glance. *Journal of Cell Science* 126, 373–378.
- Zhou, Y.N., Xu, C.P., Chen, Y., *et al.*, 2005. Alpha-catenin expression is decreased in patients with gastric carcinoma. *World Journal of Gastroenterology: WJG* 11, 3468–3472.

Extracellular Regulation of Cell-to-Matrix Adhesion

PA Janmey, University of Pennsylvania, Philadelphia, PA, USA

CA McCulloch, University of Toronto, Toronto, ON, Canada

RT Miller, University of Texas Southwestern Medical School, Dallas, TX, USA and Dallas VA Medical Center, Dallas, TX, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Connective tissue cells and matrices are distributed throughout the bodies of mammals and other metazoans and play key roles in organizing and supporting tissues and organs. Connective tissue matrices provide continuity with other cells and transduce chemical and physical stimuli that regulate cell activity. While extracellular matrices may appear to be inert, many matrices are highly dynamic and undergo remodeling as a result of cell-mediated synthesis and remodeling. The remodeling of the connective tissue matrix is essential for preserving healthy tissues and organs throughout life. Disturbances of matrix remodeling are involved in a vast array of inflammatory, neoplastic, and atrophic diseases (McCarty *et al.*, 2012; Hopps and Caimi, 2012; Eckhouse and Spinale, 2011; Lu *et al.*, 2011) which arise in part from the diverse structure and function of matrix molecules and the complex behavior of connective tissue cells that are responsible for the remodeling of the matrix.

When seen from the top, a cell that is deposited from a three-dimensional suspension onto a flat surface can bind and spread along the surface in a manner that can closely resemble that of a liquid drop wetting a surface (Sackmann and Bruinsma, 2002) or a phospholipid vesicle flattening on a surface to which it adheres (Sackmann and Smith, 2014; Bell and Torney, 1985). In all three cases objects that are approximately spherical in three-dimensional suspension become flattened as they adhere to the surface, but eventually reach a finite area that is determined by a balance of forces that reach equilibrium or at least a steady state when the adhesive energy at the interface becomes comparable to the mechanical work required to deform a viscoelastic cell or a lipid droplet with surface tension (Sackmann and Smith, 2014). In some

cases of single cells or even spherical aggregates of cells adhering to appropriate surfaces, the rates and extents to which the cells adhere and spread are similar to those of viscoelastic liquid drops wetting a surface (Doeznan *et al.*, 2011), but there are also many structural and thermodynamic differences between a cell or a cell aggregate and a simple liquid drop (Beaune *et al.*, 2014). One important difference is that even when cells adhere to molecularly flat surfaces the adhesive interface is not flat and continuous but rather is formed by small discrete regions of the basal cell membrane that may be in close contact with the substrate, leaving the rest of the cell membrane unbound to the substrate and often fluctuating distances of several hundred nanometers above the surface, as has been well documented by reflection interference microscopy (Curtis, 1964), and other methods. A second important difference between microscopic cells and macroscopic liquid drops is that the cell membrane is small and soft enough to undergo significant thermal fluctuations as well as molecular motor-driven nonthermal motions whereas a macroscopic liquid drop surface is usually too large to be governed by factors other than its surface tension. A further essential difference is that the chemistry of drops wetting the surface is often driven by a uniform set of interactions between molecular structures that are very closely spaced, and can often be reasonably modeled as continuous interfaces, whereas cell–substrate or cell–matrix interactions are dictated largely by highly specific interactions between a well-defined set of transmembrane proteins and their unique set of ligands, as discussed in examples cited below, and these specific receptor–ligand bonds are often spaced many nanometers apart even in those regions of the cell membrane that make close contact with the surface (Geiger *et al.*, 2009). A summary of the opposing interactions that control the cell–substrate interface is shown in Figure 1.

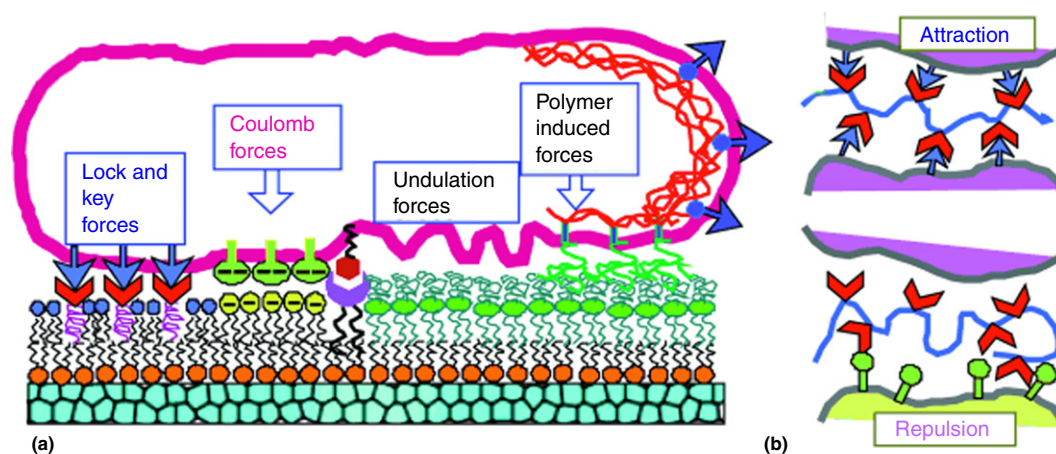


Figure 1 Summary of biochemical and mechanical effects mediating cell adhesion to surfaces. Reproduced with permission from Sackmann, E., Bruinsma, R.F., 2002. Cell adhesion as wetting transition? *ChemPhysChem* 3, 262–269.

Early Studies of Cell Adhesion

Pioneering studies of cell adhesion emphasized the role of substrate surface chemistry or topography on the ability of cells to adhere to them. Many early studies used hydrogels rather than glass or plastic surfaces as the substrates because the hydrogels formed either from native or synthetic biopolymers are better mimics of the extracellular matrix *in vivo*. Such studies, employing human tumor-derived HeLa cells and primary human fibroblasts established that these mammalian cells would not bind to alginate, agar, or silica gels unless they had been treated with human serum, but they bound to gels formed by collagen 1 or fibrin in the absence of serum, and that fibrinogen also was sufficient to transform alginate, agar, or silicate gels from nonadhesive to adhesive (Weiss, 1959). Such studies suggested that specific interactions between cell surface molecules and their ligands on the substrate were important for mediating adhesion. Later, studies showed that coating nonadhesive surfaces with polyvalent cations such as protamine (Steinhardt *et al.*, 1971) or polylysine (Maceira-Coelho and Avrameas, 1972; Mazia *et al.*, 1975) strongly facilitated attachment of cells to surfaces, by mechanisms that depended on electrostatic interaction between the positively charged surface and negatively charged polymers, such as those of the glycocalyx, that extend from the cell membrane.

Specific versus Generic Attractive and Repulsive Interactions

A great deal has been learned about the molecular biology of transmembrane adhesion receptors such as integrins, selectins, CD44, ICAM 1, and other protein complexes along with the biochemistry of their interactions with extracellular proteins such as fibronectin, collagen, fibrin, hyaluronic acid, chondroitins, and many other ligands (Heisenberg and Fassler, 2012). The fact that cell adhesion on specific surfaces can depend essentially on docking to ligands that contain the sequence RGD (arginine-glycine-aspartate) (Ruoslahti and Pierschbacher, 1987) might suggest that cell adhesion is entirely determined by the functions of these proteins and the distribution of their ligands, both of which are strictly regulated by the genetics and signaling mechanisms of the cells that produce them. In such a case resemblance of cell adhesion and surface wetting would appear to be coincidental, but there are other molecular and physical features of cell adhesion that suggest that much more is to be learned about the physics of cell-matrix interactions to accompany the wealth of information about its biochemistry.

Even when cell adhesion requires the function of a specific transmembrane protein such as an integrin, and cells detach immediately when this protein is degraded, the interface between the cell and substrate is defined by many other macromolecules at the cell surface and the substrate that modify the function of the essential adhesion protein or the morphology of the cell. The affinity and dynamics of the adhesive protein-substrate interaction can determine the strength of the binding but not necessarily the morphology of the cell or the stability of the interface. For example, many cell membranes also express membrane proteins with very large

polypeptide or carbohydrate chains that extend hundreds of nanometers or even several microns away from the lipid bilayer to form a so-called glycocalyx or pericellular matrix, and similar glycosaminoglycans form much of the extracellular matrix *in vivo* (Fotia *et al.*, 2013). The pericellular matrix is in some sense anti-adhesive because in order for the cell membrane to make close contact with a flat substrate the pericellular matrix molecules must be compressed to allow the adhesive transmembrane protein (integrin, selectin, etc.) to reach its matrix target. Recent studies have shown that the entropic cost of compressing the pericellular matrix as well as the osmotic pressure produced as the charged polyelectrolytes in the glycocalyx become concentrated along with their counterions can alter the nature of the bond between an integrin and its ligand. These effects can either stabilize the adhesion site or destabilize it depending on whether the integrin-ligand interaction contains slip or catch bonds (Fotia *et al.*, 2013; Paszek *et al.*, 2014). A further complication in cell adhesion is that the bound interface is not a passive material but rather, as adhesion proteins bind their matrix ligands, factors bound to the intracellular domains of these proteins stimulate the assembly of actin-based cytoskeletal fibers, activate myosin motors through small GTPase signals, activate enzymes that synthesize polyphosphoinositides at the site of the surface, and thereby enhance membrane fluctuations and contractile forces specifically at the site where transmembrane proteins engage the matrix (Opas, 1995; Harris *et al.*, 1980). The work done at these adhesion sites has a strong influence on the nature of the adhesive substrate that can in some cases resemble changes in surface tension at the surface of liquid drops, but that arises from different molecular mechanisms.

Stiffness, Spacing, and Curvature

The adhesive interface between cells and surfaces is influenced by the physical properties of the interface as much as by the chemical interactions between cells and substrates. Much recent work has documented the importance of physical features such as surface topography and elastic modulus on the ability of cells to bind and spread on mechanically different surfaces that have the same surface chemistry and density of ligands for transmembrane receptors. The stiffness of the planar surface of the three-dimensional matrix has a very strong affect on the adhesive interface, extent of cell spreading, and many aspects of cells, including their rate of motility, proliferation, and differentiation. Usually more rigid surfaces promote greater rates of cell spreading and larger interfaces, and promote cell proliferation, but there are significant exceptions especially in cell types such as central nervous system neurons that are derived from very soft tissues and extend their adhesive interfaces more robustly in soft materials that mimic the elastic modulus of the brain (Discher *et al.*, 2005). Surface topography can also strongly affect cell adhesion (Kim *et al.*, 2012). For example, the spacing of ligands for integrins determines both the initial attachment and eventual morphology of the adherent cell. If integrin ligands are spaced in ordered arrays at distances below a critical value (approximately 50 nm), cells adhere to such surfaces as they do on surfaces of contiguous ligand distribution, but if the ligand spacing is greater, cells

can initially attach, but their long-term spreading and eventual phenotype can vary widely, and can lead to cell detachment and death if the adhesive spacing is too large (Cavalcanti-Adam *et al.*, 2006). Curvature can also be a strong determinant of cell adhesion in a highly celltype-specific manner. For example, when non-transformed IAR-2 epithelial cells were cultured on 32- μ m-diameter glass fibers, they oriented their actin-based cytoskeletal fibers and fibronectin-based adhesion sites longitudinally along the axis of the fibers but the majority of ras-transformed cells from the same source oriented their actin fibers and adhesion sites transversely around the fiber (Levina *et al.*, 1996). Similarly, neurons extend processes and move along the axis of fibers in a fiber diameter and stiffness-dependent manner with optimal movement at curvatures in the range of axons (Smeal *et al.*, 2005). Even on flat substrates, curved edges of patterned adhesive protein influence the formation of lamellipodia and cytoskeletal assembly (James *et al.*, 2008). Here the mechanics of membrane and cytoskeletal bending rather than strength of adhesion receptors appear to dominate the morphology of the adhesive interface (Biton and Safran, 2009).

How Extracellular Signals Alter Adhesion to Matrix

Adhesive cell-matrix interactions are governed primarily by the activity of integrins, heterodimeric proteins, composed of α and β subunits that connect the cytoskeleton to components of the extracellular matrix. The activity of integrins is determined by the affinity with which they bind their ligands, and the strength of adhesion between a cell and the matrix is further determined by the number of engaged integrins on the cell surface. The activity of integrins responds to signaling pathways controlled by membrane and cytosolic signaling proteins (inside-out signaling), and integrins transmit information from the outside of cells in response to binding ligands (outside-in signaling) (Pouwels *et al.*, 2012). Integrins can exist in high- or low-affinity states for their extracellular ligands, and the inability to switch affinity states results in a range of diseases. This section will consider control of integrin function by extracellular signals. *In vivo*, inside-out, and outside-in signaling are difficult to separate because a cell's responses to its physical and chemical environment are mutually dependent, dynamic, and iterative.

Different combinations of the 18 α and eight β subunits result in 24 different integrins with binding specificities for different extracellular ligands such as collagen, laminin, fibronectin, and fibrinogen. Integrin expression is cell type specific resulting in cell type-specific interactions with matrix. Both subunits have single transmembrane domains and highly glycosylated extracellular domains that bind the ligand. The short intracellular domain of the β subunits interacts with a collection of proteins, some of which, like talin, filamin, and α -actinin, link the β subunit to the actin cytoskeleton. Proteins that interact with the C-termini of α subunits are less well understood, but include SHARPIN, MDGI (mammary-derived growth inhibitor), and RhoGAPs. The α and β subunit C-termini bind common partners and subunit-specific cytoplasmic partners with affinities that depend on the specific subunit. Additionally, talin and the Itg β subunit form the

nidus for aggregation of 50 or more proteins with structural and signaling functions, forming focal adhesions.

Talin appears to be the central protein in signaling from the cell to integrins and signaling from the matrix via integrins to the cell. Talin1 and Talin2 are related genes coding for similar proteins with similar functions and different patterns of expression. Talin is primarily present in the cytosol in an auto-inhibited conformation in which the rod domain blocks the F2-F3 regions of the head domain preventing membrane association (Das *et al.*, 2014; Song *et al.*, 2012). In response to extracellular signals that may include toll-like receptors, G protein-coupled receptors, receptor tyrosine kinases, and cytokine receptors, and activation of intracellular cell signaling pathways that can involve PIP₂, PKC α , and particularly Rap1, talin localizes to the plasma membrane at the site of integrin C-termini (Chung *et al.*, 2014). Since both Rap1 and PI3-kinase are downstream of cAMP (cyclic adenosine monophosphate) signaling, numerous extracellular signals are capable of activating this pathway. The Rap1 effector, RIAM interacts directly with talin via a high-affinity site and recruits it to the plasma membrane (Chang *et al.*, 2014; Calderwood *et al.*, 2013). In platelets, talin and kindlin are bound by ADAP (adhesion and degranulation promoting adapter protein), a membrane-associated scaffolding protein that localizes them near the β integrin C-terminus (Kasirer-Friede *et al.*, 2014). Kindlin then binds to the β integrin-talin complex reinforcing it. Talin exists as a dimer and so can initiate integrin clustering (Das *et al.*, 2014).

In the absence of stimulation, integrins exist in low-affinity, folded states where their affinity for ligands is low. Binding of the talin head domain to the C-terminus of β subunits induces a conformational change in the β subunit that appears to originate with separation of the α and β transmembrane domains and straightening of the β extracellular domain resulting in changes in the conformation of the N-terminus and an intermediate affinity binding state. A high-affinity state is produced by traction force on the integrin complex (Schurpf and Springer, 2011). Further adhesion to matrix results from integrin clustering (above) and traction as well as the binding of additional proteins to adhesion complexes.

Following initial or nascent complex formation, additional proteins such as vinculin and paxillin bind to the complex with increased binding in the presence of mechanical stress (Oakes and Gardel, 2014). Many of the proteins that join adhesion complexes following their initial formation are signaling proteins such as FAK, other Src family kinases, phosphatases, scaffolding proteins such as zyxin and p130CAS, motor proteins such as non-muscle myosins, or actin polymerizing complexes such as Arp2/3 or formins. Many of the proteins in the complex are phospho-acceptors (targets of kinases and phosphatases) or kinases and phosphatases themselves. Adhesion complexes both generate signals and respond to a range of signals from the cell, many originating as extracellular signals. The scenario described above is generalized and simplified, but is consistent with the current literature (Das *et al.*, 2014; Schurpf and Springer, 2011; Delon and Brown, 2007). The details of what proteins bind and in what order will evolve with new experimental systems and approaches, and may also depend on the cell type and specific integrins.

Cell Adhesions and Matrix Molecules in Connective Tissue Function

The synthesis and degradation of the extracellular matrix are coordinated to maintain connective tissue homeostasis (Hinz, 2013). Collagen is the most abundant molecule of mammals (Perez-Tamayo, 1978) and exerts dominant influences on the chemical and mechanical properties of connective tissues. The remodeling of collagen requires the synthesis of new molecules and the degradation of worn-out or effete molecules by collagenolysis. Because collagen is highly insoluble, specialized types of degradative systems have evolved to enable collagenolysis, such as the matrix metalloproteinase family of endopeptidases (Steffensen *et al.*, 2001; Murphy and Nagase, 2008). Another important but less well-understood system for collagen degradation is intracellular collagen degradation by phagocytic fibroblasts, which is critical for physiological matrix remodeling and wound healing (Everts *et al.*, 1996; Ten Cate, 1972). In collagen phagocytosis, fibroblasts initially adhere to collagen, followed by partial digestion of collagen fibrils extracellularly (Lee *et al.*, 2006); shorter segments are then internalized and digested intracellularly by lysosomal cathepsins (Burleigh *et al.*, 1974; Song *et al.*, 2006; Van Noorden and Everts, 1991; Nallaseeth *et al.*, 2013). In tissues with rapid matrix remodeling including the involuting uterus (Dyer and Peppler, 1977), periodontal tissues (Ten Cate, 1972), and healing wounds (McGaw and Ten Cate, 1983), collagen turnover is directly related to the volume density of collagen fibrils in phagosomes within fibroblasts (Everts *et al.*, 1996), indicating that phagocytosis is critical for collagen turnover (Everts *et al.*, 1996, 1995).

Remodeling of the extracellular matrix and the maintenance of connective tissue structures is highly dependent on the attachment of cells to matrix proteins. These attachments are important not just for the preservation of cell vitality and proliferation, but also for physiological cell migration and the degradation of collagen, which is the most abundant matrix protein of mammalian connective tissues (Perez-Tamayo, 1978). The extracellular matrix exerts multiple influences on tissue physiology, including dominant contributions to mechanical properties, reactivity to environmental changes, and the ability to regulate signaling processes. In particular, the ability of matrix interactions to influence tissue physiology is dependent on cell adhesions to the matrix, which can influence the macromolecular architecture of secreted matrix components. Conversely, matrix molecules can influence the organization of cell adhesions, which in turn propagate signals into cells (Kim *et al.*, 2012). An important element of this reciprocal signaling system is the mechanical continuity between the extracellular matrix, cell surface adhesion receptors such as integrins, and the cytoskeleton (Ingber, 1997; Wang *et al.*, 1993). The continuity of these structures is an important structural element that enables chemical and mechanical signaling, which together regulate many processes that affect matrix structure and function. At specialized attachment plaques known as focal adhesions, integrins, actin filaments, and a large group of actin-binding proteins are assembled into structures that mediate cell attachment and signal transduction (Schoenwaelder and Burridge, 1999).

Cell attachment to matrix proteins is mediated by a large variety of adhesion molecules, including integrins

(Covell, 1990). Integrins bind extracellular matrix proteins with considerable ligand specificity. For example, collagen binds preferentially to the $\alpha_2\beta_1$, $\alpha_1\beta_1$, $\alpha_{10}\beta_1$, or $\alpha_{11}\beta_1$ integrins. In connective tissues in which there are abundant collagen fibers, mechanical forces often play a central role in tissue function, such as the cardiac interstitium. In these tissues, the integrins expressed by resident myocytes and fibroblasts provide bidirectional signaling between cells and fibrillar extracellular matrix proteins. These types of signaling processes are common in many types of cells but are particularly evident in tissues where high-amplitude physical forces are applied (Hierck *et al.*, 1996). Force-related adhesion signals can promote rapid matrix remodeling (Schorb *et al.*, 1994) and abundant collagen turnover. Because integrins physically associate with signaling and actin-binding proteins such as talin, integrins provide adhesive and mechanosensory functions (Katsumi *et al.*, 2004) through focal adhesions.

Focal adhesions are actin-enriched adhesive domains found in many cultured cells including fibroblasts, endothelial cells, synovial cells, and chondrocytes (Critchley, 2000). Focal adhesions form at sites of clustered integrin receptors bound to extracellular matrix ligands. They are dynamic structures that undergo distinct maturational steps with distinct molecular composition (Wehrle-Haller and Imhof, 2002). Once thought to be strictly structural adhesive complexes, focal adhesions are now recognized for their role in signal transduction (Burridge and Chrzanowska-Wodnicka, 1996; Wehrle-Haller and Imhof, 2002). Over 50 signaling molecules are detected in nascent adhesions (Aplin and Juliano, 1999; Garrington and Johnson, 1999) one of which is the focal adhesion kinase, a 125 kDa tyrosine kinase that plays important roles in cell adhesion, contractility, migration, and signaling (Schaller, 2010). The importance of the focal adhesion kinase in adhesion-related functions is demonstrated by the behavior of focal adhesion null fibroblasts, which are highly adherent and form large, stable adhesions (Ilic *et al.*, 1995). Integrin ligation of matrix molecules like collagen or fibronectin induces auto-phosphorylation of the focal adhesion kinase, which creates binding sites and enables recruitment of other signaling proteins at focal adhesions.

Role of Cell-Matrix Adhesions in Pathological Processes: Fibrosis

As extracellular matrices in many tissues undergo continuous turnover and replacement during growth, development, and function in adults, a failure to control matrix dynamics can lead to excessive and inappropriate matrix formation or tissue destruction. These disorders of matrix biology are hallmarks of various neoplastic and fibrotic diseases (e.g., congestive heart failure (Shah and Mann, 2011)) and high prevalence inflammatory diseases such as arthritis and periodontal diseases. Cells in mechanically active environments including joint, myocardium, lung, gastrointestinal tract, and periodontium are subjected to high-amplitude exogenous forces. Under conditions of chronic high load, mechanically induced cell death can promote fibrosis, cell death, and loss of organ and tissue function as is seen in arthritis, heart failure, and periodontal diseases.

Commonly used drugs for management of hypertension, epilepsy, and organ transplantation (e.g., nifedipine, dilantin, cyclosporin A, respectively) disrupt calcium signaling and inhibit the normal adhesion of cells to matrix molecules. This disruption leads to blockade of intracellular collagen degradation in fibroblasts (McCulloch, 2004) and to significant fibrous overgrowth of periodontal tissues, a major clinical management problem in periodontology. There is no effective treatment for these disorders and there is significant morbidity (Mavrogiannis *et al.*, 2006). Disruption of cell–matrix adhesions in collagen phagocytosis is thus a critical pathway in maintaining the health and function of oral connective tissues.

In several cardiac diseases such as pressure overload, diabetic cardiomyopathy, and myocardial infarction, fibrosis is a common disorder of myocardial extracellular matrix structure and function. The clinical significance of fibrosis is that accumulation of disorganized fibrillar collagen in the cardiac interstitium can inhibit diastolic and systolic function. Fibrosis is mediated by several different cellular and extracellular processes including disruptions of fibroblast differentiation, perturbations of posttranslational processing and assembly of matrix molecules, and inappropriately organized matrix degradation by proteases and intracellular digestion. The enlargement of transformed fibroblast and myofibroblast populations in the diseased cardiac interstitium plays a critical role in the disorganized matrix remodeling that occurs after pressure overload or diabetes because these cells do not process and remodel interstitial collagen in a physiological fashion.

In addition to alterations of matrix structure and function as a result of fibrotic processes, the attachment of cells to extracellular matrix proteins can also regulate pro-inflammatory signaling that leads to matrix destruction. Interleukin-1 is a potent, pro-inflammatory cytokine that contributes to inflammatory tissue injury in lung fibrosis, arthritis, and periodontitis. In fibroblasts and chondrocytes, IL-1-induced signaling requires the formation of focal adhesions and the tight adaptation of cells to the underlying matrix (Arora *et al.*, 1995). Focal adhesions in these cells are enriched with IL-1 signaling receptors, and possibly other critical signaling elements of the IL-1 signaling pathway, such as SHP-2 and PTP α , protein tyrosine phosphatases that are found in focal adhesions and that play critical roles in IL-1-induced focal adhesion maturation, function, and signaling to enable matrix degradation (Herrera Abreu *et al.*, 2006).

Role of Cell–Matrix Adhesions in Pathological Processes: Platelets and Thrombosis

Platelet activation provides useful examples of physiology and pathophysiology in extracellular signaling and integrin activation. Reduced platelet activation results in inappropriate bleeding, and increased platelet activation results in thrombus formation and ischemic injury to organs, commonly the heart (myocardial infarction), brain (strokes), limbs, and intestines. Platelets express an integrin, $\alpha\text{IIb}/\beta 3$, that is a receptor for fibrinogen, the GPIb-IX-V complex that is a receptor for von Willebrand factor, and multiple G protein-coupled receptors for compounds including thrombin, adenosine diphosphate (ADP), and thromboxane A₂ (TxA₂), all receptors

that participate in platelet activation (Shattil *et al.*, 1998). Platelets are maintained in an inactive state in the circulation by intact endothelium, endothelial production of PGI₂, NO, and an ADP degrading enzyme (ADP activates platelets), along with the absence of platelet activators. Activated platelets are characterized by high-affinity binding of fibrinogen by $\alpha\text{IIb}/\beta 3$ with $\alpha\text{IIb}/\beta 3$ clustering and link age of multiple platelets through fibrin polymers, cytoskeletal rearrangements so that platelets spread, and degranulate to release substances that include ADP and von Willebrand factor, and increase surface expression of $\alpha\text{IIb}/\beta 3$. Platelets are normally activated in the presence of tissue injury with endothelial disruption and loss of activation inhibitors, exposure of the von Willebrand factor that binds its receptor and slows circulating platelets, and release of ADP, thrombin, and TxA₂ as well as binding of fibrinogen or collagen to $\alpha\text{IIb}/\beta 3$. However, binding of fibrinogen and collagen to $\alpha\text{IIb}/\beta 3$ alone is insufficient to activate platelets and lead to activation of $\alpha\text{IIb}/\beta 3$ with high-affinity binding of fibrinogen. Full platelet activation requires collaboration of the GPIb-IX-V complex and G protein-coupled receptors such as those for ADP. Mutations in the GPIb-IX-V complex, G protein-coupled receptors, or their signaling pathways result in platelet abnormalities with medical importance (Van Geet *et al.*, 2009). For example, reduced expression of von Willebrand factor (congenital or acquired), or loss of function mutations in its receptor, the GPIb-IX-V complex (Bernard–Soulier syndrome or Velocardiofacial syndrome (VCFS), all of which reduce platelet–matrix binding lead to increased bleeding risk.

Conclusion

Cell adhesion to extracellular matrices is an important regulator of cell function. Changes in the chemical or physical features of the matrix can have as great an impact on cell function as do biochemically and genetically controlled changes in the expression of specific proteins by which cells bind their matrix. In many cases selectivity and affinity of specific cells to specific matrices or substrates depends on expression of a small number of transmembrane adhesion proteins and their matrix ligands but many other more generic effects such as surface topography, ligand spacing, electrostatic charge, and mechanical compliance of the matrix also have significant effects on cell adhesion that can lead to changes during normal tissue development or pathology.

See also: Cell Communication: Cellular Machineries: The Molecular Architecture of Cell–Cell Adhesions

References

- Aplin, A.E., Juliano, R.L., 1999. Integrin and cytoskeletal regulation of growth factor signaling to the map kinase pathway. *Journal of Cell Science* 112 (Pt 5), 695–706.
- Arora, P.D., Ma, J., Min, W., Cruz, T., McCulloch, C.A., 1995. Interleukin-1-induced calcium flux in human fibroblasts is mediated through focal adhesions. *Journal of Biological Chemistry* 270, 6042–6049.

- Beaune, G., Stirbat, T.V., Khalifat, N., *et al.*, 2014. How cells flow in the spreading of cellular aggregates. *Proceedings of the National Academy of Sciences of the United States of America* 111, 8055–8060.
- Bell, G.I., Torney, D.C., 1985. On the adhesion of vesicles by cell-adhesion molecules. *Biophysical Journal* 48, 939–947.
- Biton, Y.Y., Safran, S.A., 2009. The cellular response to curvature-induced stress. *Physical Biology* 6, 046010.
- Burleigh, M.C., Barrett, A.J., Lazarus, G.S., 1974. Cathepsin B1. A lysosomal enzyme that degrades native collagen. *Biochemical Journal* 137, 387–398.
- Burridge, K., Chrzanoska-Wodnicka, M., 1996. Focal adhesions, contractility, and signaling. *Annual Review of Cell and Developmental Biology* 12, 463–518.
- Calderwood, D.A., Campbell, I.D., Critchley, D.R., 2013. Talins and kindlins: Partners in integrin-mediated adhesion. *Nature Reviews Molecular Cell Biology* 14, 503–517.
- Cavalcanti-Adam, E.A., Micoulet, A., Blummel, J., *et al.*, 2006. Lateral spacing of integrin ligands influences cell spreading and focal adhesion assembly. *European Journal of Cell Biology* 85, 219–224.
- Chang, Y.C., Zhang, H., Franco-Barraza, J., *et al.*, 2014. Structural and mechanistic insights into the recruitment of talin by riam in integrin signaling. *Structure* 22, 1810–1820.
- Chung, K.J., Mitroulis, I., Wiessner, J.R., *et al.*, 2014. A novel pathway of rapid tir-triggered activation of integrin-dependent leukocyte adhesion that requires rap1 gtpase. *Molecular Biology of the Cell* 25, 2948–2955.
- Covell, J.W., 1990. *Cardiac Myocyte Connective Tissue Interactions in Health and Disease*, vol. 13. Basel: Karger, pp. 99–112.
- Critchley, D.R., 2000. Focal adhesions — The cytoskeletal connection. *Current Opinion in Cell Biology* 12, 133–139.
- Curtis, A.S., 1964. The mechanism of adhesion of cells to glass. A study by interference reflection microscopy. *Journal of Cell Biology* 20, 199–215.
- Das, M., Subbaya Ithychanda, S., Qin, J., Plow, E.F., 2014. Mechanisms of talin-dependent integrin signaling and crosstalk. *Biochimica et Biophysica Acta* 1838, 579–588.
- Delon, I., Brown, N.H., 2007. Integrins and the actin cytoskeleton. *Current Opinion in Cell Biology* 19, 43–50.
- Discher, D.E., Janmey, P., Wang, Y.L., 2005. Tissue cells feel and respond to the stiffness of their substrate. *Science* 310, 1139–1143.
- Douezan, S., Guevorkian, K., Naouar, R., *et al.*, 2011. Spreading dynamics and wetting transition of cellular aggregates. *Proceedings of the National Academy of Sciences of the United States of America* 108, 7315–7320.
- Dyer, R.F., Pepler, R.D., 1977. Intracellular collagen in the nonpregnant and iud-containing rat uterus. *Anatomical Record* 187, 241–247.
- Eckhouse, S.R., Spinale, F.G., 2011. Changes in the myocardial interstitium and contribution to the progression of heart failure. *Heart Failure Clinics* 8, 7–20.
- Everts, V., Korper, W., Niehof, A., Jansen, I., Beertsen, W., 1995. Type vi collagen is phagocytosed by fibroblasts and digested in the lysosomal apparatus: Involvement of collagenase, serine proteinases and lysosomal enzymes. *Matrix Biology* 14, 665–676.
- Everts, V., van der Zee, E., Creemers, L., Beertsen, W., 1996. Phagocytosis and intracellular digestion of collagen, its role in turnover and remodelling. *Histochemical Journal* 28, 229–245.
- Fotia, C., Messina, G.M., Marletta, G., Baldini, N., Ciapetti, G., 2013. Hyaluronan-based pericellular matrix: Substrate electrostatic charges and early cell adhesion events. *European Cells & Materials* 26, 133–149 (discussion 149).
- Garrington, T.P., Johnson, G.L., 1999. Organization and regulation of mitogen-activated protein kinase signaling pathways. *Current Opinion in Cell Biology* 11, 211–218.
- Geiger, B., Spatz, J.P., Bershadsky, A.D., 2009. Environmental sensing through focal adhesions. *Nature Reviews Molecular Cell Biology* 10, 21–33.
- Harris, A.K., Wild, P., Stopak, D., 1980. Silicone rubber substrata: A new wrinkle in the study of cell locomotion. *Science* 208, 177–179.
- Heisenberg, C.P., Fassler, R., 2012. Cell-cell adhesion and extracellular matrix: Diversity counts. *Current Opinion in Cell Biology* 24, 559–561.
- Herrera Abreu, M.T., Wang, Q., Vachon, E., *et al.*, 2006. Tyrosine phosphatase shp-2 regulates il-1 signaling in fibroblasts through focal adhesions. *Journal of Cellular Physiology* 207, 132–143.
- Hierck, B.P., Poelmann, R.E., Vaniperen, L., Brouwer, A., Gittenberger-de Groot, A.C., 1996. Differential expression of alpha-6 and other subunits of laminin binding integrins during development of the murine heart. *Developmental Dynamics* 206, 100–111.
- Hinz, B., 2013. Matrix mechanics and regulation of the fibroblast phenotype. *Periodontol* 2000 (63), 14–28.
- Hopps, E., Caimi, G., 2012. Matrix metalloproteinases in metabolic syndrome. *European Journal of Internal Medicine* 23, 99–104.
- Ilic, D., Furuta, Y., Kanazawa, S., *et al.*, 1995. Reduced cell motility and enhanced focal adhesion contact formation in cells from fak-deficient mice. *Nature* 377, 539–544.
- Ingber, D.E., 1997. Tensegrity: The architectural basis of cellular mechanotransduction. *Annual Review of Physiology* 59, 575–599.
- James, J., Goluch, E.D., Hu, H., Liu, C., Mrksich, M., 2008. Subcellular curvature at the perimeter of micropatterned cells influences lamellipodial distribution and cell polarity. *Cell Motility and the Cytoskeleton* 65, 841–852.
- Kasirer-Friede, A., Kang, J., Kahner, B., *et al.*, 2014. Adap interactions with talin and kindlin promote platelet integrin alpaIIb beta3 activation and stable fibrinogen binding. *Blood* 123, 3156–3165.
- Katsumi, A., Orr, A.W., Tzima, E., Schwartz, M.A., 2004. Integrins in mechanotransduction. *Journal of Biological Chemistry* 279, 12001–12004.
- Kim, D.H., Provenzano, P.P., Smith, C.L., Levchenko, A., 2012. Matrix nanotopography as a regulator of cell function. *Journal of Cell Biology* 197, 351–360.
- Lee, H., Overall, C.M., McCulloch, C.A., Sodek, J., 2006. A critical role for the membrane-type 1 matrix metalloproteinase in collagen phagocytosis. *Molecular Biology of the Cell* 17, 4812–4826.
- Levin, E.M., Domina, L.V., Rovinsky, Y.A., Vasiliev, J.M., 1996. Cylindrical substratum induces different patterns of actin microfilament bundles in nontransformed and in ras-transformed epitheliocytes. *Experimental Cell Research* 229, 159–165.
- Lu, P., Takai, K., Weaver, V.M., Werb, Z., 2011. Extracellular matrix degradation and remodeling in development and disease. *Cold Spring Harbor Perspectives in Biology* 3 (12). doi:10.1101/cshperspect.a005058.
- Maceera-Coelho, A., Avrameas, S., 1972. Modulation of cell behavior by the substratum in fibroblastic and leukemic cell lines. *Proceedings of the National Academy of Sciences of the United States of America* 69, 2469–2473.
- Mavrogianis, M., Ellis, J.S., Thomason, J.M., Seymour, R.A., 2006. The management of drug-induced gingival overgrowth. *Journal of Clinical Periodontology* 33, 434–439.
- Mazia, D., Schatten, G., Sale, W., 1975. Adhesion of cells to surfaces coated with polylysine. Applications to electron-microscopy. *Journal of Cell Biology* 66, 198–200.
- McCarthy, S.M., Cochrane, C.A., Clegg, P.D., Percival, S.L., 2012. The role of endogenous and exogenous enzymes in chronic wounds: A focus on the implications of aberrant levels of both host and bacterial proteases in wound healing. *Wound Repair and Regeneration* 20, 125–136.
- McCulloch, C.A., 2004. Drug-induced fibrosis: Interference with the intracellular collagen degradation pathway. *Current Opinion in Drug Discovery & Development* 7, 720–724.
- McGaw, W.T., Ten Cate, A.R., 1983. A role for collagen phagocytosis by fibroblasts in scar remodeling: An ultrastructural stereologic study. *Journal of Investigative Dermatology* 81, 375–378.
- Murphy, G., Nagase, H., 2008. Progress in matrix metalloproteinase research. *Molecular Aspects of Medicine* 29, 290–308.
- Nallasetti, F.S., Lecaille, F., Li, Z., Bromme, D., 2013. The role of basic amino acid surface clusters on the collagenase activity of cathepsin k. *Biochemistry* 52, 7742–7752.
- Oakes, P.W., Gardel, M.L., 2014. Stressing the limits of focal adhesion mechanosensitivity. *Current Opinion in Cell Biology* 30, 68–73.
- Opas, M., 1995. Cellular adhesiveness, contractility, and traction: Stick, grip, and slip control. *Biochemistry and Cell Biology* 73, 311–316.
- Paszek, M.J., Dufort, C.C., Rossier, O., *et al.*, 2014. The cancer glycocalyx mechanically primes integrin-mediated growth and survival. *Nature* 511, 319–325.
- Perez-Tamayo, R., 1978. Pathology of collagen degradation. A review. *American Journal of Pathology* 92, 508–566.
- Pouwels, J., Nevo, J., Pellinen, T., Ylanne, J., Ivaska, J., 2012. Negative regulators of integrin activity. *Journal of Cell Science* 125, 3271–3280.
- Ruoslahti, E., Pierschbacher, M.D., 1987. New perspectives in cell adhesion: Rgd and integrins. *Science* 238, 491–497.
- Sackmann, E., Bruinsma, R.F., 2002. Cell adhesion as wetting transition? *ChemPhysChem* 3, 262–269.
- Sackmann, E., Smith, A.S., 2014. Physics of cell adhesion: Some lessons from cell-mimetic systems. *Soft Matter* 10, 1644–1659.
- Schaller, M.D., 2010. Cellular functions of fak kinases: Insight into molecular mechanisms and novel functions. *Journal of Cell Science* 123, 1007–1013.
- Schoenwaelder, S.M., Burridge, K., 1999. Bidirectional signaling between the cytoskeleton and integrins. *Current Opinion in Cell Biology* 11, 274–286.

- Schorb, W., Peeler, T.C., Madigan, N.N., Conrad, K.M., Baker, K.M., 1994. Angiotensin ii-induced protein tyrosine phosphorylation in neonatal rat cardiac fibroblasts. *Journal of Biological Chemistry* 269, 19626–19632.
- Schurpf, T., Springer, T.A., 2011. Regulation of integrin affinity on cell surfaces. *EMBO Journal* 30, 4712–4727.
- Shah, A.M., Mann, D.L., 2011. In search of new therapeutic targets and strategies for heart failure: Recent advances in basic science. *Lancet* 378, 704–712.
- Shattil, S.J., Kashiwagi, H., Pampori, N., 1998. Integrin signaling: The platelet paradigm. *Blood* 91, 2645–2657.
- Smeal, R.M., Rabbitt, R., Biran, R., Tresco, P.A., 2005. Substrate curvature influences the direction of nerve outgrowth. *Annals of Biomedical Engineering* 33, 376–382.
- Song, F., Wisithphrom, K., Zhou, J., Windsor, L.J., 2006. Matrix metalloproteinase dependent and independent collagen degradation. *Frontiers in Bioscience* 11, 3100–3120.
- Song, X.Q., Yang, J., Hirbawi, J., *et al.*, 2012. A novel membrane-dependent on/off switch mechanism of talin form domain at sites of cell adhesion. *Cell Research* 22, 1533–1545.
- Steffensen, B., Hakkinen, L., Larjava, H., 2001. Proteolytic events of wound-healing – Coordinated interactions among matrix metalloproteinases (mmps), integrins, and extracellular matrix molecules. *Critical Reviews in Oral Biology & Medicine* 12, 373–398.
- Steinhardt, R.A., Lundin, L., Mazia, D., 1971. Bioelectric responses of the echinoderm egg to fertilization. *Proceedings of the National Academy of Sciences of the United States of America* 68, 2426–2430.
- Ten Cate, A.R., 1972. Morphological studies of fibrocytes in connective tissue undergoing rapid remodelling. *Journal of Anatomy* 112, 401–414.
- Van Geet, C., Izzi, B., Labarque, V., Freson, K., 2009. Human platelet pathology related to defects in the g-protein signaling cascade. *Journal of Thrombosis and Haemostasis* 7 (Suppl. 1), 282–286.
- Van Noorden, C.J., Everts, V., 1991. Selective inhibition of cysteine proteinases by z-phe-alach2f suppresses digestion of collagen by fibroblasts and osteoclasts. *Biochemical and Biophysical Research Communications* 178, 178–184.
- Wang, N., Butler, J.P., Ingber, D.E., 1993. Mechanotransduction across the cell surface and through the cytoskeleton. *Science* 260, 1124–1127.
- Wehrle-Haller, B., Imhof, B., 2002. The inner lives of focal adhesions. *Trends in Cell Biology* 12, 382–389.
- Weiss, L., 1959. Studies on cellular adhesion in tissue culture. 2. The adhesion of cells to gel surfaces. *Experimental Cell Research* 17, 508–515.

Relevant Websites

- bscb.org/learning-resources/softcell-e-learning/extracellular-matrix-and-cell-adhesion-molecules/
British Society for Cell Biology.
- www.matrixdynamics.ca
Matrix Dynamics.
- www.mechanobio.info/modules/go-0007155
Mechanobiology Institute.
- www.nature.com/scitable/topicpage/cell-adhesion-and-cell-communication-14050486
Nature Education.

CELL COMMUNICATION: CELLULAR OUTCOMES

Contents

Regulation of Cell Polarity

Regulation of Cell Migration

Regulation of Cell Polarity

JD Axelrod, Stanford University School of Medicine, Stanford, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Defects in planar cell polarity (PCP) result in a range of developmental anomalies (birth defects) and diseases including neural tube closure defects (reviewed in [Simons and Mlodzik, 2008](#); [Copp and Greene, 2009](#)), polycystic kidneys (reviewed in [Simons and Walz, 2006](#); [Wallingford, 2006](#); [Wang and Nathans, 2007](#); [Simons and Mlodzik, 2008](#)), conotruncal heart defects ([Garriock et al., 2005](#); [Phillips et al., 2005, 2007](#); [Henderson et al., 2006](#)), deafness ([Curtin et al., 2003](#); [Montcouquiol et al., 2003, 2006](#); [Lu et al., 2004](#); [Davies et al., 2005](#); [Wang et al., 2005, 2006a,c](#); [Deans et al., 2007](#); [Qian et al., 2007](#); [Jones et al., 2008](#)), and situs inversus (reviewed in [Santos and Reiter, 2010](#)). PCP polarizes skin and hair ([Guo et al., 2004](#); [Wang et al., 2006b,c](#); [Devenport and Fuchs, 2008](#)), and is also believed to underlie directed migration during wound healing ([Lee and Adler, 2004](#); [Caddy et al., 2010](#)) and invasion and metastasis of malignant cells ([Weeraratna et al., 2002](#); [Lee et al., 2004](#); [Katoh, 2005](#); [Coyle et al., 2008](#); [Kuriyama and Mayor, 2008](#)). Despite the obvious importance of this pathway in human physiology and development, and considerable attention and progress in the past 15 years, major questions about the signaling mechanism remain to be answered.

Most of our mechanistic understanding of PCP signaling comes from work in *Drosophila*, in which numerous tissues, including the wing, eye, and abdomen display manifestations of PCP ([Tree et al., 2002a](#); [Zallen, 2007](#); [Simons and Mlodzik, 2008](#)). Of these, the most thoroughly studied planar polarized tissue is the fly wing, in which each cell produces a trichome (or 'hair'), that in wild type emerges from the distal side of the cell and points distally ([Figure 1](#); [Wong and Adler, 1993](#)). During polarization of hair cells, cytoskeletal regulators are recruited to proximal and distal sides of the cell and control the localization of actin polymerization and bundling to ensure distal localization of hair emergence ([Wong and Adler, 1993](#); [Strutt and Warrington, 2008](#); [Yan et al., 2009](#); [Lu et al., 2010](#)). Mutants either fail to choose a side, thus producing a hair from the center of the cell, or choose an incorrect side, resulting in an incorrectly oriented hair. Cells of the

abdominal epithelium produce multiple posteriorly oriented hairs that emerge from the posterior side of the cell and point posteriorly.

Polarity in the eye results from the differentiation of the initially equipotent R3/R4 photoreceptor progenitors into an equatorial R3 and a polar R4 ([Figure 1](#)). Here, the key distinction is not between opposite sides of the same cell, but between adjacent sides of this pair of progenitor cells. A competition for Notch signaling activation between the R3/R4 pair becomes biased by the PCP signal at the intercellular junction so that the equatorial cell always expresses low Notch levels and becomes R3. PCP mutants lead either to incorrect R3/R4 fate decisions, or in some cases, indistinctly differentiated pairs of R3/R4 cells ([Zheng et al., 1995](#); [Cooper and Bray, 1999](#)).

On the notum, PCP controls the orientation of an asymmetric cell division in the sensory organ precursor (pI) cells. pI cells differentiate within the epithelium and divide asymmetrically to produce an anterior pIIb daughter and a posterior pIIa daughter cell. The PCP pathway distinguishes anterior and posterior sides of the pI cell through asymmetric interactions with its anterior and posterior neighbors. The pI cell therefore seems to become polarized much like the surrounding epithelial cells do, but it uses this polarity to position cell fate determinants and the mitotic spindle prior to an asymmetric division. PCP mutants cause this division to be incorrectly oriented.

Mechanosensory organs are composed of a multicellular bristle, and an associated bract ([Figure 1](#)). These organs are arranged in precise patterns, and have an intrinsic polarity that is evident at several levels. The bract, elaborated by a single cell, resembles a hair, and points distally. The shaft of the bristle points distally, similar to the bracts and hairs. In addition, the base of each shaft is surrounded by a socket cell, and the vector defined by each socket-bract pair points proximally. Each of these polarities can be disturbed by PCP mutants, and may vary independently ([Held et al., 1986](#)).

The molecular machinery that directs these morphogenetic events downstream of the core module is specific to the tissue. For example, during polarization of hair cells, cytoskeletal regulators are recruited to proximal and distal sides of the cell and control the localization of actin polymerization and

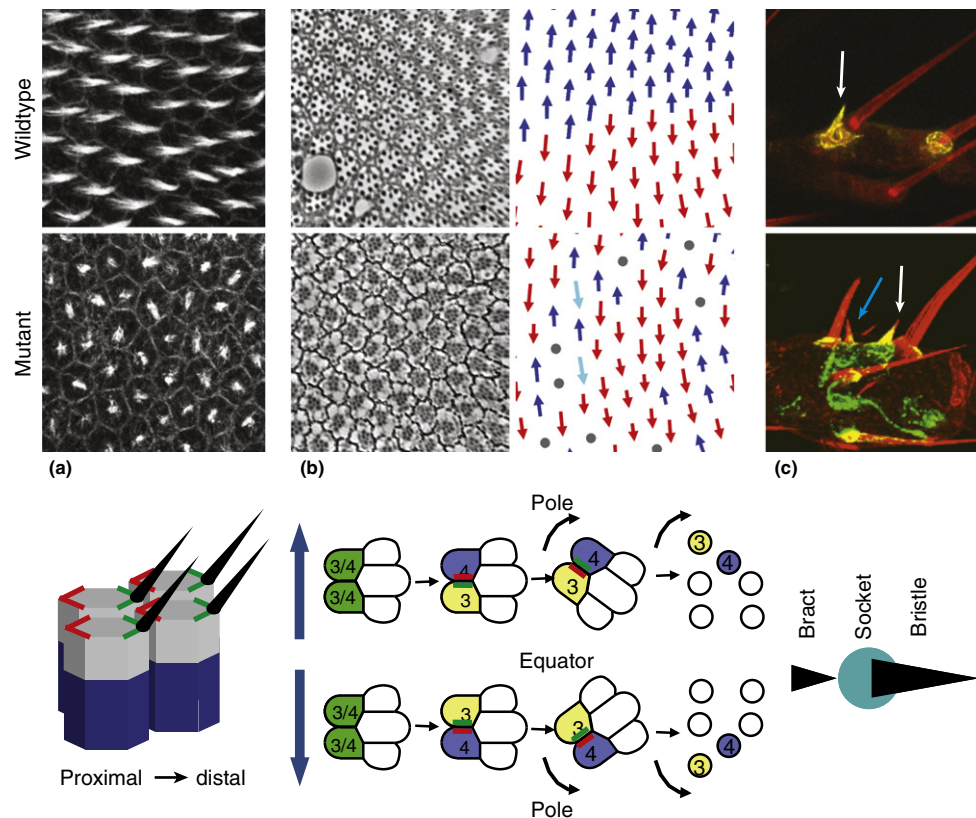


Figure 1 Manifestations of wild type and mutant PCP signaling. (a) Phalloidin stains showing the actin-based emerging hairs (prehairs) growing from the distal side of each cell and pointing distally in wild type, and growing from the center of the cell in a PCP mutant wing. Below, a cartoon of each cell, with green and red bars schematizing the subcellular localization of PCP proteins once cells are polarized. (b) Polarity in a wild-type and a mutant eye. Sections reveal the rhabdomeres of seven photoreceptor cells in each ommatidium, each forming a chiral structure with a particular orientation. In wild type, the two chiral forms are on opposite sides of an obvious equator, and point in opposite directions. In the mutant, chiralities are interspersed, and some ommatidia have symmetrical form. Many ommatidia are also misrotated. The schematic shows the PCP-dependent differentiation of the R3 and R4 cells, and subsequent rotation and morphogenesis of each ommatidium. (c) Bracts associated with mechanosensory bristles on the legs of a wild type and a PCP mutant fly. In wild type, the bract is always induced from the epithelial cell directly proximal to the socket and shaft (white arrow). In the mutant, one bract is correctly positioned (white arrow), while the other is incorrectly located on the distal side, and points proximally (blue arrow).

bundling to ensure distal localization of hair emergence (Wong and Adler, 1993; Strutt and Warrington, 2008; Yan *et al.*, 2009; Lu *et al.*, 2010). In contrast, tissue-specific control of bristle and bract polarization represents the interaction of the PCP modules with a more complex multicellular system (Peng *et al.*, 2012).

Clones of cells that are either mutant for a core PCP component or that overexpress a core component stereotypically perturb, or fail to perturb, the polarity of cells in nearby nonmutant tissue (Gubb and Garcia-Bellido, 1982; Vinson and Adler, 1987; Taylor *et al.*, 1998; Adler *et al.*, 2000). This is most obvious in the wing and eye, where arrays of polarizing cells are present (Figure 2). In the cases when polarity is perturbed, polarities are reversed to either abnormally point toward or away from the clone. This phenomenon is referred to as domineering nonautonomy, and together with the corresponding abilities of these clones to reorganize the localization of core PCP proteins in these cells (see below) has provided fertile ground for experiments aimed at understanding the mechanism of core PCP signaling.

In vertebrates, many aspects of PCP signaling identified in flies are conserved, but genetic analyses indicate that a more diverse array of morphologic events is controlled by the same components together with additional regulators not present in flies (Jones and Chen, 2007; Wang and Nathans, 2007; Zallen, 2007; Simons and Mlodzik, 2008). Here, we will focus on our emerging understanding of PCP signaling mechanisms derived from studies in flies.

A Modular Signaling System

Genetic and molecular analyses in several *Drosophila* tissues have identified components of the PCP signaling mechanism, and have suggested that they may be divided into three classes of functional modules: global directional modules, a core module, and a suite of tissue-specific effector modules that respond to the upstream modules to produce morphological asymmetry in individual tissues (Tree *et al.*, 2002a). The global directional modules link tissue-level directional information

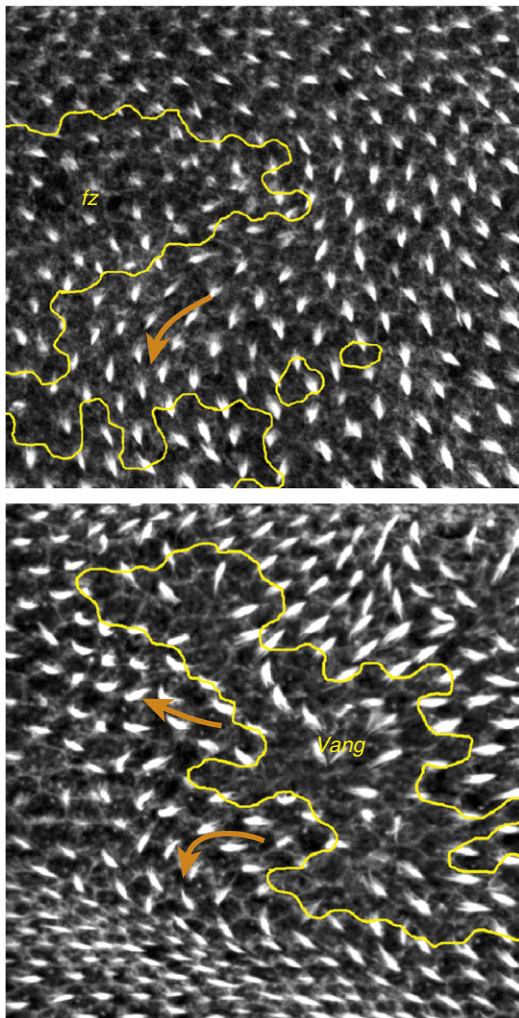


Figure 2 Domineering nonautonomy near a *fz* and a *vang* mutant clone in the wing, visualizing the emerging prehairsts with phalloidin stain. Clones were genetically marked (not shown) and are surrounded by the yellow lines. In a wild-type wing, prehairsts point distally. Note that prehairsts are reoriented to point toward the *fz* clone and away from the *vang* clone (orange arrows). They therefore perturb polarity on opposite sides of the clone.

to local cell polarization. The core module amplifies molecular asymmetry within cells while linking the polarity of cells with their neighbors, producing a coordinated polarization across the tissue. The tissue-specific modules act as effectors for a diverse group of morphological readouts of the molecular polarity produced by the upstream modules.

The first recognized and best understood global module is the Fat/Dachsous/Four-jointed (Ft/Ds/Fj) module (Figure 3). This module serves to convert tissue-level expression gradients to subcellular gradients of Ft–Ds heterodimer localization (Mahoney *et al.*, 1991; Clark *et al.*, 1995, 2002; Adler *et al.*, 1998; Yang *et al.*, 2002; Ma *et al.*, 2003; Lawrence *et al.*, 2004). It consists of the atypical cadherins Ft and Ds, that form heterodimers which may orient in either of two directions at any apical cell–cell boundary, and the Golgi resident protein Four-jointed (Fj) (Villano and Katz, 1995; Zeidler *et al.*, 2000;

Yang *et al.*, 2002; Ma *et al.*, 2003). Fj acts as an ectokinase on both Ft and Ds (Ishikawa *et al.*, 2008), to make Ft a stronger ligand, and Ds a weaker ligand, for the other (Brittle *et al.*, 2010; Simon *et al.*, 2010). As Fj and Ds are expressed in gradients across tissues (Zeidler *et al.*, 1999; Casal *et al.*, 2002; Yang *et al.*, 2002; Ma *et al.*, 2003; Lawrence *et al.*, 2004), the result is a larger fraction of Ft–Ds heterodimers in one orientation relative to the other (Ma *et al.*, 2003). This molecular asymmetry is predicted to be subtle, but can be detected at early stages of development when the Ds and Fj gradients are steeper (Ambegaonkar *et al.*, 2012; Bosveld *et al.*, 2012; Brittle *et al.*, 2012). Many elements of this relatively elegant and simple model are likely to be correct, yet it is clear that it is far from a complete picture, and important puzzles remain. This mechanism will not be discussed at length here, as it was recently reviewed (Matis and Axelrod, 2013).

For some time, Wnts have been known to participate in PCP signaling in vertebrates (Vladar *et al.*, 2009; Goodrich and Strutt, 2011; Wallingford, 2012), and recently, a potential role for Wnts, likely as global signals, has been proposed in flies, where Wnt4 and Wg have been suggested to function as redundant signals that act near the wing margin (Wu and Mlodzik, 2008). The localized expression of these diffusible molecules at the wing margin suggests the possibility that they may signal by forming gradients, but little is yet known about their mechanism of action. Studies of the *Drosophila* abdomen have led to the suggestion of yet another type of global signaling information. The Ft/Ds/Fj mechanism clearly contributes to PCP in the abdomen, but may not explain all of global signaling function. To date, there is no firm evidence pointing to the identity of an additional global signal (Lawrence *et al.*, 2002).

The core module comprises the first recognized molecular components of PCP signaling. Several decades of study have led to the understanding that it acts both to amplify asymmetry, and to coordinate polarization between neighboring cells, producing a local alignment of polarity. Proteins in the core signaling module, including the seven-pass transmembrane protein Frizzled (Fz) (Vinson and Adler, 1987; Vinson *et al.*, 1989), the multi-domain protein Dishevelled (Dsh) (Klingensmith *et al.*, 1994; Thiesen *et al.*, 1994), the Ankyrin repeat protein Diego (Dgo) (Feiguin *et al.*, 2001), the four-pass transmembrane protein Van Gogh (Vang; a.k.a. Strabismus) (Taylor *et al.*, 1998; Wolff and Rubin, 1998), the Lim domain protein Prickle (Pk) (Gubb *et al.*, 1999), and the seven-transmembrane atypical cadherin Flamingo (Fmi; aka Starry night) (Chae *et al.*, 1999; Usui *et al.*, 1999; reviewed in Zallen, 2007), adopt asymmetric subcellular localizations that predict the morphological polarity pattern such as hairs in the fly wing (Usui *et al.*, 1999; Axelrod, 2001; Feiguin *et al.*, 2001; Strutt, 2001; Tree *et al.*, 2002b; Bastock *et al.*, 2003). These proteins form intercellular complexes that communicate at cell boundaries, recruiting one group to the distal side of cells, and the other to the proximal side, through the action of a poorly understood feedback mechanism, thus aligning the polarity of adjacent cells (Tree *et al.*, 2002b; Amonlirdviman *et al.*, 2005). Though not formally established, this mechanism has all the hallmarks of bistability. The unpolarized state is predicted to be metastable and an input bias, or fluctuations caused by

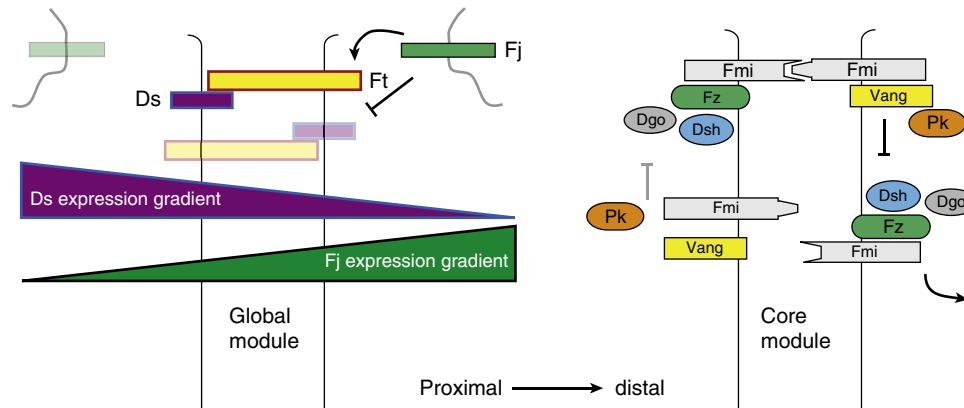


Figure 3 Schematics of the Ft/Ds/Fz global and the core modules. Organization of components at the junction between two cells is depicted. Gradients of Fz and Ds expression lead to a slight excess of Ft–Ds heterodimers at each interface. The core module is depicted as undergoing mutual inhibition, with the proximal components Vang and Pk disrupting the Fz–Dsh distal components. This particular competition was won by the upper complex, in the wild-type orientation.

noise, is expected to bias a coordinated polarization that will then amplify in magnitude. Indeed, this behavior has been captured by a variety of mathematical modeling frameworks for PCP signaling (Lawrence *et al.*, 2004; Amonlirdviman *et al.*, 2005; Le Garrec *et al.*, 2006; Le Garrec and Kerszberg, 2008; reviewed in Axelrod and Tomlin, 2011).

It is important to point out that the core module mechanism has no intrinsic directionality; allowed to polarize in the absence of a global directional signal, it is expected (and indeed seen in experiments and in simulations; reviewed in Axelrod and Tomlin, 2011) to achieve locally coordinated polarity that produces swirling patterns not respecting the global tissue axes. The global signals are therefore envisaged as providing subtle directional inputs that would both bias the initiation of polarization in a given direction and sustain the orientation of polarization during disrupting events such as cell divisions and morphogenetic rearrangements that require realignment of the polarization axis.

The segregation of proximal and distal core PCP proteins to opposite sides of the cell is a molecular marker of cell polarization, and correlates intimately with the various morphological polarization responses. For example, trichomes develop at the side of the cell where the distal proteins Fz and Dsh localize in wing cells, and Delta is activated and Notch repressed in the prospective R3 cell where Fz and Dsh localize (Figure 1).

Principles of Collective Polarization

On theoretical grounds, one can posit that generation of cell polarity requires two coordinated signaling events. First, accumulation of a cell polarity factor, P, must be self-enhancing, and second a long-range signal is required to inhibit accumulation of the same factor (Figure 4; Meinhardt, 2007). This has been modeled in single cells, and studied experimentally in systems such as budding yeast and migrating neutrophils, where, by definition, both signals are intracellular.

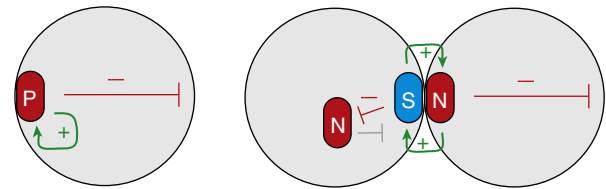


Figure 4 Cartoons of activities required for cell polarization in a single cell using only intracellular interactions (left) and using both intracellular and intercellular interactions (right).

Intracellular local self-enhancement can result from cooperative interactions between components of the polarity factor, for example. Long-range intracellular inhibition may result from a separate diffusing inhibitory molecule, or perhaps more simply, by limiting the amounts of one or more components of the polarity factor, thereby causing depletion and suppression of accumulation across the cell as a single locus stochastically becomes predominant and more rapidly accumulates the factor.

While the same principles apply in polarization of multicellular assemblies, the possibility for intercellular signaling provides additional routes to fulfill these two functions. Given a system with two polarity complexes that have the ability to recruit each other across intercellular boundaries like North and South poles of magnets, one can hypothesize that the intracellular long-range inhibition can be complemented, or even replaced, by an intercellular signal in which accumulation of North on one side of a given cell attracts components of South to the adjacent side of the neighboring cell. Since South and North repel each other, this will inhibit North from accumulating adjacent to its location in the first cell, thus providing a long-range inhibition, but in the neighbor rather than in the original cell. Furthermore, local self-enhancement can proceed by a related mechanism. North in the first cell attracts South in the neighbor, which repels North. By repelling North, accumulation of additional South in the first cell is

diminished. Physicists have thoroughly modeled these types of systems.

Feedback Control of Core Amplification

Experimental evidence suggests that the core PCP proteins function, at least in part, through intercellular interactions akin to those described above. The wild-type distributions of proximal and distal core proteins to opposite sides of the cell, and the rearrangements observed around mutant or over-expressing clones have long suggested just such a North–South relationship between proximal and distal proteins (see e.g., [Tree et al., 2002b](#); [Strutt and Warrington, 2008](#); [Wu and Mlodzik, 2008](#)), and this was eventually demonstrated formally ([Chen et al., 2008](#)). Here, North and South can be equated to Fz and Vang. Their interaction requires an additional factor, the atypical cadherin Fmi that forms homodimers that participate in the complex. Of the core PCP proteins, only Fmi is required, in both cells, to mediate this interaction.

The asymmetrically localized subcellular complexes, with Fz on the distal side and Vang on the proximal side of adjacent cells, communicate information bidirectionally between those cells ([Strutt and Strutt, 2007](#); [Chen et al., 2008](#)). This is perhaps most simply illustrated by the observation that cells on either side of adjacent *vang* and *fz* mutant clones both strongly polarize, indicating that cells with only the Fz complex and cells with only the Vang complex can strongly polarize a neighboring cell ([Strutt and Strutt, 2007](#)). Fmi homodimers are essential for this communication ([Usui et al., 1999](#); [Bastock et al., 2003](#); [Lawrence et al., 2004](#); [Strutt and Strutt, 2007](#); [Chen et al., 2008](#)). It has been shown that, rather than acting simply as a scaffold for complex assembly, information passes bidirectionally through the Fmi bridge that, although a homodimer, is functionally, and presumably structurally, asymmetric ([Chen et al., 2008](#)). This is evident from the observation that clones of cells overexpressing Fmi, but lacking both Fz and Vang retain the ability to repolarize neighboring cells. Competing models have alternately suggested that information only need flow through the Fmi homodimers ([Chen et al., 2008](#)), or that Fz and Vang directly interact to mediate signal transmission ([Wu and Mlodzik, 2008](#)).

According to this understanding, these components can provide the mutual recruitment needed, while the mutual inhibition needed for the North–South repulsion is thought to be provided by the cytosolic PCP factors Pk, Dsh, and Dgo. A more detailed discussion of how these interactions might occur follows below.

The recruitment of North and South across intercellular boundaries into complexes, in the form of Fz–Fmi:Fmi–Vang complexes, can occur in either orientation at any given intercellular boundary. Long-range intercellular inhibition requires that, in addition, North exclude South, or South exclude North, or both, within each cell. Only indirect evidence for this mechanism currently exists. This kind of model, on a conceptual level, was presciently proposed by [Adler et al. \(1997\)](#), but at the time, little molecular context was available ([Taylor et al., 1998](#)). A somewhat more explicit model was subsequently proposed based on the observation that Pk, which

accumulates with Vang (on the proximal side, in wild type), was able to block Fz-dependent recruitment of Dsh to the cell cortex in cell cultures. Coupled with the phenomenological observations that proximal and distal proteins always seem to segregate to opposite sides when polarization occurs, one could infer that mutual inhibition occurs ([Tree et al., 2002b](#); [Bastock et al., 2003](#); [Amonlirdviman et al., 2005](#)). However, the network of interactions has not been demonstrated, and it is difficult to parse out whether, at a wild-type intercellular junction, the exclusion of Fz/Dsh by Vang/Pk relies on exclusion within one cell or the loss of recruitment by the neighbor.

Asymmetric Fz–Fmi:Fmi–Vang complexes can form in the absence of the cytosolic factors, and can transmit polarity information under conditions in which differences are enforced between cells. When differences are not enforced between cells, Dsh, Pk, and Dgo functions are required for the feedback-mediated amplification of the asymmetry that develops at proximal–distal intercellular boundaries ([Lawrence et al., 2004](#); [Strutt and Strutt, 2007](#); [Chen et al., 2008](#)). In their absence, asymmetry is not amplified, and core PCP components remain more or less uniformly distributed around the cell.

As seen in wild type, core PCP proteins are not evenly distributed at intercellular junctions, but exist in uneven distributions, giving the appearance of being concentrated into aggregates, or puncta in some locations. This punctate appearance is greatly exaggerated with overexpression of Pk, one of the cytosolic proximal factors, and overall levels are increased ([Tree et al., 2002b](#); [Bastock et al., 2003](#)). Fluorescence recovery after photobleaching (FRAP) has been used to demonstrate that the puncta are indeed much more stable aggregates, at least for Fmi and Fz ([Strutt et al., 2011](#)). While membrane localization requires only Fz, Fmi, and Vang, the cytosolic factors are needed to induce stable aggregates. Unstable Fmi and Fz, or the fractions not in puncta, are subject to internalization and either recycling or degradation. Thus, the pools of core PCP components appear to exist in steady states of stable, aggregated complete complexes, less stable complexes consisting of Fz–Fmi:Fmi–Vang, and lower order, less stable subcomplexes. Pools of the unstable complexes are regulated by internalization. While additional detail about these processes is known, there is still no clear understanding of how the cytosolic factors produce amplification.

One hypothesis is that the cytosolic factors induce clustering, thereby providing the local cooperativity needed for polarization. The mechanism for clustering, or how it is enhanced by Pk, is not known. As previously proposed by [Strutt et al. \(2011\)](#), clustering may result from a scaffolding effect. These investigators dismissed the possibility of decreased endocytosis accounting for clustering.

The proximal protein Pk and the distal proteins Dgo and Dsh have been shown to engage in competitive binding interactions ([Tree et al., 2002b](#); [Jenny et al., 2003, 2005](#); [Das et al., 2004](#)). This suggests, first, that oppositely oriented complexes come into contact, and second, suggests the possibility that these interactions might result in disruption of one or the other complex. At present, this potential mechanism is speculative, but is appealing, and consistent with the early proposal by [Tree et al. \(2002b\)](#). Such a model would provide long-range intercellular inhibition that is needed for polarization.

System Architecture

Studies of polarity in the wing and eye first led to the proposal that the Ft/Ds/Fj global module provides directional information that orients the activity of the core module (Yang *et al.*, 2002; Ma *et al.*, 2003; Figure 5). This hierarchical, or series, architectural scheme derived from the observation that orientation of core PCP proteins and morphological polarity are coordinately disturbed in global mutant clones (Ma *et al.*, 2003), and from mosaic analyses in the fly eye showing that boundaries between cells expressing and not expressing *ft*, *ds*, or *fj* lost the ability to influence polarity in the absence of *fz* (Yang *et al.*, 2002). In contrast, others have argued for a parallel system architecture, with the proposal of a bypass pathway (green arrow) resting upon observations in the abdomen that clones overexpressing global components (Ds or Ds[ecto], a more active form of Ds) produce domineering non-autonomy and therefore influence polarity within tissue lacking either Fz or Fmi (Casal *et al.*, 2006; Lawrence *et al.*, 2007). The argument supporting this scheme is also not definitive (discussed at length in Axelrod, 2009). Furthermore, it was suggested that PCP in the larval epidermis depends on the global module, while the core module has no function in this context (Repiso *et al.*, 2010), although a more detailed study found that the core module does become important in this tissue as it grows (Donoughe and DiNardo, 2011). Thus, the extent to which the blue, green, and black pathways (Figure 5) contribute has been proposed to differ in different contexts. The extents to which they contribute, or whether they in fact contribute at all in some tissues, remain controversial. Furthermore, since the data leading to alternate architectures derive from different tissues, one must consider the possibility that the architecture differs in different tissues. For the purposes of this discussion, only the Ft/Ds/Fj global module will be considered, as there is little data connecting Wnt4/Wg to other systems, and the identity of other potential global systems is not known.

The Ft/Ds/Fj system has features that make it an attractive candidate for providing directional information in multiple tissues. The direction of ommatidial polarization in the *Drosophila* eye depends on Ds and Fj gradients, and these effects require an intact core module (Yang *et al.*, 2002). In the eye, flattening the gradients leads to near randomization of local polarity, and inversion of the gradients leads to reversal of local polarity (Simon, 2004). In the wing, the core PCP

proteins lose their orientation and polarize in swirling patterns in *ft* mutant clones (Ma *et al.*, 2003), and altering Ds expression patterns can reorient wing polarity (Matakatsu and Blair, 2004; Harumoto *et al.*, 2010). In the abdomen, opposing gradients of Ds and Fj are also observed, and these have opposite directions in anterior and posterior compartments of each segment (Casal *et al.*, 2002), and mutants demonstrate that they are needed for normal polarization. These observations are all consistent with a global directional function for the Ft/Ds/Fj module, acting upstream of, and in series with, the core module.

However, several observations are less simply reconciled with this model. When Ds and Fj, normally expressed in gradients along the proximal–distal (P/D) wing axis, are instead expressed uniformly, polarity is disrupted only in the proximal part of the wing (Matakatsu and Blair, 2004; Simon, 2004). In the abdomen, ectopic Ds expression can reorient tissue mutant for the core factors *fz* and *fmi* (Casal *et al.*, 2006). This would not be expected if the role of the Ft/Ds/Fj module is to provide directional information to the core module. Hence, this model is controversial, and others have proposed instead that the Ft/Ds/Fj and core modules function in parallel, acting independently on downstream effectors, and leaving open the question of what cues might orient core module function (Casal *et al.*, 2006).

In the face of the uncertainty about the relationship between the Ft/Ds/Fj module and the core module, the outlines of a unifying hypothesis that connects them in series have begun to emerge.

The first important observation is that microtubules appear to play a central role in the global alignment of PCP in the developing *Drosophila* wing. An unusual apical microtubule cytoskeleton (Eaton *et al.*, 1996; Shimada *et al.*, 2006; Harumoto *et al.*, 2010) can be seen in cells of the fly wing during polarization. During efforts to study the dynamics of Fz using a Fz::GFP transgene and live imaging, Shimada *et al.* (2006) noticed that what appeared to be vesicles can be seen departing the apical membrane, traversing the cell, and fusing to another side. Correlation to immuno-electron micrographs lends support to the notion that these were indeed Fz-containing vesicles that are associated with microtubules. A quantitative analysis showed that these vesicles move with an overall distal bias. Remarkably, using EB1::GFP, a marker of microtubule plus-ends, it was seen that the microtubules have a corresponding bias of distal plus-ends.

Vesicle movement was most active during the most dynamic period of polarization, and vesicle production depended on other core PCP proteins (Shimada *et al.*, 2006). From these data, biased trafficking could be inferred to be important in PCP acquisition, but whether this process was intrinsic to the core mechanism, or perhaps reflected a directional input, could not be discerned. However, a telltale clue was that the microtubule structure itself appeared not to depend on the core module. In a subsequent report, Harumoto *et al.* (2010) showed that the polarity of the microtubules depends on Ds, a component of the Ft/Ds/Fj module. This was consistent with evidence that Ft influences apical microtubule organization in another context (Marcinkevicius and Zallen, 2013). Plus-end bias was perturbed in *ds* mutants, and misexpression of Ds could reorganize microtubule polarity. These experiments were

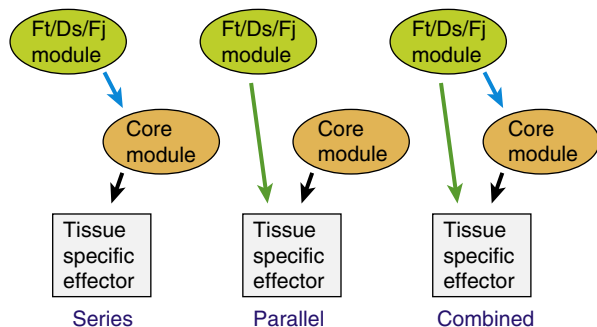


Figure 5 Possible architectures of the Ft/Ds/Fj global and the core modules in relation to the effector modules.

complicated by the lethality associated with *ft* mutants, precluding a more thorough analysis, but suggested the possibility that the Ft/Ds/Fj global module organizes microtubules that in turn directionally bias core module polarization.

It is therefore tempting to hypothesize that polarized apical microtubules that transport Fz might comprise the Ft/Ds/Fj global directional signal that orients core PCP protein segregation. This hypothesis makes a number of predictions that remain to be tested. First, the oriented microtubules should appear and be polarized prior to the onset of core polarization. Second, there should be a spatiotemporal relationship between the orientation of apical microtubules and core polarization. This is of greatest interest in the wing, where substantial morphogenetic changes occur between third instar, when polarization begins, to the conclusion of polarization in the pupal stage, when hairs begin to grow. Finally, a more detailed analysis of the contributions of Ft/Ds/Fj to microtubule organization is needed.

The above hypothesis is based on data from the wing, yet the phenotypes and expression patterns, in particular the graded expression, of Ds and Fj suggest that they might have similar roles in other compartments. However, at first glance, there is an important inconsistency. If one imagines that the gradients of Ds and Fj organize microtubules, and thereby vesicle transport, in such a way that Fz accumulates toward low Ds and high Fj, this is observed in the wing and posterior abdomen, but the relative orientations are opposite in anterior abdomen and eye (Zeidler *et al.*, 2000; Casal *et al.*, 2002; Ma *et al.*, 2003; Matakatsu and Blair, 2004; Rogulja *et al.*, 2008; Axelrod, 2009).

Because the relative orientations of the Fj and Ds gradients with respect to the direction of core protein polarization is not conserved in wing, eye, and abdomen, universality of the model described above requires that one explain how oppositely oriented inputs can be rectified to produce similarly polarized outcomes. A clue to this rectification puzzle comes from the observation that tissues in which Fz and Dsh accumulate toward high Fj (wing and posterior abdomen (P-abd)) rely on the Pk but not the Sple isoform of Pk-Sple, while tissues in which Fz and Dsh accumulate away from high Fj (eye and anterior abdomen (A-abd)) rely on the Sple isoform; in wing and P-abd, polarity is disturbed in mutants of the *pk^{pk}* but not *pk^{sple}* isoform of the *pk* gene, whereas eye and A-abd are sensitive to *pk^{sple}* but not *pk^{pk}* (Gubb *et al.*, 1999; Lawrence *et al.*, 2004). One can therefore infer that Pk is the predominantly expressed or active isoform in wing and P-abd, while Sple is predominant in A-abd and eye.

Furthermore, clues suggest that the direction of gradient interpretation might depend on the predominant isoform present in each tissue. Manipulating *pk* and *sple* expression in wing and abdomen can alter the direction of polarization with respect to the Ft/Ds/Fj gradients. More specifically, overexpression of the non-predominant isoform in wing or either abdominal compartment can at least partially reverse polarity (Lawrence *et al.*, 2004; Lin and Gubb, 2009; Doyle *et al.*, 2008; Hogan *et al.*, 2011). This suggests that Pk causes hairs to point toward high Fj, while Sple caused them to point toward low Fj. Indeed, a subsequent analysis has shown that Pk and Sple cause the opposite orientation of microtubules with respect to the Fj and Ds gradients, and therefore the trafficking of

Fz vesicles that could orient the core module is opposite (Olofsson *et al.*, 2014). The relative expression of these isoforms therefore can rectify the upstream signal to orient core module polarization.

Outlook

The past two decades have witnessed a leap in understanding the signals that control PCP, with the advances powered primarily by standard approaches employed in genetic models. Continued progress is likely to require the deployment of other approaches less commonly combined with traditional genetics, including biophysical, proteomic, cell biological, and biochemical studies. Meanwhile, the complex relationship between mutation and phenotype will demand the continued use of mathematical modeling to relate molecular interventions to pattern outcomes for a given network signaling model.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: The Hippo Pathway. Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell Polarity. Cytoskeleton and Motors: Cytoskeletal Components: Microtubules and Microtubule-Associated Proteins (MAPs). Vertical Integration: Applications: Multiscale Analysis of Morphogenesis. Vertical Integration: Modeling: Computational Approaches for Multiscale Modeling

References

- Adler, P.N., Charlton, J., Liu, J., 1998. Mutations in the cadherin superfamily member gene *dachsous* cause a tissue polarity phenotype by altering frizzled signaling. *Development* 125, 959–968.
- Adler, P.N., Krasnow, R.E., Liu, J., 1997. Tissue polarity points from cells that have higher Frizzled levels towards cells that have lower Frizzled levels. *Current Biology* 7, 940–949.
- Adler, P.N., Taylor, J., Charlton, J., 2000. The domineering non-autonomy of frizzled and van Gogh clones in the *Drosophila* wing is a consequence of a disruption in local signaling. *Mechanisms of Development* 96, 197–207.
- Ambegaonkar, A.A., Pan, G., Mani, M., Feng, Y., Irvine, K.D., 2012. Propagation of *Dachsous*-Fat planar cell polarity. *Current Biology* 22, 1302–1308.
- Amonlirdviman, K., Khare, N.A., Tree, D.R., *et al.*, 2005. Mathematical modeling of planar cell polarity to understand domineering nonautonomy. *Science* 307, 423–426.
- Axelrod, J.D., 2001. Unipolar membrane association of Dishevelled mediates Frizzled planar cell polarity signaling. *Genes & Development* 15, 1182–1187.
- Axelrod, J.D., 2009. Progress and challenges in understanding planar cell polarity signaling. *Seminars in Cell & Developmental Biology* 20, 964–971.
- Axelrod, J.D., Tomlin, C.J., 2011. Modeling the control of planar cell polarity. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine* 3, 588–605.
- Bastock, R., Strutt, H., Strutt, D., 2003. Strabismus is asymmetrically localised and binds to Prickle and Dishevelled during *Drosophila* planar polarity patterning. *Development* 130, 3007–3014.
- Bosveld, F., Bonnet, I., Guirao, B., *et al.*, 2012. Mechanical control of morphogenesis by Fat/*Dachsous*/Four-jointed planar cell polarity pathway. *Science* 336, 724–727.
- Brittle, A.L., Repiso, A., Casal, J., Lawrence, P.A., Strutt, D., 2010. Four-jointed modulates growth and planar polarity by reducing the affinity of *dachsous* for fat. *Current Biology* 20 (9), 803–810.
- Brittle, A., Thomas, C., Strutt, D., 2012. Planar polarity specification through asymmetric subcellular localization of Fat and *Dachsous*. *Current Biology* 22, 907–914.
- Caddy, J., Wilanowski, T., Darido, C., *et al.*, 2010. Epidermal wound repair is regulated by the planar cell polarity signaling pathway. *Developmental Cell* 19, 138–147.

- Casal, J., Lawrence, P.A., Struhl, G., 2006. Two separate molecular systems, Dachshous/Fat and Starry night/Frizzled, act independently to confer planar cell polarity. *Development* 133, 4561–4572.
- Casal, J., Struhl, G., Lawrence, P., 2002. Developmental compartments and planar polarity in *Drosophila*. *Current Biology* 12, 1189.
- Chae, J., Kim, M.J., Goo, J.H., *et al.*, 1999. The *Drosophila* tissue polarity gene starry night encodes a member of the protocadherin family. *Development* 126, 5421–5429.
- Chen, W.S., Antic, D., Matis, M., *et al.*, 2008. Asymmetric homotypic interactions of the atypical cadherin flamingo mediate intercellular polarity signaling. *Cell* 133, 1093–1105.
- Clark, H.F., Brentrup, D., Schneitz, K., *et al.*, 1995. Dachshous encodes a member of the cadherin superfamily that controls imaginal disc morphogenesis in *Drosophila*. *Development* 9, 1530–1542.
- Cooper, M.T., Bray, S.J., 1999. Frizzled regulation of Notch signalling polarizes cell fate in the *Drosophila* eye. *Nature* 397, 526–530.
- Copp, A.J., Greene, N.D., 2009. Genetics and development of neural tube defects. *Journal of Pathology* 220, 217–230.
- Coyle, R.C., Latimer, A., Jessen, J.R., 2008. Membrane-type 1 matrix metalloproteinase regulates cell migration during zebrafish gastrulation: Evidence for an interaction with non-canonical Wnt signaling. *Experimental Cell Research* 314, 2150–2162.
- Curtin, J.A., Quint, E., Tsiouri, V., *et al.*, 2003. Mutation of Celsr1 disrupts planar polarity of inner ear hair cells and causes severe neural tube defects in the mouse. *Current Biology* 13, 1129–1133.
- Das, G., Jenny, A., Klein, T.J., Eaton, S., Mlodzik, M., 2004. Diego interacts with Prickle and Strabismus/Van Gogh to localize planar cell polarity complexes. *Development* 131, 4467–4476.
- Davies, A., Formstone, C., Mason, I., Lewis, J., 2005. Planar polarity of hair cells in the chick inner ear is correlated with polarized distribution of c-flamingo-1 protein. *Developmental Dynamics* 233, 998–1005.
- Deans, M.R., Antic, D., Suyama, K., *et al.*, 2007. Asymmetric distribution of prickle-like 2 reveals an early underlying polarization of vestibular sensory epithelia in the inner ear. *Journal of Neuroscience* 27, 3139–3147.
- Devenport, D., Fuchs, E., 2008. Planar polarization in embryonic epidermis orchestrates global asymmetric morphogenesis of hair follicles. *Nature Cell Biology* 10, 1257–1268.
- Donoughe, S., DiNardo, S., 2011. dachshous and frizzled contribute separately to planar polarity in the *Drosophila* ventral epidermis. *Development* 138, 2751–2759.
- Doyle, K., Hogan, J., Lester, M., Collier, S., 2008. The Frizzled planar cell polarity signaling pathway controls *Drosophila* wing topography. *Developmental Biology* 317, 354–367.
- Eaton, S., Wepf, R., Simons, K., 1996. Roles for Rac1 and Cdc42 in planar polarization and hair outgrowth in the wing of *Drosophila*. *Journal of Cell Biology* 135, 1277–1289.
- Feigun, F., Hannus, M., Mlodzik, M., Eaton, S., 2001. The ankyrin repeat protein Diego mediates Frizzled-dependent planar polarization. *Developmental Cell* 1, 93–101.
- Garriock, R.J., D'Agostino, S.L., Pilcher, K.C., Krieg, P.A., 2005. Wnt11-R, a protein closely related to mammalian Wnt11, is required for heart morphogenesis in *Xenopus*. *Developmental Biology* 279, 179–192.
- Goodrich, L.V., Strutt, D., 2011. Principles of planar polarity in animal development. *Development* 138, 1877–1892.
- Gubb, D., Garcia-Bellido, A., 1982. A genetic analysis of the determination of cuticular polarity during development in *Drosophila melanogaster*. *Journal of Embryology & Experimental Morphology* 68, 37–57.
- Gubb, D., Green, C., Huen, D., *et al.*, 1999. The balance between isoforms of the prickle LIM domain protein is critical for planar polarity in *Drosophila* imaginal discs. *Genes & Development* 13, 2315–2327.
- Guo, N., Hawkins, C., Nathans, J., 2004. Frizzled6 controls hair patterning in mice. *Proceedings of the National Academy of Sciences of the United States of America* 101, 9277–9281.
- Harumoto, T., Ito, M., Shimada, Y., *et al.*, 2010. Atypical cadherins dachshous and fat control dynamics of noncentrosomal microtubules in planar cell polarity. *Developmental Cell* 19, 389–401.
- Held Jr., L.I., Duarte, C.M., Derakhshanian, K., 1986. Extra tarsal joints and abnormal cuticular polarities in various mutants of *Drosophila melanogaster*. *Roux's Archives of Developmental Biology* 195, 145–157.
- Henderson, D.J., Phillips, H.M., Chaudhry, B., 2006. Vang-like 2 and noncanonical Wnt signaling in outflow tract development. *Trends in Cardiovascular Medicine* 16, 38–45.
- Hogan, J., Valentine, M., Cox, C., Doyle, K., Collier, S., 2011. Two frizzled planar cell polarity signals in the *Drosophila* wing are differentially organized by the Fat/Dachshous pathway. *PLoS Genetics* 7, e1001305.
- Ishikawa, H.O., Takeuchi, H., Haltiwanger, R.S., Irvine, K.D., 2008. Four-jointed is a Golgi kinase that phosphorylates a subset of cadherin domains. *Science* 321, 401–404.
- Jenny, A., Darken, R.S., Wilson, P.A., Mlodzik, M., 2003. Prickle and Strabismus form a functional complex to generate a correct axis during planar cell polarity signaling. *EMBO Journal* 22, 4409–4420.
- Jenny, A., Reynolds-Kennally, J., Das, G., Burnett, M., Mlodzik, M., 2005. Diego and Prickle regulate Frizzled planar cell polarity signalling by competing for Dishevelled binding. *Nature Cell Biology* 7 (7), 691–697.
- Jones, C., Chen, P., 2007. Planar cell polarity signaling in vertebrates. *Bioessays* 29, 120–132.
- Jones, C., Roper, V.C., Foucher, I., *et al.*, 2008. Ciliary proteins link basal body polarization to planar cell polarity regulation. *Nature Genetics* 40, 69–77.
- Kato, M., 2005. WNT/PCP signaling pathway and human cancer (review). *Oncology Reports* 14, 1583–1588.
- Klingensmith, J., Nusse, R., Perrimon, N., 1994. The *Drosophila* segment polarity gene dishevelled encodes a novel protein required for response to the wingless signal. *Genes and Development* 8, 118–130.
- Kuriyama, S., Mayor, R., 2008. Molecular analysis of neural crest migration. *Proceedings of the Royal Society of London B Biological Sciences* 363, 1349–1362.
- Lawrence, P.A., Casal, J., Struhl, G., 2002. Towards a model of the organisation of planar polarity and pattern in the *Drosophila* abdomen. *Development* 129, 2749–2760.
- Lawrence, P.A., Casal, J., Struhl, G., 2004. Cell interactions and planar polarity in the abdominal epidermis of *Drosophila*. *Development* 131, 4651–4664.
- Lawrence, P.A., Struhl, G., Casal, J., 2007. Planar cell polarity: One or two pathways? *Nature Reviews Genetics* 8, 555–563.
- Lee, H., Adler, P.N., 2004. The grainy head transcription factor is essential for the function of the frizzled pathway in the *Drosophila* wing. *Mechanisms of Development* 121, 37–49.
- Lee, J.H., Park, S.R., Chay, K.O., *et al.*, 2004. KAI1 COOH-terminal interacting tetraspanin (KITENIN), a member of the tetraspanin family, interacts with KAI1, a tumor metastasis suppressor, and enhances metastasis of cancer. *Cancer Research* 64, 4235–4243.
- Le Garrec, J.F., Kerszberg, M., 2008. Modeling polarity buildup and cell fate decision in the fly eye: Insight into the connection between the PCP and Notch pathways. *Development Genes and Evolution* 218, 413–426.
- Le Garrec, J.F., Lopez, P., Kerszberg, M., 2006. Establishment and maintenance of planar epithelial cell polarity by asymmetric cadherin bridges: A computer model. *Developmental Dynamics* 235, 235–246.
- Lin, Y.Y., Gubb, D., 2009. Molecular dissection of *Drosophila* Prickle isoforms distinguishes their essential and overlapping roles in planar cell polarity. *Developmental Biology* 325, 386–399.
- Lu, Q., Yan, J., Adler, P.N., 2010. The *Drosophila* planar polarity proteins interlaced and multiple wing hairs interact physically and function together. *Genetics* 185, 549–558.
- Lu, X., Borchers, A.G., Jolicoeur, C., *et al.*, 2004. PTK7/CCK-4 is a novel regulator of planar cell polarity in vertebrates. *Nature* 430, 93–98.
- Ma, D., Yang, C.H., McNeill, H., Simon, M.A., Axelrod, J.D., 2003. Fidelity in planar cell polarity signalling. *Nature* 421, 543–547.
- Mahoney, P.A., Weber, U., Onofrechuk, P., *et al.*, 1991. The fat tumor suppressor gene in *Drosophila* encodes a novel member of the cadherin gene superfamily. *Cell* 67, 853–868.
- Marcinkevicius, E., Zallen, J.A., 2013. Regulation of cytoskeletal organization and junctional remodeling by the atypical cadherin fat. *Development* 140, 433–443.
- Matakatsu, H., Blair, S.S., 2004. Interactions between Fat and Dachshous and the regulation of planar cell polarity in the *Drosophila* wing. *Development* 131, 3785–3794.
- Matis, M., Axelrod, J.D., 2013. Regulation of PCP by the Fat signaling pathway. *Genes & Development* 27, 2207–2220.
- Meinhardt, H., 2007. Computational modelling of epithelial patterning. *Current Opinion in Genetics & Development* 17, 272–280.
- Montcouquiol, M., Rachel, R.A., Lanford, P.J., *et al.*, 2003. Identification of Vangl2 and Scrb1 as planar polarity genes in mammals. *Nature* 423, 173–177.
- Montcouquiol, M., Sans, N., Huss, D., *et al.*, 2006. Asymmetric localization of Vangl2 and Fz3 indicate novel mechanisms for planar cell polarity in mammals. *Journal of Neuroscience* 26, 5265–5275.
- Olofsson, J., Sharp, K.A., Matis, M., Cho, B., Axelrod, J.D., 2014. Prickle/Spiny-legs isoforms control the polarity of the apical microtubule network in PCP. *Development* 141 (14), 2866–2874.

- Peng, Y., Han, C., Axelrod, J.D., 2012. Planar polarized protrusions break the symmetry of EGFR signaling during *Drosophila* bract cell fate induction. *Developmental Cell* 23, 507–518.
- Phillips, H.M., Murdoch, J.N., Chaudhry, B., Copp, A.J., Henderson, D.J., 2005. Vangl2 acts via RhoA signaling to regulate polarized cell movements during development of the proximal outflow tract. *Circulation Research* 96, 292–299.
- Phillips, H.M., Rhee, H.J., Murdoch, J.N., *et al.*, 2007. Disruption of planar cell polarity signaling results in congenital heart defects and cardiomyopathy attributable to early cardiomyocyte disorganization. *Circulation Research* 101, 137–145.
- Qian, D., Jones, C., Rzadzinska, A., *et al.*, 2007. Wnt5a functions in planar cell polarity regulation in mice. *Developmental Biology* 306, 121–133.
- Repiso, A., Saavedra, P., Casal, J., Lawrence, P.A., 2010. Planar cell polarity: The orientation of larval denticles in *Drosophila* appears to depend on gradients of Dachsous and Fat. *Development* 137, 3411–3415.
- Rogulja, D., Rauskolb, C., Irvine, K.D., 2008. Morphogen control of wing growth through the Fat signaling pathway. *Developmental Cell* 15, 309–321.
- Santos, N., Reiter, J.F., 2010. Tilting at nodal windmills: Planar cell polarity positions cilia to tell left from right. *Developmental Cell* 19, 5–6.
- Shimada, Y., Yonemura, S., Ohkura, H., Strutt, D., Uemura, T., 2006. Polarized transport of Frizzled along the planar microtubule arrays in *Drosophila* wing epithelium. *Developmental Cell* 10, 209–222.
- Simon, M.A., 2004. Planar cell polarity in the *Drosophila* eye is directed by graded four-jointed and Dachsous expression. *Development* 131, 6175–6184.
- Simon, M.A., Xu, A., Ishikawa, H.O., Irvine, K.D., 2010. Modulation of Fat:Dachsous binding by the cadherin domain kinase four-jointed. *Current Biology* 20, 811–817.
- Simons, M., Mlodzik, M., 2008. Planar cell polarity signaling: From fly development to human disease. *Annual Review of Genetics* 42, 517–540.
- Simons, M., Walz, G., 2006. Polycystic kidney disease: Cell division without a c(l)ue? *Kidney International* 70, 854–864.
- Strutt, D.I., 2001. Asymmetric localization of Frizzled and the establishment of cell polarity in the *Drosophila* wing. *Molecular Cell* 7, 367–375.
- Strutt, D., Strutt, H., 2007. Differential activities of the core planar polarity proteins during *Drosophila* wing patterning. *Developmental Biology* 302, 181–194.
- Strutt, D., Warrington, S.J., 2008. Planar polarity genes in the *Drosophila* wing regulate the localisation of the FH3-domain protein Multiple Wing Hairs to control the site of hair production. *Development* 135, 3103–3111.
- Strutt, H., Warrington, S.J., Strutt, D., 2011. Dynamics of core planar polarity protein turnover and stable assembly into discrete membrane subdomains. *Developmental Cell* 20, 511–525.
- Taylor, J., Abramova, N., Charlton, J., Adler, P.N., 1998. Van Gogh: A new *Drosophila* tissue polarity gene. *Genetics* 150, 199–210.
- Thiesen, H., Purcell, J., Bennett, M., *et al.*, 1994. Dishevelled is required during wingless signaling to establish both cell polarity and cell identity. *Development* 120, 347–360.
- Tree, D.R., Ma, D., Axelrod, J.D., 2002a. A three-tiered mechanism for regulation of planar cell polarity. *Seminars in Cell & Developmental Biology* 13, 217–224.
- Tree, D.R.P., Shulman, J.M., Rousset, R., *et al.*, 2002b. Prickle mediates feedback amplification to generate asymmetric planar cell polarity signaling. *Cell* 109, 371–381.
- Usui, T., Shima, Y., Shimada, Y., *et al.*, 1999. Flamingo, a seven-pass transmembrane cadherin, regulates planar cell polarity under the control of Frizzled. *Cell* 98, 585–595.
- Villano, J.L., Katz, F.N., 1995. four-jointed is required for intermediate growth in the proximal-distal axis in *Drosophila*. *Development* 121, 2767–2777.
- Vinson, C.R., Adler, P.N., 1987. Directional non-cell autonomy and the transmission of polarity information by the *frizzled* gene of *Drosophila*. *Nature* 329, 549–551.
- Vinson, C.R., Conover, S., Adler, P.N., 1989. A *Drosophila* tissue polarity locus encodes a protein containing seven potential transmembrane domains. *Nature* 338, 263–264.
- Vladar, E.K., Antic, D., Axelrod, J.D., 2009. Planar cell polarity signaling: The developing cell's compass. *Cold Spring Harbor Perspectives in Biology* 1 (a002964), 235–253.
- Wallingford, J.B., 2006. Planar cell polarity, ciliogenesis and neural tube defects. *Human Molecular Genetics* 15 (Suppl. 2), R227–R234.
- Wallingford, J.B., 2012. Planar cell polarity and the developmental control of cell behavior in vertebrate embryos. *Annual Review of Cell and Developmental Biology* 28, 627–653.
- Wang, J., Hamblet, N.S., Mark, S., *et al.*, 2006a. Dishevelled genes mediate a conserved mammalian PCP pathway to regulate convergent extension during neurulation. *Development* 133, 1767–1778.
- Wang, J., Mark, S., Zhang, X., *et al.*, 2005. Regulation of polarized extension and planar cell polarity in the cochlea by the vertebrate PCP pathway. *Nature Genetics* 37, 980–985.
- Wang, Y., Badea, T., Nathans, J., 2006b. Order from disorder: Self-organization in mammalian hair patterning. *Proceedings of the National Academy of Sciences of the United States of America* 103, 19800–19805.
- Wang, Y., Guo, N., Nathans, J., 2006c. The role of Frizzled3 and Frizzled6 in neural tube closure and in the planar polarity of inner-ear sensory hair cells. *Journal of Neuroscience* 26, 2147–2156.
- Wang, Y., Nathans, J., 2007. Tissue/planar cell polarity in vertebrates: New insights and new questions. *Development* 134, 647–658.
- Weeraratna, A.T., Jiang, Y., Hostetter, G., *et al.*, 2002. Wnt5a signaling directly affects cell motility and invasion of metastatic melanoma. *Cancer Cell* 1, 279–288.
- Wolff, T., Rubin, G.M., 1998. Strabismus, a novel gene that regulates tissue polarity and cell fate decisions in *Drosophila*. *Development* 125, 1149–1159.
- Wong, L.L., Adler, P.N., 1993. Tissue polarity genes of *Drosophila* regulate the subcellular location for prehair initiation in pupal wing hairs. *Journal of Cell Biology* 123, 209–221.
- Wu, J., Mlodzik, M., 2008. The frizzled extracellular domain is a ligand for Van Gogh/Stbm during nonautonomous planar cell polarity signaling. *Developmental Cell* 15, 462–469.
- Yan, J., Lu, Q., Fang, X., Adler, P.N., 2009. Rho1 has multiple functions in *Drosophila* wing planar polarity. *Developmental Biology* 333, 186–199.
- Yang, C.H., Axelrod, J.D., Simon, M.A., 2002. Regulation of Frizzled by fat-like cadherins during planar polarity signaling in the *Drosophila* compound eye. *Cell* 108, 675–688.
- Zallen, J.A., 2007. Planar polarity and tissue morphogenesis. *Cell* 129, 1051–1063.
- Zeidler, M.P., Perrimon, N., Strutt, D.I., 1999. The four-jointed gene is required in the *Drosophila* eye for ommatidial polarity specification. *Current Biology* 9, 1363–1372.
- Zeidler, M.P., Perrimon, N., Strutt, D.I., 2000. Multiple roles for four-jointed in planar polarity and limb patterning. *Developmental Biology* 228, 181–196.
- Zheng, L., Zhang, J., Carthew, R.W., 1995. Frizzled regulates mirror-symmetric pattern formation in the *Drosophila* eye. *Development* 121, 3045–3055.

Regulation of Cell Migration

MA McNiven and GL Razidlo, Mayo Clinic, Rochester, MN, USA

© 2016 Elsevier Inc. All rights reserved.

Signals to Migrate

Overview

What signals trigger a cell to migrate, and instruct it in what direction to move? Cells translocate in response to cues from the extracellular environment. In general terms, ligand binding and activation of growth factor receptors on the cell surface triggers elaborate and potent intracellular signaling cascades that can induce changes in cytoskeletal and membrane dynamics, leading to cell migration. The precise signals that activate cell movement depend on the context and on the target cell. Yet the general concepts are highly conserved: first, extracellular signals are relayed through receptors on the cell surface to stimulate intracellular signaling pathways. These signals are transmitted into biochemical, cytoskeletal, and mechanical changes that ultimately induce cell migration (Figure 1).

Directed Cell Migration

Whenever cell migration is required, it is not sufficient for the cell to migrate randomly, but rather, the cell must migrate in a specific direction. In order to achieve directional migration, the cell must establish polarity (i.e., a front, or leading edge and a rear, or trailing edge). This results from the exposure to a gradient. These gradients may be quite shallow, but even these subtle differences across the body of a cell can lead to localized changes at the site of the gradient signaling resulting in polarization, and the establishment of directional migration. These polarized signals include the generation of specific

phospholipids and the activation of calcium signaling pathways that result in the generation of localized, asymmetric forces that can cause the cell to migrate in the indicated direction. In addition, in response to a mitogenic signal, cells reorient the Golgi apparatus and microtubule cytoskeleton in the direction of migration. These concerted processes of polarization regulate the persistent, directional migration of a cell (Figure 2; Rorth, 2011; Bear and Haugh, 2014; Majumdar *et al.*, 2014).

Chemotaxis

Chemotaxis is the directed migration of a cell in response to a chemical stimulus, such as a growth factor. The specific ligands and receptors used in chemotaxis vary among cell types, as do the specific mechanisms used to relay chemotactic signals. For example, migrating leukocytes can rapidly sense and respond to a chemotactic gradient to respond to tissue damage or infection. The migration used by leukocytes is generally mediated by chemokines (such as CXCL12, CXCL13, and IL8) that bind to their cognate G-protein coupled receptors (GPCRs). Signaling through GPCRs can be rapidly turned off through desensitization. This allows an adaptation to the signal, and is compatible with the rapid and early migration of neutrophils and lymphocytes in response to an insult (Jin, 2013).

Chemotaxis mediated by receptor tyrosine kinases (RTKs), which tends to be slower than that of GPCR-induced chemotaxis, is generally utilized by fibroblasts and epithelial cells. Like signaling through GPCRs, RTK signaling can induce localized changes in activation of signaling molecules that result in changes in cell adhesion and cytoskeletal remodeling. A classic example of chemotactic migration in fibroblasts

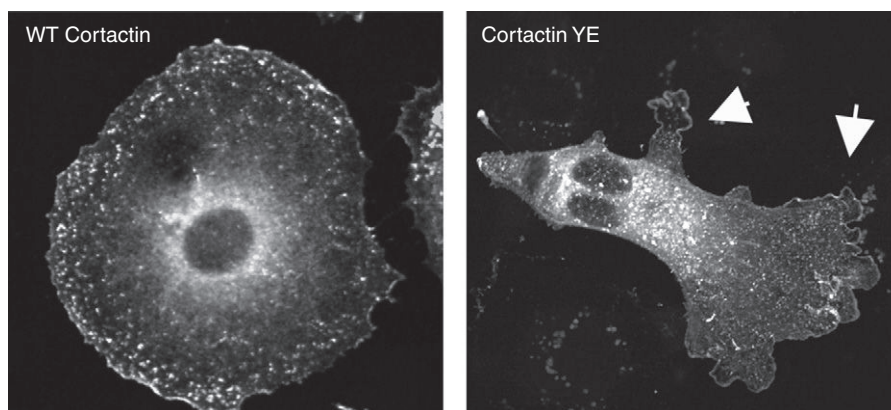


Figure 1 Cytoskeletal changes drive morphological changes and cell migration. Activation of signal transduction pathways can induce protein modifications that result in actin cytoskeletal remodeling. Here, a cultured Clone 9 rat hepatocyte is not highly migratory under steady-state conditions (left). However, expression of a mutated form of the protein cortactin, which binds and regulates actin dynamics, induces a dramatic shape change and promotes migratory behavior (right). Reproduced from Kruchten, A.E., Krueger, E.W., Wang, Y., McNiven, M.A., 2008. Distinct phospho-forms of cortactin differentially regulate actin polymerization and focal adhesions. *American Journal Physiology Cell Physiology* 295 (5), C1113–C1122, with permission from The American Physiological Society. Note the formation of multiple dynamic lamellipodia structures (arrows, described below). This mutated version of cortactin mimics phosphorylation by the kinase Src, which is a potent signaling pathway that links extracellular stimuli to cytoskeletal changes (YE=Y384E, Y429E, Y445E).

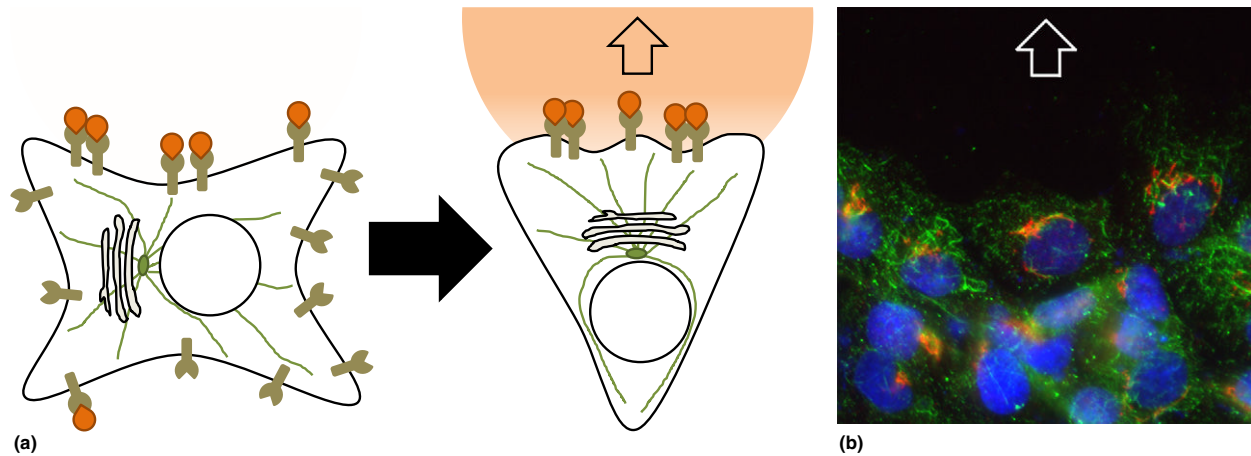


Figure 2 Directed cell migration. (a) In response to a gradient of a migratory stimulus, ligand (orange) associates with cellular receptors (brown) to induce asymmetric signal transduction. This leads to redistribution of the Golgi (gray stacks) and MTOC and microtubule cytoskeleton (green), and polarized actin cytoskeletal remodeling to drive migration in the direction of the gradient (open arrow). (b) Clone 9 rat hepatocytes exhibiting directional migration in a wound healing assay. The cells were grown to a confluent monolayer, the monolayer was then scratched, and the cells began to migrate into the gap. Migrating cells re-orient their Golgi (GM130 stain, red) and microtubules (acetylated microtubules, green) in the direction of migration (open arrow), in contrast to the cells further back, which are still nonpolarized.

occurs through the platelet-derived growth factor (PDGF). PDGF, signaling through its receptor, regulates phosphoinositide signaling through PI3K, calcium signaling through PLC γ , cell adhesion through focal adhesion kinase (FAK), and Ras and MAP kinase signal transduction (Ronnstrand and Heldin, 2001).

Haptotaxis

Whereas in chemotaxis cells migrate in response to secreted ligands, in the process of haptotaxis, cells encounter immobilized ligands, such as extracellular matrix (ECM) proteins, and these ligands promote directional migration. Cells extend protrusions that encounter such ECM proteins, including collagen, which engage proteins such as integrins on the cell membrane. Integrin activation links the ligand to the actin cytoskeleton of the migrating cell, thus regulating cellular adhesions, formation of more protrusions, and persistent migration. The molecular mechanisms of haptotaxis are similar to that of chemotaxis, but also have significant differences in regards to the machinery regulating the actin cytoskeleton (Wu *et al.*, 2012).

In a related process, cells may migrate in response to ligands that are bound in the extracellular matrix. For example, dendritic cells migrate in the direction of a gradient of the chemokine CCL21. This ligand is immobilized by binding to heparin sulfate in the extracellular matrix, and thus the dendritic cell migrates through the ECM (Weber *et al.*, 2013). While these two processes of migration following stationary cues are very similar, it is unknown if the mechanisms behind their migration are the same.

Mechanotaxis

In addition to being driven by soluble chemical signals, directional migration can also be established in response to structural or mechanical cues. These mechanical cues are often provided by the extracellular matrix, which can vary in composition and in stiffness. For example, fibroblasts tend to

migrate toward matrix of increasing stiffness, in a process called durotaxis. Migrating cells can sense mechanical properties of the matrix through integrins and focal adhesions, structures that can also exert force on the matrix (described below) (Oakes and Gardel, 2014; Plotnikov and Waterman, 2013).

Signals to Migrate During Development

Early embryonic development is a period of intensive cell division and migration, and this migration is absolutely critical for patterning and organ development. Developmental migration is precisely regulated both temporally and spatially to form the distinct germ layers of the embryo and establish specific tissue organization. In early development, cells of the neural crest undergo dramatic migration and give rise to multiple different cell types. Cells migrate either individually or collectively, with their migration regulated by secreted factors, interactions with other cells, and interactions with the extracellular environment.

Migration in neural development

As one example, during the formation of the nervous system, developing neurons can migrate over significant distances, following an extremely precise route, including turning at specific landmarks, to form the appropriate synapses (Cooper, 2013). This process is exemplified by the dynamic and highly specialized leading edge of the developing axon, called the growth cone (Vitriol and Zheng, 2012). The migration of the growth cone is guided by positive, chemoattractive cues, and also balanced by negative, chemorepulsive cues, including the Semaphorin, Slit, and ephrin families of proteins. Some protein factors, including the netrins, can be either attractive or repulsive, depending on the context. These chemical cues signal through their cognate receptors expressed on the migrating neuron to active intracellular pathways that subsequently

regulate calcium signaling and cytoskeletal changes leading to migration.

Migration in vascular development

As a second example of cell migration in development, endothelial cells undergo significant migration during both the processes of vasculogenesis (the *de novo* vascularization of an embryo) and angiogenesis, the creation of new blood vessels from existing vessels. Endothelial cells are stimulated to migrate by ligands including vascular endothelial growth factor (VEGF), the most prominent regulator of angiogenesis, and basic fibroblast growth factor (bFGF), as well as by interactions with the extracellular environment including collagen and integrins, transmembrane proteins expressed by the migrating cell (Lamallice *et al.*, 2007).

Signals to Migrate in Physiological/Pathological Processes

In addition to migration's critical developmental functions, cell migration is also essential to maintain homeostasis in the adult. Multiple cellular processes require migration, regulated by distinct chemical and mechanical signals signaling through cell-specific receptors. Even with this diverse array of upstream activators of signaling, the general mechanisms of regulation are conserved: extracellular cues are translated into intracellular signaling pathways that ultimately affect the cytoskeletal architecture, membrane dynamics, and adhesive properties that contribute to cell migration.

Wound healing

In response to wounding, there is a dramatic mobilization and recruitment of multiple cell types that will serve to repair the wound (Werner and Grose, 2003). Fibroblasts migrate into the wound in response to platelet-derived growth factor (PDGF, produced by platelets that have infiltrated the wound to promote clotting), where the fibroblasts then proliferate, and eventually lay down the collagen matrix that will be the structural support for wound healing and closure. Leukocytes including neutrophils and macrophages migrate into the wound in response to hydrogen peroxide and chemokines including RANTES/CCL5, MCP1 (CCL2), and macrophage inflammatory protein 1a (MIP1a/CCL3), respectively (Nietzhammer *et al.*, 2009; Mahdavian Delavary *et al.*, 2011). These immune cells participate in inflammation following wounding, and remove any pathogens and damaged tissue. The need to repair and regenerate blood vessels requires the pro-angiogenic migration of endothelial cells to the wound site in response to VEGF and FGF family ligands. Finally, migration of keratinocytes is required to form the new protective barrier that will ultimately close the wound (Raja *et al.*, 2007). Keratinocyte migration is stimulated by keratinocyte growth factor (KGF, in the FGF family), insulin-related growth factors, and chemicals such as nitric oxide. Keratinocytes stop migrating in response to contact inhibition, when in contact with other cells.

The immune response

Leukocytes depend on motility for a competent, systemic immune response. These cells must migrate through the

vasculature and migrate across cellular barriers in order to reach damaged or infected tissues. While in circulation, leukocytes engage in a rolling-type migration, whereby they weakly adhere to the endothelial wall through interactions between transient interactions between selectins on the endothelium and their ligands on the leukocytes (Zarbock *et al.*, 2011). To stimulate migration out of the circulation and into the target tissue, signals from the target tissue come in the form of chemokines, such as IL-1 β , TNF α , and other factors. These chemokines trigger adhesive changes in both the endothelium and leukocytes to arrest the migration of the leukocyte in the circulation. The immune cells then undergo the invasive and dynamic process of transendothelial migration called diapedesis (Heemskerk *et al.*, 2014; Muller, 2011). To cross the endothelial barrier out of the circulation, the leukocytes undergo significant deformation to pass through the gaps between cells. While few cell types are able to undergo this aggressive transmigration, the mechanisms of cytoskeletal regulation allowing the cellular changes are similar to other mechanisms of cell migration, including polarization, actin cytoskeletal remodeling, and membrane dynamics.

Cytoskeletal Changes Underlying Migration

Overview

Even with the diverse array of signals that can stimulate cell migration in different contexts, the general mechanisms for translating these motogenic signals into mechanical changes that promote cellular locomotion are conserved among cell types and species. Cell migration is primarily driven by cytoskeletal and membrane changes related to the dynamic remodeling of the actin cytoskeleton (Graziano and Weiner, 2014; Ridley, 2011). This results in the formation of transient structures that can propel a cell forward by one of several major mechanisms.

Regulation of Migration by Rho Family GTPases

The formation of dynamic, actin-based migratory structures is regulated by the activation of a family of small GTP-binding proteins, the Rho family of GTPases. These proteins, including Rac1,2,3, Cdc42, and Rho, are small, 21-kDa proteins that function as molecular switches (Figure 3). In their active form, they bind to GTP, and this GTP-bound form is able to bind to effector proteins. These effectors only bind the GTP-bound form of the protein. Once the GTP is hydrolyzed to GDP, the effectors are no longer able to bind the complex. Thus, the GTP-state of the Rho GTPases is a potent way to regulate cytoskeletal remodeling and migration.

Activation of Rho GTPases

Rho family GTPases are activated downstream of the receptor/ligand interactions described above. Rho GTPases do not have intrinsically high GTP hydrolysis activity, and thus, by themselves, would not change their activation state very readily. Rather, their GTP binding, and hence activation, is regulated by a host of associated proteins (Figure 3). Guanine nucleotide exchange factors (GEFs) are a family of proteins that

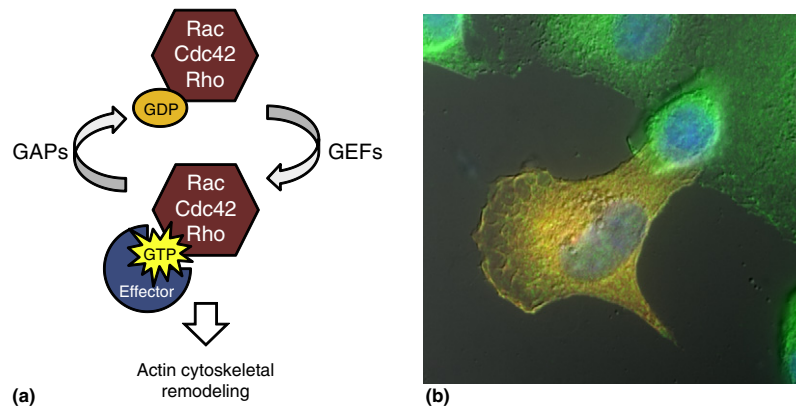


Figure 3 Rho GTPase activation cycle. (a) The Rho GTPases Rac, Cdc42, and Rho are molecular switches that are active when bound to GTP. This allows binding to effector molecules leading to polymerization or branching of the actin cytoskeleton. Rho family GTPase activation is catalyzed by GEFs, which promote GTP binding. Inactivation of Rho GTPases is due to GTP hydrolysis, regulated by GAPs. (b) Activation or overexpression of GEFs can induce GTPase activation and cell migration. A PANC1 pancreatic cancer cell overexpressing the GEF Vav1 (yellow cell) has dramatically upregulated Rac activation, lamellipodia formation, and increased migratory capacity compared to the neighboring, nonexpressing cells (green).

facilitate the release of GDP from the small G-protein, which allows the binding of GTP, which is at much higher cytoplasmic concentrations, thereby promoting activation of the GTPase. Therefore, GEFs are potent activators of Rho family GTPases. Guanine nucleotide associated proteins, or GAPs, promote the hydrolysis of GTP, leaving the protein in the inactive, GDP-bound state. Finally, GDI proteins (Rho GDP-dissociation inhibitors) can also regulate GDP/GTP exchange, as well as the targeting or localization of specific Rho family GTPases.

As activators of GTPases, GEFs have received considerable attention for a role in driving cell migration (Goicoechea *et al.*, 2014). Over 40 different GEF proteins have been identified for the Rho family. Some have cell-type specific expression patterns, but it is undefined how signals from so many different GEFs are regulated, and which are functionally redundant. GEFs contain domains that perform the exchange, such as a Dbl homology domain (DH) in the Dbl family of GEFs, or a DHR2 domain in the DOCK family. These domains coordinate the binding of GDP and a magnesium ion on the GTPase, causing a conformational change to release the GDP. The exchange activity of GEFs can be regulated acutely to lead to responsive changes in the activation of Rho GTPases. GEFs can be activated by posttranslational modifications such as phosphorylation, by changing the localization or membrane association of the GEF within the cell via phosphoinositide binding, or by regulating the protein stability. Finally, many GEFs have been found to be overexpressed in cancer, in which an upregulation of cell migration contributes to metastasis, the spread of cancer cells throughout the body.

Signaling downstream of Rho family GTPases

Active, GTP-bound Rho family GTPases bind directly to effector proteins that relay the activating signal to downstream proteins that remodel the actin cytoskeleton. For example, active Rac/Cdc42 bind to the p21 activated kinase, or PAK. This kinase has multiple substrates that regulate cell survival and migration, including a subunit of the Arp2/3 complex (Rotty *et al.*, 2013). This protein complex induces actin

nucleation and branching, thereby contributing to the formation of cellular protrusions including lamellipodia. Rho GTPases also lead to the activation of WASP/WAVE proteins, which also regulate actin dynamics via Arp2/3 (Takenawa and Suetsugu, 2007). Actin remodeling contributing to cell migration is a coordinated process involving the severing of existing filaments to generate new barbed ends, the action of actin capping proteins to restrict polymerization, the branching mediated by Arp2/3, and the bundling and cross-linking of actin filaments.

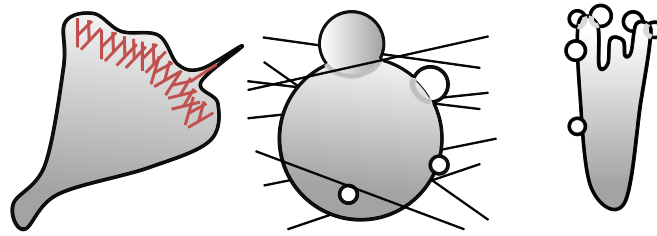
Interestingly, there is considerable cross-talk between Rho and Rac/Cdc42. Generally speaking, Rho signaling opposes Rac/Cdc42, and vice versa. Each can promote a distinct type of cell migration, and the effectors of one GTPase may directly inhibit the activation of the other through downstream effector molecules.

Different Types of Cellular Migration

Depending on the cell type and on the environment, cells can utilize multiple types of migration. These each utilize different types of migratory structures, and are activated by different signaling pathways. Importantly, cells can rapidly switch from one style of migration to another depending on the extracellular environment, including matrix density and tension, the type of substrate, or the structural barrier (Figure 4; Petrie and Yamada, 2012).

Lamellipodial migration

Mesenchymal-type cells, such as fibroblasts, migrate forward through the formation of cellular protrusions called lamellipodia. In cell culture on a flat substrate, these structures are broad and flat (the lamella) with thin, dynamic leading edges (the lamellipodia) that are the sites of active actin polymerization and branching. Cells may also extend filopodia, thin, narrow protrusions that are thought to be involved in sensing the environment. Mesenchymal migration requires the destruction or remodeling of the extracellular matrix in order to allow the cell to pass (Bisi *et al.*, 2013; Ridley, 2011).



Type of migration	Lamellipodial	Amoeboid	Lobopodial
GTPase regulation	Polarized Rac/Cdc42 activation at leading edge	RhoA-mediated Myosin II activity at blebs	Nonpolarized RhoA-mediated Myosin II activity
Actin structures	Lamellipodia and Filopodia at leading edge; Rho-mediated stress fibers	Actomyosin contraction. No stress fibers	Actomyosin contraction Rho-mediated stress fibers.
Adhesion to ECM?	High – focal contacts	Low	High – focal contacts
Requires matrix degradation?	Yes	No	No

Figure 4 Modes of cell migration in 3D. Fibroblasts can rapidly switch among cell migration methods depending on the extracellular matrix environment. Modeled after [Sixt \(2012\)](#).

Lamellipodial migration is regulated by the small G-protein Rac1. GTP-bound Rac signals downstream to effectors including PAK, N-WASP, and WAVE proteins. These proteins promote branching of the actin cytoskeleton that is driven by the Arp2/3 protein complex, which introduces a new core of nucleation along an existing actin filament. The new actin filament branches at a 70 degree angle, and this polymerization and branching drives the formation of the lamellipodia.

Filopodial formation is regulated by the small G-protein Cdc42. This signals to activate WASP/N-WASP, which then activates actin polymerization and extensions of parallel actin bundles into the protrusive filopodia, with polymerization occurring at the filopodial tip. The filaments are cross-linked by actin binding proteins such as fascin.

Amoeboid migration

Cells are also able to migrate in a way that causes significant deformation of the cell in order to squeeze through gaps in the matrix. This is called amoeboid migration, or blebbing migration, as it resembles the streaming movement of an amoeba through the formation of membrane blebs. Amoeboid migration does not require proteolytic cleavage of the matrix. Rather, the cell finds a path through the ECM using blebbing protrusions. Thus, amoeboid migration can be activated rapidly and utilized in the absence of proteolytic matrix degradation ([Wolf et al., 2003](#)).

In contrast to mesenchymal, lamellipodial-based migration, which is regulated by Rac/Cdc42, amoeboid migration is regulated by signaling through RhoA, which is associated with force generation ([Lessey et al., 2012](#)). Whereas the extension of lamellipodia is driven by actin polymerization and branching, amoeboid migration is regulated instead by actin–myosin-based contraction. Actin–myosin contraction is best known for its role in muscle contraction, where myosin heads make and break contacts with actin to ‘walk’ along the actin and induce

contraction. Similarly, non-muscle myosins can generate contraction forces to change cell tension and shape, thereby inducing a mechanical component. Rho signals to the kinase ROCK, which promotes the phosphorylation and activation of the myosin light chain kinase (MLCK). MLCK then promotes Myosin II activity to regulate actin–myosin contraction. This results in the formation of constriction rings to form blebs, followed by the streaming of cytoplasm through the contracted neck.

Lobopodial migration

More recently, a distinct type of cell migration was described through the formation of structures named lobopodia. These structures, previously described in amoebas, were found in migrating fibroblasts in 3D cultures. Under matrix conditions of linear elasticity, fibroblasts can switch to lobopodial-style migration. This structure requires high Rho activity and mechanical stress, like amoeboid migration, but also require adhesion with the extracellular matrix, like mesenchymal movement ([Petrie et al., 2012](#)).

Cell Migration Is Regulated Through Interactions with the Environment

Overview

While many studies of cell migration have been done using two-dimensional cell culture on a glass or homogenous substrate, cell migration *in vivo* in a three-dimensional environment occurs in the presence of many more physical barriers. Cells *in vivo* exist in the context of a complex extracellular matrix, and in order to migrate, must break and reform contacts with the ECM. Further, cells often need to migrate through the ECM, which may require degradation or remodeling of the extracellular matrix to allow their passage.

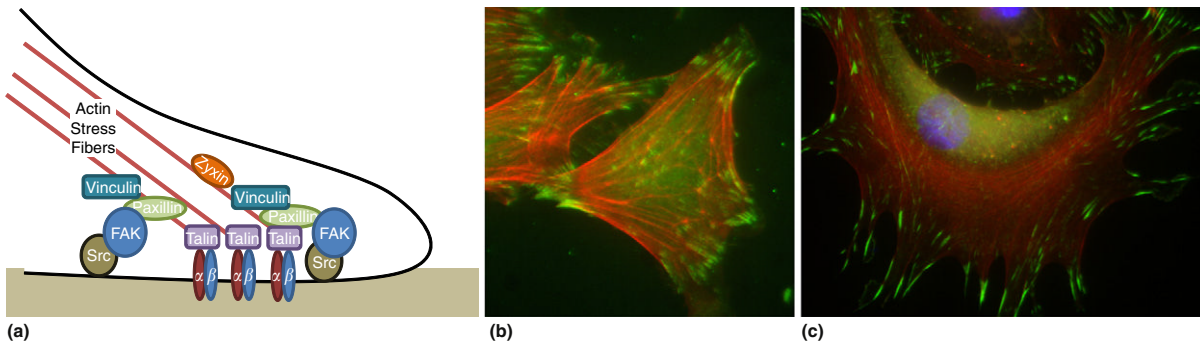


Figure 5 Focal adhesions. (a) The focal adhesion mediates contact between the cell and the extracellular matrix through integrin heterodimers (α , β). Over 150 proteins have been identified in the focal complex: examples are shown, including some proteins that couple integrin signaling to the actin cytoskeleton. Rat fibroblasts (b) or human fibroblasts (c) stained for paxillin (green) and actin (red) show prominent paxillin-positive focal adhesions that are anchoring points for actin stress fibers.

Finally, in addition to contacts with the extracellular matrix, cell migration is also regulated by interactions with other cells, which may work in concert to migrate collectively, rather than as single cells.

Focal Adhesions

Cells form specialized structures called focal adhesions that link the extracellular matrix to the cytoskeleton. Focal adhesions consist of many structural and mechanosensing proteins clustered around integrins, transmembrane proteins that bind ECM proteins on the outside of the cell, such as fibronectins (Figure 5; Huttenlocher and Horwitz, 2011; Geiger *et al.*, 2009). Inside the cell, the focal adhesion is comprised of multiple proteins including paxillin, vinculin, and talin, which are linked to actin fibers. Thus, focal adhesions are able to both sense and respond to mechanical stress from inside and outside the cell (Ciobanasi *et al.*, 2013; Kuo, 2013; Schwarz and Gardel, 2012). Focal adhesions are also regulated by the microtubule cytoskeleton, which are proposed to be involved in targeted trafficking to and from the focal adhesion (Stebbens and Wittmann, 2012). Focal adhesions form as nascent focal complexes, which then mature to larger, more stable focal adhesions. These structures form in both two-dimensional and three-dimensional contexts, and stabilize the cell's position and modulate migration (Figure 5).

To begin cell migration, cells must break these focal adhesions (Nagano *et al.*, 2012; Wehrle-Haller, 2012). This occurs in response to signal transduction pathways activated by receptor tyrosine kinases, including phosphorylation by the kinase Src and focal adhesion kinase (FAK). These potent kinases phosphorylate components of the focal adhesion and lead to its disassembly. In addition, the integrin complex is internalized by endocytosis, thereby removing the integrin from the cell surface and dissociating the interaction with the ECM.

Matrix Remodeling During Migration

Migrating cells often need to cross a barrier to migrate to their appropriate destination. Thus, cell migration may be coupled with degradation or remodeling of the extracellular matrix. In particular, mesenchymal style migration is associated with

matrix destruction. Some cells, such as macrophages and osteoclasts, form specialized structures called podosomes that are based on the actin cytoskeleton and serve as sites of localized protease activity to degrade extracellular substrates. In cancer cells, similar structures, called invadopodia, are regulated slightly differently and have different dynamics, but accomplish a similar function, which is the degradation of the matrix to allow migration across barriers (Figure 6). In addition, focal adhesions have also been described as a matrix-degrading structure. The molecular components of focal adhesions and invadopodia are highly similar, and degradation of the ECM is likely coupled with migratory structures including lamellipodial protrusions and focal adhesions (McNiven, 2013; Murphy and Courtneidge, 2011; Hoshino *et al.*, 2013).

Matrix degradation is accomplished by secreted or membrane-tethered proteases, proteins that can specifically degrade extracellular matrix proteins such as collagens or fibronectins. These proteases include the matrix metalloproteinases (MMPs) and the ADAM proteases. These proteases are highly regulated by both positive regulators, which may cleave and activate the proteases, and by negative regulators, including the TIMPs for matrix metalloproteinases. Matrix degradation allows invasive cell migration, but must be tightly regulated to maintain tissue integrity and prevent aberrant migration. Matrix-degrading proteases such as the MMPs are frequently overexpressed in tumors, which contribute to metastatic invasion (Lu *et al.*, 2011; Kessenbrock *et al.*, 2010).

In addition to matrix degradation, cells such as fibroblasts produce new matrix. This matrix remodeling is important in replacing matrix that was lost during invasion or injury, and can serve as tracks for migration. For example, invasive tumor cells can migrate along collagen fibers, which are produced by cancer-associated fibroblasts that are activated by their association with tumor cells (Egeblad *et al.*, 2010).

Collective Cell Migration

Many of the pathways described here regulate the migration of individual cells. Cells are also able to migrate collectively, or in groups, while still maintaining contact with other cell (Figure 7; Rorth, 2012; Friedl and Gilmour, 2009; Etienne-Manneville, 2014). This collective cell migration is common during development, when entire sheets of cells migrate to

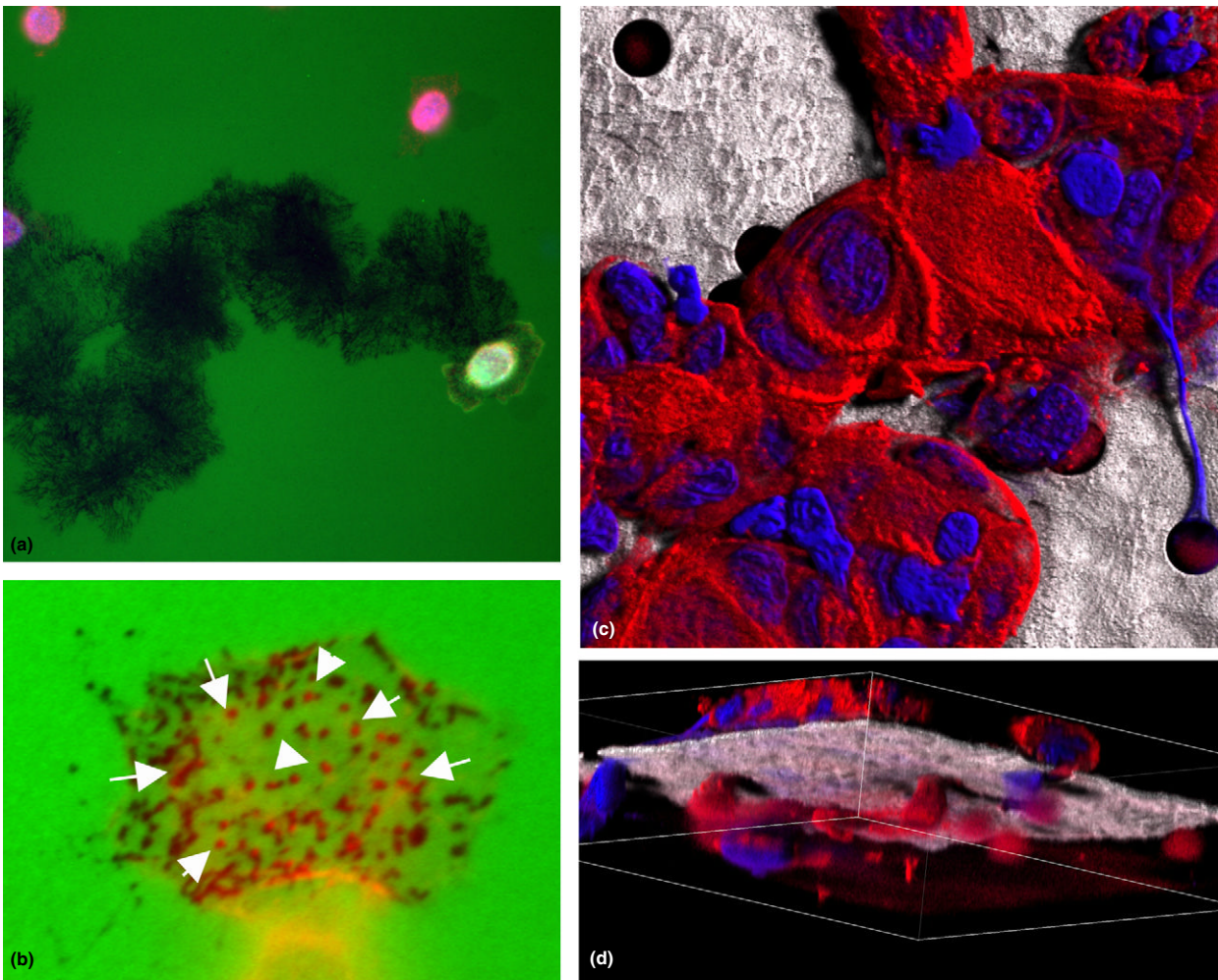


Figure 6 Matrix remodeling during migration. Degradation of the extracellular matrix is often required for cell migration *in vivo*. (a) Matrix degradation can be measured using cultured cells by plating cells on a fluorescent gelatin substrate, and measuring the loss of fluorescence (blackened areas). Here, a BxPC3 pancreatic cancer cell plated on a fluorescent gelatin matrix leaves a trail of degraded matrix in its wake as it migrates across the substrate. (b) Matrix degradation can be accomplished by actin-based protrusions called invadopodia (positive for actin, red, and arrows), that are sites for protease targeting to degrade the fluorescent gelatin matrix. (c) HPAF-II pancreatic cancer cells degrade matrix and migrate across a transwell filter, in an *in vitro* approximation of invasion that couples matrix degradation and migration. The image is a slice from a 3D rendering of a z-series using confocal microscopy of HPAF-II cells stained for actin (red) and DAPI (blue) on the underside of a filter in a transwell assay. (d) 3D reconstruction from confocal microscopy showing HPAF-II cells on the top and bottom of a transwell filter. In (c) and (d), note that the cells have degraded a barrier and migrated through the small pores in the filter (black circles).

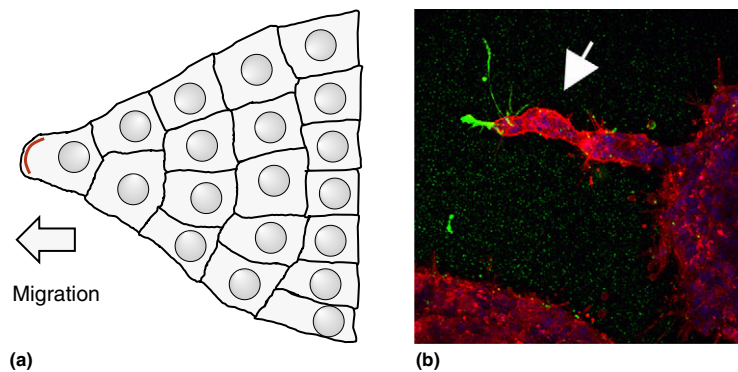


Figure 7 Collective cell migration. (a) During collective migration, many cells influence one another as they all migrate in the same direction, such as a sheet or stream of cells that remain attached through cell-cell junctions. The population is comprised of one or more 'leader' cells, which utilize different signaling pathways than the follower cells (shown by red line). (b) PANC1 pancreatic cancer cells plated in a 3D matrix of Collagen I/Matrigel exhibit collective migration away from the spheroid of cells. Multiple cells follow a leader cell that degrades the ECM with α -actinin-rich protrusions. Green: α -actinin, Red: Actin, Blue: DAPI.

support the formation of tissues. The migrating cells may be tightly held together by adherens junctions, or may be more loosely associated, yet still migrating as a population. During collective migration, the cluster of migrating cells may display some of the same polarization and molecular signaling events as during single cell migration, but with different cells in the group exhibiting distinct properties. For example, one or more cells may serve as the leader cell, extending protrusions and responding to stimuli from the extracellular environment. Leader cells have different gene expression, distinct cell shape, and activation of differential signaling molecules compared to the follower cells. This process of collective cell migration is highly conserved across species, and is also utilized by tumor cells during metastatic migration.

See also: Cytoskeleton and Motors: Complex Cytoskeletal Structures; Cell Migration; Cell Polarity; Rho GTPases

References

- Bear, J.E., Haugh, J.M., 2014. Directed migration of mesenchymal cells: Where signaling and the cytoskeleton meet. *Current Opinion in Cell Biology* 30C, 74–82.
- Bisi, S., Disanza, A., Malinverno, C., *et al.*, 2013. Membrane and actin dynamics interplay at lamellipodia leading edge. *Current Opinion in Cell Biology* 25, 565–573.
- Ciobanaru, C., Faivre, B., Le Clainche, C., 2013. Integrating actin dynamics, mechanotransduction and integrin activation: The multiple functions of actin binding proteins in focal adhesions. *European Journal of Cell Biology* 92, 339–348.
- Cooper, J.A., 2013. Cell biology in neuroscience: Mechanisms of cell migration in the nervous system. *Journal of Cell Biology* 202, 725–734.
- Egeblad, M., Rasch, M.G., Weaver, V.M., 2010. Dynamic interplay between the collagen scaffold and tumor evolution. *Current Opinion in Cell Biology* 22, 697–706.
- Etienne-Manneville, S., 2014. Neighborly relations during collective migration. *Current Opinion in Cell Biology* 30C, 51–59.
- Friedl, P., Gilmour, D., 2009. Collective cell migration in morphogenesis, regeneration and cancer. *Nature Reviews Molecular Cell Biology* 10, 445–457.
- Geiger, B., Spatz, J.P., Bershadsky, A.D., 2009. Environmental sensing through focal adhesions. *Nature Reviews Molecular Cell Biology* 10, 21–33.
- Goicoechea, S.M., Awadia, S., Garcia-Mata, R., 2014. I'm coming to GEF you: Regulation of RhoGEFs during cell migration. *Cell Adhesion & Migration* 8, 535–549.
- Graziano, B.R., Weiner, O.D., 2014. Self-organization of protrusions and polarity during eukaryotic chemotaxis. *Current Opinion in Cell Biology* 30C, 60–67.
- Heemskerk, N., Van Rijssel, J., Van Buul, J.D., 2014. Rho-GTPase signaling in leukocyte extravasation: An endothelial point of view. *Cell Adhesion & Migration* 8 (2), 67–75.
- Hoshino, D., Branch, K.M., Weaver, A.M., 2013. Signaling inputs to invadopodia and podosomes. *Journal of Cell Science* 126, 2979–2989.
- Huttenlocher, A., Horwitz, A.R., 2011. Integrins in cell migration. *Cold Spring Harbor Perspectives in Biology* 3, a005074.
- Jin, T., 2013. Gradient sensing during chemotaxis. *Current Opinion in Cell Biology* 25, 532–537.
- Kessenbrock, K., Plaks, V., Werb, Z., 2010. Matrix metalloproteinases: Regulators of the tumor microenvironment. *Cell* 141, 52–67.
- Kuo, J.C., 2013. Mechanotransduction at focal adhesions: Integrating cytoskeletal mechanics in migrating cells. *Journal of Cellular and Molecular Medicine* 17, 704–712.
- Lamalice, L., Le Boeuf, F., Huot, J., 2007. Endothelial cell migration during angiogenesis. *Circulation Research* 100, 782–794.
- Lessey, E.C., Guilluy, C., Burridge, K., 2012. From mechanical force to RhoA activation. *Biochemistry* 51, 7420–7432.
- Lu, P., Takai, K., Weaver, V.M., Werb, Z., 2011. Extracellular matrix degradation and remodeling in development and disease. *Cold Spring Harbor Perspectives in Biology* 3, a005058.
- Mahdavian Delavary, B., Van Der Veer, W.M., Van Egmond, M., Niessen, F.B., Beelen, R.H., 2011. Macrophages in skin injury and repair. *Immunobiology* 216, 753–762.
- Majumdar, R., Sixt, M., Parent, C.A., 2014. New paradigms in the establishment and maintenance of gradients during directed cell migration. *Current Opinion in Cell Biology* 30C, 33–40.
- McNiven, M.A., 2013. Breaking away: Matrix remodeling from the leading edge. *Trends in Cell Biology* 23, 16–21.
- Muller, W.A., 2011. Mechanisms of leukocyte transendothelial migration. *Annual Review of Pathology* 6, 323–344.
- Murphy, D.A., Courtneidge, S.A., 2011. The 'ins' and 'outs' of podosomes and invadopodia: Characteristics, formation and function. *Nature Reviews Molecular Cell Biology* 12, 413–426.
- Nagano, M., Hoshino, D., Koshikawa, N., Akizawa, T., Seiki, M., 2012. Turnover of focal adhesions and cancer cell migration. *International Journal of Cell Biology* 2012, 310616.
- Niethammer, P., Grabher, C., Look, A.T., Mitchison, T.J., 2009. A tissue-scale gradient of hydrogen peroxide mediates rapid wound detection in zebrafish. *Nature* 459, 996–999.
- Oakes, P.W., Gardel, M.L., 2014. Stressing the limits of focal adhesion mechanosensitivity. *Current Opinion in Cell Biology* 30C, 68–73.
- Petrie, R.J., Gavarra, N., Chadwick, R.S., Yamada, K.M., 2012. Nonpolarized signaling reveals two distinct modes of 3D cell migration. *Journal of Cell Biology* 197, 439–455.
- Petrie, R.J., Yamada, K.M., 2012. At the leading edge of three-dimensional cell migration. *Journal of Cell Science* 125, 5917–5926.
- Plotnikov, S.V., Waterman, C.M., 2013. Guiding cell migration by tugging. *Current Opinion in Cell Biology* 25, 619–626.
- Raja, S., Sivamani, K., Garcia, M.S., Isseroff, R.R., 2007. Wound re-epithelialization: Modulating keratinocyte migration in wound healing. *Frontiers in Bioscience* 12, 2849–2868.
- Ridley, A.J., 2011. Life at the leading edge. *Cell* 145, 1012–1022.
- Ronnstrand, L., Heldin, C.H., 2001. Mechanisms of platelet-derived growth factor-induced chemotaxis. *International Journal of Cancer* 91, 757–762.
- Rorth, P., 2011. Whence directionality: Guidance mechanisms in solitary and collective cell migration. *Developmental Cell* 20, 9–18.
- Rorth, P., 2012. Fellow travellers: Emergent properties of collective cell migration. *EMBO Press* 13, 984–991.
- Rotty, J.D., Wu, C., Bear, J.E., 2013. New insights into the regulation and cellular functions of the ARP2/3 complex. *Nature Reviews Molecular Cell Biology* 14, 7–12.
- Schwarz, U.S., Gardel, M.L., 2012. United we stand: Integrating the actin cytoskeleton and cell-matrix adhesions in cellular mechanotransduction. *Journal of Cell Science* 125, 3051–3060.
- Sixt, M., 2012. Cell migration: Fibroblasts find a new way to get ahead. *Journal of Cell Biology* 197, 347–349.
- Stebens, S., Wittmann, T., 2012. Targeting and transport: How microtubules control focal adhesion dynamics. *Journal of Cell Biology* 198, 481–489.
- Takenawa, T., Suetsugu, S., 2007. The WASP-WAVE protein network: Connecting the membrane to the cytoskeleton. *Nature Reviews Molecular Cell Biology* 8, 37–48.
- Vitriol, E.A., Zheng, J.Q., 2012. Growth cone travel in space and time: The cellular ensemble of cytoskeleton, adhesion, and membrane. *Neuron* 73, 1068–1081.
- Weber, M., Hauschild, R., Schwarz, J., *et al.*, 2013. Interstitial dendritic cell guidance by haptotactic chemokine gradients. *Science* 339, 328–332.
- Wehrle-Haller, B., 2012. Assembly and disassembly of cell matrix adhesions. *Current Opinion in Cell Biology* 24, 569–581.
- Werner, S., Grose, R., 2003. Regulation of wound healing by growth factors and cytokines. *Physiological Reviews* 83, 835–870.
- Wolf, K., Mazo, I., Leung, H., *et al.*, 2003. Compensation mechanism in tumor cell migration: Mesenchymal-ameboid transition after blocking of pericellular proteolysis. *Journal of Cell Biology* 160, 267–277.
- Wu, C., Asokan, S.B., Berginski, M.E., *et al.*, 2012. Arp2/3 is critical for lamellipodia and response to extracellular matrix cues but is dispensable for chemotaxis. *Cell* 148, 973–987.
- Zarbock, A., Ley, K., McEver, R.P., Hidalgo, A., 2011. Leukocyte ligands for endothelial selectins: Specialized glycoconjugates that mediate rolling and signaling under flow. *Blood* 118, 6743–6751.

CELL COMMUNICATION: PROTOTYPIC INTEGRATIVE PROCESSES

Contents

Wound Healing: An Orchestrated Process of Cell Cycle, Adhesion, and Signaling

Cancer Cell Invasion through Tissue Barriers

Neuronal Transport and Spatial Signaling Mechanisms in Neural Repair

Wound Healing: An Orchestrated Process of Cell Cycle, Adhesion, and Signaling

S Werner and M Antsiferova, Institute of Molecular Health Sciences, Zurich, Switzerland

© 2016 Elsevier Inc. All rights reserved.

Glossary

Angiogenesis Formation of new blood vessels by sprouting of existing vessels.

Blastema Group of cells capable of proliferation and regeneration into organs or body parts.

Carcinoma Cancer originating from epithelial cells.

Dermis Layer of the skin consisting mainly of connective tissue with embedded fibroblasts, some immune cells, and blood and lymphatic vessels. It is located below the epidermis and above the subcutaneous adipose tissue.

Epidermis The most upper layer of the skin, which includes keratinocytes at different stages of differentiation, some immune cells, as well as Merkel cells. The epidermis forms an inside-out and outside-in barrier to the environment.

Fibroplasia Fibroblast accumulation and their deposition of extracellular matrix.

Granulation tissue A cell-rich and highly vascularized tissue that replaces the fibrin clot in a skin wound.

Hemidesmosome Attachment site between the basal surface of certain epithelial cells and the underlying

basement membrane. Hemidesmosomes are anchored to the keratin cytoskeleton of epithelial cells.

Keloid Excessive scar tissue that extends the original boundaries of the wound.

Matrix metalloproteinases Zinc-dependent endopeptidases that cleave components of the extracellular matrix and also various other proteins.

Myofibroblast A fibroblast with features of a smooth muscle cell, including contractile properties.

Reactive oxygen species Highly reactive molecules that are formed upon incomplete one-electron reduction of molecular oxygen.

Reepithelialization Coverage of the injured skin with a new epidermis.

Scar Connective tissue replacement following wounding of the dermis.

Ulcer An inflammatory lesion of the skin or a mucous membrane that is characterized by loss of surface tissue, disintegration, and necrosis of epithelial tissue, and often pus.

Molecular and Cellular Biology of Normal Wound Healing and Scar Formation

Injury to the skin of adult mammals rapidly initiates a complex and well-coordinated series of events that ultimately result in repair of the wounded tissue. The wound healing process involves three major phases, which are partially overlapping. They include (1) formation of a blood clot and initiation of an inflammatory response, (2) formation of new tissue, including repair of the injured epidermis (reepithelialization) and replacement of the injured dermis with new

tissue (granulation tissue formation) and finally (3) tissue remodeling ([Figures 1–3](#)). Here we describe the different events that occur during these phases in a normally healing wound, and we also mention what goes awry in impaired and/or excessive wound healing. Finally, we describe the interesting parallels between wound healing and cancer development.

Blood Clotting and Inflammation

Shortly after an injury that comprises the epidermis and the dermis and thus involves vascular damage, a platelet plug and

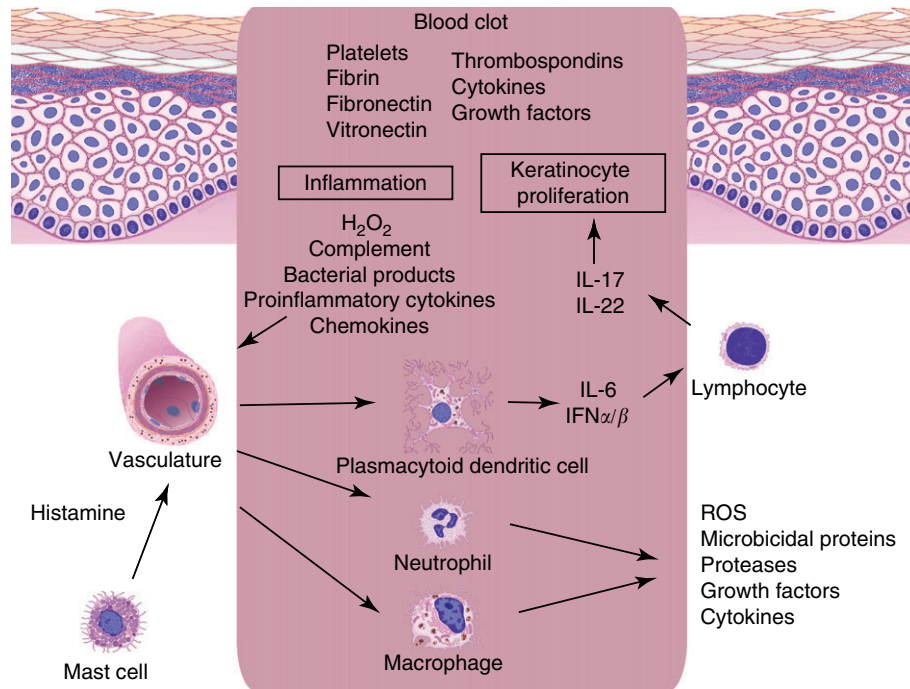


Figure 1 Schematic representation of an early skin wound in the phase of blood clotting and inflammation. Shortly after injury, blood clot formation prevents excessive bleeding. The clot also serves as a transient barrier and provides a matrix for the invasion of immune cells from the circulation and for migration of resident cells from the wound edge. Several immune cells subsequently invade into the clot. They secrete mediators, which are important for antimicrobial defense and which also initiate reepithelialization and granulation tissue formation. Major cytokines and growth factors orchestrating this phase are indicated.

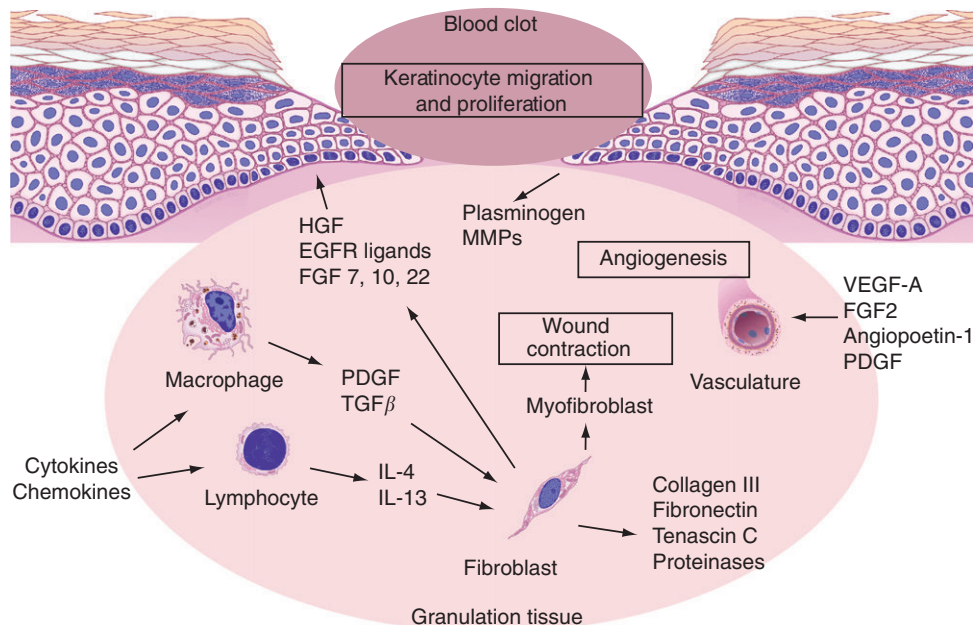


Figure 2 Schematic representation of a skin wound in the phase of new tissue formation. Three to ten days after wounding, different immune cells are present in the wound granulation tissue and massive angiogenesis occurs. Fibroblasts from the wound edge or fibroblast precursor cells from the bone marrow migrate into the wound. They proliferate and deposit extracellular matrix proteins, and many of them differentiate into contractile myofibroblasts. Keratinocytes from the injured epidermis and hair follicles at the wound edge migrate along the injured dermis and above the provisional matrix. Subsequently, hyperproliferation of the keratinocytes behind the migrating tongue occurs. Major cytokines and growth factors orchestrating this phase are indicated.

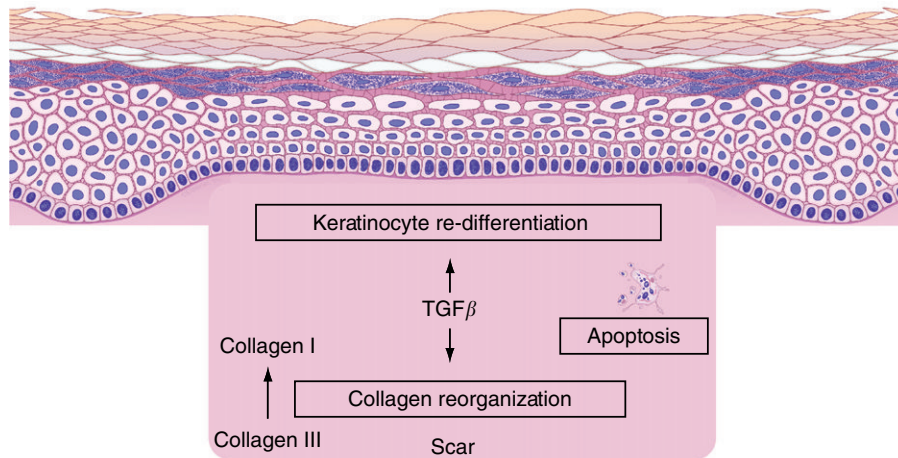


Figure 3 Schematic representation of a skin wound in the phase of tissue remodeling. One to two weeks after injury, wound reepithelialization is completed and the keratinocytes redifferentiate and reestablish an efficient barrier. Concomitantly, the cellular density of the granulation tissue gradually declines and the extracellular matrix is remodeled. This results in formation of a scar with reduced tensile strength and elasticity and lack of all appendages. TGF- β is considered as the major regulator of scar formation.

a blood clot are formed that prevent excessive and prolonged bleeding, severe water loss, contamination of the wound with microorganisms and access of toxic compounds to the injured skin. The clot consists mainly of cross-linked fibrin and plasma-derived fibronectin that is released upon injury of blood vessels, but it also includes some additional proteins such as vitronectin and thrombospondins. Furthermore, it is an important reservoir of various cytokines and growth factors that are required for the healing process, and it acts as a matrix for different cell types that are attracted to the wound from the adjacent normal skin or from the circulation (Figure 1; Martin, 1997; Gurtner *et al.*, 2008).

Inflammatory cells are then attracted to the wound site by factors produced upon complement activation or bacterial degradation or which are released from degranulation of platelets. Data from lower organisms, including *Drosophila* and zebrafish, revealed that production of hydrogen peroxide at the wound site via the action of NADPH oxidases also contributes to inflammatory cell recruitment immediately after injury (Niethammer *et al.*, 2009; Moreira *et al.*, 2010). Although a role of hydrogen peroxide in the attraction of immune cells has not been demonstrated yet for mammals, the high conservation of this process among other species and the strong expression of NADPH oxidases at the wound site in mammals, also suggest a contribution of reactive oxygen species (ROS) to the inflammatory response in mammalian wounds. At later stages of the wound healing process, chemokines/cytokines produced and/or released by various cells in the wounded tissue are particularly important for the attraction of immune cells (Figure 1). The first cells that arrive at the wound site are neutrophils due to their abundance in the circulation. They are followed by monocytes that subsequently differentiate into dendritic cells or tissue macrophages. Plasmacytoid dendritic cells (pDCs), which are specialized type I interferon-producing cells, were recently found to infiltrate mouse and human skin wounds 24 h after injury (Gregorio *et al.*, 2010). In response to nucleic acids released from damaged cells, pDCs transiently produce type I interferons and interleukin (IL)-6, and they

induce the expression of IL-17 and IL-22 in T cells, thereby promoting the inflammatory response as well as wound reepithelialization. The invasion of immune cells, in particular during the early phase of repair, is facilitated by the vasodilation and high vascular permeability that is seen in early wounds. This vascular phenotype results at least in part from histamine that is released upon degranulation of mast cells at the wound edge after injury (Figure 1). While mast cells are obviously activated adjacent to the wound, they are only recruited into the wound tissue during the late phase of the repair process (Antsiferova *et al.*, 2013; Willenborg *et al.*, 2014).

The attraction of neutrophils is particularly important for the defense against invading microorganisms, since these cells are a potent source of ROS and of proteins that help to defend microorganisms, in particular proteinases (Figure 1). However, they are not required for the normal wound healing process under sterile conditions, and a prolonged presence of neutrophils results in tissue damage and subsequent impairment of healing. Consistent with this assumption, high neutrophil content is a hallmark of chronic, non-healing ulcers (Martin and Leibovich, 2005; Eming *et al.*, 2007a).

In contrast to neutrophils, there is strong evidence for an important role of macrophages in the control of different stages of the healing process in adult mammals, and their presence is likely to be essential for the healing of large skin defects. Like neutrophils, they are required for the defense against invading microorganisms through their capacity to secrete proteinases and ROS and through their potent phagocytic function. Furthermore, they are a rich source of cytokines and growth factors. Interestingly, however, small wounds were shown to heal in the absence of neutrophils and macrophages under sterile conditions and the quality of the healed tissue was even improved. The most likely explanation for this beneficial effect is the reduced production of toxic products such as ROS in the absence of these innate immune cells, resulting in reduced cell damage. Healing of larger wounds, however, seems to depend on the release of growth factors by macrophages, which stimulate angiogenesis as well as

migration and proliferation of fibroblasts and keratinocytes (Martin and Leibovich, 2005; Eming *et al.*, 2007b).

An important role of mast cells for normal wound healing has been suggested by several studies. However, these experiments were performed with mast cell-deficient mice, which have additional abnormalities due to the lack of a functional c-Kit receptor. The recent generation of mast cell-deficient mice with wild-type c-Kit, which do not reveal any other abnormalities under unchallenged conditions, allowed a more precise analysis of mast cell function in wound healing. Surprisingly, these studies revealed that mast cells are dispensable for the healing of full-thickness excisional wounds in adult mice (Antsiferova *et al.*, 2013; Willenborg *et al.*, 2014).

In contrast to innate immune cells, little is known about the role of adaptive immune cells (B and T lymphocytes) in wound repair. The number of T cells increases in the granulation tissue within a few days after injury and their numbers remain elevated during the late stage of new tissue formation. Recent studies revealed a potential role of these cells in granulation tissue remodeling and scarring, mediated via IL-4, IL-13 and monocyte chemoattractant protein 1 (MCP-1) (Figure 2; Wong *et al.*, 2011), but the responsible T cell subsets remain to be identified. Furthermore, it is as yet unclear if and how these cells are involved in the healing process. The only exception is epidermal $\gamma\delta$ T cells, for which a role in wound reepithelialization has previously been demonstrated through their production of keratinocyte mitogens (Havran and Jameson, 2010).

The Phase of New Tissue Formation

Reepithelialization of skin wounds

The release of chemokines, cytokines, and growth factors by degranulating platelets, invading immune cells, and resident cells at the wound edge initiates the formation of new tissue. This phase starts within 1–2 days after injury by collective and coordinated migration of keratinocytes from the epidermis and injured hair follicles at the wound edge into the wounded area. Keratinocytes at the epidermal fronts initially attach to the exposed collagen and migrate down the injured dermis. Upon formation of granulation tissue they change their expression of integrin receptors, allowing them to attach to and to migrate on the newly deposited matrix proteins of the granulation tissue. To allow efficient migration, keratinocytes at the wound edge rearrange their cytoskeleton and extend lamellipodia. They reduce and modify their cell–cell contacts and the hemidesmosomes that allow them to attach to the basement membrane (Martin, 1997; Gurtner *et al.*, 2008). In addition, they secrete various types of proteinases to allow degradation of the connective tissue. These include plasminogen, which degrades the fibrin matrix, and different matrix metalloproteinases (MMP) (Figure 2). Consistent with the important function of these enzymes in wound healing, plasminogen-deficient mice showed strongly impaired reepithelialization, and healing was completely inhibited when wounds in these mice were treated with a broad-spectrum MMP inhibitor (Lund *et al.*, 1999).

The cellular alterations that occur at the migrating front of the wound epidermis are reminiscent of epithelial-

mesenchymal transition (EMT), a process that occurs during embryonic development and also in malignant carcinomas. During EMT, epithelial cells transdifferentiate into migratory individualized cells with a fibroblast-like morphology. In addition, they lose their epithelial cell markers and acquire markers characteristic for mesenchymal cells, such as the cytoskeletal protein vimentin. However, a complete EMT as seen in various cancer cells does not occur in keratinocytes of skin wounds. Rather, these cells retain some intercellular junctions and therefore show collective migration, and they still express epidermal keratins, but not vimentin. While keratinocyte migration is sufficient for the reepithelialization of small wounds, a wave of keratinocyte hyperproliferation, that occurs in keratinocytes behind the leading edge, is essential to replenish a larger wounded area with new epithelial cells. Reepithelialization is activated by various growth factors and cytokines. Functional studies in mouse models have revealed particularly important roles of hepatocyte growth factor (HGF) and of members of the fibroblast growth factor (FGF) and epidermal growth factor (EGF) families in keratinocyte migration and/or proliferation in wounded skin, although several other factors are also involved (Gurtner *et al.*, 2008; Schafer and Werner, 2008a; Figure 2).

Formation of granulation tissue

Shortly after the onset of reepithelialization, repair of the injured dermis is initiated, resulting in formation of the so-called granulation tissue (Figure 2). This is a highly vascularized and cell-rich tissue, which gradually replaces the fibrin clot. The name granulation tissue results from its granular appearance due to the presence of multiple nuclei. Initially, fibroblasts from the dermis at the wound edge are attracted to the wound site and they also start to proliferate. In addition hematopoietic stem cells and mesenchymal stem cells from the bone marrow are attracted to the wound tissue where they differentiate into collagen-producing fibroblasts (Gurtner *et al.*, 2008). Of particular importance for wound healing are fibrocytes, which are blood-borne cells expressing markers of hematopoietic and stromal cells. These cells differentiate into fibroblasts and myofibroblasts (see below) and they contribute to wound healing and pathological fibrosis (Wu *et al.*, 2007).

Activated fibroblasts in the dermis at the wound edge and in the granulation tissue produce growth factors that regulate wound reepithelialization in a paracrine manner, including HGF as well as EGF and FGF family members. In addition, they secrete proteinases to degrade the initial fibrin matrix and they produce the vast majority of the novel extracellular matrix proteins (ECM), including collagen type III, fibronectin, tenascin C, as well as proteoglycans and glycosaminoglycans. One of the major factors involved in the regulation of fibroblast migration and proliferation in skin wounds is platelet-derived growth factor (PDGF), which is initially released from degranulating platelets. During the healing process it is produced by several cell types, in particular by keratinocytes, and therefore it acts in a paracrine manner to stimulate migration, proliferation and matrix deposition by fibroblasts (Werner and Grose, 2003). Another major regulator of fibroplasia is TGF- β , which is also released from platelets. Subsequently, TGF- β is produced by many cell types in the healing wound. Furthermore, in combination with ED-A fibronectin, a splice variant

of fibronectin that is produced in wounded skin, and mechanical tension, TGF- β induces the differentiation of a large percentage of the wound fibroblasts into myofibroblasts. These cells express alpha-smooth muscle actin and therefore have efficient contractile properties. Thus, they are strongly involved in the contraction of the wound. In addition, they produce particularly large amounts of extracellular matrix proteins and they secrete proteinases, which mobilize growth factors from the matrix and which orchestrate the remodeling of the newly formed connective tissue (Tomasek *et al.*, 2002).

In parallel to the migration and proliferation of fibroblasts, new blood vessels are formed in the granulation tissue. This is essential for the supply of the cells with oxygen and nutrients (Figure 2). Neovascularization in skin wounds is mainly achieved by angiogenesis, the sprouting of new vessels from preexisting vessels. Bone marrow-derived endothelial progenitor cells can also be incorporated into new wound vessels, in particular under ischemic conditions, but their contribution to vessels in normally healing wounds appears minor. Blood vessels in wounds are initially leaky due to the presence of factors that induce vascular permeability, such as histamine and vascular endothelial growth factor A (VEGF-A). This results in continuous deposition of plasma-derived matrix proteins during the early phase of healing. Several days after healing, most of the vessels have lost the initial hyperpermeability and they become stabilized through association with pericytes and smooth muscle cells (Eming *et al.*, 2007a).

Several growth factors orchestrate the formation of new blood vessels in skin wounds. Of particular importance is VEGF-A, which is upregulated in response to the severe hypoxia that is present in early wounds. This involves hypoxia-induced stabilization of the transcription factor hypoxia-inducible factor 1 α . In addition, the related placental growth factor as well as members of the FGF family, in particular FGF2, are involved in wound angiogenesis. The subsequent stabilization of immature vessels by pericytes and smooth muscle cells is mainly achieved by angiopoietin-1 in combination with PDGF. Following the onset of blood vessel formation, lymphangiogenesis occurs in the granulation tissue to reconstruct the lymphatic vasculature. This process is predominantly controlled by VEGF-C and VEGF-D (Muller *et al.*, 2012).

Finally, nerve sprouting occurs at the wound site, resulting in reinnervation of the injured tissue. The wound is initially hyperinnervated, but this reverts during the remodeling phase (see below). Interestingly, reinnervation is important for normal wound healing as demonstrated by the severe healing impairments upon inhibition of this process (Werner and Grose, 2003).

The Phase of Tissue Remodeling

Following the phase of new tissue formation, the proliferation rate of keratinocytes returns to that of normal skin and the keratinocytes of the neo-epidermis redifferentiate, resulting in re-establishment of the epidermal barrier. At this stage the epidermis gets firmly attached to the underlying late granulation tissue/early scar tissue through reformation of the basement membrane and of hemidesmosomes, which

mediate this attachment (Figure 3). The factors, which control the shut-down of epidermal repair and the re-establishment of the keratinocyte differentiation program are still largely unknown, but members of the TGF- β superfamily and downstream effectors are likely to be involved.

In parallel to the maturation of the epidermis, remodeling of the granulation tissue is initiated, which can last up to several months. During this process the high cellularity of the granulation tissue is reduced due to apoptosis of (myo)fibroblasts and endothelial cells, resulting in vessel regression. This is most likely a consequence of the decrease in the levels of angiogenic growth factors. Since many of them are produced by macrophages, the reduction in the number of these immune cells due to apoptosis or reentry into the circulation may well contribute to vessel regression. In addition, alterations in the proteinase expression profile and the replacement of the provisional matrix by mature collagen are likely to play a role in this process. As a consequence of these apoptotic processes, the resulting scar tissue has a very low cellularity and it is mainly composed of connective tissue (Gurtner *et al.*, 2008; Figure 3).

The extracellular matrix of the granulation tissue/scar tissue undergoes a particularly extensive and long-lasting remodeling process. It involves a switch from production of collagen type III to collagen type I, which enhances the tensile strength. In addition, alterations in collagen cross-linking occur and the new fibers are aligned along tension lines. Finally, the balance between production and degradation of extracellular matrix is shifted toward the latter due to the enhanced expression of various MMPs. During the remodeling process, the tensile strength of the tissue increases continuously, but it never exceeds 70–80% of values of non-wounded skin. The elasticity of the new tissue is also significantly lower due to the lack of elastic fibers. Finally, the resulting scar tissue lacks hair follicles, sebaceous glands and sweat glands, which cannot regenerate (Gurtner *et al.*, 2008; Figure 3).

All wounds that involve dermal injury heal with a scar. The latter are normally flat and restricted to the initial wound area. However, scarring can be excessive, resulting in hypertrophic scars that are still restricted to the original wound site or keloids, which can strongly exceed this area. Hypertrophic scars and keloids are characterized by a prolonged presence of (myo)fibroblasts, which produce excessive amounts of connective tissue and by abnormal organization of collagen fibres. TGF- β is considered as the most important regulator of scarring, although additional factors contribute to this process, including activin and osteopontin. Furthermore, recent studies revealed that mechanical tension promotes scar formation through induction of Th2 cytokines and inflammatory chemokines (Wong *et al.*, 2011).

Importantly, mammalian embryos show scarless healing up to the end of the second trimester. Although not fully understood, the lack of scar formation in early embryos is thought to result from a limited inflammatory response compared to healing in adults, from reduced expression of TGF- β , as well as from differences in the proteinase expression profile. In the future it will be important to understand the cellular and molecular differences between wound healing in adults and embryos. This will be a prerequisite for the development of strategies that prevent scar formation in adults.

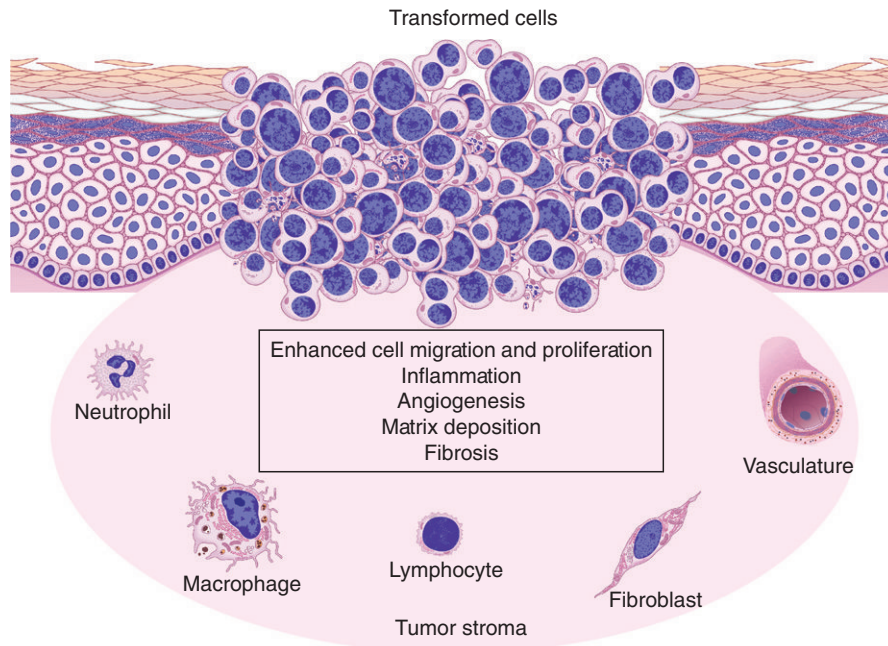


Figure 4 Cellular parallels between a malignant skin cancer and a healing skin wound. An epithelial cancer is shown schematically. Note the remarkable parallels to a skin wound in the phase of new tissue formation (Figure 2). Similar to granulation tissue of skin wounds, the tumor stroma includes different types of immune cells, newly formed blood and lymphatic vessels, and a large number of (myo)fibroblasts. In addition, hyperproliferating keratinocytes are present in the wound and in the cancer tissue. The major difference between a malignant tumor and a healing wound is the invasive and uncontrolled growth of the malignantly transformed keratinocytes.

Wound Healing versus Regeneration

While mammals heal their wounds using the above-described repair process that leads to development of scar tissue, some lower organisms, including amphibians, planarians, and hydra, have the capacity to fully regenerate, even after amputation of the head, the limbs and/or the tail. When these animals are severely injured, differentiated cells along the wound site are able to dedifferentiate and they form a so-called blastema. The blastemal cells subsequently proliferate, followed by redifferentiation according to the original lineage upon completion of wound closure. Although part of this regenerative capacity is retained in mammals, the rapid formation of scar tissue most likely prevents efficient and complete tissue regeneration (Tanaka and Reddien, 2011). A recent study, however, demonstrated that certain mammals retain a full regenerative capacity as demonstrated for the African spiny mouse. Shedding of larger skin areas and subsequent tissue regeneration were observed in these animals, involving formation of a blastema-like structure and subsequent regeneration of skin and all appendages (Seifert *et al.*, 2012). These results raise hope that strategies can be developed in humans to unlock such latent regenerative pathways, and it will therefore be important to identify the molecular and cellular mechanisms of the regenerative response of such animals and the differences to healing with scar formation.

process follows the same principles in humans. While wound healing is generally efficient in young patients, the repair capacity declines upon aging. In addition, there are many patients who suffer from severely impaired healing as a result of (1) bacterial infection of surgical or burn wounds, (2) treatment with anti-inflammatory steroids or chemotherapy or (3) systemic disease such as metabolic, vascular or autoimmune disease or a combination of these risk factors. These conditions frequently lead to the formation of chronic non-healing ulcers (Sen *et al.*, 2009). The underlying pathomechanisms are still insufficiently understood, in particular due to the lack of appropriate animal models for chronic wounds. However, vascular abnormalities, reduced levels of active growth factors, or reduced reactivity to these mitogens are suggested to be involved, together with enhanced levels/activities of various proteinases. In addition, the number of neutrophils and macrophages is frequently increased and their presence at the wound site is prolonged, resulting in excessive production of toxic ROS (Schafer and Werner, 2008b; Sindrilaru *et al.*, 2011). Future studies are required to gain insight into the molecular and cellular mechanisms underlying the pathogenesis of different types of chronic wounds. This will be a prerequisite for the development of improved therapeutic strategies.

Parallels between Wound Healing and Cancer

A dangerous consequence of chronic, non-healing wounds is the increased risk of squamous cell carcinoma development at the wound site, which is thought to result from the chronic inflammation. This association is also seen in various other

Impaired Wound Healing and Clinical Problems

The mechanisms underlying normal wound healing have been most thoroughly characterized in rodents, but the healing

tissues, such as the stomach or the intestine, where chronic inflammation due to the infection with *Helicobacter pylori* or the presence of inflammatory bowel disease, respectively, increases the risk of cancer development. These clinical observations as well as several experimental studies in mice suggested that common cellular and molecular mechanisms are active in wounds and in malignantly transformed tissues, and it was suggested by Dvorak (1986) that “cancers are wounds that do not heal.” In recent years, this hypothesis has been strongly supported by results from various studies that revealed strong cellular parallels between wound healing and cancer, including enhanced cell migration and proliferation, inflammation, angiogenesis, lymphangiogenesis, deposition of a similar type of matrix as well as fibrosis (Schafer and Werner, 2008a; Figure 4). Parallels between wound healing and cancer are also seen at the molecular level, since the gene expression profile of healing skin wounds shows strong similarities with the gene expression profile of carcinomas. Most interestingly, the presence of a gene expression signature that is similar to the one seen in early wounds allows a prediction of poor prognosis (Schafer and Werner, 2008a). In spite of these impressive parallels, a few major differences exist between wounds and tumors, in particular the presence of mutations and epigenetic alterations, which occur in tumors and prevent the temporal and spatial limitation of cell proliferation characteristic for wounded skin. In the future, it will be important to identify these differences in order to selectively inhibit tumor growth without affecting the wound healing process.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: Receptor Tyrosine Kinases and Their Ligands. Vertical Integration: Applications: A Review on Various Mathematical Modeling Approaches for Wound Healing

References

- Antsiferova, M., Martin, C., Huber, M., *et al.*, 2013. Mast cells are dispensable for normal and activin-promoted wound healing and skin carcinogenesis. *Journal of Immunology* 191, 6147–6155.
- Dvorak, H.F., 1986. Tumors: Wounds that do not heal. Similarities between tumor stroma generation and wound healing. *New England Journal of Medicine* 315, 1650–1659.
- Eming, S.A., Brachvogel, B., Odorisio, T., Koch, M., 2007a. Regulation of angiogenesis: Wound healing as a model. *Progress in Histochemistry and Cytochemistry* 42, 115–170.
- Eming, S.A., Krieg, T., Davidson, J.M., 2007b. Inflammation in wound repair: Molecular and cellular mechanisms. *Journal of Investigative Dermatology* 127, 514–525.
- Gregorio, J., Meller, S., Conrad, C., *et al.*, 2010. Plasmacytoid dendritic cells sense skin injury and promote wound healing through type I interferons. *Journal of Experimental Medicine* 207, 2921–2930.
- Gurtner, G.C., Werner, S., Barrandon, Y., Longaker, M.T., 2008. Wound repair and regeneration. *Nature* 453, 314–321.
- Havran, W.L., Jameson, J.M., 2010. Epidermal T cells and wound healing. *Journal of Immunology* 184, 5423–5428.
- Lund, L.R., Romer, J., Bugge, T.H., *et al.*, 1999. Functional overlap between two classes of matrix-degrading proteases in wound healing. *EMBO Journal* 18, 4645–4656.
- Martin, P., 1997. Wound healing — aiming for perfect skin regeneration. *Science* 276, 75–81.
- Martin, P., Leibovich, S.J., 2005. Inflammatory cells during wound repair: The good, the bad and the ugly. *Trends in Cell Biology* 15, 599–607.
- Moreira, S., Stramer, B., Evans, I., Wood, W., Martin, P., 2010. Prioritization of competing damage and developmental signals by migrating macrophages in the *Drosophila* embryo. *Current Biology: CB* 20, 464–470.
- Muller, A.K., Meyer, M., Werner, S., 2012. The roles of receptor tyrosine kinases and their ligands in the wound repair process. *Seminars in Cell & Developmental Biology* 23, 963–970.
- Niethammer, P., Grabher, C., Look, A.T., Mitchison, T.J., 2009. A tissue-scale gradient of hydrogen peroxide mediates rapid wound detection in zebrafish. *Nature* 459, 996–999.
- Schafer, M., Werner, S., 2008a. Cancer as an overhealing wound: An old hypothesis revisited. *Nature Reviews Molecular Cell Biology* 9, 628–638.
- Schafer, M., Werner, S., 2008b. Oxidative stress in normal and impaired wound repair. *Pharmacological Research: The Official Journal of the Italian Pharmacological Society* 58, 165–171.
- Seifert, A.W., Kiama, S.G., Seifert, M.G., *et al.*, 2012. Skin shedding and tissue regeneration in African spiny mice (*Acomys*). *Nature* 489, 561–565.
- Sen, C.K., Gordillo, G.M., Roy, S., *et al.*, 2009. Human skin wounds: A major and snowballing threat to public health and the economy. *Wound Repair and Regeneration: Official Publication of the Wound Healing Society [and] the European Tissue Repair Society* 17, 763–771.
- Sindrilaru, A., Peters, T., Wieschalka, S., *et al.*, 2011. An unrestrained proinflammatory M1 macrophage population induced by iron impairs wound healing in humans and mice. *Journal of Clinical Investigation* 121, 985–997.
- Tanaka, E.M., Reddien, P.W., 2011. The cellular basis for animal regeneration. *Developmental Cell* 21, 172–185.
- Tomasek, J.J., Gabbiani, G., Hinz, B., Chaponnier, C., Brown, R.A., 2002. Myofibroblasts and mechano-regulation of connective tissue remodeling. *Nature Reviews Molecular Cell Biology* 3, 349–363.
- Werner, S., Grose, R., 2003. Regulation of wound healing by growth factors and cytokines. *Physiological Reviews* 83, 835–870.
- Willenborg, S., Eckes, B., Brinckmann, J., *et al.*, 2014. Genetic ablation of mast cells redefines the role of mast cells in skin wound healing and bleomycin-induced fibrosis. *Journal of Investigative Dermatology* 134, 2005–2015.
- Wong, V.W., Paterno, J., Sorkin, M., *et al.*, 2011. Mechanical force prolongs acute inflammation via T-cell-dependent pathways during scar formation. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 25, 4498–4510.
- Wu, Y., Wang, J., Scott, P.G., Tredget, E.E., 2007. Bone marrow-derived stem cells in wound healing: A review. *Wound Repair and Regeneration: Official Publication of the Wound Healing Society [and] the European Tissue Repair Society* 15 (Suppl. 1), S18–S26.

Relevant Websites

- www.ETRS.org
European Tissue Repair Society.
- www.woundheal.org
Wound Healing Society.

Cancer Cell Invasion through Tissue Barriers

M Santiago-Medina, T Palmer, and J Yang, University of California, San Diego, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Metastasis is the most life-threatening aspect of cancer progression and is responsible for around 90% of cancer patient mortality. The metastatic process, in which cancer cells disseminate from their primary sites to form secondary tumors throughout the body, consists of a series of molecularly intricate steps. The initial step of tumor cell dissemination is invasion of the tumor cell into the surrounding tissue. During intravasation, tumor cells enter the circulatory system by breaching blood and/or lymphatic vessels. Subsequently, tumor cells must survive in circulation until they extravasate at a distant site and begin to proliferate. This proliferation at a secondary site can lead to the colonization of an essential organ, resulting in organ failure and ultimately host death.

During each of these intricate steps, cancer cells undergo the necessary cellular changes to adapt to their new environment. By adopting numerous molecular machineries to drive cell migration, invasion, and anti-apoptosis, cancer cells are

able to successfully colonize secondary sites. These processes include the epithelial–mesenchymal transition (EMT) program, which allows cancer cells to dissociate from each other and migrate away from the primary tumor, anoikis resistance, which protects cells from undergoing apoptosis in the absence of sufficient matrix attachment, and the formation of cell protrusions such as invadopodia and microtentacles, which allow cells to attach to and modify the surrounding environment. In this article, we describe the cellular programs involved in cancer cell invasion and metastasis and discuss the most widely accepted findings relevant to each molecular process (Figure 1).

EMT

Benign carcinoma cells are adhered to one another through cell–cell adhesions, such as adherens junctions and desmosomes. In addition, they may also be attached to an underlying

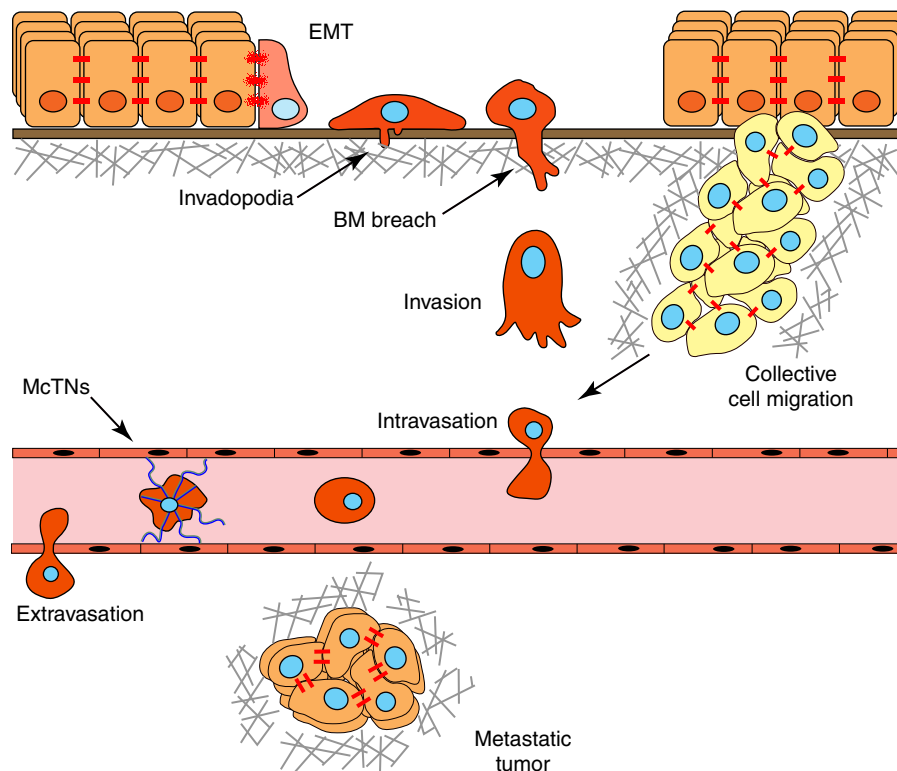


Figure 1 Cancer cell invasion and metastasis. Noninvasive epithelial cancer cells (orange cells) can undergo EMT to become migratory cancer cells (red). Cancer cells can breach the BM via collective cell migration and invasion. A subset of cancer cells can penetrate the BM via invadopodia-mediated ECM degradation. Migrating tumor cells develop resistance to anoikis, thereby allowing them to survive without appropriate ECM attachment. Cells continue to invade their surrounding tissues and intravasate into the systemic circulation via various membrane protrusions, including invadopodia, lamellipodia, and filopodia. CTCs extravasate from the blood circulation via microtentacle-mediated attachment to endothelial cells and then invadopodia-mediated endothelium breaching. In the new tissue microenvironment, cancer cells reverse EMT and proliferate to give rise to macrometastases. BM, basement membrane; EMT, epithelial–mesenchymal transition; McTNs, microtentacles.

basement membrane (BM). To metastasize, cancer cells need to separate from each other and acquire migratory and invasive capabilities. One effective way to achieve this is to activate a cellular process named EMT. During EMT, cancer cells change their adhesion profile, undergo a dramatic reorganization of the actin cytoskeleton, and form specialized membrane protrusions that are required for invasive migration. This process, which normally operates during embryo morphogenesis and wound healing, is usurped by tumor cells to achieve enhanced migration and invasion.

It is important to note that EMT is dynamically regulated during metastasis: turning on EMT promotes invasion and dissemination of tumor cells into distant organs; while turning off EMT is required for the establishment of distant metastasis (Tsai *et al.*, 2012). Furthermore, analyses of human breast cancer-circulating tumor cells suggest that a partial EMT may be sufficient to promote tumor cell dissemination (Min *et al.*, 2009; Kallergi *et al.*, 2011). Furthermore, cancer cells can also accomplish collective cell invasion in the absence of EMT (Wicki *et al.*, 2006).

Triggers of EMT

Carcinoma cells are thought to undergo EMT in response to a combination of intrinsic genetic changes within tumor cells and extracellular signals in the tumor microenvironment. Notably, oncogenic signals, such as Ras activation, have been shown to be required for tumor cells to respond to EMT-inducing extracellular signals. Under different physiological conditions, a number of extracellular signals, including Wnt, Notch, and growth factor receptor signals have been implicated in inducing some aspects of the EMT program. Growth factors known to induce EMT are transforming growth factor β (TGF β), tumor necrosis factor- α (TNF- α)/nuclear factor κ B (NF- κ B), hepatocyte growth factor (HGF), members of the epidermal growth factor (EGF) family, insulin-like growth factor (IGF), fibroblast growth factor (FGF), and platelet-derived growth factor (PDGF) (Acevedo *et al.*, 2007; Tsai and Yang, 2013). Most notably, the TGF-pathway appears to be a primary inducer of EMT under various physiological conditions (Katsuno *et al.*, 2013). In addition to growth factor signaling and inflammatory cytokines, hypoxia in the tumor microenvironment has also been shown to promote EMT (Yang *et al.*, 2008). Changes in extracellular matrix (ECM) composition and density are also known to be involved in EMT induction. Matrix proteins such as collagen I, collagen IV, fibronectin, tenascin C, and hyaluronan have been implicated in EMT activation in certain cell types (Wong and Rustgi, 2013). Studies also suggest that besides the signaling input of these ECM molecules, they may also regulate EMT and invasion by increasing ECM density and rigidity (Levental *et al.*, 2009; Provenzano *et al.*, 2009).

EMT-Inducing Transcription Factors

Although the triggers for EMT are very diverse, these upstream pathways converge onto a key group of EMT-activating transcription factors. These EMT-inducing transcription factors include the Snail family of zinc-finger transcription factors

Snail1/Snail2, the Zeb family of transcription factors Zeb1/Zeb2, and the Twist family of bHLH transcription factors Twist1/Twist2. In response to a myriad of microenvironmental stimuli, they function together as master molecular switches of the EMT program. Importantly, many of these transcription factors can induce each other in a positive feedback loop to coordinately orchestrate the entire EMT program.

Mechanistically, the Snail and Zeb family of EMT transcription factors are capable of directly binding to the E-boxes of the E-cadherin promoter, recruiting transcriptional corepressors and histone deacetylases to repress its transcription (Batlle *et al.*, 2000; Hajra *et al.*, 2002). The Zeb family proteins are also able to suppress E-cadherin transcription (Comijn *et al.*, 2001) via a double-negative feedback loop controlling Zeb1/Zeb2 and miRNA-200 family expression (Bracken *et al.*, 2008; Korpai *et al.*, 2008; Park *et al.*, 2008). Both the Snail and Zeb family of transcription factors also repress the expression of other cellular junction proteins, such as claudins and ZO-1 (Ohkubo and Ozawa, 2004; Vandewalle *et al.*, 2005). The Twist family of transcription factors, including Twist1 (Yang *et al.*, 2004) and Twist2 (Fang *et al.*, 2011) can induce EMT. Twist1 not only represses E-cadherin through induction of Snail transcription factors (Yang *et al.*, 2004; Casas *et al.*, 2011), but it also activates programs associated with tumor invasion (Eckert *et al.*, 2011), thus coordinating two major aspects of the EMT program.

As mentioned above, a subset of these alterations in EMT-transcription factor signaling can be regulated via miRNAs. Moreover, several regulatory circuits involving microRNAs (Bullock *et al.*, 2012), epigenetic machinery (Tam and Weinberg, 2013), and posttranslational control (De Craene and Berx, 2013) are crucial in enabling cancer cells to maintain a plastic cellular state or switch between the epithelial and mesenchymal states. For example, additional miRNAs that regulate EMT include miR-9 and miR-30a, which targets the E-cadherin-encoding mRNA CDH1 and Snail, respectively (Ma *et al.*, 2010; Kumarswamy *et al.*, 2012).

EMT-Associated Cellular and Molecular Changes

During EMT, epithelial cells undergo extensive morphological and cellular changes to adopt mesenchymal cell characteristics. Many of the hallmark EMT markers are subcellular structure proteins that define the epithelial or mesenchymal identity of a cell. Among them, E-cadherin is regarded as a gatekeeper of the epithelial state in various epithelial cell types. Snail and Zeb family of EMT transcription factors can directly suppress E-cadherin transcription as discussed above. Down-regulation of E-cadherin can also occur at the posttranslational level due to the activation of proteases, which cleave the protein and lead to its removal from the plasma membrane (Marambaud *et al.*, 2002; Maretzky *et al.*, 2005). Very often, in cell culture, down-regulation of E-cadherin is accompanied by the increased expression of N-cadherin, a phenomenon called the cadherin switch (Wheelock *et al.*, 2008). Additionally, other genes important in the maintenance of epithelial cell-cell adhesion such as the tight junction proteins, claudins and occludin, and the desmosome proteins, desmoplakin and plakophilin, are also repressed during this transition (Lamouille *et al.*, 2014).

Similarly, up-regulated genes include those encoding proteins involved in cell migration and invasion, such as the intermediate filament protein vimentin, and the ECM proteins fibronectin and collagen I. Importantly, through this transition, cells lose their association with epithelial cells and acquire the ability to interact and modify the surrounding ECM, thus facilitating cell migration and invasion.

Cancer progression is marked by an increase in the deposition and cross-linking of fibrillar ECM components, such as collagen and fibronectin, which results in an increase in ECM density and cell–matrix adhesions. Mirroring the cadherin switch, there are similar changes in the expression and activation of pro-invasive integrins induced by EMT-inducing transcription factors. This switch in the adhesion profile of carcinoma cells is what allows tumor cells to adhere and venture into novel territory while evading apoptosis or other cell growth inhibiting signaling pathways. As epithelial cells differentiate into mesenchymal cells, they exchange their BM-binding integrins for integrins that adhere to the surrounding collagen I-rich ECM, thus enhancing their ability to migrate away from the epithelium and invade into surrounding tissues. Migration-promoting integrins, including $\alpha1\beta1$, $\alpha2\beta1$, and $\alpha5\beta1$ are activated during EMT and increase cell adhesion to fibronectin and collagen I (Friedl and Alexander, 2011). Increased activation of mesenchymal integrins and their interactions with ECM proteins can also facilitate the disruption of E-cadherin complexes, while inducing N-cadherin expression (Koenig *et al.*, 2006; Shintani *et al.*, 2008). Other integrins associated with EMT, invasion and metastasis as well as poor prognosis in patients include $\alpha V\beta6$ (Bandyopadhyay and Raghavan, 2009) and $\alpha6\beta4$ (Evans *et al.*, 2003). Furthermore, changes to the integrin repertoire during EMT can lead to increased expression of proteases, such as the matrix metalloproteinases MMP2 and MMP9, which in turn stimulate cell migration and invasion into tissues. Studies have reported that colocalization and cooperation between $\beta1$ - and $\alpha V\beta3$ -integrin with MT1-MMP regulates cell invasion into the surrounding ECM (Galvez *et al.*, 2002).

Invadopodia

Carcinoma *in situ* are surrounded by thin, dense, and highly cross-linked ECM known as the BM. Comprised primarily of the macromolecules collagen IV and laminin and to a lesser extent perlecan and nidogen/entactin, the BM is a thin (~100 nm thick), but mechanically strong barrier. Carcinoma cells invade into nearby tissues and disseminate into the circulation by breaching through the underlying BM and surrounding ECM. Instead of using secreted MMPs, tumor cells actively remodel the ECM through the use of specialized subcellular structures termed invadopodia. Invadopodia are actin-based membrane protrusions located on the basal surface of cells at the cell–matrix interface. In two-dimensional (2D) settings, invadopodia appear as long-lived F-actin puncta with half-lives ranging from minutes to hours. Common molecular components of invadopodia include many of the same actin-regulatory proteins found in lamellipodia and filopodia, such as Arp2/Arp3, cofilin, cortactin, and WASP. As sites of cell adhesion, they also contain integrin receptors, Src,

PIXN, and FAK (Chan *et al.*, 2009). The RhoGTPase, Cdc42, and the scaffolding protein, Tks5, are thought to be crucial components for the assembly and degradation ability of invadopodia, respectively (Di Martino *et al.*, 2014). Invadopodia regulate tumor invasion by both adhering to and digesting the underlying ECM, thus allowing tumor cells to penetrate and migrate into the ECM.

BM Breach, Intravasation, and Extravasation

Studies performed both *in vitro* and *in vivo* have demonstrated a role for invadopodia-mediated BM breach. In *in vitro* assays, carcinoma cells cultured on intact BMs utilize invadopodia to degrade the matrix and subsequently assist the cell in crossing the membrane barrier (Schoumacher *et al.*, 2010). In *Caenorhabditis elegans*, the anchor cell forms a series of invadopodia-like structures, which are required for the cell to traverse the BM between cell tissues (Lohmer *et al.*, 2014).

Once metastatic tumor cells have successfully degraded their way to and through a BM, they may enter a blood vessel or lymphatic vessel through a process known as intravasation. During this process, cells enter the circulatory system by pushing and degrading their way through the BM and endothelial cell barrier that comprises the vessel. The mechanism in which carcinoma cells bypass the endothelial barrier to gain entry into circulatory vessels is known as transendothelial migration. Importantly, invadopodia have also been implicated in intravasation and metastasis *in vivo* (Gligorijevic *et al.*, 2014; Roh-Johnson *et al.*, 2014). To exit circulation, circulating tumor cells (CTCs) attach to the endothelial cells that coat the luminal side of the blood vessel and begin a process termed extravasation. As with intravasation, invadopodia also mediate extravasation and subsequent metastasis (Leong *et al.*, 2014; Sutoh Yoneyama *et al.*, 2014).

Invadopodia Induction

During EMT, epithelial cells breach the BM and migrate through the ECM by provoking the formation of invadopodia-like structures. The EMT transcription factor, Twist1, promotes invadopodia formation in response to various EMT-inducing signals (Eckert *et al.*, 2011). Likewise, a number of EMT-inducing cues are also implicated in the formation of invadopodia. For example, both EMT and invadopodia formation can be stimulated by various growth factors, including EGF (Yamaguchi *et al.*, 2005), HGF (Rajadurai *et al.*, 2012), PDGF (Eckert *et al.*, 2011), and TGF β (Pignatelli *et al.*, 2012). Additionally, environmental stimuli associated with EMT such as hypoxia, acidic pH, and matrix stiffness also promote the formation and activity of invadopodia (Alexander *et al.*, 2008a; Lucien *et al.*, 2011; Diaz *et al.*, 2013). These chemical and physical stimuli converge onto the activation of Src kinase and Rho GTPases, which are required in invadopodia formation.

Assembly of Invadopodia

Invadopodia appear as actin-rich puncta which at times are surrounded by a ring of adhesion proteins (Branch *et al.*, 2012). The formation of invadopodia follows a sequence of

distinct steps, beginning with an immature structure known as an invadopodium precursor. Invadopodium precursors begin with the local enrichment of actin and cortactin at sites of cell–ECM contact. This is followed by the recruitment of Tks5 as well as a change in the lipid composition of the plasma membrane at sites of maturing invadopodia (Sharma *et al.*, 2013). This change in lipid identity creates a docking site for Tks5, thus stabilizing the invadopodium precursor. Once stable, the invadopodium recruits proteases that mediate pericellular degradation, thus forming a fully functional mature invadopodium. One such protease, MT1-MMP, has been shown to be transported continually to mature invadopodia via recycling of endosomal vesicles (Hoshino *et al.*, 2012).

In three-dimensional (3D) matrices, invadopodia take on a different manifestation than their characteristic dot-like appearance in 2D. Invadopodia capable of forming stable 3D-structures appear as long, thin protrusions (Beatty *et al.*, 2013). Importantly, invadopodia of 3D environments maintain the same molecular signature and proteolytic function for invasion as their 2D counterparts (Yu and Machesky, 2012). In addition to 3D cell culture systems, an increasing number of recent studies have begun to characterize these structures under physiologic settings *in vivo*. For example, invadopodia-mediated cell invasion has been studied in dermis-based matrix (Tolde *et al.*, 2010), in organogenesis models of *C. elegans* (Lohmer *et al.*, 2014), and recently, in developing spinal cord neurons (Santiago-Medina *et al.*, 2015). Together, these data suggest a critical role for invadopodia in ECM degradation and remodeling in various physiological settings.

Modes of Cell Migration

To invade and metastasize successfully, cancer cells must be extremely versatile, capable of employing multiple cellular mechanisms to migrate during the metastatic process. These include mesenchymal cell migration, collective cell migration, and protease-independent amoeboid migration.

Mesenchymal Migration

During the progression of EMT, a high-degree of actin remodeling takes place. The onset of cell motility requires a relaxation of static actin structures to form dynamic membrane protrusions. In epithelial cells, actin filaments are organized in thin cortical bundles. This is contrary to mesenchymal cells, in which the actin is bundled into thick contractile stress fibers at the ventral cell surface. During cell migration, rigid actin stress fibers are disassembled and polymerized into a cortical actin meshwork, from which cell membrane protrusions like lamellipodia and filopodia emerge. Lamellipodia are flat, sheet-like protrusions composed of both unbranched and branched actin filaments. Filopodia are thin bundles of actin filaments that emanate into the surrounding environment from the leading edge. The formation of both lamellipodia and filopodia not only assist tumor cells physically migrating from one point to another, but also serve as exploratory structures that sense molecular cues in the microenvironment during the metastatic process. Actin-regulatory proteins such as Arp2/Arp3, Ena-Vasp, WASP, mDia, and cortactin are

redistributed within cells to produce these new membrane protrusions. Indeed, the increased expression of these actin-regulatory proteins correlates with poor prognosis in breast and liver carcinomas (Iwaya *et al.*, 2007). Additionally, these actin-based membrane protrusions are known to secrete proteases such as MMPs and ADAMs, revealing their potential role for the onset of cell invasion (Machesky, 2008).

The remodeling of the actin cytoskeleton that occurs during carcinoma cell migration and invasion depends on the activation of a family of molecular switches known as Rho GTPases. RhoA, Rac1, and Cdc42 are best studied among the 23 family members of Rho GTPases. With regard to migration and invasion and in a simplified view, RhoA induces actin stress-fiber formation and regulates cytoskeletal changes affecting cell–cell or cell–matrix adhesion, Rac1 is involved in lamellipodia and membrane ruffle formation, and Cdc42 is involved in filopodia formation. Based on their function in integrating and transmitting signals from chemokine, growth factor receptors and integrin-ECM binding onto the remodeling of the actin cytoskeleton, Rho proteins play a critical role in facilitating mesenchymal cell migration.

Amoeboid Migration

In addition to protease-based mesenchymal migration, tumor cells display amoeboid movement that does not require protease-mediated ECM remodeling. Amoeboid migration is characterized by cycles of expansion and contraction of the cell body and membrane blebbing mediated by Rho/ROCK-mediated actomyosin contractility. Tumor cells that adopt amoeboid migration display a characteristic rounded shape when observed in 3D matrices. This enables the cells to squeeze through preexisting gaps in ECM fibers as well as to exert sufficient force to deform the surrounding ECM without activating ECM degradation pathways.

Importantly, amoeboid and mesenchymal forms of invasion are two modes of migration that invading cells may use interchangeably as seen during mesenchymal–amoeboid transition (MAT) or amoeboid–mesenchymal transition (AMT). The low-adhesion attachment to the substrate enables amoeboid cells to translocate in the 3D environment at relatively high velocities, reaching up to $25 \mu\text{m min}^{-1}$ (Friedl *et al.*, 1994). Amoeboid cell migration can serve as a compensatory mechanism of migration upon inhibition of pericellular proteolysis.

Collective Cell Migration

Many types of epithelial cancers, including breast, prostate, pancreatic, and lung cancer, display a multicellular form of migration known as collective cell invasion. Such multicellular units consist of cells that migrate together during invasion into the surrounding stroma as well as during intravasation and metastasis. The cells comprising these mobile clusters remain bound to one another by expressing cell–cell junction molecules such as cadherins, tight junction proteins, adhesion receptors of the immunoglobulin superfamily, and gap junctions. In addition to providing cell–cell adhesion, the aforementioned junction molecules also maintain what are known as supracellular properties. These include multicellular

front–rear polarity, cytoskeletal synchronization, and juxtacrine cell–cell signaling (Friedl and Gilmour, 2009). The degree of cell–cell adhesion, cell–matrix adhesion and proteolysis can vary among cell groups giving rise to various subtypes of collective invasion (Friedl *et al.*, 2012). As a result, invading cell groups can appear as strands, monolayer sheets, or masses that form luminal structures. Cells at the periphery of the cell mass as well as at the leading edge generate traction force by actomyosin-mediated protrusion and contractility. The cells situated at sites of ECM contact engage in integrin signaling while matrix-degrading proteases, such as MT1-MMP mediate proteolytic ECM remodeling to widen tissue space during migration (Friedl and Gilmour, 2009).

Different *in vitro* models have been used to study the mechanisms of collective cell migration such as the widely used scratch-wound assay as well as tumor explants cultured in 3D matrices. By injecting 3D spheroids into the deep dermis of mice that were monitored through a window chamber, collective cell invasion has also been assessed *in vivo* (Alexander *et al.*, 2008b). Additional *in vivo* evidence for collective cancer cell invasion includes the histopathological analysis of human cancer lesions, in which neoplastic multicellular strands and masses cross tissue boundaries and extend into the stroma (Christiansen and Rajasekaran, 2006). Recent studies having found tumor cell clusters in the circulation of cancer patients suggest that collective cell invasion may also occur during intravasation (Hart, 2009; Hou *et al.*, 2011).

Anoikis

After entering the circulatory system, tumor cells in suspension must survive the lack of proper ECM attachment and the shear stresses of the vascular circulation to reach distant tissues. In the absence of cell attachment to an ECM or upon cell adhesion to an inappropriate location, normal cells undergo a specific type of apoptosis, termed anoikis (from the Greek for ‘homelessness’). In adherent cells, the activation of integrins and their downstream signaling mediators, including the non-receptor tyrosine kinase Src, FAK, and integrin-linked kinase (ILK), provide protection from anoikis (Guadamillas *et al.*, 2011). In the absence of these adhesion-based signals, cells activate caspases and endonucleases, leading to DNA fragmentation and cell death. Cancer cells are able to bypass this safeguard, and as a result, survive under suspension conditions or proliferate at ectopic sites where ECM proteins are different from the cells’ native ones. Indeed, CTCs clearly show resistance to anoikis (Yu *et al.*, 2011). This deregulation in anoikis execution is another hallmark of cancer cells and contributes to metastasis.

Cancer cells develop anoikis resistance by implementing several distinct mechanisms, including the up-regulation of integrin signaling, activation of inside–out pro-survival signals, overexpression of growth factor receptors, and activation of oncogenes. The tumor microenvironment can also contribute to anoikis resistance by modulating matrix stiffness, enhancing oxidative stress, producing pro-survival soluble factors, as well as by triggering EMT and a self-renewal ability. Carcinoma cells can avoid succumbing to anoikis by altering their integrin repertoire. For example, in squamous cell

carcinoma, a switch from $\alpha v\beta 5$ integrins to $\alpha v\beta 6$ integrins induces the intrinsic apoptotic pathway when uncoupled, and activates the pro-survival phosphoinositide 3-kinase (PI3K)–Akt pathway, respectively (Janes and Watt, 2004). Additional examples include down-regulation of $\alpha v\beta 3$ in intestinal carcinoma cells as well as up-regulation of $\alpha v\beta 3$ in melanoma cells, which suppress anoikis by modifying the anti-apoptotic Bcl-2 and pro-apoptotic Bax proteins (Paoli *et al.*, 2013).

Another mechanism employed by tumor cells to overcome anoikis is the hyperactivation of growth factor pathways and receptor tyrosine kinases (RTKs). Detached tumor cells silence the requirement for integrin signaling by constitutively activating pro-survival cascades, such as PI3K, Ras–Erk, NF- β B, and RhoGTPase signaling (Tsuji *et al.*, 2009). This is achieved through autocrine secretion of many of the same growth factors implicated in EMT, such as FGF, HGF, and PDGF. Additionally, overexpression of RTKs, such as EGFR, TrkB, HGF, and ErbB2 receptors also lead to the suppression of anoikis in tumor cells (Guadamillas *et al.*, 2011).

Downstream of RTKs, detached or migrating cancer cells modulate the PI3K–Akt pathway to compensate for the loss of integrin signals and overcome anoikis. This pathway integrates most of the pro-survival signals derived from integrins and growth factors receptors. Sustained Akt activation is achieved as a consequence of an alteration in PI3K activity, constitutive activation of receptor protein tyrosine kinases, activating Ras mutations, and loss of PTEN tumor suppressors (Guadamillas *et al.*, 2011). Interestingly, Akt activation also occurs following the switch from E-cadherin to N-cadherin that takes place during EMT. N-cadherin recruits PI3K which in turn activates Akt and induces anoikis resistance (Cavallaro and Christofori, 2004). Depletion of E-cadherin promotes mammary cell survival following loss of cell adhesion to the ECM (Derksen *et al.*, 2006), and N-cadherin expression protects melanoma cells from anoikis (Li *et al.*, 2001).

Microtentacles

To extravasate from the circulation, cancer cells latch on to endothelial cells of the inner lining of blood vessels by forming specialized structures known as microtentacles. These structures, found in free-floating tumor cells and in CTCs *in vivo*, are known to enhance metastases. Unlike actin-based invadopodia, microtentacles are microtubule-rich protrusions that do not present ECM degradation activity. As dynamic, microtubule-enriched plasma membrane extensions, microtentacles are antagonized by the actin cytoskeleton as seen by the potentiation of microtentacles in cells treated with actin depolymerization drugs (Whipple *et al.*, 2007). Similarly, cofilin-mediated actin depolymerization promotes microtentacle formation by weakening the contraction of cortical actin (Vitolo *et al.*, 2013). Analogously, microtubule stabilization as well as stabilizing proteins such as tau, enhance microtentacles, facilitating tumor cell aggregation and attachment to the capillary endothelium of the microvasculature. Interestingly, Src kinase has been shown to differentially regulate microtentacles and invadopodia, suggesting that Src activity may act as a switch between the two protrusions (Balzer *et al.*, 2010). Examination of the molecular

constitution of microtentacles demonstrates that these protrusions are enriched in detyrosinated Glu-tubulin, a post-translationally modified form of α -tubulin that is found in stabilized microtubules.

Interestingly, the ectopic expression of the EMT-transcription factors Twist and Snail not only promotes α -tubulin detyrosination but also the formation of microtentacles in suspended cells (Whipple *et al.*, 2010). During EMT, tubulin tyrosine ligase enzyme is down-regulated, which leads to an increase in detyrosinated α -tubulin (Glu-tubulin), and, subsequently, an increase in microtentacles that penetrate the vascular endothelium to facilitate tumor cell reattachment during metastasis.

Conclusion

This article covers the major cellular programs utilized by metastatic tumor cells to invade and disseminate. Among them, we focused on the EMT program given its pleiotropic impact on cell migration and invasion. Although there is some debate as to how frequently EMT is used in human cancers, an abundance of work has demonstrated the critical role of the dynamic EMT in tumor invasion and metastasis. Upon EMT, carcinoma cells lose cell–cell junctions to allow single cell migration and acquire invadopodia to promote ECM degradation. Recently, it has been demonstrated that the plasticity of tumor cells also allows them to use different modes of migration, including collective cell migration, mesenchymal cell migration, and amoeboid migration in response to different extracellular stimuli and intrinsic alterations. Other cellular programs necessary for metastasis include anoikis resistance to allow survival under unfavorable ECM attachment and the formation of cellular structure microtentacles for attachment and extravasation from the blood vessels. Interestingly, many of the same factors that regulate EMT also control these cellular programs to promote successful invasion and dissemination. Further analysis of these cellular pathways will greatly improve our understanding of the process of metastasis and reveal novel targets for therapeutic intervention.

See also: Protein Synthesis/Degradation: Protein Degradation – Intracellular: Ubiquitin, Ubiquitin-Like Proteins, and Proteasome-Mediated Degradation. Protein Synthesis/Degradation: Protein Degradation – Protease Classes: ADAMs Regulate Cell–Cell Interactions by Controlling the Function of the EGF-Receptor, TNF α and Notch; ADAMTS Proteases: Mediators of Physiological and Pathogenic Extracellular Proteolysis; Extracellular: Plasma Membrane Proteases – Serine Proteases; Matrix Metalloproteinases; Naturally-Occurring Polypeptide Inhibitors: Cystatins/Steifins, Inhibitors of Apoptosis (IAPs), Serpins, and Tissue Inhibitors of Metalloproteinases (TIMPs); The Calpain Proteolytic System

References

- Acevedo, V.D., Gangula, R.D., Freeman, K.W., *et al.*, 2007. Inducible FGFR-1 activation leads to irreversible prostate adenocarcinoma and an epithelial-to-mesenchymal transition. *Cancer Cell* 12, 559–571.
- Alexander, N.R., Branch, K.M., Parekh, A., *et al.*, 2008a. Extracellular matrix rigidity promotes invadopodia activity. *Current Biology* 18, 1295–1299.
- Alexander, S., Koehl, G.E., Hirschberg, M., Geissler, E.K., Friedl, P., 2008b. Dynamic imaging of cancer growth and invasion: A modified skin-fold chamber model. *Histochemistry and Cell Biology* 130, 1147–1154.
- Balzer, E.M., Whipple, R.A., Thompson, K., *et al.*, 2010. c-Src differentially regulates the functions of microtentacles and invadopodia. *Oncogene* 29, 6402–6408.
- Bandyopadhyay, A., Raghavan, S., 2009. Defining the role of integrin α v β 6 in cancer. *Current Drug Targets* 10, 645–652.
- Battle, E., Sancho, E., Franci, C., *et al.*, 2000. The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nature Cell Biology* 2, 84–89.
- Beatty, B.T., Sharma, V.P., Bravo-Cordero, J.J., *et al.*, 2013. Beta1 integrin regulates Arg to promote invadopodial maturation and matrix degradation. *Molecular Biology of the Cell* 24 (11), 1661–1675.
- Bracken, C.P., Gregory, P.A., Kolesnikoff, N., *et al.*, 2008. A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer Research* 68, 7846–7854.
- Branch, K.M., Hoshino, D., Weaver, A.M., 2012. Adhesion rings surround invadopodia and promote maturation. *Biology Open* 1, 711–722.
- Bullock, M.D., Sayan, A.E., Packham, G.K., Mirnezami, A.H., 2012. MicroRNAs: Critical regulators of epithelial to mesenchymal (EMT) and mesenchymal to epithelial transition (MET) in cancer progression. *Biology of the Cell* 104, 3–12.
- Casas, E., Kim, J., Bendesky, A., *et al.*, 2011. Snail2 is an essential mediator of Twist1-induced epithelial mesenchymal transition and metastasis. *Cancer Research* 71, 245–254.
- Cavallaro, U., Christofori, G., 2004. Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. *Nature Reviews Cancer* 4, 118–132.
- Chan, K.T., Cortesio, C.L., Huttenlocher, A., 2009. FAK alters invadopodia and focal adhesion composition and dynamics to regulate breast cancer invasion. *Journal of Cell Biology* 185, 357–370.
- Christiansen, J.J., Rajasekaran, A.K., 2006. Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis. *Cancer Research* 66, 8319–8326.
- Comijn, J., Berx, G., Vermassen, P., *et al.*, 2001. The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Molecular Cell* 7, 1267–1278.
- De Craene, B., Berx, G., 2013. Regulatory networks defining EMT during cancer initiation and progression. *Nature Reviews Cancer* 13, 97–110.
- Derkse, P.W., Liu, X., Saridin, F., *et al.*, 2006. Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. *Cancer Cell* 10, 437–449.
- Di Martino, J., Paysan, L., Gest, C., *et al.*, 2014. Cdc42 and Tks5: A minimal and universal molecular signature for functional invadosomes. *Cell Adhesion & Migration* 8, 280–292.
- Diaz, B., Yuen, A., Iizuka, S., Higashiyama, S., Courtneidge, S.A., 2013. Notch increases the shedding of HB-EGF by ADAM12 to potentiate invadopodia formation in hypoxia. *Journal of Cell Biology* 201, 279–292.
- Eckert, M.A., Lwin, T.M., Chang, A.T., *et al.*, 2011. Twist1-induced invadopodia formation promotes tumor metastasis. *Cancer Cell* 19, 372–386.
- Evans, R.D., Perkins, V.C., Henry, A., *et al.*, 2003. A tumor-associated beta 1 integrin mutation that abrogates epithelial differentiation control. *Journal of Cell Biology* 160, 589–596.
- Fang, X., Cai, Y., Liu, J., *et al.*, 2011. Twist2 contributes to breast cancer progression by promoting an epithelial-mesenchymal transition and cancer stem-like cell self-renewal. *Oncogene* 30, 4707–4720.
- Friedl, P., Alexander, S., 2011. Cancer invasion and the microenvironment: Plasticity and reciprocity. *Cell* 147, 992–1009.
- Friedl, P., Gilmour, D., 2009. Collective cell migration in morphogenesis, regeneration and cancer. *Nature Reviews Molecular Cell Biology* 10, 445–457.
- Friedl, P., Locker, J., Sahai, E., Segall, J.E., 2012. Classifying collective cancer cell invasion. *Nature Cell Biology* 14, 777–783.
- Friedl, P., Noble, P.B., Shields, E.D., Zanker, K.S., 1994. Locomotor phenotypes of unstimulated CD45RA^{high} and CD45RO^{high} CD4⁺ and CD8⁺ lymphocytes in three-dimensional collagen lattices. *Immunology* 82, 617–624.
- Galvez, B.G., Matias-Roman, S., Yanez-Mo, M., Sanchez-Madrid, F., Arroyo, A.G., 2002. ECM regulates MT1-MMP localization with beta1 or α v β 3 integrins at distinct cell compartments modulating its internalization and activity on human endothelial cells. *Journal of Cell Biology* 159, 509–521.
- Glgorijevic, B., Bergman, A., Condeelis, J., 2014. Multiparametric classification links tumor microenvironments with tumor cell phenotype. *PLoS Biology* 12, e1001995.

- Guadamillas, M.C., Cerezo, A., del Pozo, M.A., 2011. Overcoming anoikis – Pathways to anchorage-independent growth in cancer. *Journal of Cell Science* 124, 3189–3197.
- Hajra, K.M., Chen, D.Y., Fearon, E.R., 2002. The SLUG zinc-finger protein represses E-cadherin in breast cancer. *Cancer Research* 62, 1613–1618.
- Hart, I.R., 2009. New evidence for tumour embolism as a mode of metastasis. *Journal of Pathology* 219, 275–276.
- Hoshino, D., Koshikawa, N., Suzuki, T., *et al.*, 2012. Establishment and validation of computational model for MT1-MMP dependent ECM degradation and intervention strategies. *PLoS Computational Biology* 8, e1002479.
- Hou, J.M., Krebs, M., Ward, T., *et al.*, 2011. Circulating tumor cells as a window on metastasis biology in lung cancer. *American Journal of Pathology* 178, 989–996.
- Iwaya, K., Oikawa, K., Semba, S., *et al.*, 2007. Correlation between liver metastasis of the colocalization of actin-related protein 2 and 3 complex and WAVE2 in colorectal carcinoma. *Cancer Science* 98, 992–999.
- Janes, S.M., Watt, F.M., 2004. Switch from α 5 β 1 to α 6 β 1 integrin expression protects squamous cell carcinomas from anoikis. *Journal of Cell Biology* 166, 419–431.
- Kallergi, G., Papadaki, M.A., Politaki, E., *et al.*, 2011. Epithelial to mesenchymal transition markers expressed in circulating tumour cells of early and metastatic breast cancer patients. *Breast Cancer Research* 13, R59.
- Katsuno, Y., Lamouille, S., Derynck, R., 2013. TGF- β signaling and epithelial-mesenchymal transition in cancer progression. *Current Opinion in Oncology* 25, 76–84.
- Koenig, A., Mueller, C., Hasel, C., Adler, G., Menke, A., 2006. Collagen type I induces disruption of E-cadherin-mediated cell-cell contacts and promotes proliferation of pancreatic carcinoma cells. *Cancer Research* 66, 4662–4671.
- Korpai, M., Lee, E.S., Hu, G., Kang, Y., 2008. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. *Journal of Biological Chemistry* 283, 14910–14914.
- Kumarswamy, R., Mudduluru, G., Ceppi, P., *et al.*, 2012. MicroRNA-30a inhibits epithelial-to-mesenchymal transition by targeting Snai1 and is downregulated in non-small cell lung cancer. *International Journal of Cancer* 130, 2044–2053.
- Lamouille, S., Xu, J., Derynck, R., 2014. Molecular mechanisms of epithelial-mesenchymal transition. *Nature Reviews Molecular Cell Biology* 15, 178–196.
- Leong, H.S., Robertson, A.E., Stoleto, K., *et al.*, 2014. Invadopodia are required for cancer cell extravasation and are a therapeutic target for metastasis. *Cell Reports* 8, 1558–1570.
- Levental, K.R., Yu, H., Kass, L., *et al.*, 2009. Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell* 139, 891–906.
- Li, G., Satyamoorthy, K., Herlyn, M., 2001. N-cadherin-mediated intercellular interactions promote survival and migration of melanoma cells. *Cancer Research* 61, 3819–3825.
- Lohmer, L.L., Kelley, L.C., Hagedorn, E.J., Sherwood, D.R., 2014. Invadopodia and basement membrane invasion in vivo. *Cell Adhesion & Migration* 8, 246–255.
- Lucien, F., Brochu-Gaudreau, K., Arseneault, D., Harper, K., Dubois, C.M., 2011. Hypoxia-induced invadopodia formation involves activation of NHE-1 by the p90 ribosomal S6 kinase (p90RSK). *PLoS One* 6, e28851.
- Ma, L., Young, J., Prabhala, H., *et al.*, 2010. miR-9, a MYC/MYCIN-activated microRNA, regulates E-cadherin and cancer metastasis. *Nature Cell Biology* 12, 247–256.
- Machesky, L.M., 2008. Lamellipodia and filopodia in metastasis and invasion. *FEBS Letters* 582, 2102–2111.
- Marambaud, P., Shioi, J., Serban, G., *et al.*, 2002. A presenilin-1/gamma-secretase cleavage releases the E-cadherin intracellular domain and regulates disassembly of adherens junctions. *EMBO Journal* 21, 1948–1956.
- Maretzky, T., Reiss, K., Ludwig, A., *et al.*, 2005. ADAM10 mediates E-cadherin shedding and regulates epithelial cell-cell adhesion, migration, and beta-catenin translocation. *Proceedings of the National Academy of Sciences of the United States of America* 102, 9182–9187.
- Min, A.L., Choi, J.Y., Woo, H.Y., *et al.*, 2009. High expression of Snai1 mRNA in blood from hepatocellular carcinoma patients with extra-hepatic metastasis. *Clinical & Experimental Metastasis* 26, 759–767.
- Ohkubo, T., Ozawa, M., 2004. The transcription factor Snai1 downregulates the tight junction components independently of E-cadherin downregulation. *Journal of Cell Science* 117, 1675–1685.
- Paoli, P., Giannoni, E., Chiarugi, P., 2013. Anoikis molecular pathways and its role in cancer progression. *Biochimica et Biophysica Acta* 1833, 3481–3498.
- Park, S.M., Gaur, A.B., Lengyel, E., Peter, M.E., 2008. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes & Development* 22, 894–907.
- Pignatelli, J., Tumbarello, D.A., Schmidt, R.P., Turner, C.E., 2012. Hic-5 promotes invadopodia formation and invasion during TGF- β -induced epithelial-mesenchymal transition. *Journal of Cell Biology* 197, 421–437.
- Provenzano, P.P., Inman, D.R., Eliceiri, K.W., Keely, P.J., 2009. Matrix density-induced mechanoregulation of breast cell phenotype, signaling and gene expression through a FAK-ERK linkage. *Oncogene* 28, 4326–4343.
- Rajadurai, C.V., Havrylov, S., Zaoui, K., *et al.*, 2012. Met receptor tyrosine kinase signals through a cortactin-Gab1 scaffold complex, to mediate invadopodia. *Journal of Cell Science* 125, 2940–2953.
- Roh-Johnson, M., Bravo-Cordero, J.J., Patsialou, A., *et al.*, 2014. Macrophage contact induces RhoA GTPase signaling to trigger tumor cell intravasation. *Oncogene* 33, 4203–4212.
- Santiago-Medina, M., Gregus, K.A., Nichol, R.H., O'toole, S.M., Gomez, T.M., 2015. Regulation of ECM degradation and axon guidance by growth cone invadosomes. *Development* 142, 486–496.
- Schoumacher, M., Goldman, R.D., Louvard, D., Vignjevic, D.M., 2010. Actin, microtubules, and vimentin intermediate filaments cooperate for elongation of invadopodia. *Journal of Cell Biology* 189, 541–556.
- Sharma, V.P., Eddy, R., Entenberg, D., *et al.*, 2013. Tks5 and SHP2 regulate invadopodium maturation, but not initiation, in breast carcinoma cells. *Current Biology* 23, 2079–2089.
- Shintani, Y., Fukumoto, Y., Chaika, N., *et al.*, 2008. Collagen I-mediated up-regulation of N-cadherin requires cooperative signals from integrins and discoidin domain receptor 1. *Journal of Cell Biology* 180, 1277–1289.
- Sutoh Yoneyama, M., Hatakeyama, S., Habuchi, T., *et al.*, 2014. Vimentin intermediate filament and plectin provide a scaffold for invadopodia, facilitating cancer cell invasion and extravasation for metastasis. *European Journal of Cell Biology* 93, 157–169.
- Tam, W.L., Weinberg, R.A., 2013. The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nature Medicine* 19, 1438–1449.
- Tolde, O., Rosel, D., Vesely, P., Folk, P., Brabek, J., 2010. The structure of invadopodia in a complex 3D environment. *European Journal of Cell Biology* 89, 674–680.
- Tsai, J.H., Donaher, J.L., Murphy, D.A., Chau, S., Yang, J., 2012. Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell* 22, 725–736.
- Tsai, J.H., Yang, J., 2013. Epithelial-mesenchymal plasticity in carcinoma metastasis. *Genes & Development* 27, 2192–2206.
- Tsuji, T., Ibaragi, S., Hu, G.F., 2009. Epithelial-mesenchymal transition and cell cooperativity in metastasis. *Cancer Research* 69, 7135–7139.
- Vandewalle, C., Comijn, J., De Craene, B., *et al.*, 2005. SIP1/ZEB2 induces EMT by repressing genes of different epithelial cell-cell junctions. *Nucleic Acids Research* 33, 6566–6578.
- Vitolo, M.I., Boggs, A.E., Whipple, R.A., *et al.*, 2013. Loss of PTEN induces microtentacles through PI3K-independent activation of cofilin. *Oncogene* 32, 2200–2210.
- Wheelock, M.J., Shintani, Y., Maeda, M., Fukumoto, Y., Johnson, K.R., 2008. Cadherin switching. *Journal of Cell Science* 121, 727–735.
- Whipple, R.A., Cheung, A.M., Martin, S.S., 2007. Detyrosinated microtubule protrusions in suspended mammary epithelial cells promote reattachment. *Experimental Cell Research* 313, 1326–1336.
- Whipple, R.A., Matrone, M.A., Cho, E.H., *et al.*, 2010. Epithelial-to-mesenchymal transition promotes tubulin detyrosination and microtentacles that enhance endothelial engagement. *Cancer Research* 70, 8127–8137.
- Wicki, A., Lehenbre, F., Wick, N., *et al.*, 2006. Tumor invasion in the absence of epithelial-mesenchymal transition: Podoplanin-mediated remodeling of the actin cytoskeleton. *Cancer Cell* 9, 261–272.
- Wong, G.S., Rustgi, A.K., 2013. Matricellular proteins: Priming the tumour microenvironment for cancer development and metastasis. *British Journal of Cancer* 108, 755–761.
- Yamaguchi, H., Lorenz, M., Kempf, S., *et al.*, 2005. Molecular mechanisms of invadopodium formation: The role of the N-WASP-Arp2/3 complex pathway and cofilin. *Journal of Cell Biology* 168, 441–452.
- Yang, J., Mani, S.A., Donaher, J.L., *et al.*, 2004. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* 117, 927–939.
- Yang, M.H., Wu, M.Z., Chiou, S.H., *et al.*, 2008. Direct regulation of TWIST by HIF-1 α promotes metastasis. *Nature Cell Biology* 10, 295–305.
- Yu, M., Stott, S., Toner, M., Maheswaran, S., Haber, D.A., 2011. Circulating tumor cells: Approaches to isolation and characterization. *Journal of Cell Biology* 192, 373–382.
- Yu, X., Machesky, L.M., 2012. Cells assemble invadopodia-like structures and invade into matrigel in a matrix metalloprotease dependent manner in the circular invasion assay. *PLoS One* 7, e30605.

Neuronal Transport and Spatial Signaling Mechanisms in Neural Repair

AL Kalinski, Drexel University, Philadelphia, PA, USA

T Hines, DS Smith, and JL Twiss, University of South Carolina, Columbia, SC, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Anterograde transport Movement of cargo along neurites away from the cell body.

Axon The cytoplasmic projection from the neuron that transmits information away from the cell body.

Axon transport The process of transporting proteins, ribonucleic proteins (RNPs), vesicles, and organelles to the axon from the cell body and from the axon back to the cell body.

Cytoskeleton Network of protein polymers and protein complexes (microtubules, neurofilaments, and actin) that provide structure to the cell.

Central nervous system Components of the nervous system that include the cerebrum, brain stem and cerebellum ('brain'), and spinal cord.

Dendrite Cytoplasmic projections from the neuron that transmit information toward the cell body.

Growth cone The most distal region of a growing axon. Growth cones contain receptors that respond to guidance

cues in the extracellular environment that help direct it to its appropriate final destination, as well as dynamic cytoskeleton to promote growth.

Local protein synthesis Synthesis of new proteins at peripheral sites in the cytoplasm. For neurons, this includes translation in axons and dendrites, utilize mRNAs, translation factors, and ribosomes that are actively transported to these sites.

Motor protein Protein complexes that use ATP to move cargo and transmit tractile forces within cells.

Peripheral nervous system Components of the nervous system that are 'peripheral' to brain and spinal cord. This includes peripheral nerves, spinal ganglia, autonomic ganglia, and enteric neurons.

Retrograde transport Movement of cargo from the neurites to the cell body.

Wallerian degeneration The active process of axon degradation to remove the segment that is no longer attached to the cell body after nerve injury.

Introduction

Neurons form the circuits that control and coordinate all animal behavior. These circuits are vital for normal function of the brain, spinal cord, and peripheral nerves. Intercellular communication generated by transfer of electrical activity between neurons and from neurons to target tissues is needed for every movement, sensation, and thought produced. A typical neuron is composed of three parts: a cell body, several dendrites, and an axon. The dendrites traverse short distances (≤ 2 mm) and bring activity to the cell body, while the axons are much longer and transmit activity away from the cell body. A neuron uses membrane depolarization to rapidly convey the information that is encoded by neural activity, from 'post-synaptic' sites in dendrites to the 'presynaptic' sites at the 'termini' of axons (i.e., intracellular communication). Arrival of such an 'action potential' at the axon terminus triggers release of neurotransmitters to propagate information across neural circuits or directly to the neuron's target (i.e., intercellular communication). The human neural circuits that connect motor cortex to spinal cord to muscle cover a distance approaching 2 m in some individuals. It is the axons of the motor cortex neuron and the spinal motor neuron that cover these long distances in the spinal cord and peripheral nerves, respectively. Disruption of axons through injury or disease blocks normal transmission of this neural activity and these long axonal processes are particularly vulnerable to damage.

While propagation of action potentials across these distances occurs on a millisecond time frame, neurons also use slower forms of intracellular communication by conveying

organelles, protein complexes, and RNA-protein complexes anterogradely and retrogradely in axons and dendrites. Axonal transport has been particularly well-studied, and has historically been separated into fast and slow components based on the speed of labeled proteins moving within axons. The fast component travels around 100 mm day^{-1} , while the slow component travels around 1 mm day^{-1} (Shah and Cleveland, 2002). The cargos that move in the fast component are important for the growing tips of developing axons ('growth cones') and for the proper function of mature synapses. These include organelles, membranes, proteins, and RNAs. Fast axonal transport also conveys signals to the cell body of the neuron that can regulate gene expression. The slow component of axonal transport ferries cytoskeletal polymers and other soluble proteins for maintenance of axon integrity (Shah and Cleveland, 2002). Movement of organelles and complexes within axons and dendrites is essential for neural function and plays key roles in neural repair.

'Molecular motors' or 'motor proteins' play the role of pack animals and messenger pigeons to move cargos to and from the neuron's cell body along axons and dendrites. There are several types of motors but all share certain attributes, being multisubunit protein complexes that harness the energy of ATP hydrolysis to move along 'cytoskeleton' tracks in axons and dendrites. The rates at which motors ferry cargo depends on a number of criteria, but *in vitro* speeds of individual mammalian motors and also membrane bound vesicles in axons have been clocked at up to $4 \mu\text{m s}^{-1}$. This speed is probably not maintained consistently, as transport cargos are frequently observed to pause and even reverse directions

within axons and dendrites. Nonetheless, the movement of cellular components over very long distances in relatively short amounts of time is a rather herculean task that requires a significant expenditure of energy. Regulation of motor protein motility, cargo composition, and the organization of the cytoskeleton tracks within neurons can dramatically influence neuronal function, particularly during the repair of neural connectivity after injury. Motor-dependent transport mechanisms are also prime targets for dysregulation in disease processes that affect the nervous system.

This article will cover the contributions of this subcellular trafficking to neural repair, focusing on the axon. During development, axon growth is characterized by a rapid elongation phase until targets are reached. At this point an alternate mode of arborizing growth allows axons to contact target cells and initiate synaptogenesis. This short-distance sprouting of axons may also underlie a presynaptic contribution to synaptic plasticity. It should be noted that development and regeneration of axonal processes share many common mediators; so much of this article will also be relevant to events in developmental growth. Furthermore, many of the mechanisms outlined here are also relevant for repair of dendritic processes; although we focus on the repair of axons because of their fundamental importance in long-range communication in the nervous system, the reader is referred to 'Further Reading' section for a recent publication examining mechanisms of dendrite repair.

Constituents for Intracellular Communication and Trafficking in Neurons

The Neuronal Cytoskeleton Provides Structure, Highways for Communication, and Growth Functions

The three components of the neuronal cytoskeleton, microfilaments, intermediate filaments, and microtubules, provide structural integrity for mature axons and dendrites. Their continued maintenance throughout the life of the animal supports normal axonal and dendritic function. During developmental and regenerative events the cytoskeleton itself is highly dynamic; polymerization and depolymerization of microtubules and microfilaments in the distal ends of axons (and dendrites) contributes to both growth and pathfinding for axons and dendrites (Lowery and Van Vactor, 2009). Posttranslational modifications to the cytoskeleton and associated proteins can contribute to neuronal cytoskeletal dynamics – for example, cyclin-dependent kinase 5 (CDK5) impacts all three components of the neuronal cytoskeleton (Smith, 2003). Motor proteins can contribute to cytoskeletal dynamics by moving short segments of cytoskeletal polymers or oligomers as cargo (Shah and Cleveland, 2002). This is particularly relevant during axon growth and arborization, but even after synapses are formed, transport of cytoskeletal components is likely to be important for maintenance of axons. The bulk of this movement appears to occur much more slowly than the speeds typically associated with fast axonal transport. However, this slow movement (and perhaps all of the slow component of axon transport) is thought to be driven by 'fast' motors that move the cargo intermittently (Li *et al.*, 2012).

Microfilaments

Microfilaments are composed of two strands of actin monomers twisted into helical filaments that have intrinsic polarity. All cells, including neurons, have a microfilament-rich, mesh-like network on the cytoplasmic side of the plasma membrane. This 'cortical actin' has been proposed to provide a tethering platform for non-translocating but active motors, allowing them to function in unison to move small microtubules and attached cellular components within axons and dendrites (Myers *et al.*, 2006). Polymerization and depolymerization of microfilaments at the distal end of growing axons and along the axon shaft, respectively, also play key roles in axon growth and branching. A large number of proteins have been identified that regulate actin dynamics and the spatial organization of actin filaments in growth cones (Lowery and Van Vactor, 2009).

Spectrin is another filamentous protein that is concentrated beneath the plasma membrane, providing a cross-link for the actin cytoskeleton and anchor for complexes of signaling and structural proteins, including membrane-linked protein complexes (Machnicka *et al.*, 2014). Different spectrin isoforms are concentrated in subdomains of the axon (e.g., axon initial segment, nodes of Ranvier) that are needed for neural function (Eshed-Eisenbach and Peles, 2013). Axonal spectrins can be rapidly proteolyzed following axonal injury (Buki *et al.*, 1999), so reconstruction of these axonal domains and repopulating the axon with spectrin is a key event in transitioning a regrown axon into a functional axon.

Intermediate filaments

Neuronal intermediate filaments do not provide tracks for motors, but these proteins have important roles in both axon function and growth. Neuron-specific intermediate filaments provide structural support for growing axons, and accumulation of one class of filaments, the neurofilaments, dictates the final caliber of an individual axon (Lee and Cleveland, 1996). The pathophysiology of some human diseases is accompanied by alterations in neurofilament deposition, including amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Parkinson's disease, and several neuropathies (Perrot and Eyer, 2009). Posttranslational modification of the polypeptides that make up neurofilaments contributes to their axonal delivery (Szaro and Strong, 2010). The intermittent 'slow component' of microtubule-based transport of neurofilaments may contribute to the precise deposition of intermediate filaments along the axon (Shea and Lee, 2013). High levels of neurofilament mRNA and protein levels are typically associated with more mature neurons. In fact, synthesis of neurofilaments protein falls after axotomy of peripheral nerves and stays relatively lower during regeneration (Hoffman and Cleveland, 1988). In some cases phosphorylated neurofilaments accumulate in the soma of regenerating axons, suggesting that transport into growing axons is reduced (Goldstein *et al.*, 1987).

Another type of intermediate filament protein in neurons, peripherin, is expressed primarily in peripheral neurons. In contrast to neurofilament expression, peripherin expression increases during regenerative growth, falling back to basal levels as axons near their target tissues (Reid *et al.*, 2010). The axonal inclusions seen in some ALS model mice have been shown to include peripherin, and overexpression of the

ALS-linked protein TDP43 has been shown to cause a prolonged increase in peripherin expression after peripheral nerve injury (Swarup *et al.*, 2012). This implies that the ALS model has an altered course of neural repair.

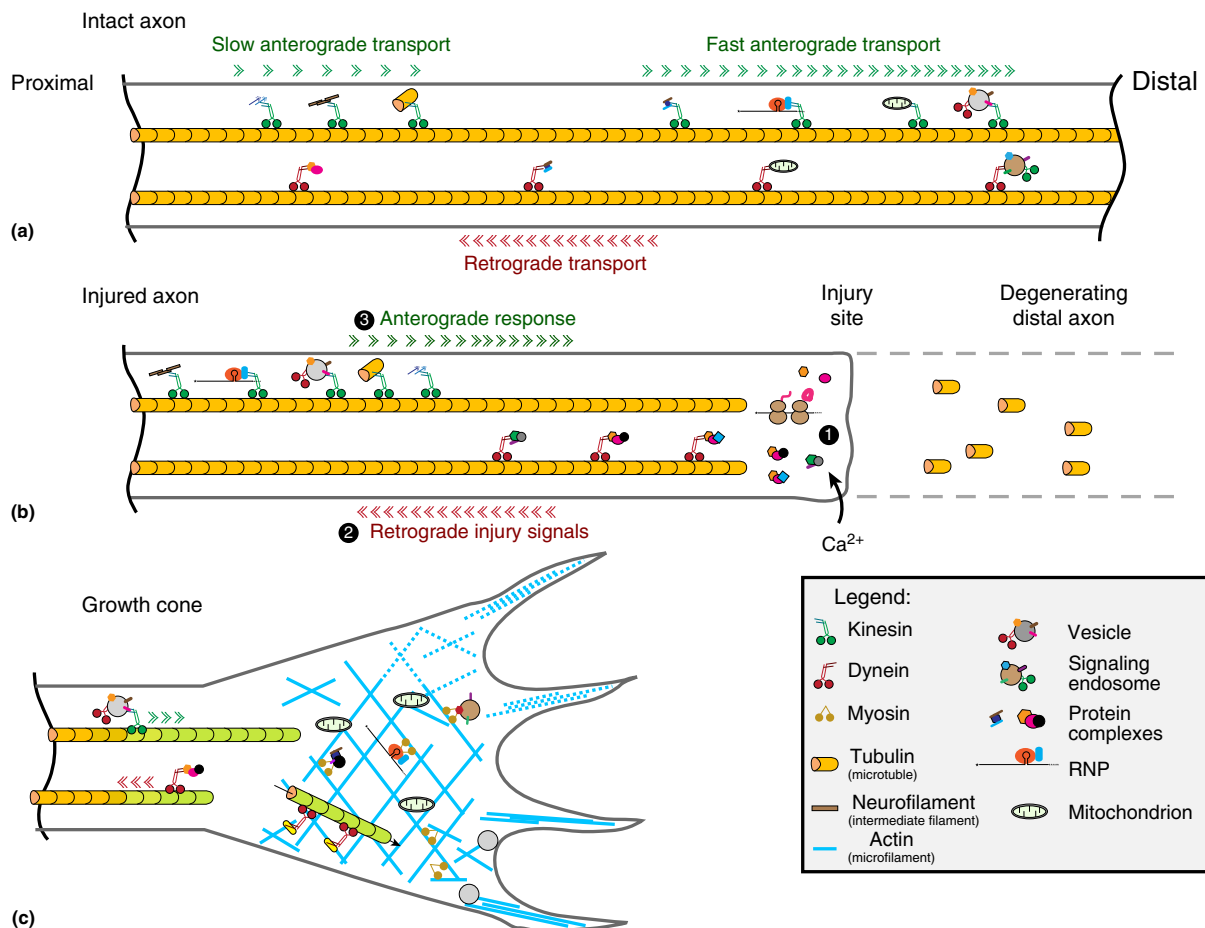
Microtubules

Microtubules are the largest of the cytoskeletal polymers, and like microfilaments, have intrinsic polarity and the potential to be highly dynamic. Each microtubule is made of 13 parallel protofilaments, which are polymers of α - and β -tubulin heterodimers (Prokop, 2013). The organization of the heterodimers imparts the polarity to microtubules, with the α subunit nearer to the 'plus end' and the β subunit is nearer to the 'minus end.' In axons, microtubules have uniform polarity with distal plus ends oriented away from the cell body, but dendrites have mixed microtubule polarity (Baas and Lin, 2011). In motor-based transport this polarity is important because motor proteins typically prefer to move in either a

plus-end or minus-end direction (Figure 1(a)). Therefore, the polarity of microtubules creates a situation in which motors exhibit unidirectional movement in axons and bidirectional movement in dendrites.

Microtubules shrink and grow due to dimer addition and loss at their plus ends in a process called dynamic instability (Mitchison and Kirschner, 1984). Plus ends are the most dynamic, and have a cadre of associated proteins, known as 'plus end tracking proteins' (+TIPs), that regulate their behavior (Schuyler and Pellman, 2001). Growth and shrinkage occur less frequently at minus ends, in part because minus-end binding proteins can both protect against microtubule depolymerization as well as prevent addition of tubulin dimers (Goodwin and Vale, 2010). With neural injury, the microtubules and neurofilaments are depolymerized in the axon near the injury site on the proximal side (Figure 1(b)).

The importance of microtubules in axon growth has been known for decades. The expression of tubulins increases



during regeneration, recapitulating developmental expression levels (Hoffman and Cleveland, 1988). Microtubule polymerization is a critical and highly regulated process in axon growth; and drugs that inhibit polymerization can block axon growth in cultured neurons (Tanaka *et al.*, 1995). Interestingly, stabilization of microtubules with taxol can augment spinal cord regeneration (Hellal *et al.*, 2011). Microtubule acetylation has been linked to microtubule dynamics and to regeneration (see below) (Cho *et al.*, 2013). Microtubule-associated proteins (MAPs) also modify microtubule structure (e.g., ‘bundling’ of microtubules along the axon shaft), microtubule stability, and motor protein processivity (Hirokawa, 1994).

The Growth Cone Is Uniquely Suited for Its Job

The growing ends of developing and regenerating axons form a specialized structure called the ‘growth cone’ (Figure 1(c)). The shape and behavior of the growth cone is dictated by both extrinsic and intrinsic factors (Gomez and Letourneau, 2014). Examples of extrinsic factors are substrate molecules in the extracellular milieu or matrix (ECM) present along the growth trajectory, and molecules secreted locally into the environment by neighboring cells. Intrinsic factors include cytoskeletal dynamics that are often regulated by proteins within the growth cone like the Rho, Rac, and CD42 small GTPases, as well as the spatial distribution of receptors for sensing environmental cues. Generation of a growth cone after an axon is injured is generally thought to be essential for regenerative growth. One of the initial events in establishing a growth cone after injury is the entrance of calcium into the damaged axon – as outlined in subsequent sections, this triggers a number of changes that not only allow this important structure to form, but also facilitates signaling to the neuronal cell body (Bradke *et al.*, 2012).

While growth cones can vary greatly in shape and behavior, a typical advancing growth cone is composed of three major domains: central (C-), transition (T-), and peripheral (P-) domains (Gomez and Letourneau, 2014). The C-domain is made up of stable microtubule bundles that come from the axon shaft. These microtubules tend to be acetylated in the axon shaft and as microtubules enter the growth cone, tyrosination becomes the prominent posttranslational modification on the tubulins (Figure 1(c)). Acetylated tubulin is associated with stable polymers, while tyrosinated tubulin is associated with more dynamic polymers (Janke, 2014). The C-domain typically contains the more dynamic microtubules that can extend linearly in sync with the advancing edge of the growth cone. However, microtubules are responsive to environmental cues, and can shrink back or become curved in order to pause advancement or to reorient direction of growth (Lowery and Van Vactor, 2009). The P-domain at the very leading edge of the growth cone is composed of actin-based projections called lamellipodia and filopodia. Microfilaments in lamellipodia form a dynamic mesh, with new filaments being added to growing edge and old filaments being broken down at the back of the mesh near the C-domain. Force-generating myosin motors (see below) control the contractility of this meshwork (Medeiros *et al.*, 2006). The filopodia, are finger-like projections outward from lamellipodia that extend and retract to sample their environment. They contain

dynamic bundles of actin filaments, which are thought to form through the concerted action of myosin-driven ‘retrograde flow’ of the actin mesh in lamellipodia and the extension of actin polymers in the filopodia (Medeiros *et al.*, 2006).

The dynamic nature of the cytoskeleton in growth cones is what allows these structures to respond to environmental cues, and growth cones ultimately must transition into synaptic endings or other types of functional axon termini. When the growth cone adheres to a substrate such as laminin, the actin cytoskeleton becomes anchored to receptors, preventing retrograde movement of actin, which causes the filopodia of the growth cone to extend or protrude via distal actin polymerization (Mitchison and Kirschner, 1988). As this happens, myosin motors create tension between F-actin filaments and an F-actin arc at the base of the growth cone closest to the axon shaft. This allows microtubules of the C-domain to move forward, leading to an engorged growth cone; after this, the F-actin arc depolymerizes and the membrane collapses around the stable microtubule bundles, consolidating this region into an extension of the axon shaft (Lowery and Van Vactor, 2009). It should be noted that localized synthesis of β -actin protein and actin-interacting proteins in growth cones may also contribute to actin dynamics in the growth cone, as well as along the axon shaft where collateral branches form (Zhang *et al.*, 1999; Spillane *et al.*, 2013). Motor protein regulators (e.g., lissencephaly protein (Lis1) and nuclear distribution E-like 1 (Ndel1), see below) are prominent in growth cones of regenerating axons in culture (Smith *et al.*, 2000; Niethammer *et al.*, 2000), and can play an active role in promoting growth cone motility.

Molecular Motors Ferry Cargo to and from Axons and Provide Force Generation for Axon Growth

Molecular motors transport of a wide range of cargos in axons including organelles, vesicles, mRNA, and signaling molecules. The mRNAs move as RNA–protein complexes termed ‘ribonucleoproteins’ (RNPs) (Gomes *et al.*, 2014). Cargo may be deposited along axons, or trafficked anterogradely to synapses and growth cones or retrogradely back to the cell body. Transport along microtubules is carried out by kinesins, which for the most part are plus-end directed, and cytoplasmic dynein, which is minus end directed (Hirokawa *et al.*, 2010). Cytoplasmic dynein also has the capacity to become tethered to the cortical actin cytoskeleton within axons and growth cones. The force generating heads that project into the axon away from the cortical actin are then able to ‘kick’ short microtubule segments, causing them to move with their plus-ends leading (Mazel *et al.*, 2014). In addition to microtubule motors, axons contain myosin motors that ferry cargo on microfilament tracks. While these three classes of motors share common features, including motor domains that require ATP, they are also quite distinct with respect to molecular organization as well as motility regulatory mechanisms.

Kinesin

There are 14 classes of kinesin superfamily proteins organized in part by where the motor domain is located within the polypeptide (Hirokawa *et al.*, 2009). ‘Conventional’ kinesins in the Kinesin-1 group are heterotetramers. Two ‘heavy chains’

(kinesin heavy chain, KHC) form a dimer with globular head domains that interact with microtubules and undergo force-producing conformational changes in response to ATP binding and hydrolysis. KHC stalk regions contain the dimerization domains, while the C-terminal tails interact with 'light chains' (kinesin light chain, KLC) that confer cargo binding (Marx *et al.*, 2006). Though the motor domains are well conserved, substantial variability in other parts of the heteromer, including KLCs, can confer cargo selectivity. Kinesins move a variety of cargo toward axon termini, so they are critical for maintenance of mature axons and synapses, as well as for delivery of construction materials, receptors, adhesion proteins, and membranes to growing axons (Figures 1(a) and 1(b)). Different kinesins in this large superfamily are responsible for anterograde transport of specific cargo types within axons. Some of the specificity is likely derived from distinct heavy chains, but cargo-specific transport by the same heavy chain can be modulated by associated light chains (Hirokawa *et al.*, 2009).

Dynein

Although a few kinesins (e.g., KIFC2) may be capable of minus end-directed transport (Hirokawa *et al.*, 2009), the predominant minus end-directed motor in axons is cytoplasmic dynein (Figures 1(a) and 1(b)). This huge protein complex contains 12–14 subunits and has 2 identical heavy chains (dynein heavy chain, DHC) that form a dimer (Vale, 2003). The large ring-shaped motor domains at one end of DHC are ATPases, and are the force-generating component of the protein complex. A stalk extending from each DHC motor domain interacts with microtubules. DHC's stem interacts with an array of other 'chains.' These include two intermediate chains, 2–3 light intermediate chains and 3 distinct light chains. It is generally believed that the accessory subunits bind to adaptors that confer cargo specificity and/or control dynein processivity (Allan, 2011). The cytoplasmic dynein holoenzyme is fairly uniform in subunit composition. However, multiple genes encode related but distinct isoforms of each accessory subunit. Moreover, alternative splicing and post-translational modifications can alter subunit composition to optimize dynein motors for different tasks (Pfister *et al.*, 1996). Because axonal dynein moves retrogradely toward the cell body, it is important for molecular communication between axon endings and the nucleus. In intact axons, dynein carries information related to synapses and target tissues, and convey signals that prevent continued axon elongation after developmental growth has ceased (Smith and Skene, 1997).

Myosin

Myosin is a force-generating ATPase that underlies muscle contraction. The classic muscle myosin remains tethered in place and brings about muscle contraction by exerting tugging forces on actin filaments. Some myosins in neurons also function by exerting forces on the subcortical actin network (Figure 1(c)) – these are particularly prominent in regulation of growth cone dynamics needed for development and regeneration (Lowery and Van Vactor, 2009). Other members of the myosin superfamily are processive and can ferry cargo, using microfilaments as transport tracks. Some myosins (e.g., myosin Va) are plus-end directed, and others (e.g., myosin VI)

are minus-end directed (Kneussel and Wagner, 2013). Myosin transport is typically thought to occur over shorter distances than the long-range axonal transport associated with microtubule motors (i.e., at branch points, in growth cones, in synapses, etc.).

Regulation of motor proteins – Activation, inhibition, and cargo selectivity

Much of what we know about motor regulation has come from *in vitro* studies with purified proteins or cultured cells. However, studies in neuronal cultures suggest that these mechanisms are likely relevant to nerve repair. Because of their force-generating capacities, their plethora of cargos and the many cellular processes influenced by motors, a strictly enforced regulatory network must be in place to coordinate motor behaviors and modulate their behaviors in response to environmental cues and intrinsic cues. At the most basic level, ATP must be in good supply for all motors to function properly. Perhaps intuitively, mitochondrial ATP was shown to be the source of ATP for fast axon transport of mitochondria (Zala *et al.*, 2013). However, the source of ATP for transport of some vesicular cargos appears to be from glycolytic machinery directly associated with the vesicle (Zala *et al.*, 2013). Thus, the cargo itself can impact activity of the motor protein carrying it, and there is a growing understanding that complexes of motor proteins, scaffolds, and regulatory proteins can modulate motor processivity and force generation. The selection of a motor's cargo is in part regulated by cargo adaptors, proteins that bind to both the motors and to a specific type of cargo (Fu and Holzbaaur, 2014). An interesting new development is that mammalian cytoplasmic dynein is incapable of moving along microtubules *in vitro* in the absence of either cargo or cargo adaptor proteins (McKenney *et al.*, 2014; Schlager *et al.*, 2014). It remains to be determined if this is also the case within the confines of the axon shaft.

Knowledge of motor protein regulation has been advanced by studies of neurological disorders. For example, mutant forms of huntingtin (HTT) protein that cause Huntington's disease form aggregates in axons and dendrites that can block transport by motors (Lee *et al.*, 2004). Wild-type HTT can bind directly to dynein intermediate chain (DIC) subunits (Caviston *et al.*, 2007), while an HTT binding partner HTT-associated protein 1 (HAP1) interacts with KLC (McGuire *et al.*, 2006). Moreover, HTT is involved in linking the glycolytic machinery to cargo vesicles to provide a ready source of ATP for fast axonal transport (Zala *et al.*, 2013). While pathogenic HTT remains capable of interacting with motor complexes, it reduces motor association with microtubules through a c-Jun N-terminal kinase 3 (JNK3)-mediated phosphorylation of kinesin-1 (Morfini *et al.*, 2009).

Most organelles have all three classes of motors associated with them, usually more than one of each kind of motor. Interactions between motors may occur via scaffolding systems (Fu and Holzbaaur, 2014). There is growing evidence that motors of one class can affect activity of other classes of motors (Ally *et al.*, 2009; Levi *et al.*, 2006). While many of these interactions have been observed, how functions are coordinated to regulate cargo movement in the same axons has not been extensively studied. The precise trafficking of an individual cargo likely involves regulation of the stoichiometry

of motor association, a tug of war between opposing motors of all three classes, the proximity of transport tracks, as well as posttranslational modification of proteins and cargos (Fu and Holzbaur, 2014; Kardon and Vale, 2009). Each of these mechanisms is operant in the mature as well as injured and regenerating axon. One known molecular control mechanism involves the JNK-interacting protein 1 (JIP1) that acts as a scaffold protein that binds to both KHC and the dynein activator, dynactin (Fu and Holzbaur, 2013). JIP1 binding to KHC relieves an intramolecular autoinhibition of kinesin motor activity. The dynactin subunit p150^{Glued} competes with KHC for binding to JIP1 so that KHC is inactive when JIP1 is bound to p150^{Glued}. Kinesin dysinhibition and dynactin binding by JIP1 is regulated by JNK-dependent phosphorylation, which can lead to directional changes in movement of amyloid precursor protein-containing vesicles in axons (Fu and Holzbaur, 2013). JIP1 is also responsible for directional selectivity of autophagosomes in axons. Autophagosomes are retrogradely transported by dynein, and are presumed important for axonal maintenance and growth. Dual mechanisms appear to prevent activation of kinesin on these organelles (Fu et al., 2014). JIP associates with autophagosomes through the adaptor, microtubule-associated protein light chain 3 (LC3), and LC3 prevents JIP binding to KHC.

Other proteins can bind directly to motors and regulate force generation and/or processivity *in vitro* and also impact axon transport in cultured cells. For example, loss of Lis1 protein reduces retrograde transport of acidic vesicles and Lis1 overexpression increases speed and run lengths of retrogradely moving vesicles in axons (Pandey and Smith, 2011). Lis1 seems to function with dynein to promote microtubule advance during growth cone remodeling and fast axon growth (Grabham et al., 2007). During embryogenesis, Lis1 contributes to neuronal migration (Reiner and Sapiro, 2013), and it interacts with two paralogous dynein-binding proteins Ndel1 and Nde1 (Bradshaw et al., 2013). Ndel1/Nde1 also serves as a platform for modulation of motor activity by posttranslational modification. Ndel1/Nde1 strengthens the Lis1 association with dynein and increases force production by promoting cooperation between multiple dynein motors (McKenney et al., 2010). Interestingly, the neuronally enriched kinase Cdk5, phosphorylates Ndel1/Nde1 (Niethammer et al., 2000). Phosphorylation of Ndel1 by Cdk5 is important for dynein-dependent organelle transport in growing axons extended by adult sensory neurons in culture (Pandey and Smith, 2011). Recent evidence suggests that Cdk5 is an important regulator of membrane vesicle dynamics in growth cones (Tojima et al., 2014; Hsiao et al., 2014). Other kinases that contribute to regulation of axonal transport have also been observed to impact regeneration. For example, glycogen synthase kinase 3 (GSK-3) exerts a negative influence on axon transport in squid and fly axons (Morfini et al., 2002; Weaver et al., 2013). GSK-3 has shown varying effects on axon regeneration, but recent work indicates that sustained activation of GSK-3 supports regeneration of axons in peripheral nerve (Gobrecht et al., 2014). Clearly, the interplay between signals regulating axon transport and those regulating axon growth is complex, and just beginning to be understood.

How extracellular stimuli regulate motor proteins is also not well understood, but some pathways are beginning to

emerge. For example, activation of semaphorin receptors along growing axons stimulates both anterograde and retrograde movement of axonal vesicles (Li et al., 2004). Semaphorin stimulation of distal axons shifts α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor subunit localization to dendrites through a coordinated retrograde signaling in axons and subsequent anterograde transport in dendrites (Yamashita et al., 2014). Neurotrophins also impact transport. These ligands increase endosome binding to dynein through DIC phosphorylation, thereby stimulating retrograde transport (Mitchell et al., 2012). Interestingly, neurotrophin receptors (Trks) are endocytosed in axons along with their ligands and can be retrogradely transported as a 'signaling endosome' that carries downstream effector kinases to the cell body (Schmieg et al., 2014). During development, many neuronal populations compete for binding to target-derived neurotrophins and other trophic agents. The neurons that do not secure an appropriate source of trophic support are eliminated through naturally occurring cell death. Culture preparations in which the neurons' cell body and axons are compartmentalized have shown qualitative differences in the response to neurotrophic stimulation of distal axons compared to the cell body (Campenot and MacInnis, 2004). This suggests that the signaling endosome conveys a signal different from Trk receptors endocytosed along the cell body.

Activation of Trk receptors along the axon has been shown to modulate transport of RNPs into axons, with accumulation and translation of mRNAs adjacent to the trophic stimulus along axons (Willis et al., 2007). Target-derived nerve growth factor (NGF) in sympathetic neurons was recently shown to coordinate transcriptional responses in the cell body with delivery of mRNA encoding a survival promoting protein, B-cell lymphoma w (Bcl_w), to the axons (Cosker et al., 2013). Given that peripheral nervous system (PNS) injury results in increased expression of neurotrophins and other trophic factors by Schwann cells in the severed nerve stump (Funakoshi et al., 1993), it is intriguing to hypothesize that signaling endosomes moving retrogradely to the cell body would be used to coordinate anterograde delivery of macromolecules to the regenerating axons. The cargos that motor proteins deliver into axons obviously play a role growth and regeneration. The axon needs new cytoskeleton and membrane components to grow, as the severed segment of the axon distal to injury site undergoes Wallerian degeneration in mammals (see below). It was also recently recognized that protein products of axonally transported mRNAs support regeneration in the PNS (Donnelly et al., 2013), and there is some evidence that this may occur after spinal cord injury as well (Willis et al., 2011).

Neural Repair Mechanisms

Injury Changes Axon Structure and Axonal Transport to Initiate a Regeneration Program

With damage and severing of an axon, the segment distal to the site of injury (i.e., the portion no longer connected to the cell body) is actively degraded through Wallerian degeneration. In the portion proximal to the site of injury, influx of Ca²⁺ from extracellular spaces promotes membrane sealing and locally activates proteases (Bradke et al., 2012). Proteases

degrade axonal proteins including cytoskeletal elements, and this is necessary for the initiation of axonal growth (Ziv and Spira, 1997). The increase in axoplasmic Ca^{2+} also triggers retrograde signaling through transport mechanisms (Rishal and Fainzilber, 2010). This shift in retrograde transport ultimately helps to turn on expression of genes that support growth (regeneration-associated genes (RAGs)) and turn off expression of genes associated more with mature neuronal functions (Rishal and Fainzilber, 2010). For example, axotomy in the PNS increases the expression of growth-associated protein-43 (GAP-43) and $\text{T}\alpha 1$ -tubulin, but decreases expression of neurofilaments. Injury to the mammalian brain or spinal cord, in contrast to peripheral nerves, often does not activate expression of RAGs, and this may in part, underlie the reduced growth capacity of mammalian central nervous system (CNS) neurons after injury (Hoffman, 2010). Many strategies to increase axon regeneration after central injuries have focused on altering the 'intrinsic growth capacity' of CNS neurons through changing gene expression programs (Mar *et al.*, 2014). ECM glycoproteins and glial-derived proteins in the injured brain and spinal cord also actively block axon growth, and overcoming these 'roadblocks' is often a complimentary strategy to support CNS neural repair (Geoffroy and Zheng, 2014).

Some of the mechanisms that underlie the shift in retrograde axonal transport are known. Axonal JNK3 is locally activated after injury, with a resulting increased association of a complex of JNK-interacting proteins 3 (JIP3) and activated JNK3 with the p150^{Glued} subunit of dynactin, and increased retrograde transport of activated JNK3 (Cavalli *et al.*, 2005). This leads to activation of c-Jun in the neuronal cell body that presumably supports expression of genes needed for regeneration (Ruff *et al.*, 2012). Interestingly, JIP3 binding to KHC can also increase kinesin motility (Sun *et al.*, 2011). It is not clear whether the JIP1-dependent retrograde transport of autophagosomes mentioned above contributes to neural repair; however, it is intriguing that both JIP1 and JIP3 can coordinate anterograde and retrograde transport (Fu and Holzbaur, 2014). This provides a powerful mechanism to coordinate transport to meet physiological needs of the neuron and potentially to recycle proteins between axons and cell body.

Calcium influx coordinates the injury response

The influx of Ca^{2+} into injured axons also triggers translation of locally stored mRNAs (Perry and Fainzilber, 2014). These mRNAs are transported into axons as RNPs and it remains unclear exactly how Ca^{2+} influx activates translation locally in axons, but it likely includes a commensurate release of Ca^{2+} stores from sites in axons (Vuppalaanchi *et al.*, 2012). Axonal mRNAs known to be translated by increased axoplasmic (Ca^{2+}) include Importin $\beta 1$, Ran specific binding protein 1 (RanBP1), and vimentin (Gomes *et al.*, 2014). These encode proteins that contribute to a retrograde injury signal in a complex highly orchestrated process (Rishal and Fainzilber, 2010). The newly translated Importin $\beta 1$ in axons heterodimerizes with Importin $\alpha 3$ (Hanz *et al.*, 2003). Localized synthesis of RanBP1 facilitates this by releasing $\alpha 3$ from an anterograde transport complex (Yudin *et al.*, 2008). Interaction of the Importin $\beta 1/\alpha 3$ complex with dynein is needed for the transport of proteins containing nuclear localization signal (NLS) to the nucleus. The cargo of the dynein-Importin $\beta 1/\alpha 3$

complex includes other proteins that are locally synthesized with injury (e.g., signal transducer and activator of transcription 3 α , Stat3 α) as well as proteins that are locally activated through injury-induced posttranslational modifications (e.g., extracellular-regulated kinase, Erk 1/2) (Ben-Yaakov *et al.*, 2012; Perlson *et al.*, 2005). Vimentin has been considered as a glial intermediate filament protein in the nervous system, but vimentin can be expressed in neurons, and its mRNA has been shown to localize to axons (Willis *et al.*, 2007). Curiously, proteolytic cleavage of vimentin in injured axons generates a 'scaffold peptide' for activated Erk 1/2, a mitogen-activated protein kinase (MAPK), to bind to Importin $\beta 1/\alpha 3$ dimer. This kinase is transported to the cell body with dynein, where it activates Elk1 (and possibly other) transcription factors (Perlson *et al.*, 2005). It is likely that such a posttranscriptional mechanism is shared with other physiological responses of axons, since neurodegenerative stimuli, neuronal survival promoting stimuli, and a neurotropic virus are now recognized to couple events in the distal axons to cell body responses through translation of axonal mRNAs (Koyuncu *et al.*, 2013; Baleriola *et al.*, 2014; Cox *et al.*, 2008; Cosker *et al.*, 2013). These and other studies indicate that axonal injury induces a localized, integrated response to shift retrograde transport toward cargos that assist in mounting a regenerative response. The full extent of these cargos has yet to be determined; the reader is referred to references in 'Further Reading' section below for additional information on cargos.

Calcium influx transforms the injured axon to support regenerative growth

Much of what we know about axonal transport and localized growth mechanisms have been derived from studies in cultured neurons, where events can be spatially and temporally monitored with high precision. Neurons cultured from invertebrate organisms have provided a platform for visualizing cytoskeletal dynamics in growing and injured axons. For example, the initial studies on Ca^{2+} signaling after axonal injury resulted from work in *Aplysia* neurons (Ambron and Walters, 1996). The Ca^{2+} gradient that forms in these axons after injury occurs in both the proximal (intact) and distal (severed) segments of the injured axon, but only the proximal end undergoes regeneration (Ziv and Spira, 1997). The Ca^{2+} dependent events including membrane sealing and cytoskeletal restructuring controlled by calcium-activated calpain, syntaxin, synaptotagmin, and synaptobrevin are critical for neural repair (Bradke *et al.*, 2012). Indeed, activation of calpain through a Ca^{2+} -dependent mechanism is required for growth cone formation (Gitler and Spira, 2002). Proteolysis of cytoskeletal elements in the cut axon results in a reorganization of microtubule polarity in *Aplysia* (Sahly *et al.*, 2006). This may generate a 'trap' for capturing both anterogradely and retrogradely transported vesicles, which are then available to fuse with the plasma membrane for efficient and effective formation of a growth cone (Erez and Spira, 2008). Work in vertebrate neurons indicates that activation of localized protein synthesis in the distal cut axon is also needed for growth cone formation (Verma *et al.*, 2005). Though the proteins synthesized were not identified, TC10 and Exoc3 mRNAs, which encode components needed for plasma membrane expansion, are locally translated in growing axons (Gracias *et al.*, 2014). Stimulation of MAPKs and

mammalian target of rapamycin (mTOR) converge on protein synthesis machinery and are needed for growth cone formation (Bradke *et al.*, 2012). Thus, localized mRNA translation supports the initial transition of an injured axon into a regenerating axon. Subsequent growth of the axon requires delivery of necessary building blocks of cytoskeleton and membrane components, which is accomplished by anterograde transport of proteins, vesicles, and RNPs.

Unexpected Roles for Proteins in Growing Axons

Axonal functions for nuclear proteins

With technical advancements in subcellular imaging, protein/RNA detection and analyses, and study of signal transduction mechanisms, previously unrecognized functions have been ascribed to proteins and protein networks that support neural repair. For example, epigenetic regulation is now well recognized as a means to modulate expression of entire gene families, and there is clear potential for such chromatin-based modulation to contribute to neural repair (Finelli *et al.*, 2013). However, recent work from two groups have uncovered intra-axonal functions for histone deacetylase (HDAC) proteins that are classically viewed as modifiers of chromatin structure and, hence, transcription. The injury-induced Ca^{2+} entry into axons noted above produces a retrograde ' Ca^{2+} wave' that triggers exit of HDAC5 from the nucleus in a protein kinase C-dependent pathway. In keeping with classic functions of HDACs, the loss of HDAC5 from the nucleus supports expression of genes whose encoded proteins have roles in regeneration (e.g., c-Fos, c-Jun, KLF4, KLF5, and Gadd 45a; see Cho *et al.*, 2013 and references within). However, HDAC5 is also anterogradely transported into the injured axon, where it supports deacetylation of microtubules. This decreases stability of microtubules in the cut axon that is needed for growth cone formation just proximal to the injury site. Curiously, this anterograde movement of HDAC5 only occurred after peripheral nerve injury and not after optic nerve injury, a lesion of CNS axons (Cho *et al.*, 2013), pointing to another distinction between the central and PNS with axonal injury (Cho *et al.*, 2013). HDAC6 protein has also been shown to function in axons, but inhibition of HDAC6, rather than activation, brings growth-promoting effects (Rivieccio *et al.*, 2009). For HDAC6 inhibition, this allows for growth on nonpermissive substrates including myelin-associated glycoprotein and CSPGs that block axonal growth after spinal cord injury. These observations suggest that HDAC5 and HDAC6 have antagonistic functions in neural repair. However, HDAC6 may contribute to axon health because it facilitates the retrograde transport of protein aggregates for destruction in autophagosomes (Lee *et al.*, 2010). It will be intriguing in the coming years to discover the mechanism of action and targets of these deacetylases in developing, regenerating the mature axons.

New functions for adenomatous polyposis coli protein, a protein associated with colorectal cancer

Other proteins have shown multifunctionality that impacts axon development and likely axon regeneration. Adenomatous polyposis coli protein (APC) is a large, multidomain microtubule-interacting protein initially discovered because

truncating mutations lead to colorectal cancer in humans (Vogelstein and Kinzler, 2004). APC has also been linked to axon growth through promoting microtubule polymerization (Chen *et al.*, 2011). In addition to its function as a +TIP, APC interacts with many different proteins including β -catenin and GSK-3 β from the canonical Wnt signaling pathway (Clevers and Nusse, 2012). Initial studies in fibroblasts raised the possibility that APC's actions may also extend to interaction with mRNAs (Mili *et al.*, 2008). More recent work has shown that APC binds directly to more than 250 neuronal mRNAs (Preitner *et al.*, 2014). Interestingly, nearly half of those mRNAs were known to localize into growing axons of PNS sensory neurons, including the Importin β 1 and β -actin mRNAs that are mentioned above. Preventing APC binding to a tubulin isoform mRNA (TUBB2B2) decreased axonal levels of this mRNA and caused growth cones to shrink, which could significantly hamper the growth cone's ability to interact with growth substrates (Preitner *et al.*, 2014). Though it is not clear whether APC is needed for transport or translation of these mRNAs (or both), the protein is perfectly positioned in the growth cone to serve as a scaffold concentrating mRNAs and translational machinery near the processive tips of microtubules. There is work in developing neurons indicating that localized mRNAs accumulate along with translation factors and ribosome subunits adjacent to cytoplasmic tails of axon guidance receptors (Tcherkezian *et al.*, 2010). It will be interesting to see if APC also contributes to transport or translation of mRNAs in injured axons where localized protein synthesis is known to contribute to regeneration.

A protein kinase that can regulate both axon regeneration and Wallerian degeneration

Genetic screens in model organisms, including flies, worms, and mice, identified the 'dual leucine zipper kinase' (DLK) as regulator of both axon regeneration and degeneration (Tedeschi and Bradke, 2013). DLK has also been called 'ZPK' in mice and 'Wallenda' in *Drosophila*. Loss of DLK delays degeneration of axons but also attenuates RAG transcription in the cell body (Shin *et al.*, 2012a). Activation of DLK can lead to downstream activation of both JNK and 38 kDa mitogen-activated protein kinase (p38^{MAPK}). As emphasized above, JNK activation is important in neural repair, and downstream c-Jun dependent transcription activates expression of RAGs in the cell body after axonal injury (Shin *et al.*, 2012a). However, blocking JNK activation in severed axons distal to the injury site delays Wallerian degeneration, indicating dual functions for both DLK and JNK (Ferraris *et al.*, 2013). In addition to JNK's roles in axonal transport mediated through JIPs, JNK is known to target microtubule binding proteins that regulate dynamic stability of these cytoskeletal elements, which can influence both axon formation and regeneration (Tedeschi and Bradke, 2013). JNK phosphorylation triggers degradation of super cervical ganglion clone 10 (SCG10, also called Stathmin-2) in axons (Shin *et al.*, 2012b). SCG10 is a membrane bound protein that destabilizes microtubules (Riederer *et al.*, 1997). Degradation of SCG10 in the severed axon is linked to Wallerian degeneration, but SCG10 is rapidly replenished in the proximal axon segment proximal to the injury and this is needed for regeneration (Shin *et al.*, 2014). Destabilization of axonal microtubules through SCG10's actions near the growth

cone contributes to axon motility. JNK is also known to phosphorylate microtubule-associated protein 1b (MAP1b) in axons, which conversely increases stability of microtubules (Chang *et al.*, 2003). Interestingly, the mRNA encoding a protein kinase that can be an intermediate between DLK and JNK, MKK7, or mitogen-activated protein kinase kinase 7 (MAP2K7), localizes to proximal neurites of differentiating NIE-115 cells and primary hippocampal neurons (Feltrin *et al.*, 2012). This may provide a means to compartmentalize the functional outcome of JNK's activation in growing axons, by concentrating substrates that differentially modify microtubule dynamics to different sites along the axon (e.g., growth cone for destabilization vs. axon shaft for stabilization).

Regulation of microtubule motor proteins can directly support axon regeneration

Interestingly, kinesin and dynein motors have recently been shown to have unexpected roles in attenuating axon growth. The levels and/or activity of kinesin and dynein motors may contribute to axonal growth by providing a gauge for how much material the axon needs in order to grow (Albus *et al.*, 2013). In cultures of adult sensory neurons, partial depletion of kinesin and/or dynein heavy chains with siRNA actually increases axon length over time (Rishal, 2012). Legs at odd angles (Loa) mice that have reduced levels of dynein heavy chain 1 (Dync1h1) show increased axon length in developing limb buds indicating that this length sensing mechanism is also relevant for axon growth *in vivo* (Rishal, 2012). These data suggest that continual anterograde and retrograde signaling, analogous to a radar signal, contributes to axon growth rates (Kam *et al.*, 2009). However, the cargos that are needed for this, both those carried anterograde into the axons and those carried retrograde back to the cell body, are not clear.

Work from the Baas group using depletion of kinesins also suggested that localized interactions of kinesin and dynein motors could affect axon length (Liu *et al.*, 2010). They observed that partial depletion of kinesin-5 or kinesin-12 increased axon length in cultured neurons, but attributed this to a shift in the tractile forces along the distal axon to favor increased growth (Myers and Baas, 2007; Liu *et al.*, 2010). They hypothesized that loss of either kinesin allows cytoplasmic dynein to pull microtubules in the opposite direction, moving more microtubules from the axon shaft into the growth cone, resulting in both increases in axon length and the inability of the growth cone to turn properly without spatial control over the microtubules in the growth cone (Liu *et al.*, 2010). Regardless of whether this is a cargo-dependent signaling mechanism or localized force differentials through altered motor stoichiometry, the end result of decreasing motor protein activity results in the counter-intuitive effect of increasing axon length. It will be of substantial interest to determine how these mechanisms can be integrated with the injury-dependent and regeneration-associated changes in axonal trafficking outlined above that are essential for repair of neural connectivity.

Perspectives

Despite all that has been learned of neural repair mechanisms and the potential for using trafficking to encourage axons to

regrow after injury, even in the PNS where regeneration is relatively robust, regenerating axons only progress at 1–2 mm per day. Thus, finding a means to increase growth rates should bring huge benefits. However, whether increased growth rate can be accomplished by modulating axon transport dynamics *in vivo* remains to be determined. Interesting potential targets for regulating axon growth through trafficking include the motors themselves, their regulatory partners, kinases, cargo adaptors, and specific cargos.

In addition to overt axon elongation driven by growth cone advance, developing neurons must also undergo a different kind of axon growth even after synapse formation. Small-scale axon growth and remodeling of synaptic connections occur throughout adult life and underlies learning and memory. But this does not compare to increased length that nerves and axon tracts accomplish during growth from a late stage embryo to late adolescent, and is unlikely to be the mechanism underlying this elongation. There is some evidence that neurons switch to a stretch-induced interstitial growth as organisms increase their rostral-to-caudal dimensions (Suter and Miller, 2011). Cultured sensory neurons have been 'stretched' in an attempt to model this process, and within limits, stretching results in increased rates of axon growth while maintaining electrophysiological function (Pfister *et al.*, 2006). The mechanisms underlying this increased growth are not clear, but it would seem that the axon must lay down more cytoskeleton and membrane or suffer compromised integrity. Adult neurons also seem to tolerate this axon stretching less well than embryonic neurons, which could relate to age acquired cytoskeletal modifications or loss of developmental gene expression programs (Loverde *et al.*, 2011). Still the concept of stretch-induced growth may be a promising lead for future neural repair strategies once the mechanism(s) can be uncovered, particularly if this might complement other growth-promoting strategies.

There are intersections between neural repair, axonal transport, and neurodegenerative mechanisms that warrant mention. As indicated above, DLK and SCG10 have dual roles in regeneration and axonal degeneration. The severed axon has lost its cell body source of proteins, RNPs, and organelles. Consequently, this severed axon segment needs to survive on existing entities, or gather what it needs from its local environment (e.g., glia). To some extent, defects in axonal transport can result in a similar situation. Reduced anterograde delivery of axonal components produced in the cell body and reduced transport of signals originating in distal targets can compromise the health of the neuron. Indeed, transport defects can be a key factor in neurological diseases and mutations or depletion of motor proteins or their regulators are emerging as causative agents in disease. Beyond Huntington's disease mentioned above, defects in axon transport in age-related dementias such as Alzheimer's disease (Kanaan *et al.*, 2013). The distances that motor axons traverse may make them particularly vulnerable to transport deficits. Consistent with this notion, a KHC mutation that disrupts microtubule binding has been found in patients with hereditary spastic paraplegias, mutations in mouse DHC lead to an ALS-like motor neuron degeneration (Fichera *et al.*, 2004; Ikenaka *et al.*, 2012) and a mutation in the p150^{Glued} dynactin subunit has been found in an autosomal dominant form of lower motor neuron disease (Puls *et al.*, 2003).

Symptoms of neurodegenerative diseases including ALS have been linked to loss of synapses that may be initiated by 'dying back' of the distal axon (Adalbert and Coleman, 2012). Blocking the degeneration of their axons can prevent the death of neurons in some neurodegenerative diseases. For example, deletion or inhibition of DLK is neuroprotective in cellular and animal models of Alzheimer's disease, Parkinson's disease and glaucoma (Ferraris *et al.*, 2013). Expression of axonally targeted Ube4b-NMNAT1 fusion protein (Wallerian degeneration slow protein, 'Wld^s') can delay Wallerian degeneration and prevent neuron loss in several different animal models of neurodegeneration (Coleman and Freeman, 2010). Delayed degeneration occurs because the Wld^s protein maintains nicotinamide adenine dinucleotide (NAD⁺) levels in the severed axons. Normally axons contain NMNAT2 that fulfills the same enzymatic function to maintain NAD⁺, but NMNAT2 has short a half-life and must be continually replenished by anterograde transport (Gilley and Coleman, 2010). If transport becomes compromised, the neuron can be placed in the vulnerable position of losing its axon. Intriguingly, work in ALS animal models pointed to a shift in the retrogradely transported cargos from 'pro-survival' cargos to 'stress-related' cargos with increasing severity of neurodegeneration (Perlson *et al.*, 2009). The similar responses in physical injury of the axon and neurodegenerative diseases suggest the possibility that neurons in both cases may be mounting, or attempting to mount, a regenerative response, although few studies have addressed this possibility. Coordination of several signal transduction pathways and intracellular trafficking/signaling is needed to mount and maintain a regenerative response. Small alterations in these pathways could effectively alter function, growth, and maintenance of the axon.

Acknowledgments

DSS is supported by grant funds from National Institutes of Health (Award # R01- NS056314). JLT is supported by grants from National Institutes of Health (Award # P01-NS055976 and R01-NS041596), Department of Defense Orthopaedic Research Program (Award # W81XWH-13-1-0308), and the Dr. Miriam and Sheldon G. Adelson Medical Research Foundation, as well as funds from the South Carolina Smart-State Endowment Program through the University of South Carolina.

See also: Vertical Integration: Applications: Neurogenesis in the Adult Brain

References

- Adalbert, R., Coleman, M.P., 2012. Axon pathology in age-related neurodegenerative disorders. *Neuropathology and Applied Neurobiology* 39, 90–108.
- Albus, C.A., Rishal, I., Fainzilber, M., 2013. Cell length sensing for neuronal growth control. *Trends in Cell Biology* 23, 305–310.
- Allan, V.J., 2011. Cytoplasmic dynein. *Biochemical Society Transactions* 39, 1169–1178.
- Ally, S., Larson, A.G., Barlan, K., Rice, S.E., Gelfand, V.I., 2009. Opposite-polarity motors activate one another to trigger cargo transport in live cells. *Journal of Cell Biology* 187, 1071–1082.
- Ambron, R.T., Walters, E.T., 1996. Priming events and retrograde injury signals. A new perspective on the cellular and molecular biology of nerve regeneration. *Molecular Neurobiology* 13, 61–79.
- Baas, P.W., Lin, S., 2011. Hooks and comets: The story of microtubule polarity orientation in the neuron. *Developmental Neurobiology* 71, 403–418.
- Baleriola, J., Walker, C.A., Jean, Y.Y., *et al.*, 2014. Axonally synthesized ATF4 transmits a neurodegenerative signal across brain regions. *Cell* 158, 1159–1172.
- Ben-Yaakov, K., Dagan, S., Segal-Ruder, Y., *et al.*, 2012. Axonal transcription factors signal retrogradely in lesioned peripheral nerve. *EMBO Journal* 31, 1350–1363.
- Bradke, F., Fawcett, J.W., Spira, M.E., 2012. Assembly of a new growth cone after axotomy: The precursor to axon regeneration. *Nature Reviews Neuroscience* 13, 183–193.
- Bradshaw, N.J., Hennah, W., Soares, D.C., 2013. NDE1 and NDEL1: Twin neurodevelopmental proteins with similar 'nature' but different 'nurture'. *Biomolecular Concepts* 4, 447–464.
- Buki, A., Siman, R., Trojanowski, J.Q., Povlishock, J.T., 1999. The role of calpain-mediated spectrin proteolysis in traumatically induced axonal injury. *Journal of Neuropathology & Experimental Neurology* 58, 365–375.
- Campanot, R.B., MacLinnis, B.L., 2004. Retrograde transport of neurotrophins: Fact and function. *Journal of Neurobiology* 58, 217–229.
- Cavalli, V., Kujala, P., Klumperman, J., Goldstein, L.S., 2005. Sunday driver links axonal transport to damage signaling. *Journal of Cell Biology* 168, 775–787.
- Caviston, J.P., Ross, J.L., Antony, S.M., Tokito, M., Holzbaur, E.L., 2007. Huntingtin facilitates dynein/dynactin-mediated vesicle transport. *Proceedings of the National Academy of Sciences of the United States of America* 104, 10045–10050.
- Chang, L., Jones, Y., Ellisman, M.H., Goldstein, L.S., Karin, M., 2003. JNK1 is required for maintenance of neuronal microtubules and controls phosphorylation of microtubule-associated proteins. *Developmental Cell* 4, 521–533.
- Chen, Y., Tian, X., Kim, W.Y., Snider, W.D., 2011. Adenomatous polyposis coli regulates axon arborization and cytoskeleton organization via its N-terminus. *PLoS One* 6, e24335.
- Cho, Y., Sloutsky, R., Naegle, K.M., Cavalli, V., 2013. Injury-induced HDAC5 nuclear export is essential for axon regeneration. *Cell* 155, 894–908.
- Clevers, H., Nusse, R., 2012. Wnt/beta-catenin signaling and disease. *Cell* 149, 1192–1205.
- Coleman, M.P., Freeman, M.R., 2010. Wallerian degeneration, wld(s), and nmnat. *Annual Review of Neuroscience* 33, 245–267.
- Cosker, K.E., Pazyra-Murphy, M.F., Fenstermacher, S.J., Segal, R.A., 2013. Target-derived neurotrophins coordinate transcription and transport of bclw to prevent axonal degeneration. *Journal of Neuroscience* 33, 5195–5207.
- Cox, L.J., Hengst, U., Gurskaya, N.G., Lukyanov, K.A., Jaffrey, S.R., 2008. Intra-axonal translation and retrograde trafficking of CREB promotes neuronal survival. *Nature Cell Biology* 10, 149–159.
- Donnelly, C.J., Park, M., Spillane, M., *et al.*, 2013. Axonally synthesized beta-actin and GAP-43 proteins support distinct modes of axonal growth. *Journal of Neuroscience* 33, 3311–3322.
- Erez, H., Spira, M.E., 2008. Local self-assembly mechanisms underlie the differential transformation of the proximal and distal cut axonal ends into functional and aberrant growth cones. *Journal of Comparative Neurology* 507, 1019–1030.
- Eshed-Eisenbach, Y., Peles, E., 2013. Axonal spectrins: All-purpose fences. *Journal of Cell Biology* 203, 381–383.
- Feltrin, D., Fusco, L., Witte, H., *et al.*, 2012. Growth cone MKK7 mRNA targeting regulates MAP1b-dependent microtubule bundling to control neurite elongation. *PLoS Biology* 10, e1001439.
- Ferraris, D., Yang, Z., Welsbie, D., 2013. Dual leucine zipper kinase as a therapeutic target for neurodegenerative conditions. *Future Medicinal Chemistry* 5, 1923–1934.
- Fichera, M., Lo Giudice, M., Falco, M., *et al.*, 2004. Evidence of kinesin heavy chain (KIF5A) involvement in pure hereditary spastic paraplegia. *Neurology* 63, 1108–1110.
- Finelli, M.J., Wong, J.K., Zou, H., 2013. Epigenetic regulation of sensory axon regeneration after spinal cord injury. *Journal of Neuroscience* 33, 19664–19676.
- Fu, M.M., Holzbaur, E.L., 2013. JIP1 regulates the directionality of APP axonal transport by coordinating kinesin and dynein motors. *Journal of Cell Biology* 202, 495–508.
- Fu, M.M., Holzbaur, E.L., 2014. Integrated regulation of motor-driven organelle transport by scaffolding proteins. *Trends in Cell Biology* 24, 564–574.
- Fu, M.M., Nirschl, J.J., Holzbaur, E.L., 2014. LC3 binding to the scaffolding protein JIP1 regulates processive dynein-driven transport of autophagosomes. *Developmental Cell* 29, 577–590.

- Funakoshi, H., Frisen, J., Barbany, G., *et al.*, 1993. Differential expression of mRNAs for neurotrophins and their receptors after axotomy of the sciatic nerve. *Journal of Cell Biology* 123, 455–465.
- Geoffroy, C.G., Zheng, B., 2014. Myelin-associated inhibitors in axonal growth after CNS injury. *Current Opinion in Neurobiology* 27, 31–38.
- Gilley, J., Coleman, M.P., 2010. Endogenous Nmnat2 is an essential survival factor for maintenance of healthy axons. *PLoS Biology* 8, e1000300.
- Gitler, D., Spira, M., 2002. Short window of opportunity for calpain induced growth cone formation after axotomy in *Aplysia* neurons. *Journal of Neurobiology* 52, 267–279.
- Gobrecht, P., Leibinger, M., Andreadaki, A., Fischer, D., 2014. Sustained GSK3 activity markedly facilitates nerve regeneration. *Nature Communications* 5, 4561.
- Goldstein, M.E., Cooper, H.S., Bruce, J., *et al.*, 1987. Phosphorylation of neurofilament proteins and chromatolysis following transection of rat sciatic nerve. *Journal of Neuroscience* 7, 1586–1594.
- Gomes, C., Merianda, T.T., Lee, S.J., Yoo, S., Twiss, J.L., 2014. Molecular determinants of the axonal mRNA transcriptome. *Developmental Neurobiology* 74, 218–232.
- Gomez, T.M., Letourneau, P.C., 2014. Actin dynamics in growth cone motility and navigation. *Journal of Neurochemistry* 129, 221–234.
- Goodwin, S.S., Vale, R.D., 2010. Patronin regulates the microtubule network by protecting microtubule minus ends. *Cell* 143, 263–274.
- Grabham, P.W., Seale, G.E., Bennecib, M., Goldberg, D.J., Vallee, R.B., 2007. Cytoplasmic dynein and LIS1 are required for microtubule advance during growth cone remodeling and fast axonal outgrowth. *Journal of Neuroscience* 27, 5823–5834.
- Gracias, N.G., Shirkey-Son, N.J., Hengst, U., 2014. Local translation of TC10 is required for membrane expansion during axon outgrowth. *Nature Communications* 5, 3506.
- Hanz, S., Perlson, E., Willis, D., *et al.*, 2003. Axoplasmic importins enable retrograde injury signaling in lesioned nerve. *Neuron* 40, 1095–1104.
- Hellal, F., Hurtado, A., Ruschel, J., *et al.*, 2011. Microtubule stabilization reduces scarring and causes axon regeneration after spinal cord injury. *Science* 331, 928–931.
- Hirokawa, N., 1994. Microtubule organization and dynamics dependent on microtubule-associated proteins. *Current Opinion in Cell Biology* 6, 74–81.
- Hirokawa, N., Niwa, S., Tanaka, Y., 2010. Molecular motors in neurons: Transport mechanisms and roles in brain function, development, and disease. *Neuron* 68, 610–638.
- Hirokawa, N., Noda, Y., Tanaka, Y., Niwa, S., 2009. Kinesin superfamily motor proteins and intracellular transport. *Nature Reviews Molecular Cell Biology* 10, 682–696.
- Hoffman, P.N., 2010. A conditioning lesion induces changes in gene expression and axonal transport that enhance regeneration by increasing the intrinsic growth state of axons. *Experimental Neurology* 223, 11–18.
- Hoffman, P.N., Cleveland, D.W., 1988. Neurofilament and tubulin expression recapitulates the developmental program during axonal regeneration: Induction of a specific beta-tubulin isotype. *Proceedings of the National Academy of Sciences of the United States of America* 85, 4530–4533.
- Hsiao, K., Bozdagi, O., Benson, D.L., 2014. Axonal cap-dependent translation regulates presynaptic p35. *Developmental Neurobiology* 74, 351–364.
- Ikenaka, K., Katsuno, M., Kawai, K., *et al.*, 2012. Disruption of axonal transport in motor neuron diseases. *International Journal of Molecular Sciences* 13, 1225–1238.
- Janke, C., 2014. The tubulin code: Molecular components, readout mechanisms, and functions. *Journal of Cell Biology* 206, 461–472.
- Kam, N., Pilpel, Y., Fainzilber, M., 2009. Can molecular motors drive distance measurements in injured neurons? *PLoS Computational Biology* 5, e1000477.
- Kanaan, N.M., Pigino, G.F., Brady, S.T., *et al.*, 2013. Axonal degeneration in Alzheimer's disease: When signaling abnormalities meet the axonal transport system. *Experimental Neurology* 246, 44–53.
- Kardon, J.R., Vale, R.D., 2009. Regulators of the cytoplasmic dynein motor. *Nature Reviews Molecular Cell Biology* 10, 854–865.
- Kneussel, M., Wagner, W., 2013. Myosin motors at neuronal synapses: Drivers of membrane transport and actin dynamics. *Nature Reviews Neuroscience* 14, 233–247.
- Koyuncu, O.O., Perlman, D.H., Enquist, L.W., 2013. Efficient retrograde transport of pseudorabies virus within neurons requires local protein synthesis in axons. *Cell Host & Microbe* 13, 54–66.
- Lee, J.Y., Koga, H., Kawaguchi, Y., *et al.*, 2010. HDAC6 controls autophagosome maturation essential for ubiquitin-selective quality-control autophagy. *EMBO Journal* 29, 969–980.
- Lee, M.K., Cleveland, D.W., 1996. Neuronal intermediate filaments. *Annual Review of Neuroscience* 19, 187–217.
- Lee, W.C., Yoshihara, M., Littleton, J.T., 2004. Cytoplasmic aggregates trap polyglutamine-containing proteins and block axonal transport in a *Drosophila* model of Huntington's disease. *Proceedings of the National Academy of Sciences of the United States of America* 101, 3224–3229.
- Levi, V., Serpinskaya, A.S., Gratton, E., Gelfand, V., 2006. Organelle transport along microtubules in *Xenopus* melanophores: Evidence for cooperation between multiple motors. *Biophysical Journal* 90, 318–327.
- Li, C., Sasaki, Y., Takei, K., *et al.*, 2004. Correlation between semaphorin3A-induced facilitation of axonal transport and local activation of a translation initiation factor eukaryotic translation initiation factor 4E. *Journal of Neuroscience* 24, 6161–6167.
- Li, Y., Jung, P., Brown, A., 2012. Axonal transport of neurofilaments: A single population of intermittently moving polymers. *Journal of Neuroscience* 32, 746–758.
- Liu, M., Nadar, V.C., Kozielski, F., *et al.*, 2010. Kinesin-12, a mitotic microtubule-associated motor protein, impacts axonal growth, navigation, and branching. *Journal of Neuroscience* 30, 14896–14906.
- Loverde, J.R., Ozoka, V.C., Aquino, R., Lin, L., Pfister, B.J., 2011. Live imaging of axon stretch growth in embryonic and adult neurons. *Journal of Neurotrauma* 28, 2389–2403.
- Lowery, L.A., Van Vactor, D., 2009. The trip of the tip: Understanding the growth cone machinery. *Nature Reviews Molecular Cell Biology* 10, 332–343.
- Machnicka, B., Czogalla, A., Hryniewicz-Jankowska, A., *et al.*, 2014. Spectrins: A structural platform for stabilization and activation of membrane channels, receptors and transporters. *Biochimica et Biophysica Acta* 1838, 620–634.
- Mar, F.M., Bonni, A., Sousa, M.M., 2014. Cell intrinsic control of axon regeneration. *EMBO Reports* 15, 254–263.
- Marx, A., Muller, J., Mandelkow, E.M., Hoenger, A., Mandelkow, E., 2006. Interaction of kinesin motors, microtubules, and MAPs. *Journal of Muscle Research and Cell Motility* 27, 125–137.
- Mazel, T., Biesemann, A., Krejczyk, M., *et al.*, 2014. Direct observation of microtubule pushing by cortical dynein in living cells. *Molecular Biology of the Cell* 25, 95–106.
- McGuire, J.R., Rong, J., Li, S.H., Li, X.J., 2006. Interaction of Huntingtin-associated protein-1 with kinesin light chain: Implications in intracellular trafficking in neurons. *Journal of Biological Chemistry* 281, 3552–3559.
- McKenney, R.J., Huynh, W., Tanenbaum, M.E., Bhabha, G., Vale, R.D., 2014. Activation of cytoplasmic dynein motility by dynactin-cargo adapter complexes. *Science* 345, 337–341.
- McKenney, R.J., Verzhinin, M., Kunwar, A., Vallee, R.B., Gross, S.P., 2010. LIS1 and NudE induce a persistent dynein force-producing state. *Cell* 141, 304–314.
- Medeiros, N.A., Burnette, D.T., Forscher, P., 2006. Myosin II functions in actin-bundle turnover in neuronal growth cones. *Nature Cell Biology* 8, 215–226.
- Mili, S., Moissoglu, K., MaCar, I.G., 2008. Genome-wide screen reveals APC-associated RNAs enriched in cell protrusions. *Nature* 453, 115–119.
- Mitchell, D.J., Blasier, K.R., Jeffery, E.D., *et al.*, 2012. Trk activation of the ERK1/2 kinase pathway stimulates intermediate chain phosphorylation and recruits cytoplasmic dynein to signaling endosomes for retrograde axonal transport. *Journal of Neuroscience* 32, 15495–15510.
- Mitchison, T., Kirschner, M., 1984. Dynamic instability of microtubule growth. *Nature* 312, 237–242.
- Mitchison, T., Kirschner, M., 1988. Cytoskeletal dynamics and nerve growth. *Neuron* 1, 761–772.
- Morfini, G., Szebenyi, G., Elluru, R., Ratner, N., Brady, S.T., 2002. Glycogen synthase kinase 3 phosphorylates kinesin light chains and negatively regulates kinesin-based motility. *EMBO Journal* 21, 281–293.
- Morfini, G.A., You, Y.M., Pollema, S.L., *et al.*, 2009. Pathogenic huntingtin inhibits fast axonal transport by activating JNK3 and phosphorylating kinesin. *Nature Neuroscience* 12, 864–871.
- Myers, K.A., Baas, P.W., 2007. Kinesin-5 regulates the growth of the axon by acting as a brake on its microtubule array. *Journal of Cell Biology* 178, 1081–1091.
- Myers, K.A., He, Y., Hasaka, T.P., Baas, P.W., 2006. Microtubule transport in the axon: Re-thinking a potential role for the actin cytoskeleton. *Neuroscientist* 12, 107–118.
- Niethammer, M., Smith, D.S., Ayala, R., *et al.*, 2000. NUDEL is a novel Cdk5 substrate that associates with LIS1 and cytoplasmic dynein. *Neuron* 28, 697–711.
- Pandey, J.P., Smith, D.S., 2011. A Cdk5-dependent switch regulates Lis1/Ndel1/dynein-driven organelle transport in adult axons. *Journal of Neuroscience* 31, 17207–17219.
- Perlson, E., Hanz, S., Ben-Yaakov, K., *et al.*, 2005. Vimentin-dependent spatial translocation of an activated MAP kinase in injured nerve. *Neuron* 45, 715–726.

- Perlson, E., Jeong, G.B., Ross, J.L., *et al.*, 2009. A switch in retrograde signaling from survival to stress in rapid-onset neurodegeneration. *Journal of Neuroscience* 29, 9903–9917.
- Perrot, R., Eyer, J., 2009. Neuronal intermediate filaments and neurodegenerative disorders. *Brain Research Bulletin* 80, 282–295.
- Perry, R.B., Fainzilber, M., 2014. Local translation in neuronal processes—*in vivo* tests of a “heretical hypothesis”. *Developmental Neurobiology* 74, 210–217.
- Pfister, B.J., Bonislawski, D.P., Smith, D.H., Cohen, A.S., 2006. Stretch-grown axons retain the ability to transmit active electrical signals. *FEBS Letters* 580, 3525–3531.
- Pfister, K.K., Salata, M.W., Dillman III, J.F., *et al.*, 1996. Differential expression and phosphorylation of the 74-kDa intermediate chains of cytoplasmic dynein in cultured neurons and glia. *Journal of Biological Chemistry* 271, 1687–1694.
- Preitner, N., Quan, J., Nowakowski, D.W., *et al.*, 2014. APC is an RNA-binding protein, and its interactome provides a link to neural development and microtubule assembly. *Cell* 158, 368–382.
- Prokop, A., 2013. The intricate relationship between microtubules and their associated motor proteins during axon growth and maintenance. *Neural Development* 8, 17.
- Puls, I., Jonnakuty, C., Lamonte, B.H., *et al.*, 2003. Mutant dynactin in motor neuron disease. *Nature Genetics* 33, 455–456.
- Reid, A.J., Welin, D., Wiberg, M., Terenghi, G., Novikov, L.N., 2010. Peripherin and ATF3 genes are differentially regulated in regenerating and non-regenerating primary sensory neurons. *Brain Research* 1310, 1–7.
- Reiner, O., Sapir, T., 2013. LIS1 functions in normal development and disease. *Current Opinion in Neurobiology* 23, 951–956.
- Riederer, B.M., Pellier, V., Antonsson, B., *et al.*, 1997. Regulation of microtubule dynamics by the neuronal growth-associated protein SCG10. *Proceedings of the National Academy of Sciences of the United States of America* 94, 741–745.
- Rishal, I., Fainzilber, M., 2010. Retrograde signaling in axonal regeneration. *Experimental Neurology* 223, 5–10.
- Rishal, I., Kam, N., Perry, R.B., *et al.*, 2012. A motor driven mechanism for cell length sensing. *Cell Reports* 1, 608–616.
- Rivieccio, M.A., Brochier, C., Willis, D.E., *et al.*, 2009. HDAC6 is a target for protection and regeneration following injury in the nervous system. *Proceedings of the National Academy of Sciences of the United States of America* 106, 19599–19604.
- Ruff, C.A., Staak, N., Patodia, S., *et al.*, 2012. Neuronal c-Jun is required for successful axonal regeneration, but the effects of phosphorylation of its N-terminus are moderate. *Journal of Neurochemistry* 121, 607–618.
- Sahly, I., Khoutorsky, A., Erez, H., Prager-Khoutorsky, M., Spira, M.E., 2006. On-line confocal imaging of the events leading to structural dedifferentiation of an axonal segment into a growth cone after axotomy. *Journal of Comparative Neurology* 494, 705–720.
- Schlager, M.A., Serra-Marques, A., Grigoriev, I., *et al.*, 2014. Bicaudal d family adaptor proteins control the velocity of Dynein-based movements. *Cell Reports* 8, 1248–1256.
- Schmieg, N., Menendez, G., Schiavo, G., Terenzio, M., 2014. Signalling endosomes in axonal transport: Travel updates on the molecular highway. *Seminars in Cell and Developmental Biology* 27, 32–43.
- Schuyler, S.C., Pellman, D., 2001. Microtubule “plus-end-tracking proteins”: The end is just the beginning. *Cell* 105, 421–424.
- Shah, J.V., Cleveland, D.W., 2002. Slow axonal transport: Fast motors in the slow lane. *Current Opinion in Cell Biology* 14, 58–62.
- Shea, T.B., Lee, S., 2013. The discontinuous nature of neurofilament transport accommodates both establishment and repair of the axonal neurofilament array. *Cytoskeleton (Hoboken)* 70, 67–73.
- Shin, J.E., Cho, Y., Beirowski, B., *et al.*, 2012a. Dual leucine zipper kinase is required for retrograde injury signaling and axonal regeneration. *Neuron* 74, 1015–1022.
- Shin, J.E., Geisler, S., Diantonio, A., 2014. Dynamic regulation of SCG10 in regenerating axons after injury. *Experimental Neurology* 252, 1–11.
- Shin, J.E., Miller, B.R., Babetto, E., *et al.*, 2012b. SCG10 is a JNK target in the axonal degeneration pathway. *Proceedings of the National Academy of Sciences of the United States of America* 109, E3696–E3705.
- Smith, D.S., 2003. Cdk5 in neuroskeletal dynamics. *Neurosignals* 12, 239–251.
- Smith, D.S., Niethammer, M., Ayala, R., *et al.*, 2000. Regulation of cytoplasmic dynein behaviour and microtubule organization by mammalian Lis1. *Nature Cell Biology* 2, 767–775.
- Smith, D.S., Skene, J.H., 1997. A transcription-dependent switch controls competence of adult neurons for distinct modes of axon growth. *Journal of Neuroscience* 17, 646–658.
- Spillane, M., Ketschek, A., Merianda, T.T., Twiss, J.L., Gallo, G., 2013. Mitochondria coordinate sites of axon branching through localized intra-axonal protein synthesis. *Cell Reports* 5, 1564–1575.
- Sun, F., Zhu, C., Dixit, R., Cavalli, V., 2011. Sunday Driver/JIP3 binds kinesin heavy chain directly and enhances its motility. *EMBO Journal* 30, 3416–3429.
- Suter, D.M., Miller, K.E., 2011. The emerging role of forces in axonal elongation. *Progress in Neurobiology* 94, 91–101.
- Swarup, V., Audet, J.N., Phaneuf, D., Kriz, J., Julien, J.P., 2012. Abnormal regenerative responses and impaired axonal outgrowth after nerve crush in TDP-43 transgenic mouse models of amyotrophic lateral sclerosis. *Journal of Neuroscience* 32, 18186–18195.
- Szaro, B.G., Strong, M.J., 2010. Post-transcriptional control of neurofilaments: New roles in development, regeneration and neurodegenerative disease. *Trends in Neurosciences* 33, 27–37.
- Tanaka, E., Ho, T., Kirschner, M., 1995. The role of microtubule dynamics in growth cone motility and axonal growth. *Journal of Cell Biology* 128, 139–155.
- Tcherkezian, J., Brittis, P.A., Thomas, F., Roux, P.P., Flanagan, J.G., 2010. Transmembrane receptor DCC associates with protein synthesis machinery and regulates translation. *Cell* 141, 632–644.
- Tedeschi, A., Bradke, F., 2013. The DLK signalling pathway — A double-edged sword in neural development and regeneration. *EMBO Reports* 14, 605–614.
- Tojima, T., Itofusa, R., Kamiguchi, H., 2014. Steering neuronal growth cones by shifting the imbalance between exocytosis and endocytosis. *Journal of Neuroscience* 34, 7165–7178.
- Vale, R.D., 2003. The molecular motor toolbox for intracellular transport. *Cell* 112, 467–480.
- Verma, P., Chierzi, S., Codd, A.M., *et al.*, 2005. Axonal protein synthesis and degradation are necessary for efficient growth cone regeneration. *Journal of Neuroscience* 25, 331–342.
- Vogelstein, B., Kinzler, K.W., 2004. Cancer genes and the pathways they control. *Nature Medicine* 10, 789–799.
- Vuppalandhi, D., Merianda, T.T., Donnelly, C., *et al.*, 2012. Lysophosphatidic acid differentially regulates axonal mRNA translation through 5'UTR elements. *Molecular and Cellular Neuroscience* 50, 136–146.
- Weaver, C., Leidel, C., Szpankowski, L., *et al.*, 2013. Endogenous GSK-3/shaggy regulates bidirectional axonal transport of the amyloid precursor protein. *Traffic* 14, 295–308.
- Willis, D.E., Van Niekerk, E.A., Sasaki, Y., *et al.*, 2007. Extracellular stimuli specifically regulate localized levels of individual neuronal mRNAs. *Journal of Cell Biology* 178, 965–980.
- Willis, D.E., Xu, M., Donnelly, C.J., *et al.*, 2011. Axonal Localization of transgene mRNA in mature PNS and CNS neurons. *Journal of Neuroscience* 31, 14481–14487.
- Yamashita, N., Usui, H., Nakamura, F., *et al.*, 2014. Plexin-A4-dependent retrograde semaphorin 3A signalling regulates the dendritic localization of GluA2-containing AMPA receptors. *Nature Communications* 5, 3424.
- Yudin, D., Hanz, S., Yoo, S., *et al.*, 2008. Localized regulation of axonal RanGTPase controls retrograde injury signaling in peripheral nerve. *Neuron* 59, 241–252.
- Zala, D., Hinckelmann, M.V., Yu, H., *et al.*, 2013. Vesicular glycolysis provides on-board energy for fast axonal transport. *Cell* 152, 479–491.
- Zhang, H.L., Singer, R.H., Bassell, G.J., 1999. Neurotrophin regulation of beta-actin mRNA and protein localization within growth cones. *Journal of Cell Biology* 147, 59–70.
- Ziv, N., Spira, M., 1997. Localized and transient elevations of intracellular Ca^{2+} induce dedifferentiation of axonal segments into growth cones. *Journal of Neuroscience* 17, 3568–3579.

Further Reading

- Franker, M.A., Hoogenraad, C.C., 2013. Microtubule-based transport — Basic mechanisms, traffic rules and role in neurological pathogenesis. *Journal of Cell Science* 126, 2319–2329.
- Freeman, M.R., 2014. Signaling mechanisms regulating Wallerian degeneration. *Current Opinion in Neurobiology* 27, 224–231.
- Holt, C.E., Schuman, E.M., 2013. The central dogma decentralized: New perspectives on RNA function and local translation in neurons. *Neuron* 80, 648–657.
- Jung, H., Gkogkas, C.G., Sonenberg, N., Holt, C.E., 2014. Remote control of gene function by local translation. *Cell* 157, 26–40.
- Michaevlevski, I., Medzihradsky, K.F., Lynn, A., Burlingame, A.L., Fainzilber, M., 2010a. Axonal transport proteomics reveals mobilization of translation machinery

- to the lesion site in injured sciatic nerve. *Molecular & Cellular Proteomics* 9, 976–987.
- Michaelievski, I., Segal-Ruder, Y., Rozenbaum, M., *et al.*, 2010b. Signaling to transcription networks in the neuronal retrograde injury response. *Science Signaling* 3, ra53.
- Nix, P., Bastiani, M., 2012. DLK: The “preconditioning” signal for axon regeneration? *Neuron* 74, 961–963.
- Park, K.K., Liu, K., Hu, Y., Kanter, J.L., He, Z., 2010. PTEN/mTOR and axon regeneration. *Experimental Neurology* 223, 45–50.
- Sengottuvel, V., Leibinger, M., Preimer, M., Andreadaki, A., Fischer, D., 2011. Taxol facilitates axon regeneration in the mature CNS. *Journal of Neuroscience* 31, 2688–2699.
- Stone, M.C., Albertson, R.M., Chen, L., Rollis, M.M., 2014. Dendrite injury triggers DLK-independent regeneration. *Cell Reports* 6, 247–253.

CELL DIVISION/DEATH: REGULATION OF CELL GROWTH

Contents

mTORC1: Upstream and Downstream

The Nucleolus

Ribosomes and Stress – Linked from Birth to Death

Extra-Ribosome Functions of Ribosomal Proteins

Ribosomal RNA Processing

Eukaryotic Ribosome Assembly and Export

Internal Ribosome Entry Site-Mediated Translation

Targeted mRNA Degradation

mTORC1: Upstream and Downstream

Y Yao, N Parnell, and K Inoki, University of Michigan, Ann Arbor, MI, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

mTOR belongs to the phosphoinositide-3-kinase (PI3K)-related family of protein kinases (Lempiäinen and Halazonetis, 2009) and primarily functions to promote cell growth and metabolism (Wullschlegel *et al.*, 2006; Laplante and Sabatini, 2012). At least two distinct multi-protein complexes, termed mTOR complex 1 (mTORC1) and mTOR complex 2 (mTORC2) (Hara *et al.*, 2002; Kim *et al.*, 2002; Loewith *et al.*, 2002; Jacinto *et al.*, 2004; Sarbassov *et al.*, 2004), are defined by their unique cofactors (Figure 1), including Raptor and PRAS40 for mTORC1 and Rictor, Sin1, and Protor for mTORC2 (Jacinto *et al.*, 2006; Yang *et al.*, 2006; Sancak *et al.*, 2007; Vander Haar *et al.*, 2007; Pearce *et al.*, 2007), while core factors common to both complexes include mLST8 (GbetaL) and Deptor (Kim *et al.*, 2003; Peterson *et al.*, 2009). It is noteworthy that novel mTORC1 and mTORC2 components and their substrates continue to emerge, expanding our knowledge of each mTOR complex.

mTORC1 and mTORC2

Raptor and PRAS40 are mTORC1 Subunits

Raptor, an essential component of mTORC1, functions as a scaffold for mTOR kinase to recruit and phosphorylate specific substrates such as ribosomal protein S6 kinases (S6Ks) and the eukaryotic translation initiation factor 4E binding proteins (e4EBPs). This specific substrate recruitment by mTORC1 occurs through a direct interaction between Raptor and the TOR signaling (TOS) motif, a conserved five amino acid sequence often found in these substrates (Nojima *et al.*, 2003; Schalm and Blenis, 2002; Schalm *et al.*, 2003). Importantly, recent studies have revealed another important function of Raptor in

guiding the subcellular localization of mTORC1. In response to amino acid availability, Raptor associates with Rag small GTPases that localize to the surface of lysosomes. This action is important for the recruitment of mTORC1 to the lysosome membrane for its activation (see Section Regulation of mTOR Signaling by Amino Acids for the details) (Sancak *et al.*, 2008, 2010). Thus, Raptor serves to link not only mTOR kinase with its substrates, but also mTOR to the lysosome membrane for its activation. In contrast to its positive roles in the regulation of mTORC1 signaling, Raptor also interacts with PRAS40, a specific negative regulator for mTORC1 (Sancak *et al.*, 2007; Vander Haar *et al.*, 2007). It has been suggested that the interaction of PRAS40 with Raptor through a TOS-like motif suppresses mTORC1 phosphorylation of other substrates by directly competing for mTORC1 (Wang *et al.*, 2007; Oshiro *et al.*, 2007). Importantly, active Akt kinase stimulates mTORC1-dependent S6K1 and 4EBP1 phosphorylation by directly phosphorylating Thr246 of PRAS40, which induces the dissociation of PRAS40 from mTORC1 and mitigates the inhibitory effect of PRAS40 on mTORC1-dependent phosphorylation of its substrates (Sancak *et al.*, 2007; Vander Haar *et al.*, 2007).

mLST8 and Deptor are Common Components of mTORC1 and mTORC2

mLST8 is a common component of both mTORC1 and mTORC2. Knockdown of mLST8 expression in HEK293T cells using siRNAs decreases S6K1 phosphorylation (Kim *et al.*, 2003) suggesting mLST8 to be required for mTORC1 activity. However, a subsequent study in *mLST8*^{-/-} mouse embryonic fibroblasts (MEFs) indicated that mLST8 loss-of-function does not affect mTORC1 activity, mTORC1 integrity, or the

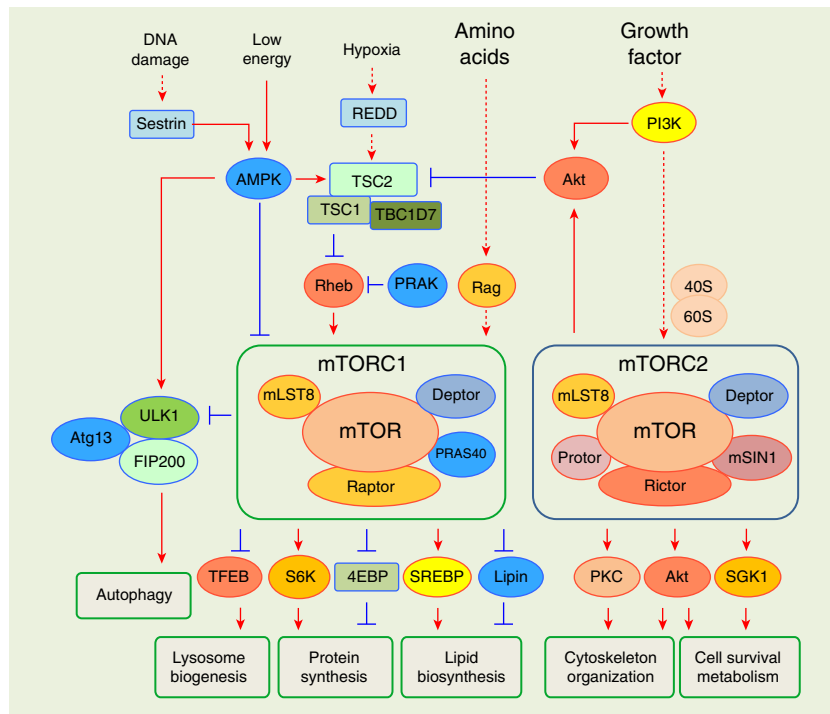


Figure 1 The mTOR signaling pathway. The key signaling nodes that regulate mTORC1 and mTORC2 and the representative outputs of the mTORC1 and mTORC2 pathways.

interaction between mTORC1 and its substrates (Guertin *et al.*, 2006). Instead, the activity of mTORC2 was completely abolished in *mLST8*^{-/-} MEFs, suggesting an essential role for mLST8 in mTORC2 function. Thus, the functional importance of mTORC1 in the regulation of mTORC2 remains to be fully resolved.

Deptor is another common component of mTORC1 and mTORC2. Initial findings indicated Deptor exerts inhibitory effects on both mTORC1 and mTORC2 (Peterson *et al.*, 2009). In a subset of multiple myeloma cells, enhanced Deptor expression leads to strong inhibition of mTORC1, which relieves mTORC1-dependent inhibition of the PI3K pathway, and results in the activation of Akt and cancer cell survival (Peterson *et al.*, 2009). However, recent studies demonstrate that phosphorylation of both Akt and S6K1 are enhanced in BAF60c transgenic mice, where Deptor is highly expressed (Meng *et al.*, 2013). Moreover, Deptor does not significantly affect mTORC1 activity in adipocytes as neither overexpression nor knockdown of DEPTOR affected phosphorylation of S6K1 (Laplante *et al.*, 2012). Taken together, these observations suggest the differential effects of Deptor on mTORC1 activity are likely to be cell or tissue type specific. Therefore, the elucidation of the molecular mechanisms underlying Deptor-dependent mTORC1 regulation requires further investigation. Intriguingly, recent studies have shown that the stability of Deptor expression is negatively regulated by mTORC1. Deptor is phosphorylated by mTORC1 (mTOR) and/or S6K, resulting in enhanced proteasome-mediated degradation of Deptor. These new observations suggest that mTORC1 exerts a positive feedback loop on its own activity by releasing the inhibitory effects of Deptor (Gao *et al.*, 2011; Zhao *et al.*, 2011).

A Brief Introduction to mTORC2

mTORC2 regulates cell survival, metabolism and cytoskeleton organization through the phosphorylation of AGC kinase family members, such as Akt, PKC, and SGK1 (Sarbasov *et al.*, 2006; Jacinto *et al.*, 2004; Ikenoue *et al.*, 2008). Relative to mTORC1, the properties of mTORC2 are less well characterized, likely due to the lack of specific mTORC2 inhibitors. mTORC2 responds to growth factors, but not to nutrients or cellular energy. The mechanism by which growth factors activate mTORC2 is not fully characterized. A recent study demonstrated that, in response to growth factors, the association of mTORC2 with ribosomes appears to be a key event for the activation of mTORC2 (Zinzalla *et al.*, 2011). Rictor is an essential and characteristic component of mTORC2. Sin1 and mLST8 are also essential for mTORC2 function, as depletion of either abolishes mTORC2 activity (Yang *et al.*, 2006; Jacinto *et al.*, 2006; Guertin *et al.*, 2006).

The Effect of Rapamycin on mTORC1 and mTORC2

mTORC1 activity is sensitive to rapamycin, a macrolide originally developed as an antifungal agent. Rapamycin strongly interacts with FK506-binding protein 12 (FKBP12), and this drug-protein complex binds to the FKBP12-rapamycin-binding (FRB) domain of mTOR kinase (Choi *et al.*, 1996). The FRB domain acts as a gatekeeper as its rapamycin binding site interacts with substrates to grant them access to the restricted active site of mTOR kinase. The rapamycin-FKBP12 therefore allosterically inhibits mTOR kinase by blocking substrate

recruitment and further restricts the accessibility of substrates to the active site of mTOR kinase (Yang *et al.*, 2013).

Unlike the effect of ATP-competitive mTOR kinase inhibitors, which completely inhibit mTORC1-dependent phosphorylation of its substrates, effects of rapamycin on the phosphorylation of mTORC1 substrates vary (Thoreen *et al.*, 2009; Feldman *et al.*, 2009). For example, the mTORC1 site on S6K1 is highly sensitive to rapamycin while sites on 4EBP1 are less sensitive. Interestingly, the dimerization of mTORC1 is required for 4EBP1 phosphorylation but not for S6K1 phosphorylation (Yip *et al.*, 2010). In addition, prolonged rapamycin treatment weakens the interaction between mTORC1 and Raptor, and disrupts the dimer formation of mTORC1. This study suggests that the contact of the rapamycin-FKBP12 complex with the FRB domain of mTORC1 may be sufficient to block access to the active site for larger sized substrates such as S6K1. In contrast, disruption of the mTORC1 dimer is required to block mTORC1-dependent phosphorylation of 4EBP1.

Although the rapamycin-FKBP12 complex directly binds and limits the accessibility of substrates to the active kinase site, it has little effect on mTOR kinase activity of mTORC2. However, prolonged rapamycin treatment inhibits mTORC2 from phosphorylating its substrates in some cell types by attenuating the formation of the mTORC2 complex (Sarbassov *et al.*, 2006). These observations suggest that specific components of mTORC2 may prevent rapamycin-FKBP12 from gaining access to the FRB domain of mTOR kinase. However, prolonged rapamycin treatment may allow the rapamycin-FKBP12 complex to access newly synthesized mTOR kinase and prevents the formation of mTORC2.

Downstream of mTORC1

Consistent with the functional roles of mTORC1 in promoting cell growth and proliferation, the effectors characterized as downstream targets of mTORC1 are largely comprised of factors that regulate protein synthesis, cell metabolism, cellular organelle synthesis, and protein degradation (Wullschlegel *et al.*, 2006; Figure 1). At the cellular level, mTORC1 functions as a critical regulator of translation initiation, the rate-limiting step of protein synthesis in which ribosomes are recruited to mRNAs (Mamane *et al.*, 2006).

S6K1 and 4EBP1 Are the Best-Characterized mTORC1 Substrates

S6K1 and 4EBP1 represent the first TOR substrates identified in metazoans and remain the best characterized. S6K1 regulates translation initiation, mRNA processing, metabolism, and cell size. Genetic evidence indicates that the ablation of dS6K in *Drosophila* reduces cell and body size (Montagne *et al.*, 1999). Mice lacking functional S6K1 also show a small body phenotype, resistance to high fat diet-induced obesity, and improved insulin sensitivity (Um *et al.*, 2004). At the molecular level, mTORC1 phosphorylates S6K1 on Thr389 and Ser371 (Burnett *et al.*, 1998; Weng *et al.*, 1998; Moser *et al.*, 1997). mTORC1 phosphorylation on Thr389 generates the binding motif on S6K1 for 3-phosphoinositide dependent

protein kinase 1 (PDK1), which further phosphorylates S6K1 on Thr229 that lies in the S6K1 activation loop (Alessi *et al.*, 1998; Pullen *et al.*, 1998). These sequential phosphorylations lead to the full activation of S6K1, thereby initiating the downstream cellular events. S6K1 interacts with a component of the eukaryotic translation initiation factor 3 (eIF3) complex when mTORC1 activity is inhibited. Upon mTORC1 activation, phosphorylated S6K1 dissociates from the eIF3 complex and begins phosphorylating its multiple substrates required for translation initiation as well as other steps in protein production (Holz *et al.*, 2005). For example, S6K1 phosphorylates eIF4B on Ser422 and stimulates its function to enhance eIF4A, an essential mRNA helicase that unwinds the secondary structure in the 5'-untranslated region (5'UTR) of mRNA for efficient translation initiation (Raught *et al.*, 2004). S6K1 also phosphorylates Ser67 of an inhibitor of eIF4A, programmed cell death 4 (PDCD4), thereby further enhancing eIF4A helicase activity (Dorrello *et al.*, 2006). In addition, S6K1 stimulates eukaryotic translation elongation factor 2 (eEF2), which plays an important role in translation elongation, by inhibiting its upstream negative regulator, eEF2 kinase through direct phosphorylation (Ser366) (Wang *et al.*, 2001). S6K1 phosphorylates ribosomal protein S6 (rpS6) on multiple sites (Ser235, Ser236, Ser240, and Ser244). However, the functional significance of rpS6 phosphorylation remains unclear (Ruvinsky *et al.*, 2005). S6K1 also phosphorylates and activates Skar, an important protein that catalyzes splicing of mRNAs (Richardson *et al.*, 2004). In addition to its role in the regulation of translation, S6K1 also phosphorylates Insulin Receptor Substrate 1 (IRS1), leading to destabilization and degradation of this substrate, thereby providing negative feedback for the PI3K-Akt pathway (Harrington *et al.*, 2004; Um *et al.*, 2004; Shah *et al.*, 2004).

e4EBPs such as 4EBP1 are also well-characterized mTORC1 substrates (Gingras *et al.*, 2001). In the absence of mTORC1 activation, hypo-phosphorylated 4EBP1 binds to and represses eIF4E, which directly interacts with the 5'-cap of mRNAs (Pause *et al.*, 1994; Richter and Sonenberg, 2005). Upon mTORC1 activation, mTORC1 binds to eIF3 and phosphorylates 4EBP1 on several sites (e.g., Thr37, Thr46, Thr70, and Ser65) (Gingras *et al.*, 1999). mTORC1-mediated 4EBP1 phosphorylation induces the dissociation of 4EBP1 from eIF4E and allows eIF4E to interact with eIF4G1/3, key adaptor proteins that interact with other initiation factors to form the active eIF4F complex (Gross *et al.*, 2003; Richter and Sonenberg, 2005).

Grb10 Is a Newly Identified mTORC1 Substrate

Although the mTORC1-mediated phosphorylation of S6K1 and 4EBP1 has been widely regarded as the gold standards for measuring cellular mTORC1 activity, mTORC1 regulates a wide array of cellular processes in addition to its role in translation. Recent large-scale phospho-proteomics studies provided a potent strategy for the identification of novel mTOR substrates. By comparing the phospho-protein profiles in the presence or absence of mTOR inhibitors, two groups independently identified Grb10 as a novel substrate for mTORC1 (Hsu *et al.*, 2011; Yu *et al.*, 2011). Grb10 binds to

and negatively regulates insulin receptors and IGF receptors (Wick *et al.*, 2003). mTORC1 phosphorylates Ser501 and Ser503 on Grb10 to promote protein stability. By phosphorylating and stabilizing Grb10, mTORC1 negatively regulates the insulin receptor-PI3K-Akt pathway, which provides another mechanism for negative feedback regulation in addition to S6K1-mediated IRS inhibition.

Sterol-Regulatory-Element-Binding Proteins Bridge mTORC1 to Lipid Synthesis

In agreement with its role as a master regulator of anabolism, mTORC1 also plays a fundamental role in lipid synthesis (Lamming and Sabatini, 2013) by regulating the cellular levels of many lipogenic genes. Sterol-regulatory-element-binding proteins (SREBPs) are transcription factors responsible for the expression of enzymes important for fatty acid and sterol synthesis (Horton *et al.*, 2002). mTORC1 stimulates the function of SREBPs at multiple steps by: (1) enhancing the mRNA and protein expression of SREBPs (Wang *et al.*, 2011; Yecies *et al.*, 2011); (2) promoting the processing of the SREBP precursor to the mature form and enhancing its nuclear localization (Yecies *et al.*, 2011; Duvel *et al.*, 2010; Owen *et al.*, 2012); and (3) phosphorylating and inhibiting Lipin1, a negative regulator of SREBPs (Peterson *et al.*, 2011). The identification of mTORC1 as a regulator of SREBPs explains how lipid synthesis is intimately coupled to the levels of cellular nutrient availability and growth factor signaling.

Transcription Factor EB Mediates the Effect of mTORC1 on Lysosome Biogenesis

mTORC1 controls lysosome biogenesis through its regulation of the bHLH leucine zipper transcription factor EB (TFEB). TFEB is a major transcription factor that stimulates transcription of genes encoding proteins related to lysosome biogenesis and autophagy (Sardiello *et al.*, 2009). Recent studies identified mTORC1 as an upstream regulator of TFEB (Roczniaik-Ferguson *et al.*, 2012; Settembre *et al.*, 2012; Martina *et al.*, 2012; Pena-Llopis *et al.*, 2011). In the nutrient-rich conditions where mTORC1 is active, TFEB is directly phosphorylated by mTORC1 at the lysosome membrane. Phosphorylated TFEB is sequestered in the cytoplasm through its interaction with 14-3-3 proteins, thus inhibiting the expression of genes required for lysosome and autophagy biogenesis. Conversely, in mTORC1 inactive conditions, non-phosphorylated TFEB translocates into the nucleus, where it stimulates lysosomal and autophagic gene expression. Intriguingly, however, transcript levels of TFEB and its target genes are enhanced in MEFs lacking functional TSC2, a key suppressor of mTORC1 (Pena-Llopis *et al.*, 2011). These observations suggest that the effect of chronic and acute mTORC1 activation on TFEB activity may not be the same.

mTORC1 Suppresses Autophagy

mTORC1 is a major suppressor of autophagy induction (Jung *et al.*, 2010; Rubinsztein *et al.*, 2007). Autophagy is an

evolutionarily conserved catabolic process in which unnecessary, dysfunctional, or toxic cellular material are captured and delivered by double-membrane autophagic vesicles (autophagosomes) to the lysosome for degradation (Rabinowitz and White, 2010; Mizushima *et al.*, 2008; Levine and Kroemer, 2008). Inhibition of mTORC1 by rapamycin stimulates the induction of autophagy even in nutrient-rich conditions where autophagy is typically repressed (Noda and Ohsumi, 1998). mTORC1 inhibits autophagy by phosphorylating several proteins required for autophagy induction. The ULK1-Atg13-FIP200 complex is required at the initial step of autophagosome formation. Under nutrient-rich conditions, mTORC1 interacts with the ULK1-Atg13-FIP200 complex and phosphorylates Atg13 and ULK1 kinases to inhibit their kinase activity, thereby mitigating autophagy induction. Under starvation or rapamycin treatment conditions, inactive mTORC1 dissociates from this complex. This leads to dephosphorylation and activation of ULK1, which promotes ULK1-dependent phosphorylation of Atg13 and FIP200, processes essential for autophagy induction (Ganley *et al.*, 2009; Hosokawa *et al.*, 2009; Jung *et al.*, 2009). A recent study indicated that AMPK, a key cellular energy sensor and activator of autophagy (Egan *et al.*, 2011), directly antagonizes mTORC1-mediated inactivation of ULK1 in the ULK1-Atg13-FIP200 complex to stimulate autophagy (Kim *et al.*, 2011). During energy shortage, active AMPK promotes autophagy by phosphorylating Ser317 and Ser777 of ULK1. In contrast, under nutrient-rich conditions, active mTORC1 phosphorylates ULK1 on Ser757, which causes the dissociation of AMPK from ULK1 (Kim *et al.*, 2011). In this regard, the ULK1-Atg13-FIP200 complex functions as a hub to transmit upstream signals that reflect cellular nutrient availability to downstream regulation of autophagy induction.

Upstream of mTORC1

Regulation of mTOR Signaling by Growth Factors

Growth factor-induced mTORC1 activation

Growth factors, such as insulin and insulin-like growth factor 1, exert strong anabolic effects and play crucial roles in promoting cell growth, proliferation, survival, and metabolism. Growth factors bind and activate the receptor tyrosine kinases, such as insulin receptor and IGF receptor, thereby engaging and activating the PI3 kinase and Akt (Alessi *et al.*, 1997). Akt mediates upstream of mTORC1 and leads to mTORC1 activation via two approaches. The first approach goes through the TSC-Rheb axis (Manning and Cantley, 2003). The tuberous sclerosis complex (TSC) is composed of TSC1, TSC2, and TBC1D7 (van Slegtenhorst *et al.*, 1998; Dibble *et al.*, 2012). TSC2 functions as a GTPase-activating protein (GAP) for the Rheb GTPase, a direct stimulator of mTORC1 (Manning and Cantley, 2003; Wulfschleger *et al.*, 2006), while TSC1 and TBC1D7 support TSC2 function by stabilizing TSC2 protein and facilitating its GAP activity toward Rheb, respectively. Thus, active TSC complex functions as a key suppressor of Rheb, while Rheb directly interacts with and stimulates mTORC1 activity (Sancak *et al.*, 2007; Long *et al.*, 2005). Previous biochemical and genetic studies demonstrated that Akt directly

phosphorylates and inhibits TSC2 function to stimulate Rheb-mTORC1 activity (see later section) (Figure 1). The second approach depends on PRAS40, a suppressor of mTORC1 (Sancak et al., 2007; Vander Haar et al., 2007). Akt phosphorylates PRAS40 on Thr224 (Sancak et al., 2007; Vander Haar et al., 2007), which decreases the binding affinity of PRAS40 for mTORC1, thereby relieving its inhibitory effect on mTORC1 activity.

In addition to Akt, Erk1/2 and RSK1 also mediate the positive effects of growth factors on mTORC1 activity through TSC (Ma et al., 2005; Roux et al., 2004). In response to growth factor stimulation, Erk1/2 phosphorylates TSC2 on Ser664, while RSK1 phosphorylates TSC2 on Ser1798. These phosphorylations lead to the inactivation of TSC2 and the activation of mTORC1. TNF signaling also activates mTORC1. IKK β , a major downstream kinase in the TNF α signaling pathway, physically interacts with and phosphorylates TSC1 at Ser487 and Ser511, resulting in suppression of the function of the TSC complex (Lee et al., 2007). Through control of the TSC-Rheb axis, the growth factor pathway and TNF pathways positively regulate mTORC1 activity.

Growth factor-induced mTORC2 activation

Similar to mTORC1, mTORC2 activation is also stimulated by growth factors-PI3K signaling. However, the exact mechanisms by which mTORC2 is activated by growth factor-PI3K signaling remains to be fully understood. By genetic screening in yeast, Zinzalla et al. (2011) found that ribosomes are required for TORC2 signaling. The observations in mammalian systems indicate that this mechanism is also conserved in mammals. Knocking down NIP7, which is required for ribosomal 60S subunit maturation, or knocking down large or small ribosome subunit proteins, invariably reduces mTORC2 activation by growth factors. Active mTORC2 physically associates with the ribosome, and insulin stimulation enhances this association in a manner dependent on PI3K activity. Because the capacity of cell growth and proliferation mainly depend on the contents of ribosomes, this study first proposed a novel mechanism by which a component that reflects cellular anabolic capacity is coordinated with extracellular stimuli to stimulate mTORC2 activity.

Regulation of mTOR Signaling by Energy

Essential cellular anabolic processes such as ribosome biogenesis, mRNA translation, and protein synthesis all consume energy and, accordingly, mTORC1 activity is sensitive to and finely tuned by levels of cellular energy. AMP-activated protein kinase (AMPK) is a major regulator of cellular and organismal energy homeostasis. It coordinates multiple metabolic pathways to balance energy supply and demand to ultimately modulate cell and organ growth. In the context of mTORC1 signaling, AMPK acts as an important suppressor of mTORC1 (Figure 1). When cellular energy level is low, active AMPK phosphorylates and activates TSC2, which leads to the inactivation of mTORC1 (Inoki et al., 2003). AMPK phosphorylates also Raptor on two well-conserved serine residues, Ser722 and Ser792, and inhibits mTORC1 (Gwinn et al., 2008). The phosphorylation of Raptor induces its binding

to the 14-3-3 protein and PRAS40, a negative regulator of mTORC1. The phosphorylation of Raptor by AMPK is also required for cell-cycle arrest induced by energy stress. A recent study indicates that cellular glucose levels also regulate mTORC1 activity. Glucose deprivation triggers the p38 MAPK-PRAK pathway and induces PRAK to phosphorylate Ser130 of the Rheb small GTPase. Rheb phosphorylation reduces its binding to guanine nucleotides (Zheng et al., 2011), thus attenuating GTPase function and its ability to activate mTORC1.

Regulation of mTOR Signaling by Cellular Stress

TSC collects various stress signals to suppress mTORC1

Hypoxia suppresses mTORC1 activity (Brugarolas et al., 2004). Oxygen is an essential regulator of cellular metabolism. Under hypoxic conditions, cells promptly activate a variety of adaptive responses that limit energy expenditure through inhibition of energy-intensive processes including protein synthesis. Inhibition of mTORC1 activity is observed when cells are exposed to the low oxygen, and mTORC1 inhibition under hypoxic conditions depends on REDD1 and TSC (DeYoung et al., 2008; Brugarolas et al., 2004). REDD1 (also known as RTP801/Dig1/DDIT4) is a member of a gene family that includes its paralog REDD2 (RTP801L, DDIT4L) and the *Drosophila* orthologs *Scylla* and *Charybdis* (Ellisen et al., 2002). Like the *Drosophila* orthologs, mammalian REDD1/2 are highly induced at the transcription level in response to a variety of stresses, including DNA damage, dexamethasone, and hypoxia (Corradetti et al., 2005). Subcellular localization of the TSC complex is critical to its signaling function (see later section). The 14-3-3 proteins have been reported to bind Akt-phosphorylated TSC2 and inhibit the function of TSC2 presumably through altering its subcellular localization. Redd1 and Redd2 stimulate TSC2 activity by repelling its association with 14-3-3 thereby inhibiting mTORC1. However, the precise molecular mechanisms by which Redd1/2 specifically block 14-3-3 association with TSC2 without affecting levels of TSC2 phosphorylation remain to be elucidated (DeYoung et al., 2008; Li et al., 2002).

Regulation of mTORC1 by stress granules

Recent studies suggest that the subcellular localization of mTORC1 is dynamically regulated and determines cellular mTORC1 activity in response to various stresses. Interestingly, the formation of RNA granules under stress conditions appears to be an important process to inhibit mTORC1 activity. RNA granules are cytosolic non-membrane bound compartments that consist of mRNAs and their associated proteins (i.e., mRNPs) (Anderson and Kedersha, 2006). A repertoire of mRNPs determines whether mRNP complexes remain silent, become translationally active, or are degraded (Buchan and Parker, 2009). Under stress conditions such as heat, oxidative and osmotic stress, virus infection, and UV irradiation, translationally silenced mRNPs accumulate in stress granules (SGs) (Anderson and Kedersha, 2008). Studies in *Saccharomyces cerevisiae* have shown that TORC1 partitions in heat-induced SGs, suggesting a physiological role of SGs in the spatiotemporal regulation of mTORC1 activity during stress (Takahara and

Maeda, 2012). Studies in mammalian cells indicate that the dual specificity tyrosine phosphorylation-regulated kinase 3 (DYRK3) plays important roles in regulating SGs dissolution and mTORC1 localization in SGs (Wippich *et al.*, 2013). When DYRK3 is active, it promotes SG dissolution and concomitantly releases mTORC1 into the cytosol, a process important for mTORC1 activation during recovery from cellular stress. In contrast, inactive DYRK3 protects SG formation and sequesters mTORC1 within SGs, thereby inhibiting mTORC1 activity in stress conditions. Furthermore, active DYRK3 directly phosphorylates PRAS40 and relieves its inhibition of mTORC1 activity, providing yet another biochemical mechanism by which DYRK3 activates mTORC1. Interestingly, another study proposed that the formation of SGs inhibits mTORC1 activity in a different manner. Thedieck *et al.* (2013) identified astrin as a novel mTORC1 binding protein. Astrin was originally identified as a kinetochore- and microtubule-associated protein, that functions to maintain microtubule-centromere attachment and separation (Mack and Compton, 2001; Thein *et al.*, 2007). Astrin is also recruited to SGs under cellular stress conditions such as arsenite treatment. Under such metabolic stress conditions, astrin interacts with Raptor and disrupts mTORC1. Thus, sequestration of Raptor within SGs through astrin decreases cellular mTORC1 activity under acute stress conditions (Thedieck *et al.*, 2013). Uncontrolled hyperactivation of mTORC1 is often observed in cancer cells and cells lacking functional TSC complex. In these cells, abundant energy is required for maintaining their cellular anabolisms. Lack of sufficient energy caused by aberrant mTORC1 activation in cancer cells or TSC-null cells causes cell death under severe stress conditions. Therefore, astrin-mediated inhibition of mTORC1 may be one mechanism underlying the anti-apoptotic nature of SG formation in these cells. Consistent with this hypothesis, the expression of astrin is enhanced in certain types of cancer cells. Thus, inhibition of aberrant mTORC1 activation through SG formation may represent a self-defense mechanism in certain cancer cells.

Genotoxic stresses suppress mTORC1

Genotoxic stress, such as DNA damage, also suppresses mTORC1 activity. DNA damage activates the major tumor suppressor p53, which blocks cell cycle progression for cell survival or executes apoptotic cell death (Levine, 1997). In addition, p53 also inhibits cell growth by attenuating certain anabolic processes including protein synthesis, a major energy consuming process, to maintain energy homeostasis under stress conditions (Vousden and Lane, 2007). Genotoxic stresses-induced p53 has been proposed to inhibit mTORC1 activity through the up-regulation of its upstream negative regulators such as PTEN, TSC2, and AMPK β 1 (Feng *et al.*, 2007). Two other genes, sestrin1 and sestrin2, have also been recently identified to mediate DNA damage-induced mTORC1 inhibition (Budanov and Karin, 2008). Both sestrin1 and sestrin2 are transcriptionally induced by p53 under stress conditions. Interestingly, enhanced expression of sestrins is capable of activating AMPK and inhibiting mTORC1 activity. Although the molecular mechanisms by which sestrins activate AMPK remain to be fully understood, many metabolic stresses converge on AMPK activation to attenuate cellular mTORC1 activity.

Regulation of mTOR Signaling by Amino Acids

Amino acids are essential for the activation of mTORC1, but not mTORC2. While it has been known for years that amino acids have crucial roles in mTORC1 activation, the exact molecular mechanisms had remained largely unknown until recently. The studies from several laboratories first discovered that evolutionarily conserved small GTPases, GTR1/2 in yeast and Rags in mammals, are required for amino acid-mediated regulation of mTORC1 (Sancak *et al.*, 2008; Kim *et al.*, 2008; Figure 2). Mammalian Rags, which belong to the Ras-related small GTPase family, consist of four proteins, RagA, B, C, and D, and localize on the lysosomal membrane. Rags form obligate heterodimers of either RagA or RagB with either RagC or RagD. Two members of the heterodimers have opposing nucleotide-loading states. In the state where RagA or RagB binds to GTP while RagC or RagD binds to GDP, the complex has its maximum activity and interacts with Raptor (Efeyan *et al.*, 2012). Upon amino acid stimulation, RagA or RagB binds GTP and the active Rag heterodimer recruits mTORC1 to the lysosomal membrane through its interactions with Raptor (Sancak *et al.*, 2008). Interestingly, accumulation of amino acids in the lysosomal lumen is required for triggering mTORC1 recruitment via the activation of the Rag heterodimer, proposing a lysosome-centric 'inside-out' model of amino acid sensing (Zoncu *et al.*, 2011). Sabatini and coworkers identified a series of key mechanisms by which the accumulation of amino acids in the lysosomal lumen stimulates Rag GTPases. These studies demonstrate that the vacuolar H⁺-adenosine triphosphate ATPase (V-ATPase) functions as a sensor for luminal amino acids and interacts with the Ragulator complex. Ragulator is a pentameric protein complex that includes the mitogen-activated protein kinase scaffold protein 1 (MP1), p14, and p18. Importantly, Ragulator resides on the late endosome/lysosome surface through p18, which has myristoylation and palmitoylation sites near its N-terminus (Nada *et al.*, 2009; Figure 2). Ragulator possesses dual functions in regulating Rags by working not only as a scaffold to anchor Rag heterodimers to the lysosome, but also as a guanine exchange factor (GEF) for RagA and RagB (Bar-Peled *et al.*, 2012). Consistently, loss of Ragulator complex components disperses the Rags from the lysosome to the cytosol and disables the lysosome recruitment of mTORC1 under amino acid-enriched conditions (Sancak *et al.*, 2010). Upon amino acid stimulation, the GEF activity of Ragulator promotes loading of RagA or B with GTP in a v-ATPase-dependent manner, which enables RagA or B to interact with Raptor in mTORC1. Taken together, these studies reveal an important role for the lysosome as an amino acids sensing organelle to initiate cellular anabolic processes through mTORC1 activation in addition to its function as a degradation and recycling center (Figure 2).

Akin to the function of the TSC complex as a GAP that inhibits the Rheb small GTPase and mTORC1 activity, RagA/B small GTPases are also subject to GAP-dependent inhibition. Bar-Peled *et al.* (2013) identified a specific GAP for RagA/B in mammalian cells. The GAP for RagA/B also exists as a multi-protein complex, termed GATOR (GAP activity toward Rags), and its functional counterparts are conserved in yeast (Panchaud *et al.*, 2013). The GATOR complex consists of at

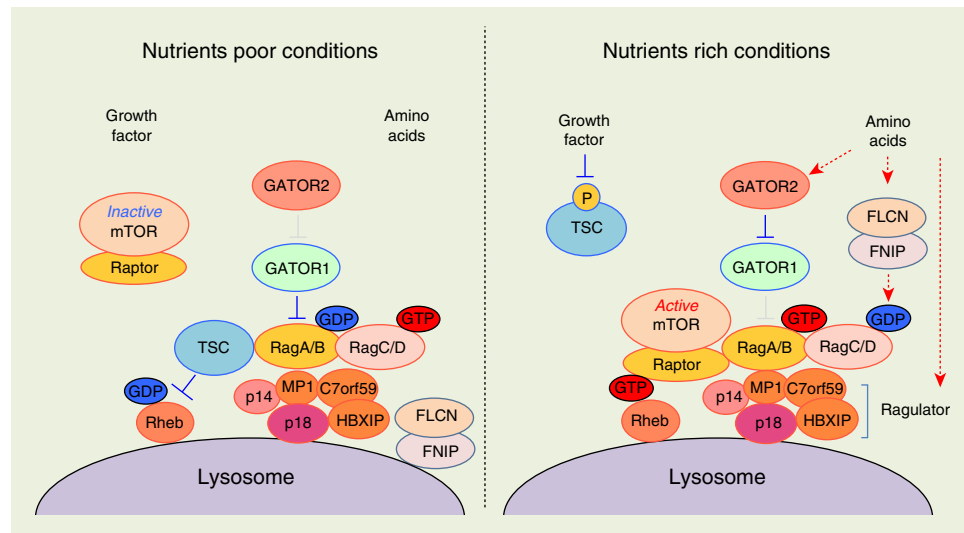


Figure 2 Spatial regulation of mTORC1 on the lysosomal membrane. Under low amino acid and growth factor conditions (left panel), the Rag heterodimer is inactivated, and mTORC1 is not recruited on the lysosomes. The TSC complex, which associates with inactive Rag heterodimer, inhibits Rheb activity. Upon amino acid stimulation (right panel), the Rag heterodimer anchored at the lysosomal membrane through Ragulator recruits mTORC1 to the lysosomes. Concomitantly, Akt stimulated by growth factors phosphorylates TSC2 and dissociates the TSC complex from the lysosomal membrane, thereby inducing Rheb activity on the lysosomal membrane. The active Rheb then stimulates mTORC1 on the lysosomes.

least eight proteins that can be divided into two sub-complexes, GATOR1 and GATOR2. Interestingly, GATOR1 (DEPDC5, Nprl2, and Nprl3) and GATOR2 (Mios, WDR24, WDR59, Seh1L, and Sec13) exert opposite effects on amino acid-induced mTORC1 activity: loss of GATOR1 subunits renders mTORC1 insensitive to amino acid starvation whereas inhibition of GATOR2 subunits suppresses mTORC1 activity. Further experiments showed that mTORC1 inhibition induced by loss of functional GATOR2 was rescued by additional deletion of GATOR1 components, indicating that GATOR2 negatively regulates GATOR1. As yeast homologs of GATOR1 have GAP activity toward Gtr1, a mammalian RagA/B ortholog, GATOR1 exerts specific GAP activity for RagA/B but not for RagC/D.

Recent studies also revealed that folliculin (FLCN), a tumor suppressor mutated in Birt-Hogg-Dube syndrome, with its interacting protein, FNIP1, function together as the RagC/D GAP (Petit *et al.*, 2013; Tsun *et al.*, 2013). For mTORC1 activation, GDP-bound RagC/D is active and necessary for the interaction between mTORC1 and the Rag heterodimer. These observations indicate that the FLCN-FNIP1 complex is an essential regulator in amino acid-induced mTORC1 activation. However, it is puzzling that a tumor suppressor such as FLCN supports mTORC1 activation, which is often observed as hyper-activated in many cancers. Indeed, mice lacking FLCN expression in heart tissue display cardiac hypertrophy with increased mTORC1 activity (Hasumi *et al.*, 2014). There are also mixed reports regarding the function of FLCN in the regulation of AMPK signaling. FLCN has been reported to interact with AMPK. In *Caenorhabditis elegans* and a human follicular thyroid carcinoma cell line, FLCN acts as a suppressor for AMPK (Possik *et al.*, 2014; Yan *et al.*, 2014); however, it functions as a positive regulator for AMPK in murine cardiac myocytes and mammary gland epithelial cells

(Goncharova *et al.*, 2014). These observations suggest that FLCN may regulate multiple steps of different nutrient sensing pathways in a positive or negative manner in order to maintain cellular energy homeostasis.

mTORC1 Associates with Rheb on Lysosomal Membranes

Amino acids and growth factors are two indispensable inputs for mTORC1 activation: lack of either one of these inputs diminishes cellular mTORC1 activation. However, the mechanism by which these essential signals are coordinated to fully activate mTORC1 has not been well understood. Amino acid-induced recruitment of mTORC1 to lysosomal membranes by Rag small GTPases predicts that lysosomal mTORC1 should also associate with Rheb, a direct activator for mTORC1. A study by Sancak *et al.* first reported this hypothesis by demonstrating that amino acid-dependent lysosomal localization of mTORC1 is required for Rheb-induced mTORC1 activation (Figure 2; Sancak *et al.*, 2008). They and others demonstrated that Rheb, which is activated by growth factors through the PI3K-Akt-TSC pathway, resides on the lysosome (Saito *et al.*, 2005; Menon *et al.*, 2014; Sancak *et al.*, 2008). The localization of Rheb to the lysosomal membrane depends on farnesylation of Rheb at its C terminal CAAX motif (Clark *et al.*, 1997). The GTP-bound form of Rheb directly interacts with mTORC1 and stimulates its kinase activity (Long *et al.*, 2005), although the precise molecular mechanism by which active Rheb enhances mTORC1 kinase activity is not fully understood. Nevertheless, these studies indicate that the surface of the lysosome provides a platform where two essential signals, amino acids and growth factors, converge to stimulate mTORC1. This model established by Sancak *et al.* raises the obvious question of how Rheb activity is regulated by growth factor signaling on lysosomes.

More recently, two groups demonstrated that the TSC complex, which is the Rheb GAP and a major upstream negative regulator of mTORC1, also localizes on lysosomes in a regulated manner (Menon *et al.*, 2014; Demetriades *et al.*, 2014; Figure 2). Interestingly, Menon and colleagues demonstrated that TSC2 dissociates from lysosomal membranes in response to insulin stimulation in a manner dependent on its phosphorylation by Akt. The phospho-defective TSC2 mutant, which cannot be phosphorylated by Akt, displays constitutive lysosomal localization and inhibitory activity even in the presence of growth factors. In PTEN deficient cancer cells where the PI3K-Akt pathway is constitutively active, Menon and colleagues further demonstrated that TSC2 remains dissociated from the lysosome, leading to constitutively active mTORC1. These observations suggest that growth factor-mediated spatial regulation of the TSC complex on lysosomes plays an essential role in modulating mTORC1 activity.

Interestingly, Demetriades and colleagues proposed that the TSC complex functions on lysosomes downstream of the nutrient-sensing pathway. TSC2 is capable of binding to RagA, and this binding depends on both an N-terminal motif and a five amino-acid C-terminal motif unique to RagA but not RagB. Upon amino acid withdrawal, TSC2 is recruited to lysosomes through the interaction with GDP-bound inactive RagA (Figure 2). Furthermore, ablation of functional Rag heterodimer on the lysosome by knocking down either RagA or Ragulator or the activation of RagA by knocking down of GATOR1 abolishes lysosomal localization of TSC2 under amino acid deprivation conditions (Sancak *et al.*, 2010; Bar-Peled *et al.*, 2013; Demetriades *et al.*, 2014). These observations indicate that inactive RagA determines TSC2 localization on the lysosomes, where TSC2 inhibits Rheb. In addition, Demetriades *et al.* demonstrated that mTORC1 does not completely dissociate from lysosomes upon amino acid starvation when active Rheb is still localized on lysosomes in TSC2 null cells. Consistently, higher mTORC1 activity could be detected in TSC2 null cells even under amino acid starved conditions. Similar observations were also made by Gao *et al.* in an earlier study (Gao *et al.*, 2002). These data suggest that complete release of mTORC1 from lysosomes requires inactivating both Rags and Rheb. Although some differences exist, these two studies propose two independent and important mechanisms by which localization of the TSC complex is regulated by either growth factor or amino acids with the lysosome serving as a central platform for the recruitment and activation of mTORC1.

Concluding Remarks

Recent studies have made great strides in our understanding of the molecular events by which numerous upstream signals coordinately stimulate or repress mTORC1 activity in response to various cellular cues. Such progress naturally gives rise to new questions in the field. As with other Ras-related small GTPases, the two essential small GTPases, Rag and Rheb, which directly regulate mTORC1, are themselves tightly regulated by their specific GAPs and GEFs. Recent studies have identified all of the key GAP proteins for Rheb, Rag A/B, and RagC/D. However, only the GEF for RagA/B was identified,

raising the question of whether GEFs for RagC/D or Rheb exist. In addition, it remains to be fully understood how amino acids coordinately regulate multiple GAPs (GATOR1 and FLCN) and GEF (Ragulator) for Rags. Although considerable evidence clearly indicates the importance of these GAPs and GEF for Rag guanyl nucleotide binding, there is little data demonstrating that cellular amino acid levels indeed alter the guanyl nucleotide charging of RagA/B or RagC/D. A recent study from the Avruch group demonstrated that withdrawal of extracellular amino acids does not alter Rag guanyl nucleotide charging, although it clearly decreases the Rag-mTORC1 interaction and cellular mTORC1 activity (Oshiro *et al.*, 2014). The authors speculated that additional factors that respond to amino acid sufficiency might bridge Rag and mTORC1 binding to enhance mTORC1 activity. Interestingly, more recent studies demonstrated that cells lacking RagA, the major mammalian homologs of yeast GTR1p, maintain an intact response to insulin for mTORC1 activation. Furthermore, in the RagA null MEFs, mTORC1 activity is maintained although mTORC1 activation or inactivation in response to amino acid stimulation or withdrawal, respectively, is significantly compromised compared to RagA expressing MEFs (Efeyan *et al.*, 2014). More intriguingly, mTORC1 activity in heart tissues lacking both RagA and RagB is not significantly changed compared to that in control heart tissues (Kim *et al.*, 2014). These observations raise the possibility that there may be additional mechanisms underlying amino acid-induced mTORC1 activation.

Another unresolved question is where mTORC1 is localized under amino acid starvation conditions and where mTORC1 phosphorylates its multiple, distinct substrates. In amino acid-deprived cells, mTOR is distributed in punctate structures throughout the cytoplasm. In contrast, upon amino acid stimulation, mTOR localizes on lysosomes while its major substrates, such as S6K1 and 4EBP1, remain diffusely distributed throughout the cytoplasm with mRNPs. Therefore, how can activated mTORC1 sequestered on lysosomal membranes contact and phosphorylate its substrates that are diffusely localized throughout the cytoplasm? One possibility is that mTORC1 primed on the surface of lysosomes or other cytoplasmic membrane compartments dissociates from these membrane platforms to gain access to its substrates in a Rag- and/or Rheb-dependent manner. Consistent with this hypothesis, sucrose density sedimentation assays reveal that the vast majority of mTORC1 as well as Rags fractionate in cytoplasmic fractions under amino acid sufficient conditions (Oshiro *et al.*, 2014). However, these subcellular fractionation experiments are inconsistent with immunocytochemistry findings that show the majority of mTOR and Rags localized to lysosomal membranes under amino acid sufficient conditions. One concern is that many laboratories in the field, including ours, use the same antibodies to detect mTOR and Rag for immunolocalization studies. It is possible that these antibodies preferentially recognize mTOR and Rag epitopes only when these factors are configured on lysosomal membranes, an idea that should be easily testable. Furthermore, the identification of additional factors that interact with Rag-mTORC1 and Rheb and the development of new methodologies to track the *in vivo* movements of mTORC1 with single-molecule precision will allow us to answer many of these remaining

questions in the field while generating the next wave of questions in the mTOR signaling pathway.

Acknowledgment

This work was supported by the NIH grants (DK083491 to KI).

See also: Organelles: Structure and Function: Conventional and Secretory Lysosomes; The Late Endosome

References

- Alessi, D.R., James, S.R., Downes, C.P., *et al.*, 1997. Characterization of a 3-phosphoinositide-dependent protein kinase which phosphorylates and activates protein kinase Balpha. *Current Biology* 7, 261–269.
- Alessi, D.R., Kozlowski, M.T., Weng, Q.P., Morrice, N., Avruch, J., 1998. 3-Phosphoinositide-dependent protein kinase 1 (PDK1) phosphorylates and activates the p70 S6 kinase *in vivo* and *in vitro*. *Current Biology* 8, 69–81.
- Anderson, P., Kedersha, N., 2006. RNA granules. *Journal of Cell Biology* 172, 803–808.
- Anderson, P., Kedersha, N., 2008. Stress granules: The Tao of RNA triage. *Trends in Biochemical Sciences* 33, 141–150.
- Bar-Peled, L., Chantranupong, L., Cherniack, A.D., *et al.*, 2013. A tumor suppressor complex with GAP activity for the Rag GTPases that signal amino acid sufficiency to mTORC1. *Science* 340, 1100–1106.
- Bar-Peled, L., Schweitzer, L.D., Zoncu, R., Sabatini, D.M., 2012. Regulator is a GEF for the rag GTPases that signal amino acid levels to mTORC1. *Cell* 150, 1196–1208.
- Brugarolas, J., Lei, K., Hurley, R.L., *et al.*, 2004. Regulation of mTOR function in response to hypoxia by REDD1 and the TSC1/TSC2 tumor suppressor complex. *Genes & Development* 18, 2893–2904.
- Buchan, J.R., Parker, R., 2009. Eukaryotic stress granules: The ins and outs of translation. *Molecular Cell* 36, 932–941.
- Budanov, A.V., Karin, M., 2008. p53 Target genes sestrin1 and sestrin2 connect genotoxic stress and mTOR signaling. *Cell* 134, 451–460.
- Burnett, P.E., Barrow, R.K., Cohen, N.A., Snyder, S.H., Sabatini, D.M., 1998. RAFT1 phosphorylation of the translational regulators p70 S6 kinase and 4E-BP1. *Proceedings of the National Academy of Sciences of the United States of America* 95, 1432–1437.
- Choi, J., Chen, J., Schreiber, S.L., Clardy, J., 1996. Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science* 273, 239–242.
- Clark, G.J., Kinch, M.S., Rogers-Graham, K., *et al.*, 1997. The Ras-related protein Rheb is farnesylated and antagonizes Ras signaling and transformation. *Journal of Biological Chemistry* 272, 10608–10615.
- Corradetti, M.N., Inoki, K., Guan, K.L., 2005. The stress-induced proteins RTP801 and RTP801L are negative regulators of the mammalian target of rapamycin pathway. *Journal of Biological Chemistry* 280, 9769–9772.
- Demetriades, C., Doumpas, N., Teleman, A.A., 2014. Regulation of TORC1 in response to amino acid starvation via lysosomal recruitment of TSC2. *Cell* 156, 786–799.
- DeYoung, M.P., Horak, P., Sofer, A., Sgroi, D., Ellisen, L.W., 2008. Hypoxia regulates TSC1/2-mTOR signaling and tumor suppression through REDD1-mediated 14-3-3 shuttling. *Genes & Development* 22, 239–251.
- Dibble, C.C., Elis, W., Menon, S., *et al.*, 2012. TBC1D7 is a third subunit of the TSC1-TSC2 complex upstream of mTORC1. *Molecular Cell* 47, 535–546.
- Dorrello, N.V., Peschiaroli, A., Guardavaccaro, D., *et al.*, 2006. S6K1- and betaTRCP-mediated degradation of PDCD4 promotes protein translation and cell growth. *Science* 314, 467–471.
- Duvel, K., Yecies, J.L., Menon, S., *et al.*, 2010. Activation of a metabolic gene regulatory network downstream of mTOR complex 1. *Molecular Cell* 39, 171–183.
- Efeyan, A., Schweitzer, L.D., Bilate, A.M., *et al.*, 2014. RagA, but Not RagB, is essential for embryonic development and adult mice. *Developmental Cell* 29, 321–329.
- Efeyan, A., Zoncu, R., Sabatini, D.M., 2012. Amino acids and mTORC1: From lysosomes to disease. *Trends in Molecular Medicine* 18, 524–533.
- Egan, D.F., Shackelford, D.B., Mihaylova, M.M., *et al.*, 2011. Phosphorylation of ULK1 (hATG1) by AMP-activated protein kinase connects energy sensing to mitophagy. *Science* 331, 456–461.
- Ellisen, L.W., Ramsayer, K.D., Johannessen, C.M., *et al.*, 2002. REDD1, a developmentally regulated transcriptional target of p63 and p53, links p63 to regulation of reactive oxygen species. *Molecular Cell* 10, 995–1005.
- Feldman, M.E., Apse, B., Uotila, A., *et al.*, 2009. Active-site inhibitors of mTOR target rapamycin-resistant outputs of mTORC1 and mTORC2. *PLoS Biology* 7, e38.
- Feng, Z., Hu, W., de Stanchina, E., *et al.*, 2007. The regulation of AMPK beta1, TSC2, and PTEN expression by p53: Stress, cell and tissue specificity, and the role of these gene products in modulating the IGF-1-AKT-mTOR pathways. *Cancer Research* 67, 3043–3053.
- Ganley, I.G., Lam du, H., Wang, J., *et al.*, 2009. ULK1-ATG13-FIP200 complex mediates mTOR signaling and is essential for autophagy. *Journal of Biological Chemistry* 284, 12297–12305.
- Gao, D., Inuzuka, H., Tan, M.K., *et al.*, 2011. mTOR drives its own activation via SCF(betaTrCP)-dependent degradation of the mTOR inhibitor DEPTOR. *Molecular Cell* 44, 290–303.
- Gao, X., Zhang, Y., Arrazola, P., *et al.*, 2002. Tsc tumour suppressor proteins antagonize amino-acid-TOR signalling. *Nature Cell Biology* 4, 699–704.
- Gingras, A.C., Gygi, S.P., Raught, B., *et al.*, 1999. Regulation of 4E-BP1 phosphorylation: A novel two-step mechanism. *Genes & Development* 13, 1422–1437.
- Gingras, A.C., Raught, B., Sonenberg, N., 2001. Regulation of translation initiation by FRAP/mTOR. *Genes & Development* 15, 807–826.
- Goncharova, E.A., Goncharov, D.A., James, M.L., *et al.*, 2014. Folliculin controls lung alveolar enlargement and epithelial cell survival through E-cadherin, LKB1, and AMPK. *Cell Reports* 7, 412–423.
- Gross, J.D., Moerke, N.J., von der Haar, T., *et al.*, 2003. Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* 115, 739–750.
- Guertin, D.A., Stevens, D.M., Thoreen, C.C., *et al.*, 2006. Ablation in mice of the mTORC components raptor, rictor, or mLST8 reveals that mTORC2 is required for signaling to Akt-FOXO and PKCalpha, but not S6K1. *Developmental Cell* 11, 859–871.
- Gwinn, D.M., Shackelford, D.B., Egan, D.F., *et al.*, 2008. AMPK phosphorylation of raptor mediates a metabolic checkpoint. *Molecular Cell* 30, 214–226.
- Hara, K., Maruki, Y., Long, X., *et al.*, 2002. Raptor, a binding partner of target of rapamycin (TOR), mediates TOR action. *Cell* 110, 177–189.
- Harrington, L.S., Findlay, G.M., Gray, A., *et al.*, 2004. The TSC1-2 tumor suppressor controls insulin-Pi3K signaling via regulation of IRS proteins. *Journal of Cell Biology* 166, 213–223.
- Hasumi, Y., Baba, M., Hasumi, H., *et al.*, 2014. Folliculin (Fln) inactivation leads to murine cardiac hypertrophy through mTORC1 deregulation. *Human Molecular Genetics* 23, 5706–5719.
- Holz, M.K., Ballif, B.A., Gygi, S.P., Blenis, J., 2005. mTOR and S6K1 mediate assembly of the translation preinitiation complex through dynamic protein interchange and ordered phosphorylation events. *Cell* 123, 569–580.
- Horton, J.D., Goldstein, J.L., Brown, M.S., 2002. SREBPs: Activators of the complete program of cholesterol and fatty acid synthesis in the liver. *Journal of Clinical Investigation* 109, 1125–1131.
- Hosokawa, N., Hara, T., Kaizuka, T., *et al.*, 2009. Nutrient-dependent mTORC1 association with the ULK1-Atg13-FIP200 complex required for autophagy. *Molecular Biology of the Cell* 20, 1981–1991.
- Hsu, P.P., Kang, S.A., Rameseder, J., *et al.*, 2011. The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. *Science* 332, 1317–1322.
- Ikenoue, T., Inoki, K., Yang, Q., Zhou, X., Guan, K.L., 2008. Essential function of TORC2 in PKC and Akt turn motif phosphorylation, maturation and signalling. *EMBO Journal* 27, 1919–1931.
- Inoki, K., Zhu, T., Guan, K.L., 2003. TSC2 mediates cellular energy response to control cell growth and survival. *Cell* 115, 577–590.
- Jacinto, E., Facchinetti, V., Liu, D., *et al.*, 2006. SIN1/MIP1 maintains rictor-mTOR complex integrity and regulates Akt phosphorylation and substrate specificity. *Cell* 127, 125–137.
- Jacinto, E., Loewith, R., Schmidt, A., *et al.*, 2004. Mammalian TOR complex 2 controls the actin cytoskeleton and is rapamycin insensitive. *Nature Cell Biology* 6, 1122–1128.
- Jung, C.H., Jun, C.B., Ro, S.H., *et al.*, 2009. ULK-Atg13-FIP200 complexes mediate mTOR signaling to the autophagy machinery. *Molecular Biology of the Cell* 20, 1992–2003.
- Jung, C.H., Ro, S.H., Cao, J., Otto, N.M., Kim, D.H., 2010. mTOR regulation of autophagy. *FEBS Letters* 584, 1287–1295.
- Kim, D.H., Sarbassov, D.D., Ali, S.M., *et al.*, 2002. mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell* 110, 163–175.

- Kim, D.H., Sarbassov, D.D., Ali, S.M., *et al.*, 2003. GbetaL, a positive regulator of the rapamycin-sensitive pathway required for the nutrient-sensitive interaction between raptor and mTOR. *Molecular Cell* 11, 895–904.
- Kim, E., Goraksha-Hicks, P., Li, L., Neufeld, T.P., Guan, K.L., 2008. Regulation of TORC1 by Rag GTPases in nutrient response. *Nature Cell Biology* 10, 935–945.
- Kim, J., Kundu, M., Viollet, B., Guan, K.L., 2011. AMPK and mTOR regulate autophagy through direct phosphorylation of Ulk1. *Nature Cell Biology* 13, 132–141.
- Kim, Y.C., Park, H.W., Sciarretta, S., *et al.*, 2014. Rag GTPases are cardioprotective by regulating lysosomal function. *Nature Communications* 5, 4241.
- Lamming, D.W., Sabatini, D.M., 2013. A Central role for mTOR in lipid homeostasis. *Cell Metabolism* 18, 465–469.
- Laplante, M., Horvat, S., Festuccia, W.T., *et al.*, 2012. DEPTOR cell-autonomously promotes adipogenesis, and its expression is associated with obesity. *Cell Metabolism* 16, 202–212.
- Laplante, M., Sabatini, D.M., 2012. mTOR signaling in growth control and disease. *Cell* 149, 274–293.
- Lee, D.F., Kuo, H.P., Chen, C.T., *et al.*, 2007. IKK beta suppression of TSC1 links inflammation and tumor angiogenesis via the mTOR pathway. *Cell* 130, 440–455.
- Lempinen, H., Halazonetis, T.D., 2009. Emerging common themes in regulation of PI3Ks and PI3Ks. *EMBO Journal* 28, 3067–3073.
- Levine, A.J., 1997. p53, the cellular gatekeeper for growth and division. *Cell* 88, 323–331.
- Levine, B., Kroemer, G., 2008. Autophagy in the pathogenesis of disease. *Cell* 132, 27–42.
- Li, Y., Inoki, K., Yeung, R., Guan, K.L., 2002. Regulation of TSC2 by 14-3-3 binding. *Journal of Biological Chemistry* 277, 44593–44596.
- Loewith, R., Jacinto, E., Wullschlegel, S., *et al.*, 2002. Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. *Molecular Cell* 10, 457–468.
- Long, X., Lin, Y., Ortiz-Vega, S., Yonezawa, K., Avruch, J., 2005. Rheb binds and regulates the mTOR kinase. *Current Biology* 15, 702–713.
- Ma, L., Chen, Z., Erdjument-Bromage, H., Tempst, P., Pandolfi, P.P., 2005. Phosphorylation and functional inactivation of TSC2 by Erk implications for tuberous sclerosis and cancer pathogenesis. *Cell* 121, 179–193.
- Mack, G.J., Compton, D.A., 2001. Analysis of mitotic microtubule-associated proteins using mass spectrometry identifies astrin, a spindle-associated protein. *Proceedings of the National Academy of Sciences of the United States of America* 98, 14434–14439.
- Mamane, Y., Petroulakis, E., Lebacquer, O., Sonenberg, N., 2006. mTOR, translation initiation and cancer. *Oncogene* 25, 6416–6422.
- Manning, B.D., Cantley, L.C., 2003. Rheb fills a GAP between TSC and TOR. *Trends in Biochemical Sciences* 28, 573–576.
- Martina, J.A., Chen, Y., Gucak, M., Puertollano, R., 2012. MTORC1 functions as a transcriptional regulator of autophagy by preventing nuclear transport of TFEB. *Autophagy* 8, 903–914.
- Meng, Z.X., Li, S., Wang, L., *et al.*, 2013. Baf60c drives glycolytic metabolism in the muscle and improves systemic glucose homeostasis through Deptor-mediated Akt activation. *Nature Medicine* 19, 640–645.
- Menon, S., Dibble, C.C., Talbot, G., *et al.*, 2014. Spatial control of the TSC complex integrates insulin and nutrient regulation of mTORC1 at the lysosome. *Cell* 156, 771–785.
- Mizushima, N., Levine, B., Cuervo, A.M., Klionsky, D.J., 2008. Autophagy fights disease through cellular self-digestion. *Nature* 451, 1069–1075.
- Montagne, J., Stewart, M.J., Stocker, H., *et al.*, 1999. Drosophila S6 kinase: A regulator of cell size. *Science* 285, 2126–2129.
- Moser, B.A., Dennis, P.B., Pullen, N., *et al.*, 1997. Dual requirement for a newly identified phosphorylation site in p70S6k. *Molecular and Cellular Biology* 17, 5648–5655.
- Nada, S., Hondo, A., Kasai, A., *et al.*, 2009. The novel lipid raft adaptor p18 controls endosome dynamics by anchoring the MEK-ERK pathway to late endosomes. *EMBO Journal* 28, 477–489.
- Noda, T., Ohsumi, Y., 1998. Tor, a phosphatidylinositol kinase homologue, controls autophagy in yeast. *Journal of Biological Chemistry* 273, 3963–3966.
- Nojima, H., Tokunaga, C., Eguchi, S., *et al.*, 2003. The mammalian target of rapamycin (mTOR) partner, raptor, binds the mTOR substrates p70 S6 kinase and 4E-BP1 through their TOR signaling (TOS) motif. *Journal of Biological Chemistry* 278, 15461–15464.
- Oshiro, N., Rapley, J., Avruch, J., 2014. Amino acids activate mammalian target of rapamycin (mTOR) complex 1 without changing Rag GTPase guanylyl nucleotide charging. *Journal of Biological Chemistry* 289, 2658–2674.
- Oshiro, N., Takahashi, R., Yoshino, K., *et al.*, 2007. The proline-rich Akt substrate of 40 kDa (PRAS40) is a physiological substrate of mammalian target of rapamycin complex 1. *Journal of Biological Chemistry* 282, 20329–20339.
- Owen, J.L., Zhang, Y., Bae, S.H., *et al.*, 2012. Insulin stimulation of SREBP-1c processing in transgenic rat hepatocytes requires p70 S6-kinase. *Proceedings of the National Academy of Sciences of the United States of America* 109, 16184–16189.
- Panchaud, N., Peli-Gulli, M.P., de Virgilio, C., 2013. Amino acid deprivation inhibits TORC1 through a GTPase-activating protein complex for the Rag family GTPase Gtr1. *Science Signaling* 6, ra42.
- Pause, A., Belsham, G.J., Gingras, A.C., *et al.*, 1994. Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. *Nature* 371, 762–767.
- Pearce, L.R., Huang, X., Boudeau, J., *et al.*, 2007. Identification of Protor as a novel Rictor-binding component of mTOR complex-2. *Biochemical Journal* 405, 513–522.
- Pena-Llopis, S., Vega-Rubin-De-Celis, S., Schwartz, J.C., *et al.*, 2011. Regulation of TFEB and V-ATPases by mTORC1. *EMBO Journal* 30, 3242–3258.
- Peterson, T.R., Laplante, M., Thoreen, C.C., *et al.*, 2009. DEPTOR is an mTOR inhibitor frequently overexpressed in multiple myeloma cells and required for their survival. *Cell* 137, 873–886.
- Peterson, T.R., Sengupta, S.S., Harris, T.E., *et al.*, 2011. mTOR complex 1 regulates lipin 1 localization to control the SREBP pathway. *Cell* 146, 408–420.
- Petit, C.S., Rocznik-Ferguson, A., Ferguson, S.M., 2013. Recruitment of folliculin to lysosomes supports the amino acid-dependent activation of Rag GTPases. *Journal of Cell Biology* 202, 1107–1122.
- Possik, E., Jalali, Z., Nouet, Y., *et al.*, 2014. Folliculin regulates ampk-dependent autophagy and metabolic stress survival. *PLoS Genetics* 10, e1004273.
- Pullen, N., Dennis, P.B., Andjelkovic, M., *et al.*, 1998. Phosphorylation and activation of p70S6k by PDK1. *Science* 279, 707–710.
- Rabinowitz, J.D., White, E., 2010. Autophagy and metabolism. *Science* 330, 1344–1348.
- Raught, B., Peiretti, F., Gingras, A.C., *et al.*, 2004. Phosphorylation of eucaryotic translation initiation factor 4B Ser422 is modulated by S6 kinases. *EMBO Journal* 23, 1761–1769.
- Richardson, C.J., Broenstrup, M., Fingar, D.C., *et al.*, 2004. SKAR is a specific target of S6 kinase 1 in cell growth control. *Current Biology* 14, 1540–1549.
- Richter, J.D., Sonenberg, N., 2005. Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* 433, 477–480.
- Rocznik-Ferguson, A., Petit, C.S., Froehlich, F., *et al.*, 2012. The transcription factor TFEB links mTORC1 signaling to transcriptional control of lysosome homeostasis. *Science Signaling* 5, ra42.
- Roux, P.P., Ballif, B.A., Anjum, R., Gygi, S.P., Blenis, J., 2004. Tumor-promoting phorbol esters and activated Ras inactivate the tuberous sclerosis tumor suppressor complex via p90 ribosomal S6 kinase. *Proceedings of the National Academy of Sciences of the United States of America* 101, 13489–13494.
- Rubinshtein, D.C., Gestwicki, J.E., Murphy, L.O., Klionsky, D.J., 2007. Potential therapeutic applications of autophagy. *Nature Reviews Drug Discovery* 6, 304–312.
- Ruvinsky, I., Sharon, N., Lerer, T., *et al.*, 2005. Ribosomal protein S6 phosphorylation is a determinant of cell size and glucose homeostasis. *Genes & Development* 19, 2199–2211.
- Saito, K., Araki, Y., Kontani, K., Nishina, H., Katada, T., 2005. Novel role of the small GTPase Rheb: Its implication in endocytic pathway independent of the activation of mammalian target of rapamycin. *Journal of Biochemistry* 137, 423–430.
- Sancak, Y., Bar-Peled, L., Zoncu, R., *et al.*, 2010. Regulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. *Cell* 141, 290–303.
- Sancak, Y., Peterson, T.R., Shaul, Y.D., *et al.*, 2008. The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science* 320, 1496–1501.
- Sancak, Y., Thoreen, C.C., Peterson, T.R., *et al.*, 2007. PRAS40 is an insulin-regulated inhibitor of the mTORC1 protein kinase. *Molecular Cell* 25, 903–915.
- Sarbassov, D.D., Ali, S.M., Kim, D.H., *et al.*, 2004. Rictor, a novel binding partner of mTOR, defines a rapamycin-insensitive and raptor-independent pathway that regulates the cytoskeleton. *Current Biology* 14, 1296–1302.
- Sarbassov, D.D., Ali, S.M., Sengupta, S., *et al.*, 2006. Prolonged rapamycin treatment inhibits mTORC2 assembly and Akt/PKB. *Molecular Cell* 22, 159–168.
- Sardiello, M., Palmieri, M., Di Ronza, A., *et al.*, 2009. A gene network regulating lysosomal biogenesis and function. *Science* 325, 473–477.
- Schalm, S.S., Blenis, J., 2002. Identification of a conserved motif required for mTOR signaling. *Current Biology* 12, 632–639.
- Schalm, S.S., Fingar, D.C., Sabatini, D.M., Blenis, J., 2003. TOS motif-mediated raptor binding regulates 4E-BP1 multisite phosphorylation and function. *Current Biology* 13, 797–806.
- Settembre, C., Zoncu, R., Medina, D.L., *et al.*, 2012. A lysosome-to-nucleus signalling mechanism senses and regulates the lysosome via mTOR and TFEB. *EMBO Journal* 31, 1095–1108.

- Shah, O.J., Wang, Z., Hunter, T., 2004. Inappropriate activation of the TSC/Rheb/mTOR/S6K cassette induces IRS1/2 depletion, insulin resistance, and cell survival deficiencies. *Current Biology* 14, 1650–1656.
- van Slegtenhorst, M., Nellist, M., Nagelkerken, B., *et al.*, 1998. Interaction between hamartin and tuberlin, the TSC1 and TSC2 gene products. *Human Molecular Genetics* 7, 1053–1057.
- Takahara, T., Maeda, T., 2012. Transient sequestration of TORC1 into stress granules during heat stress. *Molecular Cell* 47, 242–252.
- Thedieck, K., Holzwarth, B., Prentzell, M.T., *et al.*, 2013. Inhibition of mTORC1 by astrin and stress granules prevents apoptosis in cancer cells. *Cell* 154, 859–874.
- Thein, K.H., Kleylein-Sohn, J., Nigg, E.A., Gruneberg, U., 2007. Astrin is required for the maintenance of sister chromatid cohesion and centrosome integrity. *Journal of Cell Biology* 178, 345–354.
- Thoreen, C.C., Kang, S.A., Chang, J.W., *et al.*, 2009. An ATP-competitive mammalian target of rapamycin inhibitor reveals rapamycin-resistant functions of mTORC1. *Journal of Biological Chemistry* 284, 8023–8032.
- Tsun, Z.Y., Bar-Peled, L., Chantranupong, L., *et al.*, 2013. The folliculin tumor suppressor is a GAP for RagC/D GTPases that signal amino acid levels to mTORC1. *Molecular Cell* 52, 495–505.
- Um, S.H., Frigerio, F., Watanabe, M., *et al.*, 2004. Absence of S6K1 protects against age- and diet-induced obesity while enhancing insulin sensitivity. *Nature* 431, 200–205.
- Vander Haar, E., Lee, S.I., Bandhakavi, S., Griffin, T.J., Kim, D.H., 2007. Insulin signalling to mTOR mediated by the Akt/PKB substrate PRAS40. *Nature Cell Biology* 9, 316–323.
- Vousden, K.H., Lane, D.P., 2007. p53 in health and disease. *Nature Reviews Molecular Cell Biology* 8, 275–283.
- Wang, B.T., Ducker, G.S., Barczak, A.J., *et al.*, 2011. The mammalian target of rapamycin regulates cholesterol biosynthetic gene expression and exhibits a rapamycin-resistant transcriptional profile. *Proceedings of the National Academy of Sciences of the United States of America* 108, 15201–15206.
- Wang, L., Harris, T.E., Roth, R.A., Lawrence Jr., J.C., 2007. PRAS40 regulates mTORC1 kinase activity by functioning as a direct inhibitor of substrate binding. *Journal of Biological Chemistry* 282, 20036–20044.
- Wang, X., Li, W., Williams, M., *et al.*, 2001. Regulation of elongation factor 2 kinase by p90(RSK1) and p70 S6 kinase. *EMBO Journal* 20, 4370–4379.
- Weng, Q.P., Kozlowski, M., Belham, C., *et al.*, 1998. Regulation of the p70 S6 kinase by phosphorylation *in vivo*. Analysis using site-specific anti-phosphopeptide antibodies. *Journal of Biological Chemistry* 273, 16621–16629.
- Wick, K.R., Werner, E.D., Langlais, P., *et al.*, 2003. Grb10 inhibits insulin-stimulated insulin receptor substrate (IRS)-phosphatidylinositol 3-kinase/Akt signaling pathway by disrupting the association of IRS-1/IRS-2 with the insulin receptor. *Journal of Biological Chemistry* 278, 8460–8467.
- Wippich, F., Bodenmiller, B., Trajkovska, M.G., *et al.*, 2013. Dual specificity kinase DYRK3 couples stress granule condensation/dissolution to mTORC1 signaling. *Cell* 152, 791–805.
- Wulfschleger, S., Loewith, R., Hall, M.N., 2006. TOR signaling in growth and metabolism. *Cell* 124, 471–484.
- Yan, M., Gingras, M.C., Dunlop, E.A., *et al.*, 2014. The tumor suppressor folliculin regulates AMPK-dependent metabolic transformation. *Journal of Clinical Investigation* 124, 2640–2650.
- Yang, H., Rudge, D.G., Koos, J.D., *et al.*, 2013. mTOR kinase structure, mechanism and regulation. *Nature* 497, 217–223.
- Yang, Q., Inoki, K., Ikenoue, T., Guan, K.L., 2006. Identification of Sin1 as an essential TORC2 component required for complex formation and kinase activity. *Genes & Development* 20, 2820–2832.
- Yecies, J.L., Zhang, H.H., Menon, S., *et al.*, 2011. Akt stimulates hepatic SREBP1c and lipogenesis through parallel mTORC1-dependent and independent pathways. *Cell Metabolism* 14, 21–32.
- Yip, C.K., Murata, K., Walz, T., Sabatini, D.M., Kang, S.A., 2010. Structure of the human mTOR complex I and its implications for rapamycin inhibition. *Molecular Cell* 38, 768–774.
- Yu, Y., Yoon, S.O., Poulogiannis, G., *et al.*, 2011. Phosphoproteomic analysis identifies Grb10 as an mTORC1 substrate that negatively regulates insulin signaling. *Science* 332, 1322–1326.
- Zhao, Y., Xiong, X., Sun, Y., 2011. DEPTOR, an mTOR inhibitor, is a physiological substrate of SCF(betaTrCP) E3 ubiquitin ligase and regulates survival and autophagy. *Molecular Cell* 44, 304–316.
- Zheng, M., Wang, Y.H., Wu, X.N., *et al.*, 2011. Inactivation of Rheb by PRAK-mediated phosphorylation is essential for energy-depletion-induced suppression of mTORC1. *Nature Cell Biology* 13, 263–272.
- Zinzalla, V., Stracka, D., Oppliger, W., Hall, M.N., 2011. Activation of mTORC2 by association with the ribosome. *Cell* 144, 757–768.
- Zoncu, R., Bar-Peled, L., Efeyan, A., *et al.*, 2011. mTORC1 senses lysosomal amino acids through an inside-out mechanism that requires the vacuolar H(+) -ATPase. *Science* 334, 678–683.

The Nucleolus

Al Lamond and T Ly, University of Dundee, Dundee, UK

S Hutten University of Dundee, Dundee, UK and Ludwig-Maximilians-University, Munich, Germany

A Nicolas, University of Dundee, Dundee, UK

© 2016 Elsevier Inc. All rights reserved.

Nucleolar Architecture

Nucleoli are formed and maintained in a self-organizing fashion and are not bound by a membrane (Figure 1; Boisvert *et al.*, 2007; Sirri *et al.*, 2008; Shaw and Brown, 2012; Raska *et al.*, 2006). Nucleolar architecture and function are intimately linked (Stoykova *et al.*, 1985). The prevailing view is that nucleolar architecture is a consequence of the complex, dynamic processes involved in ribosomal RNA (rRNA) synthesis, processing, and ribosome subunit assembly (Melese and Xue, 1995). For example, during each cell division cycle, nucleoli are disassembled upon entry into mitosis and do not reassemble until after transcription is reestablished at the end of mitosis (Leung *et al.*, 2004).

Nucleoli form around clusters of ribosomal DNA (rDNA) genes, so called nucleolar organizing regions (NORs) (Hernandez-Verdun, 2011). Although somatic cells sharing the same genome contain an identical number of NORs, nucleoli can vary in size, shape, and number in different tissues within the same organism. Comparison of nucleoli in human cell lines derived from an embryonal carcinoma and an osteosarcoma, as shown in Figure 2, provides an illustrative example of the wide range of nucleolar morphologies.

The human nucleolus is arranged in a tripartite structure (Movie 1). Individual subnucleolar regions can be distinguished based on differences in contrast when viewed with transmission electron microscopy (TEM). As shown in the electron micrograph in Figure 3, three unique subcompartments are visible in human nucleoli: the fibrillar centre (FC), the dense fibrillar component (DFC), and the granular

component (GC). Immunofluorescence microscopy of selected markers can also be used to visualize the three nucleolar subcompartments, as shown in Figure 4 (Movie 1).

Nucleoli and Ribosome Subunit Biogenesis

Each nucleolar compartment is closely associated with specific steps in ribosome subunit biogenesis as illustrated in Figure 5.

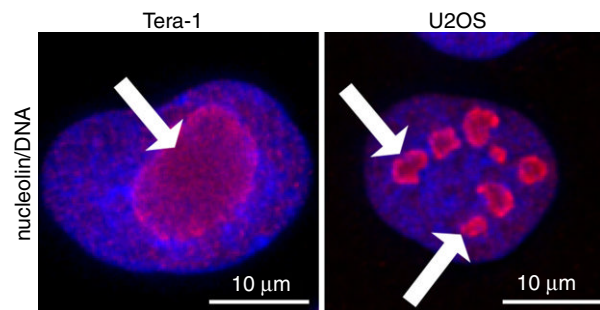


Figure 2 The size, shape, and number of nucleoli is cell-type dependent. This image compares two human cell lines derived from different types of tissues and cancer. Tera-1 cells, which are derived from a metastasis of an embryonic carcinoma of the testes, typically have a single large nucleolus. In contrast, U2OS cells, which are derived from an osteosarcoma of the bone, typically have several smaller nucleoli. Nucleoli were stained with an antinucleolin antibody (red) and DNA was stained with 4',6-diamidino-2-phenylindole (DAPI) (blue). Arrows indicate nucleoli (labeled in red).

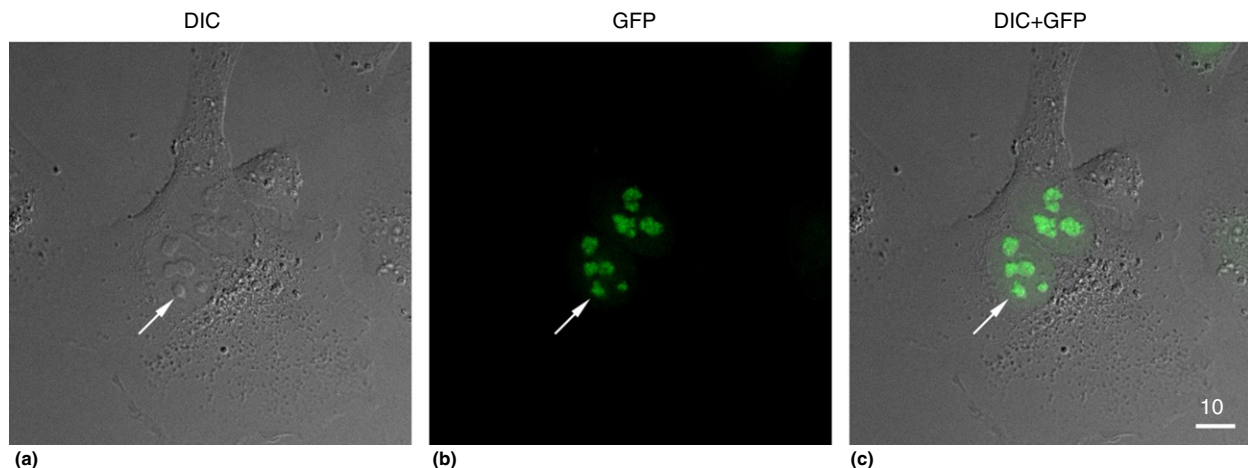


Figure 1 The nucleolus. (a) Nucleoli can be easily observed using differential interference contrast microscopy in cells as high contrast spheroids (representative nucleolus indicated by arrow) in the nucleus. This image shows human osteosarcoma (U2OS) cells. (b) Nucleolar architecture can also be visualized by transfection with a plasmid coding for a green fluorescent protein (GFP)-tagged nucleolar protein (fibrillarin shown here) and visualized under a fluorescence microscope. (c) Overlay of (a) and (b). DIC, Differential interference contrast.

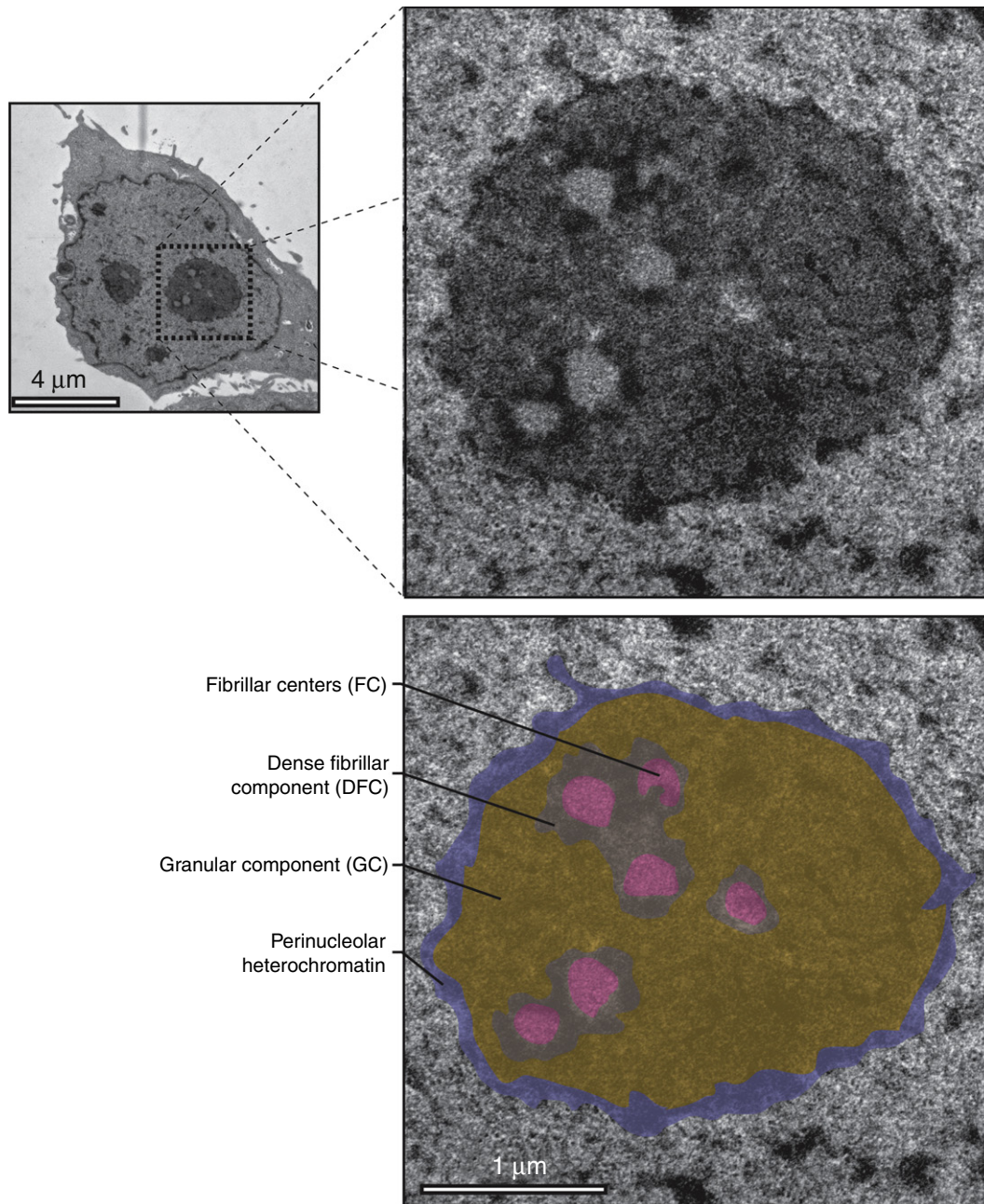


Figure 3 The tripartite structure of human nucleoli. TEM of human nucleoli reveals three nucleolar subcompartments that differ in TEM contrast: fibrillar centers, the DFC, and the GC. The nucleolus is also surrounded by heterochromatin.

Each ribosome is a ribonucleoproteic ribozyme made up of a large 60S subunit and a small 40S subunit, both of which only assemble into a full ribosome once recruited on a mature mRNA for translation. In humans the 60S subunit comprises the 28S, 5.8S, and 5S rRNAs and 47 proteins, while the 40S subunit comprises the 18S rRNA and 33 proteins (for a review of ribosome composition and structure, see [Yusupova and Yusupov, 2014](#)). The 28S, 18S, and 5.8S rRNAs are derived from cleavage and processing of a common precursor, the 47S rRNA, which is synthesized by transcription of rRNA genes by the RNA Pol I machinery. It is generally assumed that this

process takes place either at the border between the FC and DFC or within the FC. In contrast to the 47S rRNA, the 5S rRNA is transcribed in the nucleoplasm by RNA Pol III. The ribosomal proteins are generally small, basic proteins with large spans of unstructured loops, which are thought to become stabilized in ribosomes by interactions with rRNA (reviewed in [Korobeinikova et al., 2012](#)). The mRNAs that encode ribosomal proteins are synthesized by RNA Pol II. Thus, ribosome subunit biogenesis involves all three cellular transcription machineries (RNA Pol I for 47S rRNA, RNA Pol II for ribosomal protein mRNAs, and RNA Pol III for 5S rRNA).

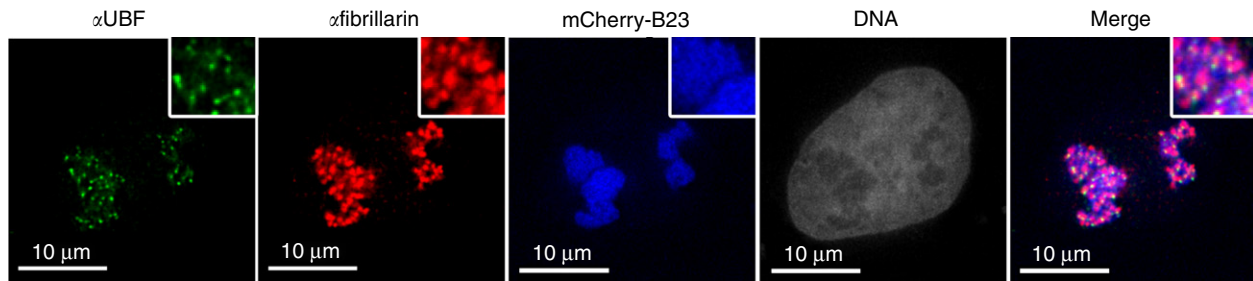


Figure 4 Immunofluorescence microscopy of the tripartite nucleolar architecture in human U2OS cells. Proteins in the FC, such as UBF, are concentrated in punctate bodies within nucleoli. Proteins in the DFC, such as fibrillarin, appear as slightly larger punctate bodies surrounding the FC. Proteins in the GC, such as B23/nucleophosmin, are heterogeneously diffuse within nucleoli. Stacks of 2D images have been reconstructed to produce a movie showing the tripartite nucleolar architecture in 3D, which is available here (Movie 1).

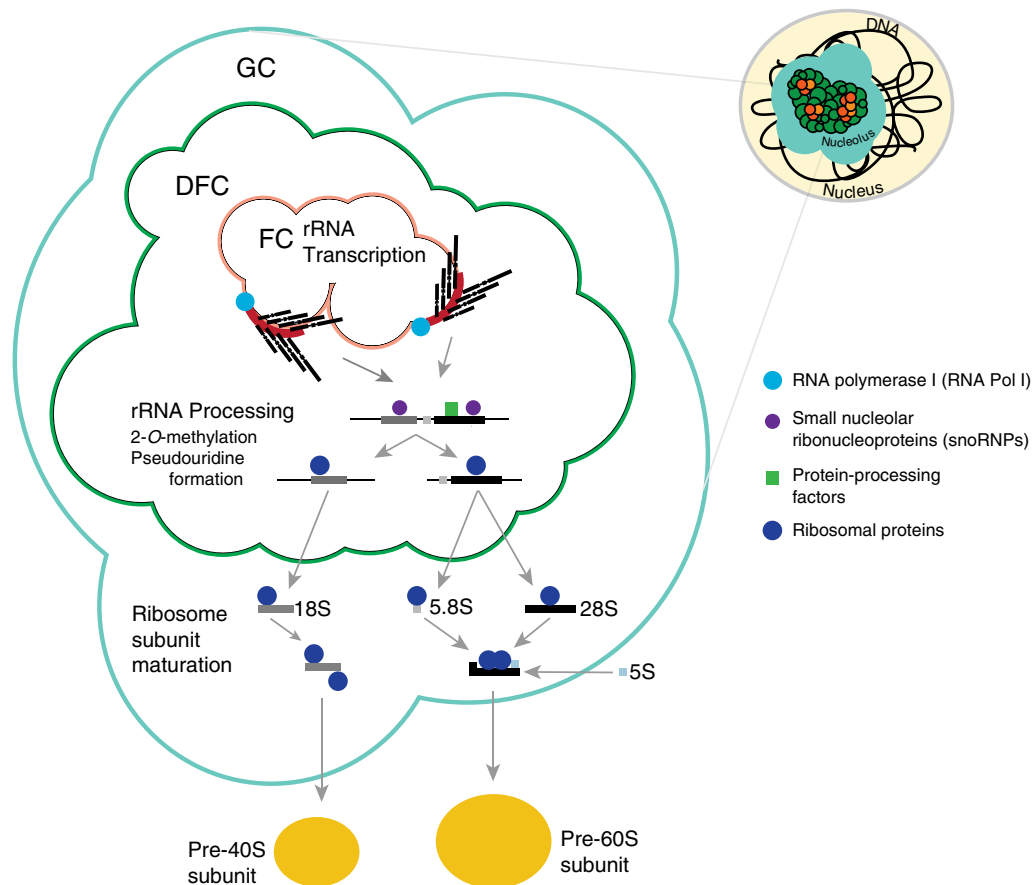


Figure 5 Illustration of ribosome subunit biogenesis in mammalian nucleoli. Transcription of rDNA occurs at active rDNA loci within the FC or at the edge of the FC/DFC. Nascent rRNA is then processed and modified in a number of ways, including cleavage, 2'-O-methylation and pseudouridination, by snoRNPs and protein-processing factors. The pre-40S and pre-60S ribosome subunits are assembled in the DFC and GC before release from the nucleolus. Final processing steps occur in the nucleoplasm prior to transport into the cytoplasm.

Once synthesized, the 47S is cleaved into 18S, 28S, and 5.8S rRNAs within the nucleolus. Factors responsible for cleavage of 47S rRNA and further posttranscriptional modifications of rRNA, such as methylation of specific base and ribose residues, and pseudouridylation, are concentrated within the DFC (for further details refer [Watkins and Bohnsack, 2012](#)). Methylation of ribose 2'-hydroxyl groups (2'-O-methylation) and conversion of uridine to pseudouridine are

guided by small nucleolar RNAs (snoRNAs) as part of small nucleolar ribonucleoprotein (snoRNP) complexes. Two classes of snoRNAs exist, the box C/D type (involved in 2'-O-methylation) and the H/ACA snoRNAs (responsible for pseudouridylation), which are distinguished by the presence of typical sequence motifs, or boxes that interact by Watson-Crick pairing with sequences to be modified (for more details on snoRNAs see [Bratkovic and Rogelj, 2014](#)). Numerous

individual modifications in rRNA are known and these are crucial for ribosome function (Watkins and Bohnsack, 2012). The Fournier laboratory has assembled a 3D ribosomal modification database as an interactive tool to visualize position and consequences of rRNA modifications (Piekna-Przybylska *et al.*, 2008; see Relevant Websites Section – University of Massachusetts Amherst).

Eventually, processed 18S, 28S, and 5.8S rRNAs are assembled with the 5S rRNA and ribosomal proteins in the GC, thus forming pre-40S and pre-60S ribosome subunits. Final processing steps, resulting in mature ribosomal subunits able to participate in translation, occur after export into the cytoplasm (reviewed in Panse and Johnson, 2010).

Ribosomal DNA

In human cells, NORs are located on the short arms (p arms) of the acrocentric chromosomes (chromosomes 13, 14, 15, 21, 22 in human) (Henderson *et al.*, 1972). As rDNA is ~10 times less condensed than neighboring DNA, NORs are visible on metaphase chromosomes as structures called secondary constrictions (as opposed to ‘primary constrictions,’ i.e., centromeres) (Heliot *et al.*, 1997). Due to the continuous presence of specific acidic/argyrophilic proteins at rDNA, even throughout mitosis, NORs can also be visualized by silver staining (Goodpasture and Bloom, 1975; Roussel *et al.*, 1996). A complementary approach to detect rDNA loci is DNA FISH

(fluorescence *in situ* hybridization) using probes specific for the rDNA locus (Figure 6).

Recent studies have shown that the rDNA sequences are sufficient to form a functional, transcriptionally competent nucleolus (Grob *et al.*, 2014). However, several other gene loci have also been shown to be associated with nucleoli (Németh *et al.*, 2010; van Koningsbruggen *et al.*, 2010). These comprise several genes embedded within the heterochromatic shell around the nucleolus and are mostly, if not entirely, transcriptionally silent. An interesting aspect relating nucleolar structure with chromatin and gene expression has been demonstrated lately, as several publications in different model organisms have highlighted the importance of nucleolar chromatin for global genome regulation, such as the balance between eu- and heterochromatin and genomic stability (Paredes and Maggert, 2009; Ide *et al.*, 2010; Kobayashi, 2008).

In total, the diploid human genome contains ~400 rDNA repeat units of ~43 kb (~45 kb in mice) that are predominantly arranged in a head-to-tail fashion (Caburet *et al.*, 2005). The general organization of the individual rDNA repeat has been studied extensively. Only ~13–14 kb of each 43 kb unit code for rDNA and these are separated from each other by ~30 kb of intergenic sequences (‘intergenic spacer’; IGS). Many elements important for regulating rDNA transcription (promoter, enhancer, and terminator sequences) can be found within the IGS. So called ‘replication fork barriers’ (RFBs), situated between individual rDNA repeats, prevent the

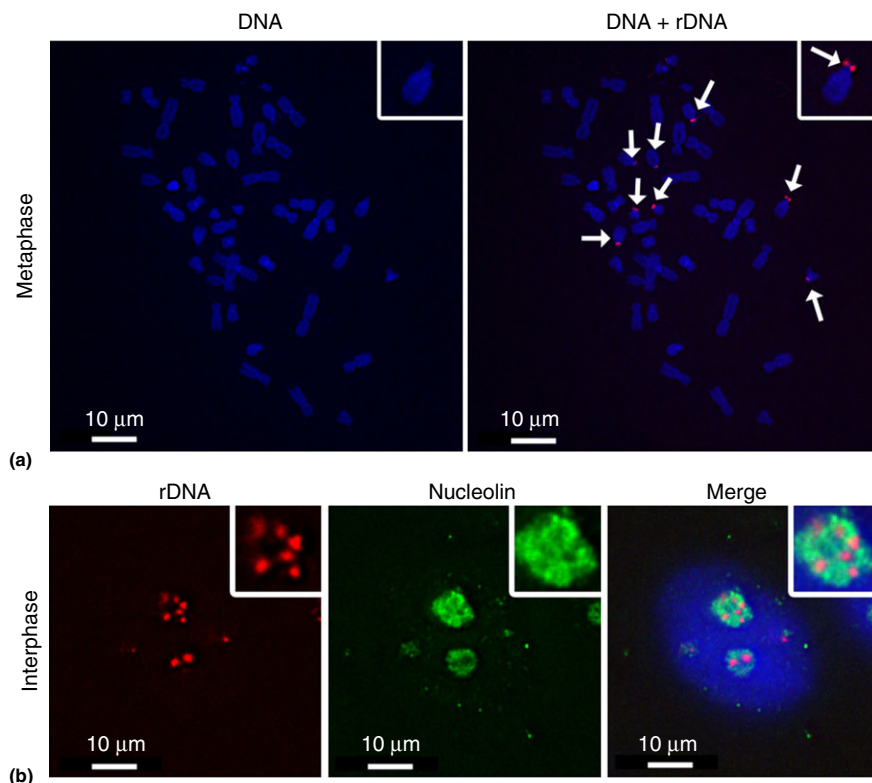


Figure 6 NORs in metaphase and interphase cells. FISH analysis using probes against rDNA loci (stained in red) reveals the subcellular localization of NORs. (a) In metaphase cells, NORs (arrows) are localized to the p-arm of acrocentric chromosomes. (b) In interphase, NORs are localized to the FC in nucleoli.

collision of the replication and transcription machineries during S-phase (Brewer and Fangman, 1988; Gerber *et al.*, 1997; Wiesendanger *et al.*, 1994; Little *et al.*, 1993; Hernandez *et al.*, 1993). In addition, the IGS codes for a small, 150–250 nucleotide long, noncoding RNA, the so-called promoter RNA (pRNA), which is involved in epigenetic regulation of the rDNA locus (see below) (Mayer *et al.*, 2006; Santoro *et al.*, 2010).

Only about half of the rDNA repeats are actively transcribed during interphase while the rest are transcriptionally inactive, but all rDNA repeats reside within nucleoli (Sullivan *et al.*, 2001). These two classes of rDNA repeats differ from each other in their chromatin organization, epigenetic marks, and associated proteins (reviewed in Nemeth and Langst, 2011; McStay and Grummt, 2008). The active repeats carry euchromatic features, such as hypomethylated DNA and euchromatic histone modifications. In contrast, the inactive repeats are organized as heterochromatin, with hypermethylated DNA and repressive histone marks. This pattern is thought to be established by allelic inactivation early during development (Schlesinger *et al.*, 2009) and is actively maintained over the cell cycle.

Although the number of active rDNA copies varies depending on nutrient availability and the rate of cell proliferation, studies in different organisms have shown that even in rapidly dividing cells about half of those repeats remain transcriptionally silent (French *et al.*, 2003; Dammann *et al.*, 1993, 1995; Michel *et al.*, 2005). Maintenance of high numbers of rDNA repeats has been suggested to protect the rDNA locus from homologous recombination events that could reduce rDNA copy number below a putative critical threshold for viability. Extensive rDNA copy number variation exists within a single species. For example, Michel *et al.* (2005) found that *Saccharomyces cerevisiae* clonal variants with 50% fewer rDNA repeats show no significant short-term defects in mitosis, meiosis, or proliferation. Stable inactivation of excess, ‘backup’ rDNA repeats is consistent with a dosage compensation mechanism. Interestingly, dosage compensation appears to be a common theme in nature. For example, in plants a different form of dosage compensation has been described, involving one set of parental NORs being inactivated by a mechanism called ‘nucleolar dominance’ (reviewed in Preuss and Pikaard, 2007; McStay, 2006).

rRNA Transcription

About half of the cellular RNA synthesized in rapidly growing cells corresponds to rRNA transcribed by RNA Pol I (Warner, 1999). Misregulation of ribosome subunit biogenesis is associated with disease, including the so-called ribosomopathies and some cancers (for reviews see, e.g., Teng *et al.*, 2013; Narla and Ebert, 2010; Hannan *et al.*, 2013; Drygin *et al.*, 2010). It is generally assumed that rDNA expression is modulated in two distinct ways: either by adjusting the transcription rate of already active rDNA repeats, as demonstrated, for example, in response to growth factor signaling and nutrient availability, or by changing the number of active rDNA repeats, for example, during differentiation and development (reviewed in McStay and Grummt, 2008; Grummt, 2013).

Synthesis of rRNA is a highly efficient and regulated process that is still being studied intensively (for a recent in-depth review we refer to Goodfellow and Zomerdijs, 2013; Németh *et al.*, 2013). In brief, rRNA transcription, with the exception of the 5S rRNA, is mediated by multiple, densely packed transcription factories of RNA polymerase I. Distinct steps in the transcription process are (a) transcription initiation in the form of formation of the pre-initiation complex (PIC), (b) promoter escape, (c) elongation, and (d) termination, with each of these four steps reported to be subject to regulation mechanisms (Grummt, 2013; Goodfellow and Zomerdijs, 2013). Topological arrangements of the rDNA locus are believed to be important for high rRNA transcription efficiency. For example, the UBF-dimer is part of a specialised nucleosome structure on rDNA called the ‘enhanceosome’ (Stefanovsky *et al.*, 2006), which has been implicated in remodeling rDNA chromatin to a more ‘open’ state that is competent for recruitment of the PIC and transcription (Grob *et al.*, 2014; Sanij *et al.*, 2008). In addition, the formation of DNA-loops and clustering of TTF1 binding sites has been proposed to enhance reinitiation (Nemeth and Langst, 2011; Diermeier *et al.*, 2013; Denissov *et al.*, 2011).

Regulation of rDNA Expression

As described above, approximately half of the rDNA repeats are transcriptionally active, while the rest is silent. Regulation of rDNA expression can be dynamic; for example, by recruiting specific factors to rDNA to prevent binding of the RNA Pol I-associated transcription machinery. Additionally, direct chemical modification of either rDNA repeats, or rDNA-associated histones, can lead either to persistent activation, or silencing, of rDNA repeats that endures upon cell division.

Several key players in this latter, ‘epigenetic’ regulation of rDNA chromatin organization have been described in mammals. At the centre of this regulatory pathway is the rDNA binding protein TTF1 (transcription termination factor 1), which can either activate, or repress, rDNA transcription depending on the protein complex that is formed with TTF1 (reviewed in McStay and Grummt, 2008; Grummt and Langst, 2013). On the one side, TTF1 can recruit NoRC (nucleolar remodeling complex) by interacting with its subunit Tip5 (TTF1-interacting protein 5) (Strohner *et al.*, 2001; Santoro *et al.*, 2002). Another member of the NoRC complex is SNF2h, an ATPase required to shift nucleosomes along rDNA and prevent PIC formation at the promoter (Li *et al.*, 2006; Langst *et al.*, 1998). In addition to TTF1, TIP5 has been described to interact with several other proteins involved in repressing transcriptional activity, such as DNA methyltransferases (DNMT 1 and 3), histone deacetylases (HDAC1/2; Zhou *et al.*, 2002), and histone methyltransferases (HMTs). Importantly, NoRC requires the presence of pRNAs transcribed from the IGS within the hypomethylated (active) rDNA locus. However, pRNAs have also been described to exert transcription-repressive functions independently of TTF1 and NoRC by mediating rDNA *de novo* methylation (Schmitz *et al.*, 2010).

TTF1 has also been shown to interact with activating factors at the rDNA, in particular CSB (Cockayne syndrome protein B), a member of the SNQ/SNF2-like DNA-dependent ATPases

(Yuan *et al.*, 2007; Bradsher *et al.*, 2002). CSB is further associated with G9a, a euchromatic HMT involved in euchromatic histone methylation (Tachibana *et al.*, 2002). Other factors promoting an open chromatin structure include histone demethylases, such as PHF8, or the rDNA promoter binding, methylation-sensitive proteins MBD3 and GADD45a (Schmitz *et al.*, 2009; Feng *et al.*, 2010; Brown and Szyf, 2007). MBD3 and GADD45a binding to the rDNA promoter protects it against *de novo* methylation and actively favours its demethylated and active state.

Interchange of Nucleoli with Other Nuclear Bodies

Nucleoli are discrete, stable structures that can be isolated to high purity from mammalian cells, as shown by scanning electron microscopy in Figure 7. Nevertheless, the proteins within nucleoli are in constant flux with the surrounding environment. This includes proteins that in steady state may appear mainly nucleolar. This is illustrated in Figure 8 for the nucleolar marker protein fibrillarin by fluorescence microscopy using Fluorescence Loss In Photobleaching (FLIP). Importantly, proteins, which under normal conditions do not obviously have a nucleolar localization, may also transiently pass through the nucleolus, for example, the paraspeckle protein PSP1 (Fox *et al.*, 2002, 2005). Stable PSP1 association with the nucleolus can only be observed under specific cellular conditions, such as

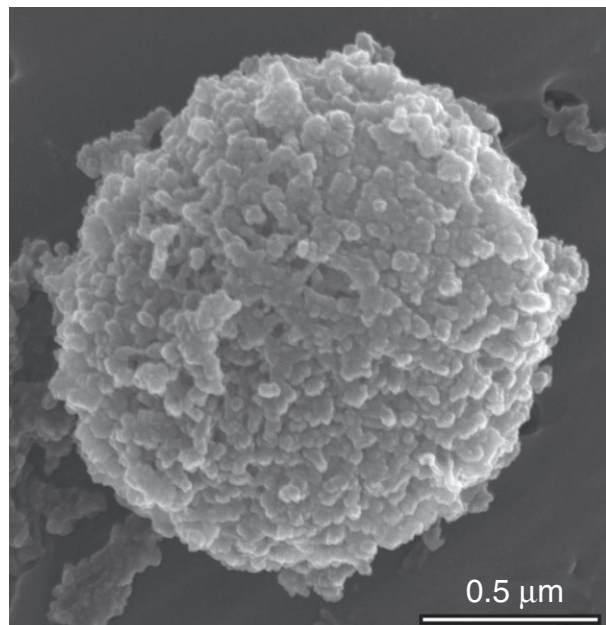


Figure 7 SEM image of a purified HeLa cell nucleolus. Relatively pure nucleolar fractions can be extracted from a variety of cell types. In a typical protocol, cells are lysed and nuclei are separated from the cytoplasm by centrifugation. Purified nuclei are submitted to mild sonication to break them open while keeping nucleoli intact and separating them from the surrounding chromatin. Free nucleoli can be purified from other nuclear material by centrifugation through sucrose cushions. For a detailed protocol, see: http://www.lamondlab.com/newwebsite/Cellular_fractionation_protocols.php

following inhibition of RNA Pol II transcription. Therefore, the nucleolar proteome is, at least partially, shared with other subcellular compartments, especially the nucleoplasm and other nuclear bodies (Box – The Nucleolus and Other Nuclear Bodies).

Out of the plethora of nuclear bodies, the Cajal body (CB) is especially well known to be closely associated with the nucleolus. Their close spatial relationship was initially recognized by Santiago Ramon y Cajal, in 1903, who first described Cajal bodies (CBs) in human neurons and named them ‘nucleolar-associated body’ (Cajal, 1903). This liaison is highlighted by the fact that nucleoli and CBs share a common pool of proteins (reviewed by Machyna *et al.*, 2013). In addition, the CB marker coilin can be found under certain physiological conditions to be associated with, or even to accumulate within, nucleoli. For example, this has been shown in the nucleoli of the hibernating dormouse (Malatesta *et al.*, 1994a, b; Ochs *et al.*, 1994). Given this close spatial relationship, it is not surprising that nucleoli and CBs are also functionally linked, such as by their role in the biogenesis and assembly of RNP complexes (reviewed by Mao *et al.*, 2011), and in the cellular stress response (Boulon *et al.*, 2010). An intensively studied pathway involves the biogenesis of snoRNPs that are required for the processing of rRNA in the nucleolus. They have been shown to assemble and mature within the CB before being transported to the nucleolus (reviewed in Verheggen and Bertrand, 2012). Another example is the biogenesis of the spliceosome component U6 snRNP (reviewed in Machyna *et al.*, 2013; Stanek and Neugebauer, 2006; Patel and Bellini, 2008). After its transcription by RNA Pol III in the nucleoplasm, the U6 snRNA will pass through the nucleolus for its maturation before reaching the CB for its final assembly into the snRNP spliceosomal subunits.

Similar to nucleoli, CBs are dynamic nuclear bodies, whose number and size varies according to the metabolic activity of the cells (Ciocce and Lamond, 2005). Also, their structure changes in response to various types of stresses (Boulon *et al.*, 2010). For example, DNA damage by UV-treatment results in dispersal of CBs and in the appearance of coilin microfoci (Ciocce *et al.*, 2006). Other perturbations highlight the close relationship of CBs with the nucleolus. A prime example here is nucleolar cap formation upon transcriptional inhibition (see below). In addition, impairment of snRNP biogenesis at various stages has been shown to result in relocation of coilin into the nucleolus (reviewed in Boulon *et al.*, 2010). Therefore, both nucleoli and CBs have been described as stress sensors, or ‘hubs,’ for the cellular stress response (Boulon *et al.*, 2010; Emmott and Hiscox, 2009).

The Dynamic Nucleolus

As described in the preceding sections, the nucleolus plays an indispensable role in ribosome subunit biogenesis, which in turn is critical for cell growth, proliferation, and cellular function. As such, the structure and function of nucleoli are tightly coupled to cellular states and can rapidly respond to changing intracellular and extracellular conditions, such as stress, cell cycle stage, and nutrient availability. Here we seek to highlight only a few examples of nucleolar dynamics,

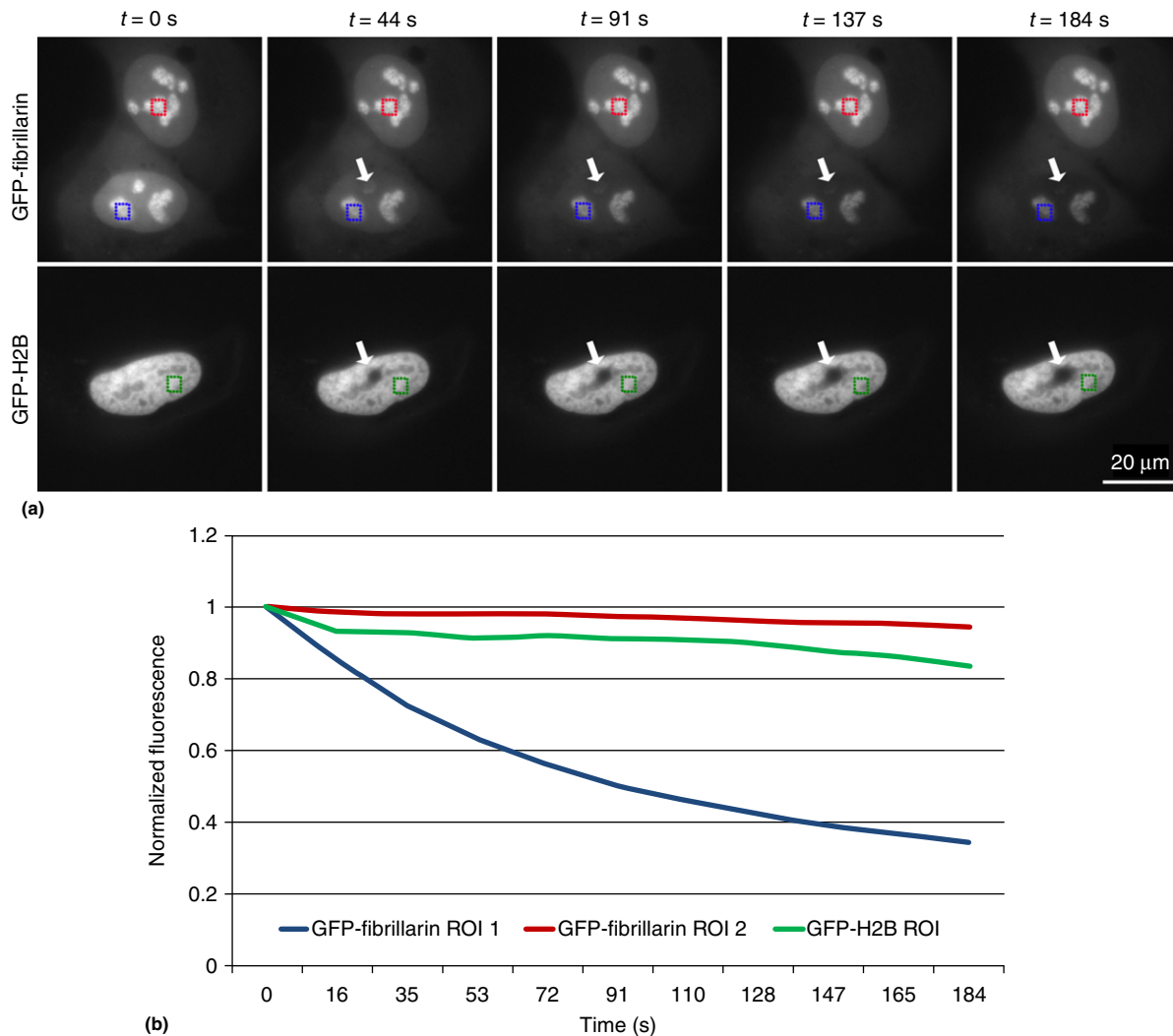


Figure 8 Nucleolar proteins are highly dynamic. (a, top) GFP-fibrillarin was photobleached in one nucleolus (arrow) and fluorescence intensities were measured in regions of interest (ROIs) in another nucleolus within the same cell (blue box) and a different cell (red box). The fluorescence intensity of GFP-fibrillarin decreases in unbleached nucleoli (blue box) and in the nucleoplasm, showing that photobleached GFP-fibrillarin rapidly exchanges with fluorescent GFP-fibrillarin during the timescale of the experiment. Photobleaching does not affect GFP-fibrillarin fluorescence in a neighboring cell (red box). (a, bottom) Photobleaching of GFP-H2B in the area indicated by the arrowhead does not significantly decrease the surrounding GFP-H2B intensity (measured in the green box). In contrast to GFP-fibrillarin, core chromatin components (such as GFP-H2B) are relatively static and do not exchange during the short timescale of the experiment. (b) Quantitation of the fluorescence intensities measured in the red, blue, and green boxes are shown. The photobleaching experiment was performed on a DeltaVision Core Restoration Microscope equipped with a Quantifiable Laser Module (Applied Precision).

including the nucleolar stress response and the assembly and disassembly of nucleoli during the cell division cycle.

Nucleolar Dynamics in Response to Transcriptional Inhibition and DNA Damage

Transcriptional inhibition and DNA damage caused by UV-damage are two well-known examples that result in changes in the protein composition of the nucleolus (Carmo-Fonseca *et al.*, 1992) and in nucleolar segregation (Al-Baker *et al.*, 2004; Govoni *et al.*, 1994), i.e., a separation of the nucleolar FC and DFC compartments, and relocalisation of nucleolar proteins

into nucleolar caps (Shav-Tal *et al.*, 2005; Carmo-Fonseca *et al.*, 1992; Figure 9(a); Movie 2).

Using combined EM and fluorescence microscopy, distinct types of nucleolar caps have been described in the case of inhibition of RNA Pol I dependent transcription, each harboring a distinct set of protein and RNA species (Shav-Tal *et al.*, 2005). It is interesting to note that multiple proteins that are normally not observed by microscopy to be accumulated in nucleoli are recruited to nucleolar caps. The list includes many proteins from other nuclear bodies, including reference markers such as coilin for CBs (Figure 9(b)), PSP1 for paraspeckles, and PML for PML bodies, all of which locate to distinct forms of nucleolar caps. (For more on nuclear bodies,

The Nucleolus and Other Nuclear Bodies

'Nuclear body' is an inclusive term used to collectively describe structures where specific sets of proteins, and often also nucleic acids, can be visualized within the nucleoplasm of a cell. The largest nuclear body is the star of this article, the nucleolus. Other well studied nuclear bodies include PML bodies, Cajal bodies (CBs), splicing speckles, and paraspeckles.

- CBs, previously called 'Coiled bodies,' are named in honor of their discoverer, the famous Spanish histologist Santiago Ramón y Cajal (Cajal, 1903). They are defined as foci enriched in the protein coilin. CBs are thought to function in the maturation of small nuclear ribonucleoproteins (snRNPs) and small nucleolar ribonucleoproteins (snoRNPs) (note that the maturation of at least one snRNP RNA, U6, also requires the nucleolus). CBs are often found in close proximity to the nucleolus with which they have a close relationship. In fact, Cajal first described CBs as 'Nucleolus Accessory Bodies.' This relationship with the nucleolus is illustrated by the presence of nucleolar proteins, such as fibrillarin, inside CBs and the propensity of CB proteins to relocate to the surface of the nucleolus, forming nucleolar 'caps,' in response to some forms of stress, for example, after transcription inhibition (see Figure 8). In rare cases, such as in hibernating dormouse hepatocytes, CBs are detected as defined bodies within the nucleolus (Malatesta *et al.*, 1994a,b).
Another type of nuclear body, Gems (Gemini of Coiled Bodies), had originally been described as foci containing the protein survival of motor neuron (SMN) (Liu and Dreyfuss, 1996). As their name implies, Gems are closely associated with CBs. In fact, in most cells CBs and Gems are undistinguishable. Interestingly, while one strain of HeLa cells was found to have distinct CBs and Gems, in another they overlapped completely (Matera and Frey, 1998). It may be better to consider that Gems are in fact a subtype of coilin-less CBs only found in certain cell types. The significance of this has not yet been understood.
- Promyelocytic leukemia (PML) bodies (Weis *et al.*, 1994), variously referred to either as Nuclear Domain 10 (ND10), PML oncogenic domains (PODs, a misleading name as they are not oncogenic), Kremer bodies, or nuclear dots. Most cells will contain between 10 and 20 PML bodies, each with a diameter in the 0.3–1 μm range. PML bodies tend to be found close to transcriptionally active DNA. They have been suggested to play a role in multiple processes, including transcriptional regulation (Zhong *et al.*, 2000), viral infection and the interferon response (Regad and Chelbi-Alix, 2001), apoptosis and senescence (Bischof *et al.*, 2002; Bernardi and Pandolfi, 2003), and tumor suppression (Gurrieri *et al.*, 2004). Many proteins in PML bodies — including PML — are SUMO-modified, and it has been suggested that SUMOylation, while not required for formation of PML-containing foci, may be essential to recruit other components of the bodies (Seeler and Dejean, 2001).
- Splicing speckles, or interchromatin granule clusters, are irregular nuclear structures of variable size and shape, each appearing as a rather loose grouping of small foci linked by fibrils, typically found in regions of the nucleoplasm with relatively low DNA density (Lamond and Spector, 2003). They contain pre-mRNA splicing factors as well as a subpopulation of RNA Pol II complexes and at least one specific poly(A)⁺ mRNA. Speckles are thought to function mainly as sites of storage or recycling of splicing factors. Thus, speckles increase in size when transcription is inhibited and splicing factors are inactive.
- Paraspeckles, as their name implies, are found in close proximity to splicing speckles (Fox *et al.*, 2002). Their function is as yet unclear, but they have been suggested to play a role in nuclear retention of large hyper A-to-I edited RNAs (Fox *et al.*, 2005). Most proteins that locate to paraspeckles have RNA binding activity, and paraspeckles contain a specific, long noncoding RNA molecule called NEAT1. Interestingly, paraspeckles were first identified as accumulations of PSPC1 (Fox *et al.*, 2002), a protein first identified in a nucleolar proteomics screen (Andersen *et al.*, 2002). PSPC1 accumulates in nucleoli if transcription is inhibited. It has subsequently become clear that most of the paraspeckle proteome transits constantly through the nucleolus. Because many paraspeckle proteins have diverse, mostly unrelated functions, it has been suggested that one of the primary roles of paraspeckles may be to sequester proteins involved in RNA metabolism to regulate their activities.

For as long as the existence of nuclear bodies has been recognized, researchers have debated about the link between their structure and function. Have proteins involved in the organization of nuclear bodies been selected during evolution as scaffolds improving the efficiency of cellular processes by locally concentrating the factors involved, or are the structures we perceive as nuclear bodies the indirect result of the local concentration/exclusion of activities? Maybe nuclear bodies are simply storage areas for excess nuclear proteins. Different explanations may apply to different forms of nuclear bodies. One thing appears certain, nuclear bodies are highly dynamic, and most, if not all, of their components can constantly shuttle in and out of the structures. There is also evidence that some component proteins are shared between different nuclear bodies and move between them in a specific order, demonstrating the existence of molecular sorting pathways in the nucleus akin to those seen in the cytoplasm. This is typified by the behavior of splicing snRNPs, which accumulate in CBs when they first enter the nucleus before relocating to splicing speckles (Sleeman and Lamond, 1999; Leung and Lamond, 2002). Another example is that of nucleolar caps discussed above and at more length in the main body of this article (Shav-Tal *et al.*, 2005).

To learn more about:

- CBs (Machyna *et al.*, 2013),
- PML bodies (Everett and Chelbi-Alix, 2007),
- splicing speckles (Spector and Lamond, 2011),
- paraspeckles (Nakagawa and Hirose, 2012; Fox and Lamond, 2010),
- or nuclear bodies in general (Mao *et al.*, 2011; Dundr, 2012; Caudron-Herger and Rippe, 2012; Sleeman and Trinkle-Mulcahy, 2014; Dundr and Misteli, 2010).

see 'Nuclear Bodies' Box and Movies 3–6). Since many of the nonnucleolar proteins seen to locate to caps have as a common feature RNA binding, it is likely that the massive changes observed upon transcription inhibition are a consequence of the constant RNA-dependent mobility of those factors. Thus, as their specific RNA targets become scarce, these proteins then accumulate at the nucleolar periphery, even if in normal cells the same proteins would be mainly detected outside of the

nucleolus. Nucleolar cap proteins may potentially bind to different RNA or protein targets depending on the nature of the nucleolar cap formed. Importantly, not all stresses have the same effect on the nucleolus and nuclear bodies, suggesting that the caps formed under different stress conditions may vary in protein composition or structure (see Table 1 in Boulon *et al.*, 2010). The phenomenon of nucleolar cap formation reflects the dynamic relationship of the nucleolus with other

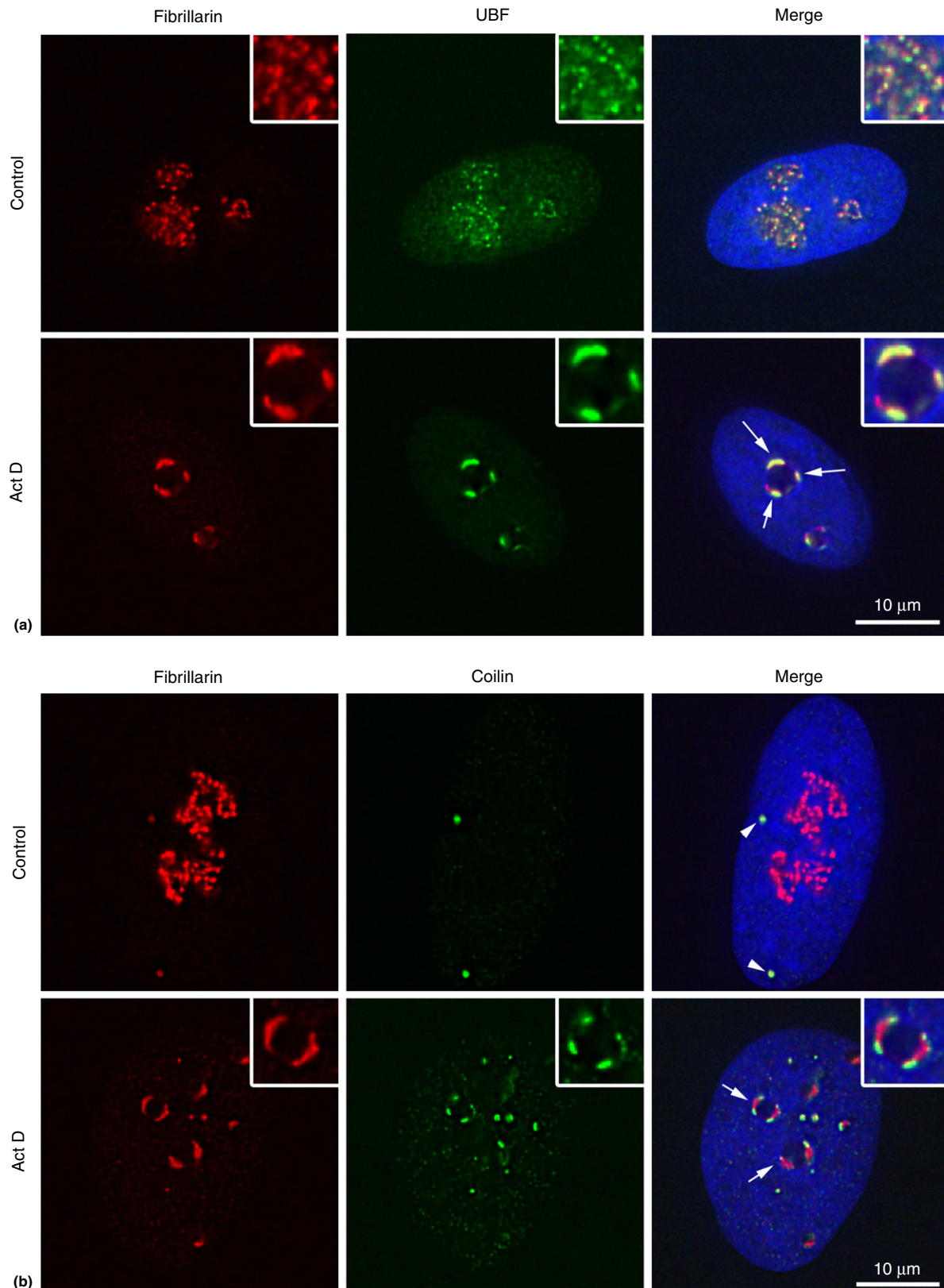


Figure 9 The nucleolar response to stress. (a) In cells stressed by inducing transcriptional inhibition with actinomycin D, nucleolar proteins, such as fibrillarin and UBF, are relocated to nucleolar caps (arrows). (b) Transcriptional inhibition also causes proteins associated with other nuclear bodies, such as the Cajal body-protein coilin (arrowheads), to localise to nucleolar caps (arrows). Importantly, while localization of nucleolar proteins and nuclear body-associated proteins to caps is frequently observed in response to diverse stresses, not all stresses are associated with cap formation, and it is likely that the exact protein composition of the nucleolar cap depends on the specific type of stress (for review see Boisvert *et al.*, 2007). See Movie 2 for a time-lapse video showing the dynamic formation of nucleolar caps in response to transcriptional inhibition.

nuclear bodies or the nucleoplasm and the fact that the structure of nuclear bodies is highly dynamic and intimately linked to their function(s).

Nucleolar Signaling Networks in the Metabolic Stress Response

Given the high proportion of cellular energy spent in ribosome subunit biogenesis (Granneman and Tollervey, 2007), it is perhaps no surprise that nucleolar function and structure are connected to mitogenic signaling and intracellular energy status. One important mechanism that dynamically controls ribosome subunit biogenesis is the regulation of rRNA synthesis. Recent studies have elucidated several ways in which rRNA synthesis is controlled in response to cellular energy levels. These include epigenetic silencing of rDNA gene loci by recruitment of chromatin remodellers (Murayama *et al.*, 2008; Tanaka *et al.*, 2010) and modulation of RNA polymerase I activity (Chen *et al.*, 2013).

The main pathway involved in regulation of cell growth in relation to nutrient sensing is centered on the serine/threonine kinase mTOR, the mammalian homolog of yeast Target of Rapamycin (TOR). mTOR is a component of two complexes, mTORC1 and mTORC2 (Loewith *et al.*, 2002). Only mTORC1 appears to be directly involved in controlling ribosome subunit synthesis, among its diverse growth regulation related functions (Powers and Walter, 1999). It has been shown – primarily in yeast – that TORC1 integrates nutrient and energy levels, as well as growth factor signaling, and affects ribosome subunit synthesis by a variety of mechanisms. For example, TORC1 has been shown in mammals and yeast to regulate the transcription of both Pol I and Pol III dependent rRNA genes (Mahajan, 1994; Cardenas *et al.*, 1999). TORC1 binds to rDNA at promoters, recruits specific transcription factors to them in a nutrient dependent manner, and regulates rDNA chromatin structure (Tsang *et al.*, 2003, 2010; Li *et al.*, 2006). TORC1 also controls the speed of RNA Pol II dependent transcription, which in turn affects the levels of ribosomal protein mRNAs (Powers and Walter, 1999; Cardenas *et al.*, 1999; Hardwick *et al.*, 1999; Beck and Hall, 1999). Finally, besides the general enhancement of RNA Pol II dependent transcription, TORC1 specifically stimulates the translation of terminal oligopyrimidine (TOP) mRNAs, the class of mRNAs that code for ribosomal proteins. However, the exact mechanism by which this occurs is still unclear (Tang *et al.*, 2001; Pende *et al.*, 2004).

Another mechanism of energy homeostasis is the reduction of energy requirements during conditions of stress by arresting cell growth and proliferation. Two key regulators of cell growth and proliferation are the tumor suppressor proteins p53 (reviewed in Lohrum and Vousden, 1999) and p21^{Cif/Waf} (reviewed in Jung *et al.*, 2010), which are both maintained at low levels during active proliferation by targeted protein degradation mechanisms. In the event of cellular stress, the levels of these proteins are rapidly upregulated to arrest proliferation and growth, which can ultimately result in either cellular senescence, autophagy, or apoptosis. This is achieved quickly, at least in part, through suppression of their targeted degradation.

Like p53, levels of the nucleolar protein p14ARF are kept low under cell proliferative states and rapidly upregulated under stress conditions (reviewed in Sherr, 2000). Nucleoplasmic p14ARF can bind and inactivate the protein Hdm2, which is the E3 ligase that targets p53 for degradation, thus leading to stabilization of the p53 protein (reviewed in Vousden, 2002; Sherr, 2001). Recently, a novel p53-independent pathway has been discovered where p14ARF can stabilize levels of p21 by upregulation of the protein FMN2 (Yamada *et al.*, 2013; Vaartjes and Boom, 1987), which is also found in nucleoli. It has also been shown that the novel nucleolar protein ARHGAP11A accumulates in the nucleus upon DNA damage and subsequently binds and stabilizes an active form of p53 (Xu *et al.*, 2013). A common theme that emerges from these studies is that conditions of stress can result in nucleoplasmic accumulation of nucleolar proteins, which stabilize levels of p53 and lead to an arrest of cell growth and proliferation (Rubbi and Milner, 2003). In addition to the mTOR pathway described above, these pathways provide crucial links between cellular energy/stress states and ribosome subunit biogenesis.

Nucleolar Dynamics during the Cell Division Cycle

The nucleolus is dynamically disassembled and reassembled with each cell division (Dundr *et al.*, 2000). Nucleolar disassembly begins during the early stages of mitosis. Simultaneously, transcription of rDNA genes is halted (Weisenberger and Scheer, 1995; Klein and Grummt, 1999), perhaps to allow for chromatin condensation and segregation. As shown in Figures 10 and 11, nucleolar assembly and disassembly is ordered with respect to the tripartite compartments FC, DFC, and GC (Leung *et al.*, 2004). Among the first components to leave nucleoli are RNA Pol I-associated proteins in the FC, followed by proteins in the DFC, and then finally the GC. Reassembly begins during late telophase with recruitment of RNA Pol I-associated factors to rDNA loci and reinitiation of rDNA transcription (i.e., FC proteins), followed by recruitment of DFC and then GC proteins. Interestingly, specific RNA Pol I factors such as UBF remain localized to rDNA loci during the entire cell cycle (Leung *et al.*, 2004; Scheer and Rose, 1984; Roussel *et al.*, 1993). It has been suggested recently that these constitutively chromatin-bound RNA Pol I factors are important in nucleolar reassembly (Grob *et al.*, 2014).

Nucleoli and Viruses

There is a large body of reports that multiple viruses – including RNA and DNA viruses and retroviruses – interact with the nucleolus (for reviews, see Wang *et al.*, 2010; Salvetti and Greco, 2014; Ni *et al.*, 2012; Hiscox *et al.*, 2010; Hiscox, 2007). Many viral proteins localize, at least partially, to nucleoli, while nucleolar proteins are in other cases relocalized from the nucleolus to virus-induced complexes. Viruses also appear to target the nucleolus because it plays a role in several antiviral responses, of which the best known is the p53 pathway. Striking examples of virus–nucleolus interactions can be found among the herpes viruses, which, like most other DNA viruses,

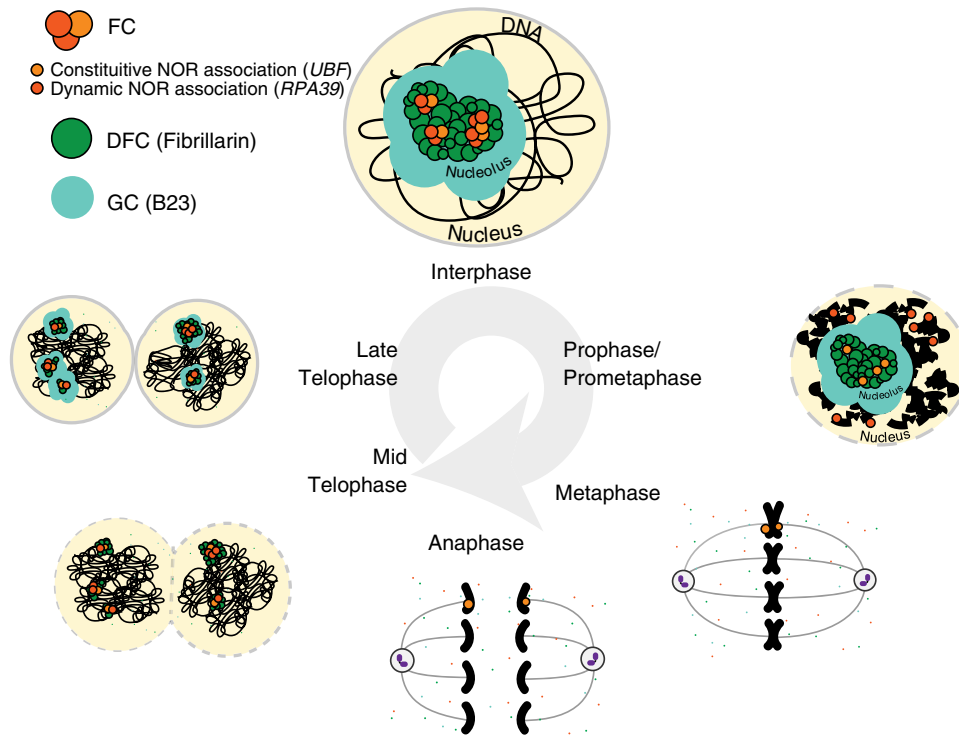


Figure 10 Nucleoli are dynamically assembled and disassembled during the cell division cycle. During prophase/prometaphase, chromatin begins to condense (represented as thick line segments) and some FC proteins, such as RPA39, exit the nucleolus, while other FC components, such as UBF, remain constitutively bound to NORs during mitosis. Between prometaphase and metaphase, the remaining nucleolar architecture is disassembled and remains so until after anaphase. By mid-telophase, FC proteins are recruited back to NORs and the DFC begins to reassemble. The last compartment to reassemble is the GC, which does so during late telophase/early G1.

replicate in the nucleus. For example, herpes simplex 1 replication is associated with first swelling, then major disruption of the nucleolus, until it virtually ceases to exist and all the cellular chromatin is 'marginalized' at the nuclear periphery (Monier *et al.*, 2000; Sirtori and Bosisio-Bestetti, 1967). Similar phenotypes have been observed for other herpes viruses (Luo *et al.*, 2007; Wang *et al.*, 2006; White, 2007). Additionally, various herpes viruses have been shown to encode nucleolus targeting proteins (reviewed in Ni *et al.*, 2012). For example, gamma herpes viruses encode a protein, ORF57, which localizes to the nucleolus where it hijacks ribosome export pathways to export unspliced viral mRNAs (Boyne and Whitehouse, 2006, 2009). However, in most cases the functional significance of virus-nucleolus interactions is still unclear. Moreover, there does not appear to be a single unifying theme behind these interactions. Rather, it is likely that diverse viruses have independently seized opportunities to exploit interactions with nucleolar factors – as would be expected to occur for such a major cellular compartment over hundreds of millions of years of coevolution.

Proteomic Studies on Nucleolar Architecture, Function, and Dynamics

Nucleoli can be successfully purified from human cells using a combination of hypotonic lysis, sonication, and centrifugation in sucrose cushions (see Figure 7; a detailed purification

protocol is available at Relevant Website Section – Lamond Laboratory; an alternate method has also recently been described, see Liang *et al.*, 2012). These methods enable biochemical analysis of the molecular composition and structure of isolated nucleoli, and complement studies performed in cells, such as immunofluorescence microscopy. In recent years, advances in large-scale protein and nucleic acid sequencing have been used to characterize the nucleolar and nucleolar-associated proteins, RNA, and DNA under diverse cellular conditions (van Koningsbruggen *et al.*, 2010; Moore *et al.*, 2011; Boisvert and Lamond, 2010; Boisvert *et al.*, 2012).

Researchers have used mass spectrometry (MS) based identification and quantification to catalog the proteins from human cells that localize to nucleoli, or are otherwise co-purified with nucleoli. These are publicly available in the nucleolar protein database (see Relevant Website Section – NopDB). The nucleolar proteome is highly enriched in proteins associated with ribosome subunit biogenesis and RNA processing, as expected from the principal function of the nucleolus. Interestingly, the nucleolar proteome is also enriched in proteins associated with other biological processes, such as cell cycle and viral replication. These auxiliary functions are consistent with a multifunctional, or plurifunctional, view of the nucleolus (Pederson, 1998; Rubbi and Milner, 2003; Olson, 2004; Pederson and Tsai, 2009).

In addition to in-depth characterization of the protein constituents of the nucleolus, MS has been used to study dynamic changes to the nucleolar proteome under various stress

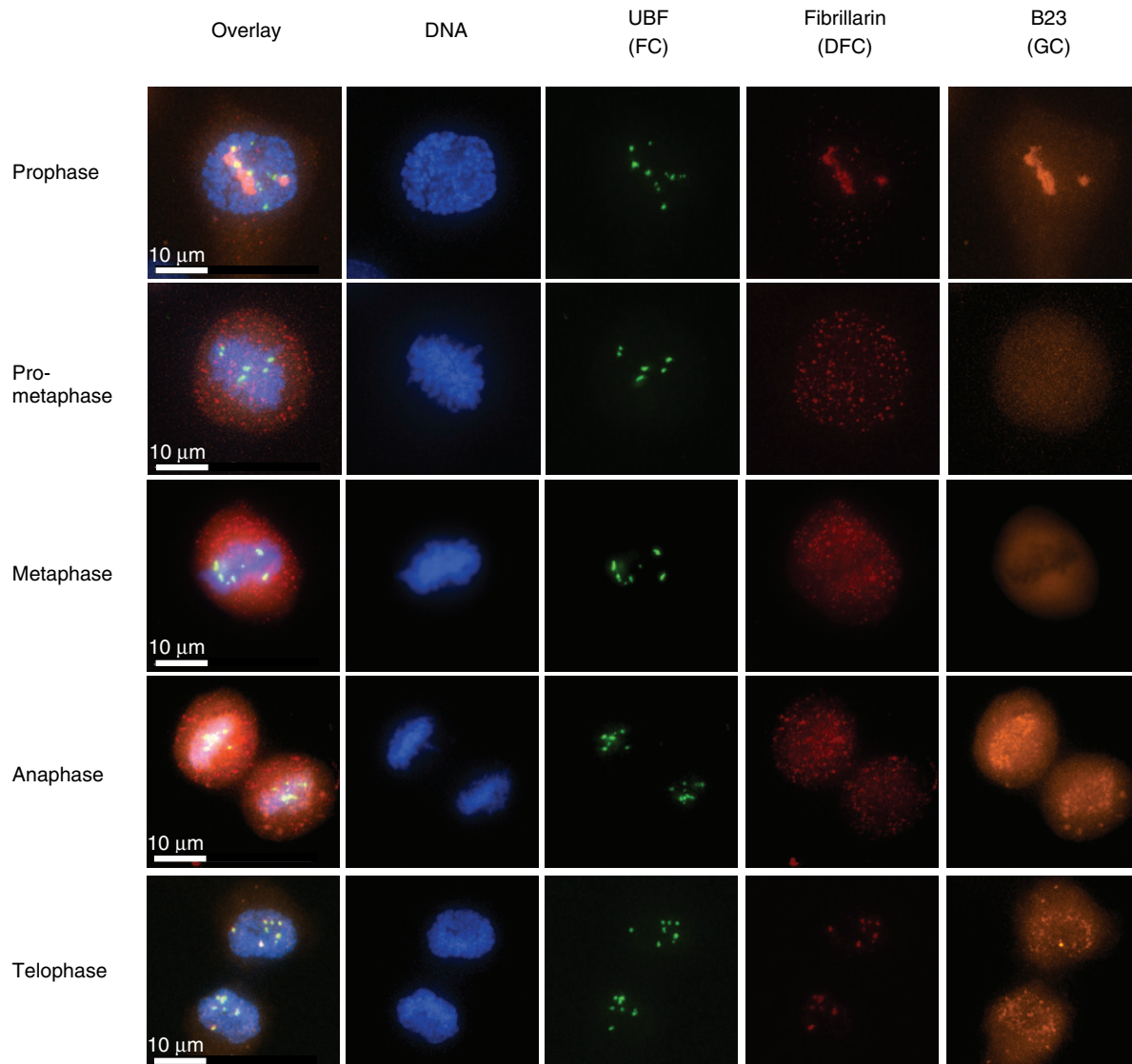


Figure 11 Immunofluorescence analysis of nucleolar disassembly and reassembly. Nucleolar compartments were visualized in asynchronous U2OS cells stained with anti-UBF (FC, green), anti-fibrillarin (DFC, red), and DAPI (DNA, blue), and transfected with a plasmid encoding mCherry-tagged B23 (GC, orange). Cells undergoing mitosis in this asynchronous population were selected for imaging and shown below. Scale bars are 10 μm .

conditions. Key to the success of this approach is accurate relative quantification of protein levels using stable isotope labeling of amino acids in cell culture (SILAC) (Ong *et al.*, 2002). MS-based proteomics has been used to examine nucleolar proteome dynamics upon transcriptional inhibition, viral infection, and DNA damage, revealing stress-specific, complex reorganization of the nucleolus during the stress response (Liang *et al.*, 2012; Westman *et al.*, 2010; Westman and Lamond, 2011; Kar *et al.*, 2011; Emmott *et al.*, 2010a,b; Lam *et al.*, 2010; Jarboui *et al.*, 2012).

In a related approach called ‘spatial proteomics,’ the relative distribution of proteins between different cellular compartments is measured. For example, the measurement of the cytoplasmic, nuclear and nucleolar distribution of proteins was used to characterize dynamic changes in protein localization following DNA damage (Boisvert and Lamond, 2010;

Boisvert *et al.*, 2010). A similar approach was used to measure compartment-specific turnover rates of over 8000 proteins in human cells and identified a subset of proteins that have nucleolar-specific turnover rates (Boisvert *et al.*, 2012; Ahmad *et al.*, 2012). These data highlight proteins that are present in cells in different functional pools and whose regulation is dependent on subcellular localization.

Conclusion

Nucleoli are striking nuclear organelles that are present in all eukaryotes and whose sole function was long thought to be to facilitate and accelerate rRNA synthesis, processing, and pre-ribosome subunit assembly by concentrating factors required for these processes in a single organelle. In the last few

decades, work by many laboratories has begun to expand our understanding of the role of the nucleolus in the cell. In this article, we have highlighted examples of additional roles for the nucleolus, such as a hiding place for signaling proteins that mediate the cellular response against stress, and for integrating cellular contextual signals, such as nutrient availability, to control rates of global gene expression. We anticipate that more mechanisms will be uncovered in the future that will link the nucleolus to other aspects of cell biology.

Methods

Unless otherwise specified, all microscopy images were acquired with a DeltaVision Core Restoration microscope mounted on an Olympus IX71 stand with a 60 × 1.42 or 1.40 NA oil immersion objective.

Acknowledgments

We thank our colleagues in the Lamond lab for advice and discussion, the Joost Zoomerdijk lab for the mCherry-CAST plasmid, Martin Kierans for help with electron microscopy, Markus Posch for help with super-resolution OMX microscopy, and Jonas Telorac (Max Planck Institute for Molecular Genetics) for performing the rDNA FISH. AIL is a Wellcome Trust Principal Research Fellow. This work was supported by grants to AIL from the Wellcome Trust (Grant#:083524/Z/07/Z, 097945/B/11/Z, 073980/Z/03/Z, 08136/Z/03/Z, and 0909444/Z/09/Z), the EU FP7 Prospects network (Grant#: HEALTH-F4-2008-201648) and EpiGeneSys network (Grant#: HEALTH-F4-2010-257082), and the BBSRC sLoLa (Grant#: BB/K003801/1).

See also: Cell Division/Death: Regulation of Cell Growth: Eukaryotic Ribosome Assembly and Export; Ribosomal RNA Processing

References

- Ahmad, Y., Boisvert, F.M., Lundberg, E., Uhlen, M., Lamond, A.I., 2012. Systematic analysis of protein pools, isoforms, and modifications affecting turnover and subcellular localization. *Molecular & Cellular Proteomics* 11 (3), M111 013680.
- Al-Baker, E.A., Boyle, J., Harry, R., Kill, I.R., 2004. A p53-independent pathway regulates nucleolar segregation and antigen translocation in response to DNA damage induced by UV irradiation. *Experimental Cell Research* 292 (1), 179–186.
- Andersen, J.S., Lyon, C.E., Fox, A.H., *et al.*, 2002. Directed proteomic analysis of the human nucleolus. *Current Biology* 12 (1), 1–11.
- Beck, T., Hall, M.N., 1999. The TOR signaling pathway controls nuclear localization of nutrient-regulated transcription factors. *Nature* 402 (6762), 689–692.
- Bernardi, R., Pandolfi, P.P., 2003. Role of PML and the PML-nuclear body in the control of programmed cell death. *Oncogene* 22 (56), 9048–9057.
- Bischof, O., Kirsh, O., Pearson, M., *et al.*, 2002. Deconstructing PML-induced premature senescence. *EMBO Journal* 21 (13), 3358–3369.
- Boisvert, F.M., Ahmad, Y., Gierliński, M., *et al.*, 2012. A quantitative spatial proteomics analysis of proteome turnover in human cells. *Molecular & Cellular Proteomics* 11 (3), M111 011429.
- Boisvert, F.M., van Koningsbruggen, S., Navascués, J., Lamond, A.I., 2007. The multifunctional nucleolus. *Nature Reviews Molecular Cell Biology* 8 (7), 574–585.
- Boisvert, F.M., Lamond, A.I., 2010. p53-Dependent subcellular proteome localization following DNA damage. *Proteomics* 10 (22), 4087–4097.
- Boisvert, F.M., Lam, Y.W., Lamont, D., Lamond, A.I., 2010. A quantitative proteomics analysis of subcellular proteome localization and changes induced by DNA damage. *Molecular & Cellular Proteomics* 9 (3), 457–470.
- Boulon, S., Westman, B.J., Hutten, S., Boisvert, F.M., Lamond, A.I., 2010. The nucleolus under stress. *Molecular Cell* 40 (2), 216–227.
- Boyne, J.R., Whitehouse, A., 2006. Nucleolar trafficking is essential for nuclear export of intronless herpesvirus mRNA. *Proceedings of the National Academy of Sciences of the United States of America* 103 (41), 15190–15195.
- Boyne, J.R., Whitehouse, A., 2009. Nucleolar disruption impairs Kaposi's sarcoma-associated herpesvirus ORF57-mediated nuclear export of intronless viral mRNAs. *FEBS Letters* 583 (22), 3549–3556.
- Bradsher, J., Auriol, J., Proietti de Santis, L., *et al.*, 2002. CSB is a component of RNA pol I transcription. *Molecular Cell* 10 (4), 819–829.
- Bratkovic, T., Rogelj, B., 2014. The many faces of small nucleolar RNAs. *Biochimica et Biophysica Acta* 1839 (6), 438–443.
- Brewer, B.J., Fangman, W.L., 1988. A replication fork barrier at the 3' end of yeast ribosomal RNA genes. *Cell* 55 (4), 637–643.
- Brown, S.E., Szyf, M., 2007. Epigenetic programming of the rRNA promoter by MBD3. *Molecular and Cellular Biology* 27 (13), 4938–4952.
- Caburet, S., Conti, C., Schurra, C., *et al.*, 2005. Human ribosomal RNA gene arrays display a broad range of palindromic structures. *Genome Research* 15 (8), 1079–1085.
- Cajal, S.R., 1903. Un sencillo metodo de coloracion selectiva del reticulo protoplasmico y sus efectos en los diversos organos nerviosos de vertebrados e invertebrados. *Trabajos del Laboratorio de Investigaciones Biológicas de la Universidad de Madrid* 2, 129–221.
- Cardenas, M.E., Cutler, N.S., Lorenz, M.C., Di Como, C.J., Heitman, J., 1999. The TOR signaling cascade regulates gene expression in response to nutrients. *Genes & Development* 13 (24), 3271–3279.
- Carmo-Fonseca, M., Pepperkok, R., Carvalho, M.T., Lamond, A.I., 1992. Transcription-dependent colocalization of the U1, U2, U4/U6, and U5 snRNPs in coiled bodies. *Journal of Cell Biology* 117 (1), 1–14.
- Caudron-Herger, M., Rippe, K., 2012. Nuclear architecture by RNA. *Current Opinion in Genetics & Development* 22 (2), 179–187.
- Chen, S., Seiler, J., Santiago-Reichert, M., *et al.*, 2013. Repression of RNA polymerase I upon stress is caused by inhibition of RNA-dependent deacetylation of PAF53 by SIRT7. *Molecular Cell* 52 (3), 303–313.
- Cioce, M., Boulon, S., Matera, A.G., Lamond, A.I., 2006. UV-induced fragmentation of Cajal bodies. *Journal of Cell Biology* 175 (3), 401–413.
- Cioce, M., Lamond, A.I., 2005. Cajal bodies: A long history of discovery. *Annual Review of Cell and Developmental Biology* 21, 105–131.
- Dammann, R., Lucchini, R., Koller, T., Sogo, J.M., 1993. Chromatin structures and transcription of rDNA in yeast *Saccharomyces cerevisiae*. *Nucleic Acids Research* 21 (10), 2331–2338.
- Dammann, R., Lucchini, R., Koller, T., Sogo, J.M., 1995. Transcription in the yeast rRNA gene locus: Distribution of the active gene copies and chromatin structure of their flanking regulatory sequences. *Molecular and Cellular Biology* 15 (10), 5294–5303.
- Denissov, S., Lessard, F., Mayer, C., *et al.*, 2011. A model for the topology of active ribosomal RNA genes. *EMBO Reports* 12 (3), 231–237.
- Diermeier, S.D., Németh, A., Rehli, M., Grummt, I., Längst, G., 2013. Chromatin-specific regulation of mammalian rDNA transcription by clustered TTF-I binding sites. *PLoS Genetics* 9 (9), e1003786.
- Drygin, D., Rice, W.G., Grummt, I., 2010. The RNA polymerase I transcription machinery: An emerging target for the treatment of cancer. *Annual Review of Pharmacology and Toxicology* 50, 131–156.
- Dundr, M., 2012. Nuclear bodies: Multifunctional companions of the genome. *Current Opinion in Cell Biology* 24 (3), 415–422.
- Dundr, M., Misteli, T., 2010. Biogenesis of nuclear bodies. *Cold Spring Harbor Perspectives in Biology* 2 (12), a000711.
- Dundr, M., Misteli, T., Olson, M.O., 2000. The dynamics of postmitotic reassembly of the nucleolus. *Journal of Cell Biology* 150 (3), 433–446.
- Emmott, E., Hiscox, J.A., 2009. Nucleolar targeting: The hub of the matter. *EMBO Reports* 10 (3), 231–238.
- Emmott, E., Smith, C., Emmett, S.R., Dove, B.K., Hiscox, J.A., 2010b. Elucidation of the avian nucleolar proteome by quantitative proteomics using SILAC and changes in cells infected with the coronavirus infectious bronchitis virus. *Proteomics* 10 (19), 3558–3562.

- Emmott, E., Wise, H., Loucaides, E.M., *et al.*, 2010a. Quantitative proteomics using SILAC coupled to LC-MS/MS reveals changes in the nucleolar proteome in influenza A virus-infected cells. *Journal of Proteome Research* 9 (10), 5335–5345.
- Everett, R.D., Chelbi-Alix, M.K., 2007. PML and PML nuclear bodies: Implications in antiviral defence. *Biochimie* 89 (6–7), 819–830.
- Feng, W., Yonezawa, M., Ye, J., Jenuwein, T., Grummt, I., 2010. PHF8 activates transcription of rRNA genes through H3K4me3 binding and H3K9me1/2 demethylation. *Nature Structural & Molecular Biology* 17 (4), 445–450.
- Fox, A.H., Bond, C.S., Lamond, A.I., 2005. P54nrb forms a heterodimer with PSP1 that localizes to paraspeckles in an RNA-dependent manner. *Molecular Biology of the Cell* 16 (11), 5304–5315.
- Fox, A.H., Lam, Y.W., Leung, A.K., *et al.*, 2002. Paraspeckles: A novel nuclear domain. *Current Biology* 12 (1), 13–25.
- Fox, A.H., Lamond, A.I., 2010. Paraspeckles. *Cold Spring Harbor Perspectives in Biology* 2 (7), a000687.
- French, S.L., Osheim, Y.N., Cioci, F., Nomura, M., Beyer, A.L., 2003. In exponentially growing *Saccharomyces cerevisiae* cells, rRNA synthesis is determined by the summed RNA polymerase I loading rate rather than by the number of active genes. *Molecular and Cellular Biology* 23 (5), 1558–1568.
- Gerber, J.K., Gögel, E., Berger, C., *et al.*, 1997. Termination of mammalian rDNA replication: Polar arrest of replication fork movement by transcription termination factor TTF-I. *Cell* 90 (3), 559–567.
- Goodfellow, S.J., Zomerijk, J.C., 2013. Basic mechanisms in RNA polymerase I transcription of the ribosomal RNA genes. *Subcellular Biochemistry* 61, 211–236.
- Goodpasture, C., Bloom, S.E., 1975. Visualization of nucleolar organizer regions in mammalian chromosomes using silver staining. *Chromosoma* 53 (1), 37–50.
- Govoni, M., Farabegoli, F., Pession, A., Novello, F., 1994. Inhibition of topoisomerase II activity and its effect on nucleolar structure and function. *Experimental Cell Research* 211 (1), 36–41.
- Granneman, S., Tollervey, D., 2007. Building ribosomes: Even more expensive than expected? *Current Biology* 17 (11), R415–R417.
- Grob, A., Collieran, C., McStay, B., 2014. Construction of synthetic nucleoli in human cells reveals how a major functional nuclear domain is formed and propagated through cell division. *Genes & Development* 28 (3), 220–230.
- Grummt, I., 2013. The nucleolus-guardian of cellular homeostasis and genome integrity. *Chromosoma* 122 (6), 487–497.
- Grummt, I., Langst, G., 2013. Epigenetic control of RNA polymerase I transcription in mammalian cells. *Biochimica et Biophysica Acta* 1829 (3–4), 393–404.
- Gurrieri, C., Capodice, P., Bernardi, R., *et al.*, 2004. Loss of the tumor suppressor PML in human cancers of multiple histologic origins. *Journal of the National Cancer Institute* 96 (4), 269–279.
- Hannan, K.M., Sanij, E., Rothblum, L.I., Hannan, R.D., Pearson, R.B., 2013. Dysregulation of RNA polymerase I transcription during disease. *Biochimica et Biophysica Acta* 1829 (3–4), 342–360.
- Hardwick, J.S., Kuruvilla, F.G., Tong, J.K., Shamji, A.F., Schreiber, S.L., 1999. Rapamycin-modulated transcription defines the subset of nutrient-sensitive signaling pathways directly controlled by the Tor proteins. *Proceedings of the National Academy of Sciences of the United States of America* 96 (26), 14866–14870.
- Heliot, L., Kaplan, H., Lucas, L., *et al.*, 1997. Electron tomography of metaphase nucleolar organizer regions: Evidence for a twisted-loop organization. *Molecular Biology of the Cell* 8 (11), 2199–2216.
- Henderson, A.S., Warburton, D., Atwood, K.C., 1972. Location of ribosomal DNA in the human chromosome complement. *Proceedings of the National Academy of Sciences of the United States of America* 69 (11), 3394–3398.
- Hernandez, P., Martín-Parras, L., Martínez-Robles, M.L., Schwartzman, J.B., 1993. Conserved features in the mode of replication of eukaryotic ribosomal RNA genes. *EMBO Journal* 12 (4), 1475–1485.
- Hernandez-Verdun, D., 2011. Assembly and disassembly of the nucleolus during the cell cycle. *Nucleus* 2 (3), 189–194.
- Hiscox, J.A., 2007. RNA viruses: Hijacking the dynamic nucleolus. *Nature Reviews Microbiology* 5 (2), 119–127.
- Hiscox, J.A., Whitehouse, A., Matthews, D.A., 2010. Nucleolar proteomics and viral infection. *Proteomics* 10 (22), 4077–4086.
- Ide, S., Miyazaki, T., Maki, H., Kobayashi, T., 2010. Abundance of ribosomal RNA gene copies maintains genome integrity. *Science* 327 (5966), 693–696.
- Jarboui, M.A., Bidoia, C., Woods, E., *et al.*, 2012. Nucleolar protein trafficking in response to HIV-1 Tat: Rewiring the nucleolus. *PLoS ONE* 7 (11), e48702.
- Jung, Y.S., Qian, Y., Chen, X., 2010. Examination of the expanding pathways for the regulation of p21 expression and activity. *Cell Signal* 22 (7), 1003–1012.
- Kar, B., Liu, B., Zhou, Z., Lam, Y.W., 2011. Quantitative nucleolar proteomics reveals nuclear re-organization during stress-induced senescence in mouse fibroblast. *BMC Cell Biology* 12, 33.
- Klein, J., Grummt, I., 1999. Cell cycle-dependent regulation of RNA polymerase I transcription: The nucleolar transcription factor UBF is inactive in mitosis and early G1. *Proceedings of the National Academy of Sciences of the United States of America* 96 (11), 6096–6101.
- Kobayashi, T., 2008. A new role of the rDNA and nucleolus in the nucleus — rDNA instability maintains genome integrity. *Bioessays* 30 (3), 267–272.
- van Koningsbruggen, S., Gierlinski, M., Schofield, P., *et al.*, 2010. High-resolution whole-genome sequencing reveals that specific chromatin domains from most human chromosomes associate with nucleoli. *Molecular Biology of the Cell* 21 (21), 3735–3748.
- Korobeynikova, A.V., Garber, M.B., Gongadze, G.M., 2012. Ribosomal proteins: Structure, function, and evolution. *Biochemistry (Mosc)* 77 (6), 562–574.
- Lam, Y.W., Evans, V.C., Heesom, K.J., Lamond, A.I., Matthews, D.A., 2010. Proteomics analysis of the nucleolus in adenovirus-infected cells. *Molecular & Cellular Proteomics* 9 (1), 117–130.
- Lamond, A.I., Spector, D.L., 2003. Nuclear speckles: A model for nuclear organelles. *Nature Reviews Molecular Cell Biology* 4 (8), 605–612.
- Langst, G., Becker, P.B., Grummt, I., 1998. TTF-I determines the chromatin architecture of the active rDNA promoter. *EMBO Journal* 17 (11), 3135–3145.
- Leung, A.K., Gerlich, D., Miller, G., *et al.*, 2004. Quantitative kinetic analysis of nucleolar breakdown and reassembly during mitosis in live human cells. *Journal of Cell Biology* 166 (6), 787–800.
- Leung, A.K., Lamond, A.I., 2002. In vivo analysis of NHPX reveals a novel nucleolar localization pathway involving a transient accumulation in splicing speckles. *Journal of Cell Biology* 157 (4), 615–629.
- Li, H., Tsang, C.K., Watkins, M., Bertram, P.G., Zheng, X.F., 2006. Nutrient regulates Tor1 nuclear localization and association with rDNA promoter. *Nature* 442 (7106), 1058–1061.
- Li, J., Langst, G., Grummt, I., 2006. NoRC-dependent nucleosome positioning silences rRNA genes. *EMBO Journal* 25 (24), 5735–5741.
- Liang, Y.M., Wang, X., Ramalingam, R., *et al.*, 2012. Novel nucleolar isolation method reveals rapid response of human nucleolar proteomes to serum stimulation. *Journal of Proteomics* 77, 521–530.
- Little, R.D., Platt, T.H., Schildkraut, C.L., 1993. Initiation and termination of DNA replication in human rRNA genes. *Molecular and Cellular Biology* 13 (10), 6600–6613.
- Liu, Q., Dreyfuss, G., 1996. A novel nuclear structure containing the survival of motor neurons protein. *EMBO Journal* 15 (14), 3555–3565.
- Loewith, R., Jacinto, E., Wulschleger, S., *et al.*, 2002. Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. *Molecular Cell* 10 (3), 457–468.
- Lohrum, M.A., Vousden, K.H., 1999. Regulation and activation of p53 and its family members. *Cell Death & Differentiation* 6 (12), 1162–1168.
- Luo, M.H., Rosenke, K., Czornak, K., Fortunato, E.A., 2007. Human cytomegalovirus disrupts both ataxia telangiectasia mutated protein (ATM)- and ATM-Rad3-related kinase-mediated DNA damage responses during lytic infection. *Journal of Virology* 81 (4), 1934–1950.
- Machyna, M., Heyn, P., Neugebauer, K.M., 2013. Cajal bodies: Where form meets function. *Wiley Interdisciplinary Reviews: RNA* 4 (1), 17–34.
- Mahajan, P.B., 1994. Modulation of transcription of rRNA genes by rapamycin. *International Journal of Immunopharmacology* 16 (9), 711–721.
- Malatesta, M., Zancanaro, C., Martin, T.E., *et al.*, 1994a. Cytochemical and immunocytochemical characterization of nuclear bodies during hibernation. *European Journal of Cell Biology* 65 (1), 82–93.
- Malatesta, M., Zancanaro, C., Martin, T.E., *et al.*, 1994b. Is the coiled body involved in nucleolar functions? *Experimental Cell Research* 211 (2), 415–419.
- Mao, Y.S., Zhang, B., Spector, D.L., 2011. Biogenesis and function of nuclear bodies. *Trends in Genetics* 27 (8), 295–306.
- Matera, A.G., Frey, M.R., 1998. Coiled bodies and gems: Janus or gemini? *American Journal of Human Genetics* 63 (2), 317–321.
- Mayer, C., Schmitz, K.M., Li, J., Grummt, I., Santoro, R., 2006. Intergenic transcripts regulate the epigenetic state of rRNA genes. *Molecular Cell* 22 (3), 351–361.
- McStay, B., 2006. Nucleolar dominance: A model for rRNA gene silencing. *Genes & Development* 20 (10), 1207–1214.
- McStay, B., Grummt, I., 2008. The epigenetics of rRNA genes: From molecular to chromosome biology. *Annual Review of Cell and Developmental Biology* 24, 131–157.
- Melese, T., Xue, Z., 1995. The nucleolus: An organelle formed by the act of building a ribosome. *Current Opinion in Cell Biology* 7 (3), 319–324.

- Michel, A.H., Kornmann, B., Dubrana, K., Shore, D., 2005. Spontaneous rDNA copy number variation modulates Sir2 levels and epigenetic gene silencing. *Genes & Development* 19 (10), 1199–1210.
- Monier, K., Armas, J.C., Etteldorf, S., Ghazal, P., Sullivan, K.F., 2000. Annexation of the interchromosomal space during viral infection. *Nature Cell Biology* 2 (9), 661–665.
- Moore, H.M., Bai, B., Boisvert, F.M., *et al.*, 2011. Quantitative proteomics and dynamic imaging of the nucleolus reveal distinct responses to UV and ionizing radiation. *Molecular & Cellular Proteomics* 10 (10), M111 009241.
- Murayama, A., Ohmori, K., Fujimura, A., *et al.*, 2008. Epigenetic control of rDNA loci in response to intracellular energy status. *Cell* 133 (4), 627–639.
- Nakagawa, S., Hirose, T., 2012. Paraspeckle nuclear bodies – useful uselessness? *Cellular and Molecular Life* 69 (18), 3027–3036.
- Narla, A., Ebert, B.L., 2010. Ribosomopathies: Human disorders of ribosome dysfunction. *Blood* 115 (16), 3196–3205.
- Németh, A., Conesa, A., Santoyo-Lopez, J., *et al.*, 2010. Initial genomics of the human nucleolus. *PLoS Genetics* 6 (3), e1000889.
- Németh, A., Langst, G., 2011. Genome organization in and around the nucleolus. *Trends in Genetics* 27 (4), 149–156.
- Németh, A., Perez-Fernandez, J., Merkl, P., *et al.*, 2013. RNA polymerase I termination: Where is the end? *Biochimica et Biophysica Acta* 1829 (3–4), 306–317.
- Ni, L., Wang, S., Zheng, C., 2012. The nucleolus and herpesviral usurpation. *Journal of Medical Microbiology* 61 (Pt 12), 1637–1643.
- Ochs, R.L., Stein Jr., T.W., Tan, E.M., 1994. Coiled bodies in the nucleolus of breast cancer cells. *Journal of Cell Science* 107 (Pt 2), 385–399.
- Olson, M.O., 2004. Sensing cellular stress: Another new function for the nucleolus? *Science's STKE* 2004 (224), pe10.
- Ong, S.E., Blagoev, B., Kratchmarova, I., *et al.*, 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & Cellular Proteomics* 1 (5), 376–386.
- Panse, V.G., Johnson, A.W., 2010. Maturation of eukaryotic ribosomes: Acquisition of functionality. *Trends in Biochemical Sciences* 35 (5), 260–266.
- Paredes, S., Maggert, K.A., 2009. Ribosomal DNA contributes to global chromatin regulation. *Proceedings of the National Academy of Sciences of the United States of America* 106 (42), 17829–17834.
- Patel, S.B., Bellini, M., 2008. The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Research* 36 (20), 6482–6493.
- Pederson, T., 1998. The plurifunctional nucleolus. *Nucleic Acids Research* 26 (17), 3871–3876.
- Pederson, T., Tsai, R.Y., 2009. In search of nonribosomal nucleolar protein function and regulation. *Journal of Cell Biology* 184 (6), 771–776.
- Pende, M., Um, S.H., Mieulet, V., *et al.*, 2004. S6K1(–/–)/S6K2(–/–) mice exhibit perinatal lethality and rapamycin-sensitive 5'-terminal oligopyrimidine mRNA translation and reveal a mitogen-activated protein kinase-dependent S6 kinase pathway. *Molecular and Cellular Biology* 24 (8), 3112–3124.
- Piekna-Przybylska, D., Decatur, W.A., Fournier, M.J., 2008. The 3D rRNA modification maps database: With interactive tools for ribosome analysis. *Nucleic Acids Research* 36 (Database issue), D178–D183.
- Powers, T., Walter, P., 1999. Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in *Saccharomyces cerevisiae*. *Molecular Biology of the Cell* 10 (4), 987–1000.
- Preuss, S., Pikaard, C.S., 2007. rRNA gene silencing and nucleolar dominance: Insights into a chromosome-scale epigenetic on/off switch. *Biochimica et Biophysica Acta* 1769 (5–6), 383–392.
- Raska, I., Shaw, P.J., Cmarko, D., 2006. Structure and function of the nucleolus in the spotlight. *Current Opinion in Cell Biology* 18 (3), 325–334.
- Regad, T., Chelbi-Alix, M.K., 2001. Role and fate of PML nuclear bodies in response to interferon and viral infections. *Oncogene* 20 (49), 7274–7286.
- Roussel, P., André, C., Comai, L., Hernandez-Verdun, D., 1996. The rDNA transcription machinery is assembled during mitosis in active NORs and absent in inactive NORs. *Journal of Cell Biology* 133 (2), 235–246.
- Roussel, P., André, C., Masson, C., Géraud, G., Hernandez-Verdun, D., 1993. Localization of the RNA polymerase I transcription factor hUBF during the cell cycle. *Journal of Cell Science* 104 (Pt 2), 327–337.
- Rubbi, C.P., Milner, J., 2003. Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *EMBO Journal* 22 (22), 6068–6077.
- Salvetti, A., Greco, A., 2014. Viruses and the nucleolus: The fatal attraction. *Biochimica et Biophysica Acta* 1842 (6), 840–847.
- Sanij, E., Poortinga, G., Sharkey, K., *et al.*, 2008. UBF levels determine the number of active ribosomal RNA genes in mammals. *Journal of Cell Biology* 183 (7), 1259–1274.
- Santoro, R., Li, J., Grummt, I., 2002. The nucleolar remodeling complex NoRC mediates heterochromatin formation and silencing of ribosomal gene transcription. *Nature Genetics* 32 (3), 393–396.
- Santoro, R., Schmitz, K.M., Sandoval, J., Grummt, I., 2010. Intergenic transcripts originating from a subclass of ribosomal DNA repeats silence ribosomal RNA genes in trans. *EMBO Reports* 11 (1), 52–58.
- Scheer, U., Rose, K.M., 1984. Localization of RNA polymerase I in interphase cells and mitotic chromosomes by light and electron microscopic immunocytochemistry. *Proceedings of the National Academy of Sciences of the United States of America* 81 (5), 1431–1435.
- Schlesinger, S., Selig, S., Bergman, Y., Cedar, H., 2009. Allelic inactivation of rDNA loci. *Genes & Development* 23 (20), 2437–2447.
- Schmitz, K.M., Mayer, C., Postepska, A., Grummt, I., 2010. Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes & Development* 24 (20), 2264–2269.
- Schmitz, K.M., Schmitt, N., Hoffmann-Rohrer, U., *et al.*, 2009. TAF12 recruits Gadd45a and the nucleotide excision repair complex to the promoter of rRNA genes leading to active DNA demethylation. *Molecular Cell* 33 (3), 344–353.
- Seeler, J.S., Dejean, A., 2001. SUMO: Of branched proteins and nuclear bodies. *Oncogene* 20 (49), 7243–7249.
- Shav-Tal, Y., Blechman, J., Darzacq, X., *et al.*, 2005. Dynamic sorting of nuclear components into distinct nucleolar caps during transcriptional inhibition. *Molecular Biology of the Cell* 16 (5), 2395–2413.
- Shaw, P., Brown, J., 2012. Nucleoli: Composition, function, and dynamics. *Plant Physiology* 158 (1), 44–51.
- Sherr, C.J., 2000. The Pezcoller lecture: Cancer cell cycles revisited. *Cancer Research* 60 (14), 3689–3695.
- Sherr, C.J., 2001. The INK4a/ARF network in tumor suppression. *Nature Reviews Molecular Cell Biology* 2 (10), 731–737.
- Sirri, V., Urcuqui-Inchima, S., Roussel, P., Hernandez-Verdun, D., 2008. Nucleolus: The fascinating nuclear body. *Histochemistry and Cell Biology* 129 (1), 13–31.
- Sirtori, C., Bosio-Bestetti, M., 1967. Nucleolar changes in KB tumor cells infected with herpes simplex virus. *Cancer Res* 27 (2), 367–376.
- Sleeman, J.E., Lamond, A.I., 1999. Newly assembled snRNPs associate with coiled bodies before speckles, suggesting a nuclear snRNP maturation pathway. *Current Biology* 9 (19), 1065–1074.
- Sleeman, J.E., Trinkle-Mulcahy, L., 2014. Nuclear bodies: New insights into assembly/dynamics and disease relevance. *Current Opinion in Cell Biology* 28C, 76–83.
- Spector, D.L., Lamond, A.I., 2011. Nuclear speckles. *Cold Spring Harbor Perspectives in Biology* 3 (2), a000646.
- Stanek, D., Neugebauer, K.M., 2006. The Cajal body: A meeting place for spliceosomal snRNPs in the nuclear maze. *Chromosoma* 115 (5), 343–354.
- Stefanovsky, V., Langlois, F., Gagnon-Kugler, T., Rothblum, L.I., Moss, T., 2006. Growth factor signaling regulates elongation of RNA polymerase I transcription in mammals via UBF phosphorylation and r-chromatin remodeling. *Molecular Cell* 21 (5), 629–639.
- Stoykova, A.S., Dabeva, M.D., Dimova, R.N., Hadjiolov, A.A., 1985. Ribosome biogenesis and nucleolar ultrastructure in neuronal and oligodendroglial rat brain cells. *Journal of Neurochemistry* 45 (6), 1667–1676.
- Strohner, R., Németh, A., Jansa, P., *et al.*, 2001. NoRC—a novel member of mammalian ISWI-containing chromatin remodeling machines. *EMBO Journal* 20 (17), 4892–4900.
- Sullivan, G.J., Bridger, J.M., Cuthbert, A.P., *et al.*, 2001. Human acrocentric chromosomes with transcriptionally silent nucleolar organizer regions associate with nucleoli. *EMBO Journal* 20 (11), 2867–2874.
- Tachibana, M., Sugimoto, K., Nozaki, M., *et al.*, 2002. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes & Development* 16 (14), 1779–1791.
- Tanaka, Y., Okamoto, K., Teye, K., *et al.*, 2010. JmjC enzyme KDM2A is a regulator of rRNA transcription in response to starvation. *EMBO Journal* 29 (9), 1510–1522.
- Tang, H., Hornstein, E., Stolovich, M., *et al.*, 2001. Amino acid-induced translation of TOP mRNAs is fully dependent on phosphatidylinositol 3-kinase-mediated signaling, is partially inhibited by rapamycin, and is independent of S6K1 and rpS6 phosphorylation. *Molecular and Cellular Biology* 21 (24), 8671–8683.
- Teng, T., Thomas, G., Mercer, C.A., 2013. Growth control and ribosomopathies. *Current Opinion in Genetics & Development* 23 (1), 63–71.
- Tsang, C.K., Bertram, P.G., Ai, W., Drenan, R., Zheng, X.F., 2003. Chromatin-mediated regulation of nucleolar structure and RNA Pol I localization by TOR. *EMBO Journal* 22 (22), 6045–6056.

- Tsang, C.K., Liu, H., Zheng, X.F., 2010. mTOR binds to the promoters of RNA polymerase I- and III-transcribed genes. *Cell Cycle* 9 (5), 953–957.
- Vaartjes, S.R., Boom, H.B., 1987. Left ventricular internal resistance and unloaded ejection flow assessed from pressure-flow relations: A flow-clamp study on isolated rabbit hearts. *Circulation Research* 60 (5), 727–737.
- Verheggen, C., Bertrand, E., 2012. CRM1 plays a nuclear role in transporting snoRNPs to nucleoli in higher eukaryotes. *Nucleus* 3 (2), 132–137.
- Vousden, K.H., 2002. Activation of the p53 tumor suppressor protein. *Biochimica et Biophysica Acta* 1602 (1), 47–59.
- Wang, L., Ren, X.M., Xing, J.J., Zheng, A.C., 2010. The nucleolus and viral infection. *Virologica Sinica* 25 (3), 151–157.
- Wang, Y., Tang, Q., Maul, G.G., Yuan, Y., 2006. Kaposi's sarcoma-associated herpesvirus ori-Lyt-dependent DNA replication: Dual role of replication and transcription activator. *Journal of Virology* 80 (24), 12171–12186.
- Warner, J.R., 1999. The economics of ribosome biosynthesis in yeast. *Trends in Biochemical Sciences* 24 (11), 437–440.
- Watkins, N.J., Bohnsack, M.T., 2012. The box C/D and H/ACA snoRNPs: Key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdisciplinary Reviews: RNA* 3 (3), 397–414.
- Weis, K., Rambaud, S., Lavau, C., *et al.*, 1994. Retinoic acid regulates aberrant nuclear localization of PML-RAR alpha in acute promyelocytic leukemia cells. *Cell* 76 (2), 345–356.
- Weisenberger, D., Scheer, U., 1995. A possible mechanism for the inhibition of ribosomal RNA gene transcription during mitosis. *Journal of Cell Biology* 129 (3), 561–575.
- Westman, B.J., Lamond, A.I., 2011. A role for SUMOylation in snoRNP biogenesis revealed by quantitative proteomics. *Nucleus* 2 (1), 30–37.
- Westman, B.J., Verheggen, C., Hutten, S., *et al.*, 2010. A proteomic screen for nucleolar SUMO targets shows SUMOylation modulates the function of Nop5/Nop58. *Molecular Cell* 39 (4), 618–631.
- White, R.E., Carline, L., Allday, M.J., 2007. Mutagenesis of the herpesvirus saimiri terminal repeat region reveals important elements for virus production. *Journal of Virology* 81 (12), 6765–6770.
- Wiesendanger, B., Lucchini, R., Koller, T., Sogo, J.M., 1994. Replication fork barriers in the *Xenopus* rDNA. *Nucleic Acids Research* 22 (23), 5038–5046.
- Xu, J., Zhou, X., Wang, J., *et al.*, 2013. RhoGAPs attenuate cell proliferation by direct interaction with p53 tetramerization domain. *Cell Reports* 3 (5), 1526–1538.
- Yamada, K., Ono, M., Bensaddek, D., Lamond, A.I., Rocha, S., 2013. FMN2 is a novel regulator of the cyclin-dependent kinase inhibitor p21. *Cell Cycle* 12 (15), 2348–2354.
- Yuan, X., Feng, W., Imhof, A., Grummt, I., Zhou, Y., 2007. Activation of RNA polymerase I transcription by cockayne syndrome group B protein and histone methyltransferase G9a. *Molecular Cell* 27 (4), 585–595.
- Yusupova, G., Yusupov, M., 2014. High-Resolution Structure of the Eukaryotic 80S Ribosome. *Annual Review of Biochemistry* 83, 467–486.
- Zhong, S., Salomoni, P., Pandolfi, P.P., 2000. The transcriptional role of PML and the nuclear body. *Nature Cell Biology* 2 (5), E85–E90.
- Zhou, Y., Santoro, R., Grummt, I., 2002. The chromatin remodeling complex NoRC targets HDAC1 to the ribosomal gene promoter and represses RNA polymerase I transcription. *EMBO Journal* 21 (17), 4632–4640.

Further Reading

Olson, M.O.J. (Ed.), 2011. *The Nucleolus*. New York: Springer. ISBN: 978-1-4614-0514-6.

Relevant Websites

http://www.lamondlab.com/newwebsite/Cellular_fractionation_protocols.php
Lamond Laboratory.

<http://www.lamondlab.com/newwebsite/protocols.php?id=0>
Lamond Laboratory.

<http://www.lamondlab.com/NOPdb3.0/>
Lamond Laboratory.

<http://www.peptracker.com/epd>
PepTracker.

<http://www.compbio.dundee.ac.uk/www-nod/>
University of Dundee.

<http://people.biochem.umass.edu/fournierlab/3dmodmap/main.php>
University of Massachusetts Amherst.

Ribosomes and Stress – Linked from Birth to Death

JRP Knight, T Sbarrato, M Stoneley, and AE Willis, University of Leicester, Leicester, UK

© 2016 Elsevier Inc. All rights reserved.

Glossary

Autophagy The regulated destruction of organelles within a cell, either following damage or to recycle biomolecules.

Cell stress The damage which occurs to cell components from external or internal factors. Cells have mechanisms to detect stress, signaling pathways to relay this detection and processes to deal with the stress.

Ribosome biogenesis The process by which ribosomes are synthesized from their constituent parts, namely, the

association of ribosomal proteins and ribosomal RNAs. The process also requires multiple non-ribosomal components called ribosome biogenesis factors.

Translation The process of making proteins, which is carried out by the mature ribosome. The term translation is used due to the need to change the genetic code in the form of mRNA into a protein sequence.

Introduction

Ribosome stress is a broad term applied to interactions between the cellular stress response and elements of the ribosome. Ribosome stress encompasses nucleolar stress – nucleoli are after all where ribosomes are made – but also pertains to stress in the context of translation by the ribosome. Despite spatial separation between nucleolar pre-ribosomes and cytoplasmic ribosomes, there are similarities in their regulation in response to stress. Cellular stress comes in many forms, from damage of cell components to reduced levels of nutrients. Genetic aberrations are also a stress, which may lead to the altered expression of genes and potentially oncogenic consequences. Importantly, ribosome stress has to be considered a two-way interaction where nucleoli and ribosomes also create and detect stress.

The mechanical activity of the ribosome is one of the most energy consuming activities in the cell. In proliferating cells the proportion of energy consumed by translation is second to only one other activity – the energy demanded for the synthesis of ribosomes themselves. This places ribosomes at the forefront of cellular energy management as their suppression (either at synthesis, during translation or both) will rapidly reduce energy expenditure. This article primarily follows the life of the mammalian ribosome (with other organisms referenced appropriately), from synthesis in nucleoli, function in the cytoplasm and finally considers ribosome degradation, showing that ribosomes are linked to stress sensing and signaling at each stage ([Figure 1](#)).

Ribosome Birth in the Nucleolus

Ribosomes consist of two subunits assembled from their constituent parts – 4 ribosomal RNAs (rRNAs) and 80 ribosomal proteins (RPs). Ribosomes are made in the nucleolus, a protein-rich subnuclear compartment which forms upon transcribed rDNA. RNA polymerase I transcribes rDNA, making an rRNA precursor that contributes to both subunits. This ensures stoichiometric synthesis of 40S and 60S subunits, but requires the activity of hundreds of auxiliary ribosome

biogenesis (RiBi) factors to cleave and modify the rRNA. As well as RNA polymerase I, ribosome biogenesis requires RNA polymerase II for synthesis of RP and RiBi protein mRNAs and RNA polymerase III for 5S rRNA and snoRNAs (a form of RiBi consisting of RNA). This makes ribosome biogenesis the ideal process to sense transcriptional health. When we also consider that the production of RPs and RiBis requires production of protein, and that nucleoli contain DNA, the nucleolus is an obvious choice to detect cell damaging stress.

Stress Detection by the Nucleolus – p53 Activation

Early studies with mouse models provided evidence for a p53 checkpoint in ribosome biogenesis, with liver-specific RPS6 knock-down resulting in cell cycle arrest ([Volarević et al., 2000](#)). Acting as a stress-responsive hub, the transcription factor p53 activates genes involved in cell cycle regulation, apoptosis, DNA repair and senescence. P53 is regulated primarily by one of its target gene products, human double minute 2 (HDM2, called MDM2 in mouse). HDM2 is an E3 ubiquitin ligase that poly-ubiquitinylates its targets, including p53. This auto-regulatory feedback loop occurs via two mechanisms: 1) inhibition of the p53 transactivation domain and 2) decreased p53 stability through ubiquitinylation and proteasomal degradation (reviewed by [Zhang and Lu, 2009](#)).

Further evidence of a link between ribosome biogenesis and p53 was displayed by dominant negative mutants of the RiBi Bop1, which resulted in both inhibition of ribosome biogenesis and p53-dependent cell cycle arrest ([Pestov et al., 2001](#)). This was supplemented by the finding that p53 activation requires nucleolar disruption caused by a multitude of stresses ([Rubbi and Milner, 2003](#)). Importantly, both 40S and 60S ribosome subunit maturation pathways can signal stress to p53 independently, as evidenced by superinduction following perturbation of both pathways ([Fumagalli et al., 2012](#)). It is not surprising that tight and sensitive surveillance is implemented due to the necessity for ribosome function within the cell.

Ribosomal proteins

Nucleolar stress releases RPs into the nucleoplasm where a subset can interact with HDM2 to abrogate p53

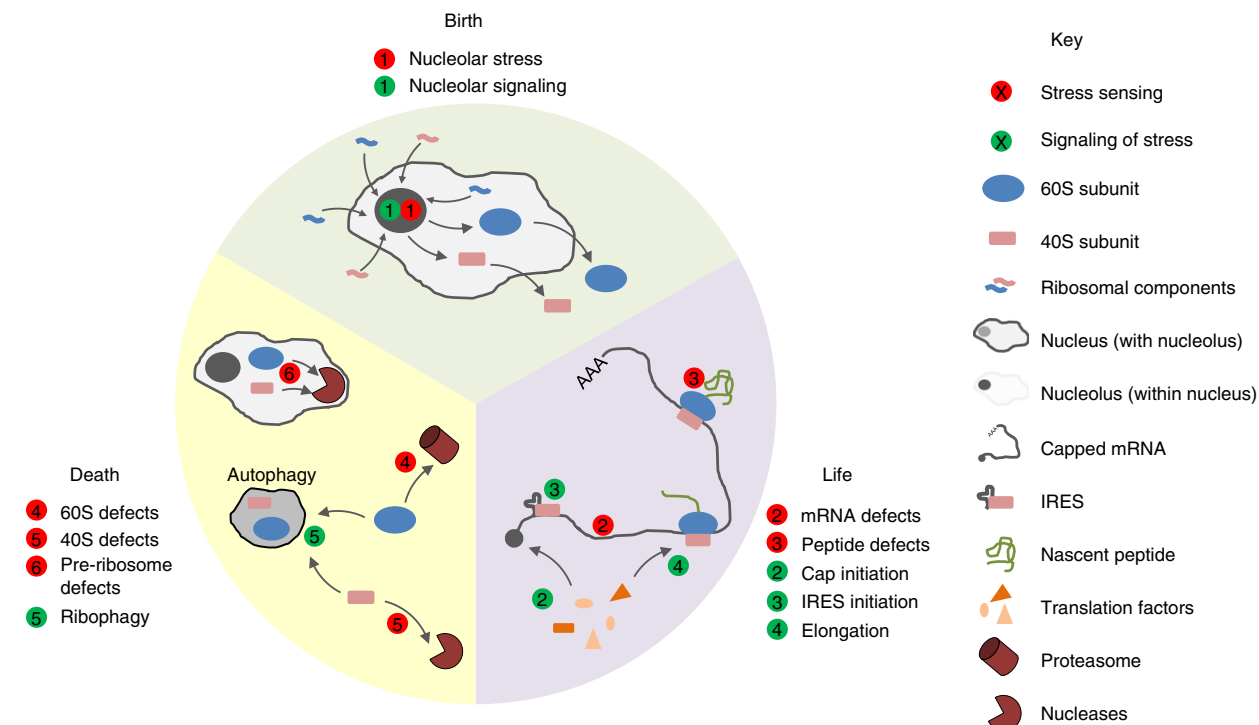


Figure 1 The life of the ribosome and its links to stress. Birth: ribosomes are born in the nucleolus, where the major rRNA precursor is synthesized and assembled with RPs. A further rRNA is imported from the nucleoplasm. Two subunits are synthesized by independent pathways, releasing mature 40S and 60S subunits into the cytoplasm to begin their life as ribosomes. Nucleoli are capable of both sensing stress and also responding to stress signaling. Life: Translating ribosomes are regulated by stress signaling during initiation and elongation, but also serve an important role in surveillance of mRNA and protein quality. Death: ribosomes, or their precursors, have to be degraded if they are synthesized with faults or gain aberrations during their life. These defects are sensed and dealt with by different degradation mechanisms – pre-ribosome degradation is driven by nucleases, as is mature 40S subunit degradation. Whereas mature 60S subunit degradation is driven by the ubiquitin/proteasome system. Stress signaling can also promote degradation of ribosomes not harboring defects via ribophagy.

ubiquitinylation and degradation. The first evidence that RPs participate in the regulation of HDM2 came from the report of the association of RPL5 with HDM2 in a 5S rRNA-RPL5-HDM2-p53 complex (Marechal *et al.*, 1994). Subsequently, it was shown that upon nucleotoxic stress or serum starvation, RPL5, RPL11, and RPL23 can bind to HDM2 leading to the accumulation and activation of p53. Since then similar negative effects upon HDM2 have been found for a total of 14 RPs, summarized in Table 1.

With 14 of the 80 RPs binding HDM2 there is high redundancy in the system. Interestingly, these RPs are partially clustered within the ribosome: near the 5S rRNA and exit channel in 60S subunits and either side of the mRNA decoding center in the 40S subunit (Figure 2), but are required throughout the different stages of ribosome biogenesis (Table 1). Redundancy may be explained by different binding sites on HDM2 and mechanisms of regulation. Domain mutation experiments have revealed that RPL5, RPL11 and RPL23 all bind the central domain of HDM2 in a different fashion (Dai *et al.*, 2004). Similarly, the ability to inhibit HDM2's ubiquitin ligase activity varies across these RPs; over-expression of RPL11 fails to suppress p53 poly-ubiquitinylation to the same extent as RPL5 or RPL23 (Dai *et al.*, 2006). Moreover, although all are capable of stabilizing p53 via HDM2, over-expression of RPL37, RPS15 or RPS20 led to the

Table 1 Stress detection by ribosomal proteins

Protein	Subunit	Biogenesis involvement	P53/HDM2 dependent
RPL5	60S	Late	Y
RPL6	60S	–	Y
RPL11	60S	Late	Y
RPL23	60S	–	Y
RPL26	60S	Early	Partial
RPL37	60S	–	Y
RPS3	40S	Early	Y
RPS6	40S	Early	Partial
RPS7	40S	Early	Y
RPS9	40S	Early	N
RPS14	40S	Early	Y
RPS15	40S	Late	Y
RPS20	40S	Late	Y
RPS25	40S	Late	Y
RPS26	40S	Late	Y
RPS27	40S	Early	Y

Note: Human ribosomal protein names are given for those with signaling capabilities. These are grouped by subunit. The role of each protein in ribosome biogenesis of their respective subunits is given, denoted as either early or late stage. The dependence of the RPs on p53/HDM2 for signaling is indicated (Y=yes; N=no).

Relevant reviews: Zhang, Y., Lu, H., 2009. Signaling to p53: Ribosomal proteins find their way. *Cancer Cell* 16, 369–377; Deisenroth, C., Zhang, Y., 2010. Ribosome biogenesis surveillance: Probing the ribosomal protein-Mdm2-p53 pathway. *Oncogene* 29, 4253–4260.

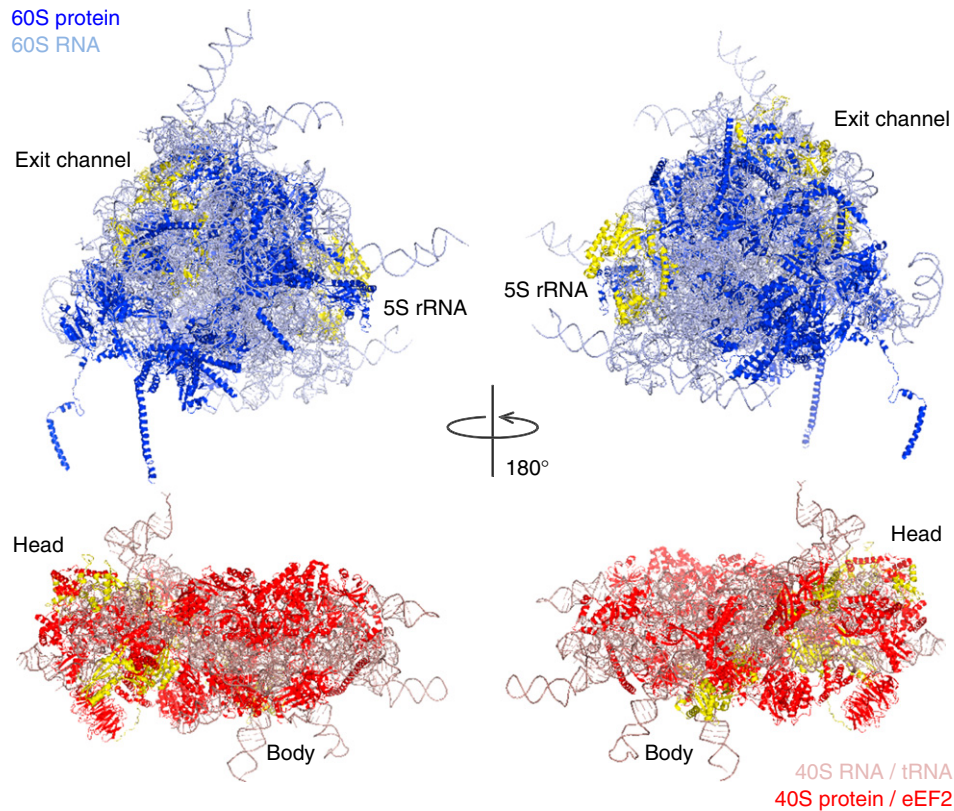


Figure 2 The location of HDM2 binding RPs in the ribosome. The human 80S ribosome is represented, with the two subunits shown individually from two opposite perspectives (Anger *et al.*, 2013). Ribosomal proteins highlighted in yellow on both subunits are able to bind HDM2 when not incorporated into the mature ribosome. These are clustered in the 60S subunit around the 5S rRNA and nascent peptide exit channel, and also in two areas the 40S subunit either side of the mRNA decoding center in the head and body.

induction of different p53 target genes, suggesting that the role of each RP is different (Daftuar *et al.*, 2013).

RPL5 and RPL11 have been identified as the master ribosomal regulators of HDM2/p53. Knockdown of either protein failed to activate p53 upon ribosome stress (Dai *et al.*, 2004; Jin *et al.*, 2004). It is possible that RPs act cooperatively upon HDM2, which has been demonstrated for RPL5 and RPL11 (Horn and Vousden, 2008). This may allow integration of defects at different stages of ribosome biogenesis (Table 1). It is also possible that redundancy may compensate for RP misregulation. We also have to consider the cumulative effect of these RP-HDM2 interactions – interestingly, under normal conditions RP expression may be restrained so that the RP-HDM2 response is prevented (Lam *et al.*, 2007). Upon stress, increased RPs in the nucleoplasm (either released from the nucleolus or aberrantly over-synthesized RPs imported from cytoplasm) can tip the balance and trigger the RP-HDM2 response. The RP-HDM2 pathway is summarized in Figure 3.

Nucleolar RNAs

RPs are not alone in signaling to p53; ribosomal RNAs also signal to HDM2. The 5S rRNA is involved in a pre-ribosomal complex together with RPL5 and RPL11, and is directly linked to HDM2 inhibition (Marechal *et al.*, 1994; Donati *et al.*, 2013). The 5S rRNA is critical within this complex and essential for p53 stabilization (Sloan *et al.*, 2013a). The MDM2

homolog MDMX, which heterodimerises with and enhances MDM2-mediated ubiquitinylation, can also be targeted by the 5S rRNA, leading to its rapid destabilization (Li and Gu, 2011). Furthermore, noncoding RNAs transcribed by RNA polymerase II from intergenic regions of rDNA have been shown to act as ‘sponges’ for factors containing a nucleolar detention signal (NoDS), including HDM2 (Audas *et al.*, 2012).

RiBi factors

RiBi factors respond to nucleolar stress and activate p53. Nucleolar multifunctional protein nucleophosmin (NPM), which is involved in the shuttling of factors for ribosome biogenesis, has a complex relationship with HDM2. Following stress, NPM relocates to the nucleoplasm and binds to HDM2 causing p53 activation (Kurki *et al.*, 2004; Itahana *et al.*, 2003). NPM can also enhance p53 target gene transactivation activity under stress conditions (Drygin *et al.*, 2009). However, in the absence of stress NPM promotes HDM2 activity by binding and sequestering the HDM2-suppressor p14^{ARF} within the nucleolus (Korgaonkar *et al.*, 2005; Colombo *et al.*, 2005).

P53-Independent Ribosome Stress

The importance of p53 in the response to ribosome stress is clear, but ribosome biogenesis is also monitored independently of p53. The yeast *Saccharomyces cerevisiae* lacks a p53

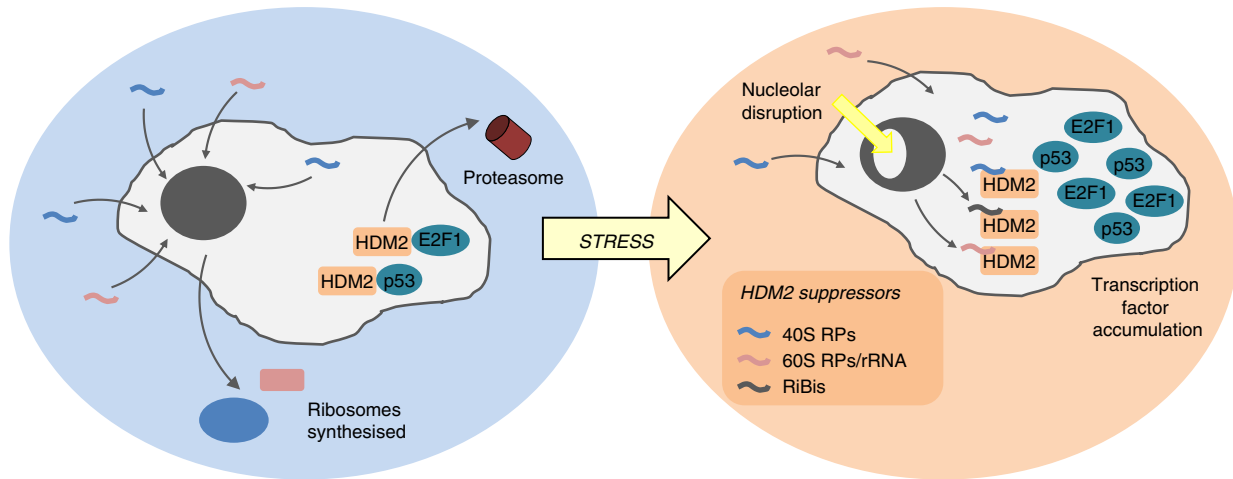


Figure 3 HDM2 suppression by the nucleolus during stress. Under normal conditions nucleoli synthesize ribosomes from their precursors. Simultaneously the HDM2/p53 axis suppresses p53, and other HDM2 targets via the ubiquitin/proteasome system. Upon stress, the nucleolus can become physically disrupted, correlating with reduced ribosome synthesis and alterations in the nucleolar proteome. Release of factors from the nucleolus can suppress HDM2 binding to its targets, allowing accumulation of transcription factors such as p53, and induction of a stress response.

homolog, but ribosome stress induced by defective ribosome biogenesis is able to disrupt the cell cycle (Bernstein *et al.*, 2007). Similarly, in *Drosophila melanogaster* where p53 has not been shown to participate in cell cycle regulation, the Minute phenotypes caused by RP haploinsufficiency are characterized by reduced proliferation linked to cell cycle regulation. In mammalian cells, somatic mutations and inactivation of p53 is the most common form of transformation occurring in cancers. As such, HeLa cells, bearing a wild-type p53 inactivated by the human papillomavirus, respond to nucleolar stress with an arrest in cell cycle and/or apoptosis (Lindström and Nistér, 2010). Also, in RP mutant mouse models, not all phenotypes are explained by p53 activation (Terzian and Box, 2013), although the ultimate test of this would be RP mutation in the context of a p53 null background.

The first confirmation in mammalian cells of a p53-independent response to ribosome stress came from the treatment of p53 null cells with actinomycin D. Inhibition of rRNA synthesis causes inhibition of HDM2 by RPL11, followed by stabilization of another HDM2 target: the transcription factor and cell cycle regulator E2F1 (Donati *et al.*, 2011). The E2F1 protein controls cell cycle progression from G1 to S phase via modulation of transcriptional targets. In addition, defects in ribosome biogenesis by RNA polymerase I inhibition or RP deficiency (RPS19, RPS6, RPL7a knock downs) decrease the expression of the proto-oncogene kinase PIM1, leading to cell cycle arrest (Iadevaia *et al.*, 2010). Upon ribosome stress, decreased PIM1 expression relieves the inhibition and degradation of cell cycle regulator p27^{Kip1}.

RPL11 also has further HDM2-independent effects, acting via the regulation of the oncogene c-Myc. C-Myc is a transcription factor involved in driving proliferation, cell growth, differentiation and stem cell renewal. RPL11 was shown to compete for binding to c-Myc transcriptional targets with its activator TRRAP (Dai *et al.*, 2007). In addition to inhibiting transcription of c-Myc targets, RPL11 can also interact with the

3'UTR of c-MYC mRNA and mediate a microRNA repression of c-Myc translation under stress (Challagundla *et al.*, 2011).

Nucleolar Signaling During Stress

The severity of stress signaling that can emanate from the nucleolus is testament to the requirement for fidelity in ribosome biogenesis. However, as well as being a cause of stress, cells use the nucleolus as part of the response to stress, mediated by dynamic localization of nucleolar proteins and multiple stress signaling pathways.

Nucleolar proteome dynamics

Mass spectrometry has proved an invaluable tool for the identification of over 700 proteins that reside in the nucleolus of cells. Importantly, localization of many of these proteins is dynamic following stress. Under normal conditions nucleoli are organized into sub-regions, which are disrupted following stress, and coupled to proteins leaving or entering the nucleolus (Andersen *et al.*, 2005). These protein dynamics signal physical nucleolar disruption to initiate a stress response. As discussed above, the exodus of RPs from the nucleolus to the nucleoplasm exemplifies these dynamics (Figure 3).

The classic example of a dynamic nucleolar protein is the tumor suppressor p14^{ARF}. P14^{ARF} promotes the accumulation of p53 and suppresses proliferation by promoting degradation of the RiBi NPM (Pomerantz *et al.*, 1998; Itahana *et al.*, 2003). Under normal conditions p14^{ARF} is an unstable protein which resides in the nucleolus (Rodway *et al.*, 2004). Activation of p14^{ARF} results in a physical interaction with HDM2, with subsequent localization apparently dependent upon the dose of p14^{ARF} but eliciting the same activation of p53. Low level induction localizes a fraction of p14^{ARF} to the nucleoplasm, which retains HDM2 in the nucleus but is subsequently degraded (Llanos *et al.*, 2001), whereas high levels of p14^{ARF}

cause nucleolar sequestration of HDM2 (Weber *et al.*, 1999; Tao and Levine, 1999).

Another example is the localization of the nucleolar protein SENP3 into the nucleoplasm following UV irradiation (Moore *et al.*, 2011). SENP3 functions as a deSUMOylation protease, which has been implicated in promoting the nucleolar localization of a 60S biogenesis complex, among other targets (Finkbeiner *et al.*, 2011). Nucleolar localization of this SENP3-regulated complex is lost with very similar kinetics to SENP3, suggesting that deSUMOylation promotes its nucleolar localization (Moore *et al.*, 2011). Interestingly, SENP3 also plays a part in the p14^{ARF} stress response, being targeted for degradation by a mechanism initiated by p14^{ARF}, and involving association with NPM (Kuo *et al.*, 2008). SENP3 degradation subsequently enhances overall SUMOylation levels, which allows p14^{ARF} to induce cell cycle arrest independent of p53.

Kinase signaling in the nucleolus

RNA polymerase I requires association with specialized transcription factors to synthesize rRNA. Suppression of ribosome biogenesis by targeting transcription ensures the maximal energy saving to the cell during stress. It is for this reason that a number of stress-responsive pathways directly target RNA polymerase I via its transcription initiation factor TIF-IA (Figure 4).

The c-Jun N-terminal kinases (JNKs) are ubiquitously expressed stress-activated protein kinases responsive to cell damage, managing the decision between cell survival and apoptosis. For example, the JNK pathway can promote p14^{ARF} nucleoplasmic relocation following genotoxic stress, to promote activation of p53 (Yogev *et al.*, 2008). The JNKs also modulate

ribosome biogenesis, with JNK2 directly phosphorylating TIF-IA in response to stress (Mayer *et al.*, 2005); phosphorylation of Thr²⁰⁰ impairs transcription machinery assembly, reducing rRNA synthesis. Phosphorylation by JNK2 also causes loss of nucleolar TIF-IA, which accumulates in the nucleoplasm, further perturbing rRNA synthesis (Mayer *et al.*, 2005).

The AMP-activated protein kinase (AMPK) senses energy availability via allosteric binding of ATP and ADP/AMP according to their intracellular ratio. Decreased ATP is indicative of reduced cellular energy production or overactive energy expenditure. AMPK promotes ATP production via phosphorylation of metabolic enzymes but also reduces ATP use, in part via the nucleolar stress response. Similar to JNK2, AMPK phosphorylates TIF-IA, disrupting interactions required for transcription initiation (Hoppe *et al.*, 2009). However, phosphorylation occurs at a different residue (Ser⁶³⁵) and has not been reported to affect TIF-IA localization.

The mammalian target of rapamycin (mTOR) protein forms two distinct complexes able to influence cell fate via kinase activity. The mTORC1 complex maintains phosphorylation of its targets when conditions are favorable for cell growth and proliferation. Upon stress mTORC1 activity is reduced, resulting in target hypo-phosphorylation. The physiological outcome of this phosphorylation switch is a reduction in energy expenditure. Suppression of mTORC1 by rapamycin modulates phosphorylation of TIF-IA, impairing the interaction with RNA polymerase I and causing localization to the cytoplasm (Mayer *et al.*, 2004). mTORC1 has roles in ribosome biogenesis beyond TIF-IA. mTORC1 nucleolar localization also promotes rRNA maturation, an effect which is lost following rapamycin-mediated inhibition of the kinase (Iadevaia *et al.*, 2012). Furthermore, S6K (a downstream target of mTORC1) is required for

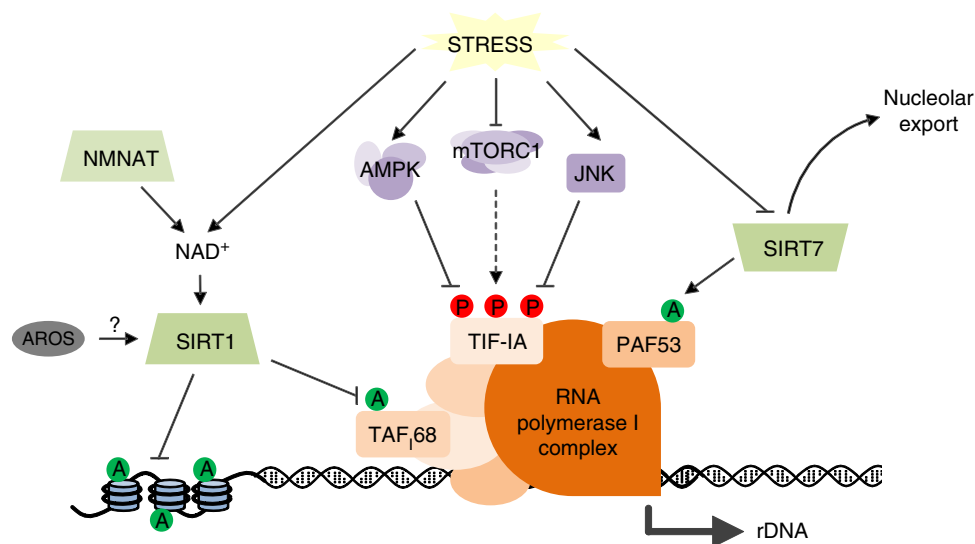


Figure 4 Stress signaling to RNA polymerase I. Three transcription factors of RNA polymerase I are subject to regulation by posttranslational modification following stress. TIF-IA is phosphorylated on multiple residues by stress signaling kinase pathways. Phosphorylation results in suppression of rRNA synthesis via disruption of the RNA polymerase I initiation complex. SIRT7 promotes rRNA synthesis performing the deacetylation of PAF53 in a constitutive cycle of acetylation required for transcription. In response to stress SIRT7 nucleolar localization is lost, disrupting this acetylation cycle. SIRT1 suppresses rRNA synthesis via reduced RNA polymerase I initiation through deacetylation of TAF₆₈, and also via epigenetic silencing of rDNA. This responds to NAD⁺ availability, which is increased by reduction in cellular energy levels. We also note that NAD⁺ may be locally increased by NMNAT in the nucleolus and that the SIRT1 allosteric activator AROS is a RiBi which has an increased effect on SIRT1 during stress. P=phosphorylation; A=acetylation. Dashed line indicates an indirect mechanism.

optimal transcription of over 100 RiBi mRNAs in a mouse model of nutrient stress (Chauvin *et al.*, 2014).

Nucleolar sirtuins

SIRT1 is an important factor in managing stress, in part via modulation of ribosome biogenesis. SIRT1 activity is metabolically responsive due to dependence upon NAD^+ as a cofactor. NAD^+ levels increase during nutrient stress, resulting in an increase in SIRT1 activity that enhances mitochondria biogenesis, promoting energy generation (Canto *et al.*, 2009; Rodgers *et al.*, 2005). Supplementary to this, SIRT1 also regulates rDNA epigenetic availability under nutrient stress, acting in a suppressor complex with the methyl-transferase SUV39H1 to reduce rRNA transcription when NAD^+ levels are high (Murayama *et al.*, 2008). SIRT1 also reverses acetylation of the transcription factor component TAF₆₈, reducing rDNA transcription and suppressing ribosome biogenesis (Muth *et al.*, 2001). Active regulator of SIRT1 (AROS) is an allosteric activator of SIRT1, with enhanced activity towards SIRT1 following stress (Kim *et al.*, 2007; Knight *et al.*, 2013a). Interestingly, AROS is also a RiBi required for 40S biogenesis (Knight *et al.*, 2013b), presenting the possibility that SIRT1 activity can be modulated following perturbations in ribosome biogenesis via signaling from AROS.

In contrast to SIRT1, SIRT7 promotes ribosome biogenesis, acting directly upon the transcription machinery. SIRT7 participates in an acetylation/deacetylation cycle of the RNA polymerase I recruitment factor PAF53, with deacetylation promoting rDNA binding (Chen *et al.*, 2013). High SIRT7 expression correlates with proliferation, likely via SIRT7-driven rRNA production (Ford *et al.*, 2006). The nucleolar localization of SIRT7 is dependent upon the presence of rRNA, with reduced rRNA synthesis causing SIRT7 to localize to the nucleoplasm (Chen *et al.*, 2013). This allows continued acetylation of PAF53 to suppress rRNA transcription. Together this suggests opposing roles for SIRT1 and SIRT7, both of which are responsive to NAD^+ . Epigenetic suppression of rDNA by SIRT1 would presumably negate any potential effects of SIRT7, which requires access to rRNA to elicit its function. Interestingly, the NAD^+ synthesis enzyme NMNAT localizes to nucleoli and leads to rRNA suppression, consistent with the role of SIRT1 (Song *et al.*, 2013). However, it is possible that both sirtuins are equally responsive – SIRT1 suppresses rDNA under conditions of nutrient stress, whereas SIRT7 is only nucleolar in the absence of stress – suggesting a resolution explained by stress exposure. The role of sirtuins in signaling to RNA polymerase I is summarized in Figure 4.

The Busy Life of the Ribosome

The primary function of the mature ribosome is translation, which like all jobs is highly regulated. This is carried out by the ribosome-associating translation factors; however, the activity of ribosomes themselves should not be overlooked. The capabilities of the ribosome that are defined during their biogenesis are essential for modulation of both global and specific protein expression. This is especially evident following stress. Adding to their busy workload further is an important role for ribosomes in quality control of both proteins and mRNAs.

Stress Signaling to the Ribosome

Cap-dependent translation

mRNAs are modified by the addition of methylated guanine to their 5' ends, creating the m⁷G cap. These caps are bound by the cap-binding protein eIF4E, which in turn binds a scaffold protein eIF4G and an RNA helicase eIF4A. This complex (called eIF4F) recruits 40S subunits to mRNA via eIF4G's association with further initiation factors, notably the multimeric eIF3 complex in a process termed cap-dependent initiation. Each physical interaction requires two parties; the 40S subunit has coevolved with the cap-dependent initiation factors to facilitate recruitment. In yeast, the 40S uses its rRNA and RPS0A (mammalian RPSA) to bind eIF3 (Valášek *et al.*, 2003).

EIF4E binding-proteins (4EBPs) are important regulators of cap-dependent initiation. In their hypo-phosphorylated state the 4EBPs compete with eIF4G for association with eIF4E. Interaction between 4EBPs or eIF4G with eIF4E is mutually exclusive such that the 4EBPs disrupts the eIF4F complex, decreasing 40S recruitment to mRNA and therefore the global rate of translation. Nutrient deprivation, hypoxia, and double stranded DNA lesions lead to suppression of mTORC1 activity, hypo-phosphorylation of the 4EBPs and diminished cap-dependent translation – see Figure 5 (reviewed by Spriggs *et al.*, 2010). Heat shock also suppresses cap-dependent translation in part via mTORC1 but also via the sequestration of eIF4G to stress granules.

To allow control by cap-dependent initiation the ribosome must be a poor initiator without assistance from eIF4F – the quantity of 40S subunits and mRNAs within the cell is such that the two will interact, yet the ribosome does not initiate on the majority of mRNAs. However, cap-independent initiation is possible on a subset of mRNAs, which contain an internal ribosome entry site (IRES) (Jang *et al.*, 1988). IRESs are structured RNA sequences able to initiate translation of their message without requirement for all of the canonical initiation factors and may be present in up to 10% of human mRNAs (Johannes *et al.*, 1999; Mitchell *et al.*, 2005). For all IRESs the need for cap-binding by eIF4E is removed, allowing expression when cap-dependent translation is suppressed, such as during stress.

40S subunits are active in recruitment of IRES containing messages since suppression of 40S subunit abundance results in reduced IRES initiation, beyond the suppression of general translation (Huang *et al.*, 2012). Furthermore, the 40S protein RPS25 is required for IRES initiation as 40S subunits lacking the protein are unable to bind IRES mRNAs (Landry *et al.*, 2009). The effect may extend to overall ribosome integrity as two ribosome biogenesis factors involved in rRNA modification – dyskerin and fibrillarin – are each required for efficient initiation from IRESs (Yoon *et al.*, 2006; Marcel *et al.*, 2013). These examples illustrate the reduced translational repertoires of aberrant ribosomes that have escaped nucleolar surveillance, an effect which will be most severe when attempting to respond to stress (Figure 5).

Regulation of elongation

The energy used in translation is predominantly consumed during elongation – 2 molecules of GTP are hydrolyzed for each amino acid added to the peptide chain. Regulation can occur during elongation via phosphorylation of eEF2, one of these

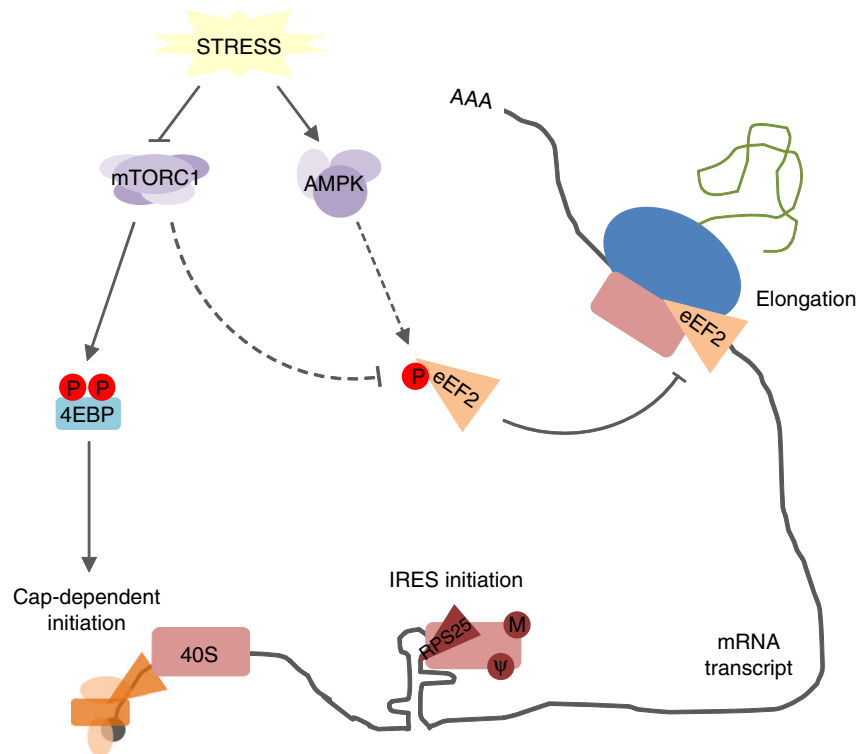


Figure 5 Stress signaling to the translating ribosome. Translation is suppressed following stress via regulation of initiation and elongation. mTORC1 promotes cap-dependent translation by suppressive phosphorylation of 4EBPs. Under stress this activity is reduced, suppressing translation initiating from the mRNA cap. mTORC1 suppression also results in phosphorylation of eEF2, which suppresses elongation by the ribosome. Likewise, AMPK also promotes eEF2 phosphorylation following the detection of stress. The ability of ribosomes to initiate upon internal ribosome entry sites (IRESs) is dependent upon fully competent 40S subunits being synthesized, including correct incorporation of RPS25 and modification of rRNAs by dyskerin (ψ) and fibrillarin (M). P=phosphorylation. Dashed lines indicate an indirect mechanism.

GTPases. The ribosome requires eEF2 to fuel movement along a message, expelling the used tRNA and allowing decoding of the next codon. However, the ribosome senses a single phosphorylation at Thr⁵⁶, which occludes eEF2 from entering the ribosome. This phosphorylation is carried out by eEF2K, which is activated following stress to reduce eEF2 activity. For example, the mTORC1 pathway induces eEF2 phosphorylation via S6K/eEF2K following nutrient deprivation and hypoxia (Wang *et al.*, 2001). Similarly, reduced ATP availability is sensed by AMPK, which signals to eEF2K to conserve energy expenditure by suppressing elongation (Horman *et al.*, 2002). In both cases occlusion of phosphorylated eEF2 is essential for translational suppression (Figure 5). Interestingly the regulation of elongation has strong parallels with the regulation of rRNA transcription, sharing two of the same modulating kinases. Thus, activation of AMPK and suppression of mTORC1 has common effects upon ribosome biogenesis in the nucleolus and cytoplasmic ribosome function.

Stress Detection by the Ribosome

Ensuring the integrity of the proteome is essential for overall cell function. Protein fidelity is promoted by chaperones, which assist protein folding, and by mechanisms to degrade faulty proteins. Translation by the ribosome acts as the perfect place for quality control assessment for both nascent mRNA

and proteins. These functions are split between the 40S (mRNA) and 60S (protein) subunits.

Protein quality control

Chaperones promote correct protein folding and are required under non-stressed conditions but also following proteotoxic stress – stress that effects protein integrity such as heat or chemical insult (Albanèse *et al.*, 2006). Peptides are synthesized deep within the ribosome and subsequently exit via a channel within the 60S subunit, protecting the peptides until they exit. The length of peptide protected is around 40 amino acids. Chaperones have long been known to associate with the nascent peptide emerging from the translating ribosome, which is believed to facilitate their folding. Following proteotoxic stress from heat or chemicals, ribosomes globally stall upon messages within the first 50–65 codons, coinciding with the emergence of peptides from the exit channel (Liu *et al.*, 2013; Shalgi *et al.*, 2013). This stalling correlates with reduced chaperone association with ribosomes, indicating that ribosomes sense the availability of chaperones and pause to allow proteotoxic stress to be combatted (Figure 6(a)).

Stalling of ribosomes has also been linked to the ubiquitinylation and degradation of aberrant nascent peptides in yeast. A ribosome quality control (RQC complex) comprising of Ltn1, Rqc1, Tae2, and Cdc48 associates with stalled ribosomes (Figure 6(b)), using Cdc48 and Ltn1 to extract and ubiquitinylate the aberrant nascent chain (Brandman *et al.*, 2012).

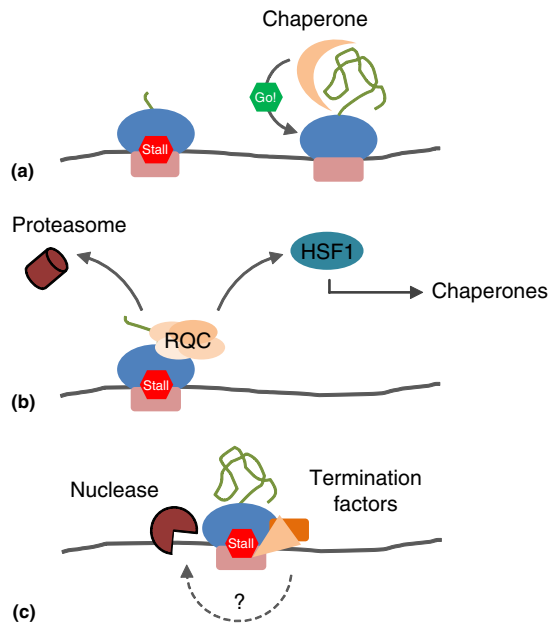


Figure 6 Ribosomes as a platform for quality control. (a) Chaperones are required for elongation after emergence of a peptide from the ribosome exit channel. In the absence of chaperones, such as during proteotoxic stress, ribosomes stall as the nascent peptide exits. This can also be interpreted as the binding of chaperones to nascent chains signaling to the ribosomes to allow elongation. (b) Ribosome stalling drives recognition and degradation of defective peptides. Protein defects can stall ribosome elongation, signaling to the ribosome quality control (RQC) complex to extract and ubiquitinate the peptide for proteasomal degradation. This complex can also enhance chaperone production via Hsf1. The RQC complex directly interacts with the 60S ribosomal subunit to elicit these functions. (c) Defective mRNAs stop ribosome elongation and are degraded by nucleases. mRNAs harboring an aberration require termination factors to halt translation upstream of the ribosome. Cleavage of the message occurs at the opposite side of the ribosome, suggesting that the ribosome acts as a platform for recruitment of nucleases.

Interestingly this requires the direct association of the RQC complex with the 60S subunit, a process which initiates prior to dissociation of the 80S ribosome but requires dissociation to complete (Defenuouillère *et al.*, 2013; Shao *et al.*, 2013). Of further note is the signaling capacity of the RQC complex, which uses Tae2 to up-regulate stress response chaperones via the transcription factor Hsf1 (Brandman *et al.*, 2012). This signal comes directly from stalled 60S subunits. The system has been characterized in yeast, and is genetically conserved in mammals, where the mammalian Ltn1 homolog is mutated in a murine model of neurodegeneration linked to defective protein quality control (Bengtson and Joazeiro, 2010).

mRNA quality control

mRNA surveillance has been categorized according to 3 different types of mRNA aberration and the degradation systems employed to eradicate them: mRNAs with an early stop codon (nonsense mediated decay), those lacking a stop codon (nonstop decay) and those whose structure inhibits elongation by the ribosome (no-go decay) (reviewed by Shoemaker and Green, 2012). For the latter two, poor quality is signaled by

stalling of the ribosome, either due to mRNA structure or to the synthesis of a stall-inducing basic polylysine tract from the polyA tail. An early stop codon can be recognized by the presence of an exon junction complex downstream of the terminating stop codon, or the distance between the stop codon and the 3'UTR.

After recognition of aberrant mRNA, endo- and exonucleases are recruited to cleave the message and degrade each fragment in both directions, efficiently removing the capacity to translate the mRNA (Eberle *et al.*, 2009; Gatfield and Izaurralde, 2004). Termination factor binding is required to initiate degradation, but mRNA cleavage occurs at the opposite side of the ribosome (Shoemaker *et al.*, 2010; Becker *et al.*, 2011). This has led to a theoretical role for the ribosome as a scaffold for recruitment of the degradation machinery (Figure 6(c)).

The End of the Line – Ribosome Degradation

Ribosome synthesis is subject to strict surveillance to ensure production of optimally functional ribosomes. Nevertheless, this process is error prone necessitating the degradation of aberrant ribosomal precursors. Ribosome surveillance does not stop in the nucleus, with proof-reading occurring in the cytoplasm and mechanisms present to eliminate aberrant ribosomes. Mature cytoplasmic ribosomes can also pay the ultimate price for cell survival by being targeted for bulk degradation by a specialized form of autophagy.

Degradation of Defective Ribosomes

Defects detected within pre-ribosomes are dealt with by the nuclear RNA degradation machinery. Catalytically this uses an exonuclease complex termed the exosome, but is likely to require polyadenylation of the defective species by exosome cofactors (Lubas *et al.*, 2011). It is interesting to note that the exosome is responsible for on-track endonucleolytic cleavage of pre-rRNAs during ribosome biogenesis (Sloan *et al.*, 2013b) as well as being the likely nuclease involved in mRNA degradation within the cytoplasm. This highlights the importance of (pre-) ribosome-exosome interactions which may be similar for each process.

Despite these nuclear surveillance mechanisms, defective ribosomal subunits can enter the cytoplasm. Here, 40S subunits undergo a proof-reading step before entering the pool of translating ribosomes (Strunk *et al.*, 2012; Lebaron *et al.*, 2012). This proof-reading was identified in yeast and is likely to be conserved to mammalian cells. However, the outcome of failure at this proof-reading step for both the defective 40S subunit and stress signaling remains to be elucidated. It is also unknown whether 60S subunits are proof-read in this way.

Both ribosomal subunits are also monitored for aberrations which may result from faulty biogenesis or stress induced damage. For the 40S subunit this process is recognized during translation elongation by stalling of the defective ribosome (Cole *et al.*, 2009). Degradation then follows a similar pathway to degradation of mRNA by no-go decay, involving endo- and exonucleases. In contrast, ubiquitinylation of aberrant 60S

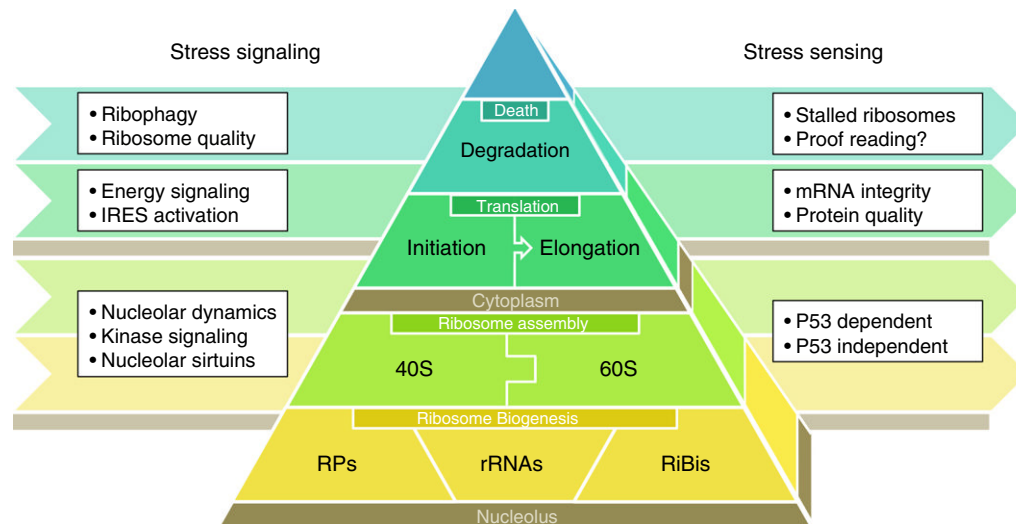


Figure 7 Ribosome biogenesis as the foundation for ribosome function and stress. Throughout their lives (running from bottom to top) ribosomes respond to stress signaling (left) and are able to detect stress at its source (right). Ribosome biogenesis in the nucleolus is the foundation of ribosome stress. This process detects stress and has a severe response mechanism, signaling to p53 and other proteins capable of deciding cell fate. Simultaneously, nucleoli are responsive to stress signaling from kinases, sirtuins and alterations in their protein components. As well as carrying out these stress functions nucleoli assemble 40S and 60S subunits, which are exported to the cytoplasm to catalyze translation. Translation is regulated by some of the stress response pathways that also govern ribosome biogenesis. Accurately synthesized ribosomes are also required for activation of cap-independent initiation via IRESs, which is common following stress. Ribosome fidelity may also be important for the emerging roles of ribosomes in mRNA and protein quality control. This perhaps explains the strict surveillance of ribosome biogenesis and of mature ribosome quality, which feed into ribosome degradation pathways. Finally, ribosomes can still contribute to the cell in death, by undergoing ribophagy in times of stress to replenish biomolecule availability.

subunits drives their degradation by the proteasome (Fujii *et al.*, 2009). Ubiquitinylation presumably occurs upon RPs of the 60S subunits. Inhibition of the proteasome stabilizes defective 60S rRNAs indicating that their degradation is dependent upon proteasomal targeting (Fujii *et al.*, 2012). As such, the paradigm of RNA and protein management being split between the 40S and 60S subunits respectively is paralleled by subunit degradation mechanisms – degradation of the 40S uses RNA degradation pathways, whereas 60S degradation is driven by protein degradation pathways.

Ribophagy

Under nutrient stress the demand for biomolecules drives the shutdown of ribosome biogenesis and translation. However, severe nutrient stress may require further measures. One such mechanism is the bulk degradation of cellular components within cytoplasmic vesicles – a process called autophagy. Ribosomes contain both nucleic and amino acids, making ribophagy (autophagy of ribosomes) an attractive source of biomolecules. In yeast, both 40S and 60S subunits are substrates for ribophagy, with the 60S mechanism dependent upon a ubiquitinylation switch involving both ubiquitin proteases and ligases (Kraft *et al.*, 2008; Kraft and Peter, 2008). Supplementing the ubiquitin modulating enzymes in ribophagy is the Cdc48 protein (Ossareh-Nazari *et al.*, 2010). Ribophagy appears to be unbiased by ribosome function, but it is interesting to note that Cdc48 is also a component of the ribosome associated protein quality control complex (Brandman *et al.*, 2012), indicating multiple Cdc48

activities at the ribosome. 40S subunits are susceptible to ribophagy, but the markers for their degradation remain to be identified.

Conclusion

Ribosome stress is a complex system of two-way regulation, where potential stresses are detected both during the biogenesis and subsequent function of ribosomes. We have seen that ribosomal energy expenditure has necessitated their synergy with energy-stress signaling pathways. This has often been considered separately for ribosome biogenesis and for translation, despite some of the same stress pathways modulating both processes. Cells have evolved mechanisms to deal with stress at the local level of the translating ribosome (aberrant protein, mRNA or ribosome decay) as well as a fail-safe activation of p53 and other cell fate regulators linked to ribosome biogenesis. Synthesis of ribosomes evidently evolved as a closely monitored pathway, to which further arms of cell stress recognition have been added (Figure 7). The severity of the response to detection of stress within nucleoli, which is linked to alterations in ribosome biogenesis, may act to ensure the full ribosomal repertoire is in place for translation. It seems that cells employ the mantra of doing things right the first time to ensure fewer problems in the future.

See also: Cell Division/Death: Regulation of Cell Growth: Internal Ribosome Entry Site-Mediated Translation; mTORC1:

Upstream and Downstream; Ribosomal RNA Processing; The Nucleolus. **Nucleic Acid Synthesis/Breakdown: RNA Synthesis/Function:** Messenger RNA (mRNA): The Link between DNA and Protein; Ribosomal RNAs and Protein Synthesis; The Interplay between Eukaryotic mRNA Degradation and Translation. **Organelles: Structure and Function:** At the Center of Autophagy: Autophagosomes. **Protein Synthesis/Degradation: Protein Degradation – Intracellular:** Ubiquitin, Ubiquitin-Like Proteins, and Proteasome-Mediated Degradation. **Protein Synthesis/Degradation: Translation:** Components, Initiation, Elongation, Termination, and Regulation

References

- Albanèse, V., Yam, A.Y.-W., Baughman, J., Parnot, C., Frydman, J., 2006. Systems analyses reveal two chaperone networks with distinct functions in eukaryotic cells. *Cell* 124, 75–88.
- Andersen, J.S., Lam, Y.W., Leung, A.K.L., *et al.*, 2005. Nucleolar proteome dynamics. *Nature* 433, 77–83.
- Anger, A.M., Armache, J.-P., Berninghausen, O., *et al.*, 2013. Structures of the human and *Drosophila* 80S ribosome. *Nature* 497, 80–85.
- Audas, T.E., Jacob, M.D., Lee, S., 2012. Immobilization of proteins in the nucleolus by ribosomal intergenic spacer noncoding RNA. *Molecular Cell* 45, 147–157.
- Becker, T., Armache, J.-P., Jarasch, A., *et al.*, 2011. Structure of the no-go mRNA decay complex Dom34–Hbs1 bound to a stalled 80S ribosome. *Nature Structural & Molecular Biology* 18, 715–720.
- Bengtson, M.H., Joazeiro, C.A.P., 2010. Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. *Nature* 467, 470–473.
- Bernstein, K.A., Bleichert, F., Bean, J.M., Cross, F.R., Baserga, S.J., 2007. Ribosome biogenesis is sensed at the start cell cycle checkpoint. *Molecular Biology of the Cell* 18, 953–964.
- Brandman, O., Stewart-Ornstein, J., Wong, D., *et al.*, 2012. A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. *Cell* 151, 1042–1054.
- Canto, C., Gerhart-Hines, Z., Feige, J.N., *et al.*, 2009. AMPK regulates energy expenditure by modulating NAD⁺ metabolism and SIRT1 activity. *Nature* 458, 1056–1060.
- Challagundla, K.B., Sun, X.-X., Zhang, X., *et al.*, 2011. Ribosomal protein L11 recruits miR-24/miRISC to repress c-Myc expression in response to ribosomal stress. *Molecular and Cellular Biology* 31, 4007–4021.
- Chauvin, C., Koka, V., Nouschi, A., *et al.*, 2014. Ribosomal protein S6 kinase activity controls the ribosome biogenesis transcriptional program. *Oncogene* 33, 474–483.
- Chen, S., Seiler, J., Santiago-Reichert, M., *et al.*, 2013. Repression of RNA polymerase I upon stress is caused by inhibition of RNA-dependent deacetylation of PAF53 by SIRT7. *Molecular Cell* 52, 303–313.
- Cole, S.E., Larivière, F.J., Merrick, C.N., Moore, M.J., 2009. A convergence of rRNA and mRNA quality control pathways revealed by mechanistic perturbation of nonfunctional rRNA decay. *Molecular Cell* 34, 440–450.
- Colombo, E., Bonetti, P., Lazzarini Denchi, E., *et al.*, 2005. Nucleophosmin is required for DNA integrity and p19Arf protein stability. *Molecular and Cellular Biology* 25, 8874–8886.
- Daftuar, L., Zhu, Y., Jacq, X., Prives, C., 2013. Ribosomal proteins RPL37, RPS15 and RPS20 regulate the Mdm2-p53-MdmX network. *PLoS One* 8, e68667.
- Dai, M.-S., Shi, D., Jin, Y., *et al.*, 2006. Regulation of the MDM2-p53 pathway by ribosomal protein L11 involves a post-ubiquitination mechanism. *Journal of Biological Chemistry* 281, 24304–24313.
- Dai, M.-S., Zeng, S.X., Jin, Y., *et al.*, 2004. Ribosomal protein L23 activates p53 by inhibiting MDM2 function in response to ribosomal perturbation but not to translation inhibition. *Molecular and Cellular Biology* 24, 7654–7668.
- Dai, M.S., Arnold, H., Sun, X.X., Sears, R., Lu, H., 2007. Inhibition of c-Myc activity by ribosomal protein L11. *EMBO Journal* 26, 3332–3345.
- Defenouillère, Q., Yao, Y., Mouaikel, J., *et al.*, 2013. Cdc48-associated complex bound to 60S particles is required for the clearance of aberrant translation products. *Proceedings of the National Academy of Sciences of the United States of America* 110, 5046–5051.
- Deisenroth, C., Zhang, Y., 2010. Ribosome biogenesis surveillance: Probing the ribosomal protein-Mdm2-p53 pathway. *Oncogene* 29, 4253–4260.
- Donati, G., Brighenti, E., Vici, M., *et al.*, 2011. Selective inhibition of rRNA transcription downregulates E2F-1: A new p53-independent mechanism linking cell growth to cell proliferation. *Journal of Cell Science* 124, 3017–3028.
- Donati, G., Peddigari, S., Mercer, C.A., Thomas, G., 2013. 5S Ribosomal RNA is an essential component of a nascent ribosomal precursor complex that regulates the Hdm2-p53 checkpoint. *Cell Reports* 4, 87–98.
- Drygin, D., Siddiqui-Jain, A., O'Brien, S., *et al.*, 2009. Anticancer activity of CX-3543: A direct inhibitor of rRNA biogenesis. *Cancer Research* 69, 7653–7661.
- Eberle, A.B., Lykke-Andersen, S., Muhlemann, O., Jensen, T.H., 2009. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nature Structural & Molecular Biology* 16, 49–55.
- Finkbeiner, E., Haindl, M., Muller, S., 2011. The SUMO system controls nucleolar partitioning of a novel mammalian ribosome biogenesis complex. *EMBO Journal* 30, 1067–1078.
- Ford, E., Voit, R., Liszt, G., *et al.*, 2006. Mammalian Sir2 homolog SIRT7 is an activator of RNA polymerase I transcription. *Genes & Development* 20, 1075–1080.
- Fujii, K., Kitabatake, M., Sakata, T., Miyata, A., Ohno, M., 2009. A role for ubiquitin in the clearance of nonfunctional rRNAs. *Genes & Development* 23, 963–974.
- Fujii, K., Kitabatake, M., Sakata, T., Ohno, M., 2012. 40S subunit dissociation and proteasome-dependent RNA degradation in nonfunctional 25S rRNA decay. *EMBO Journal* 31, 2579–2589.
- Fumagalli, S., Ivanenkov, V.V., Teng, T., Thomas, G., 2012. Suprainduction of p53 by disruption of 40S and 60S ribosome biogenesis leads to the activation of a novel G2/M checkpoint. *Genes & Development* 26, 1028–1040.
- Gatfield, D., Izaurralde, E., 2004. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in *Drosophila*. *Nature* 429, 575–578.
- Hoppe, S., Bierhoff, H., Cado, I., *et al.*, 2009. AMP-activated protein kinase adapts rRNA synthesis to cellular energy supply. *Proceedings of the National Academy of Sciences of the United States of America* 106, 17781–17786.
- Horman, S., Browne, G.J., Krause, U., *et al.*, 2002. Activation of AMP-activated protein kinase leads to the phosphorylation of elongation factor 2 and an inhibition of protein synthesis. *Current Biology* 12, 1419–1423.
- Horn, H.F., Vousden, K.H., 2008. Cooperation between the ribosomal proteins L5 and L11 in the p53 pathway. *Oncogene* 27, 5774–5784.
- Huang, J.-Y., Su, W.-C., Jeng, K.-S., Chang, T.-H., Lai, M.M.C., 2012. Attenuation of 40S ribosomal subunit abundance differentially affects host and HCV translation and suppresses HCV replication. *PLoS Pathogens* 8, e1002766.
- Iadevaia, V., Caldarola, S., Biondini, L., *et al.*, 2010. PIM1 kinase is destabilized by ribosomal stress causing inhibition of cell cycle progression. *Oncogene* 29, 5490–5499.
- Iadevaia, V., Zhang, Z., Jan, E., Proud, C.G., 2012. mTOR signaling regulates the processing of pre-rRNA in human cells. *Nucleic Acids Research* 40, 2527–2539.
- Itahana, K., Bhat, K.P., Jin, A., *et al.*, 2003. Tumor suppressor ARF degrades B23, a nucleolar protein involved in ribosome biogenesis and cell proliferation. *Molecular Cell* 12, 1151–1164.
- Jang, S.K., Kräusslich, H.G., Nicklin, M.J., *et al.*, 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during *in vitro* translation. *Journal of Virology* 62, 2636–2643.
- Jin, A., Itahana, K., O'Keefe, K., Zhang, Y., 2004. Inhibition of MDM2 and activation of p53 by ribosomal protein L23. *Molecular and Cellular Biology* 24, 7669–7680.
- Johannes, G., Carter, M.S., Eisen, M.B., Brown, P.O., Sarnow, P., 1999. Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. *Proceedings of the National Academy of Sciences of the United States of America* 96, 13118–13123.
- Kim, E.J., Kho, J.H., Kang, M.R., Um, S.J., 2007. Active regulator of SIRT1 cooperates with SIRT1 and facilitates suppression of p53 activity. *Molecular Cell* 28, 277–290.
- Knight, J.R.P., Allison, S.J., Milner, J., 2013a. Active regulator of SIRT1 is required for cancer cell survival but not for SIRT1 activity. *Open Biology* 3, 130130.
- Knight, J.R.P., Willis, A.E., Milner, J., 2013b. Active regulator of SIRT1 is required for ribosome biogenesis and function. *Nucleic Acids Research* 41, 4185–4197.
- Korgaonkar, C., Hagen, J., Tompkins, V., *et al.*, 2005. Nucleophosmin (B23) targets ARF to nucleoli and inhibits its function. *Molecular and Cellular Biology* 25, 1258–1271.
- Kraft, C., Deplazes, A., Sohrmann, M., Peter, M., 2008. Mature ribosomes are selectively degraded upon starvation by an autophagy pathway requiring the Ubp3p/Bre5p ubiquitin protease. *Nature Cell Biology* 10, 602–610.

- Kraft, C., Peter, M., 2008. Is the Rsp5 ubiquitin ligase involved in the regulation of ribophagy? *Autophagy* 4, 838–840.
- Kuo, M.-L., Den Besten, W., Thomas, M.C., Sherr, C.J., 2008. Arf-induced turnover of the nucleolar nucleophosmin-associated SUMO-2/3 protease Snp3. *Cell Cycle* 7, 3378–3387.
- Kurki, S., Peltonen, K., Latonen, L., *et al.*, 2004. Nucleolar protein NPM interacts with HDM2 and protects tumor suppressor protein p53 from HDM2-mediated degradation. *Cancer Cell* 5, 465–475.
- Lam, Y.W., Lamond, A.I., Mann, M., Andersen, J.S., 2007. Analysis of nucleolar protein dynamics reveals the nuclear degradation of ribosomal proteins. *Current Biology* 17, 749–760.
- Landry, D.M., Hertz, M.I., Thompson, S.R., 2009. RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs. *Genes & Development* 23, 2753–2764.
- Lebaron, S., Schneider, C., Van Nues, R.W., *et al.*, 2012. Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. *Nature Structural & Molecular Biology* 19, 744–753.
- Li, M., Gu, W., 2011. A critical role for noncoding 5S rRNA in regulating Mdmx stability. *Molecular Cell* 43, 1023–1032.
- Lindström, M.S., Nistér, M., 2010. Silencing of ribosomal protein S9 elicits a multitude of cellular responses inhibiting the growth of cancer cells subsequent to p53 activation. *PLoS One* 5, e9578.
- Liu, B., Han, Y., Qian, S.-B., 2013. Cotranslational response to proteotoxic stress by elongation pausing of ribosomes. *Molecular Cell* 49, 453–463.
- Llanos, S., Clark, P.A., Rowe, J., Peters, G., 2001. Stabilization of p53 by p14ARF without relocation of MDM2 to the nucleolus. *Nature Cell Biology* 3, 445–452.
- Lubas, M., Christensen, M.S., Kristiansen, M.S., *et al.*, 2011. Interaction profiling identifies the human nuclear exosome targeting complex. *Molecular Cell* 43, 624–637.
- Marcel, V., Ghayad, S.E., Belin, S., *et al.*, 2013. p53 Acts as a safeguard of translational control by regulating fibrillarin and rRNA methylation in cancer. *Cancer Cell* 24, 318–330.
- Marechal, V., Elenbaas, B., Piette, J., Nicolas, J.C., Levine, A.J., 1994. The ribosomal L5 protein is associated with mdm-2 and mdm-2-p53 complexes. *Molecular and Cellular Biology* 14, 7414–7420.
- Mayer, C., Bierhoff, H., Grummt, I., 2005. The nucleolus as a stress sensor: JNK2 inactivates the transcription factor TIF-IA and down-regulates rRNA synthesis. *Genes & Development* 19, 933–941.
- Mayer, C., Zhao, J., Yuan, X., Grummt, I., 2004. mTOR-dependent activation of the transcription factor TIF-IA links rRNA synthesis to nutrient availability. *Genes & Development* 18, 423–434.
- Mitchell, S.A., Spriggs, K.A., Bushell, M., *et al.*, 2005. Identification of a motif that mediates polypyrimidine tract-binding protein-dependent internal ribosome entry. *Genes & Development* 19, 1556–1571.
- Moore, H.M., Bai, B., Boisvert, F.-M., *et al.*, 2011. Quantitative proteomics and dynamic imaging of the nucleolus reveal distinct responses to UV and ionizing radiation. *Molecular & Cellular Proteomics* 10 (10), M111.009241.
- Murayama, A., Ohmori, K., Fujimura, A., *et al.*, 2008. Epigenetic control of rDNA loci in response to intracellular energy status. *Cell* 133, 627–639.
- Muth, V., Nadaud, S., Grummt, I., Voit, R., 2001. Acetylation of TAFI68, a subunit of TIF-IB/SL1, activates RNA polymerase I transcription. *EMBO Journal* 20, 1353–1362.
- Ossareh-Nazari, B., Bonizic, M., Cohen, M., *et al.*, 2010. Cdc48 and Ufd3, new partners of the ubiquitin protease Ubp3, are required for ribophagy. *EMBO Reports* 11, 548–554.
- Pestov, D.G., Strezoska, Ž., Lau, L.F., 2001. Evidence of p53-dependent cross-talk between ribosome biogenesis and the cell cycle: Effects of nucleolar protein Bop1 on G1/S transition. *Molecular and Cellular Biology* 21, 4246–4255.
- Pomerantz, J., Schreiber-Agus, N., Liégeois, N.J., *et al.*, 1998. The ink4a tumor suppressor gene product, p19Arf, interacts with MDM2 and neutralizes MDM2's inhibition of p53. *Cell* 92, 713–723.
- Rodgers, J.T., Lerin, C., Haas, W., *et al.*, 2005. Nutrient control of glucose homeostasis through a complex of PGC-1alpha and SIRT1. *Nature* 434, 113–118.
- Rodway, H., Llanos, S., Rowe, J., Peters, G., 2004. Stability of nucleolar versus non-nucleolar forms of human p14ARF. *Oncogene* 23, 6186–6192.
- Rubbi, C.P., Milner, J., 2003. Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *EMBO Journal* 22, 6068–6077.
- Shalgi, R., Hurt, J., Jessica, A., *et al.*, 2013. Widespread regulation of translation by elongation pausing in heat shock. *Molecular Cell* 49, 439–452.
- Shao, S., Von der Malsburg, K., Hegde, R.S., 2013. Listerin-dependent nascent protein ubiquitination relies on ribosome subunit dissociation. *Molecular Cell* 50, 637–648.
- Shoemaker, C.J., Eyler, D.E., Green, R., 2010. Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. *Science* 330, 369–372.
- Shoemaker, C.J., Green, R., 2012. Translation drives mRNA quality control. *Nature Structural & Molecular Biology* 19, 594–601.
- Sloan, K.E., Bohnsack, M.T., Watkins, N.J., 2013a. The 5S RNP couples p53 homeostasis to ribosome biogenesis and nucleolar stress. *Cell Reports* 5, 237–247.
- Sloan, K.E., Mattijssen, S., Lebaron, S., *et al.*, 2013b. Both endonucleolytic and exonucleolytic cleavage mediate ITS1 removal during human ribosomal RNA processing. *Journal of Cell Biology* 200, 577–588.
- Song, T., Yang, L., Kabra, N., *et al.*, 2013. The NAD⁺ synthesis enzyme nicotinamide mononucleotide adenylyltransferase (NMNAT1) regulates ribosomal RNA transcription. *Journal of Biological Chemistry* 288, 20908–20917.
- Spriggs, K.A., Bushell, M., Willis, A.E., 2010. Translational regulation of gene expression during conditions of cell stress. *Molecular Cell* 40, 228–237.
- Strunk, B.S., Novak, M.N., Young, C.L., Karbstein, K., 2012. A translation-like cycle is a quality control checkpoint for maturing 40S ribosome subunits. *Cell* 150, 111–121.
- Tao, W., Levine, A.J., 1999. P19(ARF) stabilizes p53 by blocking nucleocytoplasmic shuttling of Mdm2. *Proceedings of the National Academy of Sciences of the United States of America* 96, 6937–6941.
- Terzian, T., Box, N., 2013. Genetics of ribosomal proteins: “Curiouser and curiouser”. *PLoS Genetics* 9, e1003300.
- Valášek, L., Mathew, A.A., Shin, B.-S., *et al.*, 2003. The yeast eIF3 subunits TIF32/a, NIP1/c, and eIF5 make critical connections with the 40S ribosome *in vivo*. *Genes & Development* 17, 786–799.
- Volarević, S., Stewart, M.J., Ledermann, B., *et al.*, 2000. Proliferation, but not growth, blocked by conditional deletion of 40S ribosomal protein S6. *Science* 288, 2045–2047.
- Wang, X., Li, W., Williams, M., *et al.*, 2001. Regulation of elongation factor 2 kinase by p90RSK1 and p70 S6 kinase. *EMBO Journal* 20, 4370–4379.
- Weber, J.D., Taylor, L.J., Roussel, M.F., Sherr, C.J., Bar-Sagi, D., 1999. Nucleolar Arf sequesters Mdm2 and activates p53. *Nature Cell Biology* 1, 20–26.
- Yogev, O., Saadon, K., Anzi, S., Inoue, K., Shaulian, E., 2008. DNA damage-dependent translocation of B23 and p19ARF is regulated by the Jun N-terminal kinase pathway. *Cancer Research* 68, 1398–1406.
- Yoon, A., Peng, G., Brandenburg, Y., *et al.*, 2006. Impaired control of IRES-mediated translation in X-linked dyskeratosis congenita. *Science* 312, 902–906.
- Zhang, Y., Lu, H., 2009. Signaling to p53: Ribosomal proteins find their way. *Cancer Cell* 16, 369–377.

Further Reading

- Boulon, S., Westman, B.J., Hutten, S., Boisvert, F.-M., Lamond, A.I., 2010. The nucleolus under Stress. *Molecular Cell* 40, 216–227.
- Donati, G., Montanaro, L., Derenzini, M., 2012. Ribosome biogenesis and control of cell proliferation: p53 is not alone. *Cancer Research* 72, 1602–1607.
- Grummt, I., 2013. The nucleolus – Guardian of cellular homeostasis and genome integrity. *Chromosoma* 122, 487–497.
- Holmberg Olsson, K., Nistér, M., Lindström, M.S., 2012. p53-dependent and -independent nucleolar stress responses. *Cells* 1, 774–798.
- Lafontaine, D.L., 2010. A ‘garbage can’ for ribosomes: How eukaryotes degrade their ribosomes. *Trends in Biochemical Sciences* 35, 267–277.
- Pechmann, S., Willmund, F., Frydman, J., 2013. The ribosome as a hub for protein quality control. *Molecular Cell* 49, 411–421.
- Pichon, X., Wilson, L.A., Stoneley, M., *et al.*, 2012. RNA binding protein/RNA element interactions and the control of translation. *Current Protein & Peptide Science* 13, 294–304.

Relevant Websites

- <http://www.ribogenesis.com>
RIBO Genesis.
- <http://www.lamondlab.com/NOPdb/>
The Lamond Lab.

Extra-Ribosome Functions of Ribosomal Proteins

DA Franklin and Y Zhang, The University of North Carolina, Chapel Hill, NC, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction to Basic Ribosome Structure and Function

As discussed in other sections of this work, cells follow the central dogma of cell biology, which describes the conversion of template DNA to secondary messengers (mRNA), which are translated into the primary effectors (proteins). This complex process requires precise collaboration and rapid turnover of ribosomes, which are responsible for the coordination of mRNA transcripts, and charged tRNA molecules to form peptide bonds required for protein synthesis. Eukaryotes express 79 ribosomal proteins that combine with ribosomal RNA (rRNA) to form the large and small subunits of the functional ribosome (Boisvert *et al.*, 2007). The functions of ribosomal proteins in protein synthesis are discussed in other sections of this work. In this section, we explore the extra-ribosomal functions of these proteins. First, we provide a brief explanation of ribosomal assembly and canonical ribosome function. Then, we will address the regulation of ribosome biogenesis by both tumor suppressors and oncogenes. Finally, extra-ribosomal functions and the role of mutant ribosomal proteins in disease will be discussed with a focus on the ribosomal protein-MDM2-p53 pathway.

Ribosome Assembly

In actively dividing cells, ribosome biogenesis accounts for a significant portion of cellular resources. In yeast, and presumably in mammalian cells, upwards of 60% of total cellular resources are devoted to the rapid generation of ribosomes (Warner, 1999). Ribosome biogenesis occurs in the nucleolus, a condensed non-membrane bound section of the nucleus that contains a number of uncharacterized proteins and chromosomal DNA. The lack of a membrane barrier between the nucleolus and the nucleoplasm is unique among organelles, as it allows for the free exchange of molecules between these two compartments. Prior to transcription of ribosomal genes, tandem repeats of chromosomal rDNA genes are arranged in the nucleolar organizing regions (NOR). Upon establishment of the NOR, the 47S precursor rRNA is then transcribed by RNA polymerase I (RNA pol I) in the fibrillar center of the nucleolus and then further processed to form the mature 18S, 28S, and 5.8S rRNA's (Boisvert *et al.*, 2007). In contrast, the gene encoding the 5S rRNA is transcribed by RNA polymerase III (RNA pol III) prior to export from the nucleolus to the cytoplasm where it binds RPL5 to form the 5S ribonucleoprotein (RNP) (Szymanski *et al.*, 2003). Ribosomal proteins are transcribed by RNA polymerase II (RNA pol II) and are exported from the nucleus for translation. The majority of ribosomal proteins demonstrate basic isoelectric points, which facilitates their association with rRNA molecules found in the ribosome (Lempiäinen and Shore, 2009). After translation in the cytoplasm, the nascent ribosomal proteins are then imported back into the nucleolus for assembly of the large (60S) and small (40S) ribosomal subunits. In the absence of sufficient rRNA, many

ribosomal proteins are thought to form a negative feedback loop against their own transcription and translation to prevent aberrant accumulation of ribosomal proteins. For example, several ribosomal proteins including ribosomal protein L30 (RPL30), ribosomal protein S14 (RPS14), ribosomal protein L12 (RPL12) in *Caenorhabditis elegans*, and ribosomal protein S13 (RPS13) in humans have been shown to inhibit their own mRNA splicing thereby forming negative feedback loops, which can prevent excessive RP accumulation. Moreover, RPL2 and RPS28 have both been shown to shorten the half-lives of their own mRNA transcripts (Warner and McIntosh, 2009). RNA interference has been utilized in the study of various proteins after it was initially discovered that introduction of short double stranded RNAs silenced genes containing the homologous sequence in *C. elegans* (Fire *et al.*, 1998). Furthermore, short interfering RNAs (siRNA) of approximately 21–22 nucleotides can be introduced to cells to specifically decrease expression of a gene using the cells own machinery to degrade the double strand RNA complexes created by siRNA binding to specific mRNA transcripts (Elbashir *et al.*, 2001). Consistent with these observations, siRNA-mediated suppression of a single ribosomal protein within the small subunit of the ribosome (RPS) in HeLa cells can cause the concomitant suppression of the expression of other RPS proteins while having no effect on the expression of ribosomal proteins of the large subunit (RPL). The decreased expression of subunit-specific ribosomal proteins was also observed for siRNA-mediated knockdown of RPL protein expression suggesting that feedback inhibition can prevent the accumulation of subunit-specific ribosomal proteins to avoid stoichiometric imbalances (Robledo *et al.*, 2008). This method of auto-regulation could play an important role in preventing the activation of various stress pathways we will see later in this article.

Ribosome Biogenesis Can Be Inhibited Using Small Molecules

The high energetic demands of ribosome biogenesis require the cell to exert tight control over the process and allow for changes in flux through the pathway in response to a variety of cellular stresses. Low levels of the chemotherapeutic Actinomycin D (ActD) can be used to specifically inhibit rRNA transcription through RNA pol I independent of the DNA damage that is observed with higher concentrations of ActD (Iapalucci-Espinoza and Franze-Fernandez, 1979). Furthermore, another chemotherapeutic 5-fluorouracil (5-FU) can be used to inhibit all RNA synthesis, which inhibits ribosome biogenesis in a less specific manner (Longley *et al.*, 2003).

Ribosome Biogenesis Is Intrinsically Connected to Cancerous Growth

Tightly Regulated by Oncogenes

Cancer cell proliferation requires high levels of protein synthesis and correspondingly high levels of ribosome biogenesis.

A number of oncogenes such c-Myc, protein kinase B (Akt), and extracellular signal-regulated kinases (ERK) have been shown to upregulate ribosome biogenesis primarily by the regulation of rRNA transcription.

c-Myc is a general transcription factor that regulates the transcriptional activity of all three RNA polymerases (RNA pol I, II, and III) through various mechanisms. Briefly, c-Myc colocalizes with its binding partner Max to E-box sequences found within the promoter region of rDNA resulting in an increase in RNA pol I transcription (Grandori *et al.*, 2005). c-Myc also activates RNA pol III transcriptional activity through its association with TFIIB. The association between c-Myc and TFIIB is observed within the promoter regions of 5S rRNA and tRNA codons and correlates with an increase in the transcription of these genes (Gomez-Roman *et al.*, 2003). c-Myc regulation of RNA pol II transcription, which is responsible for the transcription of ribosomal proteins and other proteins associated with ribosome biogenesis, is evident based on the increased presence of c-Myc within RNA pol II promoter regions (Menssen and Hermeking, 2002). Recent studies have shown that c-Myc also regulates RNA pol II mediated transcription by increasing the rate at previously assembled complexes initiate transcription, which is known as pause and release regulation (Rahl *et al.*, 2010). Furthermore, c-Myc mediated increases in ribosome biogenesis play a central role in c-Myc-driven tumorigenesis as deletion of one allele of L24 or L38 inhibits the development of E μ -Myc driven lymphoma (Barna *et al.*, 2008).

The mechanistic target of rapamycin (mTOR) is an atypical serine/threonine protein kinase that regulates cell proliferation and survival through both mTOR complexes named mTORC1 and mTORC2 (Laplane and Sabatini, 2012). The PI3K-Akt-mTORC1 pathway directly regulates ribosomal protein and rRNA synthesis through the activation of S6 kinase (S6K), which phosphorylates RPS6. In non-Hodgkin lymphoma, RPS6 phosphorylation is associated with increased translation of 5' terminal oligo-pyrimidine (TOP) genes, which include ribosomal proteins, splicing factors, and translation elongation factors (Hagner *et al.*, 2011). mTORC1 also associates with either TIF-IA or upstream binding factor (UBF), a necessary cofactor in the initiation of transcription by RNA pol I, to upregulate rRNA transcription. mTORC1 is also involved in the processing and maturation of rRNA in the nucleolus (Mayer and Grummt, 2006).

The Ras-ERK oncogenic pathway coordinates with the c-Myc and Akt signaling axes to further increase ribosome biogenesis forming a complex network of oncogenes that positively regulate ribosome biogenesis. The Ras-ERK pathway has also been shown to increase expression of p90 ribosomal S6 kinase, which functions in a similar manner as mTOR-activated S6K to increase ribosome biogenesis (Roux *et al.*, 2007).

Repressed by Tumor Suppressors

In contrast to oncogenes, which upregulate ribosome biogenesis, tumor suppressors primarily inhibit this process. Increased expression of the tumor suppressor Rb directly inhibits ribosome biogenesis through its association with UBF in the nucleolus thereby preventing the proper assembly of the RNA Pol I transcription complex in mouse fibroblast cells (Hannan

et al., 2000). Similar to Rb, p53 also binds SL1, a member of the UBF transcription complex, in the nucleolus to inhibit RNA Pol I activity (Zhai and Comai, 2000). Other studies have shown that p53 also inhibits RNA Pol III through its direct binding of TFIIB, which results in the sequestration of TFIIB from active transcription complexes (Cairns and White, 1998).

The tumor suppressor ARF has been shown to inhibit ribosome biogenesis through multiple mechanisms. ARF can potentially activate p53 by binding and inhibiting MDM2; however, ARF can also inhibit ribosome biogenesis through p53-independent mechanisms. The gene mouse double minute 2 (MDM2) is the primary negative regulator of p53 and has been shown to bind directly to the N-terminal domain of p53 thereby physically preventing p53 association with DNA. ARF inhibits the phosphorylation of UBF independent of its ability to bind MDM2. Moreover, ARF inhibits rRNA processing by promoting the degradation of the NPM/B23 complex, a multifunctional complex in the nucleolus that plays a central role in ribosome biogenesis (Ayrault *et al.*, 2006; Itahana *et al.*, 2003).

Another tumor suppressor that inhibits ribosome biogenesis is the PTEN phosphatase, which dephosphorylates phosphatidylinositol-3,4,5-triphosphate resulting in the inhibition of the PI3K-Akt pathway (Li *et al.*, 2014). PTEN has also been shown to inhibit ribosome biogenesis in a more direct manner by blocking SL1-UBF complex recruitment to rDNA promoter regions (Zhang *et al.*, 2005). Current data suggest that cytoplasmic and nuclear pools of PTEN can coordinate the inhibition of ribosome biogenesis through two distinct mechanisms further demonstrating the importance of monitoring flux through the ribosome biogenesis pathway. Altogether, tight regulation of ribosome biogenesis is accomplished by oncogenes and tumor suppressor genes that appear to focus on the two signaling nodes S6K and UBF.

Ribosomopathies Are Associated with Increased Cancer Risk

Heritable genetic disorders associated with rDNA genes and ribosome assembly factors are collectively referred to as ribosomopathies. Studies investigating the mechanisms responsible for ribosomopathies have helped elucidate some of the non-ribosomal functions of ribosomal proteins. For example, studies on Diamond-Blackfan anemia (DBA) have revealed that the etiology of this particular disease is related to mutations in the RPS19 gene, and DBA patients also display an increased risk of acute myeloid leukemia (AML). More recent studies that have linked DBA with mutations in the genes *RPS24*, *RPS17*, *RPL35a*, *RPL5*, and *RPL11* have shown that this disease can lead to developmental defects in addition to the increased risk for cancer (Gazda *et al.*, 2008). In addition to DBA, a portion of chromosome 5 containing the *RPS14* gene is often deleted in patients with 5q syndrome, and this condition is similarly associated with an increased risk of AML (Ebert *et al.*, 2008). Ribosomopathies are not restricted to mutations in genes that encode ribosomal proteins, as mutations in factors that facilitate proper rRNA processing and subunit maturation are associated with Shwachman-Diamond syndrome and Treacher Collins syndrome (Freed *et al.*, 2010).

In the ribosomopathy cartilage–hair hypoplasia (CHH), mutations are commonly found in the *RMRP* gene, which contributes to rRNA processing in the nucleolus. A long-term follow-up study of CHH patients revealed a seven-fold increase in the occurrence of non-Hodgkin lymphoma and a ten-fold increase in the occurrence of basal cell carcinoma of the skin relative to a non-CHH population suggesting a direct connection between ribosome biogenesis and cancer initiation (Taskinen *et al.*, 2008). Furthermore, model organisms with alterations in ribosome biogenesis show an increased risk of cancer (Watson *et al.*, 1992; MacInnes *et al.*, 2008). In human tumors, ribosomal proteins are often overexpressed in breast, prostate, cervical, esophageal and liver cancers when compared with normal tissues (van Riggelen *et al.*, 2010). The deleterious effects of many ribosomopathies, particularly 5q syndrome, result in increased p53 activation, which is also observed in the context of ribosomal protein haploinsufficiency. Lenalidomide can be used to treat multiple ribosomopathies through a mechanism that is poorly understood but may involve the destabilization of p53 (Wei *et al.*, 2013). Ribosomopathies are associated with a wide array of phenotypes that cannot be fully explained by a decrease in protein synthesis. Thus, this class of diseases provides valuable physiological evidence of non-ribosomal functions for the factors involved in ribosome biogenesis.

Non-ribosomal Functions of RPs Independent of the MDM2-p53 Pathway

Studies on the non-ribosomal functions of RPs have been heavily focused on the ability of RPs to stabilize p53. However, many RP non-ribosomal functions have been described that do not directly involve p53 or its upstream regulator MDM2. Early studies of ribosomal proteins focused on RNA binding related extra-ribosomal functions as RPS1 was the first ribosomal protein identified to execute a non-ribosomal function. This observation was based on its ability to associate with a bacteriophage Q β peptide and endogenous peptides EF-Tu and EF-Ts to form the bacteriophage Q β replicase (Blumenthal and Carmichael, 1979). Subsequent studies have shown that bacterial ribosomal proteins RPS10 and RPS4 exhibit anti-termination effects on specific mRNAs when associated with the NUS complex after lambda phage infection (Luo *et al.*, 2008; Torres *et al.*, 2001). In another study, RPL4 was shown to bind to RNase E to modulate its activity in response to cellular stress (Singh *et al.*, 2009). Similarly, RPS3 has further been shown to exert a non-ribosomal function through its DNA endonuclease activity associated with the processing of UV-induced DNA damage (Kim *et al.*, 2006). Other studies of RPS3 have shown that this protein also functions as a critical component of the NF- κ B transcription complex in response to cellular stresses. The KH domain, which is present in many proteins that bind single-strand DNA or RNA, of RPS3 facilitates many of these extra-ribosomal activities (Wan *et al.*, 2007). RPL22 has been shown to bind to EBER-1 RNA of the Epstein–Barr virus; however, it remains unclear whether ribosomal proteins regulate viral RNA transcript stability or vice versa (Fok *et al.*, 2006). Through its basic N-terminus, RPL7 was shown to associate with and inhibit the

transcriptional regulation of the nuclear receptors that bind vitamin D and retinoic acid upon overexpression of RPL7 in yeast and at endogenous levels in human lymphoma cells (Berghofer-Hochheimer *et al.*, 1998).

As mentioned previously, tumor suppressors and oncogenes tightly regulate ribosomal protein levels and recent studies suggest that free ribosomal proteins also function as reciprocal regulators of cellular signaling. RPL11 has been shown to reduce cell proliferation through decreased ribosome biogenesis by inhibiting c-Myc-dependent transcription through RNA polymerases I, II, and III (Dai *et al.*, 2007). RPL23 has also been shown to bind and sequester nucleophosmin (NPM) from Miz1, a known inhibitor of c-Myc-dependent cell proliferation, to facilitate cell proliferation (Wanzel *et al.*, 2008). Another protein RACK1 was identified primarily for its role in cell signaling, but recent technical advances have demonstrated that RACK1 actually associates with the small subunit of the ribosome (Sengupta *et al.*, 2004). RACK1 is one of the few ribosomal proteins in which the non-ribosomal functions have been more thoroughly studied than its role as part of the ribosome (Warner and McIntosh, 2009). RACK1 functions as a scaffold protein that associates with a diverse array of binding partners including kinases, phosphatases, and G proteins among others to regulate a number of cellular processes, which inherently connects RACK1 to cancer development (Li and Xie, 2014). RPS3 and RPS6 bind HSP90, a protein chaperone, which prevents their own proteasomal degradation. Similarly, increased expression and nuclear localization of RPS3 appears to be involved in the induction of apoptosis in a caspase-dependent manner (Jang *et al.*, 2004); however, this function may involve p53 activation through the RP-MDM2-p53 pathway discussed later in this chapter. Overexpression of RPS29 causes an increase in apoptosis in non-small cell lung cancer cells, which coincides with increased expression of p53 and the pro-apoptotic p53 target gene Bax (Khanna *et al.*, 2003). RPL10 binds and inhibits the proto-oncogene c-Jun, a member of the AP-1 transcription factor and downstream target of the WNT and MAP kinase pathways (Montecarlo and Vogt, 1993; Nateri *et al.*, 2005).

Ribosomal proteins have also been implicated in multiple diseases such as Alzheimer's and type 1 diabetes. Studies on RPL10 have demonstrated that this RP binds presenillin-1, a regulator of β -amyloid peptides, the accumulation of which is associated with Alzheimer's disease. RPL10 binding to presenillin-1 appears to further modulate c-Jun transcriptional activity; however, whether this interaction plays a role in the development of Alzheimer's disease has not been confirmed (Imafuku *et al.*, 1999). Studies using large-scale integrative genomics to analyze human liver tissue have shown that reduced expression of RPS26 is associated with increased susceptibility to type 1 diabetes through a currently unknown mechanism (Schadt *et al.*, 2008). RPL22 has also been shown to bind histone H1 in *Drosophila melanogaster*, which could have implications in a variety of diseases, as histones are central regulators of gene transcription. Furthermore, the expression levels of RPL22 inversely correlate with overall mRNA transcription, as overexpression of RPL22 mimics the gene repression observed upon overexpression of histone H1 (Ni *et al.*, 2006; Table 1).

Table 1 Non-ribosomal functions of ribosomal proteins that occur independent of MDM2-p53

<i>Ribosomal protein</i>	<i>Function</i>	<i>References</i>
RPL4	Binds RNase E thereby modulating its activity	Singh <i>et al.</i> (2009)
RPL7	Inhibits transcription of vitamin D receptor and retinoid X receptor	Berghofer-Hochheimer <i>et al.</i> (1998)
RPL10	Binds and inhibits the proto-oncogene c-Jun	Montecarlo and Vogt (1993)
RPL10	Binds presenilin-1, biological function unknown	Imafuku <i>et al.</i> (1999)
RPL11	Inhibits c-Myc-driven RNA pol I-, II-, and III-mediated transcription	Dai <i>et al.</i> (2007)
RPL22	Binds Epstein-Barr viral RNA, which may promote viral life cycle	Fok <i>et al.</i> (2006)
RPL22	Binds histone H1 and overexpression decreases mRNA transcription	Ni <i>et al.</i> (2006)
RPL23	Inhibits Miz1 to facilitate c-Myc-driven cell proliferation	Wanzel <i>et al.</i> (2008)
RACK1	Scaffold protein involved in various cell signaling cascades	Li and Xie (2014)
RPS1	Associates with three other peptides to form the bacteriophage Q β replicase	Blumenthal and Carmichael (1979)
RPS3	DNA endonuclease activity responsible for processing UV-associated DNA damage	Jang <i>et al.</i> (2004)
RPS3	Component of NF- κ B transcription complex	Wan <i>et al.</i> (2007)
RPS4 and RPS10	Modulates transcription through its association with the NUS complex during lambda phage infection	Torres <i>et al.</i> (2001)
RPS26	Reduced expression correlates with an increased occurrence of type 1 diabetes	Schadt <i>et al.</i> (2008)
RPS29	Overexpression causes an increase in apoptosis that correlates with increased p53 expression	Khanna <i>et al.</i> (2003)

Ribosomal Protein-MDM2-p53 Pathway

MDM2-p53 Overview

Across all forms of cancer, p53 is the most commonly mutated gene with mutations observed in approximately 50% of tumors. The p53 stress response pathway results in the stabilization of p53 and an increase in p53 binding to the promoter regions of several target genes that largely suppress sources of genomic damage and tumor development. The canonical functions of p53 are cell cycle arrest, apoptosis and senescence. However, a recent knock-in mouse model harboring mutations that inhibit the acetylation of residues in p53 necessary for its activation following DNA damage has illustrated that the p53 tumor suppressive function requires more poorly understood p53 functions such as metabolic and reactive oxygen species (ROS) regulation (Li *et al.*, 2012). p53 is thought to enact diverse transcriptional programs that vary in severity according to the intensity of the cellular stress. For example, although p53 activation induces cell cycle arrest relatively quickly in response to several stresses, p53 upregulation of pro-apoptotic genes usually occurs only in response to the most severe cellular stresses. Moreover, studies have shown that the particular p53 transcriptional response can be altered depending on the duration of p53 expression (Purvis *et al.*, 2012).

As mentioned previously, MDM2 is the primary negative regulator of p53 through its ability to bind the N-terminal domain of p53 physically preventing p53 association with DNA. Furthermore, MDM2 acts as an E3 ubiquitin ligase that catalyzes the ubiquitination of p53, which leads to the nuclear export and subsequent proteasomal degradation of p53 (Pei *et al.*, 2012). The importance of MDM2-dependent regulation of p53 is evident in the fact that p53 activation requires the inhibition of MDM2-p53 association. DNA damage-mediated activation of the ATR-Chk1 or ATM-Chk2 pathways results in the phosphorylation of MDM2 and p53 in key residues that prevent their association. MDM2 inactivation can also be

triggered by oncogene activation, many of which increase the transcription of ARF, which binds MDM2 and inhibits p53 association (Hu *et al.*, 2012). p53 acetylation can also prevent MDM2-p53 binding and modulate p53 transcription through mechanisms that are not yet well understood (Li *et al.*, 2012; Figure 1).

p53 Activation Occurs in Response to Ribosome Biogenesis Inhibition

Hypoxia, heat shock and growth factor deprivation can activate p53 in a mechanism dependent on ribosome biogenesis inhibition (Zhang and Lu, 2009). DNA damage may also activate p53 through a similar mechanism as demonstrated by altered nucleolar structure following severe stress and highlighted by the disruption of nucleolar caps that form from condensation and segregation of nucleolar proteins and rRNA (Rubbi and Milner, 2003). p53 activation has been observed in response to changes in the expression levels of ribosomal proteins, suggesting that ribosome biogenesis requires precise stoichiometry of ribosomal proteins and other components involved in ribosome biogenesis to prevent p53 activation and facilitate cellular growth (Bursac *et al.*, 2014). p53 has been directly linked to ribosome biogenesis through the expression of a dominant negative form of the rRNA maturation factor Bop1, which mimics ribosomal stress and activates p53 (Pestov *et al.*, 2001). The activation of p53 was also observed after mutation or deletion of WDR12, which forms a complex with Bop1, further demonstrating that improper rRNA maturation and ribosome biogenesis signal through p53 (Holzel *et al.*, 2005). In actively dividing cells, the translation of ribosomal proteins in the cytoplasm places significant strain on nuclear import machinery to transport ribosomal proteins into the nucleolus for the proper assembly of the ribosomal subunits. Correspondingly, a transient decrease in the expression of importin 7 in human cancer cells inhibits ribosome biogenesis, disrupts nucleolar structure and activates p53 in an

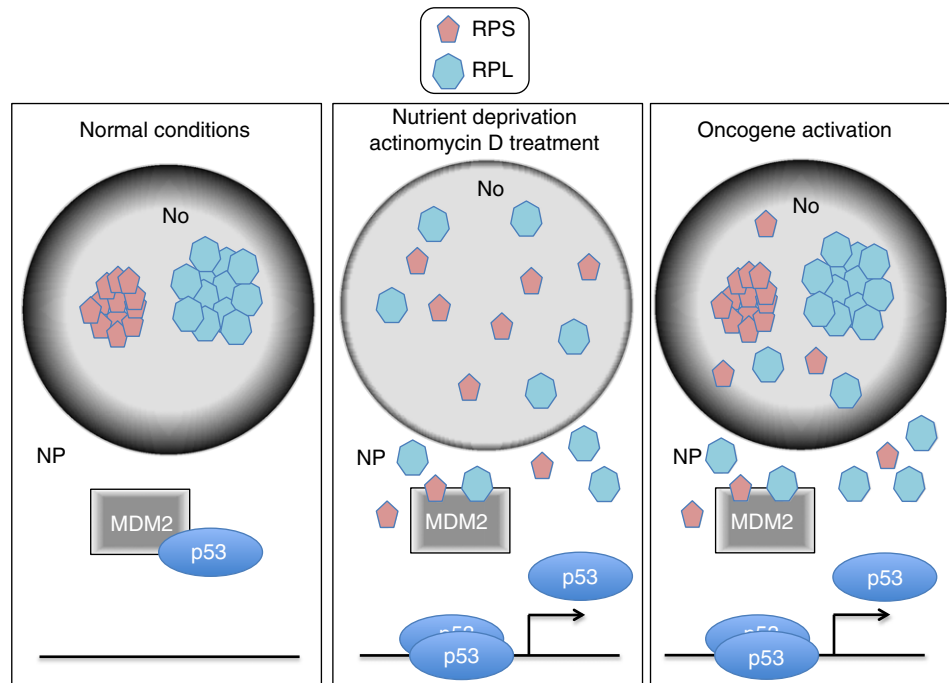


Figure 1 RP-MDM2-p53 pathway model. Ribosomal proteins of the large (blue) and small (red) subunits normally interact within the nucleolus to form the large and small subunits. Upon nutrient deprivation or inhibition of rRNA synthesis by ActD treatment, the stoichiometry of ribosomal proteins and nucleolar structure is interrupted allowing free RPs to bind MDM2 in the nucleoplasm to activate p53. Oncogene activation causes the overproduction of RPs, which can then bind MDM2 and activate p53. NO, nucleolus; NP, nucleoplasm.

RPL11-dependent manner (Golomb *et al.*, 2012). Collectively, studies over the previous decade have demonstrated that disruptions at any step in the ribosome assembly pathway are sufficient to activate p53, which strongly supports a link between p53 and ribosome biogenesis. Interestingly, 5.8S rRNA was shown to covalently bind p53 at a phosphoserine residue near the N-terminus over twenty years ago further linking p53 to ribosome biogenesis; however, the *in vivo* role and significance of this interaction remains unknown (Fontoura *et al.*, 1992).

MDM2 Binds Ribosomal Proteins

MDM2 has been shown to bind several ribosomal proteins, including L5, L11, L23, L26, L37, S3, S7, S14, S15, S20, S25, S26, and S27. Ribosomal proteins L5 and L11 are thought to play a central role in p53 activation in response to ribosomal stress, as an imbalance in the stoichiometry of these ribosomal proteins is sufficient to activate the p53 stress response (Golomb *et al.*, 2014). Two main models of RPL11-MDM2 association have been suggested. The first model proposes that disruption of the nucleolus releases all ribosomal proteins into the nucleoplasm where the basic ribosomal proteins are free to bind the acidic domain of MDM2 in accordance with their binding affinity characteristics. A second theory suggests that ribosomal stress could cause an increase in the cytoplasmic pool of free ribosomal proteins either through increased translation of RPL5 and RPL11 or through the liberation of RPL5 and RPL11 from previously assembled ribosomes, after which these proteins are imported into the nucleus where they

can bind MDM2 (Deisenroth and Zhang, 2011). Ribosomal proteins also activate p53 through mechanisms independent of MDM2 association. For example, RPL26 binds to the untranslated regions of p53 mRNA and augments translation through the inhibition of the translation inhibitory nucleolin-p53 mRNA association (Chen *et al.*, 2012). Moreover, RPL11 has been shown to bind p53 response elements upon low levels of actinomycin D treatment, which results in an increase in p53-mediated transcription through the recruitment of the p53 co-activator p300/CBP (Mahata *et al.*, 2012).

Ribosomal proteins play an irreplaceable role in protein synthesis and the p53 stress response, therefore, the fact that deletion of RPS6, RPS19, or RPL11 is embryonic lethal in a number of model organisms is not surprising (Liu and Zhang, 2014). Mice that are haploinsufficient for RPL24 are viable, but these mice display developmental deformities referred to as the belly spot and tail phenotype, which is dependent on p53 expression. Surprisingly, RPL24 heterozygosity on a p53-null background results in embryonic lethality, which suggests that another layer of complexity exists in the interplay between ribosomal proteins and the p53 stress response (Barkic *et al.*, 2009). RPL11- and RPL5-MDM2 association occurs primarily in the central zinc finger domain of MDM2 and partially overlaps with the binding region of the tumor suppressor ARF (Honda and Yasuda, 1999). Analysis of an MDM2 knock-in mouse model harboring the human cancer-associated mutation C305F has been useful in extending our understanding on the role of the RP-MDM2-p53 pathway *in vivo*. Based on the observation that several human tumor mutations selectively target the central zinc finger of MDM2, the C305F mutation has been shown to selectively prevent L5 and L11 binding to

MDM2. Although we lack definitive proof, the fact that L5/L11 binding is selectively disrupted in tumors suggests that RP-MDM2 binding plays an important role in tumor suppression. Importantly, in the C305F mouse, the p53 response to DNA damage remains intact; however, ribosomal stress induced by low-level actinomycin D or starvation is unable to activate p53 because of the inability of MDM2 C305F to bind RPL5 and RPL11. This mouse model was further shown to harbor accelerated *Eμ*-Myc-induced lymphomagenesis, indicating that the RP-MDM2-p53 plays a role in preventing tumorigenesis independently of ARF (Macias *et al.*, 2010). Additional studies on the C305F mouse model have also shown that the RP-MDM2-p53 pathway is vital in sensing nutrient deprivation and activating lipid catabolism pathways in the mouse liver in response to acute starvation stress (Liu *et al.*, 2014). Thus, the C305F mouse model represents a unique opportunity to directly study the physiological role of the RP-MDM2-p53 pathway *in vivo*.

The RP-MDM2-p53 Pathway as a Putative Target in Cancer Therapy

Because cancer cells require a high rate of ribosome biogenesis and because approximately 50% of tumors retain wild-type p53, targeted inhibition of ribosome biogenesis could represent a valuable therapeutic option for many cancer patients. Drugs that specifically inhibit RNA pol I can reactivate wild-type p53 through the ribosomal stress pathway while simultaneously inhibiting ribosome biogenesis required for rapid cell growth. This two-fold tumor suppressive effect may be more effective than some of the current therapies. The small molecule CX-5461 inhibits SL1, which is a necessary cofactor for the RNA pol I transcription complex and correspondingly for rRNA transcription (Drygin *et al.*, 2011). The small molecule BMH-21 promotes the proteasome dependent degradation of RPA194, which is a member of the RNA pol I initiation complex, thereby inhibiting ribosome biogenesis without triggering the DNA damage response (Peltonen *et al.*, 2014). The effectiveness of these compounds in tumors expressing wild-type p53 is currently unknown, and further studies on this promising therapeutic avenue are needed.

Conclusions

Ribosome biogenesis is an energetically and logistically demanding process that is tightly regulated to facilitate controlled cell growth and survival. Mutations in the components associated with proper ribosome assembly are associated with ribosomopathies that display a diverse array of phenotypes that cannot be fully explained through altered protein synthesis. These observations suggest that ribosomal proteins exert several extra-ribosomal functions. In this section, we examined studies using a wide range of model organisms and techniques in an effort to obtain an overarching view of the function of ribosomal proteins outside of their niche in translation. We discussed some of the wide-ranging effects of the ribosome network on the cell with special attention to the

connection between ribosome biogenesis and the tumor suppressor p53.

See also: Cell Division/Death: Regulation of Cell Growth: Eukaryotic Ribosome Assembly and Export; Ribosomes and Stress – Linked from Birth to Death

References

- Ayrault, O., Andrique, L., Fauvin, D., *et al.*, 2006. Human tumor suppressor p14ARF negatively regulates rRNA transcription and inhibits UBF1 transcription factor phosphorylation. *Oncogene* 25, 7577–7586.
- Barkic, M., Crnomarkovic, S., Grabusic, K., *et al.*, 2009. The p53 tumor suppressor causes congenital malformations in Rpl24-deficient mice and promotes their survival. *Molecular and Cellular Biology* 29, 2489–2504.
- Barna, M., Pusic, A., Zollo, O., *et al.*, 2008. Suppression of Myc oncogenic activity by ribosomal protein haploinsufficiency. *Nature* 456, 971–975.
- Berghofer-Hochheimer, Y., Zurek, C., Wolf, S., Hemmerich, P., Munder, T., 1998. L7 protein is a coregulator of vitamin D receptor-retinoid X receptor-mediated transactivation. *Journal of Cellular Biochemistry* 69, 1–12.
- Blumenthal, T., Carmichael, G.G., 1979. RNA replication: Function and structure of Qbeta-replicase. *Annual Review of Biochemistry* 48, 525–548.
- Boisvert, F.M., Van Koningsbruggen, S., Navascues, J., Lamond, A.I., 2007. The multifunctional nucleolus. *Nature Reviews Molecular Cell Biology* 8, 574–585.
- Bursac, S., Brdovcak, M.C., Donati, G., Volarevic, S., 2014. Activation of the tumor suppressor p53 upon impairment of ribosome biogenesis. *Biochimica et Biophysica Acta* 1842, 817–830.
- Cairns, C.A., White, R.J., 1998. p53 is a general repressor of RNA polymerase III transcription. *EMBO Journal* 17, 3112–3123.
- Chen, J., Guo, K., Kastan, M.B., 2012. Interactions of nucleolin and ribosomal protein L26 (RPL26) in translational control of human p53 mRNA. *Journal of Biological Chemistry* 287, 16467–16476.
- Dai, M.S., Arnold, H., Sun, X.X., Sears, R., Lu, H., 2007. Inhibition of c-Myc activity by ribosomal protein L11. *EMBO Journal* 26, 3332–3345.
- Deisenroth, C., Zhang, Y., 2011. The ribosomal protein-Mdm2-p53 pathway and energy metabolism: Bridging the gap between feast and famine. *Genes & Cancer* 2, 392–403.
- Drygin, D., Lin, A., Bliesath, J., *et al.*, 2011. Targeting RNA polymerase I with an oral small molecule CX-5461 inhibits ribosomal RNA synthesis and solid tumor growth. *Cancer Research* 71, 1418–1430.
- Ebert, B.L., Pretz, J., Bosco, J., *et al.*, 2008. Identification of RPS14 as a 5q-syndrome gene by RNA interference screen. *Nature* 451, 335–339.
- Elbashir, S.M., Lendeckel, W., Tuschl, T., 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes & Development* 15, 188–200.
- Fire, A., Xu, S., Montgomery, M.K., *et al.*, 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806–811.
- Fok, V., Mitton-Fry, R.M., Grech, A., Steitz, J.A., 2006. Multiple domains of EBER 1, an Epstein-Barr virus noncoding RNA, recruit human ribosomal protein L22. *RNA* 12, 872–882.
- Fontoura, B.M., Sorokina, E.A., David, E., Carroll, R.B., 1992. p53 is covalently linked to 5.8S rRNA. *Molecular and Cellular Biology* 12, 5145–5151.
- Freed, E.F., Bleichert, F., Dutca, L.M., Baserga, S.J., 2010. When ribosomes go bad: Diseases of ribosome biogenesis. *Molecular BioSystems* 6, 481–493.
- Gazda, H.T., Sheen, M.R., Vlachos, A., *et al.*, 2008. Ribosomal protein L5 and L11 mutations are associated with cleft palate and abnormal thumbs in Diamond-Blackfan anemia patients. *American Journal of Human Genetics* 83, 769–780.
- Golomb, L., Bublik, D.R., Wilder, S., *et al.*, 2012. Importin 7 and exportin 1 link c-Myc and p53 to regulation of ribosomal biogenesis. *Molecular Cell* 45, 222–232.
- Golomb, L., Volarevic, S., Oren, M., 2014. p53 and ribosome biogenesis stress: The essentials. *FEBS Letters* 588, 2571–2579.
- Gomez-Roman, N., Grandori, C., Eisenman, R.N., White, R.J., 2003. Direct activation of RNA polymerase III transcription by c-Myc. *Nature* 421, 290–294.
- Grandori, C., Gomez-Roman, N., Felton-Edkins, Z.A., *et al.*, 2005. c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. *Nature Cell Biology* 7, 311–318.

- Hagner, P.R., Mazan-Mamczarz, K., Dai, B., *et al.*, 2011. Ribosomal protein S6 is highly expressed in non-Hodgkin lymphoma and associates with mRNA containing a 5' terminal oligopyrimidine tract. *Oncogene* 30, 1531–1541.
- Hannan, K.M., Kennedy, B.K., Cavanaugh, A.H., *et al.*, 2000. RNA polymerase I transcription in confluent cells: Rb downregulates rDNA transcription during confluence-induced cell cycle arrest. *Oncogene* 19, 3487–3497.
- Holzel, M., Rohrmoser, M., Schlee, M., *et al.*, 2005. Mammalian WDR12 is a novel member of the Pes1-Bop1 complex and is required for ribosome biogenesis and cell proliferation. *Journal of Cell Biology* 170, 367–378.
- Honda, R., Yasuda, H., 1999. Association of p19^{ARF} with MDM2 inhibits ubiquitin ligase activity of MDM2 for tumor suppressor p53. *EMBO Journal* 18, 22–27.
- Hu, W., Feng, Z., Levine, A.J., 2012. The regulation of multiple p53 stress responses is mediated through MDM2. *Genes & Cancer* 3, 199–208.
- Iapalucci-Espinoza, S., Franze-Fernandez, M.T., 1979. Effect of protein synthesis inhibitors and low concentrations of actinomycin D on ribosomal RNA synthesis. *FEBS Letters* 107, 281–284.
- Imafuku, I., Masaki, T., Waragai, M., *et al.*, 1999. Presenilin 1 suppresses the function of c-Jun homodimers via interaction with QM/Jif-1. *Journal of Cell Biology* 147, 121–134.
- Itahana, K., Bhat, K.P., Jin, A., *et al.*, 2003. Tumor suppressor ARF degrades B23, a nucleolar protein involved in ribosome biogenesis and cell proliferation. *Molecular Cell* 12, 1151–1164.
- Jang, C.Y., Lee, J.Y., Kim, J., 2004. RpS3, a DNA repair endonuclease and ribosomal protein, is involved in apoptosis. *FEBS Letters* 560, 81–85.
- Khanna, N., Sen, S., Sharma, H., Singh, N., 2003. S29 ribosomal protein induces apoptosis in H520 cells and sensitizes them to chemotherapy. *Biochemical and Biophysical Research Communications* 304, 26–35.
- Kim, T.S., Jang, C.Y., Kim, H.D., *et al.*, 2006. Interaction of Hsp90 with ribosomal proteins protects from ubiquitination and proteasome-dependent degradation. *Molecular Biology of the Cell* 17, 824–833.
- Laplanche, M., Sabatini, D.M., 2012. mTOR signaling in growth control and disease. *Cell* 149, 274–293.
- Lempiäinen, H., Shore, D., 2009. Growth control and ribosome biogenesis. *Current Opinion in Cell Biology* 21, 855–863.
- Li, J.J., Xie, D., 2014. RACK1, a versatile hub in cancer. *Oncogene* 1–9. doi:10.1038/onc.2014.127.
- Li, P., Wang, D., Li, H., *et al.*, 2014. Identification of nucleolus-localized PTEN and its function in regulating ribosome biogenesis. *Molecular Biology Reports* 41, 6383–6390.
- Li, T., Kon, N., Jiang, L., *et al.*, 2012. Tumor suppression in the absence of p53-mediated cell-cycle arrest, apoptosis, and senescence. *Cell* 149, 1269–1283.
- Liu, Y., He, Y., Jin, A., *et al.*, 2014. Ribosomal protein-Mdm2-p53 pathway coordinates nutrient stress with lipid metabolism by regulating MCD and promoting fatty acid oxidation. *Proceedings of the National Academy of Sciences of the United States of America* 111, E2414–E2422.
- Liu, Y., Zhang, Y., 2014. Nucleolar signaling determines cell fate: The RP-Mdm2-p53 axis fine-tunes cellular homeostasis. In: Kumar, R. (Ed.), *Nuclear Signaling Pathways and Targeting Transcription in Cancer*. New York, NY: Springer.
- Longley, D.B., Harkin, D.P., Johnston, P.G., 2003. 5-fluorouracil: Mechanisms of action and clinical strategies. *Nature Reviews Cancer* 3, 330–338.
- Luo, X., Hsiao, H.H., Bubunenko, M., *et al.*, 2008. Structural and functional analysis of the *E. coli* NusB-S10 transcription antitermination complex. *Molecular Cell* 32, 791–802.
- Macias, E., Jin, A., Deisenroth, C., *et al.*, 2010. An ARF-independent c-MYC-activated tumor suppression pathway mediated by ribosomal protein-Mdm2 interaction. *Cancer Cell* 18, 231–243.
- MacInnes, A.W., Amsterdam, A., Whittaker, C.A., Hopkins, N., Lees, J.A., 2008. Loss of p53 synthesis in zebrafish tumors with ribosomal protein gene mutations. *Proceedings of the National Academy of Sciences of the United States of America* 105, 10408–10413.
- Mahata, B., Sundqvist, A., Xirodimas, D.P., 2012. Recruitment of RPL11 at promoter sites of p53-regulated genes upon nucleolar stress through NEDD8 and in an Mdm2-dependent manner. *Oncogene* 31, 3060–3071.
- Mayer, C., Grummt, I., 2006. Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. *Oncogene* 25, 6384–6391.
- Menssen, A., Hermeking, H., 2002. Characterization of the c-MYC-regulated transcriptome by SAGE: Identification and analysis of c-MYC target genes. *Proceedings of the National Academy of Sciences of the United States of America* 99, 6274–6279.
- Montecarlo, F.S., Vogt, P.K., 1993. A Jun-binding protein related to a putative tumor suppressor. *Proceedings of the National Academy of Sciences of the United States of America* 90, 6726–6730.
- Nateri, A.S., Spencer-Dene, B., Behrens, A., 2005. Interaction of phosphorylated c-Jun with TCF4 regulates intestinal cancer development. *Nature* 437, 281–285.
- Ni, J.Q., Liu, L.P., Hess, D., Rietdorf, J., Sun, F.L., 2006. Drosophila ribosomal proteins are associated with linker histone H1 and suppress gene transcription. *Genes & Development* 20, 1959–1973.
- Pei, D., Zhang, Y., Zheng, J., 2012. Regulation of p53: A collaboration between Mdm2 and Mdmx. *Oncotarget* 3, 228–235.
- Peltonen, K., Colis, L., Liu, H., *et al.*, 2014. A targeting modality for destruction of RNA polymerase I that possesses anticancer activity. *Cancer Cell* 25, 77–90.
- Pestov, D.G., Strezoska, Z., Lau, L.F., 2001. Evidence of p53-dependent cross-talk between ribosome biogenesis and the cell cycle: Effects of nucleolar protein Bop1 on G(1)/S transition. *Molecular and Cellular Biology* 21, 4246–4255.
- Purvis, J.E., Karhohs, K.W., Mock, C., *et al.*, 2012. p53 dynamics control cell fate. *Science* 336, 1440–1444.
- Rahl, P.B., Lin, C.Y., Seila, A.C., *et al.*, 2010. c-Myc regulates transcriptional pause release. *Cell* 141, 432–445.
- van Riggelen, J., Yetil, A., Felsner, D.W., 2010. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nature Reviews Cancer* 10, 301–309.
- Robledo, S., Idol, R.A., Crimmins, D.L., *et al.*, 2008. The role of human ribosomal proteins in the maturation of rRNA and ribosome production. *RNA* 14, 1918–1929.
- Roux, P.P., Shahbazian, D., Vu, H., *et al.*, 2007. RAS/ERK signaling promotes site-specific ribosomal protein S6 phosphorylation via RSK and stimulates cap-dependent translation. *Journal of Biological Chemistry* 282, 14056–14064.
- Rubbi, C.P., Milner, J., 2003. Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. *EMBO Journal* 22, 6068–6077.
- Schadt, E.E., Molony, C., Chudin, E., *et al.*, 2008. Mapping the genetic architecture of gene expression in human liver. *PLoS Biology* 6, e107.
- Sengupta, J., Nilsson, J., Gursky, R., *et al.*, 2004. Identification of the versatile scaffold protein RACK1 on the eukaryotic ribosome by cryo-EM. *Nature Structural & Molecular Biology* 11, 957–962.
- Singh, D., Chang, S.J., Lin, P.H., *et al.*, 2009. Regulation of ribonuclease E activity by the L4 ribosomal protein of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* 106, 864–869.
- Szymanski, M., Barciszewska, M.Z., Erdmann, V.A., Barciszewski, J., 2003. 5S rRNA: Structure and interactions. *Biochemical Journal* 371, 641–651.
- Taskinen, M., Ranki, A., Pukkala, E., *et al.*, 2008. Extended follow-up of the Finnish cartilage-hair hypoplasia cohort confirms high incidence of non-Hodgkin lymphoma and basal cell carcinoma. *American Journal of Medical Genetics Part A* 146A, 2370–2375.
- Torres, M., Condon, C., Balada, J.M., Squires, C., Squires, C.L., 2001. Ribosomal protein S4 is a transcription factor with properties remarkably similar to NusA, a protein involved in both non-ribosomal and ribosomal RNA antitermination. *EMBO Journal* 20, 3811–3820.
- Wan, F., Anderson, D.E., Barnitz, R.A., *et al.*, 2007. Ribosomal protein S3: A KH domain subunit in NF-kappaB complexes that mediates selective gene regulation. *Cell* 131, 927–939.
- Wanzel, M., Russ, A.C., Kleine-Kohlbrecher, D., *et al.*, 2008. A ribosomal protein L23-nucleophosmin circuit coordinates Miz1 function with cell growth. *Nature Cell Biology* 10, 1051–1061.
- Warner, J.R., 1999. The economics of ribosome biosynthesis in yeast. *Trends in Biochemical Sciences* 24, 437–440.
- Warner, J.R., McIntosh, K.B., 2009. How common are extraribosomal functions of ribosomal proteins? *Molecular Cell* 34, 3–11.
- Watson, K.L., Konrad, K.D., Woods, D.F., Bryant, P.J., 1992. Drosophila homolog of the human S6 ribosomal protein is required for tumor suppression in the hematopoietic system. *Proceedings of the National Academy of Sciences of the United States of America* 89, 11302–11306.
- Wei, S., Chen, X., McGraw, K., *et al.*, 2013. Lenalidomide promotes p53 degradation by inhibiting MDM2 auto-ubiquitination in myelodysplastic syndrome with chromosome 5q deletion. *Oncogene* 32, 1110–1120.
- Zhai, W., Comai, L., 2000. Repression of RNA polymerase I transcription by the tumor suppressor p53. *Molecular and Cellular Biology* 20, 5930–5938.
- Zhang, C., Comai, L., Johnson, D.L., 2005. PTEN represses RNA polymerase I transcription by disrupting the SL1 complex. *Molecular and Cellular Biology* 25, 6899–6911.
- Zhang, Y., Lu, H., 2009. Signaling to p53: Ribosomal proteins find their way. *Cancer Cell* 16, 369–377.

Ribosomal RNA Processing

K Umaer, M Ciganda, and N Williams, University at Buffalo, Buffalo, NY, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

In the early 1960s, a group at MIT led by James Darnell was trying to characterize mRNA in mammalian cells when they found evidence for 'giant RNAs' in the dense fractions of sucrose gradients (Scherrer *et al.*, 1963). These 45S and 35S species were later proven to be metabolic precursors of ribosomal RNAs. The question of how these precursors gave rise to discrete ribosomal RNA species was hotly debated in the following decades (reviewed in Scherrer, 2003). Twenty years later, Thomas Cech, at University of Colorado, Boulder was studying the maturation of an rRNA precursor in *Tetrahymena* when he discovered that this particular RNA could catalyze its own processing in the absence of protein, a major breakthrough in molecular biology (Cech *et al.*, 1981).

Ribosomes are composed of two subunits. In prokaryotes, the large ribosomal subunit (LSU or 50S) contains the 23S rRNA and the 5S rRNA as well as over 30 proteins, and the

small ribosomal subunit (SSU or 30S) contains 16S rRNA and 21 proteins. In eukaryotes, the large, 60S subunit (LSU) contains 3 rRNA molecules (25/28S, 5.8S, and 5S) (Figure 1) as well as over 40 proteins. The small, 40S subunit (SSU) contains the 18S rRNA and over 30 proteins. Size variation of rRNAs among different species is due partly to expansion segments, which are sequence insertions in divergent regions of the rRNA ancestor. Over 200 accessory factors and 75 small nucleolar RNAs associate with the pre-ribosome throughout the process of ribosomal biogenesis.

Transcription of ribosomal RNAs usually involves two compartments: the nucleolus for the major rDNA unit and the nucleoplasm for 5S rRNA. Transcription of the major rDNA unit by RNA polymerase I generates a large precursor, 35S in yeast. 5S rRNA is normally transcribed by RNA polymerase III in the nucleoplasm with the exception of *Saccharomyces cerevisiae* where it is synthesized in the nucleolus. Association of the major precursor with 5S rRNA, ribosomal proteins,

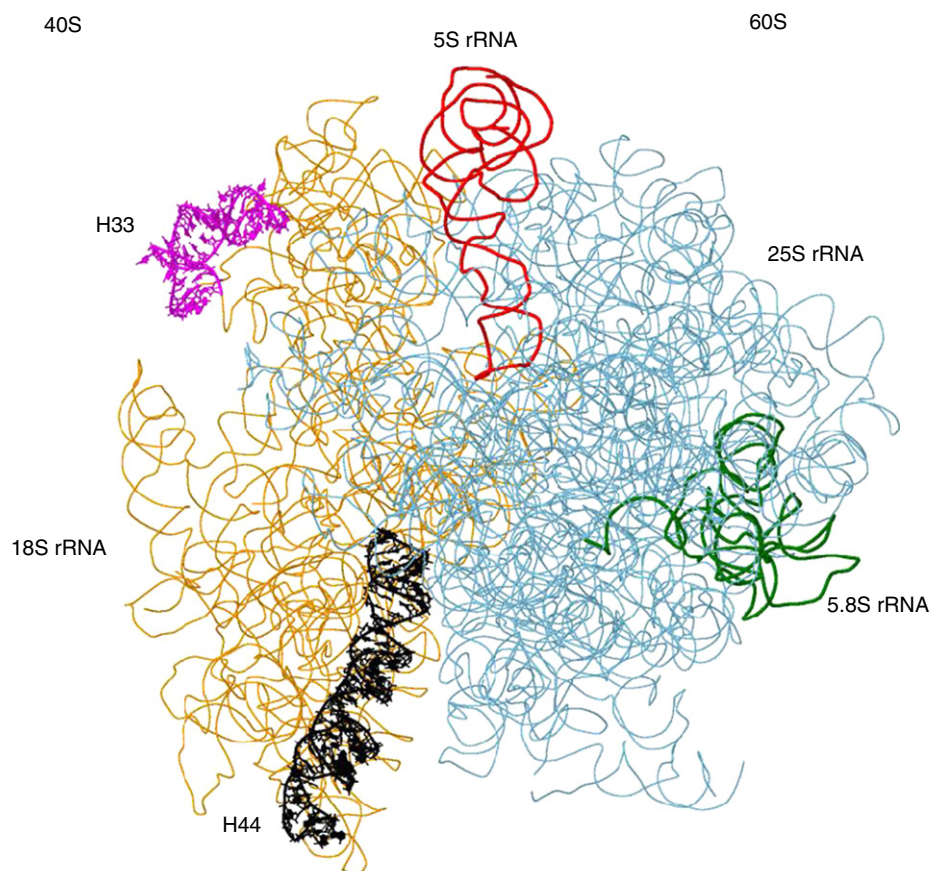


Figure 1 Structure of *S. cerevisiae* mature ribosomal RNAs. The large subunit 25S rRNA is represented in light blue, 5.8S rRNA is depicted in green, and 5S rRNA is in red. The small subunit 18S rRNA is represented in orange. Helices in 18S rRNA important during processing and discussed in the text are shown in magenta (H33) and black (H44). H33 constitutes the 'beak' that forms during cytoplasmic maturation of 18S rRNA. Site D lies within H44 and is formed only after processing of the ITS1. PDB accession number: 3j77.

accessory factors, and processing complexes in the nucleolus constitutes the 90S pre-ribosome. This 90S complex is processed by cleavage of the 35S precursor, generating a 66S precursor of the 60S subunit, and a 43S precursor of the 40S subunit. These particles continue to mature as they transit from the nucleolus to the nucleoplasm. They are then exported to the cytoplasm where they become mature 60S and 40S subunits.

The functional importance of these reactions and the question of why so much energy is invested in transcribing 'wasteful' RNA sequences has been addressed in recent years with the generation of mutant strains lacking in specific enzymes in the processing cascade. For at least for a few 'model' organisms, we now have a clear picture of rRNA processing that details specific sequential modifications and their corresponding enzymatic complexes.

Ribosomal RNA Processing in Yeast

Cleavage and Folding of Precursor rRNA

Ribosomal RNA processing has been studied extensively in the budding yeast, *S. cerevisiae* (reviewed in Woolford and Baserga, 2013; Nazar, 2004). In this organism, the major unit of the ribosomal RNA genes (rDNA) is organized in tandem consisting of the SSU gene (18S) at the 5' end followed by two LSU genes (5.8S and 25S) at the 3' end (Figure 2). There are approximately 150 repeats of this unit on chromosome XII. Not all copies are transcriptionally active at any given time, but the redundancy protects against genomic instability. The extra sequences at both ends of each tandem repeat that are transcribed but are not present in the final products are the external transcribed spacers, 5'ETS and 3'ETS. The sequences separating the rDNA genes which are also transcribed but excised during processing are the internal transcribed spacers, ITS1 (between 18S and 5.8S) and ITS2 (between 5.8S and 25S). Exonucleolytic and endonucleolytic activities are required to remove these sequences as indicated in Figure 2 (letters A to E, with subscripts added based on the overall chronology of events). Processing also involves chemical modification of bases at specific positions (as described below).

Cleavage and modification of the precursor can start during transcription with about 80% of precursors being co-transcriptionally processed in yeast. The nascent RNA associates first with a series of ribonucleoprotein complexes collectively known as the SSU processome (Perez-Fernandez et al., 2007; Bernstein et al., 2004) in the nucleolus, required for the biogenesis of 18S rRNA. These RNPs have been classified according to their composition and time of assembly into several subcomplexes: the tUTP complex, the bUTP complex, the U3 snoRNP, the cUTP complex, the Bms/Rcl1 complex, and the Mpp10 complex. Later association with a different complex known as the A₃ cluster is necessary for processing of the LSU rRNAs. The A₃ cluster comprises the Pwp1 subcomplex, 6 interdependent A₃ factors, and 2 Dead-box proteins. For a comprehensive list of yeast protein factors involved in ribosome assembly, including processing, see (Woolford and Baserga, 2013).

For a majority of transcripts, once the polymerase has transcribed ITS1, the first processing events occur on sites denoted A₀ (on the distal end of the 5'ETS), A₁ (on the proximal end of the 5'ETS), and A₂ (in ITS1) (Figure 2). The last cleavage at site A₂ separates the 18S precursor (20S) from the remaining nascent transcript. This second transcript, containing a portion of ITS1, 5.8S, ITS2, 25S, and 3'ETS, is released by cotranscriptional cleavage by the RNase III Rnt1p at the B₀ site on the 3'ETS. Processing of the 3' end of this precursor to site B₂ generates 27SA₂. Those transcripts that are not co-transcriptionally processed at A₂ constitute the 35S pool. This 35S precursor is processed at the A₀, A₁, A₂, and B₂ sites, also generating 20S and 27SA₂.

The enzyme responsible for the cleavages at A₀, A₁, and A₂ is Rcl1p. This protein does not have a discernible *in vitro* activity on RNA, although it shares homology with 3' terminal phosphate cyclase enzymes. Rcl1p associates with Bms1pn, and disruption of this association leads to accumulation of precursor (Delprato et al., 2014). Disrupting an interaction between protein factors Krr1 and Faf1 in the 90S complex impairs cleavage at sites A₀, A₁, and A₂ (Zheng et al., 2014). The C-terminal domain of ribosomal biogenesis factor Rrp5 (Lebaron et al., 2013) is necessary for cleavage at A₀ and A₂ and its N-terminal domain is necessary for cleavage at A₃. Rrp5 and Rok1, a helicase, make direct contact with RNA helices within expansion segment 6 of the SSU (Martin et al., 2014). The majority of 27SA₂ precursors are cleaved at site A₃ by endonuclease MRP (Mitochondrial RNA Processing) generating 27SA₃. Exonucleases Rat1, Rrp17, and Xrn1 act in the 5' to 3' direction to trim this precursor to site B_{1S}, generating intermediate 27SB_S. This process requires Rat1 cofactor Rai1, as well as several other interdependent proteins belonging to the A₃ factors. An alternative fate for 27SA₂ is cleavage by an endonuclease at site B_{1L}, generating a longer intermediate 27SB_L. Eventually these two pathways generate two different pools of 5.8S in yeast cells, differing by ~7 nucleotides in their 5' ends (Henry et al., 1994). Both are incorporated in ribosomes and are competent for translation. Approximately 80% of the 5.8S rRNA precursor gets processed through the 5.8S_S pathway. These alternative forms have also been described in most but not all eukaryotes. Their significance is not clear.

No further cleavage occurs in the nucleolus. Intermediate 20S is transported to the cytoplasm as part of the pre-40S complex and cleaved at site D by Nob1p, assisted by Pno1/Dim2, generating its mature 3' end. An additional event in cytoplasmic maturation of 18S involves rearrangement and formation of the 'beak,' a protrusion of helix 33 (Figure 1). This step involves Hrr25 dependent phosphorylation and dissociation of ribosomal protein Rps3, which is close to helix 33. After formation of the beak, Rps3 is reincorporated to the 40S subunit (Schafer et al., 2006).

RNA folding participates in the regulation of these cleavage events by ensuring the sequential nature of certain steps. Prior to cleavage at A₂, folding of ITS1 prevents formation of helix 44 (Figure 1) at the 3' end of the 18S sequence. Cleavage site D lies within this helix (Lamanna and Karbstein, 2011). Once the A₂ site has been cut, this constraint on helix 44 is released, and Nob1p can recognize site D and generate the mature 3' end of 18S.

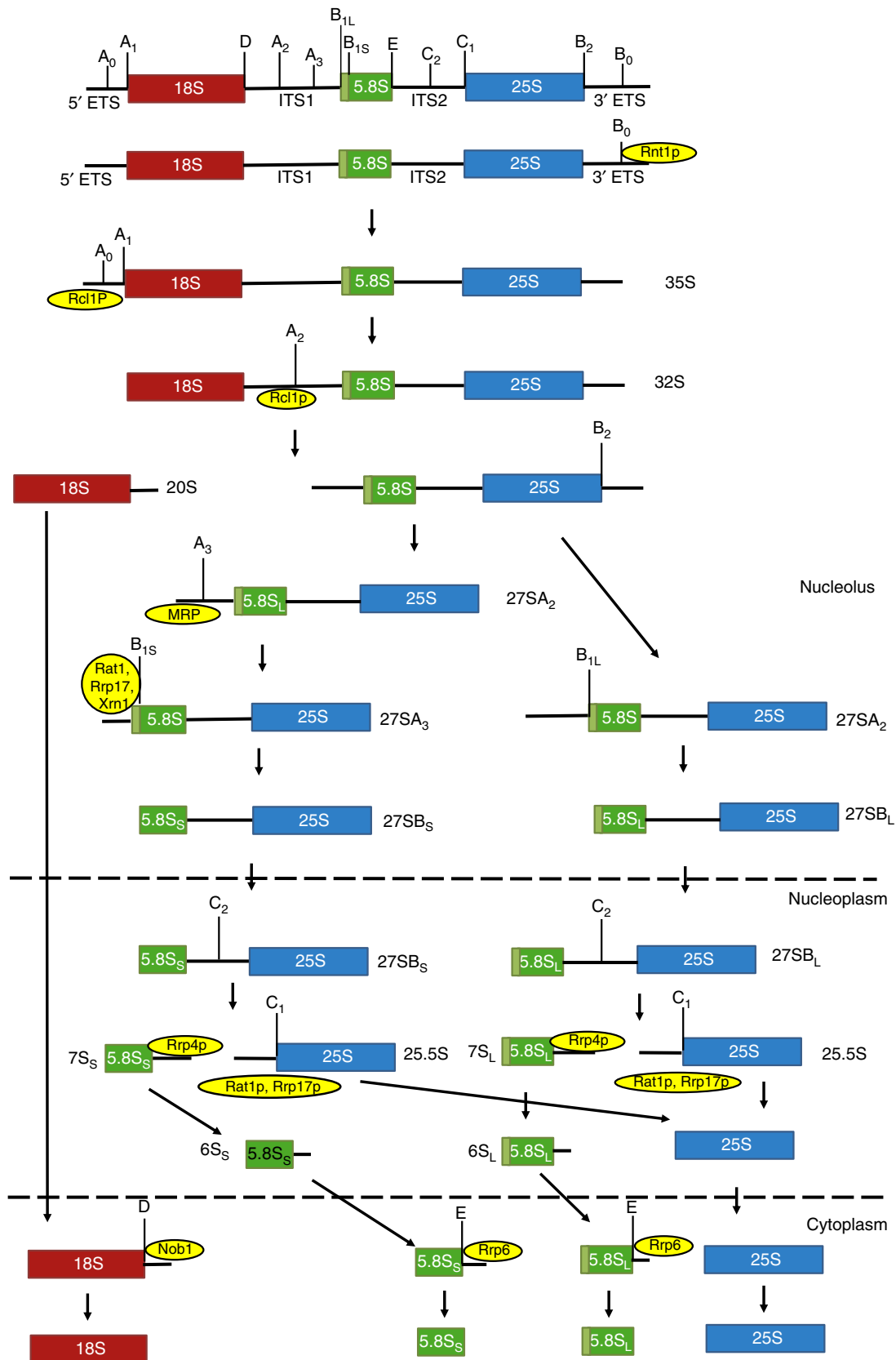


Figure 2 Processing of the major ribosomal RNA unit in *S. cerevisiae*. Sites of cleavage or exonucleolytic trimming are indicated with letters and subscript numbers. When known, enzymes responsible for these events are indicated in yellow ovals. The subcellular localization of different processing events is indicated on the right. There are two forms of 5.8S rRNA, 5.8S_s and 5.8S_L which differ by 7 nucleotides at their 5' ends. Subscript S and L denote small and long form of 5.8S rRNA. These two forms are generated by alternative pathways originating at 27SA₂. 27SB_s generates 5.8S_s and 27SB_L generates 5.8S_L. rRNA precursor species are named based on the sedimentation coefficient and the cut site that is used to generate the species.

Maturation of 27SB_L and 27SB_S continues in the nucleoplasm. A cleavage event within ITS2 at site C₂ separates the 5.8S precursor (7S_L or 7S_S) from the 25S precursor (25.5S). Two conformations of ITS2 are possible: a ring structure and a hairpin structure. The ring structure may form first and facilitate protein association. The hairpin structure contains a stem (III) that is essential for processing. This stem brings together distant sequences and allows for base pairing interactions between 5.8S rRNA and 25S rRNA (Cote *et al.*, 2002).

Exonucleolytic activity by Rat1p and Rrp17p in the 5' to 3' direction generates the mature 5' end of 25S. Several 3' to 5' exonucleases of the exosome like Rrp4p (De La Cruz *et al.*, 1998) act on the 3' end of 7S_L and 7S_S, reducing it to a point 30 bases downstream of site E. Exonuclease Rrp6p processes the remaining 30 nucleotides, generating mature 5.8S_S or 5.8S_L in the cytoplasm (Figure 2; Mitchell *et al.*, 1996).

Chemical Modifications of rRNA

The most common nucleolar modifications of rRNA are the additions of a methyl group at the 2' position of ribose (2'-O-methylation), and the linkage of uracil to ribose via carbon 5 in the pyrimidine ring, instead of nitrogen 1 (pseudouridylation). These modifications take place at specific positions, many of them in the vicinity of functionally relevant areas of the ribosome, such as the peptidyl-transferase center (Baudin-Baillieu *et al.*, 2009). There are several ways in which chemical modification of rRNA affects its biochemical properties. Methylation at the 2'-OH of the ribose moiety makes RNA more resistant to hydrolysis, favoring the 3' endo configuration. Isomerization of uracil changes the code of donor and acceptor groups for hydrogen bonds. Local base stacking can be enhanced by pseudouridylation. These chemical alterations of specific bases affect the efficiency of translation of the mature ribosome (Liang *et al.*, 2009). A number of small nucleolar RNA molecules (snoRNAs) dictates base specificity by hybridizing in the vicinity of the base to be modified (reviewed in Tollervey and Kiss, 1997). This base-pairing positions the enzymatic machinery correctly. Some snoRNPs such as U3 and U14 are essential for cleavage events in the 5'ETS at A₀ and A₁ and in ITS1 at A₂.

The class of snoRNA that leads to 2'-O-methylation possesses a conserved C/D box structure. Box C (RUGAUGA) hybridizes with box D (CUGA) forming the C/D motif. Not all box C/D elements are involved in methylation; some participate in cleavage events during processing and this may be their ancestral function. Box C/D snoRNAs associate with RNA binding protein Snu13, and proteins Nop56, Nop58, and Nop1, the catalytic methyltransferase that uses S-adenosylmethionine as donor (Watkins and Bohnsack, 2012).

snoRNAs that assist in pseudouridylation belong to the H/ACA box type. These snoRNAs fold in a typical hairpin-hinge-hairpin-tail secondary structure. Box H (ANANNA) is located within the hinge region, while box ACA (ACANNN) is at the 3' end of the tail. The rRNA to be modified hybridizes within loops in the hairpin regions. Box H/ACA snoRNAs form a complex with Nhp2 (an RNA binding protein), Nop10, Gar1 (which regulates loading), and Cbf5, the catalytic pseudouridylation synthase (Watkins and Bohnsack, 2012).

Chemical modification independent of snoRNPs also occurs in yeast rRNAs. A single C residue in the terminal helix of 18S rRNA is acetylated by Rra1p. This modification is important for the processing of the SSU precursor (Ito *et al.*, 2014). Two adenosine residues in 18S rRNA are converted to dimethyladenosines by Dim1p in the cytoplasm (Lafontaine *et al.*, 1998). This step is one of the few modifications common to eukaryotes and prokaryotes, and it yields mature 18S rRNA at the end of the processing pathway.

rRNA Processing in Other Eukaryotes

One major difference between *S. cerevisiae* and higher eukaryotes appears to be processing within ITS1. In *Xenopus*, two cleavage sites exist at the 5' and 3' ends of ITS1. Processing of these sites can occur in a different order, so there are different intermediates. One of these intermediates contains the 18S precursor as well as the ITS1 sequence in the 3', whereas a different intermediate contains the 28S precursor and the ITS1 sequence at the 5' end (Borovjagin and Gerbi, 2001, 2005).

In humans and other metazoans, an additional cleavage occurs in the 5'ETS (in the context of the SSU processome), upstream of cleavage sites A₀ and 1, 01 (Figure 3). Cleavage at this site is cotranscriptional and dependant on the assembly of the tUTP complex, requiring the bUTP complex and U3 snoRNP for high efficiency (Sloan *et al.*, 2014). Depletion of components of these complexes affects not only the rate of processing, but also the rate of transcription by polymerase I, indicating a tight connection between transcription and processing at the 5'ETS. The human homolog of Rat1p, XRN2, is involved in termination of transcription of the precursor, as well as in the degradation of fragments generated by cleavage at 01. Following cleavage at 01, a 45S precursor is generated. Cleavage at site 2 in ITS1 separates the precursor molecules for the LSU and SSU rRNAs. Alternative pathways lead to the production of a 41S intermediate via initial cuts at sites A₀ and A₁ on the 5'ETS. An additional intermediate, 18S-E, has also been described. This molecule could be generated by direct endonucleolytic cleavage at site E within the ITS1, or by the action of exonucleases acting on the 3' end of 21S in the 3' to 5' direction (Preti *et al.*, 2013). Although it is known that many factors are required for processing, only two enzymatic activities (5'-3' exonuclease Xrn2 and 3'-5' exonuclease ISG20L2) have been assigned to cleavage steps in human cells (Coute *et al.*, 2008; Wang and Pestov, 2011; Figure 3).

In the protozoan ciliate *Tetrahymena thermophila* the rRNA precursor is capable of self-cleavage to release the intervening sequence (IVS). A guanosine molecule performs a nucleophilic attack on the 5' splice site, creating a free hydroxyl group at the 3' end of the 5' exon. This group participates in a subsequent transesterification reaction at the 3' splice site, releasing the IVS (Price and Cech, 1988; Cech *et al.*, 1981). No protein cofactors are necessary for this reaction as indicated above.

Ribosomal RNA processing in early branching eukaryotes diverges farther away from the yeast pathway. In the parasite *Trypanosoma brucei*, the causative agent of African trypanosomiasis, the LSU rRNA (homolog to 25S rRNA in yeast) is cleaved to produce two large rRNAs (LSU 1 and 2) and four small rRNAs (SR1-4). In addition, the rRNAs of trypanosomes

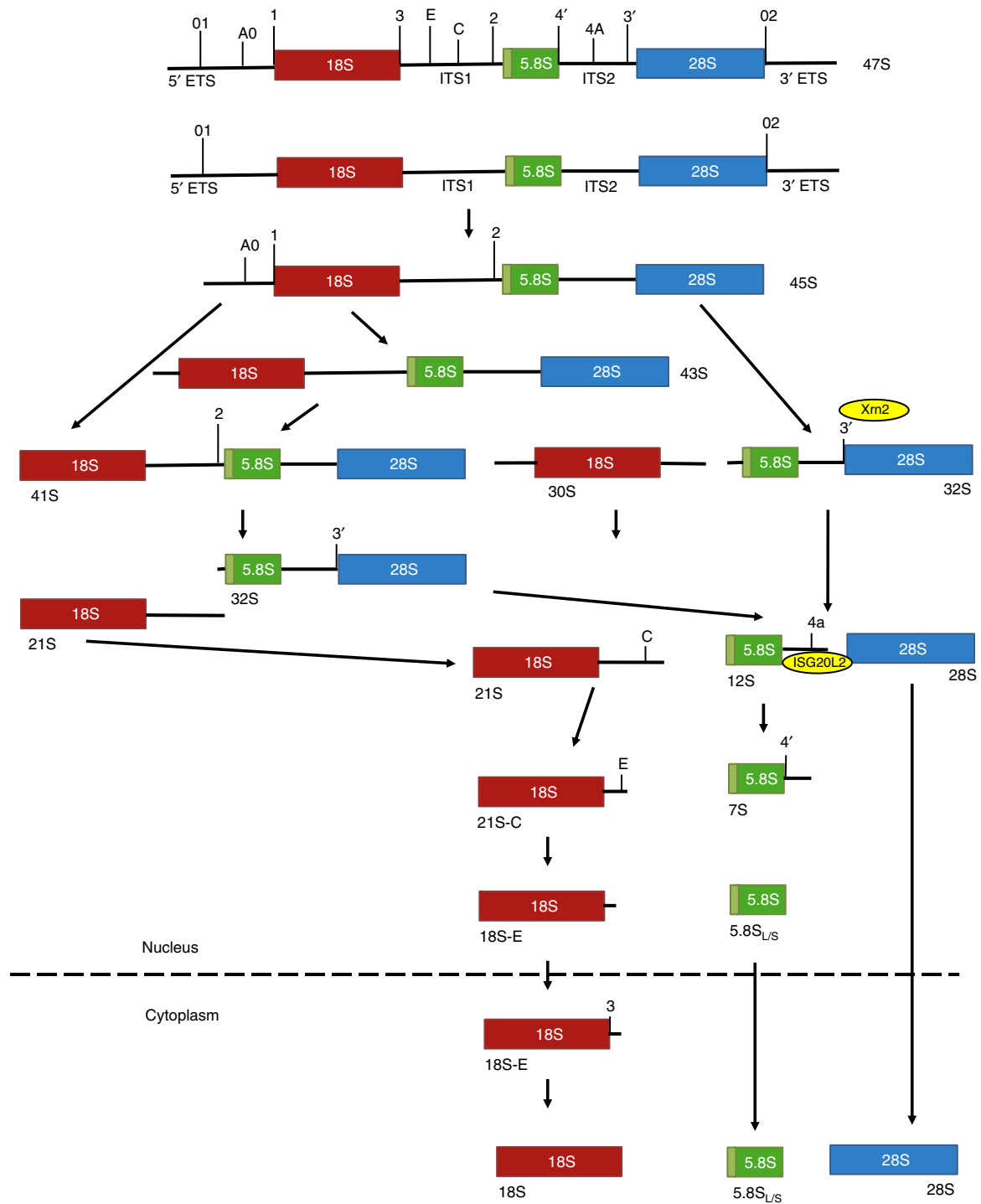


Figure 3 Processing of the major ribosomal RNA unit processing in humans. Sites of cleavage and exonucleolytic trimming are indicated with numbers and letters. An extra 01 processing site exists in the 5'ETS in human rDNA. Known enzymes, responsible for these events are indicated in yellow ovals. The two alternative pathways and their convergence further downstream are shown.

contain unusually large expansion segments (Hashem *et al.*, 2013). Furthermore, the 5.8S rRNA is processed to generate a single species (equivalent to 5.8S_S) instead of the two species present in other organisms (Sakyama *et al.*, 2013). Despite the presence of some peculiarities, ribosomal biogenesis in

T. brucei requires both conserved and trypanosome-specific factors (Umaer *et al.*, 2014). Another example of unusual rRNA processing is the extensive fragmentation of the rRNA into 16 molecules, instead of 4, which occurs in the free living eukaryote *Euglena gracilis* (Schnare and Gray, 1990).

rRNA Processing in Prokaryotes

In prokaryotes, ribosomal RNA genes are organized in tandem and transcribed from the same promoter, generating a primary 30S transcript that contains the sequences for 16S, 23S, and 5S, as well as an internal tRNA in some cases (reviewed in [Shajani et al., 2011; Figure 4](#)). In *Escherichia coli*, this precursor is never observed experimentally in wild type cells, because cleavage by RNase III occurs co-transcriptionally. This initial processing step generates a 17S precursor (16S rRNA with a 115 nucleotide 5' terminal extension and a 33 nucleotide 3' terminal extension), a 23S precursor (with a 7 nucleotide 5' terminal extension and a 9 nucleotide 3' terminal extension), and a 9S precursor (5S rRNA with an 84 nucleotide 5' terminal extension and a 42 nucleotide 3' terminal extension). RNase

III needs double stranded RNA as a substrate ([Court et al., 2013](#)), so folding of the RNA precursor is necessary prior to initial cleavage. Base pairing between the 5' and 3' flanking regions of each rRNA generates the double stranded RNA recognized by RNase III.

The 17S precursor is further modified by RNase E and RNase G in the 5' end ([Li et al., 1999b](#)). Maturation of the 3' end is more complex and involves several 3' to 5' exonucleases: RNase II, RNase R, and polynucleotide phosphorylase (PNPase) ([Sulthana and Deutscher, 2013](#)).

The 23S precursor is cleaved at the 5' end by RNase G ([Song et al., 2011](#)) and at the 3' end by RNase T ([Li et al., 1999a](#)). Premature processing by RNase T is prevented by a CC dinucleotide ([Gutgsell and Jain, 2012](#)). Both RNase T and RNase III are more efficient when their substrates are

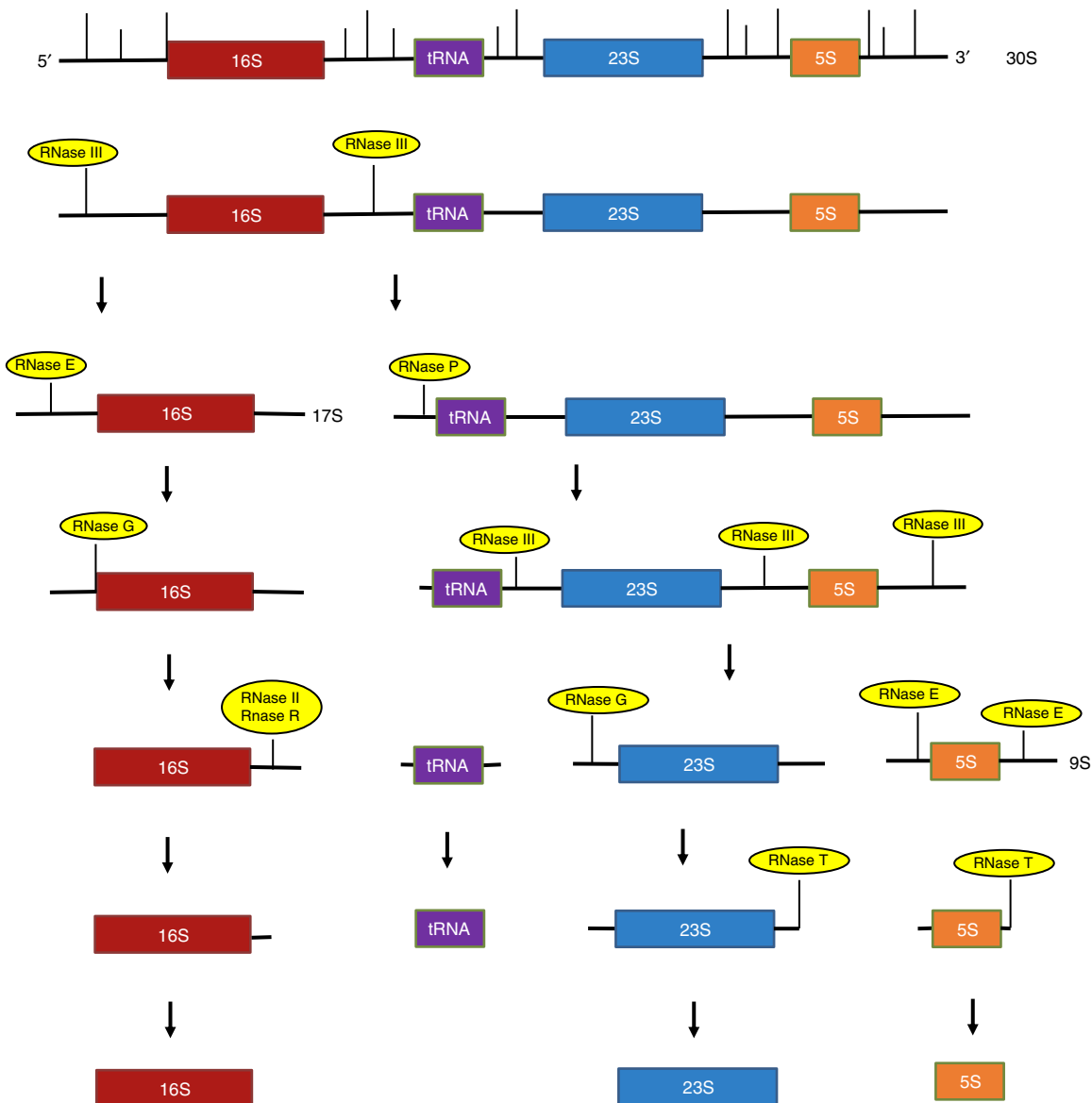


Figure 4 Ribosomal RNA processing in *E. coli*. The initial precursor contains the sequence for 5S rRNA (orange) in addition to 23S (red) and 16S rRNAs (blue), interspersed with a tRNA gene (purple). Cleavage sites are represented by vertical lines and the enzymes known to be responsible for cleavage events are indicated in yellow ovals.

precursors within a ribonucleoprotein complex rather than naked RNAs.

The 9S precursor is cleaved by RNase E (Mackie, 1998) on both ends, 3 nucleotides upstream and downstream of the mature terminal nucleotides. The 3 nucleotides at the 3' end are removed by RNase T. This step bears similarity to the processing of the 3' end of 23S precursor both in the enzyme involved as well as in the nature of the substrate, which in both cases is a short extension adjacent to a stem. The enzyme involved in the processing of the 5' extension of the 5S precursor has not been identified.

Mutant strains of *E. coli* that lack one or more of the enzymes involved in the processing of 16S rRNA are able to assemble ribosomes at the same rate as wild type cells, and these precursor-containing ribosomes are competent for translation (although their fidelity is reduced) (Sulthana and Deutscher, 2013). However, mutants that lack the enzymes involved in the processing of 23S rRNA exhibit a considerable delay in the rate of ribosome assembly, and precursors containing more than a few additional nucleotides at the 5' or 3' ends of 23S rRNA are not efficiently incorporated into ribosomes. It has been proposed that the existence of these extensions acts as a control mechanism linking rRNA processing and ribosome assembly.

Bacterial rRNAs are also chemically modified by methylation and pseudouridylation. However, prokaryotes do not use accessory RNAs equivalent to eukaryotic snoRNAs. Instead, the methyltransferases and pseudouridylate synthases recognize their substrates directly (Hage and Tollervy, 2004). Methylation of 16S rRNA by methyltransferases like ArmA and NpmA (acquired by horizontal DNA transfer) is responsible for resistance to aminoglycoside antibiotics by modifying the A site of the ribosome.

Mitochondria and plastids originated from ancient prokaryotic endosymbionts and, as such, their rRNA processing pathways share some characteristics with the prokaryotic rRNA processing pathways (reviewed in Bohne, 2014). There is great variability in genome organization of mitochondrial DNA across different organisms. In human mitochondria, two RNA species, 16S and 12S, together with approximately 80 nuclear-encoded proteins form the mitochondrial ribosome, which consists of a large 39S subunit and a small 28S subunit. The 16S and 12S genes on the circular mitochondrial genome are separated by a tRNA gene (V) and flanked on both ends by a tRNA gene (L at the 5' and F at the 3' end). The mitochondrial nucleoid associates with a ribosome assembly factor ERAL1 (Dennerlein *et al.*, 2010), proposed to act as a chaperone for nascent ribosomal RNA. A mutation in the tRNA gene between 16S and 23S can disrupt cleavage of the precursor transcript. Enzymes RNase P (Sanchez *et al.*, 2011) and ELAC2 (Brzezniak *et al.*, 2011) have been shown to be involved in the processing of the mitochondrial rRNA precursor.

5S rRNA Processing

5S rRNA is a molecule of approximately 120 nucleotides typically transcribed by RNA polymerase III. In higher eukaryotes, 5S rDNA genes are found in clusters in regions

which are distinct from the major unit rDNA repeats while in *S. cerevisiae* 5S genes are immediately adjacent to the rDNA repeats (reviewed in Ciganda and Williams, 2011). The initial transcript in yeast contains 3' extensions of 7–13 nucleotides. Rex1p is required for efficient processing of this extension (Van Hoof *et al.*, 2000). In the absence of Rex1p, other nucleases can process the 3' end of 5S rRNA with low efficiency. The trimming of the 3' end of 5S rRNA is not essential for yeast viability. An association between 5S rRNA and the La protein assists in the folding and maturation of the 3' ends of 5S rRNA (Wolin and Cedervall, 2002). Ribosomal protein L5 forms a stable complex with 5S rRNA (Lee and Nazar, 1997) prior to its incorporation to the precursor 90S particle. In *E. coli*, RNase T is also essential for the formation of the mature 3' of 5S rRNA (Li and Deutscher, 1995). In archaea, the 5' end of 5S rRNA can also be trimmed by tRNase Z, the same enzyme involved in the generation of the mature 3' end of tRNAs.

Human Diseases Caused by rRNA Processing Defects

rRNA processing is an essential and integral part of ribosome biogenesis. Given the universal importance of ribosomes in cell growth, division, and survival, it is not surprising that ribosomal biogenesis defects lead to severe congenital ribosomopathies (reviewed in Armistead and Triggs-Raine, 2014). The most common features of ribosomopathies include hematological defects, cancer predisposition, and small stature. Although ribosome assembly is an essential process in all cells, these diseases are often tissue specific. This fact highlights the underlying variability that exists in rRNA processing alternative pathways in mammalian cells, and how these pathways intersect with other developmentally regulated cellular pathways. In most cases, defects in processing lead to decreased rates of protein synthesis, and a decrease in cell proliferation. Although defects in large number of ribosomal proteins and ribosomal biogenesis factors lead to defects in rRNA maturation, here we discuss ribosomopathies caused by mutations in components known to be directly responsible for rRNA processing.

A number of disorders are caused by mutations in RMRP gene (Ridanpaa *et al.*, 2001, 2002) which encodes the RNA component of the MRP endoribonuclease. This snoRNP localizes to both mitochondria and the nucleolus and is involved in cleavage of mRNA and rRNA. The ortholog of RMRP in yeast is called *nme1*. Mutation of this gene affects endonucleolytic cleavage of 5.8S rRNA at the ITS-1 A3 site and as a result, 60S ribosomal assembly is disrupted (Schmitt and Clayton, 1993). In addition to rRNA processing defects, mutations in this gene can also affect cyclin B mRNA degradation, which leads to mitotic delay (Thiel *et al.*, 2005). Thus, mutations in *nme1* can impair both protein synthesis and cell cycle regulation.

Cartilage hair hypoplasia (CHH, MIM 250250), also known as metaphysical chondrodysplasia is one of these diseases. CHH, an autosomal recessive disorder, is a form of short limbed dwarfism with a height range of adults from 104 to 149 cm. The disease is also characterized by fine, sparse hair, defective T and/or B cell immunity, anemia, gastrointestinal anomalies, and susceptibility to cancer (Ridanpaa *et al.*, 2001;

Makitie and Kaitila, 1993). CHH was first described and is most frequently found in the Amish population where the incident rate is 1 in 1300 newborns. In people of Finnish descent, it affects about 1 in 20 000 live births. The RMRP gene defects can cause another related autosomal recessive disorder anauxetic dysplasia (AD, MIM 607095). Patients with AD showed more severe skeletal phenotypes than CHH with extreme short stature. However, they are not predisposed to malignancy (Thiel *et al.*, 2005). The RMRP mutations in CHH patients affect both rRNA and mRNA cleavage whereas AD patients are only deficient in rRNA maturation. These effects highlight the multiple functions of RMRP. In addition to CHH and AD, other mutations in RMRP have been implicated in additional rare inherited conditions such as metaphyseal dysplasia without hypotrichosis, kyphomelic dysplasia, and Omenn syndrome with a spectrum of symptoms generally involving the musculoskeletal system or the immune system (Martin and Li, 2007).

A missense mutation in another component of rRNA processing, Cirhin, causes North American Indian childhood cirrhosis (NAIC, MIM 604901) (Chagnon *et al.*, 2002). Cirhin is a homolog of a yeast tUTP component, UTP4, shown to be essential for 18S and 25S rRNA maturation in *S. cerevisiae* (Freed and Baserga, 2010). NAIC is a rare genetic disease that affects children of Native American descent in Quebec. It causes early progressive and irreversible liver damage, leading to death unless it is treated with liver transplantation.

As we gain further understanding of the ribosomopathies, we obtain more insights into the complexity of rRNA processing in humans.

Perspectives

What is the evolutionary advantage of a system that requires processing of its components to such a large extent, especially when it comes at the high energetic cost of synthesizing several RNA sequences (ETS and ITS elements) only to degrade them through maturation? One possible explanation is that processing provides directionality to ribosomal biogenesis. Formation of the 90S pre-ribosome in the nucleolus involves multiple protein–protein and protein–RNA interactions. Throughout ribosomal biogenesis, new protein–protein associations are formed and many are lost, as factors are released from the pre-60S and pre-40S particles. These interactions are essentially reversible. However, cleavage of RNA phosphodiester bonds is not, therefore making the process unidirectional and irreversible. Processing also provides a clock mechanism ensuring that enough time is given for appropriate folding and assembly of intermediates at different steps. Timing of incorporation and release of protein factors can be determined by sequential irreversible processing of the intermediates (Strunk and Karbstein, 2009).

Although our understanding of rRNA processing has greatly improved, much work is left to be done. More research is necessary to determine the structure and folding of the intermediates during processing. We are also starting to learn that alternative pathways of processing are not only organism dependent, but also vary among different cell types and metabolic states.

Conclusions

- Ribosomal RNA is transcribed as precursors with external and internal spacers.
- Processing of ribosomal RNA starts in the nucleolus, continues in the nucleoplasm, and ends in the cytoplasm.
- Exo and endonucleases remove the spacers with the help of accessory proteins.
- Specific residues are methylated or pseudouridylated.
- Chemical modifications rely on small nucleolar RNAs and their associated proteins.
- Some ribosomal RNAs can catalyze their own processing.
- Alternative pathways for the production of rRNAs are frequently used.
- Defects in rRNA processing are associated with human diseases.

Acknowledgments

This work was supported by the National Institutes of Health grants GM 092719 and GM 110227 to N.W.

See also: Cell Division/Death: Cell Cycle: Interplay between Oncogenes and Tumor Suppressor Genes in Human Disease. Cell Division/Death: Regulation of Cell Growth: Eukaryotic Ribosome Assembly and Export

References

- Armistead, J., Triggs-Raine, B., 2014. Diverse diseases from a ubiquitous process: The ribosomopathy paradox. *FEBS Letter* 588, 1491–1500.
- Baudin-Baillieu, A., Fabret, C., Liang, X.H., *et al.*, 2009. Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy. *Nucleic Acids Research* 37, 7665–7677.
- Bernstein, K.A., Gallagher, J.E., Mitchell, B.M., *et al.*, 2004. The small-subunit processome is a ribosome assembly intermediate. *Eukaryotic Cell* 3, 1619–1626.
- Bohne, A.V., 2014. The nucleoid as a site of rRNA processing and ribosome assembly. *Frontiers in Plant Science* 5, 257.
- Borovjagin, A.V., Gerbi, S.A., 2001. *Xenopus* U3 snoRNA GAC-Box A' and Box A sequences play distinct functional roles in rRNA processing. *Molecular and Cellular Biology* 21, 6210–6221.
- Borovjagin, A.V., Gerbi, S.A., 2005. An evolutionary intra-molecular shift in the preferred U3 snoRNA binding site on pre-ribosomal RNA. *Nucleic Acids Research* 33, 4995–5005.
- Brzezniak, L.K., Bijata, M., Szczesny, R.J., Stepień, P.P., 2011. Involvement of human ELAC2 gene product in 3' end processing of mitochondrial tRNAs. *RNA Biology* 8, 616–626.
- Cech, T.R., Zaug, A.J., Grabowski, P.J., 1981. *In vitro* splicing of the ribosomal RNA precursor of *Tetrahymena*: Involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell* 27, 487–496.
- Chagnon, P., Michaud, J., Mitchell, G., *et al.*, 2002. A missense mutation (R565W) in cirhin (FLJ14728) in North American Indian childhood cirrhosis. *American Journal of Human Genetics* 71, 1443–1449.
- Ciganda, M., Williams, N., 2011. Eukaryotic 5S rRNA biogenesis. *Wiley Interdisciplinary Reviews: RNA* 2, 523–533.
- Cote, C.A., Greer, C.L., Peculis, B.A., 2002. Dynamic conformational model for the role of ITS2 in pre-rRNA processing in yeast. *RNA* 8, 786–797.
- Court, D.L., Gan, J., Liang, Y.H., *et al.*, 2013. RNase III: Genetics and function; structure and mechanism. *Annual Review of Genetics* 47, 405–431.
- Coute, Y., Kindbeiter, K., Belin, S., *et al.*, 2008. ISG20L2, a novel vertebrate nucleolar exoribonuclease involved in ribosome biogenesis. *Molecular & Cellular Proteomics* 7, 546–559.

- De La Cruz, J., Kressler, D., Tollervey, D., Linder, P., 1998. Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in *Saccharomyces cerevisiae*. *EMBO Journal* 17, 1128–1140.
- Delprato, A., Kadri, Y.A., Perebaskine, N., *et al.*, 2014. Crucial role of the Rcl1p-Bms1p interaction for yeast pre-ribosomal RNA processing. *Nucleic Acids Research* 42, 10161–10172.
- Dennerlein, S., Rozanska, A., Wydro, M., *et al.*, 2010. Human ERAL1 is a mitochondrial RNA chaperone involved in the assembly of the 28S small mitochondrial ribosomal subunit. *Biochemical Journal* 430, 551–558.
- Freed, E.F., Baserga, S.J., 2010. The C-terminus of Utp4, mutated in childhood cirrhosis, is essential for ribosome biogenesis. *Nucleic Acids Research* 38, 4798–4806.
- Gutgsell, N.S., Jain, C., 2012. Role of precursor sequences in the ordered maturation of *E. coli* 23S ribosomal RNA. *RNA* 18, 345–353.
- Hage, A.E., Tollervey, D., 2004. A surfeit of factors: Why is ribosome assembly so much more complicated in eukaryotes than bacteria? *RNA Biology* 1, 10–15.
- Hashem, Y., Des Georges, A., Fu, J., *et al.*, 2013. High-resolution cryo-electron microscopy structure of the *Trypanosoma brucei* ribosome. *Nature* 494, 385–389.
- Henry, Y., Wood, H., Morrissey, J.P., *et al.*, 1994. The 5' end of yeast 5.8S rRNA is generated by exonucleases from an upstream cleavage site. *EMBO Journal* 13, 2452–2463.
- Ito, S., Akamatsu, Y., Noma, A., *et al.*, 2014. A single acetylation of 18S rRNA is essential for biogenesis of the small ribosomal subunit in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 289, 26201–26212.
- Lafontaine, D.L., Preiss, T., Tollervey, D., 1998. Yeast 18S rRNA dimethylase Dim1p: A quality control mechanism in ribosome synthesis? *Molecular and Cellular Biology* 18, 2360–2370.
- Lamanna, A.C., Karbstein, K., 2011. An RNA conformational switch regulates pre-18S rRNA cleavage. *Journal of Molecular Biology* 405, 3–17.
- Lebaron, S., Segerstolpe, A., French, S.L., *et al.*, 2013. Rrp5 binding at multiple sites coordinates pre-rRNA processing and assembly. *Molecular Cell* 52, 707–719.
- Lee, Y., Nazar, R.N., 1997. Ribosomal 5S rRNA maturation in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 272, 15206–15212.
- Li, Z., Deutscher, M.P., 1995. The rRNA processing enzyme RNase T is essential for maturation of 5S RNA. *Proceedings of the National Academy of Sciences of the United States of America* 92, 6883–6886.
- Li, Z., Pandit, S., Deutscher, M.P., 1999a. Maturation of 23S ribosomal RNA requires the exoribonuclease RNase T. *RNA* 5, 139–146.
- Li, Z., Pandit, S., Deutscher, M.P., 1999b. RNase G (CafA protein) and RNase E are both required for the 5' maturation of 16S ribosomal RNA. *EMBO Journal* 18, 2878–2885.
- Liang, X.H., Liu, Q., Fournier, M.J., 2009. Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA* 15, 1716–1728.
- Mackie, G.A., 1998. Ribonuclease E is a 5'-end-dependent endonuclease. *Nature* 395, 720–723.
- Makitie, O., Kaitila, I., 1993. Cartilage-hair hypoplasia — Clinical manifestations in 108 Finnish patients. *European Journal of Pediatrics* 152, 211–217.
- Martin, A.N., Li, Y., 2007. RNase MRP RNA and human genetic diseases. *Cell Research* 17, 219–226.
- Martin, R., Hackert, P., Ruprecht, M., *et al.*, 2014. A pre-ribosomal RNA interaction network involving snoRNAs and the Rok1 helicase. *RNA* 20, 1173–1182.
- Mitchell, P., Petfalski, E., Tollervey, D., 1996. The 3' end of yeast 5.8S rRNA is generated by an exonuclease processing mechanism. *Genes & Development* 10, 502–513.
- Nazar, R.N., 2004. Ribosomal RNA processing and ribosome biogenesis in eukaryotes. *IUBMB Life* 56, 457–465.
- Perez-Fernandez, J., Roman, A., De Las Rivas, J., *et al.*, 2007. The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism. *Molecular and Cellular Biology* 27, 5414–5429.
- Preti, M., O'Donohue, M.F., Montel-Lehry, N., *et al.*, 2013. Gradual processing of the ITS1 from the nucleolus to the cytoplasm during synthesis of the human 18S rRNA. *Nucleic Acids Research* 41, 4709–4723.
- Price, J.V., Cech, T.R., 1988. Determinants of the 3' splice site for self-splicing of the *Tetrahymena* pre-rRNA. *Genes & Development* 2, 1439–1447.
- Ridanpaa, M., Sistonen, P., Rockas, S., *et al.*, 2002. Worldwide mutation spectrum in cartilage-hair hypoplasia: Ancient founder origin of the major 70A→G mutation of the untranslated RMRP. *European Journal of Human Genetics* 10, 439–447.
- Ridanpaa, M., Van Eenennaam, H., Pelin, K., *et al.*, 2001. Mutations in the RNA component of RNase MRP cause a pleiotropic human disease, cartilage-hair hypoplasia. *Cell* 104, 195–203.
- Sakiyama, J., Zimmer, S.L., Ciganda, M., *et al.*, 2013. Ribosome biogenesis requires a highly diverged XRN family 5'->3' exoribonuclease for rRNA processing in *Trypanosoma brucei*. *RNA* 19, 1419–1431.
- Sanchez, M.I., Mercer, T.R., Davies, S.M., *et al.*, 2011. RNA processing in human mitochondria. *Cell Cycle* 10, 2904–2916.
- Schafer, T., Maco, B., Petfalski, E., *et al.*, 2006. Hrr25-dependent phosphorylation state regulates organization of the pre-40S subunit. *Nature* 441, 651–655.
- Scherrer, K., 2003. Historical review: The discovery of 'giant' RNA and RNA processing: 40 years of enigma. *Trends in Biochemical Sciences* 28, 566–571.
- Scherrer, K., Latham, H., Darnell, J.E., 1963. Demonstration of an unstable RNA and of a precursor to ribosomal RNA in HeLa cells. *Proceedings of the National Academy of Sciences of the United States of America* 49, 240–248.
- Schmitt, M.E., Clayton, D.A., 1993. Nuclear RNase MRP is required for correct processing of pre-5.8S rRNA in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 13, 7935–7941.
- Schnare, M.N., Gray, M.W., 1990. Sixteen discrete RNA components in the cytoplasmic ribosome of *Euglena gracilis*. *Journal of Molecular Biology* 215, 73–83.
- Shajani, Z., Sykes, M.T., Williamson, J.R., 2011. Assembly of bacterial ribosomes. *Annual Review of Biochemistry* 80, 501–526.
- Sloan, K.E., Bohnsack, M.T., Schneider, C., Watkins, N.J., 2014. The roles of SSU processome components and surveillance factors in the initial processing of human ribosomal RNA. *RNA* 20, 540–550.
- Song, W.S., Lee, M., Lee, K., 2011. RNase G participates in processing of the 5'-end of 23S ribosomal RNA. *Journal of Microbiology* 49, 508–511.
- Strunk, B.S., Karbstein, K., 2009. Powering through ribosome assembly. *RNA* 15, 2083–2104.
- Sulthana, S., Deutscher, M.P., 2013. Multiple exoribonucleases catalyze maturation of the 3' terminus of 16S ribosomal RNA (rRNA). *Journal of Biological Chemistry* 288, 12574–12579.
- Thiel, C.T., Horn, D., Zabel, B., *et al.*, 2005. Severely incapacitating mutations in patients with extreme short stature identify RNA-processing endoribonuclease RMRP as an essential cell growth regulator. *American Journal of Human Genetics* 77, 795–806.
- Tollervey, D., Kiss, T., 1997. Function and synthesis of small nucleolar RNAs. *Current Opinion in Cell Biology* 9, 337–342.
- Umaer, K., Ciganda, M., Williams, N., 2014. Ribosome biogenesis in African trypanosomes requires conserved and trypanosome-specific factors. *Eukaryotic Cell* 13, 727–737.
- Van Hoof, A., Lennertz, P., Parker, R., 2000. Three conserved members of the RNase D family have unique and overlapping functions in the processing of 5S, 5.8S, U4, U5, RNase MRP and RNase P RNAs in yeast. *EMBO Journal* 19, 1357–1365.
- Wang, M., Pestov, D.G., 2011. 5'-end surveillance by Xrn2 acts as a shared mechanism for mammalian pre-rRNA maturation and decay. *Nucleic Acids Research* 39, 1811–1822.
- Watkins, N.J., Bohnsack, M.T., 2012. The box C/D and H/ACA snoRNPs: Key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdisciplinary Reviews: RNA* 3, 397–414.
- Wolin, S.L., Cedervall, T., 2002. The La protein. *Annual Review of Biochemistry* 71, 375–403.
- Woolford Jr., J.L., Baserga, S.J., 2013. Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*. *Genetics* 195, 643–681.
- Zheng, S., Lan, P., Liu, X., Ye, K., 2014. Interaction between Ribosome Assembly Factors Krr1 and Faf1 is Essential for Formation of Small Ribosomal Subunit in Yeast. *Journal of Biological Chemistry* 289, 22692–22703.

Eukaryotic Ribosome Assembly and Export

VG Panse and CS Weirich, ETH Zürich, Switzerland

© 2016 Elsevier Inc. All rights reserved.

Introduction

The ribosome is responsible for carrying out the final decoding step in gene expression, in which mRNA is translated into proteins. The overall structure and function of this molecular machine is universally conserved and consists of two subunits composed of ribosomal RNA (rRNA) and ribosomal proteins (r-proteins). The small subunit, which is 40S in eukaryotes and 30S in bacteria and archae, is responsible for codon recognition and pairing mRNAs with their cognate tRNAs (Wimberly *et al.*, 2000). The large subunit, which is 60S in eukaryotes and 50S in bacteria and archae, catalyzes the peptidyl transferase reaction to synthesize polypeptide chains (Ban *et al.*, 2000). Biogenesis of these enormous ribonucleoprotein complexes is a critical and energy-consuming task for all cells.

Although the overall architecture of the ribosome is conserved, eukaryotic ribosomes are significantly larger than their prokaryotic counterparts and contain proportionally more rRNA (65%) and less r-protein (35%) mass. The large, 60S subunit of *Saccharomyces cerevisiae* contains three rRNAs (25S, 5.8S, and 5S) and 46 r-proteins. The small, 40S subunit contains a single rRNA (18S) and 33 r-proteins (Ben-Shem *et al.*, 2011; Klinge *et al.*, 2011; Rabl *et al.*, 2011). The structures of both ribosomal subunits have been solved, providing an exquisite level of detail regarding their architectures and unprecedented insights into the mechanism of translation (Ramakrishnan, 2014). Despite these advances, we still understand relatively little about the assembly and transport pathways required to generate these molecular machines. This article will focus on what is currently known and will focus on budding yeast, which has been used to generate a general framework to further dissect this conserved biosynthetic pathway.

Eukaryotic ribosome assembly takes place in multiple cellular compartments: the nucleolus, the nucleoplasm, and the cytoplasm (Figure 1). This assembly process is a complex task that requires the coordination of RNA polymerases I, II, and III, as well as an efficient splicing machinery and efficient intracellular transport machinery. Surprisingly, despite this complexity, ribosome biogenesis is an incredibly efficient process, with yeast cells producing up to 60 ribosomes per second (Warner, 1999; Tschochner and Hurt, 2003).

Ribosomal Pre-rRNA Processing

Ribosome biogenesis begins co-transcriptionally on the 35S pre-rRNA (45S in mammals), which is synthesized by RNA Polymerase I (Figure 2; Dragon *et al.*, 2002; Woolford and Baserga, 2013). This transcript contains both 5' and 3' externally transcribed spacer (ETS) sequences, as well as the three rRNAs (18S, 5.8S, and 25S), separated by the internally transcribed spacer sequences ITS1 and ITS2. The yeast 90S particle (which is homologous to the mammalian SSU (Small SubUnit processome)) assembles on this precursor rRNA, forming structures first described by electron microscopy as 'Christmas trees' forming at the rDNA loci of *Xenopus* oocytes (Miller and Beatty, 1969; Dragon *et al.*, 2002; Grandi *et al.*, 2002). This precursor is cleaved first at site A0, generating the shorter 33S pre-rRNA. The second cleavage event occurs at site A1, generating the 32S pre-rRNA. The third cleavage site occurs at site A2 and generates the 20S and 27SA2. At this point, the 20S pre-rRNA is released to be further processed within the context of the 43S, which will become the mature 40S subunit, and is

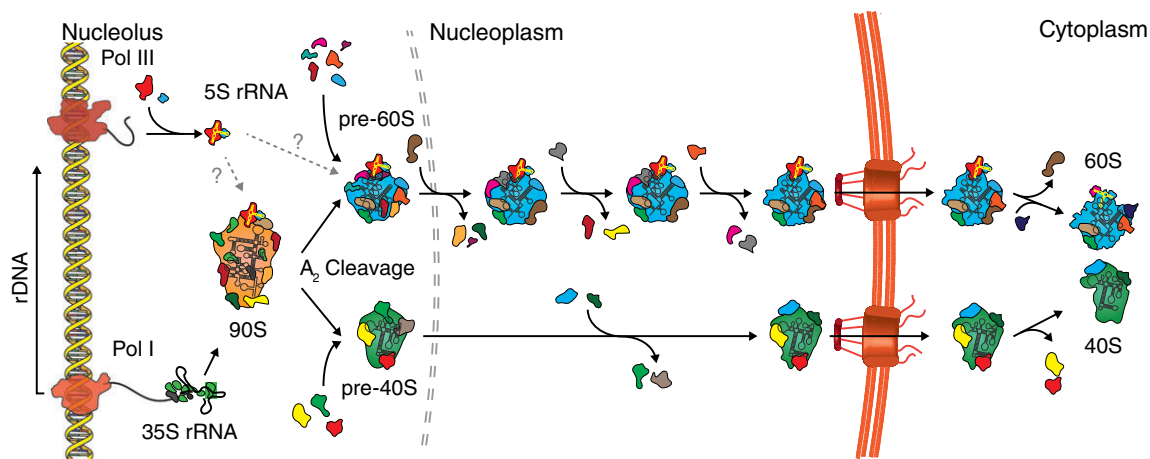
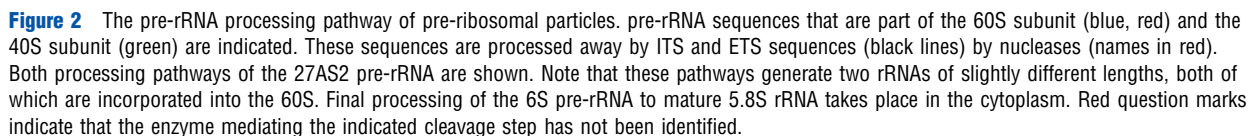


Figure 1 Eukaryotic ribosomes assemble along pathways that span the nucleolus, nucleoplasm, and cytoplasm. The 90S particle (orange) is built from the 35S pre-rRNA, which recruits assembly factors and small subunit r-proteins co-transcriptionally. Cleavage at site A2 separates the 90S into a pre-40S subunit (green) and a pre-60S subunit (blue), which follow independent maturation pathways. Assembly factors bind and drive maturation steps on pre-ribosomal particles, then are sequentially released. Final maturation of both the 40S and 60S pre-ribosomal subunits occurs in the cytoplasm. Question marks indicate uncertainty in the assembly pathway.



The remaining particles, containing 27AS2 pre-rRNA, recruit the 5S rRNA, which is transcribed by RNA polymerase III and bound by the biogenesis factors Rpf2 and Rrs1 and the ribosomal proteins uL18 (human L5) and uL5 (human L11) (Zhang *et al.*, 2007; nomenclature as in Ban *et al.*, 2014). Then, 27AS2 continues processing through one of the two pathways. The first pathway, comprising 85% of molecules, begins with

In addition to RNA cleavage events, eukaryotic ribosomes undergo multiple posttranslational RNA modifications, including isomerization of uridines to pseudouridines and

methylation (Woolford and Baserga, 2013; Lafontaine, 2015). Two classes of small nucleolar RNAs (snoRNAs) direct these modifications: the box C/D snoRNAs direct methylation, whereas the box H/ACA snoRNAs direct pseudouridylation. Although both modifications are more widespread in eukaryotes, occurring between 50 and 100 times each, they are also found in prokaryotes and archae. The functions of individual modifications remain unclear, and have been proposed to affect translational efficiency and rRNA structure (Baxter-Roshek *et al.*, 2007).

Trans-Acting Factors Involved in Ribosome Assembly

Unsurprisingly, ribosomal proteins themselves play critical roles in ribosome assembly. One study demonstrated that most 40S subunit proteins interact directly with the 35S pre-rRNA, though the interactions are weak. In addition, the small subunit proteins uS17 (yeast Rps11), uS4 (yeast Rps9), and uS15 (yeast Rps13) are required for early rRNA processing events and are homologous to the *Escherichia coli* 'primary binders' that mark the beginning of prokaryotic ribosome assembly *in vitro* (Held *et al.*, 1974). However, unlike their prokaryotic counterparts, eukaryotic ribosomes cannot self-assemble *in vitro* and require, in addition to ribosomal proteins, >200 additional *trans*-acting biogenesis factors.

Genetic screens in budding yeast and, in particular, the advent of tandem affinity purification (TAP) have yielded a long list of factors involved in ribosome assembly and export (Tschochner and Hurt, 2003). The majority of uncovered factors are essential, contrasting sharply with ribosome biogenesis in prokaryotes, which does not rely on essential biogenesis factors (Hage and Tollervey, 2004; Connolly and Culver, 2009). These evolutionarily conserved factors are generally thought to play roles in distinct maturation stages, after which they are released and recycled. In addition to scaffolds and RNA-interacting proteins, energy-consuming enzymes such as AAA-adenosine triphosphatase (ATPases), ABC-ATPases, guanosine triphosphatase (GTPases), and ATP-dependent RNA helicases are thought to be responsible for factor release, conferring directionality and irreversibility to the assembly process. Recent studies have begun to unravel the substrates, pre-ribosomal binding sites, and specific functions of these energy-consuming enzymes (Kressler *et al.*, 2012; Strunk *et al.*, 2012).

Two well-characterized examples of energy-dependent factor release are the essential AAA-ATPases Rix7 and Rea1, both required during 60S biogenesis. A related AAA-ATPase Drg1 is required for cytoplasmic maturation of pre-60S subunits (see Section Cytoplasmic Maturation of the 60S Subunit). Rix7, an early acting energy-consuming release factor, is responsible for early nucleolar 60S maturation. Although clear evidence remains elusive, this ATPase is proposed to be involved in the release of the assembly factor Nsa1. In turn, the release of Nsa1 is thought to require posttranslational modifications, which allow its recognition by Rix7. This energy-consuming release step is critical to facilitate the nucleolar to nucleoplasmic movement of 60S pre-ribosomal particles (Kressler *et al.*, 2008; Panse *et al.*, 2006).

Rea1 is involved at multiple maturation steps of the large ribosomal precursor (Ulbrich *et al.*, 2009; Bassler *et al.*, 2010;

Kressler *et al.*, 2012). Rea1 triggers the release of the biogenesis factors Ytm1 and Rsa4 (Nissan *et al.*, 2004; Ulbrich *et al.*, 2009; Bassler *et al.*, 2010). Release of Ytm1, along with its cofactors Erb1 and Nop7, has been proposed to trigger the displacement of neighboring assembly factors, conferring directionality to the ribosome maturation process (Nissan *et al.*, 2002, 2004; Ulbrich *et al.*, 2009). In addition to its nucleolar role, recent studies suggest that Rea1 is involved in a pre-60S nuclear checkpoint, releasing the GTPase Nug2 (Nog2) and allowing recruitment of the 60S export factor Nmd3. These observations suggest that Rea1 coordinates nuclear maturation and export of pre-60S particles (Matsuo *et al.*, 2014).

Many 40S biogenesis factors, such as the RNA dimethylase Dim1p, Dim2p, Enp1p, Hrr25p, Nob1p, the RNA helicase Prp43p, Rrp12p, Tsr1p, and Tsr2, required for 40S biogenesis are present within the 90S prior to A2 cleavage and release of the 20S pre-rRNA (Panse and Johnson, 2010). After A2 cleavage, additional proteins, including Ltv1p, Pfa1p/Sqs1p, Rio1p, and Rio2p, join the 20S pre-rRNA assembled within the pre-40S subunit (Panse and Johnson, 2010). Although the mechanisms by which these factors are recruited and the precise roles they play during biogenesis remain unclear, some details have been identified. For example, the kinase Hrr25 is thought to phosphorylate Enp1p, Ltv1p, and the small subunit protein uS3 (yeast Rps3p), perhaps to cause a structural change in these pre-40S particles and promote efficient transport through the nuclear pore (Schafer *et al.*, 2006). In addition, the essential RNA dimethylase Dim1p is required for early nucleolar processing events, though it also plays a later role in the cytoplasm (Lafontaine *et al.*, 1994, 1998) (see Section Cytoplasmic Maturation Events of the 40S Subunit).

In addition to assembly factors and energy-consuming release and remodeling factors, pre-ribosomal particles have been proposed to require the action of r-protein delivery factors, termed 'escortins.' One escortin, Tsr2, releases eS26 from its import receptor in a RanGTP-independent manner and then ensures its secure transfer to the earliest identified 90S pre-ribosome particle (Schutz *et al.*, 2014). Given the large influx of r-proteins into the nucleus and the inherent instability of these RNA-binding polypeptides, a network of escortins could keep r-proteins soluble until they are incorporated into pre-ribosomes.

Nuclear Export of Pre-Ribosomal Particles

Although translation is a cytoplasmic event, ribosome production begins in the nucleolus. Therefore, in addition to the import and nucleolar targeting of ribosomal proteins, ribosomal subunits must also be exported into the cytoplasm. All nucleocytoplasmic transport takes place through nuclear pore complexes (NPCs), which are macromolecular assemblies that span the double lipid bilayer of the nuclear envelope. The extraordinary size of pre-ribosomal particles (approximately 2 MDa) presents a particular challenge to the transport machinery (Ribbeck and Gorlich, 2002). In addition, the highly charged surfaces of these complexes must be shielded from the hydrophobic FG-repeats present within the NPC. Despite these obstacles, ribosome export is remarkably efficient, with growing yeast cells exporting about 25 pre-ribosomal particles

per minute per NPC (Warner, 1999). How the transport machinery manages to move these large, charged pre-ribosomal particles quickly and efficiently has been a topic of intense study, yielding the identification of multiple export factors, as described below.

In the 1990s, visualization assays based on green fluorescent protein (GFP)-tagging of ribosomal proteins and *in situ* hybridization against rRNAs were used to identify factors involved in pre-ribosomal subunit export (Hurt *et al.*, 1999; Moy and Silver, 1999; Stage-Zimmermann *et al.*, 2000; Milkereit *et al.*, 2003). These approaches identified several structural components of the NPC, the small GTPase Ran, Ran regulators, and the export receptor Crm1/Xpo1 as export factors for pre-60S and pre-40S subunits (Hurt *et al.*, 1999; Moy and Silver, 1999; Stage-Zimmermann *et al.*, 2000). However, these factors are not predicted to be sufficient for pre-ribosomal subunit export, since the rate of transport receptor shuttling is known to decrease as the size of a cargo increases (Ribbeck *et al.*, 1998). Transport of very large pre-ribosomal particles is predicted to be extremely slow, suggesting the requirement of multiple export factors to increase translocation rates.

Visual and genetic screens identified the nuclear export signal (NES)-containing adaptor Nmd3, which forms a trimeric complex with RanGTP and Xpo1 on the surface of the pre-60S subunit (Ho *et al.*, 2000; Gadal *et al.*, 2001). Using parallel approaches based on systematic RNA interference (RNAi) and ribosomal protein reporters, such visual screens have been extended to human cells. One such screen uncovered an essential role of the RanGTP-binding protein Exp5 as an export receptor for pre-60S subunits (Wild *et al.*, 2010). A structural modeling approach identified an additional export factor Rrp12, which binds both pre-ribosomal subunits and FXFG-repeat nucleoporins (Oeffinger *et al.*, 2004). Like RanGTP-dependent export factors, Rrp12 contains HEAT repeats, and therefore may interact with FXFG repeats using a similar mechanism. Intriguingly, Rrp12 was shown to interact with both GTP and guanosine diphosphate (GDP) forms of Ran. However, the mechanism by which Rrp12 promotes ribosomal subunit export and its regulation by Ran remain unknown.

Intriguingly, the essential mRNA export factor Mex67-Mtr2 (TAP-p15 in humans) also exports both pre-60S and pre-40S subunits (Segref *et al.*, 1997; Yao *et al.*, 2007; Faza *et al.*, 2012). The large subunit of the heterodimer, Mex67, consists of an N-terminal domain, a leucine-rich repeat (LRR), a nuclear transport factor 2 (NTF2)-like middle domain, and a C-terminal ubiquitin-associated (UBA-like) domain (Strasser *et al.*, 2000). Mtr2 shares structural features with NTF2 (Bayliss *et al.*, 2002) and heterodimerizes with the NTF2-like middle domain of Mex67 (Santos-Rosa *et al.*, 1998; Strasser *et al.*, 2000). Although Mex67 and Mtr2 do not rely on the RanGTP system, and have no structural similarities to karyopherins, the heterodimer also exports cargoes by interacting with FG-rich nucleoporins. To export mRNAs and pre-ribosomal subunits, Mex67-Mtr2 uses different binding surfaces. Interaction with pre-ribosomal subunits relies on loops within the NTF2-like domains of Mex67 and Mtr2 (Fribourg and Conti, 2003; Senay *et al.*, 2003; Yao *et al.*, 2007, 2008). Whether these loops directly interact with an exposed structured rRNA or a protein factor(s) on pre-ribosomal subunits still remains to be

determined. The NTF2-like domains of Mex67-Mtr2 and the UBA-like domain of Mex67 interact with FG-repeat nucleoporins to mediate mRNA and pre-ribosomal subunit export (Strasser *et al.*, 2000; Strawn *et al.*, 2001). The mRNA export factor Npl3 also plays a role in pre-60S subunit nuclear export (Hackmann *et al.*, 2011). Thus, the mRNA export machinery is partially adapted to carry both pre-ribosomal and mRNA cargoes. It is not yet known how the cellular pools of these factors are fractionated to participate in the nuclear export of mRNAs and pre-ribosomal subunits. Understanding the molecular basis of this allocation will reveal how the three export pathways cross-talk to deliver appropriate levels of mRNA and ribosomal subunits in the cytoplasm.

In addition to essential transport receptors, pre-ribosomal particles are also associated with nonessential transport factors that directly facilitate their translocation through the NPC channel, through interactions with nucleoporins. Although these factors are known to promote pre-ribosomal subunit export, it is not known whether they decorate every subunit during its export process, adding incrementally to export efficiency, or whether each subunit requires a quorum of factors, but not all, for efficient export. Two examples of these nonessential factors are described below.

The pre-60S export factor Arx1 is structurally homologous to methionine amino peptidases, which remove N-terminal methionines from nascent polypeptides as they emerge from the ribosome (Bradatsch *et al.*, 2007; Hung *et al.*, 2008). However, Arx1 lacks methionine amino peptidase activity, and mutations in the methionine-binding pocket impair pre-60S subunit export (Bradatsch *et al.*, 2007). When bound to the pre-60S, Arx1 occludes the exit tunnel of the 60S subunit and stabilizes the eukaryotic-specific rRNA expansion segment 27 (ES27) in an alternate conformation (Bradatsch *et al.*, 2012; Greber *et al.*, 2012). However, the details of how Arx1 mediates pre-60S export, as well as the function of its interaction with ES27 currently remain unknown.

The mRNA export factor Gle2 also mediates the export of pre-60S subunits, using a distinct, non-FG repeat binding mechanism (Occhipinti *et al.*, 2013). This factor interacts with late pre-60S subunits as well as the Gle2-binding-sequence (GLEBS)-motif of the nucleoporin Nup116. The Gle2-Nup116 interaction is required both to bind and export pre-60S subunits, but disruption of this interaction has no effect on mRNA export. Therefore, Gle2 appears to have dual functions in mRNA export and pre-60S export. This unique mechanism has been proposed to prevent kinetic delays during pre-60S translocation through the NPC, especially if cargoes have failed to recruit their optimal complement of export factors.

In addition to Gle2 and Arx1, Ecm1 and Bud20 have also been identified as factors involved in pre-60S biogenesis (Bradatsch *et al.*, 2007; Yao *et al.*, 2010; Altwater *et al.*, 2012; Bassler *et al.*, 2012). One major challenge in the field remains to identify the binding sites of individual export factors on pre-ribosomal particles. Together, these analyses will provide insight both into the structure and of pre-ribosomal export particles and the diversity underlying this export process.

Export of pre-40S subunits also relies on auxiliary factors such as the Xpo1 interacting proteins Ltv1, Dim2, and Rio2 (Zemp *et al.*, 2009). In addition, the conserved RanGTP-binding protein Yrb2 (RanBP3 in humans) has been identified

as a specific pre-40S export factor (Taura *et al.*, 1998; Moy and Silver, 2002). The identification of multiple, nonessential NES containing adaptors suggests that these factors play redundant roles to recruit the essential export receptor Xpo1. Such redundancy would act to guarantee Xpo1 recruitment, but remains to be demonstrated.

Cytoplasmic Events in Ribosome Biogenesis

Although the majority of ribosome biogenesis takes place in the nucleus, both subunits complete their maturation only in the cytoplasm. Completing ribosome biogenesis in this subcellular compartment potentially allows a system of checkpoints to ensure that only functional ribosomal subunits are allowed to participate in translation.

Cytoplasmic Maturation of the 60S Subunit

The cytoplasmic maturation pathway of the 60S subunit requires the release of multiple factors in a specific order (Figure 3; Lo *et al.*, 2010). These release events rely on factors present exclusively in the cytoplasm that associate only transiently with pre-60S particles (Panse and Johnson, 2010). The cytoplasmic maturation pathway begins with the release of Rlp24 by Drg1. Next, the ribosomal protein eL24 (yeast Rpl24) is incorporated onto the subunit and recruits Rei1. Together with Jjj1 and the ATPase Ssa1/Ssa2, these proteins cause subsequent Arx1 and Alb1 release. Only after the release of these two proteins, as well as formation of the acidic ribosomal stalk and correct assembly of the catalytic P-site, is the maturation factor Tif6 also released (Bussiere *et al.*, 2012).

The AAA-ATPase Drg1 initiates the earliest cytoplasmic maturation step on pre-60S particles, which is required for subsequent steps that release and recycle additional factors. Consequently, Drg1 catalytic mutants accumulate shuttling assembly factors Rlp24, Nog1, Arx1, and Tif6 on cytoplasmic pre-60S subunits (Pertschy *et al.*, 2007). A recent study demonstrated that Drg1 catalyzes Rlp24 dissociation, which is

essential for subsequent release events (Kappel *et al.*, 2012). Drg1 is recruited to a C-terminal region within Rlp24 that also stimulates its ATPase activity. Notably, Rlp24 release also appears to require the FG-nucleoporin Nup116. Thus, Drg1 has been hypothesized to provide irreversibility to the nuclear export step and initiate the cytoplasmic maturation pathway.

Release of the ribosomal-like protein Rlp24 is necessary to allow the ribosomal protein eL24 to assemble into the pre-60S subunit. This exchange allows recruitment of the zinc-finger protein Rei1, which triggers the next maturation event. Rei1 works in conjunction with the J-domain protein Jjj1 and the ATPase Ssa1/Ssa2 (Hsp70) to release Arx1 and its interacting partner Alb1 (Demoinet *et al.*, 2007; Meyer *et al.*, 2007). Interestingly, these data indicate that Arx1 has an inhibitory role in driving cytoplasmic maturation pathway of pre-60S subunits. Arx1 has evolved from methionyl amino peptidases (MetAPs) (Bradshaw *et al.*, 1998). Based on similarity of Arx1 to MetAPs, Arx1 was predicted to bind to the MetAP site on the ribosome, preventing MetAP binding. Indeed, Cryo-EM studies showed that Arx1 binds near the ribosomal protein uL25, which interacts with the signal recognition particle and the translocon, further suggesting that Arx1 plays an inhibitory role in the cytoplasmic maturation pathway (Dalley *et al.*, 2008).

Tif6 is another biogenesis factor that must be removed from the large subunit in the cytoplasm. This shuttling factor remains bound to cytoplasmic pre-60S particles, preventing their interaction with mature 40S subunits (Russell and Spremulli, 1979; Valenzuela *et al.*, 1982). The GTPase Efl1 and its cofactor, the Shwachman–Bodian Syndrome protein Sdo1, trigger release of Tif6 (Becam *et al.*, 2001; Senger *et al.*, 2001; Menne *et al.*, 2007). In *efl1* and *sdo1* mutants, Tif6 accumulates in the cytoplasm on late pre-60S subunits, and *tif6* mutations that weaken its affinity for the 60S subunit suppress the growth defects of *efl1* and *sdo1* mutants, providing strong genetic evidence that Tif6 is the primary substrate of Efl1 and Sdo1.

Interestingly, Tif6 also mislocalizes in *γvh1* mutants, which also display defects in stalk assembly (Kemmler *et al.*, 2009;

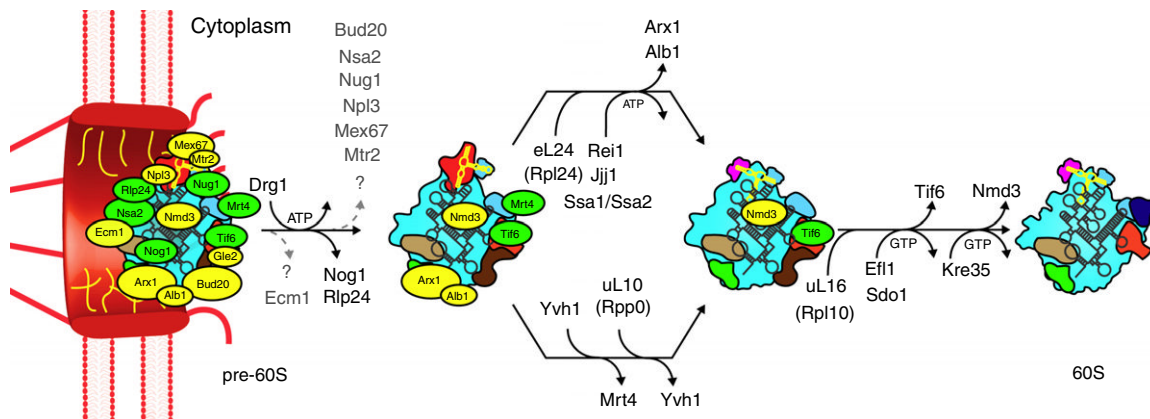


Figure 3 Cytoplasmic events required for 60S pre-ribosome maturation. Export factors are shown in yellow, and shuttling assembly factors are shown in green. Release of assembly factors Nog1 and Rlp24 by the ATPase Drg1 is thought to be the first step in the 60S pre-ribosome cytoplasmic maturation pathway. Rei1, J-domain protein Jjj1, and the ATPase Ssa1/Ssa2 are required for the release of Arx1 and Alb1. Stalk assembly relies on release of Mrt4 by Yvh1 and recruitment of uL10, which releases Yvh1. Tif6 release by Efl1 and Sdo1, as well as Nmd3 release by Kre35 follow stalk assembly. Factors that are released by unknown mechanisms (?) are listed in gray.

Lo *et al.*, 2009). During translation, the stalk functions in recruitment and activation of the GTPase eEF2 (Bargis-Surgey *et al.*, 1999). Given that Efl1 is closely related to eEF2, stalk assembly might play a similar role in biogenesis, recruiting Efl1 for the release of Tif6. These analyses suggest that the cytoplasmic maturation events in the 60S biogenesis are coupled and ordered sequentially from Drg1-dependent release of Rlp24 to Efl1-dependent release of Tif6 (Figure 3). Therefore, this pathway appears to rely on the coupled stepwise recruitment of cytoplasmic release factors to pre-60S subunits.

The essential NES containing adapter Nmd3 must also be released from pre-60S subunits in the cytoplasm. Both the ribosomal protein uL16 (yeast Rpl10) and the GTPase Kre35 (Lsg1) have been implicated in the release of Nmd3 (Karl *et al.*, 1999; Hedges *et al.*, 2005). Mutations in uL16 prevent release of Nmd3 from pre-60S subunits. Moreover, mutations in Kre35 that are predicted to disrupt its GTPase activity also block Nmd3 release in the cytoplasm. These results suggest that Kre35 triggers the binding of uL16 to the 60S, an event that is coupled to the release of Nmd3. However, the exact timing of Nmd3 release has not been elucidated.

Assembly of the stalk structure on the 60S subunit represents a critical step toward ribosome functionality. This structure is essential for the recruitment and activation of translation factors, in particular elongation factors. In yeast, the stalk region is composed of the ribosomal protein uL10 (yeast Rpp0) and two heterodimers of P1 (yeast Rpp1) and P2 (yeast Rpp2) (Berk and Cate, 2007). However, these proteins are not present in pre-60S subunits (Lo *et al.*, 2009; Kemmler *et al.*, 2009). Instead, the ribosomal-like protein Mrt4 is assembled into the pre-60S in the nucleus. Once the pre-60S is exported, the phosphatase Yvh1 catalyzes Mrt4 release, allowing the recruitment of uL10. However, the precise mechanism of Mrt4 release and the molecular events that lead to stalk assembly remains elusive (Lo *et al.*, 2009; Kemmler *et al.*, 2009).

In addition to these factors, a study combining genetic trapping, affinity purification and a targeted proteomic approach was used to characterize the proteome of 60S pre-ribosomes after nuclear export (Altwater *et al.*, 2012). This hybrid approach identified the factors Bud20, Nug1, Nsa2, and Rli1 as 60S-associated factors that are released in the cytoplasm after Drg1-mediated release of Rlp24. The functional significance of shuttling behavior of the identified assembly factors remains unknown. It could be that they participate directly in the transport and/or final functional proofreading of pre-60S subunits. The factors and mechanisms that release these shuttling assembly factors from 60S particles in the cytoplasm remain to be discovered.

Cytoplasmic Maturation Events of the 40S Subunit

The small pre-ribosomal subunit is accompanied to the cytoplasm by a handful of proteins, including Enp1, Tsr1, Ltv1, Dim1, Dim2, Nob1, Rio2, Hrr25, and Prp43 (Figure 4). These factors mediate both export and endo-nucleolytic cleavage of 20S pre-rRNA to 18S rRNA. This final cleavage event is an essential cytoplasmic maturation step that renders pre-40S particles translation-competent (Strunk *et al.*, 2011). How these shuttling factors are released from pre-40S subunits in the cytoplasm remains unclear.

Prior to 20S pre-rRNA cleavage, two adenine bases near the 3'-end of the future 18S rRNA are dimethylated by the essential dimethylase Dim1 (Lafontaine *et al.*, 1994). However, this modification is not essential, since processing of 20S pre-rRNA to 18S rRNA can still occur in the absence of dimethylation, forming functional 40S subunits. Dimethylation was suggested to play a role in fine-tuning of translation, as the catalytic inactive *dim1* mutant lacking 40S dimethylation displays increased antibiotic sensitivity (Lafontaine *et al.*, 1998). Multiple

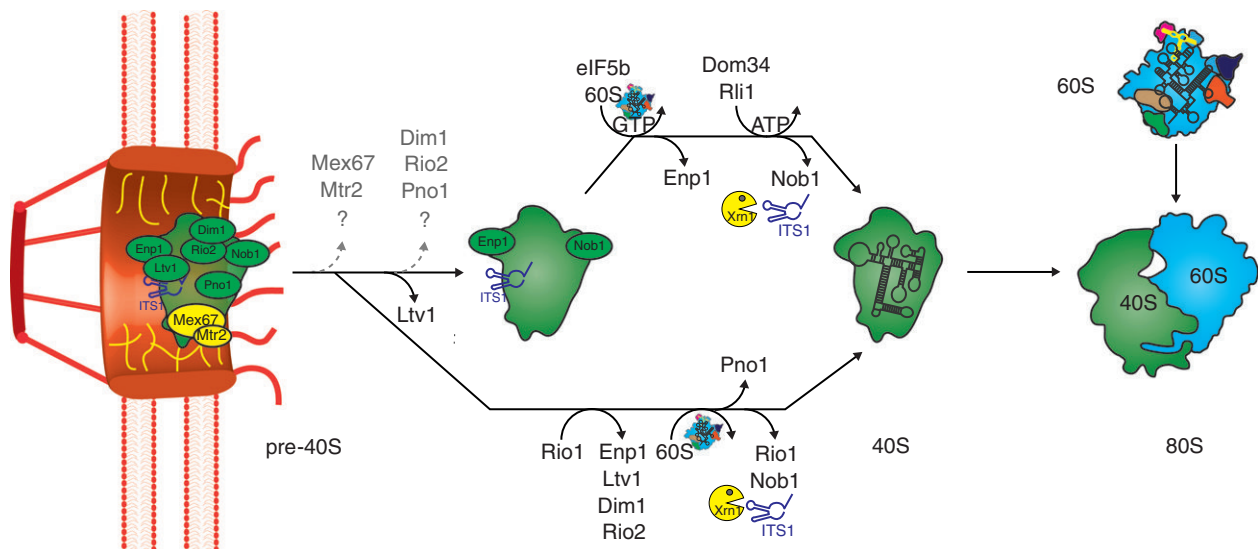


Figure 4 Cytoplasmic events required for 40S pre-ribosome maturation. Export factors are shown in yellow, and shuttling assembly factors are in green. Release of shuttling factors and final pre-rRNA processing requires the ATPase Rio1 and the GTPase eIF5b as well as interactions with a mature 60S subunit. 20S pre-rRNA is processed in the cytoplasm to the final 18S rRNA by the endonuclease Nob1 within a 80S-like particle. Factors that are released by unknown mechanisms (?) are listed in gray.

energy-consuming ATPases (Prp43, Rio2, and Fap7) and the PIN-domain endonuclease Nob1 have been implicated in this late maturation (Clissold and Ponting, 2000; Geerlings *et al.*, 2003; Vanrobays *et al.*, 2003; Jakovljevic *et al.*, 2004; Lamanna and Karbstein, 2009; Pertschy *et al.*, 2009). Intriguingly, Nob1 becomes associated with 40S pre-ribosomes in the nucleus, suggesting that this protein is specifically activated in the cytoplasm. This activation of Nob1 in the cytoplasm occurs within an 80S-like particle formed between by the joining of a pre-40S subunit with a mature 60S subunit. Nob1 activation seems to also require the GTPase activity of the translation initiation factor Fun12/eIF5b and the ATPase Rio1 (Lebaron *et al.*, 2012; Strunk *et al.*, 2012; Turowski *et al.*, 2014).

Quality Control Mechanisms for Ribosome Assembly

Because accurate protein translation is critical to the cell, pre-ribosomes must undergo quality control prior to their release into the translating pool. Quality control begins in the nucleus, where aberrant pre-ribosomes are recognized and rapidly degraded by the exosome (Mitchell *et al.*, 1997; Allmang *et al.*, 2000). Although recognition depends in part on the poly-adenylating activity of the TRAMP complex, the precise recognition mechanism remains unclear (Dez *et al.*, 2006). A second quality control step is the cytoplasmic nonfunctional RNA decay (NRD) pathway, which targets improperly assembled pre-ribosomal subunits for degradation (LaRiviere *et al.*, 2006; Cole *et al.*, 2009). In addition to these quality control mechanisms, a cytoplasmic pathway exists to proof-read pre-ribosomes in the cytoplasm, preventing the premature binding of translation factors and, most likely, actively checking pre-ribosomal subunits for functionality.

Cytoplasmic 60S proofreading depends a coterie of factors that block the binding of ribosomal proteins, interactions with translation factors, or pairing with the 40S (Figure 3). Proteins that block r-protein binding include the ribosomal-like proteins Rlp24 and Mrt4, that act as 'placeholders' for the r-proteins eL24 and P0, respectively. Tif6 prevents binding of pre-60S particles to mature 40S subunits, ensuring that only properly assembled mature 60S subunits are able to engage their partners. In addition, the GTPase Efl1, which is required for Tif6 release, shares sequence similarity with the GTPase elongation factor eEF2, and has been proposed to check the integrity of the P-site for functionality (Bussiere *et al.*, 2012). Therefore, cytoplasmic release factors appear to be responsible both for releasing shuttling assembly factors and, simultaneously, testing ribosome function.

Similar to the 60S, quality control of the 40S relies on assembly factors that are thought to prevent the premature binding of initiation factors, mRNA, tRNA, and the 60S subunit (Figure 4). Based on cryo-EM studies of a late cytoplasmic pre-40S particle, Ltv1, Enp1, Rio2, Tsr1, Dim1, Pno1, and Nob1 binding sites have been identified (Strunk *et al.*, 2011). Ltv1 and Enp1 directly bind uS3 (yeast Rps3) on its solvent side, thereby blocking the mRNA channel opening. Rio2, Tsr1, and Dim1 bind the subunit interface, thus preventing joining of the mature 60S subunit and translation initiation factor eIF1A. Nob1 and Pno1 block the binding of eIF3 and thereby interfere with translation initiation. Therefore, these factors are likely to function, in part, by preventing premature translation initiation.

After the release of Rio2, Tsr1, and Dim1, the resulting pre-40S particle has been shown to interact with mature 60S subunits, forming an 80S-like particle *in vitro* and *in vivo* (Strunk *et al.*, 2012; Lebaron *et al.*, 2012). This translation-like interaction, which could test the ability of a pre-40S subunit to engage with 60S subunits, triggers the activity of Nob1 to cleave 20S pre-rRNA to mature 18S rRNA *in vitro*. In addition, the conserved translation initiation factor eIF5b/Fun12 is important for formation of the 80S-like particle (Lebaron *et al.*, 2012). After cleavage, the ABC-ATPase Rli1 is thought to cause dissociation of the 80S-like particle and Nob1 release (Strunk *et al.*, 2012). Thus, processing of the 20S pre-rRNA has been proposed to act both as a trigger for subunit maturation and as a quality control mechanism sensing translational competence. By constantly engaging with each other in the cytoplasm, ribosomal subunits may sense their 'decoding' ability to segregate and target faulty subunits for disassembly and degradation (Schutz and Panse, 2012).

Concluding Remarks

The biogenesis and nuclear export of ribosomal subunits is one of the most important energy-consuming tasks in a eukaryotic cell. The pathways responsible for 60S and 40S biogenesis are remarkably complex, requiring hundreds of evolutionarily conserved assembly factors in addition to the RNA and protein components of the ribosomal subunits themselves. Surprisingly for such a complex pathway, ribosome biogenesis and transport is also incredibly fast and at the same time efficient. Although the factors involved in this process have been largely delineated, their functions remain unclear, and there remains much to be learned about the detailed mechanism by which eukaryotic ribosomes are generated and transported through the cell. We anticipate a combination of genetic, cell-biological, biochemical, and high-resolution structural approaches in budding yeast will shed light on this poorly understood fundamental process.

Acknowledgments

V. G. Panse is supported by grants from the Swiss National Science Foundation and the ETH Zurich, and is the recipient of a Starting Grant Award (EURIBIO260676) from the European Research Council.

See also: Cell Division/Death: Regulation of Cell Growth: Ribosomal RNA Processing

References

- Allmang, C., Mitchell, P., Petfalski, E., Tollervey, D., 2000. Degradation of ribosomal RNA precursors by the exosome. *Nucleic Acids Research* 28, 1684–1691.
- Altwater, M., Chang, Y., Melnik, A., *et al.*, 2012. Targeted proteomics reveals compositional dynamics of 60S pre-ribosomes after nuclear export. *Molecular Systems Biology* 8, 628.

- Ban, N., Beckmann, R., Cate, J.H., *et al.*, 2014. A new system for naming ribosomal proteins. *Current Opinion in Structural Biology* 24, 165–169.
- Ban, N., Nissen, P., Hansen, J., Moore, P.B., Steitz, T.A., 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. *Science* 289, 905–920.
- Bargis-Survey, P., Lavergne, J.P., Gonzalo, P., *et al.*, 1999. Interaction of elongation factor eEF-2 with ribosomal P proteins. *European Journal of Biochemistry* 262, 606–611.
- Bassler, J., Kallas, M., Pertschy, B., *et al.*, 2010. The AAA-ATPase Rea1 drives removal of biogenesis factors during multiple stages of 60S ribosome assembly. *Molecular Cell* 38, 712–721.
- Bassler, J., Klein, I., Schmidt, C., *et al.*, 2012. The conserved Bud20 zinc finger protein is a new component of the ribosomal 60S subunit export machinery. *Molecular and Cellular Biology* 32, 4898–4912.
- Baxter-Roshek, J.L., Petrov, A.N., Dinman, J.D., 2007. Optimization of ribosome structure and function by rRNA base modification. *PLoS One* 2, e174.
- Bayliss, R., Leung, S.W., Baker, R.P., *et al.*, 2002. Structural basis for the interaction between NTF2 and nucleoporin FxFG repeats. *EMBO Journal* 21, 2843–2853.
- Becam, A.M., Nasr, F., Racki, W.J., Zagulski, M., Herbert, C.J., 2001. Rla1p (Ynl163c), a protein similar to elongation factors 2, is involved in the biogenesis of the 60S subunit of the ribosome in *Saccharomyces cerevisiae*. *Molecular Genetics & Genomics* 266, 454–462.
- Ben-Shem, A., Garreau De Loubresse, N., Melnikov, S., *et al.*, 2011. The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science* 334, 1524–1529.
- Berk, V., Cate, J.H., 2007. Insights into protein biosynthesis from structures of bacterial ribosomes. *Current Opinion in Structural Biology* 17, 302–309.
- Bradatsch, B., Katahira, J., Kowalinski, E., *et al.*, 2007. Arx1 functions as an unorthodox nuclear export receptor for the 60S preribosomal subunit. *Molecular Cell* 27, 767–779.
- Bradatsch, B., Leidig, C., Granneman, S., *et al.*, 2012. Structure of the pre-60S ribosomal subunit with nuclear export factor Arx1 bound at the exit tunnel. *Nature Structural & Molecular Biology* 19, 1234–1241.
- Bradshaw, R.A., Brickley, W.W., Walker, K.W., 1998. N-terminal processing: The methionine aminopeptidase and N alpha-acetyl transferase families. *Trends in Biochemical Sciences* 23, 263–267.
- Bussiere, C., Hashem, Y., Arora, S., Frank, J., Johnson, A.W., 2012. Integrity of the P-site is probed during maturation of the 60S ribosomal subunit. *Journal of Cell Biology* 197, 747–759.
- Clissold, P.M., Ponting, C.P., 2000. PIN domains in nonsense-mediated mRNA decay and RNAi. *Current Biology* 10, R888–R890.
- Cole, S.E., LaRiviere, F.J., Merrick, C.N., Moore, M.J., 2009. A convergence of rRNA and mRNA quality control pathways revealed by mechanistic analysis of nonfunctional rRNA decay. *Molecular Cell* 34, 440–450.
- Connolly, K., Culver, G., 2009. Deconstructing ribosome construction. *Trends in Biochemical Sciences* 34, 256–263.
- Dalley, J.A., Selkirk, A., Pool, M.R., 2008. Access to ribosomal protein Rpl25p by the signal recognition particle is required for efficient cotranslational translocation. *Molecular Biology of the Cell* 19, 2876–2884.
- Demoinet, E., Jacquier, A., Iutilla, G., Fromont-Racine, M., 2007. The Hsp40 chaperone Jij1 is required for the nucleocytoplasmic recycling of preribosomal factors in *Saccharomyces cerevisiae*. *RNA* 13, 1570–1581.
- Dez, C., Houseley, J., Tollervey, D., 2006. Surveillance of nuclear-restricted pre-ribosomes within a subnuclear region of *Saccharomyces cerevisiae*. *EMBO Journal* 25, 1534–1546.
- Dragon, F., Gallagher, J.E., Compagnone-Post, P.A., *et al.*, 2002. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. *Nature* 417, 967–970.
- Faza, M.B., Chang, Y., Occhipinti, L., Kemmler, S., Panse, V.G., 2012. Role of Mex67-Mtr2 in the nuclear export of 40S pre-ribosomes. *PLoS Genetics* 8, e1002915.
- Fribourg, S., Conti, E., 2003. Structural similarity in the absence of sequence homology of the messenger RNA export factors Mtr2 and p15. *EMBO Reports* 4, 699–703.
- Gadal, O., Strauss, D., Braspenning, J., *et al.*, 2001. A nuclear AAA-type ATPase (Rix7p) is required for biogenesis and nuclear export of 60S ribosomal subunits. *EMBO Journal* 20, 3695–3704.
- Geerlings, T.H., Faber, A.W., Bister, M.D., Vos, J.C., Raue, H.A., 2003. Rio2p, an evolutionarily conserved, low abundant protein kinase essential for processing of 20S Pre-rRNA in *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* 278, 22537–22545.
- Grandi, P., Rybin, V., Bassler, J., *et al.*, 2002. 90S Pre-ribosomes include the 35S pre-rRNA, the U3 snoRNP, and 40S subunit processing factors but predominantly lack 60S synthesis factors. *Molecular Cell* 10, 105–115.
- Greber, B.J., Boehringer, D., Montellese, C., Ban, N., 2012. Cryo-EM structures of Arx1 and maturation factors Rei1 and Jij1 bound to the 60S ribosomal subunit. *Nature Structural & Molecular Biology* 19, 1228–1233.
- Hackmann, A., Gross, T., Baierlein, C., Krebber, H., 2011. The mRNA export factor Npl3 mediates the nuclear export of large ribosomal subunits. *EMBO Reports* 12, 1024–1031.
- Hage, A.E., Tollervey, D., 2004. A surfeit of factors: Why is ribosome assembly so much more complicated in eukaryotes than bacteria? *RNA Biology* 1, 10–15.
- Hedges, J., West, M., Johnson, A.W., 2005. Release of the export adapter, Nmd3p, from the 60S ribosomal subunit requires Rpl10p and the cytoplasmic GTPase Lsg1p. *EMBO Journal* 24, 567–579.
- Held, W.A., Ballou, B., Mizushima, S., Nomura, M., 1974. Assembly mapping of 30S ribosomal proteins from *Escherichia coli*. Further studies. *Journal of Biological Chemistry* 249, 3103–3111.
- Ho, J.H., Kallstrom, G., Johnson, A.W., 2000. Nascent 60S ribosomal subunits enter the free pool bound by Nmd3p. *RNA* 6, 1625–1634.
- Hung, N.J., Lo, K.Y., Patel, S.S., Helmke, K., Johnson, A.W., 2008. Arx1 is a nuclear export receptor for the 60S ribosomal subunit in yeast. *Molecular Biology of the Cell* 19, 735–744.
- Hurt, E., Hannus, S., Schmelz, B., *et al.*, 1999. A novel *in vivo* assay reveals inhibition of ribosomal nuclear export in ran-cycle and nucleoporin mutants. *Journal of Cell Biology* 144, 389–401.
- Jakovljevic, J., De Mayolo, P.A., Miles, T.D., *et al.*, 2004. The carboxy-terminal extension of yeast ribosomal protein S14 is necessary for maturation of 43S preribosomes. *Molecular Cell* 14, 331–342.
- Kappel, L., Loibl, M., Zisser, G., *et al.*, 2012. Rlp24 activates the AAA-ATPase Drg1 to initiate cytoplasmic pre-60S maturation. *Journal of Cell Biology* 199, 771–782.
- Karl, T., Onder, K., Kodzius, R., *et al.*, 1999. GRC5 and NMD3 function in translational control of gene expression and interact genetically. *Current Genetics* 34, 419–429.
- Kemmler, S., Occhipinti, L., Veisu, M., Panse, V.G., 2009. Yvh1 is required for a late maturation step in the 60S biogenesis pathway. *Journal of Cell Biology* 186, 863–880.
- Klinge, S., Voigts-Hoffmann, F., Leibundgut, M., Arpagaus, S., Ban, N., 2011. Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. *Science* 334, 941–948.
- Kressler, D., Hurt, E., Bergler, H., Bassler, J., 2012. The power of AAA-ATPases on the road of pre-60S ribosome maturation—molecular machines that strip pre-ribosomal particles. *Biochimica et Biophysica Acta* 1823, 92–100.
- Kressler, D., Linder, P., De La Cruz, J., 1999. Protein trans-acting factors involved in ribosome biogenesis in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 19, 7897–7912.
- Kressler, D., Roser, D., Pertschy, B., Hurt, E., 2008. The AAA ATPase Rix7 powers progression of ribosome biogenesis by stripping Nsa1 from pre-60S particles. *Journal of Cell Biology* 181, 935–944.
- Lafontaine, D., Delcour, J., Glasser, A.L., Desgres, J., Vandenhaute, J., 1994. The DIM1 gene responsible for the conserved m6(2)Am6(2)A dimethylation in the 3'-terminal loop of 18S rRNA is essential in yeast. *Journal of Molecular Biology* 241, 492–497.
- Lafontaine, D.L., 2015. Noncoding RNAs in eukaryotic ribosome biogenesis and function. *Nature Structural & Molecular Biology* 22, 11–19.
- Lafontaine, D.L., Preiss, T., Tollervey, D., 1998. Yeast 18S rRNA dimethylase Dim1p: A quality control mechanism in ribosome synthesis? *Molecular and Cellular Biology* 18, 2360–2370.
- Lamanna, A.C., Karbstein, K., 2009. Nob1 binds the single-stranded cleavage site D at the 3'-end of 18S rRNA with its PIN domain. *Proceedings of the National Academy of Sciences of the United States of America* 106, 14259–14264.
- Lariviere, F.J., Cole, S.E., Ferullo, D.J., Moore, M.J., 2006. A late-acting quality control process for mature eukaryotic rRNAs. *Molecular Cell* 24, 619–626.
- Lebaron, S., Schneider, C., Van Nues, R.W., *et al.*, 2012. Proofreading of pre-40S ribosome maturation by a translation initiation factor and 60S subunits. *Nature Structural & Molecular Biology* 19, 744–753.
- Lo, K.Y., Li, Z., Bussiere, C., *et al.*, 2010. Defining the pathway of cytoplasmic maturation of the 60S ribosomal subunit. *Molecular Cell* 39, 196–208.
- Lo, K.Y., Li, Z., Wang, F., Marcotte, E.M., Johnson, A.W., 2009. Ribosome stalk assembly requires the dual-specificity phosphatase Yvh1 for the exchange of Mrt4 with P0. *Journal of Cell Biology* 186, 849–862.

- Matsuo, Y., Granneman, S., Thoms, M., *et al.*, 2014. Coupled GTPase and remodelling ATPase activities form a checkpoint for ribosome export. *Nature* 505, 112–116.
- Menne, T.F., Goyenechea, B., Sanchez-Puig, N., *et al.*, 2007. The Shwachman-Bodian-Diamond syndrome protein mediates translational activation of ribosomes in yeast. *Nature Genetics* 39, 486–495.
- Meyer, A.E., Hung, N.J., Yang, P., Johnson, A.W., Craig, E.A., 2007. The specialized cytosolic J-protein, Jj1, functions in 60S ribosomal subunit biogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 104, 1558–1563.
- Milkereit, P., Strauss, D., Bassler, J., *et al.*, 2003. A Noc complex specifically involved in the formation and nuclear export of ribosomal 40S subunits. *Journal of Biological Chemistry* 278, 4072–4081.
- Miller Jr., O.L., Beatty, B.R., 1969. Visualization of nucleolar genes. *Science* 164, 955–957.
- Mitchell, P., Petfalski, E., Shevchenko, A., Mann, M., Tollervey, D., 1997. The exosome: A conserved eukaryotic RNA processing complex containing multiple 3'→5' exoribonucleases. *Cell* 91, 457–466.
- Moy, T.I., Silver, P.A., 1999. Nuclear export of the small ribosomal subunit requires the ran-GTPase cycle and certain nucleoporins. *Genes & Development* 13, 2118–2133.
- Moy, T.I., Silver, P.A., 2002. Requirements for the nuclear export of the small ribosomal subunit. *Journal of Cell Science* 115, 2985–2995.
- Nissan, T.A., Bassler, J., Petfalski, E., Tollervey, D., Hurt, E., 2002. 60S pre-ribosome formation viewed from assembly in the nucleolus until export to the cytoplasm. *EMBO Journal* 21, 5539–5547.
- Nissan, T.A., Galani, K., Maco, B., *et al.*, 2004. A pre-ribosome with a tadpole-like structure functions in ATP-dependent maturation of 60S subunits. *Molecular Cell* 15, 295–301.
- Occhipinti, L., Chang, Y., Altwater, M., *et al.*, 2013. Non-FG mediated transport of the large pre-ribosomal subunit through the nuclear pore complex by the mRNA export factor Gle2. *Nucleic Acids Research* 41, 8266–8279.
- Oeffinger, M., Dlakic, M., Tollervey, D., 2004. A pre-ribosome-associated HEAT-repeat protein is required for export of both ribosomal subunits. *Genes & Development* 18, 196–209.
- Panse, V.G., Johnson, A.W., 2010. Maturation of eukaryotic ribosomes: Acquisition of functionality. *Trends in Biochemical Sciences* 35, 260–266.
- Panse, V.G., Kressler, D., Pauli, A., *et al.*, 2006. Formation and nuclear export of preribosomes are functionally linked to the small-ubiquitin-related modifier pathway. *Traffic* 7, 1311–1321.
- Pertschy, B., Saveanu, C., Zisser, G., *et al.*, 2007. Cytoplasmic recycling of 60S preribosomal factors depends on the AAA protein Drg1. *Molecular and Cellular Biology* 27, 6581–6592.
- Pertschy, B., Schneider, C., Gnadig, M., *et al.*, 2009. RNA helicase Prp43 and its co-factor Plat promote 20 to 18S rRNA processing catalyzed by the endonuclease Nob1. *Journal of Biological Chemistry* 284, 35079–35091.
- Rabl, J., Leibundgut, M., Ataide, S.F., Haag, A., Ban, N., 2011. Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science* 331, 730–736.
- Ramakrishnan, V., 2014. The ribosome emerges from a black box. *Cell* 159, 979–984.
- Ribbeck, K., Gorlich, D., 2002. The permeability barrier of nuclear pore complexes appears to operate via hydrophobic exclusion. *EMBO Journal* 21, 2664–2671.
- Ribbeck, K., Lipowsky, G., Kent, H.M., Stewart, M., Gorlich, D., 1998. NTF2 mediates nuclear import of Ran. *EMBO Journal* 17, 6587–6598.
- Russell, D.W., Spremulli, L.L., 1979. Purification and characterization of a ribosome dissociation factor (eukaryotic initiation factor 6) from wheat germ. *Journal of Biological Chemistry* 254, 8796–8800.
- Santos-Rosa, H., Moreno, H., Simos, G., *et al.*, 1998. Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the nuclear pores. *Molecular and Cellular Biology* 18, 6826–6838.
- Schafer, T., Maco, B., Petfalski, E., *et al.*, 2006. Hrr25-dependent phosphorylation state regulates organization of the pre-40S subunit. *Nature* 441, 651–655.
- Schutz, S., Fischer, U., Altwater, M., *et al.*, 2014. A RanGTP-independent mechanism allows ribosomal protein nuclear import for ribosome assembly. *eLife* 3, e03473.
- Schutz, S., Panse, V.G., 2012. Getting ready to commit: Ribosomes rehearse translation. *Nature Structural & Molecular Biology* 19, 861–862.
- Segref, A., Sharma, K., Doye, V., *et al.*, 1997. Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A) + RNA and nuclear pores. *EMBO Journal* 16, 3256–3271.
- Senay, C., Ferrari, P., Rocher, C., *et al.*, 2003. The Mtr2-Mex67 NTF2-like domain complex. Structural insights into a dual role of Mtr2 for yeast nuclear export. *Journal of Biological Chemistry* 278, 48395–48403.
- Senger, B., Lafontaine, D.L., Grainger, J.S., *et al.*, 2001. The nucleolar Tif6p and Efl1p are required for a late cytoplasmic step of ribosome synthesis. *Molecular Cell* 8, 1363–1373.
- Stage-Zimmermann, T., Schmidt, U., Silver, P.A., 2000. Factors affecting nuclear export of the 60S ribosomal subunit *in vivo*. *Molecular Biology of the Cell* 11, 3777–3789.
- Strasser, K., Bassler, J., Hurt, E., 2000. Binding of the Mex67p/Mtr2p heterodimer to FXFG, GLFG, and FG repeat nucleoporins is essential for nuclear mRNA export. *Journal of Cell Biology* 150, 695–706.
- Strawn, L.A., Shen, T., Wentz, S.R., 2001. The GLFG regions of Nup116p and Nup100p serve as binding sites for both Kap95p and Mex67p at the nuclear pore complex. *Journal of Biological Chemistry* 276, 6445–6452.
- Strunk, B.S., Loucks, C.R., Su, M., *et al.*, 2011. Ribosome assembly factors prevent premature translation initiation by 40S assembly intermediates. *Science* 333, 1449–1453.
- Strunk, B.S., Novak, M.N., Young, C.L., Karbstein, K., 2012. A translation-like cycle is a quality control checkpoint for maturing 40S ribosome subunits. *Cell* 150, 111–121.
- Taura, T., Krebber, H., Silver, P.A., 1998. A member of the Ran-binding protein family, Yrb2p, is involved in nuclear protein export. *Proceedings of the National Academy of Sciences of the United States of America* 95, 7427–7432.
- Tschochner, H., Hurt, E., 2003. Pre-ribosomes on the road from the nucleolus to the cytoplasm. *Trends in Cell Biology* 13, 255–263.
- Turowski, T.W., Lebaron, S., Zhang, E., *et al.*, 2014. Rio1 mediates ATP-dependent final maturation of 40S ribosomal subunits. *Nucleic Acids Research* 42, 12189–12199.
- Ulbrich, C., Diepholz, M., Bassler, J., *et al.*, 2009. Mechanochemical removal of ribosome biogenesis factors from nascent 60S ribosomal subunits. *Cell* 138, 911–922.
- Valenzuela, D.M., Chaudhuri, A., Maitra, U., 1982. Eukaryotic ribosomal subunit anti-association activity of calf liver is contained in a single polypeptide chain protein of Mr=25,500 (eukaryotic initiation factor 6). *Journal of Biological Chemistry* 257, 7712–7719.
- Vanrobays, E., Gelugne, J.P., Gleizes, P.E., Caizergues-Ferrer, M., 2003. Late cytoplasmic maturation of the small ribosomal subunit requires RIO proteins in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 23, 2083–2095.
- Warner, J.R., 1999. The economics of ribosome biosynthesis in yeast. *Trends in Biochemical Sciences* 24, 437–440.
- Wild, T., Horvath, P., Wyler, E., *et al.*, 2010. A protein inventory of human ribosome biogenesis reveals an essential function of exportin 5 in 60S subunit export. *PLoS Biology* 8, e1000522.
- Wimberly, B.T., Brodersen, D.E., Clemons Jr., W.M., *et al.*, 2000. Structure of the 30S ribosomal subunit. *Nature* 407, 327–339.
- Woelford Jr., J.L., Baserga, S.J., 2013. Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*. *Genetics* 195, 643–681.
- Yao, W., Lutzmann, M., Hurt, E., 2008. A versatile interaction platform on the Mex67-Mtr2 receptor creates an overlap between mRNA and ribosome export. *EMBO Journal* 27, 6–16.
- Yao, W., Roser, D., Kohler, A., *et al.*, 2007. Nuclear export of ribosomal 60S subunits by the general mRNA export receptor Mex67-Mtr2. *Molecular Cell* 26, 51–62.
- Yao, Y., Demoinet, E., Saveanu, C., *et al.*, 2010. Ecm1 is a new pre-ribosomal factor involved in pre-60S particle export. *RNA* 16, 1007–1017.
- Zemp, I., Wild, T., O'Donoghue, M.F., *et al.*, 2009. Distinct cytoplasmic maturation steps of 40S ribosomal subunit precursors require hRio2. *Journal of Cell Biology* 185, 1167–1180.
- Zhang, J., Harpicharnchai, P., Jakovljevic, J., *et al.*, 2007. Assembly factors Rpf2 and Rrs1 recruit 5S rRNA and ribosomal proteins rPl5 and rPl11 into nascent ribosomes. *Genes & Development* 21, 2580–2592.

Further Reading

- Fernández-Pevida, A., Kressler, D., de la Cruz, J., 2014. Processing of preribosomal RNA in *Saccharomyces cerevisiae*. *Wiley Interdiscip Reviews, RNA* 6, 191–209.

- Floch, A.G., Palancade, B., Doye, V., 2014. Fifty years of nuclear pores and nucleoplasmic transport studies: Multiple tools revealing complex rules. *Methods in Cell Biology* 122, 1–40.
- Granneman, S., Baserga, S.J., 2005. Crosstalk in gene expression: Coupling and co-regulation of rDNA transcription, pre-ribosome assembly and pre-rRNA processing. *Current Opinion in Cell Biology* 17 (3), 281–286.
- Shajani, A., Sykes, M.T., Williamson, J.R., 2011. Assembly of bacterial ribosomes. *Annual Review of Biochemistry* 80, 501–526.
- Steitz, T.A., 2008. A structural understanding of the dynamic ribosome machine. *Nature Reviews Molecular Cell Biology* 9, 242–253.

Relevant Websites

- http://www.mol.biol.ethz.ch/groups/ban_group/nomenclature
ETH Zürich.
- <http://www.yeastgenome.org>
Saccharomyces Genome Database.
- <http://www.arb-silva.de>
Silva ribosomal RNA database.
- <http://people.biochem.umass.edu/fournierlab/snornadb/main.php>
snoRNA database.

Internal Ribosome Entry Site-Mediated Translation

N Sonenberg, Y Svitkin, and N Siddiqui, McGill University, Montreal, QC, Canada

© 2016 Elsevier Inc. All rights reserved.

Glossary

A site The point of entry for the amino-acyl tRNA except for Met-tRNA_i on the 40S ribosomal subunit.

Cap structure A 7-methylguanosine linked to the first nucleotide of eukaryotic mRNA by an inverted 5'-5' triphosphate bond. Its primary functions are to stabilize mRNA and to promote its translation.

Cistron A tract of DNA encoding a protein or a functional RNA molecule.

Codon A sequence of three nucleotides encoding a specific amino acid.

E site The exit site for deacylated tRNA on the ribosome.

Initiator tRNA A methionyl-tRNA (Met-tRNA_i) that base pairs with the AUG initiation codon and initiates protein

synthesis. It forms a ternary complex with eIF2 and GTP and binds the ribosomal P site.

P site The site on the 40S ribosomal subunit that accommodates Met-tRNA_i or peptidyl-tRNA.

Picornaviruses Small viruses belonging to the family Picornaviridae. Members of this family contain a single-stranded RNA genome of positive polarity. Many picornaviruses are important human and animal pathogens.

Poly(A) tail A stretch of adenine residues, generally of more than 100 nucleotides, at the 3' end of most eukaryotic mRNAs.

Translation A part of gene expression mechanism in which ribosomes synthesize proteins. It proceeds in three phases: initiation, elongation, and termination.

Introduction: The Canonical Cap-Dependent Mechanism of Translation Initiation in Eukaryotes

In eukaryotes, the formation of the 80S ribosomal initiation complex is the rate-limiting step in translation and a frequent target for translational control. On most mRNAs, translation is initiated by the scanning mechanism that involves the mRNA 5' cap structure (m⁷GpppN, where N is any nucleotide) and the cap-binding eukaryotic initiation factor (eIF) 4F (a heterotrimeric complex comprised of a cap-binding subunit, eIF4E, an ATP-dependent RNA-helicase, eIF4A, and a large scaffolding protein, eIF4G). The association of eIF4F with the 5' end of the mRNA is additionally stabilized by the poly(A) binding protein, which simultaneously interacts with eIF4G and the 3' poly(A) tail and confers on mRNA a circular conformation. The 5' end of the mRNA recruits the 40S ribosomal subunit preloaded with eIF1, eIF1A, eIF3, eIF5, and the ternary complex eIF2•GTP•Met-tRNA_i (termed as 43S complex) via the eIF3-eIF4G interaction. The 43S complex then scans the 5'-untranslated region (5'-UTR), aided by the unwinding of secondary structures using energy provided by ATP hydrolysis, until the AUG initiation codon in the optimal context (G/AXXAUGG, where X is any nucleotide) is encountered. In addition to eIF4A and its cofactor eIF4B, some other RNA helicases, such as DHX29, have been implicated in the removal of secondary structures during scanning (Jackson *et al.*, 2010; Hinnebusch, 2014). Subsequent joining of the 60S subunit accomplishes the initiation step of protein synthesis. This step requires hydrolysis of eIF2-bound GTP and the displacement of eIFs from the 40S subunit (promoted by eIF5 and eIF5B•GTP, respectively).

Although the mechanism of the 43S complex migration along the mRNA remains enigmatic even to this day, the scanning model is consistent with most experimental data on translation of eukaryotic mRNAs and nicely explains why these mRNAs are overwhelmingly monocistronic. Accordingly, this

model has not undergone substantial revision since its original proposal by Kozak (1978).

Internal Ribosome Entry Site-Mediated Translation Initiation

Discovery

The first exception to the 5' cap-dependent scanning model of translation initiation emerged from the studies of picornaviruses, which possess naturally uncapped, but yet efficiently translated RNA genomes. Other features of picornavirus RNAs that are not compatible with the cap-dependent scanning mechanism of translation are the presence of unusually long (610–1500 nt) and highly structured 5' UTR with multiple nonfunctional AUG triplets that are poorly conserved between strains. Furthermore, inactivation of the eIF4F cap-binding protein complex after picornavirus infection, which occurs either as a result of eIF4G cleavage or dephosphorylation and activation of eIF4E-binding proteins (4E-BPs), does not impede but rather stimulate the translation of viral mRNAs while effectively shutting off cap-dependent host cell protein synthesis.

In 1988, a landmark discovery explaining this distinct regulation of picornavirus genome translation was made in the laboratories of Nahum Sonenberg and Eckard Wimmer. In these experiments, internal segments of the 5' UTRs of poliovirus (PV) and encephalomyocarditis virus (EMCV), referred to as internal ribosome entry site (IRES) or ribosome landing pad, were shown capable of initiating translation independently of the cap structure (Jang *et al.*, 1988; Pelletier and Sonenberg, 1988). Soon after this discovery, IRES elements were found in the genomes of other picornaviruses as well as in many non-picornavirus mRNAs of both viral and cellular origin.

Assay

IRESs are typically identified by inserting the mRNA 5' UTR between two cistrons of a bicistronic mRNA (Figure 1). Efficient translation of the second cistron, even when the first cistron is not translated, such as when preceded by a stable stem loop structure or in poliovirus (PV) infected cells, provides a proof for 5'-end independent internal ribosome loading onto the 5' UTRs (Jang *et al.*, 1988; Pelletier and Sonenberg, 1988).

Although the use of bicistronic reporter constructs (usually dual *Renilla* luciferase and firefly luciferase reporter gene constructs) is considered to be the 'gold standard' for detecting of IRES activity, there are several caveats to be considered when interpreting the data. For example, the apparent IRES activity of some cellular mRNA 5' UTRs in the bicistronic construct assay, could be explained by the generation of aberrant monocistronic mRNAs due to cryptic transcription or the use of cryptic splice acceptor sites within the candidate IRES segments (see below).

As an alternative to the method of bicistronic constructs, IRESs can be identified by their ability to direct translation of downstream cistrons in circular RNAs (Chen and Samow, 1995).

Viral IRESs

Classification

A large number of positive strand RNA viruses, such as picornaviruses, use IRES-mediated translation to escape translational competition with the capped cellular mRNAs. Based on their structure and initiation factor requirements, viral IRESs are classified into four major types: (1) *Picornaviridae* enterovirus and rhinovirus (i.e., PV) IRESs, (2) *Picornaviridae* cardio-, aphtho-, and parechovirus (i.e., EMCV) IRESs, (3) Hepatitis C virus (HCV), pestivirus (i.e., classical swine fever (CSFV)) and other HCV-like IRESs, and (4) the dicistrovirus (i.e., cricket paralysis virus (CPV) intergenic region (IGR) IRESs (Hellen, 2009).

Unclassified viral IRESs include the picornavirus hepatitis A virus (HAV), the retrovirus human immunodeficiency virus 1 (HIV-1) and mRNAs of DNA-containing herpesviruses. Some RNA-containing plant viruses were also reported to contain IRES (Lopez-Lastra *et al.*, 2005). For representative viral IRESs, see Table 1.

Structure and mechanism

There is no structural similarity between different types of viral mRNA IRESs. However, despite their structural and functional

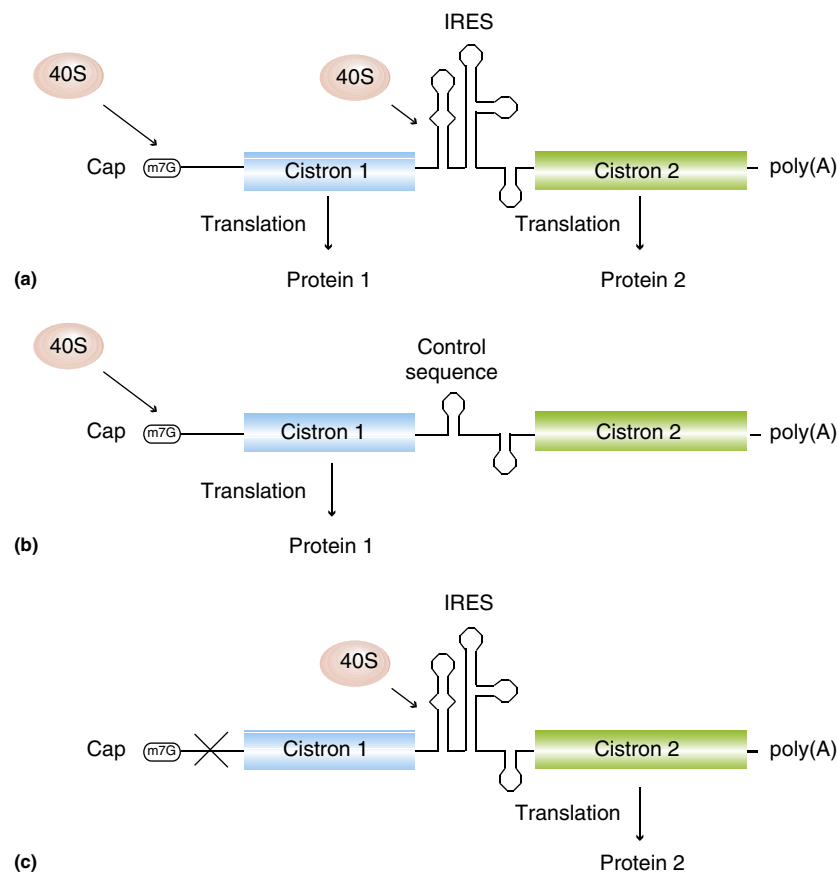


Figure 1 Schematic representation of the bicistronic construct assay. (a and b) In mRNAs transcribed from bicistronic constructs, the first cistron is translated by a cap-dependent scanning mechanism while translation of the second cistron depends on the presence of IRES in the intercistronic spacer. (c) Ribosome recruitment by IRES is independent of the first cistron expression. A 5'-UTR sequence is qualified as IRES if it significantly increases the Cistron-2/Cistron-1 expression ratio after insertion in the construct. For simplicity, the diagram does not show the 60S ribosomal subunits, eIFs and ITAFs, which participate in these processes.

Table 1 Representative examples of viral mRNA internal ribosome entry sites (IRESs)

<i>Virus and abbreviation</i>	<i>IRES</i>
<i>Picornaviridae</i>	
<i>Enterovirus</i>	Type 1
Poliovirus (PV)	
Coxsackievirus B3 (CVB3)	
Enterovirus 71 (EV71)	
Bovine enterovirus (BEV)	
<i>Rhinovirus</i>	
Human rhinovirus (HRV)	
<i>Picornaviridae</i>	
<i>Cardiovirus</i>	Type 2
Encephalomyocarditis virus (EMCV)	
Theiler's murine encephalomyelitis virus (TMEV)	
<i>Aphthovirus</i>	
Foot-and-mouth disease virus (FMDV)	
<i>Parechovirus</i>	
Human parechovirus 1 (HPEV1)	
<i>Flaviviridae</i>	
<i>Hepacivirus</i>	Type 3
Hepatitis C virus (HCV)	
GB virus B (GBV-B)	
<i>Pestivirus</i>	
Classical swine fever virus (CSFV)	
Bovine virus diarrhea virus (BVDV)	
<i>Picornaviridae</i>	
<i>Teschovirus</i>	
Porcine teschovirus 1 (PTV-1)	
<i>Enterovirus</i>	
Simian picornavirus type 9 (SPV9)	
<i>Dicistroviridae</i>	
<i>Dicistrovirus</i>	Type 4
Cricket paralysis virus (CrPV)	
<i>Plautia stali</i> intestine virus (PSIV)	
<i>Rhopalosiphum padi</i> virus (RhPV)	
<i>Picornaviridae</i>	
<i>Hepatovirus</i>	Other
Hepatitis A virus (HAV)	
<i>Kobuvirus</i>	
Aichivirus (AV)	

Source: Data compiled from Hellen, C.U., Sarnow, P., 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes & Development* 15, 1593–1612; Baird, S. D., Turcotte, M., Korneluk, R.G., Holcik, M., 2006. Searching for IRES. *RNA* 12, 1755–1785; Fitzgerald, K.L., Semler, P., 2009. Bridging IRES elements in mRNAs to the eukaryotic translation apparatus. *Biochimica et Biophysica Acta* 1789, 518–528; and Mokrejs, M., Masek, T., Vopalensky, V., *et al.*, 2010. IRESite – A tool for the examination of viral and cellular internal ribosome entry sites. *Nucleic Acids Research* 38, D131–D136.

differences, all four types of IRESs share a common feature, which is the ability to interact with one or more canonical components of the translational machinery, such as eIF4G, eIF3, or 40S ribosomal subunit (Figure 2).

Type 1 and 2 IRESs

The core IRES sequence of picornavirus RNA genomes is ~450 nt in length and consists of five domains, designated as II–VI and H–L in Type 1 and 2 IRESs, respectively. Secondary

structure predictions for picornavirus IRESs are fairly robust, as being validated by mutagenesis and functional assays. Secondary structure, and to some degree primary structure, is conserved within each type. However, sequence similarity between Type 1 and 2 IRESs is limited to a 3' terminal Y_n - X_m -AUG motif (where Y_n is an oligopyrimidine tract and X_m is a 18–20 nt spacer between Y_n and an AUG triplet). In EMCV, translation starts at the 3' terminal AUG of the IRES, while in foot-and-mouth disease virus (FMDV) both this and the next downstream initiator AUG is used. In PV and human rhinovirus (HRV), the initiator AUG is spaced out from the 3' border of the IRES by ~160 and ~30 nucleotides, respectively, so that the 43S complex must scan downstream from the silent AUG to reach the initiation codon. In Type 1 IRESs, the intact Y_n - X_m -AUG motif is important for initiation at the downstream AUG. Even small alterations of this motif impair IRES activity and viral growth (Pilipenko *et al.*, 1992). Structural elements upstream of the Y_n - X_m -AUG motif can also modulate PV IRES activity.

An important functional distinction between Type 1 and 2 IRESs concerns their ability to function in a rabbit reticulocyte lysate. While Type 2 IRESs are fully active in this system, translation from Type 1 IRESs is inefficient unless the lysate is supplemented with extracts of nucleated cells (see below).

Type 3 HCV-like IRESs

The ~300 nt-long HCV IRES encompasses three main structural domains, which adopt a tertiary fold (Honda *et al.*, 1999). The negative effect of several domain mutations on IRES activity is averted by compensatory mutations, which validates the proposed base pairing. The AUG initiation codon is located near the 3' end of the IRES. By now, many three-dimensional (3-D) structures of HCV and CSFV IRESs have been solved using NMR and X-ray crystallography (Berry *et al.*, 2011). The cryo-EM studies demonstrated that HCV IRES binds to the solvent-exposed side of the 40S subunit in extended conformation, docks its domain into the ribosomal exit (E) site and places the initiator codon in the ribosomal P site (Spahn *et al.*, 2001). Despite their ability to establish binary complexes with the 40S subunit, both HCV and CSFV IRESs require a limited set of eIFs to initiate translation (see below).

Importantly, the 40S subunit undergoes significant conformational changes following interaction with Type 3 IRESs (Fraser and Doudna, 2007). It is presumed that these changes promote entry, positioning and fixation of mRNA in the ribosomal decoding channel as well as subunit joining.

Type 4 Dicistroviridae IGR IRESs

The 190–220 nt-long IGR IRESs of dicistroviruses, which are located between the cistrons for structural and nonstructural proteins, adopt a tightly folded globular structure with two conserved stem-loops and three essential pseudoknots (Costantino *et al.*, 2008). Remarkably, these small IRESs do not require any initiation factors, the initiator Met-tRNA_i or a proper AUG initiation codon to assemble elongation-competent 80S initiation complexes.

The CrPV IRES binds directly to 40S subunits or 80S ribosomes where it occupies the inter-subunit space and mimics the base-paired initiation codon and anticodon in

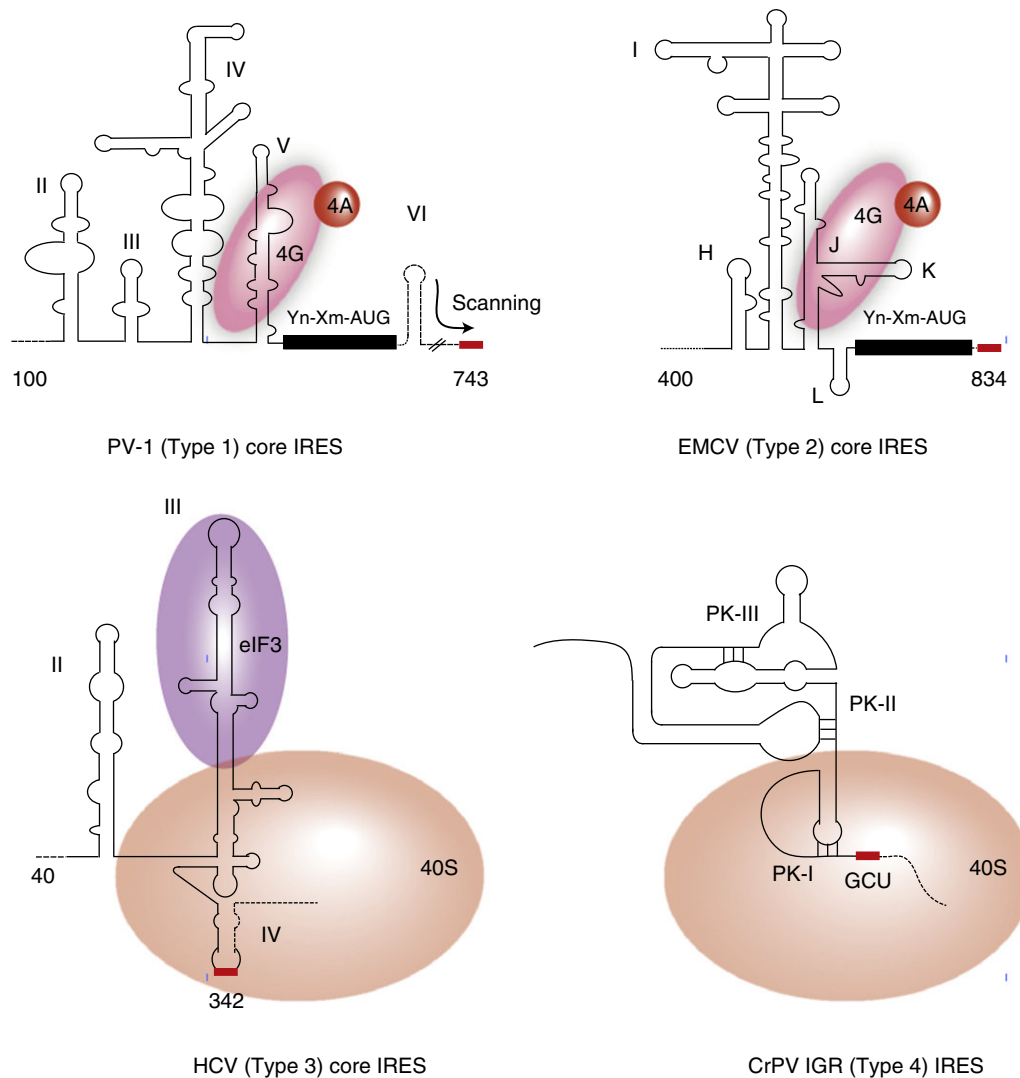


Figure 2 Simplified representation of Type 1, 2, 3, and 4 IRES secondary structure elements and their interaction with the components of the translational machinery. The IRES stem-loop structures, II–VI in PV-1, H–L in EMCV, and II–IV in HCV, and pseudoknots (PK), I–III in CrPV IGR, are indicated. Red rectangles indicate the positions of the AUG and GCU start codons in mRNAs bearing Type 1–3 and Type 4 IRES, respectively. In Type 1 and 2 IRESs, Yn-Xm-AUG represents the 3'-terminal oligopyrimidine tract-AUG motif. Data from Jackson, R.J., 2013. The current status of vertebrate cellular mRNA IRESs. *Cold Spring Harbor Perspectives in Biology* 5, a011569 and Kieft, J.S., 2008. Viral IRES RNA structures and ribosome interactions. *Trends in Biochemical Sciences* 33, 274–283.

the ribosomal P site (Wilson *et al.*, 2000; Jan, 2006). Cryo-EM visualization of the CrPV IRES bound to ribosomes along with biochemical data point to the decisive role of the ribosomal protein rpS25 for IRES binding and indicate that IRES elements manipulate the ribosome to functionally replace initiator Met-tRNA_i and translation factors (Muhs *et al.*, 2011).

During initiation on CrPV IRES, the first codon to be decoded (alanine or glutamine triplet) is located in the A site, and the cognate amino-acyl tRNA is recruited to the A site by eukaryotic elongation factor 1 (eEF1). Amino-acyl tRNA is then translocated to the P site with the aid of eEF2 (Pestova and Hellen, 2003). This step is often referred to as pseudo-translocation, since in contrast to conventional amino-acyl tRNA translocation it occurs before peptide bond formation (Figure 3).

Roles of initiation factors

Canonical eIFs

For the most part, IRES-mediated translation requires fewer eIFs, as compared to cap-dependent translation (Figure 4). Moreover, the ability of Type 4 IRESs to directly interact with the 40S ribosomal subunit circumvents the requirement for all canonical eIFs.

Type 1 (PV), Type 2 (EMCV), and Aichivirus (AV) IRESs do not interact with the 40S subunit directly, but rather through their high-affinity eIF4G-binding sites. These IRESs make use of all the canonical eIFs, with the exception of eIF4E (i.e., eIFs 1, 1A, 2, 3, 4A, 4B, 4G, 5 and 5A), to initiate translation (Pestova *et al.*, 1996; Sweeney *et al.*, 2013). The middle segment of eIF4G that retains eIF4A and eIF3 binding sites is sufficient for initiation complex formation on these IRES.

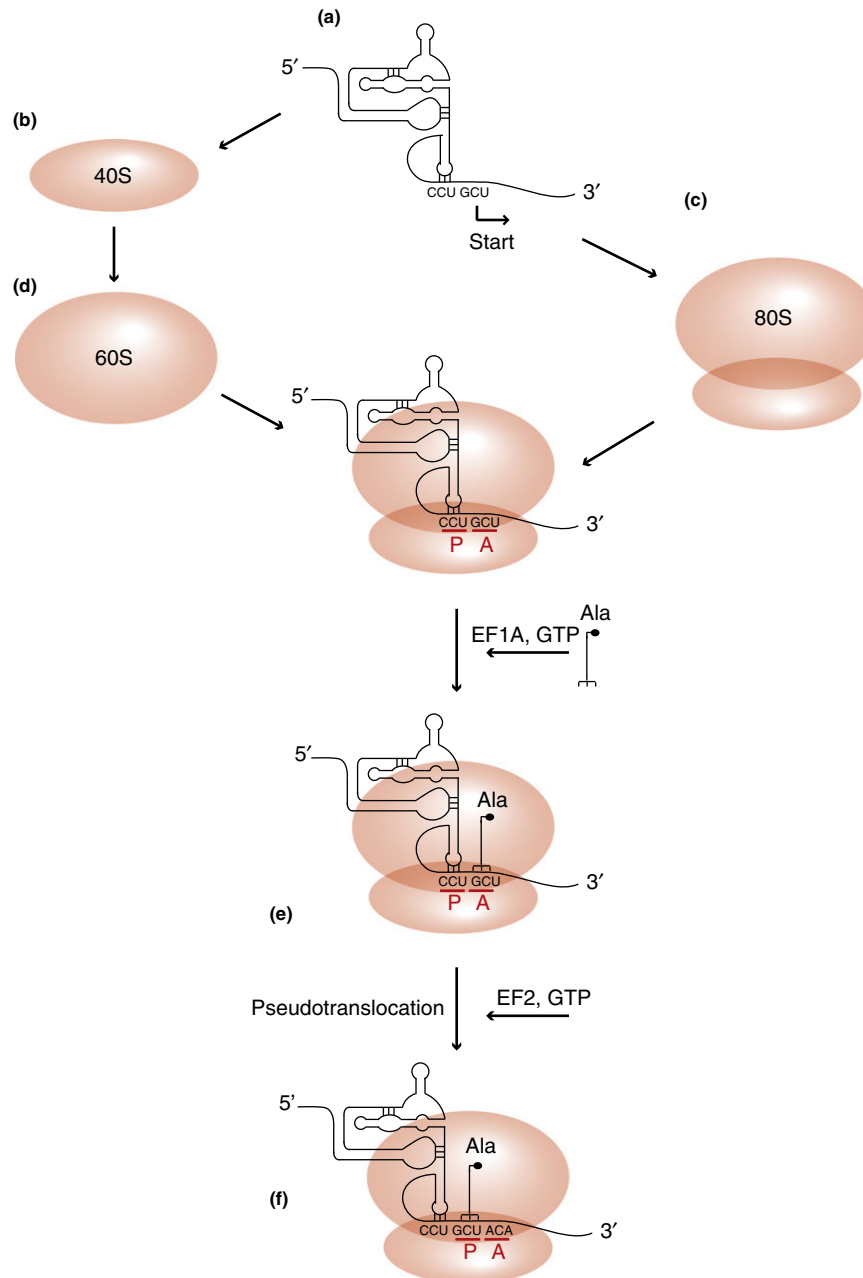


Figure 3 Translation initiation by Type 4 IRESs. (a) Simplified representation of CrPV IRES. (b) The IRES sequentially interacts with the 40S and 60S ribosomal subunits to form the 80S initiation complex. (c) Alternatively, the IRES can bind a preformed 80S ribosome. (d and e) The delivery of the Ala-tRNA to the A site promotes pseudotranslocation with the resulting formation of the elongation-competent 80S complexes. The participation of elongation factors, EF1 and EF2, in initiation of translation by CrPV IRES is also shown. For simplicity, the ribosomal E site is not shown. Data from Jan, E., 2006. Divergent IRES elements in invertebrates. *Virus Research* 119, 16–28 and Kieft, J.S., 2008. Viral IRES RNA structures and ribosome interactions. *Trends in Biochemical Sciences* 33, 274–283.

Initiation on EMCV IRES begins with specific binding of eIF4G to the J-K stem-loops, which is stimulated by eIF4A (Kolupaeva *et al.*, 2003). Initiation on PV IRESs follows a similar pathway, and involves interaction of eIF4G/4A with the basal half of Domain V (de Breyne *et al.*, 2009). The PABP/3'-end poly(A) complex strongly potentiates Type 1 IRESs, presumably by contacting eIF4G and increasing its affinity for the IRES (Svitkin *et al.*, 2001). Mutations made in

the eIF4G/4A binding domain that weaken its secondary structure decrease the efficiency of translation of PV mRNAs *in vitro* and constitute the basis of attenuation of all three Sabin PV vaccine strains (Svitkin *et al.*, 1988; Minor, 1992; Ochs *et al.*, 2003). Conservation of the stem-loop structures in other picornavirus IRESs could be also important for optimal IRES function. The recruitment of eIF4G by Type 1 and 2 IRESs, however, is not sufficient for initiation since the borders

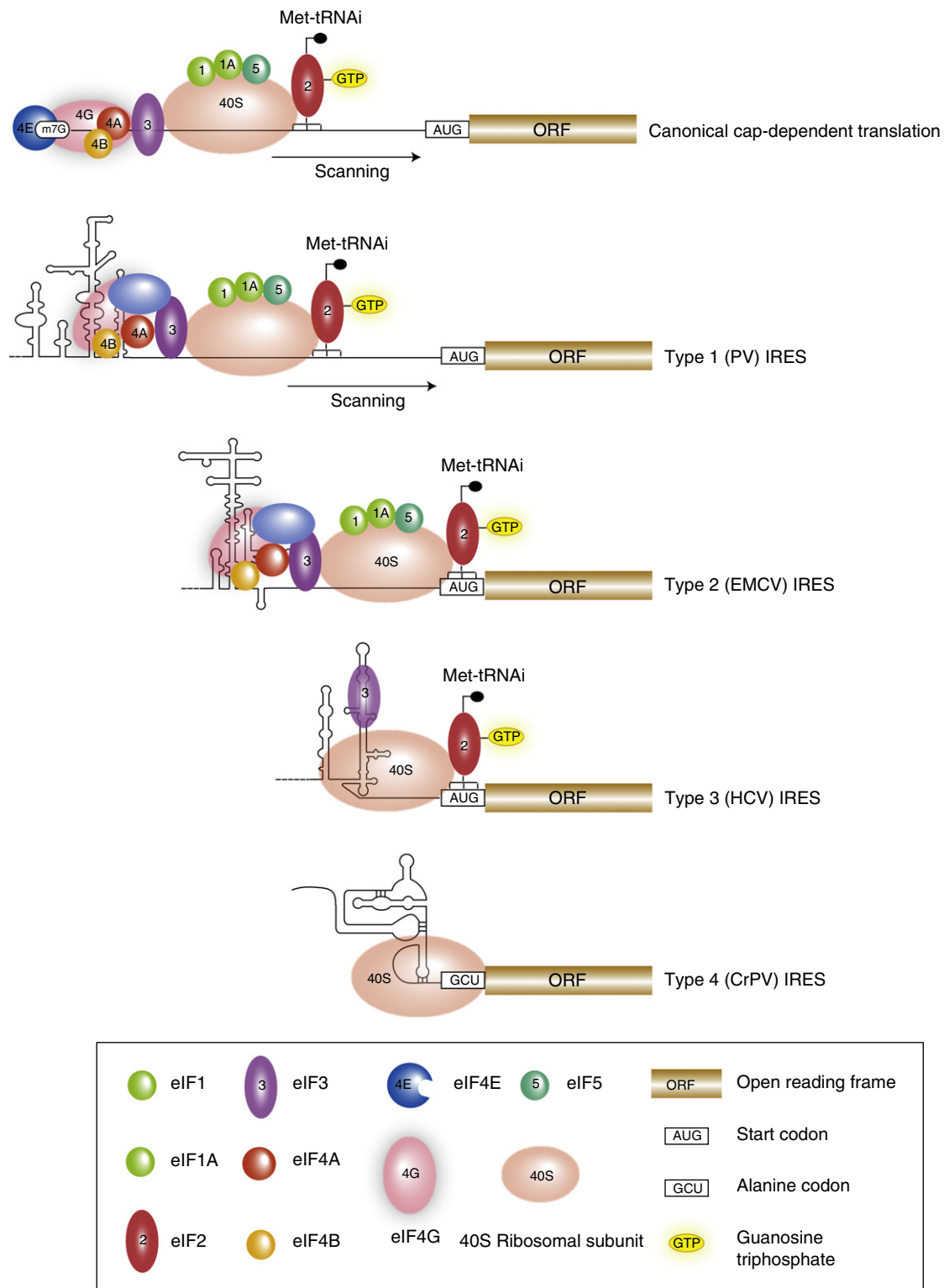


Figure 4 Different pathways of IRES-mediated translation in comparison with cap-dependent translation. For simplicity, only the first most diverse step of the pathways, that is, the recruitment of the 40S ribosomal subunit is shown.

of these IRESs extend far upstream of the eIF4G binding sites. The functions provided by the upstream regions of Type 1 and 2 IRESs remain largely unknown.

In contrast to Type 1 and 2 IRESs, Type 3 (HCV-like) IRESs recruit the 40S ribosomal subunit without the aid of the eIF4G/4A adaptor. The 48S pre-initiation complex on these

IRESs can be formed solely with a 40S ribosomal subunit, Met-tRNA_i, and eIF2 (Pestova *et al.*, 1998). The HCV-like IRESs are capable of establishing multiple contacts with the 40S subunit and eIF3 leading to the positioning of the initiation codon into the ribosomal P site. After base pairing of the initiator AUG with the anticodon of eIF2-bound Met-tRNA_i, a 48S initiator complex is formed. Thus, the factor-less recruitment of the small ribosomal subunit by Type 3 IRESs is akin to that used by prokaryotic mRNAs.

The role of eIF3 in translation initiation by Type 3 IRESs is not clear. Recent cryo-electron microscopy (cryo-EM) 3-D structures of the complex between the 40S subunit, eIF3, and HCV-like CSFV IRES indicate that eIF3 is displaced from its ribosomal position in the 43S complex to establish a new interaction with the apical region of domain III of the IRES (Hashem *et al.*, 2013).

Alternatively, eIF5B (an analog of bacterial IF2) can promote Met-tRNA_i binding to Type 3 IRES-40S complexes, resulting in the assembly of elongation-competent 80S ribosomes (Pestova *et al.*, 2008; Terenin *et al.*, 2008). This eIF2-independent mode of translation might be predominant when eIF2 is inactivated as a result of eIF2 α phosphorylation in response to viral infection or other stresses.

IRES transacting factors

Several RNA-binding proteins that are dispensable for cap-dependent translation, termed IRES transacting factors (ITAFs), modulate IRES activity by the mechanisms that are not well defined. First evidence for ITAFs was provided by the demonstration that HeLa cell extracts strongly stimulate the translation of PV RNA from the authentic initiation site, while reducing the incidence of abnormal internal initiation events, in rabbit reticulocyte lysates (Brown and Ehrenfeld, 1979; Dörner *et al.*, 1984). One of these stimulatory factors was purified and identified as La autoantigen (Meerovitch *et al.*, 1993). polyC binding protein 2 (PCBP2), SRp20, polypyrimidine tract binding protein (PTB), upstream of N-ras (Unr), and hnRNPA1 were also reported to activate Type 1 IRESs (Fitzgerald and Semler, 2009; Jackson, 2013). However, *in vitro* reconstitution of the 48S initiation complex formation on Type 1 IRESs did not reveal the requirement for these ITAFs except for PCBP2 (Sweeney *et al.*, 2013). Type 2 EMCV IRES is activated by PTB, while FMDV IRES requires PTB and the cell-cycle dependent protein ITAF45. Even the HCV IRES, which directly interacts with the 40S subunit, was suggested to be regulated by ITAFs, such as PTB, PCBP2, La, hnRNP D, and hnRNP L.

Mechanistically, ITAFs are thought to function as RNA chaperons that assist IRESs to adopt a functionally active conformation. Indeed, most ITAFs possess multiple RNA-binding domains and are able to dimerize in solution. Realization of these multiple RNA-protein and protein-protein interactions could alter the secondary and tertiary structures of IRESs and facilitate their interaction with the canonical translation components. Further information on RNA-binding proteins with their assigned ITAF's function can be found in (King *et al.*, 2010).

Cellular mRNA IRESs

Evidence for cellular mRNA IRESs

In many physiological and pathological situations, such as mitosis, differentiation, apoptosis, stress (e.g., hypoxia or heat

shock), and viral infection, cap-dependent initiation is compromised, either because of cleavage of the canonical eIFs or a change in the phosphorylation state of eIFs and their binding partners (Holcik and Sonenberg, 2005). IRESs were predicted to exist in a subset of cellular mRNAs whose translation is either maintained or enhanced under the conditions of global protein synthesis inhibition (Hellen and Sarnow, 2001).

The structural features of cellular mRNAs that indicate the existence of IRESs are long and extensively structured 5' UTRs and the presence of AUG triplets within 5' UTRs (which is generally incompatible with efficient scanning). The mRNA encoding BiP (immunoglobulin heavy-chain-binding protein) was first suggested to possess IRES because it is translated in spite of the shutoff of cap-dependent translation after PV infection (Macejak and Sarnow, 1991). Since then, more than hundred cellular mRNAs were reported to contain IRESs (up to 10% of all mRNAs, as suggested by *in silico* analyses) (King *et al.*, 2010; Mokrejs *et al.*, 2010). Proteins encoded by mRNAs possessing putative IRESs are often involved in control of cell growth or cell death (Table 2).

Cellular mRNA IRESs are not amenable to classification because of significant difference in sequence and predicted secondary structure (Baird *et al.*, 2006). In contrast to viral IRESs, whose activity is dependent on the integrity of the structural elements, cellular mRNA IRESs are not defined by overall structure, but instead contain several modules that display IRES activities on their own (Baird *et al.*, 2007).

Similar to Type 1 and 2 viral IRESs, activity of cellular mRNA IRESs generally requires all the canonical eIFs, except for eIF4E. Consistent with this presumption, efficient IRES-mediated translation of vascular endothelial growth factor (VEGF) and p120 catenin mRNAs in some breast cancer cell lines correlates with a high expression of eIF4G relative to eIF4E (Silvera *et al.*, 2009).

Activities of putative cellular mRNA IRESs also appear to be dependent on ITAFs, such as PTB, PCBP2, Unr, La, hnRNPA1 and other hnRNPs (Komar and Hatzoglou, 2005; Baird *et al.*, 2006). For example, La has been reported to activate the IRES of laminin B1 mRNA during epithelial to mesenchymal transition of hepatocellular carcinoma cells (Petz *et al.*, 2011). The list of ITAFs specific for cellular mRNA IRESs is steadily growing (King *et al.*, 2010).

As for viral IRESs, the mechanism of activation of cellular mRNA IRES by ITAFs is not known. In some cases, ITAF binding was found to destabilize certain structures within IRESs, which could be inhibitory for the recruitment of ribosomes or other factors (Mitchell *et al.*, 2003). There are also precedents of negative regulation of IRESs by ITAFs (King *et al.*, 2010).

It is noteworthy that many ITAFs shuttle between the nucleus and the cytoplasm. Changes in subcellular distribution of ITAFs under different physiological conditions can provide an important means to regulate IRES-mediated translation (Lewis and Holcik, 2008).

Critical analysis of cellular mRNA IRESs

All cellular mRNAs contain m⁷G-cap at their 5' end and therefore can be translated by the canonical cap-dependent mechanism. Although long and extensively structured 5' UTRs generally inhibit scanning, they are not incompatible with the

Table 2 Examples of cellular mRNAs reported to contain IRESs

Protein function	mRNA/IRES
Growth factors	Fibroblast growth factors (FGF) Insulin-like growth factor 2 (IGF-II) Vascular endothelial growth factor (VEGF)
Oncogenes	c-myc Cyclin-dependent protein kinase PITSRE
Transcription factors	AML1/RUNX1 Antennapedia MYT2 NF- κ B repressing factor NRF Ultrabithorax
Translation factors	Death associated protein 5 (DAP5) eIF4G1
Transporters and receptors	Cationic amino acid transporter Cat-1 Nuclear form of Notch 2
Activators of apoptosis	Apoptotic protease activating factor (Apaf-1)
Inhibitors of apoptosis	Bcl-2 HIAP2 X-linked inhibitor of apoptosis (XIAP)
Neuronal dendrite proteins	Activity-regulated cytoskeletal protein (ARC) α -subunit of calcium calmodulin-dependent kinase II dendrin Amyloid precursor protein Microtubule-associated protein 2 (MAP2) Neurogranin (RC3)
Other proteins	APC β -subunit of mitochondrial H ⁺ -ATP synthetase Connexins 26, 32 and 43 Heat shock protein 70 HIF-1 α , HIF-2 α , Immunoglobulin heavy-chain-binding protein (BiP) Ornithine decarboxylase (ODC) Upstream of N-ras (Unr)

Source: Data compiled from Hellen, C.U., Sarnow, P., 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes & Development* 15, 1593–1612; Baird, S. D., Turcotte, M., Korneluk, R.G., Holcik, M., 2006. Searching for IRES. *RNA* 12, 1755–1785; Fitzgerald, K.L., Semler, P., 2009. Bridging IRES elements in mRNAs to the eukaryotic translation apparatus. *Biochimica et Biophysica Acta* 1789, 518–528, and Mokrejs, M., Masek, T., Vopalensky, V., *et al.*, 2010. IRESite – A tool for the examination of viral and cellular internal ribosome entry sites. *Nucleic Acids Research* 38, D131–D136.

cap-dependent translation at least in some cases. It is therefore uncertain whether the IRES-mediated translation of a subset of cellular mRNAs can occur in parallel with cap-dependent scanning and even predominate over the latter mechanism in some circumstances.

First, it could be argued that the affinity of eIF4F for the 5'-end is not the same for all mRNAs. Thus, partial eIF4F inactivation under stress conditions while inhibiting the translation of bulk mRNAs could confer competitive advantage to select mRNAs that have higher affinity for eIF4F (Jackson, 2013). In addition, ribosome shunting, in which ribosomes reach the initiator AUG by bypassing large segments of the 5' UTR, is expected to be favored over scanning when eIF4F activity is reduced (Yueh and Schneider, 1996). Finally, the 'cap- and IRES-independent' scanning mechanism of translation initiation, in which scanning from the 5' end relies on the

interaction of eIF4F with specialized RNA structures called cap-independent translation enhancers, akin to initiation on some plant viral mRNAs, can potentially operate under conditions in which general protein synthesis is inhibited (Shatsky *et al.*, 2010).

With these multiple possibilities of translational persistence of select mRNAs under stress conditions, the question arises of whether initiation via cellular mRNA IRESs is of any physiological significance. Indeed, most of the evidence for activity of cellular mRNA IRESs rests on the method of bicistronic DNA constructs, which could be prone to misinterpretation (Jackson, 2013). In this assay, a fraction of monocistronic mRNAs could be transcribed from cryptic promoters within candidate IRES segments or even within vector backbones if the putative IRESs possess 3'-splice sites (Van Eden *et al.*, 2004; Baranick *et al.*, 2008; Lemp *et al.*, 2012). Common tests to detect monocistronic mRNAs, such as Northern blotting and qPCR, are not always adequate, as they do not provide sufficient sensitivity. A more rigorous test is to co-express a siRNA that targets the first cistron coding sequence (Van Eden *et al.*, 2004). If the siRNA causes no reduction in the second cistron expression, the most likely scenario is that this expression is driven by unintended monocistronic mRNAs. Reevaluation of several cellular mRNAs 5' UTRs by this and other stringent tests failed to confirm the presence of IRESs. To circumvent the above problems, some researches use RNA transfection or *in vitro* translation assays. However, many putative cellular mRNA IRESs appear to be much less active than their viral counterparts in these assays (Shatsky *et al.*, 2010).

Conclusion and Perspective

Since the original discovery of IRESs in picornaviruses (Jang *et al.*, 1988; Pelletier and Sonenberg, 1988), the cap- and 5'-end-independent mechanism of translational initiation has gained significant support in numerous publications. This alternative pathway of translation initiation has apparently evolved to ensure synthesis of select proteins under stress and other physiological and pathological conditions where cap-dependent initiation of translation is inhibited. Deregulation of IRESs occurs in diseases, as putative IRESs are present in mRNAs that encode oncogenes, growth factors, and regulators of apoptosis. Many viruses of significant medical importance use IRESs to synthesize their proteins. These IRESs exhibit a wide diversity with respect to their structure and eIFs requirements. Accordingly, viral IRESs and their ITAFs represent potential targets for the development of novel antiviral therapeutics. Finally, IRESs have emerged as biotechnology tools, in particular for gene therapy, as they function under variable physiological conditions and provide a unique means for co-expression of several proteins from one polycistronic mRNA.

See also: Cell Division/Death: Regulation of Cell Growth: Targeted mRNA Degradation. Nucleic Acid Synthesis/ Breakdown: RNA Synthesis/Function: Messenger RNA (mRNA): The Link between DNA and Protein; Transfer RNA. Protein

Synthesis/Degradation: Translation: Components, Initiation, Elongation, Termination, and Regulation

References

- Baird, S.D., Lewis, S.M., Turcotte, M., Holcik, M., 2007. A search for structurally similar cellular internal ribosome entry sites. *Nucleic Acids Research* 35, 4664–4677.
- Baird, S.D., Turcotte, M., Korneluk, R.G., Holcik, M., 2006. Searching for IRES. *RNA* 12, 1755–1785.
- Baranick, B.T., Lemp, N.A., Nagashima, J., *et al.*, 2008. Splicing mediates the activity of four putative cellular internal ribosome entry sites. *Proceedings of the National Academy of Sciences of the United States of America* 105, 4733–4738.
- Berry, K.E., Waghray, S., Mortimer, S.A., Bai, Y., Doudna, J.A., 2011. Crystal structure of the HCV IRES central domain reveals strategy for start-codon positioning. *Structure* 19, 1456–1466.
- de Breynne, S., Yu, Y., Unbehauen, A., Pestova, T.V., Hellen, C.U., 2009. Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites. *Proceedings of the National Academy of Sciences of the United States of America* 106, 9197–9202.
- Brown, B.A., Ehrenfeld, E., 1979. Translation of poliovirus RNA *in vitro*: Changes in cleavage pattern and initiation sites by ribosomal salt wash. *Virology* 97, 396–405.
- Chen, C.Y., Sarnow, P., 1995. Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. *Science* 268, 415–417.
- Costantino, D.A., Pfingsten, J.S., Rambo, R.P., Kieft, J.S., 2008. tRNA-mRNA mimicry drives translation initiation from a viral IRES. *Nature Structural & Molecular Biology* 15, 57–64.
- Doner, A.J., Semler, B.L., Jackson, R.J., *et al.*, 1984. *In vitro* translation of poliovirus RNA: Utilization of internal initiation sites in reticulocyte lysate. *Journal of Virology* 50, 507–514.
- Fitzgerald, K.D., Semler, B.L., 2009. Bridging IRES elements in mRNAs to the eukaryotic translation apparatus. *Biochimica et Biophysica Acta* 1789, 518–528.
- Fraser, C.S., Doudna, J.A., 2007. Structural and mechanistic insights into hepatitis C viral translation initiation. *Nature Reviews Microbiology* 5, 29–38.
- Hashem, Y., des Georges, A., Dhote, V., *et al.*, 2013. Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature*. doi:10.1038/nature12658.
- Hellen, C.U., 2009. IRES-induced conformational changes in the ribosome and the mechanism of translation initiation by internal ribosomal entry. *Biochimica et Biophysica Acta* 1789, 558–570.
- Hellen, C.U., Sarnow, P., 2001. Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes & Development* 15, 1593–1612.
- Hinnebusch, A.G., 2014. The scanning mechanism of eukaryotic translation initiation. *Annual Review of Biochemistry*. doi:10.1146/annurev-biochem-060713-035802.
- Holcik, M., Sonenberg, N., 2005. Translational control in stress and apoptosis. *Nature Reviews* 6, 318–327.
- Honda, M., Beard, M.R., Ping, L.H., Lemon, S.M., 1999. A phylogenetically conserved stem-loop structure at the 5' border of the internal ribosome entry site of hepatitis C virus is required for cap-independent viral translation. *Journal of Virology* 73, 1165–1174.
- Jackson, R.J., 2013. The current status of vertebrate cellular mRNA IRESs. *Cold Spring Harbor Perspectives in Biology* 5, a011569.
- Jackson, R.J., Hellen, C.U., Pestova, T.V., 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nature Reviews* 11, 113–127.
- Jan, E., 2006. Divergent IRES elements in invertebrates. *Virus Research* 119, 16–28.
- Jang, S.K., Kräusslich, H.G., Nicklin, M.J., *et al.*, 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during *in vitro* translation. *Journal of Virology* 62, 2636–2643.
- King, H.A., Cobbald, L.C., Willis, A.E., 2010. The role of IRES trans-acting factors in regulating translation initiation. *Biochemical Society Transactions* 38, 1581–1586.
- Kolupaeva, V.G., Lomakin, I.B., Pestova, T.V., Hellen, C.U., 2003. Eukaryotic initiation factors 4G and 4A mediate conformational changes downstream of the initiation codon of the encephalomyocarditis virus internal ribosomal entry site. *Molecular and Cellular Biology* 23, 687–698.
- Komar, A.A., Hatzoglou, M., 2005. Internal ribosome entry sites in cellular mRNAs: Mystery of their existence. *Journal of Biological Chemistry* 280, 23425–23428.
- Kozak, M., 1978. How do eucaryotic ribosomes select initiation regions in messenger RNA? *Cell* 15, 1109–1123.
- Lemp, N.A., Hiraoka, K., Kasahara, N., Logg, C.R., 2012. Cryptic transcripts from a ubiquitous plasmid origin of replication confound tests for cis-regulatory function. *Nucleic Acids Research* 40, 7280–7290.
- Lewis, S.M., Holcik, M., 2008. For IRES trans-acting factors, it is all about location. *Oncogene* 27, 1033–1035.
- Lopez-Lastra, M., Rivas, A., Barria, M.I., 2005. Protein synthesis in eukaryotes: The growing biological relevance of cap-independent translation initiation. *Biological Research* 38, 121–146.
- Macejak, D.G., Sarnow, P., 1991. Internal initiation of translation mediated by the 5' leader of a cellular mRNA. *Nature* 353, 90–94.
- Meerovitch, K., Svitkin, Y.V., Lee, H.S., *et al.*, 1993. La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. *Journal of Virology* 67, 3798–3807.
- Minor, P.D., 1992. The molecular biology of poliovaccines. *Journal of General Virology* 73, 3065–3077.
- Mitchell, S.A., Spriggs, K.A., Coldwell, M.J., *et al.*, 2003. The Apaf-1 internal ribosome entry segment attains the correct structural conformation for function via interactions with PTB and unr. *Molecular Cell* 11, 757–771.
- Mokrejs, M., Masek, T., Vopalensky, V., *et al.*, 2010. IRESite — A tool for the examination of viral and cellular internal ribosome entry sites. *Nucleic Acids Research* 38, D131–D136.
- Muhs, M., Yamamoto, H., Ismer, J., *et al.*, 2011. Structural basis for the binding of IRES RNAs to the head of the ribosomal 40S subunit. *Nucleic Acids Research* 39, 5264–5275.
- Ochs, K., Zeller, A., Saleh, L., *et al.*, 2003. Impaired binding of standard initiation factors mediates poliovirus translation attenuation. *Journal of Virology* 77, 115–122.
- Pelletier, J., Sonenberg, N., 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* 334, 320–325.
- Pestova, T.V., de Breynne, S., Pisarev, A.V., Abaeva, I.S., Hellen, C.U., 2008. eIF2-dependent and eIF2-independent modes of initiation on the CSFV IRES: A common role of domain II. *EMBO Journal* 27, 1060–1072.
- Pestova, T.V., Hellen, C.U., 2003. Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. *Genes & Development* 17, 181–186.
- Pestova, T.V., Shatsky, I.N., Fletcher, S.P., Jackson, R.J., Hellen, C.U., 1998. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes & Development* 12, 67–83.
- Pestova, T.V., Shatsky, I.N., Hellen, C.U., 1996. Functional dissection of eukaryotic initiation factor 4F: The 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. *Molecular and Cellular Biology* 16, 6870–6878.
- Petz, M., Them, N., Huber, H., Beug, H., Mikulits, W., 2011. La enhances IRES-mediated translation of laminin B1 during malignant epithelial to mesenchymal transition. *Nucleic Acids Research* 40, 290–302.
- Pilipenko, E.V., Gmyl, A.P., Maslova, S.V., *et al.*, 1992. Prokaryotic-like cis elements in the cap-independent internal initiation of translation on picornavirus RNA. *Cell* 68, 119–131.
- Shatsky, I.N., Dmitriev, S.E., Terenin, I.M., Andreev, D.E., 2010. Cap- and IRES-independent scanning mechanism of translation initiation as an alternative to the concept of cellular IRESs. *Molecules and Cells* 30, 285–293.
- Silvera, D., Arju, R., Darvishian, F., *et al.*, 2009. Essential role for eIF4GI overexpression in the pathogenesis of inflammatory breast cancer. *Nature Cell Biology* 11, 903–908.
- Spahn, C.M., Kieft, J.S., Grassucci, R.A., *et al.*, 2001. Hepatitis C virus IRES RNA-induced changes in the conformation of the 40S ribosomal subunit. *Science* 291, 1959–1962.
- Svitkin, Y.V., Imataka, H., Khaleghpour, K., *et al.*, 2001. Poly(A)-binding protein interaction with eIF4G stimulates picornavirus IRES-dependent translation. *RNA* 7, 1743–1752.
- Svitkin, Y.V., Pestova, T.V., Maslova, S.V., Agol, V.I., 1988. Point mutations modify the response of poliovirus RNA to a translation initiation factor: A comparison of neurovirulent and attenuated strains. *Virology* 166, 394–404.
- Sweeney, T.R., Abaeva, I.S., Pestova, T.V., Hellen, C.U., 2013. The mechanism of translation initiation on Type 1 picornavirus IRESs. *EMBO Journal*. doi:10.1002/emboj.201386124.
- Terenin, I.M., Dmitriev, S.E., Andreev, D.E., Shatsky, I.N., 2008. Eukaryotic translation initiation machinery can operate in a bacterial-like mode without eIF2. *Nature Structural & Molecular Biology* 15, 836–841.

- Van Eden, M.E., Byrd, M.P., Sherrill, K.W., Lloyd, R.E., 2004. Demonstrating internal ribosome entry sites in eukaryotic mRNAs using stringent RNA test procedures. *RNA* 10, 720–730.
- Wilson, J.E., Pestova, T.V., Hellen, C.U., Sarnow, P., 2000. Initiation of protein synthesis from the A site of the ribosome. *Cell* 102, 511–520.
- Yueh, A., Schneider, R.J., 1996. Selective translation initiation by ribosome jumping in adenovirus-infected and heat-shocked cells. *Genes & Development* 10, 1557–1567.

Further Reading

- Agol, V.I., 2002. Picornavirus genome: An overview. In: Semler, B.L., Wimmer, E. (Eds.), *Molecular Biology of Picornaviruses*. Washington, DC: ASM Press, pp. 127–148.
- Balvay, L., Soto Rifo, R., Ricci, E.P., Decimo, D., Ohlmann, T., 2009. Structural and functional diversity of viral IRESes. *Biochimica et Biophysica Acta* 1789, 542–557.
- Doudna, J.A., Sarnow, P., 2007. Translation initiation by viral internal ribosome entry sites. In: Mathews, M.B., Sonenberg, N., Hershey, J.W.B. (Eds.), *Translational Control in Biology and Medicine*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 129–153.
- Ehrenfeld, E., Teterina, N.L., 2002. Initiation of translation of picornavirus RNAs: Structure and function of the internal ribosome entry site. In: Semler, B.L., Wimmer, E. (Eds.), *Molecular Biology of Picornaviruses*. Washington, DC: ASM Press, pp. 159–169.
- Elroy-Stein, O., Merrick, W.C., 2007. Translation initiation via cellular internal ribosome entry sites. In: Mathews, M.B., Sonenberg, N., Hershey, J.W.B. (Eds.), *Translational Control in Biology and Medicine*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press, pp. 155–172.
- Filbin, M.E., Kieft, J.S., 2009. Toward a structural understanding of IRES RNA function. *Current Opinion in Structural Biology* 19, 267–276.
- Gilbert, W.V., 2010. Alternative ways to think about cellular internal ribosome entry. *Journal of Biological Chemistry* 285, 29033–29038.
- Gingras, A.C., Raught, B., Sonenberg, N., 1999. eIF4 initiation factors: Effectors of mRNA recruitment to ribosomes and regulators of translation. *Annual Review of Biochemistry* 68, 913–963.
- Holcik, M., 2004. Targeting translation for treatment of cancer — A novel role for IRES? *Current Cancer Drug Targets* 4, 299–311.
- Komar, A.A., Hatzoglou, M., 2011. Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states. *Cell Cycle* 10, 229–240.
- Lopez-Lastra, M., Ramdohr, P., Letelier, A., *et al.*, 2010. Translation initiation of viral mRNAs. *Reviews in Medical Virology* 20, 177–195.
- Malys, N., McCarthy, J.E., 2011. Translation initiation: Variations in the mechanism can be anticipated. *Cellular and Molecular Life Sciences* 68, 991–1003.
- Pestova, T.V., Kolupaeva, V.G., Lomakin, I.B., *et al.*, 2001. Molecular mechanisms of translation initiation in eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America* 98, 7029–7036.
- Sonenberg, N., Hinnebusch, A.G., 2009. Regulation of translation initiation in eukaryotes: Mechanisms and biological targets. *Cell* 136, 731–745.
- Thompson, S.R., 2012. So you want to know if your message has an IRES? *Wiley Interdisciplinary Reviews: RNA* 3, 697–705.

Relevant Website

<http://www.iresite.org>
IRESite.

Targeted mRNA Degradation

S Djuranovic and HS Zaher, Washington University, St. Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The central dogma of molecular biology postulates the flow of genetic information from DNA to RNA to protein. This implies that gene expression in cells is a multistep process that involves transcription of genetic material from DNA to RNA molecules, followed by translation of messenger RNAs (mRNAs) into proteins. These processes define normal cell states and they are subject to stringent control at all levels – essentially any step of gene expression can be controlled. Generally, the amount of mRNA in a system is tightly regulated by altering its rate of transcription; however, post-transcriptional control mechanisms fine tune amounts of both mRNA and protein through modification, stability, and degradation of these molecules. Historically, gene regulation has been mainly studied from transcriptional regulation point of view; however, recent genome wide analyses of mRNA and protein abundance in mammalian cells (Rabani *et al.*, 2011; Schwanhauser *et al.*, 2011) revealed that posttranscriptional control plays an equally important part during gene expression regulation. While changes in transcription rates indeed determine the majority of temporal changes in mRNA levels, changes in mRNA degradation rates are found to be important for shaping the dynamic nature of gene regulation which is critical for cell proliferation, cell differentiation, stress, metabolism, immune response, and apoptosis (Ghosh *et al.*, 2010; Houseley and Tollervey, 2009; Schoenberg and Maquat, 2012; Wang *et al.*, 2002). Thus, it comes as no surprise that many cellular factors and mechanisms are employed to specifically regulate and modulate mRNA degradation rates.

mRNA stability in eukaryotic cells depends on the presence of a 5' 7-methylguanosine cap (m7G) and 3' poly(A) tail, which both serve as integral stability factors for each transcript. These structures also ensure efficient translation of mRNAs through interactions with translation initiation machinery and poly(A)-binding protein (PABP). Loss of these terminal structures from mRNAs through decapping or deadenylation results in their destabilization. mRNAs lacking a m7G cap are rapidly degraded by the exoribonuclease activity of Xrn1 in the 5'→3' direction. Transcripts lacking a poly(A) tail are eliminated by the 3'→5' exonucleolytic activity of a large protein complex – exosome. These processes serve as the major pathway by which mRNAs are degraded (Figure 1). We note that initial destabilization by decapping and deadenylation is subject to additional regulation by sequence-specific RNA-binding proteins (RBPs), short RNAs, and by RNA secondary structure, making the process of mRNA degradation specific for each individual transcript.

It is worth noting that generation of mRNA through the course of transcription and RNA processing is an imperfect process resulting in a substantial amount of defective transcripts. Furthermore, mRNA is subject to a number of chemical and environmental insults that alter its chemical properties. These mRNAs are subject to quality control processes and

interestingly their decay appears to be initiated through endonucleolytic cleavage. Since the targets of these processes cause mistranslation, it makes great sense that 'mRNA surveillance' depends on the ribosome for the initial recognition of the defect, where the processes are intimately coupled to translation (Shoemaker and Green, 2012).

mRNA Surveillance Mechanisms

As discussed earlier, aberrant mRNAs are generated through multiple pathways. The translation of these defective mRNAs is likely to result in aberrant proteins that are likely to misfold and hence cause havoc on the cell. It is not surprising then that cells evolved a number of quality control processes to recognize and degrade aberrant transcripts from the cellular mRNA pool. In eukaryotes, three cytoplasmic quality control processes have so far been identified (Graille and Seraphin, 2012; Kervestin and Jacobson, 2012; Shoemaker and Green, 2012). These are collectively referred to as 'mRNA surveillance' and include nonsense-mediated decay (NMD), no-go decay (NGD), and nonstop-decay (NSD) (Figure 2). In NMD, transcripts that contain a premature stop codon (PTC) that could result from mis-splicing, for example, are rapidly degraded. The process of NGD is responsible for degrading transcripts that cause stalling during translation. Finally, NSD acts on mRNAs lacking a stop codon where the ribosome runs to the 3'-end of the mRNA.

Substrate Choice

NMD

The process of NMD was initially documented in *Saccharomyces cerevisiae* more than three decades ago, where nonsense mutations in the *URA3* gene were found to have profound effects on the stability of its transcript (Losson and Lacroute, 1979). Soon after that, Maquat *et al.* (1981) showed that β -globin mRNA from β -thalassemia patients, which encodes for a truncated ORF, has a significantly shorter half-life relative to non-thalassemic mRNA. As a result of these initial observations and several others, NMD substrates were classically characterized as transcripts harboring PTCs, which can appear as a result of genomic mutations, transcriptional errors, defective DNA rearrangements, or alternative splicing. The latter process appears to generate many NMD substrates. In particular, it has been estimated that more than 75% of the human pre-mRNA is alternatively spliced and 45% of these produce one mRNA isoform that looks like an NMD target (Lewis *et al.*, 2003).

Recent high-throughput transcriptomic analysis, however, suggests that depending on the organism and conditions, 3–15% of the cellular mRNA pool is subject to NMD (Guan

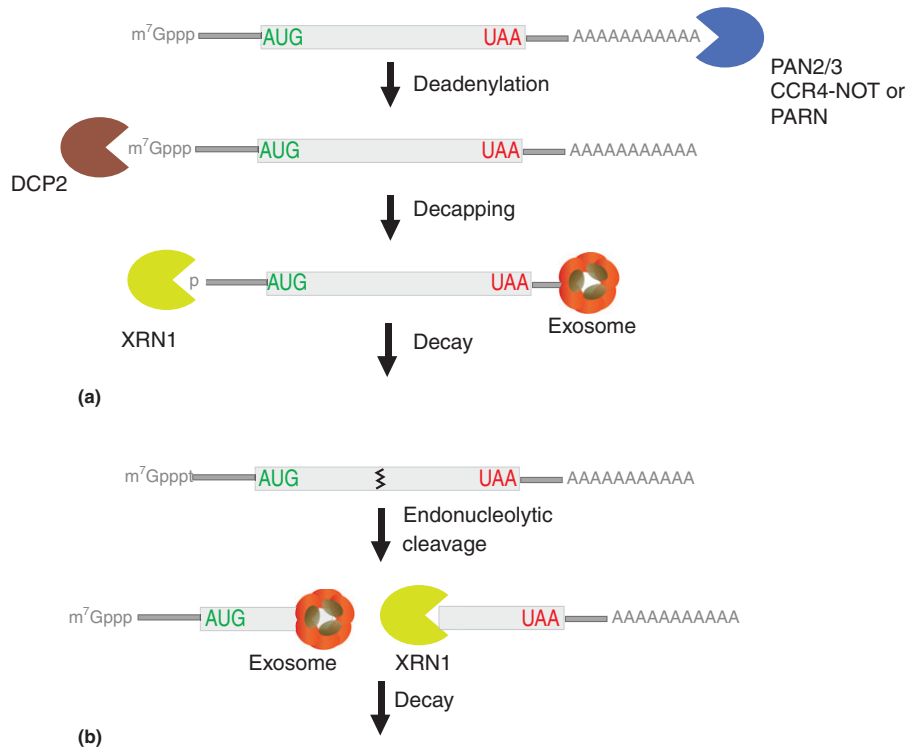


Figure 1 The main pathways of mRNA degradation in eukaryotes: (a) The canonical degradation pathway. The mRNA is initially deadenylated by the action of a deadenylase complex followed by decapping. The mRNA is then degraded through the action of the 5'–3' exonuclease Xrn1 and the exosome in the 3'–5' direction. (b) Degradation of mRNA through endonucleolytic cleavage. The cleavage creates substrates for the exosome, which degrades the 5' fragment, and Xrn1, which degrades the 3' fragment.

et al., 2006; He *et al.*, 2003; Lelivelt and Culbertson, 1999; Tani *et al.*, 2012). Furthermore, these NMD targets appear to vary from one cell type to another (Huang *et al.*, 2011; Yepiskoposyan *et al.*, 2011). These findings suggest that NMD, in addition to its role in clearing aberrant mRNA, plays a critical role in gene expression and regulation and appears to be important for a wide range of biological problems. Indeed, whereas NMD is dispensable in some lower eukaryotes such as *S. cerevisiae* and *Caenorhabditis elegans*, it is absolutely essential in mammals, zebra fish, and fruit fly. In mice, for example, the absence of NMD causes embryonic lethality (Medghalchi *et al.*, 2001).

These high-throughput studies and the observation that NMD is involved in a number of biological processes indicate that the classical definition of an NMD target as a PTC-harboring transcript is limited. Emerging from these studies is the discovery that many classes of transcripts can trigger NMD. Transcripts having upstream open reading frames (uORF) are logical targets, where a ribosome translating the uORF encounters that looks like a PTC. 3'-UTRs that have an intron also trigger NMD. Furthermore, long UTR, that are present under normal conditions or result from alternative polyadenylation, can trigger NMD (Chen *et al.*, 2006; Martins *et al.*, 2012), however, not all transcripts with long 3'-UTRs do. Collectively these studies suggest that sequence elements downstream of the stop codon, however, are responsible for triggering NMD. This unifying theme is often referred to as the 'faux 3'-UTR' model (Amrani *et al.*, 2004).

NGD

Similar to NMD, NGD was also initially observed in *S. cerevisiae*; Doma and Parker (2006) showed that mRNAs with strong secondary structures such as hairpins, pseudoknots as well as GC-rich sequences are turned over rapidly. These structures cause translating ribosomes to stall, 'not go.' While transcripts containing these roadblocks are the most effective NGD targets, the process appears to have other subtle targets that have features that can modulate the speed of the translating ribosome; these include transcripts harboring strings of rare codons or transcripts coding for peptides that, through interactions with the exit tunnel, stall the ribosome (Kuroha *et al.*, 2010; Letzring *et al.*, 2010). It is worth noting that although the use of mRNA reporters with stable secondary structures or long stretches of rare codons have contributed significantly to the dissection of the mechanism of NGD, such mRNAs for the most part do not exist in the transcriptome and hence they have been argued to have limited biological relevance. Recent data suggests that chemically damaged mRNAs, such as oxidized transcripts, are the real targets of NGD. Some of these adducts have been shown to stall the ribosome *in vitro* (Simms *et al.*, 2014).

NSD

As the name suggests, NSD evolved to degrade mRNAs lacking an in-frame stop codon (Frischmeyer *et al.*, 2002; van Hoof *et al.*, 2002). Formally such targets are of two

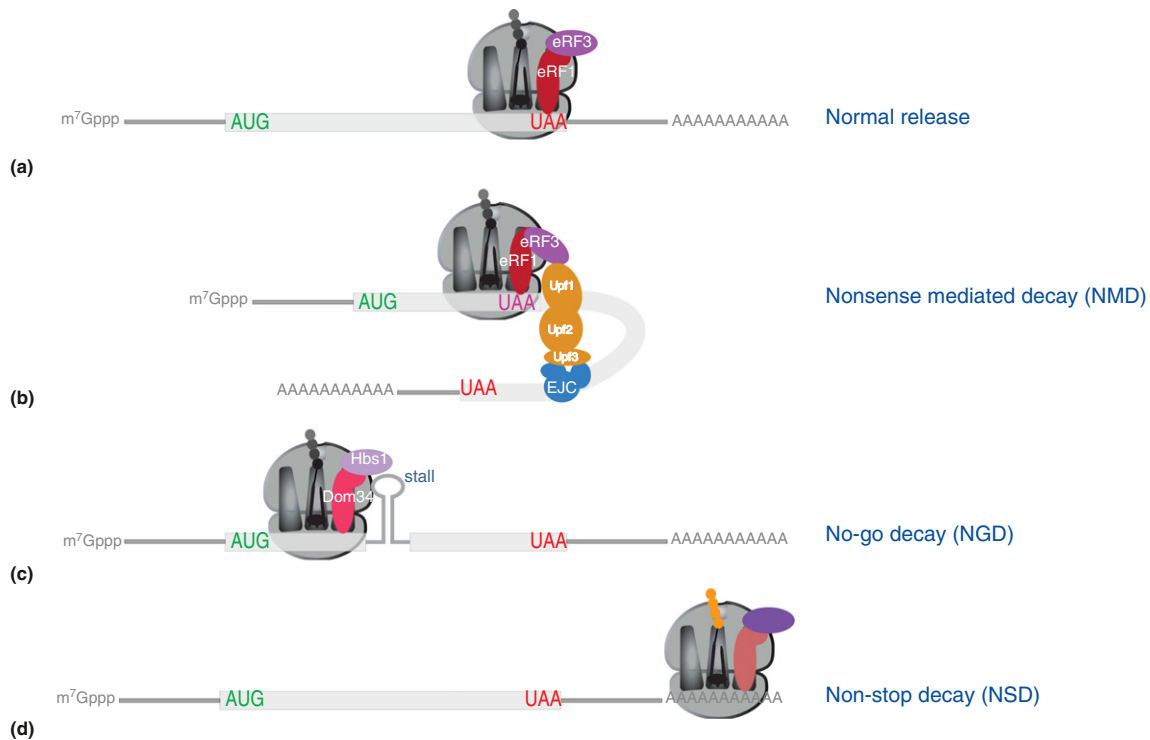


Figure 2 mRNA surveillance processes. (a) Canonical termination of protein synthesis. The stop codon is recognized by a ternary complex of eRF1, eRF3, and GTP, which promotes peptide release and initiates recycling of the ribosomal subunits. (b) The exon junction complex (EJC) model of nonsense-mediated decay (NMD). A premature stop codon (PTC) is recognized by eRF1 and eRF3, but due to the presence of an EJC downstream of the stop codon, several other interactions occur with the Upf proteins. These interactions appear to be required for the recruitment of several decay factors. (c) The process of no-go decay (NGD). The yeast eRF1, eRF3, homologues Dom34, and Hbs1 (respectively) recognize the stalled ribosome and bind in the A site. They then catalyze the dissociation of the ribosome and stimulate an endonucleolytic cleavage of the mRNA upstream of the ribosome (not shown). (d) PolyA-mediated nonstop-decay (NSD). In the absence of a stop codon, the ribosome runs to the polyA of the mRNA, synthesizing poly-Lys peptide and effectively stalling the ribosome. The process is similar to NGD in that Dom34 and Hbs1 are sometimes involved but in some organisms other factors are also involved (see text).

different types: the first is generated through truncation or endonucleolytic cleavage and as a result during translation the ribosome simply runs to the end of the message; the second is generated from mutation in the genomic sequence or an error during transcription resulting in mRNAs with no stop codon but with polyA tail. For the latter type, it has been speculated that the ribosome would translate the polyA tail synthesizing long poly-Lys peptide (Ito-Harashima *et al.*, 2007). The poly-Lys sequence accumulates an overall large positive charge, which is hypothesized to interact with the negative charge of the phosphodiester backbone of the ribosome exit tunnel and induce stalling. Hence, NSD and NGD are similar processes except for where the initial signal for the degradation along the mRNA originates; for NGD it is in the middle, whereas for NSD it is at the end. Indeed, as we shall see later the processes appear to share many of the factors in higher eukaryotes.

Substrate Recognition

NMD

As mentioned earlier, all NMD targets share the common feature of having a stop codon at a noncanonical position.

Due to the nature of the stop codon translation termination is speculated to be quite distinct from normal termination (Amrani *et al.*, 2004). The exact mechanism by which the ribosome distinguish between a normal stop codon and a PTC is not fully understood, but it involves three conserved proteins: Upf1, Upf2, and Upf3 in *S. cerevisiae* (Cui *et al.*, 1995; Leeds *et al.*, 1991, 1992). These proteins are thought to bridge a connection between downstream signals and the terminating ribosome (Chakrabarti *et al.*, 2011; Chamieh *et al.*, 2008). The connection between the Upf proteins and the ribosome is thought to be stimulatory for NMD, but inhibitory for termination and they compete with connections that are under normal conditions stimulatory for termination.

How do Upf proteins communicate signals to the ribosome? Early studies on NMD suggested that mRNAs harboring a PTC > ~50 nt upstream of exon-exon junction sites triggered a robust NMD response (Nagy and Maquat, 1998; Thermann *et al.*, 1998; Zhang *et al.*, 1998). Later studies showed that during splicing the region just upstream of the exon-exon junction sites is coated with a protein complex fittingly called the exon junction complex (EJC) (Le Hir *et al.*, 2000). During the pioneering round of translation, the EJC complexes are stripped off the mRNA by the ribosome. Given that most of the stop codons reside in the last exon of a gene,

the presence of an EJC complex downstream of a terminating ribosome provides a rational molecular signature for NMD. The core EJC complex is composed of three proteins, eIF4A3, MLN51, and a MAGOH/Y14 heterodimer, which interacts with the C-terminus of the NMD factor Upf3 (Andersen *et al.*, 2006; Bono *et al.*, 2006; Buchwald *et al.*, 2010). Upf3 has orthologues in all examined eukaryotes and consistent with its interaction with the EJC complex it shuttles between the nucleus and cytoplasm (Kim *et al.*, 2001). Upf2 acts as a bridge between Upf3 and Upf1, which in turn interacts with the release factors eRF1 and eRF3 as well as the phosphatidylinositol 3-kinase-related kinase SMG1 (Kashima *et al.*, 2006; Kunz *et al.*, 2006). Among NMD factors, Upf1 is the most conserved factor and is central to the process (Culbertson and Leeds, 2003). Sequence analysis as well as biochemical characterization of the protein revealed that it belongs to the ATP-dependent helicase superfamily (Cheng *et al.*, 2007; Czaplinski *et al.*, 1995). It binds RNA and ATP and uses the energy from ATP hydrolysis to unwind RNA in a 5' to 3' direction. In addition, the protein is subject to cycles of phosphorylation (by SMG1) and dephosphorylation on its unstructured C-terminus and N-terminus region (Grimson *et al.*, 2004; Okada-Katsuhata *et al.*, 2012; Yamashita *et al.*, 2001). On NMD substrates, Upf1–Upf2–Upf3 form a stable complex and it has been suggested that interaction between Upf2 and a cysteine/histidine (CH) rich domain on Upf1 activates the ATPase and helicase activity of Upf1 (Chamieh *et al.*, 2008). In the absence of Upf2, the CH domain of Upf1 interacts with the factor's helicase domain forming a closed conformation that inhibits the unwinding activity of Upf1. In addition to their interaction with the EJC complex, Upf proteins have been shown to interact with translation factors, namely release factors, for which they appear to modulate termination efficiency (Kashima *et al.*, 2006). Stop codon readthrough is significantly more efficient on PTCs relative to normal stop codons (Wang *et al.*, 2001).

What is clear from the many decades of studies on NMD is that the process is intimately coupled to translation termination. Termination of protein synthesis occurs when one of three nearly universal stop codons (UAG, UGA, and UAA) enter the A site of the ribosome. In eukaryotes, termination requires the two release factors eRF1 and eRF3. eRF1 is a tRNA mimic, which binds the A site of the ribosome and recognizes the stop codon. eRF3 is a GTPase and similar to EF1A (Frolova *et al.*, 1996), which binds aa-tRNAs to deliver them to the ribosome during elongation, is thought to form a ternary complex with eRF1 and GTP (Mitkevich *et al.*, 2006). Once on a terminating ribosome, GTP is hydrolyzed, eRF1 undergoes a conformational change allowing the conserved GGQ domain to engage the active site of the ribosome to initiate peptide release. It has also been suggested that eRF3 interacts with polyA-binding protein (Pab1), and that this interaction stimulates peptide release (Kononenko *et al.*, 2010). Following peptide release, the conserved recycling factor ABCE1 binds to eRF1 kicking eRF3 off; subunit dissociation ensues to complete the translation cycle (Pisarev *et al.*, 2010). Based on this relatively short description of termination, it should be apparent that the process is amenable to modulation through differential protein–protein interaction network. Indeed on NMD substrates, as detailed earlier, termination is inhibited due to

what has been suggested as direct competition between the Upf proteins and termination-stimulating factors (Ivanov *et al.*, 2008). In particular, Upf1 has been shown to interact with eRF3, but whether this interaction on its own is sufficient to elicit NMD is debated.

While the EJC model is elegant and appealing, it fails to explain all of the targets of NMD. An increasing number of NMD-targeted mRNAs do not have an intron downstream of a stop codon and as such are predicted not to harbor an EJC complex. Furthermore, NMD is robust in yeast, even though most of the pre-mRNAs are not spliced. It is worth noting that immunoprecipitation of eIF4AIII (a core component of the EJC) followed by deep sequencing suggests that the EJC can be deposited on mRNAs in a sequence-dependent and splicing-independent manner (Sauliere *et al.*, 2012; Singh *et al.*, 2012). So the possibility that the EJC is found on the UTR of mRNAs lacking an intron downstream of the stop codon cannot be ruled out. Furthermore, for all of these noncanonical targets, the mRNAs share the common feature of having a long UTR. The UTR model suggests that Upf1 binds to mRNAs non-specifically and during translation the ribosome actively displaces the factor (Hogg and Goff, 2010). Since the ribosome does not traverse the 3'-UTR, mRNAs with long UTR are expected to have a larger number of Upf1 proteins and hence are subject to NMD. Alternatively, it has been suggested that the polyA tail of mRNAs with long 3'-UTRs is distant from the stop codon, precluding any stimulatory effect PAB1 might have on termination. In contrast, others have argued that in the absence of a downstream EJC complex Upf2 and Upf3 interact with a ribosome-bound Upf1, which on its own is able to sense the efficiency of termination (Stalder and Muehle-mann, 2008). In summary, Upf1 appears to be the primary factor in determining what constitutes an NMD target whatever the initiating signal might be. Future work is needed to clarify the mechanism of the initial recognition of a PTC.

NGD

NGD is triggered by a stalled ribosome and requires the conserved protein Dom34 (or Pelota) and Hbs1. Interestingly, Dom34 and Hbs1 share structural features with eRF1 and eRF3, respectively (Atkinson *et al.*, 2008; Chen *et al.*, 2010). Consistent with these structural similarities, the factors interact with the A site of the ribosome (Becker *et al.*, 2011) but because Dom34 lacks the GGQ domain of eRF1, the complex does not promote peptide release. Instead, *in vitro* biochemical studies showed that the factors promote subunit dissociation (Shoemaker *et al.*, 2010). Moreover, the recycling factor ABCE1 (Rli1 in yeast) plays an important role during this reaction (Shoemaker and Green, 2011).

How does Dom34–Hbs1 complex recognize a stalled ribosome? How does it compete effectively with the ternary complex of EF1–aa-tRNA–GTP that normally binds the A site during elongation? Important clues into this process emerged from recent biochemical characterization showing an inverse correlation between the efficiency of subunit dissociation activity and the length of the mRNA downstream of the P site (Pisareva *et al.*, 2011; Shoemaker and Green, 2011). In particular, Dom34–Hbs1 appears to prefer ribosomal complexes

with little sequence downstream of the P site. Later structural studies showed that this length dependency is bestowed by Hbs1. The factor binds the ribosome in the mRNA entry tunnel and hence can only bind when the mRNA is short (Becker *et al.*, 2011).

NSD

As mentioned earlier, NSD is similar to NGD except that the former involves ribosomes stalling at the end of the mRNA. As a result, the processes proceed in a similar fashion utilizing the same factors. A notable exception is the requirement of Ski7 for NSD in some species of yeast (van Hoof, 2005). The factor is a translational GTPase related to eRF3 and Hbs1. However, in contrast to eRF3 and Hbs1, which have binding partners (eRF1 and Dom34, respectively), no binding partner for Ski7 has been identified. As a result, the mechanism of the recognition of NSD targets in yeast remains poorly understood. It is worth noting that Ski7 interacts with the exosome, thereby linking ribosome recognition of NSD targets to their ultimate degradation (van Hoof *et al.*, 2002).

Substrate Fate

The ultimate fate of the aberrant mRNA, regardless of the process (NMD, NGD, or NSD), is degradation. All of the pathways eventually utilize canonical decay pathways in both directions: 5'–3' Xrn1-mediated and 3'–5' exosome-mediated. Decay of metazoan NMD targets appears to initiate with an endonucleolytic cleavage catalyzed by Smg6 (Eberle *et al.*, 2009; Gatfield and Izaurralde, 2004). Following this cleavage, the 5' and 3' pieces are degraded by the exosome and Xrn1, respectively. Nonetheless, accumulating evidence suggests that alternative and potentially redundant mechanisms exist for clearing NMD targets that do not include endonucleolytic cleavage. For instance, yeast does not have a Smg6 homologue and no cleavage to date has been observed on NMD targets, yet decay is robust in this organism. Furthermore, tethering of the NMD factor Smg7 to mRNAs, in the absence of the endonuclease Smg6, results in fast degradation of the mRNA (Unterholzner and Izaurralde, 2004). Therefore, it has been suggested that decay of NMD targets can also proceed through exonucleolytic routes. The Smg6–Smg7 complex has been shown to recruit decay factors, mainly the deadenylation machinery, to accelerate degradation of NMD targets (Mitchell and Tollervey, 2003; Muhrlad and Parker, 1994).

NSD and NGD targets are subject to endonucleolytic cleavage. The cleavage has been shown to occur upstream of the stalled ribosome. For NSD, the catalytic subunit of the exosome rrp44 (Dis3), which has endonucleolytic and exonucleolytic activities, has been implicated in catalyzing the initial cleavage reaction (Schaeffer *et al.*, 2009; Schaeffer and van Hoof, 2011). For NGD, the endonuclease is yet to be identified, but Dom34–Hbs1 (although unnecessary for the cleavage activity) appears to stimulate the reaction (Tsuboi *et al.*, 2012). In summary, the three mRNA surveillance processes to a large extent exploit similar mechanisms to initiate the decay of their targets. There is a lot to learn about the

process of recognition of the targets and the mechanism of preparing them for degradation.

Targeted mRNA Degradation in Gene Regulation

In the first part we focused on the process of mRNA degradation due to interruptions in translation cycle that ultimately leads to activation of mRNA surveillance mechanisms. However, as we mentioned earlier, each mRNA transcript will have different set of bound RBPs and RNPs and differences in cellular presence of such *trans* regulatory elements will lead toward variation in mRNA degradation rates and differential gene regulation. The changes in mRNA degradation rates due to the targeted binding of *trans* elements as a subject of gene regulation control are described in nearly all cellular pathways from development to oncogenesis (Keene, 2007; Kishore *et al.*, 2010). As such eukaryotic mRNA degradation rates can vary remarkably from transcript to transcript as well as for the same transcript from cell to cell. Cellular modulators of mRNA degradation rates are mainly short RNAs and RBPs. The principle underlying the modulation of mRNA degradation rates by mRNPs involves controlled exposure of mRNA molecules to cellular RNA degradation machineries. This requires either removal of 'stabilizing' RBPs attached to mRNA transcripts and/or recruitment of 'destabilizing' RBPs and short RNAs. The vulnerable 'exposure' of mRNAs to short RNAs and RBPs dictate particular mRNA decay patterns of such mRNPs via direct or indirect interactions with the enzymes controlling mRNA degradation, mRNA deadenylation, or decapping factors (Figure 3).

mRNA Degradation Control by Short RNAs

microRNA (miRNAs) are short (about 21–25 nt long) non-coding RNAs that regulate the translation and degradation rates of mRNAs through nearly perfect complementary binding of target sequences in mRNA 3' untranslated regions (UTRs) (Bartel, 2004). While originally discovered in the nematode *C. elegans* (Lee *et al.*, 1993; Wightman *et al.*, 1993), they are now regarded as one of the key regulators of genes in most metazoans (Lagos-Quintana *et al.*, 2001; Lau *et al.*, 2001; Lee and Ambros, 2001). miRNAs are thought to regulate the expression of target genes by translational repression and targeted mRNA degradation (Bazzini *et al.*, 2012; Bethune *et al.*, 2012; Djuranovic *et al.*, 2012; Eulalio *et al.*, 2009; Guo *et al.*, 2010; Meijer *et al.*, 2013; Figure 3). These two processes are coupled in the living cell and whether the regulatory effects of miRNAs on target genes are the result of the direct induction of mRNA degradation or effects on mRNA decay are the result of an early block in translation initiation, consequentially leading to rapid mRNA degradation still remains elusive (Djuranovic *et al.*, 2011; Fabian *et al.*, 2010). In both cases miRNAs end up MukMuenhancing mRNA target degradation by directing mRNAs to the regular exonucleolytic mRNA decay involving initial deadenylation of the mRNA via the CCR4–POP2–NOT complex (Huntzinger and Izaurralde, 2011; Figure 3). Effects of miRNAs on mRNA target occur through ribonucleoprotein complexes named miRISCs, which at minimum include the

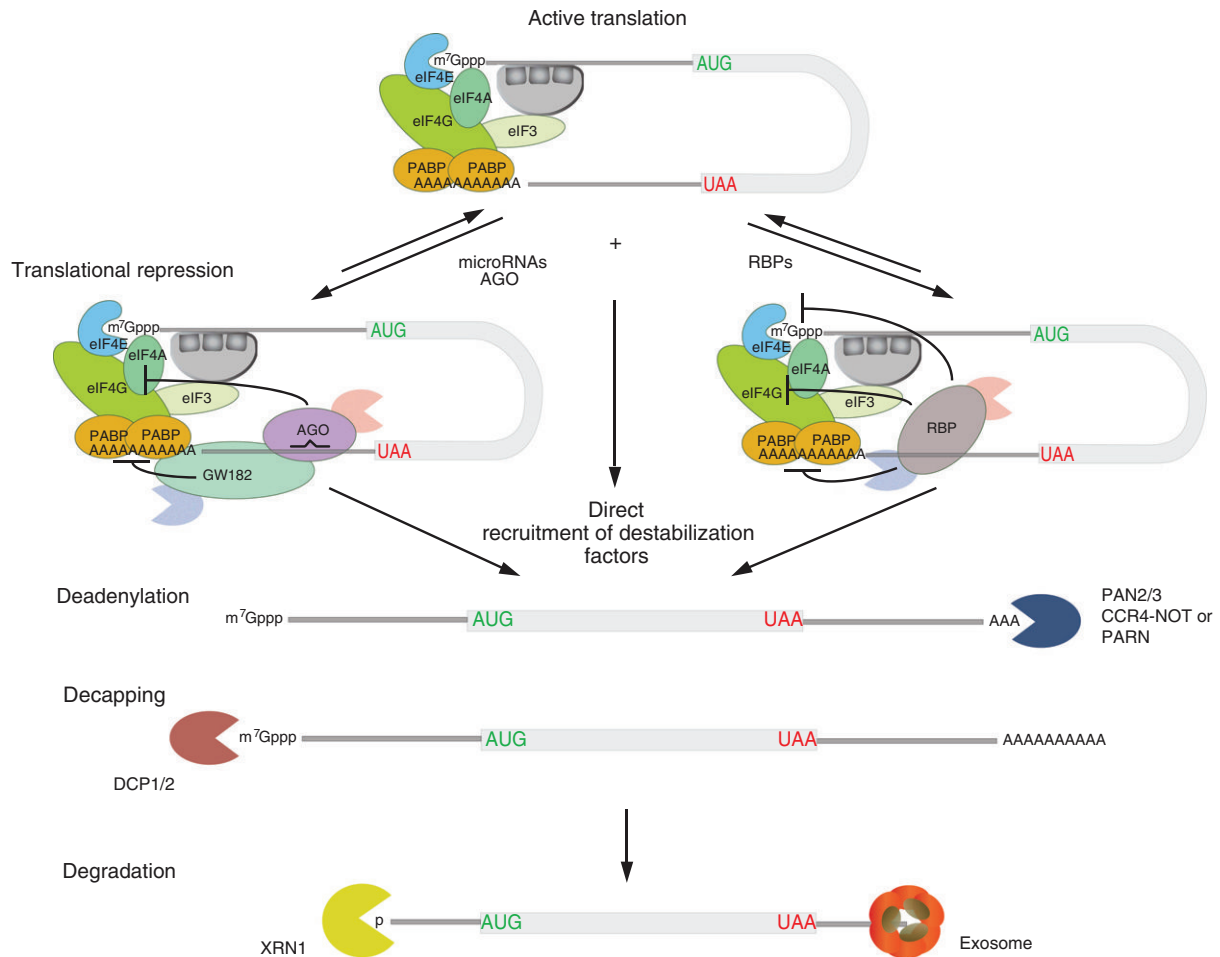


Figure 3 The basic mechanisms controlling mRNA degradation by microRNA (miRNAs) and RNA-binding proteins (RBPs). Recruitment of the mRNA regulators, such as miRNAs or RBPs, triggers either translational repression or direct destabilization of target mRNAs ultimately leading toward mRNA degradation. RBPs and miRISC (miRNA-AGO-GW182 complex) act directly on translation through interaction with translation initiation factors (eIF3, eIF4A, eIF4E, eIF4G) or PABP, preventing active translation of targeted mRNA and exposing it to mRNA degradation machinery. In alternative pathway, RNA regulators directly recruit mRNA destabilization factors, involved in decapping and deadenylation. The removal of the protective structures from either 5' (m⁷Gppp-cap) or 3'-terminus (poly(A)-tail) of target mRNA results in RNA degradation by Xrn1 (5'–3' decay) and the exosome (3'–5' decay).

miRNA, Argonaute (AGO), and GW182 protein (Wilson and Doudna, 2013). The binding of such minimal miRISC complex is sufficient for translational repression which in turn might be sufficient to make mRNA targets 'vulnerable' to mRNA degradation machinery (Djuranovic et al., 2011). Additionally it was also shown that direct interactions of AGO or GW182 proteins with both deadenylation and decapping factors generate more substantial effects on mRNA degradation (Braun et al., 2011). An interesting fact is that targeting of mRNAs by miRNAs often leads to localization of target mRNAs in cytoplasmic foci termed P-bodies which can act as a site for both storage and degradation of target mRNAs (Parker and Sheth, 2007). It is plausible to think that the complexity of miRNA biogenesis (Ha and Kim, 2014), mechanisms of interaction between miRNAs and target mRNAs (Bartel, 2009) as well as existence of different miRISC complexes (Wu et al., 2013) offer numerous possibilities for modulation of miRNA mediated gene regulation and targeted mRNA degradation.

More recently it was shown that piRNAs, another class of short RNAs, may control mRNA degradation patterns in multiple organisms (Gou et al., 2014; Kiuchi et al., 2014; Rouget et al., 2010; Watanabe et al., 2015). piRNAs are known for their function in the repression of transposable elements primarily within the germ line cells, however their function might be also important in regulation of gene expression (Thomson and Lin, 2009). piRNAs in early *Drosophila* embryos affect nanos mRNA deadenylation and degradation through the interplay with both Smaug, an RNA-binding protein, and already mentioned CCR4-POP2-NOT deadenylation complex (Rouget et al., 2010). In a similar manner, interactions of piRNA-silencing complex (piRISC), containing piRNAs and murine PIWI protein (MIWI), and CCR4-CAF1-NOT deadenylase complex in mouse testes controls expression of approximately 5000 genes through targeted mRNA destabilization through CAF1-dependent deadenylation (Gou et al., 2014).

Targeted mRNA Degradation by RBPs

The main feature of RBPs is the presence of at least one RNA-binding domain in their protein sequence (Lunde *et al.*, 2007). One of the best characterized families of RBPs linked to mRNA degradation is the Hu, ELAV proteins (Brennan and Steitz, 2001; Simone and Keene, 2013). This group of proteins specifically interacts with 3' UTRs of mRNA transcripts containing adenylate/uridylylate-rich elements (AREs), *cis*-regulatory sequences linked to mRNA degradation control (Shaw and Kamen, 1986). The addition of ARE elements derived from the human colony stimulating factor 2 (CSF2) gene to the 3' UTR sequences of a stable reporter mRNA, such as β -globin, results in an unstable chimeric transcript. Such experiments demonstrated the importance of these sequence elements in the stability of mRNAs in the cell (Chen and Shyu, 1995; Shaw and Kamen, 1986). The importance of Hu/ELAV proteins and their biological significance is especially highlighted by their function in neurons. Pathological conditions affecting the level of functional Hu/ELAV proteins lead to severe brain damage pathologies that are quite often connected to degenerative neurological disorders following paraneoplastic syndromes (Darnell, 2013; Pascale *et al.*, 2008). Bioinformatic analyses estimated that 5–10% of the human transcriptome contain AREs and appear to be essential in regulating apoptosis, immune response, and intracellular signaling to name a few (Halees *et al.*, 2008). The exact mechanism by which Hu/Elav and other ARE binding proteins lead to a change in mRNA stability are not yet fully resolved. In one particular example, HuR positively regulates the stability of the mRNA transcript of the human eukaryotic translation initiation factor 4E (eIF4E). The process involves direct competition between HuR and yet another ARE binding factor – AUF1 (Topisirovic *et al.*, 2009). AUF1 has been implicated in RNA decay regulation through its binding to eukaryotic initiation factor 4G (eIF4G) and PABP ultimately leading to mRNA destabilization (Lu *et al.*, 2006). Mechanistically, molecular interaction between AUF1, eIF4G, and PABP may lead to the interruption in translation cycle or/and displacement of PABP from poly(A) tail of AUF1 bound transcripts allowing the deadenylase complex access to the mRNA (Sagliocco *et al.*, 2006). Additionally, AUF1 may facilitate degradation of targeted mRNAs by direct recruiting of exosomes to ARE–mRNA complexes (Chen *et al.*, 2001; Torrisani *et al.*, 2007). At a molecular level, by binding to AREs, AUF1 or HuR are likely to alter some local RNA structure, which in turn provides necessary surface area for additional RBPs or RNPs and hence alters mRNA translation and stability (Wilson *et al.*, 2003; Zucconi *et al.*, 2010). Given that majority (57%) of HuR targets are also bound by AUF1, antagonizing effects and competitive binding to AREs between these two RBPs might be relevant for the stability of many mRNA targets (Lal *et al.*, 2004). In addition to already mentioned HuR/ELAV and AUF1, several other RBPs are also known to interact with AU-rich elements. These among others include – tristetraprolin (TTP) (Lai *et al.*, 1999), butyrate response factor-1 (BRF1) (Stoecklin *et al.*, 2002) and KH-type splicing regulatory protein (KSRP) (Gherzi *et al.*, 2004). The KH domains of KSRP are essential for their degradation activity via interaction with deadenylase PARN, decapping factor Dcp2, and exosome subunit Rrp4 (Chou *et al.*, 2006; Gherzi *et al.*, 2004). TTP and

BRF1 interact specifically with exosomes, the deadenylase complex (involving Ccr4), decapping enzymes, and Xrn1 to promote mRNA degradation in P-bodies through a sequential process (Fenger-Gron *et al.*, 2005; Franks and Lykke-Andersen, 2007; Lykke-Andersen and Wagner, 2005).

Other RBP such as Smaug are also important in controlling translation and degradation of targeted mRNAs. Smaug belongs to a family of conserved sterile alpha motif domain containing RBP and all members of this family are thought to be involved in posttranscriptional gene regulation (Aviv *et al.*, 2003). In *Drosophila*, Smaug plays an important role in post-transcriptional regulation of the RNA-binding protein Nanos (Nos) (Chen *et al.*, 2014). Smaug carries out this function by binding Smaug recognition elements (SREs) in the Nos 3' UTR, as well as in other 3' UTRs. This feature is seen across many RBPs since the specificity for target interaction depends on both primary sequence and secondary structure features (Smibert *et al.*, 1996). Beside the effect on translation inhibition, Smaug was later shown to control the stability of targeted mRNAs by recruiting the deadenylase complex CCR4–POP2–NOT (Semotok *et al.*, 2005).

In addition to those mentioned above, several other *cis*-regulatory elements have been related to RNA degradation control. These include U-rich, CU-rich, GC-rich, poly(C), CA-rich, and GU-rich (GRE) elements (Hamilton *et al.*, 2008; Rattenbacher *et al.*, 2010; Stoecklin *et al.*, 2002; You *et al.*, 1992), which are regulated by a wide range of RBPs. Moreover, with a recent delineation of an 'atlas of mammalian mRNA binding proteins' (Castello *et al.*, 2012) which identified more than 300 newly described RBPs and emerging variety of *cis*-regulatory mRNA degradation elements the complex regulatory landscape controlling the fate of mRNA molecules in the cell is still far from being completely explored.

Combined Action of Short RNAs and RBPs

Last but not least, it is plausible that dynamic combination of miRNAs and RBPs that are associated to most mRNAs modulates these two types of RNA regulatory interactions and that the outcome of such interactions affects the mRNA degradation patterns in yet completely novel fashion (Figure 3). It was shown that binding of HuR protein to the 3' UTR of the CAT1 mRNA in human cells attenuates the repressive effects of miR-122 on CAT1 (Bhattacharyya *et al.*, 2006). These experiments were further confirmed using reporter mRNA targets as well as *in vitro* experiments (Kundu *et al.*, 2012) showing that binding of HuR to mRNA transcripts targeted by miRNAs ultimately leads to the dissociation of miRISC from mRNA target. Furthermore, HuR association with miRNA targets can also inhibit miRNA-induced mRNA deadenylation and degradation. The possible mechanistic explanation for antagonizing actions of HuR protein on miRISC targets came from more recent study on another ARE binding protein HuD (Fukao *et al.*, 2014). While miRISC promoted efficient release of eukaryotic initiation factor 4A (eIF4A) from target mRNAs (Fukao *et al.*, 2014) causing translational repression, HuD protein counteracted this miRISC action by stabilizing binding of eIF4A protein to target mRNAs ensuring efficient translation of such transcripts. In some other cases, RBPs can actually

enhance miRNA regulation. For instance, in mammalian cells, the Pumilio RBPs Pum1 and Pum2 promote the regulatory effects of miR-221/miR-222 on the p27 mRNA (Kedde *et al.*, 2010). In regulation of p27 by Pum1 and miR-221/222, Pum1 binds to a stem-loop containing Pumilio recognition element (PRE) and the miR-221/222 target site relaxing the RNA secondary structure and allowing miRNA binding. These interactions are significant, given that p27 downregulation by miR-221/222 is essential for cell proliferation and may also have a central role in the development of cancer (Triboulet and Gregory, 2010). In the case of E2F mRNA it was demonstrated that removal of PREs or depletion of Pum1/2 prevented miRNA regulation of the E2F mRNA. It is currently unknown how Pum1/2 in conjunction with miRNAs regulates E2F transcript translation and stability (Miles *et al.*, 2012). It is possible that binding of Pum1/2 induces RNA structural rearrangements, however these changes would need to affect three different miRNA binding sites over a several hundred base-pair long distance. At the end, these examples are here illustrating the importance of the physical interactions between mRNA, RBPs, and miRISC to achieve efficient gene regulation through the modulation of RNA degradation rates.

See also: Nucleic Acid Synthesis/Breakdown: RNA Synthesis/Function: Messenger RNA (mRNA): The Link between DNA and Protein

References

- Amrani, N., Ganesan, R., Kervestin, S., *et al.*, 2004. A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* 432, 112–118.
- Andersen, C.B.F., Ballut, L., Johansen, J.S., *et al.*, 2006. Structure of the exon junction core complex with a trapped DEAD-box ATPase bound to RNA. *Science* 313, 1968–1972.
- Atkinson, G.C., Baldauf, S.L., Hauryliuk, V., 2008. Evolution of nonstop, no-go and nonsense-mediated mRNA decay and their termination factor-derived components. *BMC Evolutionary Biology* 8, 290.
- Aviv, T., Lin, Z., Lau, S., *et al.*, 2003. The RNA-binding SAM domain of Smaug defines a new family of post-transcriptional regulators. *Nature Structural Biology* 10, 614–621.
- Bartel, D.P., 2004. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* 116, 281–297.
- Bartel, D.P., 2009. MicroRNAs: Target recognition and regulatory functions. *Cell* 136, 215–233.
- Bazzini, A.A., Lee, M.T., Giraldez, A.J., 2012. Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. *Science* 336, 233–237.
- Becker, T., Armache, J.-P., Jarasch, A., *et al.*, 2011. Structure of the no-go mRNA decay complex Dom34-Hbs1 bound to a stalled 80S ribosome. *Nature Structural & Molecular Biology* 18, 715–720.
- Bethune, J., Artus-Revel, C.G., Filipowicz, W., 2012. Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells. *EMBO Reports* 13, 716–723.
- Bhattacharyya, S.N., Habermacher, R., Martine, U., Closs, E.I., Filipowicz, W., 2006. Relief of microRNA-mediated translational repression in human cells subjected to stress. *Cell* 125, 1111–1124.
- Bono, F., Ebert, J., Lorentzen, E., Conti, E., 2006. The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell* 126, 713–725.
- Braun, J.E., Huntzinger, E., Fauser, M., Izaurralde, E., 2011. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Molecular Cell* 44, 120–133.
- Brennan, C.M., Steitz, J.A., 2001. HuR and mRNA stability. *Cellular and Molecular Life Sciences: CMLS* 58, 266–277.
- Buchwald, G., Ebert, J., Basquin, C., *et al.*, 2010. Insights into the recruitment of the NMD machinery from the crystal structure of a core EJC-UPF3b complex. *Proceedings of the National Academy of Sciences of the United States of America* 107, 10050–10055.
- Castello, A., Fischer, B., Eichelbaum, K., *et al.*, 2012. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393–1406.
- Chakrabarti, S., Jayachandran, U., Bonneau, F., *et al.*, 2011. Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Molecular Cell* 41, 693–703.
- Charnieh, H., Ballut, L., Bonneau, F., Le Hir, H., 2008. NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nature Structural & Molecular Biology* 15, 85–93.
- Chen, C.Y., Gherzi, R., Ong, S.E., *et al.*, 2001. AU binding proteins recruit the exosome to degrade ARE-containing mRNAs. *Cell* 107, 451–464.
- Chen, C.Y.A., Shyu, A.B., 1995. AU-rich elements – characterization and importance in messenger-rna degradation. *Trends in Biochemical Sciences* 20, 465–470.
- Chen, J.-M., Ferec, C., Cooper, D.N., 2006. A systematic analysis of disease-associated variants in the 3' regulatory regions of human protein-coding genes I: General principles and overview. *Human Genetics* 120, 1–21.
- Chen, L., Dumelie, J.G., Li, X., *et al.*, 2014. Global regulation of mRNA translation and stability in the early *Drosophila* embryo by the Smaug RNA-binding protein. *Genome Biology* 15, R4.
- Chen, L., Muhrad, D., Hauryliuk, V., *et al.*, 2010. Structure of the Dom34-Hbs1 complex and implications for no-go decay. *Nature Structural & Molecular Biology* 17, 1233–1240.
- Cheng, Z., Muhrad, D., Lim, M.K., Parker, R., Song, H., 2007. Structural and functional insights into the human Upf1 helicase core. *EMBO Journal* 26, 253–264.
- Chou, C.-F., Mulky, A., Maitra, S., *et al.*, 2006. Tethering KSRP, a decay-promoting AU-rich element-binding protein, to mRNAs elicits mRNA decay. *Molecular and Cellular Biology* 26, 3695–3706.
- Cui, Y., Hagan, K.W., Zhang, S.A., Peltz, S.W., 1995. Identification and characterization of genes that are required for the accelerated degradation of messenger-RNAs containing a premature translational termination codon. *Genes & Development* 9, 423–436.
- Culbertson, M.R., Leeds, P.F., 2003. Looking at mRNA decay pathways through the window of molecular evolution. *Current Opinion in Genetics & Development* 13, 207–214.
- Czapinski, K., Weng, Y., Hagan, K.W., Peltz, S.W., 1995. Purification and characterization of the Upf1 protein – a factor involved in translation and messenger-RNA degradation. *RNA* 1, 610–623.
- Darnell, R.B., 2013. RNA protein interaction in neurons. *Annual Review of Neuroscience* 36, 243–270.
- Djuranovic, S., Nahvi, A., Green, R., 2011. A parsimonious model for gene regulation by miRNAs. *Science* 331, 550–553.
- Djuranovic, S., Nahvi, A., Green, R., 2012. miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. *Science* 336, 237–240.
- Doma, M.K., Parker, R., 2006. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* 440, 561–564.
- Eberle, A.B., Lykke-Andersen, S., Muhlemann, O., Jensen, T.H., 2009. SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nature Structural & Molecular Biology* 16, 49–55.
- Eulalio, A., Huntzinger, E., Nishihara, T., *et al.*, 2009. Deadenylation is a widespread effect of miRNA regulation. *RNA* 15, 21–32.
- Fabian, M.R., Sonenberg, N., Filipowicz, W., 2010. Regulation of mRNA translation and stability by microRNAs. In: Kornberg, R.D., Raetz, C.R.H., Rothman, J.E., Thorner, J.W. (Eds.), *Annual Review of Biochemistry*, vol. 79. Palo Alto, CA: Annual Reviews, pp. 351–379.
- Fenger-Gron, M., Fillman, C., Norrild, B., Lykke-Andersen, J., 2005. Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. *Molecular Cell* 20, 905–915.
- Franks, T.M., Lykke-Andersen, J., 2007. TTP and BRF proteins nucleate processing body formation to silence mRNAs with AU-rich elements. *Genes & Development* 21, 719–735.
- Frishmeyer, P.A., van Hoof, A., O'Donnell, K., *et al.*, 2002. An mRNA surveillance mechanism that eliminates transcripts lacking termination codons. *Science* 295, 2258–2261.
- Frolova, L., LeGoff, X., Zhouravleva, G., *et al.*, 1996. Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase. *RNA* 2, 334–341.

- Fukao, A., Mishima, Y., Takizawa, N., *et al.*, 2014. MicroRNAs trigger dissociation of eIF4A1 and eIF4AII from target mRNAs in humans. *Molecular Cell* 56, 79–89.
- Garfield, D., Izaurralde, E., 2004. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in *Drosophila*. *Nature* 429, 575–578.
- Gherzi, R., Lee, K.Y., Briata, P., *et al.*, 2004. A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. *Molecular Cell* 14, 571–583.
- Ghosh, S., Ganesan, R., Amrani, N., Jacobson, A., 2010. Translational competence of ribosomes released from a premature termination codon is modulated by NMD factors. *RNA—A Publication of the RNA Society* 16, 1832–1847.
- Gou, L.T., Dai, P., Yang, J.H., *et al.*, 2014. Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Research* 24, 680–700.
- Grailhe, M., Seraphin, B., 2012. Surveillance pathways rescuing eukaryotic ribosomes lost in translation. *Nature Reviews Molecular Cell Biology* 13, 727–735.
- Grimson, A., O'Connor, S., Newman, C.L., Anderson, P., 2004. SMG-1 is a phosphatidylinositol kinase-related protein kinase required for nonsense-mediated mRNA decay in *Caenorhabditis elegans*. *Molecular and Cellular Biology* 24, 7483–7490.
- Guan, Q., Zheng, W., Tang, S., *et al.*, 2006. Impact of nonsense-mediated mRNA decay on the global expression profile of budding yeast. *PLoS Genetics* 2, 1924–1943.
- Guo, H., Ingolia, N.T., Weissman, J.S., Bartel, D.P., 2010. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 466, 835–840.
- Ha, M., Kim, V.N., 2014. Regulation of microRNA biogenesis. *Nature Reviews Molecular Cell Biology* 15, 509–524.
- Halees, A.S., El-Badrawi, R., Khabar, K.S.A., 2008. ARED organism: Expansion of ARED reveals AU-rich element cluster variations between human and mouse. *Nucleic Acids Research* 36, D137–D140.
- Hamilton, B.J., Wang, X.W., Collins, J., *et al.*, 2008. Separate cis-trans pathways post-transcriptionally regulate murine CD154 (CD40 ligand) expression: A novel function for CA repeats in the 3'-untranslated region. *Journal of Biological Chemistry* 283, 25606–25616.
- He, F., Li, X.R., Spatrick, P., *et al.*, 2003. Genome-wide analysis of mRNAs regulated by the nonsense-mediated and 5' to 3' mRNA decay pathways in yeast. *Molecular Cell* 12, 1439–1452.
- Hogg, J.R., Goff, S.P., 2010. Upf1 senses 3' UTR length to potentiate mRNA decay. *Cell* 143, 379–389.
- van Hoof, A., 2005. Conserved functions of yeast genes support the duplication, degeneration and complementation model for gene duplication. *Genetics* 171, 1455–1461.
- van Hoof, A., Frischmeyer, P.A., Dietz, H.C., Parker, R., 2002. Exosome-mediated recognition and degradation of mRNAs lacking a termination codon. *Science* 295, 2262–2264.
- Houssley, J., Tollervey, D., 2009. The many pathways of RNA degradation. *Cell* 136, 763–776.
- Huang, L., Lou, C.-H., Chan, W., *et al.*, 2011. RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. *Molecular Cell* 43, 950–961.
- Huntzinger, E., Izaurralde, E., 2011. Gene silencing by microRNAs: Contributions of translational repression and mRNA decay. *Nature Reviews Genetics* 12, 99–110.
- Ito-Harashima, S., Kuroha, K., Tatematsu, T., Inada, T., 2007. Translation of the poly (A) tail plays crucial roles in nonstop mRNA surveillance via translation repression and protein destabilization by proteasome in yeast. *Genes & Development* 21, 519–524.
- Ivanov, P.V., Gehring, N.H., Kunz, J.B., Hentze, M.W., Kulozik, A.E., 2008. Interactions between UPF1, eRFs, PABP and the exon junction complex suggest an integrated model for mammalian NMD pathways. *EMBO Journal* 27, 736–747.
- Kashima, I., Yamashita, A., Izumi, N., *et al.*, 2006. Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes & Development* 20, 355–367.
- Kedde, M., van Kouwenhove, M., Zwart, W., *et al.*, 2010. A Pumilio-induced RNA structure switch in p27-3' UTR controls miR-221 and miR-222 accessibility. *Nature Cell Biology* 12, 1014–1020.
- Keene, J.D., 2007. RNA regulons: Coordination of post-transcriptional events. *Nature Reviews Genetics* 8, 533–543.
- Kervestin, S., Jacobson, A., 2012. NMD: A multifaceted response to premature translational termination. *Nature Reviews Molecular Cell Biology* 13, 700–712.
- Kim, V.N., Kataoka, N., Dreyfuss, G., 2001. Role of the nonsense-mediated decay factor hUpf3 in the splicing-dependent exon-exon junction complex. *Science* 293, 1832–1836.
- Kishore, S., Lubner, S., Zavolan, M., 2010. Deciphering the role of RNA-binding proteins in the post-transcriptional control of gene expression. *Briefings in Functional Genomics* 9, 391–404.
- Kiuchi, T., Koga, H., Kawamoto, M., *et al.*, 2014. A single female-specific piRNA is the primary determiner of sex in the silkworm. *Nature* 509, 633–636.
- Kononenko, A.V., Mitkevich, V.A., Atkinson, G.C., *et al.*, 2010. GTP-dependent structural rearrangement of the eRF1:eRF3 complex and eRF3 sequence motifs essential for PABP binding. *Nucleic Acids Research* 38, 548–558.
- Kundu, P., Fabian, M.R., Sonenberg, N., Bhattacharyya, S.N., Filipowicz, W., 2012. HuR protein attenuates miRNA-mediated repression by promoting miRISC dissociation from the target RNA. *Nucleic Acids Research* 40, 5088–5100.
- Kunz, J.B., Neu-Yilik, G., Hentze, M.W., Kulozik, A.E., Gehring, N.H., 2006. Functions of hUpf3a and hUpf3b in nonsense-mediated mRNA decay and translation. *RNA—A Publication of the RNA Society* 12, 1015–1022.
- Kuroha, K., Akamatsu, M., Dimitrova, L., *et al.*, 2010. Receptor for activated C kinase 1 stimulates nascent polypeptide-dependent translation arrest. *EMBO Reports* 11, 956–961.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W., Tuschl, T., 2001. Identification of novel genes coding for small expressed RNAs. *Science* 294, 853–858.
- Lai, A., Carballo, E., Strum, J.R., *et al.*, 1999. Evidence that tristetraprolin binds to AU-rich elements and promotes the deadenylation and destabilization of tumor necrosis factor alpha mRNA. *Molecular and Cellular Biology* 19 (6), 4311–4323.
- Lai, A., Mazan-Mamczarz, K., Kawai, T., *et al.*, 2004. Concurrent versus individual binding of HuR and AUF1 to common labile target mRNAs. *EMBO Journal* 23, 3092–3102.
- Lau, N.C., Lim, L.P., Weinstein, E.G., Bartel, D.P., 2001. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858–862.
- Le Hir, H., Izaurralde, E., Maquat, L.E., Moore, M.J., 2000. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon–exon junctions. *EMBO Journal* 19, 6860–6869.
- Lee, R.C., Ambros, V., 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294, 862–864.
- Lee, R.C., Feinbaum, R.L., Ambros, V., 1993. The *C.elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75, 843–854.
- Leeds, P., Peltz, S.W., Jacobson, A., Culbertson, M.R., 1991. The product of the yeast UPF1 gene is required for rapid turnover of messenger-RNAs containing a premature translational termination codon. *Genes & Development* 5, 2303–2314.
- Leeds, P., Wood, J.M., Lee, B.S., Culbertson, M.R., 1992. Gene-products that promote messenger-RNA turnover in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* 12, 2165–2177.
- Lelivelt, M.J., Culbertson, M.R., 1999. Yeast Upf proteins required for RNA surveillance affect global expression of the yeast transcriptome. *Molecular and Cellular Biology* 19, 6710–6719.
- Letzring, D.P., Dean, K.M., Grayhack, E.J., 2010. Control of translation efficiency in yeast by codon-anticodon interactions. *RNA—A Publication of the RNA Society* 16, 2516–2528.
- Lewis, B.P., Green, R.E., Brenner, S.E., 2003. Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proceedings of the National Academy of Sciences of the United States of America* 100, 189–192.
- Losson, R., Lacroute, F., 1979. Interference of nonsense mutations with eukaryotic messenger-rna stability. *Proceedings of the National Academy of Sciences of the United States of America* 76, 5134–5137.
- Lu, J.Y., Bergman, N., Sadri, N., Schneider, R.J., 2006. Assembly of AUF1 with eIF4G-poly(A) binding protein complex suggests a translation function in AU-rich mRNA decay. *RNA—A Publication of the RNA Society* 12, 883–893.
- Lunde, B.M., Moore, C., Varani, G., 2007. RNA-binding proteins: Modular design for efficient function. *Nature Reviews Molecular Cell Biology* 8, 479–490.
- Lykke-Andersen, J., Wagner, E., 2005. Recruitment and activation of mRNA decay enzymes by two ARE-mediated decay activation domains in the proteins TTP and BRF-1. *Genes & Development* 19, 351–361.
- Maquat, L.E., Kinniburgh, A.J., Rachmilewitz, E.A., Ross, J., 1981. Unstable beta-globin messenger-RNA in messenger RNA-deficient beta-o thalassemia. *Cell* 27, 543–553.
- Martins, R., Proenca, D., Silva, B., *et al.*, 2012. Alternative polyadenylation and nonsense-mediated decay coordinately regulate the human HFE mRNA levels. *PLoS One* 7, e35461.
- Medghalchi, S.M., Frischmeyer, P.A., Mendell, J.T., *et al.*, 2001. Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Human Molecular Genetics* 10, 99–105.

- Meijer, H.A., Kong, Y.W., Lu, W.T., *et al.*, 2013. Translational repression and eIF4A2 activity are critical for microRNA-mediated gene regulation. *Science* 340, 82–85.
- Miles, W.O., Tschop, K., Herr, A., Ji, J.Y., Dyson, N.J., 2012. Pumi1 facilitates miRNA regulation of the E2F3 oncogene. *Genes & Development* 26, 356–368.
- Mitchell, P., Tollervey, D., 2003. An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3' → 5' degradation. *Molecular Cell* 11, 1405–1413.
- Mitkevich, V.A., Kononenko, A.V., Petrushanko, I.Y., *et al.*, 2006. Termination of translation in eukaryotes is mediated by the quaternary eRF1 center dot eRF3 center dot GTP center dot Mg²⁺ complex. The biological roles of eRF3 and prokaryotic RF3 are profoundly distinct. *Nucleic Acids Research* 34, 3947–3954.
- Muhlrad, D., Parker, R., 1994. Premature translational termination triggers messenger-RNA decapping. *Nature* 370, 578–581.
- Nagy, E., Maquat, L.E., 1998. A rule for termination-codon position within intron-containing genes: When nonsense affects RNA abundance. *Trends in Biochemical Sciences* 23, 198–199.
- Okada-Katsuhata, Y., Yamashita, A., Kutsuzawa, K., *et al.*, 2012. N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. *Nucleic Acids Research* 40, 1251–1266.
- Parker, R., Sheth, U., 2007. P bodies and the control of mRNA translation and degradation. *Molecular Cell* 25, 635–646.
- Pascale, A., Amadio, M., Quattrone, A., 2008. Defining a neuron: Neuronal ELAV proteins. *Cellular and Molecular Life Sciences: CMLS* 65, 128–140.
- Pisarev, A.V., Skabkin, M.A., Pisareva, V.P., *et al.*, 2010. The role of ABCE1 in eukaryotic posttermination ribosomal recycling. *Molecular Cell* 37, 196–210.
- Pisareva, V.P., Skabkin, M.A., Hellen, C.U.T., Pestova, T.V., Pisarev, A.V., 2011. Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. *EMBO Journal* 30, 1804–1817.
- Rabani, M., Levin, J.Z., Fan, L., *et al.*, 2011. Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nature Biotechnology* 29, 436–442.
- Rattenbacher, B., Beisang, D., Wiesner, D.L., *et al.*, 2010. Analysis of CUGBP1 targets identifies GU-repeat sequences that mediate rapid mRNA decay. *Molecular and Cellular Biology* 30, 3970–3980.
- Rouget, C., Papin, C., Boureux, A., *et al.*, 2010. Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo. *Nature* 467, 1128–1132.
- Sagliocco, F., Laloo, B., Cosson, B., *et al.*, 2006. The ARE-associated factor AUH binds poly(A) *in vitro* in competition with PABP. *Biochemical Journal* 400, 337–347.
- Sauliere, J., Murigneux, V., Wang, Z., *et al.*, 2012. CLIP-seq of eIF4AIII reveals transcriptome-wide mapping of the human exon junction complex. *Nature Structural & Molecular Biology* 19, 1124–1131.
- Schaeffer, D., van Hoof, A., 2011. Different nuclease requirements for exosome-mediated degradation of normal and nonstop mRNAs. *Proceedings of the National Academy of Sciences of the United States of America* 108, 2366–2371.
- Schaeffer, D., Tsanova, B., Barbas, A., *et al.*, 2009. The exosome contains domains with specific endoribonuclease, exoribonuclease and cytoplasmic mRNA decay activities. *Nature Structural & Molecular Biology* 16, 56–62.
- Schoenberg, D.R., Maquat, L.E., 2012. Regulation of cytoplasmic mRNA decay. *Nature Reviews: Genetics* 13, 246–259.
- Schwanhauser, B., Busse, D., Li, N., *et al.*, 2011. Global quantification of mammalian gene expression control. *Nature* 473, 337–342.
- Semotok, J.L., Cooperstock, R.L., Pinder, B.D., *et al.*, 2005. Smaug recruits the CCR4/POP2/NOT deadenylase complex to trigger maternal transcript localization in the early *Drosophila* embryo. *Current Biology: CB* 15, 284–294.
- Shaw, G., Kamen, R., 1986. A conserved AU sequence from the 3' untranslated region of GM-CSF messenger-RNA mediates selective messenger-RNA degradation. *Cell* 46, 659–667.
- Shoemaker, C.J., Eyer, D.E., Green, R., 2010. Dom34:Hbs1 promotes subunit dissociation and peptidyl-tRNA drop-off to initiate no-go decay. *Science* 330, 369–372.
- Shoemaker, C.J., Green, R., 2011. Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. *Proceedings of the National Academy of Sciences of the United States of America* 108, E1392–E1398.
- Shoemaker, C.J., Green, R., 2012. Translation drives mRNA quality control. *Nature Structural & Molecular Biology* 19, 594–601.
- Simms, C.L., Hudson, B.H., Mosior, J.W., Rangwala, A.S., Zaher, H.S., 2014. An active role for the ribosome in determining the fate of oxidized mRNA. *Cell Reports* 9, 1256–1264.
- Simone, L.E., Keene, J.D., 2013. Mechanisms coordinating ELAV/Hu mRNA regulons. *Current Opinion in Genetics & Development* 23, 35–43.
- Singh, G., Kucukural, A., Cenik, C., *et al.*, 2012. The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* 151, 750–764.
- Smibert, C.A., Wilson, J.E., Kerr, K., Macdonald, P.M., 1996. Smaug protein represses translation of unlocalized nanos mRNA in the *Drosophila* embryo. *Genes & Development* 10, 2600–2609.
- Stalder, L., Muehlemann, O., 2008. The meaning of nonsense. *Trends in Cell Biology* 18, 315–321.
- Stoecklin, G., Colombi, M., Raineri, I., *et al.*, 2002. Functional cloning of BRF1, a regulator of ARE-dependent mRNA turnover. *EMBO Journal* 21, 4709–4718.
- Tani, H., Imamachi, N., Salam, K.A., *et al.*, 2012. Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. *RNA Biology* 9, 1370–1379.
- Thermann, R., Neu-Yilik, G., Deters, A., *et al.*, 1998. Binary specification of nonsense codons by splicing and cytoplasmic translation. *EMBO Journal* 17, 3484–3494.
- Thomson, T., Lin, H., 2009. The biogenesis and function of PIWI proteins and piRNAs: Progress and prospect. *Annual Review of Cell and Developmental Biology* 25, 355–376.
- Topisirovic, I., Siddiqui, N., Orolicki, S., *et al.*, 2009. Stability of eukaryotic translation initiation factor 4E mRNA is regulated by HuR, and this activity is dysregulated in cancer. *Molecular and Cellular Biology* 29, 1152–1162.
- Torrisani, J., Unterberger, A., Tendulkar, S.R., Shikimi, K., Szyf, M., 2007. AUF1 cell cycle variations define genomic DNA methylation by regulation of DNMT1 mRNA stability. *Molecular and Cellular Biology* 27, 395–410.
- Triboulet, R., Gregory, R.I., 2010. Pumi1 turns on microRNA function. *Nature Cell Biology* 12, 928–929.
- Tsuboi, T., Kuroha, K., Kudo, K., *et al.*, 2012. Dom34:Hbs1 plays a general role in quality-control systems by dissociation of a stalled ribosome at the 3' end of aberrant mRNA. *Molecular Cell* 46, 518–529.
- Unterholzer, L., Izaurralde, E., 2004. SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Molecular Cell* 16, 587–596.
- Wang, W.R., Czaplinski, K., Rao, Y., Peltz, S.W., 2001. The role of Upf proteins in modulating the translation read-through of nonsense-containing transcripts. *EMBO Journal* 20, 880–890.
- Wang, Y., Liu, C.L., Storey, J.D., *et al.*, 2002. Precision and functional specificity in mRNA decay. *Proceedings of the National Academy of Sciences of the United States of America* 99, 5860–5865.
- Watanabe, T., Cheng, E.C., Zhong, M., Lin, H., 2015. Retrotransposons and pseudogenes regulate mRNAs and lncRNAs via the piRNA pathway in the germline. *Genome Research* 25 (3), 368–380.
- Wightman, B., Ha, I., Ruvkun, G., 1993. Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855–862.
- Wilson, G.M., Lu, J., Sutphen, K., *et al.*, 2003. Phosphorylation of p40AUF1 regulates binding to A + U-rich mRNA-destabilizing elements and protein-induced changes in ribonucleoprotein structure. *Journal of Biological Chemistry* 278, 33039–33048.
- Wilson, R.C., Doudna, J.A., 2013. Molecular mechanisms of RNA interference. *Annual Review of Biophysics* 42, 217–239.
- Wu, P.H., Isaji, M., Carthew, R.W., 2013. Functionally diverse microRNA effector complexes are regulated by extracellular signaling. *Molecular Cell* 52, 113–123.
- Yamashita, A., Ohnishi, T., Kashima, I., Taya, Y., Ohno, S., 2001. Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. *Genes & Development* 15, 2215–2228.
- Yepiskoposyan, H., Aeschmann, F., Nilsson, D., Okoniewski, M., Muehlemann, O., 2011. Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. *RNA-A Publication of the RNA Society* 17, 2108–2118.
- You, Y., Chen, C.Y., Shyu, A.B., 1992. U-rich sequence-binding proteins (URBPs) interacting with a 20-nucleotide U-rich sequence in the 3' untranslated region of c-fos mRNA may be involved in the first step of c-fos mRNA degradation. *Molecular and Cellular Biology* 12, 2931–2940.
- Zhang, J., Sun, X.L., Qian, Y.M., LaDuca, J.P., Maquat, L.E., 1998. At least one intron is required for the nonsense-mediated decay of triosephosphate isomerase mRNA: A possible link between nuclear splicing and cytoplasmic translation. *Molecular and Cellular Biology* 18, 5272–5283.
- Zucconi, B.E., Ballin, J.D., Brewer, B.Y., *et al.*, 2010. Alternatively expressed domains of AU-rich element RNA-binding protein 1 (AUF1) regulate RNA-binding affinity, RNA-induced protein oligomerization, and the local conformation of bound RNA ligands. *Journal of Biological Chemistry* 285, 39127–39139.

CELL DIVISION/DEATH: APOPTOSIS

Contents

Apoptosis

Anoikis

Mitochondria in Cell Death Regulation

Tumor Necrosis Factor Signaling Pathways

Caspases

Efferocytosis in the Tumor Microenvironment

The Bcl-2 Family Proteins: Insights into Their Mechanism of Action and Therapeutic Potential

Inhibitor of Apoptosis Proteins, the Sentinels of Cell Death and Signaling

Mitotic Catastrophe

Autophagy

Apoptosis

MF Navarro and G Salvesen, Sanford-Burnham Medical Research Institute, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Apoptosis, Death as Part of Life

Ancient Egyptians regarded death as an extension of life and not the end of it. In preparation for the 'underworld journey' elaborate funeral ceremonies were planned carefully and considerable resources had to be spent in elements such as the construction of the tomb and more importantly, the coffin. Construction of coffins was commissioned to dedicated craftsmen while the final touch of decoration with texts and paints was the duty of scribes and painters. Likewise, multicellular organisms have evolved intricate mechanisms of programmed cell death and within their own ceremony take the cell for its final underworld journey. In light of their irreversible character, the onset and propagation of cell death programs are tightly regulated by a convoluted circuitry of molecules with limited protein hydrolysis, phosphorylation, and ubiquitination as the main biochemical regulatory mechanisms (Dix *et al.*, 2012; Peng *et al.*, 2011). Once the cell commits to execute a death mechanism progressive morphological changes reduce the entire cell architecture down to small fragments called apoptotic bodies. Currently four major mechanisms of cell death are described in the literature – apoptosis, necrosis, autophagy, and pyroptosis (Fink and Cookson, 2005; Maiuri *et al.*, 2007; Duprez *et al.*, 2009). The most intensively researched mechanism, apoptosis, has seen significant development in the last 20 years and thanks to this we now have greater understanding of its biomedical and biomedical implications.

Control of cell number in multicellular organism is essential, the human body is composed of approximately 10^{14}

cells and if mitosis were to proceed with no cell death, a person 80-year-old would have a gut of 16 km long and 2 ton of bone marrow and lymph nodes (Melino, 2001; Fischer and Schulze-Osthoff, 2005). In the roundworm *Caenorhabditis elegans* an adult hermaphrodite generates 1090 cells and always the same 131 cells undergo apoptosis during development (Lettre and Hengartner, 2006). Given these numbers we must understand apoptosis as a physiological process vital for organismal development, regulation of the immune system, tissue homeostasis, and elimination of damaged or infected cells (Johnson *et al.*, 2000; Portt *et al.*, 2011). In contrast to this, dysregulated apoptosis is strongly implicated in pathological conditions where the balance of cells is altered, such as cancer, AIDS, Parkinson's disease, ischemia, and heart failure (Elmore, 2007; Favaloro *et al.*, 2012).

The Extrinsic Pathway

Activation of apoptosis may occur via the extrinsic or intrinsic pathway (mitochondrial) depending of the initial stimuli (Stennicke and Salvesen, 2000); both of these pathways rely on the activity of proteases known as caspases and the subsequent cleavage of their specific substrates (Portt *et al.*, 2011; Fuentes-Prior and Salvesen, 2004; Figure 1). The extrinsic pathway is mediated by a group of transmembrane death receptors that belong to the tumor necrosis factor receptor-1 (TNFR1) superfamily in which we find TNFR1 itself, CD95/APO-1, TRAIL receptors, and other death receptors (DR) (Chang *et al.*, 2003). After binding to their cognate trimeric ligand these receptors recruit initiator caspase-8 to a

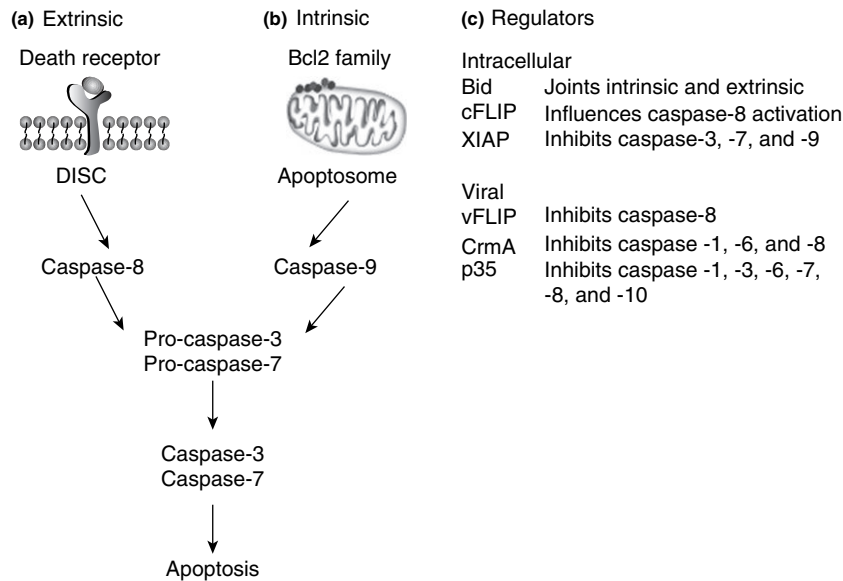


Figure 1 Apoptotic pathway. (a) Activation of the extrinsic pathway is initiated when a death receptor is bound by a death ligand triggering trimerization and recruitment of elements that form the DISC. Recruitment of pro-caspase-8 to the DISC induces its activation via dimerization, in turn caspase-8 cleaves and activates pro-caspase-3 and -7. (b) In the intrinsic pathway intracellular stimuli promote translocation of pro-apoptotic proteins to the mitochondria causing cytochrome c release and apoptosome assembly. The assembled apoptosome activates pro-caspase-9, which in turn can also activate caspase-3 and -7. Ultimately apoptosis execution via the extrinsic or intrinsic pathway dismantles the cell architecture producing small apoptotic bodies cleared phagocytes. (c) List of some intracellular and viral regulators of apoptosis discussed throughout the text.

multiprotein complex, the death-inducing signaling complex (DISC), that also includes the adaptor FAS-associated death domain (FADD) or TNFR-associated death domain (TRADD) (Boatright *et al.*, 2004; McIlwain *et al.*, 2013). Once activated at the DISC caspase-8 cleaves and activates zymogens of the executioner caspases -3, -6, -7 to directly induce death, or the BH3-only protein BID (BH3-interacting domain death agonist) to initiate the intrinsic pathway (McIlwain *et al.*, 2013). The functional significance of caspase-8 and the other components of the DISC in the context of the extrinsic pathway is highlighted in experimental settings and in cancer. Cancerous cells and cells deficient in some of these elements are resistant to apoptosis mediated by the TNFR family of receptors (Chang *et al.*, 2003; van Raam *et al.*, 2013; Eggert *et al.*, 2001).

The Intrinsic Pathway

The intrinsic pathway is fundamentally regulated and centered around the mitochondria (Figure 1). Here elements of the pro-apoptotic and anti-apoptotic machinery that reside either in the cytoplasm or in the mitochondria itself interact to resolve the cell fate (Wang and Youle, 2009). The stimuli required for activation of the intrinsic pathway are far more diverse than those required for activation of the extrinsic pathway and do not rely on receptor/ligand interactions. Some of the triggers include DNA damage, endoplasmic reticulum stress, serum starvation, chemotherapeutic agents, and UV radiation (Tait and Green, 2010; Wang and Youle, 2009). Although the precise details involved in the execution of the intrinsic pathway depend on the specific initial stimuli the following general

model is widely accepted. Prior to caspase activation (the formal initiation of apoptosis) the release of cytochrome c from the mitochondrial intermembrane space drives the assembly of a macromolecular platform termed the apoptosome (Tait and Green, 2010). Upon release from mitochondria cytochrome c interacts with the apoptotic protease-activating factor-1 (Apaf-1) to drive ATP-dependent oligomerization of the two proteins into a wheel-shaped ring (Riedl and Salvesen, 2007; Taylor *et al.*, 2008). In turn oligomerization of Apaf-1 and cytochrome c serve as a scaffold to recruit caspase-9 monomers to induce its dimerization, resulting in a catalytically competent enzyme. Once activated caspase-9 cleaves and activates caspase-3 and caspase-7 to initiate the apoptotic cascade (Riedl and Salvesen, 2007).

Notably, cross talk between the extrinsic and the intrinsic pathways is possible since both share some common elements. Once activated during execution of the extrinsic pathway, caspase-8 cleaves BID into (cBID) that promotes mitochondrial cytochrome c release, assembly of the apoptosome and activation of caspase-9. The subsequent step, as just described, is activation of caspase-3 and -7 by caspase-9 (Tait and Green, 2010). Whether apoptosis is initiated through the intrinsic or the extrinsic pathway, massive morphological changes reduce the cell architecture down to small fragments that in mammals can be removed by phagocytes to avoid undesired immune response: early cell rounding, plasma membrane blebbing, flipping of phospholipids in the cell membrane, internucleosomal DNA cleavage and degradation, nuclear condensation, detachment from the extracellular matrix, and fragmentation into small apoptotic bodies (Taylor *et al.*, 2008).

Of Life and Apoptosis: The Anti-Apoptotic versus Pro-Apoptotic Machinery

In the introduction we describe the apoptotic pathway assuming a free fall to cellular demise but several lines of evidence show that this scenario is further regulated by the existence of positive and negative effectors of apoptosis that control the onset and progression of the pathway. Generally, this may be achieved directly by targeting caspases or indirectly by antagonizing the negative effect of the caspase inhibitors. Next we describe three protein families that constitute the anti-apoptotic and pro-apoptotic machinery: caspases, the B-cell lymphoma-2 (Bcl-2) family, and inhibitors of apoptosis proteins (IAPs). Within the cell at any given time interactions between members of these three families are finely balanced and disruption of this balance may be enough to establish whether the cell enters into a proliferative state or the apoptotic pathway.

Pro-Apoptotic Proteins

Caspases

The central molecular element, and defining characteristic, of apoptosis are the caspases, a family of cysteine proteases that by means of limited proteolysis deliver and execute apoptotic stimuli (Salvesen and Riedl, 2008). The word caspase derives from cysteine-dependent aspartate-specific protease, which illustrates the remarkable specificity of these proteases for aspartic acid in position P1 of the substrate and variable specificity in positions P2, P3, and P4 (Pop and Salvesen, 2009). The catalytic triad that typifies cysteine proteases, Cys-His-X, is unique among caspases as the third residue and is not conserved in all the members of this family. In caspase-8 the third residue is Arg, in caspase-1 is Pro, and Thr in caspase-3, leading to the possibility that only a Cys-His dyad is required for proton shuttling during catalysis (Stennicke and Salvesen, 1999; Watt *et al.*, 1999). The number of caspases varies across organisms, in *C. elegans* there is one, in *Drosophila* there are four, while in mice and humans this number increases to ten and eleven (Yi and Yuan, 2009). However, the original hypothesis that the expansion in caspase number in mice and humans may be related to the role of these proteases out of the cell death framework is falsified when one considers that Cnidaria – the most primitive multicellular animals, have almost as many caspases as humans do (Zmasek *et al.*, 2007). Not all caspases perform an apoptotic function, indeed the first described member of the family that were later to be named caspases is interleukin 1 converting enzyme (ICE) (Cerretti *et al.*, 1992; Morris *et al.*, 1992). In mammals a general classification based on the physiological role distinguishes between the pro-inflammatory and apoptotic subgroups (Salvesen and Ashkenazi, 2011). Pro-inflammatory caspases include caspase-1, -4, and -5 in humans and caspase-1 and -11 in mice. The group of apoptotic caspases is further subdivided into two subgroups: initiators (-2, -8, -9, and -10) and executioners (-3, -6, and -7), importantly caspase-10 is absent in mice and caspase-11 is absent in humans (Pop and Salvesen, 2009; Tyas *et al.*, 2000; McIlwain *et al.*, 2013; Siegel, 2006). A separate classification has been pondered for

caspase-14 since this is neither pro-inflammatory nor apoptotic, instead this caspase participates in cornification, a phase during keratinocyte differentiation. Expression of caspase-14 is restricted to the skin (Van de Craen *et al.*, 1998). How and why caspases have evolved to program such apparently opposing functions as inflammation, skin cornification, and apoptosis is a mystery.

Initiator (apical) caspases and executioner caspases do not possess the same domain organization and this leads to substantial distinctions in the mechanisms of activation and regulation (Fuentes-Prior and Salvesen, 2004). The proximity-induced model introduced in 2003 (Boatright *et al.*, 2003) explains activation of apical caspases. This model postulates that apical caspases exist as monomeric zymogens that are recruited to macromolecular 'activation platforms' for their activation. Zymogens of caspase-8 are recruited to the DISC through their death effector domain (DEDs) and caspase-9 zymogens are recruited to the apoptosome through the caspase activation and recruitment domain (CARD) (Boatright *et al.*, 2003). Once caspase-8 is in complex with the DISC and caspase-9 is in complex with the apoptosome, activation of zymogen monomers to productive enzymes occurs by means of dimerization followed by cleavage at specific aspartate residues. Although caspase activation is frequently associated with cleavage, this event is not essential for the enzymes to attain full catalytic potential (Boatright *et al.*, 2004; Pop *et al.*, 2006).

Executioner caspases-3, -6, and -7 reside in the cytoplasm as inactive dimers where they are processed and activated by executioner caspases-8, -9, and -10. Cleavage of executioner caspases occurs at specific aspartate residues in the inter-subunit linker between the small and large domain. This single cleavage step induces rearrangement of mobile loops that allow formation and stabilization of the productive catalytic state (Pop and Salvesen, 2009).

Of the three executioner caspases, caspase-6 is the least studied and the most studied are caspase-3 and -7. During the biochemical characterization of caspase-3 and -7 the extensive use of synthetic peptides suggests that activity of these caspases is very similar and they may perform redundant roles. This perspective turned out to be reductionist but it was expanded latter after applying experimental approaches that included the combined use of cell-free extracts and immunodepletion assays. Results from these experiments showed that redundancy in the proteins cleaved by each caspase was rather limited as only a few substrates were cleaved by both caspase-3 and caspase-7. It was observed that caspase-3 processed a larger range of substrates implying that this caspase is the most far-reaching during the demolition phase of apoptosis when cell morphology is reduced down to debris (Slee *et al.*, 2001; Walsh *et al.*, 2008).

Caspases can be considered the Rosetta stone in the design of the apoptotic pathway; preincubation of cells with the pan-caspase inhibitor z-VAD-FMK prior to apoptosis stimulation prevents execution of the pathway (Karpnich *et al.*, 2006). Still, a growing number of publications relate caspases with diverse roles apart from those known in cell death. Caspase-1 and -11 regulate inflammation (Degtarev *et al.*, 2003; Kayagaki *et al.*, 2011), caspase-3 is necessary for erythroblast differentiation, and caspase-8 participates in macrophage differentiation (Yi and Yuan, 2009). Paradoxically,

caspase-8 also has a survival role, mice deficient in caspase-8 die during embryonic development (Varfolomeev *et al.*, 1998). Lethality in caspase-8-deficient mice has been attributed to the complex of two proteins that promote uncontrolled necrosis during development: receptor interacting kinase protein-1 and -3 (RIPK-1 and RIPK-3). Integration of biochemical and genetic data concluded that caspase-8 does not only promote apoptosis but is also a negative regulator RIP-1/RIPK-3 (Kaiser *et al.*, 2011). A valid question at this point would be how does a protein well established as a death promoter turn the switch to cell survival? The answer comes from a different molecular entity formed between caspase-8 and the cellular FLICE-like inhibitor protein (FLIP). Association of caspase-8 and FLIP to form a Csp8/FLIP heterodimer produces a protein with distinct enzymatic activity; this heterodimer displays unique kinetic properties and dissimilar substrate specificity as compared with the caspase-8 homodimer (Pop *et al.*, 2011). Biologically the caspase-8/FLIP dimer is essential for suppression of RIPK1/RIPK3-mediated necrosis during embryonic development and in immune cell proliferation (Oberst *et al.*, 2011).

Bcl-2 family

The B-cell lymphoma-2 (Bcl-2) family of proteins comprises three subfamilies associated mainly with controlling mitochondrial apoptosis: the anti-apoptotic, the pro-apoptotic, and the BH3-only proteins. The members of this family are evolutionary related and share some sequence or structural homology with Bcl-2, the founding member of the family (Taylor *et al.*, 2008). Bcl-2 members of the pro-apoptotic and the BH3-only groups may interact synergistically to promote apoptosis although a mechanistic description to explain these observations is far from complete. An appropriate example of this occurs with two pro-apoptotic proteins, BAX and BAK, and the BH3 only protein, BID. Upon cleavage by caspase-8, cBID drives translocation of BAK and/or BAX to the mitochondrial outer membrane where they oligomerize to form pores that induce mitochondrial outer membrane permeabilization (MOMP) (Youle and Strasser, 2008; Moldoveanu *et al.*, 2013). The indispensable role of Bak/Bax is exemplified in cells deficient in both of these factors as these are resistant to multiple apoptotic stimuli such as chemicals (staurosporine, etoposide, thapsigargin, and tunicamycin), physical insults (ultraviolet radiation), and suboptimal growing conditions (growth factor deprivation) (Wei *et al.*, 2001).

The contribution of MOMP to cell death is dual, first supporting apoptosis and second precipitating in the deterioration of mitochondrial physiology that severely compromises cellular function. The primary intended effect of MOMP is the escape of apoptogenic proteins from the mitochondria into the cytoplasm. As we have described cytochrome c stimulates caspase activation through the apoptosome, while other molecules such as the second mitochondria-derived activator of caspase (SMAC) has an indirect role antagonizing the negative effect of caspase inhibitors (Tait and Green, 2010; Ferrer *et al.*, 2012). The secondary contribution of MOMP is less specific, exhibits slower kinetics compared to that of caspase activation, and signifies an irreversible commitment to die as it carries severe damage in the capacity of the mitochondria to generate energy for the cell. The proposed mechanism involves ATP depletion in the mitochondria as well

as impairment of the respiratory function to generate ATP necessary for clonogenic survival (Lartigue *et al.*, 2009; Kushnareva and Newmeyer, 2010). In light of its cleavage by caspase-8 and as a promoter of cytochrome c release to activate caspase-9, BID has been regarded as an element in which the intrinsic and the extrinsic pathways intersect to potentiate the effect of the initial stimulus (Youle and Strasser, 2008).

Other apoptogenic molecules

In the late 1980s shortly before the discovery of the first caspase, Tschoopp described the isolation and characterization of eight proteases contained in the lytic granules of cytotoxic T lymphocytes (Masson and Tschoopp, 1987). These enzymes termed granzymes (Gzms) are produced by diverse populations of T lymphocytes and natural killers cells (NK) to induce death of malignant and virus-infected cells in order to prevent propagation (Lord *et al.*, 2003). Gzms are serine proteases with a catalytic triad formed by His-Ser-Asp, cleavage specificities among the members of the family are different. GzmM cleaves after Met or Leu, GzmC and GzmH are chymotrypsin-like enzymes that cleave after aromatic residues, GzmA cleaves after basic residues and GzmB acts as caspase-like protease that cleaves after Asp (Chowdhury and Lieberman, 2008; Trapani, 2001; Van Damme *et al.*, 2010). GzmB contributes to apoptosis initiation by directly cleaving and activating caspase-3, thus acting as a pseudo-apical caspase. At the same time GzmB may cleave BID that, as previously discussed for caspase-8, produces cBID to induce MOMP resulting in caspase-9 activation (Lord *et al.*, 2003). The role of the other granzymes in inducing cell death is highly debated, and a consensus is yet to arise.

Lipids

The various aspects of apoptosis explained so far rely entirely on proteins as the primary molecular effector, mounting evidence implicates lipids as prominent figures in apoptosis, noteworthy in the context of membranes. For this matter membranes should not be considered as passive witnesses that form barriers to protect the cell or limit compartments with specialized and dedicated cellular functions. Beyond membrane blebbing, a hallmark of apoptosis, mitochondrial membranes participate actively in apoptosis (Fink and Cookson, 2005). Previously, roles for lipids in cell death were proposed even before caspase participation in apoptosis was established but a mechanistic description was scarce (Obeid *et al.*, 1993).

Initial research to address the involvement of lipids in apoptosis proposed that ceramide had a direct role by way of forming stable channels in the mitochondria outer membrane (MOM) hence supporting escape of small apoptogenic proteins from the mitochondria (e.g., cytochrome c) (Siskind *et al.*, 2002, 2006). However, an alternative model emerging from recent work suggests that metabolic intermediates of ceramide provide the appropriate membrane milieu to anchor Bcl-2 proteins BAK and BAX, therefore acting synergistically in promoting their insertion and polymerization (Ganesan *et al.*, 2010; Chipuk *et al.*, 2012). Under this model sphingosine 1-phosphate (S1P), a precursor in the synthesis ceramide, is necessary for BAK activation stimulated by caspase-8-processed BID. In contrast, activation of BAX was observed to require association with hexadecenal (Hex), a molecule produced

from sphingosine, the immediate ceramide precursor (Chipuk *et al.*, 2012). In both cases regulation of BAK/BAX polymerization by lipids had two direct effects: MOMP and cytochrome c release. Consistent with this data the systematic inhibition or overexpression of enzymes that participate in the metabolism of ceramide, S1P, sphingosine, and Hex lead to either sensitization of cells to apoptosis or clonogenic survival.

Anti-Apoptotic Proteins

IAPs

Considering caspases as the directors that orchestrate the symphony of destruction during apoptosis is not unexpected that these proteases are a molecular target for apoptosis control. A direct mode of caspase regulation is through interaction with members of the IAP family. IAPs are characterized by the presence of 1–3 Baculovirus IAP repeat (BIR), a motif of approximately 70 amino acids that mediates protein–protein interactions (Edlich *et al.*, 2011). The first IAP was described in a Baculovirus strain capable of suppressing apoptosis in infected cells (Clem *et al.*, 1991) and this was followed by the identification of cellular IAPs in insects and vertebrates (Oberoi-Khanuja *et al.*, 2013). In addition to BIR repeats several IAPs contain also a RING domain (really interesting new gene) and a ubiquitin-associated (UBA) domain that endows these proteins to participate in cellular processes regulated by ubiquitination including inflammation, immunity, cell migration, and apoptosis (Gyrd-Hansen and Meier, 2010).

Although all human IAPs, of which there are eight, contain at least one BIR domain the X-linked IAP (XIAP) is the only bona fide caspase inhibitor identified to date (Gyrd-Hansen and Meier, 2010; Oberoi-Khanuja *et al.*, 2013). Specifically, the BIR2 domain inhibits caspase-3 and -7 while BIR3 inhibits caspase -9 (Eckelman *et al.*, 2006; Gyrd-Hansen and Meier, 2010). Unexpectedly, two close paralogs of XIAP, the cellular IAP1 and IAP2 (cIAP1 and cIAP2) also bind caspases -3, -7, and -9 but these are at best weak inhibitors of caspase activity. Presumably the lack of inhibitory action in cIAP1 and cIAP2 is due to mutations on critical residues forming the interacting surfaces with caspases (Eckelman and Salvesen, 2006). Currently, described functions for cIAP1 and cIAP2 relate mainly to their RING domain including regulation of tumor cell migration, extending cancer cell survival and paradoxically activation of caspase-1 through the inflammasome (Oberoi *et al.*, 2012; Bertrand *et al.*, 2008; Labbe *et al.*, 2011). The CARD domain is found only in cIAP1 and cIAP2 but is still not well described. A study that combines structural as well as functional data suggests that in cIAP1 the CARD domain regulates RING negatively preventing its activation to suppress both cell proliferation and migration (Lopez *et al.*, 2011).

Since XIAP is the most relevant IAP in the context of apoptosis we will focus our attention to this protein. XIAP is a well-rounded apoptosis regulator that has evolved several domains to counteract cell death in diverse nodes of function. XIAP is composed of three BIR domains (BIR1, BIR2, and BIR3), a UBA domain, and a RING domain (Gyrd-Hansen and Meier, 2010). The BIR1 domain acts as scaffold for the dimerization and activation of TAK1, which in turn activates the I-kappa B kinase complex (IKK). Activation of IKK

culminates in the activation of the NF-kappaB pathway that indirectly antagonizes cell death by promoting cell survival (Lu *et al.*, 2007). Inhibition of caspase-3 and -7 by BIR2 follows a model of two-site interaction that cooperatively contributes to the full inhibitory power of BIR2. An N-terminal segment that extends beyond BIR2 binds to the caspase active site most likely blocking or reducing substrate access. The second point of interaction comprises a surface groove in BIR2 and the N-terminus of the caspase small subunits (Scott *et al.*, 2005). Inhibition by BIR3 follows a different mechanism to that of BIR2; instead of interfering with the catalytic activity, BIR3 prevents caspase-9 activation. The active form of caspase-9 is an obligate homodimer, a state that is maintained by extensive interactions between β -strands in the small subunits (Renatus *et al.*, 2001). BIR3 inhibits caspase-9 by binding to the monomer in the same region that drives dimerization hence blocking the formation of active dimers (Shiozaki *et al.*, 2003).

The Bcl-2 family

The anti-apoptotic subfamily of the Bcl-2 group is small, sequence analysis of the human genome has identified a total of eight proteins including the first and founding member of this family, Bcl-2 (Youle and Strasser, 2008). Bcl-2 was originally identified as an aberrant chromosomal translocation in a patient suffering acute B-cell lymphoma (thereof its name, B-cell lymphoma-2) and was assumed to be an oncogene (Tsujimoto *et al.*, 1984). This assumption was based exclusively on the analogy with the chromosomal translocation of c-myc present in other tumors of B-cells, but the underlying mechanism was yet to be explored. At that time oncogenes were known to mediate cancer only through unregulated cellular proliferation. Subsequent studies found the connection with apoptosis, Bcl-2 functions to block cytochrome c release thus preventing apoptosis and promoting cancer through cell survival and not via cell proliferation (Vaux *et al.*, 1988; Kluck *et al.*, 1997). To play a pro-survival role Bcl-2 and other homologs like Bcl-XL heterodimerize directly with pro-apoptotic family members Bax and Bak to preclude their homoligomerization at the MOM (Michels *et al.*, 2013). By blocking Bax and Bak, Bcl-2 and Bcl-XL prevent MOMP and inhibit downstream signals in the apoptotic program: activation of caspase -9, and caspase-3 and -7. Furthermore after Bax is translocated to the mitochondria Bcl-XL inhibits Bax by promoting its retrotranslocation back to the cytoplasm, thus functioning as a constant downregulator of MOMP (Edlich *et al.*, 2011).

Members of all three subfamilies of Bcl-2 form an interaction network to regulate apoptosis; a BH3-only protein can interact with different pro-apoptotic molecules, and a Bcl-2 pro-survival protein may interact with different BH3-only proteins or different pro-apoptotic proteins (Shamas-Din *et al.*, 2013). The first model to explain the interplay of Bcl-2 proteins to decide the cell's fate was the 'rheostat' model that proposed a quantitative regulation in which the balance between proteins with opposing biological activities would determine if a cell would proliferate or enter in apoptosis (Korsmeyer *et al.*, 1993). The rheostat model was based only on the ratios of Bcl-2/Bax, however as research in the field has progressed more inclusive and comprehensive models have been proposed. Recently the 'unified' model has emerged to

incorporate the interactions of all three subfamilies of Bcl-2 proteins (Llambi *et al.*, 2011). Under this model two 'modes' of interactions are distinguished, in the first mode the pro-survival proteins sequester BH3-only proteins (mode 1) and in the second mode the pro-survival proteins bind activated proteins BAX and BAK (mode 2). During mode 1, sequestration of BH3 only proteins is less efficient to promote MOMP while mode 2 is a more direct and efficient mechanism to prevent MOMP. In addition, the unified model acknowledges from a previous model 'embedded together' the significance of mitochondrial membrane as the focal point where Bcl-2 proteins reach their active conformation to promote apoptosis (Shamas-Din *et al.*, 2013).

The application of biophysical techniques has helped to dissect with great detail the structural basis of the inhibitory interactions between the apoptotic and anti-apoptotic Bcl-2 proteins. A common feature is the formation of a heterodimer in which an amphipathic α -helix from the BH3 domain of the apoptotic protein is inserted into a hydrophobic groove in the anti-apoptotic binding partner. Examples of this include the BH3 domain from Bak that binds to the hydrophobic cleft in Bcl-XL, the BH3 domain from Bad that binds to Bcl-XL, the BH3 domain from Bim that binds to Bcl-XL and the BH3 domain from Bax that binds to Bcl-2 (Sattler *et al.*, 1997; Petros *et al.*, 2000; Liu *et al.*, 2003; Lomonosova *et al.*, 2009; Ku *et al.*, 2011). Binding of the BH3 domain from pro-apoptotic proteins to anti-apoptotic proteins is not seen exclusively in heterodimers of these subfamilies of Bcl-2 proteins. Homoligomerization of Bax or Bak, activation of Bax and Bak by Bim and Bid in order to promote apoptosis require the same structural elements (Moldoveanu *et al.*, 2013). Surprisingly the interaction of Bax with Bcl-XL during Bax retrotranslocation from the mitochondria to cytoplasm seems to involve the membrane anchoring-region of Bcl-XL with the hydrophobic groove of Bax (Todt *et al.*, 2013).

Anti-apoptosis proteins of viral origin

Viruses are one of the greatest successes in parasitic biology, to fulfill their life cycle they replicate and spread their progeny at the expense of the infected cell. To achieve this viruses have evolved a myriad of almost 'intelligent' strategies: evasion of the host's immune system, manipulation of the cell's molecular machinery and control of apoptosis in the infected cell. Undoubtedly, controlling apoptosis until enough progeny has been secured is a powerful survival advantage for the virus (Roulston *et al.*, 1999). Below we will describe briefly some of the molecular strategies employed by viruses to inhibit apoptosis.

During homeostasis apoptosis functions to maintain adequate numbers of T cells, this is mediated by the expression of the death receptor Fas on the surface of the target cell (Kaplan and Sieg, 1998). When Fas ligand binds to the Fas receptor the interaction triggers the extrinsic apoptotic cascade that activates pro-caspase-8 at the DISC (Holler *et al.*, 2000). Upon infection with the herpes simplex virus type 2 (HSV-2) cells are inhibited from expressing the Fas ligand on the surface, thus HSV-2 suppresses Fas ligand-mediated apoptosis in cells bearing the Fas receptor (Sieg *et al.*, 1996). Lower amounts of Fas ligand found on the surface of infected cells do not correlate with reduced mRNA expression or limited protein production but instead the virus seems to disrupt the

intracellular mechanism of ligand transport (Sieg *et al.*, 1996). This mechanism of cell death inhibition is specific for HSV-2 since HSV-1 modulates apoptosis through a different mechanism. Thus far seven genes from HSV-1 have been proposed to exhibit anti-apoptosis activity with U₅₃ being the most certain of all genes identified. During the first 3 h postinfection with HSV-1 cells may undergo apoptosis, nevertheless after this time U₅₃ blocks cytochrome c release and caspase-3 activation. This suggests that U₅₃ prevents apoptosis upstream of mitochondrial signaling (Nguyen and Blaho, 2007).

Blocking cells from embarking in apoptosis may not only potentiate progeny expansion, but also prevent extensive tissue damage to prolong proper organismal functioning, thus increasing the possibility of infecting other individuals. The hepatitis C virus (HCV) is the etiological agent of hepatitis C, a viral infection that causes severe liver damage and chronic infection that can last for 30 years. HCV is observed to block the extrinsic pathway of apoptosis when cells are stimulated with the TNF-related apoptosis-inducing ligand (TRAIL). In addition, when cells are treated with the chemotherapeutic etoposide HCV inhibits the intrinsic pathway, this effect seems to be mediated by blockage of cytochrome c release (Lee *et al.*, 2005). In both instances, the envelope glycoprotein E2 that controls cell entry and fusion is implicated in regulating resistance of cells to die via apoptosis (Lee *et al.*, 2005; Kong *et al.*, 2013).

Another strategy to control apoptosis is common to several viruses and relies on the direct inhibition of caspase activity by serpins (serine protease inhibitors). Serpins are potent inhibitors of serine and cysteine proteases found in all kingdoms, in mammals these proteins regulate clot formation, blood pressure, connective tissue breakdown and hormone transport (Huntington, 2011; Chen *et al.*, 2011). Viruses have engineered serpins to their own advantage; CrmA (cytokine response modifier A) from the cowpox virus inhibits caspase-1, -6, and -8 by two mechanisms. First CrmA forms an irreversible covalent thioacyl intermediate with the active site cysteine effectively trapping the enzyme to prevent catalysis. Second it induces a significant structural change that further renders enzymes catalytically incompetent. Analysis by mass spectrometry and SDS-PAGE show that the complex of CrmA with caspase-1 or -8 undergoes extensive loss of the small caspase subunit p10 (Dobo *et al.*, 2006). As the active forms of caspase-1 and -8 are obligate homodimers of the form (p10p20)₂, disruption of this stoichiometric ratio upon p10 dissociation is a bonus toward viral caspase inhibition and apoptosis control.

Other non-serpin proteins from viruses also inhibit caspases through covalent-intermediates similar to that explained from CrmA. The protein p35 from the cowpox virus inhibits caspase-1, -3, -6, -7, -8, and -10 (Zhou *et al.*, 1998). Inhibition of caspase-8 by p35 involves the formation of a non-cleavable thioester bond that prevents further cycles of catalysis (Xu *et al.*, 2001). Generally during the reaction an activated water molecule hydrolyzes the thioester intermediate to release the cleaved substrate and the free enzyme. However the N-terminus of p35 blocks access of the water molecule to the active site and the thioester intermediate is 'trapped' in catalysis to prevent regeneration of free caspase-8 (Xu *et al.*, 2001).

Acknowledgment

The work of Mario F. Navarro and Guy Salvesen are supported by NIH grant RO1GM09040.

See also: Cell Division/Death: Apoptosis: Anoikis; Caspases; Inhibitor of Apoptosis Proteins, the Sentinels of Cell Death and Signaling; The Bcl-2 Family Proteins: Insights into Their Mechanism of Action and Therapeutic Potential; Tumor Necrosis Factor Signaling Pathways. Cell Division/Death: Cell Cycle: The Restriction Point

References

- Bertrand, M.J., Milutinovic, S., Dickson, K.M., *et al.*, 2008. cIAP1 and cIAP2 facilitate cancer cell survival by functioning as E3 ligases that promote RIP1 ubiquitination. *Molecular Cell* 30, 689–700.
- Boatright, K.M., Deis, C., Denault, J.B., Sutherlin, D.P., Salvesen, G.S., 2004. Activation of caspases-8 and -10 by FLIP(L). *Biochemical Journal* 382, 651–657.
- Boatright, K.M., Renatus, M., Scott, F.L., *et al.*, 2003. A unified model for apical caspase activation. *Molecular Cell* 11, 529–541.
- Cerretti, D.P., Kozlosky, C.J., Mosley, B., *et al.*, 1992. Molecular cloning of the interleukin-1 beta converting enzyme. *Science* 256, 97–100.
- Chang, D.W., Xing, Z., Capacio, V.L., Peter, M.E., Yang, X., 2003. Interdimer processing mechanism of procaspase-8 activation. *EMBO Journal* 22, 4132–4142.
- Chen, H., Zheng, D., Davids, J., *et al.*, 2011. Viral serpin therapeutics from concept to clinic. *Methods in Enzymology* 499, 301–329.
- Chipuk, J.E., McStay, G.P., Bharti, A., *et al.*, 2012. Sphingolipid metabolism cooperates with BAK and BAX to promote the mitochondrial pathway of apoptosis. *Cell* 148, 988–1000.
- Chowdhury, D., Lieberman, J., 2008. Death by a thousand cuts: Granzyme pathways of programmed cell death. *Annual Review of Immunology* 26, 389–420.
- Clem, R.J., Fechheimer, M., Miller, L.K., 1991. Prevention of apoptosis by a baculovirus gene during infection of insect cells. *Science* 254, 1388–1390.
- Degterev, A., Boyce, M., Yuan, J., 2003. A decade of caspases. *Oncogene* 22, 8543–8567.
- Dix, M.M., Simon, G.M., Wang, C., *et al.*, 2012. Functional interplay between caspase cleavage and phosphorylation sculpts the apoptotic proteome. *Cell* 150, 426–440.
- Dobo, J., Swanson, R., Salvesen, G.S., Olson, S.T., Gettins, P.G., 2006. Cytokine response modifier 1 inhibition of initiator caspases results in covalent complex formation and dissociation of the caspase tetramer. *Journal of Biological Chemistry* 281, 38781–38790.
- Duprez, L., Wirawan, E., Vanden Berghe, T., Vandenabeele, P., 2009. Major cell death pathways at a glance. *Microbes and Infection* 11, 1050–1062.
- Eckelman, B.P., Salvesen, G.S., 2006. The human anti-apoptotic proteins cIAP1 and cIAP2 bind but do not inhibit caspases. *Journal of Biological Chemistry* 281, 3254–3260.
- Eckelman, B.P., Salvesen, G.S., Scott, F.L., 2006. Human inhibitor of apoptosis proteins: Why XIAP is the black sheep of the family. *EMBO Reports* 7, 988–994.
- Edlich, F., Banerjee, S., Suzuki, M., *et al.*, 2011. Bcl-x(L) retrotranslocates Bax from the mitochondria into the cytosol. *Cell* 145, 104–116.
- Eggert, A., Grotzer, M.A., Zuzak, T.J., *et al.*, 2001. Resistance to tumor necrosis factor-related apoptosis-inducing ligand (TRAIL)-induced apoptosis in neuroblastoma cells correlates with a loss of caspase-8 expression. *Cancer Research* 61, 1314–1319.
- Elmore, S., 2007. Apoptosis: A review of programmed cell death. *Toxicologic Pathology* 35, 495–516.
- Favaloro, B., Allocati, N., Graziano, V., Di Ilio, C., De Laurenzi, V., 2012. Role of apoptosis in disease. *Aging (Albany, NY)* 4, 330–349.
- Ferrer, P.E., Frederick, P., Gulbis, J.M., Dewson, G., Kluck, R.M., 2012. Translocation of a Bak C-terminus mutant from cytosol to mitochondria to mediate cytochrome C release: Implications for Bak and Bax apoptotic function. *PLoS One* 7, e31510.
- Fink, S.L., Cookson, B.T., 2005. Apoptosis, pyroptosis, and necrosis: Mechanistic description of dead and dying eukaryotic cells. *Infection and Immunity* 73, 1907–1916.
- Fischer, U., Schulze-Osthoff, K., 2005. New approaches and therapeutics targeting apoptosis in disease. *Pharmacological Reviews* 57, 187–215.
- Fuentes-Prior, P., Salvesen, G.S., 2004. The protein structures that shape caspase activity, specificity, activation and inhibition. *Biochemical Journal* 384, 201–232.
- Ganesan, V., Perera, M.N., Colombini, D., *et al.*, 2010. Ceramide and activated Bax act synergistically to permeabilize the mitochondrial outer membrane. *Apoptosis* 15, 553–562.
- Gyrd-Hansen, M., Meier, P., 2010. IAPs: From caspase inhibitors to modulators of NF-kappaB, inflammation and cancer. *Nature Reviews Cancer* 10, 561–574.
- Holler, N., Zaru, R., Micheau, O., *et al.*, 2000. Fas triggers an alternative, caspase-8-independent cell death pathway using the kinase RIP as effector molecule. *Nature Immunology* 1, 489–495.
- Huntington, J.A., 2011. Serpin structure, function and dysfunction. *Journal of Thrombosis and Haemostasis* 9 (Suppl. 1), 26–34.
- Johnson, V.L., Ko, S.C., Holmstrom, T.H., Eriksson, J.E., Chow, S.C., 2000. Effector caspases are dispensable for the early nuclear morphological changes during chemical-induced apoptosis. *Journal of Cell Science* 113 (Pt 17), 2941–2953.
- Kaiser, W.J., Upton, J.W., Long, A.B., *et al.*, 2011. RIP3 mediates the embryonic lethality of caspase-8-deficient mice. *Nature* 471, 368–372.
- Kaplan, D., Sieg, S., 1998. Role of the Fas/Fas ligand apoptotic pathway in human immunodeficiency virus type 1 disease. *Journal of Virology* 72, 6279–6282.
- Karpinich, N.O., Tafani, M., Schneider, T., Russo, M.A., Farber, J.L., 2006. The course of etoposide-induced apoptosis in Jurkat cells lacking p53 and Bax. *Journal of Cellular Physiology* 208, 55–63.
- Kayagaki, N., Warming, S., Lamkanfi, M., *et al.*, 2011. Non-canonical inflammasome activation targets caspase-11. *Nature* 479, 117–121.
- Kluck, R.M., Bossy-Wetzel, E., Green, D.R., Newmeyer, D.D., 1997. The release of cytochrome c from mitochondria: A primary site for Bcl-2 regulation of apoptosis. *Science* 275, 1132–1136.
- Kong, L., Giang, E., Niesuma, T., *et al.*, 2013. Hepatitis C virus E2 envelope glycoprotein core structure. *Science* 342, 1090–1094.
- Korsmeyer, S.J., Shutter, J.R., Veis, D.J., Merry, D.E., Oltvai, Z.N., 1993. Bcl-2/Bax: A rheostat that regulates an anti-oxidant pathway and cell death. *Seminars in Cancer Biology* 4, 327–332.
- Ku, B., Liang, C., Jung, J.U., Oh, B.H., 2011. Evidence that inhibition of BAX activation by BCL-2 involves its tight and preferential interaction with the BH3 domain of BAX. *Cell Research* 21, 627–641.
- Kushnareva, Y., Newmeyer, D.D., 2010. Bioenergetics and cell death. *Annals of the New York Academy of Sciences* 1201, 50–57.
- Labbe, K., McIntire, C.R., Doiron, K., Leblanc, P.M., Saleh, M., 2011. Cellular inhibitors of apoptosis proteins cIAP1 and cIAP2 are required for efficient caspase-1 activation by the inflammasome. *Immunity* 35, 897–907.
- Lartigue, L., Kushnareva, Y., Seong, Y., *et al.*, 2009. Caspase-independent mitochondrial cell death results from loss of respiration, not cytotoxic protein release. *Molecular Biology of the Cell* 20, 4871–4884.
- Lee, S.H., Kim, Y.K., Kim, C.S., *et al.*, 2005. E2 of hepatitis C virus inhibits apoptosis. *Journal of Immunology* 175, 8226–8235.
- Lettre, G., Hengartner, M.O., 2006. Developmental apoptosis in *C. elegans*: A complex CEDnario. *Nature Reviews Molecular Cell Biology* 7, 97–108.
- Liu, X., Dai, S., Zhu, Y., Marrack, P., Kappler, J.W., 2003. The structure of a Bcl-xL/Bim fragment complex: Implications for Bim function. *Immunity* 19, 341–352.
- Llambi, F., Moldoveanu, T., Tait, S.W., *et al.*, 2011. A unified model of mammalian BCL-2 protein family interactions at the mitochondria. *Molecular Cell* 44, 517–531.
- Lomonosova, E., Ryerse, J., Chinnadurai, G., 2009. BAX/BAK-independent mitoptosis during cell death induced by proteasome inhibition? *Molecular Cancer Research* 7, 1268–1284.
- Lopez, J., John, S.W., Tenev, T., *et al.*, 2011. CARD-mediated autoinhibition of cIAP1's E3 ligase activity suppresses cell proliferation and migration. *Molecular Cell* 42, 569–583.
- Lord, S.J., Rajotte, R.V., Korbitt, G.S., Bleackley, R.C., 2003. Granzyme B: A natural born killer. *Immunological Reviews* 193, 31–38.
- Lu, M., Lin, S.C., Huang, Y., *et al.*, 2007. XIAP induces NF-kappaB activation via the BIR1/Table1 interaction and BIR1 dimerization. *Molecular Cell* 26, 689–702.
- Maiuri, M.C., Zalckvar, E., Kimchi, A., Kroemer, G., 2007. Self-eating and self-killing: Crosstalk between autophagy and apoptosis. *Nature Reviews Molecular Cell Biology* 8, 741–752.
- Masson, D., Tschopp, J., 1987. A family of serine esterases in lytic granules of cytolytic T lymphocytes. *Cell* 49, 679–685.
- McIlwain, D.R., Berger, T., Mak, T.W., 2013. Caspase functions in cell death and disease. *Cold Spring Harbor Perspectives in Biology* 5, a008656.

- Melino, G., 2001. The Sirens' song. *Nature* 412, 23.
- Michels, J., Kepp, O., Senovilla, L., *et al.*, 2013. Functions of BCL-X L at the Interface between cell death and metabolism. *International Journal of Cell Biology* 2013, 10.
- Moldoveanu, T., Grace, C.R., Llambi, F., *et al.*, 2013. BID-induced structural changes in BAK promote apoptosis. *Nature Structural & Molecular Biology* 20, 589–597.
- Morris, S.W., Nelson, N., Valentine, M.B., *et al.*, 1992. Assignment of the genes encoding human interleukin-8 receptor types 1 and 2 and an interleukin-8 receptor pseudogene to chromosome 2q35. *Genomics* 14, 685–691.
- Nguyen, M.L., Blaho, J.A., 2007. Apoptosis during herpes simplex virus infection. *Advances in Virus Research* 69, 67–97.
- Obeid, L.M., Linardic, C.M., Karolak, L.A., Hannun, Y.A., 1993. Programmed cell death induced by ceramide. *Science* 259, 1769–1771.
- Oberoi, T.K., Dogan, T., Hocking, J.C., *et al.*, 2012. IAPs regulate the plasticity of cell migration by directly targeting Rac1 for degradation. *EMBO Journal* 31, 14–28.
- Oberoi-Khanuja, T.K., Murali, A., Rajalingam, K., 2013. IAPs on the move: Role of inhibitors of apoptosis proteins in cell migration. *Cell Death and Disease* 4, e784.
- Oberst, A., Dillon, C.P., Weinlich, R., *et al.*, 2011. Catalytic activity of the caspase-8-FLIP(L) complex inhibits RIPK3-dependent necrosis. *Nature* 471, 363–367.
- Peng, C., Cho, Y.Y., Zhu, F., *et al.*, 2011. Phosphorylation of caspase-8 (Thr-263) by ribosomal S6 kinase 2 (RSK2) mediates caspase-8 ubiquitination and stability. *Journal of Biological Chemistry* 286, 6946–6954.
- Petros, A.M., Nettesheim, D.G., Wang, Y., *et al.*, 2000. Rationale for Bcl-xL/Bad peptide complex formation from structure, mutagenesis, and biophysical studies. *Protein Science* 9, 2528–2534.
- Pop, C., Oberst, A., Drag, M., *et al.*, 2011. FLIP(L) induces caspase 8 activity in the absence of interdomain caspase 8 cleavage and alters substrate specificity. *Biochemical Journal* 433, 447–457.
- Pop, C., Salvesen, G.S., 2009. Human caspases: Activation, specificity, and regulation. *Journal of Biological Chemistry* 284, 21777–21781.
- Pop, C., Timmer, J., Sperandio, S., Salvesen, G.S., 2006. The apoptosome activates caspase-9 by dimerization. *Molecular Cell* 22, 269–275.
- Portt, L., Norman, G., Clapp, C., Greenwood, M., Greenwood, M.T., 2011. Anti-apoptosis and cell survival: A review. *Biochimica et Biophysica Acta* 1813, 238–259.
- van Raam, B.J., Ehrnhofer, D.E., Hayden, M.R., Salvesen, G.S., 2013. Intrinsic cleavage of receptor-interacting protein kinase-1 by caspase-6. *Cell Death and Differentiation* 20, 86–96.
- Renatus, M., Stennicke, H.R., Scott, F.L., Liddington, R.C., Salvesen, G.S., 2001. Dimer formation drives the activation of the cell death protease caspase 9. *Proceedings of the National Academy of Sciences of the United States of America* 98, 14250–14255.
- Riedl, S.J., Salvesen, G.S., 2007. The apoptosome: Signalling platform of cell death. *Nature Reviews Molecular Cell Biology* 8, 405–413.
- Roulston, A., Marcellus, R.C., Branton, P.E., 1999. Viruses and apoptosis. *Annual Review of Microbiology* 53, 577–628.
- Salvesen, G.S., Ashkenazi, A., 2011. Snapshot: Caspases. *Cell* 147, 476–476e1.
- Salvesen, G.S., Riedl, S.J., 2008. Caspase mechanisms. *Advances in Experimental Medicine and Biology* 615, 13–23.
- Sattler, M., Liang, H., Nettesheim, D., *et al.*, 1997. Structure of Bcl-xL-Bak peptide complex: Recognition between regulators of apoptosis. *Science* 275, 983–986.
- Scott, F.L., Denault, J.B., Riedl, S.J., *et al.*, 2005. XIAP inhibits caspase-3 and -7 using two binding sites: Evolutionarily conserved mechanism of IAPs. *EMBO Journal* 24, 645–655.
- Shamas-Din, A., Kale, J., Leber, B., Andrews, D.W., 2013. Mechanisms of action of Bcl-2 family proteins. *Cold Spring Harbor Perspectives in Biology* 5, a008714.
- Shiozaki, E.N., Chai, J., Rigotti, D.J., *et al.*, 2003. Mechanism of XIAP-mediated inhibition of caspase-9. *Molecular Cell* 11, 519–527.
- Sieg, S., Yildirim, Z., Smith, D., *et al.*, 1996. Herpes simplex virus type 2 inhibition of Fas ligand expression. *Journal of Virology* 70, 8747–8751.
- Siegel, R.M., 2006. Caspases at the crossroads of immune-cell life and death. *Nature Reviews Immunology* 6, 308–317.
- Siskind, L.J., Kolesnick, R.N., Colombini, M., 2002. Ceramide channels increase the permeability of the mitochondrial outer membrane to small proteins. *Journal of Biological Chemistry* 277, 26796–26803.
- Siskind, L.J., Kolesnick, R.N., Colombini, M., 2006. Ceramide forms channels in mitochondrial outer membranes at physiologically relevant concentrations. *Mitochondrion* 6, 118–125.
- Slee, E.A., Adrain, C., Martin, S.J., 2001. Executioner caspase-3, -6, and -7 perform distinct, non-redundant roles during the demolition phase of apoptosis. *Journal of Biological Chemistry* 276, 7320–7326.
- Stennicke, H.R., Salvesen, G.S., 1999. Catalytic properties of the caspases. *Cell Death and Differentiation* 6, 1054–1059.
- Stennicke, H.R., Salvesen, G.S., 2000. Caspases – Controlling intracellular signals by protease zymogen activation. *Biochimica et Biophysica Acta* 1477, 299–306.
- Tait, S.W., Green, D.R., 2010. Mitochondria and cell death: Outer membrane permeabilization and beyond. *Nature Reviews Molecular Cell Biology* 11, 621–632.
- Taylor, R.C., Cullen, S.P., Martin, S.J., 2008. Apoptosis: Controlled demolition at the cellular level. *Nature Reviews Molecular Cell Biology* 9, 231–241.
- Todt, F., Cakir, Z., Reichenbach, F., Youle, R.J., Edlich, F., 2013. The C-terminal helix of Bcl-x(L) mediates Bax retrotranslocation from the mitochondria. *Cell Death and Differentiation* 20, 333–342.
- Trapani, J.A., 2001. Granzymes: A family of lymphocyte granule serine proteases. *Genome Biology* 2, 3014.1–3014.7.
- Tsujimoto, Y., Yunis, J., Onorato-Showe, L., *et al.*, 1984. Molecular cloning of the chromosomal breakpoint of B-cell lymphomas and leukemias with the t(11;14) chromosome translocation. *Science* 224, 1403–1406.
- Tyas, L., Brophy, V.A., Pope, A., Rivett, A.J., Tavare, J.M., 2000. Rapid caspase-3 activation during apoptosis revealed using fluorescence-resonance energy transfer. *EMBO Reports* 1, 266–270.
- Van Damme, P., Maurer-Stroh, S., Hao, H., *et al.*, 2010. The substrate specificity profile of human granzyme A. *Biological Chemistry* 391, 983–997.
- Van de Craen, M., Van Loo, G., Pype, S., *et al.*, 1998. Identification of a new caspase homologue: Caspase-14. *Cell Death and Differentiation* 5, 838–846.
- Varfolomeev, E.E., Schuchmann, M., Luria, V., *et al.*, 1998. Targeted disruption of the mouse Caspase 8 gene ablates cell death induction by the TNF receptors, Fas/Apo1, and DR3 and is lethal prenatally. *Immunity* 9, 267–276.
- Vaux, D.L., Cory, S., Adams, J.M., 1988. Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells. *Nature* 335, 440–442.
- Walsh, J.G., Cullen, S.P., Sheridan, C., *et al.*, 2008. Executioner caspase-3 and caspase-7 are functionally distinct proteases. *Proceedings of the National Academy of Sciences of the United States of America* 105, 12815–12819.
- Wang, C., Youle, R.J., 2009. The role of mitochondria in apoptosis. *Annual Review of Genetics* 43, 95–118.
- Watt, W., Koepfinger, K.A., Mildner, A.M., *et al.*, 1999. The atomic-resolution structure of human caspase-8, a key activator of apoptosis. *Structure* 7, 1135–1143.
- Wei, M.C., Zong, W.X., Cheng, E.H., *et al.*, 2001. Proapoptotic BAX and BAK: A requisite gateway to mitochondrial dysfunction and death. *Science* 292, 727–730.
- Xu, G., Cirilli, M., Huang, Y., *et al.*, 2001. Covalent inhibition revealed by the crystal structure of the caspase-8/p35 complex. *Nature* 410, 494–497.
- Yi, C.H., Yuan, J., 2009. The Jekyll and Hyde functions of caspases. *Developmental Cell* 16, 21–34.
- Youle, R.J., Strasser, A., 2008. The BCL-2 protein family: Opposing activities that mediate cell death. *Nature Reviews Molecular Cell Biology* 9, 47–59.
- Zhou, Q., Krebs, J.F., Snipas, S.J., *et al.*, 1998. Interaction of the baculovirus anti-apoptotic protein p35 with caspases. Specificity, kinetics, and characterization of the caspase/p35 complex. *Biochemistry* 37, 10757–10765.
- Zmasek, C.M., Zhang, Q., Ye, Y., Godzik, A., 2007. Surprising complexity of the ancestral apoptosis network. *Genome Biology* 8, R226.

Relevant Website

<http://australianmuseum.net.au/Preparation-for-death-in-ancient-Egypt>
Australian Museum.

Anoikis

SM Frisch, West Virginia University, Morgantown, WV, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Normal epithelial cells detach from their resident extracellular matrices with high frequency. Anoikis, an ancient Greek word meaning 'homelessness,' is the apoptotic response that ensures they are eliminated (Frisch and Francis, 1994; Frisch *et al.*, 2013). This response prevents the translocation of cells with (epi)genetic changes from attaching elsewhere and exploiting their growth/survival advantages or altered differentiation states from harmful effects on their ectopic tissue of residence. Tumor metastasis requires the suppression of anoikis, through mechanisms that are under extensive investigation (Buchheit *et al.*, 2014; Charpentier and Martin, 2013; Paoli *et al.*, 2013; Simpson *et al.*, 2008).

The core machinery – i.e., death-effector mechanism – of anoikis is identical to that of apoptosis in general (Frisch and Screaton, 2001). Thus, anoikis generally depends upon mitochondrial outer membrane permeabilization (MOMP), with release of cytochrome *c* to generate a caspase-9-containing apoptosome that activates effector caspases. Upstream of mitochondria, Bcl-2 family proteins contribute pro-apoptotic and anti-apoptotic roles as with other apoptotic settings, with Mcl-1 perhaps being of exceptional importance in some contexts (Woods *et al.*, 2007). The BH3-only family genes (Bim, Bmf, Bad, Noxa, and Puma) are very responsive to integrin-assisted signal transduction and have therefore been characterized extensively in the context of anoikis, with evidence for cytoskeletal control by direct interaction in some cases (Schmelzle *et al.*, 2007).

The best-characterized signaling pathways that control the core apoptotic machinery in anoikis are the MAP kinases (ERKs, JNK, and p38 MAPK) and the PI3-Kinase/Akt pathway. The latter is of particular interest, because this pathway is highly regulated both by integrin ligation and metabolic nodes such as insulin/IGF receptors. Moreover, Akt programs downstream metabolism through mTOR, thus affecting cell growth, apoptosis, and autophagy (Ng *et al.*, 2012). Focal adhesion kinase (FAK) is a central non-receptor tyrosine kinase that, in complexes with s-src and numerous other components, integrates signals from integrins, growth factor receptors, including those devoted to metabolic regulation such as insulin and IGF (Zhang and Hochwald, 2014; Zheng *et al.*, 2010). FAK is often over-expressed or hyperactive in tumors, and, as such, is a compelling drug target (Zheng *et al.*, 2010). The aberrant expression of activated trkB – a neurotrophin receptor – in certain tumor types of nonneuronal origin can activate survival signaling through PI3K/Akt as well (Geiger and Peeper, 2007). There is additionally, some evidence for a role of death receptors or their interactants such as FAS-associated death domain protein (FADD) or cellular FLICE (FADD-like IL-1 β -converting enzyme)-inhibitory protein (c-FLIP) in controlling anoikis in certain contexts (Mawji *et al.*, 2007).

Superimposed upon these signaling mechanisms is the anoikis-resistance afforded by the oncogenic epithelial-mesenchymal transition (EMT). EMT is thought to be the gene

expression program driving tumor metastasis (even though detecting EMT *in vivo* can be challenging due to its localized, transient, and reversible properties). Anoikis-resistance conferred by EMT recently has been reviewed elsewhere (Frisch *et al.*, 2013). The loss of E-cadherin plays an important role in this process, through the cytoskeletal protein ankyrin-G, altering the localization of critical transcription factors that control cell survival-related target genes (Kumar *et al.*, 2011). Moreover, the EMT-driven loss of epithelial cell polarity complexes also alters anoikis through the highly interconnected Hippo/YAP/TAZ, TGF- β , and Wnt signaling pathways at the transcriptional level (Frisch *et al.*, 2013). The serine-threonine kinase, integrin-linked kinase (ILK), also contributes to EMT and anoikis-resistance in some contexts, through Akt activation (White *et al.*, 2001).

Why Leave?

It is unclear what drives tumor cells to detach and manifest the complex program of metastasis. What was wrong with their primary environment? The intratumoral cellular metabolic challenges described below may offer an explanation, and, importantly, also, their response to these challenges may equip the tumor cell to metastasize, by enabling them to evade anoikis.

The Challenges of Being a Tumor Cell in its Primary Microenvironment

Oncogenic mutations or epigenetic changes in tumor cells, like some university administrators, politicians, or corporate CEOs, confer short-term growth advantages that, however, pose longer-term survival challenges to their institutions.

First, oncogenic changes often endanger cell survival directly, as is famously the case with c-myc over-expression, which, in the absence of compensatory defects in apoptosis genes, will induce apoptosis (as a complex with p14ARF protein) without fail. There are many indirect challenges as well. In light of the poor vasculature of tumors preceding the angiogenic switch, the tumor cell faces deficits in all manner of nutrients, including carbon sources and oxygen, i.e., hypoxia. Rapid proliferation exacerbates these problems, in that the tumor cell must accelerate anabolic processes to keep pace with growing need for lipids, amino acids, and nucleotides. In addition, the stromal microenvironment may be altered by inflammatory responses, as well as by novel factors secreted by the tumor cells, with indirect consequences for tumor growth. Indeed, different regions of the same tumor may face altogether different problems. A successful tumor must develop region-specific adaptations, overall cooperation among the different regions, and, first and foremost, a cellular exit strategy

that involves at least temporary cell survival, i.e., resistance to anoikis.

Interestingly, some of these very challenges provoke EMT, a stable transcriptional program that, by definition, encompasses anoikis-resistance. For example, in the absence of sufficient oxygen, the transcription factor subunit HIF1- α is stabilized. HIF1 induces antioxidant genes to help neutralize the reactive oxygen species (ROS) generation that, paradoxically, accompanies hypoxia (Buchheit *et al.*, 2014). HIF1 also down-regulates mitochondrial biogenesis and respiration. Moreover, HIF1 induces EMT through induction of the transcription factors Twist, Snail, and ZEB1 as well as stimulation of TGF- β signaling (Haase, 2009; Jiang *et al.*, 2011).

Tumor cells also tend to have high ROS levels, presumably a reflection of their hypoxic environment, and they also tend to over-express antioxidant enzymes, presumably to keep the ROS below a lethal threshold concentration.

The emerging paradigm is that tumor cells respond to microenvironmental signals specific to their location within a tumor by metabolic adaptations that may include EMT, invasion and metastasis, that, to be successful, must be accompanied by anoikis-resistance (Taddei *et al.*, 2013).

ROS and Anoikis – A Double-edged Sword?

ROS arise as reaction products from superoxide, generated chiefly from two sources. First, accompanying oxidative phosphorylation in mitochondria, single electrons are transferred from the electron transport chain to oxygen at a substantial frequency (0.1–2%). This process is altered in tumor cells or in normal cells in novel microenvironments, by multiple factors. Indirectly, the facilitated import of glucose and glutamine are the master controllers of most intracellular metabolism, oxidative phosphorylation being no exception. The glucose and glutamine transporter proteins are thus pivotal regulators of not only energy but also redox potential. Interestingly, the glucose transporter 3 (GLUT3) was recently shown to be up-regulated during EMT by the transcription factor ZEB1, with possible implications for anoikis control (Masin *et al.*, 2014). In addition, the activity of pyruvate dehydrogenase (PDH), the branch point between glycolysis (pyruvate) and TCA cycle (acetyl-CoA) is a major regulatory node for oxidative phosphorylation; the enzyme is inhibited by pyruvate dehydrogenase kinases (PDKs) that are, in turn, regulated by adhesion signals, oncogenic signaling pathways and HIF1 (in the case of PDK1) (Cairns *et al.*, 2007; Grassian *et al.*, 2011; Kamarajugadda *et al.*, 2012). Thus, mitochondrial oxidative phosphorylation is both regulated the level of glucose and the branch point enzyme, PDH.

Within the mitochondria, several factors modulate the amount of superoxide that is produced per unit of substrate. One such factor is a particular isoform derived from an alternative promoter of the SHC1 gene, p66Shc protein. P66Shc resides in the mitochondria and generates superoxides in a complex with cytochrome *c* (Giorgio *et al.*, 2005); the role of this protein in anoikis is discussed below. Superimposed on this, it should be noted that individual electron transport components differ greatly in their effects on ROS levels. For example, the experimental knockdown of certain complex I

components of the electron transport chain (GRIM-19, NDUFS-3) induces ROS, as does inhibition of complex I by the diabetes drug metformin (MF), whereas complex III components are required for hypoxia-induced ROS (Guzy *et al.*, 2005; Huang *et al.*, 2007; Skinner *et al.*, 2012). The mitochondrial uncoupling protein family members (UCP1-3) generally protect against ROS generation, thus certain knock-outs actually promote longevity (Mookerjee *et al.*, 2010). Hence, the electron transport chain should be viewed as a complex machine with both positive and negative components with respect to ROS generation. Bcl-2, in addition to controlling MOMP, also regulates redox balance through interactions with mitochondrial membrane proteins (Chong *et al.*, 2014).

Secondly, superoxides are generated at the membrane by the nicotinamide adenine dinucleotide phosphate (NADPH) oxidase (NOX) complexes, related to those of inflammatory/phagocytic cells. Intriguingly, the cytoskeletal regulatory small GTPase Rac and certain protein kinase C isoforms – both responding to integrins, other cell adhesion molecules, growth factor receptors and (in the case of PKC) direct caspase-mediated proteolytic cleavage – are the major activators of most NOX complexes (Lambeth and Neish, 2014). The non-canonical NOX4 is activated by RhoA, also linked to cytoskeleton and cell adhesion (Manickam *et al.*, 2014). The implications of these connections between NOX complexes and cytoskeleton have not been explored.

However they are generated, all cells have multiple mechanisms for neutralizing ROS efficiently. Superoxides generated as above are rapidly converted to hydrogen peroxide and then to water by superoxide dismutases (SODs) and catalase, respectively; other ROS can be neutralized by other mechanisms. The reduced glutathione (GSH) and thioredoxin pools play important roles in neutralizing ROS. Importantly, their maintenance in the reduced state requires the cofactor NADPH, mainly arising from pentose phosphate pathway (PPP) (Lu and Holmgren, 2014; Singh *et al.*, 2012).

ROS As a Mediator of Anoikis

When epithelial cells lose attachment to matrix, GLUTs are diverted from the membrane, causing a loss of PPP activity, and a sharp increase in cellular ROS, owing to the lack of NADPH (Schafer *et al.*, 2009). This cell death appears to be largely caspase-independent, and its exact mechanism is not clear. The authors propose that the burst of ROS subsequent to NADPH decline ultimately induces cell death due to low-energy charge, as ROS also inhibits utilization of the alternative energy source, fatty acids. Interpretation of these studies is somewhat complicated. They report on events that occur 24 h or later after detachment, perhaps reflecting effects of cell death rather than causes; for example, the NOX complex is activated by caspase cleavage of PKC, so the late ROS burst may be a marker of cell death. The lack of detectable caspase activity in this model may result from the oxidative inactivation of the caspases' cysteine active sites by ROS occurring well after the cell has committed to anoikis (Barboudi *et al.*, 2007).

Nevertheless, this study and others clearly demonstrate that ROS arising from glucose deprivation in suspended cells

contributes to cell death. In principle, at least three survival mechanisms may counterbalance this effect: (1) upon matrix detachment, AMP-activated protein kinase (AMPK) is activated (through the liver kinase B1 (STK11)), inhibiting acetyl-CoA carboxylase, and therefore saving precious NADPH (Ng *et al.*, 2012); (2) AMPK activates autophagy, producing metabolites and aiding the effort to restore energy charge (Dunlop and Tee, 2013); and (3) Tumor cells tend to express a splice variant of pyruvate kinase, PKM2, that is potentially inhibited by growth factor receptors such as fibroblast growth factor receptors (FGFR2) as well as by ROS. This inhibition causes a buildup of glycolytic intermediates that ultimately stimulates the PPP, helping to increase NADPH production (Vander Heiden *et al.*, 2011).

Looking further downstream ROS induces apoptosis at the mitochondrial level by peroxidation of the lipid that sequesters cytochrome *c*, cardiolipin, as well as inactivating anti-apoptotic Bcl-2 family proteins (Huttemann *et al.*, 2011). It is unclear, at present, what the source(s) of this ROS actually is (are). Consistent with the pro-apoptotic role of ROS in anoikis, nevertheless, the over-expression of antioxidant genes appears to be functionally significant, in that experimental over-expression of SOD and/or catalase in normal epithelial cells partially protected against detachment-induced death, while the knockdown of these enzymes in breast cancer cell lines promoted this death response (Davison *et al.*, 2013).

Another potential positive link between ROS and anoikis is in p66Shc, noted above to be a mitochondrial ROS-generating protein, although it demonstrates antioxidant properties when cytoplasmically localized in hematopoietic cells. The knockdown of p66Shc potentially protects against anoikis; mutational analysis indicated that the pro-anoikis effect of p66Shc was not mediated by p66Shc's domain for interaction with cytochrome *c*, the best-characterized mechanism for ROS generation (Ma *et al.*, 2007). Nevertheless, p66Shc was proposed to enforce anoikis through RhoA activation (Ma *et al.*, 2007). Speculatively, because RhoA can serve as a NOX4 activator in some contexts, it remains possible that p66Shc enforces anoikis through ROS (Manickam *et al.*, 2014). Interestingly, the suppression of p66Shc expression by the transcription factor Aiolos is a major contributor to lung cancer (Li *et al.*, 2014).

Further support for the tumor suppressive function of ROS – reflecting, in part, its pro-anoikis effect – also derives from the observations about the transcription factor Nrf2 (Jaramillo and Zhang, 2013). Nrf2 is over-expressed in numerous tumor types, correlating with poor prognosis. It transactivates the expression of several PPP genes, thus maintaining NADPH, and also promotes glutathione synthesis as well as SOD and catalase expression. Nrf2 gene expression is induced by various proliferation-stimulating oncogenes including K-ras, B-raf, and myc. In a fascinating demonstration of the power of metabolic alterations, fumarate hydratase mutations lead to papillary renal cell carcinoma by causing the accumulation of malate, which constitutively activates Nrf2 by binding to its sequestration partner, Keap1 (Kinch *et al.*, 2011). Speculatively, the constitutively activated Nrf2 may protect against anoikis via transactivation of antioxidant genes.

In summary, the evidence presented above strongly implicates ROS as promoters of anoikis that, as such, protect the

organism against tumor metastasis. The concepts in this section are summarized in Figure 1.

ROS as Cell Survival Factors

Most tumors are thought to have elevated ROS levels (Liou and Storz, 2010). These studies are based, however, on comparisons of tumor tissue versus normal tissue. At least in the case of breast cancer, such comparisons are perilous because the tumor tissue is predominantly epithelial whereas the normal tissue is primarily stromal, thus differential gene expression may be attributed to cell type rather than normal versus tumor status. Thus, this conclusion should not be accepted as dogma without further investigation.

Paradoxically, the complete pharmacologic ablation of ROS causes cell cycle arrest or, in some circumstances, may even trigger apoptosis (Giannoni *et al.*, 2008). This is not surprising, in light of several considerations. First, protein tyrosine phosphatases (PTPs) are uniformly and irreversibly inhibited by ROS, due to their uniquely sensitive active sites. PTP activity, therefore, is critically regulated by physiologic ROS levels, with too little leading to global PTP hyperactivation and, consequently, global inhibition of signaling pathways that depend upon tyrosine phosphorylation (Giannoni *et al.*, 2008). Additionally, the hub-like tyrosine kinase c-src is actually activated by ROS by relief of autoinhibition (Giannoni and Chiarugi, 2014). Among other problems, activated src can cause ligand-independent activation of epidermal growth factor receptor (EGFR) and its attendant cell survival signaling. Moreover, many growth factor receptors stimulate NOX complexes to produce superoxides, presumably to attenuate the opposing action of PTPs.

A recent study (Zhu *et al.*, 2011) has proposed that integrins stimulated by the matrix protein angiopoietin-like 4 (ANGPTL4) stimulate the NOX complex to generate ROS in attached cells, activating src and promoting cell survival of attached cells; exogenous ANGPTL4 was found to protect against anoikis in detached cells, accordingly. Neutralization of ROS promoted anoikis, by the criterion of short-term caspase activation after detachment. The apparent contradiction this poses may be reconciled by the specificity of the assay used in the ANGPTL4 study, which, unlike other studies measuring total oxidation potential, was specific for superoxides. Another discrepancy may arise from short-term caspase activity assays in this study versus long-term survival assays in other studies: the former may be more influenced by ROS-mediated caspase inhibition by oxidation.

Several reports have also ascribed an indirect protective role of ROS: its purported ability to induce EMT, through, in part, activation of at least two EMT-driving transcription factors, HIF1 and NF- κ B (Cannito *et al.*, 2008; Cichon and Radisky, 2014). For example, matrix metalloproteases-2 and -9 can cleave the ectodomain of E-cadherin, inducing the expression of a Rac1 isoform, Rac1b, that activates ROS generation through NOX complexes (Radisky *et al.*, 2005). In another example, knockdown of the mitochondrial complex I subunits generates mitochondrial ROS, as mentioned above, which leads the knockdown cell lines to eventually display an EMT phenotype. Persistent activation may in certain contexts lead

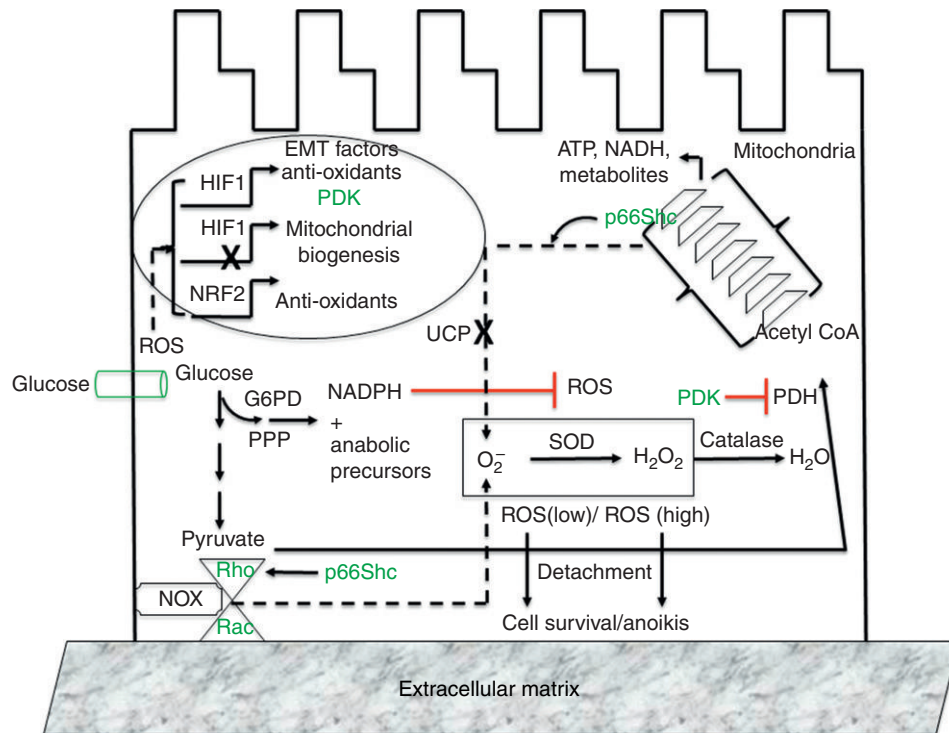


Figure 1 Control of anoikis through ROS. Factors that are known to be regulated by cell-matrix adhesion are in blue font. High levels of ROS are generated by cell-matrix detachment and commit the normal detached cell to die. Low levels of ROS (as compared with experimentally induced absence of ROS) may promote cell survival through Src activation and phosphatase inactivation (not shown in figure). ROS is generated in mitochondria during oxidative phosphorylation and also through the action of NADPH (NOX) complexes at the membrane. NOX complexes are usually activated by Rac-GTP, but NOX4 is activated by Rho-GTP. P66Shc may potentially promote ROS generation at mitochondria through an electron transfer reaction with cytochrome *c*, or at the membrane through activation of RhoA/NOX4. The Uncoupling Protein family (UCP) attenuates ROS production. The branch point between glycolysis and mitochondrial respiration is the enzyme PDH, which is negatively regulated by the adhesion-controlled kinases PDKs; thus, PDK may affect the level of ROS generated by mitochondria. In normal, attached cells, ROS levels are kept in check through growth factor-activated glucose transport, which, in addition to providing energy charge, also maintains high levels of NADPH (through the PPP); NADPH is a cofactor for reduced glutathione synthesis. At the transcriptional level, ROS activates at least two transcription factors, HIF1 and Nrf2. HIF1 induces antioxidant and EMT-related gene expression, while repressing mitochondrial biogenesis. Nrf2 also induces a battery of antioxidant genes. Both factors contribute to tumorigenesis and may do so through evasion of anoikis.

to an EMT phenotype that includes transcriptionally programmed anoikis-resistance. It is unclear from these reports, however, whether ROS directly induced a stable EMT phenotype, as opposed to selecting for EMT cells that are resistant to the toxic effects of ROS. Both models have important implications, however, for tumor evolution.

What Otto Warburg Said, How It Was Misinterpreted, and the Idea of Metabolic Coupling in Tumors

Otto Warburg stated that ‘tumor cells’ (an impossible oversimplification) continued to rely on glycolysis even under aerobic conditions that would normally suppress glycolysis. This generalization has not, however, stood the test of time. Moreover, it has been widely misinterpreted to mean that tumor cells uniformly suppress oxidative phosphorylation. As clarified in recent reviews, tumor cells vary dramatically in their relative dependence upon the two metabolic pathways (Dang, 2010; Koppenol *et al.*, 2011). In fact, a single tumor possesses great variation from one region to another, or between primary microenvironment, circulating tumor cell

environment and metastatic niche (Sonveaux *et al.*, 2008). Thus, a single tumor may contain highly oxidative as well as glycolytic regions. The profile of individual tumor cells may well depend upon their availability of oxygen concentration, nutrients, factors secreted by their neighboring stromal cells, and degree of infiltration with immune cells. In fact, the success of an individual tumor cell may well depend on its metabolic plasticity rather than its particular profile at a given time point (Taddei *et al.*, 2013). In this connection, an intimate connection between tumor and stroma, known as ‘metabolic coupling’ (Martinez-Outschoorn *et al.*, 2011) posits that cancer associated fibroblasts – under the influence of secreted factors from the tumor cells – generally operate glycolytically and therefore produce a large quantity of pyruvate and lactate. These metabolites are transported by the monocarboxylate transporter (MCT) family proteins that tend to be over-expressed in aggressive tumors, where they are used primarily in the TCA cycle to produce both metabolites for anabolic processes and energy through oxidative phosphorylation.

What are the implications of this important, emerging model for anoikis regulation? Tumor cells coupled in this way with their neighboring stroma are essentially addicted to

lactate and related monocarboxylic acids as fuel. They are unlikely to utilize the PPP efficiently, thus weakening their NADPH defense and increasing ROS. One of two scenarios may be envisioned. First, EMT is more likely to occur in the presence of increased ROS, via activation of, for example, HIF1 and NF- κ B. EMT will promote both increased invasion and anoikis-resistance, at the transcriptional level. Alternatively, this micro-environment will provide selective pressure for the outgrowth of cells that have a strong antioxidant enzyme-based defense, for example, expression of SOD, catalase. Over-expression of these enzymes protects these cells against anoikis. Thus, metabolic coupling is predicted to drive tumor evolution to generate anoikis-resistant and potentially metastatic cells. This view is supported by emerging concepts that link tumor cell reliance on mitochondrial oxidative phosphorylation with anoikis-resistance. Tumor cells that use lactate as a carbon source, and are therefore obligated to use oxidative phosphorylation, are poorly adapted to the hypoxic environment present in the primary tumor. This provides a stimulus for them to metastasize via transient EMT, suspension of anoikis and increased invasiveness. Indeed, metastatic breast cancer cells recovered from lymph nodes or brain frequently have high mitochondrial number and mitochondrial membrane potential (Amoedo *et al.*, 2014; Taddei *et al.*, 2013).

Summary

The metabolic control of anoikis is clearly established from cell culture models; *in vivo* tumor models are much more difficult to analyze from this perspective, because this would require tracing of an individual cell's metabolic profile *in vivo*, and then following that cell's ability to survive and stay in the primary tumor versus its ability to metastasize – not an easy experimental protocol. As a result, most current *in vivo* data correlate overall tumor metabolic profiles with overall tumor growth and/or chemotherapy resistance, rather an obscure way to infer much about anoikis sensitivity. Nevertheless, tumor cells responses to ROS and other metabolites clearly play a critical role in whether they will die by anoikis, and, if so, how long that process will take, i.e., will it occur before seeding a distant site, as required to prevent metastasis? Technical advances will hopefully elucidate these issues more clearly.

See also: Cell Division/Death: Apoptosis: Apoptosis

References

- Amoedo, N.D., Rodrigues, M.F., Rumjanek, F.D., 2014. Mitochondria: Are mitochondria accessory to metastasis? *International Journal of Biochemistry & Cell Biology* 51, 53–57.
- Barboui, A., Amorgianiotis, C., Kolettas, E., Kanavaros, P., Galaris, D., 2007. Hydrogen peroxide inhibits caspase-dependent apoptosis by inactivating procaspase-9 in an iron-dependent manner. *Free Radical Biology & Medicine* 43, 1377–1387.
- Buchheit, C.L., Weigel, K.J., Schafer, Z.T., 2014. Cancer cell survival during detachment from the ECM: Multiple barriers to tumour progression. *Nature Reviews Cancer* 14, 632–641.
- Cairns, R.A., Papandreou, I., Sutphin, P.D., Denko, N.C., 2007. Metabolic targeting of hypoxia and HIF1 in solid tumors can enhance cytotoxic chemotherapy. *Proceedings of the National Academy of Sciences of the United States of America* 104, 9445–9450.
- Cannito, S., Novo, E., Compagnone, A., *et al.*, 2008. Redox mechanisms switch on hypoxia-dependent epithelial-mesenchymal transition in cancer cells. *Carcinogenesis* 29, 2267–2278.
- Charpentier, M., Martin, S., 2013. Interplay of stem cell characteristics, EMT, and microtentacles in circulating breast tumor cells. *Cancers* 5, 1545–1565.
- Chong, S.J., Low, I.C., Pervaiz, S., 2014. Mitochondrial ROS and involvement of Bcl-2 as a mitochondrial ROS regulator. *Mitochondrion* 19 (Pt A), 39–48.
- Cichon, M.A., Radisky, D.C., 2014. ROS-induced epithelial-mesenchymal transition in mammary epithelial cells is mediated by NF- κ B-dependent activation of Snail. *Oncotarget* 5, 2827–2838.
- Dang, C.V., 2010. Rethinking the Warburg Effect with Myc micromanaging glutamine metabolism. *Cancer Research* 70, 859–862.
- Davison, C.A., Durbin, S.M., Thau, M.R., *et al.*, 2013. Antioxidant enzymes mediate survival of breast cancer cells deprived of extracellular matrix. *Cancer Research* 73, 3704–3715.
- Dunlop, E.A., Tee, A.R., 2013. The kinase triad, AMPK, mTORC1 and ULK1, maintains energy and nutrient homeostasis. *Biochemical Society Transactions* 41, 939–943.
- Frisch, S.M., Francis, H., 1994. Disruption of epithelial cell-matrix interactions induces apoptosis. *Journal of Cell Biology* 124, 619–626.
- Frisch, S.M., Schaller, M., Cieply, B., 2013. Mechanisms that link the oncogenic epithelial-mesenchymal transition to suppression of anoikis. *Journal of Cell Science* 126, 21–29.
- Frisch, S.M., Screaton, R.A., 2001. Anoikis mechanisms. *Current Opinion in Cell Biology* 13, 555–562.
- Geiger, T.R., Peeper, D.S., 2007. Critical role for TrkB kinase function in anoikis suppression, tumorigenesis, and metastasis. *Cancer Research* 67, 6221–6229.
- Giannoni, E., Buricchi, F., Grimaldi, G., *et al.*, 2008. Redox regulation of anoikis: Reactive oxygen species as essential mediators of cell survival. *Cell Death and Differentiation* 15, 867–878.
- Giannoni, E., Chiarugi, P., 2014. Redox circuitries driving Src regulation. *Antioxidants & Redox Signaling* 20, 2011–2025.
- Giorgio, M., Migliaccio, E., Orsini, F., *et al.*, 2005. Electron transfer between cytochrome c and p66Shc generates reactive oxygen species that trigger mitochondrial apoptosis. *Cell* 122, 221–233.
- Grassian, A.R., Colloff, J.L., Brugge, J.S., 2011. Extracellular matrix regulation of metabolism and implications for tumorigenesis. *Cold Spring Harbor Symposia on Quantitative Biology* 76, 313–324.
- Guzy, R.D., Hoyos, B., Robin, E., *et al.*, 2005. Mitochondrial complex III is required for hypoxia-induced ROS production and cellular oxygen sensing. *Cell Metabolism* 1, 401–408.
- Haase, V.H., 2009. Oxygen regulates epithelial-to-mesenchymal transition: Insights into molecular mechanisms and relevance to disease. *Kidney International* 76, 492–499.
- Huang, G., Chen, Y., Lu, H., Cao, X., 2007. Coupling mitochondrial respiratory chain to cell death: An essential role of mitochondrial complex I in the interferon-beta and retinoic acid-induced cancer cell death. *Cell Death and Differentiation* 14, 327–337.
- Huttemann, M., Pecina, P., Rainbolt, M., *et al.*, 2011. The multiple functions of cytochrome c and their regulation in life and death decisions of the mammalian cell: From respiration to apoptosis. *Mitochondrion* 11, 369–381.
- Jaramillo, M.C., Zhang, D.D., 2013. The emerging role of the Nrf2-Keap1 signaling pathway in cancer. *Genes & Development* 27, 2179–2191.
- Jiang, J., Tang, Y.L., Liang, X.H., 2011. EMT: A new vision of hypoxia promoting cancer progression. *Cancer Biology & Therapy* 11, 714–723.
- Kamarajugadda, S., Stemborski, L., Cai, Q., *et al.*, 2012. Glucose oxidation modulates anoikis and tumor metastasis. *Molecular and Cellular Biology* 32, 1893–1907.
- Kinch, L., Grishin, N.V., Brugarolas, J., 2011. Succination of Keap1 and activation of Nrf2-dependent antioxidant pathways in FH-deficient papillary renal cell carcinoma type 2. *Cancer Cell* 20, 418–420.
- Koppenol, W.H., Bounds, P.L., Dang, C.V., 2011. Otto Warburg's contributions to current concepts of cancer metabolism. *Nature Reviews Cancer* 11, 325–337.
- Kumar, S., Park, S.H., Cieply, B., *et al.*, 2011. A pathway for the control of anoikis sensitivity by E-cadherin and epithelial-to-mesenchymal transition. *Molecular and Cellular Biology* 31, 4036–4051.
- Lambeth, J.D., Neish, A.S., 2014. Nox enzymes and new thinking on reactive oxygen: A double-edged sword revisited. *Annual Review of Pathology* 9, 119–145.

- Li, X., Xu, Z., Du, W., *et al.*, 2014. Aiolos promotes anchorage independence by silencing p66Shc transcription in cancer cells. *Cancer Cell* 25, 575–589.
- Liou, G.Y., Storz, P., 2010. Reactive oxygen species in cancer. *Free Radical Research* 44, 479–496.
- Lu, J., Holmgren, A., 2014. The thioredoxin antioxidant system. *Free radical Biology & Medicine* 66, 75–87.
- Ma, Z., Myers, D.P., Wu, R.F., Nwariaku, F.E., Terada, L.S., 2007. p66Shc mediates anoikis through RhoA. *Journal of Cell Biology* 179, 23–31.
- Manickam, N., Patel, M., Griendling, K.K., Gorin, Y., Barnes, J.L., 2014. RhoA/Rho kinase mediates TGF-beta1-induced kidney myofibroblast activation through Poldip2/Nox4-derived reactive oxygen species. *American Journal of Physiology Renal Physiology* 307, F159–F171.
- Martinez-Outschoorn, U.E., Pavlides, S., Howell, A., *et al.*, 2011. Stromal-epithelial metabolic coupling in cancer: Integrating autophagy and metabolism in the tumor microenvironment. *International Journal of Biochemistry & Cell Biology* 43, 1045–1051.
- Masin, M., Vazquez, J., Rossi, S., *et al.*, 2014. GLUT3 is induced during epithelial-mesenchymal transition and promotes tumor cell proliferation in non-small cell lung cancer. *Cancer & Metabolism* 2, 11.
- Mawji, I.A., Simpson, C.D., Hurren, R., *et al.*, 2007. Critical role for Fas-associated death domain-like interleukin-1-converting enzyme-like inhibitory protein in anoikis resistance and distant tumor formation. *Journal of the National Cancer Institute* 99, 811–822.
- Mookerjee, S.A., Divakaruni, A.S., Jastroch, M., Brand, M.D., 2010. Mitochondrial uncoupling and lifespan. *Mechanisms of Ageing and Development* 131, 463–472.
- Ng, T.L., Leprieux, G., Robertson, M.D., *et al.*, 2012. The AMPK stress response pathway mediates anoikis resistance through inhibition of mTOR and suppression of protein synthesis. *Cell Death and Differentiation* 19, 501–510.
- Paoli, P., Giannoni, E., Chiarugi, P., 2013. Anoikis molecular pathways and its role in cancer progression. *Biochimica et Biophysica Acta* 1833, 3481–3498.
- Radisky, D.C., Levy, D.D., Littlepage, L.E., *et al.*, 2005. Rac1b and reactive oxygen species mediate MMP-3-induced EMT and genomic instability. *Nature* 436, 123–127.
- Schafer, Z.T., Grassian, A.R., Song, L., *et al.*, 2009. Antioxidant and oncogene rescue of metabolic defects caused by loss of matrix attachment. *Nature* 461, 109–113.
- Schmelzle, T., Mailloux, A.A., Overholtzer, M., *et al.*, 2007. Functional role and oncogene-regulated expression of the BH3-only factor Bim in mammary epithelial anoikis and morphogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 104, 3787–3792.
- Simpson, C.D., Anyiwe, K., Schimmer, A.D., 2008. Anoikis resistance and tumor metastasis. *Cancer Letters* 272, 177–185.
- Singh, S., Khan, A.R., Gupta, A.K., 2012. Role of glutathione in cancer pathophysiology and therapeutic interventions. *Journal of Experimental Therapeutics & Oncology* 9, 303–316.
- Skinner, H.D., Sandulache, V.C., Ow, T.J., *et al.*, 2012. TP53 disruptive mutations lead to head and neck cancer treatment failure through inhibition of radiation-induced senescence. *Clinical Cancer Research: An official Journal of the American Association for Cancer Research* 18, 290–300.
- Sonveaux, P., Vegran, F., Schroeder, T., *et al.*, 2008. Targeting lactate-fueled respiration selectively kills hypoxic tumor cells in mice. *Journal of Clinical Investigation* 118, 3930–3942.
- Taddei, M.L., Giannoni, E., Comito, G., Chiarugi, P., 2013. Microenvironment and tumor cell plasticity: An easy way out. *Cancer Letters* 341, 80–96.
- Vander Heiden, M.G., Lunt, S.Y., Dayton, T.L., *et al.*, 2011. Metabolic pathway alterations that support cell proliferation. *Cold Spring Harbor Symposia on Quantitative Biology* 76, 325–334.
- White, D.E., Cardiff, R.D., Dedhar, S., Muller, W.J., 2001. Mammary epithelial-specific expression of the integrin-linked kinase (ILK) results in the induction of mammary gland hyperplasias and tumors in transgenic mice. *Oncogene* 20, 7064–7072.
- Woods, N.T., Yamaguchi, H., Lee, F.Y., Bhalla, K.N., Wang, H.G., 2007. Anoikis, initiated by Mcl-1 degradation and Bim induction, is deregulated during oncogenesis. *Cancer Research* 67, 10744–10752.
- Zhang, J., Hochwald, S.N., 2014. The role of FAK in tumor metabolism and therapy. *Pharmacology & Therapeutics* 142, 154–163.
- Zheng, D., Golubovskaya, V., Kurenova, E., *et al.*, 2010. A novel strategy to inhibit FAK and IGF-1R decreases growth of pancreatic cancer xenografts. *Molecular Carcinogenesis* 49, 200–209.
- Zhu, P., Tan, M.J., Huang, R.L., *et al.*, 2011. Angiotensin-like 4 protein elevates the prosurvival intracellular $O_2(-):H_2O_2$ ratio and confers anoikis resistance to tumors. *Cancer Cell* 19, 401–415.

Mitochondria in Cell Death Regulation

I Vitale and G Manic, Regina Elena National Cancer Institute, Rome, Italy

G Kroemer, INSERM U1138, Paris, France; Hôpital Européen Georges Pompidou, Paris, France; Gustave Roussy, Villejuif, France; Centre de Recherche des Cordeliers, Paris, France; and Université Paris Descartes/Paris V, Sorbonne Paris Cité, Paris, France

L Galluzzi, INSERM U1138, Paris, France; Gustave Roussy, Villejuif, France; Centre de Recherche des Cordeliers, Paris, France; and Université Paris Descartes/Paris V – Sorbonne Paris Cité, Paris, France

© 2016 Elsevier Inc. All rights reserved.

Glossary

Caspases Family of cysteine-dependent aspartate-directed proteases involved in apoptosis, differentiation, and inflammatory responses. In the context of apoptosis, these proteases often operate as part of self-amplificatory catalytic cascades resulting in the massive activation of caspase-3, -6, and -7.

Glutamate-ammonia ligase (GLUL) Enzyme that plays an essential role in the metabolism of nitrogen by catalyzing the reversible condensation of glutamate and ammonia to form glutamine.

Glutamate dehydrogenase 1 (GLUD1) Enzyme of the mitochondrial matrix that catalyzes the oxidative deamination of glutamate to α -ketoglutarate and ammonia.

Harlequin mice Experimental mouse strain that expresses reduced amounts of AIF due to a hypomorphic, X-linked recessive mutation in *Aifm1*.

Mitochondrial transmembrane potential ($\Delta\psi_m$) In physiological conditions, respiratory complexes actively extrude protons from the mitochondrial matrix, resulting in an electrochemical gradient that sustains the catalytic activity of the F_1F_0 -ATPase. This corresponds to an electrical potential of approximately -130 mV.

Mitophagy Selective removal of damaged or supernumerary mitochondria by the autophagic machinery.

Reactive oxygen species (ROS) Highly reactive chemical species containing oxygen. The most common ROS of mitochondrial derivation are superoxide anion (O_2^-) and hydrogen peroxide (H_2O_2).

Tumor protein p53 (TP53) Major oncosuppressor protein involved in the control of cell cycle progression, cell death, intermediate metabolism, and redox balance.

Introduction

Mitochondria

Mitochondria – from the ancient Greek *μίτος*/mitos (thread) and *χόνδρος*/chondros (grain) – are double-membraned cytoplasmic organelles that can be found in virtually all eukaryotic cells (Galluzzi *et al.*, 2010, 2012b). Human mitochondria contain multiple copies of a double-stranded, circular DNA molecule of approximately 16.6 kbp, coding for 22 tRNAs, 2 rRNAs, and 13 proteins (Anderson *et al.*, 1981). Notably, all mitochondrial DNA (mtDNA)-encoded macromolecules are employed *in loco* and most mitochondrial proteins are coded by nuclear genes (Elstner *et al.*, 2009). This implies the existence of refined mechanisms of interorganellar crosstalk that ensure the maintenance of cellular homeostasis in physiological conditions as well as in the course of adaptive stress responses (Galluzzi *et al.*, 2012a; Green *et al.*, 2011; Schmidt *et al.*, 2010).

Normally, mitochondria constitute the most prominent source of intracellular adenosine triphosphate (ATP) and the vast majority of eukaryotic cells critically depends on these organelles for survival (Galluzzi *et al.*, 2012b). This central bioenergetic function, i.e., the condensation of adenosine diphosphate (ADP) with inorganic phosphate to generate ATP, is mediated by the so-called F_1F_0 -ATPase, a multiprotein complex embedded in the inner mitochondrial membrane (IMM) (Devenish *et al.*, 2008; Bonora *et al.*, 2015). To this aim, the F_1F_0 -ATPase dissipates in a controlled fashion the

electrochemical gradient built across the IMM by the sequential activity of four other supramolecular entities embedded in the IMM (respiratory complexes I–IV) and two semi-soluble electron shuttles, i.e., coenzyme Q_{10} and cytochrome *c* (CYTC) (Okuno *et al.*, 2011; Stock *et al.*, 1999; Efremov and Sazanov, 2011). This cascade of redox reactions, which is generally known as oxidative phosphorylation (OXPHOS), requires reduced nicotinamide adenine dinucleotide (NADH), operating as an electron donor, and molecular oxygen, acting as a terminal electron acceptor (which results in its reduction into H_2O) (Mitchell, 1961). Mitochondrial NADH is generated by the tricarboxylic acid (TCA) cycle (also known as Krebs cycle), which in turn feeds either on cytosolic pyruvate of glycolytic derivation or on mitochondrial acetyl-CoA produced by the oxidation of fatty acids (β oxidation). Actually, the Krebs cycle is connected to a plethora of catabolic pathways, including those that mediate the degradation of amino acids and nucleosides (Watford, 1991).

Besides constituting the cell's ‘power plants,’ mitochondria not only host several anabolic reactions (including multiple steps of the biosynthesis of heme and iron–sulfur clusters) (Ajioka *et al.*, 2006; Stehling and Lill, 2013), but are also involved in several signal transduction cascades, in particular in the context of stress responses (Galluzzi *et al.*, 2012a; Tait and Green, 2012; Nunnari and Suomalainen, 2012). At least in part, this reflects (1) the intimate crosstalk through which mitochondria and the endoplasmic reticulum (ER) control intracellular Ca^{2+} stores (Kornmann, 2013; Rowland and Voeltz, 2012); (2) the capacity of mitochondria to produce

reactive oxygen species (ROS) (Sena and Chandel, 2012; Murphy, 2009); (3) the strategic position occupied by mitochondria in the control of multiple signaling pathways leading to cell death (Galluzzi *et al.*, 2009b, 2012a; Vandenabeele *et al.*, 2010; Tait and Green, 2010; Kroemer *et al.*, 2007; Vitale *et al.*, 2011), as discussed below.

Regulated Cell Death

In response to harsh conditions (e.g., steep changes in temperature and extreme ionic strengths), cells die in a virtually instantaneous and uncontrollable manner (Kroemer *et al.*, 2009). Such accidental instances of cell death occur very rarely (if ever) in physiopathological scenarios, as these are characterized by relatively mild homeostatic perturbations. In this setting, cells initially respond to alterations of the intracellular or extracellular microenvironment by activating one or more cytoprotective mechanisms aimed at restoring homeostasis (Portt *et al.*, 2011; Green *et al.*, 2011; Kroemer *et al.*, 2010; Fulda *et al.*, 2010; Tait and Green, 2012). If this attempt fails, as it occurs when the triggering stimulus is excessively intense or durable, the cell actively initiates a genetically encoded (hence regulated) program that leads to its own demise (Galluzzi *et al.*, 2007, 2012d). Of note, similar mechanisms of cellular 'suicide' are activated throughout embryonic and post-embryonic development, providing a critical contribution to morphogenesis, as well as continuously in adult life, underlying the maintenance of tissue homeostasis (Taylor *et al.*, 2008; Jacobson *et al.*, 1997; Claveria *et al.*, 2013). These purely physiological instances of regulated cell death are generally referred to as 'programmed,' as they occur independently of exogenous triggers. Of note, regulated cell death (be it programmed or not) can manifest with an incredibly rich panel of morphologies (Vandenabeele *et al.*, 2010; Kroemer and Levine, 2008; Kroemer *et al.*, 2009; Taylor *et al.*, 2008).

For a long time, the phenotypic features of dying cells have been employed to classify the cellular demise into three classes: (1) apoptosis, manifesting with chromatin condensation, nuclear breakdown, cell shrinkage, blebbing, and the formation of discrete, plasma membrane-bound apoptotic corpses; (2) autophagy, involving the massive accumulation of cytoplasmic vacuoles, and (3) necrosis, being defined by the lack of apoptotic and autophagic characteristics (Vandenabeele *et al.*, 2010; Kroemer and Levine, 2008; Kroemer *et al.*, 2009; Taylor *et al.*, 2008). This classification originated from the poor methodological arsenal that was available for the study of cell death until the 1990s (Kepp *et al.*, 2011; Galluzzi *et al.*, 2009a), and fails to take into account important functional aspects. Thus, definitions based on measurable biochemical/molecular parameters are emerging (Galluzzi *et al.*, 2012d). Such a novel taxonomy encompasses (1) two forms of cell death that manifest with apoptotic features, both of which critically rely on caspase activation; and (2) at least three distinct forms of cellular demise exhibiting necrotic features, including necroptosis, which requires receptor-interacting protein kinase 3 (RIPK3) and mixed lineage kinase domain-like (MLKL), mitochondrial permeability transition (MPT)-dependent regulated necrosis, which obligatorily involves peptidylprolyl isomerase F (PPIF, best known as cyclophilin D,

CYPD), and parthanatos, a poly (ADP-ribose) polymerase 1 (PARP1)-dependent variant of cell death (Galluzzi *et al.*, 2012d, 2014a). Moreover, autophagic cell death has been re-defined as a type of cellular demise that critically depends on the molecular machinery for autophagy (Liu *et al.*, 2013; Galluzzi *et al.*, 2012d; Kroemer and Levine, 2008).

Here, we review the signal transduction cascades whereby mitochondria control several instances of cell death, including intrinsic apoptosis, extrinsic apoptosis, MPT-dependent regulated necrosis, and parthanatos. Moreover, we will present the main molecular mechanisms whereby mitochondria get permeabilized to seal the cell death, i.e., mitochondrial outer membrane permeabilization (MOMP) and MPT.

Mitochondrial Control of Regulated Cell Death

Mitochondrial Control of Apoptosis

Intrinsic apoptosis

Intrinsic (also known as mitochondrial) apoptosis can be initiated in response to perturbations of intracellular homeostasis as different as DNA damage, oxidative stress, cytosolic Ca^{2+} overload, dysregulated protein folding, the rupture of lysosomal membranes, and cytoskeletal alterations (Kroemer *et al.*, 2007). Mitochondria play a key regulatory role in this setting as they integrate anti- and pro-apoptotic signals to seal the cellular fate (Tait and Green, 2010; Taylor *et al.*, 2008; Kroemer *et al.*, 2007; Galluzzi *et al.*, 2012a). Thus, while in the course of adaptive responses to stress mitochondrial integrity is actively preserved via multiple mechanisms (Gomes *et al.*, 2011; Rong and Distelhorst, 2008; Rambold *et al.*, 2011), when pro-apoptotic signals predominate over their anti-apoptotic counterparts most mitochondria lose their structural (and hence functional) integrity via either of two intimately intertwined mechanisms, MOMP and MPT (see below) (Brenner and Moulin, 2012; Brenner and Grimm, 2006; Tait and Green, 2010; Kroemer *et al.*, 2007). This is catastrophic for at least two reasons. First, it coincides with the dissipation of the mitochondrial transmembrane potential ($\Delta\psi_m$), which in turn abrogates mitochondrial ATP synthesis (Mitchell, 1961; Okuno *et al.*, 2011; Rambold *et al.*, 2011), blocks the import of several proteins encoded by the nuclear genome (Neupert and Herrmann, 2007; Jin *et al.*, 2010), and impairs mitochondrial ionic homeostasis (Galluzzi *et al.*, 2012d; Kroemer *et al.*, 2007). Second, it allows for the release of potentially cytotoxic proteins that are normally sequestered within the mitochondrial intermembrane space (IMS), including CYTC (Li *et al.*, 1997; Kluck *et al.*, 1997; Liu *et al.*, 1996), apoptosis-inducing factor, mitochondrion-associated, 1 (AIFM1, best known as AIF) (Joza *et al.*, 2001; Susin *et al.*, 1999), endonuclease G (ENDOG) (Van Loo *et al.*, 2001; Li *et al.*, 2001; Parrish *et al.*, 2001), the inhibitor of apoptosis protein (IAP)-binding mitochondrial protein DIABLO (also known as SMAC) (Verhagen *et al.*, 2000; Du *et al.*, 2000), and HtrA serine peptidase 2 (HTRA2, also known as OMI) (Srinivasula *et al.*, 2003; Yang *et al.*, 2003; Martins *et al.*, 2002).

For illustrative purposes, these proteins can be classified into caspase-dependent and caspase-independent cell death effectors (Galluzzi *et al.*, 2012d; Kroemer and Martin, 2005).

As the current tendency is to view apoptosis as a caspase-dependent instance of cell death (irrespective of its morphological manifestations) (Galluzzi *et al.*, 2012d), we will here focus on IMS proteins that promote caspase activation upon release into the cytosol. Of note, most of these proteins (as well as nearly all the upstream regulators and downstream effectors of apoptosis) are evolutionarily conserved (Zhivotovskiy *et al.*, 2009; Arnoult *et al.*, 2001; Yuan, 1996) and mediate key physiological functions (Galluzzi *et al.*, 2012c; Galluzzi *et al.*, 2008a), perhaps suggesting that evolution has assembled the cell death machinery based on available components rather than building new ones.

CYTC, which normally shuttles electrons between respiratory complex III and IV (Ow *et al.*, 2008), is the prototype of mitochondrial caspase activators (Garrido *et al.*, 2006). Indeed, in the presence of deoxyATP, cytosolic CYTC actively assembles an apoptotic peptidase activating factor 1 (APAF1)-containing supramolecular platform that recruits pro-caspase-9 molecules and favors their proximity-induced activation (Twiddy *et al.*, 2004; Cain, 2003; Rodriguez and Lazebnik, 1999; Riedl and Salvesen, 2007). In turn, caspase-9 is capable of setting off an auto-amplificatory cascade (Mondragon *et al.*, 2009; Garrido *et al.*, 2006) resulting in the massive activation of caspase-3, the central executioner of apoptosis (Taylor *et al.*, 2008; Thornberry and Lazebnik, 1998; Janicke *et al.*, 1998), as well as caspase-6 and -7 (at least in some settings) (Walsh *et al.*, 2008; Lakhani *et al.*, 2006; Leblanc *et al.*, 1999). Although the deletion of the CYTC-coding gene is embryonic lethal, a refined genetic strategy has been successfully implemented to study the lethal functions of this protein. This approach confirmed previous results indicating that the apoptotic machinery is essential for the normal development of the central nervous system (Yoshida *et al.*, 1998; Kuida *et al.*, 1996, 1998) yet demonstrated a differential requirement for CYTC and APAF1 in some instances of apoptosis (Green, 2005; Hao *et al.*, 2005).

At odds with CYTC, DIABLO, and HTRA2 operate as indirect activators of caspases, at least in part as they sequester, thereby inhibiting, several members of the baculoviral IAP repeat containing (BIRC) protein family (Verhagen *et al.*, 2007; Wilkinson *et al.*, 2004; Martins *et al.*, 2002; Wu *et al.*, 2000; Verhagen *et al.*, 2000; Du *et al.*, 2000). In addition, HTRA2 catalyzes the proteolytic cleavage of BIRC family members, hence inactivating them in an irreversible fashion (Yang *et al.*, 2003). Of note, while initially BIRC-like proteins were viewed as mere inhibitors of caspases (Scott *et al.*, 2005; Tenev *et al.*, 2005; Shiozaki *et al.*, 2003; Srinivasula *et al.*, 2001), it is now clear that these factors mediate multipronged cytoprotective effects, regulating not only the execution but also the initiation of apoptosis by virtue of their E3 ubiquitin ligase activity (Yang *et al.*, 2014; Hu and Yang, 2003; Li *et al.*, 2002).

Extrinsic apoptosis

Extrinsic apoptosis is invariably initiated by plasma membrane receptors in response to perturbations of the extracellular microenvironment (Galluzzi *et al.*, 2012d; Kroemer *et al.*, 2007). For illustrative purposes, extracellular pro-apoptotic cues can be subdivided into two mutually exclusive groups: 'active' and 'passive' (Mehlen and Bredesen, 2011; Wajant, 2002). The former are conveyed by the so-called 'death

receptors,' which are activated by soluble or plasma membrane-bound agonists (Schutze *et al.*, 2008; Wajant, 2002; Guicciardi and Gores, 2009; Locksley *et al.*, 2001). Conversely, the latter are transduced by the so-called 'dependence receptors,' which emit an apoptotic signal in the absence of their ligands (Mehlen and Bredesen, 2011; Mille *et al.*, 2009; Tauszig-Delamasure *et al.*, 2007). A detailed description of the molecular events taking place at the intracellular tails of death and dependence receptors upon activation goes beyond the scope of this article and can be found in references Schutze *et al.* (2008), Strasser *et al.* (2009), and Mehlen and Bredesen (2011).

In lymphocytes (exemplifying the so-called 'Type I cells'), death receptors such as FAS (also known as CD95) trigger apoptosis via a caspase-8 (or -10)-dependent signal transduction cascade that does not require mitochondria (Ferreira *et al.*, 2012; Scaffidi *et al.*, 1998). Conversely, in hepatocytes (a prototype of 'Type II cells'), FAS is unable to dispatch an efficient pro-apoptotic signal in the absence of MOMP, which is triggered by the caspase-8-dependent cleavage of the BH3-protein BH3 interacting domain death agonist (BID, see below) (Yin *et al.*, 1999; Li *et al.*, 1998; Luo *et al.*, 1998). Of note, the expression levels of the BIRC family member X-linked inhibitor of apoptosis (XIAP) appear to determine the propensity of cells to behave as Type I or Type II cells in response to death receptor activation (Jost *et al.*, 2009). Thus, as Type II cells contain high levels of XIAP, they undergo apoptosis in response to death receptor agonists only when DIABLO and HTRA2 are released by mitochondria upon MOMP (Jost *et al.*, 2009; Kaufmann *et al.*, 2012). This said, the formal distinction between Type I and II extrinsic apoptosis in terms of mitochondrial involvement may be less clear than what thought until recently (Meng *et al.*, 2011; Maas *et al.*, 2010).

The signal transduction cascades linking dependence receptors to apoptosis are heterogeneous and generally do not involve mitochondria (Mehlen and Bredesen, 2011). For instance, while unoccupied unc-5 homolog B (UN5B) triggers apoptosis via a protein phosphatase 2A (PP2A)- and death-associated protein kinase 1 (DAPK1)-dependent mechanism (Guenebeaud *et al.*, 2010), the absence of sonic hedgehog (SHH) endows patched 1 (PTCH1) with the ability to recruit a caspase-9-activatory platform involving the adaptors DRAL and TUCAN (Mille *et al.*, 2009). As an exception, the withdrawal of neurotrophin 3, one of the main ligands of neurotrophic tyrosine kinase, receptor, type 3 (NTRK3, best known as TRKC), results in the generation of a TRKC fragment that is shuttled to mitochondria by negative elongation factor complex member B (NELFB, best known as COBRA1) to promote MOMP (Tauszig-Delamasure *et al.*, 2007; Ichim *et al.*, 2013).

Taken together, these observations indicate that mitochondria play a critical role in the regulation of both intrinsic and extrinsic instances of apoptosis (Figure 1).

Mitochondrial Control of Regulated Necrosis

Initially, mitochondria were thought to play a critical role in all main instances of regulated necrosis described to date, including necroptosis (Wang *et al.*, 2012; Galluzzi *et al.*, 2011;

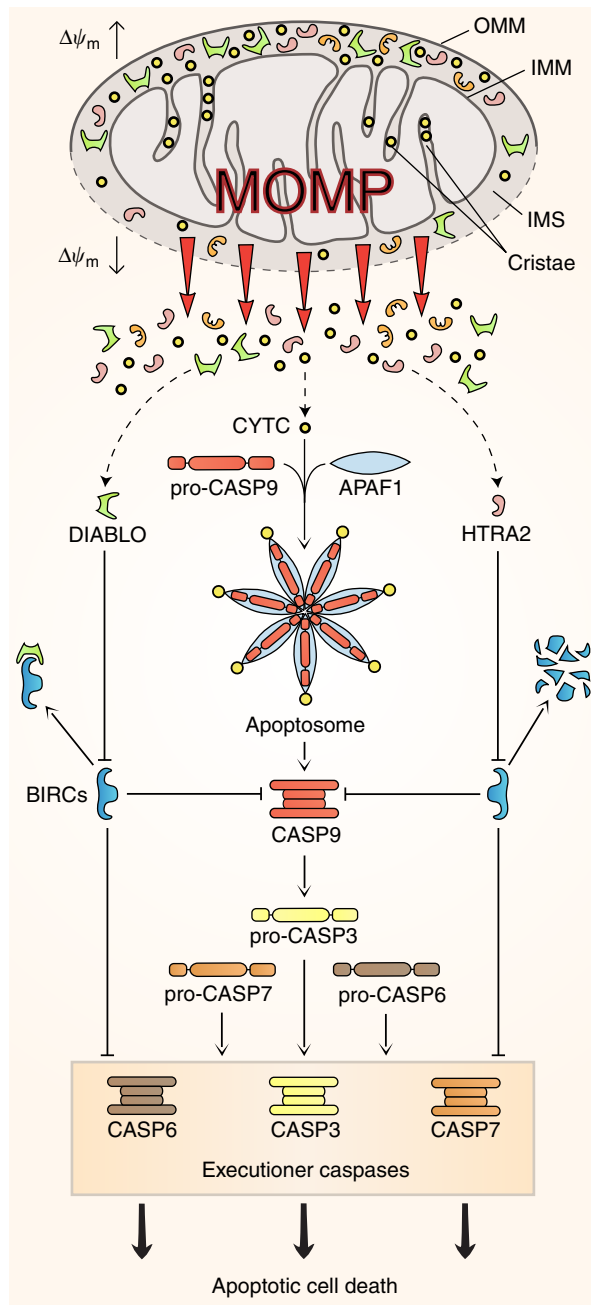


Figure 1 Mitochondrial control of apoptosis. Mitochondria constitute the central regulator of intrinsic apoptosis and contribute to several instances of its extrinsic counterpart. This key role reflects the dramatic consequences of widespread mitochondrial outer membrane permeabilization (MOMP), including (1) the arrest of ATP synthesis and other mitochondrial transmembrane potential ($\Delta\psi_m$)-dependent activities, and (2) the release of several cytotoxic proteins that are normally sequestered within the mitochondrial IMS into the cytosol. The latter include direct as well as indirect activators of caspases, globally promoting the self-amplificatory caspase (CASP) cascade that underlies the execution of apoptotic cell death. APAF1, apoptotic peptidase activating factor 1; BIRC, baculoviral IAP repeat containing; CYTC, cytochrome c; HTRA2, HtrA serine peptidase 2; IMM, inner mitochondrial membrane; OMM, outer mitochondrial membrane.

Vandenabeele *et al.*, 2010; Zhang *et al.*, 2009), MPT-driven regulated necrosis (Schinzel *et al.*, 2005; Baines *et al.*, 2005; Nakagawa *et al.*, 2005), and parthanatos (Wang *et al.*, 2003, 2011; Yu *et al.*, 2002). While the actual requirement of mitochondria for the execution of necroptosis has nowadays been reconsidered, these organelles are still regarded as key regulators of MPT-driven necrosis and parthanatos (Galluzzi *et al.*, 2014a; Vanden Berghe *et al.*, 2014).

The suspicion that mitochondria could be involved in necroptosis first surged in 2009, when RIPK3 was shown to boost the activity of several mitochondrial enzymes, including glutamate-ammonia ligase (GLUL) and glutamate dehydrogenase 1 (GLUD1) (Zhang *et al.*, 2009). Additional evidence was taken in support of a mitochondrial involvement in necroptosis, including the relatively circumstantial implication of (1) ROS (of unclear origin) (Morgan *et al.*, 2008; Kim *et al.*, 2007; Goossens *et al.*, 1999; Schulze-Osthoff *et al.*, 1992), (2) adenine nucleotide translocase (ANT) inhibition (Temkin *et al.*, 2006), and (3) the accumulation of mitochondriotoxic lipids like ceramide (Kim *et al.*, 2005; Suffys *et al.*, 1991). In that period, the concept of necroptosis had just begun to gather consensus and the molecular mechanisms responsible for this instance of regulated cell death were for the most part unknown (Galluzzi *et al.*, 2011; Vandenabeele *et al.*, 2010). The first connections between RIPK3 and the execution of necroptosis were indeed traced in 2012, when MLKL was suggested to transduce necrotic signals via a phosphoglycerate mutase family member 5 (PGAM5)-dependent, dynamin 1-like (DN1L)-mediated mechanism resulting in mitochondrial fragmentation (Wang *et al.*, 2012; Sun *et al.*, 2012). Since then, the critical function of MLKL in necroptosis has been confirmed and characterized ever more precisely (Galluzzi *et al.*, 2014b; Chen *et al.*, 2014; Cai *et al.*, 2014; Murphy *et al.*, 2013; Wu *et al.*, 2013). On the contrary, the actual implication of PGAM5, DN1L, and mitochondrial fragmentation has remained unconfirmed since. Recently, Douglas Green and collaborators have demonstrated that cells deprived of most mitochondria upon the induction of widespread mitophagy are resistant to apoptotic triggers but perfectly sensitive to the induction of necroptosis (Tait *et al.*, 2013). The same authors have also shown that the absence of (most) mitochondria significantly blunts the production of ROS, yet does not compromise the ability of butylated hydroxyanisole, a ROS scavenger, to rescue cells from necroptosis (Tait *et al.*, 2013). Taken together, these findings indicate that necroptosis does not causally involve mitochondria (and by extension mitochondrial ROS).

Conversely, mitochondria play a critical role in MPT-driven regulated necrosis (Schinzel *et al.*, 2005; Baines *et al.*, 2005; Nakagawa *et al.*, 2005). The term MPT is generally employed to indicate an abrupt increase in the permeability to ions and small solutes of the IMM, resulting in the osmotic breakdown of mitochondria (see below) (Brenner and Moulin, 2012; Brenner and Grimm, 2006). The precise composition of the supramolecular entity that is currently believed to mediate the MPT, the so-called 'permeability transition pore complex' (PTPC), has not yet been determined with precision (Galluzzi and Kroemer, 2007; Baines *et al.*, 2007; Kokoszka *et al.*, 2004). Nonetheless, robust pharmacological and genetic evidence accumulated throughout the past two decades demonstrates

that CYPD controls several instances of MPT-driven regulated necrosis in a nonredundant fashion (Schinzel *et al.*, 2005; Baines *et al.*, 2005; Nakagawa *et al.*, 2005). Thus, rodents pretreated with CYPD inhibitors (e.g., cyclosporin A and sangliferin A) as well as *Ppif*^{-/-} mice appear to be more resistant than their vehicle-treated or wild-type counterparts to a variety of insults that promote necrotic cell death *in vivo*, including ischemia-reperfusion injuries of the heart (Nakagawa *et al.*, 2005; Baines *et al.*, 2005; Kung *et al.*, 2011), brain (Schinzel *et al.*, 2005; Vaseva *et al.*, 2012), and kidneys (Linkermann *et al.*, 2013; Devalaraja-Narashimha *et al.*, 2009).

Of note, in selected instances, MPT-driven cell death can manifest with an apoptotic, rather than a necrotic, morphology (Kroemer *et al.*, 2009; Galluzzi *et al.*, 2007). The reasons underlying such a variability in the morphological manifestations of MPT, possibly reflecting the activation of distinct biochemical cascades, are poorly understood. As caspase-dependent proteolysis is an energy-consuming process and MPT is generally associated with the abrupt cessation of mitochondrial ATP synthesis, intracellular ATP stores have been suggested to entirely determine how MPT-driven cell death is executed (and hence manifests) (Leist *et al.*, 1997). More probably, several different factors influence the propensity of caspases to intervene in MPT-driven cell death, perhaps including the expression levels of BIRC family members and their endogenous inhibitors (Jost *et al.*, 2009). In addition, multiple IMS proteins released by permeabilized mitochondria mediate bona fide caspase-independent cytotoxic functions. This is the case of HTRA2, which degrades many proteins other than BIRC family members (Srinivasula *et al.*, 2003; Yang *et al.*, 2003; Martins *et al.*, 2002), as well as of ENDOG and AIF, both of which can access the nucleus and operate as endonucleases to mediate large-scale chromatin degradation (at least in some circumstances) (Parrish and Xue, 2003; Ye *et al.*, 2002; Van Loo *et al.*, 2001; Li *et al.*, 2001). In summary, mitochondria exert a crucial control on MPT-driven cell death, which most often (but not always) manifests with a necrotic morphotype.

Along similar lines, mitochondria have been critically involved in the execution of the PARP1- and AIF-dependent instance of regulated necrosis known as parthanatos (Wang *et al.*, 2011, 2003; Yu *et al.*, 2002). Interestingly, while *Parp1*^{-/-} mice are viable and develop normally (Masutani *et al.*, 1999; Wang *et al.*, 1995), the lack of *Aifm1* at the whole body level is associated with death by embryonic day E9 (Brown *et al.*, 2006). In addition, the cardiac and skeletal muscle-specific deletion of *Aifm1* results in the gross metabolic perturbations that reflect respiratory defects at the level of complex I (Pospisilik *et al.*, 2007; Joza *et al.*, 2005). Thus, AIF resembles CYTC in that it normally mediates key, non-redundant vital functions (Galluzzi *et al.*, 2008a, 2012c). Nonetheless, a large amount of experimental evidence convincingly links PARP1 hyperactivation to the calpain-independent and PAR-dependent release of AIF from mitochondria and neuronal death *in vivo* (Wang *et al.*, 2009, 2011). Thus, rodents pretreated with chemical inhibitors of PARP1, *Parp1*^{-/-} mice as well as Harlequin mice (Klein *et al.*, 2002), reportedly resist a wide panel of neurotoxic insults including multiple chemicals (Mandir *et al.*, 1999, 2000; Wang *et al.*, 2009), excitotoxic neurotransmitters (Andrabi *et al.*, 2011),

irradiation (Osato *et al.*, 2010), trauma (Piao *et al.*, 2012; Sarnaik *et al.*, 2010), retinal detachment (Hisatomi *et al.*, 2008), as well as perinatal (Hagberg *et al.*, 2004; Zhu *et al.*, 2003, 2007b), and adult cerebral ischemia (Culmsee *et al.*, 2005; Zhu *et al.*, 2007a; Thal *et al.*, 2011; Li *et al.*, 2010).

Besides promoting the release of AIF from mitochondria, a process that depends on the binding of AIF to poly(ADP-ribose) moieties (Wang *et al.*, 2011), PARP1 hyperactivation results in the rapid depletion of NAD⁺, which has a dramatic impact on both mitochondrial and glycolytic ATP synthesis (Chiarugi, 2005). The relative contribution of such a bioenergetic crisis and AIF to the execution of parthanatos remains to be established and may exhibit a significant degree of context dependency. Several factors including global metabolic fitness (Galluzzi *et al.*, 2013) as well as previous exposure to DNA-damaging conditions (Michels *et al.*, 2013) may indeed influence the propensity of cells to undergo parthanatos. Further studies are required to elucidate these aspects.

Irrespective of the above-mentioned incognita, mitochondria play a critical role in the control of MPT-driven regulated necrosis and parthanatos, while they appear to be dispensable for necroptotic cell death (Figure 2).

Mitochondrial Membrane Permeabilization

MOMP

MOMP is the major mechanism whereby mitochondria lose their structural and functional integrity in the course of apoptosis (Tait and Green, 2010; Taylor *et al.*, 2008; Kroemer *et al.*, 2007), and is specifically initiated at the outer mitochondrial membrane (OMM) by the pore-forming activity of multidomain pro-apoptotic members of the Bcl-2 protein family, notably BCL2-associated X protein (BAX) and BCL2-antagonist/killer 1 (BAK1) (Lindsten *et al.*, 2000; Wei *et al.*, 2001). Under physiological conditions, BAX exists as a cytosolic protein, while BAK1 is associated with the OMM in an inactive conformation (Czabotar *et al.*, 2014; Youle and Strasser, 2008; Schellenberg *et al.*, 2013). In response to several triggers, both BAX and BAK1 undergo a conformational shift that confers them the ability to form large, homo- and/or heterodimeric channels across the OMM (Czabotar *et al.*, 2013; Suzuki *et al.*, 2000; Griffiths *et al.*, 1999; Desagher *et al.*, 1999). These pores are generally large enough to allow for the release of soluble IMS proteins into the cytoplasm, *de facto* initiating the assembly of the apoptosome and the execution of apoptotic cell death (Bleicken *et al.*, 2013; Rehm *et al.*, 2003; Kuwana *et al.*, 2002; Antonsson *et al.*, 2000).

The pore-forming activity of BAX and BAK1 is tightly controlled by various other members of the Bcl-2 proteins, including multidomain anti-apoptotic proteins-like B-cell CLL/lymphoma 2 (BCL-2), BCL2-like 1 (BCL2L1, best known as BCL-X_L), and myeloid cell leukemia sequence 1 (MCL1) as well as small, pro-apoptotic BH3-only proteins, including BID, BCL2 binding component 3 (BBC3, best known as PUMA), BCL2-like 11 (BCL2L11, best known as BIM), phorbol-12-myristate-13-acetate-induced protein 1 (PMAIP1, best known as NOXA), and BCL2-associated agonist of cell death (BAD) (Czabotar *et al.*, 2014; Youle and Strasser, 2008; Willis and

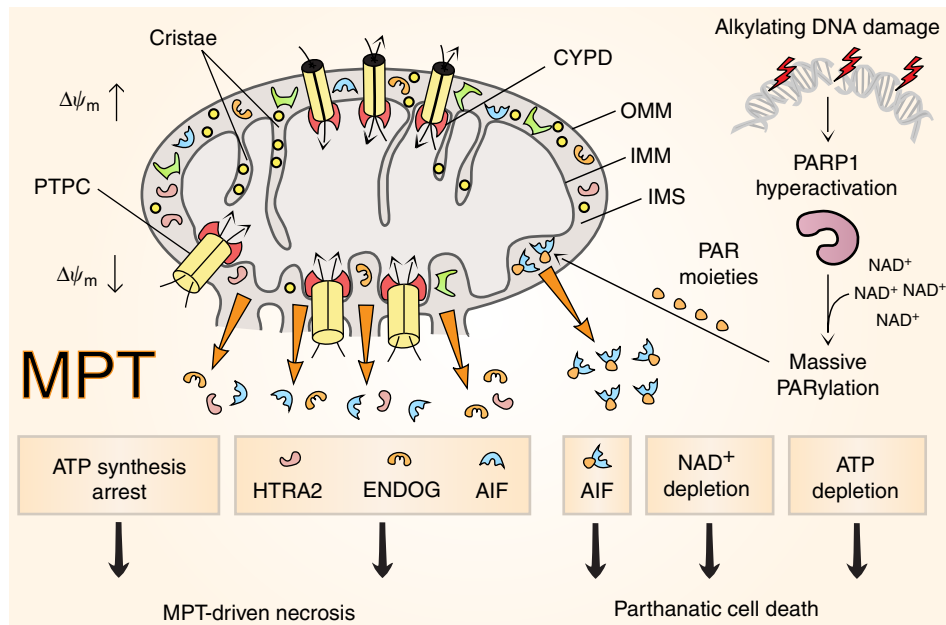


Figure 2 Mitochondrial control of regulated necrosis. While necroptosis appears to proceed irrespective of a causal mitochondrial involvement, this is not the case of mitochondrial permeability transition (MPT) driven regulated necrosis and parthanatos, both of which are controlled (at least in part) by mitochondria. MPT can be ignited by oxidative stress or cytosolic Ca^{2+} overload and results in the immediate cessation of mitochondrial transmembrane potential ($\Delta\psi_m$)-dependent functions, including ATP synthesis, as well as in the osmosis-driven structural breakdown of mitochondria. Although MPT involves the release of caspase activators and can indeed elicit apoptosis, in selected circumstances, MPT-driven necrosis is insensitive to caspase inhibition. Thus, other mechanisms are involved in the execution of this cell death subroutine, including the bioenergetic crisis resulting from the abrupt cessation of ATP synthesis and the release of caspase-independent cell death effectors normally sequestered within the mitochondrial intermembrane space (IMS). Parthanatos is elicited in the context of poly (ADP-ribose) polymerase 1 (PARP1) hyperactivation, which also causes major bioenergetic issues as it depletes NAD^+ . In addition, hyperactive PARP1 promotes the selective release of AIF from the IMS, hence unleashing its concealed endonucleolytic activity. ENDOG, endonuclease G; HTRA2, HtrA serine peptidase 2; IMM, inner mitochondrial membrane; OMM, outer mitochondrial membrane.

Adams, 2005). Anti-apoptotic BCL-2-like proteins are localized to the OMM and exert main MOMP-inhibitory functions as they sequester their pore-forming counterparts into inactive complexes (Edlich *et al.*, 2011; Fletcher *et al.*, 2008; Cheng *et al.*, 2003). In addition, BCL-2-like proteins have been shown to mediate cytoprotective effects (1) as they antagonize the activity of BAX and BAK1 at the ER, hence limiting the size of ER Ca^{2+} stores (Oakes *et al.*, 2005; Scorrano *et al.*, 2003), (2) as they promote mitochondrial ATP synthesis by physically interacting with the F_1F_0 -ATPase (Perciavalle *et al.*, 2012; Chen *et al.*, 2011; Alavian *et al.*, 2011).

Conversely, BH3-only proteins generally operate as stress sensors, becoming activated via transcriptional or post-translational mechanisms in response to multiple perturbations of intracellular homeostasis (Llambi *et al.*, 2011; Adams and Cory, 2007). Thus, BID is generally activated upon caspase-8 cleavage, representing the main link between extrinsic and intrinsic apoptosis (Li *et al.*, 1998; Luo *et al.*, 1998; Lovell *et al.*, 2008), PUMA and NOXA are transactivated in response to DNA damage by tumor protein p53 (TP53, best known as p53) (Michalak *et al.*, 2008; Nakano and Vousden, 2001; Yu *et al.*, 2001), BIM is released by inhibitory interactions with cytoskeletal components upon phosphorylation (Chen and Zhou, 2004; Puthalakath *et al.*, 1999), while the pro-apoptotic activity of BAD requires its dephosphorylation (Yan *et al.*, 2013; Kutuk and Letai, 2008). For illustrative purposes, BH3-

only proteins can be classified into two groups based on their main mechanism of action, namely, 'activators' (prototypes: BID, BIM) and 'derepressors' (prototypes: NOXA, BAD) (Kim *et al.*, 2006, 2009; Letai *et al.*, 2002; Giam *et al.*, 2008). Thus, while the former are thought to promote the pore-forming activity of BAX and BAK1 upon direct, though liable, physical interactions, a process that has also been indicated as 'kiss-and-run activation' (Chen *et al.*, 2005; Ren *et al.*, 2010), the latter do so indirectly, by displacing BAX and BAK1 from inhibitory liaisons with BCL-2-like proteins (Chen *et al.*, 2005; Willis *et al.*, 2005). Of note, MOMP as elicited by multiple stimuli (including developmental cues) critically relies on the presence of at least one of the main activator BH3-only proteins BIM, BID, and PUMA (Ren *et al.*, 2010). At odds with previous beliefs, it has recently been suggested that activator BH3-only proteins may exhibit some degree of specificity for BAX or BAK1 and hence exert non overlapping roles (Sarosiek *et al.*, 2013). Moreover, recent findings indicate that BAX is targeted to mitochondria continuously, not only upon activation by BH3-only proteins (Schellenberg *et al.*, 2013). According to this model, MOMP would be normally avoided by the continuative retrotranslocation of BAX to the cytosol, a process that would be mediated by BCL-X_L (Edlich *et al.*, 2011; Schellenberg *et al.*, 2013).

Interestingly a cytosolic pool of p53 has been ascribed with the ability to trigger MOMP in a transcription-independent

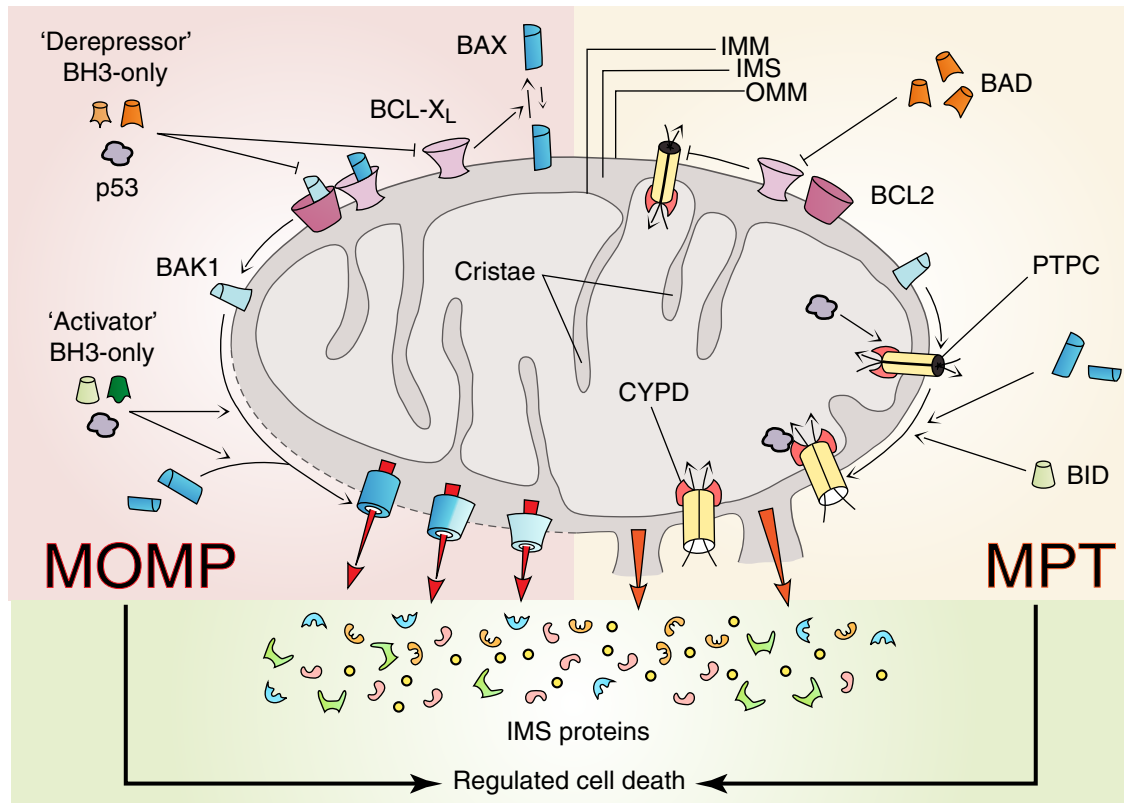


Figure 3 Mechanisms of mitochondrial membrane permeabilization. In the course of apoptosis and regulated necrosis, mitochondria can lose their structural and functional integrity via either of two non-mutually exclusive mechanisms, mitochondrial outer membrane permeabilization (MOMP) and mitochondrial permeability transition (MPT). The former is initiated at the outer mitochondrial membrane (OMM), and is controlled by several members of the Bcl-2 protein family. The latter originates at the inner mitochondrial membrane (IMM), owing to the activity of a poorly characterized, cyclophilin D (CYPD)-controlled supramolecular entity known as permeability transition pore complex (PTPC). Interestingly, the molecular machineries for MOMP and MPT mutually interact and are both regulated, via transcriptional as well as posttranslational mechanisms, by p53. Both MOMP and MPT result in the structural breakdown of mitochondria, dictating mostly apoptotic and mostly necrotic instances of cell death, respectively. BAD, BCL2-associated agonist of cell death; BAK1, BCL2-antagonist/killer 1; BAX, BCL2-associated X protein; BCL2, B-cell CLL/lymphoma 2; IMS, intermembrane space.

fashion via multiple mechanisms, including (1) a derepressor BH3-only protein-like activity, resulting in the liberation of BAX and BAK1 from their anti-apoptotic counterparts (Galluzzi *et al.*, 2008; Moll *et al.*, 2005; Mihara *et al.*, 2003); and (2) an activator BH3-only protein function, directly promoting the pore-forming activity of BAX (Chipuk *et al.*, 2003, 2004). The nuclear pool of p53 appears to play a critical role in this latter setting as it transactivates PUMA, hence causing the displacement of cytoplasmic p53 from inhibitory interactions with BCL-X_L and allowing it to mediate full-blown extranuclear apoptotic functions (Chipuk *et al.*, 2005).

Of note, in spite of nearly two decades of studies, the precise mechanisms whereby BAX- and BAK1 mediate MOMP have not yet been determined at the molecular level (Bender and Martinou, 2013; Czabotar *et al.*, 2014). It has been proposed that some mitochondrion-restricted lipids, including cardiolipin (Kuwana *et al.*, 2002; Lutter *et al.*, 2000; Choi *et al.*, 2007), would play a prominent role in this process, yet conclusive evidence in support of this hypothesis is lacking. Along similar lines, although several Bcl-2 proteins have been shown to functionally and physically interact with various components of the molecular machinery that controls

mitochondrial dynamics (Rolland *et al.*, 2009; Berman *et al.*, 2009; Sheridan *et al.*, 2008; Karbowski *et al.*, 2006; Arnoult *et al.*, 2005), to which extent mitochondrial fission and fusion actually influence MOMP remains to be determined, as several contradictory reports have been published in this respect (Van Der Bliek *et al.*, 2013; Martinou and Youle, 2011; Westermann, 2010). Interestingly, recent data suggest that BAX and BAK1 may impact on the propensity of cells to undergo MPT (see below) owing to their ability to regulate mitochondrial dynamics (Whelan *et al.*, 2012). Further studies are required to provide additional insights into this possibility.

Irrespective of these and other unresolved questions, MOMP stands out as a central step in the regulation of apoptosis (Figure 3).

MPT

MPT is a key decisional step in some apoptotic and necrotic instances of cell death (Galluzzi *et al.*, 2012d, 2014a; Kroemer *et al.*, 2007), and – contrarily to MOMP – is specifically initiated at the IMM (Brenner and Moulin, 2012; Brenner and

Grimm, 2006; Tait and Green, 2010; Kroemer *et al.*, 2007). According to current models, MPT is mediated by the opening of the PTPC, a poorly characterized multiprotein complex assembled at the interface of the IMM and OMM, exhibiting Ca^{2+} -activatable, voltage-dependent, and cyclosporin A-sensitive channel activity (Brenner and Grimm, 2006; Yu *et al.*, 2001). Throughout the past two decades, several proteins have been suspected to constitute key structural or regulatory subunits of the PTPC, including (1) various isoforms of the voltage-dependent anion channel (VDAC) and translocator protein (18 kDa) (TSPO, best known as peripheral benzodiazepine receptor), which are embedded in the OMM (Tajeddine *et al.*, 2008; Galluzzi and Kroemer, 2007; Azarashvili *et al.*, 2007; Pastorino *et al.*, 1994; Szabo *et al.*, 1993; Szabo and Zoratti, 1993; Mcenery *et al.*, 1992); (2) multiple variants of ANT and solute carrier family 25 member 3 (SLC25A3, best known as phosphate carrier), which are inserted into the IMM (Pauleau *et al.*, 2008; Alcalá *et al.*, 2008; Palmieri, 2004; Ruck *et al.*, 1998; Beutner *et al.*, 1996, 1998); (3) CYPD, a protein of the mitochondrial matrix (Martel *et al.*, 2012; Schinzel *et al.*, 2005; Baines *et al.*, 2005; Basso *et al.*, 2005); (4) creatine kinase, mitochondrial 1 (CKMT1), a protein of the IMS (Beutner *et al.*, 1996, 1998; Verrier *et al.*, 2004); as well as (5) hexokinase 1 (HK1), HK2, glycogen synthase kinase β (GSK3 β), and protein kinase C ϵ (PKC ϵ), all of which are cytosolic proteins (Beutner *et al.*, 1996, 1998; Verrier *et al.*, 2004; Chiara *et al.*, 2008; Baines *et al.*, 2003; Galluzzi *et al.*, 2008b). Robust genetic evidence argues against a key role for most of these proteins (Galluzzi and Kroemer, 2007; Baines *et al.*, 2007; Kokoszka *et al.*, 2004), with the notable exception of CYPD (Schinzel *et al.*, 2005; Baines *et al.*, 2005; Nakagawa *et al.*, 2005). However, it seems unlikely that CYPD might constitute the actual pore-forming unit of the PTPC, for at least two reasons. First, CYPD is not embedded into the IMM but localized to the mitochondrial matrix. Second, *Ppif*^{-/-} mitochondria exhibit a pronounced shift in the concentration of Ca^{2+} ions that is required to open the PTPC, yet remain capable of undergoing MPT (Schinzel *et al.*, 2005; Baines *et al.*, 2005; Nakagawa *et al.*, 2005; Basso *et al.*, 2005). Data accumulating in the last few years indicate that the F_1F_0 -ATPase, in particular the c subunit, may play a key structural role within the PTPC, perhaps even constituting the long-sought pore-forming component (Bonora *et al.*, 2013, 2015; Giorgio *et al.*, 2013; Bernardi, 2013). Formal evidence in support of this notion, however, is still lacking. Actually, the molecular composition of the PTPC appears to be very dynamic, which has led some investigators to speculate that most, if not all, mitochondrial proteins might become able to assemble a relatively unspecific channel under some circumstances (Brenner and Moulin, 2012; Brenner and Grimm, 2006), a hypothesis that is still awaiting experimental confirmation.

Also the precise molecular mechanisms underlying MPT remain elusive. Indeed, while several reports indicate that the PTPC would trigger MPT upon the assumption of a high-conductance state (Zoratti *et al.*, 2010; Zoratti and Szabo, 1994; Zorov *et al.*, 1992), others suggest that MPT ensues the closure of the PTPC, resulting in an osmotic imbalance that eventually compromises the structural integrity of mitochondria (Vander Heiden and Thompson, 1999; Vander Heiden *et al.*, 1999, 2001). A model reconciling these apparently discrepant

interpretations has not yet been put forward. Irrespective of these mechanistic incognita, the MPT as triggered by cytosolic Ca^{2+} waves and oxidative stress has been shown to constitute an etiological determinant in several pathophysiologically relevant instances of cell death (see above).

Intriguingly, several members of the Bcl-2 proteins as well as p53 have been shown to interact with, thereby regulating the activity of, several proteins that have been involved in the PTPC (Vaseva *et al.*, 2012; Shimizu *et al.*, 1999; Vander Heiden *et al.*, 1999). In particular, BCL-2 and BCL-X_L reportedly inhibit MPT by controlling the conductance of VDAC1 (Shimizu *et al.*, 1999; Vander Heiden *et al.*, 1999). Moreover, BAX, BAK1, and BID have been shown to stimulate MPT by interacting with ANT1 or VDAC1 (Karch *et al.*, 2013; Marzo *et al.*, 1998; Zamzami *et al.*, 2000; Narita *et al.*, 1998). Along similar lines, BAD appears to initiate MPT in a VDAC1-dependent, BCL-X_L-sensitive manner, possibly owing to its ability to displace VDAC1 from inhibitory interactions with BCL-X_L (Roy *et al.*, 2009). Finally, a pool of p53 localized to the mitochondrial matrix has recently been suggested to contribute to MPT-driven regulated necrosis by physically interacting with CYPD (Vaseva *et al.*, 2012). Conversely, putative PTPC constituents may mediate MOMP-regulatory functions. In particular, VDAC2 has been shown to mediate robust cytoprotective effects owing to its ability to sequester active BAK1 (Cheng *et al.*, 2003). Thus, MPT not only delineates the frontiers between a cell's life and death in several instances of apoptosis and regulated necrosis, but is also intimately intertwined with MOMP in the context of a multipronged, mutually regulatory interaction (Figure 3).

Concluding Remarks

Mitochondria not only are essential for the survival of eukaryotic cells but also control several cell death subroutines, including most instances of apoptosis, MPT-driven necrosis and parthanatos (Galluzzi *et al.*, 2014a; Kroemer *et al.*, 2007; Tait and Green, 2010; Vanden Berghe *et al.*, 2014; Vandenabeele *et al.*, 2010). In addition, several mitochondrial constituents operate as damage-associated molecular patterns, alerting the organism of a threatening condition and eliciting an innate immune response aimed at the reestablishment of systemic homeostasis (Galluzzi *et al.*, 2012a; Tait and Green, 2012). Whether the pronounced ability of mitochondria to elicit cellular and organismal adaptive responses reflects (at least to some extent) their foreign evolutionistic derivation remains to be determined. Irrespective of this conundrum, mitochondria nowadays stand out as key regulators of cellular and organismal danger signaling (Galluzzi *et al.*, 2012a; Tait and Green, 2012).

See also: Cell Division/Death: Apoptosis: Apoptosis

References

- Adams, J.M., Cory, S., 2007. The Bcl-2 apoptotic switch in cancer development and therapy. *Oncogene* 26, 1324–1337.

- Ajioka, R.S., Phillips, J.D., Kushner, J.P., 2006. Biosynthesis of heme in mammals. *Biochimica et Biophysica Acta* 1763, 723–736.
- Alavian, K.N., Li, H., Collis, L., *et al.*, 2011. Bcl-xL regulates metabolic efficiency of neurons through interaction with the mitochondrial F1F0 ATP synthase. *Nature Cell Biology* 13, 1224–1233.
- Alcala, S., Klee, M., Fernandez, J., *et al.*, 2008. A high-throughput screening for mammalian cell death effectors identifies the mitochondrial phosphate carrier as a regulator of cytochrome *c* release. *Oncogene* 27, 44–54.
- Anderson, S., Bankier, A.T., Barrell, B.G., *et al.*, 1981. Sequence and organization of the human mitochondrial genome. *Nature* 290, 457–465.
- Andrabi, S.A., Kang, H.C., Haince, J.F., *et al.*, 2011. Iduna protects the brain from glutamate excitotoxicity and stroke by interfering with poly(ADP-ribose) polymer-induced cell death. *Nature Medicine* 17, 692–699.
- Antonsson, B., Montessuit, S., Lauper, S., *et al.*, 2000. Bax oligomerization is required for channel-forming activity in liposomes and to trigger cytochrome *c* release from mitochondria. *Biochemical Journal* 345 (Pt 2), 271–278.
- Arnoult, D., Rismanchi, N., Grodet, A., *et al.*, 2005. Bax/Bak-dependent release of DDP/TIMM8a promotes Drp1-mediated mitochondrial fission and mitoptosis during programmed cell death. *Current Biology* 15, 2112–2118.
- Arnoult, D., Tatischeff, I., Estaquier, J., *et al.*, 2001. On the evolutionary conservation of the cell death pathway: Mitochondrial release of an apoptosis-inducing factor during *Dictyostelium discoideum* cell death. *Molecular Biology of the Cell* 12, 3016–3030.
- Azarashvili, T., Grachev, D., Krestina, O., *et al.*, 2007. The peripheral-type benzodiazepine receptor is involved in control of Ca^{2+} -induced permeability transition pore opening in rat brain mitochondria. *Cell Calcium* 42, 27–39.
- Baines, C.P., Kaiser, R.A., Purcell, N.H., *et al.*, 2005. Loss of cyclophilin D reveals a critical role for mitochondrial permeability transition in cell death. *Nature* 434, 658–662.
- Baines, C.P., Kaiser, R.A., Sheiko, T., *et al.*, 2007. Voltage-dependent anion channels are dispensable for mitochondrial-dependent cell death. *Nature Cell Biology* 9, 550–555.
- Baines, C.P., Song, C.X., Zheng, Y.T., *et al.*, 2003. Protein kinase Cepsilon interacts with and inhibits the permeability transition pore in cardiac mitochondria. *Circulation Research* 92, 873–880.
- Basso, E., Fante, L., Fowlkes, J., *et al.*, 2005. Properties of the permeability transition pore in mitochondria devoid of Cyclophilin D. *Journal of Biological Chemistry* 280, 18558–18561.
- Bender, T., Martinou, J.C., 2013. Where killers meet—permeabilization of the outer mitochondrial membrane during apoptosis. *Cold Spring Harbor Perspectives in Biology* 5, a011106.
- Berman, S.B., Chen, Y.B., Qi, B., *et al.*, 2009. Bcl-x L increases mitochondrial fission, fusion, and biomass in neurons. *Journal of Cell Biology* 184, 707–719.
- Bernardi, P., 2013. The mitochondrial permeability transition pore: A mystery solved? *Frontiers in Physiology* 4, 95.
- Beutner, G., Ruck, A., Riede, B., *et al.*, 1996. Complexes between kinases, mitochondrial porin and adenylate translocator in rat brain resemble the permeability transition pore. *FEBS Letters* 396, 189–195.
- Beutner, G., Ruck, A., Riede, B., Brdiczka, D., 1998. Complexes between porin, hexokinase, mitochondrial creatine kinase and adenylate translocator display properties of the permeability transition pore. Implication for regulation of permeability transition by the kinases. *Biochimica et Biophysica Acta* 1368, 7–18.
- Beicken, S., Landeta, O., Landajuela, A., *et al.*, 2013. Proapoptotic Bax and Bak proteins form stable protein-permeable pores of tunable size. *Journal of Biological Chemistry* 288, 33241–33252.
- Bonora, M., Bononi, A., De Marchi, E., *et al.*, 2013. Role of the c subunit of the FO ATP synthase in mitochondrial permeability transition. *Cell Cycle* 12, 674–683.
- Bonora, M., Wieckowski, M.R., Chinopoulos, C., *et al.*, 2015. Molecular mechanisms of cell death: Central implication of ATP synthase in mitochondrial permeability transition. *Oncogene* 34 (12), 1475–1486. doi: 10.1038/nc.2014.96.
- Brenner, C., Grimm, S., 2006. The permeability transition pore complex in cancer cell death. *Oncogene* 25, 4744–4756.
- Brenner, C., Moulin, M., 2012. Physiological roles of the permeability transition pore. *Circulation Research* 111, 1237–1247.
- Brown, D., Yu, B.D., Joza, N., *et al.*, 2006. Loss of AIF function causes cell death in the mouse embryo, but the temporal progression of patterning is normal. *Proceedings of the National Academy of Sciences of the United States of America* 103, 9918–9923.
- Cai, Z., Jitkaew, S., Zhao, J., *et al.*, 2014. Plasma membrane translocation of trimerized MLKL protein is required for TNF-induced necroptosis. *Nature Cell Biology* 16, 55–65.
- Cain, K., 2003. Chemical-induced apoptosis: Formation of the Apaf-1 apoptosome. *Drug Metabolism Reviews* 35, 337–363.
- Chen, D., Zhou, Q., 2004. Caspase cleavage of BimEL triggers a positive feedback amplification of apoptotic signaling. *Proceedings of the National Academy of Sciences of the United States of America* 101, 1235–1240.
- Chen, L., Willis, S.N., Wei, A., *et al.*, 2005. Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function. *Molecular Cell* 17, 393–403.
- Chen, X., Li, W., Ren, J., *et al.*, 2014. Translocation of mixed lineage kinase domain-like protein to plasma membrane leads to necrotic cell death. *Cell Research* 24, 105–121.
- Chen, Y.B., Aon, M.A., Hsu, Y.T., *et al.*, 2011. Bcl-xL regulates mitochondrial energetics by stabilizing the inner membrane potential. *Journal of Cell Biology* 195, 263–276.
- Cheng, E.H., Sheiko, T.V., Fisher, J.K., *et al.*, 2003. VDAC2 inhibits BAK activation and mitochondrial apoptosis. *Science* 301, 513–517.
- Chiara, F., Castellaro, D., Marin, O., *et al.*, 2008. Hexokinase II detachment from mitochondria triggers apoptosis through the permeability transition pore independent of voltage-dependent anion channels. *PLoS ONE* 3, e1852.
- Chiarugi, A., 2005. Intrinsic mechanisms of poly(ADP-ribose) neurotoxicity: Three hypotheses. *Neurotoxicology* 26, 847–855.
- Chipuk, J.E., Bouchier-Hayes, L., Kuwana, T., *et al.*, 2005. PUMA couples the nuclear and cytoplasmic proapoptotic function of p53. *Science* 309, 1732–1735.
- Chipuk, J.E., Kuwana, T., Bouchier-Hayes, L., *et al.*, 2004. Direct activation of Bax by p53 mediates mitochondrial membrane permeabilization and apoptosis. *Science* 303, 1010–1014.
- Chipuk, J.E., Maurer, U., Green, D.R., Schuler, M., 2003. Pharmacologic activation of p53 elicits Bax-dependent apoptosis in the absence of transcription. *Cancer Cell* 4, 371–381.
- Choi, S.Y., Gonzalez, F., Jenkins, G.M., *et al.*, 2007. Cardiolipin deficiency releases cytochrome *c* from the inner mitochondrial membrane and accelerates stimuli-elicited apoptosis. *Cell Death & Differentiation* 14, 597–606.
- Claveria, C., Giovino, G., Sierra, R., Torres, M., 2013. Myc-driven endogenous cell competition in the early mammalian embryo. *Nature* 500, 39–44.
- Culmsee, C., Zhu, C., Landsamer, S., *et al.*, 2005. Apoptosis-inducing factor triggered by poly(ADP-ribose) polymerase and Bid mediates neuronal cell death after oxygen-glucose deprivation and focal cerebral ischemia. *Journal of Neuroscience* 25, 10262–10272.
- Czabotar, P.E., Lessene, G., Strasser, A., Adams, J.M., 2014. Control of apoptosis by the BCL-2 protein family: Implications for physiology and therapy. *Nature Reviews Molecular Cell Biology* 15, 49–63.
- Czabotar, P.E., Westphal, D., Dewson, G., *et al.*, 2013. Bax crystal structures reveal how BH3 domains activate Bax and nucleate its oligomerization to induce apoptosis. *Cell* 152, 519–531.
- Desagher, S., Osen-Sand, A., Nichols, A., *et al.*, 1999. Bid-induced conformational change of Bax is responsible for mitochondrial cytochrome *c* release during apoptosis. *Journal of Cell Biology* 144, 891–901.
- Devalaraja-Narashimha, K., Diener, A.M., Padanilam, B.J., 2009. Cyclophilin D gene ablation protects mice from ischemic renal injury. *American Journal of Physiology - Renal Physiology* 297, F749–F759.
- Devenish, R.J., Prescott, M., Rodgers, A.J., 2008. The structure and function of mitochondrial F1F0-ATP synthases. *International Review of Cell and Molecular Biology* 267, 1–58.
- Du, C., Fang, M., Li, Y., *et al.*, 2000. Smac, a mitochondrial protein that promotes cytochrome *c*-dependent caspase activation by eliminating IAP inhibition. *Cell* 102, 33–42.
- Edlich, F., Banerjee, S., Suzuki, M., *et al.*, 2011. Bcl-x(L) retrotranslocates Bax from the mitochondria into the cytosol. *Cell* 145, 104–116.
- Efremov, R.G., Sazanov, L.A., 2011. Respiratory complex I: 'steam engine' of the cell? *Current Opinion in Structural Biology* 21, 532–540.
- Elstner, M., Andreoli, C., Klopstock, T., *et al.*, 2009. The mitochondrial proteome database: MitoP2. *Methods Enzymol* 457, 3–20.
- Ferreira, K.S., Kreutz, C., Macnelly, S., *et al.*, 2012. Caspase-3 feeds back on caspase-8, Bid and XIAP in type I Fas signaling in primary mouse hepatocytes. *Apoptosis* 17, 503–515.
- Fletcher, J.I., Meusburger, S., Hawkins, C.J., *et al.*, 2008. Apoptosis is triggered when prosurvival Bcl-2 proteins cannot restrain Bax. *Proceedings of the National Academy of Sciences of the United States of America* 105, 18081–18087.
- Fulda, S., Gorman, A.M., Hori, O., Samali, A., 2010. Cellular stress responses: Cell survival and cell death. *International Journal of Cell Biology* 2010, 214074.

- Galluzzi, L., Aaronson, S.A., Abrams, J., *et al.*, 2009a. Guidelines for the use and interpretation of assays for monitoring cell death in higher eukaryotes. *Cell Death & Differentiation* 16, 1093–1107.
- Galluzzi, L., Blomgren, K., Kroemer, G., 2009b. Mitochondrial membrane permeabilization in neuronal injury. *Nature Reviews Neuroscience* 10, 481–494.
- Galluzzi, L., Joza, N., Tasdemir, E., *et al.*, 2008a. No death without life: Vital functions of apoptotic effectors. *Cell Death & Differentiation* 15, 1113–1123.
- Galluzzi, L., Kepp, O., Krautwald, S., *et al.*, 2014a. Molecular mechanisms of regulated necrosis. *Seminars in Cell & Developmental Biology* 35, 24–32.
- Galluzzi, L., Kepp, O., Kroemer, G., 2012a. Mitochondria: Master regulators of danger signalling. *Nature Reviews Molecular Cell Biology* 13, 780–788.
- Galluzzi, L., Kepp, O., Kroemer, G., 2014b. MLKL regulates necrotic plasma membrane permeabilization. *Cell Research* 24, 139–140.
- Galluzzi, L., Kepp, O., Tajeddine, N., Kroemer, G., 2008b. Disruption of the hexokinase-VDAC complex for tumor therapy. *Oncogene* 27, 4633–4635.
- Galluzzi, L., Kepp, O., Trojel-Hansen, C., Kroemer, G., 2012b. Mitochondrial control of cellular life, stress, and death. *Circulation Research* 111, 1198–1207.
- Galluzzi, L., Kepp, O., Trojel-Hansen, C., Kroemer, G., 2012c. Non-apoptotic functions of apoptosis-regulatory proteins. *EMBO Reports* 13, 322–330.
- Galluzzi, L., Kepp, O., Vander Heiden, M.G., Kroemer, G., 2013. Metabolic targets for cancer therapy. *Nature Reviews Drug Discovery* 12, 829–846.
- Galluzzi, L., Kroemer, G., 2007. Mitochondrial apoptosis without VDAC. *Nature Cell Biology* 9, 487–489.
- Galluzzi, L., Maiuri, M.C., Vitale, I., *et al.*, 2007. Cell death modalities: Classification and pathophysiological implications. *Cell Death & Differentiation* 14, 1237–1243.
- Galluzzi, L., Morselli, E., Kepp, O., *et al.*, 2008c. Targeting p53 to mitochondria for cancer therapy. *Cell Cycle* 7, 1949–1955.
- Galluzzi, L., Morselli, E., Kepp, O., *et al.*, 2010. Mitochondrial gateways to cancer. *Molecular Aspects of Medicine* 31, 1–20.
- Galluzzi, L., Vanden Berghe, T., Vanlangenakker, N., *et al.*, 2011. Programmed necrosis from molecules to health and disease. *International Review of Cell and Molecular Biology* 289, 1–35.
- Galluzzi, L., Vitale, I., Abrams, J.M., *et al.*, 2012d. Molecular definitions of cell death subroutines: Recommendations of the Nomenclature Committee on Cell Death 2012. *Cell Death & Differentiation* 19, 107–120.
- Garrido, C., Galluzzi, L., Brunet, M., *et al.*, 2006. Mechanisms of cytochrome *c* release from mitochondria. *Cell Death & Differentiation* 13, 1423–1433.
- Giam, M., Huang, D.C., Bouillet, P., 2008. BH3-only proteins and their roles in programmed cell death. *Oncogene* 27 (Suppl. 1), S128–S136.
- Giorgio, V., Von Stockum, S., Antoniel, M., *et al.*, 2013. Dimers of mitochondrial ATP synthase form the permeability transition pore. *Proceedings of the National Academy of Sciences of the United States of America* 110, 5887–5892.
- Gomes, L.C., Di Benedetto, G., Scorrano, L., 2011. During autophagy mitochondria elongate, are spared from degradation and sustain cell viability. *Nature Cell Biology* 13, 589–598.
- Goossens, V., Stange, G., Moens, K., *et al.*, 1999. Regulation of tumor necrosis factor-induced, mitochondria- and reactive oxygen species-dependent cell death by the electron flux through the electron transport chain complex I. *Antioxidants & Redox Signaling* 1, 285–295.
- Green, D.R., 2005. Apoptotic pathways: Ten minutes to dead. *Cell* 121, 671–674.
- Green, D.R., Galluzzi, L., Kroemer, G., 2011. Mitochondria and the autophagy-inflammation-cell death axis in organismal aging. *Science* 333, 1109–1112.
- Griffiths, G.J., Dubrez, L., Morgan, C.P., *et al.*, 1999. Cell damage-induced conformational changes of the pro-apoptotic protein Bak *in vivo* precede the onset of apoptosis. *Journal of Cell Biology* 144, 903–914.
- Guenebeaud, C., Goldschneider, D., Castets, M., *et al.*, 2010. The dependence receptor UNC5H2/B triggers apoptosis via PP2A-mediated dephosphorylation of DAP kinase. *Molecular Cell* 40, 863–876.
- Guicciardi, M.E., Gores, G.J., 2009. Life and death by death receptors. *FASEB Journal* 23, 1625–1637.
- Hagberg, H., Wilson, M.A., Matsushita, H., *et al.*, 2004. PARP-1 gene disruption in mice preferentially protects males from perinatal brain injury. *Journal of Neurochemistry* 90, 1068–1075.
- Hao, Z., Duncan, G.S., Chang, C.C., *et al.*, 2005. Specific ablation of the apoptotic functions of cytochrome C reveals a differential requirement for cytochrome *c* and Apaf-1 in apoptosis. *Cell* 121, 579–591.
- Hisatomi, T., Nakazawa, T., Noda, K., *et al.*, 2008. HIV protease inhibitors provide neuroprotection through inhibition of mitochondrial apoptosis in mice. *Journal of Clinical Investigation* 118, 2025–2038.
- Hu, S., Yang, X., 2003. Cellular inhibitor of apoptosis 1 and 2 are ubiquitin ligases for the apoptosis inducer Smac/DIABLO. *Journal of Biological Chemistry* 278, 10055–10060.
- Ichim, G., Genevois, A.L., Menard, M., *et al.*, 2013. The dependence receptor TrkC triggers mitochondria-dependent apoptosis upon Cobra-1 recruitment. *Molecular Cell* 51, 632–646.
- Jacobson, M.D., Weil, M., Raff, M.C., 1997. Programmed cell death in animal development. *Cell* 88, 347–354.
- Janicke, R.U., Sprengart, M.L., Wati, M.R., Porter, A.G., 1998. Caspase-3 is required for DNA fragmentation and morphological changes associated with apoptosis. *Journal of Biological Chemistry* 273, 9357–9360.
- Jin, S.M., Lazarou, M., Wang, C., *et al.*, 2010. Mitochondrial membrane potential regulates PINK1 import and proteolytic destabilization by PARL. *Journal of Cell Biology* 191, 933–942.
- Jost, P.J., Grabow, S., Gray, D., *et al.*, 2009. XIAP discriminates between type I and type II FAS-induced apoptosis. *Nature* 460, 1035–1039.
- Joza, N., Oudit, G.Y., Brown, D., *et al.*, 2005. Muscle-specific loss of apoptosis-inducing factor leads to mitochondrial dysfunction, skeletal muscle atrophy, and dilated cardiomyopathy. *Molecular and Cellular Biology* 25, 10261–10272.
- Joza, N., Susin, S.A., Daugas, E., *et al.*, 2001. Essential role of the mitochondrial apoptosis-inducing factor in programmed cell death. *Nature* 410, 549–554.
- Karbowski, M., Norris, K.L., Cleland, M.M., *et al.*, 2006. Role of Bax and Bak in mitochondrial morphogenesis. *Nature* 443, 658–662.
- Karch, J., Kwong, J.Q., Burr, A.R., *et al.*, 2013. Bax and Bak function as the outer membrane component of the mitochondrial permeability pore in regulating necrotic cell death in mice. *eLife* 2, e00772.
- Kaufmann, T., Strasser, A., Jost, P.J., 2012. Fas death receptor signalling: Roles of Bid and XIAP. *Cell Death & Differentiation* 19, 42–50.
- Kepp, O., Galluzzi, L., Lipinski, M., *et al.*, 2011. Cell death assays for drug discovery. *Nature Reviews Drug Discovery* 10, 221–237.
- Kim, H., Rafiuddin-Shah, M., Tu, H.C., *et al.*, 2006. Hierarchical regulation of mitochondrion-dependent apoptosis by BCL-2 subfamilies. *Nature Cell Biology* 8, 1348–1358.
- Kim, H., Tu, H.C., Ren, D., *et al.*, 2009. Stepwise activation of BAX and BAK by tBID, BIM, and PUMA initiates mitochondrial apoptosis. *Molecular Cell* 36, 487–499.
- Kim, W.H., Choi, C.H., Kang, S.K., *et al.*, 2005. Ceramide induces non-apoptotic cell death in human glioma cells. *Neurochemical Research* 30, 969–979.
- Kim, Y.S., Morgan, M.J., Choksi, S., Liu, Z.G., 2007. TNF-induced activation of the Nox1 NADPH oxidase and its role in the induction of necrotic cell death. *Molecular Cell* 26, 675–687.
- Klein, J.A., Longo-Guess, C.M., Rossmann, M.P., *et al.*, 2002. The harlequin mouse mutation downregulates apoptosis-inducing factor. *Nature* 419, 367–374.
- Kluck, R.M., Bossy-Wetzel, E., Green, D.R., Newmeyer, D.D., 1997. The release of cytochrome *c* from mitochondria: A primary site for Bcl-2 regulation of apoptosis. *Science* 275, 1132–1136.
- Kokoszka, J.E., Waymire, K.G., Levy, S.E., *et al.*, 2004. The ADP/ATP translocator is not essential for the mitochondrial permeability transition pore. *Nature* 427, 461–465.
- Kornmann, B., 2013. The molecular hug between the ER and the mitochondria. *Current Opinion in Cell Biology* 25, 443–448.
- Kroemer, G., Galluzzi, L., Brenner, C., 2007. Mitochondrial membrane permeabilization in cell death. *Physiological Reviews* 87, 99–163.
- Kroemer, G., Galluzzi, L., Vandenabeele, P., *et al.*, 2009. Classification of cell death: Recommendations of the Nomenclature Committee on Cell Death 2009. *Cell Death & Differentiation* 16, 3–11.
- Kroemer, G., Levine, B., 2008. Autophagic cell death: The story of a misnomer. *Nature Reviews Molecular Cell Biology* 9, 1004–1010.
- Kroemer, G., Marino, G., Levine, B., 2010. Autophagy and the integrated stress response. *Molecular Cell* 40, 280–293.
- Kroemer, G., Martin, S.J., 2005. Caspase-independent cell death. *Nature Medicine* 11, 725–730.
- Kuida, K., Haydar, T.F., Kuan, C.Y., *et al.*, 1998. Reduced apoptosis and cytochrome *c*-mediated caspase activation in mice lacking caspase 9. *Cell* 94, 325–337.
- Kuida, K., Zheng, T.S., Na, S., *et al.*, 1996. Decreased apoptosis in the brain and premature lethality in CPP32-deficient mice. *Nature* 384, 368–372.
- Kung, G., Konstantinidis, K., Kitsis, R.N., 2011. Programmed necrosis, not apoptosis, in the heart. *Circulation Research* 108, 1017–1036.
- Kutuk, O., Letai, A., 2008. Regulation of Bcl-2 family proteins by posttranslational modifications. *Current Molecular Medicine* 8, 102–118.
- Kuwana, T., Mackey, M.R., Perkins, G., *et al.*, 2002. Bid, Bax, and lipids cooperate to form supramolecular openings in the outer mitochondrial membrane. *Cell* 111, 331–342.
- Lakhani, S.A., Masud, A., Kuida, K., *et al.*, 2006. Caspases 3 and 7: Key mediators of mitochondrial events of apoptosis. *Science* 311, 847–851.

- Leblanc, A., Liu, H., Goodyer, C., *et al.*, 1999. Caspase-6 role in apoptosis of human neurons, amyloidogenesis, and Alzheimer's disease. *Journal of Biological Chemistry* 274, 23426–23436.
- Leist, M., Single, B., Castoldi, A.F., *et al.*, 1997. Intracellular adenosine triphosphate (ATP) concentration: A switch in the decision between apoptosis and necrosis. *Journal of Experimental Medicine* 185, 1481–1486.
- Letai, A., Bassik, M.C., Walensky, L.D., *et al.*, 2002. Distinct BH3 domains either sensitize or activate mitochondrial apoptosis, serving as prototype cancer therapeutics. *Cancer Cell* 2, 183–192.
- Li, H., Zhu, H., Xu, C.J., Yuan, J., 1998. Cleavage of BID by caspase 8 mediates the mitochondrial damage in the Fas pathway of apoptosis. *Cell* 94, 491–501.
- Li, L.Y., Luo, X., Wang, X., 2001. Endonuclease G is an apoptotic DNase when released from mitochondria. *Nature* 412, 95–99.
- Li, P., Nijhawan, D., Budihardjo, I., *et al.*, 1997. Cytochrome c and dATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell* 91, 479–489.
- Li, X., Klaus, J.A., Zhang, J., *et al.*, 2010. Contributions of poly(ADP-ribose) polymerase-1 and -2 to nuclear translocation of apoptosis-inducing factor and injury from focal cerebral ischemia. *Journal of Neurochemistry* 113, 1012–1022.
- Li, X., Yang, Y., Ashwell, J.D., 2002. TNF-RII and c-IAP1 mediate ubiquitination and degradation of TRAF2. *Nature* 416, 345–347.
- Linksten, T., Ross, A.J., King, A., *et al.*, 2000. The combined functions of proapoptotic Bcl-2 family members bax and bcl-2 are essential for normal development of multiple tissues. *Molecular Cell* 6, 1389–1399.
- Linkermann, A., Brasen, J.H., Darding, M., *et al.*, 2013. Two independent pathways of regulated necrosis mediate ischemia-reperfusion injury. *Proceedings of the National Academy of Sciences of the United States of America* 110, 12024–12029.
- Liu, X., Kim, C.N., Yang, J., *et al.*, 1996. Induction of apoptotic program in cell-free extracts: Requirement for dATP and cytochrome c. *Cell* 86, 147–157.
- Liu, Y., Shoji-Kawata, S., Sumpter Jr., R.M., *et al.*, 2013. Autophagy is a Na⁺/K⁺-ATPase-regulated form of cell death triggered by autophagy-inducing peptides, starvation, and hypoxia-ischemia. *Proceedings of the National Academy of Sciences of the United States of America* 110, 20364–20371.
- Llambi, F., Moldovanu, T., Tait, S.W., *et al.*, 2011. A unified model of mammalian BCL-2 protein family interactions at the mitochondria. *Molecular Cell* 44, 517–531.
- Locksley, R.M., Killeen, N., Lenardo, M.J., 2001. The TNF and TNF receptor superfamilies: Integrating mammalian biology. *Cell* 104, 487–501.
- Lovell, J.F., Billen, L.P., Bindner, S., *et al.*, 2008. Membrane binding by tBid initiates an ordered series of events culminating in membrane permeabilization by Bax. *Cell* 135, 1074–1084.
- Luo, X., Budihardjo, I., Zou, H., *et al.*, 1998. Bid, a Bcl2 interacting protein, mediates cytochrome c release from mitochondria in response to activation of cell surface death receptors. *Cell* 94, 481–490.
- Lutter, M., Fang, M., Luo, X., *et al.*, 2000. Cardiolipin provides specificity for targeting of tBid to mitochondria. *Nature Cell Biology* 2, 754–761.
- Maas, C., Verbrugge, I., De Vries, E., *et al.*, 2010. Smac/DIABLO release from mitochondria and XIAP inhibition are essential to limit clonogenicity of Type I tumor cells after TRAIL receptor stimulation. *Cell Death & Differentiation* 17, 1613–1623.
- Mandir, A.S., Poitras, M.F., Berliner, A.R., *et al.*, 2000. NMDA but not non-NMDA excitotoxicity is mediated by Poly(ADP-ribose) polymerase. *Journal of Neuroscience* 20, 8005–8011.
- Mandir, A.S., Przedborski, S., Jackson-Lewis, V., *et al.*, 1999. Poly(ADP-ribose) polymerase activation mediates 1-methyl-4-phenyl-1, 2,3,6-tetrahydropyridine (MPTP)-induced parkinsonism. *Proceedings of the National Academy of Sciences of the United States of America* 96, 5774–5779.
- Martel, C., Huynh Le, H., Garnier, A., *et al.*, 2012. Inhibition of the mitochondrial permeability transition for cytoprotection: Direct versus indirect mechanisms. *Biochemistry Research International* 2012, 213403.
- Martinou, J.C., Youle, R.J., 2011. Mitochondria in apoptosis: Bcl-2 family members and mitochondrial dynamics. *Developmental Cell* 21, 92–101.
- Martins, L.M., Iaccarino, I., Tenen, T., *et al.*, 2002. The serine protease Omi/HtrA2 regulates apoptosis by binding XIAP through a reaper-like motif. *Journal of Biological Chemistry* 277, 439–444.
- Marzo, I., Brenner, C., Zamzami, N., *et al.*, 1998. Bax and adenine nucleotide translocator cooperate in the mitochondrial control of apoptosis. *Science* 281, 2027–2031.
- Masutani, M., Nozaki, T., Nishiyama, E., *et al.*, 1999. Function of poly(ADP-ribose) polymerase in response to DNA damage: Gene-disruption study in mice. *Molecular and Cellular Biochemistry* 193, 149–152.
- Mcenery, M.W., Snowman, A.M., Trifiletti, R.R., Snyder, S.H., 1992. Isolation of the mitochondrial benzodiazepine receptor: Association with the voltage-dependent anion channel and the adenine nucleotide carrier. *Proceedings of the National Academy of Sciences of the United States of America* 89, 3170–3174.
- Mehlen, P., Bredesen, D.E., 2011. Dependence receptors: From basic research to drug development. *Science Signaling* 4, mr2.
- Meng, X.W., Peterson, K.L., Dai, H., *et al.*, 2011. High cell surface death receptor expression determines type I versus type II signaling. *Journal of Biological Chemistry* 286, 35823–35833.
- Michalak, E.M., Villunger, A., Adams, J.M., Strasser, A., 2008. In several cell types tumour suppressor p53 induces apoptosis largely via Puma but Noxa can contribute. *Cell Death & Differentiation* 15, 1019–1029.
- Michels, J., Vitale, I., Galluzzi, L., *et al.*, 2013. Cisplatin resistance associated with PARP hyperactivation. *Cancer Research* 73, 2271–2280.
- Mihara, M., Erster, S., Zaika, A., *et al.*, 2003. p53 has a direct apoptogenic role at the mitochondria. *Molecular Cell* 11, 577–590.
- Mille, F., Thibert, C., Fombonne, J., *et al.*, 2009. The Patched dependence receptor triggers apoptosis through a DRAL-caspase-9 complex. *Nature Cell Biology* 11, 739–746.
- Mitchell, P., 1961. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature* 191, 144–148.
- Moll, U.M., Wolff, S., Speidel, D., Deppert, W., 2005. Transcription-independent pro-apoptotic functions of p53. *Current Opinion in Cell Biology* 17, 631–636.
- Mondragon, L., Galluzzi, L., Mouhamad, S., *et al.*, 2009. A chemical inhibitor of Apaf-1 exerts mitochondrioprotective functions and interferes with the intra-S-phase DNA damage checkpoint. *Apoptosis* 14, 182–190.
- Morgan, M.J., Kim, Y.S., Liu, Z.G., 2008. TNF α and reactive oxygen species in necrotic cell death. *Cell Research* 18, 343–349.
- Murphy, J.M., Czabotar, P.E., Hildebrand, J.M., *et al.*, 2013. The pseudokinase MLKL mediates necroptosis via a molecular switch mechanism. *Immunity* 39, 443–453.
- Murphy, M.P., 2009. How mitochondria produce reactive oxygen species. *Biochemical Journal* 417, 1–13.
- Nakagawa, T., Shimizu, S., Watanabe, T., *et al.*, 2005. Cyclophilin D-dependent mitochondrial permeability transition regulates some necrotic but not apoptotic cell death. *Nature* 434, 652–658.
- Nakano, K., Vousden, K.H., 2001. PUMA, a novel proapoptotic gene, is induced by p53. *Molecular Cell* 7, 683–694.
- Narita, M., Shimizu, S., Ito, T., *et al.*, 1998. Bax interacts with the permeability transition pore to induce permeability transition and cytochrome c release in isolated mitochondria. *Proceedings of the National Academy of Sciences of the United States of America* 95, 14681–14686.
- Neupert, W., Herrmann, J.M., 2007. Translocation of proteins into mitochondria. *Annual Review of Biochemistry* 76, 723–749.
- Nunnari, J., Suomalainen, A., 2012. Mitochondria: In sickness and in health. *Cell* 148, 1145–1159.
- Oakes, S.A., Scorrano, L., Opferman, J.T., *et al.*, 2005. Proapoptotic BAX and BAK regulate the type 1 inositol trisphosphate receptor and calcium leak from the endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America* 102, 105–110.
- Okuno, D., Iino, R., Noji, H., 2011. Rotation and structure of FoF1-ATP synthase. *Journal of Biochemistry* 149, 655–664.
- Osato, K., Sato, Y., Ochiishi, T., *et al.*, 2010. Apoptosis-inducing factor deficiency decreases the proliferation rate and protects the subventricular zone against ionizing radiation. *Cell Death & Disease* 1, e84.
- Ow, Y.P., Green, D.R., Hao, Z., Mak, T.W., 2008. Cytochrome c: Functions beyond respiration. *Nature Reviews Molecular Cell Biology* 9, 532–542.
- Palmieri, F., 2004. The mitochondrial transporter family (SLC25): Physiological and pathological implications. *Pflügers Archiv* 447, 689–709.
- Parrish, J., Li, L., Klotz, K., *et al.*, 2001. Mitochondrial endonuclease G is important for apoptosis in *C. elegans*. *Nature* 412, 90–94.
- Parrish, J.Z., Xue, D., 2003. Functional genomic analysis of apoptotic DNA degradation in *C. elegans*. *Molecular Cell* 11, 987–996.
- Pastorino, J.G., Simbala, G., Gilfor, E., *et al.*, 1994. Protoporphyrin IX, an endogenous ligand of the peripheral benzodiazepine receptor, potentiates induction of the mitochondrial permeability transition and the killing of cultured hepatocytes by rotenone. *Journal of Biological Chemistry* 269, 31041–31046.
- Pauleau, A.L., Galluzzi, L., Scholz, S.R., *et al.*, 2008. Unexpected role of the phosphate carrier in mitochondrial fragmentation. *Cell Death & Differentiation* 15, 616–618.
- Percivalle, R.M., Stewart, D.P., Koss, B., *et al.*, 2012. Anti-apoptotic MCL-1 localizes to the mitochondrial matrix and couples mitochondrial fusion to respiration. *Nature Cell Biology* 14, 575–583.

- Piao, C.S., Loane, D.J., Stoica, B.A., *et al.*, 2012. Combined inhibition of cell death induced by apoptosis inducing factor and caspases provides additive neuroprotection in experimental traumatic brain injury. *Neurobiology of Disease* 46, 745–758.
- Portt, L., Norman, G., Clapp, C., *et al.*, 2011. Anti-apoptosis and cell survival: A review. *Biochimica et Biophysica Acta* 1813, 238–259.
- Pospisilik, J.A., Knauf, C., Joza, N., *et al.*, 2007. Targeted deletion of AIF decreases mitochondrial oxidative phosphorylation and protects from obesity and diabetes. *Cell* 131, 476–491.
- Puthalakath, H., Huang, D.C., O'reilly, L.A., *et al.*, 1999. The proapoptotic activity of the Bcl-2 family member Bim is regulated by interaction with the dynein motor complex. *Molecular Cell* 3, 287–296.
- Rambold, A.S., Kostecky, B., Elia, N., Lippincott-Schwartz, J., 2011. Tubular network formation protects mitochondria from autophagosomal degradation during nutrient starvation. *Proceedings of the National Academy of Sciences of the United States of America* 108, 10190–10195.
- Rehm, M., Dussmann, H., Prehn, J.H., 2003. Real-time single cell analysis of Smac/DIABLO release during apoptosis. *Journal of Cell Biology* 162, 1031–1043.
- Ren, D., Tu, H.C., Kim, H., *et al.*, 2010. BID, BIM, and PUMA are essential for activation of the BAX- and BAK-dependent cell death program. *Science* 330, 1390–1393.
- Riedl, S.J., Salvesen, G.S., 2007. The apoptosome: Signalling platform of cell death. *Nature Reviews Molecular Cell Biology* 8, 405–413.
- Rodriguez, J., Lazebnik, Y., 1999. Caspase-9 and APAF-1 form an active holoenzyme. *Genes & Development* 13, 3179–3184.
- Rolland, S.G., Lu, Y., David, C.N., Conradt, B., 2009. The BCL-2-like protein CED-9 of *C. elegans* promotes FZO-1/Mfn1.2- and EAT-3/Opa1-dependent mitochondrial fusion. *Journal of Cell Biology* 186, 525–540.
- Rong, Y., Distelhorst, C.W., 2008. Bcl-2 protein family members: Versatile regulators of calcium signaling in cell survival and apoptosis. *Annual Review of Physiology* 70, 73–91.
- Rowland, A.A., Voeltz, G.K., 2012. Endoplasmic reticulum-mitochondria contacts: Function of the junction. *Nature Reviews Molecular Cell Biology* 13, 607–625.
- Roy, S.S., Madesh, M., Davies, E., *et al.*, 2009. Bad targets the permeability transition pore independent of Bax or Bak to switch between Ca²⁺-dependent cell survival and death. *Molecular Cell* 33, 377–388.
- Ruck, A., Dolder, M., Wallimann, T., Brdiczka, D., 1998. Reconstituted adenine nucleotide translocase forms a channel for small molecules comparable to the mitochondrial permeability transition pore. *FEBS Letters* 426, 97–101.
- Sarnaik, A.A., Conley, Y.P., Okonkwo, D.O., *et al.*, 2010. Influence of PARP-1 polymorphisms in patients after traumatic brain injury. *Journal of Neurotrauma* 27, 465–471.
- Sarosiek, K.A., Chi, X., Bachman, J.A., *et al.*, 2013. BID preferentially activates BAK while BIM preferentially activates BAX, affecting chemotherapy response. *Molecular Cell* 51, 751–765.
- Scaffidi, C., Fulda, S., Srinivasan, A., *et al.*, 1998. Two CD95 (APO-1/Fas) signaling pathways. *EMBO Journal* 17, 1675–1687.
- Schellenberg, B., Wang, P., Keeble, J.A., *et al.*, 2013. Bax exists in a dynamic equilibrium between the cytosol and mitochondria to control apoptotic priming. *Molecular Cell* 49, 959–971.
- Schinzl, A.C., Takeuchi, O., Huang, Z., *et al.*, 2005. Cyclophilin D is a component of mitochondrial permeability transition and mediates neuronal cell death after focal cerebral ischemia. *Proceedings of the National Academy of Sciences of the United States of America* 102, 12005–12010.
- Schmidt, O., Pfanner, N., Meisinger, C., 2010. Mitochondrial protein import: From proteomics to functional mechanisms. *Nature Reviews Molecular Cell Biology* 11, 655–667.
- Schulze-Osthoff, K., Bakker, A.C., Vanhaesebroeck, B., *et al.*, 1992. Cytotoxic activity of tumor necrosis factor is mediated by early damage of mitochondrial functions. Evidence for the involvement of mitochondrial radical generation. *Journal of Biological Chemistry* 267, 5317–5323.
- Schutze, S., Tchikov, V., Schneider-Brachert, W., 2008. Regulation of TNFR1 and CD95 signalling by receptor compartmentalization. *Nature Reviews Molecular Cell Biology* 9, 655–662.
- Scorrano, L., Oakes, S.A., Opferman, J.T., *et al.*, 2003. BAX and BAK regulation of endoplasmic reticulum Ca²⁺: A control point for apoptosis. *Science* 300, 135–139.
- Scott, F.L., Denault, J.B., Riedl, S.J., *et al.*, 2005. XIAP inhibits caspase-3 and -7 using two binding sites: Evolutionarily conserved mechanism of IAPs. *EMBO Journal* 24, 645–655.
- Sena, L.A., Chandel, N.S., 2012. Physiological roles of mitochondrial reactive oxygen species. *Molecular Cell* 48, 158–167.
- Sheridan, C., Delivani, P., Cullen, S.P., Martin, S.J., 2008. Bax- or Bak-induced mitochondrial fission can be uncoupled from cytochrome C release. *Molecular Cell* 31, 570–585.
- Shimizu, S., Narita, M., Tsujimoto, Y., 1999. Bcl-2 family proteins regulate the release of apoptogenic cytochrome c by the mitochondrial channel VDAC. *Nature* 399, 483–487.
- Shiozaki, E.N., Chai, J., Rigotti, D.J., *et al.*, 2003. Mechanism of XIAP-mediated inhibition of caspase-9. *Molecular Cell* 11, 519–527.
- Srinivasula, S.M., Gupta, S., Datta, P., *et al.*, 2003. Inhibitor of apoptosis proteins are substrates for the mitochondrial serine protease Omi/HtrA2. *Journal of Biological Chemistry* 278, 31469–31472.
- Srinivasula, S.M., Hegde, R., Saleh, A., *et al.*, 2001. A conserved XIAP-interaction motif in caspase-9 and Smac/DIABLO regulates caspase activity and apoptosis. *Nature* 410, 112–116.
- Stehling, O., Lill, R., 2013. The role of mitochondria in cellular iron-sulfur protein biogenesis: Mechanisms, connected processes, and diseases. *Cold Spring Harbor Perspectives in Biology* 3, 1–17.
- Stock, D., Leslie, A.G., Walker, J.E., 1999. Molecular architecture of the rotary motor in ATP synthase. *Science* 286, 1700–1705.
- Strasser, A., Jost, P.J., Nagata, S., 2009. The many roles of FAS receptor signaling in the immune system. *Immunity* 30, 180–192.
- Suffys, P., Beyaert, R., De Valck, D., *et al.*, 1991. Tumour-necrosis-factor-mediated cytotoxicity is correlated with phospholipase-A2 activity, but not with arachidonic acid release per se. *European Journal of Biochemistry* 195, 465–475.
- Sun, L., Wang, H., Wang, Z., *et al.*, 2012. Mixed lineage kinase domain-like protein mediates necrosis signaling downstream of RIP3 kinase. *Cell* 148, 213–227.
- Susin, S.A., Lorenzo, H.K., Zamzami, N., *et al.*, 1999. Molecular characterization of mitochondrial apoptosis-inducing factor. *Nature* 397, 441–446.
- Suzuki, M., Youle, R.J., Tjandra, N., 2000. Structure of Bax: Coregulation of dimer formation and intracellular localization. *Cell* 103, 645–654.
- Szabo, I., De Pinto, V., Zoratti, M., 1993. The mitochondrial permeability transition pore may comprise VDAC molecules. II. The electrophysiological properties of VDAC are compatible with those of the mitochondrial megachannel. *FEBS Letters* 330, 206–210.
- Szabo, I., Zoratti, M., 1993. The mitochondrial permeability transition pore may comprise VDAC molecules. I. Binary structure and voltage dependence of the pore. *FEBS Letters* 330, 201–205.
- Tait, S.W., Green, D.R., 2010. Mitochondria and cell death: Outer membrane permeabilization and beyond. *Nature Reviews Molecular Cell Biology* 11, 621–632.
- Tait, S.W., Green, D.R., 2012. Mitochondria and cell signalling. *Journal of Cell Science* 125, 807–815.
- Tait, S.W., Oberst, A., Quarato, G., *et al.*, 2013. Widespread mitochondrial depletion via mitophagy does not compromise necroptosis. *Cell Reports* 5, 878–885.
- Tajeddine, N., Galluzzi, L., Kepp, O., *et al.*, 2008. Hierarchical involvement of Bak, VDAC1 and Bax in cisplatin-induced cell death. *Oncogene* 27, 4221–4232.
- Tauszig-Delamasure, S., Yu, L.Y., Cabrera, J.R., *et al.*, 2007. The TrkC receptor induces apoptosis when the dependence receptor notion meets the neurotrophin paradigm. *Proceedings of the National Academy of Sciences of the United States of America* 104, 13361–13366.
- Taylor, R.C., Cullen, S.P., Martin, S.J., 2008. Apoptosis: Controlled demolition at the cellular level. *Nature Reviews Molecular Cell Biology* 9, 231–241.
- Temkin, V., Huang, Q., Liu, H., *et al.*, 2006. Inhibition of ADP/ATP exchange in receptor-interacting protein-mediated necrosis. *Molecular and Cellular Biology* 26, 2215–2225.
- Tenev, T., Zachariou, A., Wilson, R., *et al.*, 2005. IAPs are functionally non-equivalent and regulate effector caspases through distinct mechanisms. *Nature Cell Biology* 7, 70–77.
- Thal, S.E., Zhu, C., Thal, S.C., *et al.*, 2011. Role of apoptosis inducing factor (AIF) for hippocampal neuronal cell death following global cerebral ischemia in mice. *Neuroscience Letters* 499, 1–3.
- Thornberry, N.A., Lazebnik, Y., 1998. Caspases: Enemies within. *Science* 281, 1312–1316.
- Twiddy, D., Brown, D.G., Adrain, C., *et al.*, 2004. Pro-apoptotic proteins released from the mitochondria regulate the protein composition and caspase-processing activity of the native Apaf-1/caspase-9 apoptosome complex. *Journal of Biological Chemistry* 279, 19665–19682.

- Van Der Bliek, A.M., Shen, Q., Kawajiri, S., 2013. Mechanisms of mitochondrial fission and fusion. *Cold Spring Harbor Perspectives in Biology* 5.
- Van Loo, G., Schotte, P., Van Gorp, M., *et al.*, 2001. Endonuclease G: A mitochondrial protein released in apoptosis and involved in caspase-independent DNA degradation. *Cell Death & Differentiation* 8, 1136–1142.
- Vanden Berghe, T., Linkermann, A., Jouan-Lanhuet, S., *et al.*, 2014. Regulated necrosis: The expanding network of non-apoptotic cell death pathways. *Nature Reviews Molecular Cell Biology* 15, 135–147.
- Vandenabeele, P., Galluzzi, L., Vanden Berghe, T., Kroemer, G., 2010. Molecular mechanisms of necroptosis: An ordered cellular explosion. *Nature Reviews Molecular Cell Biology* 11, 700–714.
- Vander Heiden, M.G., Chandel, N.S., Schumacker, P.T., Thompson, C.B., 1999. Bcl-xL prevents cell death following growth factor withdrawal by facilitating mitochondrial ATP/ADP exchange. *Molecular Cell* 3, 159–167.
- Vander Heiden, M.G., Li, X.X., Gottlieb, E., *et al.*, 2001. Bcl-xL promotes the open configuration of the voltage-dependent anion channel and metabolite passage through the outer mitochondrial membrane. *Journal of Biological Chemistry* 276, 19414–19419.
- Vander Heiden, M.G., Thompson, C.B., 1999. Bcl-2 proteins: Regulators of apoptosis or of mitochondrial homeostasis? *Nature Cell Biology* 1, E209–E216.
- Vaseva, A.V., Marchenko, N.D., Ji, K., *et al.*, 2012. p53 opens the mitochondrial permeability transition pore to trigger necrosis. *Cell* 149, 1536–1548.
- Verhagen, A.M., Ekert, P.G., Pakusch, M., *et al.*, 2000. Identification of DIABLO, a mammalian protein that promotes apoptosis by binding to and antagonizing IAP proteins. *Cell* 102, 43–53.
- Verhagen, A.M., Kratina, T.K., Hawkins, C.J., *et al.*, 2007. Identification of mammalian mitochondrial proteins that interact with IAPs via N-terminal IAP binding motifs. *Cell Death & Differentiation* 14, 348–357.
- Verrier, F., Deniaud, A., Lebras, M., *et al.*, 2004. Dynamic evolution of the adenine nucleotide translocase interactome during chemotherapy-induced apoptosis. *Oncogene* 23, 8049–8064.
- Vitale, I., Galluzzi, L., Senovilla, L., *et al.*, 2011. Illicit survival of cancer cells during polyploidization and depolyploidization. *Cell Death & Differentiation* 18, 1403–1413.
- Wajant, H., 2002. The Fas signaling pathway: More than a paradigm. *Science* 296, 1635–1636.
- Walsh, J.G., Cullen, S.P., Sheridan, C., *et al.*, 2008. Executioner caspase-3 and caspase-7 are functionally distinct proteases. *Proceedings of the National Academy of Sciences of the United States of America* 105, 12815–12819.
- Wang, H., Shimoji, M., Yu, S.W., *et al.*, 2003. Apoptosis inducing factor and PARP-mediated injury in the MPTP mouse model of Parkinson's disease. *Annals of the New York Academy of Sciences* 991, 132–139.
- Wang, Y., Kim, N.S., Haince, J.F., *et al.*, 2011. Poly(ADP-ribose) (PAR) binding to apoptosis-inducing factor is critical for PAR polymerase-1-dependent cell death (parthanatos). *Science Signaling* 4, ra20.
- Wang, Y., Kim, N.S., Li, X., *et al.*, 2009. Calpain activation is not required for AIF translocation in PARP-1-dependent cell death (parthanatos). *Journal of Neurochemistry* 110, 687–696.
- Wang, Z., Jiang, H., Chen, S., *et al.*, 2012. The mitochondrial phosphatase PGAM5 functions at the convergence point of multiple necrotic death pathways. *Cell* 148, 228–243.
- Wang, Z.Q., Auer, B., Stingl, L., *et al.*, 1995. Mice lacking ADPRT and poly(ADP-ribose)ylation develop normally but are susceptible to skin disease. *Genes & Development* 9, 509–520.
- Watford, M., 1991. The urea cycle: A two-compartment system. *Essays in Biochemistry* 26, 49–58.
- Wei, M.C., Zong, W.X., Cheng, E.H., *et al.*, 2001. Proapoptotic BAX and BAK: A requisite gateway to mitochondrial dysfunction and death. *Science* 292, 727–730.
- Westermann, B., 2010. Mitochondrial fusion and fission in cell life and death. *Nature Reviews Molecular Cell Biology* 11, 872–884.
- Whelan, R.S., Konstantinidis, K., Wei, A.C., *et al.*, 2012. Bax regulates primary necrosis through mitochondrial dynamics. *Proceedings of the National Academy of Sciences of the United States of America* 109, 6566–6571.
- Wilkinson, J.C., Wilkinson, A.S., Scott, F.L., *et al.*, 2004. Neutralization of Smac/Diablo by inhibitors of apoptosis (IAPs). A caspase-independent mechanism for apoptotic inhibition. *Journal of Biological Chemistry* 279, 51082–51090.
- Willis, S.N., Adams, J.M., 2005. Life in the balance: How BH3-only proteins induce apoptosis. *Current Opinion in Cell Biology* 17, 617–625.
- Willis, S.N., Chen, L., Dewson, G., *et al.*, 2005. Proapoptotic Bak is sequestered by Mcl-1 and Bcl-xL, but not Bcl-2, until displaced by BH3-only proteins. *Genes & Development* 19, 1294–1305.
- Wu, G., Chai, J., Suber, T.L., *et al.*, 2000. Structural basis of IAP recognition by Smac/DIABLO. *Nature* 408, 1008–1012.
- Wu, J., Huang, Z., Ren, J., *et al.*, 2013. Mkl1 knockout mice demonstrate the indispensable role of Mkl1 in necroptosis. *Cell Research* 23, 994–1006.
- Yan, J., Xiang, J., Lin, Y., *et al.*, 2013. Inactivation of BAD by IKK inhibits TNF α -induced apoptosis independently of NF- κ B activation. *Cell* 152, 304–315.
- Yang, Q.H., Church-Hajduk, R., Ren, J., *et al.*, 2003. Omi/HtrA2 catalytic cleavage of inhibitor of apoptosis (IAP) irreversibly inactivates IAPs and facilitates caspase activity in apoptosis. *Genes & Development* 17, 1487–1496.
- Yang, W., Cooke, M., Duckett, C.S., *et al.*, 2014. Distinctive effects of the cellular inhibitor of apoptosis protein c-IAP2 through stabilization by XIAP in glioblastoma multiforme cells. *Cell Cycle* 13 (6), 992–1005.
- Ye, H., Cande, C., Stephanou, N.C., *et al.*, 2002. DNA binding is required for the apoptogenic action of apoptosis inducing factor. *Nature Structural & Molecular Biology* 9, 680–684.
- Yin, X.M., Wang, K., Gross, A., *et al.*, 1999. Bid-deficient mice are resistant to Fas-induced hepatocellular apoptosis. *Nature* 400, 886–891.
- Yoshida, H., Kong, Y.Y., Yoshida, R., *et al.*, 1998. Apa1 is required for mitochondrial pathways of apoptosis and brain development. *Cell* 94, 739–750.
- Youle, R.J., Strasser, A., 2008. The BCL-2 protein family: Opposing activities that mediate cell death. *Nature Reviews Molecular Cell Biology* 9, 47–59.
- Yu, J., Zhang, L., Hwang, P.M., *et al.*, 2001. PUMA induces the rapid apoptosis of colorectal cancer cells. *Molecular Cell* 7, 673–682.
- Yu, S.W., Wang, H., Poitras, M.F., *et al.*, 2002. Mediation of poly(ADP-ribose) polymerase-1-dependent cell death by apoptosis-inducing factor. *Science* 297, 259–263.
- Yuan, J., 1996. Evolutionary conservation of a genetic pathway of programmed cell death. *Journal of Cellular Biochemistry* 60, 4–11.
- Zamzami, N., El Hamel, C., Maise, C., *et al.*, 2000. Bid acts on the permeability transition pore complex to induce apoptosis. *Oncogene* 19, 6342–6350.
- Zhang, D.W., Shao, J., Lin, J., *et al.*, 2009. RIP3, an energy metabolism regulator that switches TNF-induced cell death from apoptosis to necrosis. *Science* 325, 332–336.
- Zhivotovsky, B., Galluzzi, L., Kepp, O., Kroemer, G., 2009. Adenine nucleotide translocase: A component of the phylogenetically conserved cell death machinery. *Cell Death & Differentiation* 16, 1419–1425.
- Zhu, C., Qiu, L., Wang, X., *et al.*, 2003. Involvement of apoptosis-inducing factor in neuronal death after hypoxia-ischemia in the neonatal rat brain. *Journal of Neurochemistry* 86, 306–317.
- Zhu, C., Wang, X., Deinum, J., *et al.*, 2007a. Cyclophilin A participates in the nuclear translocation of apoptosis-inducing factor in neurons after cerebral hypoxia-ischemia. *Journal of Experimental Medicine* 204, 1741–1748.
- Zhu, C., Wang, X., Huang, Z., *et al.*, 2007b. Apoptosis-inducing factor is a major contributor to neuronal loss induced by neonatal cerebral hypoxia-ischemia. *Cell Death & Differentiation* 14, 775–784.
- Zoratti, M., De Marchi, U., Biasutto, L., Szabo, I., 2010. Electrophysiology clarifies the megarriddles of the mitochondrial permeability transition pore. *FEBS Letters* 584, 1997–2004.
- Zoratti, M., Szabo, I., 1994. Electrophysiology of the inner mitochondrial membrane. *Journal of Bioenergetics and Biomembranes* 26, 543–553.
- Zorov, D.B., Kinnally, K.W., Tedeschi, H., 1992. Voltage activation of heart inner mitochondrial membrane channels. *Journal of Bioenergetics and Biomembranes* 24, 119–124.

Tumor Necrosis Factor Signaling Pathways

A Rajput and CF Ware, Sanford Burnham Medical Research Institute, La Jolla, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Biologics A class of therapeutics composed of proteins or nucleic acids.

Death domain (DD) A protein module composed of a bundle of six α -helices. The DD is essential to generate a signal leading to apoptosis.

Inflammation A localized physical condition in which part of the body becomes reddened, swollen, hot, and often painful, especially as a reaction to injury or infection.

Ligand A molecule, such as a hormone, cytokine, or drug that binds to a specific receptor.

Phosphorylation A covalent attachment of a phosphoryl group to a serine, threonine, or tyrosine residue of a target protein.

Receptor A protein molecule in a cell or on the surface of a cell to which a ligand can bind, causing a change in the activity of the cell.

Signal transduction A set of reactions in a cell that occurs when a ligand attaches to a receptor leading to a cascade of biochemical reactions inside the cell that eventually reach the target molecule or reaction.

Tumor necrosis factor (TNF) A multifunctional proinflammatory cytokine that can bind to, and thus functions through its receptors TNFR1 (aka CD120a) and TNFR2 (aka CD120b). It is involved in the regulation of various biological processes including cell proliferation, differentiation, and cell death.

Ubiquitination A covalent attachment of the ubiquitin, a 76-amino-acid-comprising protein, via its C-terminus to an amino group on a target protein.

Introduction

Tumor necrosis factor (TNF) is the namesake member of a large superfamily of structurally related cytokines that are involved in many physiologic functions (Wiens and Glenney, 2011; Bodmer *et al.*, 2002; Locksley *et al.*, 2001). Discovered and rediscovered through its antitumor, septic shock, cytotoxic, and cachectic activities (Ruddle, 2014; Cerami and Beutler, 1988), TNF has emerged as a key host defense mechanism against infections.

TNF mediates its activity by engaging two specific cell surface receptors, themselves members of a superfamily of structurally related proteins. TNF receptors (TNFRs) activate signaling pathways leading to cellular differentiation, growth, and death (Walczak, 2011; Sun, 2011). The signaling pathways activated by TNFRs depend on many factors such as the cellular expression patterns of the receptors and signaling proteins (Schrofelbauer and Hoffmann, 2011). TNF receptor-1 (TNFR1, aka p55/60) and TNF receptor-2 (TNFR2, aka p75/80) differ structurally in their cytoplasmic regions and engage different sets of signaling proteins that directly contribute to differences in cellular responses.

TNF is one component of a larger network of cytokine signaling systems that controls many aspects of the innate and adaptive arms of the immune system (Ware, 2005; Commerman *et al.*, 2014). This network is defined by the shared usage of ligands and receptors, however, each ligand-receptor system activates signaling that initiates unique cellular and physiologic responses. For example, TNFR1 and TNFR2 engage another ligand, Lymphotoxin- α (LT α), which forms yet another ligand with Lymphotoxin- β that binds specifically to LT β receptor. The LT β receptor engages yet another ligand, termed LIGHT (lymphotoxin-like, exhibits inducible

expression, and competes with herpes simplex virus glycoprotein D for HVEM, a receptor expressed by T lymphocytes) (TNFSF14), which itself engages other receptors, Herpesvirus entry mediator (HVEM, TNFRSF14) and decoy receptor 3 (DcR3) (Ware and Sedy, 2011). A pivotal point in signaling is the ability of HVEM to engage receptors in the immunoglobulin superfamily, B- and T-lymphocyte attenuator, which limits immune responses (Murphy and Murphy, 2010). Together, this network forms an integrated system that allows communication among immune cells and tissue cells in their microenvironment and at distant sites.

The clinical relevance of research in signaling pathways emerged with the realization that antibody-based inhibitors of TNF alleviate inflammation and disease symptoms in patients with certain autoimmune diseases (Feldmann *et al.*, 1992). A significant fraction of patients with moderate to severe rheumatoid arthritis, psoriasis, or inflammatory bowel disease respond to antibodies or decoy receptors that competitively block TNF from engaging its specific receptors. The impact of TNF inhibitors is a major clinical advance that has spurred research into other TNF-related cytokines and their signaling pathways in pursuit of ways to alter human diseases (Croft *et al.*, 2013).

In this article we summarize the current state of the art for signaling pathways initiated by TNF and their relationship to cellular responses and disease.

The TNF and TNFR Superfamilies

The TNF superfamily (TNFSF) is composed of 19 structurally related proteins (ligands). These ligands bind to one or more molecules from the TNFR superfamily (TNFRSF), which

Table 1 Members of TNFSF

Symbol	Approved name	Synonyms	Chromosome
CD40LG	CD40 ligand	CD40L, TRAP, gp39, hCD40L, CD154	Xq26
CD70	CD70 molecule	CD27L	19p13
EDA	Ectodysplasin A	EDA1, XLHED, HED, XHED, ED1-A1, ED1-A2, EDA-A1, EDA-A2	Xq12-q13.1
FASLG	Fas ligand (TNF superfamily, member 6)	FasL, CD178	1q23
LTA	Lymphotoxin alpha	TNFSF1, LT	6p21.3
LTB	Lymphotoxin beta	p33, TNFSF3	6p21.3
TNF	Tumor necrosis factor (TNF)	TNFSF2, DIF, TNF-alpha	6p21.3
TNFSF4	TNF superfamily, member 4	OX-40L, gp34, CD252	1q25
TNFSF8	TNF superfamily, member 8	CD153	9q33
TNFSF9	TNF superfamily, member 9	4-1BB-L	19p13.3
TNFSF10	TNF superfamily, member 10	TRAIL, Apo-2L, TL2, CD253	3q26
TNFSF11	TNF superfamily, member 11	TRANCE, RANKL, OPGL, ODF, CD254	13q14
TNFSF12	TNF superfamily, member 12	TWEAK, DR3LG, APO3L	17p13.1
TNFSF13	TNF superfamily, member 13	APRIL, CD256	17p13.1
TNFSF13B	TNF superfamily, member 13b	BAFF, THANK, BLYS, TALL-1, TALL1, CD257	13q32-q34
TNFSF14	TNF superfamily, member 14	LIGHT, CD258	19p13.3
TNFSF15	TNF superfamily, member 15	TL1, VEGI, TL1A, VEGI192A, MGC129934, MGC129935	9q32
TNFSF18	TNF superfamily, member 18	AITRL, TL6, hGITRL	1q23

Source: HUGO Gene Nomenclature Committee (www.genenames.org).

consists of 29 structurally similar receptors. The members of TNFSF and TNFRSF are listed in [Tables 1](#) and [2](#), respectively.

The members of TNFSF are generally type II transmembrane proteins (intracellular N-terminus) having a trimeric TNF homology domain (THD), comprising of three jelly-roll protomers to form a biologically active trimer. These domains are located at the C-terminal region of the protein. All members of the TNF family form homotrimers. The exception is LT β , which functions by forming heterotrimer with LT α . All except LT α (which is entirely secreted) are transmembrane proteins that either act by cell-to-cell contact and can be cleaved by proteolysis (shedding) from the cell surface into a soluble form.

Members of the TNFRSF are type I transmembrane proteins (N-terminus exterior to the cell). In contrast to the compact structure of the ligands, the receptors have an elongated extracellular ligand-binding domain defined by a cysteine-rich amino acid motif that recurs three to six times. This motif is known as CRD (cysteine-rich domain) and typically contains six cysteines forming three disulfides. The receptors also have relatively large cytosolic domains that engage different signaling proteins.

The receptor-binding site in TNF is created by interactions between adjacent subunits, thus a trimer creates three receptor-binding sites ([Banner et al., 1993](#)). High-affinity ligand binding induces clustering of the cognate cell surface receptors, which creates the cytosolic domains into a binding platform allowing recruitment of signaling molecules that initiate signal transduction. Bivalent antibodies to TNFR or overexpression can trigger receptor activation mimicking the natural ligand. Cells maintain low expression of TNFR, often only a few thousand molecules per cell, to limit spontaneous signaling.

Structural Features of TNF, TNFR1, and TNFR2

TNF is a 233 amino acid (aa) residue (26 kDa) protein, with a short 29 aa residue cytoplasmic domain, a 28 aa residue

transmembrane segment, and a 176 aa residue extracellular region containing the THD where the receptor-binding sites are located ([Wajant et al., 2003](#)). The 17 kDa TNF protomers are composed of two stacked β -pleated sheets with five anti-parallel β -strands, forming a 'jelly-roll' β -structure. The inner β -sheet is involved in contacts between adjacent subunits, which promotes assembly into a trimer. The trimer is formed such that one edge of each subunit is packed against the inner sheet of its neighbor. The outer β -sheet is surface exposed. The surface exposed loops are involved in receptor binding. The functional unit of the ligand is the assembly of THD into trimers. Membrane TNF, as a transmembrane protein expressed on the surface of cells, is cleaved by a metalloprotease, TNF-converting enzyme (TACE) ([Black et al., 1997](#)) which liberates a trimeric soluble cytokine ([Figure 1](#)). The soluble form of TNF circulates throughout the body to confer TNF with its potent endocrine function – an ability to act at distant sites, far from the site of synthesis.

TNF biosynthesis is initiated by innate immune sensor receptors, such as toll-like receptors (TLRs) as well as antigen receptors of T and B lymphocytes. TLRs are a class of proteins that play a critical role in the early innate immune response to pathogens. TLRs recognize structural motifs found associated with microorganisms, known as pathogen-associated molecular pattern molecules, such as lipopolysaccharides, bacterial DNA, and viral double-stranded RNA, or molecules associated with cellular damage, such as heat shock proteins as well as protein fragments released from necrotic or dying cells. Stimulation of TLRs by the danger molecules initiates signaling cascade that activates the transcription factor NF- κ B (nuclear factor- κ B) and other transcription factors, which in turn leads to production of proinflammatory cytokines including TNF and type 1 interferons ([Beutler, 2009](#)). The production of TNF mRNA is regulated by NF- κ B, c-Jun, activator protein 1 (AP1), and nuclear factor associated with activated T cells (NFAT) transcription factors, which is consistent with the fact that the promoter region of the TNF gene has binding sites for these

Table 2 Members of TNFRSF

<i>Symbol</i>	<i>Approved name</i>	<i>Synonyms</i>	<i>Chromosome</i>
CD27	CD27 molecule	S152, Tp55	12p13
CD40	CD40 molecule, TNF receptor superfamily member 5	p50, Bp50	20q12-q13.2
EDA2R	Ectodysplasin A2 receptor	XEDAR, EDA-A2R, EDAA2R, TNFRSF27	Xq11.1
EDAR	Ectodysplasin A receptor	ED5, EDA3, Edar, ED1R, EDA1R	2q13
FAS	Fas cell surface death receptor	CD95, APO-1	10q24.1
LTBR	Lymphotoxin beta receptor (TNFR superfamily, member 3)	TNFCR, TNFR-RP, TNFR2-RP, TNF-R-III, TNFRSF3	12p13
NGFR	Nerve growth factor receptor	TNFRSF16, CD271, p75NTR	17q21-q22
RELT	RELT tumor necrosis factor receptor	FLJ14993	11q13.2
TNFRSF1A	TNFRSF, member 1A	TNF-R, TNFAR, TNFR60, TNF-R-I, CD120a, TNF-R55	12p13.2
TNFRSF1B	TNFRSF, member 1B	TNFBFR, TNFR80, TNF-R75, TNF-R-II, p75, CD120b	1p36.22
TNFRSF4	TNFRSF, member 4	ACT35, OX40, CD134	1p36
TNFRSF6B	TNFRSF, member 6b, decoy	DcR3, DCR3, TR6, M68	20q13.33
TNFRSF8	TNFRSF, member 8	KI-1	1p36
TNFRSF9	TNFRSF, member 9	CD137, 4-1BB	1p36
TNFRSF10A	TNFRSF, member 10a	DR4, Apo2, TRAILR-1, CD261	8p21
TNFRSF10B	TNFRSF, member 10b	DR5, KILLER, TRICK2A, TRAIL-R2, TRICKB, CD262	8p22-p21
TNFRSF10C	TNFRSF, member 10c, decoy without an intracellular domain	DcR1, TRAILR3, LIT, TRID, CD263	8p22-p21
TNFRSF10D	TNFRSF, member 10d, decoy with truncated death domain	DcR2, TRUNDD, TRAILR4, CD264	8p21
TNFRSF11A	TNFRSF, member 11a, NF- κ B activator	RANK, CD265, FEO	18q22.1
TNFRSF11B	TNFRSF, member 11b	OCIF, TR1	8q24
TNFRSF12A	TNFRSF, member 12A	FN14, TweakR, CD266	16p13.3
TNFRSF13B	TNFRSF, member 13B	TACI, CD267, IGAD2	17p11.2
TNFRSF13C	TNFRSF, member 13C	BAFFR, CD268	22q13.1-q13.3
TNFRSF14	TNFRSF, member 14	HVEM, ATAR, TR2, LIGHTR, HVEA, CD270	1p36.32
TNFRSF17	TNFRSF, member 17	BCM, CD269, TNFRSF13A	16p13.1
TNFRSF18	TNFRSF, member 18	AITR, GITR, CD357	1p36.3
TNFRSF19	TNFRSF, member 19	TAJ-alpha, TROY, TAJ, TRADE	13q12.11-q12.3
TNFRSF21	TNFRSF, member 21	DR6, CD358	6p21.1
TNFRSF25	TNFRSF, member 25	DR3, TRAMP, WSL-1, LARD, WSL-LR, DDR3, TR3, APO-3	1p36.2

Source: HUGO Gene Nomenclature Committee (www.genenames.org).

transcription factors (Tuohy *et al.*, 1998). Posttranscriptional mRNA regulation also occurs by the actions of miRNAs and RNA-binding proteins, such as tristetraprolin, mRNA decay factors, and specific 3'-untranslated region AU-rich elements (Stamou and Kontoyiannis, 2010; Gherzi *et al.*, 2004).

Both soluble and membrane TNF bind to two transmembrane receptor molecules: TNFR1 and TNFR2 (Tartaglia *et al.*, 1991). TNFR1 is a 455 aa residue (55 kDa) transmembrane glycoprotein. The molecule has a 190 aa residue extracellular region, a 25 aa residue transmembrane segment, and a 220 aa residue cytoplasmic domain. The extracellular region has four cysteine-rich motifs. The cytoplasmic domain has an 80 aa residue folding into α -helical bundle known as a 'death domain' that can trigger apoptotic pathway. TNFR2 is a 461 aa residue (75 kDa) transmembrane glycoprotein. This receptor consists of a 240 aa residue extracellular region, a 27 aa residue transmembrane segment and a 173 aa residue cytoplasmic domain, which contains structurally defined sequence known as the TRAF (TNFR-associated factor)-binding motif. The

patterns of receptor expression determine the responsiveness of a cell. Most cells express low but constitutive levels of TNFR1, while only some cells (endothelial cells, dendritic cells and macrophages, and activated lymphocytes) express TNFR2 (see Immunological Genome Project and BioGPS in Relevant Websites section for more details).

TNF-Induced TNFR Signaling

Signaling begins upon binding TNF to either TNFR1 or TNFR2. The affinity of TNF binding is high with a K_d in the piconanomolar range. The propagation of signals from receptor to enzymatically active proteins is mediated by two distinct types of signaling motifs in the cytosolic domain of the TNFR: Death Domains (DD) and TRAF-binding motifs. The death domain in TNFR1 is composed of six α -helices arranged antiparallel to one another. The DD serves as a protein-interaction site and participates in homotypic interactions with the DD containing

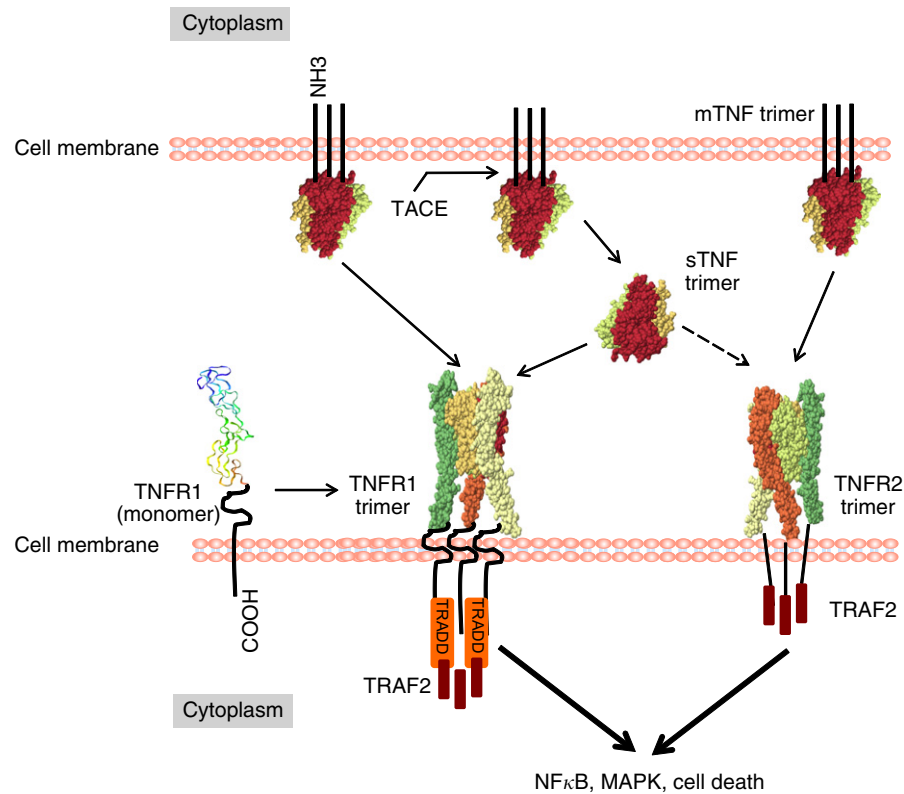


Figure 1 TNF Receptor and ligand complex. Membrane-bound TNF (mTNF) and soluble TNF (sTNF) derived by the action of TNF α -converting enzyme (TACE) bind to two members of the TNF-receptor superfamily, TNFR1 and TNFR2. Membrane TNF activates both TNFRs. Soluble TNF predominantly stimulates TNFR1 and has limited signaling capacities on TNFR2. Activation of TNFR1 and TNFR2 followed by recruitment of adaptors, TRADD and TRAF2, respectively initiates signaling events ranging from NF- κ B and MAPK signaling pathways to apoptosis.

proteins such as TNF receptor type 1-associated death domain protein (TRADD) and Fas-associated death domain containing protein (FADD). In contrast, TNFR2 initiates signaling through a TRAF-binding motif (~ 20 residues in length) that binds selectively to members of the TRAF family, TRAF2 and TRAF3.

TNFRs initiate intracellular signaling pathways that culminate in active transcription factors such as NF- κ B or AP1 (Mak and Yeh, 2002) that in turn control the expression of hundreds of genes that alter cellular differentiation. TNFR initiate the formation of active transcription factors by activating key enzymes including protein kinases, proteases, ubiquitin E3 ligases, and deubiquitinases. TNFR activation initiates several posttranslational modifications including serine phosphorylation and ubiquitination of the signaling components that fine tune the activation of transcription factors. These covalent modifications alter the properties of the target protein, thereby allowing a rapid alteration in the activity and/or the expression of signaling intermediates.

TNFR-Induced NF- κ B Activation

TNFR signaling results in the activation of several transcription factors, of which NF- κ B is critically important for induction of proinflammatory activity required for host defense. NF- κ B comprises a family of structurally related transcription factors that are involved in the control of a large number of cellular

processes, such as immune and inflammatory responses, developmental processes, cellular growth, and apoptosis. Upon ligand-dependent receptor oligomerization, the second step in the TNFR1 signaling is the recruitment of the adapter molecules, TRADD and receptor-interacting protein kinase 1 (RIPK1) to the receptor cytosolic domain, and then TRAF2 to TRADD, to form an enzymatic complex referred to as signaling complex I (Micheau and Tschopp, 2003). TRAF proteins contain an N-terminal RING domain, followed by several zinc fingers and a C-terminal domain (CTD) containing a coiled-coil (CC) and the TRAF-C region (Huxford and Ghosh, 2009). The C-terminal part of TRAFs facilitates trimerization and is involved in direct interactions with the TNFR2 and adaptor proteins, such as TRADD. The N-terminal region of TRAFs mediates dimerization (Bianchi and Meier, 2009). Mitogen-activated protein kinase kinase kinases (MAP3Ks) are then recruited and activated by the complex I. Activated MAP3Ks phosphorylate inhibitors of κ B (I κ B) to initiate the multi-protein enzyme complex called I κ B kinase (IKK) complex. The core of the IKK complex consists of two catalytic subunits, IKK α and IKK β , and a regulatory/scaffold subunit, NF- κ B essential modulator (NEMO or IKK γ). Subsequently, cIAP1/cIAP2 are recruited to the complex I through the coiled-coil region of the TRAF2 (Bianchi and Meier, 2009). Within this complex, RIPK1, TRAFs, and NEMO undergo several forms of ubiquitination, primarily conjugation of K63 or head-to tail (linear) polyubiquitin chains. K63 ubiquitination of RIPK1

and NEMO is catalyzed by the E3 ligase activity of cIAP1/cIAP2 or TRAF2 together with Uev1A and Ubc13 (Bertrand *et al.*, 2008). RIPK1 and NEMO undergo linear poly-ubiquitination by linear ubiquitin chain assembly complex (LUBAC), which consists of HOIL-1, HOIP, and SHARPIN (Tokunaga *et al.*, 2009). Ubiquitination of RIPK1 serves as a platform for the assembly of downstream IKK activating kinases. First, the transforming growth factor (TGF)- β activating kinase 1 (TAK1) is recruited by specific binding of its adaptors, TAK1-binding protein 2 (TAB2) and TAB3, through the ubiquitin chains (Ea *et al.*, 2006). TAK1 itself is K63 ubiquitinated, which triggers its autophosphorylation and activation (Scholz *et al.*, 2010).

In mammalian cells, the NF- κ B family consists of five members: RelA, RelB, c-Rel, p50/NF- κ B B1, and p52/NF- κ B B2 (proteolytic processing of p105 and p100 yields the active forms p50 and p52, respectively) (Ghosh *et al.*, 1998; Karin and Ben-Neriah, 2000). The translocation of NF- κ B leads to their activation (Schrofelbauer and Hoffmann, 2011). Homo- and hetero-dimers of NF- κ B family members are held inactive in the cytosol by I κ B, such as I κ B α , which mask nuclear localization motifs on the NF- κ B dimers. The binding of NEMO to ubiquitin chains brings IKK in close proximity to TAK1, which in turn leads to serine phosphorylation of residues in the activation loop of IKK (Tokunaga *et al.*, 2009). Activation of the NEMO-IKK complex results in phosphorylation and ubiquitin-dependent degradation of the inhibitor of I κ B proteins within the cell cytoplasm, which unmasks a nuclear translocation sequence on subunits of NF- κ B, enabling heterodimeric NF- κ B to enter into the nucleus where it binds to the enhancer of its target genes (Baldwin, 1996). Events leading to TNF-induced NF- κ B activation are diagramed in Figure 2.

TNFR2 directly interacts with TRAF2 through TRAF-binding motif and it lacks a DD indicating it does not directly activate caspases for apoptotic death. TNFR2 signaling follows the noncanonical NF- κ B pathway that regulates the activation of RelB form of NF- κ B through the NF- κ B-inducing kinase (NIK) and IKK α kinase. In unstimulated cells, NIK levels are maintained at vanishing low levels by ubiquitination and proteasome degradation through an E3 ligase complex comprised of TRAF3, TRAF2, and cIAP1/cIAP2 (Sun, 2011). Receptor ligation releases the active form of NIK by competitively displacing TRAF3 preventing further ubiquitination allowing NIK to accumulate (Sanjo *et al.*, 2010). NIK phosphorylates IKK α , which induces the proteasome-dependent processing of p100 to p52, degrading the inhibitory domain of p100 and allowing the RelB-p52 complex to activate gene transcription (Figure 2).

TNFR-Induced Cell Death

TNF-induced cell death signaling is carried out by TNFR1 (Tartaglia *et al.*, 1993). Ligation of TNFR1 promotes recruitment of intracellular 'death-inducing signaling complex' (DISC) proteins, including TRADD and FADD (Hsu *et al.*, 1996). These proteins create a scaffold permitting the recruitment and dimerization of caspase-8 (an initiator caspase), forming an active enzyme complex referred to as complex II (Micheau and Tschopp, 2003). Activated caspase-8 acts directly to cleave procaspase-3 and -7, which are known as the

executioner caspases as they directly cleave critical cellular substrates leading to apoptotic death (Ho and Hawkins, 2005; Figure 2). The activation of caspase-3, is essential for TNF-induced cell death, as it targets a latent DNase, the caspase-activated DNase (CAD), that degrades genomic DNA, a feature of apoptotic cell death (Enari *et al.*, 1998). Caspase-8 also cleaves BID, a mediator of mitochondria damage-associated cell death, an event that greatly amplifies the apoptotic process (Green *et al.*, 2011).

TNFR signaling induces a variety of genes that can inhibit cell death pathways. For example, the protease activity of caspase-8 is tightly regulated by a negative inhibitor protein FLICE (caspase-8) inhibitory protein (FLIP). cFLIP lacks a death domain but contains a death-effector domain (DED). DED allows its interactions with procaspase-8 as well as other DED containing proteins (Muzio *et al.*, 1997), which prevents constitutive procaspase-8 recruitment and activation to the TNFR1 DISC.

TNF activation-induced c-Jun-terminal kinase (JNK) pathway mediates phosphorylation and activation of the E3 ubiquitin ligase Itch, which specifically ubiquitinates cFLIP. Ubiquitination of cFLIP induces its proteasomal degradation, hence coupling JNK activation to TNF-induced cell death (Chang *et al.*, 2006). The inhibitor of apoptosis proteins (IAPs) are also important regulators of TNFR-induced cell death. IAPs are E3 ubiquitin ligases that inhibit caspases by acting with TRAF2 (Vince *et al.*, 2009). Both FLIP and IAPs are NF- κ B regulated genes. Thus, activation of TNFR1 induces a caspase-dependent apoptotic cell death that is tightly regulated by cFLIP and IAPs. Although TNF is a powerful inducer of apoptotic cell death, it induces cell death in transformed cells (Munker and Koeffler, 1987), virus infected cells (Sedger *et al.*, 1999), stressed cells, but not in most normal cells.

TNF has potent antiviral activity, which is mediated through TNFR 1 and TNFR 2. Many viruses have evolved mechanisms that specifically inhibit different steps of TNFR-induced apoptosis and NF- κ B signaling pathways (Hiscott *et al.*, 2006). Indeed most poxviruses express a pan-caspase inhibitor such as CrmA (Tewari and Dixit, 1995), and certain herpes viruses encode a FLIP ortholog that inhibits caspase-8 (Thome *et al.*, 1997).

Regulation of TNF Signaling by Ubiquitin Ligases and Deubiquitinases

The signaling complexes formed due to ligation of TNF to its receptors may be inactivated by the deubiquinating enzyme A20 (or TNFAIP3) which is thought to regulate RIP1 ubiquitination. A20 contains both an ovarian tumor (OUT) deubiquitinase domain and a coiled-coil zinc finger E3 ligase motif. A20 attenuates NF- κ B signaling in a two-step process. The A20 DUB domain first removes K63-linked ubiquitin chains from RIP1, and the A20 E3 ligase domain then promotes RIP1 K48 ubiquitination. As K63-ubiquitinated RIP1 is required for NF- κ B activation, A20-mediated RIP1 K63 de-ubiquitination and subsequent K48 ubiquitination-dependent proteasomal degradation cooperate to downregulate NF- κ B signaling (Wertz *et al.*, 2004). Upon TNF stimulation, the cellular zinc finger anti-NF- κ B protein Cezanne is also recruited to TNFR1 and promotes RIP1 de-ubiquitination, thereby attenuating NF- κ B

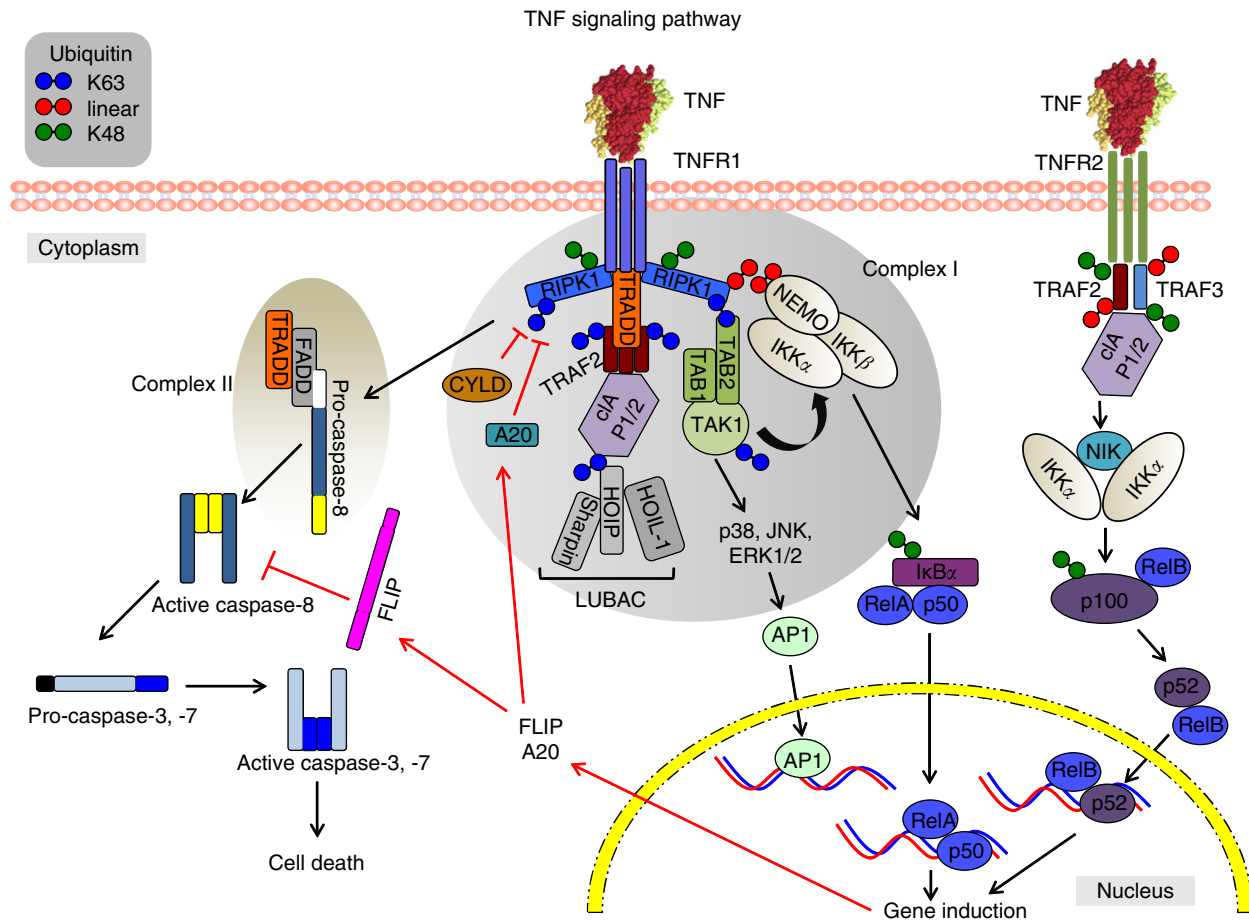


Figure 2 TNFR signaling pathways. Binding of trimeric TNF to TNF receptor-1 (TNFR1) leads to receptor trimerization and formation of TNF receptor signaling complex, also known as complex I. Then, the signaling molecules TNFR1-associated death domain (TRADD) and receptor-interacting protein kinase 1 (RIPK1) are recruited via their DDs to the complex I. Through TRADD, TNFR-associated factor 2 (TRAF2) is then recruited, which in turn recruits cIAP1 and cIAP2. The cIAPs then cause K63 (blue chains) and possibly also K48-linked ubiquitin (green chains) to different components of the complex. The LUBAC is then recruited to the complex I in a TRADD-, TRAF2-, and cIAP1/cIAP2-dependent manner. LUBAC activity in the complex I results in linear ubiquitination of RIPK1 and NF-κB essential modulator (NEMO) (red chains). The action of different E3 ubiquitin ligases facilitates the recruitment and positioning of the NEMO/IKK and TAK-binding protein/transforming growth factor β -activated kinase (TAB/TAK) complexes which enables the NF-κB and MAPK pathways to be activated. After a delay, and probably as a consequence of de-ubiquitination events at the membrane-bound TNFR signaling complex, carried out by de-ubiquitylating enzymes, A20 and CYLD, the composition of the complex changes, and a second complex, called complex II forms in the cytosol. Complex II recruits FADD and procaspase-8. This leads to activation of caspase-8 leading to its processing. Activated caspase-8 cleaves downstream effector caspases (caspase-3 and caspase-7). Proteolytic processing of caspase-3 and caspase-7 is responsible for the induction of cell death. Trimeric TNF can also bind to TNFR2 leading to recruitment of TRAF2, TRAF3, and cIAPs. This leads to degradation of TRAF2 and TRAF3 resulting in NIK accumulation. NIK in turn stimulates the IKK α , which then leads to phosphorylation of p100. This event leads to proteasome-dependent partial processing of p100 to p52, which in turn binds to RelB to form an active NF-κB dimer, hence activating an alternate NF-κB pathway. Blue, red, and green circles represent K63, linear, and K48 ubiquitin chains, respectively.

signaling (Enesa *et al.*, 2008). The tumor suppressor CYLD (cylindromatosis (turban tumor syndrome)) is also proposed to de-ubiquitinate K63-linked chains from RIPK1, NEMO, and TRAF proteins to attenuate downstream signaling cascades (Kovalenko *et al.*, 2003).

TNF-Induced Activation of MAPK Pathways

In addition to the activation of IKK, TNFR1 also leads to the potent activation of other kinases – JNK, p38, and ERK1/ERK2.

Activation of JNK depends on TRAF2 and on the formation of K63 ubiquitin chains (Fukushima *et al.*, 2007). Activation of p38 also depends on TRAF2 and RIPK1 (Lee *et al.*, 2003). Several MAP3Ks have been implicated in the activation, including MEKK1 (Takaesu *et al.*, 2000), TPL-2 (Das *et al.*, 2005), and ASK1 (Takaesu *et al.*, 2000). MAP3Ks then phosphorylate and activate MAP2Ks. While MKK3 and MKK6 specifically phosphorylate and activate p38, MKK7 activates JNK. MKK4 is less specific and can activate both p38 and JNK (Wang *et al.*, 2007). The main activator of the TNF-induced ERK1/ERK2 signal transduction pathway is TPL-2 (Eliopoulos *et al.*, 2003).

Table 3 Human genetic diseases associated with mutations in the components of TNF signaling pathway

Protein	Disease	Mutation
A20	Rheumatoid arthritis, systemic lupus erythematosus (SLE)	Phe127Cys
Caspase-8	Breast cancer, immune deficiency	Arg189Trp, Asp243His, Arg248Trp
CYLD	Multiple familial trichoepithelioma, familial cylindromatosis	Glu747Gly, Gln710X (missense mutation), Asp941Val, Arg758Ter
HOIL-1	Susceptibility to infection by pyogenic bacteria and auto-inflammation	Q185X (missense mutation)
IKK1	Cocoon syndrome	Gln422Ter
I κ B α	Ectodermal dysplasia	Ser32Ile, Glu14Ter, Ser32Ile, Trp11Ter
NEMO	Incontinentia pigmenti	Glu57Lys, Arg123Trp, Ala323Pro, Met339Val
	Ectodermal dysplasia, hypohidrotic, with immune deficiency	Leu153Arg, Leu85Arg, Leu227Pro, Ala220Gly, Asp311Asn, Gln335Ter, Glu323Ter, Cys349Arg, Cys349Phe
	Recurrent isolated invasive pneumococcal disease	Ter352Trp, Arg173Gly, Arg175Pro
	Susceptibility to Mycobacterial disease	Arg251Gln, Glu247Ala
	X-linked familial atypical micobacteriosis	Arg319Gln, Glu315Ala
	Lymphedema dysplasia, ectodermal dysplasia	Asp338Val, Asp406Val, Cys417Arg, Cys417Phe, Cys417Tyr, Met407Val
NIK	B-cell lymphopenia, decreased frequencies of class-switched memory B cells	Pro565Arg
TAB2	Congenital heart defects	Gln230Lys, Pro208Ser, Gln230Lys, Pro208Ser
TNFR1	Familial periodic fever	Mutation in first cysteine-rich domain
TRAF3	Herpes simplex encephalitis 3	Arg118Trp
XIAP	Lymphoproliferative syndrome	Glu118Ter, Cys203Tyr

Physiologic Functions of TNF and Disease Relevance

Even though most of our understanding about the physiological functions of TNF came from gene targeting approaches in mice, pre-knockout studies identified two important opposite functions of TNF. In mice, systemic bacterial infection leads to the overproduction of TNF causing a profound shock syndrome (Tracey *et al.*, 1987). In contrast to this deleterious effect, there was beneficial effect of TNF due to the stimulation of protective granuloma formation during mycobacterial infections (Flynn *et al.*, 1995). In the first knockout mice for TNFR1, two main functions of TNFR signaling were reported, its protective role in intracellular bacterial infections (*Listeria*) and its detrimental role in LPS-mediated liver toxicity (Pfeffer *et al.*, 1993; Rothe *et al.*, 1993). TNF is disproportionally produced in humans and animal models with inflammatory and autoimmune disorders, rheumatoid arthritis, psoriasis, and inflammatory bowel disease. Various human diseases are associated with genetic deficiencies or mutations in members of TNF superfamilies as well as individual components of the TNF signaling pathway (summarized in Table 3). Selective targeting of these components to develop disease-modifying therapeutics is an active area of research.

TNF Inhibitors

TNF inhibitors are used in treating several autoimmune and inflammatory diseases. TNF inhibitors are specific antibodies or soluble receptors and are classified as biologics, i.e., protein-based drugs. Both antibody- and receptor-based drugs hinder the binding of TNF ligand to receptor, hence functioning as

competitive antagonists. Currently there are five approved anti-TNF antibody-based drugs (Figure 3). Etanercept is a fusion protein of the extracellular domain of TNFR 2 with the Fc (fragment, crystallizable) region of human IgG1 forming a bivalent molecule. This bivalent molecule has higher avidity for TNF than naturally occurring soluble receptors. Infliximab is a partially humanized mouse monoclonal antibody to human TNF. Adalimumab and golimumab are two fully human monoclonal antibodies. Certolizumab is a Fab' fragment of a humanized TNF inhibitor monoclonal antibody that is chemically modified with polyethyleneglycol (PEGylation) to prolong circulation half-life in patients.

Impact of TNF Blockade in Inflammatory Diseases

TNF is centrally involved in several inflammatory conditions (Phillips *et al.*, 2004). TNF inhibitors are approved for use in the treatment of rheumatoid arthritis (RA), ankylosing spondylitis, juvenile RA, Crohn's disease, psoriatic arthritis, and ulcerative colitis. RA is a chronic autoimmune disease with inflammation in the joints affecting mobility. Treatment of murine arthritis with antibodies against TNF or with soluble TNFR ameliorates the disease (Williams *et al.*, 1992). In humans, clinical studies demonstrated the efficacy of TNF-neutralizing antibody in reducing the progression and symptoms of RA (Elliott *et al.*, 1994). Treatment of RA is usually initiated with disease-modifying antirheumatic drugs such as methotrexate or corticosteroids to improve symptoms and reduce joint damage and TNF inhibitors given to patients with moderate to severe RA. Approximately 30% of patients with RA show a positive response to TNF inhibition, however, other patients show only a

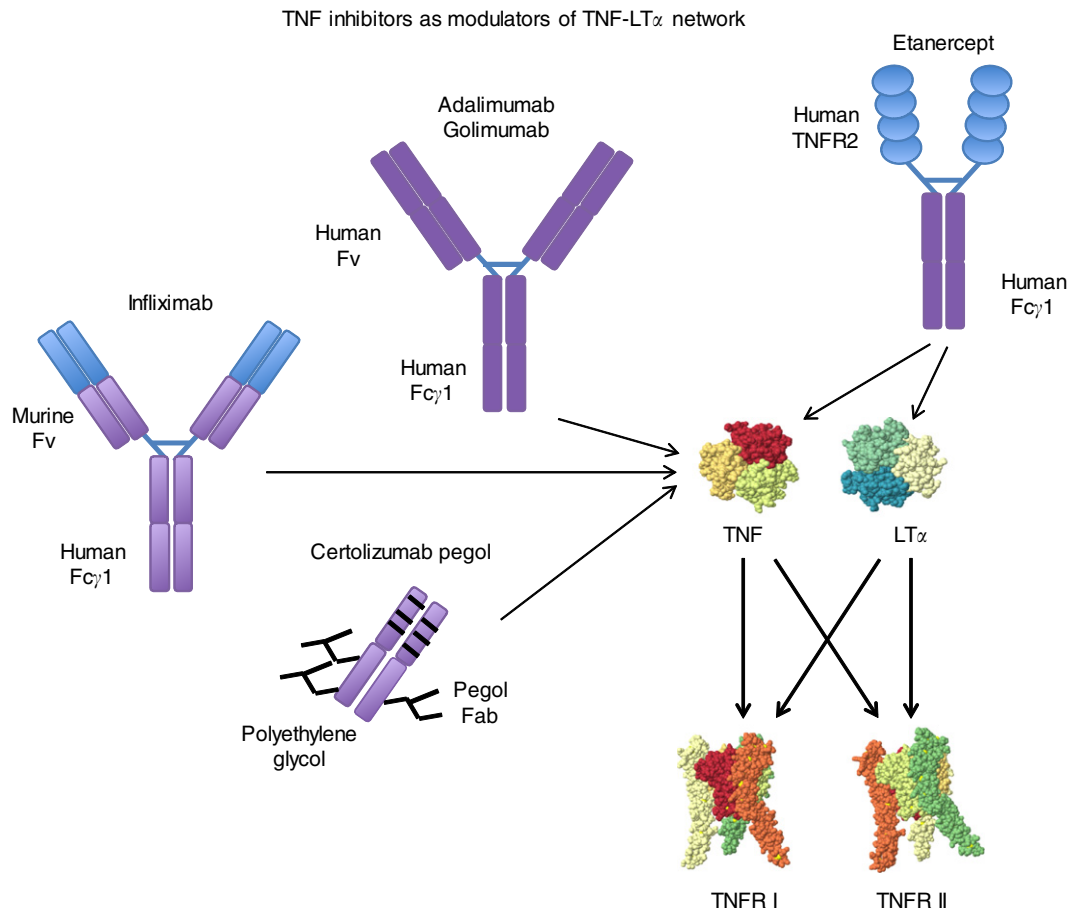


Figure 3 TNF inhibitors. The diagram depicts the biologics (antibody-based inhibitors) used as TNF inhibitors that are approved for use in patients with autoimmune diseases. Adalimumab is a bivalent fully human monoclonal antibody, whereas certolizumab is a monovalent antibody fragment with enhanced half-life due to chemical addition of polyethylene glycol moieties. In contrast, etanercept is a bivalent fusion protein containing TNFR2 and the Fc region of human IgG1. Etanercept blocks both TNF and LT α from engaging their receptors (see text for details).

partial or no response. The inflammatory bowel diseases, ulcerative colitis (limited to colon and bowel), and Crohn's disease (entire intestine) are chronic inflammatory conditions where TNF is disproportionately expressed when the disease is active. Adalimumab, a fully human anti-TNF antibody is efficacious in controlled trials in patients with moderate to severe Crohn's disease (Hanauer *et al.*, 2006). Interestingly, etanercept is not effective in treating IBD, which may be due to the inhibition of LT α limiting IgA production (Kruglov *et al.*, 2013).

Skin lesions of psoriatic patients contain elevated levels of TNF, and following treatment with TNF inhibitors, serum and skin levels of TNF decrease with remarkable clinical response (Bonifati and Ameglio, 1999; Schon and Boehncke, 2005). In clinical trials, etanercept and infliximab significantly inhibited disease activity in patients suffering with psoriatic arthritis.

Even though TNF inhibitors are generally safe because of their specificity, side effects do occur with the use of TNF inhibitors. Such side effects include Lupus-like syndrome, aplastic anemia, hepatotoxicity, interstitial lung disease, optic neuritis, and exacerbations of quiescent multiple sclerosis. TNF inhibition is associated with diminished host defense. A systematic review of randomized clinical trials using infliximab

and adalimumab found a dose-related increase of some malignancies in RA patients (Bongartz *et al.*, 2006).

See also: Cell Communication: Growth Factor Mediated Cell Signaling: Signaling and Function of Death Receptors of the Tumor Necrosis Factor Receptor Superfamily; Tumor Necrosis Factor Receptors: A Brief Digestion

References

- Baldwin Jr., A.S., 1996. The NF-kappa B and I kappa B proteins: New discoveries and insights. *Annual Review of Immunology* 14, 649–683.
- Banner, D.W., D'Arcy, A., Janes, W., *et al.*, 1993. Crystal structure of the soluble human 55 kd TNF receptor-human TNF beta complex: Implications for TNF receptor activation. *Cell* 73, 431–445.
- Bertrand, M.J., Milutinovic, S., Dickson, K.M., *et al.*, 2008. cIAP1 and cIAP2 facilitate cancer cell survival by functioning as E3 ligases that promote RIP1 ubiquitination. *Molecular Cell* 30, 689–700.
- Beutler, B.A., 2009. TLRs and innate immunity. *Blood* 113, 1399–1407.
- Bianchi, K., Meier, P., 2009. A tangled web of ubiquitin chains: Breaking news in TNF-R1 signaling. *Molecular Cell* 36, 736–742.

- Black, R.A., Rauch, C.T., Kozlosky, C.J., *et al.*, 1997. A metalloproteinase disintegrin that releases tumour-necrosis factor- α from cells. *Nature* 385, 729–733.
- Bodmer, J.L., Schneider, P., Tschopp, J., 2002. The molecular architecture of the TNF superfamily. *Trends in Biochemical Sciences* 27, 19–26.
- Bongartz, T., Sutton, A.J., Sweeting, M.J., *et al.*, 2006. Anti-TNF antibody therapy in rheumatoid arthritis and the risk of serious infections and malignancies: Systematic review and meta-analysis of rare harmful effects in randomized controlled trials. *JAMA* 295, 2275–2285.
- Bonifati, C., Ameglio, F., 1999. Cytokines in psoriasis. *International Journal of Dermatology* 38, 241–251.
- Cerami, A., Beutler, B., 1988. The role of cachectin/TNF in endotoxic shock and cachexia. *Immunology Today* 9, 28–31.
- Chang, L., Kamata, H., Solinas, G., *et al.*, 2006. The E3 ubiquitin ligase itch couples JNK activation to TNF α -induced cell death by inducing c-FLIP(L) turnover. *Cell* 124, 601–613.
- Croft, M., Benedict, C.A., Ware, C.F., 2013. Clinical targeting of the TNF and TNFR superfamily. *Nature Review Drug Discovery* 12, 147–168.
- Das, S., Cho, J., Lambert, I., *et al.*, 2005. Tpl2/cot signals activate ERK, JNK, and NF- κ B in a cell-type and stimulus-specific manner. *Journal of Biological Chemistry* 280, 23748–23757.
- Ea, C.K., Deng, L., Xia, Z.P., Pineda, G., Chen, Z.J., 2006. Activation of IKK by TNF α requires site-specific ubiquitination of RIP1 and polyubiquitin binding by NEMO. *Molecular Cell* 22, 245–257.
- Eliopoulos, A.G., Wang, C.C., Dumitru, C.D., Tschlis, P.N., 2003. Tpl2 transduces CD40 and TNF signals that activate ERK and regulates IgE induction by CD40. *EMBO Journal* 22, 3855–3864.
- Elliott, M.J., Maini, R.N., Feldmann, M., *et al.*, 1994. Randomised double-blind comparison of chimeric monoclonal antibody to tumour necrosis factor α (cA2) versus placebo in rheumatoid arthritis. *Lancet* 344, 1105–1110.
- Enari, M., Sakahira, H., Yokoyama, H., *et al.*, 1998. A caspase-activated DNase that degrades DNA during apoptosis, and its inhibitor ICAD. *Nature* 391, 43–50.
- Enesa, K., Zakkar, M., Chaudhry, H., *et al.*, 2008. NF- κ B suppression by the deubiquitinating enzyme Cezanne: A novel negative feedback loop in pro-inflammatory signaling. *Journal of Biological Chemistry* 283, 7036–7045.
- Feldmann, M., Brennan, F.M., Williams, R.O., *et al.*, 1992. Evaluation of the role of cytokines in autoimmune disease: The importance of TNF α in rheumatoid arthritis. *Progress in Growth Factor Research* 4, 247–255.
- Flynn, J.L., Goldstein, M.M., Chan, J., *et al.*, 1995. Tumor necrosis factor- α is required in the protective immune response against *Mycobacterium tuberculosis* in mice. *Immunity* 2, 561–572.
- Fukushima, T., Matsuzawa, S., Kress, C.L., *et al.*, 2007. Ubiquitin-conjugating enzyme Ubc13 is a critical component of TNF receptor-associated factor (TRAF)-mediated inflammatory responses. *Proceedings of the National Academy of Sciences of the United States of America* 104, 6371–6376.
- Gherzi, R., Lee, K.Y., Briata, P., *et al.*, 2004. A KH domain RNA binding protein, KSRP, promotes ARE-directed mRNA turnover by recruiting the degradation machinery. *Molecular Cell* 14, 571–583.
- Ghosh, S., May, M., Kopp, E., 1998. NF- κ B and REL proteins: Evolutionarily conserved mediators of immune responses. *Annual Review of Immunology* 16, 225–260.
- Gommerman, J.L., Browning, J.L., Ware, C.F., 2014. The lymphotoxin network: Orchestrating a type I interferon response to optimize adaptive immunity. *Cytokine & Growth Factor Reviews* 25, 139–145.
- Green, D.R., Galluzzi, L., Kroemer, G., 2011. Mitochondria and the autophagy-inflammation-cell death axis in organismal aging. *Science* 333, 1109–1112.
- Hanauer, S.B., Sandborn, W.J., Rutgeerts, P., *et al.*, 2006. Human anti-tumor necrosis factor monoclonal antibody (adalimumab) in Crohn's disease: The CLASSIC-I trial. *Gastroenterology* 130, 323–333. **quiz 591.**
- Hiscott, J., Nguyen, T.L., Arguello, M., Nakhaei, P., Paz, S., 2006. Manipulation of the nuclear factor- κ B pathway and the innate immune response by viruses. *Oncogene* 25, 6844–6867.
- Ho, P.K., Hawkins, C.J., 2005. Mammalian initiator apoptotic caspases. *FEBS Journal* 272, 5436–5453.
- Hsu, H., Huang, J., Shu, H.B., Baichwal, V., Goeddel, D.V., 1996. TNF-dependent recruitment of the protein kinase RIP to the TNF receptor-1 signaling complex. *Immunity* 4, 387–396.
- Huxford, T., Ghosh, G., 2009. A structural guide to proteins of the NF- κ B signaling module. *Cold Spring Harbor Perspectives in Biology* 1, a000075.
- Karin, M., Ben-Neriah, Y., 2000. Phosphorylation meets ubiquitination: The control of NF- κ B activity. *Annual Review of Immunology* 18, 621–663.
- Kovalenko, A., Chable-Bessia, C., Cantarella, G., *et al.*, 2003. The tumour suppressor CYLD negatively regulates NF- κ B signalling by deubiquitination. *Nature* 424, 801–805.
- Kruglov, A.A., Grivennikov, S.I., Kuprash, D.V., *et al.*, 2013. Nonredundant function of soluble LT α 3 produced by innate lymphoid cells in intestinal homeostasis. *Science* 342, 1243–1246.
- Lee, T.H., Huang, Q., Oikemus, S., *et al.*, 2003. The death domain kinase RIP1 is essential for tumor necrosis factor α signaling to p38 mitogen-activated protein kinase. *Molecular and Cellular Biology* 23, 8377–8385.
- Locksley, R.M., Killeen, N., Lenardo, M.J., 2001. The TNF and TNF receptor superfamilies: Integrating mammalian biology. *Cell* 104, 487–501.
- Mak, T.W., Yeh, W.C., 2002. Signaling for survival and apoptosis in the immune system. *Arthritis Research* 4 (Suppl. 3), S243–S252.
- Micheau, O., Tschopp, J., 2003. Induction of TNF receptor I-mediated apoptosis via two sequential signaling complexes. *Cell* 114, 181–190.
- Munker, R., Koeffler, P., 1987. In vitro action of tumor necrosis factor on myeloid leukemia cells. *Blood* 69, 1102–1108.
- Murphy, T.L., Murphy, K.M., 2010. Slow down and survive: Enigmatic immunoregulation by BTLA and HVEM. *Annual Review of Immunology* 28, 389–411.
- Muzio, M., Salvesen, G.S., Dixit, V.M., 1997. FLICE induced apoptosis in a cell-free system. Cleavage of caspase zymogens. *Journal of Biological Chemistry* 272, 2952–2956.
- Pfeffer, K., Matsuyama, T., Kundig, T.M., *et al.*, 1993. Mice deficient for the 55 kd tumor necrosis factor receptor are resistant to endotoxic shock, yet succumb to *L. monocytogenes* infection. *Cell* 73, 457–467.
- Phillips, K., Kedersha, N., Shen, L., Blackshear, P.J., Anderson, P., 2004. Arthritis suppressor genes TIA-1 and TTP dampen the expression of tumor necrosis factor α , cyclooxygenase 2, and inflammatory arthritis. *Proceedings of the National Academy of Sciences of the United States of America* 101, 2011–2016.
- Rothe, J., Lesslauer, W., Lotscher, H., *et al.*, 1993. Mice lacking the tumour necrosis factor receptor 1 are resistant to TNF-mediated toxicity but highly susceptible to infection by *Listeria monocytogenes*. *Nature* 364, 798–802.
- Ruddle, N.H., 2014. Lymphotoxin and TNF: How it all began—a tribute to the travelers. *Cytokine & Growth Factor Reviews* 25, 83–89.
- Sanjo, H., Zajonc, D.M., Braden, R., Norris, P.S., Ware, C.F., 2010. Allosteric regulation of the ubiquitin-NIK and ubiquitin-TRAF3 E3 ligases by the lymphotoxin- β receptor. *Journal of Biological Chemistry* 285, 17148–17155.
- Scholz, R., Sidler, C.L., Thali, R.F., *et al.*, 2010. Autoactivation of transforming growth factor β -activated kinase 1 is a sequential bimolecular process. *Journal of Biological Chemistry* 285, 25753–25766.
- Schon, M.P., Boehncke, W.H., 2005. Psoriasis. *New England Journal of Medicine* 352, 1899–1912.
- Schrofelbauer, B., Hoffmann, A., 2011. How do pleiotropic kinase hubs mediate specific signaling by TNFR superfamily members? *Immunological Reviews* 244, 29–43.
- Sedger, L.M., Shows, D.M., Blanton, R.A., *et al.*, 1999. IFN- γ mediates a novel antiviral activity through dynamic modulation of TRAIL and TRAIL receptor expression. *Journal of Immunology* 163, 920–926.
- Stamou, P., Kontoyannis, D.L., 2010. Posttranscriptional regulation of TNF mRNA: A paradigm of signal-dependent mRNA utilization and its relevance to pathology. *Current Directions in Autoimmunity* 11, 61–79.
- Sun, S.C., 2011. Non-canonical NF- κ B signaling pathway. *Cell Research* 21, 71–85.
- Takaes, G., Kishida, S., Hiyama, A., *et al.*, 2000. TAB2, a novel adaptor protein, mediates activation of TAK1 MAPKKK by linking TAK1 to TRAF6 in the IL-1 signal transduction pathway. *Molecular Cell* 5, 649–658.
- Tartaglia, L.A., Rothe, M., Hu, Y.F., Goeddel, D.V., 1993. Tumor necrosis factor's cytotoxic activity is signaled by the p55 TNF receptor. *Cell* 73, 213–216.
- Tartaglia, L.A., Weber, R.F., Figari, I.S., *et al.*, 1991. The two different receptors for tumor necrosis factor mediate distinct cellular responses. *Proceedings of the National Academy of Sciences of the United States of America* 88, 9292–9296.
- Tewari, M., Dixit, V.M., 1995. Fas- and tumor necrosis factor-induced apoptosis is inhibited by the poxvirus crmA gene product. *Journal of Biological Chemistry* 270, 3255–3260.
- Thome, M., Schneider, P., Hofmann, K., *et al.*, 1997. Viral FLICE-inhibitory proteins (FLIPs) prevent apoptosis induced by death receptors. *Nature* 386, 517–521.
- Tokunaga, F., Sakata, S., Saeki, Y., *et al.*, 2009. Involvement of linear polyubiquitylation of NEMO in NF- κ B activation. *Nature Cell Biology* 11, 123–132.
- Tracey, K.J., Fong, Y., Hesse, D.G., *et al.*, 1987. Anti-cachectin/TNF monoclonal antibodies prevent septic shock during lethal bacteraemia. *Nature* 330, 662–664.

- Tuohy, V.K., Yu, M., Yin, L., *et al.*, 1998. The epitope spreading cascade during progression of experimental autoimmune encephalomyelitis and multiple sclerosis. *Immunological Reviews* 164, 93–100.
- Vince, J.E., Pantaki, D., Feltham, R., *et al.*, 2009. TRAF2 must bind to cellular inhibitors of apoptosis for tumor necrosis factor (tnf) to efficiently activate nf- κ b and to prevent tnf-induced apoptosis. *Journal of Biological Chemistry* 284, 35906–35915.
- Wajant, H., Pfizenmaier, K., Scheurich, P., 2003. Tumor necrosis factor signaling. *Cell Death & Differentiation* 10, 45–65.
- Walczak, H., 2011. TNF and ubiquitin at the crossroads of gene activation, cell death, inflammation, and cancer. *Immunological Reviews* 244, 9–28.
- Wang, X., Destrumant, A., Tournier, C., 2007. Physiological roles of MKK4 and MKK7: Insights from animal models. *Biochimica et Biophysica Acta* 1773, 1349–1357.
- Ware, C.F., 2005. Network communications: Lymphotoxins, LIGHT, and TNF. *Annual Review of Immunology* 23, 787–819.
- Ware, C.F., Sedy, J.R., 2011. TNF superfamily networks: Bidirectional and interference pathways of the herpesvirus entry mediator (TNFSF14). *Current Opinion in Immunology* 23, 627–631.
- Wertz, I.E., O'Rourke, K.M., Zhou, H., *et al.*, 2004. De-ubiquitination and ubiquitin ligase domains of A20 downregulate NF- κ B signalling. *Nature* 430, 694–699.
- Wiens, G.D., Glenney, G.W., 2011. Origin and evolution of TNF and TNF receptor superfamilies. *Developmental and Comparative Immunology* 35, 1324–1335.
- Williams, R.O., Feldmann, M., Maini, R.N., 1992. Anti-tumor necrosis factor ameliorates joint disease in murine collagen-induced arthritis. *Proceedings of the National Academy of Sciences of the United States of America* 89, 9784–9788.

Further Reading

- Rieser, E., Cordier, S.M., Walczak, H., 2013. Linear ubiquitination: A newly discovered regulator of cell signalling. *Trends in Biochemical Sciences* 38, 94–102.
- Ubersax, J.A., Ferrell Jr., J.E., 2007. Mechanisms of specificity in protein phosphorylation. *Nature Reviews Molecular Cell Biology* 8, 530–541.

Relevant Websites

- www.biogps.org
BioGPS.
- <http://www.genenames.org>
HUGO Gene Nomenclature Committee.
- www.immgen.org
Immunological Genome Project.

Caspases

J Puccini and S Kumar, University of South Australia, Adelaide, SA, Australia

© 2016 Elsevier Inc. All rights reserved.

Introduction

Caspases are a family of aspartate-specific, cysteine proteases. Mammalian interleukin-1 β -converting enzyme (ICE, later remained caspase-1) was the first member of the caspase family (Cerretti *et al.*, 1992; Thornberry *et al.*, 1992). As the name suggests, ICE/caspase-1 is required for the enzymatic processing and secretion of the pro-inflammatory cytokine IL-1 β . The product of the *Caenorhabditis elegans* cell death gene *ced-3* was shown to have a high degree of homology with ICE (Miura *et al.*, 1993; Yuan *et al.*, 1993). Given the essential role of *ced-3* in programmed cell death (PCD) in *C. elegans*, a functional link between caspase-1 and apoptosis was proposed (Miura *et al.*, 1993; Yuan *et al.*, 1993). Shortly after caspase-1, the neural precursor cell expressed developmentally downregulated gene 2 (*Nedd2*) sharing homology with *ced-3* was identified (Kumar *et al.*, 1992). Like *ced-3*, *Nedd2* (renamed caspase-2) encodes a cysteine-protease which, when overexpressed in mammalian cells, induced cell death by apoptosis (Kumar *et al.*, 1994). Given the main function of caspase-1 in cytokine maturation, caspase-2 was initially thought to functionally emulate *ced-3* in mammalian cell death (Kumar *et al.*, 1994; Miura *et al.*, 1993).

Since the initial discovery, 18 mammalian caspases have been identified with 12 encoded in the human genome (Eckhart *et al.*, 2008). Over the last two decades the structure, function, and enzymatic properties of the caspase family has been extensively characterized in mammals and other model organisms including *Drosophila melanogaster*. Much of the knowledge of the role of caspases in developmental PCD has come from *C. elegans*, as well as *D. melanogaster* (Kumar, 2007).

Structure, Classification, and Non-Apoptotic Biological Roles

Based on their structural and functional properties, caspases have been classified into three groups. Caspase-1, -4, and -5 are known as the inflammatory caspases and are involved in cytokine processing and coordination of inflammatory response (McIlwain *et al.*, 2013). Caspase-11 (the mouse orthologue of human caspase-4) and murine caspase-12 are also members of the inflammatory caspase family (McIlwain *et al.*, 2013). However, the evolutionary function of caspase-12 is not conserved as the human orthologue is a pseudo-caspase lacking a catalytic domain (Lamkanfi *et al.*, 2004). The second group of caspases are known as the apoptotic caspases and function as the initiators (caspase-2, -8, -9, and -10) or effectors (caspase-3, -6, and -7) of apoptosis (Kumar, 2007; McIlwain *et al.*, 2013). Most caspases are ubiquitously expressed in a wide range of tissues with the exception of some, such as caspase-14 whose expression is confined to keratinocytes where it is believed to function in differentiation (McIlwain *et al.*, 2013). Additional caspase family members,

caspase-13 (the bovine homolog of human caspase-4), -15, -16, -17, and -18 have also been identified in vertebrates, including various mammalian lineages (Eckhart *et al.*, 2008). Of these, only caspase-16 (which is most similar to caspase-14 in terms of sequence homology) is conserved in humans (Eckhart *et al.*, 2008). The biological functions of all mouse caspases (except caspase-16) have been interrogated by gene knockout models (Table 1). This has revealed an essential role for many caspases in development as well as apoptotic and non-apoptotic processes. Some caspases have been shown to be able to compensate for each other's functions; therefore double knockout strains have also been generated for some caspases to determine functional redundancy in specific biological processes (Table 1).

Structural Properties

Caspases are synthesized as enzymatically inactive precursors or zymogens, whereas a fully mature caspase enzyme is a tetramer comprising two large and two small subunits (McIlwain *et al.*, 2013). Pro-caspase zymogens consist of a C-terminal catalytic domain and an N-terminal prodomain (Figure 1; Kumar, 2007). The catalytic domains of caspases (where the active site resides) are comprised of a small and large subunit, both of which are required to achieve an enzymatically active conformation (Pop and Salvesen, 2009). The catalytic cysteine residue resides within the large subunit while the small subunit contains specific residues that are essential for structuring of the substrate-binding pockets and therefore are important for substrate recognition (Pop and Salvesen, 2009).

The prodomains of caspases differ with respect to their structure and function (Kumar, 2007; Pop and Salvesen, 2009). Apoptotic effector caspases (-3, -6, and -7) and caspase-14 contain very short prodomains, typically consisting of 20–30 amino acids (Figure 1). In contrast, the inflammatory and apoptotic initiator caspases have large (up to 220 amino acids) prodomains that contain homotypic protein–protein interaction motifs such as the caspase activation and recruitment domain (CARD) in caspase-1, -2, -4, -5, and -9 and the death effector domains (DEDs) in caspase-8 and -10 (Figure 1). These protein–protein interaction motifs of initiator caspases mediate their dimerization and recruitment into caspase activation complexes (see section Caspase Activation and Function in Cell Death and Inflammation; Kumar, 2007).

Non-Apoptotic Functions

Caspases were originally grouped based on studies demonstrating their functional role in specific biological process. However, many caspases have now been shown to be involved in processes that are distinct from their originally defined functions, blurring the lines between these functional classes (Miura, 2012). In addition to their well-characterized roles in apoptosis and cytokine maturation, caspases have been shown

Table 1 Phenotypes of caspase knockout mice

Caspase knockout	Mutant phenotypes	References
Caspase-1	Normal development, increased susceptibility to bacterial/viral infection, resistant to endotoxin shock, and impaired processing and secretion of inflammatory cytokines	Kuida <i>et al.</i> , 1995; Li <i>et al.</i> , 1995; Thomas <i>et al.</i> , 2009
Caspase-2	Normal development, increased number of oocytes, premature ageing-related traits, and slight growth retardation	Bergeron <i>et al.</i> , 1998; Shalini <i>et al.</i> , 2012
Caspase-3	Perinatal lethality (incomplete penetrance) in a mixed genetic background, neuronal cell death defects (brain hyperplasia), some cell types resistant to specific apoptotic stimuli	Kuida <i>et al.</i> , 1996; Woo <i>et al.</i> , 1998
Caspase-6	Normal development and no defects in apoptosis	McIlwain <i>et al.</i> , 2013
Caspase-7	Normal development and no defects in apoptosis	Lakhani <i>et al.</i> , 2006
Caspase-8	Embryonic lethality (complete penetrance), defects in death receptor-mediated extrinsic apoptosis, and impaired development of cardiac muscle and myeloid/lymphoid lineages	Salmena <i>et al.</i> , 2003; Varfolomeev <i>et al.</i> , 1998
Caspase-9	Perinatal lethality (incomplete penetrance), neuronal cell death defects (brain hyperplasia), and some cell types resistant to specific apoptotic stimuli	Hakem <i>et al.</i> , 1998; Kuida <i>et al.</i> , 1998
Caspase-11	Normal development, resistant to LPS-induced shock, impaired caspase-1-activation and IL-1 secretion	Wang <i>et al.</i> , 1998
Caspase-12	Normal development, resistant to sepsis and ER stress-induced apoptosis	Nakagawa <i>et al.</i> , 2000
Caspase-14	Normal development, impaired terminal differentiation of keratinocytes and increased sensitivity to UVB radiation-induced photodamage	Denecker <i>et al.</i> , 2007
Caspase-2/Caspase-9	Same as <i>Casp9</i> ^{-/-}	Marsden <i>et al.</i> , 2004
Caspase-3/Caspase-7	Perinatal lethality (complete penetrance) and resistant to intrinsic and extrinsic apoptosis	Lakhani <i>et al.</i> , 2006

Note: A summary of the development and phenotypes of caspase-deficient mouse strains. Note, there are no mouse homologs of caspase-5 and -10 and caspase-11 is the mouse homolog of human caspase-4.

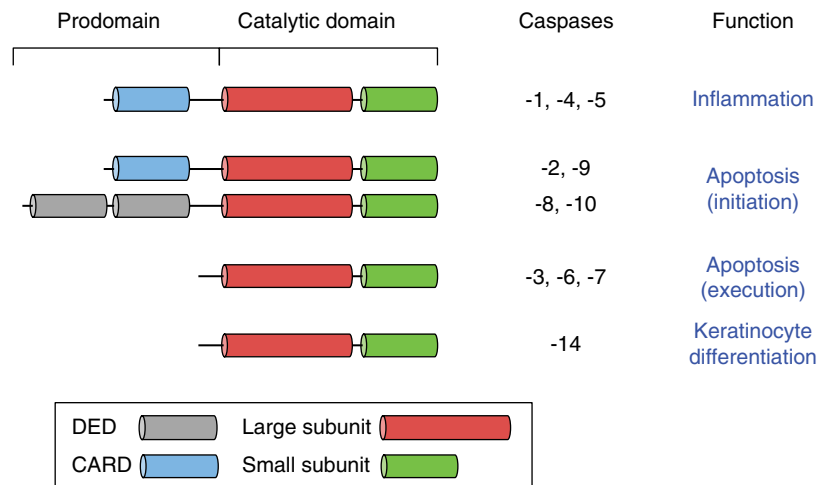


Figure 1 Domain structure of human caspases. Schematic representation of the architecture of human caspases showing the prodomains and catalytic domain. The catalytic domain is comprised of a small and large subunit. Long prodomain-containing caspases contain protein–protein interaction motifs such as a CARD or DED. The caspases have been grouped according to their biological functions.

to participate in several biological processes via both cell autonomous and non-cell autonomous mechanisms (Miura, 2012). Studies in model organisms such as mice and *D. melanogaster* using genetic or pharmacological ablation of caspase activity have revealed these novel functions of the apoptotic caspases. For example, mice with T lymphocyte-specific disruption of caspase-8 display a reduction in mature T

lymphocytes and an attenuated immune response to viral infection (Salmena *et al.*, 2003). Interestingly, these defects in genetically modified mice are also recapitulated in humans harboring caspase-8 loss-of-function mutations who display attenuated activation and proliferation of T and B lymphocytes leading to immunodeficiency (Chun *et al.*, 2002). Moreover, perturbed proliferation has also been observed following

developmental apoptosis or wound healing upon inhibition of caspase activity (Huh *et al.*, 2004). Therefore, in addition to regulation of normal cell proliferation, caspases appear to be involved in compensatory proliferation under certain conditions.

More recently, it was shown that blocking caspase activity in *D. melanogaster* disrupts the pruning of neuronal dendrites and axons, a process which is critical for the formation of neural circuits (Williams *et al.*, 2006). These observations have also been extended in mouse models which demonstrate that caspases also appear to regulate synaptic plasticity in mammals (Nikolaev *et al.*, 2009). Furthermore, several reports have shown that caspases are also involved in the process of terminal differentiation in various progenitor cell types including myoblasts and monocytes (Fernando *et al.*, 2002; Sordet *et al.*, 2002). These studies demonstrate that caspases are important for programing and coordinating developmental events independently of their apoptotic functions.

Activation Mechanisms

Procaspase zymogens require enzymatic processing in order to achieve a mature, fully active tetrameric conformation. This requirement facilitates stringent regulation of caspase activity which is essential for preventing inappropriate triggering of the caspase cascade. Depending on the presence of protein–protein interaction motifs in their prodomain, different caspases are activated by distinct biochemical mechanisms.

Initiator Caspase Activation

Long prodomain-containing initiator caspases exist as inactive zymogen monomers in unstressed cells (Chai and Shi, 2014; Pop and Salvesen, 2009). The current model for activation of

initiator caspases is proximity-induced dimerization (Chai and Shi, 2014). According to this model, dimerization of initiator caspase zymogens occurs when they are brought into close proximity via recruitment into multimeric activation complexes (see section Caspase Activation and Function in Cell Death and Inflammation) resulting in high-local concentrations of zymogen monomers (Chai and Shi, 2014; Pop and Salvesen, 2009). These conditions favor dimerization which is thought to stimulate the catalytic activity and subsequent autoproteolytic cleavage between the small and large subunits of the catalytic domain (Figure 2; Chai and Shi, 2014; Pop and Salvesen, 2009). Homodimerization is believed to be the initial event that triggers activation of initiator caspases that mediate downstream signaling to effector molecules (Chai and Shi, 2014; Pop and Salvesen, 2009). Following autoproteolytic processing, the prodomains of initiator caspases are also removed by proteolysis to promote a mature enzyme conformation with full catalytic activity (Figure 2; Chai and Shi, 2014; Pop and Salvesen, 2009). Although the induced-proximity model is consistent with most biochemical data, it offers very little insight regarding the molecular mechanisms of initiator caspase activation under physiological conditions.

Effector Caspase Activation

In contrast to initiator procaspases, effector procaspases lacking long prodomains exist as inactive dimers in unstressed cells and have a much more well-characterized activation mechanism (Chai and Shi, 2014; Pop and Salvesen, 2009). Because they do not contain protein–protein interaction motifs, effector caspases rely on proteolytic cleavage by upstream initiator caspase for their activation. As with initiator caspases, interchain cleavage within the catalytic domain of effector caspases induces conformational rearrangements in the active site which greatly increases their catalytic activity (Figure 2; Chai and Shi, 2014; Pop

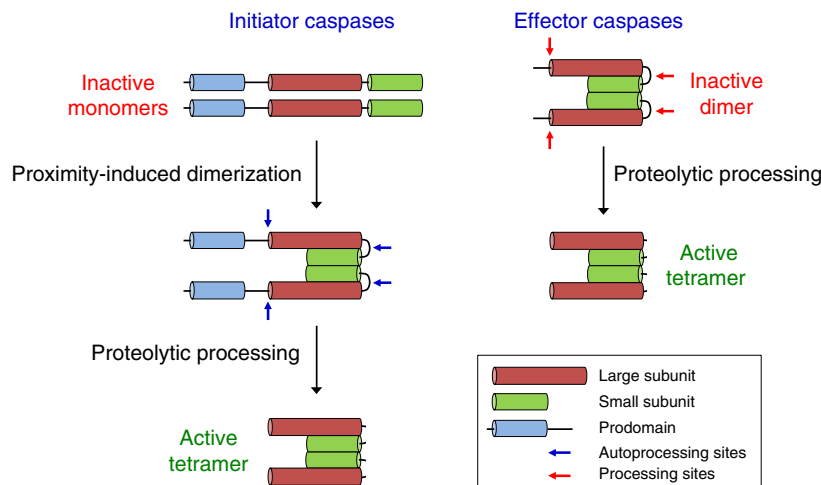


Figure 2 Caspase activation mechanisms. Depending on the structure of their prodomains, caspases are activated by different mechanisms. Long prodomain-containing initiator caspases exist as inactive monomers and are activated by proximity-induced dimerization which can sometimes be facilitated by adaptor proteins. Upon dimerization, interchain autocatalytic cleavage generates a catalytically activate enzyme. Further processing leads to the removal of the prodomain and formation of a mature tetramer assembled by the homodimerization of heterodimers. Effector caspases have short prodomains and exist as inactive dimers. They are activated by cleavage (mediated by activated initiator caspases) between the small and large subunits of the catalytic domain. As with initiator caspases, enzymatically active effector caspases are comprised of tetramers derived from two large and two small subunits.

and Salvesen, 2009). This is then followed by trimming of the prodomains of effector caspases to produce a fully mature enzyme conformation (Figure 2; Chai and Shi, 2014; Pop and Salvesen, 2009). Biochemical observations suggest that this final maturation event involving removal of prodomains is distinct from activation and is dispensable for acquisition of full catalytic activity (Chai and Shi, 2014; Pop and Salvesen, 2009). Moreover, crystal structures of mature initiator and effector caspases have revealed that enzymatically active caspase molecules exist as tetramers consisting of heterodimers of the small and large subunits of the catalytic domain, which assemble into homodimers (Figure 2; Chai and Shi, 2014; Pop and Salvesen, 2009).

Substrate Specificity and Targets

Biochemical Basis of Substrate Specificity

Caspases are defined as proteases that specifically cleave tetrapeptide substrates (P4-P3-P2-P1) C-terminal of a conserved aspartate residue (in the P1 position) using a cysteine-dependent catalytic mechanism (Poreba *et al.*, 2013). This stringent requirement for an aspartate residue in the P1 position of substrates is a characteristic biochemical property of the caspase family (Poreba *et al.*, 2013). All mammalian caspases also share a strong preference for glutamate residues in the P3 position of their substrates (Pop and Salvesen, 2009; Poreba *et al.*, 2013). Each amino acid in the tetrapeptide substrate (P1-P4) interacts with specific amino acid side chains that make up the corresponding substrate-binding pockets within the active site (named S4-S3-S2-S1) (Poreba *et al.*, 2013). The biochemical properties of these substrate-binding pockets define the specificity of each caspase. The S1 and S3 binding pockets are highly conserved between all caspases, with invariant amino acid residues in these pockets responsible for binding the conserved aspartate and glutamate residues in the P1 and P3 positions, respectively (Pop and Salvesen, 2009; Poreba *et al.*, 2013). Specificity in substrate recognition is endowed by virtue of the differential preferences of the P2 and P4 residues which is due to variability in the amino acids that make up S2 and S4 binding pockets (Pop and Salvesen, 2009; Poreba *et al.*, 2013).

Caspases belonging to the same functional classes share similar substrate recognition sequences (Table 2). This overlap in substrate preference between functionally related caspases may reflect functional redundancy between caspases that participate in the same biological processes. Given their overlapping substrate specificities, it is not surprising that caspases display a large degree of functional redundancy and are able to compensate for each other's functions (McIlwain *et al.*, 2013; Poreba *et al.*, 2013). Observations from caspase knockout mice clearly demonstrate this principle. For example, the caspase-3 substrate PARP is efficiently cleaved in caspase-3-deficient cells, suggesting that caspase-3 functions redundantly with other effector caspases, such as caspase-7 (Walsh *et al.*, 2008).

The optimal substrate recognition motif for the inflammatory caspases and caspase-14 is (W/L)EHD (Table 2; Benkova *et al.*, 2009; Mikolajczyk *et al.*, 2004). In contrast to these caspases which prefer hydrophilic or aromatic residues in the P4 position, apoptotic effectors caspase-3 and -7 have a

Table 2 Caspase substrate recognition motifs

Caspase	Substrate consensus motif (P5-P4-P3-P2-P1)	Known functions
1, 4, 5	(W/L) E X D	Inflammation
8, 9, 10, (6)	(I/V/L) E X D	Apoptosis initiation (execution)
3, 7	D E X D	Apoptosis execution
2	V D V A D	Apoptosis initiation/execution
14	W E H D	Keratinocyte differentiation

Note: The optimal substrate recognition sequence for human caspases, grouped according to their biological functions. Note the invariant aspartate residue in the P1 position (shown in bold). Proteolysis of substrates occurs immediately C-terminal of the P1 residue.

striking specificity for an aspartate residue in the P4 position of substrates (Table 2; Benkova *et al.*, 2009). Therefore, the optimal substrate recognition motif for caspase-3 and -7 is DEXD (Benkova *et al.*, 2009). Despite being functionally related to other effector caspases (-3 and -7), the optimal recognition sequence of caspase-6 is relatively less specific and resembles that of initiator caspases (-8, -9, and -10) which prefer the consensus motif (I/L/V)EXD (Table 2; Benkova *et al.*, 2009). Consistent with the biochemical properties elucidated from crystal structures, caspase-6 and the initiator caspases favor small, hydrophobic residues in the P4 position and show more flexibility in the P2 position, tolerating both hydrophobic and hydrophilic residues (Benkova *et al.*, 2009; Poreba *et al.*, 2013). Interestingly, the preferred substrate recognition motif for caspase-2 is unique among the caspase family. Unlike the other members of the caspase family which specifically cleave tetrapeptide motifs, caspase-2 has a strong preference for pentapeptide substrates (P5-P4-P3-P2-P1) with the consensus sequence VDVA D (Benkova *et al.*, 2009). *In vitro* biochemical studies show that caspase-2 has a weak affinity for DEVD substrates and that addition of an amino acid residue in the P5 position greatly enhances its catalytic activity (Benkova *et al.*, 2009; Tang *et al.*, 2011).

Caspase Substrates

Despite their specificity, a large degree of promiscuity exists between members of the caspase family. Overlapping substrate repertoires has hampered efforts to precisely define the contributions of individual caspases to specific biological processes (Poreba *et al.*, 2013). Furthermore, differences in endogenous expression levels and intrinsic catalytic activity of individual caspases have also been major obstacles to the identification of specific physiological caspase substrates *in vivo*.

Several hundred caspase substrates have been identified through biochemical and proteomic approaches. A list of caspase substrates can be found in databases such as CASBAH, CASVM, and PROSPER (see links to these databases at the end of this article). Caspase-3 is the most abundant effector caspase in cells and is responsible for cleaving most proteins that are associated with the onset of apoptosis (Poreba *et al.*, 2013). Cleavage by caspases can lead to either

inactivation or activation of substrates, however it remains unclear which of these cleaved proteins are *bona fide* caspase substrates with physiological relevance versus those that are innocent bystanders or artefacts of *in vitro* experimental systems. The most obvious targets of the initiator caspases are effector caspases, which, as discussed above require processing for activation (Pop and Salvesen, 2009). Cleavage of some caspase substrates is clearly linked to the characteristic morphological changes that take place in a cell undergoing apoptosis. For example, inhibitor of caspase activated DNase (ICAD) cleavage by caspase-3 permits activation of CAD which mediates cleavage of inter-nucleosomal linker DNA, producing characteristic DNA laddering that has been used as a marker for apoptosis for over 30 years (Enari *et al.*, 1998). In a recent finding, caspase-mediated cleavage and inactivation of the phospholipid flippase ATP11C was shown to be required for phosphatidylserine exposure which is necessary as an 'eat me' signal for engulfment of the apoptotic cells by macrophages (Segawa *et al.*, 2014). On the other hand, the inflammatory caspases mostly target pro-inflammatory cytokines for cleavage, so their activity is directly linked to their function in cytokine maturation and secretion (McIlwain *et al.*, 2013).

Caspase Activation and Function in Cell Death and Inflammation

Unlike other proteases that are involved in protein turnover and cleave a large number of proteins in a nonspecific manner, caspases function as important signaling enzymes that cleave specific repertoires of substrates which mediate downstream signaling events. Caspases have clearly defined and well-established functions in PCD and inflammation (McIlwain *et al.*, 2013). These processes are primarily controlled by upstream apical caspases whose activity is in turn regulated by specific activation complexes that assemble in response to specific stimuli (Chai and Shi, 2014). These activation complexes contain adapter proteins which recruit long prodomain-containing initiator caspases via conserved homotypic protein-protein interaction domains (see section Structure, Classification and Non-Apoptotic Biological Roles; Figure 1; Chai and Shi, 2014). These adaptor proteins serve as important molecular scaffolds that facilitate acquisition of an enzymatically active conformation by promoting autoproteolytic processing (see section Activation Mechanisms; Figure 2; Chai and Shi, 2014). The exact molecular details of the binding kinetics and stoichiometry of caspase activation platforms is still an active area of research. Hence, only the current dogma regarding assembly and function of these activation complexes is discussed here.

The Inflammasome (Inflammatory Response)

The inflammatory caspases (-1, -4, and -5) are important regulators of the inflammatory response, which plays a critical role in defense against infectious diseases as well autoimmunity and cancer (McIlwain *et al.*, 2013). In addition to noninfectious stimuli (e.g., tissue injury and environmental irritants), the inflammatory response can also be triggered by a vast range of infectious stimuli such as toxins and nucleic acids produced by

pathogenic bacteria, viruses, and fungi (McIlwain *et al.*, 2013). These pro-inflammatory stimuli are sensed by NOD-like receptor (NLR) proteins which belong to the family of highly conserved pattern recognition receptors (PRRs) that form an essential axis of the innate immune system (McIlwain *et al.*, 2013). To date, 23 NLR family members have been identified in the human genome, each of which recognizes specific repertoires of host- and pathogen-derived molecules referred to as damage-associated molecular patterns (DAMPs) and pathogen-associated molecular patterns (PAMPs), respectively (McIlwain *et al.*, 2013). In addition to NLR family members, other specialized PRRs such as the cytosolic double-stranded DNA sensor, AIM2 (absent in melanoma) can also initiate the inflammatory response (Hornung *et al.*, 2009). These specific sensory proteins initiate the assembly of multiprotein complexes called inflammasomes that activate caspase-dependent inflammatory signaling (Figure 3; Martinon and Tschopp, 2005).

Inflammasomes are multimeric protein complexes that serve as the activation platform for caspase-1, -4, and -5 in response to pro-inflammatory stimuli (Martinon *et al.*, 2002). Of the inflammatory caspases, caspase-1 is the most extensively studied in inflammasome signaling. Following recognition and binding to specific PAMPs or DAMPs, NLR or AIM2 monomers undergo conformational changes which promote their oligomerization and association with CARD-containing adaptor proteins such as apoptosis-associated speck-like protein containing a CARD (ASC) (Martinon *et al.*, 2002). This then facilitates CARD-dependent recruitment of procaspase-1 monomers, leading to their dimerization and autoproteolytic activation (Figure 3; Franchi *et al.*, 2009; Martinon *et al.*, 2002). The activity of caspase-1 is controlled by CARD-containing regulatory proteins such as COP and ICEBERG which are able to bind and sequester procaspase-1 monomers, inhibiting dimerization and autoproteolytic activation (Humke *et al.*, 2000; Lee *et al.*, 2001). Once activated, caspase-1 cleaves pro-IL-1 β and pro-IL-18, promoting their maturation and secretion (Figure 3; Franchi *et al.*, 2009; Martinon *et al.*, 2002). Upon secretion, these pro-inflammatory cytokines recruit and activate immune effector cells that drive the inflammatory response. Therefore, caspase-1-activation is crucial for initiating inflammation in response to injury or infection. The role of caspase-1 in innate immunity is further exemplified by caspase-1-deficient mice which show increased susceptibility to viral infection (Table 1; Thomas *et al.*, 2009).

Specificity in the inflammatory response is achieved by engagement of different sensory proteins (e.g., NLR family members) that recognize specific PAMPs or DAMPs and subsequently direct the assembly of different types of inflammasome complexes (McIlwain *et al.*, 2013). Therefore, the specific type of inflammatory stimulus a cell encounters, in addition to cell-specific expression of sensors and adapters, dictates which inflammasome complexes are assembled and therefore the downstream pathways that are engaged.

The Apoptosome (Intrinsic Apoptosis)

There are two distinct pathways that lead to apoptosis, namely the intrinsic (mitochondrial) pathway and the extrinsic (death receptor) pathway (see section The Death-Induced Signaling

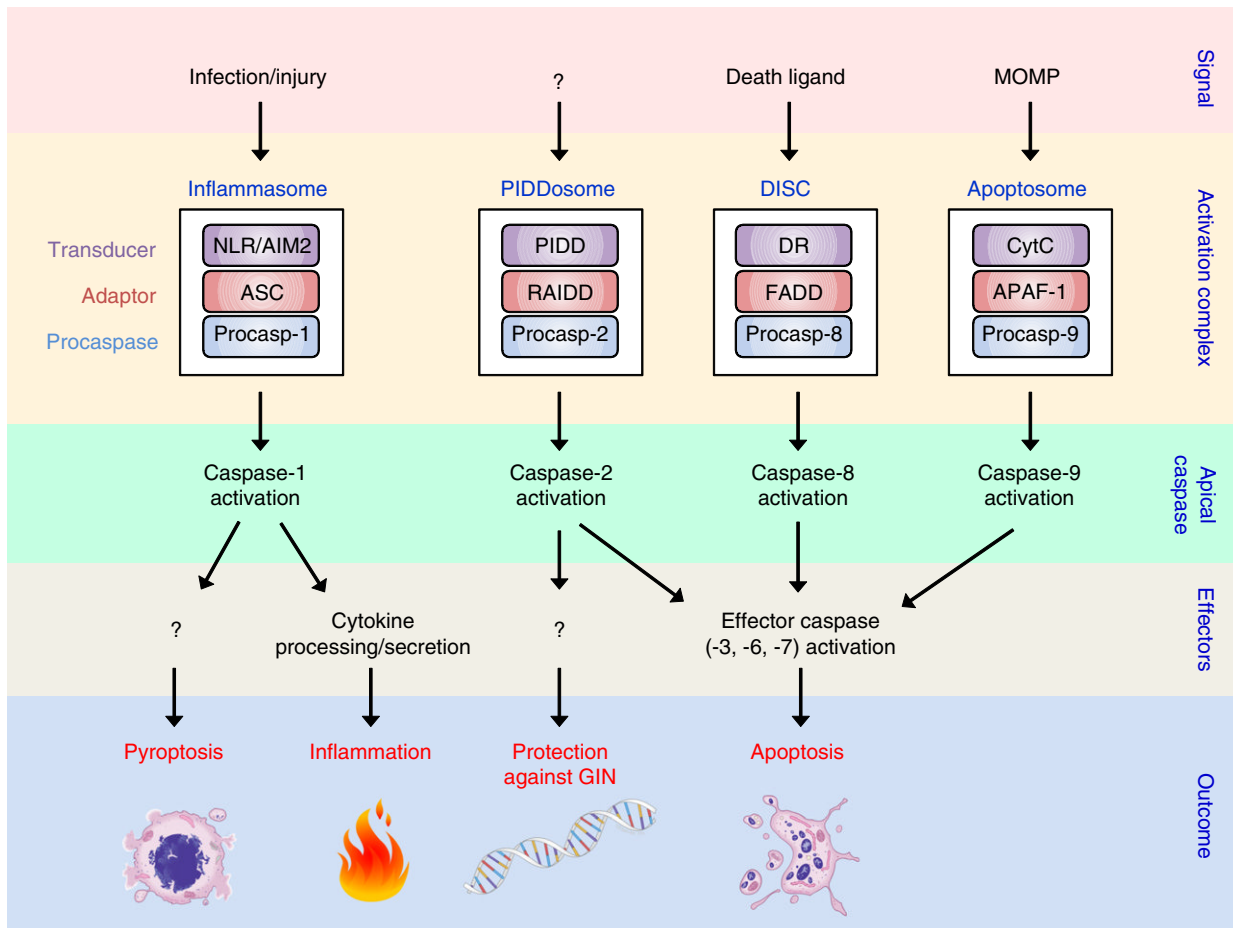


Figure 3 Caspase activation platforms. Various multiprotein platforms have been described which promote activation of specific long prodomain-containing caspases. These multiprotein complexes contain a transducer that senses specific stimuli and adaptor proteins which recruit procaspase molecules via protein–protein interaction motifs in their prodomains. Intracellular stresses such as DNA damage and hypoxia trigger MOMP and subsequently activate the intrinsic apoptosis pathway and apoptosome formation while host and pathogen-derived pro-inflammatory stimuli trigger inflammasome formation. The extrinsic pathway of apoptosis is engaged upon DISC-mediated activation of caspase-8. Caspase-2 has been shown to be activated by the PIDDosome, however caspase-2 activation is also known to occur in the absence of PIDDosome. Assembly of these activation platforms and subsequent activation of apical caspases and downstream effectors ultimately lead to initiation of the inflammatory response or cell death.

Complex (Extrinsic Apoptosis); McIlwain *et al.*, 2013). The intrinsic pathway is activated by cellular stresses such as DNA damage and reactive oxygen species (McIlwain *et al.*, 2013). The stress signals trigger mitochondrial outer membrane permeabilization (MOMP), which causes the release of mitochondrial pro-apoptotic molecules such as cytochrome c (Kluck *et al.*, 1997). When released, cytochrome c associates with the adaptor protein apoptotic protease activating factor 1 (APAF-1) which exists in an auto-inhibited state in unstressed cells (Jiang and Wang, 2000). Analogous to formation of the NLR-ASC-caspase-1 inflammasome, this binding event causes a conformational change via an ATP-dependent mechanism that facilitates APAF-1 oligomerization and formation of a heptameric complex known as the apoptosome, which functions as an activation platform for procaspase-9 (Zou *et al.*, 1999; Figure 3). Binding of procaspase-9 monomers to the apoptosome (via CARD–CARD interactions between procaspase-9 and APAF-1) stimulates their proteolytic activity and subsequent activation by autoprocessing (Zou *et al.*, 1999).

Activated caspase-9 then processes the downstream effector caspases (caspase-3, -6, and -7), which in turn cleave other effector proteins (Figure 3; Zou *et al.*, 1999).

Activation of apoptotic caspases is tightly controlled by a network of pro-apoptotic and anti-apoptotic regulators that oppose each other's function. The balance between the activities of these regulators determines a cell's decision between survival and death. The pro-apoptotic Bcl-2 (B cell lymphoma 2) family proteins Bcl-2 homologous antagonist/killer (BAK) and Bcl-2-associated protein X (BAX) regulate MOMP which in turn controls cytochrome c release (McIlwain *et al.*, 2013). In response to intrinsic apoptotic stimuli, BAX and BAK oligomerize and form pores in the outer mitochondrial membrane (Wei *et al.*, 2001). Under basal conditions, BAK/BAX function is antagonized by anti-apoptotic members of the Bcl-2 family such as BCL-2 and BCL-XL, thereby inhibiting MOMP and blocking apoptosis (Fletcher *et al.*, 2008). Inhibition of apoptosis is also achieved by BCL-2-mediated neutralization of pro-apoptotic BH3 (Bcl-2 homology 3)-only proteins that

can directly activate BAK/BAX (Kuwana *et al.*, 2005). BH3-only proteins such as p53-upregulated modulator of apoptosis (PUMA) also regulate MOMP by sequestering anti-apoptotic Bcl-2 family members, thereby relieving inhibition of BAK/BAX (Chen *et al.*, 2005). Therefore, in mammalian cells, apoptosome assembly and subsequent caspase-9 activation is predominantly controlled by upstream Bcl-2 proteins that regulate MOMP.

In response to apoptotic stimuli, pro-apoptotic proteins can also promote caspase activation by inhibiting anti-apoptotic regulators and activating other pro-apoptotic effectors. One such pro-apoptotic factor released from mitochondria following MOMP is direct IAP-binding protein with low PI (DIABLO) (Verhagen *et al.*, 2000). DIABLO sequesters anti-apoptotic proteins of the inhibitor of apoptosis (IAP) family which, under basal conditions, block apoptosis by directly inhibiting caspase activation (Verhagen *et al.*, 2000). Interestingly, genetic deletion of *Drosophila* inhibitor of apoptosis 1 (DIAP1) in *Drosophila* is sufficient to trigger apoptosis via spontaneous activation of the effector caspase Drice (Wang *et al.*, 1999). Therefore, IAP proteins serve a much more important regulatory role in fruit flies compared to mammals in which intrinsic apoptosis is predominately controlled by Bcl-2 proteins.

Caspase-2 has been shown to be involved in intrinsic apoptosis induced by various types of stimuli, although the role of this caspase in most apoptotic pathways is considered to be redundant (Puccini *et al.*, 2013a). Like other initiator caspases, caspase-2 cleavage can be facilitated by a multi-protein activation complex. This caspase-2 activation platform, termed the PIDDosome, consists of procaspase-2 complexed with p53-induced death domain protein (PIDD) and the adaptor molecule RIP-associated protein with a death domain (RAIDD) (Tinel and Tschopp, 2004). RAIDD recruits procaspase-2 via CARD–CARD interactions and subsequently associates with PIDD in a death domain (DD)-dependent manner leading to formation of the PIDDosome (Tinel and Tschopp, 2004). Although biochemical observations demonstrate that the PIDDosome can catalyze caspase-2 autoactivation, *in vitro* studies and *in vivo* evidence from knockout mice indicates caspase-2 activation can occur independently of RAIDD and PIDD, challenging the physiological relevance of the PIDDosome in caspase-2 activation (Baliga *et al.*, 2004; Manzl *et al.*, 2009; Read *et al.*, 2002). Regardless of its mechanism of activation, caspase-2 has been implicated in many processes including tumor suppression, ageing, and maintenance of genomic stability (Ho *et al.*, 2009; Puccini *et al.*, 2013b; Shalini *et al.*, 2012). However, it remains to be established whether some or any these functionalities of caspase-2 are dependent on its function in PCD.

The Death-Induced Signaling Complex (Extrinsic Apoptosis)

The extrinsic pathway of apoptosis is a critical component of the immune system that is required for the removal of damaged or infected cells (Walczak, 2013). The extrinsic pathway is engaged upon binding of extracellular ligands to transmembrane death receptors (Walczak, 2013). Several types of death receptors (e.g., Fas and TNFR) and ligands (e.g., FasL and TNF α) have been identified which belong to the tumor

necrosis factor super family (TNFSF) (Walczak, 2013). Each death receptor has cognate death ligands and specific adaptor molecules which facilitate the assembly of specific intracellular signaling complexes (Walczak, 2013). Therefore, different combinations of death ligands and adaptors give rise to specificity in death receptor signaling. In addition to the ligand binding domain in their extracellular region, death receptors contain DDs in their intracellular tails that allow interactions with DD-containing adaptor molecules (Walczak, 2013). Ligand binding stimulates the oligomerization of death receptors at the plasma membrane (e.g., Fas–FasL interaction) leading to the formation of homotrimeric death receptor complexes (Walczak, 2013). Adaptor molecules such as Fas-associated protein with death domain (FADD) are then recruited to the intracellular DDs of specific death receptor complexes (Chinnaiyan *et al.*, 1995). These adaptor molecules also contain DEDs which recruit DED-containing initiator caspases (such as procaspase-8) stimulating the formation of the death-induced signaling complex (DISC) and their activation by dimerization-dependent autoprocessing (Figure 3; Medema *et al.*, 1997). A recently proposed model suggests that DISC-mediated caspase-8 activation can also occur by a novel mechanism in which procaspase-8 monomers assemble into chain structures via DED–DED interactions (Dickens *et al.*, 2012). In addition to downstream effector caspases, activated caspase-8 can also cleave BH3 interacting-domain death agonist (BID) (Li *et al.*, 1998). This results in a truncated form of BID (tBID) that is capable of interacting with BAK/BAX, thereby stimulating MOMP and triggering procaspase-9 activation (Li *et al.*, 1998). Through this mechanism, death receptor-mediated activation of BID provides a connection between the intrinsic and extrinsic pathways.

As with other caspase activation platforms, the activity of the DISC is controlled by a host of regulatory proteins. One such well-characterized regulator of DISC assembly and extrinsic apoptosis is flc-like inhibitory protein (FLIP) (Irmeler *et al.*, 1997). FLIP contains two DEDs which compete with procaspase-8 for FADD binding and recruitment into the DISC (Irmeler *et al.*, 1997). Therefore, FLIP can act as an anti-apoptotic decoy caspase that limits procaspase-8 activation.

New Modes of Caspase-Regulated Cell Death

Over the last two decades, research on caspases has largely focused on their role as regulators of inflammation and apoptosis. More recently however, alternative modes of caspase-regulated PCD have come to light.

Pyroptosis

In addition to regulation of the inflammatory response (see section The Inflammasome (Inflammatory Response)), caspase-1-activation can also lead to cell death by a process called pyroptosis. Pyroptosis was originally described in bacteria-infected macrophages and characterized by morphological features including membrane swelling and rupture leading to the release of pro-inflammatory cell contents (Zychlinsky *et al.*, 1992). The definition of pyroptosis has since been refined to

describe a mode of caspase-1-dependent inflammatory cell death whose function is not limited to bacterial infection or macrophages (Galluzzi *et al.*, 2012). Therefore, pyroptosis has been suggested to represent a mechanism of crosstalk between inflammation and cell death, serving as a defense strategy against damaged or pathogen-infected cells. In support of this concept, various bacteria and viruses have evolved mechanisms to suppress inflammation and evade inflammatory cell death by expressing proteins that inhibit caspase-1 signaling (Galluzzi *et al.*, 2012).

Pyroptotic and apoptotic cells share some morphological features such as nuclear condensation and DNA fragmentation (Galluzzi *et al.*, 2012). However, unlike apoptotic cell death that involves degradation of internal cellular components without compromising membrane integrity, cell demise by pyroptosis occurs by osmotic lysis and release of intracellular contents (Galluzzi *et al.*, 2012). Although the morphological features of pyroptotic cell death were described almost two decades ago, the molecular events associated with this mode of cell death are poorly defined. As with initiation of the inflammatory response, pyroptosis is associated with inflammasome-dependent caspase-1 activation and processing and secretion of pro-inflammatory cytokines (IL-1 β and IL-18) (Figure 3; Galluzzi *et al.*, 2012). Under certain conditions during inflammation, secretion of these pro-inflammatory cytokines coincides with caspase-1-mediated cleavage and activation of the effector caspase, caspase-7 but not caspase-3 (Miao *et al.*, 2011). However, given that cytokine secretion and caspase-7 activation have been shown to be dispensable for caspase-1-mediated cell death, the physiological relevance of these factors to inflammatory cell death is not well understood (Miao *et al.*, 2011). Furthermore, the downstream effectors of caspase-1 that are required for pyroptosis remain largely unknown.

Recently, a novel role for inflammatory caspases-4 and -5 in pyroptosis induced by bacterial lipopolysaccharide (LPS) was described (Shi *et al.*, 2014). In this study, caspase-4 and -5 were shown to be able to directly bind LPS with high affinity (Shi *et al.*, 2014). This interaction triggered CARD-dependent oligomerization and activation of these inflammatory caspases leading to pyroptosis (Shi *et al.*, 2014). These findings revealed a previously uncharacterized mechanism of caspase activation and demonstrate that caspases not only serve as the initiators of downstream signaling following infection but can also function as pattern recognition molecules in the innate immune response. Although pyroptosis is characterized by specific morphological features and signaling events that are distinct from apoptosis, whether pyroptosis actually represents a *bona fide* mode of non-apoptotic, caspase-dependent cell death *per se*, or is simply an alternative channel of apoptotic cell death in response to pro-inflammatory or infectious stimuli is still a matter of debate. Moreover, the mechanisms and stimuli that govern the decision between pro-survival (inflammation) and pro-death (pyroptosis) outcomes in response to inflammatory stimuli is still an active area of research.

Necroptosis

Until recently, necrosis was viewed as a route of passive cell death that occurred in an unordered, random manner in

response to injury. However, pioneering work from various groups has provided clear evidence that necrosis can proceed in a highly ordered manner through well-defined molecular pathways that are genetically regulated. This type of regulated necrotic cell death is now referred to as programmed necrosis or necroptosis.

Necroptosis was originally identified as a mode of cell death induced by TNF in the presence of a pan-caspase inhibitor (Vercammen *et al.*, 1998). As discussed in section The Death-Induced Signaling Complex (Extrinsic Apoptosis), it is well-established that death ligands such as TNF induce apoptosis through a caspase-dependent pathway. These observations demonstrated that TNF was able to induce necrotic-like cell death in the absence of caspase activity. It is now well-recognized that caspases are intrinsic suppressors of necroptosis. Therefore, in stark contrast to apoptosis, necroptosis occurs in a caspase-independent manner and results in cell demise that is more akin to inflammatory cell death characterized by membrane rupture, cell swelling and lysis (Vanden Berghe *et al.*, 2014). However, it is now clear that many components of apoptotic pathways also have important regulatory functions in necroptosis signaling, suggesting that crosstalk mechanisms occur between these two apparently distinct modes of cell death.

Like inflammation, necroptosis plays an important role in innate immunity by activating various pro-inflammatory pathways in response to infection but can also have deleterious, pathological effects under certain conditions. Many viruses have evolved mechanisms that promote immune evasion and destruction. One such mechanism is the expression of virally-encoded caspase inhibitors (e.g., vFLIP) that are potent suppressors of apoptosis during infection (Thome *et al.*, 1997). Therefore, necroptosis may have evolved as a backup mechanism of PCD in vertebrate immune systems that operates to ensure cell demise when caspases activity is inhibited by invading pathogens (Vanden Berghe *et al.*, 2014).

Many different types of stimuli have been shown to trigger necroptosis including ligation of various death ligands (e.g., Fas and TNF), tissue injury, cytotoxic drugs, and infection (Vanden Berghe *et al.*, 2014). Regardless of the type of stimuli a cell encounters, engagement of the necroptotic pathway converges on stabilization and activation of receptor-interacting protein kinase 1 (RIPK1) and RIPK3 (Vanden Berghe *et al.*, 2014). Depending on the nature of the upstream signal, RIPK1 and RIPK3 associate with different regulatory factors that ultimately lead to activation of downstream effectors (Vanden Berghe *et al.*, 2014). Necroptosis induced by TNF is the most extensively studied necroptotic pathway to date. As discussed in section The Death-Induced Signaling Complex (Extrinsic Apoptosis), in response to death receptor engagement in caspase-8-proficient cells, caspase-8 is recruited to adaptor proteins such as FADD and activated by the DISC. Under these conditions, active caspase-8 directly cleaves and inactivates RIPK1 and RIPK3, thereby inhibiting necroptosis (Lin *et al.*, 1999; Oberst *et al.*, 2011). However, under conditions of caspase-8 inhibition (such as expression of virus-derived caspase inhibitors) or in caspase-8-deficient cells, RIPK1 and RIPK3 are stabilized. RIPK1 and RIPK3 activate each other and themselves by trans- and auto-phosphorylation and form a multiprotein complex called the necrosome (Cho *et al.*, 2009; Li *et al.*, 2012). Once activated in

the necrosome, RIPK1 and RIPK3 phosphorylate and activate downstream effectors, such as MLKL, that execute necroptosis (Sun *et al.*, 2012).

In summary, since the discovery of the first caspase CED-3 as a mediator of PCD, caspases have established themselves as critical enzymes that not only initiate and execute apoptosis, but play many roles in various biological processes such as immunity, inflammation, differentiation, and development. While much of the mechanistic insight into caspase enzymology and substrate cleavage is now established, the mechanisms of initiator caspase activation *in vivo* remain to be fully elucidated. Another gap in our knowledge is how many and which substrates need to be cleaved by caspases in order to ensue apoptosis. The emerging non-apoptotic and pleiotropic functions of caspases, not extensively covered in this article, suggest that targeting caspases in the treatment of diseases is far more complex than originally envisaged. On the other hand, the extensive knowledge of caspase biology and function may provide opportunities to target caspases in specific pathologies.

Acknowledgments

The caspase and cell death studies in our laboratory are supported by the National Health and Medical Council of Australia project grants (1021456, 1041807, and 1043057) and fellowship (1002863). We apologize to colleagues whose primary publications could not be cited due to the editorial policy on the number of references that can be included in this article.

See also: Cell Division/Death: Apoptosis: Apoptosis; The Bcl-2 Family Proteins: Insights into Their Mechanism of Action and Therapeutic Potential

References

- Baliga, B.C., Read, S.H., Kumar, S., 2004. The biochemical mechanism of caspase-2 activation. *Cell Death and Differentiation* 11, 1234–1241.
- Benkova, B., Lozanov, V., Ivanov, I.P., Mitev, V., 2009. Evaluation of recombinant caspase specificity by competitive substrates. *Analytical Biochemistry* 394, 68–74.
- Bergeron, L., Perez, G.I., Macdonald, G., *et al.*, 1998. Defects in regulation of apoptosis in caspase-2-deficient mice. *Genes and Development* 12, 1304–1314.
- Cerretti, D.P., Kozlosky, C.J., Mosley, B., *et al.*, 1992. Molecular cloning of the interleukin-1 beta converting enzyme. *Science* 256, 97–100.
- Chai, J., Shi, Y., 2014. Apoptosome and inflammasome: Conserved machineries for caspase activation. *National Science Review* 1, 101–118.
- Chen, L., Willis, S.N., Wei, A., *et al.*, 2005. Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function. *Molecular Cell* 17, 393–403.
- Chinnaiyan, A.M., O'Rourke, K., Tewari, M., Dixit, V.M., 1995. FADD, a novel death domain-containing protein, interacts with the death domain of Fas and initiates apoptosis. *Cell* 81, 505–512.
- Cho, Y.S., Challa, S., Moquin, D., *et al.*, 2009. Phosphorylation-driven assembly of the RIP1-RIP3 complex regulates programmed necrosis and virus-induced inflammation. *Cell* 137, 1112–1123.
- Chun, H.J., Zheng, L., Ahmad, M., *et al.*, 2002. Pleiotropic defects in lymphocyte activation caused by caspase-8 mutations lead to human immunodeficiency. *Nature* 419, 395–399.
- Denecker, G., Hoste, E., Gilbert, B., *et al.*, 2007. Caspase-14 protects against epidermal UVB photodamage and water loss. *Nature Cell Biology* 9, 666–674.
- Dickens, L.S., Boyd, R.S., Jukes-Jones, R., *et al.*, 2012. A death effector domain chain DISC model reveals a crucial role for caspase-8 chain assembly in mediating apoptotic cell death. *Molecular Cell* 47, 291–305.
- Eckhart, L., Ballaun, C., Hermann, M., *et al.*, 2008. Identification of novel mammalian caspases reveals an important role of gene loss in shaping the human caspase repertoire. *Molecular Biology and Evolution* 25, 831–841.
- Enari, M., Sakahira, H., Yokoyama, H., *et al.*, 1998. A caspase-activated DNase that degrades DNA during apoptosis, and its inhibitor ICAD. *Nature* 391, 43–50.
- Fernando, P., Kelly, J.F., Balazsi, K., Slack, R.S., Megeney, L.A., 2002. Caspase 3 activity is required for skeletal muscle differentiation. *Proceedings of the National Academy of Sciences of the United States of America* 99, 11025–11030.
- Fletcher, J.I., Meusburger, S., Hawkins, C.J., *et al.*, 2008. Apoptosis is triggered when prosurvival Bcl-2 proteins cannot restrain Bax. *Proceedings of the National Academy of Sciences of the United States of America* 105, 18081–18087.
- Franchi, L., Eigenbrod, T., Munoz-Planillo, R., Nunez, G., 2009. The inflammasome: A caspase-1-activation platform that regulates immune responses and disease pathogenesis. *Nature Immunology* 10, 241–247.
- Galluzzi, L., Vitale, I., Abrams, J.M., *et al.*, 2012. Molecular definitions of cell death subroutines: Recommendations of the Nomenclature Committee on Cell Death 2012. *Cell Death and Differentiation* 19, 107–120.
- Hakem, R., Hakem, A., Duncan, G.S., *et al.*, 1998. Differential requirement for caspase 9 in apoptotic pathways *in vivo*. *Cell* 94, 339–352.
- Ho, L.H., Taylor, R., Dorstyn, L., *et al.*, 2009. A tumor suppressor function for caspase-2. *Proceedings of the National Academy of Sciences of the United States of America* 106, 5336–5341.
- Hornung, V., Ablasser, A., Charrel-Dennis, M., *et al.*, 2009. AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* 458, 514–518.
- Huh, J.R., Guo, M., Hay, B.A., 2004. Compensatory proliferation induced by cell death in the Drosophila wing disc requires activity of the apical cell death caspase Dronc in a nonapoptotic role. *Current Biology* 14, 1262–1266.
- Humke, E.W., Shriver, S.K., Starovasnik, M.A., Fairbrother, W.J., Dixit, V.M., 2000. ICEBERG: A novel inhibitor of interleukin-1beta generation. *Cell* 103, 99–111.
- Irmrler, M., Thome, M., Hahne, M., *et al.*, 1997. Inhibition of death receptor signals by cellular FLIP. *Nature* 388, 190–195.
- Jiang, X., Wang, X., 2000. Cytochrome c promotes caspase-9 activation by inducing nucleotide binding to Apaf-1. *Journal of Biological Chemistry* 275, 31199–31203.
- Kluck, R.M., Bossy-Wetzel, E., Green, D.R., Newmeyer, D.D., 1997. The release of cytochrome C from mitochondria: A primary site for Bcl-2 regulation of apoptosis. *Science* 275, 1132–1136.
- Kuida, K., Haydar, T.F., Kuan, C.Y., *et al.*, 1998. Reduced apoptosis and cytochrome C-mediated caspase activation in mice lacking caspase 9. *Cell* 94, 325–337.
- Kuida, K., Lippke, J.A., Ku, G., *et al.*, 1995. Altered cytokine export and apoptosis in mice deficient in interleukin-1 beta converting enzyme. *Science* 267, 2000–2003.
- Kuida, K., Zheng, T.S., Na, S., *et al.*, 1996. Decreased apoptosis in the brain and premature lethality in CPP32-deficient mice. *Nature* 384, 368–372.
- Kumar, S., 2007. Caspase function in programmed cell death. *Cell Death and Differentiation* 14, 32–43.
- Kumar, S., Kinoshita, M., Noda, M., Copeland, N.G., Jenkins, N.A., 1994. Induction of apoptosis by the mouse Nedd2 gene, which encodes a protein similar to the product of the *Caenorhabditis elegans* cell death gene ced-3 and the mammalian IL-1 beta-converting enzyme. *Genes and Development* 8, 1613–1626.
- Kumar, S., Tomooka, Y., Noda, M., 1992. Identification of a set of genes with developmentally down-regulated expression in the mouse brain. *Biochemical and Biophysical Research Communications* 185, 1155–1161.
- Kuwana, T., Bouchier-Hayes, L., Chipuk, J.E., *et al.*, 2005. BH3 domains of BH3-only proteins differentially regulate Bax-mediated mitochondrial membrane permeabilization both directly and indirectly. *Molecular Cell* 17, 525–535.
- Lakhani, S.A., Masud, A., Kuida, K., *et al.*, 2006. Caspases 3 and 7: Key mediators of mitochondrial events of apoptosis. *Science* 311, 847–851.
- Lamkanfi, M., Kalai, M., Vandenabeele, P., 2004. Caspase-12: An overview. *Cell Death and Differentiation* 11, 365–368.
- Lee, S.H., Stehlik, C., Reed, J.C., 2001. Cop, a caspase recruitment domain-containing protein and inhibitor of caspase-1 activation processing. *Journal of Biological Chemistry* 276, 34495–34500.
- Li, H., Zhu, H., Xu, C.J., Yuan, J., 1998. Cleavage of BID by caspase 8 mediates the mitochondrial damage in the Fas pathway of apoptosis. *Cell* 94, 491–501.
- Li, J., McQuade, T., Siemer, A.B., *et al.*, 2012. The RIP1/RIP3 necrosome forms a functional amyloid signaling complex required for programmed necrosis. *Cell* 150, 339–350.

- Li, P., Allen, H., Banerjee, S., *et al.*, 1995. Mice deficient in IL-1 beta-converting enzyme are defective in production of mature IL-1 beta and resistant to endotoxin shock. *Cell* 80, 401–411.
- Lin, Y., Devin, A., Rodriguez, Y., Liu, Z.G., 1999. Cleavage of the death domain kinase RIP by caspase-8 prompts TNF-induced apoptosis. *Genes and Development* 13, 2514–2526.
- Manzl, C., Krumschnabel, G., Bock, F., *et al.*, 2009. Caspase-2 activation in the absence of PIDDosome formation. *Journal of Cell Biology* 185, 291–303.
- Marsden, V.S., Ekert, P.G., Van Delft, M., *et al.*, 2004. Bcl-2-regulated apoptosis and cytochrome c release can occur independently of both caspase-2 and caspase-9. *Journal of Cell Biology* 165, 775–780.
- Martinon, F., Burns, K., Tschopp, J., 2002. The inflammasome: A molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. *Molecular Cell* 10, 417–426.
- Martinon, F., Tschopp, J., 2005. NLRs join TLRs as innate sensors of pathogens. *Trends in Immunology* 26, 447–454.
- McIlwain, D.R., Berger, T., Mak, T.W., 2013. Caspase functions in cell death and disease. *Cold Spring Harbor Perspectives in Biology* 5, a008656.
- Medema, J.P., Scaffidi, C., Kischkel, F.C., *et al.*, 1997. FLICE is activated by association with the CD95 death-inducing signaling complex (DISC). *EMBO Journal* 16, 2794–2804.
- Miao, E.A., Rajan, J.V., Aderem, A., 2011. Caspase-1-induced pyroptotic cell death. *Immunological Reviews* 243, 206–214.
- Mikolajczyk, J., Scott, F.L., Krajewski, S., Sutherlin, D.P., Salvesen, G.S., 2004. Activation and substrate specificity of caspase-14. *Biochemistry* 43, 10560–10569.
- Miura, M., 2012. Apoptotic and nonapoptotic caspase functions in animal development. *Cold Spring Harbor Perspectives in Biology* 4(10): a008664.
- Miura, M., Zhu, H., Rotello, R., Hartwig, E.A., Yuan, J., 1993. Induction of apoptosis in fibroblasts by IL-1 beta-converting enzyme, a mammalian homolog of the *C. elegans* cell death gene *ced-3*. *Cell* 75, 653–660.
- Nakagawa, T., Zhu, H., Morishima, N., *et al.*, 2000. Caspase-12 mediates endoplasmic-reticulum-specific apoptosis and cytotoxicity by amyloid-beta. *Nature* 403, 98–103.
- Nikolaev, A., McLaughlin, T., O'Leary, D.D., Tessier-Lavigne, M., 2009. APP binds DR6 to trigger axon pruning and neuron death via distinct caspases. *Nature* 457, 981–989.
- Oberst, A., Dillon, C.P., Weinlich, R., *et al.*, 2011. Catalytic activity of the caspase-8-FLIP(L) complex inhibits RIPK3-dependent necrosis. *Nature* 471, 363–367.
- Pop, C., Salvesen, G.S., 2009. Human caspases: Activation, specificity, and regulation. *Journal of Biological Chemistry* 284, 21777–21781.
- Poreba, M., Strozzyk, A., Salvesen, G.S., Drag, M., 2013. Caspase substrates and inhibitors. *Cold Spring Harbor Perspectives in Biology* 5, a008680.
- Puccini, J., Dorstyn, L., Kumar, S., 2013a. Caspase-2 as a tumour suppressor. *Cell Death and Differentiation* 20, 1133–1139.
- Puccini, J., Shalini, S., Voss, A.K., *et al.*, 2013b. Loss of caspase-2 augments lymphomagenesis and enhances genomic instability in Atm-deficient mice. *Proceedings of the National Academy of Sciences of the United States of America* 110, 19920–19925.
- Read, S.H., Baliga, B.C., Ekert, P.G., Vaux, D.L., Kumar, S., 2002. A novel Apaf-1-independent putative caspase-2 activation complex. *Journal of Cell Biology* 159, 739–745.
- Salmena, L., Lemmers, B., Hakem, A., *et al.*, 2003. Essential role for caspase 8 in T-cell homeostasis and T-cell-mediated immunity. *Genes and Development* 17, 883–895.
- Segawa, K., Kurata, S., Yanagihashi, Y., *et al.*, 2014. Caspase-mediated cleavage of phospholipid flippase for apoptotic phosphatidylserine exposure. *Science* 344, 1164–1168.
- Shalini, S., Dorstyn, L., Wilson, C., *et al.*, 2012. Impaired antioxidant defence and accumulation of oxidative stress in caspase-2-deficient mice. *Cell Death and Differentiation* 19, 1370–1380.
- Shi, J., Zhao, Y., Wang, Y., *et al.*, 2014. Inflammatory caspases are innate immune receptors for intracellular LPS. *Nature* 514, 187–192.
- Sordet, O., Rebe, C., Plenchette, S., *et al.*, 2002. Specific involvement of caspases in the differentiation of monocytes into macrophages. *Blood* 100, 4446–4453.
- Sun, L., Wang, H., Wang, Z., *et al.*, 2012. Mixed lineage kinase domain-like protein mediates necrosis signaling downstream of RIP3 kinase. *Cell* 148, 213–227.
- Tang, Y., Wells, J.A., Arkin, M.R., 2011. Structural and enzymatic insights into caspase-2 protein substrate recognition and catalysis. *Journal of Biological Chemistry* 286, 34147–34154.
- Thomas, P.G., Dash, P., Aldridge Jr., J.R., *et al.*, 2009. The intracellular sensor NLRP3 mediates key innate and healing responses to influenza A virus via the regulation of caspase-1. *Immunity* 30, 566–575.
- Thornberry, N.A., Bull, H.G., Calaycay, J.R., *et al.*, 1992. A novel heterodimeric cysteine protease is required for interleukin-1 beta processing in monocytes. *Nature* 356, 768–774.
- Thome, M., Schneider, P., Hofmann, K., *et al.*, 1997. Viral FLICE-inhibitory proteins (FLIPs) prevent apoptosis induced by death receptors. *Nature* 386, 517–521.
- Tinel, A., Tschopp, J., 2004. The PIDDosome, a protein complex implicated in activation of caspase-2 in response to genotoxic stress. *Science* 304, 843–846.
- Vanden Berghe, T., Linkermann, A., Jouan-Lanhuet, S., Walczak, H., Vandenabeele, P., 2014. Regulated necrosis: The expanding network of non-apoptotic cell death pathways. *Nature Reviews Molecular Cell Biology* 15, 135–147.
- Varfolomeev, E.E., Schuchmann, M., Luria, V., *et al.*, 1998. Targeted disruption of the mouse Caspase 8 gene ablates cell death induction by the TNF receptors, Fas/Apo1, and DR3 and is lethal prenatally. *Immunity* 9, 267–276.
- Vercammen, D., Beyaert, R., Denecker, G., *et al.*, 1998. Inhibition of caspases increases the sensitivity of L929 cells to necrosis mediated by tumor necrosis factor. *Journal of Experimental Medicine* 187, 1477–1485.
- Verhagen, A.M., Ekert, P.G., Pakusch, M., *et al.*, 2000. Identification of DIABLO, a mammalian protein that promotes apoptosis by binding to and antagonizing IAP proteins. *Cell* 102, 43–53.
- Walczak, H., 2013. Death receptor-ligand systems in cancer, cell death, and inflammation. *Cold Spring Harbor Perspectives in Biology* 5, a008698.
- Walsh, J.G., Cullen, S.P., Sheridan, C., *et al.*, 2008. Executioner caspase-3 and caspase-7 are functionally distinct proteases. *Proceedings of the National Academy of Sciences of the United States of America* 105, 12815–12819.
- Wang, S.L., Hawkins, C.J., Yoo, S.J., Muller, H.A., Hay, B.A., 1999. The Drosophila caspase inhibitor DIAP1 is essential for cell survival and is negatively regulated by HID. *Cell* 98, 453–463.
- Wang, S., Miura, M., Jung, Y.K., *et al.*, 1998. Murine caspase-11, an ICE-interacting protease, is essential for the activation of ICE. *Cell* 92, 501–509.
- Wei, M.C., Zong, W.X., Cheng, E.H., *et al.*, 2001. Proapoptotic BAX and BAK: A requisite gateway to mitochondrial dysfunction and death. *Science* 292, 727–730.
- Williams, D.W., Kondo, S., Krzyzanowska, A., Hiromi, Y., Truman, J.W., 2006. Local caspase activity directs engulfment of dendrites during pruning. *Nature Neuroscience* 9, 1234–1236.
- Woo, M., Hakem, R., Soengas, M.S., *et al.*, 1998. Essential contribution of caspase 3/CPP32 to apoptosis and its associated nuclear changes. *Genes and Development* 12, 806–819.
- Yuan, J., Shaham, S., Ledoux, S., Ellis, H.M., Horvitz, H.R., 1993. The *C. elegans* cell death gene *ced-3* encodes a protein similar to mammalian interleukin-1 beta-converting enzyme. *Cell* 75, 641–652.
- Zou, H., Li, Y., Liu, X., Wang, X., 1999. An APAF-1-cytochrome c multimeric complex is a functional apoptosome that activates procaspase-9. *Journal of Biological Chemistry* 274, 11549–11556.
- Zychlinsky, A., Prevost, M.C., Sansonetti, P.J., 1992. Shigella flexneri induces apoptosis in infected macrophages. *Nature* 358, 167–169.

Relevant Websites

<http://bioinf.gen.tcd.ie/casbah/>
CASBAH.
<http://www.casbase.org/casvm/query/index.html>
CASVM.
<https://prosper.erc.monash.edu.au/help.html>
PROSPER.

Efferocytosis in the Tumor Microenvironment

DB Vaught, JC Stanford, and RS Cook, Vanderbilt University School of Medicine, Nashville, TN, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Efferocytosis The process by which a phagocyte engulfs a neighboring cell that is in early stages of cell death.

Postpartum breast cancer Breast cancers that are diagnosed in premenopausal women 2–5 years following a full-term pregnancy.

Postpartum involution Widespread cell death of the milk-producing epithelium of the breast, followed by structural remodeling of mammary microenvironment.

Introduction

Cells in the body are constantly being ‘turned over,’ to remove aged, damaged, or functionally unnecessary cells. Injury, pathogens, and physiological cues can induce widespread cell death in a tissue. For example, the milk-producing cells of the breast die when nursing ceases, while chemotherapeutic treatment of cancers would induce a profound burden of tumor cells. Regardless of the cause or extent of cell death, it is important that dying cells are rapidly removed, to prevent dying cells from releasing intracellular antigens that might trigger inflammation, tissue damage, or autoimmunity. Clearance of dying cells also stimulates production of cytokines and growth factors, that enhance cell growth and wound healing, and that dampen tissue-damaging immune responses. Given the profound burden of dying cells in the post-lactational breast, the breast is especially adept at clearance of dying cells during postpartum involution. Dying cells are removed from the breast by a process called ‘efferocytosis,’ wherein macrophages or phagocytic mammary epithelial cells engulf the dying cell. Efferocytosis is the body’s natural response to dying cells, aimed at suppressing immune responses and promoting tissue repair. Following efferocytosis, macrophages produce wound healing cytokines, adopt M2-like macrophage markers associated with wound healing, present antigens, promote immune tolerance, and dampen cytotoxic immunity. Macrophages use a cell surface protein called MerTK to recognize and engulf dying cells. The receptor tyrosine kinase (RTK) MerTK is critical for efferocytosis by wound-associated macrophages and for the cytokine modulation that follows.

A Historical Overview of Efferocytosis

Cell death occurs within most multicellular organisms. The idea that cell death within a developing or mature organism generally requires removal of those dying cells to maintain tissue homeostasis was noted more than 100 years ago when Metchnikoff coined the term ‘physiologic inflammation’ (Tauber, 2003). Since then, vast research into programmed cell death demonstrates the unique cellular packaging of dying cells into apoptotic bodies, such that the dying cell’s contents are encased within a plasma membrane that is flagged for clearance. This ensures that the intracellular contents are not

leaked into the immediate environment, which could potentially trigger immune reactions against these cellular contents. Earlier work in nematodes and mammals clearly demonstrated that apoptotic bodies were removed locally in tissues by ‘silent’ processes that did not induce inflammation, involving the uptake of apoptotic bodies by neighboring cells (Ellis *et al.*, 1991; Reddien and Horvitz, 2004; Walsh *et al.*, 1999). Since these early studies, it has become clear that all multicellular organisms, including humans, have conserved processes to engulf apoptotic bodies. Apoptotic cells are engulfed and digested by phagocytes, such as macrophages and dendritic cells (the ‘professional phagocytes’) (Fadok *et al.*, 1992; Rothlin *et al.*, 2007; Hu *et al.*, 2000). Increasing evidence suggests that many cell types in the body, including epithelial cells and fibroblasts, are capable of engulfing dying cells (Henson *et al.*, 2001; Henson and Hume, 2006; Monks *et al.*, 2008; Vandivier *et al.*, 2006; Prasad *et al.*, 2006), although nonprofessional phagocytes engulf dying neighbors with decreased efficiency as compared to macrophages and dendritic cells.

Since phagocytosis of dying cells was found to suppress inflammation, the term ‘physiological inflammation’ was revisited in an effort to reflect more accurately this process. Further, the term phagocytosis was too broad, given the many types of phagocytosis being recognized, with phagocytic clearance of dying cells using a distinct signaling pathways and producing distinct morphological features versus other modes of phagocytosis. Based on these observations, Henson and colleagues proposed the term ‘efferocytosis’ (taken from the Latin *effero*, meaning ‘to take to the grave’ or ‘to bury’) to indicate the specialized process by which phagocytic cells engulf dying cells and then modulate cytokines to dampen acute inflammation (Henson *et al.*, 2001; Henson and Hume, 2006). From these early beginnings, the importance of efferocytosis in homeostasis of multiple tissues has been explored. Additionally, new roles for efferocytosis have been discovered, including the impact of efferocytosis on intracellular microbes, antitumor immunity, and autoimmunity.

Efferocytosis Is an Established Process That Is Rapidly Deployed in the Breast

The breast endures vast changes during reproductive phases of a woman’s life (puberty, pregnancy, lactation, postpartum

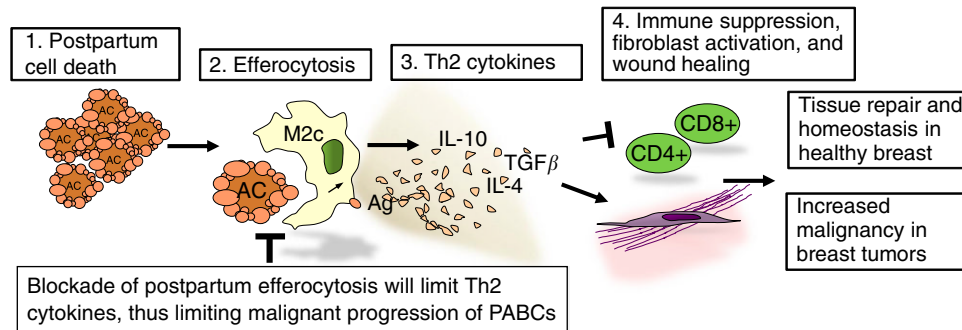


Figure 1 Schematic depicting efferocytosis during postpartum involution, and how this affects the breast microenvironment. Our hypothesis predicts that efferocytosis-induced cytokines will modulate leukocyte populations to increase tumor malignancy following episodes of widespread tumor cell death in response to postpartum involution. AC, apoptotic/dying cell; CD4, CD8, T-cells; M2, M2 macrophages.

involution, and postmenopausal involution). Although much of the patterning of the mammary ductal tree occurs during puberty, growth of milk-producing mammary epithelial cells occurs mainly during pregnancy, expanding mammary epithelial content by as much as 10-fold. Once offsprings are weaned, milk-producing mammary epithelial cells undergo rapid and widespread cell death (Figure 1), removing up to 90% of the total mammary epithelial content within the time span of just 1 week in rodent models. This returns the mammary gland to an almost pre-pregnant state, so that the changes in the mammary gland that accompany pregnancy, lactation, and involution may again occur with each successive pregnancy.

Involution of the post-lactational breast (or the postpartum breast, in the absence of nursing) is remarkable in terms of the robust burden of dying epithelial cells that require clearance. This makes the postpartum breast ideal for studying phagocytic engulfment of dying cells, or 'efferocytosis.' Efferocytosis dampens acute inflammatory responses upregulates tolerogenic cytokines and leukocytes, and promotes tissue repair and regrowth (Figure 1). In the absence of efferocytosis, dying cells undergo secondary necrotic lysis, often producing acute inflammation and immunity against self-antigens.

Efferocytosis is frequently performed by 'professional phagocytes' (macrophages and dendritic cells) but epithelial cells are also capable of efferocytosis under certain conditions. For example, mammary epithelial cells are capable of engulfing dying mammary epithelial cells during postpartum involution (Monks *et al.*, 2005, 2008; Sandahl *et al.*, 2010), although numerous bone marrow and spleen-derived macrophages are recruited to the postpartum breast in response to a growing burden of dying cells during early involution. Soluble chemo-attractants, or 'find me' signals such as monocyte chemo-attractant protein-1 (MCP-1/CCL2), CX3CL1, CCL6, CCL7, CCL8, and CXCL14, are released from dying cells during early involution (Nagaosa *et al.*, 2003), which may draw macrophages into the postpartum mammary gland (Graham *et al.*, 2011; Truman *et al.*, 2008; Clarkson *et al.*, 2004; Stein *et al.*, 2004).

Once recruited, the macrophages (or other phagocytes) identify dying cells in the local environment by engaging cell death signals that are specific to dying cells, often referred to as 'eat me' signals. External phosphatidylserine (PS) exposure is a

well-recognized marker of dying cells. Although PS is maintained on the inner plasma membrane leaflet in most healthy cells, externalized PS is how a dying cell signals to phagocytes for its engulfment (Fadok *et al.*, 1992; Borisenko *et al.*, 2003; Martin *et al.*, 1995). Externalized PS often is recognized directly by receptors, such as PS receptor (PSR), stabilin-2, TIM family receptors, and brain angiogenesis inhibitor 1 (BAI1), expressed on the surface of the phagocyte (Luciani and Chimenti, 1996; Mani *et al.*, 2010; Park *et al.*, 2007, 2008a,b; DeKruyff *et al.*, 2010; Ichimura *et al.*, 2008; Kobayashi *et al.*, 2007; Miyanishi *et al.*, 2007; Rodriguez-Manzanet *et al.*, 2010; Santiago *et al.*, 2007; Wong *et al.*, 2010). Other receptors bind to PS through extracellular bridging molecules, such as milk fat globule-epidermal growth factor-like 8 (MFG-E8), growth arrest-specific gene 6 (Gas6), Tubby, Tubby-like Protein-1 (TLP1), and Protein S, that simultaneously binds PS and the receptor (Hanayama *et al.*, 2002, 2004; Akakura *et al.*, 2004; Nakano *et al.*, 1997; Ishimoto *et al.*, 2000; Hall *et al.*, 2005). Receptors that bind to bridging molecules include MerTK, Axl, Tyro3, $\alpha\beta3$, and $\alpha\beta5$ integrins. Once the dying cell is engaged either directly or through a bridging molecule to receptors on the phagocyte, intracellular signaling events promote actin reorganization and phagocytic ingestion. The CrkII-DOCK180-ELMO complex activates Rac GTPase signaling to produce actin-dependent membrane extensions that physically engulf the target cell (Reddien and Horvitz, 2000; Gumieny *et al.*, 2001; Hasegawa *et al.*, 1996; Kiyokawa *et al.*, 1998; Van Aelst and D'Souza-Schorey, 1997), while RhoA counters the actin remodeling required for efferocytosis (Nakaya *et al.*, 2006; Riento and Ridley, 2003). Once engulfed, the dying cell is dispatched by the efferocyte, primarily through lysosomal degradation.

Although macrophages are professional phagocytes/efferocytes, mammary epithelial cells are capable of efferocytosis at the earliest stages of postpartum involution, ensuring a rapid response to the massive level of apoptosis that occurs during early involution, prior to peak accumulation of macrophages during this highly dynamic phase. Gas6-MerTK signaling is critical for mammary epithelial cell-mediated efferocytosis during post-lactational involution, as is MFG-E8, presumably through $\alpha\beta3$ and/or $\alpha\beta5$ integrins, although this has not yet been demonstrated (Sandahl *et al.*, 2010). However, the integrins $\alpha\beta3$ and $\alpha\beta5$ are expressed in the early

post-lactational mammary epithelium, when MFG-E8 is induced (Oshima *et al.*, 2002; Nakatani *et al.*, 2006; Hanayama and Nagata, 2005).

Following efferocytosis, transcriptional events in the efferocyte modulate cytokine expression in the local microenvironment (Figure 1), which orchestrate a wound healing, angiogenesis, and dampening of acute or cytotoxic inflammation (Silva, 2010). Wound healing expression signatures are robustly produced during postpartum involution (Monks *et al.*, 2005; O'Brien and Schedin, 2009; Stein *et al.*, 2004, 2009; Atabai *et al.*, 2007; Deonaraine *et al.*, 2007), consistent with transcriptional changes following efferocytosis. Recent studies demonstrate the critical nature of efferocytosis in the postpartum breast. Macrophages in MerTK-deficient mice are incapable of efferocytosis. Although mammary development occurs normally through puberty, pregnancy, and lactation in MerTK-deficient mice, impaired efferocytosis during postpartum involution allowed for secondary necrosis of mammary epithelial cells, impaired wound healing and cytokine modulation, caused inflammation and scarring, and interfered with lactation in future pregnancies (Sandahl *et al.*, 2010). Similar findings were seen upon loss of MFG-E8 signaling in the postpartum mammary gland (Hanayama and Nagata, 2005; Nakatani *et al.*, 2006). These findings underscore the need for rapid clearance of dying cells during postpartum involution.

Efferocytosis in the Tumor Microenvironment

Similar to what is seen in healthy breast, cell death is a feature of all breast tumors, but very little is known regarding efferocytosis-induced wound healing in breast cancers. Further parallels exist between sterile wounds and tumors beyond locally increased cell death, as both wounds and tumors are characterized by an abundance of M2 macrophages that express high levels of MerTK, a macrophage phenotype associated with efferocytosis (Zizzo *et al.*, 2012). Efferocytosis in the mammary tumor microenvironment heightens both MerTK expression and M2 polarization of tumor-associated macrophages (Figure 1), causing increased expression of immune suppressive Th2-like cytokines such as IL-4, IL-10, IL-13, and transforming growth factor (TGF)- β (De Oliveira Fulco *et al.*, 2014). In clinical breast cancer datasets, Th2 cytokines and M2 macrophages each correlate with decreased overall survival (Kristensen *et al.*, 2012). Consistent with this observation, tumor metastasis was increased following the induction of cell death and efferocytosis in the tumor microenvironment during postpartum/post-lactational involution, whereas inhibition of efferocytosis (through pharmacological or genetic MerTK inhibition), or inhibition of efferocytosis-induced cytokines TGF- β or IL-10 resulted in decreased distant tumor metastasis (Stanford *et al.*, 2014). These data demonstrate a mechanistic link in the tumor microenvironment between tumor cell death, efferocytosis, M2 macrophage activity, and metastasis. The precise mechanisms by which efferocytosis promotes tumor metastasis is unclear. Further investigations will fill this knowledge gap, clarifying how dynamic changes in the tumor microenvironment support tumor metastasis, and how we can combat tumor metastasis and recurrence through therapeutic strategies aimed at efferocytosis and efferocytosis-induced

cytokines. However, malignant tumor progression has long been compared to the wound healing process, due to the elevated rate of cell death, the presence of M2 macrophages, and expression of immunosuppressive cytokines. Interestingly, these same attributes characterize the post-lactational mammary gland.

Efferocytosis and Postpartum Breast Cancer

Although pregnancy at a young age decreases lifetime breast cancer risk, the first 5 years following pregnancy are associated with 'increased' breast cancer risk regardless of the woman's age, and with even greater risk with increasing age at the woman's first pregnancy. Increasingly, women are postponing child-birth, which may increase the incidence of postpartum breast cancer, defined as those breast cancers diagnosed 2–5 years after pregnancy. Currently, postpartum breast cancer accounts for nearly 25% of all breast cancers in young (premenopausal) women (Callihan *et al.*, 2013; Martinson *et al.*, 2013). In contrast to breast cancers diagnosed 'during' pregnancy, which correlate with a favorable prognosis, postpartum breast cancers are highly aggressive, more likely to be diagnosed at a metastatic stage, and display profoundly reduced 5-year survival (Lyons *et al.*, 2009; Martinson *et al.*, 2013, 2015). A series of studies in which breast cancer cells were injected into the mammary fat pads of age-matched female mice that were either never pregnant or that were at involution day 1 revealed the profound influence of postpartum involution on the capacity for the mammary microenvironment to promote malignant tumor progression (Lyons *et al.*, 2011, 2014; Maller *et al.*, 2013; Martinson *et al.*, 2015; McDaniel *et al.*, 2006; Schedin, 2006, 2007; Schedin and Borges, 2009). Detailed analysis of the postpartum breast and tumor microenvironment identified key roles played by M2 macrophages in establishing and maintaining the pro-tumor microenvironment, such as collagen grooming and cytokine regulation (Lyons *et al.*, 2011; O'Brien *et al.*, 2010, 2012; O'Brien and Schedin, 2009).

Using mouse models of postpartum breast cancer to compare mammary tumor progression in nulliparous (virgin) mice to those in age-matched but parous (single pregnancy), it was discovered that tumors from parous mice harvested during postpartum involution, but not during pregnancy, are 10-fold more metastatic, and expressed higher levels of wound healing cytokines and harbored more intra-tumoral M2 macrophage markers (Stanford *et al.*, 2014). These are the same pathological traits seen in human postpartum breast cancers (Borges and Schedin, 2012; Fornetti *et al.*, 2014). Studies have shown that widespread cell death, the hallmark of postpartum involution, triggers macrophage-mediated efferocytosis, M2 macrophage polarization, wound healing cytokine production, and TGF β -mediated tumor cell invasion in postpartum breast cancers (Stanford *et al.*, 2014). However, genetic or pharmacological MerTK inhibition impaired efferocytosis in postpartum breast cancers, blocking macrophage M2 polarization, expression of wound healing cytokines, and repressing formation of lung metastases. These findings support the idea that efferocytosis enhances metastasis of postpartum breast cancers.

Conclusion and Future Directions

Taken together, these collective findings suggest that clearance of dying cells is advantageous in sterile wounds, such as the postpartum mammary gland, for epithelial regrowth and repair, but this same process in the context of the breast tumor microenvironment may promote tumor growth and malignancy. These studies suggest that efferocytosis is a targetable process through use of MerTK-directed inhibitors, which potentially could improve the outcome for young women diagnosed with postpartum breast cancer. It remains unclear if efferocytosis-induced wound healing and metastatic acceleration similarly occurs in response to chemotherapy-induced tumor cell death. Therefore, future studies investigating efferocytosis-induced wound healing in the tumor microenvironment could have a substantial impact throughout the wider cancer research field, and may identify an opportunity to improve efficacy of currently used therapeutics that trigger widespread tumor cell death.

References

- Akakura, S., Singh, S., Spataro, M., *et al.*, 2004. The opsonin MFG-E8 is a ligand for the α v β 5 integrin and triggers DOCK180-dependent Rac1 activation for the phagocytosis of apoptotic cells. *Experimental Cell Research*, 292, 403–416.
- Atabai, K., Sheppard, D., Werb, Z., 2007. Roles of the innate immune system in mammary gland remodeling during involution. *Journal of Mammary Gland Biology and Neoplasia* 12, 37–45.
- Borges, V.F., Schedin, P.J., 2012. Pregnancy-associated breast cancer: An entity needing refinement of the definition. *Cancer* 118, 3226–3228.
- Borisenko, G.G., Matsura, T., Liu, S.X., *et al.*, 2003. Macrophage recognition of externalized phosphatidylserine and phagocytosis of apoptotic Jurkat cells – Existence of a threshold. *Archives of Biochemistry and Biophysics* 413, 41–52.
- Callihan, E.B., Gao, D., Jindal, S., *et al.*, 2013. Postpartum diagnosis demonstrates a high risk for metastasis and merits an expanded definition of pregnancy-associated breast cancer. *Breast Cancer Research and Treatment* 138, 549–559.
- Clarkson, R.W., Wayland, M.T., Lee, J., Freeman, T., Watson, C.J., 2004. Gene expression profiling of mammary gland development reveals putative roles for death receptors and immune mediators in post-lactational regression. *Breast Cancer Research* 6, R92–R109.
- De Oliveira Fulco, T., Andrade, P.R., De Mattos Barbosa, M.G., *et al.*, 2014. Effect of apoptotic cell recognition on macrophage polarization and mycobacterial persistence. *Infection and Immunity* 82, 3968–3978.
- DeKruyff, R.H., Bu, X., Ballesteros, A., *et al.*, 2010. T cell/transmembrane, Ig, and mucin-3 allelic variants differentially recognize phosphatidylserine and mediate phagocytosis of apoptotic cells. *Journal of Immunology* 184, 1918–1930.
- Deonaraine, K., Panelli, M.C., Stashower, M.E., *et al.*, 2007. Gene expression profiling of cutaneous wound healing. *Journal of Translational Medicine* 5, 11.
- Ellis, R.E., Jacobson, D.M., Horvitz, H.R., 1991. Genes required for the engulfment of cell corpses during programmed cell death in *Caenorhabditis elegans*. *Genetics* 129, 79–94.
- Fadok, V.A., Voelker, D.R., Campbell, P.A., *et al.*, 1992. Exposure of phosphatidylserine on the surface of apoptotic lymphocytes triggers specific recognition and removal by macrophages. *Journal of Immunology* 148, 2207–2216.
- Fornetti, J., Martinson, H.A., Betts, C.B., *et al.*, 2014. Mammary gland involution as an immunotherapeutic target for postpartum breast cancer. *Journal of Mammary Gland Biology and Neoplasia* 19, 213–228.
- Graham, K., Ge, X., Morenas, De Las, A., Tripathi, A., Rosenberg, C.L., 2011. Gene expression profiles of estrogen receptor-positive and estrogen receptor-negative breast cancers are detectable in histologically normal breast epithelium. *Clinical Cancer Research* 17, 236–246.
- Gumienny, T.L., Brugnera, E., Tosello-Trampont, A.C., *et al.*, 2001. CED-12/ELMO, a novel member of the Crkl/Dock180/Rac pathway, is required for phagocytosis and cell migration. *Cell* 107, 27–41.
- Hall, M.O., Obin, M.S., Heeb, M.J., Burgess, B.L., Abrams, T.A., 2005. Both protein S and Gas6 stimulate outer segment phagocytosis by cultured rat retinal pigment epithelial cells. *Experimental Eye Research* 81, 581–591.
- Hanayama, R., Nagata, S., 2005. Impaired involution of mammary glands in the absence of milk fat globule EGF factor 8. *Proceedings of the National Academy of Sciences of the United States of America* 102, 16886–16891.
- Hanayama, R., Tanaka, M., Miwa, K., *et al.*, 2002. Identification of a factor that links apoptotic cells to phagocytes. *Nature* 417, 182–187.
- Hanayama, R., Tanaka, M., Miyasaka, K., *et al.*, 2004. Autoimmune disease and impaired uptake of apoptotic cells in MFG-E8-deficient mice. *Science* 304, 1147–1150.
- Hasegawa, H., Kiyokawa, E., Tanaka, S., *et al.*, 1996. DOCK180, a major CRK-binding protein, alters cell morphology upon translocation to the cell membrane. *Molecular and Cellular Biology* 16, 1770–1776.
- Henson, P.M., Bratton, D.L., Fadok, V.A., 2001. Apoptotic cell removal. *Current Biology* 11, R795–R805.
- Henson, P.M., Hume, D.A., 2006. Apoptotic cell removal in development and tissue homeostasis. *Trends in Immunology* 27, 244–250.
- Hu, B., Sonstein, J., Christensen, P.J., Puntieri, A., Curtis, J.L., 2000. Deficient *in vitro* and *in vivo* phagocytosis of apoptotic T cells by resident murine alveolar macrophages. *Journal of Immunology* 165, 2124–2133.
- Ichimura, T., Asselton, E.J., Humphreys, B.D., *et al.*, 2008. Kidney injury molecule-1 is a phosphatidylserine receptor that confers a phagocytic phenotype on epithelial cells. *Journal of Clinical Investigation* 118, 1657–1668.
- Ishimoto, Y., Ohashi, K., Mizuno, K., Nakano, T., 2000. Promotion of the uptake of PS liposomes and apoptotic cells by a product of growth arrest-specific gene, gas6. *Journal of Biochemistry* 127, 411–417.
- Kiyokawa, E., Hashimoto, Y., Kobayashi, S., *et al.*, 1998. Activation of Rac1 by a Crk SH3-binding protein, DOCK180. *Genes & Development* 12, 3331–3336.
- Kobayashi, N., Karisola, P., Pena-Cruz, V., *et al.*, 2007. TIM-1 and TIM-4 glycoproteins bind phosphatidylserine and mediate uptake of apoptotic cells. *Immunity* 27, 927–940.
- Kristensen, V.N., Vaske, C.J., Ursini-Siegel, J., *et al.*, 2012. Integrated molecular profiles of invasive breast tumors and ductal carcinoma in situ (DCIS) reveal differential vascular and interleukin signaling. *Proceedings of the National Academy of Sciences of the United States of America* 109, 2802–2807.
- Luciani, M.F., Chimini, G., 1996. The ATP binding cassette transporter ABC1, is required for the engulfment of corpses generated by apoptotic cell death. *EMBO Journal* 15, 226–235.
- Lyons, T.R., Borges, V.F., Betts, C.B., *et al.*, 2014. Cyclooxygenase-2-dependent lymphangiogenesis promotes nodal metastasis of postpartum breast cancer. *Journal of Clinical Investigation* 124, 3901–3912.
- Lyons, T.R., O'Brien, J., Borges, V.F., *et al.*, 2011. Postpartum mammary gland involution drives progression of ductal carcinoma in situ through collagen and COX-2. *Nature Medicine* 17, 1109–1115.
- Lyons, T.R., Schedin, P.J., Borges, V.F., 2009. Pregnancy and breast cancer: When they collide. *Journal of Mammary Gland Biology and Neoplasia* 14, 87–98.
- Maller, O., Hansen, K.C., Lyons, T.R., *et al.*, 2013. Collagen architecture in pregnancy-induced protection from breast cancer. *Journal of Cell Science* 126, 4108–4110.
- Mani, O., Korner, M., Sorensen, M.T., *et al.*, 2010. Expression, localization, and functional model of cholesterol transporters in lactating and nonlactating mammary tissues of murine, bovine, and human origin. *American Journal of Physiology. Regulatory, Integrative and Comparative Physiology* 299, R642–R654.
- Martin, S.J., Reutelingsperger, C.P., McGahon, A.J., *et al.*, 1995. Early redistribution of plasma membrane phosphatidylserine is a general feature of apoptosis regardless of the initiating stimulus: Inhibition by overexpression of Bcl-2 and Abl. *Journal of Experimental Medicine* 182, 1545–1556.
- Martinson, H.A., Jindal, S., Durand-Rougely, C., Borges, V.F., Schedin, P., 2015. Wound healing-like immune program facilitates postpartum mammary gland involution and tumor progression. *International Journal of Cancer* 136 (8), 1803–1813.
- Martinson, H.A., Lyons, T.R., Giles, E.D., Borges, V.F., Schedin, P., 2013. Developmental windows of breast cancer risk provide opportunities for targeted chemoprevention. *Experimental Cell Research*, 319, 1671–1678.
- McDaniel, S.M., Rumer, K.K., Biroc, S.L., *et al.*, 2006. Remodeling of the mammary microenvironment after lactation promotes breast tumor cell metastasis. *American Journal of Pathology* 168, 608–620.
- Miyashita, M., Tada, K., Koike, M., *et al.*, 2007. Identification of Tim4 as a phosphatidylserine receptor. *Nature* 450, 435–439.

- Monks, J., Rosner, D., Geske, F.J., *et al.*, 2005. Epithelial cells as phagocytes: Apoptotic epithelial cells are engulfed by mammary alveolar epithelial cells and repress inflammatory mediator release. *Cell Death & Differentiation* 12, 107–114.
- Monks, J., Smith-Steinhart, C., Kruk, E.R., Fadok, V.A., Henson, P.M., 2008. Epithelial cells remove apoptotic epithelial cells during post-lactation involution of the mouse mammary gland. *Biology of Reproduction* 78, 586–594.
- Nagaosa, K., Shiratsuchi, A., Nakanishi, Y., 2003. Concomitant induction of apoptosis and expression of monocyte chemoattractant protein-1 in cultured rat luteal cells by nuclear factor-kappaB and oxidative stress. *Development, Growth & Differentiation* 45, 351–359.
- Nakano, T., Ishimoto, Y., Kishino, J., *et al.*, 1997. Cell adhesion to phosphatidylserine mediated by a product of growth arrest-specific gene 6. *Journal of Biological Chemistry* 272, 29411–29414.
- Nakatani, H., Aoki, N., Nakagawa, Y., *et al.*, 2006. Weaning-induced expression of a milk-fat globule protein, MFG-E8, in mouse mammary glands, as demonstrated by the analyses of its mRNA, protein and phosphatidylserine-binding activity. *Biochemical Journal* 395, 21–30.
- Nakaya, M., Tanaka, M., Okabe, Y., Hanayama, R., Nagata, S., 2006. Opposite effects of rho family GTPases on engulfment of apoptotic cells by macrophages. *Journal of Biological Chemistry* 281, 8836–8842.
- O'Brien, J., Lyons, T., Monks, J., *et al.*, 2010. Alternatively activated macrophages and collagen remodeling characterize the postpartum involuting mammary gland across species. *American Journal of Pathology* 176, 1241–1255.
- O'Brien, J., Martinson, H., Durand-Rougely, C., Schedin, P., 2012. Macrophages are crucial for epithelial cell death and adipocyte repopulation during mammary gland involution. *Development* 139, 269–275.
- O'Brien, J., Schedin, P., 2009. Macrophages in breast cancer: Do involution macrophages account for the poor prognosis of pregnancy-associated breast cancer? *Journal of Mammary Gland Biology and Neoplasia* 14, 145–157.
- Oshima, K., Aoki, N., Kato, T., Kitajima, K., Matsuda, T., 2002. Secretion of a peripheral membrane protein, MFG-E8, as a complex with membrane vesicles. *European Journal of Biochemistry* 269, 1209–1218.
- Park, D., Tosello-Trampont, A.C., Elliott, M.R., *et al.*, 2007. BAI1 is an engulfment receptor for apoptotic cells upstream of the ELMO/Dock180/Rac module. *Nature* 450, 430–434.
- Park, S.Y., Jung, M.Y., Kim, H.J., *et al.*, 2008a. Rapid cell corpse clearance by stabilin-2, a membrane phosphatidylserine receptor. *Cell Death & Differentiation* 15, 192–201.
- Park, S.Y., Kim, S.Y., Jung, M.Y., Bae, D.J., Kim, I.S., 2008b. Epidermal growth factor-like domain repeat of stabilin-2 recognizes phosphatidylserine during cell corpse clearance. *Molecular and Cellular Biology* 28, 5288–5298.
- Prasad, D., Rothlin, C.V., Burrola, P., *et al.*, 2006. TAM receptor function in the retinal pigment epithelium. *Molecular and Cellular Neuroscience* 33, 96–108.
- Reddien, P.W., Horvitz, H.R., 2000. CED-2/Crkl and CED-10/Rac control phagocytosis and cell migration in *Caenorhabditis elegans*. *Nature Cell Biology* 2, 131–136.
- Reddien, P.W., Horvitz, H.R., 2004. The engulfment process of programmed cell death in *Caenorhabditis elegans*. *Annual Review of Cell and Developmental Biology* 20, 193–221.
- Riento, K., Ridley, A.J., 2003. Rocks: Multifunctional kinases in cell behaviour. *Nature Reviews Molecular Cell Biology* 4, 446–456.
- Rodriguez-Manzanet, R., Sanjuan, M.A., Wu, H.Y., *et al.*, 2010. T and B cell hyperactivity and autoimmunity associated with niche-specific defects in apoptotic body clearance in TIM-4-deficient mice. *Proceedings of the National Academy of Sciences of the United States of America* 107, 8706–8711.
- Rothlin, C.V., Ghosh, S., Zuniga, E.I., Oldstone, M.B., Lemke, G., 2007. TAM receptors are pleiotropic inhibitors of the innate immune response. *Cell* 131, 1124–1136.
- Sandahl, M., Hunter, D.M., Strunk, K.E., Earp, H.S., Cook, R.S., 2010. Epithelial cell-directed efferocytosis in the post-partum mammary gland is necessary for tissue homeostasis and future lactation. *BMC Developmental Biology* 10, 122.
- Santiago, C., Ballesteros, A., Tamí, C., *et al.*, 2007. Structures of T Cell immunoglobulin mucin receptors 1 and 2 reveal mechanisms for regulation of immune responses by the TIM receptor family. *Immunity* 26, 299–310.
- Schedin, P., 2006. Pregnancy-associated breast cancer and metastasis. *Nature Reviews Cancer* 6, 281–291.
- Schedin, P., Borges, V., 2009. Breaking down barriers: The importance of the stromal microenvironment in acquiring invasiveness in young women's breast cancer. *Breast Cancer Research* 11, 102.
- Schedin, P., O'Brien, J., Rudolph, M., Stein, T., Borges, V., 2007. Microenvironment of the involuting mammary gland mediates mammary cancer progression. *Journal of Mammary Gland Biology and Neoplasia* 12, 71–82.
- Silva, M.T., 2010. Secondary necrosis: The natural outcome of the complete apoptotic program. *FEBS Letters* 584, 4491–4499.
- Stanford, J.C., Young, C., Hicks, D., *et al.*, 2014. Efferocytosis produces a prometastatic landscape during postpartum mammary gland involution. *Journal of Clinical Investigation* 124, 4737–4752.
- Stein, T., Morris, J.S., Davies, C.R., *et al.*, 2004. Involution of the mouse mammary gland is associated with an immune cascade and an acute-phase response, involving LBP, CD14 and STAT3. *Breast Cancer Research* 6, R75–R91.
- Stein, T., Salomonis, N., Nuyten, D.S., Van De Vijver, M.J., Gusterson, B.A., 2009. A mouse mammary gland involution mRNA signature identifies biological pathways potentially associated with breast cancer metastasis. *Journal of Mammary Gland Biology and Neoplasia* 14, 99–116.
- Tauber, A.I., 2003. Metchnikoff and the phagocytosis theory. *Nature Reviews Molecular Cell Biology* 4, 897–901.
- Truman, L.A., Ford, C.A., Pasikowska, M., *et al.*, 2008. CX3CL1/fractalkine is released from apoptotic lymphocytes to stimulate macrophage chemotaxis. *Blood* 112, 5026–5036.
- Van Aelst, L., D'Souza-Schorey, C., 1997. Rho GTPases and signaling networks. *Genes & Development* 11, 2295–2322.
- Vandivier, R.W., Henson, P.M., Douglas, I.S., 2006. Burying the dead: The impact of failed apoptotic cell removal (efferocytosis) on chronic inflammatory lung disease. *Chest* 129, 1673–1682.
- Walsh, G.M., Sexton, D.W., Blaylock, M.G., Convery, C.M., 1999. Resting and cytokine-stimulated human small airway epithelial cells recognize and engulf apoptotic eosinophils. *Blood* 94, 2827–2835.
- Wong, K., Valdez, P.A., Tan, C., *et al.*, 2010. Phosphatidylserine receptor Tim-4 is essential for the maintenance of the homeostatic state of resident peritoneal macrophages. *Proceedings of the National Academy of Sciences of the United States of America* 107, 8712–8717.
- Zizzo, G., Hilliard, B.A., Monestier, M., Cohen, P.L., 2012. Efficient clearance of early apoptotic cells by human macrophages requires M2c polarization and MerTK induction. *Journal of Immunology* 189, 3508–3520.

Further Reading

- Ravichandran, K.S., Lorenz, U., 2007. Engulfment of apoptotic cells: Signals for a good meal. *Nature Reviews Immunology* 7, 964–974.
- Reddien, P.W., Horvitz, H.R., 2004. The engulfment process of programmed cell death in *Caenorhabditis elegans*. *Annual Review of Cell and Developmental Biology* 20, 193–221.

The Bcl-2 Family Proteins: Insights into Their Mechanism of Action and Therapeutic Potential

DW Andrews and H Brahmabhatt, Sunnybrook Research Institute, Toronto, ON, Canada
B Leber, Hamilton, ON, Canada

© 2016 Elsevier Inc. All rights reserved.

Apoptosis and the Bcl-2 Family: A Historical Perspective

Apoptosis was first described in 1972 as a controlled form of programmed cell death distinct from necrosis that is required during development and in the maintenance of homeostasis of multi-cellular organisms (Kerr *et al.*, 1972). In 1985, the gene encoding B-cell lymphoma-2 (Bcl-2), the founding member of the family, was identified from a follicular lymphoma patient where a chromosomal translocation t(14;18) resulted in dysregulated Bcl-2 expression (Tsujimoto *et al.*, 1985). The link between Bcl-2, apoptosis, and cancer was established when the BCL-2 gene product was shown to block cell death rather than to enhance cell proliferation as was expected as the mode of action of oncogenes at that time (Vaux *et al.*, 1988; Hockenbery *et al.*, 1990). This seminal observation led to the concept that impaired apoptosis was critical for tumorigenesis (Hanahan and Weinberg, 2000).

The Pathway of Apoptosis

Cellular signals that initiate apoptosis originate from many different types of intracellular stresses such as DNA damage and the unfolded protein response (UPR), or from external signals like the engagement of extracellular death receptors by ligands such as binding of tumor necrosis factor alpha (TNF α) or Fas to their cognate receptors on cancer cells (Fulda and Debatin, 2006). While both these cell intrinsic and extrinsic pathways are regulated by the Bcl-2 family of proteins, the former is more tightly regulated by the Bcl-2 family proteins because these apoptotic signals ultimately converge to cause mitochondrial outer membrane permeabilization (MOMP). As a consequence of MOMP, factors such as cytochrome c are released from the mitochondrial intermembrane space into the cytoplasm, which leads to the activation of a cascade of proteases termed caspases that cleave a distinct population of subcellular substrates that yields the characteristic apoptotic phenotype (Garrido *et al.*, 2006; Chipuk *et al.*, 2006). Therefore by initiating this final effector phase of apoptosis MOMP is typically regarded as the point of no return for a cell. The crucial role of Bcl-2 family of proteins is to regulate MOMP by a coordinated series of protein–protein and protein–membrane interactions (Leber *et al.*, 2010, 2007).

The Bcl-2 Family of Proteins: Classification and Structure

More than 20 Bcl-2 family proteins have been identified and are classified into three groups based on function and sequence similarity within four conserved Bcl-2 homology (BH) regions (Figure 1(a); Sato *et al.*, 1994). The first group

comprises the anti-apoptotic proteins, Bcl-2, Bcl-XL, Bcl-w, Mcl-1, and AL, which contain all four BH regions. The pro-apoptotic family members are distributed in two groups: (1) Bax and Bak, the proteins that form the pore in the mitochondrial outer membrane to mediate MOMP share sequence similarity in the BH1-3 regions and a modified BH4 region, and (2) the BH3 proteins such as Bid, Bim, Puma, Bad, Noxa, Bmf, Bik, Beclin-1 and Hrk which function as stress sensors and only share homology in the BH3 region (Kvansakul *et al.*, 2008). A feature found in many Bcl-2 family proteins is a carboxyl-terminal (C-terminal) transmembrane region (TM) which anchors proteins to intracellular membranes such as the mitochondrial outer membrane (MOM), endoplasmic reticulum (ER) and nucleus (Shamas-Din *et al.*, 2013b).

Although Bcl-2 family proteins vary in function and in amino acid sequence, the multi-region Bcl-2 family proteins show a similar structural fold thus providing insights into binding interactions that explains their mechanism of action. The three-dimensional solution structures show a common fold where two central hydrophobic α -helices form a hairpin surrounded by 6–7 amphipathic α -helices. Additionally, a long unstructured flexible loop is present between helix 1 and helix 2. The loop acts as a regulatory region and is sensitive to posttranslational modifications. The BH1-3 regions form a hydrophobic groove of different sizes (Petros *et al.*, 2004), and these grooves serve as the docking sites for specific BH3 regions of other Bcl-2 family members in a fashion analogous to ligand-receptor interactions (Figure 1(b); Sattler *et al.*, 1997; Petros *et al.*, 2000; Liu *et al.*, 2003; Czabotar *et al.*, 2013; Dewson *et al.*, 2008). In Bax and Bcl-w, the hydrophobic C-terminal TM region occludes the hydrophobic groove from the aqueous environment of the cytoplasm (Suzuki *et al.*, 2000; Denisov *et al.*, 2003).

In contrast to the multi-region Bcl-2 proteins, NMR studies of Bim, Bad, Bmf and sequence analysis of BH3 proteins other than Bid and Bik suggest that they are intrinsically unstructured, but that the BH3 region becomes structured after interaction with the binding groove in anti-apoptotic members (Hinds *et al.*, 2007). Bid is an exception as it has a defined structure similar to the multi-region members, rather than its BH3 protein counterparts (Shamas-Din *et al.*, 2013a; Billen *et al.*, 2008b). From an evolutionary perspective, the multi-region Bcl-2 family members and Bid share a common ancestral origin whereas the BH3 proteins evolved later and separately (Aouacheria *et al.*, 2005). There is no structural information available for Bik but sequence analysis suggests that it may also be an all helical protein similar to the multi-region Bcl-2 family proteins. Unlike other family members, several BH3 proteins appear to have functions unrelated to apoptotic activity or to the functions of other BH3 proteins.

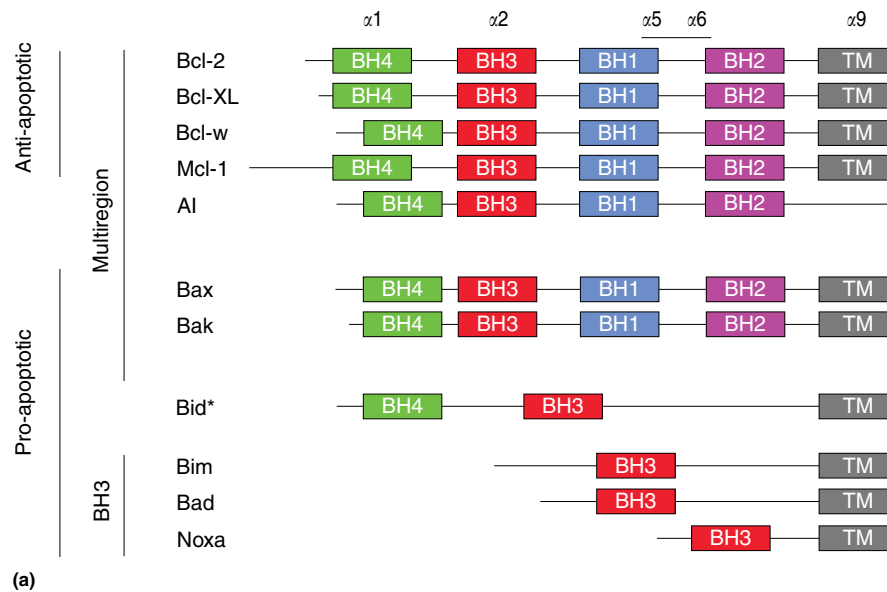


Figure 1 Schematic overview of the Bcl-2 family of proteins. (a) The Bcl-2 family of proteins is divided into three groups based on function and the presence of up to four Bcl-2 homology (BH) regions. The anti-apoptotic proteins and the multi-region pro-apoptotic proteins contain all four BH regions. The BH3 proteins contain only one conserved BH3 region. Most Bcl-2 family members contain a putative hydrophobic transmembrane (TM) region at the C-terminus that targets the proteins to intracellular membranes. *Bid is an example of an exceptional BH3 protein as it contains multiple BH regions. (b) The BH1-3 regions form a hydrophobic groove which serves as the interaction site for the BH3 region of other Bcl-2 family proteins. This figure is based on the crystal structure of Bcl-XL bound to the BH3 peptide of Bim (Liu, X., Dai, S., Zhu, Y., Marrack, P., Kappler, J.W., 2003. The structure of a Bcl-xL/Bim fragment complex: Implications for Bim function. *Immunity* 19, 341–352).

How Do Bcl-2 Family Proteins Regulate MOMP?

As the critical step in committing a cell to apoptosis, Bax and/or Bak oligomerize in the MOM forming a pore through which cytochrome c and other apoptogenic factors are released. Because both these proteins are present as monomers in cells, several models have been proposed to explain how their activation leads to oligomerization and how that is regulated by other Bcl-2 family members to ensure that MOMP only occurs in the appropriate cellular context (Figure 2(a)). According to the displacement model, Bax and Bak are constitutively active and therefore cell survival requires continuous inhibition of these proteins by binding to the anti-apoptotic proteins. During apoptosis, BH3 proteins displace activated Bax and Bak

from the anti-apoptotic proteins allowing them to oligomerize, enabling MOMP (Willis *et al.*, 2005, 2007; Chen *et al.*, 2005). In contrast, the direct activation model postulates that Bax and Bak are inactive and require activation to execute MOMP. As a consequence, BH3 proteins can be pro-apoptotic in two different ways, and can be classified as activators or sensitizers based on binding affinities for different sub-types of the multi-region Bcl-2 members (Figure 2(b)). Activator BH3 proteins bind to Bax and/or Bak directly to cause their activation, but also bind to anti-apoptotic proteins; this latter interaction does not necessarily lead to MOMP and is one mechanism by which anti-apoptotic proteins prevent apoptosis. By contrast BH3 sensitizers only bind to anti-apoptotic proteins. If the anti-apoptotic protein is already bound to an

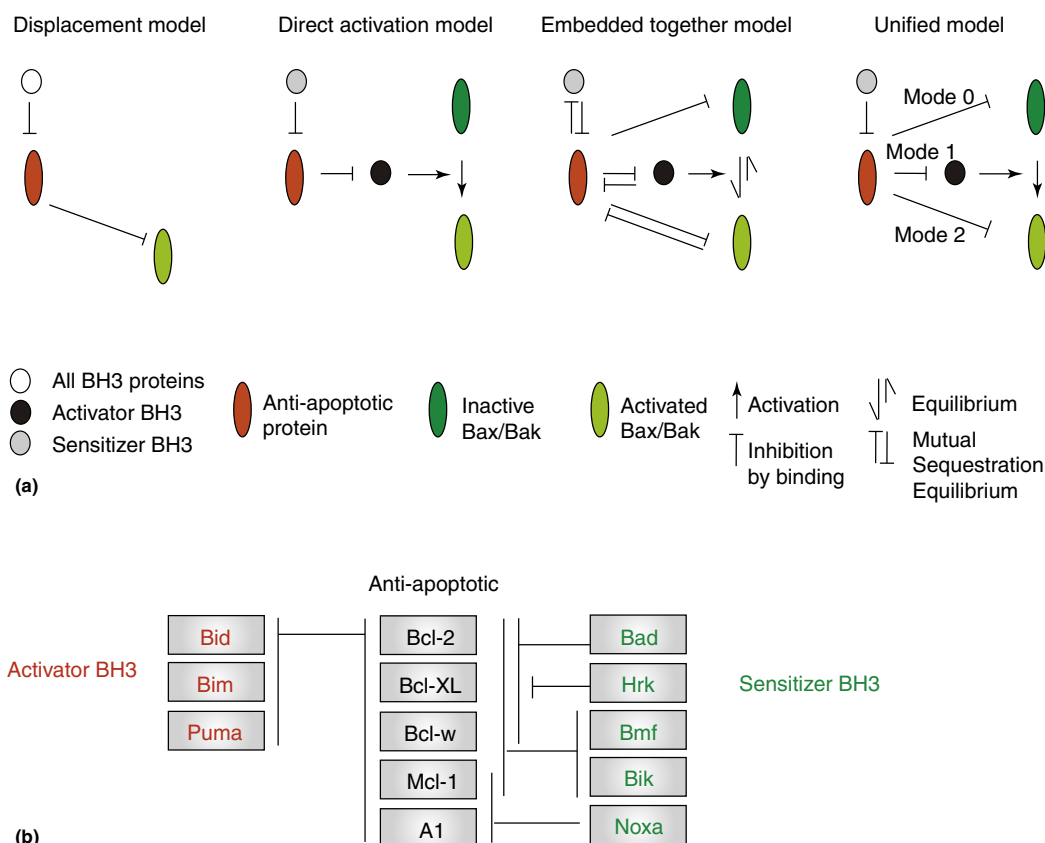


Figure 2 Models of the mechanism of action for the regulation of MOMP by Bcl-2 family proteins. (a) In the displacement model, Bax and/or Bak are constitutively active and inhibited by the anti-apoptotic proteins. BH3 proteins compete for binding to the anti-apoptotic proteins displacing Bax and/or Bak thereby allowing them to oligomerize and permeabilize membranes. In the direct activation model, BH3 proteins are classified as activators or sensitizers. Activator BH3 proteins bind to and activate Bax/Bak. Anti-apoptotic proteins inhibit the activator BH3 proteins preventing MOMP. BH3 sensitizers bind to the anti-apoptotic proteins to displace the BH3 activators whereupon they bind to and activate Bax and/or Bak. The embedded together model includes both displacement and activation mechanisms and suggests that the interactions of Bcl-2 family proteins with membranes results in conformational changes that change the affinities of the proteins for each other. Thus implicit in this model is that all interactions including those with membranes are reversible and governed equilibria. At membranes, the anti-apoptotic proteins inhibit the activator proteins as well as Bax and Bak by sequestering them or preventing their binding to the membrane. Both activator and sensitizer proteins can recruit and sequester anti-apoptotic proteins in the membrane, a term defined as mutual sequestration. The unified model defines anti-apoptotic sequestration of activator proteins and Bax or Bak as mode 1 and mode 2, respectively. Inhibition by mode 1 is less effective than mode 2. Mode 2 also prevents mitochondrial fusion and promotes mitochondrial fragmentation. (b) Classification of BH3 proteins as activators or sensitizers and selectivity of binding between BH3 proteins and anti-apoptotic members. This figure is based on binding affinities obtained from Certo, M., Del Gaizo Moore, V., Nishino, M., *et al.*, 2006. Mitochondria primed by death signals determine cellular addiction to antiapoptotic BCL-2 family members. *Cancer Cell* 9, 351–365.

activator, the activator is displaced whereupon it can bind to and activate Bax and/or Bak. Thus, sensitizers cause MOMP indirectly, and only if the anti-apoptotic protein bound has a 'pre-bound' activator (Letai *et al.*, 2002; Kuwana *et al.*, 2005). Thus, these models differ significantly as to the functional state of Bax and Bak monomers and the mechanism by which anti-apoptotic proteins inhibit MOMP (Leber *et al.*, 2007).

Based on experiments that measured these binding interactions *in vitro*, the embedded together model proposed that the distinction between activator and sensitizer BH3 proteins was valid, but incomplete. This model suggested that anti-apoptotic family members inhibit MOMP by binding to and sequestering both the activators and activated Bax and/or Bak (Leber *et al.*, 2007; Lovell *et al.*, 2008; Billen *et al.*, 2008a). Furthermore it assigns the membrane a critical role in the

process, not only as a locus of interaction that is ultimately altered during MOMP, but also as a regulator of binding interactions as after interaction with membranes Bcl-2 family members undergo conformational changes that affect subsequent binding interactions. For example, Bid binds to and activates Bax only when both proteins are membrane bound. As a result, the entire process is governed by a series of reversible equilibria of protein–protein and protein–membrane interactions and the functional outcome depends on the relative protein concentrations and binding affinities. An additional feature of this model is that it assigns a dual role to BH3 proteins in that activators and sensitizers can also 'activate' their anti-apoptotic binding partners by promoting their insertion into the membrane. As binding of a membrane-bound anti-apoptotic protein to a BH3 protein or to Bax

and/or Bak inactivates both proteins, this mechanism is termed mutual sequestration (Leber *et al.*, 2010).

The unified model also recognizes that there are multiple mechanisms whereby anti-apoptotic proteins function by designating the sequestration of activators as MODE 1 and sequestration of activated Bax and/or Bak as MODE 2. In this model, inhibition of MODE 1 is less efficient and more easily de-repressed than MODE 2. As an important extension related to, but independent of apoptosis, inhibition by MODE 2 also prevents mitochondrial fusion and promotes mitochondrial fragmentation (Llambi *et al.*, 2011). Recently, the mechanism by which in healthy cells Bcl-XL transiently binds to peripherally membrane-bound Bax causing its relocation to the cytoplasm has been ascribed MODE 0 (Billen *et al.*, 2008a; Todt *et al.*, 2013; Edlich *et al.*, 2011; Westphal *et al.*, 2014).

Conformational Changes Associated with Bax and Bak Activation

In most cells Bax and Bak are described as functionally equivalent. However, evidence from animals (described below) suggests this may not be entirely correct. After binding to activator BH3 proteins both Bax and Bak undergo conformation changes permitting oligomerization that leads to the pore formation that is the structural basis for MOMP (Degenhardt *et al.*, 2002; Hsu and Youle, 1997; Griffiths *et al.*, 1999). Both Bax and Bak are composed of nine α helices: $\alpha 5$ and $\alpha 6$ are the two central hairpin helices, $\alpha 9$ contains the C-terminal TM region, and $\alpha 2-4$ forms the hydrophobic pocket (Suzuki *et al.*, 2000; Moldoveanu *et al.*, 2006). Despite these similarities in structure, conformational changes and function, there are still important differences between them.

In healthy cells, Bax is primarily cytosolic whereas Bak is inserted in the mitochondria and ER membrane. This difference in localization is attributed differences in the structure and configuration of the C-terminal TM region in $\alpha 9$ in the two proteins. After binding to BH3 activators, $\alpha 9$ of Bax is dislodged from the hydrophobic pocket allowing Bax to insert into membranes (Wolter *et al.*, 1997; Schinzel *et al.*, 2004; Annis *et al.*, 2005). In contrast, Bak bypasses this step since it is constitutively integrated in the membrane via its $\alpha 9$; this difference in spontaneous membrane insertion may be due to the increased hydrophobicity of the TM region of Bak compared to Bax (Ferrer *et al.*, 2012).

A common conformational change that occurs after activation of both Bax and Bak is the rearrangement of an amino terminal sequence that generates a neo-epitope (Hsu and Youle, 1997; Griffiths *et al.*, 1999). A transient interaction of Bax with membranes triggers this conformation change that is detected by the 6A7 antibody. This change is largely irreversible after binding to BH3 proteins and membranes (Yethon *et al.*, 2003). Structurally this neo-epitope emerges due to the unwinding of part of the Bax $\alpha 1$ helix after activation (Peyerl *et al.*, 2007). The specific change leading to epitope exposure after Bak activation has not been elucidated, but is proposed to be similar.

Another difference (albeit controversial) between Bax and Bak is the location and number of BH3 protein binding sites. The common interaction involves the BH3 region and the hydrophobic groove of BH1-3 that leads to conformational

change favoring the formation of oligomers. While the Bid BH3 peptide binds Bak at this canonical site, an alternative binding hydrophobic pocket termed the 'rear pocket' was identified as the interaction site for the Bim BH3 peptide on Bax (Gavathiotis *et al.*, 2008; Leshchiner *et al.*, 2013). Notably this pocket is located at the 'opposite side' of the protein compared to the canonical site (termed 'front pocket'). The $\alpha 1$ and 6 helices form the rear pocket and is masked by the $\alpha 1$, 2 loop in the unbound state. Binding of the Bim BH3 peptide results in a conformation change that facilitates the displacement of C-terminal $\alpha 9$ that mediates membrane binding, and also exposes the BH3 region (Kim *et al.*, 2009; Figure 3(a)). Displacement of $\alpha 9$ is followed by the insertion of $\alpha 5$, $\alpha 6$ and $\alpha 9$ onto or into the membrane (Czabotar *et al.*, 2013; Annis *et al.*, 2005).

Regardless of whether there are one or two activation binding sites on Bax, activation results in the exposure of the hydrophobic BH3 region required for Bax homodimerization (Dewson *et al.*, 2008; George *et al.*, 2007). However, the type of dimer that is formed will depend on which binding pocket this newly exposed BH3 region finds. There are two competing models to explain the propagation of Bax dimers into large pore forming oligomers. In the asymmetric dimer model, the exposed BH3 region of one Bax molecule activates another Bax molecule through binding to the rear pocket (Figure 3(b); Gavathiotis *et al.*, 2008). In contrast, the symmetric dimer model predicts that a Bax dimer forms when the BH3 region of one Bax molecule binds to the BH3 groove of another Bax molecule. Oligomer propagation occurs as a consequence of the interaction of two dimers via the rear pocket (Figure 3(c); Czabotar *et al.*, 2013; Brouwer *et al.*, 2014).

Physiology and Regulation of Bcl-2 Family Members

Although the role of Bcl-2 family members in regulating apoptosis was initially recognized in the context of over-expression secondary to a chromosomal translocation found in a human cancer, our understanding of the physiological functions of specific Bcl-2 family members has been greatly aided by genetic knockout mice studies. Mice with a genetic deletion of *BCL-2*, *BCL-XL* or *MCL-1* are not viable. *MCL-1* $-/-$ mice display a failure in embryonic implantation (Rinkenberger *et al.*, 2000). *BCL-XL* $-/-$ mice die during embryonic development and exhibit extensive neuronal and hematopoietic progenitor cell death (Motoyama *et al.*, 1995). *BCL-2* $-/-$ mice complete embryonic development but die postnatally displaying growth retardation, renal failure, extensive lymphocyte cell death and defects in hair pigmentation (Veis *et al.*, 1993). In contrast, Bcl-w knockout males are sterile and A1 knockout mice have accelerated neutrophil apoptosis (Hamasaki *et al.*, 1998; Print *et al.*, 1998).

Pro-apoptotic proteins are also important during development. While *BAK* $-/-$ mice are developmentally normal and *BAX* $-/-$ mice have limited abnormalities restricted to hyperplasia of lymphoid tissues and male sterility, mice lacking both *BAX* and *BAK* die perinatally from the accumulation of excessive cells in the hematopoietic and central nervous system. As expected, cells from these mice are resistant to various apoptotic stimuli (Knudson *et al.*, 1995; Lindsten *et al.*, 2000).

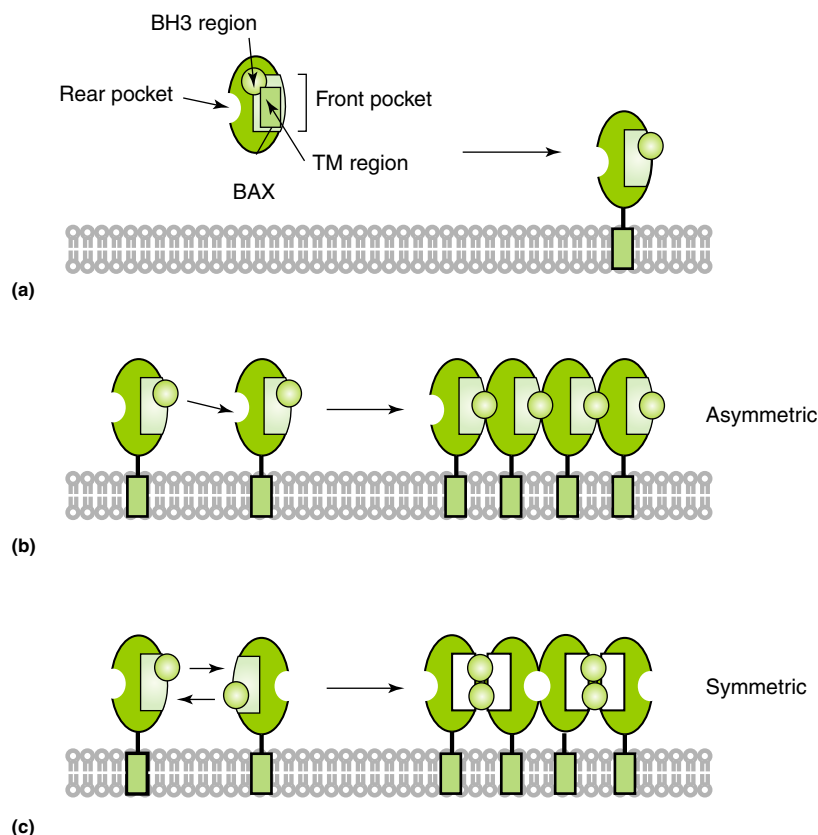


Figure 3 Models of Bax dimer formation. (a) Bax activation is triggered by interaction of a BH3 protein on the rear or front pocket. This interaction triggers the exposure of the BH3 region of Bax and insertion of the Bax into the membrane via the TM region. (b) Formation of asymmetric dimers: The exposed BH3 region of Bax and Bak interacts with the rear pocket of an adjacent Bax or Bak. Oligomer formation occurs by the subsequent interaction of the rear pocket and the BH3 regions. (c) Formation of symmetric dimers: The exposed BH3 regions of Bax and Bak interact with the hydrophobic groove of an adjacent Bax or Bak forming a homo-dimer. Pore formation occurs by subsequent oligomerization of homodimers presumably via interactions with the rear pocket.

In contrast, most mice lacking BH3 proteins are not affected developmentally, but display various tissue specific and signaling abnormalities (Table 1). As the BH3 domain is generally not required for the diverse non-apoptotic functions of BH3 proteins, it is likely that many of the defects observed in knockout animals are not due to alterations in apoptosis. Moreover, it is an open question for each family member as to which function evolved first. These proteins therefore serve as examples of intersecting nodes between cell death and metabolism, two processes that are tightly linked at many levels (Green *et al.*, 2014).

Since Bcl-2 family proteins are critical for maintaining the balance between life and death, their expression and activity is tightly controlled at multiple levels. At a transcriptional level, the expression of Bcl-2 family proteins is often increased in response to death or survival signals. For example, DNA damage results in the transcriptional activation by the tumor-suppressor P53 of pro-apoptotic NOXA, PUMA and BAX (Oda *et al.*, 2000; Nakano and Vousden, 2001; Miyashita *et al.*, 1994). By contrast, growth or survival signals cause the phosphorylation of STAT5 which acts as a transcriptional enhancer of BCL-XL (Silva *et al.*, 1999).

Function can also be regulated by alternative splicing: Bcl-XL and Bcl-XS are splice variants; the former anti-apoptotic

and the latter pro-apoptotic due to the absence of a critical domain (Grillot *et al.*, 1997; Boise *et al.*, 1993). Bim has three isoforms, of which BimS is constitutively active in cells because it lacks a dynein light chain binding domain that normally sequesters BimL and BimEL to the micro-tubule associated dynein motor complex (O'Connor *et al.*, 1998). The latter proteins are therefore held inactive but able to respond more rapidly than BimS that is regulated transcriptionally.

At the posttranslational level, many Bcl-2 proteins are regulated by posttranslational modifications including phosphorylation, myristoylation, ubiquitin-mediated degradation and proteolytic cleavage, most commonly in the unstructured loop region, BH3 region and in the TM (Kutuk and Letai, 2008). These posttranslational modifications directly affect protein localization, half-life (and therefore concentration), and binding affinity to other Bcl-2 family members. The modifications also affect binding to other regulatory proteins outside the Bcl-2 family. For example, in the presence of the IL-3 survival factor the pro-apoptotic Bad is phosphorylated, binds to 14-3-3 complexes and is thereby sequestered in the cytoplasm. Therefore the phosphorylation status of Bad can serve as a switch in function as de-phosphorylation of Bad results in dissociation from the 14-3-3 complex and binding to Bcl-XL at the membrane as a pro-apoptotic sensitizer

Table 1 Phenotypes of mice deficient in BH3 proteins

<i>BH3 protein deletion</i>	<i>Phenotypes in mice</i>	<i>References</i>
Bid	Mice are resistant to Fas mediated fatal hepatocyte apoptosis and hemorrhagic necrosis	Yin <i>et al.</i> (1999)
Bim	Mice have hyperplasia of lymphoid and myeloid cells and T cell development is perturbed. Lymphocytes are resistant to apoptosis induced by elevated Ca^{2+} by ionomycin, micro-tubule stabilization by taxol, and cytokine deprivation. With age, mice develop lymphadenopathy and a fatal systemic autoimmune disease	Bouillet <i>et al.</i> (1999)
Puma	Various cell types are resistant to apoptosis induced by DNA damage, cytokine deprivation and glucocorticoid exposure	Jeffers <i>et al.</i> (2003); Villunger <i>et al.</i> (2003)
Bad	Certain cell types display an increased resistance to epidermal or insulin-like growth factor mediated apoptosis. Lymphocytes have decreased IgG production. With age, mice develop B-cell lymphoma	Ranger <i>et al.</i> (2003)
Noxa	Mice exhibit increased resistance toward X-ray irradiation induced gastrointestinal death. Fibroblasts display mild resistance toward etoposide or γ irradiation	Villunger <i>et al.</i> (2003); Shibue <i>et al.</i> (2003)
Bmf	Mice have hyperplasia of B cells and mild hypergammaglobulinemia Thymocytes and pre-B cells are mildly resistant to apoptosis induced by glucocorticoids, and histone- deacetylase inhibition	Labi <i>et al.</i> (2008)
Bik	No obvious detectable defects	Coultas <i>et al.</i> (2005)
Beclin	Early embryonic lethality. Embryonic stem cells display abnormalities in autophagy induction	Yue <i>et al.</i> (2003)
Hrk	Some neuronal cell types displayed delayed apoptosis induced by nerve growth factor withdrawal	Coultas <i>et al.</i> (2007); Imaizumi <i>et al.</i> (2004)

(Zha *et al.*, 1996). Additionally in healthy cells, Bak is bound to mitochondrial outer membrane protein VDAC2, thereby preventing activation until it is displaced during death signal initiation by tBid, Bim or Bad (Cheng *et al.*, 2003).

Bcl-2 Proteins at the Endoplasmic Reticulum

While much research has focused on the regulation of MOMP by Bcl-2 family proteins, they have also been implicated in important processes at the ER such as modulation of calcium signaling (Oakes *et al.*, 2006; Rong and Distelhorst, 2008). During endoplasmic reticulum stress, Bax and Bak oligomerize at the endoplasmic reticulum causing the release of calcium into the cytosol (Scorrano *et al.*, 2003; Zong *et al.*, 2003; Nutt *et al.*, 2002). This can lead to both mitochondrial dependent and independent pathways of cells death although the latter is less well characterized (Heath-Engel *et al.*, 2008). As a potential mediator of this process, the BH3 protein Bik induces ER resident Bax and Bak to oligomerize and release calcium ultimately leading to MOMP and the release of cytochrome *c* (Mathai *et al.*, 2005). ER resident Bcl-2 affects calcium flux by directly or indirectly binding to the IP3 receptor calcium channel (Rong *et al.*, 2009). Additionally, Bcl-2 localized at the endoplasmic reticulum can inhibit apoptosis triggered by various stress signals (Zhu *et al.*, 1996; Bhatt *et al.*, 2008). Furthermore, Bcl-2 has been found to be present in a complex which includes the ER protein Bap31, another transmitter of apoptotic signals from the endoplasmic reticulum to mitochondria (Chami *et al.*, 2004; Rong *et al.*, 2009; Breckenridge *et al.*, 2003; Ng *et al.*, 1997).

The unfolded protein response (UPR) and autophagy are two additional complex cellular responses mediated at the ER that are also regulated by Bcl-2 family proteins. UPR occurs as a consequence of the accumulation of unfolded or misfolded

proteins in the lumen of the ER and triggers three signaling pathways mediated by ER localized ATF6, PERK and IRE-1, to restore normal function. Prolonged or excessive UPR results in apoptosis through PERK activation leading to CHOP-dependent transcriptional upregulation of *PUMA* and *BIM* (Puthalakath *et al.*, 2007; Ghosh *et al.*, 2012). A second arm of UPR mediation also requires the Bcl-2 family as Bax and Bak promote IRE-1 signaling (Hetz *et al.*, 2006).

Autophagy is a cell survival mechanism in which double membrane organelles called autophagosomes are induced to form around damaged cellular components such as misfolded proteins, protein aggregates and dysfunctional organelles. By regulated fusion with lysosomes the contents are catabolized and can be re-used to maintain cellular energy levels (Levine and Kroemer, 2008). Autophagy is therefore a survival mechanism, but autophagy can be driven to cell death through extensive degradation of critical organelles (Mukhopadhyay *et al.*, 2014). The initiation of autophagosome formation at the ER requires the BH3 protein Beclin-1 to co-ordinate a multi-protein complex (Axe *et al.*, 2008). Bcl-2 inhibits autophagy by directly binding to Beclin-1 in the presence of NAF-1 (Chang *et al.*, 2010; Rong *et al.*, 2008; Pattingre *et al.*, 2005). The BH3 protein Bad stimulates autophagy by displacing Beclin-1 from binding to Bcl-2 and Bcl-XL and allowing formation of the pre-autophagosome complex; thus there is direct cross talk between autophagy and apoptosis (Maiuri *et al.*, 2007).

Bcl-2 Family Proteins and Cancer

Cancers arise when cells acquire genetic or epigenetic abnormalities that cause uncontrolled proliferation. Apoptosis acts as a physiologic brake to this context-independent cell division (Meier *et al.*, 2000). Therefore the evasion of apoptosis is one of the required features for tumorigenesis and

many direct and indirect mediators of cell death are known to be abnormal in human and experimental cancer models, including Bcl-2 family proteins. In addition to the chromosomal translocations which led to the original discovery of *BCL-2*, many more common mechanisms have been observed including the amplification of the *BCL-2* gene as observed in small cell lung cancer and the loss of endogenous microRNAs that repress *BCL-2* gene expression as seen in chronic lymphocytic leukemia (Ikegaki *et al.*, 1994; Cimmino *et al.*, 2005). Amplification of *MCL-1* is one of the most common genetic abnormalities across cancers (Beroukhi *et al.*, 2010). Additionally, the inactivation of *BAX* by microsatellite insertions is common amongst many malignancies (Rampino *et al.*, 1997; Yamamoto *et al.*, 1997). Point mutations or posttranslational modifications altering the activity of BH3 proteins have also been reported, such as the somatic missense mutations reported in *BAD* in cases of colon adenocarcinoma. These mutations are predicted to reduce the binding affinity of Bad for Bcl-2 and Bcl-XL (Lee *et al.*, 2004).

Cancer and Addiction to Bcl-2 Family Proteins

Because of the requirement for preventing apoptosis to maintain continued proliferation of cancer cells, tumors are essentially 'addicted' to this inhibition. As critical regulators of MOMP, Bcl-2 family proteins are good (but not the only) candidates to achieve this (Certo *et al.*, 2006; Del Gaizo Moore *et al.*, 2007). Thus selective antagonism of anti-apoptotic Bcl-2 family proteins should be sufficient to kill those cancers where this represents the locus of addiction. The BH3 profiling assay is a powerful predictive tool to determine if regulation of MOMP by Bcl-2 family proteins is the locus, and to identify the nature of the addiction. To determine this, mitochondria are isolated from tumor cells and exposed to a panel of BH3 peptides representing the BH3 proteins. MOMP is measured by quantifying the release of cytochrome c from the mitochondria (Del Gaizo Moore and Letai, 2013). When MOMP is elicited, the pattern of release by effective and ineffective peptides can be divided into three classes with functional significance (Figure 4). The 'Class A' block occurs due to insufficient BH3 protein activation, and is overcome by adding activator BH3 peptides. 'Class B' block occurs due to a loss of the executors, Bax and Bak. Mitochondria with this block are insensitive to all BH3 peptides. In the 'Class C' block, MOMP is prevented by the overexpression of anti-apoptotic proteins which sequester the pro-activator molecules, and can be overcome by adding sensitizer BH3 peptides. Because of the specificity of BH3 sensitizers to different anti-apoptotic Bcl-2 family members, BH3 profiling can reveal which inhibitor(s) are present (Figure 2(b)). For example, Bad binds Bcl-2 and Bcl-XL (but not Mcl-1), whereas Noxa binds Mcl-1 and A1 (but not Bcl-2 or Bcl-XL). Thus mitochondria with a 'Class C' block are described as being 'primed' for death as an increase in the concentration of the relevant BH3 protein (or its mimetic as described in the next section) should elicit cell death in that tumor.

Bcl-2 Family Proteins as Therapeutic Targets

Targeting cell death with agents that affect Bcl-2 family proteins represents a new way to selectively target cancer cells that

contrasts with the broad toxicity of most standard anti-cancer drugs that are often selective only at concentrations that kill only dividing cells. As a result there is an emerging strategy to restore apoptosis in cancers is by antagonizing the activity of anti-apoptotic proteins using small molecules. Although several putative anti-apoptotic protein inhibitors have been identified, not all these agents have shown specificity for Bcl-2 family proteins as their mechanism of action (Vogler *et al.*, 2009).

The best characterized and furthest developed series of truly selective molecules have been designed by Abbot Laboratories and provide an instructive lesson in Bcl-2 family protein physiology. ABT 737 is a mimetic of the BH3 region of Bad and as such is a potent inhibitor of Bcl-2, Bcl-XL and Bcl-w; the closely related but better absorbed ABT-263 has been taken forward into clinical trials (Tse *et al.*, 2008). As a single agent, the results in hematological malignancies were encouraging. However, a dose limiting decrease in platelet counts (thrombocytopenia) was observed that represents an 'on-target' effect, as inhibition of Bcl-XL is known to predictably decrease platelet life-span (Vogler *et al.*, 2011). As this effect is limited to Bcl-XL among the anti-apoptotic proteins, a new compound, ABT-199 was designed which binds to Bcl-2 but not to Bcl-XL or Bcl-w (Souers *et al.*, 2013). While ABT-199 works extremely efficiently for some hematological malignancies, presumably there will be some tumors that are resistant because they are addicted to Bcl-XL rather than Bcl-2. Furthermore, solid tumors are also more resistant, and this phenomenon may be attributed to resistance conferred by Mcl-1 which is not affected by ABT-199 (or ABT-263), and thus represents another important target (van Delft *et al.*, 2006; Varin *et al.*, 2010). An important unanswered question for all approaches that inhibit anti-apoptotic Bcl-2 family proteins is whether specificity for cytotoxicity against cancer cells will be retained when these agents are combined with other cancer therapies.

While restoring apoptosis in addicted cancers by antagonizing the activity of the anti-apoptotic proteins using small molecules is promising and in clinical development, the pharmacologic modulation of pro-apoptotic Bax and Bak activity is another promising approach that is being explored. The small molecules BAM7, SMBA1 and compound 106, have been identified as potential candidates with differing putative mechanisms of action. While BAM7 activates Bax by binding the rear pocket similar to the Bim peptide, compound 106 is predicted to bind to the front pocket (Gavathiotis *et al.*, 2012; Zhao *et al.*, 2014). In contrast, SMBA1 is proposed to block the phosphorylation of Bax at serine 184. This phosphorylation is a negative regulator of Bax function, and preventing it should de-repress Bax (Xin *et al.*, 2014).

Summary

The Bcl-2 family proteins are critical regulators of cell death at the mitochondria and the ER. Bcl-2 family proteins regulate this process by a series of protein-protein and protein-membrane interactions that are associated with complex conformational changes in the proteins that ultimately control

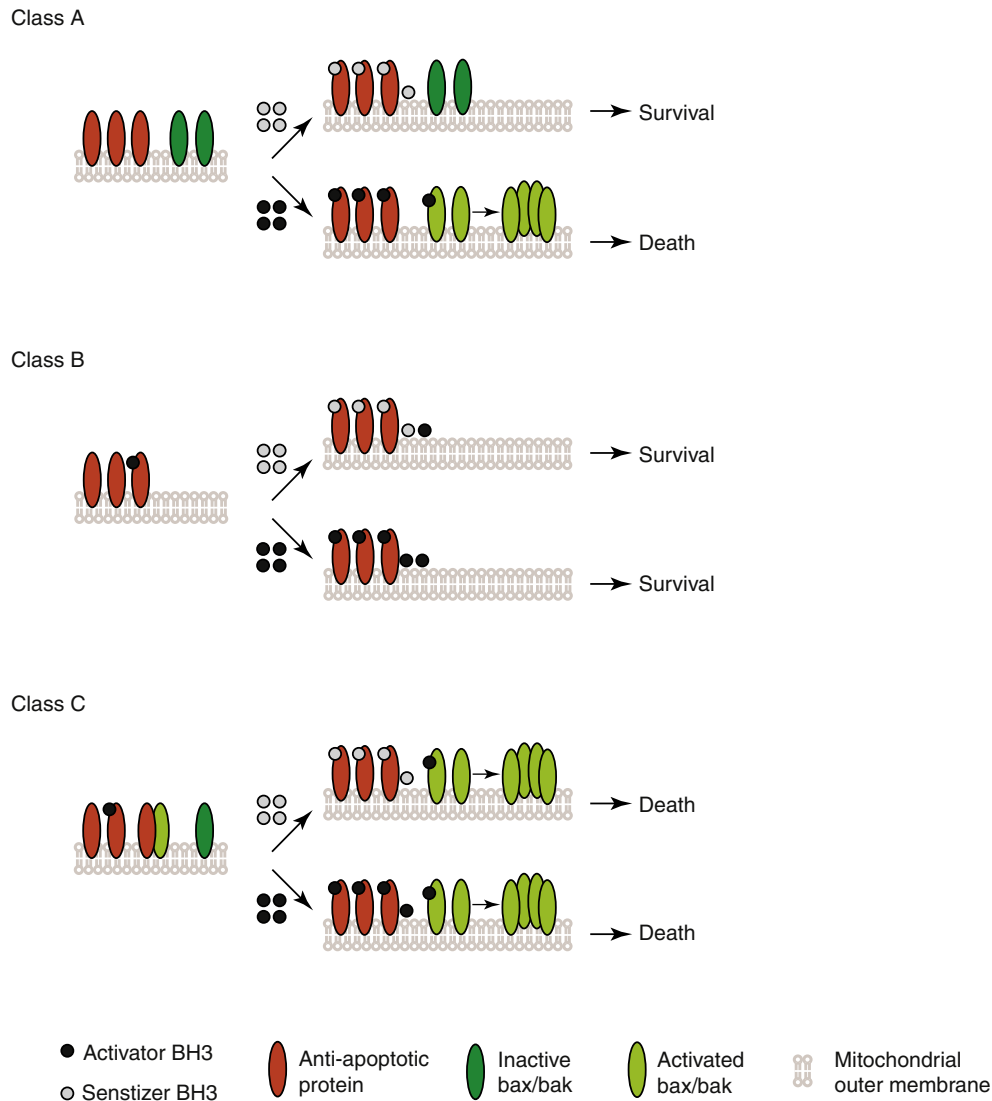


Figure 4 BH3 profiling to identify blocks in apoptosis. In BH3 profiling, mitochondria are incubated with BH3 peptides as representatives of their cognate proteins to examine the pattern of apoptosis response as measured by MOMP *in vitro*. In the ‘Class A’ block, Bax and Bak are inactive and functional BH3 proteins are low or missing and hence MOMP occurs only in response to added activators but not sensitizers. In the ‘Class B’ block, Bax and/or Bak are absent or nonfunctional. Thus, the mitochondria do not respond to the addition of any BH3 peptide. In the ‘Class C’ block, mitochondria are sequestering activator BH3 proteins or activated Bax or Bak. In this scenario, sensitizer peptides induce MOMP based on their specific binding affinities to different anti-apoptotic members. These cancers are said to be ‘primed’ and are candidates for therapies based on inhibitors the anti-apoptotic proteins.

the formation of pores in organelle membranes. Although most, if not all, of the players have been identified, rigorous quantitative modeling of these interactions is still in its infancy and promises great insights in the future (Gaudet *et al.*, 2012). From a therapeutic point of view the concept that ‘addiction’ to anti-apoptotic Bcl-2 members is relevant for cancer maintenance has led to the development of small molecule Bcl-2 inhibitors designed based on structural and functional insights derived from two decades of research. From this perspective, it will be exciting to investigate small molecules that target other interactions among Bcl-2 family members, and to see how such agents fit into treatment as monotherapy or combined with other drugs.

See also: Cell Division/Death: Apoptosis: Apoptosis; Caspases

References

- Annis, M.G., Soucie, E.L., Dlugosz, P.J., *et al.*, 2005. Bax forms multispanning monomers that oligomerize to permeabilize membranes during apoptosis. *EMBO Journal* 24, 2096–2103.
- Aouacheria, A., Brunet, F., Gouy, M., 2005. Phylogenomics of life-or-death switches in multicellular animals: Bcl-2, BH3-Only, and BNip families of apoptotic regulators. *Molecular Biology and Evolution* 22, 2395–2416.

- Axe, E.L., Walker, S.A., Maniava, M., *et al.*, 2008. Autophagosome formation from membrane compartments enriched in phosphatidylinositol 3-phosphate and dynamically connected to the endoplasmic reticulum. *Journal of Cell Biology* 182, 685–701.
- Beroukhi, R., Mermel, C.H., Porter, D., *et al.*, 2010. The landscape of somatic copy-number alteration across human cancers. *Nature* 463, 899–905.
- Bhatt, K., Feng, L., Pabla, N., *et al.*, 2008. Effects of targeted Bcl-2 expression in mitochondria or endoplasmic reticulum on renal tubular cell apoptosis. *American Journal of Physiology. Renal Physiology* 294, F499–F507.
- Billen, L.P., Kokoski, C.L., Lovell, J.F., Leber, B., Andrews, D.W., 2008a. Bcl-XL inhibits membrane permeabilization by competing with Bax. *PLoS Biology* 6, e147.
- Billen, L.P., Shamas-Din, A., Andrews, D.W., 2008b. Bid: A Bax-like BH3 protein. *Oncogene* 27 (Suppl. 1), S93–104.
- Boise, L.H., Gonzalez-Garcia, M., Postema, C.E., *et al.*, 1993. bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death. *Cell* 74, 597–608.
- Bouillet, P., Metcalf, D., Huang, D.C., *et al.*, 1999. Proapoptotic Bcl-2 relative Bim required for certain apoptotic responses, leukocyte homeostasis, and to preclude autoimmunity. *Science* 286, 1735–1738.
- Breckenridge, D.G., Stojanovic, M., Marcellus, R.C., Shore, G.C., 2003. Caspase cleavage product of BAP31 induces mitochondrial fission through endoplasmic reticulum calcium signals, enhancing cytochrome c release to the cytosol. *Journal of Cell Biology* 160, 1115–1127.
- Brouwer, J.M., Westphal, D., Dewson, G., *et al.*, 2014. Bak core and latch domains separate during activation, and freed core domains form symmetric homodimers. *Molecular Cell* 55, 938–946.
- Certo, M., Del Gaizo Moore, V., Nishino, M., *et al.*, 2006. Mitochondria primed by death signals determine cellular addiction to antiapoptotic BCL-2 family members. *Cancer Cell* 9, 351–365.
- Chami, M., Prandini, A., Campanella, M., *et al.*, 2004. Bcl-2 and Bax exert opposing effects on Ca²⁺ signaling, which do not depend on their putative pore-forming region. *Journal of Biological Chemistry* 279, 54581–54589.
- Chang, N.C., Nguyen, M., Germain, M., Shore, G.C., 2010. Antagonism of Beclin 1-dependent autophagy by BCL-2 at the endoplasmic reticulum requires NAF-1. *EMBO Journal* 29, 606–618.
- Chen, L., Willis, S.N., Wei, A., *et al.*, 2005. Differential targeting of prosurvival Bcl-2 proteins by their BH3-only ligands allows complementary apoptotic function. *Molecular Cell* 17, 393–403.
- Cheng, E.H., Sheiko, T.V., Fisher, J.K., Craigen, W.J., Korsmeyer, S.J., 2003. VDAC2 inhibits BAK activation and mitochondrial apoptosis. *Science* 301, 513–517.
- Chipuk, J.E., Bouchier-Hayes, L., Green, D.R., 2006. Mitochondrial outer membrane permeabilization during apoptosis: The innocent bystander scenario. *Cell Death & Differentiation* 13, 1396–1402.
- Cimmino, A., Calin, G.A., Fabbri, M., *et al.*, 2005. miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proceedings of the National Academy of Sciences of the United States of America* 102, 13944–13949.
- Coultas, L., Bouillet, P., Loveland, K.L., *et al.*, 2005. Concomitant loss of proapoptotic BH3-only Bcl-2 antagonists Bik and Bim arrests spermatogenesis. *EMBO Journal* 24, 3963–3973.
- Coultas, L., Terzano, S., Thomas, T., *et al.*, 2007. Hrk/DP5 contributes to the apoptosis of select neuronal populations but is dispensable for haematopoietic cell apoptosis. *Journal of Cell Science* 120, 2044–2052.
- Czabotar, P.E., Westphal, D., Dewson, G., *et al.*, 2013. Bax crystal structures reveal how BH3 domains activate Bax and nucleate its oligomerization to induce apoptosis. *Cell* 152, 519–531.
- van Delft, M.F., Wei, A.H., Mason, K.D., *et al.*, 2006. The BH3 mimetic ABT-737 targets selective Bcl-2 proteins and efficiently induces apoptosis via Bak/Bax if Mcl-1 is neutralized. *Cancer Cell* 10, 389–399.
- Degenhardt, K., Sundararajan, R., Lindsten, T., Thompson, C., White, E., 2002. Bax and Bak independently promote cytochrome C release from mitochondria. *Journal of Biological Chemistry* 277, 14127–14134.
- Del Gaizo Moore, V., Brown, J.R., Certo, M., *et al.*, 2007. Chronic lymphocytic leukemia requires BCL2 to sequester prodeath BIM, explaining sensitivity to BCL2 antagonist ABT-737. *Journal of Clinical Investigation* 117, 112–121.
- Del Gaizo Moore, V., Letai, A., 2013. BH3 profiling — Measuring integrated function of the mitochondrial apoptotic pathway to predict cell fate decisions. *Cancer Letters* 332, 202–205.
- Denisov, A.Y., Madiraju, M.S., Chen, G., *et al.*, 2003. Solution structure of human BCL-w: Modulation of ligand binding by the C-terminal helix. *Journal of Biological Chemistry* 278, 21124–21128.
- Dewson, G., Kratina, T., Sim, H.W., *et al.*, 2008. To trigger apoptosis, Bak exposes its BH3 domain and homodimerizes via BH3:groove interactions. *Molecular Cell* 30, 369–380.
- Edlich, F., Banerjee, S., Suzuki, M., *et al.*, 2011. Bcl-x(L) retrotranslocates Bax from the mitochondria into the cytosol. *Cell* 145, 104–116.
- Ferrer, P.E., Frederick, P., Gulbis, J.M., Dewson, G., Kluck, R.M., 2012. Translocation of a Bak C-terminus mutant from cytosol to mitochondria to mediate cytochrome C release: Implications for Bak and Bax apoptotic function. *PLoS One* 7, e31510.
- Fulda, S., Debatin, K.M., 2006. Extrinsic versus intrinsic apoptosis pathways in anticancer chemotherapy. *Oncogene* 25, 4798–4811.
- Garrido, C., Galluzzi, L., Brunet, M., *et al.*, 2006. Mechanisms of cytochrome c release from mitochondria. *Cell Death & Differentiation* 13, 1423–1433.
- Gaudet, S., Spencer, S.L., Chen, W.W., Sorger, P.K., 2012. Exploring the contextual sensitivity of factors that determine cell-to-cell variability in receptor-mediated apoptosis. *PLoS Computational Biology* 8, e1002482.
- Gavathiotis, E., Reyna, D.E., Bellairs, J.A., Leshchiner, E.S., Walensky, L.D., 2012. Direct and selective small-molecule activation of proapoptotic BAX. *Nature Chemical Biology* 8, 639–645.
- Gavathiotis, E., Suzuki, M., Davis, M.L., *et al.*, 2008. BAX activation is initiated at a novel interaction site. *Nature* 455, 1076–1081.
- George, N.M., Evans, J.J., Luo, X., 2007. A three-helix homo-oligomerization domain containing BH3 and BH1 is responsible for the apoptotic activity of Bax. *Genes & Development* 21, 1937–1948.
- Ghosh, A.P., Klocke, B.J., Ballester, M.E., Roth, K.A., 2012. CHOP potentially co-operates with FOXO3a in neuronal cells to regulate PUMA and BIM expression in response to ER stress. *PLoS One* 7, e39586.
- Green, D.R., Galluzzi, L., Kroemer, G., 2014. Cell biology. Metabolic control of cell death. *Science* 345, 1250256.
- Griffiths, G.J., Dubrez, L., Morgan, C.P., *et al.*, 1999. Cell damage-induced conformational changes of the pro-apoptotic protein Bak *in vivo* precede the onset of apoptosis. *Journal of Cell Biology* 144, 903–914.
- Grillot, D.A., Gonzalez-Garcia, M., Ekhterae, D., *et al.*, 1997. Genomic organization, promoter region analysis, and chromosome localization of the mouse bcl-x gene. *Journal of Immunology* 158, 4750–4757.
- Hamasaki, A., Sando, F., Nakayama, K., *et al.*, 1998. Accelerated neutrophil apoptosis in mice lacking A1-a, a subtype of the bcl-2-related A1 gene. *Journal of Experimental Medicine* 188, 1985–1992.
- Hanahan, D., Weinberg, R.A., 2000. The hallmarks of cancer. *Cell* 100, 57–70.
- Heath-Engel, H.M., Chang, N.C., Shore, G.C., 2008. The endoplasmic reticulum in apoptosis and autophagy: Role of the BCL-2 protein family. *Oncogene* 27, 6419–6433.
- Hetz, C., Bernasconi, P., Fisher, J., *et al.*, 2006. Proapoptotic BAX and BAK modulate the unfolded protein response by a direct interaction with IRE1alpha. *Science* 312, 572–576.
- Hinds, M.G., Smits, C., Fredericks-Short, R., *et al.*, 2007. Bim, Bad and Bmf: intrinsically unstructured BH3-only proteins that undergo a localized conformational change upon binding to prosurvival Bcl-2 targets. *Cell Death & Differentiation* 14, 128–136.
- Hockenbery, D., Nunez, G., Millman, C., Schreiber, R.D., Korsmeyer, S.J., 1990. Bcl-2 is an inner mitochondrial membrane protein that blocks programmed cell death. *Nature* 348, 334–336.
- Hsu, Y.T., Youle, R.J., 1997. Nonionic detergents induce dimerization among members of the Bcl-2 family. *Journal of Biological Chemistry* 272, 13829–13834.
- Ikegaki, N., Katsumata, M., Minna, J., Tsujimoto, Y., 1994. Expression of bcl-2 in small cell lung carcinoma cells. *Cancer Research* 54, 6–8.
- Imaizumi, K., Benito, A., Kiryu-Seo, S., *et al.*, 2004. Critical role for DP5/Haraki, a Bcl-2 homology domain 3-only Bcl-2 family member, in axotomy-induced neuronal cell death. *Journal of Neuroscience* 24, 3721–3725.
- Jeffers, J.R., Parganas, E., Lee, Y., *et al.*, 2003. Puma is an essential mediator of p53-dependent and -independent apoptotic pathways. *Cancer Cell* 4, 321–328.
- Kerr, J.F., Wyllie, A.H., Currie, A.R., 1972. Apoptosis: A basic biological phenomenon with wide-ranging implications in tissue kinetics. *British Journal of Cancer* 26, 239–257.
- Kim, H., Tu, H.C., Ren, D., *et al.*, 2009. Stepwise activation of BAX and BAK by tBID, BIM, and PUMA initiates mitochondrial apoptosis. *Molecular Cell* 36, 487–499.
- Knudson, C.M., Tung, K.S., Tourtellotte, W.G., Brown, G.A., Korsmeyer, S.J., 1995. Bax-deficient mice with lymphoid hyperplasia and male germ cell death. *Science* 270, 96–99.
- Kutuk, O., Letai, A., 2008. Regulation of Bcl-2 family proteins by posttranslational modifications. *Current Molecular Medicine* 8, 102–118.

- Kuwana, T., Bouchier-Hayes, L., Chipuk, J.E., *et al.*, 2005. BH3 domains of BH3-only proteins differentially regulate Bax-mediated mitochondrial membrane permeabilization both directly and indirectly. *Molecular Cell* 17, 525–535.
- Kvansakul, M., Yang, H., Fairlie, W.D., *et al.*, 2008. Vaccinia virus anti-apoptotic F1L is a novel Bcl-2-like domain-swapped dimer that binds a highly selective subset of BH3-containing death ligands. *Cell Death & Differentiation* 15, 1564–1571.
- Labi, V., Erlacher, M., Kiessling, S., *et al.*, 2008. Loss of the BH3-only protein Bim impairs B cell homeostasis and accelerates gamma irradiation-induced thymic lymphoma development. *Journal of Experimental Medicine* 205, 641–655.
- Leber, B., Lin, J., Andrews, D.W., 2007. Embedded together: The life and death consequences of interaction of the Bcl-2 family with membranes. *Apoptosis* 12, 897–911.
- Leber, B., Lin, J., Andrews, D.W., 2010. Still embedded together binding to membranes regulates Bcl-2 protein interactions. *Oncogene* 29, 5221–5230.
- Lee, J.W., Soung, Y.H., Kim, S.Y., *et al.*, 2004. Inactivating mutations of proapoptotic Bad gene in human colon cancers. *Carcinogenesis* 25, 1371–1376.
- Leshchiner, E.S., Braun, C.R., Bird, G.H., Walensky, L.D., 2013. Direct activation of full-length proapoptotic BAK. *Proceedings of the National Academy of Sciences of the United States of America* 110, E986–E995.
- Letai, A., Bassik, M.C., Walensky, L.D., *et al.*, 2002. Distinct BH3 domains either sensitize or activate mitochondrial apoptosis, serving as prototype cancer therapeutics. *Cancer Cell* 2, 183–192.
- Levine, B., Kroemer, G., 2008. Autophagy in the pathogenesis of disease. *Cell* 132, 27–42.
- Lindsten, T., Ross, A.J., King, A., *et al.*, 2000. The combined functions of proapoptotic Bcl-2 family members bak and bax are essential for normal development of multiple tissues. *Molecular Cell* 6, 1389–1399.
- Liu, X., Dai, S., Zhu, Y., Marrack, P., Kappler, J.W., 2003. The structure of a Bcl-xL/Bim fragment complex: Implications for Bim function. *Immunity* 19, 341–352.
- Liambi, F., Moldoveanu, T., Tait, S.W., *et al.*, 2011. A unified model of mammalian BCL-2 protein family interactions at the mitochondria. *Molecular Cell* 44, 517–531.
- Lovell, J.F., Billen, L.P., Bindner, S., *et al.*, 2008. Membrane binding by tBid initiates an ordered series of events culminating in membrane permeabilization by Bax. *Cell* 135, 1074–1084.
- Maiuri, M.C., Le Toumelin, G., Criollo, A., *et al.*, 2007. Functional and physical interaction between Bcl-X(L) and a BH3-like domain in Beclin-1. *EMBO Journal* 26, 2527–2539.
- Mathai, J.P., Germain, M., Shore, G.C., 2005. BH3-only BIK regulates BAX, BAK-dependent release of Ca²⁺ from endoplasmic reticulum stores and mitochondrial apoptosis during stress-induced cell death. *Journal of Biological Chemistry* 280, 23829–23836.
- Meier, P., Finch, A., Evan, G., 2000. Apoptosis in development. *Nature* 407, 796–801.
- Miyashita, T., Krajewski, S., Krajewska, M., *et al.*, 1994. Tumor suppressor p53 is a regulator of bcl-2 and bax gene expression *in vitro* and *in vivo*. *Oncogene* 9, 1799–1805.
- Moldoveanu, T., Liu, Q., Tocili, A., *et al.*, 2006. The X-ray structure of a BAK homodimer reveals an inhibitory zinc binding site. *Molecular Cell* 24, 677–688.
- Motoyama, N., Wang, F., Roth, K.A., *et al.*, 1995. Massive cell death of immature hematopoietic cells and neurons in Bcl-x-deficient mice. *Science* 267, 1506–1510.
- Mukhopadhyay, S., Panda, P.K., Sinha, N., Das, D.N., Bhutia, S.K., 2014. Autophagy and apoptosis: where do they meet? *Apoptosis* 19, 555–566.
- Nakano, K., Voutsden, K.H., 2001. PUMA, a novel proapoptotic gene, is induced by p53. *Molecular Cell* 7, 683–694.
- Ng, F.W., Nguyen, M., Kwan, T., *et al.*, 1997. p28 Bap31, a Bcl-2/Bcl-XL- and procaspase-8-associated protein in the endoplasmic reticulum. *Journal of Cell Biology* 139, 327–338.
- Nutt, L.K., Pataer, A., Pahler, J., *et al.*, 2002. Bax and Bak promote apoptosis by modulating endoplasmic reticular and mitochondrial Ca²⁺ stores. *Journal of Biological Chemistry* 277, 9219–9225.
- O'Connor, L., Strasser, A., O'Reilly, L.A., *et al.*, 1998. Bim: a novel member of the Bcl-2 family that promotes apoptosis. *EMBO Journal* 17, 384–395.
- Oakes, S.A., Lin, S.S., Bassik, M.C., 2006. The control of endoplasmic reticulum-initiated apoptosis by the BCL-2 family of proteins. *Current Molecular Medicine* 6, 99–109.
- Oda, E., Ohki, R., Murasawa, H., *et al.*, 2000. Noxa, a BH3-only member of the Bcl-2 family and candidate mediator of p53-induced apoptosis. *Science* 288, 1053–1058.
- Pattingre, S., Tassa, A., Qu, X., *et al.*, 2005. Bcl-2 antiapoptotic proteins inhibit Beclin 1-dependent autophagy. *Cell* 122, 927–939.
- Petros, A.M., Nettesheim, D.G., Wang, Y., *et al.*, 2000. Rationale for Bcl-xL/Bad peptide complex formation from structure, mutagenesis, and biophysical studies. *Protein Science* 9, 2528–2534.
- Petros, A.M., Olejniczak, E.T., Fesik, S.W., 2004. Structural biology of the Bcl-2 family of proteins. *Biochimica et Biophysica Acta* 1644, 83–94.
- Peyerl, F.W., Dai, S., Murphy, G.A., *et al.*, 2007. Elucidation of some Bax conformational changes through crystallization of an antibody-peptide complex. *Cell Death & Differentiation* 14, 447–452.
- Print, C.G., Loveland, K.L., Gibson, L., *et al.*, 1998. Apoptosis regulator bcl-w is essential for spermatogenesis but appears otherwise redundant. *Proceedings of the National Academy of Sciences of the United States of America* 95, 12424–12431.
- Puthalakath, H., O'Reilly, L.A., Gunn, P., *et al.*, 2007. ER stress triggers apoptosis by activating BH3-only protein Bim. *Cell* 129, 1337–1349.
- Rampino, N., Yamamoto, H., Ionov, Y., *et al.*, 1997. Somatic frameshift mutations in the BAX gene in colon cancers of the microsatellite mutator phenotype. *Science* 275, 967–969.
- Ranger, A.M., Zha, J., Harada, H., *et al.*, 2003. Bad-deficient mice develop diffuse large B cell lymphoma. *Proceedings of the National Academy of Sciences of the United States of America* 100, 9324–9329.
- Rinkenberger, J.L., Horning, S., Klocke, B., Roth, K., Korsmeyer, S.J., 2000. Mcl-1 deficiency results in peri-implantation embryonic lethality. *Genes & Development* 14, 23–27.
- Rong, Y., Distelhorst, C.W., 2008. Bcl-2 protein family members: versatile regulators of calcium signaling in cell survival and apoptosis. *Annual Review of Physiology* 70, 73–91.
- Rong, Y.P., Aromolaran, A.S., Bultynck, G., *et al.*, 2008. Targeting Bcl-2-IP3 receptor interaction to reverse Bcl-2's inhibition of apoptotic calcium signals. *Molecular Cell* 31, 255–265.
- Rong, Y.P., Bultynck, G., Aromolaran, A.S., *et al.*, 2009. The BH4 domain of Bcl-2 inhibits ER calcium release and apoptosis by binding the regulatory and coupling domain of the IP3 receptor. *Proceedings of the National Academy of Sciences of the United States of America* 106, 14397–14402.
- Sato, T., Irie, S., Krajewski, S., Reed, J.C., 1994. Cloning and sequencing of a cDNA encoding the rat Bcl-2 protein. *Gene* 140, 291–292.
- Sattler, M., Liang, H., Nettesheim, D., *et al.*, 1997. Structure of Bcl-xL-Bak peptide complex: recognition between regulators of apoptosis. *Science* 275, 983–986.
- Schinzl, A., Kaufmann, T., Schuler, M., *et al.*, 2004. Conformational control of Bax localization and apoptotic activity by Pro168. *Journal of Cell Biology* 164, 1021–1032.
- Scorrano, L., Oakes, S.A., Opferman, J.T., *et al.*, 2003. BAX and BAK regulation of endoplasmic reticulum Ca²⁺ : A control point for apoptosis. *Science* 300, 135–139.
- Shamas-Din, A., Bindner, S., Zhu, W., *et al.*, 2013a. tBid undergoes multiple conformational changes at the membrane required for Bax activation. *Journal of Biological Chemistry* 288, 22111–22127.
- Shamas-Din, A., Kale, J., Leber, B., Andrews, D.W., 2013b. Mechanisms of action of Bcl-2 family proteins. *Cold Spring Harbor Perspectives in Biology* 5, a008714.
- Shibue, T., Takeda, K., Oda, E., *et al.*, 2003. Integral role of Noxa in p53-mediated apoptotic response. *Genes & Development* 17, 2233–2238.
- Silva, M., Benito, A., Sanz, C., *et al.*, 1999. Erythropoietin can induce the expression of bcl-x(L) through Stat5 in erythropoietin-dependent progenitor cell lines. *Journal of Biological Chemistry* 274, 22165–22169.
- Souers, A.J., Levenson, J.D., Boghaert, E.R., *et al.*, 2013. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. *Nature Medicine* 19, 202–208.
- Suzuki, M., Youle, R.J., Tjandra, N., 2000. Structure of Bax: Coregulation of dimer formation and intracellular localization. *Cell* 103, 645–654.
- Todt, F., Cakir, Z., Reichenbach, F., Youle, R.J., Edlich, F., 2013. The C-terminal helix of Bcl-x(L) mediates Bax retrotranslocation from the mitochondria. *Cell Death & Differentiation* 20, 333–342.
- Tse, C., Shoemaker, A.R., Adickes, J., *et al.*, 2008. ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. *Cancer Research* 68, 3421–3428.
- Tsujimoto, Y., Cossman, J., Jaffe, E., Croce, C.M., 1985. Involvement of the bcl-2 gene in human follicular lymphoma. *Science* 228, 1440–1443.
- Varin, E., Denoyelle, C., Brodin, E., *et al.*, 2010. Downregulation of Bcl-xL and Mcl-1 is sufficient to induce cell death in mesothelioma cells highly refractory to conventional chemotherapy. *Carcinogenesis* 31, 984–993.
- Vaux, D.L., Cory, S., Adams, J.M., 1988. Bcl-2 gene promotes haemopoietic cell survival and cooperates with c-myc to immortalize pre-B cells. *Nature* 335, 440–442.

- Veis, D.J., Sorenson, C.M., Shutter, J.R., Korsmeyer, S.J., 1993. Bcl-2-deficient mice demonstrate fulminant lymphoid apoptosis, polycystic kidneys, and hypopigmented hair. *Cell* 75, 229–240.
- Villunger, A., Michalak, E.M., Coultas, L., *et al.*, 2003. p53- and drug-induced apoptotic responses mediated by BH3-only proteins puma and noxa. *Science* 302, 1036–1038.
- Vogler, M., Dinsdale, D., Dyer, M.J., Cohen, G.M., 2009. Bcl-2 inhibitors: small molecules with a big impact on cancer therapy. *Cell Death & Differentiation* 16, 360–367.
- Vogler, M., Hamali, H.A., Sun, X.M., *et al.*, 2011. BCL2/BCL-X(L) inhibition induces apoptosis, disrupts cellular calcium homeostasis, and prevents platelet activation. *Blood* 117, 7145–7154.
- Westphal, D., Kluck, R.M., Dewson, G., 2014. Building blocks of the apoptotic pore: How Bax and Bak are activated and oligomerize during apoptosis. *Cell Death & Differentiation* 21, 196–205.
- Willis, S.N., Chen, L., Dewson, G., *et al.*, 2005. Proapoptotic Bak is sequestered by Mcl-1 and Bcl-xL, but not Bcl-2, until displaced by BH3-only proteins. *Genes & Development* 19, 1294–1305.
- Willis, S.N., Fletcher, J.I., Kaufmann, T., *et al.*, 2007. Apoptosis initiated when BH3 ligands engage multiple Bcl-2 homologs, not Bax or Bak. *Science* 315, 856–859.
- Wolter, K.G., Hsu, Y.T., Smith, C.L., *et al.*, 1997. Movement of Bax from the cytosol to mitochondria during apoptosis. *Journal of Cell Biology* 139, 1281–1292.
- Xin, M., Li, R., Xie, M., *et al.*, 2014. Small-molecule Bax agonists for cancer therapy. *Nature Communications* 5, 4935.
- Yamamoto, H., Sawai, H., Perucho, M., 1997. Frameshift somatic mutations in gastrointestinal cancer of the microsatellite mutator phenotype. *Cancer Research* 57, 4420–4426.
- Yethon, J.A., Epand, R.F., Leber, B., Epand, R.M., Andrews, D.W., 2003. Interaction with a membrane surface triggers a reversible conformational change in Bax normally associated with induction of apoptosis. *Journal of Biological Chemistry* 278, 48935–48941.
- Yin, X.M., Wang, K., Gross, A., *et al.*, 1999. Bid-deficient mice are resistant to Fas-induced hepatocellular apoptosis. *Nature* 400, 886–891.
- Yue, Z., Jin, S., Yang, C., Levine, A.J., Heintz, N., 2003. Beclin 1, an autophagy gene essential for early embryonic development, is a haploinsufficient tumor suppressor. *Proceedings of the National Academy of Sciences of the United States of America* 100, 15077–15082.
- Zha, J., Harada, H., Yang, E., Jockel, J., Korsmeyer, S.J., 1996. Serine phosphorylation of death agonist BAD in response to survival factor results in binding to 14-3-3 not BCL-X(L). *Cell* 87, 619–628.
- Zhao, G., Zhu, Y., Eno, C.O., *et al.*, 2014. Activation of the proapoptotic Bcl-2 protein Bax by a small molecule induces tumor cell apoptosis. *Molecular and Cellular Biology* 34, 1198–1207.
- Zhu, W., Cowie, A., Wasfy, G.W., *et al.*, 1996. Bcl-2 mutants with restricted subcellular location reveal spatially distinct pathways for apoptosis in different cell types. *EMBO Journal* 15, 4130–4141.
- Zong, W.X., Li, C., Hatzivassiliou, G., *et al.*, 2003. Bax and Bak can localize to the endoplasmic reticulum to initiate apoptosis. *Journal of Cell Biology* 162, 59–69.

Further Reading

- Chi, X., Kale, J., Leber, B., Andrews, D.W., 2014. Regulating cell death at, on, and in membranes. *Biochimica et Biophysica Acta* 1843, 2100–2113.
- Czabotar, P.E., Lessene, G., Strasser, A., Adams, J.M., 2014. Control of apoptosis by the BCL-2 protein family: Implications for physiology and therapy. *Nature Reviews Molecular Cell Biology* 15, 49–63.
- Galluzzi, L., Aaronson, S.A., Abrams, J., *et al.*, 2009. Guidelines for the use and interpretation of assays for monitoring cell death in higher eukaryotes. *Cell Death & Differentiation* 16, 1093–1107.
- Kale, J., Liu, Q., Leber, B., Andrews, D.W., 2012. Shedding light on apoptosis at subcellular membranes. *Cell* 151, 1179–1184.
- Kroemer, G., Marino, G., Levine, B., 2010. Autophagy and the integrated stress response. *Molecular Cell* 40, 280–293.
- Tait, S.W., Green, D.R., 2012. Mitochondria and cell signalling. *Journal of Cell Science* 125, 807–815.

Relevant Websites

- <http://www.cellsignal.com/common/content/content.jsp?id=science-pathways-apoptosis> Cell Signalling. 'Apoptosis resources'.
- <http://www.biooncology.com/research-education/apoptosis/apoptosis> Genetech Biooncology. 'Apoptosis: A Critical Process in Homeostasis'.
- http://www.rndsystems.com/molecule_group.aspxg=464 R&D systems. 'Bcl-2 Proteins and Regulators'.
- <http://www.nature.com/scitable/topicpage/the-discovery-of-lysosomes-and-autophagy-14199828> Susana Castro-Obregon. 'The Discovery of Lysosomes and Autophagy'.
- <http://www.asbmb.org/asbmbtoday/201408/JournalNews/JBC/UPR/> Teodora Donisan. 'Unfolded Protein response signaling and metabolic diseases'.

Inhibitor of Apoptosis Proteins, the Sentinels of Cell Death and Signaling

M Cristina de Almagro and D Vucic, Genentech, South San Francisco, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Apoptosis It is a programmed type of cell death that requires caspases activation. Apoptosis is characterized by cellular shrinkage, nuclear fragmentation, chromosome condensation, blebbing of the plasma membrane, and late loss of plasma membrane integrity.

Necroptosis It is a regulated form of cell death that does not involve caspase activation but requires kinases RIP1 and

RIP3. This type of cell death is characterized by organelle swelling, lack of nuclear fragmentation, early loss of plasma membrane integrity and inflammation.

Ubiquitination It is a posttranslational modification where ubiquitin is attached to a lysine or N-terminal methionine on a substrate protein. Ubiquitination is carried out by three enzymes: E1, E2, and E3 with E3 ubiquitin ligases providing substrate specificity.

Proper regulation of cell death and survival is crucial for cellular and organismal homeostasis. To this end, eukaryotes have evolved numerous proteins to achieve a fine-tuned and precise control of cell fate. Importantly, dysregulated cell death can lead to severe diseases such as cancer or neurodegeneration. Inhibitors of Apoptosis, or IAP, proteins constitute a family of 8 proteins that are critically involved in inhibition of cell death, but also in proinflammatory signaling, cell division and migration (Salvesen and Duckett, 2002; Fulda, 2007). Interestingly, even though all the IAP family members are involved in the regulation of cell survival, they do it in quite different ways.

Structural Organization of IAPs

IAP proteins and their signature Baculovirus IAP repeat (BIR) domains were first identified in baculoviruses. Infection of insect cells with baculoviruses triggers cell death that can be inhibited by virally encoded antiapoptotic proteins (Birnbaum *et al.*, 1994; Clem *et al.*, 1991). Genetic screens based on rescuing cell death caused by infection with baculovirus strains deficient for caspase inhibitor p35 lead to the identification of baculovirus inhibitors of apoptosis as the mediators of cell death resistance (Birnbaum *et al.*, 1994; Clem *et al.*, 1991; Crook *et al.*, 1993).

BIR Domain

The evolutionarily conserved IAP family is characterized by the presence of at least one BIR domain (Figure 1). Some members of the IAP family have three BIR domains such as c-IAP1, c-IAP2, XIAP, and NAIP, while others have one: ML-IAP, Survivin, ILP-2, and Bruce. Apart from the BIR domain, IAP proteins have a diversity of other domains that confer the specific functions and characteristics of each member of the IAP family (Varfolomeev and Vucic, 2011).

The BIR domain is approximately 70–80 amino acids, and contains a zinc-binding site (Birnbaum *et al.*, 1994). Functionally, the BIR domains allow protein–protein interactions, and for XIAP, they mediate interactions with caspases (Srinivasula and Ashwell, 2008; Vaux and Silke, 2005). Many

of the BIR-mediated protein–protein interactions occur through a pocket on the BIR domains that binds to an N-terminal tetrapeptide motif called IAP-binding motifs (IBMs). But, the specific BIR sequence of each IAP protein is very important since small variations in the sequence can lead to different interaction partners:

- XIAP: binds caspase-3 and -7 via BIR2 and the preceding linker region; and SMAC/Diablo (Secondary mitochondria-derived activator of caspases) and caspase-9 through BIR3
- c-IAP1/2: bind TRAF2 via BIR1 and SMAC through BIR3
- ML-IAP, ILP-2, and Bruce: bind SMAC.

RING, UBC, and UBA Domains

The IAP proteins c-IAP1, c-IAP2, XIAP, and ML-IAP contain a really interesting new gene (RING) domain that confers E3 ligase ubiquitin activity. The RING domain allows these IAPs to regulate their binding partners by ubiquitination, as well as to promote their autoubiquitination. Ubiquitination is a three-step process initiated by an E1 ubiquitin-activating enzyme followed by charging an ubiquitin-conjugating E2 enzyme and finalized an E3 ligase that allows the transfer of the ubiquitin moiety to a lysine on the substrate protein forming monoubiquitins or ubiquitin chains (Hershko and Ciechanover, 1998; Pickart and Fushman, 2004). When ubiquitin chains are assembled, different linkages between ubiquitins can be formed depending on the ubiquitin lysine used or the N-terminal methionine. There are eight different kinds of ubiquitin chains: linear (through N-terminus), K6, K11, K27, K29, K33, K48, K63. K48 chains usually target proteins for proteasomal degradation, while K11 and K63 polyubiquitin chains trigger the formation of signaling complexes. The RING domains of IAP proteins mostly collaborate with UbcH5 and UbcH6 E2 enzymes to promote ubiquitination (Dynek *et al.*, 2010; Varfolomeev *et al.*, 2012a,b).

Instead of a RING domain, Bruce has a ubiquitin-conjugating (UBC) domain that allows the transfer of ubiquitin moieties to its substrates (Hauser *et al.*, 1998). XIAP, c-IAP-1/2, and ILP-2 also contain a ubiquitin-associated (UBA) domain that allows the binding to various ubiquitin moieties (Dynek

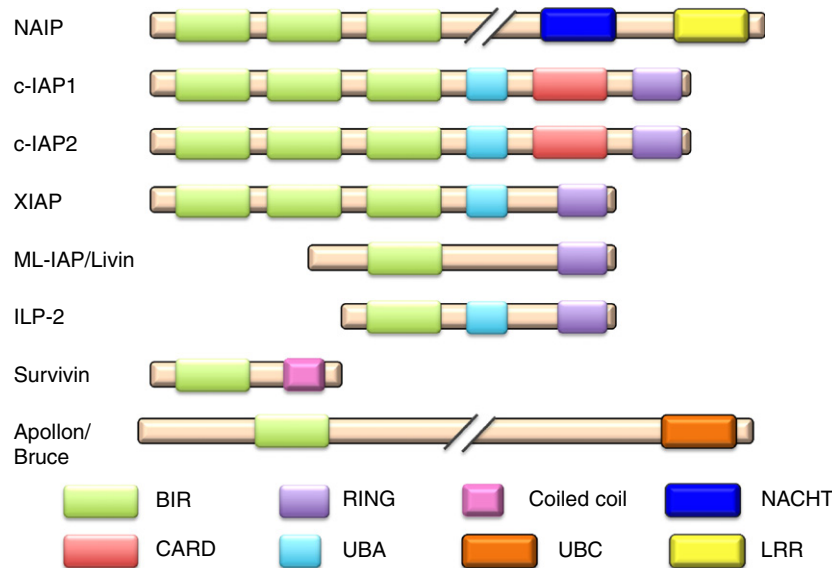


Figure 1 Inhibitor of apoptosis protein family members. The structural organization of the eight IAP protein.

et al., 2010; Blankenship *et al.*, 2009; Gyrd-Hansen *et al.*, 2008).

CARD, NATCH, and LRR Domains

The caspase activating and recruitment domain (CARD) present in c-IAP1 and 2 does not have a clear function yet (Hofmann *et al.*, 1997). Usually the CARD domain allows the interaction with other CARD-containing proteins, but c-IAP CARD binding partners have proven elusive thus far.

The only IAP protein with a NATCH domain is NAIP. The NATCH domain is found in proteins involved in immunity and NOD (nuclear binding and oligomerization domain)-mediated signaling pathways that sense intracellular pathogens. It consists of a NTPase domain and its name arises from the first letters of the proteins NAIP, CIITA, HET-E, and TP1 (Koonin and Aravind, 2000). NAIP also possesses a leucine-rich repeat (LRR) domain that gives it the ability to sense intracellular pathogens.

IAP Protein Characteristics and Function

XIAP (BIRC4)

Apoptosis regulated by XIAP

X-linked Inhibitor of apoptosis protein was found through sequence homology with baculovirus IAPs (Duckett *et al.*, 1996). XIAP primarily blocks apoptosis by directly inhibiting several caspases. XIAP can regulate either extrinsic apoptosis (induced by death ligands) or intrinsic apoptosis (mitochondria dependent intracellular signals) (Figure 2). The BIR3 domain of XIAP binds and inhibits caspase-9 by blocking its dimerization, which is needed for autocleavage. BIR2 and the linker region between BIR1 and BIR2 are responsible for the binding and inhibition of the active caspases-3 and -7 (Chai *et al.*, 2001; Riedl *et al.*, 2001; Eckelman *et al.*, 2006; Scott *et al.*, 2005). The E3 ligase activity of XIAP may also contribute to

the inhibition of apoptosis by targeting SMAC, caspases and other pro-apoptotic proteins for proteasomal degradation (Vucic *et al.*, 2011).

IAP proteins and NOD2 signaling

The NOD1/2 pathway is involved in innate immunity as NOD proteins sense pathogens and regulate the expression of genes involved in pathogen-host defense. NOD1/2 are intracellular sensors that oligomerize upon binding to their ligands, and recruit the kinase RIP2 (receptor TNFRSF-interacting serine-threonine kinase 2), XIAP and c-IAPs (Bertrand *et al.*, 2009; Krieg *et al.*, 2009; Damgaard *et al.*, 2012). The binding of these E3 ligases triggers the ubiquitination of RIP2 allowing the recruitment of TAK1, TAB1/2/3, Nemo (NF-kappa-B essential modulator) and LUBAC (linear ubiquitin chain assembly complex) leading to the NF- κ B, MAP kinases p38 and c-Jun N-terminal kinases (JNK) activation (Hasegawa *et al.*, 2008).

Mutations in XIAP were recently found in patients with the X-linked lymphoproliferative syndrome type-2 (XLP2). These mutations were found in the BIR2 and RING domains of XIAP, and resulted in impaired NOD1/2 (nucleotide-binding oligomerization domain) signaling. XLP2-related mutations found in XIAP prevent the binding of XIAP to RIP2 and subsequent ubiquitination of RIP2 by XIAP. Without ubiquitination of RIP2, NF- κ B, and MAPK signaling pathways cannot get activated thus preventing the expression of genes important for innate immunity and pathogen defense (Damgaard *et al.*, 2013).

c-IAP1 (BIRC2) and c-IAP2 (BIRC3)

Cellular inhibitors of apoptosis proteins are probably the best-studied members of the IAP family. These two proteins share the same domain structure and a noticeable sequence homology. Not surprisingly, their activities are highly redundant (Varfolomeev and Vucic, 2008). One explanation for the existence of these two proteins, which arose likely from gene

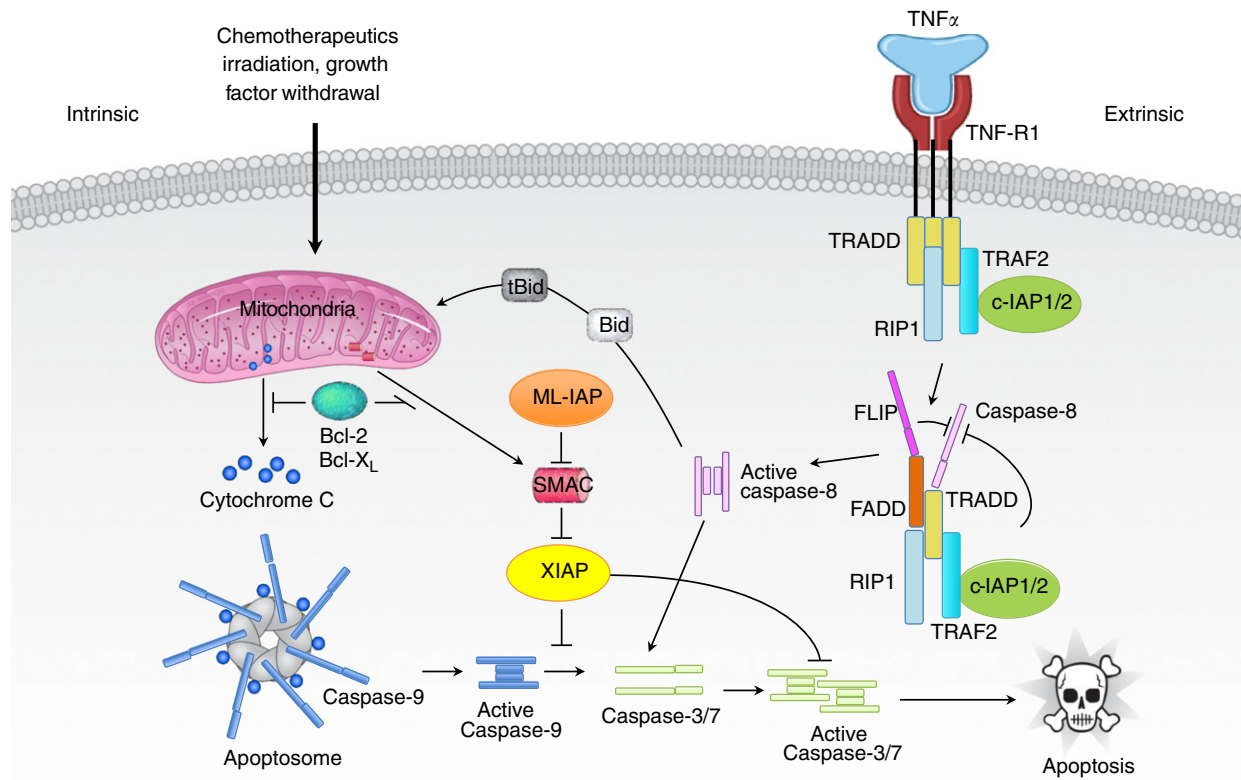


Figure 2 The intrinsic and extrinsic apoptotic pathways. IAPs can inhibit apoptosis triggered by external or intracellular stimuli.

duplication, is the critical nature of the cellular processes that they regulate. However, in some specific cell lineages such as B cells and T cells, differential roles between c-IAP proteins have been observed (Conze *et al.*, 2010; Giardino Torchia *et al.*, 2013).

It was initially thought that c-IAPs could bind to caspases and inhibit them, but it is clear now that although they interact with some caspases, they have very little inhibitory activity (Eckelman and Salvesen, 2006). The main mechanism by which c-IAPs inhibit apoptosis is likely the regulation of caspase-8 activation through RIP1 ubiquitination, as well as by the modulation of NF- κ B signaling pathways. NF- κ B signaling induces the transcription of pro-survival and cell growth genes via two different pathways: canonical and noncanonical.

IAP proteins and NF- κ B pathways

When the cytokine TNF α (Tumor Necrosis Factor α) binds to the death receptor TNFRI (TNF receptor I) in the plasma membrane, it induces the trimerization of TNFRI that triggers the assembly of TNFRI signaling complex (Figure 3). First to bind the intracellular domain of TNFRI is the adaptor protein TRADD (TNF receptor-associated death domain protein), followed by the protein kinase RIP1 and TRAF2 (tumor necrosis factor receptor-associated factor 2). TRAF2 also trimerizes, and it behaves as the adaptor protein that brings c-IAP1 and c-IAP2 to the TNFRI complex. c-IAPs are the only members of the IAP family capable of binding TRAF1 and TRAF2 through the first two α helices in their BIR1 domains (Samuel *et al.*, 2006; Varfolomeev *et al.*, 2006). Proximity within TNFRI complex permits c-IAP proteins dimerization, leading to the activation

of their E3 ubiquitin ligase activity. c-IAPs ubiquitinate RIP1 with K11- and K63-linked polyubiquitin chains, which allows the recruitment of downstream mediators of the NF- κ B signaling pathway to the TNFRI complex (Varfolomeev *et al.*, 2008). The essential role of c-IAPs in preventing apoptosis is mainly mediated by ubiquitination as these ubiquitination events prevent the translocation of RIP1 to the caspase-8-containing complex and cell death induction. Besides the key inhibition of Complex II formation (Caspase-8-containing complex), RIP1 ubiquitination also contributes to cell survival by NF- κ B activation. c-IAPs can ubiquitinate other molecules of the NF- κ B pathway as well, such as NEMO and TRAF2, and that ubiquitination modulates the activity of signaling complexes (Vucic *et al.*, 2011). Upon RIP1 and NEMO ubiquitination by c-IAPs, the IKK, TAB2/3-TAK1 and the LUBAC (HOIP, HOIL and Sharpin) complexes bind to ubiquitinated RIP1 and NEMO (Shim *et al.*, 2005; Wang *et al.*, 2001) leading to their activation, which results in the phosphorylation of I κ B α (inhibitor of kappa B) by the kinase IKK β . I κ B α sequesters NF- κ B transcription factors in the cytoplasm, but once it gets phosphorylated it undergoes ubiquitination and proteasomal degradation. This frees the NF- κ B transcription factors p50 and RelA/p65 to move into the nucleus and induce the transcription of cell survival genes (Scheidereit, 2006). In a positive feedback loop, c-IAP2 expression is upregulated by NF- κ B activation (Wang *et al.*, 1998).

Noncanonical NF- κ B pathway uses a different mechanism to activate the translocation of the NF- κ B transcription factors to the nucleus. Noncanonical NF- κ B signaling is slower than canonical NF- κ B: while the canonical pathway takes minutes

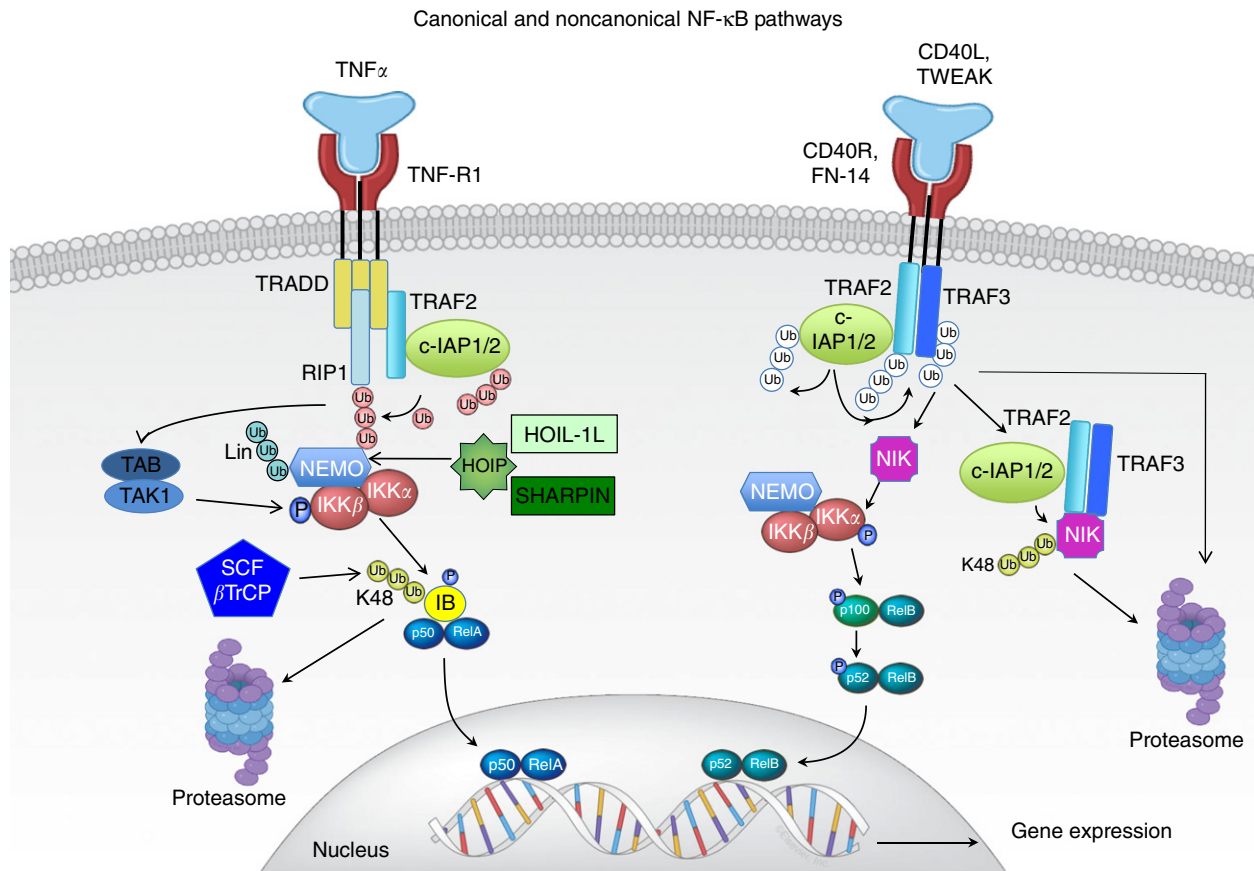


Figure 3 Canonical and noncanonical NF- κ B pathways. c-IAPs regulate the two different activation routes of NF- κ B signaling pathway.

for activation, noncanonical signaling takes hours to get activated (Dejardin, 2006). Interestingly, c-IAPs have opposite roles in the activation of these pathways: in canonical NF- κ B signaling c-IAPs are needed for the pathway activation, whereas noncanonical signaling requires loss of c-IAP proteins (Varfolomeev and Vucic, 2008).

Several TNF family ligands (TWEAK, LIGHT) can induce noncanonical NF- κ B signaling upon binding to their respective receptors. Cellular IAPs are recruited to these receptor complexes through the adaptor protein TRAF2 (Varfolomeev *et al.*, 2012a,b). The E3 ligase activity of c-IAPs constitutively ubiquitinates NIK (NF- κ B Inducing Kinase) to target it for proteasomal degradation and keep the NF- κ B signaling switched off (Vallabhapurapu *et al.*, 2008). When c-IAPs are not present, newly synthesized NIK accumulates in a period of a couple of hours. Once NIK protein reaches a certain threshold level, it phosphorylates IKK α /IKK β , leading to p100 phosphorylation and partial proteasomal processing to yield the p52 form (Dejardin, 2006). Transcription factor p52 then translocates to the nucleus in complex with RelB leading to the transcription of NF- κ B dependent genes.

Cell death ligand driven apoptosis and necroptosis

Cellular IAP proteins regulate two different types of cell death: apoptosis and necroptosis (Silke and Meier, 2013). Apoptosis is a caspase-dependent programmed type of cell death, while necroptosis is also a regulated form of cell death, but caspase

independent. In addition to being caspase-independent, necroptosis can be differentiated from apoptosis because it undergoes with organelle swelling, no nuclear fragmentation, early loss of plasma membrane integrity, inflammation, and requires the expression of the protein kinase RIP3, while apoptosis does not present these features (Zong and Thompson, 2006).

The mechanism by which c-IAPs regulate extrinsic apoptosis and necroptosis is essentially the same, as both pathways share the first signaling steps. When TNF α binds and activates the TNFR1, and c-IAPs are not present to ubiquitinate RIP1, RIP1 leaves the TNFR1-associated complex, or Complex I, and binds the FAS-associated protein with death domain (FADD)-caspase-8/10 complex (Feoktistova *et al.*, 2011; Figure 2). Once this Complex II is formed, cells may undergo apoptosis or necroptosis based on the activation of caspases and the presence of the kinase RIP3. After c-IAP depletion, cells could die of apoptosis if caspase-8/10 gets activated to start the caspase signaling cascade that will lead to apoptotic cell death. However, when caspases are not active or are inhibited cells find an alternative way to die: necroptosis. In this case, the kinase RIP3 will bind to the kinase RIP1 and they will get phosphorylated, leading to the phosphorylation of the pseudokinase MLKL (Mixed lineage kinase like) (Sun *et al.*, 2012; Zhao *et al.*, 2012; Figure 4). RIP3-mediated phosphorylation of MLKL causes MLKL oligomerization and translocation to the plasma membrane to mediate cell death (Cai *et al.*, 2014).

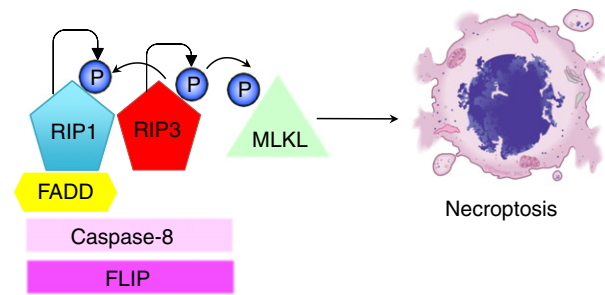


Figure 4 Necroptosis. RIP1, RIP3, and MLKL are crucial proteins for the necroptotic pathway. Their activation is based on auto- and cross-phosphorylation.

Therefore, c-IAPs are not just inhibitors of apoptosis proteins; they can also inhibit necroptotic cell death.

c-IAPs in immunity

Apart from the functions of c-IAPs in cell death, many groups have been trying to elucidate the importance of cellular IAP proteins in innate immunity and inflammation (Silke and Brink, 2010). In addition to their seminal roles in NF- κ B activation and the transcription of cytokines and inflammatory consequences of cell death, c-IAPs have been found to be involved in pathogen sensor pathways mediated by Toll-like receptor (TLR) 4 and retinoic acid-inducible gene I (RIG-I) (Vandenabeele and Bertrand, 2012).

TLR4 is a membrane bound protein that senses the bacterial lipopolysaccharide (LPS). Once it gets activated, TLR4 forms a complex with myeloid differentiation primary response 88 (MyD88) and TIR-domain-containing adapter-inducing interferon- β (TRIF), and recruits TRAF3 and TRAF6. c-IAPs regulate this pathway by ubiquitinating TRAF3 and targeting it for proteasomal degradation (Tseng *et al.*, 2010). This allows MyD88 to leave the membrane and go to the cytosol to activate the MAPK pathway that would lead to cytokine expression.

RIG-I is an intracellular sensor of dsRNA that is involved in viral defense (Yoneyama *et al.*, 2004). c-IAPs can be recruited to the RIG-I complex together with TRAF3, and ubiquitination of TRAF3 by c-IAPs participates in the activation of the interferon I response (Mao *et al.*, 2010).

ML-IAP (Livin or BIRC7)

Melanoma-IAP, with its single BIR domain can bind SMAC and with its RING domain it can promote SMAC ubiquitination and proteasomal degradation (Vucic *et al.*, 2005). ML-IAP is a very weak inhibitor of caspases but it binds SMAC with high affinity. Thus, it has been postulated that its role in apoptosis is not due to the direct inhibition of caspases, but to its competition with XIAP for the binding of SMAC (Vucic *et al.*, 2005, 2002). By acting as a SMAC sink, ML-IAP can free XIAP from SMAC inhibition and facilitate the inhibition of caspases (Figure 2). Apart from its role in inhibiting apoptosis through caspases and SMAC, ML-IAP (similarly to XIAP and c-IAP1) has been implicated in the migration of melanoma cells by binding to c-RAF and activation of the signaling pathway of MAPK (Oberoi-Khanuja *et al.*, 2012). Also, ML-IAP has a role

in cell proliferation inhibition, since its absence can induce cell cycle arrest (Wang *et al.*, 2007).

As its name indicates, ML-IAP is overexpressed in melanomas, but also in other tumors (Vucic *et al.*, 2000). Microphthalmia-Associated Transcription Factor (MITF), a protein responsible for the lineage differentiation of melanocytes, regulates ML-IAP expression in melanomas (Dynek *et al.*, 2008). Interestingly, physiological expression of ML-IAP seems to be predominantly in the eyes, although the functional relevance of eye expression has not been fully understood yet (Varfolomeev *et al.*, 2012a,b).

NAIP (BIRC1)

Neuronal apoptosis inhibitory protein was the first human IAP discovered and it was suspected that NAIP was responsible for the severe form of spinal muscular atrophy (SMA) (Roy *et al.*, 1995). Later, it was determined that the neighboring gene, survival motor neuron (SMN), was the culprit (Campbell *et al.*, 1997; Sertic *et al.*, 1997). Nevertheless, NAIP was shown to block cell death induced by a variety of stimuli (Liston *et al.*, 1996).

The presence of the NBD and LRR domains makes NAIP not only a member of the IAP family, but also of the NLR (nucleotide-binding domain [NBD]- and leucine-rich repeat [LRR]-containing) family of proteins. These proteins are involved in the immune response as intracellular sensors of pathogens.

In mice, NAIP is represented by different paralogs that are responsible for detecting pathogen components such as flagellin, a component of the flagella of bacteria (Tenthorey *et al.*, 2014; Kofoed and Vance, 2011). After the detection of the pathogen, several NAIPs can bind the NLR family CARD-containing 4 (NLRC4) protein and oligomerize it. NLRC4 then uses its CARD domain to activate caspase-1 (von Moltke *et al.*, 2013). Caspase-1 cleaves and activates the inflammatory cytokines IL-1 β and IL-18, and initiates pyroptotic cell death (von Moltke *et al.*, 2012). This form of cell death is one of the main mechanism by which infected cells are cleared (Lightfield *et al.*, 2008). It is still not clear how this pathway operates in humans, but so far it has been observed that overexpression of NAIP makes cells less permissive to pathogen growth (Vinzing *et al.*, 2008).

Survivin (BIRC5)

Survivin is the smallest of the IAP family members with a strong expression pattern in human cancers. Survivin is mainly expressed in tumors and not in normal tissue (with the exception of embryonic development, stem cells, thymus, testes, regenerating hepatocytes, and endothelial cells) representing a very promising target for drug development (Fukuda and Pelus, 2006; Chiou *et al.*, 2003). The expression of survivin correlates with tumor resistance to chemotherapy and radiotherapy, so drugs targeting survivin could have a great impact in cancer treatment.

Survivin has been implicated in numerous cellular pathways: mitosis, apoptosis, autophagy, DNA damage, and its cellular localization plays an important role for its many

functions (Wang *et al.*, 2011; Altieri, 2003). However, the physiological role of Survivin is in regulation of cell cycle. Survivin binds microtubules to contribute to the chromatin-associated spindle formation, and together with other proteins it forms the chromosomal passenger complex (CPC) (Altieri, 2008). Furthermore, the levels of survivin change during the cell cycle: they peak in M-phase but are very low in cells under G1-arrest (Li *et al.*, 1998).

Survivin has different splicing variants; from one mRNA five proteins are produced: survivin, survivin-2 α , survivin-2B (with duplicated 2nd exon), survivin- Δ -Ex-3 (with missing 3rd exon), and survivin-3B. These proteins have similar and distinctive functions, for example survivin-2B is an apoptotic inhibitor, while survivin- Δ -Ex-3 is not (Mahotka *et al.*, 2002). Survivin forms homodimers, so some of these splice variants could compete with survivin to form heterodimers with different functions.

Apollon or Bruce (BIRC6)

Bruce (BIR repeat containing ubiquitin-conjugating enzyme) is the biggest and the only member of the IAP family that is found on the membranes (Hauser *et al.*, 1998). Apart from the BIR domain, Bruce contains a UBC domain that allows ubiquitin conjugation. The E2 activity is very important for Bruce-mediated inhibition of apoptosis. Besides its auto-ubiquitination and proteosomal degradation, Bruce can promote ubiquitination of important apoptotic proteins such as SMAC/Diablo, caspase-9 and HtrA2 (Bartke *et al.*, 2004; Hao *et al.*, 2004; Sekine *et al.*, 2005).

Bruce is critical for cell survival, at least in embryonic development, since knockout mice die during embryogenesis or early after the birth. But the mechanism of action by which Bruce regulates cell death is not clear. Some studies show that Bruce knockout embryos do not have extensive apoptosis (Lotz *et al.*, 2004). Nevertheless, other studies indicate that the absence of Bruce induces apoptosis in the placenta because it stabilizes p53, which translocates to the nucleus and modulates the expression of genes involved in cell death leading to the activation of caspases (Lopergolo *et al.*, 2009; Ren *et al.*, 2005).

The importance of Bruce as an antiapoptotic protein is also supported by its role in cell cycle. In order to enter into mitosis, the proteins Cyclin A and B need to be degraded, mostly due to the E3 ligase APC/C (anaphase-promoting complex/cyclosome). However, Bruce also contributes to the ubiquitination and degradation of Cyclin A, allowing the cell to continue through the cell cycle and undergo mitosis (Kikuchi *et al.*, 2014).

Unlike other IAPs, the role of Bruce in cancer is not clear – it has been found overexpressed in some cancers, and it may confer resistance to some DNA damage chemotherapeutics, but more studies are needed to resolve this question (Sung *et al.*, 2007).

ILP-2 (BIRC8)

The expression of inhibitor of apoptosis protein-like 2 is limited to testis and lymphoblastoid cells. The amino acid sequence of ILP-2 is very similar to the C-terminal region of XIAP, making it likely that ILP-2 originated through a partial

duplication of XIAP (Lagace *et al.*, 2001; Richter *et al.*, 2001). ILP-2 can bind and inhibit caspase-9 but very weakly, probably because ILP-2 is a very unstable protein (Shin *et al.*, 2005). In addition, ILP-2 is not capable of blocking cell death triggered by Fas or TNF α .

Targeting IAP Proteins

IAP proteins are important modulators of cell death and thus, they are attractive therapeutic targets for diseases with dysregulated cell death such as cancer (Fulda and Vucic, 2012).

IAP Antagonists

The crucial role of c-IAPs and XIAP in inhibiting apoptosis has led to the idea that their antagonism could have a therapeutic benefit. These IAP proteins are often overexpressed in cancers, so their inhibition could be used to trigger apoptosis in tumors (Tamm *et al.*, 2000).

SMAC mimetics – IAP antagonists are small molecules that have been developed to mimic the way endogenous SMAC/Diablo binds IAP proteins. SMAC, as a dimer, is an endogenous IAP antagonist that is released from the mitochondria intermembrane space into the cytosol after apoptotic stimuli. The N-terminal tetrapeptide AVPI (Ala-Val-Pro-Ile) is responsible for binding the peptide-binding groove called IBM on the surface of select BIR domains of IAP proteins (Wu *et al.*, 2000). SMAC binds with high affinity to the BIR3 of XIAP, c-IAP1 and c-IAP2, the single BIR domain of ML-IAP, and it can bind the BIR2 of XIAP and the other IAPs with a lower affinity (Wu *et al.*, 2000). The binding of IAP antagonists to the select BIR domains of XIAP and c-IAP proteins can lead to cell death, but through different mechanisms.

In basal conditions, c-IAPs are folded in a way that occludes the RING domain. Binding of IAP antagonists to c-IAPs causes a conformational change that triggers the opening of c-IAP1 structure and exposes the RING domain (Dueber *et al.*, 2011). In this open conformation, the RING domain is free to dimerize with another RING domain. When this occurs, the E3 ubiquitin ligase activity is activated and c-IAPs get autoubiquitinated very rapidly, targeting themselves for proteasomal degradation (Dueber *et al.*, 2011; Feltham *et al.*, 2011). When c-IAPs are no longer present, RIP1, a crucial mediator of the pro-survival NF- κ B signaling pathway, cannot get ubiquitinated to recruit LUBAC, NEMO, TAK and TAB (Varfolomeev *et al.*, 2008; Bertrand *et al.*, 2008). Non-ubiquitinated RIP1 leaves the TNFRI-associated signaling complex and binds to caspases-8/10 and FADD forming the complex II (Bertrand *et al.*, 2008; Varfolomeev *et al.*, 2007; Vince *et al.*, 2007). Once complex II is formed, caspases-8/10 get activated and processed initiating the caspases cascade that leads to apoptotic cell death.

XIAP mechanistically inhibits apoptosis differently from c-IAPs, as XIAP directly inhibits several caspases. The binding of IAP antagonists to XIAP blocks the binding site to caspases, thus preventing caspase inhibition and resulting in cell death (Gao *et al.*, 2007; Varfolomeev *et al.*, 2009). Some IAP antagonist can promote slow XIAP ubiquitination and degradation, but that is probably not the main mechanism of action of IAP antagonists in XIAP neutralization.

One of the problems facing the therapeutic development of IAP antagonists is the relative insensitivity of numerous cancer cells to their action. IAP antagonists rely on the synthesis and secretion of TNF α to activate TNFRI, which, in the absence of c-IAPs, leads to the migration of RIP1 to caspase-8 complex. IAP antagonists usually stimulate TNF α production by activating NF- κ B signaling pathways (Petersen *et al.*, 2007). However, some cells do not activate NF- κ B signaling sufficiently to upregulate TNF α and activate TNFRI, thus not allowing IAP antagonists as single agents to induce cell death in many cancer cell types. Nevertheless, combination of IAP antagonists with TNF α , produced by neighboring cells or tissues can overcome this hurdle. Currently, IAP antagonists are tested in clinical trials to investigate their therapeutic benefit in human cancers (Fulda and Vucic, 2012; Flygare and Fairbrother, 2010).

Survivin Antisense Oligonucleotides

Survivin is an attractive therapeutic target since it is expressed mainly in cancer cells and not in normal tissues. One of the approaches that has been explored for targeting survivin utilizes gene therapy. An antisense oligonucleotide, LY2181308, is presently examined in cancer clinical trials with no critical toxicity observed so far (Tanioka *et al.*, 2011; Erba *et al.*, 2013). LY2181308 is a second-generation 2'-O-methoxyethyl, phosphorothioate antisense oligonucleotide. This antisense oligonucleotide binds to Survivin mRNA and suppresses protein translation, thus causing cell death.

Concluding Remarks

The careful regulation of cell death and survival is essential for proper maintenance of organismal homeostasis, and deregulated cell death can lead to neurological defects, infections and cancer. IAP proteins are evolutionarily conserved pro-survival mediators that regulate cell death and various signaling pathways (e.g., NF- κ B and NOD) to control cellular fate. Potent anti-cell death activity, elevated expression in many cancer types, and contribution to survival signaling pathways make IAP proteins promising targets for therapeutic intervention in human malignancies. However, our understanding of IAP proteins and cellular processes they regulate is far from complete and future studies, both in basic research and in clinical setting, should focus on fully unraveling the biological roles of IAP proteins and thereby open up additional therapeutic strategies.

See also: Cell Division/Death: Apoptosis: Apoptosis; Caspases; The Bcl-2 Family Proteins: Insights into Their Mechanism of Action and Therapeutic Potential

References

- Altieri, D.C., 2003. Survivin, versatile modulation of cell division and apoptosis in cancer. *Oncogene* 22 (53), 8581–8589.
- Altieri, D.C., 2008. Survivin, cancer networks and pathway-directed drug discovery. *Nature Reviews Cancer* 8 (1), 61–70.
- Bartke, T., Pohl, C., Pyrowolakis, G., Jentsch, S., 2004. Dual role of BRUCE as an antiapoptotic IAP and a chimeric E2/E3 ubiquitin ligase. *Molecular Cell* 14 (6), 801–811.
- Bertrand, M.J., Doiron, K., Labbé, K., *et al.*, 2009. Cellular inhibitors of apoptosis cIAP1 and cIAP2 are required for innate immunity signaling by the pattern recognition receptors NOD1 and NOD2. *Immunity* 30 (6), 789–801.
- Bertrand, M.J., Milutinovic, S., Dickson, K.M., *et al.*, 2008. cIAP1 and cIAP2 facilitate cancer cell survival by functioning as E3 ligases that promote RIP1 ubiquitination. *Molecular Cell* 30 (6), 689–700.
- Birnbaum, M.J., Clem, R.J., Miller, L.K., 1994. An apoptosis-inhibiting gene from a nuclear polyhedrosis virus encoding a polypeptide with Cys/His sequence motifs. *Journal of Virology* 68 (4), 2521–2528.
- Blankenship, J.W., Varfolomeev, E., Goncharov, T., *et al.*, 2009. Ubiquitin binding modulates IAP antagonist-stimulated proteasomal degradation of c-IAP1 and c-IAP2(1). *Biochemical Journal* 417 (1), 149–160.
- Cai, Z., Jitkaew, S., Zhao, J., *et al.*, 2014. Plasma membrane translocation of trimerized MLKL protein is required for TNF-induced necroptosis. *Nature Cell Biology* 16 (1), 55–65.
- Campbell, L., Potter, A., Ignatius, J., Dubowitz, V., Davies, K., 1997. Genomic variation and gene conversion in spinal muscular atrophy: implications for disease process and clinical phenotype. *American Journal of Human Genetics* 61 (1), 40–50.
- Chai, J., Shiozaki, E., Srinivasula, S.M., *et al.*, 2001. Structural basis of caspase-7 inhibition by XIAP. *Cell* 104 (5), 769–780.
- Chiou, S.K., Jones, M.K., Tarnawski, A.S., 2003. Survivin – An anti-apoptosis protein: Its biological roles and implications for cancer and beyond. *Medical Science Monitor* 9 (4), P125–P129.
- Clem, R.J., Fechheimer, M., Miller, L.K., 1991. Prevention of apoptosis by a baculovirus gene during infection of insect cells. *Science* 254 (5036), 1388–1390.
- Conze, D.B., Zhao, Y., Ashwell, J.D., 2010. Non-canonical NF- κ B activation and abnormal B cell accumulation in mice expressing ubiquitin protein ligase-inactive c-IAP2. *PLoS Biology* 8 (10), e1000518.
- Crook, N.E., Clem, R.J., Miller, L.K., 1993. An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif. *Journal of Virology* 67 (4), 2168–2174.
- Damgaard, R.B., Fiil, B.K., Speckmann, C., *et al.*, 2013. Disease-causing mutations in the XIAP BIR2 domain impair NOD2-dependent immune signaling. *EMBO Molecular Medicine* 5 (8), 1278–1295.
- Damgaard, R.B., Nachbur, U., Yabal, M., *et al.*, 2012. The ubiquitin ligase XIAP recruits LUBAC for NOD2 signaling in inflammation and innate immunity. *Molecular Cell* 46 (6), 746–758.
- Dejardin, E., 2006. The alternative NF- κ B pathway from biochemistry to biology: pitfalls and promises for future drug development. *Biochemical Pharmacology* 72 (9), 1161–1179.
- Duckett, C.S., Nava, V.E., Gedrich, R.W., *et al.*, 1996. A conserved family of cellular genes related to the baculovirus iap gene and encoding apoptosis inhibitors. *EMBO Journal* 15 (11), 2685–2694.
- Dueber, E.C., Schoeffler, A.J., Lingel, A., *et al.*, 2011. Antagonists induce a conformational change in cIAP1 that promotes autoubiquitination. *Science* 334 (6054), 376–380.
- Dynek, J.N., Chan, S.M., Liu, J., *et al.*, 2008. Microphthalmia-associated transcription factor is a critical transcriptional regulator of melanoma inhibitor of apoptosis in melanomas. *Cancer Research* 68 (9), 3124–3132.
- Dynek, J.N., Goncharov, T., Dueber, E.C., *et al.*, 2010. c-IAP1 and UbcH5 promote K11-linked polyubiquitination of RIP1 in TNF signaling. *EMBO Journal* 29 (24), 4198–4209.
- Eckelman, B.P., Salvesen, G.S., 2006. The human anti-apoptotic proteins cIAP1 and cIAP2 bind but do not inhibit caspases. *Journal of Biological Chemistry* 281 (6), 3254–3260.
- Eckelman, B.P., Salvesen, G.S., Scott, F.L., 2006. Human inhibitor of apoptosis proteins: Why XIAP is the black sheep of the family. *EMBO Reports* 7 (10), 988–994.
- Erba, H.P., Sayar, H., Juckett, M., *et al.*, 2013. Safety and pharmacokinetics of the antisense oligonucleotide (ASO) LY2181308 as a single-agent or in combination with idarubicin and cytarabine in patients with refractory or relapsed acute myeloid leukemia (AML). *Investigational New Drugs* 31 (4), 1023–1034.
- Feltham, R., Bettjerman, B., Budhidarmo, R., *et al.*, 2011. Smac mimetics activate the E3 ligase activity of cIAP1 protein by promoting RING domain dimerization. *Journal of Biological Chemistry* 286 (19), 17015–17028.
- Feoktistova, M., Geserick, P., Kellert, B., *et al.*, 2011. cIAPs block Ripoptosome formation, a RIP1/caspase-8 containing intracellular cell death

- complex differentially regulated by cIAP isoforms. *Molecular Cell* 43 (3), 449–463.
- Flygare, J.A., Fairbrother, W.J., 2010. Small-molecule pan-IAP antagonists: a patent review. *Expert Opinion on Therapeutic Patents* 20 (2), 251–267.
- Fukuda, S., Pelus, L.M., 2006. Survivin, a cancer target with an emerging role in normal adult tissues. *Molecular Cancer Therapeutics* 5 (5), 1087–1098.
- Fulda, S., 2007. Inhibitor of apoptosis proteins as targets for anticancer therapy. *Expert Review of Anticancer Therapy* 7 (9), 1255–1264.
- Fulda, S., Vucic, D., 2012. Targeting IAP proteins for therapeutic intervention in cancer. *Nature Reviews Drug Discovery* 11 (2), 109–124.
- Gao, Z., Tian, Y., Wang, J., *et al.*, 2007. A dimeric Smac/diablo peptide directly relieves caspase-3 inhibition by XIAP. Dynamic and cooperative regulation of XIAP by Smac/Diablo. *Journal of Biological Chemistry* 282 (42), 30718–30727.
- Giardino Torchia, M.L., Conze, D.B., Ashwell, J.D., 2013. c-IAP1 and c-IAP2 redundancy differs between T and B cells. *PLoS One* 8 (6), e66161.
- Gyrd-Hansen, M., Darding, M., Miasari, M., *et al.*, 2008. IAPs contain an evolutionarily conserved ubiquitin-binding domain that regulates NF- κ B as well as cell survival and oncogenesis. *Nature Cell Biology* 10 (11), 1309–1317.
- Hao, Y., Sekine, K., Kawabata, A., *et al.*, 2004. Apollon ubiquitinates SMAC and caspase-9, and has an essential cytoprotection function. *Nature Cell Biology* 6 (9), 849–860.
- Hasegawa, M., Fujimoto, Y., Lucas, P.C., *et al.*, 2008. A critical role of RICK/RIP2 polyubiquitination in Nod-induced NF- κ B activation. *EMBO Journal* 27 (2), 373–383.
- Hauser, H.P., Bardroff, M., Pyrowolakis, G., Jentsch, S., 1998. A giant ubiquitin-conjugating enzyme related to IAP apoptosis inhibitors. *Journal of Cell Biology* 141 (6), 1415–1422.
- Hershko, A., Ciechanover, A., 1998. The ubiquitin system. *Annual Review of Biochemistry* 67, 425–479.
- Hofmann, K., Bucher, P., Tschopp, J., 1997. The CARD domain: A new apoptotic signaling motif. *Trends in Biochemical Sciences* 22 (5), 155–156.
- Kikuchi, R., Ohata, H., Ohoka, N., Kawabata, A., Naito, M., 2014. APOLLON protein promotes early mitotic CYCLIN A degradation independent of the spindle assembly checkpoint. *Journal of Biological Chemistry* 289 (6), 3457–3467.
- Kofoed, E.M., Vance, R.E., 2011. Innate immune recognition of bacterial ligands by NALPs determines inflammasome specificity. *Nature* 477 (7366), 592–595.
- Koonin, E.V., Aravind, L., 2000. The NACHT family — A new group of predicted NTPases implicated in apoptosis and MHC transcription activation. *Trends in Biochemical Sciences* 25 (5), 223–224.
- Krieg, A., Correa, R.G., Garrison, J.B., *et al.*, 2009. XIAP mediates NOD signaling via interaction with RIP2. *Proceedings of the National Academy of Sciences of the United States of America* 106 (34), 14524–14529.
- Lagace, M., Xuan, J.Y., Young, S.S., *et al.*, 2001. Genomic organization of the X-linked inhibitor of apoptosis and identification of a novel testis-specific transcript. *Genomics* 77 (3), 181–188.
- Li, F., Ambrosini, G., Chu, E.Y., *et al.*, 1998. Control of apoptosis and mitotic spindle checkpoint by survivin. *Nature* 396 (6711), 580–584.
- Lightfield, K.L., Persson, J., Brubaker, S.W., *et al.*, 2008. Critical function for Naip5 in inflammasome activation by a conserved carboxy-terminal domain of flagellin. *Nature Immunology* 9 (10), 1171–1178.
- Liston, P., Roy, N., Tamai, K., *et al.*, 1996. Suppression of apoptosis in mammalian cells by NAIP and a related family of IAP genes. *Nature* 379 (6563), 349–353.
- Lopergolo, A., Pennati, M., Gandellini, P., *et al.*, 2009. Apollon gene silencing induces apoptosis in breast cancer cells through p53 stabilisation and caspase-3 activation. *British Journal of Cancer* 100 (5), 739–746.
- Lotz, K., Pyrowolakis, G., Jentsch, S., 2004. BRUCE, a giant E2/E3 ubiquitin ligase and inhibitor of apoptosis protein of the trans-Golgi network, is required for normal placenta development and mouse survival. *Molecular and Cellular Biology* 24 (21), 9339–9350.
- Mahotka, C., Liebmann, J., Wenzel, M., *et al.*, 2002. Differential subcellular localization of functionally divergent survivin splice variants. *Cell Death & Differentiation* 9 (12), 1334–1342.
- Mao, A.P., Li, S., Zhong, B., *et al.*, 2010. Virus-triggered ubiquitination of TRAF3/6 by cIAP1/2 is essential for induction of interferon- β (IFN- β) and cellular antiviral response. *Journal of Biological Chemistry* 285 (13), 9470–9476.
- von Moltke, J., Ayres, J.S., Kofoed, E.M., Chavarría-Smith, J., Vance, R.E., 2013. Recognition of bacteria by inflammasomes. *Annual Review of Immunology* 31, 73–106.
- von Moltke, J., Trinidad, N.J., Moayeri, M., *et al.*, 2012. Rapid induction of inflammatory lipid mediators by the inflammasome in vivo. *Nature* 490 (7418), 107–111.
- Oberoi-Khanuja, T.K., Karreman, C., Larisch, S., Rapp, U.R., Rajalingam, K., 2012. Role of melanoma inhibitor of apoptosis (ML-IAP) protein, a member of the baculoviral IAP repeat (BIR) domain family, in the regulation of C-RAF kinase and cell migration. *Journal of Biological Chemistry* 287 (34), 28445–28455.
- Petersen, S.L., Wang, L., Yalcin-Chin, A., *et al.*, 2007. Autocrine TNF α signaling renders human cancer cells susceptible to Smac-mimetic-induced apoptosis. *Cancer Cell* 12 (5), 445–456.
- Pickart, C.M., Fushman, D., 2004. Polyubiquitin chains: Polymeric protein signals. *Current Opinion in Chemical Biology* 8 (6), 610–616.
- Ren, J., Shi, M., Liu, R., *et al.*, 2005. The Birc6 (Bruce) gene regulates p53 and the mitochondrial pathway of apoptosis and is essential for mouse embryonic development. *Proceedings of the National Academy of Sciences of the United States of America* 102 (3), 565–570.
- Richter, B.W., Mir, S.S., Eiben, L.J., *et al.*, 2001. Molecular cloning of ILP-2, a novel member of the inhibitor of apoptosis protein family. *Molecular and Cellular Biology* 21 (13), 4292–4301.
- Riedl, S.J., Renatus, M., Schwarzenbacher, R., *et al.*, 2001. Structural basis for the inhibition of caspase-3 by XIAP. *Cell* 104 (5), 791–800.
- Roy, N., Mahadevan, M.S., McLean, M., *et al.*, 1995. The gene for neuronal apoptosis inhibitory protein is partially deleted in individuals with spinal muscular atrophy. *Cell* 80 (1), 167–178.
- Salvesen, G.S., Duckett, C.S., 2002. IAP proteins: Blocking the road to death's door. *Current Opinion in Chemical Biology* 3 (6), 401–410.
- Samuel, T., Welsh, K., Lober, T., *et al.*, 2006. Distinct BIR domains of cIAP1 mediate binding to and ubiquitination of tumor necrosis factor receptor-associated factor 2 and second mitochondrial activator of caspases. *Journal of Biological Chemistry* 281 (2), 1080–1090.
- Scheidereit, C., 2006. I κ B kinase complexes: Gateways to NF- κ B activation and transcription. *Oncogene* 25 (51), 6685–6705.
- Scott, F.L., Denault, J.B., Riedl, S.J., *et al.*, 2005. XIAP inhibits caspase-3 and -7 using two binding sites: Evolutionarily conserved mechanism of IAPs. *EMBO Journal* 24 (3), 645–655.
- Sekine, K., Hao, Y., Suzuki, Y., *et al.*, 2005. HtrA2 cleaves Apollon and induces cell death by IAP-binding motif in Apollon-deficient cells. *Biochemical and Biophysical Research Communications* 330 (1), 279–285.
- Sertic, J., Barisic, N., Sostarko, M., *et al.*, 1997. Deletions in the SMN and NAIP genes in patients with spinal muscular atrophy in Croatia. *Collegium Antropologicum* 21 (2), 487–492.
- Shim, J.H., Xiao, C., Paschal, A.E., *et al.*, 2005. TAK1, but not Table1 or Table2, plays an essential role in multiple signaling pathways in vivo. *Genes & Development* 19 (22), 2668–2681.
- Shin, H., Renatus, M., Eckelman, B.P., *et al.*, 2005. The BIR domain of IAP-like protein 2 is conformationally unstable: Implications for caspase inhibition. *Biochemical Journal* 385 (Pt 1), 1–10.
- Silke, J., Brink, R., 2010. Regulation of TNFRSF and innate immune signaling complexes by TRAFs and cIAPs. *Cell Death & Differentiation* 17 (1), 35–45.
- Silke, J., Meier, P., 2013. Inhibitor of apoptosis (IAP) proteins—modulators of cell death and inflammation. *Cold Spring Harbor Perspectives in Biology* 5, 2.
- Srinivasula, S.M., Ashwell, J.D., 2008. IAPs: What's in a name? *Molecular Cell* 30 (2), 123–135.
- Sun, L., Wang, H., Wang, Z., *et al.*, 2012. Mixed lineage kinase domain-like protein mediates necrosis signaling downstream of RIP3 kinase. *Cell* 148 (1–2), 213–227.
- Sung, K.W., Choi, J., Hwang, Y.K., *et al.*, 2007. Overexpression of Apollon, an antiapoptotic protein, is associated with poor prognosis in childhood de novo acute myeloid leukemia. *Clinical Cancer Research* 13 (17), 5109–5114.
- Tamm, I., Kornblau, S.M., Segall, H., *et al.*, 2000. Expression and prognostic significance of IAP-family genes in human cancers and myeloid leukemias. *Clinical Cancer Research* 6 (5), 1796–1803.
- Tanioka, M., Nohara, H., Yamamoto, N., *et al.*, 2011. Phase I study of LY2181308, an antisense oligonucleotide against survivin, in patients with advanced solid tumors. *Cancer Chemotherapy and Pharmacology* 68 (2), 505–511.
- Tenthorey, J.L., Kofoed, E.M., Daugherty, M.D., Malik, H.S., Vance, R.E., 2014. Molecular Basis for Specific Recognition of Bacterial Ligands by NAIP/NLRC4 Inflammasomes. *Molecular Cell* 54 (1), 17–29.
- Tseng, P.H., Matsuzawa, A., Zhang, W., *et al.*, 2010. Different modes of ubiquitination of the adaptor TRAF3 selectively activate the expression of type I interferons and proinflammatory cytokines. *Nature Immunology* 11 (1), 70–75.
- Vallabhupuram, S., Matsuzawa, A., Zhang, W., *et al.*, 2008. Nonredundant and complementary functions of TRAF2 and TRAF3 in a ubiquitination cascade that activates NIK-dependent alternative NF- κ B signaling. *Nature Immunology* 9 (12), 1364–1370.
- Vandenabeele, P., Bertrand, M.J., 2012. The role of the IAP E3 ubiquitin ligases in regulating pattern-recognition receptor signaling. *Nature Reviews Immunology* 12 (12), 833–844.

- Varfolomeev, E., Aliche, B., Elliott, J.M., *et al.*, 2009. X chromosome-linked inhibitor of apoptosis regulates cell death induction by proapoptotic receptor agonists. *Journal of Biological Chemistry* 284 (50), 34553–34560.
- Varfolomeev, E., Blankenship, J.W., Wayson, S.M., *et al.*, 2007. IAP antagonists induce autoubiquitination of c-IAPs, NF-kappaB activation, and TNFalpha-dependent apoptosis. *Cell* 131 (4), 669–681.
- Varfolomeev, E., Goncharov, T., Fedorova, A.V., *et al.*, 2008. c-IAP1 and c-IAP2 are critical mediators of tumor necrosis factor alpha (TNFalpha)-induced NF-kappaB activation. *Journal of Biological Chemistry* 283 (36), 24295–24299.
- Varfolomeev, E., Goncharov, T., Maecker, H., *et al.*, 2012a. Cellular inhibitors of apoptosis are global regulators of NF-kappaB and MAPK activation by members of the TNF family of receptors. *Science Signaling* 5 (216), ra22.
- Varfolomeev, E., Moradi, E., Dynek, J.N., *et al.*, 2012b. Characterization of ML-IAP protein stability and physiological role in vivo. *Biochemical Journal* 447 (3), 427–436.
- Varfolomeev, E., Vucic, D., 2008. (Un)expected roles of c-IAPs in apoptotic and NF- κ B signaling pathways. *Cell Cycle* 7 (11), 1511–1521.
- Varfolomeev, E., Vucic, D., 2011. Inhibitor of apoptosis proteins: Fascinating biology leads to attractive tumor therapeutic targets. *Future Oncology* 7 (5), 633–648.
- Varfolomeev, E., Wayson, S.M., Dixit, V.M., Fairbrother, W.J., Vucic, D., 2006. The inhibitor of apoptosis protein fusion c-IAP2.MALT1 stimulates NF-kappaB activation independently of TRAF1 AND TRAF2. *Journal of Biological Chemistry* 281 (39), 29022–29029.
- Vaux, D.L., Silke, J., 2005. IAPs, RINGs and ubiquitylation. *Current Opinion in Chemical Biology* 6 (4), 287–297.
- Vince, J.E., Wong, W.W., Khan, N., *et al.*, 2007. IAP antagonists target cIAP1 to induce TNFalpha-dependent apoptosis. *Cell* 131 (4), 682–693.
- Vinzing, M., Eitel, J., Lippmann, J., *et al.*, 2008. NAIP and Ipaf control Legionella pneumophila replication in human cells. *Journal of Immunology* 180 (10), 6808–6815.
- Vucic, D., Deshayes, K., Ackerly, H., *et al.*, 2002. SMAC negatively regulates the anti-apoptotic activity of melanoma inhibitor of apoptosis (ML-IAP). *Journal of Biological Chemistry* 277 (14), 12275–12279.
- Vucic, D., Dixit, V.M., Wertz, I.E., 2011. Ubiquitylation in apoptosis: A post-translational modification at the edge of life and death. *Current Opinion in Chemical Biology* 12 (7), 439–452.
- Vucic, D., Franklin, M.C., Wallweber, H.J., *et al.*, 2005. Engineering ML-IAP to produce an extraordinarily potent caspase 9 inhibitor: implications for Smac-dependent anti-apoptotic activity of ML-IAP. *Biochemical Journal* 385 (Pt 1), 11–20.
- Vucic, D., Stennicke, H.R., Pisabarro, M.T., Salvesen, G.S., Dixit, V.M., 2000. ML-IAP, a novel inhibitor of apoptosis that is preferentially expressed in human melanomas. *Current Biology – Cell* 10 (21), 1359–1366.
- Wang, C., Deng, L., Hong, M., *et al.*, 2001. TAK1 is a ubiquitin-dependent kinase of MKK and IKK. *Nature* 412 (6844), 346–351.
- Wang, C.Y., Mayo, M.W., Korneluk, R.G., Goeddel, D.V., Baldwin Jr., A.S., 1998. NF- κ B antiapoptosis: Induction of TRAF1 and TRAF2 and c-IAP1 and c-IAP2 to suppress caspase-8 activation. *Science* 281 (5383), 1680–1683.
- Wang, H., Tan, S.S., Wang, X.Y., *et al.*, 2007. Silencing livin gene by siRNA leads to apoptosis induction, cell cycle arrest, and proliferation inhibition in malignant melanoma LiBr cells. *Acta Pharmacologica Sinica* 28 (12), 1968–1974.
- Wang, Q., Chen, Z., Diao, X., Huang, S., 2011. Induction of autophagy-dependent apoptosis by the survivin suppressant YM155 in prostate cancer cells. *Cancer Letters* 302 (1), 29–36.
- Wu, G., Chai, J., Suber, T.L., *et al.*, 2000. Structural basis of IAP recognition by Smac/DIABLO. *Nature* 408 (6815), 1008–1012.
- Yoneyama, M., Kikuchi, M., Natsukawa, T., *et al.*, 2004. The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nature Immunology* 5 (7), 730–737.
- Zhao, J., Jitkaew, S., Cai, Z., *et al.*, 2012. Mixed lineage kinase domain-like is a key receptor interacting protein 3 downstream component of TNF-induced necrosis. *Proceedings of the National Academy of Sciences of the United States of America* 109 (14), 5322–5327.
- Zong, W.X., Thompson, C.B., 2006. Necrotic death as a cell fate. *Genes & Development* 20 (1), 1–15.

Mitotic Catastrophe

RYC Poon, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

© 2016 Elsevier Inc. All rights reserved.

Introduction

Proper control of cell number is dictated by a delicate balance between cell division and cell death. Dysregulation of this equilibrium underlies many human disorders including cancer. Progress in the past decade has unraveled some of the underlying principles of a specific form of cell death termed mitotic catastrophe. Although its biological significance and mechanism remain to be fully defined, the prevailing view is that mitotic catastrophe plays an essential role in preventing genome instability.

Mitosis is associated with profound changes in the cell's morphology and biochemistry. Although occupying just a fraction of the time of the whole cell cycle, mitosis is a period particularly prone to irreversible errors. Elaborated mechanisms have developed to ensure that chromosomes are segregated timely and accurately. Defects of these checks and balances are believed to be the major driving forces of aneuploidy. Accordingly, the role of mitotic catastrophe in eliminating cells that are unable to execute mitosis properly is crucial for maintaining genome stability.

Another salient aspect of mitotic catastrophe is that many current and emerging anticancer therapies are based on targeting mitosis. These range from classic spindle poisons to inhibitors that target mitotic regulators such as Eg5, Aurora kinases, and polo-like kinases. Other anticancer approaches such as abrogation of the G₂ DNA damage checkpoint also relies on forcing cells into mitosis. Thus understanding mitotic catastrophe may uncover principles of tumorigenesis as well as insights for novel therapeutic designs.

Mitotic Catastrophe

Historically, the term 'mitotic catastrophe' was first used (Russell and Nurse, 1987) to describe the lethal phenotype associated with the fission yeast strain that contain an activated allele of *cdc2* (*cdc2-3w*) and a recessive temperature-sensitive *wee1* mutation (*wee1-50*) (Russell and Nurse, 1986). These mutations overactivate the mitotic engine, resulting in premature and aberrant mitosis, particularly with respect to chromosome segregation and septum formation (Molz *et al.*, 1989).

Mitotic catastrophe has since been used for cell death associated with inappropriate entry into mitosis. Because the term 'mitotic catastrophe' has been adopted rather loosely in the literature, there are disagreements about the precise causes, mechanisms, and outcomes of the process. Some studies adopted a relatively simple definition to mean a mode of cell death associated with aberrant mitotic activity (Vakifahmetoglu *et al.*, 2008). More recent attempts to standardize different types of cell death defined mitotic catastrophe to be associated with aberrant mitotic activity that ultimately triggers cell death or irreversible cell cycle arrest (Galluzzi *et al.*, 2012). In this

definition, cell death can occur either during or after the defective mitosis.

One of the better characterized events that lead to mitotic catastrophe is when cells are challenged with antimitotic drugs. Under these conditions, cells typically undergo prolonged mitotic arrest. Whether this and other abnormal mitotic events represent mitotic catastrophe depends on the eventual fate of the cells (Figure 1). On the one hand, some cells are able to divide and continue to propagate (many cancer cells, in particular, can tolerate certain number of mis-segregated chromosomes). These types of impaired mitosis are not generally regarded as mitotic catastrophe. On the other hand, cells that undergo cell death during the aberrant mitosis are generally regarded as undergoing mitotic catastrophe. Mitotic catastrophe is also applicable when cells lose growth capacity after exiting the aberrant mitosis. For example, after a prolonged block in mitosis, some cells can exit mitosis without chromosome segregation through mitotic slippage. These interphase cells can then either undergo cell death or enter a permanent senescence-like arrest state.

A major challenge for studying mitotic catastrophe is to be able to distinguish it from cells undergoing normal mitosis or other forms of cell death. It should be noted that the events leading to cell death are decisive for determining if a cell undergoes mitotic catastrophe. A major drawback of endpoint analyses such as apoptosis assays is that they do not provide information on the history of the cell before cell death. Temporal information obtained from approaches such as time-lapse imaging is paramount in establishing if cell death is a direct consequence of mitosis.

Inducing Mitotic Catastrophe

Unscheduled Entry into Mitosis

Cyclin-dependent kinase 1 (CDK1) is the major protein kinase that drives cells into normal mitosis. CDK1 is activated by binding to B-type cyclins (mainly cyclin B1), which then phosphorylates substrates critical for entry into mitosis. Destruction of cyclin B1 provides a mechanism to rapidly inactivate CDK1 and allow the cell to exit mitosis (Fung and Poon, 2005).

CDK1 is present throughout the cell cycle. In contrast, cyclin B1 accumulates and forms a complex with CDK1 after entering S phase. The complex is kept inactive before mitosis by MYT1 and WEE1 through phosphorylation of CDK1^{Thr14/Tyr15}. At the end of G₂ phase, the stockpile of inactive cyclin B1-CDK1 is activated abruptly by members of the CDC25 phosphatase family. Cyclin B1-CDK1 catalyzes its own activation by simultaneously stimulating CDC25 activation and WEE1 inactivation (Lindqvist *et al.*, 2009). This bistable system is believed to be kick-started by PLK1, initiating the activation of CDC25 and inactivation of WEE1/MYT1

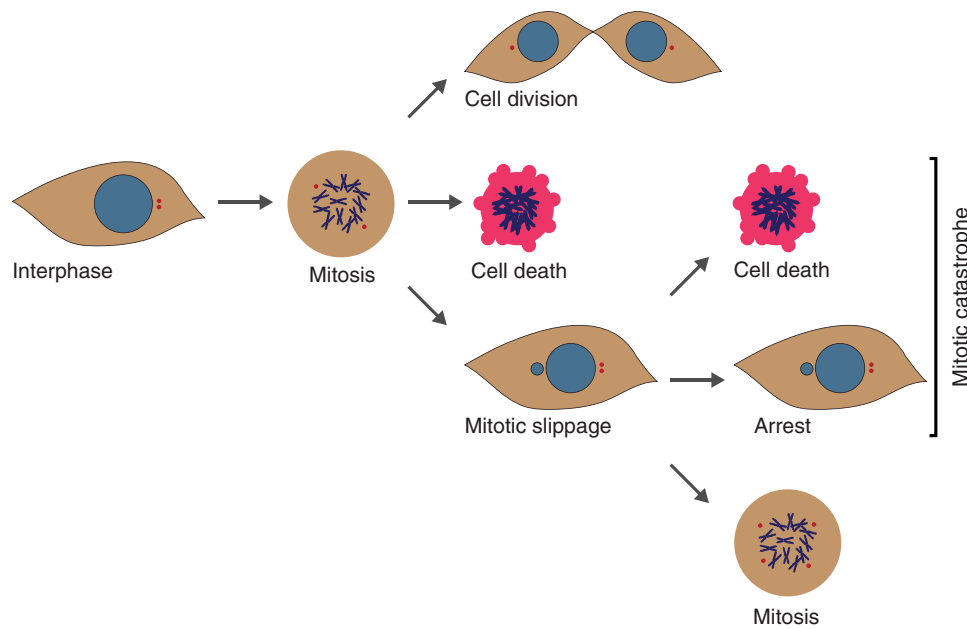


Figure 1 Different cell fates can occur after aberrant mitosis. Aberrant mitosis can give rise to multiple cell fates. Cell death can occur during the mitotic block (mitotic catastrophe). Alternatively, cells can exit mitosis without division by mitotic slippage, forming tetraploid cells (typically containing multiple nuclei). These cells can then undergo cell death or permanent cell cycle arrest (mitotic catastrophe). However, tetraploid cells can also re-replicate the DNA and centrosomes (red dots), entering another round of mitosis. Depending on the levels of defects during mitosis, it is possible that mitotic cells can perform cell division relatively normally (top of the diagram). Cell death or arrest can occur in the daughter cells (not shown in the diagram) and the process can also be regarded as mitotic catastrophe.

(van Vugt and Medema, 2005). The activation of PLK1 in turn requires phosphorylation of PLK1^{Thr210} by Aurora A, an event that is assisted by Bora. Binding of Bora to PLK1 is stimulated by cyclin B1-CDK1-dependent phosphorylation, creating yet another positive feedback loop in the activation of CDK1 (Lindqvist *et al.*, 2009).

To maintain genome stability, it is vital to prevent mitotic entry before DNA replication or DNA repair is completed. Several checkpoints participate in monitoring DNA integrity and prevent precocious entry into mitosis. It is now clear that similar underlying principles are involved in these checkpoints to prevent the activation of CDK1 (Kastan and Bartek, 2004).

Following exposure to ionizing radiation or other genotoxic insults that elicit DNA double-strand breaks, ATM is autophosphorylated on Ser1981, leading to release of active monomers from homodimer complexes. The ATM-related kinase ATR is activated by a broader spectrum of stress including hypoxia and replication stress. ATM/ATR then phosphorylates residues in the SQ/TQ domain of CHK1/CHK2, thereby stimulating the activity of these effector kinases (Smith *et al.*, 2010).

Stalled replication forks, which can be induced by insufficient nucleotide supply or lesions and obstacles on the DNA, mainly activate the ATR-CHK1 pathway. This pathway is essential even in the absence of exogenous stress during unperturbed S phase, probably for maintaining high rates of replication fork progression (Petermann and Caldecott, 2006).

Once the ATM/ATR-CHK1/CHK2 cascade is activated, CDK1 activity is suppressed by inhibitory phosphorylation (Chen and Poon, 2008). CHK1 phosphorylates and activates

WEE1 by promoting 14-3-3 binding. In addition, CHK1/CHK2 also represses all three isoforms of the CDC25 phosphatase family. Specifically, phosphorylation of CDC25C by CHK1/CHK2 inactivates its phosphatase activity both directly and indirectly through the creation of a 14-3-3 binding site. Binding of 14-3-3 masks a proximal nuclear localization sequence and anchors CDC25C in the cytoplasm, preventing efficient access of CDC25C to its substrate cyclin B1-CDK1. CDC25B, which possesses a unique role in activating cyclin B1-CDK1 at the centrosome, is also inactivated by 14-3-3 binding following CHK1-dependent phosphorylation. Finally, CDC25A is targeted for rapid degradation by SCF ^{β -TrCP}-dependent ubiquitination after CHK1-dependent phosphorylation. Collectively, these mechanisms suppress CDK1 activity when DNA replication or repair is not completed.

It is anticipated that untimely mitosis from distinct periods of the cell cycle may exert different effects on mitotic catastrophe. While normal G₁ cells probably contain insufficient cyclin B1-CDK1 complexes to support precocious mitosis, cells in S phase start to accumulate cyclin B1. Stimulating mitosis in cells with incompletely replicated DNA should generate extensive defects in chromosome segregation. Similarly, the presence of unrepaired DNA damage should affect mitosis profoundly. Conceptually, it is less clear if mitotic catastrophe is induced by accelerating G₂ cells into mitosis.

Uncoupling of the DNA integrity checkpoints promotes unscheduled entry into mitosis. It has become increasingly clear that mitotic catastrophe is a major outcome of unscheduled mitosis (Chan *et al.*, 1999; Castedo *et al.*, 2004; Vogel *et al.*, 2007). Cells that possess defective checkpoints,

such as those derived from ataxia-telangiectasia, often exhibit radio-resistant DNA synthesis and mitotic catastrophe (Lavin and Khanna, 1999). Since the ultimate effect of the checkpoints is an accumulation of CDK1 inhibitory phosphorylation, it is not surprising that expression of a non-phosphorylatable mutant of CDK1 can also trigger mitotic catastrophe (Blasina *et al.*, 1997; Niida *et al.*, 2005; Chow *et al.*, 2003).

Agents that target components of the ATM/ATR–CHK1/CHK2 axis can induce mitotic catastrophe. Indeed, a number of these small-molecule inhibitors are promising anticancer agents (Furgason and Bahassi *et al.*, 2013). The spindle-assembly checkpoint is required for mitotic catastrophe induced by abrogation of the DNA integrity checkpoints (On *et al.*, 2011; Nitta *et al.*, 2004; Vogel *et al.*, 2004; Chen *et al.*, 2012), suggesting that a trap in mitosis is indispensable for these types of cell death.

Mitotic Block

At the end of mitosis, cyclin B1 is removed by a ubiquitin ligase called the anaphase-promoting complex/cyclosome (APC/C) loaded with the targeting subunit CDC20 (Yu, 2007). Activated cyclin B1–CDK1 catalyzes its own destruction by stimulating the activity of APC/C^{CDC20}. APC/C^{CDC20} is also responsible for degrading several proteins that are important for mitotic exit. For example, degradation of securin by APC/C^{CDC20} releases separase, which in turn cleaves cohesin to facilitate sister chromatid separation (Yanagida, 2000). Proteolysis of geminin releases CDT1, enabling CDT1 to form the pre-replication complex required for the next round of DNA replication (Seo and Kroll, 2006).

Activation of APC/C^{CDC20} is initiated only when all the chromosomes have achieved bipolar attachment to the mitotic spindles. Unattached kinetochores or the absence of tension between the paired kinetochores activates a surveillance mechanism termed the spindle-assembly checkpoint (Figure 2). This checkpoint inhibits APC/C^{CDC20}, thereby

maintaining a high concentration of active cyclin B1–CDK1 (Musacchio and Salmon, 2007).

Mechanistically, unattached kinetochores attract the components of the checkpoint machinery (including BUB1, BUB3, MAD1, MAD2, MAD3/BUBR1, MPS1, and CENP-E), catalyzing the formation of a diffusible mitotic checkpoint complex (MCC, components include MAD2, BUBR1, and BUB3), which in turn inhibit APC/C^{CDC20}. The binding to CDC20 involves a conformational change in MAD2 from a less stable open conformation (known as O-MAD2) to the more stable close conformation (C-MAD2). According to a currently favored MAD2-template model, the conversion of O-MAD2 to C-MAD2 is autocatalytic, supporting the cytosolic propagation of the checkpoint signal away from the kinetochores (Musacchio and Salmon, 2007). Once all the kinetochores are properly attached, the spindle-assembly checkpoint is silenced (through mechanisms that are still not clearly defined) to enable APC/C^{CDC20} activation and mitotic exit.

In normal mitosis, the dynamic nature of microtubules facilitates efficient capturing of kinetochores by the spindles. Moreover, incorrectly attached kinetochores (such as both kinetochores of a chromosome are attached to the same spindle pole) are corrected through mechanisms that monitor tension between paired kinetochores. However, genetic alterations of components of spindle formation or the spindle-assembly checkpoint can result in severe delay in mitosis.

Agents that suppress microtubules dynamics can also lead to protracted activation of the spindle-assembly checkpoint (Weaver and Cleveland, 2005). These include chemicals that attenuate microtubules depolymerization (e.g., Taxol) or polymerization (e.g., vinca alkaloid). The stabilization of APC/C targets and persistent activation of cyclin B1–CDK1 ensure that the cell remains in mitosis. Other antimitotic chemicals that target essential mitotic regulators such as Aurora A, PLK1, or the kinesin Eg5 (KIF11) can also trigger protracted mitotic arrest.

The responses to prolonged mitotic block vary greatly between different cancer cell lines as well as between individual

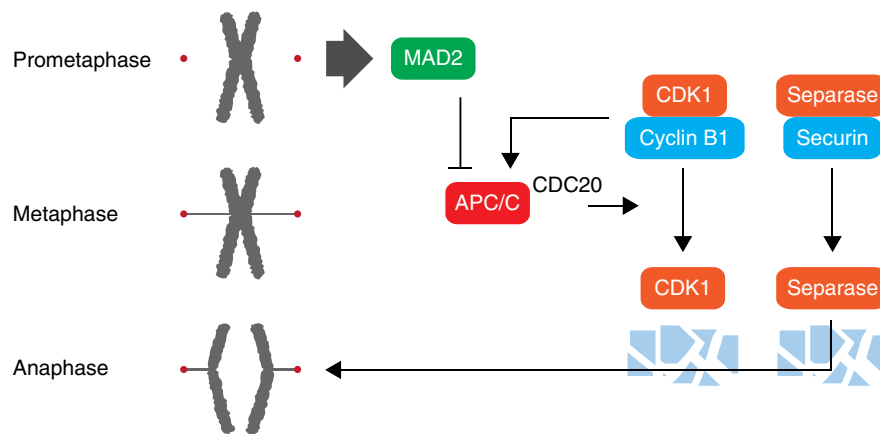


Figure 2 Regulation of mitotic exit by APC/C^{CDC20}. During early mitosis, APC/C^{CDC20} is stimulated by cyclin B–CDK1. But concurrently, the activity of APC/C^{CDC20} is suppressed by the spindle-assembly checkpoint (MAD2). The checkpoint is activated by either the presence of unattached kinetochores or the absence of tension between paired kinetochores. Once all kinetochores are properly attached (metaphase), the spindle-assembly checkpoint is silenced to enable APC/C^{CDC20} activation. APC/C^{CDC20} then targets several proteins including cyclin B1 and securin to ubiquitin-mediated degradation. Proteolysis of securin releases separase, which in turn cleaves cohesin to promote sister chromatid separation and anaphase.

cells from the same cell line (Gascoigne and Taylor, 2008). Cells can either execute cell death directly during mitosis or undergo mitotic slippage. During mitotic slippage, cyclin B1–CDK1 is downregulated precociously and cells exit mitosis without chromosome segregation and cytokinesis. The nuclear envelope then reforms randomly around groups of chromosome. In this connection, the resulting micronucleation after mitotic slippage has been used in some studies as one of the morphological features associated with mitotic catastrophe.

The central event of mitotic slippage appears to be a slow but continuous degradation of cyclin B1 (Brito and Rieder, 2006). Whether a cell dies during prolonged mitotic block is believed to be determined by a competition between the slow degradation of cyclin B1 (which affects the rate of mitotic slippage) and the activation of cell death pathways (accumulation of death activators or loss of death inhibitors). Hence cell death during mitotic arrest is the consequence of the cell breaching the death threshold before the mitotic slippage threshold is reached.

Although mitotic cell death is prevented immediately by mitotic slippage, the G_1 cells generated are less fit to propagate than normal cells. Mitotic slippage generates cells containing tetraploid DNA contents and two centrosomes, both of which can be further duplicated during the subsequent S phase. As centrosomes are microtubule organization centers, there is a possibility that cells with supernumerary centrosomes can form multipolar spindles during the subsequent mitosis. The uneven segregation of genetic material into daughter cells precipitates different cell fates, including mitotic catastrophe, aneuploidy, and transformation (Storchova and Kuffer, 2008).

It has been proposed that after mitotic slippage, a p53-dependent ‘tetraploidy checkpoint’ that detects abnormal number of chromosomes or centrosomes is activated and prevents S phase entry. However, the existence of the tetraploidy checkpoint has been disputed. The fact that many cells undergo p53-dependent cell cycle arrest or apoptosis after mitotic slippage is likely to be due to DNA damage or centrosomal stress during the aberrant mitosis (Hayashi and Karlseder, 2013).

Mechanisms of Mitotic Catastrophe

Mitotic catastrophe involves cell death or arrest during or after a defective mitosis. Although cell death during mitotic catastrophe frequently displays a phenotypic characteristic of apoptosis and shares several molecular hallmarks with apoptosis such as caspase activation, there is evidence that it can also occur by necrosis (Vakifahmetoglu *et al.*, 2008).

The molecular basis of cell death during mitotic catastrophe remains incompletely understood. As discussed above, whether cell death occurs during mitosis is likely to be determined by a competition between the slow degradation of cyclin B1 and the accumulation of cell death signals. There is good but not definitive evidence that cyclin B1–CDK1 activity is involved in activating the death signals. While the underlying principles of CDK1-mediated toxicity remain largely unresolved, a few targets that affect survival have been identified. For example, several members of the B-cell lymphoma 2 (BCL-2) family are phosphorylated by cyclin B1–CDK1.

Phosphorylation of Bcl-2-associated death promoter (BAD) by cyclin B1–CDK1 reduces the interaction with 14-3-3, allowing BAD to translocate to mitochondria and induce cell death (Berndtsson *et al.*, 2005; Konishi *et al.*, 2002). BCL-2 is also phosphorylated by cyclin B1–CDK1, but the role in mitotic catastrophe remains to be clarified (Brichese *et al.*, 2002; Furukawa *et al.*, 2000; Ling *et al.*, 1998; Scatena *et al.*, 1998). B-cell lymphoma-extra large (BCL-XL) is phosphorylated by cyclin B1–CDK1, leading to the inactivation of its anti-apoptotic activity (Terrano *et al.*, 2010; Sakurikar *et al.*, 2012). Finally, cyclin B1–CDK1 also phosphorylates the anti-apoptotic protein MCL-1, inducing the rapid degradation of MCL-1 in an APC/C^{CDK20}-dependent manner (Harley *et al.*, 2010; Chu *et al.*, 2012).

If there are large gaps in our knowledge regarding the mechanisms of mitotic cell death, even less is known about the mechanisms tipping the balance toward cell death after mitotic slippage. The DNA damage or centrosomal stress generated during the aberrant mitosis is likely to be responsible for activating p53 (Hayashi and Karlseder, 2013). This at least explains the apoptosis and senescence after mitotic slippage in cells containing wild-type p53.

An important implication is that cells that evade mitotic catastrophe can enter a tetraploidy G_1 state. If additional checkpoints are not present (such as being p53 defective), aneuploid cells can further undergo DNA replication and centrosome duplication, paving the way to multipolar mitosis and further genome instability. It is likely that the ability to execute cytokinesis reduces as the ploidy increases, giving rise to multinucleated giant cells (King, 2008).

The presence of polyploid giant cells that fail to be killed by mitotic catastrophe may also account for resistance to cancer therapy. For example, following DNA damage, many polyploid cells appear after an initial phase of mitotic catastrophe and survive for weeks as mono- or multinucleated giant cells (Blagosklonny, 1999). Although whether giant cells still retain proliferative potential is controversial, there is evidence that they can undergo multipolar mitosis to return to near diploid state (Erenpreisa and Cragg, 2001). These studies suggest that the multistep process of escaping mitotic catastrophe through polyploidization followed by depolyploidization may account for tumor relapse after initial efficient cancer therapy.

Concluding Remarks

Mitotic catastrophe is characterized by an unscheduled or prolonged activation of mitotic drivers followed by cell death. In some cases, permanent cell cycle arrest following the aberrant mitosis is also regarded as an outcome of mitotic catastrophe. Causes of mitotic catastrophe include unscheduled entry into mitosis from interphase or prolonged mitotic arrest caused by the activation of the spindle-assembly checkpoint.

Addressing the outstanding issues concerning mitotic catastrophe, including the various conditions that can trigger the process, mechanisms that regulate cell death and mitotic slippage, and the genetic determinants that modulate various cell fates will transform our understanding of both tumorigenesis and antimitotic drug treatments.

Acknowledgment

Related works at the author laboratory are supported in part by the Research Grants Council.

References

- Berndtsson, M., Konishi, Y., Bonni, A., *et al.*, 2005. Phosphorylation of BAD at Ser-128 during mitosis and paclitaxel-induced apoptosis. *FEBS Letters* 579, 3090–3094.
- Blagosklonny, M.V., 1999. Drug-resistance enables selective killing of resistant leukemia cells: Exploiting of drug resistance instead of reversal. *Leukemia* 13, 2031–2035.
- Blasina, A., Paegle, E.S., McGowan, C.H., 1997. The role of inhibitory phosphorylation of CDC2 following DNA replication block and radiation-induced damage in human cells. *Molecular Biology of the Cell* 8, 1013–1023.
- Brichese, L., Barboule, N., Helier, C., Valette, A., 2002. Bcl-2 phosphorylation and proteasome-dependent degradation induced by paclitaxel treatment: Consequences on sensitivity of isolated mitochondria to Bid. *Experimental Cell Research* 278, 101–111.
- Brito, D.A., Rieder, C.L., 2006. Mitotic checkpoint slippage in humans occurs via cyclin B destruction in the presence of an active checkpoint. *Current Biology* 16, 1194–1200.
- Castedo, M., Perfettini, J.-L., Roumier, T., *et al.*, 2004. The cell cycle checkpoint kinase Chk2 is a negative regulator of mitotic catastrophe. *Oncogene* 23, 4353–4361.
- Chan, T.A., Hermeking, H., Lengauer, C., Kinzler, K.W., Vogelstein, B., 1999. 14-3-3Sigma is required to prevent mitotic catastrophe after DNA damage. *Nature* 401, 616–620.
- Chen, Y., Chow, J.P., Poon, R.Y., 2012. Inhibition of Eg5 acts synergistically with checkpoint abrogation in promoting mitotic catastrophe. *Molecular Cancer Research* 10, 626–635.
- Chen, Y., Poon, R.Y., 2008. The multiple checkpoint functions of CHK1 and CHK2 in maintenance of genome stability. *Frontiers in Bioscience* 13, 5016–5029.
- Chow, J.P., Siu, W.Y., Ho, H.T., *et al.*, 2003. Differential contribution of inhibitory phosphorylation of CDC2 and CDK2 for unperturbed cell cycle control and DNA integrity checkpoints. *Journal of Biological Chemistry* 278, 40815–40828.
- Chu, R., Terrano, D.T., Chambers, T.C., 2012. Cdk1/cyclin B plays a key role in mitotic arrest-induced apoptosis by phosphorylation of Mcl-1, promoting its degradation and freeing Bak from sequestration. *Biochemical Pharmacology* 83, 199–206.
- Erenpreisa, J., Cragg, M.S., 2001. Mitotic death: A mechanism of survival? A review. *Cancer Cell International* 1, 1.
- Fung, T.K., Poon, R.Y., 2005. A roller coaster ride with the mitotic cyclins. *Seminars in Cell & Developmental Biology* 16, 335–342.
- Furgason, J.M., Bahassi el, M., 2013. Targeting DNA repair mechanisms in cancer. *Pharmacology & Therapeutics* 137, 298–308.
- Furukawa, Y., Iwase, S., Kikuchi, J., *et al.*, 2000. Phosphorylation of Bcl-2 protein by CDC2 kinase during G2/M phases and its role in cell cycle regulation. *Journal of Biological Chemistry* 275, 21661–21667.
- Galluzzi, L., Vitale, I., Abrams, J.M., *et al.*, 2012. Molecular definitions of cell death subroutines: Recommendations of the nomenclature committee on cell death 2012. *Cell Death & Differentiation* 19, 107–120.
- Gascoigne, K.E., Taylor, S.S., 2008. Cancer cells display profound intra- and interline variation following prolonged exposure to antimitotic drugs. *Cancer Cell* 14, 111–122.
- Harley, M.E., Allan, L.A., Sanderson, H.S., Clarke, P.R., 2010. Phosphorylation of Mcl-1 by CDK1-cyclin B1 initiates its Cdc20-dependent destruction during mitotic arrest. *EMBO Journal* 29, 2407–2420.
- Hayashi, M.T., Karlseder, J., 2013. DNA damage associated with mitosis and cytokinesis failure. *Oncogene* 32, 4593–4601.
- Kastan, M.B., Bartek, J., 2004. Cell-cycle checkpoints and cancer. *Nature* 432, 316–323.
- King, R.W., 2008. When 2 + 2 = 5: The origins and fates of aneuploid and tetraploid cells. *Biochimica et Biophysica Acta* 1786, 4–14.
- Konishi, Y., Lehtinen, M., Donovan, N., Bonni, A., 2002. Cdc2 phosphorylation of BAD links the cell cycle to the cell death machinery. *Molecular Cell* 9, 1005–1016.
- Lavin, M.F., Khanna, K.K., 1999. ATM: The protein encoded by the gene mutated in the radiosensitive syndrome ataxia-telangiectasia. *International Journal Radiation Biology* 75, 1201–1214.
- Lindqvist, A., Rodriguez-Bravo, V., Medema, R.H., 2009. The decision to enter mitosis: Feedback and redundancy in the mitotic entry network. *Journal of Cell Biology* 185, 193–202.
- Ling, Y.H., Tornos, C., Perez-Soler, R., 1998. Phosphorylation of Bcl-2 is a marker of M phase events and not a determinant of apoptosis. *Journal of Biological Chemistry* 273, 18984–18991.
- Molz, L., Booher, R., Young, P., Beach, D., 1989. cdc2 and the regulation of mitosis: Six interacting mcs genes. *Genetics* 122, 773–782.
- Musacchio, A., Salmon, E.D., 2007. The spindle-assembly checkpoint in space and time. *Nature Reviews Molecular Cell Biology* 8, 379–393.
- Niida, H., Tsuge, S., Katsuno, Y., *et al.*, 2005. Depletion of Chk1 leads to premature activation of Cdc2-cyclin B and mitotic catastrophe. *Journal of Biological Chemistry* 280, 39246–39252.
- Nitta, M., Kobayashi, O., Honda, S., *et al.*, 2004. Spindle checkpoint function is required for mitotic catastrophe induced by DNA-damaging agents. *Oncogene* 23, 6548–6558.
- On, K.F., Chen, Y., Ma, H.T., Chow, J.P., Poon, R.Y., 2011. Determinants of mitotic catastrophe on abrogation of the G2 DNA damage checkpoint by UCN-01. *Molecular Cancer Therapeutics* 10, 784–794.
- Petermann, E., Caldecott, K.W., 2006. Evidence that the ATR/Chk1 pathway maintains normal replication fork progression during unperturbed S phase. *Cell Cycle* 5, 2203–2209.
- Russell, P., Nurse, P., 1986. cdc25 + Functions as an inducer in the mitotic control of fission yeast. *Cell* 45, 145–153.
- Russell, P., Nurse, P., 1987. The mitotic inducer nim1 + functions in a regulatory network of protein kinase homologs controlling the initiation of mitosis. *Cell* 49, 569–576.
- Sakurikar, N., Eichhorn, J.M., Chambers, T.C., 2012. Cyclin-dependent kinase-1 (Cdk1)/cyclin B1 dictates cell fate after mitotic arrest via phosphoregulation of antiapoptotic Bcl-2 proteins. *Journal of Biological Chemistry* 287, 39193–39204.
- Scatena, C.D., Stewart, Z.A., Mays, D., *et al.*, 1998. Mitotic phosphorylation of Bcl-2 during normal cell cycle progression and Taxol-induced growth arrest. *Journal of Biological Chemistry* 273, 30777–30784.
- Seo, S., Kroll, K.L., 2006. Geminin's double life: Chromatin connections that regulate transcription at the transition from proliferation to differentiation. *Cell Cycle* 5, 374–379.
- Smith, J., Tho, L.M., Xu, N., Gillespie, D.A., 2010. The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. *Advances in Cancer Research* 108, 73–112.
- Storchova, Z., Kuffer, C., 2008. The consequences of tetraploidy and aneuploidy. *Journal of Cell Science* 121, 3859–3866.
- Terrano, D.T., Upreti, M., Chambers, T.C., 2010. Cyclin-dependent kinase 1-mediated Bcl-xL/Bcl-2 phosphorylation acts as a functional link coupling mitotic arrest and apoptosis. *Molecular and Cellular Biology* 30, 640–656.
- Vakifahmetoglu, H., Olsson, M., Zhivotovsky, B., 2008. Death through a tragedy: Mitotic catastrophe. *Cell Death & Differentiation* 15, 1153–1162.
- Vogel, C., Hager, C., Bastians, H., 2007. Mechanisms of mitotic cell death induced by chemotherapy-mediated G2 checkpoint abrogation. *Cancer Research* 67, 339–345.
- Vogel, C., Kienitz, A., Hofmann, I., Muller, R., Bastians, H., 2004. Crosstalk of the mitotic spindle assembly checkpoint with p53 to prevent polyploidy. *Oncogene* 23, 6845–6853.
- van Vugt, M.A., Medema, R.H., 2005. Getting in and out of mitosis with Polo-like kinase-1. *Oncogene* 24, 2844–2859.
- Weaver, B.A., Cleveland, D.W., 2005. Decoding the links between mitosis, cancer, and chemotherapy: The mitotic checkpoint, adaptation, and cell death. *Cancer Cell* 8, 7–12.
- Yanagida, M., 2000. Cell cycle mechanisms of sister chromatid separation; roles of Cut1/separin and Cut2/securin. *Genes to Cells* 5, 1–8.
- Yu, H., 2007. Cdc20: A WD40 activator for a cell cycle degradation machine. *Molecular Cell* 27, 3–16.

Autophagy refers to the group of cellular processes that mediate degradation of intracellular components in lysosomes and the subsequent recycling of their constituents (Mizushima *et al.*, 2008). As such, autophagy shares a common final compartment with endocytosis and phagocytosis, the lysosome. Continuous turnover of intracellular proteins and organelles allows for their renewal and serves as a mechanism for quality control of newly synthesized cellular structures. Furthermore, the regulated balance between synthesis and degradation of proteins and organelles permits for rapid modification of their cellular levels to accommodate to environmental changes.

Historically, the discovery of autophagy was temporally close to that of lysosomes, the subcellular compartment with higher content of hydrolases in the cell (Ohsumi, 2014). Protein degradation by autophagy was initially linked mainly to nutritional changes. Starvation or conditions in which nutrients become scarce often lead to maximal activation of autophagy to facilitate recycling of intracellular components for anabolic purposes. Although nutrient deprivation is one of the best characterized activators of autophagy, better understanding of the molecular basis of the autophagic pathways and the development of tools and markers that permit tracking flux of substrates through the autophagic system have revealed that autophagy is activated in response to a multiplicity of stimuli ranking from cellular stressors (oxidative stress, DNA damage, mitochondria depolarization, etc.) to extracellular cues and inputs (growth factors, cytokines, pathogens, etc.). This better understanding of the conditions associated with autophagy activation, and direct studies analyzing the consequences of upregulating or downregulating autophagy in whole organisms have unveiled the participation of autophagy in an impressive number of cellular functions: energetic balance quality control, cell and tissue remodeling, cell differentiation, cellular defense, innate immunity, cell death, cellular senescence, among others (Green and Levine, 2014; Kenific and Debnath, 2015; Marino *et al.*, 2014). The multiple physiological roles of autophagy along with the connection of autophagy malfunctioning to common human disorders (neurodegeneration, diabetes, muscle waste, infectious diseases, etc.) have motivated the recent interest in the understanding of the autophagic processes at the molecular level (Jiang and Mizushima, 2014; Schneider and Cuervo, 2014).

Autophagic Pathways

Differences in the molecular effectors that contribute to delivery of cargo (material to be degraded) into lysosomes are the bases for the distinction among autophagic pathways. When cargo is sequestered in double membrane vesicles in the cytosol that then acquire the hydrolases required for degradation through lysosomal fusion, autophagy is referred to as macroautophagy (Yang and Klionsky, 2010; Figure 1). In the case of

microautophagy, cargo is directly sequestered through invaginations or tubulations of the lysosomal or late endosomal membrane and are released into the lumen of these organelles after pinching off (de Waal *et al.*, 1986; Figure 2). Not all forms of autophagy require vesicles for cargo sequestration; for example, in chaperone-mediated autophagy (CMA) cytosolic proteins reach the lysosomal lumen upon translocation across the lysosomal membrane (Kaushik and Cuervo, 2012; Figure 3). Different variants have already been described for each of these mechanisms of cargo delivery to lysosomes.

Studies in yeast were key for the identification of a subset of about 35 genes (ATG or autophagy-related genes) that participate in the process of macroautophagy as well as some microautophagy specific genes, whereas the characterization of CMA and some specialized forms of macro- and microautophagy was performed in mammalian cells. For simplicity we will use the mammalian nomenclature for all the autophagy-related genes in this entry.

Autophagy Step-by-Step

Induction of Autophagy

Although multiple signaling mechanisms lead to activation of macroautophagy, the better characterized is the mammalian target of rapamycin (mTor) signaling pathway, closely linked to starvation-induced autophagy. The energy-sensing kinase complex target of rapamycin complex 1 (TORC1) displays maximal activity in the presence of nutrients, leading to the phosphorylation and sequestration of ULK1, a kinase required for autophagy initiation (Kanazawa *et al.*, 2004; Hara *et al.*, 2008; Figure 1). Upon deprivation of nutrients, inactivation of mTOR results in release of ULK1 and its localization at the sites of autophagosome formation. The limiting membrane of these organelles forms through combination of proteins and lipids from different cellular compartments such as the endoplasmic reticulum (ER), mitochondria, plasma membrane, Golgi, and sorting endosomes. In all these cases, the area in the membrane from where the limiting membrane or phagophore originates is 'marked' through phosphorylation of lipids in these membranes. Besides ULK1, Vps34 is the second lipid kinase commonly recruited to sites of autophagosome biogenesis as part of the Beclin/Vps15/Vps34 complex (Figure 1). Elongation of the membrane occurs through two ubiquitin-like conjugation processes (Ichimura *et al.*, 2000): conjugation of two proteins Atg12 to Atg5 and conjugation of a protein LC3 to the lipid phosphatidylethanolamine. A common ligase, Atg7, is shared by both conjugation events. Tethering of Atg16L-enriched micelles with conjugate LC3 also contributes to growth of the limiting membrane (Ravikumar *et al.*, 2010).

The first signaling mechanism identified for CMA was through the retinoid-acid receptor alpha that acts as an endogenous repressor of this pathway (Anguiano *et al.*, 2013).

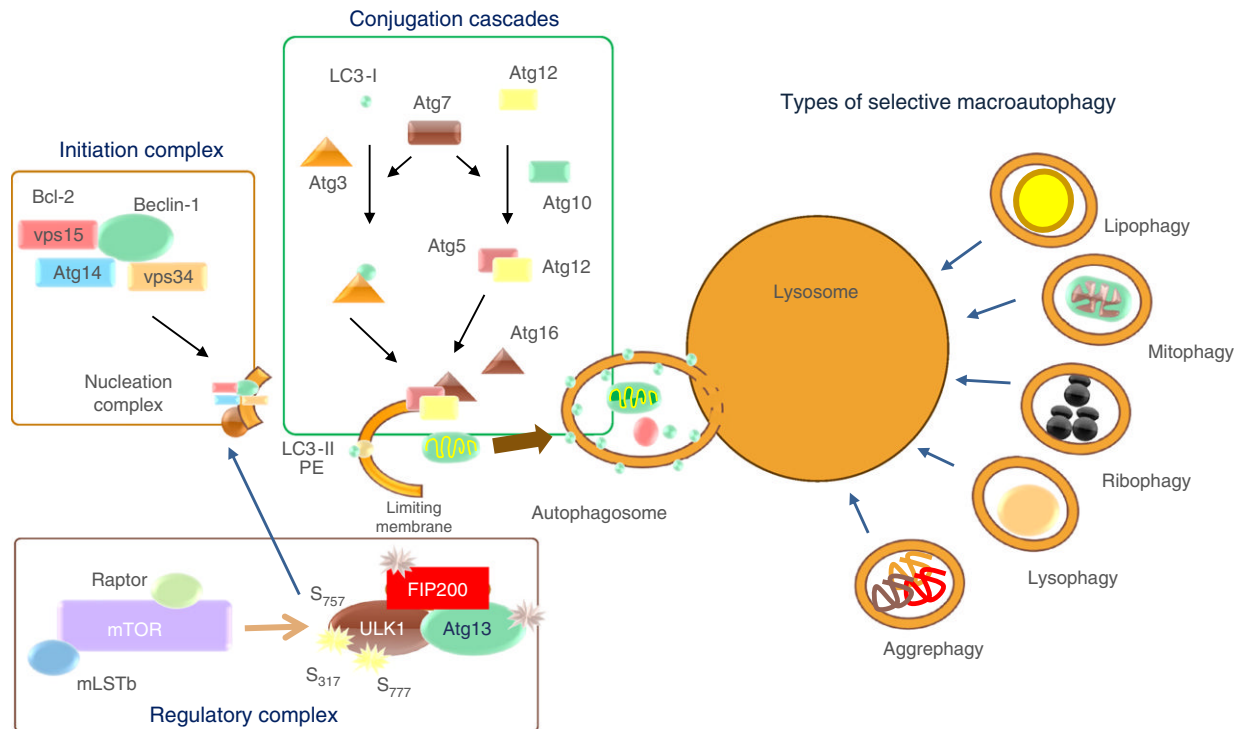


Figure 1 Macroautophagy and selective macroautophagy variants. Left: simplified schematic of the steps of macroautophagy and protein complexes that modulate some of these steps. Right: types of selective macroautophagy.

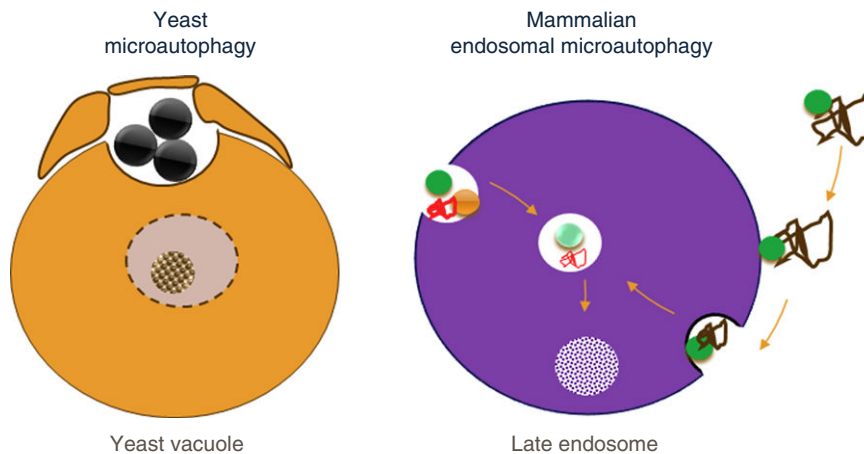


Figure 2 Yeast and mammalian microautophagy. Left: yeast microautophagy occurs at the vacuole membrane in a multistep process that allows for cargo recognition, formation of the microphagy membrane apparatus (MIPA), and internalization into single-membrane vesicles that pinch off and are subsequently degraded. Right: in mammals, microautophagy has only been described in late endocytic compartments, where cargo can be sequestered in bulk or selectively enriched upon hsc70 binding in vesicles formed in these membranes through assembly of the ESCRT proteins.

Recently, studies in T cells have unveiled activation of CMA by reactive oxygen species (ROS)-dependent signaling through the calcineurin–NFAT (nuclear factor of activated T cells) axis (Valdor *et al.*, 2014). The signaling mechanism that initiates microautophagy remains unknown.

Cargo Recognition

Macroautophagy was considered an ‘in bulk’ process in which all the cytosolic components present in the area surrounding

the site of autophagosome biogenesis were sequestered at once. However, the selectivity of this process has now been demonstrated after the identification of a subset of proteins, cargo receptors, that have the ability to simultaneously bind to cargo molecules and Atg proteins (Stolz *et al.*, 2014). This dual recruitment facilitates formation of the limiting membrane selectively around mitochondria (mitophagy), peroxisomes (pexophagy), lipid droplets (lipophagy), ribosomes (ribophagy), pathogens (xenophagy), aggregates (aggrephagy), etc. (Figure 1). Some of these cargo receptors, p62 or NBr1, are

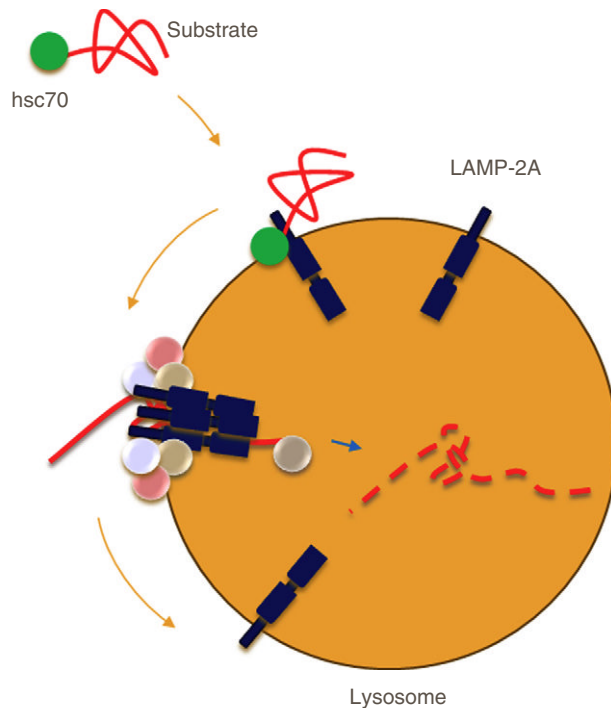


Figure 3 Chaperone-mediated autophagy. Schematic of the main steps of CMA. Upon recognition of the targeting motif by hsc70, the protein-chaperone complex is targeted to the lysosomal membrane where it binds to LAMP-2A and induces the multimerization and formation of the translocation complex. After substrate unfolding, the luminal chaperone completes translocation of the substrate.

shared among these processes due to their ability to recognize common tagging molecules such as ubiquitin (Novak and Dikic, 2011). A pair of cytosolic chaperones, hsc70 and Bag1, can also serve as cargo recognition molecules for aggregate proteins in a variant of conventional macroautophagy known as chaperone-assisted selective autophagy (CASA) (Arndt *et al.*, 2010).

In yeast microautophagy, sequestration of cytosolic components such as peroxisomes (micropexophagy), nucleus (piecemeal autophagy), or pathogens (microxenophagy) by the invaginations at the vacuole membrane occurs at specific contact sides through interaction of cargo and vacuole membrane proteins (Sakai *et al.*, 1998; Roberts *et al.*, 2003). Although microautophagy has not been described in mammalian cells, a similar process of invagination and internalization in single-membrane proteins occurs in the late endocytic compartment/multivesicular bodies and has been named endosomal microautophagy (e-MI) (Sahu *et al.*, 2011). Both in bulk and selective targeting of cytosolic proteins by hsc70 occur through e-MI (Figure 2).

Recognition of the cytosolic proteins that undergo degradation by CMA is mediated by binding of hsc70 to a pentapeptide motif in the protein sequence biochemically related to KFERQ (Dice, 1990). Mutagenesis of this targeting motif abolishes CMA of the protein, and addition of this pentapeptide sequence to proteins that usually do not undergo degradation by CMA reroutes them toward this pathway, supporting that this motif is necessary and sufficient for

targeting to CMA. Binding of hsc70 to this region is also required for selective e-MI of cytosolic proteins, but addition of this motif is not sufficient to drive e-MI of a protein (Sahu *et al.*, 2011; Figure 3).

Lysosomal Delivery

After sealing of the limiting membrane to form a double membrane vesicle in a SNARE (soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor)-mediated manner, the autophagosome moves through microtubules tracks, often toward the perinuclear region, to meet the lysosomes (Moreau *et al.*, 2011). Kiss-and-run events infuse part of the lysosomal hydrolases in autophagosomes forming a hybrid acid compartment known as autolysosome, where cargo is degraded.

In the case of microautophagy, delivery occurs through vesicles in the surface of the vacuole (forming a structure known as MIPA) in yeast, or through assembly of a series of protein complexes known as endosomal sorting complexes required for transport (ESCRT) proteins in the surface of the late endosome (in mammals). ESCRT proteins are well known because of their ability to form multivesicular bodies in the late endosomes for membrane protein degradation (Figure 2). Whether the two degradative processes, of cytosolic proteins and of plasma membrane proteins, occur in dedicate multivesicular bodies for each purpose or if they occur simultaneously in the same compartments remains unknown.

Once the complex hsc70/cytosolic proteins targeted for CMA reaches the lysosomal membrane they bind to the cytosolic tail of the lysosome-associated membrane protein type 2A (L2A) that is a spliced variant of the LAMP-2 gene. Substrate binding induces multimerization of L2A to form a 700 kDa complex required for protein translocation after unfolding (Cuervo and Dice, 1996; Bandyopadhyay *et al.*, 2008; Figure 3). Translocation also requires a luminal chaperone, a variant of cytosolic hsc70, but whether this chaperone is required for actual pulling of substrates inside lysosomes or just to prevent their retrotranslocation to the cytosol is not known. Translocation complexes are highly dynamic structures that deassemble into monomeric units of L2A right after the substrate crosses the lysosomal membrane, to allow a new cycle of binding and translocation. Chaperones in the cytosolic side of the lysosomal membrane are required for substrate unfolding and they also assist in L2A disassembly, whereas chaperones in the luminal side of the membrane participate in substrate translocation and contribute to maintain L2A stability when transitioning between monomers and multimeric complexes (Bandyopadhyay *et al.*, 2008; Figure 3).

Recycling, an Integral Part of Autophagy

The autophagy process does not end with the degradation of the cytosolic material in the lysosomal lumen; instead, both cargo and functional lysosomal components can be retrieved out of the autolysosome. Permeases dedicated to cytosolic export of amino acids and sugars have been identified in the vacuole membrane in yeast (Sekito *et al.*, 2008). Less is known about the mammalian lysosome permeases but specific amino

acid transporters have been known to exist in this membrane for a long time. Recycling of cargo components helps sustain cellular anabolic processes even in the absence of nutrients. Since the three autophagic pathways share the lysosome as a common site of degradation, these permeases and transporters also recycle amino acids from proteins degraded by CMA and microautophagy. In the case of e-MI, the cargo internalized in the multivesicular bodies can be degraded either in these late endosomes, transferred for lysosomal degradation, or exported extracellularly as exosomes upon fusion of lysosomes with the plasma membrane (Figure 2).

Lysosomal components can also be retrieved out of the lysosome into vesicles or tubules that form at the limiting membrane of the lysosome and then pinch off to constitute the basis of a new cargo-free lysosome. Vesiculation, tubulation, and sorting of lysosomal membrane proteins and enzymes into these structures is regulated by a family of lipid kinases that act in a coordinate manner with clathrin and motor proteins that bind to areas of vesicle formation (Yu *et al.*, 2010). This recycling is maximally activated after starvation conditions, when a large fraction of lysosomes has fused to autophagosomes.

Autophagy-Related Cellular Compartments

As described in the previous section, the cellular compartments that participate in macroautophagy are well characterized and include the phagophore or limiting membrane, autophagosomes, autolysosomes, and lysosomes. Autophagosomes can also fuse first with late endosomes, forming an amphisome, where extra- and intracellular material converge to ultimately reach the lysosome through endosome/lysosome fusion (Fengsrud *et al.*, 2000). Autophagic flux occurs, to some extent, even in the absence of *de novo* synthesis, thanks to a reservoir of Atg protein complexes already assembled but that remain in an inactive state through sequestration in different cellular compartments such as acetylated microtubules or plasma membrane connexins, in the case of the Beclin/Vps34 complexes (Geeraert *et al.*, 2010; Bejarano *et al.*, 2014), or lysosomal TORC1, in the case of ULK1 (Kanazawa *et al.*, 2004; Hara *et al.*, 2008). Whenever autophagy activation occurs, these complexes can be readily released and relocate to sites of autophagosome formation to contribute to formation of the limiting membrane. Sustained activation of autophagy requires an expansion of the lysosomal pool to meet the autophagic needs, which is attained in part through the above-mentioned recycling of lysosomal components out of autolysosomes, and in part through activation of *de novo* lysosomal biogenesis. In fact, most of the genes that code for lysosomal proteins and for Atgs are under the control of a family of transcription factors that recognizes a CLEAR motif in the promoter region of these genes. The first and best characterized member of this family is the transcription factor EB (TFEB) that is normally retained in the cytosol through phosphorylation by different kinases, including the TORC1 complex that sits in the surface of lysosomes (Settembre *et al.*, 2011). Amino acids exported from lysosomes maintain TORC1 active, TFEB phosphorylated, and the autophagic/lysosomal transcriptional program attenuated. However, when rates of degradation through lysosomes decrease, the absence of exiting

amino acids leads to TORC1 inactivation and the consequent reduced phosphorylation of TFEB facilitates its nuclear translocation and the activation of the transcriptional program responsible for autophagy and lysosome biogenesis. A newly identified member of this family, TFN3, seems to also undergo lysosomal mTOR phosphorylation and activate the lysosomal/autophagic program through comparable mechanisms (Martina *et al.*, 2014). The most ubiquitous distribution of TFN3 and its presence in organs such as the central nervous system, where levels of TFEB are negligible, support a more generalized regulatory role for this new protein.

Mammalian microautophagy has only been demonstrated so far in late endosomes/multivesicular bodies, compartments shared with endocytosis. However, the impact of expanded or reduced endocytic flux in e-MI remains unknown.

CMA occurs in a specific lysosomal subpopulation that bears in its lumen the lysosomal hsc70 chaperone required for substrate translocation. CMA-active lysosomes are secondary lysosomes that still retain the ability to fuse with autophagic and endocytic vesicles, although less efficiently than other lysosomal subgroups (Cuervo *et al.*, 1997). They correspond to the most acidic group of lysosomes, since lys-hsc70 is more stable at acidic pH and slight reduction in luminal acidification is sufficient to induce its rapid degradation. Rather than biogenically independent lysosomal subgroups, CMA lysosomes and lysosomes engaged in other vesicular processes rapidly convert into each other allowing for expansion or shrinkage of the pool of lysosomes most efficient for each autophagic pathway on demand.

The Autophagic Cargo and Purpose of Its Degradation

Activation of macroautophagy during starvation in yeast and mammals was considered necessary to replenish the intracellular pool of amino acids and sustain synthesis of proteins required for adaptation to starvation. However, more recently, macroautophagy has shown the capability to promote degradation of energetically more favorable molecules such as lipids and glycogen. Degradation of lipid droplets (lipophagy) starts with the formation of the limiting membrane *in situ*, as lipid droplets also contribute lipids for autophagosome biogenesis. This membrane sequesters a whole or a portion of lipid droplets and delivers them to lysosomes through autophagosome/lysosome fusion (Singh *et al.*, 2009). The lipases in the lysosomal lumen break down lipid droplet triglycerides into free fatty acids that, once exported from the lysosome, are oxidized in the mitochondria to generate energy. A similar process mediates degradation of glycogen stores into lysosomes upon autophagic sequestration (Kotoulas *et al.*, 2006).

Macroautophagy can also degrade portions or whole organelles such as mitochondria, ER, peroxisomes, and even lysosomes. Continuous basal degradation of organelles contributes to their renewal. Furthermore, selective macroautophagy can be activated to: (1) eliminate organelles when they are damaged or do no longer function, such as depolarized mitochondria or leaking lysosomes; (2) regulate the size of pool of cellular organelles and modulate their function, such as excess of peroxisomes or mitochondria; and (3) allow

for cellular differentiation or cell-type conversion, such as elimination of organelles in reticulocytes as they convert into red cells, or elimination of the spermatozoid mitochondria preovule fecundation (Singh and Cuervo, 2011).

Degradation of cytosolic content by macroautophagy is often used as a mechanism of cellular defense, such as activation of xenophagy, to eliminate pathogens that reach the cytosol or of aggrephagy to eliminate protein aggregates that could become toxic when accumulating inside cells (Levine, 2005).

Specialized regions of the plasma membrane are also targeted for autophagic degradation, often with regulatory purposes. For example, macroautophagy can have an impact in cell-to-cell communication, through degradation of connexins/gap junctions or of neuronal dendritic spines and synaptic components; it can also impact cellular motility through degradation of motile cilia (ciliophagy), environmental sensing through degradation of the primary cilia components, or cell adhesion through degradation of adhesion foci, all present at the plasma membrane.

Yeast microautophagy can also degrade proteins, organelles, and pathogens, but the reasons why this pathway is chosen in specific conditions over macroautophagy remain unknown. Mammalian endosomal microautophagy has only shown so far to degrade cytosolic proteins, but contrary to CMA, unfolding is not a prerequisite for their degradation through this pathway, making it potentially suitable for the elimination of oligomeric proteins before they consolidate into large aggregates.

Specific characteristics of cargo delivery by CMA make this pathway only suitable for degradation of single-soluble proteins that can undergo complete unfolding. In principle, all the CMA substrates come from the cytosol in order to bind the cytosolic tail of L2A. Thus, degradation of organelle proteins by CMA only occurs if these proteins relocate to the cytosol or after their *de novo* synthesis before they reach their final compartment. An example of this type of CMA degradation has been recently observed in a transgenic mouse model unable to perform CMA in the liver, where nuclearly encoded mitochondria enzymes accumulate in part in the cytosol but also inside the mitochondria due to their reduced degradation through CMA (Schneider *et al.*, 2014). This observation opens the intriguing possibility of CMA contributing to regulate the mitochondria enzymatic load. The fact that CMA substrate proteins should bind to the cytosolic tail of the receptor makes this pathway unsuitable for degradation of membrane proteins, at least as far as they are integrated in membranes, since retrotranslocation will not be possible in that case. However, it is possible that the CMA degradation recently proposed for a plasma membrane receptor occurs as quality control before membrane insertion of this protein, or that degradation of this type of proteins by CMA is preceded by a membrane-extraction step, comparable to the one proposed for chaperones before the degradation of membrane proteins by AAA proteases. Soluble cytosolic proteins, the canonical CMA cargo, undergo degradation through this pathway both with regulatory purposes or as a mechanism for protein quality control. CMA has been recently demonstrated to contribute to regulation of glucose and lipid metabolic pathways in the liver (Schneider *et al.*, 2014) and to the proteome remodeling

required for T cell activation, since selective degradation of specific subsets of proteins by CMA reduces their intracellular levels, and subsequently their function (Valdor *et al.*, 2014). The role of CMA in protein quality control has been well demonstrated during oxidative stress where CMA allows targeting of only oxidized forms of the proteins without affecting functional ones.

Pathophysiology of Autophagy

The multiplicity of functions attributed to the autophagic process speaks to their already proven physiological relevance. Although studies in unicellular models and mammalian cells in culture have been key to understanding the interplay of autophagy with other cellular processes, importance of proper autophagy functioning at the whole organism level has been inferred from genetically modified mouse models. Animals with whole body blockage of macroautophagy are born, but die few hours after birth, due to their inability to adapt to the starvation period that precedes lactancy (Kuma *et al.*, 2004). Tissue-specific conditional knockout mouse models for specific Atg have offered a better understanding of organ-specific functions of macroautophagy. Thus, complete loss of macroautophagy in the liver leads to accumulation of damaged organelles, lipids (steatosis), and liver dysfunction, whereas blockage of macroautophagy in specific areas of the brain leads to early neurodegeneration with accumulation of intraneuronal ubiquitinated protein aggregates that end killing neuronal cells (Komatsu *et al.*, 2005; Hara *et al.*, 2006). Blockage of autophagy in cardiomyocytes, when inflicted early in life, is well tolerated and defective cardiac function can only be elicited upon stress. However, blockage of autophagy in the heart of an adult animal progresses rapidly to heart failure (Nakai *et al.*, 2007). These regulatable models in which autophagy is eliminated at different times in life support the existence of early compensatory mechanisms in response to macroautophagy failure that are lost with aging. Elimination of macroautophagy genes in specific cell types has also connected autophagy to innate and acquired immunity, pancreatic beta-cell homeostasis and insulin secretion, muscle atrophy after immobilization, among others. These models are also revealing Atg functions independent of autophagy as in the case of Atg16L function in the Paneth cells of the intestine (Cadwell *et al.*, 2008). Overexpression of Atg proteins have been a successful approach to increase life span and health span in invertebrates. Although, to date, only one mouse model expressing a copy of Atg5 has been generated, analysis of life span in this model has revealed an encouraging increase (Pyo *et al.*, 2013). Future studies are needed to discriminate the specific functions of macroautophagy responsible for this life-span extension.

Studies on CMA are now also taking advantage of conditional transgenic mouse models with reduced CMA by selectively eliminating the exon that encodes for the L2A variant. Blockage of CMA in the liver leads to major metabolic aberrations and organism problems in metabolism of lipids and glycogen (Schneider *et al.*, 2014). Although none of these macromolecules can be degraded by CMA, many of the enzymes that regulate these metabolic pathways have been

validated as bona fide CMA substrates. Failure to undergo degradation by CMA in this model leads to an increase in their intracellular levels and the subsequent aberrant increase in function. Mouse models with defective CMA in T cells have shown that this autophagy pathway is required for the proteome changes that lead to the T-cell activation program (Valdor *et al.*, 2014). Failure to degrade inhibitory proteins of this program prevents T cell activation in the CMA-incompetent animals. Interestingly, studies in the T cell model have also revealed that the reduction in CMA with age in these cells is behind their functional failure in old organism (immuno-senescence) and that restoration of L2A in old T cells to the levels observed in young animals makes them functional again. Similar expression of an exogenous copy of L2A in mouse liver is sufficient to restore liver homeostasis and function.

Considering this extensive array of physiological functions attributed to autophagy, it was expected that malfunctioning of autophagy underlies the basis of severe human disorders (reviewed in detail elsewhere (Cuervo and Macian, 2014; Jiang and Mizushima, 2014; Lee, 2014; Martin *et al.*, 2015; Schneider and Cuervo, 2014)).

Concluding Remarks

The field of autophagy has undergone major expansion during the last two decades that has resulted in important paradigm shifts in this area. Although much remains to be discovered on the basis of selectivity, autophagy is no longer considered an in-bulk nonselective process. Instead, multiple autophagy variants have emerged and selective removal of specific cellular components is being identified. Furthermore, autophagy is no longer a process for destruction of unwanted cellular material, but rather a fine-tuned mechanism that contributes to maintain cellular homeostasis through selective degradation and recycling. Much is left to learn about the regulation of the different types of autophagy, the rules of their cellular co-existence, and about how their activities are coordinated in the context of disease to compensate one for another. It is precisely this connection between autophagy and human disease that has motivated the growing interest in the developing of efficient ways to manipulate the autophagic process with therapeutic purposes.

See also: Organelles: Structure and Function: At the Center of Autophagy: Autophagosomes

References

- Anguiano, J., Garner, T.P., Mahalingam, M., *et al.*, 2013. Chemical modulation of chaperone-mediated autophagy by retinoic acid derivatives. *Nature Chemical Biology* 9, 374–382.
- Arndt, V., Dick, N., Tawo, R., *et al.*, 2010. Chaperone-assisted selective autophagy is essential for muscle maintenance. *Current Biology* 20, 143–148.
- Bandyopadhyay, U., Kaushik, S., Varticovski, L., Cuervo, A.M., 2008. The chaperone-mediated autophagy receptor organizes in dynamic protein complexes at the lysosomal membrane. *Molecular and Cellular Biology* 28, 5747–5763.
- Bejarano, E., Yuste, A., Patel, B., *et al.*, 2014. Connexins modulate autophagosome biogenesis. *Nature Cell Biology* 16 (5), 401–414.
- Cadwell, K., Liu, J.Y., Brown, S.L., *et al.*, 2008. A key role for autophagy and the autophagy gene Atg16l1 in mouse and human intestinal Paneth cells. *Nature* 456, 259–263.
- Cuervo, A.M., Dice, J.F., 1996. A receptor for the selective uptake and degradation of proteins by lysosomes. *Science* 273, 501–503.
- Cuervo, A.M., Dice, J.F., Knecht, E., 1997. A population of rat liver lysosomes responsible for the selective uptake and degradation of cytosolic proteins. *Journal of Biological Chemistry* 272, 5606–5615.
- Cuervo, A.M., Macian, F., 2014. Autophagy and the immune function in aging. *Current Opinion in Immunology* 29, 97–104.
- Dice, J.F., 1990. Peptide sequences that target cytosolic proteins for lysosomal proteolysis. *Trends in Biochemical Sciences* 15, 305–309.
- Fengsrud, M., Erichsen, E.S., Berg, T.O., Raiborg, C., Seglen, P.O., 2000. Ultrastructural characterization of the delimiting membranes of isolated autophagosomes and amphisomes by freeze-fracture electron microscopy. *European Journal of Cell Biology* 79, 871–882.
- Geeraert, C., Ratier, A., Pfisterer, S.G., *et al.*, 2010. Starvation-induced hyperacetylation of tubulin is required for the stimulation of autophagy by nutrient deprivation. *Journal of Biological Chemistry* 285, 24184–24194.
- Green, D.R., Levine, B., 2014. To be or not to be? How selective autophagy and cell death govern cell fate. *Cell* 157, 65–75.
- Hara, T., Nakamura, K., Matsui, M., *et al.*, 2006. Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature* 441, 885–889.
- Hara, T., Takamura, A., Kishi, C., *et al.*, 2008. FIP200, a ULK-interacting protein, is required for autophagosome formation in mammalian cells. *Journal of Cell Biology* 181, 497–510.
- Ichimura, Y., Kirisako, T., Takao, T., *et al.*, 2000. A ubiquitin-like system mediates protein lipidation. *Nature* 408, 488–492.
- Jiang, P., Mizushima, N., 2014. Autophagy and human diseases. *Cell Research* 24, 69–79.
- Kanazawa, T., Taneike, I., Akaishi, R., *et al.*, 2004. Amino acids and insulin control autophagic proteolysis through different signaling pathways in relation to mTOR in isolated rat hepatocytes. *Journal of Biological Chemistry* 279, 8452–8459.
- Kaushik, S., Cuervo, A.M., 2012. Chaperone-mediated autophagy: A unique way to enter the lysosome world. *Trends in Cell Biology* 22, 407–417.
- Kenific, C.M., Debnath, J., 2015. Cellular and metabolic functions for autophagy in cancer cells. *Trends in Cell Biology* 25, 37–45.
- Komatsu, M., Waguri, S., Ueno, T., *et al.*, 2005. Impairment of starvation-induced and constitutive autophagy in Atg7-deficient mice. *Journal of Cell Biology* 169, 425–434.
- Kotoulas, O.B., Kalamidas, S.A., Kondomerkos, D.J., 2006. Glycogen autophagy in glucose homeostasis. *Pathology, Research and Practice* 202, 631–638.
- Kuma, A., Hatano, M., Matsui, M., *et al.*, 2004. The role of autophagy during the early neonatal starvation period. *Nature* 432, 1032–1036.
- Lee, M.S., 2014. Role of islet beta cell autophagy in the pathogenesis of diabetes. *Trends in Endocrinology & Metabolism* 25 (12), 620–627.
- Levine, B., 2005. Eating oneself and uninvited guests: Autophagy-related pathways in cellular defense. *Cell* 120, 159–162.
- Marino, G., Niso-Santano, M., Baehrecke, E.H., Kroemer, G., 2014. Self-consumption: The interplay of autophagy and apoptosis. *Nature Reviews Molecular Cell Biology* 15, 81–94.
- Martin, D.D., Ladha, S., Ehrnhoefer, D.E., Hayden, M.R., 2015. Autophagy in Huntington disease and Huntingtin in autophagy. *Trends in Neurosciences* 38, 24–36.
- Martina, J.A., Diab, H.I., Lishu, L., *et al.*, 2014. The nutrient-responsive transcription factor TFE3 promotes autophagy, lysosomal biogenesis, and clearance of cellular debris. *Science Signaling* 7 (309), ra9.
- Mizushima, N., Levine, B., Cuervo, A.M., Klionsky, D.J., 2008. Autophagy fights disease through cellular self-digestion. *Nature* 451, 1069–1075.
- Moreau, K., Ravikumar, B., Renna, M., Puri, C., Rubinstein, D.C., 2011. Autophagosome precursor maturation requires homotypic fusion. *Cell* 146, 303–317.
- Nakai, A., Yamaguchi, O., Takeda, T., *et al.*, 2007. The role of autophagy in cardiomyocytes in the basal state and in response to hemodynamic stress. *Nature Medicine* 13, 619–624.
- Novak, I., Dikic, I., 2011. Autophagy receptors in developmental clearance of mitochondria. *Autophagy* 7, 301–303.
- Ohsumi, Y., 2014. Historical landmarks of autophagy research. *Cell Research* 24, 9–23.
- Pyo, J.O., Yoo, S.M., Ahn, H.H., *et al.*, 2013. Overexpression of Atg5 in mice activates autophagy and extends lifespan. *Nature Communications* 4.2300.

- Ravikumar, B., Moreau, K., Jahreiss, L., Puri, C., Rubinshtein, D., 2010. Plasma membrane contributes to the formation of pre-autophagosomal structures. *Nature Cell Biology* 12, 747–757.
- Roberts, P., Moshitch-Moshkovitz, S., Kvam, E., *et al.*, 2003. Piecemeal microautophagy of nucleus in *Saccharomyces cerevisiae*. *Molecular Biology of the Cell* 14, 129–141.
- Sahu, R., Kaushik, S., Cannizzo, E., *et al.*, 2011. Microautophagy of cytosolic proteins by late endosomes. *Developmental Cell* 20, 131–139.
- Sakai, Y., Koller, A., Rangell, L., Keller, G., Subramani, S., 1998. Peroxisome degradation by microautophagy in *Pichia pastoris*. Identification of specific steps and morphological intermediates. *Journal of Cell Biology* 141, 625–636.
- Schneider, J.L., Cuervo, A.M., 2014. Autophagy and human disease: Emerging themes. *Current Opinion in Genetics and Development* 26, 16–23.
- Schneider, J.L., Suh, Y., Cuervo, A.M., 2014. Deficient chaperone-mediated autophagy in liver leads to metabolic dysregulation. *Cell Metabolism* 20, 417–432.
- Sekito, T., Fujiki, Y., Ohsumi, Y., Kakinuma, Y., 2008. Novel families of vacuolar amino acid transporters. *International Union of Biochemistry and Molecular Biology Life* 60, 519–525.
- Settembre, C., Di Malta, C., Polito, V.A., *et al.*, 2011. TFEB links autophagy to lysosomal biogenesis. *Science* 332, 1429–1433.
- Singh, R., Cuervo, A.M., 2011. Autophagy in the cellular energetic balance. *Cell Metabolism* 13, 495–504.
- Singh, R., Kaushik, S., Wang, Y., *et al.*, 2009. Autophagy regulates lipid metabolism. *Nature* 458, 1131–1135.
- Stolz, A., Ernst, A., Dikic, I., 2014. Cargo recognition and trafficking in selective autophagy. *Nature Cell Biology* 16, 495–501.
- Valdor, R., Mocholi, E., Botbol, Y., *et al.*, 2014. Chaperone-mediated autophagy regulates T cell responses through targeted degradation of negative regulators of T cell activation. *Nature Immunology* 15, 1046–1054.
- de Waal, E.J., Vreeling-Sindelarova, H., Schellens, J.P., Houtkooper, J.M., James, J., 1986. Quantitative changes in the lysosomal vacuolar system of rat hepatocytes during short-term starvation. A morphometric analysis with special reference to macro- and microautophagy. *Cell and Tissue Research* 243, 641–648.
- Yang, Z., Klionsky, D.J., 2010. Mammalian autophagy: Core molecular machinery and signaling regulation. *Current Opinion in Cell Biology* 22, 124–131.
- Yu, L., McPhee, C.K., Zheng, L., *et al.*, 2010. Termination of autophagy and reformation of lysosomes regulated by mTOR. *Nature* 465, 942–946.

CELL DIVISION/DEATH: CELL CYCLE

Contents

Interplay between Oncogenes and Tumor Suppressor Genes in Human Disease
Cyclins and Cyclin-Dependent Kinases
The Restriction Point
CDK Inhibitors in Normal and Malignant Cells
The *INK4a/ARF* Locus
S Phase
Traveling through Mitosis with the Chromosomal Passenger Complex
Mitosis in Animal Cells
Regulating Cytokinesis
Regulation of the p53 Pathway
Cellular Senescence

Interplay between Oncogenes and Tumor Suppressor Genes in Human Disease

SJ Parsons, University of Virginia Health System, Charlottesville, VA, USA

JO DaSilva, Regeneron Pharmaceuticals Inc., Tarrytown, NY, USA

J Thomas Parsons, University of Virginia Health System, Charlottesville, VA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

AKT Serine-threonine kinase activated by PI3K.

AMPK Serine-threonine kinase activated by AMP.

BRCA1,2 Breast cancer tumor suppressor genes 1 and 2.

ERK Serine-threonine kinase activated by MEK.

FYN Cytoplasmic tyrosine kinase of the SRC family.

LKB Serine-threonine kinase, enhancer of AMPK activity.

MEK Multi-specificity kinase activated by RAF.

mTOR Serine-threonine kinase indirectly activated by AKT.

MYC Transcription factor.

NF1 GAP for RAS-GTP.

PTEN Lipid (PI-3) phosphatase.

PTP1B Cytosolic protein tyrosine phosphatase.

PTPN11 RTK phosphatase.

RAF Serine-threonine kinase activated by RAS-GTP.

RAS Small guanine nucleotide binding protein.

RB Retinoblastoma tumor suppressor – transcriptional regulator.

SRC Cytoplasmic tyrosine kinase.

TP53 Tumor suppressor and transcription factor.

Introduction: Scope and Intent of Article

Over the past five decades, the concept that mutation of discrete cellular genes can drive the formation and progression of human cancers has evolved. These genes, termed oncogenes and tumor suppressor genes, constitute a defining paradigm for our understanding of cancer biology. The elucidation of their function in both normal and malignant cell physiology has provided a rational platform for the development of therapeutic strategies to treat the many forms of the disease. This article provides a brief overview of (1) the role of oncogenes and tumor suppressor genes in human cancers, (2) how the functional interaction between proteins encoded by oncogenes and tumor suppressor genes alters normal cellular

processes such as proliferation, metabolism/energetic, and cell survival, and (3) the role of oncogenes and tumor suppressor genes in diseases other than cancer, namely human developmental syndromes.

Oncogenes and Tumor Suppressor Genes and Human Cancer

During the first half of the twentieth century, two parallel tracks of research emerged to provide evidence for mutations of defined genes as major causes of cancer: one was the identification of animal viruses that induced rapid formation of cancers in chickens and rodents (retroviruses), and the other

was the discovery of chemical compounds that caused mutations in cellular genes (carcinogens), which would in turn promote cancer formation. In the 1970s these two lines of research merged, when it was discovered that tumor-forming retroviruses carried mutated genes (oncogenes) that were derived from normal cellular genes (proto-oncogenes) (Bishop, 1983, 1985), and that human cancers contained mutated proto-oncogenes that arose in the absence of any known viral etiology (Barbacid, 1987). These findings gave rise to the concepts that cancers can derive from multiple mechanisms and by mutation of one or more cellular genes.

As cancer cell biologists started to catalogue and compare the properties of normal and cancer cells, it became clear that cancer cells exhibited a set of cellular and physical properties, known as the 'transformed' phenotype, that were distinct from normal cells (e.g., altered morphology, unrestricted cell growth, altered metabolism, and multiple and varied chromosomal aberrations). Surprisingly, it was observed that if one 'fused' a normal cell with a cancer cell, the resultant hybrid cell containing a complete complement of host and donor chromosomes exhibited a 'normal' cellular phenotype (Harris *et al.*, 1969; Stanbridge, 1976). Acquisition of the 'transformed' phenotype required the loss of certain cellular chromosomes from the normal host, leading to the idea that certain normal cellular genes could 'suppress' the tumorigenic properties of cancer cells. Hence the concept of tumor suppressor genes emerged in the cancer literature.

Identification of Specific Oncogenes and Tumor Suppressor Genes in Human Cancers

Oncogenes

'Genes that, when activated by mutation, increase the selective growth advantage of the cell in which they reside.'

One of the first oncogenes to be discovered is harbored in the chicken retrovirus, Rous sarcoma virus, that was first described by Peyton Rous in the early 1900s (Martin, 2004). This virus causes transformation of avian and rodent cells in cell culture and induces tumors in birds and rodents. The oncogenic powers of Rous sarcoma virus reside in the viral gene denoted SRC, which is not truly a viral gene but a mutated form of a cellular gene denoted c-SRC, present in all vertebrate species (Stehelin *et al.*, 1976). This discovery led to the hypotheses that retroviruses can 'capture' and mutate cellular genes during their replication in host cells and that expression of these mutated cellular genes during subsequent rounds of host cell infection is sufficient to induce transformation of cells and tumor formation. This discovery led immediately to the realization that expression of other cellular genes altered by carcinogens or X-rays may mimic the effects of virus infection and lead to cancer formation.

Evidence that human cancers contain oncogenes came from observations that introduction of DNA purified from human cancer cells induced the transformation of normal rodent cells (Krontiris and Cooper, 1981; Perucho *et al.*, 1981; Shih and Weinberg, 1982). The molecular characterization of one of the genes responsible for this transformation unveiled a mutated version of the cellular RAS gene, which was found to encode a protein that exhibited constitutive enzymatic activity,

providing the first clue that oncogenes, of either viral or cellular origin exhibit unregulated activities (see below).

Tumor suppressor genes

'Genes that, when inactivated by mutation, increase the selective growth advantage of the cell in which they reside.'

In human cancers, inactivation of tumor suppressor genes occurs via two prominent mechanisms, loss of a defined region of a chromosome leading to complete loss of protein expression or expression of truncated forms of the protein, or mutations within a tumor suppressor gene altering protein function (Hinds and Weinberg, 1994; Weinberg, 1991). Direct evidence for the former in human cancer emanated from the study of retinoblastoma, a rare, inheritable, childhood tumor of the eye. Genetic studies showed that tumors arising in afflicted children either lack the RB gene or carry a mutated form of it. The RB gene encodes a transcription regulator (pRB) involved in normal growth control (see below) (Hollingsworth *et al.*, 1993). Subsequent genetic studies have identified a number of inheritable cancers that are linked to the loss of function of defined cellular genes (e.g., Wilm's tumor and the NF1 gene, familial adenomatous polyposis and the APC gene) (Cichowski and Jacks, 2001; Fearon and Vogelstein, 1990).

The gene encoding the p53 protein, TP53, is an example of how mutations alter normal protein function, leading to tumor initiation and progression (Hinds and Weinberg, 1994; Rivlin *et al.*, 2011; Vousden and Prives, 2009). TP53 is a transcription factor that regulates a cell's response to stress by inducing cell cycle arrest, DNA repair, or apoptosis, or by regulating cell metabolism. The importance of p53 in human cancers is underscored by the fact that a very large number of human cancers exhibit TP53 mutations and by the finding that the Li-Fraumeni syndrome, a rare cancer predisposition syndrome, is characterized by germline mutations in TP53. Interestingly, studies of rodent and human oncogenic DNA viruses have revealed that normal p53 function is neutralized in virus-infected cells by binding to discrete virus-encoded proteins. For example, the binding of the human papilloma virus-encoded protein, E6, to p53 inhibits p53 activity and is essential for human papilloma virus transformation of human cells (Moody and Laimins, 2010).

Number of mutations in a human cancer

In the past 10 years, genome-wide DNA sequencing has cataloged millions of mutations that arise in different human cancers. For example, solid tumors of the colon, breast, brain, or pancreas display somatic mutations that are predicted to alter the structure of as many as 60–70 proteins (Vogelstein *et al.*, 2013; Wood *et al.*, 2007). Interestingly, some tumors, such as melanomas and lung cancers, contain significantly more somatic mutations (approximately 200), consistent with the well-documented roles of the carcinogens, UV light, and cigarette smoke, in the etiology of these cancers.

How do we know which mutations are important for development and progression of a given cancer? It has been suggested that mutated genes which confer selective growth to the cancer are referred to as 'driver' genes, whereas mutations that do not affect function are referred to as 'passenger' mutations. An analysis of more than 18 000 mutated genes

containing over 400 000 subtle mutations in over 3000 tumors has defined 125 'driver' genes (Vogelstein *et al.*, 2013). Of these, 54 appear to function as oncogenes and 71 as tumor suppressor genes. Interestingly, a large number of these genes were previously identified as oncogenes or tumor suppressor genes using more classical approaches.

Functional Roles of Key Oncogene and Tumor Suppressor Gene Products

Oncogenes and tumor suppressor genes regulate the balance between cell proliferation, metabolism and energetics, and cell survival. They do this by encoding proteins that transmit signals from the extracellular milieu, which in turn stimulates the transcriptional apparatus of the cells. The changes in gene expression mediated by such intracellular signals trigger the synthesis of a host of cellular proteins that drive cell cycle progression, modulate the metabolic and energetic state of the cell, and control cell death. Some of the most important pathways regulated by oncogenes and tumor suppressors are briefly described below.

Transmission of extracellular cues

The ability of cells to sense and respond to their environmental milieu is dependent on a wide variety of transmembrane cell surface receptors. One class of such receptors, the receptor tyrosine kinases (RTKs), specifically engages ligands, such as polypeptide growth factors, cytokines, and hormones, with high affinity and transmits signals dependent on this binding to the inside of cells, thus initiating a cascade of biochemical reactions (pathways) that control the cell's response to its environment (Hubbard and Miller, 2007; Figure 1(a)). In the human genome over 50 unique genes encode RTKs, including receptors for the epidermal growth factor (EGF), platelet-derived growth factor, fibroblast growth factor, vascular endothelial growth factor (VEGF), hepatocyte growth factor/scatter factor, and insulin/insulin-like growth factors (IGF). Because of their central roles in promoting cell proliferation, many RTKs have been implicated as etiological agents in various cancers, either through receptor gain-of-function mutations (EGFRs) or through receptor/ligand over-expression (EGFR, VEGFR, IGF) (Blume-Jensen and Hunter, 2001). In cancer cells, the mutational activation of RTKs and/or their over-expression lead to elevated and constitutive downstream signaling, resulting in uncontrolled cell proliferation (Hubbard and Miller, 2007).

The RAS pathway

The RAS pathway (Figure 1(a)) is a major intracellular signaling pathway in normal and cancerous cells of adult organisms and developing embryos alike (Bos, 1989). It is activated downstream of RTKs by guanine nucleotide exchange proteins (GEFs), which mediate the exchange of guanine nucleotide diphosphate (GDP) for guanine nucleotide triphosphate (GTP) on the small G protein, RAS (Bos *et al.*, 2007). Once activated by GTP, RAS stimulates the catalytic activity of the serine/threonine protein kinase RAF (Dhillon *et al.*, 2007). GAPs (GTPase-activating proteins) promote the return of RAS to the inactive GDP-bound state (McCormick, 1989).

Activated RAF leads to the stimulation of a multi-specificity kinase, MEK, which in turn phosphorylates and activates the serine/threonine kinase, ERK. Activated ERK translocates to the nucleus, where it catalyzes the phosphorylation of a myriad of nuclear substrates, including transcription factors, transcriptional regulatory proteins, and other substrates involved in DNA replication and cell division. RAS proteins were among the first oncogenes to be identified, and mutated RAS is found in about 15% of human cancer (Davies *et al.*, 2002). Mutation of RAF occurs in 60–70% of malignant melanomas and at a lower frequency in a wide range of other human cancers (Davies *et al.*, 2002). Interestingly few, if any, mutations in MEK or ERK have been identified in human cancers.

The PI3K pathway

A second important signaling pathway coupled to RTK activation is the phosphatidylinositol-3-kinase (PI3K) pathway (Cantley, 2002; Figure 1(b)). PI3Ks are lipid kinases that play central roles in cell cycle progression, apoptosis, DNA repair, and cellular metabolism. PI3Ks transduce signals from cell surface receptors by generating phosphorylated phosphatidylinositols, which in turn activate effector kinase pathways, principally those involving AKT and mTOR (Figure 1). These pathways play key roles in promoting cancer cell survival and proliferation (Cantley, 2002; Osaki *et al.*, 2004). The activity of the PI3K pathway in normal cells is tightly regulated by the phosphatidylinositol-3,4,5-trisphosphate 3-phosphatase, PTEN (phosphatase and tensin homolog deleted from chromosome 10). PTEN dephosphorylates phosphatidylinositol-3,4,5-trisphosphate, thereby negatively regulating signaling through all the downstream targets of PI3K, including AKT and mTOR (Chalhoub and Baker, 2009; Di Cristofano and Pandolfi, 2000; Osaki *et al.*, 2004). In human cancers the deregulation of the PI3K signaling pathway is linked to the development of approximately one-third of human cancers (Samuels and Waldman, 2010). Somatic mutations in the PTEN gene are among the most prevalent of these genetic changes (Yin and Shen, 2008).

Transcriptional regulation

Oncogenes encode a large repertoire of transcription factors, which can be activated by chromosomal rearrangements, gene fusions, gene amplifications, and to a lesser extent somatic mutations. A prime example is the MYC oncogene (Dang, 2012). Altered expression of the MYC gene occurs in over half of human cancers (Vita and Henriksson, 2006). In many cancers the level of MYC becomes elevated as a consequence of chromosomal rearrangements that place the MYC gene under the control of a constitutively activated promoter, such as an immunoglobulin gene promoter (Burkett's lymphoma). MYC can also be the direct target of signal transduction pathways activated by RTKs in response to nutrients, growth factors, and mitogenic stimuli. ERK, the downstream effector of the RAS/RAF/MEK/ERK pathway, stabilizes MYC protein by phosphorylation (Gustafson and Weiss, 2010), thus preventing its turnover within the nucleus. Thus activation of this pathway either by RTK activation or mutation of RAS or RAF can lead to the stabilization of MYC and increased transcription of MYC target genes (Figure 1(a)).

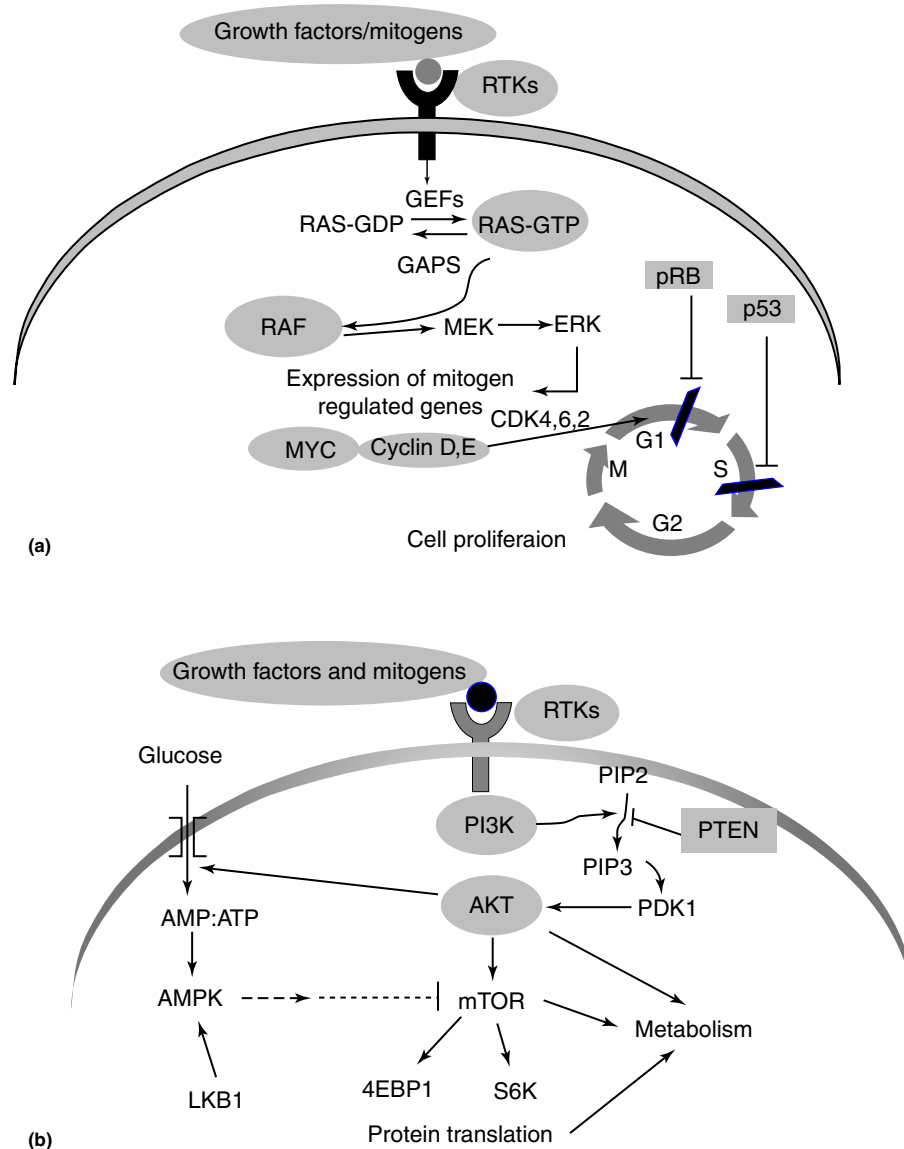


Figure 1 Key intracellular signaling pathways regulated by oncogenes and tumor suppressor genes. (a) Growth factor/mitogen binding to cell surface receptors (RTKs) leads to intracellular activation of the RAS-RAF-MEK-ERK pathway which culminates in transcriptional stimulation of genes required for cells to advance through the cell cycle. Progression through the cell cycle is dependent upon the activities of cyclin-CDKs which are needed for movement through 'restriction' points. Many pathways that stimulate the cell to divide are populated with oncogenes whose actions are counteracted by tumor suppressors (notably pRB and p53) (see text for details). The balance between these two classes of cell cycle regulators determines whether a cell completes division. (b) Growth factor/mitogen binding to RTKs also stimulates the PI3K pathway, which leads to the activation of AKT-mTOR and stimulation of protein synthesis. AKT and mTOR also play major roles in metabolism by regulating intracellular levels of glucose and AMP/ATP ratios. Filled ovals denote oncogenes, and filled rectangles represent tumor suppressors.

Interplay between Oncogenes and Tumor Suppressor Genes in Human Cancer

In cancer, the activation of oncogenes and compromised function of tumor suppressor genes alter the finely tuned balance between positive and negative signals that regulate cell proliferation, metabolism, and cell survival. This section describes key examples of how proteins encoded by oncogenes and tumor suppressor genes functionally interact with one another, the effects these interactions have on normal

physiological processes of the cell, and how the changes in such processes provide a growth advantage to the cancer cell.

Oncogenes, Tumor Suppressor Genes, and Cell Cycle Regulation

Cell proliferation is a highly controlled process that involves passage through the four main phases of the cell cycle (Ho, 2009): (1) G1, or interphase, during which cellular constituents (other than DNA) are increased in preparation for the

generation of two new daughter cells, (2) S phase, DNA synthesis, (3) G₂, the second interphase stage in which the integrity of the newly synthesized DNA is ascertained, among other things, and (4) M phase (mitosis) in which equal distribution of cellular components into two daughter cells and cell division takes place. To pass from one phase to another, the cell must traverse one or more checkpoints (Weinberg, 2007). Not unexpectedly, these checkpoints are frequent sites of regulation by proteins encoded by oncogenes and tumor suppressor genes. If cells exit the cell cycle, they assume a resting or quiescent state, termed G₀. The majority of cells in adult organisms, cells that form the fully differentiated organs of the human body, reside in the G₀ state. When injury occurs and/or cell replacement is needed, growth factors (such as those that activate RTKs) initiate the exit of cells from G₀ into G₁ and their continued presence is required for passage through the first checkpoint within G₁ (termed the restriction point).

Ligand activation of growth factor receptors stimulates intracellular signaling cascades (Figure 1(a)), which culminate in new gene transcription. Among the genes activated by this process is the cyclin D gene family, which is expressed in early G₁. Cyclin D binds and activates the cyclin-dependent protein kinases (cdks), 4 and 6, which phosphorylate pRB, leading to the release of the transcription factor, E2F, and progression through the cell cycle. pRB is a transcriptional repressor that blocks progression through early to mid-G₁ and must be inactivated by phosphorylation for passage through the mid-G₁ restriction point. Specifically, in early G₁, unphosphorylated pRB binds to and inhibits the transcription factor, E2F. Phosphorylation of pRB by cyclinD/cdk4/cdk6 leads to the release of E2F and renders E2F active transcriptionally. E2F, in turn, stimulates production of cyclin E, which binds and activates cdk2. The cyclin E/cdk2 complex completes inactivation of Rb by fully phosphorylating it, thereby allowing the cell to complete G₁ and enter S phase.

In cancer, pRB is inactivated by deletion/mutation or by DNA tumor viruses, such as papillomaviruses, which produce viral proteins that bind and sequester pRB (Moody and Laimins, 2010), leading to the release of E2F and activation of the transcriptional machinery. Without pRB functioning as the gate-keeper of the early checkpoint, cancer cells are free to divide without submitting themselves to quality controls throughout G₁. Such a situation fosters accumulated mistakes in DNA replication (S phase), uncontrolled growth, and tumor development. Although pRB itself is rarely mutated in human cancers, its pathway is dysregulated in a large majority of patients, stressing the importance of the pathway to oncogenesis. One of the most common alterations is overexpression of Cyclin D, cdk 4 and cdk 6. Furthermore, growth factor pathways are frequently hyperactivated in the same cells that harbor pRB pathway disruptions, resulting in amplified dysregulation of G₁ transit.

The transcription factor p53 monitors the strength and/or fidelity of events related to cell cycle progression, including strength of oncogenic signaling, hypoxia, DNA damage, and transcription (Weinberg, 2007). When abnormalities are sensed, a variety of pathways can trigger signals that increase levels of p53 (Harris, 1996). In turn, p53 activates pathways that regulate whether the cell should arrest progression while

repairing cellular damage (DNA) or undergo programmed cell death or apoptosis. If the cell is able to repair the abnormality sensed by p53, it can reenter and complete the cell cycle or enter a state of dormancy/senescence from which it does not return. When damage is too great for the cell to survive, p53 initiates apoptosis.

The TP53 gene is the most commonly mutated tumor suppressor in human cancers, a testament to its importance. Loss of p53 function results in continual cycling of cells even in the presence of metabolic imbalances, DNA damage, and transcriptional irregularities, giving rise in many instances to cells that accumulate mutations in key cell signaling pathways, that harbor chromosomal abnormalities and that have altered their metabolic equilibrium – all hallmarks of malignancy.

Oncogenes, Tumor Suppressor Genes and Cell Metabolism

Fully differentiated, nondividing cells have low nutrient requirements and derive the majority of their ATP from catabolism of glucose to pyruvate, which is further processed through the TCA cycle and the electron transport chain (oxidative phosphorylation) (Vander Heiden *et al.*, 2009). In contrast, rapidly dividing cells (normal or cancerous) have high nutrient requirements and an increased need for synthesis of macromolecules (proteins, RNA and DNA, carbohydrates, and complex lipids) to support cell division (Figure 2(a)). To satisfy these needs, rapidly dividing cells take up more glucose and glutamine than nondividing cells, and process them through glycolysis and glutamine catabolism, thereby generating both ATP and macromolecules, and reducing their dependence on oxidative phosphorylation. Cancer cells accomplish this conversion from oxidative phosphorylation to glycolysis even in the presence of high oxygen tension, termed aerobic glycolysis, a hallmark of the Warburg effect. Oncogenes and tumor suppressor genes antagonistically regulate this conversion through altered production of catabolic and anabolic enzymes, posttranslational modifications, and protein–protein interactions (Dang and Semenza, 1999; Iurlaro *et al.*, 2014). We will briefly discuss three examples of such genes, namely, PI3K, MYC, and p53.

The PI3K pathway

AKT, a major effector of PI3K (Figure 1(b)), is an important stimulator of the tumor glycolytic phenotype. Activation of AKT increases the expression and translocation of glucose transporters, phosphorylation of key glycolytic enzymes, and activation of mTOR through phosphorylation of negative inhibitors of mTOR (Cairns *et al.*, 2011). The serine/threonine kinase mTOR, in turn directly stimulates mRNA translation and new protein synthesis by phosphorylating/activating S6 kinase and phosphorylating/inactivating 4EBP1. Through an indirect transcriptional mechanism, mTOR also increases levels of HIF1 (Denko, 2008), a key sensor of oxygen levels and a transcription factor whose levels are augmented under hypoxic conditions, conditions frequently observed in centers of human tumors. HIF1 increases synthesis of multiple glycolytic enzymes, thereby significantly contributing to the conversion from oxidative phosphorylation to glycolysis.

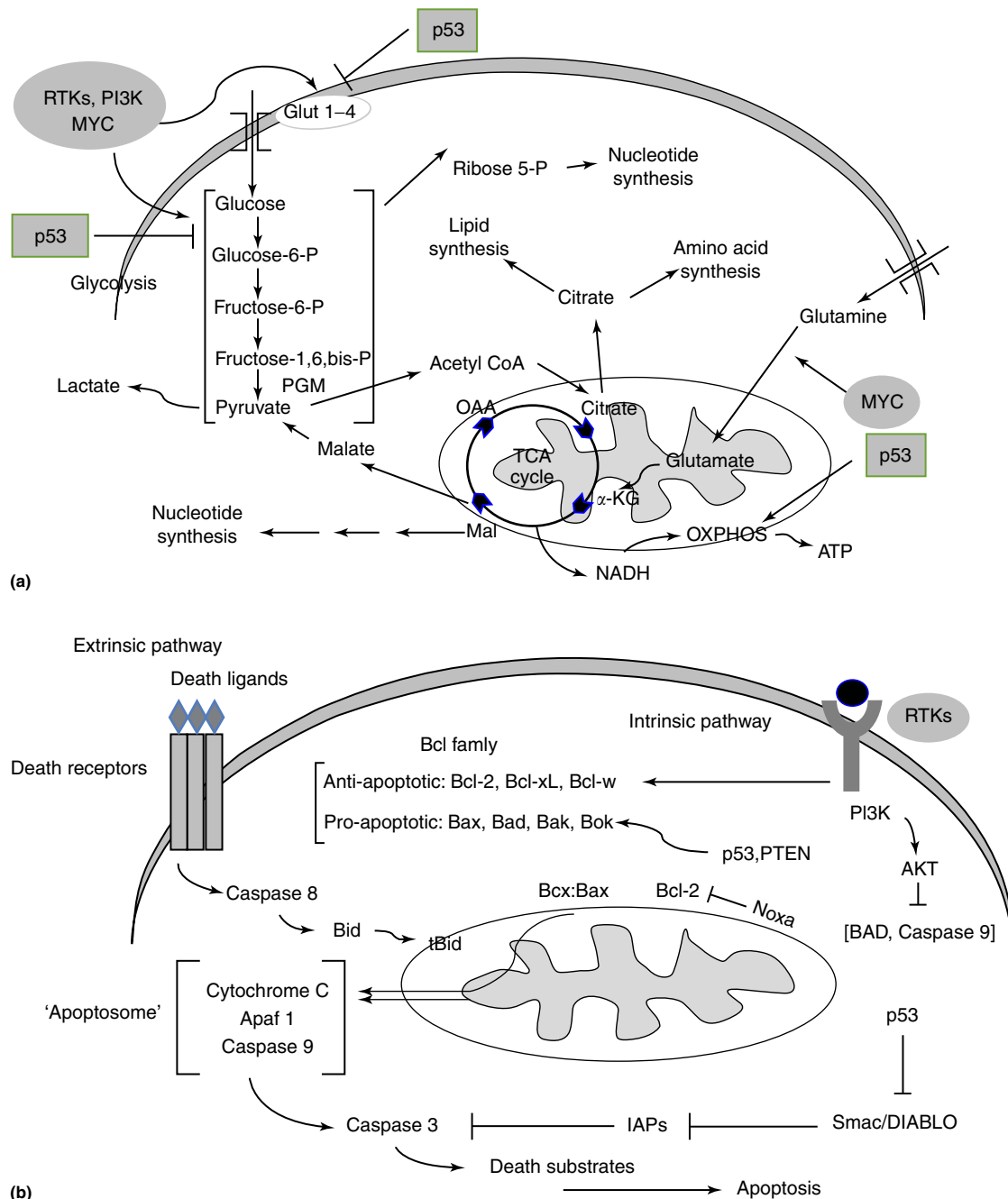


Figure 2 Oncogenes and tumor suppressor genes in metabolism and cell survival/apoptosis. (a) Glycolysis and oxidative phosphorylation (OXPHOS) are major pathways for energy generation, principally via the utilization of glucose and glutamine. The major points of regulation by oncogenes (ovals) and tumor suppressors (rectangles) are indicated (see text for details). (b) Cell survival is mediated by two major pathways, the extrinsic pathway, which is stimulated by extracellular death ligands, and the intrinsic pathway, which is activated by intracellular cues that regulate the balance between pro- and anti-apoptotic Bcl-2 family proteins. In normal cells, RTK signaling upregulates anti-apoptotic Bcl-2 family proteins, whereas the tumor suppressors p53 and pTEN upregulate the activity of pro-apoptotic Bcl-2 family members. In cancer cells, the normal control of these pathways is perturbed (see text for details). Filled ovals denote oncogenes, and filled rectangles represent tumor suppressors.

mTOR also plays a central role in the integration of growth factor signals and nutrient availability through regulation by adenosine monophosphate activated kinase (AMPK), a sensor of adenosine triphosphate (ATP) concentrations available to the cell (Figure 1(b)). A high AMP:ATP (Hardie, 2007) ratio activates AMPK, which then inactivates mTOR, thereby shifting

from glycolysis to ATP production through the oxidative phosphorylation pathway. AMPK activity itself is positively regulated by LKB1, a tumor suppressor mutated in liver and lung cancers. LKB1 is a serine-threonine kinase that phosphorylates AMPK to enhance its negative effect on mTOR (Shaw et al., 2004). Thus the LKB1/AMPK axis functions

in opposition to PI-3 kinase to regulate cellular energetics. Mutation of LKB1 and high levels of oncogenic signaling can override the negative control of LKB1/AMPK through hyperactivation of PI3K, permitting cells to divide under unfavorable energetic conditions.

MYC

The MYC oncogene regulates cell metabolism in multiple ways, including increased expression of enzymes involved in glutamine metabolism (Dang and Semenza, 1999; Figure 2(a)). MYC also collaborates transcriptionally with HIF1 to elevate levels of glucose transporters and glycolytic enzymes, thereby, reinforcing glycolysis, particularly under hypoxic conditions. Under normoxic conditions, HIF1 protein levels are reduced by oxygen-dependent hydroxylation and ubiquitination, mediated in part by the tumor suppressor, von Hippel–Lindau (VHL) protein, a ubiquitin ligase. Thus, HIF1 is regulated positively by oncogenes such as tyrosine kinases and MYC and negatively by VHL (Kim and Dang, 2006).

TP53

The tumor suppressor p53 plays a major role in regulating metabolism largely through its transcriptional activity (Levine and Puzio-Kuter, 2010). In normal cells, p53 helps to maintain appropriate levels of signaling proteins and metabolic enzymes that favor oxidative phosphorylation and limit glycolysis (Figure 2(a)). In cancer cells that harbor p53 mutations, conversion from oxidative phosphorylation to the glycolytic phenotype is favored. Specifically, mutated p53 results in increased expression of glucose transporters, enhanced nucleotide biosynthesis, and reduced synthesis of cytochrome *c* oxidase and negative regulators of PI3K. Thus, the rigorous control p53 exercises over metabolism in normal cells (opposing growth factor action) is diminished in cancer cells by mutational inactivation, which permits conversion of the cell to a metabolic phenotype that favors rapid growth.

Oncogenes, Tumor Suppressor Genes, and Apoptosis

Oncogenes and tumor suppressor genes play major roles in determining whether cells survive or die by regulating the balance between pro-survival and apoptosis-inducing pathways, regulated by the Bcl-2 family (Adams and Cory, 2007). This large family of genes (total of 25) can be either pro-apoptotic (e.g., Bax, Bad, Bak, and Bok) or anti-apoptotic (e.g., Bcl-2, Bcl-xL, and Bcl-w) and together regulate the major pathways (extrinsic and intrinsic) that activate apoptosis (Figure 2(b)). Members of this family of proteins associate with one another at the mitochondrial membrane and regulate the organelle's redox potential and membrane permeability. An overabundance of pro-apoptotic proteins at the membrane results in pore formation, release of cytochrome *c* into the cytoplasm, formation of the apoptosome, and activation of the apoptotic process. Pro-survival oncogenes, such as growth factors, their receptors, and components of their signaling pathways, increase levels of anti-apoptotic Bcl-2 family members, whereas tumor suppressors, such as p53, PTEN, and IGF-1 binding proteins, increase levels of pro-apoptotic family members.

Another function of pro-survival oncogenes is to activate the PI3K and NF κ B pro-survival pathways. PI3K activation leads to AKT phosphorylation and inactivation of the Bad, caspase 9, and I κ B pro-apoptotic proteins (Manning and Cantley, 2007), whereas NF κ B activation leads to enhanced expression of many pro-survival genes, most notably Bcl-2 (Cogswell *et al.*, 2000). PI3K and NF κ B effects are opposed by PTEN and I κ B, respectively. The tumor suppressor p53 can directly induce apoptosis if the cell has incurred too much damage to divide (Fridman and Lowe, 2003). This is accomplished by transcriptionally regulating many pro-apoptotic genes, such as increasing expression of the Fas death-inducing receptor, IGFBP3 (which sequesters IGF-1 and prevents it from binding its receptor), Bax (which oligomerizes and induces cytochrome *c* release from the mitochondria), and NOXA (which binds and inhibits the anti-apoptotic protein Bcl-2). P53 also inhibits the expression of pro-survival genes, such as Smac/DIABLO, which bind and inhibit inhibitors of apoptosis (IAPs). Finally, p53 can act by directly binding members of the Bcl-2 family to promote apoptosis (Haupt *et al.*, 2003). Loss of p53 function in cancer can lead to survival of cells that have sustained multiple mutations and damage, thereby permitting unregulated growth.

Role of Oncogenes and Tumor Suppressor Genes in Human Developmental Diseases

Sustained proliferative signaling resulting from activated oncogenes and inactivated tumor suppressor genes are fundamental traits of cancer. However, oncogenes and tumor suppressors also play important roles during embryonic development, as highlighted by the diverse set of human syndromes resulting from congenital mutations affecting not only the predisposition to neoplasms but also bone growth, neurodegeneration, and glucose homeostasis.

Developmental Syndromes of RAS-ERK and PI3K-AKT Dysregulation

Noonan/LEOPARD, neurofibromatosis type1, PTEN hamartomatous, and Proteus/Cowden syndromes are autosomally inherited familial syndromes that involve members of the RAS-ERK or PI3K-AKT signaling pathways (Figure 3(a)). Dysregulation of these pathways during embryonic development and neural crest differentiation results in common features, including craniofacial dysmorphism, cardiac malformations and cutaneous, musculoskeletal and ocular abnormalities, neurocognitive delay, and in some syndromes a predisposition to the development of cancer.

Noonan syndrome

Noonan syndrome (NS) is an autosomal dominant disorder that is characterized by postnatally reduced growth, distinctive craniofacial features, congenital cardiac defects, bleeding disorders, and variable cognitive deficits (Tartaglia *et al.*, 2011). Mutation of *PTPN11*, which encodes the non-receptor protein tyrosine phosphatase SHP2, accounts for the majority of cases. SHP2 is a negative regulator of RTKs and the RAS-ERK pathways, which when mutated and inactivated in NS patients

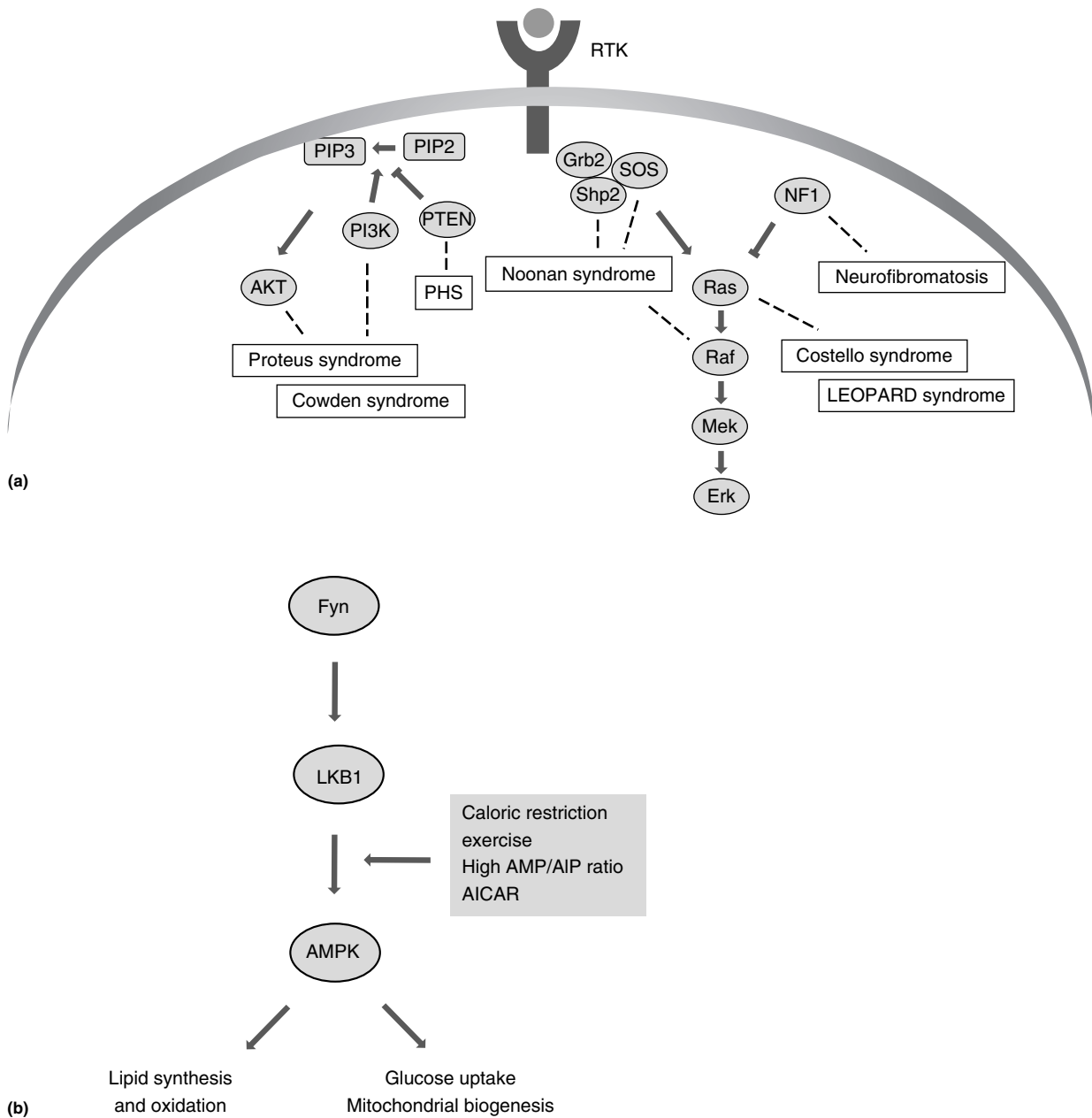


Figure 3 Oncogenes and tumor suppressor genes in developmental and metabolic disorders. (a) Developmental abnormalities associated with the RAS-ERK and PI3K-AKT pathways. Components of the RAS-ERK and PI3K-AKT pathways activated downstream of RTK are dysregulated in various human developmental syndromes. (b) Factors that affect the FYN-LKB1-AMPK pathway regulation of cellular metabolism and diabetes. The FYN-LKB1-AMPK pathway activity is increased by diverse factors, such as low circulating glucose, sedentary lifestyle, and AICAR (an analog of AMP and activator of AMPK) to regulate multiple metabolic processes. See text for details.

leads to increased signaling through the RAS-ERK pathway (Keilhack *et al.*, 2005). *SOS1* missense mutations are the second most common cause of NS occurring in approximately 13% of cases. *SOS1* is a RASGEF that becomes mutationally activated in NS, resulting in increased GTP-bound Ras and enhanced downstream signaling (Roberts *et al.*, 2007).

Neurofibromatosis

The main characteristics of Neurofibromatosis type1 (NF1) include cutaneous, subcutaneous and plexiform neurofibromas,

café-au-lait spots, and Lisch nodules in the iris. Changes of the skeletal system, vascular defects, learning disabilities, poor social skill, and predisposition for the development of malignant neoplasia are also observed (Friedman and Birch, 1997). NF1 is caused by heterozygous loss-of-function mutations or deletions of the *NF1* gene, encoding a GTPase-activating protein that functions as a negative regulator of RAS GTPase. Loss of NF1 results in hyperactivation of Ras signaling pathways, which contributes to the development of the defects observed in neurofibromatosis.

PTEN-opathies

Germline mutations of the tumor suppressor, PTEN, have been identified in patients with diverse clinical features ranging from overgrowth to malignancy that are collectively classified as PTEN hamartomatous syndromes (PHS). These include Cowden syndrome (CS) and Bannayan–Riley–Ruvalcaba syndrome (BRRS).

AKT1 and PIK3CA mutations in Proteus and Cowden syndrome

Approximately 11% of CS patients carry germline AKT1 or PI3K mutations. Cell lines generated from patients that have these mutations display increased phosphorylation of AKT1 and PIP3 levels, providing evidence that mutation of AKT1 or PI3K is a predisposition for CS (Orloff *et al.*, 2013). An activating AKT1 mutation was also found to cause Proteus syndrome (Lindhurst *et al.*, 2011), characterized by segmental overgrowth of the brain, bone, skin, and other tissues.

Overall, the relatively high prevalence of some of these disorders and the frequency of defects in RAS-ERK and PI3K-AKT signaling associated with them suggest that these pathways represent one of the most common events affecting the developmental process.

Oncogenes and Tumor Suppressor Genes in Neurodegenerative Diseases

In mammals, most neurons are generated before birth, but it is now well accepted that neurogenesis continues into adulthood. Accumulating evidence indicates that impaired neurogenesis is associated with dysregulation of oncogenic signaling pathways and tumor suppressor function leading to development of diseases, such as Alzheimer's and Parkinson's diseases. In these neurodegenerative disorders, one of the main etiological features is enhanced cell death, accompanied by dysregulation of the tumor suppressor, p53.

Parkinson's disease

In Parkinson's disease (PD) the main locus of degeneration is a small area in the ventral midbrain, the *pars compacta* of the substantia nigra that is mostly composed of dopaminergic neurons (Barzilai and Melamed, 2003). PD-affected substantia nigra harbors intracellular inclusions named Lewy bodies containing α -synuclein. Autosomal dominant inherited cases of PD are linked with mutations in α -synuclein that abrogate its ability to protect neuronal cells challenged with pro-apoptotic effectors by lowering p53 expression and transcriptional activity (da Costa *et al.*, 2000). Transgenic mice expressing α -synuclein with the pathogenic A53T mutation exhibit accumulation of p53 in the mitochondria and induction of cell death, suggesting a critical role of p53 in the development of the disease (Giaime *et al.*, 2006).

Several lines of evidence indicate that proteins associated with autosomal recessive cases of PD also functionally interact with p53. PD cases of recessive inheritance result from mutation in three genes encoding parkin, PINK-1, or DJ-1. Parkin physically interacts with the p53 promoter and represses p53 transcription (da Costa and Checler, 2010). Mutation of Parkin prevents repression, allowing elevated transcription of p53 and induction of apoptosis. Mutations in DJ-1 also result

in accumulation of p53, as well as p53's transcriptional target Bax (Fan *et al.*, 2008).

Alzheimer's disease

The pathology of Alzheimer's disease (AD) is characterized by two histological lesions, the senile plaques and the neurofibrillary tangles. The severity of these lesions correlates with the level of cognitive decline. Tangles result from the intracellular accumulation of hyperphosphorylated tau protein, while senile plaques are extracellular deposits mainly due to the aggregation of various hydrophobic amyloid- β peptides (Mandelkow, 1999; Masters *et al.*, 1985). That cell death in AD is p53-dependent is supported by several observations. APP-derived fragments can modulate p53 by directly transactivating its promoter. Additionally, oxidative DNA damage induces the extracellular translocation of APP, concomitantly with increased p53 mRNA levels. Finally, in both mouse models of AD and in AD brains, neurons with altered morphology display both APP and p53 accumulation (Ohya *et al.*, 2005). Comparison of p53 immunoreactivity in postmortem AD-affected brains revealed a consistent increase in p53 expression in the temporal cortex areas as compared to normal controls (Kitamura *et al.*, 1997). The increased expression of p53 coincides with increased DNA fragmentation and Fas expression, indicating that p53 can contribute to cell death in AD brains (de la Monte *et al.*, 1997).

The tumor suppressors p21, p27, PTEN, and BRCA1 have also been implicated in the pathology of neurodegenerative disease. BRCA1, for example, co-localizes with neurofibrillary tangles. Analysis of clinically normal, aged brain tissue also revealed less BRCA1 than in cases of AD. This finding suggests that neurofibrillary tangles of normal aging differ in origin from those present in neurodegenerative disease. In addition to the suppression of cell cycle, BRCA1 is implicated in the maintenance of genomic integrity. In cells with BRCA1 deletion, a loss of telomere repeats was observed, a feature present in degenerating neurons of AD (McPherson *et al.*, 2006). Studies of patients and animal models of PD have also demonstrated a direct role for PTEN in neurodegeneration associated with oxidative stress. Oxidative stress in rat hippocampal cells leads to mitochondrial accumulation of PTEN resulting in the induction of apoptosis. Conversely, knockdown of PTEN prevents neuronal apoptosis (Morris *et al.*, 2010).

Oncogenic signaling pathway activation leading to aberrant cell cycle entry of terminally differentiated neuronal cells has been proposed as a model of pathogenesis in neurodegenerative disorders. Cell cycle reentry driven by adenoviral-expression of c-MYC and RAS oncogenes in post-mitotic primary cortical neurons leads to tau protein phosphorylation and conformational changes similar to those observed in AD (McShea *et al.*, 2007). Evidence of cell cycle reentry is also present in mutant APP and tau transgenic mice, occurring prior to the development of pathological features such as neurofibrillary tangles and senile plaques (Lee *et al.*, 2009).

Oncogenes and Tumor Suppressor Genes in Diabetes

Type 2 diabetes mellitus (T2DM) is a metabolic disorder characterized by hyperglycemia and altered lipid metabolism

due to peripheral insulin resistance and defects of insulin secretion by pancreatic islet β -cells. The insulin receptor activates multiple downstream pathways to control energy homeostasis; however, the PI3K-AKT pathway is considered its major effector. The insulin-independent LKB1-AMPK pathway also contributes to metabolic control. Dysfunction of these pathways can lead to impaired glucose homeostasis and diabetes.

PI3K-AKT pathway

Genetic defects in loci encoding insulin pathway components, transcription factors or rate-limiting enzymes of glucose metabolism account for only 1–5% of diabetes cases (Permutt *et al.*, 2005). Nonetheless, the effects of the genetic alterations in insulin signaling downstream of PI3K-AKT has been extensively studied. For example, the amino acid substitution R409Q in the catalytic p85 α subunit of PI3K has been shown to compromise insulin-stimulated PI3K activity, while the mutation R274H within the AKT2 kinase domain results in severe autosomal dominant insulin resistance. This substitution greatly impairs AKT kinase activity and functions in a dominant-negative fashion to inhibit adipocyte differentiation *in vitro*. As is the case in humans, mice lacking AKT2 are insulin resistant, hyperglycemic, and hyperinsulinemic (George *et al.*, 2004). Conversely, mice lacking negative regulators of PI3K-AKT signaling exhibit improved glucose homeostasis. Both PTP1B (a phosphoserine-phosphothreonine phosphatase) deficiency and PTEN hemizygosity lead to improved glucose tolerance and insulin sensitivity in mice (Elchebly *et al.*, 1999; Wong *et al.*, 2007).

LKB1-AMPK pathway

Although some cases of obesity and T2DM result from genetic dysfunction, the increased incidence of these disorders is strongly correlated with sedentary lifestyle and over-nutrition. Dietary changes and regular exercise are the first therapies for the management of T2DM. However, when lifestyle modifications are insufficient to control serum glucose and lipid levels, drugs that increase insulin action and/or induce body energy consumption are used for treatment. Two commonly used drugs that potentiate insulin action, metformin and thiazolidinediones, have been shown to activate the LKB1-AMPK pathway (described above; Figure 3(b)).

AMPK activity is increased during exercise and could, at least in part, mediate the favorable effects of physical activity on insulin sensitivity, lipid and glucose utilization in skeletal muscle, adipose tissue, and liver (Ruderman *et al.*, 2003). Studies with pharmacological inhibitors of AMPK and in animal models harboring AMPK deletions indicate that AMPK activation positively regulates glucose uptake and metabolism. The serum glucose lowering effects of LKB-AMPK pathway activation occurs through CREB-dependent transcription of PGC- α and subsequent inhibition of the glucogenic enzymes, phosphoenolpyruvate carboxykinase and glucose-6-phosphatase (Koo *et al.*, 2005).

In addition to its role in glucose metabolism, the LKB1-AMPK pathway plays a critical role in the regulation of fatty acid oxidation in skeletal muscle and liver. AMPK increases fatty acid oxidation by phosphorylating the inhibitory site of acetyl-CoA-carboxylase, one of the main rate-controlling enzymes for the synthesis of malonyl-CoA, a potent inhibitor

of mitochondrial fatty acid oxidation. AMPK also represses pyruvate kinase and fatty acid synthase expression leading to decreased lipogenesis (Cook and Gamble, 1987; Leclerc *et al.*, 1998). These actions can contribute to the prevention of obesity and insulin resistance.

Although studies of LKB1-AMPK function in animal models of T2DM indicate that pathway activation may prevent disease onset, data collected to date from individuals with obesity and insulin resistance, with or without diabetes, have not detected aberrant AMPK activity in skeletal muscle of these patients (Hojlund *et al.*, 2004). These studies suggest that impaired insulin action on glycogen synthesis and lipid oxidation in skeletal muscle of these patients is not associated with altered AMPK expression or activity. However, it has also been reported that exercise-induced activation of AMPK is impaired in muscle of obese individuals with or without T2DM suggesting that activation above basal level could be effective in the treatment of diabetes and metabolic syndromes (Steinberg *et al.*, 2004).

FYN tyrosine kinase as a novel regulator of LKB1/AMPK

Recently, FYN, a non-RTK of the SRC family, was shown to play a role in insulin sensitivity and lipid utilization by regulating LKB1 function (Figure 3(b)). FYN knockout mice exhibit marked reductions in adiposity, fasting glucose and insulin levels, as well as improved insulin sensitivity (Bastie *et al.*, 2007). FYN deletion also improved plasma and tissue triglycerides, increased energy expenditure and enhanced fatty acid oxidation. The increased catabolism coincided with increased mitochondrial markers and AMPK phosphorylation in adipose tissue and skeletal muscle. Pharmacological inhibition of FYN in wild type mice induces weight loss, due to increased fatty acid oxidation associated with whole-body energy expenditure. FYN was also shown to modulate LKB1 subcellular localization through direct phosphorylation of LKB1 tyrosines 261 and 365. Mutation of these residues within LKB1 results in cytoplasmic localization and subsequent activation of AMPK1 in muscle cells (Yamada *et al.*, 2010). These findings establish the proto-oncogene FYN as a novel nutrient sensor and regulator of adipose tissue mass and insulin sensitivity. Taken together these observations indicate that activation of oncogenic signaling pathways through mutations in oncogenes or loss of tumor suppressors can impair or improve metabolic control.

See also: Cell Division/Death: Cell Cycle: Regulation of the p53 Pathway

References

- Adams, J.M., Cory, S., 2007. The Bcl-2 apoptotic switch in cancer development and therapy. *Oncogene* 26, 1324–1337.
- Barbacid, M., 1987. ras Genes. *Annual Review of Biochemistry* 56, 779–827.
- Barzilai, A., Melamed, E., 2003. Molecular mechanisms of selective dopaminergic neuronal death in Parkinson's disease. *Trends in Molecular Medicine* 9, 126–132.
- Bastie, C.C., Zong, H., Xu, J., *et al.*, 2007. Integrative metabolic regulation of peripheral tissue fatty acid oxidation by the SRC kinase family member Fyn. *Cell Metabolism* 5, 371–381.

- Bishop, J.M., 1983. Cellular oncogenes and retroviruses. *Annual Review of Biochemistry* 52, 301–354.
- Bishop, J.M., 1985. Viral oncogenes. *Cell* 42, 23–38.
- Blume-Jensen, P., Hunter, T., 2001. Oncogenic kinase signalling. *Nature* 411, 355–365.
- Bos, J.L., 1989. ras Oncogenes in human cancer: A review. *Cancer Research* 49, 4682–4689.
- Bos, J.L., Rehmann, H., Wittinghofer, A., 2007. GEFs and GAPs: Critical elements in the control of small G proteins. *Cell* 129, 865–877.
- Cairns, R.A., Harris, I.S., Mak, T.W., 2011. Regulation of cancer cell metabolism. *Nature Reviews Cancer* 11, 85–95.
- Cantley, L.C., 2002. The phosphoinositide 3-kinase pathway. *Science* 296, 1655–1657.
- Chalhoub, N., Baker, S.J., 2009. PTEN and the PI3-kinase pathway in cancer. *Annual Review of Pathology* 4, 127–150.
- Cichowski, K., Jacks, T., 2001. NF1 tumor suppressor gene function: Narrowing the GAP. *Cell* 104, 593–604.
- Cogswell, P.C., Guttridge, D.C., Funkhouser, W.K., Baldwin Jr., A.S., 2000. Selective activation of NF-kappa B subunits in human breast cancer: Potential roles for NF-kappa B2/p52 and for Bcl-3. *Oncogene* 19, 1123–1131.
- Cook, G.A., Gamble, M.S., 1987. Regulation of carnitine palmitoyltransferase by insulin results in decreased activity and decreased apparent K_i values for malonyl-CoA. *Journal of Biological Chemistry* 262, 2050–2055.
- da Costa, C.A., Ancolio, K., Checler, F., 2000. Wild-type but not Parkinson's disease-related Ala-53→Thr mutant alpha-synuclein protects neuronal cells from apoptotic stimuli. *Journal of Biological Chemistry* 275, 24065–24069.
- da Costa, C.A., Checler, F., 2010. A novel parkin-mediated transcriptional function links p53 to familial Parkinson's disease. *Cell Cycle* 9, 16–17.
- Dang, C.V., 2012. MYC on the path to cancer. *Cell* 149, 22–35.
- Dang, C.V., Semenza, G.L., 1999. Oncogenic alterations of metabolism. *Trends in Biochemical Sciences* 24, 68–72.
- Davies, H., Bignell, G.R., Cox, C., *et al.*, 2002. Mutations of the BRAF gene in human cancer. *Nature* 417, 949–954.
- Denko, N.C., 2008. Hypoxia, HIF1 and glucose metabolism in the solid tumour. *Nature Reviews Cancer* 8, 705–713.
- Dhillon, A.S., Hagan, S., Rath, O., Kolch, W., 2007. MAP kinase signalling pathways in cancer. *Oncogene* 26, 3279–3290.
- Di Cristofano, A., Pandolfi, P.P., 2000. The multiple roles of PTEN in tumor suppression. *Cell* 100, 387–390.
- Elchebly, M., Payette, P., Michaliszyn, E., *et al.*, 1999. Increased insulin sensitivity and obesity resistance in mice lacking the protein tyrosine phosphatase-1B gene. *Science* 283, 1544–1548.
- Fan, J., Ren, H., Jia, N., *et al.*, 2008. DJ-1 decreases Bax expression through repressing p53 transcriptional activity. *Journal of Biological Chemistry* 283, 4022–4030.
- Fearon, E.R., Vogelstein, B., 1990. A genetic model for colorectal tumorigenesis. *Cell* 61, 759–767.
- Fridman, J.S., Lowe, S.W., 2003. Control of apoptosis by p53. *Oncogene* 22, 9030–9040.
- Friedman, J.M., Birch, P.H., 1997. Type 1 neurofibromatosis: A descriptive analysis of the disorder in 1728 patients. *American Journal of Medical Genetics* 70, 138–143.
- George, S., Rochford, J.J., Wolfrum, C., *et al.*, 2004. A family with severe insulin resistance and diabetes due to a mutation in AKT2. *Science* 304, 1325–1328.
- Giaime, I., Sunyach, C., Herrant, M., *et al.*, 2006. Caspase-3-derived C-terminal product of synphilin-1 displays antiapoptotic function via modulation of the p53-dependent cell death pathway. *Journal of Biological Chemistry* 281, 11515–11522.
- Gustafson, W.C., Weiss, W.A., 2010. Myc proteins as therapeutic targets. *Oncogene* 29, 1249–1259.
- Hardie, D.G., 2007. AMP-activated/SNF1 protein kinases: Conserved guardians of cellular energy. *Nature Reviews Molecular Cell Biology* 8, 774–785.
- Harris, C.C., 1996. Structure and function of the p53 tumor suppressor gene: Clues for rational cancer therapeutic strategies. *Journal of the National Cancer Institute* 88, 1442–1455.
- Harris, H., Miller, O.J., Klein, G., Worst, P., Tachibana, T., 1969. Suppression of malignancy by cell fusion. *Nature* 223, 363–368.
- Haupt, S., Berger, M., Goldberg, Z., Haupt, Y., 2003. Apoptosis — the p53 network. *Journal of Cell Science* 116, 4077–4085.
- Hinds, P.W., Weinberg, R.A., 1994. Tumor suppressor genes. *Current Opinion in Genetics & Development* 4, 135–141.
- Ho, C.M., 2009. PTEN promoter methylation and LOH of 10q22–23 locus in PTEN expression of ovarian clear cell adenocarcinomas. *Gynecologic Oncology* 112, 307–313.
- Hojlund, K., Mustard, K.J., Staehr, P., *et al.*, 2004. AMPK activity and isoform protein expression are similar in muscle of obese subjects with and without type 2 diabetes. *American Journal of Physiology — Endocrinology and Metabolism* 286, E239–E244.
- Hollingsworth Jr., R.E., Hensley, C.E., Lee, W.H., 1993. Retinoblastoma protein and the cell cycle. *Current Opinion in Genetics & Development* 3, 55–62.
- Hubbard, S.R., Miller, W.T., 2007. Receptor tyrosine kinases: Mechanisms of activation and signaling. *Current Opinion in Cell Biology* 19, 117–123.
- Iurlaro, R., Leon-Annicchiarico, C.L., Munoz-Pinedo, C., 2014. Regulation of cancer metabolism by oncogenes and tumor suppressors. *Methods in Enzymology* 542, 59–80.
- Keilhack, H., David, F.S., McGregor, M., Cantley, L.C., Neel, B.G., 2005. Diverse biochemical properties of Shp2 mutants. Implications for disease phenotypes. *Journal of Biological Chemistry* 280, 30984–30993.
- Kim, J.W., Dang, C.V., 2006. Cancer's molecular sweet tooth and the Warburg effect. *Cancer Research* 66, 8927–8930.
- Kitamura, Y., Shimohama, S., Kamoshima, W., *et al.*, 1997. Changes of p53 in the brains of patients with Alzheimer's disease. *Biochemical and Biophysical Research Communications* 232, 418–421.
- Koo, S.H., Flechner, L., Qi, L., *et al.*, 2005. The CREB coactivator TORC2 is a key regulator of fasting glucose metabolism. *Nature* 437, 1109–1111.
- Krontiris, T.G., Cooper, G.M., 1981. Transforming activity of human tumor DNAs. *Proceedings of the National Academy of Sciences of the United States of America* 78, 1181–1184.
- Leclerc, I., Kahn, A., Doiron, B., 1998. The 5'-AMP-activated protein kinase inhibits the transcriptional stimulation by glucose in liver cells, acting through the glucose response complex. *FEBS Letters* 431, 180–184.
- Lee, H.G., Casadesus, G., Zhu, X., *et al.*, 2009. Cell cycle re-entry mediated neurodegeneration and its treatment role in the pathogenesis of Alzheimer's disease. *Neurochemistry International* 54, 84–88.
- Levine, A.J., Puzio-Kuter, A.M., 2010. The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. *Science* 330, 1340–1344.
- Lindhurst, M.J., Sapp, J.C., Teer, J.K., *et al.*, 2011. A mosaic activating mutation in AKT1 associated with the Proteus syndrome. *New England Journal of Medicine* 365, 611–619.
- Mandelkow, E., 1999. Alzheimer's disease. The tangled tale of tau. *Nature* 402, 588–589.
- Manning, B., Cantley, L., 2007. AKT/PKB signaling: Navigating downstream. *Cell* 129, 1261–1274.
- Martin, G.S., 2004. The road to Src. *Oncogene* 23, 7910–7917.
- Masters, C.L., Simms, G., Weinman, N.A., *et al.*, 1985. Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proceedings of the National Academy of Sciences of the United States of America* 82, 4245–4249.
- McCormick, F., 1989. ras GTPase activating protein: Signal transmitter and signal terminator. *Cell* 56, 5–8.
- McPherson, J.P., Hande, M.P., Poonepalli, A., *et al.*, 2006. A role for Brca1 in chromosome end maintenance. *Human Molecular Genetics* 15, 831–838.
- McShea, A., Lee, H.G., Petersen, R.B., *et al.*, 2007. Neuronal cell cycle re-entry mediates Alzheimer disease-type changes. *Biochimica et Biophysica Acta* 1772, 467–472.
- de la Monte, S.M., Sohn, Y.K., Wands, J.R., 1997. Correlates of p53- and Fas (CD95)-mediated apoptosis in Alzheimer's disease. *Journal of the Neurological Sciences* 152, 73–83.
- Moody, C.A., Laimins, L.A., 2010. Human papillomavirus oncoproteins: Pathways to transformation. *Nature Reviews Cancer* 10, 550–560.
- Morris, L.G., Veeriah, S., Chan, T.A., 2010. Genetic determinants at the interface of cancer and neurodegenerative disease. *Oncogene* 29, 3453–3464.
- Ohyagi, Y., Asahara, H., Chui, D.H., *et al.*, 2005. Intracellular Abeta42 activates p53 promoter: A pathway to neurodegeneration in Alzheimer's disease. *FASEB Journal* 19, 255–257.
- Orloff, M.S., He, X., Peterson, C., *et al.*, 2013. Germline PIK3CA and AKT1 mutations in Cowden and Cowden-like syndromes. *American Journal of Human Genetics* 92, 76–80.
- Osaki, M., Oshimura, M., Ito, H., 2004. PI3K-Akt pathway: Its functions and alterations in human cancer. *Apoptosis* 9, 667–676.
- Permutt, M.A., Wasson, J., Cox, N., 2005. Genetic epidemiology of diabetes. *Journal of Clinical Investigation* 115, 1431–1439.
- Perucho, M., Goldfarb, M., Shimizu, K., *et al.*, 1981. Human-tumor-derived cell lines contain common and different transforming genes. *Cell* 27, 467–476.
- Rivlin, N., Brosh, R., Oren, M., Rotter, V., 2011. Mutations in the p53 tumor suppressor gene: Important milestones at the various steps of tumorigenesis. *Genes & Cancer* 2, 466–474.
- Roberts, A.E., Araki, T., Swanson, K.D., *et al.*, 2007. Germline gain-of-function mutations in SOS1 cause Noonan syndrome. *Nature Genetics* 39, 70–74.
- Ruderman, N.B., Park, H., Kaushik, V.K., *et al.*, 2003. AMPK as a metabolic switch in rat muscle, liver and adipose tissue after exercise. *Acta Physiologica Scandinavica* 178, 435–442.

- Samuels, Y., Waldman, T., 2010. Oncogenic mutations of PIK3CA in human cancers. *Current Topics in Microbiology and Immunology* 347, 21–41.
- Shaw, R.J., Bardeesy, N., Manning, B.D., *et al.*, 2004. The LKB1 tumor suppressor negatively regulates mTOR signaling. *Cancer Cell* 6, 91–99.
- Shih, C., Weinberg, R.A., 1982. Isolation of a transforming sequence from a human bladder carcinoma cell line. *Cell* 29, 161–169.
- Stanbridge, E.J., 1976. Suppression of malignancy in human cells. *Nature* 260, 17–20.
- Stehelin, D., Varmus, H.E., Bishop, J.M., Vogt, P.K., 1976. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. *Nature* 260, 170–173.
- Steinberg, G.R., Smith, A.C., Van Denderen, B.J., *et al.*, 2004. AMP-activated protein kinase is not down-regulated in human skeletal muscle of obese females. *Journal of Clinical Endocrinology & Metabolism* 89, 4575–4580.
- Tartaglia, M., Gelb, B.D., Zenker, M., 2011. Noonan syndrome and clinically related disorders. *Best Practice & Research Clinical Endocrinology* 25, 161–179.
- Vander Heiden, M.G., Cantley, L.C., Thompson, C.B., 2009. Understanding the Warburg effect: The metabolic requirements of cell proliferation. *Science* 324, 1029–1033.
- Vita, M., Henriksson, M., 2006. The Myc oncoprotein as a therapeutic target for human cancer. *Seminars in Cancer Biology* 16, 318–330.
- Vogelstein, B., Papadopoulos, N., Velculescu, V.E., *et al.*, 2013. Cancer genome landscapes. *Science* 339, 1546–1558.
- Vousden, K.H., Prives, C., 2009. Blinded by the light: The Growing complexity of p53. *Cell* 137, 413–431.
- Weinberg, R.A., 1991. Tumor suppressor genes. *Science* 254, 1138–1146.
- Weinberg, R.A., 2007. *The Biology of Cancer*. New York, NY: Garland Science.
- Wong, J.T., Kim, P.T., Peacock, J.W., *et al.*, 2007. Pten (phosphatase and tensin homologue gene) haploinsufficiency promotes insulin hypersensitivity. *Diabetologia* 50, 395–403.
- Wood, L.D., Parsons, D.W., Jones, S., *et al.*, 2007. The genomic landscapes of human breast and colorectal cancers. *Science* 318, 1108–1113.
- Yamada, E., Pessin, J.E., Kurland, I.J., Schwartz, G.J., Bastie, C.C., 2010. Fyn-dependent regulation of energy expenditure and body weight is mediated by tyrosine phosphorylation of LKB1. *Cell Metabolism* 11, 113–124.
- Yin, Y., Shen, W.H., 2008. PTEN: A new guardian of the genome. *Oncogene* 27, 5443–5453.

Relevant Websites

- <http://cancergenome.broadinstitute.org>
Broadinstitute.
- <https://cansar.icr.ac.uk>
canSAR.
- <http://cancer.sanger.ac.uk/cancergenome/projects/cosmic/>
Catalogue of Somatic Mutations in Cancer.
- <http://www.cellsignal.com/common/content/content.jspid=science-pathways-mapk>
Cell Signaling Technology, Inc.
- [http://www.cellsignal.com/common/content/content.jspid=science-pathways-apoptosis&utm_source=2013-14 + Catalog + US&utm_medium=offline&utm_content=Apoptosis&utm_campaign=Offline + Media](http://www.cellsignal.com/common/content/content.jspid=science-pathways-apoptosis&utm_source=2013-14+Catalog+US&utm_medium=offline&utm_content=Apoptosis&utm_campaign=Offline+Media)
Cell Signaling Technology, Inc.
- <http://www.ebioscience.com/resources/pathways/p53-mediated-apoptosis-pathway-htm>
eBioscience, Inc.
- <http://www.sabiosciences.com/pathwaymagazine/minireview/metareprogram.php>
SABiosciences.
- <http://www.sabiosciences.com/pathwaycentral.php>
SABiosciences.
- <http://www.proteinatlas.org/cancer>
The Human Protein Atlas.
- http://www.biology.arizona.edu/cell_bio/tutorials/cell_cycle/cells2.html
The University of Arizona.
- <http://lsresearch.thomsonreuters.com/maps/443/>
Thomsonreuters.

Cyclins and Cyclin-Dependent Kinases

NH Lents, The City University of New York, New York, NY, USA

JJ Baldassare, Saint Louis University School of Medicine, St. Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Interest in the cell cycle dates all the way back to the mid-nineteenth century and the work of Walther Flemming, one of the founding fathers of microbiology. One of Flemming's most lasting contributions to science was his innovative use of aniline dyes, which allow the visualization of structures inside cells, especially the DNA of eukaryotes. Once it became possible to see chromatin and chromosomes, two terms coined by Flemming, the events of mitosis were first observed. Because these were the most visually conspicuous activities that were observable to early microbiologists, attention quickly focused on them and Flemming (1882) became the first biologist to study the cell cycle (Figure 1).

By observing and describing the events of mitosis, another term he coined, Flemming was the first to suggest that new cells get their nuclei from the division of the nucleus of the parent cell. This claim heavily influenced the development of modern cell theory, which states that all cells come from the division of a preexisting cell.

It was not until Morgan (1919) confirmed the role of chromosomes in conveying genetic information did biologists begin to understand the purpose and function of mitosis. By this time, the work of Mendel was being rediscovered, and the modern evolutionary synthesis was joining molecular biology with field biology under the grand unifying theme of evolution by natural selection. The central role of meiosis and mitosis in the propagation of genetic material through the ages thenceforth became clear.

The Discovery of Regulators of the Cell Cycle

After mitosis, the next phase of the cell cycle to be discovered was S phase, the DNA synthesis phase when chromosomes are

duplicated (Swift, 1950). However, it was noted that, in eukaryotes, DNA synthesis does not happen right before mitosis, as might be predicted. Instead, the synthesis of DNA occurs in a discrete, predictable, and relatively short period of time distinct from the events of mitosis. The period of time from the birth of a new cell until it begins synthesizing new DNA was named the first 'gap phase,' later shortened to G₁. The period of time after the completion of DNA synthesis and before the beginning of mitosis was called the second gap phase, or G₂. These two gap phases comprise the large majority of time in a given cell cycle, even in a cell population that is dividing as rapidly as it can. (Exceptions to this rule are found during embryogenesis and the development of multinucleate tissues such as skeletal muscle and fungal hyphae.) For an extensive review of the history of cell cycle research, see (Nurse, 2000; Figure 2).

Research into the cell cycle entered the modern era with the work of Tim Hunt using sea urchin eggs in 1982 (Evans *et al.*, 1983). Not coincidentally, sea urchin eggs were also the model system in which meiosis had been discovered a century earlier. It had previously been shown that protein synthesis inhibitors will almost immediately halt the growth and division of sea urchin embryos, even though these embryos are not engaging in a large amount of protein synthesis during the first several rounds of cell division. In all dueterostomes, the first phase of embryonic development, cleavage, is marked by rapid cycling through S phase, M phase, and cytokinesis, skipping the gap phases and with little or no growth in the overall size of the embryo. The result is a 128-cell embryo that is the same size as the original unfertilized egg.

That protein synthesis is required for cleavage to continue, even though little or no growth is taking place, suggested that a specific protein or proteins are required for cells to proceed through the cell cycle. This is what Tim Hunt was searching for in the summer of 1982. Sure enough, using radioactive pulse-chase experiments, he observed a protein that appeared and disappeared in a regular cyclic pattern that corresponded to the



Figure 1 Drawing by Walther Flemming of a cell from the salivary gland of a chironomid larva, showing the polytene chromosomes.

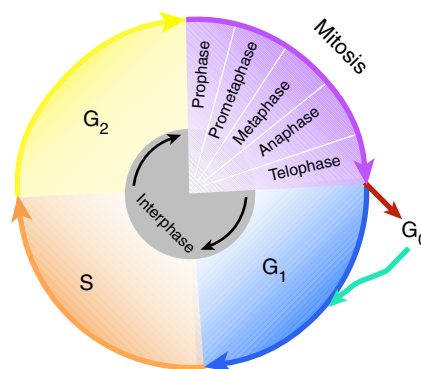


Figure 2 The phases of the eukaryotic cell cycle.

timing of the divisions of the sea urchin embryos. Specifically, the band disappeared rather suddenly about 10 minutes before the cells began the visible process of division (cytokinesis). Hunt (2004) named this protein cyclin in homage to his favorite hobby, cycling. The name stuck, given the obvious connection to the cycling of growing cells.

The cyclin that Hunt first observed was almost certainly what is now known as cyclin A (Evans *et al.*, 1983). Cyclins were subsequently observed in a variety of organisms. While cyclin A tends to gradually accumulate as a cell grows and then suddenly disappears as it prepares to divide, other cyclins have different patterns through the cell cycle.

Hunt's work on cyclins soon joined with work done by Paul Nurse and colleagues on *cdc2* (Nurse and Bissett, 1981) and that of Marc Kirschner and colleagues on the Maturation Promoting Factor, MPF (Cyert and Kirschner, 1988). It was eventually realized that cyclins are regulatory components of larger multi-protein enzyme complexes. The catalytic part of the complex, first discovered by Nurse as *cdc2p* in fission yeast (*cdc28p* in budding yeast), is a kinase that requires association with the cyclin to become active. Human *cdc2* was then re-named cyclin-dependent kinase 1 (CDKs), the founding member of the CDK family (Figure 2).

The Cyclins

Cyclins are found in every cell throughout all eukaryotes. Given the long history and frequent gene duplication, the cyclin family is quite large, with 30 identified members to date. Different cyclins function in different phases of the cell cycle in concert with their associated CDK, and different subtypes show specific tissue-distribution patterns in multicellular organisms. Nevertheless, the proteins are highly conserved, as evidenced by the fact that human cyclins can substitute for yeast cyclins to form functional cyclin-CDK complexes (Lew *et al.*, 1991).

Cyclin A, the founding member of the cyclin family of proteins, begins to accumulate in cells in late G1 phase, peaks in mid-G2 phase, and disappears completely at the beginning of mitosis (Sherr, 1994). In mammals, two genes encode slightly different forms of cyclin A: *CCNA1*, or cyclin A1, which is expressed almost exclusively in germ cells and *CCNA2*, or cyclin A2, which is expressed ubiquitously. Cyclin A partners with CDK2 in late G1 and S phases and with CDK1(*cdc2*) in late S and G2 phase (Pagano *et al.*, 1992). As is true for most all cyclins, the association of cyclin A with its partner CDKs induces a conformational change in the kinases to allow catalytic activity and to help direct substrate preference.

Other cyclins show a different pattern of cell cycle expression and complex with different CDKs. For example, D-type cyclins are present throughout the cell cycle in cells that are actively dividing and disappear only when cells exit the cell cycle temporarily during quiescence or permanently during senescence (Sherr, 1993). E-type cyclins, on the other hand, are expressed during a narrow window of time from mid-G1 to mid-S phase (Dulic *et al.*, 1992). B-type cyclins, also called the mitotic cyclins, only appreciably accumulate in G2 phase, reaching their peak in mid-mitosis before suddenly disappearing before cytokinesis (Fung and Poon, 2005).

Key to the mechanism of how cyclins execute cell cycle control is their carefully timed and sequential accumulation and degradation (Koepp *et al.*, 1999). Although transcriptional activation/repression plays a role, the greater part of this tight temporal control of protein expression is achieved through regulation of the stability of the cyclin proteins. Cyclins are markedly unstable proteins, often exhibiting a turnover rate that is among the highest of any known class of proteins. Accumulation is only achieved through the temporary stabilization of the proteins, an effect which is easily reversible, and the mechanism of which is particular to each cyclin. Thus, it should come as no surprise that a great deal of groundbreaking work on ubiquitin-mediated proteolysis was conducted while studying the very rapid and tightly regulated destruction of yeast cyclins during the cell cycle (Bai *et al.*, 1996).

At the same time, many cyclins have functions entirely separate from the cell cycle. For example, although cyclin C shares substantial sequence identity with the other cyclins and was identified in the same screen that identified cyclins D and E, it does not fluctuate through the cell cycle. Instead, cyclin C forms a complex with CDK8, which then phosphorylates the C-terminus of RNA polymerase II, a key event for transcriptional initiation (Rickert *et al.*, 1996). Similarly, cyclin H, together with CDK7 and another protein, MAT1, form a ternary complex that is an essential part of TFIID, a central component of the basal transcription machinery (Shiekhattar *et al.*, 1995). Although cyclin H-CDK7-MAT1 may also have a role in activating other CDKs, it appears constitutively active through the cycle and in quiescent cells.

Because not all cyclins fluctuate in the cell cycle, nor do all seem to bind to a CDK partner, the recognition and naming of members of the cyclin superfamily is now based on sequence similarity. The characteristic structure of a cyclin is found in the cyclin box, a large domain of about 100 amino acids that forms a characteristic folding of five alpha-helices (Noble *et al.*, 1997). All cyclins have one or two cyclin boxes, but the domain is found in other proteins as well, such as pRb and TFIIB. The cyclin box is key to the two most important features of the cyclins with respect to cell cycle control: binding to CDKs and the rapid turnover of cyclins by ubiquitin-mediated proteolysis. Table 1 below lists the cyclins known to date.

The Cyclin-Dependent Kinases

The cyclins do not have enzymatic activity. Instead, they appear to exert their function through altering the activity, localization, stability, and substrate preference of their catalytic binding partners, the cyclin-dependent kinases (CDKs) (Moragan, 1997). Accordingly, yeast cells have only one CDK: *cdc2p* in the fission yeast *Schizosaccharomyces pombe*, and *cdc28p* in the budding yeast *Saccharomyces cerevisiae*. The diversity of functions borne by the sole CDK is achieved by the sequential accumulation and degradation of various cyclins. For example, in budding yeast, the G1 cyclin Cln3p binds to and activates *cdc28p* to mediate G1 phase progression. The actions of Cln3p-*cdc28p* lead to accumulation of the next cyclins, Cln1 and Cln2. These cyclins replace Cln3 in the complex with *cdc28p* and direct it toward new targets then lead to the accumulation of the S-phase cyclins, Clb5 and Clb6. The targets

Table 1 The cyclin superfamily of proteins

Cyclin family	Genes	CDK partner	Other binding partners	Known roles
A	CCNA1, CCNA2	CDK1, CDK2	CDC6	S-G2 transition, G2 phase progression (Pagano <i>et al.</i> , 1992)
B	CCNB1, CCNB2, CCNB3	CDK1	TGFB2	M phase entry and exit (Hayles <i>et al.</i> , 1994)
C	CCNC	CDK8	MED proteins	Transcriptional initiation (Rickert <i>et al.</i> , 1996)
D	CCND1, CCND2, CCND3	CDK4, CDK6	p21, p16INK4a	Phosphorylation of pRb, G1 progression (Matsushime <i>et al.</i> , 1994)
E	CCNE1, CCNE2	CDK2	p21, p27	Phosphorylation of pRb, G1-S phase transition (Koff <i>et al.</i> , 1992)
F	CCNF	None	SKP1, CUL1, CP110	G2-M transition (Kong <i>et al.</i> , 2000)
G	CCNG1, CCNG2	None	p16INK4a, Mdm2, p53	Unknown
H	CCNH	CDK7	MAT1	TFIIH function (Feaver <i>et al.</i> , 1994), CDK-activating kinase (Makela <i>et al.</i> , 1994)
I	CCNI1, CCNI2	CDK5	Unknown	Differentiation/anti-apoptosis (Griffin <i>et al.</i> , 2006)
J	CCNJ, CCNJL	Unknown	Unknown	Unknown
K	CCNK	CDK9, 12, 13	RNA pol II	Transcriptional elongation (Edwards <i>et al.</i> , 1998)
L	CCNL1, CCNL2	CDK11	Unknown	mRNA splicing (Dickinson <i>et al.</i> , 2002)
M	FAM58A, FAM58B	CDK10	Unknown	Development, apoptosis (Guen <i>et al.</i> , 2013)
O	CCNO	Unknown	Unknown	Unknown
T	CCNT1, CCNT2	CDK9	RNA pol II	Transcriptional elongation (Peng <i>et al.</i> , 1998)
Y	CCNY, CCNYL1, CCNYL2	CDK14, CDK16	LRP6	Wnt signaling (Jiang <i>et al.</i> , 2009)

and functions of the same CDK, cdc28p, change throughout the cell cycle based on which cyclin is bound to it. For a review of yeast cyclins and CDKs, see Mendenhall and Hodge (1998).

The situation in metazoans is similar, but much more complicated. The prototypical CDK, CDK1 (previously known as cdc2 from its initial discovery as the human homolog of the *S. pombe* gene of the same name) is accompanied by more than 20 other CDKs, not all of which function in the cell cycle. In general, the protein levels of the CDKs do not fluctuate through the cell cycle as much as do the cyclins, though they do accumulate to moderately higher levels in the phases in which they are most enzymatically active. While most cyclins are medium-sized proteins at 50–60 kDa, the CDKs are substantially smaller at 30–40 kDa, underscoring the notion that the cyclins are the regulatory nodes of the enzyme complexes and the CDKs are the catalytic component.

There are more than 20 members of the CDK family in humans. Like the cyclins, the CDKs are highly conserved in eukaryotes. Human CDK1 shares more 60% sequence identity with the yeast cdc28p. Also like the cyclins, discoveries of non-cyclin-dependent and non-cell cycle-related functions have led to a structure-based definition of the CDK family of proteins (Malumbres *et al.*, 2009). To that end, the PSTAIRE helix (named for the conserved amino acids), which can also take the form PFTAIRE and other variations thereof, has been proposed as the family-defining feature of the CDKs (Table 2).

While the cyclins are larger and display a great deal of structural diversity, most CDKs are relatively compact proteins and have much less divergent structures among them. The majority of their structure is devoted to the formation of the kinase domain and only a few regulatory motifs are present.

Near the N-terminus, the CDKs contain an ATP-binding pocket, conserved among all kinase proteins, which includes one or two phospho-regulatory sites that usually inactivate the kinase when phosphorylated (Pavletich, 1999). In the C-terminal half of the protein is a characteristic T-loop structure that contains both the cyclin-binding domain (PSTAIRE box) and another phospho-regulatory threonine residue. In contrast, phosphorylation within the T-loop is required for enzymatic activity. Phosphorylation of the T-loop, together with cyclin binding, swings open the T-loop, making the active site accessible. The CDKs are proline-driven serine/threonine kinases and the consensus protein sequence for substrate recognition is S/T-P-X-K/R, although variations have been noted (Endicott *et al.*, 1999; Figure 3).

Functions of Cyclin-CDK in the Cell Cycle and Quiescence

Perhaps the most understood function of cyclins and CDKs is their control of cell cycle progression. Cyclin-CDK complexes appear to be the ‘master regulators’ of the cell cycle, controlling the events and timing that define each phase (Figure 4).

Cell cycle control begins with the D-type cyclins (Sherr, 1994). Cells that have temporarily exited the cell cycle and are not dividing are said to be quiescent, a state often referred to as G0 phase. While CDKs may be present in quiescent cells, they are not active because the cell cycle-related cyclins are not detectable. However, when cells are stimulated to reenter the cell cycle and begin dividing, either through addition of a growth factor or loss of contact inhibition, the D-type cyclins

Table 2 The cyclin-dependent kinase superfamily of proteins

CDK	Other names	Cyclin(s)	Major target(s)	Known roles
CDK1	CDC2, CDC28A, CDKN1	A, B	pRb family, MCM family, ORC1, APC, FOXM1, FOXO1	G2 progression, initiation of mitosis (Draetta and Beach, 1988)
CDK2	CDKN2	A, E	pRb family, MCM family, NPM, MYC, NPAT	G1 and S-phase progression (Koff <i>et al.</i> , 1992)
CDK3	CDKN3	C	ATF1	Transcriptional initiation (Zheng <i>et al.</i> , 2008)
CDK4	NCLK, PSSALRE	D	pRb family	G1 progression (Kato <i>et al.</i> , 1993)
CDK5		–		Neuronal signaling and development (Nikolic <i>et al.</i> , 1996)
CDK6	PLSTIRE	D	pRb family	G1 progression (Meyerson and Harlow, 1994)
CDK7	MO15	H	RNA pol II, T-loop of CDKs	TFIIH function (Roy <i>et al.</i> , 1994), CDK-activating kinase (Solomon <i>et al.</i> , 1993)
CDK8		C	RNA pol II, cyclin H	Transcriptional initiation (Rickert <i>et al.</i> , 1996)
CDK9	P-TEFb, PITALRE, CDC2L4, TAK	T, K	RNA pol II	Transcriptional elongation (Wei <i>et al.</i> , 1998)
CDK10	PISSLRE	M	ETS2	Cell growth in development, apoptosis (Kasten and Giordano, 2001)
CDK11	PITSLRE, PITSVRE	L	Unknown	Apoptosis, formation of RNA pol II L-mediator complex (Hu <i>et al.</i> , 2003)
CDK12	CRK7	K	RNA pol II, SRSF1/SF2	Transcriptional elongation (Bartkowiak <i>et al.</i> , 2010), RNA splicing (Chen <i>et al.</i> , 2006)
CDK13	CDC2L, CDC2L5, CHED	K	RNA pol II, SRSF1/SF2	Transcriptional elongation, RNA splicing (Chen <i>et al.</i> , 2007)
CDK14	PFTK1, PFTAIRES1	Y	LRP6	Wnt signaling (Davidson <i>et al.</i> , 2009)
CDK15	PFTAIRES2	Unknown	Unknown	Unknown
CDK16	PCTAIRES1, PCTK1	Y	NSF	Neural vesicle transport and exocytosis, neurite outgrowth (Graeser <i>et al.</i> , 2002)
CDK17	PCTAIRES2, PCTK2	Unknown	Unknown	Neuronal differentiation (Hirose <i>et al.</i> , 1997)
CDK18	PCTAIRES3, PCTK3	Unknown	Unknown	Unknown
CDK19	CDC2L6	Unknown	Unknown	Formation of mediator complex of RNA pol II (Ender <i>et al.</i> , 2008)
CDK20	PNQALRE, CCRK, CAK-kinase-p42	None	CDKs	Shh signaling, Cilia formation (Ko <i>et al.</i> , 2010)

(Cln3 in yeast) begin to accumulate in 2–3 h. In mammals, the D cyclins then bind to and activate CDK4 and/or CDK6. The only bona fide catalytic targets of D-cyclin complexes are the three pocket proteins: the retinoblastoma protein, pRb (RB1); p107 (RBL1); and p130 (RBL2) (Weinberg, 1995). Unlike that of other cell cycle cyclins, the levels of D-type cyclins remains high through the cell cycle, as long as the cells are stimulated and actively progressing through the cycle. Only upon withdrawal of growth factor, contact inhibition, or other anti-mitogenic signaling do the D-type cyclins disappear (Won *et al.*, 1992).

The initial phosphorylation of pocket proteins by cyclin D-CDKs in early G1 phase results in a conformational change that allows its binding partner, E2F, to become partially active. This event, which typically takes 3–5 h to be detectable, marks the progression of cells into mid-G1 phase. E2F is a transcription factor that trans-activates many cell cycle-related genes. The net effect of D-cyclin activity on E2F is to allow the

expression of the next G1 phase cyclin: cyclin E (Harbour and Dean, 2000; Ohtani *et al.*, 1995).

E cyclins then complex with CDK2 (in mammalian cells). The Cyclin E-CDK2 complex has several known targets for its kinase activity, which can be collectively understood as preparation for S-phase entry. For example, E-CDK2 phosphorylates nucleophosmin and CP110, triggering centrosome duplication (Chen *et al.*, 2002; Okuda *et al.*, 2000). E-CDK2 also activates p220(NPAT), which promotes transcription of the histone genes, in preparation for chromosome duplication (Zhao *et al.*, 2000). Finally, and perhaps most importantly, E-CDK2 continues the phosphorylation of pRb and the other pocket proteins that was previously begun by D-CDK4/6 (Harbour *et al.*, 1999). The hyper-phosphorylation of pRb completes its inactivation, fully releasing E2F to transcribe many S-phase genes such as DNA polymerase α , thymidine kinase, dihydrofolate reductase, and the next cyclin, cyclin A (Ren *et al.*, 2002). A final function of cyclin E is to promote the

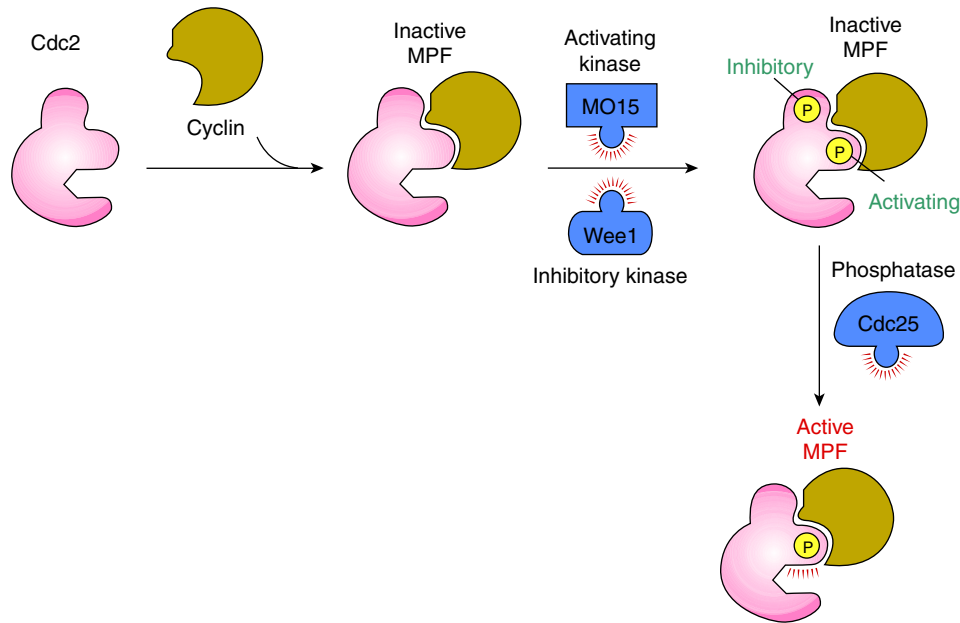


Figure 3 The major activating steps of the MPF complex.

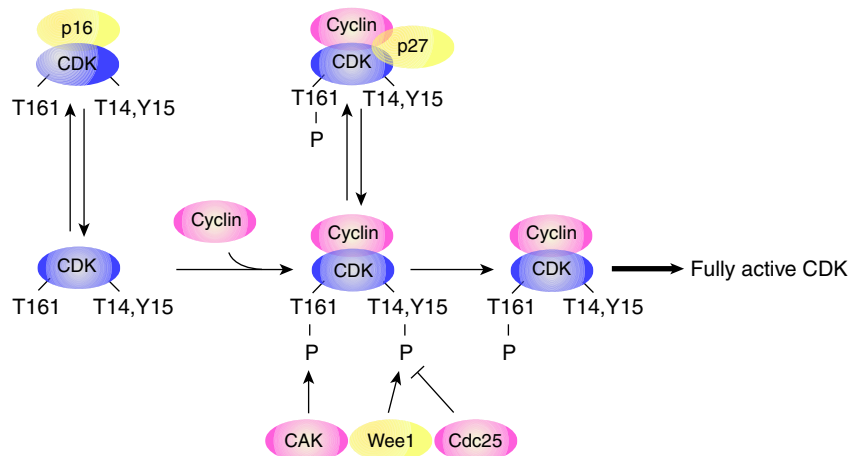


Figure 4 Associations and modifications regulating cyclin-CDK activity.

assembly of the pre-replication complex (Coverley *et al.*, 2002). Thus, the activity of E-CDK2 typically marks the transition from G1 phase into S-phase. By mid-S phase, levels of cyclin E drop dramatically (Figure 5).

The A-type cyclins begin to appear in late G1 and build throughout S phase. Initially, they bind to CDK2, replacing cyclin E after S phase has begun. Temporally speaking, the first function of A-CDK2 is to phosphorylate the ORC (origin recognition complex) and MCM (mini-chromosome maintenance) proteins in order to allow firing of replication origins (Lei and Tye, 2001). This begins DNA synthesis. Soon after, A-CDK2 then phosphorylates cdc6, which promotes disassembly of the ORC complexes (Petersen *et al.*, 1999). This second function of A-CDK2 ensures that each origin fires once and only once.

As cells complete DNA replication and begin G2 phase, cyclin A begins to join with CDK1, although A-CDK2

complexes persist into mitosis and likely play a role in spindle formation (Pagano *et al.*, 1992). A-CDK1 complexes, on the other hand, are active for a narrow window between mid-S phase and late G2 phase. A-CDK1 appears to have roles analogous to A-CDK2 in preventing multiple firing of replication origins, an important consideration given that re-replication of DNA can lead to genomic catastrophe and neoplasia. G2 functions of A-CDK1 are the execution of checkpoint controls, discussed below.

Levels of cyclin B build more slowly than the other cyclins, beginning in late G1 and only reaching appreciable levels in mid-G2, when it begins to replace cyclin A in complexes with CDK1. The cyclin C-CDK1 complex is also known as maturation promoting factor (MPF), which begins to form well before it comes enzymatically active (Murray *et al.*, 1989). The final activating step, which takes place in late G2 is the

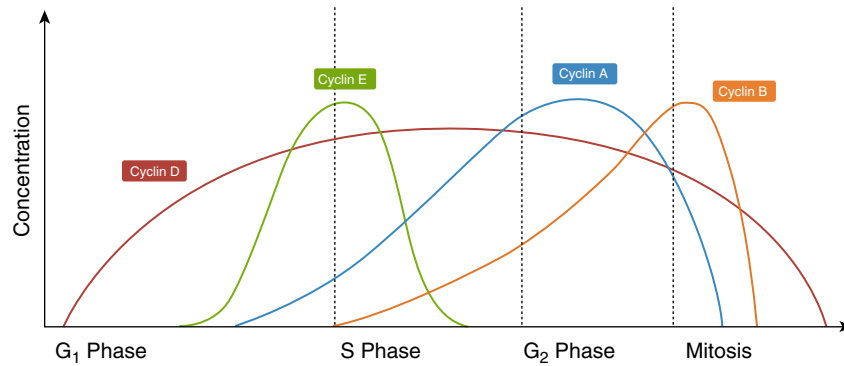


Figure 5 Timing of cyclin accumulation through the mammalian cell cycle.

removal of an inhibitory phosphorylation by the phosphatase Cdc25 (Gu *et al.*, 1992). Once active, MPF induces chromosome condensation via phosphorylation of condensin, dissolution of the nuclear lamina via phosphorylation of the lamin proteins, fragmentation of the organelles of the endomembrane system through phosphorylation of GM130, and the formation of the mitotic spindle through several targets. Cyclin B levels, and MPF activity, peaks at the moment mitosis begins, after which cyclin B is degraded to undetectable levels rather quickly. Cyclin B is targeted for degradation by the Anaphase Promoting Complex, an E3 ubiquitin ligase (Peters, 2002). Only when cyclin B is degraded do the sister chromatids separate and begin migrating to opposite centrosomes. The remaining steps of mitosis occur without regulation by cyclin-CDK complexes.

CKIs: CDK Inhibitors

Regulation of cyclin-CDK activity is multilayered. Besides cyclin binding, the kinase activity of CDKs is affected by both inhibitory and activating phosphorylation, discussed above. In addition, there are several small proteins that act as inhibitors to cyclin-CDK complexes called CKIs, cyclin-dependent kinase inhibitors (Sherr and Roberts, 1999). These small proteins lack extensive secondary or tertiary structure and instead act merely to disable kinase activity through various mechanisms.

There are two families of CKIs, the INK4 proteins and the CIP/KIP proteins. The INK4 proteins were so-named for their ability to inhibit CDK4. (They similarly inhibit the CDK4 analog CDK6.) These proteins bind to CDK4/6 in a manner that is directly competitive with the D-cyclins. By doing so, they prevent activation of CDK4/6 kinase activity and subsequent phosphorylation of the pRb family of pocket proteins (Hengst *et al.*, 1994). Without the action of D-cyclin kinase activity, the pocket proteins accumulate in their un- and under-phosphorylated forms, which then repress the E2F transcription factors. As such, the INK4 proteins are effective at halting cell cycle progression and will be discussed below for their role as tumor suppressors.

The CIP/KIP proteins function more broadly as CKIs, capable of inhibiting all or nearly all of the cyclin-CDK complexes that function in cell cycle regulation (Sherr and Roberts, 1999). In addition, their function is more nuanced. CIP/KIP

Table 3 The CDK inhibitors (CKIs)

CKI type	Member (gene name)	CDKs inhibited
INK4	p16Ink4A (CDKN2A) p15INK4B (CDKN2B) p18INK4C (CDKN2C) p19INK4D (CDKN2D)	CDK4, CDK6
CIP/KIP	p21CIP1/WAF1 (CDKN1A) p27KIP1 (CDKN1B) p57KIP2 (CDKN1C)	All cell cycle-related CDKs

proteins bind both to the cyclins and the CDKs and, when present in high concentration, can disrupt and/or inhibit these proteins even after they have formed. However, at lower concentrations, the CIP/KIP proteins may actually promote D-, E-, and possibly A- cyclin-CDK activity, either through a role in recruitment/complex formation, or nuclear translocation. Nevertheless, the CIP/KIP proteins, through their inhibition of cyclin-CDK complexes, can also put the breaks on cell cycle progression and thus function as tumor suppressors (Table 3).

Cyclins, CDKs, and CKIs in Cancer

Mutations of Cyclins and CDKs in Cancer

Given their role in promoting cell growth and division, it is not surprising that a great deal of attention has been paid to the role of the cyclins and CDKs in the genesis of cancer. However, CDK and cyclin genes are not common targets of mutation in cancer. Of all the CDK genes, only *CDK4* has been reported as a mutated oncogene and only in a single type of cancer, melanoma, in less than 5% of cases (Goldstein *et al.*, 2000). This is not altogether surprising given that full activation of the CDKs requires several regulatory steps and cannot be achieved through overexpression or mutation alone. Perhaps more surprising is that neither are the cyclins frequent targets of gene mutation in cancer, other than through gene duplication. Exceptions are that *CCND1* gene amplifications are occasionally found in a variety of cancers including that of breast, uterine, colon, and the head, while mutations in cyclin E1 sometimes appear in hepatocellular carcinoma. In both cases, mutations or duplications in cyclin genes seem to be the

result of the advanced progression of cancer, not its initiation (Sherr and McCormick, 2002).

Overexpression of Cyclins in Cancer

More commonly, however, the D-, E-, and A-cyclins are observed to be overexpressed in human cancers, owing at least in part to the fact that the most common proto-oncogenes (K-RAS, MYC, BCR-ABL, HER-2/neu, C-KIT, EGFR, B-Raf) all induce the expression of the cyclins as part of their common mechanism of action (Croce, 2008). The vast majority of proto-oncogenes encode proteins that operate in the normal extracellular signaling pathways that promote cell growth and proliferation. As such, they sit directly upstream of the cyclins in their mechanism of action. There is an advantage for cancer to activate genes that act upstream in the mitogenic signaling cascade, rather than downstream, because the latter form of misregulation is often sensed as unscheduled cell division, which trips the anti-proliferative defense mechanisms of apoptosis or senescence. Mutations in growth factor receptors, however, often promote both cell division and survival, making them much more powerful than a cyclin mutation in the genesis of neoplasia.

Because cancer cells proliferate faster than normal cells and subvert cues to restrain their growth, the cyclins and CDKs are typically overexpressed or overactive in these cells. Because cyclin-CDK complexes are the vehicles of, and rate-limiting factors in, cell proliferation, they are attractive targets for cancer treatments. To date, no CDK inhibitors have begun human trials as anticancer drugs, despite years of research thereof. Although the FDA has recently approved an inhibitor of CDK4/6 for human trials (Brower, 2014), this has come nearly three decades after the discovery of CDKs and after similar inhibitors of CDK1 and CDK2 have failed miserably in animal trials. Presumably, this owes to the extreme toxicity of these drugs, given that specificity toward specific CDKs or even just to CDKs over other kinases, has been elusive. No pharmaceuticals have yet been reported that are able to specifically disrupt the function of cyclins, though attempts to do so with miRNAs and monoclonal antibodies are ongoing. In contrast, some inhibitors of tyrosine kinase receptors have been in the pharmacopeia of cancer treatment for nearly two decades.

The CKIs: Bona Fide Tumor Suppressors

Cyclin-dependent kinase inhibitors (CKIs), on the other hand, are very frequently mutated in human cancers. This is because the natural function of CKIs is to halt or slow cellular proliferation. These genes are thus referred to as tumor suppressors, a label that is borne out by mouse models demonstrating that loss of the CKIs accelerates tumorigenesis dramatically.

INK4a is one of the most frequently mutated human tumor suppressor genes, with deletions and mutations having been reported in a wide variety of human cancers. In addition, epigenetic silencing of the locus has also been reported, usually through DNA methylation. Interpretation of mutations and gene silencing is complicated by the fact that a different tumor suppressor, p19ARF, is also encoded by the *CDKN2A* locus in an alternate reading frame (Sherr, 2001). Two distinct

tumor suppressors, indeed two of the most important tumor suppressors in our defense against cancer, are collocated within the same stretch of DNA, meaning that a single deletion or mutation event could disrupt both gene products. This is a remarkable evolutionary conundrum, given that a simple gene duplication event with subsequent divergence of the two loci would solve this genetic hyper-vulnerability.

The CIP/KIP proteins are also bona fide tumor suppressors as their deletion in mice lead to increased development of malignancy (Denicourt and Dowdy, 2004). Even more dramatically, loss or mutation of CIP/KIP genes exacerbates tumorigenesis in other cancer-prone genetic backgrounds. This argues that CKIs function as key operators in the cancer protective mechanisms of other tumor suppressors. Indeed, discovery of p21CIP1/WAF1 was announced in back-to-back papers, one describing its ability to suppress the activities of the CDKs and one describing it as a primary means by which the p53 'master tumor suppressor' mediates cell cycle arrest (Harper *et al.*, 1993; El-Deiry *et al.*, 1993). Nevertheless, loss-of-function mutations in the *CDKN1* loci are not common in human cancers, as they are with other tumor suppressor genes such as *RB1* and *TP53*. Indeed, CIP/KIP proteins are frequently observed to be overexpressed in human tumors and correlate with poor prognosis. Although oncogenic roles for these CKIs have been postulated, in many cases, this overexpression is likely the result of a failed attempt by damaged cells to contain their own malignancy.

Although the cyclins, CDKs, and CKIs are not among the list of the most frequently mutated or deleted proto-oncogenes or tumor suppressors, they are intimately connected with all those that are, as effectors, activators, or inhibitors. Neither tumor suppression nor tumorigenesis can occur without the direct involvement of the cyclins and CDKs. Cancer is a disease of the cell cycle.

Beyond the Cell Cycle: Other Functions of Cyclins and CDKs

In addition to their most famous role in cell cycle regulation, the cyclins and CDKs are being linked to an ever-increasing number of non-cell cycle-related functions. In mammals, this is likely the result of frequent gene duplication and redundancy among cyclins and CDKs, which has allowed new functions to emerge. The levels of some cyclins and most CDKs do not fluctuate appreciably through the various phases of the cell cycle, being impervious even to cell cycle exit. In the early 2000s, the cancer biology community was shocked to learn that many cyclins and CDKs were genetically dispensable in mice, presumably due to compensation and substitution by other family members (Satyanarayana and Kaldis, 2009). However, mice lacking individual cyclins or CDKs often have unexpected specific phenotypes unrelated to cellular proliferation, revealing novel functions of these gene products.

The connection between the cell cycle and transcription runs deep. Although many of the cell cycle roles of the cyclin-CDKs are executed through the phosphorylation of transcriptional activators or repressors, many cyclins and CDKs also function within the basal transcription machinery itself. Most often, cyclin-CDK complexes act to phosphorylate the C-terminal domain (CTD) of RNA polymerase II in a

sequential fashion, leading to the progressive recruitment of the basal machinery, transcriptional initiation, and elongation (Dymlach, 1997). The sequential phosphorylation of the CTD by cyclin-CDK is highly reminiscent of the sequential phosphorylation of the pocket proteins (or vice versa). Roles for cyclin-CDKs in pre-mRNA splicing have also been discovered.

Even more surprising are the discoveries that cyclins and CDKs may have functions that do not require association with a respective partner. For example, CDK10 is known to bind directly to the transcription factor ETS2 in the absence of a cyclin partner (Kasten and Giordano, 2001). Its repression of ETS2-mediated transcription of proto-oncogenic *RAF1*, which requires no kinase activity, can even promote tumorigenicity despite CDK10 itself being regulated wholly separate from the cell cycle. CDK5 has long been an outlier of the CDK family in that it seems to function only in terminally differentiated neuronal cells and does so entirely without the assistance of a cyclin partner (Dhavan and Tsai, 2001). Further, cyclins, CDKs, and even CKIs have been shown to operate in DNA damage checkpoints and repair mechanisms in an entirely cell cycle-independent manner.

More recently, the cyclins and CDKs are popping up in the study of differentiation, cell fate determination, and cell cycle exit as well as tissue renewal, stem cell proliferation, and meiosis (Lim and Kaldis, 2013). Importantly, the newly discovered roles for cyclin-CDKs in these proliferative and anti-proliferative processes are not always tied directly to cell cycle progression. Highly tissue- and context-specific functions are being discovered for the less studied cyclins and CDKs, as are new roles for the well known ones via new effectors. While it is tempting to interpret these roles as purely a function of cell cycle regulation, it does not appear to be so.

The summary of this is that the cyclin-CDK complex has formed a modular unit that has proven highly useful for eukaryotic cells. Cyclin-CDK units can be regulated by a variety of overlapping mechanisms such as complex assembly, the binding of a CKI, subcellular localization, protein stability, and posttranslational modification. In particular, the labile nature of the cyclin protein lends itself to exquisite temporal regulation. Furthermore, substrate and effector preferences of the cyclin-CDK pair are malleable and can be altered by protein-protein interactions. As such, through gene duplication events, the course of evolutionary time has seen the proliferation of these complexes and their subsequent adaptation for a variety of cellular uses. From the cellular proliferation of early embryos to the recruitment of stem cells for tissue renewal, cyclin-CDK complexes are the engines that drive, or the brakes that stop, the growth of all eukaryotic cells.

See also: Cell Division/Death: Cell Cycle: CDK Inhibitors in Normal and Malignant Cells. Functional Cell Biology: An Overview

References

- Bai, C., Sen, P., Hofmann, K., *et al.*, 1996. SKP1 connects cell cycle regulators to the ubiquitin proteolysis machinery through a novel motif, the F-box. *Cell* 86, 263–274.
- Bartkowiak, B., Liu, P., Phatnani, H.P., *et al.*, 2010. CDK12 is a transcription elongation-associated CTD kinase, the metazoan ortholog of yeast Ctk1. *Genes & Development* 24, 2303–2316.
- Brower, V., 2014. Cell cycle inhibitors make progress. *Journal of the National Cancer Institute* 106, dju221.
- Chen, H.-H., Wang, Y.-C., Fann, M.-J., 2006. Identification and characterization of the CDK12/cyclin L1 complex involved in alternative splicing regulation. *Molecular and Cellular Biology* 26, 2736–2745.
- Chen, H.-H., Wong, Y.-H., Genevieve, A.-M., Fann, M.-J., 2007. CDK13/CDC2L5 interacts with L-type cyclins and regulates alternative splicing. *Biochemical and Biophysical Research Communications* 354, 735–740.
- Chen, Z., Indjeian, V.B., McManus, M., Wang, L., Dynlacht, B.D., 2002. CP110, a cell cycle-dependent CDK substrate, regulates centrosome duplication in human cells. *Developmental Cell* 3, 339–350.
- Coverley, D., Laman, H., Laskey, R.A., 2002. Distinct roles for cyclins E and A during DNA replication complex assembly and activation. *Nature Cell Biology* 4, 523–528.
- Croce, C.M., 2008. Oncogenes and cancer. *New England Journal of Medicine* 358, 502–511.
- Cyert, M.S., Kirschner, M.W., 1988. Regulation of MPF activity *in vitro*. *Cell* 53, 185–195.
- Davidson, G., Shen, J., Huang, Y.-L., *et al.*, 2009. Cell cycle control of wnt receptor activation. *Developmental Cell* 17, 788–799.
- Denicourt, C., Dowdy, S.F., 2004. Cip/Kip proteins: More than just CDKs inhibitors. *Genes & Development* 18, 851–855.
- Dhavan, R., Tsai, L.-H., 2001. A decade of CDK5. *Nature Reviews Molecular Cell Biology* 2, 749–759.
- Dickinson, L.A., Edgar, A.J., Ehley, J., Gottesfeld, J.M., 2002. Cyclin L is an RS domain protein involved in pre-mRNA splicing. *Journal of Biological Chemistry* 277, 25465–25473.
- Draetta, G., Beach, D., 1988. Activation of *cdc2* protein kinase during mitosis in human cells: Cell cycle-dependent phosphorylation and subunit rearrangement. *Cell* 54, 17–26.
- Dulic, V., Lees, E., Reed, S.I., 1992. Association of human cyclin E with a periodic G1-S phase protein kinase. *Science* 257, 1958–1961.
- Dynlacht, B.D., 1997. Regulation of transcription by proteins that control the cell cycle. *Nature* 389, 149–152.
- Edwards, M.C., Wong, C., Elledge, S.J., 1998. Human cyclin K, a novel RNA polymerase II-associated cyclin possessing both carboxy-terminal domain kinase and Cdk-activating kinase activity. *Molecular and Cellular Biology* 18, 4291–4300.
- El-Deiry, W.S., Tokino, T., Velculescu, V.E., *et al.*, 1993. *WAF1*, a potential mediator of p53 tumor suppression. *Cell* 75, 817–825.
- Ender, C., Krek, A., Friedländer, M.R., *et al.*, 2008. A human snoRNA with microRNA-like functions. *Molecular Cell* 32, 519–528.
- Endicott, J.A., Noble, M.E., Tucker, J.A., 1999. Cyclin-dependent kinases: Inhibition and substrate recognition. *Current Opinion in Structural Biology* 9, 738–744.
- Evans, T., Rosenthal, E.T., Youngblom, J., Distel, D., Hunt, T., 1983. Cyclin: A protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. *Cell* 33, 389–396.
- Feaver, W.J., Svestrup, J.Q., Henry, N.L., Kornberg, R.D., 1994. Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIF/TFIIK. *Cell* 79, 1103–1109.
- Flemming, W., 1882. *Zellsubstanz, kern und zelltheilung*. Leipzig, Germany: Vogel.
- Fung, T.K., Poon, R.Y., 2005. A roller coaster ride with the mitotic cyclins. *Seminars in Cell & Developmental Biology* 16, 335–342.
- Goldstein, A.M., Struwing, J.P., Chidambaram, A., Fraser, M.C., Tucker, M.A., 2000. Genotype–phenotype relationships in US melanoma-prone families with CDKN2A and CDK4 mutations. *Journal of the National Cancer Institute* 92, 1006–1010.
- Graeser, R., Gannon, J., Poon, R.Y., *et al.*, 2002. Regulation of the CDK-related protein kinase PCTAIRE-1 and its possible role in neurite outgrowth in Neuro-2A cells. *Journal of Cell Science* 115, 3479–3490.
- Griffin, S.V., Olivier, J.P., Pippin, J.W., Roberts, J.M., Shankland, S.J., 2006. Cyclin I protects podocytes from apoptosis. *Journal of Biological Chemistry* 281, 28048–28057.
- Gu, Y., Rosenblatt, J., Morgan, D.O., 1992. Cell cycle regulation of CDK2 activity by phosphorylation of Thr160 and Tyr15. *EMBO Journal* 11, 3995–4005.
- Guen, V.J., Gamble, C., Flajolet, M., *et al.*, 2013. CDK10/cyclin M is a protein kinase that controls ETS2 degradation and is deficient in STAR syndrome. *Proceedings of the National Academy of Sciences* 110, 19525–19530.
- Harbour, J.W., Dean, D.C., 2000. The Rb/E2F pathway: Expanding roles and emerging paradigms. *Genes & Development* 14, 2393–2409.

- Harbour, J.W., Luo, R.X., Dei, S.A., Postigo, A.A., Dean, D.C., 1999. Cdk phosphorylation triggers sequential intramolecular interactions that progressively block Rb functions as cells move through G1. *Cell* 98, 859–869.
- Harper, J.W., Adami, G.R., Wei, N., Keyomarsi, K., Elledge, S.J., 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* 75, 805–816.
- Hayles, J., Fisher, D., Woollard, A., Nurse, P., 1994. Temporal order of S phase and mitosis in fission yeast is determined by the state of the p34^{cdc2}-mitotic B cyclin complex. *Cell* 78, 813–822.
- Hengst, L., Dulic, V., Slingerland, J.M., Lees, E., Reed, S.I., 1994. A cell cycle-regulated inhibitor of cyclin-dependent kinases. *Proceedings of the National Academy of Sciences of the United States of America* 91, 5291–5295.
- Hirose, T., Tamaru, T., Okumura, N., Nagai, K., Okada, M., 1997. PCTAIRE 2, a Cdc2-related serine/threonine kinase, is predominantly expressed in terminally differentiated neurons. *European Journal of Biochemistry* 249, 481–488.
- Hu, D., Mayeda, A., Trembley, J.H., Lahti, J.M., Kidd, V.J., 2003. CDK11 complexes promote pre-mRNA splicing. *Journal of Biological Chemistry* 278, 8623–8629.
- Hunt, T., 2004. The discovery of cyclin (I). *Cell* 116, S63–S64.
- Jiang, M., Gao, Y., Yang, T., Zhu, X., Chen, J., 2009. Cyclin Y, a novel membrane-associated cyclin, interacts with PFTK1. *FEBS Letters* 583, 2171–2178.
- Kasten, M., Giordano, A., 2001. Cdk10, a Cdc2-related kinase, associates with the Ets2 transcription factor and modulates its transactivation activity. *Oncogene* 20, 1832–1838.
- Kato, J., Matsushime, H., Hiebert, S.W., Ewen, M.E., Sherr, C.J., 1993. Direct binding of cyclin D to the retinoblastoma gene product (pRb) and pRb phosphorylation by the cyclin D-dependent kinase CDK4. *Genes & Development* 7, 331–342.
- Ko, H.W., Norman, R.X., Tran, J., *et al.*, 2010. Broad-minded links cell cycle-related kinase to cilia assembly and hedgehog signal transduction. *Developmental Cell* 18, 237–247.
- Koepp, D.M., Harper, J.W., Elledge, S.J., 1999. How the cyclin became a cyclin: Regulated proteolysis in the cell cycle. *Cell* 97, 431–434.
- Koff, A., Giordano, A., Desai, D., *et al.*, 1992. Formation and activation of a cyclin E-cdk2 complex during the G1 phase of the human cell cycle. *Science* 257, 1689–1694.
- Kong, M., Barnes, E.A., Ollendorff, V., Donoghue, D.J., 2000. Cyclin F regulates the nuclear localization of cyclin B1 through a cyclin–cyclin interaction. *EMBO Journal* 19, 1378–1388.
- Lei, M., Tye, B.K., 2001. Initiating DNA synthesis: From recruiting to activating the MCM complex. *Journal of Cell Science* 114, 1447–1454.
- Lew, D.J., Dulic, V., Reed, S.I., 1991. Isolation of three novel human cyclins by rescue of G1 cyclin (Cln) function in yeast. *Cell* 66, 1197–1206.
- Lim, S., Kaldis, P., 2013. Cdks, cyclins and CKIs: Roles beyond cell cycle regulation. *Development* 140, 3079–3093.
- Makela, T.P., Tassan, J.P., Nigg, E.A., *et al.*, 1994. A cyclin associated with the CDK-activating kinase MO15. *Nature* 371, 254–257.
- Malumbres, M., Harlow, E., Hunt, T., *et al.*, 2009. Cyclin-dependent kinases: A family portrait. *Nature Cell Biology* 11, 1275–1276.
- Matsushime, H., Quelle, D.E., Shurtleff, S.A., *et al.*, 1994. D-type cyclin-dependent kinase activity in mammalian cells. *Molecular and Cellular Biology* 14, 2066–2076.
- Mendenhall, M.D., Hodge, A.E., 1998. Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast *Saccharomyces cerevisiae*. *Microbiology and Molecular Biology Reviews* 62, 1191–1243.
- Meyerson, M., Harlow, E., 1994. Identification of G1 kinase activity for cdk6, a novel cyclin D partner. *Molecular and Cellular Biology* 14, 2077–2086.
- Morgan, D.O., 1997. Cyclin-dependent kinases: Engines, clocks, and microprocessors. *Annual Review of Cell and Developmental Biology* 13, 261–291.
- Morgan, T.H., 1919. *The Physical Basis of Heredity*. Philadelphia, PA: JB Lippincott Company.
- Murray, A.W., Solomon, M.J., Kirschner, M.W., 1989. The role of cyclin synthesis and degradation in the control of maturation promoting factor activity. *Nature* 339, 280–286.
- Nikolic, M., Dudek, H., Kwon, Y.T., Ramos, Y., Tsai, L.-H., 1996. The cdk5/p35 kinase is essential for neurite outgrowth during neuronal differentiation. *Genes & Development* 10, 816–825.
- Noble, M.E., Endicott, J.A., Brown, N.R., Johnson, L.N., 1997. The cyclin box fold: Protein recognition in cell-cycle and transcription control. *Trends in Biochemical Sciences* 22, 482–487.
- Nurse, P., 2000. A long twentieth century of the cell cycle and beyond. *Cell* 100, 71–78.
- Nurse, P., Bissett, Y., 1981. Gene required in G1 for commitment to cell cycle and in G2 for control of mitosis in fission yeast. *Nature* 292, 558–560.
- Ohtani, K., Degregori, J., Nevins, J.R., 1995. Regulation of the cyclin E gene by transcription factor E2F1. *Proceedings of the National Academy of Sciences of the United States of America* 92, 12146–12150.
- Okuda, M., Horn, H.F., Tarapore, P., *et al.*, 2000. Nucleophosmin/B23 is a target of CDK2/cyclin E in centrosome duplication. *Cell* 103, 127–140.
- Pagano, M., Pepperkok, R., Verde, F., Ansorge, W., Draetta, G., 1992. Cyclin A is required at two points in the human cell cycle. *EMBO Journal* 11, 961–971.
- Pavletich, N.P., 1999. Mechanisms of cyclin-dependent kinase regulation: Structures of cdks, their cyclin activators, and cip and INK4 inhibitors^{1,2}. *Journal of Molecular Biology* 287, 821–828.
- Peng, J., Zhu, Y., Milton, J.T., Price, D.H., 1998. Identification of multiple cyclin subunits of human P-TEFb. *Genes & Development* 12, 755–762.
- Peters, J.-M., 2002. The anaphase-promoting complex: Proteolysis in mitosis and beyond. *Molecular Cell* 9, 931–943.
- Petersen, B.O., Lukas, J., Sorensen, C.S., Bartek, J., Helin, K., 1999. Phosphorylation of mammalian CDC6 by cyclin A/CDK2 regulates its subcellular localization. *EMBO Journal* 18, 396–410.
- Ren, B., Cam, H., Takahashi, Y., *et al.*, 2002. E2F integrates cell cycle progression with DNA repair, replication, and G(2)/M checkpoints. *Genes & Development* 16, 245–256.
- Rickett, P., Seghezzi, W., Shanahan, F., Cho, H., Lees, E., 1996. Cyclin C/CDK8 is a novel CTD kinase associated with RNA polymerase II. *Oncogene* 12, 2631–2640.
- Roy, R., Adamczewski, J.P., Seroz, T., *et al.*, 1994. The MO15 cell cycle kinase is associated with the TFIH transcription-DNA repair factor. *Cell* 79, 1093–1101.
- Satyanarayanan, A., Kaldis, P., 2009. Mammalian cell-cycle regulation: Several Cdks, numerous cyclins and diverse compensatory mechanisms. *Oncogene* 28, 2925–2939.
- Sherr, C.J., 1993. Mammalian G¹ cyclins. *Cell* 73, 1059–1065.
- Sherr, C.J., 1994. G1 phase progression: Cycling on cue. *Cell* 79, 551–555.
- Sherr, C.J., 2001. The INK4a/ARF network in tumour suppression. *Nature Reviews Molecular Cell Biology* 2, 731–737.
- Sherr, C.J., McCormick, F., 2002. The RB and p53 pathways in cancer. *Cancer Cell* 2, 103–112.
- Sherr, C.J., Roberts, J.M., 1999. CDK inhibitors: Positive and negative regulators of G1-phase progression. *Genes & Development* 13, 1501–1512.
- Shiekhhattar, R., Mermelstein, F., Fisher, R.P., *et al.*, 1995. Cdk-activating kinase complex is a component of human transcription factor TFIH. *Nature* 374, 283–287.
- Solomon, M.J., Harper, J.W., Shuttleworth, J., 1993. CAK, the p34cdc2 activating kinase, contains a protein identical or closely related to p40MO15. *EMBO Journal* 12, 3133–3142.
- Swift, H., 1950. The constancy of deoxyribose nucleic acid in plant nuclei. *Proceedings of the National Academy of Sciences of the United States of America* 36, 643–654.
- Wei, P., Garber, M.E., Fang, S.-M., Fischer, W.H., Jones, K.A., 1998. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. *Cell* 92, 451–462.
- Weinberg, R.A., 1995. The retinoblastoma protein and cell cycle control. *Cell* 81, 323–330.
- Won, K.-A., Xiong, Y., Beach, D., Gilman, M.Z., 1992. Growth-regulated expression of D-type cyclin genes in human diploid fibroblasts. *Proceedings of the National Academy of Sciences* 89, 9910–9914.
- Zhao, J., Kennedy, B.K., Lawrence, B.D., *et al.*, 2000. NPAT links cyclin E–Cdk2 to the regulation of replication-dependent histone gene transcription. *Genes & Development* 14, 2283–2297.
- Zheng, D., Cho, Y.-Y., Lau, A.T., *et al.*, 2008. Cyclin-dependent kinase 3-mediated activating transcription factor 1 phosphorylation enhances cell transformation. *Cancer Research* 68, 7650–7660.

The Restriction Point

JA DeCaprio, Dana-Farber Cancer Institute, Boston, MA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Competence A quiescent cell is induced to enter into a state after a brief pulse of growth factors that allows entrance into S phase after a second pulse of growth factors.

DREAM A multi-protein complex that includes DP, RB-related, E2F, and MuvB and plays a direct role in establishing and maintaining the quiescent state.

G1/S The period in G1 after a cell has passed the restriction point but before DNA synthesis begins.

G1/S checkpoint A DNA damage checkpoint induced by activation of p53 and p21.

p53 A major cellular tumor suppressor (TP53) that is involved in many processes including apoptosis, DNA damage response, and cellular senescence.

Quiescence (G0) A viable state when normal cells are cultured without growth factors and remain capable of reentering the cell cycle when growth factors are restored.

RB The product of the retinoblastoma tumor suppressor gene RB1 that regulates E2F dependent gene expression.

RB-related proteins p107 (RBL1) and p130 (RBL2).

Restriction point A point in the G1 phase of the cell cycle when progression into S phase becomes independent of growth factors.

Senescence A viable but permanent state when cells have become committed to entering S phase but are unable due to increased p53, p21, and p16 levels.

Terminal differentiation A viable but permanent state when cells have exited from the cell cycle and display a differentiated phenotype.

The Restriction Point

The mammalian cell cycle is divided into four phases that include mitosis (M), deoxyribonucleic acid (DNA) synthesis (S), G1, and G2. Mitosis is recognized when cells visibly undergo cell division and includes prophase, metaphase, anaphase, and telophase. During S phase, cells duplicate their entire genome by DNA replication. G2 occurs after DNA synthesis has been completed but before chromosomal condensation in mitosis. The G1 phase occurs immediately after mitosis has been completed and before DNA synthesis begins. While the duration of S, G2, and M phases is relatively constant from one cell cycle to the next, there can be large variability in the duration of G1. One of the first observations regarding generation time was that by varying the growth conditions the length of a cell division cycle could change with the duration of G1 responsible for most of this variability. For example, while cells progressed through S, G2, and M phases in relatively invariable time periods, the length of the G1 phase was highly variable and this variability was dependent on the presence of growth factors (Sisken and Morasca, 1965). An important question that this observation raised was whether distinct points within the G1 phase of the cell cycle can be identified.

In 1974, Arthur Pardee published the first report on the restriction point (Pardee, 1974). While a scholar of the American Cancer Society and funded by the NIH, he performed a series of experiments at the Imperial Cancer Research Fund Laboratories in London. It was known that normal cells could be shifted from a proliferative to a quiescent state if serum or nutrients were removed from the culture. Quiescent cells could respond to growth factor addition and begin to proliferate again. However, it was not known whether the quiescent cells had arrested at a distinct point in G1 phase or had they exited from the cell cycle and entered into a new

quiescent phase. To address this question and determine whether distinct points within the G1 phase could be identified, he performed a series of experiments shifting cells from one nutrient deprivation condition to another and observed the length of time it would take for the cells to begin DNA synthesis. He determined that cells starved of serum, amino acids (isoleucine or glutamine), or phosphate or cells cultured at high density were positioned at the same point in the G1 phase relative to S phase. When he shifted cells from one nutrient starved condition to another he found that cells would not enter S phase until complete media was given. Similarly, contact-inhibited cells could not enter S phase if they were re-plated in low serum or amino-deficient media. However, if quiescent or contact-inhibited cells were fed with serum or nutrients for 9 h, they would advance into S phase and into mitosis even after serum or nutrients were removed, revealing a point during G1 when cells no longer needed growth factors to enter S phase. Pardee defined the restriction point for normal cells as a point when they were no longer dependent on nutrients to begin DNA synthesis (Figure 1(a)). He also correctly predicted that cancer cells underwent changes that lost the dependency on growth factors and were not dependent on the restriction point.

The concept that growth factors were required to induce a cell to leave the quiescent state and enter S phase was advanced by demonstrating that distinct growth factors contributed to this ability. Using BALB/c 3T3 cells that were particularly sensitive to contact inhibition, Pledger and colleagues observed that addition of a factor enriched from platelets could induce quiescent cells to advance to a state competent for DNA synthesis (Pledger *et al.*, 1977). In contrast, platelet-poor plasma was not capable of inducing quiescent cells to enter S phase. Platelets provided a growth factor that could induce a state called competence that enabled cells to progress to S phase when fed with platelet-poor plasma.

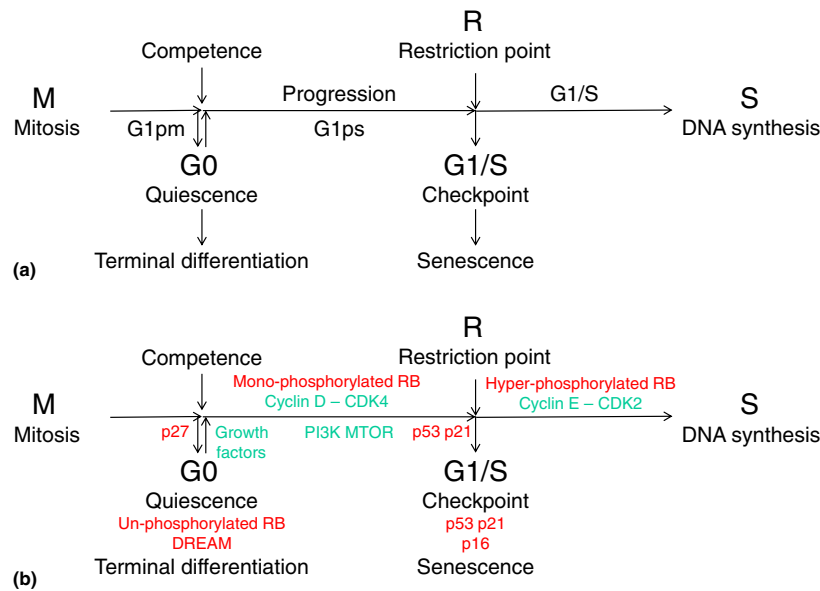


Figure 1 The G1 phase of the cell cycle. (a) Landmarks of events that occur between mitosis (M) and DNA synthesis (S). When normal cells exit mitosis, they progress through a brief period post mitosis (G1pm) when they are sensitive to growth factors. Absence of growth factors during G1pm results in entry into quiescence known as G0. A brief pulse of growth factors will enable a quiescent cell to move to competence. Addition of growth factors to competent cells enables progression to S phase. The restriction point is reached when cells are no longer dependent on growth factors to enter S phase. The G1ps period extends from quiescence until S phase. The G1/S phase refers to the time in G1 after cells have past the restriction point but before they synthesize DNA. The G1/S checkpoint occurs when cells are fully committed to entering S phase but are unable due to DNA damage signaling. Cells can permanently exit the cell cycle but remain viable by undergoing terminal differentiation or entering into senescence. (b) Factors regulating events during the G1 phase of the cell cycle. The CDK inhibitor p27 is expressed during quiescence and inhibits CDK activity. Un-phosphorylated RB and the DREAM complex repress E2F-dependent gene expression during quiescence. Addition of growth factors to a quiescent cell is necessary to become competent to re-enter the cell cycle. The second addition of growth factors induces cyclin D-CDK4/6 activity that mono-phosphorylates RB, activates PI3K signaling, and represses p53 activity. Cyclin E-CDK2 activity hyper-phosphorylates RB thereby activating E2F-dependent gene expression enabling DNA synthesis. DNA damage can activate p53 and p21 to inhibit CDK2 activity and force entry into the G1/S checkpoint. Cells that repair the DNA damage can reenter the cell cycle. Cells that cannot repair the DNA damage enter into a senescent state with an increase in p53, p21, and p16 levels.

This platelet-derived growth factor was identified as PDGF and shown to be capable of inducing the competent state (Figure 1(a); Pledger *et al.*, 1978). Platelet-poor plasma was capable of advancing competent cells to S phase in a phase called progression (Figure 1(a)). Epidermal growth factor (EGF) and insulin could replace platelet-poor plasma by promoting progression of cells made competent by PDGF. Other labs demonstrated that distinct cell types could use PDGF or other growth factors to induce progression and competence. However, it was noted that not all factors were interchangeable. For most cell types, EGF or insulin was not sufficient to induce a quiescent cell to become competent.

In addition to the decision for a quiescent cell to reenter the cell cycle, it was not known where in G1 would proliferating cells exit the cell cycle when growth factors were removed from the culture. Anders Zetterberg and colleagues used live cell imaging of proliferating Swiss 3T3 cells to determine sensitivity of cellular protein and DNA synthesis after serum withdrawal (Zetterberg and Larsson, 1985). They observed that removal of serum for 1 h at any point within 3 h of completing mitosis significantly extended the time to enter into S phase. They defined this period of sensitivity to serum withdrawal as G1pm (post mitosis) (Figure 1(a)). Notably, EGF or insulin but not PDGF could substitute for serum during the G1pm period indicating that progression but not competence

growth factors could prevent entry into the quiescent state. Once past this point, cells were no longer sensitive to serum withdrawal and could enter into S phase. If growth factors were removed during the G1pm period after mitosis, the cells entered a quiescent state. These quiescent cells could reenter the cell cycle when stimulated with PDGF plus EGF and insulin. However, the length of time for these quiescent or G0 cells to reach S phase was significantly longer than for cells that remained in complete media during the critical G1pm period. The period between quiescence and S phase was termed G1ps (pre-DNA synthesis) (Figure 1(a)). Furthermore, they determined that nearly all of the variability in the length of G1 could be accounted for by the G1ps interval.

The next advance in the restriction point concept was reported in 2001 when it was demonstrated that the requirement for prolonged continuous exposure to growth factors could be replaced with two short intervals of treatments separated by several hours. Although addition of PDGF for 30 min was sufficient to induce a quiescent cell to enter the competent state, the progression growth factors had been added for approximately 7–9 h before reaching the restriction point when growth factors could be removed and continued to S phase. Andrius Kazlauskas and colleagues observed that if quiescent NIH 3T3 or HepG2 cells were treated continuously with PDGF for 10 h, it was sufficient to reach the restriction

point and become committed to the S phase. They analysed how long after these cells were treated with an initial pulse of PDGF did they remain sensitive to addition of a second pulse of PDGF. They observed that cells made competent remained responsive to the second pulse for 8 h and would enter into S phase soon afterwards. If the second pulse was not given for 12 h after the first pulse, S phase entry was delayed for 14 h or as long as it would take if cells had been continuously treated (Jones and Kazlauskas, 2001). They also confirmed the studies of Pledger and colleagues that BALB/c 3T3 cells required PDGF treatment during the first pulse but required EGF or insulin for the second pulse. Furthermore, they confirmed that an initial single pulse with both competence and progression factors were insufficient to induce DNA synthesis. They noted that combined treatment with competence (PDGF) and progression (insulin) factors reduced the time to enter the S phase only by the time lag in the interval between the first and second pulse. They concluded that the competent state was rapidly induced by the initial treatment with competence factors but that it took several hours for competent cells to become responsive to the second pulse of progression factors.

At the same time these studies on the restriction point were taking place, many important discoveries in the RAS and mitogen-activated protein kinase (MAPK) signaling pathways triggered by PDGF, EGF, and insulin were made. In addition, significant insights were reported on the role of MYC in promoting cell growth and the cyclins and cyclin-dependent kinases (CDK) in enabling cell cycle progression. Kazlauskas and colleagues observed that the initial treatment of PDGF triggered immediate tyrosine phosphorylation of the PDGF receptor beta (PDGFRB) but that it quickly vanished after the initial pulse of PDGF was removed. RAS and MAP kinase signaling was also rapidly activated by PDGF treatment. However, PDGF treatment did not immediately lead to increased levels of MYC and cyclin D1 (CCND1). Instead, MYC and cyclin D were detectable soon after the addition of the second pulse of growth factors (Jones and Kazlauskas, 2001; Figure 1(b)). They observed that activation of MAP kinase signaling was required during the first pulse. Treatment with an MEK inhibitor (PD098059) during the first PDGF pulse inhibited entry into S phase even when the second pulse was given without the MEK inhibitor. However, MEK activation was not sufficient to drive cells into S phase. Expression of a constitutively active form of MEK1 (MAP2K1) was not sufficient to drive growth factor-starved cells into S phase (Jones and Kazlauskas, 2001). Notably, constitutively active MEK1 was capable of inducing cyclin D1 expression and kinase activity (Cheng *et al.*, 1998). This indicates that cyclin D expression was also not sufficient to drive cells from quiescence to S phase. However, overexpression of MYC combined with activated MEK1 could induce serum-starved cells to enter S phase.

The Kazlauskas group also explored the signaling events that occurred after the second pulse of growth factors. They determined that introduction of synthetic phosphatidylinositol 3-kinase (PI3K) lipid products could substitute for the second PDGF pulse but not the first. Therefore, the competent state induced by the first pulse of PDGF treatment induced a cellular state sensitive to PI3K signaling (Figure 1(b)). Deciphering the events triggered by the growth factor addition at these two distinct timepoints led to insights into key signaling pathways.

The molecular and signaling events during the two pulses was further explored by the Yarden group (Zwang *et al.*, 2011). Using a clone of normal human mammary epithelial cells, they determined that two 1-hour pulses of EGF given 7 h apart was sufficient to induce serum-starved cells to enter S phase and complete cell division. They observed that while ERK activation occurred immediately after the first and second pulses, the amplitude of the signaling events was much stronger after the second pulse compared to the first pulse. In these cells, the MAPK inhibitor U0126 given during the second pulse could block entry into S phase. In addition, PI3K signaling as evidenced by AKT and S6 phosphorylation was elevated after the second pulse. Inhibition of PI3K and MTOR (mechanistic target of rapamycin) activity with LY294002 during the second pulse could block entry into S phase (Figure 1(b)). This group also performed messenger RNA (mRNA) expression profiling of cells treated with the double pulse protocol. The first pulse induced the gradual and sustained expression of 35 genes involved in lipid biosynthesis that remained elevated until S phase entry. Treatment of cells with an inhibitor of HMG-CoA reductase blocked DNA synthesis indicating that lipid biosynthesis was required during the progression phase. A relatively recent review of the restriction point explores the contributions of biosynthetic events that occur during the progression phase and restriction point in G1 (Foster *et al.*, 2010).

In addition, the second pulse of growth factors specifically down-regulated a second set of genes that were activated by the initial pulse of EGF. Importantly, their down-regulation was dependent on PI3K but not ERK signaling. Among these down-regulated genes was p27 (CDKN1B) or CDK inhibitor. In addition, seven of the ten down-regulated genes were known to be responsive to p53 activation (Zwang *et al.*, 2011; Allen *et al.*, 2014). Importantly, RNAi (RNA interference) knockdown of p53 abrogated the need for the second pulse both in the mammary epithelial cells used as well as in NIH 3T3 cells used by the Kazlauskas lab. These results indicate that MAPK, PI3K, lipid biosynthesis, MYC, p53, and CDK played critical roles in establishing the restriction point (Figure 1(b)). Importantly, they also provided some insight into the events that occurred during progression that enabled a competent cell to be able to pass the restriction point.

In addition to studies of cell lines in culture, there are examples of cells *in vivo* with characteristics of competence that are primed to enter into S phase and proliferate more rapidly in response to a second signal. For example, muscle satellite cells represent the stem cells in adult mice that are capable of regenerating skeletal muscle cells in response to appropriate stimuli. Normally, the satellite stem cells reside in a quiescent state capable of regeneration. They can be induced to enter an activated state by injury to tissue at a distant site through hepatocyte growth factor (HGF) signaling to the MET receptor tyrosine kinase (Rodgers *et al.*, 2014). The activated stem cells were much more likely to enter S phase than quiescent stem cells and also entered S phase much more rapidly. Furthermore, these activated stem cells were larger and had significantly increased ATP levels and mitochondrial activity than quiescent stem cells consistent with elevated PI3K and MTOR activity. Thus, these activated stem cells appeared to have activated PI3K signaling *in vivo* similar to competent cells *in vitro*.

The various studies that explored the restriction point were initially focused on identifying landmarks within the G1 phase of the cell cycle that could account for the requirement for growth factors to advance a normal cell into S phase. In the 40 years that have passed since the initial description of the restriction point, many important insights revealed the signaling events that contribute to proliferation and growth. In addition to the key contributions of signaling by MAPK, PI3K, and MYC to enable cell growth, it also became clear that there were restraining activities that could inhibit cell cycle entry and progression.

The Role of RB and p53 in Establishing the Restriction Point

The product (RB) of the retinoblastoma tumor suppressor gene (RB1) is expressed in all normal cells including quiescent and proliferating cells. RB functions in part to repress the E2F transcription factors. There are two classes of E2F genes that are responsive to RB. The activating E2Fs including E2F1, E2F2, and E2F3 promote expression of a large class of genes required for DNA synthesis. The repressor E2Fs include E2F4 and E2F5 and serve to repress E2F-dependent genes during quiescence. The activating and repressing E2Fs bind to the dimerization partner DP (TFDP1-3) to form sequence-specific DNA binding transcription factors. RB binding to the activating E2Fs serves to repress their transcriptional activity. RB can also bind to the repressor E2Fs. In addition to RB, p107 (RBL1) and p130 (RBL2) are two RB-like proteins that have strong preference for binding to the repressor E2F4 and E2F5. While bound to a repressor E2F, p130 and p107 also bind to a five-protein complex termed MuvB or LINC to form the DREAM (DP, RB-like, E2F and MuvB) complex (Sadasivam and DeCaprio, 2013). During quiescence, the DREAM complex functions to repress all E2F-dependent genes including genes induced by the activating E2Fs during the G1/S transition as well as genes expressed during the G2 and M phases of the cell cycle that are dependent on B-MYB (MYBL2) and FOXM1 (Figure 1(b)).

The repressor activity of the RB proteins is regulated by CDKs. During G1, cyclin D (CCND1, CCND2, and CCND3) heterodimerizes with CDK4 or CDK6 to form the active kinase. Cyclin D expression is dependent on RAS and MYC signaling (Yu *et al.*, 2005). As mentioned above, cyclin D levels increase during the progression phase of G1 (Jones and Kazlauskas, 2001). Later in G1, cyclin E (CCNE1) levels increase. Cyclin E binds specifically to CDK2 and also contributes to RB phosphorylation. Cyclin E-CDK2 phosphorylates RB on multiple CDK consensus sites resulting in release of hyper-phosphorylated RB from the activating E2Fs thereby enabling E2F-dependent gene expression. Once past this point, growth factors are no longer required for S phase entry. Therefore, expression and activation of cyclin E-CDK2 results in the hyper-phosphorylation of RB that is required to pass the restriction point (Figure 1(b)) (Planas-Silva and Weinberg, 1997).

Recently, it was reported that cyclin D-CDK4 could specifically phosphorylate RB on any one of 15 different serine or threonine residues that form a CDK consensus site (S/T)P. Strikingly, in the presence of active cyclin D-CDK4, a given molecule of RB is only phosphorylated on one of these 15 sites (Narasimha *et al.*, 2014). Apparently, mono-phosphorylation

of RB by cyclin D-CDK4 precludes additional phosphorylation by cyclin D-CDK4. In contrast, cyclin E-CDK2 can phosphorylate a single molecule of RB on many of these (S/T)P sites resulting in a hyper-phosphorylated RB (Figure 1(b)). Notably, mono-phosphorylated RB (cyclin D-CDK4 dependent) could bind to the activating E2Fs and repress their transcriptional activity while hyper-phosphorylated RB (cyclin E-CDK2 dependent) could not. Furthermore, it was demonstrated that mono-phosphorylated RB was specifically capable of repressing the activating E2Fs when cells had undergone DNA damage signaling. Although some of these models need to be further validated, the concept of singly phosphorylated RB, previously referred to as under- or hypo-phosphorylated RB, brings a new perspective to how RB responds to cyclin D-CDK4 during G1.

In any case, it is likely that un-phosphorylated RB as well as the DREAM complex is present in quiescent cells. However, once cells have been rendered competent for cell cycle entry, cyclin D-CDK4 activity promotes the mono-phosphorylation of RB during the progression phase. Later, cyclin E-CDK2 hyper-phosphorylates RB thereby releasing it from the activating E2Fs and enabling expression of genes required for DNA synthesis. Thus cyclin E activity directed toward RB enables a cell to pass the restriction point and become committed to cell cycle entry. This period of G1 when cells have committed to S phase entry after the restriction point, when RB is hyper-phosphorylated, but before DNA synthesis begins is referred to as the G1/S phase of the cell cycle (Figure 1(b)).

In addition to cyclins and CDKs, the CDK inhibitors play an important role in establishing and regulating the restriction point. There are two general classes of CDK inhibitors: the CDKN2 inhibitors including p16 INK4a (CDKN2A), CDKN2B, CDKN2C, and CDKN2D which inhibit CDK4/6 by preventing their binding to cyclin D, and the p21 (CDKN1A) class that includes p27 (CDKN1B) and p57 (CDKN1C) which bind to and inhibit kinase activity of the cyclin D-CDK4/6 and cyclin E-CDK2 complexes. During quiescence, p27 is highly expressed and binds to cyclin D-CDK4 to inhibit its kinase activity. p27 Levels decrease when cells are made competent for cell cycle progression through a pulse of PDGF (Agrawal *et al.*, 1996). As cyclin D levels increase during the progression phase they titrate the available p27, thereby enabling cyclin E-CDK2 to become active. Therefore, p27 plays a critical role in establishing quiescence and competence through its inhibition of cyclin D-CDK4/6 (Figure 1(b); Coats *et al.*, 1996). After cells progress past the restriction point, p27 becomes phosphorylated by cyclin E-CDK2 followed by ubiquitination by the SCF (SKP1, CUL1, F-Box)–SKP2 complex and subsequent proteasome mediated destruction during G1/S phase (Vlach *et al.*, 1997; Carrano *et al.*, 1999). In addition, p27 can become deactivated during the G1 progression phase when it is transported to the cytoplasm and undergoes ubiquitination by a distinct ubiquitin ligase (Kamura *et al.*, 2004).

The p21 CDK inhibitor also plays a critical role during the G1 phase of the cell cycle. Elevated p21 levels can bind to and inactivate cyclin E-CDK2 (Harper *et al.*, 1993). p21 Undergoes a significant increase in levels and fold change in response to p53 activation during cellular stress (Allen *et al.*, 2014). For example, shortened telomeres or double-stranded DNA breaks will activate ataxia telangiectasia mutated (ATM) to phosphorylate p53 which in turn decreases its affinity for the MDM2 ubiquitin ligase and

increases it to the co-activators p300 (EP300) and CBP (CREBBP). Increased p53 levels can be induced by shortened telomeres. Indeed, many studies of the restriction point used immortalized cells that presumably have increased telomerase activity or reduced p53 activity. Thus, elimination of the p53 and p21 restraining activity can overcome the need for the second pulse of growth factors (Zwang *et al.*, 2011).

The G1/S DNA damage checkpoint can be viewed as a point in the cell cycle when the cell has become fully committed to enter into S phase and past the restriction point, but is unable to enter S phase because cyclin E-CDK2 and cyclin D-CDK4 are inactivated by the p21 and p16, respectively and RB remains capable of binding to and repressing the activating E2Fs, thereby decreasing levels of factors required for DNA synthesis (Figure 1(b)). The DNA damage that induces p53 can be derived from a variety of sources. If the DNA damage or stress signals are repaired, cells can exit the G1/S checkpoint and reenter the cell cycle. If the damage is not repaired in a timely manner, cells will enter a senescent state where they remain viable but not capable of reentering the cell cycle. RB and p53 participate in establishing the senescent state. During senescence, cells have committed to proliferation and presumably have passed the restriction point with activation of the ERK signaling pathway but perhaps have not integrated PI3K signaling resulting in activation of p53 (Figure 1(b)).

In this discussion of the restriction point, emphasis was made on the events that occur in normal cells in response to growth factors. The timely and integrated coordination of signaling events is required for a normal cell to reenter the cell cycle from quiescence. The signaling pathways downstream of receptor tyrosine kinases, MAPK, PI3K, and CDKs combined with the restraining activities of RB and p53 must be highly regulated to enable cell cycle progression. These events are highly regulated in normal fibroblasts in culture but also in all of the many cell types present in tissues. Furthermore, components of these pathways are nearly always perturbed in cancer. Therefore, study of the events that contribute to establishment of the restriction point as well as what is required to enable cellular proliferation will continue to bring insights into normal development and disease processes.

Conclusions

1. The concept of the restriction point has evolved considerably since its initial description.
2. The restriction point occurs at a point in G1 when a normal cell no longer requires growth factors to enter S phase.
3. Cells can exit the cell cycle if cultured without growth factors during a brief period after mitosis.
4. Exit from the cell cycle can lead toward quiescence and terminal differentiation.
5. Growth factors at two steps are required to pass the restriction point.
6. The first step promotes RAS and MAP kinase signaling and leads to competence.
7. Competent cells are capable of slowly activating biosynthesis pathways.
8. The second step involves cyclin D, mono-phosphorylation of RB, and PI3K signaling.

9. Competence activates while progression deactivates p53 signaling.
10. Once cyclin E is activated and RB becomes hyper-phosphorylated, the cells have entered G1/S and made the commitment to DNA synthesis.

See also: Functional Cell Biology: An Overview

References

- Agrawal, D., Hauser, P., Mcpherson, F., *et al.*, 1996. Repression of p27kip1 synthesis by platelet-derived growth factor in BALB/c 3T3 cells. *Molecular and Cellular Biology* 16, 4327–4336.
- Allen, M.A., Andrysiak, Z., Dengler, V.L., *et al.*, 2014. Global analysis of p53-regulated transcription identifies its direct targets and unexpected regulatory mechanisms. *eLife* 3, e02200.
- Carrano, A.C., Eytan, E., Hershko, A., Pagano, M., 1999. SKP2 is required for ubiquitin-mediated degradation of the CDK inhibitor p27. *Nature Cell Biology* 1, 193–199.
- Cheng, M., Sexl, V., Sherr, C.J., Roussel, M.F., 1998. Assembly of cyclin D-dependent kinase and titration of p27Kip1 regulated by mitogen-activated protein kinase kinase (MEK1). *Proceedings of the National Academy of Sciences of the United States of America* 95, 1091–1096.
- Coats, S., Flanagan, W.M., Nourse, J., Roberts, J.M., 1996. Requirement of p27Kip1 for restriction point control of the fibroblast cell cycle. *Science* 272, 877–880.
- Foster, D.A., Yellen, P., Xu, L., Saqcena, M., 2010. Regulation of G1 cell cycle progression: Distinguishing the restriction point from a nutrient-sensing cell growth checkpoint(s). *Genes Cancer* 1, 1124–1131.
- Harper, J.W., Adami, G.R., Wei, N., Keyomarsi, K., Elledge, S.J., 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* 75, 805–816.
- Jones, S.M., Kazlauskas, A., 2001. Growth-factor-dependent mitogenesis requires two distinct phases of signalling. *Nature Cell Biology* 3, 165–172.
- Kamura, T., Hara, T., Matsumoto, M., *et al.*, 2004. Cytoplasmic ubiquitin ligase KPC regulates proteolysis of p27(Kip1) at G1 phase. *Nature Cell Biology* 6, 1229–1235.
- Narasimha, A.M., Kaulich, M., Shapiro, G.S., *et al.*, 2014. Cyclin D activates the Rb tumor suppressor by mono-phosphorylation. *eLife* 3, e02872.
- Pardee, A.B., 1974. A restriction point for control of normal animal cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America* 71, 1286–1290.
- Planas-Silva, M.D., Weinberg, R.A., 1997. The restriction point and control of cell proliferation. *Current Opinion in Cell Biology* 9, 768–772.
- Pledger, W.J., Stiles, C.D., Antoniades, H.N., Scher, C.D., 1977. Induction of DNA synthesis in BALB/c 3T3 cells by serum components: Reevaluation of the commitment process. *Proceedings of the National Academy of Sciences of the United States of America* 74, 4481–4485.
- Pledger, W.J., Stiles, C.D., Antoniades, H.N., Scher, C.D., 1978. An ordered sequence of events is required before BALB/c-3T3 cells become committed to DNA synthesis. *Proceedings of the National Academy of Sciences of the United States of America* 75, 2839–2843.
- Rodgers, J.T., King, K.Y., Brett, J.O., *et al.*, 2014. mTORC1 controls the adaptive transition of quiescent stem cells from G0 to G(alert). *Nature* 510, 393–396.
- Sadasivam, S., DeCaprio, J.A., 2013. The DREAM complex: Master coordinator of cell cycle-dependent gene expression. *Nature Reviews Cancer* 13, 585–595.
- Sisken, J.E., Morasca, L., 1965. Intrapopulation kinetics of the mitotic cycle. *Journal of Cell Biology* 25, 179–189.
- Vlach, J., Hennecke, S., Amati, B., 1997. Phosphorylation-dependent degradation of the cyclin-dependent kinase inhibitor p27. *EMBO Journal* 16, 5334–5344.
- Yu, Q., Ciernerych, M.A., Sicinski, P., 2005. Ras and Myc can drive oncogenic cell proliferation through individual D-cyclins. *Oncogene* 24, 7114–7119.
- Zetterberg, A., Larsson, O., 1985. Kinetic analysis of regulatory events in G1 leading to proliferation or quiescence of Swiss 3T3 cells. *Proceedings of the National Academy of Sciences of the United States of America* 82, 5365–5369.
- Zwang, Y., Sas-Chen, A., Drier, Y., *et al.*, 2011. Two phases of mitogenic signaling unveil roles for p53 and EGR1 in elimination of inconsistent growth signals. *Molecular Cell* 42, 524–535.

CDK Inhibitors in Normal and Malignant Cells

A Besser and J Slingerland, University of Miami Miller School of Medicine, Miami, FL, USA

© 2016 Elsevier Inc. All rights reserved.

Cyclins, CDKs, and Cell Cycle Progression

The mammalian cell cycle is composed of four phases directed by the actions of Cyclin-cyclin-dependent kinase (CDK) complexes. The intensity of mitogenic or antimitogenic signals received by the cell directs a binary decision to proliferate or not. A checkpoint occurs at a restriction point late in the first growth phase (G1), and is largely governed by the gatekeeper protein Rb. Once this checkpoint is passed, normal cells irreversibly commit to replicate their DNA and will divide if replication is successful and die if it is not. The Cyclin-CDK holoenzyme complex comprises a regulatory Cyclin and a catalytic CDK that, once active will phosphorylate its protein substrates. There are four major mammalian Cyclins: Cyclin E, A, D, and B, each of which has more than one isoform and the CDK family contains 11 members. Sequential, tightly regulated increases in Cyclin gene expression and protein stability lead to successive peaks of Cyclin-CDK activity that direct cell cycle progression (see Figure 1). CDK protein and RNA levels remain largely constant across the cell cycle (Sherr and Roberts, 1999).

In response to mitogenic signals, such as epidermal growth factor receptor (EGFR) or Ras activation, genes encoding D-type Cyclins are induced, and Cyclin D-CDK4/6 complexes assemble in the cytoplasm (Sherr, 1993). D-type Cyclins act as growth factor sensors to direct the cell to exit quiescence and proliferate under favorable conditions. In quiescent cells, hypophosphorylated retinoblastoma protein (pRb) sequesters transcription factors E2F1-4 (Dyson, 1998) and inhibits induction of target genes required for DNA replication in S phase. Once Cyclin D-CDK4/6 complexes translocate to the nucleus, they initiate phosphorylation and inactivation of pRb. Later in G1, activated Cyclin E-CDK2 complexes further

phosphorylate pRb mediating release and activation of E2F to allow progression into S phase. During the second growth phase (G2), Cyclin A- and Cyclin B-CDK2 complexes maintain pRb in a hyperphosphorylated state until the cell exits mitosis (reviewed in Sherr and Roberts, 1999).

There are two CDK inhibitor protein families: the inhibitor of CDK4/6 (INK4) family, which includes homologous proteins p15^{INK4b} (Hannon and Beach, 1994), p16^{INK4a} (Serrano et al., 1993), p18^{INK4c} (Guan et al., 1994; Hirai et al., 1995), and p19^{INK4d} (Hirai et al., 1995; Chan et al., 1995) and kinase inhibitor proteins (KIP), comprising p21^{CIP1}, p27^{KIP1}, and p57^{KIP2}. In states of terminal cellular differentiation or unfavorable growth conditions, several inhibitors cooperate to halt cell cycle progression at various stages. INK4 family members act early in G1 to bind and inhibit Cyclin D-CDK4/6 complexes. Members of the CIP/KIP family of inhibitors have a broader target range and act largely to bind Cyclin E-CDK2 in G1 phase and Cyclin A-CDK2 in S phase, but may also inhibit Cyclin B/A-CDK2 in G2 phase under situations of cellular stress or DNA damage (reviewed in Sherr and Roberts, 1999; Sherr, 1993; Kim and Sharpless, 2006; Gil and Peters, 2006). The following will briefly review the function and regulation of the best understood CDK inhibitors, p15^{INK4b}, p16^{INK4a}, p21, and p27 in normal cells, their deregulation and potential prognostic significance in human cancers and their potential as markers/predictors of oncogenic pathway activation and responsiveness to targeted therapies.

The INK4 Family

Normal Function

p15^{INK4b} and p16^{INK4a} bind to CDK4 and CDK6 and allosterically inhibit the binding of Cyclin D, thus preventing activation of the complex. INK4 proteins interact with CDK4 and CDK6 through several ankyrin repeats which mediate protein-protein interaction. p15^{INK4b} and p16^{INK4a} are encoded at the *INK4a/ARF/INK4b* locus on chromosome 9p21, which is a common site of deletion or mutational inactivation in cancers (reviewed in Kim and Sharpless, 2006; Gil and Peters, 2006). Identification of the p16^{INK4a} tumor suppressor protein was prompted by genetic linkage studies showing that an inherited predisposition to melanoma could be traced back to a deletion at the *INK4a/ARF/INK4b* locus (Kamb et al., 1994; Nobori et al., 1994). p15^{INK4b} was discovered as a TGF- β induced mediator of cell cycle arrest (Hannon and Beach, 1994). Interestingly, the ARF, or alternate reading frame protein is also encoded at this locus, but translated in a different reading frame. ARF is also a cell cycle inhibitor, but largely controls p53 function in response to aberrant growth or oncogenic stress and does not function as a Cyclin-CDK inhibitor (reviewed in Kim and Sharpless, 2006; Gil and Peters, 2006).

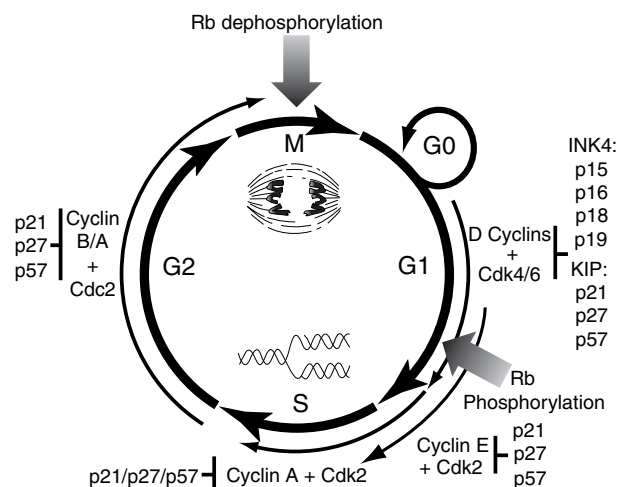


Figure 1 Sequential upregulation of Cyclin CDKs governs progression through the cell cycle. CDKs of the INK4 and CIP/KIP family critically regulate cell cycle transitions.

Regulation of p15^{INK4b} and p16^{INK4a} Levels

p15^{INK4b} and p16^{INK4a} are encoded by the *CDKN2B* and *CDKN2A* genes on the *INK4a/ARF/INK4b* locus, and both are tightly regulated by transcriptional mechanisms. The genes encoding p15^{INK4b} and p16^{INK4a} are induced in response to oncogenic Ras signaling, and bring about p16^{INK4a}-mediated senescence as a means of protecting cellular integrity. Ras activation leads to subsequent MEK/ERK activation, and ERK-mediated induction of Ets1/2 may mediate in part the induction and increase in p16 expression following Ras activation (Ohtani *et al.*, 2001). In response to TGF- β signaling, SMADs 2 and 4 induce *CDKN2B* to mediate growth arrest in TGF- β -responsive cells (Hannon and Beach, 1994). These genes are also subject to epigenetic regulation: to prevent premature senescence, the genes encoding p16^{INK4a}, p15^{INK4b}, and ARF are repressed by T-box proteins and by the polycomb group genes (PcG) BMI1, Cbx7, and Me18, which epigenetically silence genes by remodeling chromatin. Loss of BMI1 has been associated with reduced stem cell populations due to increased expression of p16^{INK4a} (reviewed in Kim and Sharpless, 2006; Gil and Peters, 2006). Methylation of the gene encoding p16^{INK4a} is significantly correlated with blastic bone marrow involvement, and sequential analyses have shown that *CDKN2B* gene methylation increases with myelodysplastic disease evolution toward AML. These data suggest that *CDKN2B* methylation may allow leukemic cells to escape inhibitory signals from the bone marrow environment (Quesnel *et al.*, 1998).

In normal hematopoietic cells, TGF- β signaling inhibits proliferation. Compared to solid malignancies, leukemic malignancies show more frequent *CDKN2B* gene methylation and inactivation compared to *CDKN2A*. Leukemic cells that lose sensitivity to TGF- β signaling generally have *CDKN2B* gene methylation, and this suggests that p15^{INK4b} inactivation is important for marrow blast proliferation in myeloid malignancies (reviewed in Quesnel and Fenaux, 1999).

Functions Independent of CDK Inhibition

Senescence/stem cell self-renewal

Several groups have shown that p16^{INK4a} levels increase markedly in aging tissues, providing evidence that pre-programmed senescence and aging are intimately linked to the cell machinery (Kim and Sharpless, 2006). p16^{INK4a} mediates G1 arrest during normal prostate epithelial cell senescence (Sandhu *et al.*, 2000). Aging is characterized by a reduced ability of stem cells to regenerate lost or damaged cells, and it has been suggested that an increase in p16^{INK4a} leads to the decline of replicative potential of stem cells (reviewed in Kim and Sharpless, 2006; Sherr and Roberts 1999; Warfel and El-Deiry 2013). Hemopoietic stem cells (HSCs) from *CDN2A* $-/-$ *ARF* $-/-$ mice have an increased ability to serially transplant and repopulate irradiated recipient mice (Stepanova and Sorrentino, 2005). *CDKN2A* knockout causes loss of the age-induced decline in proliferation in HSCs, neural stem cells, and pancreatic islets (reviewed in Kim and Sharpless, 2006).

Mouse models have provided powerful tools to elucidate CDK inhibitor function in development and tumorigenesis. *CDKN2A* $-/-$ animals develop normally, but form spontaneous tumors at an early age (Serrano *et al.*, 1996), while

p15^{INK4b} $-/-$ animals also develop normally, but do not display a predisposition to tumor formation (Roussel, 1999). Notably, mutations or selective deletions of the *CDKN2B* gene have not been frequently identified in solid human tumors. Thus, it appears that p15^{INK4b} may not be a tumor suppressor, but serves a redundant function to inhibit cell cycle.

The CIP/KIP Family

Normal Function

p21^{CIP1}, encoded by *CDKN1A*, and p27^{KIP1} encoded by *CDKN1B* (hereafter p21 and p27), are highly homologous proteins that regulate the cell cycle by binding directly to both their Cyclin partners, largely E and A-type Cyclins, and to the catalytic cleft of their target CDKs, acting as competitive inhibitors for CDK substrates. Both CDK and Cyclin binding motifs are located in the N-terminal regions of p21 and p27 and are highly conserved (reviewed in Sherr and Roberts, 1999). Efficient CIP/KIP-mediated complex inhibition requires inhibitor binding to both its target Cyclin and CDK simultaneously (Vlach *et al.*, 1997; Chu *et al.*, 2007). p21 and p27 both bear a conserved C-terminal nuclear localization sequence, ensuring that the newly translated protein is directed to CDK targets in the nucleus (Zeng *et al.*, 2000; Rodriguez-Vilarrupla *et al.*, 2002). p21 was discovered as a growth inhibitor in senescent cells (Noda *et al.*, 1994), in cells arrested by p53 in response to DNA damage (Dulic *et al.*, 1994; El-Deiry *et al.*, 1993), and through a yeast two-hybrid screen for inhibitors of CDK2 (Harper *et al.*, 1993). p21 is abundantly expressed in proliferating cells, and is usually reduced or absent in quiescent cells (see Sherr and Roberts, 1999 for review). p21 is a transcriptional target of p53 and is called into action to halt cell cycle progression in response to DNA damage. While p27 appears to govern cell cycle progression in all cells, p21 appears to play a CDK inhibitor role largely as a stress reader, and it is induced by DNA damage, hypoxia, and redox stressors (Sherr and Roberts, 1999). Interestingly, p21 null mice develop normally, although p21 null embryonic fibroblasts display a reduced ability to promote G1 arrest following DNA damage. This is consistent with a role for p21 primarily in checkpoint control, rather than as regulator of progenitor commitment.

p27 was first identified as a CDK2-inhibitor in TGF- β -treated cells and is regulated by growth inhibitory cytokines and by contact-inhibition (Koff *et al.*, 1993; Polyak *et al.*, 1994a,b; Slingerland *et al.*, 1994; Hengst *et al.*, 1994). Both p21 and p27 appear to play critical roles to govern tissue proliferation during development and are upregulated to mediate cell cycle arrest. Unlike p21, p27 is strongly expressed in nonproliferating cells and critically regulates both quiescence and G1 progression. Mice lacking p27 display multi-organ hyperplasia and increased body size, and haploinsufficient animals are susceptible to carcinogen-induced tumors. Thus, p27 acts to control both cell proliferation and tissue expansion (reviewed in Vidal and Koff, 2000). Knock-in mice expressing a Cyclin-CDK-binding defective mutant p27 (p27^{CK-}) show increases in progenitor cell self-renewal in multiple tissues including pituitary, intestine, and bronchial

epithelium. p27^{CK-}-knock-in animals have increased bronchoalveolar stem cells and develop spontaneous lung tumors (Besson *et al.*, 2007).

p21/p27 double null mice display higher levels of spontaneous pituitary adenomas, pheochromocytomas, and thyroid adenomas, and have delayed incidence of radiation-induced thymic lymphoma (Garcia-Fernandez *et al.*, 2011). p15/p21 null mice display higher fibrosarcomas and rhabdomyosarcomas (Martin-Caballero *et al.*, 2004). p16/p19/p27 null mice show a significant increase in mortality due to accelerated development and metastasis of T-cell lymphoma (Martin-Caballero *et al.*, 2004). In contrast, p16/p19/p21 null mice do not have increased mortality but do develop fibrosarcomas (Martin-Caballero *et al.*, 2004). These phenotypes demonstrate the oncogenic cooperation between defects in INK4 and CIP/KIP family members.

Regulation

For both p21 and p27, key phosphorylation events have been identified that regulate their levels, activities, and intracellular localization. Both p21 and p27 act in the nucleus to inhibit CDKs (Connor *et al.*, 2003; Cmielova and Rezacova, 2011). During G1 progression, different phosphorylation events regulate degradation of these CIP/KIPs. Extracellular regulated MAP kinase (ERK) phosphorylates p21 at threonine 57 (T57) and serine 130 (S130) to regulate p21 degradation (Hwang *et al.*, 2009). p21 stability is increased by phosphorylations at T145 and S146 by AKT, and by p38 or JNK (Jun N-terminal kinase) at S130 (Hwang *et al.*, 2009). There are at least two major pathways governing p27 proteolysis. In early G1, mitogens promote S10 phosphorylation of p27 (Boehm *et al.*, 2002; Deng *et al.*, 2004), which facilitates CRM1-mediated p27 nuclear export (Connor *et al.*, 2003; Ishida *et al.*, 2000, 2002; Rodier *et al.*, 2001) and subsequent degradation. The best characterized proteolytic mechanism governing p27 is mediated by SCF^{Skp2} (S-phase kinase associated protein 1 (SKP1)/Cullin/F-Box protein: S-phase kinase associated protein 2 (Skp2)) complex, whose binding to p27 is stimulated in late G1 by p27 phosphorylation by Cyclin E or A-Cdk2 at threonine 187 (T187) (Nakayama and Nakayama, 2006). While free, active Cyclin E or A-CDK2 can phosphorylate CDK-bound p27 at T187 efficiently *in vitro*, p27-bound CDK2 is catalytically inactive (Larrea *et al.*, 2008). p27 contains three tyrosines at residues 74, 88, and 89 within its Cdk-inhibitory domain. Y88 is part of a 310-helix in p27 that inserts into the CDK2 catalytic cleft and displaces ATP (Russo *et al.*, 1996). p27 phosphorylation by Src (Chu *et al.*, 2007) and Abl (Grimmler *et al.*, 2007) at tyrosine 74 (Y74), Y88, and Y89 causes its ejection from the catalytic cleft of Cdk2 (Chu *et al.*, 2007). Thus, p27^{Y88}-bound Cyclin E-Cdk2 is no longer inhibited and can phosphorylate p27 at T187 to activate SCF^{Skp2}-mediated p27 proteolysis in late G1/S (Chu *et al.*, 2007; Grimmler *et al.*, 2007). While Y88 is highly conserved among all CIP/KIP family CDK inhibitors, including human p21 and p57 proteins, it is not clear whether Src family kinase-mediated phosphorylation governs p21 and p57 stability in an analogous manner.

C-terminal phosphorylations also critically regulate levels and localization of both p21 and p27. p27 can be phosphorylated by PI3K-effector kinases, AKT, SGK, and RSK at

threonine residues T157 and T198 (Wander *et al.*, 2011). T157 of p27 is located within its nuclear localization sequence and phosphorylation at this site delays its nuclear import into the nucleus, shifting the pool of p27 toward cytoplasmic localization (Wander *et al.*, 2011). Phosphorylation at T198 stabilizes p27 in the cytoplasm (reviewed in Larrea *et al.*, 2009). Similarly, p21 can be phosphorylated by AKT at T145 (the site homologous to T157 in p27), leading to accumulation in the cytoplasm and loss of its CDK-inhibitory action in the nucleus (reviewed in Warfel and El-Deiry, 2013; Weiss, 2003). Growing evidence suggests that the cytoplasmic mislocalization of both of these proteins leads to pro-oncogenic gain of function (Wander *et al.*, 2011; Besson *et al.*, 2004, 2008) and these are reviewed below.

Although p27 mRNA levels are constant throughout G1-S phase, its translation decreases dramatically upon G0 exit (reviewed in Chu *et al.*, 2008); p21 mRNA levels, however, are tightly controlled by transcriptional regulators, most of which induce p21 transcription through Sp1-3 sites in the p21 promoter. These include p53, TGF- β , NGF (nerve growth factor), and BCRA1. c-Myc, however, inhibits p21 transcription to blunt p53-dependent induction of p21 in response to DNA damage. Under conditions of cellular stress, p21 mRNA can be stabilized to prevent inappropriate cell cycle entrance through binding of the transcription factor ZONAB to its 3' UTR. Following translation, the p21 protein is stabilized by HSP90, preventing its immediate degradation if otherwise left unbound (reviewed in Warfel and El-Deiry, 2013).

Functions Independent of CDK Inhibition

D-type Cyclin/CDK4/6 assembly and nuclear import

Paradoxically, p21 and p27 both act to enhance Cyclin-CDK complex formation and thereby promote proliferation. Cytoplasmic p21 and p27 both increase assembly of Cyclin D-CDK4/6 and stabilize these complexes (LaBaer *et al.*, 1997). Since neither CDKs4/6, nor D-type Cyclins bear a nuclear localization sequence (NLS), the associated KIP is required for nuclear import (reviewed in Warfel and El-Deiry, 2013; Weiss, 2003). Cyclin D-bound and Cyclin E-bound p27 have different phospho-isoform patterns (Ciarallo *et al.*, 2002) and p27 phosphorylation by AKT at T157 and T198 facilitates p27 assembly of Cyclin D-CDK4/6 (Larrea *et al.*, 2008; James *et al.*, 2007). While high p21 levels are required to inhibit G1-S Cyclin-CDKs, low p21 levels are sufficient to promote Cyclin D-CDK4 complex formation following mitogenic stimulation (Warfel and El-Deiry 2013; LaBaer *et al.*, 1997).

p21 enhances proliferation as a target of PI3K and was shown to mediate the pro-proliferative effects of IGF-1 (Dupont *et al.*, 2003). In addition, cytosolic, AKT-phosphorylated p21 increases proliferation in Her2/neu over-expressing cells (reviewed in Weiss, 2003). Introduction of an NLS-deficient p21 localized primarily to the cytoplasm, increased cell cycle progression in vascular smooth muscle cells and caused protection from apoptosis in monocytes. Furthermore, Paclitaxel treatment can increase cytosolic p21 levels in squamous carcinoma cells, thus resulting in increased proliferation (reviewed in Weiss, 2003). p27 also appears to promote tumor progression in PI3K activated cancers.

Pro-oncogenic actions of p27^{pT157pT198} in cancers are reviewed below.

Apoptosis

Following DNA damage, p53 is upregulated and directs either a cytostatic response, inducing p21 to halt proliferation, or a response leading to programmed cell death. Factors mediating this decision by p53 are not completely understood, however, Myc and its corepressor Miz-1 can be recruited to the p21 promoter to repress its transcription and switch the p53 response from cytostatic to apoptotic in response to DNA damage (Seoane *et al.*, 2002). p21 may also coordinate cell cycle arrest permitting time for DNA repair, by inhibiting the activation of pro-apoptotic factors p38, JNK, ASK1, and SAP (reviewed in Weiss, 2003). ETS-1 cooperates with p300 to activate p21 transcription and protect vascular smooth muscle cells from undergoing apoptosis (Zhang *et al.*, 2003). p21 itself is targeted during apoptosis, where caspase 3 cleaves off its NLS and relocates p21 to the cytoplasm to drive growth-arrested cells into apoptosis (Jin *et al.*, 2000).

Differentiation/senescence

Differentiation and senescence are both characterized by a stable exit from the cell cycle. While p27 activates differentiation pathways and p16^{INK4a} activates senescence pathways, p21 appears to play a role in both. The differentiation hormone vitamin D3, via its receptor, induces p21 transcription, and overexpression of both p21 and p27 can lead to terminal differentiation in monocytes/macrophages (Liu *et al.*, 1996). Cytoplasmic localization of p21 appears to drive differentiation in monocytes and neurites through poorly defined mechanisms (Tanaka *et al.*, 2002; Asada *et al.*, 1999). p21 can act as an initial, temporary mediator of senescence-like arrest in response to cellular stress or DNA damage, followed by a more permanent upregulation of p16^{INK4a} if the cell does not reenter cell cycle (Roninson, 2002; Muthna *et al.*, 2010). p21 is a key mediator of G1 arrest during the proliferative senescence of normal human mammary epithelial cells accompanying telomere shortening (Sandhu *et al.*, 2000; Noda *et al.*, 1994). p21 overexpression induces a senescence-like state in HeLa cells, and its attenuation inhibits senescence (reviewed in Weiss, 2003).

p27 appears to regulate progenitor expansion and differentiation in various normal tissue types. It is shown to be critical for xenopus neuronal differentiation (Vernon *et al.*, 2003, 2006; Nguyen *et al.*, 2006, 2007), and knockdown of Xic1, the p27 homolog in *Xenopus*, causes loss of muscle differentiation. Xic1 is required at a point prior to terminal muscle differentiation and cell cycle arrest, and may act through modulation of transcription. Xic1 increases the activity of myogenic transcription factor, MyoD (Vernon and Philpott, 2003), to drive myogenic commitment through cell cycle-independent actions (Vernon and Philpott, 2003; Messina *et al.*, 2005). Xic1 also plays a Cyclin-CDK-independent role to drive cardiac myotome differentiation (Movassagh and Philpott, 2008). During bone differentiation, p27 cooperates with p130 to promote endochondral ossification (Yeh *et al.*, 2007).

Transcriptional regulation

There is increasing data to suggest that both p21 and p27 may modulate gene expression by binding a number of transcription factors. Transfection of a Cyclin-CDK-binding defective p21 led to p21 binding to and inhibiting E2F1 activity, suggesting that p21 may function to modulate this important transcription factor. Following stimulation with leukemia inhibitor factor (LIF), p21 can bind and inhibit STAT3/CBP (CREB binding protein) transcriptional activity (reviewed in Warfel and El-Deiry, 2013).

p27 was recently shown to act as a transcriptional corepressor in a complex with other transcriptional repressors p130, E2F4, HDAC1, and SIN3A (Pippa *et al.*, 2012). Putative p27-repressed target genes identified by ChIP/CHIP assays included targets that govern cell proliferation and tissue expansion (Pippa *et al.*, 2012). Sox2 is a critical embryonic stem cell (ES) transcription factor that maintains ES self-renewal (Marson *et al.*, 2008) and may also drive cancer stem cell expansion (Bass *et al.*, 2009; Leis *et al.*, 2011). SOX2 is corepressed by p27 in that p27 binds the main SOX2 regulator site, SOX2-SRR2, located 4KB downstream of the coding exon to corepress SOX2 together with p130, E2F4, and SIN3A (Li *et al.*, 2012). Thus, p27 may carry out a spectrum of developmental roles via transcriptional actions. The gene targets of p27 that govern differentiation and how they may be disrupted in cancer remain to be elucidated.

Stem cell self-renewal

p21 and p27 may contribute to regulation of stem cell populations. p21 plays a role in maintaining a quiescent stem cell population and preventing epithelial-mesenchymal transition (EMT) and cancer stem cells (CSC) expansion *in vivo* in Ras- or c-Myc-driven tumorigenesis (reviewed in Warfel and El-Deiry, 2013). p21 knockout mice had increased hematopoietic stem cell proliferation, indicating that p21 may restrain HSC self-renewal. It is noteworthy that p27 null animals exhibit multi-organ hyperplasia and hypertrophy and spontaneous neoplasia (reviewed in Chu *et al.*, 2008), suggesting p27 normally acts to limit organ size in embryogenesis. Notably, this phenotype is phenocopied in p27^{CK-} knock-in, reflecting the loss of CDK-inhibitory action. In addition, in p27CK-knock-in animals, loss of CDK inhibition and a CDK-independent gain of p27 function appear to perturb normal stem cell homeostasis and cause expansion of progenitor/stem cells in several tissues, increased bronchoalveolar stem cells and spontaneous lung tumor formation (Besson *et al.*, 2007). These intriguing findings suggesting that p27 may have cell cycle-independent roles to regulate stem expansion in many tissues (Besson *et al.*, 2007).

Prognostic Significance of CDK Inhibitor Deregulation in Human Cancers

Deregulation of the CDK inhibitors is one of the hallmarks of cancer. While mutations/deletions of genes encoding p15^{INK4b}/p16^{INK4a} are frequent in hemopoietic leukemia, loss of p16^{INK4a} in solid tumors is also reported. Malignancies of many tissues exhibit loss of CIP/KIP proteins and genetic

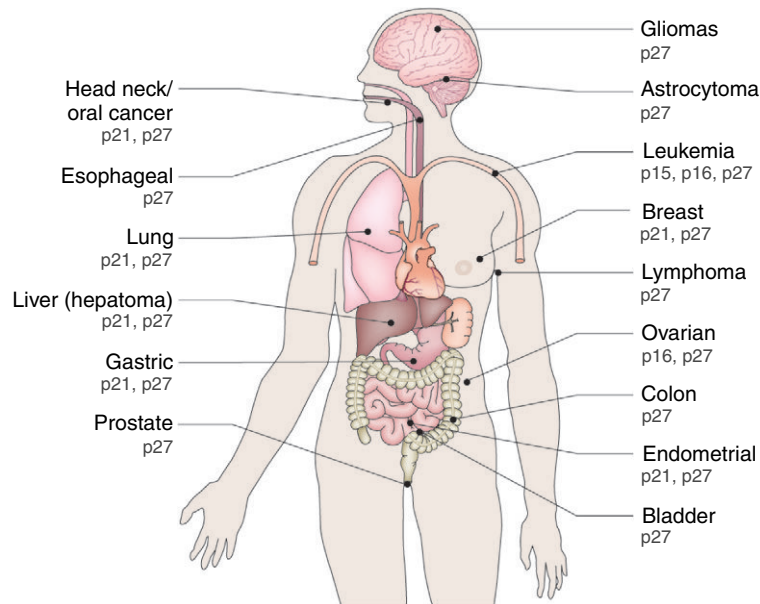


Figure 2 Schematic showing different human malignancies affected by changes in CDK inhibitors. The *INK4a* and *INK4b* alterations occur mostly at the genetic level (deletions/mutations), while changes in KIP/CIP levels and localization have potential for prognostic impact in cancers arising from the tissue sites shown.

changes leading to loss of $p15^{INK4b}/p16^{INK4a}$, and these are reviewed below (Figure 2).

INK4 Family

$p16^{INK4a}$ plays a role as a classic tumor suppressor, and homozygous deletion of chromosome 9p21, encoding the *INK4a/ARF/INK4b* locus, is one of the most frequent cytogenetic events in a wide variety of human malignancies (Kim and Sharpless, 2006; Beroukhi et al., 2010; Bignell et al., 2010). *INK4a/ARF/INK4b* locus deletion has been observed in melanoma (Kamb et al., 1994), pancreatic adenocarcinoma (Caldas et al., 1994; Wilentz et al., 1998; Rozenblum et al., 1997), glioblastoma (Moulton et al., 1995), acute lymphoblastic leukemia (Okuda et al., 1995), non-small cell lung cancer (Belinsky et al., 1998), and bladder carcinoma (Williamson et al., 1995). Its deletion or mutation leads to cellular immortalization (Hara et al., 1996; Alcorta et al., 1996), and can predispose to melanoma. In melanomas, the mutant form of B-Raf strongly induces expression of *CDKN2A*, and as a result, melanomas with increased B-Raf mutation usually have deletions at the *INK4A* locus (Pollock et al., 2003). In human cancers, deletions leading to $p16^{INK4a}$ loss most always cause loss of $p15^{INK4b}$ also and to date it is not clear how each may contribute independently to tumor formation or progression.

Since genetic loss of $p16^{INK4a}$ is tumorigenic in mice, and mutation or deletion of $p15^{INK4b}$ and $p16^{INK4a}$ occurs frequently in human cancers, several studies have investigated if loss of $p15^{INK4b}/p16^{INK4a}$, either genetically or at the protein level, is a prognostic indicator of poor overall survival or shortened disease-free survival. In acute lymphoblastic leukemia, homozygous deletions in either $p15^{INK4b}/p16^{INK4a}$, or $p16^{INK4a}$ alone, were identified as an independent predictor of relapse and death, but a second study failed to confirm this

finding (Tsihlias et al., 1999). In non-Hodgkin's lymphoma, NSCLC, and melanoma, loss of either the $p16^{INK4a}$ gene or protein has an independent prognostic significance for poor overall survival and disease-free survival (reviewed in Tsihlias et al., 1999).

CIP/KIP Family

Loss of p21, and particularly p27, have been frequently reported to be of prognostic importance in many human cancers. The genes encoding these CDK inhibitors are rarely deleted or mutated in cancer (reviewed in Chu et al., 2008; Vidal and Koff, 2000). Notably, novel mutations in *CDKN1B* have been recently brought to light in whole genome sequencing of breast cancers, but their functional significance remains to be characterized (Stephens et al., 2012). While loss of these CDK inhibitors through deregulated proteolysis or translational inhibition (Chu et al., 2008) would promote tumor cell growth, both p21 and p27 appear to acquire pro-oncogenic roles when misregulated in cancers (Warfel and El-Deiry, 2013; Chu et al., 2008; Abbas and Dutta, 2009). The mislocalization of p27 to the cytoplasm, seen in up to 60% of cancers, leads to a gain of function in which cells acquire increased motility, invasion, and metastasis capability (Larrea et al., 2008; Besson et al., 2004; Wander et al., 2013). Mislocalization of p21 to the cytoplasm has been reported in several tumor types, including breast (Xia et al., 2004; Winters et al., 2003; reviewed in Abbas and Dutta, 2009), hepatocellular carcinoma (reviewed in Abbas and Dutta, 2009), testicular cancer (Koster et al., 2010), and acute leukemia (Schepers et al., 2003). Both p21 and p27 can bind to and inhibit RhoA activity in the cytoplasm. The same phosphorylation events that promote mislocalization of p21 and p27 enhance their binding to RhoA, which impairs RhoA-ROCK activity leading

to loss of acto-myosin stability and increased cell motility (Besson *et al.*, 2004; Lee and Helfman, 2004; Larrea *et al.*, 2009; see Figure 3).

The RhoA binding effect of cytoplasmic p27 is not the only mechanism whereby p27 drives cell invasiveness. AKT is a downstream effector of PI3K, which is commonly activated in many cancers (Wander *et al.*, 2011). Not only does oncogenic PI3K/mTOR activation drive p27 phosphorylation at T157 and T198 to increase cell motility (Larrea *et al.*, 2009), cytoplasmic p27 is a major mediator of metastasis downstream of PI3K/mTOR signaling. p27 knockdown had the same effect as PI3K/mTOR inhibition to dramatically reduce bone metastasis (Wander *et al.*, 2013). We recently reported a novel mechanism whereby C-terminally phosphorylated p27 contributes to cancer progression, linking PI3K activation to STAT3 activation, EMT and cancer metastasis. We describe an oncogenic feed-forward loop in which p27^{T157}p27^{T198} binds JAK2 promoting STAT3 recruitment and activation. STAT3-mediated *TWIST1* induction drives a p27 dependent EMT and feeds forward to further activate AKT contributing thereby to acquisition and maintenance of metastatic potential (Zhao *et al.*, 2015; see Figure 3).

The loss of nuclear p27 protein predicts poor prognosis in many cancers (Chu *et al.*, 2008). Indeed, virtually all human cancers show either a reduced nuclear p27 level or mislocalization of p27 to the cytoplasm (reviewed in Larrea *et al.*, 2009). p27 immunohistochemical staining showed low nuclear p27 staining is prognostic of poor overall survival or reduced disease-free survival in breast, prostate, colon, gastric, non-small cell lung cancer, esophageal cancer, glioma, sarcoma, and hematopoietic malignancies. Further work has shown that cytoplasmic p27 mislocalization may also have prognostic importance (Wander *et al.*, 2011). A recent analysis of p27 in node positive breast cancers showed increased cytoplasmic p27 correlated with increased node positivity and decreased survival (Wander *et al.*, 2013).

Although p21 loss has been reported to enhance metastatic spread (Warfel and El-Deiry, 2013), the role of p21 as a prognostic marker for disease progression is unclear. Several studies in breast cancer, melanoma, gastric cancer, endometrial cancer, and head and neck cancer have reported conflicting outcomes based on p21 staining (reviewed in Tsihlias *et al.*, 1999). However, cytosolic p21 staining was recently correlated with both increased Her2 (c-erb2/neu) expression and poor patient outcome (Winters *et al.*, 2003). Since the prognostic value of p21 expression in cancer may be tumor type specific, further evaluation is needed before it can be applied in the clinic.

CDK Inhibitors as Predictors of Drug Response

Since p21 and p27 restrain the cell cycle in normal cells, one would not want to target their normal function therapeutically for cancer. However, there may be circumstances, particularly where these CDK inhibitors gain oncogenic function through aberrant C-terminal phosphorylations, in which targeting these proteins or their upstream activators might be therapeutically useful. Since expression of p21 delays or prevents apoptosis, targeting p21 in DNA-damaged cells may prove

synthetically lethal with chemotherapeutic agents in cancer cells with deregulated p53 or activated anti-apoptotic mechanisms. Antisense oligodeoxynucleotide (ODN)-mediated p21 knockdown has been tested *in vitro* and subcutaneous injection of p21 ODN caused attenuation of Met-1 breast cancer growth *in vivo* (Fan *et al.*, 2003). Further work is required to permit therapeutic application of p21 knockdown by ODN or gene therapy in the clinic. p21 expression levels and localization may also help to determine which drugs to give for individual tumor types. Antisense p21 ODN or gene therapy may prove beneficial, particularly in combination with AKT inhibition in tumors with cytoplasmic p21 on immunohistochemical (IHC) staining. Cancers with disrupted p21 were found to be more sensitive to chemotherapeutic agents or DNA-damaging stimuli (Alpan and Pardee, 1996).

Predictive factors for response to specific therapies are arguably more important than prognostic factors since they influence treatment decisions. Although there is abundant evidence that p27 loss or cytoplasmic mislocalization are of poor prognostic importance, p27 staining may have its greatest utility in directing the choice of targeted therapy for human malignancies (Martin-Caballero *et al.*, 2004). Several molecular targeting drugs impact p27 by inhibiting upstream signaling. The BCR-ABL kinase, commonly expressed in chronic myelogenous leukemia, reduces p27 levels through constitutive tyrosine phosphorylation and subsequent accelerated proteolysis. Treatment of chronic myeloid leukemia (CML) cells with imatinib restores p27 levels and inhibits proliferation. Furthermore, p27 levels are stabilized in Her2-overexpressing breast cancers and EGFR-overexpressing lung cancers following treatment with trastuzumab and gefitinib, respectively. In a recent study, treatment of breast cancer cells with the EGFR/ErbB-2 inhibitor lapatinib increased p27 and reduced tumor proliferation (D'Alessio *et al.*, 2010). MEK inhibition increased p27 to induce growth arrest in melanoma models, as well as human pancreatic cancer lines driven by activated Ras. Taken together, these findings suggest that effective drug-mediated interdiction of pathways driving Src activation and p27 proteolysis may be effectively monitored by testing for posttreatment upregulation of nuclear p27.

In hormone-sensitive tumors, drug resistance is usually an inevitable outcome, and p27 may be involved in mediating resistance to antiestrogen therapies. Both preclinical studies and analysis of p27 in breast cancers of patients in clinical trials suggest that high nuclear p27 may predict responsiveness to hormonal therapy (Porter *et al.*, 2006; Pohl *et al.*, 2003). In estrogen-responsive tumors, estrogen binding to the ER promotes p27 proteolysis and Cyclin E-CDK2 activation, driving cell cycle progression. In sensitive ER-positive breast cancers, ER-blocking drugs (tamoxifen and fulvestrant) and aromatase inhibitors oppose estrogen signaling and stabilize p27 to inhibit proliferation and growth. Interestingly, high Src activation correlates with reduced p27 in human breast cancers, and Src inhibition increased p27 and restored tamoxifen sensitivity in resistant breast cancer lines. Dual targeting of estrogen receptor, with antiestrogens, and oncogenic Src signaling increases p27 levels and caused greater G1 arrest and xenograft tumor growth inhibition *in vivo*, and delayed emergence of drug resistance compared with that observed with either drug alone in both breast (Chen *et al.*, 2009) and ovarian cancer models (Simpkins

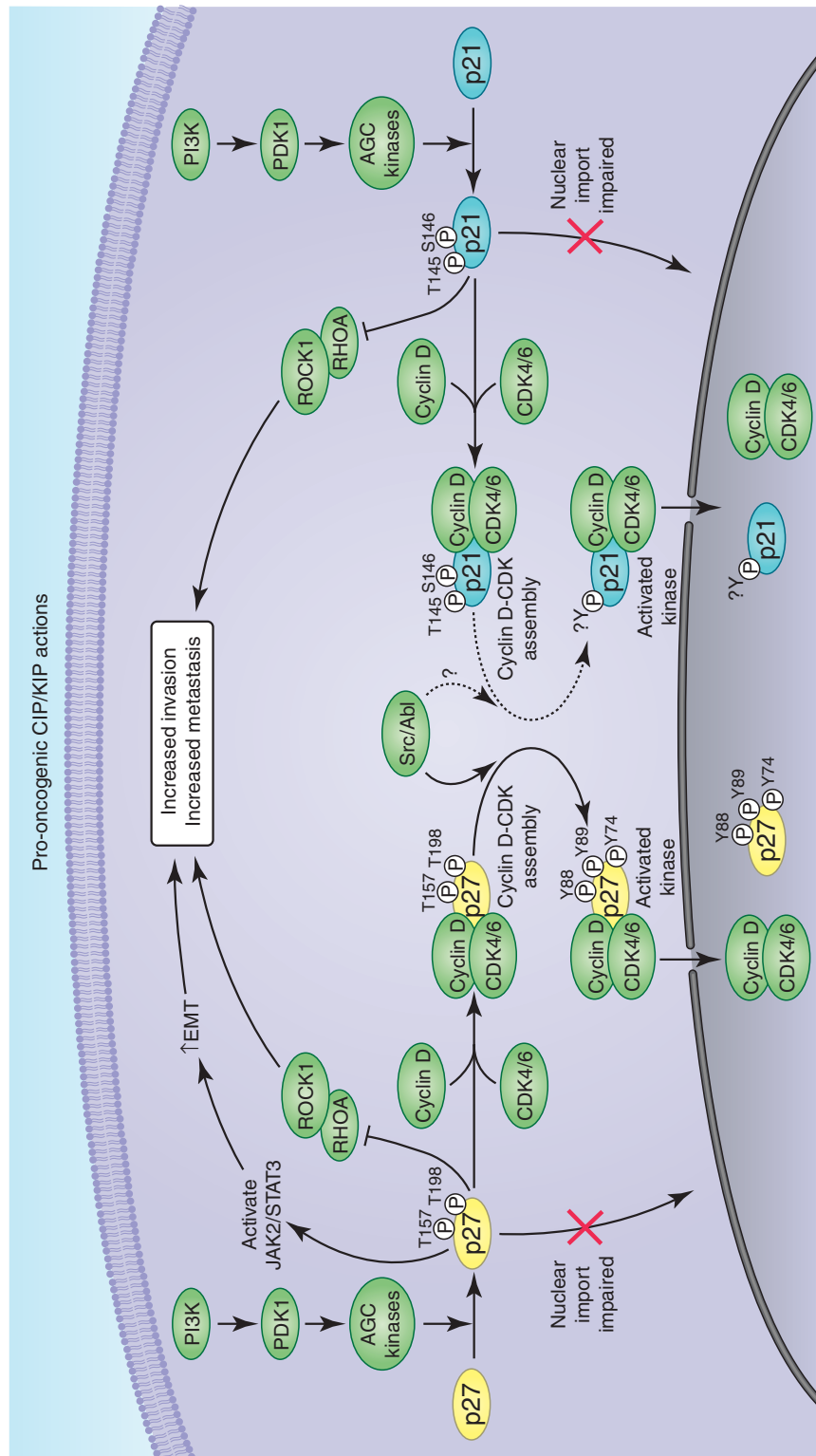


Figure 3 Both p21 and p27 are phosphorylated by oncogenic PI3K pathway effector kinases. These phosphorylations lead to cytoplasmic mislocalization (p27pT157, p27pT198, and p21pT145), protein stabilization (p27pT198 and p21pS146), and to acquisition of oncogenic properties including RhoA-ROCK1 inactivation and increased cell motility (both p21 and p27). Cyclin D-CDK activation and cell cycle progression and, in the case of p27, induction of an EMT phenotype to promote tumor metastasis.

et al., 2012). These data provided a novel rationale for the use of Src inhibitors together with antiestrogens for ER+ breast and ovarian cancers. Currently, clinical trials are underway to test the potential efficacy of combined Src and aromatase inhibition in postmenopausal ER+ breast cancer (Pegram and Slingerland, unpublished data).

In up to 60% of human breast cancers, hyperactivation of PI3K/mTOR signaling drives C-terminal phosphorylation of p27 at T157 and T198, which promotes cytoplasmic p27 mislocalization, increased invasiveness, EMT activation, and greater metastatic potential in a variety of cancers (Larrea *et al.*, 2008; Wander *et al.*, 2013). While early clinical trials with rapamycin and its analogs were met with limited success, due to incomplete inhibition of mTORC1 signaling, pure mTOR catalytic site inhibitors or dual PI3K/mTOR kinase inhibitors have shown great potential in preclinical models and early clinical trials (reviewed in Martin-Caballero *et al.*, 2004). The presence of cytoplasmic p27 in human cancer tissue may herald PI3K pathway activation and potential for PI3K/mTOR inhibitors to restore normal p27 function.

Concluding Remarks

Mammalian CDK inhibitors were first described a little over two decades ago and have illuminated fundamentals of cell cycle regulation and their study has yielded critical insights into mechanisms leading to malignant cell growth. Both INK4 family members and CIP/KIP family members are frequently functionally misregulated in human cancers, either through genetic changes (*CDKN2A* and *CDKN2B*) or through changes in phosphorylation, stability, localization, and function in the case of p21 and p27. While many studies have investigated the potential prognostic value of these alterations in various human malignancies, changes in CDK inhibitors may prove to be useful predictors of oncogenic pathway activation and potential for targeted therapy responses.

See also: Cell Division/Death: Cell Cycle: Mitosis in Animal Cells; The *INK4a/ARF* Locus; The Restriction Point. Functional Cell Biology: An Overview

References

- Abbas, T., Dutta, A., 2009. p21 in cancer: Intricate networks and multiple activities. *Nature Reviews Cancer* 9 (6), 400–414.
- Alcorta, D.A., Xiong, Y., Phelps, D., *et al.*, 1996. Involvement of the cyclin-dependent kinase inhibitor p16 (INK4a) in replicative senescence of normal human fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America* 93 (24), 13742–13747.
- Alpan, R.S., Pardee, A.B., 1996. p21WAF1/CIP1/SD1 is elevated through a p53-independent pathway by mimosine. *Cell Growth and Differentiation* 7 (7), 893–901.
- Asada, M., Yamada, T., Ichijo, H., *et al.*, 1999. Apoptosis inhibitory activity of cytoplasmic p21(Cip1/WAF1) in monocytic differentiation. *EMBO Journal* 18 (5), 1223–1234.
- Bass, A.J., Watanabe, H., Mermel, C.H., *et al.*, 2009. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nature Genetics* 41 (11), 1238–1242.
- Belinsky, S.A., Nikula, K.J., Palmisano, W.A., *et al.*, 1998. Aberrant methylation of p16(INK4a) is an early event in lung cancer and a potential biomarker for early diagnosis. *Proceedings of the National Academy of Sciences of the United States of America* 95 (20), 11891–11896.
- Beroukhi, R., Mermel, C.H., Porter, D., *et al.*, 2010. The landscape of somatic copy-number alteration across human cancers. *Nature* 463 (7283), 899–905.
- Besson, A., Assoian, R.K., Roberts, J.M., 2004. Regulation of the cytoskeleton: An oncogenic function for CDK inhibitors? *Nature Reviews Cancer* 4 (12), 948–955.
- Besson, A., Dowdy, S.F., Roberts, J.M., 2008. CDK inhibitors: Cell cycle regulators and beyond. *Developmental Cell* 14 (2), 159–169.
- Besson, A., Gurian-West, M., Schmidt, A., Hall, A., Roberts, J.M., 2004. p27Kip1 modulates cell migration through the regulation of RhoA activation. *Genes and Development* 18 (8), 862–876.
- Besson, A., Hwang, H.C., Cicero, S., *et al.*, 2007. Discovery of an oncogenic activity in p27Kip1 that causes stem cell expansion and a multiple tumor phenotype. *Genes and Development* 21 (14), 1731–1746.
- Bignell, G.R., Greenman, C.D., Davies, H., *et al.*, 2010. Signatures of mutation and selection in the cancer genome. *Nature* 463 (7283), 893–898.
- Boehm, M., Yoshimoto, T., Crook, M.F., *et al.*, 2002. A growth factor-dependent nuclear kinase phosphorylates p27(Kip1) and regulates cell cycle progression. *EMBO Journal* 21 (13), 3390–3401.
- Caldas, C., Hahn, S.A., de Costa, L.T., *et al.*, 1994. Frequent somatic mutations and homozygous deletions of the p16 (MTS1) gene in pancreatic adenocarcinoma. *Nature Genetics* 8, 27–32.
- Chan, K.M.F., Zhang, J., Cheng, L., Shapiro, D.N., Winoto, A., 1995. Identification of human and mouse p19, a novel CDK4 and CDK6 inhibitor with homology to p16INK4. *Molecular and Cellular Biology* 15, 2682–2688.
- Chen, Y., Guggisberg, N., Jorda, M., *et al.*, 2009. Combined Src and aromatase inhibition impairs human breast cancer growth *in vivo* and bypass pathways are activated in AZD0530-resistant tumors. *Clinical Cancer Research* 15 (10), 3396–3405.
- Chu, I., Sun, J., Arnaout, A., *et al.*, 2007. p27 phosphorylation by Src regulates inhibition of cyclin E-Cdk2. *Cell* 128 (2), 281–294.
- Chu, I.M., Hengst, L., Slingerland, J.M., 2008. The Cdk inhibitor p27 in human cancer: Prognostic potential and relevance to anticancer therapy. *Nature Reviews Cancer* 8 (4), 253–267.
- Ciarallo, S., Subramaniam, V., Hung, W., *et al.*, 2002. Altered p27(Kip1) phosphorylation, localization, and function in human epithelial cells resistant to transforming growth factor beta-mediated G(1) arrest. *Molecular and Cellular Biology* 22 (9), 2993–3002.
- Cmielova, J., Rezavova, M., 2011. p21Cip1/Waf1 protein and its function based on a subcellular localization (corrected). *Journal of Cellular Biochemistry* 112 (12), 3502–3506.
- Connor, M.K., Kotchetkov, R., Cariou, S., *et al.*, 2003. CRM1/Ran-mediated nuclear export of p27(Kip1) involves a nuclear export signal and links p27 export and proteolysis. *Molecular Biology of the Cell* 14 (1), 201–213.
- D'Alessio, A., De, L.A., Maiello, M.R., *et al.*, 2010. Effects of the combined blockade of EGFR and ErbB-2 on signal transduction and regulation of cell cycle regulatory proteins in breast cancer cells. *Breast Cancer Research and Treatment* 123 (2), 387–396.
- Deng, X., Mercer, S.E., Shah, S., Ewton, D.Z., Friedman, E., 2004. The cyclin-dependent kinase inhibitor p27Kip1 is stabilized in G(0) by Mirk/dyrk1B kinase. *Journal of Biological Chemistry* 279 (21), 22498–22504.
- Dulic, V., Kaufman, W.K., Wilson, S.J., *et al.*, 1994. p53-dependent inhibition of cyclin-dependent kinase activities in human fibroblasts during radiation-induced G1 arrest. *Cell* 76, 1013–1023.
- Dupont, J., Karas, M., LeRoith, D., 2003. The cyclin-dependent kinase inhibitor p21CIP/WAF is a positive regulator of insulin-like growth factor I-induced cell proliferation in MCF-7 human breast cancer cells. *Journal of Biological Chemistry* 278 (39), 37256–37264.
- Dyson, N., 1998. The regulation of E2F by pRB-family proteins. *Genes and Development* 12 (15), 2245–2262.
- El-Deiry, W.S., Tokino, T., Velculescu, V.E., *et al.*, 1993. WAF1, a potential mediator of p53 tumor suppression. *Cell* 75, 817–825.
- Fan, Y., Borowsky, A.D., Weiss, R.H., 2003. An antisense oligodeoxynucleotide to p21(Waf1/Cip1) causes apoptosis in human breast cancer cells. *Molecular Cancer Therapeutics* 2 (8), 773–782.
- Garcia-Fernandez, R.A., Garcia-Palencia, P., Sanchez, M.A., *et al.*, 2011. Combined loss of p21(waf1/cip1) and p27(kip1) enhances tumorigenesis in mice. *Laboratory Investigation* 91 (11), 1634–1642.
- Gil, J., Peters, G., 2006. Regulation of the INK4b-ARF-INK4a tumour suppressor locus: All for one or one for all. *Nature Reviews Molecular Cell Biology* 7 (9), 667–677.

- Grimmler, M., Wang, Y., Mund, T., *et al.*, 2007. Cdk-inhibitory activity and stability of p27Kip1 are directly regulated by oncogenic tyrosine kinases. *Cell* 128 (2), 269–280.
- Guan, K.L., Jenkins, C.W., Li, Y., *et al.*, 1994. Growth suppression by p18, a p16INK4/MTS1 and p14INK4B/MTS2-related CDK6 inhibitor, correlates with wild-type pRb function. *Genes and Development* 8 (24), 2939–2952.
- Hannon, G.J., Beach, D., 1994. p15INK4B is a potential effector of TGF- β -induced cell cycle arrest. *Nature* 371 (6494), 257–261.
- Hara, E., Smith, R., Parry, D., *et al.*, 1996. Regulation of p16CDKN2 expression and its implications for cell immortalization and senescence. *Molecular and Cellular Biology* 16, 859–867.
- Harper, J.W., Adami, G.R., Wei, N., Keyomarsi, K., Elledge, S.J., 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependant kinases. *Cell* 75, 805–816.
- Hengst, L., Dulic, V., Slingerland, J.M., Lees, E., Reed, S.I., 1994. A cell cycle-regulated inhibitor of cyclin-dependent kinases. *Proceedings of the National Academy of Sciences of the United States of America* 91 (12), 5291–5295.
- Hirai, H., Roussel, M.F., Kato, J.Y., Ashmun, R.A., Sherr, C.J., 1995. Novel INK4 proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6. *Molecular and Cellular Biology* 15, 2672–2681.
- Hwang, C.Y., Lee, C., Kwon, K.S., 2009. Extracellular signal-regulated kinase 2-dependent phosphorylation induces cytoplasmic localization and degradation of p21Cip1. *Molecular and Cellular Biology* 29 (12), 3379–3389.
- Ishida, N., Hara, T., Kamura, T., *et al.*, 2002. Phosphorylation of p27Kip1 on serine 10 is required for its binding to CRM1 and nuclear export. *Journal of Biological Chemistry* 277 (17), 14355–14358.
- Ishida, N., Kitagawa, M., Hatakeyama, S., Nakayama, K., 2000. Phosphorylation at Serine 10, a major phosphorylation site of p27Kip1, increases its protein stability. *Journal of Biological Chemistry* 275 (33), 25146–25154.
- James, M., Ray, A., Leznova, D., Blain, S.W., 2007. Differential modification of p27Kip1 controls its cyclin D-cdk4 inhibitory activity. *Molecular and Cellular Biology* 28 (1), 498–510.
- Jin, Y.H., Yoo, K.J., Lee, Y.H., Lee, S.K., 2000. Caspase 3-mediated cleavage of p21WAF1/CIP1 associated with the cyclin A-cyclin-dependent kinase 2 complex is a prerequisite for apoptosis in SK-HEP-1 cells. *Journal of Biological Chemistry* 275 (39), 30256–30263.
- Kamb, A., Shattuch-Eidens, D., Eeles, R., *et al.*, 1994. Analysis of the p16 gene (CDKN2) as a candidate for the chromosome 9p melanoma susceptibility locus. *Nature Genetics* 8, 23–26.
- Kim, W.Y., Sharpless, N.E., 2006. The regulation of INK4/ARF in cancer and aging. *Cell* 127 (2), 265–275.
- Koff, A., Ohtsuki, M., Polyak, K., Roberts, J.M., Massague, J., 1993. Negative regulation of G1 in mammalian cells: Inhibition of cyclin E-dependent kinase by TGF- β . *Science* 260 (5107), 536–539.
- Koster, R., Di, P.A., Timmer-Bosscha, H., *et al.*, 2010. Cytoplasmic p21 expression levels determine cisplatin resistance in human testicular cancer. *Journal of Clinical Investigation* 120 (10), 3594–3605.
- LaBaer, J., Garrett, M.D., Stevenson, L.F., *et al.*, 1997. New functional activities for the p21 family of CDK inhibitors. *Genes and Development* 11 (7), 847–862.
- Larrea, M.D., Hong, F., Wander, S.A., *et al.*, 2009. RSK1 drives p27Kip1 phosphorylation at T198 to promote RhoA inhibition and increase cell motility. *Proceedings of the National Academy of Sciences of the United States of America* 106 (23), 9268–9273.
- Larrea, M.D., Liang, J., Da, S.T., *et al.*, 2008. Phosphorylation of p27Kip1 regulates assembly and activation of cyclin D1-Cdk4. *Molecular and Cellular Biology* 28 (20), 6462–6472.
- Larrea, M.D., Wander, S.A., Slingerland, J.M., 2009. p27 as Jekyll and Hyde: Regulation of cell cycle and cell motility. *Cell Cycle* 8 (21), 3455–3461.
- Lee, S., Helfman, D.M., 2004. Cytoplasmic p21Cip1 is involved in Ras-induced inhibition of the ROCK/LIMK/cofilin pathway. *Journal of Biological Chemistry* 279 (3), 1885–1891.
- Leis, O., Eguirra, A., Lopez-Arribilla, E., *et al.*, 2011. Sox2 expression in breast tumours and activation in breast cancer stem cells. *Oncogene* 31 (11), 1354–1365.
- Li, H., Collado, M., Villasante, A., *et al.*, 2012. p27(Kip1) directly represses Sox2 during embryonic stem cell differentiation. *Cell Stem Cell* 11 (6), 845–852.
- Liu, M., Lee, M.H., Cohen, M., Bommakanti, M., Freedman, L.P., 1996. Transcriptional activation of the Cdk inhibitor p21 by vitamin D3 leads to the induced differentiation of the myelomonocytic cell line U937. *Genes and Development* 10 (2), 142–153.
- Marson, A., Levine, S.S., Cole, M.F., *et al.*, 2008. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 134 (3), 521–533.
- Martin-Caballero, J., Flores, J.M., Garcia-Palencia, P., Collado, M., Serrano, M., 2004. Different cooperating effect of p21 or p27 deficiency in combination with INK4a/ARF deletion in mice. *Oncogene* 23 (50), 8231–8237.
- Messina, G., Blasi, C., La Rocca, S.A., *et al.*, 2005. p27Kip1 acts downstream of N-cadherin-mediated cell adhesion to promote myogenesis beyond cell cycle regulation. *Molecular Biology of the Cell* 16 (3), 1469–1480.
- Moulton, T., Samara, G., Chung, W.Y., *et al.*, 1995. MTS1/p16/CDKN2 lesions in primary glioblastoma multiforme. *American Journal of Pathology* 146 (3), 613–619.
- Movassagh, M., Philpott, A., 2008. Cardiac differentiation in *Xenopus* requires the cyclin-dependent kinase inhibitor, p27Xic1. *Cardiovascular Research* 79 (3), 436–447.
- Muthna, D., Soukup, T., Vavrova, J., *et al.*, 2010. Irradiation of adult human dental pulp stem cells provokes activation of p53, cell cycle arrest, and senescence but not apoptosis. *Stem Cells and Development* 19 (12), 1855–1862.
- Nakayama, K.I., Nakayama, K., 2006. Ubiquitin ligases: Cell-cycle control and cancer. *Nature Reviews Cancer* 6 (5), 369–381.
- Nguyen, L., Besson, A., Heng, J.I., *et al.*, 2006. p27kip1 independently promotes neuronal differentiation and migration in the cerebral cortex. *Genes and Development* 20 (11), 1511–1524.
- Nguyen, L., Besson, A., Heng, J.I., *et al.*, 2007. p27Kip1 independently promotes neuronal differentiation and migration in the cerebral cortex. *Bulletin et memoires de l'Academie royale de medecine de Belgique* 162 (5–6), 310–314.
- Nobori, T., Miura, K., Wu, D.J., *et al.*, 1994. Deletions of the cyclin-dependent kinase-4 inhibitor in multiple human cancers. *Nature* 368, 753–756.
- Noda, A., Ning, Y., Venable, S.F., Pereira-Smith, O.M., Smith, J.R., 1994. Cloning of senescent cell-derived inhibitors of DNA synthesis using an expression screen. *Experimental Cell Research* 211, 90–98.
- Ohtani, N., Zebedee, Z., Huot, T.J., *et al.*, 2001. Opposing effects of Ets and Id proteins on p16INK4a expression during cellular senescence. *Nature* 409 (6823), 1067–1070.
- Okuda, T., Shurtleff, S.A., Valentine, M.B., *et al.*, 1995. Frequent deletion of p16INK4a/MTS1 and p15^{INK4b}/MTS2 in pediatric acute lymphoblastic leukemia. *Blood* 85 (9), 2321–2330.
- Pippa, R., Espinosa, L., Gundem, G., *et al.*, 2012. p27Kip1 represses transcription by direct interaction with p130/E2F4 at the promoters of target genes. *Oncogene* 31 (38), 4207–4220.
- Pohl, G., Rudas, M., Dietze, O., *et al.*, 2003. High p27Kip1 expression predicts superior relapse-free and overall survival for premenopausal women with early-stage breast cancer receiving adjuvant treatment with tamoxifen plus goserelin. *Journal of Clinical Oncology* 21 (19), 3594–3600.
- Pollock, P.M., Harper, U.L., Hansen, K.S., *et al.*, 2003. High frequency of BRAF mutations in nevi. *Nature Genetics* 33 (1), 19–20.
- Polyak, K., Kato, J.Y., Solomon, M.J., *et al.*, 1994a. p27Kip1, a cyclin-Cdk inhibitor, links transforming growth factor- β and contact inhibition to cell cycle arrest. *Genes and Development* 8 (1), 9–22.
- Polyak, K., Lee, M.H., Erdjument-romage, H., *et al.*, 1994b. Cloning of p27KIP1, a cyclin-dependent kinase inhibitor and a potential mediator of extracellular antimitogenic signals. *Cell* 78, 59–66.
- Porter, P.L., Barlow, W.E., Yeh, I.T., *et al.*, 2006. p27(Kip1) and cyclin E expression and breast cancer survival after treatment with adjuvant chemotherapy. *Journal of the National Cancer Institute* 98 (23), 1723–1731.
- Quesnel, B., Fenaux, P., 1999. p15INK4b gene methylation and myelodysplastic syndromes. *Leukemia & Lymphoma* 35 (5–6), 437–443.
- Quesnel, B., Guillemin, G., Vereecque, R., *et al.*, 1998. Methylation of the p15 (INK4b) gene in myelodysplastic syndromes is frequent and acquired during disease progression. *Blood* 91 (8), 2985–2990.
- Rodier, G., Montagnoli, A., Di Marcotullio, L., *et al.*, 2001. p27 cytoplasmic localization is regulated by phosphorylation on Ser10 and is not a prerequisite for its proteolysis. *EMBO Journal* 20 (23), 6672–6682.
- Rodriguez-Villarrupia, A., Diaz, C., Canela, N., *et al.*, 2002. Identification of the nuclear localization signal of p21(cip1) and consequences of its mutation on cell proliferation. *FEBS Letters* 531 (2), 319–323.
- Roninson, I.B., 2002. Oncogenic functions of tumour suppressor p21(Waf1/Cip1/Sdi1): Association with cell senescence and tumour-promoting activities of stromal fibroblasts. *Cancer Letters* 179 (1), 1–14.
- Roussel, M.F., 1999. The INK4 family of cell cycle inhibitors in cancer. *Oncogene* 18 (38), 5311–5317.
- Rozenblum, E., Schutte, M., Goggins, M., *et al.*, 1997. Tumor-suppressive pathways in pancreatic carcinoma. *Cancer Research* 57 (9), 1731–1734.
- Russo, A.A., Jeffrey, P.D., Patten, A.K., Massague, J., Pavletich, N.P., 1996. Crystal structure of the p27Kip1 cyclin-dependent-kinase inhibitor bound to the cyclin A-Cdk2 complex. *Nature* 382 (6589), 325–331.

- Sandhu, C., Donovan, J., Bhattacharya, N., *et al.*, 2000. Reduction of Cdc25A contributes to cyclin E1-Cdk2 inhibition at senescence in human mammary epithelial cells. *Oncogene* 19 (47), 5314–5323.
- Sandhu, C., Peehl, D.M., Slingerland, J., 2000. p16INK4A mediates cyclin dependent kinase 4 and 6 inhibition in senescent prostatic epithelial cells. *Cancer Research* 60 (10), 2616–2622.
- Schepers, H., Geugien, M., Eggen, B.J., Vellenga, E., 2003. Constitutive cytoplasmic localization of p21(Waf1/Cip1) affects the apoptotic process in monocytic leukaemia. *Leukemia* 17 (11), 2113–2121.
- Seoane, J., Le, H.V., Massague, J., 2002. Myc suppression of the p21(Cip1) Cdk inhibitor influences the outcome of the p53 response to DNA damage. *Nature* 419 (6908), 729–734.
- Serrano, M., Hannon, G.J., Beach, D., 1993. A new regulatory motif in cell cycle control causing specific inhibition of cyclin D/Cdk4. *Nature* 366, 704–707.
- Serrano, M., Lee, H., Chin, L., *et al.*, 1996. Role of the INK4a locus in tumor suppression and cell mortality. *Cell* 85 (1), 27–37.
- Sherr, C.J., 1993. Mammalian G1 cyclins. *Cell* 73, 1059–1065.
- Sherr, C.J., Roberts, J.M., 1999. CDK inhibitors: Positive and negative regulators of G1-phase progression. *Genes and Development* 13 (12), 1501–1512.
- Simpkins, F., Hevia-Paez, P., Sun, J., *et al.*, 2012. Src inhibition with saracatinib reverses fulvestrant resistance in ER-positive ovarian cancer models *in vitro* and *in vivo*. *Clinical Cancer Research* 18 (21), 5911–5923.
- Slingerland, J.M., Hengst, L., Pan, C.H., *et al.*, 1994. A novel inhibitor of cyclin-Cdk activity detected in transforming growth factor beta-arrested epithelial cells. *Molecular and Cellular Biology* 14 (6), 3683–3694.
- Stepanova, L., Sorrentino, B.P., 2005. A limited role for p16Ink4a and p19Arf in the loss of hematopoietic stem cells during proliferative stress. *Blood* 106 (3), 827–832.
- Stephens, P.J., Tarpey, P.S., Davies, H., *et al.*, 2012. The landscape of cancer genes and mutational processes in breast cancer. *Nature* 486 (7403), 400–404.
- Tanaka, H., Yamashita, T., Asada, M., *et al.*, 2002. Cytoplasmic p21(Cip1/WAF1) regulates neurite remodeling by inhibiting Rho-kinase activity. *Journal of Cell Biology* 158 (2), 321–329.
- Tsihlias, J., Kapusta, L., Slingerland, J., 1999. The prognostic significance of altered cyclin-dependent kinase inhibitors in human cancer. *Annual Review of Medicine* 50, 401–423.
- Vernon, A.E., Devine, C., Philpott, A., 2003. The cdk inhibitor p27Xic1 is required for differentiation of primary neurones in *Xenopus*. *Development* 130 (1), 85–92.
- Vernon, A.E., Movassagh, M., Horan, I., *et al.*, 2006. Notch targets the Cdk inhibitor Xic1 to regulate differentiation but not the cell cycle in neurons. *EMBO Reports* 7 (6), 643–648.
- Vernon, A.E., Philpott, A., 2003. A single cdk inhibitor, p27Xic1, functions beyond cell cycle regulation to promote muscle differentiation in *Xenopus*. *Development* 130 (1), 71–83.
- Vidal, A., Koff, A., 2000. Cell-cycle inhibitors: Three families united by a common cause. *Gene* 247 (1–2), 1–15.
- Vlach, J., Hennecke, S., Amati, B., 1997. Phosphorylation-dependent degradation of the cyclin-dependent kinase inhibitor p27. *EMBO Journal* 16 (17), 5334–5344.
- Wander, S., Zhao, D., Besser, A., *et al.*, 2013. PI3K/mTOR inhibition can impair tumor invasion and metastasis *in vivo* despite a lack of antiproliferative action *in vitro*. Implications for targeted therapy. *Breast Cancer Research and Treatment* 138 (2), 369–381.
- Wander, S.A., Hennessy, B.T., Slingerland, J.M., 2011. Next generation mTOR inhibitors in clinical oncology: How pathway complexity informs therapeutic strategy. *Journal of Clinical Investigation* 121 (4), 1231–1241.
- Wander, S.A., Zhao, D., Slingerland, J.M., 2011. p27: A barometer of signaling deregulation and potential predictor of response to targeted therapies. *Clinical Cancer Research* 17 (1), 12–18.
- Warfel, N.A., El-Deiry, W.S., 2013. p21WAF1 and tumorigenesis: 20 years after. *Current Opinion in Oncology* 25 (1), 52–58.
- Weiss, R.H., 2003. p21Waf1/Cip1 as a therapeutic target in breast and other cancers. *Cancer Cell* 4 (6), 425–429.
- Wilentz, R.E., Geradts, J., Maynard, R., *et al.*, 1998. Inactivation of the p16 (INK4A) tumor-suppressor gene in pancreatic duct lesions: Loss of intranuclear expression. *Cancer Research* 58 (20), 4740–4744.
- Williamson, M.P., Elder, P.A., Shaw, M.E., Devlin, J., Knowles, M.A., 1995. p16 (CDKN2) is a major deletion target at 9p21 in bladder cancer. *Human Molecular Genetics* 4 (9), 1569–1577.
- Winters, Z.E., Leek, R.D., Bradburn, M.J., Norbury, C.J., Harris, A.L., 2003. Cytoplasmic p21WAF1/CIP1 expression is correlated with HER-2/ neu in breast cancer and is an independent predictor of prognosis. *Breast Cancer Research* 5 (6), R242–R249.
- Xia, W., Chen, J.S., Zhou, X., *et al.*, 2004. Phosphorylation/cytoplasmic localization of p21Cip1/WAF1 is associated with HER2/neu overexpression and provides a novel combination predictor for poor prognosis in breast cancer patients. *Clinical Cancer Research* 10 (11), 3815–3824.
- Yeh, N., Miller, J.P., Gaur, T., *et al.*, 2007. Cooperation between p27 and p107 during endochondral ossification suggests a genetic pathway controlled by p27 and p130. *Molecular and Cellular Biology* 27 (14), 5161–5171.
- Zeng, Y., Hirano, K., Hirano, M., Nishimura, J., Kanaide, H., 2000. Minimal requirements for the nuclear localization of p27Kip1, a cyclin-dependent kinase inhibitor. *Biochemical and Biophysical Research Communications* 274 (1), 37–42.
- Zhang, C., Kavurma, M.M., Lai, A., Khachigian, L.M., 2003. Ets-1 protects vascular smooth muscle cells from undergoing apoptosis by activating p21WAF1/Cip1: ETS-1 regulates basal and inducible p21WAF1/Cip1 transcription via distinct cis-acting elements in the p21WAF1/Cip1 promoter. *Journal of Biological Chemistry* 278 (30), 27903–27909.
- Zhao, D., Besser, A., Wander, S., *et al.*, 2015. Cytoplasmic p27 promotes epithelial-mesenchymal transition and tumor metastasis via STAT3-mediated twist1 upregulation. *Oncogene*. doi:10.1038/onc.2014.473.

The *INK4a/ARF* Locus

DE Quelle, University of Iowa, Iowa City, IA, USA and University of Iowa Hospitals and Clinics, Iowa City, IA, USA

J Nteeba, University of Iowa, Iowa City, IA, USA

BW Darbro, University of Iowa, Iowa City, IA, USA and University of Iowa Hospitals and Clinics, Iowa City, IA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The cell cycle is a series of ordered biochemical events resulting in division of a mother cell into two identical daughter cells. This process requires the accurate duplication of DNA and equal segregation of chromosomes into each daughter cell, and it must be tightly regulated throughout the life span of complex organisms in order to prevent the unrestrained proliferation of damaged cells (i.e., cancer). Proper cell cycle progression is monitored by checkpoints, surveillance mechanisms that ensure the completion of each cell cycle phase and maintenance of genome integrity (Malumbres and Barbacid, 2009; Sherr, 2004). They are enforced by tumor suppressor proteins that arrest the cell cycle, enabling time for the cell to repair mistakes before reentering the cycle, or eliminate cells (e.g., by programmed cell death (apoptosis)) if the damage is intolerable. Two of the most important tumor suppressors are the retinoblastoma (RB1) and p53 proteins, which are activated by the products of the *INK4a/ARF* locus, p16INK4a and the alternative reading frame protein (ARF), respectively (Levine and Oren, 2009; Sherr, 2001). This review takes a historical look at the discovery and study of *INK4a/ARF*. It begins with a consideration of fundamental mechanisms controlling the cell cycle, highlighting where and how p16INK4a and ARF exert their anti-proliferative effects, followed by an analysis of their individual and interconnected roles in cancer. Growing evidence that the *INK4a/ARF* locus plays a pivotal role in other important aspects of normal biology will also be discussed.

Control and Significance of G1–S Phase Progression in Mammals

The 1990s was a time of rapid advancement in our understanding of the mammalian cell cycle, particularly the events controlling the first Gap period (G1 phase) and mechanisms by which cells begin DNA synthesis (S phase). Building of earlier studies of cyclins and cyclin-dependent kinases (Cdks) in yeast and *Xenopus*, we learned that the mammalian D-type cyclins (D1, D2, and D3) are regulatory subunits whose expression is maximally induced by mitogens in mid-G1 phase, at which time they form active serine/threonine kinases with their catalytic partners, Cdk4 or Cdk6 (Hunter and Pines, 1994; Malumbres and Barbacid, 2009; Sherr and Roberts, 2004). Those holoenzymes promote G1 to S phase progression by phosphorylating and inactivating the RB1 tumor suppressor protein (Figure 1(a)). RB1 phosphorylation, first by cyclin D-Cdk4/6 followed by cyclin E-Cdk2 and cyclin A-Cdk2 kinases, results in the release of E2F transcription factors from inhibitory RB1 complexes (Harbour and Dean, 2000). Once free, activated E2F proteins stimulate the expression of

numerous genes required for DNA replication, permitting cells to enter and progress through S phase. Interestingly, RB1 has two relatives, p107 and p130, that act similarly although their tissue distribution varies and only RB1 is commonly mutated in human cancers (Burkhardt and Sage, 2008).

G1 phase cyclins (principally D and E) and their partners (Cdks 2, 4, and 6) are rate-limiting in promoting cell cycle progression. Their overexpression in cultured cells shortens G1 phase and accelerates S phase entry, whereas their loss, depending on the cell type, prevents cell cycle progression (Sherr and Roberts, 2004). Moreover, many human cancers have gene amplifications, mutations, or translocations that elevate the expression and activity of these cyclins and Cdks (Malumbres and Barbacid, 2009). Mimicking those effects in mice causes tumor development *in vivo*. For example, artificial overexpression of cyclin D1 in mammary epithelial cells using the mouse mammary tumor virus promoter (MMTV) causes breast cancer in MMTV-D1 transgenic mice (Wang *et al.*, 1994). Conversely, mice lacking cyclin D1 are protected from breast cancer in animals with activated *Ras* or *Her2/neu/ErbB2* oncogenes (Lee and Sicinski, 2006). Such studies have established that elevated cyclin D-Cdk4/6 and E-Cdk2 activities are key drivers of tumorigenesis.

As a result, investigators have sought to understand the mechanisms that normally restrict cyclin D-Cdk4/6 and E-Cdk2 kinase activities in non-transformed cells and define how

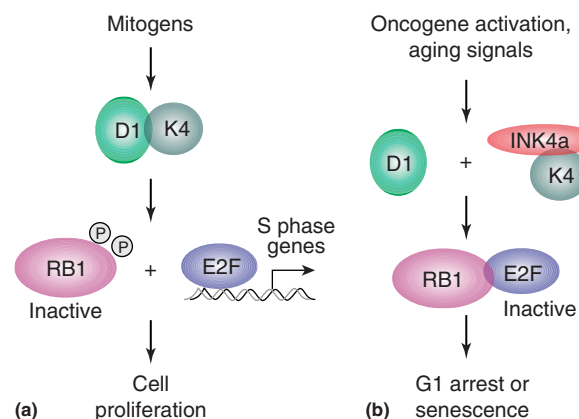


Figure 1 The *INK4a*–RB1 anti-proliferative pathway. (a) The upregulation of cyclin D1 by mitogens promotes formation of D1–Cdk4 kinases. Phosphorylation and inactivation of RB1 enables E2F to bind DNA and transcribe genes required for S phase entry and cell proliferation. (b) p16INK4a expression is induced by hyper-proliferative signals from activated oncogenes or events during cellular aging that lead to *INK4a* derepression. Cdk4 binding by p16INK4a disrupts D1–Cdk4 association and leads to accumulation of hypophosphorylated RB1 and inactive E2F. Cells become reversibly arrested in G1 phase or undergo a permanent arrest termed senescence.

those events are compromised in human tumors. Summarizing over 20 years of research, studies show that the changing phosphorylation status of each cyclin and Cdk subunit plays an essential role in controlling their subcellular localization, stability, conformation, and binding to each other. Disrupting any of those controls can promote tumor formation by enhancing the activity of cyclin D- and E-dependent kinases. An additional level of negative control involves small proteins that bind and inhibit those kinases. These cyclin-dependent kinase inhibitors (CKIs) fall into two major classes, the Cip/Kip and INK4 proteins, and their in-depth characterization has been motivated by evidence that their inactivation occurs in a significant fraction of human cancers and promotes tumorigenesis (Malumbres and Barbacid, 2009).

The Cip/Kip family ('Cdk inhibitory' or 'Kinase inhibitory' proteins) includes p21Cip1 (*CDKN1A*), p27Kip1 (*CDKN1B*), and p57Kip2 (*CDKN1C*). Low levels of these proteins are actually critical for cyclin D-Cdk4/6 assembly and cell cycling, and their sequestration by D-Cdk4/6 complexes permits activation of the Cdk2 holoenzymes (Sherr and Roberts, 2004). However, their elevated expression strongly inhibits cyclin D-dependent kinases and halts cell proliferation, whereas even low concentrations of Cip/Kip proteins inhibit cyclin E-Cdk2 and A-Cdk2 as well as the M phase kinase, cyclin B-Cdk1. While Cip/Kip factors act broadly, the INK4 proteins selectively bind Cdk4 (or its homolog, Cdk6) and act as specific 'Inhibitors of Cdk4' (Figure 1(b)). The four members of the INK4 family (p16INK4a, p15INK4b, p18INK4c, and p19INK4d) were named based on their size (kDa) and order of discovery. These proteins are defined by the presence of four ankyrin repeats, which are required for their Cdk binding and inhibitory activity.

p16INK4a and RB1-Mediated Senescence

p16INK4a is the first identified, prototypical member of the INK4 family (Serrano *et al.*, 1993). p16INK4a and the other INK4 proteins each bind and inhibit Cdk4 (or Cdk6), and when expressed *in vivo* they primarily exist in binary INK4-Cdk complexes. Structural studies of p16INK4a and p18INK4c showed they function by blocking ATP binding and distorting the cyclin binding site in the Cdk, effectively preventing or disrupting the formation of cyclin D-Cdk4/6 holoenzymes (Jeffrey *et al.*, 2000; Russo *et al.*, 1998). In proliferating cells, the INK4 proteins are either absent (p16INK4a) or expressed at very low levels (e.g., p18INK4c). However, their expression is greatly increased in response to various cellular stresses or anti-proliferative signals, such as TGF- β , and their consequent inhibition of D-Cdk4/6 kinases potently arrests the cells in G1 phase. Importantly, the INK4 proteins have no growth inhibitory activity in cells that lack functional RB1, illustrating the strict linearity of the INK4-cyclin D-Cdk4/6-RB1 pathway (Sharpless and DePinho, 1999).

The primary differences between p16INK4a and its relatives occur in their tissue distribution, signals regulating their expression, and role in biology. For instance, p18INK4c and p19INK4d are normally induced by differentiation and expressed in non-proliferating cells within various tissues (e.g., spermatocytes in the testes and sensory hair cells in the

corti) (Zindy *et al.*, 1997; Krishnamurthy *et al.*, 2004). Their expression is vital for maintaining a quiescent state of differentiated cells, which has dire consequences when lost as indicated by the sterility of p18INK4c/p19INK4d-null male mice and deafness of p19INK4d-null animals (Zindy *et al.*, 2001; Chen *et al.*, 2003). By comparison, p16INK4a is undetectable in embryos and young mammals and becomes increasingly expressed in aging mice and humans (Zindy *et al.*, 1997; Krishnamurthy *et al.*, 2004; Ressler *et al.*, 2006). That expression pattern correlates with p16INK4a playing a role in cellular senescence, an irreversible cell cycle withdrawal associated with cell and organismal aging. Senescence is also important for tumor suppression, not only by promoting the permanent arrest of damaged cells but also by inducing their immune clearance (Xue *et al.*, 2007). p16INK4a is upregulated in cells undergoing senescence, induces telomere-independent senescence when expressed, and is required for senescence in cultured human cells (Darbro *et al.*, 2005; Collado *et al.*, 2007). Since loss of senescence is a key step in tumorigenesis, it is not surprising that *INK4a* gene inactivation is seen in nearly half of all human tumors (Ruas and Peters, 1998). While most of the INK4 and Cip/Kip proteins (except p19INK4d) have tumor suppressive activity, the frequency of *INK4a* inactivation in cancers far exceeds that of the other CKIs (Malumbres and Barbacid, 2009), suggesting a more prominent contribution of *INK4a* to tumor suppression.

INK4a/ARF – a Dual Coding Locus in Mammals

The role of *INK4a* in cancer became clouded by the startling realization that another transcript is encoded by the locus. Several groups made this discovery at the same time (Stone *et al.*, 1995; Mao *et al.*, 1995; Duro *et al.*, 1995; Quelle *et al.*, 1995), and while some thought the alternate transcript was noncoding it became clear that a protein was generated and it was a powerful cell cycle inhibitor in its own right (Quelle *et al.*, 1995). The locus is unusual because it encodes two unrelated proteins derived from separate first exons, exon 1 α for p16INK4a and exon 1 β for ARF, that are spliced to the same second exon (Figure 2). Each first exon supplies its own ATG start codon that initiates translation in alternative reading frames through shared sequences in exons 2 and 3, hence the name 'ARF' for the product of the β transcript. The resulting p16INK4a and ARF (mouse p19ARF, human p14ARF) proteins share no amino acid identity, and their structures and mechanisms of action are entirely distinct. While p16INK4a acts in the RB1 pathway, ARF activates the p53 tumor suppressor as well as other cancer pathways (Sherr, 2006; Kim and Sharpless, 2006; Gil and Peters, 2006), as described in following sections. This dual coding region was then renamed *INK4a/ARF* although the original locus designation, *CDKN2A* (which fails to convey the locus complexity), is still used prevalently in databases.

The *INK4a/ARF* locus is interesting from an evolutionary perspective. It is conserved in mammals, but chickens lack *INK4a* and only express a truncated ARF protein (p7) from an orthologous exon 1 β while neither *INK4a* nor *ARF* are found in fish (Gil and Peters, 2006). However, fish and chickens do retain an *INK4b* gene, which in the mammalian genome

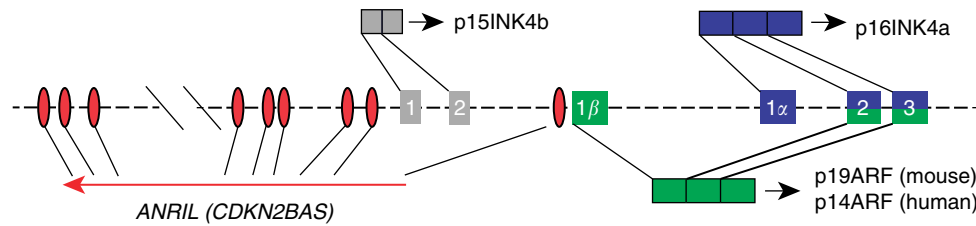


Figure 2 Schematic of the *INK4a/ARF* locus. The *ARF* and *INK4a* genes, depicted with green and blue exons, respectively, span ~20 kb on human chromosome 9p21.3 or mouse chromosome 4. The α transcript for p16INK4a (blue) and β transcript for ARF (green) are generated from distinct promoters upstream of their first exons. Mouse and human ARF proteins differ in size due to C-terminal variations and are either 19 kDa (mouse) or 14 kDa (human). The *INK4b* gene (gray exons) encoding p15INK4b resides centromeric to exon 1 β of *ARF*. Intronic sequences are represented by the dashed black horizontal line. Representative exons (thought to total 19 and spanning ~120 kb) of the long noncoding RNA, *ANRIL* (also called *CDKN2BAS*), are depicted with red ovals and the transcript shown as a red arrow. *ANRIL* is transcribed in an antisense direction from a promoter element near that for exon 1 β of *ARF*. This image is not drawn to scale, but to give some perspective exon 1 β resides ~8 kb from p15INK4b exon 2 and ~13 kb from exon 1 α . Notably, this locus is a genomic hotspot for single nucleotide polymorphisms (SNPs) associated with several cancers and a variety of other age-related diseases in humans.

resides proximal to *ARF* exon 1 β (see Figure 2). Based on sequence analyses across species, it is thought that *INK4a* arose fairly recently in evolution by duplication of *INK4b* followed by subsequent acquisition of exon 1 β . The more recent appearance of ARF is consistent with the notion that dual coding allows for the expression of novel, intrinsically disordered proteins that perform new functions (Kovacs *et al.*, 2010). Indeed, p16INK4a has a well-defined α -helical structure and binds Cdk4/6 (Byeon *et al.*, 1998), while ARF is largely unstructured (especially in its C-terminal half derived from the alternative reading frame of exon 2) and participates in other pathways (Kovacs *et al.*, 2010; Sherr, 2006).

Dual coding of unrelated proteins from overlapping reading frames is typical in the compact genomes of prokaryotes and viruses, but it is rare in eukaryotes. For years, *INK4a/ARF* was thought to be the only genuine dual coding gene in mammals but other loci of biological importance have been identified. For instance, the XBP1^S transcription factor and its inhibitor, XBP1^U, arise from alternative reading frames in the *XBP1* (X-box binding protein 1) gene and their coordinated activity is required for such fundamental processes as plasma cell differentiation and the unfolded protein response (Yoshida *et al.*, 2001; Iwakoshi *et al.*, 2003). Likewise, *GNAS1* (G protein alpha subunit) encodes the unrelated G proteins, XL α s and ALEX, whose interaction is critical for multiple signaling pathways and whose loss of function is linked with several human disorders (Freson *et al.*, 2003; Turan and Bastepe, 2013; Klemke *et al.*, 2001). An intriguing theme is that the dual products perform related tasks, either by binding and regulating each other (e.g., *XBP1* and *GNAS1*) or by performing common functions (e.g., tumor suppression by *INK4a/ARF*). At least 40 additional dual coding genes, which have been dubbed hallmarks of fascinating biology, are thought to exist in eukaryotes (Kovacs *et al.*, 2010; Chung *et al.*, 2007).

***INK4a/ARF* – One Locus, Two Important Tumor Suppressors**

Chromosome 9p21 where *INK4a/ARF* resides is the second most frequently inactivated locus in human cancers (Ruas and Peters, 1998). Since their discovery, the relative role of

p16INK4a versus ARF in cancer has been a matter of intense investigation and debate. The cumulative data show that *INK4a* and *ARF* are most often co-inactivated by silencing or homozygous deletion of the entire locus, frequently including *INK4b*. In some cancers, such as glioblastoma, loss of *INK4a/ARF* is almost always achieved by gene deletion whereas in other tumor types, such as renal cell carcinomas, promoter methylation and silencing is the predominant mechanism (Figure 3). Inactivating mutations (nonsense, missense, and frameshift) of the locus are also common in certain cancers, especially sporadic and familial melanoma, as well as stomach and pancreatic adenocarcinomas. Notably, The Cancer Genome Atlas (TCGA) and other databases compile tumor data (copy number, mutation, methylation, RNA sequencing, etc.) for *CDKN2A*, which fails to distinguish between *INK4a*- and *ARF*-specific alterations unless data for each gene in each tumor are carefully examined. Such analyses reveal that ARF is selectively disabled in various cancers by alterations of exon 1 β or ARF's open reading frame in exon 2 (Ozenne *et al.*, 2010; Maggi *et al.*, 2014). Nonetheless, the vast majority of tumor mutations specifically target p16INK4a sequences in exons 1 α and 2 (Kim and Sharpless, 2006; Gil and Peters, 2006). This fostered a lingering perception that p16INK4a would play a larger role in cancer than ARF.

Specific disruption of exon 1 β or exon 1 α in mice conclusively showed that both ARF and p16INK4a are tumor suppressors (Kamijo *et al.*, 1997; Kamijo *et al.*, 1999; Sharpless *et al.*, 2001, 2004; Krimpenfort *et al.*, 2001). In fact, *ARF*-null mice develop spontaneous tumors with a slightly shorter latency than *INK4a*-null animals (Table 1). *ARF*-null mouse embryo fibroblasts (MEFs) also evade replicative and Ras oncogene-induced senescence, unlike MEFs lacking *INK4a*, indicating that ARF loss removes key checkpoints that normally prevent cell immortalization and transformation. The fact that p16INK4a loss had no effect on MEF senescence was perplexing given its apparent importance in human cell senescence (Kim and Sharpless, 2006; Campisi, 2005). An explanation was provided by studies of *INK4a/ARF* and *INK4a* knockout mice that lacked or retained *INK4b*, which showed p15INK4b is upregulated in the absence of p16INK4a and can compensate for its loss (Krimpenfort *et al.*, 2007). Redundancy among the INK4 proteins is also illustrated by the robust

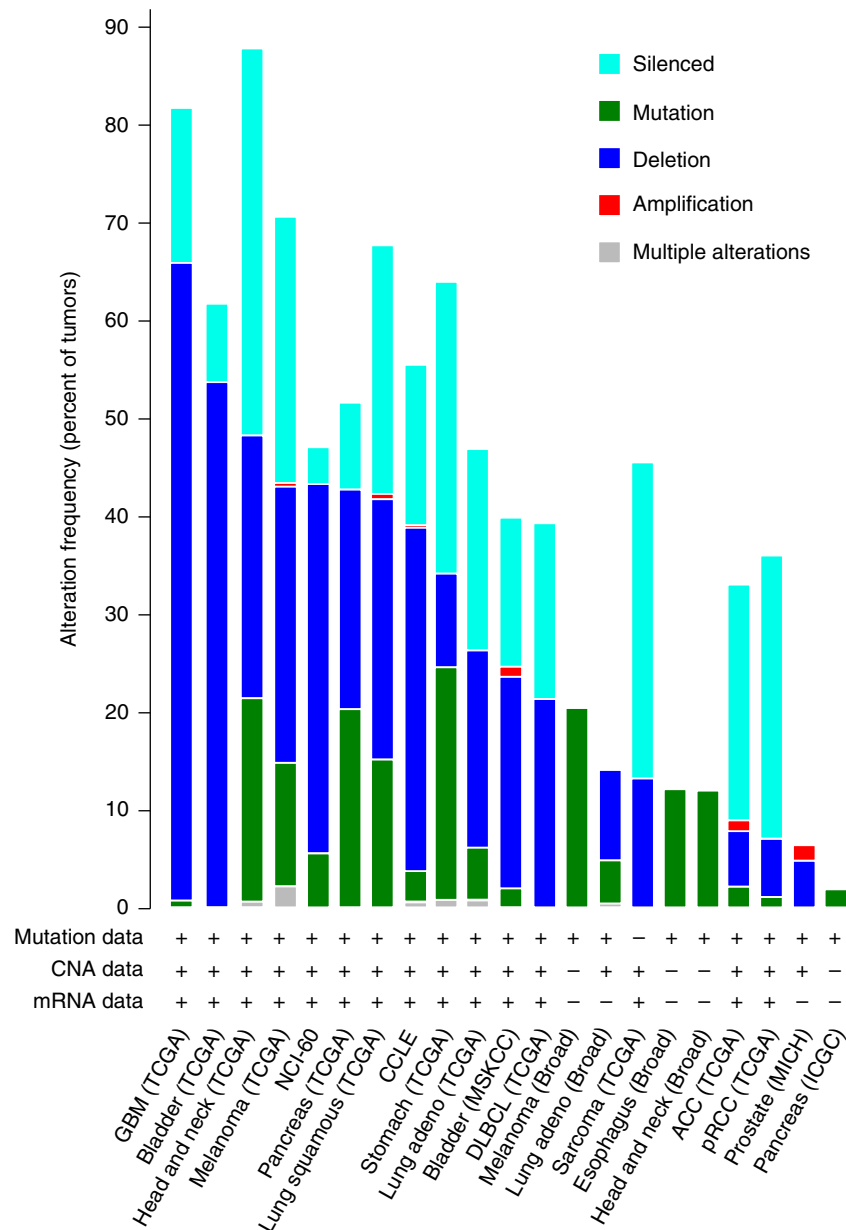


Figure 3 Frequency of genetic alterations at the *CDKN2A* locus in selected human malignancies. The relative frequency of gene deletion, mutation, silencing, and amplification occurring at the *CDKN2A* locus harboring INK4a and ARF is displayed for various cancers. The percentage of tumors with silencing of the locus was calculated based on the difference between the percent of tumors showing reduced *CDKN2A* mRNA (primarily obtained from RNA-Seq data but also some microarray results) and those with gene deletion. A z-score threshold of $+/- 0.5$ was used to determine if *CDKN2A* mRNA expression was down-regulated. Separate data sources interrogating the same tumor type can display different types and frequencies of *CDKN2A* genetic inactivation (e.g., compare results from two separate sources for melanoma, pancreatic, or bladder cancers). Such variation may reflect differences in patient populations examined (number, ethnicity, etc.) or simply that some studies solely examine copy number aberrations (CNA data) or sequence level mutation data (Mutation data), as indicated, with or without accompanying mRNA data. This graphic is derived from multiple data sources (listed) including TCGA (The Cancer Genome Atlas) and was compiled with the aid of the cBioPortal (Gao *et al.*, 2013; Cerami *et al.*, 2012). NCI-60, cell line database compiled by the U.S. National Cancer Institute; CCLL, Cancer Cell Line Encyclopedia; MSKCC, Memorial Sloan Kettering Cancer Center; Broad, Broad Institute of MIT and Harvard; MICH, Michigan Center for Translational Pathology; ICGC, International Cancer Genome Consortium; GBM, glioblastoma multiforme; Adeno, adenocarcinoma; ACC, adenoid cystic carcinoma; and pRCC, papillary renal cell carcinoma.

tumor phenotype (and lack of MEF senescence) in mice expressing a Cdk4 mutant (R24C) that fails to bind all INK4 proteins and escapes their inhibition (Sotillo *et al.*, 2001a, b; Rane *et al.*, 2002). Thus, ARF and p16INK4a are important

regulators of cellular proliferation and senescence whose loss facilitates cancer.

Interestingly, tumor predisposition is enhanced in mice lacking both p16INK4a and ARF relative to the single

Table 1 Phenotypes of *ARF*, *INK4a*, and *INK4a/ARF* deficient or transgenic mice^a

Genotype	Alteration	Tumor phenotype	References
<i>INK4a/ARF</i> ^{-/-}	Exon 2 and 3 deletion	Sarcomas and lymphomas within 8.5 months; MEFs divide rapidly, are immortal and transformed by oncogenic Ras	Serrano <i>et al.</i> (1996)
<i>ARF</i> ^{-/-}	Exon 1 β deletion	Sarcomas, lymphomas, and carcinomas within 9.5 months; MEFs divide rapidly, are immortal and transformed by oncogenic Ras	Kamijo <i>et al.</i> (1997, 1999)
<i>INK4a</i> ^{-/-}	Exon 1 α deletion	Sarcomas and lymphomas within 11 months; increased T cell proliferation; MEFs senesce, not transformed by oncogenic Ras	Sharpless <i>et al.</i> (2001)
<i>INK4a</i> ^{mt/mt}	Knock-in of exon 2 point mutant ^b (encodes a defective p16INK4a protein)	Low incidence of spontaneous tumors (B-cell lymphoma) within 17 months; melanomas develop upon loss of one allele of <i>ARF</i> ; MEFs senesce and not transformed by oncogenic Ras	Krimpenfort <i>et al.</i> (2001)
<i>Cdk4</i> ^{R24C/R24C}	Knock-in of Cdk4 mutant insensitive to INK4 proteins ^b	Various tumors with complete penetrance, including invasive melanoma; MEFs immortal and easily transformed	Sotillo <i>et al.</i> (2001a, b), Rane <i>et al.</i> (2002)
<i>INK4b</i> ^{-/-} ; <i>INK4a</i> ^{mt/mt} ; <i>ARF</i> ^{-/-}	Deletion of exon 1 α , 1 β , and <i>INK4b</i> exons plus <i>INK4a</i> exon 2 point mutant ^b	More tumor prone and wider tumor spectrum than <i>INK4a/ARF</i> ^{-/-} mice; MEFs more sensitive to oncogenic transformation than <i>INK4a/ARF</i> ^{-/-} cells	Krimpenfort <i>et al.</i> (2007)
<i>ARF</i> ^{-/-} ; E μ -Myc	Exon 1 β deletion; Myc transgene in B cells	Accelerated B-cell lymphomagenesis relative to E μ -Myc mice ^c ; mice die by 7 weeks of age versus 6 months in E μ -Myc mice	Eischen <i>et al.</i> (1999)
<i>INK4a/ARF</i> ^{-/-} ; Tyr-Ras ^{G12V}	Exon 2/3 deletion; mutant KRas in melanocytes	Cutaneous melanoma with short latency and high penetrance	Chin <i>et al.</i> (1997)
<i>ARF</i> ^{-/-} ; <i>p53</i> ^{-/-} ; <i>Mdm2</i> ^{-/-}	Triple knockout of <i>ARF</i> exon 1 β , <i>p53</i> and <i>Mdm2</i>	Much broader spectrum of tumors, and multiple tumors per animal, compared to mice lacking <i>ARF</i> or <i>p53</i> alone	Weber <i>et al.</i> (2000)
<i>ARF</i> ^{-/-} ; RIP-Tag	Exon 1 β deletion; SV40 T antigen expression in β cells	Accelerated formation of pancreatic neuroendocrine tumors with increased angiogenesis relative to RIP-Tag mice	Ulanet and Hanahan (2010)
<i>Pdx1-Cre</i> ; <i>KRas</i> ^{G12D} ; <i>INK4a/ARF</i> ^{fllox/fllox}	Tissue-specific expression of mutant KRas and <i>INK4a/ARF</i> exon 2/3 deletion	Accelerated development of pancreatic intraepithelial neoplasias (PanINs); metastatic pancreatic ductal adenocarcinoma (PDAC) and death by 2–3 months of age	Aguirre <i>et al.</i> (2003)
<i>ARF</i> ^{-/-} or <i>INK4a</i> ^{-/-} on <i>Ntv-a</i> or <i>Gtv-a</i> background	Neural- or astrocyte-specific expression of PDGFB in exon 1 β or 1 α deleted mice	Development of PDGFB-induced gliomas; much greater incidence of high-grade tumors in <i>ARF</i> ^{-/-} mice than <i>INK4a</i> ^{-/-} mice	Tchougounova <i>et al.</i> (2007)
<i>INK4/ARF</i> -tg	One extra allele of <i>INK4b-ARF-INK4a</i>	Increased cancer resistance, reduced production of sperm, normal aging	Matheu <i>et al.</i> (2004)
<i>INK4/ARF</i> -tg/tg	Two extra alleles of <i>INK4b-ARF-INK4a</i>	Further enhanced cancer resistance, absence of sperm, decrease in aging markers, and extended longevity	Matheu <i>et al.</i> (2009)

^aIncomplete listing of mouse models with altered *INK4a* and *ARF*.^bNaturally occurring point mutations found in human tumors, including melanomas.^cLoss of *INK4a* or *RB1* does not accelerate E μ -Myc lymphomagenesis (Lowe and Sherr, 2003).

knockout animals (Sharpless *et al.*, 2004; Serrano *et al.*, 1996). The synergism of dual p16INK4a and ARF loss on mouse tumorigenesis is consistent with the distinct molecular functions of each protein and with human clinical data showing their frequent co-inactivation. Still, it was disappointing that mice lacking ARF, p16INK4a, or both factors did not frequently develop the types of cancer commonly seen in humans bearing *INK4a/ARF* alterations (e.g., melanoma, glioblastoma, and pancreatic adenocarcinoma). Most mice lacking ARF and/or p16INK4a develop sarcomas or

lymphomas. It is widely appreciated that human tumors sustain multiple genetic changes concomitant with *INK4a/ARF* loss (Hanahan and Weinberg, 2011). Therefore, many studies have since tested the cooperative effects of *INK4a* and/or *ARF* loss when coupled with oncogene activation (e.g., Ras or Myc) or tumor suppressor inactivation (e.g., p53) (see Table 1). Such compound mouse mutants successfully model the tumor types seen in humans and demonstrate a critical role for p16INK4a and/or ARF inactivation in those cancers. Conversely, 'super-*INK4a/ARF*' mice expressing one or

two extra alleles of the *INK4a/ARF/INK4b* locus display a dose-dependent resistance to cancer (Matheu *et al.*, 2007, 2009).

ARF, a Multi-Faceted Tumor Suppressor

Unlike p16INK4a, which has one target (Cdk4/6) and acts within a single pathway, ARF binds to many proteins and regulates numerous cancer pathways (Figure 4), as reviewed in (Sherr, 2006; Ozenne *et al.*, 2010; Maggi *et al.*, 2014; Gil and Peters, 2006). The dominant mechanism by which ARF prevents cancer is through activation of p53. Widely considered the 'guardian of the genome,' p53 is the most commonly inactivated tumor suppressor gene in human cancers (Levine and Oren, 2009; Vousden and Prives, 2009). As a transcription factor, p53 regulates over a hundred genes involved in cell proliferation, survival, senescence, metabolism, and migration, among other processes in response to cellular stress. In so doing, p53 activates checkpoints that protect cells from hypoxia, telomere attrition, oncogene activation, DNA damage, and other insults that threaten the genome. It is normally kept at low levels in cells by Mdm2, an E3 ubiquitin ligase and p53 transcriptional target that promotes p53 ubiquitination and degradation (Senturk and Manfredi, 2012). Mdm2 can also block p53 transactivation and facilitate its translocation out of the nucleus. ARF inhibits Mdm2 and disrupts the Mdm2–p53 negative feedback loop, thereby stabilizing and activating p53. It does so by binding Mdm2 in either the nucleoplasm or nucleoli, physically sequestering Mdm2 away from p53 while also antagonizing Mdm2 ubiquitin ligase activity. As a result, ARF potently stimulates p53-dependent transcription and growth inhibition.

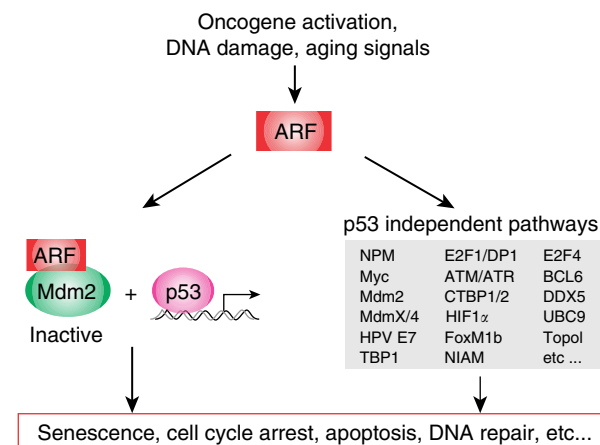


Figure 4 The ARF anti-proliferative pathways. ARF expression is induced by aberrant hyper-proliferative signals elicited by activated oncogenes or events during cellular aging, including DNA damage, that lead to *INK4a/ARF* derepression. Once expressed, ARF activates p53-mediated transcription through its inhibition of Mdm2. ARF also binds numerous proteins independently of p53 and consequently regulates many other cancer pathways. Through both p53-dependent and p53-independent mechanisms, ARF promotes senescence, cell cycle arrest, apoptosis, and DNA repair, among other antitumor activities, thereby helping to preserve genomic integrity.

Early models suggested that ARF acted solely through p53 to suppress cancer. This was based on initial studies suggesting mutually exclusive loss of ARF or p53 in human tumors, as well as observations that ARF caused robust cell cycle arrest and death in p53-positive (but not p53-negative) cells. However, many studies over the years have shown that ARF retains substantial tumor suppressive activity in the absence of p53 (Sherr, 2006; Ozenne *et al.*, 2010; Maggi *et al.*, 2014). Specifically, ARF expression can cause delayed apoptosis, cell cycle arrest or senescence in cells with nonfunctional p53. It can also inhibit ribosome biogenesis, reduce tumor cell migration and angiogenesis, activate pathways that repair damaged DNA, and promote autophagy ('self-eating'). For the latter, full-length ARF and a short mitochondrial ARF (smARF; translated from an internal methionine) isoform have been shown to localize to mitochondria and induce autophagy in a p53- and caspase-independent manner (Balaburski *et al.*, 2010). Collectively, these p53-independent activities of ARF reflect its binding to numerous proteins besides Mdm2. Approximately 40 ARF partners have been identified to date (Figure 4).

There is still much to learn about the biological significance of ARF's many interactions. Nevertheless, the diversity of ARF binding proteins underscores the breadth of ARF's p53-dependent and p53-independent tumor suppressive activities (Sherr, 2006; Maggi *et al.*, 2014; Ozenne *et al.*, 2010). ARF associates with transcriptional activators (e.g., Myc, HIF1α, FoxM1b, E2F1, DP1, YY1, p53, and p63), transcriptional repressors (e.g., E2F4, BCL6, CtBP1/2), checkpoint kinases (e.g., ATM and ATR), ligases involved in ubiquitination or sumoylation (e.g., Mdm2, ARF-BP1, and UBC9), factors controlling ribosomal RNA and DNA synthesis (e.g., NPM, nucleolin, TTF-1, DDX5, and UBF), DNA modifying enzymes (e.g., WRN and Topoisomerase I (TopoI)), acetyltransferases (e.g., Tip60), and viral proteins (e.g., TBP1 and human papilloma virus (HPV) 16 E7), among other proteins. Many of those partners are oncogenic proteins, such as Myc and FoxM1b, whose tumor-promoting activities are inhibited by ARF. Other partners suppress cancer (e.g., Tip60 and ATM) and are activated by ARF as part of checkpoints that protect against hyper-proliferative signaling (i.e., oncogene activation) or DNA damage.

Notably, cooperative effects of ARF's p53-dependent and p53-independent activities are expected since many ARF partners affect p53 signaling while also influencing other cellular processes separate from p53. Mdm2 and its relative, MdmX (also called Mdm4), for example, both inhibit p53 yet can also function independently of p53 in the DNA damage response to impede DNA repair and increase genome instability (Melo and Eischen, 2012; Senturk and Manfredi, 2012). Similarly, NIAM is a 'nuclear interactor of ARF and Mdm2' that is able to activate p53 through inhibition of Mdm2 (Reed *et al.*, 2014a), yet it also acts through undefined mechanisms to suppress proliferation and promote chromosomal stability in the absence of p53 (Tompkins *et al.*, 2007). Perhaps the best example is nucleophosmin (NPM1), one of the better validated ARF interacting proteins (Grisendi *et al.*, 2006). NPM1 is an abundant nucleolar phosphoprotein that inhibits ARF-mediated p53 activation by retaining ARF in nucleoli and blocking its sequestration of Mdm2 away from p53. Independently of p53, NPM1 is also required for ARF stability and nucleolar

localization (where most ARF protein resides) while ARF upregulation blocks NPM1 nucleocytoplasmic shuttling, thereby interfering with ribosome nuclear export (Sherr, 2006; Maggi *et al.*, 2014). ARF–NPM1 co-regulation therefore impacts p53 signaling as well as other pathways affecting ribosome function and cell proliferation.

The physiological importance of ARF's many p53-independent activities has been demonstrated by the striking phenotypes of mouse tumor models specifically lacking ARF on a p53-deficient background (see Table 1). The first realization that ARF suppresses cancer in living animals independent of p53 came from a seminal study showing that *ARF/Mdm2/p53* triple knockout mice develop a wider range of tumor types and multiple primary tumors per animal relative to *ARF*^{-/-}, *p53*^{-/-} or *Mdm2/p53*^{-/-} mice (Weber *et al.*, 2000). Others showed that ARF (but not p53) is required for oncogene-induced senescence in benign nevi, and its loss combined with N-Ras activation promotes melanocyte transformation and melanoma development (Ha *et al.*, 2007). In a mouse model of pancreatic neuroendocrine tumors in which p53 and RB1 are both inactivated by the SV40 large T antigen, ARF deficiency accelerates tumor formation by promoting tumor angiogenesis (Ulanet and Hanahan, 2010). These *in vivo* studies exemplify how ARF can suppress cancer independent of p53 through multiple mechanisms, such as promoting senescence in benign neoplasias and inhibiting blood vessel formation within existing tumors.

Regulation and Loss of *INK4a/ARF* in Cancer

Genetic inactivation of *INK4a/ARF* is a frequent event in tumor development, observed in nearly half of all human malignancies. However, functional loss of p16INK4a and/or ARF in tumors can be achieved through a variety of other mechanisms

including enhanced promoter methylation and silencing, loss of transcriptional activators, reduced translation, increased protein degradation, and altered expression of their target/effector proteins (Figure 5). In fact, alterations of *INK4a/ARF* or its regulators and effector pathways are thought to occur in essentially all human cancers. In many cases, the entire locus is lost by genetic deletion or coordinated transcriptional repression; however, the separate control of *INK4a* and *ARF* by unique factors or signals reveals how a significant number of tumors with an intact *INK4a/ARF* locus may lack ARF function while retaining p16INK4a activity (or vice versa).

In normal cells, *INK4a/ARF* is tightly silenced by chromatin-associated Polycomb repressive complexes (PRC1 and PRC2) and histone deacetylase (HDAC) complexes (Gil and Peters, 2006; Kim and Sharpless, 2006; Aguilo *et al.*, 2011). This is important for normal cell proliferation, stem cell self-renewal, and tissue homeostasis since aberrant derepression of *INK4a/ARF* (as seen in mouse knockouts lacking PRC components Bmi1, CBX7 or EZH2) causes significant hematological, pancreatic, and neurological defects. PRC recruitment to *INK4a/ARF* is mediated by binding of the Polycomb group protein, CBX7, to the long noncoding antisense RNA *ANRIL* (Yap *et al.*, 2010), which spans the *INK4b/INK4a/ARF* locus (see Figure 2). As normal cells senesce or encounter stresses that induce senescence, such as oncogene activation or DNA damage, the expression of *ANRIL* and *EZH2* declines leading to displacement of the repressive complexes and activated transcription of *INK4a/ARF*. The consequent induction of p16INK4a and ARF signaling provides a powerful defense against cellular transformation and tumor progression. In cancer cells, however, the levels of *ANRIL*, CBX7, and other PRC components can be elevated causing improper maintenance of *INK4a/ARF* repression.

Induction of *INK4a/ARF* is principally stimulated by hyperproliferative signals emanating from activated oncogenes, such

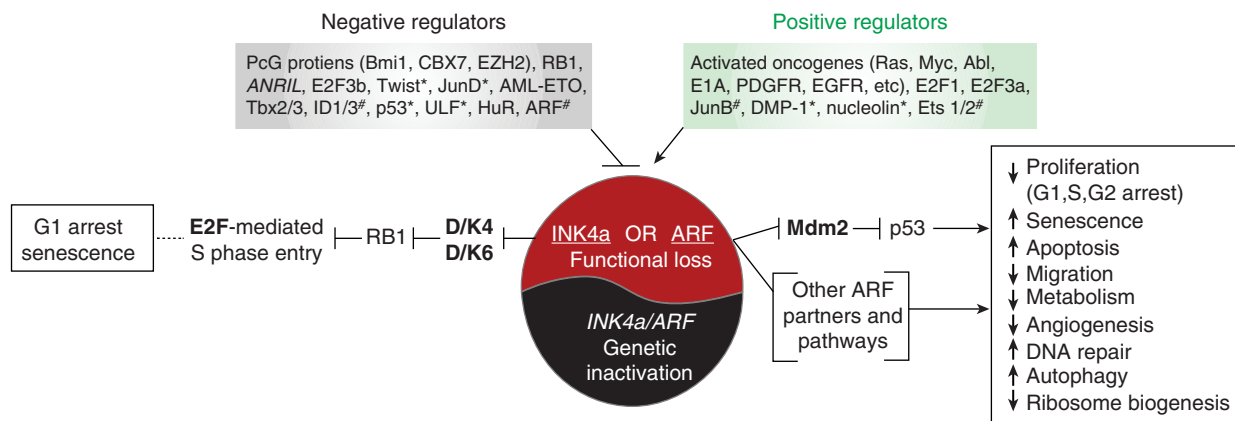


Figure 5 Multiple targets for disruption of p16INK4a and ARF function in human cancers. In normal cells, the *INK4a/ARF* locus is tightly silenced by repressive complexes containing Polycomb group (PcG) proteins, such as Bmi1, CBX7, and EZH2. Derepression of the locus is induced by hyper-proliferative signals emitted by activated oncogenes, such as Ras or Myc. This leads to p16INK4a and ARF expression, activation of their respective tumor suppressor pathways, and induction of senescence or apoptosis (acute response) or reduced cell proliferation, DNA repair, migration, metabolism, etc. Most cells maintain this ‘inhibited state.’ However, the few cells that escape those constraints can do so through a variety of mechanisms: (1) genetic deletion or mutation of *INK4a/ARF*, (2) promoter silencing due to upregulation of factors in the PcG repressive complexes and DNA hyper-methylation, (3) loss of key activators (e.g., the DMP-1 transcription factor for ARF) or effectors (e.g., RB1 for p16INK4a or p53 for ARF), and (4) hyper-activation of oncogenic target proteins (examples are bolded). An incomplete listing of negative and positive regulators of *INK4a/ARF* is shown, with specific regulators of ARF (*) and p16INK4a (#) indicated.

as Ras or Myc, thereby engaging RB1, p53, and ARF/p53-independent checkpoints that protect against unrestrained cell proliferation, genomic mutation, and cancer (Kim and Sharpless, 2006; Gil and Peters, 2006; Ozenne *et al.*, 2010; Sherr, 2012). For the most part, oncogenic stress causes the coordinated upregulation of both p16INK4a and ARF by transcriptional activators, such as E2F1. The role of *INK4a/ARF* in DNA damage checkpoints is less clear and more controversial. Various types of DNA damage induce p16INK4a, ARF, or both proteins; however, their upregulation is generally delayed and may be more important for mediating long-term responses (including DNA repair or senescence) following chronic and/or low-level DNA damage as opposed to acute DNA damage. Consistent with that notion, ARF was found to be required for p53-dependent tumor suppressive activities in irradiated mice once the acute, pathological response to DNA damage had ended (Christophorou *et al.*, 2006).

While *INK4a* and *ARF* are often co-regulated, a large number of factors selectively regulate the individual genes or transcripts. Several *INK4a*-specific transcriptional activators (e.g., JunB, Ets1/2) and repressors (e.g., Id1, Id3) have been identified, as have *ARF*-specific transcriptional activators (e.g., DMP-1, AML1) and repressors (e.g., Twist, AML-ETO). At the posttranscriptional level, the RNA binding protein HuR (which is upregulated in various cancers and inhibits senescence) negatively regulates *INK4a* and *ARF* mRNAs but through distinct mechanisms that depend on cell context (Maggi and Weber, 2013). In human diploid fibroblasts, HuR binds and destabilizes the *INK4a* mRNA whereas in MEFs it binds the 5' UTR of the *ARF* mRNA and blocks its translation by preventing nucleolin-*ARF* mRNA association and export to the cytoplasm. The basis for differential HuR regulation of *INK4a* and *ARF* in mouse and human cells is unclear, but it is not entirely unusual since Ras activation selectively induces p16INK4a in human cells despite upregulating both p16INK4a and ARF in rodent cells (Gil and Peters, 2006; Kim and Sharpless, 2006).

Other mechanisms distinctly controlling ARF or p16INK4a may play an important role in cancer. UFL, an 'ubiquitin ligase for ARF,' is upregulated in tumors and lowers ARF levels by promoting its lysine-independent ubiquitination and proteasomal degradation (Chen *et al.*, 2010). Intriguingly, ARF can promote p16INK4a degradation via its interaction with the REG γ subunit of the proteasome (Kobayashi *et al.*, 2013). Such crosstalk between the *INK4a/ARF* products may constitute an autoregulatory feedback loop to reduce p16INK4a levels, enabling cell cycle reentry following a successful checkpoint response and thereby delaying aging (see below). ARF is also subject to negative feedback regulation but it occurs at the transcriptional level due to p53-mediated repression. Specifically, ARF is induced by activated oncogenes and activates p53; subsequently, p53 binds the promoter for *ARF* (not *INK4a*) and recruits HDACs and Polycomb group proteins to repress its transcription (Robertson and Jones, 1998; Zeng *et al.*, 2011). Those results explain why cancer cells lacking p53 often express high levels of ARF.

Increased expression of ARF in cells lacking p53 is a beneficial barrier against tumorigenesis due to its many p53-independent activities, as discussed earlier. This was effectively illustrated in a recent study that showed ARF upregulation

following acute p53 loss inhibits interferon β (IFN- β) production and signaling, markedly restricting the proliferation and tumorigenicity of those cells (Forys *et al.*, 2014). Unfortunately, there is strong selective pressure for cells to escape such constraints and remove ARF to enable proliferation, facilitating the outgrowth of tumor cells with activated IFN- β signaling. In fact, analyses of human triple-negative breast cancer samples revealed that more than 50% have co-inactivation of ARF and p53 as well as hyperactive IFN signaling. Those findings may be useful in the clinic since they suggest that therapies targeting IFN- β effectors, such as STAT1, would be appropriate for treating tumors with combined p53 and ARF inactivation.

INK4a/ARF in Aging

Genome-wide association studies (GWAS) in millions of patients implicate the *INK4a/ARF* locus in a range of aging-related diseases (Jeck *et al.*, 2012). The data identify a hotspot of single nucleotide polymorphisms (SNPs) on chromosome 9p21.3 spanning the *INK4a/ARF* locus (and the sequences encoding p15INK4b and *ANRIL*) that is associated with increased risk for type 2 diabetes, glaucoma, late onset Alzheimer's, frailty, and atherosclerotic pathologies including stroke, myocardial infarction, and aortic aneurysm (see Figure 2). Exactly how the different SNPs affect *ANRIL* and *INK4a/ARF* expression, however, is not well understood. Some SNPs associated with a higher risk of atherosclerosis increase the expression of *ANRIL* isoforms and circular species, thereby enhancing repression of *INK4a/ARF* in peripheral blood T cells and causing reduced p15INK4b, p16INK4a, and ARF levels (Burd *et al.*, 2010). Other SNPs may reduce *ANRIL* expression, which would be predicted to increase p16INK4a and ARF levels. That seems likely for SNPs associated with type 2 diabetes since precocious expression of p16INK4a and ARF in pancreatic β cells of *EZH2*-null mice impairs β cell proliferation, reduces insulin production, and promotes type 2 diabetes (Chen *et al.*, 2009). Rescue of those defects by *INK4a/ARF* deletion established a causative role for p16INK4a and/or ARF in that process.

The observations above suggest that loss and gain of *INK4a/ARF* expression can each promote aging phenotypes, so does the locus accelerate or delay aging? The answer is complicated and still unfolding. On the one hand, substantial evidence suggests *INK4a/ARF* enhances aging (reviewed in Sorrentino *et al.*, 2014; Sherr, 2012). This includes findings that p16INK4a and ARF levels increase in aging mouse and human tissues, promote cellular senescence (which accumulates with organismal aging), and are regulated by age-modifying stimuli such as caloric restriction. Also, elevated p16INK4a and ARF levels have been linked to the age-dependent decline in normal stem cell function in the neural, pancreatic, muscle, and hematopoietic stem cell compartments. On the other hand, transgenic mice expressing two extra copies of the *INK4b/ARF/INK4a* locus (*INK4/ARF*-tg/tg) have reduced signs of aging and a longer life span, with the delayed aging phenotype distinct from the reduced cancer incidence seen in those animals (Matheu *et al.*, 2009). How can such apparently contradictory findings be reconciled?

One explanation is that p16INK4a and ARF have separate, opposing roles in aging and that the pro-aging effects of p16INK4a may be countered by antiaging activities of ARF. For instance, p16INK4a (not ARF) expression is increased in aging T-cell progenitors, significantly reduces their number and promotes thymic involution (Berent-Maoz *et al.*, 2012), a well-recognized change in the elderly linked with impaired immunity. Also, selective ablation of *INK4a* or *ARF* in a progeroid mouse model of premature aging (*BubR1*-deficient mice) showed that p16INK4a promotes senescence and aging, whereas ARF suppresses those processes (Baker *et al.*, 2008, 2011). The antiaging effect of ARF in this setting may be partly mediated through negative regulation of p16INK4a levels since *ARF* deletion increases p16INK4a expression in tissues displaying accelerated aging (muscle, fat, and eye). ARF may also protect against aging by inducing a p53 transcriptional response involving antioxidant gene expression which promotes quiescence, DNA repair, and cell survival. The fact that p53 loss phenocopies ARF loss in *BubR1* mutant mice supports that notion (Baker *et al.*, 2008), as does the upregulation of antioxidant p53 targets and delayed aging of *ARF/p53* transgenic mice (Matheu *et al.*, 2007).

It is also possible that both *INK4a* and *ARF* have antiaging effects although the timing and context of their expression may be important. Although acute *INK4a/ARF* upregulation during oncogenesis may cause excessive cell apoptosis or senescence that compromises tissue integrity and accelerates aging, the progressive low-level activation of the locus during aging could be beneficial by slowing cell proliferation rates or triggering a reversible quiescent arrest, thereby delaying the exhaustion of stem cell pools (Matheu *et al.*, 2009). The mitochondrial smARF form of ARF may also inhibit aging. Studies in neurons showed that smARF is a critical upstream activator of a Parkin/PINK1 mitophagy pathway that eliminates damaged mitochondria from cells (Grenier *et al.*, 2014). This is relevant because impaired Parkin/PINK1-mediated mitophagy in neurons is thought to cause the accumulation of damaged mitochondria that occurs during aging and is associated with Parkinson's disease (PD). Whether or not smARF levels and activity are impaired in PD brains and how this pathway contributes to neurodegeneration in PD remains to be determined. Indeed, more work is needed to define the individual and combined roles of p16INK4a, ARF, and smARF in aging and age-related disorders. Nonetheless, the cumulative evidence establishes that *INK4a/ARF* is a bona fide aging locus.

Unique Roles of ARF in Development

INK4a/ARF repression is necessary and widespread throughout most tissues during development. However, there are a few cell types where transient ARF expression is critical for proper development, particularly in the eye and testes (Gromley *et al.*, 2009). *ARF*-null mice are blind with smaller eyes than *ARF*-positive animals, mimicking a human eye disease called persistent hyperplastic primary vitreous (PHPV) (McKeller *et al.*, 2002). The blindness results from elevated expression of platelet-derived growth factor receptor β (PDGFR β) in perivascular cells within the vitreous of the eye, which causes their

aberrant proliferation and prevents the regression of the hyaloid vascular system required for proper retina and lens formation (Reed *et al.*, 2014b). At the molecular level, ARF suppresses PDGFR β expression through complementary mechanisms that are p53-dependent (repression of PDGFR β transcription) and p53-independent (upregulation of the miR-34a microRNA and inhibition of PDGFR β translation).

Paradoxically, ARF expression in male germ cells has no effect on proliferation and instead promotes the survival of spermatogonia by blocking p53-mediated apoptosis, contrasting with its role as a p53 activator in cancer (Churchman *et al.*, 2011). *INK4a*-null mice do not display developmental defects. However, it is notable that overexpression of p16INK4a, p15INK4b, and ARF in transgenic *INK4/ARF*-tg/tg mice causes male sterility (see Table 1; Matheu *et al.*, 2009). This is thought to reflect a proliferative defect of the spermatogonia due to increased *INK4a/b*-RB1 mediated growth inhibition rather than ARF-p53 effects.

Conclusions

In summary, ARF and p16INK4a are major barriers to the outgrowth of cancer cells. Their induction by various cell stresses, particularly hyper-proliferative stimuli and aging signals, and consequent activation of the p53 and RB1 tumor suppressors is critical for their anticancer function. In the absence of p53, ARF can also function through many other pathways to induce senescence and suppress tumor cell survival, migration, angiogenesis, etc. Through those pathways, each protein engages protective checkpoints that inhibit cancer cell outgrowth by reversible cell cycle arrest (and DNA repair) of mildly damaged cells or irreversible removal of severely damaged cells via senescence or apoptosis, thereby ensuring genomic integrity.

The importance of these pathways in human cancer is demonstrated by their frequent disruption in tumors. Indeed, functional inactivation of p16INK4a, ARF, p53, and/or RB1 is thought to occur in most, if not all, human tumors due to genetic/epigenetic inactivation and alterations targeting their signaling pathways (Figure 5). Advancing our understanding of these pathways and employing that knowledge to improve tumor diagnosis, prognosis, and treatment is the ultimate goal. For example, preclinical studies suggest that ARF peptides targeting the FoxM1b oncogenic protein have beneficial anti-tumor activity in mouse models of lymphoma and sarcoma (Wang *et al.*, 2013). Selective inhibitors of Cdk4/6 are further along in development and show significant clinical promise in treating RB1-positive tumors that lack functional p16INK4a (Flaherty *et al.*, 2012).

Derepression of the *INK4a/ARF* locus is also important in cellular and organismal aging although its impact on age-related pathologies and longevity is complex. Whether or not *INK4a/ARF* is a benefit or liability to aging is controversial and still under investigation. The context (timing, level, cause, and tissue) of *INK4a/ARF* regulation and whether p16INK4a, ARF or both proteins are expressed may be critical in dictating outcome. That said, the fact that increased *INK4/ARF* gene dosage in mice leads to delayed aging and an extended life span (as well as increased cancer resistance) makes a

persuasive case that more *INK4a/ARF* is better, at least if you are a mouse (and female, since the male transgenic mice are sterile). For humans, any potential benefit of manipulating *INK4a/ARF* to prolong life (beyond the protection against cancer) requires a firmer grasp on how, when, and where p16^{INK4a} and/or ARF expression influences the aging process.

See also: Nucleic Acid Synthesis/Breakdown: RNA Synthesis/
Function: Comparison of Bacterial and Eukaryotic Replisome
Components

References

- Aguilo, F., Zhou, M.-M., Walsh, M.J., 2011. Long noncoding RNA, polycomb, and the ghosts haunting *INK4b-ARF-INK4a* expression. *Cancer Research* 71, 5365–5369.
- Aguirre, A.J., Bardeesy, N., Sinha, M., *et al.*, 2003. Activated *Kras* and *Ink4a/Arf* deficiency cooperate to produce metastatic pancreatic ductal adenocarcinoma. *Genes & Development* 17, 3112–3126.
- Baker, D.J., Perez-Terzic, C., Jin, F., *et al.*, 2008. Opposing roles for p16^{INK4a} and p19^{Arf} in senescence and ageing caused by BubR1 insufficiency. *Nature Cell Biology* 10, 825–836.
- Baker, D.J., Wijshake, T., Tchkonja, T., *et al.*, 2011. Clearance of p16^{INK4a}-positive senescent cells delays ageing-associated disorders. *Nature* 479, 232–236.
- Balaburski, G.M., Hontz, R.D., Murphy, M.E., 2010. p53 and ARF: Unexpected players in autophagy. *Trends in Cell Biology* 20, 363–369.
- Berent-Maoz, B., Montecino-Rodriguez, E., Signer, R.A.J., Dorshkind, K., 2012. Fibroblast growth factor-7 partially reverses murine thymocyte progenitor aging by repression of *Ink4a*. *Blood* 119, 5715–5721.
- Burd, C.E., Jeck, W.R., Liu, Y., *et al.*, 2010. Expression of linear and novel circular forms of an *INK4/ARF*-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genetics* 6 (12), e1001233.
- Burkhardt, D.L., Sage, J., 2008. Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nature Reviews Cancer* 8, 671–682.
- Byeon, I.J.L., Li, J.N., Ericson, K., *et al.*, 1998. Tumor suppressor p16^{INK4a}: Determination of solution structure and analyses of its interaction with cyclin-dependent kinase 4. *Molecular Cell* 1, 421–431.
- Campisi, J., 2005. Senescent cells, tumor suppression, and organismal aging: Good citizens, bad neighbors. *Cell* 120, 513–522.
- Cerami, E., Gao, J., Dogrusoz, U., *et al.*, 2012. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discovery* 2, 401–404.
- Chen, D., Shan, J., Zhu, W.G., Qin, J., Gu, W., 2010. Transcription-independent ARF regulation in oncogenic stress-mediated p53 responses. *Nature* 464, 624–627.
- Chen, H., Gu, X., Su, I.H., *et al.*, 2009. Polycomb protein Ezh2 regulates pancreatic beta-cell *Ink4a/Arf* expression and regeneration in diabetes mellitus. *Genes & Development* 23, 975–985.
- Chen, P., Zindy, F., Abdala, C., *et al.*, 2003. Progressive hearing loss in mice lacking the cyclin-dependent kinase inhibitor *Ink4d*. *Nature Cell Biology* 5, 422–426.
- Chin, L., Pomerantz, J., Polsky, D., *et al.*, 1997. Cooperative effects of *INK4a* and *ras* in melanoma susceptibility *in vivo*. *Genes & Development* 11, 2822–2834.
- Christophorou, M.A., Ringshausen, I., Finch, A.J., Swigart, L.B., Evan, G.I., 2006. The pathological response to DNA damage does not contribute to p53-mediated tumour suppression. *Nature* 443, 214–217.
- Chung, W.-Y., Wadhawan, S., Szklarczyk, R., Pond, S.K., Nekrutenko, A., 2007. A first look at ARFome: Dual-coding genes in mammalian genomes. *Plos Computational Biology* 3, 855–861.
- Churchman, M.L., Roig, I., Jasin, M., Keeney, S., Sherr, C.J., 2011. Expression of Arf tumor suppressor in spermatogonia facilitates meiotic progression in Male germ cells. *PLoS Genetics* 7, e1002157.
- Collado, M., Blasco, M.A., Serrano, M., 2007. Cellular senescence in cancer and aging. *Cell* 130, 223–233.
- Darbro, B.W., Schneider, G.B., Klingelutz, A.J., 2005. Co-regulation of p16^{INK4a} and migratory genes in culture conditions that lead to premature senescence in human keratinocytes. *Journal of Investigative Dermatology* 125, 499–509.
- Duro, D., Bernard, O., Della Valle, V., Berger, R., Larsen, C.J., 1995. A new type of p16^{INK4a/MTS1} gene transcript expressed in B-cell malignancies. *Oncogene* 11, 21–29.
- Eischen, C.M., Weber, J.D., Roussel, M.F., Sherr, C.J., Cleveland, J.L., 1999. Disruption of the ARF-Mdm2-p53 tumor suppressor pathway in Myc-induced lymphomagenesis. *Genes & Development* 13, 2658–2669.
- Flaherty, K.T., Lorusso, P.M., Demichele, A., *et al.*, 2012. Phase I, dose-escalation trial of the oral cyclin-dependent kinase 4/6 inhibitor PD 0332991, administered using a 21-day schedule in patients with advanced cancer. *Clinical Cancer Research* 18, 568–576.
- Forys, J.T., Kuzmicki, C.E., Saporita, A.J., *et al.*, 2014. ARF and p53 coordinate tumor suppression of an oncogenic IFN- β -STAT1-ISG15 signaling axis. *Cell Reports* 7, 514–526.
- Freson, K., Jaeken, J., Van Helvoirt, M., *et al.*, 2003. Functional polymorphisms in the paternally expressed XL alpha s and its cofactor ALEX decrease their mutual interaction and enhance receptor-mediated cAMP formation. *Human Molecular Genetics* 12, 1121–1130.
- Gao, J., Aksoy, B.A., Dogrusoz, U., *et al.*, 2013. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Science Signaling* 6, p1.
- Gil, J., Peters, G., 2006. Regulation of the *INK4b-ARF-INK4a* tumour suppressor locus: All for one or one for all. *Nature Reviews Molecular Cell Biology* 7, 667–677.
- Grenier, K., Kontogiannia, M., Fon, E.A., 2014. Short mitochondrial ARF triggers Parkin/PINK1-dependent mitophagy. *Journal of Biological Chemistry* 289, 29519–29530.
- Grisendi, S., Mecucci, C., Falini, B., Pandolfi, P.P., 2006. Nucleophosmin and cancer. *Nature Reviews Cancer* 6, 493–505.
- Gromley, A., Churchman, M.L., Zindy, F., Sherr, C.J., 2009. Transient expression of the Arf tumor suppressor during male germ cell and eye development in Arf-Cre reporter mice. *Proceedings of the National Academy of Sciences of the United States of America* 106, 6285–6290.
- Ha, L., Ichikawa, T., Anver, M., *et al.*, 2007. ARF functions as a melanoma tumor suppressor by inducing p53-independent senescence. *Proceedings of the National Academy of Sciences of the United States of America* 104, 10968–10973.
- Hanahan, D., Weinberg, R.A., 2011. Hallmarks of cancer: The next generation. *Cell* 144, 646–674.
- Harbour, J.W., Dean, D.C., 2000. The Rb/E2F pathway: Expanding roles and emerging paradigms. *Genes & Development* 14, 2393–2409.
- Hunter, T., Pines, J., 1994. Cyclins and cancer. II: Cyclin D and CDK inhibitors come of age. *Cell* 79, 573–582.
- Iwakoshi, N.N., Lee, A.H., Glimcher, L.H., 2003. The X-box binding protein-1 transcription factor is required for plasma cell differentiation and the unfolded protein response. *Immunological Reviews* 194, 29–38.
- Jeck, W.R., Siebold, A.P., Sharpless, N.E., 2012. Review: A meta-analysis of GWAS and age-associated diseases. *Aging Cell* 11, 727–731.
- Jeffrey, P.D., Tong, L., Pavletich, N.P., 2000. Structural basis of inhibition of CDK-cyclin complexes by *INK4* inhibitors. *Genes & Development* 14, 3115–3125.
- Kamijo, T., Bodner, S., Van De Kamp, E., Randle, D.H., Sherr, C.J., 1999. Tumor spectrum in ARF-deficient mice. *Cancer Research* 59, 2217–2222.
- Kamijo, T., Zindy, F., Roussel, M.F., *et al.*, 1997. Tumor suppression at the mouse *INK4a* locus mediated by the alternative reading frame product p19^{ARF}. *Cell* 91, 649–659.
- Kim, W.Y., Sharpless, N.E., 2006. The regulation of *INK4/ARF* in cancer and aging. *Cell* 127, 265–275.
- Klemke, M., Kehlenbach, R.H., Huttner, W.B., 2001. Two overlapping reading frames in a single exon encode interacting proteins – a novel way of gene usage. *EMBO Journal* 20, 3849–3860.
- Kobayashi, T., Wang, J., Al-Ahmadie, H., Abate-Shen, C., 2013. ARF regulates the stability of p16 protein via REGγ-dependent proteasome degradation. *Molecular Cancer Research* 11, 828–833.
- Kovacs, E., Tompa, P., Liliom, K., Kalmar, L., 2010. Dual coding in alternative reading frames correlates with intrinsic protein disorder. *Proceedings of the National Academy of Sciences* 107, 5429–5434.
- Krimpenfort, P., Ijpenberg, A., Song, J.-Y., *et al.*, 2007. p15^{INK4b} is a critical tumour suppressor in the absence of p16^{INK4a}. *Nature* 448, 943–946.
- Krimpenfort, P., Quon, K.C., Mooi, W.J., Loonstra, A., Berns, A., 2001. Loss of p16^{INK4a} confers susceptibility to metastatic melanoma in mice. *Nature* 413, 83–86.
- Krishnamurthy, J., Torrice, C., Ramsey, M.R., *et al.*, 2004. *Ink4a/Arf* expression is a biomarker of aging. *Journal of Clinical Investigation* 114, 1299–1307.
- Lee, Y.-M., Sicinski, P., 2006. Targeting cyclins and cyclin-dependent kinases in cancer – lessons from mice, hopes for therapeutic applications in human. *Cell Cycle* 5, 2110–2114.

- Levine, A.J., Oren, M., 2009. The first 30 years of p53: Growing ever more complex. *Nature Reviews Cancer* 9, 749–758.
- Lowe, S.W., Sherr, C.J., 2003. Tumor suppression by Ink4a-Arf: Progress and puzzles. *Current Opinion in Genetics & Development* 13, 77–83.
- Maggi Jr., L.B., Weber, J.D., 2013. Forget transcription: Translation is where the action is. *Molecular and Cellular Biology* 33, 1884–1885.
- Maggi Jr., L.B., Winkler, C.L., Miceli, A.P., *et al.*, 2014. ARF tumor suppression in the nucleolus. *Biochimica et Biophysica Acta (BBA) – Molecular Basis of Disease* 1842, 831–839.
- Malumbres, M., Barbacid, M., 2009. Cell cycle, CDKs and cancer: A changing paradigm. *Nature Reviews Cancer* 9, 153–166.
- Mao, L., Merlo, A., Bedi, G., *et al.*, 1995. A novel p16INK4A transcript. *Cancer Research* 55, 2995–2997.
- Matheu, A., Maraver, A., Collado, M., *et al.*, 2009. Anti-aging activity of the Ink4/Arf locus. *Aging Cell* 8, 152–161.
- Matheu, A., Maraver, A., Klatt, P., *et al.*, 2007. Delayed ageing through damage protection by the Arf/p53 pathway. *Nature* 448, 375–379.
- Matheu, A., Pantoja, C., Efeyan, A., *et al.*, 2004. Increased gene dosage of Ink4a/Arf results in cancer resistance and normal aging. *Genes & Development* 18, 2736–2746.
- McKeller, R.N., Fowler, J.L., Cunningham, J.J., *et al.*, 2002. The Arf tumor suppressor gene promotes hyaloid vascular regression during mouse eye development. *Proceedings of the National Academy of Sciences of the United States of America* 99, 3848–3853.
- Melo, A.N., Eischen, C.M., 2012. Protecting the genome from Mdm2 and Mdmx. *Genes & Cancer* 3, 283–290.
- Ozenne, P., Eymin, B., Brambilla, E., Gazzeri, S., 2010. The ARF tumor suppressor: Structure, functions and status in cancer. *International Journal of Cancer* 127, 2239–2247.
- Quelle, D.E., Zindy, F., Ashmun, R.A., Sherr, C.J., 1995. Alternative reading frames of the INK4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest. *Cell* 83, 993–1000.
- Rane, S.G., Cosenza, S.C., Mettus, R.V., Reddy, E.P., 2002. Germ line transmission of the Cdk4(R24C) mutation facilitates tumorigenesis and escape from cellular senescence. *Molecular and Cellular Biology* 22, 644–656.
- Reed, S.M., Hagen, J., Tompkins, V.S., *et al.*, 2014a. Nuclear interactor of ARF and Mdm2 regulates multiple pathways to activate p53. *Cell Cycle* 13, 1288–1298.
- Reed, S.M., Quelle, F.W., Quelle, D.E., 2014b. ARF sees Pdgfr beta through the miR. *Cell Cycle* 13, 1520–1521.
- Ressler, S., Bartkova, J., Niederegger, H., *et al.*, 2006. p16(INK4A) is a robust *in vivo* biomarker of cellular aging in human skin. *Aging Cell* 5, 379–389.
- Robertson, K.D., Jones, P.A., 1998. The human ARF cell cycle regulatory gene promoter is a CpG island which can be silenced by DNA methylation and down-regulated by wild-type p53. *Molecular and Cellular Biology* 18, 6457–6473.
- Ruas, M., Peters, G., 1998. The p16INK4a/CDKN2A tumor suppressor and its relatives. *Biochimica et Biophysica Acta* 1378, F115–F177.
- Russo, A.A., Tong, L., Lee, J.O., Jeffrey, P.D., Pavletich, N.P., 1998. Structural basis for inhibition of the cyclin-dependent kinase Cdk6 by the tumour suppressor p16 (INK4a). *Nature* 395, 237–243.
- Senturk, E., Manfredi, J.J., 2012. Mdm2 and tumorigenesis: Evolving theories and unsolved mysteries. *Genes & Cancer* 3, 192–198.
- Serrano, M., Hannon, G.J., Beach, D., 1993. A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. *Nature* 366, 704–707.
- Serrano, M., Lee, H., Chin, L., *et al.*, 1996. Role of the INK4a locus in tumor suppression and cell mortality. *Cell* 85, 27–37.
- Sharpless, N.E., Bardeesy, N., Lee, K.H., *et al.*, 2001. Loss of p16(Ink4a) with retention of p19(Arf) predisposes mice to tumorigenesis. *Nature* 413, 86–91.
- Sharpless, N.E., DePinho, R.A., 1999. The INK4A/ARF locus and its two gene products. *Current Opinion in Genetics & Development* 9, 22–30.
- Sharpless, N.E., Ramsey, M.R., Balasubramanian, P., Castrillon, D.H., Depinho, R.A., 2004. The differential impact of p16(INK4a) or p19(ARF) deficiency on cell growth and tumorigenesis. *Oncogene* 23, 379–385.
- Sherr, C.J., 2001. The INK4a/ARF network in tumour suppression. *Nature Reviews Molecular Cell Biology* 2, 731–737.
- Sherr, C.J., 2004. Principles of tumor suppression. *Cell* 116, 235–246.
- Sherr, C.J., 2006. Divorcing ARF and p53: an unsettled case. *Nature Reviews Cancer* 6, 663–673.
- Sherr, C.J., 2012. Ink4-Arf locus in cancer and aging. *Wiley Interdisciplinary Reviews: Developmental Biology* 1, 731–741.
- Sherr, C.J., Roberts, J.M., 2004. Living with or without cyclins and cyclin-dependent kinases. *Genes & Development* 18, 2699–2711.
- Sorrentino, J.A., Sanoff, H.K., Sharpless, N.E., 2014. Defining the toxicology of aging. *Trends in Molecular Medicine* 20, 375–384.
- Sotillo, R., Dubus, P., Martin, J., *et al.*, 2001a. Wide spectrum of tumors in knock-in mice carrying a Cdk4 protein insensitive to INK4 inhibitors. *EMBO Journal* 20, 6637–6647.
- Sotillo, R., Garcia, J.F., Ortega, S., *et al.*, 2001b. Invasive melanoma in Cdk4-targeted mice. *Proceedings of the National Academy of Sciences of the United States of America* 98, 13312–13317.
- Stone, S., Jiang, P., Dayananth, P., *et al.*, 1995. Complex structure and regulation of the P16 (MTS1) locus. *Cancer Research* 55, 2988–2994.
- Tchougounova, E., Kastemar, M., Brasater, D., *et al.*, 2007. Loss of Arf causes tumor progression of PDGFB-induced oligodendroglioma. *Oncogene* 26, 6289–6296.
- Tompkins, V.S., Hagen, J., Frazier, A.A., *et al.*, 2007. A novel nuclear interactor of ARF and MDM2 (NIAM) that maintains chromosomal stability. *Journal of Biological Chemistry* 282, 1322–1333.
- Turan, S., Bastepe, M., 2013. The GNAS complex locus and human diseases associated with loss-of-function mutations or epimutations within this imprinted gene. *Hormone Research in Paediatrics* 80, 229–241.
- Ulanet, D.B., Hanahan, D., 2010. Loss of p19^{Arf} facilitates the angiogenic switch and tumor initiation in a multi-stage cancer model via p53-dependent and independent mechanisms. *PLoS ONE* 5, e12454.
- Vousden, K.H., Prives, C., 2009. Blinded by the light: the growing complexity of p53. *Cell* 137, 413–431.
- Wang, T.C., Cardiff, R.D., Zukerberg, L., *et al.*, 1994. Mammary hyperplasia and carcinoma in MMTV-cyclin D1 transgenic mice. *Nature* 369, 669–671.
- Wang, Z., Zheng, Y., Park, H.J., *et al.*, 2013. Targeting FoxM1 effectively retards p53-null lymphoma and sarcoma. *Molecular Cancer Therapeutics* 12, 759–767.
- Weber, J.D., Jeffers, J.R., Reh, J.E., *et al.*, 2000. p53-Independent functions of the p19ARF tumor suppressor. *Genes & Development* 14, 2358–2365.
- Xue, W., Zender, L., Miething, C., *et al.*, 2007. Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature* 445, 656–660.
- Yap, K.L., Li, S., Munoz-Cabello, A.M., *et al.*, 2010. Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Molecular Cell* 38, 662–674.
- Yoshida, H., Matsui, T., Yamamoto, A., Okada, T., Mori, K., 2001. XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell* 107, 881–891.
- Zeng, Y., Kotake, Y., Pei, X.-H., Smith, M.D., Xiong, Y., 2011. p53 Binds to and is required for the repression of Arf tumor suppressor by HDAC and polycomb. *Cancer Research* 71, 2781–2792.
- Zindy, F., Den Besten, W., Chen, B., *et al.*, 2001. Control of spermatogenesis in mice by the cyclin D-dependent kinase inhibitors p18(Ink4c) and p19(Ink4d). *Molecular and Cellular Biology* 21, 3244–3255.
- Zindy, F., Quelle, D.E., Roussel, M.F., Sherr, C.J., 1997. Expression of the p16 (INK4a) tumor suppressor versus other INK4 family members during mouse development and aging. *Oncogene* 15, 203–211.

Relevant Websites

<http://www.cancer.gov/clinicaltrials/search/resultsprotocolsearchid=6184393>

List of NCI-sponsored clinical trials involving the use of selective Cdk4 and Cdk6 inhibitors, such as palbociclib (PD-0332991), to treat various malignancies.

<http://www.cbioportal.org/>

The cBioPortal for Cancer Genomics.

<http://cancergenome.nih.gov>

Website for the Cancer Genome Atlas (TCGA).

S Phase

SJ Smith, RG Lingeman, CM Li, L Gu, RJ Hickey, and LH Malkas, Beckman Research Institute at City of Hope, Duarte, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The S (synthesis) phase occurs between the G1 and G2 phases of the cell cycle. It is during this phase that the genomic DNA within the nucleus is replicated. The steps involved in faithfully executing the duplication of the genomic information of the cell include: signaling the G1 to S phase transition, replicating both DNA strands, and repairing DNA damage prior to the G2 phase.

S Phase Activation

Cyclin-Dependent Kinases

The transition from G1 to S phase is controlled by the interaction of cyclin-dependent kinases (CDKs). Prior to the initiation of cell division, CDK4 and CDK6 are inhibited by p16 (a CDK inhibitor). Upregulation of cyclin D expression displaces p16 binding to CDKs, thereby activating the CDKs. Activated CDK4 and CDK6 are then able to phosphorylate and inactivate retinoblastoma (Rb) protein, subsequently activating E2F transcription factors which ultimately initiate DNA replication. Progress has been made in understanding how CDK-phosphorylated substrates can elicit origin firing. Through the study of budding yeast, it has been demonstrated that replication origin firing requires two essential CDK targets, Sld2 and Sld3. Sld3 must be phosphorylated to bind DpbII (TopBP1 orthologue), while phosphorylated Sld2 promotes DpbII binding. This trimeric complex acts as a docking platform for other proteins required in the assembly and firing of replication origins (Tanaka *et al.*, 2007; Zegerman and Diffley, 2007).

In contrast, the CDK inhibitor, p27, inhibits CDK2 and CDK4 activities and arrests the cell cycle prior to entry into the S phase. The involvement of a network of cellular components enables cells to respond to various intracellular and extracellular conditions and integrate these input signals to determine the initiation of DNA replication.

Myc Control

Myc, a transcription factor encoded by the MYC proto-oncogene, is a key regulator of G1–S progression and plays a direct role in the control of DNA replication. Myc is broadly expressed during embryogenesis, but it is also expressed in highly proliferative adult tissues (epidermis and gut) (Spencer and Groudine, 1991). Enforced expression of Myc is sufficient to stimulate quiescent cells to proliferate (Eilers *et al.*, 1991), and dysregulation of Myc expression has been found in about 50% of all solid tumors (Nesbit *et al.*, 1999). Therefore, physiological and pathological expression of Myc closely correlates to the proliferative status of cells. Structurally, Myc belongs to the helix–loop–helix/leucine zipper protein family.

It binds to the DNA consensus sequences of the E-box (enhancer box) in complex with Max proteins and activates transcription by recruiting transcriptional coactivators to its target genes (Blackwood and Eisenman, 1991). Myc target genes include CCND2, which encodes cyclin D2, CDK4, CUL1, and CKS (Pelengaris *et al.*, 2002). Together, the proteins encoded by these genes are responsible for the sequestration of CDK inhibitor p27 in cyclin-D2–CDK4 complexes and the subsequent degradation of p27 (O'Hagan *et al.*, 2000), thereby allowing activation of the cyclin-E–CDK2 complex. In addition to activating genes involved in cell cycle progression, the MYC–MAX heterodimers can repress the expression of CDK inhibitors, such as INK4B and WAF1/p21, by interacting with transcription factors MIZ1 and/or SP1 (Brenner *et al.*, 2005; Gartel *et al.*, 2001). By activating or repressing its target genes, Myc plays a critical role in activating CDKs and promoting G1–S progression.

DNA Replication

During the S phase of the cell cycle, DNA must be replicated error-free to ensure the correct genomic information is transferred to the newly created cells. In order to successfully complete this process, specific proteins and enzymes must coordinately activate. As cells increase in complexity (from bacteria to mammals), the process of DNA replication becomes more elaborate. However, the basic mechanism for each cell type is similar. Numerous biochemical and genetic studies have revealed the basic mechanisms involved in this semi-conserved process. The replication of double-stranded DNA (dsDNA) involves the coordination of a helicase for DNA unwinding, the protection of the single-strand with DNA-binding proteins, and synthesis involving clamp loaders, polymerases, a DNA ligase, and DNA topoisomerases.

Replication Fork Formation

Upon the initiation of DNA replication, the helical dsDNA must be unwound and separated into two separate parental strands that act as a template for replication. This step is accomplished with the help of a ring-shaped structure that contains a nucleotide-binding (AAA+) domain, often referred to as a helicase. The most commonly characterized of these is DnaB from *Escherichia coli*. This homohexamer possesses the ability to self-load onto single-strand DNA (ssDNA), and is large enough to fit dsDNA through to unwind it. The helicases of archaea and eukaryotes, however, are composed of minichromosome maintenance (MCM) proteins. While most archaeal helicases consist of one MCM protein, most eukaryotes have helicases consisting of six distinct MCM proteins, Mcm2–7.

What is unique to these higher order cells is the MCM helicases cannot unwind the DNA strands alone. Once the

origin replication complex (ORC) is activated, Cdc6 and Cdt1 are signaled, which in turn load Mcm2-7, followed by helicase activation (Cdc45–Mcm2-7–GINS (CMG) complex) (Tanaka and Araki, 2010). Mammalian helicases exhibit features structurally distinct from bacterial helicases: the N-terminal domains of Mcm2-7 are structurally unrelated to DnaB, and the C-terminal motor domain of Mcm is an AAA+ protein, which is evolutionarily and organizationally distinct from DnaB with its RecA-type motors (Costa *et al.*, 2013). Single-particle electron microscopy (EM) reconstructions of *Drosophila melanogaster* reveal that the Mcm2-7 relies on the weak Mcm2–Mcm5 interaction to spontaneously open up the hexamer (Costa *et al.*, 2011). For each ORC–Cdc6 interaction, Mcm2-7 is loaded as a double hexamer in a head-to-head configuration on the dsDNA (Costa *et al.*, 2013). Eukaryotes will use an AAA+ -type initiator similar to that observed in bacteria in order to catalyze the DNA loading. GINS and Cdc45 form a complex with Mcm2-7, significantly increasing adenosine triphosphatase (ATPase) activity, allowing for 3'→5' DNA unwinding (Costa *et al.*, 2011). The CMG complex is thought to directly promote duplex opening and strand separation by splitting the dsDNA between the pore of the helicase and the pore formed by GINS–Cdc45–Mcm235 (Costa *et al.*, 2011).

The Replication Fork

The separation of dsDNA results in the formation of a Y-shaped junction, which is more commonly referred to as the replication fork. Single-strand binding (SSB) proteins protect the exposed ssDNA from premature annealing and nuclease digestion. Most SSB proteins contain at least one conserved DNA-binding oligonucleotide/oligosaccharide binding (OB) fold. These folds can range from 70 to 150 amino acids in length, forming β -barrels capped with an α -helix (Marceau, 2012). Most eukaryotic cells use replication protein A (RPA) as the SSB protein. RPAs form heterotrimers (RPA70, RPA34, RPA14), while most bacterial SSB proteins form homotetramers. Archaeal SSBs resemble either bacterial SSBs or eukaryotic RPAs.

DNA is replicated in a manner that results in a leading strand synthesized continuously and a lagging strand synthesized discontinuously. The leading-strand DNA is processed by a polymerase in the same direction of helicase unwinding, while the lagging strand is synthesized in the opposite direction. The most thoroughly characterized model, *E. coli*, utilizes the DNA polymerase III holoenzyme and other essential proteins, including the DnaB helicase and DnaG primase (Baranska *et al.*, 2013). The most commonly used model for eukaryotic systems is *Schizosaccharomyces pombe* (Aves *et al.*, 2012).

Leading-strand synthesis

At the leading strand, SSB proteins help to load and stabilize the DNA pol α -primase complex (Dornreiter *et al.*, 1992). The pol α -primase synthesizes an RNA primer using complementary base pairing to associate with ssDNA. Following primer synthesis, replication factor C (RFC) competitively binds RPA to disrupt the pol α interaction and load proliferating cell

nuclear antigen (PCNA) (Waga and Stillman, 1998). PCNA is a 29 kDa protein that forms a homotrimer to encircle the DNA strand, thus acting as a loading scaffold for other proteins and enzymes within the replication process, such as DNA polymerase δ and ϵ (Burgers, 1988). PCNA/RPA function to tether the replicative polymerase pol δ to the DNA, thereby acting as a 'polymerase switch' to replace the primase with polymerase, which further processes the new strand of DNA (Forsburg, 2008; Trujillo and Osley, 2008). With the polymerase tethered to the DNA, the leading strand is synthesized in long, continuous segments in the 5'→3' direction.

Lagging-strand synthesis

Owing to the antiparallel nature of DNA replication, the lagging strand is synthesized in a much different manner than the leading strand. The lagging strand is synthesized discontinuously in short segments growing away from the replication fork. These segments are called Okazaki fragments (Rossi *et al.*, 2006). Pol α -primase synthesizes an RNA primer at the 5' end of each Okazaki fragment. After initial primer synthesis, polymerase switching occurs, whereby pol α is removed and pol δ binds (Rossi *et al.*, 2006). Pol δ then adds nucleotides to the 5' end to extend the Okazaki fragment (Alberts, 2003). Upon the convergence of two fragments, the RNA primer is displaced and a single-stranded flap is formed (Aguilera and Gomez-Gonzalez, 2008; Rossi *et al.*, 2006). This flap is then cleaved by flap endonuclease 1 (FEN1) and RNase H, resulting in a nick, which is subsequently sealed by DNA ligase I (Aguilera and Gomez-Gonzalez, 2008; Rossi *et al.*, 2006).

Strand coordination

DNA is replicated in a manner that results in a leading strand synthesized continuously and a lagging strand synthesized discontinuously. In order for this to be a simultaneous process, it is thought that two different DNA polymerases are coupled, as confirmed in both phage and bacterial systems (Johnson and O'Donnell, 2005). These polymerases coordinate with other associated factors to complete this task in both prokaryotic (McHenry, 1988) and eukaryotic (Tsurimoto and Stillman, 1989) replication. This coupling, however, creates a topological problem, where both DNA polymerases generate helical DNA strands that must either travel a helical path or turn behind the complex. It is most likely that the coordination of the leading and lagging-strand synthesis involves a DNA looping mechanism (Ashton *et al.*, 2013). This mechanism, referred to as the trombone model, requires the lagging strand to fold back and form a loop through the complex, allowing parallel polymerization to proceed together in the 5' to 3' direction, thus preventing replication fork stalling. The trombone model has been observed and verified by EM (Alberts *et al.*, 1983). Figure 1(b) depicts the trombone model.

Replication Factories

Because DNA replication is highly coordinated, it is thought that the process is completed by replication complexes that have all the required components. It is still under debate when these complexes are formed. Evidence from biochemical data indicates that multiprotein replication complexes containing

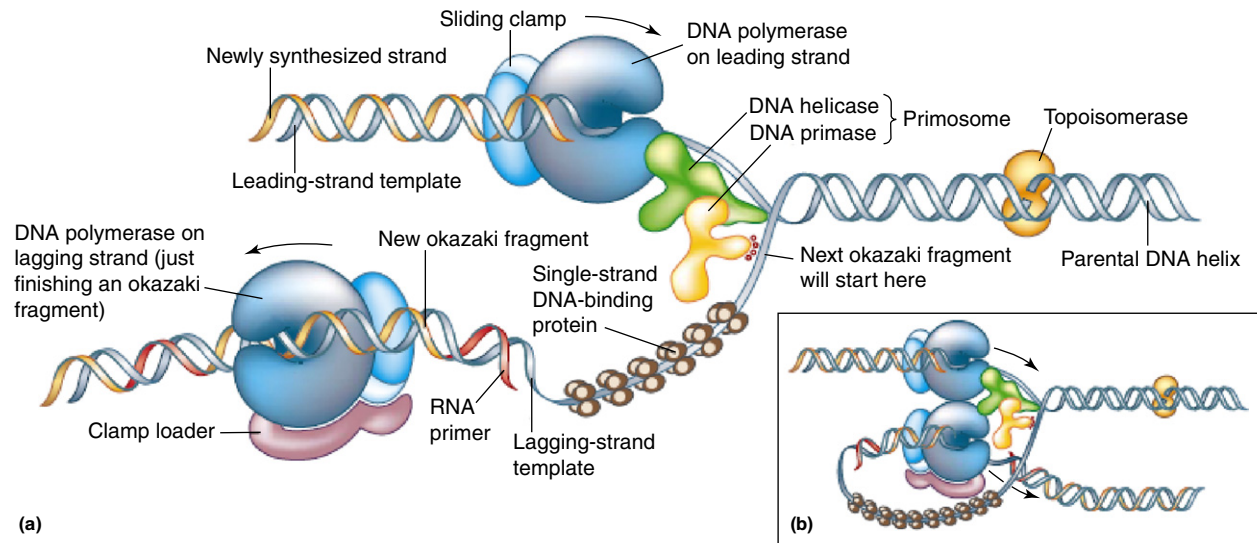


Figure 1 (a) The coordinated activity of proteins and enzymes at the replication fork to process both leading and lagging strand daughter DNA strands. After the parental DNA is unwound by the primosome (DNA helicase and DNA primase), two DNA polymerases synthesize the leading and lagging strand. The lagging strand requires synthesis between separate Okazaki fragments. (b) DNA synthesis with coupled polymerases, resulting in lagging strand looping. Reprinted from Alberts, B., 2003. DNA replication and recombination. Nature 421, 431–435, with permission from Macmillan Publishers Ltd, © 2003.

all required elements form prior to initiation of DNA replication (Tom *et al.*, 1996), while images from live-cell microscopy suggest the interaction of each component within the complex to be highly dynamic (Sporbert *et al.*, 2002).

Replication factories appear to be fixed on a diffuse nucleoskeleton, and the newly synthesized DNA appears to move through replication factories. Images obtained with EM confirm this theory, as the DNA helicase remains intact throughout the replication process, rather than splitting between the two newly synthesized daughter DNA strands (Wessel *et al.*, 1992). Time-lapse fluorescence microscopy in live cells further validates the anchoring of replication factories in the nucleus, as the labeled foci do not appear to change position (Leonhardt *et al.*, 2000).

The genetic information of the eukaryotic and mammalian cell is contained within the chromatin structure of the nucleus. In order to fully replicate the encoded DNA, each cell must contain enough origins to fire once during the process, as re-firing can result in unbalanced genes and fragile replication forks. These origins are grouped into ‘replicons’ that simultaneously replicate large sections of the chromosome. Replicons further cluster together into foci at distinct regions of the nuclear matrix or nucleoskeleton. Each focus contains the proteins and enzymes necessary to replicate DNA, which are themselves organized as ‘replication factories.’ These complexes have been identified microscopically as electron-dense bodies in HeLa cells (Hozak *et al.*, 1993), with replication factories approximately 150 nm in diameter.

DNA Damage Response

During the S phase of every cell cycle, each cell must accurately duplicate DNA to ensure genomic integrity. At the same time, these cells are constantly bombarded by insults that threaten

to compromise the DNA template and, consequently, replication. DNA damage is a continuous phenomenon that is the by-product of these endogenous and exogenous processes and exposures. Exogenously, chemicals or radiation from ionizing (IR) or ultraviolet (UV) sources can damage DNA, while the endogenous damage results from oxidative stress induced by the reactive oxygen species (ROS) of cellular metabolites. Approximately 10^{-10} mistakes are generated per base pair in each DNA replication cycle (Polycove and Feinendegen, 2003). DNA damage leads to stalled replication forks, and failure to repair DNA lesions can result in mutations or large-scale genomic changes that may endanger cellular, and perhaps organism, viability. Consequently, cells have an array of methods to cope with DNA damage to avert fork stalling or facilitate fork restart when stalling occurs.

DNA damage response is a cascade of events involving sensors, transducers, and effectors (Marechal and Zou, 2013). Once a disruption in the DNA structure is identified, signal transduction pathways trigger a response, which often involve slowing down the cell cycle to allow for proper DNA repair. The most common responses involve activation of the ATM (ataxia telangiectasia mutated) and ATR (ATM and Rad3 regulated) protein kinases. Double-stranded breaks in DNA from IR and ROS damage often activate an ATM-related response, while DNA lesions and other types of DNA damage that create ssDNA trigger the ATR pathway. The viability of mammalian cells generally requires ATR, but not ATM. However, they are both important for genomic stability (Marechal and Zou, 2013; Smith *et al.*, 2010).

Damage Signaling

ATM-related damage response

When a double-stranded break on the DNA is sensed, the MRN complex (Mre11, Rad50, and Nbs1) binds and recruits

ATM to the damage site (Figure 2). ATM switches from an inactive homodimer to a particularly active monomer via autophosphorylation to initiate binding near the DNA damage site (Smith *et al.*, 2010). Once bound ATM is able to phosphorylate histone H2AX at Ser139, referred to as γ H2AX. MDC1 binds to γ H2AX and Nbs1 to indirectly associate with ATM, and possibly allow phosphorylation of additional H2AX proteins at nearby nucleosomes along the chromatin structure (Marechal and Zou, 2013).

Activated ATM can signal many other proteins involved in both DNA repair and other cellular checkpoint processes, including Chk2, p53, SMC1, BRCA1, NBS1, CtIP, MDC1, and p53BP1 (Smith *et al.*, 2010). For example, after ATM phosphorylates p53 at Ser15, the kinase inhibitor p21 inhibits cyclin E/A-Cdk2 leading to G1 to S phase stalling. The G2/M

checkpoint also involves p53 and transactivation of p21 (Bolderson *et al.*, 2009). Both p53 and p21 regulate cell cycle arrest, repair, and apoptosis.

ATR-related damage response

ATR is known to play an essential role in replication fork progression. Figure 3 illustrates an example of an ATR-related response. When DNA polymerases encounter damage, such as a lesion, they will often halt DNA synthesis. RPA will coat the long stretches of ssDNA that are created as the DNA helicase continues to unwind the parental strand. Nucleotide excision repair, mismatch repair (MMR), long-patch base excision repair (BER), postreplication repair, end resection at double-strand breaks, or telomere erosion can also expose ssDNA to RPA. The Rad17–RFC complex senses the ssDNA and dsDNA junction and loads the 9–1–1 complex (Rad9–Rad1–Hus1) to dsDNA. ATR-interacting protein (ATRIP) can bind to RPA, thereby recruiting ATR to damaged DNA. Mediator proteins such as TopBP1, Claspin, Timeless, and Tipin stimulate ATR activity. ATR phosphorylates Chk1, resulting in replication checkpoint and DNA damage responses. ATM can also activate other proteins such as BRCA1, MCM, p53, and parts of the RPA complex (Marechal and Zou, 2013; Smith *et al.*, 2010; Tibbetts *et al.*, 1999).

Effector proteins

Chk1 and Chk2

Effector proteins are important to the regulation of DNA replication and repair in the S phase. Chk1 and Chk2 are two such proteins that have been well-characterized. Chk1 is phosphorylated by ATR at Ser319 and Ser345, while Chk2 is phosphorylated by ATM at Thr68 within a Ser-Gln/Thr-Gln (SQ/TQ) rich domain. Phosphorylated Chk1 and Chk2 are

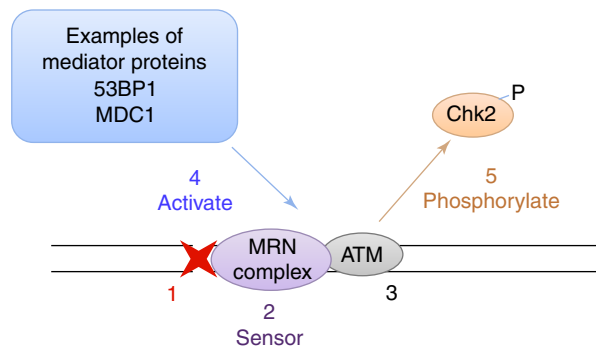


Figure 2 Example of an ATM-related response. (1) In the event of a double strand break (red X), the MRN complex (2) senses the damage and recruits ATM (3). Mediator proteins (4) aid the response for ATM to phosphorylate and activate the effector protein Chk2 (5).

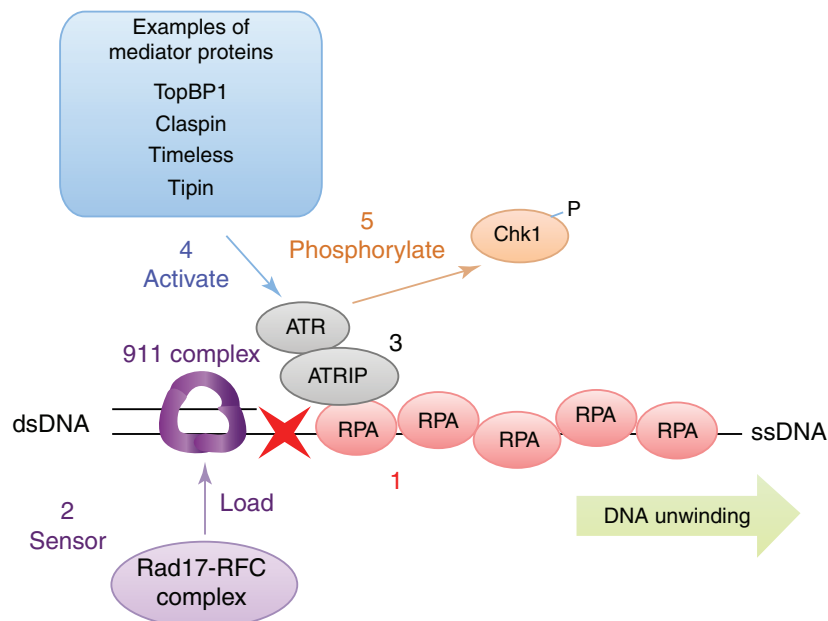


Figure 3 Example of an ATR-related response. (1) In response to DNA damage (red X) from DNA helicase unwinding faster than DNA synthesis, RPA covers stretches of ssDNA. (2) The Rad17–RFC2–5 complex loads the RFC 911 complex (Rad9, Rad1, and Hus1) to dsDNA. (3) ATRIP recruits ATR. (4) With help of mediator proteins, such as TopBP1, Claspin, Timeless, and Tipin, (5) ATR phosphorylates effector proteins, such as Chk1.

thought to detach from DNA and influence proteins involved in cell cycle progression, apoptosis, and gene transcription. For example, they inactivate types of Cdc25 phosphatases that inhibit Cdk1/cyclin B and Cdk2-cyclin E kinases, leading to G₂ to M-phase and G₁ to S-phase arrest, respectively. The resulting cell cycle arrest allows time for DNA repair. Activated Chk1 regulates Wee1 kinase (cell cycle control) and histone H3 phosphorylation (regulation of gene transcription due to DNA damage), and plays a role in recombination through Rad51 and BRCA2 phosphorylation (Smith *et al.*, 2010). Activated Chk2 regulates p53, BRCA1, and transcription factors, including FOXM1 and E2F1.

p53

p53 is a key responder to DNA damage and a major safeguard against the introduction of inheritable mutations into the genome through the replication of damaged DNA (Harris and Levine, 2005). In the absence of DNA damage, p53 is inactivated by its negative regulator, Mdm2. Upon DNA damage or other stresses, various pathways, including ATM and ATR, are activated. Once activated, the ATM and ATR kinases activate their downstream kinases, including Chk1, Chk2, and the MAP kinase p38, which phosphorylate p53; causing it to dissociate from Mdm2. Activated p53 binds DNA and activates expression of several genes, including the AF1/CIP1 gene encoding for p21. p21 then forms a complex with and inhibits CDK2 activity. The ensuing cell cycle arrest gives cells time to repair the damaged DNA. When DNA damage is beyond repair, p53 promotes programmed cell death or apoptosis to prevent accumulation and proliferation of cells harboring potentially oncogenic mutations in their genomes.

The importance of p53 in monitoring DNA damage and regulating DNA repair in normal cells is highlighted by the prevalence of mutations in the gene (TP53) encoding the p53 protein in tumors. Mutations in TP53 are present in about 50% of all human malignancies (Muller and Vousden, 2013). An inherited TP53 mutation has been found to be responsible for the cancer predisposing Li-Fraumeni syndrome (Muller and Vousden, 2013). These individuals develop tumors in their early adulthood. Mutant p53 is unable to bind to DNA, making it an ineffective transcriptional regulator. Therefore, it cannot induce p21 protein expression to stop cell cycle progression in response to DNA damage. Consequently, DNA synthesis continues through the damaged sites, creating inheritable mutations in daughter cells. Accumulation of these mutations in the genome eventually leads to malignancy. Therefore, a functional p53 plays an essential role in maintaining genome integrity and suppressing tumorigenesis.

DNA Repair Mechanisms

Direct reversal

Some modifications to nucleotides can simply be reversed chemically without requiring a template or breaking the phosphodiester backbone of DNA. This type of repair is referred to as direct reversal. An example of this type of repair is the alkylation of guanine that creates an O⁶-methylguanine lesion. O⁶-alkylguanine DNA alkyltransferase (AGT, MGMT, or AGAT) removes the methyl group from O⁶-methylguanine by

transferring the methyl group to a cysteine residue present in MGMT. As a result the lesion reverts back to the correct guanine nucleotide.

BER

The BER pathway is a process that repairs the simplest types of lesions due to DNA damage. BER utilizes a number of specialized glycosylases that can recognize specific nucleotide modifications that require repair. Base modifications recognized by this pathway include nucleotide deaminations, alkylations, oxidations, and spontaneously formed apurinic (AP) sites. Typical deamination reactions that occur during the cell cycle result in the conversion of cytosine to uracil, adenine to hypoxanthine, guanine to xanthine, and 5-methylcytosine to thymine within the DNA structure. Alkylation most commonly results in the addition of methyl groups to the nitrogen residues found in nucleotides. The oxidation of nucleotides is the result of oxygen metabolism, producing superoxide and hydroxyl free radicals, and generating ROS that modify nucleotides. The most common oxidation product is 8-oxo-deoxyguanosine, which can base pair with adenosine and introduce a point mutation if not repaired prior to replication. AP sites result from the hydrolytic cleavage of the glycosidic bond that holds a base in place in the DNA strand. BER is essential for the repair of an estimated loss of ~5000 purines per cell per day.

To remove lesions, the BER pathway utilizes modification-specific glycosylases to scan DNA for errors. Once a lesion is found, the glycosylase 'flips' the base into its damage-specific binding pocket where the N-glycosidic bond is cleaved and the base is removed from the DNA strand, leaving an AP site in its place. A single-strand break is then introduced at the AP site by AP endonuclease, or by the glycosylase itself if it has AP lyase activity. The broken ends are processed by end-processing enzymes such as polynucleotide-kinase-phosphatase (PNKP). DNA repair synthesis then replaces the missing base. There are two modes of DNA repair synthesis: short patch and long patch. In short patch BER, just the single missing base is replaced. The DNA polymerase that performs this function is pol β . In long-patch BER, 2–10 nucleotides are replaced. The first missing base is still replaced by pol β , but then pol δ and pol ϵ in concert with PCNA replace an additional series of nucleotides. In this process, the 5' end of the broken strand is displaced by the new synthesis, and the resulting flap of displaced DNA is removed by FEN1. To complete BER, the remaining nick in the repaired strand is ligated by DNA ligase III and the cofactor XRCC1 in short-patch repair, and by DNA ligase I in long-patch repair (Figure 4).

Nuclear excision repair

The nuclear excision repair (NER) pathway finds and removes bulky modifications that may be present in DNA and interfere with replication (de Laat *et al.*, 1999; Dijk *et al.*, 2014; Le May *et al.*, 2010). These modifications include the addition of methylation groups, oxygen groups, heterocyclic amines found in cooked food, and benzo(a)pyrene-diolepoxide (BPDE) adducts. BPDE is a noted carcinogen derived from the metabolic processing of benzo(a)pyrene present in automobile exhaust, cigarette smoke, and burned meat. NER also removes adducts acquired during chemotherapy, such as those created

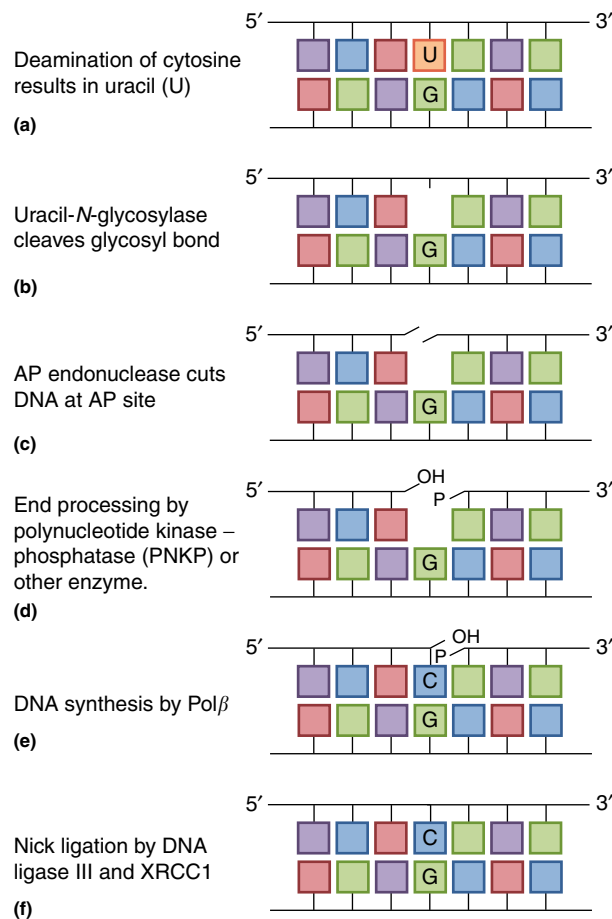


Figure 4 The base excision repair (BER) pathway.

by cisplatin. As such, upregulation of proteins involved in NER are often associated with acquired chemoresistance by patients treated with lesion inducing drugs (Li *et al.*, 1998; Li and Melton, 2012). Perhaps the most essential damage repaired by NER is the damage due to UV exposure. The two major UV photoproducts are cyclobutane pyrimidine dimers (CPD) and pyrimidine (6-4) pyrimidone photoproducts. A deficiency in any one of the XP proteins (XPA through XPG) required for NER results in the condition known as xeroderma pigmentosum (XP), which, like XP 'variant' (XPV) is characterized by extreme sensitivity to sunlight. Patients with XP must avoid sunlight and, when unprotected, have a 10 000-fold higher incidence of non-melanoma skin tumors and a 2000-fold higher incidence of melanoma before the age of 20 than people without the deficiency (Kraemer *et al.*, 1994).

As opposed to the numerous glycosylases employed in BER, NER relies on the XPC–RAD23B protein complex and the accessory DDB1/DDB2 heterodimer (XPE) to constantly scan the genome for specific distortions in the DNA. XPE is important for the identification of some types of distortions that XPC–RAD23B is unable to detect on its own (Dijk *et al.*, 2014). When the complex identifies a distortion, transcription factor II H (TFIIH), XPA, and XPG are recruited to the site. The XPB and XPD subunits of TFIIH act as a helicase to unwind 22–29 nucleotides surrounding the adduct. RPA binds and protects the resulting ssDNA. XPG makes an incision on the 3'

side of the DNA strand containing the modification, while the XPF–ERCC1 complex makes an incision on the 5' side. The damaged single strand is removed, and the 22–29 nucleotide gap is filled in by the PCNA replication complex. The remaining nick is ligated and NER is complete (Figure 5).

MMR

MMR is a repair pathway that addresses the DNA mismatches that occur during replication and other stages of the cell cycle (Hsieh and Yamane, 2008; Iyer *et al.*, 2006; Li, 2008). A common source of mismatches is the misincorporation of nucleotides during replication, which occurs only once in every 10^7 bp. In the event of an incorporation error, MMR repairs the misincorporations and, in doing so, increases the fidelity of replication by 50- to 1000-fold (Iyer *et al.*, 2006). In addition to the normal replicative misincorporation, there are other sources of mismatches repaired through the MMR pathway. These include rare tautomeric forms of nucleotides that can form noncanonical base pairs. For example, the rare imino form of cytosine can base pair with the common amino form of adenosine, and the rare enol form of guanine can mispair with the common keto form of thymine. Other anomalous base pairings can result with ionized nucleotides or with atypical nucleotide conformations (syn-anti and wobble pairing) (Echols and Goodman, 1990; Friedberg *et al.*, 1995). Often mismatches occur in regions of the genome with short 1–5 bp repeats. These mismatches are the result of replication slippage, also known as slipped-strand mispairing, where regions of repeats on the newly synthesized strand are duplicated, or repeats on the template strand are omitted, causing an insertion or deletion (Tautz and Schlotterer, 1994). Other mismatches can result from damage to a nucleotide, including alkylation of nucleotides, adduct formation from carcinogens or chemotherapeutic agents (such as cisplatin) (Hang, 2004; Wozniak and Blasiak, 2002), and the creation of UV photoproducts (Reichrath and Rass, 2014). Finally, genetic recombination can result in mismatches in heteroduplex strands, which are due to sequence differences present in the two strands (Lichten *et al.*, 1990; Reenan and Kolodner, 1992).

The process of MMR in *E. Coli* has been fairly well defined (Figure 6; Iyer *et al.*, 2006; Joseph *et al.*, 2006). The polypeptide MutS initiates MMR by recognizing and binding as a dimer to mismatch regions. This binding signals the recruitment of additional dimeric protein, MutL. The MutS–MutL complex activates several downstream repair processes, including the activation of the latent endonuclease, MutH. Activated MutH incises the unmethylated strand of newly synthesized DNA at hemi-methylated GATC sites. The single-strand break serves to distinguish the mismatch strand from the template, allowing excision repair to begin on the broken strand. The site of the nick on the new strand is not significant, as it can be either 5' or 3' to the site of the mismatch, and does not have to be made by MutH or be within a GATC sequence.

The MutS–MutL complex can also activate DNA helicase II and several single-strand-specific exonucleases. DNA helicase II unwinds the DNA starting at the site of the single-strand break. Unwinding occurs preferentially toward the site of the mismatch, whether the single-strand break is upstream or downstream of the mismatch, suggesting that the MutS–MutL complex plays a role in coordinating the helicase. As the

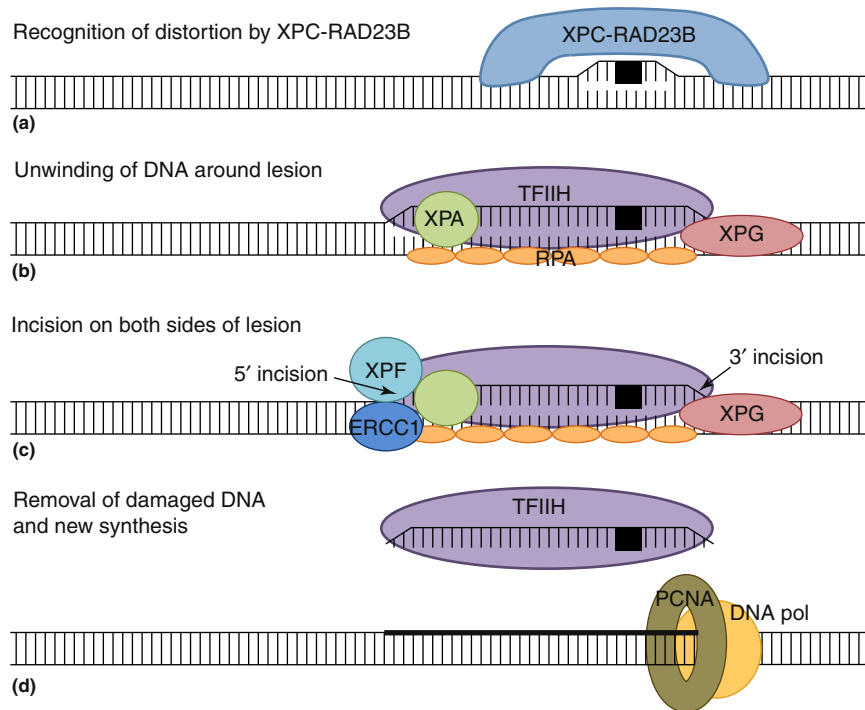


Figure 5 The nuclear excision repair (NER) pathway.

broken strand is unwound, it is digested by several exonucleases. When unwinding occurs 5' of the mismatch, ExoVII and RecJ utilize their 5' → 3' exonuclease activity to hydrolyze the strand. When unwinding starts 3' of the mismatch, the strand is hydrolyzed by the 3' → 5' exonuclease activity of ExoI, ExoVII, or ExoX. After removal of the mismatch, DNA polymerase III holoenzyme synthesizes the missing segment, and DNA ligase seals the remaining nick.

MMR in eukaryotes is not as well characterized as in bacteria (Modrich, 2006). Several homologues for MutS have been found in eukaryotes, and have been designated MSH1–MSH6. These proteins are active as heterodimers. In humans, MSH2 pairs with MSH6 (MutS α) or with MSH3 (MutS β). Homologues for MutL have also been identified. They have been named MLH1, MLH2, MLH3, PMS1, and PMS2. These MutL homologues also form heterodimers. In humans, the MLH1–PMS2 heterodimer (MutL α) is capable of supporting repair initiated by MutS α or MutS β . A MLH1–PMS1 heterodimer (MutL β) has also been identified, but its function is unknown. No homologue for MutH has been identified in eukaryotes, and it is not known how excision repair is targeted to the strand containing the mismatch. It has not been shown that helicase activity is involved in eukaryotic MMR. Additionally, only one exonuclease (ExoI) has thus far been shown to be involved in the eukaryote system. PCNA and at least DNA polymerase δ are involved in the DNA synthesis required to replace the excised sequence.

Translesion DNA synthesis

Translesion DNA synthesis (TLS) represents an important survival mechanism for cells. It is a DNA damage tolerant process rather than a DNA repair process, as it enables the replication fork to bypass the encountered damage. It acts as a

quick fix for the damage by minimizing the length of time that replication fork stalling can threaten cell viability, while creating a double-strand product and a sister chromatid. This product can be used as a substrate or template for subsequent repair by mechanisms such as NER, BER, and homologous recombination (Prakash *et al.*, 2005; Saugar *et al.*, 2014; Waters *et al.*, 2009).

During conventional replication, DNA priming and elongation are conducted by highly accurate replication polymerases (Pol α , Pol δ , Pol ϵ , and Pol γ) that are present in the PCNA-based replication complex. The specificities of these polymerases is due in large part to the concise fit with which the polymerase encloses the template DNA. This exacting fit ensures that the active site of the polymerase reliably inserts the correct nucleotide into the newly synthesized DNA strand. However, in the event a DNA lesion is encountered, the polymerase is unable to accommodate such a bulky modification, which results in stalling at the replication fork. In order to accommodate these lesions, special TLS polymerases (Table 1) are required.

In order to switch between replicative polymerases and TLS polymerases, ubiquitinated PCNA (at lysine 164 by the Rad6–Rad18 protein duplex) is essential (Yang *et al.*, 2013). The TLS polymerases, as listed in Table 1, have a less restrictive conformation with DNA, allowing them to handle the unusual size of the lesion and still insert nucleotides (Prakash *et al.*, 2005; Yang, 2014). Even though the replication complex can continue synthesis beyond the lesion, a mutation has now been introduced into the DNA. The degree of error involved is dependent upon both the type of lesion and the TLS polymerase used in the bypass (Prakash *et al.*, 2005).

Often, more than one TLS polymerase is required to move past a lesion. For instance, in bypassing some thymidine

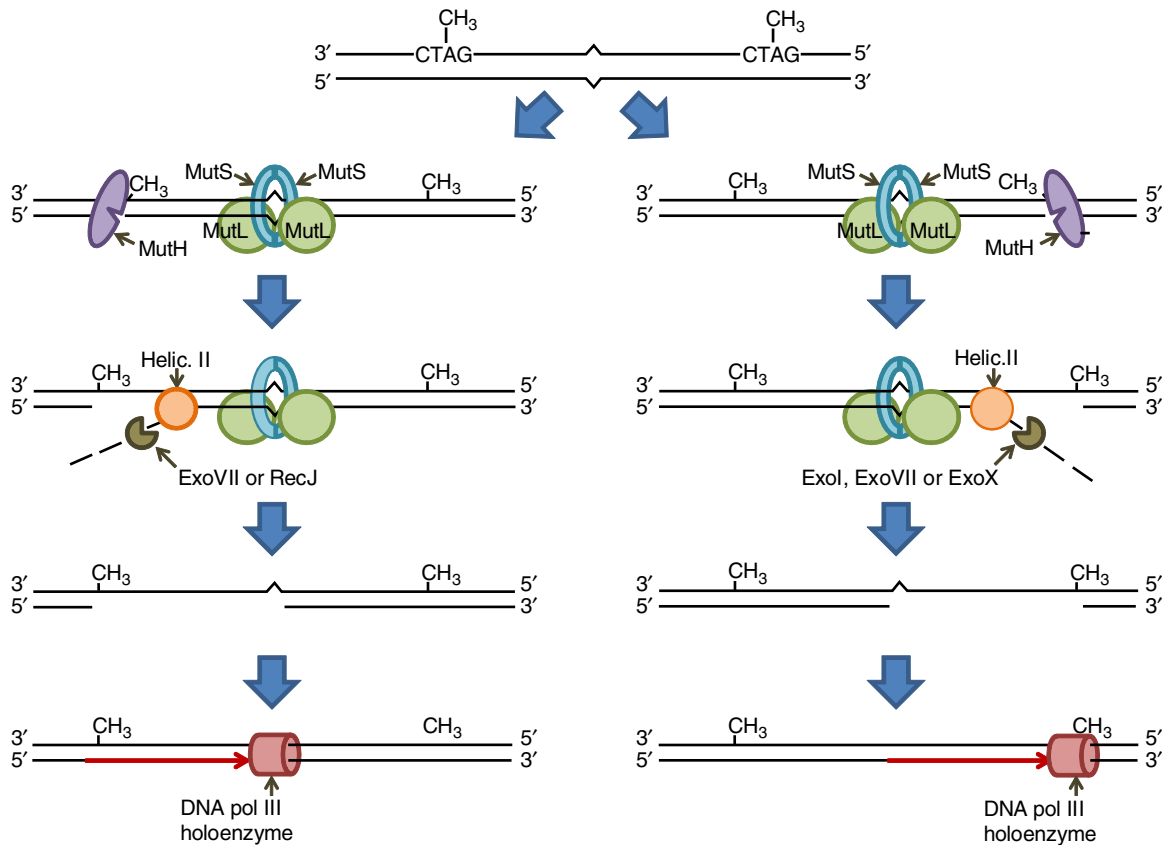


Figure 6 Mismatch repair (MMR) in *E. Coli*. Newly synthesized DNA containing a mismatch is recognized by MutS dimer. MutS recruits MutL to the site of mismatch. The MutS–MutL complex recruits and activates latent MutH. MutH incises the newly synthesized strands at hemi-methylated (5')-GATC-(3') sites in duplex DNA. The incision used in repair can be either 5' or 3' to the site of the mismatch. Helicase II begins unwinding DNA from the site of incision toward the site of mismatch. Single-strand-specific exonucleases digest the unwound strand containing the mismatch. If the incised strand is unwound in the 5' to 3' direction, ExoVII or RecJ are the exonucleases that digest the broken strand. If the incised strand is unwound in the 3' to 5' direction, ExoI, ExoVII, or ExoX are the exonucleases involved in digestion. DNA polymerase III holoenzyme resynthesizes the removed segment of DNA and DNA ligase seals the remaining nick (not shown).

Table 1 Identified polymerases involved in translesion synthesis (TLS)

Family	Gene	Protein
A	POLQ	Pol θ
B	REV3 and REV7	Pol ζ
Y	REV1	Rev1
Y	POLH	Pol η
Y	POLI	Pol ι
Y	POLK	Pol κ
A	POLN	Pol ν

dimers, such as 6-4 photoproducts, the TLS polymerase pol η will accurately position an adenosine base opposite the first thymidine base with the correct Watson–Crick base pairing but will insert the next adenosine in a *syn* orientation that results in Hoogsteen pairing with the second thymidine. In this situation, pol η will subsequently be replaced by a more error-prone polymerase, such as pol γ , to extend the DNA past the lesion, where the more accurate replicative polymerase can then continue normal synthesis (Chou, 2011). When bypass results in the wrong base being inserted in the newly

synthesized strand, subsequent DNA repair of the original lesion may use the incorrect sequence as a template for repair, which results in a point mutation within the genome. Multiple point mutations have been shown to drive normal cells to become cancerous, or cancer cells to acquire chemoresistance (Figure 7).

The importance of TLS is best illustrated in XPV patients. These patients lack one of the TLS polymerases (i.e., Pol η), which enables bypass of CPD typically caused by UV exposure. Even though XPV patients have an intact NER pathway, they suffer from the same condition as patients without NER – that is, they have acute sensitivity to UV exposure, which results in a dramatic increase in mutations and incidence of skin cancer (Kraemer *et al.*, 1994; Menck and Munford, 2014).

Double-strand break repair

The most serious, but less common, form of DNA damage is a double-stranded break. Not repairing these damages leads to apoptotic cell death. In a model involving both the ATM and ATR pathway, the ATM pathway first responds to dsDNA breaks. DNA-end resection, involving ATM, MRN complex, CtIP, and BRCA1, occurs when specialized nucleases create

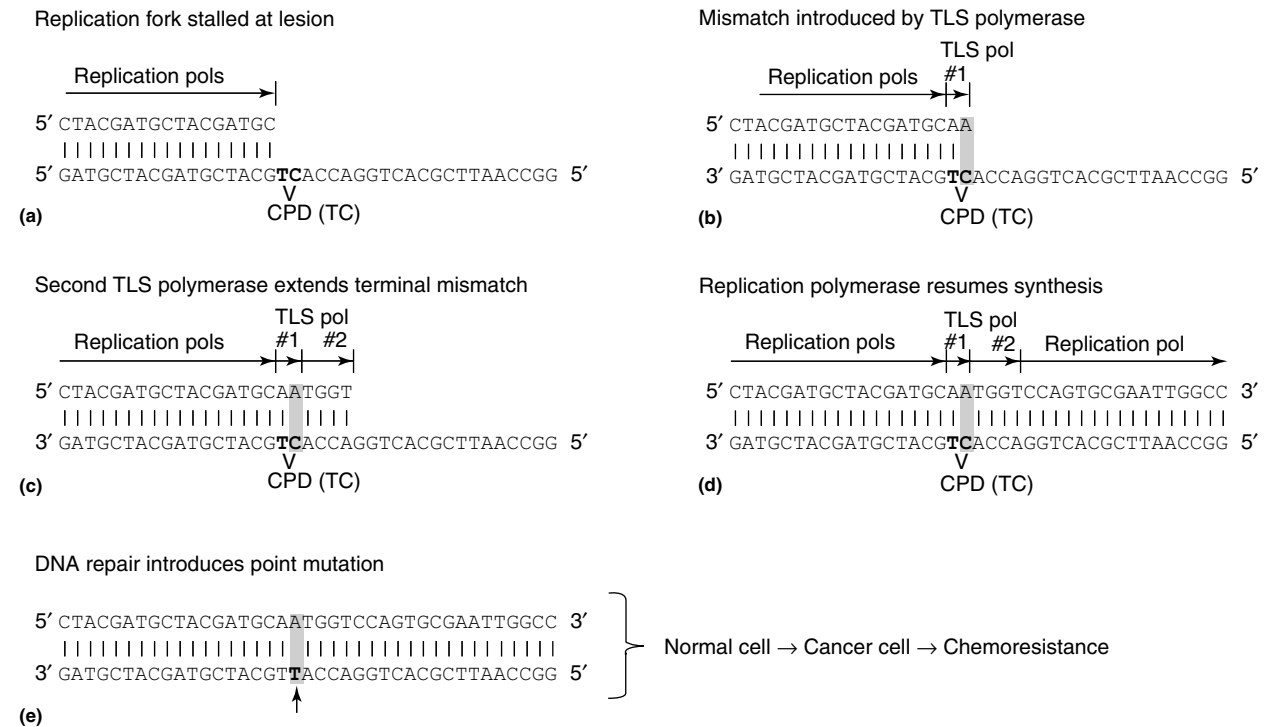


Figure 7 The TLS lesion repair mechanism.

ssDNA. RPA also binds to the newly created ssDNA (Huertas, 2010) activating the ATR–Chk1 pathway.

Double-strand break repair mechanisms

Mechanisms for repairing double-strand breaks mainly involve (1) NHEJ (nonhomologous end-joining) and (2) HR (homologous recombination repair) (Huertas, 2010). NHEJ occurs during any phase of the cell cycle, and was thought to be error-prone because it involves rejoining broken dsDNA regardless of the sequence, but emerging evidence shows otherwise (Betermier *et al.*, 2014). HR, on the other hand, requires a homologous template and occurs during the S and G₂ phase of the cell cycle (Smith *et al.*, 2010).

NHEJ repairs most double-strand breaks caused by ionizing radiation. In the canonical model for NHEJ, the Ku70/80 heterodimer recruits DNA-dependent protein kinase to the two ends of dsDNA in a ring-like structure. This allows the XRCC4-ligase IV complex to ligate the ends of DNA with the aid of the XRCC4-like factor (Betermier *et al.*, 2014; Bolderson *et al.*, 2009). Alternative models of NHEJ have also been described (Betermier *et al.*, 2014).

In a common model of HR, the ends of the ssDNA, such as from sister chromatids, are used as templates for repair. In vertebrates, Cdk appears to phosphorylate CtIP, which controls DNA strand resectioning via cell cycle regulation. Yeast studies reveal that HR occurs mainly in the S and G₂ phase of the cell cycle, and HR in vertebrates is thought to be regulated similarly. DNA strand resection by the Mre11 endonuclease of the MRN complex helps create tracks of ssDNA/RPA, with the aid of BRCA1 and CtIP. More single strands are formed with other nucleases and helicases, such as

Exo1 and BLM, where RPA binds. These events can activate the ATR pathway. After strand resectioning, BRCA2 loads Rad51 recombinase onto the ssDNA, which displaces RPA and starts DNA damage repair using the sister chromatid as a template. BRCA1 responds to DNA damage through its role in DNA repair, checkpoint activation, and gene transcription (Smith *et al.*, 2010).

Conclusions

- The synthesis (S) phase of the cell cycle is of critical importance to precisely replicating the genomic information encoded in the nucleus of the cell.
- The G1 to S phase transition requires the activation and signaling of a number of specific components, including CDKs and Myc.
- The actual process of DNA replication in mammalian cells is complex, requiring the coordinated activity of specific proteins and enzymes. The exact mechanism involved in the simultaneous, dual processing of the DNA strands, however, is currently unknown.
- DNA damage is a constant threat to the genomic integrity of the cell. As a result, the cells have a number of mechanisms to respond to the damage incurred. This will allow the cell cycle to continue.

See also: Functional Cell Biology: An Overview

References

- Aguilera, A., Gomez-Gonzalez, B., 2008. Genome instability: A mechanistic view of its causes and consequences. *Nature Reviews Genetics* 9, 204–217.
- Alberts, B., 2003. DNA replication and recombination. *Nature* 421, 431–435.
- Alberts, B.M., Barry, J., Bedinger, P., *et al.*, 1983. Studies on DNA replication in the bacteriophage T4 *in vitro* system. Cold Spring Harbor Symposia on Quantitative Biology 47 (Pt 2), 655–668.
- Ashton, N.W., Bolderson, E., Cubeddu, L., *et al.*, 2013. Human single-stranded DNA binding proteins are essential for maintaining genomic stability. *BMC Molecular Biology* 14, 9.
- Aves, S.J., Liu, Y., Richards, T.A., 2012. Evolutionary diversification of eukaryotic DNA replication machinery. *Sub-Cellular Biochemistry* 62, 19–35.
- Baranska, S., Glinkowska, M., Herman-Antosiewicz, A., *et al.*, 2013. Replicating DNA by cell factories: Roles of central carbon metabolism and transcription in the control of DNA replication in microbes, and implications for understanding this process in human cells. *Microbial Cell Factories* 12, 55.
- Betermier, M., Bertrand, P., Lopez, B.S., 2014. Is non-homologous end-joining really an inherently error-prone process? *PLoS Genetics* 10, e1004086.
- Blackwood, E.M., Eisenman, R.N., 1991. Max: A helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc. *Science* 251, 1211–1217.
- Bolderson, E., Richard, D.J., Zhou, B.B., Khanna, K.K., 2009. Recent advances in cancer therapy targeting proteins involved in DNA double-strand break repair. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research* 15, 6314–6320.
- Brenner, C., Deplus, R., Didelot, C., *et al.*, 2005. Myc represses transcription through recruitment of DNA methyltransferase corepressor. *EMBO Journal* 24, 336–346.
- Burgers, P.M., 1988. Mammalian cyclin/PCNA (DNA polymerase delta auxiliary protein) stimulates processive DNA synthesis by yeast DNA polymerase III. *Nucleic Acids Research* 16, 6297–6307.
- Chou, K.M., 2011. DNA polymerase eta and chemotherapeutic agents. *Antioxidants & Redox Signaling* 14, 2521–2529.
- Costa, A., Hood, I.V., Berger, J.M., 2013. Mechanisms for initiating cellular DNA replication. *Annual Review of Biochemistry* 82, 25–54.
- Costa, A., Ilves, I., Tamberg, N., *et al.*, 2011. The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nature Structural & Molecular Biology* 18, 471–477.
- Dijk, M., Tymas, D., Mullenders, L., Pines, A., 2014. Insight in the multilevel regulation of NER. *Experimental Cell Research* 329, 116–123.
- Dornreiter, I., Erdile, L.F., Gilbert, I.U., *et al.*, 1992. Interaction of DNA polymerase alpha-primease with cellular replication protein A and SV40 T antigen. *EMBO Journal* 11, 769–776.
- Echols, H., Goodman, M.F., 1990. Mutation induced by DNA damage: A many protein affair. *Mutation Research* 236, 301–311.
- Eilers, M., Schirm, S., Bishop, J.M., 1991. The MYC protein activates transcription of the alpha-prothymosin gene. *EMBO Journal* 10, 133–141.
- Forsburg, S.L., 2008. The MCM helicase: Linking checkpoints to the replication fork. *Biochemical Society Transactions* 36, 114–119.
- Friedberg, E.C., Walker, G.C., Siede, W., 1995. *DNA Repair and Mutagenesis*. Washington, DC: ASM Press.
- Gartel, A.L., Ye, X., Goufman, E., *et al.*, 2001. Myc represses the p21(WAF1/CIP1) promoter and interacts with Sp1/Sp3. *Proceedings of the National Academy of Sciences of the United States of America* 98, 4510–4515.
- Hang, B., 2004. Repair of exocyclic DNA adducts: Rings of complexity. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology* 26, 1195–1208.
- Harris, S.L., Levine, A.J., 2005. The p53 pathway: Positive and negative feedback loops. *Oncogene* 24, 2899–2908.
- Hozak, P., Hassan, A.B., Jackson, D.A., Cook, P.R., 1993. Visualization of replication factories attached to nucleoskeleton. *Cell* 73, 361–373.
- Hsieh, P., Yamane, K., 2008. DNA mismatch repair: Molecular mechanism, cancer, and ageing. *Mechanisms of Ageing and Development* 129, 391–407.
- Huertas, P., 2010. DNA resection in eukaryotes: Deciding how to fix the break. *Nature Structural & Molecular Biology* 17, 11–16.
- Iyer, R.R., Pluciennik, A., Burdett, V., Modrich, P.L., 2006. DNA mismatch repair: Functions and mechanisms. *Chemical Reviews* 106, 302–323.
- Johnson, A., O'Donnell, M., 2005. Cellular DNA replicases: Components and dynamics at the replication fork. *Annual Review of Biochemistry* 74, 283–315.
- Joseph, N., Duppatla, V., Rao, D.N., 2006. Prokaryotic DNA mismatch repair. *Progress in Nucleic Acid Research and Molecular Biology* 81, 1–49.
- Kraemer, K.H., Lee, M.M., Andrews, A.D., Lambert, W.C., 1994. The role of sunlight and DNA repair in melanoma and nonmelanoma skin cancer. The xeroderma pigmentosum paradigm. *Archives of Dermatology* 130, 1018–1021.
- de Laat, W.L., Jaspers, N.G., Hoeijmakers, J.H., 1999. Molecular mechanism of nucleotide excision repair. *Genes & Development* 13, 768–785.
- Le May, N., Egly, J.M., Coin, F., 2010. True lies: The double life of the nucleotide excision repair factors in transcription and DNA repair. *Journal of Nucleic Acids* 2010, 616342.
- Leonhardt, H., Rahn, H.P., Weinzierl, P., *et al.*, 2000. Dynamics of DNA replication factories in living cells. *Journal of Cell Biology* 149, 271–280.
- Li, G.M., 2008. Mechanisms and functions of DNA mismatch repair. *Cell Research* 18, 85–98.
- Li, Q., Gardner, K., Zhang, L., *et al.*, 1998. Cisplatin induction of ERCC-1 mRNA expression in A2780/CP70 human ovarian cancer cells. *Journal of Biological Chemistry* 273, 23419–23425.
- Li, W., Melton, D.W., 2012. Cisplatin regulates the MAPK kinase pathway to induce increased expression of DNA repair gene ERCC1 and increase melanoma chemoresistance. *Oncogene* 31, 2412–2422.
- Lichten, M., Goyon, C., Schultes, N.P., *et al.*, 1990. Detection of heteroduplex DNA molecules among the products of *Saccharomyces cerevisiae* meiosis. *Proceedings of the National Academy of Sciences of the United States of America* 87, 7653–7657.
- Marceau, A.H., 2012. Functions of single-strand DNA-binding proteins in DNA replication, recombination, and repair. *Methods in Molecular Biology* 922, 1–21.
- Marechal, A., Zou, L., 2013. DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harbor Perspectives in Biology* 5, a012716.
- McHenry, C.S., 1988. DNA polymerase III holoenzyme of *Escherichia coli*. *Annual Review of Biochemistry* 57, 519–550.
- Menck, C.F., Munford, V., 2014. DNA repair diseases: What do they tell us about cancer and aging? *Genetics and Molecular Biology* 37, 220–233.
- Modrich, P., 2006. Mechanisms in eukaryotic mismatch repair. *Journal of Biological Chemistry* 281, 30305–30309.
- Muller, P.A., Vousden, K.H., 2013. p53 mutations in cancer. *Nature Cell Biology* 15, 2–8.
- Nesbit, C.E., Tersak, J.M., Prochownik, E.V., 1999. MYC oncogenes and human neoplastic disease. *Oncogene* 18, 3004–3016.
- O'Hagan, R.C., Ohh, M., David, G., *et al.*, 2000. Myc-enhanced expression of Cul1 promotes ubiquitin-dependent proteolysis and cell cycle progression. *Genes and Development* 14, 2185–2191.
- Pelengaris, S., Khan, M., Evan, G., 2002. c-MYC: More than just a matter of life and death. *Nature Reviews Cancer* 2, 764–776.
- Pollycove, M., Feinendegen, L.E., 2003. Radiation-induced versus endogenous DNA damage: Possible effect of inducible protective responses in mitigating endogenous damage. *Human and Experimental Toxicology* 22, 290–306. **Discussion 307, 315–297, 319–223.**
- Prakash, S., Johnson, R.E., Prakash, L., 2005. Eukaryotic translesion synthesis DNA polymerases: Specificity of structure and function. *Annual Review of Biochemistry* 74, 317–353.
- Reenan, R.A., Kolodner, R.D., 1992. Characterization of insertion mutations in the *Saccharomyces cerevisiae* MSH1 and MSH2 genes: Evidence for separate mitochondrial and nuclear functions. *Genetics* 132, 975–985.
- Reichrath, J., Rass, K., 2014. Ultraviolet damage, DNA repair and vitamin D in nonmelanoma skin cancer and in malignant melanoma: An update. *Advances in Experimental Medicine and Biology* 810, 208–233.
- Rossi, M.L., Purohit, V., Brandt, P.D., Bambara, R.A., 2006. Lagging strand replication proteins in genome stability and DNA repair. *Chemical Reviews* 106, 453–473.
- Saugar, I., Ortiz-Bazan, M.A., Tercero, J.A., 2014. Tolerating DNA damage during eukaryotic chromosome replication. *Experimental Cell Research* 329, 170–177.
- Smith, J., Tho, L.M., Xu, N., Gillespie, D.A., 2010. The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. *Advances in Cancer Research* 108, 73–112.
- Spencer, C.A., Groudine, M., 1991. Control of c-myc regulation in normal and neoplastic cells. *Advances in Cancer Research* 56, 1–48.
- Sporbert, A., Gahl, A., Ankerhold, R., *et al.*, 2002. DNA polymerase clamp shows little turnover at established replication sites but sequential de novo assembly at adjacent origin clusters. *Molecular Cell* 10, 1355–1365.
- Tanaka, S., Araki, H., 2010. Regulation of the initiation step of DNA replication by cyclin-dependent kinases. *Chromosoma* 119, 565–574.
- Tanaka, S., Umemori, T., Hirai, K., *et al.*, 2007. CDK-dependent phosphorylation of Sld2 and Sld3 initiates DNA replication in budding yeast. *Nature* 445, 328–332.
- Tautz, D., Schlotterer, C., 1994. Simple sequences. *Current Opinion in Genetics and Development* 4, 832–837.

- Tibbetts, R.S., Brumbaugh, K.M., Williams, J.M., *et al.*, 1999. A role for ATR in the DNA damage-induced phosphorylation of p53. *Genes and Development* 13, 152–157.
- Tom, T.D., Malkas, L.H., Hickey, R.J., 1996. Identification of multiprotein complexes containing DNA replication factors by native immunoblotting of HeLa cell protein preparations with T-antigen-dependent SV40 DNA replication activity. *Journal of Cellular Biochemistry* 63, 259–267.
- Trujillo, K.M., Osley, M.A., 2008. INO80 meets a fork in the road. *Nature Structural and Molecular Biology* 15, 332–334.
- Tsurimoto, T., Stillman, B., 1989. Multiple replication factors augment DNA synthesis by the two eukaryotic DNA polymerases, alpha and delta. *EMBO Journal* 8, 3883–3889.
- Waga, S., Stillman, B., 1998. The DNA replication fork in eukaryotic cells. *Annual Review of Biochemistry* 67, 721–751.
- Waters, L.S., Minesinger, B.K., Wiltout, M.E., *et al.*, 2009. Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. *Microbiology and Molecular Biology Reviews* 73, 134–154.
- Wessel, R., Schweizer, J., Stahl, H., 1992. Simian virus 40T-antigen DNA helicase is a hexamer which forms a binary complex during bidirectional unwinding from the viral origin of DNA replication. *Journal of Virology* 66, 804–815.
- Wozniak, K., Blasiak, J., 2002. Recognition and repair of DNA-cisplatin adducts. *Acta Biochimica Polonica* 49, 583–596.
- Yang, K., Weinacht, C.P., Zhuang, Z., 2013. Regulatory role of ubiquitin in eukaryotic DNA translesion synthesis. *Biochemistry* 52, 3217–3228.
- Yang, W., 2014. An overview of Y-Family DNA polymerases and a case study of human DNA polymerase eta. *Biochemistry* 53, 2793–2803.
- Zegerman, P., Diffley, J.F., 2007. Phosphorylation of Sld2 and Sld3 by cyclin-dependent kinases promotes DNA replication in budding yeast. *Nature* 445, 281–285.

Traveling through Mitosis with the Chromosomal Passenger Complex

ME Bekier II, University of Toledo, Toledo, OH, USA and University of Michigan, Ann Arbor, MI, USA
WR Taylor, University of Toledo, Toledo, OH, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The chromosomal passenger complex (CPC) consisting of Aurora B, Borealin, INCENP, and Survivin is essential for mitosis and meiosis (Carmena *et al.*, 2012). Activation of Aurora B occurs upon its interaction with the C-terminus of INCENP while the other passenger proteins ensure that Aurora B is properly localized. The CPC shows a dynamic pattern of localization allowing it to regulate multiple mitotic events. During prometaphase and metaphase, the CPC concentrates at the inner centromere where it is essential for error correction, a process that destabilizes inappropriate kinetochore–spindle attachments (Carmena *et al.*, 2012). Aurora B is also essential for phosphorylation of histone H3 at serine 10 which releases HP1 proteins from heterochromatin during mitosis to allow chromosome compaction. Aurora B also contributes to release of cohesin from chromosome arms during prophase, a process that helps to maintain genomic stability. At anaphase the CPC relocates to the spindle midzone where it coordinates cleavage furrow formation. At cytokinesis, CPC at the midbody inhibits abscission in response to chromosome bridges. Evidence exists for functional conservation of the CPC in diverse organisms including fungi and amoebozoans (Chen *et al.*, 2007). A potential homologue of INCENP has also been identified in plants (Kirioukhova *et al.*, 2011). The CPC appears to be a relatively early addition to the eukaryotic lineage and an essential component of animal cell division.

Identification of the Core CPC Members

Discovery of CPC started with INCENP. By raising antibodies against the bulk proteins of the chicken mitotic chromosome scaffold, an antibody recognizing two proteins concentrated at the pericentromere was identified. The proteins were termed inner centromere proteins (INCENPs) and subsequently found to arise from alternative splicing of the incenp mRNA (Cooke *et al.*, 1987). INCENPs displayed dynamic localization during mitosis, localizing to inner centromeres before chromosome segregation and to the spindle midzone and midbody during cytokinesis and telophase, respectively. This pattern of localization led to the ‘chromosomal passenger’ hypothesis which stated that various mitotic events, such as chromosome segregation and cytokinesis, were regulated by proteins that localize to chromosomes during early mitosis and transfer to the spindle midzone during late mitosis (Cooke *et al.*, 1987). Further functional analysis revealed that INCENP played a role in chromosome congression and segregation (Ainsztein *et al.*, 1998). Aurora B was the next CPC family member identified, but not in that capacity. Aurora B was identified as the budding yeast mutant *Ipl1* which underwent whole genome polyploidization at the restrictive temperature (Chan and Botstein, 1993). Studies from another direction identified

Survivin as an antiapoptotic protein overexpressed in cancer. Localization patterns of Aurora B, Survivin, and INCENP were similar and it was not long before all three were found to be in a complex providing a direct link between chromosomal passenger proteins and mitotic checkpoint control (Kim *et al.*, 1999). Shortly thereafter, a proteomic screen for novel human proteins associated with histone-depleted mitotic chromosomes led to the discovery of Borealin, providing a glimpse of the core members of the CPC (Gassmann *et al.*, 2004).

Putting the CPC Together

INCENP, Survivin, and Borealin can be thought of as both a delivery and activation system for Aurora B. Binding of Aurora B to INCENP activates the kinase. Localization of this activity is just as important for CPC function, a role played by all CPC members. Ultimately, it is the catalytic activity of Aurora B that is responsible for the effects of the CPC. Aurora B is a member of the Aurora family of kinases that are related to the cAMP (cyclic adenosine monophosphate)-dependent, cGMP (cyclic guanosine monophosphate)-dependent, protein kinase C family of serine/threonine kinases (Carmena *et al.*, 2012). The family of Aurora kinases comprised of Aurora A, Aurora B, and Aurora C share the consensus motif $([R/K] \times [S/T]\Phi)$ but differ in substrates mainly by specific distribution patterns within the cell. For instance, Aurora A localizes to the mitotic spindle, centrosomes, and midbody while Aurora B localizes to chromosome arms, inner centromeres, the central spindle, and midbody. Localization of Aurora C is similar to Aurora B; however, Aurora C is mainly expressed in germ cells and the morula during embryonic development. Interestingly, the differential function of Aurora A and Aurora B is achieved by a single amino acid difference (Eyers *et al.*, 2005). Basal activation of Aurora B is initially mediated by binding the conserved C-terminal IN-box of INCENP. Aurora B then phosphorylates INCENP at conserved threonine-serine-serine (TSS) motif on INCENP and threonine 232 in the activation loop of Aurora B resulting in full activation of Aurora B (Bishop and Schumacher, 2002). Microtubules can also activate Aurora B and microtubule-mediated activation of Aurora B is stimulated by telophase disk protein 60 (TD-60) (Rosasco-Nitcher *et al.*, 2008).

To understand the critical role of CPC subunits in localizing Aurora B activity, a view of how the subunits are assembled is essential (Figure 1). The C-terminus of Survivin and the N-terminus of Borealin bind the N-terminus of INCENP in a three-helix bundle to facilitate complex formation and mediate targeting of the CPC to various intracellular localizations (Jeyaprakash *et al.*, 2007). While the N-terminus of INCENP binds Borealin and Survivin, the C-terminus binds Aurora B. The N- and C-terminus of INCENP are separated by a large, unstructured region which contains

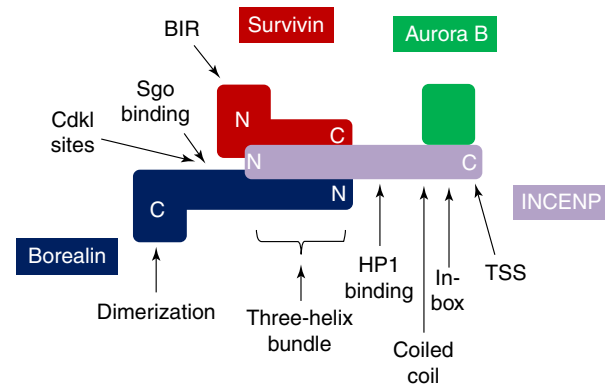


Figure 1 Subunits of the CPC. The CPC is composed of INCENP, Aurora B, Survivin, and Borealin. Complex formation is facilitated by a three-helix bundle between INCENP, Survivin, and Borealin. The central region of INCENP contains an HP-1 binding site and a coiled-coil domain which can interact with microtubules. The C-terminus of INCENP contains a highly conserved IN-box motif that mediates the interaction with Aurora B. The C-terminus of INCENP also contains a conserved TSS motif which is phosphorylated by Aurora B and consequently stimulates Aurora B activity. The N-terminus of Survivin contains a conserved Zn^{2+} -binding BIR domain. The central region of Borealin contains multiple CDK1-mediated phosphorylation sites and the C-terminus of Borealin contains a conserved dimerization domain.

the heterochromatin binding protein-1 (HP-1) binding motif and multiple CDK1 phosphorylation sites (Carmena *et al.*, 2012). Survivin is a well-studied, multifunctional protein that contains an N-terminal baculoviral inhibitor of apoptosis (BIR) domain bound to Zn^{2+} . The C-terminal α -helix of Survivin interacts with INCENP and Borealin (Jeyaprakash *et al.*, 2007). Survivin was originally described as an inhibitor of apoptosis protein (IAP) that was proposed to negatively regulate cell death in mitosis (reviewed in Altieri, 2010). Numerous studies have connected high levels of Survivin with a reduction of apoptosis. Also, Survivin is overexpressed in many types of human cancer. However, knockdown or knockout of Survivin does not initially sensitize to apoptosis, but does ultimately result in apoptosis secondary to failed cytokinesis and accumulation of DNA damage (Yang *et al.*, 2004). Therefore, Survivin, at endogenous levels does not suppress apoptosis. Overexpressed Survivin, either in an experiment setting, or as a result of upregulation in cancer antagonizes pro-apoptotic signals. Survivin forms a dimer *in vitro*; however, the dimerization interface is disrupted when Survivin binds Borealin, thus inhibiting homodimer formation when in complex with the CPC (Jeyaprakash *et al.*, 2007; Verdecia *et al.*, 2000; Chantalat *et al.*, 2000). Survivin is phosphorylated by multiple kinases including Aurora B, CDK1, Plk1, and Casein kinase-2 providing multiple levels of regulation (Nogueira-Ferreira *et al.*, 2013).

Borealin interacts with Survivin and INCENP via two N-terminal α -helices located between amino acids 15 and 76 (Jeyaprakash *et al.*, 2007). Interestingly, the yeast homolog of Borealin, Nbl1, is roughly half the size of the vertebrate Borealin proteins; however, all Borealin homologs contain the N-terminal helix that mediates triple-helix bundle formation with INCENP and Survivin (Carmena *et al.*, 2012). The central region of Borealin (amino acids 106–219) contains multiple

CDK1 phosphorylation sites which have been linked to centromere localization of the CPC (Kaur *et al.*, 2010; Tsukahara *et al.*, 2010). The central region of Borealin also interacts with the ESCRT-III subunit Shrb/CHMP4C to regulate cytokinesis (Capalbo *et al.*, 2012; Carlton *et al.*, 2012). Deletion of part of the central region of Borealin created a protein that was more dramatically upregulated in response to proteasome inhibition, suggesting that this region may contribute to Borealin metabolism (Date *et al.*, 2012). Like Survivin, Borealin is overexpressed in a diverse group of human cancers, possible due to its regulation by the E2F family of cell cycle regulated transcription factors (Date *et al.*, 2007). The C-terminus of Borealin (amino acids 224–280) contains a dimerization domain whose structure has been solved (Bourhis *et al.*, 2009). Threonine 230 within the dimerization domain is phosphorylated by the mitotic kinase Mps1 (Bourhis *et al.*, 2009; Jelluma *et al.*, 2008). Mutation of T230 to either alanine or valine impairs Aurora B activity and chromosome alignment *in vivo*. However, the T230A mutant disrupts dimer formation while the T230V stabilizes it (Bourhis *et al.*, 2009). Therefore, the role of dimerization in Borealin function is unresolved. Borealin is also sumoylated in a RanBP2-dependent manner during early mitosis and the isopeptidase SENP3 catalyzes removal of SUMO2/SUMO3 from Borealin (Klein *et al.*, 2009). However, the function of sumoylation is unclear given that mutation of all lysines on Borealin does not result in obvious mitotic defects.

CPC during the Approach to Mitosis

During interphase, the CPC localizes to pericentromeric heterochromatin via INCENP binding to HP-1 (Ainsztein *et al.*, 1998). HP1 is recruited to heterochromatin by binding to histone H3 that is di- or tri-methylated at lysine 9 (H3K9me2/H3K9me3). Phosphorylation of serine 10 on H3, an event that begins in G2, displaces HP1 possibly to facilitate concentration of INCENP at centromeres during prometaphase (Carmena *et al.*, 2012). In an early study on isolated mitotic chromosomes phosphorylation of serine 10 of histone H3 was found as one of the most abundant modifications (Paulson and Taylor, 1982). Subsequent studies identified Aurora B as the major kinase responsible, and PP1 as the phosphatase counteracting this modification (Hsu *et al.*, 2000). During interphase, a number of kinases including Msk1 and Msk2 phosphorylate H3 at serine 10 to modulate transcription (Duncan *et al.*, 2006). Nonetheless, G2 phosphorylation of H3 at serine 10 by Aurora B marks the beginning of CPC activity in preparation for mitosis and provides one example of how the CPC controls its own localization.

Release of HP1 by serine 10 phosphorylation may also impact the pericentromeric heterochromatin environment. Centromeres are enriched in specific modified histones including H3K4me2, H3K36me2, and H3K9me2/H3K9me3. H3K9me2/H3K9me3 marks a large heterochromatin domain that extends beyond CenPA nucleosomes (Wang and Higgins, 2012). Evidence in fission yeast suggests that release of HP1 may be required to allow transcription of centromeric repeats to maintain a specific balance of histone modifications required for centromere function (Chen *et al.*, 2008). A similar

type of regulation may occur in mammals. For example, targeting the LSD1 H3K4-specific methyltransferase to a neocentromere resulted in the gradual loss of CenPA nucleosomes, reduction in centromere-associated noncoding transcripts, and loss of kinetochore markers (Bergmann *et al.*, 2011). Thus, the heterochromatin environment is essential for proper function of the centromere and phosphorylation of H3 at serine 10 likely contributes to this regulation. H3 phosphorylated at serine 10 plays a second key role in mitotic chromatin condensation. S10-phosphorylated H3 recruits Hst2 which deacetylates H4K16 allowing tighter nucleosome packing required for mitotic chromosome condensation (Wilkins *et al.*, 2014). Related to this finding is the observation that Aurora B is required for axial shortening of chromosomes which continues into until anaphase in mammalian cells (Mora-Bermudez *et al.*, 2007). Therefore, one of the most abundant Aurora B targets, serine-10 phosphorylated H3, plays multiple important roles in mitosis.

CPC in Mitosis

From prophase to telophase, localization of the CPC is dynamic and is an indication of the multiple roles the CPC plays in regulating mitotic progression and cell division. During early mitosis, the CPC is found at centromeres and diffusely localized along chromosome arms. Another key mitotic role of Aurora B is carried out during prophase. Along with Cdk1, Aurora B contributes to sister chromatid resolution by phosphorylating the cohesion-stabilizing protein Sororin (Losada, 2014; Dreier *et al.*, 2011). Phosphorylated Sororin dissociates from the cohesion subunit Pds5, which ultimately results in Wapl-mediated release of cohesion from chromosome arms. Polo-like kinase is also essential for prophase removal of cohesin, and works by phosphorylating cohesin subunits (Losada, 2014). Cells in which the prophase removal pathway is inhibited show an increase in chromosome loss upon completion of mitosis, indicating the importance of this pathway in maintaining genomic stability (Haarhuis *et al.*, 2013). As mitosis progresses, the CPC concentrates at inner centromeres where it participates in an essential, evolutionarily conserved surveillance mechanism required for high-fidelity chromosome segregation. The spindle assembly checkpoint (also known as the mitotic checkpoint) blocks entry into anaphase until all chromosomes attain a bipolar attachment to the spindle. The best recognized role of the CPC in this process is in 'error correction' where inappropriate spindle-kinetochore attachments are converted to unattached kinetochores that trigger the SAC.

The Spindle Assembly Checkpoint

Observations of mitotic animal cells indicated that initiation of chromosome segregation correlated with alignment of chromosomes at the metaphase plate. Furthermore, inhibiting spindle assembly with small molecules that disrupt the dynamic instability of microtubules ultimately halted cell cycle progression at mitosis. These observations suggested that a checkpoint coupled mitotic exit to chromosomal alignment

and spindle assembly (Rieder, 2011). Two seminal papers describing yeast mutants that fail to arrest in mitosis when exposed to microtubule-disrupting agents, termed mitotic-arrest deficient (*mad*) and budding-uninhibited by benzimidazole (*bub*), led to the idea that mitosis is indeed under checkpoint control (Li and Murray, 1991; Hoyt *et al.*, 1991). Additional research established that the fidelity of centromeres and kinetochores, the multi-protein complexes that connect chromosomes to the spindle, was essential for proper chromosome segregation (e.g., Spencer and Hieter, 1992). An elegant study that realized the importance of kinetochores in regulating the spindle assembly checkpoint showed that laser ablation of the last unattached kinetochore rapidly induced mitotic exit indicating that unattached kinetochores are the primary signal that activates the spindle assembly checkpoint (Rieder *et al.*, 1994). Not surprisingly, subsequent biochemical and cytological evaluation of *Mad* and *Bub* gene products indicated that many *Mad* and *Bub* proteins associate with kinetochores (Cheeseman, 2014). In particular, the homologous *mad2* gene product in *Xenopus laevis*, XMAD2, was described as a dynamic kinetochore-associated protein that bound to unattached kinetochores, where XMAD2 localization to kinetochores was reduced upon chromosome alignment, suggesting that *Mad2* was at the apex of spindle assembly checkpoint activation (Chen *et al.*, 1996). Additional genetic and biochemical studies indicated that *Mad2* interacted with an E3 ubiquitin ligase known as the anaphase promoting complex/cyclosome (APC/C) (Murray, 1992). The APC/C targets Cyclin B for proteasomal degradation to extinguish Cdk1 activity and drive mitotic exit. *Mad2* inhibits the APC/C to keep cells in a mitotic state until all chromosomes attain bipolarity (He *et al.*, 1997; Li *et al.*, 1997). Further characterization of homologous *mad* and *bub* genes in higher eukaryotes eventually led to the general conclusion that unattached kinetochores catalyze the formation of an inhibitory complex known as the mitotic checkpoint complex (MCC), composed of *Mad2*, Bub-related 1 (*BubR1*), *Bub3*, and the APC/C co-activator *Cdc20* (Figure 2). The MCC inhibits the APC/C by sequestering *Cdc20*, which leads to Cyclin B stabilization and spindle assembly checkpoint activation via sustained CDK1 activity (Figure 2). Formation of the MCC ceases only when sister kinetochores bind to microtubules from opposing spindle poles, which allows *Cdc20* to activate the APC/C. Consequently, activation of the APC/C leads to Cyclin B degradation, which inactivates CDK1 and induces mitotic exit.

Spindle Assembly Checkpoint and Inter-Kinetochore Tension

Because the spindle assembly checkpoint appeared to ensure proper attachment of kinetochores to spindle microtubules, it was not clear how improper kinetochore-microtubule attachments were sensed and regulated. For instance, when a single sister kinetochore is bound to the both spindle poles (merotelic) or both sister kinetochores are bound to microtubules from the same spindle pole (syntelic), initiating mitotic exit when merotelic or syntelic attachments persists often leads to chromosome mis-segregation resulting in the

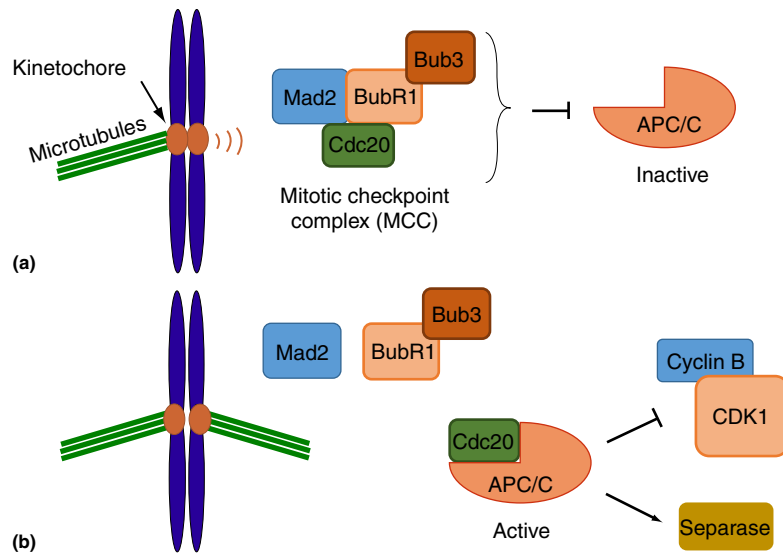


Figure 2 The spindle assembly checkpoint. (a) Prior to anaphase initiation, unattached kinetochores generate an inhibitory complex composed of Mad2, BubR1, Bub3, and Cdc20, known as the MCC. The MCC effectively sequesters Cdc20, which is a co-activator of the E3 ubiquitin ligase APC/C. (b) When chromosomes become bi-oriented and kinetochores are stably attached to spindle microtubules, the MCC disassembles and Cdc20 is free to activate the APC/C. Active APC/C targets essential mitotic substrates such as Cyclin B and securin for proteasomal degradation ultimately leading to anaphase initiation and mitotic exit.

generation of daughter cells with unequal DNA content, known as aneuploidy (Kops *et al.*, 2005). Initially it was hypothesized that the physical tension that arises from bi-orientation of chromosomes, or sister-kinetochore attachment to opposite spindle poles, regulated the spindle assembly checkpoint given that bi-orientation is consistent with anaphase initiation and successful cell division (Rieder *et al.*, 1994). Indeed, application of tension on improperly attached chromosomes leads to anaphase initiation earlier compared to the absence of tension (Li and Nicklas, 1995). Additionally, it was observed that increased tension between sister kinetochores and anaphase initiation correlated with dephosphorylation of kinetochore proteins indicating that kinetochore chemistry is sensitive to tension and may link mitotic forces to checkpoint activation (Nicklas *et al.*, 1995). Further characterization of Mad and Bub proteins indicated that their kinetochore localization was regulated differentially in response to either kinetochore-microtubule attachment or sister-kinetochore tension (Waters *et al.*, 1998). Together, these observations led to speculation that either microtubule occupancy or physical tension induced by microtubule capture negatively regulated spindle assembly checkpoint activation. Yet again, observations of cell cycle dynamics in yeast mutants led to a molecular link between tension and spindle assembly checkpoint activation. The observation that yeast *cdc6* mutants, which are allowed to progress into mitosis without replicating their genome, activate the spindle assembly checkpoint despite attaching all kinetochores to microtubules was evidence that lack of tension between sister kinetochores was the primary activator of the spindle assembly checkpoint (Stern and Murray, 2001). Furthermore, the discovery that mutating the gene *Ipl1*, which had been described as a regulator of microtubule turnover at kinetochores, overcame the

tensionless arrest in *cdc6* mutants suggested that tension was sensed by the *Ipl1* gene product (Biggins and Murray, 2001). Altogether, these observations led to a more refined model of spindle assembly checkpoint regulation where unattached kinetochores initially activate the spindle assembly checkpoint and the checkpoint signal is inactivated when microtubules bind kinetochores (Figure 3). However, kinetochore-microtubule interactions that do not result in the physical separation of sister kinetochores are destabilized resulting in MCC formation and spindle assembly checkpoint activation (Figure 3). Only when kinetochores capture microtubules from opposing spindle poles is inter-kinetochore tension high enough to separate sister kinetochores and silence MCC production (Figure 3). Thus, activation of the spindle assembly checkpoint is sensitive to kinetochore-microtubule attachment, but also inter-kinetochore tension. The tension-regulated error-correction mechanism was ultimately identified as the CPC.

CPC and Error Correction

An attractive model that explains the role of the CPC in error correction is known as the spatial gradient model and suggests that Aurora B phosphorylates microtubule-binding kinetochore proteins under low inter-kinetochore tension which leads to dissociation of kinetochore-microtubule interactions (Figure 4; Carmena *et al.*, 2012). As chromosomes become bi-oriented and sister kinetochores are under high tension, Aurora B is spatially separated from the kinetochore and can no longer destabilize kinetochore-microtubule interactions. Indeed, sister chromosomes that have not bi-oriented display relatively unstable kinetochore-microtubule interactions,

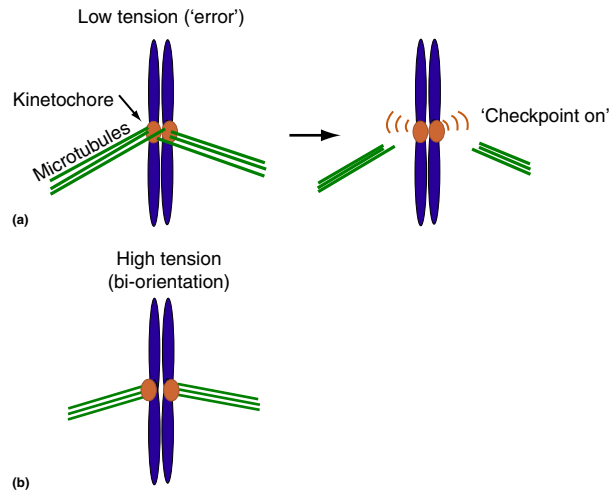


Figure 3 Inter-kinetochore tension and error correction (a) Kinetochore–microtubule interactions that do not result in high tension between sister kinetochores (inter-kinetochore tension), such as merotelic or syntelic attachments, are erroneous and selectively destabilized, which enables MCC production and activation of the mitotic checkpoint. (b) When sister kinetochores attach to microtubules from opposing spindle poles, or bi-orientation, inter-kinetochore tension is high and kinetochore–microtubule interactions are stabilized ultimately silencing MCC production which promotes mitotic exit.

compared to bi-oriented chromosomes, which requires Aurora B kinase activity (Carmena *et al.*, 2012). Additionally, the core microtubule-binding kinetochore proteins of the KMN network, which consists of the Knl1 complex, the Mis12 complex, and the Ndc80/Hec1 complex, have been identified as Aurora B substrates and Aurora B-mediated phosphorylation reduces their affinity for microtubules *in vitro* (Welburn *et al.*, 2010). Moreover, an Aurora B FRET-sensor detects a gradient of Aurora B activity that is highest near inner centromeres (Figure 4; Fuller *et al.*, 2008). Thus, according to this spatial gradient model of CPC-mediated checkpoint activation, inner centromere localization of the CPC is paramount for regulating error correction and spindle assembly checkpoint activation. Localization of the CPC to inner centromeres during mitosis is common to yeast, zebra fish, frog, chicken, mouse, and human cells, suggesting that centromere accumulation is essential for the function of the CPC.

Recent studies in yeast suggest that the role of CPC in error correction may be more complicated than spatial restriction due to an inner centromere location. For example, yeast cells expressing the INCENP mutant (Sli15ΔN) are viable and exhibit normal chromosome segregation despite the fact that this mutant fails to bind Survivin and is unable to localize Aurora B to centromeres (Campbell and Desai, 2013). Therefore, yeast and animal cells may differ in the dependency on CPC centromere location for proper function. Some studies have detected CPC family members at the kinetochore, a finding that may be at odds with the spatial gradient model (Petsalaki *et al.*, 2011). However, PP1-dependent dephosphorylation of kinetochore substrates may help to counteract CPC that spreads from the inner centromere, especially when kinetochores are not under tension.

The Two-Histone Model of CPC Localization to Centromeres

Localization of the CPC to centromeres reportedly intersects with two histone modifications that overlap at inner centromeres: phosphorylated histone H3 at threonine 3 (pH3T3) and phosphorylated histone 2A at threonine 120 (pH2AT120) (Figure 5; Yamagishi *et al.*, 2010). Haspin kinase generates pH3T3, which is in turn bound by Survivin. The BIR domain of Survivin directly interacts with pH3T3 to recruit the CPC and this interaction can be regulated by pH, suggesting that the *in vivo* interaction could be regulated by the local environment (Jeyaparakash *et al.*, 2011). The CPC engages in positive feedback to maintain pH3T3 levels during mitosis. For instance, Aurora B directly phosphorylates Haspin at multiple residues during mitosis and mutation of these residues reduces pH3T3 (Wang *et al.*, 2011). Consistent with this observation, inhibiting Aurora B during mitosis leads to a reduction in pH3T3 levels. However, the mechanism of Aurora B-mediated regulation of Haspin activity is unknown because mutation of Aurora B phosphorylation sites does not reduce intrinsic kinase activity of Haspin. Aurora B phosphorylation might regulate the release of a Haspin inhibitor or facilitate substrate binding (Wang *et al.*, 2011). Dephosphorylation of pH3T3 is mediated by the phosphatase PP1γ–Repo-Man complex and Aurora B-mediated phosphorylation negatively regulates PP1γ–Repo-Man activity toward pH3T3 (Qian *et al.*, 2011). Thus, Aurora B positively regulates pH3T3 abundance by promoting Haspin kinase activity toward H3 and negatively regulating removal of pH3T3 by phosphorylating PP1γ–Repo-Man.

The second histone mark that dictates CPC localization to centromeres is generated by the kinetochore-localized kinase Bub1. Kinetochore-bound Bub1 generates pH2AT120 at the kinetochore, which recruits the Shugoshin proteins (Sgo1/Sgo2) (Yamagishi *et al.*, 2010). Sgo1/Sgo2 can bind CDK1-phosphorylated Survivin in yeast and Borealin in humans, which may explain why centromere accumulation of the CPC begins in late prophase when CDK1 activity is highest in the disassembled nucleus (Tsukahara *et al.*, 2010). Interestingly, the BIR domain of Survivin can bind the N-terminus of human Sgo1 *in vitro*, suggesting cross talk between the two-histone pathways (Jeyaparakash *et al.*, 2011). The CPC also indirectly engages in feedback with Bub1 to generate pH2AT120. For example, Aurora B kinase activity is essential for kinetochore targeting and activation of the mitotic kinase Monopolar Spindle 1 (Mps1) (Saurin *et al.*, 2011). Mps1 phosphorylates the kinetochore protein Knl1 at Met-Glu-Leu-Thr (MELT) motifs, which recruits Bub1 (Primorac *et al.*, 2013). In turn, Bub1 generates pH2AT120 to recruit Sgo1/Sgo2 and consequently the CPC via the Borealin–Sgo1/Sgo2 interaction. Additionally, Aurora B phosphorylates ataxia telangiectasia mutated (ATM) which enhances the kinase activity of Bub1 (Yang *et al.*, 2011).

Modified View of CPC Targeting

Whether or not the CPC contacts both modified histones in the same complex has not been formally demonstrated. Also, the role of both pH3T3 and pH2AT120 in mediating CPC

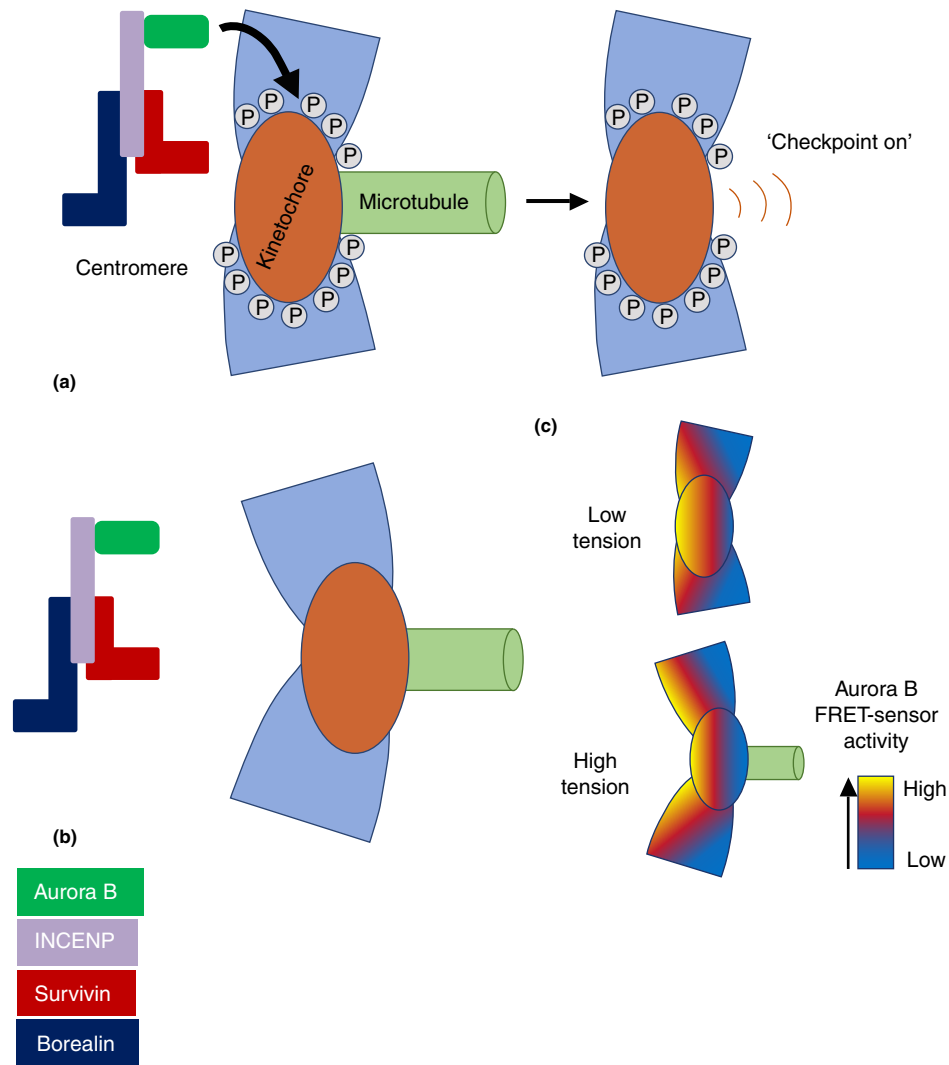


Figure 4 Spatial gradient model of CPC-mediated mitotic checkpoint activation. (a) Under low inter-kinetochore tension, the CPC is positioned in close proximity to kinetochores where Aurora B can phosphorylate microtubule-binding kinetochore proteins and reduce their affinity for microtubules. The reduced affinity for microtubules results in dissociation of erroneous kinetochore-microtubule interactions and production of the MCC which activates the mitotic checkpoint. (b) When sister kinetochores come under high tension from bi-orientation, kinetochores are physically separated from Aurora B at centromeres and become dephosphorylated, which stabilizes the kinetochore-microtubule interactions. (c) An Aurora B FRET-sensor placed at centromeres and various locations along the kinetochore reveals a gradient of Aurora B activity along kinetochores. Under low tension, a gradual gradient of Aurora B activity emanates from the inner centromere along the kinetochore axis. However, under high tension the Aurora B phosphorylation gradient is steep.

localization to centromeres may be more complicated. For example, Aurora B localization is unaffected when either Mps1 or Aurora B are inhibited despite large reductions in pH2AT120 (van der Waal *et al.*, 2012). Structure-function analysis of Borealin has also suggested a more complicated view of CPC targeting to the centromere. A small fragment of Borealin from amino acids 1–110 that lacks the reported Sgo-binding region still localizes, although transiently to centromeres (Bekier *et al.*, 2015). Similar transient localization of Borealin to centromeres is observed with a truncated form (amino acids 1–221) containing the central region but lacking the dimerization domain. Replacing the Borealin dimerization domain with FK506-binding protein (FKBP) creates a protein that localizes to centromeres in the absence of a dimeric

rapalogue, but much more strongly when dimerization is induced (Bekier *et al.*, 2015). Interestingly, a small population of Borealin is observed at kinetochores when the bulk of pH3T3 is reduced with a small molecule Haspin inhibitor. Furthermore, monomeric Borealin1-221-FKBP fails to localize to kinetochores after Haspin inhibition, but localization is rescued by adding dimerizer (Bekier *et al.*, 2015). In contrast, Borealin1-110-FKBP fails to localize to kinetochores even when dimerization is induced. Kinetochore localization of full-length Borealin partially overlaps with residual pH2AT120 after Haspin inhibition, although other binding sites cannot be ruled out. Based on these results, a modified model of CPC targeting suggests that two pools of CPC may exist, one at the inner centromere and a second at the kinetochore, the latter of

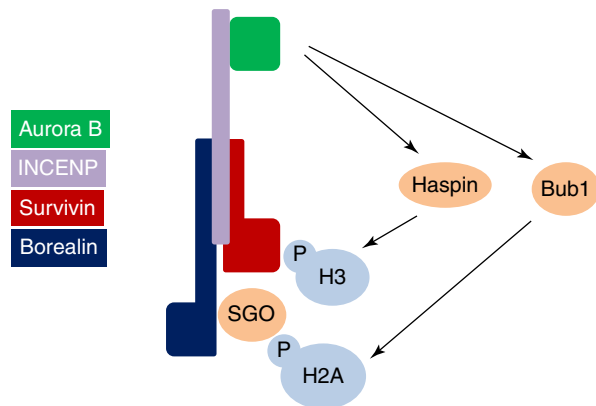


Figure 5 Two-histone model of CPC localization to centromeres. The CPC engages feedback with Bub1 and Haspin to generate two histone modifications that mediate CPC localization to inner centromeres: Bub1 generates pH2AT120 which recruits Sgo1 and binds the phosphorylated central region of Borealin while Haspin generates pH3T3 which binds the BIR domain of Survivin.

which is most evident in taxol-arrested cells (Bekier *et al.*, 2015). Dimerization of the CPC via Borealin is essential for optimal centromere association. CPC monomers may associate with the inner centromere transiently via a single Survivin/pH3T3 contact, whereas two such contacts allow stable binding in the context of a CPC dimer. Association of the CPC with kinetochores only occurs in a dimer which may allow dual Borealin/pH2AT120 contacts. These results also imply that Survivin binds stronger to its centromere target than the Borealin central region binds its kinetochore target.

CPC during Anaphase, Telophase, and Cytokinesis

The CPC translocates to the spindle midzone at anaphase to help regulate formation and constriction of the cleavage furrow (Carmena, 2008). This transition is under several levels of control. In the budding yeast, transition of the CPC from centromeres to microtubules of the spindle midzone is regulated by Ipl1 (Aurora B)- and Cdc28 (Cdk1)-dependent phosphorylation of the microtubule binding region of Sli15 (INCENP). This phosphorylation, which inhibits microtubule binding is reversed by the phosphatase Cdc14 at the metaphase/anaphase transition. Neither Cdk1 nor Aurora B phosphorylate the microtubule binding site in mammalian INCENP, however Cdk1-dependent phosphorylation of T59 does inhibit association of the CPC with the spindle midzone (Carmena *et al.*, 2012; Hummer and Mayer, 2009). Dephosphorylation of T59 at the metaphase/anaphase transition allows INCENP to interact with the kinesin MKLP2. Both proteins are mutually dependent for midzone targeting, with MKLP2 contributing to microtubule bundling (Hummer and Mayer, 2009). At the spindle midzone, the CPC appears to regulate the cleavage furrow at multiple levels. Aurora B phosphorylates MKLP1 (mitotic kinesin-like protein-1) which is part of the centralspindlin complex composed of two molecules of MKLP1 and two of MgcRacGap (male germ cell Rac-GTPase

activating protein) (Guse *et al.*, 2005). Aurora B-dependent phosphorylation of MKLP1 helps to recruit centralspindlin to the spindle midzone where it is needed to induce furrowing. MgcRacGap, possibly via its Gap activity, inhibits Rac (Canman *et al.*, 2008). Centralspindlin also binds Ect2, a guanine exchange factor that activates Rho in a localized manner to induce actomyosin-dependent contraction in a band that transects the cell equator (marked by the position of the spindle midzone) (Yuce *et al.*, 2005). Additional evidence in *Caenorhabditis elegans* suggests that the CPC has an additional function in inducing constriction of the cleavage furrow, independent of regulating centralspindlin (Lewellyn *et al.*, 2011). The targets for this additional role of the CPC have not been defined; however, depletion of the Septin UNC-59 alleviated the constriction defect observed upon inactivation of CPC but not centralspindlin (Lewellyn *et al.*, 2011). Septins are GTP-binding proteins that form a variety of higher order filamentous structures. Septins are well known as being essential for division in budding yeast, while their contribution to mammalian cell division appears to be cell-type dependent. That CPC might regulate Septin function is an intriguing prospect.

One of the characteristic features of all CPC members is their retention at the midbody during cytokinesis. Experiments in budding yeast first indicated that midbody localization is related to an additional key CPC function in regulating abscission. The NoCut pathway (also known as the abscission checkpoint) blocks abscission in cells that have progressed to cytokinesis with lagging chromosomes in the cleavage plane. In budding yeast, the blockade requires anillin-related proteins Boi1 and Boi2, and can be overridden by inactivating either Ipl1 (Aurora B) or Sli15 (INCENP) (Norden *et al.*, 2006). Furthermore, Boi1 and Boi2 are recruited to the bud neck in cells undergoing abscission delay, but only when Ipl1 is functional. These results suggest a model in which the CPC recruits Boi1 and Boi2 to the cleavage plane when lagging DNA is present. Boi1 and Boi2 then inhibit abscission (Norden *et al.*, 2006).

Although some of the players differ in mammalian cells, it is clear that the CPC also delays abscission in response to DNA in the cleavage furrow. Abscission involves highly coordinated membrane rearrangements and depends on the endosomal sorting complex required for transport (ESCRT) machinery and the VPS4 AAA ATPase. Borealin interacts with an additional protein charged multivesicular body (MVB) protein 4C (CHMP4C), an ESCRT subunit. CHMP4C is required for the abscission delay. Aurora B phosphorylates CHMP4C at serine 210, possibly while it is bound to Borealin (Carlton *et al.*, 2012). S210 phosphorylation by Aurora B is required for CHMP4C to inhibit abscission. To inhibit abscission, phosphorylated CHMP4C collaborates with a new cytokinesis regulator, abscission/NoCut checkpoint regulator (ANCHR) to retain VPS4 at the midbody ring (Thoresen *et al.*, 2014). Only when S210 phosphorylation is reduced is VPS4 released from the CHMP4C-ANCHR complex to catalyze abscission.

Conclusion

Mitosis must occur accurately to ensure that daughter cells contain an intact copy of the genome. This process is

fundamental to the life of all cells, yet we are early in our understanding of how this is all controlled. The CPC contributes in many ways to the accurate segregation of the genome. Starting at its concentration at the centromere where it controls attachment of chromosomes to the spindle, to its monitoring of wayward chromosomes in the cleavage plane that would be fragmented if abscission were allowed to proceed, the CPC is essential for proper movement through mitosis. Given its central role in dividing cells, many attempts are being made to target CPC members in cancer. Inhibitors of Aurora B and Survivin are at various clinical trial stages with some evidence that these treatments may outperform classical chemotherapy regimens (Kantarjian *et al.*, 2013). Future studies that uncover deeper aspect of CPC function may provide new drug targets to investigate.

See also: Functional Cell Biology: An Overview

References

- Ainsztein, A.M., Kandels-Lewis, S.E., Mackay, A.M., Earnshaw, W.C., 1998. INCENP centromere and spindle targeting: Identification of essential conserved motifs and involvement of heterochromatin protein HP1. *Journal of Cell Biology* 143 (7), 1763–1774.
- Altieri, D.C., 2010. Survivin and IAP proteins in cell-death mechanisms. *Biochemical Journal* 430 (2), 199–205.
- Bekier, M.E., Mazur, T., Rashid, M.S., Taylor, W.R., 2015. Borealin dimerization mediates optimal CPC checkpoint function by enhancing localization to centromeres and kinetochores. *Nature Communications* 6, 6775. doi: 10.1038/ncomms7775.
- Bergmann, J.H., Rodriguez, M.G., Martins, N.M., *et al.*, 2011. Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *EMBO Journal* 30 (2), 328–340.
- Biggins, S., Murray, A.W., 2001. The budding yeast protein kinase Ipl1/Aurora allows the absence of tension to activate the spindle checkpoint. *Genes & Development* 15 (23), 3118–3129.
- Bishop, J.D., Schumacher, J.M., 2002. Phosphorylation of the carboxyl terminus of inner centromere protein (INCENP) by the Aurora B Kinase stimulates Aurora B kinase activity. *Journal of Biological Chemistry* 277 (31), 27577–27580.
- Bourhis, E., Lingel, A., Phung, Q., Fairbrother, W.J., Cochran, A.G., 2009. Phosphorylation of a borealin dimerization domain is required for proper chromosome segregation. *Biochemistry* 48 (29), 6783–6793.
- Campbell, C.S., Desai, A., 2013. Tension sensing by Aurora B kinase is independent of survivin-based centromere localization. *Nature* 497 (7447), 118–121.
- Canman, J.C., Lewellyn, L., Laband, K., *et al.*, 2008. Inhibition of Rac by the GAP activity of centralspindlin is essential for cytokinesis. *Science* 322 (5907), 1543–1546.
- Capalbo, L., Montembault, E., Takeda, T., *et al.*, 2012. The chromosomal passenger complex controls the function of endosomal sorting complex required for transport-III Snf7 proteins during cytokinesis. *Open Biology* 2 (5), 120070.
- Carlton, J.G., Caballe, A., Agromayor, M., Kloc, M., Martin-Serrano, J., 2012. ESCRT-III governs the Aurora B-mediated abscission checkpoint through CHMP4C. *Science* 336 (6078), 220–225.
- Carmena, M., 2008. Cytokinesis: The final stop for the chromosomal passengers. *Biochemical Society Transactions* 36 (Pt 3), 367–370.
- Carmena, M., Wheelock, M., Funabiki, H., Earnshaw, W.C., 2012. The chromosomal passenger complex (CPC): From easy rider to the godfather of mitosis. *Nature Reviews Molecular Cell Biology* 13 (12), 789–803.
- Chan, C.S., Botstein, D., 1993. Isolation and characterization of chromosome-gain and increase-in-ploidy mutants in yeast. *Genetics* 135 (3), 677–691.
- Chantalat, L., Skoufias, D.A., Kleman, J.P., *et al.*, 2000. Crystal structure of human survivin reveals a bow tie-shaped dimer with two unusual alpha-helical extensions. *Molecular Cell* 6 (1), 183–189.
- Cheeseman, I.M., 2014. The kinetochore. *Cold Spring Harbor Perspectives in Biology* 6 (7), a015826.
- Chen, E.S., Zhang, K., Nicolas, E., *et al.*, 2008. Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* 451 (7179), 734–737.
- Chen, Q., Lakshmikanth, G.S., Spudich, J.A., De Lozanne, A., 2007. The localization of inner centromeric protein (INCENP) at the cleavage furrow is dependent on Kif12 and involves interactions of the N terminus of INCENP with the actin cytoskeleton. *Molecular Biology of the Cell* 18 (9), 3366–3374.
- Chen, R.H., Waters, J.C., Salmon, E.D., Murray, A.W., 1996. Association of spindle assembly checkpoint component XMAP2 with unattached kinetochores. *Science* 274 (5285), 242–246.
- Cooke, C.A., Heck, M.M., Earnshaw, W.C., 1987. The inner centromere protein (INCENP) antigens: Movement from inner centromere to midbody during mitosis. *Journal of Cell Biology* 105 (5), 2053–2067.
- Date, D., Dreier, M.R., Borton, M.T., Bekier, M.E., Taylor, W.R., 2012. Effects of phosphatase and proteasome inhibitors on Borealin phosphorylation and degradation. *Journal of Biochemistry* 151 (4), 361–369.
- Date, D.A., Jacob, C.J., Bekier, M.E., *et al.*, 2007. Borealin is repressed in response to p53/Rb signaling. *Cell Biology International* 31 (12), 1470–1481.
- Dreier, M.R., Bekier, M.E., Taylor, W.R., 2011. Regulation of sororin by Cdk1-mediated phosphorylation. *Journal of Cell Science* 124 (Pt 17), 2976–2987.
- Duncan, E.A., Anest, V., Cogswell, P., Baldwin, A.S., 2006. The kinases MSK1 and MSK2 are required for epidermal growth factor-induced, but not tumor necrosis factor-induced, histone H3 Ser10 phosphorylation. *Journal of Biological Chemistry* 281 (18), 12521–12525.
- Eyers, P.A., Churchill, M.E., Maller, J.L., 2005. The Aurora A and Aurora B protein kinases: A single amino acid difference controls intrinsic activity and activation by TPX2. *Cell Cycle* 4 (6), 784–789.
- Fuller, B.G., Lampson, M.A., Foley, E.A., *et al.*, 2008. Midzone activation of aurora B in anaphase produces an intracellular phosphorylation gradient. *Nature* 453 (7198), 1132–1136.
- Gassmann, R., Carvalho, A., Henzing, A.J., *et al.*, 2004. Borealin: A novel chromosomal passenger required for stability of the bipolar mitotic spindle. *Journal of Cell Biology* 166 (2), 179–191.
- Guse, A., Mishima, M., Glotzer, M., 2005. Phosphorylation of ZEN-4/MKLP1 by aurora B regulates completion of cytokinesis. *Current Biology* 15 (8), 778–786.
- Haarhuis, J.H., Elbatsh, A.M., van den Broek, B., *et al.*, 2013. WAPL-mediated removal of cohesin protects against segregation errors and aneuploidy. *Current Biology* 23 (20), 2071–2077.
- He, X., Patterson, T.E., Sazer, S., 1997. The Schizosaccharomyces pombe spindle checkpoint protein mad2p blocks anaphase and genetically interacts with the anaphase-promoting complex. *Proceedings of the National Academy of Sciences of the United States of America* 94 (15), 7965–7970.
- Hoyt, M.A., Totis, L., Roberts, B.T., 1991. *S. cerevisiae* genes required for cell cycle arrest in response to loss of microtubule function. *Cell* 66 (3), 507–517.
- Hsu, J.Y., Sun, Z.W., Li, X., *et al.*, 2000. Mitotic phosphorylation of histone H3 is governed by Ipl1/aurora kinase and Glc7/PP1 phosphatase in budding yeast and nematodes. *Cell* 102 (3), 279–291.
- Hummer, S., Mayer, T.U., 2009. Cdk1 negatively regulates midzone localization of the mitotic kinesin Mklp2 and the chromosomal passenger complex. *Current Biology* 19 (7), 607–612.
- Jelluma, N., Brenkman, A.B., van den Broek, N.J., *et al.*, 2008. Mps1 phosphorylates Borealin to control Aurora B activity and chromosome alignment. *Cell* 132 (2), 233–246.
- Jeyaprakash, A.A., Basquin, C., Jayachandran, U., Conti, E., 2011. Structural basis for the recognition of phosphorylated histone h3 by the survivin subunit of the chromosomal passenger complex. *Structure* 19 (11), 1625–1634.
- Jeyaprakash, A.A., Klein, U.R., Lindner, D., *et al.*, 2007. Structure of a Survivin-Borealin-INCENP core complex reveals how chromosomal passengers travel together. *Cell* 131 (2), 271–285.
- Kantarjian, H.M., Martinelli, G., Jabbour, E.J., *et al.*, 2013. Stage I of a phase 2 study assessing the efficacy, safety, and tolerability of barasertib (AZD1152) versus low-dose cytosine arabinoside in elderly patients with acute myeloid leukemia. *Cancer* 119 (14), 2611–2619.
- Kaur, H., Bekier, M.E., Taylor, W.R., 2010. Regulation of Borealin by phosphorylation at serine 219. *Journal of Cellular Biochemistry* 111 (5), 1291–1298.
- Kim, J.H., Kang, J.S., Chan, C.S., 1999. Sli15 associates with the ipl1 protein kinase to promote proper chromosome segregation in *Saccharomyces cerevisiae*. *Journal of Cell Biology* 145 (7), 1381–1394.
- Kirioukhova, O., Johnston, A.J., Kleen, D., *et al.*, 2011. Female gametophytic cell specification and seed development require the function of the putative *Arabidopsis* INCENP ortholog WYRD. *Development* 138 (16), 3409–3420.

- Klein, U.R., Haindl, M., Nigg, E.A., Muller, S., 2009. RanBP2 and SENP3 function in a mitotic SUMO2/3 conjugation-deconjugation cycle on Borealin. *Molecular Biology of the Cell* 20 (1), 410–418.
- Kops, G.J., Weaver, B.A., Cleveland, D.W., 2005. On the road to cancer: Aneuploidy and the mitotic checkpoint. *Nature Reviews Cancer* 5 (10), 773–785.
- Lewellyn, L., Carvalho, A., Desai, A., Maddox, A.S., Oegema, K., 2011. The chromosomal passenger complex and centralspindlin independently contribute to contractile ring assembly. *Journal of Cell Biology* 193 (1), 155–169.
- Li, R., Murray, A.W., 1991. Feedback control of mitosis in budding yeast. *Cell* 66 (3), 519–531.
- Li, X., Nicklas, R.B., 1995. Mitotic forces control a cell-cycle checkpoint. *Nature* 373 (6515), 630–632.
- Li, Y., Gorbea, C., Mahaffey, D., Rechsteiner, M., Benezra, R., 1997. MAD2 associates with the cyclosome/anaphase-promoting complex and inhibits its activity. *Proceedings of the National Academy of Sciences of the United States of America* 94 (23), 12431–12436.
- Losada, A., 2014. Cohesin in cancer: Chromosome segregation and beyond. *Nature Reviews Cancer* 14 (6), 389–393.
- Mora-Bermudez, F., Gerlich, D., Ellenberg, J., 2007. Maximal chromosome compaction occurs by axial shortening in anaphase and depends on Aurora kinase. *Nature Cell Biology* 9 (7), 822–831.
- Murray, A.W., 1992. Creative blocks: Cell-cycle checkpoints and feedback controls. *Nature* 359 (6396), 599–604.
- Nicklas, R.B., Ward, S.C., Gorbisky, G.J., 1995. Kinetochore chemistry is sensitive to tension and may link mitotic forces to a cell cycle checkpoint. *Journal of Cell Biology* 130 (4), 929–939.
- Nogueira-Ferreira, R., Vitorino, R., Ferreira-Pinto, M.J., Ferreira, R., Henriques-Coelho, T., 2013. Exploring the role of post-translational modifications on protein-protein interactions with survivin. *Archives of Biochemistry and Biophysics* 538 (2), 64–70.
- Norden, C., Mendoza, M., Dobbelaere, J., *et al.*, 2006. The NoCut pathway links completion of cytokinesis to spindle midzone function to prevent chromosome breakage. *Cell* 125 (1), 85–98.
- Paulson, J.R., Taylor, S.S., 1982. Phosphorylation of histones 1 and 3 and nonhistone high mobility group 14 by an endogenous kinase in HeLa metaphase chromosomes. *Journal of Biological Chemistry* 257 (11), 6064–6072.
- Petsalaki, E., Akoumianaki, T., Black, E.J., Gillespie, D.A., Zachos, G., 2011. Phosphorylation at serine 331 is required for Aurora B activation. *Journal of Cell Biology* 195 (3), 449–466.
- Primorac, I., Weir, J.R., Chirol, E., *et al.*, 2013. Bub3 reads phosphorylated MELT repeats to promote spindle assembly checkpoint signaling. *eLife* 2, e01030.
- Qian, J., Lesage, B., Beullens, M., Van Eynde, A., Bollen, M., 2011. PP1/Repo-man dephosphorylates mitotic histone H3 at T3 and regulates chromosomal aurora B targeting. *Current Biology* 21 (9), 766–773.
- Rieder, C.L., 2011. Mitosis in vertebrates: The G2/M and M/A transitions and their associated checkpoints. *Chromosome Research* 19 (3), 291–306.
- Rieder, C.L., Schultz, A., Cole, R., Sluder, G., 1994. Anaphase onset in vertebrate somatic cells is controlled by a checkpoint that monitors sister kinetochore attachment to the spindle. *Journal of Cell Biology* 127 (5), 1301–1310.
- Rosasco-Nitcher, S.E., Lan, W., Khorasanizadeh, S., Stukenberg, P.T., 2008. Centromeric Aurora-B activation requires TD-60, microtubules, and substrate priming phosphorylation. *Science* 319 (5862), 469–472.
- Saurin, A.T., van der Waal, M.S., Medema, R.H., Lens, S.M., Kops, G.J., 2011. Aurora B potentiates Mps1 activation to ensure rapid checkpoint establishment at the onset of mitosis. *Nature Communications* 2, 316.
- Spencer, F., Hieter, P., 1992. Centromere DNA mutations induce a mitotic delay in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences of the United States of America* 89 (19), 8908–8912.
- Stern, B.M., Murray, A.W., 2001. Lack of tension at kinetochores activates the spindle checkpoint in budding yeast. *Current Biology* 11 (18), 1462–1467.
- Thoresen, S.B., Campsteijn, C., Vietri, M., *et al.*, 2014. ANCHR mediates Aurora-B-dependent abscission checkpoint control through retention of VPS4. *Nature Cell Biology* 16 (6), 550–560.
- Tsukahara, T., Tanno, Y., Watanabe, Y., 2010. Phosphorylation of the CPC by Cdk1 promotes chromosome bi-orientation. *Nature* 467 (7316), 719–723.
- Verdecia, M.A., Huang, H., Dutil, E., *et al.*, 2000. Structure of the human anti-apoptotic protein survivin reveals a dimeric arrangement. *Nature Structural Biology* 7 (7), 602–608.
- van der Waal, M.S., Saurin, A.T., Vromans, M.J., *et al.*, 2012. Mps1 promotes rapid centromere accumulation of Aurora B. *EMBO Reports* 13 (9), 847–854.
- Wang, F., Higgins, J.M., 2012. Histone modifications and mitosis: Countermarks, landmarks, and bookmarks. *Trends in Cell Biology* 23 (4), 175–184.
- Wang, F., Ulyanova, N.P., van der Waal, M.S., *et al.*, 2011. A positive feedback loop involving Haspin and Aurora B promotes CPC accumulation at centromeres in mitosis. *Current Biology* 21 (12), 1061–1069.
- Waters, J.C., Chen, R.H., Murray, A.W., Salmon, E.D., 1998. Localization of Mad2 to kinetochores depends on microtubule attachment, not tension. *Journal of Cell Biology* 141 (5), 1181–1191.
- Welburn, J.P., Vleugel, M., Liu, D., *et al.*, 2010. Aurora B phosphorylates spatially distinct targets to differentially regulate the kinetochore-microtubule interface. *Molecular Cell* 38 (3), 383–392.
- Wilkins, B.J., Rall, N.A., Ostwal, Y., *et al.*, 2014. A cascade of histone modifications induces chromatin condensation in mitosis. *Science* 343 (6166), 77–80.
- Yamagishi, Y., Honda, T., Tanno, Y., Watanabe, Y., 2010. Two histone marks establish the inner centromere and chromosome bi-orientation. *Science* 330 (6001), 239–243.
- Yang, C., Tang, X., Guo, X., *et al.*, 2011. Aurora-B mediated ATM serine 1403 phosphorylation is required for mitotic ATM activation and the spindle checkpoint. *Molecular Cell* 44 (4), 597–608.
- Yang, D., Welm, A., Bishop, J.M., 2004. Cell division and cell survival in the absence of survivin. *Proceedings of the National Academy of Sciences of the United States of America* 101 (42), 15100–15105.
- Yuce, O., Piekny, A., Glotzer, M., 2005. An ECT2-centralspindlin complex regulates the localization and function of RhoA. *Journal of Cell Biology* 170 (4), 571–582.

Mitosis in Animal Cells

ATY Lian and M Chircop, Children's Medical Research Institute, Westmead, NSW, Australia and University of Sydney, Sydney, NSW, Australia

© 2016 Elsevier Inc. All rights reserved.

Introduction

Mitosis is the final phase of the cell cycle (Figure 1). It is a relatively short phase. For example, mitosis occupies approximately 10% of the total cell cycle time of a typical human proliferative cell (Alberts, 2002). Mitosis is the division process of the cell cycle whereby the duplicated chromosomes and cytoplasmic organelles are separated into two independent daughter cells (Nurse, 2000). As shown in Figure 2, this process is easily visible by conventional light microscopy. Interphase makes up the remainder of the cell cycle and proceeds mitosis. It consists of G1, S (synthesis), and G2 phases, whereby the chromosomes and centrosomes are duplicated during S phase.

Thus, interphase prepares the cell for division. Centrosomes are the microtubule organizing centers of the cell and are duplicated and divided in a cycle that is tightly coupled to the cell cycle (Figure 1). Its duplication is essential for equal segregation of the duplicated chromosomes as the centrosomes form the poles of the mitotic spindle. Mitosis only occurs in eukaryotic cells and this process differs amongst specific groups. For example, animal cells undergo an 'open' mitosis – nuclear envelope breaks down prior to chromosome segregation. Whereas fungi such as *Aspergillus nidulans* and *Saccharomyces cerevisiae* (yeast) undergo a 'closed' mitosis – chromosomes divide within an intact nucleus.

Mitosis is a highly complex process consisting of six distinct sequential stages – prophase, prometaphase, metaphase,

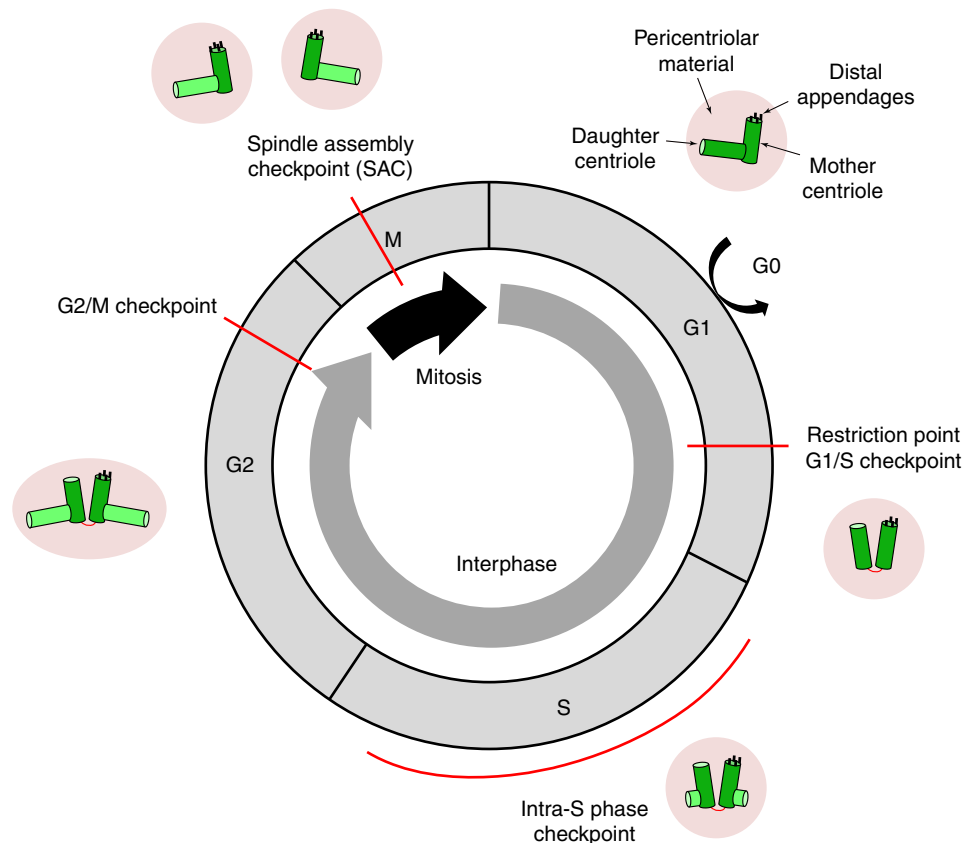


Figure 1 The animal cell cycle. The animal cell cycle is made up of four phases: (1) G1 phase, where cellular contents are duplicated; (2) S phase, where chromosomes are replicated; (3) G2 phase, where the cell ensures proper DNA replication; and (4) M phase (mitosis), where the cell is physically separated into two daughter cells. Together, G1, S, G2 phases are also known as interphase. The cell in G1 phase can enter into G0 phase, a quiescent phase. Cell cycle progression is tightly regulated by a series of checkpoint mechanisms that delay progression until errors are corrected. After completion of mitosis, the cell reenters G1, ready for the next cycle. The cell cycle is tightly coordinated with the centrosome cycle. A centrosome contains a mother centriole and a daughter centriole (dark green). The mother centriole is distinguishable by its additional distal appendages. The centrosomes are surrounded by the PCM (pink). The centrosomes are separated in late G1/early S phase in preparation of duplication, followed by the formation of procentrioles (daughter centrioles) at the proximal ends (light green). The procentrioles continue to elongate during S and G2 phases and reach maturity in late G2 phase. Centrosome separation occurs during mitosis and upon completion, each daughter cell inherits one centrosome.

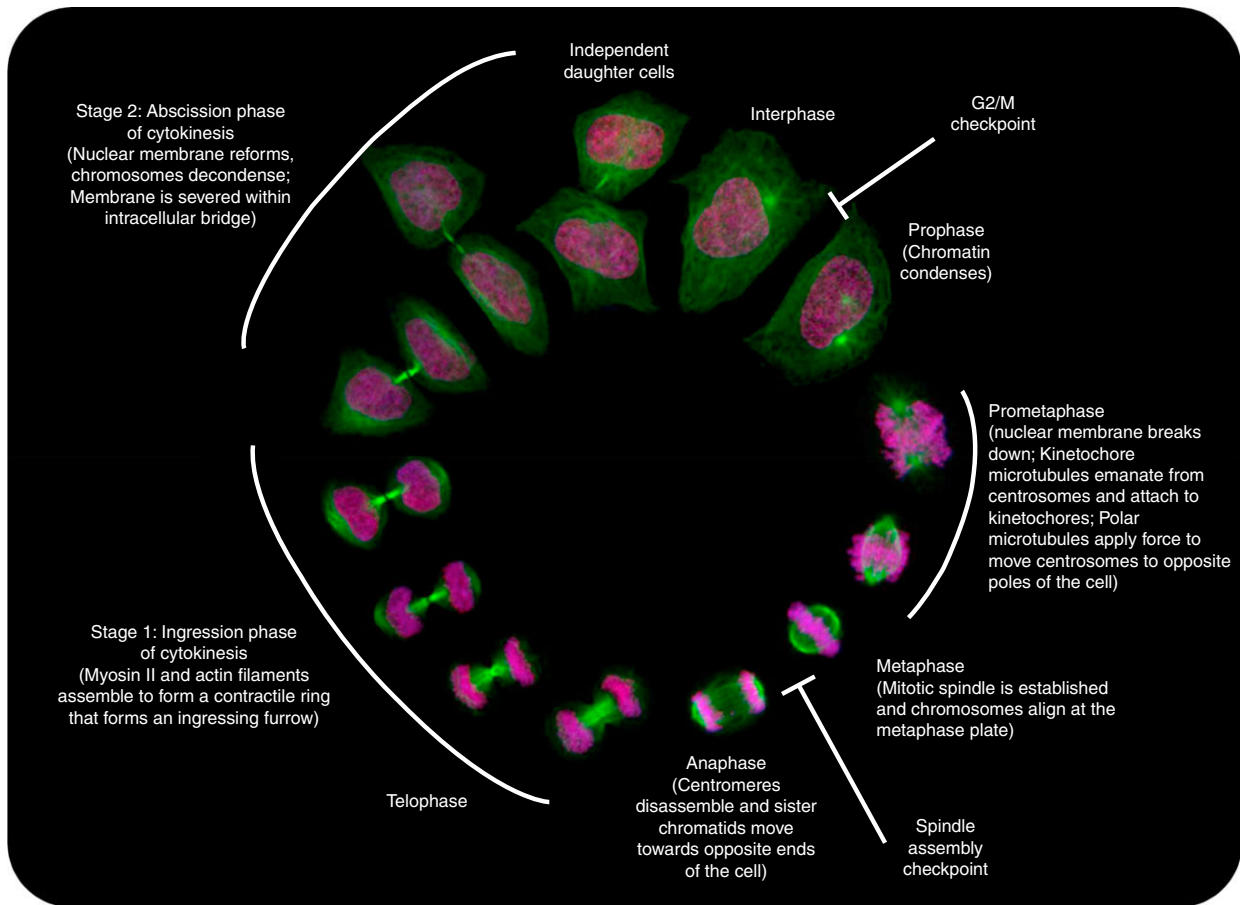


Figure 2 The multiple stages of mitosis. Mitosis begins with prophase where chromatin condenses into chromosomes (purple). The nuclear envelope breaks down in prometaphase and duplicated centrosomes move to the opposite poles of the cell. During metaphase, paired sister chromatids attach to the microtubules (green) from opposite poles at the centromere and align at the equatorial plane. The sister chromatids are separated and pulled toward opposite poles during anaphase. During telophase, a new nuclear envelope starts to form around each set of the separated sister chromatids and they condense and unfold back into the chromatin. Cytokinesis heavily overlaps with late anaphase and early telophase and requires two events: membrane ingression and membrane abscission. The formation of the actin–myosin II contractile ring begins in early telophase and its contraction triggers membrane ingression, forming the cleavage furrow. The furrow is fully ingressed in late telophase and the spindle midzone is compressed into a thin intercellular bridge with a protein-dense midbody ring in the center. Cytokinesis is completed when abscission physically splits daughter cells apart.

anaphase, telophase, and cytokinesis (**Figure 2**). Mitotic progression is highly regulated and involves dynamic modulation of the actin cytoskeleton, microtubule network (formation of mitotic spindle), as well as membranes to align and then equally segregate the genome and cytoplasmic contents into two daughter cells. Protein kinases play a pivotal role in executing cell division. The major protein kinases that mediate key mitotic events include the serine/threonine protein kinases, cyclin-dependent kinases (Cdks) whose catalytic activities are dependent on their regulatory cyclin subunits, polo and polo-like kinases (Plk), Aurora kinases, Nek (NIMA-related kinase), and Bub families, as well as Greatwall and Mps1/TTK ([Ma and Poon, 2011](#)). Spatial and temporal activation of protein kinases is essential for regulating centrosome maturation and separation, formation of the bipolar mitotic spindle, chromosome capture (microtubule–kinetochore attachment), sister chromatid segregation, and cytokinesis. Thus ensuring mitosis proceeds in an ordered manner and completes successfully, protecting cells against genetic instability during cell division. The key events

that occur at each stage of mitosis and those molecular components involved as well as how they are regulated in animal cells, specifically by protein kinases, are discussed.

Stages of Mitosis

Prophase

Prophase is characterized morphologically by profound changes in the cell's architecture to detach from the substrate and become round in shape as well as condensation of chromatin into chromosomes. Cell rounding involves extensive rearrangement of the actin cytoskeleton, de-adhesion, and an increase in cortical rigidity ([Dao et al., 2009](#); [Kunda et al., 2008](#)). It is important for subsequent mitotic events such as spindle assembly and positioning as well as chromosome capture. During this time, the two sister chromatids from the replicated chromosomes are bound together at the centromere

by a protein complex known as cohesin. Time-lapse microscopy analysis of HeLa cells reveals that cell rounding is one of the earliest mitotic events in prophase occurring prior to centrosome separation and visible chromatin condensation followed by nuclear envelope breakdown (NEB) (Matthews *et al.*, 2012). Cell rounding as well as cortical stiffening is primarily driven by the small GTPase RhoA, which is activated upon mitotic entry (Maddox and Burridge, 2003; Matthews *et al.*, 2012) and concentrates at the cell cortex (Mali *et al.*, 2010; Matthews *et al.*, 2012). RhoA is regulated by the Cdk1/cyclin B1 protein kinase, which phosphorylates the RhoA regulators, the guanine exchange factor (GEF) Ect2 (Hara *et al.*, 2006; Matthews *et al.*, 2012; Niiya *et al.*, 2006), and the GTPase activating protein (GAP) p190RhoGAP (Maddox and Burridge, 2003). The result is a global increase in RhoA activity (Maddox and Burridge, 2003). Active RhoA mediates cell rounding and cortical stiffening through activation of its downstream effectors, Rho kinase (ROCK) and subsequent phosphorylation of myosin II in an analogous pathway for contractile ring assembly and function during the ingression phase of cytokinesis (Maddox and Burridge, 2003; Matthews *et al.*, 2012).

Prior to separation, centrosomes undergo maturation, which is executed by Plk1. Centrosomes consist of two perpendicularly orientated centrioles that are surrounded by pericentriolar material (PCM; Figure 1). Plk1 phosphorylates the PCM component, pericentrin (Lee and Rhee, 2011a), which subsequently recruits additional PCM components to coordinate microtubule nucleation for mitotic spindle assembly (Lee and Rhee, 2011b). Following centrosome maturation, the two centrosomes dissociate and move to opposite ends of the cell, forming the poles of the mitotic spindle (Meraldi and Nigg, 2002). At the spindle poles, nucleated microtubules originating from the centrosomes are organized into dense, radial arrays over the nuclear envelope. The end of prophase corresponds to NEB and again this is thought to be triggered by Cdk1/cyclin B1 (Lindqvist *et al.*, 2009) as well as Cdk1/cyclin A2 (Gong *et al.*, 2007).

Prometaphase

Prometaphase begins when the nuclear envelope disassembles, exposing nuclear structures and its contents to the cytoplasm. The nuclear envelope is composed of two membrane barriers, the inner nuclear membrane (INM) and the outer nuclear membrane (ONM). The nuclear envelope is stabilized by nuclear lamina (polymerization of lamin proteins) underlying the INM. NEB is thought to involve early spindle microtubules directly piercing the nuclear envelope causing folds and invaginations that create mechanical tension in the nuclear lamina (Beaudouin *et al.*, 2002; Georgatos *et al.*, 1997). Lamins are subsequently phosphorylated by Cdk1/cyclin B1 triggering their depolymerization, causing the nuclear envelope to vesiculate (Gallant and Nigg, 1992; Heald and McKeon, 1990; Peter *et al.*, 1990). Concurrently, nuclear pore complexes (NPCs) are disassembled by phosphorylation of the core components, nucleoporins, resulting in complete dissolve of the membrane permeability barrier (Terasaki *et al.*, 2001).

Following NEB, the mitotic spindle is assembled for alignment of chromosomes along the center of the cell

(metaphase plate) and subsequent equal segregation of sister chromatids (anaphase) into daughter cells (Helmke *et al.*, 2013; Rieder, 2005). The mitotic spindle is a bipolar structure composed of microtubules and associated motor proteins. Microtubules initially emanate from the spindle poles/centrosomes toward the cell equator and with their plus-ends attach to a large protein assembly on the centrometric chromatin of a chromosome called the kinetochore (Figure 3). This population of microtubules is termed kinetochore- or K-fibres. This attachment allows rapid movement of the bound chromosome such that the sister kinetochore attaches to a microtubule growing from the opposing centrosome. The result is a correctly bi-orientated chromosome at the metaphase plate with a stable microtubule–kinetochore attachment. Once all chromosomes have achieved this, the cell is considered to be in metaphase and chromosome segregation during anaphase can proceed. The mitotic spindle consists of two other populations of microtubules: (1) microtubules that do not attach to kinetochores, but also emanate toward the cell center and are called central spindle or non-kinetochore microtubules and (2) microtubules that emanate from the centrosomes circumferentially and anchor at the cell cortex are called astral microtubules, which function to maintain correct spindle orientation. Proteomics of purified mitotic spindles has identified approximately 800 spindle-associated proteins (Sauer *et al.*, 2005), indicating the breadth of signaling pathways required to establish and maintain this structure for correct chromosome attachment and alignment. Future investigation of these proteins in functional studies will aid in our understanding.

Metaphase

Metaphase is typified by the alignment of chromosomes at the center of cell, commonly referred to as the metaphase plate. The chromosomes are now highly coiled and condensed. The kinetochore microtubules from opposing spindle poles capture each pair of sister chromatids at their kinetochores so that the chromosomes are attached in a bi-orientated manner. During this capture process, the chromosomes oscillate gently at the metaphase plate due to the pulling forces and tension across sister kinetochores that are generated by the kinetochore microtubules. The signal for chromosome separation/segregation is governed by the spindle assembly checkpoint (SAC; Box 1; Figure 4). Premature or aberrant chromosome segregation results in aneuploidy, and is avoided by activity of the SAC. This signaling pathway consists of a number of protein complexes that monitor mitotic spindle assembly, delaying anaphase onset until all chromosomes have formed stable bipolar attachments to the spindle via kinetochores (Musacchio and Salmon, 2007).

Astral microtubules play a key role during metaphase by regulating spindle stability and orientation. Spindle orientation is achieved through signaling pathways that link the plus-ends of astral microtubules to the cell cortex (Kotak and Gonczy, 2013). The core machinery necessary for spindle orientation has been identified in most cell contexts (*Caenorhabditis elegans*, *Drosophila*, and cultured mammalian cells) and involves four components (Bergstrahl *et al.*, 2013; Kotak and Gonczy, 2013). In human cells, this includes the

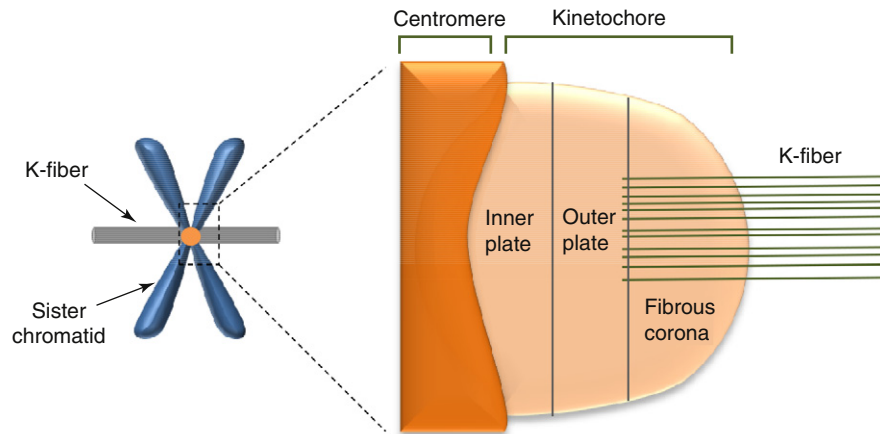


Figure 3 The centromere–kinetochore association. The centromere is composed of borealin, survivin, Aurora B, inner centromere protein (INCENP), and mitotic centromere-associated kinesin (MCAK), which function to regulate stability of microtubule–kinetochore attachments. The kinetochore is a protein matrix attached to the centromere and is composed of several layers (Musacchio and Salmon, 2007). The inner plate is in contact with the centromere (or centromeric DNA) and is marked by the specialized histone variant CENP-A (which substitutes histone H3 in this region) as well as additional auxiliary proteins. Adjacent to the inner plate is the outer plate, which consists primarily of proteins that assemble at the kinetochore upon NEB. Evolutionarily conserved proteins such as Spc105, MIS12, MSM21, and Hdc80 occupy this space to form attachment sites for bundles of microtubules known as K-fibers. In animal cells, kinetochores contain approximately 20 anchoring sites for the plus-ends of microtubules. The outer most region of the kinetochore is the fibrous corona and consists of proteins that are highly dynamic in their concentration. These proteins are either high in their concentration at the kinetochore before (molecular motors CENP-E and dynein as well as their target components ZW10 and ROD, and the spindle checkpoint proteins such as Mad1, Mad2, BubR1, and Cdc20) or after (EB1, APC, and proteins in the Ran pathway such as RanGap1 and RanBP2) K-fiber capture.

motor protein dynein, which is cortically anchored by an evolutionary conserved ternary protein complex consisting of G-protein signaling modulator-2 (LGN), nuclear mitotic apparatus (NuMA), and Gai (Kotak and Gonczy, 2013). Gai binds to the plasma membrane and this may involve the GEF, Ric8 (Woodard *et al.*, 2010). GDP-bound Gai recruits and anchors LGN at the cell cortex by binding to its C-terminal GoLoCo motifs. The N-terminal TPR domains of LGN forms a docking platform for NuMA. NuMA associates with minus-end-directed microtubule motor, dynein via its N-terminus. The LGN-dynein interaction is specific to the cortically localized dynein as the depletion of LGN does not disrupt other pools of dynein on kinetochore or spindle microtubules (Kiyomitsu and Cheeseman, 2013). Collectively, this complex functions in spindle-capture as well as in force generation, whereby it is hypothesized that dynein attempts to move along astral microtubules toward the spindle poles, but instead pulls on the astral microtubules, generating tension (Hendricks *et al.*, 2012; Laan *et al.*, 2012). Disruption of the ternary complex alters cortical dynein levels and thus spindle positioning (Kotak *et al.*, 2012; Zheng *et al.*, 2013).

Anaphase

The separation of sister chromatids marks the commencement of anaphase followed by the movement of chromosomes toward the spindle poles. The chromosomes are pulled apart by shortening of the kinetochore microtubules at the centromere attachment sites, and this is thought to be triggered by the loss of tension following cleavage of the cohesin complex. Simultaneously, the non-kinetochore spindle microtubules elongate, thereby pushing the centrosomes farther apart. This is regulated by two forces: (1) plus-end-directed motor

proteins, such as end-binding protein-1 (EB1), at the central spindle and between the overlapping microtubules from both poles (Tamura and Draviam, 2012). These motor proteins slide the microtubules at the plus-ends past one another, creating the forces to push the spindle poles farther apart. (2) Minus-end directed motor proteins, such as dynein, that interact with the cell cortex and the astral microtubules to create pulling forces on the spindle poles (Tamura and Draviam, 2012). By the end of anaphase, the kinetochore microtubules have disassembled and a complete set of chromosomes has assembled at each pole and begun to decondense. During this time, cyclin B1 is degraded, resulting in the inactivation of Cdk1 and this is necessary for driving mitotic exit (Hershko, 1999).

Telophase

During telophase, the cell continues to elongate due to lengthening of the microtubules. The chromosomes at both poles begin to uncoil and the nuclear membrane vesicles (disassembled in prophase) reform the nuclear envelope around each set of chromosomes, forming two daughter interphase nuclei. The membrane at the cell equator also begins to fold into the cell, forming an ingression furrow that results in compression of the central spindle.

Cytokinesis

Cytokinesis is the final stage of cell division that generates two independent daughter cells and is often the longest stage of cell division. Cytokinesis in animal cells can be subdivided broadly into two steps: (1) membrane ingression and (2) membrane abscission (Figure 2; Glotzer, 2005). These steps

Box 1 The spindle assembly checkpoint

The 'spindle assembly checkpoint' or SAC is a term used to describe the complexes of proteins and signaling pathways that prevent anaphase onset until all chromosomes have achieved bi-oriented attachment via their kinetochores to the mitotic spindle and are aligned at the metaphase plate (**Figure 4**). As such, the SAC is activated upon mitotic entry due to the presence of unattached kinetochores. In addition to microtubule–kinetochore attachment, tension is also important for SAC inactivation (*Nicklas et al., 1995*). Microtubule–kinetochore tension is initially low but high upon bi-orient attachment resulting in an increase in kinetochore-to-kinetochore distance between sister chromatids, which ultimately aids in their segregation during anaphase. Incorrectly attached chromosomes have low tension and this is detected by Aurora B (Ipl1 in *Saccharomyces cerevisiae*) kinase (*Tanaka et al., 2002*). Cells that have an impaired SAC enter anaphase prematurely and are genomically unstable.

The core SAC signaling components reside at the kinetochore (**Figure 3**). Many of these proteins localize to unattached kinetochores in prometaphase and are removed immediately following attachment of microtubule plus-ends, indicating that events at the kinetochores contribute to SAC activity (*Lara-Gonzalez et al., 2012*). This decrease in protein localization at kinetochores as more attachment sites bind to K-fibres is called occupancy. For example, MAD2 kinetochore localization is reduced at metaphase kinetochores by 50–100-fold compared with unattached prometaphase kinetochores. The SAC targets CDC20, a cofactor of the ubiquitin ligase anaphase-promoting complex/cyclosome (APC/C) (*Hwang et al., 1998; Kim et al., 1998*). Specifically, the SAC negatively regulates the ability of CDC20 to activate APC/C-mediated polyubiquitylation of two key substrates, cyclin B1 and securin, thereby preventing their destruction by the 26S proteasome. Separase cleaves the cohesin complex that holds sister chromatids together, which is required to execute anaphase, whilst cyclin B1 proteolysis inactivates the mitotic kinase, Cdk1, which promotes exit from mitosis (*Peters, 2006*). A mitotic checkpoint complex (MCC) is a SAC effector that contains three SAC proteins, MAD2, the protein kinase BUBR1/Mad3, and BUB3, as well as CDC20 itself (*Sudakin et al., 2001*). The MCC binds the APC/C rendering it inactive. Other 'core' SAC components include MAD1 and the kinases BUB1, multipolar spindle-1 (MPS1), and Aurora B. These proteins are required to amplify the SAC signal and the rate of MCC formation (*Morrow et al., 2005*). Additional proteins that regulate SAC activity in higher eukaryotes include the ROD (rough deal)–ZW10 (zeste white-10)–ZWILCH (RZZ) complex, p31^{comet}, several protein kinases including mitogen activated protein kinase (MAPK), Cdk1/cyclin B1, NEK2, and PIK1; the MT motors centromere protein (CENP)-E (also known as kinesin-7) and dynein; and dynein-associated proteins such as dynactin, cytoplasmic linker protein (CLIP)170, and lissencephaly-1 (LIS1) (*Musacchio and Salmon, 2007*).

are highly ordered processes that begin immediately following chromosome segregation (anaphase) and need to occur in a sequential manner for cytokinesis to be executed correctly. Cytokinesis failure can occur at any one of the steps in this process due to an array of errors such as protein mislocalization, deregulated protein phosphorylation, and erroneous signaling. Depending on the stage of cytokinesis that the error occurs, membrane ingression may not be initiated, or if

partially completed will regress or if the midbody has formed then the nascent daughter cells may remain connected due to persistent bridge formation. Nevertheless, the resultant outcome is the generation of a binucleated (or aneuploid cell), which can lead to genomic instability and thus contribute to the initiation or progression of tumorigenesis.

Membrane ingression

Membrane ingression involves two phases. Firstly, the site of division is determined by recruitment of RhoA. Secondly, the membrane ingresses, which requires the assembly and activity of the actin–myosin II contractile ring (**Figure 5; Glotzer, 2005**).

Immediately after chromosome segregation, cells must ensure that the spindle poles are on opposing sides of the division plane so that the segregated chromosomes are equally distributed into each of the two newly forming cells. In animal cells, the division plane is determined at the metaphase to anaphase transition and its position is dictated in part by signaling pathways mediated by three separate pools of microtubules that collectively form the mitotic spindle (*Burgess and Chang, 2005; Glotzer, 2004*): (1) polar astral microtubules (emanating from the spindle pole to sites away from the ingression site – most likely aid in positioning of the ingression site by inhibiting cortical contraction (*Danowski, 1989; Pletjushkina et al., 2001; Ren et al., 1999*) by spatially regulating myosin recruitment) (*Canman and Bement, 1997; Werner et al., 2007*), (2) equatorial astral microtubules (emanating from the spindle poles toward the ingression site – are thought to contribute to stabilization of the equatorial cortical region (*Canman et al., 2003*), as well as provide positive signals that stimulate formation and contraction for membrane ingression) (*Rappaport, 1986*), and (3) central spindle (a set of antiparallel microtubule bundles between the spindle poles – deliver positive signals to the midzone region, which are especially important for later cytokinetic stages). The roles of each microtubule population are partly redundant but collectively they ensure that a robust division site is established (*Bringmann and Hyman, 2005; Dechant and Glotzer, 2003*).

The key molecule that defines the site to initiate formation of the ingressing furrow in animal cells is active RhoA (*Drechsel et al., 1997; Jantsch-Plunger et al., 2000; Kamijo et al., 2006; Kishi et al., 1993; Piekny et al., 2005*). Its activation is spatially restricted to the equatorial cell cortex as disruption of this restricted spatial regulation results in a wider zone of activation and failure to initiate membrane ingression (*Yuce et al., 2005*). Central spindle microtubules are critical in assembling a signaling network to activate RhoA in a restricted spatially narrow zone at the equatorial cell cortex upon anaphase onset (*Bement et al., 2005; Nishimura and Yonemura, 2006; Yonemura et al., 2004; Yoshizaki et al., 2003; Yuce et al., 2005*). The central spindle components include protein regulator of cytokinesis-1 (PRC1), kinesin-4 family member (KIF4), the centralspindlin complex (composed of kinesin-6 family member MKLP1 and the Rho GTPase activating protein MgcRacGAP) and the chromosomal passenger complex (CPC, composed of Aurora B kinase and its regulatory subunits inner centromere protein INCENP, Survivin, and Borealin) (*Gassmann et al., 2004; Kurasawa et al., 2004;*

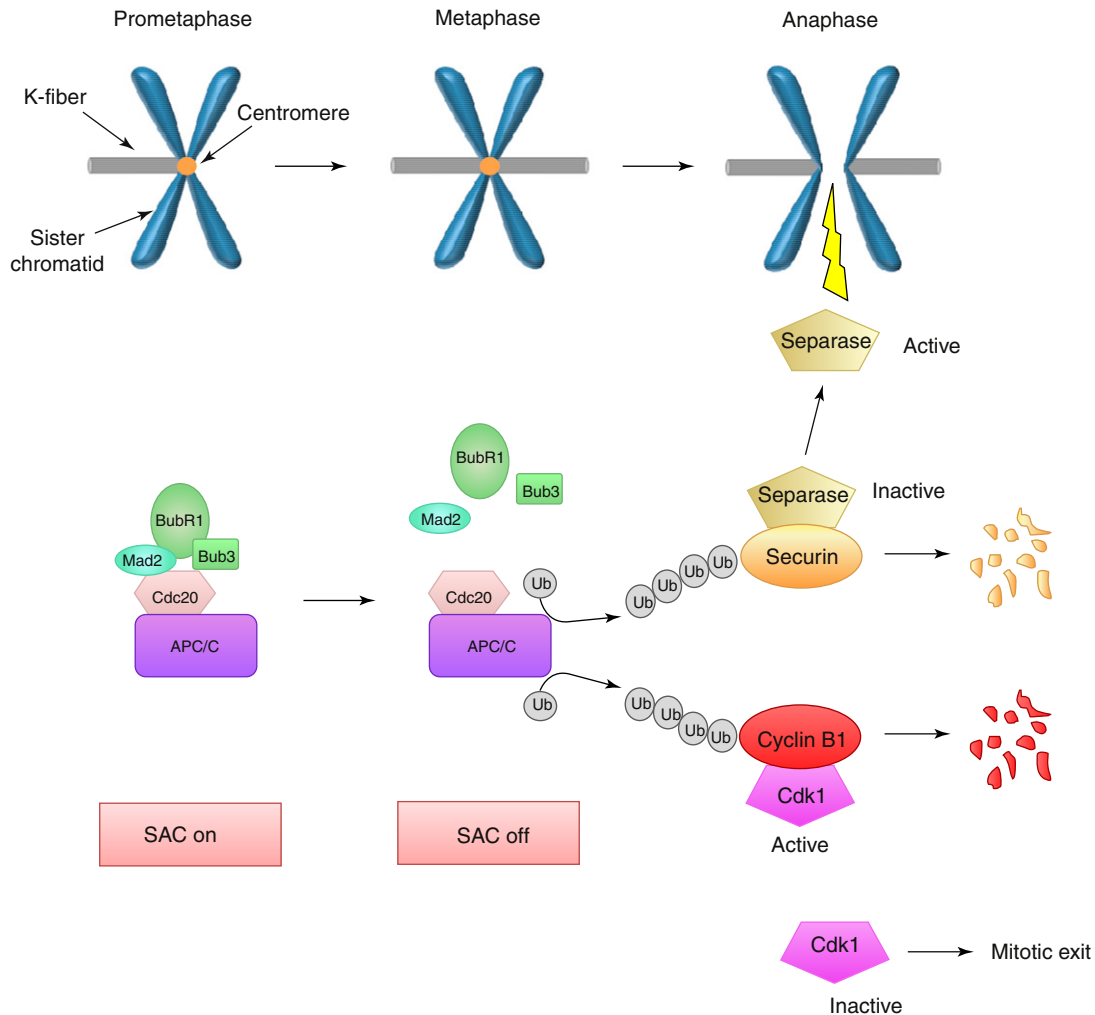


Figure 4 SAC signaling and inhibition of the APC/C. During prometaphase, the SAC is activated due to the presence of unattached kinetochores. Mad2 is activated together with BubR1 and Bub3 bind to Cdc20, thereby inhibiting the activity of the APC/C. In metaphase, following correct bi-orient attachment of all chromosomes to the kinetochore microtubules and alignment at the metaphase plate, the SAC is deactivated. This involves release of Mad2, BubR1, and Bub3 from Cdc20, allowing it to activate the APC/C. APC/C subsequently mediates the polyubiquitination of many substrates leading to their degradation. The two key substrates include cyclin B1, to deactivate Cdk1 for mitotic exit, and securin, which releases separase to cleave the cohesin complex that binds sister chromatids. Sister chromatids are now no longer tethered together, leading to their separation during anaphase.

Mishima *et al.*, 2002, 2004; Mollinari *et al.*, 2002; Severson *et al.*, 2000; van der Waal *et al.*, 2012; Wheatley *et al.*, 2001; Zhu and Jiang, 2005). PRC1 is localized at the central spindle, where it interacts and recruits KIF4 (Kurasawa *et al.*, 2004; Zhu and Jiang, 2005). The CPC then phosphorylates MKLP1 and promotes the recruitment of the centralspindlin complex to the central spindle (Douglas *et al.*, 2010; Guse *et al.*, 2005). The centralspindlin component MgcRacGAP together with the RhoGEF, Ect2, contribute to the equatorial activation of RhoA (Kamijo *et al.*, 2006), which subsequently drives assembly and activity of the contractile ring (Figure 5).

The contractile ring drives membrane ingression and invagination of the cell membrane. It is composed of parallel filaments of actin and non-muscle myosin II (Maupin and Pollard, 1986). Myosin comprises two heavy chains, two essential light chains and two regulatory myosin light chains

(MLC). Phosphorylation of MLC on T18 and S19 releases myosin II from an autoinhibitory state, allowing myosin II to assemble into filaments (Ikebe *et al.*, 1988) and subsequently triggering activity of the myosin II motor (Komatsu *et al.*, 2000; Yamakita *et al.*, 1994). Myosin filaments then interact with actin filaments, where it is thought to either cross-link actin filaments or accelerate actin disassembly and together, the contractile ring constricts and the membrane ingresses (Guha *et al.*, 2005; Ma *et al.*, 2012b; Murthy and Wadsworth, 2005; Scholey *et al.*, 1980). The membrane continues to ingress as the contractile ring constricts to its fullest extent, whereby the antiparallel microtubules forming the mitotic spindle (midzone) and associated proteins are compressed to form an intracellular bridge that is approximately 1–1.5 μm in diameter between nascent daughter cells (Eggert *et al.*, 2006). The midzone plays a key role in maintaining the segregated

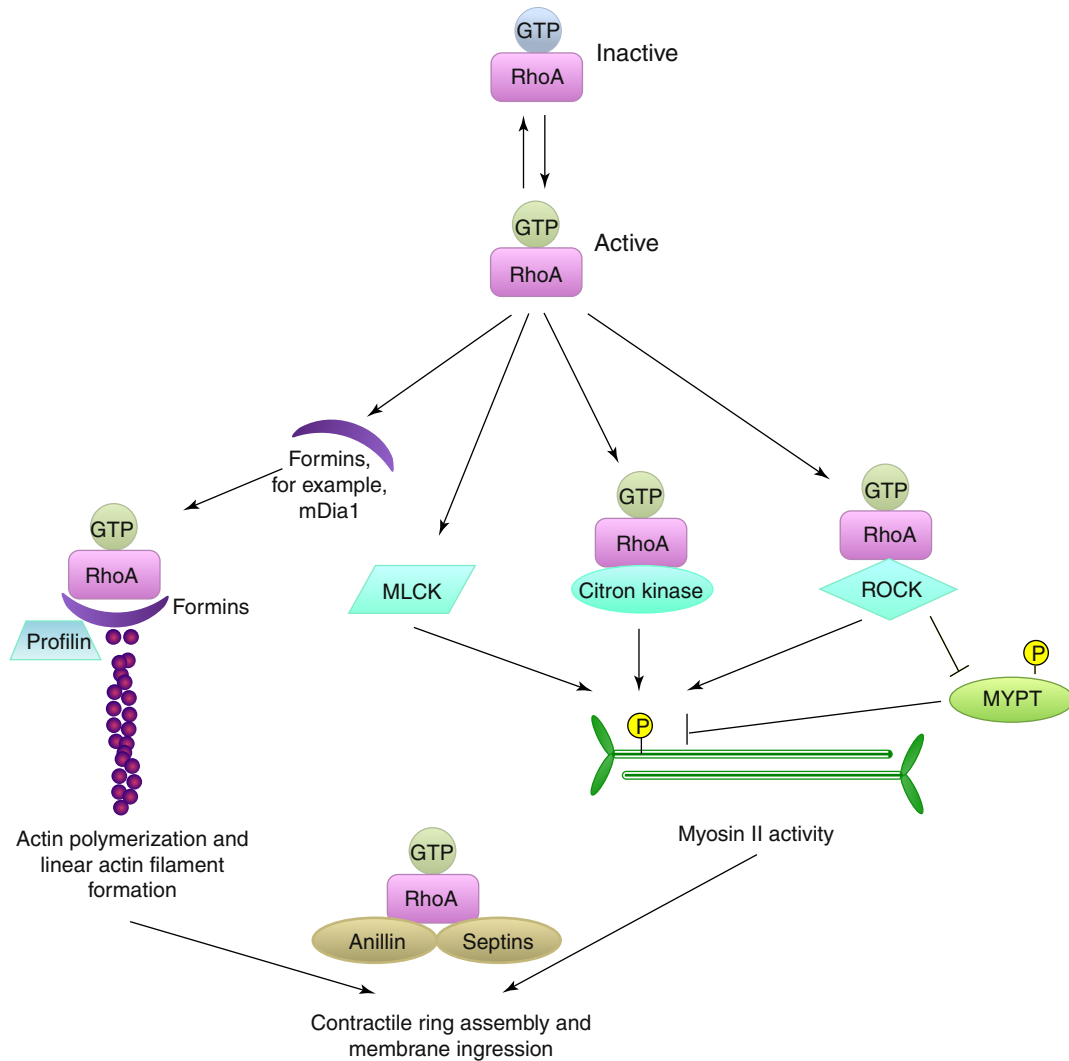


Figure 5 Membrane ingression is driven by RhoA-mediated signaling pathways. Upon anaphase onset, the site of the division plane is established and RhoA is activated within a narrow zone of the cell cortex. Here, RhoA activates the effectors, mDia1 for unbranched linear actin filament polymerization (Li and Higgs, 2003), as well as the kinases, myosin regulatory light chain kinase (MLCK), ROCK, and citron kinase, to activate myosin II. Myosin II activity is also regulated in part by myosin phosphatase (MYPT1), which down-regulates myosin II contractility by dephosphorylating S19 on MLC. ROCK (Yamashiro *et al.*, 2003), also acts to maintain active myosin by phosphorylating T853 on a subunit of MYPT1, thereby inhibiting the phosphatase activity of MYPT1 by causing dissociation of its catalytic subunit (Kawano *et al.*, 1999; Yokoyama *et al.*, 2005). Collectively, the outcome is assembly and activation of the actin–myosin II contractile ring to mediate membrane ingression. RhoA also interacts with the scaffold proteins anillin and septins, which function to maintain all furrow components within a narrow zone for efficient and complete membrane ingression to ensure a robust cytokinesis.

chromosomes separated during cytokinesis as depolymerization of microtubules results in the newly forming nuclei returning to the center of the cell and fusing (Straight *et al.*, 2003).

Membrane abscission

Membrane abscission also involves two phases: (1) transition from contractile ring closure to formation and maturation of a midbody ring (MR) at the center of the intracellular bridge. This is followed by (2) a secondary ingression or thinning of the intracellular bridge to approximately 100 nm at a site on one side of the MR. This is thought to be the site of abscission, which subsequently follows this secondary ingression

(Figure 6; Elia *et al.*, 2011; Guizetti *et al.*, 2011a; Schiel *et al.*, 2011). In HeLa cells, this process takes approximately 60 mins with daughter cells remaining connected for approximately 40 mins until secondary ingression is initiated and abscission physically separates the two cells (Steigemann and Gerlich, 2009a).

The MR results from compression of the midzone and consists of several proteins (Skop *et al.*, 2004), such as the core structural component γ -tubulin (Julian *et al.*, 1993; Shu *et al.*, 1995), centriolin (Gromley *et al.*, 2003), and the exocyst complex (Gromley *et al.*, 2005). Transition from contractile ring closure (membrane ingression) to MR formation involves anillin (Kechad *et al.*, 2012). Specifically, it is the removal and

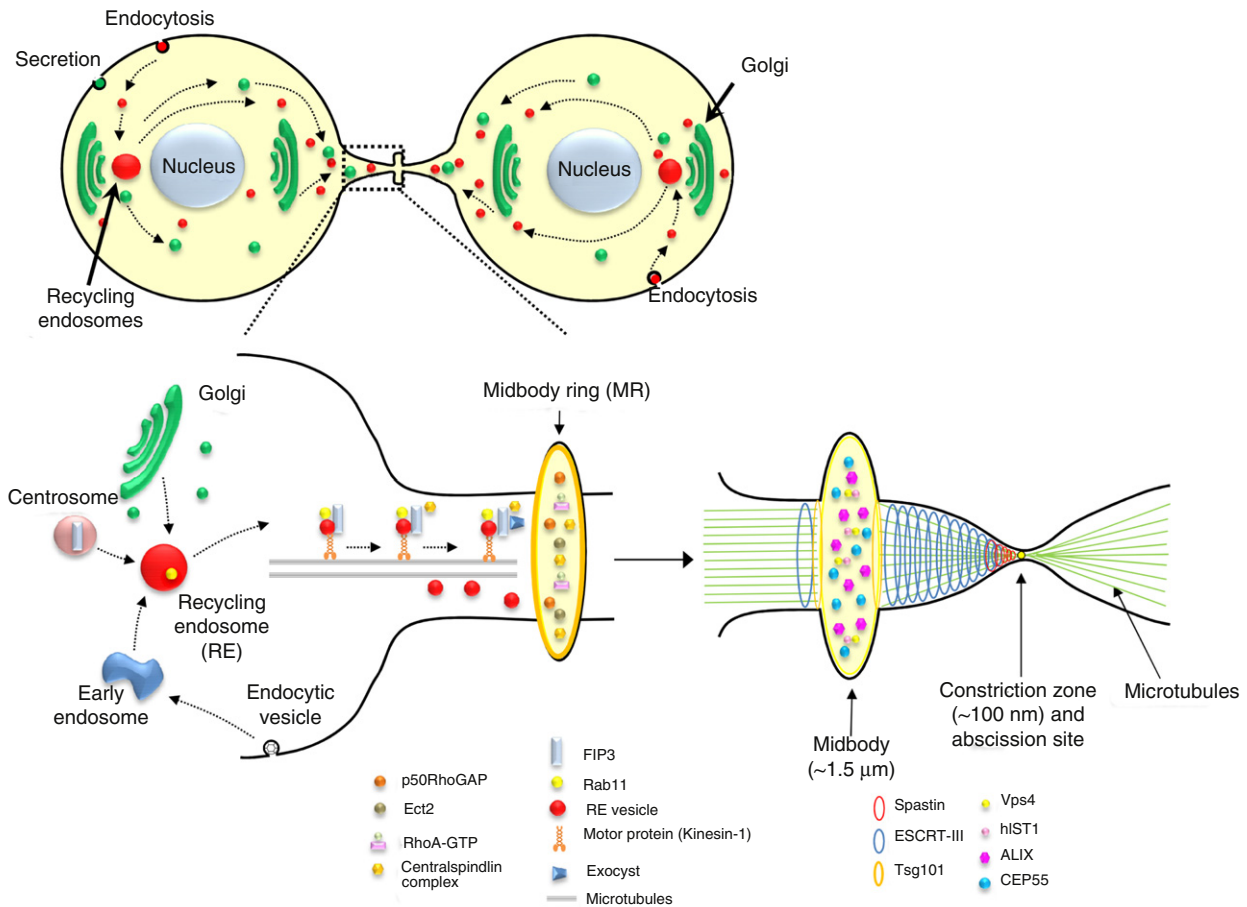


Figure 6 Membrane trafficking pathways involved in the abscission phase of cytokinesis. The abscission phase of cytokinesis commences upon formation of the intracellular bridge. In HeLa cells, this phase takes approximately 60 mins to achieve membrane abscission and daughter cell separation. The first approximate 40 mins of the abscission phase is characterized by extensive membrane trafficking events whereby vesicles derived from internal membrane stores including the Golgi, lysosomes, and early and recycling endosomes (Chen *et al.*, 2006; Fielding *et al.*, 2005; Goss and Toomre, 2008; Gromley *et al.*, 2005; Pohl and Jentsch, 2008; Schweitzer and D'Souza-Schorey, 2005; Wilson *et al.*, 2005), are recruited from both daughter cells to the intracellular bridge (Goss and Toomre, 2008). Many proteins involved in vesicle trafficking and membrane fusion locate within the intracellular bridge and specifically dock at the MR, such as Rab11 and SNARE. ALIX and TSG101 (ESCRT-I) accumulate at the midbody and recruit ESCRT-III to the intracellular bridge (Carlton *et al.*, 2008; Elia *et al.*, 2011; Morita *et al.*, 2007). The final approximate 20 mins is marked by constriction of the membrane at a secondary ingression site adjacent to the MR followed by abscission. ESCRT-III components have been associated with recruiting the microtubule severing enzyme spastin, which appears to be required for shearing the microtubules during abscission (Connell *et al.*, 2009; Elia *et al.*, 2011), and this most likely occurs after completion of the secondary ingression but prior to membrane abscission. There are 11 ESCRT-III proteins in mammals and are often called charged MVB proteins (CHMPs). CHMP4B and other ESCRT-III components extend toward the constriction zone while Cep55, TSG101, and ALIX remain confined to the midbody (Guizetti *et al.*, 2011b). ESCRT-III proteins polymerize to form 17 nm filaments within the intracellular bridge that spans a distance of 1 μm from the MR toward the constriction zone. The ESCRT disassembly factor, vacuolar protein sorting-4 (Vps4), is then recruited to the constriction zone and abscission follows immediately (Elia *et al.*, 2011). Membrane abscission is thought to occur due to the inward curving properties of the ESCRT-III filaments or the ability of Vps4 to remodel the cytoskeletal membrane at the constriction site.

retention of anillin at the closing contractile ring and MR, respectively that corresponds with this transition (El *et al.*, 2013). The MR is not contractile and is unable to account for abscission on its own. It appears to act as a scaffold platform for recruiting or activating signaling proteins that contribute to the regulation of abscission and post-mitotic cell fate (Ettinger *et al.*, 2011; Green *et al.*, 2013; Kuo *et al.*, 2011). Proteomics and cell biology studies have indicated that a large number of proteins (>100) locate to the MR itself or to flanking regions within the intracellular bridge (Skop *et al.*, 2004; Steigemann and Gerlich, 2009a). The molecular details of the role of these

proteins, and the interaction between pathways involved in the abscission process, are only starting to emerge. Proteins that accumulate at the MR and flanking regions within the intracellular bridge include those derived from (1) the central spindle such as kinesin-like motor proteins and chromosomal passenger proteins, that travel along the central spindle toward microtubule plus-ends, accumulating in the midzone, (2) the cell equator such as MgcRacGAP, Ect2, and anillin, (3) the equatorial cortex such as RhoA and sept2, and (4) proteins previously located within the nascent daughter cells that traffic to the intracellular bridge upon formation such as

Rab11 and endosomal sorting complexes required for transport (ESCRT)-III. Thus, in human cell lines several lines of evidence suggest that recruitment of key abscission proteins to the midbody are dependent on microtubules and protein trafficking. Indeed, in mammalian cells, vesicles derived from internal membrane stores including the Golgi, lysosomes, and early and recycling endosomes (Chen *et al.*, 2006; Fielding *et al.*, 2005; Goss and Toomre, 2008; Gromley *et al.*, 2005; Pohl and Jentsch, 2008; Schweitzer and D'Souza-Schorey, 2005; Wilson *et al.*, 2005), are recruited from both daughter cells to the intracellular bridge (Goss and Toomre, 2008) and many proteins involved in vesicle trafficking and membrane fusion are required for cytokinesis completion, including the soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor (SNARE) proteins, centriolin, exocyst components, dynamin II, and sorting nexin-9 (Chircop *et al.*, 2010; Echard *et al.*, 2004; Gromley *et al.*, 2005; Ma and Chircop, 2012a; Neto *et al.*, 2013a,b; Skop *et al.*, 2004; Smith and Chircop, 2012).

The secondary ingression and abscission process involves depolymerization of cortical actin located within the intracellular bridge (Schiel *et al.*, 2012), which appears to require inactivation of RhoA (Bement *et al.*, 2005; Yuce *et al.*, 2005) as well as depolymerization of microtubules (Elia *et al.*, 2011; Guizetti *et al.*, 2011a; Schiel *et al.*, 2011), trafficking of Rab11-FIP3 recycling endosomes to the MR (Fielding *et al.*, 2005; Schiel *et al.*, 2011, 2012; Wilson *et al.*, 2005), and subsequent recruitment of the ESCRT complexes for abscission (Carlton *et al.*, 2008; Morita *et al.*, 2007). The ESCRT machinery consists of four protein complexes (ESCRT-0, -I, -II, -III) that function in a sequential order and are involved in intraluminal vesicle formation and fission during viral budding and to sort membrane proteins into the endosome lumen to create multivesicular bodies (Raiborg and Stenmark, 2009). In an analogous manner, ESCRT-III proteins polymerize to form 17 nm filaments within the intracellular bridge, spiraling toward the secondary ingression site and driving membrane scission (Figure 6; Guizetti *et al.*, 2011a). After abscission the MR is inherited by one of the daughter cells (Elia *et al.*, 2011; Guizetti *et al.*, 2011a; Pohl and Jentsch, 2009; Schiel *et al.*, 2011).

Although the sequential targeting of the ESCRT machinery is now widely accepted as the final events leading to abscission, the mechanisms regulating this process are still poorly understood. It is proposed that septin filaments drive elongation of the intercellular bridge and generates the secondary ingression site, priming the intracellular bridge for ESCRT-III assembly and abscission (Renshaw *et al.*, 2014). Furthermore, the daughter cells are held together by a high tension pulling force due to cell contractility. This is alleviated when cells stop moving apart and this loss of tension leads to the assembly of the ESCRT-III components at the abscission site, which is shortly followed by abscission itself (Lafaurie-Janvore *et al.*, 2013). Aurora B contributes to regulating abscission timing whereby its activity declines gradually during cytokinesis leading to its inactivation within the intracellular bridge, which subsequently triggers abscission (Steigemann *et al.*, 2009b). Aurora B activity is maintained in cells that contain lagging chromosomes trapped in the intracellular bridge (Mendoza *et al.*, 2009; Steigemann *et al.*, 2009b).

This involves Aurora B-mediated phosphorylation of the ESCRT-III component CHMP4C at S210, which presumably prevents assembly of the abscission complex (Carlton *et al.*, 2012). Thus, Aurora B is a key regulatory kinase that functions to protect the newly formed daughter cells from acquiring DNA damage by preventing premature abscission.

Mitosis As an Anticancer Drug Target

Cancer is characterized by excessive proliferation of cells. This is often due to underlying genetic mutations that result in misregulation of components of the cell division machinery (Chi and Jeang, 2007; Kops *et al.*, 2005; Nicholson and Cimini, 2011; Vogelstein and Kinzler, 2004). Thus causing errors that increase chromosome instability and oncogenic potential, which can subsequently drive tumorigenesis. Therefore it is not surprising that aneuploidy arising from unstable tetraploid cells (Ganem *et al.*, 2007) is a common histopathological hallmark of many cancers, such as the presence of large multinucleated cells observed in glioblastoma multiforme. These tetraploid cells are primarily caused by cell division errors arising from mitotic slippage (premature mitotic exit after prolonged SAC activation), chromosome mis-segregation, and cytokinesis failure.

One of the most successful classes of anticancer chemotherapeutic drugs currently used in the clinic targets mitosis, causing cell death following its disruption (Jackson *et al.*, 2007; Schmidt and Bastians, 2007). These drugs, such as the Vinca alkaloids, target tubulin, by interfering with microtubule dynamics and thus disrupt the mitotic spindle, resulting in prolonged SAC activation that leads to mitotic slippage and cell death. However, microtubules also play important roles in nondividing and noncancer cells. Thus, anti-microtubule drugs have severe side effects in patients such as peripheral neuropathy. Current research is focused on discovering new targets to develop inhibitors of proteins that play critical and exclusive roles in mitosis such as Cdk, Aurora, and Polo kinases as well as kinesin spindle proteins (KSPs) (Jackson *et al.*, 2007; Schmidt and Bastians, 2007). This new targeted class of antimitotic agents are predicted to be as efficacious as anti-microtubule agents, while avoiding many of the adverse side effects. Indeed, many inhibitors of these mitotic proteins have been developed and are currently being tested in clinical trials of various different types of human cancer with promising success (Table 1). Consistent with the mitotic function of these proteins, these inhibitors either inhibit or activate the SAC or cause cytokinesis failure and have been shown to subsequently induce cell death and reduce tumor volume in *in vivo* mouse models of cancer. However, cell death is not always the result as different cancer cell types have varied responses such as senescence and continued cell proliferation. Both of which are undesirable as it would not result in tumor regression but potentially tumor progression. Therefore, a thorough understanding of the molecular mechanisms that control mitotic progression as well as the mechanisms governing cell fate, specifically how cell death pathways are activated, will provide invaluable knowledge to translate new antimitotic compounds into the clinic as improved cancer therapies.

Table 1 Targeted antimetabolic compounds and their effects in cell culture, animal models, and clinical trials

Compound	Cultured cells	Mouse models	Clinical trials	Notes	References
<i>Aurora A inhibitors</i>					
ENMD-2076	Anti-proliferative Myeloma cells	Xenograft tumors H929 tumor Human HT-29 colorectal cancer	Phase I/II Relapsed or refractory leukemia	2010:Orphan drug designation for treatment of acute myeloid leukemia (AML)	Cheung <i>et al.</i> (2011)
MLN8054	Reduced cell viability and promotes apoptosis in p53 deficient cells	Antitumor activity	Phase I Advanced solid tumor, no anti-proliferative affect	Aurora A ATP competitive Affected by Aurora A T217D mutation	Cheung <i>et al.</i> (2011)
MLN8237	Anti-proliferative HCT-116, PC3, SKOV-3, LY-3, MM, and acute lymphoblastic leukemia (ALL)	Xenograft tumors Neuroblastoma and ALL	Phase I/II	Phase I/II	Cheung <i>et al.</i> (2011)
VX-689/MK-5108	Anti-proliferative MCF-7, AU565, SW48, Colo205, SKOV-3, T47D, HCC1419, and BT474 cancer cell lines	Synergistic antitumor activity	Phase I/II CML and ALL	Phase I/II	Cheung <i>et al.</i> (2011)
<i>Pan-Aurora Inhibitors</i>					
AMG-900	Anti-proliferative Effective against MDR1 cells	Xenograft tumors Human HCT-116 tumor and NCI-H460-PTX (MDR over expressing) tumor models	Discontinued	Inhibits auto-phosphorylation of Aurora A and B	Cheung <i>et al.</i> (2011)
AT9283	Anti-proliferative Targets imatinib-resistant BCR-ABL cancer cells		Phase I Refractory solid tumors	Targets Aurora A/B# Orphan drug designation for AML treatment USA and Europe	Cheung <i>et al.</i> (2011)
CYC116	Anti-proliferative	P388/0 murine leukemia mouse model	Phase I Discontinued unknown reason	Aurora A and B	Cheung <i>et al.</i> (2011)
GSK1070916	Anti-proliferative Colo205 cells	Xenograft tumors Lung, colon, and breast cancers	Phase I	Inhibits Aurora B/C-INCENP interaction ATP competitive	Cheung <i>et al.</i> (2011)
PHA-680632	Anti-proliferative	Xenograft tumors HL-60 cells, A2780, HCT-116, and breast cancer models	Phase I Disease stabilization	Targets other kinases	Kitzen <i>et al.</i> (2010)
R763/AS703569	Anti-proliferative A549, DU145, HeLa, Colo205, and U2OS	Xenograft tumors Human pancreatic and leukemia cancers		Targets other kinases	Cheung <i>et al.</i> (2011)
VX680/MK-0457	Anti-proliferative	Xenograft tumors Human AML, colorectal and pancreatic cancer	Phase I/II Suspended		Cheung <i>et al.</i> (2011)
XL228			Phase I	Targets other kinases	Cheung <i>et al.</i> (2011)

(Continued)

Table 1 Continued

Compound	Cultured cells	Mouse models	Clinical trials	Notes	References
<i>Aurora B inhibitors</i> AZD1152	Anti-proliferative T-cell leukemia Hepatocellular carcinoma Breast cancer Irradiated cells Anti-proliferative	Xenograft tumor Human colon cancer Irradiated mice	Phase I Solid malignant tumors showed no response Target of MDR1		Cheung <i>et al.</i> (2011)
BI 811283		Xenograft tumor Human Non-small cell lung cancer (NSCLC) Xenograft tumors	Phase I Advanced solid tumors		Cheung <i>et al.</i> (2011)
PF-03814735			Phase I Disease stabilization Phase I/II Advanced solid tumors Malignant solid tumors Disease stabilization	Inhibits Aurora A at high concentrations ATP competitive	Cheung <i>et al.</i> (2011) and Kitzen <i>et al.</i> (2010)
PHA-739358	Anti-proliferative	Xenograft tumor Huh-7			Cheung <i>et al.</i> (2011)
<i>APC/C Inhibitor</i> TAME	Induced SAC dependent mitotic arrest			Prevents activation by Cdc20 or Cdh1	Zeng <i>et al.</i> (2010)
<i>CENP-E Inhibitors</i> GSK923295A	Anti-proliferative	Antitumor in Ewing sarcoma, rhabdoid, and rhabdomyosarcoma xenografts	Phase I Solid tumors Disease stabilization		Chung <i>et al.</i> (2012), Janssen and Medema (2011), and Lock <i>et al.</i> (2012)
Lonafarnib	Anti-proliferative	Spontaneous cancers in mouse models	Phase I Phase II – No response in head and neck carcinomas	Also inhibits CENP-F	Hanrahan <i>et al.</i> (2009) and Janssen and Medema (2011)
<i>Eg5/KSP inhibitors</i> Ispinesib	Anti-proliferative	Xenograft models	Phase I/II Metastatic breast cancers Stabilization of disease		Jackson <i>et al.</i> (2007)
MK-0731 SB-743921	Anti-proliferative Anti-proliferative	Xenograft models Xenograft models	Phase I Phase I Cholangiocarcinoma		Jackson <i>et al.</i> (2007) Jackson <i>et al.</i> (2007)
<i>Msp1 inhibitors</i> NMS-P715	Anti-proliferative	Xenograft tumors Nude mice; established A2780 tumors or A375 melanoma tumors			Colombo <i>et al.</i> (2010)

Msp-IN-1 and Msp1-IN-2	Premature mitotic exit and gross aneuploidy leading to cell death	Kwiatkowski <i>et al.</i> (2010)
<i>Plk Inhibitors</i> BI2536	Anti-proliferative	Jackson <i>et al.</i> (2007) and Schoffski (2009)
GSK-461364		Phase I/II Head and neck cancers NSCLC platinum failures HRPC Disease stabilization Phase I Solid tumors and NHL ongoing
HMMN-241		Lens <i>et al.</i> (2010) and Schoffski (2009)
NMS-1286937		Phase I Advanced solid tumors Disease stabilization
ON01910	Anti-proliferative	Lens <i>et al.</i> (2010)
		Phase I Advanced solid tumors and hematological cancers
		Phase I Advanced solid tumors Disease stabilization
		Jackson <i>et al.</i> (2007) and Schoffski (2009)

Conclusion

Mitosis is a fundamental biological process essential for cellular growth and tissue development. The mechanisms of yeast cell division are well understood. In the last 10 years, a focused effort from many laboratories has begun to expand our understanding of the mechanisms required for animal cell division. In this article, we highlight the morphological events that take place in each sequential stage of mitosis in animal cells and provide an overview of the key signaling proteins/pathways that regulate these events. Specific focus is on cytokinesis, which is the stage of mitosis that researchers have made the most significant advancements in recent years. We also provide some examples of how our understanding of the cell division process is critical to gaining insight into cancer biology mechanisms and how these discoveries may translate into new anticancer therapeutic strategies for better treatment of cancer patients.

Acknowledgments

Work from the Chircop lab is supported by grants from the National Health and Medical Research Council (NH&MRC) of Australia, Cancer Council NSW, Australian Research Council, and an NH&MRC Career Development Award Fellowship.

See also: Cell Division/Death: Apoptosis: Mitotic Catastrophe

References

- Alberts, B., 2002. *Molecular Biology of the Cell*, fourth ed. New York, NY: Garland Science.
- Beaudouin, J., Gerlich, D., Daigle, N., Eils, R., Ellenberg, J., 2002. Nuclear envelope breakdown proceeds by microtubule-induced tearing of the lamina. *Cell* 108 (1), 83–96. PM:11792323.
- Bement, W.M., Benink, H.A., von Dassow, G., 2005. A microtubule-dependent zone of active RhoA during cleavage plane specification. *Journal of Cell Biology* 170 (1), 91–101. PM:15998801.
- Bergstrahl, D.T., Haack, T., St Johnston, D., 2013. Epithelial polarity and spindle orientation: Intersecting pathways. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 368 (1629), 20130291. PM:24062590.
- Bringmann, H., Hyman, A.A., 2005. A cytokinesis furrow is positioned by two consecutive signals. *Nature* 436 (7051), 731–734. PM:16079852.
- Burgess, D.R., Chang, F., 2005. Site selection for the cleavage furrow at cytokinesis. *Trends in Cell Biology* 15 (3), 156–162. PM:15752979.
- Canman, J.C., Bement, W.M., 1997. Microtubules suppress actomyosin-based cortical flow in *Xenopus* oocytes. *Journal of Cell Science* 110 (Pt 16), 1907–1917. PM:9296390.
- Canman, J.C., Cameron, L.A., Maddox, P.S., et al., 2003. Determining the position of the cell division plane. *Nature* 424 (6952), 1074–1078. PM:12904818.
- Carlton, J.G., Agromayor, M., Martin-Serrano, J., 2008. Differential requirements for Axl and ESCRT-III in cytokinesis and HIV-1 release. *Proceedings of the National Academy of Sciences of the United States of America* 105 (30), 10541–10546. PM:18641129.
- Carlton, J.G., Caballe, A., Agromayor, M., Kloc, M., Martin-Serrano, J., 2012. ESCRT-III governs the Aurora B-mediated abscission checkpoint through CHMP4C. *Science* 336 (6078), 220–225. PM:22422861.
- Chen, X.W., Inoue, M., Hsu, S.C., Saltiel, A.R., 2006. RalA-exocyst-dependent recycling endosome trafficking is required for the completion of cytokinesis. *Journal of Biological Chemistry* 281 (50), 38609–38616. PM:17028198.
- Cheung, C.H., Coumar, M.S., Chang, J.Y., Hsieh, H.P., 2011. Aurora kinase inhibitor patents and agents in clinical testing: An update (2009–10). *Expert Opinion on Therapeutic Patents* 21 (6), 857–884. PM:21591849.
- Chi, Y.H., Jeang, K.T., 2007. Aneuploidy and cancer. *Journal of Cellular Biochemistry* 102 (3), 531–538. PM:17661351.
- Chircop, M., Malladi, C.S., Lian, A.T., et al., 2010. Calcineurin activity is required for the completion of cytokinesis. *Cellular and Molecular Life Sciences* 67, 3725–3737. PM:20496096.
- Chung, V., Heath, E.I., Schelman, W.R., et al., 2012. First-time-in-human study of GSK923295, a novel antimitotic inhibitor of centromere-associated protein E (CENP-E), in patients with refractory cancer. *Cancer Chemotherapy and Pharmacology* 69 (3), 733–741. PM:22020315.
- Colombo, R., Caldarelli, M., Mennecozzi, M., et al., 2010. Targeting the mitotic checkpoint for cancer therapy with NMS-P715, an inhibitor of MPS1 kinase. *Cancer Research* 70 (24), 10255–10264. PM:21159646.
- Connell, J.W., Lindon, C., Luzio, J.P., Reid, E., 2009. Spastin couples microtubule severing to membrane traffic in completion of cytokinesis and secretion. *Traffic* 10 (1), 42–56. PM:19000169.
- Danowski, B.A., 1989. Fibroblast contractility and actin organization are stimulated by microtubule inhibitors. *Journal of Cell Science* 93 (Pt 2), 255–266. PM:2482296.
- Dao, V.T., Dupuy, A.G., Gavet, O., Caron, E., de Gunzburg, J., 2009. Dynamic changes in Rap1 activity are required for cell retraction and spreading during mitosis. *Journal of Cell Science* 122 (Pt 16), 2996–3004. PM:19638416.
- Dechant, R., Glotzer, M., 2003. Centrosome separation and central spindle assembly act in redundant pathways that regulate microtubule density and trigger cleavage furrow formation. *Developmental Cell* 4 (3), 333–344. PM:12636915.
- Douglas, M.E., Davies, T., Joseph, N., Mishima, M., 2010. Aurora B and 14-3-3 coordinately regulate clustering of centralspindlin during cytokinesis. *Current Biology* 20 (10), 927–933. PM:20451386.
- Drechsel, D.N., Hyman, A.A., Hall, A., Glotzer, M., 1997. A requirement for Rho and Cdc42 during cytokinesis in *Xenopus* embryos. *Current Biology* 7 (1), 12–23. PM:8999996.
- Echard, A., Hickson, G.R., Foley, E., O'Farrell, P.H., 2004. Terminal cytokinesis events uncovered after an RNAi screen. *Current Biology* 14 (18), 1685–1693. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15380073 (accessed 24.02.15).
- Eggert, U.S., Mitchison, T.J., Field, C.M., 2006. Animal cytokinesis: From parts list to mechanisms. *Annual Review of Biochemistry* 75, 543–566. ISI:000239807600021.
- El, A.N., Kechad, A., Jananji, S., Hickson, G.R., 2013. Opposing actions of septins and Anillin promote the transition from contractile to midbody ring. *Journal of Cell Biology* 203 (3), 487–504. PM:24217622.
- Elia, N., Sougrat, R., Spurlin, T.A., Hurley, J.H., Lippincott-Schwartz, J., 2011. Dynamics of endosomal sorting complex required for transport (ESCRT) machinery during cytokinesis and its role in abscission. *Proceedings of the National Academy of Sciences of the United States of America* 108 (12), 4846–4851. PM:21383202.
- Ettinger, A.W., Wilsch-Brauninger, M., Marzesso, A.M., et al., 2011. Proliferating versus differentiating stem and cancer cells exhibit distinct midbody-release behaviour. *Nature Communications* 2, 503. PM:22009035.
- Fielding, A.B., Schonteich, E., Matheson, J., et al., 2005. Rab11-FIP3 and FIP4 interact with Arf6 and the exocyst to control membrane traffic in cytokinesis. *EMBO Journal* 24 (19), 3389–3399. PM:16148947.
- Gallant, P., Nigg, E.A., 1992. Cyclin B2 undergoes cell cycle-dependent nuclear translocation and, when expressed as a non-destructible mutant, causes mitotic arrest in HeLa cells. *Journal of Cell Biology* 117 (1), 213–224. PM:1532584.
- Ganem, N.J., Storchova, Z., Pellman, D., 2007. Tetraploidy, aneuploidy and cancer. *Current Opinion in Genetics & Development* 17 (2), 157–162. PM:17324569.
- Gassmann, R., Carvalho, A., Henzing, A.J., et al., 2004. Borealin: A novel chromosomal passenger required for stability of the bipolar mitotic spindle. *Journal of Cell Biology* 166 (2), 179–191. PM:15249581.
- Georgatos, S.D., Pyrasopoulou, A., Theodoropoulos, P.A., 1997. Nuclear envelope breakdown in mammalian cells involves stepwise lamina disassembly and microtubule-drive deformation of the nuclear membrane. *Journal of Cell Science* 110 (Pt 17), 2129–2140. PM:9378763.
- Glotzer, M., 2004. Cleavage furrow positioning. *Journal of Cell Biology* 164 (3), 347–351. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14757750 (accessed 24.02.15).
- Glotzer, M., 2005. The molecular requirements for cytokinesis. *Science* 307 (5716), 1735–1739. ISI:000227883900034.

- Gong, D., Pomeroy, J.R., Myers, J.W., *et al.*, 2007. Cyclin A2 regulates nuclear-envelope breakdown and the nuclear accumulation of cyclin B1. *Current Biology* 17 (1), 85–91. PM:17208191.
- Goss, J.W., Toomre, D.K., 2008. Both daughter cells traffic and exocytose membrane at the cleavage furrow during mammalian cytokinesis. *Journal of Cell Biology* 181 (7), 1047–1054. PM:18573914.
- Green, R.A., Mayers, J.R., Wang, S., *et al.*, 2013. The midbody ring scaffolds the abscission machinery in the absence of midbody microtubules. *Journal of Cell Biology* 203 (3), 505–520. PM:24217623.
- Gromley, A., Jurczyk, A., Silibourne, J., *et al.*, 2003. A novel human protein of the maternal centriole is required for the final stages of cytokinesis and entry into S phase. *Journal of Cell Biology* 161 (3), 535–545. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12732615 (accessed 24.02.15).
- Gromley, A., Yeaman, C., Rosa, J., *et al.*, 2005. Centriolin anchoring of exocyst and SNARE complexes at the midbody is required for secretory-vesicle-mediated abscission. *Cell* 123 (1), 75–87. ISI:000232536800010.
- Guha, M., Zhou, M., Wang, Y.L., 2005. Cortical actin turnover during cytokinesis requires myosin II. *Current Biology* 15 (8), 732–736. PM:15854905.
- Guizetti, J., Schermelleh, L., Mantler, J., *et al.*, 2011a. Cortical constriction during abscission involves helices of ESCRT-III-dependent filaments. *Science* 331 (6024), 1616–1620. PM:21310966.
- Guizetti, J., Schermelleh, L., Mantler, J., *et al.*, 2011b. Cortical constriction during abscission involves helices of ESCRT-III-dependent filaments. *Science* 331 (6024), 1616–1620. PM:21310966.
- Guse, A., Mishima, M., Glotzer, M., 2005. Phosphorylation of ZEN-4/MKLP1 by aurora B regulates completion of cytokinesis. *Current Biology* 15 (8), 778–786. PM:15854913.
- Hanrahan, E.O., Kies, M.S., Glisson, B.S., *et al.*, 2009. A phase II study of Lonafarnib (SCH66336) in patients with chemorefractory, advanced squamous cell carcinoma of the head and neck. *American Journal of Clinical Oncology* 32 (3), 274–279. PM:19433965.
- Hara, T., Abe, M., Inoue, H., *et al.*, 2006. Cytokinesis regulator ECT2 changes its conformation through phosphorylation at Thr-341 in G2/M phase. *Oncogene* 25 (4), 566–578. PM:16170345.
- Heald, R., McKeon, F., 1990. Mutations of phosphorylation sites in lamin A that prevent nuclear lamina disassembly in mitosis. *Cell* 61 (4), 579–589. PM:2344612.
- Helmke, K.J., Heald, R., Wilbur, J.D., 2013. Interplay between spindle architecture and function. *International Review of Cell and Molecular Biology* 306, 83–125. PM:24016524.
- Hendricks, A.G., Lazarus, J.E., Perlson, E., *et al.*, 2012. Dynein tethers and stabilizes dynamic microtubule plus ends. *Current Biology* 22 (7), 632–637. PM:22445300.
- Hershko, A., 1999. Mechanisms and regulation of the degradation of cyclin B. *Proceedings of the Royal Society of London B Biological Sciences* 354 (1389), 1571–1575. PM:10582242.
- Hwang, L.H., Lau, L.F., Smith, D.L., *et al.*, 1998. Budding yeast Cdc20: A target of the spindle checkpoint. *Science* 279 (5353), 1041–1044. PM:9461437.
- Ikebe, M., Koretz, J., Hartshorne, D.J., 1988. Effects of phosphorylation of light chain residues threonine 18 and serine 19 on the properties and conformation of smooth muscle myosin. *Journal of Biological Chemistry* 263 (13), 6432–6437. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=2966156 (accessed 24.02.15).
- Jackson, J.R., Patrick, D.R., Dar, M.M., Huang, P.S., 2007. Targeted anti-mitotic therapies: Can we improve on tubulin agents? *Nature Reviews Cancer* 7 (2), 107–117. PM:17251917.
- Janssen, A., Medema, R.H., 2011. Mitosis as an anti-cancer target. *Oncogene* 30 (25), 2799–2809. PM:21339734.
- Jantsch-Plunger, V., Gonczy, P., Romano, A., *et al.*, 2000. CYK-4: A Rho family GTPase activating protein (GAP) required for central spindle formation and cytokinesis. *Journal of Cell Biology* 149 (7), 1391–1404. PM:10871280.
- Julian, M., Tollon, Y., Lajoie-Mazenc, I., *et al.*, 1993. Gamma-Tubulin participates in the formation of the midbody during cytokinesis in mammalian cells. *Journal of Cell Science* 10 (Pt 1), 145–156. PM:8360269.
- Kamijo, K., Ohara, N., Abe, M., *et al.*, 2006. Dissecting the role of Rho-mediated signaling in contractile ring formation. *Molecular Biology of the Cell* 17 (1), 43–55. PM:16236794.
- Kawano, Y., Fukata, Y., Oshiro, N., *et al.*, 1999. Phosphorylation of myosin-binding subunit (MBS) of myosin phosphatase by Rho-kinase *in vivo*. *Journal of Cell Biology* 147 (5), 1023–1038. PM:10579722.
- Kechad, A., Jananji, S., Ruella, Y., Hickson, G.R., 2012. Anillin acts as a bifunctional linker coordinating midbody ring biogenesis during cytokinesis. *Current Biology* 22 (3), 197–203. PM:22226749.
- Kim, S.H., Lin, D.P., Matsumoto, S., Kitazono, A., Matsumoto, T., 1998. Fission yeast Slp1: an effector of the Mad2-dependent spindle checkpoint. *Science* 279 (5353), 1045–1047. PM:9461438.
- Kishi, K., Sasaki, T., Kuroda, S., Itoh, T., Takai, Y., 1993. Regulation of cytoplasmic division of *Xenopus* embryo by rho p21 and its inhibitory GDP/GTP exchange protein (rho GDI). *Journal of Cell Biology* 120 (5), 1187–1195. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8436590 (accessed 24.02.15).
- Kitzen, J.J., de Jonge, M.J., Verweij, J., 2010. Aurora kinase inhibitors. *Critical Reviews in Oncology/Hematology* 73 (2), 99–110.
- Kiyomitsu, T., Cheeseman, I.M., 2013. Cortical dynein and asymmetric membrane elongation coordinately position the spindle in anaphase. *Cell* 154 (2), 391–402. PM:23870127.
- Komatsu, S., Yano, T., Shibata, M., Tuft, R.A., Ikebe, M., 2000. Effects of the regulatory light chain phosphorylation of myosin II on mitosis and cytokinesis of mammalian cells. *Journal of Biological Chemistry* 275 (44), 34512–34520. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10944522 (accessed 24.02.15).
- Kops, G.J.P.L., Weaver, B.A.A., Cleveland, D.W., 2005. On the road to cancer: Aneuploidy and the mitotic checkpoint. *Nature Reviews Cancer* 5 (10), 773–785. ISI:000232522800014.
- Kotak, S., Busso, C., Gonczy, P., 2012. Cortical dynein is critical for proper spindle positioning in human cells. *Journal of Cell Biology* 199 (1), 97–110. PM:23027904.
- Kotak, S., Gonczy, P., 2013. Mechanisms of spindle positioning: Cortical force generators in the limelight. *Current Opinion in Cell Biology* 25 (6), 741–748. PM:23958212.
- Kunda, P., Pelling, A.E., Liu, T., Baum, B., 2008. Moesin controls cortical rigidity, cell rounding, and spindle morphogenesis during mitosis. *Current Biology* 18 (2), 91–101. PM:18207738.
- Kuo, T.C., Chen, C.T., Baron, D., *et al.*, 2011. Midbody accumulation through evasion of autophagy contributes to cellular reprogramming and tumorigenicity. *Nature Cell Biology* 13 (10), 1214–1223. PM:21909099.
- Kurasawa, Y., Earnshaw, W.C., Mochizuki, Y., Dohmae, N., Todokoro, K., 2004. Essential roles of KIF4 and its binding partner PRC1 in organized central spindle midzone formation. *EMBO Journal* 23 (16), 3237–3248. PM:15297875.
- Kwiatkowski, N., Jelluma, N., Filippakopoulos, P., *et al.*, 2010. Small-molecule kinase inhibitors provide insight into Mps1 cell cycle function. *Nature Chemical Biology* 6 (5), 359–368. PM:20383151.
- Laan, L., Pavin, N., Husson, J., *et al.*, 2012. Cortical dynein controls microtubule dynamics to generate pulling forces that position microtubule asters. *Cell* 148 (3), 502–514. PM:22304918.
- Lafaurie-Janvore, J., Maiuri, P., Wang, I., *et al.*, 2013. ESCRT-III assembly and cytokinetic abscission are induced by tension release in the intercellular bridge. *Science* 339 (6127), 1625–1629. PM:23539606.
- Lara-Gonzalez, P., Westhorpe, F.G., Taylor, S.S., 2012. The spindle assembly checkpoint. *Current Biology* 22 (22), R966–R980. PM:23174302.
- Lee, K., Rhee, K., 2011a. PLK1 phosphorylation of pericentrin initiates centrosome maturation at the onset of mitosis. *Journal of Cell Biology* 195 (7), 1093–1101. PM:22184200.
- Lee, K., Rhee, K., 2011b. PLK1 phosphorylation of pericentrin initiates centrosome maturation at the onset of mitosis. *Journal of Cell Biology* 195 (7), 1093–1101. PM:22184200.
- Lens, S.M., Voest, E.E., Medema, R.H., 2010. Shared and separate functions of polo-like kinases and aurora kinases in cancer. *Nature Reviews Cancer* 10 (12), 825–841. PM:21102634.
- Li, F., Higgs, H.N., 2003. The mouse Formin mDia1 is a potent actin nucleation factor regulated by autoinhibition. *Current Biology* 13 (15), 1335–1340. PM:12906795.
- Lindqvist, A., Rodriguez-Bravo, V., Medema, R.H., 2009. The decision to enter mitosis: Feedback and redundancy in the mitotic entry network. *Journal of Cell Biology* 185 (2), 193–202. PM:19364923.
- Lock, R.B., Carol, H., Morton, C.L., *et al.*, 2012. Initial testing of the CENP-E inhibitor GSK923295A by the pediatric preclinical testing program. *Pediatric Blood & Cancer* 58 (6), 916–923.
- Ma, H.T., Poon, R.Y., 2011. How protein kinases co-ordinate mitosis in animal cells. *Biochemical Journal* 435 (1), 17–31. PM:21406064.
- Ma, M.P., Chircop, M., 2012a. SNX9, SNX18 and SNX33 are required for progression through and completion of mitosis. *Journal of Cell Science* 125 (Pt 18), 4372–4382. PM:22718350.
- Ma, X., Kovacs, M., Conti, M.A., *et al.*, 2012b. Nonmuscle myosin II exerts tension but does not translocate actin in vertebrate cytokinesis. *Proceedings of the National Academy of Sciences of the United States of America* 109 (12), 4509–4514. PM:22393000.

- Maddox, A.S., Burridge, K., 2003. RhoA is required for cortical retraction and rigidity during mitotic cell rounding. *Journal of Cell Biology* 160 (2), 255–265. PM:12538643.
- Mali, P., Wirtz, D., Searson, P.C., 2010. Interplay of RhoA and motility in the programmed spreading of daughter cells postmitosis. *Biochemical Journal* 99 (11), 3526–3534.
- Matthews, H.K., Delabre, U., Rohn, J.L., *et al.*, 2012. Changes in Ect2 localization couple actomyosin-dependent cell shape changes to mitotic progression. *Developmental Cell* 23 (2), 371–383. PM:22898780.
- Maupin, P., Pollard, T.D., 1986. Arrangement of actin filaments and myosin-like filaments in the contractile ring and of actin-like filaments in the mitotic spindle of dividing HeLa cells. *Journal of Ultrastructure and Molecular Structure Research* 94 (1), 92–103. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3772181 (accessed 24.02.15).
- Mendoza, M., Norden, C., Durrer, K., *et al.*, 2009. A mechanism for chromosome segregation sensing by the NoCut checkpoint. *Nature Cell Biology* 11 (4), 477–483. PM:19270692.
- Meraldi, P., Nigg, E.A., 2002. The centrosome cycle. *FEBS Letters* 521 (1–3), 9–13. PM:12067716.
- Mishima, M., Kaitna, S., Glotzer, M., 2002. Central spindle assembly and cytokinesis require a kinesin-like protein/RhoGAP complex with microtubule bundling activity. *Developmental Cell* 2 (1), 41–54. PM:11782313.
- Mishima, M., Pavicic, V., Gruneberg, U., Nigg, E.A., Glotzer, M., 2004. Cell cycle regulation of central spindle assembly. *Nature* 430 (7002), 908–913. PM:15282614.
- Mollinari, C., Kleman, J.P., Jiang, W., *et al.*, 2002. PRC1 is a microtubule binding and bundling protein essential to maintain the mitotic spindle midzone. *Journal of Cell Biology* 157 (7), 1175–1186. PM:12082078.
- Morita, E., Sandrin, V., Chung, H.Y., *et al.*, 2007. Human ESCRT and ALIX proteins interact with proteins of the midbody and function in cytokinesis. *EMBO Journal* 26 (19), 4215–4227. PM:17853893.
- Morrow, C.J., Tighe, A., Johnson, V.L., *et al.*, 2005. Bub1 and aurora B cooperate to maintain BubR1-mediated inhibition of APC/CCdc20. *Journal of Cell Science* 118 (Pt 16), 3639–3652. PM:16046481.
- Murthy, K., Wadsworth, P., 2005. Myosin-II-dependent localization and dynamics of F-actin during cytokinesis. *Current Biology* 15 (8), 724–731. PM:15854904.
- Musacchio, A., Salmon, E.D., 2007. The spindle-assembly checkpoint in space and time. *Nature Reviews Molecular Cell Biology* 8 (5), 379–393. PM:17426725.
- Neto, H., Balmer, G., Gould, G., 2013a. Exocyst proteins in cytokinesis: Regulation by Rab11. *Communicative & Integrative Biology* 6 (6), e27635. PM:24563720.
- Neto, H., Kaupisch, A., Collins, L.L., Gould, G.W., 2013b. Syntaxin 16 is a master recruitment factor for cytokinesis. *Molecular Biology of the Cell* 24 (23), 3663–3674. PM:24109596.
- Nicholson, J.M., Cimini, D., 2011. How mitotic errors contribute to karyotypic diversity in cancer. *Advances in Cancer Research* 112, 43–75. PM:21925301.
- Nicklas, R.B., Ward, S.C., Gorbosky, G.J., 1995. Kinetochore chemistry is sensitive to tension and may link mitotic forces to a cell cycle checkpoint. *Journal of Cell Biology* 130 (4), 929–939. PM:7642708.
- Niia, F., Tatsumoto, T., Lee, K.S., Miki, T., 2006. Phosphorylation of the cytokinesis regulator ECT2 at G2/M phase stimulates association of the mitotic kinase Plk1 and accumulation of GTP-bound RhoA. *Oncogene* 25 (6), 827–837. PM:16247472.
- Nishimura, Y., Yonemura, S., 2006. Centralspindlin regulates ECT2 and RhoA accumulation at the equatorial cortex during cytokinesis. *Journal of Cell Science* 119 (Pt 1), 104–114. PM:16352658.
- Nurse, P., 2000. A long twentieth century of the cell cycle and beyond. *Cell* 100 (1), 71–78. PM:10647932.
- Peter, M., Nakagawa, J., Doree, M., Labbe, J.C., Nigg, E.A., 1990. *In vitro* disassembly of the nuclear lamina and M phase-specific phosphorylation of lamins by cdc2 kinase. *Cell* 61 (4), 591–602. PM:2188731.
- Peters, J.M., 2006. The anaphase promoting complex/cyclosome: A machine designed to destroy. *Nature Reviews Molecular Cell Biology* 7 (9), 644–656. PM:16896351.
- Piekny, A., Werner, M., Glotzer, M., 2005. Cytokinesis: Welcome to the Rho zone. *Trends in Cell Biology* 15 (12), 651–658. PM:16243528.
- Pletjushkina, O.J., Rajfur, Z., Pomorski, P., *et al.*, 2001. Induction of cortical oscillations in spreading cells by depolymerization of microtubules. *Cell Motility and the Cytoskeleton* 48 (4), 235–244. PM:11276073.
- Pohl, C., Jentsch, S., 2008. Final stages of cytokinesis and midbody ring formation are controlled by BRUCE. *Cell* 132 (5), 832–845. PM:18329369.
- Pohl, C., Jentsch, S., 2009. Midbody ring disposal by autophagy is a post-abscission event of cytokinesis. *Nature Cell Biology* 11 (1), 65–70.
- Raiborg, C., Stenmark, H., 2009. The ESCRT machinery in endosomal sorting of ubiquitinated membrane proteins. *Nature* 458 (7237), 445–452. PM:19325624.
- Rappaport, R., 1986. Establishment of the mechanism of cytokinesis in animal cells. *International Review of Cytology* 105, 245–281. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3539854 (accessed 24.02.15).
- Ren, X.D., Kiosses, W.B., Schwartz, M.A., 1999. Regulation of the small GTP-binding protein Rho by cell adhesion and the cytoskeleton. *EMBO Journal* 18 (3), 578–585. PM:9927417.
- Renshaw, M.J., Liu, J., Lavoie, B.D., Wilde, A., 2014. Anillin-dependent organization of septin filaments promotes intercellular bridge elongation and Chmp4B targeting to the abscission site. *Open Biology* 4, 130190. PM:24451548.
- Rieder, C.L., 2005. Kinetochore fiber formation in animal somatic cells: Dueling mechanisms come to a draw. *Chromosoma* 114 (5), 310–318.
- Sauer, G., Korner, R., Hanisch, A., *et al.*, 2005. Proteome analysis of the human mitotic spindle. *Molecular & Cellular Proteomics* 4 (1), 35–43. PM:15561729.
- Schiel, J.A., Park, K., Morphew, M.K., *et al.*, 2011. Endocytic membrane fusion and buckling-induced microtubule severing mediate cell abscission. *Journal of Cell Science* 124 (Pt 9), 1411–1424. PM:21486954.
- Schiel, J.A., Simon, G.C., Zaharris, C., *et al.*, 2012. FIP3-endosome-dependent formation of the secondary ingression mediates ESCRT-III recruitment during cytokinesis. *Nature Cell Biology* 14 (10), 1068–1078. PM:23000966.
- Schmidt, M., Bastians, H., 2007. Mitotic drug targets and the development of novel anti-mitotic anticancer drugs. *Drug Resistance Updates* 10 (4–5), 162–181. PM:17669681.
- Schoffski, P., 2009. Polo-like kinase (PLK) inhibitors in preclinical and early clinical development in oncology. *Oncologist* 14 (6), 559–570. PM:19474163.
- Scholey, J.M., Taylor, K.A., Kendrick-Jones, J., 1980. Regulation of non-muscle myosin assembly by calmodulin-dependent light chain kinase. *Nature* 287 (5779), 233–235. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=6893621 (accessed 24.02.15).
- Schweitzer, J.K., D'Souza-Schorey, C., 2005. A requirement for ARF6 during the completion of cytokinesis. *Experimental Cell Research* 311 (1), 74–83. ISI:000233268100007.
- Severson, A.F., Hamill, D.R., Carter, J.C., Schumacher, J., Bowerman, B., 2000. The aurora-related kinase AIR-2 recruits ZEN-4/CeMKLP1 to the mitotic spindle at metaphase and is required for cytokinesis. *Current Biology* 10 (19), 1162–1171. PM:11050384.
- Shu, H.B., Li, Z., Palacios, M.J., Li, Q., Joshi, H.C., 1995. A transient association of gamma-tubulin at the midbody is required for the completion of cytokinesis during the mammalian cell division. *Journal of Cell Science* 108 (Pt 9), 2955–2962. PM:8537435.
- Skop, A.R., Liu, H., Yates, J., Meyer, B.J., Heald, R., 2004. Dissection of the mammalian midbody proteome reveals conserved cytokinesis mechanisms. *Science* 305 (5680), 61–66. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15166316 (accessed 24.02.15).
- Smith, C.M., Chircop, M., 2012. Clathrin-mediated endocytic proteins are involved in regulating mitotic progression and completion. *Traffic* 13 (12), 1628–1641. PM:22901037.
- Steigemann, P., Gerlich, D.W., 2009a. Cytokinetic abscission: Cellular dynamics at the midbody. *Trends in Cell Biology* 19 (11), 606–616. PM:19733077.
- Steigemann, P., Wurzenberger, C., Schmitz, M.H., *et al.*, 2009b. Aurora B-mediated abscission checkpoint protects against tetraploidization. *Cell* 136 (3), 473–484. PM:19203582.
- Straight, A.F., Cheung, A., Limouze, J., *et al.*, 2003. Dissecting temporal and spatial control of cytokinesis with a myosin II inhibitor. *Science* 299 (5613), 1743–1747. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12637748 (accessed 24.02.14).
- Sudakin, V., Chan, G.K., Yen, T.J., 2001. Checkpoint inhibition of the APC/C in HeLa cells is mediated by a complex of BUBR1, BUB3, CDC20, and MAD2. *Journal of Cell Biology* 154 (5), 925–936. PM:11535616.
- Tamura, N., Draviam, V.M., 2012. Microtubule plus-ends within a mitotic cell are 'moving platforms' with anchoring, signaling and force-coupling roles. *Open Biology* 2 (11), 120132. PM:23226599.
- Tanaka, T.U., Rachidi, N., Janke, C., *et al.*, 2002. Evidence that the Ipl1-Sli15 (Aurora kinase-INCENP) complex promotes chromosome bi-orientation by altering kinetochore-spindle pole connections. *Cell* 108 (3), 317–329. PM:11853667.
- Terasaki, M., Campagnola, P., Rolls, M.M., *et al.*, 2001. A new model for nuclear envelope breakdown. *Molecular Biology of the Cell* 12 (2), 503–510. PM:11179431.

- Vogelstein, B., Kinzler, K.W., 2004. Cancer genes and the pathways they control. *Nature Medicine* 10 (8), 789–799. PM:15286780.
- van der Waal, M.S., Hengeveld, R.C., van der Horst, A., Lens, S.M., 2012. Cell division control by the chromosomal passenger complex. *Experimental Cell Research* 318 (12), 1407–1420. PM:22472345.
- Werner, M., Munro, E., Glotzer, M., 2007. Astral signals spatially bias cortical myosin recruitment to break symmetry and promote cytokinesis. *Current Biology* 17 (15), 1286–1297. PM:17669650.
- Wheatley, S.P., Carvalho, A., Vagnarelli, P., Earnshaw, W.C., 2001. INCENP is required for proper targeting of Survivin to the centromeres and the anaphase spindle during mitosis. *Current Biology* 11 (11), 886–890. PM:11516652.
- Wilson, G.M., Fielding, A.B., Simon, G.C., *et al.*, 2005. The FIP3-Rab11 protein complex regulates recycling endosome targeting to the cleavage furrow during late cytokinesis. *Molecular Biology of the Cell* 16 (2), 849–860. PM:15601896.
- Woodard, G.E., Huang, N.N., Cho, H., *et al.*, 2010. Ric-8A and Gi alpha recruit LGN, NuMA, and dynein to the cell cortex to help orient the mitotic spindle. *Molecular Biology of the Cell* 30 (14), 3519–3530. PM:20479129.
- Yamakita, Y., Yamashiro, S., Matsumura, F., 1994. *In vivo* phosphorylation of regulatory light chain of myosin II during mitosis of cultured cells. *Journal of Cell Biology* 124 (1–2), 129–137. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8294496 (accessed 24.02.15).
- Yamashiro, S., Totsukawa, G., Yamakita, Y., *et al.*, 2003. Citron kinase, a Rho-dependent kinase, induces di-phosphorylation of regulatory light chain of myosin II. *Molecular Biology of the Cell* 14 (5), 1745–1756. PM:12802051.
- Yokoyama, T., Goto, H., Izawa, I., Mizutani, H., Inagaki, M., 2005. Aurora-B and Rho-kinase/ROCK, the two cleavage furrow kinases, independently regulate the progression of cytokinesis: Possible existence of a novel cleavage furrow kinase phosphorylates ezrin/radixin/moesin (ERM). *Genes Cells* 10 (2), 127–137. PM:15676024.
- Yonemura, S., Hirao-Minakuchi, K., Nishimura, Y., 2004. Rho localization in cells and tissues. *Experimental Cell Research*, 295 (2), 300–314. PM:15093731.
- Yoshizaki, H., Ohba, Y., Kurokawa, K., *et al.*, 2003. Activity of Rho-family GTPases during cell division as visualized with FRET-based probes. *Journal of Cell Biology* 162 (2), 223–232. PM:12860967.
- Yuce, O., Piekny, A., Glotzer, M., 2005. An ECT2-centralspindlin complex regulates the localization and function of RhoA. *Journal of Cell Biology* 170 (4), 571–582. PM:16103226.
- Zeng, X., Sigoillot, F., Gaur, S., *et al.*, 2010. Pharmacologic inhibition of the anaphase-promoting complex induces a spindle checkpoint-dependent mitotic arrest in the absence of spindle damage. *Cancer Cell* 18 (4), 382–395. PM:20951947.
- Zheng, Z., Wan, Q., Liu, J., *et al.*, 2013. Evidence for dynein and astral microtubule-mediated cortical release and transport of Galphai/LGN/NuMA complex in mitotic cells. *Molecular Biology of the Cell* 24 (7), 901–913. PM:23389635.
- Zhu, C., Jiang, W., 2005. Cell cycle-dependent translocation of PRC1 on the spindle by Kif4 is essential for midzone formation and cytokinesis. *Proceedings of the National Academy of Sciences of the United States of America* 102 (2), 343–348. PM:15625105.

Relevant Website

<http://microkit.biocuckoo.org/>
Huazhong University of Science & Technology.

Regulating Cytokinesis

G Zachos, University of Crete, Heraklion, Greece

© 2016 Elsevier Inc. All rights reserved.

Cytokinesis, i.e., the process by which the cytoplasm of a cell is separated giving rise to two daughter cells at the end of cell division, is spatially and temporally coupled to chromosome segregation and can be divided into four stages: specification of the cleavage plane, ingression of the cleavage furrow, formation of the midbody and abscission, the final cut of the narrow cytoplasmic canal that connects the two daughter cells (Figure 1; Barr and Gruneberg, 2007; Normand and King, 2010).

In early anaphase, spindle microtubules are of critical importance for selecting the site of cell division (Figure 1). Furthermore, kinase signaling networks, kinesin motors and scaffolding proteins regulate formation of the contractile ring that includes actin and myosin proteins (the actomyosin ring) to promote cleavage furrow ingression. At later stages, membrane fusion events and remodeling of the plasma membrane at the midbody region promote completion of cytokinesis by abscission (Figure 1). Failure of cytokinesis results in formation of tetraploid cells which can lead to carcinogenesis

(Fujiwara *et al.*, 2005). Understanding cytokinesis will therefore give us insight into the process of carcinogenesis and could lead to novel therapeutic strategies to prevent tumor growth by inhibiting cancer cell division. In this article, we discuss recent advances in identifying the mechanisms that regulate cytokinesis in animal somatic cells.

Positioning the Site of Cell Division

The Role of Spindle Microtubules

Central issues in cytokinesis are to ensure that furrowing begins after sister chromatid segregation and the division plane is placed between the two sets of sister chromatids, to make sure that each daughter cell receives a complete set. Temporal regulation of cytokinesis will be discussed later in this article. Correct positioning of the cleavage furrow is controlled by mechanisms that require spindle microtubules (Glotzer, 2004; Burgess and Chang, 2005). There are three models of how microtubules of the early anaphase spindle direct positioning of the cleavage furrow which are not mutually exclusive (Figure 2).

First, equatorial astral microtubules, which extend from the spindle poles across the equator, stimulate formation of the cleavage furrow at the equatorial cortex (Figure 2(a); Canman *et al.*, 2003). Second, polar astral microtubules, which radiate from the centrosomes and anchor the spindle poles to the cell membrane, inhibit cortical contraction by preventing myosin recruitment to the asters to promote contraction at the equatorial region (Figure 2(b); Werner *et al.*, 2007). Finally, central spindle microtubules, i.e., microtubules that partially overlap at the equator with those emanating from the opposing pole, deliver signals that promote positioning of the cleavage furrow at the cell equator (Figure 2(c); Bringmann and Hyman, 2005). It is proposed that these signals provided by different populations of spindle microtubules are partially redundant to ensure efficient formation of the cleavage furrow (Dechant and Glotzer, 2003; Bringmann and Hyman, 2005; Werner *et al.*, 2007).

The RhoA Pathway

What are the positive signals delivered by microtubules that promote positioning of the cleavage furrow at the cell equator? The small GTPase RhoA is the main regulator of actin dynamics in interphase or mitotic cells and several studies suggest that RhoA is a critical target for furrow formation (Piekny *et al.*, 2005). RhoA is activated by the Rho GDP-GTP exchange factor (GEF) ECT2 and inactivated by the GTPase-activating protein (GAP) MgcRacGAP (Piekny *et al.*, 2005). MgcRacGAP forms a complex with the kinesin motor protein Mklp1 (the centralspindlin complex) and is restricted by Mklp1 to the central spindle. Centralspindlin binds ECT2 resulting in accumulation of the active GTP-bound form of RhoA in a

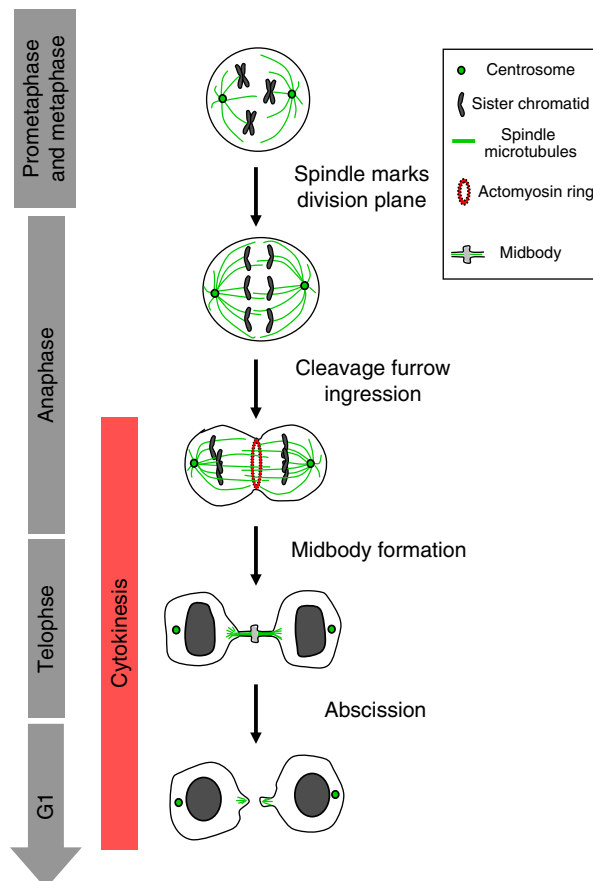


Figure 1 Stages of cytokinesis. Cytokinesis is spatially and temporally coupled to chromosome segregation and can be divided into four stages: specification of the division plane, ingression of the cleavage furrow, formation of the midbody, and abscission.

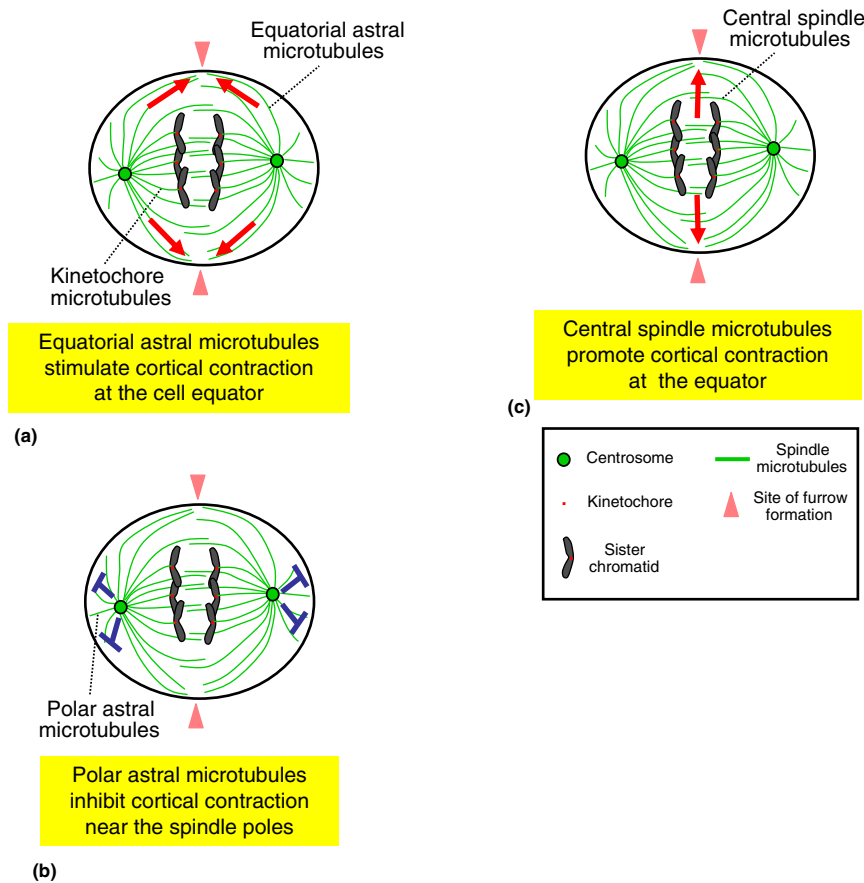


Figure 2 Spindle microtubules regulate positioning of the cleavage furrow. (a) Equatorial astral microtubules stimulate formation of the cleavage furrow. (b) Polar astral microtubules inhibit cortical contraction at the poles. (c) Central spindle microtubules promote positioning of the cleavage furrow at the cell equator.

narrow zone inside the central spindle (Figure 3; Yuce *et al.*, 2005). It is therefore proposed that microtubules deliver proteins that lead to local RhoA activation. In turn, activated RhoA marks the future cleavage site and organizes furrowing events.

Formation of the Cleavage Furrow

Activated RhoA stimulates actin polymerization and myosin II activity to promote assembly of the actomyosin ring and cleavage furrow ingression. Scaffolding proteins such as anillin and septins are also required for organizing the cleavage furrow and promoting cytokinesis. In addition, the Serine/Threonine kinases Aurora B and polo-like kinase 1 (Plk1) regulate RhoA activity and early furrowing events in animal cells.

RhoA Activity Promotes Actomyosin Ring Assembly

Active RhoA promotes assembly of actin filaments in the contractile ring by binding to formins, a group of proteins involved in actin polymerization (Figure 3; Pollard, 2007). Also, RhoA binds to ROCK and Citron kinases to induce their catalytic activity (Figure 3). ROCK and Citron kinase phosphorylate the myosin light chain (MLC) at positive regulatory

sites and these phosphorylations stimulate localization and motor activity of myosin II at the future cleavage site (Kosako *et al.*, 2000; Yamashiro *et al.*, 2003). Furthermore, ROCK phosphorylates the targeting subunit of myosin phosphatase to inactivate the phosphatase and promote positive MLC phosphorylation (Figure 3; Kawano *et al.*, 1999). As a result, active RhoA directs actin polymerization and myosin II activity at the cell equator. It is proposed that accumulation of actin and myosin II at the cell cortex leads to formation of contractile filaments that are drawn together like a 'purse string' and perform furrow ingression (Wang, 2005).

Scaffolding Proteins: Anillin and Septins

Anillin is a scaffold protein that binds actin filaments, myosin, septins, and activated RhoA (Normand and King, 2010). Anillin localizes to the cleavage furrow at early stages of cytokinesis, directs assembly of septin filaments and contributes to furrow stability during ingression, midbody formation, and abscission (Normand and King, 2010).

Septins are polymerizing GTPases that form filaments and assemble into two rings at the cell cortex on either side of the actomyosin ring (Dobbelaere and Barral, 2004). Septin rings act as barriers to compartmentalize the cortex around the cleavage site and maintain diffusible cortical factors to the site

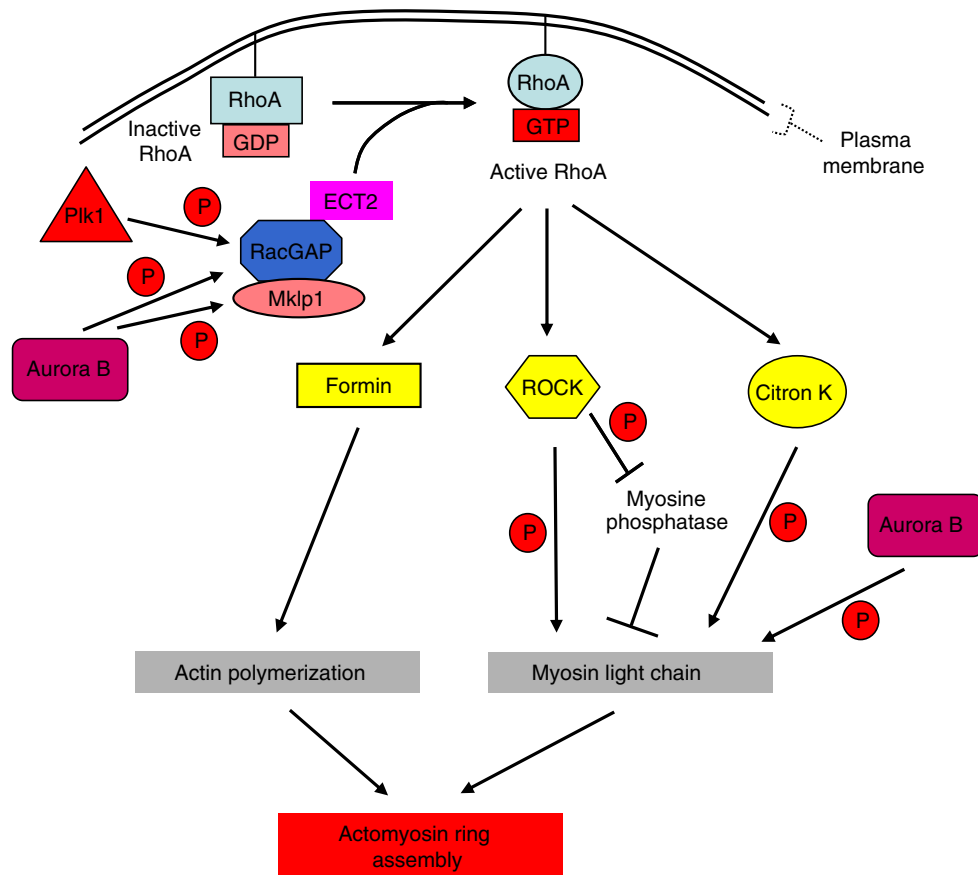


Figure 3 RhoA activity regulates assembly of the actomyosin ring. RhoA is activated by interaction with the ECT2/RacGAP/Mklp1 complex. Active RhoA stimulates actin polymerization by formin. Furthermore, active RhoA binds to ROCK and Citron kinase to induce activatory phosphorylations of MLC and promote assembly of the actomyosin ring. Also, ROCK inhibits removal of MLC activatory phosphorylations by myosin phosphatase. In addition, Plk1 and Aurora B regulate ECT2/RacGAP/Mklp1 and MLC (see text for details). P, phosphorylation.

of cleavage. In turn, such factors are required for actomyosin ring function and abscission (Dobbelaere and Barral, 2004).

Aurora B and the Chromosomal Passenger Complex

Aurora B kinase is the catalytic subunit of the chromosomal passenger complex (CPC) which also includes regulatory proteins INCENP, Survivin, and Borealin (van der Waal *et al.*, 2012). The CPC associates with chromosome arms in early prophase, localizes to centromeres in prometaphase and metaphase, and then transfers to the central spindle and the midbody in anaphase and telophase respectively (van der Waal *et al.*, 2012). Several studies demonstrate that the CPC is required for central spindle formation, furrow ingression, and late stages of cytokinesis (reviewed in Barr and Gruneberg, 2007). Aurora B phosphorylates Mklp1 and this phosphorylation could be important for central spindle assembly (Figure 3; Kaitna *et al.*, 2000). Also, Aurora B regulates myosin activity by phosphorylating MLC (Figure 3; Murata-Hori *et al.*, 2000). Furthermore, Aurora B phosphorylates MgcRacGAP to activate GAP RhoA activity and this phosphorylation could be important for restricting RhoA activity during furrow ingression or terminating RhoA activity in late cytokinesis

(Figure 3; Minoshima *et al.*, 2003). Also, Aurora B phosphorylates the intermediate filament protein vimentin and is required for filament disassembly during completion of cytokinesis (Goto *et al.*, 2003). In addition, Aurora B is required for the abscission checkpoint as discussed later in this article.

The Role of Plk1

Similar to the CPC, Plk1 also changes localization to various mitotic structures during cell division. Plk1 associates with centrosomes in prophase, is recruited to kinetochores in prometaphase and metaphase, localizes to the central spindle on anaphase and to the midbody in telophase (Petronczki *et al.*, 2008). Acute inactivation of Plk1 at the metaphase-to-anaphase transition prevented ingression of the cleavage furrow and initiation of cytokinesis (Brennan *et al.*, 2007; Burkard *et al.*, 2007). Plk1 phosphorylates MgcRacGAP and this phosphorylation generates a phospho-epitope recognized by ECT2 (Petronczki *et al.*, 2007; Wolfe *et al.*, 2009). Interaction of ECT2 with MgcRacGAP is required for recruitment of ECT2 to the central spindle and subsequent RhoA activation in the cleavage furrow (Figure 3). In addition, Plk1 negatively regulates abscission by phosphorylating Cep-55 (centrosomal protein of 55 kDa) as discussed later in this article.

Formation of the Midbody

After furrowing is completed, the two daughter cells remain connected by a narrow (diameter of 0.1–1 μm) cytoplasmic bridge known as the midbody (Hu *et al.*, 2012). The principal function of the midbody is to localize abscission which physically separates the two daughter cells however, understanding how midbody assembly is regulated is under active investigation. The midbody contains microtubules from the central spindle (also known as the midzone) and includes kinases, actin-binding proteins, microtubule-binding and membrane-associated proteins (Skop *et al.*, 2004). Midbody proteins can fall into three categories depending on their localization (Figure 4; Hu *et al.*, 2012). First, the ‘bulge’ proteins localize to the lump at the center of the midbody and include centralspindlin, ECT2, RhoA, anillin, and Cep-55, a protein required for abscission. Second, the ‘dark zone’ proteins localize to a narrow region on the microtubule bundle in the center of the midbody where tubulin staining is blocked. These include the kinesin protein KIF4, Plk1 kinase and PRC1, a microtubule bundling protein that is essential for central spindle formation. PRC1 is targeted to the central spindle by

KIF4 and in turn recruits the centralspindlin complex and additional mitotic kinesins (Hu *et al.*, 2012). In the absence of PRC1, the cleavage furrow is formed but abscission fails. Furthermore, PRC1 serves as an important docking site for Plk1 in the central spindle, as will be described in a later paragraph (Neef *et al.*, 2007). Third, the ‘flanking zone’ proteins localize to two broad bands of microtubules outside the dark zone and include the kinesin Mklp2 and its cargo Aurora B kinase. Once the midbody is assembled it organizes abscission, the final event of cytokinesis leading to two separate cells.

Abscission

A wide variety of proteins involved in vesicle trafficking and tethering, membrane fusion and plasma membrane remodeling are required for abscission. At this stage, the intercellular bridge has narrowed to approximately 0.2 μm in diameter and human cultured cells may remain connected for several hours before daughter cells are finally separated (Normand and King, 2010).

Vesicle Trafficking

Completion of cytokinesis requires remodeling of the plasma membrane and increased membrane surface at the site of abscission (Figure 5). Membrane trafficking and fusion are essential for those events and both the endoplasmic reticulum (ER)–Golgi and the recycling endosome pathways provide vesicles destined for delivery to the intercellular bridge during abscission (Montagnac *et al.*, 2008; Neto *et al.*, 2011). Noticeably, vesicles accumulate only on one side of the midbody suggesting that delivery is asymmetric (Gromley *et al.*, 2005).

Secretory vesicles derived from the ER–Golgi apparatus are delivered along microtubules down the intercellular bridge in late cytokinesis (Figure 5(a); Gromley *et al.*, 2005). Those vesicles carry vesicle-SNARE (v-SNARE) proteins at their surface which are essential for vesicle tethering and membrane fusion (Gromley *et al.*, 2005). Centriolin, a protein that localizes to the maternal centriole during interphase, is required for tethering vesicles to the midbody (Gromley *et al.*, 2003). Specifically, centriolin recruits components of the exocyst protein complex and, in turn, exocyst proteins tether secretory vesicles to the midbody via interacting with v-SNAREs at their surface (Figure 5(a); Gromley *et al.*, 2005). Furthermore, recruitment of centriolin and exocyst to the midbody is dependent on the Mklp1/MgcRacGAP (centralspindlin) complex (Gromley *et al.*, 2005).

Endocytosis is another source of vesicles for completion of cytokinesis (Figure 5(a); Boucrot and Kirchhausen, 2007). These vesicles are delivered to the midbody through the recycling endosome pathway in the cell body however, endocytosis adjacent to the midbody may also contribute to membrane remodeling during abscission (Figure 5(a), arrowhead). In interphase cells, the small GTPase protein Rab11 localizes to the recycling endosome and is required for the recycling of vesicles to the plasma membrane (Montagnac *et al.*, 2008). During cytokinesis, Rab11-positive vesicles are

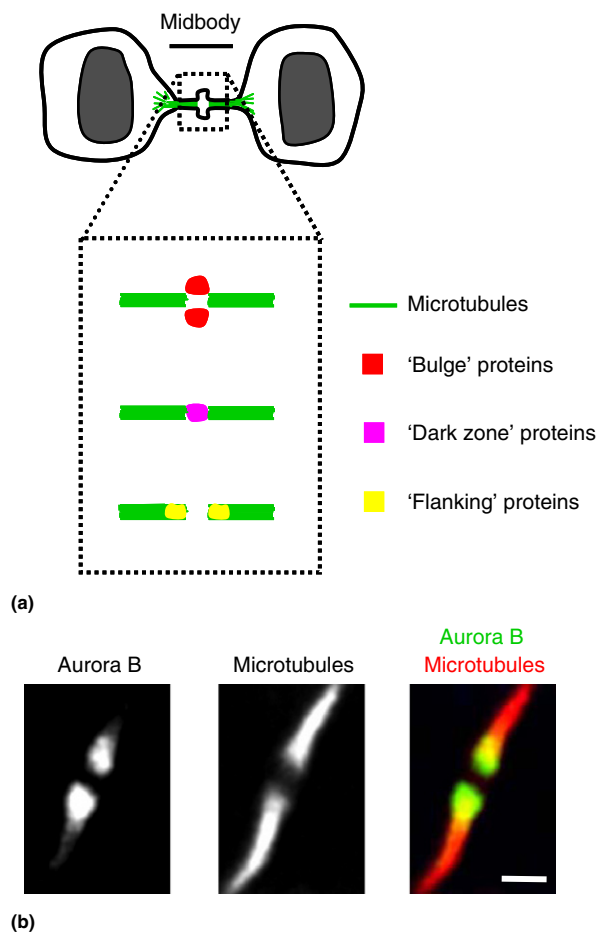


Figure 4 Midbody proteins. (a) Depending on their localization, midbody proteins can be classified as ‘bulge,’ ‘dark zone,’ or ‘flanking.’ (b) Midbody localization of the ‘flanking’ protein Aurora B. Confocal microscopy image of a midbody from BE human colon carcinoma cells. Green: Aurora B, red: α -tubulin. Bar, 3 μm .

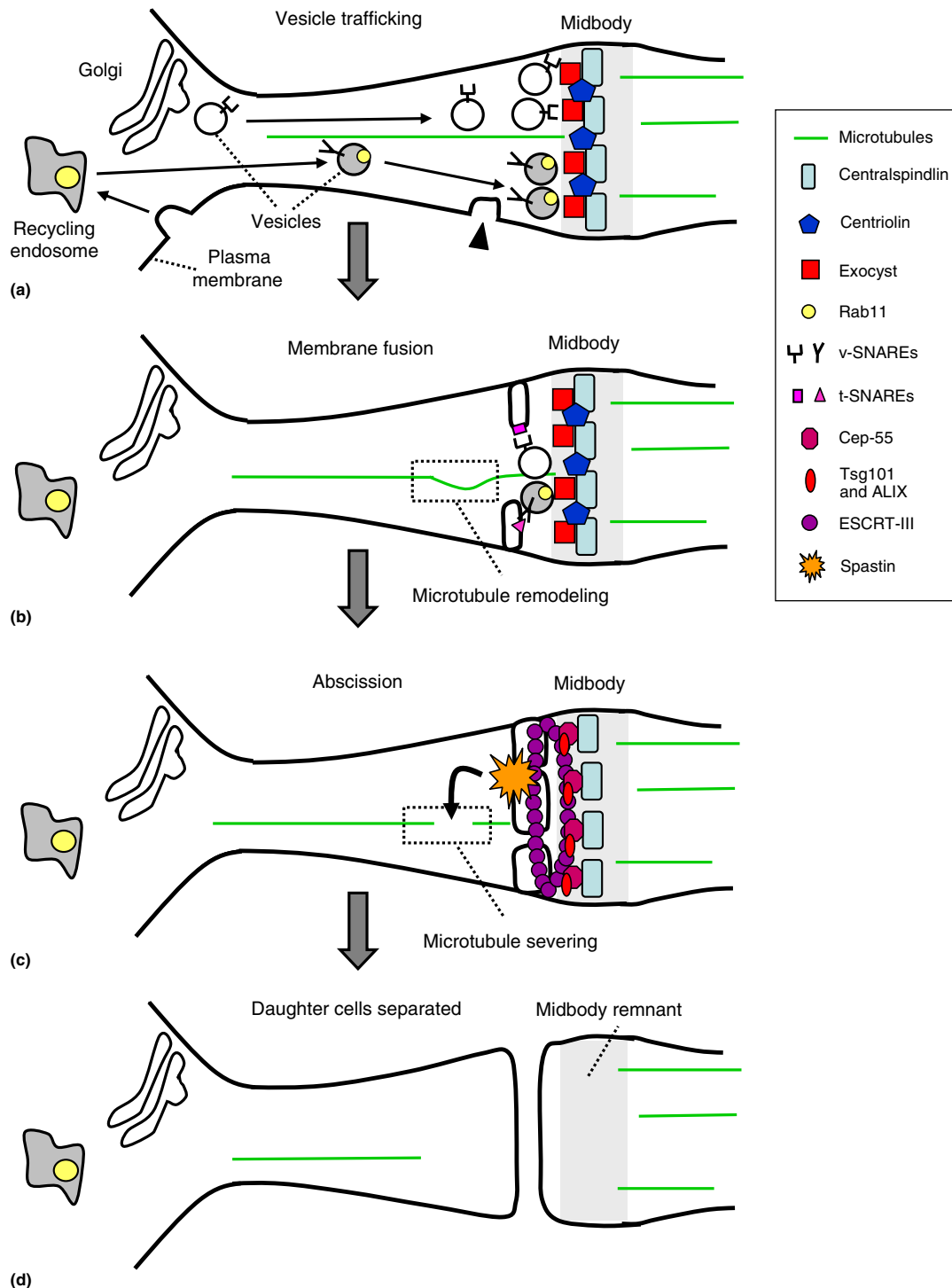


Figure 5 Late stages of cytokinesis in animal cells. (a) A model for vesicle trafficking and tethering. Secretory (white) or endocytic (gray) vesicles are anchored to the midbody through interactions with exocyst proteins. Endocytosis inside the intercellular bridge is marked by an arrowhead. (b) Membrane fusion. Tethered vesicles fuse with one another or with the plasma membrane through interactions of SNARE proteins. Remodeling of the central spindle causes midbody microtubules to 'bend.' (c) A model for abscission. ESCRT-III proteins form rings around the abscission site and make the plasma membrane curve to separate the daughter cells. Spastin cleaves 'curved' midbody microtubules. (d) After completion of cytokinesis, a midbody remnant is attached to one daughter cell.

transferred along microtubules inside the intercellular bridge by the motor protein kinesin-I (Montagnac *et al.*, 2009). Endocytic vesicles are then recruited to the midbody through

interaction of Rab11-binding proteins, such as FIP3 and FIP4 (not shown), with components of the exocyst complex (Figure 5(a); Fielding *et al.*, 2005).

Membrane Fusion

Following trafficking and anchoring of vesicles to the midbody, SNARE proteins mediate membrane fusion (Montagnac *et al.*, 2008). SNAREs are incorporated into membranes and share a cytosolic segment that is capable of reversibly assembling into bundles of alpha-helices. Vesicle-SNAREs (v-SNAREs) are incorporated in the membranes of transport vesicles whereas target-SNAREs (t-SNAREs) are integrated in target compartments such as the plasma membrane (Montagnac *et al.*, 2008). During abscission, v-SNAREs from recycling or secretory vesicles anchored to the midbody form tight complexes with t-SNAREs (such as syntaxin and endobrevin) into the plasma membrane or other vesicles causing membranes to fuse and generate new sections of the plasma membrane (Figure 5(b); Low *et al.*, 2003; Gromley *et al.*, 2005). Furthermore, accumulation of phosphatidylinositol lipids to endocytic vesicles and the plasma membrane inside the furrow has been proposed to enhance SNARE-dependent fusion and organize abscission (Neto *et al.*, 2011). In addition, vesicle accumulation in late telophase coincides with reorganization of the central spindle and generation of microtubule 'buckles' that may facilitate microtubule severing later in abscission (Figure 5(b); Schiel *et al.*, 2011).

ESCRT Proteins

Recently, endosomal sorting complex required for transport (ESCRT) proteins which are involved in multivesicular body formation and retroviral budding, have been also implicated in abscission (Neto and Gould, 2011). ESCRT subunits form protein complexes (ESCRTs -0, -I, -II, and -III; ALIX; and Vps4) that transiently assemble on specific target membranes and execute membrane scission in a direction away from the cytoplasm, which is topologically opposite to 'classical' vesicular transport involving membrane budding from a donor membrane into the cytoplasm (Henne *et al.*, 2011). In late cytokinesis, Mklp1 kinesin of the centralspindlin complex interacts with Cep-55 which mediates recruitment of ESCRT proteins to the late midbody (Bastos and Barr, 2010). Specifically, Cep-55 interacts with Tsg101 (an ESCRT-I component) and ALIX which in turn recruit ESCRT-III, the complex responsible for scission activity (Figure 5(c); Lee *et al.*, 2008; Neto and Gould, 2011). ESCRT-III proteins assemble in filaments and form rings around the abscission site (Figure 5(c)). These ESCRT-III rings make the plasma membrane curve, eventually driving the final cut that separates the daughter cells (Elia *et al.*, 2011; Guizetti *et al.*, 2011). In addition, ESCRT-III proteins recruit spastin (Yang *et al.*, 2008), a microtubule-severing protein that cleaves 'curved' midbody microtubules to complete cytokinesis (Figure 5(c); Guizetti *et al.*, 2011; Schiel *et al.*, 2011).

The Midbody Remnant

After completion of cytokinesis, a midbody remnant is attached to one daughter cell (Figure 5(d)). This implies asymmetric membrane trafficking and fusion that mediates abscission, however, how this is achieved is not fully understood (Barr and Gruneberg, 2007). The midbody remnant can

be shed to the extracellular space for degradation or internalized into the cytoplasm where it can be degraded by autophagy (Chai *et al.*, 2012).

Temporal Regulation of Cytokinesis

Thus far, we have outlined basic components and mechanisms that control different stages of cytokinesis. These mechanisms have to be temporally regulated to ensure that cytokinesis begins after chromosome segregation. In this section, we will discuss the role of mitotic kinases Cdk1 (Cyclin-dependent kinase 1), Aurora B, and Plk1 in regulating proper timing of cytokinesis.

Cdk1 Inactivation is the Trigger for Initiation of Cytokinesis

Cdk1, in complex with cyclin B (Cdk1/cyclin B), is a negative regulator of cytokinesis and Cdk1/cyclin B catalytic activity must be switched off to allow initiation of cytokinesis (Potapova *et al.*, 2006). In early mitosis, Cdk1/cyclin B activity is high to promote assembly of the mitotic apparatus and chromosome alignment. When all chromosomes become properly attached to spindle microtubules and aligned at the metaphase plate, the anaphase-promoting complex/cyclosome (APC/C) is activated. Activated APC/C is an E3 ubiquitin ligase that targets key proteins, including cyclin B and securin, for degradation, leading to Cdk1 inactivation and sister chromatid segregation (Kops *et al.*, 2005). At the end of mitosis, APC/C also targets cytokinesis proteins, such as Plk1 and Aurora B, for degradation to terminate cytokinesis (Normand and King, 2010).

Active Cdk1 inhibits central spindle formation and furrowing by several mechanisms (Figure 6(a)): First, Cdk1 prevents activation of RhoA by phosphorylating ECT2 at a site that inhibits its association with centralspindlin (Yuce *et al.*, 2005). Also, Cdk1 phosphorylates MLC to inhibit myosin activation (Satterwhite *et al.*, 1992). Furthermore, Cdk1 phosphorylates Mklp1 to reduce its microtubule-binding activity (Mishima *et al.*, 2004). Also, phosphorylation of PRC1 by Cdk1 inhibits its activity to bundle microtubules thus preventing central spindle formation (Jiang *et al.*, 1998). In addition, Cdk1 prevents localization of Aurora B and Plk1 to the central spindle as described in the following paragraphs (Figure 6(b)). After cyclin B degradation (Figure 6(c)), the above inhibitory Cdk1 phosphorylations are lifted to initiate cytokinesis.

Regulation of Aurora B by Cdk1

INCENP and Aurora B mediate recruitment of the CPC to the central spindle after anaphase onset (van der Waal *et al.*, 2012). Cdk1 phosphorylates INCENP and this phosphorylation prevents interaction of both INCENP and Aurora B with Mklp2 and docking of the CPC onto the central spindle (Figure 6(b); Hummer and Mayer, 2009). Because Aurora B is a positive regulator of RhoA, MLC activity and central spindle formation, inhibition of Aurora B localization to the central spindle by active Cdk1 restrains onset of cytokinesis.

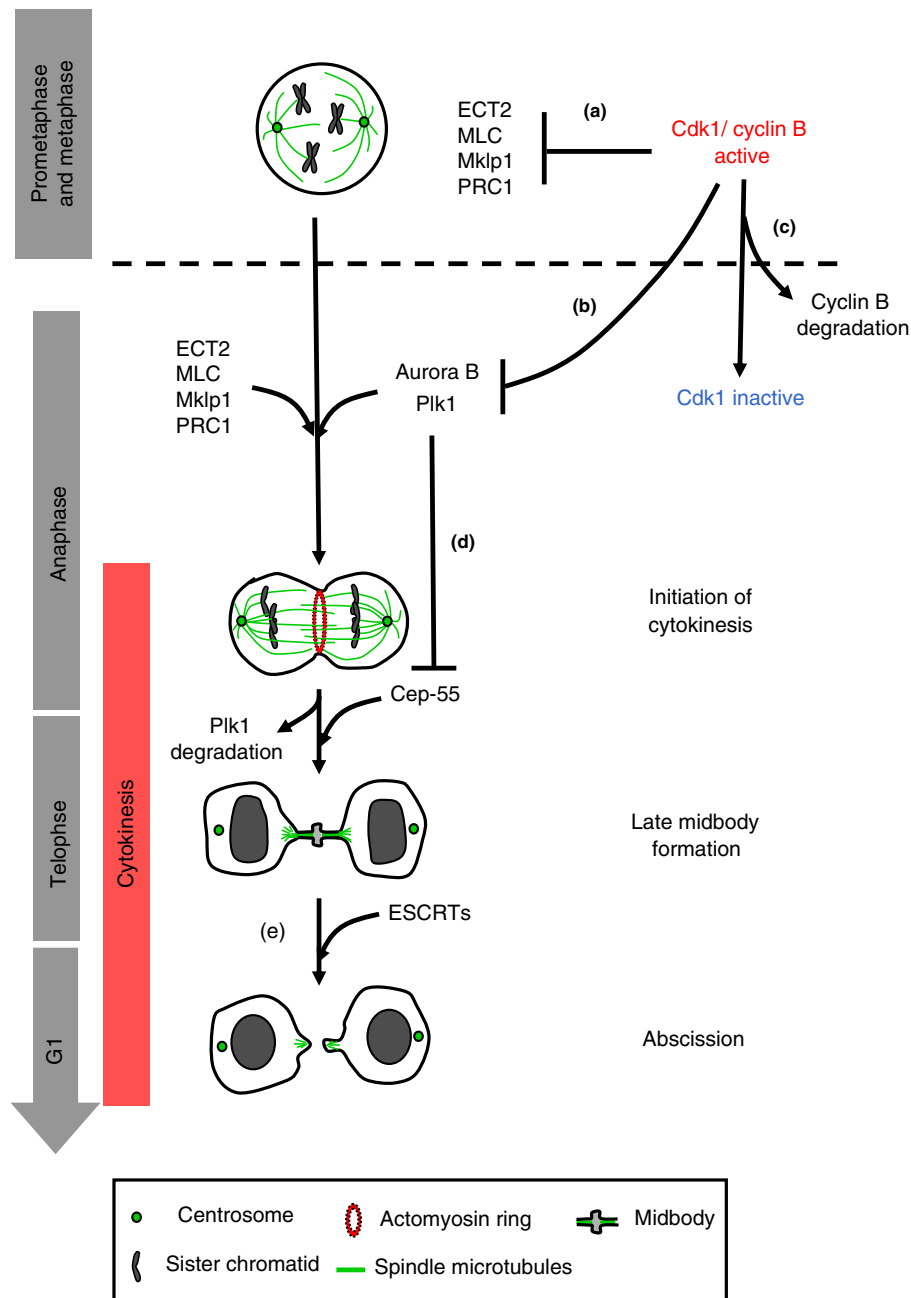


Figure 6 Cdk1 kinase activity prevents initiation of cytokinesis. (a) Cdk1 phosphorylates and inhibits ECT2, MLC, Mklp1, and PRC1 in prometaphase and metaphase (see text for details). (b) Cdk1 prevents recruitment of Aurora B and Plk1 to the central spindle. (c) After cyclin B degradation at the start of anaphase, Cdk1 is inactive and cytokinesis begins. (d) Plk1 prevents localization of Cep-55 to the central spindle. (e) At the end of cytokinesis, Plk1 is degraded and Cep-55 accumulates in the midbody to allow recruitment of ESCRTs and abscission. The horizontal dotted line indicates the metaphase to anaphase transition.

Temporal Regulation of Plk1

Plk1 recruitment to different subcellular structures during mitosis progression is mediated by a conserved sequence element at the C-terminus of the protein, termed the Polo-box domain (Petronczki *et al.*, 2008). The Polo-box domain preferentially binds to phosphorylated peptides suggesting that a kinase has to first 'prime' the target site before Plk1 docking

(Petronczki *et al.*, 2008). PRC1 is the major binding partner for Plk1 on anaphase spindles (Neef *et al.*, 2007). In prometaphase and metaphase, Cdk1 primes proteins to target Plk1 to centrosomes or kinetochores. Furthermore, Cdk1 phosphorylates PRC1 at sites that prevent Plk1 binding to PRC1 and docking onto the central spindle (Figure 6(b); Neef *et al.*, 2007). After Cdk1 inactivation at the beginning of anaphase, the inhibitory PRC1 phosphorylations are removed. As a

result, Plk1 phosphorylates PRC1 to self-prime its docking site and associates with PRC1 on central spindle where it plays an essential role in the initiation of cytokinesis (Neef *et al.*, 2007).

Also, Plk1 phosphorylates Cep-55 and this phosphorylation negatively regulates the interaction of Cep-55 with Mklp1 (Figure 6(d); Bastos and Barr, 2010). Only when Plk1 is degraded at the end of mitosis can Cep-55 interact with Mklp1 and accumulate in the midbody to allow recruitment of ESCRT proteins and abscission (Figure 6(e)).

The Abscission Checkpoint Protects against the Deleterious Effects of Chromatin Bridges

Chromatin Bridges

Chromatin bridges are strings of chromatin connecting the two segregating masses of chromosomes in anaphase or daughter nuclei in cytokinesis (Figures 7(a) and 7(b); Gisselsson, 2008). They can be caused by failure to eliminate replication or recombination intermediates during interphase and have been linked to chromosomal instability in human tumors and tumorigenesis in mice (Gisselsson, 2008). Vertebrate cells

recruit DNA helicases, such as the Bloom's syndrome protein helicase (BLM) and the Plk1-interacting checkpoint helicase (PICH), to chromatin bridges to unravel chromatin and promote bridge resolution (Ke *et al.*, 2011; Petsalaki *et al.*, 2014). However, if chromatin bridges are not resolved, attempting cytokinesis in their presence can lead to chromosome breakage and aneuploidy, or cytokinesis can fail and the cell will become tetraploid (Norden *et al.*, 2006; Steigemann *et al.*, 2009; Carlton *et al.*, 2012).

The Abscission Checkpoint

To avoid progression of cytokinesis in the presence of chromatin bridges, eukaryotic cells have evolved mechanisms that delay abscission until the bridges have been resolved, termed 'the abscission checkpoint' (Norden *et al.*, 2006; Steigemann *et al.*, 2009; Carlton *et al.*, 2012). Recent reports show that, in mammalian cells, Aurora B kinase regulates abscission by phosphorylating CHMP4C, a subunit of the ESCRT-III complex that forms rings around the abscission site to execute membrane scission (Capalbo *et al.*, 2012; Carlton *et al.*, 2012). It has been proposed that Aurora B phosphorylation drives localization of CHMP4C from the site of abscission to the

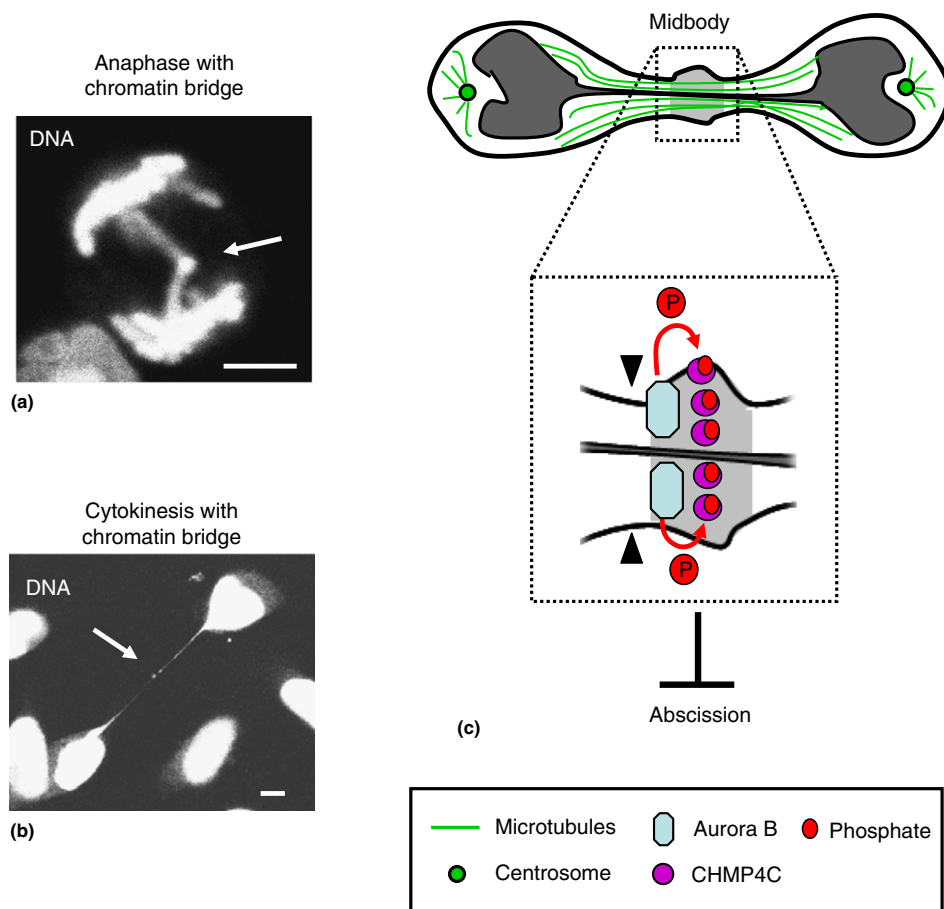


Figure 7 Aurora B delays abscission in the presence of chromatin bridges. (a, b) Examples of chromatin bridges in anaphase (a) or cytokinesis (b). Confocal microscopy images of DT40 avian β -lymphoma (a) or BE human colon carcinoma cells (b). Chromatin bridges are indicated by arrows. Bar, 5 μ m. (c) Model for the role of Aurora B in the abscission checkpoint (see text for details). Arrowheads indicate the abscission site. P, phosphorylation.

center of the midbody where it can stop abscission (**Figure 7 (c)**; [Carlton *et al.*, 2012](#)). However, exactly how this happens is a matter of active investigation.

Conclusions and Future Perspectives

Cytokinesis is a remarkably complex process that involves multiple components and regulatory pathways. Although many cytokinesis proteins have been identified, the exact mechanism of their function is under active investigation. Failure of cytokinesis can lead to tetraploidy and genomic instability, a prelude of carcinogenesis ([Fujiwara *et al.*, 2005](#)). Furthermore, pharmacologically induced cytokinesis failure in cancer cells may be important as an anticancer strategy and inhibitors of kinases that regulate cytokinesis, such as Plk1 and Aurora B, are being tested as anticancer drugs ([Petronczki *et al.*, 2008](#); [Gascoigne and Taylor, 2009](#)). Because inhibition of the above kinases is likely to induce cytokinesis failure in normal dividing cells in addition to cancer cells, those drugs are expected to induce myelosuppression ([Petronczki *et al.*, 2008](#); [Gascoigne and Taylor, 2009](#)). It therefore remains a challenge to identify conditions under which mitotic inhibitors are able to kill tumor cells while still being tolerated by normal proliferating cells, for example, in the bone marrow.

Acknowledgments

We thank Worldwide Cancer Research and Fondation Santé for financial support.

See also: Cell Division/Death: Cell Cycle: Mitosis in Animal Cells

References

- Barr, F.A., Gruneberg, U., 2007. Cytokinesis: Placing and making the final cut. *Cell* 131, 847–860.
- Bastos, R.N., Barr, F.A., 2010. Plk1 negatively regulates Cep55 recruitment to the midbody to ensure orderly abscission. *Journal of Cell Biology* 191, 751–760.
- Boucrot, E., Kirchhausen, T., 2007. Endosomal recycling controls plasma membrane area during mitosis. *Proceedings of the National Academy of Sciences of the United States of America* 104, 7939–7944.
- Brennan, I.M., Peters, U., Kapoor, T.M., Straight, A.F., 2007. Polo-like kinase controls vertebrate spindle elongation and cytokinesis. *PLoS One* 2, e409.
- Bringmann, H., Hyman, A.A., 2005. A cytokinesis furrow is positioned by two consecutive signals. *Nature* 436, 731–734.
- Burgess, D.R., Chang, F., 2005. Site selection for the cleavage furrow at cytokinesis. *Trends in Cell Biology* 15, 156–162.
- Burkard, M.E., Randall, C.L., Larochelle, S., *et al.*, 2007. Chemical genetics reveals the requirement for Polo-like kinase 1 activity in positioning RhoA and triggering cytokinesis in human cells. *Proceedings of the National Academy of Sciences of the United States of America* 104, 4383–4388.
- Canman, J.C., Cameron, L.A., Maddox, P.S., *et al.*, 2003. Determining the position of the cell division plane. *Nature* 424, 1074–1078.
- Capalbo, L., Montebault, E., Takeda, T., *et al.*, 2012. The chromosomal passenger complex controls the function of endosomal sorting complex required for transport-III Snf7 proteins during cytokinesis. *Open Biology* 2, 120070.
- Carlton, J.G., Caballe, A., Agromayor, M., Kloc, M., Martin-Serrano, J., 2012. ESCRT-III governs the Aurora B-mediated abscission checkpoint through CHMP4C. *Science* 336, 220–225.
- Chai, Y., Tian, D., Yang, Y., *et al.*, 2012. Apoptotic regulators promote cytokinetic midbody degradation in *C. elegans*. *Journal of Cell Biology* 199, 1047–1055.
- Dechant, R., Glotzer, M., 2003. Centrosome separation and central spindle assembly act in redundant pathways that regulate microtubule density and trigger cleavage furrow formation. *Developmental Cell* 4, 333–344.
- Dobbelaere, J., Barral, Y., 2004. Spatial coordination of cytokinetic events by compartmentalization of the cell cortex. *Science* 305, 393–396.
- Elia, N., Sougrat, R., Spurlin, T.A., Hurley, J.H., Lippincott-Schwartz, J., 2011. Dynamics of endosomal sorting complex required for transport (ESCRT) machinery during cytokinesis and its role in abscission. *Proceedings of the National Academy of Sciences of the United States of America* 108, 4846–4851.
- Fielding, A.B., Schonteich, E., Matheson, J., *et al.*, 2005. Rab11-FIP3 and FIP4 interact with Arf6 and the exocyst to control membrane traffic in cytokinesis. *EMBO Journal* 24, 3389–3399.
- Fujiwara, T., Bandi, M., Nitta, M., *et al.*, 2005. Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature* 437, 1043–1047.
- Gascoigne, K.E., Taylor, S.S., 2009. How do anti-mitotic drugs kill cancer cells? *Journal of Cell Science* 122, 2579–2585.
- Gisselsson, D., 2008. Classification of chromosome segregation errors in cancer. *Chromosoma* 117, 511–519.
- Glotzer, M., 2004. Cleavage furrow positioning. *Journal of Cell Biology* 164, 347–351.
- Goto, H., Yasui, Y., Kawajiri, A., *et al.*, 2003. Aurora-B regulates the cleavage furrow-specific vimentin phosphorylation in the cytokinetic process. *Journal of Biological Chemistry* 278, 8526–8530.
- Gromley, A., Jurczyk, A., Sillibourne, J., *et al.*, 2003. A novel human protein of the maternal centriole is required for the final stages of cytokinesis and entry into S phase. *Journal of Cell Biology* 161, 535–545.
- Gromley, A., Yeaman, C., Rosa, J., *et al.*, 2005. Centriolin anchoring of exocyst and SNARE complexes at the midbody is required for secretory-vesicle-mediated abscission. *Cell* 123, 75–87.
- Guizetti, J., Schermelleh, L., Mantler, J., *et al.*, 2011. Cortical constriction during abscission involves helices of ESCRT-III-dependent filaments. *Science* 331, 1616–1620.
- Henne, W.M., Buchkovich, N.J., Emr, S.D., 2011. The ESCRT pathway. *Developmental Cell* 21, 77–91.
- Hu, C.K., Coughlin, M., Mitchison, T.J., 2012. Midbody assembly and its regulation during cytokinesis. *Molecular Biology of the Cell* 23, 1024–1034.
- Hummer, S., Mayer, T.U., 2009. Cdk1 negatively regulates midzone localization of the mitotic kinesin Mklp2 and the chromosomal passenger complex. *Current Biology* 19, 607–612.
- Jiang, W., Jimenez, G., Wells, N.J., *et al.*, 1998. PRC1: A human mitotic spindle-associated CDK substrate protein required for cytokinesis. *Molecular Cell* 2, 877–885.
- Kaitna, S., Mendoza, M., Jantsch-Plunger, V., Glotzer, M., 2000. Incenp and an aurora-like kinase form a complex essential for chromosome segregation and efficient completion of cytokinesis. *Current Biology* 10, 1172–1181.
- Kawano, Y., Fukata, Y., Oshiro, N., *et al.*, 1999. Phosphorylation of myosin-binding subunit (MBS) of myosin phosphatase by Rho-kinase *in vivo*. *Journal of Cell Biology* 147, 1023–1038.
- Ke, Y., Huh, J.W., Warrington, R., *et al.*, 2011. PICH and BLM limit histone association with anaphase centromeric DNA threads and promote their resolution. *EMBO Journal* 30, 3309–3321.
- Kops, G.J., Weaver, B.A., Cleveland, D.W., 2005. On the road to cancer: Aneuploidy and the mitotic checkpoint. *Nature Reviews Cancer* 5, 773–785.
- Kosako, H., Yoshida, T., Matsumura, F., *et al.*, 2000. Rho-kinase/ROCK is involved in cytokinesis through the phosphorylation of myosin light chain and not ezrin/radixin/moesin proteins at the cleavage furrow. *Oncogene* 19, 6059–6064.
- Lee, H.H., Elia, N., Ghirlando, R., Lippincott-Schwartz, J., Hurley, J.H., 2008. Midbody targeting of the ESCRT machinery by a noncanonical coiled coil in CEP55. *Science* 322, 576–580.
- Low, S.H., Li, X., Miura, M., *et al.*, 2003. Syntaxin 2 and endobrevin are required for the terminal step of cytokinesis in mammalian cells. *Developmental Cell* 4, 753–759.
- Minoshima, Y., Kawashima, T., Hirose, K., *et al.*, 2003. Phosphorylation by aurora B converts MgcRacGAP to a RhoGAP during cytokinesis. *Developmental Cell* 4, 549–560.
- Mishima, M., Pavicic, V., Gruneberg, U., Nigg, E.A., Glotzer, M., 2004. Cell cycle regulation of central spindle assembly. *Nature* 430, 908–913.
- Montagnac, G., Echard, A., Chavrier, P., 2008. Endocytic traffic in animal cell cytokinesis. *Current Opinion in Cell Biology* 20, 454–461.

- Montagnac, G., Sibarita, J.B., Loubery, S., *et al.*, 2009. ARF6 Interacts with JIP4 to control a motor switch mechanism regulating endosome traffic in cytokinesis. *Current Biology* 19, 184–195.
- Murata-Hori, M., Fumoto, K., Fukuta, Y., *et al.*, 2000. Myosin II regulatory light chain as a novel substrate for AIM-1, an aurora/lpl1p-related kinase from rat. *Journal of Biochemistry* 128, 903–907.
- Neef, R., Gruneberg, U., Kopajtich, R., *et al.*, 2007. Choice of Plk1 docking partners during mitosis and cytokinesis is controlled by the activation state of Cdk1. *Nature Cell Biology* 9, 436–444.
- Neto, H., Collins, L.L., Gould, G.W., 2011. Vesicle trafficking and membrane remodeling in cytokinesis. *Biochemical Journal* 437, 13–24.
- Neto, H., Gould, G.W., 2011. The regulation of abscission by multi-protein complexes. *Journal of Cell Science* 124, 3199–3207.
- Norden, C., Mendoza, M., Dobbelaere, J., *et al.*, 2006. The NoCut pathway links completion of cytokinesis to spindle midzone function to prevent chromosome breakage. *Cell* 125, 85–98.
- Normand, G., King, R.W., 2010. Understanding cytokinesis failure. *Advances in Experimental Medicine and Biology* 676, 27–55.
- Petronczki, M., Glotzer, M., Kraut, N., Peters, J.M., 2007. Polo-like kinase 1 triggers the initiation of cytokinesis in human cells by promoting recruitment of the RhoGEF Ect2 to the central spindle. *Developmental Cell* 12, 713–725.
- Petronczki, M., Lenart, P., Peters, J.M., 2008. Polo on the rise—from mitotic entry to cytokinesis with Plk1. *Developmental Cell* 14, 646–659.
- Petsalaki, E., Dandoulaki, M., Morrice, N., Zachos, G., 2014. Chk1 protects against chromatin bridges by constitutively phosphorylating BLM serine 502 to inhibit BLM degradation. *Journal of Cell Science* 127, 3902–3908.
- Piekny, A., Werner, M., Glotzer, M., 2005. Cytokinesis: Welcome to the Rho zone. *Trends in Cell Biology* 15, 651–658.
- Pollard, T.D., 2007. Regulation of actin filament assembly by Arp2/3 complex and formins. *Annual Review of Biophysics and Biomolecular Structure* 36, 451–477.
- Potapova, T.A., Daum, J.R., Pittman, B.D., *et al.*, 2006. The reversibility of mitotic exit in vertebrate cells. *Nature* 440, 954–958.
- Satterwhite, L.L., Lohka, M.J., Wilson, K.L., *et al.*, 1992. Phosphorylation of myosin-II regulatory light chain by cyclin-p34cdc2: A mechanism for the timing of cytokinesis. *Journal of Cell Biology* 118, 595–605.
- Schiel, J.A., Park, K., Morphew, M.K., *et al.*, 2011. Endocytic membrane fusion and buckling-induced microtubule severing mediate cell abscission. *Journal of Cell Science* 124, 1411–1424.
- Skop, A.R., Liu, H., Yates 3rd, J., Meyer, B.J., Heald, R., 2004. Dissection of the mammalian midbody proteome reveals conserved cytokinesis mechanisms. *Science* 305, 61–66.
- Steigemann, P., Wurzenberger, C., Schmitz, M.H., *et al.*, 2009. Aurora B-mediated abscission checkpoint protects against tetraploidization. *Cell* 136, 473–484.
- van der Waal, M.S., Hengeveld, R.C., van der Horst, A., Lens, S.M., 2012. Cell division control by the chromosomal passenger complex. *Experimental Cell Research* 318, 1407–1420.
- Wang, Y.L., 2005. The mechanism of cortical ingression during early cytokinesis: Thinking beyond the contractile ring hypothesis. *Trends in Cell Biology* 15, 581–588.
- Werner, M., Munro, E., Glotzer, M., 2007. Astral signals spatially bias cortical myosin recruitment to break symmetry and promote cytokinesis. *Current Biology* 17, 1286–1297.
- Wolfe, B.A., Takaki, T., Petronczki, M., Glotzer, M., 2009. Polo-like kinase 1 directs assembly of the HsCdk-4 RhoGAP/Ect2 RhoGEF complex to initiate cleavage furrow formation. *PLoS Biology* 7, e1000110.
- Yamashiro, S., Totsukawa, G., Yamakita, Y., *et al.*, 2003. Citron kinase, a Rho-dependent kinase, induces di-phosphorylation of regulatory light chain of myosin II. *Molecular Biology of the Cell* 14, 1745–1756.
- Yang, D., Rismanchi, N., Renvoise, B., *et al.*, 2008. Structural basis for midbody targeting of spastin by the ESCRT-III protein CHMP1B. *Nature Structural & Molecular Biology* 15, 1278–1286.
- Yuce, O., Piekny, A., Glotzer, M., 2005. An ECT2-centralspindlin complex regulates the localization and function of RhoA. *Journal of Cell Biology* 170, 571–582.

Regulation of the p53 Pathway

CL Brooks, Stemline Therapeutics, Inc., New York, NY, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The p53 tumor suppressor gene has been called the ‘guardian of the genome’ and continues to exhibit complex and exquisite regulatory control over cellular functions as a sequence-specific transcription factor (Lane, 1992; Levine, 1997). Great progress has been made since it was first cloned and described in 1979, yet the knowledge gained to date paints an even more complex and intricate role of the gene. Recent findings have shown some surprising and unexpected functions of p53 in normal cells as well, which potentially have broad implications on therapeutics that target this pathway.

As a transcription factor, p53 is highly controlled through a multitude of transcriptional, translational, and posttranslational mechanisms. Normal homeostatic levels of the protein are predominantly regulated through the ubiquitin-proteasome pathway, which maintains p53 at low levels in a cell during normal, unstressed conditions. However, upon sensing of virtually all types of cellular stress, the protein levels of p53 are quickly stabilized and increased, and p53 is activated through numerous posttranslational modifications that can lead to specific outcomes. It has been postulated that the type and magnitude of the stress that a cell experiences guides the type of modifications that occur on p53 and the consequent downstream effect: low to moderate levels of stress can lead to specific acetylation events that drive the expression of genes related to cell growth arrest, thus providing a cell time to repair the damage and survive. However, high levels of stress, which often leads to irreparable damage, is thought to drive acetylation at distinct sites that leads to the expression of genes governing apoptosis and cell death. Nevertheless, this hypothesized model has become more complex with numerous recent findings of p53 involvement in metabolism and normal energy homeostasis.

This article will discuss recent findings into the mechanisms of p53 regulation and activation. It will highlight our current understanding of the p53 pathway and summarize the role of p53 in cellular metabolism. Although there remains much to be learned on this fascinating and complex gene, the package of findings have begun to provide a working model of how p53 governs cellular processes and identifies aspects of the pathway that we may exploit therapeutically. The goal of this section is to provide a larger and more ‘global’ view of the p53 pathway with guidance to recent reviews and work on specific aspects of the biology.

Control of p53 through the Ubiquitin-Proteasome Pathway

Ubiquitination

The cell maintains protein levels and recycles others in part through the efficient and well-controlled ubiquitin-

proteasome pathway. This enzymatically regulated process begins with an E1-activating enzyme, which binds the 76 amino acid ubiquitin and catalyzes acyl adenylation through an ATP-dependent reaction (Pickart, 2001). The ubiquitin moiety is then transferred to an E2-conjugating enzyme through an energy-driven reaction, which subsequently complexes with an E3 ubiquitin ligase. This latter class of proteins provides substrate specificity for proteins targeted for ubiquitination, and numerous E3 ligases have been described for p53 (Lee and Gu, 2010). The most well-known and highly characterized E3 ligase of p53 is mouse double minute 2 (Mdm2) (Hdm2 in humans), which is regulated by p53 through a negative feedback loop. During normal, unstressed conditions, p53 drives the expression of Mdm2; however, during cellular stress signals, the transcription of Mdm2 by p53 is abrogated, leading to an accumulation of the protein that is available for further activation modification. The Mdm2 knockout mouse underscores this intricate feedback relationship between the two proteins, as embryonic lethality is observed at day E6.5, but completely rescued in the double knockout of p53 and Mdm2 (Jones *et al.*, 1995; Montes de Oca Luna *et al.*, 1995; Sharpless, 2005; Sherr, 2001).

Abrogation of the p53–Mdm2 interaction during times of cellular stress can occur through a number of mechanisms, including phosphorylation at Ser15 and Ser20 by Chk1, Chk2, ataxia telangiectasia mutated (ATM), ATR, and DNA-PK (DNA-dependent protein kinase) (Kruse and Gu, 2009). This posttranslational event is thought to physically hinder protein–protein interactions. Acetylation of p53 at the C-terminus of the protein, which is a critical region for transactivation, has also been shown to block ubiquitination at shared lysine residues (Ito *et al.*, 2001; Li *et al.*, 2002a,b). Another mechanism for disruption of the p53–Mdm2 axis involves the tumor suppressor p14ARF, which is a nucleolar protein that binds to and sequesters Mdm2 from p53 in response to oncogenic stress (Sharpless and DePinho, 2004; Sherr, 2001). Moreover, this activity can occur both in and out of the nucleolus (Korgaonkar *et al.*, 2005; Sharpless, 2005). Mdm2 destabilization seems to be an obligatory step for sufficient p53 stabilization and activation (Ohkubo *et al.*, 2006). Cells undergoing a DNA damage response that were treated with proteasome inhibitors to block Mdm2 degradation failed to induce several p53 target genes. Mdm2 destabilization seems to involve DNA damage-induced kinases such as ATM, since treatment with the general kinase inhibitor Wortmannin blocked Mdm2 destabilization in response to stress. More intriguing, however, is the notion that reducing Mdm2 stability is as important as inhibiting its function during a DNA damage response. The half-life of Mdm2 is as short as 5 min during this period, and it is alluring to speculate that phosphorylated Mdm2 may have enhanced self-ubiquitination catalysis. It is possible that the abundance of ubiquitinated Mdm2 during DNA damage is incapable of interacting with p53 due to the sheer speed of degradation, as indicated by its short half-life.

Ubiquitin E3 ligase	Effect on p53
Mdm2	Degradation/nuclear translocation/transcriptional repression
MdmX	Transcriptional inhibition
Ubc13	Degradation
E4F1*	Transcriptional activation
ARF-BP1	Degradation
COP1	Degradation
WWP1	Nuclear translocation
PIRH2	Degradation
CHIP	Degradation
CARP1	Degradation
CARP2	Degradation
TOPORS	Degradation
P300/CBP*	Degradation
MSL2	Nuclear translocation
YY1*	Degradation
Gankyrin*	Degradation

Figure 1 Key ubiquitin E3 ligases of the p53 pathway. Proteins with an asterisk do not function as a classical E3 ligase; rather, these E4 ligases polyubiquitinate a monoubiquitinated substrate.

Additionally, Mdm2 can be degraded within the nucleus under these conditions and this may provide the avenue for quick degradation.

E3 Ligases in the p53 Pathway

A number of other E3 ligases have been described as having specificity for p53, including Ubc13, E4F1, ARF-BP1, COP1, WWP1, and Pirh2 (Lee and Gu, 2010; Figure 1). COP1 and Pirh2 both ubiquitinate p53 independently of Mdm2, and Topors can both ubiquitinate and sumoylate p53 as well (Rajendra *et al.*, 2004; Weger *et al.*, 2005; Dornan *et al.*, 2004; Leng *et al.*, 2003). ARF-BP1 was first identified as a p14ARF binding protein and subsequently shown to directly ubiquitinate p53 as well (Chen *et al.*, 2005; Zhong *et al.*, 2005). Importantly, however, this protein appears to have both p53-dependent and p53-independent activity. In addition, CARP1 and CARP2 ubiquitinate phosphorylated p53 in stressed cells at position S20 (Yang *et al.*, 2007). Taken together, the functions of the various E3 ligases with activity against p53 underscore the critical importance of tight regulation of p53 at all times in a cell and suggest that, although Mdm2 is a critical factor during embryonic development, it is likely not the only factor controlling p53 levels under normal and stressed conditions in mature cells.

Monoubiquitination versus Polyubiquitination

In addition to driving the regulation and degradation of p53 through the 26S proteasome, ubiquitin also has temporal and functional consequences for p53. Monoubiquitination, or the placement of a single ubiquitin or multiple single moieties at distinct lysine residues, has been shown to play an important role in the signaling and localization of several proteins,

including epidermal growth factor (EGF) receptor pathway substrate clone 15 (Eps15) as well as histones H2A and H2B (Ramanathan and Ye, 2012). This process can lead to events such as endocytosis, cell signaling, or transcriptional regulation. As p53 is a nuclear transcription factor, the process of monoubiquitination acts as a unique signal that drives p53 into the cytoplasm through nuclear export for various activities (Li *et al.*, 2003). Mdm2 is capable of inducing both monoubiquitination and polyubiquitination of p53, and monoubiquitination appears to be dependent on Mdm2 levels. Mechanistically, p53-mediated transcription of Mdm2 during times of low or no stress may provide yet another form of regulation for the protein as continuous monoubiquitination of p53 would lead to its translocation into the cytoplasm and away from transcriptional targets in the nucleus. Upon activation of a stress signal, this process would stop and lead to the accumulation of p53 protein levels in the nucleus for further modification and activation. Given the negative feedback loop between the two proteins, high levels of p53 would subsequently drive expression of increased levels of Mdm2, thus leading to the polyubiquitination of p53 and degradation to shut down activity once repairs are completed. Although a simplistic model, this delicate balance of protein levels and precise regulation of p53 activity and location may very well provide the regulatory foundation on which a plethora of other signaling cascades and interacting partners can act.

Cytoplasmic-bound p53 has also been shown to host a range of cellular activities that are independent of transcription. For example, p53 has been shown to bind to apoptotic factors of the Bcl family, including Bcl-2 and Bcl-XL, which mediate the oligomerization of Bak and Bax and play a role in the efflux of apoptotic triggers from mitochondria, including cytochrome *c* (Chipuk *et al.*, 2004; Mihara *et al.*, 2003; Tomita *et al.*, 2006). In addition, p53 has been shown to be further polyubiquitinated in the cytoplasm by p300/CBP thus promoting its degradation (Grossman *et al.*, 2003; Shi *et al.*, 2009). This mechanism of cytoplasmic p300/CBP is distinct from that of protein located in the nucleus, which acetylates p53 for transcriptional activation (see Section Transactivation of p53 through Posttranslational Modifications). Another protein, Synoviolin, is located in the endoplasmic reticulum (ER) and mediates ER-associated degradation (ERAD) of p53 (Yamasaki *et al.*, 2007). Similarly, the anti-apoptotic protein NAPA, which has been shown to promote resistance to the chemotherapeutic cisplatin in cancer cells, can decrease the ubiquitination-mediated degradation of Synoviolin and lead to a decrease in p53 levels (Wu *et al.*, 2011). Although evidence suggests a regulatory involvement of several ER proteins in controlling p53, the specific roles and net outcome of these interactions remain to be elucidated.

Deubiquitination

The ubiquitination pathway is highly dynamic in nature, and mechanisms exist that allow for full reversal of the process. Deubiquitination enzymes (DUBs) have been described in numerous cellular processes, and their role in the p53 pathway has important implications on overall p53 control (Hershko and Ciechanover, 1998; Laney and Hochstrasser, 1999; Pickart

and Eddins, 2004). There are approximately 95 DUBs in the human genome that are divided into five distinct classes: JAB1/MPN/MOV34 (JAMMs), ubiquitin C-terminal proteases (OTUs), Machado-Joseph disease (MJDs), ubiquitin C-terminal hydrolases (UCHs), and ubiquitin-specific proteases (USPs) (Pickart and Eddins, 2004). The first p53-specific DUB identified was Ubiquitin-Specific Protease 7 (USP7), which is also called Herpesvirus-specific Ubiquitin-Specific Protease (HAUSP). USP7 was shown to deubiquitinate and stabilize p53 protein (Li *et al.*, 2002a, 2004). However, soon thereafter it was shown that a USP7 HCT116 knockout cell line exhibited high levels of p53, suggesting a more complex interplay between these two proteins (Cummins *et al.*, 2004). It was subsequently shown that USP7 can also deubiquitinate Mdm2, which exhibits autoubiquitination activity and is a highly unstable protein (Li *et al.*, 2004). Removing USP7 leads to destabilization of Mdm2 and consequently the stabilization of p53, indicating the preferential interaction of USP7 with Mdm2. Tight control of p53 is critical in close proximity to transcriptional targets, yet USP7 has also been shown to localize near mitochondria where it can presumably deubiquitinate monoubiquitinated p53 for further functions required in the cytoplasm. It has also been shown that USP7 can be both ubiquitinated and neddylated, though the consequences of these posttranslational modifications remain unclear (Lee *et al.*, 2005). When N-terminal co-crystal structures were solved for USP7 together with Mdm2, p53, and EBNA peptides, USP7 was shown to have a binding preference for the consensus sequence motif P/AXXS (Sheng *et al.*, 2006). These three proteins all compete for a binding region within a pocket of the N-terminus MATH domain. The mutually exclusive manner of binding, particularly between USP7-p53 and USP7-Mdm2, has important implications in p53 pathway regulation, as discussed below. Interestingly, USP7 has more extensive interactions and forms a more stable complex with Mdm2 even in the presence of excess p53 (Hu *et al.*, 2006).

The p53–Mdm2–HAUSP Connection

The profound effect of DNA damage on Mdm2 temporal stability and the rapid stabilization of p53 suggest an intimate interplay between p53, Mdm2, and HAUSP. Since the half-life of Mdm2 drastically decreases upon stimulation with DNA damage reagents and the binding affinity between HAUSP and Mdm2 goes down upon stimulation with NCS, it is quite likely that HAUSP has a direct role in these observations. Under normal, unstressed, homeostatic cell conditions, HAUSP may preferentially bind to Mdm2 and facilitate its stability at a level sufficient for maintaining p53 at low levels. It has been shown that the binding affinity between Mdm2 and HAUSP is slightly higher ($K_d = 10 \mu\text{M}$) than that between p53 and HAUSP ($K_d = 8 \mu\text{M}$) (Sheng *et al.*, 2006). HAUSP also preferentially binds to Mdm2 *in vitro* when incubated with one molar equivalence of Mdm2 and 10-fold molar equivalence of p53, suggesting a stronger affinity for Mdm2 when the three proteins exist together (Hu *et al.*, 2002). Considering the fact that p53 and Mdm2 share the same binding site within the N-terminal domain of HAUSP, it is unlikely that these three proteins exist as a three-protein complex at the same time.

Rather, the resting state may have a shift in equilibrium in favor of HAUSP–Mdm2 complexes over HAUSP–p53 complexes. This shift would ensure relatively stable Mdm2 that would maintain p53 at low levels when not needed. Upon DNA damage and cellular stresses, however, this shift would be reversed and allow for two mechanisms to take place. First, the binding affinity between HAUSP and Mdm2 would decrease, allowing for free HAUSP molecules to interact with p53 and stabilize the protein by directly removing ubiquitin moieties. Second, the unbound Mdm2 would continue to undergo self-ubiquitination and degradation, shortening its half-life drastically and decreasing the available pool able to ubiquitinate p53. In support of this notion, when HAUSP is removed by somatic knockout in HCT116 cells or transiently reduced in stably transfected HAUSP siRNA cells, the half-life of Mdm2 decreases drastically with a concomitant increase in p53 levels. A similar response also occurs after DNA damage treatment. At an undefined point in the DNA damage response, if DNA repairs were made and signaling pathways allowed for it, the interaction between HAUSP and Mdm2 would be restored and p53 levels would then return to their resting levels.

DUBs in the p53 Pathway

Other DUBs can also affect p53 activity as well as other members of the pathway. USP2a specifically deubiquitinates Mdm2, but not p53, and promotes Mdm2-dependent p53 degradation (Stevenson *et al.*, 2007). Moreover, reduction of USP2a in cells also stabilized p53 and led to its subsequent activation, underscoring the important role of this DUB in regulating Mdm2. USP7 can also deubiquitinate MdmX, which is an Mdm2 homolog that aids the E3 ligase in regulating p53 (Gu *et al.*, 2002; Sharp *et al.*, 1999; Tanimura *et al.*, 1999). MdmX can also promote transcription of p53 by Mdm2, thus leading to the downregulation of the transcription factor (Biderman *et al.*, 2012). USP10 is a cytoplasmic DUB that is activated through ATM phosphorylation after DNA damage occurs, translocates to the nucleus, and deubiquitinates p53 to stabilize and activate the transcription factor (Yuan *et al.*, 2010). More recently, it has been shown that oxidative stress can promote the upregulation of USP29 transcription, which is a DUB that stabilizes p53 (Liu *et al.*, 2011). This activity subsequently leads to p53 activation and the induction of apoptosis. The first OTU family member was recently shown to specifically deubiquitinate p53. The OTU domain-containing ubiquitin aldehyde-binding protein 1 (Otub1) stabilizes and activates p53 to induce p53-dependent apoptosis and cell growth arrest (Sun *et al.*, 2012). However, the activity does not seem to be dependent on its catalytic domain based on mutational analyses. Rather, the protein was shown to bind to an E2-conjugating enzyme specific for Mdm2, UbcH5, thus suppressing Mdm2-mediated ubiquitination of p53 indirectly (Sun *et al.*, 2012). This observation is the first demonstration of a ‘noncanonical’ mechanism of p53 inhibition by a DUB (Sun and Dai, 2014). Taken together, the number of DUBs with specific activity against p53 underscores the importance of this process in balancing p53 protein levels through the ubiquitin-proteasome pathway and reversing the signals marking the protein for degradation.

Transactivation of p53 through Posttranslational Modifications

Acetylation

Acetylation has been well described as a posttranslational event that regulates histones and protein translation with precision, but p53 was the first nonhistone protein shown to be acetylated as well (Gu and Roeder, 1997). The level of p53 acetylation has a direct correlation with protein stabilization and stress-induced activation (Barlev *et al.*, 2001; Ito *et al.*, 2001; Knights *et al.*, 2006; Li *et al.*, 2007; Luo *et al.*, 2004; Rodriguez *et al.*, 2000) and this process stimulates sequence-specific DNA binding (Gu and Roeder, 1997; Liu *et al.*, 1999; Luo *et al.*, 2004; Sakaguchi *et al.*, 1998; Zhao *et al.*, 2006). A number of HATs can acetylate p53 at C-terminal lysine residues (positions 120, 164, 305, 370, 372, 373, 381, and 382), including p300/CREB-binding protein (CBP), HAT p300/CBP-associated factor (PCAF), MYST, hMOF, and TIP60 (Gu and Roeder, 1997; Liu *et al.*, 1999; Sakaguchi *et al.*, 1998; Sykes *et al.*, 2006; Tang *et al.*, 2006; Wang *et al.*, 2003). Acetylation at C-terminal residues also promotes the recruitment and binding of additional transcription cofactors, and it has been shown that some HATs can enhance the acetylation and transcription of proximal target genes (Goodman and Smolik, 2000).

The finding that p53 could be acetylated on a lysine located outside of the C-terminal region (K120) was the first indication that this process may have other consequences than transactivation. The hMOF and TIP60 HATs, which are part of the MYST family, acetylate this unique residue but do not have an effect on DNA binding (Sykes *et al.*, 2006; Tang *et al.*, 2006). Rather, acetylation of K120 was shown to induce the activation of proapoptotic proteins Puma and Bax, which provides a distinct mechanism for inducing apoptosis. The p53-mediated induction of apoptosis through this process occurs after DNA damage signals or p14ARF activation, and preventing acetylation at the site (through introduction of an arginine in place of the lysine) blocked the ability of p53 to induce apoptosis but not cell growth arrest (Tang *et al.*, 2006). These data suggest that acetylation of lysine 120 is crucial for the differential activation of p53 target genes that promote either an apoptosis or cell growth arrest pathway.

Recently, a new acetylation site was described for p53 (Tang *et al.*, 2008). Lysine 164 was identified as a p300/CBP acetylation site by a mass spectrometric analysis of all acetylation sites of p53. Interestingly, mutation of K164 together with seven other C-terminal acetylation sites (creating a p53-8KR expression construct) completely blocked the induction of p21-mediated cell growth arrest. Importantly, single mutations introduced into these sites did not have an overall effect on p53 transactivation activity, suggesting that some compensation can occur by other acetylation sites. However, the combination of all eight sites may be required for specific activation of downstream events. Together, these results underscore the importance of acetylation as an indispensable posttranslational modification for p53 activation.

Deacetylation

Acetylation is a critical process that regulates p53 activity, and as such deacetylation mechanisms exist to tightly control activity.

It has been well established that histone deacetylases (HDACs) can reverse p53 acetylation, including HDAC1 and the class III NAD⁺-dependent SIRT1 (Luo *et al.*, 2001, 2000; Vaziri *et al.*, 2001). Interactions between p53 and these HDACs reduce p53 acetylation levels, leading to termination of p53-dependent transcription of target genes and subsequent cell growth arrest. In addition, the class III NAD⁺-dependent deacetylase SIRT1 was shown to deacetylate p53 (Luo *et al.*, 2001; Vaziri *et al.*, 2001). SIRT1-mediated deacetylation of p53 inactivates the sequence-specific transcriptional activity of the protein and represses p53-mediated cell growth arrest and apoptosis in response to DNA damage and oxidative stress. This process was also shown to prevent p53-dependent transactivation of p21 (Luo *et al.*, 2001). SIRT has been shown to play a role in several other cellular processes in addition to the p53 pathway, which range from metabolism to cellular senescence and aging (Bordone and Guarente, 2005). Although SIRT1 negatively regulates p53 transcriptional activity, the knockout of Sir2a, which is the mouse homolog of SIRT1, shows defects in p53 function in embryonic fibroblasts (Wang *et al.*, 2008). Although p53 is hyperacetylated in these cells, they exhibit defects in DNA repair ability, suggesting that deacetylation and SIRT1 may play additional roles in the p53 pathway. Indeed, SIRT1 is known to be involved in cellular senescence and aging (Bordone and Guarente, 2005). From a survival point of view, these mechanisms may be quite important in response to low levels of cell stress and DNA damage by providing the cell time to make repairs and continue to survive. Thus, deacetylation of p53 by SIRT1 in cells that are experiencing chronic exposure to low levels of damage may push cells into senescence in these mice; these cells have an increased capacity to proliferate in response to chronic, sublethal doses of oxidative stress and fail to mount an adequate DNA damage response (Cheng *et al.*, 2003; Chua *et al.*, 2005). Moreover, SIRT1 has been shown to relocate to sites of DNA double-stranded breaks to help promote DNA repair and can deacetylate the DNA repair protein NBS1 (Yuan *et al.*, 2007). Embryonic fibroblasts from Sirt1-null mice show defects in DNA repair, despite a high steady-state level of acetylated p53. These findings, together with the seemingly opposing results of p53 regulation, suggest that SIRT1 and deacetylation may have an important role in p53-mediated cellular senescence (Tyner *et al.*, 2002; Wang *et al.*, 2008). It is well established that SIRT1 has critical functions in longevity and aging in other organisms such as yeast (Bordone and Guarente, 2005). SIRT1-mediated deacetylation of p53 in response to low levels of DNA damage may therefore promote cellular senescence and help cellular recovery from nonlethal genotoxic events (Wang *et al.*, 2008). However, additional exploration of SIRT1-mediated deacetylation and the net effect on p53 is needed, as the HDAC does not fit the classical definition of a tumor suppressor and mutations in the gene have not been identified to date (Huang *et al.*, 2008; Luo *et al.*, 2001). Nevertheless, current data suggest that SIRT1 may provide a link between the p53 pathway and cellular longevity.

Involvement of p53 in Metabolism

Research of p53 function and regulation has intensely focused on cell growth arrest and apoptosis in response to a multitude

of damaging hits that a cell experiences, which is driven by key p53 responsive genes such as p21, Bax, and Puma. Yet, knockout mice lacking p21 and Puma do not show any increase in the incidence of tumorigenesis, suggesting that p53 may have other functions for maintaining normal cellular homeostasis likely involving metabolic pathways (Choudhury *et al.*, 2007; Michalak *et al.*, 2008; Valente *et al.*, 2013). Through the series of events that occur in dysplastic and neoplastic cells, numerous metabolic pathways are altered that allow a cell to adapt to the increased need for energy and other factors to promote increased proliferation. Indeed, cancer cells are known to have dysregulation of aerobic glycolysis and biosynthesis of macromolecules (Cairns *et al.*, 2011). Notably, p53 has been shown to counteract many of the metabolic changes occurring in tumor cells, such as upregulating mitochondrial OXPHOS and promoting fatty acid oxidation (Maddocks and Vousden, 2011). A recent study of mice carrying arginine mutations in place of lysines at key acetylation sites on p53 (K117, 161, and 162) showed a lack of functions for inducing apoptosis and cell growth arrest, yet still maintained the ability to induce several downstream metabolic targets (Li *et al.*, 2012). This finding indicates that distinct from the well-accepted tumor-suppressive functions through cell growth and apoptosis, p53 has the ability to regulate metabolic pathways as well.

Several metabolic changes occur within a tumor cell that allow it to survive at increased rates of cell proliferation, including an increased glycolytic rate, antioxidant responses, lipid metabolism, and nucleotide biosynthesis (Cairns *et al.*, 2011;

Hsu and Sabatini, 2008). p53 is known to mediate an increase in mitochondrial oxidative phosphorylation by activating Sco2, Gls2, and Parkin (Bensaad *et al.*, 2006; Hu *et al.*, 2010; Kawauchi *et al.*, 2008; Matoba *et al.*, 2006; Schwartzberg-Bar-Yoseph *et al.*, 2004; Suzuki *et al.*, 2010; Zhang *et al.*, 2011). These three genes each play a role in mediating this process: Sco2 increases respiration by regulating cytochrome *c* oxidase within the electron transport chain, Parkin upregulates pyruvate dehydrogenase E1a1 thereby increasing acetyl-CoA levels, and Gls2 can convert glutamine to glutamate as a substrate for glutathione synthesis. More recently, p53 was shown to induce the expression of Tigar, which functions to decrease aerobic glycolysis (Bensaad *et al.*, 2006). Together, these p53-regulated genes promote oxidative phosphorylation and inhibit aerobic glycolysis, which blocks the primary methods of energy utilization in tumor cells. Interestingly, other evidence has shown that Tigar has tumor-protective functions as well by preventing reactive oxygen species (ROS) build up and upregulating nucleotide synthesis (Cheung *et al.*, 2013). In addition, one study found that Tigar promoted the survival of glioma cells under hypoxic and glucose starvation conditions (Wanka *et al.*, 2012). These results raise an interesting paradox, since p53-mediated induction of Tigar may aid tumor cells in adapting to increased metabolic needs and promoting survival. However, the same pathways that promote the survival of tumor cells also aid in protecting normal cells from metabolic stress, and therefore there is likely a fine balance and transitional range when p53 switches from promoting metabolic efficiency to induction of tumor-suppressive functions (Wang and Gu, 2014). This can

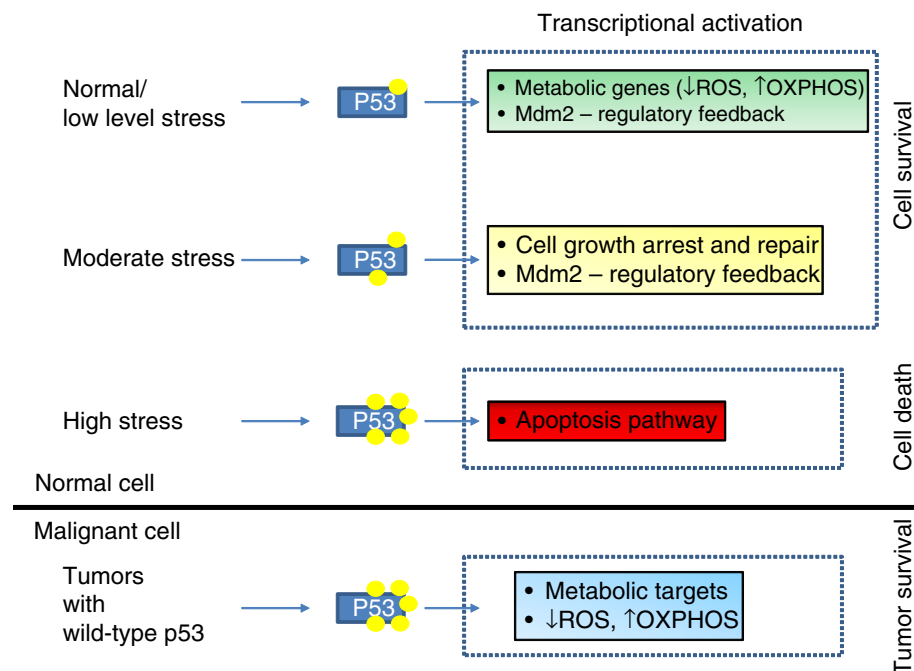


Figure 2 Proposed model of p53 regulation. Under normal or low stress conditions, p53 is activated through specific posttranslational modifications that promote the transactivation of metabolic genes and Mdm2, which serves as a regulator in a negative feedback loop. Upon moderate levels of DNA damage that require repair, further modifications of p53, and particularly acetylation at K164 and C-terminal residues, drive the expression of genes that induce cell growth arrest (p21) that allow the cell to repair damage and survive. However, when damage is too extensive for repair, p53 is further posttranslationally modified (K120) to commit the cell to an apoptosis pathway. In cells that have undergone malignant transformation, it is possible that a functional p53 could promote a metabolically favorable environment for the tumor cell to survive.

also be viewed as a 'hierarchy' of response depending on the type and extent of damage a cell is experiencing (Figure 2). When a cell experiences low levels of DNA damage or cellular stress, such as hypoxic conditions, p53 may promote the expression of various metabolic genes as well as Mdm2, which balances active p53 levels below a threshold that would otherwise trigger cell growth arrest or apoptosis. A moderate level of cellular stress, however, may induce acetylation at specific sites of p53 that drive the protein to induce cell growth arrest through p21 induction. Severe and irreparable damage would induce additional acetylation and other posttranslational events that would shift the p53 response down an apoptotic pathway (Wang and Gu, 2014). Further work is needed to understand the precise mechanisms by which p53 regulates metabolic genes and the specific ways in which it transitions from promoting metabolic pathways to inducing tumor-suppressive pathways.

Conclusions

The knowledge base of p53 regulation continues to evolve, and it is becoming evident that this critical tumor suppressor protein has an impact on numerous cellular pathways, including cell growth arrest, apoptosis, and metabolism. *In vivo* studies have indicated that numerous posttranslational modifications are required for p53 to be fully activated, and the precise modifications that occur may dictate the downstream activation of particular pathways. Acetylation is a process that is clearly required for p53 activation, yet we are still elucidating the exact mechanisms by which p53 is activated and how particular acetylation events drive its activity. In addition, the involvement and control of cellular metabolic pathways by p53 has clear implications on normal and tumor cell growth, and therefore a greater understanding of these events and how these processes can be exploited therapeutically is needed. Several human clinical trials of compounds targeting factors of the p53 pathway are currently underway, and therefore future findings will likely provide even greater insight into the impact of this pathway on tumorigenesis.

See also: Cell Communication: Prototypic Integrative Processes: Cancer Cell Invasion through Tissue Barriers

References

- Barlev, N.A., Liu, L., Chehab, N.H., *et al.*, 2001. Acetylation of p53 activates transcription through recruitment of coactivators/histone acetyltransferases. *Molecular Cell* 8, 1243–1254.
- Bensaad, K., Tsuruta, A., Selak, M.A., *et al.*, 2006. TIGAR, a p53-inducible regulator of glycolysis and apoptosis. *Cell* 126, 107–120.
- Biderman, L., Poyurovsky, M.V., Assia, Y., Manley, J.L., Prives, C., 2012. MdmX is required for p53 interaction with and full induction of the Mdm2 promoter after cellular stress. *Molecular and Cellular Biology* 32, 1214–1225.
- Bordone, L., Guarente, L., 2005. Calorie restriction, SIRT1 and metabolism: Understanding longevity. *Nature Reviews Molecular Cell Biology* 6, 298–305.
- Cairns, R.A., Harris, I.S., Mak, T.W., 2011. Regulation of cancer cell metabolism. *Nature Reviews Cancer* 11, 85–95.
- Chen, D., Kon, N., Li, M., *et al.*, 2005. ARF-BP1/Mule is a critical mediator of the ARF tumor suppressor. *Cell* 121, 1071–1083.
- Cheng, H.L., Mostoslavsky, R., Saito, S., *et al.*, 2003. Developmental defects and p53 hyperacetylation in Sir2 homolog (SIRT1)-deficient mice. *Proceedings of the National Academy of Sciences of the United States of America* 100, 10794–10799.
- Cheung, E.C., Athineos, D., Lee, P., *et al.*, 2013. TIGAR is required for efficient intestinal regeneration and tumorigenesis. *Developmental Cell* 25, 463–477.
- Chipuk, J.E., Kuwana, T., Bouchier-Hayes, L., *et al.*, 2004. Direct activation of Bax by p53 mediates mitochondrial membrane permeabilization and apoptosis. *Science* 303, 1010–1014.
- Choudhury, A.R., Ju, Z., Djojicubroto, M.W., *et al.*, 2007. Cdkn1a deletion improves stem cell function and lifespan of mice with dysfunctional telomeres without accelerating cancer formation. *Nature Genetics* 39, 99–105.
- Chua, K.F., Mostoslavsky, R., Lombard, D.B., *et al.*, 2005. Mammalian SIRT1 limits replicative life span in response to chronic genotoxic stress. *Cell Metabolism* 2, 67–76.
- Cummins, J.M., Rago, C., Kohli, M., *et al.*, 2004. Tumour suppression: Disruption of HAUSP gene stabilizes p53. *Nature* 428.1 p following 486.
- Dornan, D., Wertz, I., Shimizu, H., *et al.*, 2004. The ubiquitin ligase COP1 is a critical negative regulator of p53. *Nature* 429, 86–92.
- Goodman, R.H., Smolik, S., 2000. CBP/p300 in cell growth, transformation, and development. *Genes & Development* 14, 1553–1577.
- Grossman, S.R., Deato, M.E., Brignone, C., *et al.*, 2003. Polyubiquitination of p53 by a ubiquitin ligase activity of p300. *Science* 300, 342–344.
- Gu, J., Kawai, H., Nie, L., *et al.*, 2002. Mutual dependence of MDM2 and MDMX in their functional inactivation of p53. *Journal of Biological Chemistry* 277, 19251–19254.
- Gu, W., Roeder, R.G., 1997. Activation of p53 sequence-specific DNA binding by acetylation of the p53 C-terminal domain. *Cell* 90, 595–606.
- Hershko, A., Ciechanover, A., 1998. The ubiquitin system. *Annual Review of Biochemistry* 67, 425–479.
- Hsu, P.P., Sabatini, D.M., 2008. Cancer cell metabolism: Warburg and beyond. *Cell* 134, 703–707.
- Hu, M., Gu, L., Li, M., *et al.*, 2006. Structural basis of competitive recognition of p53 and MDM2 by HAUSP/USP7: Implications for the regulation of the p53-MDM2 pathway. *PLOS Biology* 4, e27.
- Hu, M., Li, P., Li, M., *et al.*, 2002. Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* 111, 1041–1054.
- Hu, W., Zhang, C., Wu, R., *et al.*, 2010. Glutaminase 2, a novel p53 target gene regulating energy metabolism and antioxidant function. *Proceedings of the National Academy of Sciences of the United States of America* 107, 7455–7460.
- Huang, J., Gan, Q., Han, L., *et al.*, 2008. SIRT1 overexpression antagonizes cellular senescence with activated ERK/S6k1 signaling in human diploid fibroblasts. *PLoS One* 3, e1710.
- Ito, A., Lai, C.H., Zhao, X., *et al.*, 2001. p300/CBP-mediated p53 acetylation is commonly induced by p53-activating agents and inhibited by MDM2. *EMBO Journal* 20, 1331–1340.
- Jones, S.N., Roe, A.E., Donehower, L.A., Bradley, A., 1995. Rescue of embryonic lethality in Mdm2-deficient mice by absence of p53. *Nature* 378, 206–208.
- Kawauchi, K., Araki, K., Tobiume, K., Tanaka, N., 2008. p53 Regulates glucose metabolism through an IKK-NF-kappaB pathway and inhibits cell transformation. *Nature Cell Biology* 10, 611–618.
- Knight, C.D., Catania, J., Di Giovanni, S., *et al.*, 2006. Distinct p53 acetylation cassettes differentially influence gene-expression patterns and cell fate. *Journal of Cell Biology* 173, 533–544.
- Korgaonkar, C., Hagen, J., Tompkins, V., *et al.*, 2005. Nucleophosmin (B23) targets ARF to nucleoli and inhibits its function. *Molecular and Cellular Biology* 25, 1258–1271.
- Kruse, J.P., Gu, W., 2009. Modes of p53 regulation. *Cell* 137, 609–622.
- Lane, D.P., 1992. Cancer. p53, guardian of the genome. *Nature* 358, 15–16.
- Laney, J.D., Hochstrasser, M., 1999. Substrate targeting in the ubiquitin system. *Cell* 97, 427–430.
- Lee, H.J., Kim, M.S., Kim, Y.K., Oh, Y.K., Baek, K.H., 2005. HAUSP, a deubiquitinating enzyme for p53, is polyubiquitinated, polyubiquitinated, and dimerized. *FEBS Letters* 579, 4867–4872.
- Lee, J.T., Gu, W., 2010. The multiple levels of regulation by p53 ubiquitination. *Cell Death & Differentiation* 17, 86–92.
- Leng, R.P., Lin, Y., Ma, W., *et al.*, 2003. Pirh2, a p53-induced ubiquitin-protein ligase, promotes p53 degradation. *Cell* 112, 779–791.
- Levine, A.J., 1997. p53, the cellular gatekeeper for growth and division. *Cell* 88, 323–331.

- Li, A.G., Piluso, L.G., Cai, X., *et al.*, 2007. An acetylation switch in p53 mediates holo-TFIID recruitment. *Molecular Cell* 28, 408–421.
- Li, M., Brooks, C.L., Kon, N., Gu, W., 2004. A dynamic role of HAUSP in the p53-Mdm2 pathway. *Molecular Cell* 13, 879–886.
- Li, M., Brooks, C.L., Wu-Baer, F., *et al.*, 2003. Mono- versus polyubiquitination: Differential control of p53 fate by Mdm2. *Science* 302, 1972–1975.
- Li, M., Chen, D., Shiloh, A., *et al.*, 2002a. Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nature* 416, 648–653.
- Li, M., Luo, J., Brooks, C.L., Gu, W., 2002b. Acetylation of p53 inhibits its ubiquitination by Mdm2. *Journal of Biological Chemistry* 277, 50607–50611.
- Li, T., Kon, N., Jiang, L., *et al.*, 2012. Tumor suppression in the absence of p53-mediated cell-cycle arrest, apoptosis, and senescence. *Cell* 149, 1269–1283.
- Liu, J., Chung, H.J., Vogt, M., *et al.*, 2011. JTV1 co-activates FBP to induce USP29 transcription and stabilize p53 in response to oxidative stress. *EMBO Journal* 30, 846–858.
- Liu, L., Scolnick, D.M., Trievel, R.C., *et al.*, 1999. p53 Sites acetylated *in vitro* by PCAF and p300 are acetylated *in vivo* in response to DNA damage. *Molecular and Cellular Biology* 19, 1202–1209.
- Luo, J., Li, M., Tang, Y., *et al.*, 2004. Acetylation of p53 augments its site-specific DNA binding both *in vitro* and *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America* 101, 2259–2264.
- Luo, J., Nikolaev, A.Y., Imai, S., *et al.*, 2001. Negative control of p53 by Sir2alpha promotes cell survival under stress. *Cell* 107, 137–148.
- Luo, J., Su, F., Chen, D., Shiloh, A., Gu, W., 2000. Deacetylation of p53 modulates its effect on cell growth and apoptosis. *Nature* 408, 377–381.
- Maddocks, O.D., Vousden, K.H., 2011. Metabolic regulation by p53. *Journal of Molecular Medicine (Berlin)* 89, 237–245.
- Matoba, S., Kang, J.G., Patino, W.D., *et al.*, 2006. p53 Regulates mitochondrial respiration. *Science* 312, 1650–1653.
- Michalak, E.M., Villunger, A., Adams, J.M., Strasser, A., 2008. In several cell types tumour suppressor p53 induces apoptosis largely via Puma but Noxa can contribute. *Cell Death & Differentiation* 15, 1019–1029.
- Mihara, M., Erster, S., Zaika, A., *et al.*, 2003. p53 has a direct apoptogenic role at the mitochondria. *Molecular Cell* 11, 577–590.
- Montes de Oca Luna, R., Wagner, D.S., Lozano, G., 1995. Rescue of early embryonic lethality in mdm2-deficient mice by deletion of p53. *Nature* 378, 203–206.
- Ohkubo, S., Tanaka, T., Taya, Y., Kitazato, K., Prives, C., 2006. Excess HDM2 impacts cell cycle and apoptosis and has a selective effect on p53-dependent transcription. *Journal of Biological Chemistry* 281, 16943–16950.
- Pickart, C.M., 2001. Mechanisms underlying ubiquitination. *Annual Review of Biochemistry* 70, 503–533.
- Pickart, C.M., Eddins, M.J., 2004. Ubiquitin: Structures, functions, mechanisms. *Biochimica et Biophysica Acta* 1695, 55–72.
- Rajendra, R., Malegaonkar, D., Pungalija, P., *et al.*, 2004. Topors functions as an E3 ubiquitin ligase with specific E2 enzymes and ubiquitinates p53. *Journal of Biological Chemistry* 279, 36440–36444.
- Ramanathan, H.N., Ye, Y., 2012. Cellular strategies for making monoubiquitin signals. *Critical Reviews in Biochemistry and Molecular Biology* 47, 17–28.
- Rodriguez, M.S., Desterro, J.M., Lain, S., Lane, D.P., Hay, R.T., 2000. Multiple C-terminal lysine residues target p53 for ubiquitin-proteasome-mediated degradation. *Molecular and Cellular Biology* 20, 8458–8467.
- Sakaguchi, K., Herrera, J.E., Saito, S., *et al.*, 1998. DNA damage activates p53 through a phosphorylation-acetylation cascade. *Genes & Development* 12, 2831–2841.
- Schwartzberg-Bar-Yoseph, F., Armoni, M., Karnieli, E., 2004. The tumor suppressor p53 down-regulates glucose transporters GLUT1 and GLUT4 gene expression. *Cancer Research* 64, 2627–2633.
- Sharp, D.A., Kratowicz, S.A., Sank, M.J., George, D.L., 1999. Stabilization of the MDM2 oncoprotein by interaction with the structurally related MDMX protein. *Journal of Biological Chemistry* 274, 38189–38196.
- Sharpless, N.E., 2005. INK4a/ARF: A multifunctional tumor suppressor locus. *Mutation Research* 576, 22–38.
- Sharpless, N.E., DePinho, R.A., 2004. Telomeres, stem cells, senescence, and cancer. *Journal of Clinical Investigation* 113, 160–168.
- Sheng, Y., Saridakis, V., Sarkari, F., *et al.*, 2006. Molecular recognition of p53 and MDM2 by USP7/HAUSP. *Nature Structural & Molecular Biology* 13, 285–291.
- Sherr, C.J., 2001. The INK4a/ARF network in tumour suppression. *Nature Reviews Molecular Cell Biology* 2, 731–737.
- Shi, D., Pop, M.S., Kulikov, R., *et al.*, 2009. CBP and p300 are cytoplasmic E4 polyubiquitin ligases for p53. *Proceedings of the National Academy of Sciences of the United States of America* 106, 16275–16280.
- Stevenson, L.F., Sparks, A., Allende-Vega, N., *et al.*, 2007. The deubiquitinating enzyme USP2a regulates the p53 pathway by targeting Mdm2. *EMBO Journal* 26, 976–986.
- Sun, X.X., Challagundla, K.B., Dai, M.S., 2012. Positive regulation of p53 stability and activity by the deubiquitinating enzyme Otubain 1. *EMBO Journal* 31, 576–592.
- Sun, X.X., Dai, M.S., 2014. Deubiquitinating enzyme regulation of the p53 pathway: A lesson from Otub1. *World Journal of Biological Chemistry* 5, 75–84.
- Suzuki, S., Tanaka, T., Poyurovsky, M.V., *et al.*, 2010. Phosphate-activated glutaminase (GLS2), a p53-inducible regulator of glutamine metabolism and reactive oxygen species. *Proceedings of the National Academy of Sciences of the United States of America* 107, 7461–7466.
- Sykes, S.M., Mellert, H.S., Holbert, M.A., *et al.*, 2006. Acetylation of the p53 DNA-binding domain regulates apoptosis induction. *Molecular Cell* 24, 841–851.
- Tang, Y., Luo, J., Zhang, W., Gu, W., 2006. Tip60-dependent acetylation of p53 modulates the decision between cell-cycle arrest and apoptosis. *Molecular Cell* 24, 827–839.
- Tang, Y., Zhao, W., Chen, Y., Zhao, Y., Gu, W., 2008. Acetylation is indispensable for p53 activation. *Cell* 133, 612–626.
- Tanimura, S., Ohtsuka, S., Mitsui, K., *et al.*, 1999. MDM2 interacts with MDMX through their RING finger domains. *FEBS Letters* 447, 5–9.
- Tomita, Y., Marchenko, N., Erster, S., *et al.*, 2006. WT p53, but not tumor-derived mutants, bind to Bcl2 via the DNA binding domain and induce mitochondrial permeabilization. *Journal of Biological Chemistry* 281, 8600–8606.
- Tyner, S.D., Venkatachalam, S., Choi, J., *et al.*, 2002. p53 Mutant mice that display early ageing-associated phenotypes. *Nature* 415, 45–53.
- Valente, L.J., Gray, D.H., Michalak, E.M., *et al.*, 2013. p53 Efficiently suppresses tumor development in the complete absence of its cell-cycle inhibitory and proapoptotic effectors p21, Puma, and Noxa. *Cell Reports* 3, 1339–1345.
- Vaziri, H., Dessain, S.K., Ng Eaton, E., *et al.*, 2001. hSIR2(SIRT1) functions as an NAD-dependent p53 deacetylase. *Cell* 107, 149–159.
- Wang, R.H., Sengupta, K., Li, C., *et al.*, 2008. Impaired DNA damage response, genome instability, and tumorigenesis in SIRT1 mutant mice. *Cancer Cell* 14, 312–323.
- Wang, S.J., Gu, W., 2014. To be, or not to be: Functional dilemma of p53 metabolic regulation. *Current Opinion in Oncology* 26, 78–85.
- Wang, Y.H., Tsay, Y.G., Tan, B.C., Lo, W.Y., Lee, S.C., 2003. Identification and characterization of a novel p300-mediated p53 acetylation site, lysine 305. *Journal of Biological Chemistry* 278, 25568–25576.
- Wanka, C., Steinbach, J.P., Rieger, J., 2012. Tp53-induced glycolysis and apoptosis regulator (TIGAR) protects glioma cells from starvation-induced cell death by up-regulating respiration and improving cellular redox homeostasis. *Journal of Biological Chemistry* 287, 33436–33446.
- Weger, S., Hammer, E., Heilbronn, R., 2005. Topors acts as a SUMO-1 E3 ligase for p53 *in vitro* and *in vivo*. *FEBS Letters* 579, 5007–5012.
- Wu, Z.Z., Sun, N.K., Chien, K.Y., Chao, C.C., 2011. Silencing of the SNARE protein NAPA sensitizes cancer cells to cisplatin by inducing ERK1/2 signaling, synoviolin ubiquitination and p53 accumulation. *Biochemical Pharmacology* 82, 1630–1640.
- Yamasaki, S., Yagishita, N., Sasaki, T., *et al.*, 2007. Cytoplasmic destruction of p53 by the endoplasmic reticulum-resident ubiquitin ligase 'Synoviolin'. *EMBO Journal* 26, 113–122.
- Yang, W., Rozan, L.M., McDonald 3rd, E.R., *et al.*, 2007. CARPs are ubiquitin ligases that promote MDM2-independent p53 and phospho-p53ser20 degradation. *Journal of Biological Chemistry* 282, 3273–3281.
- Yuan, J., Luo, K., Zhang, L., Cheville, J.C., Lou, Z., 2010. USP10 regulates p53 localization and stability by deubiquitinating p53. *Cell* 140 (3), 384–396.
- Yuan, Z., Zhang, X., Sengupta, N., Lane, W.S., Seto, E., 2007. SIRT1 regulates the function of the Nijmegen breakage syndrome protein. *Molecular Cell* 27, 149–162.
- Zhang, C., Lin, M., Wu, R., *et al.*, 2011. Parkin, a p53 target gene, mediates the role of p53 in glucose metabolism and the Warburg effect. *Proceedings of the National Academy of Sciences of the United States of America* 108, 16259–16264.
- Zhao, Y., Lu, S., Wu, L., *et al.*, 2006. Acetylation of p53 at lysine 373/382 by the histone deacetylase inhibitor depsipeptide induces expression of p21(Waf1/Cip1). *Molecular and Cellular Biology* 26, 2782–2790.
- Zhong, Q., Gao, W., Du, F., Wang, X., 2005. Mule/ARF-BP1, a BH3-only E3 ubiquitin ligase, catalyzes the polyubiquitination of Mcl-1 and regulates apoptosis. *Cell* 121, 1085–1095.

Cellular Senescence

S Da Silva-Álvarez and M Collado, IDIS, Clinical University Hospital (CHUS), SERGAS, Santiago de Compostela, Spain

© 2016 Elsevier Inc. All rights reserved.

Introduction

Cellular senescence is a state of permanent cell cycle arrest that was initially defined for cells grown in cell culture. Later on, this cell response was identified both *in vitro* and *in vivo* for cells subjected to different forms of stress, and more recently it has also been involved in physiological situations during development.

The initial observation of cell senescence was done for human primary fibroblasts while trying to establish the optimal conditions of cell culture (Hayflick and Moorhead, 1961). The aim was to have access to a safe source of human unperturbed cells that could be used for vaccine development and other biotechnological applications. At that time, early cell culture technology relied on the use of established cell lines derived from human and mouse tumors. In contrast to the robust and permanent proliferation observed for the tumor cells, human primary fibroblasts could be adapted to *in vitro* conditions of cell culture and proliferated at a constant rate during the initial phases of the culture. However, invariably as cultures accumulated passages, cells changed to a flattened and enlarged morphology and ceased to proliferate. This growth arrest was stable since addition of new growth factors did not result in resumed cell division.

These observations led the authors to relate cell senescence with the process of aging. Proliferative exhaustion was seen as the cellular reflection of organismal decay. In contrast, tumor cells were immortal and did not show any signs of cell senescence. One key element derived from this theoretical framework is that cells need to have a counting mechanism responsible for triggering the senescence program. This prompted many laboratories to initiate the search for factors that could be involved in regulating the entry into senescence as a way to decipher the molecular mechanisms underlying aging.

Decades later, while trying to model cancer initiation after oncogene activation, cell senescence was identified as a response opposing oncogenic transformation (Serrano *et al.*, 1997). Cells carrying an activated oncogene were shown to either die through apoptosis or enter the characteristic stable cell cycle arrest that defines cell senescence. These observations were substantiated *in vivo* using animal models and human samples undergoing oncogenic stress, and cell senescence proved to be a powerful barrier blocking tumor development (Collado and Serrano, 2010).

Oncogenic activation was the first example of cell senescence induction as a response against a stress that could potentially alter the cellular integrity. Other stresses soon followed and cell senescence started to be seen as a common protective response preventing proliferation of damaged cells (Shay and Roninson, 2004).

Recently, the identification of cell senescence during embryonic development as a remodeling mechanism opens up

the possibility that senescence originated as a physiological process that was co-opted in adulthood to respond in stressful conditions (Campisi, 2014). The implications of cell senescence for the normal physiology of tissues in healthy conditions, or during aging or disease are still a matter of intense investigation, but the relevance of this cellular response is undisputed.

What Defines Cellular Senescence?

No definitive marker of cell senescence exists and cumulative evidence is required to define a cell as senescent. From the initial identification of cell senescence, the main feature of this process is the stable cell cycle arrest. Even though cells remain metabolically active, they can no longer re-enter the cell cycle. This growth arrest is imposed by cell cycle regulators such as the cyclin-dependent kinase inhibitors of the INK (p15, p16, p18, and p19) and CIP/KIP (p21, p27, and p57) families. Identification of high levels of these proteins together with the concomitant absence of proliferation markers, are indicative of a possible cell senescence response. However, quiescence or transitory growth arrest can also result in increased levels of cell cycle inhibitors and lack of proliferative markers. *In vitro*, cells adopt a typical appearance characterized by enlarged and flatten morphology although the relevance of this change in morphology has never been proved *in vivo*.

The most widely used marker of cell senescence is the enzymatic determination of lysosomal beta-galactosidase (Dimiri *et al.*, 1995) known as SA-beta-gal (for senescence-associated beta-galactosidase) activity. For not totally understood reasons, the lysosomal content of senescent cells increase. In parallel, mRNA transcription and protein synthesis from the GLB1 gene also increase. Lysosomal beta-galactosidase activity is usually assessed through a chromogenic reaction employing a galactose analog, X-gal. The reaction is optimal at pH 4.0–4.5 but when performed at a suboptimal pH of 6.0, only cells with increased lysosomal content and high beta-galactosidase activity, that is, senescent cells, are detected (Kurz *et al.*, 2000). This situation provides a useful, although limited marker to detect senescent cells. The beta-galactosidase activity is not functionally related with the senescent process, since experimentally decreased expression or absence in patients with autosomal recessive G(M1)-gangliosidosis, carrying a defective lysosomal beta-galactosidase enzyme, do not show any defect in senescence induction (Lee *et al.*, 2006). Besides, some reports showed positive artefactual reactions not linked to senescence.

Numerous gene expression changes have been proposed to underlie cell senescence, although the generality and relevance of each individual change is difficult to define. Particular molecular signaling pathways, epigenetic changes, and secreted factors have been identified and might play important

roles during senescence, but these features could be only taking place in particular situations and will be discussed later.

Molecular Signaling Pathways

The description of cell senescence after proliferative exhaustion in culture implied the existence of a molecular mechanism capable of counting cell divisions, or more generally speaking, a mechanism that could sense the proliferative history of the cell. At the end of the last century, with the detailed study of the structure and function of telomeres, nucleoprotein structures protecting the tip of the chromosomes, they became an obvious candidate to be playing this role in the cell. Cell division leads to the shortening of the repetitive DNA at telomeres due to the end-replication problem, the impossibility of the DNA replication machinery to completely copy both strands of DNA. The accumulation of cell divisions thus results in attrition of telomeres up to a point in which they lose their protective capacity and the extremes of chromosomes end up being recognized as sites of DNA damage by the repair machinery (Harley *et al.*, 1990). This damage triggers a tumor suppressive response with cell senescence as one of the possible outcomes, being apoptosis another potential consequence of DNA damage at telomeres. The nature of the cellular decision remains elusive.

Another consequence of consecutive cell divisions is the accumulation of unresolved genomic damage at sites different from telomeric regions. The inevitable consequence of errors derived from an imperfect DNA replication process, or not completely repaired DNA damage, is a genome that piles up increasing levels of lesions. This situation cannot be tolerated by the cell, due to the risk that injured genomic information would pose for the integrity of the message that has to be passed from cell to cell. Unrepaired DNA damage results in the activation of p53 tumor suppressor and, again, in the induction of senescence or apoptosis (d'Adda di Fagagna, 2008).

Finally, another important player in determining the entry into the senescent growth arrest after serial passage is the derepression of the *Ink4a/ARF* locus (Gil and Peters, 2006). This locus is a remarkable genomic region that encodes for two structurally unrelated tumor suppressors, p16 and ARF (p14 in humans and p19 in mice). This locus stays silent in primary cells and becomes activated only after cumulative cell divisions (Figure 1).

Interestingly, other triggers of cell senescence result also in the activation of these molecular pathways (Collado *et al.*, 2007). Aged tissues show telomere shortening and DNA damage signaling that emanates from dysfunctional telomeres. Oncogene induction activates p53 and the *Ink4a/ARF* locus, and can cause replicative stress that signals through the DNA damage response pathway.

Epigenetic Control of Senescence

The irreversibility of the growth arrest imposed by cell senescence has an epigenetic basis. De novo heterochromatin formation of genomic regions carrying genes with crucial roles promoting cell proliferation is a hallmark of senescence. These regions are known as senescence-associated heterochromatin

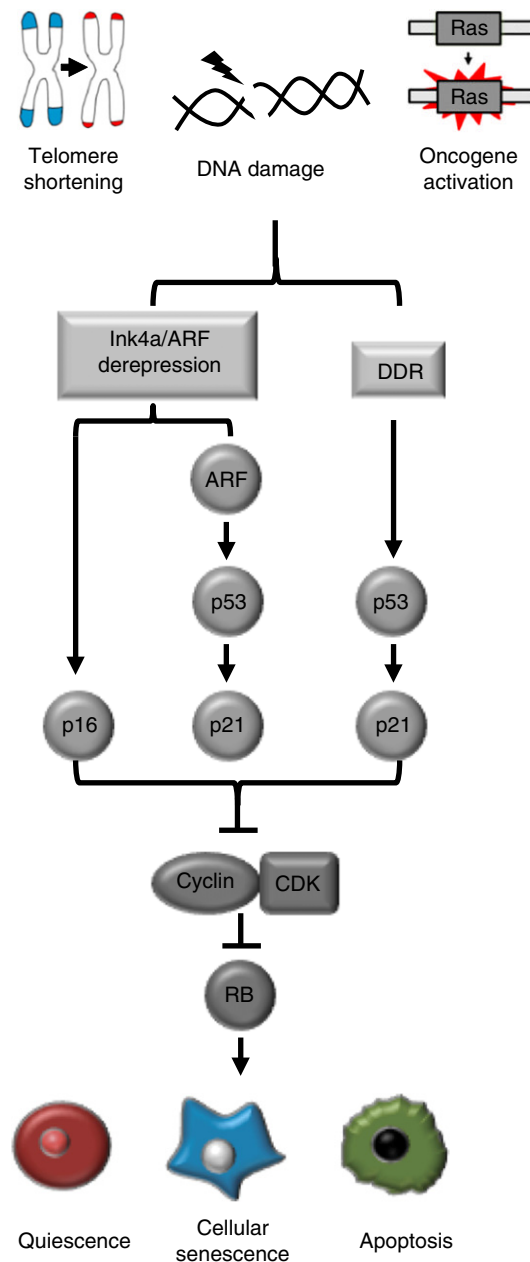


Figure 1 Molecular pathways leading to cell senescence. Different stress signals might engage the cell senescence machinery. The main inducers of cell senescence are telomere shortening, unresolved DNA damage accumulation, and oncogene activation. These different stressors converge on common cellular pathways governed by the *Ink4a/ARF* and p53 tumor suppressor genes. The activation of these molecules imposes a stable cell cycle arrest, although other alternative end points, such as reversible growth arrest (quiescence) or cell death (apoptosis), are also possible.

foci (SAHFs) (Narita *et al.*, 2003) and can be visualized through different methodological approaches. Compacted chromatin, visualized as intense 4',6-diamidino-2-phenylindole (DAPI)-stained DNA or the presence of Lysine 9 trimethylated histone H3 and HP1 γ , are features of the heterochromatin nature of these genomic areas. SAHFs are also

characterized by their accumulation of HMGA proteins (non-histone chromatin architectural proteins) and depletion of linker Histone H1 (Narita *et al.*, 2006; Funayama *et al.*, 2006). Efforts to decipher the sequence of events that culminate with the establishment of this epigenetic regulation and the molecular players involved has provided us with a better understanding of this strict control (Chan *et al.*, 2005; Zhang *et al.*, 2005; Funayama *et al.*, 2006; Narita *et al.*, 2006; Ye *et al.*, 2007). However, there is no proof for a direct causal relationship between SAHF formation and the particular regulation of specific genes involved in senescence.

Secreted Factors

Since senescent cells are characterized by their stable exit from the cell cycle, it was assumed that this response was an end point that enabled efficient management of stress responses. However, senescent cells remain metabolically active once they stop dividing and in recent years it has become clear that they specifically produce a high number of secreted factors. Senescent cells release a plethora of factors mainly extracellular proteases, cytokines and chemokines, and perhaps counter intuitively growth factors (Coppé *et al.*, 2010). These secreted factors are known collectively as senescence-associated secretory phenotype (SASP). The SASP has a wide range of potential activities, with both autocrine and paracrine effects proposed, and with both beneficial and detrimental activities. In the initial description of the SASP, secreted factors were identified as reinforcing the senescence cell cycle arrest in an autocrine manner (Kuilman *et al.*, 2008; Acosta *et al.*, 2008). From this point of view, the SASP would act in an autocrine manner and be beneficial. Some reports also suggested that senescent cells use the SASP as a chemical message to alert neighboring cells and extend the senescence phenotype (Acosta *et al.*, 2013) acting in a paracrine manner to prepare the tissue to confront a potential damage. Even though *in vitro* senescent cells have always been considered to remain alive for prolonged periods of time, the situation *in vivo* might be different and the SASP has also been pointed as responsible for attracting immune cells to the sites of damage to clear up damaged growth-arrested cells (Hoenicke and Zender, 2012). Strikingly, some of the secreted factors have obvious links with pro-inflammatory signals that have been traditionally considered to exert pro-tumoral actions. These led many investigators to propose that a potential detrimental action of senescent cells accumulating at sites of damage or during aging could be the result of a constant low level inflammatory reaction. Some authors have proposed that in the context of oncogene activation cell senescence exerts an antitumor action with a SASP that can switch from cancer-inhibitory to cancer-promoting depending on the integrity of tumor suppressor p53 (Pribluda *et al.*, 2013). This pro-inflammatory signals derived from senescent cells have also been proposed to be responsible for promoting proliferation of premalignant or fully tumorigenic neighboring cells (Davalos *et al.*, 2010; Coppé *et al.*, 2010). Some authors even proposed that the SASP could be responsible for promoting a more malignant and metastatic phenotype in cancer cells (Yu *et al.*, 2013; Angelini *et al.*, 2013), something that would imply that potential therapies aimed at

inducing cancer cell senescence could be even dangerous (Figure 2).

Senescence in Aging

Originally, the finite proliferative capacity of cells was seen as a reflection of the process of aging. Understanding the molecular mechanisms that control this process was considered a way to reveal the limit of the life span of an organism and the key to reverse aging. In human cells, telomere shortening was identified as responsible for triggering senescence after prolonged culture (Harley *et al.*, 1990) and the regulation of telomeres has become a crucial field in aging biology. Different studies showed accumulation of SA-beta-gal-positive cells in aged tissues as well as an age-dependent increase in the expression of senescence markers (Herbig *et al.*, 2006; Wang *et al.*, 2009). Signs of DNA damage at telomeric regions increase with age and this has been seen as an *in vivo* proof of the *in vitro* observations. All these 'signs' of senescence are more prominent in tissues from animal models and human patients suffering premature aging syndromes (Gonzalo, 2014). However, a causative relationship between senescence and aging is not completely clear. A correlation between donor age and *in vitro* proliferative potential is controversial (Cristofalo *et al.*, 2004). At the same time, cells derived from premature aging syndrome patients become senescent more readily in culture compared with young or healthy individuals (McClintock *et al.*, 2006).

Some authors proposed that the age-dependent accumulation of senescent cells occurs in stem cells and progenitor compartments in various tissues, limiting the regenerative capacity of those organs during aging (Sharpless and DePinho, 2007). However, there is no formal proof that these cells were actually undergoing senescence since this description relied only on the identification of increased p16 in these cells (Molofsky *et al.*, 2006; Krishnamurthy *et al.*, 2006; Janzen *et al.*, 2006). Using mouse models of premature aging, other authors proposed a role for an exacerbated cell senescence response in the premature aging phenotype (Baker *et al.*, 2006). Deletion of p16 was shown to improve some of the aging phenotypes but again, it is not clear that this was causing a decrease in senescence (Baker *et al.*, 2008). Recently, selective killing of p16-positive cells in this same premature aging mouse model ameliorated the aging phenotype, but had no effect on global life span (Baker *et al.*, 2011). To add more confusion, mice carrying extra gene dosages of the whole Ink4a/ARF locus (coding for p15, p16, and p19ARF) actually showed prolonged life span (Matheu *et al.*, 2009), similarly to what was also shown for animals with reinforced p53 and Ink4a/ARF (Matheu *et al.*, 2007).

Senescence in Cancer

Senescence of human diploid fibroblasts after long-term cell culture *in vitro* was received with incredulity due to the experience with human tumor derived cell cultures that can be maintained essentially indefinitely. This different behavior implies that tumor cells evade this response as part of the essential changes that make them tumorigenic (Hanahan and Weinberg, 2000). Investigation of the molecular players

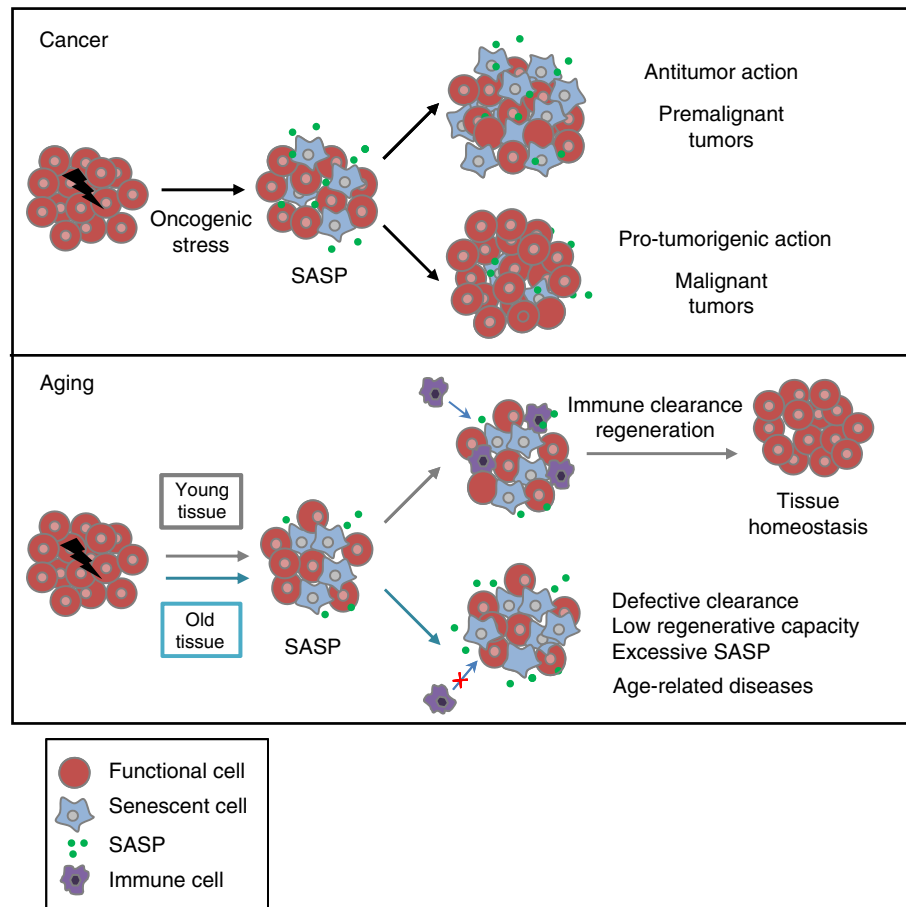


Figure 2 Role of senescence-associated secretory phenotype (SASP) in cancer and aging. The SASP can have both beneficial and detrimental effects on senescent cells and surrounding tissue. In cancer, after cell senescence triggered by oncogenic stress, cells secrete various SASP components that might result in enhanced cell senescence or escape from the growth arrest depending on other not well-defined circumstances. During aging, there is a shift from a beneficial contribution of the SASP to tissue homeostasis after cell senescence induced by sporadic damage, to a detrimental action exerted in old tissue due to accumulated damage that is not effectively resolved by immune clearance. Excessive SASP in aged tissue might contribute to age-related diseases.

involved in senescence showed that the cell cycle arrest was mediated by the main tumor suppressor networks of the cell, ARF/p53 and INK4a/RB (Hara *et al.*, 1991; Shay *et al.*, 1991; Bond *et al.*, 1994; Lin *et al.*, 1998; Schmitt *et al.*, 2002; Beauséjour *et al.*, 2003; Collins and Sedivy, 2003; Oren, 2003; Herbig *et al.*, 2004; Jacobs and de Lange, 2004; Ohtani *et al.*, 2004; Chen *et al.*, 2005; Campisi and d'Adda di Fagagna, 2007; Rodier *et al.*, 2007). Loss of these tumor suppressors is an essential requisite to render cells insensitive to the senescent arrest, that is, to become immortal (Hahn *et al.*, 2002; Hahn and Weinberg, 2002), and an extremely frequent event in human cancer (Hollstein *et al.*, 1991; Ruas and Peters, 1998; Sharpless and DePinho, 1999). Thus, cell senescence could be considered as a tumor suppressor mechanism. This idea was experimentally exposed when an activated oncogene was introduced into normal human diploid fibroblasts (Serrano *et al.*, 1997). Expression of the oncogene triggered an immediate phenotypic change that was identical to the one observed after serial passage, and the permanent cell cycle arrest, mediated by p16 and p53, was considered an efficient tumor suppressive barrier. *In vivo* evidence for the existence and

relevance of this oncogene-induced senescence was provided later on using animal models of oncogenic activation and human samples derived from cancer patients. Premalignant tumors arising from experimental oncogenic activation in genetically modified animals (Braig *et al.*, 2005; Chen *et al.*, 2005; Collado *et al.*, 2005) or samples obtained from cancer patients (Michaloglou *et al.*, 2005) were shown to be restricted by senescence induction. Avoiding the senescence response by deletion of the main tumor suppressor genes is fundamental for tumor progression and this correlates with the absence of the senescence response *in vivo* (Chen *et al.*, 2005). The growing list of studies showing similar results (Collado and Serrano, 2010) clearly argue for cell senescence as a basic mechanism restraining cancer through a cell-autonomous block of the proliferation of oncogenically primed cells.

Despite the universal requirement for a tumor cell to bypass the senescence growth arrest, some chemotherapy treatments have the ability to cause senescence of cultured tumor cells (Shay and Roninson, 2004), and *in vivo* grown tumors (Schmitt *et al.*, 2002; Coppé *et al.*, 2008). To which extend this is also occurring in human tumors after standard chemotherapy is

currently unknown, with only limited evidence described for some tumors (Poele *et al.*, 2002; Roberson *et al.*, 2005; Haugstetter *et al.*, 2010), but it would be very important to clearly establish the potential contribution of therapy-induced senescence. Interestingly, experimental reactivation of tumor suppressor p53 in tumors generated in its absence has the potential to control tumor growth either by inducing apoptosis or senescence depending on the tissue (Ventura *et al.*, 2007; Xue *et al.*, 2007). Furthermore, the induction of senescence is rapidly triggered even after short transient re-expression of p53 and is associated with tumor regression. This tumor regression after senescence induction is achieved through the action of the immune system. Again, the SASP seems an obvious candidate to be mediating such a process.

Given the theoretical irreversibility of cell senescence and the requisite to escape from this control for a tumor to grow, it is not clear how tumor cells manage to avoid the action of senescence. Either tumor cells acquire senescence deactivating mutations early on in their path toward full neoplastic transformation, or they suffer mutations, epigenetic alterations, or activate signaling pathways that can reverse the senescent growth arrest. One possible source of signals that could promote escape from senescence comes from the immune system. Pro-tumorigenic inflammation has been shown to cooperate with oncogene activation to dismantle the defense response of the cell in animal models of oncogenic induction in the pancreas (Guerra *et al.*, 2011) or colorectal epithelium (Pribluda *et al.*, 2013). Interestingly, recent data show that Gr1 + myeloid cells (a particular immune cell type) are attracted to the tumors by signals released by transformed cells. These myeloid cells are recruited to the tumors to take advantage of their senescence antagonizing activity (Di Mitri *et al.*, 2014).

Senescence and Tissue Repair

As already discussed, senescent cells are characterized by the release of multiple factors to the extracellular microenvironment, the SASP. Among them, several molecules have a clear involvement in tissue repair. Several growth factors and proteases that participate in wound healing, chemoattractants for immune cells, and proteins that mobilize stem or progenitor cells are in the lists of SASP components. Thus, senescent cells communicate with the surrounding tissue through the SASP to alert neighboring cells and the immune system about cellular damage and to stimulate repair. This function has been exemplified in recent work that revealed the role of senescence during tissue damage repair in liver. Acute liver damage activates stellate cells that, as a result, proliferate and secrete extracellular matrix (ECM) components that produce a fibrotic scar. After this initial reaction, stellate cells undergo senescence, stop producing ECM and in turn secrete SASP components that degrade the ECM, limiting excessive liver damage. When the senescence response is compromised by the lack of crucial senescence mediators such as p53 or *Ink4a/ARF*, the fibrotic reaction is disproportionate resulting in detrimental severe fibrosis that limits tissue repair (Krizhanovsky *et al.*, 2008). In another study, CCN1 was identified as a senescence mediator increased in human cirrhotic liver and in acutely damaged liver in mice. When CCN1 was deleted in genetically modified mice, the senescence process was impaired

and the fibrotic response was concomitantly exacerbated after liver damage (Kim *et al.*, 2013). Finally, another factor, IL22, was also shown to be hepatoprotective due to its ability to contribute to senescence induction of stellate cells, ameliorating excessive liver fibrogenesis (Kong *et al.*, 2012).

Similarly, skin damage leads to the induction of cell senescence of myofibroblasts during wound healing, restricting the fibrotic response, a process in which CCN1 has also been shown to play a similar role than the one played in liver. Interestingly, topical application of CCN1 protein to wounds induces senescence of fibroblasts and this response functions to curb fibrosis during skin repair (Jun and Lau, 2010).

In summary, mounting evidence point to a positive role for the senescence response as a process involved in normal tissue repair. In this context, the SASP seems to operate as a beneficial mechanism to favor tissue remodeling and regeneration.

Senescence in Development

Senescence gained a reputation as a stress response after its identification in cell culture and through its involvement as an anticancer barrier. Many saw the senescent growth arrest as a mechanism that evolved to secure the integrity of the genome and its putative role in aging as an undesired unselected detrimental side effect. With mounting evidence of senescence being involved also in beneficial responses related with tissue repair and remodeling, the question of a possible role for senescence in physiological embryo development was recently addressed by two groups. The identification of SA-beta-gal-positive cells in the transitory kidney, the endolymphatic sac of the inner ear, the closing neural tube and the apical ectodermal ridge, among other structures in the embryo, proved that cell senescence operates during embryogenesis (Muñoz-Espín *et al.*, 2013; Storer *et al.*, 2013). All these structures seem to trigger senescence through the action of the cell cycle inhibitor p21, and not through the common senescence regulator p16, and surprisingly, in a p53 independent manner. Senescent cells in the embryo seem to be cleared by the action of the immune system, very much like what has been described for adult tissues undergoing a stress response. Also, at least for some of these structures, released factors seem to play a crucial role promoting cell growth and ECM remodeling to allow tissue development, resembling the role of cell senescence in tissue repair. All these similarities raise the question of a possible evolutionary origin of cell senescence operating during embryo development and being co-opted during adulthood for its positive activities.

Conclusion

Once considered the basis of the process of aging, the causative relationship of cell senescence in aging still remains a matter of debate nowadays. The description of oncogene-induced senescence offered a plausible mechanism of tumor suppression based on the implementation of a permanent cell cycle arrest, and marked the way for a new view of senescence as a stress response. This positive effect protecting us from cancer was mirrored by the putative detrimental action that the accumulation of senescent cells could have with aging. This was seen by

many as an example of the antagonistic pleiotropy hypothesis of evolution. According to this idea, cell senescence could have been selected during evolution to serve a positive function stopping damaged cells from dividing and forming a tumor. Later in life, senescence would show its negative face, once the reproductive age was passed, and accumulation of senescent cells would interfere with the proper activity of the tissue. However, giving its function avoiding the emergence of damaged cells, and as promoter of repair and remodeling to help maintain tissue homeostasis, the accumulation of senescent cells with age could have new interpretations. The high numbers of senescent cells during aging could turn out to be considered the reflection of a futile effort to restore normal tissue function. The recent identification of cell senescence during embryonic development opens

yet a new avenue of research and defects in this process could have unexplored pathological consequences in adulthood. To which extent cell senescence has similar positive functions in normal unstressed conditions is something that remains to be addressed. Further studies into this exciting field would provide us with a better understanding of basic cell biology mechanisms and opportunities for more effective therapies (Figure 3).

See also: Nucleic Acid Synthesis/Breakdown: DNA Synthesis/Repair: Telomere Biology; Telomeres and Telomerase

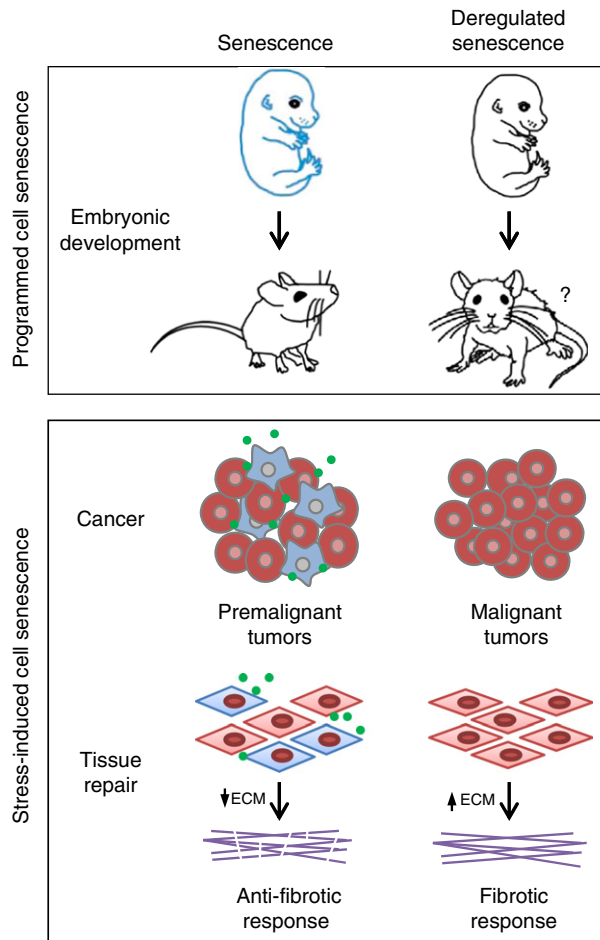


Figure 3 Programmed and stress-induced senescence. Recent work has identified programmed cell senescence as part of the developmental program, helping shape different tissues, and structures during embryogenesis. This discovery opens new possibilities regarding the pathological consequences that might derive from a defective or excessive cell senescence response during development. In the adult, stress-induced cell senescence has been shown to provide a barrier against tumor development and to contribute to tissue repair. The absence of an effective cell senescence process can result in tumor progression or an excessive fibrotic response. Similarly, an exacerbated or prolonged cell senescence response might promote tumor cell growth or result in dysfunctional aged tissues.

References

- Acosta, J.C., Banito, A., Wuestefeld, T., *et al.*, 2013. A complex secretory program orchestrated by the inflammasome controls paracrine senescence. *Nature Cell Biology* 15 (8), 978–990.
- Acosta, J.C., O’Loghlen, A., Banito, A., *et al.*, 2008. Chemokine signaling via the CXCR2 receptor reinforces senescence. *Cell* 133 (6), 1006–1018.
- d’Adda di Fagagna, F., 2008. Living on a break: Cellular senescence as a DNA-damage response. *Nature Reviews. Cancer* 8 (7), 512–522.
- Angelini, P.D., Zacarias Fluck, M.F., Pedersen, K., *et al.*, 2013. Constitutive HER2 signaling promotes breast cancer metastasis through cellular senescence. *Cancer Research* 73 (1), 450–458.
- Baker, D.J., Jeganathan, K.B., Malureanu, L., *et al.*, 2006. Early aging-associated phenotypes in Bub3/Rae1 haploinsufficient mice. *Journal of Cell Biology* 172 (4), 529–540.
- Baker, D.J., Perez-Terzic, C., Jin, F., *et al.*, 2008. Opposing roles for p16Ink4a and p19Arf in senescence and ageing caused by BubR1 insufficiency. *Nature Cell Biology* 10 (7), 825–836.
- Baker, D.J., Wijshake, T., Tchkonja, T., *et al.*, 2011. Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders. *Nature* 479 (7372), 232–236.
- Beauséjour, C.M., Krtolica, A., Galimi, F., *et al.*, 2003. Reversal of human cellular senescence: Roles of the p53 and p16 pathways. *EMBO Journal* 22 (16), 4212–4222.
- Bond, J.A., Wyllie, F.S., Wynford-Thomas, D., 1994. Escape from senescence in human diploid fibroblasts induced directly by mutant p53. *Oncogene* 9 (7), 1885–1889.
- Braig, M., Lee, S., Loddenkemper, C., *et al.*, 2005. Oncogene-induced senescence as an initial barrier in lymphoma development. *Nature* 436 (7051), 660–665.
- Campisi, J., 2014. Cell biology: The beginning of the end. *Nature* 505 (7481), 35–36.
- Campisi, J., d’Adda di Fagagna, F., 2007. Cellular senescence: When bad things happen to good cells. *Nature Reviews. Molecular Cell Biology* 8 (9), 729–740.
- Chan, H.M., Narita, M., Lowe, S.W., Livingston, D.M., 2005. The p400 E1A-associated protein is a novel component of the p53 → p21 senescence pathway. *Genes & Development* 19 (2), 196–201.
- Chen, Z., Trotman, L.C., Shaffer, D., *et al.*, 2005. Crucial role of p53-dependent cellular senescence in suppression of Pten-deficient tumorigenesis. *Nature* 436 (7051), 725–730.
- Collado, M., Blasco, M.A., Serrano, M., 2007. Cellular senescence in cancer and aging. *Cell* 130 (2), 223–233.
- Collado, M., Gil, J., Efeyan, A., *et al.*, 2005. Tumour biology: Senescence in premalignant tumours. *Nature* 436 (7051), 642.
- Collado, M., Serrano, M., 2010. Senescence in tumours: Evidence from mice and humans. *Nature Reviews. Cancer* 10 (1), 51–57.
- Collins, C.J., Sedivy, J.M., 2003. Involvement of the INK4a/Arf gene locus in senescence. *Aging Cell* 2 (3), 145–150.
- Coppé, J.-P., Desprez, P.-Y., Krtolica, A., Campisi, J., 2010. The senescence-associated secretory phenotype: The dark side of tumor suppression. *Annual Review of Pathology* 5, 99–118.
- Coppé, J.-P., Patil, C.K., Rodier, F., *et al.*, 2008. Senescence-associated secretory phenotypes reveal cell-nonautonomous functions of oncogenic RAS and the p53 tumor suppressor. *PLoS Biology* 6 (12), 2853–2868.
- Cristofalo, V.J., Lorenzini, A., Allen, R.G., Torres, C., Tresini, M., 2004. Replicative senescence: A critical review. *Mechanisms of Ageing and Development* 125 (10–11), 827–848.

- Davalos, A.R., Coppe, J.-P., Campisi, J., Desprez, P.-Y., 2010. Senescent cells as a source of inflammatory factors for tumor progression. *Cancer Metastasis Reviews* 29 (2), 273–283.
- Dimri, G., Lee, X., Basile, G., 1995. A biomarker that identifies senescent human cells in culture and in aging skin *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America* 92, 9363–9367.
- Di Mitri, D., Toso, A., Chen, J.J., *et al.*, 2014. Tumour-infiltrating Gr-1+ myeloid cells antagonize senescence in cancer. *Nature* 515 (7525), 134–137.
- Funayama, R., Saito, M., Tanobe, H., Ishikawa, F., 2006. Loss of linker histone H1 in cellular senescence. *Journal of Cell Biology* 175 (6), 869–880.
- Gil, J., Peters, G., 2006. Regulation of the INK4b-ARF-INK4a tumour suppressor locus: All for one or one for all. *Nature Reviews. Molecular Cell Biology* 7 (9), 667–677.
- Gonzalo, S., 2014. DNA damage and lamins. *Advances in Experimental Medicine and Biology* 773, 377–399.
- Guerra, C., Collado, M., Navas, C., *et al.*, 2011. Pancreatitis-induced inflammation contributes to pancreatic cancer by inhibiting oncogene-induced senescence. *Cancer Cell* 19 (6), 728–739.
- Hahn, W.C., Dessain, S.K., Brooks, M.W., *et al.*, 2002. Enumeration of the simian virus 40 early region elements necessary for human cell transformation. *Molecular and Cellular Biology* 22 (7), 2111–2123.
- Hahn, W.C., Weinberg, R.A., 2002. Rules for making human tumor cells. *New England Journal of Medicine* 347 (20), 1593–1603.
- Hanahan, D., Weinberg, R.A., 2000. The hallmarks of cancer. *Cell* 100 (1), 57–70.
- Hara, E., Tsurui, H., Shinozaki, A., Nakada, S., Oda, K., 1991. Cooperative effect of antisense-Rb and antisense-p53 oligomers on the extension of life span in human diploid fibroblasts, TIG-1. *Biochemical and Biophysical Research Communications* 179 (1), 528–534.
- Harley, C.B., Futcher, A.B., Greider, C.W., 1990. Telomeres shorten during ageing of human fibroblasts. *Nature* 345 (6274), 458–460.
- Haugstetter, A.M., Loddenkemper, C., Lenze, D., *et al.*, 2010. Cellular senescence predicts treatment outcome in metastasised colorectal cancer. *British Journal of Cancer* 103 (4), 505–509.
- Hayflick, L., Moorhead, P.S., 1961. The serial cultivation of human diploid cell strains. *Experimental Cell Research* 25, 585–621.
- Herbig, U., Ferreira, M., Condel, L., Carey, D., Sedivy, J.M., 2006. Cellular senescence in aging primates. *Science* (New York, NY) 311 (5765), 1257.
- Herbig, U., Jobling, W.A., Chen, B.P.C., Chen, D.J., Sedivy, J.M., 2004. Telomere shortening triggers senescence of human cells through a pathway involving ATM, p53, and p21(CIP1), but not p16(INK4a). *Molecular Cell* 14 (4), 501–513.
- Hoenicke, L., Zender, L., 2012. Immune surveillance of senescent cells — biological significance in cancer- and non-cancer pathologies. *Carcinogenesis* 33 (6), 1123–1126.
- Hollstein, M., Sidransky, D., Vogelstein, B., Harris, C.C., 1991. p53 mutations in human cancers. *Science* (New York, NY) 253 (5015), 49–53.
- Jacobs, J.J.L., de Lange, T., 2004. Significant role for p16INK4a in p53-independent telomere-directed senescence. *Current Biology: CB* 14 (24), 2302–2308.
- Janzen, V., Forkert, R., Fleming, H.E., *et al.*, 2006. Stem-cell ageing modified by the cyclin-dependent kinase inhibitor p16INK4a. *Nature* 443 (7110), 421–426.
- Jun, J.-I., Lau, L.F., 2010. The matricellular protein CCN1 induces fibroblast senescence and restricts fibrosis in cutaneous wound healing. *Nature Cell Biology* 12 (7), 676–685.
- Kim, K.-H., Chen, C.-C., Monzon, R.I., Lau, L.F., 2013. Matricellular protein CCN1 promotes regression of liver fibrosis through induction of cellular senescence in hepatic myofibroblasts. *Molecular and Cellular Biology* 33 (10), 2078–2090.
- Kong, X., Feng, D., Wang, H., *et al.*, 2012. Interleukin-22 induces hepatic stellate cell senescence and restricts liver fibrosis in mice. *Hepatology* (Baltimore, MD) 56 (3), 1150–1159.
- Krishnamurthy, J., Ramsey, M.R., Ligon, K.L., *et al.*, 2006. p16INK4a induces an age-dependent decline in islet regenerative potential. *Nature* 443 (7110), 453–457.
- Krizhanovsky, V., Yon, M., Dickins, R.A., *et al.*, 2008. Senescence of activated stellate cells limits liver fibrosis. *Cell* 134 (4), 657–667.
- Kuilman, T., Michaloglou, C., Vredeveld, L.C.W., *et al.*, 2008. Oncogene-induced senescence relayed by an interleukin-dependent inflammatory network. *Cell* 133 (6), 1019–1031.
- Kurz, D.J., Decary, S., Hong, Y., Erusalimsky, J.D., 2000. Senescence-associated (beta)-galactosidase reflects an increase in lysosomal mass during replicative ageing of human endothelial cells. *Journal of Cell Science* 113 (Pt 2), 3613–3622.
- Lee, B.Y., Han, J.A., Im, J.S., *et al.*, 2006. Senescence-associated beta-galactosidase is lysosomal beta-galactosidase. *Aging Cell* 5 (2), 187–195.
- Lin, A.W., Barradas, M., Stone, J.C., *et al.*, 1998. Premature senescence involving p53 and p16 is activated in response to constitutive MEK/MAPK mitogenic signaling. *Genes & Development* 12 (19), 3008–3019.
- Matheu, A., Maraver, A., Collado, M., *et al.*, 2009. Anti-aging activity of the Ink4/Arf locus. *Aging Cell* 8 (2), 152–161.
- Matheu, A., Maraver, A., Klatt, P., *et al.*, 2007. Delayed ageing through damage protection by the Arf/p53 pathway. *Nature* 448 (7151), 375–379.
- McClintock, D., Gordon, L.B., Djabali, K., 2006. Hutchinson-Gilford progeria mutant lamin A primarily targets human vascular cells as detected by an anti-Lamin A G608G antibody. *Proceedings of the National Academy of Sciences of the United States of America* 103 (7), 2154–2159.
- Michaloglou, C., Vredeveld, L.C.W., Soengas, M.S., *et al.*, 2005. BRAF600-associated senescence-like cell cycle arrest of human naevi. *Nature* 436 (7051), 720–724.
- Molofsky, A.V., Slutsky, S.G., Joseph, N.M., *et al.*, 2006. Increasing p16INK4a expression decreases forebrain progenitors and neurogenesis during ageing. *Nature* 443 (7110), 448–452.
- Muñoz-Espín, D., Cañamero, M., Maraver, A., *et al.*, 2013. Programmed cell senescence during mammalian embryonic development. *Cell* 155 (5), 1104–1118.
- Narita, M., Narita, M., Krizhanovsky, V., *et al.*, 2006. A novel role for high-mobility group A proteins in cellular senescence and heterochromatin formation. *Cell* 126 (3), 503–514.
- Narita, M., Nunez, S., Heard, E., *et al.*, 2003. Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* 113 (6), 703–716.
- Ohtani, N., Yamakoshi, K., Takahashi, A., Hara, E., 2004. The p16INK4a-RB pathway: Molecular link between cellular senescence and tumor suppression. *Journal of Medical Investigation: JMI* 51 (3–4), 146–153.
- Oren, M., 2003. Decision making by p53: Life, death and cancer. *Cell Death and Differentiation* 10 (4), 431–442.
- Poele, R.H., Okorokov, A.L., Jardine, L., Cummings, J., Joel, S.P., 2002. DNA damage is able to induce senescence in tumor cells *in vitro* and *in vivo*. *Cancer Research* 62 (6), 1876–1883.
- Pribluda, A., Elyada, E., Wiener, Z., *et al.*, 2013. A senescence-inflammatory switch from cancer-inhibitory to cancer-promoting mechanism. *Cancer Cell* 24 (2), 242–256.
- Roberson, R.S., Kussick, S.J., Vallieres, E., Chen, S.-Y.J., Wu, D.Y., 2005. Escape from therapy-induced accelerated cellular senescence in p53-null lung cancer cells and in human lung cancers. *Cancer Research* 65 (7), 2795–2803.
- Rodier, F., Campisi, J., Bhaumik, D., 2007. Two faces of p53: Aging and tumor suppression. *Nucleic Acids Research* 35 (22), 7475–7484.
- Ruas, M., Peters, G., 1998. The p16INK4a/CDKN2A tumor suppressor and its relatives. *Biochimica et Biophysica Acta* 1378 (2), F115–F177.
- Schmitt, C.A., Fridman, J.S., Yang, M., *et al.*, 2002. A senescence program controlled by p53 and p16INK4a contributes to the outcome of cancer therapy. *Cell* 109 (3), 335–346.
- Serrano, M., Lin, A., McCurrach, M., Beach, D., Lowe, S., 1997. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16 INK4a. *Cell* 88, 593–602.
- Sharpless, N.E., DePinho, R.A., 1999. The INK4A/ARF locus and its two gene products. *Current Opinion in Genetics & Development* 9 (1), 22–30.
- Sharpless, N.E., DePinho, R.A., 2007. How stem cells age and why this makes us grow old. *Nature Reviews. Molecular Cell Biology* 8 (9), 703–713.
- Shay, J.W., Pereira-Smith, O.M., Wright, W.E., 1991. A role for both RB and p53 in the regulation of human cellular senescence. *Experimental Cell Research* 196 (1), 33–39.
- Shay, J.W., Roninson, I.B., 2004. Hallmarks of senescence in carcinogenesis and cancer therapy. *Oncogene* 23 (16), 2919–2933.
- Storer, M., Mas, A., Robert-Moreno, A., *et al.*, 2013. Senescence is a developmental mechanism that contributes to embryonic growth and patterning. *Cell* 155 (5), 1119–1130.
- Ventura, A., Kirsch, D.G., McLaughlin, M.E., *et al.*, 2007. Restoration of p53 function leads to tumour regression *in vivo*. *Nature* 445 (7128), 661–665.
- Wang, C., Jurk, D., Maddick, M., *et al.*, 2009. DNA damage response and cellular senescence in tissues of aging mice. *Aging Cell* 8 (3), 311–323.
- Xue, W., Zender, L., Miething, C., *et al.*, 2007. Senescence and tumour clearance is triggered by p53 restoration in murine liver carcinomas. *Nature* 445 (7128), 656–660.
- Ye, X., Zerlanko, B., Kennedy, A., *et al.*, 2007. Downregulation of Wnt signaling is a trigger for formation of facultative heterochromatin and onset of cell senescence in primary human cells. *Molecular Cell* 27 (2), 183–196.
- Yu, Y.-C., Yang, P.-M., Chuah, Q.-Y., *et al.*, 2013. Radiation-induced senescence in securin-deficient cancer cells promotes cell invasion involving the IL-6/STAT3 and PDGF-BB/PDGFR pathways. *Scientific Reports* 3, 1675.
- Zhang, R., Poustovoitov, M.V., Ye, X., *et al.*, 2005. Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. *Developmental Cell* 8 (1), 19–30.

CELLULAR IMMUNOLOGY: OVERVIEW

Introduction to Functional Cell Biology of Immunity

K Cadwell, New York University School of Medicine, New York, NY, USA

ML Dustin, New York University School of Medicine, New York, NY, USA and The University of Oxford, Headington, UK

© 2016 Elsevier Inc. All rights reserved.

Glossary

Antibody The secreted splice variant of the antigen-binding effector molecule generated by B lymphocytes. It is the product of three somatic gene modifications: recombination of variable, diversity and joining gene segments, somatic hypermutation and class switch recombination. Antibodies are secreted in large amounts by plasma cells.

B cell receptor The cell surface splice variant of the antigen-binding receptor on B cells that is generated through three somatic gene modifications: recombination of variable, diversity and joining gene segments, somatic hypermutation and class switch recombination. Alternatively spliced forms are expressed on the cell surface where they serve primary roles of signaling and antigen capture or secreted forms that are involved in immune effector function. Also referred to as immunoglobulin or antibody.

Caspase A family of proteases that are involved in control of cell death pathways.

Cluster of differentiation The common nomenclature system that is used to disambiguate surface molecules in the human and mouse based on monoclonal antibody staining patterns across stages of leukocyte differentiation.

Immunoglobulin superfamily A large family of proteins that share a common globular domain of ~100 amino acids with two β sheets of 3–5 strands often joined by an intrachain disulfide bond. The level of homology can be low ($\geq 20\%$) such that criteria for inclusion are typically based on a number of criteria including structural prediction, intron-exon boundaries, and, most definitively, X-ray crystallographic structures.

Immunological synapse A cell–cell junction stabilized by adhesion molecules and directed by innate or adaptive immune recognition with effector functions carried out by polarized secretion or extracellular vesicle budding.

Inflammasome A cytoplasmic complex that mediates induced innate immune reactions associated with injury or infection.

Integrin A family large heterodimeric adhesion molecules that display activation regulated, divalent cation dependent binding of ligands and create an integral membrane bridge to the cytoskeleton.

Interleukin Small proteins that mediate communication between cells in the immune system. There are currently 36 interleukins with the designation IL-1 through IL-36.

Isotype class The variable domain of an antibody can be linked to different effector domains through ‘class switching,’ which is regulated by activation induced cytosine deaminase (AID) and transcription from promoters of gene segments encoding different ‘constant’ domains.

Peripheral node addressin A sulphated sialyl lewis X carbohydrate antigen associated with sialyl mucins such as CD34 expressed on high endothelial venules of lymph nodes. It's the ligand for CD62L on leukocytes.

T cell receptor The antigen-binding receptor on T cells.

Toll-like receptor A family of type I transmembrane proteins with leucine rich repeats in the extracellular domain and cytoplasmic domains that link to signal transduction to stimulated induced innate immunity. The ligands include conserved macromolecules from pathogens and some endogenous molecules released or generated in conditions of tissue injury.

Variable lymphocyte receptor Leucine rich repeat type receptors that are generated by somatic gene rearrangements in jawless fish that are responsible for adaptive immunity. An alternative structural scaffold to the immunoglobulin superfamily for generation of combinatorial diversity to recognize diverse pathogens.

Virulence factor Gene products that code for proteins that enhance the fitness of a pathogen to survive in the host and spread to new hosts.

Introduction

One of the fundamental concepts in cell biology is that cells build barriers to create compartmentalization of life processes and infection takes place when these barriers are compromised by another organism at the expense of the host. The

compartments can be within the cell itself or formed by co-operation of many cells in a multicellular organism. We will define immunity as mechanisms that reinforce these standing barriers to protect host compartments and the response of the host once these barriers are breached or functionally compromised. These same mechanisms may also manage

colonization with helpful microbes to achieve optimal fitness in the community. All self-replicating organisms engage in activities that could be considered a form of immunity to infection. Human infectious agents are subdivided into three major groups (1) viruses and intracellular bacteria, (2) extracellular bacteria and fungi, and (3) protozoan and multicellular parasites. These divisions are not taxonomic, but are based on studying how the host organizes its immune response to these categories of invaders. All self-replicating organisms need to cope with infection – including bacteria, plants, protists, and invertebrate animals. This section on the functional cell biology will focus attention on human immunity and model organisms that are studied to better understand principles of human immunity to infection. Model organisms are quite diverse and include flies, jawless and jawed fish, and other mammals. We will further focus on how functional immunity relates to cell biology principles learned from studying model cells and developmental processes.

We can divide the immune response into interconnected innate and adaptive branches. Innate immunity refers to evolved strategies to prevent barrier compromise and mount a rapid response to infection once a barrier breach occurs. In the first stage of an innate immune response, the body detects the presence of a threat, which is typically the presence of a molecule derived from an infectious agent. This ability to discriminate an entity as foreign or out of place, referred to as ‘self versus nonself,’ is essential for both innate and adaptive immunity. Innate immunity takes advantage of evolutionarily conserved characteristics of pathogens, which includes the ability to survive in particular physical environments and conserved molecular signatures of the pathogen that are not found in the host, such as components of the bacterial cell wall. The innate immune system can also be triggered by the presence of a host-derived molecule that is foreign to a particular compartment, such as genomic DNA that is present outside of the nucleus or cell. The mislocalization of host molecules is often a consequence of an injury that coincides with infection. These molecules that are foreign or perceived as danger signals can be recognized by germline-encoded receptor systems to initiate inflammation. Inflammation is a coordinated local and systemic response that rapidly acts to prevent the spread of the pathogen and create conditions that will destroy most invading organisms. All multicellular organisms have elaborate innate immune mechanisms that lead to inflammatory response; even bacteria have counter-measures to control propagation of phages (viruses that infect bacteria) such as the Crispr-Cas system, which also has practical utility in genetic engineering. Vertebrates, including jawed fish, reptiles, amphibians, birds, and mammals, have an additional layer of adaptive immunity which is defined by the ability to select new receptors to recognize nonconserved macromolecules of pathogens that more efficiently focus inflammatory mechanisms to eradicate pathogens that are not stopped by innate immunity alone. Failure of innate immunity is one of the alarms that activate adaptive immune mechanisms and thus these systems are tightly interlinked. Furthermore, after a first encounter with a pathogen, adaptive immunity imparts immunological memory, which allows a more efficient response in future encounters with the same pathogen.

Elements of Innate and Adaptive Immunity

Innate immunity has standing and induced components. Standing components of innate immunity are in place continuously and respond immediately to challenges, and include barriers such as the skin, transport mechanisms to avoid colonization like the mucociliary escalator in the airways, and constitutive chemical defenses such as the acidity of the stomach or the complement proteins in the blood. Other responses of innate immunity are carried out by recruited immune cells and induced translational/transcriptional programs that alter the capabilities of the tissue in combating an invader within minutes to hours. Some induced responses in long-lived innate cells can result in modification of the innate immune system such that it can mount a better response to a similar pathogen in the future, but this is different than adaptive immunity mediated by B and T lymphocytes. In the case of B and T lymphocytes, the respective surface receptors undergo recombination such that each cell has a unique receptor. The component that is recognized by the receptor is referred to as the antigen. The entire naïve repertoire of billions of lymphocytes may contain only a 1000 cells specific for antigens corresponding to a small pathogen like a virus. These rare antigen specific cells must undergo several rounds of cell division in order to mount an effector response. After eradication of the infectious agent the number of cells is again reduced to an intermediate sized memory pool that can persist for many years. T cells do not recognize antigen directly, but instead recognize antigenic peptides presented on proteins of the major histocompatibility complex (MHC). T cells have powerful regulatory role in responses, with some direct effector functions, but equally or more powerful roles based on organizing responses by B cells and innate cells. B cells make high-affinity antibodies in a manner dependent on T cells for all the highly repetitive carbohydrate antigens. Plasma cells that develop from B cells secrete antibodies and can live for years following an appropriate initial stimulation.

Barriers and Tissue Cells

An often neglected aspect of innate immunity is that role of tissue cells in forming barriers and establishing environments in which infections can be isolated and destroyed. The barrier function of the keratinized epithelium of the skin is illustrated by the vulnerability of burn victims to infection, whereas the single cell layer of the gut epithelium stands between us and trillions of microorganisms. When kept in place our gut microbiota is beneficial, but if they breach the epithelial barrier the resulting systemic infection or sepsis can be devastating. Stromal cells also play critical roles in innate immunity as endothelial barriers and stromal networks control the access of innate cells such as neutrophils and monocytes to sites of infection.

Chemical Defenses

The innate immune system uses chemical warfare against invading microbes. One example of this is the nicotinamide

adenine dinucleotide phosphate (NADPH) oxidase of neutrophils and phagocytes. This system allows neutrophils and other activated innate cells to generate reactive oxygen species that are the biological equivalent of 'bleach' to kill microbes. This type of activity must be carefully controlled to avoid toxicity to the host.

Important chemical defenses also include antimicrobial peptides made at mucosal interfaces to control colonization and the complement cascade in the blood. These have direct toxicity to bacteria, often by disrupting membrane integrity.

Complement is an effector mechanism that was initially studied as a 'complement' to antibodies that imparted microbicidal activity following antigen recognition. Complement is a proenzyme (zymogen) cascade similar in this respect to the clotting cascade. A series of proteins that circulate in the blood as inactive pro-forms assemble on microbial surfaces into active 'convertases' that proteolyse downstream components to generate effector molecules. There are two major outcome of complement activation – the most important is that some of the small proteolytic fragments generated in forming the convertases are potent chemoattractants that recruit neutrophils and monocytes from the blood into the tissues. In addition, the 'terminal' complement components form pores on the surface of microbes and can directly kill bacteria. There are many steps to this pathway that allow for regulation. One characteristic of complement is that the cascade is activated at a low rate on any surface with appropriate attachment sites for highly reactive thioester groups. Host cells are protected from this spontaneous activation of the cascade by host proteins that inactivate the convertase or block formation of pores. Pathogens that lack these protective molecules are attacked – an example of what is referred to as 'missing self' recognition. Microbes can also activate the complement cascade based on mannan-binding proteins that target some microbial cell wall components and through antibodies.

Antimicrobial peptides and complement proteins are always present at some levels, but their production at surfaces (peptide) or in the liver (complement) is dramatically increased in the setting of inflammation. The increased production of complement and other antimicrobial peptides is known as the 'acute phase response.'

Sensing Mechanisms

Sensing mechanisms that use receptors specific for evolutionarily conserved pathogen-associated molecular patterns (PAMPs) or danger-associated molecular patterns (DAMPs) are central to innate immunity and are themselves evolutionarily ancient. The first insights into the molecular basis of pathogen recognition came from classical genetics. Mutants in structurally related proteins in flies (Toll) and mice (Toll-like receptors) are critical for defense against pathogens. There are now 10 Toll-like receptors (TLRs) in mammals that are coupled to activation of the nuclear factor- κ B (NF κ B) transcription factor family to orchestrate induced innate immunity and, as will be discussed below, the link between innate and adaptive immunity for many extracellular pathogens. TLRs are type I transmembrane proteins with extracellular domains that

possess a leucine rich repeat (LRR) that assembles into long, sickle shaped domains that undergo oligomerization in response to specific ligands. Subsets of TLRs operate to sense extracellular microbes at the cell surface with TLR 7 and 8 residing in endosomes to sense viral nucleic acids that are generated as viral particles are uncoated at acidic pH. The cytoplasmic domains of the TLRs possess motifs that undergo oligomerization with cytoplasmic adapters that link through myeloid differentiation primary response gene 88 (MyD88) to activation of NF- κ B and AP1 transcription factors. Viral nucleic acid recognition additionally leads to induction of type I interferon (IFN) (α/β) through interferon regulatory factors (IRF3/7). It has become clear that different signaling adapters are prelocalized in plasma membrane domains and endosomes such that a single plasma membrane TLR can generate different signals before and after endocytosis. Thus, the exact signal is dependent upon the compartment in which ligand is recognized (Kagan, 2012).

Cytoplasmic recognition of PAMPs is achieved through retinoic acid inducible gene I (RIG-I) like receptors (RLRs), nucleotide-binding oligomerization domain-like receptors (NLRs), and other sensors that are less characterized or remain to be identified. Upon binding viral RNA species through RNA helicase-like domains, the RLRs RIG-I and MDA5 associate with mitochondrial antiviral-signaling protein (MAVS) to activate antiviral transcriptional responses. MAVS localizes to mitochondrial surfaces to induce IRFs and NF- κ B transcription factors that mediate expression of type I IFN and other immune genes. Cytoplasmic bacteria and DNA viruses are sensed through the cytoplasmic receptor stimulator of interferon genes (STING) and other factors to generate type I interferon. DNA in the cytoplasm can be converted to cyclic GMP-AMP (cGAMP) by the cGAMP synthase (cGAS) to activate STING localized to the ER surface, subsequently activating similar transcription factors as MAVS. A variety of ligands have been demonstrated to activate LRR domain-containing NLRs, but how these ligands are generated under physiological situations and whether they are sensed through direct binding remains an area of active investigation. The NLR family members NOD1 and NOD2 can localize to the cytosolic surface of endosome and are activated by peptidoglycan motifs derived from internalized bacteria, such as following phagocytosis. NOD2 has been reported to bind RNA from viruses as well. NOD1 and NOD2 signal through the adaptor molecule RIP2, again activating NF- κ B, AP1, and IRF transcription factors. Other NLR family members, such as NOD-like receptor family pyrin domain-containing 3 (NLRP3), oligomerize into large complexes referred to as inflammasomes in the presence of flagellin, pore-forming toxins, adenosine triphosphate (ATP), extracellular glucose, monosodium urate crystals, and other signs of cellular distress. Upon activation, the inflammasome aggregates with pro-caspase-1 to induce auto-cleavage into the active caspase-1 form, which in turn cleaves pro-IL-1 β and IL-18 to generate the biologically active form of these cytokines.

Pathogens interfere with innate sensing mechanisms. For example, hepatitis C virus (HCV), which causes chronic infections in humans, has a protease, NS3-4A, that targets innate signaling adapters MAVS and TIR-domain-containing adapter-inducing interferon- β (TRIF) thus inactivates multiple innate

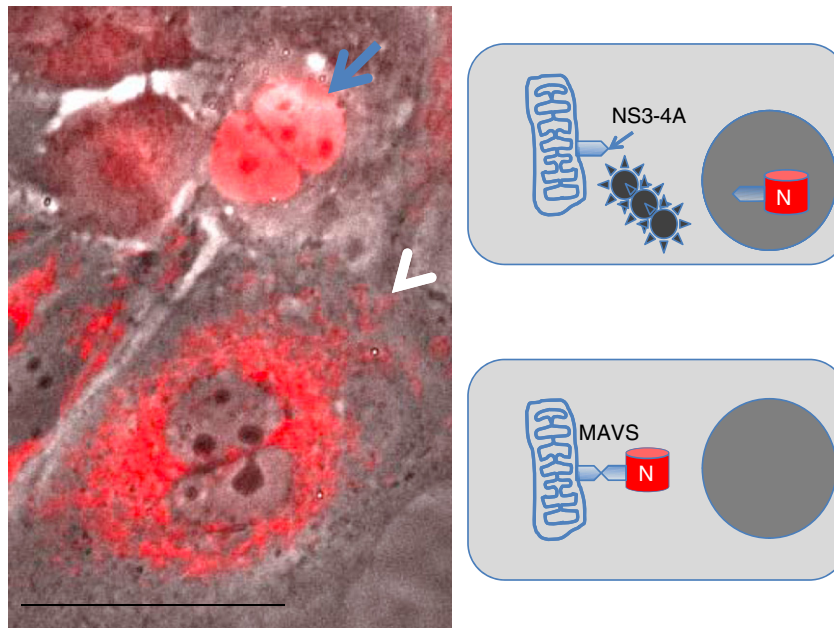


Figure 1 Viral immune evasion by targeting innate cytoplasmic sensor. The NS3-4A protease of hepatitis C-virus cleaves the mitochondrial antiviral-signaling protein, MAVS, releasing it from mitochondrial and mitochondrial-associated membranes. The hepatoma cells shown in this image express a reporter comprised of the MAVS transmembrane domain fused to red fluorescent protein and a nuclear localization sequence (Jones *et al.*, 2010). RFP is localized to the mitochondria in noninfected hepatocytes (white arrowhead), but is released from the mitochondria by the action of NS3-4A in hepatitis C-infected hepatocytes (blue arrow). Left-fluorescence signal in red overlaid on phase contrast micrograph to show relationship to cell morphology. Scale bar=50 μ m. Micrograph courtesy of Stephen M. Laidlaw and Lynn B. Dustin Right-Schematic of the MAVS-based construct and basis for RFP translocation (N=nuclear localization sequence).

sensing pathways. A sensor for *in vitro* detection of HCV infection has been constructed by linking red fluorescent protein with a nuclear localization sequence to MAVS (Jones *et al.*, 2010; Figure 1). The RFP localizes to mitochondria, the normal location of MAVS, in noninfected cells (white Arrowhead) and, after action of NS3-4A, to the nucleus in infected cells (blue Arrow). Even with this evasion mechanism, HCV-infected hepatocytes produce an innate response including the production of the cytokine interferon- λ , highlighting the depth of viral sensing mechanisms (Marukian *et al.*, 2011). NS3-4A also processes the viral polyprotein that shares a sequence motif with the host adapters and is targeted by a protease inhibitor that is part of a highly successful drug combination used to treat HCV.

Phagocytes and Their Receptors

Phagocytes include neutrophils, macrophages, and dendritic cells (DCs), which have the capacity to engulf and digest relatively large particles on the order of 1–10 μ m and even larger. In adults, these cells are generated from hematopoietic stem cells in the bone marrow. Phagocytosis can be activated by receptors that share structural features with T and B cell antigen receptors, in which case the process takes on characteristics of cell–cell interactions with specialized junctions referred to as phagocytic synapses. Phagocytes express many type of ‘scavenger’ receptors that allow them to participate in clearance of aged host proteins and cells or microbe-derived materials. Pathogens may also exploit phagocytes as a port of

entry to colonize the host. One example is the gram positive bacterium *Listeria monocytogenes*, which generates a number of interesting virulence factors. One is called listeriolysin O, which allows the bacterium to lyse the phagocytic vacuole and directly enter the phagocyte’s cytoplasm. A second virulence factor, called Act A, then induces dynamic actin polymerization by activating the Actin-related protein 2/3 complex to generate ‘actin rockets’ that allow *Listeria* to move from cell to cell without passing through the extracellular space, thus avoiding extracellular immune effectors. Due to their ability to take up extracellular particles, phagocytes such as macrophages and DCs are excellent antigen presenting cells (APCs) that activate lymphocytes. B cells are also potent APCs because they can internalize antigens bound to the B cell receptor, which is essentially a plasma membrane-bound antibody. Therefore, although phagocytosis can lead to destruction of internalized microbes, the process also functions as a bridge between innate and adaptive immunity.

Innate Lymphoid Cells

Innate lymphoid cells (ILC) look physically and, to some extent, transcriptionally like lymphocytes, but they lack antigen receptors. They develop in the bone marrow from hematopoietic progenitor cells and are found in the blood, lymphoid tissues, and peripheral tissue in small numbers compared to T and B lymphocytes. The best known of these cells are the natural killer (NK) cells as these are relatively abundant and have direct effector function in killing tumor cells and some

virally infected cells. In the current nomenclature system for these cells there are three broad types of ILC: the group 1 ILCs that express the master transcription factor Tbet and make the signature cytokine IFN- γ ; the group 2 ILCs that express high levels of the transcription factor GATA-3 and make the signature cytokine IL-5, IL-13, and amphiregulin; and group 3 ILCs that express the master transcription factor ROR γ t and make signature cytokines IL-17 and IL-22 (Spits *et al.*, 2013). Each group of ILCs can be further divided into specialized subsets that express different combinations of their signature cytokines and effector molecules. As becomes clear below, the types of ILC parallel the types of T cells and the three programs are currently believed to reflect the coordinated responses needed to respond to the three major categories of pathogens—type 1 being viruses and intracellular bacteria; type 2 being parasites; and type 3 being extracellular bacteria. As with T cells, the ability of ILC to kill is imparted by the master transcriptional regulator eomesodermin in combination with Tbet, which also reflects a major transcriptional difference between type 1 helper T cells and cytotoxic T cells. Perhaps due to their highly similar effector programs and extensive specialization for cell-mediated killing, NK cells and cytotoxic T cells are more transcriptionally related to each other than NK cells are to other ILCs (Best *et al.*, 2013).

Lymphocytes

Lymphocytes are the core of adaptive immunity and were so named because they are abundant in thoracic duct lymph – the common return path of lymph fluid from the tissues to the blood. Removal of lymphocytes from an animal through thoracic duct cannulation severely impairs adaptive ‘humoral’ and ‘cellular’ immune responses. Lymphocytes can be divided into two types as briefly mentioned above. B lymphocytes or B cells (B is for Bursa, the hematopoietic site in birds) develop in the bone marrow of mammals and their effector function is to develop into antibody-secreting cells referred to as plasma cells. T lymphocytes or T cells (T is for thymus) develop in the thymus (see lymphoid tissues below) and in the periphery they engage in cellular immunity. T helper cells produce cytokines that orchestrate antibody production and other aspects of the immune response, in many cases through direct cell–cell contact. The key common feature of T and B cells is that they both have antigen receptors that exist in pieces in the germline that need to be assembled by a somatic gene rearrangement process mediated in part by recombinase-activating genes that are expressed in two precise periods in B cell and T cell development corresponding to sequential recombination of two loci for each antigen receptor. Once the T cell antigen receptor (TCR) and BCR are made these cells undergo selection processes to eliminate auto-reactive cells (negative selection) and in the case of T cells a process of positive selection that ensures that T cells are predisposed to recognize appropriately processed antigens. Antigen processing is discussed in the context of DCs. In the thymus, both DCs and epithelial cells engage in the negative and positive selection processes with the thymocytes. The mature, selected B and T cells leave the primary lymphoid tissues to begin recirculation through the blood and secondary lymphoid tissues (see below).

Immunological Synapses

Adaptive immune responses require at least three different central cell types – DCs and the two types of lymphocytes. In order to be able to help B cells make antibodies, T cells must be ‘primed’ and this is most efficiently done by a DC. The priming process requires that the DC takes antigens into its endocytic pathway, and through partial proteolytic processing, generate peptides that bind the MHC proteins. Dendritic cells can introduce antigens taken up from outside the cell into two different types of MHC molecules. MHC class Ia molecules bind peptides in the endoplasmic reticulum that are generated in the cytoplasm by the proteasome. Dendritic cells are able to ‘cross-present’ in that they can introduce proteins that are taken into the endocytic pathway into the cytoplasm to generate peptides that are then presented on MHC class Ia. In other cell types, the presentation of a peptide derived from an infectious agent on MHC class Ia would indicate that the cell had the pathogen in its cytoplasm and it could thus become a target for killing, but DCs have the ability to present peptides on MHC class Ia even when they are not infected directly (Bevan, 2006). MHC class II molecules bind peptides in the endocytic pathway because they are blocked from binding peptides in the endoplasmic reticulum by a chaperone protein called invariant chain. MHC proteins are germline encoded, so they are fixed within an individual, but they are highly polymorphic across animal populations. This is important because the six MHC class Ia and six MHC class II genes in each individual human, for example, may not cover all antigenic peptides and this diversity in a population can be important for survival of a species faced with an epidemic of a particular deadly pathogen like *Variola major* (smallpox virus) or *Yersinia pestis* (bubonic plague bacterium). MHC class Ib genes include β 2 microglobulin-associated peptide-binding proteins like H2-M3 and Qa-1 in the mouse and human leukocyte antigen (HLA-E) in human. HLA-E binds signal peptides derived from MHC class Ia proteins and serves as a ligand for inhibitory receptor CD94/NKG2A on NK cells and CD8⁺ T cells, sensing pathogen interference with MHC class Ia protein expression. This is a specific example of missing self-recognition – the manner in which NK cells sense the presence of viruses that interfere with MHC class Ia expression to avoid having infected cells killed by cytotoxic T cells (Karre, 2008). The cell biology of MHC class I and II presentation has been intensively studied, but surprisingly, the underlying mechanisms that allow cross-presentation by DCs are controversial.

Once a peptide–MHC complex is expressed on an APC, it may potentially activate a T cell with a complementary TCR. This process requires close membrane apposition during recognition in the form of an immunological synapse or a kinapse. The TCR and peptide–MHC will span about 15 nm when engaged and this close proximity is facilitated by adhesion molecules in close coordination with the actin cytoskeleton. The TCR interaction with the peptide–MHC complex is weak in solution, but seems perfectly adapted to operate in a synaptic junction. The basic chemistry and physics of this environment is an area of active study. The downstream signal transduction process is well defined, although there are many nuances to this including how distinct co-stimulatory signals are integrated with TCR signals. We would define the

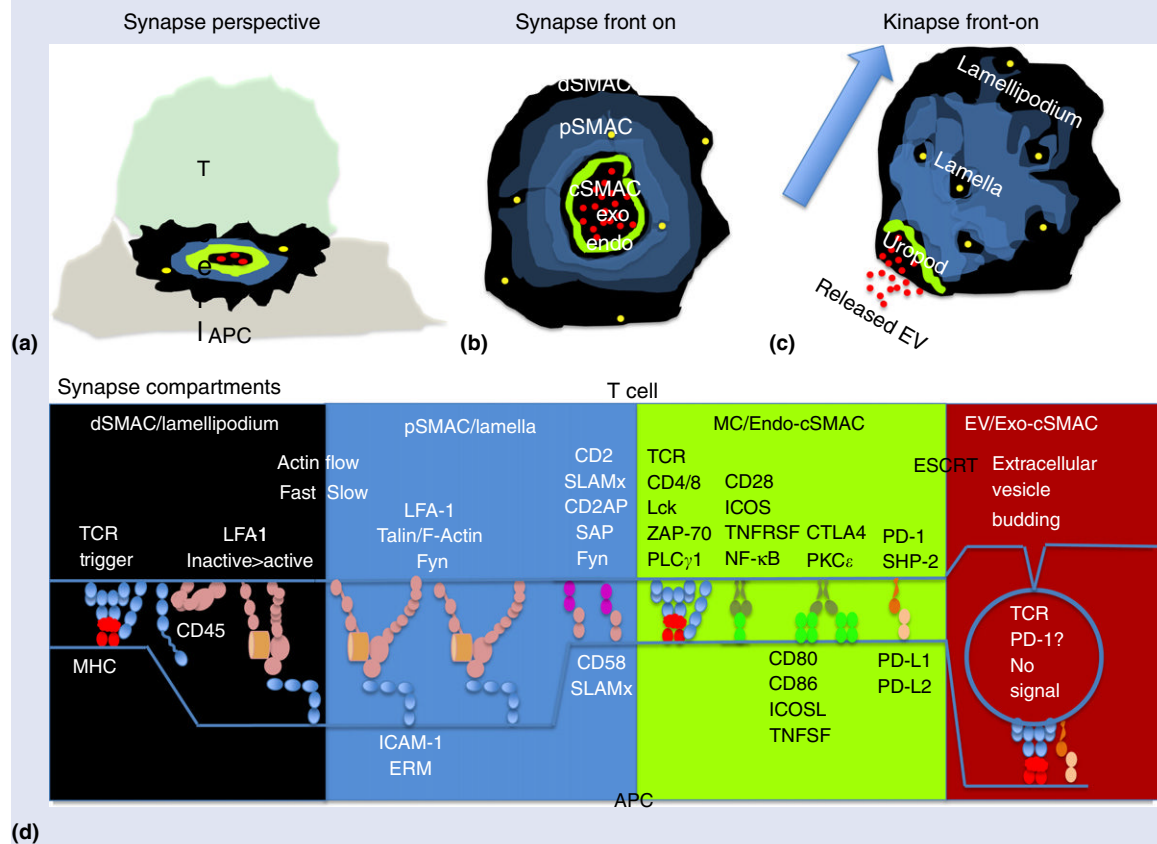
immunological synapse as a (1) stationary cell–cell junction, (2) built from bona fide adhesion molecules, (3) with contiguous membranes, and (4) vectorial communication by secretion (Dustin and Colman, 2002; Dustin *et al.*, 1998). It has been observed *in vitro* and *in vivo* that T cells can recognize antigen through a stationary or moving junction (Dustin, 2007).

Thus we have stationary synapses and mobile kinapses (Box 1). Early work on stationary junctions defined three compartments referred to as supramolecular activation clusters (SMACs) (Freiberg *et al.*, 2002; Monks *et al.*, 1998). Analysis of the F-actin dynamics in the immunological synapse suggests that the distal SMAC (dSMAC) corresponds to a lamellipodium, the peripheral SMAC (pSMAC) to a lamella and the central SMAC (cSMAC) to the uropod of a migrating mesenchymal cell (Sims *et al.*, 2007; Yi *et al.*, 2012). The major adhesion molecules on human lymphocytes are CD2, which is a member of the signaling lymphocyte activation molecule (SLAM) family, and lymphocyte function-associated antigen-1 (LFA-1), which is a member of the integrin family (Springer,

1990). The ligand for CD2 is CD58 (also called LFA-3) and the ligand for LFA-1 is intercellular adhesion molecule-1 (ICAM-1). These molecules concentrate in different compartments of the pSMAC during immunological synapse formation, whereas the TCR and co-stimulatory molecules like CD28 accumulate first in microclusters (MC) that form in the dSMAC and traverse the pSMAC in gaps within the adhesion ring to converge on the cSMAC. In the cSMAC ubiquitinated TCR are segregated from the CD28 signaling complexes forming two compartments by the endosomal sorting complexes required for transport (ESCRT) – an endo cSMAC that is active in signaling and an exo-cSMAC that is a site of TCR signal termination and either endocytosis for degradation or budding of extracellular vesicles (EV) into the synaptic cleft for transfer to the APC (Choudhuri *et al.*, 2014). In the stationary synapse, the motile machinery of the lamellipodium-like dSMAC is prevented from moving the T cell by the overall radial symmetry, which results in balanced forces and focusing of molecules toward the center. The conversion of a synapse to a kinapse requires a symmetry breaking event in the pSMAC,

Box 1 Model immunological synapse

(a) Schematic of T cell–APC interface with a centralized immunological synapse (b) and (c). Front-on view of model for central synapse organization and how this related to kinapse – a mobile junction with similar compartments. (d) Schematics with zones showing a small subset of adhesion and recognition molecules involved. Color code: black, F-actin lamellipodium; blue, LFA-1-ICAM-1 adhesion; red, TCR-MHC antigen recognition; and green, CD28-CD80 costim SMACs are labeled. Note that the cSMAC is an intensive site for membrane trafficking including a directed budding of T cell receptor enriched extracellular microvesicles in the immunological synapse



which is the force generating element of the synapse/kinapse system (Sims *et al.*, 2007). In the kinapse, the TCR microclusters form similarly, but the T cell continues to move and the TCR enriched extracellular vesicles (EV) are left behind with the antigen presenting surface (Choudhuri *et al.*, 2014). There is also important vesicular trafficking of TCR that contributes to synapse formation (Das *et al.*, 2004).

This model for formation of a radially symmetric immunological synapse is largely based on work with supported planar bilayers rather than live APC. Not all live APC induce T cells to form the radially organized SMACs, yet T cells can still form stationary synapses with these APC. A prominent example is DCs, which forms multifocal synapses that lack obvious radial symmetry in many studies (Brossard *et al.*, 2005; Gunzer *et al.*, 2004; Tseng *et al.*, 2008). This has been observed with multiple DC types including splenic DC and bone marrow-derived DC generated with granulocyte-macrophage colony stimulating factor and IL-4. This issue is explored in two other chapters of this encyclopedia. Studies by Wulfigg and colleagues define a deep lamellar F-actin component even in the T-B junctions that do form SMACs (Roybal *et al.*, 2013). Parallel work from Carman and colleagues define multiple invadopodia-like protrusions (ILP) that push into endothelial APC to form a multifocal synapse (Sage *et al.*, 2012). Recently we have found working with Carman that TCR microclusters formed on supported lipid bilayers have similar molecular machinery to ILPs, particularly requiring Wiscott-Aldrich syndrome protein (WASP) and containing phosphorylated hematopoietic lineage cell-specific protein 1 (HS1) (Kumari *et al.*, 2015). Figure 2 suggests a unifying model in which a similarly equipped T cell can interact with two different APC to

form a radially symmetric synapse based on MC or a multifocal synapse based on ILP. While the multifocal synapse lacks radial symmetry, the junction is still stationary because the ILP protrusions effectively entangle the T cell and APC. While symmetry break is required for conversion of a MC type synapse to a kinapse, the ILPs would need to be withdrawn to allow the T cell to disengage from the ILP type synapse. We know that TCR signaling by MC is terminated in ~ 2 min in the cSMAC, but the natural lifetime of the ILP and how they are turned over remains to be determined.

Antigens presented on MHC class Ia are recognized by CD8⁺ cytotoxic T cell precursors that undergo priming on DCs, expand rapidly and then leave the secondary lymphoid tissue (assuming this is not the site of infection) and migrate through the blood to sites of infection where they extravasate and kill peptide-MHC expressing cells. The radially symmetric central synapses have a well-defined secretory domain and may be the most effective at killing (Stinchcombe *et al.*, 2001). Killing is threefold more effective through a stationary synapse compared to a kinapse (Beal *et al.*, 2008). They also make cytokines and chemokines and these can recruit additional cells that support pathogen eradication and the transition to repair. The number of CD8 T cells then contracts and if antigen is cleared a population of memory cells remains to respond more effectively to future encounters. If the infection is not resolved then chronic infection may result with some balance between control of the infection and immunopathology resulting from the chronic immune response.

Antigenic peptides presented on MHC class II complexes are recognized by CD4⁺ helper T cells that undergo priming on DCs, expand and then differentiate into diverse helper T cell

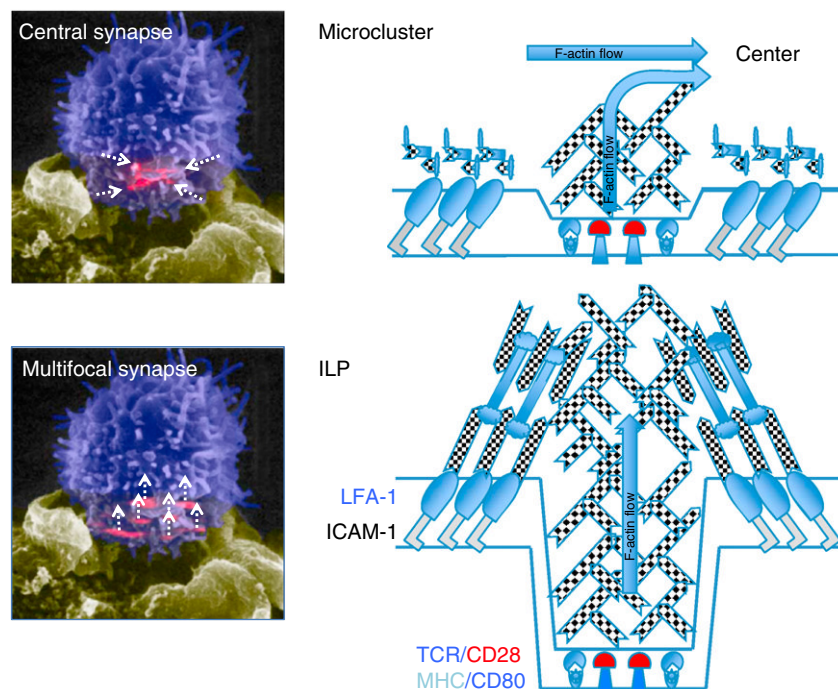


Figure 2 Immunological synapse formation on stiff and soft APC. TCR/CD28 signaling is associated with central clusters that are transported by F-actin and myosin II-mediated centripetal transport on round, relatively stiff APC like B cells (*in vitro*). Paradoxically, softer cells in which the TCR clusters can protrude through the cortical actin of APC will not permit lateral transport and instead the dominant F-actin flow is away from the synapse, rather than toward the center. In this case the TCR/CD28 signaling foci are multifocal.

types. The helper types Th1, Th2, and Th17 correspond roughly to the type 1, 2, and 3 ILC, respectively. A fourth subset of helper T cells, T follicular helper cells, or Tfh cells, express the chemokine receptor CXCR5 and migrate to B cell follicles to provide contact dependent help to B cells to form germinal centers. Engagement of the ICOS co-stimulatory receptor further reinforces this pathway (Liu *et al.*, 2014). Germinal center B cells engage in cycles of interaction with Tfh in the 'light zone' and cell division and somatic hypermutation and class switching in the 'dark zone' of the germinal center. The other Th types influence the process of class switching and generate effector cells that migrate to the site of infection to orchestrate the immune response. However, the innate recognition of the pathogen by DCs and other tissue cells, as well as some information associated with peptide-MHC levels, determine which effector type the non-follicular Th assume. Type I interferon and IL-12 stimulated by viral and intracellular bacterial infection promote the differentiation to Th1 cells that make interferon- γ and promote macrophage activation and promote B cell class switching to Fc receptor-binding cytotoxic antibody isotypes. Proteases and other signatures associated with parasitic infection favor generation of Th2 cells, which make IL-4, IL-5, and IL-13 and promote antiparasitic responses in tissues and isotype switching to IgE. In the presence of inflammasome activation leading to IL-1 β and IL-6 production that favor the differentiation Th17 cells, which make cytokines IL-17 and IL-22 that induce neutrophil recruiting chemokines and reinforce epithelial barrier function as a defense against extracellular bacterial infections. These phenotypes can be quite plastic with cells making both IFN- γ and IL-17 observed in some immune response contexts. CD4 T cell memory may thus be more diverse due to the different effector types. Th cells also receive inputs from a number of TNF receptor family members that influence effector function and can promote memory development across different helper types. Due to the antigen specificity of these effector functions, Th may be able to act more aggressively than innate like cells because their activity can be more precisely controlled in both space and time.

A subset of CD4⁺ T cells are referred to as regulatory T cells that express the master transcription factor FoxP3. These cells play an essential role in maintaining immune homeostasis, at least in part by managing the immunogenicity of DCs through trans-endocytosis, a process in which the regulatory T cells strip DCs of co-stimulatory molecules that are necessary for activating other T cells. A population of regulatory T cells is generated in the thymus and another population is generated in the periphery from naïve CD4⁺ T cells in response to self-antigens or to control responses to benign environmental antigens (Weiss *et al.*, 2012). The induction of regulatory T cells *in vitro* can be achieved with high concentrations of transforming growth factor- β , which is also one of the effector molecules employed by these cells to regulate other immune cells. Regulatory T cells are marked by high expression of the CD25 subunit of the growth promoting cytokine IL-2 receptor. This is of functional relevance as FoxP3⁺ regulatory T cells require IL-2 to survive, and one of the ways in which they suppress growth and differentiation of effector T cells is to consume IL-2 in a competitive manner (Furtado *et al.*, 2002). FoxP3⁺ regulatory T cells seem to shadow the various effector

helper types such that there are regulatory T cells in all tissue compartments including the germinal centers. There is evidence that CD8⁺ T cells can also be involved in regulation of immune responses, but this may be a more general feedback process related to killing. Individuals with defects in killing mechanisms shared by CD8⁺ T cells and NK cells are vulnerable to hemophagocytic syndrome, which is immunopathology caused by uncontrolled activation of phagocytes (Menasche *et al.*, 2005). A specific subpopulation of CD8 T cells that is restricted by the Qa1/HLA-E MHC class Ib molecules are also involved in regulation of recall responses by CD4⁺ T cells, probably by direct killing of the the Qa-1/HLA-E positive T cells (Kim and Cantor, 2011).

Lymphoid Tissues

Lymphoid tissues (LT) are specialized for development or function of immune cells. There are three types of LT: primary LT where developmental steps takes place, secondary LT that are involved in steady state surveillance, and tertiary LT that are generated *de novo* to support local immune responses associated with chronic infection or autoimmunity. In mammals, the major primary lymphoid tissues are the bone marrow and thymus. The bone marrow functions in generation of red blood cells, myeloid cells of innate immunity, and B lymphocytes. The output of neutrophils is the most intensive activity and perhaps also the most responsive process to signals associated with inflammation. Production of B cells is tightly linked to the rearrangement of antibody genes. T cell progenitors leave the bone marrow and colonize the thymus. The thymus is an organ that sits on top of the heart and has extensive epithelial components that are organized within the cortical and medullary regions with post-capillary venules at the cortico-medullary junction. Developing T cells, called thymocytes, rearrange the TCR genes in the cortex and initiate selection processes. Selection is completed in the medulla, where specialized epithelial cells express the auto-immune regulator transcription factor (AIRE), which allows these cells to express large segments of tissue-specific genes in the thymus to facilitate negative selection or thymic regulatory T cell development in response to self-antigens (Anderson *et al.*, 2002).

Secondary lymphoid tissues include the lymph nodes, spleen, tonsils, Peyer's patches, and bronchi associated lymphoid tissues. Lymphoid tissue development depends upon a type 3 ILC referred to as a lymphoid tissue inducer cell, which uses TNF family cytokines such as lymphotoxin- β to induced reorganization of at least four stromal cell types to generate distinct, but adjacent T cell and B cell zones with site specific entry points for antigens and entry and exit points for DCs and lymphocytes. T and B cell zones are defined by distinct stromal cells that make chemoattractants CCL19 and CCL21 for T cell zones and CXCL13 for B cell zones. Within the T cell zones the fibroblastic reticular cells (FRC) form a CCL19 positive network that drives rapid migration in the steady state. This brings DCs that form an overlying network on the FRC in contact with ~1000 T cells per hour. Within the B cell zones, also known as follicles, the follicular dendritic cells (FDC) form a CXCL13 positive network for follicular B

cell migration and germinal center formation. The FDC in the light zone of the germinal center because further specialized and express complement receptor such that they can capture and present antigen containing immune complexes to provide a source of antigen to drive the necessary mutation/selection cycles in the germinal center (Victoria and Nussenzweig, 2012).

The lymph nodes are closely associated with the blood and lymphatic vasculature such that they effectively create a tissue island within an expanded lymphatic vessel that is supported by a blood capillary network with specialized post-capillary venules. Lymph nodes samples antigens from the lymph draining most vascularized tissues and organs and are thus important sites for function of subcutaneous and intramuscular injected vaccines. These can reach the lymph nodes as soluble antigens, immune complexes (with antibodies and complement), or associated with migratory DCs. The secondary lymphoid tissues of the spleen are referred to as the white pulp and form as a periarteriolar lymphatic sheath. Arterioles from this central arterioles open into the marginal zone and red pulp sinuses. Marginal zone B cells are specialized to shuttle antigens from the marginal zone to follicular DCs. The FDCs then to make these accessible to capture by BCR of follicular and germinal center B cells, which process these antigens to peptide-MHC complexes for presentation to follicular helper T cells. Antigens reach the T cell zone of the spleen associated with DCs that migrate from the marginal zone. The Peyer's patches, tonsils, and bronchi associated lymphoid tissues samples antigens across the mucosal epithelium. The best studied mechanism involves M-cells over the epithelium of the Peyer's patch that allows sampling of food and microbial components from the small intestine and transfer of these to DCs.

T and B cells enter lymph nodes from the blood through high endothelial post-capillary venules that express peripheral lymph node addressin (PNAd), chemokine CCL21, and ICAM-1. T and B lymphocyte enter lymph nodes by initially undergoing rolling adhesion in blood flow using CD62L to latch onto PNAd, then being activated by chemokine receptor CCR7 interaction with chemokine CCL21 that activates LFA-1 mediated adhesion to ICAM-1 to stabilize interaction with the endothelial cells and allow extravasation. This is referred to as the multistep paradigm for blood cell homing (Springer, 1994). Variations on this are used for entry into other lymphoid tissues. For example, in the gut, the role of PNAd and ICAM-1 is largely replaced by the mucosal addressin cell adhesion molecule (MAdCAM), which is a ligand for the lymphocyte Peyer's patch adhesion molecule (LPAM-1), which is a distinct integrin family member from LFA-1. Entry into the spleen is more complex and involves capture in the red pulp venous sinuses and migration through bridging channels into the T cell zones, or migration through the marginal zone directly into follicles for B cells. Once T and B cells have crossed the endothelium, they migrate toward the respective zones.

Tertiary lymphoid tissues emerge in sites of chronic immune responses, including in autoimmunity. Well known examples include the lymphoid aggregates in synovium of patients with rheumatoid arthritis and chronic infections such as *Helicobacter pylori*. In the cancer setting, the presence of

tertiary lymphoid tissue in colorectal cancer is a good prognostic indicator. On the other hand, stromal cells making CXCL12 in pancreatic cancer may divert tumor infiltrating lymphocytes from mounting effective responses.

Stromal cells can also participate directly in regulation of immune responses. FRC can suppress immune responses through interferon- γ stimulated nitric oxide production (Lukacs-Kornek *et al.*, 2011). On the other hand, endothelial cells can up-regulate MHC class II in response to interferon- γ and participate in activation of helper T cells.

Evolution

Innate immunity is an ancient process that is common to all multicellular organisms. Adaptive immunity based on immunoglobulin superfamily antigen receptors is shared by jawed fish, reptiles, amphibians, birds, and mammals. Interestingly, larval forms of jawless fish possess an adaptive immune system based on somatic cell gene rearrangement of LRR-based receptors referred to as variable lymphocyte receptors (VLRs) (Boehm *et al.*, 2012). There are two types of VLR expressing lymphocytes that appear to correspond to B cell-like and T cell-like roles, although the equivalent of an MHC for the TCR-like VLR has not been identified. Thus, the evolution of adaptive immune systems with humoral and cellular components has devolved at least twice on top of existing innate components with cells that have specialized secretory and cellular effector functions.

Perspective

The immune system provides a treasure trove for functional cell biology. This is in part because in this system pathogens act as our teachers and call our attention to functionally significant and nonredundant elements in the immune system that are often shared with other developmental and homeostatic processes in the body. It has also become clear that improvements in treatment for autoimmune diseases and better immunotherapies will only be possible with a better understanding of the functional cell biology of the immune system. This section of the *Encyclopedia of Cell Biology* defines our current knowledge about cellular aspects of the immune system, and certainly also exposes gaps in knowledge that can be filled using tools of cell biology.

See also: Cellular Immunology: Cell-Cell Interactions and the Immune System: Cellular Structures Controlling T Cell Signaling in Time and Space; CD28 Costimulation and Regulatory T Cells; CXCL12/SDF-1 and Hematopoiesis; Lymphocyte-Endothelial Interactions; Roles of Stromal Cells in the Immune System; Single Molecule Methods to Measure Receptor-Ligand Interaction in Immunological Synapses; T Cell Receptor Triggering. Cellular Immunology: Innate Immunity: CD4 T Cell Memory and Role of TNF Receptor Family; Dendritic Cells; Function of Epithelial Barriers; Neutrophil Biology; Phagocytic Synapses; Scavenger Receptors; Specialized Subsets of Tissue-Resident Macrophages in

Secondary Lymphoid Organs; The Phagocyte NADPH Oxidase: Structure and Assembly of the Key Multicomponent Enzyme of Innate Immunity. **Cellular Immunology: Transcriptional Basis of B and T cell Lineages and Memory Cells: NF-kappaB and the Immune System; T Cell Memory to Viral Infections; T Follicular Helper Cells. Stem Cell Biology: Structure and Function: Pluripotent Cells: New Tools for Disease Research and Therapies**

References

- Anderson, M.S., Venanzi, E.S., Klein, L., *et al.*, 2002. Projection of an immunological self shadow within the thymus by the aire protein. *Science* 298, 1395–1401.
- Beal, A.M., Anikeeva, N., Varma, R., *et al.*, 2008. Protein kinase C(θ) regulates stability of the peripheral adhesion ring junction and contributes to the sensitivity of target cell lysis by CTL. *Journal of Immunology* 181, 4815–4824.
- Best, J.A., Blair, D.A., Knell, J., *et al.*, 2013. Transcriptional insights into the CD8(+) T cell response to infection and memory T cell formation. *Nature Immunology* 14, 404–412.
- Bevan, M.J., 2006. Cross-priming. *Nature Immunology* 7, 363–365.
- Boehm, T., McCurley, N., Sutoh, Y., *et al.*, 2012. VLR-based adaptive immunity. *Annual Review of Immunology* 30, 203–220.
- Brossard, C., Feuillet, V., Schmitt, A., *et al.*, 2005. Multifocal structure of the T cell – Dendritic cell synapse. *European Journal of Immunology* 35, 1741–1753.
- Choudhuri, K., Llodra, J., Roth, E.W., *et al.*, 2014. Polarized release of T-cell-receptor-enriched microvesicles at the immunological synapse. *Nature* 507, 118–123.
- Das, V., Nal, B., Dujancourt, A., *et al.*, 2004. Activation-induced polarized recycling targets T cell antigen receptors to the immunological synapse; involvement of SNARE complexes. *Immunity* 20, 577–588.
- Dustin, M.L., 2007. Cell adhesion molecules and actin cytoskeleton at immune synapses and kinapses. *Current Opinion in Cell Biology* 19, 529–533.
- Dustin, M.L., Colman, D.R., 2002. Neural and immunological synaptic relations. *Science* 298, 785–789.
- Dustin, M.L., Olszowy, M.W., Holdorf, A.D., *et al.*, 1998. A novel adapter protein orchestrates receptor patterning and cytoskeletal polarity in T cell contacts. *Cell* 94, 667–677.
- Freiberg, B.A., Kupfer, H., Maslanik, W., *et al.*, 2002. Staging and resetting T cell activation in SMACs. *Nature Immunology* 3, 911–917.
- Furtado, G.C., Curotto de Lafaille, M.A., Kutchukhidze, N., Lafaille, J.J., 2002. Interleukin 2 signaling is required for CD4(+) regulatory T cell function. *Journal of Experimental Medicine* 196, 851–857.
- Gunzer, M., Weishaupt, C., Hillmer, A., *et al.*, 2004. A spectrum of biophysical interaction modes between T cells and different antigen-presenting cells during priming in 3-D collagen and *in vivo*. *Blood* 104, 2801–2809.
- Jones, C.T., Catanese, M.T., Law, L.M., *et al.*, 2010. Real-time imaging of hepatitis C virus infection using a fluorescent cell-based reporter system. *Nature Biotechnology* 28, 167–171.
- Kagan, J.C., 2012. Defining the subcellular sites of innate immune signal transduction. *Trends in Immunology* 33, 442–448.
- Karre, K., 2008. Natural killer cell recognition of missing self. *Nature Immunology* 9, 477–480.
- Kim, H.J., Cantor, H., 2011. Regulation of self-tolerance by Qa-1-restricted CD8(+) regulatory T cells. *Seminars in Immunology* 23, 446–452.
- Kumari, S., Depoil, D., Martinelli, R., *et al.*, 2015. Actin foci facilitate activation of the phospholipase C- γ in primary T lymphocytes via the WASP pathway. *eLife* 4, e04953.
- Liu, D., Xu, H., Shih, C., *et al.*, 2014. T-B-cell entanglement and ICOSL-driven feed-forward regulation of germinal centre reaction. *Nature* 517, 214–218.
- Lukacs-Kornek, V., Malhotra, D., Fletcher, A.L., *et al.*, 2011. Regulated release of nitric oxide by nonhematopoietic stroma controls expansion of the activated T cell pool in lymph nodes. *Nature Immunology* 12, 1096–1104.
- Marukian, S., Andrus, L., Sheahan, T.P., *et al.*, 2011. Hepatitis C virus induces interferon-lambda and interferon-stimulated genes in primary liver cultures. *Hepatology* 54, 1913–1923.
- Menasche, G., Feldmann, J., Fischer, A., de Saint Basile, G., 2005. Primary hemophagocytic syndromes point to a direct link between lymphocyte cytotoxicity and homeostasis. *Immunological Reviews* 203, 165–179.
- Monks, C.R., Freiberg, B.A., Kupfer, H., Sciaky, N., Kupfer, A., 1998. Three-dimensional segregation of supramolecular activation clusters in T cells. *Nature* 395, 82–86.
- Roybal, K.T., Sinai, P., Verkade, P., Murphy, R.F., Wulfig, C., 2013. The actin-driven spatiotemporal organization of T-cell signaling at the system scale. *Immunological Reviews* 256, 133–147.
- Sage, P.T., Varghese, L.M., Martinelli, R., *et al.*, 2012. Antigen recognition is facilitated by invadosome-like protrusions formed by memory/effector T cells. *Journal of Immunology* 188, 3686–3699.
- Sims, T.N., Soos, T.J., Xenias, H.S., *et al.*, 2007. Opposing effects of PKC θ and WASp on symmetry breaking and relocation of the immunological synapse. *Cell* 129, 773–785.
- Spits, H., Artis, D., Colonna, M., *et al.*, 2013. Innate lymphoid cells – A proposal for uniform nomenclature. *Nature Reviews Immunology* 13, 145–149.
- Springer, T.A., 1990. Adhesion receptors of the immune system. *Nature* 346, 425–434.
- Springer, T.A., 1994. Traffic signals for lymphocyte recirculation and leukocyte emigration: The multistep paradigm. *Cell* 76, 301–314.
- Stinchcombe, J.C., Bossi, G., Booth, S., Griffiths, G.M., 2001. The immunological synapse of CTL contains a secretory domain and membrane bridges. *Immunity* 15, 751–761.
- Tseng, S.Y., Waite, J.C., Liu, M., Vardhana, S., Dustin, M.L., 2008. T cell-dendritic cell immunological synapses contain TCR-dependent CD28-CD80 clusters that recruit protein kinase C θ . *Journal of Immunology* 181, 4852–4863.
- Victoria, G.D., Nussenzweig, M.C., 2012. Germinal centers. *Annual Review of Immunology* 30, 429–457.
- Weiss, J.M., Bilate, A.M., Gobert, M., *et al.*, 2012. Neuropilin 1 is expressed on thymus-derived natural regulatory T cells, but not mucosa-generated induced Foxp3+ T reg cells. *Journal of Experimental Medicine* 209, 1723–1742, S1721.
- Yi, J., Wu, X.S., Crites, T., Hammer III, J.A., 2012. Actin retrograde flow and actomyosin II arc contraction drive receptor cluster dynamics at the immunological synapse in Jurkat T cells. *Molecular Biology of the Cell* 23, 834–852.

CELLULAR IMMUNOLOGY: TRANSCRIPTIONAL BASIS OF B AND T CELL LINEAGES AND MEMORY CELLS

Contents

The T-Cell Receptor Signalosome

T Follicular Helper Cells

V(D)J Recombination: Orchestrating Diversity without Damage

T Cell Memory to Viral Infections

Nuclear Factor of Activated T Cells and Tolerance

NF-kappaB and the Immune System

The T-Cell Receptor Signalosome

M Mingueneau, Biogen Idec, Cambridge, MA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

All jawed vertebrates have lymphoid organs like the thymus and spleen and genes encoding T and B cell receptor chains. T lymphocytes mature in the thymus and play a central role in cellular immunity, by regulating other immune cells and by differentiating into cytotoxic effector cells that can also make cytokines. T cells express a clonally distributed receptor composed of two main chains, $\alpha\beta$ or $\gamma\delta$ T-cell receptor (TCR). Ligands of $\gamma\delta$ T cell receptors are not well-characterized and this article will focus mainly on $\alpha\beta$ T lymphocytes. With the exception of a few TCRs, which utilize invariant $\alpha\beta$ TCR to recognize bacterial glycolipids associated to CD1d or sense bacterial metabolites associated with MR1, most ligands of $\alpha\beta$ TCRs are peptides (p) associated to class I and class II major histocompatibility molecules (p-MHC). Major histocompatibility complex (MHC) class I binds peptides derived from endogenously synthesized proteins, whereas MHC class II presents peptides derived from proteins taken into the cell from the extracellular space. The $\alpha\beta$ TCR endows T cells with the ability to recognize foreign peptide antigens (e.g., from bacteria, virus, or tumor origin) with high sensitivity while being tolerant to the individual self-peptide antigens in complex with the same MHC molecules. How does the cellular machinery recognize and tolerate self-p-MHC complexes, and differentiate them from foreign p-MHC complexes to trigger the necessary immunological responses?

The TCR belongs to a family of immunoreceptors which also include the B-cell receptor (BCR) and the receptors for the Fc portion of immunoglobulins (FcR). These receptors share a similar membrane organization with multiple transmembrane chains having distinct ligand-binding and signal transduction functions. No intrinsic enzymatic activity is associated with

these receptors. Instead, non-covalently associated kinases from the SRC and spleen tyrosine kinase (SYK) families initiate signaling cascades and transcriptional programs associated with T cell activation and differentiation. Signal initiation requires the conversion of the initial ligand recognition event into an intracellular biochemical modification. This signal conversion step involves the assembly of a large complex of multifunctional proteins in the vicinity of the TCR, collectively designed as TCR signalosome. The signaling code used to transform the properties of the ligand into biochemical signals has to account for the high discriminatory potential of the TCR for foreign and self-antigens, i.e., ligands of similar structure. The coding, conversion, and transmission of receptor–ligand properties by cell-surface receptors are problems of general interest in cell biology. Unlike most receptors, the TCR is not germline-encoded due to DNA recombination events occurring at the TCR locus during T cell development in the thymus. As a result, each of the millions of different TCRs expressed by each individual does not recognize one single ligand, but a limited series of p-MHC ligands with different affinities. The coding of the information associated with these ligands by the TCR signalosome is thus the focus of much interest and is the object of this article.

The TCR and Its Ligands: Structural and Functional Topology

The TCR, a Tyrosine Kinase-Coupled Receptor

Receptors that signal by activating tyrosine kinase activities either contain a kinase domain within their cytoplasmic tail (receptor tyrosine kinases) or are coupled to cytosolic tyrosine

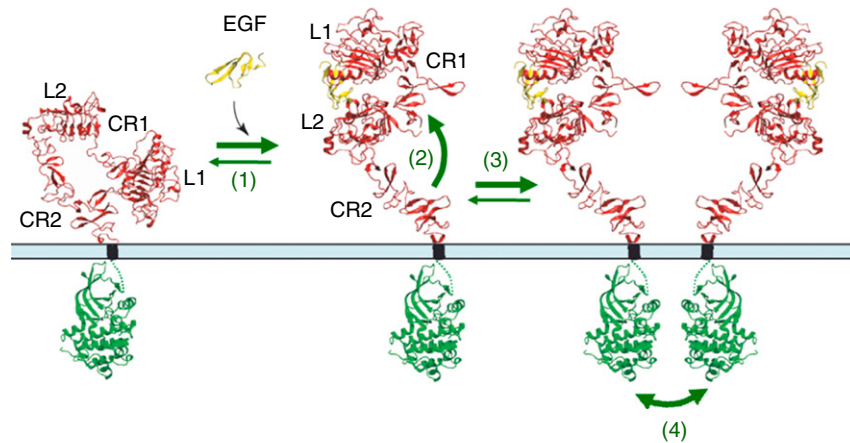


Figure 1 EGF receptor activation involves both conformational and oligomerization events. EGF binding induces a conformational change of the receptor (1), which exposes a dimerization loop (CR1) (2) required for receptor oligomerization (3). Full activation requires additional conformational changes within the kinase domains following dimerization (4). Adapted with permission from Ward, C.W., Lawrence, M.C., Streltsov, V.A., *et al.*, 2007. The insulin and EGF receptor structures: New insights into ligand-induced receptor activation. Trends in Biochemical Sciences 32, 129–137. © 2007 Elsevier.

kinases (tyrosine kinase-linked receptors). The TCR belongs to the latter category. Numerous studies explored the signal transduction properties of these receptors and revealed multiple signaling mechanisms.

Activation by dimerization is the most common signal transduction mechanism. Dimerization usually puts into close proximity the kinase domains within the cytoplasmic tails. *Trans*-phosphorylation events then lead to full activation of the kinase activity and additional phosphorylation events nucleate the recruitment of multiprotein signaling complexes. This mechanism applies to platelet derived and fibroblast growth factor receptors, to name two examples. Conformational changes can also participate in signal transduction. For instance, epidermal growth factor (EGF) induces a conformational change in the extracellular domain of its receptor and this change is required for dimerization and activation (Figure 1; Ward *et al.*, 2007). Finally, the insulin receptor exists as an inactive dimer and ligand induced conformational changes alter the proximity of the kinases domains to allow *trans*-phosphorylation. Altogether, these examples illustrate that the signaling code used by cell-surface receptors is complex and can involve both dimerization and conformational events.

Subunit Composition, Stoichiometry, and Valency of the TCR

The TCR is composed of two ligand-binding chains (α and β or γ and δ) and six CD3 chains. The CD3 subunits form three dimers ($\delta\epsilon$, $\gamma\epsilon$, and $\zeta\zeta$) and are required for cell-surface expression and intracellular signaling (Malissen *et al.*, 1999). The TCR complex is thus characterized by spatially separated ligand-binding and signal transduction subunits. The assembly of the receptor relies on interactions between the extracellular immunoglobulin domains of the CD3 subunits and $\alpha\beta$ TCR chains, and the presence of nine conserved residues within the transmembrane domains (Call and Wucherpfennig, 2007; Figure 2). None of the TCR receptor chains have intrinsic kinase activity. However, SRC-family kinases are associated with CD8 and CD4 TCR co-receptors. These co-receptors bind

to non-polymorphic epitopes on MHC class I and class II molecules, respectively, upon recognition of activating ligands by the TCR. Co-receptors thus couple the TCR to SRC kinases by co-localizing them in the vicinity of the TCR chains.

Despite the absence of enzymatic activity, CD3 chains all contain immunoreceptor tyrosine-based activation motif(s) (ITAMs), which are also present in other immunoreceptors (Figure 3). With 2 tyrosines per ITAM and 10 ITAM motifs per TCR complex, there are a total of 20 tyrosine residues susceptible to become phosphorylated upon TCR triggering. The combinatorial mutation of the ITAMs of the TCR demonstrated their crucial importance for T cell development and functions (Holst *et al.*, 2008).

The stoichiometry of the receptor is a crucial parameter as it determines the number of ligand-binding subunits per receptor complex, i.e., the receptor valency. The stoichiometry of the TCR was generally thought to be one $\alpha\beta$ TCR heterodimer and three CD3 dimers: $\zeta\zeta$, $\delta\epsilon$, and $\gamma\epsilon$ (Call and Wucherpfennig, 2005). This monovalent model largely relies on the biochemical analysis of TCR complexes neosynthesized *in vitro* on microsomes (Call *et al.*, 2004), but also on multiple other experimental approaches (reviewed in Schamel and Alarcon, 2013). Other studies proposed bivalent or multivalent models for the TCR complex, based on the measurement of sedimentation coefficients, native gel electrophoresis, immunological analysis of detergent solubilized receptors, and the observation of higher order TCR oligomers by electron microscopy (reviewed in Schamel and Alarcon, 2013). An even higher level of organization has been suggested by analysis of 'protein islands' containing hundreds of TCR in domains ranging from 30 nm to 300 nm – preexisting prior to activation events. However, analysis of TCR diffusion in quiescent T cells suggests that each TCR moves independently. It has been suggested that some of the discrepancies may arise from changes in TCR organization at different stages of T cell differentiation, for example, that the TCR complex is differently organized in T cells that have only recognized self-antigens (naïve) and those that have been previously activated by a

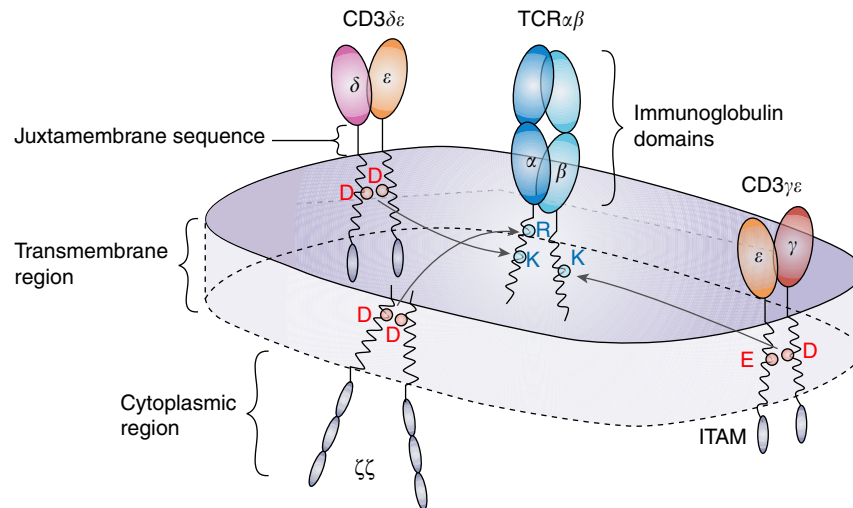


Figure 2 Structure of the $\alpha\beta$ TCR. TCR α and β chains only contain five amino acids in their cytoplasmic domains. Intracellular signals are transduced by $\alpha\beta$ TCR-associated CD3 subunits. Their interaction with α and β chains involves acidic (red) and basic residues (blue) in the transmembrane domains of CD3 and $\alpha\beta$ TCR chains. Adapted with permission from Call, M.E., Wucherpfennig, K.W., 2007. Common themes in the assembly and architecture of activating immune receptors. *Nature Reviews Immunology* 7, 841–850. © 2007 Nature Publishing Group.

CD3 ϵ	GQNKERPPPPVNPDPY	E	P	I	R	K	G	Q	R	L	-	Y	S	G	L			
CD3 δ	GAAEVQALLKNEQL	Y	Q	P	L	R	D	R	E	T	Q	-	Y	S	R	L		
CD3 γ	RASDKQTLLQNEQL	Y	Q	P	L	K	D	R	E	T	Q	-	Y	S	H	L		
CD3 ζ (a)	SAETAANLQDPNQ	L	Y	N	E	L	N	L	G	R	R	E	-	Y	D	V	L	
CD3 ζ (b)	GGKQARRRNPQEG	V	N	A	L	Q	K	D	K	M	A	E	A	Y	S	E	I	
CD3 ζ (c)	TKGERRRGKGHD	G	L	Y	Q	G	L	S	T	A	T	K	T	-	Y	D	A	L
Ig α	FGVDMPPDYEDEN	L	Y	E	G	L	N	L	D	D	C	S	M	-	Y	E	D	I
Ig β	DKDDGKAGMEEDHT	Y	E	G	L	N	I	D	Q	T	A	T	-	Y	E	D	I	
Fc ϵ R γ	RKAAIASREKADAV	T	G	L	N	T	R	S	Q	E	T	-	Y	E	T	L	-	-
Fc ϵ R β	GQELESKKVPDDR	L	Y	E	E	L	N	V	Y	S	P	I	-	-	Y	S	E	L
DAP12	EGTRKQHIAETES	P	Y	Q	E	L	G	Q	R	P	E	V	-	Y	S	D	L	-
								</										

D/Ex₀₋₂ Yxx L/I x₆₋₈ Y xx L/I

Figure 3 ITAM motifs in multiple-chain immunoreceptors. Alignment of cytoplasmic domains from multiple ITAM-containing immunoreceptors. Conserved and partially conserved amino acids are highlighted in yellow and green, respectively. The ITAM consensus sequence is indicated below the alignment.

higher affinity foreign antigen (memory). A general solution to the long-standing question of the stoichiometry of the TCR remains to be identified.

Properties of TCR Ligands

The ability of $\alpha\beta$ TCRs to recognize a diverse universe of pathogens is based on peptides non-covalently associated with class I and class II MHC molecules, respectively recognized by CD8⁺ and CD4⁺ T cells. There are multiple genes encoding class I and class II MHC proteins and each of these genes has a large number of alleles across a population. As most of the allelic diversity is focused on the pocket binding the peptides (Bjorkman *et al.*, 1987), the combination of MHC molecules restricts the repertoire of peptides that bind to each individual's

MHC molecules and thus largely defines the immunological self and the ability to respond to pathogens.

In the event of an infection, peptides generated by proteolysis of the proteins of the pathogen are loaded onto MHC molecules in a similar manner to endogenous peptides. The number of MHC molecules forming complexes with foreign peptides during an infection is estimated to range from 10² to 10⁴, thus only representing 0.1–10% of the total number of MHC molecules expressed on the cell surface. Therefore, the first property of activating TCR ligands is their low representation on the cell surface, in comparison with endogenous p-MHC ligands.

The second property of these ligands is their structural complexity. Unlike most conventional receptors, TCRs recognize ligands that they have not previously encountered. In most cases, only a few side chains of the peptide are exposed to the solvent above the peptide binding groove defined by the two MHC helices, whereas about half the side chains serve as anchors that bind the peptide to the MHC (Figure 4; Stern and Wiley, 1994). The distinction between self- and foreign-p-MHC complexes thus relies on a few specific bonds with exposed side chains of the peptide plus small changes in the positioning of the peptide backbone due to differences in anchoring interactions.

A third property of TCR ligands is their low affinity for their cognate receptor, with K_D ranging from 1 μ M to effectively unmeasurable for some self-p-MHC (> 1 mM) (Davis *et al.*, 1998). This affinity range is much lower than the affinity range for many immunoglobulins and their ligands, which undergo a process of 'somatic hypermutation' to achieve K_D less than 1 nM. TCR undergo no similar mutation process to achieve higher affinity and engineering of high affinity TCR:p-MHC for tumor immunotherapy has yielded mixed results. TCR:p-MHC interactions have an affinity range that is typical for adhesion receptors, where confinement of interactions in the interface leads to efficient binding in this affinity range. Thus, the observed physiological TCR:p-MHC affinity range is optimal, but

there are still many questions about how they operate in natural immune responses.

Finally, TCR ligands also have this intriguing property of inducing very pleiotropic responses in T cells, ranging from survival to apoptosis, partial or full activation, and acquisition of various effector functions (cytotoxicity, cytokine production, proliferation, and differentiation). Although these decisions are influenced by the context of the activation event, ligand properties also impact on the kinetics, magnitude, and qualitative nature of the response (reviewed in [Corse et al., 2011](#)). As an example, the polarization of CD4⁺ T cells toward a Th1 (characterized by interferon- γ , interleukin (IL)-2, tumor necrosis factor production, and cell-mediated immunity) versus Th2 effector phenotype (characterized by IL-4, IL-5, IL-10, and IL-13 production and strong antibody responses) is

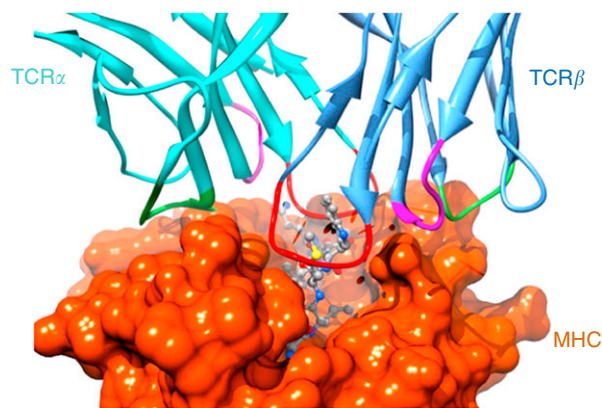


Figure 4 3D-structure of a TCR:p-MHC complex. The recognition by the TCR (blue) involves interactions with the MHC molecule (orange) and with the few exposed side chains of the peptide (displayed in ball and stick). Adapted from Zoete, V., Irving, M., Ferber, M., Cuendet, M. A., Michelin, O., 2013. Structure-based, rational design of T cell receptors. *Frontiers in Immunology* 4, 268.

strongly influenced by the strength of TCR:p-MHC interactions ([Van Panhuys et al., 2014](#)).

Altogether, the unique properties of activating TCR ligands raise the following question: what is the signaling code used by the TCR signalosome to convert and propagate the information associated with ligand recognition?

The Signaling Code Used by the TCR Signalosome to Convert Extracellular Signals into Biological Responses

The Kinetic Model of T Cell Activation

Thermodynamic and kinetic parameters of TCR:p-MHC interactions are known for several pairs of TCR:p-MHC complexes ([Figure 5](#); [Kalergis et al., 2001](#); [Davis et al., 1998](#)).

Combined with the observation that TCR cross-linking via anti-TCR antibodies induces T cell activation ([Kaye et al., 1983](#)), these data point to a kinetic model of T cell activation in which the K_D of the TCR:p-MHC interaction, or more accurately the half-life of the complex ($t_{1/2}$), dictates the degree of TCR clustering with its co-receptors and associated kinases, and thus T cell activation ([Choudhuri and Van Der Merwe, 2007](#); [Figure 6](#)). Note that TCR clustering could involve either the aggregation of monovalent TCR complexes and/or multivalent preformed TCR nanoclusters into larger microclusters.

Supporting this model, quasi-elastic light scattering measurements of the hydrodynamic radius of TCR:p-MHC complexes in solution showed that recognition of the cognate p-MHC ligand by the TCR leads to oligomerization in solution, an event which was strictly dependent on the presence of a specific activating peptide ([Reich et al., 1997](#)). Accordingly, the kinetics of the TCR:p-MHC interaction measured by surface plasmon resonance are compatible with a dimerization model ([Alam et al., 1999](#)). Finally, dynamic imaging of receptor complexes during T cell activation also confirmed the

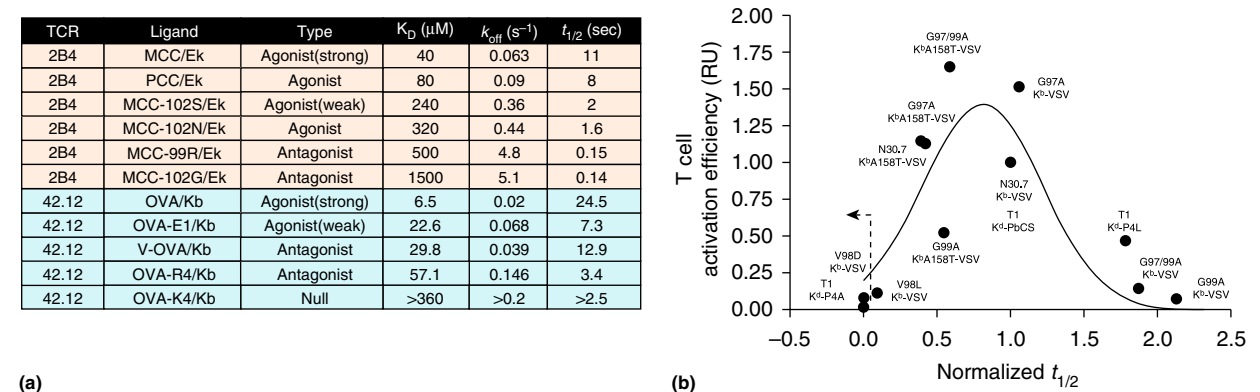


Figure 5 Thermodynamic and kinetic parameters of TCR:p-MHC interactions. (a) Thermodynamic and kinetic dissociation constants for two TCRs and their ligands (reviewed in [Davis et al., 1998](#)). Ligands are categorized according to their functional impact: strong and weak agonists induce full and partial T cell responses, respectively, whereas antagonists inhibit agonist responses. The substitution of one single residue in the peptide (e.g., MCC 102S/E^k) converts the agonist into a partial agonist with biological responses attenuated by a 20 000-fold factor, while the $t_{1/2}$ of the interaction decreases from 11 to 2 s, i.e., only by a 5-fold factor. (b) Correlation between TCR:p-MHC half-lives and T cell activation. Each dot represents a TCR:p-MHC complex. Adapted with permission from Kalergis, A.M., Boucheron, N., Doucey, M.A., et al., 2001. Efficient T cell activation requires an optimal dwell-time of interaction between the TCR and the pMHC complex. *Nature Immunology* 2, 229–234. © 2001 Elsevier.

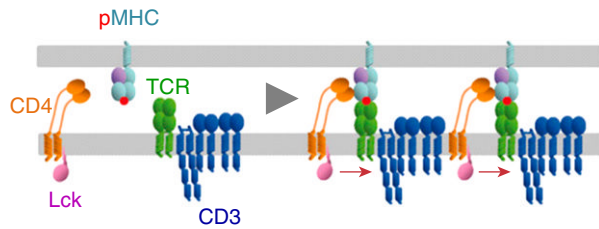


Figure 6 Kinetic model of T cell activation. In this model, TCR aggregation is dictated by thermodynamic and kinetic properties of the TCR:p-MHC interaction. Longer half-lives favor the recruitment of co-receptor associated kinases and ITAM phosphorylation. Adapted with permission from Choudhuri, K., Van Der Merwe, P.A., 2007. Molecular mechanisms involved in T cell receptor triggering. *Seminars in Immunology* 19, 255–261. © 2007 Elsevier.

formation of supramolecular activation clusters or SMACs (Campi *et al.*, 2005; Varma *et al.*, 2006).

In this kinetic model, the signaling code is thus based on the kinetic parameters of the TCR:p-MHC interaction, which translate into various degrees of TCR and co-receptor clustering, and thus various levels of SRC kinase activation and downstream ITAM phosphorylation. This model explains well why the distance between TCR ligands, but not their orientation, impacts on T cell activation (Cochran *et al.*, 2001; Delcassian *et al.*, 2013).

However, this model does not account for the sensitivity of T cells since 90–99.9% of p-MHC complexes are endogenous and not activating p-MHC, and the probability of TCR oligomerization is thus very low. As a potential explanation to this paradox, the oligomerization property of TCR:p-MHC complexes observed in solution could favor TCR clustering. Another explanation resides in the observation that activating dimerization events can involve TCRs that do not recognize the activating p-MHC. In this pseudo-dimer model, the low affinity of the interaction between the co-receptor and MHC allows for its dissociation from the activating p-MHC and rebinding to a neighboring endogenous p-MHC engaged with another TCR. These pseudo-dimers of activating/endogenous p-MHC complexes induce activating events, whereas monomers or dimers of endogenous p-MHC do not (Krogsgaard *et al.*, 2005). Finally, clustering events take place in the 2D-plane of the cellular membrane and more particularly in the delimited area corresponding to the zone of contact between two cells or immunological synapse (Dustin and Depoil, 2011). This synapse locally concentrates TCR:p-MHC complexes and thus increases the probability of TCR clustering events. Altogether, these various mechanisms explain why a low number or even a single activating TCR:p-MHC interaction can induce TCR oligomerization and drive T cell activation.

The fact that the TCR:p-MHC interaction takes place in a dynamic interface between two cells also provides an opportunity to use physical forces to sharpen ligand discrimination (Klotzsch and Schutz, 2013). A critical parameter in this context is the response of the kinetic off-rate to applied forces. Recently, Liu *et al.* (2014) have found that agonist TCR:p-MHC form 'catch-bonds,' which become stronger when a force is applied. The structural basis of this behavior is not known, but this could explain situations in which solution measurements

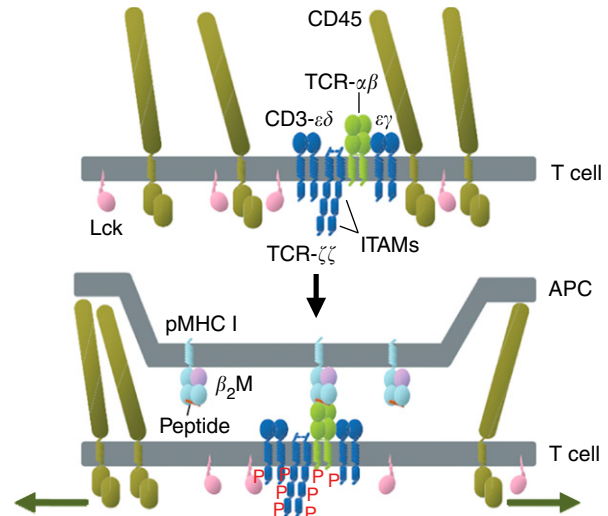


Figure 7 The kinetic-segregation model of T cell activation. The formation of TCR:p-MHC complexes requires the apposition of the two cell membranes at a distance of 15 nm. This binding event induces the spatial segregation of proteins with large extracellular domains, like CD45, from the synaptic contact. This local exclusion tips the phosphatase-kinase balance in favor of kinases, and triggers ITAM phosphorylation. Adapted with permission from Choudhuri, K., Wiseman, D., Brown, M.H., *et al.*, 2005. T-cell receptor triggering is critically dependent on the dimensions of its peptide-MHC ligand. *Nature* 436, 578–582. © 2005 Nature Publishing Group.

of TCR:p-MHC interactions failed to predict functional ranking of ligand potency.

How to Account for Specificity and Sensitivity of TCR Responses?

The TCR signaling code has to account for the specificity and sensitivity of T cell responses. The kinetic model of T cell activation does not fully account for these two properties since the half-lives of the TCRs for their full agonist, partial agonist, and antagonist p-MHC ligands are relatively similar (Figure 5). Small differences in half-lives yet translate into very different biological outcomes and raise the question as to how T cells discriminate between such similar ligands.

The specificity of T cell responses can be explained by the amplification of small kinetic differences by the TCR signalingome. This amplification results from the existence of multiple sequential biochemical steps in the signaling cascade. One example is the exclusion of molecules with large ectodomains like CD45 phosphatases from the synaptic contact between cells. If the interaction is not long enough, the exclusion of phosphatases is not efficient, which leads to the rapid reversion of proximal biochemical events to the basal state via dephosphorylation. This mechanism describes a variation of the kinetic model of T cell activation, called kinetic-segregation model (Figure 7) and illustrates the multiple cycles of molecular association–dissociation, phosphorylation–dephosphorylation which occur at each step of the TCR signaling cascade. These cycles are coupled to energy-consuming reactions and associated with a waiting time before the completion of these biochemical reactions leads to a productive

signal. By increasing the discrimination potential of T cells for thermodynamically similar p-MHC ligands, these linear amplification steps constitute the basis of the kinetic-proof-reading model of T cell activation (Mckeithan, 1995).

In addition to this linear amplification mechanism, the feedback control model of T cell activation proposes that the $t_{1/2}$ of the TCR:p-MHC interaction also dictates the initiation of nonlinear signaling pathways (Chan *et al.*, 2004). Accordingly, this nonlinear property of TCR signaling networks which magnifies small initial differences during signal propagation was recently observed using mass cytometry, a high-dimensional technique that can probe multiple signaling nodes at the single-cell level (Mingueneau *et al.*, 2014). One example of such nonlinear mechanisms is the negative regulatory loop exerted by the phosphatase SH2-containing phosphatase 1 (SHP-1) on the activity of the SRC kinase lymphocyte-specific protein tyrosine kinase (LCK) (Altan-Bonnet and Germain, 2005). This regulation increases the cell activation threshold and reduces nonspecific LCK activation events. In contrast, only activating ligands induce the phosphorylation of LCK on Ser59 by extracellular signal regulated kinases (ERK), an event which prevents its association with SHP-1. The deactivation of this negative regulatory loop positively reinforces biochemical activities along the whole TCR signaling cascade. Therefore, these nonlinear signaling events increase the signal-to-noise ratio within the TCR signaling network and enhance the ligand discrimination potential of T cells.

If the above models account for the specificity of T cell responses, they fail to explain the sensitivity of T cells to low-abundance ligands. By introducing temporal delays within the signaling cascade, these models actually limit ligand screening capabilities, and thus T cell sensitivity. The serial triggering model of T cell activation accounts for this property by taking into account the relative $t_{1/2}$ of the ternary p-MHC:TCR complexes – in the range of a few seconds (Figure 5) – relatively to the $t_{1/2}$ of p-MHC ligands – in the range of hours (Valitutti *et al.*, 1995). Based on this difference in half-lives, this model proposes that a single activating p-MHC serially engages multiple TCRs. This model explains well why two to three p-MHC complexes are sufficient to trigger T cell activation (Boniface *et al.*, 1998; Cochran *et al.*, 2000), and why even one single p-MHC can induce ZAP70 recruitment and cytokine secretion by promoting TCR clustering (Huang *et al.*, 2013).

Altogether, the kinetic-proofreading, feedback control, and serial triggering models of T cell activation illustrate different mechanisms at play within the TCR signalosome and help explain the existence of an optimal dwell time for the TCR:p-MHC interaction (Figure 5). This optimal dwell time likely reflects a compromise between: (1) a $t_{1/2}$ which is long enough to trigger the linear and nonlinear biochemical events ensuring specific responses; and (2) a dissociation constant which is fast enough to allow for serial triggering of multiple receptors, and thus ensuring high sensitivity of T cell responses.

The Conformational Model of T Cell Activation

Even though kinetic models of T cell activation account for the biological properties of a large number of TCR:p-MHC complexes, there are some thermodynamic exceptions for which

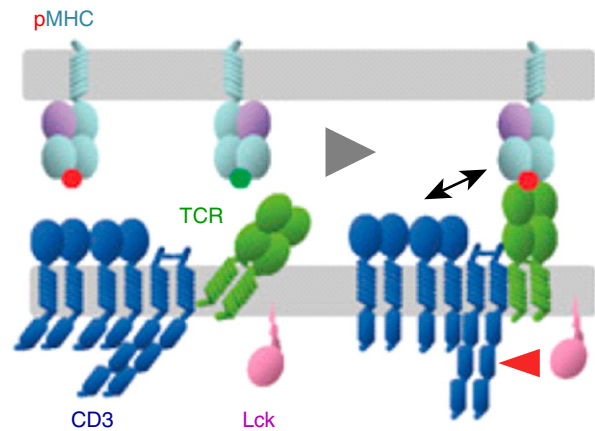


Figure 8 The conformational model of T cell activation. In this model, p-MHC binding triggers a conformational change of the TCR (black arrow). This structural modification would lead to an increased accessibility of the ITAMs to kinases (red arrow). Adapted with permission from Choudhuri, K., Van Der Merwe, P.A., 2007. Molecular mechanisms involved in T cell receptor triggering. *Seminars in Immunology* 19, 255–261. © 2007 Elsevier.

no correlation is observed between the K_D of the interaction and the biological outcome (Yoon *et al.*, 1994; Al-Ramadi *et al.*, 1995). Based on these observations, Janeway (1995) proposed a conformational model of T cell activation almost 20 years ago (Figure 8). Since then, different variations of this model have been proposed, involving rotation, piston-like movements, or deformation mechanisms driven by the forces associated with the binding or detachment of the p-MHC (reviewed in van der Merwe and Dushek, 2011).

3D structures of $\alpha\beta$ TCRs in complex or not with their cognate p-MHC ligands did not show evidence of conformational changes upon ligand binding, except at the TCR:p-MHC interface (Rudolph *et al.*, 2006). One single study identified a conformational change in a domain of the TCR α chain which was later proposed to play a role in TCR oligomerization (Beddoe *et al.*, 2009; Kuhns *et al.*, 2010). These results open the possibility that both conformational changes and receptor clustering participate to TCR signal transduction, similarly to the EGF receptor (Figure 1). However, this model remains to be validated. Conformational change models may also incorporate physical force as one of the actuators of conformational change (Kim *et al.*, 2009).

Additional arguments have been proposed in support of the conformational model of T cell activation. In particular, Aivazian and Stern (2000) proposed that CD3 ζ cytoplasmic domain adopts a folded structure in presence of acidic phospholipids which is refractory to phosphorylation. Building on this observation, other groups proposed that the ITAM motifs are sequestered within the lipid bilayer in steady-state and that TCR triggering makes them accessible to tyrosine phosphorylation (Risueno *et al.*, 2008b; Xu *et al.*, 2008; Figure 9). Yet, these results were later challenged (Fernandes *et al.*, 2010) by the result that mutation of the acidic lipid-binding, basic sequences in CD3 had no effect on proximal TCR signaling (Deford-Watts *et al.*, 2009, 2011). Similarly, a proline-rich sequence within the cytoplasmic tail of CD3 ϵ was initially proposed to be inducibly exposed upon TCR

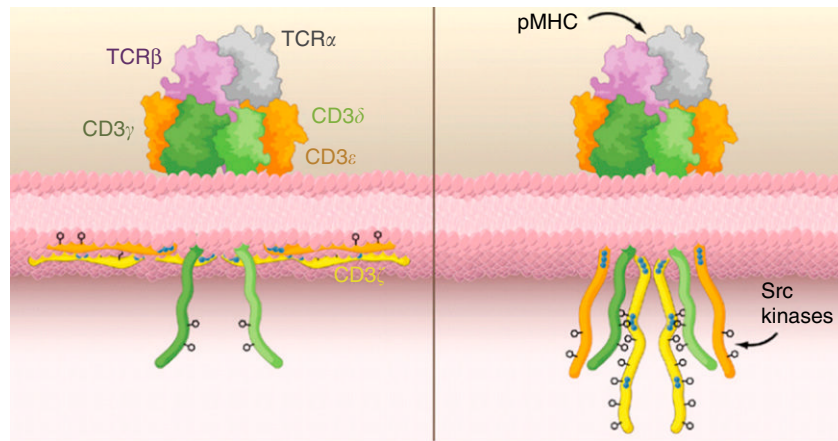


Figure 9 Sequestration of CD3 ϵ and CD3 ζ ITAMs within the plasma membrane and inducible exposure upon TCR triggering. In this model, basic residues in CD3 ϵ and CD3 ζ (blue dots) interact with the acidic inner leaflet of the plasma membrane, while the phenyl rings of the ITAMs are buried in the bilayer. Ligand binding exposes CD3 cytoplasmic domains. Adapted with permission from Kuhns, M.S., Davis, M.M., 2008. The safety on the TCR trigger. Cell 135, 594–596. © 2008 Elsevier.

triggering, prior to ITAM phosphorylation (Gil *et al.*, 2002). This observation led to a series of follow-up reports (reviewed in Risueno *et al.* (2008a)), but the targeted mutation of this sequence in mice showed that it is largely dispensable for T cell activation (Mingueneau *et al.*, 2008; Borroto *et al.*, 2014).

In conclusion, even if it is not excluded that conformational changes participate to TCR signaling, available evidence for such mechanism remains so far elusive. Recently, the artificial reconstitution of the TCR signalosome into a non-immune cell type showed that TCR triggering could be recapitulated using various engineered receptor–ligand pairs in place of the TCR:p-MHC complex, thus suggesting that conformational changes might not constitute the fundamental mechanism of T cell activation (James and Vale, 2012).

Biochemical Modules and Assembly of the TCR Signalosome

The TCR signalosome can be described by a series of biochemical modules, which initiate the biochemical cascade, amplify, and diversify the initial signal into a complex, multifunctional assembly of proteins with diverse enzymatic and adaptor activities (Figure 10).

Signal Initiation Module

The first detectable biochemical event following p-MHC binding by the TCR is the activation of co-receptor associated tyrosine kinases of the SRC family. In T cells, LCK is the major SRC kinase positively regulating T cell activation and its regulation relies on the existence of two kinase conformations (reviewed in Boggon and Eck, 2004; Salmond *et al.*, 2009; Figure 11). The open conformation of the kinase is the major form of LCK found in T cells in steady-state conditions. Upon TCR triggering, CD4 or CD8 co-receptors interact with the MHC molecule bound by the TCR. Since the affinity of the co-receptor for the MHC is significantly lower than the affinity of

the TCR for the p-MHC ($K_D > 200\mu\text{M}$), the role of the co-receptor is mainly to co-localize LCK in the proximity of the ITAMs and, to a much lesser extent, to stabilize the TCR:p-MHC interaction. Once fully activated by *trans*-phosphorylation of the Y394, LCK dissociates from CD4 and catalyzes the transfer of phosphate groups from adenosine triphosphate (ATP) to the hydroxyl residues of CD3 ITAM tyrosines.

SRC kinase activation leads to full or partial phosphorylation of the 10 ITAMs from the CD3 complex. The main downstream signaling node is the SYK-family kinase ZAP70. Unlike SRC kinases, the primary event driving ZAP70 activation is not the dephosphorylation of an inhibitory tyrosine, but a conformational change occurring upon binding of ZAP70 SH2 domains onto a doubly phosphorylated ITAM motif (reviewed in Wang *et al.*, 2010; Figure 12). The resulting flexibility in the kinase domain facilitates the phosphorylation of several tyrosine residues, which leads to full activation.

Signal Amplification and Diversification Module

The main substrate of ZAP70 is the transmembrane adaptor LAT (Linker for Activation of T cells) (reviewed in Roncagalli *et al.*, 2010; Malissen *et al.*, 2014). The key role of LAT in the TCR signalosome was clearly demonstrated in *Lat*-deficient Jurkat cells, which show almost completely abrogated signaling events downstream of the TCR (Finco *et al.*, 1998), and in *Lat*-deficient mice (Sommers *et al.*, 2005), in which T cell development is arrested at an early stage in the thymus. LAT counts nine tyrosine residues in its cytoplasmic tail. Knock-in mice in which the last four C-terminal tyrosines are mutated to phenylalanine display a phenotype that recapitulates the phenotype seen in *Lat*-knockout mice (Sommers *et al.*, 2001). Therefore, most LAT signaling functions are funneled through phosphorylation of these four tyrosine residues.

Once phosphorylated, LAT recruits cytosolic adaptors including SH2 domain-containing leukocyte phosphoprotein of 76 kDa (SLP76), growth factor receptor-bound protein 2 (GRB2), GRB2-related adaptor protein (GRAP2), as well as enzymes like PhosphoLipase C (PLC γ 1) and IL-2 inducible T

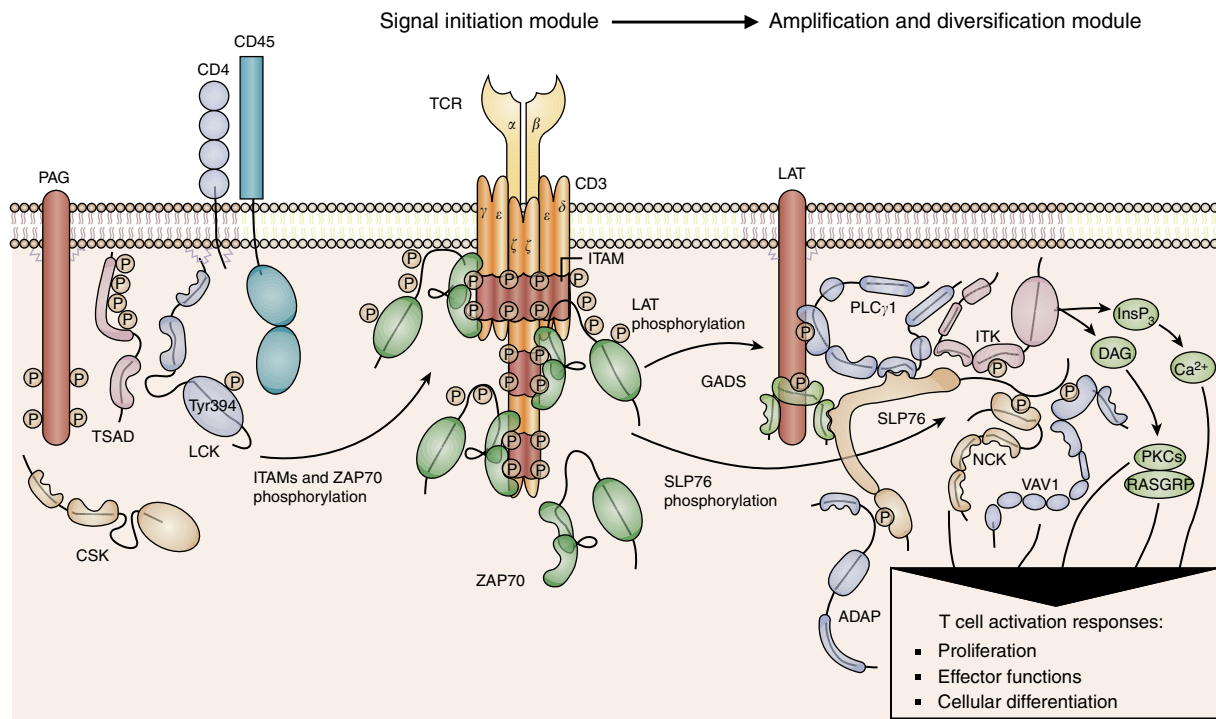


Figure 10 Modular organization of the TCR signalosome. The signal initiation module converts the kinetic properties of the p-MHC ligand into an intracellular biochemical signal, i.e., various levels of CD3 ITAM phosphorylation. This module involves the SRC kinase LCK and the SYK-family kinase ZAP70. The amplification and diversification module recruits multifunctional protein complexes nucleated around the adaptor molecules LAT and SLP76. See text for details. Adapted with permission from Acuto, O., Bartolo, V.D., Michel, F., *et al.*, 2008. Tailoring T-cell receptor signals by proximal negative feedback mechanisms. *Nature Reviews Immunology* 8, 699–712. © 2008 Nature Publishing Group.

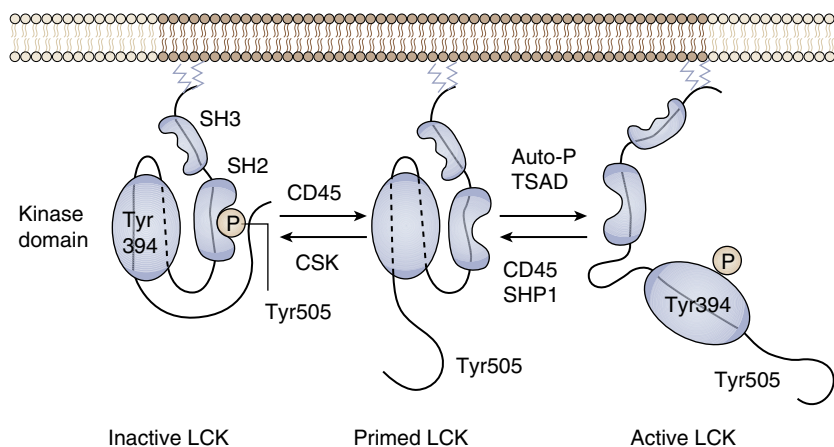


Figure 11 Mechanistic of LCK activation. In its closed conformation, the C-terminal Y505 residue is phosphorylated by CSK and forms an intramolecular interaction with the SH2 domain. The dephosphorylation of this residue by CD45 leads to an open conformation. In T cells, 80% of LCK kinases are in this primed conformation, likely due to the abundance of CD45 on the cell surface. Full activation of the kinase requires the additional *trans*-phosphorylation of the activation loop (Y394). Adapted with permission from Acuto, O., Bartolo, V.D., Michel, F., *et al.*, 2008. Tailoring T-cell receptor signals by proximal negative feedback mechanisms. *Nature Reviews Immunology* 8, 699–712. © 2008 Nature Publishing Group.

cell kinase (ITK). The formation of this multifunctional complex is a highly cooperative process involving protein–protein interactions, phosphorylation events, and other enzymatic activities (Figure 13).

Affinity purification coupled to mass spectrometry recently identified 90 interacting partners in the ZAP70-LAT-SLP76 signalosome, most of which were previously unknown (Roncagalli *et al.*, 2014), thus redefining the breadth of the

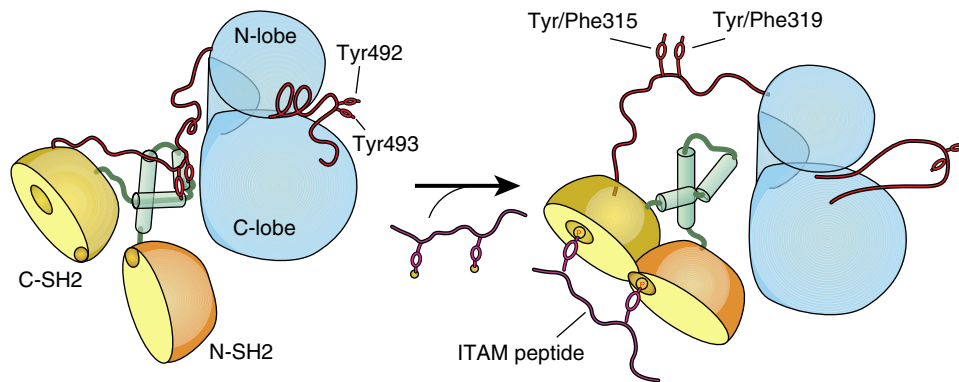


Figure 12 ZAP70 activation requires conformational changes and phosphorylation events. Prior to ITAM-binding, ZAP-70 is in an inactive conformation due to hydrophobic interactions between the SH2-kinase inter-domain, the kinase domain, and the SH2-SH2 inter-domain. Binding to a doubly phosphorylated ITAM induces a conformational change, which releases the tandem-SH2 domains from the kinase domain and allows for LCK-mediated phosphorylation of Y315 and Y319, as well as Y492 and Y493 residues in the activation loop. These phosphorylation events lead to maximal activation of the kinase. Adapted with permission from Deindl, S., Kadlecsek, T.A., Brdicka, T., *et al.*, 2007. Structural basis for the inhibition of tyrosine kinase activity of ZAP-70. *Cell* 129, 735–746. © 2007 Elsevier.

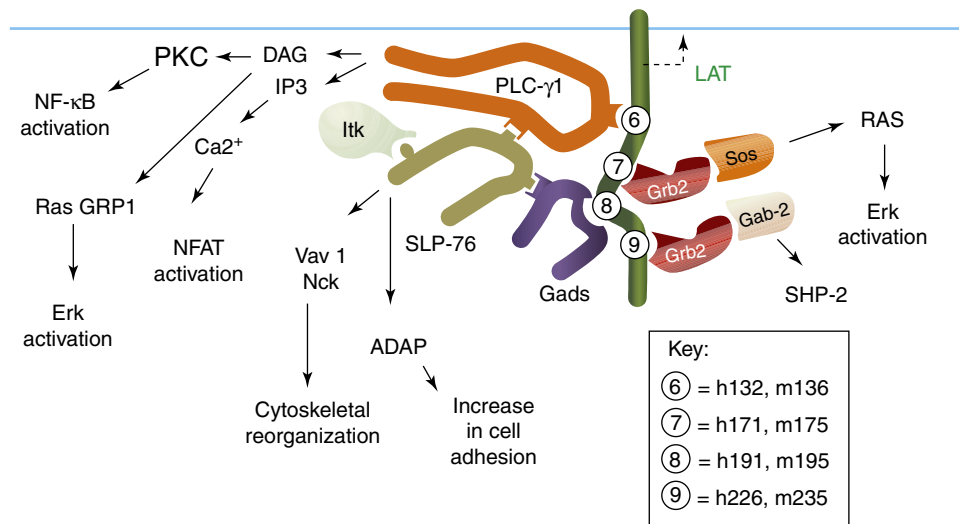


Figure 13 The transmembrane adaptor LAT is a key molecular platform nucleating the assembly of a multifunctional protein complex. The four more distal tyrosines of LAT are numbered. Phosphorylated Y#6 recruits PLC γ 1, whereas GRAP2/GRB2-related adaptor downstream of SHC (GADS) binds LAT on phosphorylated Y#7 and Y#8. The cytosolic adaptor SLP76 binds LAT via its association to GRAP2/GADS and PLC γ 1. SLP76 phosphorylation by ZAP70 leads to the recruitment of the TEC-family tyrosine kinase ITK, which can then phosphorylate PLC γ 1 and enhance its activity. The resulting increase in inositol triphosphate (IP3) and di-acyl glycerol (DAG) triggers calcium-dependent signaling pathways, including nuclear factor of activated T cells (NFAT) nuclear translocation, and protein kinase C-nuclear factor kappa B (PKC-NF κ B)/mitogen-activated protein kinase (MAPK) signaling pathways. Phosphorylated SLP76 also recruits VAV, non-catalytic region of tyrosine kinase adapter protein (NCK), Wiskott–Aldrich syndrome protein (WASP), and adhesion and degranulation promoting adaptor (ADAP) which participate to the remodeling of the actin cytoskeleton and cellular adhesion. The adaptor GRB2 can bind several tyrosines of LAT and recruits the guanine exchange factor 'son of sevenless' (SOS) to promote MAPK signaling. The p85A regulatory subunit of the phosphatidylinositol 3-kinase (PI3K) is also recruited by LAT and leads to the formation of the second messenger phosphatidyl 3,4,5-triphosphate (PIP3), which recruits molecules with pleckstrin homology (PH) domains, among which are ITK and protein kinase B (PKB)/AKT. Adapted with permission from Roncagalli, R., Mingueau, M., Gregoire, C., *et al.*, 2010. LAT signaling pathology: An 'autoimmune' condition without T cell self-reactivity. *Trends in Immunology* 31, 253–259. © 2010 Elsevier.

unknown within the TCR signalosome. Also adding to this complexity is the recent observation that LAT adaptors can form protein islands on the cell surface, or, in an alternative model, accumulate within sub-synaptic compartments where

they would participate to TCR signaling (reviewed in [Dustin and Depoil, 2011](#)). The functional importance of these sub-cellular compartments during TCR triggering remains to be understood.

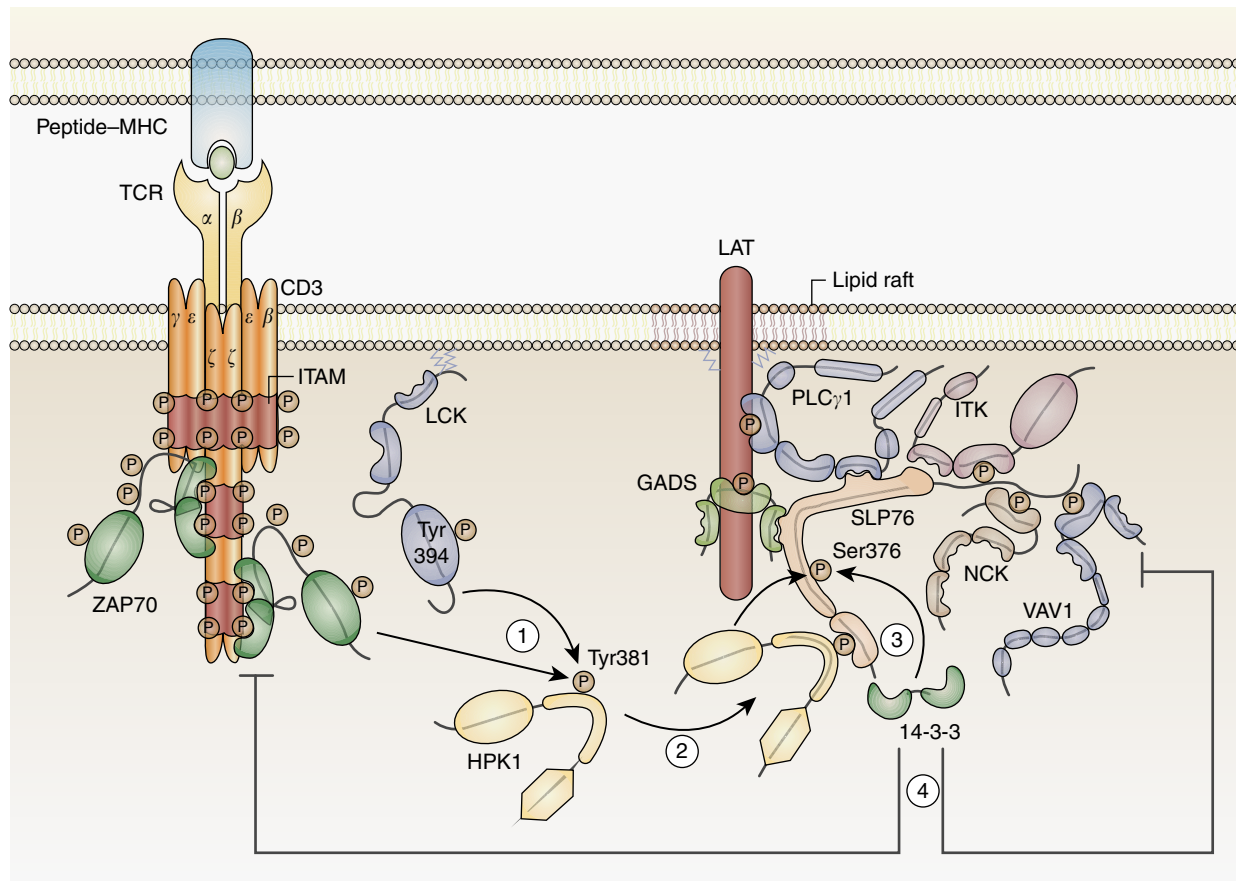


Figure 14 Negative feedback loop operated by the kinase HPK1. HPK1 is phosphorylated upon TCR activation (1), and inducibly recruited on the LAT signalosome by SLP76 (2). HPK1 then phosphorylates SLP76 on the S376 residue which recruits 14-3-3-family molecules (3), leading to the negative regulation of the signalosome (4). S376 phosphorylation peaks between 10 and 15 min after the initiation of TCR signaling. This regulation is thus relatively delayed compared with more proximal feedback loops mediated by SHP1, CD45, or CSK. Adapted with permission from Acuto, O., Bartolo, V.D., Michel, F., *et al.*, 2008. Tailoring T-cell receptor signals by proximal negative feedback mechanisms. *Nature Reviews Immunology* 8, 699–712. © 2008 Nature Publishing Group.

Negative Regulation Modules

In addition to positively regulating signaling, the TCR signalosome also turns on negative regulation loops to keep in check cellular responses (reviewed in Acuto *et al.*, 2008). The existence of such negative regulation modules is clearly demonstrated by the genetic ablation of LAT in peripheral T cells, which leads to hyperactivated and hyperproliferative T cells (Mingueneau *et al.*, 2009; Shen *et al.*, 2010). The TCR-LAT signaling axis is thus more than a positive regulator of TCR signaling.

These negative loops act at different levels of the TCR signaling cascade. Some feedback loops involve kinases, like the Ser/Thr kinase hematopoietic progenitor kinase 1 (HPK1). In its absence, an increased phosphorylation of zeta-associated protein 70 (ZAP-70), SLP-76, PLCγ1, LAT, VAV-1, and ERK kinases is observed. The mechanism of action of HPK1 involves 14-3-3 proteins which negatively regulate TCR signaling by ubiquitination and degradation (Wang *et al.*, 2012; Figure 14).

Phosphatases also mediate negative regulatory events. For instance, CD45 is not only involved in the priming of LCK by dephosphorylating its inhibitory Y505 residue, but also in the downregulation of TCR signaling by dephosphorylating the Y384 residue of LCK as well as tyrosines of the ITAMs. The phosphatase SHP1 also plays a key role in the deactivation of proximal signaling pathways (reviewed in Lorenz, 2009), as demonstrated by the phenotype of mice expressing loss-of-function mutations in the corresponding gene. These mice indeed develop severe autoimmune conditions associated to the prolonged activation of SRC kinases. In addition to SHP1, other phosphatases like PEST domain-enriched tyrosine phosphatase (PEP) and protein tyrosine phosphatase-PEST participate in negative feedback loops through their recruitment by C-terminal SRC kinase (CSK). Both phosphatases dephosphorylate Y394 of LCK. Therefore, the cooperative action of CSK on the inhibitory Y505 residue and of CSK-associated phosphatases on the activating Y394 residue likely explains the tight regulation of LCK activity upon TCR triggering.

Multiple other mechanisms participate to the regulation of the TCR signalosome. These include TCR internalization and degradation (reviewed in Balagopalan *et al.*, 2009), c-casitas B lymphoma (CBL) and CBL-B-dependent ubiquitination and degradation of TCR signaling nodes (reviewed in Mohapatra *et al.*, 2013), lipid phosphatases and kinases (reviewed in Huang and Sauer, 2010), as well as regulations mediated by co-inhibitory receptors that often recruit phosphatases SHP-1 or SHP-2 (reviewed in Chen and Flies, 2013). Altogether, these mechanisms tightly control T cell activation responses and avoid chronic or exacerbated T cell responses. Although these control mechanisms are usually effective, there is a significant burden of autoimmune diseases, which impact a few percent of the world population and allergies to harmless environmental antigens that are even more common.

Conclusions

- The antigenic receptor expressed by T cells is a multiple-chain receptor with separate ligand recognition and signal transduction subunits. The TCR lacks intrinsic kinase activity and non-covalently associates with the tyrosine kinase LCK and ZAP-70 upon TCR triggering. LCK initiates recruitment of ZAP-70 and the assembly of a multi-functional protein complex, designated as the TCR signalosome.
- A large body of evidence supports the role of receptor/co-receptor clustering during T cell activation, even though it remains to determine if clustering involves monovalent receptor complexes and/or preassembled TCR nanoclusters. Available evidence for conformational changes remains so far elusive.
- The properties of activating TCR ligands impose unique constraints on the mechanism of signal transduction. The TCR signaling code has indeed to account for both the exquisite sensitivity of T cells to rare activating p-MHC ligands and the specificity of their responses, only triggered by foreign p-MHC complexes, and not their structurally similar endogenous p-MHC counterparts. Kinetic-proof-reading and feedback control mechanisms contribute to the specificity of T cell responses. Serial triggering of multiple TCRs by the same activating p-MHC ligand increases the sensitivity of T cells to rare ligands.
- Biochemical events driving the assembly of the TCR signalosome involve the activation of SRC and SYK-family kinases, as well as multiple other enzymatic activities, transmembrane and cytosolic adaptor proteins. LAT and SLP76 are the key adaptors which nucleate the cooperative assembly of large multiprotein complexes that funnel most TCR signaling activity.
- Despite remarkable progress made in our understanding of this signaling system since the identification of TCR proteins in the 1980s (reviewed in Weiss, 2005), recent results indicate that many of the newly identified components of the TCR signalosome are not characterized. In addition, T cell activation changes the expression of more than 2000 genes within 24 h following TCR triggering (Mingueneau *et al.*, 2013). Therefore, understanding how proximal signaling events triggered by the TCR signalosome are spatially and temporally integrated within the cell to induce nuclear programs and functional responses will constitute one of the next challenges. Answering these questions will allow us to draw a more complete and dynamic map of the signaling code used by T cells, and ultimately help us design better therapeutics to achieve selective modulation of T cell responses.

See also: Organelles: Structure and Function: Extracellular Vesicles

References

- Acuto, O., Bartolo, V.D., Michel, F., 2008. Tailoring T-cell receptor signals by proximal negative feedback mechanisms. *Nature Reviews Immunology* 8, 699–712.
- Aivazian, D., Stern, L.J., 2000. Phosphorylation of T cell receptor zeta is regulated by a lipid dependent folding transition. *Nature Structural & Molecular Biology* 7, 1023–1026.
- Alam, S.M., Davies, G.M., Lin, C.M., *et al.*, 1999. Qualitative and quantitative differences in T cell receptor binding of agonist and antagonist ligands. *Immunity* 10, 227–237.
- Al-Ramadi, B.K., Jelonek, M.T., Boyd, L.F., *et al.*, 1995. Lack of strict correlation of functional sensitization with the apparent affinity of MHC/peptide complexes for the TCR. *Journal of Immunology* 155, 662–673.
- Altan-Bonnet, G., Germain, R.N., 2005. Modeling T cell antigen discrimination based on feedback control of digital ERK responses. *PLoS Biology* 3, e356.
- Balogopalan, L., Barr, V.A., Samelson, L.E., 2009. Endocytic events in TCR signaling: Focus on adaptors in microclusters. *Immunological Reviews* 232, 84–98.
- Beddoe, T., Chen, Z., Clements, C.S., *et al.*, 2009. Antigen ligation triggers a conformational change within the constant domain of the alpha-beta T cell receptor. *Immunity* 30, 777–788.
- Bjorkman, P.J., Saper, M.A., Samraoui, B., *et al.*, 1987. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* 329, 512–518.
- Boggon, T.J., Eck, M.J., 2004. Structure and regulation of Src family kinases. *Oncogene* 23, 7918–7927.
- Boniface, J.J., Rabinowitz, J.D., Wulfig, C., *et al.*, 1998. Initiation of signal transduction through the T cell receptor requires the multivalent engagement of peptide/MHC ligands [corrected]. *Immunity* 9, 459–466.
- Borrotto, A., Arellano, I., Blanco, R., *et al.*, 2014. Relevance of Nck-CD3 epsilon interaction for T cell activation *in vivo*. *Journal of Immunology* 192, 2042–2053.
- Call, M.E., Pyrdol, J., Wucherpfennig, K.W., 2004. Stoichiometry of the T-cell receptor-CD3 complex and key intermediates assembled in the endoplasmic reticulum. *EMBO Journal* 23, 2348–2357.
- Call, M.E., Wucherpfennig, K.W., 2005. The T cell receptor: Critical role of the membrane environment in receptor assembly and function. *Annual Review of Immunology* 23, 101–125.
- Call, M.E., Wucherpfennig, K.W., 2007. Common themes in the assembly and architecture of activating immune receptors. *Nature Reviews Immunology* 7, 841–850.
- Campi, G., Varma, R., Dustin, M.L., 2005. Actin and agonist MHC-peptide complex-dependent T cell receptor microclusters as scaffolds for signaling. *Journal of Experimental Medicine* 202, 1031–1036.
- Chan, C., Stark, J., George, A.J., 2004. Feedback control of T-cell receptor activation. *Proceedings of the Royal Society B: Biological Sciences* 271, 931–939.
- Chen, L., Flies, D.B., 2013. Molecular mechanisms of T cell co-stimulation and co-inhibition. *Nature Reviews Immunology* 13, 227–242.
- Choudhuri, K., Van Der Merwe, P.A., 2007. Molecular mechanisms involved in T cell receptor triggering. *Seminars in Immunology* 19, 255–261.
- Cochran, J.R., Cameron, T.O., Stern, L.J., 2000. The relationship of MHC-peptide binding and T cell activation probed using chemically defined MHC class II oligomers. *Immunity* 12, 241–250.

- Cochran, J.R., Cameron, T.O., Stone, J.D., *et al.*, 2001. Receptor proximity, not intermolecular orientation, is critical for triggering T-cell activation. *Journal of Biological Chemistry* 276, 28068–28074.
- Corse, E., Gottschalk, R.A., Allison, J.P., 2011. Strength of TCR-peptide/MHC interactions and *in vivo* T cell responses. *Journal of Immunology* 186, 5039–5045.
- Davis, M.M., Boniface, J.J., Reich, Z., *et al.*, 1998. Ligand recognition by alpha beta T cell receptors. *Annual Review of Immunology* 16, 523–544.
- Deford-Watts, L.M., Dougall, D.S., Belkaya, S., *et al.*, 2011. The CD3 zeta subunit contains a phosphoinositide-binding motif that is required for the stable accumulation of TCR-CD3 complex at the immunological synapse. *Journal of Immunology* 186, 6839–6847.
- Deford-Watts, L.M., Tassin, T.C., Becker, A.M., *et al.*, 2009. The cytoplasmic tail of the T cell receptor CD3 epsilon subunit contains a phospholipid-binding motif that regulates T cell functions. *Journal of Immunology* 183, 1055–1064.
- Delcassian, D., Depoil, D., Rudnicka, D., *et al.*, 2013. Nanoscale ligand spacing influences receptor triggering in T cells and NK cells. *Nano Letters* 13, 5608–5614.
- Dustin, M.L., Depoil, D., 2011. New insights into the T cell synapse from single molecule techniques. *Nature Reviews Immunology* 11, 672–684.
- Fernandes, R.A., Yu, C., Carmo, A.M., *et al.*, 2010. What controls T cell receptor phosphorylation? *Cell* 142, 668–669.
- Finco, T.S., Kadlecsek, T., Zhang, W., *et al.*, 1998. LAT is required for TCR-mediated activation of PLCgamma1 and the Ras pathway. *Immunity* 9, 617–626.
- Gil, D., Schamel, W.W., Montoya, M., *et al.*, 2002. Recruitment of Nck by CD3 epsilon reveals a ligand-induced conformational change essential for T cell receptor signaling and synapse formation. *Cell* 109, 901–912.
- Holst, J., Wang, H., Eder, K.D., *et al.*, 2008. Scalable signaling mediated by T cell antigen receptor-CD3 ITAMs ensures effective negative selection and prevents autoimmunity. *Nature Immunology* 9, 658–666.
- Huang, J., Brameshuber, M., Zeng, X., *et al.*, 2013. A single peptide-major histocompatibility complex ligand triggers digital cytokine secretion in CD4(+) T cells. *Immunity* 39, 846–857.
- Huang, Y.H., Sauer, K., 2010. Lipid signaling in T-cell development and function. *Cold Spring Harbor Perspectives in Biology* 2, a002428.
- James, J.R., Vale, R.D., 2012. Biophysical mechanism of T-cell receptor triggering in a reconstituted system. *Nature* 487, 64–69.
- Janeway Jr., C.A., 1995. Ligands for the T-cell receptor: Hard times for avidity models. *Immunology Today* 16, 223–225.
- Kalergis, A.M., Boucheron, N., Doucey, M.A., *et al.*, 2001. Efficient T cell activation requires an optimal dwell-time of interaction between the TCR and the pMHC complex. *Nature Immunology* 2, 229–234.
- Kaye, J., Porcelli, S., Tite, J., *et al.*, 1983. Both a monoclonal antibody and antisera specific for determinants unique to individual cloned helper T cell lines can substitute for antigen and antigen-presenting cells in the activation of T cells. *Journal of Experimental Medicine* 158, 836–856.
- Kim, S.T., Takeuchi, K., Sun, Z.Y., *et al.*, 2009. The alphabeta T cell receptor is an anisotropic mechanosensor. *Journal of Biological Chemistry* 284, 31028–31037.
- Klotzsch, E., Schutz, G.J., 2013. Improved ligand discrimination by force-induced unbinding of the T cell receptor from peptide-MHC. *Biophysical Journal* 104, 1670–1675.
- Krogsgaard, M., Li, Q.J., Sumen, C., *et al.*, 2005. Agonist/endogenous peptide-MHC heterodimers drive T cell activation and sensitivity. *Nature* 434, 238–243.
- Kuhns, M.S., Girvin, A.T., Klein, L.O., *et al.*, 2010. Evidence for a functional sidedness to the alphabetaTCR. *Proceedings of the National Academy of Sciences of the United States* 107, 5094–5099.
- Liu, B., Chen, W., Evavold, B.D., Zhu, C., 2014. Accumulation of dynamic catch bonds between TCR and agonist peptide-MHC triggers T cell signaling. *Cell* 157, 357–368.
- Lorenz, U., 2009. SHP-1 and SHP-2 in T cells: Two phosphatases functioning at many levels. *Immunological Reviews* 228, 342–359.
- Malissen, B., Ardouin, L., Lin, S.Y., *et al.*, 1999. Function of the CD3 subunits of the pre-TCR and TCR complexes during T cell development. *Advances in Immunology* 72, 103–148.
- Malissen, B., Gregoire, C., Malissen, M., Roncagalli, R., 2014. Integrative biology of T cell activation. *Nature Immunology* 15, 790–797.
- McKeithan, T.W., 1995. Kinetic proofreading in T-cell receptor signal transduction. *Proceedings of the National Academy of Sciences of the United States of America* 92, 5042–5046.
- Mingueneau, M., Kreslavsky, T., Gray, D., *et al.*, 2013. The transcriptional landscape of alphabeta T cell differentiation. *Nature Immunology* 14, 619–632.
- Mingueneau, M., Krishnaswamy, S., Spitzer, M.H., *et al.*, 2014. Single-cell mass cytometry of TCR signaling: Amplification of small initial differences results in low ERK activation in NOD mice. *Proceedings of the National Academy of Sciences of the United States of America* 111, 16466–16471.
- Mingueneau, M., Roncagalli, R., Gregoire, C., *et al.*, 2009. Loss of the LAT adaptor converts antigen-responsive T cells into pathogenic effectors that function independently of the T cell receptor. *Immunity* 31, 197–208.
- Mingueneau, M., Sansoni, A., Gregoire, C., *et al.*, 2008. The proline-rich sequence of CD3epsilon controls T cell antigen receptor expression on and signaling potency in preselection CD4+ CD8+ thymocytes. *Nature Immunology* 9, 522–532.
- Mohapatra, B., Ahmad, G., Nadeau, S., *et al.*, 2013. Protein tyrosine kinase regulation by ubiquitination: Critical roles of Cbl-family ubiquitin ligases. *Biochimica et Biophysica Acta* 1833, 122–139.
- Reich, Z., Boniface, J.J., Lyons, D.S., *et al.*, 1997. Ligand-specific oligomerization of T-cell receptor molecules. *Nature* 387, 617–620.
- Risueno, R.M., Ortiz, A.R., Alarcon, B., 2008a. Conformational model. *Advances in Experimental Medicine and Biology* 640, 103–112.
- Risueno, R.M., Schamel, W.W., Alarcon, B., 2008b. T cell receptor engagement triggers its CD3epsilon and CD3zeta subunits to adopt a compact, locked conformation. *PLoS One* 3, e1747.
- Roncagalli, R., Hauri, S., Fiore, F., *et al.*, 2014. Quantitative proteomics analysis of signalosome dynamics in primary T cells identifies the surface receptor CD6 as a Lat adaptor-independent TCR signaling hub. *Nature Immunology* 15, 384–392.
- Roncagalli, R., Mingueneau, M., Gregoire, C., *et al.*, 2010. LAT signaling pathology: An “autoimmune” condition without T cell self-reactivity. *Trends in Immunology* 31, 253–259.
- Rudolph, M.G., Stanfield, R.L., Wilson, I.A., 2006. How TCRs bind MHCs, peptides, and coreceptors. *Annual Review of Immunology* 24, 419–466.
- Salmond, R.J., Filby, A., Qureshi, I., *et al.*, 2009. T-cell receptor proximal signaling via the Src-family kinases, Lck and Fyn, influences T-cell activation, differentiation, and tolerance. *Immunological Reviews* 228, 9–22.
- Schamel, W.W., Alarcon, B., 2013. Organization of the resting TCR in nanoscale oligomers. *Immunological Reviews* 251, 13–20.
- Shen, S., Chuck, M.I., Zhu, M., *et al.*, 2010. The importance of LAT in the activation, homeostasis, and regulatory function of T cells. *Journal of Biological Chemistry* 285, 35393–35405.
- Sommers, C.L., Lee, J., Steiner, K.L., *et al.*, 2005. Mutation of the phospholipase C-gamma1-binding site of LAT affects both positive and negative thymocyte selection. *Journal of Experimental Medicine* 201, 1125–1134.
- Sommers, C.L., Menon, R.K., Grinberg, A., *et al.*, 2001. Knock-in mutation of the distal four tyrosines of linker for activation of T cells blocks murine T cell development. *Journal of Experimental Medicine* 194, 135–142.
- Stern, L.J., Wiley, D.C., 1994. Antigenic peptide binding by class I and class II histocompatibility proteins. *Structure* 2, 245–251.
- Valitutti, S., Muller, S., Cella, M., *et al.*, 1995. Serial triggering of many T-cell receptors by a few peptide-MHC complexes. *Nature* 375, 148–151.
- Van Der Merwe, P.A., Dushek, O., 2011. Mechanisms for T cell receptor triggering. *Nature Reviews Immunology* 11, 47–55.
- Van Panhuys, N., Klauschen, F., Germain, R.N., 2014. T-Cell-receptor-dependent signal intensity dominantly controls CD4(+) T cell polarization *in vivo*. *Immunity* 41, 63–74.
- Varma, R., Campi, G., Yokosuka, T., *et al.*, 2006. T cell receptor-proximal signals are sustained in peripheral microclusters and terminated in the central supramolecular activation cluster. *Immunity* 25, 117–127.
- Wang, H., Kadlecsek, T.A., Au-Yeung, B.B., *et al.*, 2010. ZAP-70: An essential kinase in T-cell signaling. *Cold Spring Harbor Perspectives in Biology* 2, a002279.
- Wang, X., Li, J.P., Chiu, L.L., *et al.*, 2012. Attenuation of T cell receptor signaling by serine phosphorylation-mediated lysine 30 ubiquitination of SLP-76 protein. *Journal of Biological Chemistry* 287, 34091–34100.
- Ward, C.W., Lawrence, M.C., Streltsov, V.A., *et al.*, 2007. The insulin and EGF receptor structures: New insights into ligand-induced receptor activation. *Trends in Biochemical Sciences* 32, 129–137.
- Weiss, A., 2005. Discovering the TCR beta-chain by subtraction. *Journal of Immunology* 175, 2769–2770.
- Xu, C., Gagnon, E., Call, M.E., *et al.*, 2008. Regulation of T cell receptor activation by dynamic membrane binding of the CD3epsilon cytoplasmic tyrosine-based motif. *Cell* 135, 702–713.
- Yoon, S.T., Dianzani, U., Bottomly, K., Janeway Jr., C.A., 1994. Both high and low avidity antibodies to the T cell receptor can have agonist or antagonist activity. *Immunity* 1, 563–569.

Further Reading

- Brownlie, R.J., Zamoyska, R., 2013. T cell receptor signalling networks: Branched, diversified and bounded. *Nature Reviews Immunology* 13, 257–269.
- Chakraborty, A.K., Weiss, A., 2014. Insights into the initiation of TCR signaling. *Nature Immunology* 15, 798–807.
- Germain, R.N., 2010. Computational analysis of T cell receptor signaling and ligand discrimination—past, present, and future. *FEBS Letters* 584, 4814–4822.
- Sherman, E., Barr, V., Samelson, L.E., 2013. Super-resolution characterization of TCR-dependent signaling clusters. *Immunological Reviews* 251, 21–35.
- Xie, J., Tato, C.M., Davis, M.M., 2013. How the immune system talks to itself: The varied role of synapses. *Immunological Reviews* 251, 65–79.

T Follicular Helper Cells

MM Shannack and MA Linterman, Babraham Institute, Cambridge, UK

© 2016 Elsevier Inc. All rights reserved.

Introduction

After infection or immunization, $CD4^+$ T cells respond by differentiating into different subsets of helper T cells that each has a distinct function. Initially, two types of effector $CD4^+$ T cells were described (Mosmann *et al.*, 1986), T helper (Th) type 1 cells that form during infection by intracellular bacteria or protozoa, and Th type 2 cells that form in response to extracellular parasites. It is now recognized that there are numerous T cell subsets that act to facilitate the clearance of pathogens and generate immunological memory. A common feature of the responses to most pathogens, and vaccination, is the production of antibodies that act to neutralize infectious agents and trigger activation of other immune cells. Thus, antibodies can provide protection against subsequent infection by blocking the ability of pathogens to infect the host. Antibodies can be produced via two different cellular pathways. The first is the short-lived extrafollicular response that produces an initial burst of antibodies rapidly after infection. The extrafollicular response can occur with or without T cell help. The second pathway is the germinal center (GC) response, which produces long-lived plasma cells that secrete antibodies that bind pathogens with high affinity. These GC-derived plasma cells are able to provide prolonged protection against subsequent infection. The formation and function of the GC is absolutely dependent on 'help' from a specialized subset of $CD4^+$ helper T cells known as T follicular helper (Tfh) cells. Within the GC, Tfh cells provide help to GC B cells which is essential for mediating the processes of affinity maturation, driven by somatic hypermutation (SHM) and subsequent selection. Tfh cells are instrumental in selecting the highest-affinity GC B cells for differentiation into long-lived plasma cells or memory B cells, and are therefore critical for the output of the GC response. Consequently, Tfh cells are critical for vaccine-mediated immune responses and have been connected with immune pathology in autoimmune diseases and even cancer (Crotty, 2014). In this article we cover the stages of Tfh cell differentiation, their role in the GC response, and how they have been implicated in health and disease.

T Follicular Helper Cell Differentiation

Initiation of Tfh Development

Tfh cell differentiation is a multistage process that is initiated when naïve $CD4^+$ T cells are activated by antigen-presenting dendritic cells (DCs) in the T zone of secondary lymphoid tissues (Moreau and Bousso, 2014). The interactions between DCs and naïve $CD4^+$ T cells are necessary for T cell activation, and it is during these interactions that DCs provide specific cues to T cells that trigger their differentiation into Tfh cells (Ballesteros-Tato and Randall, 2014). As with all other T cell subsets, the first step in Tfh cell differentiation is for naïve T

cells to be primed, requiring two signals to be delivered in tandem. The first is the presentation of processed antigen peptide fragments to helper T cells on class II Major Histocompatibility Complex (MHCII) molecules by DCs, which are recognized by the T cell receptor (TCR) (Grakoui *et al.*, 1999). The second signal is a co-stimulatory signal that involves ligation of the receptor CD28 on T cells by either of its ligands, CD80 or CD86, expressed by a DC that has received the appropriate 'danger' signals (Sharpe and Freeman, 2002). When co-stimulation occurs in concert with MHCII/TCR ligation, T cells are primed and can undergo further development into specialized helper T cell subsets (Deenick and Ma, 2011; Figure 1). The divergence of activated T cells into different T cell subsets is a process that relies on the integration of multiple signals that culminate in the expression of the transcription factors that drive specific T cell fates.

The transcriptional repressor B cell lymphoma-6 (Bcl-6) is essential for Tfh cell differentiation (Nurieva *et al.*, 2009; Yu *et al.*, 2009; Johnston *et al.*, 2009). The initial induction of Bcl-6 expression in Tfh precursor (pre-Tfh) cells requires signals from multiple cell-surface receptors and cytokines delivered by DCs to pre-Tfh cells. Once expressed in pre-Tfh cells, Bcl-6 is able to suppress the expression or function of transcription factors that promote alternative cell fates, thereby skewing the cell toward a pro-Tfh cell lineage (Johnston *et al.*, 2009; Nurieva *et al.*, 2009; Yu *et al.*, 2009).

One of the key requirements for Bcl-6 expression in pre-Tfh cells is ligation of the inducible co-stimulator (ICOS) by ICOS ligand (ICOSL) on DCs (Simpson *et al.* 2010a,b; Bossaller *et al.* 2006), which activates the phosphatidylinositol-3-kinase (PI3K) pathway to facilitate both Tfh differentiation and the subsequent migration of pre-Tfh cells to the border of the B cell follicle and T cell zone (T:B border) (Simpson *et al.*, 2010a,b; Rolf *et al.*, 2010; Choi *et al.*, 2011). Various cytokines produced by DCs synergize with ICOS signaling to promote Tfh formation. Intriguingly, the pro-Tfh cytokines differ between mice and humans. In mice, interleukin (IL)-6 triggers a signaling cascade that results in the phosphorylation of signal transducer and activator of transcription 3 (STAT3), which promotes Tfh differentiation (Ma *et al.*, 2012). In humans, IL-12, IL-21, IL-23, and transforming growth factor- β (TGF- β) are important mediators of Tfh development (Eto *et al.*, 2011; Choi *et al.*, 2013a,b; Nurieva *et al.*, 2008; Ma *et al.*, 2009; Schmitt *et al.*, 2014). After these cytokines bind their respective receptors on activated human T cells, both STAT3 and STAT4 are activated (Choi *et al.*, 2013a,b; Ma *et al.*, 2012). Activated STAT3 and STAT4 then translocate to the nucleus where they contribute to the induction of Bcl-6 (Choi *et al.*, 2013a,b). STAT3 can promote the production and release of IL-21 creating an autocrine signaling loop to further potentiate Bcl-6 expression and Tfh differentiation (Caprioli *et al.*, 2008).

In addition to stimuli that can induce Tfh differentiation, there are also signals that skew activated T cells away from this cell fate. IL-2 is a potent inhibitor of Tfh differentiation,

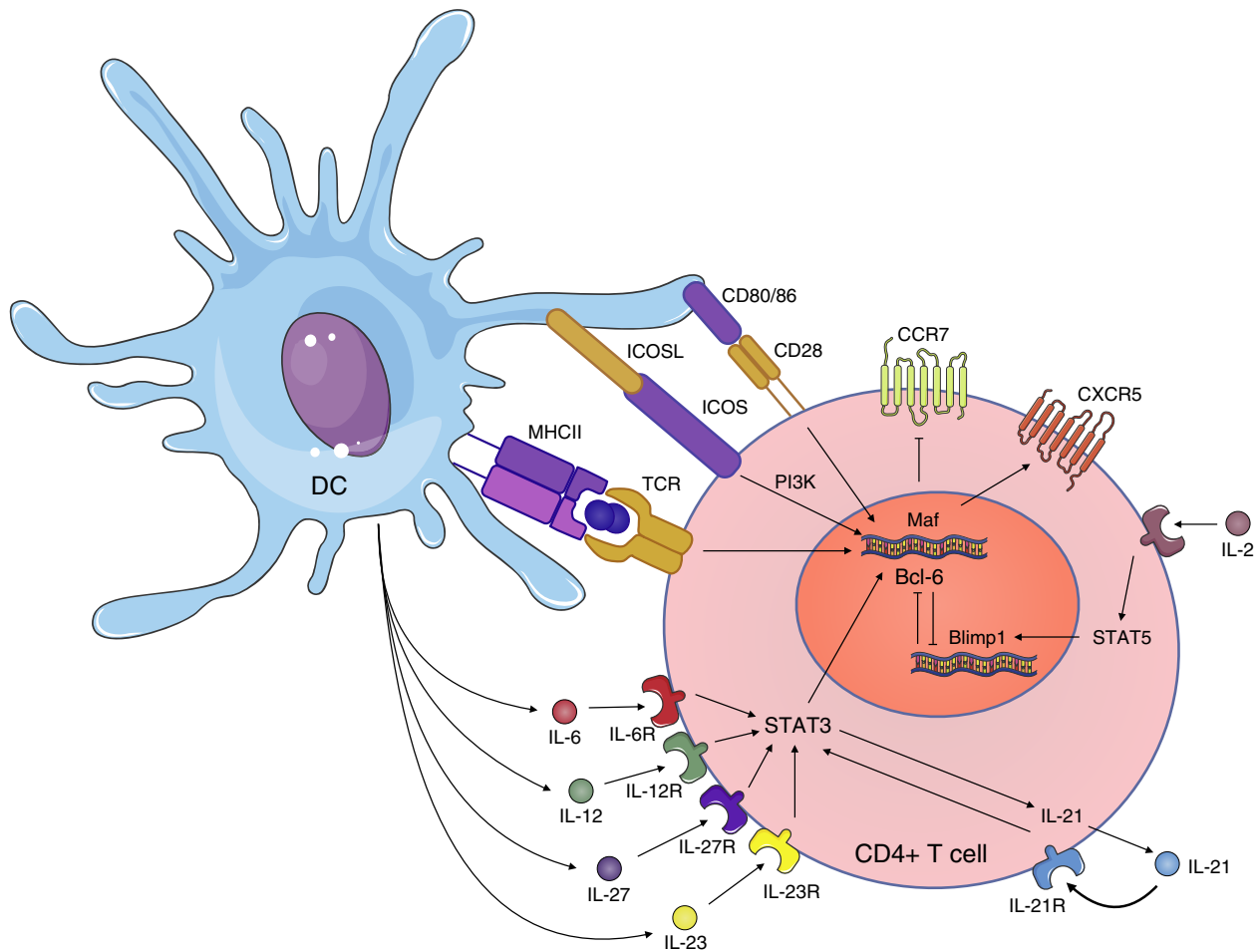


Figure 1 T cell priming and activation. Naïve CD4⁺ T cells are primed by DCs in the T zone of secondary lymphoid organs through a multitude of cell-surface interactions and external cytokine signals. Two critical activation signals include antigen presentation through MHCII binding to the TCR, and CD28 ligation by CD80/86. The downstream effects of this co-stimulatory activation are induction of Bcl-6 and Maf expression, both of which are known to be important in the development of the Tfh cell phenotype. ICOS/ICOSL signaling contributes to pre-Tfh cell positioning along the T:B border and is also involved with inducing expression of Bcl-6/Maf through the PI3K signaling pathway. DCs produce a range of cytokines that drive Tfh cell differentiation such as IL-6 (mouse), IL-12 (human), IL-21, IL-23 (human), and IL-27. Cytokine binding triggers STAT3 activation and the subsequent induction of Bcl-6, Maf. Negative regulation of Tfh development may occur through the IL-2/STAT5/Blimp1 axis, which is antagonistic to Bcl-6 activity.

therefore the expression of its receptor on activated T cells and/or the bioavailability of IL-2 can polarize T cell differentiation. When IL-2 binds its receptor it triggers a STAT5-dependent activation of B lymphocyte-induced maturation protein 1 (Blimp-1). Blimp-1 is a negative regulator of Tfh cell differentiation as it acts as a potent repressor of Bcl-6 (Johnston *et al.*, 2009; Pepper *et al.* 2011; Johnston *et al.*, 2012).

While Bcl-6 is critical for pre-Tfh cell development, and is often considered to be the ‘master regulator’ of Tfh cell formation, it is not the only known transcriptional regulator involved in this cell fate. These include Maf, interferon regulatory factor 4 (IRF4), basic leucine zipper transcription factor (Batf), and achaete-scute homologue 2 (Ascl-2), which synergize with Bcl-6 for Tfh formation. Maf is induced by ICOS ligation and is central in supporting Tfh cell differentiation and it also supports IL-21 production by Tfh cells (Bauquet *et al.*, 2009; Vogelzang *et al.*, 2008; Kroenke *et al.*, 2012). IRF4 and Batf are

produced at an early stage of T cell activation and contribute significantly to Tfh cell differentiation by ensuring the expression and function of the pro-Tfh lineage factors Bcl-6 and Maf (Ise *et al.*, 2011; Bollig *et al.*, 2012). More recently, the basic helix-loop-helix transcription factor Ascl-2 has been implicated in the initiation of Tfh development (Liu *et al.*, 2014). Ascl-2, like Bcl-6, can repress the transcription of alternative lineage genes and was suggested to precede Bcl-6 upregulation. However, unlike Bcl-6, which is necessary for Tfh formation, T cells that lack Ascl-2 are still able to become Tfh cells, suggesting redundancy in the transcription factors required for Tfh differentiation (Liu *et al.*, 2014).

After being primed by DCs and initiating pre-Tfh cell differentiation, Tfh-precursors require further rounds of antigen presentation to potentiate full Tfh cell development. Although Bcl-6 is a transcriptional repressor, it functions to increase the expression of molecules that are crucial for Tfh migration to

the T-B border or to the interfollicular zone where they encounter B cells (Kroenke *et al.*, 2012). Interactions between pre-Tfh cells and antigen-activated B cells provide further differentiation signals to secure commitment to the Tfh cell lineage.

Stabilization of Tfh Cell Development

Pre-Tfh cells must initiate further interactions with antigen-presenting cells to complete their differentiation. At this stage in the response the majority of mature DCs have died, while the number of antigen-activated B cells, that can also act as antigen-presenting cells, has greatly increased due to clonal expansion. To interact with B cells, pre-Tfh cells need to localize to distinct regions of the secondary lymphoid tissues where they have access to B cells. The Tfh differentiation program, initiated by Bcl-6 and Ascl-2, endows pre-Tfh cells with a chemokine receptor expression profile that facilitates migration to the T-B border and interfollicular zones, where they are able to encounter B cells. Pre-Tfh cells express C-X-C chemokine receptor type 5 (CXCR5), which promotes migration toward the C-X-C motif chemokine 13 (CXCL13)-rich B cell follicle. In addition, pre-Tfh have downregulated C-C chemokine receptor type 7 (CCR7) and P selectin glycoprotein ligand 1 (PSGL1), both of which promote localization of T cells to the CCL19/21-rich T zone (Breitfeld *et al.*, 2000; Yu *et al.*, 2009). In addition to its role as a pro-Tfh co-stimulatory molecule, ICOS is important for the migration of pre-Tfh cells toward the B cell follicle by engaging ICOSL expressed by bystander B cells that do not present antigen to pre-Tfh cells (Xu *et al.*, 2013). In this way, these bystander B cells help localize Tfh cells to the follicular border where they can engage in interactions with, and provide help to, cognate B cells.

The cognate interactions between pre-Tfh and activated B cells at the edge of the B cell follicle are a bidirectional exchange that is essential for the migration of B cells into the GC and for pre-Tfh cells to complete their development into GC Tfh cells (Okada *et al.*, 2005; Kerfoot *et al.*, 2011). These interactions involve a series of co-stimulatory signals that stabilize Bcl-6 expression in pre-Tfh cells. Although B cells are the major type of antigen-presenting cell at this phase of the immune response, they do not deliver a distinct 'B cell specific' signal for Tfh formation as this can be substituted by other cell types in some circumstances (Deenick *et al.*, 2010). The stability of conjugate formation between pre-Tfh and B cells is critical for the transmission of signals between these cells (Cannons *et al.*, 2011). This cellular adhesion between T and B cells is mediated by the signaling lymphocytic activation molecule (SLAM) family members CD84 and Ly108 that are coupled to intracellular signaling pathways by the SRC homology 2 (SH2) domain of SLAM-associated protein (SAP) (Cannons *et al.*, 2011; Kageyama *et al.*, 2012). In the absence of SAP, pre-Tfh/B cells cannot form stable conjugates and Tfh cells do not form (Qi *et al.*, 2008). This results in impaired GC formation and long-lived antibody production, although the extrafollicular short-lived antibody response is intact (Linterman *et al.*, 2009).

During interactions at the T-B border, T and B cells exchange signals through cell-surface molecules and cytokines

that are important for Tfh cell commitment and B cell development (Figure 2). CD40L expressed on Tfh cells binds to its receptor CD40 on cognate B cells promoting B cell survival, antibody class-switching, and inducing ICOSL and CD80/86 expression on the B cell surface, which in turn provide survival signals to Tfh cells (Liang *et al.*, 2002; Foy *et al.*, 1996; Elgueta *et al.*, 2009; Gigoux *et al.*, 2009; Rolf *et al.*, 2010; Linterman *et al.*, 2014). In addition, Tfh differentiation is also supported by various cytokines. In mice, IL-21 and IL-6 support Tfh differentiation, as does IL-27 in both mice and humans (Nurieva *et al.*, 2007; Batten *et al.*, 2010).

As a result of these interactions at the T-B border, there is a fulfillment of Tfh cell differentiation and sustained Bcl-6 expression. These Tfh cells then enter the GC where they provide survival and differentiation cues to GC B cells, and play a central role in selecting which B cells will exit the GC as long-lived plasma cells or memory B cells.

The GC

After receiving T cell help at the T-B border, activated B cells either commit to the extrafollicular pathway, where they develop into short-lived antibody-producing plasma cells, or enter the follicles, where they initiate GC formation (MacLennan *et al.*, 2003). The B cells which will seed the GC downregulate Epstein-Barr virus-induced G-protein coupled receptor 2 (EBI2), enabling them to migrate into the B cell follicle where they undergo extensive proliferation (Pereira *et al.*, 2009). Following an initial rapid proliferative phase, the GC structure becomes anatomically polarized. Proliferating B cells, called centroblasts, localize to the region of the GC closest to the T cell zone termed the dark zone (DZ). This polarization is facilitated by expression of CXCR4 on the centroblasts enabling chemotaxis toward the CXCL12-rich areas in the DZ (Allen *et al.*, 2004). Adjacent to the DZ, and close to sites of antigen entry into the lymphoid tissue, is the light zone (LZ) that contains a network of follicular dendritic cells (FDCs), the nondividing progeny of centroblasts known as centrocytes, and GC Tfh cells (Allen *et al.*, 2004; Victora *et al.*, 2010). Each zone serves a distinct purpose: B cell proliferation and SHM in the DZ and selection by GC Tfh cells and FDCs in the LZ (Figure 3; Allen *et al.*, 2007).

Within the DZ, B cells divide and mutate their B cell receptor (BCR) genes through the directed process of SHM, which introduces random point mutations into the variable (V) region of immunoglobulin genes in a process mediated by the enzyme activation-induced deaminase (AID) (Noia and Neuberger, 2007). SHM results in the generation of a spectrum of BCRs that vary in antigen affinity. The random nature of the mutations induced by SHM requires that centroblasts undergo a process of selection prior to differentiation into long-lived antibody-secreting plasma cells and memory B cells. Following SHM, centroblasts exit the cell cycle, becoming centrocytes that migrate to the LZ where they scan antigen complexes presented by FDCs (Miller *et al.*, 2002). This serves as an opportunity for GC B cells to test their BCR binding affinity. B cells use their newly mutated BCRs to bind and internalize these antigens, then process and present them on MHCII to GC Tfh cells (Figure 4). Only a small fraction of GC B cells that pass FDC

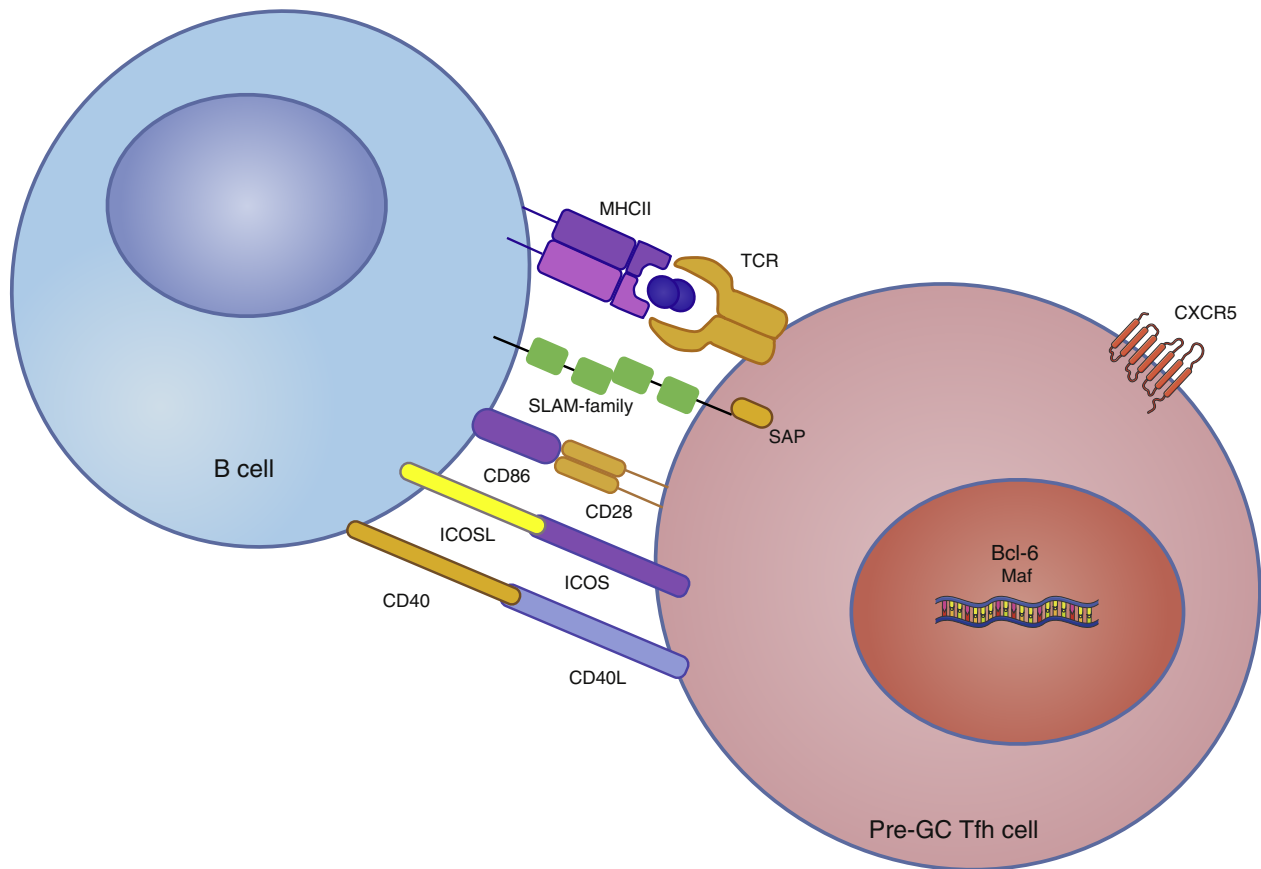


Figure 2 Interactions at the T:B border. After migration to the T:B border, pre-GC Tfh cells complete their commitment to the Tfh lineage through a series of interactions with cognate B cells. B cells take over the antigen-presenting role of DCs by providing an MHCII/TCR signal. ICOS/ICOSL, CD40/CD40L, and CD28/CD86 interactions provide important developmental cues for both pre-GC Tfh cells and B cells. The duration of cognate interactions is crucial for successful signal transduction and is ensured through SLAM-family/SAP-mediated adhesion and stability.

scanning receive sufficient ‘help’ to survive and go on to form stable interactions with cognate GC Tfh cells (Allen *et al.*, 2007). It is thought that GC Tfh cells recognize the relative amounts of peptide:MHCII on the B cell surface as a proxy for the antigen affinity of a given B cell (Victora *et al.*, 2010; Gitlin *et al.*, 2014). A mutated BCR with greater affinity can acquire and process more of an antigen from FDCs and therefore present more peptide:MHCII to GC Tfh cells, thereby out-competing a B cell with a lower affinity receptor that presents less. Thus, GC B cells with mutated BCRs are subject to a Darwinian-like selection that enriches for a population of B cells with enhanced affinity for antigen (Allen *et al.*, 2007; Tarlinton, 2008). Positively selected B cells can differentiate into antibody secreting plasma cells or memory B cells that are exported from the GC, or they can reenter the DZ for further proliferation and BCR diversification (Victora and Mesin, 2014). In this way, GC B cells can experience reiterative cycling between BCR diversification in the DZ and positive selection in the LZ, yielding increasingly higher-affinity cells (Camacho *et al.*, 1998). Occasionally, SHM can produce BCR variants that lose antigen binding completely or have had their antigen specificity altered to recognize self-antigens. In such instances, these cells do not receive help from GC Tfh cells, undergo

apoptosis and are subsequently cleared from the GC by resident macrophages (Victora *et al.*, 2010). Therefore, help from GC Tfh cells is an important mechanism of peripheral tolerance that, if perturbed, can lead to humoral autoimmunity (Linterman *et al.*, 2009).

GC Tfh cells provide help to GC B cells through a range of cytokines and cell-surface molecules and also maintain the expression of SAP that is essential for adhesion and stable interactions (Qi *et al.*, 2008). The key signals from the GC Tfh cells to B cells include CD40L, IL-21, and IL-4 (Crotty, 2014). In mice, IL-21 is required for optimal Bcl-6 expression in GC B cells, which is the transcription factor essential for maintaining the GC B cell phenotype (Linterman *et al.*, 2010; Zotos *et al.*, 2010). In response to SLAM-mediated signaling, GC Tfh cells express elevated levels of Maf, which promotes both IL-21 and IL-4 production by GC Tfh cells (Yusuf *et al.*, 2010; Bauquet *et al.*, 2009; Kim *et al.*, 1999). The cytokines produced by GC Tfh cells support the expression of the molecular machinery involved with class-switch recombination and SHM in GC B cells that are engaged in conjugates with GC Tfh cells (Reinhardt *et al.*, 2009; Basso *et al.*, 2012; Avery *et al.*, 2008). The dialog between GC Tfh and B cells is bidirectional, as during these interactions B cells also deliver signals to GC

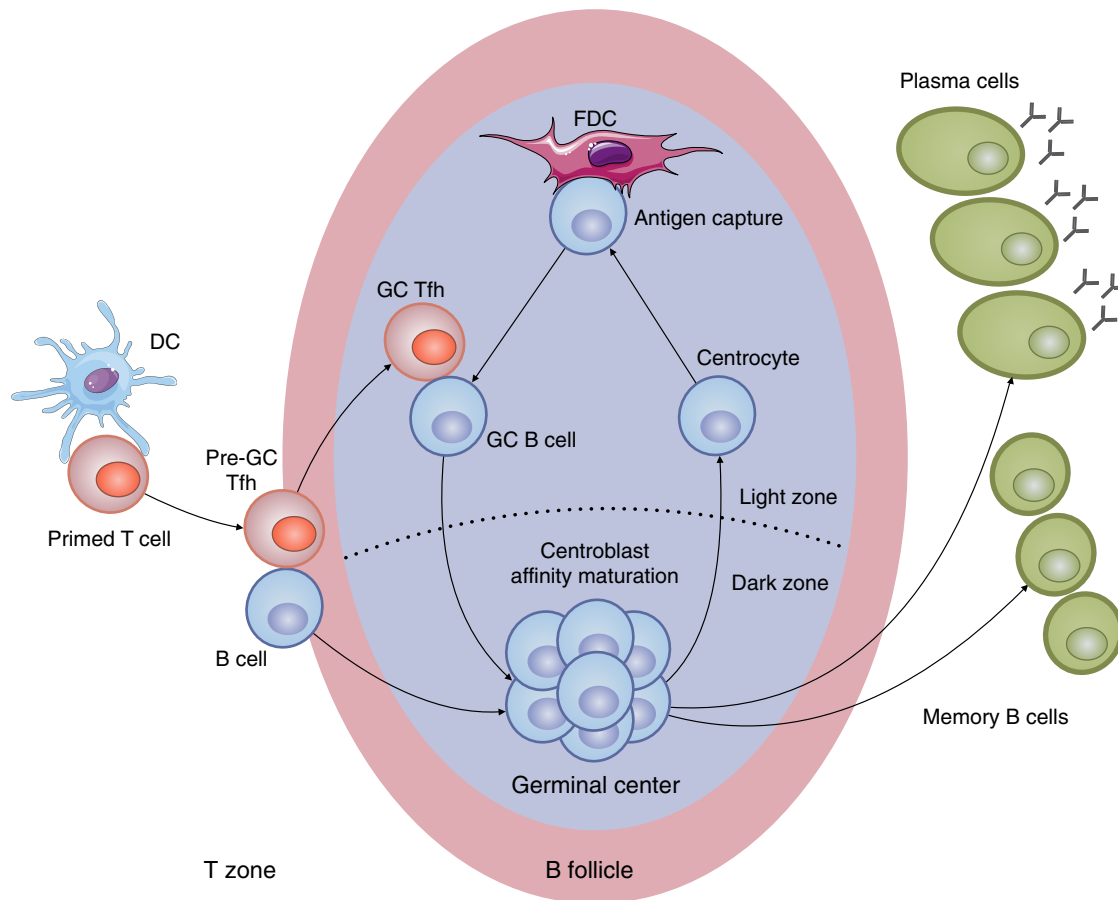


Figure 3 Overview of Tfh cell activation, development, and role in the GC response. After naïve $CD4^{+}$ T cells in the T zone are primed by antigen-presenting DCs, they begin the process of differentiation into pre-Tfh cells. Changes in cell-surface chemokine receptors cause migration of primed T cells to the T:B border where they undertake a series of interactions with cognate B cells culminating in Tfh cell differentiation. After receiving cognate T cell help, B cells migrate to the LZ of the GC and become rapidly dividing centroblasts. These centroblasts undergo SHM of their BCR genes before exiting the cell cycle and migrating from the LZ as nondividing centrocytes that localize to the LZ. The centrocytes then scan FDCs and capture presented antigens, which are then processed into peptide fragments. Next, the centrocytes compete to receive help from GC Tfh cells. Those cells with the greatest antigen-presenting capacity, due to their ability to capture the most antigens from FDCs, outcompete lower affinity B cells for help from GC Tfh cells. In the absence of T cell help, GC B cells undergo apoptosis. Those high-affinity GC B cells that receive T cell help then reenter the LZ where they can undergo expansion and further SHM before being exported as either long-lived, high-affinity plasma cells or memory B cells.

Tfh cells that enhance their survival, such as CD28 stimulation that prevents apoptosis of Tfh cells (Good-Jacobson *et al.*, 2010; Linterman *et al.*, 2014).

In addition to positive regulation, GC Tfh cells can also be recipients of negative regulation in the GC. This control of Tfh cell numbers is likely to be critical for a high-quality GC response. Indiscriminate GC Tfh cell help to GC B cells results in a population of GC B cells that have not increased their affinity for antigen (Victoria *et al.*, 2010). By limiting Tfh numbers relative to the number of B cells, Tfh cell help is restricted to only higher-affinity GC B cells that can outcompete lower affinity cells for Tfh cell help. GC Tfh cells express high levels of programmed cell death 1 (PD-1), a CD28 family co-inhibitory receptor (Haynes *et al.*, 2007) that controls the size of the Tfh cell population (Good-Jacobson *et al.*, 2010; Kawamoto *et al.*, 2012; Qi *et al.*, 2014).

The GC is also subject to negative regulation by a defined subset of Foxp3⁺ T cells within the GC known as follicular

regulatory T (Tfr) cells. These cells have been reported as having an inhibitory role in GC size and output as well as repressing GC Tfh cell numbers (Wollenberg *et al.*, 2011; Chung *et al.*, 2011; Linterman *et al.*, 2011). The precise role of Tfr cells in the GC has yet to be revealed, but it is clear that they are important for controlling the size of the GC response.

After the demise of the GC, Tfh cells are able to become memory cells in both mice and humans (Locci *et al.*, 2013; Choi *et al.*, 2013a,b; Hale *et al.*, 2013). These can be resident in secondary lymphoid tissues or circulate in the blood (Fazilleau *et al.*, 2007). The population of circulating Tfh-like cells is a valuable tool for investigating Tfh cells in humans. Upon restimulation, these memory Tfh cells can engage in cognate interactions with memory B cells and support new GC responses (Hale *et al.*, 2013). This, coupled with the formation of memory B cells from the GC, facilitates the establishment of a rapid GC response following future exposure to the same antigen(s). It is this creation of T and B cell memory that is the

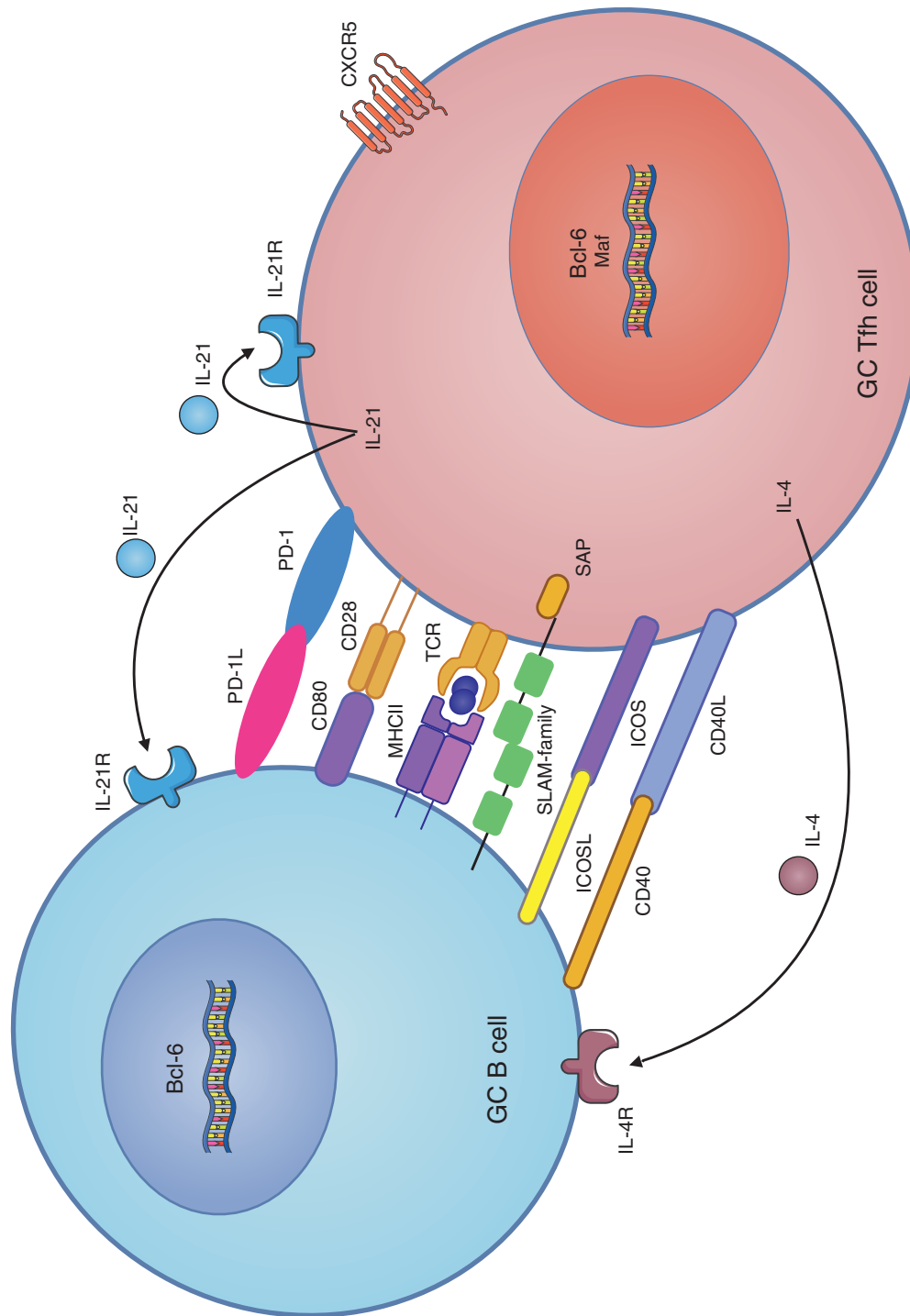


Figure 4 Provision of help to GC B cells by GC Tfh cells. GC B cells present antigen to GC Tfh cells via peptide MHCII interacting with the TCR. During these interactions, T and B cells also exchange signals through ICOS/ICOSL, CD40/CD40L, and CD28/CD80 interactions, which provide maintenance signals for both GC Tfh and GC B cells. The SLAM-family/SAP provides the necessary stability for sufficient duration of contact to elicit successful signals. PD-1/PD-1L interactions deliver an inhibitory signal to prevent GC Tfh cells from becoming excessive in number. IL-4 and IL-21 produced by GC Tfh cells provide support to GC B cells, in part, by IL-21 promoting Bcl-6 expression in B cells.

basis for successful vaccination, and that creates the potential for life-long autoimmunity to arise from the GC if the response goes awry.

Tfh Cells in Health and Disease

Given their central role in the selection of high-affinity B cells in the GC, Tfh cells are required for the generation of long-lived, high-affinity plasma cells that are involved in the protective immunity conferred by vaccination. Indeed, most licensed human vaccines are designed to generate long-term antibody responses that block pathogens from establishing an infection. There is now considerable evidence to identify Tfh cell function as the limiting factor for vaccine-induced antibody production in humans (Pallikkuth *et al.*, 2012; Bentebibel *et al.*, 2013; Duan *et al.*, 2014). Additionally, boosting Tfh function in a mouse malaria model provided a significantly higher antibody response (Butler *et al.*, 2012). Taken together, this evidence indicates that Tfh-boosting approaches could be employed to therapeutically enhance vaccination responses.

Tfh cells have been implicated in a number of autoimmune diseases, particularly those in which autoantibody formation is involved in disease pathogenesis. In mice, it is clear that Tfh cells are able to drive a lupus-like phenotype due to dysregulation of the GC response (Linterman *et al.*, 2009). In humans, an increased frequency of circulating Tfh-like cells has been observed in multiple autoimmune diseases such as systemic lupus erythematosus (Simpson *et al.*, 2010a,b), Sjögren's syndrome (Szabo *et al.*, 2013), and juvenile dermatomyositis (Morita *et al.*, 2011). Moreover, increased frequencies of Tfh-like cells in systemic lupus erythematosus and rheumatoid arthritis patients correlate with greater autoantibody levels and more severe disease profiles (Simpson *et al.*, 2010a,b; He *et al.*, 2013). Together, these and many other studies demonstrate that Tfh cells play a central role in humoral autoimmunity, and represent a potential therapeutic target for treatment of these diseases.

One of the more surprising roles for Tfh cells in recent years is their relevance in some types of cancer. An increased presence of Tfh cells and ectopic B cell follicles near the tumor was found to be a biomarker of greater survival in two independent studies of colorectal and breast cancer (Bindea *et al.*, 2013; Gu-trantien *et al.*, 2013). Although the precise role of Tfh cells in tumor immunity is not known, these studies indicated that IL-21 and CXCR13 may be important potentiators of Tfh cell infiltration and their levels positively correlated with patient survival. How Tfh cells might mediate these protective effects will likely be an exciting research area in the future.

Tfh cells are also central to the production of high-affinity IgA in the gut that controls the homeostasis of the commensal microflora there (Kato *et al.*, 2014). Within the gut, IgA can either form in a T-independent fashion or can arise from GCs within the Peyer's patches and give rise to IgA secreting plasma cells with mutated BCRs. If Tfh cell numbers are not controlled in the gut, then the quality of the IgA response is compromised, leading to dysbiosis of the commensal microbiota (Kawamoto *et al.*, 2014).

With advancing age, there is a steady decline in immune function, leading to increased risk of infection, and consequent

morbidity and mortality associated with infectious disease. Despite the success of vaccination programs in younger individuals, vaccines are rather ineffective in older persons (Goronzy and Weyand, 2013). In mice, a decline in GC size occurs with increasing age, as do antibody titers in humans (Stiasny *et al.*, 2012). It appears that aging of the T cells, rather than B cells, is primarily responsible for GC dysfunction with age (Eaton *et al.*, 2004), suggesting that impaired Tfh formation and/or function with advancing age is probably responsible for this decline. In addition, the aged micro-environment also reduces Tfh cell differentiation, suggesting that both T cell intrinsic and T cell extrinsic factors can alter the Tfh cell population in older individuals (Lefebvre *et al.*, 2012).

It is clear that Tfh cells have multifaceted and varied roles in health and disease. Although there has been a great deal of progress in studying Tfh cell biology, further advances in our knowledge may allow Tfh cell manipulation to enhance protective immunity for vaccines and cancer, and inhibit Tfh cells in autoimmune diseases.

Acknowledgments

Michelle A. Linterman is supported by the BBSRC. Michael M. Shannack is the recipient of a BBSRC CASE award. Servier Medical Art used in figures is licensed under Creative Commons Attribution 3.0 Unported.

See also: Cellular Immunology: Cell–Cell Interactions and the Immune System: T Cell Receptor Triggering. Cellular Immunology: Transcriptional Basis of B and T Cell Lineages and Memory Cells: T Cell Memory to Viral Infections

References

- Allen, C.D.C., Ansel, K.M., Low, C., *et al.*, 2004. Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nature Immunology* 5 (9), 943–952. doi:10.1038/ni1100.
- Allen, C.D.C., Okada, Takaharu, Tang, H.L., Cyster, J.G., 2007. Imaging of germinal center selection events during affinity maturation. *Science* (New York, NY) 315 (5811), 528–531. doi:10.1126/science.1136736.
- Avery, D.T., Bryant, V.L., Ma, C.S., de Waal Malefyt, R., Tangye, S.G., 2008. IL-21-induced isotype switching to IgG and IgA by human naive B cells is differentially regulated by IL-4. *Journal of Immunology* 181 (3), 1767–1779. doi:10.4049/jimmunol.181.3.1767.
- Ballesteros-Tato, A., Randall, T.D., 2014. Priming of T follicular helper cells by dendritic cells. *Immunology and Cell Biology* 92 (1), 22–27. doi:10.1038/icb.2013.62.
- Basso, K., Schneider, C., Shen, Q., *et al.*, 2012. BCL6 positively regulates AID and germinal center gene expression via repression of miR-155. *Journal of Experimental Medicine* 209 (13), 2455–2465. doi:10.1084/jem.20121387.
- Batten, M., Ramamoorthi, N., Kijavini, N.M., *et al.*, 2010. IL-27 supports germinal center function by enhancing IL-21 production and the function of T follicular helper cells. *Journal of Experimental Medicine* 207 (13), 2895–2906. doi:10.1084/jem.2010064.
- Bauquet, A.T., Jin, H., Paterson, A.M., *et al.*, 2009. The costimulatory molecule ICOS regulates the expression of c-Maf and IL-21 in the development of follicular T helper cells and TH-17 cells. *Nature Immunology* 10 (2), 167–175. doi:10.1038/ni.1690.
- Bentebibel, S.E., Lopez, S., Obermoser, G., *et al.*, 2013. Induction of ICOS⁺ CXCR3⁺ CXCR5⁺ TH cells correlates with antibody responses to influenza

- vaccination. *Science Translational Medicine* 5 (176), 1–19. doi:10.1126/scitranslmed.3005191.Induction.
- Bindea, G., Mlecnik, B., Tosolini, M., *et al.*, 2013. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 39 (4), 782–795. doi:10.1016/j.immuni.2013.10.003.
- Bollig, N., Brüstle, A., Kellner, K., *et al.*, 2012. Transcription factor IRF4 determines germinal center formation through follicular T-helper cell differentiation. *Proceedings of the National Academy of Sciences of the United States of America* 109 (22), 8664–8669. Available at: <http://dx.doi.org/doi:10.1073/pnas.1205834109/-DCSupplemental>.
- Bossaller, L., Burger, J., Draeger, R., *et al.*, 2006. ICOS deficiency is associated with a severe reduction of CXCR5 + CD4 germinal center Th cells. *Journal of Immunology* 177, 4927–4932.
- Breitfeld, D., Ohi, L., Kremmer, E., *et al.*, 2000. Follicular B helper T cells express CXC chemokine receptor 5, localize to B cell follicles, and support immunoglobulin production. *Journal of Experimental Medicine* 192, 11.
- Butler, N.S., Moebius, J., Pewe, L., *et al.*, 2012. Therapeutic blockade of PD-L1 and LAG-3 rapidly clears established blood-stage plasmodium infection. *Nature Immunology* 13 (2), 188–195. doi:10.1038/ni.2180.
- Camacho, S.A., Kosco-Vilbois, M.H., Berek, C., 1998. The dynamic structure of the germinal center. *Immunology Today* no. 11, 511–514.
- Cannons, J.L., Qi, H., Lu, K.T., *et al.*, 2011. Optimal germinal center responses require a multi-stage T:B cell adhesion process involving integrins, SLAM-associated protein and CD84. *Immunity* 32 (2), 253–265. doi:10.1016/j.immuni.2010.01.010.Optimal.
- Caprioli, F., Sarra, M., Caruso, R., *et al.*, 2008. Autocrine regulation of IL-21 production in human T lymphocytes. *Journal of Immunology*. doi:10.4049/jimmunol.180.3.1800.
- Choi, Y.S., Danelle, Eto, Yang, J., Lao, C., Crotty, S., 2013a. Cutting edge: STAT1 is required for IL-6-mediated Bcl6 induction for early follicular helper cell differentiation. *Journal of Immunology* (Baltimore, MD: 1950) 190 (7), 3049–3053. doi:10.4049/jimmunol.1203032.
- Choi, Y.S., Kageyama, R., Eto, D., 2011. Bcl6 dependent T follicular helper cell differentiation diverges from effector cell differentiation during priming and depends on the gene ICOS. *Immunity* 34 (6), 932–946. doi:10.1016/j.immuni.2011.03.023.Bcl6.
- Choi, Y.S., Yang, J.A., Yusuf, I., *et al.*, 2013b. Bcl6 expressing follicular helper CD4 T cells are fate committed early and have the capacity to form memory. *Journal of Immunology* 190 (8), 4014–4026. doi:10.4049/jimmunol.1202963.Bcl6.
- Chung, Y., Tanaka, S., Chu, F., *et al.*, 2011. Follicular regulatory T cells expressing Foxp3 and Bcl-6 suppress germinal center reactions. *Nature Medicine* 17 (8), 983–988. doi:10.1038/nm.2426.
- Crotty, S., 2014. T follicular helper cell differentiation, function, and roles in disease. *Immunity* 41 (4), 529–542. doi:10.1016/j.immuni.2014.10.004.
- Deenick, E.K., Chan, A., Ma, C.S., Gatto, D., 2010. Follicular helper T cell differentiation requires continuous antigen presentation that is independent of unique B cell signaling. *Immunity* 33 (2), 241–253. doi:10.1016/j.immuni.2010.07.015.Follicular.
- Deenick, E.K., Ma, C.S., 2011. The regulation and role of T follicular helper cells in immunity. *Immunology* 134 (4), 361–367. doi:10.1111/j.1365-2567.2011.03487.x.
- Duan, X., Chen, X., Liang, Z., *et al.*, 2014. Genetic polymorphisms of CXCR5 and CXCL13 are associated with non-responsiveness to the hepatitis B vaccine. *Vaccine* 32 (41), 5316–5322. Elsevier Ltd. doi:10.1016/j.vaccine.2014.07.064.
- Eaton, S.M., Burns, E.M., Kusser, K., Randall, T.D., Haynes, L., 2004. Age-related defects in CD4 T cell cognate helper function lead to reductions in humoral responses. *Journal of Experimental Medicine* 200 (12), 1613–1622. doi:10.1084/jem.20041395.
- Elgueta, R., Benson, M.J., de Vries, V.C., *et al.*, 2009. Molecular mechanism and function of CD40/CD40L engagement in the immune system. *Immunological Reviews* 229 (1), 152–172. doi:10.1111/j.1600-065X.2009.00782.x.
- Eto, D., Lao, C., DiToro, D., *et al.*, 2011. IL-21 and IL-6 are critical for different aspects of B cell immunity and redundantly induce optimal follicular helper CD4 T cell (T_{fh}) differentiation. *PLoS One* 6 (3), e17739. doi:10.1371/journal.pone.0017739.
- Fazilleau, N., Eisenbraun, M.D., Malherbe, L., *et al.*, 2007. Lymphoid reservoirs of antigen-specific memory T helper cells. *Nature Immunology* 8 (7), 753–761. doi:10.1038/ni1472.
- Foy, T.M., Aruffo, A., Bajorath, J., Buhlmann, J.E., Noelle, R.J., 1996. Immune regulation by Cd40 and its ligand Gp39. *Annual Review of Immunology* 14 (1), 591–617. doi:10.1146/annurev.immunol.14.1.591.
- Gigoux, M., Shang, J., Pak, Y., *et al.*, 2009. Inducible costimulator promotes helper T-cell differentiation through phosphoinositide 3-kinase. *Proceedings of the National Academy of Sciences of the United States of America* 106 (48), 20371–20376. doi:10.1073/pnas.0911573106.
- Gitlin, A.D., Shulman, Z., Nussenzweig, M.C., 2014. Clonal selection in the germinal centre by regulated proliferation and hypermutation. *Nature* 509 (7502), 637–640. doi:10.1038/nature13300.
- Good-Jacobson, K.L., Szumilas, C.G., Chen, L., *et al.*, 2010. PD-1 regulates germinal center B cell survival and the formation and affinity of long-lived plasma cells. *Nature Immunology* 11 (6), 535–542. doi:10.1038/ni.1877.
- Goronzy, J.J., Weyand, C.M., 2013. Understanding immunosenescence to improve responses to vaccines. *Nature Immunology* 14 (5), 428–436.
- Grakoui, A., Bromley, S.K., Sumen, C., *et al.*, 1999. The immunological synapse: A molecular machine controlling T cell activation. *Science* 285 (July), 221–227.
- Gu-trantien, C., Loi, S., Garaud, S., *et al.*, 2013. CD4⁺ follicular helper T cell infiltration predicts breast cancer survival. *Journal of Clinical Investigation* 123 (7), 2873–2892. doi:10.1172/JCI67428.Traditionally.
- Hale, J.S., Ben Youngblood, D.R., Latner, L., *et al.*, 2013. Distinct memory CD4⁺ T cells with commitment to t follicular helper- and T helper 1-cell lineages are generated after acute viral infection. *Immunity* 38 (4), 805–817. doi:10.1016/j.immuni.2013.02.020.Distinct.
- Haynes, N.M., Allen, C.D.C., Lesley, R., *et al.*, 2007. Role of CXCR5 and CCR7 in follicular Th cell positioning and appearance of a programmed cell death gene-high germinal center-associated subpopulation. *Journal of Immunology* 179 (8), 5099–5108. doi:10.4049/jimmunol.179.8.5099.
- He, Jing, Tsai, L.M., Leong, Y.A., *et al.*, 2013. Circulating precursor CCR7(lo)PD-1 (hi) CXCR5⁺ CD4⁺ T cells indicate T_{fh} cell activity and promote antibody responses upon antigen reexposure. *Immunity* 39 (4), 770–781. doi:10.1016/j.immuni.2013.09.007.
- Ise, W., Kohyama, M., Schraml, B.U., *et al.*, 2011. The transcription factor BATF controls the global regulators of class-switch recombination in both B cells and T cells. *Nature Immunology* 12 (6), 536–543. doi:10.1038/ni.2037.
- Johnston, R.J., Choi, Y.S., Diamond, J.A., Yang, J.A., Crotty, S., 2012. STAT5 is a potent negative regulator of TFH cell differentiation. *Journal of Experimental Medicine* 209 (2), 243–250.
- Johnston, R.J., Poholek, A.C., DiToro, D., *et al.*, 2009. Bcl6 and Blimp-1 are reciprocal and antagonistic regulators of T follicular helper cell differentiation. *Science* (New York, NY) 325 (5943), 1006–1010. doi:10.1126/science.1175870.
- Kageyama, R., Cannons, J.L., Zhao, F., *et al.*, 2012. The receptor Ly108 functions as a SAP adaptor dependent on-off switch for T cell help to B cells and NKT cell development. *Immunity* 36 (6), 986–1002. doi:10.1016/j.immuni.2012.05.016.
- Kato, L.M., Kawamoto, S., Maruya, M., Fagarasan, S., 2014. The role of the adaptive immune system in regulation of gut microbiota. *Immunological Reviews* 260 (1), 67–75. doi:10.1111/immr.12185.
- Kawamoto, S., Maruya, M., Kato, L.M., *et al.*, 2014. Foxp3 + T cells regulate immunoglobulin A selection and facilitate diversification of bacterial species responsible for immune homeostasis. *Immunity* 41 (1), 152–165. doi:10.1016/j.immuni.2014.05.016.
- Kawamoto, S., Tran, T.H., Maruya, M., Suzuki, K., 2012. The inhibitory receptor PD-1 regulates IgA selection and bacterial composition in the gut. *Science* 321 (April), 485–490.
- Kerfoot, S.M., Yaari, G., Patel, J.R., 2011. Germinal center B cell and T follicular helper cell development initiates in the interfollicular zone. *Immunity* 34 (6), 947–960. doi:10.1016/j.immuni.2011.03.024.Germinal.
- Kim, J.I., I-Cheng Ho, Grusby, M.J., Glimcher, L.H., 1999. The transcription factor c-Maf controls the production of interleukin-4 but not other Th2 cytokines. *Immunity* 10 (6), 745–751. doi:10.1016/S1074-7613(00)80073-4.
- Kroenke, M., Eto, D., Locci, M., *et al.*, 2012. Bcl6 and Maf cooperate to instruct human follicular helper CD4 T cell differentiation. *Journal of Immunology* 188 (8), 3734–3744. doi:10.4049/jimmunol.1103246.
- Lefebvre, J.S., Maue, A.C., Eaton, S.M., *et al.*, 2012. The aged microenvironment contributes to the age-related functional defects of CD4 T cells in mice. *Aging Cell* 11 (5), 732–740. doi:10.1111/j.1474-9726.2012.00836.x.
- Liang, L., Porter, E.M., Sha, W.C., 2002. Constitutive expression of the B7h ligand for inducible costimulator on naive B cells is extinguished after activation by distinct B cell receptor and interleukin 4 receptor-mediated pathways and can be rescued by CD40 signaling. *Journal of Experimental Medicine* 196 (1), 97–108. doi:10.1084/jem.20020298.
- Linterman, M.A., Denton, A.E., Divekar, D.P., 2014. CD28 expression is required after T cell priming for helper T cell responses and protective immunity to infection. *eLife*.
- Linterman, M., Beaton, L., Yu, D., *et al.*, 2010. IL-21 acts directly on B cells to regulate Bcl-6 expression and germinal center responses. *Journal of Experimental Medicine* 207 (2), 353–363. doi:10.1084/jem.20091738.

- Linterman, M., Pierson, W., Lee, S.K., *et al.*, 2011. Foxp3⁺ follicular regulatory T cells control the germinal center response. *Nature Medicine* 17 (8), 975–982. doi:10.1038/nm.2425.
- Linterman, M., Rigby, R.J., Wong, R.K., *et al.*, 2009. Follicular helper T cells are required for systemic autoimmunity. *Journal of Experimental Medicine* 206 (3), 561–576. doi:10.1084/jem.20081886.
- Liu, X., Chen, X., Zhong, B., *et al.*, 2014. Transcription factor achaete-scute homologue 2 initiates follicular t-helper-cell development. *Nature* 507 (7493), 513–518. doi:10.1038/nature12910.
- Locci, M., Havenar-Daughton, C., Landais, E., *et al.*, 2013. Human CXCR3-CXCR5⁺ memory T_H cells are highly functional and correlate with broadly neutralizing HIV antibody responses. *Immunity* 39 (4), 758–769. doi:10.1016/j.immuni.2013.08.031.
- Ma, C.S., Avery, D.T., Chan, A., *et al.*, 2012. Functional STAT3 deficiency compromises the generation of human T follicular helper cells. *Blood* 119 (17), 3997–4008. doi:10.1182/blood-2011-11-392985.
- Ma, C.S., Suryani, S., Avery, D.T., *et al.*, 2009. Early commitment of naïve human CD4⁺ T cells to the T follicular helper (T_{FH}) cell lineage is induced by IL-12. *Immunology and Cell Biology* 87 (8), 590–600. doi:10.1038/icb.2009.64.
- MacLennan, I.C.M., Toellner, K.-M., Cunningham, A.F., *et al.*, 2003. Extrafollicular antibody responses. *Immunological Reviews* 194 (1), 8–18. doi:10.1034/j.1600-065X.2003.00058.x.
- Miller, M.J., Wei, S.H., Parker, I., Cahalan, M.D., 2002. Two-photon imaging of lymphocyte motility and antigen response in intact lymph node. *Science* (New York, NY) 296 (5574), 1869–1873. doi:10.1126/science.1070051.
- Moreau, H.D., Bousso, P., 2014. Visualizing how T cells collect activation signals *in Vivo*. *Current Opinion in Immunology* 26 (February), 56–62. doi:10.1016/j.coi.2013.10.013.
- Morita, R., Schmitt, N., Bentebibel, S.E., 2011. Human blood CXCR5⁺ CD4⁺ T cells are counterparts of T follicular cells and contain specific subsets that differentially support antibody secretion. *Immunity* 34 (1), 108–121. doi:10.1016/j.immuni.2010.12.012.Human.
- Mosmann, T.R., Chenwinski, H., Bond, M.W., Giedlin, M.A., Coffman, R.L., 1986. Two types of murine helper T cell clone. I. definition according to profiles of lymphokine activities and secreted proteins. *Journal of Immunology* 136 (7), 2348–2357.
- Noia, J.M. D., Neuberger, M.S., 2007. Molecular mechanisms of antibody somatic hypermutation. *Annual Review of Biochemistry* 76 (January), 1–22. doi:10.1146/annurev.biochem.76.061705.090740.
- Nurieva, R.I., Chung, Y., Hwang, D., *et al.*, 2008. Generation of follicular helper T cells is mediated by IL-21 but independent of TH1, TH2 or TH17 lineages. *Immunity* 29 (1), 138–149. doi:10.1016/j.immuni.2008.05.009.Generation.
- Nurieva, R.I., Chung, Y., Martinez, G.J., *et al.*, 2009. Bcl6 mediates the development of T follicular helper cells. *Science* (New York, NY) 325 (5943), 1001–1005. doi:10.1126/science.1176676.
- Nurieva, R., Yang, X.O., Martinez, G., *et al.*, 2007. Essential autocrine regulation by IL-21 in the generation of inflammatory T cells. *Nature* 448 (7152), 480–483. doi:10.1038/nature05969.
- Okada, T., Miller, M.J., Parker, I., *et al.*, 2005. Antigen-engaged B cells undergo chemotaxis toward the T Zone and form motile conjugates with helper T cells. *PLoS Biology* 3 (6), e150. doi:10.1371/journal.pbio.0030150.
- Pallikuth, S., Parmigiani, A., Silva, S.Y., *et al.*, 2012. Impaired peripheral blood T-follicular helper cell function in HIV-infected nonresponders to the 2009 H1N1/09 vaccine. *Blood* 120 (5), 985–993. doi:10.1182/blood-2011-12-396648.
- Pepper, M., Pagán, A.J., Igyártó, B.Z., Taylor, J.J., Jenkins, M.K., 2011. Opposing signals from the Bcl6 transcription factor and the interleukin-2 receptor generate T helper-1 central and effector memory cells. *Immunity* 35 (4), 583–595. doi:10.1016/j.immuni.2011.09.009.Opposing.
- Pereira, J.P., Kelly, L.M., Xu, Y., Cyster, J.G., 2009. EB12 mediates B cell segregation between the outer and centre follicle. *Nature* 460 (7259), 1122–1126. doi:10.1038/nature08226.
- Qi, H., Cannons, J.L., Klauschen, F., Schwartzberg, P.L., Germain, R.N., 2008. SAP-controlled T-B cell interactions underlie germinal centre formation. *Nature* 455 (7214), 764–769. doi:10.1038/nature07345.
- Qi, H., Chen, X., Chu, C., *et al.*, 2014. Follicular T-helper cells: Controlled localization and cellular interactions. *Immunology and Cell Biology* 92 (1), 28–33. Nature Publishing Group. doi:10.1038/icb.2013.59.
- Reinhardt, R.L., Liang, H.-E., Locksley, R.M., 2009. Cytokine-secreting follicular T cells shape the antibody repertoire. *Nature Immunology* 10 (4), 385–393. doi:10.1038/ni.1715.
- Rolf, J., Bell, S.E., Kovacs, D., *et al.*, 2010. Phosphoinositide 3-kinase activity in T cells regulates the magnitude of the germinal center reaction. *Journal of Immunology* 185 (7), 4042–4052. doi:10.4049/jimmunol.1001730.
- Schmitt, N., Liu, Y., Bentebibel, S.-E., *et al.*, 2014. The cytokine TGF- β co-opts signaling via STAT3-STAT4 to promote the differentiation of human TFH cells. *Nature Immunology* 15 (9), 856–865. doi:10.1038/ni.2947.
- Sharpe, A.H., Freeman, G.J., 2002. The B7–CD28 superfamily. *Nature Reviews Immunology* 2 (2), 116–126. doi:10.1038/nri727.
- Simpson, N., Gatenby, P., Wilson, A., *et al.*, 2010a. Expansion of circulating T cells resembling follicular helper T cells is a fixed phenotype that identifies a subset of severe systemic lupus erythematosus. *Arthritis and Rheumatism* 62 (1), 234–244. doi:10.1002/art.25032.
- Simpson, T.R., Quezada, S., Allison, J.P., 2010b. Regulation of CD4 T cell activation and effector function by inducible costimulator (ICOS). *Current Opinion in Immunology* 22 (3), 326–332. doi:10.1016/j.coi.2010.01.001.
- Stiasny, K., Aberle, J.H., Keller, M., Grubeck-Loebenstein, B., Heinz, F.X., 2012. Age affects quantity but not quality of antibody responses after vaccination with an inactivated flavivirus vaccine against tick-borne encephalitis. *PLoS One* 7 (3), e34145. doi:10.1371/journal.pone.0034145.
- Szabo, K., Papp, G., Barath, S., *et al.*, 2013. Follicular helper T cells may play an important role in the severity of primary sjögren's syndrome. *Clinical Immunology (Orlando, FL)* 147 (2), 95–104. doi:10.1016/j.clim.2013.02.024.
- Tarlington, D.M., 2008. Evolution in miniature: selection, survival and distribution of antigen reactive cells in the germinal centre. *Immunology and Cell Biology* 86 (2), 133–138. doi:10.1038/sj.icb.7100148.
- Victoria, G.D., Mesin, L., 2014. Clonal and cellular dynamics in germinal centers. *Current Opinion in Immunology* 28 (June), 90–96. doi:10.1016/j.coi.2014.02.010.
- Victoria, G.D., Schwickert, T.A., Fooksman, D.R., *et al.*, 2010. Germinal center dynamics revealed by multiphoton microscopy using a photoactivatable fluorescent reporter. *Cell* 143 (4), 592–605. doi:10.1016/j.cell.2010.10.032.Germinal.
- Vogelzang, A., McGuire, H.M., Yu, D., *et al.*, 2008. A fundamental role for interleukin-21 in the generation of T follicular helper cells. *Immunity* 29 (1), 127–137. doi:10.1016/j.immuni.2008.06.001.
- Wollenberg, I., Agua-Doce, A., Hernández, A., *et al.*, 2011. Regulation of the germinal center reaction by Foxp3⁺ follicular regulatory T cells. *Journal of Immunology (Baltimore, MD: 1950)* 187 (9), 4553–4560. doi:10.4049/jimmunol.1101328.
- Xu, H., Li, X., Liu, D., *et al.*, 2013. Follicular T-helper cell recruitment governed by bystander B cells and ICOS-driven motility. *Nature* 496 (7446), 523–527. doi:10.1038/nature12058.
- Yu, D., Rao, S., Tsai, L.M., *et al.*, 2009. The transcriptional repressor Bcl-6 directs T follicular helper cell lineage commitment. *Immunity* 31 (3), 457–468. doi:10.1016/j.immuni.2009.07.002.
- Yusuf, I., Kageyama, R., Monticelli, L., *et al.*, 2010. Germinal center T follicular helper cell IL-4 production is dependent on signaling lymphocytic activation molecule receptor (CD150). *Journal of Immunology* 185 (1), 190–202. doi:10.4049/jimmunol.0903505.
- Zotos, D., Coquet, J.M., Zhang, Y., *et al.*, 2010. IL-21 regulates germinal center B cell differentiation and proliferation through a B cell-intrinsic mechanism. *Journal of Experimental Medicine* 207 (2), 365–378.

V(D)J Recombination: Orchestrating Diversity without Damage

C Lescale and L Deriano, Institut Pasteur, Paris, France

© 2016 Elsevier Inc. All rights reserved.

"RAGs are the end-point of the strategies developed through evolution to achieve the anticipatory "Promethean" recognition of all universal molecular forms." – A.A. de Freitas, Tractus Immuno-Logicus

Introduction

How does the organism cope with all the various infectious agents that may attack it during a lifetime? The answer resides in the ability of the B and T lymphocytes of the adaptive arm of the immune system to express a vast enough repertoire of antigen-binding molecules (Immunoglobulin – Ig and T-cell receptor – TCR) that any pathogen will face an immune response. Each Ig or TCR polypeptide consists of a constant region that serves structural, signaling, and effector functions, and a variable region that determines antigen recognition and specificity. Ig can be secreted or expressed as a transmembrane surface receptor on B cells (B-cell receptor – BCR). Ig and TCR are heterodimeric proteins with the Ig composed of dimers of an Ig heavy chain (IgH) combined with either a κ or λ light chain, and the TCR composed of $\alpha\beta$ or $\gamma\delta$ dimers (Figure 1).

In vertebrates, including humans and mice, antigen receptor diversity is generated in immature lymphocytes and depends on a process of somatic DNA rearrangement termed V(D)J recombination, a cut and paste mechanism which shuffles a large array of preexisting variable (V), diversity (D), and joining (J) gene fragments to create functional genes encoding for the variable regions of Ig and TCR proteins. This combinatorial approach generates an almost unlimited repertoire of discrete receptor molecules ($\sim 1 \times 10^7$ – without taking into account junctional diversity)

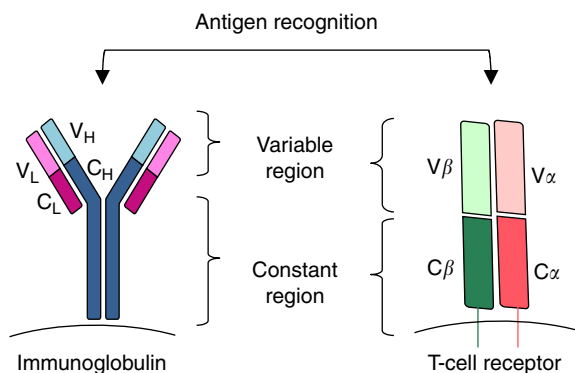


Figure 1 Structure of the antigen receptors. Protein structure of immunoglobulin/BCR and TCR; each polypeptide consists of a constant region (C) that serves structural, signaling, and effector functions, and a variable region (V) that determines antigen recognition and specificity. Immunoglobulins can be produced as either surface or secreted receptor molecules.

from a relatively modest investment in germline coding capacity (in mammals, a few hundred V segments recombining approximately a hundred D and/or J segments). Mammals have seven distinct genomic loci undergoing somatic recombination in a lineage-, stage-, and allele-specific manner: the *IgH*, light κ (*IgLκ*), and λ (*IgLλ*) chain genes in B lymphocytes; and the *Tcr α*, β , γ , and δ chains in T lymphocytes (Figure 2).

Discrete V, D, and J segments are flanked by recombination signal sequences (RSS) that are recognized by a lymphoid-specific recombinase (also called V(D)J recombinase), which consists of the protein products of the recombination activating genes *Rag1* and *Rag2*. RAG proteins constitute a site-specific endonuclease that introduces DNA double-strand breaks (DSBs) adjacent to RSSs. Cleavage at a pair of RSSs (e.g., one adjacent to a V segment and one adjoining a J segment) generates four broken DNA ends: two covalently sealed (hairpin) coding ends, which terminate in the coding segment, and two blunt signal ends, which terminate in the RSS. These DNA ends activate a general DNA damage response (DDR) and are joined by the ubiquitous non-homologous end-joining (NHEJ) proteins to form a coding joint that encodes the variable portion of the antigen receptor protein and a reciprocal product termed a signal joint that is usually excised from the chromosome. Coding joints are typically imprecise, as the coding-end hairpins undergo processing with gain or loss of nucleotides and, subsequent to this, the lymphoid-specific enzyme, terminal deoxynucleotidyl transferase (TdT), adds non-templated nucleotides. This junctional variability contributes further to antigen receptor diversity and is considered characteristic of repair by NHEJ.

The process of V(D)J rearrangement relies on repeated cleavage and rejoining of chromosomal gene segments before a complete Ig/BCR or TCR molecule can be expressed at the cell surface, thus endangering the integrity of the genome, if not the whole organism. As millions of lymphocytes undergo antigen receptor rearrangement every day, there are many opportunities for genome integrity to be ruptured over the course of an individual's lifetime. It is therefore not surprising that multiple restriction mechanisms, often redundant, have also emerged to carefully control when, where, and how recombination occurs (Table 1).

In this article, we will first present the developmental (lymphocyte differentiation) and epigenetic (locus accessibility and substrate selection) aspects of V(D)J recombination regulation. We will then present recent and earlier work that has shed light on the molecular mechanisms of RAG-mediated V(D)J recombination, with an emphasis on the intricate roles of the recombinase (RAG1 and RAG2), DDR and DNA repair proteins. Finally, we will discuss V(D)J recombination in the context of genome maintenance and malignancies of the immune system.

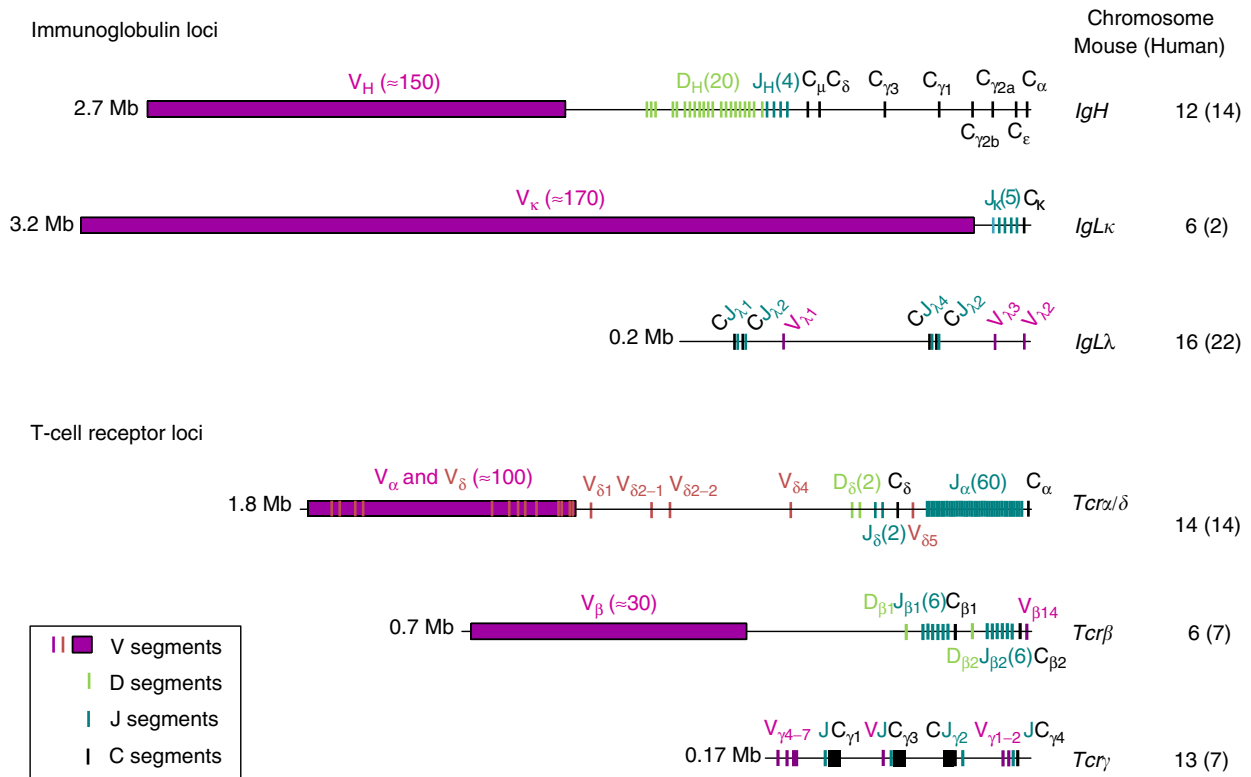


Figure 2 Structure of the antigen receptor genes. Germline organization of antigen receptor loci; all loci contain V and J segments, linked to a constant region (C). The *IgH* and *Tcr β* and δ loci also comprise D segments between the V's and J's. Number of V, D, and J gene segments is indicated in brackets. Genomic length of each locus in megabases is indicated on the left. Adapted from Johnson, K., Chaumeil, J., Skok, J.A., 2010. Epigenetic regulation of V(D)J recombination. *Essays in Biochemistry* 48, 221–243.

Table 1 Levels of restriction for V(D)J recombination

Restriction	Description
Lineage specificity	T-cell receptors are assembled only in T cells, and Ig assembly is restricted to B cells
Temporal-ordered rearrangement	V(D)J recombination follows a precise order: • Between the loci: in B cells, the <i>IgH</i> locus is rearranged before <i>IgL</i> (κ or λ); in T cells, <i>Tcrβ</i> is rearranged before <i>Tcrα</i> • Within each locus: at the <i>IgH</i> and <i>Tcrβ</i> loci, D to J rearrangement precedes V-to-DJ.
Allelic exclusion	At most loci (<i>IgH</i> , <i>Igκ</i> , <i>Igλ</i> , <i>Tcrβ</i> , and <i>Tcrγ</i>), only one allele can be productively rearranged, ensuring that each lymphocyte expresses one receptor molecule with unique antigen-binding properties
RAG expression	RAG1 and RAG2 are co-expressed only in lymphoid cells and only at specific stages in lymphocyte development. RAG expression is cell cycle regulated, with highest levels in G1
Substrate selection	The <i>Tcr</i> and <i>Ig</i> V, D, and J coding segments are flanked by recombination signal sequences (RSS) that are recognized by the RAG proteins. RSS comprise a heptamer sequence, a nonamer sequence, and an intervening spacer of either 12 (12-RSS) or 23 nucleotides (23-RSS). Efficient recombination is only permitted between paired gene segments flanked by a 12- and a 23-RSS (12/23 rule)
Repair pathway choice	RAG DNA ends are repaired exclusively by the nonhomologous end-joining (NHEJ) pathway, which is mostly operative in G0/G1 stage of the cell cycle, when V(D)J recombination occurs. This restriction is also assured by the RAG proteins themselves, which handoff DNA ends from the post-cleavage complex to the NHEJ machinery

Source: Adapted from Johnson, K., Chaumeil, J., Skok, J.A., 2010. Epigenetic regulation of V(D)J recombination. *Essays in Biochemistry* 48, 221–243.

Developmental and Epigenetic Regulation of V(D)J Recombination

In adult mammals, B and T lymphocytes are continuously generated from bone marrow hematopoietic stem cells (HSC). While B lymphocytes develop exclusively in the bone marrow,

T lymphocytes undergo most of their development in the thymus (Figure 3). During lymphopoiesis, multipotent hematopoietic progenitors undergo progressive restriction of lineage potential, resulting in B and T lymphoid lineage commitment (Nutt and Kee, 2007; Singh and Pongubala, 2006). To complete their development, all lymphocytes must

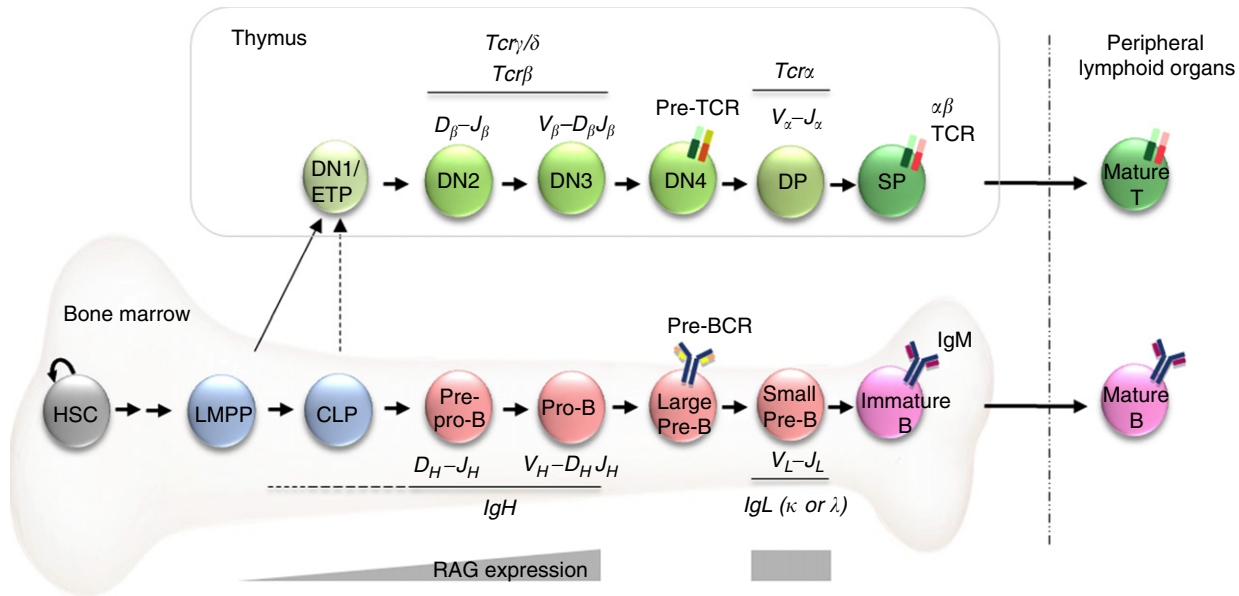


Figure 3 Lymphocyte development and V(D)J recombination. RAG expression and rearrangement of V, D, and J gene segments occur at specific stages of lymphocyte development. During B-cell development (lower panel), the *IgH* locus is rearranged at the pre-pro-B and pro-B cell stage and then the *IgL* (κ or λ) in pre-B cells. During T-cell development (upper panel), the *Tcrβ* and *Tcrγ/δ* loci are rearranged at the DN2 and DN3 stages, before *Tcrα* in DP cells. For both *IgH* and *Tcrβ* loci, D to J rearrangement occurs first and is followed by V to DJ. CLP, Common lymphoid progenitor; DN, CD4⁻ and CD8⁻ double negative; DP, CD4⁺ and CD8⁺ double positive; HSC, hematopoietic stem cell; LMPP, lymphoid primed progenitor; SP, single positive (either CD4⁺ or CD8⁺).

assemble a functional antigen receptor (a BCR or Ig for B lymphocytes and a TCR for T lymphocytes). Therefore, V(D)J recombination represents a defining feature of B- and T-cell development (Figures 1 and 3). Importantly, lineage restriction ensures that, although the recombinase (i.e., the RAG1 and RAG2 proteins) is expressed in both cell lineages, *Tcr* genes fully recombine only in T cells and *Ig* genes only in B cells. Ultimately, each lymphocyte expresses a receptor bearing a unique antigen specificity (i.e., monospecificity), which is achieved by limiting the number of functional Ig or TCR alleles to one per cell, a process known as allelic exclusion.

It is important to note that mature B cells, after being stimulated with a cognate antigen, undergo two additional antigen receptor diversification reactions in secondary peripheral lymphoid organs. These two processes are beyond the scope of this article and will therefore not be detailed here, but briefly, class switch recombination (CSR) and somatic hypermutation (SHM), are initiated by the lymphoid-specific enzyme activation-induced cytidine deaminase (AID). SHM introduces point mutations into *IgH* and *IgL* variable region exons, allowing selection of B cells that produce high-affinity antibodies. CSR exchanges the constant region of the *IgH* chain, thereby producing antibodies with distinct effector classes and functions (Dudley et al., 2005).

Ordered Rearrangement of TCR and BCR Gene Segments

In both B and T-cell lineages, V(D)J recombination occurs in a strict temporally ordered manner, with D to J recombination occurring prior to V to DJ rearrangement (Figure 3). B lymphopoiesis depends on the sequential productive (i.e., in frame and translated into a protein) DNA rearrangement of

the *Ig* heavy (*IgH*) chain locus and the *Ig* light chain loci (*IgLκ* followed by *IgLλ*), and their subsequent expression and assembly into BCRs. Rearrangement of the *IgH* locus involves recombination of *D_H* to *J_H* gene segments in uncommitted common lymphoid progenitors (CLP) and pre-pro-B cells, followed by *V_H* to *DJ_H* gene segments in B-cell lineage committed pro-B cells (Borghesi et al., 2004; Hardy et al., 1991). Productive *V_H*-*D_H*-*J_H* rearrangement leads to the expression of the *IgH* chain and its assembly with the surrogate light chain, comprised of the $\lambda 5$ and *VpreB* proteins, and the signaling subunits *Igα* and *Igβ* to form a pre-BCR (Figure 3). The pre-BCR promotes the expansion of a population of large pre-B cells and serves as a critical checkpoint that ensures the expression of a signaling-competent *IgH* chain before rearrangement of the *IgL* chain V and J gene segments in small resting pre-B cells (Figure 3). Successful *V_L* to *J_L* rearrangement results in the expression of either an *IgLκ* or *IgLλ* chain that pairs with the heavy chain to form a functional Ig (Figures 1 and 3).

T lymphocytes undergo an analogous developmental program to produce TCR $\alpha\beta$ or $\gamma\delta$ thymocytes with the majority of mature T cells expressing a $\alpha\beta$ TCR (~95–99% in mice) (Figure 3; Krangel, 2009; Rothenberg and Taghon, 2005). The *Tcrβ*, γ , and δ loci all recombine in immature T cells that are CD4⁻CD8⁻ double negative (DN), with D to J rearrangement occurring prior to V to DJ rearrangement. Successful recombination of *Tcrγ* and γ loci leads to the expression of a $\gamma\delta$ TCR, whereas successful recombination of *Tcrβ* promotes assembly of a pre-TCR, composed of a TCR β chain paired with an invariant pre-TCR α chain. Pre-TCR signaling induces several rounds of proliferation and differentiation of $\alpha\beta$ -lineage cells to the CD4⁺CD8⁺ double positive (DP) cell stage, during

which the *Tcr α* locus is rearranged (Figure 3). At the DP stage, $\alpha\beta$ T cells must successfully undergo positive and negative selection before they can leave the thymus. $\alpha\beta$ TCRs able to bind major histocompatibility complex (MHC) class I or II molecules with at least a weak affinity will provide survival signals to the cell (positive selection) while $\alpha\beta$ TCRs that bind with high affinity to self peptides and/or MHC molecules will provide apoptotic signals to the cell (negative selection). Positively selected $\alpha\beta$ TCR⁺ T cells will exit the thymus as mature single positive (SP) CD4⁺ or CD8⁺ T cells (Figure 3).

Expression of the V(D)J Recombinase

Transcriptional control

RAG expression is restricted to B and T lymphocytes although a small fraction of early hematopoietic progenitors express RAG mRNA (Mansson *et al.*, 2007; Igarashi *et al.*, 2002). In both lymphoid lineages, RAG expression occurs in two waves that correlate with the ordered rearrangement of antigen receptor loci (Figure 3). The first wave results in the assembly of the *IgH* or *Tcr β* chain genes, in pro-B or pro-(DN3) T cells, respectively. Subsequently, the expression of a pre-BCR or a pre-TCR at the cell surface induces a proliferative burst that is associated with the downregulation of RAG expression in cycling large pre-B and pre-(DN4) T cells. Pre-B and DN4 T cells then exit the cell cycle and activate a second wave of RAG

expression leading to the assembly of *IgL* or *Tcr α* chain genes, respectively (Kuo and Schlissel, 2009; Hillion *et al.*, 2009; Figure 3). Importantly, regulation of proliferation versus RAG reexpression and light chain assembly seems to involve a complex interplay between cytokine receptor and pre-BCR or pre-TCR signaling pathways (Clark *et al.*, 2014). Immature pre-B cells continue to transcribe the *Rag* genes if they express a nonfunctional or an autoreactive antigen receptor leading to receptor editing (Figure 4 and below). Immature DP T cells continue to express RAGs and rearrange the *Tcr α* locus until they undergo positive selection and transit to the SP stage.

Many of the transcription factors initiating commitment to the B and T-cell lineages have also been implicated in the regulation of *Rag* transcription, including, for example, E2A, Ikaros, Pax5, Foxp1, FoxO1, and NF- κ B in B cells. It is thought that the aforementioned transcription factors regulate RAG expression by binding to the RAG1 and RAG2 promoters and/or other *Rag* regulatory elements. Interestingly, deletion of certain *cis*-regulatory elements leads to different RAG expression levels in T and B cells, suggesting that each cell type regulates the *Rag* locus using different transcription factors and *cis*-regulatory elements (Kuo and Schlissel, 2009).

Posttranscriptional control

An additional level of regulation of the V(D)J recombinase is at the level of protein stability (Desiderio, 2010). In dividing

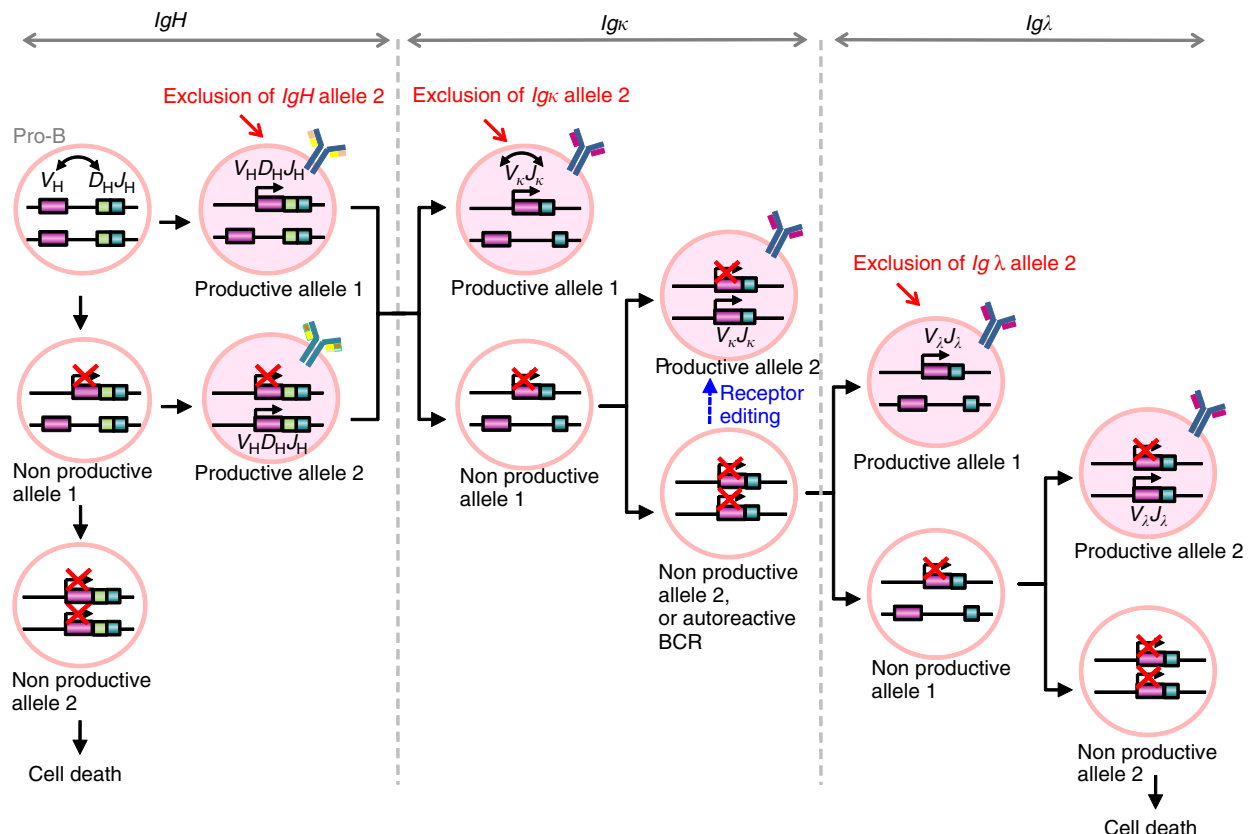


Figure 4 Allelic exclusion and V(D)J recombination. During B-cell development, D_H - J_H rearrangement occurs on both alleles while V_H - D_H - J_H , V_K - J_K , and V_L - J_L are subjected to allelic exclusion. Receptor editing can replace nonproductive or autoreactive V_K - J_K exons by reassembling upstream V_K segments with downstream J_K segments to create a different BCR. If this fails, the *IgL* is targeted for recombination.

cells, RAG2 is degraded during the G1 to S transition of the cell cycle and accumulates again upon reentry into G1, therefore limiting the RAG enzyme activity to the G0/G1 phases. It was demonstrated that the cell-cycle regulator cyclinA/CDK2, which promotes entry into S phase, phosphorylates RAG2 at residue Threonine 490, leading to association with the Skp2-SCF ubiquitin ligase and subsequent proteasomal degradation. This level of regulation restricts recombination to noncycling cells and might help ensure that RAG-mediated DNA breaks do not interfere with the replication machinery. Thymocytes from mice carrying a RAG2 hypomorphic mutation (T490A) that prevents its degradation during G1-S transition accumulate signal ends in S-G2-M and TCR-associated chromosomal translocations (Zhang *et al.*, 2011). Therefore, segregation of proliferation and V(D)J recombination is a crucial regulatory process to prevent aberrant recombination events and translocations.

Allelic Exclusion

One of the axioms of Burnet's clonal selection theory (1957) states that a B- or T-cell expresses a single 'specificity' of antigen receptor (Jung and Alt, 2004; Pernis *et al.*, 1965), and indeed, allelic exclusion ensures that this is the case (In mice, less than 1% of splenic B cells display allelic inclusion – the expression of two different IgH chains (Barreto and Cumano, 2000)). In B-cell progenitors, D_H–J_H rearrangement of the *IgH* locus occurs on both alleles but successful V_H to D_HJ_H rearrangement on one allele and expression of a pre-BCR prevents a second D_HJ_H allele from undergoing V_H to D_HJ_H recombination (Figure 4). Only in cases where rearrangement of the first *IgH* allele is nonproductive (i.e., out-of-frame rearrangement or expression of a premature stop codon), the second allele is targeted for recombination, which occurs in two-thirds of pro-B cells (Mostoslavsky *et al.*, 2004). Failure to produce a functional rearrangement on both *IgH* alleles leads to cell death (Figure 4). In pre-B cells, the *Igk* locus is rearranged preferentially prior to the *Igl* locus (although this can vary between species) and both loci are subjected to allelic exclusion. If rearrangement of the first *Igk* allele is in-frame, a BCR comprised of the heavy and light chains is expressed at the cell surface and, during a process termed negative selection, is tested for auto-reactivity (i.e., recognizing a self antigen). Autoreactive BCRs may be rescued by further rearrangement, which leads to the recombination of unused V_k and J_k segments and to the expression of a novel, potentially non-autoreactive BCR. This process is termed receptor editing and is a major contributor to the maintenance of B-cell self-tolerance. Of note, only five total J_k segments (Figure 2) are available for recombination, therefore limiting the number of secondary rearrangements that can occur in a single B-cell. In case of nonfunctional rearrangement on both *Igk* alleles, the *Igl* locus is targeted for recombination. Failure to produce non-autoreactive BCRs after both *Igk* and *Igl* ultimately leads to cell death (Figure 4).

Similarly, functional rearrangement of the *Tcrβ* locus in DN3 cells leads to the expression of a pre-TCR and inhibits further *Tcrβ* rearrangement. The *Tcrα* locus is less stringently regulated and can undergo multiple rounds of recombination (Krangel, 2009); rearrangement of V_α and J_α gene segments in DP T cells continues until the production of a TCR that is able

to recognize MHC molecules. Unlike the *Igk* and *Igl* loci, the *Tcrα* locus comprises a large array of J_α and V_α gene segments that enable multiple rounds of rearrangements (Figure 2).

The phenomenon of allelic exclusion implies the existence of mechanisms that control the initiation of V(D)J recombination on one allele only, and the exclusion of the other allele from being accessible to the recombinase. Although the precise chain of events remain unclear, theoretical models, most notably the instructive and stochastic models (not necessarily mutually exclusive), have been proposed to explain allelic exclusion during V(D)J recombination (Vettermann and Schlissel, 2010; Mostoslavsky *et al.*, 2004; Jung and Alt, 2004). Additionally, experimental evidence points to cooperation of multiple mechanisms to regulate the initiation and maintenance of allelic exclusion.

The instructive model refers to labeling of the two alleles with different epigenetic marks, such that only one will be prone or 'primed' to rearrange first. In this model, asynchronous rearrangement of the two alleles might result from the low efficiency of the recombination process due to limitations of chromatin accessibility and consequently of the RAG recombinase to access its target gene segments (see below for molecular details of accessibility).

The stochastic model postulates that the probability of productive rearrangement on both alleles in a single cell is very low. In this model, since no coordination is required between the two alleles, asynchrony of allelic recombination or feedback inhibition mechanisms are not relevant. By taking into account the low probability of rearranging one allele in the correct reading frame that encodes a pairing Ig chain, one can calculate the theoretical upper limit for *IgH* allelic inclusion in newly generated B cells to be approximately 20%. However, as aforementioned, the observed frequencies of *IgH* and *IgL* allelic inclusion among peripheral B cells do not exceed 1%. Clearly, additional mechanisms must exist to prevent bi-allelic Ig rearrangement.

Asynchronous replication timing of Ig alleles, which is established and maintained as an epigenetic mark, might function to provide a single allele accessible for initial rearrangement, a prerequisite for additional feedback regulation of allelic exclusion (Farago *et al.*, 2012; Vettermann and Schlissel, 2010; Mostoslavsky *et al.*, 2001).

One such feedback inhibition comes from the pre-BCR or pre-TCR that provides signals inhibiting further rearrangement of the *IgH* and *Tcrβ* loci. The precise signaling pathways involved in this feedback mechanism remain largely unknown but result in the reduced accessibility of the unrearranged V_H and V_β segments to the RAG recombinase (Vettermann and Schlissel, 2010; Brady *et al.*, 2010).

A more recent feedback inhibition model is based on the observation that the two *IgH* alleles are located in close proximity prior to or during the initiation of V(D)J recombination. In this model, pairing of the two alleles is dependent on the presence of the RAG proteins but independent of DNA cleavage. This physical proximity might help communication in *trans* between the two alleles. In fact, DNA cleavage on one allele promotes repositioning of the second allele to pericentromeric heterochromatin (Figure 5). Nuclear repositioning to a repressive environment is dependent on the C-terminal regulatory domain of RAG2 and

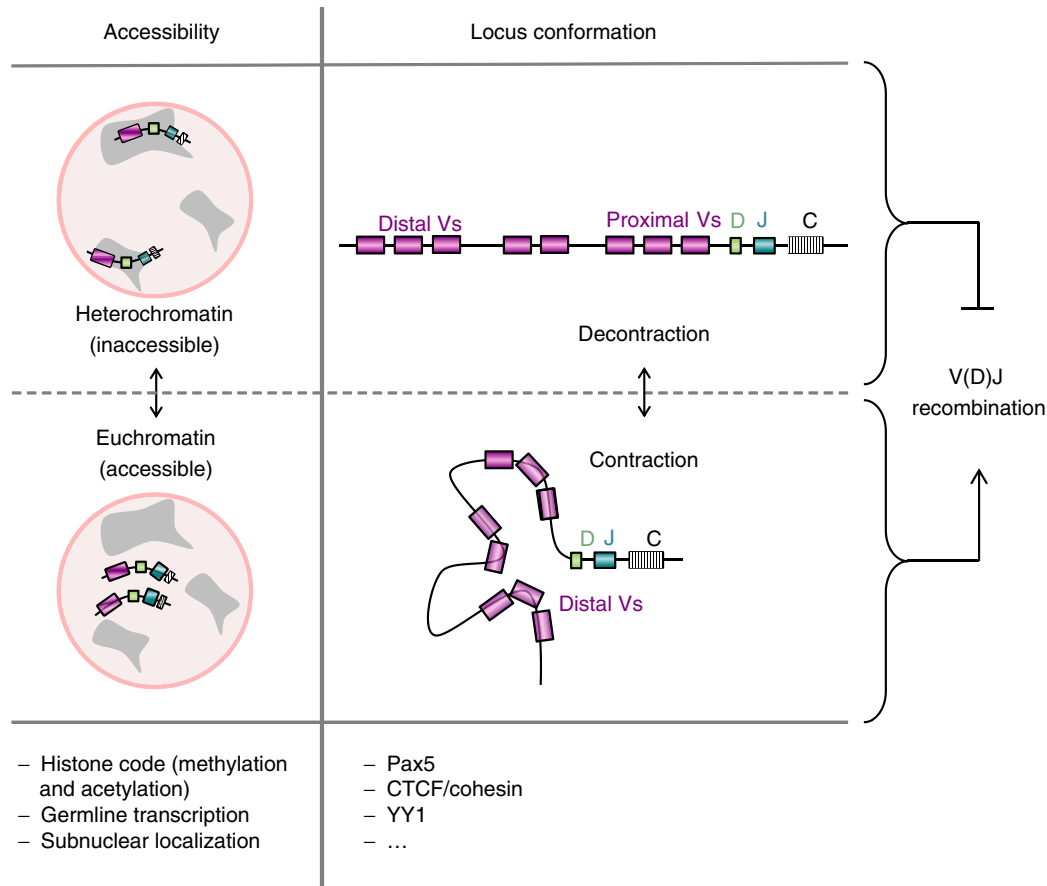


Figure 5 Locus accessibility/conformation and V(D)J recombination. Antigen receptor locus accessibility is a central process governing V(D)J recombination. It is regulated at multiple levels that include sterile transcription, subnuclear localization, and histone modifications. Moreover, locus contraction brings together distant segments, allowing the equal use of the more distal and the proximal V segments. Importantly, these changes in chromatin structure and loci conformation are dynamic/reversible.

the action of the DDR factor ataxia telangiectasia mutated (ATM) (see Section Molecular mechanisms of RAG-mediated V(D)J recombination) and likely limits access of the RAG recombinase to the un-cleaved allele during repair of the first break (Chaumeil and Skok, 2013; Chaumeil *et al.*, 2013a,b; Brandt *et al.*, 2010; Hewitt *et al.*, 2009). This mechanism inhibits bi-allelic cleavage and therefore minimizes the presence of multiple DNA breaks in the cell and the risk of genomic instability. Indeed, mice expressing a truncated version of RAG2 or deficient for ATM display increase frequency of bi-allelic breaks, which might provide a direct mechanism for the generation of Ig- and TCR-associated translocations observed in these animals. Additionally, it was recently shown that ATM-deficient mice harbor increased frequency of lymphocytes expressing IgH, IgL κ or TCR β chains from both alleles, further supporting the role of ATM in enforcing allelic exclusion (Steinel *et al.*, 2013).

Accessibility and Chromosomal Organization

How does the recombinase find its specific target sequences within antigen receptor loci during appropriate developmental windows? In the nucleus, the DNA molecule is packaged into chromatin by histones and other nuclear proteins, generating a

3D structure that can evolve both ways from a closed conformation (heterochromatin) to an open conformation (euchromatin) that is accessible to the enzymatic machineries (Figure 5). Upon B- and T-cell lineage commitment, developmental signals lead to chromatin opening that renders V(D)J gene segments accessible to the RAG recombinase. These epigenetic and structural changes are highly dynamic and regulated at multiple levels that include germline transcription (also called sterile or noncoding transcription), chromatin modifications and remodeling, *trans*-acting factors, and alterations in locus conformation (and many others of which, certainly, remain to be discovered).

Transcription and epigenetic marks

Based on the correlation between transcription and V(D)J recombination, accessibility of particular V, D, and J segments was (correctly) proposed 30 years ago as the basis for developmental stage-, lineage-, and allele-specific regulation of V(D)J recombination (Sleckman and Oltz, 2012; Jung and Alt, 2004; Yancopoulos and Alt, 1985; Table 1). Germline transcripts generated from unrearranged V, D, and J segments have since been described for all Ig and Tcr loci, and their stage- and lineage-specific expression strongly correlates with accessibility of these transcribed elements to the V(D)J

recombinase (Sleckman and Oltz, 2012; Matheson and Corcoran, 2012; Jung and Alt, 2004). Germline transcription also includes antisense intergenic transcription that occurs throughout the *IgH* locus (Matheson and Corcoran, 2012). Though the causal link between germline transcripts and recombination has been debated, targeted deletion of promoters and enhancers that drive germline transcripts abrogate recombination, and inhibiting transcription via the insertion of a transcriptional terminator in the *J α* cluster suppresses V-J rearrangement downstream of the terminator (Abarrategui and Krangel, 2006). Additionally, many features of open chromatin, including DNase I sensitivity, DNA methylation, and histone modifications, are connected to germline transcription and recombination (Jaeger *et al.*, 2013; Sleckman and Oltz, 2012; Cedar and Bergman, 2011; Abarrategui and Krangel, 2009). All of this points to the idea that recombinationally active V(D)J loci are associated with open chromatin, while inactive loci are associated with closed chromatin in non-lymphoid cells or recombinationally inactive lymphocytes. In support of this hypothesis, combined action of the transcriptional machinery with specific chromatin remodelers such as the SWI/SNF complex has been shown to modify nucleosome structure such that RSSs become accessible to the RAG recombinase (Bevington and Boyes, 2013; Osipovich *et al.*, 2007, 2009; Golding *et al.*, 1999; Kwon *et al.*, 1998). Last but not least, a direct link between chromatin, transcription and V(D)J recombination came from the discovery of a non-canonical plant homeodomain (PHD) finger in the C-terminal region of RAG2 (Figure 6). The RAG2 PHD finger specifically recognizes hypermethylated histone H3 at residue K4 (H3K4me3) and is important for efficient V(D)J recombination (Matthews *et al.*, 2007; Liu *et al.*, 2007; West *et al.*, 2005; Elkin *et al.*, 2005; Callebaut and Mornon, 1998). H3K4me3 is

tightly associated with transcriptionally active chromatin and also marks antigen receptor loci undergoing or poised to undergo recombination (Desiderio, 2010). Strikingly, genome wide ChIP-seq analyses indicate that RAG2 recruitment mirrors H3K4me3 footprints, with thousands of binding sites in the genome, in contrast to RAG1 which binds predominantly at conserved RSSs (Schatz and Ji, 2011; Ji *et al.*, 2010). Whether RAG2's ability to read the histone code (Jenuwein and Allis, 2001; Strahl and Allis, 2000) is biologically relevant outside recombination at antigen receptor loci is not known.

Chromosomal organization

In addition to RSS accessibility, nuclear location and structural reorganization of the antigen receptor loci provide another important layer of regulation for proper orchestration of V(D)J recombination. Both the peripheral nuclear lamina and regions of pericentric heterochromatin are associated with gene silencing (Towbin *et al.*, 2013). It has been observed that antigen receptor loci are located in these repressive compartments in non-lymphoid and early hematopoietic progenitors but move to euchromatin regions in lymphocytes undergoing V(D)J recombination (Hewitt *et al.*, 2009; Kosak *et al.*, 2002; Figure 5). Although the mechanisms by which antigen receptor loci associate to these repressive regions remain unknown, nuclear repositioning of the *Ig* and *Tcr* loci to repressive regions has been implicated in the initiation and maintenance of allelic exclusion (Johnson *et al.*, 2010; Hewitt *et al.*, 2009; Skok *et al.*, 2007; Roldan *et al.*, 2005).

Most antigen receptor loci encompass large genomic regions (for example, up to 3.2 Mb for the *Ig κ*), with some V segments located at greater distances than others from the D and J segments (Figure 2). The equal use of distal and proximal V segments is essential to insure a complete

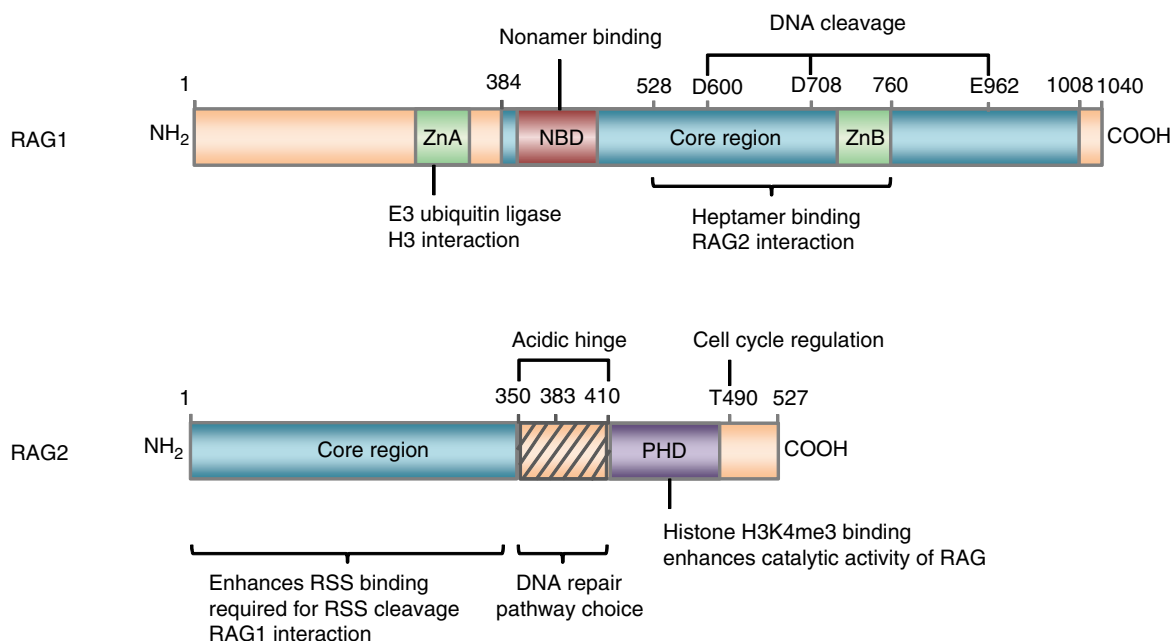


Figure 6 Structure of RAG1 and RAG2 proteins. NBD, nonamer-binding domain; PHD, plant homeodomain; Zn, zing finger domain. See text for details (Adapted from Schatz, D.G., Ji, Y., 2011. Recombination centres and the orchestration of V(D)J recombination. *Nature Reviews Immunology* 11, 251–263.).

diversification of antigen receptors. This implies that a mechanism must exist to physically bring together the most distal V segments to the D/J segments before recombination (Figure 5). Fluorescence *in situ* hybridization experiments using locus-specific DNA probes revealed that the *Ig* and *Tcr* loci undergo reversible contraction, facilitating recombination of the locus at appropriate developmental stage (Skok *et al.*, 2007; Sayegh *et al.*, 2005; Roldan *et al.*, 2005; Fuxa *et al.*, 2004; Kosak *et al.*, 2002). More recently, based on an extended 3D structure and predicted dynamic of the *IgH* locus during B-cell development (Lucas *et al.*, 2014; Jhunjhunwala *et al.*, 2008), a model was proposed in which antigen receptor loci are organized in a dynamic rosette-like structure made of multiple loops that undergo variable compaction (Bossen *et al.*, 2012; Schatz and Ji, 2011; Okada and Comings, 1979). Of note, contraction of the *IgH* locus and rearrangement of the distal V_H genes with proximal D_HJ_H segments require key factors that are also involved in B lineage cell priming and commitment, including Ikaros and Pax5 (Reynaud *et al.*, 2008; Fuxa *et al.*, 2004), in addition to ubiquitous factors such as YY1, CTCF and its cofactor cohesin (Ebert *et al.*, 2013; Chaumeil and Skok, 2012). Recent findings also suggest that sterile transcription within the *Ig* loci might also play an active role in locus conformation and looping (Stubbington and Corcoran, 2013; Verma-Gaur *et al.*, 2012).

Molecular Mechanisms of RAG-Mediated V(D)J Recombination

RAG Proteins: Core and Noncore Regions

The RAG recombinase is the only lymphoid-specific protein complex required for V(D)J recombination (Oettinger *et al.*, 1990; Schatz *et al.*, 1989). The mouse and human *Rag1* and *Rag2* genes lie on chromosome 2 and chromosome 11 respectively, immediately adjacent to one another with the entire open reading frame of each gene contained in a single exon. Full-length murine RAG1 (1040 amino acids) and RAG2 (527 amino acids) are conserved polypeptides with species-specific variations of only 3–10 amino acids (Jones and Simkus, 2009). Briefly, RAG1 possesses site-specific DNA binding activity and catalytic residues. RAG2 has little or no DNA binding activity on its own but is crucial to the reaction as it contains regulatory domains that control DNA cleavage, DNA binding, post-cleavage DNA end stabilization, and DNA repair pathway choice (Figure 6; Schatz and Swanson, 2011; Jones and Simkus, 2009; Gellert, 2002). The RAG1 protein contains a nonamer-binding domain, a central region that is responsible for interactions with the RSS heptamer and RAG2, and an active site for DNA cleavage, which includes three essential acidic amino acids (D600, D708, and E962) that coordinate divalent metal ions and are essential for catalysis. RAG1 also contains two zinc finger regions (Zn) important for homodimerization (ZnA) and interaction with RAG2 (ZnB). Interestingly, ZnA also has ubiquitin ligase activity and interacts with and ubiquitylates histone H3 (Schatz and Swanson, 2011; Jones and Simkus, 2009; Gellert, 2002).

Early mutagenesis studies defined the catalytic core of RAG1 and RAG2 proteins to residues 384–1008 and 1–383,

respectively (Jones and Simkus, 2009; Sadofsky *et al.*, 1993, 1994). More recent work has reduced the minimal region of RAG2 required to support recombination to residues 1–350 (Coussens *et al.*, 2013), coinciding with the end of a region containing six kelch-like motifs predicted to adopt a six-bladed β -propeller-like structure (Callebaut and Mornon, 1998) termed the core domain. This core RAG2 region is essential for DNA cleavage activity, enhances RSS binding and interacts with RAG1. The C-terminal noncore region of RAG2 contains multiple regulatory motifs (Jones and Simkus, 2009): a PHD finger that interacts H3K4me3, a cell cycle regulated protein degradation signal in which the phosphorylation of an essential threonine (T490) signals the periodic destruction of RAG2 at the G1-to-S transition, and a highly conserved acidic 'hinge' region that participates in post-cleavage DNA ends stabilization and DNA repair pathway choice (Coussens *et al.*, 2013) through an unknown mechanism. The C-terminus can also bind to phosphatidylinositol phosphate (PIP) as well as core histones (Jones and Simkus, 2009), but the significance of these interactions remains unclear.

DNA Binding and Cleavage

The *Tcr* and *Ig* V, D, and J segments are flanked by RSS that consist of relatively conserved heptamer and nonamer elements, with respective consensus sequences of CACAGTG and ACAAAAACC, separated by an intervening spacer of either 12 (12-RSS) or 23 (23-RSS) nucleotides (Figures 7 and 8(a)). The cleavage starts with the binding of the RAG proteins to one RSS and the formation of a synapsis in which a partner RSS is captured. HMG1 (high mobility group 1), a nonspecific DNA bending protein, facilitates synapsis formation and subsequent cleavage (Dai *et al.*, 2005; Swanson, 2002; van Gent *et al.*, 1997). After binding, the RAG proteins nick one DNA strand precisely between the RSS heptamer and the coding segment. This generates a free 3'-OH that is used to attack the opposite strand in a *trans*-esterification reaction, forming a DNA DSB (Schatz and Swanson, 2011). This paired cleavage step results in the formation of four broken DNA ends: two covalently sealed (hairpin) coding ends and two blunt (signal end) DNA ends (Schatz and Swanson, 2011; Figure 7).

Regulation of Cleavage – 12/23 Restriction and Beyond

Although nicking has been shown to operate at isolated RSSs (and at RSS-like sequences *in vitro*), it is modulated by a distinct synapsis of a 12-RSS and 23-RSS pair (Schatz and Swanson, 2011). Additionally, efficient DSB formation requires 12/23-RSS pair assembly into a synaptic complex with the RAG proteins. This restriction, referred to as the 12/23 rule, directs recombination between appropriate gene segments within a given locus (Figure 8(b)) and likely participates in limiting DSB formation to pre-identified and preassembled RSS partners, therefore preventing aberrant rearrangements. Various lines of evidence indicate that another level of regulation resides in the biochemistry of the RAG-DNA recombination reaction. Though theoretically 12/23 compatible, the direct joining of some gene segments, for example, the *TcrV β* (23-RSS) and *TcrJ β* (12-RSS) segments rarely occurs

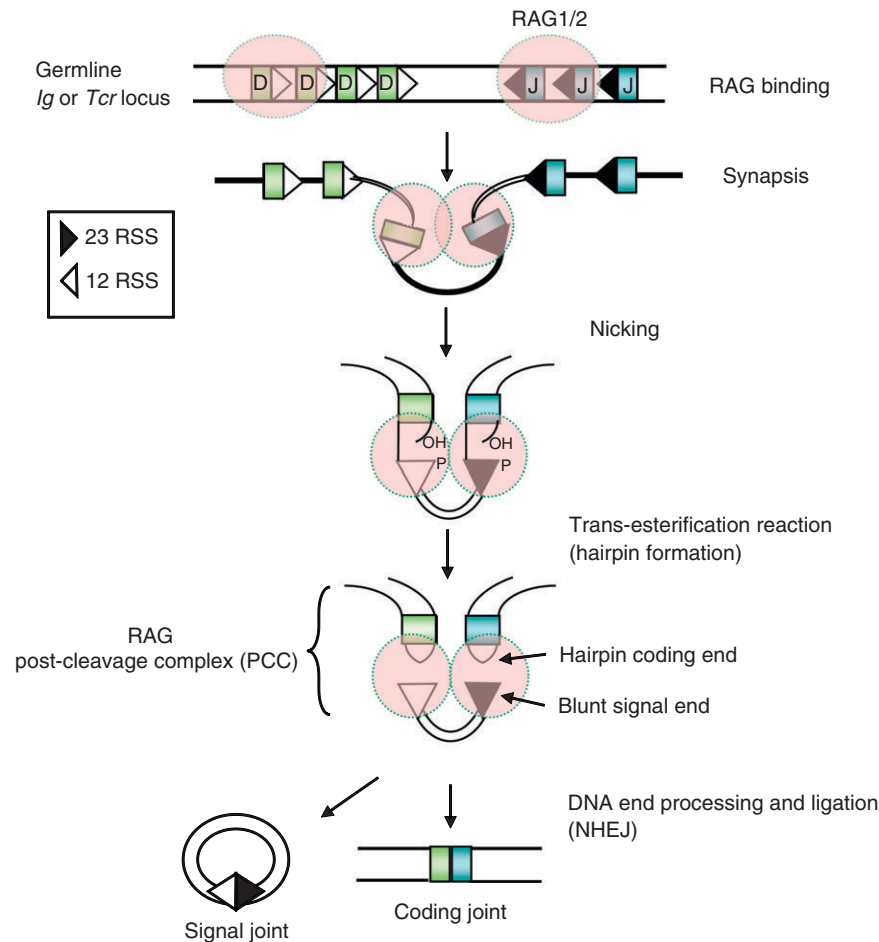


Figure 7 RAG cleavage. After RSS recognition and synapsis, RAG1/2 proteins introduce single-strand nicks precisely between the RSS heptamer and the coding segment. The resulting 3'-OH is used to catalyze a direct *trans*-esterification reaction on the opposite DNA strand, generating a DNA DSB. This results in the formation of paired hairpin (coding end) and blunt (signal end) DNA ends that are held in a RAG PCC. DNA ends are then processed and ligated by the NHEJ pathway to form coding and signal joints.

(Figure 8(c)). This restriction is referred to as 'Beyond 12/23' or 'B12/23' and seems to be enforced by the ability of the RAG proteins to initiate cleavage at certain RSSs within the synaptic complex (Banerjee and Schatz, 2014; Schatz and Swanson, 2011; Tillman *et al.*, 2004).

Joining the Broken DNA Ends

After cleavage, the RAG proteins stay associated with the DNA ends in a post-cleavage complex (PCC), yet providing another layer of regulation to be enforced at the joining step (Figure 8, also see below). It is not known whether the four DNA ends all stay associated in a single complex or whether the two coding ends and two signal ends are maintained separately. The fact that the RAGs bind much more avidly to signal-end pairs than coding-end pairs indicates that RAG-signal end PCCs and RAG-coding end PCCs may have different properties (Hiom and Gellert, 1997; Agrawal and Schatz, 1997), and perhaps reflects the need for coding ends to be accessible for processing before ligation. The ubiquitous NHEJ machinery joins pairs of RSS signal ends and pairs of coding segment ends to form signal and coding joints, respectively (Figures 8 and 9).

The V(D)J recombination reaction therefore provides an inherent chromosomal directionality to the rearrangements that is dependent on relative orientation of participating gene segments and their associated RSSs.

NHEJ pathway

In mammalian cells, three specialized repair systems exist to detect, process, and ligate broken DNA ends (Deriano and Roth, 2013). Homologous recombination (HR) and classical NHEJ (cNHEJ) are the major DSB repair pathways. HR functions in the post-replicative S and G2 phases of the cell cycle and employs a template – the sister chromatid or homolog – to direct repair so that DSBs or even gaps can be repaired seamlessly, without loss of genomic information. As opposed to HR, cNHEJ is active throughout the cell cycle, particularly in G0/G1, and utilizes direct joining which does not require sequence homology although very short sequence homologies of a few nucleotides can appear at the junction and might help to align DNA ends. A third and most recently discovered pathway goes by the name of alternative NHEJ (aNHEJ) and was identified in yeast and mammalian cells as a 'backup' system capable of repairing DSBs in cells genetically deficient

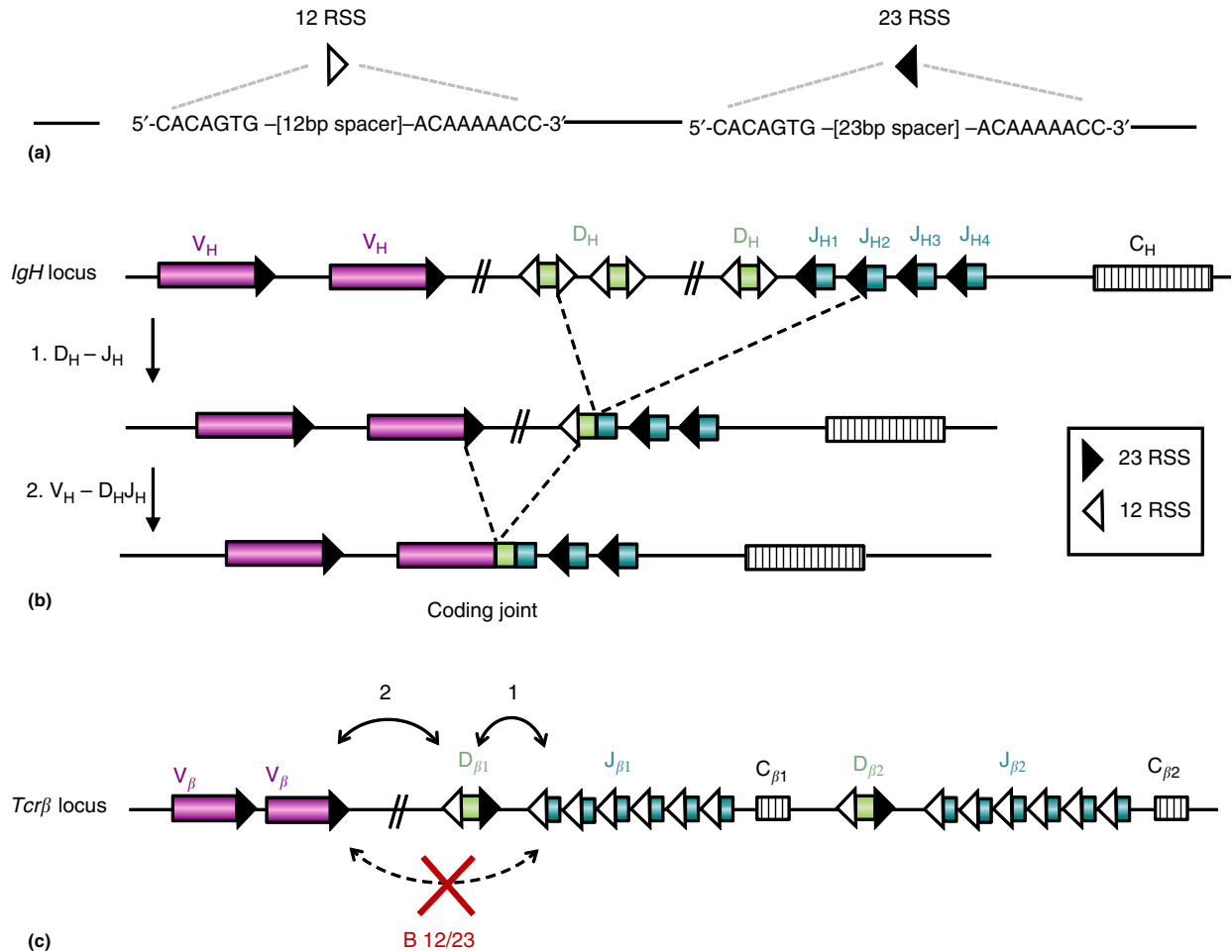


Figure 8 Recognition signal sequences, 12/23 rule and beyond. (a) RSS. The consensus sequences of the heptamer and nonamer are shown with either a 12 base pair (12-RSS) or 23 base pair (23-RSS) intervening spacer. (b and c) Schematics of *IgH* (b) and *Tcrβ* (c) loci depicting the order of V, D, and J gene segments rearrangements. Efficient recombination occurs only between gene segments flanked by a 12- and 23-RSS (12/23 rule). Though theoretically 12/23 compatible, the joining of some gene segments, such as the V_β (23-RSS) and J_β (12-RSS) segments in the *Tcrβ* locus, never occurs (beyond 12/23 rule or B12/23), implying that additional mechanisms regulate RAG-mediated cleavage.

for one (or more) factors critical for cNHEJ. aNHEJ is characterized by excessive deletions and frequent, although not invariable, microhomologies at the repair junctions. It is also much less reliable than cNHEJ, as it commonly leads to loss of genetic information and to chromosome translocations (at least in mice).

DSB repair requires DNA end recognition, processing if needed, and ligation (Deriano and Roth, 2013; Helmink and Sleckman, 2012). For cNHEJ, the Ku70-Ku80 heterodimer (Ku) recognizes DNA ends and the DNA ligase complex XRCC4-Ligase IV joins them (Figure 9). cNHEJ of all types of broken DNA ends depends *ad minima* on these 'core' activities. As RAG-cleaved signal ends are perfect blunt ended cNHEJ substrates, they are usually directly ligated with little or no nucleotide additions or losses. RAG-cleaved coding ends however terminate in hairpins that must be processed before joining. After binding to broken DNA ends, Ku recruits the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) to form the DNA-PK holoenzyme. DNA-PK phosphorylates multiple substrates, promoting synapsis of DNA ends and

facilitating the recruitment of end processing and ligation enzymes. One such enzyme is Artemis, discovered as the gene mutated in certain radiosensitive severe combined immunodeficiency (RS-SCID) patients. The endonuclease activity of Artemis is stimulated by DNA-PK and carries out hairpin opening at coding ends. This hairpin opening is often asymmetric from the apex, leading to the addition of short palindromic nucleotides ('P' nucleotides) at the junction (Helmink and Sleckman, 2012). After hairpin opening, the TdT DNA polymerase proceeds to template-independent nucleotide addition ('N' nucleotides) to 3' DNA ends, which, along with P nucleotides, increase junctional diversity at coding joints (Komori *et al.*, 1993; Gilfillan *et al.*, 1993; Desiderio *et al.*, 1984). TdT is only expressed in developing lymphocytes undergoing V(D)J recombination and adds nucleotides primarily to coding ends and not signal ends. Interestingly, TdT expression is absent during fetal life, which leads to a smaller T- and B-cell antigen receptor repertoire in neonates than in adults (Feeney, 1992a,b; Bogue *et al.*, 1991; Gu *et al.*, 1990). Mice genetically deficient for TdT are less sensitive to

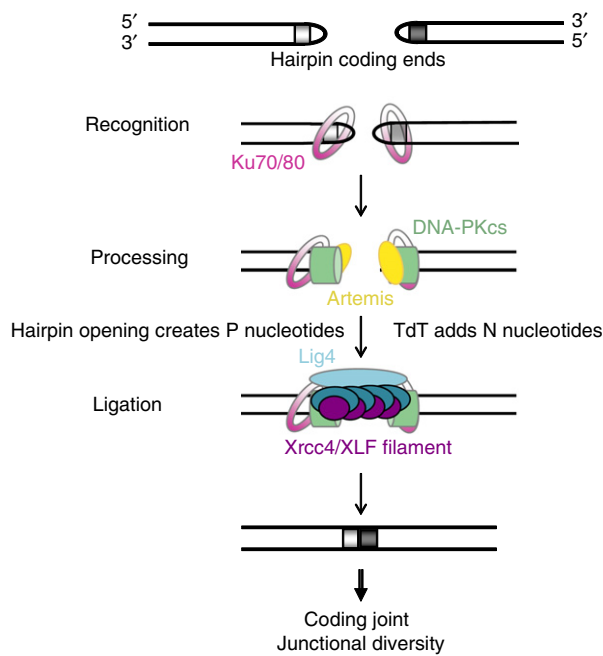


Figure 9 Repair of RAG DNA breaks by the NHEJ pathway: DNA end recognition, processing, and ligation. See text for details.

autoimmune diseases (Robey *et al.*, 2004), indicating that the level of antigen receptor diversification must be tightly controlled (at least during fetal and early life) to avoid the appearance of excessive autoreactive specificities in the antigen receptor repertoire. The final step in the repair of RAG-mediated DNA breaks is joining of the DNA ends by the DNA Ligase IV/XRCC4/Cernunnos-XLF complex (Figure 9).

Genetic ablation of the cNHEJ factors Ku, DNA-PKcs, Artemis, DNA Ligase IV, or XRCC4 in mice prevents the completion of V(D)J recombination, arresting B- and T-cell development at an early stage and leading to SCID phenotype (Deriano and Roth, 2013; Boboila *et al.*, 2012). Therefore, although multiple end-joining pathways exist, V(D)J recombination strictly relies on cNHEJ. This restriction is explained both by the limitation of RAG expression to the G1 phase of the cell cycle and by the ability of the wild type RAG complex to guide DNA ends to the cNHEJ pathway (Table 1 and see Section ‘The RAG recombinase: Domestication of a transposase’).

DDR

DSBs activate ATM kinase-dependent DDR in which ATM phosphorylates numerous substrates that mediate cell-cycle checkpoints, chromatin changes and DNA repair (Ciccio and Elledge, 2010). ATM-dependent p53 phosphorylation mediates G1/S checkpoint that arrest or eliminates cells with unrepaired DSBs. ATM also phosphorylates chromatin- and/or DNA-associated proteins, including H2AX, 53BP1, MDC1, and factors of the MRE11 complex (MRE11, NBS1, and RAD50) that assemble over large DNA regions of the chromatin on both sides of DNA breaks to form so-called nuclear DNA repair foci.

Interestingly, these factors, although present at RAG-mediated DSBs (Chen *et al.*, 2000), seem to have only modest

roles in overall V(D)J recombination. The most striking example is 53BP1, which, through participating in cNHEJ, helps to repair AID-dependent DNA breaks occurring during the process of CSR in mature B cells. Although critical for CSR, loss of 53BP1 has only a very modest impact on lymphocyte development and V(D)J recombination, indicating that 53BP1 functions might be compensated by other accessory proteins during the repair of RAG-DNA breaks (Boboila *et al.*, 2012). A similar situation is observed in the case of XRCC4-like factor (XLF-also called Cernunnos). XLF, the most recently identified cNHEJ factor, was discovered simultaneously from patients that display growth retardation, immunodeficiency, and radiosensitivity (Buck *et al.*, 2006) and by a yeast two-hybrid screen for proteins that interact with XRCC4 (Ahnesorg *et al.*, 2006). XLF-deficient mice, unlike other cNHEJ-deficient mice, are not markedly immunodeficient and pro-B cell lines derived from these animals perform nearly normal V(D)J recombination, leading to the speculation that unknown lymphocyte-specific factors/pathways might compensate for XLF function during V(D)J recombination (Li *et al.*, 2008). The answer came from the analysis of mice with combined deficiencies of DDR (ATM, H2AX, or 53BP1) factors with XLF. In these mice, lymphocyte development is blocked at an early stage due to a significant defect in the repair of RAG-mediated DSBs, indicating that XLF serves more major roles in cNHEJ during V(D)J recombination, which are masked by functional redundancies with the ATM-dependent DDR (Kumar *et al.*, 2014).

Studies in pro-B cell lines demonstrated that in the absence of ATM, while V(D)J recombination is not abrogated, a fraction of RAG-dependent DSBs is released from the PCCs, leading to unrepaired DNA ends that can persist in the cell and engage chromosomal instability, including translocations (Callen *et al.*, 2007; Bredemeyer *et al.*, 2006). These findings indicate that ATM, and likely factors of the ATM-dependent DDR, assist in tethering of the DNA ends in the RAG PCC prior to their ligation, a function that might overlap with XLF (Kumar *et al.*, 2014; Deriano and Roth, 2013; Vera *et al.*, 2012). Recent structural analysis of the XLF–XRCC4 complex supports this assumption; XLF forms with XRCC4 long heterofilaments readily observable by electron microscopy (Andres *et al.*, 2012; Ropars *et al.*, 2011). One possible common function between these filaments and the ATM-DDR might be end-tethering/stabilization of the RAG PCC.

Another example of functional redundancy between cNHEJ and DDR factors comes from the analysis of signal joint formation in DNA-PKcs and ATM-deficient lymphocytes. Recent reports pointed to overlapping functions of DNA-PKcs and ATM, mediated through their kinase activities, in promoting efficient signal joint formation and preventing accumulation of signal ends (Zha *et al.*, 2011; Capud and Sleckman, 2011). In response to DSBs, DNA-PKcs and ATM phosphorylate a large number of shared substrates including chromatin-associated factors, cNHEJ factors and potentially the RAG proteins. It is difficult to attribute the flaws in signal joint formation in this situation to defective tethering. Nevertheless, it is likely that cNHEJ and DDR (and RAG proteins – see below) collaborate in maintaining pairs DNA ends together to facilitate proper repair and minimize aberrant end-joining events (Figure 10).

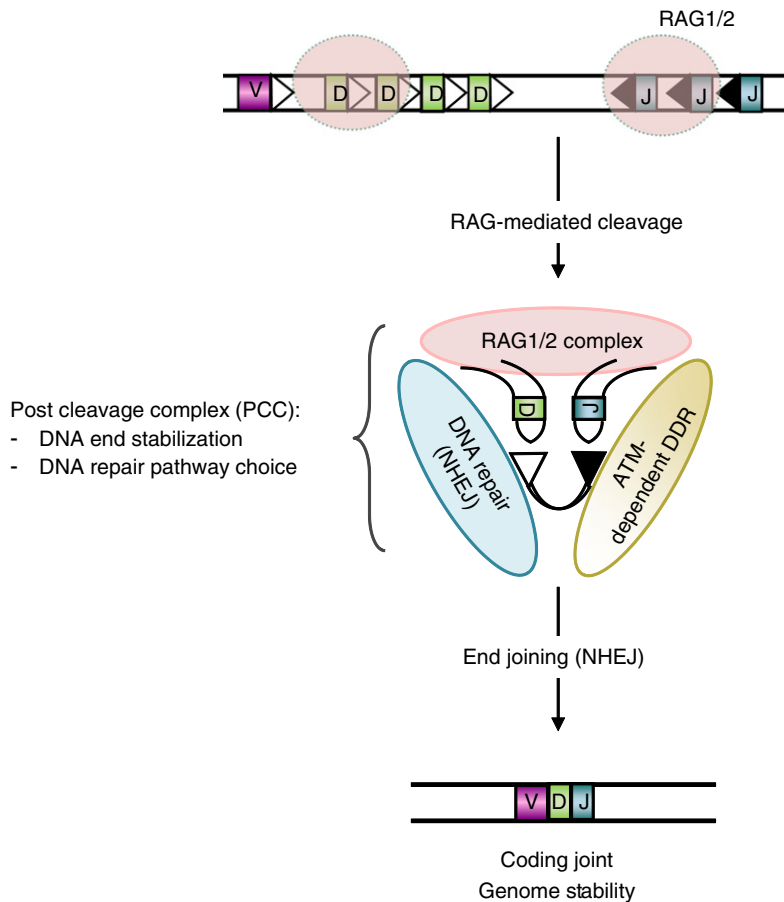


Figure 10 RAG recombinase, DDR, and DNA repair: cooperation to maintain genome stability during V(D)J recombination. See text for details.

The RAG recombinase: Domestication of a transposase

The structure of the *Rag* genes (i.e., DDE motif in RAG1) and similarities in the reaction chemistry (i.e., *trans*-esterification mechanism) of V(D)J recombination and transposition support the idea that the RAG recombinase derives from a transposable element that entered the genome of a common ancestor of all jawed vertebrates (Litman *et al.*, 2010; Fugmann, 2010). However, transposition, the process of moving genetic information from one position in the host genome into another, is substantially divergent from RAG-mediated V(D)J recombination. In the latter, the RAG complex participates in resealing of DNA breaks through cNHEJ and joins the two signal ends together instead of using them to integrate into a new target location. Interestingly, Transib and Transib-like transposition elements contain a single open reading frame that shares similarities with RAG1 (Fugmann, 2010). RAG2, however, lacks any apparent sequence similarity to any known transposon encoded protein (Fugmann, 2010). It is therefore tempting to speculate that RAG2 might have largely contributed to the evolution of an ancient RAG1 transposase into a domesticated V(D)J recombinase that hands off the DNA ends to the cNHEJ machinery, assuring full diversification of the Ig and TCR repertoires (Figure 9) while preserving the genome and organism integrity (Figure 10 and see below).

During the past 15 years, combination of biochemical, genetic, and *in vivo* studies have shed light on a novel role for

the RAG proteins in the joining step of V(D)J recombination. The discovery of separation of function mutants in RAG1 and RAG2 that are capable of cleavage but exhibit severe joining defects provided the first compelling evidence that the RAG PCC might serve a crucial function in joining both coding and signal ends (Roth, 2003). Additional studies showed that destabilization of the RAG PCC in certain RAG mutants leads to repair by HR and aNHEJ (Coussens *et al.*, 2013; Corneo *et al.*, 2007; Lee *et al.*, 2004). These results led to a model in which the quite stable RAG PCC actively directs DNA ends to the cNHEJ machinery for repair, thus protecting them from error-prone end-joining pathways and aberrant recombination events. This model predicts that removal of such RAG regulatory function would affect genome stability during V(D)J recombination. This is in fact the case; mice harboring certain mutations in the C-terminus of RAG2 display a destabilized RAG-DNA end complex, allow increased aNHEJ accompanied by genomic instability and accelerated lymphomagenesis in the absence of p53, generating tumors bearing a complex landscape of translocations, deletions, inversions, and duplications (Deriano and Roth, 2013; Deriano *et al.*, 2011). Interestingly, the lymphomas and translocations seen in RAG2 mutant/p53 deficient mice closely resemble those of ATM-deficient mice (also see below – V(D)J Recombination and Malignancies of the Immune System), suggesting that a similar DNA end release mechanism underlies genomic instability

and lymphoma genesis in both mouse models (Deriano *et al.*, 2011). These results also revealed a genome guardian role for RAG2 during V(D)J recombination and lymphocyte development.

V(D)J Recombination and Malignancies of the Immune System

V(D)J Recombination-Mediated Genomic Instability and Lymphoid Cancers

Theoretical models

There are at least two general models (Figure 11) to explain aberrant V(D)J recombination events (Brandt and Roth,

2009): (1) cleavage errors in which the RAG proteins recognize one or more cryptic RSSs, or other target sequences, generating a chromosomal translocation (e.g., between an antigen receptor locus and another chromosome at a cryptic RSS – Figure 11(a), left panel) or a chromosomal deletion (e.g., between two cryptic RSSs on the same chromosome – Figure 11(a), right panel); (2) aberrant joining which could result from either release of DNA ends from the protection of the RAG PCC and/or absence of DDR or NHEJ factors, both resulting in an increased susceptibility of the free DNA ends to engage genomic instability (Figure 11(b)).

Body of evidence in humans

During *Tcr* or *Ig* rearrangements, other genetic loci are present in an open chromatin configuration and therefore susceptible

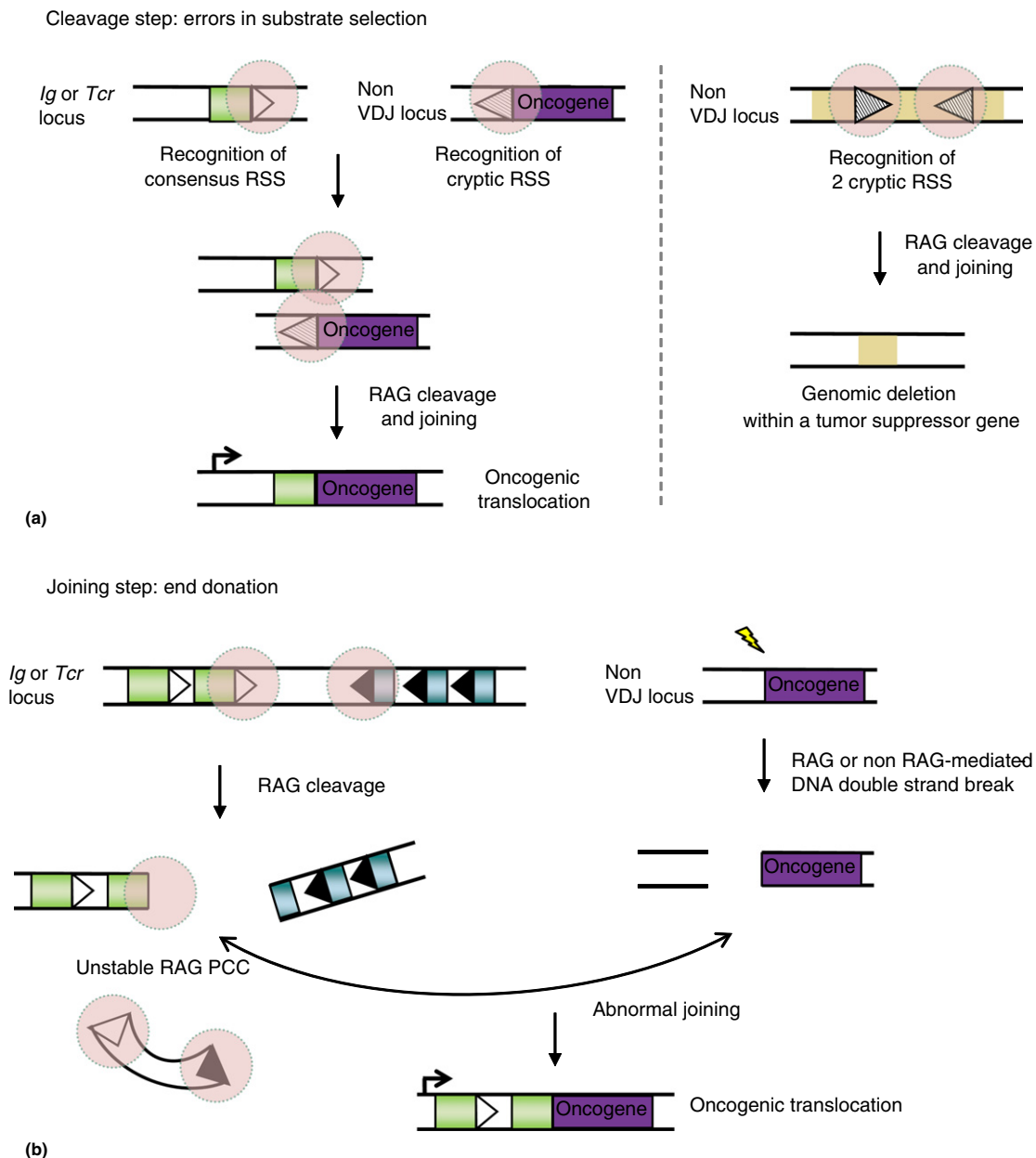


Figure 11 Theoretical models of RAG-mediated oncogenic rearrangements. See text for details.

to the activity of the RAG recombinase thus putting the stability of the genome as a whole at risk. The V(D)J recombinase was first suspected to be involved in human lymphoid neoplasm when observations of translocation breakpoints in certain tumors showed that an antigen receptor locus had been fused with other regions of chromosomal DNA, often placing an oncogene next to a highly expressed antigen receptor gene. Examples of this are the *IgH/MYC* translocation in Burkitt lymphoma and the *IgH/BCL2* translocation in follicular lymphoma (Kuppers and Dalla-Favera, 2001; Tsujimoto *et al.*, 1985; Erikson *et al.*, 1982). Aberrant V(D)J recombination activities have also been implicated in T-cell leukemia and lymphoma (Onozawa and Aplan, 2012; Aifantis *et al.*, 2008). In approximately 35% of T-acute lymphoblastic leukemia (ALL), the *Tcr* loci are involved in translocations that probably occur during T-cell differentiation when V(D)J recombination takes place. In addition to translocations, aberrant V(D)J recombination has been shown to generate oncogenic inducing deletions/fusions such as the classic oncogenic *SIL-TAL1* fusion (interstitial deletions involving pairs of RSS-like motifs also called cryptic RSS) in 25% of pediatric T-ALL (Onozawa and Aplan, 2012). Deletions leading to the activation of Notch1 are frequently mediated by dysfunctional RAG activity in murine T lymphomas, and perhaps in adult T-ALL (Aifantis *et al.*, 2008). Lymphomagenic genome rearrangements attributed to RAG activities are also seen in B-ALL, including *IKZ1* deletions and alterations of, for example, *PAX5* (Mullighan and Downing, 2009; Kuppers, 2005; Kuppers and Dalla-Favera, 2001). These genomic aberrations are normally associated with multiple other chromosomal abnormalities, often concealing the underlying mechanism(s) implicated in the onset and in the progression of the disease. Nevertheless, a recent study clearly points to RAG-mediated recombination as the predominant driver of oncogenic rearrangement in a subtype of ALL (ETV6-RUNX1-positive ALL) (Papaemmanuil *et al.*, 2014). The authors performed genome and exome sequencing on over 50 leukemic samples and resolved 354 genomic structural variations (SVs) at base-pair resolution. Using this data, they searched for typical features of RAG activity and found that RSSs were identified in 12% of the SVs, close to 40% of the breakpoint regions contained cryptic RSSs, and non-templated sequences were found at 70% of the breakpoints. Of note, these motifs were highly enriched compared to genomic data from several solid tumors, which presumably do not originate from RAG-expressing cells. A possible explanation for the increased generation of SVs at RAG target sites could be that in various ALL subtypes, including ETV6-RUNX1-positive ALL, the RAG genes are constitutively expressed. However, in opposition to this idea, in ETV6-RUNX1-positive ALL, RAG-motifs were also enriched compared to the haplodiploid ALL and early T-cell precursor ALL subtypes. These observations suggest that ETV6-RUNX1-positive ALL cells might harbor a unique deregulation of the V(D)J recombinase that drives genomic instability and cell transformation. Additional experimental approaches will likely provide insight into the underlying mechanisms of V(D)J recombinase-induced SVs in cancer genomes.

Experimental evidences

Direct evidence that the RAG recombinase mediates genomic instability *in vivo* comes from the analysis of mice deficient in

factors involved in detecting and responding to DNA breaks (Gostissa *et al.*, 2011; Nussenzweig and Nussenzweig, 2010; Jankovic *et al.*, 2007). The clearest examples come from mice deficient for various cNHEJ factors (Ku80, XRCC4, Ligase IV, DNA-PKcs, or Artemis) that are also deficient for p53 tumor suppressor. These mice invariably develop pro-B cell lymphomas harboring oncogenic translocations that link the *IgH* locus to sequences near *c-myc* on chromosome 15 (or *N-myc* on chromosome 12 in the case of Artemis deficiency), and both genes are commonly amplified. The majority of breakpoints in the *IgH* locus are near the J_H cluster, supporting the idea that failed V(D)J recombination initiates these translocations. Additionally, breeding these mice onto a RAG-deficient background leads to the appearance of B-cell lymphomas that lack these typical translocations and amplicons. Similarly, elimination of RAG in ATM-deficient mice, which routinely develop T-cell lymphomas with recurrent translocations and amplifications involving the *Tcr α / δ* locus on chromosome 14, leads to the kinetically delayed appearance of thymic lymphomas that lack the *Tcr α / δ* translocation. Thus, abundant direct experimental evidence links RAG activity to the generation of recurrent oncogenic translocations (Deriano and Roth, 2013; Gostissa *et al.*, 2011; Nussenzweig and Nussenzweig, 2010; Jankovic *et al.*, 2007; Roth, 2003). These findings also demonstrate that the cNHEJ pathway acts as a genome guardian by promoting faithful joining and that p53 activation protects against RAG-mediated translocations (Deriano and Roth, 2013; Roth and Gellert, 2000).

Although RAG involvement in generating *Ig* or *Tcr* breaks in human and mouse lymphomas has been well characterized, the origin of the DSB on partner chromosomal loci, such as the *BCL2* or *MYC* oncogenes, is more speculative. As aforementioned, the RAG proteins might cleave cryptic RSSs at non-V(D)J loci (Papaemmanuil *et al.*, 2014; Marculescu *et al.*, 2002; Lewis *et al.*, 1997). RAG also recognizes sequences that adopt a non-B DNA structure, as found in the main breakpoint region of *BCL2*, and nicks DNA at sites of transition from double stranded-DNA to single stranded-DNA; such nicks could be converted to DSBs by several mechanisms and are involved in the generation of chromosomal translocations (Raghavan *et al.*, 2004; Lee *et al.*, 2004). Finally, translocations might involve one RAG-mediated DSB and a DSB outside antigen receptor locus generated by other mechanisms (Bunting and Nussenzweig, 2013; Alt *et al.*, 2013). In any case, RAG recombinational activities need to be tightly kept in check to avoid genome scrambling and appearance of potentially oncogenic mutations.

V(D)J Recombination and Immune Deficiency Disorders

As previously mentioned, V(D)J recombination is a prerequisite to lymphocyte differentiation and the generation of a full repertoire of different *Ig* and *TCR* molecules. In humans, naturally occurring gene mutations have been identified that cause a complete or partial loss of physiological V(D)J recombination activity (Woodbine *et al.*, 2014; Schatz and Swanson, 2011; de Villartay, 2009). The most striking examples come from patients deficient in NHEJ proteins, such as Ligase IV, XLF, Artemis, or DNA-PKcs, who typically display radiosensitivity (as a consequence of the role of NHEJ in the

rejoining of radiation-induced DSBs) and severe combined immunodeficiency (RS-SCID). Of note, NHEJ deficiency can also confer additional developmental abnormalities, including microcephaly (Woodbine *et al.*, 2014; de Villartay, 2009). RAG mutations have also been associated with a spectrum of severe immune deficiency disorders. Complete RAG deficiency is associated with classical SCID and absence of T- and B cells. RAG deficiency with residual V(D)J recombinational activity is associated with several clinical and immunological subtypes such as complete or incomplete Omenn syndrome, $\gamma\delta$ T-cell expansion, granuloma formation, and maternofetal engraftment (Schatz and Swanson, 2011; Niehues *et al.*, 2010; Villa *et al.*, 2008).

Acknowledgments

Research in the Deriano lab is supported by the Institut Pasteur, the CNRS, as well as by the European Research Council under the ERC starting grant agreement #310907.

References

- Abarregui, I., Krangel, M.S., 2006. Regulation of T cell receptor- α gene recombination by transcription. *Nature Immunology* 7, 1109–1115.
- Abarregui, I., Krangel, M.S., 2009. Germline transcription: A key regulator of accessibility and recombination. *Advances in Experimental Medicine and Biology* 650, 93–102.
- Agrawal, A., Schatz, D.G., 1997. RAG1 and RAG2 form a stable postcleavage synaptic complex with DNA containing signal ends in V(D)J recombination. *Cell* 89, 43–53.
- Ahnesorg, P., Smith, P., Jackson, S.P., 2006. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell* 124, 301–313.
- Aifantis, I., Raetz, E., Buonamici, S., 2008. Molecular pathogenesis of T-cell leukaemia and lymphoma. *Nature Reviews Immunology* 8, 380–390.
- Alt, F.W., Zhang, Y., Meng, F.L., Guo, C., Schwer, B., 2013. Mechanisms of programmed DNA lesions and genomic instability in the immune system. *Cell* 152, 417–429.
- Andres, S.N., Vergnes, A., Ristic, D., *et al.*, 2012. A human XRCC4-XLF complex bridges DNA. *Nucleic Acids Research* 40, 1868–1878.
- Banerjee, J.K., Schatz, D.G., 2014. Synapsis alters RAG-mediated nicking at Tcrb RSSs: Implications for the beyond 12/23 rule. *Molecular and Cellular Biology* 34, 2566–2580.
- Barreto, V., Cumano, A., 2000. Frequency and characterization of phenotypic Ig heavy chain allelically included IgM-expressing B cells in mice. *Journal of Immunology* 164, 893–899.
- Bevington, S., Boyes, J., 2013. Transcription-coupled eviction of histones H2A/H2B governs V(D)J recombination. *EMBO Journal* 32, 1381–1392.
- Boboila, C., Alt, F.W., Schwer, B., 2012. Classical and alternative end-joining pathways for repair of lymphocyte-specific and general DNA double-strand breaks. *Advances in Immunology* 116, 1–49.
- Bogue, M., Candeias, S., Benoist, C., Mathis, D., 1991. A special repertoire of α and β T cells in neonatal mice. *EMBO Journal* 10, 3647–3654.
- Borghesi, L., Hsu, L.Y., Miller, J.P., *et al.*, 2004. B lineage-specific regulation of V(D)J recombinase activity is established in common lymphoid progenitors. *Journal of Experimental Medicine* 199, 491–502.
- Bossen, C., Mansson, R., Murre, C., 2012. Chromatin topology and the regulation of antigen receptor assembly. *Annual Review of Immunology* 30, 337–356.
- Brady, B.L., Steinell, N.C., Bassing, C.H., 2010. Antigen receptor allelic exclusion: An update and reappraisal. *Journal of Immunology* 185, 3801–3808.
- Brandt, V.L., Hewitt, S.L., Skok, J.A., 2010. It takes two: Communication between homologous alleles preserves genomic stability during V(D)J recombination. *Nucleus* 1, 23–29.
- Brandt, V.L., Roth, D.B., 2009. Recent insights into the formation of RAG-induced chromosomal translocations. *Advances in Experimental Medicine and Biology* 650, 32–45.
- Bredemeyer, A.L., Sharma, G.G., Huang, C.-Y., *et al.*, 2006. ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. *Nature* 442, 466–470.
- Buck, D., Malivert, L., de Chasseval, R., *et al.*, 2006. Cernunnos, a novel nonhomologous end-joining factor, is mutated in human immunodeficiency with microcephaly. *Cell* 124, 260–262.
- Bunting, S.F., Nussenzweig, A., 2013. End-joining, translocations and cancer. *Nature Reviews Cancer* 13, 443–454.
- Callebaut, I., Mornon, J.P., 1998. The V(D)J recombination activating protein RAG2 consists of a six-bladed propeller and a PHD fingerlike domain, as revealed by sequence analysis. *Cellular and Molecular Life Sciences* 54, 880–891.
- Callen, E., Jankovic, M., Difilippantonio, S., *et al.*, 2007. ATM prevents the persistence and propagation of chromosome breaks in lymphocytes. *Cell* 130, 63–75.
- Cedar, H., Bergman, Y., 2011. Epigenetics of haematopoietic cell development. *Nature Reviews Immunology* 11, 478–488.
- Chaumeil, J., Micsinai, M., Ntziachristos, P., *et al.*, 2013a. Higher-order looping and nuclear organization of Tcr facilitate targeted rag cleavage and regulated rearrangement in recombination centers. *Cell Reports* 3, 359–370.
- Chaumeil, J., Micsinai, M., Ntziachristos, P., *et al.*, 2013b. The RAG2 C-terminus and ATM protect genome integrity by controlling antigen receptor gene cleavage. *Nature Communications* 4, 2231.
- Chaumeil, J., Skok, J.A., 2012. The role of CTCF in regulating V(D)J recombination. *Current Opinion in Immunology* 24, 153–159.
- Chaumeil, J., Skok, J.A., 2013. A new take on v(dj) recombination: Transcription driven nuclear and chromatin reorganization in RAG-mediated cleavage. *Frontiers in Immunology* 4, 423.
- Chen, H.T., Bhandoola, A., Difilippantonio, M.J., *et al.*, 2000. Response to RAG-mediated VDJ cleavage by NBS1 and gamma-H2AX. *Science* 290, 1962–1965.
- Ciccia, A., Elledge, S.J., 2010. The DNA damage response: Making it safe to play with knives. *Molecular Cell* 40, 179–204.
- Clark, M.R., Mandal, M., Ochial, K., Singh, H., 2014. Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. *Nature Reviews Immunology* 14, 69–80.
- Corneo, B., Wendland, R.L., Deriano, L., *et al.*, 2007. Rag mutations reveal robust alternative end joining. *Nature* 449, 483–486.
- Coussens, M.A., Wendland, R.L., Deriano, L., *et al.*, 2013. RAG2's acidic hinge restricts repair-pathway choice and promotes genomic stability. *Cell Reports* 4, 870–878.
- Dai, Y., Wong, B., Yen, Y.M., *et al.*, 2005. Determinants of HMGB proteins required to promote RAG1/2-recombination signal sequence complex assembly and catalysis during V(D)J recombination. *Molecular and Cellular Biology* 25, 4413–4425.
- Deriano, L., Chaumeil, J., Coussens, M., *et al.*, 2011. The RAG2 C-terminus suppresses genomic instability and lymphomagenesis. *Nature* 471, 119–123.
- Deriano, L., Roth, D.B., 2013. Modernizing the nonhomologous end-joining repertoire: Alternative and classical NHEJ share the stage. *Annual Review of Genetics* 47, 433–455.
- Desiderio, S., 2010. Temporal and spatial regulatory functions of the V(D)J recombinase. *Seminars in Immunology* 22, 362–369.
- Desiderio, S.V., Yancopoulos, G., Paskind, M., *et al.*, 1984. Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxynucleotidyl transferase in B cells. *Nature* 311, 752–755.
- Dudley, D.D., Chaudhuri, J., Bassing, C.H., Alt, F.W., 2005. Mechanism and control of V(D)J recombination versus class switch recombination: Similarities and differences. *Advances in Immunology* 86, 43–112.
- Ebert, A., Medvedovic, J., Tagoh, H., Schwickert, T.A., Busslinger, M., 2013. Control of antigen receptor diversity through spatial regulation of V(D)J recombination. *Cold Spring Harbor Symposia on Quantitative Biology* 78, 11–21.
- Elkin, S.K., Ivanov, D., Ewalt, M., *et al.*, 2005. A PHD finger motif in the C terminus of RAG2 modulates recombination activity. *Journal of Biological Chemistry* 280, 28701–28710.
- Erikson, J., Finan, J., Nowell, P.C., Croce, C.M., 1982. Translocation of immunoglobulin VH genes in Burkitt lymphoma. *Proceedings of the National Academy of Sciences of the United States of America* 79, 5611–5615.
- Farago, M., Rosenbluh, C., Tevlin, M., *et al.*, 2012. Clonal allelic predetermination of immunoglobulin-kappa rearrangement. *Nature* 490, 561–565.
- Feeney, A.J., 1992a. Comparison of junctional diversity in the neonatal and adult immunoglobulin repertoires. *International Reviews of Immunology* 8, 113–122.
- Feeney, A.J., 1992b. Predominance of VH-D-JH junctions occurring at sites of short sequence homology results in limited junctional diversity in neonatal antibodies. *Journal of Immunology* 149, 222–229.
- Fugmann, S.D., 2010. The origins of the Rag genes—from transposition to V(D)J recombination. *Seminars in Immunology* 22, 10–16.

- Fuxa, M., Skok, J., Souabni, A., *et al.*, 2004. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes and Development* 18, 411–422.
- Gapud, E.J., Sleckman, B.P., 2011. Unique and redundant functions of ATM and DNA-PKcs during V(D)J recombination. *Cell Cycle* 10, 1928–1935.
- Gellert, M., 2002. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annual Review of Biochemistry* 71, 101–132.
- van Gent, D.C., Hiom, K., Paull, T.T., Gellert, M., 1997. Stimulation of V(D)J cleavage by high mobility group proteins. *EMBO Journal* 16, 2665–2670.
- Gilfillan, S., Dierich, A., Lemeur, M., Benoist, C., Mathis, D., 1993. Mice lacking TdT: Mature animals with an immature lymphocyte repertoire. *Science* 261, 1175–1178.
- Golding, A., Chandler, S., Ballestar, E., Wolffe, A.P., Schlissel, M.S., 1999. Nucleosome structure completely inhibits *in vitro* cleavage by the V(D)J recombinase. *EMBO Journal* 18, 3712–3723.
- Gostissa, M., Alt, F.W., Chiarle, R., 2011. Mechanisms that promote and suppress chromosomal translocations in lymphocytes. *Annual Review of Immunology* 29, 319–350.
- Gu, H., Forster, I., Rajewsky, K., 1990. Sequence homologies, N sequence insertion and JH gene utilization in VHDJH joining: Implications for the joining mechanism and the ontogenetic timing of Ly1 B cell and B-CLL progenitor generation. *EMBO Journal* 9, 2133–2140.
- Hardy, R.R., Carmack, C.E., Shinton, S.A., Kemp, J.D., Hayakawa, K., 1991. Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *Journal of Experimental Medicine* 173, 1213–1225.
- Helmink, B.A., Sleckman, B.P., 2012. The response to and repair of RAG-mediated DNA double-strand breaks. *Annual Review of Immunology* 30, 175–202.
- Hewitt, S.L., Yin, B., Ji, Y., *et al.*, 2009. RAG-1 and ATM coordinate monoallelic recombination and nuclear positioning of immunoglobulin loci. *Nature Immunology* 10, 655–664.
- Hillion, S., Rochas, C., Youinou, P., Jamin, C., 2009. Signaling pathways regulating RAG expression in B lymphocytes. *Autoimmunity Reviews* 8, 599–604.
- Hiom, K., Gellert, M., 1997. A stable RAG1-RAG2-DNA complex that is active in V(D)J cleavage. *Cell* 88, 65–72.
- Igarashi, H., Gregory, S.C., Yokota, T., Sakaguchi, N., Kincade, P.W., 2002. Transcription from the RAG1 locus marks the earliest lymphocyte progenitors in bone marrow. *Immunity* 17, 117–130.
- Jaeger, S., Fernandez, B., Ferrier, P., 2013. Epigenetic aspects of lymphocyte antigen receptor gene rearrangement or 'when stochasticity completes randomness'. *Immunology* 139, 141–150.
- Jankovic, M., Nussenzweig, A., Nussenzweig, M.C., 2007. Antigen receptor diversification and chromosome translocations. *Nature Immunology* 8, 801–808.
- Jenuwein, T., Allis, C.D., 2001. Translating the histone code. *Science* 293, 1074–1080.
- Jhunjhunwala, S., van Zelm, M.C., Peak, M.M., *et al.*, 2008. The 3D structure of the immunoglobulin heavy-chain locus: Implications for long-range genomic interactions. *Cell* 133, 265–279.
- Ji, Y., Resch, W., Corbett, E., *et al.*, 2010. The *in vivo* pattern of binding of RAG1 and RAG2 to antigen receptor loci. *Cell* 141, 419–431.
- Johnson, K., Chaumeil, J., Skok, J.A., 2010. Epigenetic regulation of V(D)J recombination. *Essays in Biochemistry* 48, 221–243.
- Jones, J.M., Simkus, C., 2009. The roles of the RAG1 and RAG2 "non-core" regions in V(D)J recombination and lymphocyte development. *Archivum Immunologiae et Therapia Experimentalis* 57, 105–116.
- Jung, D., Alt, F.W., 2004. Unraveling V(D)J recombination; insights into gene regulation. *Cell* 116, 299–311.
- Komori, T., Okada, A., Stewart, V., Alt, F.W., 1993. Lack of N regions in antigen receptor variable region genes of TdT-deficient lymphocytes. *Science* 261, 1171–1175.
- Kosak, S.T., Skok, J.A., Medina, K.L., *et al.*, 2002. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* 296, 158–162.
- Krangel, M.S., 2009. Mechanics of T cell receptor gene rearrangement. *Current Opinion in Immunology* 21, 133–139.
- Kumar, V., Alt, F.W., Oksenyk, V., 2014. Functional overlaps between XLF and the ATM-dependent DNA double strand break response. *DNA Repair* 16, 11–22.
- Kuo, T.C., Schlissel, M.S., 2009. Mechanisms controlling expression of the RAG locus during lymphocyte development. *Current Opinion in Immunology* 21, 173–178.
- Kuppers, R., 2005. Mechanisms of B-cell lymphoma pathogenesis. *Nature Reviews Cancer* 5, 251–262.
- Kuppers, R., Dalla-Favera, R., 2001. Mechanisms of chromosomal translocations in B cell lymphomas. *Oncogene* 20, 5580–5594.
- Kwon, J., Imbalzano, A.N., Matthews, A., Oettinger, M.A., 1998. Accessibility of nucleosomal DNA to V(D)J cleavage is modulated by RSS positioning and HMG1. *Molecular Cell* 2, 829–839.
- Lee, G.S., Neiditch, M.B., Salus, S.S., Roth, D.B., 2004. RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. *Cell* 117, 171–184.
- Lewis, S.M., Agard, E., Suh, S., Czyzyk, L., 1997. Cryptic signals and the fidelity of V(D)J joining. *Molecular and Cellular Biology* 17, 3125–3136.
- Li, G., Alt, F.W., Cheng, H.L., *et al.*, 2008. Lymphocyte-specific compensation for XLF/cernunnos end-joining functions in V(D)J recombination. *Molecular Cell* 31, 631–640.
- Litman, G.W., Rast, J.P., Fugmann, S.D., 2010. The origins of vertebrate adaptive immunity. *Nature Reviews Immunology* 10, 543–553.
- Liu, Y., Subrahmanyam, R., Chakraborty, T., Sen, R., Desiderio, S., 2007. A plant homeodomain in RAG-2 that binds Hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. *Immunity* 27, 561–571.
- Lucas, J.S., Zhang, Y., Dudko, O.K., Murre, C., 2014. 3D trajectories adopted by coding and regulatory DNA elements: First-passage times for genomic interactions. *Cell* 158, 339–352.
- Mansson, R., Hultquist, A., Luc, S., *et al.*, 2007. Molecular evidence for hierarchical transcriptional lineage priming in fetal and adult stem cells and multipotent progenitors. *Immunity* 26, 407–419.
- Marculescu, R., Le, T., Simon, P., Jaeger, U., Nadel, B., 2002. V(D)J-mediated translocations in lymphoid neoplasms: A functional assessment of genomic instability by cryptic sites. *Journal of Experimental Medicine* 195, 85–98.
- Matheson, L.S., Corcoran, A.E., 2012. Local and global epigenetic regulation of V(D)J recombination. *Current Topics in Microbiology and Immunology* 356, 65–89.
- Matthews, A.G., Kuo, A.J., Ramon-Maiques, S., *et al.*, 2007. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* 450, 1106–1110.
- Mostoslavsky, R., Alt, F.W., Rajewsky, K., 2004. The lingering enigma of the allelic exclusion mechanism. *Cell* 118, 539–544.
- Mostoslavsky, R., Singh, N., Tenzen, T., *et al.*, 2001. Asynchronous replication and allelic exclusion in the immune system. *Nature* 414, 221–225.
- Mullighan, C.G., Downing, J.R., 2009. Global genomic characterization of acute lymphoblastic leukemia. *Seminars in Hematology* 46, 3–15.
- Niehues, T., Perez-Becker, R., Schuetz, C., 2010. More than just SCID—the phenotypic range of combined immunodeficiencies associated with mutations in the recombinase activating genes (RAG) 1 and 2. *Clinical Immunology* 135, 183–192.
- Nussenzweig, A., Nussenzweig, M.C., 2010. Origin of chromosomal translocations in lymphoid cancer. *Cell* 141, 27–38.
- Nutt, S.L., Kee, B.L., 2007. The transcriptional regulation of B cell lineage commitment. *Immunity* 26, 715–725.
- Oettinger, M.A., Schatz, D.G., Gorka, C., Baltimore, D., 1990. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* 248, 1517–1523.
- Okada, T.A., Comings, D.E., 1979. Higher order structure of chromosomes. *Chromosoma* 72, 1–14.
- Onozawa, M., Aplan, P.D., 2012. Illegitimate V(D)J recombination involving nonantigen receptor loci in lymphoid malignancy. *Genes, Chromosomes and Cancer* 51, 525–535.
- Osipovich, O., Cobb, R.M., Oestreich, K.J., *et al.*, 2007. Essential function for SWI-SNF chromatin-remodeling complexes in the promoter-directed assembly of Tcrb genes. *Nature Immunology* 8, 809–816.
- Osipovich, O.A., Subrahmanyam, R., Pierce, S., Sen, R., Oltz, E.M., 2009. Cutting edge: SWI/SNF mediates antisense IgH transcription and locus-wide accessibility in B cell precursors. *Journal of Immunology* 183, 1509–1513.
- Papaemmanuil, E., Rapado, I., Li, Y., *et al.*, 2014. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nature Genetics* 46, 116–125.
- Pernis, B., Chiappino, G., Kelus, A.S., Gell, P.G., 1965. Cellular localization of immunoglobulins with different allotypic specificities in rabbit lymphoid tissues. *Journal of Experimental Medicine* 122, 853–876.
- Raghavan, S.C., Swanson, P.C., Wu, X., Hsieh, C.L., Lieber, M.R., 2004. A non-B-DNA structure at the Bcl-2 major breakpoint region is cleaved by the RAG complex. *Nature* 428, 88–93.
- Reynaud, D., Demarco, I.A., Reddy, K.L., *et al.*, 2008. Regulation of B cell fate commitment and immunoglobulin heavy-chain gene rearrangements by Ikaros. *Nature Immunology* 9, 927–936.
- Robey, I.F., Peterson, M., Horwitz, M.S., *et al.*, 2004. Terminal deoxynucleotidyltransferase deficiency decreases autoimmune disease in diabetes-prone nonobese diabetic mice and lupus-prone MRL-Fas(lpr) mice. *Journal of Immunology* 172, 4624–4629.
- Roldan, E., Fuxa, M., Chong, W., *et al.*, 2005. Locus 'decontraction' and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. *Nature Immunology* 6, 31–41.

- Ropars, V., Drevet, P., Legrand, P., *et al.*, 2011. Structural characterization of filaments formed by human Xrcc4-Cernunnos/XLF complex involved in nonhomologous DNA end-joining. *Proceedings of the National Academy of Sciences of the United States of America* 108, 12663–12668.
- Roth, D.B., 2003. Restraining the V(D)J recombinase. *Nature Reviews Immunology* 3, 656–666.
- Roth, D.B., Gellert, M., 2000. New guardians of the genome. *Nature* 404, 823–825.
- Rothenberg, E.V., Taghon, T., 2005. Molecular genetics of T cell development. *Annual Review of Immunology* 23, 601–649.
- Sadofsky, M.J., Hesse, J.E., Gellert, M., 1994. Definition of a core region of RAG-2 that is functional in V(D)J recombination. *Nucleic Acids Research* 22, 1805–1809.
- Sadofsky, M.J., Hesse, J.E., McBlane, J.F., Gellert, M., 1993. Expression and V(D)J recombination activity of mutated RAG-1 proteins. *Nucleic Acids Research* 21, 5644–5650.
- Sayegh, C.E., Jhunjhunwala, S., Riblet, R., Murre, C., 2005. Visualization of looping involving the immunoglobulin heavy-chain locus in developing B cells. *Genes & Development* 19, 322–327.
- Schatz, D.G., Ji, Y., 2011. Recombination centres and the orchestration of V(D)J recombination. *Nature Reviews Immunology* 11, 251–263.
- Schatz, D.G., Oettinger, M.A., Baltimore, D., 1989. The V(D)J recombination activating gene, RAG-1. *Cell* 59, 1035–1048.
- Schatz, D.G., Swanson, P.C., 2011. V(D)J recombination: Mechanisms of initiation. *Annual Review of Genetics* 45, 167–202.
- Singh, H., Pongubala, J.M., 2006. Gene regulatory networks and the determination of lymphoid cell fates. *Current Opinion in Immunology* 18, 116–120.
- Skok, J.A., Gisler, R., Novatchkova, M., *et al.*, 2007. Reversible contraction by looping of the Tcr α and Tcr β loci in rearranging thymocytes. *Nature Immunology* 8, 378–387.
- Sleckman, B.P., Oltz, E.M., 2012. Preparing targets for V(D)J recombinase: Transcription paves the way. *Journal of Immunology* 188, 7–9.
- Steinel, N.C., Lee, B.S., Tubbs, A.T., *et al.*, 2013. The ataxia telangiectasia mutated kinase controls Igkappa allelic exclusion by inhibiting secondary V κ to-J κ rearrangements. *Journal of Experimental Medicine* 210, 233–239.
- Strahl, B.D., Allis, C.D., 2000. The language of covalent histone modifications. *Nature* 403, 41–45.
- Stubbington, M.J., Corcoran, A.E., 2013. Non-coding transcription and large-scale nuclear organisation of immunoglobulin recombination. *Current Opinion in Genetics & Development* 23, 81–88.
- Swanson, P.C., 2002. Fine structure and activity of discrete RAG-HMG complexes on V(D)J recombination signals. *Molecular and Cellular Biology* 22, 1340–1351.
- Tillman, R.E., Wooley, A.L., Hughes, M.M., Khor, B., Sleckman, B.P., 2004. Regulation of T-cell receptor beta-chain gene assembly by recombination signals: The beyond 12/23 restriction. *Immunological Reviews* 200, 36–43.
- Towbin, B.D., Gonzalez-Sandoval, A., Gasser, S.M., 2013. Mechanisms of heterochromatin subnuclear localization. *Trends in Biochemical Sciences* 38, 356–363.
- Tsujiimoto, Y., Gorham, J., Cossman, J., Jaffe, E., Croce, C.M., 1985. The t(14;18) chromosomal translocations involved in B cell neoplasms result from mistakes in V(D)J joining. *Science* 229, 1390–1393.
- Vera, G., Rivera-Munoz, P., Abramowski, V., *et al.*, 2012. Cernunnos deficiency reduces thymocyte lifespan and alters the T cell repertoire in mice and humans. *Molecular and Cellular Biology* 33, 701–711.
- Verma-Gaur, J., Torkamani, A., Schaffer, L., *et al.*, 2012. Noncoding transcription within the Igh distal V(H) region at PAIR elements affects the 3D structure of the Igh locus in pro-B cells. *Proceedings of the National Academy of Sciences of the United States of America* 109, 17004–17009.
- Vettermann, C., Schlissel, M.S., 2010. Allelic exclusion of immunoglobulin genes: Models and mechanisms. *Immunological Reviews* 237, 22–42.
- Villa, A., Marrella, V., Rucci, F., Notarangelo, L.D., 2008. Genetically determined lymphopenia and autoimmune manifestations. *Current Opinion in Immunology* 20, 318–324.
- de Villartay, J.P., 2009. V(D)J recombination deficiencies. *Advances in Experimental Medicine and Biology* 650, 46–58.
- West, K.I., Singha, N.C., De Ioannes, P., *et al.*, 2005. A direct interaction between the RAG2 C terminus and the core histones is required for efficient V(D)J recombination. *Immunity* 23, 203–212.
- Woodbine, L., Gennery, A.R., Jeggo, P.A., 2014. Reprint of "The clinical impact of deficiency in DNA non-homologous end-joining". *DNA Repair* 17, 9–20.
- Yancopoulos, G.D., Alt, F.W., 1985. Developmentally controlled and tissue-specific expression of unrearranged VH gene segments. *Cell* 40, 271–281.
- Zha, S., Jiang, W., Fujiwara, Y., *et al.*, 2011. Ataxia telangiectasia-mutated protein and DNA-dependent protein kinase have complementary V(D)J recombination functions. *Proceedings of the National Academy of Sciences of the United States of America* 108, 2028–2033.
- Zhang, L., Reynolds, T.L., Shan, X., Desiderio, S., 2011. Coupling of V(D)J recombination to the cell cycle suppresses genomic instability and lymphoid tumorigenesis. *Immunity* 34, 163–174.

Further Reading

- Corfe, S.A., Paige, C.J., 2012. The many roles of IL-7 in B cell development; mediator of survival, proliferation and differentiation. *Seminars in Immunology* 24, 198–208.
- Di Noia, J.M., Neuberger, M.S., 2007. Molecular mechanisms of antibody somatic hypermutation. *Annual Review of Biochemistry* 76, 1–22.
- Keim, C., Kazadi, D., Rothschild, G., Basu, U., 2013. Regulation of AID, the B-cell genome mutator. *Genes & Development* 27, 1–17.
- Lai, A.Y., Kondo, M., 2008. T and B lymphocyte differentiation from hematopoietic stem cell. *Seminars in Immunology* 20, 207–212.
- Matthews, A.J., Zheng, S., Dimenna, L.J., Chaudhuri, J., 2014. Regulation of immunoglobulin class-switch recombination: Choreography of noncoding transcription, targeted DNA deamination, and long-range DNA repair. *Advances in Immunology* 122, 1–57.
- Wong, G.W., Zuniga-Pflucker, J.C., 2010. Gammadelta and alphabeta T cell lineage choice: Resolution by a stronger sense of being. *Seminars in Immunology* 22, 228–236.

T Cell Memory to Viral Infections

P Klenerman and JR Fergusson, University of Oxford, Oxford, UK

© 2016 Elsevier Inc. All rights reserved.

Background

The ability of the immune system to develop long-lived memory is critical for its function. Perhaps the most important development in this respect was the observation by Edward Jenner that exposure to cowpox led to acquired resistance to smallpox, laying the foundations for vaccination. Since that time much has been learned about the mechanisms underlying protective immunologic memory – in other words immunity. Much of the recent drive to understand this in more depth has been the need to develop vaccines against more complex pathogens, especially persistent virus infections such as human immunodeficiency virus (HIV) and hepatitis C virus (HCV). Defining the rules that govern the control of T cell memory populations is a central part of this effort. Although we still lack effective vaccines for HIV and HCV, much has been learned about the nature of T cell responses against such pathogens, as well as those in apparently successful immune responses. These responses will be described in this article, together with discussion of some of the unanswered questions that remain.

Memory Formation

T lymphocytes develop in the thymus, until they are released as mature T cells which circulate in the bloodstream, migrating through successive lymph nodes scanning antigen presenting cells (APCs) within for presentation of their cognate antigen. At this stage of differentiation, until they have encountered antigen, they are described as 'naïve.' Upon recognition of antigen, a specific T cell is activated to undergo clonal expansion and differentiation into a pool of 'effector' cells that perform specialised functions. Classically, these functions have been divided based upon coreceptor expression, with CD8 + T cells differentiating into cytotoxic T lymphocytes (CTLs) with the ability to kill infected cells through the release of cytotoxic molecules such as perforin and granzymes, and CD4 + T cells differentiating into T helper (Th) cells of which there are several lineages, including Th1, Th2, Th17 and T regulatory (Treg) cells. These lineages are defined based upon the predominant cytokines produced and the characteristic transcription factor expression, and are thought to differentially develop according to innate immune cues in the milieu during antigen presentation that communicate the type of infection to which the antigen is associated. For example, the recently identified Th17 lineage is characterized by production of interleukin (IL)-17 and expression of the master transcription factor RAR-related orphan receptor (ROR) γ t, and is thought to develop in the presence of innate cues triggered by extracellular bacterial and fungal infections (Cosmi *et al.*, 2008). IL-17 acts on tissue cells to produce mediators that increase neutrophil recruitment and inflammation – an appropriate response to extracellular pathogens.

Following clonal expansion, and generally after removal of the invading microorganism, T lymphocytes undergo extensive contraction. Numbers of pathogen-reactive cells reduce rapidly due to activation induced cell death (AICD) before reaching a plateau, with a percentage of effector cells remaining at increased numbers compared to their naïve precursors (see Figure 1). This pool of long-lasting antigen-specific T cells form a defining attribute of the adaptive immune system, namely their ability to develop immunological memory. Memory is typically defined by the persistence of increased numbers of antigen-reactive lymphocytes after removal of antigenic stimulation (Lau *et al.*, 1994; Murali-Krishna *et al.*, 1999; Swain, 1999) together with the capacity to expand and differentiate rapidly upon secondary challenge (Lau *et al.*, 1994; Veiga-Fernandes *et al.*, 2000). This definition and analysis becomes more complex in situations where the pathogen may persist – even at relatively low levels or at specific sites.

Phenotypes of Memory T Cells

Human memory T cells can be identified by a switch in isoform from CD45RA to CD45RO, yet with reversion to CD45RA possible, especially in the CD8 + T cell subset. Further phenotypic analysis of human memory cells also revealed a dichotomous heterogeneity based upon expression of the lymph node-homing chemokine receptor CCR7 (Sallusto *et al.*, 1999). Expression of CCR7 by 'central memory' T cells (T_{CM}), together with CD62L, enables them to home to the lymph node where they can proliferate to initiate secondary responses. T_{CM} cells may produce few cytokines except IL-2. On the other hand, CCR7 negative 'effector memory' T cells (T_{EM}) typically produce high levels of several cytokines, including IL-4 and interferon (IFN)- γ and circulate through non-lymphoid tissues. Characterization based upon expression of

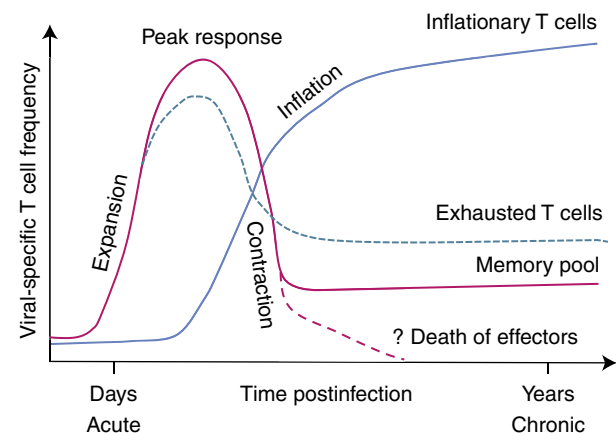


Figure 1 Schematic graph representing the different developmental pathways of memory T cells following viral infection.

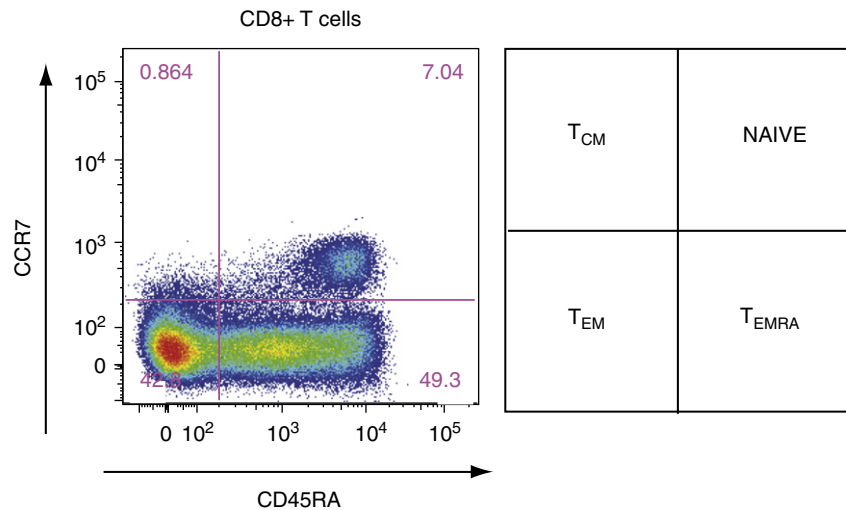


Figure 2 Flow cytometry plot (left) showing CD8 + T cells divided into memory subsets (right) based upon expression of CCR7 and CD45RA. T_{CM}, central memory; T_{EM}, effector memory; T_{EMRA}, terminally differentiated memory.

both CD45RA and CCR7 divides T cells into three CD4 + subsets (CCR7 + CD45RA + naïve, CCR7 + CD45RA – and CCR7 – CD45RA – memory) and four CD8 + subsets (with an additional CCR7 – CD45RA +), with re-expression of CD45RA associated with terminal differentiation (T_{EMRA}). These subsets are shown in **Figure 2**, although very few T_{CM} CD8 + T cells are present in the circulation.

There has been some discussion regarding the differentiation pathways of these distinct memory subsets. Originally it was thought that activation of naïve T cells generated T_{CM} cells, which then provided a source of T_{EM}, with more immediate effector function (Sallusto *et al.*, 1999; **Figure 3(a)**). Conversely, studies of adoptive transfer with T cell receptor (TCR)-transgenic mice suggested a linear differentiation pathway, from naïve to effector to T_{EM} cell, which may subsequently convert to T_{CM} (Wherry *et al.*, 2003; Bouneaud *et al.*, 2005) (**Figure 3(b)**). Yet, it has been suggested that this conversion from T_{EM} to T_{CM} is an artefact of the TCR-transgenic system, with T_{EM} cells instead representing a stable memory phenotype (Marzo *et al.*, 2005). Still further studies have suggested that the two subsets are in fact distinct and independent (Baron *et al.*, 2003), whereas others suggest that an effector stage may not even be required to generate memory cells (Manjunath *et al.*, 2001; **Figure 3(c)**).

Another key question is whether those cells which develop into memory cells – so called ‘memory precursors’ – can be distinguished from those which will die during the contraction phase. Indeed, expression of the IL-7 receptor (IL-7R) (Kaeche *et al.*, 2003), together with downregulated expression of CD25 (Kalia *et al.*, 2010) and KLRG1 (Sarkar *et al.*, 2008), is thought to distinguish those effector CD8 + T cells which will differentiate into memory cells. These phenotypic changes are accompanied by profound changes within the cell in relation to metabolism, with the switch from acute expansion of effector cells to memory formation being associated with a switch away from the aerobic glycolysis required for rapid proliferation.

Tissue-Resident Memory

In addition to the two T cell memory subsets already described, a population of memory T cells that, unlike T_{EM}, do not recirculate, has been identified in certain peripheral tissues, including the skin, gut, and lung. This population is stably maintained within the respective tissue, independently of their circulating counterparts, and have hence been described as ‘tissue-resident memory’ T cells (T_{RM}). Like innate immune cells, T_{RM} cells reside in tissues that form a barrier to invading pathogens, and can provide potent protection against re-infection (Jiang *et al.*, 2012). Transcriptional studies have demonstrated CD8 + T_{RM} to be distinct from circulating memory CD8 + T cells (Wakim *et al.*, 2012; Mackay *et al.*, 2013), yet few definitive markers of T_{RM} have been identified to date. However, expression of CD103 (integrin α E), which is expressed on memory CD8 + T cells within many tissues yet low on circulating T cells, identifies T_{RM} within human mucosal tissues such as gut and lung (Sathaliyawala *et al.*, 2013), and has hence been used as a marker of T_{RM} cells. CD103 binds E-cadherin, a molecule expressed on epithelial cells, and therefore may function to retain T_{RM} cells within peripheral tissues. However, it should be noted that CD103 has also been described as a marker of regulatory T cells, and therefore it is the combination of CD103 expression together with their tissue localization that identifies T_{RM} cells. In a similar way, CD69 has also been used as a marker of T_{RM} (Sathaliyawala *et al.*, 2013; Turner *et al.*, 2014), but can also be expressed by activated T cells within the circulation. CD69 expression leads to down-regulation of the sphingosine-1 phosphate receptor 1, which mediates egress from tissue to the lymph and blood. The differentiation pathway of this unique T_{RM} population is largely unknown, with questions such as whether there are distinct T_{RM} precursors within the effector cell pool, or whether T_{EM} cells can differentiate into T_{RM} once recruited into tissues remaining to be answered, although it is

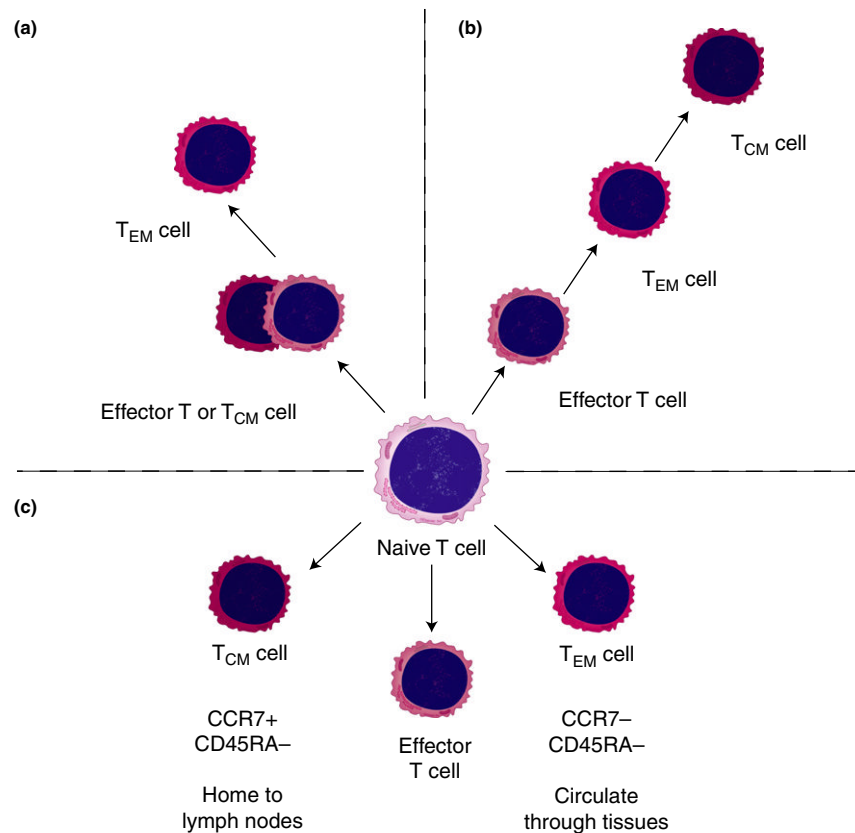


Figure 3 Potential pathways of differentiation of memory T cells.

clear that the specific tissue microenvironment plays an important role.

Different Phenotypes Characterize Different Viral Infections

Although the above model of expansion, contraction, and long-lived memory does apply to many viral infections, it is by no means the only pathway of differentiation available to T cells (see Figure 1). Most of the work in this area is based on studies of CD8⁺ T cells. This is partly for technical reasons, since the expansion of these cells following virus infection tends to be very large, and thus the populations are relatively easier to track and characterize. The process of tracking T cells was transformed in the mid-1990s by the introduction of major histocompatibility complex (MHC) peptide tetramer technology. 'Tetramers,' or more generally 'multimers,' are constructs based on soluble MHC-peptide complexes generated *in vitro* according to the antigen of interest. Such soluble complexes can be bound as multimers around fluorescent molecules such as tagged streptavidin, and their avidity is such that they will bind and be internalized by T cells bearing the appropriate TCR for that complex (see Figure 4). This allows the cells to be tracked using flow cytometry, and additionally allows the surface phenotype and intracellular molecules to be readily assessed on the same antigen-specific cell (Altman *et al.*, 1996).

The advent of this technology revealed the diversity of T cells targeting different viruses, especially those with some degree of persistence, such as the herpesvirus group, HIV and HCV. One way of visualizing this diversity is along a spectrum based on antigen re-encounter, maturation, or linear differentiation over time (Appay *et al.*, 2008). In those virus infections where exposure is very limited, the responses contract as above and long-lived memory is maintained in an antigen-independent manner (central memory pools dominating). In settings where antigen exposure is more prolonged, the predominant cell seen tends to acquire 'effector' memory characteristics. This includes loss of lymph node homing molecules such as CCR7 and CD62L, loss of co-stimulator dependence through CD28 and CD27 down-regulation, and acquisition of effector functions through up-regulation of perforin, for example. At one end of this spectrum is Epstein-Barr virus (EBV), which has a relatively limited maturation, while at the other is cytomegalovirus (CMV) in mouse and man, which has the most extreme of the differentiated phenotypes (see 'memory inflation' below).

The idea that these memory pools represent a spectrum has been reinforced recently by the advent of a new technology – cytometry by time of flight (CyTOF) (Newell *et al.*, 2012). This technique is based on the use of antibodies conjugated to metal isotopes, which are bound to the cell surface as in normal fluorescence-based methods. Cells are passed singly into a chamber where they are vaporized and their metal ion cargo released for analysis by mass spectrometry. The resulting

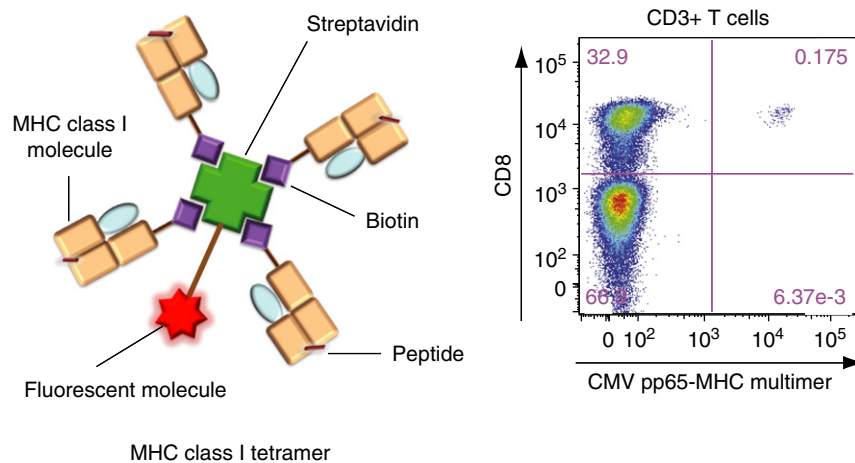


Figure 4 Diagram of MHC molecules complexed around fluorescently-tagged streptavidin to form an MHC class I Tetramer (left) used to identify viral-specific cells by flow cytometry (right).

information allows analysis of multiple antibodies (30–60 in many cases), which can all be independently assessed, overcoming the limitations of fluorescence caused by spectral overlap. By analysis of multiple memory markers, it is clearly shown that T cell memory phenotypes exist as a broad range, or continuum, without a clearly defined barrier between them. Increasing resolution of this type will ultimately allow better description of these memory pools, beyond the somewhat arbitrary markers currently used.

Immune Exhaustion

One further feature of antiviral memory populations, which has emerged from the study of persistent viruses – notably lymphocytic choriomeningitis virus (LCMV) in the mouse – is the phenomenon of ‘exhaustion.’ This is most evident in certain strains of the virus such as LCMV Clone 13 and DOCILE strain, which, when inoculated intravenously at high dose, lead to an overwhelming infection which can persist. This is accompanied by down-regulation of the T cell response to the virus, most dramatically seen as the emergence of CD8⁺ T cell responses that are poorly functional, with up-regulation of multiple inhibitory receptors. Since LCMV itself is non-cytopathic, this down-regulation of T cell reactivity has been rationalized as a mechanism induced to limit immune-mediated pathology. The most well-studied of these inhibitory molecules is programmed death 1 or PD-1. Blockade of PD-1 can relieve the inhibition of CD8⁺ T cells and lead to increased control over the viral load (Barber *et al.*, 2006). Downstream of PD-1, down-regulation of the critical transcription factor T-bet (TBX21) is responsible for loss of many effector functions (Kao *et al.*, 2011).

Immune exhaustion has also been observed in human persistent virus infections such as HIV, HCV, and hepatitis B virus (HBV). In these infections, high levels of PD-1 have been observed on virus-specific T cell populations, as well as other inhibitory receptors such as CD160, Lag-3, and TIM3. This would be consistent with the marked down-regulation of antiviral population size against HCV in chronic infection, compared to maintenance of classic memory populations in the subset of patients who control in-

fection after acute exposure. One further phenomenon serving to limit the efficacy of such antiviral populations is the emergence of T cell escape mutations within the virus (Phillips *et al.*, 1991). Variable viruses, where RNA replication is error-prone and proof-reading is limited, generate multiple mutations which can serve to impair T cell recognition via loss of peptide binding to MHC, lack of processing at the proteasome, or loss of TCR recognition. Under immune selection such variants come to dominate and lead to both loss of immune control and loss of TCR-triggered T cell maturation.

Memory Inflation

At the other end of the spectrum, as mentioned above, persistent infection with CMV in both mouse and man is associated with the development of responses which show markers associated with ‘late’ differentiation. This phenomenon is also associated with numerical expansion of these ‘effector memory’ pools, and accumulation of large antiviral populations in organs such as lung and liver, as well as in blood. In the mouse model, where T cells can be tracked after infection, it is apparent that such populations may not dominate initially, but expand gradually over time, hence the phenomenon has been termed memory ‘inflation’ (O’Hara *et al.*, 2012). The same phenomenon has been observed in responses against other persistent viruses, including viruses such as those of the parvovirus family, without a classic latency program. Interestingly the phenomenon is restricted to limited subsets of T cells responding to epitopes that appear to be presented at sufficient levels during the memory phase. Responses to many other epitopes, including those that elicit dominant responses during acute infection, undergo a classical contraction and generate memory pools indistinguishable from those seen against acute nonpersistent infections. It is believed that these peptides are not presented efficiently during memory formation, either through competition, or due to the requirement for the immunoproteasome for processing, which is restricted to certain cell types under noninflammatory conditions.

Atypical Memory

The ability to track memory populations using MHC-peptide tetramers has revealed some unexpected responses in addition to those described above. The first of these is the identification of T cell responses with apparent memory characteristics in naïve individuals. These have been described in the case of HIV and HCV in humans, and to a range of pathogens in mice. This is thought to be due to the natural cross-reactive capacity of TCRs, which cannot be uniquely specific to a single MHC-peptide complex. In the case of HIV, peptides have been described from the gut microbiome that could act to prime HIV specific T cells in non-HIV exposed individuals. In the case of HCV, a cross-reactive peptide from influenza has been suggested, although the avidity of this interaction is rather lower than that of most conventional TCR-peptides. In both cases, the question of whether such antigen-experienced pools can contribute to protection or immunopathology remains open, but it is certainly clear that such cross-reactivity, or heterologous immunity, even at low level, could serve to mold the TCR pool in unexposed individuals and/or T cell memory after infection (Welsh *et al.*, 2010).

The reverse situation has also been observed. Very sensitive assays can identify naïve precursors in unexposed individuals at extremely low frequencies (around 1 in a million T lymphocytes) (Alanio *et al.*, 2010). Naïve T cell pools can also be identified in some individuals with persistent infections such as HCV. Additionally, there are populations of T cells with 'naïve' markers which can be observed after infection or vaccination (CD45RA⁺, CCR7⁺), which have not been well characterized but clearly are antigen responsive, have expanded in response to antigen and remain functional (Barnes *et al.*, 2012).

One recent surprise in this field is that a substantial proportion of CD8⁺ T cells (about 1 in 6) in human blood do not represent a typical antiviral response, but rather an innate type of lymphocyte termed a mucosal associated invariant T cells or MAIT cell. These T cells possess many markers typical of 'effector' memory pools, and home to organs, especially the liver, where they represent around 50% of the CD8⁺ T cell pool. However, their TCRs are not diverse like conventional T cells, but rather semi-invariant (possessing a restricted Va7.2-Ja33 rearrangement and limited beta chain pairing), and this TCR is bacterially responsive rather than antiviral (recognizing a bacterial metabolic product presented on the conserved MR1 molecule rather than MHC) (Le Bourhis *et al.*, 2013). Yet, MAIT cells possess the ability to produce Interferon (IFN)- γ , a potent antiviral cytokine, in a TCR-independent manner to a combination of IL-12 and IL-18 (Ussher *et al.*, 2014), suggesting a further role in antiviral defense. Furthermore, MAIT cells are specifically depleted in HIV infection (Cosgrove *et al.*, 2013; Leeansyah *et al.*, 2013) and have been implicated in the control of influenza through the production of IL-17 (Hamada *et al.*, 2009).

Impact on Vaccine Design

As indicated above, much of the incentive to understand T cell memory in more detail stems from a desire to develop novel vaccines, not only for complex viral targets such as HIV, but also for cancers, where induction of specific CD8⁺ T cell pools can be particularly effective. In the case of the latter, the

observation that PD-1 and similar inhibitory molecules can restrict effector functions during a persistent viral stimulus has been very informative in guiding novel therapies which block inhibitory functions *in vivo* (e.g., biologic agents blocking PD-1 and CTLA4 for use in human cancer immunotherapy) (Coulzin-Frankel, 2013).

The main aim of modern vaccine design is to induce specific immune responses that are long-lived and protective. In the case of T cell responses, this initiative is often hampered by a lack of knowledge of the 'correlates of protection.' This compares to protection mediated by antibodies, where reaching a certain titer of neutralizing antibody in the blood can be linked readily to protection against challenge. A number of features have been proposed which could contribute to protection, including: the frequency of antigen-specific cells, their functionality (notably secretion of multiple cytokines, including IL-2), and acquisition of markers of long-term memory. For some virus infections such as HIV, which infect mucosally, it has also been shown that induction of high frequencies of effector memory T cell populations – including those induced by novel CMV vectors – can be linked to immediate protective function (Hansen *et al.*, 2011). However, this remains a challenging area and each viral infection may require its own approach. Historically, the best vaccines have typically been live attenuated strains, where induction of T cells mirrors most closely the natural setting.

In the absence of a live attenuated virus, current technology to prime CD8⁺ T cell responses *in vivo* typically involve delivery of viral antigen from DNA – in the form of DNA vaccination, or delivered as a viral vector. A range of approaches have been used effectively in mice, but in humans currently the most efficient vectors appear to be those based on viral vectors such as primate adenoviruses. These are highly immunogenic, although in the case of HIV vaccines some of these approaches have been halted by lack of efficacy and concerns that immune activation induced by such vectors may increase the risk of HIV infection.

Conclusions and Future Directions

The understanding of T cell memory has developed enormously from a phenomenological approach to a much better molecular understanding of the transcriptional drivers and metabolic changes associated with generation of long-lived antigen-specific lymphocyte pools. Further work is required in the future to link these data – which nicely describe the cellular changes required for memory maintenance and function – to protection against challenge. However, as illustrated in this article, the rules that govern protection may differ subtly between different virus challenges and thus it may prove difficult to provide a one-size-fits-all model for protective memory. Furthermore, much of the focus to date has been on CD8⁺ T cell pools and much more detail is required on the parallel CD4⁺ T cell memory pools to provide a complete picture. If, as current data suggest, such antiviral (or vaccine-induced) memory for both CD4⁺ and CD8⁺ T cells exists as a continuum, then perhaps defining the protective qualities of T cells at different regions within this continuum could be an achievable aim.

See also: Cellular Immunology: Cell–Cell Interactions and the Immune System: T Cell Receptor Triggering

References

- Alanio, C., Lemaitre, F., Law, H.K.W., Hasan, M., Albert, M.L., 2010. Enumeration of human antigen-specific naive CD8 + T cells reveals conserved precursor frequencies. *Blood* 115 (18), 3718–3725.
- Altman, J.D., Moss, P.A., Goulder, P.J., *et al.*, 1996. Phenotypic analysis of antigen-specific T lymphocytes. *Science* (New York, NY) 274 (5284), 94–96.
- Appay, V., van Lier, R.A., Sallusto, F., Roederer, M., 2008. Phenotype and function of human T lymphocyte subsets: Consensus and issues. *Cytometry. Part A: The Journal of the International Society for Analytical Cytology* 73 (11), 975–983.
- Barber, D.L., Wherry, E.J., Masopust, D., *et al.*, 2006. Restoring function in exhausted CD8 T cells during chronic viral infection. *Nature* 439 (7077), 682–687.
- Barnes, E., Folgori, A., Capone, S., *et al.*, 2012. Novel adenovirus-based vaccines induce broad and sustained T cell responses to HCV in man. *Science Translational Medicine* 4 (115), 1–11.
- Baron, V., Bouneaud, C., Cumano, A., *et al.*, 2003. The repertoires of circulating human CD8(+) central and effector memory T cell subsets are largely distinct. *Immunity* 18 (2), 193–204.
- Bouneaud, C., Garcia, Z., Kourilsky, P., Pannetier, C., 2005. Lineage relationships, homeostasis, and recall capacities of central- and effector-memory CD8 T cells *in vivo*. *Journal of Experimental Medicine* 201 (4), 579–590.
- Cosgrove, C., Ussher, J.E., Rauch, A., *et al.*, 2013. Early and nonreversible decrease of CD161++/MAIT cells in HIV infection. *Blood* 121 (6), 951–961.
- Cosmi, L., De Palma, R., Santarlasci, V., *et al.*, 2008. Human interleukin 17–producing cells originate from a CD161 + CD4 + T cell precursor. *Journal of Experimental Medicine* 205 (8), 1903–1916.
- Couzin-Frankel, J., 2013. Cancer Immunotherapy. *Science* 342 (December), 1432–1433.
- Hamada, H., Garcia-Hernandez, M.D.L.L., Reome, J.B., *et al.*, 2009. Tc17, a unique subset of CD8 T cells that can protect against lethal influenza challenge. *Journal of Immunology* (Baltimore, MD: 1950) 182 (6), 3469–3481.
- Hansen, S.G., Ford, J.C., Lewis, M.S., *et al.*, 2011. Profound early control of highly pathogenic SIV by an effector memory T-cell vaccine. *Nature* 473 (7348), 523–527.
- Jiang, X., Clark, R.A., Liu, L., *et al.*, 2012. Skin infection generates non-migratory memory CD8 + T(RM) cells providing global skin immunity. *Nature* 483 (7388), 227–231.
- Kaech, S.M., Tan, J.T., Wherry, E.J., *et al.*, 2003. Selective expression of the interleukin 7 receptor identifies effector CD8 T cells that give rise to long-lived memory cells. *Nature Immunology* 4 (12), 1191–1198.
- Kalia, V., Sarkar, S., Subramaniam, S., *et al.*, 2010. Prolonged interleukin-2/alpha expression on virus-specific CD8 + T cells favors terminal-effector differentiation *in vivo*. *Immunity* 32 (1), 91–103.
- Kao, C., Oestreich, K.J., Paley, M.A., *et al.*, 2011. Transcription factor T-bet represses expression of the inhibitory receptor PD-1 and sustains virus-specific CD8 + T cell responses during chronic infection. *Nature Immunology* 12 (7), 663–671.
- Lau, L.L., Jamieson, B.D., Somasundaram, T., Ahmed, R., 1994. Cytotoxic T-cell memory without antigen. *Nature* 369, 648–652.
- Leeansyah, E., Ganesh, A., Quigley, M.F., *et al.*, 2013. Activation, exhaustion, and persistent decline of the antimicrobial MR1-restricted MAIT-cell population in chronic HIV-1 infection. *Blood* 121 (7), 1124–1135.
- Le Bourhis, L., Mburu, Y.K., Lantz, O., 2013. MAIT cells, surveyors of a new class of antigen: Development and functions. *Current Opinion in Immunology* 25 (2), 174–180.
- Mackay, L.K., Rahimpour, A., Ma, J.Z., *et al.*, 2013. The developmental pathway for CD103(+)CD8 + tissue-resident memory T cells of skin. *Nature Immunology* 14 (12), 1294–1301.
- Manjunath, N., Shankar, P., Wan, J., *et al.*, 2001. Effector differentiation is not prerequisite for generation of memory cytotoxic T lymphocytes. *Journal of Clinical Investigation* 108 (6), 805–806.
- Marzo, A.L., Klonowski, K.D., Le Bon, A., *et al.*, 2005. Initial T cell frequency dictates memory CD8 + T cell lineage commitment. *Nature Immunology* 6 (8), 793–799.
- Murali-Krishna, K., Lau, L.L., Sambhara, S., *et al.*, 1999. Persistence of memory CD8 T cells in MHC class I-deficient mice. *Science* 286 (5443), 1377–1381.
- Newell, E.W., Sigal, N., Bendall, S.C., Nolan, G.P., Davis, M.M., 2012. Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8 + T cell phenotypes. *Immunity* 36 (1), 142–152.
- O'Hara, G.A., Welten, S.P.M., Klenerman, P., Arens, R., 2012. Memory T cell inflation: Understanding cause and effect. *Trends in Immunology* 33 (2), 84–90.
- Phillips, R.E., Rowland-Jones, S., Nixon, D.F., *et al.*, 1991. Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature* 354, 453–459.
- Sallusto, F., Lenig, D., Forster, R., Lipp, M., Lanzavecchia, A., 1999. Two subsets of memory T lymphocytes with distinct homing potentials. *Nature* 401, 708–712.
- Sarkar, S., Kalia, V., Haining, W.N., *et al.*, 2008. Functional and genomic profiling of effector CD8 T cell subsets with distinct memory fates. *Journal of Experimental Medicine* 205 (3), 625–640.
- Sathaliyawala, T., Kubota, M., Yudanin, N., *et al.*, 2013. Distribution and compartmentalization of human circulating and tissue-resident memory T cell subsets. *Immunity* 38 (1), 187–197.
- Swain, S.L., 1999. Class II-independent generation of CD4 memory T cells from effectors. *Science* 286 (5443), 1381–1383.
- Turner, D.L., Bickham, K.L., Thome, J.J., *et al.*, 2014. Lung niches for the generation and maintenance of tissue-resident memory T cells. *Mucosal Immunology* 7 (3), 501–510.
- Ussher, J.E., Bilton, M., Attwod, E., *et al.*, 2014. CD161++CD8 + T cells, including the MAIT cell subset, are specifically activated by IL-12 + IL-18 in a TCR-independent manner. *European Journal of Immunology* 44, 195–203.
- Weiga-Fernandes, H., Walter, U., Bourgeois, C., McLean, A., Rocha, B., 2000. Response of naive and memory CD8 + T cells to antigen stimulation *in vivo*. *Nature Immunology* 1 (1), 47–53.
- Wakim, L.M., Woodward-Davis, A., Liu, R., *et al.*, 2012. The molecular signature of tissue resident memory CD8 T cells isolated from the brain. *Journal of Immunology* (Baltimore, MD: 1950) 187 (7), 3462–3471.
- Welsh, R.M., Che, J.W., Brehm, M.A., Selin, L.K., 2010. Heterologous immunity between viruses. *Immunological Reviews* 235 (1), 244–266.
- Wherry, E.J., Teichgräber, V., Becker, T.C., *et al.*, 2003. Lineage relationship and protective immunity of memory CD8 T cell subsets. *Nature Immunology* 4 (3), 225–234.

Further Reading

- Ahmed, R., Bevan, M.J., Reiner, S.L., Fearon, D.T., 2009. The precursors of memory: Models and controversies. *Nature Reviews. Immunology* 9 (9), 662–668.
- Angelosanto, J., Wherry, E.J., 2010. Transcription factor regulation of CD8 + T-cell memory and exhaustion. *Immunological Reviews* 236, 167–175.
- Shin, H., Iwasaki, A., 2013. Tissue-resident memory T cells. *Immunological Reviews* 255 (1), 165–181.
- Su, L., Davis, M.M., 2013. Antiviral memory phenotype T cells in unexposed adults. *Immunological Reviews* 255 (1), 95–109.
- van der Windt, G., Pearce, E.L., 2012. Metabolic switching and fuel choice during T-cell differentiation and memory development. *Immunological Reviews* 249 (1), 27–42.
- Zinkernagel, R.M., 1996. Immunology taught by viruses. *Science* 271 (5246), 173–178.

Nuclear Factor of Activated T Cells and Tolerance

F Macian, Albert Einstein College of Medicine, Bronx, NY, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The Nuclear Factor of activated T cells (NFAT) family of transcription factors encompasses five different proteins. Four of them are regulated by calcium signaling and termed NFATc1 (also known as NFAT2), NFATc2 (NFAT1 or NFATp), NFATc3 (NFAT4 or NFATx), and NFATc4 (NFAT3), while the activation of a fifth NFAT protein, NFAT5 (TonEBP), responds predominantly to hypertonic stress (Hogan *et al.*, 2003; Macian, 2005). At least one NFAT protein is expressed in any cell type or tissue to regulate a number of cell processes that range from developmental programs, cell differentiation or activation to modulation of responses to the extracellular environment (Crabtree and Olson, 2002). In T cells NFAT proteins were initially described as essential factors required to upregulate the expression of cytokine genes induced following T cell activation. Further studies started to uncover the wide range of cellular functions in which NFAT plays a regulatory role, which include almost all aspects of T cell biology. This article will focus on the different ways in which NFAT controls immune tolerance by regulating distinct aspects of the programs that control T cell activation.

NFAT Domain Structure

All calcium-regulated NFAT proteins contain three distinct domains. An N-terminal NFAT-homology region constitutes the regulatory domain of NFAT. This domain is moderately conserved among the different NFAT proteins and contains the regions that modulate NFAT activation, including phosphorylation sites, calcineurin docking sites, nuclear localization signals, and nuclear export signals; as well as the major transactivation domain. Following the N-terminal domain lies the Rel-homology region, which shares structural homology with the DNA-binding domain of Rel proteins and is highly conserved among all NFAT family members. This Rel-homology region not only contains the DNA-binding domain, which determines the binding specificity of NFAT, but also most of the residues involved in supporting many of the interactions of NFAT proteins with their transcriptional copartners, as well as the amino acids that maintain the contacts that allow for the formation of NFAT dimers. The C-terminus is not conserved among NFAT proteins and has been shown in some family members to be involved in mediating interactions with other transcription factors (Rao *et al.*, 1997; Macian, 2005). As opposed to the calcium-regulated NFAT proteins, NFAT5 is activated in response to osmotic stress and forms an obligated homodimer, much like the Rel proteins. Nevertheless, the DNA-binding domain of NFAT5 has 43% homology with the DNA-binding domain of the other NFAT family members and recognizes similar binding motifs on the DNA (Lopez-Rodriguez *et al.*, 2001). However, as no role has been attributed so far to NFAT5 in the regulation of immune tolerance, this article will focus on the calcium-regulated NFAT

factors expressed in T cells, which have been implicated in regulation of T cell activation and tolerance.

Regulation of NFAT Activation

The activation of all calcium-dependent NFAT proteins is regulated by their phosphorylation status (Okamura *et al.*, 2000). NFAT is maintained in the cytosol as a heavily phosphorylated protein. Phosphorylation occurs at a series of serine residues localized in the N-terminal regulatory domain and it is mediated by cytosolic kinases, such as casein kinase 1 (CK1), that maintain the phosphorylation of cytosolic NFAT (Okamura *et al.*, 2004); and export nuclear kinases, including CK1 and glycogen synthase kinase 3 (GSK3), which requires prior priming of NFAT through phosphorylation mediated by protein kinase A (PKA) or the dual-specificity tyrosine phosphorylation regulated kinases (DYRK) (Beals *et al.*, 1997; Gwack *et al.*, 2006). In the cytosol, NFAT and NFAT-kinases associate in a supramolecular complex that includes among other proteins several members of the importin beta family of nuclear import proteins, the leucine-rich repeat kinase 2 (LRKK2) and the calmodulin-binding protein IQGAP1. The noncoding RNAs NRON and lincRNA act as scaffolds for that complex (Sharma *et al.*, 2011; Willingham *et al.*, 2005).

Several calcium-coupled receptors can activate NFAT in different cell types. In T cells it is the engagement of the T cell receptor (TCR) that constitutes the main signaling event that leads to the activation of NFAT. As a result of recognition of antigen by the TCR, activation of phospholipase C gamma (PLC γ) causes the hydrolyzation of phosphatidylinositol-4,5-bisphosphate into Inositol 1,4,5-trisphosphate (IP3) and diacylglycerol. IP3 binds its receptor in the endoplasmic reticulum (ER), which causes exit of calcium from the ER into the cytosol. Depletion of the ER intracellular calcium stores is sensed by stromal interaction molecules 1 and 2 (STIM1 and 2) at the ER membrane, which move then closer to the plasma membrane to activate calcium-release activated calcium (CRAC) channels, which are formed by oligomers of ORAI1 proteins and are ultimately responsible for the store-operated calcium entry of extracellular calcium into the cytosol (Feske, 2007; Feske *et al.*, 2006). Docking of calcium-bound calmodulin on calcineurin causes activation of this protein-phosphatase that is responsible for the concerted dephosphorylation of multiple serine residues in the N-terminal domain of NFAT. This causes a conformational change in NFAT that leads to the exposure of one or more nuclear localization signals to allow the nuclear translocation of NFAT (Okamura *et al.*, 2000). It has been recently shown that activation-induced generation of reactive oxygen species plays also an important role in the regulation of the amplitude of the TCR-induced calcium signaling that mediates NFAT activation (Sena *et al.*, 2013). Once in the nucleus NFAT proteins bind sites in the promoters or enhancers of genes where they will form either cooperative complexes with other transcription

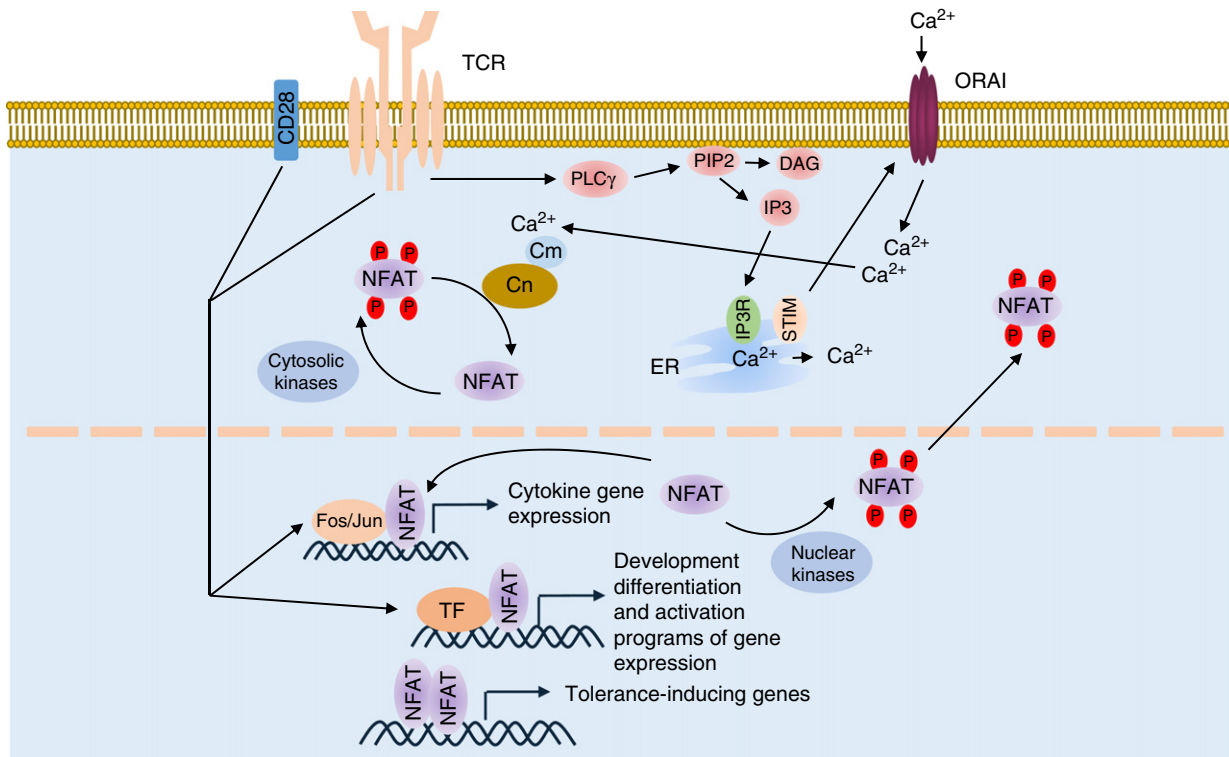


Figure 1 Regulation of nuclear factor of activated T cells (NFAT) activation. NFAT is maintained in the cytosol as heavily phosphorylated protein through the activity of cytosolic kinases and export nuclear kinases. Engagement of the T cell receptor (TCR) activates phospholipase C gamma ($PLC\gamma$) that converts phosphatidylinositol-4,5-bisphosphate (PIP2) into Inositol 1,4,5-trisphosphate (IP3) and diacylglycerol (DAG). IP3 binds its receptor (IP3R) in the endoplasmic reticulum (ER) which depletes this intracellular calcium store. This is sensed by stromal interaction molecules (STIM) that activate the calcium-release activated calcium channels that are formed by ORAI proteins. This leads to entry of extracellular calcium into the cell and activation of the phosphatase calcineurin (Cn) by calcium-bound calmodulin (Cm) to dephosphorylate multiple serine residues in the N-terminal domain of NFAT. This causes a conformational change that exposes nuclear localization signals and nuclear translocation of NFAT. In the nucleus NFAT proteins bind sites in the promoters or enhancers of genes where they will form complexes with other transcription factors (TF) to direct the expression of genes responsible for the development, differentiation or activation of T cells. NFAT may also form dimers to regulate the expression of genes that induce T cell inactivation and lead to the establishment of tolerance.

factors or NFAT dimers, to induce the expression of genes involved in the regulation of T cell development, differentiation, activation or tolerance (Feske *et al.*, 2001; Muller and Rao, 2010; Figure 1). Additional levels of regulation have also been reported to control NFAT activity. For instance, NFAT sumoylation regulates NFAT subnuclear localization and transcriptional activity, while poly-ADP-ribose polymerase (PARP)-dependent ADP-ribosylation increases NFAT DNA-binding activity (Olabisi *et al.*, 2008; Terui *et al.*, 2004).

A unique mechanism is also in place to regulate the expression of the short isoform of NFATc1 (NFATc1/A). The expression of most NFAT proteins, including the long isoforms of NFATc1, is constitutive in T cells. However, a positive autoregulatory loop controls the transcription of NFATc1/A, which accounts for the increased expression of this protein in activated T cells (Chuvpilo *et al.*, 2002).

NFAT-Regulated Transcription

The transcriptional activity of NFAT proteins is modulated by the partners with which they interact. NFAT does not bind DNA nor activate gene transcription as an independent

monomer, relying on interaction with other transcription partners to form active complexes. This property positions NFAT as a major hub for signal integration. Coordinated activation of calcium signaling with other signaling pathways to regulate specific cell functions is frequently attained in T cells through cooperation of calcium-activated NFAT with other transcriptional factors activated in response to specific signals (Macian, 2005). For instance, during T cell activation heterotrimeric NFAT/Activator protein 1 (AP-1) complexes, formed by an NFAT molecule interacting with Fos and Jun, are required for efficient induction of the expression of many cytokine genes that contain NFAT/AP-1 composite binding sites on their promoter or enhancer regions (Jain *et al.*, 1993; Macian *et al.*, 2000). This complex allows for integration of signals mediated by the TCR and costimulatory receptors, which lead to concomitant engagement of calcium signaling and the protein kinase C (PKC)/mitogen activated protein kinase (MAPK) signaling pathways, required for the activation of NFAT and AP-1, respectively. Similar mechanism of signal integration are responsible for the formation of complexes of NFAT with many other transcription factors, such as IRF4, GATA3 or p300, to regulate a wide range of cell functions that include, among others, T helper cell differentiation, cell

migration, cell cycle regulation, and cell death and survival (Avni *et al.*, 2002; Garcia-Rodriguez and Rao, 1998; Macian, 2005; Rengarajan *et al.*, 2002a).

As mentioned above the DNA-binding domain of NFAT proteins show structural identity with the DNA-binding domain of Rel proteins (Serfling *et al.*, 2004). Similar to members of the Rel family of transcription factors, that form obligated dimers to bind DNA, NFAT proteins have also been shown to have the ability to dimerize and bind to specific sites that contain two tandem NFAT-binding sites separated by a spacer formed by one or two bases (Jin *et al.*, 2003). These dimers direct the expression of distinct programs of gene expression that, as opposed to the classically described integrative nature of NFAT-regulated transcription, depend solely on the activation of NFAT without requiring other signals. As it will be discussed below, NFAT dimers play an important role in the expression of genes that induce the establishment of T cell tolerance (Soto-Nieves *et al.*, 2009).

Functions of NFAT in T cells

Although this article focuses on the role of NFAT in the regulation of T cell tolerance, the functions of this family of transcription factors in the modulation of the immune response are not limited to T cells and extend to the regulation of the activation and differentiation of many other cells that are involved in the generation of innate and adaptive immune responses, including B cells, dendritic cells, NKT cells, NK cells, macrophages, mast cells, or neutrophils (Muller and Rao, 2010).

The first NFAT protein to be cloned, NFATc2, was initially described as a transcription factor that bound the *Il2* gene promoter during T cell activation to, in cooperation with Fos and Jun, induce the expression of that cytokine gene (Jain *et al.*, 1992). Subsequent studies have identified NFAT proteins as ultimate targets of the signaling pathways that control the expression of many cytokine genes in T cells, therefore becoming central players in the regulation of T cell activation. Initial studies analyzing the phenotypes of mice with deletions of single NFAT proteins revealed smaller effects than expected on the responses of mature T cells to activation, suggesting certain level of functional redundancy (Ranger *et al.*, 1998a; Xanthoudakis *et al.*, 1996). The existence of this redundancy became more evident when mice deficient in two or more NFAT proteins were generated (Peng *et al.*, 2001; Rengarajan *et al.*, 2002b). For instance, the loss of NFATc1 and NFATc2 in T cells resulted in a profound defect on the ability of those cells to produce cytokines or generate cytolytic activity following TCR engagement (Peng *et al.*, 2001).

By controlling the expression of lineage specific cytokines as well as by acting in functional cooperation with transcription factors that influence lineage commitment, NFAT proteins also participate in the regulation of T helper cell differentiation. Sustained IL-4 production in activated naïve NFATc2 or NFATc2 + NFATc3-deficient T cells skews their differentiation toward a T helper type 2 (T_H2) phenotype, whereas T cells with a deletion of the gene encoding for NFATc1 appear to tilt the differentiation process toward T_H1 (Kiani *et al.*, 1997; Ranger *et al.*, 1998a; Rengarajan *et al.*, 2002b). NFAT has

also been shown to regulate the expression of IFN γ , IL-17 or IL-21, and therefore modulate T_H1 and T_H17 differentiation (Kiani *et al.*, 2001; Kim *et al.*, 2005; Liu *et al.*, 2004). Furthermore, cooperative interactions between NFAT and transcription factors that mediate T helper cell lineage commitment or direct NFAT-dependent regulation the expression of such transcription factors, such as GATA3, Tbet or ROR γ t, also underlie the central role of NFAT in the regulation of T helper cell differentiation (Avni *et al.*, 2002; Hermann-Kleiter and Baier, 2010; Kim *et al.*, 2014).

This positive effect of NFAT on T cell activation and effector differentiation is thought to account for the immunosuppressive activity of the commonly used inhibitors of calcineurin, such as cyclosporine A, which, by blocking calcium-calcineurin dependent NFAT dephosphorylation, prevent NFAT nuclear translocation and the activation of NFAT-dependent gene expression in T cells (Jain *et al.*, 1993).

NFAT and T Cell Tolerance

Studies performed in the last 10 years have revealed that NFAT proteins not only regulate the induction of positive effector T cell functions but they also play an essential role in the induction of tolerance (Baine *et al.*, 2009; Serfling *et al.*, 2006). The ability of NFAT to control two opposite functional programs in T cells resides in specific properties of these transcription factors that include their ability to regulate thymocyte selection, direct the differentiation of regulatory T cells and their ability to cooperate with many other transcription factors and form alternative transcriptional complexes activated under specific conditions in T cells to direct the expression of programs of gene expression that mediate a diverse set of responses including self-antigen induced T cell inactivation (Figure 2).

Thymocyte Development and Positive Selection

Several NFAT proteins are expressed early during thymocyte development, and regulate key steps in thymocyte maturation. Survival signals are transduce through NFAT activation and are required to ensure double negative to double positive transition (Adachi *et al.*, 2000; Oukka *et al.*, 1998). This is accomplished not only through canonical calcium-calcineurin-mediated activation of NFAT but also by other pathways, which include IL-7-receptor induced activation of NFATc1 or IKK-regulated expression of NFAT5. Activation of those NFAT proteins guarantees the survival of double negative thymocytes and their progression to CD4 + CD8 + double positive cells (Berga-Bolanos *et al.*, 2013; Patra *et al.*, 2013). The important role of NFAT proteins during thymocyte positive selection is also underscored by studies that have analyzed the consequences of the deletion of calcineurin or specific NFAT proteins in thymocytes, which result in marked defects in positive selection (Neilson *et al.*, 2004; Oukka *et al.*, 1998; Ranger *et al.*, 1998b). In some of these models, such as in the NFATc3- or the double NFATc2/NFATc3-deficient mice, peripheral mature T cells show a hyperactivated state that, while it may respond in part to defects in mechanisms of peripheral

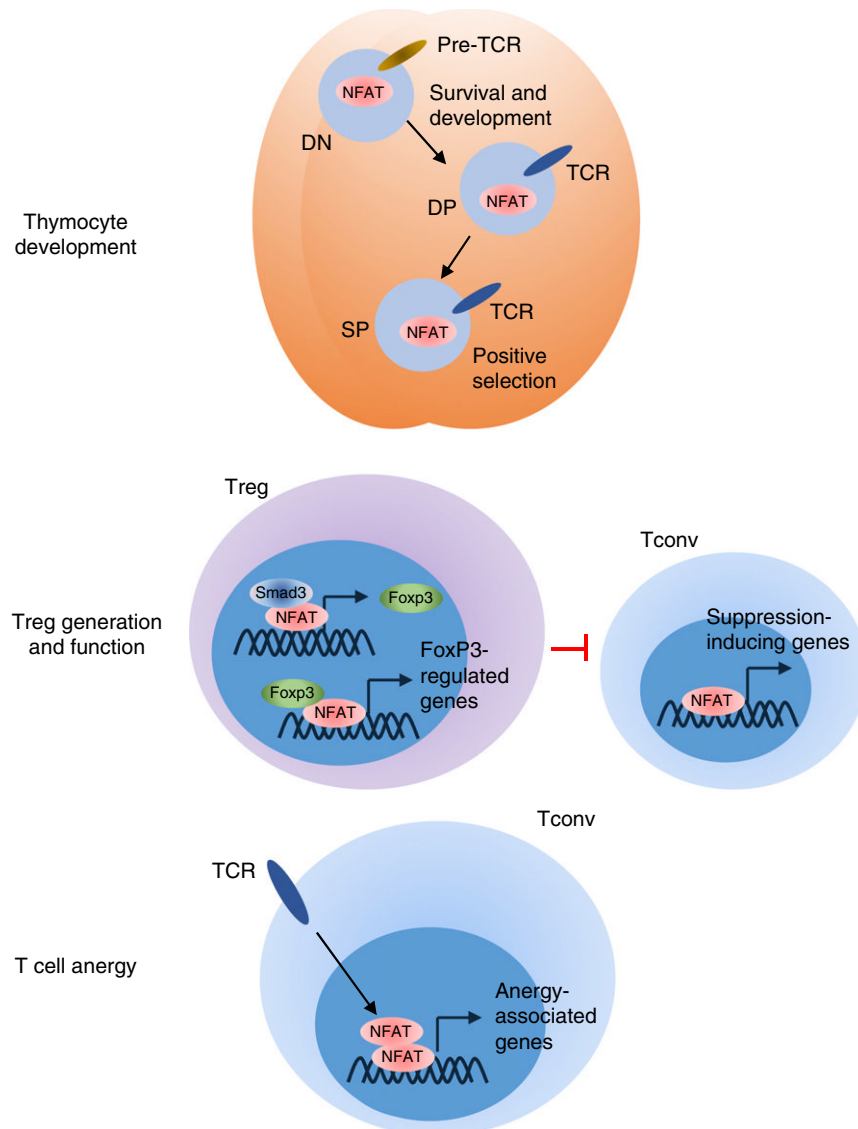


Figure 2 Nuclear factor of activated T cells (NFAT) regulates T cell tolerance at different levels. NFAT proteins are required to ensure double negative (DN) thymocyte survival and the transition to double CD4⁺ CD8⁺ (DP) stage. NFAT also controls positive selection contributing to shaping the peripheral T cell receptor (TCR) repertoire. Expression of Foxp3 in regulatory T cells (Treg) is regulated by NFAT, which forms complexes with Smad3 to direct the expression of that gene. NFAT also cooperates with FoxP3 to regulate the expression of many of the FoxP3-controlled genes that define the Treg phenotype. Moreover, expression of NFAT-dependent genes induced by the presence of Treg during conventional T cell (Tconv) activation ensures efficient inhibition of effector functions in Tconv. NFAT dimers formed in response to suboptimal stimulation in CD4⁺ or CD8⁺ T cells are responsible for the expression of inhibitory gene that cause the induction of anergy in Tconv.

tolerance, it is likely also due to altered thymocyte selection in those mice.

Regulatory T Cells

Regulatory T cells (Treg) constitute a complex population of T cells that have a fundamental role in the maintenance of immune tolerance. Natural or thymic Treg, which develop in the thymus, and peripheral Treg, generated from naïve CD4⁺ T cells in the periphery, have the ability to suppress the activation of self-reactive T cells and several other cell types, and by doing so, prevent the development of uncontrolled

immune reactions against self (Wing and Sakaguchi, 2010). FoxP3 is the transcription factor that acts as the master regulator of Treg development and its function and expression are regulated by NFAT. NFAT forms part of the complex of transcription factors that bind to the enhancers that direct the expression of FoxP3 during Treg development (Tone *et al.*, 2008). Consequently mice bearing T cells which either fail to efficiently activate store-operated calcium entry or have deletions in the genes that encode for one or more NFAT proteins, have reduced levels of Tregs (Oh-Hora *et al.*, 2008; Vaeth *et al.*, 2012). NFAT involvement in the regulation of Treg function extends to at least two other aspects. First, NFATc2

has been shown to form complexes with FoxP3 to either repress or activate the expression of different genes, and, therefore, cooperate with FoxP3 in establishing a Treg-specific program of gene expression (Wu *et al.*, 2006). Furthermore, the suppressive activity of Tregs on effector CD4+ T cells appears to be dependent on the ability of Tregs to induce the expression of NFAT-dependent genes, which inhibit effector cytokine production and cell proliferation in suppressed T cells (Bopp *et al.*, 2005; Shin *et al.*, 2014).

T Cell Anergy

In response to suboptimal stimulation, T cells do not engage a 'full activation' gene expression program but instead upregulate the expression of a series of genes that encode proteins that have the ability to either inhibit signaling pathways downstream of the TCR or directly suppress the transcription of cytokine genes. As a consequence cells become anergic and respond poorly to new encounters with antigen (Macian *et al.*, 2002). The expression of this program of gene expression is directly or indirectly dependent on NFATc2 and includes genes such as those encoding for the ubiquitin ligases Grail, Cbl-b or Itch, which through targeting of specific signaling intermediates induce the degradation or functional inactivation of different signaling pathways downstream of the TCR; caspase 3, that selectively cleaves signaling proteins in anergic cells without inducing apoptosis; diacylglycerol kinase, which inhibits diacylglycerol-dependent signaling in T cells; the histone deacetylase Sirt1, which inhibits AP-1 activation; or the transcriptional repressors Ikaros and TLE4, which directly suppress IL-2 and IFN γ expression, respectively (Valdor and Macian, 2013). The expression of the transcription factors Egr2 and 3 is also upregulated in an NFAT-dependent manner in anergic T cells, where they directly regulate the expression of several anergy inducing factors, including Cbl-b and Sirt1 (Safford *et al.*, 2005; Zhang *et al.*, 2009). As discussed above, the ability of NFAT to direct two opposite programs, T cell activation and tolerance, results from the coordinated activation or lack of thereof of NFAT-transcriptional partners. This will determine which program will be engaged in each situation. During a full activation, likely to occur when pathogen-derived antigens are presented by activated dendritic cells, engagement of TCR and costimulatory receptors, especially CD28, will cause the activation of, among other transcription factors, NFAT and AP-1. These form a high-affinity complex that binds composite sites present in promoters and/or enhancers of many genes expressed in activated T cells. On the other hand, suboptimal activation due to reduced expression of costimulatory ligands on dendritic cells that have not been activated by danger signals (as it is likely to happen when self-antigens are processed) and/or the engagement of coinhibitory receptors in T cells and/or the presence of Treg, leads to the activation of NFAT without concomitant full AP-1 activation. In this context NFAT forms dimers that bind to NFAT-dimer sites present in the regulatory regions of anergy-associated genes (Soto-Nieves *et al.*, 2009). These dimers are not formed in activated T cells due to the much higher affinity of NFAT for Fos and Jun to form NFAT/AP-1 complexes, which prevents the expression of anergy-associated genes in activated T cells (Baine *et al.*, 2009). Although most studies performed so far have tried to decipher

how NFAT controls the induction of anergy in CD4+ T cells, evidence also supports that NFAT activity is required to induce tolerance in CD8+ T cells and B cells (Barrington *et al.*, 2006; Fehr *et al.*, 2010).

Recently, the physiological relevance of this process has been demonstrated in the context of cancer-induced tolerance. Induction of T cell tolerance is a common mechanism of immune evasion in many tumors. Studies using NFATc2 or Egr2 deficient mouse models have shown that T cells from those mice become resistant to tumor-induced tolerance, which results in efficient control of tumor growth (Abe *et al.*, 2012; Zheng *et al.*, 2012). The activation of the NFAT-dependent anergy-inducing program of gene expression in tumor infiltrating T cells likely results from the existence of inefficient transient contacts between dendritic and T cells in the tumor microenvironment, as has been shown for CD8+ T cells (Marangoni *et al.*, 2013). These fast contacts would result in maintained NFAT activation, due to delayed nuclear export, in the absence of AP-1, which requires more stable contacts to become activated.

Conclusions

- NFAT proteins play central roles not only in the regulation of positive programs of T cell differentiation and activation but also in the control of T cell tolerance.
- Integration of multiple signaling inputs results from the ability of NFAT proteins to form transcriptional complexes with different transcription factors, which allows NFAT to regulate specific programs of gene expression through partner selection and control T cell development, differentiation, activation, and tolerance.
- Members of the NFAT family of transcription factors participate at different levels in the regulation of immune tolerance, including the regulation of thymocyte selection, the differentiation of Treg, and the induction of T cell anergy.
- The multiple, and frequently opposite, roles that NFAT proteins play in the regulation of the adaptive immune response highlight the importance of the need to design targeted therapies that can affect certain functions without altering others.

See also: Nucleic Acid Synthesis/Breakdown: Transcription: Distant Activation of Transcription by Enhancers; Eukaryotic Transcriptional Regulation

References

- Abe, B.T., Shin, D.S., Mocholi, E., Macian, F., 2012. NFAT1 supports tumor-induced anergy of CD4(+) T cells. *Cancer Research* 72, 4642–4651.
- Adachi, S., Amasaki, Y., Miyatake, S., Arai, N., Iwata, M., 2000. Successive expression and activation of NFAT family members during thymocyte differentiation. *Journal of Biological Chemistry* 275, 14708–14716.
- Avni, O., Lee, D., Macian, F., *et al.*, 2002. T(H) cell differentiation is accompanied by dynamic changes in histone acetylation of cytokine genes. *Nature Immunology* 3, 643–651.
- Baine, I., Abe, B.T., Macian, F., 2009. Regulation of T-cell tolerance by calcium/NFAT signaling. *Immunological Reviews* 231, 225–240.

- Barrington, R.A., Borde, M., Rao, A., Carroll, M.C., 2006. Involvement of NFAT1 in B cell self-tolerance. *Journal of Immunology* 177, 1510–1515.
- Beals, C.R., Sheridan, C.M., Turck, C.W., Gardner, P., Crabtree, G.R., 1997. Nuclear export of NF-ATc enhanced by glycogen synthase kinase-3. *Science* 275, 1930–1934.
- Berga-Bolanos, R., Alberdi, M., Buxade, M., Aramburu, J., Lopez-Rodriguez, C., 2013. NFAT5 induction by the pre-T-cell receptor serves as a selective survival signal in T-lymphocyte development. *Proceedings of the National Academy of Sciences of the United States of America* 110, 16091–16096.
- Bopp, T., Palmethofer, A., Serfling, E., *et al.*, 2005. NFATc2 and NFATc3 transcription factors play a crucial role in suppression of CD4+ T lymphocytes by CD4+ CD25+ regulatory T cells. *Journal of Experimental Medicine* 201, 181–187.
- Chuvpilo, S., Jankevics, E., Tyrstin, D., *et al.*, 2002. Autoregulation of NFATc1/A expression facilitates effector T cells to escape from rapid apoptosis. *Immunity* 16, 881–895.
- Crabtree, G.R., Olson, E.N., 2002. NFAT signaling: Choreographing the social lives of cells. *Cell* 109, S67–S79.
- Fehr, T., Lucas, C.L., Kurtz, J., *et al.*, 2010. A CD8 T cell-intrinsic role for the calcineurin-NFAT pathway for tolerance induction *in vivo*. *Blood* 115, 1280–1287.
- Feske, S., 2007. Calcium signalling in lymphocyte activation and disease. *Nature Reviews Immunology* 7, 690–702.
- Feske, S., Giltman, J., Dolmetsch, R., Staudt, L.M., Rao, A., 2001. Gene regulation mediated by calcium signals in T lymphocytes. *Nature Immunology* 2, 316–324.
- Feske, S., Gwack, Y., Prakriya, M., *et al.*, 2006. A mutation in Orai1 causes immune deficiency by abrogating CRAC channel function. *Nature* 441, 179–185.
- Garcia-Rodriguez, C., Rao, A., 1998. Nuclear factor of activated T cells (NFAT)-dependent transactivation regulated by the coactivators p300/CREB-binding protein (CBP). *Journal of Experimental Medicine* 187, 2031–2036.
- Gwack, Y., Sharma, S., Nardone, J., *et al.*, 2006. A genome-wide *Drosophila* RNAi screen identifies DYRK-family kinases as regulators of NFAT. *Nature* 441, 646–650.
- Hermann-Kleiter, N., Baier, G., 2010. NFAT pulls the strings during CD4+ T helper cell effector functions. *Blood* 115, 2989–2997.
- Hogan, P.G., Chen, L., Nardone, J., Rao, A., 2003. Transcriptional regulation by calcium, calcineurin, and NFAT. *Genes & Development* 17, 2205–2232.
- Jain, J., McCaffrey, P.G., Miner, Z., *et al.*, 1993. The T-cell transcription factor NFATp is a substrate for calcineurin and interacts with Fos and Jun. *Nature* 365, 352–355.
- Jain, J., McCaffrey, P.G., Valge-Archer, V.E., Rao, A., 1992. Nuclear factor of activated T cells contains Fos and Jun. *Nature* 356, 801–804.
- Jin, L., Sliz, P., Chen, L., *et al.*, 2003. An asymmetric NFAT1 dimer on a pseudo-palindromic kappa B-like DNA site. *Nature Structural & Molecular Biology* 10, 807–811.
- Kiani, A., Garcia-Cozar, F.J., Habermann, I., *et al.*, 2001. Regulation of interferon-gamma gene expression by nuclear factor of activated T cells. *Blood* 98, 1480–1488.
- Kiani, A., Viola, J.P., Lichtman, A.H., Rao, A., 1997. Down-regulation of IL-4 gene transcription and control of Th2 cell differentiation by a mechanism involving NFAT1. *Immunity* 7, 849–860.
- Kim, H.P., Korn, L.L., Gamero, A.M., Leonard, W.J., 2005. Calcium-dependent activation of interleukin-21 gene expression in T cells. *Journal of Biological Chemistry* 280, 25291–25297.
- Kim, K.D., Srikanth, S., Tan, Y.V., *et al.*, 2014. Calcium signaling via Orai1 is essential for induction of the nuclear orphan receptor pathway to drive Th17 differentiation. *Journal of Immunology* 192, 110–122.
- Liu, X.K., Lin, X., Gaffen, S.L., 2004. Crucial role for nuclear factor of activated T cells in T cell receptor-mediated regulation of human interleukin-17. *Journal of Biological Chemistry* 279, 52762–52771.
- Lopez-Rodriguez, C., Aramburu, J., Jin, L., *et al.*, 2001. Bridging the NFAT and NF-kappaB families: NFAT5 dimerization regulates cytokine gene transcription in response to osmotic stress. *Immunity* 15, 47–58.
- Macian, F., 2005. NFAT proteins: Key regulators of T-cell development and function. *Nature Reviews Immunology* 5, 472–484.
- Macian, F., Garcia-Cozar, F., Im, S.H., *et al.*, 2002. Transcriptional mechanisms underlying lymphocyte tolerance. *Cell* 109, 719–731.
- Macian, F., Garcia-Rodriguez, C., Rao, A., 2000. Gene expression elicited by NFAT in the presence or absence of cooperative recruitment of Fos and Jun. *EMBO Journal* 19, 4783–4795.
- Marangoni, F., Murooka, T.T., Manzo, T., *et al.*, 2013. The transcription factor NFAT exhibits signal memory during serial T cell interactions with antigen-presenting cells. *Immunity* 38, 237–249.
- Muller, M.R., Rao, A., 2010. NFAT, immunity and cancer: a transcription factor comes of age. *Nature Reviews Immunology* 10, 645–656.
- Neilson, J.R., Winslow, M.M., Hur, E.M., Crabtree, G.R., 2004. Calcineurin B1 is essential for positive but not negative selection during thymocyte development. *Immunity* 20, 255–266.
- Oh-Hora, M., Yamashita, M., Hogan, P.G., *et al.*, 2008. Dual functions for the endoplasmic reticulum calcium sensors STIM1 and STIM2 in T cell activation and tolerance. *Nature Immunology* 9, 432–443.
- Okamura, H., Aramburu, J., Garcia-Rodriguez, C., *et al.*, 2000. Concerted dephosphorylation of the transcription factor NFAT1 induces a conformational switch that regulates transcriptional activity. *Molecular Cell* 6, 539–550.
- Okamura, H., Garcia-Rodriguez, C., Martinson, H., *et al.*, 2004. A conserved docking motif for CK1 binding controls the nuclear localization of NFAT1. *Molecular and Cellular Biology* 24, 4184–4195.
- Olabis, O.A., Soto-Nieves, N., Nieves, E., *et al.*, 2008. Regulation of transcription factor NFAT by ADP-ribosylation. *Molecular and Cellular Biology* 28, 2860–2871.
- Oukka, M., Ho, I.C., de la Brousse, F.C., *et al.*, 1998. The transcription factor NFAT4 is involved in the generation and survival of T cells. *Immunity* 9, 295–304.
- Patra, A.K., Avots, A., Zahedi, R.P., *et al.*, 2013. An alternative NFAT-activation pathway mediated by IL-7 is critical for early thymocyte development. *Nature Immunology* 14, 127–135.
- Peng, S.L., Gerth, A.J., Ranger, A.M., Glimcher, L.H., 2001. NFATc1 and NFATc2 together control both T and B cell activation and differentiation. *Immunity* 14, 13–20.
- Ranger, A.M., Hodge, M.R., Gravalles, E.M., *et al.*, 1998a. Delayed lymphoid repopulation with defects in IL-4-driven responses produced by inactivation of NF-ATc. *Immunity* 8, 125–134.
- Ranger, A.M., Oukka, M., Rengarajan, J., Glimcher, L.H., 1998b. Inhibitory function of two NFAT family members in lymphoid homeostasis and Th2 development. *Immunity* 9, 627–635.
- Rao, A., Luo, C., Hogan, P.G., 1997. Transcription factors of the NFAT family: Regulation and Function. *Annual Review of Immunology* 15, 707–747.
- Rengarajan, J., Mowen, K.A., McBride, K.D., *et al.*, 2002a. Interferon regulatory factor 4 (IRF4) interacts with NFATc2 to modulate interleukin 4 gene expression. *Journal of Experimental Medicine* 195, 1003–1012.
- Rengarajan, J., Tang, B., Glimcher, L.H., 2002b. NFATc2 and NFATc3 regulate T(H)2 differentiation and modulate TCR-responsiveness of naive T(H) cells. *Nature Immunology* 3, 48–54.
- Safford, M., Collins, S., Lutz, M.A., *et al.*, 2005. Egr-2 and Egr-3 are negative regulators of T cell activation. *Nature Immunology* 6, 472–480.
- Sena, L.A., Li, S., Jairaman, A., *et al.*, 2013. Mitochondria are required for antigen-specific T cell activation through reactive oxygen species signaling. *Immunity* 38, 225–236.
- Serfling, E., Berberich-Siebelt, F., Avots, A., *et al.*, 2004. NFAT and NF-kappaB factors—the distant relatives. *International Journal of Biochemistry & Cell Biology* 36, 1166–1170.
- Serfling, E., Klein-Hessling, S., Palmethofer, A., *et al.*, 2006. NFAT transcription factors in control of peripheral T cell tolerance. *European Journal of Immunology* 36, 2837–2843.
- Sharma, S., Findlay, G.M., Bandukwala, H.S., *et al.*, 2011. Dephosphorylation of the nuclear factor of activated T cells (NFAT) transcription factor is regulated by an RNA-protein scaffold complex. *Proceedings of the National Academy of Sciences of the United States of America* 108, 11381–11386.
- Shin, D.S., Jordan, A., Basu, S., *et al.*, 2014. Regulatory T cells suppress CD4+ T cells through NFAT-dependent transcriptional mechanisms. *EMBO Reports* 15, 991–999.
- Soto-Nieves, N., Puga, I., Abe, B.T., *et al.*, 2009. Transcriptional complexes formed by NFAT dimers regulate the induction of T cell tolerance. *Journal of Experimental Medicine* 206, 867–876.
- Terui, Y., Saad, N., Jia, S., McKeon, F., Yuan, J., 2004. Dual role of sumoylation in the nuclear localization and transcriptional activation of NFAT1. *Journal of Biological Chemistry* 279, 26–36.
- Tone, Y., Furuchi, K., Kojima, Y., *et al.*, 2008. Smad3 and NFAT cooperate to induce Foxp3 expression through its enhancer. *Nature Immunology* 9, 194–202.
- Vaeth, M., Schliesser, U., Muller, G., *et al.*, 2012. Dependence on nuclear factor of activated T-cells (NFAT) levels discriminates conventional T cells from Foxp3+ regulatory T cells. *Proceedings of the National Academy of Sciences of the United States of America* 109, 16258–16263.
- Valdor, R., Macian, F., 2013. Induction and stability of the anergic phenotype in T cells. *Seminars in Immunology* 25, 313–320.

- Willingham, A.T., Orth, A.P., Batalov, S., *et al.*, 2005. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* 309, 1570–1573.
- Wing, K., Sakaguchi, S., 2010. Regulatory T cells exert checks and balances on self tolerance and autoimmunity. *Nature Immunology* 11, 7–13.
- Wu, Y., Borde, M., Heissmeyer, V., *et al.*, 2006. FOXP3 controls regulatory T cell function through cooperation with NFAT. *Cell* 126, 375–387.
- Xanthoudakis, S., Viola, J.P., Shaw, K.T., *et al.*, 1996. An enhanced immune response in mice lacking the transcription factor NFAT1. *Science* 272, 892–895.
- Zhang, J., Lee, S.M., Shannon, S., *et al.*, 2009. The type III histone deacetylase Sirt1 is essential for maintenance of T cell tolerance in mice. *Journal of Clinical Investigation* 119, 3048–3058.
- Zheng, Y., Zha, Y., Driessens, G., Locke, F., Gajewski, T.F., 2012. Transcriptional regulator early growth response gene 2 (*Egr2*) is required for T cell anergy *in vitro* and *in vivo*. *Journal of Experimental Medicine* 209, 2157–2163.

Further Reading

- Feske, S., 2007. Calcium signalling in lymphocyte activation and disease. *Nature Reviews Immunology* 7, 690–702.
- Macian, F., 2005. NFAT proteins: Key regulators of T-cell development and function. *Nature Reviews Immunology* 5, 472–484.
- Muller, M.R., Rao, A., 2010. NFAT, immunity and cancer: A transcription factor comes of age. *Nature Reviews Immunology* 10, 645–656.
- Serfling, E., Klein-Hessling, S., Palmethofer, A., *et al.*, 2006. NFAT transcription factors in control of peripheral T cell tolerance. *European Journal of Immunology* 36, 2837–2843.

NF-kappaB and the Immune System

TD Gilmore and TW Siggers, Boston University, Boston, MA, USA

S Gerondakis, Monash University Central Clinical School, Melbourne, VIC, Australia

© 2016 Elsevier Inc. All rights reserved.

Glossary

Acquired immunity An antigen-specific immune response mounted by B and T cells, which is tailored to specific pathogens (e.g., virus, bacteria). Also called adaptive immunity.

Apoptosis Programmed cell death, which can be part of normal developmental or regulatory processes, and can be altered in disease processes (e.g., many leukemias/lymphomas have mutations that enhance cell survival).

B cells Lymphocytes that play a key role in humoral immunity, where their main function is to produce antibodies.

Chemokines A subset of cytokines, which primarily induce immune cell chemotaxis.

Cytokines Small, secreted proteins that coordinate the immune response by affecting immune cell function.

Gene amplification An increase in the copy number of a given gene from the normal diploid state, a situation that occurs frequently for select genes in cancer cells.

Immunoglobulin Another name for an antibody.

Inflammation A complex biological response to pathogens, cell damage, or irritants. Characterized by vascular changes leading to swelling, heat, pain, and soreness.

Innate immunity First line of defense against pathogens, mediated primarily by macrophages and dendritic cells, which recognize conserved molecules on pathogens.

Isotype switching The process by which the type of antibody produced by a B cell changes from one class to another (e.g., switches from IgM to IgG).

Oncogene A viral or cellular gene that can transform a cell from a normal to a cancerous state. These genes are usually mutant versions of a normal cellular gene (the proto-oncogene).

Signal transduction A multicomponent process by which a biochemical signal (often an extracellular ligand) is converted through a series of cytoplasmic protein intermediates into a change in gene expression within the nucleus of a cell, which then alters some process of cellular physiology.

Single nucleotide polymorphism (SNP) A single nucleotide sequence variation in genomic DNA that distinguishes groups of individuals within a population. SNPs are often used as predictors of disease susceptibility or other biological differences between individuals.

T cells Lymphocytes that play the central role in cell-mediated immunity.

Transformation The process by which a normal cell is converted into a malignant cell.

Transcription factor A protein that can bind to specific sites on DNA to increase or decrease expression of a nearby gene.

Introduction

The mammalian NF- κ B (nuclear factor of κ B site binding) transcription factor family comprises the related cellular proteins p50/p105, p52/p100, c-Rel, p65 (aka RelA), and RelB (Table 1). These proteins form homodimers or heterodimers, which can alter the expression of nuclear genes, often ones involved in cellular processes, such as development, cell survival, cell proliferation, and immunity (Gilmore, 2006).

The N-terminal halves of all NF-kappaB (NF- κ B) proteins consist of a conserved DNA-binding/dimerization domain called the Rel homology domain (RHD) (Figure 1(a)). Sequences C-terminal to the RHD can be used to subdivide the NF- κ B family into those that contain a transactivation domain (c-Rel, RelA, RelB), which is required for activation of target gene expression, versus those that contain C-terminal ankyrin repeats (p50/p105, p52/p100), which inhibit the DNA-binding activity of these transcription factors. NF- κ B proteins have also been found in invertebrates, including insects (e.g., *Drosophila melanogaster*), cnidarians (e.g., sea anemones, corals, and hydra), sponges, and some single-celled premetazoans (e.g., choanoflagellates and Capsaspora) (Gilmore and Wolenski, 2012).

In most normal cells, NF- κ B dimers are located in the cytoplasm in an inactive state bound to an inhibitor called I κ B (inhibitor of NF- κ B) (Figure 1(b)). There are several mammalian I κ B proteins (Table 1), which have different affinities for the various NF- κ B dimers and different cell-specific expression patterns. A large number of extracellular signals initiate intracellular signal transduction pathways that lead to I κ B degradation, which enables NF- κ B to enter the nucleus to bind specific DNA sequences and activate or, less often, repress target gene expression. Activation of NF- κ B can be divided into two general pathways, which are distinguished by the upstream activating kinases and by the NF- κ B dimers that are induced. In the classical (or canonical) NF- κ B pathway, inducers activate cytoplasmic I κ B kinase β (IKK β), which phosphorylates an I κ B protein (e.g., I κ B α , β , ϵ), leading to its degradation by the proteasome; this process induces nuclear translocation of specific NF- κ B complexes, generally p50-RelA and p50-c-Rel. In the alternative (or noncanonical) NF- κ B pathway, inducers activate IKK α , which phosphorylates residues on the p52 precursor p100, leading to degradation of the p100 inhibitory ankyrin repeat domain by the proteasome; this process induces nuclear translocation of p52-RelB dimers. In many cases, these two NF- κ B pathways are simultaneously

Table 1 Core NF- κ B signaling gene and protein names

Mouse gene	Mouse protein	Human gene	Human protein	RefSeq accession number
Transcription factors				
<i>nfk1</i>	p50/p105	<i>NFKB1</i>	p50/p105	NP_001158884
<i>nfk2</i>	p52/p100	<i>NFKB2</i>	p52/p100	NP_002493.3
<i>c-rel</i>	c-Rel	<i>REL</i>	REL	NP_002899.1
<i>rela</i>	RelA (or p65)	<i>RELA</i>	RELA (or p65)	NP_001138610.1
<i>relb</i>	RelB	<i>RELB</i>	RELB	NP_006500.2
I κ B proteins				
<i>nfk1a</i>	I κ B α	<i>NFKB1A</i>	I κ B α	NM_020529.2
<i>nfk1b</i>	I κ B β	<i>NFKB1B</i>	I κ B β	NM_002503.4
<i>nfk1e</i>	I κ B ϵ	<i>NFKB1E</i>	I κ B ϵ	NM_004556.2
<i>nfk1d</i>	I κ BNS	<i>NFKB1D</i>	I κ BNS	NP_640332.1
<i>bcl3</i>	Bcl-3	<i>BCL3</i>	BCL3	NM_005178
<i>nfk1z</i>	I κ B ζ	<i>NFKB1Z</i>	I κ B ζ	NM_031419
IKK signaling proteins				
<i>ikb1a</i>	IKK α	<i>IKBKA</i>	IKK α	NM_001278
<i>ikb1b</i>	IKK β	<i>IKKBK</i>	IKK β	NM_001556.2
<i>ikb1g</i>	Nemo	<i>IKBKG</i>	NEMO	NM_001161422.1

induced, but in others, specific activators induce only one of the pathways.

As described in this article, NF- κ B plays a key developmental and functional role in innate and acquired immune cell function as well as inflammation. Importantly, nuclear NF- κ B activation is normally transient in almost all cells, a state achieved by a wide variety of regulatory mechanisms (Hayden and Ghosh, 2011, 2014). Consistent with the need to tightly control NF- κ B activity, in a variety of chronic inflammatory and malignant diseases, NF- κ B is constitutively active and located in the nucleus. Similarly, in certain genetically based human immunodeficiencies, NF- κ B signaling is defective.

NF- κ B in Normal Immune Cell Function and Development

NF- κ B Expression and Knockout Mice

NF- κ B plays key roles in B- and T-cell development, the generation of secondary lymphoid organs (e.g., the spleen and lymph nodes), and the function of many immune cell types (e.g., B cells, T cells, natural killer T cells, dendritic cells (DCs), and macrophages) (Hayden and Ghosh, 2011; Vallabhapurapu and Karin, 2009). NF- κ B transcription factors are expressed in all cell types; however, in many cases their highest levels of expression are in immune cells. In general, p50 and RelA are expressed in all cell types, and the p50–RelA dimer is the NF- κ B complex that is induced in most non-hematopoietic cells. On the other hand, high levels of p52/p100, c-Rel, and RelB expression are restricted to a variety of hematopoietic cells. That is, p52/p100 is expressed at the highest levels in the thymus and spleen, c-Rel levels are high in certain B- and T-cell subtypes, myeloid cells, and erythroid cells, while high levels of RelB are found in lymphoid cells of the thymus, spleen, and small intestine (Peyer's patches).

Much of what is known about the function of NF- κ B in immune cells comes from the study of genetically modified mice (knockouts or KOs) lacking specific NF- κ B subunits

(Gerondakis *et al.*, 2006). With the exception of RelA KO mice, which die embryonically, mice with KOs of other individual NF- κ B subunits are viable with their most prominent phenotypes being alterations in immune cell function (Table 2). The distinct phenotypes of mice with KOs of each NF- κ B subunit demonstrate that each subunit has a unique function. Several mouse models missing multiple NF- κ B subunits have also been created, and generally the phenotypes of mice missing multiple NF- κ B subunits are exaggerated compared to the single KO strains (Gerondakis *et al.*, 2006).

nfk1 KO mice

Mice with a KO of the *nfk1* gene encoding p50/p105 are viable, but have a variety of defects in acquired and innate immune cells (Sha *et al.*, 1995). Various B-cell subtypes (e.g., marginal zone and B1 cells) are reduced in number or have impaired function (e.g., follicular B cells). As a consequence, *nfk1* KO mice mount a poor humoral immune response to certain microbial molecules (e.g., lipopolysaccharide (LPS)) and have abnormal immunoglobulin (Ig) isotype switching. Most types of T cells develop normally in *nfk1* KO mice, although T helper-cell differentiation is partially defective and CD8 memory T-cell development is enhanced in these mice. Innate immune cells are also differentially affected in *nfk1* KO mice: for example, macrophages show reduced expression of certain cytokines (e.g., IL-6, IL-10), whereas natural killer cells have enhanced function. Because of these mixed immune cell phenotypes, *nfk1* KO mice show a range of responses to different pathogens and stresses as compared to wild-type mice: in some cases, *nfk1* KO mice are more susceptible to a given pathogen, and in other cases, they have unchanged or enhanced responses.

nfk2 KO mice

nfk2 KO mice lack the mature p52 protein, as well as the p100 precursor protein. These mice develop normally and are viable and fertile as adults (Caamaño *et al.*, 1998; Franzoso *et al.*, 1998). Nevertheless, they have defective spleen and

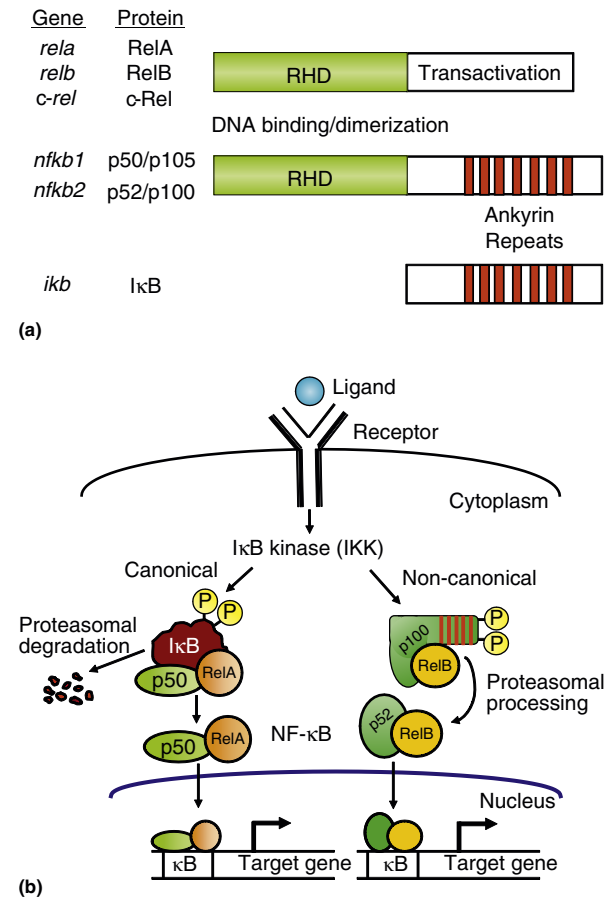


Figure 1 (a) Structures of NF- κ B superfamily proteins. Shown are the general structures of the two subclasses of NF- κ B proteins. All NF- κ B proteins contain an N-terminal DNA-binding/dimerization domain (green) called the Rel homology domain (RHD). At the C-terminus of the RHD there is also a sequence essential for nuclear localization (N). The C-terminal halves of the NF- κ B precursors p105 and p100 contain an ankyrin repeat (ANK) domain (red bars), which must be removed by proteolysis to activate the DNA-binding activity of the NF- κ B subunits p50 and p52, respectively. The C-terminal halves of Rel proteins contain transactivation domains, which are not removed upon activation. The stand-alone I κ B proteins consist largely of ANK repeats, which enable them to bind to and inhibit NF- κ B dimers. (b) General outline of the canonical and noncanonical pathways for activation of NF- κ B transcription factors (see text for more details).

lymph node architecture, impaired B-cell maturation and survival, as well as defective T cell, macrophage, dendritic cell, and natural killer cell functions. Given that p52–RelB is the primary dimer of the alternative NF- κ B pathway, it is not surprising that many of these immune cell defects appear to be due to loss of alternative NF- κ B signaling. Thus, many of the defects seen in *nfkb2* KO mice are also seen in KO mice lacking upstream activators of alternative NF- κ B signaling (e.g., BAFF, the lymphotoxin β receptor or NIK (NF- κ B-inducing kinase)) or missing RelB, the dimeric partner of p52.

c-rel KO mice

The c-Rel protein is expressed most highly in B and T cells. Mice with a *c-rel* gene knockout are viable and fertile, but have

multiple mature immune cell defects that make these mice more sensitive to infection (Gilmore and Gerondakis, 2011). Defects in *c-rel* KO mice that contribute to this phenotype include impaired B-cell proliferation, survival, and isotype switching in response to B cell-activating signals such as antigen receptor and CD40 stimulation. These B-cell deficiencies are due in part to c-Rel regulation of specific genes important for cell division (e.g., E2F, cyclin E, c-Myc) and the inhibition of cell death (e.g., A1, Bcl-XL).

Although T cells in *c-rel* knockout mice have proliferation defects due to impaired antigen and co-stimulatory (CD28) receptor signaling, the extent of the T-cell effects are immune signal dependent (Gerondakis et al., 2006). For example, T cells in *c-rel* KO mice do not proliferate in response to *Toxoplasma gondii*, but do proliferate in response to influenza. c-Rel is also important for the development of immune-suppressive CD4 regulatory T (Treg) cells, with the numbers of Treg cells reduced by approximately sixfold in *c-rel* KO mice (Gerondakis et al., 2014). Nevertheless, the Treg cells that do develop in c-Rel KO mice have normal T cell-suppressive functions, which may explain why these mice fail to develop autoimmune disease despite reduced numbers of Treg cells (Gilmore and Gerondakis, 2011). Of note, the loss of c-Rel function can also be beneficial. For example, the immune response defects seen in *c-rel* KO mice render them less susceptible to certain experimental inducers of arthritis (Gerondakis et al., 2006), and pathology arising from IL-17-dependent inflammatory responses controlled by T helper cells is reduced in *c-rel* KO mice due to c-Rel's role in Th17 differentiation (Chen et al., 2011).

rela KO mice

Full animal KO of the *rela* gene is embryonically lethal because of extensive liver apoptosis, which occurs due to the sensitivity of embryonic hepatocytes to the cell-killing effects of circulating tumor necrosis factor (TNF) in the absence of expression of RelA-dependent anti-apoptotic protective genes (Beg et al., 1995; Beg and Baltimore, 1996). However, a role for RelA in other cell types has been demonstrated through the use of RelA/TNF-receptor double KO mice or by transplanting RelA KO hemopoietic stem cells into wild-type mice. Through such studies, RelA has also been found to be important for the survival of macrophages, B cells, and T cells, as well as for the development of secondary lymphoid organs including the spleen and lymph nodes (Gerondakis et al., 2006).

relb KO mice

Mice with the *relb* gene knocked out are plagued by a large number of primary and secondary lymphoid cell and organ defects and multiple immune cell defects (Weih et al., 1995). For example, *relb* KO mice lack Peyer's patches or germinal centers in the spleen, due to the failure of the TNF superfamily of cytokines to activate p52–RelB dimers. *relb* KO mice also lack certain types of dendritic cells and natural killer T cells, while B cells exhibit functional defects. Moreover, many organs show extensive T-cell and innate immune cell infiltration that coincides with inflammation-induced tissue damage. Collectively, these multiple defects in the immune system lead to a reduction in the life span of mice lacking RelB (Weih et al., 1995).

Table 2 Immune defects of NF- κ B subunit knockout mice

Gene knockout	Protein	Cell type		References
		Affected	Defect	
<i>nfkb1</i>	p50/p105	B cell	Reduced subtypes and proliferation	Sha <i>et al.</i> (1995) Artis <i>et al.</i> (2005) Banerjee <i>et al.</i> (2006) Tato <i>et al.</i> (2006)
		T helper	Reduced differentiation	
		Macrophage	Reduced response to TLR	
		NK	Enhanced proliferation, IFN γ secretion	
<i>nfkb2</i>	p52/p100	B cell	Impaired development	Caamaño <i>et al.</i> (1998) Franzoso <i>et al.</i> (1998) Speirs <i>et al.</i> (2004)
		Spleen	Defective development	
		DC	Enhanced function	
<i>c-rel</i>	c-Rel	B cell	Reduced proliferation/survival	Gerondakis <i>et al.</i> (1996) Isomura <i>et al.</i> (2009) Grumont <i>et al.</i> (2001)
		Tregs	Reduced numbers	
		DC	Reduced numbers, impaired function	
<i>rela</i>	RelA	Spleen	Reduced size	Beg <i>et al.</i> (1995) Beg and Baltimore (1996) Doi <i>et al.</i> (1997) Heise <i>et al.</i> (2014)
		DC	Reduced subtypes, impaired function	
		B cell	Impaired isotype switching and proliferation	
		Plasma cells	Impaired development	
<i>relb</i>	RelB	T cells	Reduced proliferation/survival	Weih <i>et al.</i> (1995) Burkly <i>et al.</i> (1995)
		T helper	Impaired differentiation	
		Spleen	Disorganized architecture	Caamaño <i>et al.</i> (1998)
		DC, NK	Reduced cell numbers, impaired function	

Abbreviations: DC, dendritic cells; IFN, interferon; NK, natural killer cells; TLR, toll-like receptor; Treg, T regulatory cell.

Important NF- κ B-Activating Pathways in Acquired Immunity

B- and T-cell receptor pathways

Several induced processes that are important for B- and T-cell function require activation of NF- κ B. In particular, stimulation of the B-cell receptor (BCR) and T-cell receptor (TCR) by antigen promotes immune cell expansion by turning on cell proliferation, cell survival, and lymphocyte maturation pathways (Schulze-Luehrmann and Ghosh, 2006). Although both of these antigen receptor pathways proceed through the canonical NF- κ B pathway, the molecules upstream of the NEMO (NF- κ B essential modulator)–IKK complex are distinct (Hayden *et al.*, 2006).

CD40

CD40 is a TNF receptor family member that is expressed on the surface of B cells, which are activated during antigen-dependent B- and T-cell interactions through CD40 ligand–CD40 engagement. This interaction results in activation of NF- κ B, which promotes B-cell division, survival, isotype switching, and plasma cell differentiation (Hayden and Ghosh, 2014). Unlike most receptors, CD40 engages both the classical and alternate NF- κ B pathways in B cells.

BAFF receptor pathway

The TNF superfamily cytokine B cell-activating factor (BAFF) binds to three distinct receptors (BAFF-R, TACI, and BCMA) mainly expressed to varying degrees on B cells of different maturity. Signaling through BAFF-R and BCMA promotes the survival and proliferation of B cells by activating NF- κ B dimers downstream of the classical and alternate pathways (Rickert *et al.*, 2011).

CD28 receptor pathway

T-cell activation and survival is enhanced by CD28, a receptor for B7.1 and B7.2, ligands expressed on antigen presenting cells. CD28–B7 interactions enhance T-cell function in part by promoting transcription of the gene encoding the cytokine IL-2, which depends on classical NF- κ B dimers comprising c-Rel, RelA, and NF- κ B1 (Cheng *et al.*, 2011).

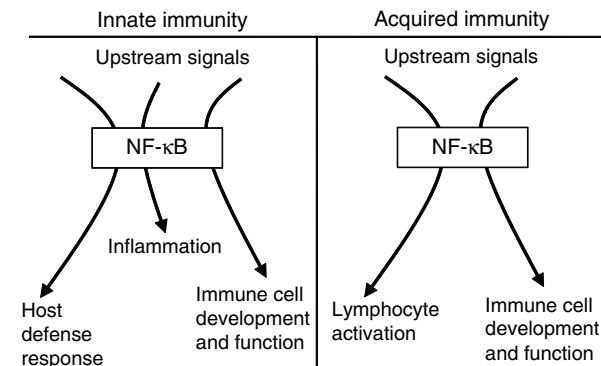
Role of NF- κ B in Innate Immunity

The innate immune system is the first line of defense against microbial pathogens, functioning to identify infectious agents and rapidly eliminate them. This initial response to infection is characterized by numerous activities including cellular communication via secreted cytokines, homing of immune cells (e.g., neutrophils and natural killer (NK) cells) along chemokine gradients to sites of infection, secretion of acute-phase proteins by the liver, and production of pro-inflammatory molecules (e.g., nitric oxide and prostaglandins). NF- κ B plays a central role in many of these responses, acting early in macrophages and dendritic cells that first sense pathogens to drive expression of inflammatory and immune regulatory genes, while also functioning in other cell types (e.g., liver cells and vascular endothelial cells) to mediate various cellular responses to pro-inflammatory cytokines (Dev *et al.*, 2011). The function of NF- κ B in the innate immune response is conserved throughout vertebrate evolution, and indeed, some of the components of innate immunity are found in insects, such as *Drosophila*, suggesting that the control of the host-defense responses is one of the earliest functions of NF- κ B signaling. The receptors of the human innate immune system that activate NF- κ B are listed in Table 3.

Table 3 Innate immune receptors that activate NF-κB

Receptor family	Cellular location	Number in humans	Ligands
TLR	Plasma membrane	9	Lipids, proteins, nucleic acids
NLR	Cytosolic	6	Peptidoglycans, microbial toxins
RLR	Cytosolic	2	Viral RNA

Abbreviations: NLR, NOD-like receptor; RLR, RIG-I-like receptor; TLR, toll-like receptor.

**Figure 2** Summary of major classes of innate and acquired immune processes controlled by activation of NF-κB (see text for more details).

A defining feature of the innate immune response is the recognition of invading pathogens through a host of conserved pathogen-associated molecules. Sentinel cells in our immune system, such as macrophages and dendritic cells, express a range of specialized receptor molecules for detecting pathogen components. The Toll-like receptors (TLRs) are a family of plasma membrane-spanning proteins that bind a range of pathogen-association molecules, such as the bacterial cell-wall component lipopolysaccharide or unmethylated DNA (Kawasaki and Kawai, 2014). In the cytoplasm, NOD-like receptors (NLR) (Saxena and Yeretssian, 2014) bind bacterial molecules such as peptidoglycans and RIG-I-like receptors (Reikine *et al.*, 2014) recognize viral RNA. Upon ligand binding, all of these pathogen-detection receptors activate NF-κB, which in turn increase the expression of target genes encoding effectors of the innate immune response, such as the pro-inflammatory signaling molecules (e.g., TNF and interleukin 1-beta (IL-1β)), proteins involved in microbial killing (e.g., defensins), chemokines that coordinate immune cell migration (e.g., interleukin 8 (IL-8)), and proteins involved in inflammation (e.g., nitric oxide synthase (iNOS) and cyclooxygenase enzyme 2 (COX2)). This central role of NF-κB in the initial response to pathogen detection remains conserved in other organisms; in *Drosophila*, TLR engagement can activate an NF-κB-like pathway that activates a number of antimicrobial genes called cecropins (Ganesan *et al.*, 2011).

In addition to initiating the innate immune response, NF-κB functions in a variety of cells that respond to the pro-inflammatory cytokines by mediating signaling downstream of IL-1 and TNF receptors. Liver cells respond to IL-1 and TNF by producing and secreting acute-phase proteins that have broad immune effects including pathogen opsonization by complement proteins and inhibition of microbial growth. Vascular endothelial cells respond to IL-1 and TNFα by up-regulating

Table 4 Some NF-κB target genes important in immune cell processes

Functional category	Target gene product
Anti-apoptosis	Bcl-2, Bcl-xL, A1/Bfl-1, IAP
Cytokines	TNF-α, IL-1β, IL-6, IL-8, IFNβ
Chemokines	CCL5
Cell cycle/cell proliferation	Cyclin D1, E2Fa
Adhesion	VCAM, ICAM, E-selectin
Enzymes	MMP1, iNOS, COX-2
Transcription factors	Foxp3, c-Myc, c-Rel, IRF4, NFAT

Abbreviations: Bcl, B-cell lymphoma; CCL5, chemokine (C-C motif) ligand 5; COX-2, cyclooxygenase-2; IAP, inhibitor of apoptosis; ICAM, intercellular adhesion molecule; IFN, interferon; IL, interleukin; iNOS, inducible nitric oxidase synthetase; IRF, interferon regulatory factor; MMP1, matrix metalloproteinase-1; NFAT, nuclear factor of activated T cells; TNF, tumor necrosis factor; VCAM, vascular cell adhesion molecule.

molecules, such as E-selectin, ICAM-1, and VCAM-1, needed for leukocyte trafficking to sites of infection. These diverse cellular responses, all mediated by NF-κB in response to pro-inflammatory cytokines, are central to the coordinated response to infection. Thus, NF-κB plays a key role at the center of the innate immune response (Figure 2).

Misregulation of NF-κB in Human Immune Diseases

Given the important role that NF-κB plays in normal immune cell function, it is perhaps not surprising that it is disrupted in many diseases that involve altered immunity or inflammation. In general, these diseases rely on defective or chronic NF-κB signaling, which results in changes in the normal NF-κB-dependent regulation of target gene expression. Some of the key genes that are involved in normal and pathological NF-κB-dependent immune and inflammatory responses are listed in Table 4.

Increased NF-κB Activity in Chronic Inflammatory Diseases

The key role that NF-κB plays in initiating and maintaining normal acute inflammatory responses explains its prominent role in many chronic inflammatory diseases. Most, if not all, chronic inflammatory diseases show increased NF-κB activity, which clearly contributes to the disease. One of the best characterized examples of NF-κB's role in inflammatory disease is Crohn's disease, which can be caused by mutations in a cytoplasmic pathogen receptor gene, *NOD2*. These disease-associated mutations cause *NOD2* to become a constitutive activator of NF-κB-dependent inflammation (Saxena and Yeretssian, 2014).

Single nucleotide polymorphisms (SNPs) at the *REL*, *NFKB1*, and *NFKBIA* (encoding I κ B) locus have been associated with several human autoimmune diseases, including arthritis, psoriasis, celiac disease, and inflammatory bowel disease (Table 5; Gilmore and Gerondakis, 2011; Sun and Zhang, 2007). However, whether these genetic changes play an active role in these diseases is not clear.

Decreased NF- κ B Activity in Human Immunodeficiency Diseases

Mutations that abrogate activation of NF- κ B have been found in several human immunodeficiencies. In particular, mutations in several genes that lead to defective activation of the TLR pathway are associated with reduced responsiveness to pathogens (Frans *et al.*, 2014). In one of the best characterized examples, mutations in the *IKBKG* gene lead to immunodeficiencies. The *IKBKG* gene encodes a scaffold protein (NEMO) that interacts with and is required for activation of the I κ B kinase by a variety of upstream signals to lead to activation of NF- κ B. Mutations that inactivate NEMO function, and thus lessen NF- κ B activation, are primarily associated with the immunodeficiency syndromes of ectodermal dysplasia and incontinentia pigmenti (Courtois and Israël, 2011).

Increased NF- κ B Activity in Immune Cell Cancers

Many B- and T-cell leukemias and lymphomas have constitutively high levels of nuclear NF- κ B DNA-binding activity and consequently increased NF- κ B target gene expression (see Table 5). The increased NF- κ B activity can arise from a variety of mechanisms: (1) point mutations in genes that activate or

inhibit NF- κ B, (2) gene amplification, (3) autocrine or paracrine stimulation, and (4) possibly epigenetic mechanisms.

Activation of NF- κ B by point mutations in signaling pathway genes has been studied extensively in diffuse large B-cell lymphoma (DLBCL) and multiple myeloma (Staudt, 2010). In DLBCL, mutations that lead to chronic activation of NF- κ B are commonly found amongst proteins in the BCR-to-NF- κ B pathway, and include activating mutations in the BCR, adaptor proteins (CARD11, MALT1, BCL10), and inactivating mutations in I κ B (Staudt, 2010). Moreover, one subclass of DLBCL, called the activated B-cell (ABC) subtype, is characterized by activated NF- κ B that results in cells from ABC-DLBCL being particularly sensitive to growth inhibition by a variety of NF- κ B inhibitors (Staudt, 2010). Similarly, multiple myeloma cells have a variety of mutations in cytoplasmic regulators of NF- κ B signaling and anti-NF- κ B compounds inhibit their growth (Gilmore, 2007).

The *REL* gene (encoding human c-Rel) is frequently amplified in several types of human B-cell lymphoma, including DLBCL, Hodgkin's lymphoma and follicular lymphoma, and less frequently in T-cell lymphoma (Gilmore and Gerondakis, 2011). Mutations in *REL*, *RELA*, and *NFKB2* have also been described in B-cell lymphomas, multiple myeloma, and T-cell lymphoma, respectively (Courtois and Gilmore, 2006). Moreover, some human lymphoma cells have inactivating mutations in the gene encoding I κ B, and such mutations cause NF- κ B to be constitutively active. Of note, the *v-rel* oncogene carried by the avian Rev-T retrovirus rapidly transforms a variety of avian hematopoietic and lymphoid cells *in vitro* and *in vivo* (Gilmore, 1999), and overexpression of human c-Rel can directly transform animal lymphoid cells in culture (Gilmore and Gerondakis, 2011).

Human Immune Cell Viruses Use NF- κ B for Replication and Pathogenesis

Several human viruses utilize NF- κ B for their replication or pathogenesis in immune cells (Gilmore and Mosialos, 2003). In one of the best characterized examples, human immunodeficiency virus type-1 (HIV-1), the causative agent of acquired immunodeficiency syndrome (AIDS), has two NF- κ B binding sites in its promoter and these sites are required for efficient expression of viral genes in several cell types (Chan and Greene, 2012). Moreover, activation of NF- κ B in immune cells following infection by other pathogens may, at least in part, be responsible for reactivation of HIV-1 gene expression, leading to the escape from viral latency seen in many HIV-1-infected individuals. In other cases, viruses encode proteins that can intrude into cellular signaling pathways to activate NF- κ B. For example, Epstein-Barr Virus (EBV), which is associated with Burkitt's B-cell lymphoma, encodes the LMP1 protein that mimics the ability of CD40 to activate NF- κ B, which is required for EBV's B-cell transforming effect (Ersing *et al.*, 2013). Similarly, the Tax protein of the human T-lymphotrophic virus type 1 (HTLV-1) interacts with several cellular proteins to activate IKK and cause chronic activation of NF- κ B, a process that is required for the T-cell transforming activity of HTLV-1 (Cheng *et al.*, 2012).

Table 5 Human immune-related diseases associated with alterations in NF- κ B signaling

Disease	NF- κ B activity	Molecular changes
B-cell lymphoma	Increased	<i>REL</i> amplification; <i>NKFB2</i> , <i>BCL3</i> , <i>REL</i> mutation; Upstream signaling mutations
T-cell lymphoma	Increased	<i>REL</i> amplification; <i>NKFB2</i> , <i>BCL3</i> , <i>REL</i> mutation
Multiple myeloma	Increased	<i>RELA</i> , <i>NFKB1</i> , <i>NFKB2</i> mutations; Upstream signaling mutations
Chronic inflammatory (arthritis, psoriasis, celiac disease, inflammatory bowel disease)	Increased	<i>REL</i> , <i>NFKB1</i> , <i>NFKBIA</i> SNPs; <i>NOD2</i> mutations
Immunodeficiencies (incontinentia pigmenti; ectodermal dysplasia)	Decreased	<i>IKBKG</i> and <i>IKBA</i> mutations; Upstream signaling mutations

Targeting NF- κ B for Therapy in Immune Diseases

There are approximately 1000 molecules that have been reported to inhibit NF- κ B signaling and ones that block almost every step in the NF- κ B pathway have been described (Gilmore and Garbati, 2010). Many natural products that have been ascribed anti-inflammatory, antiaging, and anticancer properties have been shown to block NF- κ B activation in cell-based studies (Luqman and Pezzuto, 2010). In most cases, it is not clear whether such substances exert disease protective effects through inhibition of NF- κ B. To cite a few examples, natural products with proposed biological activity and anti-NF- κ B activity include an active green tea component epigallocatechin-3-gallate, curcumin, and many other food ingredients with antioxidant activity.

Between approximately 1990 and 2005, there was considerable interest in developing anti-NF- κ B pharmaceuticals, primarily for the treatment of cancer and inflammation. Many of these studies sought to target IKK, as a central mediator of NF- κ B signaling and because of successes in targeting other protein kinases. Although several such compounds showed success in cell-based or animal studies, they have not been particularly successful in clinical trials. At least in part, this is because inhibition of IKK can promote liver cancer and unexpectedly also induced systemic inflammation (Luedde *et al.*, 2007). However, proteasome inhibitors, which probably act in part by inhibiting NF- κ B activation through blocking I κ B degradation, have shown some success in the treatment of multiple myeloma (Lawasut *et al.*, 2012). NF- κ B subunit-specific inhibitors may hold more promises for future therapy. While such compounds do not currently exist for each NF- κ B subunit, a recent report on a c-Rel-specific inhibitor that blocks graft versus host disease in mouse recipients of bone marrow transplants (Shono *et al.*, 2014) lends considerable weight to this therapeutic strategy.

See also: Nucleic Acid Synthesis/Breakdown: Transcription: Distant Activation of Transcription by Enhancers; Eukaryotic Transcriptional Regulation

References

- Artis, D., Kane, C.M., Fiore, J., *et al.*, 2005. Dendritic cell-intrinsic expression of NF- κ B1 is required to promote optimal Th2 cell differentiation. *Journal of Immunology* 174, 7154–7159.
- Banerjee, A., Gugasyan, R., McMahon, M., Gerondakis, S., 2006. Diverse Toll-like receptors utilize Tpl2 to activate extracellular signal-regulated kinase (ERK) in hematopoietic cells. *Proceedings of the National Academy of Sciences of the United States of America* 103, 3274–3279.
- Beg, A.A., Baltimore, D., 1996. An essential role for NF- κ B in preventing TNF- α -induced cell death. *Science* 274, 782–784.
- Beg, A.A., Sha, W.C., Bronson, R.T., Ghosh, S., Baltimore, D., 1995. Embryonic lethality and liver degeneration in mice lacking the RelA component of NF- κ B. *Nature* 376, 167–170.
- Burkly, L., Hession, C., Ogata, L., *et al.*, 1995. Expression of relB is required for the development of thymic medulla and dendritic cells. *Nature* 373, 531–536.
- Caamaño, J.H., Rizzo, C.A., Durham, S.K., *et al.*, 1998. Nuclear factor (NF)- κ B2 (p100/p52) is required for normal splenic microarchitecture and B cell-mediated immune responses. *Journal of Experimental Medicine* 187, 185–196.
- Chan, J.K., Greene, W.C., 2012. Dynamic roles for NF- κ B in HTLV-1 and HIV-1 retroviral pathogenesis. *Immunological Reviews* 246, 286–310.
- Chen, G., Hardy, K., Pagler, E., *et al.*, 2011. The NF- κ B transcription factor c-Rel is required for Th17 effector cell development in experimental autoimmune encephalomyelitis. *Journal of Immunology* 187, 4483–4491.
- Cheng, H., Ren, T., Sun, S.-C., 2012. New insight into the oncogenic mechanism of the retroviral oncoprotein Tax. *Protein & Cell* 3, 581–589.
- Cheng, J., Montecalvo, A., Kane, L.P., 2011. Regulation of NF- κ B induction by TCR/CD28. *Immunologic Research* 50, 113–117.
- Courtois, G., Gilmore, T.D., 2006. Mutations in the NF- κ B signaling pathway: Implications for human disease. *Oncogene* 25, 6831–6843.
- Courtois, G., Israëli, A., 2011. IKK regulation and human genetics. *Current Topics in Microbiology and Immunology* 349, 73–95.
- Dev, A., Iyer, S., Razani, B., Cheng, D., 2011. NF- κ B and innate immunity. *Current Topics in Microbiology and Immunology* 349, 115–143.
- Doi, T.S., Takahashi, T., Taguchi, O., Azuma, T., Obata, Y., 1997. NF- κ B RelA-deficient lymphocytes: Normal development of T cells and B cells, impaired production of IgA and IgG1 and reduced proliferative responses. *Journal of Experimental Medicine* 185, 953–961.
- Ersing, I., Bernhardt, K., Gewurz, B.E., 2013. NF- κ B and IRF7 pathway activation by Epstein-Barr virus latent membrane protein 1. *Viruses* 5, 1587–1606.
- Frans, G., Meyts, I., Picard, C., *et al.*, 2014. Addressing diagnostic challenges in primary immunodeficiencies: Laboratory evaluation of Toll-like receptor- and NF- κ B-mediated immune responses. *Critical Reviews in Clinical Laboratory Sciences* 51, 112–123.
- Franzoso, G., Carlson, L., Poljak, L., *et al.*, 1998. Mice deficient in nuclear factor (NF)- κ B/p52 present with defects in humoral responses, germinal center reactions, and spleen microarchitecture. *Journal of Experimental Medicine* 187, 147–159.
- Ganesan, S., Aggarwal, K., Paquette, N., Silverman, N., 2011. NF- κ B/Rel proteins and the humoral immune responses of *Drosophila melanogaster*. *Current Topics in Microbiology and Immunology* 349, 25–60.
- Gerondakis, S., Fulford, T.S., Messina, N.L., Grumont, R.J., 2014. NF- κ B control of T cell development. *Nature Immunology* 15, 15–25.
- Gerondakis, S., Grumont, R., Gugasyan, R., *et al.*, 2006. Unravelling the complexities of the NF- κ B signaling pathway using mouse knockout and transgenic models. *Oncogene* 25, 6781–6799.
- Gerondakis, S., Strasser, A., Metcalf, D., *et al.*, 1996. Rel-deficient T cells exhibit defects in production of interleukin 3 and granulocyte-macrophage colony-stimulating factor. *Proceedings of the National Academy of Sciences of the United States of America* 93, 3405–3409.
- Gilmore, T.D., 1999. Multiple mutations contribute to the oncogenicity of the retroviral oncoprotein v-Rel. *Oncogene* 18, 6925–6937.
- Gilmore, T.D., 2006. Introduction to NF- κ B: Players, pathways, perspectives. *Oncogene* 25, 6680–6684.
- Gilmore, T.D., 2007. Multiple myeloma: Lusting for NF- κ B. *Cancer Cell* 12, 95–97.
- Gilmore, T.D., Garbati, M.R., 2010. Inhibition of NF- κ B signaling as a strategy in disease therapy. *Current Topics in Microbiology and Immunology* 349, 245–263.
- Gilmore, T.D., Gerondakis, S., 2011. The role of the c-Rel transcription factor in development and disease. *Genes and Cancer* 7, 685–711.
- Gilmore, T.D., Mosialos, G., 2003. Viruses as intruders in the Rel/NF- κ B signaling pathway. In: Beyaert, R. (Ed.), *Nuclear Factor- κ B: Regulation and Role in Disease*. The Netherlands: Kluwer Academic Publishers, pp. 91–115.
- Gilmore, T.D., Wolenski, F.S., 2012. Evolution of NF- κ B: Where did it come from and why? *Immunological Reviews* 246, 14–35.
- Grumont, R., Hochrein, H., O’Keeffe, M., *et al.*, 2001. c-Rel regulates interleukin 12 p70 expression in CD8(+) dendritic cells by specifically inducing p35 gene transcription. *Journal of Experimental Medicine* 194, 1021–1032.
- Hayden, M.S., Ghosh, S., 2011. NF- κ B in immunobiology. *Cell Research* 21, 223–244.
- Hayden, M.S., Ghosh, S., 2014. Regulation of NF- κ B by TNF family cytokines. *Seminars in Immunology* 26, 253–266.
- Hayden, M.S., West, A.P., Ghosh, S., 2006. NF- κ B and the immune response. *Oncogene* 25, 6758–6780.
- Heise, N., De Silva, N.S., Silva, K., *et al.*, 2014. Germinal center B cell maintenance and differentiation are controlled by distinct NF- κ B transcription factor subunits. *Journal of Experimental Medicine* 211, 2103–2118.
- Isomura, I., Palmer, S., Grumont, R.J., *et al.*, 2009. c-Rel is required for the development of thymic Foxp3⁺ regulatory T cells. *Journal of Experimental Medicine* 206, 3001–3014.
- Kawasaki, T., Kawai, T., 2014. Toll-like receptor signaling pathways. *Frontiers in Immunology* 5, 461.

- Lawasut, P., Chauhan, D., Laubach, J., *et al.*, 2012. New proteasome inhibitors in myeloma. *Current Hematologic Malignancy Reports* 7, 258–266.
- Luedde, T., Beraza, N., Kotsikoris, V., *et al.*, 2007. Deletion of NEMO/I κ B γ in liver parenchymal cells causes steatohepatitis and hepatocellular carcinoma. *Cancer Cell* 11, 119–132.
- Lugman, S., Pezzuto, J.M., 2010. NF κ B: A promising target for natural products in cancer chemoprevention. *Phytotherapy Research* 24, 949–963.
- Reikine, S., Nguyen, J.B., Modis, Y., 2014. Pattern recognition and signaling mechanisms of RIG-I and MDA5. *Frontiers in Immunology* 5, 342.
- Rickert, R.C., Jellusova, J., Miletic, A.V., 2011. Signaling by the tumor necrosis factor receptor superfamily in B-cell biology and disease. *Immunological Reviews* 244, 115–133.
- Saxena, M., Yeretssian, G., 2014. NOD-like receptors: Master regulators of inflammation and cancer. *Frontiers in Immunology* 5, 327.
- Schulze-Luehrmann, J., Ghosh, S., 2006. Antigen-receptor signaling to nuclear factor κ B. *Immunity* 25, 701–715.
- Sha, W.C., Liou, H.-C., Tuomanen, E.I., Baltimore, D., 1995. Targeted disruption of the p50 subunit of NF- κ B leads to multifocal defects in immune responses. *Cell* 80, 321–330.
- Shono, Y., Tuckett, A.Z., Ouk, S., *et al.*, 2014. A small-molecule c-Rel inhibitor reduces alloactivation of T cells without compromising antitumor activity. *Cancer Discovery* 4, 578–591.
- Speirs, K., Lieberman, L., Caamano, J., Hunter, C.A., Scott, P., 2004. Cutting edge: NF- κ B2 is a negative regulator of dendritic cell function. *Journal of Immunology* 172, 752–756.
- Staudt, L.M., 2010. Oncogenic activation of NF- κ B. *Cold Spring Harbor Perspectives in Biology* 2, a000109.
- Sun, X.-F., Zhang, H., 2007. *NFKB* and *NFKB1* polymorphisms in relation to susceptibility of tumor and other diseases. *Histology and Histopathology* 22, 1387–1398.
- Tato, C.M., Mason, N., Artis, D., *et al.*, 2006. Opposing roles of NF- κ B family members in the regulation of NK cell proliferation and production of IFN- γ . *International Immunology* 18, 505–513.
- Vallabhapurapu, S., Karin, M., 2009. Regulation and function of NF- κ B transcription factors in the immune system. *Annual Review of Immunology* 27, 693–733.

Weih, F., Carrasco, D., Durham, S.K., *et al.*, 1995. Multiorgan inflammation and hematopoietic abnormalities in mice with targeted disruption of RelB, a member of the NF- κ B/Rel family. *Cell* 80, 331–340.

Further Reading

- Baker, R.G., Hayden, M.S., Ghosh, S., 2011. NF- κ B, inflammation, and metabolic disease. *Cell Metabolism* 13, 11–22.
- Bhatt, D., Ghosh, S., 2014. Regulation of the NF- κ B-mediated transcription of inflammatory genes. *Frontiers in Immunology* 5, 71.
- Gerondakis, S., Siebenlist, U., 2010. Roles of the NF- κ B pathway in lymphocyte development and function. *Cold Spring Harbor Perspectives in Biology* 2, 1000182.
- Ghosh, S., Hayden, M.S., 2008. New regulators of NF- κ B in inflammation. *Nature Reviews Immunology* 8, 837–848.
- Hayden, M.S., Ghosh, S., 2008. Shared principles in NF- κ B signaling. *Cell* 132, 344–362.
- Hayden, M.S., Ghosh, S., 2012. NF- κ B, the first quarter-century: Remarkable progress and outstanding questions. *Genes & Development* 26, 203–234.
- Oh, H., Ghosh, S., 2013. NF- κ B: Roles and regulation in different CD4(+) T-cell subsets. *Immunological Reviews* 252, 41–51.
- Sun, S.-C., 2012. The noncanonical NF- κ B pathway. *Immunological Reviews* 246, 125–140.

Relevant Websites

- www.nf-kb.org
 Boston University.
[http://en.wikipedia.org/wiki/NF- \$\kappa\$ B](http://en.wikipedia.org/wiki/NF-κB)
 Wikipedia.

CELLULAR IMMUNOLOGY: CELL–CELL INTERACTIONS AND THE IMMUNE SYSTEM

Contents

Functional Specialization of Dendritic Cell Subsets

CD28 Costimulation and Regulatory T Cells

Roles of Stromal Cells in the Immune System

CXCL12/SDF-1 and Hematopoiesis

Lymphocyte–Endothelial Interactions

T Cell Receptor Triggering

Single Molecule Methods to Measure Receptor–Ligand Interaction in Immunological Synapses

Single-Cell Interrogation of the Immune System Using Microtools

Cellular Structures Controlling T Cell Signaling in Time and Space

Functional Specialization of Dendritic Cell Subsets

GF Heidkamp, CHK Lehmann, and L Heger, University Hospital of Erlangen, Erlangen, Germany

A Baransk, University Hospital of Erlangen, Erlangen, Germany and Aix Marseille Université, Marseille-Luminy, France

A Hoffmann, J Lühr, and D Dudziak, University Hospital of Erlangen, Erlangen, Germany

© 2016 Elsevier Inc. All rights reserved.

Discovery of Dendritic Cells

Ralph Steinman was the first to describe the functional properties of a very rare cell type, the dendritic cells (DCs) (Steinman and Cohn, 1973; Steinman and Nussenzweig, 1980). In 2011, Steinman was awarded the Nobel Prize for his discoveries (Banchereau *et al.*, 2011; Steinman, 2012). DCs are localized throughout the body, equipped with a variety of Fc receptors (FcR) and specialized pattern recognition receptors (PRR) for sensing and recognition of invading pathogens. These PRRs include Toll-like receptors (TLR), C-type lectin receptors (CLR), nucleotide-binding oligomerization domain receptors (NOD), and retinoic-inducible gene 1-like receptors (RIG-I) (Kawai and Akira, 2008, 2011; Akira *et al.*, 2006; Figdor *et al.*, 2002; Chopin *et al.*, 2012; Elinav *et al.*, 2011; Yoneyama and Fujita, 2009; Nimmerjahn and Ravetch, 2006, 2008; Guillemins *et al.*, 2014a). Upon antigen encounter, DCs engulf antigens from their surroundings by macropinocytosis, phagocytosis, or receptor-mediated endocytosis (Burgdorf and Kurts, 2008; Roche and Furuta, 2015; Trombetta and Mellman, 2005; Guemnonprez *et al.*, 2002; Savina and Amigorena, 2007). Direct contact to microbial products induces DC maturation processes including the upregulation of MHC class II co-stimulatory molecules (e.g., CD80 (B7-1), CD86 (B7-2)), and activation markers (e.g., CD40 and CD83) (Kamath *et al.*, 2000; Cao *et al.*, 2005) as well as the secretion of cytokines (Vega-Ramos *et al.*, 2014; Kurts *et al.*, 2010). In addition, chemokine receptors, such as CCR7, are upregulated and facilitate DC migration to secondary lymphoid

tissues, for the presentation of processed antigens (Förster *et al.*, 1999; Dieu *et al.*, 1998). As most important antigen presenting cells, DCs are able to bridge the gap between innate and acquired immunity by inducing primary immune responses, thereby initiating lifelong immunological memory to the invading pathogen (Banchereau and Steinman, 1998). On the other hand, DCs are able to maintain peripheral tolerance against self-antigens (Hawiger *et al.*, 2001; Steinman and Nussenzweig, 2002). DC antigen presentation under steady state conditions leads to elimination or anergy (durable non-responsiveness) of the interacting, potentially auto-reactive T cells and thereby to protection from autoimmune responses (Steinman *et al.*, 2003; Steinman and Nussenzweig, 2002; Schwartz, 2003).

The presentation of antigens occurs in the immunological synapse in a complex either with the major histocompatibility molecule type I or II (MHC-I and MHC-II), thus allowing for recognition of the peptide antigen either by CD8⁺ T cells or CD4⁺ T cells, respectively. Importantly, only professional antigen presenting cells (or certain types of nonprofessional antigen presenting cells) are able to present antigens as peptide MHC class II complexes (pMHC-II) on their cell surface (Shortman and Liu, 2002; Heath and Carbone, 2001; Mellman and Steinman, 2001; Villadangos and Schnorrer, 2007; Villadangos, 2001; Joffre *et al.*, 2012; Merad *et al.*, 2013; Konermann *et al.*, 2012). The specific combination of the antigen presenting DC subtype, origin of activation stimulus, and the secretion of DC subset specific cytokines directs T cell priming and determines the fate of a CD4⁺ T helper cell to

become a T_H1 , T_H2 , T_H17 , or regulatory T cell, thereby indirectly influencing the assembly and production of antigen-specific antibodies by B cells. In the following sections, the major classical and alternative antigen presentation pathways will be discussed in more detail.

Antigen Processing and Presentation

Classical MHC Class I Processing

Almost all nucleated cell populations express MHC class I (MHC-I) molecules on their surfaces. Cells that have down-regulated MHC-I or express high levels of NK cell activating receptors are eliminated by NK cells (Lanier, 2005; Diefenbach *et al.*, 2001). This can be observed during carcinogenic transformation of normal cells into tumor cells or as a consequence of viral infection. The MHC-I molecule is of pivotal importance for the presentation of antigens to cytotoxic $CD8^+$ T cells. In order to present antigens on the cell surface as peptide MHC class I complexes (pMHC-I), cytoplasmic antigens have to be proteolytically cleaved within the proteasome, transported into the lumen of the endoplasmic reticulum (ER), and subsequently loaded onto freshly synthesized MHC-I molecules (Bjorkman *et al.*, 1987). These pMHC-I complexes are transported via the Golgi network to the cell surface (Joyce, 1997).

Under steady state conditions, a multicatalytic protease, the so-called proteasome, generates MHC-I peptides from different substrates, such as defective ribosomal products (DRiPs), viral components, as well as error-free short- and long-lived proteins (e.g., transcription factors). The quantitative contribution of DRiPs to the pool of degradable proteins is still under debate (Rock *et al.*, 2014; Yewdell *et al.*, 1996; Schubert *et al.*, 2000). The 26S proteasome contains a 20S core and two 19S outer compartments (Arrigo *et al.*, 1988; Voges *et al.*, 1999). The core consists of 28 subunits, which are structured into four heptameric rings containing the outer α -subunits and inner proteolytically active β -subunits, namely $\beta1$, $\beta2$, and $\beta5$ (Kloetzel and Ossendorp, 2004). The polyubiquitination of misfolded proteins flags them for recognition by the 19S compartment (Loureiro and Ploegh, 2006; Kloetzel and Ossendorp, 2004; Lecker *et al.*, 2006; Metzger *et al.*, 2014). Unfolded proteins reach the inner 20S compartment for initial degradation (Cascio *et al.*, 2001). Most of the resulting peptide fragments are reduced to peptides <8 residues in length, in particular by the cytosolic metalloendopeptidase thimet oligopeptidase (TOP) (Saric *et al.*, 2001; York *et al.*, 2003). As a result, the majority of peptides are too small for MHC-I loading, therefore lost for antigen presentation and might be rather reused for protein biosynthesis (Rock *et al.*, 2004). However, a small fraction of peptides escapes destruction or just experiences cytosolic peptide-trimming. The membrane-bound heterodimeric TAP-1/TAP-2 complex (adenosine-triphosphate (ATP) dependent transporter associated with antigen processing (TAP)) transports these peptides from the cytosol into the lumen of the endoplasmic reticulum (ER). Here TAP associates with Tapasin (TAP-binding protein) and the MHC-I complex, consisting of an α -chain and a β_2 -microglobulin molecule (Wearsch and Cresswell, 2008). Only

a minor percentage of the translocated antigens has been fully processed by the proteasome (Lucchiari-Hartz *et al.*, 2000). Evidence arises that most epitope precursor peptides, which have entered the ER, require further N-terminal trimming to the appropriate amino acid (aa) length of 8–9 residues by the heterodimeric complex of ER-resident aminopeptidases 1 and 2 (ERAP1; ERAP2) (Chang *et al.*, 2005; Evnouchidou *et al.*, 2014; Weimershaus *et al.*, 2013; Schatz *et al.*, 2008; Kanaseki *et al.*, 2006; Rock and Goldberg, 1999; Kisselev *et al.*, 1999). Noteworthy, both peptide length and sequence are crucial for a successful $CD8^+$ T cell response (Ekeruche-Makinde *et al.*, 2013; Burrows *et al.*, 2006; Cole *et al.*, 2010; Schumacher *et al.*, 1991). Before loading of the processed peptides onto the MHC-I molecule, the latter has to be assembled. Therefore, the two chaperones Calnexin and Calreticulin stabilize the MHC-I α -chain after its co-translational insertion into the ER membrane. Further, Erp57 joins this complex and ensures proper folding of the MHC-I α -chain and association of the β_2 microglobulin (Zhang *et al.*, 2006b). All these molecules belong to the peptide-loading complex (PLC) (Wearsch and Cresswell, 2008). Loading of the peptide onto an MHC-I molecule results in a stable conformation and the dissociation from the PLC. The pMHC-I complex enters the Golgi network to be transported to the cell surface, where it is presented to $CD8^+$ T cells (Jensen, 2007; Janeway, 1992). The presentation of pMHC-I complexes is strictly dependent on the TAP transporter. Only in some rare cases a pMHC-I presentation might be possible in DCs without TAP (Joffe *et al.*, 2012).

Classical MHC Class II Processing

In contrast to the formation of pMHC-I complexes, only professional antigen presenting cells, especially DCs, are able to process and present peptides of antigenic proteins as peptide MHC class II complexes (pMHC-II) on their cell surface for the induction of $CD4^+$ T cell responses (Cella *et al.*, 1997; Pooley *et al.*, 2001; Shortman and Liu, 2002; Mellman and Steinman, 2001; Villadangos and Schnorrer, 2007; Merad *et al.*, 2013). Professional antigen presenting cells are specialized in the uptake of extracellular antigens and the formation of so-called intracellular phagosomes or early endosomes (Burgdorf and Kurts, 2008; Roche and Furuta, 2015; Trombetta and Mellman, 2005; Guernonprez *et al.*, 2002; Savina and Amigorena, 2007). Based on our current knowledge, at least four different internalization routes of exogenous antigens can be described: clathrin-mediated or caveolae-dependent endocytosis, phagocytosis, and macropinocytosis (Trombetta and Mellman, 2005; Roche and Furuta, 2015).

Cell surface receptors involved in receptor-mediated endocytosis recycle back to the plasma membrane (recycling endosomes), while the engulfed antigens are further processed (Roche and Furuta, 2015). Therefore, the antigen-containing early endosomes (as well as macropinosomes and phagosomes) fuse with a late endosomal/lysosomal antigen processing compartment. Characteristically, this organelle is a multivesicular body (MVB), with many small intraluminal vesicles generated by the endosomal sorting complexes required for transport (ESCRT) that contain ubiquitinated membrane proteins. Additionally, the low pH and the

presence of proteolytic enzymes foster the assembly of pMHC-II complexes (Roche and Furuta, 2015; Trombetta and Mellman, 2005; Luzio *et al.*, 2010).

Similar to the proteolytic cleavage of proteins by the proteasome for the subsequent loading onto MHC class I molecules in the ER, exogenous antigens also require particular pre-processing steps. First, the gamma interferon-inducible lysosomal thiolreductase (GILT), which shows a maximum activity in an acidic environment of pH 4–5, reduces protein disulfide bonds to free thiols and therefore enhances further antigen processing (Phan *et al.*, 2000; Maric *et al.*, 2001; Arunachalam *et al.*, 2000). Depending on the antigen presenting cell subpopulation (thymic epithelial cells, DCs, monocytes, macrophages, B cells, among others), a set of endosomal/lysosomal proteases has been described to be involved in antigen processing and function for MHC class II presentation (Chapman, 2006) such as cathepsin B, C, D, E, F, H, K, L, S, V, W, X, as well as the asparaginyl endopeptidase (AEP, also called legumain) (Chapman, 2006; Riese and Chapman, 2000; Manoury *et al.*, 1998; Conus and Simon, 2010; Honey and Rudensky, 2003). This group of proteases, comprising endopeptases, carboxy-, and aminopeptidases, fulfills various functions in the context of antigen processing in professional antigen presenting cells. Cathepsins are of crucial importance, since their dysregulation generates major perturbations of the immune system homeostasis (Nakagawa *et al.*, 1999; Conus *et al.*, 2008; Mohamed and Sloane, 2006). Therefore, cathepsins have become interesting targets for clinical treatment approaches (Yasuda *et al.*, 2005).

MHC-II molecules, consisting of an α -chain and a β -chain, are assembled within the ER, where they form nonamers of three MHC-II α/β heterodimers and three invariant chains (I_i , CD74) (Newcomb *et al.*, 1996). The latter is not only needed to stabilize the newly synthesized MHC-II protein chains, but also to protect the nascent MHC-II molecule from premature association with peptides in the MHC-II binding cleft before reaching the endosome. Further, the invariant chain routes the MHC-II- I_i complex into endosomal organelles (Bakke and Dobberstein, 1990). These compartments are also referred to as the MHC-II compartment (MIIC). In these compartments, I_i is proteolytically cleaved by AEP and cathepsin S in DCs, to yield the CLIP fragment (class II-associated invariant chain peptide). AEP has been shown to be involved in the initiation phase of I_i cleavage, while cathepsin S is responsible for further degradation of I_i and CLIP fragment formation (Nakagawa *et al.*, 1999; Manoury *et al.*, 1998; Watts *et al.*, 2005; Riese *et al.*, 1996; Riese and Chapman, 2000). As this last step should be controlled to protect from unwanted processing and potential loading of self-antigens onto MHC-II, the endogenous inhibitor Cystatin C (CyC) regulates the function of cathepsin S (Pierre and Mellman, 1998; El-Sukari *et al.*, 2003). Interestingly, it could be shown that murine splenic CD11c⁺ DCs contain lower levels of lysosomal proteases (e.g., cathepsin L, S, D, B, and AEP) in comparison to peritoneal lavage macrophages, which was correlating with a prolonged presentation of immunogenic peptides (Delamarre *et al.*, 2005).

Finally, the CLIP fragment is removed by the chaperone HLA-DM, and thereby peptides within the late endosome/lysosome gain access to the binding cleft of MHC-II (Riese and Chapman, 2000; Conus and Simon, 2010; Colbert *et al.*, 2009). In DCs, B cells, and thymic epithelial cells, the non-

polymorphic HLA-DR (mouse: MHC-II) like molecule HLA-DO influences the function of HLA-DM, thereby promoting self-tolerance (Fallas *et al.*, 2004; Poluektov *et al.*, 2013; Denzin, 2013). Upon DC stimulation pMHC-II complexes, now in a stable conformation, follow a microtubular transport from the antigen processing compartment to the cell surface for the presentation to CD4⁺ T cells (Chow *et al.*, 2002). Interestingly, this transport seems to be adjustable as a proportion of pMHC-II complexes can be specifically ubiquitinated by the E3 ligase MARCH 1. Thereby, pMHC-II complexes are transferred to the lysosomal compartments for their degradation, while only a minor fraction of the pMHC-II complexes reach the surface (Mintern *et al.*, 2015; Moffat *et al.*, 2013a).

Cross-Presentation

Cross-presentation describes a process uniquely found in subsets of DCs, in which extracellular antigens, such as viral or tumor antigens, which were engulfed by pinocytosis, phagocytosis or receptor-mediated endocytosis, are presented on MHC-I molecules to CD8⁺ T cells (Dudziak *et al.*, 2007; den Haan *et al.*, 2000; Pooley *et al.*, 2001; Brossart and Bevan, 1997; Villadangos *et al.*, 2007; Rock and Shen, 2005; Joffre *et al.*, 2012; Segura and Amigorena, 2014; Kamphorst *et al.*, 2010; Mintern *et al.*, 2015). This process was first described by Bevan (Bevan, 1976). Due to the shorter half-life of pMHC-I complexes (in comparison to pMHC-II) on the cell surface, the necessary DC activation and migration to draining lymph nodes, phagocytosed antigens require a delayed processing (Gil-Torregrosa *et al.*, 2004). van Montfoort *et al.* (2009) described an MHC-II[−]EEA1[−]LAMP1⁺ lysosome-like compartment in murine bone marrow derived DCs (BMDCs), different from the MIIC, in which exogenous antigen can be stored for several days, therefore creating a depot for sustained peptide delivery for cross-presentation (Martín-Fontecha *et al.*, 2003; van Montfoort *et al.*, 2009; Kreer *et al.*, 2012). In addition, attenuated antigen processing is facilitated in (not further defined) splenic and BMDCs due to lower concentrations of lysosomal proteases and inhibition of acidification (Delamarre *et al.*, 2005; Hotta *et al.*, 2006; Kreer *et al.*, 2012).

Overall, the detailed mechanisms of cross-presentation remain poorly understood, especially how engulfed antigens ‘escape’ the endosomal/lysosomal compartment for proteasomal degradation (Kreer *et al.*, 2012; Rodriguez *et al.*, 1999). Two major pathways of cross-presentation are under current debate: the ‘cytosolic’ and the ‘vacuolar pathway’ (Joffre *et al.*, 2012). For the first theory of cross-presentation, the ‘vacuolar pathway’ via the endocytic compartment, it was shown that this route can be abrogated by lysosomal inhibitors, but is resistant to proteasomal blockade and TAP-independent (Rock and Shen, 2005; Shen and Rock, 2004; Song and Harding, 1996). Therefore, it is believed that antigens following the ‘vacuolar pathway’ of cross-presentation are enzymatically dismantled within the endocytic compartment and loaded onto MHC-I molecules, which have been recycled from the cell surface.

The second mechanism, the ‘cytosolic pathway,’ is characterized by an egress of internalized antigens from the endosomal compartment into the cytosol. This process of antigen export is potentially conducted by the ATPase p97 and the

heterotrimeric SEC61 complex, which have been previously reported to be crucial for the retrotranslocation of misfolded proteins from the ER (ER associated protein degradation, ERAD) into the cytosol (Ackerman *et al.*, 2006; Smith *et al.*, 2011). After proteasomal degradation peptides are subsequently loaded in a TAP-dependent manner onto MHC-I molecules in the ER (Wearsch and Cresswell, 2008). In addition, several research groups have postulated that antigenic peptides are TAP-dependently transported back into the phagosomal/endosomal compartment (Guermonprez *et al.*, 2003; Houde *et al.*, 2003; Burgdorf and Kurts, 2008). It is further believed that regulatory protein SEC22b affects antigen cross-presentation; SEC22b is involved in membrane fusion processes and localized in the ER-Golgi intermediate compartment (ERGIC) as an ER resident soluble *N*-ethylmaleimide-sensitive-factor attachment receptor (SNARE). Interaction of SEC22b with the phagosomal syntaxin 4 facilitates the transfer of ER components, for example, the TAP complex (Cebrian *et al.*, 2011; Joffre *et al.*, 2012). Consequently, the cytosolic pathway of cross-presentation is sensitive to proteasome inhibitors (Kovacovics-Bankowski and Rock, 1995).

In the following further findings correlated with cross-presentation of antigens will be discussed. Data indicate that I α can colocalize with a nascent MHC-I molecule in the ER, instead of MHC-II, followed by a TAP-dependent transfer into the endolysosomal compartment and therefore facilitating cross-presentation of phagocytosed antigens (Basha *et al.*, 2012; Duan and Srivastava, 2012). Moreover, in recent years the existence of the endosomal insulin-regulated aminopeptidase (IRAP) has been elucidated. In analogy to the ER-resident ERAP molecules, endosomal IRAP is specialized in aminoterminal trimming of peptides destined for cross-presentation in DCs (Saveanu *et al.*, 2009; Weimershaus *et al.*, 2012; Segura and Villadangos, 2011).

Unexpectedly, Singh and Cresswell found that the gamma interferon-inducible lysosomal thiolreductase (GILT, also Ifi30), which is described to mediate pMHC-II presentation due to its ability to reduce protein disulfide bounds of engulfed antigens, is critically involved in a TAP-dependent cross-presentation of viral antigens. This activity seemed to be dependent on the protein sequence of the engulfed antigen (Singh and Cresswell, 2010). Whether this is a DC subset specific feature still needs to be determined.

Another way of regulating cross-presentation of antigens has been shown to be influenced by the activity of an alternative proteasome in interferon- γ or interferon- α treated human monocyte derived DCs (moDCs) or BMDCs, but also primary CD11c⁺ DCs, in which the 19S subunit is replaced by the 11S cap, and the active β -subunits β 1, β 2, and β 5 of the 20S subunit are replaced by LMP2, MECL1 (LMP10), and LMP7, respectively (Hisamatsu and Shimbara, 1996; Fröh *et al.*, 1994; Tosello *et al.*, 2009; Bandoh *et al.*, 2005). This catalytical complex is called immunoproteasome and generates peptides with high affinity for MHC class I (Gaczynska *et al.*, 1994).

Autophagy

Autophagy is a crucial player in the maintenance of cellular homeostasis. The term 'autophagy' describes the engulfment

and subsequent endosomal/lysosomal proteolysis of long-lived proteins and the recycling of nutrients or defective organelles (Ravikumar *et al.*, 2010; Ciechanover, 2005; Filomeni *et al.*, 2014; Kroemer *et al.*, 2010). The observation that the cytoplasmic protein neomycin phosphotransferase II (NeoR) was no longer presented as peptide MHC-II complex on EBV transformed human B cells after blockade of the autophagosomal pathway indicated that autophagy is involved in antigen presentation of cytoplasmic proteins on MHC-II (Nimmerjahn *et al.*, 2003). In addition, evidence emerged that also cytosolic pathogens can follow the destruction in these acidic compartments, accompanied by a presentation on MHC-II complex by professional antigen presenting cells (Crotzer and Blum, 2010; Roche and Furuta, 2015). Up to date, there are three major mechanisms defined for lysosomal degradation of intracellular material: microautophagy, macroautophagy, and chaperone-mediated autophagy (CMA) (Mizushima *et al.*, 2008; Münz, 2009).

In the latter pathway the cytosolic chaperone heat shock cognate protein of 70 kD (HSC70) recognizes target peptide sequences. The unfolded target protein binds to the lysosomal membrane mediated by the help of a chaperone complex, including HSC70, HSP90, HSP40, Bag-1, Hip, and Hop (Münz, 2009). Multimeric lysosome-associated membrane protein-2A (LAMP-2A) and lysosomal HSC70 finally facilitate the protein translocation into the lumen of the lysosome for subsequent degradation (Cuervo and Dice, 1996; Agarraberes *et al.*, 1997).

The second mechanism, microautophagy, describes a process of direct inclusion of cytosolic material into the lumen of the late endosomal compartment followed by enzymatic hydrolysis. It further relies on the activity of ESCRT I and III, which conduct the invagination of the endosomal membrane for the formation of MVBs. The transport of the cytosolic cargo is mediated by the chaperone HSC70 (Sahu *et al.*, 2011).

The third mechanism of autophagy, the so-called macroautophagy, describes a process of vesicle formation around cytoplasmic constituents. A set of autophagy-related genes (Atg) such as Atg8/LC3 are involved in the assembly of the autophagosomal double-walled membrane and thereby in an increased presentation of peptides as pMHC-II complex (Schmid *et al.*, 2007). Researchers have identified Beclin-1 as a fundamental key player in the regulation of autophagy and apoptosis (Kang *et al.*, 2011). In addition, Atg5, Atg12, and Atg16L1 form another complex under the influence of the ubiquitin-like enzymes Atg7 (E1-like) and Atg10 (E2-like), which is associated with the outer membrane of the nascent autophagosome until completion (Mizushima *et al.*, 1998). The newly built autophagosome then fuses with endosomal and lysosomal compartments for the generation of amphisomes and autolysosomes, respectively (Fader and Colombo, 2009; Münz, 2009).

In professional antigen presenting cells, most of all in DCs, autophagy is a specific means to deliver cytoplasmic viral particles to intracellular compartments with subsequent MHC-II presentation, TLR stimulation, and proinflammatory cytokine production (Ireland and Unanue, 2011; Severa *et al.*, 2013; Li *et al.*, 2005; Morris *et al.*, 2011). Most knowledge on autophagy in DCs is based on data generated from BMDCs or MoDCs. Interestingly, Lee *et al.* could demonstrate in splenic

murine CD11c⁺ DCs that cross-presentation of peptides remained unaffected in the absence of Atg5, while MHC-II presentation and thereby CD4⁺ T cell induction was impossible due to an impaired phagosome-to-lysosome fusion (Lee *et al.*, 2010). This is a very interesting finding, which needs further clarification of DC subset specific differences.

Besides the classical MHC class I and II processing and presentation pathways, CD1 restricted, MHC-independent presentation mechanisms of lipid antigens have been described, but are excluded here (Barral and Brenner, 2007; Silk *et al.*, 2008).

DC Development and DC Characterization

The rapid progress in flow cytometry, mass cytometry, histocytometry, the conjugation of antibodies to varieties of fluorochromes and tandem conjugates, but also the development of multiple engineered fluorescent proteins as well as transcriptional and proteome profiling allowed for the discovery of a variety of cell subpopulations with specialized immune functions in the field of DC research (O'Donnell *et al.*, 2013; Perfetto *et al.*, 2004; Chattopadhyay and Roederer, 2012; Bendall *et al.*, 2012; Newell *et al.*, 2013; Eissing *et al.*, 2014; Gerner *et al.*, 2012; Edwards *et al.*, 2003; Lindquist *et al.*, 2004; Chattopadhyay *et al.*, 2008; Piatkevich and Verkhusha, 2011; Pandey *et al.*, 2013; Lundberg *et al.*, 2013; Xue *et al.*, 2014; Meissner and Mann 2014; Heim *et al.*, 1995; Tsien 1998; Vander Lugt *et al.*, 2014; Miller *et al.*, 2012; Watchmaker *et al.*, 2014). Today it is known that all (murine) DC subpopulations originate from hematopoietic progenitor cells generated in the bone marrow, or yolk sac/fetal liver at earlier developmental stages (Geissmann *et al.*, 2010; Doulatov *et al.*, 2010; Shortman and Naik, 2007; Liu and Nussenzweig, 2010). In this process the hematopoietic stem cell (HSC) differentiates into a common lymphoid (CLP) and a common myeloid progenitor (CMP) (Liu and Nussenzweig, 2010; Shortman and Naik, 2007). CMPs give rise to a monocyte macrophage DC progenitor (MDP), which generates different monocyte subpopulations and a common DC progenitor (CDP) (Liu and Nussenzweig, 2010; Karsunky *et al.*, 2003; Fogg *et al.*, 2006; Waskow *et al.*, 2008). At this point the CDP can further differentiate into terminally differentiated plasmacytoid DCs (pDCs) as well as precursors of classical/conventional DCs (preDCs). The latter ones leave the bone marrow into the blood and populate lymphoid tissues (lymph nodes, spleen, or Peyer's patches, but also bone marrow, and thymus) as well as non-lymphoid tissues (skin, mucosa, intestine, and lamina propria) (Geissmann *et al.*, 2010; Liu and Nussenzweig, 2010; Merad and Manz, 2009).

Besides their anatomical localization, lifespan, and origin, DCs can be phenotypically distinguished from other cell subpopulations by the absence of markers of the B (mouse: CD19 or human: CD19/CD20) and T cell lineage (CD3) as well as NK cell (NKp46, in humans also: CD56), erythrocyte (mouse: Ter119) and granulocyte/neutrophil lineage (mouse: Ly6G, human: CD66b), in combination with an expression of MHC-II in mice or of HLA-DR in humans, which are all of utmost importance for the identification and definition of DCs (Liu and Nussenzweig, 2010; Waskow *et al.*, 2008; Shortman

and Naik, 2007; Merad and Manz, 2009). Additional cell surface markers such as CD4, CD8, CD11b, CD11c, CD13, CD103, EpCAM, Langerin (human: CD1c (BDCA1 (Blood DC antigen)), CD123, CD141 (BDCA3), CD207 (Langerin) CD303 (BDCA2), CD370 (Clec9A)) are used for the definition of primary DC subpopulations (Liu and Nussenzweig, 2010; Merad and Manz, 2009; Figdor *et al.*, 2002; Satpathy *et al.*, 2012). Recent progress in DC ontogeny research supports the notion that DCs might be better distinguished from the monocyte and macrophage lineages based on the expression of transcripts for the tyrosine kinase receptor fms-like tyrosine kinase Flt3 (also Flk2 or CD135) and the zinc finger transcription factor (zbtb46) (Schraml and Reis e Sousa, 2015; Guilleams *et al.*, 2014b; Merad *et al.*, 2013).

Functional Specialization in Antigen Presentation of Mouse DC Subpopulations

DCs consist of a great family of different DC subpopulations expressing differential sets of cell surface receptors such as TLR, CLR, scavenger receptors, and FcR, all of them important for antigen recognition and/or antigen engulfment. On top, data indicate a DC subset specialization in antigen processing and presentation and thereby activation of differential cellular and humoral immune responses. Most information we have obtained on DC subset specialization is based on the original definition of DCs by cell surface marker expression. According to Merad and Manz, DCs can be separated into two groups. The first group consists of conventional or classical DCs (cDCs), which can be further distinguished into non-lymphoid tissue-migratory and lymphoid tissue-resident DCs. Overall, a clear definition of a cDC subset should consider its functional specialization and developmental origin. The second group is built by plasmacytoid DCs (Merad and Manz, 2009).

Plasmacytoid DCs

Plasmacytoid DCs (pDCs), also known as natural interferon-producing cells, separate early in development from the classical DC branch and are pivotal in the production of type I interferons after viral infections with distinct differences in their migratory capacities (Villadangos and Young, 2008). The pDC development is dependent on common regulators of DC development, such as Stat3, PU.1, E2-2, or Irf8 (Miller *et al.*, 2012; Satpathy *et al.*, 2011; Reizis *et al.*, 2011; Sawai *et al.*, 2013), while their lineage determination is dependent on the E2-2 (Tcf4) transcription factor (Reizis *et al.*, 2011). Recently, it could be shown that the migration of pDCs from the bone marrow into the peripheral organs strictly relies on the transcription factor Runx2 (Sawai *et al.*, 2013).

Besides their plasma cell morphology, characteristic murine pDC cell surface molecules are PDCA-1, SiglecH, and CD45RA/B220, with an intermediate or absent expression of CD11c and a complete absence of most conventional DC lineage markers (Zhang *et al.*, 2006a; Merad and Manz, 2009). As pDCs are the main producers of type I interferons in response to viral infection, TLR7 and 9 ligands or specific

microbial stimuli, they have an important role in the activation of cells of the innate and the adaptive immune system such as T cells (Asselin-Paturel *et al.*, 2005; Swiecki and Colonna, 2010; Cervantes-Barragan *et al.*, 2012; Reizis *et al.*, 2011). pDCs can be found in bone marrow, blood, and lymphoid organs, and they can accumulate in diseased tissues. Studies have implicated that dysregulated pDCs might be involved in the pathogenesis of Sjögren's disease and lupus erythematosus (Rönnblom, 2010).

In contrast to cDCs, murine pDCs are regarded to be incapable of presenting pMHC-II complexes derived from exogenous antigens for a longer period of time to specific T cells, which is in contrast to data derived from human pDCs (Segura and Amigorena, 2014). This is due to the fact that MHC-II turnover and degradation are not downregulated in activated murine pDCs (Young *et al.*, 2008). However, this issue is under current debate, since evidence is emerging that murine pDCs can stimulate profound humoral and cellular immune responses in the presence of certain TLR stimulation (Loschko *et al.*, 2011). Interestingly, pDCs might be further separated into mature CCR9 negative and immature CCR9 positive pDCs (Hadeiba *et al.*, 2008; Wendland *et al.*, 2007). The latter ones have been described to be potent inducers of antigen specific regulatory T cells leading to suppression of antigen-specific immune responses under steady state conditions (Hadeiba *et al.*, 2008). Interestingly, in dependency on pMHC-II presentation the same subset was described to induce mixed Th1/Th2 CD4⁺ responses after *in vivo* delivery of antigens to the pDC surface receptor bone marrow stromal antigen-2 (Bst-2/CD317/PDCA-1) under immunostimulatory conditions. Furthermore, in a model for the depletion of conventional DCs, without affecting pDCs, the delivery of antigens to Bst-2 on pDCs indicated that these cells might only be capable of inducing CD4⁺ T cell responses (Sapozhnikov *et al.*, 2007). This is in contrast to data published by Loschko *et al.*, who clearly demonstrated an antigen cross-presentation and CD8⁺ T cell response under stimulatory conditions (Loschko *et al.*, 2011b; Moffat *et al.*, 2013b). pDC specific depletion models might help to shed more light on the antigen processing and presentation capabilities of pDCs.

Most data on murine pDCs indicate an overall tolerogenic role in a variety of different tissues and autoimmune diseases (Bailey-Bucktrout *et al.*, 2008; de Heer *et al.*, 2004; Hadeiba *et al.*, 2008; Jongbloed *et al.*, 2009; Kang *et al.*, 2007; Ochando *et al.*, 2006; Goubier *et al.*, 2008). In line with these correlations, antigen delivery to the inhibitory sialic acid binding Ig-like lectin H (Siglec-H) molecule on pDCs induced rather tolerogenic antigen specific immune responses even in the presence of a strong immune stimulation, but without involvement of regulatory T cells (Loschko *et al.*, 2011b). Overall, the mechanisms how pDCs might process and present engulfed antigens still remain to be resolved.

Lymphoid Tissue-Resident DCs

Two major resident non-migrating conventional DC subsets, both expressing MHC-II and CD11c, can be found in the murine spleen and lymph nodes: CD8⁺CD11b[−] and CD8[−]CD11b⁺ DCs (Banchereau *et al.*, 2000; McLellan *et al.*,

2002; Merad and Manz, 2009; Vremec *et al.*, 1992; Pulendran *et al.*, 1999). They are reported to be also present with variations in their surface molecules in bone marrow, mucosal tissues, thymus, and liver (Iwasaki, 2007; Merad and Manz, 2009; Pulendran *et al.*, 2008). In the following both subset specifications in the process of antigen processing and presentation will be described in more detail.

CD8⁺CD11b[−] DCs

The CD8⁺CD11b[−] resident DC subset expresses DEC205 (CD205, Ly75), XCR 1 and Clec9A (DNNGR-1), is mainly localized in the white pulp T cell area (periarterial lymphoid sheets (PALS)), but also in the marginal zone of the murine spleen as well as in the lymph node T cell region (Gerner *et al.*, 2012; Kraal *et al.*, 1986; Idoyaga *et al.*, 2009; Dudziak *et al.*, 2007). The development of CD8⁺CD11b[−] DCs is strictly dependent on the presence of the transcription factors BATF3, IRF8, and Id2, while both CD8⁺CD11b[−] as well as CD8[−]CD11b⁺ DCs depend on the transcription factor Zbtb46 (Edelson *et al.*, 2011; Hildner *et al.*, 2008; Hashimoto *et al.*, 2011; Hacker *et al.*, 2003; McLellan *et al.*, 2002; Satpathy *et al.*, 2012; Meredith *et al.*, 2012). Interestingly, it was demonstrated that spleen resident CD8⁺CD11b[−] DCs can be further subdivided into CD8⁺CD103[−] and CD8⁺CD103⁺ DCs (McLellan *et al.*, 2002; Qiu *et al.*, 2009). Controversial data on the function of the latter ones described their ability to uptake blood borne apoptotic cells, which is a general feature of CD8⁺CD11b[−] DCs (Qiu *et al.*, 2009; Iyoda *et al.*, 2002). Notably, only in Balb/c mice CD8⁺CD103⁺ cells were also positive for Langerin (Idoyaga *et al.*, 2009; Qiu *et al.*, 2009). Kissenpfennig *et al.* (2005) also detected Langerin expression on thymic DCs and blood-derived CD8⁺CD11b[−] DCs in secondary lymphoid organs. Antibody immune complex engulfment, antigen protein uptake, but also *in vivo* antigen targeting studies using chimeric non-Fc receptor binding antibodies fused to an antigen of choice, revealed that in contrast to CD8[−]CD11b⁺ DCs murine CD8⁺CD11b[−] DCs are superior in cross-presentation and activation of CD8⁺ cytotoxic T cell responses (Iyoda *et al.*, 2002; Pooley *et al.*, 2001; Schnorrer *et al.*, 2006; Schulz and Reis E Sousa, 2002; Bonifaz *et al.*, 2002; Bonifaz *et al.*, 2004; Dudziak *et al.*, 2007; Idoyaga *et al.*, 2008; Soares *et al.*, 2007; den Haan *et al.*, 2000; Vander Lugt *et al.*, 2014). Notably, this was found to be correlating to an increased expression of molecules involved in the processing of antigens for MHC class I in CD8⁺CD11b[−] DCs (such as TAP1, TAP2, Tapasin, ERp57, Calreticulin, Calnexin, Sec61, and ERAAP). pMHC-I presentation of targeted antigens needed TAP transport as TAP deficient CD8⁺CD11b[−] DCs were unable to induce CD8⁺ cytotoxic T cell activation (Bonifaz *et al.*, 2002; den Haan *et al.*, 2000). Cystatin C is described as an inhibitor of cathepsin S and thereby a potential suppressor of MHC class II protein translocation and peptide loading. Although Cystatin C is overexpressed in MHC-II⁺LAMP1⁺ compartments of CD8⁺CD11b[−] DCs (Dudziak *et al.*, 2007; El-Sukkari *et al.*, 2003), data indicate that I_i degradation and MHC-II peptide loading of processed protein antigens was similar in isolated and cultured Cystatin C deficient CD8⁺CD11b[−] and CD8[−]CD11b⁺ DCs (El-Sukkari *et al.*, 2003). These data suggest that antigen cross-presentation might be superior in

CD8⁺CD11b[−] DCs and not directly dependent on Cystatin C expression.

Besides cell intrinsic features, the endocytic antigen uptake receptor might play a role in directing the processing and presentation capabilities in the DCs. Antigen targeting studies either using chimeric antibody antigen constructs or other antigen sources suggest that the binding site of the targeted receptor indirectly dictates the stability of the antigen, which might be due to differential routing of the antigens within the cell (Kreer *et al.*, 2012; Mintern *et al.*, 2015). It was further shown that both the CD8⁺CD11b[−] and the CD8[−]CD11b⁺ DC subset of human DEC205 receptor transgenic mice are able to cross-present ovalbumin antigen coupled to an anti-human DEC205 antibody, thereby indicating a specific role of the receptor in routing antigens (Kamphorst *et al.*, 2010). This is notable because CD8[−]CD11b⁺ DCs otherwise seem to be specialized for pMHC-II presentation (see below).

Overall, it is widely accepted that CD8⁺CD11b[−] DCs only use the cytosolic pathway for cross-presentation of antigens (Joffre *et al.*, 2012). In addition to the molecules involved in antigen processing for MHC-I, cross-presentation activity was demonstrated to be also dependent on the alkaline pH in the endosomal/phagosomal compartments in CD8⁺CD11b[−] DCs, thereby antigen degradation remains limited in this subset (Joffre *et al.*, 2012; Savina *et al.*, 2009). Data indicate that the cytosolic (p47phox) and membrane subunits (gp91phox and gp22phox) of the NADPH oxidase NOX2 assemble in CD8⁺CD11b[−] phagosomes, which leads to the production of reactive oxygen species (ROS) and thereby inhibition of phagosomal acidification. NOX2 recruitment was directly controlled by the small GTPase Rac2 (Savina *et al.*, 2009). All these data correlate with the finding that BATF3 deficient mice, which are missing the CD8⁺CD11b[−] DC subset, reveal less functional anti-viral and anti-tumor immune responses directly correlating with inefficient CD8⁺ T cell responses (Hildner *et al.*, 2008; Edelson *et al.*, 2010). Notably, antibody responses were not affected in these mice indicating a less pronounced function of CD8⁺CD11b[−] DCs in pMHC-II antigen presentation (Hildner *et al.*, 2008). Similar findings have been obtained by the use of a Langerin-DTR mouse, in which not only Langerin⁺ DCs from skin were depleted upon diphtheria toxin treatment, but also CD103⁺ and CD8⁺CD11b[−] lymphoid tissue DCs. In this mouse model, it became clear that those DCs are important for the priming of influenza, but also Leishmania, specific CD8⁺ T cell responses (GeurtsvanKessel *et al.*, 2008; Brewig *et al.*, 2009; Merad *et al.*, 2013). In addition to their ability to cross-present antigens to CD8⁺ cytotoxic T cells, CD8⁺CD11b[−] DCs are prominent producers of IL-12 upon stimulation, which is important for the differentiation of CD4⁺ T cells into T_H1 cells. Interestingly, even in T_H2 prone Balb/c mice T_H1 responses were promoted by CD8⁺CD11b[−] DCs (Soares *et al.*, 2007; Pulendran *et al.*, 1999; Maldonado-López *et al.*, 1999). As a 'signal 3' IL-12 is also crucial for 'classical cognate licensing' of CD8⁺ T cell responses, a process that involves the simultaneous recognition of linked antigens by a CD4⁺ helper T cell and a cytotoxic CD8⁺ T cell, thereby avoiding misguided cytotoxic CD8⁺ T cell activation and memory (Thaïss *et al.*, 2011; Schulz *et al.*, 2000). Moreover, under non-inflammatory conditions it was demonstrated that CD8⁺CD11b[−] DCs are

able to release TGF- β , which regulates the production of regulatory CD4⁺ T cells in the periphery, and thereby identifying CD8⁺CD11b[−] DCs as key players for the maintenance of peripheral T cell tolerance (Yamazaki *et al.*, 2008).

CD8[−]CD11b⁺ DCs

In contrast, CD8[−]CD11b⁺ DCs are strictly dependent on RbpJ/CBF1 (recombining binding protein suppressor of hairless), NF κ B, IRF2, and IRF4 transcription factor expression, but also on the presence of Zbtb46 (Satpathy *et al.*, 2012; Suzuki *et al.*, 2004; Caton *et al.*, 2007; Schlitzer *et al.*, 2013; Vander Lugt *et al.*, 2014). CD8[−]CD11b⁺ DCs can be found in the splenic marginal zone and red pulp as well as in the lymphatic zone of the lymph node (Germer *et al.*, 2012; Dudziak *et al.*, 2007). These cells additionally coexpress of CLR, including DCIR2 (mostly in the spleen), which is recognized by the monoclonal antibody 33D1, but also other immunoregulatory molecules such as the molecule signal-regulatory protein alpha (SIRP α /CD172a) (Dudziak *et al.*, 2007; Vremec *et al.*, 2000; Shortman and Liu 2002). CD8[−]CD11b⁺ DCs can be further subdivided into CD4⁺CD8[−]CD11b⁺ and CD4[−]CD8[−]CD11b⁺ DCs. However, subset specific differences have not been clearly defined yet (Vremec *et al.*, 2000).

The direct delivery of antigen using antigen targeting strategies indicates that in contrast to CD8⁺CD11b[−] DCs, CD8[−]CD11b⁺ DCs excel in MHC class II presentation and are superior in presenting endocytosed antigens to CD4⁺ helper T cells (Dudziak *et al.*, 2007; Soares *et al.*, 2007; Pulendran *et al.*, 2008; Yamazaki *et al.*, 2008; Boscardin *et al.*, 2006). This is correlated with an increased expression of molecules important for MHC-II antigen processing, for example, GILT, cathepsins, Legumain, and H2-DM (Dudziak *et al.*, 2007). Recently, it was postulated that IRF4 might be the most important factor not only for the development of CD8[−]CD11b⁺ DCs, but also for controlling the transcriptional program in this DC subset (Vander Lugt *et al.*, 2014). Interestingly, and in line with the differential expression pattern of the two main DC subpopulations, upon antigen engulfment CD8[−]CD11b⁺ DCs rapidly acidify their endosomal/lysosomal compartments, thus allowing for a more rapid proteolytic cleavage of the uptaken antigens and allocation of the peptides for MHC-II presentation. This process is regulated by the recruitment of the small GTPase Rac1 (Rac2 in CD8⁺CD11b[−] DCs), which is directing NOX2 not to the endosomal/lysosomal compartments, but rather to the plasma membrane (Savina *et al.*, 2009).

Antigen presentation by CD8[−]CD11b⁺ DCs demonstrates their ability to especially prime CD4⁺ T cell responses. This is leading to a CD4⁺ T cell derived secretion of the T_H2 cytokines IL4, IL5, and IL13, especially demonstrated in T_H2 prone Balb/c mice (Soares *et al.*, 2007; Maldonado-López *et al.*, 1999). Further, by measurement of the cytokine IFN γ but also by the indirect measurement of IgG subclasses, CD8[−]CD11b⁺ DCs have been demonstrated to be also able to induce CD4⁺ T_H1T cell responses (Do *et al.*, 2010; Soares *et al.*, 2007; Neubert *et al.*, 2014; Boscardin *et al.*, 2006). Interestingly, CD8[−]CD11b⁺ DCs might not directly secrete cytokines themselves. Data indicate that NOTCH ligands (such as DLL4) were found to be specifically generated by CD8[−]CD11b⁺ DCs, thereby enabling this DC subset to generate CD4⁺ T_H1 T cell responses even in the absence of IL12 (Skokos and

Nussenzweig, 2007; Tu *et al.*, 2005; Amsen *et al.*, 2004). Under non-stimulatory conditions the ability of CD8⁺CD11b⁺ DCs has been described to further expand inducible T_{reg} cells, thereby indicating a specific role in the maintenance of peripheral CD4⁺ T cell tolerance (Yamazaki *et al.*, 2008).

Importantly, under specific immunoactivatory conditions, even CD8⁺CD11b⁺ DCs might be capable to present exogenous antigens, for example, antibody immune complexes, for the induction of CD8⁺ T cell responses, as demonstrated by den Haan and Bevan (den Haan and Bevan, 2002). How this is regulated in this DC subset is under current investigation.

Non-Lymphoid Tissue DC

Non-lymphoid tissue and skin DCs display a heterogeneous population characterized by a differential expression of CD11b, CD103, and CD207 (Henri *et al.*, 2010; Malissen *et al.*, 2014). In addition, they exhibit select patterns of TLRs, CLRs, chemokine receptors, and production of distinct cytokines. In the following, we will only provide a brief overview of those DC subsets. Further details can be obtained from excellent reviews by the group of Miriam Merad (Merad *et al.*, 2013; Helft *et al.*, 2010).

CD103⁺CD11b[−] DCs

The integrin α chain (CD103) positive CD103⁺CD11b[−] DCs are phenotypically and transcriptionally similar to lymphoid tissue CD8⁺CD11b[−] DCs as they have the same developmental origin. It was demonstrated that CD103⁺CD11b[−] DCs are also dependent on IRF8, BATF3, and Id2 transcription factors (Ginhoux *et al.*, 2009; Edelson *et al.*, 2010). CD103⁺CD11b[−] DCs can be found at the interface to the environment in intestinal, pancreatic, and pulmonary tissues. They accumulate in Peyer's patches but overall they never reach more than 20–30% of CD11b⁺ DCs (Merad *et al.*, 2013). CD103⁺CD11b[−] DCs coexpress MHC-II, partly Langerin (beside pancreatic and intestinal tissues), CD8 (beside lamina propria tissue) as well as Flt3, but lack macrophage markers such as CD115 (c-kit), CX3CR1, CD11b, or F4/80 (Merad *et al.*, 2013).

Upon antigen encounter CD103⁺CD11b[−] DCs migrate to the T cell zones of draining lymph nodes (Gerner *et al.*, 2012). It was demonstrated that also CD103⁺CD11b[−] DCs, similar to CD8⁺CD11b[−] lymphoid resident DCs, are capable to cross-present antigens to CD8⁺ cytotoxic T cells. This is correlating with the finding that CD103⁺CD11b[−] as well as CD8⁺CD11b[−] DCs coexpress the chemokine receptor XCR1, which interacts with the XCL1 ligand produced by CD8⁺ T cells for the interaction of the antigen presenting DC with the specific T cell. If there are differences between the CD103⁺CD11b[−] and CD8⁺CD11b[−] DCs in the escape of the engulfed antigens, the pH regulation in the endosomal/lysosomal compartments, or the antigen processing machinery is not clear at the moment. As described above in BATF3^{−/−} mice, both CD8⁺CD11b[−] and CD103⁺CD11b[−] DCs are diminished leading to reduced antiviral and antitumor immune responses. However, antibody responses were not abrogated in this mouse model (Hildner *et al.*, 2008; Edelson *et al.*, 2010).

At the moment a specific mouse model to decipher the roles of lymphoid tissue-resident CD8⁺CD11b[−] and non-lymphoid tissue CD103⁺CD11b[−] DCs is not available.

CD11b⁺ DCs

Until today it is still challenging to truly define the function of non-lymphoid tissue CD11b⁺ DCs. They are localized in similar tissues as CD103⁺CD11b[−] DCs, but can be less accurately distinguished from tissue macrophages. This is due to the finding that tissue macrophages also express the surface molecule CD11c. Thereby divergent functional descriptions of CD11b⁺ DCs exist. In muscle CD64 (Fc γ RI) was described to be an excellent marker for the definition of monocyte derived DCs, thereby distinguishing them from tissue CD11b⁺ DCs (Langlet *et al.*, 2012). In the lamina propria CD11b⁺ DCs can be separated into CD11b⁺CD103⁺ and CD11b⁺CD103[−] DCs (Bogunovic *et al.*, 2009; Schulz *et al.*, 2009). CD11b⁺ DCs migrate to the lymph nodes when they have engulfed pathogenic antigens (Gerner *et al.*, 2012). Those migratory DCs (see below) have been described to induce mainly CD4⁺ T cell responses either upon virus infection or antigen encounter in combination with TLR ligation. Further data indicate that non-lymphoid tissue resident CD11b⁺ DCs have also tolerogenic functions as they express aldehyde dehydrogenase (ALDH), which is necessary to metabolize vitamin A into retinoic acid. Retinoic acid is described to be important for the induction of regulatory T cells (Merad *et al.*, 2013; Guillemins *et al.*, 2010). A mouse model in which only non-resident CD11b⁺ DCs are deleted is not available yet.

Langerhans cells and migratory DCs

In the murine skin as a non-lymphoid tissue five different CD11c⁺MHC-II⁺ DC subpopulations have been described so far (Malissen *et al.*, 2014; Merad *et al.*, 2008; Tamoutounour *et al.*, 2013; Mollah *et al.*, 2014; Henri *et al.*, 2010). Beside CD8, CD11b, and CD103, the cell adhesion molecule EpCAM as well as XCR1, CD24, and the C-type lectin receptor Langerin (CD207) are used for a proper subdivision of skin DCs. The skin consists of epidermis, dermis, and subcutaneous tissue. Langerhans cells (CD103[−]EpCAM⁺Langerin⁺) reside in the murine epidermis. In the dermis, beside CD103[−]CD11b⁺EpCAM⁺Langerin⁺ migratory Langerhans cells, four other DC subpopulations can be defined (Pulendran *et al.*, 2008; Romani *et al.*, 2010; Schulz and Steinman, 1985). These are CD103⁺CD11b^{low}EpCAM[−]Langerin⁺ dermal DCs (CD103⁺Langerin⁺ dermal DCs), CD103[−]CD11b^{high}EpCAM[−]Langerin[−] dermal DCs (CD11b⁺ dermal DCs), CD103[−]CD11b^{low}EpCAM[−]Langerin⁺ dermal DCs (CD103[−]Langerin⁺ dermal DCs), CD103[−]CD11b^{low}EpCAM[−]Langerin[−] dermal DCs (CD103[−]Langerin[−] dermal DCs), and CD103[−]CD11b⁺EpCAM⁺Langerin⁺ migratory Langerhans cells (Poulin *et al.*, 2007; Nagao *et al.*, 2009; Stoitzner *et al.*, 2003). Migratory Langerhans cells can also be visualized in skin-draining lymph nodes (Stoitzner *et al.*, 2003). CD103⁺Langerin⁺ dermal DCs can be distinguished from migratory Langerhans cells by their lack of Birbeck granules and the absence of the cell adhesion molecule EpCAM. Also the two Langerin[−]CD103[−]EpCAM[−] dermal DCs lack the formation of Birbeck granules (Merad *et al.*, 2008).

CD11b⁺ dermal DCs have been found to migrate into the interfollicular zone of the draining lymph node (Gerner *et al.*,

2012). Further, dermal and pulmonary migratory CD11b⁺ DCs have been described to efficiently prime CD4⁺ T cell responses upon infection with HSV-1 or influenza virus *ex vivo* (Bedoui *et al.*, 2009; Kim and Braciale, 2009), as well as against *Candida albicans* (Igyártó *et al.*, 2011). These findings are in line with the observation that Langerin⁺ dermal DCs are able to induce CD4⁺ responses, but also regulatory T cell responses, when antigens (either protein or by antigen targeting to Langerin⁺ cells) were injected subcutaneously in combination with an adjuvant (Kastenmüller *et al.*, 2011; Idoyaga *et al.*, 2008). Moreover, CD103⁺Langerin⁺ dermal DCs seem to be related to splenic CD8⁺CD11b⁺ cDCs in regard to their cross-presentation capacity and Th1 CD4⁺ T cell priming potential (Ginhoux *et al.*, 2009; Bedoui *et al.*, 2009). In addition, also CD103[−] dermal DCs were described in murine skin-draining lymph nodes, with the capacity to produce retinoic acid and a subsequent induction of Foxp3⁺ regulatory T cells (Guilliams *et al.*, 2010). Interestingly, the depletion of this DC subset in Langerin-DTR mice abolished the induction of CD4⁺ Th1 responses in a model of autoimmune encephalomyelitis (King *et al.*, 2010). In contrast, although CD8⁺ T cell responses are strongly reduced in this mouse model against *Leishmania major*, CD4⁺ T cell responses were not affected. These findings demonstrate that both the nature of the encountered pathogen and the mouse system used are important for the DC-induced immune response.

Langerhans cells have been described as the most efficient DCs in T cell priming. Overall, Langerhans cells are less efficient than previously thought after the analysis of a variety of mouse models, in which Langerhans cells are depleted. In contact hypersensitivity analyses, evidence was raised that dermal CD103⁺ DCs are able to take over Langerhans cell function (Kaplan *et al.*, 2005; Kaplan, 2010; Edelson *et al.*, 2010). Further, HSV-1 infection of Langerhans cells did not induce an efficient CD8⁺ T cell response, while other dermal DCs did. Interestingly, Langerhans cells might play a role in tolerance induction but also in the induction of Th17 responses induced by *Candida albicans* (Igyártó *et al.*, 2011; Merad *et al.*, 2013).

Functional Specialization in Antigen Presentation of Human DC Subpopulations

Lymphoid Tissue DCs

In contrast to the murine DC system, the picture of human DC subpopulations is rather incomplete. Up to date, most knowledge is obtained from blood DCs or MoDCs (Sallusto and Lanzavecchia, 1994; Shortman and Liu, 2002; MacDonald *et al.*, 2002; Dzionek *et al.*, 2000). However, in recent years insight into the biology of human tissue DCs has steadily increased (Haniiffa *et al.*, 2012; Segura *et al.*, 2012; Segura and Amigorena, 2014; Klechevsky *et al.*, 2008; Paessens *et al.*, 2008; Mittag *et al.*, 2011; McGovern *et al.*, 2014a). Limiting factors are ethical constraints, difficult availability of healthy human organs, and low frequencies of the various human tissue DC subpopulations. For example, human blood DCs account for only 0.16–0.68% of circulating peripheral blood leukocytes (Haller Hasskamp *et al.*, 2005). DCs are commonly referred to as lineage negative (LIN[−]), which discriminates them from

other cell types such as T cells, B cells, NK cells, and erythrocytes by the lack of expression of the cell surface molecules CD3, CD19 and CD20, CD56, and CD235a, respectively (Merad *et al.*, 2013). Further, anti-CD14 and anti-CD16 (FcγRIII) antibodies are frequently included in the lineage cocktail. However, attention has to be paid on the analyzed tissue, since CD14⁺ lymph node and dermis-resident DC subpopulations have been postulated (Segura *et al.*, 2012; Nestle *et al.*, 1994). However, based on phenotypic and transcriptome studies these cells might rather be of monocyte/macrophage origin (Segura *et al.*, 2012; Schlitzer and Ginhoux, 2014). In addition, the existence of a CD16⁺ slanDC population in human peripheral blood and skin has been reported, but is still under debate (MacDonald *et al.*, 2002; Schäkel *et al.*, 2002; Hänsel *et al.*, 2011; Döbel *et al.*, 2013; Cros *et al.*, 2010). Nevertheless, the lineage cocktail composition needs further clarification and tissue-specific refinement.

Efforts have been made to unify the nomenclature of human DC subpopulations (Ziegler-Heitbrock *et al.*, 2010; Guilliams *et al.*, 2014). As a consequence, various alternatives can be found in the literature. Under non-inflammatory conditions two major groups of DC subpopulations (conventional DCs (cDCs) and pDCs), both expressing HLA-DR (human leukocyte antigen DR, the human equivalent to murine MHC-II), have been described in human lymphoid tissues such as blood, spleen, tonsils, lymph nodes, and bone marrow (Merad *et al.*, 2013; Dzionek *et al.*, 2000; McIlroy *et al.*, 2001; Poulin *et al.*, 2010; Mittag *et al.*, 2011; Segura *et al.*, 2013, 2012; Jongbloed *et al.*, 2010; Nizzoli *et al.*, 2013). The phenotype of human DC subpopulations in further lymphoid tissues (e.g., thymus or cord blood) is under current investigation. Conventional DCs contribute two subpopulations. The first, named CD1c⁺ DCs (also referred to as cDC2 DCs, myeloid-derived DCs type 1 (mDC1 DCs), CD11b-like DCs, or BDCA1⁺ DCs), shows high levels of CD1c, SIRPα (CD172a), and CD11c. The second conventional DC subpopulation is represented by CD141⁺ DCs (also called cDC1 DCs, mDC2 DCs, CD8-like DCs, or BDCA3⁺ DCs), which are characterized by the expression of CD141 (BDCA3), Clec9A, XCR1, Nect2, an intermediate level of CD11c, and by the absence of CD1c (BDCA-1), CD14, SIRPα, and CD123 (Merad *et al.*, 2013; Dzionek *et al.*, 2000; Huysamen *et al.*, 2008; McGovern *et al.*, 2014b; Bachem *et al.*, 2010). The third major blood DC subpopulation are plasmacytoid DCs (pDCs), which are CD1c[−] and CD11c^{low}, but can be identified by the expression of CD303 (BDCA2), CD304 (BDCA-4), and CD123 (Merad *et al.*, 2013; Dzionek *et al.*, 2000). As a key feature, pDCs rapidly respond to viral infection by the secretion of type I interferons (Cella *et al.*, 1999). In addition to these three major DC subpopulations, Segura *et al.* have identified three skin-derived migratory DC subpopulations in human axillary lymph nodes, namely migratory Langerhans cells, CD1⁺ DCs, and CD206⁺ DCs (Segura *et al.*, 2012).

The identification of the human functional ortholog of mouse CD8⁺ DCs proved to be challenging due to the lack of CD8 expression on human DCs (Shortman and Liu, 2002). In 2008 and 2010, several research groups independently declared CD141⁺ DCs as the human counterpart of the cross-presenting murine CD8⁺ DC subpopulation (Robbins *et al.*, 2008; Poulin *et al.*, 2010; Bachem *et al.*, 2010; Jongbloed *et al.*, 2010; Crozat *et al.*, 2010). The authors collectively pointed out similarities of mouse CD8⁺ DCs and human CD141⁺ DCs

based on cell surface marker expression (e.g., XCR1, Clec9A (DNCR-1)) and on the superiority of CD141⁺ DCs in the cross-presentation of soluble or cell-associated HCMV protein pp65 or necrotic cell antigens (including viral proteins). Haniffa *et al.* (2012) demonstrated a pronounced capacity of blood CD141⁺ DCs to cross-present soluble hepatitis B surface antigen (HBsAg) after TLR3 specific poly(I:C) stimulation. The assumption that human CD141⁺ DCs are orthologous to murine CD8⁺ DCs gained further support by the description of the transcription factors BATF3 and IRF8, which are expressed in both cell types (Poulin *et al.*, 2010; Hildner *et al.*, 2008). However, in contrast to the mouse system the development of all human DC subpopulations as well as monocytes is abrogated due to a specific IRF8 mutation, which leads to impaired transcriptional activity, suggesting differential pathways in these two species (Hambleton *et al.*, 2011; Segura and Amigorena, 2014).

It was further documented in *in vitro* culture systems that CD141⁺ blood DCs display a superior XCL1-dependent migratory behavior (Bachem *et al.*, 2010; Crozat *et al.*, 2010). In addition, both the murine and the human cell surface molecule Clec9A seem to be responsible for sensing and cross-presentation of dead-cell-associated material *in vitro* and *in vivo* (mouse) via SYK kinase signaling (Sancho *et al.*, 2009; Poulin *et al.*, 2010). Thus, the C-type lectin receptor Clec9A does not seem to be directly involved in the uptake mechanism, but rather in routing endocytosed dead-cell material into non-lysosomal compartments, therefore facilitating cross-presentation (Sancho *et al.*, 2009). Accordingly, it was reported by Crozat *et al.* that the Clec9A expressing human CD141⁺ blood DC subpopulation excels in the cross-presentation ability of apoptotic cell-associated HIV peptides under poly(I:C) stimulated conditions *in vitro* (Crozat *et al.*, 2010).

However, evidence has emerged in recent years that most if not all human DC subpopulations are capable of cross-presenting antigen. Segura and Amigorena have nicely summarized recent findings on the capacity of human blood and lymphoid organ DCs as well as skin DC subpopulations for the cross-presentation of long peptides, soluble antigens, cell-associated antigens, or antigens delivered via receptor targeting (Segura and Amigorena, 2014). Thus, it seems that cross-presentation is not restricted to intrinsic properties of a single DC subpopulation in humans, but is rather dependent on a combination of circumstances. The type of antigen, and thus the uptake mechanism as well as pathogen-intrinsic properties play a role in the route of antigen presentation. For example, *Listeria monocytogenes* produce listeriolysin O (LLO), facilitating lysosomal escape and leading to MHC-I presentation after proteasomal degradation (Cohn *et al.*, 2013). Further data indicate that the kind of stimulatory environment as well as the timing and stage of the immune reaction are decisive for the cross-presentation capacity as summarized by Nierkens *et al.* (Nierkens *et al.*, 2013). Many earlier studies have used poly(I:C) as the stimulatory adjuvant, which favors CD141⁺ DCs due to their high TLR3 expression (Jongbloed *et al.*, 2010). However, more appropriate TLR stimuli resulted in efficient cross-presentation of CMV pp65 protein by blood CD1c⁺ and CD141⁺ DCs (Nizzoli *et al.*, 2013). Of note, a comprehensive analysis of the cross-presentation capacity of the various DC subpopulations from human lymphoid

and non-lymphoid tissues considering different stimuli and administered antigens remains missing. These *in vitro* data can only provide limited insight into human DC cross-presentation, since different contributions of the various human DC subpopulations in the resolution of an infection add more complexity to the topic. Regarding the T helper cell polarization potential accumulating data suggest that CD141⁺ DCs are specialized in the induction of T_H1 T cell responses, whereas CD1c⁺ DCs predominantly elicit T_H2 and T_H17 responses (Schlitzer and Ginhoux, 2014; McGovern *et al.*, 2014a). In contrast, other research groups believe that CD141⁺ DCs mainly induce T_H2 T cell responses, while CD1c⁺ DCs prefer the initiation of T_H1 T cell polarization (Yu *et al.*, 2014). It was further demonstrated that human blood and tonsillar CD1c⁺ DCs exceeded in the production of IL-12 and were highly efficient in cross-presenting soluble CMV pp65 protein under appropriate TLR stimulation in comparison to the corresponding CD141⁺ DC or pDC subpopulations (Nizzoli *et al.*, 2013). In conclusion, further in-depth analyses will have to shed light on the T helper cell polarization potential of human lymphoid (and non-lymphoid) tissue DC subpopulations.

Non-Lymphoid Tissue (NLT) DCs (Skin) and Migratory DCs

In line with murine DCs, human DC subpopulations exhibit sentinel functions not only in the blood and primary lymphoid organs, but also in non-lymphoid tissues. Therefore, human DCs have been described in the lungs, liver, gut, and skin (e.g., Haniffa *et al.*, 2012; Watchmaker *et al.*, 2014; Boltjes and van Wijk, 2014). Romani and colleagues described the skin as the home of epidermal CD1⁺CD207⁺EpCAM⁺ Langerhans cells (with their characteristic Birbeck granules) and dermal Langerin⁺ DCs, ideally positioned to take up antigen for a subsequent presentation to antigen-specific T cells in a skin-draining lymph node. *In vitro* studies have demonstrated that Langerhans cells are capable of initiating T_H1, T_H2, T_H17, and T_H22 CD4⁺ T cell polarization as well as CD8⁺ T cell responses via cross-presentation (Klechevsky, 2013).

Dermal Langerin⁺ DCs can be further subdivided into CD14⁺ and CD1⁺ dermal DCs, of which the latter population harbors CD1c⁺ DCs and CD141⁺ DCs (Romani *et al.*, 2010b; Boltjes and van Wijk, 2014). Skin-derived CD141⁺ DCs were shown to exhibit a superior capacity to cross-present soluble antigens, in comparison to CD1c⁺ DCs (Haniffa *et al.*, 2012). Of note, transcriptomic analyses of dermal CD14⁺ DCs have led to the assumption that these cells are of rather monocytic origin (McGovern *et al.*, 2014a). Under inflammatory conditions, additional DC subpopulations migrate into the skin, such as pDCs, 'Tip-DCs', and inflammatory dendritic epidermal cells (IDEC) (Romani *et al.*, 2010a,b; Boltjes and van Wijk, 2014).

Conclusions

- DCs are important antigen presenting cells
- DCs are key regulators for the induction of immune responses, but also for maintaining peripheral tolerance

- DCs can be stimulated upon recognition of microbial and viral antigens leading to migration, upregulation of co-stimulatory molecules, and antigen processing
- DCs present antigenic peptides as pMHC-I or pMHC-II complexes to CD8⁺ or CD4⁺ T cells
- At least four different ways of antigen processing with (1) classical MHC-I presentation, (2) classical MHC-II presentation, (3) cross-presentation, and (4) autophagy are described
- Cross-presentation is defined as most important processing way for the induction of immune responses against viral and tumor antigens and is tightly regulated
- Variety of DC subsets with differential function in antigen processing and presentation
- Overall, distinction into (1) classical/conventional resident and migratory DCs in lymphoid and non-lymphoid tissues and (2) plasmacytoid DCs
- In the mouse system: CD8⁺CD11b[−] (also CD103⁺) DCs efficient in cross-presentation to CD8⁺ T cells
- In the human system: various reports with different findings on DC subset functions indicating (so far) no specialization of a defined DC subset for cross-presentation.

Acknowledgments

This work was supported by the German Research Foundation (DU548/2-1, CRC643, RTG1962, RTG1660 to D.D.), the Bavarian Genome Research Network and by an intramural ELAN grant to G.H. (DE-14-10-17-1-Heidkamp) and is partially based on the authors' PhD or Masters' theses.

See also: Cellular Immunology: Innate Immunity: Dendritic Cells

References

- Ackerman, A.L., Giodini, A., Cresswell, P., 2006. A role for the endoplasmic reticulum protein retrotranslocation machinery during crosspresentation by dendritic cells. *Immunity* 25, 607–617.
- Agarraberes, F.A., Terlecky, S.R., Dice, J.F., 1997. An intralysosomal hsp70 is required for a selective pathway of lysosomal protein degradation. *Journal of Cell Biology* 137 (4), 825–834.
- Akir, S., Uematsu, S., Takeuchi, O., 2006. Pathogen recognition and innate immunity. *Cell* 124 (4), 783–801. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16497588>.
- Amsen, D., Blander, J.M., Lee, G.R., *et al.*, 2004. Instruction of distinct CD4 T helper cell fates by different notch ligands on antigen-presenting cells. *Cell* 117 (4), 515–526.
- Arrigo, A.P., Tanaka, K., Goldberg, A.L., *et al.*, 1988. Identity of the 19S “prosome” particle with the large multifunctional protease complex of mammalian cells (the proteasome). *Nature* 331, 192–194.
- Arunachalam, B., Phan, U.T., Geuze, H.J., Cresswell, P., 2000. Enzymatic reduction of disulfide bonds in lysosomes: Characterization of a gamma-interferon-inducible lysosomal thiol reductase (GILT). *Proceedings of the National Academy of Sciences of the United States of America* 97 (2), 745–750.
- Asselin-Paturel, C., Brizard, G., Chemin, K., *et al.*, 2005. Type I interferon dependence of plasmacytoid dendritic cell activation and migration. *Journal of Experimental Medicine* 201 (7), 1157–1167.
- Bachem, A., Güttler, S., Hartung, E., *et al.*, 2010. Superior antigen cross-presentation and XCR1 expression define human CD11c + CD141 + cells as homologues of mouse CD8 + dendritic cells. *Journal of Experimental Medicine* 207 (6), 1273–1281.
- Bailey-Bucktrout, S.L., Caulkins, S.C., Goings, G., *et al.*, 2008. Cutting edge: Central nervous system plasmacytoid dendritic cells regulate the severity of relapsing experimental autoimmune encephalomyelitis. *Journal of Immunology* 180 (10), 6457–6461.
- Bakke, O., Dobberstein, B., 1990. MHC class II-associated invariant chain contains a sorting signal for endosomal compartments. *Cell* 63, 707–716.
- Banchereau, J., Briere, F., Caux, C., *et al.*, 2000. Immunobiology of dendritic cells. *Annual Review of Immunology* 18, 767–811.
- Banchereau, J., Cohn, F., Inaba, K., *et al.*, 2011. Remembering Ralph Steinman. *Journal of Experimental Medicine* 208, 2343–2347.
- Banchereau, J., Steinman, R.M., 1998. Dendritic cells and the control of immunity. *Nature* 392 (6673), 245–252.
- Bandoh, N., Ogino, T., Cho, H.S., *et al.*, 2005. Development and characterization of human constitutive proteasome and immunoproteasome subunit-specific monoclonal antibodies. *Tissue Antigens* 66 (3), 185–194.
- Barral, D.C., Brenner, M.B., 2007. CD1 antigen presentation: How it works. *Nature Reviews. Immunology* 7 (12), 929–941.
- Basha, G., Omilusik, K., Chavez-Steenbock, A., *et al.*, 2012. A CD74-dependent MHC class I endolysosomal cross-presentation pathway. *Nature Immunology* 13 (3), 237–245.
- Bedoui, S., Whitney, P.G., Waithman, J., *et al.*, 2009. Cross-presentation of viral and self antigens by skin-derived CD103 + dendritic cells. *Nature Immunology* 10 (5), 488–495.
- Bendall, S.C., Nolan, G.P., Roederer, M., Chattopadhyay, P.K., 2012. A deep profiler's guide to cytometry. *Trends in Immunology* 33 (7), 323–332.
- Bevan, M.J., 1976. Minor H antigens introduced on H-2 different stimulating cells cross-react at the cytotoxic T cell level during *in vivo* priming. *Journal of Immunology* 117 (6), 2233–2238.
- Bjorkman, P., Saper, M.A., Samraoui, B., *et al.*, 1987. The foreign antigen binding site and T cell recognition regions of class I histocompatibility antigens. *Nature* 329, 512–518.
- Bogunovic, M., Ginhoux, F., Helft, J., *et al.*, 2009. Origin of the lamina propria dendritic cell network. *Immunity* 31 (3), 513–525.
- Boltjes, A., van Wijk, F., 2014. Human dendritic cell functional specialization in steady-state and inflammation. *Frontiers in Immunology* 5, 131.
- Bonifaz, L.C., Bonnyay, D., Mahnke, K., *et al.*, 2002. Efficient targeting of protein antigen to the dendritic cell receptor DEC-205 in the steady state leads to antigen presentation on major histocompatibility complex class I products and peripheral CD8 + T cell tolerance. *Journal of Experimental Medicine* 196 (12), 1627–1638.
- Bonifaz, L.C., Bonnyay, D.P., Charalambous, A., *et al.*, 2004. *In vivo* targeting of antigens to maturing dendritic cells via the DEC-205 receptor improves T cell vaccination. *Journal of Experimental Medicine* 199 (6), 815–824.
- Boscardin, S.B., Hafalla, J.C., Masilamani, R.F., *et al.*, 2006. Antigen targeting to dendritic cells elicits long-lived T cell help for antibody responses. *Journal of Experimental Medicine* 203 (3), 599–606.
- Brewig, N., Kissenpennig, A., Malissen, B., *et al.*, 2009. Priming of CD8 + and CD4 + T cells in experimental leishmaniasis is initiated by different dendritic cell subtypes. *Journal of Immunology (Baltimore, MD: 1950)* 182 (2), 774–783.
- Brossart, P., Bevan, M.J., 1997. Presentation of exogenous protein antigens on major histocompatibility complex class I molecules by dendritic cells: Pathway of presentation and regulation by cytokines. *Blood* 90, 1594–1599.
- Burgdorf, S., Kurts, C., 2008. Endocytosis mechanisms and the cell biology of antigen presentation. *Current Opinion in Immunology* 20 (1), 89–95.
- Burrows, S.R., Rossjohn, J., McCluskey, J., 2006. Have we cut ourselves too short in mapping CTL epitopes? *Trends in Immunology* 27 (1), 11–16.
- Cao, W., Lee, S.H., Lu, J., 2005. CD83 is preformed inside monocytes, macrophages and dendritic cells, but it is only stably expressed on activated dendritic cells. *Biochemical Journal* 385 (Pt 1), 85–93.
- Cascio, P., Hilton, C., Kisselev, A.F., Rock, K.L., Goldberg, A.L., 2001. 26S proteasomes and immunoproteasomes produce mainly N-extended versions of an antigenic peptide. *EMBO Journal* 20 (10), 2357–2366.
- Caton, M.L., Smith-Raska, M.R., Reizis, B., 2007. Notch-RBP-J signaling controls the homeostasis of CD8- dendritic cells in the spleen. *Journal of Experimental Medicine* 204 (7), 1653–1664.
- Cebrian, I., Visentin, G., Blanchard, N., *et al.*, 2011. Sec22b regulates phagosomal maturation and antigen crosspresentation by dendritic cells. *Cell* 147 (6), 1355–1368.

- Cella, M., Engering, A., Pinet, V., Pieters, J., Lanzavecchia, A., 1997. Inflammatory stimuli induce accumulation of MHC class II complexes on dendritic cells. *Nature* 388 (6644), 782–787.
- Cella, M., Jarrossay, D., Facchetti, F., *et al.*, 1999. Plasmacytoid monocytes migrate to inflamed lymph nodes and produce large amounts of type I interferon. *Nature Medicine* 5 (8), 919–923.
- Cervantes-Barragan, L., Lewis, K.L., Firner, S., *et al.*, 2012. Plasmacytoid dendritic cells control T-cell response to chronic viral infection. *Proceedings of the National Academy of Sciences* 109 (8), 3012–3017.
- Chang, S.-C., Momberg, F., Bhutani, N., Goldberg, A.L., *et al.*, 2005. The ER aminopeptidase, ERAP1, trims precursors to lengths of MHC class I peptides by a “molecular ruler” mechanism. *Proceedings of the National Academy of Sciences of the United States of America* 102 (47), 17107–17112.
- Chapman, H.A., 2006. Endosomal proteases in antigen presentation. *Current Opinion in Immunology* 18 (1), 78–84.
- Chattopadhyay, P.K., Hogerkerp, C.M., Roederer, M., 2008. A chromatic explosion: The development and future of multiparameter flow cytometry. *Immunology* 125, 441–449.
- Chattopadhyay, P.K., Roederer, M., 2012. Cytometry: Today’s technology and tomorrow’s horizons. *Methods* 57, 251–258.
- Chopin, M., Allan, R.S., Belz, G.T., 2012. Transcriptional regulation of dendritic cell diversity. *Frontiers in Immunology* 3, 1–12.
- Chow, A., Toomre, D., Garrett, W., Mellman, I., 2002. Dendritic cell maturation triggers retrograde MHC class II transport from lysosomes to the plasma membrane. *Nature* 418, 988–994.
- Ciechanover, A., 2005. Proteolysis: From the lysosome to ubiquitin and the proteasome. *Nature Reviews. Molecular Cell Biology* 6, 79–87.
- Cohn, L., Chatterjee, B., Esselborn, F., *et al.*, 2013. Antigen delivery to early endosomes eliminates the superiority of human blood BDCA3+ dendritic cells at cross presentation. *Journal of Experimental Medicine* 210 (5), 1049–1063.
- Colbert, J.D., Matthews, S.P., Miller, G., Watts, C., 2009. Diverse regulatory roles for lysosomal proteases in the immune response. *European Journal of Immunology* 39, 2955–2965.
- Cole, D.K., Edwards, E.S., Wynn, K.K., *et al.*, 2010. Modification of MHC anchor residues generates heteroclitic peptides that alter TCR binding and T cell recognition. *Journal of Immunology (Baltimore, MD: 1950)* 185, 2600–2610.
- Conus, S., Perozzo, R., Reinheckel, T., *et al.*, 2008. Caspase-8 is activated by cathepsin D initiating neutrophil apoptosis during the resolution of inflammation. *Journal of Experimental Medicine* 205 (3), 685–698.
- Conus, S., Simon, H.U., 2010. Cathepsins and their involvement in immune responses. *Swiss Medical Weekly* 140, 1–8.
- Cros, J., Cagnard, N., Woollard, K., *et al.*, 2010. Human CD14dim monocytes patrol and sense nucleic acids and viruses via TLR7 and TLR8 receptors. *Immunity* 33 (3), 375–386.
- Crotzer, V.L., Blum, J.S., 2010. Autophagy and adaptive immunity. *Immunology* 131 (1), 9–17.
- Crozat, K., Guiton, R., Contreras, V., *et al.*, 2010. The XC chemokine receptor 1 is a conserved selective marker of mammalian cells homologous to mouse CD8alpha+ dendritic cells. *Journal of Experimental Medicine* 207 (6), 1283–1292.
- Cuervo, A.M., Dice, J.F., 1996. A receptor for the selective uptake and degradation of proteins by lysosomes. *Science (New York, NY)* 273 (5274), 501–503.
- de Heer, H.J., Hammad, H., Soullié, T., *et al.*, 2004. Essential role of lung plasmacytoid dendritic cells in preventing asthmatic reactions to harmless inhaled antigen. *Journal of Experimental Medicine* 200 (1), 89–98.
- Delamarre, L., Pack, M., Chang, H., Mellman, I., Trombetta, E.S., 2005. Differential lysosomal proteolysis in antigen-presenting cells determines antigen fate. *Science (New York, NY)* 307 (5715), 1630–1634.
- Denzin, L.K., 2013. Inhibition of HLA-DM mediated MHC class II peptide loading by HLA-DO promotes self tolerance. *Frontiers in Immunology* 4, 1–5.
- Diefenbach, A., Jensen, E.R., Jamieson, A.M., Raulet, D.H., 2001. Rae1 and H60 ligands of the NKG2D receptor stimulate tumour immunity. *Nature* 413 (1991), 165–171.
- Dieu, M.C., Vanbervliet, B., Vicari, A., *et al.*, 1998. Selective recruitment of immature and mature dendritic cells by distinct chemokines expressed in different anatomic sites. *Journal of Experimental Medicine* 188 (2), 373–386.
- Do, Y., Koh, H., Park, C.G., *et al.*, 2010. Targeting of LcrV virulence protein from *Yersinia pestis* to dendritic cells protects mice against pneumonic plague. *European Journal of Immunology* 40 (10), 2791–2796.
- Döbel, T., Kunze, A., Babatz, J., *et al.*, 2013. FcγRIII (CD16) equips immature 6-sulfo LacNAc-expressing dendritic cells (sIaDCs) with a unique capacity to handle IgG-complexed antigens. *Blood* 121 (18), 3609–3618.
- Doulatov, S., Notta, F., Eppert, K., *et al.*, 2010. Revised map of the human progenitor hierarchy shows the origin of macrophages and dendritic cells in early lymphoid development. *Nature Immunology* 11 (7), 585–593.
- Duan, F., Srivastava, P.K., 2012. An invariant road to cross-presentation. *Nature Immunology* 13 (3), 207–208.
- Dudzziak, D., Kamphorst, A.O., Heidkamp, G.F., *et al.*, 2007. Differential antigen processing by dendritic cell subsets *in vivo*. *Science* 315 (5808), 107–111.
- Dzionek, A., Fuchs, A., Schmidt, P., *et al.*, 2000. BDCA-2, BDCA-3, and BDCA-4: Three markers for distinct subsets of dendritic cells in human peripheral blood. *Journal of Immunology (Baltimore, MD: 1950)* 165 (11), 6037–6046.
- Edelson, B.T., Bradstreet, T.R., Wumesh, K.C., *et al.*, 2011. Batf3-dependent CD11b low/- peripheral dendritic cells are GM-CSF-independent and are not required for th cell priming after subcutaneous immunization. *PLoS One* 6 (10), 1–10.
- Edelson, B.T., Wumesh, K.C., Juang, R., *et al.*, 2010. Peripheral CD103+ dendritic cells form a unified subset developmentally related to CD8alpha+ conventional dendritic cells. *Journal of Experimental Medicine* 207 (4), 823–836.
- Edwards, A.D., Chaussabel, D., Tomlinson, S., *et al.*, 2003. Relationships among murine CD11c high dendritic cell subsets as revealed by baseline gene expression patterns. *Journal of Immunology* 171, 47–60.
- Eissing, N., Heger, L., Baranska, A., *et al.*, 2014. Easy performance of 6-color confocal immunofluorescence with 4-laser line microscopes. *Immunology Letters* 161 (1), 1–5.
- Ekeruche-Makinde, J., Miles, J.J., van den Berg, H.A., *et al.*, 2013. Peptide length determines the outcome of TCR/peptide-MHCI engagement. *Blood* 121 (7), 1112–1123.
- El-Sukkari, D., Wilson, N.S., Hakansson, K., *et al.*, 2003. The protease inhibitor cystatin C is differentially expressed among dendritic cell populations, but does not control antigen presentation. *Journal of Immunology (Baltimore, MD: 1950)* 171 (10), 5003–5011.
- Elinav, E., Strowig, T., Henao-Mejia, J., Flavell, R.A., 2011. Regulation of the antimicrobial response by NLR proteins. *Immunity* 34 (5), 665–679.
- Evnouchidou, I., Weimershaus, M., Saveanu, L., van Ender, P., 2014. ERAP1-ERAP2 dimerization increases peptide-trimming efficiency. *Journal of Immunology (Baltimore, MD: 1950)* 193 (2), 901–908.
- Fader, C.M., Colombo, M.I., 2009. Autophagy and multivesicular bodies: Two closely related partners. *Cell Death and Differentiation* 16, 70–78.
- Fallas, J.L., Tobin, H.M., Lou, O., *et al.*, 2004. Ectopic expression of HLA-DO in mouse dendritic cells diminishes MHC class II antigen presentation. *Journal of Immunology (Baltimore, MD: 1950)* 173 (3), 1549–1560.
- Figdor, C.G., van Kooyk, Y., Adema, G.J., 2002. C-type lectin receptors on dendritic cells and Langerhans cells. *Nature Reviews. Immunology* 2 (2), 77–84.
- Filomeni, G., De Zio, D., Cecconi, F., 2014. Oxidative stress and autophagy: The clash between damage and metabolic needs. *Cell Death and Differentiation* 22 (3), 377–388.
- Fogg, D.K., Sibon, C., Miled, C., *et al.*, 2006. A clonogenic bone marrow progenitor specific for macrophages and dendritic cells. *Science* 311 (5757), 83–87.
- Förster, R., Schubel, A., Breitfeld, D., *et al.*, 1999. CCR7 coordinates the primary immune response by establishing functional microenvironments in secondary lymphoid organs. *Cell* 99 (1), 23–33.
- Früh, K., Gossen, M., Wang, K., *et al.*, 1994. Displacement of housekeeping proteasome subunits by MHC-encoded LMPs: A newly discovered mechanism for modulating the multicatalytic proteinase complex. *EMBO Journal* 13 (14), 3236–3244.
- Gaczynska, M., Rock, K.L., Spies, T., Goldberg, A.L., 1994. Peptidase activities of proteasomes are differentially regulated by the major histocompatibility complex-encoded genes for LMP2 and LMP7. *Proceedings of the National Academy of Sciences of the United States of America* 91 (20), 9213–9217.
- Geissmann, F., Manz, M.G., Jung, S., *et al.*, 2010. Development of monocytes, macrophages, and dendritic cells. *Science* 327 (5966), 656–661.
- Gerner, M.Y., Kastenmuller, W., Ifrim, I., Kabat, J., Germain, R.N., 2012. Histocytometry: A method for highly multiplex quantitative tissue imaging analysis applied to dendritic cell subset microanatomy in lymph nodes. *Immunity* 37 (2), 364–376.
- GeurtsvanKessel, C.H., Willart, M.A., van Rijt, L.S., *et al.*, 2008. Clearance of influenza virus from the lung depends on migratory langerin+ CD11b- but not plasmacytoid dendritic cells. *Journal of Experimental Medicine* 205 (7), 1621–1634.
- Gil-Torregrosa, B.C., Lennon-Duménil, A.M., Kessler, B., *et al.*, 2004. Control of cross-presentation during dendritic cell maturation. *European Journal of Immunology* 34 (2), 398–407.
- Ginhoux, F., Liu, K., Helft, J., *et al.*, 2009. The origin and development of nonlymphoid tissue CD103+ DCs. *Journal of Experimental Medicine* 206 (13), 3115–3130.

- Goubier, A., Dubois, B., Gheit, H., *et al.*, 2008. Plasmacytoid dendritic cells mediate oral tolerance. *Immunity* 29 (3), 464–475.
- Guermonprez, P., Saveanu, L., Kleijmeer, M., *et al.*, 2003. ER-phagosome fusion defines an MHC class I cross-presentation compartment in dendritic cells. *Nature* 425 (6956), 397–402.
- Guermonprez, P., Valladeau, J., Zitvogel, L., Théry, C., Amigorena, S., 2002. Antigen presentation and T cell stimulation by dendritic cells. *Annual Review of Immunology* 20, 621–667.
- Guilliams, M., Bruhns, P., Saeys, Y., Hammad, H., Lambrecht, B.N., 2014a. The function of Fcγ receptors in dendritic cells and macrophages. *Nature Reviews. Immunology* 14, 94–108.
- Guilliams, M., Crozat, K., Henri, S., *et al.*, 2010. Skin-draining lymph nodes contain dermis-derived CD103(–) dendritic cells that constitutively produce retinoic acid and induce Foxp3(+) regulatory T cells. *Blood* 115 (10), 1958–1968.
- Guilliams, M., Ginhoux, F., Jakubczak, C., *et al.*, 2014b. Dendritic cells, monocytes and macrophages: A unified nomenclature based on ontogeny. *Nature Reviews. Immunology* 14 (8), 571–578.
- den Haan, J.M.M., Bevan, M.J., 2002. Constitutive versus activation-dependent cross-presentation of immune complexes by CD8+ and CD8– dendritic cells *in vivo*. *Journal of Experimental Medicine* 196 (6), 817–827.
- den Haan, J.M.M., Lehar, S.M., Bevan, M.J., 2000. CD8(+) but not CD8(–) dendritic cells cross-prime cytotoxic T cells *in vivo*. *Journal of Experimental Medicine* 192 (12), 1685–1696.
- Hacker, C., Kirsch, R.D., Ju, X.S., *et al.*, 2003. Transcriptional profiling identifies Id2 function in dendritic cell development. *Nature Immunology* 4 (4), 380–386.
- Hadeiba, H., Sato, T., Habtezion, A., *et al.*, 2008. CCR9 expression defines tolerogenic plasmacytoid dendritic cells able to suppress acute graft-versus-host disease. *Nature Immunology* 9 (11), 1253–1260.
- Haller Hasskamp, J., Zapas, J.L., Elias, E.G., 2005. Dendritic cell counts in the peripheral blood of healthy adults. *American Journal of Hematology* 78 (4), 314–315.
- Hambleton, S., Salem, S., Bustamante, J., *et al.*, 2011. IRF8 mutations and human dendritic-cell immunodeficiency. *New England Journal of Medicine* 365 (2), 127–138.
- Haniffa, M., Shin, A., Bigley, V., *et al.*, 2012. Human tissues contain CD141(hi) cross-presenting dendritic cells with functional homology to mouse CD103(+) nonlymphoid dendritic cells. *Immunity* 37 (1), 60–73.
- Hänsel, A., Günther, C., Ingwersen, J., *et al.*, 2011. Human slan (6-sulfo LacNAc) dendritic cells are inflammatory dermal dendritic cells in psoriasis and drive strong Th17/Th1 T-cell responses. *Journal of Allergy and Clinical Immunology* 127 (3), 787–794.
- Hashimoto, D., Miller, J., Merad, M., 2011. Dendritic cell and macrophage heterogeneity *In Vivo*. *Immunology* 35 (3), 323–335.
- Hawiger, D., Inaba, K., Dorsett, Y., *et al.*, 2001. Dendritic cells induce peripheral T cell unresponsiveness under steady state conditions *in vivo*. *Journal of Experimental Medicine* 194 (6), 769–779.
- Heath, W.R., Carbone, F.R., 2001. Cross-presentation, dendritic cells, tolerance and immunity. *Annual Review of Immunology* 19, 47–64.
- Heim, R., Cubitt, A.B., Tsien, R.Y., 1995. Improved green fluorescence. *Nature* 373 (6516), 663–664.
- Helft, J., Ginhoux, F., Bogunovic, M., Merad, M., 2010. Origin and functional heterogeneity of non-lymphoid tissue dendritic cells in mice. *Immunological Reviews* 234 (1), 55–75.
- Henri, S., Poulin, L., Tamoutounour, S., *et al.*, 2010. CD207+ CD103+ dermal dendritic cells cross-present keratinocyte-derived antigens irrespective of the presence of Langerhans cells. *Journal of Experimental Medicine* 207, 189–206.
- Hildner, K., Edelson, B.T., Purtha, W.E., *et al.*, 2008. Batf3 deficiency reveals a critical role for CD8α+ dendritic cells in cytotoxic T cell immunity. *Science* 322 (5904), 1097–1100.
- Hisamatsu, H., Shimbara, N., 1996. Newly identified pair of proteasomal subunits regulated reciprocally by interferon γ. *Journal of Experimental Medicine* 183 (4), 1097–1100.
- Honey, K., Rudensky, A.Y., 2003. Lysosomal cysteine proteases regulate antigen presentation. *Nature Reviews. Immunology* 3, 472–482.
- Hotta, C., Fujimaki, H., Yoshinari, M., Nakazawa, M., Minami, M., 2006. The delivery of an antigen from the endocytic compartment into the cytosol for cross-presentation is restricted to early immature dendritic cells. *Immunology* 117 (1), 97–107.
- Houde, M., Bertholet, S., Gagnon, E., *et al.*, 2003. Phagosomes are competent organelles for antigen cross-presentation. *Nature* 425, 402–406.
- Huysamen, C., Willment, J.A., Dennehy, K.M., Brown, G.D., 2008. CLEC9A is a novel activation C-type lectin-like receptor expressed on BDCA3+ dendritic cells and a subset of monocytes. *Journal of Biological Chemistry* 283 (24), 16693–16701.
- Itoyaga, J., Cheong, C., Suda, K., *et al.*, 2008. Cutting edge: Langerin/CD207 receptor on dendritic cells mediates efficient antigen presentation on MHC I and II products *in vivo*. *Journal of Immunology* (Baltimore, MD: 1950) 180 (6), 3647–3650.
- Itoyaga, J., Suda, N., Suda, K., Park, C.G., Steinman, R.M., 2009. Antibody to Langerin/CD207 localizes large numbers of CD8α+ dendritic cells to the marginal zone of mouse spleen. *Proceedings of the National Academy of Sciences of the United States of America* 106 (5), 1524–1529.
- Igyártó, B.Z., Haley, K., Ortner, D., *et al.*, 2011. Skin-resident murine dendritic cell subsets promote distinct and opposing antigen-specific T helper cell responses. *Immunity* 35 (2), 260–272.
- Ireland, J.M., Unanue, E.R., 2011. Autophagy in antigen-presenting cells results in presentation of citrullinated peptides to CD4 T cells. *Journal of Experimental Medicine* 208 (13), 2625–2632.
- Iwasaki, A., 2007. Mucosal dendritic cells. *Annual Review of Immunology* 25, 381–418.
- Iyoda, T., Shimoyama, S., Liu, K., *et al.*, 2002. The CD8+ dendritic cell subset selectively endocytoses dying cells in culture and *in vivo*. *Journal of Experimental Medicine* 195 (10), 1289–1302.
- Janeway, C.A., 1992. The T cell receptor as a multicomponent signalling machine: CD4/CD8 coreceptors and CD45 in T cell activation. *Annual Review of Immunology* 10, 645–674.
- Jensen, P.E., 2007. Recent advances in antigen processing and presentation. *Nature Immunology* 8 (10), 1041–1048.
- Joffre, O.P., Segura, E., Savina, A., Amigorena, S., 2012. Cross-presentation by dendritic cells. *Nature Reviews. Immunology* 12 (8), 557–569.
- Jongbloed, S.L., Benson, R.A., Nickdel, M.B., *et al.*, 2009. Plasmacytoid dendritic cells regulate breach of self-tolerance in autoimmune arthritis. *Journal of Immunology* 182 (2), 963–968.
- Jongbloed, S.L., Kassianos, A.J., McDonald, K.J., *et al.*, 2010. Human CD141+ (BDCA-3)+ dendritic cells (DCs) represent a unique myeloid DC subset that cross-presents necrotic cell antigens. *Journal of Experimental Medicine* 207 (6), 1247–1260.
- Joyce, S., 1997. Traffic control of completely assembled MHC class I molecules beyond the endoplasmic reticulum. *Journal of Molecular Biology* 266 (5), 993–1001.
- Kamath, A.T., Pooley, J., O’Keeffe, M.A., *et al.*, 2000. The development, maturation, and turnover rate of mouse spleen dendritic cell populations. *Journal of Immunology* 165 (12), 6762–6770.
- Kamphorst, A.O., Guermonprez, P., Dudziak, D., Nussenzweig, M.C., 2010. Route of antigen uptake differentially impacts presentation by dendritic cells and activated monocytes. *Journal of Immunology* (Baltimore, MD: 1950) 185 (6), 3426–3435.
- Kanaseki, T., Blanchard, N., Hammer, G.E., Gonzalez, F., Shastri, N., 2006. ERAAP synergizes with MHC class I molecules to make the final cut in the antigenic peptide precursors in the endoplasmic reticulum. *Immunity* 25 (5), 795–806.
- Kang, H.-K., Liu, M., Datta, S.K., 2007. Low-dose peptide tolerance therapy of lupus generates plasmacytoid dendritic cells that cause expansion of autoantigen-specific regulatory T cells and contraction of inflammatory Th17 cells. *Journal of Immunology* 178 (12), 7849–7858.
- Kang, R., Zeh, H.J., Lotze, M.T., Tang, D., 2011. The Beclin 1 network regulates autophagy and apoptosis. *Cell Death and Differentiation* 18 (4), 571–580.
- Kaplan, D.H., 2010. *In vivo* function of Langerhans cells and dermal dendritic cells. *Trends in Immunology* 31 (12), 446–451.
- Kaplan, D.H., Jenison, M.C., Saeland, S., Shlomchik, W.D., Shlomchik, M.J., 2005. Epidermal langerhans cell-deficient mice develop enhanced contact hypersensitivity. *Immunity* 23 (6), 611–620.
- Karsunky, H., Merad, M., Cozzio, A., Weissman, I.L., Manz, M.G., 2003. Flt3 ligand regulates dendritic cell development from Flt3+ lymphoid and myeloid-committed progenitors to Flt3+ dendritic cells *in vivo*. *Journal of Experimental Medicine* 198 (2), 305–313.
- Kastenmüller, K., Wille-Reece, U., Lindsay, R.W., *et al.*, 2011. Protective T cell immunity in mice following protein-TLR7/8 agonist-conjugate immunization requires aggregation, type I IFN, and multiple DC subsets. *Journal of Clinical Investigation* 121 (5), 1782–1796.
- Kawai, T., Akira, S., 2008. Toll-like receptor and RIG-I-like receptor signaling. *Annals of the New York Academy of Sciences* 1143, 1–20.
- Kawai, T., Akira, S., 2011. Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. *Immunity* 34 (5), 637–650.
- Kim, T.S., Braciale, T.J., 2009. Respiratory dendritic cell subsets differ in their capacity to support the induction of virus-specific cytotoxic CD8+ T cell responses. *PLoS One* 4, e4204.

- King, I.L., Kroenke, M.A., Segal, B.M., 2010. GM-CSF-dependent, CD103+ dermal dendritic cells play a critical role in Th effector cell differentiation after subcutaneous immunization. *Journal of Experimental Medicine* 207 (5), 953–961.
- Kisselev, A.F., Akopian, T.N., Woo, K.M., Goldberg, A.L., 1999. The sizes of peptides generated from protein by mammalian 26 and 20S proteasomes. Implications for understanding the degradative mechanism and antigen presentation. *Journal of Biological Chemistry* 274, 3363–3371.
- Kissenpfennig, A., Henri, S., Dubois, B., *et al.*, 2005. Dynamics and function of Langerhans cells *in vivo*: Dermal dendritic cells colonize lymph node areas distinct from slower migrating Langerhans cells. *Immunity* 22 (5), 643–654.
- Klechevsky, E., 2013. Human dendritic cells — Stars in the skin. *European Journal of Immunology* 43 (12), 3147–3155.
- Klechevsky, E., Morita, R., Liu, M., *et al.*, 2008. Functional specializations of human epidermal Langerhans cells and CD14+ dermal dendritic cells. *Immunity* 29 (3), 497–510.
- Kloetzel, P.M., Ossendorp, F., 2004. Proteasome and peptidase function in MHC-class-I-mediated antigen presentation. *Current Opinion in Immunology* 16, 76–81.
- Konermann, A., Deschner, J., Allam, J.P., *et al.*, 2012. Antigen-presenting cell marker expression and phagocytotic activity in periodontal ligament cells. *Journal of Oral Pathology & Medicine: Official Publication of the International Association of Oral Pathologists and the American Academy of Oral Pathology* 41 (4), 340–347.
- Kovacovics-Bankowski, M., Rock, K.L., 1995. A phagosome-to-cytosol pathway for exogenous antigens presented on MHC class I molecules. *Science (New York, NY)* 267, 243–246.
- Kraal, B.Y.G., Breel, M., Janse, M., Bruin, G., 1986. Langerhans' cells, veiled cells, and interdigitating cells in the mouse recognized by a monoclonal antibody. *Journal of Experimental Medicine* 163 (April), 981–997.
- Kreer, C., Rauen, J., Zehner, M., Burgdorf, S., 2012. Cross-presentation: How to get there — Or how to get the ER. *Frontiers in Immunology* 2, 87.
- Kroemer, G., Mariño, G., Levine, B., 2010. Autophagy and the integrated stress response. *Molecular Cell* 40 (2), 280–293.
- Kurts, C., Robinson, B.W.S., Knolle, P.A., 2010. Cross-priming in health and disease. *Nature Reviews Immunology* 10 (6), 403–414.
- Langlet, C., Tamoutounour, S., Henri, S., *et al.*, 2012. CD64 expression distinguishes monocyte-derived and conventional dendritic cells and reveals their distinct role during intramuscular immunization. *Journal of Immunology* 188 (4), 1751–1760.
- Lanier, L.L., 2005. NK cell recognition. *Annual Review of Immunology* 23, 225–274.
- Lecker, S.H., Goldberg, A.L., Mitch, W.E., 2006. Protein degradation by the ubiquitin-proteasome pathway in normal and disease states. *Journal of the American Society of Nephrology: JASN* 17 (11), 1807–1819.
- Lee, H.K., Mattei, L.M., Steinberg, B.E., *et al.*, 2010. *In Vivo* Requirement for Atg5 in antigen presentation by dendritic cells. *Immunity* 32 (2), 227–239.
- Li, P., Gregg, J.L., Wang, N., *et al.*, 2005. Compartmentalization of class II antigen presentation: Contribution of cytoplasmic and endosomal processing. *Immunological Reviews* 207, 206–217.
- Lindquist, R.L., Shakhar, G., Dudziak, D., *et al.*, 2004. Visualizing dendritic cell networks *in vivo*. *Nature Immunology* 5 (12), 1243–1250.
- Liu, K., Nussenzweig, M.C., 2010. Origin and development of dendritic cells. *Immunological Reviews* 234 (1), 45–54.
- Loschko, J., Heink, S., Hackl, D., *et al.*, 2011b. Antigen targeting to plasmacytoid dendritic cells via Siglec-H inhibits Th cell-dependent autoimmunity. *Journal of Immunology (Baltimore, MD: 1950)* 186 (12), 6346–6356.
- Loschko, J., Schlitzer, A., Dudziak, D., *et al.*, 2011a. Antigen delivery to plasmacytoid dendritic cells via BST2 induces protective T cell-mediated immunity. *Journal of Immunology (Baltimore, MD: 1950)* 186 (12), 6718–6725.
- Loureiro, J., Ploegh, H.L., 2006. Antigen presentation and the ubiquitin-proteasome system in host-pathogen interactions. *Advances in Immunology* 92 (06), 225–305.
- Lucchini-Hartz, M., van Endert, P.M., Lauvau, G., *et al.*, 2000. Cytotoxic T lymphocyte epitopes of HIV-1 Nef: Generation of multiple definitive major histocompatibility complex class I ligands by proteasomes. *Journal of Experimental Medicine* 191 (2), 239–252.
- van der Lugt, B., Khan, A.A., Hackney, J.A., *et al.*, 2014. Transcriptional programming of dendritic cells for enhanced MHC class II antigen presentation. *Nature Immunology* 15 (2), 161–167.
- Lundberg, K., Albrekt, A.S., Nelissen, I., *et al.*, 2013. Transcriptional profiling of human dendritic cell populations and models—unique profiles of *in vitro* dendritic cells and implications on functionality and applicability. *PLoS One* 8 (1), e52875.
- Luzio, J.P., Gray, S.R., Bright, N.A., 2010. Endosome-lysosome fusion. *Biochemical Society Transactions* 38, 1413–1416.
- MacDonald, K.P.A., Munster, D.J., Clark, G.J., *et al.*, 2002. Characterization of human blood dendritic cell subsets. *Blood* 100 (13), 4512–4520.
- Maldonado-López, R., De Smedt, T., Michel, P., *et al.*, 1999. CD8alpha+ and CD8alpha- subclasses of dendritic cells direct the development of distinct T helper cells *in vivo*. *Journal of Experimental Medicine* 189 (3), 587–592.
- Malissen, B., Tamoutounour, S., Henri, S., 2014. The origins and functions of dendritic cells and macrophages in the skin. *Nature Reviews Immunology* 14 (6), 417–428.
- Manoury, B., Hewitt, E.W., Morrice, N., *et al.*, 1998. An asparaginyl endopeptidase processes a microbial antigen for class II MHC presentation. *Nature* 396 (1997), 695–699.
- Maric, M., Arunachalam, B., Phan, U.T., *et al.*, 2001. Defective antigen processing in GILT-free mice. *Science (New York, NY)* 294 (5545), 1361–1365.
- Martin-Fontecha, A., Sebastiani, S., Höpken, U.E., *et al.*, 2003. Regulation of dendritic cell migration to the draining lymph node: Impact on T lymphocyte traffic and priming. *Journal of Experimental Medicine* 198 (4), 615–621.
- McGovern, N., Chan, J.K.Y., Ginhoux, F., 2014b. Dendritic cells in humans — From fetus to adult. *International Immunology* 27 (2), 65–72.
- McGovern, N., Schlitzer, A., Gunawan, M., *et al.*, 2014a. Human dermal CD14+ cells are a transient population of monocyte-derived macrophages. *Immunity* 41 (3), 465–477.
- McLellan, A.D., Kapp, M., Eggert, A., *et al.*, 2002. Anatomic location and T-cell stimulatory functions of mouse dendritic cell subsets defined by CD4 and CD8 expression. *Blood* 99 (6), 2084–2093.
- McIlroy, D., Troade, C., Grassi, F., *et al.*, 2001. Investigation of human spleen dendritic cell phenotype and distribution reveals evidence of *in vivo* activation in a subset of organ donors. *Blood* 97 (11), 3470–3477.
- Meissner, F., Mann, M., 2014. Quantitative shotgun proteomics: Considerations for a high-quality workflow in immunology. *Nature Immunology* 15 (2), 112–117.
- Mellman, I., Steinman, R.M., 2001. Dendritic cells: Specialized and regulated antigen processing machines. *Cell* 106 (3), 255–258.
- Merad, M., Ginhoux, F., Collin, M., 2008. Origin, homeostasis and function of Langerhans cells and other langerin-expressing dendritic cells. *Nature Reviews Immunology* 8 (12), 935–947.
- Merad, M., Manz, M., 2009. Dendritic cell homeostasis. *Blood* 113 (15), 3418–3427.
- Merad, M., Sathe, P., Helft, J., Miller, J., Mortha, A., 2013. The dendritic cell lineage: Ontogeny and function of dendritic cells and their subsets in the steady state and the inflamed setting. *Annual Review of Immunology* 31, 563–604.
- Meredith, M.M., Liu, K., Kamphorst, A.O., *et al.*, 2012. Zinc finger transcription factor ZDC is a negative regulator required to prevent activation of classical dendritic cells in the steady state. *Journal of Experimental Medicine* 209 (9), 1583–1593.
- Mezger, M.B., Pruneda, J.N., Klevit, R.E., Weissman, A.M., 2014. RING-type E3 ligases: Master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. *Biochimica et Biophysica Acta — Molecular Cell Research* 1843 (1), 47–60.
- Miller, J.C., Brown, B.D., Shay, T., *et al.*, 2012. Deciphering the transcriptional network of the dendritic cell lineage. *Nature Immunology* 13 (9), 888–899.
- Mintern, J.D., Macri, C., Villadangos, J.A., 2015. Modulation of antigen presentation by intracellular trafficking. *Current Opinion in Immunology* 34, 16–21.
- Mittag, D., Proietto, A.I., Loudovaris, T., *et al.*, 2011. Human dendritic cell subsets from spleen and blood are similar in phenotype and function but modified by donor health status. *Journal of Immunology (Baltimore, MD: 1950)* 186 (11), 6207–6217.
- Mizushima, N., Levine, B., Cuervo, A.M., Klionsky, D.J., 2008. Autophagy fights disease through cellular self-digestion. *Nature* 451 (7182), 1069–1075.
- Mizushima, N., Noda, T., Yoshimori, T., *et al.*, 1998. A protein conjugation system essential for autophagy. *Nature* 395 (6700), 395–398.
- Moffat, J.M., Mintern, J.D., Villadangos, J.A., 2013b. Control of MHC II antigen presentation by ubiquitination. *Current Opinion in Immunology* 25 (1), 109–114.
- Moffat, J.M., Segura, E., Khoury, G., *et al.*, 2013a. Targeting antigen to bone marrow stromal cell-2 expressed by conventional and plasmacytoid dendritic cells elicits efficient antigen presentation. *European Journal of Immunology* 43 (3), 595–605.
- Mohamed, M.M., Sloane, B.F., 2006. Cysteine cathepsins: Multifunctional enzymes in cancer. *Nature Reviews Cancer* 6, 764–775.
- Mollah, S.A., Dobrin, J.S., Feder, R.E., *et al.*, 2014. Flt3L dependence helps define an uncharacterized subset of murine cutaneous dendritic cells. *Journal of Investigative Dermatology* 134 (5), 1265–1275.

- van Montfoort, N., Camps, M.G., Khan, S., *et al.*, 2009. Antigen storage compartments in mature dendritic cells facilitate prolonged cytotoxic T lymphocyte cross-priming capacity. *Proceedings of the National Academy of Sciences of the United States of America* 106 (16), 6730–6735.
- Morris, S., Swanson, M.S., Lieberman, A., *et al.*, 2011. Autophagy-mediated dendritic cell activation is essential for innate cytokine production and APC function with respiratory syncytial virus responses. *Journal of Immunology* (Baltimore, MD: 1950) 187, 3953–3961.
- Münz, C., 2009. Enhancing immunity through autophagy. *Annual Review of Immunology* 27, 423–449.
- Nagao, K., Ginhoux, F., Leitner, W.W., *et al.*, 2009. Murine epidermal Langerhans cells and langerin-expressing dermal dendritic cells are unrelated and exhibit distinct functions. *Proceedings of the National Academy of Sciences of the United States of America* 106 (9), 3312–3317.
- Nakagawa, T.Y., Brissette, W.H., Lira, P.D., *et al.*, 1999. Impaired invariant chain degradation and antigen presentation and diminished collagen-induced arthritis in cathepsin S null mice. *Immunity* 10, 207–217.
- Nestle, F.O., Turka, L.A., Nickoloff, B.J., 1994. Characterization of dermal dendritic cells in psoriasis. Autostimulation of T lymphocytes and induction of Th1 type cytokines. *Journal of Clinical Investigation* 94 (1), 202–209.
- Neubert, K., Lehmann, C.H., Heger, L., *et al.*, 2014. Antigen delivery to CD11c + CD8- dendritic cells induces protective immune responses against experimental melanoma in mice *In Vivo*. *Journal of Immunology* 192 (12), 5830–5838.
- Newcomb, J.R., Carboy-Newcomb, C., Cresswell, P., 1996. Trimeric interactions of the invariant chain and its association with major histocompatibility complex class II alpha beta dimers. *Journal of Biological Chemistry* 271 (39), 24249–24256.
- Newell, E.W., Sigal, N., Nair, N., *et al.*, 2013. Combinatorial tetramer staining and mass cytometry analysis facilitate T-cell epitope mapping and characterization. *Nature Biotechnology* 31 (7), 623–629.
- Nierkens, S., Tel, J., Janssen, E., Adema, G.J., 2013. Antigen cross-presentation by dendritic cell subsets: One general or all sergeants? *Trends in Immunology*, 1–10.
- Nimmerjahn, F., Milosevic, S., Behrends, U., *et al.*, 2003. Major histocompatibility complex class II-restricted presentation of a cytosolic antigen by autophagy. *European Journal of Immunology* 33 (5), 1250–1259.
- Nimmerjahn, F., Ravetch, J.V., 2006. Fcgamma receptors: Old friends and new family members. *Immunity* 24 (1), 19–28.
- Nimmerjahn, F., Ravetch, J.V., 2008. Fcgamma receptors as regulators of immune responses. *Nature Reviews. Immunology* 8 (1), 34–47.
- Nizzoli, G., Krietsch, J., Weick, A., *et al.*, 2013. Human CD1c + dendritic cells secrete high levels of IL-12 and potently prime cytotoxic T-cell responses. *Blood* 122 (6), 932–942.
- Ochando, J.C., Homma, C., Yang, Y., *et al.*, 2006. Alloantigen-presenting plasmacytoid dendritic cells mediate tolerance to vascularized grafts. *Nature Immunology* 7 (6), 652–662.
- O'Donnell, E.A., Ernst, D.N., Hingorani, R., 2013. Multiparameter flow cytometry: Advances in high resolution analysis. *Immune Network* 13 (2), 43–54.
- Paessens, L., Fluitsma, D., van Kooyk, Y., 2008. Haematopoietic antigen-presenting cells in the human thymic cortex: Evidence for a role in selection and removal of apoptotic thymocytes. *Journal of Pathology* 214 (1), 96–103.
- Pandey, G., Cohain, A., Miller, J., Merad, M., 2013. Decoding dendritic cell function through module and network analysis. *Journal of Immunological Methods* 387 (1–2), 71–80.
- Perfetto, S.P., Chattopadhyay, P.K., Roederer, M., 2004. Seventeen-colour flow cytometry: Unravelling the immune system. *Nature Reviews. Immunology* 4 (8), 648–655.
- Phan, U.T., Arunachalam, B., Cresswell, P., 2000. Gamma-interferon-inducible lysosomal thiol reductase (GILT). Maturation, activity, and mechanism of action. *Journal of Biological Chemistry* 275 (34), 25907–25914.
- Piatkevich, K.D., Verkhusa, V.V., 2011. Guide to red fluorescent proteins and biosensors for flow cytometry. *Methods in Cell Biology* 102, 431–461.
- Pierre, P., Mellman, I., 1998. Developmental regulation of invariant chain proteolysis controls MHC class II trafficking in mouse dendritic cells. *Cell* 93, 1135–1145.
- Poluektov, Y.O., Kim, A., Hartman, I.Z., Sadegh-Nasseri, S., 2013. HLA-DO as the optimizer of epitope selection for MHC class II antigen presentation. *PLoS One* 8, e71228.
- Pooley, J.L., Heath, W.R., Shortman, K., 2001. Cutting edge: Intravenous soluble antigen is presented to CD4 T cells by CD8- dendritic cells, but cross-presented to CD8 T cells by CD8+ dendritic cells. *Journal of Immunology* (Baltimore, MD: 1950) 166 (9), 5327–5330.
- Poulin, L.F., Henri, S., de Bovis, B., *et al.*, 2007. The dermis contains langerin + dendritic cells that develop and function independently of epidermal Langerhans cells. *Journal of Experimental Medicine* 204 (13), 3119–3131.
- Poulin, L.F., Salio, M., Griessinger, E., *et al.*, 2010. Characterization of human DNGR-1 + BDCA3 + leukocytes as putative equivalents of mouse CD8 {alpha} + dendritic cells. *Journal of Experimental Medicine* 207 (6), 1261–1271.
- Pulendran, B., Smith, J.L., Caspary, G., *et al.*, 1999. Distinct dendritic cell subsets differentially regulate the class of immune response *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America* 96 (3), 1036–1041.
- Pulendran, B., Tang, H., Denning, T.L., 2008. Division of labor, plasticity, and crosstalk between dendritic cell subsets. *Current Opinion in Immunology* 20 (1), 61–67.
- Qiu, C.-H., Miyake, Y., Kaise, H., *et al.*, 2009. Novel subset of CD8{alpha} + dendritic cells localized in the marginal zone is responsible for tolerance to cell-associated antigens. *Journal of Immunology* (Baltimore, MD: 1950) 182 (7), 4127–4136.
- Ravikumar, B., Sarkar, S., Davies, J.E., *et al.*, 2010. Regulation of mammalian autophagy in physiology and pathophysiology. *Physiological Reviews* 90, 1383–1435.
- Reizis, B., Bunin, A., Ghosh, H.S., Lewis, K.L., Sisirak, V., 2011. Plasmacytoid dendritic cells: Recent progress and open questions. *Annual Review of Immunology* 29, 163–183.
- Riese, R.J., Chapman, H.A., 2000. Cathepsins and compartmentalization in antigen presentation. *Current Opinion in Immunology* 12 (1), 107–113.
- Riese, R.J., Wolf, P.R., Brömme, D., *et al.*, 1996. Essential role for cathepsin S in MHC class II-associated invariant chain processing and peptide loading. *Immunity* 4 (4), 357–366.
- Robbins, S.H., Walzer, T., Dembélé, D., *et al.*, 2008. Novel insights into the relationships between dendritic cell subsets in human and mouse revealed by genome-wide expression profiling. *Genome Biology* 9 (1), R17.
- Roche, P.A., Furuta, K., 2015. The ins and outs of MHC class II-mediated antigen processing and presentation. *Nature Reviews Immunology* 15, 203–216.
- Rock, K.L., Farfán-Arribas, D.J., Colbert, J.D., Goldberg, A.L., 2014. Re-examining class-I presentation and the DRiP hypothesis. *Trends in Immunology* 35 (4), 144–152.
- Rock, K.L., Goldberg, A.L., 1999. Degradation of cell proteins and the generation of MHC class I-presented peptides. *Annual Review of Immunology* 17, 739–779.
- Rock, K.L., Shen, L., 2005. Cross-presentation: Underlying mechanisms and role in immune surveillance. *Immunological Reviews* 207, 166–183.
- Rock, K.L., York, I.A., Goldberg, A.L., 2004. Post-proteasomal antigen processing for major histocompatibility complex class I presentation. *Nature Immunology* 5 (7), 670–677.
- Rodríguez, A., Regnault, A., Kleijmeer, M., *et al.*, 1999. Selective transport of internalized antigens to the cytosol for MHC class I presentation in dendritic cells. *Nature Cell Biology* 1 (6), 362–368.
- Romani, N., Clausen, B.E., Stoitzner, P., 2010b. Langerhans cells and more: Langerin-expressing dendritic cell subsets in the skin. *Immunological Reviews* 234 (1), 120–141.
- Romani, N., Thurnher, M., Idoyaga, J., Steinman, R.M., Flacher, V., 2010a. Targeting of antigens to skin dendritic cells: Possibilities to enhance vaccine efficacy. *Immunology and Cell Biology* 234, 1–7.
- Rönnblom, L., 2010. Potential role of IFN α in adult lupus. *Arthritis Research & Therapy* 12 (Suppl. 1), S3.
- Sahu, R., Kaushik, S., Clement, C.C., *et al.*, 2011. Microautophagy of cytosolic proteins by late endosomes. *Developmental Cell* 20 (1), 131–139.
- Sallusto, F., Lanzavecchia, A., 1994. Efficient presentation of soluble antigen by cultured human dendritic cells is maintained by granulocyte/macrophage colony-stimulating factor plus interleukin 4 and downregulated by tumor necrosis factor α . *Journal of Experimental Medicine* 179 (1), 1109–1118.
- Sancho, D., Joffre, O.P., Keller, A.M., *et al.*, 2009. Identification of a dendritic cell receptor that couples sensing of necrosis to immunity. *Nature* 458 (7240), 899–903.
- Sapozhnikov, A., Fischer, J.A., Zaft, T., *et al.*, 2007. Organ-dependent *in vivo* priming of naive CD4 +, but not CD8 +, T cells by plasmacytoid dendritic cells. *Journal of Experimental Medicine* 204 (8), 1923–1933.
- Saric, T., Beninga, J., Graef, C.I., *et al.*, 2001. Major histocompatibility complex class I-presented antigenic peptides are degraded in cytosolic extracts primarily by thimet oligopeptidase. *Journal of Biological Chemistry* 276 (39), 36474–36481.
- Satpathy, A.T., Murphy, K.M., Wumesh, K.C., 2011. Transcription factor networks in dendritic cell development. *Seminars in Immunology* 23 (5), 388–397.
- Satpathy, A.T., Wu, X., Albring, J.C., *et al.*, 2012. Re(De)fining the dendritic cell lineage. *Nature Immunology* 13 (12), 1145–1154.

- Saveanu, L., Carroll, O., Weimershaus, M., *et al.*, 2009. IRAP identifies an endosomal compartment required for MHC class I cross-presentation. *Science* (New York, NY) 325 (5937), 213–217.
- Savina, A., Amigorena, S., 2007. Phagocytosis and antigen presentation in dendritic cells. *Immunological Reviews* 219, 143–156.
- Savina, A., Peres, A., Cebrian, I., *et al.*, 2009. The small GTPase Rac2 controls phagosomal alkalization and antigen crosspresentation selectively in CD8 + dendritic cells. *Immunity* 30 (4), 544–555.
- Sawai, C.M., Sisirak, V., Ghosh, H.S., *et al.*, 2013. Transcription factor Runx2 controls the development and migration of plasmacytoid dendritic cells. *Journal of Experimental Medicine* 210 (11), 2151–2159.
- Schäkel, K., Kannagi, R., Knip, B., *et al.*, 2002. 6-Sulfo LacNAc, a novel carbohydrate modification of PSGL-1, defines an inflammatory type of human dendritic cells. *Immunity* 17 (3), 289–301.
- Schatz, M.M., Peters, B., Akkad, N., *et al.*, 2008. Characterizing the N-terminal processing motif of MHC class I ligands. *Journal of Immunology* (Baltimore, MD: 1950) 180, 3210–3217.
- Schlitzer, A., Ginhoux, F., 2014. Organization of the mouse and human DC network. *Current Opinion in Immunology* 26 (1), 90–99.
- Schlitzer, A., McGovern, N., Teo, P., *et al.*, 2013. IRF4 transcription factor-dependent CD11b + dendritic cells in human and mouse control mucosal IL-17 cytokine responses. *Immunity* 38 (5), 970–983.
- Schmid, D., Pypaert, M., Münz, C., 2007. Antigen-loading compartments for major histocompatibility complex class II molecules continuously receive input from autophagosomes. *Immunity* 26 (1), 79–92.
- Schnorrer, P., Behrens, G.M., Wilson, N.S., *et al.*, 2006. The dominant role of CD8 + dendritic cells in cross-presentation is not dictated by antigen capture. *Proceedings of the National Academy of Sciences of the United States of America* 103 (28), 10729–10734.
- Schraml, B.U., Reis e Sousa, C., 2015. Defining dendritic cells. *Current Opinion in Immunology* 32, 13–20.
- Schubert, U., Antón, L.C., Gibbs, J., *et al.*, 2000. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature* 404, 770–774.
- Schuler, G., Steinman, R.M., 1985. Murine epidermal Langerhans cells mature into potent immunostimulatory dendritic cells *in vitro*. *Journal of Experimental Medicine* 161, 526–546.
- Schulz, O., Edwards, A.D., Schito, M., *et al.*, 2000. CD40 triggering of heterodimeric IL-12 p70 production by dendritic cells *in vivo* requires a microbial priming signal. *Immunity* 13 (4), 453–462.
- Schulz, O., Jaensson, E., Persson, E.K., *et al.*, 2009. Intestinal CD103 +, but not CX3CR1 +, antigen sampling cells migrate in lymph and serve classical dendritic cell functions. *Journal of Experimental Medicine* 206 (13), 3101–3114.
- Schulz, O., Reis E Sousa, C., 2002. Cross-presentation of cell-associated antigens by CD8?? + dendritic cells is attributable to their ability to internalize dead cells. *Immunology* 107 (2), 183–189.
- Schumacher, T.N., De Bruijn, M.L., Vernie, L.N., *et al.*, 1991. Peptide selection by MHC class I molecules. *Nature* 350, 703–706.
- Schwartz, R.H., 2003. T cell anergy. *Annual Review of Immunology* 21 (1), 305–334.
- Segura, E., Amigorena, S., 2014. Cross-presentation by human dendritic cell subsets. *Immunology Letters* 158 (1–2), 73–78.
- Segura, E., Durand, M., Amigorena, S., 2013. Similar antigen cross-presentation capacity and phagocytic functions in all freshly isolated human lymphoid organ-resident dendritic cells. *Journal of Experimental Medicine* 210 (5), 1035–1047.
- Segura, E., Valladeau-Guilemond, J., Donnadieu, M.H., *et al.*, 2012. Characterization of resident and migratory dendritic cells in human lymph nodes. *Journal of Experimental Medicine* 209 (4), 653–660.
- Segura, E., Villadangos, J.A., 2011. A modular and combinatorial view of the antigen cross-presentation pathway in dendritic cells. *Traffic* (Copenhagen, Denmark) 12 (12), 1677–1685.
- Severa, M., Giacomini, E., Gafa, V., *et al.*, 2013. EBV stimulates TLR- and autophagy-dependent pathways and impairs maturation in plasmacytoid dendritic cells: Implications for viral immune escape. *European Journal of Immunology* 43, 147–158.
- Shen, L., Rock, K.L., 2004. Cellular protein is the source of cross-priming antigen *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America* 101 (9), 3035–3040.
- Shortman, K., Liu, Y.-J., 2002. Mouse and human dendritic cell subtypes. *Nature Reviews Immunology* 2 (3), 151–161.
- Shortman, K., Naik, S.H., 2007. Steady-state and inflammatory dendritic-cell development. *Nature Reviews Immunology* 7 (1), 19–30.
- Silk, J.D., Salio, M., Brown, J., Jones, E.Y., Cerundolo, V., 2008. Structural and functional aspects of lipid binding by CD1 molecules. *Annual Review of Cell and Developmental Biology* 24, 369–395.
- Singh, R., Cresswell, P., 2010. Defective cross-presentation of viral antigens in GILT-free mice. *Science* (New York, NY) 328 (5984), 1394–1398.
- Skokos, D., Nussenzweig, M.C., 2007. CD8- DCs induce IL-12-independent Th1 differentiation through Delta 4 Notch-like ligand in response to bacterial LPS. *Journal of Experimental Medicine* 204 (7), 1525–1531.
- Smith, M.H., Ploegh, H.L., Weissman, J.S., 2011. Road to ruin: Targeting proteins for degradation in the endoplasmic reticulum. *Science* 334 (6059), 1086–1090.
- Soares, H., Waechter, H., Glaichenhaus, N., *et al.*, 2007. A subset of dendritic cells induces CD4 + T cells to produce IFN-gamma by an IL-12-independent but CD70-dependent mechanism *in vivo*. *Journal of Experimental Medicine* 204 (5), 1095–1106.
- Song, R., Harding, C.V., 1996. Roles of proteasomes, transporter for antigen presentation (TAP), and beta 2-microglobulin in the processing of bacterial or particulate antigens via an alternate class I MHC processing pathway. *Journal of Immunology* (Baltimore, MD: 1950) 156 (34), 4182–4190.
- Steinman, R.M., 2012. Decisions about dendritic cells: Past, present, and future. *Annual Review of Immunology* 30, 1–22.
- Steinman, R.M., Cohn, Z.A., 1973. Identification of a novel cell type in peripheral lymphoid organs of mice I. Morphology, quantitation, tissue distribution. *Journal of Experimental Medicine* 137, 1142–1162.
- Steinman, R.M., Hawiger, D., Nussenzweig, M.C., 2003. Tolerogenic dendritic cells. *Annual Review of Immunology* 21, 685–711.
- Steinman, R.M., Nussenzweig, M.C., 1980. Dendritic cells: Features and functions. *Immunological Reviews* 53, 127–147.
- Steinman, R.M., Nussenzweig, M.C., 2002. Avoiding horror autotoxicus: The importance of dendritic cells in peripheral T cell tolerance. *Proceedings of the National Academy of Sciences of the United States of America* 99 (1), 351–358.
- Stoitzner, P., Holzmann, S., McLellan, A.D., *et al.*, 2003. Visualization and characterization of migratory Langerhans cells in murine skin and lymph nodes by antibodies against Langerin/CD207. *Journal of Investigative Dermatology* 120 (2), 266–274.
- Suzuki, S., Honma, K., Matsuyama, T., *et al.*, 2004. Critical roles of interferon regulatory factor 4 in CD11bhighCD8alpha- dendritic cell development. *Proceedings of the National Academy of Sciences of the United States of America* 101 (24), 8981–8986.
- Swiecki, M., Colonna, M., 2010. Unraveling the functions of plasmacytoid dendritic cells during viral infections, autoimmunity, and tolerance. *Immunological Reviews* 234 (1), 142–162.
- Tamoutounour, S., Guillems, M., Montanana Sanchis, F., *et al.*, 2013. Origins and functional specialization of macrophages and of conventional and monocyte-derived dendritic cells in mouse skin. *Immunity* 39 (5), 925–938.
- Thaiss, C.A., Semmling, V., Franken, L., Wagner, H., Kurts, C., 2011. Chemokines: A new dendritic cell signal for cell activation. *Frontiers in Immunology* 2, 1–9.
- Tosello, V., Zamarchi, R., Merlo, A., *et al.*, 2009. Differential expression of constitutive and inducible proteasome subunits in human monocyte-derived DC differentiated in the presence of IFN-alpha or IL-4. *European Journal of Immunology* 39 (1), 56–66.
- Trombetta, S., Mellman, I., 2005. Cell biology of antigen processing *in vitro* and *in vivo*. *Annual Review of Immunology* 23, 975–1028.
- Tsien, R.Y., 1998. The green fluorescent protein. *Annual Review of Biochemistry* 67, 509–544.
- Tu, L., Fang, T.C., Artis, D., *et al.*, 2005. Notch signaling is an important regulator of type 2 immunity. *Journal of Experimental Medicine* 202 (8), 1037–1042.
- Vega-Ramos, J., Roquilly, A., Asehnoune, K., Villadangos, J.A., 2014. Modulation of dendritic cell antigen presentation by pathogens, tissue damage and secondary inflammatory signals. *Current Opinion in Pharmacology* 17, 64–70.
- Villadangos, J.A., 2001. Presentation of antigens by MHC class II molecules: Getting the most out of them. *Molecular Immunology* 38, 329–346.
- Villadangos, J.A., Heath, W.R., Carbone, F.R., 2007. Outside looking in: The inner workings of the crosspresentation pathway within dendritic cells. *Trends in Immunology* 28 (2), 45–47.
- Villadangos, J.A., Schnorrer, P., 2007. Intrinsic and cooperative antigen-presenting functions of dendritic-cell subsets *in vivo*. *Nature Reviews Immunology* 7 (7), 543–555.
- Villadangos, J.A., Young, L.J., 2008. Antigen-presentation properties of plasmacytoid dendritic cells. *Immunity* 29 (3), 352–361.
- Voges, D., Zwickl, P., Baumeister, W., 1999. The 26S proteasome: A molecular machine designed for controlled proteolysis. *Annual Review of Biochemistry* 68, 1015–1068.
- Vremec, D., Pooley, J., Hochrein, H., Wu, L., Shortman, K., 2000. CD4 and CD8 expression by dendritic cell subtypes in mouse thymus and spleen. *Journal of Immunology* (Baltimore, MD: 1950) 164 (6), 2978–2986.

- Vremec, D., Zorbas, M., Scollay, R., *et al.*, 1992. The surface phenotype of dendritic cells purified from mouse thymus and spleen: Investigation of the CD8 expression by a subpopulation of dendritic cells. *Journal of Experimental Medicine* 176 (1), 47–58.
- Waskow, C., Liu, K., Darrasse-Jèze, G., *et al.*, 2008. The receptor tyrosine kinase Flt3 is required for dendritic cell development in peripheral lymphoid tissues. *Nature Immunology* 9 (6), 676–683.
- Watchmaker, P.B., Lahl, K., Lee, M., *et al.*, 2014. Comparative transcriptional and functional profiling defines conserved programs of intestinal DC differentiation in humans and mice. *Nature Immunology* 15 (1), 98–108.
- Watts, C., Matthews, S.P., Mazzeo, D., Manoury, B., Moss, C.X., 2005. Asparaginyl endopeptidase: Case history of a class II MHC compartment protease. *Immunological Reviews* 207, 218–228.
- Wearsch, P.A., Cresswell, P., 2008. The quality control of MHC class I peptide loading. *Current Opinion in Cell Biology* 20, 624–631.
- Weimershaus, M., Evnouchidou, I., Saveanu, L., van Endert, P., 2013. Peptidases trimming MHC class I ligands. *Current Opinion in Immunology* 25 (1), 90–96.
- Weimershaus, M., Maschalidi, S., Sepulveda, F., *et al.*, 2012. Conventional dendritic cells require IRAP-Rab14 endosomes for efficient cross-presentation. *Journal of Immunology* (Baltimore, MD: 1950) 188, 1840–1846.
- Wendland, M., Czeloth, N., Mach, N., *et al.*, 2007. CCR9 is a homing receptor for plasmacytoid dendritic cells to the small intestine. *Proceedings of the National Academy of Sciences of the United States of America* 104 (15), 6347–6352.
- Xue, J., Schmidt, S.V., Sander, J., *et al.*, 2014. Transcriptome-Based Network Analysis Reveals a Spectrum Model of Human Macrophage Activation. *Immunity* 40 (2), 274–288.
- Yamazaki, S., Dudziak, D., Heidkamp, G.F., *et al.*, 2008. CD8 + CD205 + splenic dendritic cells are specialized to induce Foxp3 + regulatory T cells. *Journal of Immunology* 181 (10), 6923–6933.
- Yasuda, Y., Kaleta, J., Brömme, D., 2005. The role of cathepsins in osteoporosis and arthritis: Rationale for the design of new therapeutics. *Advanced Drug Delivery Reviews* 57, 973–993.
- Yewdell, J.W., Antón, L.C., Bennink, J.R., 1996. Defective ribosomal products (DRiPs): A major source of antigenic peptides for MHC class I molecules? *Journal of Immunology* (Baltimore, MD: 1950) 157, 1823–1826.
- Yoneyama, M., Fujita, T., 2009. RNA recognition and signal transduction by RIG-I-like receptors. *Immunological Reviews* 227 (1), 54–65.
- York, I.A., *et al.*, 2003. The cytosolic endopeptidase, thimet oligopeptidase, destroys antigenic peptides and limits the extent of MHC class I antigen presentation. *Immunity* 18, 429–440.
- Young, L.J., Wilson, N.S., Schnorrer, P., *et al.*, 2008. Differential MHC class II synthesis and ubiquitination confers distinct antigen-presenting properties on conventional and plasmacytoid dendritic cells. *Nature Immunology* 9 (11), 1244–1252.
- Yu, C.I., Becker, C., Metang, P., *et al.*, 2014. Human CD141+ Dendritic Cells Induce CD4+ T Cells To Produce Type 2 Cytokines. *Journal of Immunology* 193 (12), 6210.
- Zhang, J., Raper, A., Sugita, N., *et al.*, 2006a. Characterization of siglec-H as a novel endocytic receptor expressed on murine plasmacytoid dendritic cell precursors. *Blood* 107 (9), 3600–3608.
- Zhang, Y., Baig, E., Williams, D.B., 2006b. Functions of ERp57 in the folding and assembly of major histocompatibility complex class I molecules. *Journal of Biological Chemistry* 281 (21), 14622–14631.
- Ziegler-Heitbrock, L., Ancuta, P., Crowe, S., *et al.*, 2010. Nomenclature of monocytes and dendritic cells in blood. *Blood* 116 (16), e74–e80.

CD28 Costimulation and Regulatory T Cells

B Soskic and DM Sansom, UCL Institute of Immunity and Transplantation, London, UK

© 2016 Elsevier Inc. All rights reserved.

Introduction

An important feature of a mammalian organism is its ability to protect itself against infectious agents such as viruses and bacteria present in the environment. Over millions of years of evolution, a highly specialized system consisting of different cells and molecules that enable immunity against harmful organisms has developed. Simultaneously, this requires mechanisms in order to prevent an immune response developing against host tissue. Innate immune cells like macrophages, are equipped with receptors which can recognize molecules that are not normally present in the mammalian organism (e.g., lipopolysaccharide (LPS) in bacterial walls), minimizing inappropriate activation. The limitation of this response is that it remains unaltered even in the face of repeated infections, responding in the same manner toward any microorganism that contains a particular recognized pattern. On the other hand, cells of the adaptive immune system such as T cells provide specific immunity, which is capable of long-term 'memory' and more vigorous and rapid responses in the case of reinfection. However, T cells recognize peptide fragments derived from proteins, which have a large, but nonetheless limited, range of possible amino acid configurations. Thus, although the nature of an antigen that T cell recognize enables a vigorous response to specific microorganisms, it also provides a constant threat to host tissues that may contain the same or similar amino acid sequences as part of their normal physiology. Therefore, both T cell development and T cell activation are tightly controlled in order to prevent autoimmune diseases, which can result from such recognition.

How Is the Self-Tolerance Established?

The discrimination between host peptides and those derived from the pathogen proteins is not straightforward and presents a major challenge to adaptive immunity. To protect the host tissue from harmful T cell responses, T cells that recognize self-antigens with a high affinity are not released into the immune system. Instead, such cells are deleted during development in the thymus. This clonal deletion of highly self-reactive T cells represents the basis of central tolerance, steering the immune response away from self-antigens. However, although such developmental controls are important for establishing immune tolerance, they are not sufficient and T cells that survive this 'negative selection' are not completely purged of self-reactivity. In addition, not all host proteins in the periphery are expressed in thymus and T cells therefore emerge that can still recognize self-antigens albeit with the lower affinity. Finally, a large number of new harmless proteins are constantly introduced into our system (e.g., through food). In order to prevent unwanted T cell responses to such proteins a set of mechanisms has evolved to control T cell activation.

Central Tolerance

Using their highly variable T cell receptor (TCR), T cells recognize antigens in the context of major histocompatibility complex (MHC). Given that the number of potential antigens far exceeds the number of TCR genes in the human genome, the main problem is how to generate a TCR repertoire broad enough to recognize all conceivable antigens present in our environment (Sewell, 2012). Evolution has solved this problem by introducing random gene rearrangement of TCR genes during T cell development, which drives enormous variability. However, the random nature of the process introduces two major problems. To begin with, the majority of T cells that express randomly rearranged TCRs cannot recognize MHC molecules and are deleted in the thymus with only T cells that can bind to peptide-MHC complex being positively selected. Interestingly, a wide range of tissue-restricted proteins (e.g., insulin) are also processed and presented by thymic antigen-presenting cells (APCs) and T cells that recognize these antigens with high affinity also undergo cell death (negative selection). As such negative selection eliminates the most dangerous T cell clones preventing imminent tissue destruction. Whilst the detailed mechanisms of negative selection are still under investigation it is believed that strong TCR signals generated upon antigen recognition induce pro-apoptotic proteins that mediate cell suicide (Hogquist *et al.*, 2005; Kyewski and Derbinski, 2004).

Negative selection against ectopically expressed self-antigens undoubtedly has a major role in establishing tolerance and mutation in the transcription factor AIRE, which is responsible for driving the expression of tissue-restricted antigens in thymic epithelial cells, leads to a systemic autoimmunity (autoimmune polyendocrine syndrome) (Mathis and Benoist, 2009). However, it is equally clear that despite rigorous selection in the thymus, the peripheral T cell pool still contains T cells capable of reacting to our own tissues, which require additional specific mechanisms of control in a process generically termed peripheral tolerance.

Peripheral Tolerance

Deleting every T cell that can recognize a self-antigen would result in an extremely limited repertoire and most likely would not be compatible with sufficient diversity of pathogen recognition. Therefore, a set of additional mechanisms has been established in the periphery in order to allow the existence of weakly self-reactive T cells that are potentially useful against pathogens (Walker and Abbas, 2002). In the last decade, a variety of molecules implicated in the peripheral regulation of T cell responses have been described. For instance, defects in FoxP3 (a master transcription factor in regulatory T cells), CTLA-4 (a negative regulator of CD28 pathway), TGF- β (T cell growth factor), and IL-10 and IL-2 (anti-inflammatory cytokines) all lead to a severe autoimmunity. Below we discuss in

detail the CD28/CTLA-4 pathway as an essential component of peripheral T cell tolerance.

Responses to pathogens versus responses to self – the role of CD28

Since the T cell repertoire inevitably contains self-reactive T cells, the question arises: how are responses to self-proteins discriminated from the responses to foreign antigens? It is currently thought that activation of CD4⁺ T cells needs at least two signals provided by specialized APCs, such as dendritic cells (DCs), macrophages, and B cells (Figure 1). The first signal is generated when TCR recognizes the antigen displayed by MHC II complex. The requirement for the second signal, known as costimulation, ensures that T cell activation does not occur in the absence of additional information. This is due to the fact that recognition of pathogens (e.g., via Toll-like receptors) or the presence of inflammation specifically upregulates the costimulatory ligands CD80 (B7-1) and CD86 (B7-2) which bind to CD28. Via their interaction with costimulatory receptor CD28, on T cells, these ligands provide signals necessary for a successful T cell activation, proliferation, and differentiation. In this way costimulatory signals are only triggered by the presence of inflammation or 'danger' associated with infection, limiting activation of potentially self-reactive T cells to such settings.

Studies using TCR-transgenic mice that respond to myelin basic protein (MBP), demonstrated that mice housed in standard animal facilities develop experimental autoimmune encephalomyelitis (EAE). In contrast, the same TCR-transgenic mice housed in pathogen-free facilities fail to develop the same disease (Goverman *et al.*, 1993). A simple interpretation of these results is that microbes in the environment are

important in driving the maturation of APCs, enhancing their ability to costimulate. Although these groups of mice have the same TCR specificity, the outcome of immune reaction is dependent on the CD28 engagement. These data along with numerous other studies have firmly established that, CD28 engagement presents a key immune checkpoint that influences self-tolerance: initiating T cell responses not based on the source of peptide antigen, but based on sensing the inflammatory context of recognition. Naturally, regulation of the CD28 pathway then serves as critical point in prevention of autoimmunity.

Regulatory T cells – a specialized cell type that regulate immune responses

Regulatory T cells (Tregs) are a subset of CD4⁺ T cells whose role is to regulate immune responses especially those against self-tissues. This is revealed by diseases such as IPEX (immune dysregulation, polyendocrinopathy, enteropathy, X linked) in humans and the Scurfy phenotype in mice due to loss of Treg function (Bennett *et al.*, 2001; Brunkow *et al.*, 2001). Tregs are characterized by high expression of IL-2 receptor α (CD25) and low level of IL-7 receptor α (CD127). Tregs are generated in thymus upon recognition of self-antigen (called thymus-derived or natural Tregs), or in the periphery upon activation of naïve T cells under tolerogenic conditions (peripherally derived or induced Tregs). The similarities and differences between nTregs and iTregs have been extensively studied and they are reviewed elsewhere (Curotto de Lafaille and Lafaille, 2009; Bilate and Lafaille, 2012).

After discovery of the suppressive activity of Tregs, a major hurdle was how to distinguish them from activated T cells due to the fact that they express high level of CD25, which is

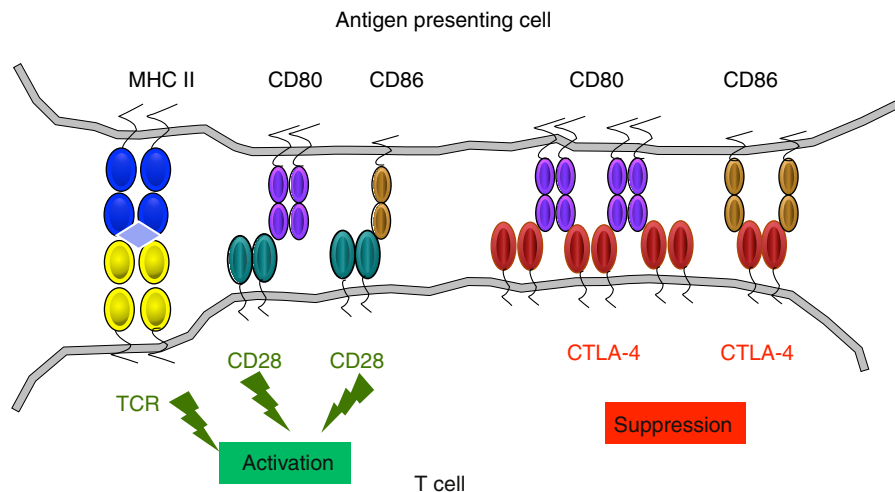


Figure 1 CD28 costimulation controls T cell activation. The goal of T cell activation is to generate a large number of antigen-specific T cells with effector properties that can eliminate pathogens. It is currently thought that activation of CD4⁺ T cells needs at least two signals provided by a specialized APC that captures, processes, and presents antigens as a part of MHC. Simultaneously, recognition of pathogen-related molecules or the presence of inflammation specifically upregulates the costimulatory ligands CD80 and CD86. The first signal is generated when TCR recognizes the antigen displayed by MHC complex. The requirement for the second signal, known as costimulation, ensures that T cell activation does not occur in the absence of additional inflammatory or 'danger' information. Via interaction with CD28, on T cells, CD80 and CD86 provide translate such danger signals to T cells and are therefore necessary for a successful T cell activation, proliferation, and differentiation. The second receptor for the CD80 and CD86 is cytotoxic T lymphocyte antigen 4 (CTLA-4). Expression of CTLA-4 is limited to activated conventional T cells and regulatory T cells (Tregs). CTLA-4 has higher affinity and avidity for both ligands, and serves as a negative regulator of T cell responses.

also upregulated upon the activation of conventional T cells (Tconv). The study of Tregs was therefore greatly facilitated by the discovery of transcription factor FoxP3, which appears to be specific to this lineage (Fontenot *et al.*, 2005). It has been demonstrated that this member of forkhead family of transcription factors is critical for Treg development (Hori *et al.*, 2003; Fontenot *et al.*, 2003). Loss of FoxP3 function causes fatal systemic autoimmune phenotype in both mice and humans characterized by uncontrolled proliferation of T cells, which leads to multiorgan infiltration and failure. Taken together, these studies indicate that Tregs are indispensable for immune homeostasis.

Since the importance of Tregs has been discovered, much work has been performed to elucidate mechanisms by which Tregs regulate immune response. In short, various studies have demonstrated that Tregs can either directly suppress responder T cells, or indirectly modulate the stimulatory capacity of APCs (Shevach, 2009; Josefowicz *et al.*, 2012). Direct mechanisms include production of anti-inflammatory cytokines (e.g., IL-10 or TGF- β), IL-2 consumption, granzyme-mediated apoptosis, and cell-cycle arrest. Of particular interest here is targeting the ability of APCs to provide efficient CD28 costimulation thereby preventing the initiation of T cell response. Evidence from both *in vitro* and *in vivo* experiments shows that Tregs express high levels of CTLA-4 (Takahashi *et al.*, 2000), which is a nonredundant regulator of immune response and is tightly coupled with CD28 costimulation. CTLA-4 appears to be critical for Treg function (Wing *et al.*, 2008; Schmidt *et al.*, 2009) and the molecular basis of this process is described below.

The Biology of the CD28/CTLA-4 Pathway

CD28 is 44 kDa glycoprotein constitutively expressed on the surface of the majority of T cells. It consists of a single extracellular Ig-V-like domain that forms a homodimer. Although other costimulatory pathways have been discovered (e.g., ICOS pathway), CD28 appears to be the apical costimulatory molecule governing T cell activation. As such, mice deficient in CD28 or its ligands CD80 and CD86 show defects in IL-2 production and T cell proliferation, as well as B cell functions such as germinal center formation, antibody levels, and immunoglobulin class switching.

Upon the recognition of CD80 and CD86 ligands, CD28 relocates to the immune synapse where it initiates various costimulatory signaling events (Figure 1). The exact pathways involved in the signal transduction are still under investigation. To date the importance of various kinases in CD28 signaling has been demonstrated, such as PI3K, ITK, Lck, and PKC- θ (Rudd *et al.*, 2009). These molecules initiate a signaling cascade responsible for the activation of genes involved in survival and cell-cycle progression. However, the outcome and the importance of specific signaling pathways initiated upon CD28 engagement remains controversial. Recent data have indicated that CD28 costimulation may be important in actin reorganization within the T cells in order to facilitate activation signals, in a manner that may not be compensated by simply increasing TCR signals (Tan *et al.*, 2014; Dustin and Davis, 2014).

Functions of CD28

It is still unclear whether CD28 engagement causes unique changes in T cell signaling or simply amplifies TCR-based signals. Over years it has been established that costimulation serves to lower the threshold necessary to achieve the activation of T cells, particularly enabling responses to weak agonist TCR ligands. Data indicate that CD28 engagement ensures more rapid division and enhances production of different cytokines, particularly a T cell growth factor IL-2. In addition, it was demonstrated that CD28 increases the expression of the anti-apoptotic protein BCL-X_L thus promoting the survival of T cells (Acuto and Michel, 2003).

Furthermore, it has been suggested that CD28 signaling influences T cell differentiation. The most prominent defect in CD28 knockout mice is the lack of germinal centers and consequently the production of high-affinity antibodies is impaired (Ferguson *et al.*, 1996; Shahinian *et al.*, 1993). This could be due to an indispensable role of CD28 in driving inducible T cell costimulator (ICOS) upregulation that is necessary for the development of follicular helper T cell (Linterman *et al.*, 2009). Moreover, there are indications that high-CD28 signaling in the context of weak-TCR signals drives the production of Th2 cytokines (Lenschow *et al.*, 1996; Smeets *et al.*, 2012), although CD28 does not seem to be absolutely required for Th2 differentiation (Brown *et al.*, 1996).

In addition to its role in conventional T cell activation, studies in CD28 knockout mice revealed that CD28 is also implicated in the generation and homeostasis of thymic regulatory T cells (Sansom and Walker, 2006). Surprisingly, it has been shown that CD28 is required for Treg generation in thymus independently of its role to stimulate IL-2 production. In addition, Gogishvili *et al.* (2013) have shown that CD28 deletion significantly lowered the number of Tregs, but not of conventional T cells in the periphery. More importantly, they showed that IL-2 is not able to compensate for the role of CD28 costimulation in maintaining the number of Tregs in the periphery. Therefore, it appears that in addition to IL-2, there is a nonredundant cell-intrinsic role of CD28 costimulation in peripheral Treg homeostasis (Sansom and Walker, 2013). A recent study showed slightly different results using a Treg-specific CD28 KO mice system (Zhang *et al.*, 2013). Here, except in the thymus, the number of Tregs was fairly well preserved, however, their function was notably impaired. These Tregs had a significantly decreased expression of CTLA-4 and PD-1, leading to a robust autoimmune phenotype.

Although a role for CD28 in Treg selection in the thymus is clear, the role of CD28 in peripheral Treg induction is still controversial. It has been reported that CD28 signaling negatively regulates differentiation of induced Tregs (Semple *et al.*, 2011). However, others have shown that CD28 promotes differentiation of induced Tregs (Gabrysova *et al.*, 2011). Thus, it is likely that the role of CD28 in Treg generation depends greatly on the other signals that T cell receives, particularly the strength of TCR, as well the cytokine milieu in which the activation occurs.

Overall, the CD28 pathway is critical for T cells activation, differentiation, and function. Moreover, given that T cells interact with a number of immune cell types, regulation of

CD28 pathway is therefore crucial for successful B cell, macrophage, and cytotoxic T cell responses. This places CD28/CTLA-4 pathway at the heart of immune regulation.

CTLA-4 – A Key Negative Regulator of T cell Activation

The second receptor for the CD80 and CD86 is cytotoxic T lymphocyte antigen 4 (CTLA-4) (Figure 1). CTLA-4 is 34 kDa type 1 transmembrane glycoprotein expressed by T cells. Genes for CD28 and CTLA-4 are both located on the same region of the chromosome and they share around 30% of homology at the amino acid level indicating that CD28 and CTLA-4 have evolved from the same ancestral gene. Nonetheless, CD28 and CTLA-4 proteins have a very different expression pattern, biological properties, and effects upon T cell responses. Whereas CD28 is constitutively expressed on the surface of most T cells, CTLA-4 expression is limited to activated conventional T cells and more importantly, Tregs. Upon T cell activation, CTLA-4 is synthesized and directed toward the cell membrane, where it binds to CD80 and CD86 with a higher affinity than CD28. Also unlike CD28, CTLA-4 is a highly

endocytic molecule and the amount of time in which it is resident on the cell surface is very limited (Figure 2).

Whereas CD28 engagement promotes T cell activation, CTLA-4 serves as a negative regulator of T cell responses. The importance of CTLA-4 has been widely demonstrated in CTLA-4 KO mice (Tivol *et al.*, 1995) and recently in humans with heterozygous nonsense/missense CTLA-4 mutations (Schubert *et al.*, 2014). T cells from CTLA-4 KO mice exhibit a strong response to self-antigens, which leads to lymphoproliferative disease and these mice die after 3–4 weeks of age from multiorgan failure. In addition, various polymorphisms in CTLA-4 gene have been implicated in different autoimmune diseases in humans (Gough *et al.*, 2005). These phenotypes provide strong evidence that CTLA-4 evolved to prevent autoimmunity, however, the molecular mechanisms used to achieve this have been controversial.

CTLA-4 and CD28 Bind Same Ligands with Different Affinities

CD80 and CD86 are ligands for CD28 and CTLA-4 and both expressed by APCs (Figure 1). CD28 and CTLA-4 are both dimers and they share a conserved MYPPPY amino acid motif that forms the ligand binding domain (Yu *et al.*, 2011).

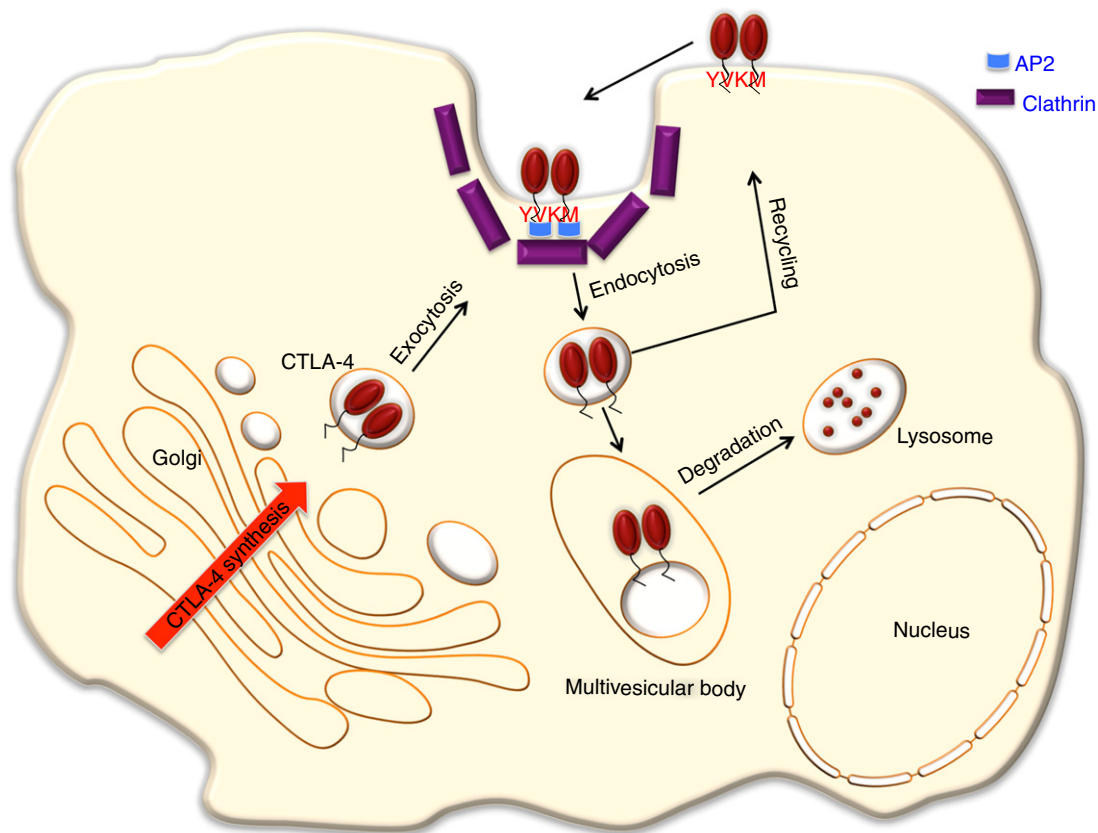


Figure 2 Cell biology of CTLA-4. CTLA-4 is synthesized on endoplasmic reticulum and matures in Golgi forming the functional homodimer. Upon T cell activation vesicles containing CTLA-4 proteins are targeted toward cell membrane. At the cell surface a YVKM motif in the cytoplasmic domain of CTLA-4 is recognized by the clathrin adaptor protein AP-2 leading to a rapid clathrin-mediated endocytosis. Following internalization, CTLA-4 either recycles back to the cell surface or is degraded in lysosomes. Therefore, CTLA-4 is a highly dynamic protein with short half-life, which recycles and is rapidly degraded in lysosomes.

Importantly, CTLA-4 has higher affinity and avidity for both ligands, which allows CTLA-4 to out-compete CD28 for CD80/86 binding. Thus it has been proposed that CTLA-4 can act by trapping CD80 and CD86 and preventing their interactions with CD28 (Walker and Sansom, 2011).

Surprisingly, although CD80 and CD86 share the same receptors, these two ligands have only around 25% of identity at the protein level. In addition, CD80 exists as a dimer while CD86 appears to be a monomer (Collins *et al.*, 2002). Based on chromosomal location it seems likely that CD80 and CD86 have evolved from the same ancestral gene probably in the process of gene duplication (Zhu *et al.*, 2014), and then acquired specific functions. However, the precise functional differences between CD80 and CD86 are not yet understood.

In terms of interaction with their receptors, CD80 has a significantly higher affinities and avidities for both CD28 and CTLA-4 than CD86, thus suggesting that CD80 is a more potent ligand (Collins *et al.*, 2002). However, studies in CD80 and CD86 knockout mice have revealed that CD86 is more important in T cell activation than CD80 (Borriello *et al.*, 1997). The biology behind the superior capacity of CD86 to costimulate is still unclear and further studies are required to distinguish between effects of expression level and those of affinity and structure. More importantly these results raise another question: if CD86 alone is sufficient to costimulate what is then the role of CD80 in T cell differentiation and function? A more comprehensive understanding of exact roles of CD80 and CD86 during different phases of immune response is awaited and will almost certainly increase our knowledge of T cell biology and enable a more efficient use of immunomodulatory drugs that target CD28/CTLA-4 pathway.

How Does CTLA-4 Function?

The autoimmune phenotype in CTLA-4 deficient mice can be prevented by blocking CD80 and CD86 with CTLA4-Ig (Tivol *et al.*, 1997) or by crossing CTLA-4 KO mice with mice deficient for both ligands (Mandelbrot *et al.*, 1999). This clearly demonstrates that CTLA-4 functions as a regulator of CD28 signaling. Accordingly, CD28 independent T cell activation is not controllable by CTLA-4.

The Mechanism of CTLA-4-Mediated Suppression Is Cell-Extrinsic

Numerous studies in CTLA-4 biology have suggested that upon ligation CTLA-4 delivers negative signal (Rudd *et al.*, 2009) by recruiting protein phosphatases PP2A and SYP to TCR complex, which inhibit TCR signaling by turning off T cell activation and IL-2 production. Other studies have shown that CTLA-4 recruits PI3K or competes with CD28 for PKC θ binding. However, there is still debate as to whether CTLA-4 delivers an inhibitory signal, particularly *in vivo*. To begin with, the idea of negative signaling was predominantly based on experiments using agonistic anti-CTLA-4 antibody, however, there is no natural ligand that solely binds CTLA-4 but not CD28, making such reagents

questionable. Moreover, the precise signaling mechanisms still remain unclear, for example, it has been shown that the interaction of CTLA-4 with PP2A is not necessary for the inhibition of T cell responses (Teff *et al.*, 2009). Other pathways proposed to be initiated by CTLA-4 (i.e., PI3K and PKC θ pathway) are also implicated in CD28 signaling and generally are thought to have positive effect upon T cell activation. The fact that CTLA-4 is not well expressed on the cell surface and possesses a highly dynamic trafficking itinerary (see Section Cell biology of CTLA-4), that limits amount of time in which it is expressed on cell surface is also not easily reconciled with a signaling function (Qureshi *et al.*, 2012).

The most persuasive argument to consider mechanisms other than intrinsic inhibitory signals comes from *in vivo* experiments with chimeric mice. These mice contain both CTLA-4^{+/+} and CTLA-4^{-/-} T cells and have challenged the cell-intrinsic view of CTLA-4 function. Unlike CTLA-4^{-/-} mice that die from lymphoproliferative disease, mice containing both CTLA-4^{-/-} and CTLA-4^{+/+} cells fail to develop this pathology (Bachmann *et al.*, 1999). The most likely explanation for this finding is that CTLA-4^{+/+} cells are able to control expansion and effector functions of CTLA-4^{-/-} T cells. This led to the suggestion that CTLA-4 dominantly suppresses T cell activation via modifying phenotype and function of other cell types such as APCs, consistent with a function on a ‘suppressor’ cell such as a Treg. Until recently a molecular mechanism by which CTLA-4 can control T cell activation in cell-extrinsic manner has been difficult to elucidate (Walker and Sansom, 2011). However, we recently found that CTLA-4 was capable of carrying out a novel cellular process we term transendocytosis (Soskic *et al.*, 2014; Qureshi *et al.*, 2011).

Transendocytosis: A Molecular Mechanism of CTLA-4 Function

From a cell biology perspective, the most fundamental feature of the CTLA-4 protein is its highly endocytic nature (Figure 2). This is somewhat surprising giving that CTLA-4 operates on the cell surface by interacting with CD80 and CD86 expressed on APCs. However, the endocytic behavior of CTLA-4 appears to be crucial for CTLA-4 function. Here we explore cell biology of CTLA-4 and how endocytosis is employed as an effector pathway in order to control T cell responses.

Cell biology of CTLA-4

Upon TCR engagement, the level of CTLA-4 is rapidly increased (Egen and Allison, 2002) and vesicles containing CTLA-4 proteins are targeted toward immune synapse in ADP-ribosylation factor 1 (ARF1) and phospholipase D (PLD) dependant fashion (Mead *et al.*, 2005). At the cell surface, a YVKM motif in the cytoplasmic domain of CTLA-4 is recognized by the clathrin adaptor protein AP-2 leading to a rapid clathrin-mediated endocytosis (Shiratori *et al.*, 1997). Earlier studies have proposed that upon activation tyrosine in YVKM motif is phosphorylated, preventing binding of AP-2 (Shiratori *et al.*, 1997). This would cause stabilization of CTLA-4 at the cell surface. Although evidence exists that *in vitro* phosphorylation of CTLA-4 or overexpression of Src kinase prevents AP-2 binding (Follows *et al.*, 2001; Chuang *et al.*, 1999), it is

unclear whether this happens under physiological stimulation. Indeed our group has found that clathrin-mediated endocytosis generally persists during T cell activation, and at any given time approximately 90% of CTLA-4 molecules are expressed intracellularly (Qureshi *et al.*, 2012). Therefore the increase of CTLA-4 at the cell surface following T cell activation is not due to the disruption of CTLA-4 internalization, but rather to the increased delivery of CTLA-4 containing vesicles to the cell surface resulting from the increased protein levels during T cell activation.

Upon internalization, the fate of CTLA-4 is still rather unclear. Our lab has demonstrated that CTLA-4 co-localize with Rab11 (Qureshi *et al.*, 2012), which is known to regulate recycling of endocytosed proteins and there is evidence that CTLA-4 can recycle back to the cell surface. Further precise studies are still required to elucidate motifs and molecular mechanisms involved in CTLA-4 recycling. Multiple studies have also shown that CTLA-4 is degraded in lysosomal compartments and neutralization of lysosomal pH using ammonium chloride (NH_4Cl) increases the expression of CTLA-4 (Kaur *et al.*, 2013; Egen and Allison, 2002). On the other hand, blocking protein synthesis with cycloheximide (CHX) results in the rapid decrease of CTLA-4 level giving a half-life of CTLA-4 of around 2 h (Egen and Allison, 2002). The above data provide evidence that CTLA-4 is a highly dynamic protein with short half-life, which recycles and is rapidly degraded in lysosomes. Understanding the cell biology of CTLA-4 in the context of its function is therefore clearly important (Figure 2).

Overall, although CTLA-4 shares the same ligands with CD28 it has strikingly different biology. The highly dynamic nature of CTLA-4 has presented a major conundrum in the field. Interestingly, it has been shown that cytoplasmic domain of CTLA-4 is highly conserved among all mammals, suggesting that there is a strong evolutionary pressure to maintain a complex CTLA-4 trafficking across species (Walker and Sansom, 2011). We have therefore taken the view that CTLA-4 trafficking is key to its function.

Transendocytosis: cell biology at the heart of immune regulation

In recent studies we have observed that CTLA-4 effectively can act as a molecular Hoover that removes CD80 and CD86 ligands from the surface of APCs, by ‘transendocytosis’ (Qureshi *et al.*, 2011; Figure 3). By depleting ligands, CTLA-4 reduces the capacity of APCs to provide effective costimulation through CD28, thereby directly coupling the function of CTLA-4 to CD28 as observed in *in vivo* experiments. To elucidate the mechanism of CTLA-4-mediated depletion of ligands, we have generated cell lines that expressed CTLA-4 and CD80 or CD86 tagged with green fluorescent protein (GFP) (Figure 4). Perhaps most strikingly, we have observed that the entire receptor–ligand complex (including cytoplasmic domain of ligands tagged with GFP) gets internalized into CTLA-4 expressing cell. The complex is subsequently targeted to lysosomes where ligands are rapidly degraded while CTLA-4 may recycle back to the cell surface to capture more ligands (Figure 3). This latter point still requires more research but it is clear that ligands are degraded and that this is inhibited by agents such as NH_4Cl . Further, to test the relevance of transendocytosis *in vivo* we generated mice that express CD86-GFP on the surface of their APCs. Strikingly, capture

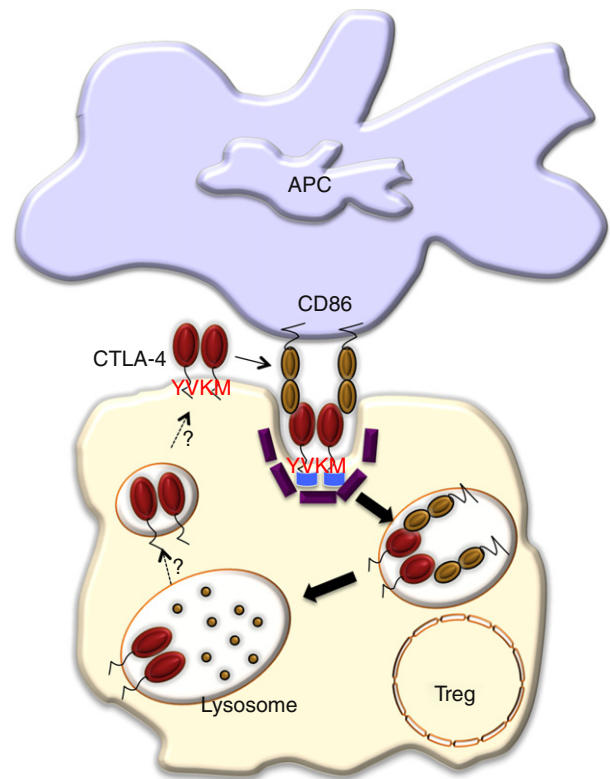


Figure 3 Transendocytosis – A molecular mechanism of CTLA-4 action. Upon activation CTLA-4 containing vesicles are targeted to the immune synapse (the site of T cell–APC interaction), where it binds CD80 or CD86. Subsequently, the entire receptor–ligand complex gets internalized into T cell. The process of CTLA-4 dependent CD80/86 internalization is termed transendocytosis. The complex is targeted to lysosomes where ligands are rapidly degraded while CTLA-4 may recycle back to the cell surface to capture more ligands. As such CTLA-4 expressed by Tregs can remove ligands from APC preventing the activation of all T cells that subsequently recognize antigens presented by that APC. This provides an efficient mechanism of immune suppression and controls the level of CD80 and CD86 expression.

of ligands was observed only in $\text{CD}4^+ \text{CD}25^+$ T cells (activated conventional T cells or Tregs) and $\text{CD}4^+ \text{CD}25^+$ T cells from $\text{CTLA-4}^{-/-}$ mice did not acquire ligands demonstrating that the transfer of ligand is CTLA-4 dependent *in vivo*. More importantly, the transendocytosis occurred only upon peptide stimulation indicating that the process is antigen specific. These features are highly consistent with a role for CTLA-4 on regulatory T cells as an effector mechanism for reducing costimulation via CD28.

Although the exact molecular players involved in CTLA-4 transendocytosis are currently unknown, the concept of intercellular exchange of membrane proteins has been observed elsewhere. A number of studies have shown that various cell types are capable of exchanging their membrane proteins via transendocytosis (Rechavi *et al.*, 2009). A well-characterized example is Notch, which upon the interaction with its ligands is cleaved and internalized into the ligand-expressing cell (Klug and Muskavitch, 1999). This process is necessary to initiate Notch receptor signaling which is implicated in cell differentiation during development. Several other examples of

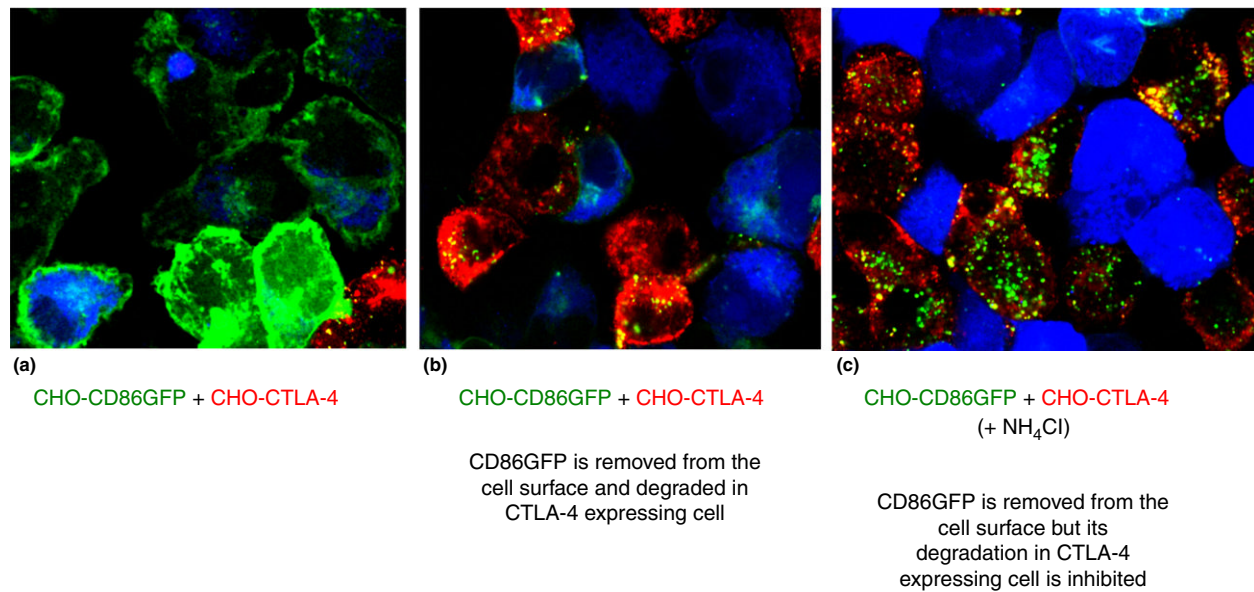


Figure 4 Confocal microscopy of transendocytosis. Chinese hamster ovary (CHO) cells stably expressing CD86-GFP (green) or CTLA-4 (red) were co-incubated for 8 h. In areas of the coculture where CTLA4-expressing cells are absent (a) CD86-GFP is preserved at the plasma membrane. However, where ligand-expressing cells make close contact with CTLA-4 expressing cells CD86 ligand is taken up by the CTLA4-expressing cells and destroyed. Accordingly most of the GFP signal is lost (b). In the presence of lysosomal inhibitor NH₄Cl (c) the degradation of CD86-GFP is blocked and the captured ligand accumulates in intracellular vesicles within the CTLA4-expressing cells.

transendocytosis have been described, including the bi-directional transfer of CD47–SHPS-1 complex (Kusakari *et al.*, 2008), which controls T cell activation, and transendocytosis of ephrin B that regulates contact repulsion (Marston *et al.*, 2003). While transendocytosis involves transfer of ligands directly into receptor-expressing cell, it has some distinctions from similar processes such as Trogocytosis (Daubeuf *et al.*, 2010) which is common in the immune system (Davis, 2007). It has been identified that within the minutes of co-incubation of NK cells with their targets, NK cell receptor KIR is transferred to target cell in MHC I dependent fashion by trogocytosis. Similarly, several studies have shown that CD4⁺ T cells can acquire MHC II, CD80, and CD86 from APCs. Interestingly, CD28-expressing cells can also acquire CD80 and CD86 ligands, however, CD28-mediated transfer is by trogocytosis. In our experience whilst trogocytosed proteins remain detectable on the cell surface after transfer, in transendocytosis transferred proteins are directly routed to an intracellular compartment, which precludes surface detection.

Transendocytosis as a mechanism of immune regulation

A long-term conundrum in immunology has been why a key stimulatory protein (CD28) and a key inhibitory molecule (CTLA-4) would share the same ligands. If CTLA-4 is viewed as a receptor that delivers negative signal, then it is difficult to imagine the evolutionary logic behind controlling activation and inhibition of T cell responses with the same triggers. However, from transendocytosis perspective the core concept behind the model is that CTLA-4 must share the ligands with CD28 in order to function. Furthermore, in order for suppression to be efficient, CTLA-4 needs to be able to compete with CD28 for CD80 and CD86 binding, which is in agreement with biophysical characteristics of interaction between

CD28/CTLA-4 and its ligands. The transendocytosis model therefore describes a form of competition between CD28 and CTLA-4 that is cell-extrinsic and where competition between CD28 and CTLA-4 can be temporally separated. As such CTLA-4 expressed by Tregs can remove ligands from APC preventing the activation of all T cells that subsequently recognize antigens presented by the same APC.

In transendocytosis, the ability of CTLA-4 to internalize CD80 and CD86 from APC membrane is dependent on both the number of ligands that are expressed by APC and number of CTLA-4 molecules. Therefore a key feature of CTLA-4-mediated immune control is that it is quantitative. As such, if the inflammatory environment (i.e., due to the presence of pathogen) drives increased expression of ligands, CTLA-4-mediated suppression has a diminished efficacy. In these settings, if the production of CD80/86 is greater than their downregulation, this allows the engagement with CD28. On the other hand, if the level of CD80/86 ligands is lower, CTLA-4 can act to keep the level below the threshold necessary for costimulation, thus preventing the activation of self-reactive T cells.

CTLA-4: a key mechanism of Treg suppression

Numerous studies have suggested that CTLA-4 plays a major role in Treg-mediated suppression of immune response (Wing *et al.*, 2008; Schmidt *et al.*, 2009; Onishi *et al.*, 2008). Indeed, there is a remarkable similarity between CTLA-4 KO and Scurfy mice, suggesting that the Scurfy phenotype seen in FoxP3 deficient mice is at least partially a result of the loss of CTLA-4 due to absence of Tregs. Furthermore, removal of CD28 or injection of CTLA4-Ig significantly improves survival of Scurfy mice (Singh *et al.*, 2007).

Transendocytosis offers a simple explanation for CTLA-4 dependent Treg function. Upon the antigen recognition, the

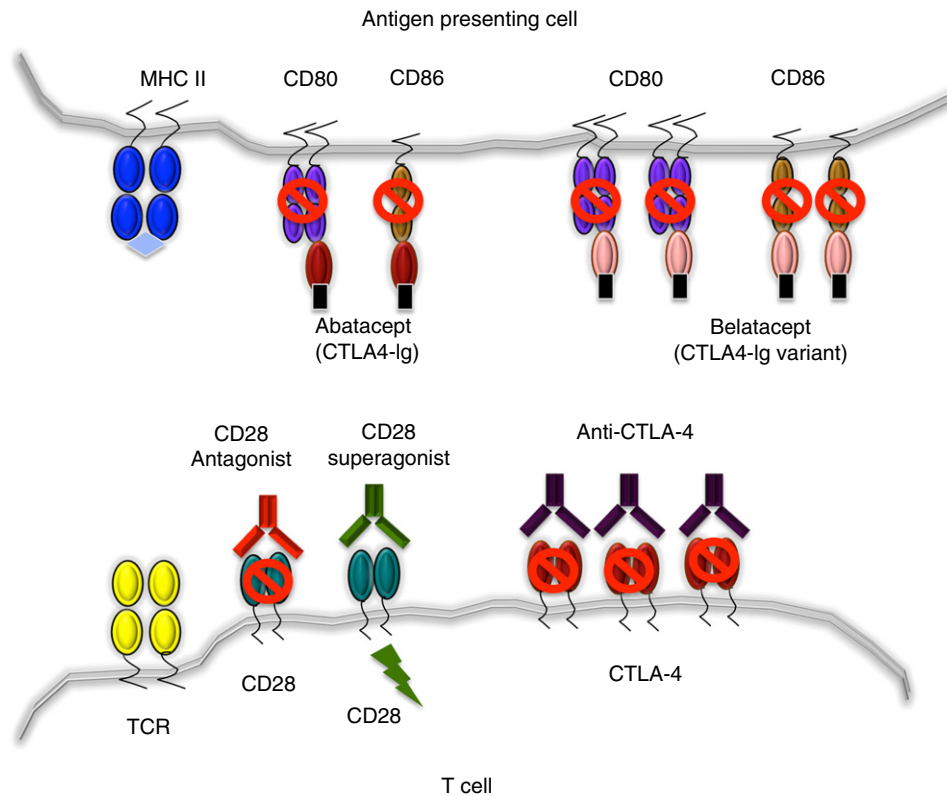


Figure 5 Manipulation of CD28/CTLA-4 pathway. Multiple therapeutic compounds have been designed to manipulate the CD28/CTLA-4 pathway. CTLA4-Ig binds to CD80 and CD86 expressed on APCs preventing their interaction with CD28 and therefore the initiation of T cell responses. CTLA4-Ig is clinically available in two forms: Abatacept and Belatacept. Compared to Abatacept, Belatacept has an increased affinity for CD80 and CD86, with an increased bias toward CD86 blockade. Both these compounds are used as immunosuppressive drugs. As an alternative approach to ligand blockade, CD28-blocking antibodies have been developed for induction of transplantation tolerance and treatment of autoimmunity. In contrast, CD28 superagonistic antibodies have been used to induce immune tolerance by preferential expansion of FoxP3⁺ Tregs. Several CTLA-4-blocking antibodies have also been developed and demonstrated antitumor effects. Thus the system is capable of being manipulated to promote immune tolerance or immune activation depending on the components that are targeted.

expression level and cycling of CTLA-4 dramatically increases in regulatory T cells. As functional consequence, CTLA-4 removes CD80 and CD86 ligands from the surface of APCs and efficiently prevents the activation of self-reactive T cells. In line with this, several studies have demonstrated that Tregs downregulate CD80 and CD86 from APCs in CTLA-4 dependent fashion (Oderup *et al.*, 2006; Onishi *et al.*, 2008; Wing *et al.*, 2008).

Another important concept that arises from this model is that any T cell that expresses CTLA-4 (activated conventional T cells and Tregs) has regulatory capacity (Corse and Allison, 2012; Wang *et al.*, 2012). Indeed it was recently demonstrated that the combination of CTLA-4 expression and IL-2 suppression is sufficient to convert conventional T cells into regulatory T cells with potent immunoregulatory capacity (Yamaguchi *et al.*, 2013).

Manipulation of CD28/CTLA-4 Pathway: Antagonist, Agonists, and Superagonists

The last decade of research in immunology has taught us that the relationship between CD80- and CD86-CD28-CTLA-4 is

extremely well balanced and any disturbance may have a profound influence on the control of the immune response. Because of its central importance, several approaches have been used to manipulate T cell activation through CD28/CTLA-4 pathway (Figure 5).

Blocking CD80 and CD86 is widely used as a strategy for blocking CD28/CTLA-4 pathway. The higher affinity of CTLA-4 for its ligands led to development of CTLA4-Ig fusion protein as an inhibitor of T cell responses. CTLA4-Ig binds to CD80 and CD86 expressed on APCs preventing their interaction with CD28 and accordingly the initiation of T cell responses. CTLA4-Ig is clinically available in two forms: Abatacept and Belatacept. Abatacept has been approved for the treatment of rheumatoid arthritis in patients who failed to respond to antitumor necrosis factor (TNF) therapy (Genovese *et al.*, 2012) and polyarticular juvenile idiopathic arthritis (Goldzweig and Hashkes, 2011). However, a significant proportion of patients do not respond to Abatacept treatment leading to a second generation of artificially engineered CTLA4-Ig variant, LEA29Y (Belatacept). Two amino acid substitutions in Abatacept sequence increases the affinity for CD80 and CD86 with an increased bias toward CD86 blockade. Belatacept has now been approved for clinical use as an immunosuppressant

following kidney transplantation (Gardner *et al.*, 2014; Melvin *et al.*, 2012). Given the above discussions, it is interesting to reflect that removal of ligands by transendocytosis and blocking of ligands with Abatacept may ultimately be functionally equivalent strategies.

Although ligand blockade has been widely accepted as a strategy for manipulating T cell responses, there are several issues that remain to be addressed. To begin with, it is widely appreciated that CD28 costimulation is necessary for generation and homeostasis of regulatory T cells. As such, reducing the number of Tregs by ligand blockade in patients with autoimmunity can be counter-productive. Therefore the impact on both Tregs and effector T cells may be significant for the outcome of the therapy. Another important issue is that CTLA4-Ig can only inhibit CD28-dependent responses and it is still unclear why some T cell responses appear to be less dependent on CD28 signaling (Gardner *et al.*, 2014). As an alternative approach to ligand blockade CD28-blocking antibodies have been developed for induction of transplantation tolerance and treatment of autoimmunity (Poirier *et al.*, 2012). The major challenge of using CD28 antibodies to suppress T cell activation is avoiding potential agonistic effects (Gardner *et al.*, 2014), since CD28 agonists have had a somewhat chequered history. Initially, CD28 superagonistic antibodies were shown to induce immune tolerance by preferential expansion of FoxP3⁺ Tregs in humanized mouse model of graft-versus-host disease (Kitazawa *et al.*, 2009). This data along with numerous other studies have suggested that CD28 superagonistic antibodies may specifically target Treg cells *in vivo* and maintain their suppressive function. However in humans, CD28 superagonistic antibody (TGN1412) caused the rapid release of cytokines from effector memory T cells leading to multiorgan failure (Hunig, 2012). Despite this major setback, recent data again indicates that the selective expansion of Tregs may be possible by low dose of CD28 agonists (Tabares *et al.*, 2014).

Several CTLA-4-blocking antibodies have also been developed and efficacy as antitumor therapies has been tested with remarkable success. Here, the concept is that by blocking natural tolerance mechanisms to our own tissues it may be possible to unleash the immune system against tumors. Ipilimumab, which is the leading CTLA-4 antibody in its class, appears to be effective in treatment of melanoma with a small subset of patients demonstrating remarkable responses. Further understanding of the role of CTLA-4 on both Tregs and activated T cells in response to treatment will almost certainly lead to advancement of anti-CTLA-4 therapy (Callahan *et al.*, 2010). The development of this CTLA-4 blockade approach has now pioneered the rapid development of several tumor immunotherapy approaches (Pardoll, 2012).

Conclusions

- CD28 signaling is essential for T cells activation, differentiation, and function. Moreover, given that T cells interact with a number of immune cell types, CD28 appears to be an apical control point for successful B cell, macrophage, and cytotoxic T cell responses.

- CTLA-4 functions as a direct regulator of CD28 signaling.
- CD28 and CTLA-4 bind the same ligands CD80 and CD86; CTLA-4 has higher affinity and avidity for both ligands.
- CTLA-4 is a highly endocytic protein and has a complex recycling behavior.
- CTLA-4 effectively can act as a 'molecular hoover' that removes CD80 and CD86 ligands from the surface of APCs, by 'transendocytosis.' By depleting ligands, CTLA-4 reduces the capacity of APCs to provide effective costimulation through CD28 and maintains immune tolerance.

See also: Cellular Immunology: Transcriptional Basis of B and T Cell Lineages and Memory Cells: T Follicular Helper Cells

References

- Acuto, O., Michel, F., 2003. CD28-mediated co-stimulation: A quantitative support for TCR signalling. *Nature Reviews Immunology* 3, 939–951.
- Bachmann, M.F., Kohler, G., Ecabert, B., Mak, T.W., Kopf, M., 1999. Cutting edge: Lymphoproliferative disease in the absence of CTLA-4 is not T cell autonomous. *Journal of Immunology* 163, 1128–1131.
- Bennett, C.L., Christie, J., Ramsdell, F., *et al.*, 2001. The immune dysregulation, polyendocrinopathy, enteropathy, X-linked syndrome (IPEX) is caused by mutations of FOXP3. *Nature Genetics* 27, 20–21.
- Bilate, A.M., Lafaille, J.J., 2012. Induced CD4⁺ Foxp3⁺ regulatory T cells in immune tolerance. *Annual Review of Immunology* 30, 733–758.
- Borriello, F., Sethna, M.P., Boyd, S.D., *et al.*, 1997. B7-1 and B7-2 have overlapping, critical roles in immunoglobulin class switching and germinal center formation. *Immunity* 6, 303–313.
- Brown, D.R., Green, J.M., Moskowitz, N.H., *et al.*, 1996. Limited role of CD28-mediated signals in T helper subset differentiation. *Journal of Experimental Medicine* 184, 803–810.
- Brunkow, M.E., Jeffery, E.W., Hjerrild, K.A., *et al.*, 2001. Disruption of a new forkhead/winged-helix protein, scurf, results in the fatal lymphoproliferative disorder of the scurfy mouse. *Nature Genetics* 27, 68–73.
- Callahan, M.K., Wolchok, J.D., Allison, J.P., 2010. Anti-CTLA-4 antibody therapy: Immune monitoring during clinical development of a novel immunotherapy. *Seminars in Oncology* 37, 473–484.
- Chuang, E., Lee, K.M., Robbins, M.D., *et al.*, 1999. Regulation of cytotoxic T lymphocyte-associated molecule-4 by Src kinases. *Journal of Immunology* 162, 1270–1277.
- Collins, A.V., Brodie, D.W., Gilbert, R.J., *et al.*, 2002. The interaction properties of costimulatory molecules revisited. *Immunity* 17, 201–210.
- Corse, E., Allison, J.P., 2012. Cutting edge: CTLA-4 on effector T cells inhibits in trans. *Journal of Immunology* 189, 1123–1127.
- Curotto de Lafaille, M.A., Lafaille, J.J., 2009. Natural and adaptive foxp3⁺ regulatory T cells: More of the same or a division of labor? *Immunity* 30, 626–635.
- Daubeuf, S., Aucher, A., Bordier, C., *et al.*, 2010. Preferential transfer of certain plasma membrane proteins onto T and B cells by trogocytosis. *PLoS One* 5, e8716.
- Davis, D.M., 2007. Intercellular transfer of cell-surface proteins is common and can affect many stages of an immune response. *Nature Reviews Immunology* 7, 238–243.
- Dustin, M.L., Davis, S.J., 2014. TCR signaling: The barrier within. *Nature Immunology* 15, 136–137.
- Egen, J.G., Allison, J.P., 2002. Cytotoxic T lymphocyte antigen-4 accumulation in the immunological synapse is regulated by TCR signal strength. *Immunity* 16, 23–35.
- Ferguson, S.E., Han, S.H., Kelsoe, G., Thompson, C.B., 1996. CD28 is required for germinal center formation. *Journal of Immunology* 156, 4576–4581.
- Follows, E.R., McPheat, J.C., Minshall, C., *et al.*, 2001. Study of the interaction of the medium chain mu 2 subunit of the clathrin-associated adapter protein

- complex 2 with cytotoxic T-lymphocyte antigen 4 and CD28. *Biochemical Journal* 359, 427–434.
- Fontenot, J.D., Gavin, M.A., Rudensky, A.Y., 2003. Foxp3 programs the development and function of CD4+ CD25+ regulatory T cells. *Nature Immunology* 4, 330–336.
- Fontenot, J.D., Rasmussen, J.P., Williams, L.M., *et al.*, 2005. Regulatory T cell lineage specification by the forkhead transcription factor foxp3. *Immunity* 22, 329–341.
- Gabrysova, L., Christensen, J.R., Wu, X., *et al.*, 2011. Integrated T-cell receptor and costimulatory signals determine TGF- β -dependent differentiation and maintenance of Foxp3+ regulatory T cells. *European Journal of Immunology* 41, 1242–1248.
- Gardner, D., Jeffery, L.E., Sansom, D.M., 2014. Understanding the CD28/CTLA-4 (CD152) pathway and its implications for costimulatory blockade. *American Journal of Transplantation* 14, 1985–1991.
- Genovese, M.C., Schiff, M., Luggen, M., *et al.*, 2012. Longterm safety and efficacy of abatacept through 5 years of treatment in patients with rheumatoid arthritis and an inadequate response to tumor necrosis factor inhibitor therapy. *Journal of Rheumatology* 39, 1546–1554.
- Gogishvili, T., Luhder, F., Goebbels, S., *et al.*, 2013. Cell-intrinsic and -extrinsic control of Treg-cell homeostasis and function revealed by induced CD28 deletion. *European Journal of Immunology* 43, 188–193.
- Goldzweig, O., Hashkes, P.J., 2011. Abatacept in the treatment of polyarticular JIA: Development, clinical utility, and place in therapy. *Drug Design, Development and Therapy* 5, 61–70.
- Gough, S.C., Walker, L.S., Sansom, D.M., 2005. CTLA4 gene polymorphism and autoimmunity. *Immunological Reviews* 204, 102–115.
- Goverman, J., Woods, A., Larson, L., *et al.*, 1993. Transgenic mice that express a myelin basic protein-specific T cell receptor develop spontaneous autoimmunity. *Cell* 72, 551–560.
- Hogquist, K.A., Baldwin, T.A., Jameson, S.C., 2005. Central tolerance: Learning self-control in the thymus. *Nature Reviews Immunology* 5, 772–782.
- Hori, S., Nomura, T., Sakaguchi, S., 2003. Control of regulatory T cell development by the transcription factor Foxp3. *Science* 299, 1057–1061.
- Hunig, T., 2012. The storm has cleared: Lessons from the CD28 superagonist TGN1412 trial. *Nature Reviews Immunology* 12, 317–318.
- Josefowicz, S.Z., Lu, L.F., Rudensky, A.Y., 2012. Regulatory T cells: Mechanisms of differentiation and function. *Annual Review of Immunology* 30, 531–564.
- Kaur, S., Qureshi, O.S., Sansom, D.M., 2013. Comparison of the intracellular trafficking itinerary of ctla-4 orthologues. *PLoS One* 8, e60903.
- Kitazawa, Y., Fujino, M., Li, X.K., *et al.*, 2009. Superagonist CD28 antibody preferentially expanded Foxp3-expressing nTreg cells and prevented graft-versus-host diseases. *Cell Transplantation* 18, 627–637.
- Klueg, K.M., Muskavitch, M.A., 1999. Ligand-receptor interactions and trans-endocytosis of delta, serrate and notch: Members of the notch signalling pathway in *Drosophila*. *Journal of Cell Science* 112 (Pt 19), 3289–3297.
- Kusakari, S., Ohnishi, H., Jin, F.J., *et al.*, 2008. Trans-endocytosis of CD47 and SHPS-1 and its role in regulation of the CD47-SHPS-1 system. *Journal of Cell Science* 121, 1213–1223.
- Kyewski, B., Derbinski, J., 2004. Self-representation in the thymus: An extended view. *Nature Reviews Immunology* 4, 688–698.
- Lenschow, D.J., Herold, K.C., Rhee, L., *et al.*, 1996. CD28/B7 regulation of Th1 and Th2 subsets in the development of autoimmune diabetes. *Immunity* 5, 285–293.
- Linterman, M.A., Rigby, R.J., Wong, R., *et al.*, 2009. Roquin differentiates the specialized functions of duplicated T cell costimulatory receptor genes CD28 and ICOS. *Immunity* 30, 228–241.
- Mandelbrot, D.A., McAdam, A.J., Sharpe, A.H., 1999. B7-1 or B7-2 is required to produce the lymphoproliferative phenotype in mice lacking cytotoxic T lymphocyte-associated antigen 4 (CTLA-4). *Journal of Experimental Medicine* 189, 435–440.
- Marston, D.J., Dickinson, S., Nobes, C.D., 2003. Rac-dependent trans-endocytosis of ephrinBs regulates Eph-ephrin contact repulsion. *Nature Cell Biology* 5, 879–888.
- Mathis, D., Benoist, C., 2009. Aire. *Annual Review of Immunology* 27, 287–312.
- Mead, K.I., Zheng, Y., Manzotti, C.N., *et al.*, 2005. Exocytosis of CTLA-4 is dependent on phospholipase D and ADP ribosylation factor-1 and stimulated during activation of regulatory T cells. *Journal of Immunology* 174, 4803–4811.
- Melvin, G., Sandhiya, S., Subraja, K., 2012. Belatacept: A worthy alternative to cyclosporine? *Journal of Pharmacology and Pharmacotherapeutics* 3, 90–92.
- Oderup, C., Cederbom, L., Makowska, A., Cilio, C.M., Ivars, F., 2006. Cytotoxic T lymphocyte antigen-4-dependent down-modulation of costimulatory molecules on dendritic cells in CD4+ CD25+ regulatory T-cell-mediated suppression. *Immunology* 118, 240–249.
- Onishi, Y., Fehervari, Z., Yamaguchi, T., Sakaguchi, S., 2008. Foxp3+ natural regulatory T cells preferentially form aggregates on dendritic cells in vitro and actively inhibit their maturation. *Proceedings of the National Academy of Sciences of the United States of America* 105, 10113–10118.
- Pardoll, D.M., 2012. Immunology beats cancer: A blueprint for successful translation. *Nature Immunology* 13, 1129–1132.
- Poirier, N., Mary, C., Dilek, N., *et al.*, 2012. Preclinical efficacy and immunological safety of FR104, an antagonist anti-CD28 monovalent Fab' antibody. *American Journal of Transplantation* 12, 2630–2640.
- Qureshi, O.S., Kaur, S., Hou, T.Z., *et al.*, 2012. Constitutive clathrin-mediated endocytosis of CTLA-4 persists during T cell activation. *Journal of Biological Chemistry* 287, 9429–9440.
- Qureshi, O.S., Zheng, Y., Nakamura, K., *et al.*, 2011. Trans-endocytosis of CD80 and CD86: A molecular basis for the cell-extrinsic function of CTLA-4. *Science* 332, 600–603.
- Rechavi, O., Goldstein, I., Kloog, Y., 2009. Intercellular exchange of proteins: The immune cell habit of sharing. *FEBS Letters* 583, 1792–1799.
- Rudd, C.E., Taylor, A., Schneider, H., 2009. CD28 and CTLA-4 coreceptor expression and signal transduction. *Immunological Reviews* 229, 12–26.
- Sansom, D.M., Walker, L.S., 2006. The role of CD28 and cytotoxic T-lymphocyte antigen-4 (CTLA-4) in regulatory T-cell biology. *Immunological Reviews* 212, 131–148.
- Sansom, D.M., Walker, L.S., 2013. CD28 costimulation: Walking the immunological tightrope. *European Journal of Immunology* 43, 42–45.
- Schmidt, E.M., Wang, C.J., Ryan, G.A., *et al.*, 2009. Ctla-4 controls regulatory T cell peripheral homeostasis and is required for suppression of pancreatic islet autoimmunity. *Journal of Immunology* 182, 274–282.
- Schubert, D., Bode, C., Kenefick, R., *et al.*, 2014. Autosomal dominant immune dysregulation syndrome in humans with CTLA4 mutations. *Nature Medicine* 20, 1410–1416.
- Semple, K., Nguyen, A., Yu, Y., *et al.*, 2011. Strong CD28 costimulation suppresses induction of regulatory T cells from naive precursors through Lck signaling. *Blood* 117, 3096–3103.
- Sewell, A.K., 2012. Why must T cells be cross-reactive? *Nature Reviews Immunology* 12, 669–677.
- Shahinian, A., Pfeffer, K., Lee, K.P., *et al.*, 1993. Differential T cell costimulatory requirements in CD28-deficient mice. *Science* 261, 609–612.
- Shevach, E.M., 2009. Mechanisms of foxp3+ T regulatory cell-mediated suppression. *Immunity* 30, 636–645.
- Shiratori, T., Miyatake, S., Ohno, H., *et al.*, 1997. Tyrosine phosphorylation controls internalization of CTLA-4 by regulating its interaction with clathrin-associated adaptor complex AP-2. *Immunity* 6, 583–589.
- Singh, N., Chandler, P.R., Seki, Y., *et al.*, 2007. Role of CD28 in fatal autoimmune disorder in scurfy mice. *Blood* 110, 1199–1206.
- Smeets, R.L., Fleuren, W.W., He, X., *et al.*, 2012. Molecular pathway profiling of T lymphocyte signal transduction pathways; Th1 and Th2 genomic fingerprints are defined by TCR and CD28-mediated signaling. *BMC Immunology* 13, 12.
- Soskic, B., Qureshi, O.S., Hou, T., Sansom, D.M., 2014. A transendocytosis perspective on the CD28/CTLA-4 pathway. *Advances in Immunology* 124, 95–136.
- Tabares, P., Berr, S., Romer, P.S., *et al.*, 2014. Human regulatory T cells are selectively activated by low-dose application of the CD28 superagonist TGN1412/ Table08. *European Journal of Immunology* 44, 1225–1236.
- Takahashi, T., Tagami, T., Yamazaki, S., *et al.*, 2000. Immunologic self-tolerance maintained by CD25(+) CD4(+) regulatory T cells constitutively expressing cytotoxic T lymphocyte-associated antigen 4. *Journal of Experimental Medicine* 192, 303–310.
- Tan, Y.X., Manz, B.N., Freedman, T.S., *et al.*, 2014. Inhibition of the kinase Csk in thymocytes reveals a requirement for actin remodeling in the initiation of full TCR signaling. *Nature Immunology* 15, 186–194.
- Teft, W.A., Chau, T.A., Madrenas, J., 2009. Structure-Function analysis of the CTLA-4 interaction with PP2A. *BMC Immunology* 10, 23.
- Tivol, E.A., Borriello, F., Schweitzer, A.N., *et al.*, 1995. Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA-4. *Immunity* 3, 541–547.
- Tivol, E.A., Boyd, S.D., McKeon, S., *et al.*, 1997. CTLA4lg prevents lymphoproliferation and fatal multiorgan tissue destruction in CTLA-4-deficient mice. *Journal of Immunology* 158, 5091–5094.
- Walker, L.S., Abbas, A.K., 2002. The enemy within: Keeping self-reactive T cells at bay in the periphery. *Nature Reviews Immunology* 2, 11–19.

- Walker, L.S., Sansom, D.M., 2011. The emerging role of CTLA4 as a cell-extrinsic regulator of T cell responses. *Nature Reviews Immunology* 11, 852–863.
- Wang, C.J., Kenefeck, R., Wardzinski, L., *et al.*, 2012. Cutting edge: Cell-extrinsic immune regulation by CTLA-4 expressed on conventional T cells. *Journal of Immunology* 189, 1118–1122.
- Wing, K., Onishi, Y., Prieto-Martin, P., *et al.*, 2008. CTLA-4 control over Foxp3 + regulatory T cell function. *Science* 322, 271–275.
- Yamaguchi, T., Kishi, A., Osaki, M., *et al.*, 2013. Construction of self-recognizing regulatory T cells from conventional T cells by controlling CTLA-4 and IL-2 expression. *Proceedings of the National Academy of Sciences of the United States of America* 110, E2116–E2125.
- Yu, C., Sonnen, A.F., George, R., *et al.*, 2011. Rigid-body ligand recognition drives cytotoxic T-lymphocyte antigen 4 (CTLA-4) receptor triggering. *Journal of Biological Chemistry* 286, 6685–6696.
- Zhang, R., Huynh, A., Whitcher, G., *et al.*, 2013. An obligate cell-intrinsic function for CD28 in Tregs. *Journal of Clinical Investigation* 123, 580–593.
- Zhu, L.Y., Lin, A.F., Shao, T., *et al.*, 2014. B Cells in teleost fish act as pivotal initiating APCs in priming adaptive immunity: An evolutionary perspective on the origin of the B-1 cell subset and B7 molecules. *Journal of Immunology* 192 (6), 2699–2714.

Roles of Stromal Cells in the Immune System

EW Roberts, University of California, San Francisco, CA, USA

AE Denton, University of Cambridge, Cambridge, UK

DT Fearon, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA and Weill Cornell Medical College, New York, NY, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction to Stromal Cells

Originally described as noncancerous cells within a tumor, stromal cells are now typically defined as connective tissue cells of an organ. Stroma is comprised of the extracellular matrix and a number of cell types, including mesenchymal cells such as fibroblasts and pericytes, as well as endothelial cells. Stromal cells provide structural support for organs, producing extracellular matrix proteins and basement membrane components. Increasingly, their important role in actively regulating organ homeostasis and function is now becoming better known and in this article we will review the roles of stromal cells in regulating the immune system in both lymphoid and non-lymphoid tissues.

Stromal Cell Regulation of Lymphoid Tissue Structure and Function

Structure of Secondary Lymphoid Organs

Secondary lymphoid organs (SLOs) form a functionally diverse network of tissues that survey the lymph and blood for evidence of foreign organisms in need of clearance from the body. While the spleen and lymph node (LN) are the best studied SLOs, other lymphoid structures such as Peyer's patches, gut- and nasal-associated lymphoid tissues and tertiary lymphoid tissues (TLOs) have similar structures. For the purposes of this article we will focus on the LN, although the stromal cells of the splenic white pulp are essentially functionally equivalent to that present in the LN.

LNs are situated throughout the body, scanning the lymph for pathogens as the lymph drains to be returned to the blood stream. LNs are also encapsulated, and can be divided into three distinct areas: the subcapsular sinus (SCS), cortex and medulla. The SCS is lined with macrophages, which capture antigens and particles that enter the LN via the lymph. The cortex is divided into the outer and para cortex; the former contains the B cell follicles and the latter the T cell zone. Within the T cell zone, specialized blood vessels called high endothelial venules (HEVs) mediate lymphocyte entry into the LN. The lymphocytes, depending on their identity as T or B cells, then localize to the T cell zone, which also contains antigen-presenting dendritic cells (DCs), or to the B cell zone, which contains follicular dendritic cells (FDCs), which have an antigen-presenting function. The localization of lymphocytes in the LN and spleen is controlled by chemokines produced by distinct stromal cell types that populate each area.

Stromal Cells in SLOs

Multiple stromal cell types control the localization of immune cells within SLOs. There are several distinct stromal cell types

within the LN, including endothelial and mesenchymal cells. Endothelial cells line the blood (BEC) and lymphatic (LEC) vessels and not only control the flow of blood and lymph through the LN, but also mediate entry and exit of lymphocytes into the LN. There are five types of mesenchymal stromal cells that populate the LN, which have distinct roles in regulating LN structure and function, but are likely to share a common origin ([Castagnaro et al., 2013](#)). Within the T cell zone, the fibroblastic reticular cells (FRCs) form a dense network of stromal cells that direct the migration of T cells and DCs through interactions between CC chemokine ligand (CCL)19 and CCL21, produced by FRCs, and CC chemokine receptor (CCR)7, expressed on T cells and DCs ([Bajenoff et al., 2006](#); [Forster et al., 1999](#); [Link et al., 2007](#)). FRCs produce interleukin (IL)-7, and thus serve as an important source of T cell survival factors ([Link et al., 2007](#)), and surround and form an important conduit network that shuttles antigens, chemokines and other small molecules through the LN ([Sixt et al., 2005](#)). In the B cell follicle, the FDCs form an intricate stromal network that attracts and retains CXC chemokine receptor (CXCR)5-expressing B cells through expression of CXC chemokine ligand (CXCL)13, and captures and presents antigens and immune complexes during an immune reaction ([Allen and Cyster, 2008](#)). Marginal reticular cells (MRCs) line the SCS and are closely associated with FDCs. Until recently, their function was not well understood but they are now thought to be a precursor for FDCs in the LN ([Jarjour et al., 2014](#)). Lastly, stromal cells called double-negative cells (DNCs) have been described in the LN, so named because they express neither gp38 nor CD31. Among these cells is a population of stromal cells that expresses integrin alpha 7 (ITGA7). These DNCs are closely associated with HEVs, sitting between the BECs and FRCs, and have contractile properties that define them as pericyte-like cells ([Malhotra et al., 2012](#); [Figure 1](#)).

Stromal Cell Regulation of Lymphocyte Migration

The control of lymphocyte and DC migration in LNs is critical to the capacity of the immune system to survey for infection and generate immune responses required for pathogen clearance. Lymphocytes traffic between lymph nodes using the blood and lymph ([Girard et al., 2012](#)). Lymphocytes typically enter LNs via the blood at HEVs. Through the expression of CD62L ligands (in particular glycosylation-dependent cell adhesion molecule-1, GlyCAM-1) and CCL21, HEVs encourage lymphocyte rolling, adhesion and transmigration of CD62L- and CCR7-expressing lymphocytes, resulting in their recruitment into the LN. After entry into the LN, T and B cells then localize to their specific regions, directed by the stromal components therein: FRCs make CCL19 and 21 and thus retain T cells and DCs ([Link et al., 2007](#)); FDCs produce

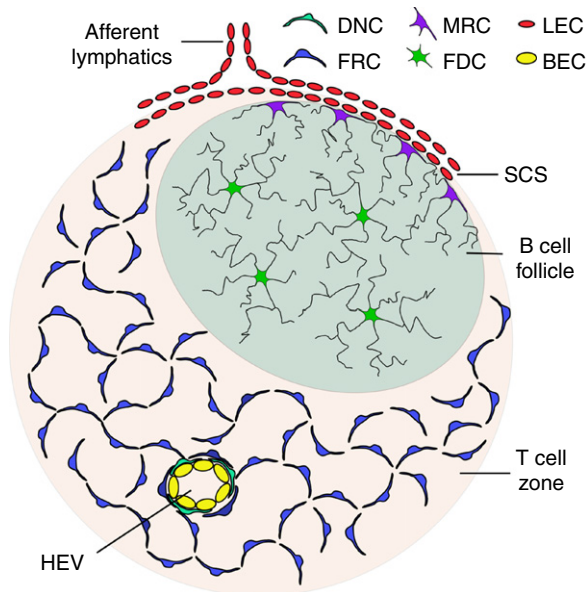


Figure 1 Stromal cells control the structure of lymphoid tissues. Lymph nodes (LNs) are situated throughout the body and filter the lymph as it drains from peripheral tissues. LNs are critical structures for the maintenance of T and B cells and the generation of immune responses. T and B cells enter the LN from the blood through specialized blood vessels called high endothelial venules (HEVs). The blood endothelial cells (BECs, yellow) of the HEVs express the chemokine CCL21 and various adhesion molecules that promote lymphocyte adhesion, rolling, and transmigration into the LN. Once inside the LN, T cells and B cells localized to specific regions, guided by the stromal cells therein. The T cell zone retains T cells and dendritic cells (DCs), and is populated by the mesenchymal stromal cell, the fibroblastic reticular cell (FRC, blue). FRCs produce CCL19 and CCL21 to attract CCR7-expressing T cells and DCs and T cell survival factors such as IL-7; secrete extracellular matrix components that are the scaffold of the T cell zone; and form conduit networks that shuttle small molecules through the LN. CXCR5-expressing B cells localize to the B cell follicle, attracted by the chemokine CXCL13, which is produced by the follicular dendritic cells (FDCs, green) that form the stromal network of the follicle. Marginal reticular cells (MRCs, purple) are found between the subcapsular sinus (SCS) and follicle and involved in antigen transport between the SCS and the FDCs in the follicle. T and B cells reside in the LN for 12–18 h, after which they egress at cortical sinuses into efferent lymphatic vessels for recirculation through the blood. Lymphatic vessels are lined by lymphatic endothelial cells (LECs, red) and express signals that promote lymphocyte egress for recirculation, and chemokines such as CCL21, which promotes the transmigration of antigen-bearing DCs as they traffic from peripheral sites in the lymph.

CXCL13 to attract B cells, which initially enter LNs via CCR7-mediated mechanisms. FDCs also express B cell activating factor (BAFF), which helps recruit and activate B cells in the follicle (Allen and Cyster, 2008). It has recently been shown that FRCs in or near the B cell follicle are also a significant source of BAFF (Cremasco *et al.*, 2014). Naïve lymphocytes are retained in the LN through a combination of CCR7-mediated retention and sphingosine-1 phosphate receptor-1 (S1PR1)-mediated egress signals; lymphocytes transiently downregulate S1PR1 in the blood and regain expression of this receptor within tissues (Matloubian *et al.*, 2004). Reexpression of

S1PR1 is required for lymphocyte egress, which occurs at cortical sinuses where the LECs lining this structure produce and export sphingosine-1 phosphate (S1P), the ligand for S1PR1 (Grigorova *et al.*, 2009). The major source of S1P in blood is red blood cells, whereas tissue stromal and myeloid cells express enzymes that degrade S1P, keeping the S1P concentration low in tissues. Therefore, LECs are critical in the egress of lymphocytes from the LN. Peripheral tissue LECs, through their expression of CCL21, also play an important role in controlling T cell and DC migration from the tissues, into lymphatics and the LN. Activated antigen-bearing DCs trafficking from peripheral sites follow a CCL21 gradient toward the LN (Weber *et al.*, 2013). Once DCs reach the LN, they migrate across the lymphatic vessel and use the CCL19/21 gradients supplied by the FRCs in the cortex to find T cells (Bajenoff *et al.*, 2006), and begin an immune reaction. LECs also support the transmigration of CCR7-expressing T cells via the lymph into the LN (Girard *et al.*, 2012).

Initiation and Regulation of Adaptive Immune Responses by Stromal Cells

Stromal cells also play important roles in coordinating the initiation of T cell responses to infection. The FRC network is absolutely essential to maximize the opportunity for T cells to interact with antigen-bearing DCs that have drained from peripheral sites. The maximization of T cell–DC interaction is chiefly due to the production of the chemokines CCL19/21, which bring CCR7-expressing T cells and DCs together to initiate immune responses (Bajenoff *et al.*, 2006). The paucity of LN T cells (*plt* mouse strain is a line with a spontaneous mutation in the CCL19/21 genes, rendering the mice null for CCL19/21 expression in SLOs (although some expression of CCL21-Leu still occurs in the LECs) (Nakano *et al.*, 1997). Immunization or infection of *plt* mice results in delayed but enhanced immune responses, some of which is mislocalized in the splenic red pulp (Mori *et al.*, 2001; Nakano *et al.*, 2009). Further studies have since shown that mice deficient in CCR7 or CCL19 fail to generate adequate immune responses to challenge (Forster *et al.*, 1999; Link *et al.*, 2007), most likely due to the lack of normal LN architecture in the absence of these ligands. A recent study has since shown that acute depletion of FRCs from normal intact adult LNs also results poor initiation of immune responses due to a lack of T cells and DCs in the absence of FRCs, showing that the presence of FRCs, and not just their ability to create the LN structure, is important for the initiation of immune responses (Cremasco *et al.*, 2014; Denton *et al.*, 2014). Interestingly, Ludewig and colleagues have shown that, while a FRC-like network of stromal cells is formed in the absence of FRC-specific lymphotoxin beta receptor (LT β R) signaling, the ‘maturation’ of FRCs through LT β R during development is essential for the initiation of antiviral immune responses (Chai *et al.*, 2013).

It is well known that FRCs actively respond to infection and immunization, proliferating and increasing in number as the LN swells in size (Denton *et al.*, 2014; Yang *et al.*, 2014). Curiously, the FRC expansion is typically delayed relative to that of the CD45⁺ cells, and continues with the ongoing expansion of antigen-specific T cells, suggesting that the early

increase in naïve lymphocytes in the LN during an immune reaction may instruct the expansion of stromal cells (Yang *et al.*, 2014). Stromal cells undergo a number of transcriptional changes in response to inflammation (Malhotra *et al.*, 2012), and a decrease in CCL19/21 and CXCL13 expression has been observed in responding lymph nodes during infection (Mueller *et al.*, 2007), suggesting that stromal cell responses to infection or immunization may help shape the immune response. For example, the expression of retinoic acid by mesenteric LN stromal cells preferentially induces a gut-homing phenotype in activated T cells (Hammerschmidt *et al.*, 2008). Interestingly, recent studies suggested that LN stromal cells, in particular LECs and FRCs, can suppress T cell proliferation *in vitro* (Khan *et al.*, 2011; Lukacs-Kornek *et al.*, 2011; Siegert *et al.*, 2011). This is dependent on nitric oxide synthase (Nos)2 production by the stromal cells, induced by T cell-derived interferon (IFN)- γ . IFN γ -mediated upregulation of Nos2 expression by FRCs has also been described *in vivo*, and this occurs very early in a T cell-mediated response (Lukacs-Kornek *et al.*, 2011), or within 24–28 h of infection (Siegert *et al.*, 2011). While bone marrow chimeric experiments could not directly demonstrate that stromal cell-derived Nos2 could suppress endogenous antiviral T cell responses, another study has since shown that depletion of FRCs once an immune response is established, and during a phase of active T cell proliferation and differentiation, did not significantly alter the ongoing immune response or result in release T cells from any suppressive effect of FRCs (Denton *et al.*, 2014). It is important to note, however, that these depletion experiments were done at a later phase of the immune response than when virally-induced Nos2 upregulation was observed, suggesting that at this stage of the response, FRC-mediated suppression would not occur. This latter study also demonstrated that, while it has been predicted that relinquishment of stromal control over T cell responses may be important for the optimal differentiation of activated T cells (Mueller *et al.*, 2007), depletion of the stromal cells themselves does not alter the resultant immune response. Taken together, studies thus far suggest that once initiated, T cell responses are independent of the FRC network, although the responses of stromal cells early in infection may help guide the development of immune responses.

The stromal cells in the B cell follicle also undergo significant changes during an immune reaction, and these changes promote the formation of germinal centers (GCs), in which develop long-lived high-affinity antibody-secreting B cells. In addition to the expression of CXCL13, FDCs influence B cell biology through the expression of complement (CR) and Fc receptors, BAFF and a range of adhesion molecules (reviewed in (Allen and Cyster, 2008)). During the immune response, the B cell follicle swells and the stromal network undergoes considerable changes, including an increase in FDCs and the conversion of T cell zone-resident stromal cells into CXCL13-expressing FDC-like cells that delineate the new follicle boundary (Mionnet *et al.*, 2013). As the GC develops, B cells migrate toward the T-B border where they present antigen to and receive costimulatory signals from T helper cells. The B cells that received further stimulation then migrate back into the GC, proliferate, undergo somatic hypermutation and develop into long-lived memory B cells or plasma cells.

The GC reaction requires interactions with a unique subset of CD4 + T cells, termed T follicular helper cells (Tfh), which promote the selection of high-affinity antibody-secreting B cells in the GC (Linterman *et al.*, 2012). Depletion of FDCs prior to immunization reduces B cell motility and prevents GC formation, while the disruption of FDC networks during the GC reaction leads to disruption of B cell clustering and results in B cell death (Wang *et al.*, 2011), suggesting that FDCs are important to both the initiation and continued function of GCs. Moreover, the recruitment of Tfh is dependent on FDC-derived chemokines: as Tfh differentiate from CD4 + T cell in the T cell zone they upregulate CXCR5 and lose expression of CCR7, and become responsive to CXCL13-mediated migration signals from FDCs in the GC (Ansel *et al.*, 1999). The GC separates into two distinct zones, termed the light and dark zone. Movement of B cells between these zones is required for proper B cell differentiation, and the differential expression of CXCL13 and CXCL12 in the light and dark zones, respectively, is thought to control B cell movement between these zones (Allen and Cyster, 2008). FDCs also directly present antigen to B cells, and thus are integral to the selection of high-affinity B cell clones, while their ability to trap antigen in the form of immune complexes, through expression of CRs, allows them to act as a long-term reservoir for antigens. FDCs are thought to acquire antigen from multiple sources, including through the lymph, which drains via the SCS and conduit network to FDCs; direct capture from antigen-bearing DCs and macrophages; and ‘hand-off’ from B cells via CRs (Allen and Cyster, 2008). Bajenoff and colleagues recently demonstrated that while FDC networks expand during GC reactions, the new cells are generated by the differentiation of MRCs into FDCs, rather than proliferation of FDCs themselves (Jarjour *et al.*, 2014). Given that MRCs line the region between the follicle and the SCS, this suggests that MRCs may capture antigen at the SCS and maintain the antigen depot as they differentiate into FDCs in the GC.

Regulation of Peripheral Tolerance by Stromal Cells

Tolerance is an important mechanism for the prevention of autoimmunity, and is chiefly achieved by deletion, or inactivation, of self-reactive T cell clones. Tolerance can occur in two distinct ways: central tolerance, which occurs in the thymus; and peripheral tolerance, which is extrathymic. Lymph node stromal cells have recently been shown to be an important contributor to peripheral tolerance. Central tolerance is chiefly mediated by autoimmune regulator (AIRE)-mediated expression of peripheral tissue antigens (PTAs), which induces deletion of reactive clones during thymic maturation. While a population of AIRE-expressing stromal cells has been identified in SLOs (Gardner *et al.*, 2008), AIRE-independent expression of PTAs in LN stromal cells has also been described. Both FRCs and LECs are adept at presenting model PTAs and inducing tolerance to endogenous antigens, in both inflammatory and steady-state conditions (Cohen *et al.*, 2010; Fletcher *et al.*, 2010), although LECs seem more proficient in this capacity than FRCs. LN stromal cells have recently been shown to acquire peptide/MHC class II complexes from DCs and present these complexes to induce antigen-specific

tolerance of CD4 + T cells (Dubrot *et al.*, 2014). Considering that LN FRCs are specialized mesenchymal stromal cell, it is tempting to speculate that mesenchymal stromal cells in other tissues may exhibit similar tolerizing behavior, in both steady-state and inflammatory conditions, which may contribute to poor T cell-mediated clearance of inflammatory lesions.

Stromal Cell Regulation of Immune Responses in Peripheral Tissues

Stromal Cells in Induced Tertiary Lymphoid Organs

Tertiary lymphoid organs (TLOs) are typically defined as lymphoid-like structures, containing clearly demarcated T and B cell zones that develop after birth. TLOs form during viral and bacterial infections, chronic transplant rejection, autoimmune disorders, cancer, and other detrimental pathologies (reviewed in Neyt *et al.* (2012)). TLOs have structure highly similar to SLOs, including FDC and FRC populations that control the localization of B and T cells through expression of CXCL13 and CCL21, respectively. In response to influenza A virus infection, a TLO forms in the lung, termed inducible bronchus-associated lymphoid tissue (iBALT), and this structure is particularly well-studied. The iBALT generates robust T and GC B cell responses, good enough to clear an infection in the absence of SLOs (Moyron-Quiroz *et al.*, 2004). Moreover, the presence of iBALT prior to infection significantly improves survival, viral clearance and antibody production and decreases lung immunopathology and infection-associated morbidity (Wiley *et al.*, 2009). While iBALT formation is likely important for early antigen-specific immunity to a subsequent infection (Moyron-Quiroz *et al.*, 2006), the maintenance of immune reactions by local TLOs can also be detrimental to the host. In a number of autoimmune diseases TLOs increase with disease pathogenesis; are associated with more severe immunopathology; and promote disease-associated mortality in mouse models (Astorri *et al.*, 2010; Lochner *et al.*, 2011), presumably through promoting the production of auto-reactive T and B cells. While antigen-specific responses are detected in TLOs, it is not known whether TLOs directly induce autoimmunity or merely promote the established disease.

Initiation of Tissue Inflammation

Whilst discussion to this point has been limited to stromal cells within lymphoid organs, BECs and fibroblasts in peripheral tissues also have important roles in modulating the immune system. Upon initiation of an inflammatory response cytokines such as tumor necrosis factor (TNF α) and IL-1 act on BECs inducing the expression of vasodilatory agents such as Nos, the upregulation of adhesion molecules including P-selectin, vascular cell adhesion molecule-1 (VCAM-1) and intercellular adhesion molecule-1 (ICAM-1) and chemokines (Mai *et al.*, 2013). The BEC response to infection does not consist of a binary switch between resting and activated; adhesion molecule and chemokine expression are modulated by the surrounding milieu, for example, expression of CXCL10 and E-selectin is increased in T helper type 1 (Th1) biased environments (Austrup *et al.*, 1997). Whilst these may be seen

to be responses to the surrounding cytokine environment rather than direct organization of the immune response by BECs, BECs can directly sense pathogen associated molecular patterns (PAMPs) through pattern recognition receptors (PRRs) and can direct appropriate immune responses. Toll-like receptor (TLR)4 expression by endothelium has been shown to be important for the recruitment of neutrophils to sites of gram-negative bacterial infections (Andonegui *et al.*, 2003). BECs have also been shown to express major histocompatibility complex (MHC) I and II (von Willebrand *et al.*, 1985) and costimulatory molecules (Khayyamian *et al.*, 2002) leading to the suggestion that BECs may interact with memory cells and may recruit antigen-specific memory CD4 + T cells during recall responses in peripheral tissues (Manes and Pober, 2008). More recently it has been shown that once CD4 + T effector cells have been recruited to a site, BECs upregulate MHCII in response to IFN γ and this allows recruitment of antigen-specific regulatory T (Treg) cells which can then lead to resolution of the inflammatory response (Fu *et al.*, 2014). Under steady state conditions, however, endothelial cells should not be seen as quiescent. In the liver BECs capture and cross present self antigens and recruit CD8 + T cells and induce tolerance in a manner analogous to that seen by LECs in the LN (von Oppen *et al.*, 2009; Cohen *et al.*, 2010). It is, however, difficult to ascribe a role to resting state BECs due to the great degree of diversity in BEC phenotype throughout an organism both in surface markers and in function (Abbott *et al.*, 2006; Kumar *et al.*, 1987). As such, it is possible that BECs in different organs have entirely different roles in regulating the immune response at steady state which are yet to be uncovered.

Whilst BECs regulate the interface between the blood and the parenchyma and thus have an obvious role in regulating the immune response, fibroblasts found within the tissue also have important roles to play. At the most basic level fibroblasts are responsible for deposition of the extracellular matrix (ECM) along which immune cells travel within tissues once they extravasate (Kalluri and Zeisberg, 2006). ECM is not, however, a passive scaffold and modifications in the tensile strength and composition of ECM have dramatic impacts on cellular behavior (Levental *et al.*, 2009). However, a review of the effect of ECM on immune responses is beyond the scope of this article and has been reviewed relatively recently (Sorokin, 2010). Fibroblasts, like BECs, also express PRRs and have roles in initiating immune responses. Synovial fibroblasts have been shown to respond to both Tenascin-C and lipopolysaccharide (LPS) in a TLR4-dependent manner to produce distinct combinations of IL-6, IL-8 and TNF α (Midwood *et al.*, 2009). While fibroblasts can exhibit remarkably distinct transcriptomic profiles (Chang *et al.*, 2002), fibroblasts identified by expression of fibroblast activation protein-alpha (FAP) have very similar gene expression profiles (Roberts *et al.*, 2013). Indeed, this transcriptome is remarkably pro-inflammatory, and has much in common with that of cancer-associated fibroblasts (CAFs), which also express FAP (Erez *et al.*, 2010; Feig *et al.*, 2013; Roberts *et al.*, 2013). The fact that FAP + fibroblasts are primed for expression of inflammatory mediators, coupled to their apparent lack of TLR expression, suggests that the coordinated response to multiple stromal cell types is important for the initiation of inflammation, and this

inflammatory process is essential for the wound healing response in multiple tissues (Eming *et al.*, 2007).

While stromal cells are important for the initiation of inflammation, they also play a role later in the response. Once an inflammatory infiltrate has been established fibroblasts can produce survival factors such as type I IFN, IL-15, and BAFF due to TLR engagement which prolong the survival of the inflammatory lesion (Benito-Miguel *et al.*, 2012). Indeed, the removal of these survival factors for inflammatory cells has been suggested as a mechanism by which fibroblasts promote the resolution of inflammation (Buckley *et al.*, 2001). Fibroblasts have been posited to have a more active role in resolving inflammation through active suppression of immune responses at sites of inflammation. Mesenchymal stem cells (MSCs), a type of stromal cell, have been defined by their ability to differentiate into adipocytes, chondrocytes, and osteoblasts. They have been isolated based on a variety of different surface markers which vary by strain of mouse (Peister *et al.*, 2004). Most work demonstrating an active role for MSCs has been carried out *in vitro* with ex vivo-expanded MSCs, which once activated by monocyte derived cytokines, suppress T cell proliferation (Groh *et al.*, 2005). Indeed this further work suggested that these effects could differentially affect antigen-specific T cells (Karlsson *et al.*, 2008). It has also been shown *in vitro* that these cells can alter macrophage activation through production of soluble mediators (Maggini *et al.*, 2010). These roles, if representative of what occurs *in vivo*, could indicate important functions for fibroblasts in actively switching the phenotype of an inflammatory lesion to that of an immune suppressive environment suitable for regenerative wound healing rather than through repair. Studies of MSC function *in vivo*, however, do not address the normal role for MSCs as they rely on transfer of high numbers of ex vivo-expanded MSCs (Yanez *et al.*, 2006). As such it is unclear whether these functions are important in normal wound healing responses. As with BECs, however, it is important to recognize the incredible heterogeneity found amongst fibroblasts (Aubin, 2001) whilst ascribing specific roles to them. Indeed within the skin fibroblasts from different sites showed as much transcriptional variation as seen between leukocyte populations (Chang *et al.*, 2002). However, fibroblasts isolated from distinct sites based on expression of FAP had similar transcriptomic profiles suggesting that populations common to diverse sites may be parsed out through the use of diverse surface markers (Roberts *et al.*, 2013). As such diverse populations of fibroblasts may have distinct, even opposing, roles within tissues during the initiation and resolution of inflammation.

Inflammatory Diseases

During chronic inflammatory conditions the roles played by BECs and fibroblasts can become maladaptive. As with macrophage polarization, BECs and fibroblasts can ‘polarize’ in different ways, which may contribute to pathology by various means. For example, stromal cells can both prevent an effective immune response or support an excessive response. Cancer provides a model where both BECs and fibroblasts have been shown to have immunosuppressive roles that promote tumor growth (Buckanovich *et al.*, 2008; Kraman *et al.*, 2010).

Whilst many tumors are immunogenic and therapeutic vaccines have been shown to induce adequate T cell responses, clinical efficacy of these responses has been limited (DuPage *et al.*, 2011). Endothelial cells microdissected from human ovarian tumors lacking T cell infiltration were shown to have upregulated Endothelin B receptor and blockade of this receptor led to increased tumor infiltration and increased efficacy of vaccination (Buckanovich *et al.*, 2008). Fibroblasts also have been shown to have effects on the ability of lymphocytes to infiltrate tumors. Fibroblasts activated in the tumor microenvironment express matrix components, proteases and other ECM modifying enzymes altering the mechanical properties of the tissue (Kalluri and Zeisberg, 2006; Levental *et al.*, 2009). Using real time imaging of slices of human lung tumors Salmon *et al.* showed that dense collagen fibers aligned around tumor nests excluded T cells from the cancer cells and that disrupting these structures allowed lymphocytes to infiltrate these structures (Salmon *et al.*, 2012). Recent studies have also identified CXCL12 as having a role in the exclusion of CD8 + T cells from tumors and this could be overcome by administering AMD3100, a CXCR4 antagonist (Feig *et al.*, 2013). Fibroblasts in LLC tumors and autochthonous pancreatic tumors were found to be the major source of CXCL12 indicating that fibroblasts can, similarly to BECs, assume a phenotype excluding T cells from the tumor (Feig *et al.*, 2013). Interestingly, despite this apparent role in immune cell exclusion when CAFs, isolated by their expression of platelet-derived growth factor receptor- α (PDGFR α), were isolated from skin carcinomas, pancreatic and mammary tumors and their transcriptome analyzed by microarray whilst they were found to be relatively heterogeneous, a pro-inflammatory gene signature did emerge (Erez *et al.*, 2010). This indicates that CAFs drive tumor promoting inflammation whilst excluding the potentially tumoricidal tumor-reactive lymphocytes. In fact CAFs have been shown to express monocyte chemotactic protein 1 (MCP-1) and that not only recruits monocytes but also biases their differentiation toward the pro-tumor M2 tumor-associated macrophage (TAM) phenotype (Comito *et al.*, 2014). The pro-inflammatory gene signature observed by Erez *et al.* could be induced in dermal fibroblasts by coculturing them with conditioned media from carcinoma cells indicating a close cross regulation between carcinoma cells, CAFs and TAMs (Erez *et al.*, 2010). Although this signature was present, it varied between PDGFR α + cells from different sources, and these may not represent the total pool of fibroblasts. Indeed the heterogeneity in tumor fibroblast populations may not simply be due to the heterogeneity observed in tissue resident fibroblasts previously alluded to (Chang *et al.*, 2002), but may also be due to heterogeneous origins of different CAF populations. Kidd *et al.* recently looked at different mesenchymal subsets within the tumor microenvironment and characterized their origins. They found that some cells marked by fibroblast specific protein 1 (FSP-1) originated from the bone marrow whilst others such as pericytes appeared to originate from adjacent tissue (Kidd *et al.*, 2012). As such it is again important to bear in mind that the markers used in individual studies will affect the populations of fibroblasts being examined.

BECs and fibroblasts also have important roles in driving pathological immune responses such as in rheumatoid

arthritis (RA). BECs are persistently activated in the inflammatory milieu of the RA synovium and express high levels of adhesion molecules and chemokines (Koch *et al.*, 1991). However, the changes seen in BECs in RA appear to be similar to those occurring in inflammatory conditions, however, these become persistent and maladaptive when the inflammatory environment persists and BECs continue to recruit further inflammatory cells. Fibroblasts within the synovium can also become activated in similar ways to those mentioned earlier, by TLR stimulation or in response to leukocyte-derived signals such as cytokines or microparticles, and they may then express a multitude of pro-inflammatory cytokines and chemokines, including IL-6, type I IFN, BAFF, which promote the inflammatory process (Benito-Miguel *et al.*, 2012). However, in RA and other chronic inflammatory diseases fibroblasts do not stop producing these factors and as such the damaging inflammatory process continues (Buckley *et al.*, 2004). Indeed implantation of diseased cartilage in severe combined immunodeficiency (SCID) mice showed that these activated fibroblasts could travel through the vasculature and spread to unaffected healthy cartilage implanted at a distal site potentially explaining the tendency of RA to spread from affected to previously unaffected joints indicating a critical orchestrating role for stromal cells in this system (Lefevre *et al.*, 2009). Indeed in many chronic inflammatory conditions such as rheumatoid arthritis, colitis, systemic fibrosis, and kidney fibrosis it appears that there is epigenetic modulation of the fibroblast cells giving them a more stable pro-inflammatory phenotype (Bechtel *et al.*, 2010). This may allow cells which do migrate to spread the inflammatory state even in the absence of new TLR stimulation or lymphocyte-derived signals. Whether the activation of the stromal cells drives or merely contributes to the disease progression is, again, unknown although fibroblasts remain an attractive therapeutic target due to their pro-inflammatory profile.

Targeting Stroma

With so many roles in inflammation it is tempting to view stromal cells as attractive therapeutic targets. As stated, infusion of ex vivo-expanded fibroblast populations, either FRCs or MSCs, have been shown to have positive effects in many inflammatory conditions (Fletcher *et al.*, 2014; Yanez *et al.*, 2006). Strategies able to modulate the pro-immune functions of fibroblasts at sites of chronic inflammation would potentially allow resolution of autoimmune lesions whilst skewing CAFs away from their immune suppressive phenotype could have benefits in cancer therapy, and indeed could complement current checkpoint blockade therapies (Feig *et al.*, 2013). In order to capitalize on these promising therapeutic directions, however, more work is needed to understand the heterogeneity of the populations being studied. Furthermore, if we are to avoid off-target effects it is important to understand the roles played by stromal cells in other tissues, which could be affected by therapeutic interventions and their relationships to the pathologic stromal cells being targeted. Previous work suggested that fibroblasts expressing FAP were restricted to sites of wound healing and cancer, however, ablation of these cells experimentally led to cachexia and anemia (Roberts *et al.*, 2013; Tran *et al.*, 2013). It is possible that even with better

understanding of markers of stromal populations such blunt approaches, targeting a single marker, will be too nonspecific to provide direct benefit. Future work identifying pathways important for the immune modulatory roles of fibroblasts is also important and has shown promising results in cancer therapy (Feig *et al.*, 2013).

See also: Cellular Immunology: Cell–Cell Interactions and the Immune System: Lymphocyte–Endothelial Interactions

References

- Abbott, N.J., Ronnback, L., Hansson, E., 2006. Astrocyte-endothelial interactions at the blood-brain barrier. *Nature Reviews Neuroscience* 7, 41–53.
- Allen, C.D., Cyster, J.G., 2008. Follicular dendritic cell networks of primary follicles and germinal centers: Phenotype and function. *Seminars in Immunology* 20, 14–25.
- Andonegui, G., Bonder, C.S., Green, F., *et al.*, 2003. Endothelium-derived Toll-like receptor-4 is the key molecule in LPS-induced neutrophil sequestration into lungs. *Journal of Clinical Investigation* 111, 1011–1020.
- Ansel, K.M., Mcheyzer-Williams, L.J., Ngo, V.N., Mcheyzer-Williams, M.G., Cyster, J.G., 1999. *In vivo*-activated CD4 T cells upregulate CXC chemokine receptor 5 and reprogram their response to lymphoid chemokines. *Journal of Experimental Medicine* 190, 1123–1134.
- Astorri, E., Bombardieri, M., Gabba, S., *et al.*, 2010. Evolution of ectopic lymphoid neogenesis and *in situ* autoantibody production in autoimmune nonobese diabetic mice: Cellular and molecular characterization of tertiary lymphoid structures in pancreatic islets. *Journal of Immunology* 185, 3359–3368.
- Aubin, J.E., 2001. Regulation of osteoblast formation and function. *Reviews in Endocrine and Metabolic Disorders* 2, 81–94.
- Austrup, F., Vestweber, D., Borges, E., *et al.*, 1997. P- and E-selectin mediate recruitment of T-helper-1 but not T-helper-2 cells into inflamed tissues. *Nature* 385, 81–83.
- Bajenoff, M., Egen, J.G., Koo, L.Y., *et al.*, 2006. Stromal cell networks regulate lymphocyte entry, migration, and territoriality in lymph nodes. *Immunity* 25, 989–1001.
- Bechtel, W., McGoochan, S., Zeisberg, E.M., *et al.*, 2010. Methylation determines fibroblast activation and fibrogenesis in the kidney. *Nature Medicine* 16, 544–550.
- Benito-Miguel, M., Garcia-Carmona, Y., Balsa, A., *et al.*, 2012. IL-15 expression on RA synovial fibroblasts promotes B cell survival. *PLoS One* 7, e40620.
- Buckanovich, R.J., Facciabene, A., Kim, S., *et al.*, 2008. Endothelin B receptor mediates the endothelial barrier to T cell homing to tumors and disables immune therapy. *Nature Medicine* 14, 28–36.
- Buckley, C.D., Filer, A., Haworth, O., Parsonage, G., Salmon, M., 2004. Defining a role for fibroblasts in the persistence of chronic inflammatory joint disease. *Annals of the Rheumatic Diseases* 63 (Suppl. 2), ii92–ii95.
- Buckley, C.D., Pilling, D., Lord, J.M., *et al.*, 2001. Fibroblasts regulate the switch from acute resolving to chronic persistent inflammation. *Trends in Immunology* 22, 199–204.
- Castagnaro, L., Lenti, E., Maruzzelli, S., *et al.*, 2013. Nkx2-5(+) islet1(+) mesenchymal precursors generate distinct spleen stromal cell subsets and participate in restoring stromal network integrity. *Immunity* 38, 782–791.
- Chai, Q., Onder, L., Scandella, E., *et al.*, 2013. Maturation of lymph node fibroblastic reticular cells from myofibroblastic precursors is critical for antiviral immunity. *Immunity* 38, 1013–1024.
- Chang, H.Y., Chi, J.T., Dudoit, S., *et al.*, 2002. Diversity, topographic differentiation, and positional memory in human fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America* 99, 12877–12882.
- Cohen, J.N., Guidi, C.J., Tewalt, E.F., *et al.*, 2010. Lymph node-resident lymphatic endothelial cells mediate peripheral tolerance via Aire-independent direct antigen presentation. *Journal of Experimental Medicine* 207, 681–688.
- Cornito, G., Giannoni, E., Segura, C.P., *et al.*, 2014. Cancer-associated fibroblasts and M2-polarized macrophages synergize during prostate carcinoma progression. *Oncogene* 33, 2423–2431.
- Cremasco, V., Woodruff, M.C., Onder, L., *et al.*, 2014. B cell homeostasis and follicle confines are governed by fibroblastic reticular cells. *Nature Immunology* 15, 973–981.

- Denton, A.E., Roberts, E.W., Linterman, M.A., Fearon, D.T., 2014. Fibroblastic reticular cells of the lymph node are required for retention of resting but not activated CD8 + T cells. *Proceedings of the National Academy of Sciences of the United States of America* 111, 12139–12144.
- Dubrot, J., Duraes, F.V., Potin, L., *et al.*, 2014. Lymph node stromal cells acquire peptide-MHCII complexes from dendritic cells and induce antigen-specific CD4 (+) T cell tolerance. *Journal of Experimental Medicine* 211, 1153–1166.
- DuPage, M., Cheung, A.F., Mazumdar, C., *et al.*, 2011. Endogenous T cell responses to antigens expressed in lung adenocarcinomas delay malignant tumor progression. *Cancer Cell* 19, 72–85.
- Eming, S.A., Krieg, T., Davidson, J.M., 2007. Inflammation in wound repair: Molecular and cellular mechanisms. *Journal of Investigative Dermatology* 127, 514–525.
- Erez, N., Truitt, M., Olson, P., Arron, S.T., Hanahan, D., 2010. Cancer-associated fibroblasts are activated in incipient neoplasia to orchestrate tumor-promoting inflammation in an NF-kappaB-dependent manner. *Cancer Cell* 17, 135–147.
- Feig, C., Jones, J.O., Kraman, M., *et al.*, 2013. Targeting CXCL12 from FAP-expressing carcinoma-associated fibroblasts synergizes with anti-PD-L1 immunotherapy in pancreatic cancer. *Proceedings of the National Academy of Sciences of the United States of America* 110, 20212–20217.
- Fletcher, A.L., Elman, J.S., Astarita, J., *et al.*, 2014. Lymph node fibroblastic reticular cell transplants show robust therapeutic efficacy in high-mortality murine sepsis. *Science Translational Medicine* 6, 249ra109.
- Fletcher, A.L., Lukacs-Kornek, V., Reynoso, E.D., *et al.*, 2010. Lymph node fibroblastic reticular cells directly present peripheral tissue antigen under steady-state and inflammatory conditions. *Journal of Experimental Medicine* 207, 689–697.
- Forster, R., Schubel, A., Breitfeld, D., *et al.*, 1999. CCR7 coordinates the primary immune response by establishing functional microenvironments in secondary lymphoid organs. *Cell* 99, 23–33.
- Fu, H., Kishore, M., Gittens, B., *et al.*, 2014. Self-recognition of the endothelium enables regulatory T-cell trafficking and defines the kinetics of immune regulation. *Nature Communications* 5, 3436.
- Gardner, J.M., Devoss, J.J., Friedman, R.S., *et al.*, 2008. Deletional tolerance mediated by extrathymic Aire-expressing cells. *Science* 321, 843–847.
- Girard, J.P., Moussion, C., Forster, R., 2012. HEVs, lymphatics and homeostatic immune cell trafficking in lymph nodes. *Nature Reviews Immunology* 12, 762–773.
- Grigороva, I.L., Schwab, S.R., Phan, T.G., *et al.*, 2009. Cortical sinus probing, S1P1-dependent entry and flow-based capture of egressing T cells. *Nature Immunology* 10, 58–65.
- Groh, M.E., Maitra, B., Szekely, E., Koc, O.N., 2005. Human mesenchymal stem cells require monocyte-mediated activation to suppress alloreactive T cells. *Experimental Hematology* 33, 928–934.
- Hammerschmidt, S.I., Ahrendt, M., Bode, U., *et al.*, 2008. Stromal mesenteric lymph node cells are essential for the generation of gut-homing T cells *in vivo*. *Journal of Experimental Medicine* 205, 2483–2490.
- Jarjour, M., Jorquera, A., Mondor, I., *et al.*, 2014. Fate mapping reveals origin and dynamics of lymph node follicular dendritic cells. *Journal of Experimental Medicine* 211, 1109–1122.
- Kalluri, R., Zeisberg, M., 2006. Fibroblasts in cancer. *Nature Reviews Cancer* 6, 392–401.
- Karlsson, H., Samarasinghe, S., Ball, L.M., *et al.*, 2008. Mesenchymal stem cells exert differential effects on alloantigen and virus-specific T-cell responses. *Blood* 112, 532–541.
- Khan, O., Headley, M., Gerard, A., *et al.*, 2011. Regulation of T cell priming by lymphoid stroma. *PLoS One* 6, e26138.
- Khayamian, S., Hutloff, A., Buchner, K., *et al.*, 2002. ICOS-ligand, expressed on human endothelial cells, costimulates Th1 and Th2 cytokine secretion by memory CD4 + T cells. *Proceedings of the National Academy of Sciences of the United States of America* 99, 6198–6203.
- Kidd, S., Spaeth, E., Watson, K., *et al.*, 2012. Origins of the tumor microenvironment: Quantitative assessment of adipose-derived and bone marrow-derived stroma. *PLoS One* 7, e30563.
- Koch, A.E., Burrows, J.C., Haines, G.K., *et al.*, 1991. Immunolocalization of endothelial and leukocyte adhesion molecules in human rheumatoid and osteoarthritic synovial tissues. *Laboratory Investigation* 64, 313–320.
- Kraman, M., Bambrough, P.J., Arnold, J.N., *et al.*, 2010. Suppression of antitumor immunity by stromal cells expressing fibroblast activation protein- α . *Science* 330, 827–830.
- Kumar, S., West, D.C., Ager, A., 1987. Heterogeneity in endothelial cells from large vessels and microvessels. *Differentiation* 36, 57–70.
- Lefevre, S., Knedla, A., Tennie, C., *et al.*, 2009. Synovial fibroblasts spread rheumatoid arthritis to unaffected joints. *Nature Medicine* 15, 1414–1420.
- Levental, K.R., Yu, H., Kass, L., *et al.*, 2009. Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell* 139, 891–906.
- Link, A., Vogt, T.K., Favre, S., *et al.*, 2007. Fibroblastic reticular cells in lymph nodes regulate the homeostasis of naive T cells. *Nature Immunology* 8, 1255–1265.
- Linterman, M.A., Liston, A., Vinuesa, C.G., 2012. T-follicular helper cell differentiation and the co-option of this pathway by non-helper cells. *Immunological Reviews* 247, 143–159.
- Lochner, M., Ohnmacht, C., Presley, L., *et al.*, 2011. Microbiota-induced tertiary lymphoid tissues aggravate inflammatory disease in the absence of RORgamma t and LTi cells. *Journal of Experimental Medicine* 208, 125–134.
- Lukacs-Kornek, V., Malhotra, D., Fletcher, A.L., *et al.*, 2011. Regulated release of nitric oxide by nonhematopoietic stroma controls expansion of the activated T cell pool in lymph nodes. *Nature Immunology* 12, 1096–1104.
- Maggini, J., Mirkin, G., Bognanni, I., *et al.*, 2010. Mouse bone marrow-derived mesenchymal stromal cells turn activated macrophages into a regulatory-like profile. *PLoS One* 5, e9252.
- Mai, J., Virtue, A., Shen, J., Wang, H., Yang, X.F., 2013. An evolving new paradigm: Endothelial cells—conditional innate immune cells. *Journal of Hematology & Oncology* 6, 61.
- Malhotra, D., Fletcher, A.L., Astarita, J., *et al.*, 2012. Transcriptional profiling of stroma from inflamed and resting lymph nodes defines immunological hallmarks. *Nature Immunology* 13, 499–510.
- Manes, T.D., Pober, J.S., 2008. Antigen presentation by human microvascular endothelial cells triggers ICAM-1-dependent transendothelial protrusion by, and fractalkine-dependent transendothelial migration of, effector memory CD4 + T cells. *Journal of Immunology* 180, 8386–8392.
- Matloubian, M., Lo, C.G., Cinamon, G., *et al.*, 2004. Lymphocyte egress from thymus and peripheral lymphoid organs is dependent on S1P receptor 1. *Nature* 427, 355–360.
- Midwood, K., Sacre, S., Piccinini, A.M., *et al.*, 2009. Tenascin-C is an endogenous activator of Toll-like receptor 4 that is essential for maintaining inflammation in arthritic joint disease. *Nature Medicine* 15, 774–780.
- Mionnet, C., Mondor, I., Jorquera, A., *et al.*, 2013. Identification of a new stromal cell type involved in the regulation of inflamed B cell follicles. *PLoS Biology* 11, e1001672.
- Mori, S., Nakano, H., Aritomi, K., *et al.*, 2001. Mice lacking expression of the chemokines CCL21-ser and CCL19 (plt mice) demonstrate delayed but enhanced T cell immune responses. *Journal of Experimental Medicine* 193, 207–218.
- Moyron-Quiroz, J.E., Rangel-Moreno, J., Hartson, L., *et al.*, 2006. Persistence and responsiveness of immunologic memory in the absence of secondary lymphoid organs. *Immunity* 25, 643–654.
- Moyron-Quiroz, J.E., Rangel-Moreno, J., Kusser, K., *et al.*, 2004. Role of inducible bronchus associated lymphoid tissue (iBALT) in respiratory immunity. *Nature Medicine* 10, 927–934.
- Mueller, S.N., Hosiawa-Meagher, K.A., Konieczny, B.T., *et al.*, 2007. Regulation of homeostatic chemokine expression and cell trafficking during immune responses. *Science* 317, 670–674.
- Nakano, H., Lin, K.L., Yanagita, M., *et al.*, 2009. Blood-derived inflammatory dendritic cells in lymph nodes stimulate acute T helper type 1 immune responses. *Nature Immunology* 10, 394–402.
- Nakano, H., Tamura, T., Yoshimoto, T., *et al.*, 1997. Genetic defect in T lymphocyte-specific homing into peripheral lymph nodes. *European Journal of Immunology* 27, 215–221.
- Neyt, K., Perros, F., Geurtsvankessel, C.H., Hammad, H., Lambrecht, B.N., 2012. Tertiary lymphoid organs in infection and autoimmunity. *Trends in Immunology* 33, 297–305.
- von Oppen, N., Schurich, A., Hegenbarth, S., *et al.*, 2009. Systemic antigen cross-presented by liver sinusoidal endothelial cells induces liver-specific CD8 T-cell retention and tolerization. *Hepatology* 49, 1664–1672.
- Peister, A., Mellad, J.A., Larson, B.L., *et al.*, 2004. Adult stem cells from bone marrow (MSCs) isolated from different strains of inbred mice vary in surface epitopes, rates of proliferation, and differentiation potential. *Blood* 103, 1662–1668.
- Roberts, E.W., Deonaraine, A., Jones, J.O., *et al.*, 2013. Depletion of stromal cells expressing fibroblast activation protein- α from skeletal muscle and bone marrow results in cachexia and anemia. *Journal of Experimental Medicine* 210, 1137–1151.
- Salmon, H., Franciszewicz, K., Damotte, D., *et al.*, 2012. Matrix architecture defines the preferential localization and migration of T cells into the stroma of human lung tumors. *Journal of Clinical Investigation* 122, 899–910.
- Siebert, S., Huang, H.Y., Yang, C.Y., *et al.*, 2011. Fibroblastic reticular cells from lymph nodes attenuate T cell expansion by producing nitric oxide. *PLoS One* 6, e27618.

- Sixt, M., Kanazawa, N., Selg, M., *et al.*, 2005. The conduit system transports soluble antigens from the afferent lymph to resident dendritic cells in the T cell area of the lymph node. *Immunity* 22, 19–29.
- Sorokin, L., 2010. The impact of the extracellular matrix on inflammation. *Nature Reviews Immunology* 10, 712–723.
- Tran, E., Chinnasamy, D., Yu, Z., *et al.*, 2013. Immune targeting of fibroblast activation protein triggers recognition of multipotent bone marrow stromal cells and cachexia. *Journal of Experimental Medicine* 210, 1125–1135.
- Wang, X., Cho, B., Suzuki, K., *et al.*, 2011. Follicular dendritic cells help establish follicle identity and promote B cell retention in germinal centers. *Journal of Experimental Medicine* 208, 2497–2510.
- Weber, M., Hauschild, R., Schwarz, J., *et al.*, 2013. Interstitial dendritic cell guidance by haptotactic chemokine gradients. *Science* 339, 328–332.
- Wiley, J.A., Richert, L.E., Swain, S.D., *et al.*, 2009. Inducible bronchus-associated lymphoid tissue elicited by a protein cage nanoparticle enhances protection in mice against diverse respiratory viruses. *PLoS One* 4, e7142.
- von Willebrand, E., Lautenschlager, I., Inkinen, K., *et al.*, 1985. Distribution of the major histocompatibility complex antigens in human and rat kidney. *Kidney International* 27, 616–621.
- Yanez, R., Lamana, M.L., Garcia-Castro, J., *et al.*, 2006. Adipose tissue-derived mesenchymal stem cells have *in vivo* immunosuppressive properties applicable for the control of the graft-versus-host disease. *Stem Cells* 24, 2582–2591.
- Yang, C.Y., Vogt, T.K., Favre, S., *et al.*, 2014. Trapping of naive lymphocytes triggers rapid growth and remodeling of the fibroblast network in reactive murine lymph nodes. *Proceedings of the National Academy of Sciences of the United States of America* 111, E109–E118.

CXCL12/SDF-1 and Hematopoiesis

ML Capitano and HE Broxmeyer, Indiana University School of Medicine, Indianapolis, IN, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Chemokines, the largest family of cytokines, are named for their chemoattractant activities. Defined by structure not function, chemokines have a highly conserved tertiary structure due to their conservatively spaced disulfide-bonded cysteines. Chemokines are classified by numbers and location of cysteine in their amino acid sequence, with subfamilies distinguished by CXC, CC, XC, and CX3C motifs followed by the letter 'L' and a number. They can be further classified by additional motifs such as the Glu-Leu-Arg (ELR) N-terminal motif of the CXC subfamily or the position of the sixth cysteine found either in the C-terminal domain or between C-2 and C-3 of the CC subfamily (all of which have 6 cysteines). Chemokines can exist as monomers, dimers, tetramers, and even complex quaternary structures which occur when they bind to glycosaminoglycans (GAGs), carbohydrates containing amino sugars, on the cell surface. Tethering to the extracellular matrix (ECM) or GAGs on the cell surface (except in the case of CX3CL1/fractalkine and CXCL16 which have transmembrane domains), chemokines bind to chemokine receptors. Chemokine receptors are members of the seven-transmembrane domain superfamily of G protein-coupled receptors; their function stems from a single polypeptide chain which is responsible for chemokine binding, membrane anchoring, and carrying signaling domains. Chemokine receptor nomenclature is mostly determined by the ligand subfamily (e.g., CXC or CC) followed by R and a number (Bacon *et al.*, 2002; Baggiolini *et al.*, 1997; Murphy, 1996).

Chemokines coordinate hematopoietic cell trafficking covering the range of mature leukocytes and hematopoietic stem cell (HSC) and progenitor cell (HPC) populations. The function of chemokines within these cell populations allows for further classification into two more defined chemokine systems: the homeostatic and inflammatory systems (Le *et al.*, 2004). The homeostatic system of chemokines regulates hematopoiesis, the process by which HSCs self-renew and differentiate to form the peripheral blood lineages, and coordinate immune surveillance, accomplished through constitutively expressed chemokine ligands and receptors within specific tissues. Examples of homeostatic chemokine receptors include: CXCR4 (a receptor for CXCL12) on hematopoietic stem and progenitor cells, CXCR5 on naïve B cells, CCR7 on mature dendritic cells and naïve T cells, and gut and skin-specific T cell homing receptors CCR9 and CCR10, respectively. The inflammatory system is characterized by the ability of 'danger signals' (produced, for example, following tissue damage or infection) to regulate chemokine and chemokine receptor expression where chemokines are inducible and chemokine receptors can be either inducible (in adaptive immunity) or constitutively expressed (in innate immunity).

Here we examine roles chemokines play within the bone marrow (BM) microenvironment focusing primarily on the

role of the homeostatic chemokine CXCL12 and one of its receptors, CXCR4, on regulating HSC/HPC homing to and maintenance within their preferred niches in the BM. We also examine roles of CXCL12 in controlling neutrophil and B cell production within BM.

Role of CXCL12 in Homing to BM

HSCs and HPCs circulate through the bloodstream and extravasate into the BM. The ability of HSCs to home to the BM was first demonstrated in mice whose spleen was shielded from a dose of lethal total body irradiation. These mice still exhibited hematopoietic recovery suggesting that repopulating cells residing within the spleen could home to the BM (Jacobson *et al.*, 1949; Lorenz *et al.*, 1951); but how did these cells know where to go? Like most of their mature counterparts, homing of HSCs and HPCs to a specific tissue site requires the process of cell tethering/rolling along the vascular endothelial as controlled by the interaction of selectins and carbohydrate ligands, the activation of the cell by chemokine binding to G protein-coupled receptors to mediate firm adhesion to the vascular wall by interactions of integrins with adhesion molecules of the Ig superfamily, and by the extravasation of the cell through the basal lamina ECM mediated via chemokine signaling (Muller, 2014).

As HSCs and HPCs circulate through the blood, they encounter and interact with BM vascular endothelial cells of BM venous sinuses. The sinus wall is composed of endothelial cells, basement membrane and a layer of adventitial cells (Inoue and Osmond, 2001). Within the endothelial cell layer are multiple sites in which abluminal and luminal membranes are fused, forming diaphragmed fenestrae (Burdon *et al.*, 2008; Levesque *et al.*, 2002). It is at this site that HSCs and HPCs are believed to migrate through the blood-BM barrier. BM sinusoidal endothelial cells constitutively express vascular cell adhesion molecule-1 (VCAM-1), E- and P-selectins (Frenette *et al.*, 1996, 1998; Mazo *et al.*, 1998; Miyake *et al.*, 1991; Papayannopoulou *et al.*, 1995). When HSCs and HPCs expressing P-selectin glycoprotein ligand-1 (PSGL-1) and high levels of very late antigen-4 (VLA-4/integrin $\alpha 4 \beta 1$) interact with VCAM-1, E- and P-selectin, they are able to roll along the vascular endothelial layer until they encounter the proper chemokine signal. CXCL12 is responsible for activation of HSCs and HPCs allowing for firm, shear-resistant adhesion to BM endothelium, although stem cell factor (SCF) and FMS-like tyrosine kinase-3 (FLT-3) can synergize enhancing migration. Interaction of BM endothelial CXCL12 with CXCR4 on HSC or HPC surfaces mediate activation of VLA-4, VLA-5 (integrin $\alpha 5 \beta 1$) and lymphocyte function-associated antigen-1 (LFA-1/integrin $\alpha L \beta 2$) via the PI-PLC and PI3-K pathways. This leads to ERK and Pyk2 activation, allowing for firm adhesion to take place (Bonig *et al.*, 2008; Petit *et al.*, 2005).

Integrins exist on the surface of HSC/HPC in a low-affinity, 'bent' conformation which upon activation by CXCL12: CXCR4 signaling alters integrin conformation allowing for integrin binding to its ligand so that the cell can extravasate and migrate into the BM. Adhesion molecules believed to be involved in the extravasation process include platelet-endothelial cell adhesion molecule-1 (PCAM-1/CD31), CD99, CD31, VLA-4, VLA-5, and VLA-6 (Imbert *et al.*, 2006; Yong *et al.*, 1998).

There is substantial evidence for the importance of CXCL12: CXCR4 interaction in HSC and HPC migration. For example, CXCR4^{-/-} and CXCL12^{-/-} mice demonstrate profound defects in the ability of fetal liver HSCs and HPCs to migrate/populate BM suggesting that CXCL12: CXCR4 signaling is important in marrow colonization and, therefore, in homing of these cells to the proper BM microenvironment during development (Ma *et al.*, 1998; Nagasawa *et al.*, 1996; Zou *et al.*, 1998). In adult models, CXCL12: CXCR4 plays a major role in homing to and mobilization from BM of HSCs and HPCs. Blocking CXCR4 function with neutralizing anti-CXCR4 antibodies prevents hematopoietic recovery following transplantation of human CD34⁺ cells in non-obese diabetic/severe combined immunodeficient (NOD/SCID) mice (Peled *et al.*, 1999). Overexpression of CXCR4 on human CD34⁺ cells enhanced the migration of these cells into NOD/SCID mice thus enhancing hematopoietic recovery following transplantation (Broxmeyer *et al.*, 2003; Kahn *et al.*, 2004).

Since HSC and HPC homing to BM following transplantation is a critical step in engraftment, CXCL12 signaling modulators have been utilized. One CXCL12 modulator is the enzyme CD26/dipeptidylpeptidase IV (DPP4). CD26 is expressed on surfaces of multiple cell types, and functions as a cell surface serine protease that selectively cleaves the N-terminal penultimate proline or alanine amino acids of various growth factors and cytokines including CXCL12 (Broxmeyer, 2008; O'Leary *et al.*, 2013; Ou *et al.*, 2013). Inhibition of DPP4 with diprotin A (Ile-Pro-Ile peptide) enhances chemotaxis of both human and mouse HSCs/HPCs *in vitro* (Christopherson *et al.*, 2002). *In vivo* pretreatment of recipient mice with diprotin A enhances engraftment of donor HSCs and HPCs following transplantation further supporting the role of CD26 as a modulator of CXCL12-mediated migration/homing to BM (Christopherson *et al.*, 2004). Another method to enhance migration of HSCs and HPCs towards a CXCL12 gradient is by priming the cells to respond more robustly to lower concentrations of CXCL12 (Levesque *et al.*, 2010; Moll and Ransohoff, 2010). Fibrinogen, fibronectin, complement C1q, complement C3a, platelet-derived microvesicles, hyaluronic acid, thrombin, and bioreactive lipids prime HSCs and HPCs to better migrate toward low concentrations of CXCL12. This is thought to occur by recruiting CXCR4 into specialized plasma membrane domains called lipid rafts which allows for the receptor to be in closer proximity to downstream signaling components. The use of these 'primers' is also associated with a highly proteolytic microenvironment due to increased matrix metalloproteinase (MMP)-2 and MMP-9 expression. These results, amongst others, strongly suggest that CXCL12: CXCR4 signaling is required for optimal migration of HSCs and HPCs into BM; however, once in the BM CXCL12 also plays a role.

Role of CXCL12 in HSC and HPC Maintenance within BM Niches

Within the BM microenvironment exist regions enriched with CXCL12 where HSCs home to, are retained in, and which help to maintain their self-renewal and differentiation potential through inhibiting G₀→G₁ cell cycle progression and by increasing the survival of HSCs and HPCs (Broxmeyer, 2001, 2008; Broxmeyer *et al.*, 2003; Mendelson and Frenette, 2014). HSCs and HPCs that home to the BM microenvironment through endothelial sinusoids from blood must navigate the ECM and lodge in marrow parenchyma. Cellular and extracellular signals provided within CXCL12-enriched regions define the potential regulatory function of that particular niche. The two primary HSC niches are the perivascular niche and the osteoblastic or endosteal niche which is located at the bone surface (Mendelson and Frenette, 2014). There is some disagreement as to the role of each niche in the maintenance of HSCs and HPCs with some believing that the endosteal region of the BM is more important in maintaining quiescent HSCs, while others suggest that the perivascular region of the BM plays a larger role in HSC maintenance (Levesque *et al.*, 2010; Mendelson and Frenette, 2014; Moll and Ransohoff, 2010).

Recent studies have attempted to define the cells responsible in maintaining HSC and HPC numbers and function within BM, thus hinting at the importance of each particular niche. It is clear that CXCL12 is essential in maintaining and retaining HSCs and HPCs within the BM. Key cell types were found to play an important role in producing or regulating CXCL12 in the BM. This includes perivascular mesenchymal stem and progenitor cells (MSPCs), perivascular stromal cells (e.g., CXCL12-abundant reticular or CAR cells), endothelial cells, hematopoietic cells (e.g., macrophages), and sympathetic neurons (Ding and Morrison, 2013; Greenbaum *et al.*, 2013; Mendelson and Frenette, 2014). Perivascular MSPCs/stromal cells express the highest concentrations of CXCL12 followed by osteoprogenitors, endothelial cells, osteoblasts, and finally hematopoietic cells. Animal models in which specific HSC/HPC niche cells have conditionally deleted *Cxcl12* were utilized to determine the role of these different CXCL12-producing niche components on HSC and HPC maintenance and retention.

To examine perivascular MSPCs and stromal cells, the genetic models *Prx1-cre*, *Nestin-Gfp*, and *Lepr-cre* were utilized. *Cxcl12* deletion in nestin-negative mesenchymal progenitors using *Prx1-cre* mice demonstrated a significant loss of HSC and HPC numbers in BM associated with increased HSC number in spleen, decreased long-term repopulating activity and loss of HSC quiescence (Ding and Morrison, 2013; Greenbaum *et al.*, 2013). Deletion of *Cxcl12* in *Lepr-cre* cells significantly induced mobilization of HSCs and HPCs to peripheral blood and spleen, but had no effect on HSC maintenance, suggesting that these cells play a role in HSC/HPC retention within the BM (Ding and Morrison, 2013; Ding *et al.*, 2012). Thus, CXCL12 production by MSPCs plays an important role in HSC and HPC retention, maintenance and self-renewal. Examining the role of CXCL12 production in endothelial cells by deleting *Cxcl12* in Tie2⁺ cells, demonstrated modest decreases in HSC numbers and repopulating capacity without changes to HPC numbers, suggesting that the CXCL12 produced by endothelial

cells may play a role in HSC maintenance but has little effect on HPCs. Upon examining the role of CXCL12 production by osteoprogenitors (Osterix-*cre* mouse model) and osteoblasts (Col2.3-*cre* mouse model) by deleting *Cxcl12*, it was discovered that CXCL12 from osteoprogenitors and osteoblast had no role on HSC maintenance (Ding and Morrison, 2013). However, both cell types assist in maintaining HPC numbers and helped with HPC retention within the BM, suggesting that CXCL12 production by MSPCs, and to lesser extent endothelial cells, plays an important role in maintaining HSCs and that the osteoblastic lineage cells play a larger role in HPC maintenance.

Another method of regulating CXCL12 expression within the BM microenvironment is through the sympathetic nervous system. Sympathetic nerves are in close contact with nestin-expressing MSPCs. Sympathetic nerves create an adrenergic signal (which responds to circadian rhythms) that down-regulates CXCL12, Kit Ligand, and VCAM-1 gene expression, leading to HSC/HPC mobilization (Katayama *et al.*, 2006; Mendez-Ferrer *et al.*, 2010). Interestingly, macrophages within the BM microenvironment seem to have an antagonistic effect compared to adrenergic signals. Like the sympathetic neurons, macrophages are also in close contact with the nestin-expressing MSPCs and through receiving an unknown signal can induce CXCL12 expression (Chow *et al.*, 2011; Christopher *et al.*, 2011). Thus, both sympathetic nerves and macrophages regulate HSC/HPC retention within the BM.

CXCL12 in HSC and HPC Mobilization

Utilizing the knowledge that CXCL12 is an important homing signaling for HSC and HPC to migrate towards adult BM, studies were performed that either increased CXCL12 levels in the periphery or decreased the ability for CXCL12 to bind or signal through CXCR4. When CXCL12 levels in the peripheral circulation were upregulated through intravenous injection of an adenoviral vector expressing CXCL12 into SCID mice, there were significant increases in numbers of HPC (i.e., colony-forming units-granulocyte-macrophage) and increases in cells with stem cell potential (i.e., colony-forming units-spleen and long-term reconstituting HSCs) mobilized into the periphery (Hattori *et al.*, 2001). Mobilized human CD34⁺ cells also demonstrated a decrease in CXCR4 expression suggesting an important role in CXCL12: CXCR4 for regulating the mobilization of HSCs and HPCs (Dar *et al.*, 2006; Vagima *et al.*, 2011).

A major method of stimulating HSC/HPC mobilization is through the use of intravenous granulocyte colony-stimulating factor (G-CSF), the most widely used mobilization agent in the clinic. G-CSF expression can interrupt CXCL12 retention signals of these cells to remain within their niches. One possible mechanism in which this occurs is through the mediation of proteolytic inactivation of both CXCL12 and CXCR4 (Alvarez *et al.*, 2013; Broxmeyer, 2008; Mendelson and Frenette, 2014). Following activation by G-CSF, neutrophils release elastase and lactoferrin, allowing for further degranulation of neutrophils within the BM stroma. This increase in granulocytic proteases (including elastase and cathepsin) may be responsible for cleavage of important retention mediators such as

VCAM-1, CXCL12, CXCR4, and c-KIT receptor. In addition, increased concentrations of G-CSF was associated with increased expression of CD26 peptidase in CD34⁺CD38[−] human cord blood cell population which correlated with decreased ability of these cells to home towards CXCL12 (Christopherson *et al.*, 2002, 2006). Further supporting these findings, CD26^{−/−} mice demonstrated diminished responses to G-CSF-induced HSC/HPC mobilization, suggesting that CXCL12 cleavage by CD26/DPP4 may be a key mechanistic step (Christopherson *et al.*, 2003). G-CSF also enhances mobilization by inducing expression of plasminogen, a glycoprotein present in the blood plasma and other extracellular fluids that plays a role in dissolution of blood clots following thrombosis (Alvarez *et al.*, 2013; Broxmeyer, 2008; Mendelson and Frenette, 2014). Plasminogen is able to bind BM ECM, be converted into active protease plasmin, and then degrade the ECM. It can also activate other proteases including those within the MMP family including MMP-3,-9,-12, and -13 which can further degrade the ECM and collagen. Therefore, G-CSF plays a critical role in regulating proteases not only by inhibiting the ability of CXCL12 to signal through its receptor but also by enhancing the ability of HSCs/HPCs to migrate by enhancing degradation of collagen and the ECM.

G-CSF also functions to augment CXCL12 expression via indirect means. G-CSF under steady-state conditions and following therapeutic treatment is associated with decreased expression of CXCL12 mRNA in osteoblasts (Semerád *et al.*, 2005). However, it is unclear how this regulation occurs since the osteoblast does not express G-CSF receptors. *In vivo*, therapeutic treatment with G-CSF is associated with increased osteoblast apoptosis due to inhibition of osteoblast differentiation which may contribute to decreased CXCL12 secretion; however, whether this occurs under physiological conditions remains to be seen (Christopher and Link, 2008). *In vivo* studies suggest that macrophages and sympathetic neurons may play a role in G-CSF-mediated increases in osteoblast apoptosis. Following G-CSF exposure, there is an increase not just in osteoblast apoptosis but a decrease in macrophage numbers. This was associated with activation of sympathetic neurons to release norepinephrine in the BM microenvironment (Christopher *et al.*, 2011; Katayama *et al.*, 2006; Lucas *et al.*, 2012; Winkler *et al.*, 2010). Macrophages and sympathetic neurons express G-CSF receptors and by blocking either G-CSF receptor signaling or norepinephrine signaling HSC/HPC mobilization is inhibited, suggesting that macrophages and sympathetic neurons may be regulating osteoblasts. G-CSF signaling also leads to increased CXCL12 expression by BM endothelial cells and megakaryocytes (Eash *et al.*, 2010; Kohler *et al.*, 2011). Thus, this is another way in which G-CSF may regulate HSC and HPC numbers in the periphery.

HSC/HPC mobilization can also be stimulated through the use of CXCR4 antagonists and agonists. AMD3100, a small organic molecule consisting of two cyclam rings connected by 1,4-phenylenebis(methylene) linker, is an antagonist of CXCL12/CXCR4 signaling by binding to CXCR4 (Debnath *et al.*, 2013; Fricker, 2013). Blocking of CXCL12 ability to signal through CXCR4 results in an increased mobilization of HSCs and HPCs to the blood (Broxmeyer, 2008; Broxmeyer *et al.*, 2005). AMD3100 works synergistically with G-CSF to

mobilize HSCs and HPCs. HSCs and HPCs mobilized using both AMD3100 and G-CSF demonstrated increased expression of genes associated with better engraftment following transplantation when compared to cells mobilized with G-CSF alone. CTCE-0021, a CXCR4 agonist peptide, also enhances HSC/HPC mobilization; an effect believed to be mediated by the downregulation of CXCR4 expression on the cell surface (Pelus *et al.*, 2005).

Role of CXCL12 in Maintenance of Neutrophils in BM

CXCL12-depleted adult mice as well as CXCL12^{-/-} embryonic mice demonstrate that CXCL12 is essential in maintaining hematopoiesis (Ma *et al.*, 1998; Nagasawa *et al.*, 1996; Zou *et al.*, 1998). In both models, numbers of HPCs was severely reduced in BM. Furthermore, HPCs from mice that transgenically overexpress CXCL12 displayed enhanced hematopoiesis in both BM and spleen (Broxmeyer *et al.*, 2003). This was associated with increased HPC cycling. *In vitro* studies of HPCs from CXCL12 transgenic mice demonstrated enhanced survival, which was dependent upon CXCL12 signaling through its receptor CXCR4 since utilizing neutralizing antibodies against CXCR4 and AMD3100 blocked these effect of CXCL12 (Broxmeyer *et al.*, 2003). In addition, the use of pertussis toxin, which inhibits G α i protein-linked receptor signaling, also blocked effects of CXCL12 on HPC survival, offering further evidence for the importance of CXCL12 signaling on HPC maintenance (Broxmeyer *et al.*, 2003).

Although CXCL12 is important in maintaining hematopoiesis, the exact BM niche component that regulates this process is unclear. Decreasing numbers of macrophages and osteoblasts in BM decreased numbers of HPC suggesting that the endosteal region of BM is important in maintaining hematopoiesis. A relationship between the osteoblastic niche and hematopoiesis is supported by histological studies demonstrating that HPCs are in close proximity to endosteum (Ueda *et al.*, 2005, 2004). Furthermore, stimulation of osteoclasts with the cytokine RANKL reduced expression of important HPC survival and growth factors such as CXCL12, SCF, and osteopontin along the endosteum resulted in fewer HPCs in BM (Kollet *et al.*, 2006). However, these findings do not correlate with recent work demonstrating no changes in HPC maintenance or numbers when *Cxcl12* was deleted in osteoprogenitors or osteoblasts (Ding and Morrison, 2013; Greenbaum *et al.*, 2013). Further work needs to be performed to understand the roles niche cells play in regulating hematopoiesis.

CXCL12 is important not only in regulation of hematopoiesis, but in maintenance of neutrophil stores as well. Under homeostatic conditions, only 1–2% of total neutrophils at any given time are in the peripheral blood; the rest are stored in the BM (Broxmeyer *et al.*, 1974; Semerad *et al.*, 2002). During emergency situations (such as infection or tissue injury), neutrophil stores in the BM act as a large reservoir of neutrophils that can be released when needed. Like HSC and HPC, neutrophils utilize the venous sinuses as the primary migration site to and from BM. Neutrophils exit from BM and their entrance to the bloodstream is controlled by chemokine signaling from the local and systemic cytokine milieu, binding

and release from various adhesion molecules, and through regulation of protease activity.

Three members of the CXC family of chemokines are thought to control neutrophil egress from the BM: CXCL1 (KC), CXCL2 (MIP-2), and CXCL12. CXCL1 and CXCL2 are members of the ELR-subfamily of CXC chemokines and are expressed by osteoblasts in BM, megakaryocytes, macrophages, neutrophils, and endothelial cells. CXCL1 and CXCL2 bind to CXCR2 on the neutrophil surface causing rapid release of neutrophils into the circulation (Burdon *et al.*, 2005, 2008; Wengner *et al.*, 2008). This occurs under steady-state conditions to replenish neutrophil numbers and under stress conditions such as infection or tissue damage to increase numbers of circulating neutrophils. Neutrophil egress from BM under steady-state conditions is mediated by the VLA-4 on the neutrophil surface and by signaling through CXCR2. VLA-4 integrin engages VCAM-1 on the endothelial surface, stabilizing neutrophil adhesion, inducing neutrophil shape change, and supporting transmigration across the endothelial cell monolayer (Burdon *et al.*, 2005, 2008). Mice that have CXCR2-deficient neutrophils selectively retain neutrophils in the BM and exhibit neutropenia in the circulation, suggesting that signaling through CXCR2 plays a role in neutrophil egress from BM to peripheral blood (Eash *et al.*, 2010).

While in BM, neutrophils express low, but detectable, levels of surface CXCR4. When CXCL12 binds CXCR4 on the neutrophil cell surface, CXCR4 is internalized through a process of receptor recycling; therefore, cells have a large intracellular store of CXCR4 so that a constant concentration can remain on the cell surface during the time in which neutrophils remain in the BM (Martin *et al.*, 2003). CXCR4-deficiency, when specific to the myeloid lineage, causes neutrophilia in mice, as did conditional deletion of CXCL12 (Eash *et al.*, 2009). Thus, CXCL12:CXCR4 signaling is thought to be the key retention signal for neutrophils to remain in BM.

The CXCL12:CXCR4 retention signal and the CXCL1/2:CXCR2 release signal regulate neutrophil numbers in BM in two key sites: the vasculature endothelium (which secretes the highest concentrations of CXCL1 and CXCL2) and the BM stroma (which is where the majority of CXCL12 is produced) (Day and Link, 2012). Steady-state conditions favor retention signals, allowing for large stores of neutrophils to remain within the marrow. Utilizing CXCR2 knockout, CXCR4 knockout and double knockout mice, it was shown that CXCR4 signaling is more dominant than CXCR2 signaling (Eash *et al.*, 2010, 2009). However, *in vitro* treatment of neutrophils with CXCL1 or CXCL2 results in downregulation of cell surface CXCR4 (Eash *et al.*, 2009; Richardson *et al.*, 2003; Suratt *et al.*, 2004). Therefore, enough CXCL1 or CXCL2 to significantly reduce CXCR4 expression on neutrophils favor neutrophil release from BM.

G-CSF signaling, under both steady-state conditions and in response to therapeutic treatment, also sways the balance in favor of release by regulating CXCL12:CXCR4 signaling. As noted earlier, G-CSF under steady-state conditions and following therapeutic treatment is known to directly regulate the function of CXCL12 by inducing proteolytic cleavage and by indirectly regulating CXCL12 secretion by osteoblasts (Christopher and Link, 2008; Semerad *et al.*, 2005). G-CSF treatment *in vitro* and *in vivo* is associated with downregulation

of neutrophil CXCR4 (De La Luz Sierra *et al.*, 2007). This is accomplished by G-CSF signal-dependent production of the transcription factor Gfi-1, a zinc-finger protein that binds a DNA sequence upstream of the *Cxcr4* gene and represses expression of CXCR4. Thus, another way in which G-CSF regulates neutrophil numbers in the periphery is through regulation of neutrophil release from the BM.

Once neutrophils exit the BM, they move into the peripheral blood where they may remain until they become senescent or spontaneously apoptose or they may move into a secondary site to perform their innate immune function and die. When neutrophils become senescent or spontaneously apoptose, they are primarily cleared in BM, liver, and spleen (Furze and Rankin, 2008a,b; Suratt *et al.*, 2001; Thakur *et al.*, 1977). It is believed that homeostatic CXCL12 expression in these tissues attracts neutrophils. However, when the neutrophil exit the BM and enter the peripheral blood, CXCR4 expression is downregulated or the receptor has been cleaved off the cell surface (Gordy *et al.*, 2011; Kim *et al.*, 2006; Levesque *et al.*, 2003, 2003; Martin *et al.*, 2003; Nagase *et al.*, 2002). So then how are neutrophils attracted to these tissue sites? Freshly isolated neutrophils lack CXCR4 expression on the cell surface, but after 20 h in culture, this expression increases suggesting that as neutrophils age CXCR4 levels increase (Martin *et al.*, 2003; Nagase *et al.*, 2002). How and whether this occurs under physiological conditions remains unclear.

Once older or senescent neutrophils home into tissue sites, they undergo apoptosis. Apoptotic neutrophils are then phagocytosed by tissue phagocytes (e.g., macrophages, Kupffer cells, and dendritic cells) usually found within the reticular endothelial system. These phagocytes are in direct contact with the blood for better recognition of apoptotic/senescent neutrophils (Furze and Rankin, 2008a,b; Gordy *et al.*, 2011; Suratt *et al.*, 2001). It is hypothesized that the process of engulfing apoptotic neutrophils curbs further production/release of new neutrophils. Macrophages and dendritic cells produce IL-23 which can stimulate $\gamma\delta$ T cells, natural killer (NK) T cells, and Th17 cells to produce IL-17 (Aggarwal *et al.*, 2003). IL-17 stabilizes G-CSF mRNA by increasing its half-life, ultimately enhancing systemic concentrations of G-CSF and driving granulopoiesis and release of neutrophils into the peripheral blood (Cai *et al.*, 1998). However, several papers dispute these findings suggesting that it may be an artifact of the mouse model used or may only occur following stress/inflammation. These latter papers describe a process that does not involve IL-23 or IL-17 (Furze and Rankin, 2008a,b; Gordy *et al.*, 2011). Rather, the data suggest that in BM and spleen, absence of phagocytosis of apoptotic neutrophils results in production of G-CSF, IL-1 β , and CCL3 (MIP-1 α) in an IL-17-independent manner. Further experiments utilized a mouse model that lacked marginal zone and BM stromal macrophages. These mice exhibited neutrophilia that was G-CSF-dependent and IL-1 β and CCL3-independent (Gordy *et al.*, 2011). Mechanisms regulating G-CSF expression are not clear. Thus, under steady-state conditions, circulating neutrophil numbers, as well as new neutrophil production in BM, may be regulated by numbers of apoptotic/senescent neutrophils being phagocytosed by tissue phagocytes in BM, spleen and liver and alterations to the chemokine/cytokine milieu that this process creates.

Role of CXCL12 in Maintenance of the B cell Lineage in BM

Development of lymphoid cells or lymphopoiesis is regulated by CXCL12. CXCL12^{-/-} or CXCR4^{-/-} mice demonstrate reduced B cell lymphopoiesis in the fetal livers and an almost complete loss of lymphopoiesis in BM (Kawabata *et al.*, 1999; Ma *et al.*, 1998, 1999; Nagasawa *et al.*, 1996; Zou *et al.*, 1998). Transplantation of CXCL12 or CXCR4 knockout fetal liver cells into lethally irradiated wild type mice led to greatly reduced reconstitution of adult BM B cells (Kawabata *et al.*, 1999; Ma *et al.*, 1999). These transplanted mice also exhibited increased numbers of B cell precursors in the bloodstream suggesting that CXCL12:CXCR4 signaling is required for retaining B cell precursors within the BM (Ma *et al.*, 1999).

It was discovered that like HSCs, early lymphoid progenitors had preferred niches within the BM (Ding and Morrison, 2013; Greenbaum *et al.*, 2013). Upon deleting *Cxcl12* from osteoblasts, using a Col2.3-*cre* mouse model, there was a modest loss in common lymphoid progenitor cells (CLPs). The importance of the osteoblast lineage in lymphopoiesis was further explored by deleting *Cxcl12* in osteoprogenitor cells utilizing the Osterix-*cre* mouse model. Lack of CXCL12 expression in osteoprogenitor cells led to a significant loss of B lymphoid progenitors and B cell precursors, suggesting that the osteoblast lineage plays a role in B cell lymphopoiesis. This is further supported by early lymphoid progenitors accumulating near the endosteum (Ding and Morrison, 2013). Deletion of CAR cells, an adipo-osteogenic progenitor, led to reduced frequency of CLPs in BM (Omatsu *et al.*, 2010). In addition, deletion of *Cxcl12* from perivascular MSPCs, utilizing the Prx1-*cre* mouse model, also led to significant reduction in lymphoid progenitors. However, deleting *Cxcl12* from endothelial cells, utilizing the Tie2-*cre* mouse model, resulted in unaltered committed lymphoid phenotype. This suggested that MSPCs, osteoprogenitor cells, and to lesser extent osteoblasts play roles in lymphoid lineage maintenance, but endothelial cells do not (Ding and Morrison, 2013). Thus, early lymphoid progenitors, in particular those of the B cell lineage, prefer the endosteal niche. The close proximity of CLP and B cell precursors to niche cells allows these cells to be in a microenvironment that provides essential growth factors for B cell differentiation and proliferation such as IL-7, insulin-like growth factor 1, SCF, and FLT-3 amongst others (D'Apuzzo *et al.*, 1997; Dorshkind *et al.*, 1992; Kincade, 1992).

CXCL12:CXCR4 signaling is an important BM retention signal for B cell precursors. As mentioned above, HSC/HPC and neutrophils rely on altered CXCL12/CXCR4 expression to help determine whether to remain within the BM. Unlike neutrophils, whose CXCR4 expression is downregulated or lost when the cell matures, allowing for its migration out of the BM, immature and naïve B cells that need to migrate out of the BM do not have reduced CXCR4 expression compared to CLP or B cell precursors (Glodek *et al.*, 2003; Stein and Nombela-Arrieta, 2005). Instead, B cell precursors and immature/naïve B cells alter the intensity of VLA-4:VCAM-1 adhesion to stromal cells by regulating the duration of activation of downstream signaling molecules within the CXCL12:CXCR4 signaling pathway. Upon CXCL12:CXCR4 signaling,

focal adhesion kinase (FAK), an important regulator of cell adhesion and migration, is activated by phosphorylation which occurs in a PI3-K- and PKC-dependent process (Glodek *et al.*, 2003). However, the duration of this phosphorylation event is longer in B cell progenitors than in their more mature counterparts. As progenitor B cells differentiate into a more mature phenotype, CXCL12 responsiveness overtime diminishes, causing loss of the sustained adhesion response to stroma cells, thus allowing more mature B cell to migrate out of the BM (Glodek *et al.*, 2003).

Once having migrated into the periphery, B cells undergo antigen-driven differentiation into antibody-secreting cells (ASCs) in peripheral lymphoid tissue (e.g., lymph nodes). This results in downregulation of chemokine receptors that allow B cells to migrate into and be retained within peripheral lymphoid tissues. The type of antigen-driven differentiation helps regulate cell surface chemokine receptor expression. Both systemically induced IgG ASCs and IgM ASCs are found in large number within BM, as well as a small number of intestinally induced IgA ASCs, suggesting that BM is an important site for production of serum antibody responses (Kunkel and Butcher, 2003). The quantity and type of ASCs found in BM is dependent on cell expression of CXCR4 and responds to CXCL12. Intestinally induced IgA, systemically induced IgG and IgM ASCs all express CXCR4. Following immunization in mice that were transplanted with CXCR4-deficient fetal liver cells, antigen-specific ASCs were found in larger numbers in spleen and peripheral blood. This correlated with fewer ASCs within BM, suggesting that CXCR4 plays a role in the ability of ASCs to home to BM. Since there was not a complete loss of CXCR4-deficient ASCs within BM, this suggests that other chemokines/cytokines may play a role in this homing. ASCs found within peripheral blood as well as BM also express CXCR6. The ligand for CXCR6 is CXCL16 which is also expressed in BM compartment. It is important to note that the BM is not the only CXCL12-enriched tissue. Spleen and epithelium in mucosal tissue also express CXCL12, and are known to recruit ASCs. BM, as well as other CXCL12-enriched sites, provides survival signals needed for these cells to perform their function so that long lasting humoral immunity is established.

The Role of CXCL12 in Balancing Granulopoiesis and Lymphopoiesis during Inflammation

CXCL12 plays an important role in maintenance of both myeloid- and lymphoid-lineage cells within BM under homeostatic conditions, but what role does CXCL12 play in regulating granulopoiesis and lymphopoiesis following danger signals? Proinflammatory cytokines, especially tumor necrosis factor α (TNF- α) and IL-1 β , mobilize lymphocytes (e.g., T cells, B cells, and B cell progenitors) in great numbers out of the BM and into the periphery (Nagaoka *et al.*, 2000; Ueda *et al.*, 2005, 2004). Depletion of lymphocyte numbers correlate with reduction in CXCL12 expression in BM (Ueda *et al.*, 2004). This reduction in CXCL12 occurs both by reducing its expression as demonstrated through reduction in message levels, and through posttranslational mechanisms, most likely through proteolytic cleavage (Ueda *et al.*, 2004). Loss of BM lymphocytes is associated with increased extramedullary

lymphopoiesis in the spleen; however, the importance of this event is unknown (Ueda *et al.*, 2004).

Loss of B cell lymphopoiesis and reduction in lymphoid cell numbers in BM is associated with a large expansion in granulopoiesis in BM and increased neutrophil numbers in the periphery (Nagaoka *et al.*, 2000; Ueda *et al.*, 2004, 2005). Expansion of BM granulocytes appears to occur within the same niche that the lymphocytes abandoned. Under homeostatic conditions, granulopoiesis and lymphopoiesis are balanced by a sufficient supply of growth resources that are shared in both processes. The importance of increasing granulopoiesis during inflammation is high since once neutrophils exit the BM its lifespan in the periphery is only hours to days; therefore, it is essential that during inflammation, the equilibrium in the BM should initially shift in favor of granulopoiesis. However, it is also important that the equilibrium between granulopoiesis and lymphopoiesis be re-established following inflammation. Normal lymphoid numbers in the BM following resolution of inflammation corresponds to reduced IL-1 β expression and overexpression of lymphoid growth factors, suggesting that the shared granulocyte and lymphoid niche may require a bias toward lymphopoiesis to reverse increased granulopoiesis.

If CXCL12-depleted mice demonstrate loss of both myeloid and B cell progenitors, why when CXCL12 concentrations are reduced in the BM following inflammation do we observe a loss of lymphopoiesis but an increase in granulopoiesis? One hypothesis is that myeloid progenitors are more sensitive to smaller concentrations of CXCL12 than lymphoid progenitors. Several studies support this rationale. First, CXCL12 is still present in BM following exposure of mice to TNF- α and IL-1 β ; however, it is reduced when compared to normal levels. Second, depleting CXCL12 production from one hematopoietic niche cell at a time was insufficient to reduce myeloid progenitors. Depleting either osteoprogenitor cells or osteoblasts was sufficient to reduce CLP and B cell precursor numbers. Another possibility is that there is another retention signal that works synergistically with lower levels of CXCL12 to help maintain myeloid progenitors in the BM.

Conclusion

CXCL12:CXCR4 signaling is important in maintaining HSC, HPC, and mature leukocyte numbers in the BM and in the periphery. Investigating the role of this key chemokine in the BM microenvironment is essential to understanding hematopoiesis. Not only is this important in discovering new options for enhancing engraftment of HSCs and HPCs or in enhancing numbers of mobilized HSCs and HPCs to collect for transplantation, but also for understanding how leukocyte numbers in the periphery are regulated. By better understanding these processes, it is possible that new therapeutics can be developed for treating different types of blood disorders as well as infection and tissue injury.

References

- Aggarwal, S., Ghilardi, N., Xie, M.H., de Sauvage, F.J., Gurney, A.L., 2003. Interleukin-23 promotes a distinct CD4 T cell activation state characterized by the production of interleukin-17. *Journal of Biological Chemistry* 278, 1910–1914.

- Alvarez, P., Carrillo, E., Velez, C., *et al.*, 2013. Regulatory systems in bone marrow for hematopoietic stem/progenitor cells mobilization and homing. *BioMed Research International* 2013, 312656.
- Bacon, K., Baggiolini, M., Broxmeyer, H., *et al.*, 2002. Chemokine/chemokine receptor nomenclature. *Journal of Interferon & Cytokine Research* 22, 1067–1068.
- Baggiolini, M., Dewald, B., Moser, B., 1997. Human chemokines: An update. *Annual Review of Immunology* 15, 675–705.
- Bonig, H., Wundes, A., Chang, K.H., Lucas, S., Papayannopoulou, T., 2008. Increased numbers of circulating hematopoietic stem/progenitor cells are chronically maintained in patients treated with the CD49d blocking antibody naltizumab. *Blood* 111, 3439–3441.
- Broxmeyer, H., Van Zant, G., Zucali, J.R., LoBue, J., Gordon, A.S., 1974. Mechanisms of leukocyte production and release. XII. A comparative assay of the leukocytosis-inducing factor (LIF) and the colony-stimulating factor (CSF). *Proceedings of the Society for Experimental Biology and Medicine* 145, 1262–1267.
- Broxmeyer, H.E., 2001. Regulation of hematopoiesis by chemokine family members. *International Journal of Hematology* 74, 9–17.
- Broxmeyer, H.E., 2008. Chemokines in hematopoiesis. *Current Opinion in Hematology* 15, 49–58.
- Broxmeyer, H.E., Kohli, L., Kim, C.H., *et al.*, 2003. Stromal cell-derived factor-1/CXCL12 directly enhances survival/antiapoptosis of myeloid progenitor cells through CXCR4 and G(alpha)i proteins and enhances engraftment of competitive, repopulating stem cells. *Journal of Leukocyte Biology* 73, 630–638.
- Broxmeyer, H.E., Orschi, C.M., Clapp, D.W., *et al.*, 2005. Rapid mobilization of murine and human hematopoietic stem and progenitor cells with AMD3100, a CXCR4 antagonist. *Journal of Experimental Medicine* 201, 1307–1318.
- Burdon, P.C., Martin, C., Rankin, S.M., 2005. The CXC chemokine MIP-2 stimulates neutrophil mobilization from the rat bone marrow in a CD49d-dependent manner. *Blood* 105, 2543–2548.
- Burdon, P.C., Martin, C., Rankin, S.M., 2008. Migration across the sinusoidal endothelium regulates neutrophil mobilization in response to ELR + CXC chemokines. *British Journal of Haematology* 142, 100–108.
- Cai, X.Y., Gommoll Jr., C.P., Justice, L., Narula, S.K., Fine, J.S., 1998. Regulation of granulocyte colony-stimulating factor gene expression by interleukin-17. *Immunology Letters* 62, 51–58.
- Chow, A., Lucas, D., Hidalgo, A., *et al.*, 2011. Bone marrow CD169+ macrophages promote the retention of hematopoietic stem and progenitor cells in the mesenchymal stem cell niche. *Journal of Experimental Medicine* 208, 261–271.
- Christopher, M.J., Link, D.C., 2008. Granulocyte colony-stimulating factor induces osteoblast apoptosis and inhibits osteoblast differentiation. *Journal of Bone and Mineral Research* 23, 1765–1774.
- Christopher, M.J., Rao, M., Liu, F., Woloszynek, J.R., Link, D.C., 2011. Expression of the G-CSF receptor in monocytic cells is sufficient to mediate hematopoietic progenitor mobilization by G-CSF in mice. *Journal of Experimental Medicine* 208, 251–260.
- Christopherson 2nd, K.W., Hangoc, G., Broxmeyer, H.E., 2002. Cell surface peptidase CD26/dipeptidylpeptidase IV regulates CXCL12/stromal cell-derived factor-1 alpha-mediated chemotaxis of human cord blood CD34+ progenitor cells. *Journal of Immunology* 169, 7000–7008.
- Christopherson 2nd, K.W., Hangoc, G., Mantel, C.R., Broxmeyer, H.E., 2004. Modulation of hematopoietic stem cell homing and engraftment by CD26. *Science* 305, 1000–1003.
- Christopherson 2nd, K.W., Uralil, S.E., Porecha, N.K., *et al.*, 2006. G-CSF- and GM-CSF-induced upregulation of CD26 peptidase downregulates the functional chemotactic response of CD34+ CD38- human cord blood hematopoietic cells. *Experimental Hematology* 34, 1060–1068.
- Christopherson, K.W., Cooper, S., Hangoc, G., Broxmeyer, H.E., 2003. CD26 is essential for normal G-CSF-induced progenitor cell mobilization as determined by CD26^{-/-} mice. *Experimental Hematology* 31, 1126–1134.
- D'Apuzzo, M., Rolink, A., Loetscher, M., *et al.*, 1997. The chemokine SDF-1, stromal cell-derived factor 1, attracts early stage B cell precursors via the chemokine receptor CXCR4. *European Journal of Immunology* 27, 1788–1793.
- Dar, A., Kollet, O., Lapidot, T., 2006. Mutual, reciprocal SDF-1/CXCR4 interactions between hematopoietic and bone marrow stromal cells regulate human stem cell migration and development in NOD/SCID chimeric mice. *Experimental Hematology* 34, 967–975.
- Day, R.B., Link, D.C., 2012. Regulation of neutrophil trafficking from the bone marrow. *Cellular and Molecular Life Sciences* 69, 1415–1423.
- De La Luz Sierra, M., Gasperini, P., McCormick, P.J., Zhu, J., Tosato, G., 2007. Transcription factor Gfi-1 induced by G-CSF is a negative regulator of CXCR4 in myeloid cells. *Blood* 110, 2276–2285.
- Debnath, B., Xu, S., Grande, F., Garofalo, A., Neamati, N., 2013. Small molecule inhibitors of CXCR4. *Theranostics* 3, 47–75.
- Ding, L., Morrison, S.J., 2013. Haematopoietic stem cells and early lymphoid progenitors occupy distinct bone marrow niches. *Nature* 495, 231–235.
- Ding, L., Saunders, T.L., Enikolopov, G., Morrison, S.J., 2012. Endothelial and perivascular cells maintain haematopoietic stem cells. *Nature* 481, 457–462.
- Dorshkind, K., Narayanan, R., Landreth, K.S., 1992. Regulatory cells and cytokines involved in primary B lymphocyte production. *Advances in Experimental Medicine and Biology* 323, 119–123.
- Eash, K.J., Greenbaum, A.M., Gopalan, P.K., Link, D.C., 2010. CXCR2 and CXCR4 antagonistically regulate neutrophil trafficking from murine bone marrow. *Journal of Clinical Investigation* 120, 2423–2431.
- Eash, K.J., Means, J.M., White, D.W., Link, D.C., 2009. CXCR4 is a key regulator of neutrophil release from the bone marrow under basal and stress granulopoiesis conditions. *Blood* 113, 4711–4719.
- Frenette, P.S., Mayadas, T.N., Rayburn, H., Hynes, R.O., Wagner, D.D., 1996. Susceptibility to infection and altered hematopoiesis in mice deficient in both P- and E-selectins. *Cell* 84, 563–574.
- Frenette, P.S., Subbarao, S., Mazo, I.B., von Andrian, U.H., Wagner, D.D., 1998. Endothelial selectins and vascular cell adhesion molecule-1 promote hematopoietic progenitor homing to bone marrow. *Proceedings of the National Academy of Sciences of the United States of America* 95, 14423–14428.
- Fricker, S.P., 2013. Physiology and pharmacology of plerixafor. *Transfusion Medicine and Hemotherapy* 40, 237–245.
- Furze, R.C., Rankin, S.M., 2008a. Neutrophil mobilization and clearance in the bone marrow. *Immunology* 125, 281–288.
- Furze, R.C., Rankin, S.M., 2008b. The role of the bone marrow in neutrophil clearance under homeostatic conditions in the mouse. *FASEB Journal* 22, 3111–3119.
- Glodek, A.M., Honczarenko, M., Le, Y., Campbell, J.J., Silberstein, L.E., 2003. Sustained activation of cell adhesion is a differentially regulated process in B lymphopoiesis. *Journal of Experimental Medicine* 197, 461–473.
- Gordy, C., Pua, H., Sempowski, G.D., He, Y.W., 2011. Regulation of steady-state neutrophil homeostasis by macrophages. *Blood* 117, 618–629.
- Greenbaum, A., Hsu, Y.M., Day, R.B., *et al.*, 2013. CXCL12 in early mesenchymal progenitors is required for haematopoietic stem-cell maintenance. *Nature* 495, 227–230.
- Hattori, K., Heissig, B., Tashiro, K., *et al.*, 2001. Plasma elevation of stromal cell-derived factor-1 induces mobilization of mature and immature hematopoietic progenitor and stem cells. *Blood* 97, 3354–3360.
- Imbert, A.M., Belaoui, G., Bardin, F., *et al.*, 2006. CD99 expressed on human mobilized peripheral blood CD34+ cells is involved in transendothelial migration. *Blood* 108, 2578–2586.
- Inoue, S., Osmond, D.G., 2001. Basement membrane of mouse bone marrow sinusoids shows distinctive structure and proteoglycan composition: A high resolution ultrastructural study. *Anatomical Record* 264, 294–304.
- Jacobson, L.O., Marks, E.K., Gaston, E.O., Robson, M.J., Zirkle, R.E., 1949. The role of the spleen in radiation injury. *Proceedings of the Society for Experimental Biology and Medicine* 70, 740–742.
- Kahn, J., Byk, T., Jansson-Sjostrand, L., *et al.*, 2004. Overexpression of CXCR4 on human CD34+ progenitors increases their proliferation, migration, and NOD/SCID repopulation. *Blood* 103, 2942–2949.
- Katayama, Y., Battista, M., Kao, W.M., *et al.*, 2006. Signals from the sympathetic nervous system regulate hematopoietic stem cell egress from bone marrow. *Cell* 124, 407–421.
- Kawabata, K., Ujikawa, M., Egawa, T., *et al.*, 1999. A cell-autonomous requirement for CXCR4 in long-term lymphoid and myeloid reconstitution. *Proceedings of the National Academy of Sciences of the United States of America* 96, 5663–5667.
- Kim, H.K., De La Luz Sierra, M., Williams, C.K., Gulino, A.V., Tosato, G., 2006. G-CSF down-regulation of CXCR4 expression identified as a mechanism for mobilization of myeloid cells. *Blood* 108, 812–820.
- Kincade, P.W., 1992. Cell interaction molecules and cytokines which participate in B lymphopoiesis. *Baillière's Clinical Haematology* 5, 575–598.
- Kohler, A., De Filippo, K., Hasenberg, M., *et al.*, 2011. G-CSF-mediated thrombopoietin release triggers neutrophil motility and mobilization from bone marrow via induction of Cxcr2 ligands. *Blood* 117, 4349–4357.
- Kollet, O., Dar, A., Shviti, S., *et al.*, 2006. Osteoclasts degrade endosteal components and promote mobilization of hematopoietic progenitor cells. *Journal home* 12, 657–664.
- Kunkel, E.J., Butcher, E.C., 2003. Plasma-cell homing. *Nature Reviews Immunology* 3, 822–829.
- Le, Y., Zhou, Y., Iribarren, P., Wang, J., 2004. Chemokines and chemokine receptors: Their manifold roles in homeostasis and disease. *Cellular and Molecular Immunology* 1, 95–104.

- Levesque, J.P., Helwani, F.M., Winkler, I.G., 2010. The endosteal 'osteoblastic' niche and its role in hematopoietic stem cell homing and mobilization. *Leukemia* 24, 1979–1992.
- Levesque, J.P., Hendy, J., Takamatsu, Y., Simmons, P.J., Bendall, L.J., 2003. Disruption of the CXCR4/CXCL12 chemotactic interaction during hematopoietic stem cell mobilization induced by G-CSF or cyclophosphamide. *Journal of Clinical Investigation* 111, 187–196.
- Levesque, J.P., Hendy, J., Takamatsu, Y., *et al.*, 2002. Mobilization by either cyclophosphamide or granulocyte colony-stimulating factor transforms the bone marrow into a highly proteolytic environment. *Experimental Hematology* 30, 440–449.
- Levesque, J.P., Hendy, J., Winkler, I.G., Takamatsu, Y., Simmons, P.J., 2003. Granulocyte colony-stimulating factor induces the release in the bone marrow of proteases that cleave c-KIT receptor (CD117) from the surface of hematopoietic progenitor cells. *Experimental Hematology* 31, 109–117.
- Lorenz, E., Uphoff, D., Reid, T.R., Shelton, E., 1951. Modification of irradiation injury in mice and guinea pigs by bone marrow injections. *Journal of the National Cancer Institute* 12, 197–201.
- Lucas, D., Bruns, I., Battista, M., *et al.*, 2012. Norepinephrine reuptake inhibition promotes mobilization in mice: Potential impact to rescue low stem cell yields. *Blood* 119, 3962–3965.
- Ma, Q., Jones, D., Borghesani, P.R., *et al.*, 1998. Impaired B-lymphopoiesis, myelopoiesis, and derailed cerebellar neuron migration in CXCR4- and SDF-1-deficient mice. *Proceedings of the National Academy of Sciences of the United States of America* 95, 9448–9453.
- Ma, Q., Jones, D., Springer, T.A., 1999. The chemokine receptor CXCR4 is required for the retention of B lineage and granulocytic precursors within the bone marrow microenvironment. *Immunity* 10, 463–471.
- Martin, C., Burdon, P.C., Bridger, G., *et al.*, 2003. Chemokines acting via CXCR2 and CXCR4 control the release of neutrophils from the bone marrow and their return following senescence. *Immunity* 19, 583–593.
- Mazo, I.B., Gutierrez-Ramos, J.C., Frenette, P.S., *et al.*, 1998. Hematopoietic progenitor cell rolling in bone marrow microvessels: Parallel contributions by endothelial selectins and vascular cell adhesion molecule 1. *Journal of Experimental Medicine* 188, 465–474.
- Mendelson, A., Frenette, P.S., 2014. Hematopoietic stem cell niche maintenance during homeostasis and regeneration. *Journal home* 20, 833–846.
- Mendez-Ferrer, S., Michurina, T.V., Ferraro, F., *et al.*, 2010. Mesenchymal and haematopoietic stem cells form a unique bone marrow niche. *Nature* 466, 829–834.
- Miyake, K., Medina, K., Ishihara, K., *et al.*, 1991. A VCAM-like adhesion molecule on murine bone marrow stromal cells mediates binding of lymphocyte precursors in culture. *Journal of Cell Biology* 114, 557–565.
- Moll, N.M., Ransohoff, R.M., 2010. CXCL12 and CXCR4 in bone marrow physiology. *Expert Review of Hematology* 3, 315–322.
- Muller, W.A., 2014. How endothelial cells regulate transmigration of leukocytes in the inflammatory response. *American Journal of Pathology* 184, 886–896.
- Murphy, P.M., 1996. Chemokine receptors: Structure, function and role in microbial pathogenesis. *Cytokine & Growth Factor Reviews* 7, 47–64.
- Nagaoka, H., Gonzalez-Aseguinolaza, G., Tsuji, M., Nussenzweig, M.C., 2000. Immunization and infection change the number of recombination activating gene (RAG)-expressing B cells in the periphery by altering immature lymphocyte production. *Journal of Experimental Medicine* 191, 2113–2120.
- Nagasawa, T., Hirota, S., Tachibana, K., *et al.*, 1996. Defects of B-cell lymphopoiesis and bone-marrow myelopoiesis in mice lacking the CXC chemokine PBSF/SDF-1. *Nature* 382, 635–638.
- Nagase, H., Miyamasu, M., Yamaguchi, M., *et al.*, 2002. Cytokine-mediated regulation of CXCR4 expression in human neutrophils. *Journal of Leukocyte Biology* 71, 711–717.
- O'Leary, H., Ou, X., Broxmeyer, H.E., 2013. The role of dipeptidyl peptidase 4 in hematopoiesis and transplantation. *Current Opinion in Hematology* 20, 314–319.
- Omatsu, Y., Sugiyama, T., Kohara, H., *et al.*, 2010. The essential functions of adipogenic progenitors as the hematopoietic stem and progenitor cell niche. *Immunity* 33, 387–399.
- Ou, X., O'Leary, H.A., Broxmeyer, H.E., 2013. Implications of DPP4 modification of proteins that regulate stem/progenitor and more mature cell types. *Blood* 122, 161–169.
- Papayannopoulou, T., Craddock, C., Nakamoto, B., Priestley, G.V., Wolf, N.S., 1995. The VLA4/VCAM-1 adhesion pathway defines contrasting mechanisms of lodgement of transplanted murine hemopoietic progenitors between bone marrow and spleen. *Proceedings of the National Academy of Sciences of the United States of America* 92, 9647–9651.
- Peled, A., Petit, I., Kollet, O., *et al.*, 1999. Dependence of human stem cell engraftment and repopulation of NOD/SCID mice on CXCR4. *Science* 283, 845–848.
- Pelus, L.M., Bian, H., Fukuda, S., *et al.*, 2005. The CXCR4 agonist peptide, CTCE-0021, rapidly mobilizes polymorphonuclear neutrophils and hematopoietic progenitor cells into peripheral blood and synergizes with granulocyte colony-stimulating factor. *Experimental Hematology* 33, 295–307.
- Petit, I., Goichberg, P., Spiegel, A., *et al.*, 2005. Atypical PKC-zeta regulates SDF-1-mediated migration and development of human CD34+ progenitor cells. *Journal of Clinical Investigation* 115, 168–176.
- Richardson, R.M., Tokunaga, K., Marjoram, R., Sata, T., Snyderman, R., 2003. Interleukin-8-mediated heterologous receptor internalization provides resistance to HIV-1 infectivity. Role of signal strength and receptor desensitization. *Journal of Biological Chemistry* 278, 15867–15873.
- Semerad, C.L., Christopher, M.J., Liu, F., *et al.*, 2005. G-CSF potently inhibits osteoblast activity and CXCL12 mRNA expression in the bone marrow. *Blood* 106, 3020–3027.
- Semerad, C.L., Liu, F., Gregory, A.D., Stumpf, K., Link, D.C., 2002. G-CSF is an essential regulator of neutrophil trafficking from the bone marrow to the blood. *Immunity* 17, 413–423.
- Stein, J.V., Nombela-Arrieta, C., 2005. Chemokine control of lymphocyte trafficking: A general overview. *Immunology* 116, 1–12.
- Suratt, B.T., Petty, J.M., Young, S.K., *et al.*, 2004. Role of the CXCR4/SDF-1 chemokine axis in circulating neutrophil homeostasis. *Blood* 104, 565–571.
- Suratt, B.T., Young, S.K., Lieber, J., *et al.*, 2001. Neutrophil maturation and activation determine anatomic site of clearance from circulation. *American Journal of Physiology: Lung Cellular and Molecular Physiology* 281, L913–921.
- Thakur, M.L., Lavender, J.P., Arnot, R.N., Silvester, D.J., Segal, A.W., 1977. Indium-111-labeled autologous leukocytes in man. *Journal of Nuclear Medicine* 18, 1014–1021.
- Ueda, Y., Kondo, M., Kelsoe, G., 2005. Inflammation and the reciprocal production of granulocytes and lymphocytes in bone marrow. *Journal of Experimental Medicine* 201, 1771–1780.
- Ueda, Y., Yang, K., Foster, S.J., Kondo, M., Kelsoe, G., 2004. Inflammation controls B lymphopoiesis by regulating chemokine CXCL12 expression. *Journal of Experimental Medicine* 199, 47–58.
- Vagima, Y., Lapid, K., Kollet, O., *et al.*, 2011. Pathways implicated in stem cell migration: The SDF-1/CXCR4 axis. *Methods in Molecular Biology* 750, 277–289.
- Wengner, A.M., Pitchford, S.C., Furze, R.C., Rankin, S.M., 2008. The coordinated action of G-CSF and ELR + CXC chemokines in neutrophil mobilization during acute inflammation. *Blood* 111, 42–49.
- Winkler, I.G., Sims, N.A., Pettit, A.R., *et al.*, 2010. Bone marrow macrophages maintain hematopoietic stem cell (HSC) niches and their depletion mobilizes HSCs. *Blood* 116, 4815–4828.
- Yong, K.L., Watts, M., Shaun Thomas, N., *et al.*, 1998. Transmigration of CD34+ cells across specialized and nonspecialized endothelium requires prior activation by growth factors and is mediated by PECAM-1 (CD31). *Blood* 91, 1196–1205.
- Zou, Y.R., Kottmann, A.H., Kuroda, M., Taniuchi, I., Littman, D.R., 1998. Function of the chemokine receptor CXCR4 in haematopoiesis and in cerebellar development. *Nature* 393, 595–599.

Lymphocyte–Endothelial Interactions

CV Carman and R Martinelli, Harvard Medical School, Boston, MA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Antigen Molecules that can elicit an immune response.

Antigen presentation The display of peptides bound to major histocompatibility complex on the surface of an antigen presenting or target cell, permitting specific recognition by T cells and their activation.

Co-inhibitory molecule A molecule on the surface of an antigen presenting or target cell that provides a negative signal for the antigen-specific activation of T cells.

Co-stimulatory molecule A molecule on the surface of an antigen presenting or target cell that provides a positive signal for the antigen-specific activation of T cells.

Diapedesis The process by which a cell breaches and migrates across an endothelial barrier in order to either exit or enter the blood circulation.

Immune surveillance The process of the immune cells moving throughout the body in search of invading pathogens and aberrant (e.g., cancerous) host cells.

Immunological synapse (IS) Interface between a lymphocyte and an antigen-presenting cell responsible for coordination of antigen recognition and response.

Intravasation The process through which a cell (leukocytes and cancer cells) breaches the basal lamina of the endothelium from the interstitium into the lumen of a blood or lymphatic vessel, in order to enter the circulation.

Invadopodia Cylindrical cellular structures with an actin-rich core that establish close contact with the extracellular matrix (ECM), typically formed in cancer cells. Invadopodia have a diameter of approximately 8 μm and a height of approximately 5 μm , they are very stable and are associated with matrix degradation.

Invadosomes Term including both podosomes and invadopodia.

Invadosome-like protrusions (ILPs) Submicron-scale, actin-rich cylindrical protrusions related to invadosomes through which lymphocytes actively probe the surface of the endothelium by dynamic insertion and retraction. These were demonstrated to function in supporting migratory pathfinding and to enforce close T cell–EC apposition, which seems to facilitate Ag recognition and TCR signaling.

Major histocompatibility complex (MHC) Surface molecules involved in the process of antigen presentation.

Podosome Highly dynamic cylindrical cellular structures with an actin-rich core that establish close contact with the ECM, typically formed in monocytic cells (monocytes, macrophages, dendritic cells, and osteoclasts), endothelial cells, and smooth muscle cells. Podosomes have a diameter of approximately 1 μm and a height of approximately 0.4 μm .

Podo-prints Endothelial invaginations generated by protrusion of lymphocyte ILPs.

Podo-synapse A novel IS architecture dominated by dense arrays of calcium-stabilized ILPs.

Secondary lymphoid organs (SLOs) SLOs include lymph nodes, tonsils, spleen, Peyer's patches, and mucosa-associated lymphoid tissue (MALT). These serve as locations for lymphocytes become in contact with the antigen to initiate and regulate adaptive immune responses.

T cell receptor T cell surface molecule that recognizes the antigen in the context of MHC.

Introduction

The endothelia of the contiguous vascular and lymphatic circulatory systems form the essential barrier that partitions the body's most fundamental compartments: the tissues and the blood/lymph. Immune cells have the mandate to continuously move throughout the body and pass in and out of the tissues in order to conduct immune surveillance and mediate immune responses. The basic barrier versus invasive functions of endothelial and immune cells, respectively, would seem to be fundamentally opposed. However, highly cooperative interactions between these cell types promote extremely efficient immune cell trafficking and surveillance while retaining barrier functions. The nature of this cooperation seems to be deeply rooted in evolutionary and developmental origins. Indeed, the first appearance of both a closed endothelial-lined circulatory system (Monahan-Earley *et al.*, 2013) and of an adaptive immune system (Bajoghli *et al.*, 2009;

Kasahara and Sutoh, 2014; Hirano *et al.*, 2011) occurred simultaneously ~500 million years ago in the ancestors of the most phylogenetically distant living vertebrates from mammals, lampreys, and hagfish. Additionally, in mammalian development both endothelial and hematopoietic lineages derive from a common hemangioblast precursor (Choi *et al.*, 1998; Lancrin *et al.*, 2009; Risau and Flamme, 1995; Tavian *et al.*, 2000), which would seem to underlie the unique multitasking ability of the endothelium to function as both a barrier and an antigen-presenting cell (APC). In this article we provide an overview of the architecture of the vascular and lymphatic circulation, as well as the basic contexts and mechanisms for lymphocyte–endothelial interactions during trafficking. After establishing this background, we summarize the known functional contributions of endothelial cells (ECs) to antigen (Ag)-dependent regulation of adaptive immunity and the unique cellular and molecular mechanisms involved in this information exchange.

The Endothelium as a Barrier for the Vascular and Lymphatic Systems

Vascular Endothelium

The vascular endothelium is an anatomically unique structure. A defining feature of higher multicellular organisms is the presence of a closed cardiovascular circulatory system that delivers oxygen and nutrient-rich blood to the tissues (Monahan-Earley *et al.*, 2013). This system is organized as a series of large arterial vessels (macrovascular) that carry blood away from the heart and progressively branch into smaller diameter

arteriole vessels and ultimately capillaries in which oxygen and nutrient exchange occurs (Figure 1). Further downstream, individual capillaries merge into progressively larger post-capillary venules and veins that eventually return to the heart (Figure 1). The vascular endothelium is the monolayer of cells (blood endothelial cells; BECs) that form the inner lining of this circulatory system and distinctly serves as the physical barrier that maintains organization between the blood and tissue compartments.

Altogether the blood vasculature is an enormous organ system that is comprised of ~60 trillion BECs, that make up

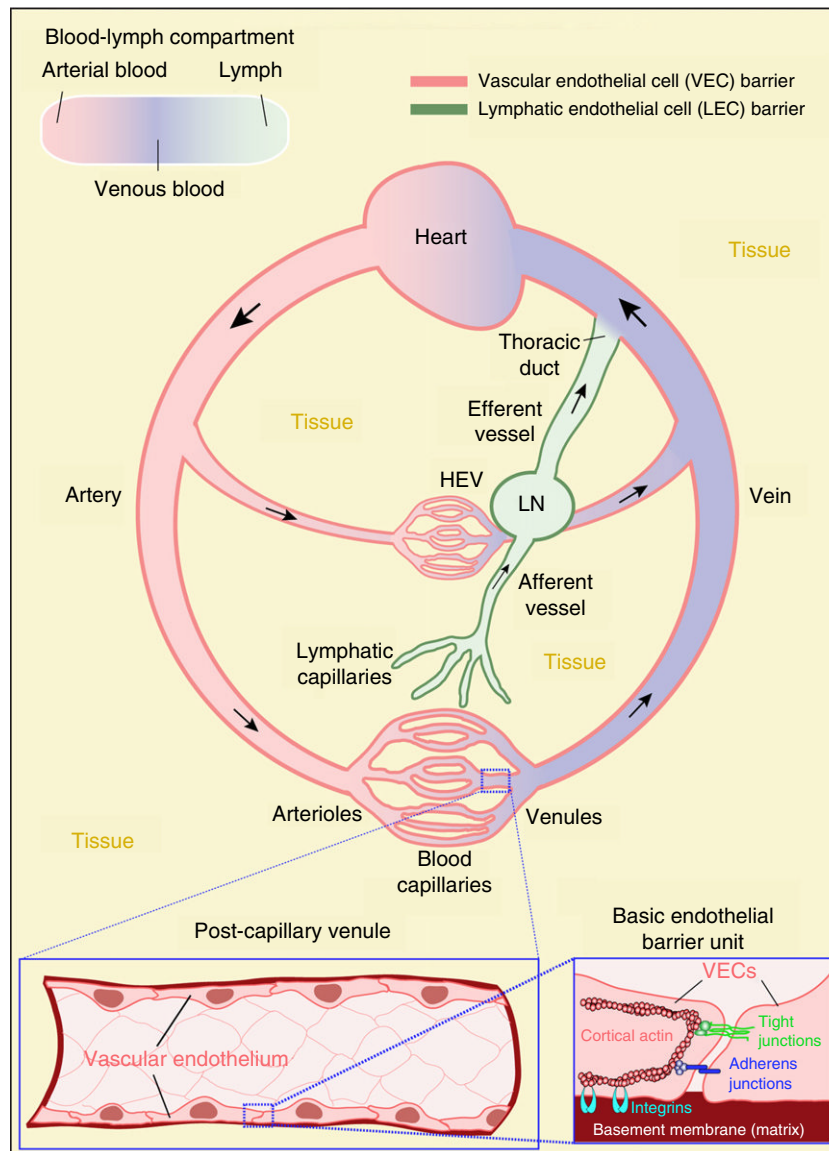


Figure 1 Blood-lymph circulatory system schematic depicts the contiguous blood-lymph circulatory system. Arterial, oxygen-rich, blood (pink) flows away from the heart into the microvasculature (arterioles, capillaries, and venules). Oxygen-depleted blood (blue) flows away from the microvasculature back to the heart. Fluid (lymph; green) leaked into the tissue (yellow) from the microvasculature is taken up by the lymphatic capillaries, and progressively flows through the afferent lymphatic vessels, LNs, efferent vessels, and back into the blood circulation via the lymphatic duct. Local microvasculature of the lymph node (i.e., high endothelial venules (HEV)) serves as a location for lymphocytes to enter the LN. Dark pink and dark green lines indicate the vascular and lymphatic endothelial barriers, respectively. Boxed region (lower panels) depicts a segment of postcapillary venule (right) and the fundamental composition of the minimal endothelial barrier unit: an endothelial cell monolayer; intercellular adhesions, adhesions to the basement membrane.

~165 000 km of vessels with a total surface area of ~4000–7000 m² (most of which are the microvascular vessels; i.e., arterioles, capillaries, and venules) (Aird, 2005; Hwa *et al.*, 2005). The overall architecture of the blood vasculature is such that microvessels are densely interdigitated throughout all other tissues and organs of the body. As a result, most tissue cells lie within just tens of microns of a vessel. This allows for highly effective delivery of oxygen and nutrients (Aird, 2007a,b). An additional consequence of this design is that every tissue microenvironment has a local interface (i.e., BECs) with the circulation. As we discuss below, evolution has apparently taken advantage of this arrangement to provide BECs with effective ‘sentinel’ functions that are able to communicate local interstitial information (e.g., infection, damage) directly to circulating immune cells in circulation.

Lymphatic Endothelium

The blood circulatory system works in concert with an adjunct lymphatic system. Rather than forming a closed circulation, the lymphatic vessels begin blindly in the intercellular spaces of the soft tissue (not unlike the roots of a tree) in close proximity to microvascular beds (Figure 1; Girard *et al.*, 2012). Such ‘terminal lymphatic capillaries’ progressively merge to form afferent collecting vessels that feed into lymph nodes (secondary lymphoid tissues that receives antigens from lymph), which are connected downstream to efferent lymphatic vessels. These merge into larger collecting vessels and ultimately the thoracic duct that connects directly to the venous blood circulation at the subclavian vein (Figure 1). The lymphatic system serves to collect fluids (lymph fluid), which leak from capillaries into the tissue, and return them back to the vascular circulation. This is required to maintain an appropriate interstitial fluid balance. In addition, the lymphatics provide a route for trafficking of interstitial APCs (e.g., dendritic cells; DCs) to lymph nodes (Girard *et al.*, 2012; Figure 1). As with the blood circulation, the lymphatic system is also lined by a monolayer of lymphatic endothelial cells (LECs) that provide the barrier that separates the tissue from the lymph.

Basic Features of Endothelia

LECs and BECs are highly similar in many regards, yet also exhibit important phenotypic and functional distinctions (Aird, 2007a,b; Pfeiffer *et al.*, 2008). Moreover, significant heterogeneity is found among ECs based on the size of a vessel and the tissue in which they reside (Aird, 2007a,b). Nonetheless, the minimal barrier unit of all endothelia fundamentally consist of a monolayer of endothelial cells that are held to each other by adherens and tight junction proteins (e.g., VE-cadherin, platelet endothelial cell adhesion molecule-1 (PECAM-1), junctional adhesion molecule (JAM-1), CD99, claudins, and occludins) and to the underlying basement membrane by integrin adhesion receptors (e.g., $\alpha\beta3$ and $\alpha5\beta1$) (Bazzoni and Dejana, 2004; Pepper and Skobe, 2003; Baluk *et al.*, 2007; Figure 1).

Contexts for Lymphocyte–Endothelial Interactions

In order to conduct their basic function of searching for pathogens to eliminate, immune cells must continuously traffic throughout the body (von Andrian and Mackay, 2000). While movement within the vascular circulation and lymphatic systems provides a means of rapid and organized transit, immune cells must also repeatedly move in and out of the tissues by adhering to and migrating across the BEC and LEC barriers (a process known as ‘transendothelial migration’ or ‘diapedesis’) (Ley *et al.*, 2007; von Andrian and Mackay, 2000). Before we discuss the mechanisms driving immune cell adhesion and diapedesis, we will first consider the diverse contexts for such events, focusing on just one subtype of immune cell, the T lymphocyte (or T cell).

T Lymphocytes Defined

T lymphocytes are a branch of the adaptive immune system, which are able to detect the presence of infecting pathogens (e.g., virus or bacteria) or aberrant host cells (e.g., tumor cells) by identifying unique molecular features that distinguish them from the host (termed foreign Ag). Specifically, these Ag are derived from protein fragments (i.e., ‘peptide Ag’) and recognized by specialized cell-surface T cell receptors (TCRs). Importantly, peptide Ag can only be detected when presented on the surface of an APC by major histocompatibility complex (MHC) proteins (Schwartz, 1992; Germain, 1994). T cells that have never encountered their specific cognate Ag are termed ‘naïve.’ Those that have been activated by Ag become either ‘effector cells’ that participate actively in immune responses to eradicate the detected pathogen or long-lived ‘memory’ cells that serve as the basis for ‘acquired’ or ‘adaptive’ immunity. Additionally, two major kinds of T cells exist. Those that bear the cell surface marker CD4 only respond to Ag presented on class II MHC molecules (MHC-II) and participate in immune reactions to exogenous antigens, mostly through secretion of cytokines that ‘help’ other immune cells to conduct their functions (termed ‘CD4 T helper or Th cells’) (Schwartz, 1992; Germain, 1994). Those that bear CD8 respond to endogenously produced Ag presented on MHC-I and function in direct killing of Ag bearing cells (termed ‘CD8 T cells’ or ‘cytotoxic T cells’; CTL) (Schwartz, 1992; Germain, 1994).

Trafficking during T Lymphocyte Maturation

The life cycle of all immune cells starts in the bone marrow. In the case of T lymphocytes, immature T cell progenitors first cross the bone marrow sinusoidal endothelium in order to enter the blood stream (i.e., intravasate) (De Bruyn *et al.*, 1971; Chamberlain and Lichtman, 1978; Wolosewick, 1984; Figure 2). These immediately home to the thymus where they exit the blood circulation (i.e., extravasate) (Figure 2) via diapedesis across the thymic postcapillary vessels in order to enter the thymus. Here, strongly self-reactive cells are deleted and weakly self-reactive ones mature. Weak self-reactivity generates a repertoire of mature T cells that generally do not become activated in response to self-antigens and have a possibility of reacting to foreign antigen on self-MHC proteins.

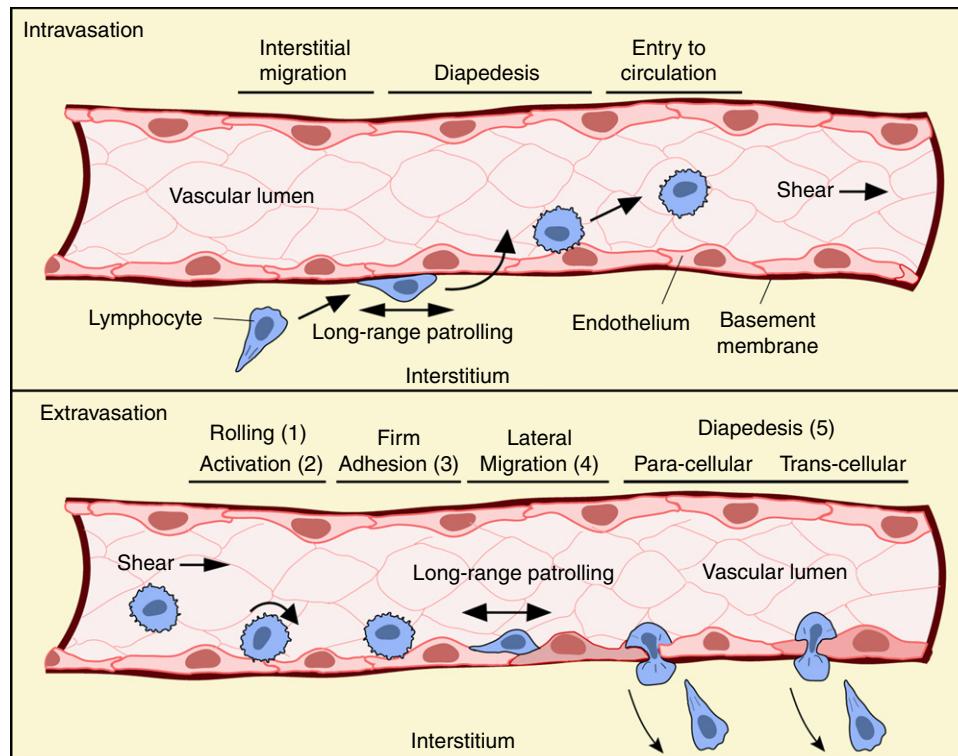


Figure 2 Stages of lymphocyte trafficking across endothelial barriers: Intravasation and extravasation trafficking of lymphocytes throughout the body requires their movement out of (intravasation) and into (extravasation) the tissue. Upper Panel: Stages of intravasation. The details of the process by which lymphocytes (blue) enter vascular and lymphatic structures are not well characterized, but generally characterized to involve steps of interstitial migration toward the vessel, migration (sometimes long-range) on the basal-lateral side of the endothelium, migration across the basement membrane and endothelial barriers (diapedesis), and ultimately, entry of lymphocytes into the blood vessel lumen and the circulation. Lower Panel: Stages of extravasation. The lymphocyte extravasation process has been intensively studied and shown to involve, in the first instance, transient rolling-type interactions mediated predominantly by selectins. This facilitates chemokine-dependent activation and firm arrest, which is mediated by the binding of leukocyte integrins (e.g., LFA1, Mac1, and VLA4) to endothelial cell-adhesion molecules (e.g., ICAM1, ICAM2, and VCAM1). Subsequently, leukocytes migrate laterally over the surface of the endothelium, probing for a site to penetrate through it (see also **Figure 3**). Finally lymphocytes cross the endothelial barrier (diapedese) and enter the interstitium either by opening a gap between two adjacent endothelial cells (paracellular route) or by migrating directly through the body of a single endothelial cell (transcellular route).

A small subset of T cells with intermediate self-reactivity become regulatory T cells (Tregs). Mature naïve T lymphocytes exit the thymus via diapedesis across postcapillary venules (to reenter the blood circulation directly) or across lymphatic vessels of the thymus (to enter the lymphatic system, and eventually the blood circulation) (Toro and Olah, 1967; Ushiki, 1986; Weinreich and Hogquist, 2008; Kotani *et al.*, 1966; Ernstom *et al.*, 1965).

Trafficking during T Lymphocyte Immune Surveillance

Once in the blood, naïve T cells begin constitutive cycles of migrating into and out of secondary lymphoid organs (SLOs) in search of cognate Ag. This starts as T cells pass through specialized high endothelial (postcapillary) venules (HEVs) that constitutively expresses adhesion molecules (e.g., peripheral lymph node addressin (PNAd), ICAM-1, ICAM-2, and mucosal vascular addressin cell adhesion molecule 1 (MAdCAM-1)) and chemokines (e.g., CC chemokine ligands (CCL19 and CCL21)) necessary to promote naïve and memory

lymphocyte diapedesis (Girard *et al.*, 2012; Miyasaka and Tanaka, 2004; von Andrian and Mackay, 2000). The exception to this is the spleen, where there are no HEV and T and B cells enter through the specialized blood endothelium of the splenic marginal zone or through 'bridging channels' connecting the venous sinuses of the splenic red pulp (RP) from the SLO compartment referred to as the white pulp (WP). Once inside SLO, T cells migrate to 'T cell area' where high numbers of tissue-derived DCs are found (Bajenoff *et al.*, 2007). DCs are a type of 'professional' APC that can efficiently phagocytize extracellular material, process proteins and present peptide Ag on their surface via both MHC-I and MHC-II molecules. DCs also have a unique ability (along with LEC, see below) to ingest exogenous antigens and present them on the MHC-I, such that DCs do not need to be infected with viruses or intracellular microbes to be able to initiate activation of CD8⁺ T cells. DCs also express a wide range of 'co-stimulatory' molecules and cytokines that are also needed, together with MHC-Ag, to activate responses in naïve T cells (Schwartz, 1992; Janeway and Bottomly, 1994; Paul and Seder, 1994; Germain, 1994) (discussed further below). As T cells migrate

throughout the T cell area they form a series of short-lived intimate contacts with the DCs that allow them to survey the cell surface MHC–Ag complexes (von Andrian and Mempel, 2003). When these interactions fail to identify a cognate Ag, T cells leave the T cell area and undergo random migration within the cortical and medullary sinuses of the lymph node. During this process they mediate direct interactions with the LECs that line these sinuses. Eventually, after tens of minutes these T cells transmigrating across the LECs barrier to enter the efferent lymphatic vessels, which ultimately carry them back into the blood circulation (Figure 1; Grigorova *et al.*, 2009, 2010; Farr *et al.*, 1980; Olah and Glick, 1985).

Trafficking during T Lymphocyte Immune Responses

In cases when T cells do encounter cognate Ag on lymph node DCs, the two cells engage in sustained (~30 min to 8 h) intimate cell–cell interactions, with precisely regulated sub-cellular architecture and dynamics (Mempel *et al.*, 2004; Fooksman *et al.*, 2010). The result of such ‘immunological synapses’ (ISs) (Grakoui *et al.*, 1999) is the triggering of sustained calcium signaling in the T cells that leads to T cell activation, proliferation, and differentiation into of Ag-specific effector and memory lymphocytes (Dustin, 2003). These cells reenter the circulation now reprogrammed (e.g., with decreased CD62L and CC chemokine receptor (CCR7) and increased LFA-1, very late antigen 4 (VLA4), CCR3, CCR5, and CXCR3 expression) to adhere preferentially to, and migrate across, activated endothelia of infected/inflamed peripheral tissues (Springer, 1994; von Andrian and Mackay, 2000) (discussed further below). Of note, any tissue can become inflamed and the resident vascular beds may exhibit extensive phenotypic and functional diversity that may strongly influence the outcome of these interactions (Aird, 2007a,b).

Trafficking-Independent Interactions between T Lymphocytes and EC

In addition to the trafficking events described above, there is a range of trafficking/diapedesis-independent settings whereby T cells and ECs engage in intimate contacts. First, it must be noted that the average diameter of a lymphocyte (~8–12 μm) exceeds that of capillaries (~5–10 μm). Thus, every trip through the circulation includes a phase of extremely close T cell–endothelial cell (EC) contacts as the T cell pass through a capillary bed. In addition, emerging intravital imaging technologies have begun to reveal more sustained intra- and extravascular T cell–BEC interactions. CD4 Th1 and Th17 lymphocytes were discovered to undergo repeated transient (~20 min duration) arrest on resting liver sinusoidal endothelium (Carambia *et al.*, 2013). Certain immune cells have also been shown to undergo constitutive long-range intravascular migrations (Figure 2). Such intravascular surveillance was first noted for natural killer T cells patrolling the liver sinusoidal endothelium (Geissmann *et al.*, 2005; Velazquez *et al.*, 2008). This patrolling was arrested by exposure to glycolipid antigens, presented by CD1d, recognized by these ‘innate-like’ T cells. Subsequently, subsets of monocytes were seen to crawl continuously within noninflamed peripheral

microvasculature of the skin, mesentery, and central nervous system, apparently searching for signs of infection or damage in the underlying tissues in a form of innate patrolling (Auf-fray *et al.*, 2007; Audoy-Remus *et al.*, 2008). Recently, effector T lymphocytes were shown to undergo extended both luminal and abluminal migration on brain microvascular endothelium in the setting of autoimmune reaction (Bartholomaeus *et al.*, 2009).

Altogether the preceding discussion demonstrates that lymphocytes undergo extensive interactions with diverse endothelia. These take place at different time points during lymphocyte maturation and differentiation, each with distinct functional roles. These represent important opportunities for functional information exchange between T cells and ECs. Indeed, exquisitely sophisticated local intercellular communication between T cells and ECs exist that are central to effective immune functions, as discussed below.

Mechanisms for T Lymphocyte Adhesion to and Diapedesis across Endothelia

Molecular Basis for Adhesion and Diapedesis

Lymphocyte–endothelial interactions are essential determinants of lymphocyte trafficking. The process of lymphocyte egress from the blood into the tissues (extravasation) involves a highly orchestrated dynamics on the part of both the lymphocyte and the endothelium, as we discuss in the following section. The process of intravasation is equally important and probably shares many of the mechanisms of extravasation (Huttenlocher and Poznansky, 2008; Carman, 2009a; Wolosewick, 1984). However, since this process remains far less studied, it will not be discussed further herein.

Whether for constitutive (e.g., to thymus or SLOs) or inflammation-induced trafficking, the first active events in extravasation are the expression and cell surface presentation of adhesion molecules and chemoattractants by ECs. In response to local tissue-derived signals (e.g., cytokines), BECs present cell surface adhesion molecules including selectins (E-selectin, P-selectin), selectin ligands (PNAd), and cellular adhesion molecules (CAMs; ICAM-1, ICAM-2, VCAM-1, MAdCAM-1) (Ley *et al.*, 2007; von Andrian and Mackay, 2000). Some of these molecules, such as P-selectin, are mobilized to the cell surface from preexisting intracellular pools (i.e., specialized granules termed Weibel–Palade bodies). Others are transcriptionally upregulated. In addition, endothelial cells present chemoattractants bound the glycoproteins and glycolipids (collectively termed the glycocalyx) on their luminal surface (Proudfoot *et al.*, 2003). These may be expressed by the endothelial cells themselves (e.g., GM-CSF, Gro-regulated oncogenes (GRO- α , GRO- γ), IL-6, IL-8, monocyte chemoattractant protein (MCP-1), regulated on activation, normal T cell expressed and secreted (RANTES)) (Goebeler *et al.*, 1997; Card *et al.*, 2014) and/or be produced in the interstitium, internalized by BECs abluminally and actively transcytosed to the luminal surface (Middleton *et al.*, 1997, 2002). By creating carefully controlled local expression/presentation of the appropriate adhesion molecules and chemoattractants,

endothelia generate a type of ‘address code’ to promote adhesion and trafficking of target subsets of immune cells (Springer, 1994; von Andrian and Mackay, 2000; Butcher, 1991). As just one simple example we can consider the tissue of the gastro-intestinal tract or ‘gut’. The specific immunologic demands of this environment require specialized subsets of gut lymphocytes that uniquely express the integrin $\alpha 4\beta 7$ (Gorfu *et al.*, 2009). Only the HEVs of gut-associated lymphoid tissues or microvasculature of inflamed gut tissues express the specific ligand for $\alpha 4\beta 7$, MAdCAM, thus providing a basis for selective recruitment of the gut-specific lymphocyte subset.

Once the address code is in place (for acute inflammation this can be induced in just a few hours) a well-ordered series of events take place (Butcher, 1991; Springer, 1994; Luscinskas *et al.*, 1994; Ley *et al.*, 2007; Figure 2). This process begins with transient rolling type adhesions mediated by the selectin family of adhesion molecules binding to their glycoprotein/glycolipid ligands (Step 1). The resulting increase in cell–cell contacts facilitates T cell sensing of chemokines presented on the endothelial surface, which induces intracellular signaling and ‘activation’ responses (Step 2). This, in turn, triggers high affinity interaction of lymphocyte integrin adhesion receptors (e.g., LFA-1, Mac-1, and VLA-4) with their endothelial ligands (e.g., ICAM-1, ICAM-2, and VCAM-1) resulting in firm lymphocyte arrest (Step 3) (Luo *et al.*, 2007). Subsequently, lymphocytes undergo actin-dependent spreading, polarization and integrin/CAM-dependent lateral migration over the luminal surface of the endothelium (covering tens of microns over several minutes) (Step 4). This activity seems to allow lymphocytes to search out sites permissive for endothelial barrier penetration (Schenkel *et al.*, 2004; Phillipson *et al.*, 2006). Finally, the lymphocyte formally breaches the endothelium either at the intercellular junctions (i.e., para-cellularly) or directly through individual endothelial cells (i.e., the ‘trans-cellular route’) (Step 5) (Muller, 2003; Sage and Carman, 2009; Ley *et al.*, 2007; Figure 2). In addition to integrins and CAM ligands, other adhesion molecules (e.g., PECAM-1, JAM-1, and CD99) also contribute to Step 5.

Specialized Cytoskeletal Remodeling during Adhesion and Diapedesis

During the above cascade avid cytoskeletal remodeling takes place in both the lymphocyte and the endothelium. These act in concert to create more intimate interactions and ultimately guide the process of diapedesis while preserving endothelial barrier properties.

Endothelial remodeling: Docking, guidance and barrier maintenance

In response to lymphocyte adhesion, the endothelium rapidly assembles F-actin-rich microvilli-like structures at the site of T cell attachment (Barreiro *et al.*, 2002; Carman *et al.*, 2003; Carman and Springer, 2004; Carman, 2009b). These form as a consequence of direct ligation and clustering of ICAM-1 and VCAM-1 by lymphocyte integrins, which triggers signaling to the actin adaptor/regulatory proteins, ezrin and moesin (Barreiro *et al.*, 2002). During rapid lateral T cell migration, these microvilli contacts serve as tethers at the trailing edge of the

lymphocyte to encourage lateral migration arrest (Carman and Springer, 2004). As lymphocytes slow their lateral migration, large assemblies of microvilli form symmetrically around the T cells that effectively embrace them, termed ‘docking structures’ or ‘transmigratory-cups’ (Barreiro *et al.*, 2002; Carman and Springer, 2004). The resulting increase in cell–cell contact area (which is enriched endothelial ICAM-1 and VCAM-1 and lymphocyte LFA-1 and VLA-4 integrins) promotes adhesion strengthening to better resist T cell detachment by fluid shear forces of the circulation. This structure also provides a traction scaffold that is oriented perpendicular to the plane of the endothelium (i.e., parallel to the ultimate plane of diapedesis) that functionally guides T cells to breach and migrate across the endothelial barrier (Carman, 2009b; Figure 3). At the same time, this architecture helps to meet the need for functional barrier maintenance. Though significant size ($\sim 4\text{--}6\ \mu\text{m}$) pores and gaps must form in the endothelial barrier, which serve as passageways that accommodate diapedesis, the transmigratory cup maintains extensive close cell–cell contacts that minimize leakage of blood fluids or proteins during the diapedesis event. At the final conclusion of diapedesis (which is typically several minutes after the process began) the lymphocyte ultimately exits the migration passageway. At this point, the endothelium rapidly mobilizes a distinct type of Rac-1- and Arp2/3-dependent actin remodeling activity (termed ‘ventral lamellipodia’), which serve the specialized role of rapidly resealing the barrier (Martinelli *et al.*, 2013). Thus, endothelial cells actively support and guide lymphocyte egress across it while maintaining barrier integrity through actin remodeling dynamics and intimate adhesions.

T Lymphocyte remodeling: Spreading, migration, probing, and piercing

Concomitant with the above endothelial remodeling, T cells engage in their own complex cytoskeletal and shape-change dynamics. Shortly following initial arrest, the actin cortex of the T cell breaks down in favor of Arp2/3-mediated lamellipodia. These initially radiate outward symmetrically and then rapidly (within tens of seconds) polarized into a single leading edge. Commensurate with the appearance of lamellipodia is the formation of orthogonal F-actin assemblies termed invadosome-like protrusions (ILPs) (Carman *et al.*, 2007; Carman, 2009a; Shulman *et al.*, 2009; Gerard *et al.*, 2009) that are highly related to podosomes and invadopodia found in other cells types (Linder, 2009). Each ILP is a cylindrically-shaped column of F-actin with both depth and diameter each of $\sim 0.5\text{--}1\ \mu\text{m}$. These protrude from the bottom of the T cells and exert physical forces against the endothelium that promote extremely intimate cell–cell contacts and drive deformation/indentations in the endothelial cell surface (termed ‘podo-prints’) (Carman *et al.*, 2007). ILPs are enriched in, and dependent on, LFA-1, the actin regulatory proteins WAS protein (WASp) and HS1 (a hematopoietic homolog of cortactin) and src kinase (Carman, 2009a). During lateral migration, T cells extend dozens of ILPs (preferentially near the leading edge) against the endothelia surface that dynamically turnover, each with lifetimes of $\sim 20\ \text{s}$. In this way the lymphocytes are able to physically probe or ‘palpate’ the local biomechanical properties, specifically the stiffness, of the endothelial substrate (Figure 3). Such ‘biomechanical scanning’ facilitates

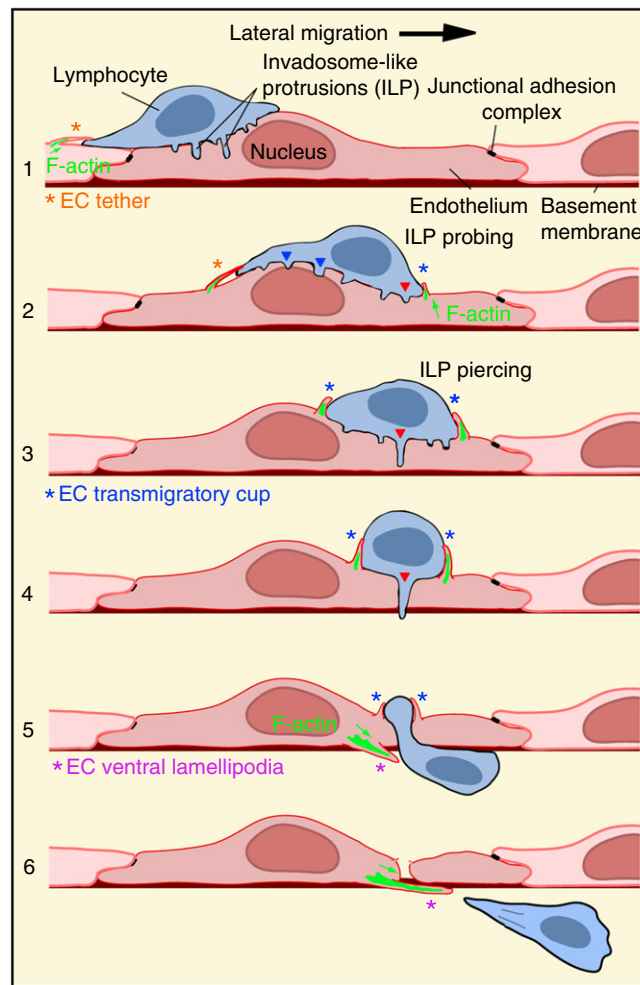


Figure 3 Dynamic remodeling in endothelium and lymphocytes during diapedesis. Schematic depicts EC (pink) and lymphocyte (blue) dynamics during T cell lateral migration over and transcellular diapedesis across the endothelium. Numbered frames show successive time points that are intended to represent intervals of ~30–60 s. Dynamic insertion (~0.2–1 μm in depth) and retraction of multiple lymphocyte ILPs into the apical surface of the endothelium, is shown during lateral migration in frames 1–2. ILPs that form over the nucleus of the endothelial cell are impeded by physical resistance from the nuclear lamina and remain shallow (blue arrowheads in frame 2). At a location of sufficiently low endothelial resistance, an ILP progressively extends several micrometers in depth, ultimately breaching the endothelium transcellularly (red arrowheads in panels 2–4). Also shown is the previously described ‘trans migratory cup’ structure (asterisks), which consists of vertical endothelial microvilli-like projections (rich in F-actin; green), ICAM1, VCAM1, PECAM1, and JAM1 that surround the periphery of adherent lymphocytes. During rapid T cell lateral migration the projections serve largely as tethers to slow lymphocyte migration (orange asterisks). At later stages this structure serves to guide diapedesis by providing a vertical traction substrate for the protrusion of leukocytes against the endothelial surface (blue asterisks). As the lymphocyte exits the transmigration passageway, endothelial cells mobilize ‘ventral lamellipodia’ (purple asterisks) that rapidly restore endothelial barrier integrity.

identification of sufficiently soft EC regions, perhaps also sensing the underlying basement membrane, at which to initiate diapedesis, where ILPs can progressively extend and formally breach the endothelial barrier (Carman *et al.*, 2007; Martinelli *et al.*, 2014). Thus, lymphocytes use ILPs to guide an active process of seeking out the path-of-least-physical-resistance for egress into the tissue (termed ‘tender-taxis’) (Martinelli *et al.*, 2014). Lymphocytes from patients with the genetic immune-deficiency known as Wiskott–Aldrich syndrome (WAS), lack WASp and are unable to effectively form ILPs, which results in defective tender-taxis and diapedesis (Carman *et al.*, 2007).

Additionally, studies suggest that the intimate contacts generated by ILPs may facilitate lymphocytes’ ability to detect discrete pools of chemokine held deeply beneath the thick glycocalyx close to the endothelial plasma membrane (Shulman *et al.*, 2012). Thus, lymphocyte ILPs may function in both biomechanical and biochemical probing of the endothelial surface (as discussed further below). Importantly, such structures form in essentially every type of lymphocyte–endothelial interaction setting (e.g., bone marrow, thymus, HEVs, SLOs, and diverse inflamed tissues), including both intravasation and extravasation events (Sage and Carman, 2009; De Bruyn *et al.*, 1971; Becker and De Bruyn, 1976; Farr *et al.*, 1980; Azzali

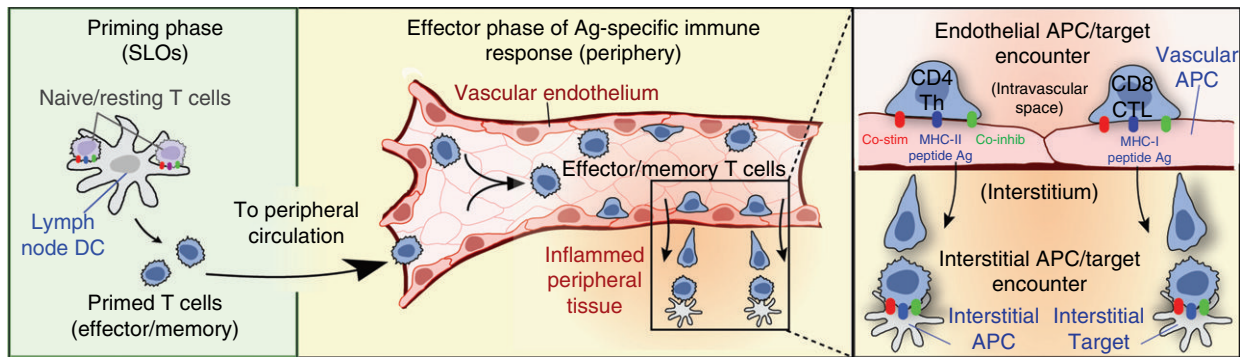


Figure 4 Endothelial cells: A peripheral ‘semi-professional’ APC compartment. While priming of naïve T lymphocytes occurs within SLO, such as LN, by professional APCs (e.g., DC; left panel), they then egress into the peripheral circulation and diapedese into inflamed tissues to interact with interstitial APCs and targets (middle and right panels). This progression requires direct, intimate interactions with the vascular endothelium. The endothelium is established as a ‘semi-professional’ APC (see Figure 5). Thus the first peripheral APC that lymphocytes are likely to encounter after leaving SLO are endothelial cells.

et al., 2008; Barron *et al.*, 1974; Astrom *et al.*, 1968; Lossinsky and Shivers, 2004; Raine *et al.*, 1990; Fujita *et al.*, 1991; Wolosewick, 1984; Wolburg *et al.*, 2005; Olah and Glick, 1985; Greenwood *et al.*, 1994; Bamforth *et al.*, 1997; De Bruyn *et al.*, 1989). This would suggest that ILPs represent an important sensory organelle that is broadly used by lymphocytes during their trafficking to probe their local cellular environment.

Information Exchange at the T Lymphocyte–Endothelial Interface

While the endothelium clearly plays an essential role (both by delivering biochemical signals and by actively remodeling) in driving appropriate immune cell migration, the same and/or additional endothelial cues can instruct immune cell activation and differentiation states. Indeed, studies with a wide range of innate and adaptive immune cells have collectively established that diapedesis across inflamed endothelium generally has a pro-inflammatory or ‘priming’ effect on these cells (Nourshargh and Marelli-Berg, 2005). For example, effector functions of neutrophils (e.g., degranulation and respiratory burst) and APC functions of monocytes and DCs (e.g., MHC-II expression) are both upregulated by extravasation *in vivo*, as well as by diapedesis across activated endothelia *in vitro*. Such studies suggest that the endothelium communicates to immune cells, in specific ways, that they are entering a new environment and helps them to adapt their differentiation and activation states to be prepared for this new setting.

Endothelial Function in Antigen-Specific Immunity

An important, but still incompletely understood, mode of EC communication with immune cells includes Ag-specific instruction of T lymphocytes through APC function. As discussed above, activation of initial responses in naïve T cells occurs in the T cell area of the lymph node (LN) through interactions with professional APCs (i.e., DCs). DCs not only efficiently take up, process and present Ag on MHC-I and -II (‘signal 1’), but also provide a wide range of potent co-stimulatory and co-

inhibitory molecules (that bind cognate receptors on the T cell; ‘signal 2’) and secrete critical cytokines (‘signal 3’). The appropriate combination of all three of these inputs are required to activate naïve T cells (Schwartz, 1992; Janeway and Bottomly, 1994; Paul and Seder, 1994; Germain, 1994; Guermontprez *et al.*, 2002). The quality of these signals shapes the strength and type of responses generated (e.g., pro-inflammatory Th1 versus anti-inflammatory/tolerogenic CD4 Treg). Once expanded and primed, effectors T cells reenter the circulation to home to peripheral sites of inflammation in search of cognate Ag on interstitial APCs and targets (Figure 4). Thus an obligate preceding step of the effector T cell is the adhesion to and diapedesis across the endothelium. As described below, ECs are now recognized as a type of ‘semi-professional’ APCs that might exert peripheral regulation over adaptive immunity at this stage.

Antigen presenting properties of endothelial cells

Like most cells, ECs constitutively express class I MHC molecules. Additionally, they also constitutively express MHC II, which are generally restricted to professional APCs (e.g., DCs) (Choi *et al.*, 2004; Pober *et al.*, 1983). Inflammatory cytokines further upregulate MHC I and II on endothelium (Choi *et al.*, 2004). Moreover, endothelial cells express Ag processing machinery (e.g., LMP2, 7, transporter associated with antigen processing 1 (TAP1), 2, invariant chain and HLA-DM) and have been shown to efficiently take up, process and present/cross-present Ag *in vitro* and *in vivo* (Greening *et al.*, 2003; Savage *et al.*, 1995; Vora *et al.*, 1994; von Oppen *et al.*, 2009; Schurich *et al.*, 2009; Limmer *et al.*, 2000; Hirose *et al.*, 2014). Additionally, ECs express a range of co-stimulatory (e.g., ICAM-1, VCAM-1, CD40, LFA3, ICOSL, 4-1BB, OX40L, TL1A) and co-inhibitory (e.g., programmed death ligand-1 (PD-L1, PD-L2)) molecules, as well as cytokines, both of which are strongly modulated by inflammatory cues (Choi *et al.*, 2004; Hollenbaugh *et al.*, 1995; Grabie *et al.*, 2007; Rodig *et al.*, 2003; Eppihimer *et al.*, 2002; Card *et al.*, 2014; Figure 5). Critically, however, ECs generally do not express CD80 (murine B7-1) or CD86 (B7-2) co-stimulators expressed by professional APCs that are necessary for activation of naïve lymphocytes (Figure 5). As a result, ECs are incapable of activating naïve

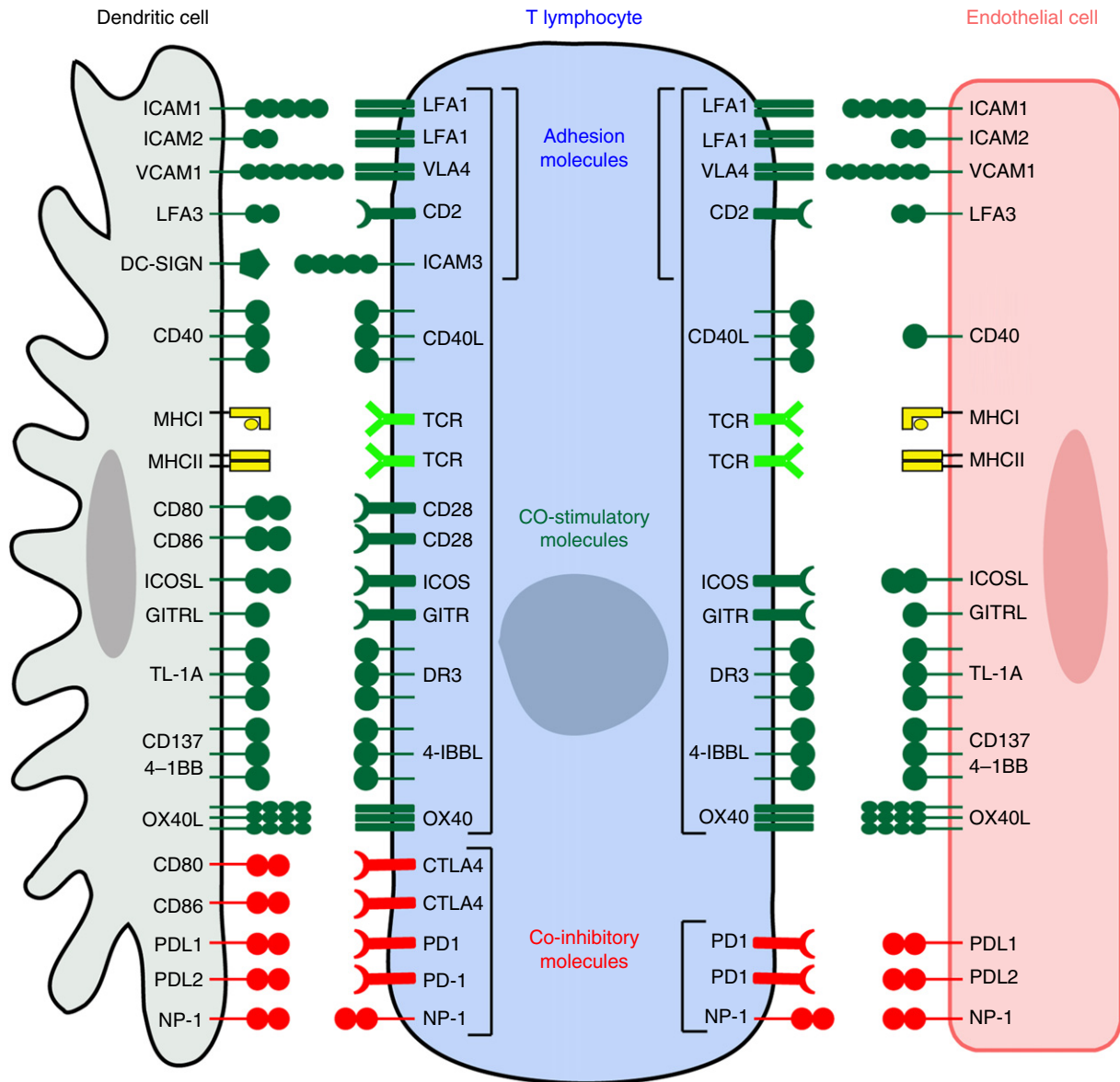


Figure 5 Endothelial cells are 'semi-professional' APCs, schematic depicts comparison of the Ag presentation, co-stimulatory, co-inhibitory and adhesion molecules expressed by DCs and ECs. Note that endothelial cells, unlike most other non-hematopoietic cells, express most of the critical molecules found in DCs. The notable exceptions include CD80 and CD86 that are critical for activation of naive T cells, as well as the co-stimulatory/adhesion molecules DC-SIGN.

lymphocytes. However, they can readily mediate Ag-specific stimulation of Ag-experienced (i.e., effector/memory) CD4 and CD8 lymphocytes and are therefore regarded as 'semi-professional' APCs (Marelli-Berg *et al.*, 1996; Ma and Pober, 1998; Perez *et al.*, 1998; Epperson and Pober, 1994; Dengler and Pober, 2000; Sage *et al.*, 2012).

Emerging and Putative Roles for Endothelial Cell Ag Presentation

The overall functions of ECs in Ag-specific immunity remain to be completely elucidated. Clearly, such functions will be ancillary to those of professional APCs (Choi *et al.*, 2004; Pober

and Sessa, 2007). Yet, the large collective surface area of the BEC-APC compartment of $\sim 6000 \text{ m}^2$ (which, in fact, outnumbers the professional APC compartment by ~ 1000 -fold (Mai *et al.*, 2013; Yano *et al.*, 2006)) and its frequent contacts with recirculating lymphocytes provide significant opportunity to influence Ag-specific responses. Moreover, the unique and strategic anatomical positioning of BECs would allow them to exert qualitatively distinct and locally tuned regulation within the periphery. Below we summarize most well-established APC functions of BECs in both healthy and pathologic immunity. These highly varied, and in some cases opposing (e.g., activation vs. tolerance/anergy), functions reflect highly heterogeneous contexts in which they develop.

Reactivation of effector/memory T cells

Many studies have demonstrated *in vitro* that human umbilical vein BECs (HUVEC; a classically used *in vitro* BEC model), as well as dermal and lung microvascular BECs can activate proliferation and inflammatory cytokine secretion (e.g., IFN- γ , IL-2, IL-4, IL-10) in resting CD4 memory cells in a manner that is enhanced by co-stimulators such as ICOS-L and CD40 (Marelli-Berg *et al.*, 1996; Ma and Pober, 1998; Perez *et al.*, 1998; Sage *et al.*, 2012; Khayyamian *et al.*, 2002). Similarly, effector/memory CD4 lymphocytes can be directly activated by allogeneic endothelium both *in vitro* and *in vivo* in an MHC-II-dependent manner (Shiao *et al.*, 2005; Pober, 1999). Additionally, presence of MHC-I bearing cognate Ag on BECs promotes CD8 CTL activation and EC killing, though these responses are partially mitigated by concomitant BEC expression co-inhibitors (i.e., PD-L1, PD-L2) (Grabie *et al.*, 2007; Sage *et al.*, 2012; Rodig *et al.*, 2003; Dengler and Pober, 2000). Studies such as these suggest a general ability of BECs to promote Ag-dependent pro-inflammatory responses as lymphocytes home to tissues.

Maintenance/promotion of tolerance

The idea that BECs can provide a form of peripheral tolerance has been evidenced in diverse settings (Marelli-Berg and Lechler, 1999). For example, early *in vitro* work introduced the idea of transmigration anergy, whereby Ag encounter during diapedesis in the context of insufficient co-stimulation, led to a state of nonresponsiveness to a second APC/Ag encounter (Kawai *et al.*, 2000). This concept has been particularly well developed in settings of liver sinusoidal endothelial cell (LSEC) and LECs.

LSECs

The liver is known to provide an important tolerogenic environment (Thomson and Knolle, 2010; Crispe, 2009). Among the essential APCs in this organ (i.e., liver DCs, Kupffer cells, and LSECs), LSECs exhibit unique tolerogenic functions (Thomson and Knolle, 2010; Crispe, 2009). At steady-state circulating effector lymphocytes are able to make repeated transient (~10–20 min) direct contacts with LSEC *in vivo* without the need for diapedesis (Carambia *et al.*, 2013). LSEC have been shown to possess potent scavenger functions (that are comparatively more efficient than those of liver DCs) (von Oppen *et al.*, 2009; Schurich *et al.*, 2009; Sorensen *et al.*, 2012) and can effectively cross-present exogenous Ag to CD8 T cells to promote tolerance (Limmer *et al.*, 2000). LSEC also suppress pro-inflammatory Th1 and Th17 responses in an Ag-specific and IL-10- and PD-L1-dependent manner (Carambia *et al.*, 2013; Knolle *et al.*, 1999; Klugewitz *et al.*, 2002). Finally, LSEC exhibit superior efficacy (i.e., as compared to liver DCs and Kupffer cells) in promoting differentiation of CD4 Tregs that suppress inflammation and induce tolerance. This is accomplished through the unique composition of the LSEC glycocalyx (which includes the glycoprotein GARP) that uniquely enables cell surface capture and presentation of the critical Treg cytokine TGF- β (Carambia *et al.*, 2014).

LECs

LECs are also emerging as an important player in shaping immunity and tolerance (Tewalt *et al.*, 2012a; Card *et al.*,

2014). Lymphatic vessels and the contiguous subcapillary, cortical, and medullary sinuses of the SLOs provide a conduit for movement of lymphocytes. LECs of the lymph node sinuses constitutively express MHC-I and -II and under noninflamed conditions take up and cross-present peripheral tissue Ags to CD8 T cells in a lysosome- and TAP1-dependent manner (Hirose *et al.*, 2014). As a result of the absent co-stimulatory and high co-inhibitory (i.e., PD-L1) molecule expression on these LECs, such cross-presentation induces abortive proliferation and clonal deletion of self-reactive CD8 lymphocytes (Cohen *et al.*, 2010; Tewalt *et al.*, 2012b). Since lymphocytes must cross the cortical and/or medullary LECs in order to egress from the lymph node, these ECs are proposed to serve as anatomical checkpoint to enforce self-tolerance in lymphocytes attempting to exit the lymph node (Tewalt *et al.*, 2012a). While MHC-II on LECs is functional (as evidenced by the ability of peptide pulsed LECs to induce proliferation of CD4 cells) the overall roles of these APCs in class II immune responses remain to be determined (Tewalt *et al.*, 2012a; Card *et al.*, 2014).

Altered trafficking

A range of studies have demonstrated influences of EC Ag presentation on T cell trafficking behaviors. For instance, recognition of MHC-II/cognate Ag presented on EC was shown to cause a migratory stop signal in CD4 T cells that could prevent or transiently delay diapedesis across EC *in vitro* (Manes *et al.*, 2007; Sage *et al.*, 2012; Tay *et al.*, 2003). Other *in vitro* and *in vivo* studies support the intriguing hypothesis that Ag presented on the endothelium promotes the selective diapedesis of Ag-reactive CD4 and CD8 T cells. This seems in part to be mediated through an ability of TCR signaling events (i.e., activation of tyrosine kinase Zap-70) to cross talk with cell migration machinery (e.g., Rac-1, myosin-IIA) (Manes and Pober, 2008, 2013). In this way, Ag-specific effectors may more effectively home to the original tissue source Ag where pathogens have infected (Marelli-Berg *et al.*, 1999, 2004, 2007; Savinov *et al.*, 2003; Manes and Pober, 2008). It has been shown that the specific co-stimulatory/co-inhibitory molecules' expression patterns on the endothelium can strongly influence this process (David and Marelli-Berg, 2007). Additionally, recent studies demonstrate that recruitment of Treg to sites of inflammation is strongly favored by their ability to recognize self-Ags presented by endothelium (Fu *et al.*, 2014). As a result, the normally low numbers of Tregs that exist in the body can be optimally recruited to ensure sufficient numbers to balance out pro-inflammatory effector lymphocytes during resolution of inflammation (Fu *et al.*, 2014).

Roles for Endothelial Antigen Presentation in Pathology

Although roles in protective immunity are still emerging, strong connections of EC Ag presentation to immune pathology have been established. For example, endothelial expression of MHC-II molecules is significantly upregulated in a range of autoimmune/inflammatory diseases including multiple sclerosis, rheumatoid arthritis, diabetes, lupus, Crohn's disease, allograft rejection, dilated cardiomyopathy, myocarditis, and vasculitis (Turesson, 2004; Barkley *et al.*, 1989).

Additionally, alteration in the endothelial expression of critical co-stimulatory molecules and co-inhibitory molecules (e.g., CD40 and PD-L1) has been implicated in the pathogenesis of autoimmune diseases (Hollenbaugh *et al.*, 1995; Grabie *et al.*, 2007; Rodig *et al.*, 2003; Eppihimer *et al.*, 2002). Particularly, compelling functional evidence for a pathogenic role of EC Ag presentation exists in:

Allograft rejection

Perhaps the clearest evidence for functional roles of EC MHC is seen in the extensive *in vitro*, *in vivo* and clinical investigation of transplantation and allo-responses (Denton *et al.*, 2000; Pober, 1999; Taflin *et al.*, 2011; Rose, 1998). Effector/memory CD4 and CD8 lymphocytes are well-evidenced to be activated directly by allogeneic endothelium *in vitro* (Shiao *et al.*, 2005; Pober, 1999). Similarly, *in vivo* the endothelium of solid organ transplants (the first cell surfaces that host lymphocyte encounter) plays a pivotal MHC-dependent role in stimulating allo-responses and also can become a major target for CD8-mediated destruction. Moreover, in studies examining graft versus host disease, associated with stem cell transplantation, allo-reactive donor CD8 CTLs were demonstrated to be directly responsible for endothelial injury (Biedermann *et al.*, 2002).

Multiple sclerosis

Early studies using an animal model of MS (Experimental Autoimmune Encephalitis) demonstrated that MHC II expression on brain microvascular endothelium preceded the formation of T cell infiltrates, suggesting a role for EC Ag presentation in this process (Sobel *et al.*, 1984). Subsequent studies demonstrated that MHC-I/Ag presented on the luminal surface of the blood–brain barrier was responsible for Ag-specific trafficking of T cells to the CNS (Galea *et al.*, 2007). *In vitro* studies have also demonstrated that human brain ECs constitutively co-express of MHC-IMHC II, CD40, and ICOSL, take up fluorescently labeled Ags via macropinocytosis and promote Ag-dependent proliferation of CD4 and CD8 T cells (Wheway *et al.*, 2013).

Diabetes

In vitro studies have demonstrated that human ECs can process the diabetes type I islet autoantigen GAD65 and functionally present it on MHC-II molecules to promote selective transmigration of Ag-specific Th1 lymphocytes (Greening *et al.*, 2003). Similarly, using murine model of type-I diabetes, insulin-specific CD8 lymphocytes were demonstrated to home to the pancreas in a manner that was dependent on endothelial presentation *in vivo* (Savinov *et al.*, 2003).

Bacterial super-antigen

Bacterial super-antigens (SAGs) can functionally mimic ‘high affinity’ agonistic peptide Ag by concomitantly binding (and effectively ‘bridging’) cell surface MHC-II and TCR independently of the peptide binding sites (Redpath *et al.*, 1999; Proft and Fraser, 2003). SAGs are effective tools (model Ags) that are widely used to study T cell activation. More importantly, these are also pathogens in a wide range of diseases, including arthritis, diabetes, multiple sclerosis/EAE, sepsis, toxic shock, Kawasaki disease, and idiopathic dilated cardiomyopathy that accumulate (Renno and Acha-Orbea, 1996). During many infections,

significant levels of SAGs appear in the blood where they can readily be captured on EC MHC-II and presumably signal circulating T cells. Compelling evidence exists, in at least one setting (Kawasaki disease) (Leung *et al.*, 1995), that SAG presented specifically by endothelial MHC-II is important for pathogenesis.

Antigen Recognition and Responses during Lymphocyte–Endothelial Interactions

ISs with APC models and professional APCs

It has become clear that many of the keys to understanding protective and pathologic Ag-specific responses are in understanding the fundamental cell biology of the T cell–APC interaction (Dustin, 2005). This has been intensely investigated for activation of lymphocytes by professional APCs or artificial APC models. Such studies show that initially, Ag recognition delivers a calcium-dependent migratory stop signal that allows formation of a more sustained interaction with the APC (i.e., IS) in which micron-scale clusters of TCR mediate active signaling at the periphery of the contact (Dustin *et al.*, 1997). Studies have also identified several cytoskeletal modulators (Cdc42, WASp, WAVE2, Vav1, Arp2/3, HS1) that are critical for IS formation (Huang and Burkhardt, 2007; Billadeau *et al.*, 2007).

T Lymphocyte ILP: Unique Ag-recognition sensory organelles

Detailed investigation of the coordinated cellular, molecular and signaling dynamics of responses to endothelial APCs are only beginning to emerge. In one study lung and dermal microvascular BECs presenting MHC-II/cognate Ag were shown to induce proliferation and stimulate INF- γ production in resting CD4 effector/memory lymphocytes (Sage *et al.*, 2012). Detailed dynamic imaging studies showed that such responses were initiated by a rapid flux of intercellular calcium that is coupled to a migratory stop signal, which was initiated within tens of seconds of contacting BEC bearing MHC-II. This was followed by nuclear translocation of NFAT within minutes (Sage *et al.*, 2012). Calcium signaling persisted for about 30–60 min, after which T cells resumed lateral migration and then underwent diapedesis, still bearing nuclear NFAT. Closer characterization of the subcellular dynamics and architecture of the T cell–BEC IS showed that T cells initiated ILP probing that enforced extremely close contacts with EC that preceded, and were required for, Ag-dependent calcium responses (Figure 6). The same was found for CD8 CTL calcium signaling and killing responses to EC presentation of MHC-I/cognate Ag (Sage *et al.*, 2012). Thus, the roles discussed of above for ILP in promoting efficient ‘informational scanning’ for EC surface chemokines (Shulman *et al.*, 2012) is also relevant for detection of MHC/Ag.

The critical role for ILP in Ag recognition is evident when we consider the topological features of the requisite cell–cell interaction. The surface of all cells are modified by a gel-like polysaccharide coating, termed the glycocalyx, that is made up of diverse proteoglycans and glycoproteins (Reitsma *et al.*, 2007; Weinbaum *et al.*, 2007). On a typical cell the glycocalyx is at least 45 nm thick, but may be substantially thicker (up to ~500 nm) on others, such as the endothelium (Reitsma *et al.*,

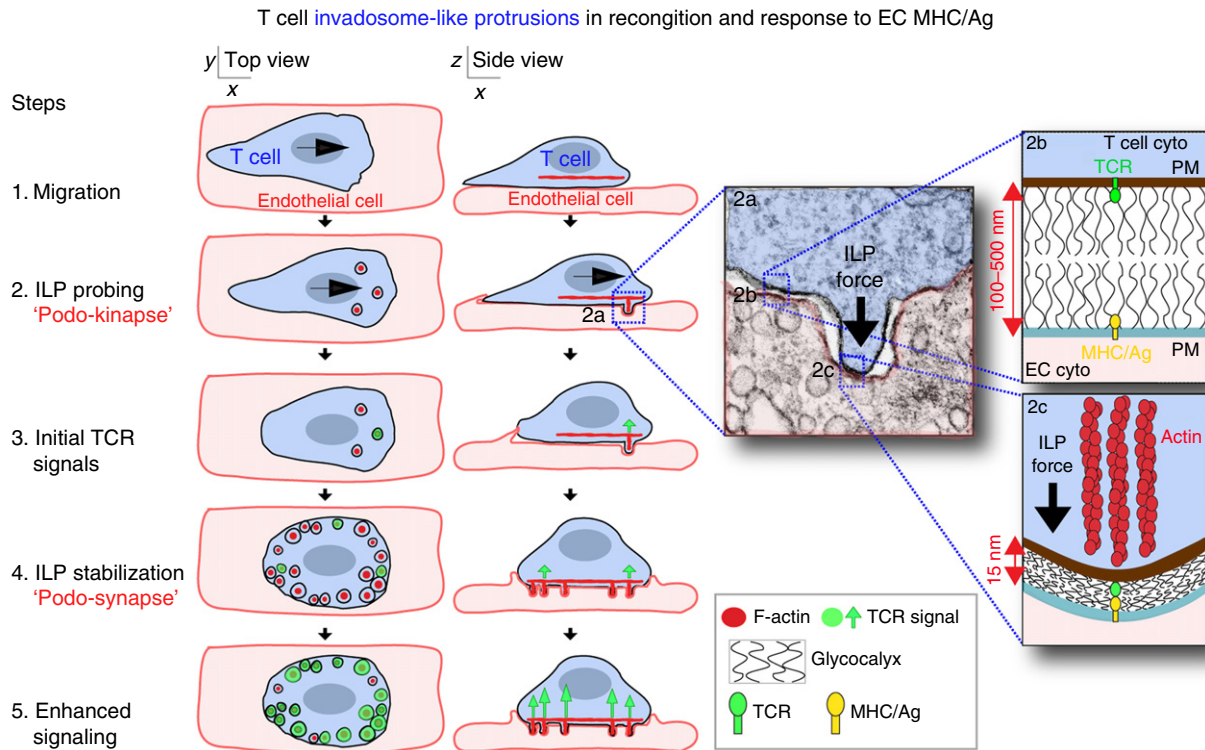


Figure 6 Model for ILPs function in Ag recognition and response schematic depicts top, down and side views of a memory/effector T cell (blue) interacting with an endothelium (pink) presenting cognate Ag. Lymphocytes initiate lateral migration (Step 1) and begin to dynamically drive ILPs against the opposing cell (Step 2, inset 2a). Close interactions between T and APC/target cells, which are partially opposed by the cell glycocalyx (inset 2b), form at ILP tips (inset 2c). Thus TCR/MHC interactions may be facilitated in these zones. Initial calcium released upon Ag recognition (green overlay/arrows; Step 3) seems to be coupled to stabilization/accumulation of ILP arrays ('podo-synapses'; Step 4), help sustain/enhance signaling (Step 5). The peripheral arrays of ILP seen in this setting as highly analogous to podosome belts that serve as a type cell–cell sealing apparatus for compartmentalized delivery of extracellular material, which in this case may include cytokine or cytotoxic agents released by the T cells.

2007; Weinbaum *et al.*, 2007). The glycocalyx serves several important functions, such as protection from proteases and stabilizing the orientation of adhesion and signaling receptors (Rudd *et al.*, 2001). At the same time the glycocalyx also provides a formidable energy barrier to close membrane–membrane encounter between cells, such that relatively small cell surface adhesion and signaling molecules (e.g., TCR and MHC, each ~7 nm tall (Rudolph *et al.*, 2006)) are effectively shielded (Bell, 1978; Bell *et al.*, 1984; Springer, 1990; Reitsma *et al.*, 2007; Weinbaum *et al.*, 2007; Figure 6). In this way, the glycocalyx negatively modulates immune cell adhesion and immune recognition (Gubbels *et al.*, 2010; Komatsu *et al.*, 1999; Tsuboi and Fukuda, 2001; Tsuboi and Fukuda, 1998; van de Wiel-van Kemenade *et al.*, 1993; Sabri *et al.*, 2000; Manjunath *et al.*, 1995; Stockton *et al.*, 1998; Onami *et al.*, 2002; McFarland *et al.*, 1995). Forces provided by ILPs appear to provide sufficient energy to overcome this barrier, driving close membrane–membrane apposition (Figure 6) and thereby promoting molecular interactions that might otherwise be inefficient or impossible.

Though the functional contribution of ILP have only recently been elucidated, it was previously established that the leading edge lamellipodia of migrating lymphocytes possess heightened Ag recognition capabilities (Valitutti *et al.*, 1995;

Negulescu *et al.*, 1996; Wei *et al.*, 1999). It would now appear that the ILP, which from preferentially within lamellipodia (Carman *et al.*, 2007; Sage *et al.*, 2012; Martinelli *et al.*, 2014), represent the direct 'actuators' of the special immune surveillance capabilities found at the T cell leading edge.

Stabilized ILP arrays function in signal amplification

Following initial Ag recognition, the resulting intracellular calcium flux concomitantly drives a migratory stop signal and stabilization of ILPs. This causes ILPs to accumulate preferentially around the periphery of the T cell–EC interface in dense arrays forming a unique IS topology termed a 'podo-synapse' (Sage *et al.*, 2012; Figure 6). The individual Ag stabilized lymphocyte ILPs are enriched in actin, TCR and molecules suggestive of active signaling (e.g., PKC- θ , phospho-tyrosine and HS1), whereas the cognate endothelial contact sites (i.e., podo-prints) are enriched in MHC (Figure 7). Thus, ILPs may function to promote sustained TCR signaling by providing specialized sub-micron-scale cylindrical volumes to increase signaling efficiency in analogy to well-developed concepts for 'signalosomes' (Burbach *et al.*, 2007; Perez-Moreno *et al.*, 2003; Harwood and Batista, 2008; Vicente-Manzanares and Sanchez-Madrid, 2004; Miranti and Brugge, 2002). This feature is generally not unlike TCR signaling micro-clusters defined using

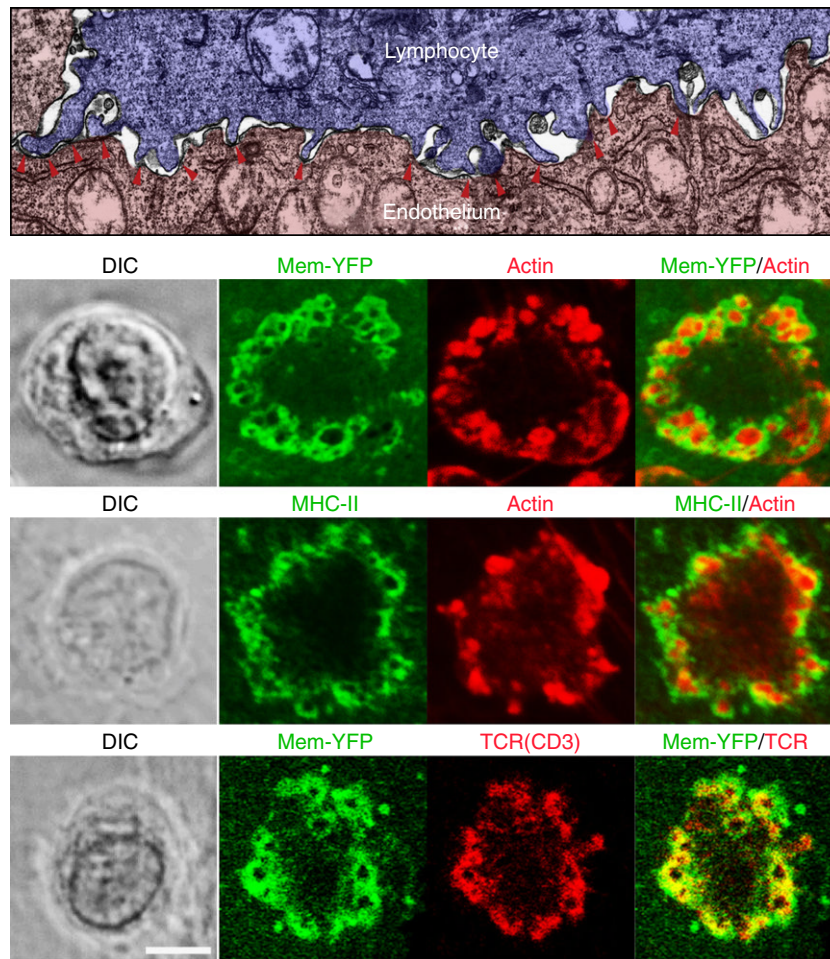


Figure 7 Ultrastructure and fluorescence imaging of the podosynapse images show CD4 effector/memory lymphocytes interacting with microvascular endothelial cells presenting MHC-II/cognate Ag. Upper panels show ultrastructural imaging of a lymphocyte (blue) extending multiple ILP into the endothelium, generating zones of close cell–cell interaction, and EC deformation (i.e., creating multiple EC ‘podo-prints’). Lower panels depict light microscopy imaging of T cells (by differential interference contrast (DIC) and fluorescence imaging). T cell ILP is enriched in actin and TCR (Red). Podo-prints in endothelium are evidenced as rings of plasma membrane (mem-YFP) and are enriched in MHC-II (green).

planar lipid bilayer APC model substrates (Dustin *et al.*, 2006; Saito and Yokosuka, 2006; Brossard *et al.*, 2005; Dustin, 2009; Seminario and Bunnell, 2008; Gomez *et al.*, 2006; Yokosuka *et al.*, 2005; Bunnell *et al.*, 2002) and ‘multi-focal’ ISs noted to form on DC (Dustin *et al.*, 2006; Saito and Yokosuka, 2006; Brossard *et al.*, 2005).

At a higher level of organization, the ILP arrays may provide additional functions in promoting T cell–EC information exchange. Specifically, the podo-synapse exhibits striking topological similarity to osteoclast ‘podosome-belts’ (Luxenburg *et al.*, 2007; Destaing *et al.*, 2008; Figure 6), structures that form sealing zones for directed secretion of bone degrading enzymes (Saltel *et al.*, 2004). Thus, an additional function of the podo-synapse may be similarly to direct secretion of cytokines (or cytotoxic agents) toward the target cell, while limiting bystander effects.

Lessons learned from BEC-APC as unique planar APC

The preceding discussion highlights ILP functions that could be envisioned as being broadly relevant ones. Whereas

commonly used APC models (planar lipid bilayer or coated glass/plastic model APCs substrates) provide ideal features (i.e., a planar surface oriented parallel to the optimal x-y imaging plane) for visualizing micron-scale topological dynamics, their inherently rigid nature would most likely frustrate and/or mask ILP extension in the z-plane. Alternatively, imaging of most physiologic and deformable APCs/targets is hindered by orientation issues that profoundly limit resolution (Oddos *et al.*, 2008; Balagopalan *et al.*, 2011; Cebecauer *et al.*, 2010). As an experimental model, BEC combine the strengths (i.e., near planarity and orientation) of model substrates and cellular APCs (deformability and active responses) forming a ‘planar cellular APC’ surface that is optimized for detection of ILPs, but not necessarily the only setting for their formation. Cross-sectional fluorescence and EM images of ISs formed with a range of (non-endothelial) APCs and targets show topologies consistent with presence ILPs (Bunnell *et al.*, 2006; Yokosuka *et al.*, 2005; Wetzel *et al.*, 2002; Butler *et al.*, 2008; Barcia *et al.*, 2008a,b, 2009; Davis, 2009; Sage *et al.*, 2012). Thus, ILP may

have broad function in Ag recognition and responses that are uniquely revealed the setting of EC Ag presentation, rather being a unique feature of this setting.

Summary

The endothelium represents a critical anatomical partition between the tissues and the blood/lymph system and also play essential proactive sentinel roles that regulate immunity. In particular, BECs and LECs act as coordinators of immune cell trafficking during maturation, surveillance and responses. Additionally, BECs and LECs co-express MHC-I, MHC-II and a wide array of co-stimulatory molecules and exhibit ‘semi-professional’ APC functions that can communicate Ag-specific information to effector/memory T cells in the periphery. Although these functions, by definition, will be secondary to those of professional APCs, the enormous scale and strategic anatomical positioning of the endothelial APC compartment suggest it may have important and unique functions in adaptive immunity, as well as immune pathologies. The ability of lymphocyte-endothelial interactions to so effectively orchestrate information exchange, trafficking events and Ag-specific responses is coupled with the mutual ability of these cells to engage in extremely intimate and dynamically coordinated interactions. Active embrace of T cells by endothelial cup-like structures promotes enhanced adhesion and guidance for diapedesis, while EC ventral lamellipodia restore barrier function following extravasation events. T cell ILPs have been uncovered as a critical sensory organelle that allows T cells to effectively survey both the biomechanical and biochemical properties of the endothelium (and likely other cells). In this way, ILP are both critical for T cell migratory pathfinding across the endothelial barrier (i.e., ‘tender-taxis’) and for detection of MHC-Ag on the BEC surface. With respect to the latter, stabilized arrays of ILPs (‘podo-synapses’) provides a unique IS architecture that functionally contributes to immune signaling at this interface in several ways. The preceding suggests that more clearly defining the diverse context-dependent roles of EC in Ag-specific immunity and the underlying cellular/molecular regulation represents areas of significant biomedical significance.

Supplementary Material

The following are the supplementary data to this article: Video 1.

Acknowledgment

This work was supported by a grant from the NIH (HL 104006 to CVC).

See also: Cellular Immunology: Cell–Cell Interactions and the Immune System: Roles of Stromal Cells in the Immune System

References

- Aird, W.C., 2005. Spatial and temporal dynamics of the endothelium. *Journal of Thrombosis and Haemostasis* 3, 1392–1406.
- Aird, W.C., 2007a. Phenotypic heterogeneity of the endothelium: I. Structure, function, and mechanisms. *Circulation Research* 100, 158–173.
- Aird, W.C., 2007b. Phenotypic heterogeneity of the endothelium: II. Representative vascular beds. *Circulation Research* 100, 174–190.
- von Andrian, U.H., Mackay, C.R., 2000. T-cell function and migration. Two sides of the same coin. *New England Journal of Medicine* 343, 1020–1034.
- von Andrian, U.H., Mempel, T.R., 2003. Homing and cellular traffic in lymph nodes. *Nature Reviews Immunology* 3, 867–878.
- Astrom, K.E., Webster, H.D., Arnason, B.G., 1968. The initial lesion in experimental allergic neuritis. A phase and electron microscopic study. *Journal of Experimental Medicine* 128, 469–495.
- Audoy-Remus, J., Richard, J.F., Soulet, D., *et al.*, 2008. Rod-Shaped monocytes patrol the brain vasculature and give rise to perivascular macrophages under the influence of proinflammatory cytokines and angiotensin-2. *Journal of Neuroscience* 28, 10187–10199.
- Auffray, C., Fogg, D., Garfa, M., *et al.*, 2007. Monitoring of blood vessels and tissues by a population of monocytes with patrolling behavior. *Science* 317, 666–670.
- Azzali, G., Arcari, M.L., Caldara, G.F., 2008. The “mode” of lymphocyte extravasation through HEV of Peyer’s patches and its role in normal homing and inflammation. *Microvascular Research* 75, 227–237.
- Bajenoff, M., Egen, J.G., Qi, H., *et al.*, 2007. Highways, byways and breadcrumbs: Directing lymphocyte traffic in the lymph node. *Trends in Immunology* 28, 346–352.
- Bajoghli, B., Aghaallaei, N., Hess, I., *et al.*, 2009. Evolution of genetic networks underlying the emergence of thymopoiesis in vertebrates. *Cell* 138, 186–197.
- Balogopalan, L., Sherman, E., Barr, V.A., Samelson, L.E., 2011. Imaging techniques for assaying lymphocyte activation in action. *Nature Reviews Immunology* 11, 21–33.
- Baluk, P., Fuxe, J., Hashizume, H., *et al.*, 2007. Functionally specialized junctions between endothelial cells of lymphatic vessels. *Journal of Experimental Medicine* 204, 2349–2362.
- Barnforth, S.D., Lightman, S.L., Greenwood, J., 1997. Ultrastructural analysis of interleukin-1 beta-induced leukocyte recruitment to the rat retina. *Investigative Ophthalmology & Visual Science* 38, 25–35.
- Barcia, C., Gomez, A., DePablos, V., *et al.*, 2008a. CD20, CD3, and CD40 ligand microclusters segregate three-dimensionally *in vivo* at B-cell-T-cell immunological synapses after viral immunity in primate brain. *Journal of Virology* 82, 9978–9993.
- Barcia Jr., C., Gomez, A., Gallego-Sanchez, J.M., *et al.*, 2009. Infiltrating CTLs in human glioblastoma establish immunological synapses with tumorigenic cells. *American Journal of Pathology* 175, 786–798.
- Barcia, C., Sanderson, N.S., Barrett, R.J., *et al.*, 2008b. T cells’ immunological synapses induce polarization of brain astrocytes *in vivo* and *in vitro*: A novel astrocyte response mechanism to cellular injury. *PLoS One* 3, e2977.
- Barkley, D., Allard, S., Feldmann, M., Maini, R.N., 1989. Increased expression of HLA-DQ antigens by interstitial cells and endothelium in the synovial membrane of rheumatoid arthritis patients compared with reactive arthritis patients. *Arthritis & Rheumatology* 32, 955–963.
- Barreiro, O., Yanez-mo, M., Serrador, J.M., *et al.*, 2002. Dynamic interaction of VCAM-1 and ICAM-1 with moesin and ezrin in a novel endothelial docking structure for adherent leukocytes. *Journal of Cell Biology* 157, 1233–1245.
- Barron, K.D., Means, E.D., Feng, T., Harris, H., 1974. Ultrastructure of retrograde degeneration in thalamus of rat. 2. Changes in vascular elements and transvascular migration of leukocytes. *Experimental and Molecular Pathology* 20, 344–362.
- Bartholomaeus, I., Kawakami, N., Odoardi, F., *et al.*, 2009. Effector T cell interactions with meningeal vascular structures in nascent autoimmune CNS lesions. *Nature* 462, 94–98.
- Bazzoni, G., Dejana, E., 2004. Endothelial cell-to-cell junctions: Molecular organization and role in vascular homeostasis. *Physiological Reviews* 84, 869–901.
- Becker, R.P., De Bruyn, P.P., 1976. The transmural passage of blood cells into myeloid sinusoids and the entry of platelets into the sinusoidal circulation; a scanning electron microscopic investigation. *American Journal of Anatomy* 145, 183–205.
- Bell, G.I., 1978. Models for the specific adhesion of cells to cells. *Science* 200, 618–627.
- Bell, G.I., Dembo, M., Bongrand, P., 1984. Cell adhesion. Competition between nonspecific repulsion and specific bonding. *Biophysical Journal* 45, 1051–1064.

- Biedermann, B.C., Sahner, S., Gregor, M., *et al.*, 2002. Endothelial injury mediated by cytotoxic T lymphocytes and loss of microvessels in chronic graft versus host disease. *Lancet* 359, 2078–2083.
- Billadeau, D.D., Nolz, J.C., Gomez, T.S., 2007. Regulation of T-cell activation by the cytoskeleton. *Nature Reviews Immunology* 7, 131–143.
- Brossard, C., Feuillet, V., Schmitt, A., *et al.*, 2005. Multifocal structure of the T cell – Dendritic cell synapse. *European Journal of Immunology* 35, 1741–1753.
- Bunnell, S.C., Hong, D.I., Kardon, J.R., *et al.*, 2002. T cell receptor ligation induces the formation of dynamically regulated signaling assemblies. *Journal of Cell Biology* 158, 1263–1275.
- Bunnell, S.C., Singer, A.L., Hong, D.I., *et al.*, 2006. Persistence of cooperatively stabilized signaling clusters drives T-cell activation. *Molecular and Cellular Biology* 26, 7155–7166.
- Burbach, B.J., Medeiros, R.B., Mueller, K.L., Shimizu, Y., 2007. T-cell receptor signaling to integrins. *Immunological Reviews* 218, 65–81.
- Butcher, E.C., 1991. Leukocyte-endothelial cell recognition: Three (or more) steps to specificity and diversity. *Cell* 67, 1033–1036.
- Butler, B., Kastendieck, D.H., Cooper, J.A., 2008. Differently phosphorylated forms of the cortactin homolog HS1 mediate distinct functions in natural killer cells. *Nature Immunology* 9, 887–897.
- Carambia, A., Frenzel, C., Bruns, O.T., *et al.*, 2013. Inhibition of inflammatory CD4 T cell activity by murine liver sinusoidal endothelial cells. *Journal of Hepatology* 58, 112–118.
- Carambia, A., Freund, B., Schwinge, D., *et al.*, 2014. TGF-beta-dependent induction of CD4CD25Foxp3 Tregs by liver sinusoidal endothelial cells. *Journal of Hepatology* 61, 594–599.
- Card, C.M., Yu, S.S., Swartz, M.A., 2014. Emerging roles of lymphatic endothelium in regulating adaptive immunity. *Journal of Clinical Investigation* 124, 943–952.
- Carman, C.V., 2009a. Mechanisms for transcellular diapedesis: Probing and pathfinding by 'invadosome-like protrusions'. *Journal of Cell Science* 122, 3025–3035.
- Carman, C.V., 2009b. Transmigratory cups and invadosome-like protrusions: New aspects of diapedesis. In: Ley, K. (Ed.), *Leukocyte Adhesion*. Current Topics in Membranes 64. Amsterdam: Elsevier, pp. 297–333.
- Carman, C.V., Jun, C.D., Salas, A., Springer, T.A., 2003. Endothelial cells proactively form microvilli-like membrane projections upon intercellular adhesion molecule 1 engagement of leukocyte LFA-1. *Journal of Immunology* 171, 6135–6144.
- Carman, C.V., Sage, P.T., Sciuto, T.E., *et al.*, 2007. Transcellular diapedesis is initiated by invasive podosomes. *Immunity* 26, 784–797.
- Carman, C.V., Springer, T.A., 2004. A transmigratory cup in leukocyte diapedesis both through individual vascular endothelial cells and between them. *Journal of Cell Biology* 167, 377–388.
- Cebecauer, M., Spitaler, M., Serge, A., Magee, A.I., 2010. Signalling complexes and clusters: Functional advantages and methodological hurdles. *Journal of Cell Science* 123, 309–320.
- Chamberlain, J.K., Lichtman, M.A., 1978. Marrow cell egress: Specificity of the site of penetration into the sinus. *Blood* 52, 959–968.
- Choi, J., ENIS, D.R., Koh, K.P., Shiao, S.L., Pober, J.S., 2004. T lymphocyte-endothelial cell interactions. *Annual Review of Immunology* 22, 683–709.
- Choi, K., Kennedy, M., Kazarov, A., Papadimitriou, J.C., Keller, G., 1998. A common precursor for hematopoietic and endothelial cells. *Development* 125, 725–732.
- Cohen, J.N., Guidi, C.J., Tewalt, E.F., *et al.*, 2010. Lymph node-resident lymphatic endothelial cells mediate peripheral tolerance via Aire-independent direct antigen presentation. *Journal of Experimental Medicine* 207, 681–688.
- Crispe, I.N., 2009. The liver as a lymphoid organ. *Annual Review of Immunology* 27, 147–163.
- David, R., Marelli-Berg, F.M., 2007. Regulation of T-cell migration by co-stimulatory molecules. *Biochemical Society Transactions* 35, 1114–1118.
- Davis, D.M., 2009. Mechanisms and functions for the duration of intercellular contacts made by lymphocytes. *Nature Reviews Immunology* 9, 543–555.
- De Bruyn, P.P., Cho, Y., Michelson, S., 1989. Endothelial attachment and plasmalemmal apposition in the transcellular movement of intravascular leukemic cells entering the myeloid parenchyma. *American Journal of Anatomy* 186, 115–126.
- De Bruyn, P.P., Michelson, S., Thomas, T.B., 1971. The migration of blood cells of the bone marrow through the sinusoidal wall. *Journal of Morphology* 133, 417–437.
- Dengler, T.J., Pober, J.S., 2000. Human vascular endothelial cells stimulate memory but not naive CD8 + T cells to differentiate into CTL retaining an early activation phenotype. *Journal of Immunology* 164, 5146–5155.
- Denton, M.D., Davis, S.F., Baum, M.A., *et al.*, 2000. The role of the graft endothelium in transplant rejection: Evidence that endothelial activation may serve as a clinical marker for the development of chronic rejection. *Pediatric Transplantation* 4, 252–260.
- Destaing, O., Sanjay, A., Itzstein, C., *et al.*, 2008. The tyrosine kinase activity of c-Src regulates actin dynamics and organization of podosomes in osteoclasts. *Molecular Biology of the Cell* 19, 394–404.
- Dustin, M.L., 2003. Coordination of T cell activation and migration through formation of the immunological synapse. *Annals of the New York Academy of Sciences* 987, 51–59.
- Dustin, M.L., 2005. A dynamic view of the immunological synapse. *Seminars in Immunology* 17, 400–410.
- Dustin, M.L., 2009. The cellular context of T cell signaling. *Immunity* 30, 482–492.
- Dustin, M.L., Bromley, S.K., Kan, Z., Peterson, D.A., Unanue, E.R., 1997. Antigen receptor engagement delivers a stop signal to migrating T lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America* 94, 3909–3913.
- Dustin, M.L., Tseng, S.Y., Varma, R., Campi, G., 2006. T cell-dendritic cell immunological synapses. *Current Opinion in Immunology* 18, 512–516.
- Epperson, D.E., Pober, J.S., 1994. Antigen-presenting function of human endothelial cells. Direct activation of resting CD8 T cells. *Journal of Immunology* 153, 5402–5412.
- Eppihimer, M.J., Gunn, J., Freeman, G.J., *et al.*, 2002. Expression and regulation of the PD-L1 immunoinhibitory molecule on microvascular endothelial cells. *Microcirculation* 9, 133–145.
- Ernst, U., Glynn, L., Larsson, B., 1965. Venous output of lymphocytes from the thymus. *Nature* 207, 540–541.
- Farr, A.G., Cho, Y., De Bruyn, P.P., 1980. The structure of the sinus wall of the lymph node relative to its endocytic properties and transmural cell passage. *American Journal of Anatomy* 157, 265–284.
- Fooksman, D.R., Vardhana, S., Vasiliver-Shamis, G., *et al.*, 2010. Functional anatomy of T cell activation and synapse formation. *Annual Review of Immunology* 28, 79–105.
- Fu, H., Kishore, M., Gittens, B., *et al.*, 2014. Self-recognition of the endothelium enables regulatory T-cell trafficking and defines the kinetics of immune regulation. *Nature Communications* 5, 3436.
- Fujita, S., Puri, R.K., Yu, Z.X., Travis, W.D., Ferrans, V.J., 1991. An ultrastructural study of *in vivo* interactions between lymphocytes and endothelial cells in the pathogenesis of the vascular leak syndrome induced by interleukin-2. *Cancer* 68, 2169–2174.
- Galea, I., Bernardes-Silva, M., Forse, P.A., *et al.*, 2007. An antigen-specific pathway for CD8 T cells across the blood–brain barrier. *Journal of Experimental Medicine* 204, 2023–2030.
- Geissmann, F., Cameron, T.O., Sidobre, S., *et al.*, 2005. Intravascular immune surveillance by CXCR6 + NKT cells patrolling liver sinusoids. *PLoS Biology* 3, e113.
- Gerard, A., Van Der Kammen, R.A., Janssen, H., Ellenbroek, S.I., Collard, J.G., 2009. The Rac activator Tiam1 controls efficient T-cell trafficking and route of transendothelial migration. *Blood* 113, 6138–6147.
- Germain, R.N., 1994. MHC-dependent antigen processing and peptide presentation: Providing ligands for T lymphocyte activation. *Cell* 76, 287–299.
- Girard, J.P., Moussion, C., Forster, R., 2012. HEVs, lymphatics and homeostatic immune cell trafficking in lymph nodes. *Nature Reviews Immunology* 12, 762–773.
- Goebeler, M., Yoshimura, T., Toksoy, A., *et al.*, 1997. The chemokine repertoire of human dermal microvascular endothelial cells and its regulation by inflammatory cytokines. *Journal of Investigative Dermatology* 108, 445–451.
- Gomez, T.S., McCarney, S.D., Carrizosa, E., *et al.*, 2006. HS1 functions as an essential actin-regulatory adapter protein at the immune synapse. *Immunity* 24, 741–752.
- Gorfu, G., Rivera-Nieves, J., Ley, K., 2009. Role of beta7 integrins in intestinal lymphocyte homing and retention. *Current Molecular Medicine* 9, 836–850.
- Grabie, N., Gotsman, I., Dacosta, R., *et al.*, 2007. Endothelial programmed death-1 ligand 1 (PD-L1) regulates CD8 + T-cell mediated injury in the heart. *Circulation* 116, 2062–2071.
- Grakoui, A., Bromley, S.K., Sumen, C., *et al.*, 1999. The immunological synapse: A molecular machine controlling T cell activation. *Science* 285, 221–227.
- Greening, J.E., Tree, T.I., Kotowicz, K.T., *et al.*, 2003. Processing and presentation of the islet autoantigen GAD by vascular endothelial cells promotes transmigration of autoreactive T-cells. *Diabetes* 52, 717–725.
- Greenwood, J., Howes, R., Lightman, S., 1994. The blood-retinal barrier in experimental autoimmune uveoretinitis. Leukocyte interactions and functional damage. *Laboratory Investigation* 70, 39–52.

- Grigorova, I.L., Pantelev, M., Cyster, J.G., 2010. Lymph node cortical sinus organization and relationship to lymphocyte egress dynamics and antigen exposure. *Proceedings of the National Academy of Sciences of the United States of America* 107, 20447–20452.
- Grigorova, I.L., Schwab, S.R., Phan, T.G., *et al.*, 2009. Cortical sinus probing, S1P1-dependent entry and flow-based capture of egressing T cells. *Nature Immunology* 10, 58–65.
- Gubbels, J.A., Felder, M., Horibata, S., *et al.*, 2010. MUC16 provides immune protection by inhibiting synapse formation between NK and ovarian tumor cells. *Molecular Cancer* 9, 11.
- Guermonprez, P., Valladeau, J., Zitvogel, L., Thery, C., Amigorena, S., 2002. Antigen presentation and T cell stimulation by dendritic cells. *Annual Review of Immunology* 20, 621–667.
- Harwood, N.E., Batista, F.D., 2008. New insights into the early molecular events underlying B cell activation. *Immunity* 28, 609–619.
- Hirano, M., Das, S., Guo, P., Cooper, M.D., 2011. The evolution of adaptive immunity in vertebrates. *Advances in Immunology* 109, 125–157.
- Hirosue, S., Vokali, E., Raghavan, V.R., *et al.*, 2014. Steady-state antigen scavenging, cross-presentation, and CD8 + T cell priming: A new role for lymphatic endothelial cells. *Journal of Immunology* 192, 5002–5011.
- Hollenbaugh, D., Mischel-Petty, N., Edwards, C.P., *et al.*, 1995. Expression of functional CD40 by vascular endothelial cells. *Journal of Experimental Medicine* 182, 33–40.
- Huang, Y., Burkhardt, J.K., 2007. T-cell-receptor-dependent actin regulatory mechanisms. *Journal of Cell Science* 120, 723–730.
- Huttenlocher, A., Poznansky, M.C., 2008. Reverse leukocyte migration can be attractive or repulsive. *Trends in Cell Biology* 18, 298–306.
- Hwa, C., Sebastian, A., Aird, W.C., 2005. Endothelial biomedicine: Its status as an interdisciplinary field, its progress as a basic science, and its translational bench-to-bedside gap. *Endothelium* 12, 139–151.
- Janeway Jr., C.A., Bottomly, K., 1994. Signals and signs for lymphocyte responses. *Cell* 76, 275–285.
- Kasahara, M., Sutoh, Y., 2014. Two forms of adaptive immunity in vertebrates: Similarities and differences. *Advances in Immunology* 122, 59–90.
- Kawai, T., Seki, M., Watanabe, H., *et al.*, 2000. T(h)1 transmigration anergy: A new concept of endothelial cell-T cell regulatory interaction. *International Immunology* 12, 937–948.
- Khayamian, S., Hutloff, A., Buchner, K., *et al.*, 2002. ICOS-ligand, expressed on human endothelial cells, costimulates Th1 and Th2 cytokine secretion by memory CD4 + T cells. *Proceedings of the National Academy of Sciences of the United States of America* 99, 6198–6203.
- Klugewitz, K., Blumenthal-Barby, F., Schrage, A., *et al.*, 2002. Immunomodulatory effects of the liver: Deletion of activated CD4 + effector cells and suppression of IFN- γ -producing cells after intravenous protein immunization. *Journal of Immunology* 169, 2407–2413.
- Knolle, P.A., Schmitt, E., Jin, S., *et al.*, 1999. Induction of cytokine production in naive CD4(+) T cells by antigen-presenting murine liver sinusoidal endothelial cells but failure to induce differentiation toward Th1 cells. *Gastroenterology* 116, 1428–1440.
- Komatsu, M., Yee, L., Carraway, K.L., 1999. Overexpression of sialomucin complex, a rat homolog of MUC4, inhibits tumor killing by lymphokine-activated killer cells. *Cancer Research* 59, 2229–2236.
- Kotani, M., Seiki, K., Yamashita, A., Horii, I., 1966. Lymphatic drainage of thymocytes to the circulation in the guinea pig. *Blood* 27, 511–520.
- Lancrin, C., Sroczynska, P., Stephenson, C., *et al.*, 2009. The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. *Nature* 457, 892–895.
- Leung, D.Y., Meissner, C., Fulton, D., Schlievert, P.M., 1995. The potential role of bacterial superantigens in the pathogenesis of Kawasaki syndrome. *Journal of Clinical Immunology* 15, 11S–17S.
- Ley, K., Laudanna, C., Cybulsky, M.I., Nourshargh, S., 2007. Getting to the site of inflammation: The leukocyte adhesion cascade updated. *Nature Reviews Immunology* 7, 678–689.
- Limmer, A., Ohl, J., Kurts, C., *et al.*, 2000. Efficient presentation of exogenous antigen by liver endothelial cells to CD8 + T cells results in antigen-specific T-cell tolerance. *Nature Medicine* 12, 1348–1354.
- Linder, S., 2009. Invadosomes at a glance. *Journal of Cell Science* 12, 3009–3013.
- Lossinsky, A.S., Shivers, R.R., 2004. Structural pathways for macromolecular and cellular transport across the blood-brain barrier during inflammatory conditions. *Histology and Histopathology* 19, 535–564.
- Luo, B.H., Carman, C.V., Springer, T.A., 2007. Structural basis of integrin regulation and signaling. *Annual Review of Immunology* 25, 619–647.
- Luscinskas, F.W., Kansas, G.S., Ding, H., *et al.*, 1994. Monocyte rolling, arrest and spreading on IL-4-activated vascular endothelium under flow is mediated via sequential action of L-selectin, β_1 -integrins, and β_2 -integrins. *Journal of Cell Biology* 125, 1417–1427.
- Luxenburg, C., Geblinger, D., Klein, E., *et al.*, 2007. The architecture of the adhesive apparatus of cultured osteoclasts: From podosome formation to sealing zone assembly. *PLoS One* 2, e179.
- Ma, W., Pober, J.S., 1998. Human endothelial cells effectively costimulate cytokine production by, but not differentiation of, naive CD4 + T cells. *Journal of Immunology* 161, 2158–2167.
- Mai, J., Virtue, A., Shen, J., Wang, H., Yang, X.F., 2013. An evolving new paradigm: Endothelial cells—conditional innate immune cells. *Journal of Hematology & Oncology* 6, 61.
- Manes, T.D., Pober, J.S., 2008. Antigen presentation by human microvascular endothelial cells triggers ICAM-1-dependent transendothelial protrusion by, and fractalkine-dependent transendothelial migration of, effector memory CD4 + T cells. *Journal of Immunology* 180, 8386–8392.
- Manes, T.D., Pober, J.S., 2013. TCR-driven transendothelial migration of human effector memory CD4 T cells involves Vav, Rac, and myosin IIA. *Journal of Immunology* 190, 3079–3088.
- Manes, T.D., Shiao, S.L., Dengler, T.J., Pober, J.S., 2007. TCR signaling antagonizes rapid IP-10-mediated transendothelial migration of effector memory CD4 + T cells. *Journal of Immunology* 178, 3237–3243.
- Manjunath, N., Correa, M., Ardman, M., Ardman, B., 1995. Negative regulation of T-cell adhesion and activation by CD43. *Nature* 377, 535–538.
- Marelli-Berg, F.M., Frasca, L., Weng, L., Lombardi, G., Lechler, R.I., 1999. Antigen recognition influences transendothelial migration of CD4 + T cells. *Journal of Immunology* 162, 696–703.
- Marelli-Berg, F.M., Hargreaves, R.E., Carmichael, P., *et al.*, 1996. Major histocompatibility complex class II-expressing endothelial cells induce allospecific nonresponsiveness in naive T cells. *Journal of Experimental Medicine* 183, 1603–1612.
- Marelli-Berg, F.M., James, M.J., Dangerfield, J., *et al.*, 2004. Cognate recognition of the endothelium induces HY-specific CD8 + T-lymphocyte transendothelial migration (diapedesis) *in vivo*. *Blood* 103, 3111–3116.
- Marelli-Berg, F.M., Lechler, R.I., 1999. Antigen presentation by parenchymal cells: A route to peripheral tolerance? *Immunological Reviews* 172, 297–314.
- Marelli-Berg, F.M., Okkenhaug, K., Mirenda, V., 2007. A two-signal model for T cell trafficking. *Trends in Immunology* 28, 267–273.
- Martinelli, R., Kamei, M., Sage, P.T., *et al.*, 2013. Release of cellular tension signals self-restorative ventral lamellipodia to heal barrier micro-wounds. *Journal of Cell Biology* 201, 449–465.
- Martinelli, R., Zeiger, A.S., Whitfield, M., *et al.*, 2014. Probing the biomechanical contribution of the endothelium to lymphocyte migration: Diapedesis by the path of least resistance. *Journal of Cell Science* 127, 3020–3034.
- McFarland, T.A., Ardman, B., Manjunath, N., Fabry, J.A., Lieberman, J., 1995. CD43 diminishes susceptibility to T lymphocyte-mediated cytotoxicity. *Journal of Immunology* 154, 1097–1104.
- Mempel, T.R., Henrickson, S.E., Von Andrian, U.H., 2004. T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. *Nature* 427, 154–159.
- Middleton, J., Neil, S., Wintle, J., *et al.*, 1997. Transcytosis and surface presentation of IL-8 by venular endothelial cells. *Cell* 91, 1001–1011.
- Middleton, J., Patterson, A.M., Gardner, L., Schmutz, C., Ashton, B.A., 2002. Leukocyte extravasation: Chemokine transport and presentation by the endothelium. *Blood* 100, 3853–3860.
- Miranti, C.K., Brugge, J.S., 2002. Sensing the environment: A historical perspective on integrin signal transduction. *Nature Cell Biology* 4, E83–E90.
- Miyasaka, M., Tanaka, T., 2004. Lymphocyte trafficking across high endothelial venules: Dogmas and enigmas. *Nature Reviews Immunology* 4, 360–370.
- Monahan-Earley, R., Dvorak, A.M., Aird, W.C., 2013. Evolutionary origins of the blood vascular system and endothelium. *Journal of Thrombosis and Haemostasis* 11 (Suppl. 1), 46–66.
- Muller, W.A., 2003. Leukocyte-endothelial-cell interactions in leukocyte transmigration and the inflammatory response. *Trends in Immunology* 24, 327–334.
- Negulescu, P.A., Krasieva, T.B., Khan, A., Kerschbaum, H.H., Cahalan, M.D., 1996. Polarity of T cell shape, motility, and sensitivity to antigen. *Immunity* 4, 421–430.
- Nourshargh, S., Marelli-Berg, F.M., 2005. Transmigration through venular walls: A key regulator of leukocyte phenotype and function. *Trends in Immunology* 26, 157–165.
- Oddos, S., Dunsby, C., Purbhoo, M.A., *et al.*, 2008. High-speed high-resolution imaging of intercellular immune synapses using optical tweezers. *Biophysical Journal* 95, L66–L68.
- Olah, I., Glick, B., 1985. Lymphocyte migration through the lymphatic sinuses of the chicken's lymph node. *Poultry Science* 64, 159–168.

- Onami, T.M., Harrington, L.E., Williams, M.A., *et al.*, 2002. Dynamic regulation of T cell immunity by CD43. *Journal of Immunology* 168, 6022–6031.
- von Oppen, N., Schurich, A., Hegenbarth, S., *et al.*, 2009. Systemic antigen cross-presented by liver sinusoidal endothelial cells induces liver-specific CD8 T-cell retention and tolerization. *Hepatology* 49, 1664–1672.
- Paul, W.E., Seder, R.A., 1994. Lymphocyte responses and cytokines. *Cell* 76, 241–251.
- Pepper, M.S., Skobe, M., 2003. Lymphatic endothelium: Morphological, molecular and functional properties. *Journal of Cell Biology* 163, 209–213.
- Perez, V.L., Henault, L., Lichtman, A.H., 1998. Endothelial antigen presentation: Stimulation of previously activated but not naive TCR-transgenic mouse T cells. *Cell Immunology* 189, 31–40.
- Perez-Moreno, M., Jamora, C., Fuchs, E., 2003. Sticky business: Orchestrating cellular signals at adherens junctions. *Cell* 112, 535–548.
- Pfeiffer, F., Kumar, V., Butz, S., *et al.*, 2008. Distinct molecular composition of blood and lymphatic vascular endothelial cell junctions establishes specific functional barriers within the peripheral lymph node. *European Journal of Immunology* 38, 2142–2155.
- Phillipson, M., Heit, B., Colarusso, P., *et al.*, 2006. Intraluminal crawling of neutrophils to emigration sites: A molecularly distinct process from adhesion in the recruitment cascade. *Journal of Experimental Medicine* 203, 2569–2575.
- Pober, J.S., 1999. Immunobiology of human vascular endothelium. *Immunologic Research* 19, 225–232.
- Pober, J.S., Collins, T., Gimbrone Jr., M.A., *et al.*, 1983. Lymphocytes recognize human vascular endothelial and dermal fibroblast la antigens induced by recombinant immune interferon. *Nature* 305, 726–729.
- Pober, J.S., Sessa, W.C., 2007. Evolving functions of endothelial cells in inflammation. *Nature Reviews Immunology* 7, 803–815.
- Proft, T., Fraser, J.D., 2003. Bacterial superantigens. *Clinical & Experimental Immunology* 133, 299–306.
- Proudfoot, A.E., Handel, T.M., Johnson, Z., *et al.*, 2003. Glycosaminoglycan binding and oligomerization are essential for the *in vivo* activity of certain chemokines. *Proceedings of the National Academy of Sciences of the United States of America* 100, 1885–1890.
- Raine, C.S., Cannella, B., Duijvestijn, A.M., Cross, A.H., 1990. Homing to central nervous system vasculature by antigen-specific lymphocytes. II. Lymphocyte/endothelial cell adhesion during the initial stages of autoimmune demyelination. *Laboratory Investigation* 63, 476–489.
- Redpath, S., Alam, S.M., Lin, C.M., O'Rourke, A.M., Gascoigne, N.R., 1999. Cutting edge: Trimolecular interaction of TCR with MHC class II and bacterial superantigen shows a similar affinity to MHC:peptide ligands. *Journal of Immunology* 163, 6–10.
- Reitsma, S., Slaaf, D.W., Vink, H., Van Zandvoort, M.A., Oude Egbrink, M.G., 2007. The endothelial glycocalyx: Composition, functions, and visualization. *Pflügers Arch* 454, 345–359.
- Renno, T., Acha-Orbea, H., 1996. Superantigens in autoimmune diseases: Still more shades of gray. *Immunological Reviews* 154, 175–191.
- Risau, W., Flamme, I., 1995. Vasculogenesis. *Annual Review of Cell and Developmental Biology* 11, 73–91.
- Rodig, N., Ryan, T., Allen, J.A., *et al.*, 2003. Endothelial expression of PD-L1 and PD-L2 down-regulates CD8 + T cell activation and cytotoxicity. *European Journal of Immunology* 33, 3117–3126.
- Rose, M.L., 1998. Endothelial cells as antigen-presenting cells: Role in human transplant rejection. *Cellular and Molecular Life Sciences* 54, 965–978.
- Rudd, P.M., Elliott, T., Cresswell, P., Wilson, I.A., Dwek, R.A., 2001. Glycosylation and the immune system. *Science* 291, 2370–2376.
- Rudolph, M.G., Stanfield, R.L., Wilson, I.A., 2006. How TCRs bind MHCs, peptides, and coreceptors. *Annual Review of Immunology* 24, 419–466.
- Sabri, S., Soler, M., Foa, C., *et al.*, 2000. Glycocalyx modulation is a physiological means of regulating cell adhesion. *Journal of Cell Science* 113 (Pt. 9), 1589–1600.
- Sage, P.T., Carman, C.V., 2009. Settings and mechanisms for trans-cellular diapedesis. *Frontiers in Bioscience* 14, 5066–5083.
- Sage, P.T., Varghese, L.M., Martinielli, R., *et al.*, 2012. Antigen recognition is facilitated by invadosome-like protrusions formed by memory/effector T cells. *Journal of Immunology* 188, 3686–3699.
- Saito, T., Yokosuka, T., 2006. Immunological synapse and microclusters: The site for recognition and activation of T cells. *Current Opinion in Immunology* 18, 305–313.
- Saltel, F., Destaing, O., Bard, F., Eichert, D., Jurdic, P., 2004. Apatite-mediated actin dynamics in resorbing osteoclasts. *Molecular Biology of the Cell* 15, 5231–5241.
- Savage, C.O., Brooks, C.J., Harcourt, G.C., *et al.*, 1995. Human vascular endothelial cells process and present autoantigen to human T cell lines. *International Immunology* 7, 471–479.
- Savinov, A.Y., Wong, F.S., Stonebraker, A.C., Chervonsky, A.V., 2003. Presentation of antigen by endothelial cells and chemoattraction are required for homing of insulin-specific CD8 + T cells. *Journal of Experimental Medicine* 197, 643–656.
- Schenkel, A.R., Mamdouh, Z., Muller, W.A., 2004. Locomotion of monocytes on endothelium is a critical step during extravasation. *Nature Immunology* 5, 393–400.
- Schurich, A., Bottcher, J.P., Burgdorf, S., *et al.*, 2009. Distinct kinetics and dynamics of cross-presentation in liver sinusoidal endothelial cells compared to dendritic cells. *Hepatology* 50, 909–919.
- Schwartz, R.H., 1992. Costimulation of T lymphocytes: The role of CD28, CTLA-4, and B7/BB1 in interleukin-2 production and immunotherapy. *Cell* 71, 1065–1068.
- Seminario, M.C., Bunnell, S.C., 2008. Signal initiation in T-cell receptor microclusters. *Immunological Reviews* 221, 90–106.
- Shiao, S.L., Mcniff, J.M., Pober, J.S., 2005. Memory T cells and their costimulators in human allograft injury. *Journal of Immunology* 175, 4886–4896.
- Shulman, Z., Cohen, S.J., Roediger, B., *et al.*, 2012. Transendothelial migration of lymphocytes mediated by intraendothelial vesicle stores rather than by extracellular chemokine depots. *Nature Immunology* 13, 67–76.
- Shulman, Z., Shinder, V., Klein, E., *et al.*, 2009. Lymphocyte crawling and transendothelial migration require chemokine triggering of high-affinity LFA-1 integrin. *Immunity* 30, 384–396.
- Sobel, R.A., Blanchette, B.W., Colvin, R.B., 1984. Pre-inflammatory expression of fibronectin (Fn) and Ia in acute experimental allergic encephalomyelitis (EAE): Modulation of endothelial cells (EC) in the immune response detected by quantitative immunoperoxidase studies using monoclonal antibodies (MAB). *Progress in Clinical and Biological Research* 146, 81–86.
- Sorensen, K.K., McCourt, P., Berg, T., *et al.*, 2012. The scavenger endothelial cell: A new player in homeostasis and immunity. *American Journal of Physiology – Regulatory, Integrative, and Comparative Physiology* 303, R1217–R1230.
- Springer, T.A., 1990. Adhesion receptors of the immune system. *Nature* 346, 425–434.
- Springer, T.A., 1994. Traffic signals for lymphocyte recirculation and leukocyte emigration: The multistep paradigm. *Cell* 76, 301–314.
- Stockton, B.M., Cheng, G., Manjunath, N., Ardman, B., Von Andrian, U.H., 1998. Negative regulation of T cell homing by CD43. *Immunity* 8, 373–381.
- Taffin, C., Charron, D., Glotz, D., Mooney, N., 2011. Immunological function of the endothelial cell within the setting of organ transplantation. *Immunology Letters* 139, 1–6.
- Tavian, M., Cortes, F., Robin, C., *et al.*, 2000. The hemangioblast, common precursor of endothelial and hematopoietic cells. *Transfusion Clinique et Biologique* 7, 238–241.
- Tay, S.S., McCormack, A., Lawson, C., Rose, M.L., 2003. IFN-gamma reverses the stop signal allowing migration of antigen-specific T cells into inflammatory sites. *Journal of Immunology* 170, 3315–3322.
- Tewalt, E.F., Cohen, J.N., Rouhani, S.J., Engelhard, V.H., 2012a. Lymphatic endothelial cells – key players in regulation of tolerance and immunity. *Frontiers in Immunology* 3, 305.
- Tewalt, E.F., Cohen, J.N., Rouhani, S.J., *et al.*, 2012b. Lymphatic endothelial cells induce tolerance via PD-L1 and lack of costimulation leading to high-level PD-1 expression on CD8 T cells. *Blood* 120, 4772–4782.
- Thomson, A.W., Knolle, P.A., 2010. Antigen-presenting cell function in the tolerogenic liver environment. *Nature Reviews Immunology* 10, 753–766.
- Toro, I., Olah, I., 1967. Penetration of thymocytes into the blood circulation. *Journal of Ultrastructure Research* 17, 439–451.
- Tsuboi, S., Fukuda, M., 1998. Overexpression of branched O-linked oligosaccharides on T cell surface glycoproteins impairs humoral immune responses in transgenic mice. *Journal of Biological Chemistry* 273, 30680–30687.
- Tsuboi, S., Fukuda, M., 2001. Roles of O-linked oligosaccharides in immune responses. *Bioessays* 23, 46–53.
- Turesson, C., 2004. Endothelial expression of MHC class II molecules in autoimmune disease. *Current Pharmaceutical Design* 10, 129–143.
- Ushiki, T., 1986. A scanning electron-microscopic study of the rat thymus with special reference to cell types and migration of lymphocytes into the general circulation. *Cell and Tissue Research* 244, 285–298.
- Valitutti, S., Dessing, M., Aktories, K., Gallati, H., Lanzavecchia, A., 1995. Sustained signaling leading to T cell activation results from prolonged T cell receptor occupancy. Role of T cell actin cytoskeleton. *Journal of Experimental Medicine* 181, 577–584.
- Velazquez, P., Cameron, T.O., Kinjo, Y., *et al.*, 2008. Cutting edge: Activation by innate cytokines or microbial antigens can cause arrest of natural killer T cell patrolling of liver sinusoids. *Journal of Immunology* 180, 2024–2028.

- Vicente-Manzanares, M., Sanchez-Madrid, F., 2004. Role of the cytoskeleton during leukocyte responses. *Nature Reviews Immunology* 4, 110–122.
- Vora, M., Yssel, H., de Vries, J.E., Karasek, M.A., 1994. Antigen presentation by human dermal microvascular endothelial cells. Immunoregulatory effect of IFN- γ and IL-10. *Journal of Immunology* 152, 5734–5741.
- Wei, X., Tromberg, B.J., Cahalan, M.D., 1999. Mapping the sensitivity of T cells with an optical trap: Polarity and minimal number of receptors for Ca^{2+} signaling. *Proceedings of the National Academy of Sciences of the United States of America* 96, 8471–8476.
- Weinbaum, S., Tarbell, J.M., Damiano, E.R., 2007. The structure and function of the endothelial glycocalyx layer. *Annual Review of Biomedical Engineering* 9, 121–167.
- Weinreich, M.A., Hogquist, K.A., 2008. Thymic emigration: When and how T cells leave home. *Journal of Immunology* 181, 2265–2270.
- van de Wiel-Van Kemenade, E., Ligtenberg, M.J., De Boer, A.J., *et al.*, 1993. Episialin (MUC1) inhibits cytotoxic lymphocyte-target cell interaction. *Journal of Immunology* 151, 767–776.
- Wetzel, S.A., McKeithan, T.W., Parker, D.C., 2002. Live-cell dynamics and the role of costimulation in immunological synapse formation. *Journal of Immunology* 169, 6092–6101.
- Wheway, J., Obeid, S., Couraud, P.O., Combes, V., Grau, G.E., 2013. The brain microvascular endothelium supports T cell proliferation and has potential for alloantigen presentation. *PLoS One* 8, e52586.
- Wolburg, H., Wolburg-Buchholz, K., Engelhardt, B., 2005. Diapedesis of mononuclear cells across cerebral venules during experimental autoimmune encephalomyelitis leaves tight junctions intact. *Acta Neuropathologica* 109, 181–190.
- Wolosewick, J.J., 1984. Distribution of actin in migrating leukocytes *in vivo*. *Cell and Tissue Research* 236, 517–525.
- Yano, K., Liaw, P.C., Mullington, J.M., *et al.*, 2006. Vascular endothelial growth factor is an important determinant of sepsis morbidity and mortality. *Journal of Experimental Medicine* 203, 1447–1458.
- Yokosuka, T., Sakata-Sogawa, K., Kobayashi, W., *et al.*, 2005. Newly generated T cell receptor microclusters initiate and sustain T cell activation by recruitment of Zap70 and SLP-76. *Nature Immunology* 6, 1253–1262.

T Cell Receptor Triggering

B Alarcón and HM van Santen, Centro Biología Molecular Severo Ochoa, CSIC/UAM, Madrid, Spain

© 2016 Elsevier Inc. All rights reserved.

Introduction

Activation of T cells of the $\alpha\beta$ lineage depends on specific binding of the T cell antigen receptor (TCR) to its ligand, a cognate peptide bound to a major histocompatibility complex molecule (pMHC), at the cell surface of antigen presenting cells (APCs). Their TCR is composed of a clonotypic, antigen specific TCR $\alpha\beta$ heterodimer, and the invariant and signal transducing CD3 $\gamma\epsilon$, CD3 $\delta\epsilon$, and CD3 $\zeta\zeta$ dimers (Kuhns and Badgandi, 2012; Figure 1). The recognition and signaling components of the TCR are therefore located in different protein subunits and signal transduction via the TCR is therefore strictly dependent on communication between these components. Binding of TCR $\alpha\beta$ to its pMHC ligand leads to the phosphorylation of the intracellular immunoreceptor tyrosine-based activation motifs (ITAMs) of the CD3 $\gamma\epsilon$, CD3 $\delta\epsilon$, and CD3 $\zeta\zeta$ dimers by the CD4 or CD8 co-receptor-associated src-family kinase Lck, allowing recruitment of the ZAP70 protein tyrosine kinase (PTK) via its tandem SH2 domains, which in turn phosphorylates a number of downstream signaling components (Smith-Garvin *et al.*, 2009). However, not all TCR signaling is mediated by ITAMs: other sequences within the intracellular domains, such as the basic-rich stretches (BRS) of CD3 ϵ and CD3 ζ and the proline-rich sequence (PRS) of CD3 ϵ have more recently been shown to be critical for full functionality of the TCR (Boroto *et al.*, 2013, 2014; Brodeur *et al.*, 2009; Deford-Watts *et al.*, 2009; Mingueneau *et al.*, 2008; Bettini *et al.*, 2014).

A difficulty to understand how the TCR transmits outside-in signals emerges from the unavailability of a high-resolution structure of the entire complex. In lack of this essential information, partial structures of isolated dimers (ectodomains of TCR $\alpha\beta$, ectodomains of CD3 $\gamma\epsilon$, ectodomains of CD3 $\delta\epsilon$, transmembrane domain of CD3 $\zeta\zeta$, and cytoplasmic tail of CD3 ϵ) have been characterized. Most crystal structures of the TCR $\alpha\beta$ extracellular domain in free form or bound to its pMHC ligand show little or no differences between these structures (Rudolph *et al.*, 2006). However, the crystal structure of the $\alpha\beta$ LC13 TCR showed a shift in the AB loop of the α domain induced by pMHC ligand binding (Kjer-Nielsen *et al.*, 2003). This shift was later confirmed by fluorescence spectroscopy methods (Beddoe *et al.*, 2009). To what extent is this AB loop shift a peculiarity of this TCR, or a generalized phenomenon is not uncovered due to poor resolution of the α domain, and to what level such shift intervenes in the transmission of TCR-to-CD3 conformational changes remains to be clarified. Outside-in transmission of the occurrence of a TCR $\alpha\beta$ -pMHC interaction may therefore rely on conformational changes that reorientate TCR $\alpha\beta$ in relation to the CD3 components, on TCR clustering provoked by pMHC ligand-induced crosslinking or on the recruitment or exclusion of other molecules. All of these mechanisms have to account for the unusual combination of properties of TCR-mediated activation of T cells: a very low affinity of the TCR-pMHC

interaction (Alam *et al.*, 1999), very high sensitivity and specificity for antigen (Huang *et al.*, 2013; Irvine *et al.*, 2002; Reay *et al.*, 2000; Sykulev *et al.*, 1996), and ability to induce different signaling outputs and cell fates in response to different pMHC ligands both during T cell differentiation in the thymus and activation in the periphery (Daniels *et al.*, 2006; Sloan-Lancaster *et al.*, 1994). In this article we will focus on experimental evidence that explain the molecular mechanisms

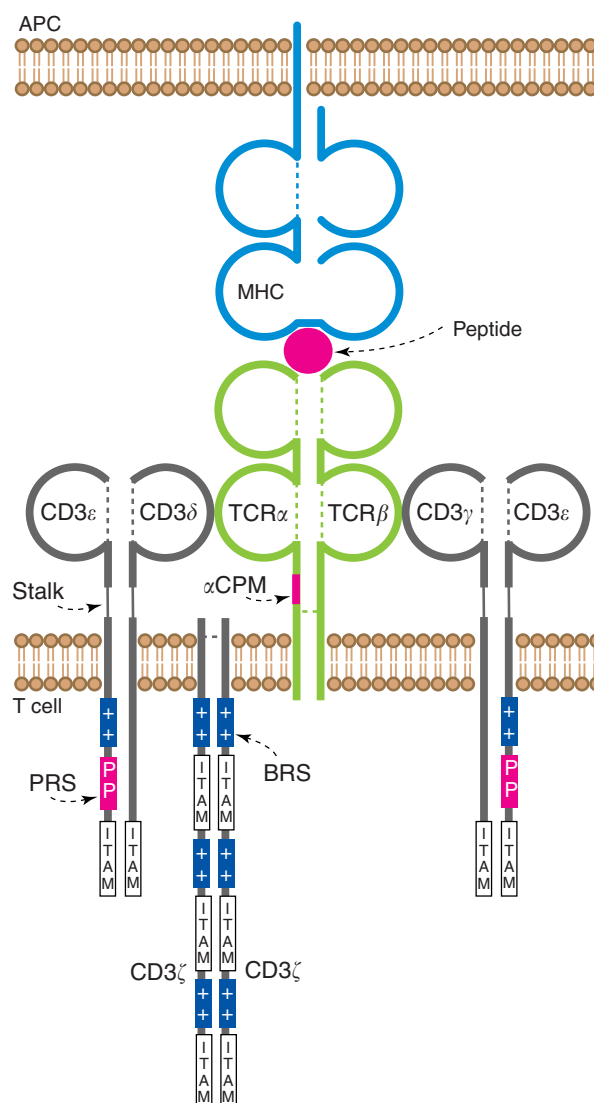


Figure 1 Schematic representation of the TCR complex. Domains discussed in this article are indicated. Semicircles represent the Ig superfamily immunoglobulin-like extracellular domains of the TCR and CD3 chains. Intra- and interchain disulfide bonds are represented by thin dashed lines. Colors of the chains and domains are maintained in all figures, although for clarity sake simplified representations omitting certain domains and chains are used.

that transmit the occurrence and quality of the extracellular pMHC-TCR $\alpha\beta$ interaction to the intracellular domains of the CD3 and CD3 $\zeta\zeta$ dimers.

Time of Interaction and Recruitment of Modulators

One of the basic tenets of T cell activation has been that the quality of the TCR-pMHC ligand interaction determines the quality of the T cell response. Measurements of kinetic parameters for the TCR-pMHC interaction in solution by surface plasmon resonance showed that the degree of activation of T cells by related pMHC ligands correlated to a reasonable extent with the off-rate of the interaction between their TCRs with these pMHC ligands (Alam *et al.*, 1999; Rosette *et al.*, 2001). These findings lead to the formulation of kinetic proofreading (McKeithan, 1995) and kinetic discrimination (Rabinowitz *et al.*, 1996) models that postulate that the time of the TCR-pMHC interaction (dwell time, dependent on the off-rate of the interaction) determines the time-dependent extent of TCR activation (e.g., phosphorylation of ITAMs, recruitment of downstream effectors such as ZAP70), which in turn determines the degree of response. These models did, however, not provide a mechanistic explanation of how the occurrence and duration of TCR-pMHC binding was transmitted to the downstream effectors.

A mechanistic insight has been provided by Palmer and colleagues who studied a correlation between kinetic parameters of TCR-pMHC interaction and positive and negative selections during T cell maturation in the thymus (Palmer and Naeher, 2009). The fate of thymocytes is determined by the interaction of their TCR with self-pMHC complexes expressed by thymic stroma: thymocytes expressing TCRs with low affinity for self-pMHC receive signals that activate a survival and differentiation program (positive selection) leading to maturation into MHC-restricted, functional T cells, whereas thymocytes expressing TCRs with high affinity for self-pMHC activate signaling pathways that lead to apoptosis (negative selection) and thus prevent the presence of strongly self-reactive T cells in the repertoire (Starr *et al.*, 2003). Palmer and colleagues found for three different TCRs that a very discrete threshold separates the longer dwell times of pMHC ligands that cause negative selection from the shorter dwell times of positively selecting pMHC ligands (Naeher *et al.*, 2007). This threshold dwell time (around 4 s) seems to be a constant, suggesting a general time-determined underlying molecular mechanism. This mechanism has been proposed to rely on CD4 or CD8 co-receptor recruitment to the TCR which would be more efficiently recruited upon ligation by high- than by low-affinity pMHC (Mallaun *et al.*, 2008). This differential recruitment of the co-receptor would result in differential recruitment of co-receptor-associated Lck (Rudd *et al.*, 1988; Veillette *et al.*, 1988) to the intracellular domains of the TCR and therefore more extensive phosphorylation of the ITAMs. The extent of ITAM phosphorylation would then in turn enable more or less efficient recruitment of downstream effectors and therefore intensity of signaling. Interestingly, co-receptor recruitment to the TCR is not only mediated by co-receptor binding to the constant domains of pMHC (Doyle and Strominger, 1987; Norment *et al.*, 1988) but also depends on

the extracellular, membrane-proximal connecting peptide motif (α CPM) domain of the TCR α chain, since T hybridomas lacking the α CPM show reduced pMHC-dependent co-receptor recruitment and reduced ITAM phosphorylation (Mallaun *et al.*, 2008), and thymocytes without this domain cannot respond to positively selecting pMHC ligands, as evidenced by impaired ERK phosphorylation, and DP arrest (Werlen *et al.*, 2000). These findings have led to the proposal that close association of the co-receptor (and therefore Lck) with the TCR is a step-wise, time-dependent 'zipper' process (Figure 2) in which the available time, depending on the duration (i.e., affinity) of the TCR-pMHC interaction, determines the degree of co-receptor-mediated Lck recruitment. This would in turn determine the extent of ITAM phosphorylation, which has been shown to regulate the biological response (Holst *et al.*, 2008; Kersh *et al.*, 1998).

The correlation between dwell time and dissociation constant (k_{off}) on one hand and biological activity on the other is good when kinetic data is calculated in tridimensional (3D) settings, such as via surface plasmon resonance in which at least one of the components (TCR or pMHC) is in solution. However, this correlation is lost when two different approaches are used that measure the affinity and kinetics of TCR-pMHC interaction with both components bound to biological membranes and restricted to diffusion in two dimensions (2D). These two systems, that have the capacity of single receptor-ligand pair resolution, have shown that the 2D affinity and 2D on-rate of this interaction corresponds very well to the stimulatory capacity of these pMHC ligands (Huang *et al.*, 2010; Huppa *et al.*, 2010). Surprisingly, the average dwell time for strongly activating ligands was shorter than for weakly activating ligands (due to very fast off-rates) (Huang *et al.*, 2010). This would contradict the above-described time of interaction models and might hint at models where rapid pMHC sampling by the TCR and the degree of accumulation of a transient secondary messenger were the underlying determinants of T cell activation.

Exclusion of Negative Regulators

In a classic review on adhesion receptors of the immune system, Timothy Springer pointed out that the extracellular domains of the various receptor-ligand pairs that form in the contact area between the T cell and APC showed important size differences and that this might lead to regions within the contact area with different intermembrane spaces (Springer, 1990). He suggested that as a consequence proteins with large extracellular domains, such as the protein tyrosine phosphatase CD45, might be separated from or reoriented with respect to smaller receptor-ligand pairs such as TCR-pMHC and that this might influence signaling. The relevance of protein tyrosine phosphatase activity in regulating TCR signaling was underscored by experiments that showed that treatment of T cells with protein tyrosine phosphatase inhibitors such as pervanadate gave rise to a series of proximal and distal T cell activation events that closely resemble those induced by TCR-dependent stimuli (Secrist *et al.*, 1993). Thus the resting state of the TCR was actively maintained by tyrosine phosphatases and part of TCR-mediated signal transduction might depend

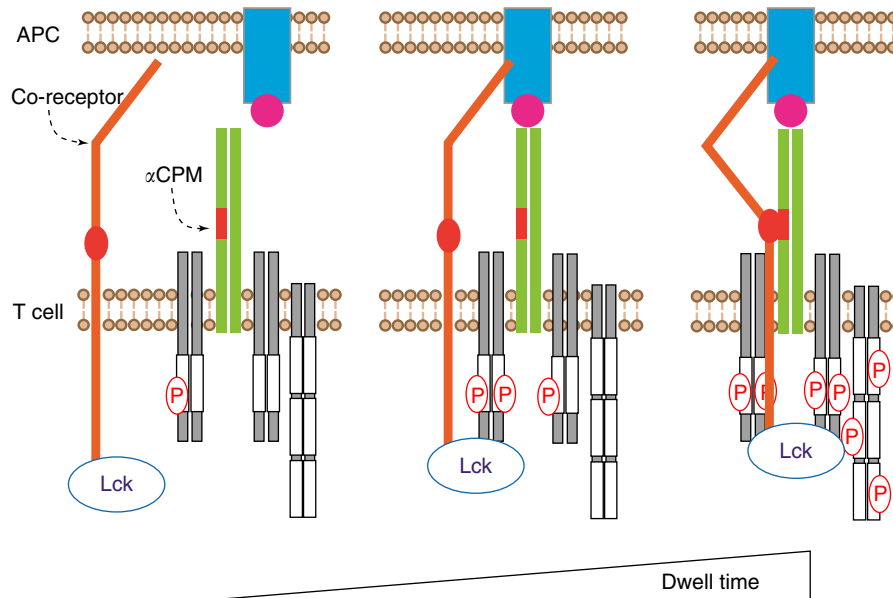


Figure 2 Zipper model of TCR triggering. Upon engagement with a cognate pMHC (peptide represented by magenta-colored circle) the co-receptor (represented in orange; CD4 for MHC class II restricted T cells, CD8 for MHC class I-restricted T cells) will bind to the MHC molecule (represented in blue). This will bring Lck, associated with the intracellular domain of the co-receptor, closer to the ITAM domains of the CD3 chains causing partial phosphorylation (red circled P). If the dwell time of the TCR–cognate pMHC interaction is sufficiently long, the co-receptor will also be able to bind the α CPM domain of TCR α (indicated in red) allowing for a closer proximity of Lck to the ITAM domains of the CD3 chains and hence more extensive ITAM phosphorylation.

on inhibition of the TCR-proximal phosphatase activities. This control of tyrosine kinases could be exerted by membrane-bound and/or cytosolic phosphatases, but the focus was put on membrane-bound phosphatases such as CD45, one of the most abundant membrane-associated phosphatases on T cells with both inhibitory and stimulatory effects on T cell activation (Rhee and Veillette, 2012), as it was shown to be absent from actively signaling TCRs of T cells stimulated via supported lipid bilayers (SLB) containing cognate pMHC complexes and ICAM-1 (Varma *et al.*, 2006) or, less physiologically, stimulated on coverslips coated with immobilized anti-CD3 antibodies (Bunnell *et al.*, 2002). These observations have led to the definition of the kinetic segregation model (Davis and van der Merwe, 2006), which proposes that, due to their large extracellular domain, CD45 and the CD148 phosphatase are excluded from the relatively narrow contact zone established between the T cell and the APC by the interaction between the TCR and pMHC and other receptor–ligand pairs that all have much shorter extracellular domains (Figure 3). Evidence in favor of this model is the observation that transfection of cell lines with recombinant CD45 molecules with shorter extracellular domains inhibits early (ZAP70 phosphorylation) and late (IL-2 production) T cell activation steps (Cordoba *et al.*, 2013). Vice versa, APCs expressing pMHC with enlarged extracellular domains are less stimulatory than APCs expressing MHC of ‘normal’ length (Choudhuri *et al.*, 2005, 2009). These observations suggest that the size of the extracellular domain of the CD45 and CD148 phosphatases relative to the dimensions of the extracellular pMHC and TCR domains determines their degree of exclusion and therefore, the degree of tyrosine phosphorylation. However, it has been

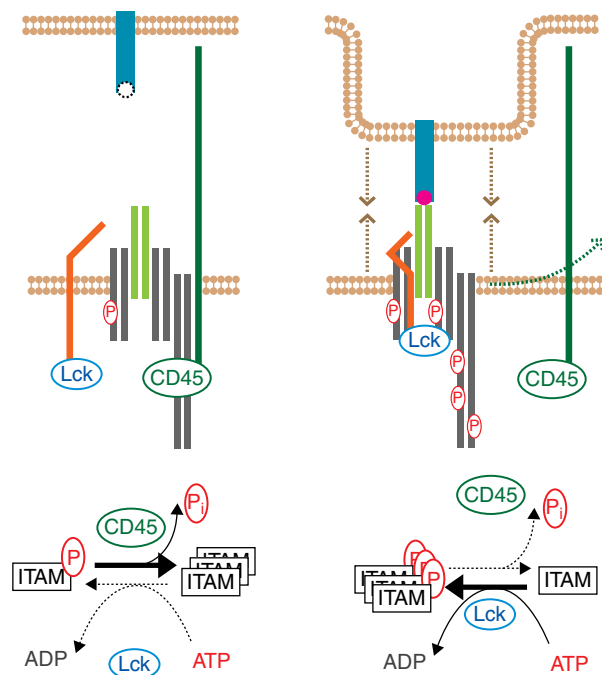


Figure 3 Kinetic segregation model. Recognition of the cognate pMHC on the APC generates a TCR-dependent force (represented by brown dashed arrows) that brings the opposing membranes of the T cell and APC in closer contact. This forces the segregation of the CD45 phosphatase (dark green) from the contact area, due to its long and rigid extracellular domain (top). Concurrent recruitment of Lck results in a shift in the balance of kinase and phosphatase activity and leads to more extensive ITAM phosphorylation (bottom).

shown in a reconstituted T cell/B cell interaction model that a CD45 recombinant with a short CD86-derived extracellular domain as well as unbound CD58 are significantly excluded from the contact area (James and Vale, 2012), indicating that size alone may not be the only determinant for exclusion from the contact area. Regardless of the relevance of ectodomain size-based phosphatase exclusion on TCR signaling, it is still pending to clarify if phosphatase exclusion is a mechanism that 'initiates' TCR signaling or just a mechanism for sustainment and amplification of the initial signal.

Changes in Conformation and Accessibility of Intracellular Domains

Engagement of the TCR ectodomains with antigen-loaded APCs, soluble pMHCs and antibodies against the extracellular domains of TCR $\alpha\beta$ or CD3 $\gamma\epsilon$ and CD3 $\delta\epsilon$ leads to exposure of the intracellular, membrane-proximal PRS of the CD3 ϵ chain, permitting the recruitment of the SH2- and SH3-domain containing adaptor protein Nck to the PRS (Gil *et al.*, 2002; Minguet *et al.*, 2007; Risueno *et al.*, 2005; Figure 4(a)). Ligand-induced binding of CD3 ϵ to Nck is independent of tyrosine phosphorylation, is induced at 0 °C, and can be triggered on cell-free purified TCR complex. This represents all hallmarks of a conformational change in the TCR complex. Transmission of the extracellular engagement to the CD3 ϵ PRS critically depends on the extracellular, membrane-proximal stalk region of CD3 ϵ . Molecular dynamics analysis of the extracellular domain of CD3 $\delta\epsilon$ dimer in presence and absence of a conformational change-inducing antibody showed that antibody binding made the stalk region more rigid (Martinez-Martin *et al.*, 2009). The molecular dynamics analysis also identified amino acid residues in the stalk, in close proximity to or part of an evolutionary conserved CXXC motif (Wang *et al.*, 2009), that appeared critical for the formation of the rigid stalk. Single mutations of these residues in full length CD3 ϵ chains impaired the antibody-mediated induction of the conformational change when these mutated CD3 ϵ chains were expressed in a murine T cell hybridoma (Martinez-Martin *et al.*, 2009). This coincided with reduced phosphorylation of the CD3 ζ chain and ERK and a diminished effector response upon activation with APCs loaded with graded amounts of the antigenic peptide. This means that transmission of the occurrence of the physiologically relevant interaction between cognate pMHC and TCR $\alpha\beta$ to the downstream signaling pathways is quantitatively dependent on the stalk region of the CD3 dimers. The mutated CD3 ϵ chains exerted a strong dominant negative effect as they were co-expressed with the wild type CD3 ϵ chains of the murine T hybridoma. This can only be explained by assuming the need of extensive collaboration between individual TCR complexes during TCR triggering. In an alternative study it was shown that mice, expressing CD3 ϵ chains with both of the cysteines of the CXXC motif mutated to serines, had defects in TCR signaling dependent thymocyte differentiation steps and reduced mature T cell responses (Wang *et al.*, 2009), underscoring the importance of the stalk domain in proper TCR triggering. It was not reported whether this mutated CD3 ϵ chain was able to undergo the conformational change. More recently, it has been shown that the CXXC

motif in the stalk of CD3 ϵ is oxidized, forming a small intra-chain disulfide loop (Brazin *et al.*, 2014). A CXXC motif is also present in the stalk regions of CD3 γ and CD3 δ . Interestingly, Cys-to-Ser mutation of the CXXC motif in CD3 δ leads to changes in NMR spectra of amino acids of the transmembrane domain and even cytoplasmic amino acids as far down as the ITAM motif (Brazin *et al.*, 2014). These results suggest that the four CXXC motifs present in the CD3 $\gamma\epsilon$ and CD3 $\delta\epsilon$ dimers intervene in the outside-in transmission of conformational changes.

Efficient induction of the conformational change and downstream signaling events by soluble ligands depends on multivalent ligands and close spacing of the engaged TCRs (Minguet *et al.*, 2007). Monovalent antibody (Fab) fragments specific for CD3 can induce the conformational change to a certain extent, but are unable to induce downstream signaling events. Upon combination of these Fab fragments with artificial ligands that bring TCRs in close proximity but by themselves cannot induce the conformational change or downstream signaling, an increased level of conformational change and downstream signaling is detected. These findings, together with the fact that all natural TCR $\alpha\beta$ heterodimers bind pMHC with an approximately 30° angle (Rudolph *et al.*, 2006), have led to a TCR triggering model (the permissive geometry model; Figure 4(b)) in which binding of multivalent pMHC to a clustered TCR provides both anchoring points and force that allow the engaged TCR $\alpha\beta$'s to reposition relative to each other, which in turn generates force on the associated CD3 dimers and induces the conformational change (Minguet *et al.*, 2007). Finally, maintenance of the conformational change is dependent on continued ligand binding (Minguet *et al.*, 2007), providing a potential role for ligand affinity in regulating the conformational change-dependent signaling pathways. It was found in two mouse models of thymic differentiation that only high affinity, negative selection inducing interactions gave rise to a conformational change, providing a potential mechanism via which TCRs in the thymus can activate different signaling pathways during positive and negative selection (Risueno *et al.*, 2006). Studies on another model of thymic selection indicated that this discriminating capacity is developmentally regulated, as in preselection thymocytes both low- and high-affinity pMHC ligands induced the conformational change, whereas in positively selected thymocytes and mature T cells only high-affinity ligands induced the conformational change (Gil *et al.*, 2005, 2008).

Regulation of the accessibility of the ITAM domains of the CD3 chains for Lck and ZAP70 may represent another mechanism of TCR signaling that could be part or not of the outside-in conformational change described above. Interestingly, it has been described that the activity of Lck does not change upon TCR triggering, thus leaving regulation of ITAM phosphorylation to a matter of Lck localization or, alternatively, to a regulation of exposure of the ITAM sequences to the kinase (Nika *et al.*, 2010). *In vitro* studies showed that recombinant proteins representing the intracellular domains of CD3 ϵ and CD3 ζ specifically associate with unilamellar vesicles composed of acidic phospholipids (Aivazian and Stern, 2000; Sigalov *et al.*, 2006; Figure 4(c)). These studies also showed that ITAM-phosphorylated versions of these proteins lost their lipid binding capacity and vice versa: once associated to the lipid

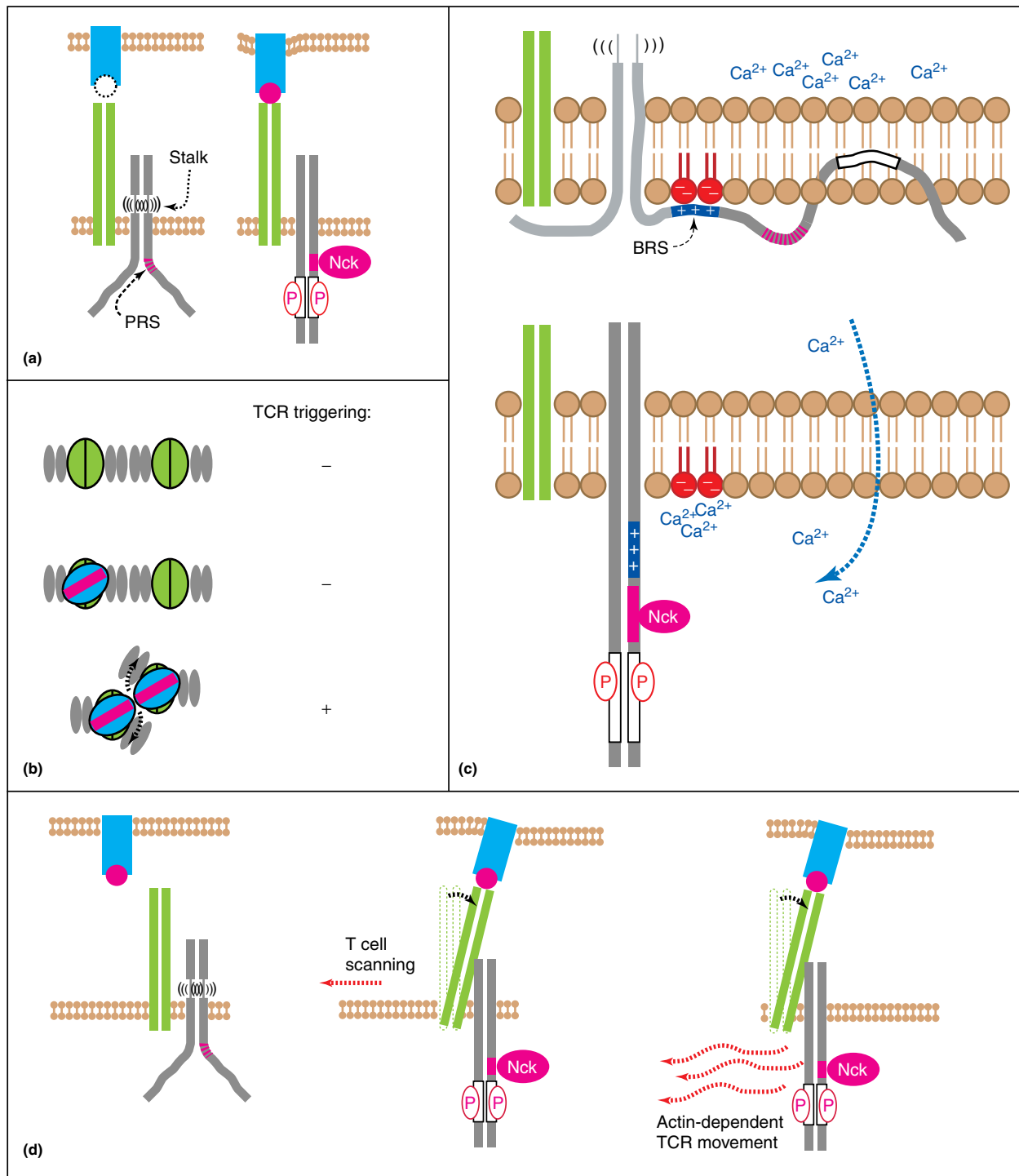


Figure 4 Conformation and accessibility changes and force as mediators of TCR triggering. (a) Interaction of the TCR with cognate pMHC leads to stiffening of the extracellular stalk domain and a conformational change in the intracellular tail of CD3 ϵ leading to exposure of the PRS. This allows for recruitment of the adaptor protein Nck. (b) The permissive geometry model proposes that binding of minimally bivalent pMHC is necessary to force a rearrangement of the TCR and CD3 chains that will result in the conformational change (top view of TCR-CD3 and pMHC). (c) The intracellular tails of CD3 ϵ chain and ζ are embedded in the inner leaflet of the plasma membrane of resting T cells in a phospholipid and BRS-dependent fashion, impeding access of Lck to the ITAMs (top). TCR triggering leads to release from the membrane, involving a TCR triggering-dependent influx of Ca^{2+} ions and competition between these Ca^{2+} ions and the BRS for the negatively charged phospholipids (bottom). (d) Interaction of the TCR with cognate pMHC provides an anchoring point to the TCR which allows it to exert force on the other components of the complex, either due to T cell movement during scanning of the APC or due to actin-dependent transport of TCR microclusters toward the immunological synapse.

vesicles, they became refractory to src kinase-mediated ITAM phosphorylation. NMR analysis indicated that the non-phosphorylated tyrosine residues of these recombinant proteins were deeply buried in the lipid bilayer, suggesting a mechanism to regulate their accessibility to PTKs and hence T cell activation (Xu *et al.*, 2008). Additional studies showed that binding of the CD3 ϵ and ζ intracellular domains to the unilamellar vesicles depended on clustered basic residues, the BRS, in these intracellular domains (DeFord-Watts *et al.*, 2009, 2011; Xu *et al.*, 2008). Recombinant proteins containing the CD3 ϵ or ζ intracellular domains with a C-terminally associated fluorescent probe were generated to study by Förster resonance energy transfer (FRET) the relevance and mechanisms of the membrane association and dissociation in T cell lines (Gagnon *et al.*, 2012; Zhang *et al.*, 2011). The recombinant CD3 ζ protein associated with the endogenous TCR components and its intracellular domain was associated with the plasma membrane in a BRS-dependent manner under resting conditions (Zhang *et al.*, 2011). TCR triggering by anti-CD3-coated beads caused dissociation from the membrane in the contact area between the bead and the T cell. This dissociation was dependent on phosphorylation of the ITAM sequences, as constructs with mutations of the Tyr residues to Phe remained associated to the membrane upon stimulation. In contrast, the study of the CD3 ϵ recombinant, for which it was not shown that it incorporated into the TCR complex and which had both Tyr residues of the ITAM replaced by Phe, showed that this chain was released from the plasma membrane upon pMHC-mediated TCR triggering independent from phosphorylation (Gagnon *et al.*, 2012). It is not clear whether this apparent contradiction reflects possible chain-specific differences under physiological conditions or is due to the properties of the recombinant proteins used in these studies. Finally, intracellular Ca²⁺ ions were shown to release the intracellular domains of recombinant CD3 ϵ from the plasma membrane by competing with the BRS for binding to the phosphate groups of the anionic phospholipids in the inner leaflet of the plasma membrane (Shi *et al.*, 2013), suggesting a positive feedback role of the TCR triggering-dependent Ca²⁺ influx in T cell activation.

TCRs as Mechanoreceptors

Application of force on TCRs changes the T cell's response to antigen. A non-stimulatory, bead-bound CD3-specific antibody could be rendered stimulatory, as measured by induction of Ca²⁺ influx, by applying force on the bead via optical tweezers (Kim *et al.*, 2009). In this setting, force was effective if the bead was moved parallel to the cell surface of the T cell, not when moved perpendicular to the cell surface, leading to the suggestion that the TCR was a directionally dependent mechanosensor. This tangentially applied force on the CD3 dimer was speculated to mimic the sideward push applied by the adjacent TCR $\alpha\beta$ when, due to its binding to its pMHC ligand on the APC, it was either bent because of the pull generated by the scanning movement of the T cell over the APC (Friedl and Brouck, 2002), or because of the actin-dependent movement of TCR microclusters (Yokosuka and Saito, 2010; Figure 4(d)). Another study showed that a weakly

stimulating artificial TCR ligand, an anti-CD3 single chain fragment coupled to the extracellular domain of CD43 expressed at the cell surface of a fibroblast, became a much more potent stimulator in combination with a tangential shear force applied to the ligand-bound T cell, or when the ligand-bound T cell was gently pulled away from the fibroblast (Li *et al.*, 2010). This would suggest that the direction of the force was less restricted than proposed by Kim and coworkers. However, these were in both cases cell-based experiments with numerous TCR–ligand interactions involved at the same time. Given that the T cell–APC interface is not necessarily a perfect plane (Roybal *et al.*, 2013), it is not straightforward to extrapolate the direction of the force applied to the whole cell to the direction of the force on individual TCR–ligand pairs.

2D affinity measurements with a biological force probe (BFP) under zero-force conditions had, as previously mentioned, shown that the biological potency of a range of pMHC ligand correlated directly with the affinity of the TCR–pMHC interaction but correlated inversely with the dwell time (Huang *et al.*, 2010). Repetition of these interaction studies under increasing pulling force conditions (applied via the BFP) showed that weak ligands formed 'slip bonds': dwell time decreased when the amount of force applied to the interaction was increased. In contrast, strong ligands formed 'catch bonds' with the TCR: dwell time reached an optimum when increased force was applied (Liu *et al.*, 2014). Furthermore, this maximal dwell time correlated with a maximal T cell response (measured as a Ca²⁺ influx). Interestingly, applying pulling force conditions resulted in a perfect correlation between dwell time and biological potency of a panel of pMHC ligands. Applying the same range of force via a TCR $\alpha\beta$ -specific antibody (which formed slip bonds but, due to its high affinity, resulted under all force conditions in dwell times longer than those obtained with the strongest pMHC) showed that maximal induction of response (Ca²⁺ flux) was obtained under the same force conditions observed for the strong pMHC. This coincidence suggests that both pMHC and the anti-TCR $\alpha\beta$ antibody act on the same force-requiring process in TCR triggering. The molecular basis of the pMHC–TCR $\alpha\beta$ catch bond formation (a functional description) is not known, but might be revealing with regard to the understanding of the mechanism of TCR triggering. Bearing all this data in mind, one has to consider that pulling or tangential forces could just be a mechanism to produce piston-like or rotational movements in the TCR and thereby provoke conformational changes.

In the above-mentioned experiments the role of force in TCR triggering was studied by applying an artificial force to the TCR–ligand pair. In support of the physiological relevance of force in TCR triggering, it has been shown that T cells can also generate force themselves upon TCR engagement. Using an alternative BFP setup that permits measurement of the force generated by the T cell, Husson *et al.* (2011) showed that T cells first apply an actin-dependent pushing force on an anti-CD3-coated bead, followed by a pulling force that eventually leads to engulfment of the bead. This pushing and pulling action is TCR-specific, as engagement of LFA-1 via an anti-LFA-1-coated bead did not lead to pushing and to marginal pulling. Pushing tended to start before detection of a Ca²⁺ influx, suggesting that it was a cause of TCR triggering although it does not exclude that it was a consequence.

Clustering

Whereas it has long been recognized that TCRs form microclusters upon antigen engagement (Krummel *et al.*, 2000; Campi *et al.*, 2005; Yokosuka *et al.*, 2005) it has only more recently been established by a variety of techniques that in the absence of ligands TCRs at the cell surface are present as a combination of preexisting oligomers with a variable size of up to 25 TCR complexes per cluster (Kumar *et al.*, 2011; Lillemeier *et al.*, 2010; Schamel *et al.*, 2005; Sherman *et al.*, 2011). These preformed TCR nanoclusters, whose existence is dependent on the presence of cholesterol in the plasma membrane, have been shown to enhance the sensitivity of the T cell for antigen (Kumar *et al.*, 2011; Schamel *et al.*, 2005). It has been proposed that this occurs via mechanisms implying cooperative ligand binding or signal transduction spreading (Figure 5; Alarcon *et al.*, 2006; Castro *et al.*, 2014; Martinez-Martin *et al.*, 2009). Direct experimental evidence for the molecular mechanisms that underlie this enhancement of sensitivity has not yet been reported. However, the 2D affinities of TCR–pMHC interactions are reduced when the interacting T cells are pretreated with cholesterol-depleting agents (Huang *et al.*, 2010; Huppa *et al.*, 2010), suggesting that cooperative binding of a cluster of pMHC ligands to a TCR nanocluster could occur.

Quantity and Valency of Cognate pMHC Ligands

Early studies on the number of pMHC complexes needed on an APC to stimulate a T cell reached the conclusion that a few to 100 cognate pMHC were sufficient to activate T cells (Demotz *et al.*, 1990; Harding and Unanue, 1990; Reay *et al.*, 2000; Sykulev *et al.*, 1996). These numbers were however based on an estimation of the average number of peptides on the APCs and using populations of T cells and APCs, leaving open the possibility that the actually activating APCs presented

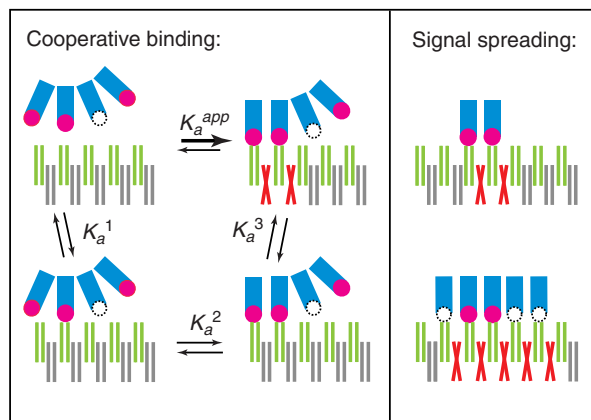


Figure 5 Enhancement of signaling via TCR nanoclusters. Binding of a first cognate pMHC facilitates binding of a second cognate pMHC within a pMHC cluster, resulting in an apparent affinity that is higher than the sum of individual binding affinities (left panel). Binding of a pMHC cluster containing both cognate and self pMHCs to a TCR nanocluster permits spreading of TCR triggering to self-pMHC engaged TCRs (right panel).

more peptides than the average number. It was also known that soluble TCR triggering molecules, such as CD3- and TCR $\alpha\beta$ -specific antibodies and soluble pMHC complexes, need to be at least bivalent in order to be able to stimulate T cells (Boniface *et al.*, 1998; Kaye and Janeway, 1984). Extrapolation of these findings to the cognate pMHCs on APCs led to an apparent problem: how likely was it that two adjacent pMHC complexes would present the same, cognate peptide when only few cognate pMHC complexes were present amongst an overwhelming majority of self-pMHC complexes (Figure 6)?

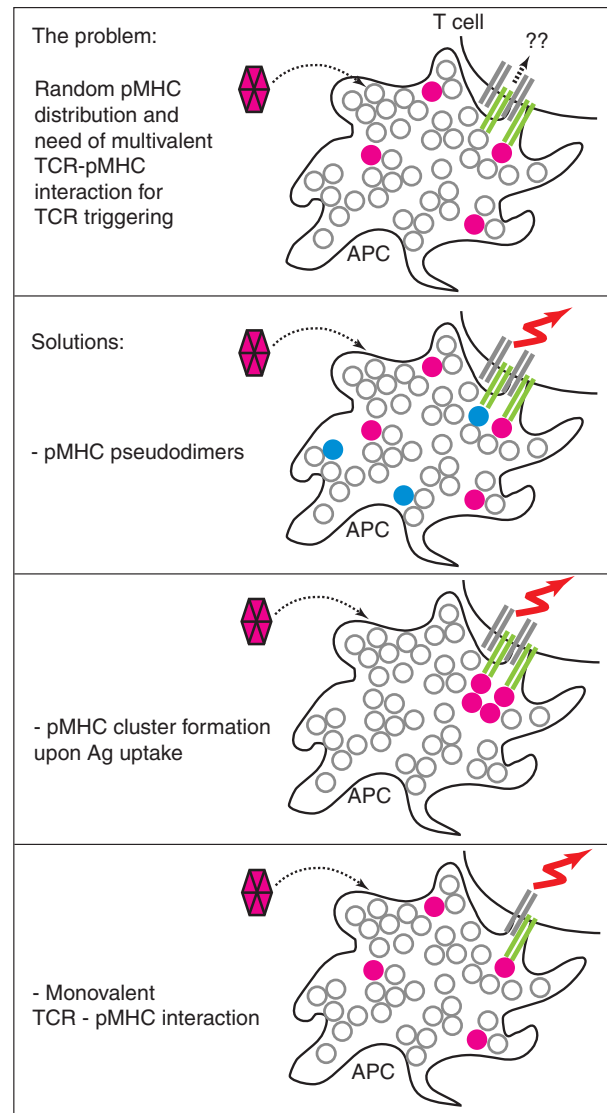


Figure 6 Cognate pMHC distribution and valency of stimulatory ligands. Representation of the proposed problem of T cell activation depending on multivalent cognate pMHC interaction and the random distribution of cognate pMHC upon Ag-uptake (top panel) and experimentally supported solutions (bottom panel). Antigenic entity (e.g., bacterium, virus, and foreign protein) represented by magenta-colored hexagonal structure; cognate pMHC, self-pMHC with co-stimulatory capacity, and other self-pMHC represented by filled magenta circles, filled blue circles, and grey-lined open circles, respectively.

It was proposed that a cognate pMHC and an adjacent self-pMHC with some minimal TCR-binding capacity might, in combination with the co-receptor, form a pseudodimer that could activate the T cell (Irvine *et al.*, 2002). Artificial, antibody-based soluble pMHC pseudodimers were indeed shown to be able to activate T cells in a co-receptor-dependent manner (Krogsgaard *et al.*, 2005). On the other hand, recent work has shown that clusters of cognate pMHC are readily detectable on APCs that have been infected with virus or have internalized a foreign antigen (Bosch *et al.*, 2013; Ferez *et al.*, 2014; Lu *et al.*, 2012), suggesting that under physiological conditions pMHC is forming multivalent complexes and can therefore be ready to bind TCR clusters in a multivalent fashion.

A more recent study, analyzing pMHC number and T cell activation in single T cell-APC pairs, suggests that when incorporated in the plasma membrane one pMHC is indeed sufficient to induce a cytokine secretion response in both naïve and memory T cells (Huang *et al.*, 2013). This would implicitly mean that the currently held dogma that memory T cells are intrinsically more sensitive than naïve T cells is not correct. One caveat to these studies is the assumption that the technique employed to quantify the pMHCs on the cell surface is truly 100% efficient and that the tetravalent fluorescent probe used to detect the pMHCs recognizes them on a one-on-one basis under the conditions employed. In an alternative approach, using the arguably less physiological but more easily quantifiable SLB model, pMHCs were confined in artificial corrals whose size was much smaller than the T cell-SLB contact area (Manz *et al.*, 2011). It was found that, while keeping the pMHC density constant (and thus the total number of pMHC available for the T cell), distribution of the pMHCs in smaller and smaller corrals led to a point where having on average less than 2 pMHCs in a corral abrogated T cell activation. This suggested that the relevant signaling unit is a TCR cluster that would need more than one pMHC to become activated.

Conclusions

The past decade has seen important progress in understanding the molecular basis of TCR triggering. It has evolved from a very TCR $\alpha\beta$ -centered description to a much more integrated model of TCR triggering where we have started to define the domains and mechanisms involved in transmitting the pMHC-TCR $\alpha\beta$ interaction event to the intracellular domains of the CD3 dimers. In addition to affinity, we now have good evidence that conformational change, force, clustering, and relocation are critically involved in TCR triggering. Further molecular analysis of the mechanisms underlying these aspects of TCR triggering may provide us with new targets for therapeutic intervention. Part of these advances in our understanding have been obtained with classical biochemical and cellular approaches, but a strong boost has been given by the introduction of the super-resolution light microscopy and BFP techniques in the field. The challenge will be to integrate our knowledge on the separate parts into a single coherent mechanism of TCR triggering. This may ultimately need a crystal structure of the complete,

membrane-embedded TCR complex in presence and absence of its pMHC ligand.

See also: Cellular Immunology: Transcriptional Basis of B and T Cell Lineages and Memory Cells: The T-Cell Receptor Signalingosome

References

- Aivazian, D., Stern, L.J., 2000. Phosphorylation of T cell receptor zeta is regulated by a lipid dependent folding transition. *Nature Structural & Molecular Biology* 7, 1023–1026.
- Alam, S.M., Davies, G.M., Lin, C.M., *et al.*, 1999. Qualitative and quantitative differences in T cell receptor binding of agonist and antagonist ligands. *Immunity* 10, 227–237.
- Alarcon, B., Swamy, M., Van Santen, H.M., Schamel, W.W., 2006. T-cell antigen-receptor stoichiometry: Pre-clustering for sensitivity. *EMBO Reports* 7, 490–495.
- Beddoe, T., Chen, Z., Clements, C.S., *et al.*, 2009. Antigen ligation triggers a conformational change within the constant domain of the alphabeta T cell receptor. *Immunity* 30, 777–788.
- Bettini, M.L., Guy, C., Dash, P., *et al.*, 2014. Membrane association of the CD3epsilon signaling domain is required for optimal T cell development and function. *Journal of Immunology* 193, 258–267.
- Boniface, J.J., Rabinowitz, J.D., Wulfig, C., *et al.*, 1998. Initiation of signal transduction through the T cell receptor requires the multivalent engagement of peptide/MHC ligands [corrected]. *Immunity* 9, 459–466.
- Boroto, A., Arellano, I., Blanco, R., *et al.*, 2014. Relevance of Nck-CD3 epsilon interaction for T cell activation *in vivo*. *Journal of Immunology* 192, 2042–2053.
- Boroto, A., Arellano, I., Dopfer, E.P., *et al.*, 2013. Nck recruitment to the TCR required for ZAP70 activation during thymic development. *Journal of Immunology* 190, 1103–1112.
- Bosch, B., Heipertz, E.L., Drake, J.R., Roche, P.A., 2013. Major histocompatibility complex (MHC) class II-peptide complexes arrive at the plasma membrane in cholesterol-rich microclusters. *Journal of Biological Chemistry* 288, 13236–13242.
- Brazin, K.N., Mallis, R.J., Li, C., *et al.*, 2014. Constitutively oxidized CXXC motifs within the CD3 heterodimeric ectodomains of the T cell receptor complex enforce the conformation of juxtaposed segments. *Journal of Biological Chemistry* 289, 18880–18892.
- Brodeur, J.F., Li, S., Da Silva Martins, M., Larose, L., Dave, V.P., 2009. Critical and multiple roles for the CD3epsilon intracytoplasmic tail in double negative to double positive thymocyte differentiation. *Journal of Immunology* 182, 4844–4853.
- Bunnell, S.C., Hong, D.I., Kardon, J.R., *et al.*, 2002. T cell receptor ligation induces the formation of dynamically regulated signaling assemblies. *Journal of Cell Biology* 158, 1263–1275.
- Campi, G., Varma, R., Dustin, M.L., 2005. Actin and agonist MHC-peptide complex-dependent T cell receptor microclusters as scaffolds for signaling. *Journal of Experimental Medicine* 202, 1031–1036.
- Castro, M., Van Santen, H.M., Ferez, M., *et al.*, 2014. Receptor pre-clustering and T cell responses: Insights into molecular mechanisms. *Frontiers in Immunology* 5, 132.
- Cordoba, S.P., Choudhuri, K., Zhang, H., *et al.*, 2013. The large ectodomains of CD45 and CD148 regulate their segregation from and inhibition of ligated T-cell receptor. *Blood* 121, 4295–4302.
- Choudhuri, K., Parker, M., Milicic, A., *et al.*, 2009. Peptide-major histocompatibility complex dimensions control proximal kinase-phosphatase balance during T cell activation. *Journal of Biological Chemistry* 284, 26096–26105.
- Choudhuri, K., Wiseman, D., Brown, M.H., Gould, K., Van Der Merwe, P.A., 2005. T-cell receptor triggering is critically dependent on the dimensions of its peptide-MHC ligand. *Nature* 436, 578–582.
- Daniels, M.A., Teixeira, E., Gill, J., *et al.*, 2006. Thymic selection threshold defined by compartmentalization of Ras/MAPK signalling. *Nature* 444, 724–729.
- Davis, S.J., van der Merwe, P.A., 2006. The kinetic-segregation model: TCR triggering and beyond. *Nature Immunology* 7, 803–809.

- DeFord-Watts, L.M., Dougall, D.S., Belkaya, S., *et al.*, 2011. The CD3 zeta subunit contains a phosphoinositide-binding motif that is required for the stable accumulation of TCR-CD3 complex at the immunological synapse. *Journal of Immunology* 186, 6839–6847.
- DeFord-Watts, L.M., Tassin, T.C., Becker, A.M., *et al.*, 2009. The cytoplasmic tail of the T cell receptor CD3 epsilon subunit contains a phospholipid-binding motif that regulates T cell functions. *Journal of Immunology* 183, 1055–1064.
- Demotz, S., Grey, H.M., Sette, A., 1990. The minimal number of class II MHC-antigen complexes needed for T cell activation. *Science* 249, 1028–1030.
- Doyle, C., Strominger, J.L., 1987. Interaction between CD4 and class II MHC molecules mediates cell adhesion. *Nature* 330, 256–259.
- Ferez, M., Castro, M., Alarcon, B., Van Santen, H.M., 2014. Cognate peptide-MHC complexes are expressed as tightly apposed nanoclusters in virus-infected cells to allow TCR crosslinking. *Journal of Immunology* 192, 52–58.
- Friedl, P., Brocker, E.B., 2002. TCR triggering on the move: Diversity of T-cell interactions with antigen-presenting cells. *Immunological Reviews* 186, 83–89.
- Gagnon, E., Schubert, D.A., Gordo, S., Chu, H.H., Wuchertpfennig, K.W., 2012. Local changes in lipid environment of TCR microclusters regulate membrane binding by the CD3epsilon cytoplasmic domain. *Journal of Experimental Medicine* 209, 2423–2439.
- Gil, D., Schamel, W.W., Montoya, M., Sanchez-Madrid, F., Alarcon, B., 2002. Recruitment of Nck by CD3 epsilon reveals a ligand-induced conformational change essential for T cell receptor signaling and synapse formation. *Cell* 109, 901–912.
- Gil, D., Schrum, A.G., Alarcon, B., Palmer, E., 2005. T cell receptor engagement by peptide-MHC ligands induces a conformational change in the CD3 complex of thymocytes. *Journal of Experimental Medicine* 201, 517–522.
- Gil, D., Schrum, A.G., Daniels, M.A., Palmer, E., 2008. A role for CD8 in the developmental tuning of antigen recognition and CD3 conformational change. *Journal of Immunology* 180, 3900–3909.
- Harding, C.V., Unanue, E.R., 1990. Quantitation of antigen-presenting cell MHC class II/peptide complexes necessary for T-cell stimulation. *Nature* 346, 574–576.
- Holst, J., Wang, H., Eder, K.D., *et al.*, 2008. Scalable signaling mediated by T cell antigen receptor-CD3 ITAMs ensures effective negative selection and prevents autoimmunity. *Nature Immunology* 9, 658–666.
- Huang, J., Brameshuber, M., Zeng, X., *et al.*, 2013. A single peptide-major histocompatibility complex ligand triggers digital cytokine secretion in CD4(+) T cells. *Immunity* 39, 846–857.
- Huang, J., Zarnitsyna, V.I., Liu, B., *et al.*, 2010. The kinetics of two-dimensional TCR and pMHC interactions determine T-cell responsiveness. *Nature* 464, 932–936.
- Huppa, J.B., Axmann, M., Mortelmaier, M.A., *et al.*, 2010. TCR-peptide-MHC interactions in situ show accelerated kinetics and increased affinity. *Nature* 463, 963–967.
- Husson, J., Chemin, K., Bohineust, A., HIVROZ, C., Henry, N., 2011. Force generation upon T cell receptor engagement. *PLoS One* 6, e19680.
- Irvine, D.J., Purbhoo, M.A., Krogsaard, M., Davis, M.M., 2002. Direct observation of ligand recognition by T cells. *Nature* 419, 845–849.
- James, J.R., Vale, R.D., 2012. Biophysical mechanism of T-cell receptor triggering in a reconstituted system. *Nature* 487, 64–69.
- Kaye, J., Janeway Jr., C.A., 1984. The Fab fragment of a directly activating monoclonal antibody that precipitates a disulfide-linked heterodimer from a helper T cell clone blocks activation by either allogeneic Ia or antigen and self-Ia. *Journal of Experimental Medicine* 159, 1397–1412.
- Kersh, E.N., Shaw, A.S., Allen, P.M., 1998. Fidelity of T cell activation through multistep T cell receptor zeta phosphorylation. *Science* 281, 572–575.
- Kim, S.T., Takeuchi, K., Sun, Z.Y., *et al.*, 2009. The alphabeta T cell receptor is an anisotropic mechanosensor. *Journal of Biological Chemistry* 284, 31028–31037.
- Kjer-Nielsen, L., Clements, C.S., Purcell, A.W., *et al.*, 2003. A structural basis for the selection of dominant alphabeta T cell receptors in antiviral immunity. *Immunity* 18, 53–64.
- Krogsaard, M., Li, Q.J., Sumen, C., *et al.*, 2005. Agonist/endogenous peptide-MHC heterodimers drive T cell activation and sensitivity. *Nature* 434, 238–243. **Epub 2005 Feb 20.**
- Krummel, M.F., Sjaastad, M.D., Wülfing, C., Davis, M.M., 2000. Differential clustering of CD4 and CD3 ζ during T cell recognition. *Science* 289, 1349–1352.
- Kuhns, M.S., Badgandi, H.B., 2012. Piecing together the family portrait of TCR-CD3 complexes. *Immunological Reviews* 250, 120–143.
- Kumar, R., Ferez, M., Swamy, M., *et al.*, 2011. Increased sensitivity of antigen-experienced T cells through the enrichment of oligomeric T cell receptor complexes. *Immunity* 35, 375–387.
- Li, Y.C., Chen, B.M., Wu, P.C., *et al.*, 2010. Cutting edge: Mechanical forces acting on T cells immobilized via the TCR complex can trigger TCR signaling. *Journal of Immunology* 184, 5959–5963.
- Lillemeier, B.F., Mortelmaier, M.A., Forstner, M.B., *et al.*, 2010. TCR and Lat are expressed on separate protein islands on T cell membranes and concatenate during activation. *Nature Immunology* 11, 90–96.
- Liu, B., Chen, W., Evavold, B.D., Zhu, C., 2014. Accumulation of dynamic catch bonds between TCR and agonist peptide-MHC triggers T cell signaling. *Cell* 157, 357–368.
- Lu, X., Gibbs, J.S., Hickman, H.D., *et al.*, 2012. Endogenous viral antigen processing generates peptide-specific MHC class I cell-surface clusters. *Proceedings of the National Academy of Sciences of the United States of America* 109, 15407–15412.
- Mallaun, M., Naehrer, D., Daniels, M.A., *et al.*, 2008. The T cell receptor's alpha-chain connecting peptide motif promotes close approximation of the CD8 coreceptor allowing efficient signal initiation. *Journal of Immunology* 180, 8211–8221.
- Manz, B.N., Jackson, B.L., Petit, R.S., Dustin, M.L., Groves, J., 2011. T-cell triggering thresholds are modulated by the number of antigen within individual T-cell receptor clusters. *Proceedings of the National Academy of Sciences of the United States of America* 108, 9089–9094.
- Martinez-Martin, N., Risueno, R.M., Morreale, A., *et al.*, 2009. Cooperativity between T cell receptor complexes revealed by conformational mutants of CD3epsilon. *Science Signaling* 2, ra43.
- McKeithan, T.W., 1995. Kinetic proofreading in T-cell receptor signal transduction. *Proceedings of the National Academy of Sciences of the United States of America* 92, 5042–5046.
- Mingueneau, M., Sansoni, A., Gregoire, C., *et al.*, 2008. The proline-rich sequence of CD3epsilon controls T cell antigen receptor expression on and signaling potency in preselection CD4+ CD8+ thymocytes. *Nature Immunology* 9, 522–532.
- Minguet, S., Swamy, M., Alarcon, B., Luescher, I.F., Schamel, W.W., 2007. Full activation of the T cell receptor requires both clustering and conformational changes at CD3. *Immunity* 26, 43–54. **Epub 2006 Dec 21.**
- Naehrer, D., Daniels, M.A., Hausmann, B., *et al.*, 2007. A constant affinity threshold for T cell tolerance. *Journal of Experimental Medicine* 204, 2553–2559.
- Nika, K., Soldani, C., Salek, M., *et al.*, 2010. Constitutively active Lck kinase in T cells drives antigen receptor signal transduction. *Immunity* 32, 766–777. **Epub 2010 Jun 11.**
- Norment, A.M., Salter, R.D., Parham, P., Engelhard, V.H., Littman, D.R., 1988. Cell-cell adhesion mediated by CD8 and MHC class I molecules. *Nature* 336, 79–81.
- Palmer, E., Naehrer, D., 2009. Affinity threshold for thymic selection through a T-cell receptor-co-receptor zipper. *Nature Reviews Immunology* 9, 207–213.
- Rabinowitz, J.D., Beeson, C., Lyons, D.S., Davis, M.M., McConnell, H.M., 1996. Kinetic discrimination in T-cell activation. *Proceedings of the National Academy of Sciences of the United States of America* 93, 1401–1405.
- Reay, P.A., Matsui, K., Haase, K., *et al.*, 2000. Determination of the relationship between T cell responsiveness and the number of MHC-peptide complexes using specific monoclonal antibodies. *Journal of Immunology* 164, 5626–5634.
- Rhee, I., Veillette, A., 2012. Protein tyrosine phosphatases in lymphocyte activation and autoimmunity. *Nature Immunology* 13, 439–447.
- Risueno, R.M., Gil, D., Fernandez, E., Sanchez-Madrid, F., Alarcon, B., 2005. Ligand-induced conformational change in the T-cell receptor associated with productive immune synapses. *Blood* 106, 601–608. **Epub 2005 Mar 24.**
- Risueno, R.M., Van Santen, H.M., Alarcon, B., 2006. A conformational change senses the strength of T cell receptor-ligand interaction during thymic selection. *Proceedings of the National Academy of Sciences of the United States of America* 103, 9625–9630. **Epub 2006 Jun 9.**
- Rosette, C., Werlen, G., Daniels, M.A., *et al.*, 2001. The impact of duration versus extent of TCR occupancy on T cell activation: A revision of the kinetic proofreading model. *Immunity* 15, 59–70.
- Roybal, K.T., Sinai, P., Verkade, P., Murphy, R.F., Wulfig, C., 2013. The actin-driven spatiotemporal organization of T-cell signaling at the system scale. *Immunological Reviews* 256, 133–147.
- Rudd, C.E., Trevillyan, J.M., Dasgupta, J.D., Wong, L.L., Schlossman, S.F., 1988. The CD4 receptor is complexed in detergent lysates to a protein-tyrosine kinase (pp58) from human T lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America* 85, 5190–5194.
- Rudolph, M.G., Stanfield, R.L., Wilson, I.A., 2006. How TCRs bind MHCs, peptides, and coreceptors. *Annual Review of Immunology* 24, 419–466.
- Schamel, W.W., Arechaga, I., Risueno, R.M., *et al.*, 2005. Coexistence of multivalent and monovalent TCRs explains high sensitivity and wide range of response. *Journal of Experimental Medicine* 202, 493–503. **Epub 2005 Aug 8.**

- Secrist, J.P., Burns, L.A., Karnitz, L., Koretzky, G.A., Abraham, R.T., 1993. Stimulatory effects of the protein tyrosine phosphatase inhibitor, pervanadate, on T-cell activation events. *Journal of Biological Chemistry* 268, 5886–5893.
- Sherman, E., Barr, V., Manley, S., *et al.*, 2011. Functional nanoscale organization of signaling molecules downstream of the T cell antigen receptor. *Immunity* 35, 705–720.
- Shi, X., Bi, Y., Yang, W., *et al.*, 2013. Ca²⁺ regulates T-cell receptor activation by modulating the charge property of lipids. *Nature* 493, 111–115.
- Sigalov, A.B., Aivazian, D.A., Uversky, V.N., Stern, L.J., 2006. Lipid-binding activity of intrinsically unstructured cytoplasmic domains of multichain immune recognition receptor signaling subunits. *Biochemistry* 45, 15731–15739.
- Sloan-Lancaster, J., Shaw, A.S., Rothbard, J.B., Allen, P.M., 1994. Partial T cell signaling: Altered phospho-zeta and lack of zap70 recruitment in APL-induced T cell anergy. *Cell* 79, 913–922.
- Smith-Garvin, J.E., Koretzky, G.A., Jordan, M.S., 2009. T cell activation. *Annual Review of Immunology* 27, 591–619.
- Springer, T.A., 1990. Adhesion receptors of the immune system. *Nature* 346, 425–434.
- Starr, T.K., Jameson, S.C., Hogquist, K.A., 2003. Positive and negative selection of T cells. *Annual Review of Immunology* 21, 139–176. Epub 2002 Oct 16.
- Sykulev, Y., Joo, M., Vturina, I., Tsomides, T.J., Eisen, H.N., 1996. Evidence that a single peptide-MHC complex on a target cell can elicit a cytolytic T cell response. *Immunity* 4, 565–571.
- Varma, R., Campi, G., Yokosuka, T., Saito, T., Dustin, M.L., 2006. T cell receptor-proximal signals are sustained in peripheral microclusters and terminated in the central supramolecular activation cluster. *Immunity* 25, 117–127.
- Veillette, A., Bookman, M.A., Horak, E.M., Bolen, J.B., 1988. The CD4 and CD8 T cell surface antigens are associated with the internal membrane tyrosine-protein kinase p56lck. *Cell* 55, 301–308.
- Wang, Y., Becker, D., Vass, T., *et al.*, 2009. A conserved CXXC motif in CD3epsilon is critical for T cell development and TCR signaling. *PLOS Biology* 7, e1000253.
- Werlen, G., Hausmann, B., Palmer, E., 2000. A motif in the alphabeta T-cell receptor controls positive selection by modulating ERK activity. *Nature* 406, 422–426.
- Xu, C., Gagnon, E., Call, M.E., *et al.*, 2008. Regulation of T cell receptor activation by dynamic membrane binding of the CD3epsilon cytoplasmic tyrosine-based motif. *Cell* 135, 702–713.
- Yokosuka, T., Saito, T., 2010. The immunological synapse, TCR microclusters, and T cell activation. *Current Topics in Microbiology and Immunology* 340, 81–107.
- Yokosuka, T., Sakata-Sogawa, K., Kobayashi, W., *et al.*, 2005. Newly generated T cell receptor microclusters initiate and sustain T cell activation by recruitment of Zap70 and SLP-76. *Nature Immunology* 6, 1253–1262.
- Zhang, H., Cordoba, S.P., Dushek, O., Van Der Merwe, P.A., 2011. Basic residues in the T-cell receptor zeta cytoplasmic domain mediate membrane association and modulate signaling. *Proceedings of the National Academy of Sciences of the United States of America* 108, 19323–19328.

Single Molecule Methods to Measure Receptor–Ligand Interaction in Immunological Synapses

JB Huppa, Medical University of Vienna, Vienna, Austria

GJ Schütz, Vienna University of Technology, Vienna, Austria

© 2016 Elsevier Inc. All rights reserved.

Introduction

When measured in solution, T-cell antigen receptors (TCRs) bind their cognate ligand with surprisingly low affinity (micro molar range). Nonetheless, T-cells are exquisitely sensitive as they can sense the presence of even a single antigenic pMHC among thousands of structurally very similar but nonstimulatory endogenous pMHCs on the APC surface (Huang *et al.*, 2013). How is this possible? We think that much of the solution to this problem lies in studying TCR–pMHC in the right place, i.e. within the immunological synapse, the transient cell contact between a T-cell and a professional APC or target cell (Norcross, 1984) (reviewed in Bromley *et al.*, 2001; Huppa and Davis, 2003), to which T-cell antigen recognition is invariably linked.

Diversity of Immunological Synapses

This intercellular membrane structure was first resolved in its mature form, where it is highly organized with a bull's eye morphology (Monks *et al.*, 1998): its center, termed 'central supra-molecular activation cluster' (cSMAC), accumulates TCR-CD3 and is bordered by the ring-like 'peripheral supra-molecular activation cluster' (pSMAC), which contains the integrin LFA-1. An outer most ring harboring large glycoproteins such as CD43, CD44 and CD45 was later termed distal SMAC (dSMAC) (Johnson *et al.*, 2000; Leupin *et al.*, 2000). Depending on the interacting partners, immunological synapses are highly diverse in composition and function (reviewed in Friedl *et al.*, 2005) and therefore the term 'synapse' is nowadays used to describe any transient bi-membranous junction between T-cells and APCs (Huppa and Davis, 2003). The common denominator of all immune synapses is that they serve as a physical platform for one-on-one conversations between two conjugated cells: through engagement of cell-specific receptor–ligand pairs and directed secretion of effector molecules.

Molecular Size Hypothesis

Regardless of its physiology, each type of immunological synapse sets the stage for all interactions taking place within. While we know in considerable detail the biochemical and structural basis of many of such interactions in solution (Garcia *et al.*, 2009; Marrack *et al.*, 2008), the problem of how to place and evaluate them under the influence of their native environment is not trivial and far from being solved (for review see Dustin *et al.*, 2001). Synaptic binding is quite different from binding in solution in many respects. Molecular partners are pre-oriented on flexible surfaces, which can adjust their morphology in response to cell stimuli and acting forces.

This might greatly increase the on-rate of binding, but only if the two participating membranes are properly aligned. As was first proposed and experimentally pursued by Anton van der Merwe and Michael Dustin, molecular size is likely to influence binding significantly within the crowded synaptic confines (Irles *et al.*, 2003; Choudhuri *et al.*, 2005; Shaw and Dustin, 1997). This hypothesis is difficult to verify but is in line with theoretical predictions (Weikl and Lipowsky, 2004; Burroughs and Wulfig, 2002) and was recently even further supported by experiments involving a reconstituted T-cell-free cell system (James and Vale, 2012). In such artificial synapses the bulky and highly expressed protein tyrosine phosphatase CD45 was observed to be excluded from the more central parts of the cell contact only by means of its size. As a consequence CD45 had no longer access to its target, the much smaller TCR, which shifted the equilibrium toward TCR-CD3 complex phosphorylation driven by accessory kinases.

Forces Acting at Synapses

Cytoskeletal forces are also crucially involved in T-cell antigen recognition. Their existence is not only implied by membrane ruffling and a continuous flow of actin during synapse formation (Kaizuka *et al.*, 2007), but also by the energy-driven formation of TCR microclusters and their directed movement from the periphery to the cSMAC (Kaizuka *et al.*, 2007; Varma *et al.*, 2006). Depending on the strength of the antigenic stimulus and costimulation, synapse maturation – i.e. the concerted formation of cSMAC-, pSMAC- and dSMAC – occurs at a fast pace, i.e. within one to 30 min (Beemiller and Krummel, 2010; Beemiller *et al.*, 2012; Burkhardt *et al.*, 2008; Wulfig and Davis, 1998; Wulfig *et al.*, 2002). CD8⁺ cytolytic T-cells (CTLs) were even shown to rip off antigenic MHC class I molecules from the surface of the target cell (Huang *et al.*, 1999). Traction forces, which were measured on T-cells in contact with stimulatory arrays of antibodies reactive against CD3, reach up to 100 pN (Bashour *et al.*, 2014). Directly or indirectly, such forces may indeed influence the function of integrins such as LFA-1, which switch in response to pN-range shear forces from a low affinity to a high affinity conformation (Kong *et al.*, 2009).

Synaptic forces might also be instrumental to help the TCR and pMHC overcome steric barriers posed by larger proteins of the glycocalyx. However, they are also bound to destabilize low affinity interactions between TCR and pMHC. More stable TCR–pMHC interactions might just benefit from cellular motility due an increased on-rate, but weaker interactions could suffer considerably because of forbidding off-rates. Along this rationale, forces may well contribute to T-cell ligand discrimination merely by enforcing kinetic proofreading (Klotzsch and Schütz, 2013). Various lines of experimentation

even support the idea that forces acting directly on the TCR trigger T-cells through deformation of the TCR-CD3 complex (Kim *et al.*, 2009; Adams *et al.*, 2011; Lim *et al.*, 2012; Judokusumo *et al.*, 2012; Xie *et al.*, 2012).

Membranes Are Compartmentalized

It is exceedingly difficult to predict the effect of membrane compartmentalization on the TCR–pMHC binding dynamics. After all, the organization of plasma membranes is still controversially discussed (Simons and Gerl, 2010). However, several groups have reported the residence of TCRs and pMHCs in specific membrane microdomains (Lillemeier *et al.*, 2006; Schamel *et al.*, 2005; Williamson *et al.*, 2011). Live-cell high speed Photoactivation Localization Microscopy (hsPALM) studies designed to analyze the surface distribution below the diffraction limit of visible light revealed TCR-enriched domains of 20–70 nm in diameter (Lillemeier *et al.*, 2010). These entities were stable for several seconds, contained 1500 to 2200 TCR molecules per μm^2 , which were still mobile. Based on structural estimates a single TCR-CD3 complex covers roughly 40 nm² (Kuhns *et al.*, 2010). TCRs should therefore occupy between 6 and 9% of the surface accessible to them and bounce as a result into one another at random at a high frequency. Whether this might lead to the formation of higher order TCR structures with consequences for pMHC binding is still open to debate. Regardless of this issue, the mere existence of TCR- and pMHC enriched plasma membrane domains should affect the on-rate of TCR–pMHC binding, especially if both types of domains are properly aligned.

Measuring Binding *In Situ*

To validate synapse-specific parameters affecting TCR–pMHC binding there is no other way than to compare TCR–pMHC binding as it occurs *in situ* with that measured *in vitro*. The latter reveals binding properties inherent to the molecules under investigation. Surface plasmon resonance and to some extent calorimetry are the methods of choice. Deviations from binding *in vitro*, as they are found under defined synaptic condition, can then be used to assess synapse-specific impacts on binding. At current there are two experimental approaches for this. One involves a mechanical assay with single molecule resolution, the other single molecule fluorescence microscopy with the use of glass-supported artificial lipid bilayers (SLBs).

Mechanical Assays to Measure 2D-kinetics

One way to determine the 2-dimensional kinetics of TCR–pMHC binding was approached by Zhu and colleagues, who applied for this purpose the so-called adhesion frequency assay and a thermofluctuation assay (Huang *et al.*, 2010). Here, a T-cell aspirated to a micropipette is confronted with a pMHC-functionalized biomembrane force probe (BFP), either a red blood cell or a bead attached to the apex of a red blood cell. When the T-cell is displaced from the BFP either directly through translational movement away from the BFP (adhesion

frequency assay) or through random thermic motion (thermofluctuation assay), the BFP becomes briefly deformed as the result of molecular binding, which can be observed by conventional light microscopy. Remarkably, such deformations can be reflective of single TCR–pMHC bonds, and hence, on- and off-rates as well as the equilibrium constant of the binding process can be extracted for the special two-dimensional scenario, in which the aspirated T-cell contacts the BFP (see Figure 1).

Zhu and colleagues found for the OT-1 TCR system, that the stimulatory potency of pMHCs correlated very well with the measured 2D-affinity, which was actually much higher than would have been predicted based on 3D-binding measurements. An exceedingly high on-rate of binding caused this strong 2D-binding. Unexpectedly, the off-rate was found to be accelerated too, with agonist pMHC dissociating the fastest, i.e., more than 8000 times faster than when measured for 3D-binding.

Hence, the geometrical constraints appear to have a strong influence on the molecular dynamics of ligand interactions, and, at least in the case of OT-1 CD8⁺ T-cells, T-cell activation is likely caused by serial TCR engagement. Unlike what has been previously proposed based on 3D measurements, the stimulatory potency of the pMHC did not correlate with the drastically reduced stability of the TCR–pMHC complex. Such unexpected binding behavior must be reflective of T-cell-specific parameters, in particular the organization, orientation and also conformation of the TCR on the cell surface. These are not at all accounted for in 3D-binding assays involving purified soluble proteins, and hence have been largely ignored for this reason.

The use of wild-type MHC and a mutant, which is no longer capable of binding the CD8 coreceptor, allowed dissecting the contribution of CD8-MHC binding to TCR–pMHC binding (Jiang *et al.*, 2011). The kinetics of adhesion were found to proceed through two stages, one involving predominantly pMHC and the TCR followed by a second stage, which coincided with a marked increase in adhesion and which required both CD8-binding and the activity of its associated TCR-proximal tyrosine kinase LCK. This rapid increase in affinity is probably best explained through cooperative binding of TCR and CD8 to pMHC. If this holds true, then cell adhesion synergizes with the stimulatory potency of the pMHC–TCR engagement: more potent ligands would lead to a stronger participation of CD8 in the overall binding process. After all, this process might provide another means for ligand discrimination.

Imaging TCR–pMHC Interactions *In Situ*

Motivated for very similar reasons as Zhu and colleagues, we had devised a noninvasive ultrasensitive live-cell imaging approach to measure TCR–pMHC interactions in their native synaptic environment. Our methodology is based on single molecule microscopy and Förster resonance energy transfer (FRET) (Huppa *et al.*, 2010). FRET is a quantum-mechanical phenomenon involving the radiation-less interaction of dipole momenta of two corresponding fluorescent dyes (for a review on single molecule FRET see Ha, 2001). Because the maximal

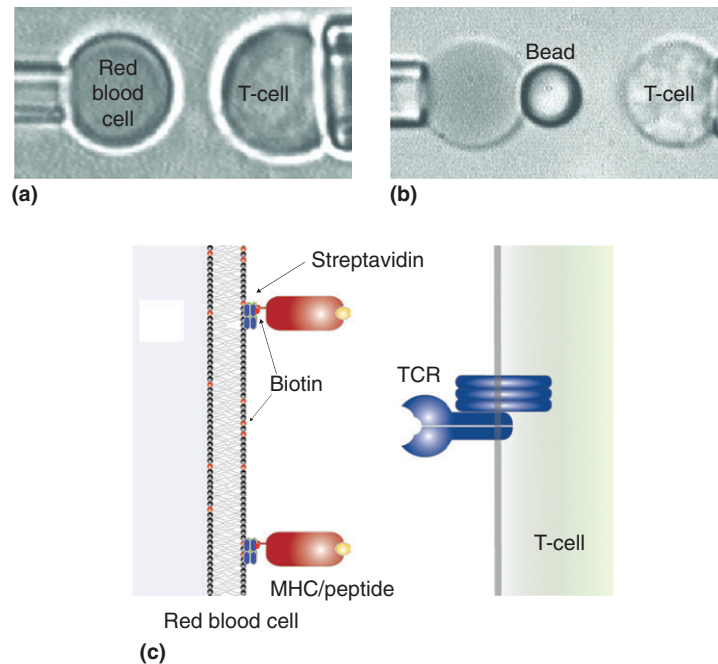


Figure 1 Micropipette and bioforce probe measurements of 2D-binding constants. (a, b) Micrographs of the micropipette and bioforce probe (BFP) measurements. A T-cell (right) is aspirated by a pipette and aligned with a red blood cell held stationary by another pipette (left) without (a) or with a bead attached (b) to the apex. These experimental setups are employed to measure binding events as visualized by the deformation of the BFP in adhesion frequency assays (a) or thermofluctuation assays (b). The sensitivity of these systems reaches down to single TCR–pMHC bonds. (c) Scheme illustrating RBCs (left) are decorated via streptavidin with biotinylated monomeric pMHC to interact with TCRs on T-cells (right). Panels (a) and (b) reprinted by Huang, J., Zarnitsyna, V.I., Liu, B., *et al.*, 2010. The kinetics of two-dimensional TCR and pMHC interactions determine T-cell responsiveness. *Nature* 464, 932–936, with permission from Macmillan Publishers Ltd.

energy transfer rate is inversely proportional to the sixth power of the inter-dye distance, FRET efficiency can be used to assess whether and at what time two labeled proteins are interacting. We applied this methodology to detect and quantify TCR–pMHC binding events directly *in situ*. For this, 5c.c7 TCR transgenic helper T-cells were decorated with a site-specifically fluorescently labeled anti-TCR single chain fragment (FRET donor) and confronted with a functionalized fluid lipid bilayer, which served as a surrogate APC as it presents site-specifically labeled pMHC (FRET acceptor) as well as ICAM-1 and B7-1 (Figure 2).

Both, the choice of recombinant proteins as FRET partners, which were bioengineered to carry a fluorophore at a defined site, and the use of planar lipid bilayers as APCs proved essential for the success of the study. Planar supported lipid bilayers had been pioneered decades ago by Harden McConnell (Brian and McConnell, 1984) but were really brought to the attention of immunologists by Michael Dustin with the first spatiotemporal description of synapse formation and maturation (Grakoui *et al.*, 1999). Its use provides many operational advantages: it allows for the defined reconstitution of the APC membrane and reduces a three-dimensional problem to two dimensions, thereby enabling the recording of fast and transient processes such as TCR–pMHC binding. Importantly, planar bilayers are compatible with total internal reflection microscopy (TIRF), which reduces background fluorescence substantially (Axelrod, 2008). Prior to our study, TIRF microscopy had helped to visualize the synaptic formation of synaptic entities, such as TCR microclusters, and to deduce

cell-biological mechanisms underlying synaptic signaling (Campi *et al.*, 2005; Varma *et al.*, 2006; Yokosuka *et al.*, 2005; Bunnell *et al.*, 2006; Huse *et al.*, 2007). In our study TIRF microscopy improved the signal to noise ratio to such an extent that we could resolve single fluorophores and single molecule FRET events and track individual pMHC–TCR interactions over time.

By doing so we found that, compared to solution binding, synaptic binding of the TCR was significantly increased in strength with both an accelerated association (about 100-fold) and an accelerated dissociation (about 4–12-fold) (Huppa *et al.*, 2010). Interestingly, depolymerization of the T-cell's cortical actin resulted in a remarkable stabilization of synaptic TCR–pMHC complexes, an observation that implicates mechanical cellular forces in disrupting TCR–pMHC complexes. It is conceivable that such forces indeed contribute to TCR–ligand discrimination and to TCR-proximal signaling for reasons discussed below.

In agreement with previous studies, addition of a CD4 blocking-antibody interfered with TCR-proximal signaling at all times, yet had, in contrast to CD8 blockade on CTLs, only a minor effect on the TCR–pMHC interaction. Hence, other avenues, for example the visualization of CD4–MHC interactions, will be needed to shed light on the enigmatic mechanism, by which CD4 boosts T-cell sensitivity.

We complemented our FRET study using single molecule tracking of fluorescently labeled pMHC (Axmann *et al.*, 2012). Individual pMHC molecules were recorded as they moved via Brownian motion within the synapse, with occasional stops

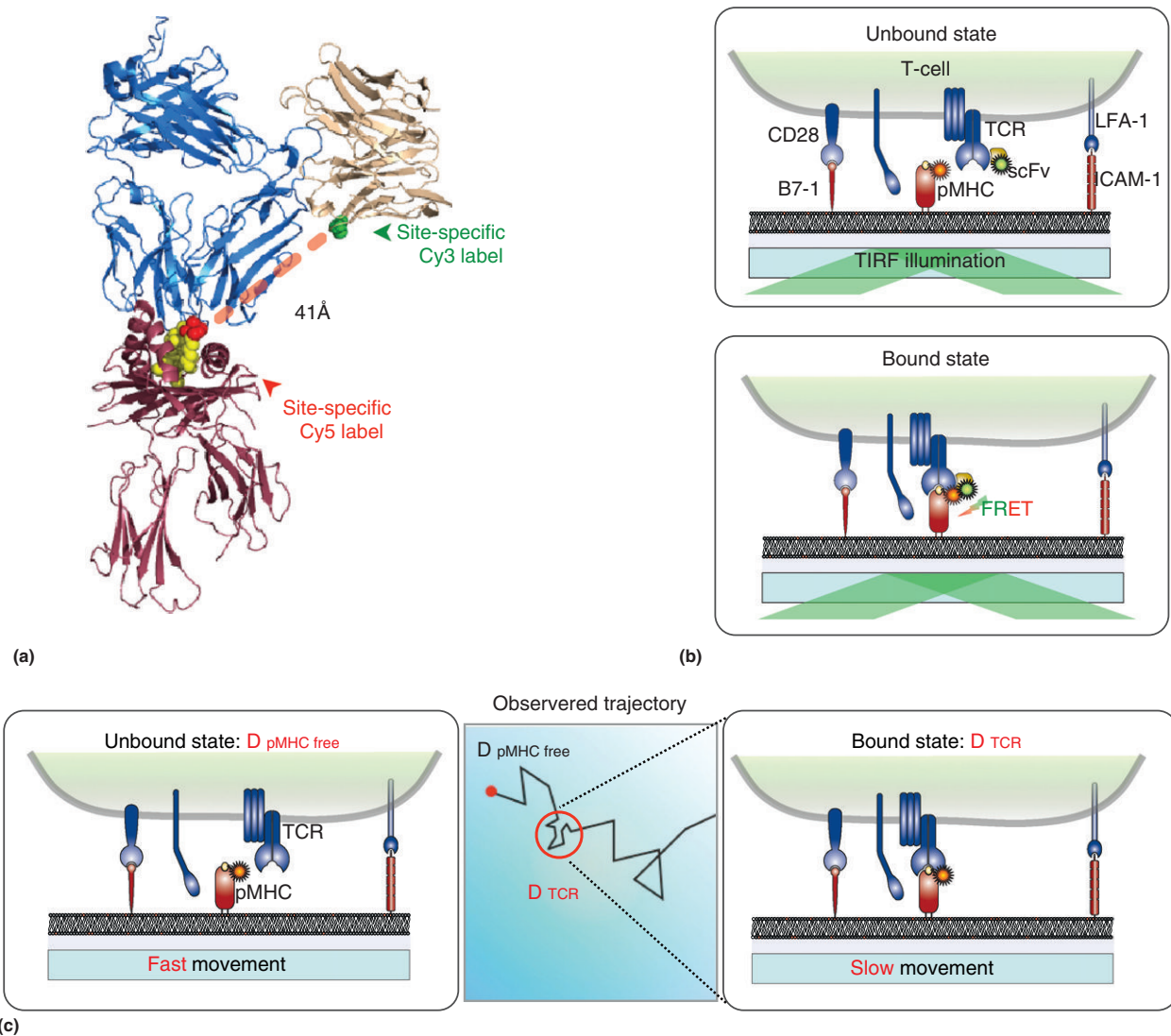


Figure 2 Förster resonance energy transfer (FRET)-based methodology (a, b) and single molecule tracking of fluorescently labeled pMHC (c) to measure TCR–pMHC interaction within the immunological synapse. (a) Shown is a composite model of the FRET system based on the TCR–pMHC and TCR–H57 Fab structures. The TCR is shown in blue, the MHC molecule in red, the peptide in yellow, and the scFv in light brown. The distance between the Cy5 (alternatively Alexa Fluor 647, FRET acceptor) attachment site (red) and the Cy3 (alternatively Alexa Fluor 555, FRET donor) attachment site is about 41 Å, less than the Förster radius of the Cy3–Cy5 (Alexa Fluor 555–Alexa Fluor 647) system, the distance of half-maximal energy transfer. (b) Schematic illustration of the experimental setup: TCR-transgenic T-cells interact with a planar glass-supported lipid bilayer, which harbors B7-1 (costimulator), ICAM1 and site-specifically labeled pMHCs. A FRET signal can only be produced when TCR and pMHC are close enough to bind to each other. The use of planar lipid bilayers allows the application of total internal reflection fluorescence microscopy (TIRF), in which only a thin slice of the T-cell contacting the bilayer (about 100 nm in thickness) is illuminated, which reduces background fluorescence substantially. (c) Illustrated is the rationale underlying the single molecule tracking approach to measure TCR–pMHC interaction times. Compared to bilayer-embedded pMHCs, T-cell-associated TCR–CD3 complexes diffuse laterally with a ~ 10 -times lower diffusion constant D . As a consequence fluorescence-labeled pMHCs slow down significantly in their lateral movement when they bind to a TCR. We have exploited this property to determine TCR–pMHC off-rates in the synaptic 2D-case.

due to binding to the TCR at the opposing T-cell surface. We quantified the immobilization lifetimes, which yielded similar interaction durations as those determined via the FRET-based method. This was *a priori* not expected, as pMHC rebinding to different TCRs in the immediate vicinity – for example, within a TCR microcluster – would have prolonged the detected pMHC immobilizations. Together, the two data sets thus indicate that – at least under the applied conditions of

high TCR–pMHC interaction probability – immediate pMHC rebinding does not occur at statistically relevant levels.

There are more implications associated with the results of our studies: first, interactions between TCRs and endogenous ligands in the periphery and within the thymus, which are too weak to be detected *in vitro*, could indeed occur in the synapse. And secondly, in light of the synaptic geometry and its influence on molecular interactions, the degree, to which low

affinity interactions take place in the synapse, is no longer only encoded within the molecular properties of the two binding partners involved but might very well depend on the properties and behavior of the two conjugated cells. Consequently, synaptic binding can no longer be comprehensively expressed with one single binding constant (as is the case for *in vitro* studies) but only described with the use of effective parameters, which are valid in a given synaptic region at a given time. The latter point is especially emphasized in our FRET-based study: TCR affinity varied substantially, that is up to 250-fold, within different synaptic regions and also between different cells. In other words, TCR–pMHC binding, and probably all other interactions of low to moderate strength, if not the entire process of antigen recognition are controlled through not well-understood cell biological and cellular parameters, which could also be regulated in T-cell development, in response to environmental cues and under disease conditions.

It should be noted that a recent publication from the Groves lab challenged the view of increased TCR–pMHC unbinding kinetics (O'Donoghue *et al.*, 2013): in their study, O'Donoghue *et al.* also employed single molecule tracking of fluorescently labeled pMHC in T-cell synapses with functionalized SLBs, however, they used rather long illumination times so that only bound pMHC molecules were imaged as individual spots, whereas unbound molecules generated a rather homogeneous background. From their data, the lifetime of pMHC immobilizations could be directly determined by

measuring the observation time of the spots, yielding approximately 30-fold slower interaction kinetics than our single molecule FRET and tracking data. At this stage, it is difficult to clarify this discrepancy: it may involve different stimulation conditions, or the presence of statistical subsets of TCR–pMHC pairs with different unbinding kinetics.

Pulling Forces Enhance Discriminative Power of Peptide Recognition

As outlined above, the two-dimensional cell-to-cell interface within the synapse not only affects the on-rate of protein–protein trans-interactions, but – due to pulling forces – also the off-rate. It is surprising that pulling forces have hardly been recognized as modulators of binding properties of cell surface receptors in general, given the large amount of literature on adhesion receptors (for review see e.g., Evans and Calderwood, 2007). Dembo *et al.* first discussed how pulling may either strengthen or weaken a bond, giving rise to the terminology ‘catch bonds’ versus ‘slip bonds,’ respectively (Dembo *et al.*, 1988) (Figure 3). Various conceptual models were proposed for catch-bond behavior (Thomas *et al.*, 2008); the most intelligible concept may be the hook model, in which receptor and ligand are connected like two hooks: with increasing force, the bond tightens and the bond lifetime gets prolonged. In contrast, slip bonds break with increasing force. This behavior can be best explained by a force-induced reduction of the

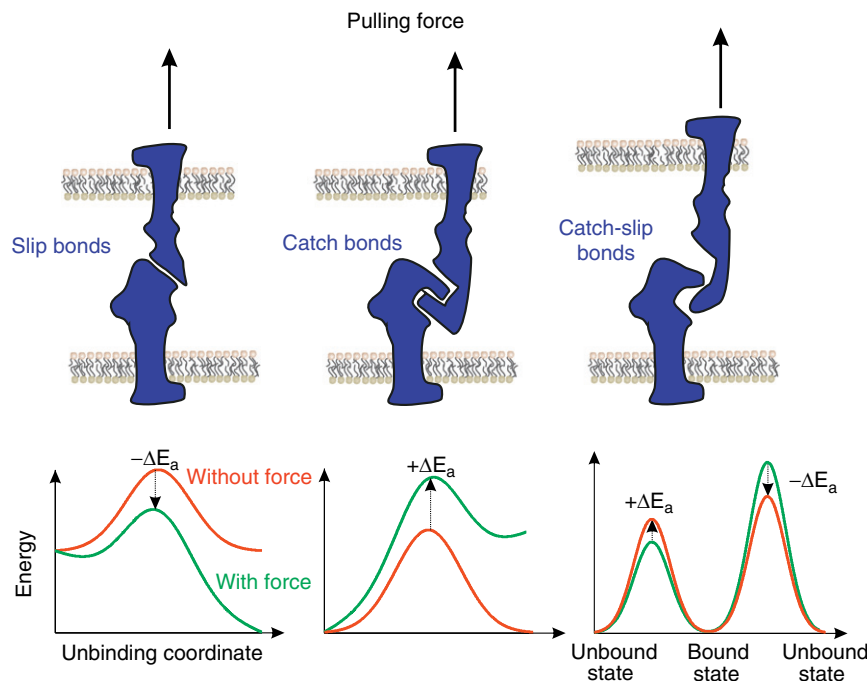


Figure 3 Effect of forces on receptor–ligand unbinding. The top row shows in a cartoon how cell surface receptors may contact each other, the bottom row the according energy landscape. In case of slip bonds (left), pulling forces reduce the activation energy for unbinding, E_a , thereby reducing the average bond lifetime. In contrast, catch bonds (middle) lead to an increase in the activation energy, so that the binding tightens. Such behavior can be explained, for example, by a hook-model, as indicated in the sketch above. Catch-slip bonds (right) show catch-bond behavior at low forces, and slip bond behavior at high forces. In that case, two force-dependent unbinding pathways have to be considered: at low forces, the catch-bond behavior is dominant ($+\Delta E_a$; unbinding to the left in the energy diagram), whereas at higher forces the slip bond behavior gets dominant ($-\Delta E_a$; unbinding to the right).

activation barrier, so that the average bond lifetime is decreased. Also two-pathway models are discussed, for example, with catch-bonds behavior at low forces and slip bond behavior at high forces (catch-slip bonds) (Thomas *et al.*, 2008). Indeed, for the TCR–pMHC interaction both catch bond and slip bond behavior has been reported (Robert *et al.*, 2012; Liu *et al.*, 2014).

Force-induced unbinding makes it thus essentially impossible to predict kinetic rate constants from *in vitro* experiments: the obtained values may be too short in case of catch bonds, or too long for slip bonds. Quantitatively, the effect on particular receptor–ligand pairs may vary substantially for different peptides, with up to 10-fold accelerated or decelerated kinetics being absolutely possible (Klotzsch and Schütz, 2013).

In the kinetic proofreading model, TCR–pMHC bond lifetimes determine the degree of ITAM phosphorylation, and thus ultimately the activation state of the T-cell. Force modulated unbinding could thus influence the power of ligand discrimination: At the extreme, two different TCR–pMHC bonds – a catch bond and a slip bond – with identical off-rate at zero force will separate into a high off-rate and a low off-rate bond at high force. Indeed, Liu *et al.* reported differential behavior of agonistic peptides (catch-slip bonds) and antagonistic peptides (slip bonds) for CD8⁺ T cells (Liu *et al.*, 2014). Pulling forces thus can generate additional contrast in the rupture behavior of different peptides.

We recently raised the question whether force-induced unbinding may also improve the specificity of pMHC discrimination in case of slip bonds (Klotzsch and Schütz, 2013). Our rationale was inspired by rupture curves of receptor–ligand pairs recorded via atomic force microscopy. In such experiments, force ramps of constant pulling velocity are applied, which result in the continuous reduction of the activation barrier and a concomitant increase in unbinding probability. The effect on the statistics of the bond lifetimes is dramatic: without force an exponential lifetime distribution is obtained; the pulling results in a pronounced peak of the lifetime and a steep slope toward higher values. In other words, higher lifetimes than the peak are extremely rare. This curve shape makes discrimination extremely reliable, with false positives being highly unlikely.

Conclusion and Outlook

We have made a case that synaptic receptor–ligand binding dynamics are not merely governed by structural and biochemical properties of the binding partners involved. *In situ* interactions are moreover subjected to geometrical constraints present in the narrow synaptic cleft and also to cell-biological parameters controlled by the conjugated cells. This means that the decision whether a T-cell remains quiescent or whether it becomes activated – especially when faced with few antigenic pMHCs – depends not only on the pMHC identity but also on the cell surface context, in which TCRs are active and their ligands are expressed. We anticipate, that the single molecule studies presented herein are only the beginning of an exciting scientific expedition into the synaptic ‘nanocosm,’ a journey, which is ultimately driven by the quest for understanding the foundation of the phenomenal T-cell antigen sensitivity.

In addition to what we have discussed above, a number of factors come immediately into mind: cell adhesion, costimulation, the nanoscale organization of TCRs, pMHCs and accessory factors within the plasma membrane are likely to alter the TCR–pMHC binding dynamics and hence T-cell antigen sensitivity. The action of sugar-binding galectins and posttranslational modification of N-linked glycans, a self-remodeling cytoskeleton and co-expression of membrane-organizing elements such as tetraspanins could well influence membrane architecture and consequently antigen recognition. The extent to which such factors matter can only be determined through experimental work, in which we deem single molecule microscopy to be instrumental.

Acknowledgments

This work was supported by the Vienna Research Fund (WWTF) project LS13-030.

See also: Cellular Immunology: Cell–Cell Interactions and the Immune System: T Cell Receptor Triggering

References

- Adams, J.J., Narayanan, S., Liu, B., *et al.*, 2011. T cell receptor signaling is limited by docking geometry to peptide-major histocompatibility complex. *Immunity* 35, 681–693.
- Axelrod, D., 2008. Chapter 7: Total internal reflection fluorescence microscopy. *Methods in Cell Biology* 89, 169–221.
- Axmann, M., Huppa, J.B., Davis, M.M., Schütz, G.J., 2012. Determination of interaction kinetics between the T cell receptor and peptide-loaded MHC class II via single-molecule diffusion measurements. *Biophysical Journal* 103, L17–L19.
- Bashour, K.T., Gondarenko, A., Chen, H., *et al.*, 2014. CD28 and CD3 have complementary roles in T-cell traction forces. *Proceedings of the National Academy of Sciences of the United States of America* 111, 2241–2246.
- Beemiller, P., Jacobelli, J., Krummel, M.F., 2012. Integration of the movement of signaling microclusters with cellular motility in immunological synapses. *Nature Immunology* 13, 787–795.
- Beemiller, P., Krummel, M.F., 2010. Mediation of T-cell activation by actin meshworks. *Cold Spring Harbor Perspectives in Biology* 2, a002444.
- Brian, A.A., McConnell, H.M., 1984. Allogeneic stimulation of cytotoxic T cells by supported planar membranes. *Proceedings of the National Academy of Sciences of the United States of America* 81, 6159–6163.
- Bromley, S.K., Iaboni, A., Davis, S.J., *et al.*, 2001. The immunological synapse and CD28–CD80 interactions. *Nature Immunology* 2, 1159–1166.
- Bunnell, S.C., Singer, A.L., Hong, D.I., *et al.*, 2006. Persistence of cooperatively stabilized signaling clusters drives T-cell activation. *Molecular and Cellular Biology* 26, 7155–7166.
- Burkhardt, J.K., Carrizosa, E., Shaffer, M.H., 2008. The actin cytoskeleton in T cell activation. *Annual Review of Immunology* 26, 233–259.
- Burroughs, N.J., Wulfig, C., 2002. Differential segregation in a cell-cell contact interface: The dynamics of the immunological synapse. *Biophysical Journal* 83, 1784–1796.
- Campi, G., Varma, R., Dustin, M.L., 2005. Actin and agonist MHC-peptide complex-dependent T cell receptor microclusters as scaffolds for signaling. *Journal of Experimental Medicine* 202, 1031–1036.
- Choudhuri, K., Wiseman, D., Brown, M.H., Gould, K., Van Der Merwe, P.A., 2005. T-cell receptor triggering is critically dependent on the dimensions of its peptide–MHC ligand. *Nature* 436, 578–582.
- Dembo, M., Torney, D.C., Saxman, K., Hammer, D., 1988. The reaction-limited kinetics of membrane-to-surface adhesion and detachment. *Proceedings of the Royal Society of London B* 234, 55–83.

- Dustin, M.L., Bromley, S.K., Davis, M.M., Zhu, C., 2001. Identification of self through two-dimensional chemistry and synapses. *Annual Review of Cell and Developmental Biology* 17, 133–157.
- Evans, E.A., Calderwood, D.A., 2007. Forces and bond dynamics in cell adhesion. *Science* 316, 1148–1153.
- Friedl, P., Den Boer, A.T., Gunzer, M., 2005. Tuning immune responses: Diversity and adaptation of the immunological synapse. *Nature Reviews. Immunology* 5, 532–545.
- Garcia, K.C., Adams, J.J., Feng, D., Ely, L.K., 2009. The molecular basis of TCR germline bias for MHC is surprisingly simple. *Nature Immunology* 10, 143–147.
- Grakoui, A., Bromley, S.K., Sumen, C., *et al.*, 1999. The immunological synapse: A molecular machine controlling T cell activation. *Science* 285, 221–227.
- Ha, T., 2001. Single-molecule fluorescence resonance energy transfer. *Methods* 25, 78–86.
- Huang, J., Brameshuber, M., Zeng, X., *et al.*, 2013. A single peptide-major histocompatibility complex ligand triggers digital cytokine secretion in CD4+ T cells. *Immunity* 39, 1–12.
- Huang, J., Zarnitsyna, V.I., Liu, B., *et al.*, 2010. The kinetics of two-dimensional TCR and pMHC interactions determine T-cell responsiveness. *Nature* 464, 932–936.
- Huang, J.F., Yang, Y., Sepulveda, H., *et al.*, 1999. TCR-mediated internalization of peptide-MHC complexes acquired by T cells. *Science* 286, 952–954.
- Huppa, J.B., Axmann, M., Mortelmaier, M.A., *et al.*, 2010. TCR-peptide-MHC interactions in situ show accelerated kinetics and increased affinity. *Nature* 463, 963–967.
- Huppa, J.B., Davis, M.M., 2003. T-cell-antigen recognition and the immunological synapse. *Nature Reviews Immunology* 3, 973–983.
- Huse, M., Klein, L.O., Girvin, A.T., *et al.*, 2007. Spatial and temporal dynamics of T cell receptor signaling with a photoactivatable agonist. *Immunity* 27, 76–88.
- Irles, C., Symons, A., Michel, F., *et al.*, 2003. CD45 ectodomain controls interaction with GEMs and Lck activity for optimal TCR signaling. *Nature Immunology* 4, 189–197.
- James, J.R., Vale, R.D., 2012. Biophysical mechanism of T-cell receptor triggering in a reconstituted system. *Nature* 487, 64–69.
- Jiang, N., Huang, J., Edwards, L.J., *et al.*, 2011. Two-stage cooperative T cell receptor-peptide major histocompatibility complex-CD8 trimolecular interactions amplify antigen discrimination. *Immunity* 34, 13–23.
- Johnson, K.G., Bromley, S.K., Dustin, M.L., Thomas, M.L., 2000. A supramolecular basis for CD45 tyrosine phosphatase regulation in sustained T cell activation. *Proceedings of the National Academy of Sciences of the United States of America* 97, 10138–10143.
- Judokusumo, E., Tabdanov, E., Kumari, S., Dustin, M.L., Kam, L.C., 2012. Mechanosensing in T lymphocyte activation. *Biophysical Journal* 102, L5–L7.
- Kaizuka, Y., Douglass, A.D., Varma, R., Dustin, M.L., Vale, R.D., 2007. Mechanisms for segregating T cell receptor and adhesion molecules during immunological synapse formation in Jurkat T cells. *Proceedings of the National Academy of Sciences of the United States of America* 104, 20296–20301.
- Kim, S.T., Takeuchi, K., Sun, Z.Y., *et al.*, 2009. The alphabeta T cell receptor is an anisotropic mechanosensor. *Journal of Biological Chemistry* 284, 31028–31037.
- Klotzsch, E., Schütz, G.J., 2013. Improved ligand discrimination by force-induced unbinding of the T cell receptor from peptide-MHC. *Biophysical Journal* 104, 1670–1675.
- Kong, F., Garcia, A.J., Mould, A.P., Humphries, M.J., Zhu, C., 2009. Demonstration of catch bonds between an integrin and its ligand. *Journal of Cell Biology* 185, 1275–1284.
- Kuhns, M.S., Girvin, A.T., Klein, L.O., *et al.*, 2010. Evidence for a functional sidedness to the alphabetaTCR. *Proceedings of the National Academy of Sciences of the United States of America* 107, 5094–5099.
- Leupin, O., Zaru, R., Laroche, T., Muller, S., Valitutti, S., 2000. Exclusion of CD45 from the T-cell receptor signaling area in antigen-stimulated T lymphocytes. *Current Biology* 10, 277–280.
- Lillemeier, B.F., Mortelmaier, M.A., Forstner, M.B., *et al.*, 2010. TCR and Lat are expressed on separate protein islands on T cell membranes and concatenate during activation. *Nature Immunology* 11, 90–96.
- Lillemeier, B.F., Pfeiffer, J.R., Surviladze, Z., Wilson, B.S., Davis, M.M., 2006. Plasma membrane-associated proteins are clustered into islands attached to the cytoskeleton. *Proceedings of the National Academy of Sciences of the United States of America* 103, 18992–18997.
- Lim, T.S., Goh, J.K., Mortellaro, A., *et al.*, 2012. CD80 and CD86 differentially regulate mechanical interactions of T-cells with antigen-presenting dendritic cells and B-cells. *PLoS One* 7, e45185.
- Liu, B., Chen, W., Evavold, B.D., Zhu, C., 2014. Accumulation of dynamic catch bonds between TCR and agonist peptide-MHC triggers T cell signaling. *Cell* 157, 357–368.
- Marrack, P., Scott-Browne, J.P., Dai, S., Gapin, L., Kappler, J.W., 2008. Evolutionarily conserved amino acids that control TCR-MHC interaction. *Annual Review of Immunology* 26, 171–203.
- Monks, C.R., Freiberg, B.A., Kupfer, H., Sciaky, N., Kupfer, A., 1998. Three-dimensional segregation of supramolecular activation clusters in T cells. *Nature* 395, 82–86.
- Norcross, M.A., 1984. A synaptic basis for T-lymphocyte activation. *Annual Review of Immunology (Paris)* 135, 113–134.
- O'Donoghue, G.P., Pielak, R.M., Smoligovets, A.A., Lin, J.J., Groves, J.T., 2013. Direct single molecule measurement of TCR triggering by agonist pMHC in living primary T cells. *Elife* 2, e00778.
- Robert, P., Aleksic, M., Dushek, O., *et al.*, 2012. Kinetics and mechanics of two-dimensional interactions between T cell receptors and different activating ligands. *Biophysical Journal* 102, 248–257.
- Schamel, W.W., Arechaga, I., Risueno, R.M., *et al.*, 2005. Coexistence of multivalent and monovalent TCRs explains high sensitivity and wide range of response. *Journal of Experimental Medicine* 202, 493–503.
- Shaw, A.S., Dustin, M.L., 1997. Making the T cell receptor go the distance: A topological view of T cell activation. *Immunity* 6, 361–369.
- Simons, K., Gerl, M.J., 2010. Revitalizing membrane rafts: new tools and insights. *Nature Reviews Molecular Cell Biology* 11, 688–699.
- Thomas, W.E., Vogel, V., Sokurenko, E., 2008. Biophysics of catch bonds. *Annual Review of Biophysics and Biomolecular Structure* 37, 399–416.
- Varma, R., Campi, G., Yokosuka, T., Saito, T., Dustin, M.L., 2006. T cell receptor-proximal signals are sustained in peripheral microclusters and terminated in the central supramolecular activation cluster. *Immunity* 25, 117–127.
- Weikl, T.R., Lipowsky, R., 2004. Pattern formation during T-cell adhesion. *Biophysical Journal* 87, 3665–3678.
- Williamson, D.J., Owen, D.M., Rossy, J., *et al.*, 2011. Pre-existing clusters of the adaptor Lat do not participate in early T cell signaling events. *Nature Immunology* 12, 655–662.
- Wulfig, C., Davis, M.M., 1998. A receptor/cytoskeletal movement triggered by costimulation during T cell activation. *Science* 282, 2266–2269.
- Wulfig, C., Sumen, C., Sjaastad, M.D., *et al.*, 2002. Costimulation and endogenous MHC ligands contribute to T cell recognition. *Nature Immunology* 3, 42–47.
- Xie, J., Huppa, J.B., Newell, E.W., *et al.*, 2012. Photocrosslinkable pMHC monomers stain T cells specifically and cause ligand-bound TCRs to be 'preferentially' transported to the cSMAC. *Nature Immunology* 13, 674–680.
- Yokosuka, T., Sakata-Sogawa, K., Kobayashi, W., *et al.*, 2005. Newly generated T cell receptor microclusters initiate and sustain T cell activation by recruitment of Zap70 and SLP-76. *Nature Immunology* 6, 1253–1262.

Single-Cell Interrogation of the Immune System Using Microtools

C Albayrak and S Tay, ETH Zürich, Basel, Switzerland

© 2016 Elsevier Inc. All rights reserved.

The immune system comprises a highly complex and dynamic network of different cell types, and it negotiates responses to a vast array of challenges over different time scales (Brennan *et al.*, 1989; Egen *et al.*, 2008; Shulman *et al.*, 2013). As such, studying the immune system necessitates technologies with high sensitivity, resolution, and throughput commensurate with its complexity and dynamic nature (Junkin and Tay, 2014). Furthermore, biological systems including the immune system are inherently noisy. Measurements from bulk populations (which amount to ensemble averages) mask the underlying intercellular heterogeneity that arises either from stochasticity of processes such as binding events (intrinsic noise, e.g., RNA polymerase binding to its promoter) or from variations in levels of auxiliary factors that indirectly affect that process (extrinsic noise, e.g., differences in ribosome or chaperone concentrations) (Eldar and Elowitz, 2010; Elowitz *et al.*, 2002). This heterogeneity underlies important biological phenomena such as T cell response to vaccines (Flatz *et al.*, 2011) and nuclear factor kappa-light-chain-enhancer of activated B cells (NF- κ B) activation dynamics (Tay *et al.*, 2010), among others. Single-cell analysis will thus lead to a more complete and accurate understanding of immune cell dynamics, i.e., how individual cells in the immune system interact to create the observed phenotypes. In this brief article, we will summarize the relevant advances in microscale technology for studying immune cell activities, such as secretion, migration, and intercellular communication, at the single-cell level.

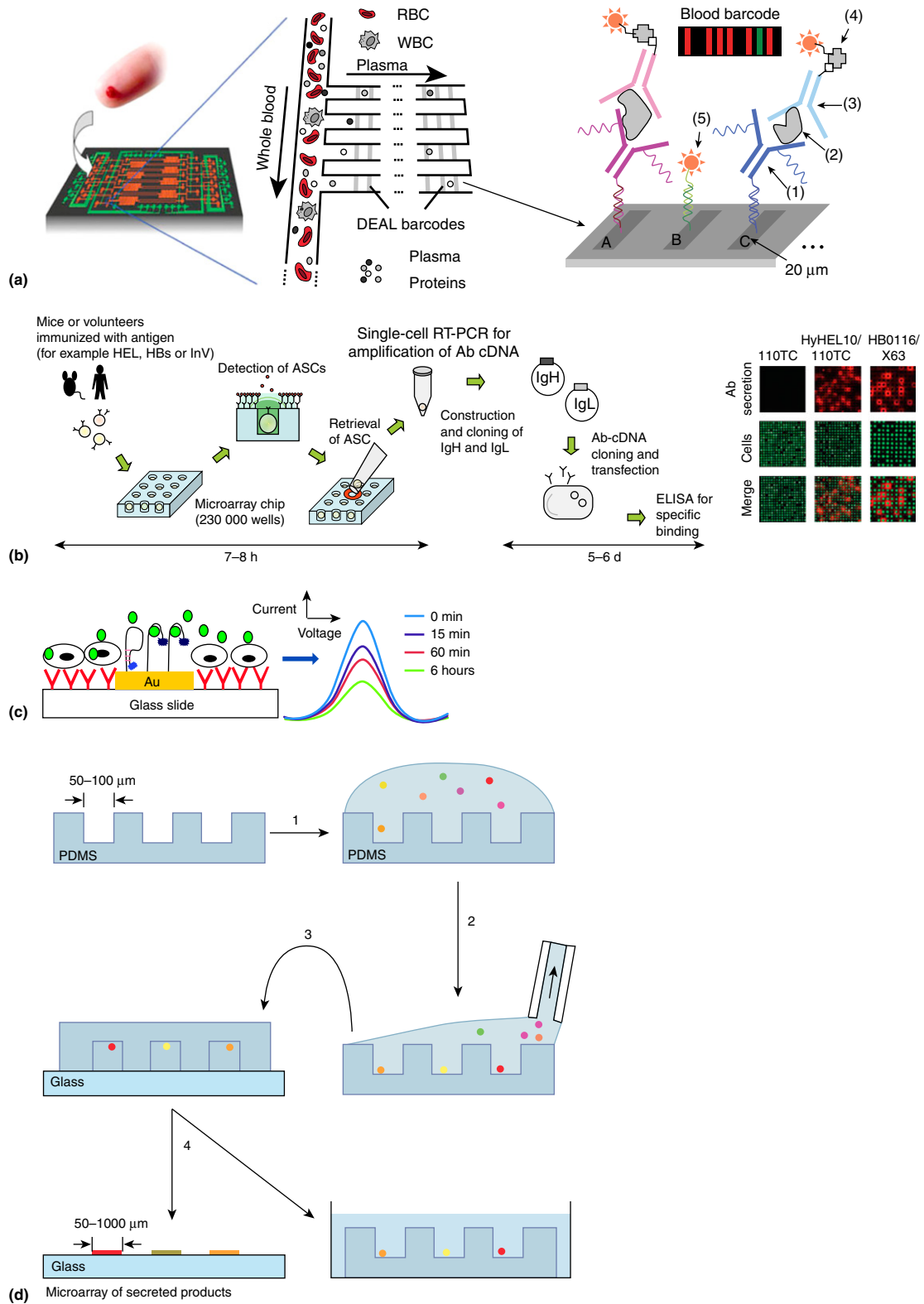
Single-Cell Secretion

Both integrated microfluidic devices (Araci and Quake, 2012; Thorsen *et al.*, 2002) and well arrays have been used for secretion analysis and selection of single cells. Well arrays are typically used in one of two configurations: (1) open configuration, where the secreted proteins are captured on the functionalized well surface, and (2) closed configuration, where the well surfaces are inert (with respect to protein-protein interactions), and the secreted proteins bind the functionalized glass lid. The latter approach, also called microengraving, enables repeated interrogation of the secreted proteins using multiple lids and separate analyses of the captured single cells and the lids.

The single-cell barcode chip (SCBC) developed by the Heath group is one example of an integrated microfluidic device for quantifying single-cell secretion. The device contains 1040 microchambers, each of which can hold one or several cells and contains immobilized primary antibodies that can bind up to 12 different antigens. Each primary antibody is covalently coupled to a different single-stranded DNA oligomer (Bailey *et al.*, 2007); these oligonucleotide-antibody conjugates are then flowed in microfluidic channels (a

different bioconjugate in each channel) over immobilized DNA oligomers with complementary sequences (Fan *et al.*, 2008). The primary antibodies are thus spatially patterned in stripes (hence the name, the 'barcode' device), and secreted proteins are quantified by measuring fluorescence intensity from the immune sandwich assays (Ma *et al.*, 2011; Figure 1(a)). The SCBC was used to compare the secretion profile of various single antigen-specific cytotoxic T lymphocytes (CTLs) sourced from healthy donors and a metastatic melanoma patient; up to 45 functional CTL subpopulations were identified by grouping cells that secreted the same proteins (Ma *et al.*, 2011). The SCBC was also used in conjunction with information theory-based entropy analysis to first construct a covariance matrix that reproduces fluctuations observed in secretions of unperturbed single macrophages, and then successfully predict perturbations in secreted protein levels due to neutralizing antibody treatment or the presence of multiple secretory cells (Shin *et al.*, 2011). Variations (e.g., the nanowire chip, pin-spotted arrays) and improved versions of the SCBC were later reported (i.e., higher throughput and degree of multiplexing) (Elitas *et al.*, 2014; Kwak *et al.*, 2014; Lu *et al.*, 2013); a most recent report features a device that can detect as many as 42 proteins secreted from single macrophages using antibodies distinguishable by both location and fluorophore color (Lu *et al.*, 2015).

Drop-based microfluidic devices have also been used for single-cell secretion analysis. Fluidic control and device manufacture are considerably simpler than integrated microfluidic devices, and single-cell isolation within drops enables higher assay throughput. Emulsification speed and assay (drop) volume can be easily modified by adjusting the flow rates of the aqueous and oil streams. This method is more suitable for batch assays, however, since controlled fluid addition to or perfusion of drops after emulsification is technically challenging. For example, a drop-based device enabled the monitoring of both surface marker CD86 (cluster of differentiation 86) production and interleukin 6 (IL-6) secretion from single dendritic cells for over 2 h using fluorescently tagged antibodies (Konry *et al.*, 2013). In an earlier study, IL-10 secretion from single T cells was analyzed by encapsulating them with beads displaying fluorescent capture antibodies (Konry *et al.*, 2011). An interesting variation was recently reported by Chokkalingam and colleagues, who captured the molecules secreted by individual Jurkat cells in agarose beads suitable for conventional flow cytometry. As in mentioned examples, single cells are encapsulated in drops with antibody-displaying beads and, crucially, ultra-low temperature gelling agarose. Upon cooling to 4 °C, the agarose gels and forms a bead that captures the secreted molecules along with the single cell; these beads are then washed, interrogated with fluorescent antibodies, and analyzed by flow cytometry (custom in-device cytometry is not needed since oil no longer comprises the continuous phase). Different groups of cells secreting IL-2,



interferon gamma (IFN γ), tumor necrosis factor alpha (TNF- α), and their combinations were identified (Chokkalingam *et al.*, 2013).

Open well arrays constitute an easily operable system and enable experimentation with and processing of hundreds of single cells. Capture reagents (i.e., typically antibodies specific to the secreted analytes) are immobilized at the bottom of the well array, and single cells are deposited in microwells. Once captured, single cells can be stimulated and the secreted molecules quantified by fluorescence (Zhu *et al.*, 2009). The immunospot array on a chip method pioneered by the Muraguchi group enables simultaneous detection of single antibody-secreting cells (ASCs) in $\sim 230\,000$ wells within several hours; antigen-specific ASCs can then be recovered, and the monoclonal antibodies obtained by subsequent cloning in a week (Jin *et al.*, 2011, 2009; Figure 1(b)).

Aptamers have also been used as capture reagents in microfluidic well arrays to detect secreted molecules. In this 'aptasensing' method, a redox reporter is covalently attached to the aptamer, and the aptamer specific to the secreted molecule is immobilized on miniature electrodes. Secreted cytokines lead to a conformational change in the aptamer upon binding, which leads to a change in the recorded current (Liu *et al.*, 2011; Figure 1(c)). Unlike antibody-based detection, this method does not need washing and rebinding of the analyte for multiple detections, thereby increasing time resolution of a kinetic secretion study. Liu *et al.* (2015, 2012) were able to monitor TNF- α and IFN- γ secretion from several hundred cells over 2–4 h with 20-minute time resolution. However, single-cell resolution is yet to be demonstrated; cytokine secretion from as low as 90 cells could be detected by aptasensing (Liu *et al.*, 2011).

Microengraving was originally demonstrated by Love and colleagues as a fast and high-throughput method for screening and isolation of ASCs. Individual cells could be viably captured in printed wells by limiting occupancy (i.e., using Poisson distribution), grown in the 0.1–1-nl wells, and their antibody secretions captured on glass lids functionalized either with the cognate antigen or secondary antibody (Bradshaw *et al.*, 2008; Love *et al.*, 2006; Story *et al.*, 2008; Figure 1(d)). Dense packing of wells increased throughput, while volume miniaturization drastically reduced the number of molecules required to beat the affinity threshold for efficient non-

covalent capture of secreted antibodies, and thus overall duration and cost of the screen. The supernatant can be sampled multiple times to monitor single-cell secretion over time (Han *et al.*, 2010). By measuring single-cell cytokine secretion every 2 h, researchers showed that stimulated T cells predominantly secrete cytokines one at a time as opposed to secreting multiple molecules simultaneously (Han *et al.*, 2012). Microcapillary arrays (i.e., the DiCAST platform (Fitzgerald *et al.*, 2015)) further increase screening capability by enabling interrogation of single-cell secretions using two functionalized lids simultaneously (rather than one as in microengraving). Many microcapillaries ($\sim 10^6$) with open top and bottom surfaces were densely packed in an array; proteins secreted by single cells in these capillaries queried using dual antibody-containing lids; and the desired cells subsequently recovered by automated analysis and pneumatic tools. Using this high-throughput method, antigen-specific hybridoma cells could be recovered without cloning by limiting dilution (Fitzgerald *et al.*, 2015).

Cell Migration

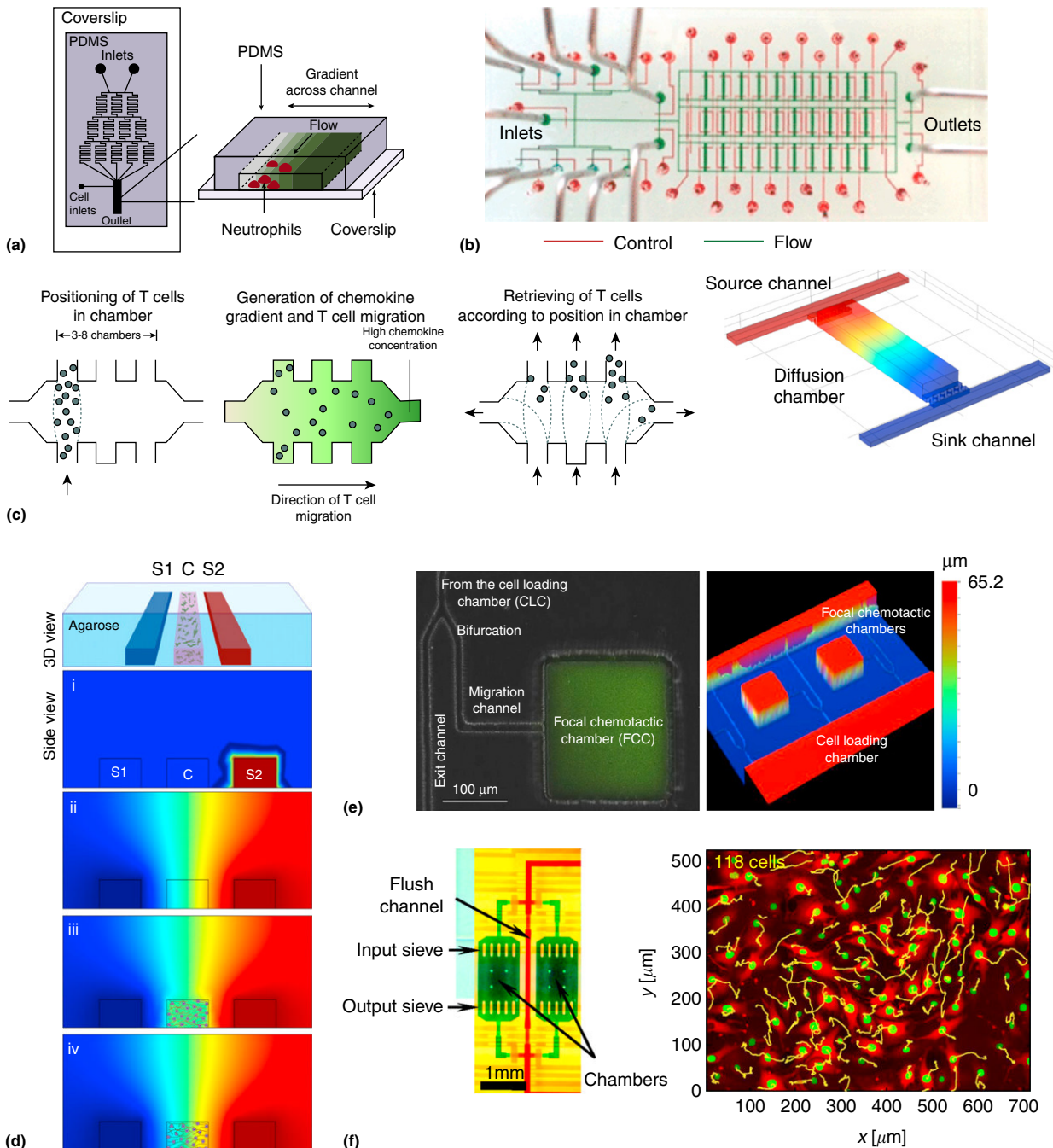
Microfluidic gradient devices coupled with time-lapse microscopy enables the study of single-cell migration under various chemokine gradients. Earlier designs depended on continuous flow to generate gradients, where solutions containing different chemokine concentrations were flowed side-by-side in a laminar regime. Cell migration is observed under flow in response to the transverse chemokine gradient. One of the earliest flow-dependent gradient devices was reported by Dertinger *et al.* (2001), and was used to study single neutrophil chemotaxis in different IL-8 gradients (Li Jeon *et al.*, 2002; Figure 2(a)). The device contains a branching network of serpentine channels that alternate with single wide channels where the gradients are formed perpendicular to the direction of laminar flow. The number of serpentine channels increases downstream, thereby enabling the formation of smooth linear gradients; the numbers of branching channel networks and liquid inlets can also be modified to create gradients with more complex shapes (Dertinger *et al.*, 2001). In this gradient device, the response of individual neutrophils to complex IL-8 gradients was traced. The neutrophils overshot the maximum

Figure 1 Microfluidic methods for measuring single-cell secretion. (a) Single-cell barcode chip (SCBC). Letters denote the different DNA-antibody conjugates, and the numbers indicate components required for the fluorescent sandwich assay. The result is a set of parallel fluorescent stripes, each specific to a secreted protein. (Reprinted with permission from Fan, R., Vermesh, O., Srivastava, A., *et al.*, 2008. Integrated barcode chips for rapid, multiplexed analysis of proteins in microliter quantities of blood. *Nature Biotechnology* 26, 1373–1378. © 2008 Nature Publishing Group.) (b) Immunospot array on a chip. Schematic (left) describes the detection and isolation of single antigen-secreting cells (ASCs), and the panels (right) show the orthogonal enumeration of single cells and secreted antibodies by fluorescence. (Adapted with permission from Jin, A., Ozawa, T., Tajiri, K., *et al.*, 2009. A rapid and efficient single-cell manipulation method for screening antigen-specific antibody-secreting cells from human peripheral blood. *Nature Medicine* 15, 1088–1092. © 2009 Nature Publishing Group.) (c) Aptamer-based electrochemical sensing ('aptasensing'). Secreted molecules (green circles) bind the aptamer, which triggers a conformational change in the aptamer and leads to a decrease in the detected current. (Reprinted with permission from Liu, Y., Yan, J., Howland, M.C., Kwa, T., Revzin, A., 2011. Micropatterned aptasensors for continuous monitoring of cytokine release from human leukocytes. *Analytical Chemistry* 83, 8286–8292. © 2011 American Chemical Society.) (d) Microengraving. After depositing cells onto the well array in bulk (1), single cells settle in wells and excess cells are removed (2). A glass lid containing immobilized affinity reagents is then put in contact with the well array (3), and the secreted molecules are captured on the lid (4) and quantified by fluorescence microscopy. (Reprinted with permission from Love, J.C., Ronan, J.L., Grotenbreg, G.M., van der Veen, A.G., Ploegh, H. L., 2006. A microengraving method for rapid selection of single cells producing antigen-specific antibodies. *Nature Biotechnology* 24, 703–707. © 2006 Nature Publishing Group.)

IL-8 concentration in a gentle (hill) gradient, but did not overshoot when the IL-8 concentration decreased abruptly (i.e., in a cliff gradient), thereby demonstrating their ability to respond to complex chemotactic cues (Li Jeon *et al.*, 2002). These cells were later shown to oscillate in their migration direction under flow when placed in opposing gradients of IL-8 and leukotriene B4 (LTB₄), both of which lie below formyl-methionyl-leucyl-phenylalanine (fMLP) in signaling hierarchy (Byrne *et al.*, 2014).

Flow-dependent gradient devices are inherently limited, however, since continuous flow is required to maintain the

chemokine gradient and may itself perturb cell migration and signaling (Walker *et al.*, 2005), thereby introducing additional variables unrelated to cell physiology. Flow-free devices were thus used by us and others to study immune cell migration under various scenarios; gradients in these devices are established by diffusion between solutions with and without the chemotactic agent (i.e., source and sink). For example, our group recently reported a modular gradient device that can generate up to 10 different programmable gradients in triplicate by alternating flow in the source and sink channels across a total of 30 chambers (Frank and Tay, 2013; Figure 2(b)).



Unlike in flow-dependent gradient devices, the absence of cross-flow in this device enables the observation of single or multiple cells in suspension and communication between individual cells via secreted factors. A different flow-free gradient device was used by us to study migration of single human T cells under a CXCL12 (chemokine (C-X-C motif) ligand 12) gradient. This device used a similar source-sink configuration, but enabled the retrieval of cells with different chemotactic characteristics (fast vs. slow, migrating vs. stationary; **Figure 2(c)**). Integrated analysis of the single T cells revealed that ones that migrate strongly in a CXCL12 gradient had elevated levels of CXCL4 (chemokine (C-X-C motif) ligand 4) expression (**Mehling et al., 2015**). **Keenan et al. (2006)** designed a flow-free gradient device (i.e., the Microjets device) to study neutrophil desensitization, the process by which neutrophils distinguish different chemotactic cues and respond to ones that take them to the infection or damage site effectively; fluorescent dyes were mixed in the chemokine solutions to simultaneously verify gradient profiles. Different subsets of neutrophils were identified based on their migration behavior (**Keenan et al., 2010**).

Three-dimensional (3D) gradient devices contain a gel matrix between the source and sink compartments and thus better mimic a tissue microenvironment by allowing the cells to migrate within the matrix in 3D. **Abhyankar et al. (2008)** designed a 3D gradient device, where two reservoirs are placed at opposite ends and on top of a 3D collagen gel. A steady gradient can be maintained by replenishing the solutions in the reservoirs, and chemotaxis of individual neutrophils in a temporally evolving fMLP gradient was quantified. Haessler and colleagues reported another device, which contains a biological extracellular matrix in a channel between the source and sink channels (**Figure 2(d)**). Dendritic cells were observed to preferentially migrate toward a CCL21 (chemokine (C-C motif) ligand 21) rather than a CCL19 (chemokine (C-C motif) ligand 19) gradient in the 3D environment, both of which bind CCR7 (chemokine (C-C motif) receptor 7) with

similar affinity (**Haessler et al., 2011**). This preference was also observed in T cells under flow (**Nandagopal et al., 2011**).

Microfluidic devices with small channels were also employed to reveal more complex migration behavior. For example, single T-lymphocytes migrated upstream on flow microchannels coated with ICAM-1 (intercellular adhesion molecule 1), but downstream on ones coated with vascular cell adhesion protein 1 (VCAM-1), and they preferred to migrate upstream when under increasing shear rate (**Dominguez et al., 2015**). **Sundt et al. (2012)** studied neutrophil rolling under high shear stress, observed the formation of slings (i.e., long cell protrusions) downstream in the direction of flow, and postulated that these slings originate from tethers behind the cell and appear in front as a result of neutrophil rolling. The slings also stabilize neutrophil rolling under high shear stress conditions such as in acute inflammation. In a separate study, forward and reverse migration (retroaxis) of individual neutrophils were quantified using a microfluidic device with U-shaped channels under various chemokine gradients; cells displayed three different retroaxis phenotypes (**Hamza et al., 2014**). **Boneschansker et al. (2014)** used a flow-free microfluidic device with cell traps to quantitatively study neutrophil and lymphocyte migration. Once trapped, individual cells could either migrate toward or away from the chemokine in narrow microfluidic channels; the direction, speed, persistence, and percentage of motile cells were quantified from time-lapsed images, thereby yielding a more comprehensive picture of cell migration in chemokine gradients. Jones and colleagues mimicked a focal inflammation site using a microfluidic device, where focal chemotactic chambers (FCC) that simulate foci of inflammation were connected by narrow channels to a central cell loading chamber (CLC). The CLC also acted as a sink for the chemokines diffusing from the FCC. Cells that chemotactically migrated toward the FCC could be distinguished from those exhibiting chemokinesis by the bifurcation in the small channel; the former would only follow the migration channel, while the latter would follow both

Figure 2 Microtools for interrogating immune cell migration. (a) Flow-dependent gradient chip. Serpentine channels are used to develop a chemokine gradient transverse to flow direction. After gradient generation, cells are introduced to and migrate in the rectangular observation channel under continuous flow. (Adapted with permission from Li Jeon, N., Baskaran, H., Dertinger, S.K.W., et al., 2002. Neutrophil chemotaxis in linear and complex gradients of interleukin-8 formed in a microfabricated device. *Nature Biotechnology* 20, 826–830. © 2002 Nature Publishing Group.) (b) Integrated flow-free gradient chip. Upper panel shows the flow and control layers of the device that contains 30 gradient chambers and can generate 10 independent gradients, while lower panel shows the simulation of a diffusion-based gradient. (Adapted from Frank, T., Tay, S., 2013. Flow-switching allows independently programmable, extremely stable, high-throughput diffusion-based gradients. *Lab on a Chip* 13, 1273–1281 by permission of The Royal Society of Chemistry.) (c) Schematic of an individual migration chamber. T cells can be separately collected according to their migration behavior. (Reproduced from Mehling, M., Frank, T., Albayrak, C., Tay, S., 2015. Real-time tracking, retrieval and gene expression analysis of migrating human T cells. *Lab on a Chip* 15, 1276–1283 by permission of The Royal Society of Chemistry.) (d) Top panel shows one functional unit of the 3D gradient device. Cells are placed within the 3D matrix in the center channel (C), which is flanked by the chemokine source/sink channels (S1 and S2). (i–iv) Simulation of a CCL19 gradient in the 3D gradient device. (Reprinted from Haessler, U., Pisano, M., Wu, M., Swartz, M.A., 2011. Dendritic cell chemotaxis in 3D under defined chemokine gradients reveals differential response to ligands CCL21 and CCL19. *Proceedings of the National Academy of Sciences of the United States of America* 108, 5614–5619. © 2011 National Academy of Sciences, USA.) (e) Chemokine is placed in the focal chemotactic chamber (FCC), and a gradient is formed along the migration channel (left panel). Profilometer image of two FCCs and migration channels connected to the central cell loading chamber (CLC, right panel). (Adapted from Jones, C.N., Dalli, J., Dimisko, L., et al., 2012. Microfluidic chambers for monitoring leukocyte trafficking and humanized nano-proresolving medicines interactions. *Proceedings of the National Academy of Sciences of the United States of America* 109, 20560–20565. © 2012 National Academy of Sciences, USA.) (f) Two chambers of the automated cell culture device (left panel). Migration trajectories (yellow) of individual fibroblasts in a cell culture chamber (right panel). (Adapted with permission from Gomez-Sjöberg, R., Leyrat, A.A., Pirone, D.M., Chen, C.S., Quake, S.R., 2007. Versatile, Fully automated, microfluidic cell culture system. *Analytical Chemistry* 79, 8557–8563 and Vedel, S., Tay, S., Johnston, D.M., Bruus, H., Quake, S.R., 2013. Migration of cells in a social context. *Proceedings of the National Academy of Sciences of the United States of America* 110, 129–134. © 2007 American Chemical Society. © 2013 National Academy of Sciences, USA.)

migration and exit channels (Figure 2(e)). A fluorescent elastase assay was implemented in-chip to distinguish inflammation-induced recruitment of monocytes, in which elastase is up-regulated. This microfluidic setup enabled the researchers to obtain a more comprehensive understanding of the inflammation scenario in the presence of different leukocyte subpopulations, chemoattractants and other inflammation-mediating agents (Jones *et al.*, 2012). In addition to these devices, the automated cell culture chip (Gomez-Sjöberg *et al.*, 2007) was used by us to quantitatively analyze fibroblast migration in a population context, but on a single-cell level (Figure 2(f)). Cell motility was observed to be influenced by both internal factors such as directional bias of pseudopod formation and social factors such as intercellular collisions and secreted chemokines (Vedel *et al.*, 2013).

Finally, several reports were focused on simplifying the implementation of these microtools for cell migration studies, thereby making them accessible to a broader group of researchers. For example, Sackmann *et al.* (2012) developed the kit-on-a-lid-assay (KOALA) that can be conducted with only a micropipette and an optical microscope; migration of individual primary neutrophils was studied using a few microliters of blood. Wu *et al.* (2014) reported the USB microscope-based microfluidic chemotaxis analysis system (UMCAS) that automates cell tracking analysis and provides quantitative cell migration data. Sip *et al.* (2014) designed a simple microfluidic transwell insert for standard 6-well plates that can generate a stable gradient over a large portion of the well surface area. Chemotaxis of individual HL-60 cells due to an fMLP gradient in a conventional well plate could be quantified using this microfluidic insert and time-lapse microscopy. Examples such as these will hopefully facilitate the adoption of these powerful microtools by researchers not specialized in microfluidics.

Cell-Cell Communication

In this context, microfluidic devices have been used to analyze interactions (via direct contact or paracrine signaling) between different cell types on a single-cell level or to capture the physiological complexity within a more simplified experimental setting by mimicking *in vivo* scenarios (e.g., tumor microenvironments, blood vessels).

We have recently designed an integrated coculture device that ensures precise control of intercellular communication via paracrine molecules (Figure 3(a)). First and second responder cells (macrophages and fibroblasts, in this case) were cultured in adjacent chambers; each of these chambers could be independently fed and washed; and intercellular communication via paracrine signaling was permitted by opening the valve that divides the adjacent culture chambers. Single or multiple macrophages were stimulated by bacterial lipopolysaccharide (LPS); single or multiple fibroblasts were exposed to the TNF- α secreted by macrophages upon LPS activation; and nuclear localization of NF- κ B in individual fibroblasts was measured by fluorescence microscopy (Frank and Tay, 2015). Another recent study by Shalek and colleagues used Fluidigm's C₁ single-cell capture device with single-cell RNA sequencing to study the effect of paracrine signaling on the response of primary dendritic cells to different pathogenic components. The C₁ device

allowed the researchers to stimulate individual dendritic cells in isolation and compare these to cells collectively stimulated in a test tube, thereby decoupling the effects of paracrine agents from those of other factors (Figure 3(b)). The researchers discovered a small subset of 'precocious' dendritic cells that first express antiviral genes and activate these genes in other cells by IFN- β signaling, thereby mounting a population response. This is followed by a second wave of paracrine signaling independent of the initial IFN- β that downregulates the inflammatory genes (Shalek *et al.*, 2014). Elitas and colleagues compared the profile of secreted proteins from cocultures of glioma cells and macrophages to those from single glioma or macrophages; 16 secreted proteins were quantified using a barcode chip (Lu *et al.*, 2013). Statistically significant differences between the coculture secretion profile and the synthetic addition of secreted protein profiles from individual cell types indicated paracrine effect (Elitas *et al.*, 2014). The aptasensing platform was also used to analyze paracrine communication between different groups (active and quiescent) of monocyte-like cells via cytokines. About 600 cells were captured in microwells, and the researchers were able to quantify the TNF- α secreted by different cell groups and calculate secretion rates (Kwa *et al.*, 2014).

Kim and Haynes used a flow-dependent device to investigate single neutrophil chemotaxis and migration on endothelial cells, which mimicks neutrophil movement in blood vessels prior to extravasation. The presence of endothelial cells and molecules expressed by these cells (such as surface adhesion molecules CD11b (cluster of differentiation 11b, also known as integrin alpha M (ITGAM)) and CD66b (cluster of differentiation 66b, also known as carcinoembryonic antigen-related cell adhesion molecule 8 (CEACAM8))) drastically altered the migratory behavior of neutrophils, demonstrating the importance of cell-cell interactions in neutrophil migration (Kim and Haynes, 2013). In the same vein, Mahmood and coworkers analyzed migration of single natural killer (NK) cells in dendritic cell (DC)-conditioned medium, which mimics the presence of DCs and NK cells in the same microenvironment. The DC-conditioned medium (containing factors secreted by DCs) increased NK cell chemotaxis, and this chemotactic activation was partially dependent on CXCR3 (Chemokine (C-X-C motif) receptor 3) signaling (Mahmood *et al.*, 2014). Immunological synapse formation between mature DC and T cells were analyzed on a single-cell level using a drop-based device. The two cells were encapsulated together in drops, and cytoskeleton remodeling and tubulin polymerization could be observed upon contact between mature DC and T cells (Konry *et al.*, 2013).

Lytic interactions between two different cell types were also investigated using microfluidics. Varadarajan and colleagues used microengraving to analyze both secretion and cytolytic activity of CD8⁺ T cells from HIV-infected patients on a single-cell level. Their high-throughput analysis showed that when T cells encounter viral antigen presenting cells (APCs), a majority of them either secrete IFN- α or lyse their targets, but not both (Varadarajan *et al.*, 2011). Other microfluidic studies on NK cells revealed more single-cell heterogeneity in their interactions with target cells. For example, a small subset of IL-2-activated NK cells in the studied population killed the majority of target kidney cells; these 'serial killer' cells were larger than others and induced fast necrotic cell death

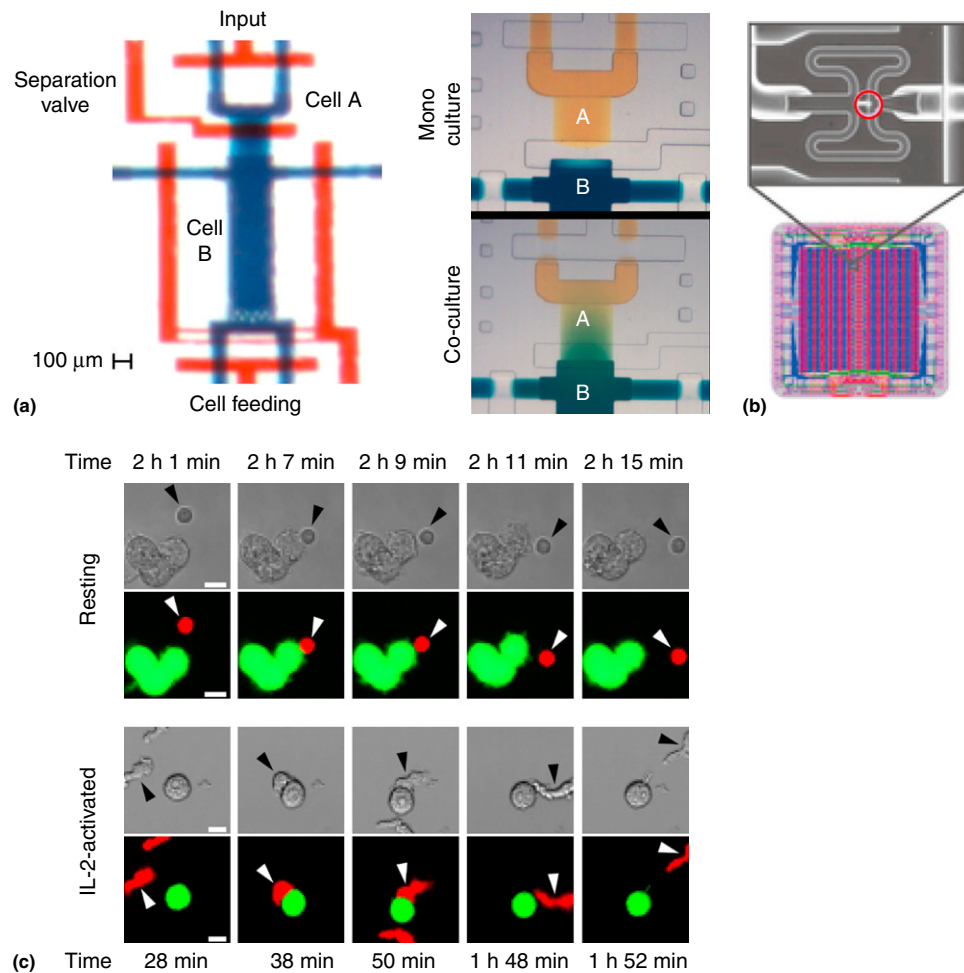


Figure 3 Intercellular communication at the single-cell level. (a) One functional unit of our new coculture device. Each cell (A and B) can be fed and washed independently, and the separation valve can be opened to trigger intercellular communication. (Adapted from Frank, T., Tay, S., 2015. Automated coculture system for spatiotemporal analysis of cell-to-cell communication. *Lab on a Chip* 15, 2192–2200 by permission of The Royal Society of Chemistry.) (b) Isolation of a single dendritic cell (red circle, upper panel) using Fluidigm's C₁ device (lower panel). (Reproduced with permission from Shalek, A.K., Satija, R., Shuga, J., *et al.*, 2014. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* 509, 363–369. © 2014 Nature Publishing Group.) (c) Time-lapse light (top panels) and fluorescent (bottom panels) microscopy of resting or activated NK cells (red) interacting with target cells (green). (Reproduced with permission from Olofsson, P.E., Forslund, E., Vanherberghen, B., *et al.*, 2014. Distinct migration and contact dynamics of resting and IL-2-activated human natural killer cells. *Frontiers in Immunology* 5, 80. © 2014 Olofsson, Forslund, Vanherberghen, Chechet, Mickelin, Ahlin, Everhorn and Önfelt.)

(Vanherberghen *et al.*, 2013). They also seem to spread over their targets upon contact (Olofsson *et al.*, 2014). In contrast, resting NK cells were less mobile and less lethal, formed fewer contacts with target kidney cells, and maintained their spherical shape upon contact (Olofsson *et al.*, 2014; **Figure 3(c)**). An earlier study utilizing microwell arrays also revealed that individual IL-2-activated NK cells secreted MIP-1 β and TNF- α upon contacting target cells (Yamanaka *et al.*, 2012).

Other Applications

Medical Applications: Microtools as Part of an Integrated Workflow

The mentioned examples utilized microtools to reveal different aspects of immune cell physiology on a single-cell level.

However, microfluidic devices can also comprise a powerful unit in an integrated workflow; they can either be used to screen and identify desired cells in a high-throughput fashion or to analyze them on a single-cell level after selection. Unlike methods such as intracellular cytokine staining (ICS) (Jung *et al.*, 1993), single cells can be evaluated in these devices and desired ones later recovered for further characterization (Varadarajan *et al.*, 2012). For example, Ogunniyi and colleagues combined microengraving with on-device cytometry and single-cell RT-PCR to profile immune cells from HIV patients. Plasma cells and memory B cells were scored in parallel for viability, marker phenotype, and secreted antibody isotype and specificity; some cells were later recovered for RT-PCR and sequencing of the antibody-encoding genes (Ogunniyi *et al.*, 2014). In another study, researchers combined microengraving with whole-exome sequencing to determine somatic

single-nucleotide variants (SSNVs) among circulating tumor cells (CTCs) and compare the mutations in CTCs to those from different tumors of a single patient with metastatic prostate cancer (Lohr *et al.*, 2014).

Microfluidic tools were also used in one recent study, where isolating and singly analyzing rare cell populations revealed new disease mechanisms and therapeutic target molecules. Zhou and colleagues identified immunotherapy targets by administering a small hairpin RNA (shRNA) library and sequencing ones that enable T cell proliferation *in vivo* in the tumor microenvironment. After discovering *Ppp2r2d* as a potential target, the researchers used a well array to analyze at a single-cell level the cytokines secreted by *Ppp2r2d*-silenced T cells harvested from a melanoma tumor (Zhou *et al.*, 2014).

Dissecting Biological Noise

One of the greatest contributions of single-cell analysis has been in the discovery and analysis of biological noise (Elowitz *et al.*, 2002; Figure 4(a)). Traditional bulk measurements

inherently overlooked this fundamental aspect of biology, since they represent averages of isogenic yet fluctuating cells. Microfluidics and fluorescence microscopy have been used by us and others to facilitate high-throughput interrogation of signaling pathways at the single-cell level.

Our group has learned much from single-cell analysis of the NF- κ B signaling pathway, dysfunction of which plays a role in inflammation, immune diseases, and cancer (Courtois and Gilmore, 2006). In a new study, Kellogg and Tay used an integrated valve-based microfluidic device (Gomez-Sjöberg *et al.*, 2007) to study NF- κ B expression in hundreds of individual mammalian cells for >20 h under various stimulation scenarios (sustained vs. oscillating input) and understand how noise and input oscillations affect signal transduction (Kellogg and Tay, 2015; Kellogg *et al.*, 2014; Figure 4(b)). Nuclear NF- κ B oscillations were entrained (i.e., phase-locked with input oscillations) or disrupted depending on the oscillating input period, which affected expression of NF- κ B target genes; entrained NF- κ B oscillations led to higher levels of target expression. Both single-cell kinetic data and simulations based on a hybrid NF- κ B model

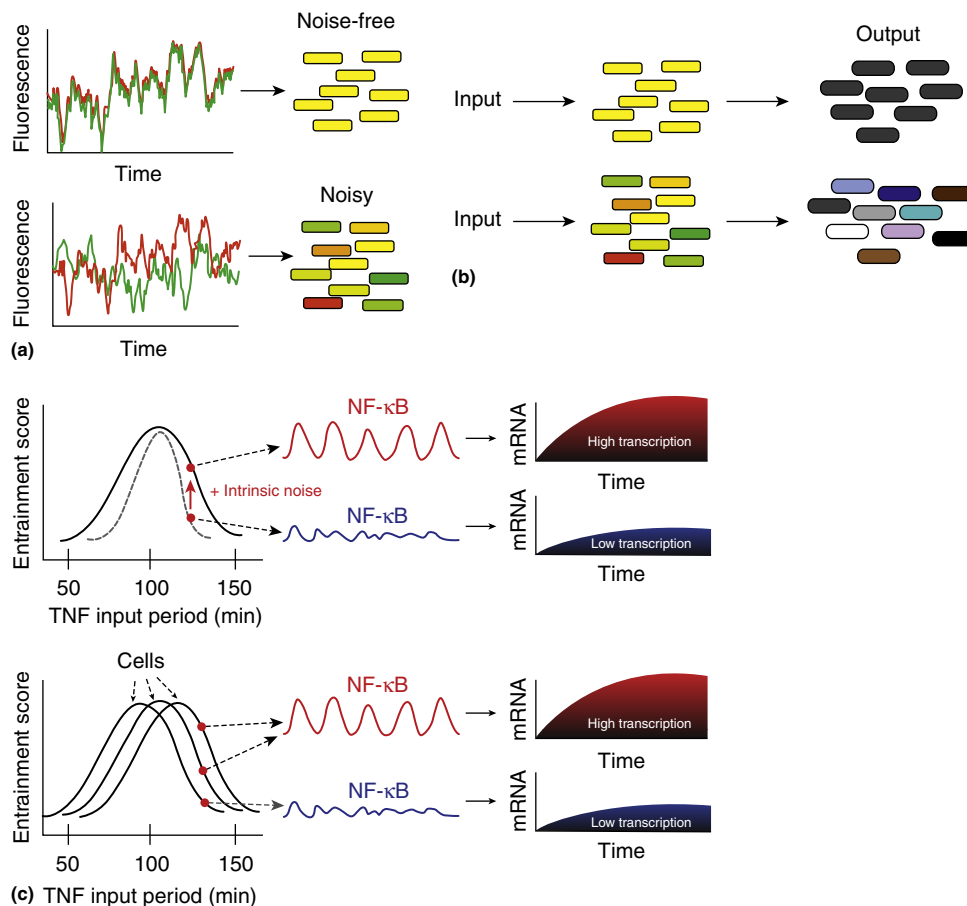


Figure 4 Biological noise and its consequences. (a) Intrinsic noise leads to differential expression of red and green fluorescent proteins, resulting in a heterogeneous population of isogenic bacteria. (b) Members of a noisy cell population will also process a given signal (input) differently and exhibit a wide range of responses (output). (Reprinted with permission from Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S., 2002. Stochastic gene expression in a single cell. *Science* 297, 1183–1186. © 2002 American Association for the Advancement of Science.) (c) Roles of intrinsic and extrinsic noise in transcriptional control and signaling dynamics. Intrinsic noise increases NF- κ B oscillation and entrainment (which, in turn, augments targets gene expression), while extrinsic noise broadens the oscillation frequency that can entrain cells, thereby increasing the chance of entrainment within the cell population. (Reprinted with permission from Kellogg, R.A., Tay, S., 2015. Noise facilitates transcriptional control under dynamic inputs. *Cell* 160, 381–392. © 2015 Elsevier.)

(Tay *et al.*, 2010) showed that intrinsic noise broadens the range of input periods that lead to and sustain NF- κ B oscillations required for entrainment, while extrinsic noise increases the chance that a subset of cells would entrain to a given dynamic input (analogous to bacterial bet-hedging (Veening *et al.*, 2008)). As a result, the combination of intrinsic and extrinsic noise enables output (i.e., nuclear NF- κ B oscillation) entrainment, which, in turn, leads to strong transcription of target genes in response to a range of dynamic inputs (Figure 4(c)). In the earlier study describing the hybrid model, the same high-throughput microfluidic device was used to reveal the digital nature of NF- κ B activation upon constant (non-oscillating) inputs at the single-cell level, with fewer cells digitally activated at lower input dose. Despite the digital nature of the response, some features such as maximum nuclear NF- κ B amount, response time to input and number of NF- κ B oscillations depended on input dose and thus were analogue (Tay *et al.*, 2010).

Acknowledgments

This work was partially supported by Schweizerischer Nationalfonds, a SystemsX Grant (NeuroStemX), and an ERC Starting Grant (SingleCellDynamics) to S.T. This publication is also supported by the Swiss National Science Foundation as part of the NCCR Molecular Systems Engineering.

References

- Abhyankar, V.V., Toepke, M.W., Cortesio, C.L., *et al.*, 2008. A platform for assessing chemotactic migration within a spatiotemporally defined 3D microenvironment. *Lab on a Chip* 8, 1507–1515.
- Araci, I.E., Quake, S.R., 2012. Microfluidic very large scale integration (mVLSI) with integrated micromechanical valves. *Lab on a Chip* 12, 2803–2806.
- Bailey, R.C., Kwong, G.A., Radu, C.G., Witte, O.N., Heath, J.R., 2007. DNA-encoded antibody libraries: A unified platform for multiplexed cell sorting and detection of genes and proteins. *Journal of the American Chemical Society* 129, 1959–1967.
- Boneschansker, L., Yan, J., Wong, E., Briscoe, D.M., Irimia, D., 2014. Microfluidic platform for the quantitative analysis of leukocyte migration signatures. *Nature Communications* 5, 1–12.
- Bradshaw, E.M., Kent, S.C., Tripuraneni, V., *et al.*, 2008. Concurrent detection of secreted products from human lymphocytes by microengraving: Cytokines and antigen-reactive antibodies. *Clinical Immunology* 129, 10–18.
- Brennan, F.M., Chantry, D., Jackson, A., Maini, R., Feldmann, M., 1989. Inhibitory effect of TNF alpha antibodies on synovial cell interleukin-1 production in rheumatoid arthritis. *Lancet* 2, 244–247.
- Byrne, M.B., Kimura, Y., Kapoor, A., *et al.*, 2014. Oscillatory behavior of neutrophils under opposing chemoattractant gradients supports a winner-take-all mechanism. *PLoS ONE* 9, e85726.
- Chokkalingam, V., Tel, J., Wimmers, F., *et al.*, 2013. Probing cellular heterogeneity in cytokine-secreting immune cells using droplet-based microfluidics. *Lab on a Chip* 13, 4740–4744.
- Courtis, G., Gilmore, T.D., 2006. Mutations in the NF- κ B signaling pathway: Implications for human disease. *Oncogene* 25, 6831–6843.
- Dertinger, S.K.W., Chiu, D.T., Jeon, N.L., Whitesides, G.M., 2001. Generation of gradients having complex shapes using microfluidic networks. *Analytical Chemistry* 73, 1240–1246.
- Dominguez, G.A., Anderson, N.R., Hammer, D.A., 2015. The direction of migration of T-lymphocytes under flow depends upon which adhesion receptors are engaged. *Integrative Biology* 7, 345–355.
- Egen, J.G., Rothfuchs, A.G., Feng, C.G., *et al.*, 2008. Macrophage and T Cell dynamics during the development and disintegration of mycobacterial granulomas. *Immunity* 28, 271–284.
- Eldar, A., Elowitz, M.B., 2010. Functional roles for noise in genetic circuits. *Nature* 467, 167–173.
- Elitas, M., Brower, K., Lu, Y., Chen, J.J., Fan, R., 2014. A microchip platform for interrogating tumor-macrophage paracrine signaling at the single-cell level. *Lab on a Chip* 14, 3582–3588.
- Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S., 2002. Stochastic gene expression in a single cell. *Science* 297, 1183–1186.
- Fan, R., Vermesh, O., Srivastava, A., *et al.*, 2008. Integrated barcode chips for rapid, multiplexed analysis of proteins in microliter quantities of blood. *Nature Biotechnology* 26, 1373–1378.
- Fitzgerald, V., Manning, B., O'Donnell, B., *et al.*, 2015. Exploiting highly ordered subnanoliter volume microcapillaries as microtools for the analysis of antibody producing cells. *Analytical Chemistry* 87, 997–1003.
- Flatz, L., Roychoudhuri, R., Honda, M., *et al.*, 2011. Single-cell gene-expression profiling reveals qualitatively distinct CD8 T cells elicited by different gene-based vaccines. *Proceedings of the National Academy of Sciences of the United States of America* 108, 5724–5729.
- Frank, T., Tay, S., 2013. Flow-switching allows independently programmable, extremely stable, high-throughput diffusion-based gradients. *Lab on a Chip* 13, 1273–1281.
- Frank, T., Tay, S., 2015. Automated co-culture system for spatiotemporal analysis of cell-to-cell communication. *Lab on a Chip* 15, 2192–2200.
- Gomez-Sjöberg, R., Leyrat, A.A., Pirone, D.M., Chen, C.S., Quake, S.R., 2007. Versatile, Fully automated, microfluidic cell culture system. *Analytical Chemistry* 79, 8557–8563.
- Haessler, U., Pisano, M., Wu, M., Swartz, M.A., 2011. Dendritic cell chemotaxis in 3D under defined chemokine gradients reveals differential response to ligands CCL21 and CCL19. *Proceedings of the National Academy of Sciences of the United States of America* 108, 5614–5619.
- Hamza, B., Wong, E., Patel, S., *et al.*, 2014. Retrotaxis of human neutrophils during mechanical confinement inside microfluidic channels. *Integrative Biology* 6, 175–183.
- Han, Q., Bagheri, N., Bradshaw, E.M., *et al.*, 2012. Polyfunctional responses by human T cells result from sequential release of cytokines. *Proceedings of the National Academy of Sciences of the United States of America* 109, 1607–1612.
- Han, Q., Bradshaw, E.M., Nilsson, B., Hafner, D.A., Love, J.C., 2010. Multidimensional analysis of the frequencies and rates of cytokine secretion from single cells by quantitative microengraving. *Lab on a Chip* 10, 1391–1400.
- Jin, A., Ozawa, T., Tajiri, K., *et al.*, 2009. A rapid and efficient single-cell manipulation method for screening antigen-specific antibody-secreting cells from human peripheral blood. *Nature Medicine* 15, 1088–1092.
- Jin, A., Ozawa, T., Tajiri, K., *et al.*, 2011. Rapid isolation of antigen-specific antibody-secreting cells using a chip-based immunospot array. *Nature Protocols* 6, 668–676.
- Jones, C.N., Dalli, J., Dimisko, L., *et al.*, 2012. Microfluidic chambers for monitoring leukocyte trafficking and humanized nano-proresolving medicines interactions. *Proceedings of the National Academy of Sciences of the United States of America* 109, 20560–20565.
- Jung, T., Schauer, U., Heusser, C., Neumann, C., Rieger, C., 1993. Detection of intracellular proteins by flow cytometry. *Journal of Immunological Methods* 159, 197–207.
- Junkin, M., Tay, S., 2014. Microfluidic single-cell analysis for systems immunology. *Lab on a Chip* 14, 1246–1260.
- Keenan, T.M., Frevert, C.W., Wu, A., Wong, V., Folch, A., 2010. A new method for studying gradient-induced neutrophil desensitization based on an open microfluidic chamber. *Lab on a Chip* 10, 116–122.
- Keenan, T.M., Hsu, C.H., Folch, A., 2006. Microfluidic “jets” for generating steady-state gradients of soluble molecules on open surfaces. *Applied Physics Letters* 89, 114103.
- Kellogg, R.A., Gómez-Sjöberg, R., Leyrat, A.A., Tay, S., 2014. High-throughput microfluidic single-cell analysis pipeline for studies of signaling dynamics. *Nature Protocols* 9, 1713–1726.
- Kellogg, R.A., Tay, S., 2015. Noise facilitates transcriptional control under dynamic inputs. *Cell* 160, 381–392.
- Kim, D., Haynes, C.L., 2013. On-chip evaluation of neutrophil activation and neutrophil – Endothelial cell interaction during neutrophil chemotaxis. *Analytical Chemistry* 85, 10787–10796.
- Konry, T., Dominguez-Villar, M., Baecher-Allan, C., Hafner, D.A., Yarmush, M.L., 2011. Droplet-based microfluidic platforms for single T cell secretion analysis of IL-10 cytokine. *Biosensors & Bioelectronics* 26, 2707–2710.
- Konry, T., Golberg, A., Yarmush, M., 2013. Live single cell functional phenotyping in droplet nano-liter reactors. *Scientific Reports* 3, 3179.
- Kwa, T., Zhou, Q., Gao, Y., *et al.*, 2014. Reconfigurable microfluidics with integrated aptasensors for monitoring intercellular communication. *Lab on a Chip* 14, 1695–1704.

- Kwak, M., Kim, D.-J., Lee, M.-R., *et al.*, 2014. Nanowire array chips for molecular typing of rare trafficking leukocytes with application to neurodegenerative pathology. *Nanoscale* 6, 6537–6550.
- Li Jeon, N., Baskaran, H., Dertinger, S.K.W., *et al.*, 2002. Neutrophil chemotaxis in linear and complex gradients of interleukin-8 formed in a microfabricated device. *Nature Biotechnology* 20, 826–830.
- Liu, Y., Kwa, T., Revzin, A., 2012. Simultaneous detection of cell-secreted TNF- α and IFN- γ using micropatterned aptamer-modified electrodes. *Biomaterials* 33, 7347–7355.
- Liu, Y., Liu, Y., Matharu, Z., Rahimian, A., Revzin, A., 2015. Detecting multiple cell-secreted cytokines from the same aptamer-functionalized electrode. *Biosensors & Bioelectronics* 64, 43–50.
- Liu, Y., Yan, J., Howland, M.C., Kwa, T., Revzin, A., 2011. Micropatterned aptasensors for continuous monitoring of cytokine release from human leukocytes. *Analytical Chemistry* 83, 8286–8292.
- Lohr, J.G., Adalsteinsson, V.A., Cibulskis, K., *et al.*, 2014. Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. *Nature Biotechnology* 32, 479–484.
- Love, J.C., Ronan, J.L., Grotenbreg, G.M., van der Veen, A.G., Ploegh, H.L., 2006. A microengraving method for rapid selection of single cells producing antigen-specific antibodies. *Nature Biotechnology* 24, 703–707.
- Lu, Y., Chen, J.J., Mu, L., *et al.*, 2013. High-throughput secretomic analysis of single cells to assess functional cellular heterogeneity. *Analytical Chemistry* 85, 2548–2556.
- Lu, Y., Xue, Q., Eisele, M.R., *et al.*, 2015. Highly multiplexed profiling of single-cell effector functions reveals deep functional heterogeneity in response to pathogenic ligands. *Proceedings of the National Academy of Sciences of the United States of America* 112 (7), E607–E615.
- Ma, C., Fan, R., Ahmad, H., *et al.*, 2011. A clinical microchip for evaluation of single immune cells reveals high functional heterogeneity in phenotypically similar T cells. *Nature Medicine* 17, 738–743.
- Mahmood, S., Nandagopal, S., Sow, I., Lin, F., Kung, S.K.P., 2014. Microfluidic-based, live-cell analysis allows assessment of NK-cell migration in response to cross-talk with dendritic cells. *European Journal of Immunology* 44, 2737–2748.
- Mehling, M., Frank, T., Albayrak, C., Tay, S., 2015. Real-time tracking, retrieval and gene expression analysis of migrating human T cells. *Lab on a Chip* 15, 1276–1283.
- Nandagopal, S., Wu, D., Lin, F., 2011. Combinatorial guidance by CCR7 ligands for T lymphocytes migration in co-existing chemokine fields. *PLoS ONE* 6, e18183.
- Ogunniyi, A.O., Thomas, B.A., Politano, T.J., *et al.*, 2014. Profiling human antibody responses by integrated single-cell analysis. *Vaccine* 32, 2866–2873.
- Olofsson, P.E., Forslund, E., Vanherberghen, B., *et al.*, 2014. Distinct migration and contact dynamics of resting and IL-2-activated human natural killer cells. *Frontiers in Immunology* 5, 80.
- Sackmann, E.K., Berthier, E., Young, E.W.K., *et al.*, 2012. Microfluidic kit-on-a-lid: A versatile platform for neutrophil chemotaxis assays. *Blood* 120, 45–54.
- Shalek, A.K., Satija, R., Shuga, J., *et al.*, 2014. Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. *Nature* 509, 363–369.
- Shin, Y.S., Remacle, F., Fan, R., *et al.*, 2011. Protein signaling networks from single cell fluctuations and information theory profiling. *Biophysical Journal* 100, 2378–2386.
- Shulman, Z., Gitlin, A.D., Targ, S., *et al.*, 2013. T follicular helper cell dynamics in germinal centers. *Science* 341, 673–677.
- Sip, C.G., Bhattacharjee, N., Folch, A., 2014. Microfluidic transwell inserts for generation of tissue culture-friendly gradients in well plates. *Lab on a Chip* 14, 302–314.
- Story, C.M., Papa, E., Hu, C.-C.A., *et al.*, 2008. Profiling antibody responses by multiparametric analysis of primary B cells. *Proceedings of the National Academy of Sciences of the United States of America* 105, 17902–17907.
- Sundd, P., Gutierrez, E., Koltsova, E.K., *et al.*, 2012. "Slings" enable neutrophil rolling at high shear. *Nature* 488, 399–403.
- Tay, S., Hughey, J.J., Lee, T.K., *et al.*, 2010. Single-cell NF-kappaB dynamics reveal digital activation and analogue information processing. *Nature* 466, 267–271.
- Thorsen, T., Maerkl, S.J., Quake, S.R., 2002. Microfluidic large-scale integration. *Science* 298, 580–584.
- Vanherberghen, B., Olofsson, P.E., Forslund, E., *et al.*, 2013. Classification of human natural killer cells based on migration behavior and cytotoxic response. *Blood* 121, 1326–1334.
- Varadarajan, N., Julg, B., Yamanaka, Y.J., *et al.*, 2011. A high-throughput single-cell analysis of human CD8 + T cell functions reveals discordance for cytokine secretion and cytotoxicity. *Journal of Clinical Investigation* 121, 4322–4331.
- Varadarajan, N., Kwon, D.S., Law, K.M., *et al.*, 2012. Rapid, efficient functional characterization and recovery of HIV-specific human CD8 + T cells using microengraving. *Proceedings of the National Academy of Sciences of the United States of America* 109, 3885–3890.
- Vedel, S., Tay, S., Johnston, D.M., Bruus, H., Quake, S.R., 2013. Migration of cells in a social context. *Proceedings of the National Academy of Sciences of the United States of America* 110, 129–134.
- Veening, J.-W., Smits, W.K., Kuipers, O.P., 2008. Bistability, epigenetics, and bet-hedging in bacteria. *Annual Review of Microbiology* 62, 193–210.
- Walker, G.M., Sai, J., Richmond, A., *et al.*, 2005. Effects of flow and diffusion on chemotaxis studies in a microfabricated gradient generator. *Lab on a Chip* 5, 611–618.
- Wu, J., Ouyang, L., Wadhawan, N., *et al.*, 2014. A compact microfluidic system for cell migration studies. *Biomedical Microdevices* 16, 521–528.
- Yamanaka, Y.J., Berger, C.T., Sips, M., *et al.*, 2012. Single-cell analysis of the dynamics and functional outcomes of interactions between human natural killer cells and target cells. *Integrative Biology* 4, 1175–1184.
- Zhou, P., Shaffer, D.R., Alvarez Arias, D.A., *et al.*, 2014. *In vivo* discovery of immunotherapy targets in the tumour microenvironment. *Nature* 506, 52–57.
- Zhu, H., Stybayeva, G., Silangcruz, J., *et al.*, 2009. Detecting cytokine release from single T-cells. *Analytical Chemistry* 81, 8150–8156.

Cellular Structures Controlling T Cell Signaling in Time and Space

C Wülfing, University of Bristol, Bristol, UK

KT Roybal, University of California, San Francisco, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

T cells activate in a cellular interaction with antigen presenting cells (APC). The central activating receptor is the T cell receptor (TCR) that recognizes antigenic peptides presented by major histocompatibility complex (MHC) molecules on the surface of the APC (Davis *et al.*, 2003). A substantial number of costimulatory receptors regulate the efficiency and functional outcomes of T cell activation. The ligands for most of these receptors are also expressed on the APC surface (Davis *et al.*, 2003). A large number of ligand-engaged receptors thus initiate T cell signaling at the cellular interface between a T cell and an APC in a polarized fashion. This triggers the activation of an even larger number of associated signaling intermediates. Understanding the functional integration of such complex signaling is an important unresolved challenge that is common among multiple cell types. Critical data to address this challenge are provided by the polarized nature of the subcellular organization of T cell signaling: During T cell activation, receptors and signaling intermediates are not necessarily evenly distributed within the T cell but can enrich at particular regions at specific times, creating complex spatiotemporal signaling distributions. The resulting highly structured interface between the T cell and the APC is often referred to as an 'immunological synapse'. As co-enrichment of signaling intermediates enhances the efficiency of their interaction, uneven spatiotemporal signaling distributions control the flow of information through a signaling network and thus govern the efficiency of cellular activation. Understanding signaling organization thus provides unique insight into the function of complex signaling systems. While uneven signaling distributions are widespread in the activation of many cell types, a wealth of genetic, biochemical, and cell biological data makes T cell activation particularly well-suited to understand how subcellular signaling distributions are controlled and how they regulate cellular activation. Here we discuss approaches to study the organization of T cell signaling, principal structures driving it as the focus of this article, and difficulties in the determination of the functional significance of uneven subcellular signaling distributions that have limited its analysis to date.

Studying the Organization of T Cell Signaling

Approaches to Activate T Cells

T cell signaling plays a variety of roles in the *in vivo* life cycle of a T cell. T cells develop in the thymus in response to modest strength engagement of the TCR resulting in various distinct subsets, prominently CD4 helper and CD8 cytotoxic T cells. To ensure a diverse recognition repertoire, virtually every T cell expresses a different TCR. Upon completion of the thymic

development a 'naïve' T cell is highly motile, but remains largely quiescent from a cell cycle standpoint until it is triggered by strong engagement of its TCR. Naïve T cell activation triggers extensive functional differentiation, for example, into various helper subsets. Activated or 'primed' T cells execute effector functions within secondary lymphoid organs or at sites of infection in response to reengagement of their TCRs. The vast majority of activated T cells die within about a week with few T cells surviving as memory T cells. Throughout their life T cells interact with a variety of different cell types as APCs, for example, with the thymic epithelium during development, with dendritic cells during naïve T cell activation, or with B cells in support of their differentiation. This complex life cycle sets up a bewildering variety of scenarios to study T cell activation and the specific signaling organization associated with it. Here we will focus on the activation of primed CD4 T cells as the most widely studied experimental model.

As practically every T cell expresses a different TCR, a population of *in* or *ex vivo* T cells cannot be activated through engagement of their TCRs by a single MHC peptide complex. There are various approaches to circumvent this problem. Mice can be genetically engineered to have all of their T cells express the same TCR with a defined MHC/peptide recognition specificity. Much of the work reviewed here uses this approach. Human T cells with a given recognition specificity can be maintained through periodic restimulation with their cognate antigen *in vitro* or they can be immortalized. Alternatively, antibodies or superantigens can be used to activate T cells irrespective of their TCR recognition specificity, as commonly used to study human T cells. Different activation regimes can be executed in the context of different APCs, B cells are widely used, or by employing reductionist approaches as introduced in the following paragraph. Given the large number of different scenarios and methods to activate a T cell, it is of utmost importance to keep in mind that the experimental constraints of the particular approach of T cell activation chosen differentially shape T cell signaling and its subcellular organization (Brossard *et al.*, 2005; Schubert *et al.*, 2012; Richie *et al.*, 2002; Ebert *et al.*, 2008; Hailman *et al.*, 2002; Zanin-Zhorov *et al.*, 2010).

Methods to Study T Cell Organization

Physiological activation of T cells occurs in the cellular interaction of T cells with various types of APCs in secondary lymphoid organs or at a site of infection. Signaling organization thus should ideally be studied in this context. Two-photon microscopy has been incredibly successful in tracking the motion of immune cells inside intact organs (Okada *et al.*, 2005; Mempel *et al.*, 2004; Bajenoff *et al.*, 2006; Celli *et al.*, 2005; Witt *et al.*, 2005; Skokos *et al.*, 2007). However, technical constraints, in particular difficulties in distinguishing productive from accidental T cell:APC contacts and in

delineating T cell:APC interfaces as the site of signaling initiation, have made it very difficult to study subcellular signaling organization *in vivo*. With the exception of a few pioneering studies (Azar *et al.*, 2010; McGavern *et al.*, 2002; Melichar *et al.*, 2013), most of the work on the organization of T cell signaling has therefore been performed *in vitro*, the focus of this article. Supporting the utility of *in vitro* studies, cell coupling inside secondary lymphoid organs shows comparable dynamics and responses to varying T cell stimuli (Henrickson *et al.*, 2008; Zheng *et al.*, 2008; Azar *et al.*, 2010; Moreau *et al.*, 2012; Henrickson *et al.*, 2013) as cell coupling *in vitro*.

T cells are a difficult cell type to address with cell biological means, as T cells are non-adherent motile small cells (7–9 μm diameter) with most of the T cell volume taken up by the nucleus. As they are activated in a cellular interaction, only the subset of T cells in tight cell couples with APCs yields relevant data. T cell:APC couples can be observed by live cell imaging using, for example, spinning disk confocal microscopy. This requires the rapid acquisition of three-dimensional volumes, limiting resolution. Therefore, from its beginning T cell organization has also been studied in reductionist systems, where the APC is substituted by a flat surface, either supported lipid bilayers (SLB) or antibody-coated cover slips (Grakoui *et al.*, 1999). This allows precise control of the T cell stimulus and access to single z-plane high-resolution imaging, for example, in tracking diffraction-limited objects by total internal reflection fluorescence microscopy or by applying recent superresolution approaches. Glass cover slips can be coated with ligands for receptors on the T cell surface, commonly antibodies against the TCR. In this setting the location of receptor engagement is fixed as antibodies cannot move on the glass surface and signaling organization is determined in relation to this localized receptor engagement. Alternatively, ligands for receptors on the T cell surface can be inserted into SLB generated by spreading small unilamellar vesicles on glass. In this setting the ligands remain mobile, allowing for receptor:ligand transport at the interface between the T cell and the SLB, similar to a T cell:APC interface. However in contrast to a T cell:APC interface, the SLB cannot be deformed to allow for extensive membrane undulations and internalization of receptor:ligand couples into the T cell. Finally, only electron microscopy can resolve T cell organization at the scale of a few nm, as for example, required in analyzing membrane topology at the plasma membrane and in vesicles. However, this requires extensive sample processing.

The fundamental role of signaling organization is to regulate when and where signaling intermediates interact with each other. The more signaling intermediates are studied in the same experimental system the more signaling interactions can be addressed. Systems biology approaches thus are particularly powerful in understanding signaling organization. Initial studies with more than ten signaling intermediates analyzed with the same T cell:APC combination have revealed a remarkably diverse organization of T cell signaling in time and space and substantial sensitivity to alterations in T cell stimuli (Singleton *et al.*, 2011, 2009), establishing that T cell signaling organization can be highly information rich.

In summary, with the large variety of approaches to activate and image T cells, each individual study can only hope to capture a small part of signaling organization and its function.

Technical constraints of a particular approach used need to be kept in mind and every attempt should be made to integrate as much of the available data as possible under consideration of these constraints to slowly approach the scope of system insight required to understand something as complex as T cell signaling organization and its function.

Signaling Organization and the Structures Driving It

Signaling organization is governed by a large set of elements of cellular architecture that cross-regulate each other in a complex fashion. Critical elements include membrane topology, cytoskeletal structures, and complexes of receptors and/or signaling intermediates. Membrane topology at the cellular interface and through vesicular trafficking, determines how many receptors can engage their ligands at which location. Cytoskeletal structures control membrane topology and can constrain and move receptors and signaling intermediates. Receptor complexes control the efficiency of signaling initiation. Together with membrane topology and cytoskeletal structures they shape signaling distributions. Signaling complexes can stabilize such distributions and the resulting localization of signaling intermediates feeds back on the structures driving it, in particular through local control of cytoskeletal dynamics and vesicular trafficking. Individual elements of signaling organization thus cannot be comprehensively understood in isolation. Figure 1 presents all elements of cellular architecture discussed in context.

T Cell Shape and Plasma Membrane Topology

In general, the dynamics of cell coupling, for example, the duration of cell couples, the frequency of their often sequential formation, or the number of cells bound simultaneously, constrain T cell signal initiation and transduction, an important question that is beyond the scope of this article and addressed elsewhere (Davis, 2009; Batista and Dustin, 2013). At the level of individual cell couples as discussed here, the topology of the cellular interface between a T cell and an APC constitutes a primary determinant of the initiation of signaling through receptor engagement.

Even prior to binding to an APC, a T cell is polarized. Two defining elements of T cell polarity are an actin-rich leading edge and a posterior extension with the microtubule organizing center (MTOC) at its base, the uropod. Some receptors and signaling intermediates are enriched in the uropod, such as ICAM-2, ezrin, and moesin (Yonemura *et al.*, 1998; Savage *et al.*, 2002). A network of PDZ-containing proteins plays an important role in such polarization (Ludford-Menting *et al.*, 2005). Upon formation of a tight cell couple some receptors including the TCR are partially and transiently recruited to the uropod before being released to the interface when the uropod is retracted as a consequence of T cell activation (Singleton *et al.*, 2009). Other receptor and signaling intermediates are moved to or remain behind the nucleus in a tight cell couple, forming the distal pole complex, as reviewed elsewhere (Cullinan *et al.*, 2002). Nevertheless, upon formation of a tight cell couple, most receptors move to the T cell:APC interface (Singleton *et al.*, 2009).

Dynamic cellular structures that regulate the spatiotemporal organization of T cell signaling and activation

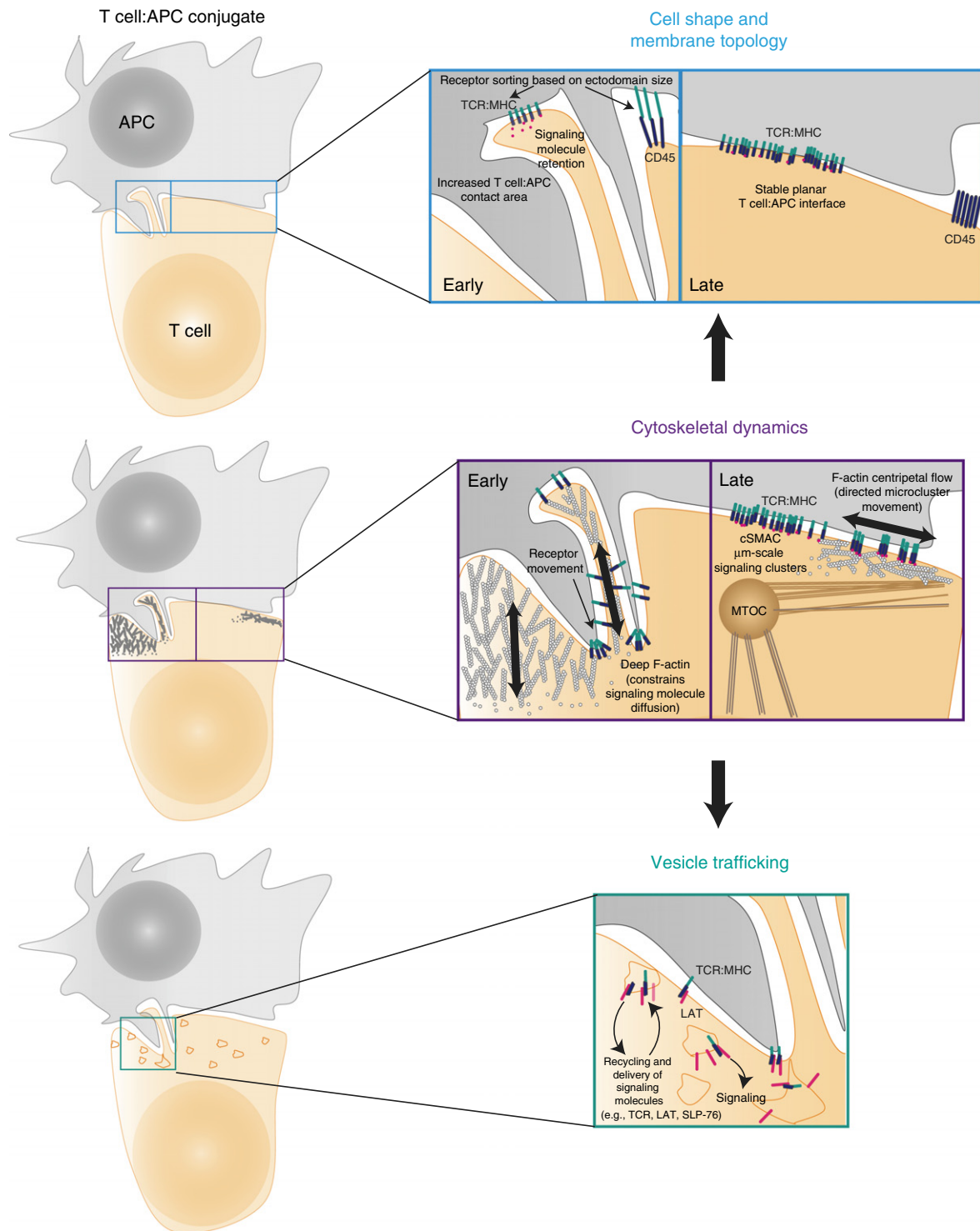


Figure 1 The figure summarizes the cellular structures that control the organization of T cell signaling in T cell:APC couples. The T cell is indicated in light brown, the APC in gray. In the three panels from top to bottom different structures are highlighted as indicated and panels on the right zoom into the cellular interface. As T cell cellular structures are dynamically regulated, two zoomed-in panels are shown, one representing early T cell activation, i.e., within the first 5 min of tight cell coupling at the peak of biochemical signaling activity, and the other representing late T cell activation thereafter. Receptors are indicated in blue, their ligands in turquoise, actin in light grey, microtubules in dark brown, and signaling intermediates in pink. Thick black arrows indicate direction of actin protrusion and retrograde transport.

In a T cell:APC couple the cellular interface contains numerous smaller and larger undulations, with T cell protrusions reaching into APCs and APC membrane being pulled into the T cell (Ueda *et al.*, 2011; Singleton *et al.*, 2006; Sage *et al.*, 2012). The cumulative effect of these undulations can increase the area of T cell membrane at the T cell:APC interface to up to 4-fold over the area that a completely flat interface would occupy. Membrane undulations thus dramatically enhance the area available to receptor:ligand engagement. The extent of these membrane undulations is dynamically regulated. The undulations are most dramatic within the first couple of minutes of tight T cell:APC interface formation (Roybal *et al.*, 2013) when biochemical signaling activity peaks but are sustained somewhat less prominently over hours (Ueda *et al.*, 2011). At least the early burst in membrane undulations is likely driven by actin dynamics, when numerous small actin structures are aligned perpendicular to the cellular interface, possibly driving it outward. Directly linking membrane topology to signaling organization, the TCR can enrich in small membrane protrusions (Sage *et al.*, 2012). A distinct element of membrane undulations is a large T cell invagination that forms selectively within the first couple of minutes of tight cell coupling at the center of the T cell:APC interface and that can reach several μm deep into the T cell with APC membrane still attached (Singleton *et al.*, 2006). Signaling intermediates, for example, active Rho, and receptors such as CD2 or even the TCR are enriched at the base of the transient invagination (Singleton *et al.*, 2006, 2011). However, their subsequent fate is unresolved and the suggested function of the invagination in resetting early T cell signaling needs to be corroborated.

The distance between the T cell and APC plasma membranes at the even smaller scale of tens of nm is critical for receptor function and localization. Experimental and theoretical studies show that receptor ligand confinement at a cellular interface shapes the affinity of receptor ligand binding, likely as a function of the membrane distance (Huppa *et al.*, 2010; Hu *et al.*, 2013; Dustin *et al.*, 1996). In addition, it has long been suggested that the distance between the T cell and the APC membrane is critical for sorting receptor:ligand couples into distinct areas by size. Strong support for this notion comes from experiments where changing the dimensions of receptors on the T cell surface dramatically impacts their function (Wild *et al.*, 1999; Choudhuri *et al.*, 2005). In addition, receptors with a particularly large extracellular domain, prominently CD43 and the phosphatases CD45 and CD148 tend to be removed from the cellular interface upon cell coupling (Lin and Weiss, 2003; Leupin *et al.*, 2000; Allenspach *et al.*, 2001; Delon *et al.*, 2001; Roumier *et al.*, 2001; Johnson *et al.*, 2000) or from the vicinity of ligand-engaged TCRs when analyzed in T cells activated on SLBs (Varma *et al.*, 2006), consistent with a size-based sorting mechanism.

The function of the different forms of membrane undulations is uncertain, as it is virtually impossible to establish causality by changing T cell membrane topology experimentally without affecting various other elements of cellular architecture and signaling. Nevertheless, from biophysical first principles these structures should shape T cell signaling to a substantial extent. For example, altered membrane curvature and the dramatically increased membrane surface to

cytoplasm volume ratio in T cell membrane protrusions should alter signal progression (Meyers *et al.*, 2006; Neves *et al.*, 2008; Rangamani *et al.*, 2013; Schmick and Bastiaens, 2014). Plasma membrane undulations as highly deformed membrane structures should affect molecular interactions at the membrane, as for example, Bin/Amphiphysin/Rvs (BAR) proteins bind to and maintain curved membrane structures providing a link to various actin regulators (Antonny, 2011).

Vesicle Trafficking

Continuously, secretory and recycling vesicles fuse with the plasma membrane and endocytic vesicles internalize at the plasma membrane, controlling the area, and composition of the plasma membrane. In addition, exocytic vesicles can bud from the plasma membrane. The number of receptors accessible to ligand at the T cell:APC interface thus can be carefully controlled. As a prominent example, even prior to APC contact the TCR continuously cycles through the endocytic pathway. Upon T cell activation, a substantial decrease in TCR plasma membrane expression is mediated through lysosomal rerouting of the TCR in endosomes (Liu *et al.*, 2000; Das *et al.*, 2004; Finetti *et al.*, 2009). In addition, vesicles containing TCRs are secreted in an ESCRT-dependent fashion by the T cell at the center of the T cell:APC interface, again preventing the TCRs from participation in signal transduction (Choudhuri *et al.*, 2014). Inversely to vesicular TCR trafficking, the critical inhibitory receptors CTLA-4 and Lag-3 are largely contained in vesicles inside the T cell and only become inserted into the plasma membrane upon the activation of a naïve T cell (Woo *et al.*, 2010; Jansson *et al.*, 2005; Linsley *et al.*, 1996). The control of interface localization through vesicular trafficking also extends to signaling intermediates, in particular receptor proximal signaling intermediates that are membrane associated and their interaction partners. As prominent examples, the tyrosine kinase Lck that phosphorylates the TCR and the adaptor proteins Linker of activated T cells (LAT) and SLP-76 that constitutes the principal signaling platforms immediately downstream of the TCR are inserted into the plasma membrane from vesicles in distinctly regulated fashions and are internalized as a consequence of TCR engagement (Soares *et al.*, 2013; Anton *et al.*, 2011; Barr *et al.*, 2006). Interestingly, internalized TCR can remain active as long as it remains associated with Lck (Yudushkin and Vale, 2010) even though it is unresolved, whether the MHC/peptide ligand or membrane patches from the APC surface containing it can be pulled out of the APC surface to provide continued TCR engagement in the endocytic vesicles. The question of how much of LAT activity during T cell activation is provided by the plasma membrane pool of LAT versus LAT coming from secretory and recycling vesicles is discussed controversially (Purbhoo *et al.*, 2010; Williamson *et al.*, 2011; Balagopalan *et al.*, 2013). Again, it is exceedingly difficult to generate causal experimental support for functional relevance of specific trafficking events. Nevertheless corroborating the principal importance of vesicular trafficking, interference with the function of general regulators of vesicle fusion and trafficking has consistently shown that vesicular trafficking shapes T cell signaling to a substantial extent (Larghi *et al.*, 2013; Vardhana *et al.*, 2010; Gomez *et al.*, 2005).

Cytoskeletal Dynamics

For more than two decades it is well established that the T cell cytoskeleton with its two principal components, the microtubule network originating at the MTOC and the actin network, repolarizes upon formation of a T cell:APC couple (Kupfer and Singer, 1989). In a migrating T cell prior to cell coupling the MTOC is located behind the nucleus at the base of the uropod. Upon cell couple formation the MTOC swings around the nucleus within a few minutes to become positioned right behind the center of the T cell:APC interface (Tskvitaria-Fuller *et al.*, 2003; Kuhn and Poenie, 2002). MTOC reorientation is driven by TCR signaling (Lowin-Kropf *et al.*, 1998; Quann *et al.*, 2011). The mechanism of MTOC reorientation is still being debated. Suggested mechanisms include attachment of microtubule ends in a ring around the interface periphery with centripetal pulling to center it (Kuhn and Poenie, 2002) and attachment of microtubule ends at the interface center with capture-shrinkage to position the MTOC (Yi *et al.*, 2013). MTOC localization is important for signaling organization as the vast majority of the T cell vesicular machinery moves with the MTOC (Quintana *et al.*, 2007).

Actin dynamics extensively shape signaling organization by direct and indirect means: They control cell shape, membrane topology, and short-distance vesicle trafficking, signaling intermediates can position themselves through direct actin binding and can be transported by actin motion, and dense actin matrices can trap signaling complexes. Actin dynamics are therefore likely the most studied element of signaling organization in T cell activation. There are four principal T cell actin structures. Within seconds of the initiation of T cell coupling, be it to an APC or a planar substitute thereof, actin forms a peripheral ring that pushes the edge of the interface outward in the formation of a tight cell couple (Bunnell *et al.*, 2001; Tskvitaria-Fuller *et al.*, 2003). In T cell:APC couples the peripheral actin ring is prominent only for a few minutes, in T cell activation on planar APC substitutes it last considerably longer. In T cell:APC couples the peripheral actin ring can be accompanied by an equally transient second structure, a dense actin matrix that reaches from the undulating cellular interface deep into the T cell across the entire interface (Roybal *et al.*, 2013). A similar matrix is seen in the coupling of natural killer cells to their targets (Rak *et al.*, 2011; Brown *et al.*, 2011). During the entire duration of cell coupling, the cortical actin network underlying the T cell plasma membrane similar to other cell types enriches at the cellular interface. Finally, after a few minutes of cell coupling, actin often enriches in individual lamellar extensions that can occur at the edge of the interface to maintain it or away from the interface when a T cell repolarizes to another APC.

Actin is highly dynamic. The use of planar APC substitutes has been instrumental in providing access to F-actin motion. When T cells are activated on planar APC substitutes, actin moves from the peripheral actin ring inward, possibly with decreasing velocity as it approaches the interface center (Yi *et al.*, 2012; Babich *et al.*, 2012). This actin motion provides a centripetal transport machinery for the organization of T cell signaling (Babich *et al.*, 2012; Yi *et al.*, 2012). The role of the motor protein Myosin II in this actin motion is still being debated (Hammer and Burkhardt, 2013). It needs to be

established how this centripetal actin motion is oriented in a T cell:APC couple, where in particular early during cell coupling large membrane undulations impair the formation of a continuous actin network in the interface plane, rather breaking actin up into discreet structures aligned with the undulations (Figure 1). Actin dynamics are controlled by TCR signaling in conjunction with costimulation (Tskvitaria-Fuller *et al.*, 2003; Wülfing and Davis, 1998). Some interface actin enrichment accompanies T cell coupling in response to virtually every stimulus. However, the extent and dynamics of actin turnover are likely used by the T cells as a critical means to tune to efficiency of T cell signaling (Roybal *et al.*, 2013). Illustrating the general importance of actin dynamics for T cell activation, genetic deletion of virtually every protein that controls actin turnover, for example, WAVE2, the hematopoietic cortactin homolog HS1, Cofilin, or Coronin, leads to substantial defects in signaling organization and cellular function (Billadeau *et al.*, 2007; Gomez *et al.*, 2006; Nolz *et al.*, 2006; Badour *et al.*, 2003; Snapper *et al.*, 1998; Cannon and Burkhardt, 2004; Cannon *et al.*, 2001; Labno *et al.*, 2003; Zipfel *et al.*, 2006; Foger *et al.*, 2006; Mugnier *et al.*, 2008; Eibert *et al.*, 2004). However, again it is technically exceedingly difficult to causally link distinct actin structures to signaling organization and the regulation of T cell activation.

Signaling Complexes – The Central Supramolecular Activation Cluster

Already in the first studies characterizing T cell signaling organization, accumulation of receptors, for example, ligand-engaged TCR, and signaling intermediates, for example, protein kinase C (PKC) Θ , at the center of the T cell:APC interface was described (Monks *et al.*, 1998, 1997; Grakoui *et al.*, 1999). Kupfer termed such accumulation the central supramolecular activation complex or central supramolecular activation cluster (cSMAC). Work using planar APC substitutes has established that ligand-engaged receptors move centripetally in the T cell:APC interface to enrich in the cSMAC (Varma *et al.*, 2006; Yokosuka *et al.*, 2005; Kaizuka *et al.*, 2007). The velocity of this receptor transport is similar to that of the centripetal actin motion, suggesting that actin drives receptor transport. Proteins in the cSMAC, be it in T cells activated with APCs or on SLB, display dramatically reduced mobility, consistent with the idea that the cSMAC is a highly cross-linked protein complex with potentially distinct physical phase properties (Rocha-Perugini *et al.*, 2013; Li *et al.*, 2012). The importance of the cSMAC for T cell activation has long been controversial. A large number of receptors and signaling intermediates accumulate at the interface center when T cells receive a strong stimulus (Singleton *et al.*, 2009). However, efficient T cell activation can occur in the absence of cSMAC formation and in response to weaker stimuli receptor clustering in the cSMAC is commonly lost even though accumulation of PKC Θ is retained (Singleton *et al.*, 2009; Purdie *et al.*, 2005; Lee *et al.*, 2002; Purbhoo *et al.*, 2004). Curiously the accumulation of active signaling intermediates in the cSMAC is less prominent than that of ligand-engaged receptors (Singleton *et al.*, 2009). This is particularly pronounced in T cells activated on SLB, where hardly any signaling activity resides in the cSMAC (Varma *et al.*, 2006). Based on these

observations, it seems likely that the mutual clustering of numerous receptors and signaling intermediates with high valence interaction motifs drives cSMAC formation as enhanced by centripetal actin transport. Turnover of receptors and signaling intermediates in the cSMAC as likely influenced by vesicular trafficking then determines how much signaling activity resides in the cSMAC. When the cSMAC can turnover efficiently to allow continuous exchange of molecules with the remainder of the T cell signaling system, it may enhance signaling by allowing signaling intermediates to dynamically concentrate, as possibly the case in T cell:APC couples. When the cSMAC becomes sufficiently immobile to prevent exchange with the remainder of the cell or when parts of it become separated from the cytoplasm through sorting into extracellular vesicles (Choudhuri *et al.*, 2014), the cSMAC may impair signaling by sequestering receptors and signaling intermediates, as possibly the case in T cell activation on planar surfaces. The sorting of TCR into extracellular microvesicles spares engaged CD28, which continues to signal through PKC- θ . Independent of the amount of signaling provided by the cSMAC, a substantial part of T cell signaling occurs outside the cSMAC in smaller receptor and signaling intermediate assemblies.

Signaling Complexes – Microclusters

The use of planar surfaces as APC substitutes has been instrumental in identifying smaller receptor and signaling clusters. Even prior to T cell activation in naïve and even more so primed T cells, TCRs are localized in clusters of dozens of molecules (Boyle *et al.*, 2011; Fahmy *et al.*, 2001; Kumar *et al.*, 2011). Upon T cell activation on SLB it can be readily observed that such clusters grow to up to hundreds of molecules, giving rise to ‘microclusters’ (Yokosuka *et al.*, 2005; Varma *et al.*, 2006). A single agonist peptide/MHC complex is sufficient to drive the clustering of hundreds of TCRs (Huang *et al.*, 2013). Microclusters form at the periphery of the interface and upon initial TCR signaling associate with a substantial part of known proximal signaling intermediates, starting with ZAP-70, the adaptor proteins LAT and SLP-76 all the way to Nck and WASP (Bunnell *et al.*, 2002; Barda-Saad *et al.*, 2005). Ligand-engaged TCRs then move toward the interface center in an actin-dependent process (Xie *et al.*, 2012). Microclusters need to have a minimal size to be transported centripetally (Hartman *et al.*, 2009). During transport, more distal signaling intermediates dissociate from the microclusters over time, SLP-76 to move to a perinuclear structure, Nck and WASP to move into the actin-rich region at the periphery of the interface between the T cell and the planar surface (Bunnell *et al.*, 2002; Barda-Saad *et al.*, 2005). Upon reaching the cSMAC, much of the remaining associated signaling activity is lost (Varma *et al.*, 2006). When T cell activation is disrupted either by blocking TCR engagement or actin dynamics, receptor clustering in the microclusters rapidly dissolves, suggesting that microclusters are dynamically maintained (Varma *et al.*, 2006). Super-resolution and electron microscopy have provided further detail in the characterization of signaling clusters. Prior to T cell activation, clusters of T cell surface molecules, including the TCR, LAT and lipid markers were found to be mostly separate. Upon TCR engagement, clusters grow in size and become adjacent to each other, but only partially mix

(Sherman *et al.*, 2011; Lillemeier *et al.*, 2006, 2010). Similar spatial separation at the nm scale of ligand-engaged receptors and LAT is found in mast cell activation (Wilson *et al.*, 2001).

Together, the work using planar APC substitutes strongly suggests that the control of the clustering of receptors and signaling intermediates, overlapping but also with clear spatial distinctions, in combination with actin-dependent transport of signaling clusters plays a major role in shaping the organization of T cell signaling. It remains to be seen, how the more undulating plasma membrane topology in T cell:APC couples with its increased membrane flexibility and more local actin structures shape the formation and transport of signaling cluster at the T cell:APC interface. Curiously, the spatial segregation of proximal and distal signaling observed in T cells activated on planar APC substitutes is substantially more pronounced in T cell:APC couples (Roybal *et al.*, 2013). Moreover, it remains unresolved how much signaling activity is mediated by even smaller groups or individual receptors outside the microclusters with some such signaling activity likely (Schamel *et al.*, 2005).

Summary of Signaling Structures

A large number of structures govern signaling organization in T cell activation, membrane topology, vesicular trafficking, actin dynamics, and large and smaller protein complexes. Importantly, the extent to which these structures form varies depending on how the T cell is activated. Moreover, the mapping of signaling intermediates onto these structures is to date only been executed in a highly fragmented fashion, with generally only few signaling intermediates mapped to few structures in different experimental systems. As a field we are thus still far away from comprehensively understanding how T cell signaling is organized in time and space. Nevertheless, the remarkable extent of the control of the localization of receptors and signaling intermediates inside a T cell discussed here makes it is safe to assume that the organization of signaling is a principal regulator of T cell function.

Signaling Organization Varies with T Cell Activation Conditions

Even while a comprehensive picture of the organization of T cell signaling is still lacking, individual smaller data sets already strongly suggest that signaling organization is highly sensitive to T cell activation conditions. As an illustration, a few almost random examples are listed here. In T cell maturation thymocytes undergoing positive and negative selection display hardly any or distinctly less patterning (Richie *et al.*, 2002; Ebert *et al.*, 2008). In T cell differentiation Th2-polarized T cells show less central TCR patterning and lipid raft accumulation than Th0 or Th1 cells (Balamuth *et al.*, 2001; Singleton *et al.*, 2011). Regulatory T cells have PKC θ shifted from the T cell:APC interface to the distal pole (Zanin-Zhorov *et al.*, 2010) and recruitment of a large GIT/Pix complex to CTLA-4 is restricted to regulatory T cells (Kong *et al.*, 2014). T cells expressing a self-reactive TCRs yield less pronounced spatiotemporal organization upon activation than those expressing a TCR recognizing foreign antigens (Schubert *et al.*, 2012). Engagement of the

TCR with a lower affinity peptide widely impairs accumulation at the interface center (Monks *et al.*, 1998; Singleton *et al.*, 2009). While these examples provide an idea of the impressive variability of signaling organization, much work needs to be done to move from these individual observations to a more comprehensive understanding of how T cell activation conditions govern signaling organization.

The Function of Signaling Organization

By controlling when and where signaling interactions can efficiently proceed, signaling organization should be of fundamental importance in determining, how the activity of dozens of receptor and signaling intermediates is integrated in a signaling network. This should be critical in tuning cellular activation, in T cells for example in the adjustment of T cell reactivity in the distinction between protection of our own body versus effective attack on cancer cells and infectious agents. However the comprehensive verification of this prediction of the functional importance of signaling organization is still lacking, as signaling organization is uniquely difficult to study in an unambiguous causal fashion. When one wants to understand the function of a particular protein, knockout mice, RNAi or overexpression provide straightforward approaches for causal manipulation. However, required for a definite assessment of the importance of a particular signaling structure are experiments where this structure is changed while others are not directly targeted and while the activity of signaling intermediates is not directly affected. Such experimental regimens are exceedingly difficult to generate, let alone in a consistent fashion. Therefore signaling organization has to be studied slightly differently. When a structure cannot be interfered with specifically, then it should be carefully characterized what happens to other structures in any chosen interference regimen. For example, to assess the importance of the peripheral actin ring, it should be evaluated what the effect of a drug that effectively blocks peripheral actin ring formation is on other actin structures, such as cortical or lamellar actin. Thus it can be established that the intended structure is indeed the major target of the drug treatment and in retrospect, consequences of the drug treatment could be linked to its smaller effects on other structures if prompted by emerging data. In addition, multiple such studies are likely needed to understand the importance of a single structure: Only when the loss of a structure is consistently associated with the same outcome, some certainty is gained about its function. This can be accomplished within a single study, where multiple interference regimes are tested or it can happen across multiple studies from different laboratories. A good example for the integrative effect of multiple studies is the investigation of the importance of the cSMAC, where initial studies (Monks *et al.*, 1997; Lee *et al.*, 2003, 2002; Stinchcombe *et al.*, 2001) reached widely divergent conclusions, that the cSMAC is required for signaling, not required for signaling, required for directed vesicular secretion. Only over many years a more comprehensive picture, as discussed above, has emerged. Across the field the function of T cell signaling organization and the various cellular structures associated with it thus remains widely uncertain. Because of the fragmented and incomplete

knowledge of the function of signaling organization, a comprehensive review of the data addressing the functional significance of the different elements of subcellular architecture discussed here is well beyond the scope of this manuscript. However, many methods and data addressing T cell organization have been discussed in two entire recent journal volumes of reviews dedicated to the cell coupling and the cytoskeleton (Burkhardt, 2013; Batista and Dustin, 2013).

See also: Cellular Immunology: Cell–Cell Interactions and the Immune System: T Cell Receptor Triggering

References

- Allenspach, E.J., Cullinan, P., Tong, J., *et al.*, 2001. ERM-dependent movement of CD43 defines a novel protein complex distal to the immunological synapse. *Immunity* 15, 739–750.
- Anton, O.M., Andres-Delgado, L., Reglero-Real, N., Batista, A., Alonso, M.A., 2011. MAL protein controls protein sorting at the supramolecular activation cluster of human T lymphocytes. *Journal of Immunology* 186, 6345–6356.
- Antonny, B., 2011. Mechanisms of membrane curvature sensing. *Annual Review of Biochemistry* 80, 101–123.
- Azar, G.A., Lemaitre, F., Robey, E.A., Bousso, P., 2010. Subcellular dynamics of T cell immunological synapses and kinapses in lymph nodes. *Proceedings of the National Academy of Sciences of the United States of America* 107, 3675–3680.
- Babich, A., Li, S., O'Connor, R.S., *et al.*, 2012. F-actin polymerization and retrograde flow drive sustained PLCgamma1 signaling during T cell activation. *Journal of Cell Biology* 197, 775–787.
- Badour, K., Zhang, J., Siminovich, K.A., 2003. The Wiskott-Aldrich syndrome protein: Forging the link between actin and cell activation. *Immunological Reviews* 192, 98–112.
- Bajenoff, M., Egen, J.G., Koo, L.Y., *et al.*, 2006. Stromal cell networks regulate lymphocyte entry, migration, and territoriality in lymph nodes. *Immunity* 25, 989–1001.
- Balogopalan, L., Barr, V.A., Kortum, R.L., Park, A.K., Samelson, L.E., 2013. Cutting edge: Cell surface linker for activation of T cells is recruited to microclusters and is active in signaling. *Journal of Immunology* 190, 3849–3853.
- Balamuth, F., Leitenberg, D., Unteraehrer, J., Mellman, I., Bottomly, K., 2001. Distinct patterns of membrane microdomain partitioning in Th1 and Th2 cells. *Immunity* 15, 729–738.
- Barda-Saad, M., Braiman, A., Titerence, R., *et al.*, 2005. Dynamic molecular interactions linking the T cell antigen receptor to the actin cytoskeleton. *Nature Immunology* 6, 80–89.
- Barr, V.A., Balagopalan, L., Barda-Saad, M., *et al.*, 2006. T-cell antigen receptor-induced signaling complexes: Internalization via a cholesterol-dependent endocytic pathway. *Traffic* 7, 1143–1162.
- Batista, F.D., Dustin, M.L., 2013. Cell:Cell interactions in the immune system. *Immunological Reviews* 251, 7–12.
- Billadeau, D.D., Nolz, J.C., Gomez, T.S., 2007. Regulation of T-cell activation by the cytoskeleton. *Nature Reviews Immunology* 7, 131–143.
- Boyle, S., Kolin, D.L., Bieler, J.G., *et al.*, 2011. Quantum dot fluorescence characterizes the nanoscale organization of T cell receptors for antigen. *Biophysical Journal* 101, L57–L59.
- Brossard, C., Feuillet, V., Schmitt, A., *et al.*, 2005. Multifocal structure of the T cell – Dendritic cell synapse. *European Journal of Immunology* 35, 1741–1753.
- Brown, A.C., Oddos, S., Dobbie, I.M., *et al.*, 2011. Remodelling of cortical actin where lytic granules dock at natural killer cell immune synapses revealed by super-resolution microscopy. *PLoS Biology* 9, e1001152.
- Bunnell, S.C., Hong, D.I., Kardon, J.R., *et al.*, 2002. T cell receptor ligation induces the formation of dynamically regulated signaling assemblies. *Journal of Cell Biology* 158, 1263–1275.
- Bunnell, S.C., Kapoor, V., Tribble, R.P., Zhang, W., Samelson, L.E., 2001. Dynamic actin polymerization drives T cell receptor-induced spreading: A role for the signal transduction adaptor LAT. *Immunity* 14, 315–329.

- Burkhardt, J.K., 2013. Cytoskeletal function in the immune system. *Immunological Reviews* 256, 5–9.
- Cannon, J.L., Burkhardt, J.K., 2004. Differential roles for Wiskott-Aldrich syndrome protein in immune synapse formation and IL-2 production. *Journal of Immunology* 173, 1658–1662.
- Cannon, J.L., Labno, C.M., Bosco, G., *et al.*, 2001. Wasp recruitment to the T cell: APC contact site occurs independently of Cdc42 activation. *Immunity* 15, 249–259.
- Celli, S., Garcia, Z., Bousso, P., 2005. CD4 T cells integrate signals delivered during successive DC encounters *in vivo*. *Journal of Experimental Medicine* 202, 1271–1278.
- Choudhuri, K., Llodra, J., Roth, E.W., *et al.*, 2014. Polarized release of T-cell-receptor-enriched microvesicles at the immunological synapse. *Nature* 507, 118–123.
- Choudhuri, K., Wiseman, D., Brown, M.H., Gould, K., van der Merwe, P.A., 2005. T-cell receptor triggering is critically dependent on the dimensions of its peptide-MHC ligand. *Nature* 436, 578–582.
- Cullinan, P., Sperling, A.I., Burkhardt, J.K., 2002. The distal pole complex: A novel membrane domain distal to the immunological synapse. *Immunological Reviews* 189, 111–122.
- Das, V., Nal, B., Dujeancourt, A., *et al.*, 2004. Activation-induced polarized recycling targets T cell antigen receptors to the immunological synapse; involvement of SNARE complexes. *Immunity* 20, 577–588.
- Davis, D.M., 2009. Mechanisms and functions for the duration of intercellular contacts made by lymphocytes. *Nature Reviews Immunology* 9, 543–555.
- Davis, S.J., Ikemizu, S., Evans, E.J., *et al.*, 2003. The nature of molecular recognition by T cells. *Nature Immunology* 4, 217–224.
- Delon, J., Kaibuchi, K., Germain, R.N., 2001. Exclusion of CD43 from the immunological synapse is mediated by phosphorylation-regulated relocation of the cytoskeletal adaptor moesin. *Immunity* 15, 691–701.
- Dustin, M.L., Ferguson, L.M., Chan, P.Y., Springer, T.A., Golan, D.E., 1996. Visualization of CD2 interaction with LFA-3 and determination of the two-dimensional dissociation constant for adhesion receptors in a contact area. *Journal of Cell Biology* 132, 465–474.
- Ebert, P.J., Ehrlich, L.I., Davis, M.M., 2008. Low ligand requirement for deletion and lack of synapses in positive selection enforce the gauntlet of thymic T cell maturation. *Immunity* 29, 734–745.
- Eibert, S.M., Lee, K.H., Pipkorn, R., *et al.*, 2004. Cofilin peptide homologs interfere with immunological synapse formation and T cell activation. *Proceedings of the National Academy of Sciences of the United States of America* 101, 1957–1962.
- Fahmy, T.M., Bieler, J.G., Edidin, M., Schneek, J.P., 2001. Increased TCR avidity after T cell activation: A mechanism for sensing low-density antigen. *Immunity* 14, 135–143.
- Finetti, F., Paccani, S.R., Riparbelli, M.G., *et al.*, 2009. Intraflagellar transport is required for polarized recycling of the TCR/CD3 complex to the immune synapse. *Nature Cell Biology* 11, 1332–1339.
- Foger, N., Rangell, L., Danilenko, D.M., Chan, A.C., 2006. Requirement for coronin 1 in T lymphocyte trafficking and cellular homeostasis. *Science* 313, 839–842.
- Gomez, T.S., Hamann, M.J., McCarney, S., *et al.*, 2005. Dynamin 2 regulates T cell activation by controlling actin polymerization at the immunological synapse. *Nature Immunology* 6, 261–270.
- Gomez, T.S., McCarney, S.D., Carrizosa, E., *et al.*, 2006. HS1 functions as an essential actin-regulatory adaptor protein at the immune synapse. *Immunity* 24, 741–752.
- Grakoui, A., Bromley, S.K., Sumen, C., *et al.*, 1999. The immunological synapse: A molecular machinery controlling T cell activation. *Science* 285, 221–226.
- Hailman, E., Burack, W.R., Shaw, A.S., Dustin, M.L., Allen, P.M., 2002. Immature CD4(+)CD8(+) thymocytes form a multifocal immunological synapse with sustained tyrosine phosphorylation. *Immunity* 16, 839–848.
- Hammer 3RD, J.A., Burkhardt, J.K., 2013. Controversy and consensus regarding myosin II function at the immunological synapse. *Current Opinion in Immunology* 25, 300–306.
- Hartman, N.C., Nye, J.A., Groves, J.T., 2009. Cluster size regulates protein sorting in the immunological synapse. *Proceedings of the National Academy of Sciences of the United States of America* 106, 12729–12734.
- Henrickson, S.E., Mempel, T.R., Mazo, I.B., *et al.*, 2008. T cell sensing of antigen dose governs interactive behavior with dendritic cells and sets a threshold for T cell activation. *Nature Immunology* 9, 282–291.
- Henrickson, S.E., Perro, M., Loughhead, S.M., *et al.*, 2013. Antigen availability determines CD8(+) T cell-dendritic cell interaction kinetics and memory fate decisions. *Immunity* 39, 496–507.
- Hu, J., Lipowsky, R., Weikl, T.R., 2013. Binding constants of membrane-anchored receptors and ligands depend strongly on the nanoscale roughness of membranes. *Proceedings of the National Academy of Sciences of the United States of America* 110, 15283–15288.
- Huang, J., Brameshuber, M., Zeng, X., *et al.*, 2013. A single peptide-major histocompatibility complex ligand triggers digital cytokine secretion in CD4(+) T cells. *Immunity* 39, 846–857.
- Huppa, J.B., Axmann, M., Mortelmaier, M.A., *et al.*, 2010. TCR-peptide-MHC interactions *in situ* show accelerated kinetics and increased affinity. *Nature* 463, 963–967.
- Jansson, A., Barnes, E., Klenerman, P., *et al.*, 2005. A theoretical framework for quantitative analysis of the molecular basis of costimulation. *Journal of Immunology* 175, 1575–1585.
- Johnson, K.G., Bromley, S.K., Dustin, M.L., Thomas, M.L., 2000. A supramolecular basis for CD45 tyrosine phosphatase regulation in sustained T cell activation. *Proceedings of the National Academy of Sciences of the United States of America* 97, 10138–10143.
- Kaizuka, Y., Douglass, A.D., Varma, R., Dustin, M.L., Vale, R.D., 2007. Mechanisms for segregating T cell receptor and adhesion molecules during immunological synapse formation in Jurkat T cells. *Proceedings of the National Academy of Sciences of the United States of America* 104, 20296–20301.
- Kong, K.F., Fu, G., Zhang, Y., *et al.*, 2014. Protein kinase C- ϵ controls CTLA-4-mediated regulatory T cell function. *Nature Immunology* 15, 465–474.
- Kuhn, J.R., Poenie, M., 2002. Dynamic polarization of the microtubule cytoskeleton during CTL-mediated killing. *Immunity* 16, 111–121.
- Kumar, R., Perez, M., Swamy, M., *et al.*, 2011. Increased sensitivity of antigen-experienced T cells through the enrichment of oligomeric T cell receptor complexes. *Immunity* 35, 375–387.
- Kupfer, A., Singer, S.J., 1989. The specific interaction of helper T cells and antigen-presenting B cells. IV. Membrane and cytoskeletal reorganizations in the bound T cell as a function of antigen dose. *Journal of Experimental Medicine* 170, 1697–1713.
- Labno, C.M., Lewis, C.M., You, D., *et al.*, 2003. Itk functions to control actin polymerization at the immune synapse through localized activation of Cdc42 and WASP. *Current Biology* 13, 1619–1624.
- Larghi, P., Williamson, D.J., Carpi, J.M., *et al.*, 2013. VAMP7 controls T cell activation by regulating the recruitment and phosphorylation of vesicular Lat at TCR-activation sites. *Nature Immunology* 14, 723–731.
- Lee, K.H., Dinner, A.R., Tu, C., *et al.*, 2003. The immunological synapse balances T cell receptor signaling and degradation (see comment). *Science* 302, 1218–1222.
- Lee, K.H., Holdorf, A.D., Dustin, M.L., *et al.*, 2002. T cell receptor signaling precedes immunological synapse formation. *Science* 295, 1539–1542.
- Leupin, O., Zaru, R., Laroche, T., Muller, S., Valitutti, S., 2000. Exclusion of CD45 from the T-cell receptor signaling area in antigen-stimulated T lymphocytes. *Current Biology* 10, 277–280.
- Li, P., Banjade, S., Cheng, H.C., *et al.*, 2012. Phase transitions in the assembly of multivalent signalling proteins. *Nature* 483, 336–340.
- Lillemeier, B.F., Mortelmaier, M.A., Forstner, M.B., *et al.*, 2010. TCR and Lat are expressed on separate protein islands on T cell membranes and concatenate during activation. *Nature Immunology* 11, 90–96.
- Lillemeier, B.F., Pfeiffer, J.R., Surviladze, Z., Wilson, B.S., Davis, M.M., 2006. Plasma membrane-associated proteins are clustered into islands attached to the cytoskeleton. *Proceedings of the National Academy of Sciences of the United States of America* 103, 18992–18997.
- Lin, J., Weiss, A., 2003. The tyrosine phosphatase CD148 is excluded from the immunologic synapse and down-regulates prolonged T cell signaling. *Journal of Cell Biology* 162, 673–682.
- Linsley, P.S., Bradshaw, J., Greene, J., *et al.*, 1996. Intracellular trafficking of CTLA-4 and focal localization towards sites of TCR engagement. *Immunity* 4, 535–543.
- Liu, H., Rhodes, M., Wiest, D.L., Vignali, D.A., 2000. On the dynamics of TCR:CD3 complex cell surface expression and downmodulation. *Immunity* 13, 665–675.
- Lowin-Kropf, B., Shapiro, V.S., Weiss, A., 1998. Cytoskeletal polarization of T cells is regulated by an immunoreceptor tyrosine-based activation motif-dependent mechanism. *Journal of Cell Biology* 140, 861–871.
- Ludford-Menting, M.J., Oliaro, J., Sacirbegovic, F., *et al.*, 2005. A network of PDZ-containing proteins regulates T cell polarity and morphology during migration and immunological synapse formation. *Immunity* 22, 737–748.
- McGavern, D.B., Christen, U., Oldstone, M.B., 2002. Molecular anatomy of antigen-specific CD8(+) T cell engagement and synapse formation *in vivo*. *Nature Immunology* 3, 918–925.
- Melichar, H.J., Ross, J.O., Herzmark, P., Hogquist, K.A., Robey, E.A., 2013. Distinct temporal patterns of T cell receptor signaling during positive versus negative selection *in situ*. *Science Signaling* 6, ra92.

- Mempel, T.R., Henrickson, S.E., von Andrian, U.H., 2004. T-cell priming by dendritic cells in lymph nodes occurs in three distinct phases. *Nature* 427, 154–159.
- Meyers, J., Craig, J., Odde, D.J., 2006. Potential for control of signaling pathways via cell size and shape. *Current Biology* 16, 1685–1693.
- Monks, C.R., Freiberg, B.A., Kupfer, H., Sciaky, N., Kupfer, A., 1998. Three-dimensional segregation of supramolecular activation clusters in T cells. *Nature* 395, 82–86.
- Monks, C.R., Kupfer, H., Tamir, I., Barlow, A., Kupfer, A., 1997. Selective modulation of protein kinase C- θ during T-cell activation. *Nature* 385, 83–86.
- Moreau, H.D., Lemaître, F., Terriac, E., *et al.*, 2012. Dynamic in situ cytometry uncovers T cell receptor signaling during immunological synapses and kinapses *in vivo*. *Immunity* 37, 351–363.
- Mugnier, B., Nal, B., Verthuy, C., *et al.*, 2008. Coronin-1A links cytoskeleton dynamics to TCR alpha beta-induced cell signaling. *PLoS One* 3, e3467.
- Neves, S.R., Tsokas, P., Sarkar, A., *et al.*, 2008. Cell shape and negative links in regulatory motifs together control spatial information flow in signaling networks. *Cell* 133, 666–680.
- Nolz, J.C., Gomez, T.S., Zhu, P., *et al.*, 2006. The WAVE2 complex regulates actin cytoskeletal reorganization and CRAC-mediated calcium entry during T cell activation. *Current Biology* 16, 24–34.
- Okada, T., Miller, M.J., Parker, I., *et al.*, 2005. Antigen-engaged B cells undergo chemotaxis toward the T zone and form motile conjugates with helper T cells. *PLoS Biology* 3, e150.
- Purbhoo, M.A., Irvine, D.J., Huppa, J.B., Davis, M.M., 2004. T cell killing does not require the formation of a stable mature immunological synapse. *Nature Immunology* 5, 524–530.
- Purbhoo, M.A., Liu, H., Oddo, S., *et al.*, 2010. Dynamics of subsynaptic vesicles and surface microclusters at the immunological synapse. *Science Signaling* 3, ra36.
- Purtic, B., Pitcher, L.A., van Oers, N.S., Wülfing, C., 2005. T cell receptor (TCR) clustering in the immunological synapse integrates TCR and costimulatory signaling in selected T cells. *Proceedings of the National Academy of Sciences of the United States of America* 102, 2904–2909.
- Quann, E.J., Liu, X., Altan-Bonnet, G., Huse, M., 2011. A cascade of protein kinase C isozymes promotes cytoskeletal polarization in T cells. *Nature Immunology* 12, 647–654.
- Quintana, A., Schwindling, C., Wenning, A.S., *et al.*, 2007. T cell activation requires mitochondrial translocation to the immunological synapse. *Proceedings of the National Academy of Sciences of the United States of America* 104, 14418–14423.
- Rak, G.D., Mace, E.M., Banerjee, P.P., Svitkina, T., Orange, J.S., 2011. Natural killer cell lytic granule secretion occurs through a pervasive actin network at the immune synapse. *PLoS Biology* 9, e1001151.
- Rangamani, P., Lipshtat, A., Azeloglu, E.U., *et al.*, 2013. Decoding information in cell shape. *Cell* 154, 1356–1369.
- Richie, L.I., Ebert, P.J., Wu, L.C., *et al.*, 2002. Imaging synapse formation during thymocyte selection: Inability of CD3zeta to form a stable central accumulation during negative selection. *Immunity* 16, 595–606.
- Rocha-Perugini, V., Zamai, M., Gonzalez-Granado, J.M., *et al.*, 2013. CD81 controls sustained T cell activation signaling and defines the maturation stages of cognate immunological synapses. *Molecular and Cellular Biology* 33, 3644–3658.
- Roumier, A., Olivo-Marin, J.C., Arpin, M., *et al.*, 2001. The membrane-microfilament linker ezrin is involved in the formation of the immunological synapse and in T cell activation. *Immunity* 15, 715–728.
- Roybal, K.T., Sinai, P., Verkade, P., Murphy, R.F., Wülfing, C., 2013. The actin-driven spatiotemporal organization of T-cell signaling at the system scale. *Immunological Reviews* 256, 133–147.
- Sage, P.T., Varghese, L.M., Martinelli, R., *et al.*, 2012. Antigen recognition is facilitated by invadosome-like protrusions formed by memory/effector T cells. *Journal of Immunology* 188, 3686–3699.
- Savage, N.D., Kimzey, S.L., Bromley, S.K., *et al.*, 2002. Polar redistribution of the sialoglycoprotein CD43: Implications for T cell function. *Journal of Immunology* 168, 3740–3746.
- Schamel, W.W., Arechaga, I., Risueno, R.M., *et al.*, 2005. Coexistence of multivalent and monovalent TCRs explains high sensitivity and wide range of response. *Journal of Experimental Medicine* 202, 493–503.
- Schmick, M., Bastiaens, P.J., 2014. The interdependence of membrane shape and cellular signal processing. *Cell* 156, 1132–1138.
- Schubert, D.A., Gordo, S., Sabatino JR, J.J., *et al.*, 2012. Self-reactive human CD4 T cell clones form unusual immunological synapses. *Journal of Experimental Medicine* 209, 335–352.
- Sherman, E., Barr, V., Manley, S., *et al.*, 2011. Functional nanoscale organization of signaling molecules downstream of the T cell antigen receptor. *Immunity* 35, 705–720.
- Singleton, K.L., Gosh, M., Dandekar, R.D., *et al.*, 2011. Itk controls the spatiotemporal organization of T cell activation. *Science Signaling* 4, ra66.
- Singleton, K.L., Parvaze, N., Dama, K.R., *et al.*, 2006. A large T cell invagination with CD2 enrichment resets receptor engagement in the immunological synapse. *Journal of Immunology* 177, 4402–4413.
- Singleton, K.L., Roybal, K.T., Sun, Y., *et al.*, 2009. Spatiotemporal patterning during T cell activation is highly diverse. *Science Signaling* 2, ra15.
- Skokos, D., Shakhar, G., Varma, R., *et al.*, 2007. Peptide-MHC potency governs dynamic interactions between T cells and dendritic cells in lymph nodes. *Nature Immunology* 8, 835–844.
- Snapper, S.B., Rosen, F.S., Mizoguchi, E., *et al.*, 1998. Wiskott-Aldrich syndrome protein-deficient mice reveal a role for WASP in T but not B cell activation. *Immunity* 9, 81–91.
- Soares, H., Henriques, R., Sachse, M., *et al.*, 2013. Regulated vesicle fusion generates signaling nanoterritories that control T cell activation at the immunological synapse. *Journal of Experimental Medicine* 210, 2415–2433.
- Stinchcombe, J.C., Bossi, G., Booth, S., Griffiths, G.M., 2001. The immunological synapse of CTL contains a secretory domain and membrane bridges. *Immunity* 15, 751–761.
- Tskvartaria-Fuller, I., Rozelle, A.L., Yin, H.L., Wülfing, C., 2003. Regulation of sustained actin dynamics by the TCR and costimulation as a mechanism of receptor localization. *Journal of Immunology* 171, 2287–2295.
- Ueda, H., Mophew, M.K., McIntosh, J.R., Davis, M.M., 2011. CD4+ T-cell synapses involve multiple distinct stages. *Proceedings of the National Academy of Sciences of the United States of America* 108, 17099–17104.
- Vardhana, S., Choudhuri, K., Varma, R., Dustin, M.L., 2010. Essential role of ubiquitin and TSG101 protein in formation and function of the central supramolecular activation cluster. *Immunity* 32, 531–540.
- Varma, R., Campi, G., Yokosuka, T., Saito, T., Dustin, M.L., 2006. T cell receptor-proximal signals are sustained in peripheral microclusters and terminated in the central supramolecular activation cluster. *Immunity* 25, 117–127.
- Wild, M.K., Cambiaggi, A., Brown, M.H., *et al.*, 1999. Dependence of T cell antigen recognition on the dimensions of an accessory receptor-ligand complex. *Journal of Experimental Medicine* 190, 31–41.
- Williamson, D.J., Owen, D.M., Rossy, J., *et al.*, 2011. Pre-existing clusters of the adaptor Lat do not participate in early T cell signaling events. *Nature Immunology* 12, 655–662.
- Wilson, B.S., Pfeiffer, J.R., Surviladze, Z., Gaudet, E.A., Oliver, J.M., 2001. High resolution mapping of mast cell membranes reveals primary and secondary domains of Fc(epsilon)RI and LAT. *Journal of Cell Biology* 154, 645–658.
- Witt, C.M., Raychaudhuri, S., Schaefer, B., Chakraborty, A.K., Robey, E.A., 2005. Directed migration of positively selected thymocytes visualized in real time. *PLoS Biology* 3, e160.
- Woo, S.R., Li, N., Bruno, T.C., *et al.*, 2010. Differential subcellular localization of the regulatory T-cell protein LAG-3 and the coreceptor CD4. *European Journal of Immunology* 40, 1768–1777.
- Wülfing, C., Davis, M.M., 1998. A receptor/cytoskeletal movement triggered by costimulation during T cell activation. *Science* 282, 2266–2270.
- Xie, J., Huppa, J.B., Newell, E.W., *et al.*, 2012. Photocrosslinkable pMHC monomers stain T cells specifically and cause ligand-bound TCRs to be 'preferentially' transported to the cSMAC. *Nature Immunology* 13, 674–680.
- Yi, J., Wu, X., Chung, A.H., *et al.*, 2013. Centrosome repositioning in T cells is biphasic and driven by microtubule end-on capture-shrinkage. *Journal of Cell Biology* 202, 779–792.
- Yi, J., Wu, X.S., Crites, T., Hammer 3RD, J.A., 2012. Actin retrograde flow and actomyosin II arc contraction drive receptor cluster dynamics at the immunological synapse in Jurkat T cells. *Molecular Biology of the Cell* 23, 834–852.
- Yokosuka, T., Sakata-Sogawa, K., Kobayashi, W., *et al.*, 2005. Newly generated T cell receptor microclusters initiate and sustain T cell activation by recruitment of Zap70 and SLP-76. *Nature Immunology* 6, 1253–1262.
- Yonemura, S., Hirao, M., Doi, Y., *et al.*, 1998. Ezrin/radixin/moesin (ERM) proteins bind to a positively charged amino acid cluster in the juxta-membrane cytoplasmic domain of CD44, CD43, and ICAM-2. *Journal of Cell Biology* 140, 885–895.
- Yudushkin, I.A., Vale, R.D., 2010. Imaging T-cell receptor activation reveals accumulation of tyrosine-phosphorylated CD3zeta in the endosomal compartment. *Proceedings of the National Academy of Sciences of the United States of America* 107, 22128–22133.

- Zanin-Zhorov, A., Ding, Y., Kumari, S., *et al.*, 2010. Protein kinase C- θ mediates negative feedback on regulatory T cell function. *Science* 328, 372–376.
- Zheng, H., Jin, B., Henrickson, S.E., *et al.*, 2008. How antigen quantity and quality determine T-cell decisions in lymphoid tissue. *Molecular and Cellular Biology* 28, 4040–4051.
- Zipfel, P.A., Bunnell, S.C., Witherow, D.S., *et al.*, 2006. Role for the Abi/wave protein complex in T cell receptor-mediated proliferation and cytoskeletal remodeling. *Current Biology* 16, 35–46.

Further Reading

- Batista, F.D., Dustin, M.L. (Eds.), 2013. *Immunological Reviews* (Special issue — Cell: Cell Contacts), vol. 251, issue 1. Hoboken, NJ: John Wiley and sons Ltd.

- Burkhardt, J.K. (Ed.), 2013. *Immunological Reviews* (Special issue — The cytoskeleton), vol. 256, issue 1. Hoboken, NJ: John Wiley and sons Ltd.

Relevant Websites

- <http://mmbios.org/index.php/software#cellorganizer>
MMBioS.
- <http://www.niaid.nih.gov/LabsAndResources/labs/aboutlabs/lstb/Pages/simimmuneproject.aspx>
National Institutes of Health.
- <http://www.bristol.ac.uk/cellmolmed/research/infect-immune/wuelfing/spatiotemporal-patterning/>
University of Bristol.

CELLULAR IMMUNOLOGY: INNATE IMMUNITY

Contents

Function of Epithelial Barriers

Phagocytic Synapses

The Phagocyte NADPH Oxidase: Structure and Assembly of the Key Multicomponent Enzyme of Innate Immunity

Cytoplasmic Sensing in Innate Immunity

Scavenger Receptors

Dendritic Cells

Neutrophil Biology

Specialized Subsets of Tissue-Resident Macrophages in Secondary Lymphoid Organs

CD4 T Cell Memory and Role of TNF Receptor Family

Function of Epithelial Barriers

K Cadwell and AM Marchiando, New York University School of Medicine, New York, NY, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Apical region The side of a cell that faces either the external environment or the lumen.

Basolateral region The side of a cell in contact with the underlying tissue.

Epithelium Animal tissue comprised of a network of cells that line the cavities and surfaces of an animal.

Lumen Interior side of a tubular structure.

Microvilli A membrane protrusion on the apical surface of an epithelial cell that increases surface area.

Simple epithelium Epithelium consisting of a single layer of cells found where secretion and absorption occur.

Stratified epithelium Epithelium consisting of two or more layers of cells found where abrasion may occur.

Introduction

Compartmentalization is a fundamental feature of life. One of the first lessons of biology is that we are an aggregate of cells with lipid bilayer membranes, and that these cells themselves contain membrane-bound compartments termed organelles. In animals, compartmentalization at the whole organism level is achieved by the epithelium, which refers to the network of cells that line the cavities and surfaces of our body. In addition to the skin that serves as the boundary between our internal environment and the outside world, an epithelial layer lines the digestive tract, lungs, kidneys, urinary tract, and other organs to maintain separation of their contents. However, the function of the epithelium is not limited to that of a static physical barrier. Like the membranes that surround cells and organelles, the epithelium mediates the exchange of chemicals and other materials between the internal and external environment. Epithelia are also responsible for sensing the external environment as well as serving as a primary defense mechanism. Consequently, the cells that make up the epithelium are as diverse in their appearance and function as the organs in which they are found.

This article will present an overview of the variety and function of epithelia, and use examples to illustrate the dynamic nature of this major type of animal tissue. In the first section, we will discuss structural features of the epithelium. The second section will examine how individual epithelial cells reflect the specialized function of this tissue and the associated organs. Many features of the epithelium reflect the tremendous evolutionary pressure conferred by infectious disease. Therefore, in the third section we will use the role of epithelia in immunity to highlight the complexity of barriers. Finally, we will discuss the epithelium in the context of disease and discuss the recent advances in research in this area. The endothelium, the layer of cells that line the inside of blood and lymphatic vessels, is frequently included in discussion of the epithelium and also serves as a barrier. Although many of the same principles apply, the endothelium has its own unique properties and will not be discussed further. In many parts of this article, we will use the intestinal epithelium as an example. Many of the recent advances in epithelial biology highlight the extraordinary mechanisms utilized by the intestinal epithelium to maintain coexistence with the vast amount of bacteria present in close proximity. By discussing the key

cell biological features of epithelia, we hope that the reader will appreciate that the epithelial barrier is every bit as dynamic as the underlying organs and cavities they surround.

Structure of the Epithelial Layer

Epithelia are derived from all three primary germ layers, and cells derived from each layer establish distinct compartments. For example, the external skin is derived from the ectoderm, while the endoderm forms the linings of the gastrointestinal and respiratory tracts, and the mesoderm gives rise to renal epithelium. The structure of the epithelium reflects their location and function. The gastrointestinal tract, lungs, and kidneys are lined with a simple epithelia, which consists of a single layer of cells, while the outer epithelium, the skin, is a stratified epithelium that is several layers thick (Alberts *et al.*, 2008). Simple epithelia are found where secretion and absorption or passive diffusion occur, and frequently have cilia or microvilli, a membrane protrusion that increases surface area. In contrast, stratified epithelium are thicker, consisting of two or more layers of cells, and usually have a protective role.

The simple and stratified epithelium can be further divided into squamous, cuboidal, and columnar based on the shape of the cells that comprise the layer (Bryant and Mostov, 2008). Squamous epithelial cells are flattened cells with irregular scale-like shapes. Simple squamous epithelia are found lining the cavities of the body including the pericardial, pleural, and peritoneal cavities, or in areas where passive diffusion occurs, such as glomeruli in the kidney and alveoli in the respiratory tract. Stratified squamous epithelium is comprised of several layers of epithelial cells, the bottom of which is in contact with a basement membrane that underlies the epithelium. Above this layer are additional layers of cells of different shapes and number, with the top layer being a flattened squamous layer of epithelium. This type of epithelium is found in areas of the body subject to frequent abrasion, as underlying tissues can remain protected even when the surface layer of cells is lost, and are found lining the anal canal, cervix and vagina, pharynx, and esophagus. As their names imply, cuboidal epithelial cells are roughly shaped like cubes, and columnar epithelial cells are taller than wider. Simple cuboidal epithelial cells are located in the ovaries and parts of the kidney including nephrons and renal tubules. Stratified cuboidal epithelial cells are less common and are found lining salivary, mammary, and sweat glands. Large portions of the gastrointestinal tract are lined with simple columnar epithelia. While comprised of only a single layer of cells, parts of the upper respiratory tract are lined with pseudostratified epithelium, which at first glance appears to have more than one layer due to an uneven positioning of nuclei in neighboring cells. Stratified columnar epithelia are less common and found in reproductive organs such as the uterus and vas deferens, as well as parts of the gastrointestinal tract like the pharynx and anus.

These structural differences reflect the local environment and organ function. The protective barrier provided by the stratified epithelium of the skin is essential for preventing contamination of the underlying tissue from noxious chemical and infectious material present in the external environment

(Tsukita and Furuse, 2002). In stark contrast to the multi-layered epithelium of the skin, the single layer of cells of the alveoli allows for exchange of carbon dioxide and oxygen by diffusion (Hasenberg *et al.*, 2013). The simple epithelium of the renal tubule both forms a barrier between the renal interstitium and the sterile tubular interior (lumen), and maintains a gradient that allows for regulated nutrient and water transport and urine production (Field *et al.*, 1989). The structural integrity of these diverse epithelia relies on the polarization of the individual cells. Epithelial cells have an apical region, which faces either the lumen or the external environment in the case of the skin, and a basolateral region in contact with the underlying tissue (Bryant and Mostov, 2008). The tight junction (discussed below) and everything above is considered the apical portion of the cell, and everything below is the basolateral portion (Baum and Georgiou, 2011). Lateral membranes, where the junctional complexes are located, exist between neighboring epithelial cells. The apical and basolateral surfaces of epithelial cells have different features, such as microvilli on the apical surface of many epithelial cells. Also, the distribution of proteins, intracellular organelles, and cytoskeletal components can differ significantly between the two regions (Bryant and Mostov, 2008). Polarity is achieved with assistance from the cytoskeleton, which serves as a structural framework for shape and supports membrane traffic. Rather than creating an inflexible structure, the cytoskeleton allows dramatic remodeling when necessary. For example, contraction of the actomyosin circumferential belt has been shown to be a mechanism by which apoptotic cells in the intestinal epithelium are extruded from the monolayer without compromising the barrier (Rosenblatt *et al.*, 2001).

The ability of epithelia to maintain polarity and create uninterrupted boundaries is made possible by the junctions between cells. While good for providing a strong barrier, impermeant junctions would not allow diffusion or transport across the epithelium. Thus a highly regulated, selectively permeable barrier exists in simple epithelia such as the one found in the intestine. Intercellular junctions seal the space between neighboring epithelial cells, preventing unregulated movement across the epithelium between cells (Turner, 2006). The three main junctional complexes between epithelial cells are the tight junction, the adherens junction, and desmosomes. The tight junction is the most apical of these complexes, and though it is the primary determinant of epithelial barrier function, it does not provide much strength to the intercellular bond (Pappenheimer and Reiss, 1987). Proteins that comprise the tight junction include transmembrane proteins such as occludin, marvelD3, and many claudin family members that interact with peripheral membrane proteins in the cytoplasm including ZO-1, -2, and -3 and link to the actin cytoskeleton (Mitic and Anderson, 1998). Tight junction proteins must maintain an intact barrier during physiological events including epithelial cell shedding, apoptosis, and pathophysiological events-like small injuries (Rosenblatt *et al.*, 2001; Bement *et al.*, 1993). Below the tight junction complex is the adherens junction. The adherens junctions do not contribute to the barrier function of the epithelium, but are instead a structural component of the junctional complex, providing strength needed to hold the neighboring epithelial cells together to form a continuous sheet (Baum and

Georgiou, 2011). Proteins of the adherens junction are predominately the E-cadherin adhesion molecules and the cytoplasmically associated catenins (Hermiston and Gordon, 1995). Catenins interact with F-actin and other components of the actin cytoskeleton to form and stabilize puncta. In addition to conferring strength, these junctions are important for epithelial cell differentiation and polarization. On the lateral membrane of the epithelial cells and basal to the adherens junction are the desmosomes, which connect to keratin intermediate filaments and provide further mechanical stabilization of the epithelial sheet (Turner, 2006). The transmembrane proteins that comprise the desmosomes are cell adhesion molecules, primarily from the cadherin family. The basal interaction of the epithelial cells with the basement membrane (an extracellular matrix sheet) is mediated by integrin family adhesion molecules that link to the actin and intermediate filament (hemidesmosomes) cytoskeletons. Together, these junctional complexes maintain a mechanically robust and chemically selective barrier.

Epithelial Cell Types

The individual cells that form epithelia are incredibly diverse in morphology and function. The epithelial cells of the gastrointestinal tract exemplify this diversity. The fundus of the stomach contains specialized epithelial cells that secrete hydrochloric acid and intrinsic factor, a glycoprotein necessary for absorption of vitamin B¹² (Schubert, 2011). The resulting acidic environment causes denaturation of ingested proteins, and together with chief cells that release enzymes, mediates proteolysis and ultimately absorption of nutrients. In the small intestine, finger-like projections called villi are covered by a monolayer of epithelial cells that maximize the surface area available for absorption (Bjerknes and Cheng, 2006). At the base of each villus is an invagination into the wall of the intestine, the crypt of Lieberkuhn. The epithelium of the colon does not contain villi, but is similar in that a monolayer of epithelial cells cover the crypts that emanate from the gut wall. Though the intestinal epithelium is made up of numerous cell types designed for different functions, these cell types have a common origin. Intestinal epithelial cells arise from Lgr5-expressing columnar stem cells located within the crypts (Barker *et al.*, 2007). Epithelial cells undergo three or four cell divisions and migrate away from the crypt base toward the villi in the small intestine or apical surface of the crypts in the colon. Cells that reach the tip of the villus or crypt are continuously shed during this process of renewal, resulting in turn over of the entire surface of the intestinal epithelium every 5–7 days (Bullen *et al.*, 2006).

Due to recent technical advances, signaling events that allow differentiation of functionally diverse intestinal epithelial cells are becoming clear. In addition to the stem cells, the base of the crypt is also home to Paneth cells, specialized cells of the small intestine which do not migrate up the villus but instead are longer lived residents of the crypts (Ouellette, 2010). Formation of Paneth cells relies on Wnt signaling, including the transcription factor Sox9 (Mori-Akiyama *et al.*, 2007). These secretory cells have two main functions, to provide growth factors important for maintaining neighboring

stem cells, and to aid in barrier maintenance and defense against bacteria in the intestinal lumen by secretion of antimicrobial factors (Ouellette, 2010). Another secretory cell important for maintaining homeostasis in the intestine is the goblet cell (Figure 1). These cells arise in the absence of Notch signaling, and rely on the transcription factors Spdef and Klf4 for their differentiation (Noah *et al.*, 2011; Katz *et al.*, 2002). Goblet cells are simple columnar epithelial cells found interspersed among other epithelial cell types in the gastrointestinal and respiratory tracts. Goblet cells secrete mucin, the main component of the mucus layer which coats the apical surface of the epithelium and provides protection against microorganisms and cellular damage from luminal contents (Noah *et al.*, 2011). A third secretory cell found scattered among other epithelial cell types in the gastrointestinal tract are the enteroendocrine cells. The transcription factor neurogenin3 is required for differentiation of these cells, which in response to various stimuli can secrete hormones that regulate digestion into the local environment or into the bloodstream (Jenny *et al.*, 2002). The hormones produced define the subtypes, including K-cells and L-cells in the small intestine, and G-cells in the stomach (May and Kaestner, 2010). Another rare cell type located in the respiratory and intestinal epithelium are the tuft cells, which are characterized by thick, long projections of microvilli from the apical surface and corresponding dense arrays of apical actin filaments. Beyond the transcription factor Gfi1b, few cell-specific markers have been identified (Bjerknes *et al.*, 2012). Assignment of tuft cells to the secretory lineage

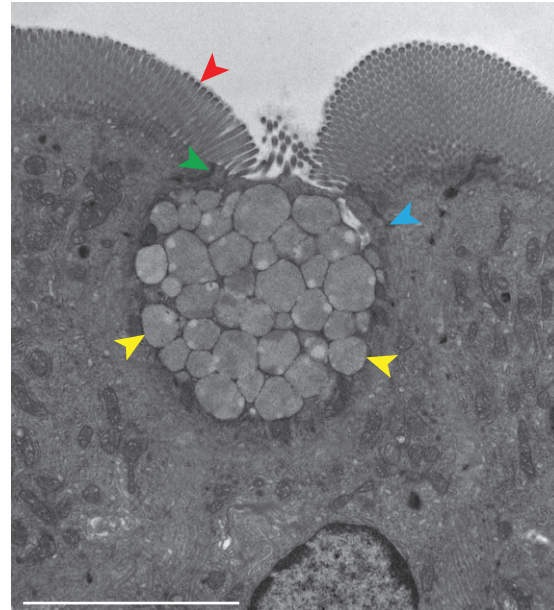


Figure 1 Transmission electron micrograph of a goblet cell in the small intestine. Goblet cells contain granules near the apical surface that contain mucin (yellow arrow heads), which are secreted into the intestinal lumen to generate the protective mucus layer. The tight junctions (green arrow head) and desmosomes (blue arrow head) of the apical junctional complex can be seen between the goblet cell and neighboring enterocytes, the absorptive epithelial cell type of the intestine. The enterocytes display prominent microvilli (red arrow head) on the apical surface. Scale bar = 5 μ m.

remains controversial, as development of tuft cells is independent of transcription factors essential for other secretory cell types, such as *Neurog3*, *Sox9*, and *Spdef* (Gerbe *et al.*, 2012). Tuft cells are potentially chemosensory cells, but their role remains largely undefined (Gerbe *et al.*, 2012).

In addition to secretory cells, other cell types can be found in the gastrointestinal tract including M cells. Microfolds, rather than microvilli, cover the apical surface of these cells and give them their name. Differentiation of these cells is dependent on the cytokine RankL and the transcription factor SpiB (De Lau *et al.*, 2011). M cells are found in specialized sections of the intestinal epithelium, the Peyer's patches, which are discrete collections of epithelial cells located above the gut-associated lymphoid tissue (GALT) (Randall *et al.*, 2008). The M cells within the Peyer's patches are critically involved in immunity in that they are an interface between the luminal contents and the tissue-associated immune cells.

The final and most prevalent epithelial cell type in the small intestine and colon is the enterocyte. As these simple columnar cells migrate out of the crypt base, they maintain active Notch signaling and repress the transcription factor Math1 (Barker *et al.*, 2007). In this way, they avoid differentiation into a secretory lineage and become absorptive cells. Microvilli cover the apical surface of enterocytes (Figure 1), greatly increasing the surface area by which these cells can transport and absorb nutrients from the lumen. Enterocytes are responsible for the uptake of ions, lipids, peptides, sugars, and water and also maintain a physical barrier between the microorganisms in the intestinal lumen and the underlying tissue.

Defense

A basic defense mechanism is to block entry of the pathogen into the body, thereby preventing the establishment of an infection. As discussed earlier, the skin is a formidable wall composed of multiple layers of cells packed closely together. The importance of this defense strategy is illustrated by the dire consequences that follow a breach in this barrier. Prior to the development of the tetanus vaccine, the release of the toxin produced by *Clostridium tetani* into the bloodstream was a major cause of lethality following wounding (Bizzini, 1979). Insect bites represent another common type of breach in the skin. Diseases such as malaria and dengue fever are caused by parasites and viruses carried by mosquitoes and account for a staggering number of deaths each year (Murray *et al.*, 2012; Guzman *et al.*, 2010). Still, the skin blocks the entry of many infectious agents into our body and prevents the rest of the immune system from being overwhelmed. Simple epithelium must also engage in defense. In the lungs, the ciliated cells of the respiratory epithelium work together with goblet cells as part of the mucociliary escalator (Lillehoj *et al.*, 2013). The cilia beat in concert to push pathogens and harmful particles trapped in the mucus out of the airway. The mucus-trapped material is either expectorated or swallowed into the stomach where the low pH can kill pathogens. Thus, the parietal cells of the stomach that maintain this low pH can also be considered part of the epithelial defense strategy.

These defense mechanisms are rudimentary in comparison to the chemical warfare that is prominent in the lower gastrointestinal tract. The intestinal epithelium produces or transports an immense amount of mucus, antibodies (immunoglobulins), and antimicrobial peptides – approximately 3 g of immunoglobulin A (IgA) is secreted into the intestinal lumen daily and accounts for the majority of immunoglobulins in the human body (Hooper, 2009; Fagarasan and Honjo, 2003). All epithelium engage in this type of chemical warfare to a certain extent, but to understand why the intestine in particular relies heavily on these long-range weapons requires discussion of its unique vulnerabilities. As mentioned earlier, the intestinal epithelium is one-cell-thick to accommodate nutrient acquisition, and the enterocytes turn over rapidly. This thin epithelial layer is in constant proximity to a large number of an assortment of microbes in the intestinal lumen. The human body is home to 10^{14} or greater commensal bacteria, the majority of which is found in the distal parts of the intestine (Honda and Littman, 2012). While collectively referred to as commensals, these microbes are heterogeneous and represent >1000 species. The gram-positive anaerobic *Lactobacillus* and *Bifidobacterium* species are well known for assisting with digestion of carbohydrates and producing vitamins that cannot be synthesized by human cells (LeBlanc *et al.*, 2013). The gram-negative *Bacteroides* genus and *Escherichia coli* are also ubiquitous and provide metabolites for the host, but certain strains express virulence factors and are associated with inflammatory bowel disease (IBD) and colon cancer (Packey and Sartor, 2009; Wu *et al.*, 2009; Arthur *et al.*, 2012). The intestinal epithelium is also exposed to a continual flow of foreign material that can include pathogenic bacteria such as *Salmonella enterica* and *Listeria monocytogenes*. These and many other infectious agents enter the digestive tract when we consume food and water, and cause diarrhea as well as more severe complications upon dissemination to other organs (Preidis *et al.*, 2011). Therefore, the intestinal epithelium is a barrier between the rest of our body and a spectrum of microbes that can be transient or stable, and range from pathogenic to innocuous, or even beneficial.

Despite the obvious danger that penetration of this barrier represents, the intestinal immune system cannot indiscriminately remove all bacteria since commensals provide benefits beyond digestion. Commensal bacteria as a whole are necessary for occupying a niche that would otherwise be exposed to pathogenic bacteria, a phenomenon referred to as colonization resistance. This function of commensals is exemplified by *Clostridium difficile* infection. *Clostridium difficile* is a gram-positive spore-forming bacterium that can be found in the intestines of healthy individuals. However, treatment with antibiotics, particularly during hospital visits, can remove bacteria in the colon and create opportunities for *C. difficile* to overgrow (Pacheco and Johnson, 2013). *Clostridium difficile* produces enterotoxins that cause diarrhea and pseudomembranous colitis, and is the leading cause of death due to gastrointestinal infection in the United States (Hall *et al.*, 2012). Remarkably, the most promising strategy for treating recurrent *C. difficile* infection is fecal transplantation involving the transfer of intestinal bacteria from a healthy donor (Kassam *et al.*, 2013). The efficacy of this intervention provides additional evidence that commensal bacteria are critical for

preventing opportunistic infections. For this and other reasons, the goal of the intestinal epithelium is to maintain a mutually beneficial coexistence with commensal bacteria rather than creating a sterile environment that leaves the host susceptible.

How does the intestinal epithelium fulfill this task while functioning as a barrier for a surface area that is 10 times larger than that of the skin? Two general strategies are apparent – maintaining distance with the bacterial flora and calibrating the immune response. A 50 μm ‘demilitarized zone’ has been described in mice that separates most commensal bacteria from the underlying epithelium (Vaishnava *et al.*, 2011). This separation is achieved in part by the secretion of an antimicrobial molecule termed Reg3 γ by Paneth cells and enterocytes. Also, both cell surface-associated and secreted mucins create the mucus layer of the intestine that extends this distance further by forming a net (Sheng *et al.*, 2012). Mice deficient in Muc2, the predominant mucin produced by intestinal goblet cells, are highly susceptible to intestinal inflammation in a variety of instances due to a breakdown in this extended barrier (Heazlewood *et al.*, 2008; Van Der Sluis *et al.*, 2006; Bergstrom *et al.*, 2010). Epithelial cells also transport antibodies to the lumen that are produced by plasma B cells in the lamina propria, the connective tissue below the epithelium. The polymeric IgA receptor (pIgR) can bind IgA on the basolateral side, undergo endocytosis, and release IgA into the lumen upon transport to the apical surface (Fagarasan and Honjo, 2003). These and many other secreted molecules create a chemical gradient that reduces the density of bacteria near the epithelium.

This strategy of creating distance is complemented by a second strategy, which is to ignore innocuous microbes and act swiftly in the face of a bona fide threat. The epithelium plays an active role in this process as well by helping to maintain a tolerogenic adaptive immune response under steady-state conditions. Adaptive immune responses are generated when lymphocytes (B and T cells) become activated upon recognition of foreign material presented by antigen-presenting cells. M cells actively transport luminal contents to CD103⁺ dendritic cells, a type of antigen-presenting cell that is abundant in the small intestinal lamina propria (Honda and Littman, 2012). At the same time M cells are delivering antigens to dendritic cells in the underlying GALT, enterocytes are absorbing and processing vitamin A precursors (Von Lintig, 2012). CD103⁺ dendritic cells display increased retinal dehydrogenase activity to convert this vitamin A (retinal) into retinoic acid (Agace and Persson, 2012). Retinoic acid, along with TGF β , favors the conversion of lymphocytes into noninflammatory IgA secreting plasma cells and regulatory T cells (Tregs) (Agace and Persson, 2012; Honda and Littman, 2012). Thus, in the absence of inflammatory signals, the system is rigged to recognize commensal bacteria and other foreign material as innocuous agents rather than a threat. There are most likely many additional mechanisms by which epithelial cells induce a tolerogenic state that are yet to be uncovered. For example, recent studies have revealed unexpected properties of goblet cells and mucins in delivering antigens to CD103⁺ dendritic cells and inhibiting inflammatory gene expression (McDole *et al.*, 2012; Shan *et al.*, 2013).

In contrast to this immuno-suppressive function, the intestinal epithelium promotes immune responses by expressing

cell adhesion molecules that attract dendritic cells as well as intra-epithelial lymphocytes (IELs). Unlike traditional T cells, IELs are imbedded in the epithelial layer and poised to respond to threats without a prolonged maturation process (Cheroutre *et al.*, 2011). Also, toxins and virulence factors produced by pathogens can act on the epithelium and trigger inflammatory signals leading to recruitment and differentiation of pro-inflammatory dendritic cells (Honda and Littman, 2012). For instance, bacterial adhesion is a virulence strategy that can directly trigger inflammatory signaling in intestinal epithelial cells (Boyle and Finlay, 2003). Through these various mechanisms, many of which involve the epithelium, the intestine can exist in a paradoxical state where it is generally tolerogenic but simultaneously primed for a swift inflammatory response toward pathogens.

Loss of Barrier Function in Disease

The importance of the epithelial barrier becomes obvious when we observe the consequences of failure during disease conditions. A ubiquitous virulence strategy utilized by food- and water-borne bacterial pathogens is the secretion of toxins that disrupt tight junctions, which allows the microorganism access to the favorable environment on the other side of the barrier (Berkes *et al.*, 2003). These toxins also interfere with ion transporters and directly disrupt the fluid absorbing properties of enterocytes. For instance, the cholera toxin produced by *Vibrio cholerae* causes efflux of chloride ions via phosphorylation of the cystic fibrosis transmembrane conductance regulator (CFTR), subsequently leading to secretion of water, sodium, and potassium (Field *et al.*, 1989). Therefore, enteric infections are associated with diarrhea and dehydration, which remains one of the leading causes of infant mortality (Preidis *et al.*, 2011). As mentioned above, attachment to the intestinal epithelium represents another common virulence strategy that can lead to effacement, disorganization, and hyperproliferation of the epithelium, and requires quick resolution to prevent catastrophic breakdown of the barrier (Figure 2). Some pathogens readily disseminate throughout the host upon ingestion. In the case of typhoid fever, a deadly disease that affects over 16 million people annually, *Salmonella typhi* travels from the intestine through lymph and blood to reach the liver, spleen, and kidneys (Parry *et al.*, 2002). Epithelial barriers other than that of the intestine are also commonly overcome by pathogens. Urinary tract infections can progress to pyelonephritis when bacteria and viruses that infect the bladder, most commonly *E. coli*, spread to one or both kidneys (Ragnarsdottir and Svanborg, 2012). During sexual transmission of human immunodeficiency virus (HIV), the virus is thought to reach its target T cells by hitching a ride on dendritic cells in the mucosa (such as the vaginal surface) that migrate to lymphoid tissues (Wu and KewalRamani, 2006). In developing countries, over 200 million individuals are infected by Schistosome parasites, whose larvae found in freshwater burrow through the skin and mature in various organ tissue to cause a range of symptoms (King *et al.*, 2005).

Mendelian genetic disorders that directly disrupt the barrier are rare, perhaps reflecting the general importance of maintaining

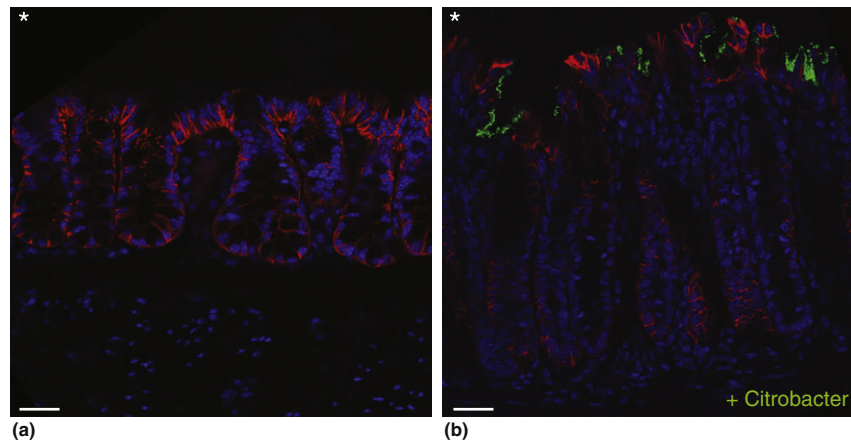


Figure 2 Bacterial adhesion can disrupt the epithelium. A single layer of epithelial cells demarcated by E-cadherin form invaginations referred to as crypts. *Citrobacter rodentium* is a model mouse pathogen related to *E. coli* that attaches to the apical surface of the intestinal epithelium. This adhesion causes an uneven appearance in the surface epithelium due to effacement as well as a hyper-proliferative response that leads to crypt elongation. The left panel shows an immuno-fluorescence microscopy image of colonic crypts from uninfected mice. The right panel shows the dramatic restructuring of the colonic crypts following infection with *C. rodentium* expressing green fluorescence protein (GFP). Red = E-cadherin, blue = nuclei, green = GFP-*C. rodentium*, and asterisks = lumen. Scale bar = 40 μ m.

epithelial barriers. The tight junction complex in the inner ear epithelium is required for maintaining differential ion concentrations between the inner and outer fluids of the cochlea, and thus recessive mutations in *Claudin-14* lead to congenital deafness (Wilcox *et al.*, 2001; Ben-Yosef *et al.*, 2003). Also, *Claudin-16* (*paracellin-1*) mutations lead to familial hypomagnesemia with hypercalciuria and nephrocalcinosis due to defects in paracellular transport in the kidney (Simon *et al.*, 1999). In contrast, IBD is a prominent example of barrier dysfunction involving complex genetic and environmental factors (McGuckin *et al.*, 2009; Khor *et al.*, 2011). IBD, which includes Crohn's disease and ulcerative colitis, is linked to over a hundred polymorphisms that are thought to disrupt the tolerogenic environment in the intestine discussed above (Jostins *et al.*, 2012). Although the involvement of any specific intestinal microbe remains controversial (Packey and Sartor, 2009), there is consensus that a loss of tolerance toward commensals leads to recurring inflammatory destruction of the epithelium. In addition to direct epithelial damage, inflammatory cytokines such as tumor necrosis factor (TNF)- α and interferon (IFN)- γ can increase intestinal permeability by inducing myosin light chain kinase (MLCK) to destabilize tight junctions (McGuckin *et al.*, 2009). Commensal bacteria that are otherwise innocuous may gain access past the epithelium to exacerbate or sustain ongoing inflammation. Additionally, patients frequently display altered levels of Paneth cell defensins and goblet cell mucins (Wehkamp *et al.*, 2005; Sheng *et al.*, 2012). Thus, epithelial dysfunction appears to be a key feature of IBD, although some of these abnormalities may be secondary to defects in lymphocytes and other white blood cells.

Perspective

Major research efforts in the field of epithelial biology include examination of how epithelia in various organs differ in their developmental program, interactions with microbes, and contributions to disease. These efforts have been greatly

facilitated by technological innovations such as organ cultures, high-throughput sequencing, increased computational capacities to analyze complex data, and genetic models that allow cell type-specific manipulations. The application of these techniques to IBD has certainly improved our understanding of disease pathogenesis, but has also increased our knowledge of cell biology. For instance, large-scale population genetic studies have led to the identification of a highly common polymorphism in *Atg16L1*, a gene that is essential for macroautophagy. Macroautophagy (or autophagy) is a pathway by which double membrane vesicles sequester cytosolic material and target the contents for degradation upon fusion with the lysosome (Levine *et al.*, 2011). Mice deficient in *Atg16L1* and other autophagy genes in the intestinal epithelium display profound defects in Paneth cell granule packaging (Cadwell *et al.*, 2008, 2009), mucin secretion by goblet cells (Patel *et al.*, 2013), and degradation of bacteria in enterocytes (Benjamin *et al.*, 2013; Conway *et al.*, 2013). Importantly, the recent generation of mice that delete loxP flanked alleles in Paneth cells (defensin 6-Cre mice) has allowed investigators to take this a step forward. Dual deletion of *Atg16L1* and *Xbp1*, an ER stress gene that has also been linked to IBD, leads to dramatic intestinal inflammation (Adolph *et al.*, 2013). Hence, a cell-intrinsic defect in a secretory epithelial cell type can directly lead to intestinal disease. Nevertheless, basic research of the autophagy pathway reminds us that the epithelium does not exist in a vacuum. Among other immune functions, autophagy has also been shown to have a role in dampening inflammatory cytokine expression in macrophages (Saitoh *et al.*, 2008). Also, the Paneth cell dysfunction in *Atg16L1* deficient mice is only apparent in the presence of murine norovirus (MNV), an intestinal virus that is endemic in mouse colonies (Cadwell *et al.*, 2010). Surprisingly, these same *Atg16L1* deficient mice are incredibly resistant to intestinal bacterial infection due to an enhanced immune response (Marchiando *et al.*, 2013). Thus, it is still unclear why a basic membrane trafficking pathway is linked to a disorder that features the breakdown of

the intestinal epithelial barrier following inflammation. One intriguing possibility is that the same cell biological pathway is important for barrier function in both epithelial and non-epithelial cells, but the exact contribution and mechanism are distinct.

These challenges facing the IBD field highlight the need for multidisciplinary approaches in investigating epithelial biology. In addition to technological progress, investigators are increasingly engaging in both small-scale collaborations as well as giant multi-institutional efforts. The generation of a reference catalog of commensal bacteria that colonize epithelial surfaces, the human microbiome project, will be an incredible resource for research into how distinct epithelia interact with microorganisms ([Human Microbiome Project Consortium, 2012](#)). In comparison, we still lack an appreciation of how fungi, parasites, and viruses maintain or disrupt epithelial function. Another topic that requires better resolution is the contribution of epithelia to interindividual variability. The answer to this question will have significant medical impact in the context of personalized medicine and spread of infectious agents such as *Staphylococcus aureus*. Antibiotic resistant strains of *S. aureus* have devastated the healthcare system by causing a range of diseases once it penetrates the epithelial barrier, but this bacterium also colonizes the skin as an innocuous commensal in about one-third of individuals ([Lowy, 1998](#)). Advances in technology and the way investigators approach difficult problems through teamwork will solve many of these tough challenges in epithelial biology and barrier function in the years to come.

See also: Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell–Cell Adhesion and the Cytoskeleton

References

- Adolph, T.E., Tomczak, M.F., Niederreiter, L., *et al.*, 2013. Paneth cells as a site of origin for intestinal inflammation. *Nature* 503, 272–276.
- Agace, W.W., Persson, E.K., 2012. How vitamin A metabolizing dendritic cells are generated in the gut mucosa. *Trends in Immunology* 33, 42–48.
- Alberts, B., Wilson, J.H., Hunt, T., 2008. *Molecular Biology of the Cell*. New York, NY: Garland Science.
- Arthur, J.C., Perez-Chanona, E., Muhlbauer, M., *et al.*, 2012. Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science* 338, 120–123.
- Barker, N., Van Es, J.H., Kuipers, J., *et al.*, 2007. Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* 449, 1003–1007.
- Baum, B., Georgiou, M., 2011. Dynamics of adherens junctions in epithelial establishment, maintenance, and remodeling. *Journal of Cell Biology* 192, 907–917.
- Bement, W.M., Forscher, P., Mooseker, M.S., 1993. A novel cytoskeletal structure involved in purse string wound closure and cell polarity maintenance. *Journal of Cell Biology* 121, 565–578.
- Ben-Yosef, T., Belyantseva, I.A., Saunders, T.L., *et al.*, 2003. Claudin 14 knockout mice, a model for autosomal recessive deafness DFNB29, are deaf due to cochlear hair cell degeneration. *Human Molecular Genetics* 12, 2049–2061.
- Benjamin, J.L., Sumpter Jr., R., Levine, B., Hooper, L.V., 2013. Intestinal epithelial autophagy is essential for host defense against invasive bacteria. *Cell Host & Microbe* 13, 723–734.
- Bergstrom, K.S., Kisoos-Singh, V., Gibson, D.L., *et al.*, 2010. Muc2 protects against lethal infectious colitis by disassociating pathogenic and commensal bacteria from the colonic mucosa. *PLoS Pathogens* 6, e1000902.
- Berkes, J., Viswanathan, V.K., Savkovic, S.D., Hecht, G., 2003. Intestinal epithelial responses to enteric pathogens: Effects on the tight junction barrier, ion transport, and inflammation. *Gut* 52, 439–451.
- Bizzini, B., 1979. Tetanus toxin. *Microbiological Reviews* 43, 224–240.
- Bjerknes, M., Cheng, H., 2006. Intestinal epithelial stem cells and progenitors. *Methods in Enzymology* 419, 337–383.
- Bjerknes, M., Khandanpour, C., Moroy, T., *et al.*, 2012. Origin of the brush cell lineage in the mouse intestinal epithelium. *Developmental Biology* 362, 194–218.
- Boyle, E.C., Finlay, B.B., 2003. Bacterial pathogenesis: Exploiting cellular adherence. *Current Opinion in Cell Biology* 15, 633–639.
- Bryant, D.M., Mostov, K.E., 2008. From cells to organs: Building polarized tissue. *Nature Reviews Molecular Cell Biology* 9, 887–901.
- Bullen, T.F., Forrest, S., Campbell, F., *et al.*, 2006. Characterization of epithelial cell shedding from human small intestine. *Laboratory Investigation* 86, 1052–1063.
- Cadwell, K., Liu, J.Y., Brown, S.L., *et al.*, 2008. A key role for autophagy and the autophagy gene *Atg16l1* in mouse and human intestinal Paneth cells. *Nature* 456, 259–263.
- Cadwell, K., Patel, K.K., Komatsu, M., Virgin, H.W.T., Stappenbeck, T.S., 2009. A common role for *Atg16L1*, *Atg5* and *Atg7* in small intestinal Paneth cells and Crohn disease. *Autophagy* 5, 250–252.
- Cadwell, K., Patel, K.K., Maloney, N.S., *et al.*, 2010. Virus-plus-susceptibility gene interaction determines Crohn's disease gene *Atg16L1* phenotypes in intestine. *Cell* 141, 1135–1145.
- Cheroutre, H., Lambolez, F., Mucida, D., 2011. The light and dark sides of intestinal intraepithelial lymphocytes. *Nature Reviews Immunology* 11, 445–456.
- Conway, K.L., Kuballa, P., Song, J.H., *et al.*, 2013. *Atg16l1* is required for autophagy in intestinal epithelial cells and protection of mice from *Salmonella* infection. *Gastroenterology* 145, 1347–1357.
- De Lau, W., Barker, N., Low, T.Y., *et al.*, 2011. *Lgr5* homologues associate with Wnt receptors and mediate R-spondin signalling. *Nature* 476, 293–297.
- Fagarasan, S., Honjo, T., 2003. Intestinal IgA synthesis: Regulation of front-line body defences. *Nature Reviews Immunology* 3, 63–72.
- Field, M., Rao, M.C., Chang, E.B., 1989. Intestinal electrolyte transport and diarrheal disease (2). *New England Journal of Medicine* 321, 879–883.
- Gerbe, F., Legraverend, C., Jay, P., 2012. The intestinal epithelium tuft cells: Specification and function. *Cellular and Molecular Life Sciences* 69, 2907–2917.
- Guzman, M.G., Halstead, S.B., Artsob, H., *et al.*, 2010. Dengue: A continuing global threat. *Nature Reviews Microbiology* 8, S7–16.
- Hall, A.J., Curns, A.T., McDonald, L.C., Parashar, U.D., Lopman, B.A., 2012. The roles of *Clostridium difficile* and norovirus among gastroenteritis-associated deaths in the United States, 1999–2007. *Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America* 55, 216–223.
- Hasenberg, M., Stegemann-Koniszewski, S., Gunzer, M., 2013. Cellular immune reactions in the lung. *Immunological Reviews* 251, 189–214.
- Heazlewood, C.K., Cook, M.C., Eri, R., *et al.*, 2008. Aberrant mucin assembly in mice causes endoplasmic reticulum stress and spontaneous inflammation resembling ulcerative colitis. *PLoS Medicine* 5, e54.
- Hermiston, M.L., Gordon, J.I., 1995. *In vivo* analysis of cadherin function in the mouse intestinal epithelium: Essential roles in adhesion, maintenance of differentiation, and regulation of programmed cell death. *Journal of Cell Biology* 129, 489–506.
- Honda, K., Littman, D.R., 2012. The microbiome in infectious disease and inflammation. *Annual Review of Immunology* 30, 759–795.
- Hooper, L.V., 2009. Do symbiotic bacteria subvert host immunity? *Nature Reviews Microbiology* 7, 367–374.
- Human Microbiome Project Consortium, 2012. A framework for human microbiome research. *Nature* 486, 215–221.
- Jenny, M., Uhl, C., Roche, C., *et al.*, 2002. Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *EMBO Journal* 21, 6338–6347.
- Jostins, L., Ripke, S., Weersma, R.K., *et al.*, 2012. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 491, 119–124.
- Kassam, Z., Lee, C.H., Yuan, Y., Hunt, R.H., 2013. Fecal microbiota transplantation for *Clostridium difficile* infection: Systematic review and meta-analysis. *American Journal of Gastroenterology* 108, 500–508.
- Katz, J.P., Perreault, N., Goldstein, B.G., *et al.*, 2002. The zinc-finger transcription factor *Klf4* is required for terminal differentiation of goblet cells in the colon. *Development* 129, 2619–2628.
- Khor, B., Gardet, A., Xavier, R.J., 2011. Genetics and pathogenesis of inflammatory bowel disease. *Nature* 474, 307–317.
- King, C.H., Dickman, K., Tisch, D.J., 2005. Reassessment of the cost of chronic helminth infection: A meta-analysis of disability-related outcomes in endemic schistosomiasis. *Lancet* 365, 1561–1569.
- LeBlanc, J.G., Milani, C., Giori, D.E., *et al.*, 2013. Bacteria as vitamin suppliers to their host: A gut microbiota perspective. *Current Opinion in Biotechnology* 24, 160–168.

- Levine, B., Mizushima, N., Virgin, H.W., 2011. Autophagy in immunity and inflammation. *Nature* 469, 323–335.
- Lillehoj, E.P., Kato, K., Lu, W., Kim, K.C., 2013. Cellular and molecular biology of airway mucins. *International Review of Cell and Molecular Biology* 303, 139–202.
- Lowy, F.D., 1998. *Staphylococcus aureus* infections. *New England Journal of Medicine* 339, 520–532.
- Marchiando, A.M., Ramanan, D., Ding, Y., *et al.*, 2013. A deficiency in the autophagy gene Atg16L1 enhances resistance to enteric bacterial infection. *Cell Host & Microbe* 14, 216–224.
- May, C.L., Kaestner, K.H., 2010. Gut endocrine cell development. *Molecular and Cellular Endocrinology* 323, 70–75.
- McDole, J.R., Wheeler, L.W., McDonald, K.G., *et al.*, 2012. Goblet cells deliver luminal antigen to CD103+ dendritic cells in the small intestine. *Nature* 483, 345–349.
- McGuckin, M.A., Eri, R., Simms, L.A., Florin, T.H., Radford-Smith, G., 2009. Intestinal barrier dysfunction in inflammatory bowel diseases. *Inflammatory Bowel Disease* 15, 100–113.
- Mitic, L.L., Anderson, J.M., 1998. Molecular architecture of tight junctions. *Annual Review of Physiology* 60, 121–142.
- Mori-Akiyama, Y., Van Den Born, M., Van Es, J.H., *et al.*, 2007. SOX9 is required for the differentiation of paneth cells in the intestinal epithelium. *Gastroenterology* 133, 539–546.
- Murray, C.J., Rosenfeld, L.C., Lim, S.S., *et al.*, 2012. Global malaria mortality between 1980 and 2010: A systematic analysis. *Lancet* 379, 413–431.
- Noah, T.K., Donahue, B., Shroyer, N.F., 2011. Intestinal development and differentiation. *Experimental Cell Research* 317, 2702–2710.
- Ouellette, A.J., 2010. Paneth cells and innate mucosal immunity. *Current Opinion in Gastroenterology* 26, 547–553.
- Pacheco, S.M., Johnson, S., 2013. Important clinical advances in the understanding of *Clostridium difficile* infection. *Current Opinion in Gastroenterology* 29, 42–48.
- Packey, C.D., Sartor, R.B., 2009. Commensal bacteria, traditional and opportunistic pathogens, dysbiosis and bacterial killing in inflammatory bowel diseases. *Current Opinion in Infectious Diseases* 22, 292–301.
- Pappenheimer, J.R., Reiss, K.Z., 1987. Contribution of solvent drag through intercellular junctions to absorption of nutrients by the small intestine of the rat. *Journal of Membrane Biology* 100, 123–136.
- Parry, C.M., Hien, T.T., Dougan, G., White, N.J., Farrar, J.J., 2002. Typhoid fever. *New England Journal of Medicine* 347, 1770–1782.
- Patel, K.K., Miyoshi, H., Beatty, W.L., *et al.*, 2013. Autophagy proteins control goblet cell function by potentiating reactive oxygen species production. *EMBO Journal* 32, 3130–3144.
- Preidis, G.A., Hill, C., Guerrant, R.L., *et al.*, 2011. Probiotics, enteric and diarrheal diseases, and global health. *Gastroenterology* 140, 8–14.
- Ragnarsdottir, B., Svanborg, C., 2012. Susceptibility to acute pyelonephritis or asymptomatic bacteriuria: Host-pathogen interaction in urinary tract infections. *Pediatric Nephrology* 27, 2017–2029.
- Randall, T.D., Carragher, D.M., Rangel-Moreno, J., 2008. Development of secondary lymphoid organs. *Annual Review of Immunology* 26, 627–650.
- Rosenblatt, J., Raff, M.C., Cramer, L.P., 2001. An epithelial cell destined for apoptosis signals its neighbors to extrude it by an actin- and myosin-dependent mechanism. *Current Biology* 11, 1847–1857.
- Saitoh, T., Fujita, N., Jang, M.H., *et al.*, 2008. Loss of the autophagy protein Atg16L1 enhances endotoxin-induced IL-1 β production. *Nature* 456, 264–268.
- Schubert, M.L., 2011. Gastric secretion. *Current Opinion in Gastroenterology* 27, 536–542.
- Shan, M., Gentile, M., Yeiser, J.R., *et al.*, 2013. Mucus enhances gut homeostasis and oral tolerance by delivering immunoregulatory signals. *Science* 342 (6157), 447–453.
- Sheng, Y.H., Hasnain, S.Z., Florin, T.H., McGuckin, M.A., 2012. Mucins in inflammatory bowel diseases and colorectal cancer. *Journal of Gastroenterology and Hepatology* 27, 28–38.
- Simon, D.B., Lu, Y., Choate, K.A., *et al.*, 1999. Paracellin-1, a renal tight junction protein required for paracellular Mg $^{2+}$ resorption. *Science* 285, 103–106.
- Tsukita, S., Furuse, M., 2002. Claudin-based barrier in simple and stratified cellular sheets. *Current Opinion in Cell Biology* 14, 531–536.
- Turner, J.R., 2006. Molecular basis of epithelial barrier regulation: From basic mechanisms to clinical application. *American Journal of Pathology* 169, 1901–1909.
- Vaishnav, S., Yamamoto, M., Severson, K.M., *et al.*, 2011. The antibacterial lectin RegIII γ promotes the spatial segregation of microbiota and host in the intestine. *Science* 334, 255–258.
- Van Der Sluis, M., De Koning, B.A., De Bruijn, A.C., *et al.*, 2006. Muc2-deficient mice spontaneously develop colitis, indicating that MUC2 is critical for colonic protection. *Gastroenterology* 131, 117–129.
- Von Lintig, J., 2012. Provitamin A metabolism and functions in mammalian biology. *American Journal of Clinical Nutrition* 96, 1234S–1244S.
- Wehkamp, J., Salzman, N.H., Porter, E., *et al.*, 2005. Reduced Paneth cell alpha-defensins in ileal Crohn's disease. *Proceedings of the National Academy of Sciences of the United States* 102, 18129–18134.
- Wilcox, E.R., Burton, Q.L., Naz, S., *et al.*, 2001. Mutations in the gene encoding tight junction claudin-14 cause autosomal recessive deafness DFNB29. *Cell* 104, 165–172.
- Wu, L., KewalRamani, V.N., 2006. Dendritic-cell interactions with HIV: Infection and viral dissemination. *Nature Reviews Immunology* 6, 859–868.
- Wu, S., Rhee, K.J., Albesiano, E., *et al.*, 2009. A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17T cell responses. *Nature Medicine* 15, 1016–1022.

Phagocytic Synapses

HS Goodridge and N Hassanzadeh-Kiabi, Cedars-Sinai Medical Center, Los Angeles, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Adaptive immune response The later phase of an immune response that involves T cells and B cells and is responsible for generating immune memory; each T or B cell expresses receptors with a unique molecular structure that are only activated by specific molecules (antigens), which allows them to induce antigen-specific responses.

Antigen presenting cells Cells that display antigens such as peptides acquired from degraded microbes on their surface to activate T cells; dendritic cells are particularly effective at antigen presentation.

Innate immune response The early phase of an immune response when microbes and other threats are detected by cells such as neutrophils, macrophages, and dendritic cells to initiate an immune response; involves nonspecific

detection of molecular patterns commonly associated with microbial components by 'pattern recognition receptors'; microbes acquired during this phase are destroyed and processed to derive antigen molecules for presentation to T cells to activate the 'adaptive immune response.'

Opsonization Binding of serum proteins such as complement and antibodies to the surface of microbes and cells to enhance their detection and uptake by phagocytic cells.

Pattern recognition receptors Receptors (such as Toll-like receptors and Dectin-1) used by immune cells to detect microbial components.

Phagocytosis Receptor-mediated uptake of particulate targets $>0.5\ \mu\text{m}$ by a process of membrane engulfment driven by the actin cytoskeleton.

Introduction

Cell-cell contact mediated by the interaction of proteins on the surface of neighboring cells enables intercellular communication. Large multi-protein complexes can form stable connections or 'synapses' between cells that in some cases can last hours, days, or even years. Neurons, for example, transmit electrical and chemical signals by forming synapses with muscle cells or other neurons. In the immune system, activation of an adaptive immune response is dependent on the formation of 'immunological synapses' between antigen presenting cells (APC) (such as dendritic cells) and T cells (Figure 1). Recent studies have shown that the responses of phagocytic cells of the innate immune system are also controlled by synapses: 'phagocytic synapses' (Dustin, 2012; Goodridge *et al.*, 2011; Yamauchi *et al.*, 2012; Figure 1).

Phagocytosis is the process by which cells such as macrophages, dendritic cells, and neutrophils engulf large ($>0.5\ \mu\text{m}$) particulate matter, including microbes, apoptotic cells, and cellular/extracellular matrix debris. Phagocytic receptors are activated when they detect target particles. Their intracellular signals instruct cytoskeletal rearrangement to coordinate particle internalization. Several receptors that detect microbial components can induce phagocytosis of microbes (Figure 1). For example, when Dectin-1 on the surface of a macrophage detects β -glucan in the wall of a fungal yeast cell, it triggers phagocytic internalization of the yeast by the macrophage. Macrophages are also instructed to phagocytose 'opsonized' targets that are coated with antibodies (immunoglobulins, Ig) or complement proteins found in serum. For example, the Fc γ receptor (Fc γ R) detects IgG-opsonized microbes to trigger internalization and microbial killing.

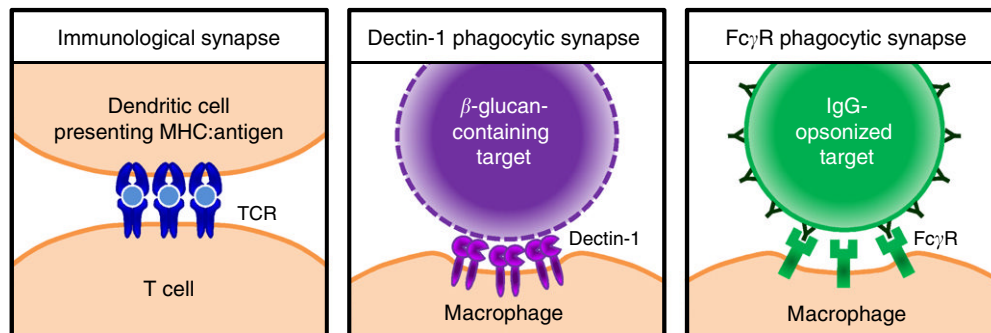


Figure 1 Immunological and phagocytic synapses. Formation of immunological synapses between dendritic cells (or other APCs) and T cells are required for T cell activation. The TCR is only activated when it detects its cognate antigen presented by MHC molecules on the dendritic cell surface. Similarly, phagocytic synapses permit the activation of the phagocytic receptors Dectin-1 and Fc γ R to induce target phagocytosis and cytokine production. Dectin-1 on the surface of a macrophage (or neutrophil or dendritic cell) detects β -glucans in fungal cell walls. Fc γ R detects the Fc portion of IgG coating target particles such as microbes.

In addition to phagocytosis, intracellular signals generated by Dectin-1 and Fc γ R also induce the production of inflammatory mediators such as cytokines. Moreover, macrophages and dendritic cells use phagocytosis to acquire antigens for presentation to T cells to initiate adaptive immune responses. As discussed below, phagocytic synapses formed when Dectin-1 and Fc γ R encounter target particles control the initiation of phagocytosis and cytokine production to ensure that these responses are not inadvertently triggered.

Dectin-1 Activation by β -Glucans

Dectin-1 detects fungal β -glucans (specifically β -1,3-glucans) and plays key roles in immune defense against fungi such as *Candida albicans* and *Aspergillus fumigatus* (Drummond and Brown, 2011; Goodridge *et al.*, 2009). β -Glucans are large polymers of D-glucose and are a major structural component of fungal cell walls. Upon detection of β -glucans in fungal cell walls, Dectin-1 signaling triggers phagocytic uptake of fungal cells, the production of reactive oxygen species (oxidative burst response) to kill the internalized fungi, and the generation of inflammatory mediators to recruit and activate other immune cells. Unlike most microbial receptors, however, Dectin-1 is only activated when it detects its agonist in particulate form, for example, β -glucan in a yeast cell wall or β -glucan-coated latex beads (Goodridge *et al.*, 2011). Soluble β -glucans fail to activate the receptor. Even very large, highly polymerized soluble β -glucans that simultaneously bind and cross-link lots of receptor molecules are unable to activate Dectin-1. This is in contrast to Toll-like receptor 2, for example, which is activated in response to microbial lipopeptides in soluble form as well as in the context of a microbial cell wall.

This feature of Dectin-1 is thought to limit the responses of macrophages to direct encounters with β -glucan-containing fungal cells. Dectin-1 instructs responses that other microbial receptors such as Toll-like receptors do not (phagocytosis and a strong oxidative burst response). It is therefore important to prevent collateral damage caused by the initiation of powerful antimicrobial and inflammatory responses in the absence of an immediate threat, for example, when macrophages detect β -glucans shed from yeast that are not in the immediate vicinity. The phagocytic synapse model was proposed to explain how Dectin-1 can distinguish between soluble and particulate β -glucans (Goodridge *et al.*, 2011).

Dectin-1 signals in a similar manner to the T cell antigen receptor (TCR), which transmits signals to activate T cells by associating with CD3 molecules. CD3 molecules have immunoreceptor tyrosine-based activation motifs (ITAMs) which become phosphorylated on tyrosine residues by members of the Src family of tyrosine kinases (SFKs) when the antigen-specific TCR detects its cognate antigen (Figure 2). Zap70 kinases are then recruited to the phosphorylated ITAM tyrosines to propagate the signal. Dectin-1 has an ITAM-like motif (called a hemITAM) in its cytoplasmic tail, which is phosphorylated by SFKs when Dectin-1 detects particulate or immobilized β -glucans (Figure 2). The Zap70-related kinase Syk is then recruited to Dectin-1 to transduce downstream signaling.

As noted above, Dectin-1 does not induce responses when it binds soluble β -glucans. Likewise, the TCR does not respond to

soluble antigen, but is only activated when it detects its antigen in the context of major histocompatibility complex (MHC) molecules on the surface of APCs such as dendritic cells. When an antigen is presented to the TCR by MHC, TCR molecules cluster and a large multi-protein complex forms at the contact site between the dendritic cell and the T cell, resulting in a stable interaction between the two cells – an immunological synapse (Figure 1; Dustin, 2012; Dustin and Depoil, 2011). Within the immunological synapse, membrane-intrinsic and recruited proteins are arranged in a bull's-eye pattern to permit TCR signaling and T cell activation. One of the key events in the formation of the immunological synapse is the isolation of CD45 molecules away from TCR molecules. CD45 is a membrane-intrinsic tyrosine phosphatase that is initially required for SFK activation but subsequently plays a negative regulatory role by dephosphorylating the ITAM tyrosines and other signaling molecules activated by the SFKs and Zap70. Sequestration of CD45 away from the TCR therefore permits the generation of a productive TCR signal.

A similar mechanism controls the initiation of Dectin-1 signaling – the phagocytic synapse (Goodridge *et al.*, 2011; Figure 3). In macrophages, CD45 and a related tyrosine phosphatase CD148 regulate Dectin-1 activation by first permitting SFK activation and subsequently negatively regulating Dectin-1 signaling via SFKs and Syk. When a macrophage engages a β -glucan-containing particle (e.g., a yeast cell or a β -glucan-coated bead), Dectin-1 receptor molecules cluster at the contact site and their activation is evidenced by the recruitment of activated SFKs and Syk. This is accompanied by the movement of CD45 and CD148 away from the clustered receptors to permit continued Dectin-1 signaling. Confocal microscopy reveals a bull's-eye pattern of molecules on the surface of the cell with clustered Dectin-1 and active SFKs and Syk in the central zone and excluded CD45/CD148 around the outside (Figures 4 and 5).

Fc γ R Activation by Immunoglobulin G (IgG)

Phagocytic synapses also control Fc γ R-mediated uptake of antibody-opsonized targets (Yamauchi *et al.*, 2012; Figures 3 and 4). Like Dectin-1, Fc γ R signaling triggers actin-driven phagocytosis of target particles, as well as the induction of an oxidative burst response and production of inflammatory mediators. Therefore, Fc γ R must also distinguish between soluble and particulate forms of its agonist so that it will transduce signals upon detection of IgG-opsonized cells (e.g., microbes, tumor cells) but ignore free IgG molecules. When macrophages engage IgG-opsonized cells or IgG-coated beads, Fc γ R clustering is observed at the contact site. Signaling is initiated by SFK phosphorylation of ITAMs to permit Syk recruitment and activation (Figure 2). CD45 and CD148 regulate Fc γ R signaling (Zhu *et al.*, 2008) and CD45 exclusion is a feature of Fc γ R phagocytic synapse formation (Yamauchi *et al.*, 2012). CD148 is also likely to be excluded from the Fc γ R phagocytic synapse.

Studies of Fc γ R activation revealed another key player in phagocytic synapse formation – myosin II (Yamauchi *et al.*, 2012). Myosin II, an ATP-driven motor protein that generates contractile forces along actin filaments, regulates Fc γ R

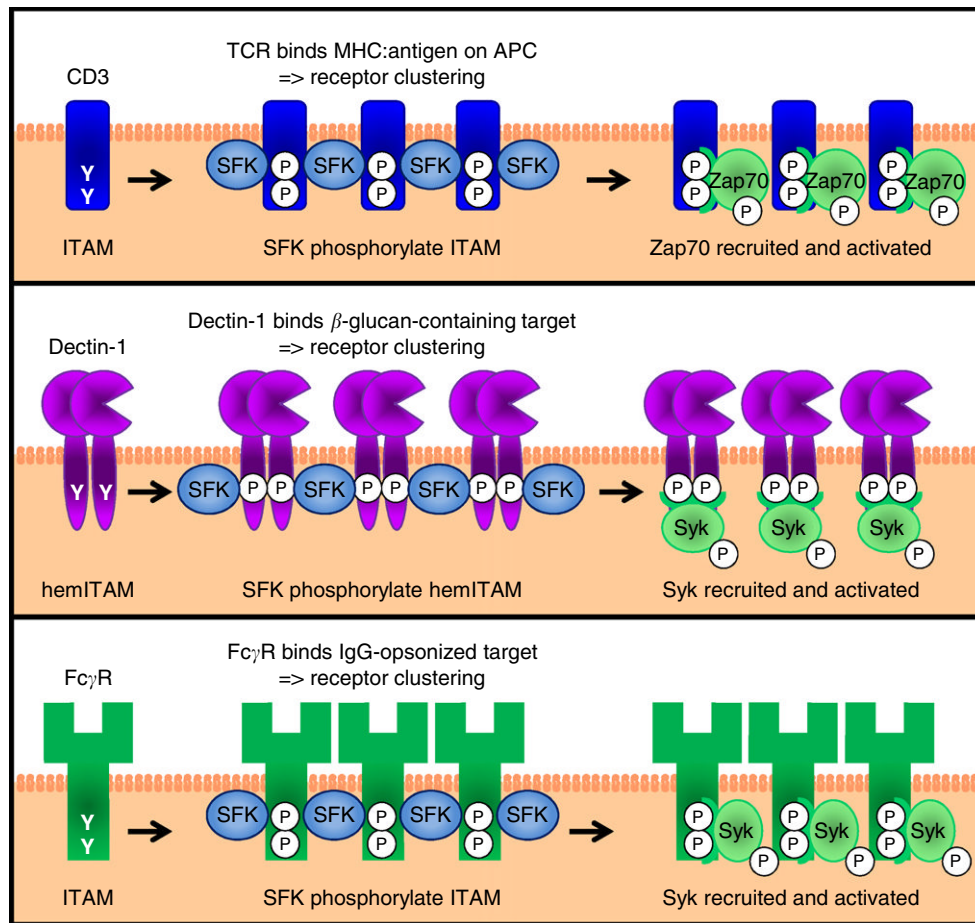


Figure 2 ITAM signaling by the TCR, Dectin-1, and Fc γ R. The TCR, Dectin-1, and Fc γ R signal via ITAMs or ITAM-like motifs. The TCR signals via ITAM-containing CD3 molecules. ITAMs are motifs containing dual tyrosines that are phosphorylated by SFK upon ligand binding, receptor clustering, and synapse formation. A dual phosphorylated CD3 ITAM forms a docking site for the tyrosine kinase Zap70, which binds via its dual SH2 domains. Zap70 is then activated by phosphorylation to initiate downstream signaling. Dectin-1 and Fc γ R on the macrophage surface signal in a similar manner, but they recruit a related kinase called Syk instead of Zap70. Dectin-1 does not possess a conventional ITAM. Instead, an ITAM-like motif (called a hemITAM) in its intracellular tail contains a single tyrosine residue that is phosphorylated by SFK. Syk is then recruited to phosphorylated Dectin-1 homodimers.

phagocytosis by controlling the actin dynamics responsible for pseudopod extension and target engulfment (Araki, 2006). It also plays a key role in the formation of the immunological synapse (Dustin, 2012), so its involvement in Fc γ R phagocytic synapse formation illustrates another parallel between the two synapse models. Following Fc γ R clustering and SFK activation, myosin II is required for ITAM phosphorylation and Syk activation to coordinate actin polymerization (Yamauchi *et al.*, 2012).

CD45 associates with actin filaments via spectrin and ankyrin, which control lateral movement of CD45 in the plasma membrane (Cairo *et al.*, 2010; Pradhan and Morrow, 2002). Thus it seems likely that actin dynamics are responsible for CD45 exclusion, although phagocytic synapse formation probably precedes pseudopod extension, since the signals that drive large-scale cytoskeletal movement to form pseudopods are presumably dependent on CD45 exclusion. Spectrin is localized to the plasma membrane via interactions with PI-4,5-bisphosphate (PIP₂), which is also thought to participate

in the formation of the Fc γ R phagocytic synapse. Given the parallels between Dectin-1 and Fc γ R signaling (Goodridge *et al.*, 2012), it is likely that myosin II regulates the Dectin-1 phagocytic synapse in a comparable manner.

The Phagocytic Synapse Model

The current phagocytic synapse model (Figure 3) therefore suggests that when Dectin-1/Fc γ R molecules on the surface of a macrophage initially detect a soluble or particulate ligand, a small number of receptor molecules cluster, triggering low-level SFK activation and ITAM phosphorylation. Although initially required to permit SFK activation, CD45 and CD148 subsequently limit receptor signaling by ITAM dephosphorylation, thereby preventing receptor activation in response to soluble agonists. However, if a particle is engaged by a large number of receptor molecules and myosin II is recruited, actin will begin to polymerize near the clustered receptors.

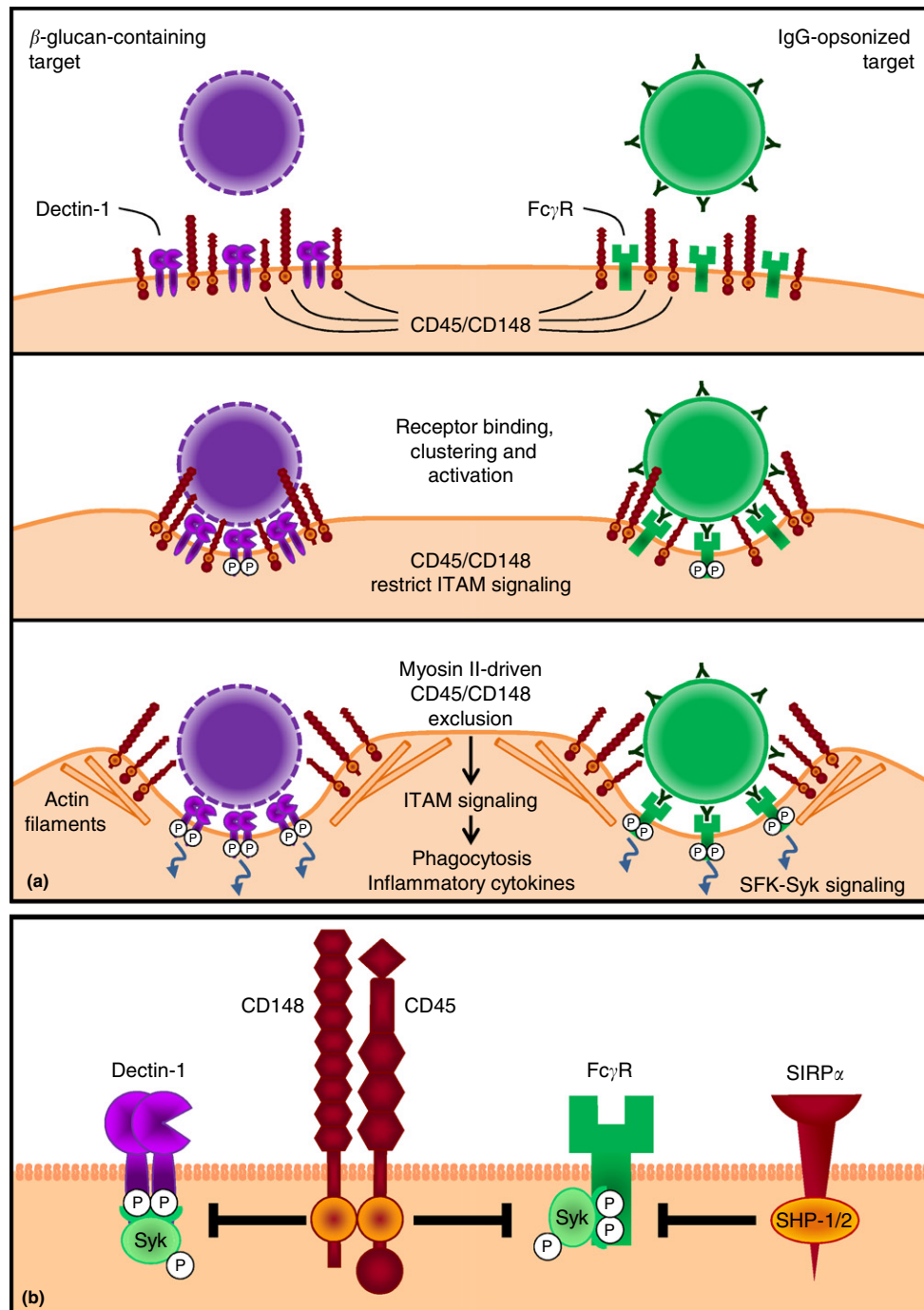


Figure 3 Model of phagocytic synapse formation to permit Dectin-1 and Fc γ R activation. (a) Receptors begin to cluster upon initial recognition of particulate targets by Dectin-1/Fc γ R. At the early stages, CD45 and CD148 are initially required for SFK activation (they remove an inhibitory tyrosine phosphorylation from SFK). However, the continued presence of CD45 and CD148 prohibits productive ITAM signaling because they dephosphorylate ITAMs and inactivate SFK, Syk, and downstream signaling molecules. Therefore, CD45 and CD148 must be isolated from the clustered receptors to permit productive signaling, which results in target internalization and other responses such as cytokine production. The isolation of CD45 (and probably CD148) away from clustered Fc γ R is controlled by myosin II. It is likely that myosin II also controls CD45/CD148 movement during formation of the Dectin-1 phagocytic synapse. When soluble ligands such as β -glucan fragments shed from fungal cell walls bind to the receptors they do not induce phagocytic synapse formation. Consequently, they do not trigger receptor activation, phagocytosis, and cytokine production. (b) In addition to CD45/CD148, SIRP α also negatively regulates Fc γ R-mediated phagocytosis upon detection of its ligand CD47 on target cells. SIRP α has an ITIM in its cytoplasmic tail, which recruits the tyrosine phosphatases SHP-1 and SHP-2 following CD47 binding. The SHP phosphatases antagonize ITAM signaling by dephosphorylating ITAMs/ITAM-signaling molecules to block Fc γ R signaling and prevent phagocytosis.

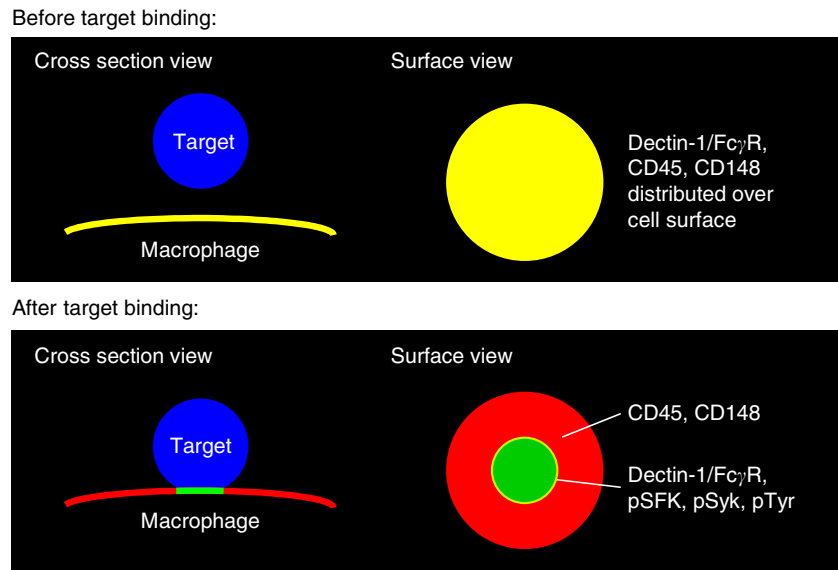


Figure 4 Phagocytic synapse ‘bull’s-eyes.’ A ‘bull’s-eye’ pattern of staining is observed when phagocytic synapses are visualized by confocal microscopy using antibodies against the receptors, intracellular signaling molecules, and CD45/CD148. Prior to target binding, the receptors and CD45/CD148 are distributed evenly over the cell surface. Upon target binding, the receptors cluster in the center of the ‘bull’s-eye’ and are shown to be actively signaling by the presence of phosphorylated SFK and Syk, and detection of tyrosine-phosphorylated proteins (likely the receptor ITAMs, SFK, Syk, and downstream signaling proteins). CD45 and CD148 are now excluded from this central region to allow productive receptor signaling to drive phagocytosis and cytokine production.

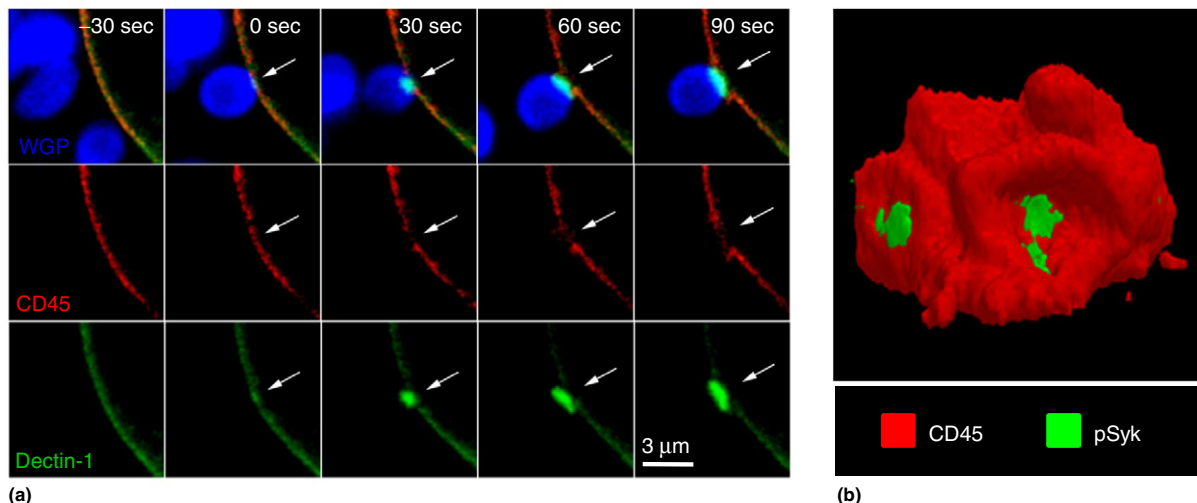


Figure 5 Confocal imaging of Dectin-1 phagocytic synapses. (a) Frames taken at 30 s intervals from a movie of phagocytic synapse formation upon detection of a β -glucan-containing yeast particle (whole glucan particle, WGP) by a macrophage. When the particle (blue) is detected by the cell, CD45 (red) is cleared from the site of particle attachment as Dectin-1 receptors (green) cluster. (b) 3-D isosurface modeling of two Dectin-1 phagocytic synapses reveals ‘bull’s-eye’ distribution of CD45 (red) and active Syk (phosphorylated Syk, pSyk; green) at the macrophage surface. Bound particles are not shown. Reproduced with permission from Goodridge, H.S., *et al.*, 2011. Activation of the innate immune receptor Dectin-1 upon formation of a ‘phagocytic synapse.’ *Nature* 472, 471–475.

As myosin II further orchestrates actin dynamics, CD45 and CD148 are dragged away from the clustered receptors and pseudopods begin to extend around the particle. This establishes a positive feedback loop to amplify Dectin-1/Fc γ R signaling by preventing dephosphorylation of the receptor ITAMs. This mechanism is also likely to regulate signaling by these receptors in other phagocytic cells such as dendritic cells and neutrophils.

The early events that drive phagocytic synapse formation are still unclear. The phagocytic synapse may originate from the coalescence of small clusters of Dectin-1/Fc γ R molecules interacting with an immobilized ligand on the same particle. Soluble agonists may be able to trigger receptor clustering and perhaps even a low level of signaling, but are ultimately unable to activate sufficient signaling to induce a productive phagocytic response because the ligand molecules they detect

are not physically connected to one another and so are unable to induce the receptor aggregation required for synapse formation and isolation of the suppressive CD45/CD148 phosphatases. Another possibility, hypothesized from studies of the immunological synapse, is that Dectin-1 and Fc γ R might act as mechanosensors that detect forces generated by interaction with their immobilized ligands.

It is likely that other molecules are also segregated when phagocytic synapses form. Biological targets (microbes, apoptotic cells, tumor cells, etc.) are complex in composition, so multiple receptors are usually involved in their detection. Even though these receptors may not instruct phagocytic internalization, their presence in the synapse may affect the efficiency of target binding and the rate of uptake (Goodridge *et al.*, 2012). In particular, integrins act as adhesion molecules and as sensors of microbial components and opsonizing complement proteins, and can themselves instruct phagocytosis. Indeed, integrins are involved in the formation of the immunological synapse. It is therefore likely that they also participate in phagocytic synapse formation and organization, and since they signal via tyrosine kinases they may also be

regulated by CD45/CD148. It is not yet clear whether other phagocytic receptors are regulated in the same way as Dectin-1 and Fc γ R. However, as discussed below, evidence from studies of phagocytosis of apoptotic cells and tumor cells suggests that this may be the case and provides another mechanism by which the phagocytic synapse can be regulated.

Regulation of the Phagocytic Synapse by CD47

Phagocytosis is not only important for the control of microbial invaders, but also for clearance of 'self' cells to maintain tissue homeostasis by removing senescent or damaged cells (including tumor cells), and for tissue remodeling. To distinguish between cells that they should phagocytose and cells that they should ignore, macrophages look for 'eat me' and 'don't eat me' signals. For example, calreticulin is normally located in the endoplasmic reticulum, but apoptotic cells and tumor cells display it on the cell surface where it represents an 'eat me' signal that is detected by low-density lipoprotein-related protein (LRP, also known as CD91) (Chao *et al.*, 2012).

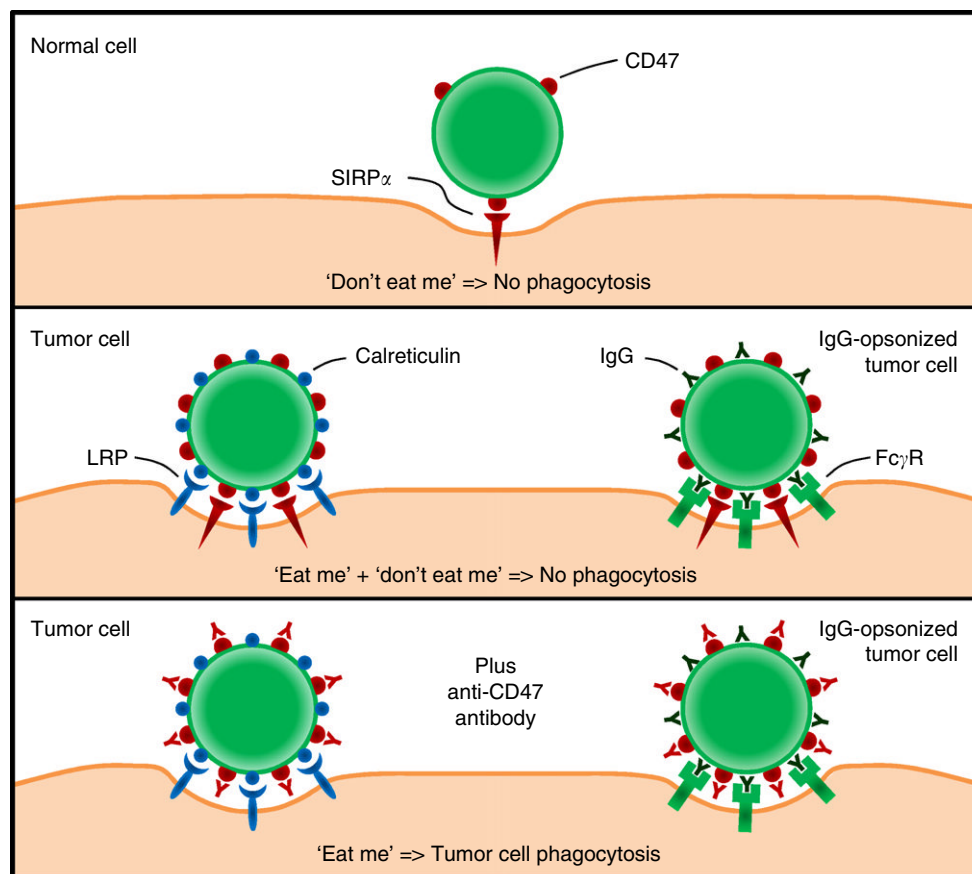


Figure 6 Regulation of the Fc γ R phagocytic synapse by CD47-SIRP α . CD47 on the surface of target cells delivers a 'don't eat me' signal to macrophages via binding to SIRP α to prevent phagocytic uptake of the CD47-bearing cell. Senescent and apoptotic cells decrease CD47 expression and/or increase the expression of 'eat me' signals such as calreticulin to encourage their clearance. However, tumor cells up-regulate CD47 expression to avoid being phagocytosed, even in the presence of 'eat me' signals such as calreticulin or opsonizing antibodies. Therapeutic strategies that target CD47 are being explored for cancer treatment to prevent SIRP α -mediated suppression of phagocytic synapse formation and promote tumor cell uptake. Coadministration of anti-CD47 antibodies that block the 'don't eat me' signal along with opsonizing antitumor antibodies that deliver an 'eat me' signal to activate Fc γ R phagocytic synapses may be particularly effective.

When immunoglobulins and complement proteins attach to target cells (opsonization) they also deliver 'eat me' signals to instruct phagocytosis via Fc receptor- and complement receptor-mediated uptake, respectively. CD47, on the other hand, delivers a 'don't eat me' signal to macrophages (Oldenborg, 2013; Oldenborg *et al.*, 2000). CD47 is expressed on all cells but at variable levels. Healthy red blood cells, for example, avoid being phagocytosed by splenic macrophages due to their high expression of CD47. As red blood cells age, however, their expression of CD47 decreases and they are cleared by splenic macrophages.

The macrophage receptor that detects CD47 is called signal regulatory protein α (SIRP α). SIRP α signals via an immunoreceptor tyrosine-based inhibition motif (ITIM). Activated ITIM-containing receptors oppose signals generated by ITAM-containing receptors by recruiting the tyrosine phosphatases SHP-1 and SHP-2 to dephosphorylate the tyrosine-phosphorylated proteins that transduce ITAM signals (Figure 3(b)). Indeed, CD47-SIRP α signaling has been shown to restrict Fc γ R-mediated phagocytic uptake of IgG-opsonized cells by decreasing the tyrosine phosphorylation and accumulation of myosin II, thereby restricting phagocytic synapse formation (Tsai and Discher, 2008). CD47-SIRP α also suppresses target cell uptake by other phagocytic receptors, including complement receptor and LRP (Chao *et al.*, 2010; Tsai and Discher, 2008). These receptors do not appear to signal via ITAMs but the phagocytic signals they transduce have been reported to be driven by tyrosine kinases, which are presumably antagonized by the SIRP α -activated phosphatases. It is possible that CD45 and CD148 also regulate these receptors and that the phagocytic synapse model is not just restricted to phagocytic receptors that signal via ITAMs.

Interestingly, tumor cells often exploit the CD47-SIRP α interaction by up-regulating CD47 expression and using it to evade immune detection (Chao *et al.*, 2012). However, dying tumor cells and even some viable tumor cells also have higher surface expression of the calreticulin 'eat me' signal (Chao *et al.*, 2010, 2012). Blockade of CD47 using antibodies is therefore being investigated as an anticancer treatment to silence the 'don't eat me' signal. Anti-CD47 antibodies block the interaction of CD47 with SIRP α , which allows for the delivery of the calreticulin-LRP 'eat me' signal to trigger tumor cell phagocytosis (Figure 6). Normal cells are not cleared, even in the presence of anti-CD47 antibodies, because they do not express calreticulin at the cell surface so they do not deliver an 'eat me' signal via LRP. Anti-CD47 antibodies could be particularly effective at enhancing phagocytic uptake and destruction of tumor cells when delivered in combination with antibodies specific for tumor antigens, which would opsonize tumor cells to simultaneously deliver an 'eat me' signal to Fc receptors (Chao *et al.*, 2012). Blocking the CD47 signal would favor phagocytic synapse formation to instruct uptake of tumor cells, which is required for tumor antigen acquisition for presentation to T cells to initiate an antitumor immune response.

Summary

The phagocytic synapse provides a model mechanism whereby phagocytic cells are able to identify particulate targets that they need to clear, perhaps kill, and ultimately degrade. It permits phagocytes to ignore soluble material that does not require such treatment, thereby conserving energy and limiting collateral damage that might be induced by powerful inflammatory and immune responses. A more complete understanding of how this mechanism regulates phagocytic receptor signaling will inform the development of novel therapeutic strategies to promote or limit phagocytosis and other downstream immune responses.

See also: Intracellular Infectiology: Cell Processes: Phagocytosis

References

- Araki, N., 2006. Role of microtubules and myosins in Fc gamma receptor-mediated phagocytosis. *Frontiers in Bioscience* 11, 1479–1490.
- Cairo, C.W., Das, R., Albohy, A., *et al.*, 2010. Dynamic regulation of CD45 lateral mobility by the spectrin-ankyrin cytoskeleton of T cells. *Journal of Biological Chemistry* 285, 11392–11401.
- Chao, M.P., Jaiswal, S., Weissman-Tsukamoto, R., *et al.*, 2010. Calreticulin is the dominant pro-phagocytic signal on multiple human cancers and is counterbalanced by CD47. *Science Translational Medicine* 2, 63ra94.
- Chao, M.P., Weissman, I.L., Majeti, R., 2012. The CD47-SIRPalpha pathway in cancer immune evasion and potential therapeutic implications. *Current Opinion in Immunology* 24, 225–232.
- Drummond, R.A., Brown, G.D., 2011. The role of Dectin-1 in the host defence against fungal infections. *Current Opinion in Microbiology* 14, 392–399.
- Dustin, M.L., 2012. Signaling at neuro/immune synapses. *Journal of Clinical Investigation* 122, 1149–1155.
- Dustin, M.L., Depoil, D., 2011. New insights into the T cell synapse from single molecule techniques. *Nature Reviews Immunology* 11, 672–684.
- Goodridge, H.S., Reyes, C.N., Becker, C.A., *et al.*, 2011. Activation of the innate immune receptor Dectin-1 upon formation of a 'phagocytic synapse'. *Nature* 472, 471–475.
- Goodridge, H.S., Underhill, D.M., Touret, N., 2012. Mechanisms of Fc receptor and dectin-1 activation for phagocytosis. *Traffic* 13, 1062–1071.
- Goodridge, H.S., Wolf, A.J., Underhill, D.M., 2009. Beta-glucan recognition by the innate immune system. *Immunological Reviews* 230, 38–50.
- Oldenborg, P.A., 2013. CD47: A cell surface glycoprotein which regulates multiple functions of hematopoietic cells in health and disease. *International Scholarly Research Notices Hematology* 2013, 614–619.
- Oldenborg, P.A., Zheleznyak, A., Fang, Y.F., *et al.*, 2000. Role of CD47 as a marker of self on red blood cells. *Science* 288, 2051–2054.
- Pradhan, D., Morrow, J., 2002. The spectrin-ankyrin skeleton controls CD45 surface display and interleukin-2 production. *Immunity* 17, 303–315.
- Tsai, R.K., Discher, D.E., 2008. Inhibition of "self" engulfment through deactivation of myosin-II at the phagocytic synapse between human cells. *Journal of Cell Biology* 180, 989–1003.
- Yamauchi, S., Kawauchi, K., Sawada, Y., 2012. Myosin II-dependent exclusion of CD45 from the site of Fc gamma receptor activation during phagocytosis. *FEBS Letters* 586, 3229–3235.
- Zhu, J.W., Brdicka, T., Katsumoto, T.R., Lin, J., Weiss, A., 2008. Structurally distinct phosphatases CD45 and CD148 both regulate B cell and macrophage immunoreceptor signaling. *Immunity* 28, 183–196.

The Phagocyte NADPH Oxidase: Structure and Assembly of the Key Multicomponent Enzyme of Innate Immunity

J El-Benna and P My-Chan Dang, INSERM U1149, Centre de Recherche sur l'Inflammation CRI, Paris, France and Université Paris Diderot, Paris, France

MB Yaffe, Broad Institute Massachusetts Institute of Technology, Cambridge, MA, USA

© 2016 Elsevier Inc. All rights reserved.

The Phagocyte Oxidative Burst

Phagocytes such as polymorphonuclear neutrophils (PMN), monocytes, and macrophages are the first line of host defenses against pathogens such as bacteria and fungi (Nathan, 2006). At the infection site phagocytes recognize and engulf the pathogen. This process involves the binding of specific pathogen-associated molecular patterns (PAMPs) such as peptidoglycans, lipopolysaccharides (LPS), flagellin to their respective cell receptors, namely Toll-like receptors (TLR), and the binding of opsonins (immunoglobulins G (IgG), C3b, and C3bi fragments) to immunoglobulin receptors and integrins (Witko-Sarsat *et al.*, 2000; Lee *et al.*, 2003). The recognition is generally followed by engulfment of the particle which will be surrounded by a membrane envelope which will constitute a vacuole called the phagosome. Engulfment of the bacteria is accompanied by the activation of neutrophils which will lead to their death and destruction following the release of anti-microbial peptides, proteases, and reactive oxygen species (ROS) oxidizing molecules (Roos *et al.*, 2003; Nauseef, 2007). ROS include superoxide anion ($O_2^{\cdot-}$), hydrogen peroxide (H_2O_2), the hydroxyl radical ($OH^{\cdot}$), and hypochlorous acid (HOCl). They are produced by phagocytes in a powerful 'oxidative burst,' characterized by a rapid, cyanide-insensitive increase in oxygen uptake, an increase in glucose consumption, and abrupt ROS production (Hampton *et al.*, 1998). Superoxide anion ($O_2^{\cdot-}$) is the first molecule produced by nicotinamide adenine dinucleotide phosphate-oxidase (NADPH oxidase) and is the precursor of all the other ROS generated in the phagosome (Figure 1). It is produced by monovalent reduction of oxygen in the following reaction (Babior, 2000):



The Phagocyte NADPH Oxidase NOX2

The enzyme responsible for ROS production is called the phagocyte NADPH oxidase (NOX2) or respiratory burst oxidase (Babior, 1999). The active NADPH oxidase is a multicomponent enzyme system located in the phagosome and plasma membranes of activated cells. It consists of several components, a heterodimer composed of two transmembrane proteins, gp91phox and p22phox (phox: phagocyte oxidase) also called cytochrome b558, and four cytosolic components, p47phox, p67phox, p40phox, and either Rac1 (in monocytes) or rac2 (in neutrophils) (Chanock *et al.*, 1994; Figure 2). The membrane-bound and cytosolic components are separated in the membranes and cytosol respectively in resting cells,

ensuring a dormant unactivated state of NOX2, this last one becoming active only if cells are stimulated (El-Benna *et al.*, 2005). In intact cells, NADPH oxidase activation is accompanied by three important events: (1) the phosphorylation of p47phox, p67phox, p40phox, p22phox, and gp91phox; (2) Rac1/2 activation; (3) cytosolic protein translocation toward the membrane and changes in 'protein-protein' and 'proteins-lipids' interactions (Groemping and Rittinger, 2005; DeLeo and Quinn, 1996).

The vital importance of ROS production by the phagocyte NADPH oxidase was demonstrated by a genetic disease called

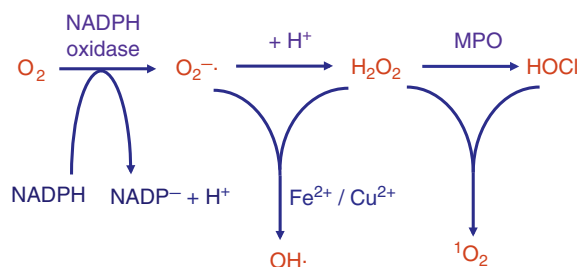


Figure 1 ROS reactions in the phagosome. Molecular oxygen (O_2) is reduced by the NADPH oxidase complex to produce superoxide anion ($O_2^{\cdot-}$). Superoxide anions can react with protons to produce hydrogen peroxide (H_2O_2), which is used by the myeloperoxidase enzyme (MPO) to produce hypochlorous acid (HOCL). Superoxide can react with H_2O_2 to produce hydroxyl radical ($OH^{\cdot}$). This reaction can occur in the presence of H_2O_2 and metals such as Fe^{2+} or Cu^{2+} .

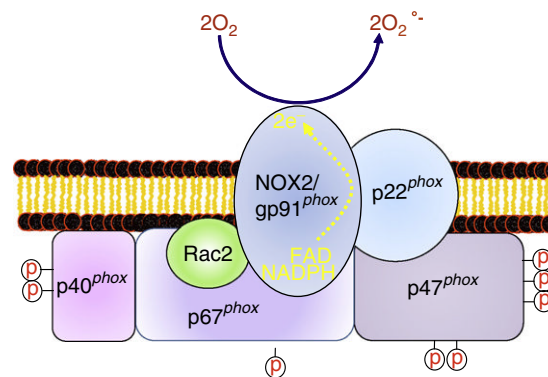


Figure 2 The active NADPH oxidase complex. The active NADPH oxidase is composed of a membrane-bound flavocytochrome b558 (composed of gp91phox and p22phox) which binds several cytosolic proteins (p67phox, p47phox, p40phox, and rac2). Activated NADPH oxidase uses cytosolic NADPH to reduce oxygen. (P) denotes phosphorylation.

'chronic granulomatous disease' (CGD) (Kannengiesser *et al.*, 2008). It is an inherited immune deficiency disorder in which phagocytes are unable to produce ROS and thus to eliminate pathogens. Recurrent, often life-threatening bacterial and fungal infections usually start during childhood. CGD results from mutations in the NADPH oxidase component genes, namely the CYBB gene (Xp21) that encodes the gp91phox subunit, the CYBA gene (16q24) that encodes the p22phox subunit, the NCF1 gene (7q11) and NCF2 gene (1q25), which encode p47phox, p67phox, respectively (Roos *et al.*, 1996). The most frequent form of CGD (approximately 70% of all cases) is the X-linked gp91phox-deficient form, followed by the autosomal form deficient in p47phox (approximately 25%). Less frequent forms are autosomal CGD deficient in p67phox (<5%) and p22phox (<5%). Only one case of rac2 gene mutation and one case of p40phox gene mutation have been described, possibly due to the relatively mild clinical phenotype which may preclude diagnostic recognition.

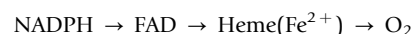
Structure of the Membrane Component, Flavocytochrome b₅₅₈ in Resting Cells

The flavocytochrome b₅₅₈ is the central catalytic membrane-bound component of the NADPH oxidase (Vignais, 2002). It is composed of a glycosylated 91-kDa protein subunit (gp91phox) and a non-glycosylated 22-kDa subunit (p22phox) in a 1:1 complex (Huang *et al.*, 1995; Wallach and Segal, 1996; Figure 3). Both proteins stabilize each other in neutrophils (Yu *et al.*, 1997). The heterodimer is called cytochrome b₅₅₈ for its characteristic spectral absorption peak at 558 nm. In resting cells, the flavocytochrome b₅₅₈ is mostly localized in membranes of specific granules, gelatinase granules, and secretory granules and only 10–20% is present in the plasma membrane (El-Benna *et al.*, 2008). In intact cells, the

flavocytochrome b₅₅₈ alone, without the cytosolic components is in an inactive state not able to transfer electrons.

gp91phox

Gp91phox also called the β -subunit of the flavocytochrome b₅₅₈, is composed of 570 amino acids. This protein of a molecular weight of 65 kDa runs as a broad smear around 91 kDa on sodium dodecyl sulfate-polyacrylamide gel (SDS-PAGE) because it is very glycosylated on extracytoplasmic sites. Gp91phox has a short cytoplasmic N-terminal sequence (Paclet *et al.*, 2004), predicted six transmembrane α -helices which bind two hemes and a C-terminal cytosolic tail which contains one NADPH binding site, and one FAD binding site (Royer-Pokora *et al.*, 1986; Rotrosen *et al.*, 1992). Thus, gp91phox contains all the elements necessary to transfer electrons from NADPH to oxygen, and is the transfer chain of the enzyme. Electrons transfer is believed to occur in this order:



The NADPH oxidase family has been identified by homology search to gp91phox. The family consists of seven members, five NOXs (NOX-1 to NOX-5) and two DUOXs (DUOX1 and DUOX2), (DUOX for dual oxidase) (Bedard and Krause, 2007). Gp91phox was renamed NOX2 which is also used now to designate the phagocyte NADPH oxidase complex. The new NOX and DUOX are expressed in epithelial cells, endothelial cells, vascular smooth muscle cells, and fibroblast of several tissues such as colon, lung, kidney, heart, and thyroid. ROS production by the NOX/DUOX proteins is very low and is believed to be involved in the regulation of several cellular functions such as local tissue-specific bactericidal activity, vascular tone, thyroid-hormones production, and intracellular signaling (Lambeth, 2007; Dupuy *et al.*, 1991).

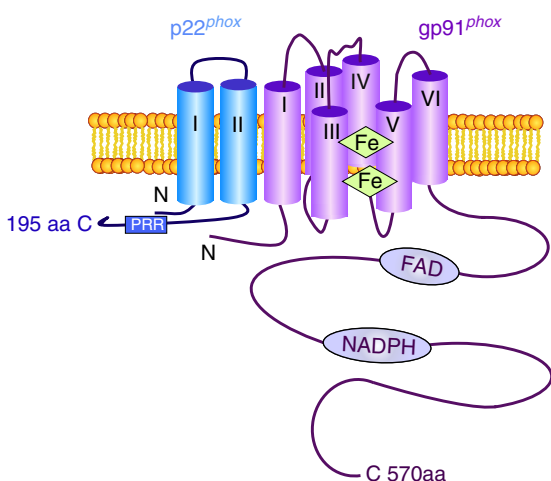


Figure 3 Structure of the flavocytochrome b₅₅₈. The flavocytochrome b₅₅₈ is composed of two transmembrane proteins gp91phox (or NOX2) and p22phox. The gp91phox protein has a short N-terminal cytosolic sequence, six transmembrane helices, and one long C-terminal cytosolic tail containing the FAD binding site and one NADPH binding site. The p22phox protein has an N-terminal cytosolic sequence, two transmembrane helices, and one long C-terminal cytosolic tail containing a PRR which binds to p47phox SH3 domains.

p22phox

P22phox also called the α -subunit of the flavocytochrome b₅₅₈, is composed of 195 amino acids and runs at 22 kDa on SDS-PAGE (Parkos *et al.*, 1988). It has two or three predicted transmembrane α -helices and a C-terminal cytosolic tail which contains one proline rich region (PRR) (Pro156-Pro-Arg-Pro159) which binds to p47phox-src homology 3 (SH3) domains during NADPH oxidase activation (de Mendez *et al.*, 1997). Thus, in addition to its electron transfer activity, cytochrome b₅₅₈ is also the central docking site for the cytosolic components. P22phox by associating with its partner gp91phox is essential for the maturation and stabilization of the flavocytochrome b₅₅₈ (DeLeo *et al.*, 2000).

Structure of the Cytosolic Components in Resting Cells

The cytosolic components of phagocyte NADPH oxidase are p47phox, p67phox, p40phox, and a small G protein (rac1 or rac2). In resting cells p47phox, p67phox, and p40phox interact together to form a cytosolic complex (Iyer *et al.*, 1994; Park *et al.*, 1994), while a guanosine diphosphate (GDP) bound Rac1/2 is in a complex with its inhibitor Rho-GDP-dissociation inhibitor (GDI) (Abo *et al.*, 1994). During activation, approximately 10% of these proteins migrate to the plasma

and phagosome membranes to assemble with cytochrome 558 and to induce its activation (Clark *et al.*, 1990; Quinn *et al.*, 1993; El Benna *et al.*, 1994b).

p47phox

P47phox also called NCF1 (neutrophil cytosolic factor 1), is a protein composed of 390 amino acids which migrates at almost 47 kDa on SDS-PAGE (Lomax *et al.*, 1989; Volpp *et al.*, 1989). Its N-terminal amino acid sequence has one PX (phox homology) domain (amino acids 4–121), its middle amino acid sequence contains two SH3 domains (amino acids 159–214 (SH3A) and amino acids 229–284 (SH3B)). The C-terminal sequence is very basic, is rich of serines and arginines and has one region called the autoinhibitory region (AIR) (amino acids 292–340) and a PRR (amino acids 363–368) (Groemping and Rittinger, 2005; El-Benna *et al.*, 2009). In resting cells, some of p47phox may be in a complex with p67phox and p40phox and some is free in the cytosol. In the free form, the two p47phox-SH3 domains interact intramolecularly with the C-terminal region of the non-phosphorylated protein to keep p47phox in an auto-inhibited state (Leto *et al.*, 1994; Sumimoto *et al.*, 1994; Groemping *et al.*, 2003). For this reason the C-terminal region is called the AIR. The X-ray structure of the auto-inhibited form of p47phox (lacking the PX domain) reveals that both SH3 domains form a common interface to create a single shallow groove that constitutes the peptide binding surface (Groemping *et al.*, 2003; Yuzawa *et al.*, 2003; Ogura *et al.*, 2006), thereby acting as a single ‘super’ SH3 domain. In the auto-inhibited form, this site is occupied by the AIR polybasic region of p47phox. Simultaneously, this intramolecular interaction allows the formation of a binding surface on SH3A for the PX domain (Groemping and Rittinger, 2005; Marcoux *et al.*, 2010) to hold the entire multi-domain protein in an inactive conformation. P47phox binds to cytochrome b₅₅₈ during activation. In the activated form the p47phox tandem SH3 domains are occupied by the PRR of the cytoplasmic domain of p22phox instead of the AIR region. The crystal structure of the isolated p47phox-PX domain simultaneously bound to phosphatidylinositol 3,4-bisphosphate and phosphatidic acid revealed that the two lipids bound to two distinct basic pockets in the p47phox domain (Karathanassis *et al.*, 2002) and explained the specificity of this binding (Yaffe, 2002). This study further showed that a point mutation of the second SH3 domain or the phosphorylation of p47phox induced a transition from a closed to an open conformation that binds membranes through the PX domain.

p67phox

P67phox also called NCF2 (Neutrophil Cytosolic Factor 2), is a protein composed of 526 amino acids which migrates at almost 67 kDa on SDS-PAGE (Leto *et al.*, 1990). Its N-terminal amino acid sequence has four tetratricopeptide-rich regions (TPR) (residues (3–36), (37–70), (71–104), (121–154)) that contribute to Rac binding; its middle amino acid sequence contains an activation domain (AD) (residues 199–210), a PRR (residues 226–236), and one SH3 domains (amino acids 243–298(SH3A)) (Han *et al.*, 1998; Yuzawa *et al.*, 2009). The C-terminal sequence has a PB1(phox and Bem1) domain or PC (Phox and Cdc24p) domain (amino acids 351–429) and a second SH3 domain (amino acids 460–514(SH3B)). In resting

cells, all of p67phox is believed to be in a complex with p47phox in the cytosol (Park *et al.*, 1994; Ellson *et al.*, 2001; Kanai *et al.*, 2001). The X-ray structure of full-length p67phox is not known, however, the crystal structure the N-terminal domain in free form and in a complex with Rac–guanosine-5'-triphosphate (GTP) have been described (Lapouge *et al.*, 2000; Grizot *et al.*, 2001a,b). Both studies show that the TPR domains adopt a superhelix fold with a hydrophobic groove ending with a short helix. Recently, structural analysis of p67phox by small angle X-ray scattering (SAXS) has suggested the existence of an intramolecular interaction which might engage the N-terminal SH3 domain (SH3A) and the TPR domains of p67phox (Durand *et al.*, 2010).

p40phox

P40phox also called NCF4 (Neutrophil Cytosolic Factor 4), is a 336-amino-acid protein which migrates at 40 kDa on SDS-PAGE (Wientjes *et al.*, 1993; Tsunawaki *et al.*, 1994). Its N-terminal amino acid sequence has a PX domain (residues 18–136), its middle amino acid sequence contains one SH3 domains (amino acids 173–228) and the C-terminal sequence has a PB1/PC domain (Matute *et al.*, 2005). In resting cells, all of p40phox is in a complex with p67phox in the cytosol. Indeed, it was initially identified through its binding to p67phox. P40phox binding to p67phox strongly stabilizes the protein as its expression is significantly reduced in p67phox deficient CGD patients (Tsunawaki *et al.*, 1994; Kuribayashi *et al.*, 2002). The crystal structure of the full-length p40phox has been described (Honbou *et al.*, 2007). It revealed an intramolecular interaction between the PB1 and the PX domain, however, the interaction interface localizes on the opposite side of that for the p67phox PB1 domain, thus still allowing the PB1–PB1 interaction between p40phox and p67phox. The crystal structure of the p40phox-PX domain bound to phosphatidylinositol 3-phosphate reveals the phospholipid interaction site in p40phox and explained the specificity for 3-phosphoinositide (Bravo *et al.*, 2001; Ellson *et al.*, 2001; Kanai *et al.*, 2001).

Rac

The small GTPase Rac also called (NCF3) is composed of 192 amino acids. In human neutrophils, p21-rac2 is the most abundant rac isoform, but p21-rac1 is also present. However, in monocytes and macrophages Rac1 is the major expressed isoform (Knaus *et al.*, 1991; Abo *et al.*, 1991). Rac1 and Rac2 are 92% identical, they differ by only 14 residues and both can activate the NADPH oxidase. Like cytochrome b₅₅₈, p47phox and p67phox, Rac is essential for optimal NADPH oxidase activation in a cell-free system and in intact neutrophils. In resting cells, all of Rac1 and Rac2 in their GDP-bound form, are in a complex with their natural inhibitor Rho-GDI (Abo *et al.*, 1994; Grizot *et al.*, 2001a,b).

The ‘p47phox–p67phox–p40phox’ complex in resting cells

In the cytosol of resting cells, p47phox, p67phox, and p40phox can associate in a trimer complex with a 1:1:1 stoichiometry (Iyer *et al.*, 1994; Park *et al.*, 1994; Lapouge *et al.*, 2002). However p47phox may exist separately from the p67phox–p40phox dimer complex (Ellson *et al.*, 2001; Kanai *et al.*, 2001). In the complex, p67phox plays a central role as it bridges

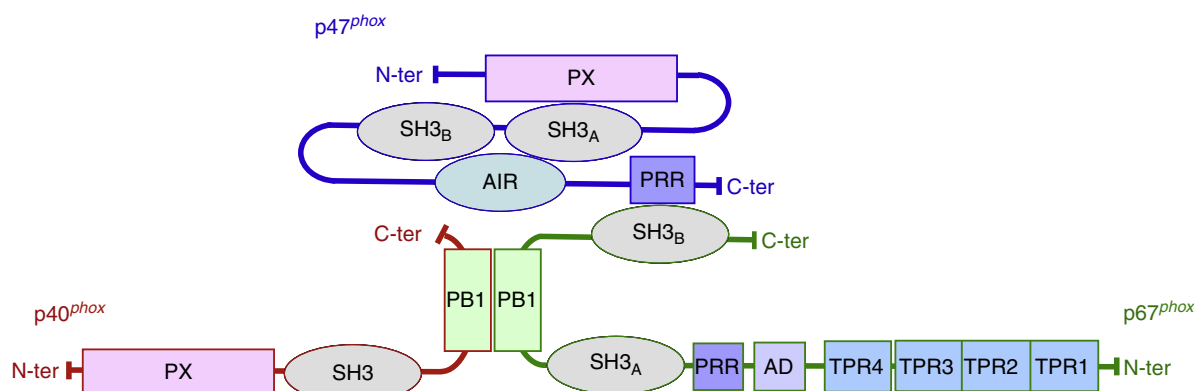


Figure 4 Structure of the cytosolic complex. p67phox interacts with p47phox and p40phox. The p47phox PRR interacts with the p67phox-SH3 domain while p40phox interacts with p67phox via a constitutive PB1/PB1 interaction.

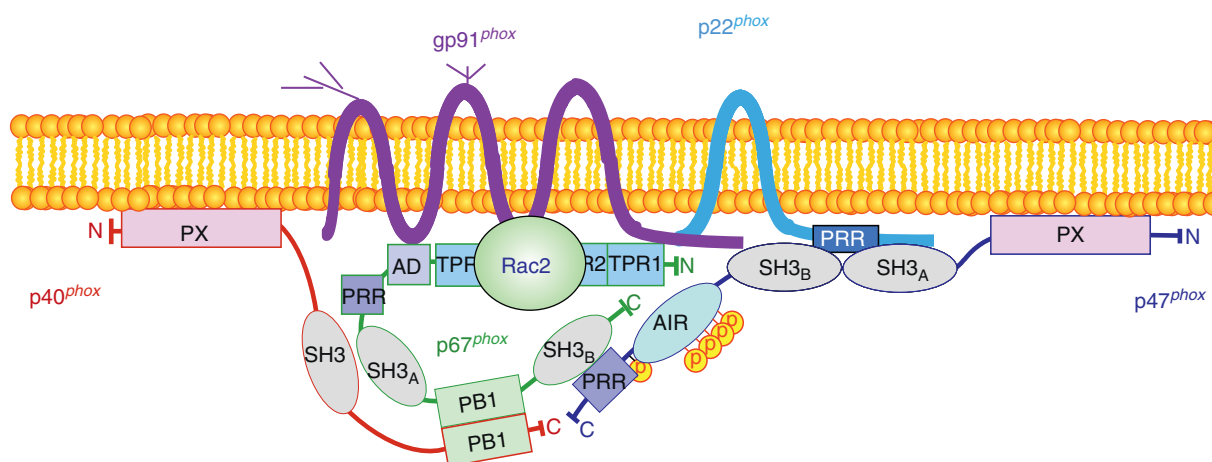


Figure 5 Interactions in the assembled NADPH oxidase complex. Activated NADPH oxidase is composed gp91phox, p22phox, p47phox, p67phox, p40phox, and Rac1/2. The p22phox binds to p47phox-SH3 domains. P67phox binds to p47phox, p40phox, Rac, and gp91phox. Rac binds to gp91phox and p67phox. The p47phox- and the p40-phox PX domains interact with membrane phospholipids. p67phox via its AD promotes gp91phox enzymatic activity.

p40phox to p47phox, the PB1 domain of p67phox binds to the PB1 domain of p40phox while the C-terminal SH3 domain of p67phox binds to the PRR of p47phox (Figure 4). As we described above, the free and complexed p47phox is in a closed auto-inhibited state since its AIR (amino acids 292–340) interacts with the two p47phox-SH3 domains to keep p47phox and NOX2 in a resting state which prevents binding of p47phox to p22phox (Groemping *et al.*, 2003).

Assembly and Activation of the NADPH Oxidase Complex

In neutrophils and monocytes, NADPH oxidase activation can be induced by a large number of soluble and particulate agents such as opsonized bacteria, opsonized zymosan, formylated peptides such as formyl-Met-Leu-Phe (fMLF), calcium ionophores (ionomycin, A23187), and the protein kinase C (PKC) activator phorbol myristate acetate (PMA) (El-Benna *et al.*, 2008). These agents induce p47phox

phosphorylation on serines localized in the AIR region and translocation of cytosolic components to the membranes (Figure 5).

NADPH oxidase can be also activated *in vitro* by a mixture of neutrophil cytosol and plasma membranes or the corresponding recombinant proteins in the presence of Mg^{2+} , GTP, and an activating anionic amphiphile such as arachidonate or SDS, which are believed to mimic phosphorylation by providing negative charges (Bromberg and Pick, 1985). This cell-free system can also be activated by protein kinase C instead of anionic amphiphile agents (El-Benna *et al.*, 1995).

Indeed introduction of negative charges by phosphorylation prevents the 'AIR/tandem SH3 domains' interaction and converts p47phox to an open active complex. The tandem SH3 domains then bind the PRR localized in the cytosolic tail of p22phox (Fontayne *et al.*, 2002), while the p47phox PX domain binds to phospholipids in the membranes (Kanai *et al.*, 2001; Ellson *et al.*, 2001) and other p47phox-regions interact

with gp91phox (DeLeo *et al.*, 1995a,b) in order to ensure optimal positioning of the p67phox AD which induces the activation of electron transfer by gp91phox. Indeed p67phox is able to bind gp91phox *in vitro* and p47phox increases this binding (Dang *et al.*, 2001, 2002; Figure 5).

P47phox is the first cytosolic component to interact with cytochrome b₅₅₈ during activation in order to bring the cytosolic complex (p47phox–p67phox–p40phox) to the membrane (Heyworth *et al.*, 1991). Simultaneously, the Rac protein is believed to bind to gp91phox by its Rac insert region (Diebold and Bokoch, 2001) then to p67phox via the TPR domains. These interactions are believed to stimulate electron transfer by gp91phox as does the direct interaction of the AD of p67phox to gp91phox (Han *et al.*, 1998; Dang *et al.*, 2001). The p40phox–PX domain binds also to phospholipids in the membranes in order to stabilize, and localize the assembled complex in internalized phagosomes (Ellson *et al.*, 2001; Kanai *et al.*, 2001; Zhan *et al.*, 2002).

Regulation of NADPH Oxidase Assembly and Activation by Phosphorylation

Phosphorylation of p47phox

Activation of the NADPH oxidase in cells is accompanied by phosphorylation of its cytosolic subunits p47phox, p67phox, and p40phox (El Benna *et al.*, 1994a; El-Benna *et al.*, 1997; Bouin *et al.*, 1998) and also its membrane subunits p22phox and gp91phox (Regier *et al.*, 1999; Raad *et al.*, 2009). However, only the role of p47phox phosphorylation is extensively studied (El-Benna *et al.*, 2009). Indeed, p47phox was first identified because its phosphorylated form was missing in neutrophils isolated from some CGD-patients (Segal *et al.*, 1985). In resting cells, p47phox is not phosphorylated, in stimulated cells eight to nine p47phox isoforms were seen due to its phosphorylation (Rotrosen and Leto, 1990). Phosphorylation analysis showed that p47phox is phosphorylated on multiple sites in the C-terminal portion of the protein, including serines 303 to 379 (El Benna *et al.*, 1994a,1996). Combined mutation of all these serines was used to show that p47phox phosphorylation is required for NADPH oxidase activation in intact cells (Faust *et al.*, 1995; Belambri *et al.*, 2012). Individual mutation of each serine showed that only serine 379 was important for oxidase activation (Faust *et al.*, 1995). Double mutation analysis showed that two pairs of phosphorylated serines, Ser(303 + 304) and Ser(359 + 370), are necessary for NADPH oxidase activation (Johnson *et al.*, 1998; Inanami *et al.*, 1998).

The ‘phosphorylated–p47phox–Pin1’ axis

Phosphorylation of p47phox or anionic amphiphiles that mimics phosphorylation, induce a conformational changes of the protein that prevent the interaction between the AIR and the SH3 domains (Fontayne *et al.*, 2002; Huang and Kleinberg, 1999). This open-p47phox state initiates assembly of the active enzyme via interaction of the SH3 domains with the PRR of p22phox. Phosphorylated p47phox also increases the interaction between p67phox and gp91phox (Dang *et al.*, 2002). Most of these results were obtained in a cell free system with recombinant p47phox which was heavily phosphorylated

in vitro by PKC (Fontayne *et al.*, 2002). In the context of intact neutrophils, NOX2 hyper-activation occurs when cells encounter first a priming agent then the stimulating agent. We have shown that the proline isomerase Pin1 is likely to be a major enzyme involved in NOX2 hyperactivation through the regulation of p47phox hyper-phosphorylation (Boussetta *et al.*, 2010). Indeed, the pro-inflammatory cytokine tumor necrosis factor- α (TNF α), does not activate NADPH oxidase but primes its activation in response to secondary stimulus such as N-Formyl-Met-Leu-Phe (FMLP) (El-Benna *et al.*, 2008; Brown *et al.*, 2004). This phenomenon is critical in the physiological regulation of microbicidal activity and during tissue injury in inflammatory diseases (Condliffe *et al.*, 1998; El-Benna *et al.*, 2008). TNF α induces partial phosphorylation of p47phox on Ser345 and this mechanism is critical for the priming of ROS production in neutrophils (Dang *et al.*, 2006). TNF α also induced the activation of the proline isomerase Pin1 followed by its binding to p47phox, only when this last one is phosphorylated on Ser-345. Pin1 then catalyzed a conformational change of p47phox that facilitate subsequent phosphorylation of the protein on other sites by PKC and consequently allowed NOX2 hyperactivation (Boussetta *et al.*, 2010). Thus the ‘p47phox–Ser345 phosphorylation–Pin1’ axis controls TNF α -induced NADPH oxidase hyperactivation in human neutrophils. This pathway was also involved in TLR7/8-mediated neutrophil NADPH oxidase priming (Makni-Maalej *et al.*, 2012).

Phosphorylation of p67phox and p40phox

The p67phox is also phosphorylated in human neutrophils, although at lower level than p47phox (El-Benna *et al.*, 1997). P67phox is constitutively phosphorylated in resting human neutrophils, a process that can be enhanced by stimulation of cells (Dang *et al.*, 2011). This phosphorylation occurs on serine and threonine residues (El-Benna *et al.*, 1997; Forbes *et al.*, 1999). Several protein kinases such as PKC, ERK1/2, and p38MAPK can phosphorylate p67phox (Dang *et al.*, 2003). Up until now one p67phox phosphorylated site corresponding to threonine 233 has been identified (Forbes *et al.*, 1999), however, phosphorylation of other sites has not been entirely ruled out. Interestingly phosphorylation studies have suggested that p67phox may exist in an auto-inhibited state (Dang *et al.*, 2003). Indeed cryptic phosphorylation sites localized in the C-terminal portion of p67phox become accessible to ERK2 only after the removal of the N-terminal fragment of the protein. Furthermore, the addition of the N-terminal fragment inhibited the phosphorylation reaction (Dang *et al.*, 2003). The presence of the intramolecular interaction in p67phox was also suggested by the SAXS analysis of p67phox (Durand *et al.*, 2010). Phosphorylation of p67phox does not affect its binding to gp91phox *in vitro*. Its role can be further investigated by the identification of all the phosphorylated sites and site directed mutagenesis.

P40phox is also phosphorylated in neutrophils and neutrophil-like HL-60 cells. It is phosphorylated on two major sites Thr154 and Ser315 (Bouin *et al.*, 1998). Mutation of these residues did not affect the interaction with p67phox and p47phox in the two hybrid system (Bouin *et al.*, 1998). However, it was shown that phosphorylated p40phox increased the binding of p67phox to cytochrome b₅₅₈ (Dang *et al.*, 2002). In another study, the addition of phosphorylated

p40phox in a PKC-activated cell-free system inhibited NADPH oxidase activation (Lopes *et al.*, 2004). It was suggested that phosphorylation of Thr154 could induce a conformational changes that induce an inhibitory form of p40phox. However, recent studies using wild type and mutated version of p40phox introduced into fully differentiated mouse neutrophils by retroviral transduction of p40phox $-/-$ bone marrow progenitors has shown that phosphorylation of threonine 154 in p40phox is an important new element in the physiological activation of NOX2 (Chessa *et al.*, 2010) that remains to be further explored.

Phosphorylation of gp91phox and p22phox

Gp91phox/NOX2 is phosphorylated in human neutrophils activated by several agents known to induce NADPH oxidase activation (PMA, fMLP, and opsonized zymosan) (Raad *et al.*, 2009). In addition, this phosphorylation occurs in the C-terminal region containing the flavoprotein domain. PKC phosphorylation increased electron transfer activity of the gp91phox flavoprotein cytosolic domain and its binding to Rac2, p67phox, and p47phox (Raad *et al.*, 2009). These data suggest that phosphorylation of gp91phox could play a stimulatory role for NADPH oxidase. The p22phox subunit is phosphorylated in stimulated human neutrophils (Regier *et al.*, 1999). The major p22phox phosphorylated site was identified as Thr147 (Lewis *et al.*, 2010). Mutation of this site to Ala induced inhibition of NADPH oxidase activation and p22phox/p47phox interaction. These results indicate that phosphorylation of p22phox on Thr147 is critical for complex assembly and activation.

Conclusion

ROS production by the phagocyte NOX2 is essential for host defense against invading pathogens. However, its hyperactivation can induce tissue injury involved in inflammatory disease. This multicomponent enzyme is tightly regulated to ensure activation when required, and deactivation to limit excessive ROS production. The subunits of this multicomponent enzyme are today fully identified. At the exception of the Rac protein and p40phox, only the structure of some domains of the other subunits have been described. Some of the signaling pathways have been identified. Future studies are required to determine the structure of full-length gp91phox, p22phox, p47phox, and p67phox, and the assembled complex, the role of the phosphorylation of p67phox, p40phox, and gp91phox, the regulation of the activation/deactivation in intact cells, the signaling pathways that control the relative amounts of intracellular versus extracellular ROS, and the identification of novel specific inhibitors for anti-inflammatory therapeutic purposes.

Acknowledgments

This work was supported by grants from Agence Nationale de la Recherche (ANR), Arthritis Fondation Courtin and VLM (Vaincre La Mucoviscidose), Labex Inflamex and the National Institutes of Health.

See also: Intracellular Infectiology: Cell Processes: Microbicidal Mechanisms

References

- Abo, A., Pick, E., Hall, A., *et al.*, 1991. Activation of the NADPH oxidase involves the small GTP-binding protein p21rac1. *Nature* 353, 668–670.
- Abo, A., Webb, M.R., Grogan, A., Segal, A.W., 1994. Activation of NADPH oxidase involves the dissociation of p21rac from its inhibitory GDP/GTP exchange protein (rhoGDI) followed by its translocation to the plasma membrane. *Biochemical Journal* 298, 585–591.
- Babior, B.M., 1999. NADPH oxidase: An update. *Blood* 93, 1464–1476.
- Babior, B.M., 2000. Phagocytes and oxidative stress. *American Journal of Medicine* 109, 33–44.
- Bedard, K., Krause, K.H., 2007. The NOX family of ROS-generating NADPH oxidases: Physiology and pathophysiology. *Physiological Reviews* 87, 245–313.
- Belambri, S.A., Hurtado-Nedelec, M., Senator, A., *et al.*, 2012. Phosphorylation of p47phox is required for receptor-mediated NADPH oxidase/NOX2 activation in Epstein-Barr virus-transformed human B lymphocytes. *American Journal of Blood Research* 2, 187–193.
- Bouin, A.P., Grandvaux, N., Vignais, P.V., Fuchs, A., 1998. p40(phox) is phosphorylated on threonine 154 and serine 315 during activation of the phagocyte NADPH oxidase. Implication of a protein kinase C-type kinase in the phosphorylation process. *Journal of Biological Chemistry* 273, 30097–30103.
- Boussetta, T., Gougerot-Pocidalo, M.A., Hayem, G., *et al.*, 2010. The prolyl isomerase Pin1 acts as a novel molecular switch for TNF-alpha-induced priming of the NADPH oxidase in human neutrophils. *Blood* 116, 5795–5802.
- Bravo, J., Karathanassis, D., Pacold, C.M., *et al.*, 2001. The crystal structure of the PX domain from p40(phox) bound to phosphatidylinositol 3-phosphate. *Molecular Cell* 8, 829–839.
- Bromberg, Y., Pick, E., 1985. Activation of NADPH-dependent superoxide production in a cell-free system by sodium dodecyl sulfate. *Journal of Biological Chemistry* 260, 13539–13545.
- Brown, G.E., Stewart, M.Q., Bissonnette, S.A., *et al.*, 2004. Distinct ligand-dependent roles for p38 MAPK in priming and activation of the neutrophil NADPH oxidase. *Journal of Biological Chemistry* 279, 27059–27068.
- Chanock, S.J., El Benna, J., Smith, R.M., Babior, B.M., 1994. The respiratory burst oxidase. *Journal of Biological Chemistry* 269, 24519–24522.
- Chessa, T.A., Anderson, K.E., Hu, Y., *et al.*, 2010. Phosphorylation of threonine 154 in p40phox is an important physiological signal for activation of the neutrophil NADPH oxidase. *Blood* 116, 6027–6036.
- Clark, R.A., Volpp, B.D., Leidal, K.G., Nauseef, W.M., 1990. Two cytosolic components of the human neutrophil respiratory burst oxidase translocate to the plasma membrane during cell activation. *Journal of Clinical Investigation* 85, 714–721.
- Condliffe, A.M., Kitchen, E., Chilvers, E.R., 1998. Neutrophil priming: Pathophysiological consequences and underlying mechanisms. *Clinical Science (London)* 94, 461–471.
- Dang, P.M., Morel, F., Gougerot-Pocidalo, M.A., Benna, J.E., 2003. Phosphorylation of the NADPH oxidase component p67(PHOX) by ERK2 and P38MAPK: Selectivity of phosphorylated sites and existence of an intramolecular regulatory domain in the tetratricopeptide-rich region. *Biochemistry* 42, 4520–4526.
- Dang, P.M., Raad, H., Derkawi, R.A., *et al.*, 2011. The NADPH oxidase cytosolic component p67phox is constitutively phosphorylated in human neutrophils: Regulation by a protein tyrosine kinase, MEK1/2 and phosphatases 1/2 A. 1. *Biochemical Pharmacology* 82, 1145–1152.
- Dang, P.M., Stensballe, A., Boussetta, T., *et al.*, 2006. A specific p47phox -serine phosphorylated by convergent MAPKs mediates neutrophil NADPH oxidase priming at inflammatory sites. *Journal of Clinical Investigation* 116, 2033–2043.
- Dang, P.M.C., Cross, A.R., Babior, B.M., 2001. Assembly of the neutrophil respiratory burst oxidase: A direct interaction between p67phox and cytochrome b558. *Proceedings of the National Academy of Sciences of the United States of America* 98, 3001–3005.
- Dang, P.M.C., Cross, A.R., Quinn, M.T., Babior, B.M., 2002. Assembly of the neutrophil respiratory burst oxidase: A direct interaction between p67phox and cytochrome b558 II. *Proceedings of the National Academy of Sciences of the United States of America* 99, 4262–4265.

- DeLeo, F.R., Burritt, J.B., Yu, L., *et al.*, 2000. Processing and maturation of flavocytochrome b558 include incorporation of heme as a prerequisite for heterodimer assembly. *Journal of Biological Chemistry* 275, 13986–13993.
- DeLeo, F.R., Nauseef, W.M., Jesaitis, A.J., *et al.*, 1995a. A domain of p47phox that interacts with human neutrophil flavocytochrome b558. *Journal of Biological Chemistry* 270, 26246–26251.
- DeLeo, F.R., Quinn, M.T., 1996. Assembly of the phagocyte NADPH oxidase: Molecular interaction of oxidase proteins. *Journal of Leukocyte Biology* 60, 677–690.
- DeLeo, F.R., Yu, L., Burritt, J.B., *et al.*, 1995b. Mapping sites of interaction of p47-phox and flavocytochrome b with random-sequence peptide phage display libraries. *Proceedings of the National Academy of Sciences of the United States of America* 92, 7110–7114.
- Diebold, B.A., Bokoch, G.M., 2001. Molecular basis for Rac2 regulation of phagocyte NADPH oxidase. *Nature Immunology* 2, 211–215.
- Dupuy, C., Virion, A., Ohayon, R., *et al.*, 1991. Mechanism of hydrogen peroxide formation catalyzed by NADPH oxidase in thyroid plasma membrane. *Journal of Biological Chemistry* 266, 3739–3743.
- Durand, D., Vivès, C., Cannella, D., *et al.*, 2010. NADPH oxidase activator p67 (phox) behaves in solution as a multidomain protein with semi-flexible linkers. *Journal of Structural Biology* 169, 45–53.
- El-Benna, J., Dang, P.M., Gaudry, M., *et al.*, 1997. Phosphorylation of the respiratory burst oxidase subunit p67(phox) during human neutrophil activation. Regulation by protein kinase C-dependent and independent pathways. *Journal of Biological Chemistry* 272, 17204–17208.
- El-Benna, J., Dang, P.M., Gougerot-Pocidalo, M.A., 2008. Priming of the neutrophil NADPH oxidase activation: Role of p47phox phosphorylation and NOX2 mobilization to the plasma membrane. *Seminars in Immunopathology* 30, 279–289.
- El-Benna, J., Dang, P.M., Gougerot-Pocidalo, M.A., Marie, J.C., Braut-Boucher, F., 2009. p47phox, the phagocyte NADPH oxidase/NOX2 organizer: Structure, phosphorylation and implication in diseases. *Experimental Molecular Medicine* 41, 217–225.
- El-Benna, J., Dang, P.M.C., Gougerot-Pocidalo, M.A., Elbim, C., 2005. Phagocyte NADPH oxidase: A multicomponent enzyme essential for host defenses. *Archivum Immunologiae et Therapiae Experimentalis* 53, 199–206.
- El Benna, J., Faust, L.P., Babior, B.M., 1994a. The phosphorylation of the respiratory burst oxidase component p47phox during neutrophil activation. Phosphorylation of sites recognized by protein kinase C and by proline-directed kinases. *Journal of Biological Chemistry* 269, 23431–23436.
- El Benna, J., Faust, L.R.P., Johnson, J., Babior, B.M., 1996. Phosphorylation of the respiratory burst oxidase subunit p47phox as determined by two-dimensional phosphopeptide mapping. Phosphorylation by protein kinase C, protein kinase A and a mitogen activated protein kinase. *Journal of Biological Chemistry* 271, 6374–6378.
- El-Benna, J., Park, J.W., Ruedi, J.M., Babior, B.M., 1995. Cell-free activation of the respiratory burst oxidase by protein kinase C. *Blood Cells, Molecules and Diseases* 21, 201–206.
- El Benna, J., Ruedi, J.M., Babior, B.M., 1994b. Cytosolic guanine nucleotide-binding protein Rac 2 operates *in vivo* as a component of the neutrophil respiratory burst oxidase. Transfer of Rac 2 and the cytosolic oxidase components p47 (phox) and p67 (phox) to the submembranous actin cytoskeleton during oxidase activation. *Journal of Biological Chemistry* 269, 6729–6734.
- Ellson, C.D., Gobert-Gosse, S., Anderson, K.E., *et al.*, 2001. PtdIns(3)P regulates the neutrophil oxidase complex by binding to the PX domain of p40(phox). *Nature Cell Biology* 3, 679–682.
- Faust, L.P., El Benna, J., Babior, B.M., Chanock, S.J., 1995. The phosphorylation targets of p47 phox a subunit of the respiratory burst oxidase. Functions of the individual target serines as evaluated by site-directed mutagenesis. *Journal of Clinical Investigation* 96, 1499–1505.
- Fontayne, A., Dang, P.M., Gougerot-Pocidalo, M.A., El-Benna, J., 2002. Phosphorylation of p47phox sites by PKC alpha, beta II, delta, and zeta: Effect on binding to p22phox and on NADPH oxidase activation. *Biochemistry* 41, 7743–7750.
- Forbes, L.V., Truong, O., Wientjes, F.B., Moss, S.J., Segal, A.W., 1999. The major phosphorylation site of the NADPH oxidase component p67phox is Thr233. *Biochemical Journal* 338, 99–105.
- Grizot, S., Fauré, J., Fieschi, F., *et al.*, 2001a. Crystal structure of the Rac1-RhoGDI complex involved in nadph oxidase activation. *Biochemistry* 40, 10007–10013.
- Grizot, S., Fieschi, F., Dagher, M.C., Pebay-Peyroula, E., 2001b. The active N-terminal region of p67phox. Structure at 1.8 Å resolution and biochemical characterizations of the A128V mutant implicated in chronic granulomatous disease. *Journal of Biological Chemistry* 276, 21627–21631.
- Groemping, Y., Lapouge, K., Smerdon, S.J., Rittinger, K., 2003. Molecular basis of phosphorylation-induced activation of the NADPH oxidase. *Cell* 113, 343–355.
- Groemping, Y., Rittinger, K., 2005. Activation and assembly of the NADPH oxidase: A structural perspective. *Biochemical Journal* 386, 401–416.
- Hampton, M.B., Kettle, A.J., Winterbourn, C.C., 1998. Inside the neutrophil phagosome: Oxidants, myeloperoxidase, and bacterial killing. *Blood* 92, 3007–3017.
- Han, C.H., Freeman, J.L., Lee, T., Motalebi, S.A., Lambeth, J.D., 1998. Regulation of the neutrophil respiratory burst oxidase. Identification of an activation domain in p67(phox). *Journal of Biological Chemistry* 273, 16663–16668.
- Heyworth, P.G., Curnutte, J.T., Nauseef, W.M., *et al.*, 1991. Neutrophil nicotinamide adenine dinucleotide phosphate oxidase assembly. Translocation of p47-phox and p67-phox requires interaction between p47-phox and cytochrome b558. *Journal of Clinical Investigation* 87, 352–356.
- Honbou, K., Minakami, R., Yuzawa, S., *et al.*, 2007. Full-length p40phox structure suggests a basis for regulation mechanism of its membrane binding. *EMBO Journal* 26, 1176–1186.
- Huang, J., Hitt, N.D., Kleinberg, M.E., 1995. Stoichiometry of p22-phox and gp91-phox in phagocyte cytochrome b558. *Biochemistry* 34, 16753–16757.
- Huang, J., Kleinberg, M.E., 1999. Activation of the phagocyte NADPH oxidase protein p47(phox). Phosphorylation controls SH3 domain-dependent binding to p22(phox). *Journal of Biological Chemistry* 274, 19731–19737.
- Inanami, O., Johnson, J.L., McAdara, J.K., *et al.*, 1998. Activation of the leukocyte NADPH oxidase by phorbol ester requires p47phox phosphorylation on serine 303 or 304. *Journal of Biological Chemistry* 273, 9539–9543.
- Iyer, S.S., Pearson, D.W., Nauseef, W.M., Clark, R.A., 1994. Evidence for a readily dissociable complex of p47phox and p67phox in cytosol of unstimulated human neutrophils. *Journal of Biological Chemistry* 269, 22405–22511.
- Johnson, J.L., Park, J.W., El Benna, J., *et al.*, 1998. Activation of p47phox, a cytosolic subunit of the leukocyte NADPH oxidase. Phosphorylation of ser359 or 370 precedes phosphorylation at other sites and is required for activity. *Journal of Biological Chemistry* 273, 35147–35152.
- Kanai, F., Liu, H., Field, S.J., *et al.*, 2001. The PX domains of p47phox and p40phox bind to lipid products of PI(3)K. *Nature Cell Biology* 3, 675–678.
- Kannengiesser, C., Gérard, B., El Benna, J., *et al.*, 2008. Molecular epidemiology of chronic granulomatous disease in a series of 80 kindreds: Identification of 31 novel mutations. *Human Mutation* 29, E132–E149.
- Karathanassis, D., Stahelin, R.V., Bravo, J., *et al.*, 2002. Binding of the PX domain of p47(phox) to phosphatidylinositol 3,4-bisphosphate and phosphatidic acid is masked by an intramolecular interaction. *EMBO Journal* 21, 5057–5068.
- Knaus, U.G., Heyworth, P.G., Evans, T., Curnutte, J.T., Bokoch, G.M., 1991. Regulation of phagocyte oxygen radical production by the GTP-binding protein Rac 2. *Science* 254, 1512–1515.
- Kuribayashi, F., Nunoi, H., Wakamatsu, K., *et al.*, 2002. The adaptor protein p40 (phox) as a positive regulator of the superoxide-producing phagocyte oxidase. *EMBO Journal* 21, 6312–6320.
- Lambeth, J.D., 2007. Nox enzymes, ROS, and chronic disease: An example of antagonistic pleiotropy. *Free Radical Biology and Medicine* 43, 332–347.
- Lapouge, K., Smith, S.J., Groemping, Y., Rittinger, K., 2002. Architecture of the p40-p47-p67phox complex in the resting state of the NADPH oxidase. A central role for p67phox. *Journal of Biological Chemistry* 277, 10121–10128.
- Lapouge, K., Smith, S.J., Walker, P.A., *et al.*, 2000. Structure of the TPR domain of p67phox in complex with RacGTP. *Molecular Cell* 6, 899–907.
- Lee, W.L., Harrison, R.E., Grinstein, S., 2003. Phagocytosis by neutrophils. *Microbes Infection* 5, 1299–1306.
- Leto, T.L., Adams, A.G., de Mendez, I., 1994. Assembly of the phagocyte NADPH oxidase: Binding of Src homology 3 domains to proline-rich targets. *Proceedings of the National Academy of Sciences of the United States of America* 91, 10650–10654.
- Leto, T.L., Lomax, K.J., Volpp, B.D., *et al.*, 1990. Cloning of a 67-kD neutrophil oxidase factor with similarity to a noncatalytic region of p60c-src. *Science* 248, 727–730.
- Lewis, E.M., Sergeant, S., Ledford, B., *et al.*, 2010. Phosphorylation of p22phox on threonine 147 enhances NADPH oxidase activity by promoting p47phox binding. *Journal of Biological Chemistry* 285, 2959–2967.
- Lomax, K.J., Leto, T.L., Nunoi, H., Gallin, J.I., Malech, H.L., 1989. Recombinant 47-kilodalton cytosol factor restores NADPH oxidase in chronic granulomatous disease. *Science* 245, 409–412. Erratum in: *Science* 1989 Nov 24; 246(4933): 987.
- Lopes, L.R., Dagher, M.C., Gutierrez, A., *et al.*, 2004. Phosphorylated p40PHOX as a negative regulator of NADPH oxidase. *Biochemistry* 43, 3723–3730.
- Makni-Maalej, K., Boussetta, T., Hurtado-Nedelec, M., *et al.*, 2012. The TLR7/8 agonist CL097 primes N-formyl-methionyl-leucyl-phenylalanine-stimulated

- NADPH oxidase activation in human neutrophils: Critical role of p47phox phosphorylation and the proline isomerase Pin1. *Journal of Immunology* 189, 4657–4665.
- Marcoux, J., Man, P., Petit-Haertlein, I., *et al.*, 2010. p47phox molecular activation for assembly of the neutrophil NADPH oxidase complex. *Journal of Biological Chemistry* 285, 28980–28990.
- Matute, J.D., Arias, A.A., Dinuer, M.C., Patiño, P.J., 2005. p40phox: The last NADPH oxidase subunit. *Blood Cells Molecules and Diseases* 35, 291–302.
- de Mendez, I., Homayounpour, N., Leto, T.L., 1997. Specificity of p47phox SH3 domain interactions in NADPH oxidase assembly and activation. *Molecular and Cellular Biology* 17, 2177–2185.
- Nathan, C., 2006. Neutrophils and immunity: Challenges and opportunities. *Nature Review of Immunology* 6, 173–182.
- Nauseef, W.M., 2007. How human neutrophils kill and degrade microbes: An integrated view. *Immunological Reviews* 219, 88–102.
- Ogura, K., Nobuhisa, I., Yuzawa, S., *et al.*, 2006. NMR solution structure of the tandem Src homology 3 domains of p47phox complexed with a p22phox-derived proline-rich peptide. *Journal of Biological Chemistry* 281, 3660–3668.
- Paclet, M.H., Henderson, L.M., Campion, Y., Morel, F., Dagher, M.C., 2004. Localization of Nox2 N-terminus using polyclonal antipeptide antibodies. *Biochemical Journal* 382, 981–986.
- Park, J.W., El Benna, J., Scott, K.E., *et al.*, 1994. Isolation of a complex of respiratory burst oxidase components from resting neutrophil cytosol. *Biochemistry* 33, 2907–2911.
- Parkos, C.A., Dinuer, M.C., Walker, L.E., *et al.*, 1988. Primary structure and unique expression of the 22-kilodalton light chain of human neutrophil cytochrome b. *Proceedings of the National Academy of Sciences of the United States of America* 85, 3319–3323.
- Quinn, M.T., Evans, T., Loetterle, L.R., Jesaitis, A.J., Bokoch, G.M., 1993. Translocation of Rac correlates with NADPH oxidase activation. Evidence for equimolar translocation of oxidase components. *Journal of Biological Chemistry* 268, 20983–20987.
- Raad, H., Paclet, M.H., Boussetta, T., *et al.*, 2009. Regulation of the phagocyte NADPH oxidase activity: Phosphorylation of gp91phox/NOX2 by protein kinase C enhances its diaphorase activity and binding to Rac2, p67phox, and p47phox. *FASEB Journal* 23, 1011–1022.
- Regier, D.S., Waite, K.A., Wallin, R., McPhail, L.C., 1999. A phosphatidic acid-activated protein kinase and conventional protein kinase C isoforms phosphorylate p22(phox), an NADPH oxidase component. *Journal of Biological Chemistry* 274, 36601–36608.
- Roos, D., De Boer, M., Kuribayashi, F., *et al.*, 1996. Mutations in the X-linked and autosomal recessive forms of chronic granulomatous disease. *Blood* 87, 1663–1681.
- Roos, D., van Bruggen, R., Meischl, C., 2003. Oxidative killing of microbes by neutrophils. *Microbes Infection* 5, 1307–1315.
- Rotrosen, D., Leto, T.L., 1990. Phosphorylation of neutrophil 47-kDa cytosolic oxidase factor. Translocation to membrane is associated with distinct phosphorylation events. *Journal of Biological Chemistry* 265, 19910–19915.
- Rotrosen, D., Yeung, C.L., Leto, T.L., Malech, H.L., Kwong, C.H., 1992. Cytochrome b558: The flavin-binding component of the phagocyte NADPH oxidase. *Science* 256, 1459–1462.
- Royer-Pokora, B., Kunkel, L.M., Monaco, A.P., *et al.*, 1986. Cloning the gene for an inherited human disorder—chronic granulomatous disease—on the basis of its chromosomal location. *Nature* 322, 32–38.
- Segal, A.W., Heyworth, P.G., Cockcroft, S., Barrowman, M., 1985. Stimulated neutrophils from patients with autosomal recessive chronic granulomatous disease fail to phosphorylate a Mr-44 000 protein. *Nature* 316, 547–549.
- Sumimoto, H., Kage, Y., Nuno, H., *et al.*, 1994. Role of Src homology 3 domains in assembly and activation of the phagocyte NADPH oxidase. *Proceedings of the National Academy of Sciences of the United States of America* 91, 5345–5349.
- Tsunawaki, S., Mizunari, H., Nagata, M., Tatsuzawa, O., Kuratsui, T., 1994. A novel cytosolic component, p40phox, of respiratory burst oxidase associates with p67phox and is absent in patients with chronic granulomatous disease who lack p67phox. *Biochemical and Biophysical Research Communications* 199, 1378–1387.
- Vignais, P.V., 2002. The superoxide-generating NADPH oxidase: Structural aspects and activation mechanism. *Cellular and Molecular Life Sciences* 59, 1428–1459.
- Volpp, B.D., Nauseef, W.M., Donelson, J.E., Moser, D.R., Clark, R.A., 1989. Cloning of the cDNA and functional expression of the 47-kilodalton cytosolic component of human neutrophil respiratory burst oxidase. *Proceedings of the National Academy of Sciences of the United States of America* 86, 7195–7199. *Erratum in: Proceedings of the National Academy of Sciences of the United States of America* 86, 9563.
- Wallach, T.M., Segal, A.W., 1996. Stoichiometry of the subunits of flavocytochrome b558 of the NADPH oxidase of phagocytes. *Biochemical Journal* 320, 33–38.
- Wientjes, F.B., Hsuan, J.J., Totty, N.F., Segal, A.W., 1993. p40phox, a third cytosolic component of the activation complex of the NADPH oxidase to contain src homology 3 domains. *Biochemical Journal* 296, 557–561.
- Witko-Sarsat, V., Rieu, P., Descamps-Latscha, B., Lesavre, P., Halbwachs-Mecarelli, L., 2000. Neutrophils: Molecules, functions and pathophysiological aspects. *Laboratory Investigation* 80, 617–653.
- Yaffe, M.B., 2002. The p47phox PX domain: Two heads are better than one! *Structure* 10, 1288–1290.
- Yu, L., Zhen, L., Dinuer, M.C., 1997. Biosynthesis of the phagocyte NADPH oxidase cytochrome b558. Role of heme incorporation and heterodimer formation in maturation and stability of gp91phox and p22phox subunits. *Journal of Biological Chemistry* 272, 27288–27294.
- Yuzawa, S., Miyano, K., Honbou, K., Inagaki, F., Sumimoto, H., 2009. The domain organization of p67 phox, a protein required for activation of the superoxide-producing NADPH oxidase in phagocytes. *Journal of Innate Immunity* 1, 543–555.
- Yuzawa, S., Suzuki, N.N., Fujioka, Y., *et al.*, 2003. Crystallization and preliminary crystallographic analysis of the autoinhibited form of the tandem SH3 domain of p47(phox). *Acta Crystallographica D: Biological Crystallography* 59, 1479–1480.
- Zhan, Y., Virbasius, J.V., Song, X., Pomerleau, D.P., Zhou, W., 2002. The p40phox and p47phox PX domains of NADPH oxidase target cell membranes via direct and indirect recruitment by phosphoinositides. *Journal of Biological Chemistry* 277, 4512–4518.

Cytoplasmic Sensing in Innate Immunity

C Hutcheon, P Paulvannan, and N Subramanian, Institute for Systems Biology, Seattle, WA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The vertebrate immune system is composed of two major subdivisions, the innate or inborn immune system and the adaptive or acquired immune system. The innate immune system is our first line of defense against pathogens, has anatomical features that serve as barriers to infection (such as the skin and epithelia or mucous membranes lining the respiratory, gastrointestinal, and urinary tracts) and includes defenses that are constitutively present and ready to be deployed upon infection. The adaptive immune system on the other hand is antigen specific and confers protection against reexposure to the same pathogen. Cells of the innate immune system such as epithelial cells, macrophages, neutrophils, and dendritic cells (DCs) detect pathogens by employing a network of germline-encoded sensors that are present on the cell surface or in the cytoplasm. These are collectively called pattern recognition receptors (PRRs) to denote their ability to recognize conserved microbial structures or pathogen-associated molecular patterns (PAMPs). While evolutionarily believed to sense ligands of microbial origin, some PRRs can also detect endogenous danger signals released from damaged tissues (danger-associated molecular patterns; DAMPs). The major classes of innate sensors include Toll-like receptors (TLRs), nucleotide oligomerization domain leucine-rich repeat receptors (NLRs), C-type lectin receptors (CLRs), retinoic acid-inducible gene I-like receptors (RLRs), as well as several putative sensors of intracellular DNA such as absent in melanoma 2 (AIM2), stimulator of interferon genes (STING) and DNA-dependent activator of interferon regulatory factors (DAI) (Ishii *et al.*, 2008). Upon activation by PAMPs or DAMPs, PRRs initiate innate immune activities and instruct subsequent adaptive immune responses, enabling both short-term effector function as well as long-lasting immunological memory. The PAMPs recognized and the subsequent immune responses initiated vary depending upon the invading microbe and the PRR triggered; however signaling generally converges on activation of nuclear transcription factors such as nuclear factor- κ B (NF- κ B) and interferon regulatory factors (IRFs) leading to pro-inflammatory cytokine production, or formation of oligomeric signaling complexes called inflammasomes that activate caspase-1 leading to generation of bioactive interleukin (IL)-1 β , IL-18 and pyroptosis. Exquisite coordination of several innate immune pathways determines the quality, magnitude, and duration of the ensuing host response, ultimately leading to containment of invading microorganisms. Apart from its well appreciated role in host defense, innate signaling is also emerging as a critical factor in human inflammatory disease. Indeed, the unrestrained inflammatory response propagated by innate sensors can be detrimental and has been implicated in the development of a variety of immune disorders including atherosclerosis, Type 1 diabetes, inflammatory bowel disease, and systemic lupus erythematosus (SLE) (Li *et al.*, 2009; Holm *et al.*, 2013; Zhong *et al.*, 2013). In this article we review

current information on innate immune signaling with a focus on cytoplasmic sensing. In 'Classes, Structure, and Activation of Innate Sensors,' we introduce the major classes of innate sensors, their signaling pathways and role in mediating protective host defense and/or autoimmunity. In 'Role of Intracellular Organelles and Spatial Relocation,' we discuss the role of subcellular structures and spatial patterns of intracellular movement in generation and/or propagation of the innate immune response. In 'Combinatorial Sensing and PRR Crosstalk,' we discuss the crosstalk between different PRR pathways and integration of signals from multiple ligands displayed by a microbe. A holistic and predictive understanding of innate immunity requires systems-level integration of information from multiple molecular levels, an aspect that is discussed in 'Multiscale Regulation of Cytosolic Sensing Pathways.'

Classes, Structure, and Activation of Innate Sensors

CLRs

CLRs are a large, functionally diverse group of soluble and transmembrane proteins that are primarily involved in detecting a wide range of carbohydrate structures on fungal pathogens, but can also recognize a diverse repertoire of structurally dissimilar microbe-associated or endogenous ligands (Hoving *et al.*, 2014). CLRs are characterized by a C-type lectin domain (CTLN) also referred to as a carbohydrate recognition domain (CRD) in cases where carbohydrates are recognized (Figure 1). The CLR family encompasses more than 1000 identified members with diverse functions including endocytosis, phagocytosis, cell adhesion, complement activation, tissue remodeling, antimicrobial, pro-inflammatory, and anti-inflammatory responses. However, based on their molecular structure CLRs can be classified into three notable groups: two groups of membrane-bound CLRs namely Type I CLRs that contain multiple CRDs, and Type II CLRs that contain a single CRD, and a group of soluble CLRs (Table 1). Prominent examples of CLRs from these groups include dectin-1, dectin-2, DC-specific ICAM3-grabbing nonintegrin (DC-SIGN), mincle, DNGR-1 (CLEC9A), and mannose-binding lectin (MBL). CLRs may be activatory or inhibitory based upon their ability to associate with signaling molecules or the presence of specific motifs in their cytoplasmic tails, which can be one of at least four different types: immunoreceptor tyrosine-based activation motifs (ITAMs), immunoreceptor tyrosine-based inhibitory motifs (ITIMs), a single tyrosine-based motif (hemiITAM), or tyrosine-independent motif. Upon ligand binding of activatory CLRs, the tyrosine residues of an ITAM are phosphorylated by Src family kinases which in turn promote the recruitment of Syk family kinases culminating in the activation of various cellular responses, notably the activation of NF- κ B and production of reactive oxygen species (ROS).

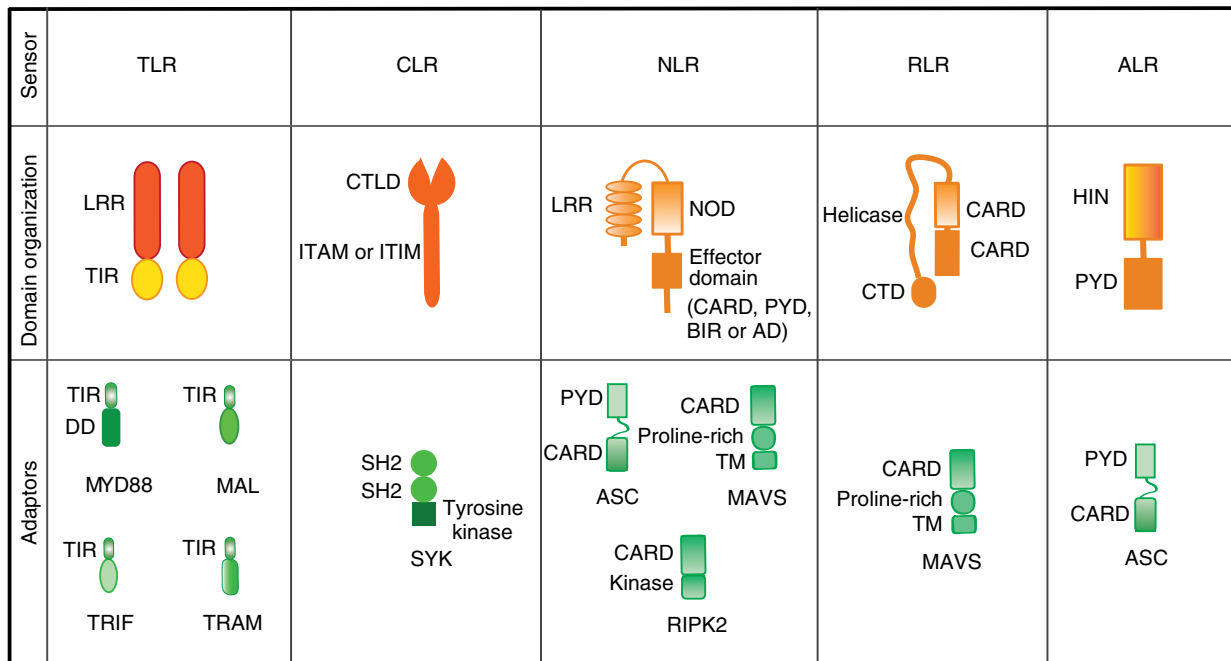


Figure 1 Chart showing the major classes of innate immune sensors, their general domain organization, and adaptors utilized for signaling. These include the TLRs, NLRs, RLRs, and ALRs. In the case of NLRs, the N-terminal effector region can be a caspase activation and recruitment domain (CARD), pyrin domain (PYD), baculoviral inhibitor of apoptosis repeat (BIR), or transactivator domain (AD) depending upon the NLR. ALR, AIM2-like receptor; ASC, Apoptosis-associated speck-like protein containing CARD; CLR, C-type lectin receptor; CTD, C-terminal domain; CTLD, C-type lectin domain (CTLD); DD, death domain; HIN, hematopoietic expression; interferon-inducible nature; and nuclear localization; ITAM, immunoreceptor tyrosine-based activation motif; ITIM, immunoreceptor tyrosine-based inhibitory motif; LRR, Leucine-rich repeats; MAL, MYD88-adaptor-like protein; MAVS, mitochondrial antiviral signaling protein; MYD88, myeloid differentiation primary-response protein 88; NLR, NOD-like receptor; NOD, Nucleotide-binding oligomerization domain; RIPK2, receptor-interacting serine/threonine kinase 2; RLR, RIG-I-like receptor; Syk, Spleen tyrosine kinase; TIR, Toll/interleukin-1 receptor; TLR, Toll-like receptor; TRAM, TRIF-related adaptor molecule; TRIF, TIR domain-containing adaptor protein inducing interferon- β .

Table 1 Summary of select CLRs, their ligands and responses

Type	Sensor	Ligand(s)	Response
Type I	DEC 205	CpG oligonucleotides, ligands expressed by dying apoptotic or necrotic cells	Antigen presentation, cytokine production
	MMR	Mannose, <i>N</i> -acetylglucosamines, fucose residues <i>Pneumocystis carinii</i> , <i>Leishmania donovani</i>	Phagocytosis, cytokine production, IL-8 in cooperation with TLR2
Type II	Dectin-1	β -glucans	Phagocytosis, ROS production, activation of NF- κ B and AP-1 through SYK adaptor, ROS production and potassium efflux resulting in NLRP3 inflammasome activation
	Dectin-2	High mannose-type carbohydrates	
	Mincle	Fungal α -mannose, mycobacterial glycolipid trehalose-6'-6'-dimycolate	Activation of NF- κ B and AP-1 through the Fc receptor common γ -chain (FcR γ) and SYK
	DC-SIGN	Human immunodeficiency virus (HIV)-1, Hepatitis C virus, dengue virus, cytomegalovirus, ebola virus, <i>Leishmania</i> , and <i>Candida</i> species	Modulation of TLR signaling by acetylation of active NF- κ B p65 through the serine threonine kinase Raf-1 resulting in anti-inflammatory responses
	CLEC9A	DAMPs such as exposed actin filaments of damaged or dead cells	Antigen presentation
Soluble	MBL	Repetitive mannose and/or <i>N</i> -acetylglucosamine residues, including those on HIV	Phagocytosis, anti-inflammatory cytokines

Ligand engagement of inhibitory CLRs commonly results in ITIM tyrosine phosphorylation by Src kinases, the recruitment and activation of protein tyrosine phosphatases such as SHP-1 and SHP-2 and the dephosphorylation of substrates leading to

the inhibition of cellular activation (Osorio *et al.*, 2011; Redelinghuys and Brown, 2011). CLRs have been well-reviewed elsewhere (Hoving *et al.*, 2014; Osorio *et al.*, 2011) and because they primarily function in recognition of extracellular ligands,

will not be covered in detail in this review. A summary of select CLRs, including their known ligands and signaling responses, is shown in [Table 1](#).

TLRs

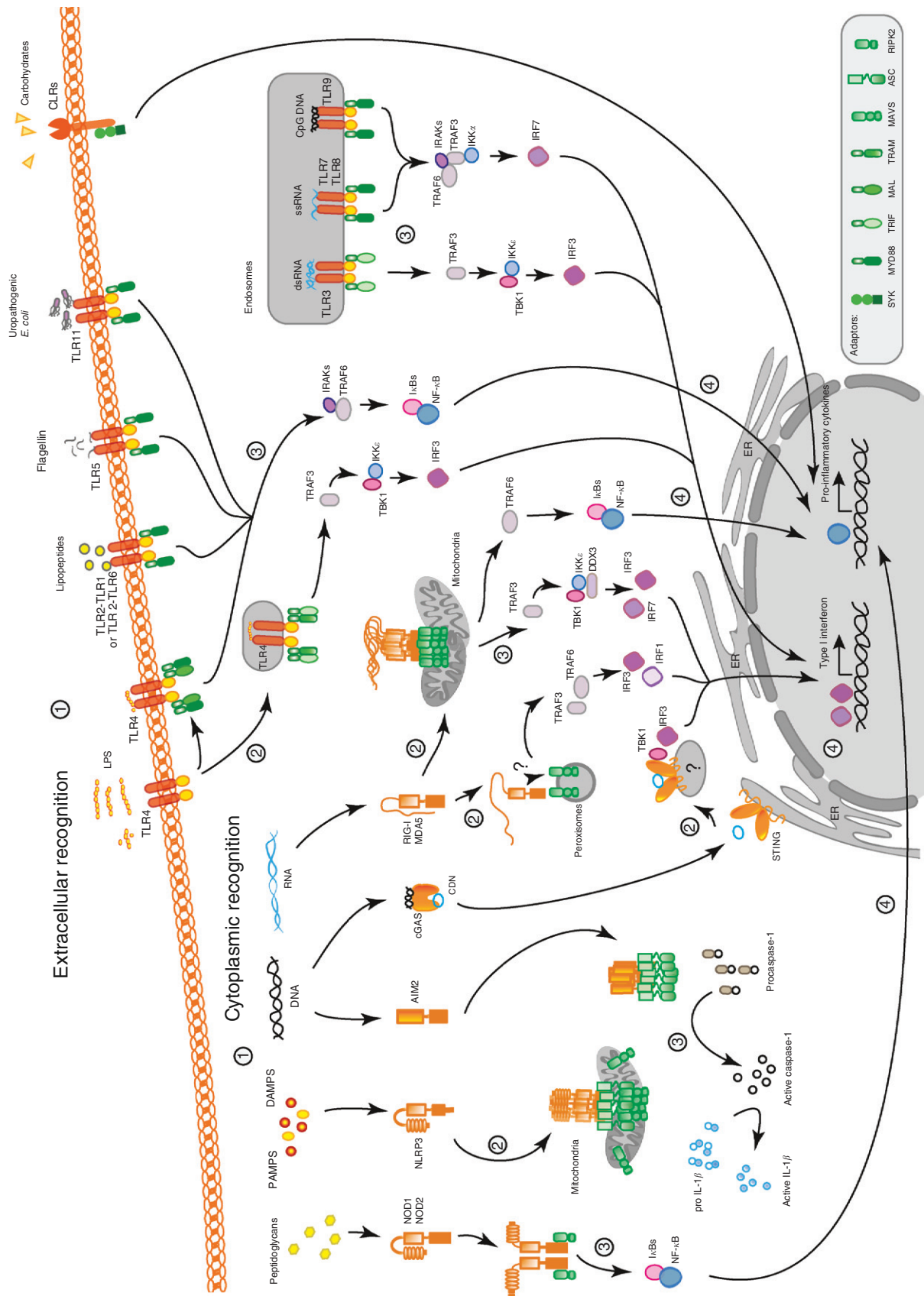
TLRs were the first class of innate sensors to be identified and are the most well characterized PRRs. To date, 13 members of the TLR family have been identified in mammals. These recognize a wide range of PAMPs including LPS (recognized by TLR4); bacterial peptidoglycan, lipoproteins, and fungal zymosan (recognized by TLR2-TLR1 or TLR2-TLR6 heterodimers); bacterial flagellin (recognized by TLR5); unmethylated bacterial or viral CpG DNA (recognized by TLR9); viral single-stranded RNA (recognized by TLR7 and TLR8) and viral double-stranded RNA (recognized by TLR3); and bacterial 23S ribosomal RNA (recognized by TLR13) ([Table 2](#); [Akira et al., 2006](#); [Kawai and Akira, 2011](#)). TLRs are type I transmembrane proteins characterized by three distinct domains: an extracellular ligand sensing domain-containing varying numbers of leucine-rich repeat (LRR) motifs that mediate the recognition of PAMPs, a transmembrane region, and a cytoplasmic domain homologous to that of the interleukin 1 receptor (IL-1R), termed the Toll/IL-1R homology (TIR) domain that activates downstream signaling pathways ([Figure 1](#)).

TLRs are expressed on myeloid and lymphoid cells of hematopoietic origin, including macrophages, DCs, B cells, specific types of T cells, and even on non-hematopoietic cells such as fibroblasts and epithelial cells. Moreover, TLRs may be expressed extra- or intracellularly depending upon their PAMP recognition properties. While certain TLRs (TLRs 1, 2, 4, 5, and 6) are expressed and bind to their respective ligands on the cell surface, others (TLRs 3, 7, 8, 9, and 13) are found almost exclusively in intracellular compartments such as endosomes where they sense microbial and host-derived nucleic acids ([O'Neill et al., 2013](#); [Blasius and Beutler, 2010](#)). TLRs 11 and 12 also localize within endosomal compartments where they are involved in the detection of *Toxoplasma gondii* profilin protein ([Yarovinsky, 2014](#); [Andrade et al., 2013](#); [Pifer et al., 2011](#)). All the TLRs are expressed in mice; however, humans lack functional TLR11, TLR12, and TLR13 genes.

TLR signaling is initiated by ligand-induced homo-dimerization or heterodimerization of receptors. Following dimerization, the TIR domains of TLRs associate with TIR domain-containing adaptor proteins, either myeloid differentiation primary-response protein 88 (MYD88), and MYD88-adaptor-like protein (MAL) also called TIR domain-containing adaptor protein (TIRAP), or TIR domain-containing adaptor protein inducing IFN β (TRIF) and TRIF-related adaptor molecule (TRAM). Engagement of these adaptor molecules

Table 2 Summary of human and mouse TLR ligands, adaptors, responses, and associated human diseases

Sensor	Ligand(s)	Adaptor	Response	Associated diseases
TLR 1/TLR2	Triacyl lipopeptides (Pam3CSK4)	MyD88 TIRAP/MAL	Inflammatory cytokines	
TLR 2/TLR6	Diacyl lipopeptides Lipoteichoic acid Zymosan Glycolipids GPI anchor Phospholipomannan	MyD88 TIRAP/MAL	Inflammatory cytokines Type 1 IFN	Atherosclerosis
TLR 3	Synthetic dsRNA Viral dsRNA	TRIF	Type I IFN	Virus-triggered autoimmune disease
TLR 4	LPS Mannan Glucuronoxylomannan Glycoinositolphospholipids GPI anchor	MyD88 TIRAP/MAL TRAM TRIF	Inflammatory cytokines Type 1 IFN	Atherosclerosis
TLR 5	Flagellin	MyD88 TRIF	Inflammatory cytokines	
TLR 7	ssRNA imidazoquinoline derivatives (R848)	MyD88	Inflammatory cytokines Type 1 IFN	Psoriasis, multiple sclerosis, systemic lupus erythematosus
TLR 8	ssRNA imidazoquinoline derivatives (R848)	MyD88		Multiple sclerosis
TLR 9	Unmethylated CpG DNA Synthetic CpG oligodinucleotides Hemozoin	MyD88	Inflammatory cytokines Type 1 IFN	Psoriasis, systemic lupus erythematosus
TLR 10 (humans)	Unknown	MyD88	Activation of NF- κ B and ENA-789 promoters	
TLR11	<i>Toxoplasma gondii</i> profilin protein, uropathogenic <i>Escherichia coli</i> , <i>Salmonella</i> flagellin	MyD88	NF- κ B and IRF8 dependent IL-12 production	
TLR12	<i>T. gondii</i> profilin protein	MyD88	NF- κ B and IRF8 dependent IL-12 production	
TLR13	Bacterial 23S ribosomal RNA	MyD88	Activation of NF- κ B	



organizes downstream signaling that involves interactions between IL-1R-associated kinases (IRAKs) and the TNF receptor-associated factors (TRAFs), leading to the activation of the mitogen-activated protein kinases (MAPKs), JUN N-terminal kinase (JNK), p38, and to the activation of transcription factors (Figure 2). Two important families of transcription factors that are activated downstream of TLR signaling are NF- κ B and the IRFs, but other transcription factors, such as cyclic AMP-responsive element-binding protein (CREB) and activator protein 1 (AP1) may also be activated. A major consequence of TLR signaling is the induction of pro-inflammatory cytokines such as tumor necrosis factor α (TNF- α), IL-1, IL-6, and IL-12, and in the case of the endosomal TLRs, the induction of type 1 interferon (IFN) (Figure 2; Blasius and Beutler, 2010; O'Neill *et al.*, 2013).

Activation of TLRs on antigen-presenting cells such as macrophages, DCs and B cells, strongly influences the development of adaptive immune responses by controlling several essential processes such as the upregulation of co-stimulatory or accessory molecules like CD80 and CD86, enhancement of antigen presentation, B-cell proliferation and maturation, activation of T cells, and suppression of regulatory T cell activity (Iwasaki and Medzhitov, 2010), thereby mediating effective host defense to a wide range of pathogens. However, an unchecked TLR-induced inflammatory response can be deleterious to the host resulting in a variety of immune disorders. TLRs are implicated in the pathogenesis of autoimmune diseases, such as SLE, rheumatoid arthritis, atherosclerosis, hepatitis, diabetes, kidney disease, and certain other disorders (Li *et al.*, 2009). Recent studies have shown that functional TLRs are expressed not only on immune cells, but also on cancer cells, implicating a role of TLRs in tumor biology. Analogous to their favorable and detrimental roles in inflammation, increasing evidence suggests that TLRs act as a double-edged sword in cancer cells too, pointing to a dual function of TLRs as anti- and pro-tumor modulators (Basith *et al.*, 2012). While unrestrained TLR signaling provides a microenvironment conducive for tumor growth and evasion of the immune response, TLRs can also trigger an antitumor response that inhibits tumor progression and promotes tumor clearance. Thus, apart from their protective role against various pathogens, the involvement of TLRs in autoimmune diseases and tumorigenesis indicates the need for further studies to delve into the potential of targeting TLR signaling for therapeutic benefit under these conditions.

NLRs

While TLRs detect microbial signatures in the extracellular environment or intracellularly in the lumen of endosomes, NLRs are the largest known family of innate sensors that survey the cytosolic environment. In humans, 23 NLRs have been identified (Harton *et al.*, 2002; Inohara and Nunez, 2003; Table 3). Their structural hallmark is a modular organization of three domains with distinct function: a variable C-terminal LRR domain for ligand binding, a conserved nucleotide binding or oligomerization domain (NBD or NOD), and a variable N-terminal effector or protein interaction domain that mediates downstream signaling cascades (Figure 1). The N-terminal domain mediates homotypic protein-protein interactions and can be one of four different types depending upon the receptor subclass: NLRA or Class II transactivator (CIITA) contains an acidic transactivation domain, NLRBs or neuronal apoptosis inhibitor proteins (NAIPs) contain a baculoviral inhibitor of apoptosis repeat (BIR) domain, NLRs have a caspase activation and recruitment domain (CARD), and NLRPs possess a pyrin domain (PYD). The NLRX1 effector domain bears no homology to the N-terminal region of any of the above four subclasses and is instead categorized as a CARD-related X effector domain. PAMPs and DAMPs are believed to be sensed by the C-terminal LRR domain. Upon ligand binding, the auto-inhibitory LRR undergoes a conformational change, allowing for protein oligomerization through the central NOD domain. This in turn exposes the N-terminal effector domain, allowing for recruitment of downstream signaling adaptors and effector proteins resulting in formation of an oligomeric complex. The activating stimuli, mechanisms of activation, and response to ligand sensing vary within the NLR family (Table 3) and have been well defined only for a few family members; the physiologic functions and relevant signaling pathways of most members of the NLR family are either unknown or poorly defined. Among the NLRs that have been studied intensively, NLRP3, NLRP1, and NLRC4 form a macromolecular signaling complex called the inflammasome that acts as a scaffold for activation of caspase-1, a key step that leads to processing and secretion of the potent pro-inflammatory cytokines IL-1 β and IL-18, while NOD1 and NOD2 form signaling platforms that activate NF- κ B (Martinon *et al.*, 2009).

Figure 2 Overview of cytoplasmic innate sensing and signaling mechanisms. Major signaling cascades triggered by extracellular and intracellular PRRs are shown. Pathogen- or danger-associated stimuli are detected by innate sensors (1), often through combinatorial sensing in which multiple receptors sense ligands from the same microbe. Ligand binding results in conformational changes, followed by self-association and/or autophosphorylation of the receptors, allowing them to bind adaptor proteins. Additional trafficking proteins (not depicted) may guide the active oligomeric complexes to secondary compartments within the cell (2). For instance, TLR4 may be trafficked to early endosomal compartments where it binds the adaptors TRIF and TRAM. Activated RIG-I and NLRP3 associate with MAVS on the mitochondrial surface; RIG-I may additionally associate with MAVS on the peroxisomes. STING moves from the ER to a secondary structure that has, as yet, not been fully identified and is therefore depicted with a '?'. Adaptor proteins may allow for assembly of large, oligomeric, multiprotein signaling complexes capable of activating downstream mediators and transcription factors (TFs) resulting in a functional response (3). For example, activated NLRP3 and AIM2 form structures called inflammasomes that serve as platforms for activation of caspase-1, which in turn cleaves pro-IL-1 β and pro-IL-18 to produce active cytokines. Activated RIG-I induces polymerization of MAVS into large self-propagating prion-like structures that trigger phosphorylation cascades ultimately leading to activation of the TFs IRF3 and IRF7. The nature of RIG-I and MDA5 interactions with adaptor proteins on the peroxisome is not yet fully understood and is therefore depicted with a '?'. IRF3 and IRF7 translocate to the nucleus and initiate transcription of type 1 IFN (4). The TF NF- κ B is also activated by RIG-I-MAVS interaction, ultimately leading to transcription of a number of pro-inflammatory cytokines. Combinatorial sensing of diverse ligands can result in simultaneous or sequential activation of multiple pathways, which may be important for generating a tailored, specific response to a potential pathogen.

Table 3 Summary of human NLRs, their activators, responses, and associated diseases

Effector domain	Sensor	Activator(s)	Response	Associated diseases
CARD	CIITA	Unknown	MHC class II transcriptional activation	Bone density defects, Hodgkin's lymphoma, B-cell lymphoma, celiac disease, arthritis, multiple sclerosis, primary adrenal insufficiency, systemic lupus erythematosus, diabetes
	NOD1	Meso-diaminopimelic acid (DAP), γ -D-glutamyl-meso-diaminopimelic acid (iE-DAP), meso-lanthionine, GM-tripeptide, FK156, FK565	NF- κ B activation through IKK complex; MAP kinase pathway activation leading to pro-inflammatory responses	Inflammatory bowel disease, atopic eczema, asthma
	NOD2	Muramyl dipeptide (MDP), M-TR _{LVS}	NF- κ B activation through IKK complex; MAP kinase pathway activation leading to pro-inflammatory responses	GVHD, Blau disease, Crohn's disease, asthma, atopic dermatitis, arthritis, sarcoidosis, prostate and endometrial cancer, gastric lymphoma, leprosy
	NLRC4	Flagellin, Type III secretion system components	Inflammasome formation resulting in caspase-1 activation and cleavage of IL-1 β and IL-18 to their active forms	Susceptibility to bacterial infections
	NLRC3	Unknown	Negative regulator of T cell activation, TLR, and STING signaling	
	NLRC5	Unknown; NLRP3 agonists, including bacterial PAMPs and crystals have been reported as elicitors in cell culture systems	Transcriptional regulator of MHC class I. Other functions are not fully understood because response appears to vary with species, context, and cell type	Susceptibility to viral infections
	NLRX1	Unknown	Activates NF- κ B and ROS through JNK pathway, negatively regulates RIG-I-mediated IFN response and LPS-elicited TRAF6-IKK signaling pathways	Susceptibility to chronic hepatitis B infection, gastric cancer
PYD	NLRP1	Muramyl dipeptide, <i>Bacillus anthracis</i> lethal toxin	Inflammasome assembly resulting in caspase-1 activation and production of active IL-1 β and IL-18	Skin autoimmune disorders, vitiligo, Addison's disease, diabetes, celiac disease, autoimmune thyroid disorders, systemic lupus erythematosus, systemic sclerosis, giant cell arteritis, congenital toxoplasmosis, Alzheimer's disease, corneal intraepithelial dyskeratosis, arthritis
	NLRP2	Unknown		Beckwith–Wiedemann syndrome, GVHD
	NLRP3	Crystals (monosodium urate, calcium pyrophosphate dihydrate, cholesterol, asbestos, silica, hydroxyapatite), amyloid β , mitochondrial DNA, cardiolipin, ceramides, ATP, ROS, RNA viruses, bacterial toxins	Inflammasome assembly resulting in caspase-1 activation and production of active IL-1 β and IL-18	Gout, Cryopyrin-associated periodic fever syndromes (CAPS), diabetes, celiac disease, psoriasis, inflammatory bowel diseases, Alzheimer's disease, atherosclerosis
	NLRP4	Unknown	Unknown, possible suppressor of NF- κ B	
	NLRP5	Unknown	Unknown	Familial biparental hydatidiform moles

(Continued)

Table 3 Continued

Effector domain	Sensor	Activator(s)	Response	Associated diseases
	NLRP6	Unknown	Unknown, possible negative regulator of NF- κ B and IL-1 β , an NLRP6 inflammasome has been proposed in mouse colonic epithelial cells	Metabolic syndrome-associated abnormalities, colitis, and colon cancer
	NLRP7	Bacterial acylated lipoproteins (acLP)	Unknown, possible negative regulator of IL-1 β , inflammasome assembly in response to bacterial acLP	Familial biparental hydatidiform moles, abnormal pregnancies and embryonic development, testicular and endometrial cancer
	NLRP8	Unknown	Unknown	Alzheimer's disease, colorectal cancer
	NLRP9	Unknown	Unknown	Abnormal embryonic development
	NLRP10	Unknown	Negatively regulates caspase-1 activation, regulation of DC emigration from inflamed tissues in mice	Atopic dermatitis
	NLRP11	Unknown	Unknown	
	NLRP12	Unknown	Negatively regulates NF- κ B activation	Hereditary fever syndromes, dermatitis
	NLRP13	Unknown	Unknown	Abnormal embryonic development
	NLRP14	Unknown	Unknown	Familial biparental hydatidiform moles, spermatogenic failure
BIR	NAIP	Type III secretion system needle proteins	Formation of a large oligomeric complex with NLRC4 resulting in caspase-1 activation and production of IL-1 β	Spinal muscle atrophy

The current paradigm for NLR activation is self-oligomerization followed by recruitment of adaptor proteins that mediate activation of downstream effectors. In the case of NOD1 and NOD2, two well-studied NLRs that sense different components of peptidoglycan from Gram-negative and Gram-positive bacteria (Elinav *et al.*, 2011), ligand binding allows for the recruitment of the receptor-interacting serine-threonine kinase 2 (RIPK2) adaptor protein through the exposed CARD subsequently leading to polyubiquitination of the inhibitor of NF- κ B kinase subunit gamma (IKK γ). This in turn culminates in activation of the I κ B kinase (IKK) complex, I κ B α phosphorylation followed by its degradation and activation of NF- κ B (Windheim *et al.*, 2007; Abbott *et al.*, 2007; Figure 2). NOD1 or NOD2 activation also results in the activation of MAP kinases such as ERK-1, ERK-2, JNK, and p38, which coordinate with NF- κ B to upregulate the expression of pro-inflammatory molecules (Girardin *et al.*, 2001; Pauleau and Murray, 2003; Park *et al.*, 2007). In addition, activation of NOD2 by single-stranded viral RNA from respiratory syncytial virus (RSV), vesicular stomatitis virus (VSV), and influenza virus has been shown to trigger a noncanonical signaling pathway that requires the mitochondrial antiviral signaling protein (MAVS) and induces IRF3 activity, leading to the production of type 1 IFN (Sabbah *et al.*, 2009). In the case of NLRP3, conformational changes in the molecule following ligand sensing expose the PYD domain resulting in interaction with a similar domain on the adaptor protein apoptosis-associated speck-like protein containing a CARD (ASC), which through homotypic CARD domain interactions associates with

procaspase-1 resulting in its auto-catalytic cleavage to enzymatically active caspase-1 (Agostini *et al.*, 2004). The latter then cleaves the pro forms of IL-1 β and IL-18 to produce active cytokine. This process is amplified by a second adaptor protein MAVS in response to noncrystalline NLRP3 activators like poly I:C (Subramanian *et al.*, 2013). In contrast to NLRP3, NLRP1, and NLRC4 contain a CARD domain that allows them to associate directly with procaspase-1. NLRC4 and NLRP1 may also recruit ASC, resulting in augmented inflammasome activity; however ASC is not absolutely required for inflammasome assembly (Lamkanfi and Dixit, 2014; Schroder and Tschopp, 2010; Hsu *et al.*, 2008).

Among the inflammasome-forming NLRs, NLRP3 is perhaps the best studied. NLRP3 inflammasome formation appears to require at least two distinct signals. The first signal is an NF- κ B activating stimulus, which leads to transcriptional upregulation of pro-IL-1 β and NLRP3. This is often referred to as a 'priming' signal and may consist of a TLR activator such as LPS (Bauernfeind *et al.*, 2009). A second signal then activates NLRP3 resulting in inflammasome assembly (Lamkanfi and Dixit, 2014). The nature of this activating stimulus can be diverse. A variety of signals associated with infection or metabolic dysfunction can activate NLRP3, ranging from crystalline substances like monosodium urate, cholesterol, calcium pyrophosphate dehydrate, silica or alum crystals, to noncrystalline activators such as RNA or bacterial toxins, and endogenous danger signals like ATP, mitochondrial DNA, cardiolipin, ceramides, low cytosolic potassium, and increased intracellular calcium (Dostert *et al.*, 2008; Cassel *et al.*, 2008;

Cruz *et al.*, 2007; Lee *et al.*, 2012). How these chemically and structurally distinct stimuli elicit a response from a single sensor is unclear (Gross *et al.*, 2011). One proposed common feature of NLRP3 activators is induction of ROS that may lead to the generation of a potential ligand of NLRP3 or may modify NLRP3 or associated proteins directly (Cruz *et al.*, 2007; Zhou *et al.*, 2010). However, a study also argues that ROS are required for transcriptional induction of NLRP3 expression, but not for NLRP3 activation (Bauernfeind *et al.*, 2011); these divergent conclusions may be due to potentially off-target effects of pharmacological scavengers or inducers of ROS, and their usage at differing doses or duration that may cause varying outcomes depending on the cell type. Another reported pathway involves lysosome disruption by crystals such as those formed by cholesterol, uric acid, or alum, causing release of active cathepsin B that presumably acts as an upstream activator of NLRP3 (Hornung *et al.*, 2008). Interestingly, deubiquitination of NLRP3 can prime NLRP3 non-transcriptionally and is essential for ASC aggregation and inflammasome activity, however the precise purpose of this deubiquitination is not known (Juliana *et al.*, 2012; Lopez-Castejon *et al.*, 2013). The high level of regulation of the NLRP3 inflammasome may be due in part to the prion-like nature of NLRP3 inflammasome assembly. Nucleation of ASC oligomerization through the PYD has been shown to promote energetically-favorable polymerization into large, stable, amyloid-like structures which may serve to quickly amplify signaling (Cai *et al.*, 2014), yet which may be difficult for the cell to reverse once initiated.

The NLRC4 inflammasome has mainly been implicated in the host defense to bacterial pathogens such as *Salmonella typhimurium*, *Legionella pneumophila*, *Burkholderia pseudomallei*, *Pseudomonas aeruginosa*, and *Klebsiella pneumoniae* (Corridoni *et al.*, 2014). Posttranslational phosphorylation of Ser533 in NLRC4 by PKC δ has been proposed to be crucial for NLRC4 inflammasome function (Qu *et al.*, 2012). NLRC4 is expressed mostly on myeloid cells and activates caspase-1 dependent production of bioactive IL-1 β upon cytosolic detection of bacterial flagellin, and components of the type III secretion system (T3SS) (Miao *et al.*, 2006, 2010b; Franchi *et al.*, 2006). These bacterial ligands trigger oligomerization of NLRC4 with NAIPs, which have been identified as critical components of the NLRC4 inflammasome. In mice, NAIP5 and NAIP6 in complex with NLRC4 confer specificity for recognition of bacterial flagellin, whereas a NAIP2–NLRC4 complex confers specificity for recognition of the *S. typhimurium* rod protein PrgJ. In humans, NLRC4 and the sole human NAIP serve as sensors for the bacterial T3SS needle protein (Kofoed and Vance, 2011; Yang *et al.*, 2013; Zhao *et al.*, 2011). Activation of caspase-1 through a NLRC4 dependent pathway has also been associated with subsequent cell death, termed pyroptosis, which can take place independent of ASC (Miao *et al.*, 2010a; Broz *et al.*, 2010).

NLRP1 was the first inflammasome-forming NLR to be described (Martinon *et al.*, 2002). NLRP1 is expressed in diverse cell types including myeloid cells, T and B cells. In humans, NLRP1 is a single copy gene which encodes a N-terminal PYD, a NBD/NOD, a LRR domain, a function-to-find domain (FIIND), and a C-terminal CARD. Mice encode three polymorphic forms of NLRP1 proteins (*Nlrp1a*, *Nlrp1b*,

and *Nlrp1c*), which lack functional PYD and FIIND domains. Genetic studies have shown that the *NLRP1b* gene is the primary mediator of mouse macrophage susceptibility to *Bacillus anthracis* lethal toxin (LT) in sensitive strains such as BALB/c. Upon exposure to LT, NLRP1 activates caspase-1 resulting in production of mature IL-1 β and pyroptosis. Some proposed mechanisms of LT-induced cell death include cleavage of mitogen-activated protein kinase kinases (MEKs), impairment of mitochondrial function, ATP leakage, lysosomal membrane permeabilization, and cathepsin B release, proteasome-mediated protein degradation, potassium efflux, caspase-1-mediated macrophage necrosis, and inflammasome activation with release of IL-1 β and IL-18; however, the mechanism of activation of the NLRP1 inflammasome by LT remains unclear (Alileche *et al.*, 2006; Averette *et al.*, 2009; Muehlbauer *et al.*, 2007; Wickliffe *et al.*, 2008). In addition, both human and mouse NLRP1 can be activated by muramyl dipeptide (MDP). MDP was suggested to induce conformational changes in NLRP1, which enable its oligomerization, thus creating a platform for caspase-1 activation (Faustin *et al.*, 2007). A NLRP1–NOD2 interaction has also been shown to be required for optimal responses to LT and MDP *in vivo* (Hsu *et al.*, 2008).

Recent studies have uncovered host regulatory mechanisms and pathogen immune-evasion strategies targeting inflammasome function, further emphasizing the importance of inflammasomes in the antimicrobial response. For instance, mouse effector and memory CD4⁺ T cells have been shown to suppress NLRP1 and NLRP3 inflammasome-mediated caspase-1 activation and subsequent IL-1 β release by select TNF family ligands in a cognate manner (Guarda *et al.*, 2009), and type 1 IFNs repress the activity of NLRP1 and NLRP3 inflammasomes through a STAT-1-dependent mechanism (Guarda *et al.*, 2011). Human NLRP1 is regulated by interactions with the anti-apoptotic proteins Bcl-2 and Bcl-X_L, which bind and suppress NLRP1, reducing caspase-1 activation and IL-1 β production (Bruey *et al.*, 2007). NLRP1 is also targeted by viruses to evade innate immunity. For instance, Vaccinia virus F1L protein, a viral homolog of the cellular Bcl-2 protein, has been demonstrated to bind and inhibit NLRP1 through a small hexapeptide in F1L thereby promoting viral virulence (Gerlic *et al.*, 2013). In another example, Kaposi's sarcoma herpes virus (KSHV) Orf63 was found to be a viral homolog of human NLRP1 that interacts with NLRP1 and blocks NLRP1-dependent innate immune responses, thereby contributing to reactivation and generation of progeny virus. KSHV Orf63 was also shown to inhibit the NLRP3 inflammasome (Gregory *et al.*, 2011).

Many newly emerging aspects of the biology of NLR proteins point to functions that transcend pathogen detection and inflammation to include autophagy (e.g., autophagic responses to bacteria are regulated by NOD2, NLRC4), tissue homeostasis (e.g., NLRP3 functions as a key regulator of intestinal homeostasis), transcriptional regulation (e.g., CIITA and NLRC5 function as transcriptional regulators of major histocompatibility complex I and II respectively), as well as nonimmune functions such as embryonic development (NLRP5, NLRP2), and reproduction (NLRP7, NLRP14) (Kuffer and Sansonetti, 2011). NLRs can act as negative or positive regulators of a variety of signaling pathways. For instance, NLRC3 acts as a negative regulator of T cell activation, NLRP2,

NLRP4, NLRP6, NLRP12, and NLRC5 inhibit NF- κ B activation, and NLRC5 and NLRX1 negatively regulate type 1 IFN signaling pathways. On the other hand, activating roles have been reported for NLRP12 and NLRP2 in activation of caspase-1, NLRX1 in ROS production, and NLRC5 in IFN-dependent antiviral responses (Lich and Ting, 2007). Thus NLRs may have diverse roles depending upon the cell type and the infectious context (Table 3).

The critical role of NLRs in physiology is further underscored by their clear implication in several human autoinflammatory and autoimmune diseases. Mutations in NLRP3 and NLRP12 lead to hereditary periodic fever syndromes, while mutations in NOD2 are linked to inflammatory bowel disease, Crohn's disease or Blau syndrome (Zhong *et al.*, 2013). The NLRP3 inflammasome has also been implicated in diseases associated with metabolic dysfunction such as obesity, gout, type 2 diabetes and insulin resistance, atherosclerosis, and Alzheimer's disease due to increased activation of caspase-1 and IL-1 β downstream of NLRP3 triggered by diverse danger signals depending upon the disease context (Vandanmagsar *et al.*, 2011; Masters *et al.*, 2010; Duewell *et al.*, 2010; Heneka *et al.*, 2013). Polymorphisms in NLRP1 are associated with a variety of autoimmune disorders including vitiligo, celiac disease, type 1 diabetes, autoimmune thyroid disorders, systemic lupus erythematosus, rheumatoid arthritis, and

Alzheimer's disease (Zhong *et al.*, 2013). Genome-wide association studies (GWAS) have identified linkage of polymorphisms and/or mutations in NLR genes with a host of other human diseases; these are summarized in Table 3. Given the essential role of NLRs in human disease, future studies investigating the activating ligands, physiological roles and signaling pathways triggered by diverse NLRs will be key to providing fundamental insights into not only how NLRs normally function but also for addressing the molecular basis of aberrant NLR function in patients.

RLRs

The RLRs serve as cytosolic sensors of RNA. RLR family members include retinoic acid-inducible gene 1 (RIG-I), melanoma differentiation factor 5 (MDA5), and laboratory of genetics and physiology 2 (LGP2) which differentially regulate viral RNA detection (Table 4). They are characterized by three distinct domains: a C-terminal domain (CTD), a central DExD/H box helicase domain for RNA binding, and two N-terminal CARD domains, that mediate downstream signaling; these N-terminal CARDs are present in RIG-I and MDA5 but not in LGP2 (Figure 1). RIG-I preferentially binds to short (<300 bp) double-stranded (ds) RNAs that have blunt ends and a 5' triphosphate (5'-ppp) moiety while MDA5

Table 4 Summary of select human nucleic acid sensors, their ligands, responses and associated diseases

Sensor class	Nucleic acid sensor	Ligand(s)	Response	Associated diseases
RLR	RIG-I	5'ppp-dsRNA Viruses: Sendai, Influenza, Newcastle disease, Japanese encephalitis, measles, Vesicular stomatitis, Dengue, Hepatitis C, Rabies	Type 1 IFN	Multiple sclerosis, Type 1 diabetes, psoriasis, rheumatoid arthritis
RLR	MDA5	dsRNA, Poly(I:C) Viruses: Picornavirus, Sendai, Encephalomyocarditis, Rotavirus, Dengue, Rabies	Type 1 IFN	Type 1 diabetes
Other	DAI	B-form dsDNA Pathogens: Human cytomegalovirus, HSV-1, <i>Streptococcus pneumoniae</i>	Type 1 IFN through IRF and TBK1 pathways.	
Other	DDX41	B-form dsDNA Viruses: Adenovirus, HSV-1, Vaccinia	Type 1 IFN through STING pathway	
Other	RNA pol III	B-form dsDNA Pathogens: Epstein-Barr virus, Adenovirus, <i>Legionella pneumophila</i>	Type 1 IFN production through RIG-I pathway	
ALR	AIM2	B-form dsDNA Pathogens: Vaccinia virus, HSV-1, <i>Francisella tularensis</i> , <i>Listeria monocytogenes</i>	Inflammasome formation resulting in IL-1 β and IL-18 and cell death by pyroptosis; NF- κ B activation	Systemic lupus erythematosus
ALR	IFI16	dsDNA, sequence-independent, 70 >> 50 bp Viruses: HSV-1	Type 1 IFN through STING pathway, inflammasome formation in nucleus of Kaposi-sarcoma associated herpes virus (KSHV) infected endothelial cells	Systemic lupus erythematosus
ALR	PYHIN1		STING-dependent IFN production	Asthma in African Americans

preferentially recognizes long dsRNA strands (>1000 bp) with no end specificity (Kato *et al.*, 2008; Hornung *et al.*, 2006; Pichlmair *et al.*, 2006). Both may recognize chemically synthesized dsRNA such as poly I:C or RNA analogs. While 5'ppp-dsRNA exhibits maximal binding affinity, RIG-I is also capable of binding single-stranded (ss) RNA molecules, although with lower affinity (Wang *et al.*, 2010). RIG-I and MDA5 have been reported to have both differential and redundant roles in sensing RNA viruses. RIG-I recognizes mainly negative-sense ssRNA viruses or positive-sense ssRNA/dsDNA viruses including those in the paramyxoviridae, orthomyxoviridae, rhabdoviridae, bunyaviridae families, and some flaviviruses such as Japanese encephalitis and hepatitis C virus; in contrast, MDA5 was shown to primarily detect the dsRNA intermediate form generated during replication of positive ssRNA viruses such as those in the picornaviridae family (Kato *et al.*, 2005, 2006; Loo *et al.*, 2008). However, both are required for an optimal response to West Nile virus, dengue virus, reoviruses, and some paramyxoviruses such as Sendai and measles virus. While RIG-I and MDA5 may differentially detect RNA viruses, they are thought to trigger similar downstream signaling cascades (Yoneyama *et al.*, 2005). In both cases, the ligand-bound RLR engages the mitochondrial or peroxisomal adaptor protein MAVS to activate the cytosolic protein kinases IKK and TANK-binding kinase 1 (TBK1), which in turn activate the transcription factors NF- κ B and IRF3, resulting in production of type 1 IFN and other inflammatory antimicrobial cytokines (Figure 2).

Recent structural studies of RIG-I bound to dsRNA have contributed toward a better understanding of ligand binding and activation of RIG-I. In its inactive form, RIG-I is in an auto-repressed, open conformation whereby the RIG-I CARD motifs sequester the helicase domain. Upon binding of the helicase and CTD with dsRNA and ATP, the protein rearranges into a closed conformation that exposes the CARDs which then interact with the N-terminal CARD of the adaptor protein MAVS to induce NF- κ B and IRF3 activation and establish an antiviral state (Kowalinski *et al.*, 2011). The CARDs of RIG-I have been shown to nucleate the aggregation of MAVS into self-propagating prion-like structures that activate IRF3 (Hou *et al.*, 2011). It is hypothesized that such MAVS polymers may serve to quickly amplify signaling. The prion-like nature of MAVS polymerization bears resemblance to ASC aggregation observed during inflammasome assembly discussed previously, and is under tight regulation to achieve the optimal balance between the antiviral response and unchecked tissue damaging inflammation (Reikine *et al.*, 2014). These regulatory mechanisms include posttranslational modifications such as polyubiquitination and phosphorylation/dephosphorylation of RIG-I and MDA5 that are critical for their immune signaling ability (Gack *et al.*, 2010; Zeng *et al.*, 2010; Wies *et al.*, 2013). Furthermore, signaling may be modulated by virally encoded factors that manipulate these posttranslational modifications (PTMs) or cause proteolytic degradation and cleavage of MAVS (Gack, 2014), as well as by numerous host proteins including the third RLR LGP2 that has been reported to have stimulatory and inhibitory effects on MDA5 and RIG-I signaling respectively (Zhu *et al.*, 2014). Further investigations are needed to clarify the role of LGP2 in antiviral immunity.

AIM2-Like Receptors and Other Intracellular DNA Sensors

Cytoplasmic DNA derived from infectious agents like viruses that replicate in the cytosol or from dying host cells can act as a danger signal to alert the immune system to infection or altered cellular homeostasis. A number of proteins that mediate cytoplasmic surveillance of host- or pathogen-derived DNA have been identified in recent years, however their role in host defense and/or autoimmunity and the mechanistic aspects of their activation are just beginning to be elucidated. These include the AIM2-like receptors (ALRs) (Hornung *et al.*, 2009), DAI (Takaoka *et al.*, 2007), and multiple DNA sensors that engage the ER-localized adaptor protein STING (Ishikawa and Barber, 2008; Holm *et al.*, 2013) resulting in the production of type 1 IFN.

STING binds cyclic dinucleotides (CDNs), second messengers synthesized by bacteria or by the receptor cyclic GMP-AMP synthase (cGAS), which synthesizes noncanonical CDNs from dsDNA (Yin *et al.*, 2012; Burdette *et al.*, 2011; Zhang *et al.*, 2014). STING contains an N-terminal transmembrane region and a cytosolic C-terminal domain (CTD) that interact in a dimerized complex in the absence of CDNs. Upon ligand binding, the CTD is released allowing it to interact with TBK1 and IRF3 (Ouyang *et al.*, 2012). Phosphorylated IRF3 in turn binds to the *Irnf* promoter, resulting in IFN- β synthesis (Figure 2). Several DNA sensors have been shown to act through the STING pathway, including several members of the ALR family discussed below.

DAI was the first TLR-independent cytosolic DNA sensor to be reported. It recognizes long strands (~500-1 kb in size) of dsDNA in its canonical B-form. Upon sensing of dsDNA by DAI, a TBK1 and IRF3-dependent signaling cascade is initiated, resulting in the production of type 1 IFN (Takaoka *et al.*, 2007). Whether or not STING is also required has not yet been fully elucidated. Activated DAI has also been shown to recruit the receptor-interacting protein kinases RIP1 and RIP3, leading to NF- κ B activation (Rebsamen *et al.*, 2009). Subsequently, the DDX41 DExD/H box helicase was also shown to bind intracellular B-form DNA or CDNs leading to a type 1 IFN response (Parvatiyar *et al.*, 2012; Zhang *et al.*, 2011). The signal is transduced by direct binding of activated DDX41 to STING which results in phosphorylation of IRF3 through TBK1. Cytosolic DNA may also be detected through the actions of RNA Polymerase III, which transcribes dsDNA to 5'-ppp RNA. The RNA is in turn sensed by RIG-I, resulting in a RIG-I-mediated IFN and NF- κ B response (Ablasser *et al.*, 2009; Chiu *et al.*, 2009). This pathway has been implicated in sensing intracellular bacteria, like *L. pneumophila*, and DNA viruses, like Epstein-Barr virus, Herpes simplex virus (HSV)-1, and adenovirus.

Five AIM2-like receptors (ALRs) have been identified in humans and 13 in mice. These are members of the Pyrin and HIN domain (PYHIN) family encoded by an IFN-inducible gene cluster on chromosome 1. In humans, they comprise of AIM2, IFN- γ inducible protein 16 (IFI16), myeloid cell nuclear differentiation antigen (MNDA), pyrin and HIN domain family member 1 (PYHIN1), and the recently identified pyrin domain only protein 3 (POP3) (Table 4). They are characterized by a N-terminal PYD and/or a C-terminal hematopoietic IFN-inducible nuclear protein (HIN)

domain (Figure 1). The HIN domain contains partially conserved repeats that allow for binding of DNA ligands resulting in activation of the inflammasome, type 1 IFNs or both depending upon the ALR. For instance, detection of DNA by AIM2 leads to assembly of an inflammasome resulting in activation of caspase-1 and production of active IL-1 β (Hornung *et al.*, 2009; Fernandes-Alnemri *et al.*, 2009), PYHIN1 activates STING-dependent IFN production (Brunette *et al.*, 2012), while IFI16 can lead to inflammasome activation or STING-dependent activation of type 1 IFN (Kerur *et al.*, 2011; Unterholzner *et al.*, 2010). AIM2 is perhaps the best studied of the ALRs. Recent structural studies have advanced our understanding of how AIM2 binds DNA to activate immune responses. Electrostatic interaction between positively charged HIN domain residues and the dsDNA sugar-phosphate backbone allows for nonsequence-specific DNA recognition by AIM2 (Jin *et al.*, 2012). In the inactive state, the HIN and PYD domains complex to maintain the molecule in an auto-inhibited conformation. Binding of dsDNA to the HIN domain opens the protein conformation releasing the PYD domain, which is then free to engage the adaptor protein ASC resulting in assembly of the AIM2 inflammasome that culminates in production of active IL-1 β and IL-18. However, the AIM2 inflammasome differs from the NLR inflammasomes in that the protein scaffold centers not on protein oligomerization domains but on multiple AIM2 molecules binding to the same dsDNA ligand. Nevertheless, AIM2 interaction with the adaptor ASC may initiate prion-like self-polymerization of ASC into large, stable filaments which have been shown to be both necessary and sufficient for AIM2 inflammasome responses (Cai *et al.*, 2014; Lu *et al.*, 2014). Given the polymeric nature of inflammasome assembly, it is anticipated that AIM2 activation must be under strict regulation; however, the regulatory mechanisms are poorly understood to date. Recently POP3, a newly identified member of the ALR gene cluster has been shown to negatively regulate DNA virus induced activation of the AIM2 and IFI16 inflammasomes by inhibiting recruitment of ASC to these ALRs (Khare *et al.*, 2014). Additionally, activation of IFI16 is also regulated by viral proteins that degrade IFI16 or block its oligomerization thereby interfering with the antiviral response (Orzalli *et al.*, 2012; Li *et al.*, 2013b).

Emerging data suggest that apart from host defense, DNA sensors may also be involved in development of autoimmune disease resulting from persistent triggering of these sensors by accumulated endogenous DNA (Holm *et al.*, 2013). Inefficient removal of chromatin due to defects in deoxyribonuclease 1 is associated with SLE (Napirei *et al.*, 2006), while mutations in the gene encoding three prime repair exonuclease 1 (TREX1), a negative regulator of STING signaling, result in Aicardi-Goutières syndrome (AGS), a severe neurological brain disease (Crow *et al.*, 2006). Thus, innate immune detection of DNA results in different functional outcomes, i.e., either host defense or pathology, depending upon the biological context. The field of DNA sensing is still burgeoning, and it will be interesting to gain a fuller understanding of the mechanisms underlying the beneficial and pathological consequences of innate recognition of DNA, while exploring the prospect of using nucleic acids as agents to promote antiviral immunity.

Role of Intracellular Organelles and Spatial Relocation

Activation of innate sensors following ligand binding may be manifested not only through changes in protein conformation and phosphorylation, but also through their spatial relocation (Figure 2). Following activation, the cytoplasmic sensors RIG-I, MDA5, and NLRP3 move from the cytoplasm to the mitochondria where they associate with MAVS (Seth *et al.*, 2005; Subramanian *et al.*, 2013). This relocation is essential for signal propagation; dislocation of MAVS from mitochondria abolishes RIG-I-dependent type 1 IFN and NLRP3-dependent inflammasome activation in response to RNA viruses (Seth *et al.*, 2005; Park *et al.*, 2013). AIM2 relocates from the cytosol to ASC foci upon activation by DNA. Likewise, all other murine ALRs translocate from the nucleus to puncta containing the adaptors ASC, STING, or both in discrete patterns under co-expression conditions (Brunette *et al.*, 2012). STING itself relocates within the cell following activation. In its inactive state, STING is present on the endoplasmic reticulum (ER), the mitochondria, and/or the mitochondria-associated ER membranes, which tether the ER to the mitochondrion (Ishikawa and Barber, 2008; Zhong *et al.*, 2008). Upon activation, STING moves to 'concentrated puncta' within the cell. These foci have been proposed to be associated with autophagy; however the association between STING activation and autophagy remains unclear (Burdette and Vance, 2013).

In addition to signal transduction, spatial relocation may also influence the outcome of signaling by determining the nature of the downstream response. Perhaps the best studied example is that of TLR4, which senses extracellular LPS. Activation of TLR4 on the plasma membrane results in its dimerization followed by binding to MyD88 and TIRAP/MAL adaptor proteins. This complex is trafficked to lipid rafts ultimately resulting in activation of NF- κ B and expression of pro-inflammatory cytokines. However, TLR4 may also engage the adaptors TRIF and TRAM resulting in phosphorylation of IRF3 and ultimately in production of type 1 IFN. For this latter pathway to be initiated, complex assembly must occur in early endosomes (Gangloff, 2012). Thus movement of TLR4 from the plasma membrane to the endosomes switches the outcome of signaling from production of NF- κ B-dependent cytokines to type 1 IFN. Likewise, TLR9 in most cells traffics from the ER through the golgi to proteolytically active early endosomal compartments where it is activated by cleavage of its ectodomain, facilitating signaling through NF- κ B to activate pro-inflammatory cytokines. In plasmacytoid DCs however, cleaved TLR9 in endosomes interacts with the adaptor protein (AP)-3 complex which delivers it to specialized lysosome-related organelles, where TLR9 can engage the TRAF3 and IRF7 signaling pathway resulting in induction of type 1 IFN (Honda *et al.*, 2005; Sasai *et al.*, 2010). A third example is that of the cytoplasmic RNA sensor RIG-I that can engage its adaptor MAVS from both peroxisomal and mitochondrial locations. Following viral infection, peroxisomal MAVS induces IRF1-dependent but type 1 IFN-independent rapid expression of IFN-stimulated genes (ISGs) and defense factors that may provide immediate protection, whereas mitochondrial MAVS activates IRF3 and type 1 IFN-dependent expression of ISGs

with delayed kinetics, that may serve to amplify and stabilize the antiviral response (Dixit *et al.*, 2010). More recently, it has been demonstrated that signaling via mitochondrial MAVS induces IFN- β and IFN- λ , while peroxisomal MAVS selectively activates an IFN- λ response (Odendall *et al.*, 2014). In all these instances, the intracellular movement of the innate sensor and/or its signaling adaptor to specific subcellular compartments determines the nature of the response.

The significance of this spatial relocation is not completely understood. It seems possible that the cellular location of ligand detection or pathogen replication is not always the best location for signal propagation. Nuclear ALRs, for instance, may be appropriately positioned to sense dsDNA in the nucleus where viral DNA is undergoing replication, however once activated move to a site where the downstream signaling components are located (Brunette *et al.*, 2012). Organelle membranes could provide ideal platforms for solid-phase oligomerization of innate sensors and/or their adaptors leading to assembly of active complexes, thereby allowing the input signal to be quickly amplified to ensure a robust innate response. Additional, as yet unidentified, molecules on organelle membranes might also be required for a complete functional response. Mitochondria, for instance, have been shown to directly contribute to bactericidal activity through the generation of ROS upon stimulation of a subset of TLRs. This response involves translocation of the TLR signaling adaptor, TRAF6, to mitochondria where it engages evolutionarily conserved signaling intermediate in Toll pathways (ECSIT), a protein that participates in mitochondrial respiratory chain assembly, resulting in increased mitochondrial ROS generation (West *et al.*, 2010). We speculate that relocation of innate immune signaling components to organelles such as mitochondria may contribute to or enhance microbicidal and/or other yet unknown responses. Mitochondria also play a role in regulating the apoptotic response to stress signals (Wang, 2001), and localization of innate signaling complexes on mitochondria could allow for the integration of signals from innate immune and cell death pathways.

Combinatorial Sensing and PRR Crosstalk

Innate sensing of pathogens is a complex, interconnected system of receptors, adaptors, and signaling pathways. A single pathogen may be detected by multiple receptors; for example, *Mycobacterium paratuberculosis* infection is sensed by TLR2, TLR4, and NOD2 (Ferwerda *et al.*, 2007) while rhinovirus infection of bronchial epithelium is sensed by TLR3, RIG-I, and MDA5 (Slater *et al.*, 2010). Activation of each of these receptors initiates distinct as well as common signaling cascades, which must be synthesized by the cell into an effective immune response. Crosstalk between innate immune signaling pathways is therefore an essential determinant of the quality and magnitude of the ensuing host response. Cooperation between multiple signaling pathways may be important for mounting an appropriate response to an elicitor. For example, concomitant activation of the CLR Dectin-1 and TLR2, both of which sense fungal cell wall components, is required to induce optimal cytokine responses in macrophages. This collaboration requires Dectin-1 signaling through

the Syk pathway, and can also occur with TLRs 4, 5, 7, and 9 that like TLR2, signal via the adaptor MyD88. Activation of Dectin-1 alone induces no TNF and activation TLR2 induces a low level of TNF. However activation of both Dectin-1 and TLR2 results in sustained degradation of I κ B, enhanced nuclear translocation of NF- κ B, and a synergistic TNF response (Dennehy *et al.*, 2008). In another study, collaboration between TLR2 and TLR9 was required for optimal host resistance to *Mycobacterium tuberculosis* (Bafica *et al.*, 2005). Signal integration from multiple pathways may thus be important for the cell to mount a scaled, robust, specifically tailored response to a pathogen, while decreasing potential 'noise' from commensal microbes and/or homeostatic cellular functions.

Sequential activation of multiple PRR pathways may enable the cell to generate an optimal response to contain a potential pathogen while minimizing immunopathology. For instance, in conventional DCs, a subset of herpes simplex viruses is first detected by cell surface TLR2 interacting with the virions, and then by intracellular TLR9 that senses viral genomic DNA (Sato *et al.*, 2006). In another study, sequential activation of MyD88-independent and MyD88-dependent cytokine responses by the intracellular bacterium *Listeria monocytogenes* was required for monocyte recruitment to infected tissue and monocyte activation respectively (Serbina *et al.*, 2003). Finally, NLRP3 inflammasome activation requires successive activation of the TLRs and NLRP3 (Netea *et al.*, 2009; Bauernfeind *et al.*, 2009). A TLR-activating stimulus first induces NF- κ B-dependent upregulation of pro-IL-1 β and NLRP3. A second NLRP3-activating stimulus then triggers assembly of the inflammasome resulting in processing of procaspase-1 to its active form, which in turn cleaves the pro form of IL-1 β to generate active cytokine. Both TLR and NLR pathways are therefore essential for induction of a complete IL-1 β response in macrophages.

Innate immune signaling pathways may also negatively regulate each other. For instance, the RLR-stimulated response has been shown to interfere with TLR signaling. The RLR-activated transcription factor IRF3 dominantly binds to the promoter region of the interleukin-12b (*Il12b*) gene, out-competing TLR-induced IRF5. As a consequence, activation of RLRs in mice attenuates TLR-induced T helper type 1 (Th1) and interleukin 17-producing helper T cell (Th17) responses, and pre-infection of mice with a virus reduces their ability to respond to bacterial pathogens effectively (Negishi *et al.*, 2012). Likewise, NLRX1 has been proposed to interact with TRAF6 thereby inhibiting NF- κ B signaling in response to LPS stimulation. Additionally, NLRX1 acts as a negative regulator of RIG-I signaling by interfering with RIG-I–MAVS interaction. Consequently, *Nlrp1* $-/-$ mice exhibit increased expression of antiviral signaling molecules and pro-inflammatory IL-6 after influenza virus infection, and this increased inflammation is associated with marked morbidity and tissue pathology in *Nlrp1* $-/-$ mice (Allen *et al.*, 2011). Negative regulation may also occur by sensors within the same PRR family. For example, within the TLR pathway, the adaptor MyD88 inhibits the TLR3-TRIF dependent response, which may protect the host from immunopathology associated with excessive production of IFN- β (Johnson *et al.*, 2008; Siednienko *et al.*, 2011). Within the NLR family, NLRP7 has been shown to negatively regulate the NLRP3 inflammasome by

direct interaction with procaspase-1 and pro IL-1 β (Radian *et al.*, 2013). Finally, although not as well characterized, several members of the NLR family have been proposed to have anti-inflammatory functions. NLRP6, NLRP10, NLRP12, NLRC3, and NLRX1 have all been reported to negatively regulate canonical or noncanonical NF- κ B activation and subsequent cytokine responses (Allen, 2014).

Crosstalk occurs not only between innate signaling pathways, but also between innate and other cellular pathways. For example, mTOR, which is usually involved in cell growth and metabolism, has also been shown to limit pro-inflammatory responses to bacterial stimuli by blocking NF- κ B activation and conversely increase anti-inflammatory cytokines like IL-10 by enhancing STAT3 activity, thereby limiting the priming of Th1 and Th17 cells (Weichhart *et al.*, 2008). Moreover, in the context of an infection, crosstalk also occurs between the host and pathogen. Many pathogens can modulate the host response, allowing them to overcome or bypass host defense mechanisms. For instance, the human cytomegalovirus actively cleaves host DNA and disrupts the ability of the infected cell to repair it, while selectively maintaining nucleotide excision repair of the viral genome (O'Dowd *et al.*, 2012). Likewise, many viral pathogens actively cleave or sequester components of innate immune pathways to disrupt signaling, while intracellular bacterial pathogens such as *Salmonella enterica* and *Mycobacterium tuberculosis* have developed strategies to avoid, modulate, or hijack the host immune response in their favor (Forrellad *et al.*, 2013; Gack, 2014; de Jong *et al.*, 2012; Figure 3). Interplay between PRR and other cellular pathways as well as between the host and the pathogen, are therefore essential parameters that influence the outcome of infection.

Multiscale Regulation of Cytosolic Sensing Pathways

Successful host defense against pathogens requires development of an appropriate, scaled immune response commensurate with the nature of the threat, a process that involves dynamic, feedback-controlled interactions between immune system components at multiple molecular and cellular levels. Information is integrated from the genomic, transcriptional and translational levels, posttranslational controls, protein–protein interactions, assembly of functional multiprotein complexes and their appropriate cellular localization, to coordinate a successful effector response against the potential pathogen, followed by induction of appropriate negative regulatory mechanisms to restore cellular homeostasis (Figure 3). Information from heterogeneous cells must further be integrated into an organ- and organism-level response. Regulation at each level affects the response at many others; therefore a functional immune response is the result of an orchestrated set of events, with combinatorial regulatory potential. Deciphering the intricate physical and functional interactions that underlie innate immunity will therefore require a global, systems-level understanding of the molecular and cellular networks that comprise the processes at all levels (Subramanian *et al.*, 2015).

This seems a daunting endeavor; however, recent technological advances have allowed for the collection of comprehensive, large-scale ‘omics’ datasets at multiple molecular and cellular levels. Coupled with major developments in

computational modeling and network analysis methodologies, they hold the potential to greatly advance our understanding of innate immune regulation by enabling a systems approach to innate immunity. Genomics, which involves sequencing of entire genomes, has enabled GWAS that permit unbiased association of genetic variants commonly arising due to single nucleotide polymorphisms with disease traits (Xavier and Rioux, 2008). Such studies have revealed association of polymorphisms in cytosolic sensors, particularly the NLRs, with a plethora of autoimmune and autoinflammatory diseases (Zhong *et al.*, 2013; Table 3). A notable example is the association of NOD2 genetic variants with susceptibility to Crohn's disease (Xavier and Rioux, 2008). Transcriptomics, the global analysis of gene expression pathways, has resulted in the identification of novel transcription factors, *cis*-regulatory elements, and previously unknown signaling modules involved in innate immune responses, while ChIP-Seq (chromatin immunoprecipitation followed by sequencing) approaches have allowed for identification of genome-wide binding sites for these transcription factors (Gilchrist *et al.*, 2006; Elkon *et al.*, 2007; Ramsey *et al.*, 2008; Gottschalk *et al.*, 2013). Together with network perturbation strategies such as RNAi, these technologies have allowed for elucidation of gene-regulatory networks controlling the innate immune response (Amit *et al.*, 2011). Transcript profiling has also been used to elucidate immunological responses to vaccine antigens, with the goal of predicting vaccine efficacy and potentially optimizing the immunogenicity of vaccines (Li *et al.*, 2013a). For instance, systems biology approaches have been used to determine the mechanisms involved in conferring immunity to yellow fever (Querec *et al.*, 2009; Gaucher *et al.*, 2008), seasonal influenza (Nakaya *et al.*, 2011) and Human immunodeficiency virus (HIV) (Zak *et al.*, 2012) and to develop computational models that can predict vaccine efficacy in subsequent, independent trials.

Advances in mass spectrometry have further enabled collection of large-scale proteomic datasets for quantification and characterization of proteins and posttranslational modifications (PTMs) in response to various stimuli. For instance, new developments in proteomics technologies have resulted in exciting discoveries of phosphorylation sites and led to an increased understanding of the role of phosphorylation in the immune response (Gottschalk *et al.*, 2013). Formation of functional multiprotein complexes (involving receptor dimerization/oligomerization, interaction with adaptor proteins, and other protein–protein interactions necessary for assembly of the signaling complex) is an essential step that controls signal propagation and critically determines the quality and/or magnitude of the innate response. Protein expression, protein–protein interaction, and PTM data can provide insights into assembly of such functional complexes, and when analyzed using network analysis tools, can also help elucidate complex signaling circuits underlying the innate response (Subramanian *et al.*, 2015). Finally, metabolomics, the comprehensive analysis of metabolites such as metabolic intermediates, hormones, and secondary metabolites, has been used to better understand host–pathogen interactions (Olszewski *et al.*, 2009) and to develop metabolic biomarkers for diseases (Monteiro *et al.*, 2013). Although metabolomics has not been widely employed in the study of innate immunity, it is

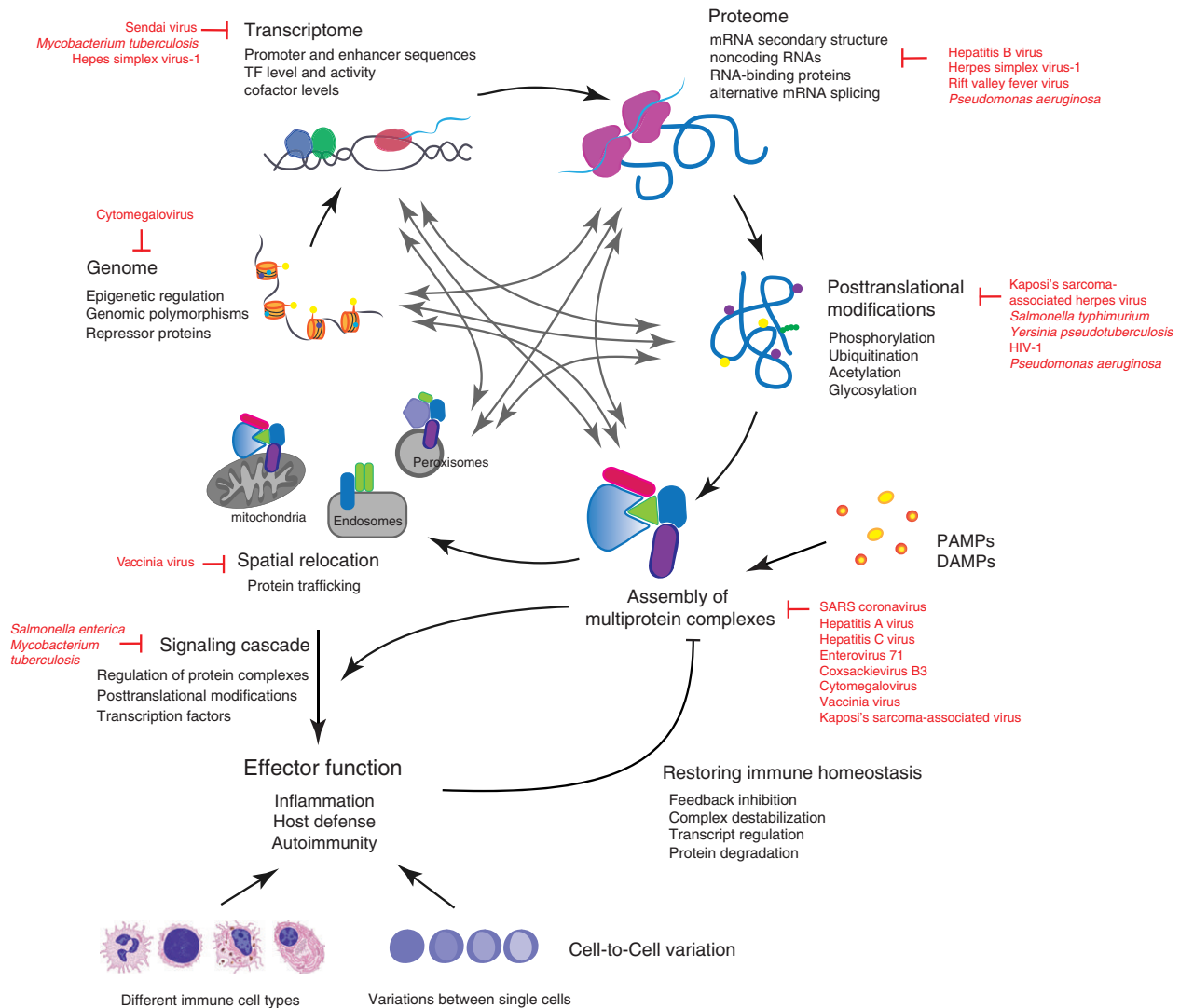


Figure 3 Systems-level overview of innate immunity. Development of a functional immune response is dependent on interactions between cellular components and processes at multiple levels. Modifications at the genomic level, such as genetic polymorphisms, affect the sequence of the gene being transcribed, while epigenetic modifications or other effects on DNA regulation influence the amplitude of transcription. An RNA transcript may be regulated by miRNAs, lncRNAs, and other noncoding RNAs, mRNA secondary structure or splicing events that may have profound effects on translation and protein expression. A protein may be further modified posttranslationally by addition of biochemical functional groups (such as phosphates or acetates) or through structural changes (such as formation of disulfide bonds). Sensing of PAMPs and DAMPs by innate sensors initiates the assembly of multiprotein signaling complexes, and the signal may be further propagated through trafficking or spatial relocation of the protein(s) or protein complex to intracellular compartments. Formation of functional signaling complexes is therefore a highly regulated process that results from an integrated set of intracellular events like transcription, translation, posttranslational modifications (PTMs), proper protein trafficking, and localization. The nature and magnitude of the response may also vary depending upon the type of innate immune cell encountered (e.g., macrophage, neutrophil, and B cell) and stochastic variations between single cells of the same type. Pathogens may further modulate or evade the immune response by interfering with host defense mechanisms at multiple levels; some examples are indicated in red. The regulation and integration of these events occurring at various levels determines the eventual effector response that may either result in host protection and restoration of immune homeostasis or unrestrained inflammation leading to tissue damage and/or autoimmunity.

now appreciated that cytosolic sensors, particularly the NLRs play an unequivocal role in sensing danger signals produced during metabolic stress. A role for metabolism in controlling inflammation has also emerged (Kominsky *et al.*, 2010), and future studies should provide greater insight into the metabolic networks controlling innate immunity.

Integrative analysis of high-dimensional data acquired from multiple molecular levels in this manner will be key to

elucidating the complex spatiotemporal interactions that underlie the development of a functional immune response. Advances in network biology have facilitated representation of the intricate interactions at each level as physical or functional networks of interconnected components, which enables recognition of emergent properties arising from these interactions (Subramanian *et al.*, 2015). When substantiated by focused experimentation, this approach holds great promise for accelerating

our understanding of the transcriptional, gene-regulatory, and signaling networks controlling the innate immune response. Integration of the aforementioned global approaches with advances in network biology, computational, and mathematical tools, may thus in future endeavors permit the generation of quantitative and predictive models of innate immunity, which in turn, will empower vaccinology, drug discovery, and other therapeutic possibilities for treatment of human immune disorders.

Acknowledgments

This work was supported by the Institute for Systems Biology.

See also: Cellular Immunology: Innate Immunity: Scavenger Receptors. Cellular Immunology: Overview: Introduction to Functional Cell Biology of Immunity

References

- Abbott, D.W., Yang, Y., Hutti, J.E., *et al.*, 2007. Coordinated regulation of toll-like receptor and NOD2 signaling by K63-linked polyubiquitin chains. *Molecular and Cellular Biology* 27, 6012–6025.
- Ablasser, A., Bauernfeind, F., Hartmann, G., *et al.*, 2009. RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate. *Nature Immunology* 10, 1065–1072.
- Agostini, L., Martinon, F., Burns, K., *et al.*, 2004. NALP3 forms an IL-1 β -processing inflammasome with increased activity in Muckle-Wells autoinflammatory disorder. *Immunity* 20, 319–325.
- Akira, S., Uematsu, S., Takeuchi, O., 2006. Pathogen recognition and innate immunity. *Cell* 124, 783–801.
- Alileche, A., Squires, R.C., Muehlbauer, S.M., Lisanti, M.P., Brojtsch, J., 2006. Mitochondrial impairment is a critical event in anthrax lethal toxin-induced cytolysis of murine macrophages. *Cell Cycle* 5, 100–106.
- Allen, I.C., 2014. Non-inflammasome forming NLRs in inflammation and tumorigenesis. *Frontiers in Immunology* 5, 169.
- Allen, I.C., Moore, C.B., Schneider, M., *et al.*, 2011. NLRX1 protein attenuates inflammatory responses to infection by interfering with the RIG-I-MAVS and TRAF6-NF- κ B signaling pathways. *Immunity* 34, 854–865.
- Amit, I., Regev, A., Hacohen, N., 2011. Strategies to discover regulatory circuits of the mammalian immune system. *Nature Reviews Immunology* 11, 873–880.
- Andrade, W.A., Souza Mdo, C., Ramos-Martinez, E., *et al.*, 2013. Combined action of nucleic acid-sensing Toll-like receptors and TLR11/TLR12 heterodimers imparts resistance to *Toxoplasma gondii* in mice. *Cell Host & Microbe* 13, 42–53.
- Averette, K.M., Pratt, M.R., Yang, Y., *et al.*, 2009. Anthrax lethal toxin induced lysosomal membrane permeabilization and cytosolic cathepsin release is Nlrp1b/Nalp1b-dependent. *PLoS One* 4, e7913.
- Bafica, A., Scanga, C.A., Feng, C.G., *et al.*, 2005. TLR9 regulates Th1 responses and cooperates with TLR2 in mediating optimal resistance to *Mycobacterium tuberculosis*. *Journal of Experimental Medicine* 202, 1715–1724.
- Basith, S., Manavalan, B., Yoo, T.H., Kim, S.G., Choi, S., 2012. Roles of toll-like receptors in cancer: A double-edged sword for defense and offense. *Archives of Pharmacological Research* 35, 1297–1316.
- Bauernfeind, F., Bartok, E., Rieger, A., *et al.*, 2011. Cutting edge: Reactive oxygen species inhibitors block priming, but not activation, of the NLRP3 inflammasome. *Journal of Immunology* 187, 613–617.
- Bauernfeind, F.G., Horvath, G., Stutz, A., *et al.*, 2009. Cutting edge: NF- κ B activating pattern recognition and cytokine receptors license NLRP3 inflammasome activation by regulating NLRP3 expression. *Journal of Immunology* 183, 787–791.
- Blasius, A.L., Beutler, B., 2010. Intracellular Toll-like receptors. *Immunity* 32, 305–315.
- Broz, P., Von Moltke, J., Jones, J.W., Vance, R.E., Monack, D.M., 2010. Differential requirement for Caspase-1 autoproteolysis in pathogen-induced cell death and cytokine processing. *Cell Host & Microbe* 8, 471–483.
- Bruey, J.M., Bruey-Sedano, N., Luciano, F., *et al.*, 2007. Bcl-2 and Bcl-XL regulate proinflammatory caspase-1 activation by interaction with NALP1. *Cell* 129, 45–56.
- Brunette, R.L., Young, J.M., Whitley, D.G., *et al.*, 2012. Extensive evolutionary and functional diversity among mammalian AIM2-like receptors. *Journal of Experimental Medicine* 209, 1969–1983.
- Burdette, D.L., Monroe, K.M., Sotelo-Troha, K., *et al.*, 2011. STING is a direct innate immune sensor of cyclic di-GMP. *Nature* 478, 515–518.
- Burdette, D.L., Vance, R.E., 2013. STING and the innate immune response to nucleic acids in the cytosol. *Nature Immunology* 14, 19–26.
- Cai, X., Chen, J., Xu, H., *et al.*, 2014. Prion-like polymerization underlies signal transduction in antiviral immune defense and inflammasome activation. *Cell* 156, 1207–1222.
- Cassel, S.L., Eisenbarth, S.C., Iyer, S.S., *et al.*, 2008. The Nalp3 inflammasome is essential for the development of silicosis. *Proceedings of the National Academy of Sciences* 105, 9035–9040.
- Chiu, Y.-H., Macmillan, J.B., Chen, Z.J., 2009. RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. *Cell* 138, 576–591.
- Corridoni, D., Arseneau, K.O., Cifone, M.G., Cominelli, F., 2014. The dual role of nod-like receptors in mucosal innate immunity and chronic intestinal inflammation. *Frontiers in Immunology* 5, 317.
- Crow, Y.J., Hayward, B.E., Parmar, R., *et al.*, 2006. Mutations in the gene encoding the 3′-5′ DNA exonuclease TREX1 cause Aicardi-Goutieres syndrome at the AGS1 locus. *Nature Genetics* 38, 917–920.
- Cruz, C.M., Rinna, A., Forman, H.J., *et al.*, 2007. ATP activates a reactive oxygen species-dependent oxidative stress response and secretion of proinflammatory cytokines in macrophages. *Journal of Biological Chemistry* 282, 2871–2879.
- Dennehy, K.M., Ferwerda, G., Faro-Trindade, I., *et al.*, 2008. Syk kinase is required for collaborative cytokine production induced through Dectin-1 and Toll-like receptors. *European Journal of Immunology* 38, 500–506.
- Dixit, E., Boulant, S., Zhang, Y., *et al.*, 2010. Peroxisomes are signaling platforms for antiviral innate immunity. *Cell* 141, 668–681.
- Dostert, C., Petrilli, V., Van Bruggen, R., *et al.*, 2008. Innate immune activation through Nalp3 inflammasome sensing of asbestos and silica. *Science* 320, 674–677.
- Duewell, P., Kono, H., Rayner, K.J., *et al.*, 2010. NLRP3 inflammasomes are required for atherogenesis and activated by cholesterol crystals. *Nature* 464, 1357–1361.
- Elinav, E., Strowig, T., Henao-Mejia, J., Flavell, R., 2011. Regulation of the antimicrobial response by NLR proteins. *Immunity* 34, 665–679.
- Elkon, R., Linhart, C., Halperin, Y., Shiloh, Y., Shamir, R., 2007. Functional genomic delineation of TLR-induced transcriptional networks. *BMC Genomics* 8, 394.
- Faustin, B., Lartigue, L., Bruey, J.M., *et al.*, 2007. Reconstituted NALP1 inflammasome reveals two-step mechanism of caspase-1 activation. *Molecular Cell* 25, 713–724.
- Fernandes-Alnemri, T., Yu, J.-W., Datta, P., Wu, J., Alnemri, E.S., 2009. AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. *Nature* 458, 509–513.
- Ferwerda, G., Kullberg, B.J., De Jong, D.J., *et al.*, 2007. *Mycobacterium paratuberculosis* is recognized by Toll-like receptors and NOD2. *Journal of Leukocyte Biology* 82, 1011–1018.
- Forrellad, M.A., Klepp, L.I., Goffre, A., *et al.*, 2013. Virulence factors of the *Mycobacterium tuberculosis* complex. *Virulence* 4, 3–66.
- Franchi, L., Amer, A., Body-Malapel, M., *et al.*, 2006. Cytosolic flagellin requires Ipaf for activation of caspase-1 and interleukin 1 β in salmonella-infected macrophages. *Nature Immunology* 7, 576–582.
- Gack, M.U., 2014. Mechanisms of RIG-I-like receptor activation and manipulation by viral pathogens. *Journal of Virology* 88, 5213–5216.
- Gack, M.U., Nistal-Villan, E., Inn, K.S., Garcia-Sastre, A., Jung, J.U., 2010. Phosphorylation-mediated negative regulation of RIG-I antiviral activity. *Journal of Virology* 84, 3220–3229.
- Gangloff, M., 2012. Different dimerisation mode for TLR4 upon endosomal acidification? *Trends in Biochemical Sciences* 37, 92–98.
- Gaucher, D., Therrien, R., Kettaf, N., *et al.*, 2008. Yellow fever vaccine induces integrated multilineage and polyfunctional immune responses. *Journal of Experimental Medicine* 205, 3119–3131.
- Gerlic, M., Faustin, B., Postigo, A., *et al.*, 2013. Vaccinia virus F1L protein promotes virulence by inhibiting inflammasome activation. *Proceedings of the National Academy of Sciences of the United States of America* 110, 7808–7813.
- Gilchrist, M., Thorsson, V., Li, B., *et al.*, 2006. Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. *Nature* 441, 173–178.
- Girardin, S.E., Tournibize, R., Mavris, M., *et al.*, 2001. CARD4/Nod1 mediates NF- κ B and JNK activation by invasive *Shigella flexneri*. *EMBO Reports* 2, 736–742.

- Gottschalk, R.A., Martins, A.J., Sjoelund, V.H., *et al.*, 2013. Recent progress using systems biology approaches to better understand molecular mechanisms of immunity. *Seminars in Immunology* 25, 201–208.
- Gregory, S.M., Davis, B.K., West, J.A., *et al.*, 2011. Discovery of a viral NLR homolog that inhibits the inflammasome. *Science* 331, 330–334.
- Gross, O., Thomas, C.J., Guarda, G., Tschopp, J., 2011. The inflammasome: An integrated view. *Immunological Reviews* 243, 136–151.
- Guarda, G., Braun, M., Staehli, F., *et al.*, 2011. Type I interferon inhibits interleukin-1 production and inflammasome activation. *Immunity* 34, 213–223.
- Guarda, G., Dostert, C., Staehli, F., *et al.*, 2009. T cells dampen innate immune responses through inhibition of NLRP1 and NLRP3 inflammasomes. *Nature* 460, 269–273.
- Harton, J.A., Linhoff, M.W., Zhang, J., Ting, J. P.-Y., 2002. Cutting edge: CATERPILLER: a large family of mammalian genes containing CARD, pyrin, nucleotide-binding, and leucine-rich repeat domains. *Journal of Immunology* 169, 4088–4093.
- Heneka, M.T., Kummer, M.P., Stutz, A., *et al.*, 2013. NLRP3 is activated in Alzheimer's disease and contributes to pathology in APP/PS1 mice. *Nature* 493, 674–678.
- Holm, C.K., Paludan, S.R., Fitzgerald, K.A., 2013. DNA recognition in immunity and disease. *Current Opinion in Immunology* 25, 13–18.
- Honda, K., Ohba, Y., Yanai, H., *et al.*, 2005. Spatiotemporal regulation of MyD88-IRF-7 signalling for robust type-I interferon induction. *Nature* 434, 1035–1040.
- Hornung, V., Ablasser, A., Charrel-Dennis, M., *et al.*, 2009. AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* 458, 514–518.
- Hornung, V., Bauernfeind, F., Halle, A., *et al.*, 2008. Silica crystals and aluminum salts activate the NALP3 inflammasome through phagosomal destabilization. *Nature Immunology* 9, 847–856.
- Hornung, V., Ellegast, J., Kim, S., *et al.*, 2006. 5'-Triphosphate RNA is the ligand for RIG-I. *Science* 314, 994–997.
- Hou, F., Sun, L., Zheng, H., *et al.*, 2011. MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response. *Cell* 146, 448–461.
- Hoving, J.C., Wilson, G.J., Brown, G.D., 2014. Signalling C-Type lectin receptors, microbial recognition and immunity. *Cellular Microbiology* 16, 185–194.
- Hsu, L.C., Ali, S.R., McGivillray, S., *et al.*, 2008. A NOD2-NALP1 complex mediates caspase-1-dependent IL-1 β secretion in response to *Bacillus anthracis* infection and muramyl dipeptide. *Proceedings of the National Academy of Sciences of the United States of America* 105, 7803–7808.
- Inohara, N., Nunez, G., 2003. NODs: Intracellular proteins involved in inflammation and apoptosis. *Nature Reviews Immunology* 3, 371–382.
- Ishii, K.J., Koyama, S., Nakagawa, A., Coban, C., Akira, S., 2008. Host innate immune receptors and beyond: Making sense of microbial infections. *Cell Host & Microbe* 3, 352–363.
- Ishikawa, H., Barber, G.N., 2008. STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature* 455, 674–678.
- Iwasaki, A., Medzhitov, R., 2010. Regulation of adaptive immunity by the innate immune system. *Science* 327, 291–295.
- Jin, T., Perry, A., Jiang, J., *et al.*, 2012. Structures of the HIN domain:DNA complexes reveal ligand binding and activation mechanisms of the AIM2 inflammasome and IFI16 receptor. *Immunity* 36, 561–571.
- Johnson, A.C., Li, X., Pearlman, E., 2008. MyD88 functions as a negative regulator of TLR3/TRIF-induced corneal inflammation by inhibiting activation of c-Jun N-terminal kinase. *Journal of Biological Chemistry* 283, 3988–3996.
- de Jong, H.K., Parry, C.M., Van Der Poll, T., Wiersinga, W.J., 2012. Host-pathogen interaction in invasive Salmonellosis. *PLoS Pathogens* 8, e1002933.
- Juliana, C., Fernandes-Alnemri, T., Kang, S., *et al.*, 2012. Non-transcriptional priming and deubiquitination regulate NLRP3 inflammasome activation. *Journal of Biological Chemistry* 287, 36617–36622.
- Kato, H., Sato, S., Yoneyama, M., *et al.*, 2005. Cell type-specific involvement of RIG-I in antiviral response. *Immunity* 23, 19–28.
- Kato, H., Takeuchi, O., Mikamo-Sato, E., *et al.*, 2008. Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. *Journal of Experimental Medicine* 205, 1601–1610.
- Kato, H., Takeuchi, O., Sato, S., *et al.*, 2006. Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* 441, 101–105.
- Kawai, T., Akira, S., 2011. Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. *Immunity* 34, 637–650.
- Kerur, N., Veettil, M.V., Sharma-Walia, N., *et al.*, 2011. IFI16 acts as a nuclear pathogen sensor to induce the inflammasome in response to Kaposi Sarcoma-associated herpesvirus infection. *Cell Host & Microbe* 9, 363–375.
- Khare, S., Ratsimandresy, R.A., Almeida, D.e., *et al.*, 2014. The PYRIN domain-only protein POP3 inhibits ALR inflammasomes and regulates responses to infection with DNA viruses. *Nature Immunology* 15, 343–353.
- Kofoed, E.M., Vance, R.E., 2011. Innate immune recognition of bacterial ligands by NALPs determines inflammasome specificity. *Nature* 477, 592–595.
- Kominsky, D.J., Campbell, E.L., Colgan, S.P., 2010. Metabolic shifts in immunity and inflammation. *Journal of Immunology* 184, 4062–4068.
- Kowalinski, E., Lunardi, T., McCarthy, Andrew A., *et al.*, 2011. Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. *Cell* 147, 423–435.
- Kufer, T.A., Sansonetti, P.J., 2011. NLR functions beyond pathogen recognition. *Nature Immunology* 12, 121–128.
- Lamkanfi, M., Dixit, V.M., 2014. Mechanisms and functions of inflammasomes. *Cell* 157, 1013–1022.
- Lee, G.S., Subramanian, N., Kim, A.I., *et al.*, 2012. The calcium-sensing receptor regulates the NLRP3 inflammasome through Ca²⁺ and cAMP. *Nature* 492, 123–127.
- Li, M., Zhou, Y., Feng, G., Su, S.B., 2009. The critical role of Toll-like receptor signaling pathways in the induction and progression of autoimmune diseases. *Current Molecular Medicine* 9, 365–374.
- Li, S., Nakaya, H.I., Kazmin, D.A., Oh, J.Z., Pulendran, B., 2013a. Systems biological approaches to measure and understand vaccine immunity in humans. *System Immunology* 25, 209–218.
- Li, T., Chen, J., Cristea, I.M., 2013b. Human cytomegalovirus tegument protein pUL83 inhibits IFI16-mediated DNA sensing for immune evasion. *Cell Host & Microbe* 14, 591–599.
- Lich, J.D., Ting, J.P., 2007. Monarch-1/PYPAF7 and other CATERPILLER (CLR, NOD, NLR) proteins with negative regulatory functions. *Microbes and Infection* 9, 672–676.
- Loo, Y.-M., Fornek, J., Crochet, N., *et al.*, 2008. Distinct RIG-I and MDA5 signaling by RNA viruses in innate immunity. *Journal of Virology* 82, 335–345.
- Lopez-Castejon, G., Luhschi, N.M., Compan, V., *et al.*, 2013. Deubiquitinases regulate the activity of caspase-1 and interleukin-1 β secretion via assembly of the inflammasome. *Journal of Biological Chemistry* 288, 2721–2733.
- Lu, A., Magupalli, V.G., Ruan, J., *et al.*, 2014. Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes. *Cell* 156, 1193–1206.
- Martinon, F., Burns, K., Tschopp, J., 2002. The inflammasome: A molecular platform triggering activation of inflammatory caspases and processing of proIL- β . *Molecular Cell* 10, 417–426.
- Martinon, F., Mayor, A., Tschopp, J., 2009. The inflammasomes: Guardians of the body. *Annual Review of Immunology* 27, 229–265.
- Masters, S.L., Dunne, A., Subramanian, S.L., *et al.*, 2010. Activation of the NLRP3 inflammasome by islet amyloid polypeptide provides a mechanism for enhanced IL-1 β in type 2 diabetes. *Nature Immunology* 11, 897–904.
- Miao, E.A., Alpujch-Aranda, C.M., Dors, M., *et al.*, 2006. Cytoplasmic flagellin activates caspase-1 and secretion of interleukin 1 β via Ipaf. *Nature Immunology* 7, 569–575.
- Miao, E.A., Leaf, I.A., Treuting, P.M., *et al.*, 2010a. Caspase-1-induced pyroptosis is an innate immune effector mechanism against intracellular bacteria. *Nature Immunology* 11, 1136–1142.
- Miao, E.A., Mao, D.P., Yudkovsky, N., *et al.*, 2010b. Innate immune detection of the type III secretion apparatus through the NLRC4 inflammasome. *Proceedings of the National Academy of Sciences of the United States of America* 107, 3076–3080.
- Monteiro, M.S., Carvalho, M., Bastos, M.L., Guedes De Pinho, P., 2013. Metabolomics analysis for biomarker discovery: Advances and challenges. *Current Medicinal Chemistry* 20, 257–271.
- Muehlbauer, S.M., Evering, T.H., Bonuccelli, G., *et al.*, 2007. Anthrax lethal toxin kills macrophages in a strain-specific manner by apoptosis or caspase-1-mediated necrosis. *Cell Cycle* 6, 758–766.
- Nakaya, H.I., Wrammert, J., Lee, E.K., *et al.*, 2011. Systems biology of vaccination for seasonal influenza in humans. *Nature Immunology* 12, 786–795.
- Napirei, M., Gultekin, A., Kloeckl, T., *et al.*, 2006. Systemic lupus-erythematosus: Deoxyribonuclease 1 in necrotic chromatin disposal. *International Journal of Biochemistry & Cell Biology* 38, 297–306.
- Negishi, H., Yanai, H., Nakajima, A., *et al.*, 2012. Cross-interference of RLR and TLR signaling pathways modulates antibacterial T cell responses. *Nature Immunology* 13, 659–666.
- Netea, M.G., Nold-Petry, C.A., Nold, M.F., *et al.*, 2009. Differential requirement for the activation of the inflammasome for processing and release of IL-1 β in monocytes and macrophages. *Blood* 113, 2324–2335.

- O'Dowd, J.M., Zavala, A.G., Brown, C.J., Mori, T., Fortunato, E.A., 2012. HCMV-infected cells maintain efficient nucleotide excision repair of the viral genome while abrogating repair of the host genome. *PLoS Pathogens* 8, e1003038.
- O'Neill, L.A., Golenbock, D., Bowie, A.G., 2013. The history of Toll-like receptors – Redefining innate immunity. *Nature Reviews Immunology* 13, 453–460.
- Odendall, C., Dixit, E., Stavru, F., *et al.*, 2014. Diverse intracellular pathogens activate type III interferon expression from peroxisomes. *Nature Immunology* 15, 717–726.
- Oliszewski, K.L., Morrissey, J.M., Wilinski, D., *et al.*, 2009. Host-parasite interactions revealed by *Plasmodium falciparum* metabolomics. *Cell Host & Microbe* 5, 191–199.
- Orzalli, M.H., Deluca, N.A., Knipe, D.M., 2012. Nuclear IFI16 induction of IRF-3 signaling during herpesviral infection and degradation of IFI16 by the viral ICP0 protein. *Proceedings of the National Academy of Sciences of the United States of America* 109, E3008–E3017.
- Osorio, F., Reis, E., Sousa, C., 2011. Myeloid C-type lectin receptors in pathogen recognition and host defense. *Immunity* 34, 651–664.
- Ouyang, S., Song, X., Wang, Y., *et al.*, 2012. Structural analysis of the STING adaptor protein reveals a hydrophobic dimer interface and mode of cyclic di-GMP binding. *Immunity* 36, 1073–1086.
- Park, J.-H., Kim, Y.-G., Shaw, M., *et al.*, 2007. Nod1/RICK and TLR signaling regulate chemokine and antimicrobial innate immune responses in mesothelial cells. *Journal of Immunology* 179, 514–521.
- Park, S., Juliana, C., Hong, S., *et al.*, 2013. The mitochondrial antiviral protein MAVS associates with NLRP3 and regulates its inflammasome activity. *Journal of Immunology* 191, 4358–4366.
- Parvatiyar, K., Zhang, Z., Teles, R.M., *et al.*, 2012. The helicase DDX41 recognizes the bacterial secondary messengers cyclic di-GMP and cyclic di-AMP to activate a type I interferon immune response. *Nature Immunology* 13, 1155–1161.
- Pauleau, A.L., Murray, P.J., 2003. Role of nod2 in the response of macrophages to Toll-like receptor agonists. *Molecular and Cellular Biology* 23, 7531–7539.
- Pichlmair, A., Schulz, O., Tan, C.P., *et al.*, 2006. RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. *Science* 314, 997–1001.
- Pifer, R., Benson, A., Sturge, C.R., Yarovsky, F., 2011. UNC93B1 is essential for TLR11 activation and IL-12-dependent host resistance to *Toxoplasma gondii*. *Journal of Biological Chemistry* 286, 3307–3314.
- Qu, Y., Misaghi, S., Izrael-Tomasevic, A., *et al.*, 2012. Phosphorylation of NLRC4 is critical for inflammasome activation. *Nature* 490, 539–542.
- Querec, T.D., Akondy, R.S., Lee, E.K., *et al.*, 2009. Systems biology approach predicts immunogenicity of the yellow fever vaccine in humans. *Nature Immunology* 10, 116–125.
- Radian, A.D., De Almeida, L., Dorfleutner, A., Stehlik, C., 2013. NLRP7 and related inflammasome activating pattern recognition receptors and their function in host defense and disease. *Microbes and Infection* 15, 630–639.
- Ramsey, S.A., Klemm, S.L., Zak, D.E., *et al.*, 2008. Uncovering a macrophage transcriptional program by integrating evidence from motif scanning and expression dynamics. *PLoS Computational Biology* 4, e1000021.
- Rebsamen, M., Heinz, L.X., Meylan, E., *et al.*, 2009. DAI/ZBP1 recruits RIP1 and RIP3 through RIP homotypic interaction motifs to activate NF- κ B. *EMBO Reports* 10, 916–922.
- Redelinghuys, P., Brown, G.D., 2011. Inhibitory C-type lectin receptors in myeloid cells. *Immunology Letters* 136, 1–12.
- Reikine, S., Nguyen, J.B., Modis, Y., 2014. Pattern recognition and signaling mechanisms of RIG-I and MDA5. *Frontiers in Immunology* 5, 342.
- Sabbah, A., Chang, T.H., Harnack, R., *et al.*, 2009. Activation of innate immune antiviral responses by Nod2. *Nature Immunology* 10, 1073–1080.
- Sasai, M., Linehan, M.M., Iwasaki, A., 2010. Bifurcation of Toll-like receptor 9 signaling by adaptor protein 3. *Science* 329, 1530–1534.
- Sato, A., Linehan, M.M., Iwasaki, A., 2006. Dual recognition of herpes simplex viruses by TLR2 and TLR9 in dendritic cells. *Proceedings of the National Academy of Sciences* 103, 17343–17348.
- Schroder, K., Tschopp, J., 2010. The inflammasomes. *Cell* 140, 821–832.
- Serbina, N.V., Kuziel, W., Flavell, R., *et al.*, 2003. Sequential MyD88-independent and -dependent activation of innate immune responses to intracellular bacterial infection. *Immunity* 19, 891–901.
- Seth, R.B., Sun, L., Ea, C.-K., Chen, Z.J., 2005. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF- κ B and IRF3. *Cell* 122, 669–682.
- Siednienko, J., Gajanayake, T., Fitzgerald, K.A., Moynagh, P., Miggin, S.M., 2011. Absence of MyD88 results in enhanced TLR3-dependent phosphorylation of IRF3 and increased IFN- β and RANTES production. *Journal of Immunology* 186, 2514–2522.
- Slater, L., Bartlett, N.W., Haas, J.J., *et al.*, 2010. Co-ordinated role of TLR3, RIG-I and MDA5 in the innate response to rhinovirus in bronchial epithelium. *PLoS Pathogens* 6, e1001178.
- Subramanian, N., Natarajan, K., Clatworthy, M.R., Wang, Z., Germain, R.N., 2013. The adaptor MAVS promotes NLRP3 mitochondrial localization and inflammasome activation. *Cell* 153, 348–361.
- Subramanian, N., Torabi-Parizi, P., Gottschalk, R.A., Germain, R.N., Dutta, B., 2015. Network representations of immune system complexity. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 7, 13–38.
- Takaoka, A., Wang, Z., Choi, M.K., *et al.*, 2007. DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. *Nature* 448, 501–505.
- Unterholzner, L., Keating, S.E., Baran, M., *et al.*, 2010. IFI16 is an innate immune sensor for intracellular DNA. *Nature Immunology* 11, 997–1004.
- Vandanmagsar, B., Youm, Y.H., Ravussin, A., *et al.*, 2011. The NLRP3 inflammasome instigates obesity-induced inflammation and insulin resistance. *Nature Medicine* 17, 179–188.
- Wang, X., 2001. The expanding role of mitochondria in apoptosis. *Genes & Development* 15, 2922–2933.
- Wang, Y., Ludwig, J., Schuberth, C., *et al.*, 2010. Structural and functional insights into 5[prime]-ppp RNA pattern recognition by the innate immune receptor RIG-I. *Nature Structural & Molecular Biology* 17, 781–787.
- Weichhart, T., Costantino, G., Poglitsch, M., *et al.*, 2008. The TSC-mTOR signaling pathway regulates the innate inflammatory response. *Immunity* 29, 565–577.
- West, A.P., Brodsky, I.E., Rahner, C., *et al.*, 2010. TLR signalling augments macrophage bactericidal activity through mitochondrial ROS. *Nature* 472, 476–480.
- Wickliffe, K.E., Leppla, S.H., Moayeri, M., 2008. Anthrax lethal toxin-induced inflammasome formation and caspase-1 activation are late events dependent on ion fluxes and the proteasome. *Cellular Microbiology* 10, 332–343.
- Wies, E., Wang, M.K., Maharaj, N.P., *et al.*, 2013. Dephosphorylation of the RNA sensors RIG-I and MDA5 by the phosphatase PP1 is essential for innate immune signaling. *Immunity* 38, 437–449.
- Windheim, M., Lang, C., Pegg, M., Plater, L.A., Cohen, P., 2007. Molecular mechanisms involved in the regulation of cytokine production by muramyl dipeptide. *Biochemical Journal* 404, 179–190.
- Xavier, R.J., Rioux, J.D., 2008. Genome-wide association studies: A new window into immune-mediated diseases. *Nature Reviews Immunology* 8, 631–643.
- Yang, J., Zhao, Y., Shi, J., Shao, F., 2013. Human NAIIP and mouse NAIIP1 recognize bacterial type III secretion needle protein for inflammasome activation. *Proceedings of the National Academy of Sciences of the United States of America* 110, 14408–14413.
- Yarovinsky, F., 2014. Innate immunity to *Toxoplasma gondii* infection. *Nature Reviews Immunology* 14, 109–121.
- Yin, Q., Tian, Y., Kabaleeswaran, V., *et al.*, 2012. Cyclic di-GMP sensing via the innate immune signaling protein STING. *Molecular Cell* 46, 735–745.
- Yoneyama, M., Kikuchi, M., Matsumoto, K., *et al.*, 2005. Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity. *Journal of Immunology* 175, 2851–2858.
- Zak, D.E., Andersen-Nissen, E., Peterson, E.R., *et al.*, 2012. Merck Ad5/HIV induces broad innate immune activation that predicts CD8⁺ T-cell responses but is attenuated by preexisting Ad5 immunity. *Proceedings of the National Academy of Sciences* 109, E3503–E3512.
- Zeng, W., Sun, L., Jiang, X., *et al.*, 2010. Reconstitution of the RIG-I pathway reveals a signaling role of unanchored polyubiquitin chains in innate immunity. *Cell* 141, 315–330.
- Zhang, X., Wu, J., Du, F., *et al.*, 2014. The cytosolic DNA sensor cGAS forms an oligomeric complex with DNA and undergoes switch-like conformational changes in the activation loop. *Cell Reports* 6, 421–430.
- Zhang, Z., Yuan, B., Bao, M., *et al.*, 2011. The helicase DDX41 senses intracellular DNA mediated by the adaptor STING in dendritic cells. *Nature Immunology* 12, 959–965.
- Zhao, Y., Yang, J., Shi, J., *et al.*, 2011. The NLRC4 inflammasome receptors for bacterial flagellin and type III secretion apparatus. *Nature* 477, 596–600.
- Zhong, B., Yang, Y., Li, S., *et al.*, 2008. The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation. *Immunity* 29, 538–550.
- Zhong, Y., Kinio, A., Saleh, M., 2013. Functions of NOD-like receptors in human diseases. *Frontiers in Immunology* 4, 333.
- Zhou, R., Tardivel, A., Thorens, B., Choi, I., Tschopp, J., 2010. Thioredoxin-interacting protein links oxidative stress to inflammasome activation. *Nature Immunology* 11, 136–140.
- Zhu, Z., Zhang, X., Wang, G., Zheng, H., 2014. The laboratory of genetics and physiology 2: Emerging insights into the controversial functions of this RIG-I-like receptor. *BioMed Research International* 2014, 960190.

Scavenger Receptors

S Gordon, Sir William Dunn School of Pathology, Oxford, UK and Stellenbosch Institute for Advanced Study, Stellenbosch, South Africa

C Neyen, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

© 2016 Elsevier Inc. All rights reserved.

Introduction

Scavenger receptors (SRs) belong to the vast family of innately encoded pattern recognition receptors – first-line defense molecules as ancient as host–microbe interactions themselves. Given their extreme structural and functional diversity, it may be useful to define what constitutes a true SR. A recent effort to standardize SR nomenclature delineates four major defining features: SRs are cell surface receptors (i.e., transmembrane molecules, as opposed to secreted or intracellular) that share mostly myeloid expression (with an extension to professional scavenging cells like scavenger endothelial cells) (Sørensen *et al.*, 2012); they typically bind multiple ligands that are mostly nonself and modified self, and functional outcomes of ligand recognition include endocytosis, phagocytosis, adhesion, and signaling (PrabhuDas *et al.*, 2014). The present article summarizes structure–function characteristics of SRs throughout evolution, focuses more specifically on innate immune mechanisms and finally links physiological functions of SRs to disease phenotypes and therapeutic advances.

SR-Specific Domains Share Ancient Origins

SRs are grouped in 10 classes based on domain architecture as depicted in Figure 1. Mammalian members are listed in Table 1 alongside related invertebrate members for classes where their existence has been described. Although most domains with SR-defining functions have arisen early during evolution they may not always occur in the same combinations as found in mammalian SRs. A direct consequence of this combinatorial domain organization is the broad binding specificity of SRs, which earned them the nickname ‘molecular flypaper’ (Krieger *et al.*, 1992). Despite the multitude of recognition domains, many receptors share the same ligands, for example, Gram-negative lipopolysaccharide (LPS) or Gram-positive lipoteichoic acid (LTA), and all bind the defining ligand modified LDL. This shared ligand specificity despite structural complexity has been proposed to stem from a similar distribution of electrostatic potential in the three-dimensional structure of many receptors (Canton *et al.*, 2013). Electrostatic interactions are destabilized and permit ligand dissociation in the lower pH environment of endolysosomes, as shown for several SRs (Nielsen *et al.*, 2013; Doi *et al.*, 1994).

The scavenger receptor cysteine rich (SRCR) domain found in class A and class I SRs is part of a larger superfamily based on a ~100 amino acid domain that includes 6 (group A) to 8 (group B) cysteines forming an immunoglobulin-like fold held together by intramolecular disulfide bridges (Hohenester *et al.*, 1999; Sarrias *et al.*, 2004). Group A SRCRs are found in the first identified macrophage SR, SR-AI, and in CD163 (Freeman *et al.*, 1990; Resnick *et al.*, 1994), whereas group B SRCRs are part of lymphocyte surface receptors CD5 and CD6

currently not classified as SRs. The SRCR is required for ligand binding in macrophage receptor with collagenous structure (MARCO) but dispensable for ligand binding in SR-AI, since the shorter isoform, SR-AII, lacks the SRCR but has identical ligand specificity (Brannstrom *et al.*, 2002). SRCRs are often found together with coiled-coils or collagen-like repeats that mediate ligand binding in SR-AI and SR-AII (Rohrer *et al.*, 1990; Kodama *et al.*, 1990), or with C-type lectin-like domains (CTLDs) that bind ligands in COLEC12/SRCL (Nakamura *et al.*, 2001). In *Drosophila*, SRCRs are mostly found in secreted multi-domain immune response proteins that contain enzymatic effector domains such as serine proteases involved in the humoral immune response, suggesting molecular coupling of pathogen detection and immune effector mechanisms (Munier *et al.*, 2004).

The CTLD is a Ca^{2+} -dependent carbohydrate recognition domain found in class A and E receptors. While CTLDs are part of the larger carbohydrate recognition domain family and found in bona fide innate immune receptors for glycans like the mannose receptor or DC-SIGN, the CTLDs in SRs do not necessarily bind glycans, but can also recognize proteins, lipids, or nucleic acids (Zelensky and Gready, 2005). CTLDs are part of surface receptors in lower organisms like nematodes and insects, but combine with functionally different domains compared to vertebrate SRs (Drickamer and Dodd, 1999; Kanost *et al.*, 2004). The LINK domain found in class H SRs shares structural homology with C-type lectins, but does not bind glycosaminoglycans (Prevo *et al.*, 2004).

Another recurrent domain is the 30–40 amino acid EGF-like fold found in class F and class H receptors. Like the SRCR, the EGF-like domain is cross-linked by disulfide bridges and occurs in a variety of cell surface and secreted molecules. Laminin, found in class H SRs, is a subtype of EGF-like domains distinguished by four instead of three disulfide bridges (Wouters *et al.*, 2005). A common feature of EGF-like and EGF-laminin domains is their occurrence in tandem or multicopy repeats and a role in extracellular matrix adhesion (see Section SRs Contribute to Cell Adhesion). In class H SRs, EGF-like domains are interspersed with another extracellular matrix (ECM) adhesion domain, the evolutionarily ancient fasciclin domain (FAS1) found both in plants and animals.

Additional domains specific to single SR classes include immunoglobulin-like domains found in class J SRs, and the lysosome-associated membrane glycoprotein (LAMP) and mucin domains found in CD68, the single class D SR. LAMPs are primarily found in lysosomal and endosomal proteins, reflected by the preferred endosomal localization of CD68.

Finally, the CD36 fold exclusive to class B SRs is found both in plasma membrane receptors and lysosomal membrane proteins (Calvo *et al.*, 1995). Its characteristic ‘hairpin’ topology with two transmembrane passes forms a hydrophobic tunnel that channels neutral lipids across

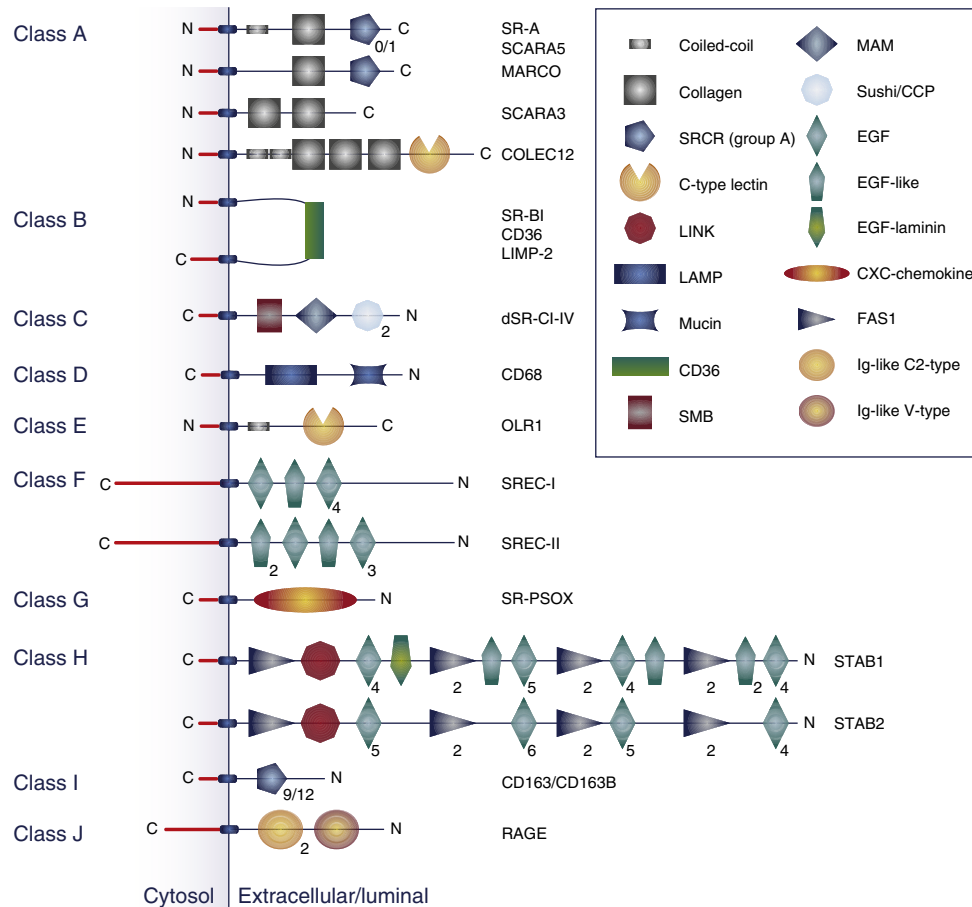


Figure 1 Human scavenger receptor classes and functional domains. Adapted from Canton, J., Neculai, D., Grinstein, S., 2013. Scavenger receptors in homeostasis and immunity. *Nature Reviews Immunology* 13, 621–634; PrabhuDas, M., Bowdish, D., Drickamer, K., *et al.*, 2014. Standardizing scavenger receptor nomenclature. *Journal of Immunology* 192, 1997–2006, based on SMART and InterPro domain representations for human members of scavenger receptor families. Cytosolic tails are represented by red bars, transmembrane domains by blue boxes. Domain repeats are indicated by subscript numbers; no number means the domain is present as single copy. CCP, complement control protein; EGF, epithelial growth factor; FAS1, fasciclin 1; Ig, immunoglobulin; LAMP, lysosome-associated membrane glycoprotein; MAM, meprin, A-5 protein, and receptor protein-tyrosine phosphatase mu; SMB, somatomedin-B; SRCR, scavenger receptor cysteine rich. See also [Table 1](#) for members of each class.

biomembranes (see Sections Class B SRs Mediate Physiological Lipid Transport and SRs as Therapeutic Targets), while its heavily glycosylated extracellular (CD36, SR-BI) or luminal (SCARB2/LIMP2) domain interacts with a variety of ligands ([Neculai *et al.*, 2013](#)).

Class C SRs have so far only been described in insects, and contain a number of domains with dimerization or adhesive properties, including the vitronectin-derived somatomedin-B (SMB) domain, a meprin, A-5 protein, and receptor protein-tyrosine phosphatase mu (MAM) domain, and Sushi/complement control protein domains (CCP). Whether invertebrate class C SRs can functionally substitute class A SRs which are conspicuously absent from invertebrate species ([Whelan *et al.*, 2012](#)), remains to be investigated.

Cell Biology of SRs: Endocytosis and Signaling

Contrary to their extremely varied ectodomains, the cytosolic tail structure of SRs, with the exception of class F, is relatively

simple. Internalization signals have been described in the cytosolic tails of most receptors ([Murphy *et al.*, 2005](#)), and modified LDL uptake was shown early on to proceed through clathrin-coated pits ([Fukuda *et al.*, 1986](#); [Figure 2](#)). Indeed, the dileucine motifs required for interaction with clathrin adaptors have subsequently been identified in a number of SRs. SR-AI is internalized via clathrin-coated pits ([Chen *et al.*, 2006](#)), and internalization rate is modulated by phosphorylation of serine residues in the cytosolic tail ([Fong and Le, 1999](#)). STAB1 is mostly intracellular, with a subpopulation cycling between the plasma membrane and early endosomes ([Prevo *et al.*, 2004](#)), and both STAB1 and STAB2 associate with the clathrin adaptor AP2 for early endosome localization ([Hansen *et al.*, 2005](#)). Further, two binding sites for the Golgi-localized, γ -adaptin ear-containing, adenosine diphosphate ribosylation factor (ARF)-binding (GGA) adaptor mediate shuttling between Golgi and lysosomes but not endocytosis of STAB1 ([Zhang *et al.*, 2009](#)). CD68 resides almost exclusively in endolysosomal compartments and undergoes rapid endocytosis based on a tyrosine motif in its cytosolic tail ([Kurushima *et al.*, 2000](#)).

Table 1 Mammalian scavenger receptor nomenclature with ancestral class members

Class	Proposed name	Current NCBI gene name	Alternative names	Ancestral proteins in this class
A	SR-A1 ^a	MSR1	SR-AI, CD204, SCARA1	Speract receptor in <i>Strongylocentrotus</i>
	SR-A3	SCARA3	MSRL1, APC7 ^b	
	SR-A4	COLEC12	SCARA4, SRCL, CL-P1	
	SR-A5	SCARA5	TESR, NET33	
	SR-A6	MARCO	SCARA2, Ly112 ^b	
B	SR-B1 ^a	SCARB1	SR-BI, CD36L1	Lipid transporters ninaD and Santa-maria in <i>Drosophila</i> Odorant receptor SNMP in <i>Drosophila</i> Phospholipid-binding protein Lmp in <i>Dictyostelium</i> dSR-CI to dSR-CIV, phagocytic receptors in <i>Drosophila</i>
	SR-B2	CD36	SCARB3, FAT, GPIV, PAS4	
	SR-B3	SCARB2	LIMP2, CD36L2, LGP85	
C	SR-C ^c			
D	SR-D1	CD68	gp110, SCARD1, LAMP4, macrophage ^b	
E	SR-E1 ^a	OLR1	LOX-1, SCARE1, CLEC8A	Eater, Nimrod, Draper phagocytic/adhesion receptors in <i>Drosophila</i>
F	SR-F1	SCARF1	SREC-I	
	SR-F2	SCARF2	SREC-II	
G	SR-F3	MEGF10	EMARDD	
	SR-G1	CXCL16	SR-PSOX	
H	SR-H1	STAB1	FEEL-1	
	SR-H2	STAB2	FEEL-2	
I	SR-I1	CD163	M130	
	SR-I2	CD163L1	M160, CD163B	
J	SR-J1 ^a	RAGE	AGER	

^aIndicates a splice variant exists which will be labeled SR-X1.1.

^bIndicates name exclusive to murine protein.

^cMembers of this class have only been described in *Drosophila*. Throughout this article, the NCBI names will be used as the nomenclature discussion is ongoing.

The class B SR SCARB2/LIMP2 lacks tyrosine motifs but is nonetheless targeted directly from Golgi to lysosomes, based on a [DE]XXXL[L] dileucine motif that interacts with the late endosome adaptor AP3 (Vega *et al.*, 1991; Janvier *et al.*, 2003).

LOX-1 undergoes clathrin-independent, dynamin-dependent endocytosis through caveolae, which is dependent on an endocytic motif and palmitoylation of its cytosolic tail (Kumano-Kuramochi *et al.*, 2013; Vohra *et al.*, 2009; Murphy *et al.*, 2008). Similarly, both SCARB1/SR-BI and CD36 are sorted into lipid rafts by a palmitoyl anchor and associate with caveolin (Babitt *et al.*, 1997), but the role of caveolin in cholesterol exchange is debated. SR-BI surface localization and function are posttranslationally controlled in a tissue-specific manner by the PZDK1 scaffold protein interacting with its cytosolic tail (Fenske *et al.*, 2009). Only the cytosolic tail splice variant SR-BII contains a dileucine motif and undergoes clathrin-mediated endocytosis (Eckhardt *et al.*, 2006). A recent effort in phosphoproteomics has identified a heteromeric protein complex associated with CD36 internalization: β 1/2 integrins and the tetraspanins CD9 and CD81 all associate with CD36 in receptor–ligand complexes and promote recruitment of FcR γ , whose ITAM motif is the actual signaling component that engages Src kinases for endocytosis (Heit *et al.*, 2013).

An interesting concept is that of ‘frustrated phagocytosis,’ which postulates that macrophage adhesion mediated by SR interactions with immobilized ligands, such as amyloid- β in plaques, for example, are in fact attempts to endocytose these normally soluble ligands. This concept was tested in SR-A,

which proved that the cytosolic tail motifs required for endocytosis are uncoupled from those required for adhesion (Kosswig *et al.*, 2003). Adhesion requires a short membrane-proximal motif and recruitment of SR-A to cholesterol rich membrane domains (Vadali and Post, 2014). Moreover, initial engagement of SR-A to substrate activates PI3K signaling which promotes a surface increase of SR-A and enhances macrophage adhesion (Cholewa *et al.*, 2010).

Adhesion via SRs induces cell spreading due to cytoskeletal rearrangements, as observed for MARCO and SR-A. SR-A adhesion requires sequential activation of the Src kinase Lyn and PI3K, which leads to paxillin phosphorylation (Nikolic *et al.*, 2007). Likewise, CD36 signals to the actin cytoskeleton and to focal adhesion components by interaction with paxillin (Stuart *et al.*, 2007). SCARF1/SREC-I induces neurite-like protrusions when overexpressed in fibroblasts, through interaction with the actin-binding protein advillin (Shibata *et al.*, 2004).

Signaling downstream of SRs is cell-type specific as exemplified by CD36 engaging different signaling adaptors in endothelial cells, macrophages, and platelets (Silverstein and Febbraio, 2009). Signaling is also context-dependent, and is fine-tuned by differential pairing of SRs with co-receptors when facing bacteria or modified self (see Section SRs Collaborate with Other PRR Families to Fine-Tune Responses). RAGE, LOX-1, and SR-PSOX are reported to trigger pro-inflammatory NF- κ B signaling, with the latter two also activating PI3K/Akt pathways. SR-A and CD36 activate the stress-related p38, JNK, and MAPK modules in response to certain ligands. Interestingly, location defines signal output for

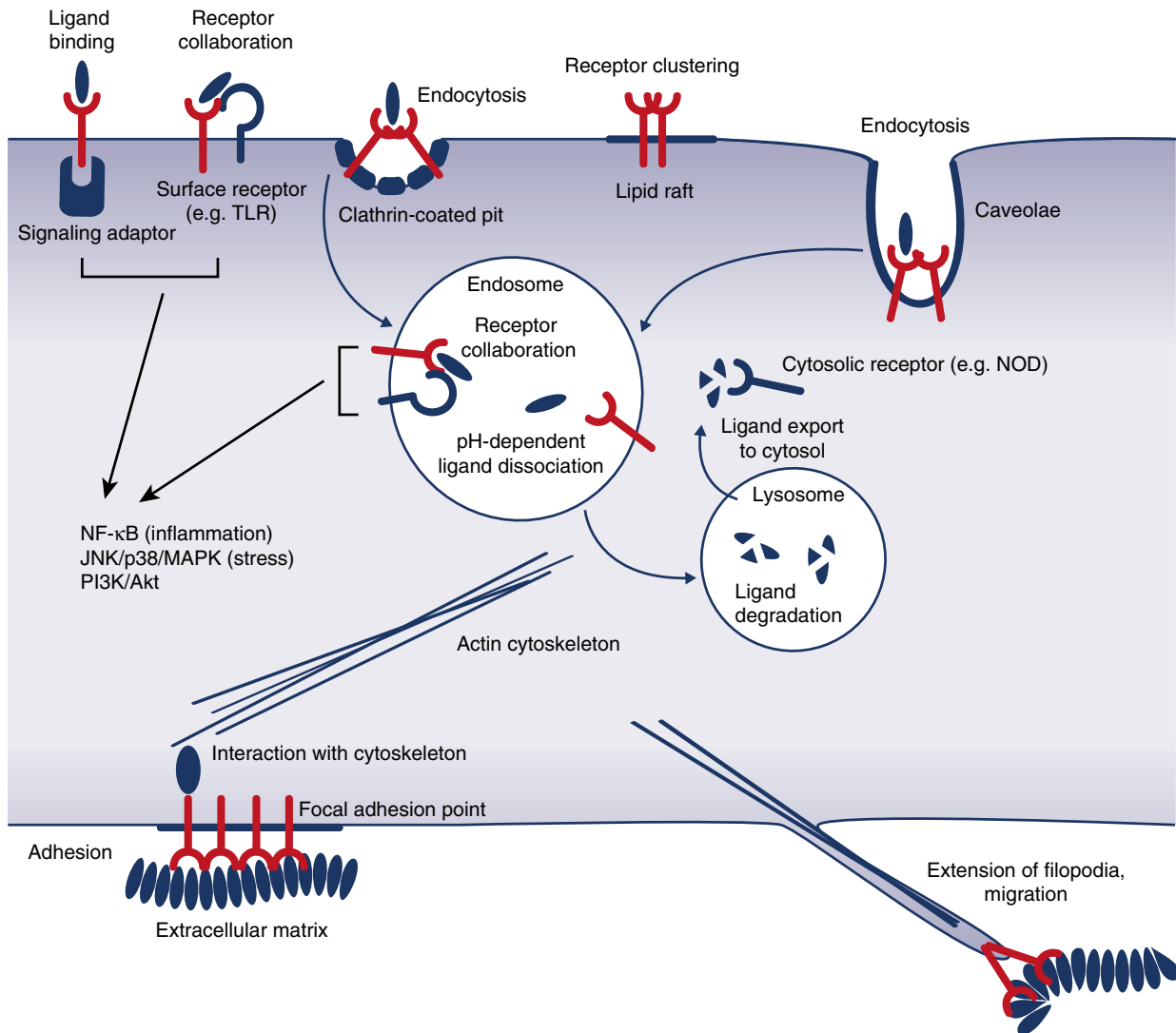


Figure 2 Cell biology of scavenger receptors. Scavenger receptors occupy a range of cellular compartments. Receptors localizing to the plasma membrane can signal upon ligand binding and pair with non-SR receptors to extend their signaling spectrum. Ligand-binding triggers endocytosis via clathrin-coated pits or caveolae. Some receptors are found in lipid rafts. Endocytosed receptors access the endosomal compartment where they can pair with endosomal Toll-like receptors (TLRs) and signal. Maturation of endosomes and fusion with lysosomes leads to acidification and pH-dependent release of ligands from receptors. Ligand degradation products can be exported from endosomes into the cytosol and activate cytosolic sensors like NOD. SRs involved in adhesion to substratum organize focal adhesion points or filopodia by interacting with the actin cytoskeleton. See also main text for details.

SR-A: clathrin-mediated endocytosis triggers ERK signaling, while caveolae-mediated endocytosis triggers p38 and JNK signaling (Zhu *et al.*, 2011). Induction of apoptosis upon ligand-binding has only been reported for SR-A and LOX-1 so far, and may depend on co-stimulation of additional pathways (Seimon *et al.*, 2006).

Transcriptional Control of SR Expression

A defining feature of SRs is their expression on myeloid cells, particularly on resident tissue macrophages, but many are expressed on additional cells, reflecting specialized functions (Figure 3). Class A SRs, for example, are enriched on alveolar

macrophages in the lung, where they contribute to innate immunity against airway pathogens (MARCO and *Streptococcus pneumoniae*, see Section SRs Collaborate with Other PRR Families to Fine-Tune Responses) or against environmental particles (SR-A and silica or asbestos). The class J SR RAGE is uniquely enriched on lung epithelial cells and is involved in lung cancer and fibrosis, but the underlying mechanisms are still largely unclear (Buckley and Ehrhardt, 2010). Class B SRs which mediate lipid transport are highly expressed in tissues of fat metabolism (CD36 in adipose tissue, SR-BI in hepatocytes, and liver Kupffer cells) or steroidogenesis (SR-BI in adrenal glands and testis) (Rigotti *et al.*, 2003; Silverstein and Febbraio, 2009). Many SRs are also particularly enriched in tissues with an immune or first-line defense function, like the

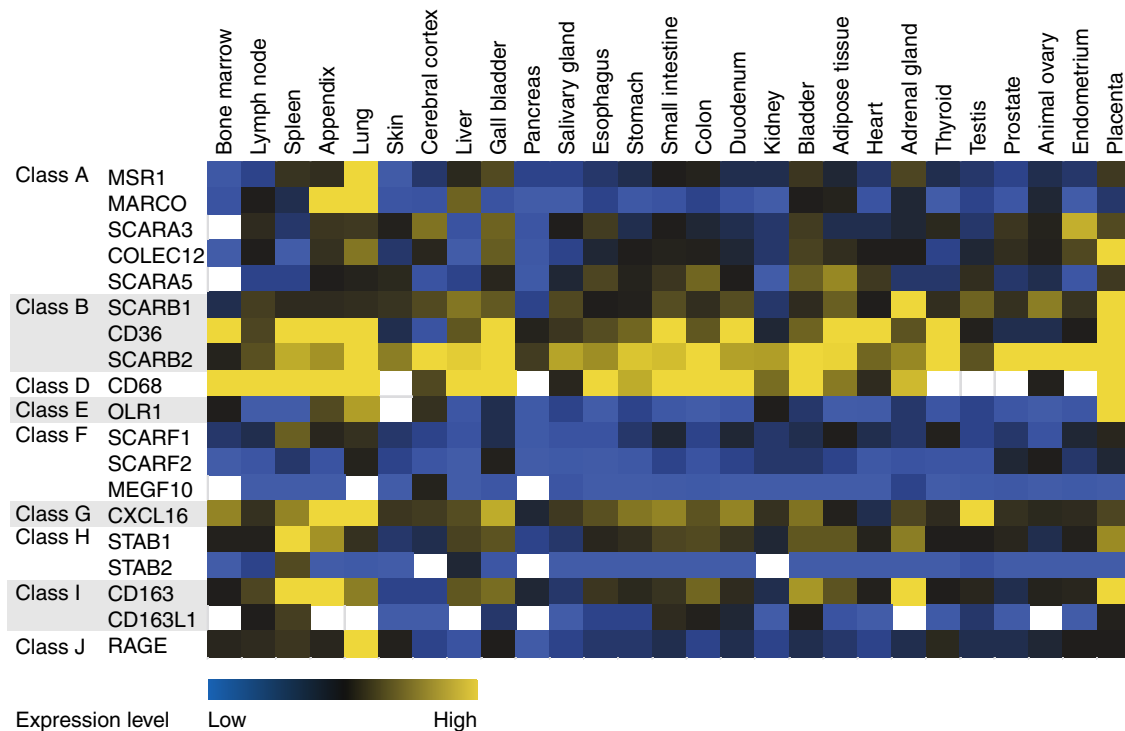


Figure 3 Baseline expression of scavenger receptors in 27 human tissues. The Expression Atlas at the European Bioinformatics Institute (EBI) was used to retrieve gene expression patterns across tissues for all SRs. Shown is a heat-map based on FPKM (fragments per kilobase of exon per million fragments mapped) values obtained by RNA sequencing of coding RNA from tissue samples of 95 human individuals representing 27 different tissues (Fagerberg *et al.*, 2014).

spleen, lung, or appendix. With few exceptions, SRs are up-regulated during monocyte to macrophage differentiation (compare bone marrow to lymph node or spleen expression).

Myeloid expression of SRs is dependent on colony-stimulating factor 1 (CSF-1, also M-CSF) which drives differentiation in the mononuclear phagocyte system (Cecchini *et al.*, 1994). Studies in CSF-1 deficient *op/op* mice (*op* stands for osteopetrotic, given that CSF-1 deficient mice lack osteoclasts and suffer from bone morphogenic defects) identified CSF-1-dependent and -independent macrophage populations, with peritoneal, splenic marginal zone and lymph node sinus macrophage populations fully dependent on CSF-1, and tissue macrophages like liver Kupffer cells, lung alveolar, and intestinal lamina propria macrophages or brain microglia largely independent of CSF-1 (Witmer-Pack *et al.*, 1993). Beyond transcription, CSF-1 can enhance function of certain SRs: CSF-1 not only increases SR-A expression at transcript level, but also stabilizes the receptor and drives its relocation from internal membranes to the cell surface, thereby promoting endocytic and adhesive functions of SR-A expressing macrophages (de Villiers *et al.*, 1994). Major transcriptional regulators of SRs in macrophages are the lipid-responsive nuclear receptors, peroxisome proliferator-activated receptors (PPARs) (Chinetti *et al.*, 2000; Ricote *et al.*, 1998; Tontonoz *et al.*, 1998). SRs and PPARs therefore constitute a double lipid sensing mechanism at the cell surface and in the nucleus that acts in concert to define macrophage differentiation. In addition, immune stimuli such as cytokines modulate SR

expression and define states of macrophage immune activation (Murphy *et al.*, 2005). SR-A in particular is a marker of alternative, T_H2 activated macrophages, which exhibit enhanced endocytic and phagocytic capacities and an anti-inflammatory profile (Gordon and Martinez, 2010).

Evolutionary Origins of SRs and Functions in Host Physiology

Historically, SRs originate from the quest for an endocytic receptor of modified LDL in mammalian macrophages (Goldstein *et al.*, 1979). Evolutionarily, however, SRs carry out a variety of other physiological functions, often related to detection, transport, or clearance of endogenous molecules. SRs are conserved across phyla and found in slime molds, nematodes, fruit flies, sea urchins, and in the primitive chordate amphioxus (Huang *et al.*, 2008; Janssen *et al.*, 2001; Messier-Solek *et al.*, 2010; Rast *et al.*, 2006). Below we list some examples of functional conservation across phyla that highlight roles in host physiology.

Class B SRs Mediate Physiological Lipid Transport

The lipid-binding capacity of class B SRs represents a fascinating example of functional conservation across evolution. In the slime mold *Dictyostelium discoideum*, lysosomal membrane

proteins of the CD36/LIMP2 family bind phosphatidylinositol diphosphates, a type of membrane phospholipid that gives functional identity to endomembranes. By bridging lysosomal compartments and cytoskeletal proteins, the Lmp proteins are thought to facilitate vesicular transport (Karakesisoglou *et al.*, 1999; Janssen *et al.*, 2001). The fruit fly *Drosophila* harbors a relatively high number of CD36-like genes compared to humans (14 in flies vs. 3 in men), which predicts physiologically important functions. Indeed, flies are cholesterol auxotrophs, meaning they rely on cholesterol from their diet. Tissue-specific expression of CD36 genes in *Drosophila* suggests roles in intestinal lipid transport, steroid hormone synthesis, lipid storage, and satiety sensing (Herboso *et al.*, 2011). In particular, the fly CD36 homologs *ninaD* and *Santamaria* were identified as specific docking sites for the uptake and transport of β -carotene from the diet (Kiefer *et al.*, 2002; Wang *et al.*, 2007a). Carotenoids, like cholesterol, are water-insoluble isoprenoid lipids transported on lipoproteins, and are essential for eye development. Fly CD36-like receptors function analogously to mammalian SR-BI, which acts as a docking site for lipoproteins and facilitates cholesterol transport in and out of cells (Rigotti *et al.*, 1997). Moreover, the fly CD36 members *croquemort*, *peste* and *SNMP* are upregulated in steroidogenic tissues during development, where they support ecdysone biogenesis (Talamillo *et al.*, 2013; Herboso *et al.*, 2011), much like SR-BI expression is regulated by hormones in mammalian adrenal glands, where it supports steroid hormone metabolism (Rigotti *et al.*, 2003). SNMPs, other fly CD36 relatives, are expressed in olfactory neurons and detect fatty acid derived pheromones that define social preferences in flies (Benton *et al.*, 2007). A similar sensory function is attributed to intestinal CD36 which detects dietary fatty acids in mammals, translating a preference for fatty foods and the feeling of satiety (Martin *et al.*, 2011).

SRs Contribute to Cell Adhesion

A highly conserved and wide-spread role among SR classes is cell–cell and cell–extracellular matrix adhesion. Many domains described above confer adhesive properties in primitive multicellular organisms: SRCR domains allow reaggregation of conspecific cells in sponges (Blumbach *et al.*, 1998), and a CD36 family member mediates sponge cell interactions with extracellular matrix. The sponge CD36 interaction is particularly intriguing as it is blocked by a protein domain found in thrombospondin, a ligand for mammalian CD36 (Muller *et al.*, 2004). CD36 also mediates fusion between macrophages during giant cell formation, an immune process accompanying granulomatous infections. Here, CD36 is required on one fusion partner for recognition of exposed phosphatidylserine (PS) on the other (Helming *et al.*, 2009). Mammalian SR-A mediates macrophage adhesion to a variety of substrata, including extracellular matrix proteoglycans, apolipoproteins occurring in atherosclerotic lesions, and amyloid deposits in Alzheimer's disease affected brains (Santiago-Garcia *et al.*, 2003; Neyen *et al.*, 2013a; Alarcon *et al.*, 2005).

Our focus below is on EGF-like repeat containing SRs, which present the best documented examples of cross-

evolutionary functional conservation in cell–cell and cell–matrix adhesion.

Class F and H SRs comprise multi-EGF-like repeat receptors, interspersed with EGF-like laminin and fasciclin (FAS1) domains. FAS1 domains appear before the plant–animal split and their adhesive role is exemplified beautifully in the flagellate green alga *Volvox*, which forms spherical colonies of adherent cells. Here, FAS1 domains are part of the algal cell adhesion molecule that tethers together embryonic algal cells (Huber and Sumper, 1994). FAS1 domains are exclusive to mammalian class H SRs (STAB1/FEEL1 and 2), which also play roles in cell–cell interaction and more specifically in vascular epithelial cell adhesion during angiogenesis (Adachi and Tsujimoto, 2002). Invertebrates seem to lack homologs of class H SRs but homologs of class F SRs are well-described in *Drosophila*. They encompass Nimrods and Eater, mainly known for their phagocytic capacity in bacterial clearance (see Section SRs in Innate Immunity), and Draper, involved in apoptotic cell clearance (see Section SRs in Corpse Clearance; Stuart and Ezekowitz, 2008). By domain homology, Nimrods and Draper closely resemble class F SRs based on relatively few extracellular EGF-like domains (1–15) and a longer cytosolic tail that may favor signaling interactions. Eater in particular, with its multiple EGF repeats and short cytoplasmic tail strongly resembles MEGF family members and less class F SRs (Somogyi *et al.*, 2008). Tellingly, *Eater* mutants show a complete deficiency of hemocyte homing to hematopoietic stem cell niches, highlighting a specific role in cell–cell or cell–matrix interaction (Bretscher *et al.*, 2015). Comparable roles in organizing immune tissues are ascribed to non-EGF mammalian SR-AI and MARCO, whose presence is required to maintain the microarchitecture of the spleen and lymph nodes (Chen *et al.*, 2005; Hughes *et al.*, 1995).

Interestingly, detecting self (in the form of a SR family member) can trigger aggregation or repulsion: heterotypic interactions between class F SRs SREC-I and SREC-II contribute to aggregation of fibroblasts (Ishii *et al.*, 2002), while homotypic interactions between the ectodomains of class F receptor MEGF10 or non-SR MEGF11/Jedi ensure correct spacing and subtype distribution of neurons during neuronal patterning in the eye (Kay *et al.*, 2012).

Multiple SRs Mediate Clearance of Normal Self

All SRs are defined by their ability to recognize and endocytose modified lipoproteins, a feature exacerbated by high-fat, high-sugar diets that bring about the inflammatory and oxidative conditions generating these ligands (see Section SRs in Inflammatory Diseases: Innate Immunity to Modified Self). However, in a healthy, normally fed organism, SRs contribute to clearance of a range of physiological ligands.

Class A SCARA5 contributes to iron metabolism by mediating transferrin-independent ferritin uptake (Li *et al.*, 2009). The structurally unrelated class I receptor CD163 also participates in iron metabolism by removing hemoglobin/haptoglobin complexes, thereby facilitating the breakdown of heme and recycling of iron (Kristiansen *et al.*, 2001). In addition to recycling iron, clearance of hemoglobin prevents oxidative damage in the vasculature. Another damage-preventing

clearance mechanism is provided by SR-AI, which binds to and endocytoses calciprotein particles, assemblies of fetuin-A protein and calcium-containing mineral debris that threaten to calcify the vasculature (Herrmann *et al.*, 2012). Class H SRs STAB1 and 2 bear homology to hyaluronan receptors and function as clearance receptors for circulating glycosaminoglycans in vertebrates (Harris and Weigel, 2008).

Finally, several SRs primarily reside and function within endolysosomes. An example is the class B SR SCARB2/LIMP2, which functions as a lysosomal receptor for the lysosomal enzyme β -glucocerebrosidase (Reczek *et al.*, 2007). Deficiency of lysosomal enzymes causes lysosomal storage disorders, for example, Gaucher's type diseases where the efficient cellular breakdown of waste products within lysosomes is impaired (see Section Rare Monogenic Diseases Attributed to SRs).

SRs in Corpse Clearance

SRs, including members from classes A, B, D, E, and F, feature prominently in the removal of apoptotic self during development or tissue injury (Fadok *et al.*, 2001). Alongside other recognition signals, apoptotic cells flip PS from inner to outer membrane leaflet, a process dependent on ABC transporters and other phospholipid scramblases (Ravichandran and Lorenz, 2007). PS and other anionic phospholipids are the most frequently recognized apoptotic cell ligands for SRs, but SRs can mediate both PS-dependent and PS-independent apoptotic cell clearance. Interestingly, PS exposure on viable cells does not constitute an 'eat-me' signal for macrophages, supporting the notion that additional apoptotic signatures are required in PS-dependent cell clearance (Segawa *et al.*, 2011). Apoptotic cell recognition processes are conserved throughout evolution, reflected by usage of the class B SR croquemort (crq) in *Drosophila* (Franc *et al.*, 1999) and SR-BI and CD36 in mammals (Rigotti *et al.*, 1995). An additional, vertebrate-specific corpse clearance system is based on the TAM family of receptor-tyrosine kinases (RTK) (Tyro3, Axl, and Mer) that act alongside SRs and integrins in removing apoptotic cells. TAMs also recognize PS via bridging molecules Gas6 and ProS, which activates RTK signaling. Efficient clearance of apoptotic cells is thought to rely on coupling of RTK signaling via TAMs and phagocytosis via non-signaling SRs, as exemplified by the Mer/SR-A couple (Lemke and Burstyn-Cohen, 2010).

Invertebrate homologs of class F SRs, CED-1 in *Caenorhabditis elegans* and Draper in *Drosophila melanogaster*, are phagocytic receptors that clear apoptotic neurons and guide axons during development (Zhou *et al.*, 2001; Manaka *et al.*, 2004). This role is conserved in mammalian class F receptor MEGF10 and non-SR MEGF12/PEAR1 in phagocytic glial cells (Wu *et al.*, 2009). Microglia also use the non-SR complement receptor CR3 to prune inactive synaptic connections during postnatal neuronal development (Schafer *et al.*, 2012).

Oxidative modifications on circulating lipoproteins bear resemblance to apoptotic cell epitopes, which is why SRs involved in apoptotic cell clearance are usually also contributing to oxLDL uptake in macrophages, causing atherosclerosis (see Section SRs in Inflammatory Diseases: Innate Immunity to Modified Self).

SRs in Innate Immunity

Beyond the defined physiological roles delineated above, all SRs are bona fide pattern recognition receptors (PRRs) and contribute to host defense by recognizing pathogen-associated molecular patterns (PAMPs) (Krieger, 1997). SR-mediated non-opsonic phagocytosis has been shown for hemocyte SRs in invertebrates, and includes phagocytic removal of Gram-positive and Gram-negative bacteria by *Drosophila* class C and H SRs (Kocks *et al.*, 2005; Pearson *et al.*, 1995; Stuart and Ezekowitz, 2008; R  met *et al.*, 2001). Microbial ligands for SRs have been extensively reviewed before (Areschoug and Gordon, 2008, 2009; Mukhopadhyay *et al.*, 2009; Pl  ddemann *et al.*, 2006), and databases exist to match receptors with their microbial ligands (Lata and Raghava, 2008). We chose to highlight some interesting functional aspects of SR-microbe interaction that pin-point contributions of SRs in infectious disease.

SRs Collaborate with Other PRR Families to Fine-Tune Responses

As described above, SRs have generally short cytosolic tails and lack a comprehensive signaling adaptor arsenal. Instead, SRs frequently associate with surface and intracellular PRRs to detect microbes and activate signaling (Underhill and Ozinsky, 2002; Figure 2). Most receptor interactions involve Toll-like receptors (TLRs) and occasionally NOD-like receptors (NLRs), with SRs providing the phagocytic and TLRs/NLRs the sensing/signaling part (Mukhopadhyay *et al.*, 2004).

MARCO contributes to the detection and pro-inflammatory response against mycobacterial trehalose dimycolate by tethering this glycolipid to receptor complexes composed of MARCO, TLR2, and CD14 (Bowdish *et al.*, 2009). In addition, MARCO is required for efficient immune responses against *S. pneumonia* infection in the nasopharynx: by mediating phagocytosis of *S. pneumonia*, MARCO delivers ligands for TLR2 and NOD2 that activate cytokine signaling (Dorrington *et al.*, 2013). Similarly, CD36 collaborates with TLRs in ligand-specific signaling assemblies. Macrophages lacking CD36 are unable to activate antibacterial TLR2/6 signaling in response to diacylated lipopeptides and LTA, revealing a crucial helper function of CD36 in TLR signaling to and clearance of Gram-positive bacteria (Hoebe *et al.*, 2005; Stuart *et al.*, 2005). The Gram-negative bacterium *Neisseria meningitidis* marks an exception: here SR-A-mediated phagocytosis is independent of LPS and uncoupled from TLR4/LPS induced cytokine production (Peiser *et al.*, 2002). Clearance of *Neisseria* in a mouse model of meningococcal septicemia highlights the exclusive role of SR-A in clearing the pathogen and protecting the host from endotoxic shock (Pl  ddemann *et al.*, 2009). Mechanistically, this is thought to be linked to trafficking of TLR ligands by SRs and differential activation of TLR signaling from different cellular compartments. A comprehensive combinatorial study uses mouse mutants of class A SRs SR-AI and MARCO, surface TLR4 and endosomal TLR3, and cytosolic NOD2 and NLRP3 to dissect individual contributions of each receptor. The authors show that class A SRs prevent surface activation of TLR4 signaling by rapid endocytosis of the shared

ligand LPS, but enhance activation of intracellular receptors by increasing ligand availability within the cell (Mukhopadhyay *et al.*, 2011).

Abundant evidence exists on modulation of LPS responses by SRs, which are globally protective against endotoxic shock. A shared mechanism is likely the highly efficient binding and clearance of LPS from the bloodstream (Hampton *et al.*, 1991), although this is mainly due to hepatic SR expression. SR-AI was initially shown to protect against sepsis due to its up-regulation on activated macrophages and efficient clearance of LPS and correlated reduced tumor necrosis factor- α (TNF- α) production (Haworth *et al.*, 1997). A second mode of action is modulation of downstream pro-inflammatory signaling pathways. Quantitative trait linkage (QTL) analysis in laboratory mouse strains has identified the SR-AI locus as relevant for protective, anti-inflammatory IL-10 production after LPS injection (Fulton *et al.*, 2006). Moreover, in endoplasmic reticulum (ER)-stressed macrophages combined activation of SR-AI and TLR4 by LPS leads to pro-apoptotic rather than pro-survival signaling, eliminating a source of pro-inflammatory cytokines (Seimon *et al.*, 2006). Likewise, macrophage expression of SR-BI modulates systemic NF- κ B and stress pathway signaling in response to Gram-negative LPS *in vivo*, a function that also extends to Gram-positive LTA in cultured cells (Cai *et al.*, 2012).

SRs as Microbial Entry Points

Pathogens have coevolved with host surface receptors and abuse their select specificity to gain access to cells. Prominent examples are viruses that attach to surface receptors and use the endocytic machinery to enter and replicate within defined cell types. Class B family members in particular are targeted by virus particles: Hepatitis C virus (HCV) uses SR-BI in concert with CD81 as an entry point into hepatocytes, through specific binding of the viral envelope glycoprotein E2 to SR-BI (Scarselli *et al.*, 2002). Another class B family member, SCARB2/LIMP2, is an entry point for enterovirus and coxsackievirus that cause hand, foot, and mouth disease (Yamayoshi *et al.*, 2009). Beyond virus entry, class B SRs contribute to HCV infection by promoting assembly of lipoproteins that are hijacked by viral particles for dissemination (Burlone and Budkowska, 2009). In a process termed 'apoptotic mimicry' first described for Vaccinia virus (Mercer and Helenius, 2008), a variety of viral families including filovirus, flavivirus, New World arena virus, and alphavirus families exploit the hard-wired mechanisms of PS-dependent apoptotic cell clearance. PS displayed on viral membranes is recognized by PS receptors including TIMs (T-cell Immunoglobulin and Mucin-domain containing proteins) and potentially other PS receptors such as SRs expressed on macrophages, providing a broad nonspecific viral entry mechanism (Jemielity *et al.*, 2013).

SRs also promote entry of bacteria and parasites. In flies, the CD36 family member peste mediates uptake of *Mycobacterium fortuitum*; and in mice, CD36 is required for intracellular survival of *Mycobacterium tuberculosis* (Hawkes *et al.*, 2010; Philips *et al.*, 2005), suggesting ancient coadaptation between mycobacteria and their hosts. The malaria parasite *Plasmodium falciparum* interacts with class B SRs at multiple

stages. SR-BI, along with the tetraspannin CD81, is used by the motile sporozoites to gain entry into hepatocytes and also to provide lipids that support the transformation from sporozoites to the infectious exoerythrocytic form (Rodrigues *et al.*, 2008; Yalaoui *et al.*, 2008). CD36 contributes to the adhesion of Plasmodium-infected red blood cells to vascular epithelium, which allows the parasite to proliferate in the bloodstream and avoid phagocytosis by macrophages. However, the roles of CD36 in *Plasmodium* infection are pleiotropic and include noninflammatory phagocytosis of infected erythrocytes by macrophages, a host-protective feature that has proven a worthwhile therapeutic target (Cabrera *et al.*, 2014; see Section SRs in Metabolic Disease Association Studies and SRs as Therapeutic Targets).

SRs in Inflammatory Diseases: Innate Immunity to Modified Self

It may be puzzling that the host innate immune system should turn on itself. On closer inspection though, the 'self' ligands recognized by SRs are remarkably similar to PAMPs and DAMPs (pattern and danger associated molecular patterns, respectively). Indeed, oxidized lipid epitopes, apoptotic cell epitopes, and certain bacterial lipids are structurally equivalent, and form the basis for the defining feature of SRs – recognition of modified lipoproteins (Binder *et al.*, 2002). Inflammation and oxidation are tightly linked through the release of reactive oxygen species triggered by pro-inflammatory signaling. As a result, chronic inflammation-dependent generation of SR ligands drives disease progression in conditions such as atherosclerosis or Alzheimer's.

All SRs are defined as endocytic receptors for oxidized LDL, and knock-out studies in mice have collectively shown their contributions to cardiovascular disease. Because SRs are not feedback-controlled by intracellular lipid levels as is the case for the LDL receptor, SR-mediated lipid clearance leads to lipid overload in macrophages. Free cholesterol accumulates in the ER and causes macrophage apoptosis, in a process involving ER stress sensing, p38 and JNK signaling, and engagement of SR-AI (DeVries-Seimon *et al.*, 2005). Additional SR-dependent pro-atherogenic mechanisms include pro-inflammatory signaling, angiogenesis, macrophage adhesion to atherosclerotic plaques, and modulation of insulin metabolism, best exemplified in the multitasking SR CD36 (Silverstein and Febbraio, 2009).

Early knock-out studies in mice implicated SR-AI, CD36, and LOX-1 in the development of atherosclerosis, mainly because of their contribution to cholesterol loading (Febbraio *et al.*, 2000; Mehta *et al.*, 2007; Suzuki *et al.*, 1997). However, SR functions in inflammatory/metabolic diseases are more complex than simple cholesterol loading, precluding a clear-cut verdict (Moore and Freeman, 2006).

In addition to lipid overload, high-fat diets also drive insulin resistance and concomitant rises in circulating blood sugar, which causes nonenzymatic glycation of proteins in circulation. Tissue accumulation of these advanced glycation end-products (AGEs) causes oxidative stress and pro-inflammatory signaling (Yan *et al.*, 1994). AGE removal and renal clearance are essential for tissue homeostasis, and are

mediated primarily by class J SRs, termed RAGE (receptors for AGE), seconded by SR-AI, CD36, LOX-1, and STAB1/FEEL-1 and 2 (Horiuchi *et al.*, 2003). Binding to clearance receptors is a double-edged sword, as RAGE activates pro-inflammatory signaling and leads to more ligand generation and exacerbation of the disease (Schmidt *et al.*, 2001).

Like atherosclerosis, neurodegenerative diseases progress through a combination of chronic inflammation, ligand generation, and SR-mediated clearance and immune signaling. CD36 expressed on microglia mediates adhesion to β -amyloid deposits, which triggers pro-inflammatory cytokine and reactive oxygen species release, which further exacerbates neurodegeneration (El Khoury *et al.*, 1996, 2003). CD36 acts together in a receptor complex with integrins which enhances microglial adhesion to amyloid fibrils (Bamberger *et al.*, 2003). Ideally, microglia clear accumulating amyloid before plaques assemble, but the clearance receptors, including SR-AI, CD36, and RAGE, decline with age and exposure to a pro-inflammatory environment (Hickman *et al.*, 2008).

Receptor collaboration also underlies this sterile inflammation, as exemplified by CD36 cooperation with TLRs in response to oxLDL and amyloid- β . CD36 is required to endocytose ligands and assemble endosomal receptor-ligand complexes containing TLR4/TLR6 heterodimers (Stewart *et al.*, 2010). In addition to activating endosomal TLRs, CD36 (and to a lesser extent, SR-A, which binds the same ligands) promotes ligand aggregation in endosomes, and the resulting cholesterol crystals or amyloid fibrils destabilize lysosomal membranes and trigger NLRP3 inflammasome activation in the cytosol (Sheedy *et al.*, 2013).

SRs as Causes and Therapeutic Targets in Disease

SRs in Metabolic Disease Association Studies

Given their prominent role in oxidized lipoprotein uptake and macrophage foam cell formation, SRs are candidate risk genes for cardiovascular disease and metabolic syndrome. Redundancy in oxLDL uptake mechanisms however means that single SR loci have so far failed to reach significance in genome-wide association studies.

SR-BI is involved in reverse cholesterol transport, consisting in the removal of cholesterol from foamy macrophages in peripheral arteries, loading onto circulating lipoproteins and reuptake in the liver with subsequent excretion through the bile (Tall, 1998). Although SR-BI promotes cholesterol efflux from macrophages *in vitro*, the main efflux receptors on vascular macrophages *in vivo* are ABC transporters ABCA1 and ABCG1 (Wang *et al.*, 2007b). SR-BI is athero-protective through its contribution in the liver, where it helps to absorb cholesterol from the circulation (Varban *et al.*, 1998). Accordingly, a host of studies, both large-scale and targeted to families with altered lipid profiles, have found polymorphisms in SR-BI to be associated with variations in plasma high-density lipoprotein (HDL) cholesterol levels (Osgood *et al.*, 2003; Tai *et al.*, 2003; Teslovich *et al.*, 2010; McCarthy *et al.*, 2003; Roberts *et al.*, 2007). A SNP (rs10846744) in SR-BI is predictive of cardiovascular events across ethnicities (Naj *et al.*, 2010; Manichaikul *et al.*, 2012). Additional SNPs correlate well with the

pleiotropic effects of SR-BI in steroid hormone metabolism, fertility, and platelet function (Chadwick and Sahoo, 2013).

CD36 deficiency, also called Platelet glycoprotein IV deficiency as CD36 was originally identified as a platelet surface molecule, frequently affects Asian and African populations, and causes cardiomyopathies (Hirano *et al.*, 2003). As CD36 is a transporter for long-chain fatty acids, which are a major energy source for the heart muscle, patients that lack CD36 develop heart muscle failure. Additional abnormalities in fatty acid metabolism cause defects in lipid distribution and insulin resistance, commonly termed 'metabolic syndrome.' Unsurprisingly, CD36 polymorphisms therefore also associate with type 2 diabetes in population studies (Banerjee *et al.*, 2010; Gautam *et al.*, 2013). The geographical prevalence of CD36 deficiency may be linked to its role in immunity, in particular, interactions with *M. tuberculosis* and *Plasmodium*-infected erythrocytes. Tuberculosis and malaria affect geographical areas where CD36 deficiency is prevalent, but whether mutations are under selection by infectious diseases is so far unclear (Fry *et al.*, 2009; Aitman *et al.*, 2000).

SRs in Cancer

Germline mutations and variants in SR-AI are associated with increased risk for prostate cancer (Xu *et al.*, 2002), and for Barrett's esophagus, an acid reflux condition leading to esophageal cancer (Orloff *et al.*, 2011). While it is unclear by which mechanism SR-A predisposes to development of these specific types of cancer, a relevant starting point may be the observation that SR-A is upregulated on tumor-associated macrophages (TAMs), and that SR-A on TAMs is a pro-oncogenic driver in tumor cell-macrophage cross-talk (Hagemann *et al.*, 2006). This has been successfully exploited to target a toxin to SR-A expressing TAMs and prevent tumor progression in a mouse model of ovarian cancer. Similarly, blocking SR-A by small peptide mimetics of SR-A ligands in the same model seems to interfere with pro-oncogenic TAM-tumor cell cross-talk and halt metastasis (Neyen *et al.*, 2013b). The pro-oncogenic role of SR-A might however be tumor-type or environment specific as studies in lung tumorigenesis report an association between SNPs lowering SR-A expression and bad prognosis (Ben *et al.*, 2012). An interesting link to investigate would be the prominent contribution of class A SRs to detection of environmental irritants in the lung.

For class A member SCARA5 pro-oncogenic mechanisms are clearer: identified in a human genome-wide screen for candidate tumor suppressor genes, SCARA5 is involved in maintaining adhesion signaling in normal cells. SCARA5 down-regulation in tumor cells licenses focal adhesion kinase signaling and promotes metastasis (Huang *et al.*, 2010).

A comprehensive survey of SR expression levels in lymph node sinusoidal cells shows distinct but overlapping profiles of SRs on epithelial cells and macrophages, suggesting a variety of retention mechanisms for metastatic tumor cells within lymph nodes (Martens *et al.*, 2006). Metastasis relies in part on the ability of tumor cells to adhere to extracellular matrix and leave the circulation. Hyaluronic acid prevents tumor metastasis, but is efficiently cleared from the circulation by STAB1 and 2. Two promising antitumor strategies target STAB1

interaction with hyaluronan. Antibody-mediated blocking of STAB1 resulted in elevated hyaluronan in circulation and decreased metastasis in a murine model of metastatic melanoma (Hirose *et al.*, 2012). In a murine prostate cancer model, the same approach decreased interaction between sinusoidal epithelium and tumor cells mediated by their pericellular hyaluronan matrix (Simpson *et al.*, 2012).

Taken together, these examples suggest that the main role of SRs in tumorigenesis is probably to be found in oncogenic misappropriation of physiological roles in cell–cell and cell–matrix adhesion.

Rare Monogenic Diseases Attributed to SRs

Since many SRs are functionally redundant, it may be unsurprising that lack-of-function mutations in SR genes rarely give striking phenotypes. However, some rare deficiencies give insight into gene-exclusive roles for select SRs.

Rare mutations in the class B member SCARB2/LIMP2 cause myoclonic epilepsy and renal failure at an early adult age. The underlying pathology derives from incorrect targeting of the lysosomal enzyme β -glucocerebrosidase, which is chaperoned by SCARB2 in normal cells. In the absence of functional receptor, unprocessed enzyme substrate accumulates in lysosomes (see Section Multiple SRs Mediate Clearance of Normal Self; Dardis *et al.*, 2009).

Lack of class F SCARF2/SREC-II causes a rare autosomal recessive disorder called Van Den Ende–Gupta syndrome, which is characterized by skeletal abnormalities that do not affect intelligence and are not life-threatening (Anastasio *et al.*, 2010). The fact that this deficiency is not compensated by other SRs points toward a critical role of SCARF2 in epidermal differentiation and bone morphogenesis during development.

SRs as Therapeutic Targets

Beyond the oncostatic approaches above, SR-based therapies are generally harnessing the myeloid-restricted expression of SRs to target macrophages in disease contexts. In particular, SRs are highly expressed on macrophages and smooth muscle cells in atherosclerotic plaques but not in healthy arteries, and are highly efficient in internalizing ligands. These properties can be exploited to target ligand-coated nanoparticles to macrophage-rich atherosclerotic lesions, to improve their contrast on magnetic resonance imaging (MRI) scans and allow better predictive assessment of plaque stability (Lipinski *et al.*, 2004). One approach successfully applied in atherosclerotic mice used iron oxide nanoparticles coated with ligands for SR-AI, for example, dextran sulfate or an SR-AI reactive peptide (Segers *et al.*, 2013; Tu *et al.*, 2011). Another approach enclosed the contrast agent gadolinium in phospholipid micelles coated with anti-SR-A antibodies which directed uptake into macrophages and improved plaque contrast in MRI (Amirbekian *et al.*, 2007). The advantage of lipid micelles or liposomes is their versatility in delivering cargo. In addition to contrast agents for MRI, a cargo of choice are chemotherapeutic agents targeted to TAMs (Rensen *et al.*, 2006; Etzerodt *et al.*, 2012). An exciting new avenue has opened up with the discovery of naturally occurring transport

of microRNAs on circulating lipid-based carriers (Vickers and Remaley, 2012). In particular, HDL carrying miRNAs is delivered to SR-BI expressing cells and effects gene expression changes in target cells (Tabet *et al.*, 2014; Vickers *et al.*, 2011).

Another class of therapeutic molecules targeting SRs are BLTs (Blocker of Lipid Transport), small molecule inhibitors which block selective cholesterol uptake by occupying the lipid channel in class B receptors (Niemand *et al.*, 2008, 2013). Beyond their obvious applications in modulating cholesterol transport, BLTs have been applied to other SR-B-dependent etiologies. BLT-1, 2, and 4 can block *Plasmodium* sporozoite infection of human hepatocytes in culture, by interfering with parasite docking to SR-BI (Rodrigues *et al.*, 2008). Similarly, HCV uses SR-BI as a co-receptor to infect hepatocytes, and BLT-2 and 4 are efficient in blocking HDL-aided HCV entry into hepatocytes (Bartosch *et al.*, 2005; Voisset *et al.*, 2005).

Finally, class A and B SRs are able to bind to small apolipoproteins A and E which share an amphipathic α -helical structure and perform athero-protective functions (Williams *et al.*, 2000; Neyen *et al.*, 2009). Small peptides that mimic this amphipathic helical structure are versatile therapeutics with a broad range of applications (Navab *et al.*, 2006), some of which specifically target SRs. The apo A-I mimetic D4F blocks SR-A function in tumor cell–macrophage cross-talk and prevents metastasis of ovarian and pancreatic cancer cells in a murine model (Neyen *et al.*, 2013b). Like BLTs, apo E peptide mimetics block access of HCV to SR-BI (Liu *et al.*, 2012), and a similarly structured peptide blocks LPS binding to and internalization by SR-BI, preventing downstream pro-inflammatory signaling (Bocharov *et al.*, 2004). Administration of apoAI or apoE mimetic peptides to murine models of neurodegenerative disease reduces brain inflammation and microglial accumulation (Handattu *et al.*, 2009, 2013). Although a direct action of peptides on microglial adhesion through SRs is not proven, it remains a tempting mechanistic explanation.

In summary, understanding the multifaceted biology of SRs in homeostasis and disease is expected to yield future advances in therapy, particularly in macrophage-driven inflammatory diseases.

See also: Cellular Immunology: Innate Immunity: Dendritic Cells

References

- Adachi, H., Tsujimoto, M., 2002. FEEL-1, a novel scavenger receptor with *in vitro* bacteria-binding and angiogenesis-modulating activities. *Journal of Biological Chemistry* 277, 34264–34270.
- Aitman, T.J., Cooper, L.D., Norsworthy, P.J., *et al.*, 2000. Malaria susceptibility and CD36 mutation. *Nature* 405, 1015–1016.
- Alarcon, R., Fuenzalida, C., Santibanez, M., Von Bernhardi, R., 2005. Expression of scavenger receptors in glial cells. Comparing the adhesion of astrocytes and microglia from neonatal rats to surface-bound beta-amyloid. *Journal of Biological Chemistry* 280, 30406–30415.
- Amirbekian, V., Lipinski, M.J., Briley-Saebo, K.C., *et al.*, 2007. Detecting and assessing macrophages *in vivo* to evaluate atherosclerosis noninvasively using molecular MRI. *Proceedings of the National Academy of Sciences of the United States of America* 104, 961–966.

- Anastasio, N., Ben-Omran, T., Teebi, A., *et al.*, 2010. Mutations in SCARF2 are responsible for Van Den Ende-Gupta syndrome. *American Journal of Human Genetics* 87, 553–559.
- Areschoug, T., Gordon, S., 2008. Pattern recognition receptors and their role in innate immunity: Focus on microbial protein ligands. In: Egesten, A., Schmidt, A., Herwald, H. (Eds.), *Trends in Innate Immunity. Contributions to Microbiology*, vol. 15. Basel: Karger, pp. 45–60.
- Areschoug, T., Gordon, S., 2009. Scavenger receptors: Role in innate immunity and microbial pathogenesis. *Cellular Microbiology* 11, 1160–1169.
- Babitt, J., Trigatti, B., Rigotti, A., *et al.*, 1997. Murine SR-BI, a high density lipoprotein receptor that mediates selective lipid uptake, is N-glycosylated and fatty acylated and colocalizes with plasma membrane caveolae. *Journal of Biological Chemistry* 272, 13242–13249.
- Bamberger, M.E., Harris, M.E., McDonald, D.R., Husemann, J., Landreth, G.E., 2003. A cell surface receptor complex for fibrillar β -amyloid mediates microglial activation. *Journal of Neuroscience* 23, 2665–2674.
- Banerjee, M., Gautam, S., Saxena, M., Bid, H.K., Agrawal, C.G., 2010. Association of CD36 gene variants rs1761667 (G>A) and rs1527483 (C>T) with Type 2 diabetes in North Indian population. *International Journal of Diabetes Mellitus* 2, 179–183.
- Bartosch, B., Verney, G., Dreux, M., *et al.*, 2005. An interplay between hypervariable region 1 of the hepatitis C virus E2 glycoprotein, the scavenger receptor BI, and high-density lipoprotein promotes both enhancement of infection and protection against neutralizing antibodies. *Journal of Virology* 79, 8217–8229.
- Ben, J., Jin, G., Zhang, Y., *et al.*, 2012. Class A scavenger receptor deficiency exacerbates lung tumorigenesis by cultivating a procarcinogenic microenvironment in humans and mice. *American Journal of Respiratory and Critical Care Medicine* 186, 763–772.
- Benton, R., Vannice, K.S., Voshall, L.B., 2007. An essential role for a CD36-related receptor in pheromone detection in *Drosophila*. *Nature* 450, 289–293.
- Binder, C.J., Chang, M.K., Shaw, P.X., *et al.*, 2002. Innate and acquired immunity in atherosclerosis. *Nature Medicine* 8, 1218–1226.
- Blumbach, B., Pancer, Z., Diehl-Seifert, B., *et al.*, 1998. The putative sponge aggregation receptor. Isolation and characterization of a molecule composed of scavenger receptor cysteine-rich domains and short consensus repeats. *Journal of Cell Science* 111 (Pt 17), 2635–2644.
- Bocharov, A.V., Baranova, I.N., Vishnyakova, T.G., *et al.*, 2004. Targeting of scavenger receptor class B type I by synthetic amphipathic α -helical-containing peptides blocks lipopolysaccharide (LPS) uptake and LPS-induced pro-inflammatory cytokine responses in THP-1 monocyte cells. *Journal of Biological Chemistry* 279, 36072–36082.
- Bowdish, D.M.E., Sakamoto, K., Kim, M.J., *et al.*, 2009. MARCO, TLR2, and CD14 are required for macrophage cytokine responses to mycobacterial trehalose dimycolate and *Mycobacterium tuberculosis*. *PLoS Pathogens* 5, e1000474.
- Brannstrom, A., Sankala, M., Tryggvason, K., Pikkarainen, T., 2002. Arginine residues in domain V have a central role for bacteria-binding activity of macrophage scavenger receptor MARCO. *Biochemical and Biophysical Research Communications* 290, 1462–1469.
- Bretscher, A.J., Honti, V., Binggeli, O., *et al.*, 2015. The Nimrod transmembrane receptor Eater is required for hemocyte attachment to the sessile compartment in *Drosophila melanogaster*. *Biology Open* 4 (3), 355–363. doi:10.1242/bio.201410595.
- Buckley, S.T., Ehrhardt, C., 2010. The receptor for advanced glycation end products (RAGE) and the lung. *Journal of Biomedicine and Biotechnology* 2010, 917108.
- Burlone, M.E., Budkowska, A., 2009. Hepatitis C virus cell entry: Role of lipoproteins and cellular receptors. *Journal of General Virology* 90, 1055–1070.
- Cabrera, A., Neculai, D., Kain, K.C., 2014. CD36 and malaria: Friends or foes? A decade of data provides some answers. *Trends in Parasitology* 30, 436–444.
- Cai, L., Wang, Z., Meyer, J.M., Ji, A., Van Der Westhuyzen, D.R., 2012. Macrophage SR-BI regulates LPS-induced pro-inflammatory signaling in mice and isolated macrophages. *Journal of Lipid Research* 53, 1472–1481.
- Calvo, D., Dopazo, J., Vega, M.A., 1995. The CD36, CLA-1 (CD36L1), and LIMPII (CD36L2) gene family: Cellular distribution, chromosomal location, and genetic evolution. *Genomics* 25, 100–106.
- Canton, J., Neculai, D., Grinstein, S., 2013. Scavenger receptors in homeostasis and immunity. *Nature Reviews Immunology* 13, 621–634.
- Cecchini, M.G., Dominguez, M.G., Mocchi, S., *et al.*, 1994. Role of colony stimulating factor-1 in the establishment and regulation of tissue macrophages during postnatal development of the mouse. *Development* 120, 1357–1372.
- Chadwick, A.C., Sahoo, D., 2013. Functional genomics of the human high-density lipoprotein receptor scavenger receptor BI: An old dog with new tricks. *Current Opinion in Endocrinology, Diabetes and Obesity* 20, 124–131.
- Chen, Y., Pikkarainen, T., Elomaa, O., *et al.*, 2005. Defective microarchitecture of the spleen marginal zone and impaired response to a thymus-independent type 2 antigen in mice lacking scavenger receptors MARCO and SR-A. *Journal of Immunology* 175, 8173–8180.
- Chen, Y., Wang, X., Ben, J., *et al.*, 2006. The di-leucine motif contributes to class A scavenger receptor-mediated internalization of acetylated lipoproteins. *Arteriosclerosis, Thrombosis, and Vascular Biology* 26, 1317–1322.
- Chinetti, G., Gbaguidi, F.G., Griglio, S., *et al.*, 2000. CLA-1/SR-BI is expressed in atherosclerotic lesion macrophages and regulated by activators of peroxisome proliferator-activated receptors. *Circulation* 101, 2411–2417.
- Cholewa, J., Nikolic, D., Post, S.R., 2010. Regulation of class A scavenger receptor-mediated cell adhesion and surface localization by PI3K: Identification of a regulatory cytoplasmic motif. *Journal of Leukocyte Biology* 87, 443–449.
- Dardis, A., Filocamo, M., Grossi, S., *et al.*, 2009. Biochemical and molecular findings in a patient with myoclonic epilepsy due to a mistarget of the beta-glucosidase enzyme. *Molecular Genetics and Metabolism* 97, 309–311.
- De Villiers, W.J., Fraser, I.P., Hughes, D.A., Doyle, A.G., Gordon, S., 1994. Macrophage-colony-stimulating factor selectively enhances macrophage scavenger receptor expression and function. *Journal of Experimental Medicine* 180, 705–709.
- DeVries-Seimon, T., Li, Y., Pin, M.Y., *et al.*, 2005. Cholesterol-induced macrophage apoptosis requires ER stress pathways and engagement of the type A scavenger receptor. *Journal of Cell Biology* 171, 61–73.
- Doi, T., Kurasawa, M., Higashino, K., *et al.*, 1994. The histidine interruption of an α -helical coiled coil allosterically mediates a pH-dependent ligand dissociation from macrophage scavenger receptors. *Journal of Biological Chemistry* 269, 25598–25604.
- Dorrington, M.G., Roche, A.M., Chauvin, S.E., *et al.*, 2013. MARCO is required for TLR2- and Nod2-mediated responses to *Streptococcus pneumoniae* and clearance of pneumococcal colonization in the murine nasopharynx. *Journal of Immunology* 190, 250–258.
- Drickamer, K., Dodd, R.B., 1999. C-Type lectin-like domains in *Caenorhabditis elegans*: Predictions from the complete genome sequence. *Glycobiology* 9, 1357–1369.
- Eckhardt, E.R., Cai, L., Shetty, S., *et al.*, 2006. High density lipoprotein endocytosis by scavenger receptor SR-BII is clathrin-dependent and requires a carboxyl-terminal dileucine motif. *Journal of Biological Chemistry* 281, 4348–4353.
- El Khoury, J., Hickman, S.E., Thomas, C.A., *et al.*, 1996. Scavenger receptor-mediated adhesion of microglia to β -amyloid fibrils. *Nature* 382, 716–719.
- El Khoury, J.B., Moore, K.J., Means, T.K., *et al.*, 2003. CD36 mediates the innate host response to β -amyloid. *Journal of Experimental Medicine* 197, 1657–1666.
- Etzerodt, A., Maniecki, M.B., Graversen, J.H., *et al.*, 2012. Efficient intracellular drug-targeting of macrophages using stealth liposomes directed to the hemoglobin scavenger receptor CD163. *Journal of Controlled Release* 160, 72–80.
- Fadok, V.A., Bratton, D.L., Henson, P.M., 2001. Phagocyte receptors for apoptotic cells: Recognition, uptake, and consequences. *Journal of Clinical Investigation* 108, 957–962.
- Fagerberg, L., Hallstrom, B.M., Oksvold, P., *et al.*, 2014. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Molecular & Cellular Proteomics* 13, 397–406.
- Febbraio, M., Podrez, E.A., Smith, J.D., *et al.*, 2000. Targeted disruption of the class B, scavenger receptor CD36 protects against atherosclerotic lesion development in mice. *Journal of Clinical Investigation* 105, 1049–1056.
- Fenske, S.A., Yesilaltay, A., Pal, R., *et al.*, 2009. Normal hepatic cell surface localization of the high density lipoprotein receptor, scavenger receptor class B, type I, depends on all four PDZ domains of PDZK1. *Journal of Biological Chemistry* 284, 5797–5806.
- Fong, L.G., Le, D., 1999. The processing of ligands by the class A scavenger receptor is dependent on signal information located in the cytoplasmic domain. *Journal of Biological Chemistry* 274, 36808–36816.
- Franc, N.C., Heitzler, P., Ezekowitz, R.A.B., White, K., 1999. Requirement for croquemort in phagocytosis of apoptotic cells in *Drosophila*. *Science* 284, 1991–1994.
- Freeman, M., Ashkenas, J., Rees, D.J., *et al.*, 1990. An ancient, highly conserved family of cysteine-rich protein domains revealed by cloning type I and type II murine macrophage scavenger receptors. *Proceedings of the National Academy of Sciences of the United States of America* 87, 8810–8814.
- Fry, A.E., Ghansai, A., Small, K.S., *et al.*, 2009. Positive selection of a CD36 nonsense variant in sub-Saharan Africa, but no association with severe malaria phenotypes. *Human Molecular Genetics* 18, 2683–2692.
- Fukuda, S., Horiuchi, S., Tomita, K., *et al.*, 1986. Acetylated low-density lipoprotein is endocytosed through coated pits by rat peritoneal macrophages. *Virchows Archiv B, Cell Pathology Including Molecular Pathology* 52, 1–13.

- Fulton, W.B., Reeves, R.H., Takeya, M., De Maio, A., 2006. A quantitative trait loci analysis to map genes involved in lipopolysaccharide-induced inflammatory response: Identification of macrophage scavenger receptor 1 as a candidate gene. *Journal of Immunology* 176, 3767–3773.
- Gautam, S., Pirabu, L., Agrawal, C.G., Banerjee, M., 2013. CD36 gene variants and their association with type 2 diabetes in an Indian population. *Diabetes Technology and Therapeutics* 15, 680–687.
- Goldstein, J.L., Ho, Y.K., Basu, S.K., Brown, M.S., 1979. Binding site on macrophages that mediates uptake and degradation of acetylated low density lipoprotein, producing massive cholesterol deposition. *Proceedings of the National Academy of Sciences of the United States of America* 76, 333–337.
- Gordon, S., Martinez, F.O., 2010. Alternative activation of macrophages: Mechanism and functions. *Immunity* 32, 593–604.
- Hagemann, T., Wilson, J., Burke, F., et al., 2006. Ovarian cancer cells polarize macrophages toward a tumor-associated phenotype. *Journal of Immunology* 176, 5023–5032.
- Hampton, R.Y., Golenbock, D.T., Penman, M., Krieger, M., Raetz, C.R., 1991. Recognition and plasma clearance of endotoxin by scavenger receptors. *Nature* 352, 342–344.
- Handattu, S.P., Garber, D.W., Monroe, C.E., et al., 2009. Oral apolipoprotein A-I mimetic peptide improves cognitive function and reduces amyloid burden in a mouse model of Alzheimer's disease. *Neurobiology of Disease* 34, 525–534.
- Handattu, S.P., Monroe, C.E., Nayyar, G., et al., 2013. *In vivo* and *in vitro* effects of an apolipoprotein E mimetic peptide on amyloid- β pathology. *Journal of Alzheimer's Disease* 36, 335–347.
- Hansen, B., Longati, P., Elvevold, K., et al., 2005. Stabilin-1 and stabilin-2 are both directed into the early endocytic pathway in hepatic sinusoidal endothelium via interactions with clathrin/AP-2, independent of ligand binding. *Experimental Cell Research* 303, 160–173.
- Harris, E.N., Weigel, P.H., 2008. The ligand-binding profile of HARE: Hyaluronan and chondroitin sulfates A, C, and D bind to overlapping sites distinct from the sites for heparin, acetylated low-density lipoprotein, dermatan sulfate, and CS-E. *Glycobiology* 18, 638–648.
- Hawkes, M., Li, X., Crockett, M., et al., 2010. CD36 deficiency attenuates experimental mycobacterial infection. *BMC Infectious Diseases* 10, 299.
- Haworth, R., Platt, N., Keshav, S., et al., 1997. The macrophage scavenger receptor type A is expressed by activated macrophages and protects the host against lethal endotoxic shock. *Journal of Experimental Medicine* 186, 1431–1439.
- Heit, B., Kim, H., Cosio, G., et al., 2013. Multimolecular signaling complexes enable Syk-mediated signaling of CD36 internalization. *Developmental Cell* 24, 372–383.
- Helming, L., Winter, J., Gordon, S., 2009. The scavenger receptor CD36 plays a role in cytokine-induced macrophage fusion. *Journal of Cell Science* 122, 453–459.
- Herboso, L., Talamillo, A., Pérez, C., Barrio, R., 2011. Expression of the scavenger receptor class B type I (SR-BI) family in *Drosophila melanogaster*. *International Journal of Developmental Biology* 55, 603–611.
- Herrmann, M., Schäfer, C., Heiss, A., et al., 2012. Clearance of fetuin-A-containing calciprotein particles is mediated by scavenger receptor-A. *Circulation Research* 111, 575–584.
- Hickman, S.E., Allison, E.K., El Khoury, J., 2008. Microglial dysfunction and defective β -amyloid clearance pathways in aging Alzheimer's disease mice. *Journal of Neuroscience* 28, 8354–8360.
- Hirano, K., Kuwasako, T., Nakagawa-Toyama, Y., et al., 2003. Pathophysiology of human genetic CD36 deficiency. *Trends in Cardiovascular Medicine* 13, 136–141.
- Hirose, Y., Saijou, E., Sugano, Y., et al., 2012. Inhibition of Stabilin-2 elevates circulating hyaluronic acid levels and prevents tumor metastasis. *Proceedings of the National Academy of Sciences of the United States of America* 109, 4263–4268.
- Hoebe, K., Georgel, P., Rutschmann, S., et al., 2005. CD36 is a sensor of diacylglycerides. *Nature* 433, 523–527.
- Hohenester, E., Sasaki, T., Timpl, R., 1999. Crystal structure of a scavenger receptor cysteine-rich domain sheds light on an ancient superfamily. *Nature Structural & Molecular Biology* 6, 228–232.
- Horiuchi, S., Sakamoto, Y., Sakai, M., 2003. Scavenger receptors for oxidized and glycated proteins. *Amino Acids* 25, 283–292.
- Huang, J., Zheng, D.L., Qin, F.S., et al., 2010. Genetic and epigenetic silencing of SCARA5 may contribute to human hepatocellular carcinoma by activating FAK signaling. *Journal of Clinical Investigation* 120, 223–241.
- Huang, S., Yuan, S., Guo, L., et al., 2008. Genomic analysis of the immune gene repertoire of amphioxus reveals extraordinary innate complexity and diversity. *Genome Research* 18, 1112–1126.
- Huber, O., Sumper, M., 1994. Algal-CAMs: Isoforms of a cell adhesion molecule in embryos of the alga *Volvox* with homology to *Drosophila* fasciclin I. *EMBO Journal* 13, 4212–4222.
- Hughes, D.A., Fraser, L.P., Gordon, S., 1995. Murine macrophage scavenger receptor: *In vivo* expression and function as receptor for macrophage adhesion in lymphoid and non-lymphoid organs. *European Journal of Immunology* 25, 466–473.
- Ishii, J., Adachi, H., Aoki, J., et al., 2002. SREC-II, a new member of the scavenger receptor type F family, trans-interacts with SREC-I through its extracellular domain. *Journal of Biological Chemistry* 277, 39696–39702.
- Janssen, K.P., Rost, R., Eichinger, L., Schleicher, M., 2001. Characterization of CD36/LIMP-II Homologues in *Dictyostelium discoideum*. *Journal of Biological Chemistry* 276, 38899–38910.
- Janvier, K., Kato, Y., Boehm, M., et al., 2003. Recognition of dileucine-based sorting signals from HIV-1 Nef and LIMP-II by the AP-1 gamma-sigma1 and AP-3 delta-sigma3 hemicomplexes. *Journal of Cell Biology* 163, 1281–1290.
- Jemielity, S., Wang, J.J., Chan, Y.K., et al., 2013. TIM-family proteins promote infection of multiple enveloped viruses through virion-associated phosphatidylserine. *PLoS Pathogens* 9, e1003232.
- Kanost, M.R., Jiang, H., Yu, X.Q., 2004. Innate immune responses of a lepidopteran insect, *Manduca sexta*. *Immunological Reviews* 198, 97–105.
- Karakesisoglou, I., Janssen, K.P., Eichinger, L., Noegel, A.A., Schleicher, M., 1999. Identification of a suppressor of the *Dictyostelium* profilin-minus phenotype as a CD36/LIMP-II homologue. *Journal of Cell Biology* 145, 167–181.
- Kay, J.N., Chu, M.W., Sanes, J.R., 2012. MEGF10 and MEGF11 mediate homotypic interactions required for mosaic spacing of retinal neurons. *Nature* 483, 465–469.
- Kiefer, C., Sumser, E., Wernet, M.F., Von Lintig, J., 2002. A class B scavenger receptor mediates the cellular uptake of carotenoids in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America* 99, 10581–10586.
- Kocks, C., Cho, J.H., Nehme, N., et al., 2005. Eater, a transmembrane protein mediating phagocytosis of bacterial pathogens in *Drosophila*. *Cell* 123, 335–346.
- Kodama, T., Freeman, M., Rohrer, L., et al., 1990. Type I macrophage scavenger receptor contains alpha-helical and collagen-like coiled coils. *Nature* 343, 531–535.
- Kosswig, N., Rice, S., Daugherty, A., Post, S.R., 2003. Class A scavenger receptor-mediated adhesion and internalization require distinct cytoplasmic domains. *Journal of Biological Chemistry* 278, 34219–34225.
- Krieger, M., 1997. The other side of scavenger receptors: Pattern recognition for host defense. *Current Opinion in Lipidology* 8, 275–280.
- Krieger, M., Abrams, J.M., Lux, A., Steller, H., 1992. Molecular flypaper, atherosclerosis, and host defense: Structure and function of the macrophage scavenger receptor. *Cold Spring Harbor Symposia on Quantitative Biology* 57, 605–609.
- Kristiansen, M., Graversen, J.H., Jacobsen, C., et al., 2001. Identification of the haemoglobin scavenger receptor. *Nature* 409, 198–201.
- Kumano-Kuramochi, M., Xie, Q., Kajiwara, S., et al., 2013. Lectin-like oxidized LDL receptor-1 is palmitoylated and internalizes ligands via caveolae/raft-dependent endocytosis. *Biochemical and Biophysical Research Communications* 434, 594–599.
- Kurushima, H., Ramprasad, M., Kondratenko, N., et al., 2000. Surface expression and rapid internalization of macrophage (mouse CD68) on elicited mouse peritoneal macrophages. *Journal of Leukocyte Biology* 67, 104–108.
- Lata, S., Raghava, G.P., 2008. PRRDB: A comprehensive database of pattern-recognition receptors and their ligands. *BMC Genomics* 9, 180.
- Lemke, G., Burstyn-Cohen, T., 2010. TAM receptors and the clearance of apoptotic cells. *Annals of the New York Academy of Sciences* 1209, 23–29.
- Li, J.Y., Paragas, N., Ned, R.M., et al., 2009. Scaras5 is a ferritin receptor mediating non-transferrin iron delivery. *Developmental Cell* 16, 35–46.
- Lipinski, M.J., Fuster, V., Fisher, E.A., Fayad, Z.A., 2004. Technology insight: Targeting of biological molecules for evaluation of high-risk atherosclerotic plaques with magnetic resonance imaging. *Nature Clinical Practice Cardiovascular Medicine* 1, 48–55.
- Liu, S., McCormick, K.D., Zhao, W., et al., 2012. Human apolipoprotein E peptides inhibit hepatitis C virus entry by blocking virus binding. *Hepatology* 56, 484–491.
- Manaka, J., Kuraishi, T., Shiratsuchi, A., et al., 2004. Draper-mediated and phosphatidylserine-independent phagocytosis of apoptotic cells by *Drosophila* hemocytes/macrophages. *Journal of Biological Chemistry* 279, 48466–48476.
- Manichaikul, A., Naj, A.C., Herrington, D., et al., 2012. Association of SCARB1 variants with subclinical atherosclerosis and incident cardiovascular disease: The multi-ethnic study of atherosclerosis. *Arteriosclerosis, Thrombosis, and Vascular Biology* 32, 1991–1999.
- Martens, J.H., Kzyshkowska, J., Falkowski-Hansen, M., et al., 2006. Differential expression of a gene signature for scavenger/lectin receptors by endothelial cells

- and macrophages in human lymph node sinuses, the primary sites of regional metastasis. *Journal of Pathology* 208, 574–589.
- Martin, C., Chevrot, M., Poirier, H., *et al.*, 2011. CD36 as a lipid sensor. *Physiology and Behavior* 105, 36–42.
- McCarthy, J.J., Lehner, T., Reeves, C., *et al.*, 2003. Association of genetic variants in the HDL receptor, SR-B1, with abnormal lipids in women with coronary artery disease. *Journal of Medical Genetics* 40, 453–458.
- Mehta, J.L., Sanada, N., Hu, C.P., *et al.*, 2007. Deletion of LOX-1 reduces atherogenesis in LDLR knockout mice fed high cholesterol diet. *Circulation Research* 100, 1634–1642.
- Mercer, J., Helenius, A., 2008. Vaccinia virus uses macropinocytosis and apoptotic mimicry to enter host cells. *Science* 320, 531–535.
- Messier-Solek, C., Buckley, K.M., Rast, J.P., 2010. Highly diversified innate receptor systems and new forms of animal immunity. *Seminars in Immunology* 22, 39–47.
- Moore, K.J., Freeman, M.W., 2006. Scavenger receptors in atherosclerosis: Beyond lipid uptake. *Arteriosclerosis, Thrombosis, and Vascular Biology* 26, 1702–1711.
- Mukhopadhyay, S., Herre, J., Brown, G.D., Gordon, S., 2004. The potential for Toll-like receptors to collaborate with other innate immune receptors. *Immunology* 112, 521–530.
- Mukhopadhyay, S., Plüddemann, A., Gordon, S., 2009. Macrophage pattern recognition receptors in immunity, homeostasis and self tolerance. In: Kishore, U. (Ed.), *Target Pattern Recognition in Innate Immunity*. Advances in Experimental Medicine and Biology, vol. 653. New York, NY: Springer, pp. 1–14.
- Mukhopadhyay, S., Varin, A., Chen, Y., *et al.*, 2011. SR-A/MARCO-mediated ligand delivery enhances intracellular TLR and NLR function, but ligand scavenging from cell surface limits TLR4 response to pathogens. *Blood* 117, 1319–1328.
- Muller, W.E., Thakur, N.L., Ushijima, H., *et al.*, 2004. Matrix-mediated canal formation in primmorphs from the sponge *Suberites domuncula* involves the expression of a CD36 receptor-ligand system. *Journal of Cell Science* 117, 2579–2590.
- Munier, A.I., Medzhitov, R., Janeway Jr., C.A., *et al.*, 2004. Graal: A *Drosophila* gene coding for several mosaic serine proteases. *Insect Biochemistry and Molecular Biology* 34, 1025–1035.
- Murphy, J.E., Tedbury, P.R., Homer-Vanniasinkam, S., Walker, J.H., Ponnambalam, S., 2005. Biochemistry and cell biology of mammalian scavenger receptors. *Atherosclerosis* 182, 1–15.
- Murphy, J.E., Vohra, R.S., Dunn, S., *et al.*, 2008. Oxidised LDL internalisation by the LOX-1 scavenger receptor is dependent on a novel cytoplasmic motif and is regulated by dynamin-2. *Journal of Cell Science* 121, 2136–2147.
- Naj, A.C., West, M., Rich, S.S., *et al.*, 2010. Association of scavenger receptor class B type I polymorphisms with subclinical atherosclerosis: The multi-ethnic study of atherosclerosis. *Circulation: Cardiovascular Genetics* 3, 47–52.
- Nakamura, K., Funakoshi, H., Miyamoto, K., Tokunaga, F., Nakamura, T., 2001. Molecular cloning and functional characterization of a human scavenger receptor with C-type lectin (SRCL), a novel member of a scavenger receptor family. *Biochemical and Biophysical Research Communications* 280, 1028–1035.
- Navab, M., Anantharamaiah, G.M., Reddy, S.T., Fogelman, A.M., 2006. Apolipoprotein A-I mimetic peptides and their role in atherosclerosis prevention. *Nature Clinical Practice Cardiovascular Medicine* 3, 540–547.
- Neculai, D., Schwake, M., Ravichandran, M., *et al.*, 2013. Structure of LIMP-2 provides functional insights with implications for SR-BI and CD36. *Nature* 504, 172–176.
- Neyen, C., Mukhopadhyay, S., Gordon, S., Hagemann, T., 2013a. An apolipoprotein A-I mimetic targets scavenger receptor A on tumor-associated macrophages: A prospective anticancer treatment? *Oncoimmunology* 2, e24461.
- Neyen, C., Plüddemann, A., Mukhopadhyay, S., *et al.*, 2013b. Macrophage scavenger receptor A promotes tumor progression in murine models of ovarian and pancreatic cancer. *Journal of Immunology* 190, 3798–3805.
- Neyen, C., Plüddemann, A., Roversi, P., *et al.*, 2009. Macrophage scavenger receptor A mediates adhesion to apolipoproteins A-I and E. *Biochemistry* 48, 11858–11871.
- Nielsen, T.J.F., Shaw, J.T., Jaipuri, F.A., *et al.*, 2008. Identification of the molecular target of small molecule inhibitors of HDL receptor SR-BI activity. *Biochemistry* 47, 460–472.
- Nielsen, M.J., Andersen, C.B., Moestrup, S.K., 2013. CD163 binding to haptoglobin-hemoglobin complexes involves a dual-point electrostatic receptor-ligand pairing. *Journal of Biological Chemistry* 288, 18834–18841.
- Nikolic, D.M., Cholewa, J., Gass, C., Gong, M.C., Post, S.R., 2007. Class A scavenger receptor-mediated cell adhesion requires the sequential activation of Lyn and PI3-kinase. *American Journal of Physiology – Cell Physiology* 292, C1450–C1458.
- Orloff, M., Peterson, C., He, X., *et al.*, 2011. Germline mutations in MSR1, ASCC1, and CTHRC1 in patients with Barrett esophagus and esophageal adenocarcinoma. *JAMA* 306, 410–419.
- Osgood, D., Corella, D., Demissie, S., *et al.*, 2003. Genetic variation at the scavenger receptor class B type I gene locus determines plasma lipoprotein concentrations and particle size and interacts with type 2 diabetes: The Framingham study. *Journal of Clinical Endocrinology and Metabolism* 88, 2869–2879.
- Pearson, A., Lux, A., Krieger, M., 1995. Expression cloning of dSR-CI, a class C macrophage-specific scavenger receptor from *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences of the United States of America* 92, 4056–4060.
- Peiser, L., Winther, D.E., Makepeace, M.P.J., *et al.*, 2002. The class A macrophage scavenger receptor is a major pattern recognition receptor for *Neisseria meningitidis* which is independent of lipopolysaccharide and not required for secretory responses. *Infection and Immunity* 70, 5346–5354.
- Philips, J.A., Rubin, E.J., Perrimon, N., 2005. *Drosophila* RNAi screen reveals CD36 family member required for mycobacterial infection. *Science* 309, 1251–1253.
- Plüddemann, A., Hoe, J.C., Makepeace, K., Moxon, E.R., Gordon, S., 2009. The macrophage scavenger receptor A is host-protective in experimental meningococcal septicemia. *PLoS Pathogens* 5, e1000297.
- Plüddemann, A., Mukhopadhyay, S., Gordon, S., 2006. The interaction of macrophage receptors with bacterial ligands. *Expert Reviews in Molecular Medicine* 8, 1–25.
- PrabhuDas, M., Bowdish, D., Drickamer, K., *et al.*, 2014. Standardizing scavenger receptor nomenclature. *Journal of Immunology* 192, 1997–2006.
- Prevo, R., Banerji, S., Ni, J., Jackson, D.G., 2004. Rapid plasma membrane-endosomal trafficking of the lymph node sinus and high endothelial venule scavenger receptor/homing receptor stabilin-1 (FEEL-1/CLEVER-1). *Journal of Biological Chemistry* 279, 52580–52592.
- Rämet, M., Pearson, A., Manfrulli, P., *et al.*, 2001. *Drosophila* scavenger receptor CI is a pattern recognition receptor for bacteria. *Immunity* 15, 1027–1038.
- Rast, J.P., Smith, L.C., Loza-Coll, M., Hibino, T., Litman, G.W., 2006. Genomic insights into the immune system of the sea urchin. *Science* 314, 952–956.
- Ravichandran, K.S., Lorenz, U., 2007. Engulfment of apoptotic cells: Signals for a good meal. *Nature Reviews Immunology* 7, 964–974.
- Reczek, D., Schwake, M., Schroder, J., *et al.*, 2007. LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase. *Cell* 131, 770–783.
- Rensen, P.C.N., Gras, J.C.E., Lindfors, E.K., *et al.*, 2006. Selective targeting of liposomes to macrophages using a ligand with high affinity for the macrophage scavenger receptor class A. *Current Drug Discovery Technologies* 3, 135–144.
- Resnick, D., Pearson, A., Krieger, M., 1994. The SRCR superfamily: A family reminiscent of the Ig superfamily. *Trends in Biochemical Sciences* 19, 5–8.
- Ricote, M., Li, A.C., Willson, T.M., Kelly, C.J., Glass, C.K., 1998. The peroxisome proliferator-activated receptor- γ is a negative regulator of macrophage activation. *Nature* 391, 79–82.
- Rigotti, A., Acton, S.L., Krieger, M., 1995. The class B scavenger receptors SR-BI and CD36 are receptors for anionic phospholipids. *Journal of Biological Chemistry* 270, 16221–16224.
- Rigotti, A., Miettinen, H.E., Krieger, M., 2003. The role of the high-density lipoprotein receptor SR-BI in the lipid metabolism of endocrine and other tissues. *Endocrine Reviews* 24, 357–387.
- Rigotti, A., Trigatti, B.L., Penman, M., *et al.*, 1997. A targeted mutation in the murine gene encoding the high density lipoprotein (HDL) receptor scavenger receptor class B type I reveals its key role in HDL metabolism. *Proceedings of the National Academy of Sciences of the United States of America* 94, 12610–12615.
- Roberts, C.G.P., Shen, H., Mitchell, B.D., *et al.*, 2007. Variants in scavenger receptor class B type I gene are associated with HDL cholesterol levels in younger women. *Human Heredity* 64, 107–113.
- Rodrigues, C.D., Hannus, M., Prudêncio, M., *et al.*, 2008. Host scavenger receptor SR-BI plays a dual role in the establishment of malaria parasite liver infection. *Cell Host and Microbe* 4, 271–282.
- Rohrer, L., Freeman, M., Kodama, T., Penman, M., Krieger, M., 1990. Coiled-coil fibrous domains mediate ligand binding by macrophage scavenger receptor type II. *Nature* 343, 570–572.
- Santiago-Garcia, J., Kodama, T., Pitas, R.E., 2003. The class A scavenger receptor binds to proteoglycans and mediates adhesion of macrophages to the extracellular matrix. *Journal of Biological Chemistry* 278, 6942–6946.
- Sarrias, M.R., Gronlund, J., Padilla, O., *et al.*, 2004. The Scavenger Receptor Cysteine-Rich (SRCR) domain: An ancient and highly conserved protein module of the innate immune system. *Critical Reviews in Immunology* 24, 1–37.
- Scarselli, E., Ansuini, H., Cerino, R., *et al.*, 2002. The human scavenger receptor class B type I is a novel candidate receptor for the hepatitis C virus. *EMBO Journal* 21, 5017–5025.

- Schafer, D.P., Lehrman, E.K., Kautzman, A.G., *et al.*, 2012. Microglia sculpt postnatal neural circuits in an activity and complement-dependent manner. *Neuron* 74, 691–705.
- Schmidt, A.M., Yan, S.D., Yan, S.F., Stern, D.M., 2001. The multiligand receptor RAGE as a progression factor amplifying immune and inflammatory responses. *Journal of Clinical Investigation* 108, 949–955.
- Segawa, K., Suzuki, J., Nagata, S., 2011. Constitutive exposure of phosphatidylserine on viable cells. *Proceedings of the National Academy of Sciences of the United States of America* 108, 19246–19251.
- Segers, F.M.E., Den Adel, B., Bot, I., *et al.*, 2013. Scavenger receptor-AI-targeted iron oxide nanoparticles for *in vivo* MRI detection of atherosclerotic lesions. *Arteriosclerosis, Thrombosis, and Vascular Biology* 33, 1812–1819.
- Seimon, T.A., Obstfeld, A., Moore, K.J., Golenbock, D.T., Tabas, I., 2006. Combinatorial pattern recognition receptor signaling alters the balance of life and death in macrophages. *Proceedings of the National Academy of Sciences of the United States of America* 103, 19794–19799.
- Sheedy, F.J., Grebe, A., Rayner, K.J., *et al.*, 2013. CD36 coordinates NLRP3 inflammasome activation by facilitating intracellular nucleation of soluble ligands into particulate ligands in sterile inflammation. *Nature Immunology* 14, 812–820.
- Shibata, M., Ishii, J., Koizumi, H., *et al.*, 2004. Type F scavenger receptor SREC-I interacts with advillin, a member of the gelsolin/villin family, and induces neurite-like outgrowth. *Journal of Biological Chemistry* 279, 40084–40090.
- Silverstein, R.L., Febbraio, M., 2009. CD36, a scavenger receptor involved in immunity, metabolism, angiogenesis, and behavior. *Science Signaling* 2, re3.
- Simpson, M.A., Weigel, J.A., Weigel, P.H., 2012. Systemic blockade of the hyaluronan receptor for endocytosis prevents lymph node metastasis of prostate cancer. *International Journal of Cancer* 131, E836–E840.
- Somogyi, K., Sipos, B., Péntzes, Z., *et al.*, 2008. Evolution of genes and repeats in the Nimrod superfamily. *Molecular Biology and Evolution* 25, 2337–2347.
- Sørensen, K.K., McCourt, P., Berg, T., *et al.*, 2012. The scavenger endothelial cell: A new player in homeostasis and immunity. *American Journal of Physiology – Regulatory Integrative and Comparative Physiology* 303, R1217–R1230.
- Stewart, C.R., Stuart, L.M., Wilkinson, K., *et al.*, 2010. CD36 ligands promote sterile inflammation through assembly of a Toll-like receptor 4 and 6 heterodimer. *Nature Immunology* 11, 155–161.
- Stuart, L.M., Bell, S.A., Stewart, C.R., *et al.*, 2007. CD36 signals to the actin cytoskeleton and regulates microglial migration via a p130Cas complex. *Journal of Biological Chemistry* 282, 27392–27401.
- Stuart, L.M., Deng, J., Silver, J.M., *et al.*, 2005. Response to *Staphylococcus aureus* requires CD36-mediated phagocytosis triggered by the COOH-terminal cytoplasmic domain. *Journal of Cell Biology* 170, 477–485.
- Stuart, L.M., Ezekowitz, R.A., 2008. Phagocytosis and comparative innate immunity: Learning on the fly. *Nature Reviews Immunology* 8, 131–141.
- Suzuki, H., Kurihara, Y., Takeya, M., *et al.*, 1997. A role for macrophage scavenger receptors in atherosclerosis and susceptibility to infection. *Nature* 386, 292–296.
- Tabet, F., Vickers, K.C., Cuesta Torres, L.F., *et al.*, 2014. HDL-transferred microRNA-223 regulates ICAM-1 expression in endothelial cells. *Nature Communications* 5, 3292.
- Tai, E.S., Adiconis, X., Ordovas, J.M., *et al.*, 2003. Polymorphisms at the SRBI locus are associated with lipoprotein levels in subjects with heterozygous familial hypercholesterolemia. *Clinical Genetics* 63, 53–58.
- Talamillo, A., Herboso, B., Pirone, L., *et al.*, 2013. Scavenger receptors mediate the role of SUMO and Ftz-I1 in *Drosophila steroidogenesis*. *PLoS Genetics* 9, e1003473.
- Tall, A.R., 1998. An overview of reverse cholesterol transport. *European Heart Journal* 19, A31–A35.
- Teslovich, T.M., Musunuru, K., Smith, A.V., *et al.*, 2010. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 466, 707–713.
- Tontonoz, P., Nagy, L., Alvarez, J.G.A., Thomazy, V.A., Evans, R.M., 1998. PPAR γ promotes monocyte/macrophage differentiation and uptake of oxidized LDL. *Cell* 93, 241–252.
- Tu, C., Ng, T.S.C., Sohi, H.K., *et al.*, 2011. Receptor-targeted iron oxide nanoparticles for molecular MR imaging of inflamed atherosclerotic plaques. *Biomaterials* 32, 7209–7216.
- Underhill, D.M., Ozinsky, A., 2002. Phagocytosis of microbes: Complexity in action. *Annual Review of Immunology* 20, 825–852.
- Vadali, S., Post, S.R., 2014. Lipid rafts couple class A scavenger receptors to phospholipase A2 activation during macrophage adhesion. *Journal of Leukocyte Biology* 96, 873–881.
- Varban, M.L., Rinninger, F., Wang, N., *et al.*, 1998. Targeted mutation reveals a central role for SR-BI in hepatic selective uptake of high density lipoprotein cholesterol. *Proceedings of the National Academy of Sciences of the United States of America* 95, 4619–4624.
- Vega, M.A., Rodriguez, F., Segui, B., *et al.*, 1991. Targeting of lysosomal integral membrane protein LIMP II. The tyrosine-lacking carboxyl cytoplasmic tail of LIMP II is sufficient for direct targeting to lysosomes. *Journal of Biological Chemistry* 266, 16269–16272.
- Vickers, K.C., Palmisano, B.T., Shoucri, B.M., Shamburek, R.D., Remaley, A.T., 2011. MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nature Cell Biology* 13, 423–435.
- Vickers, K.C., Remaley, A.T., 2012. Lipid-based carriers of microRNAs and intercellular communication. *Current Opinion in Lipidology* 23, 91–97.
- Vohra, R.S., Walker, J.H., Howell, G.J., Homer-Vanniasinkam, S., Ponnambalam, S., 2009. The LOX-1 scavenger receptor cytoplasmic domain contains a transplantable endocytic motif. *Biochemical and Biophysical Research Communications* 383, 269–274.
- Voisset, C., Callens, N., Blanchard, E., *et al.*, 2005. High density lipoproteins facilitate hepatitis C virus entry through the scavenger receptor class B type I. *Journal of Biological Chemistry* 280, 7793–7799.
- Wang, T., Jiao, Y., Montell, C., 2007a. Dissection of the pathway required for generation of vitamin A and for *Drosophila* phototransduction. *Journal of Cell Biology* 177, 305–316.
- Wang, X., Collins, H.L., Ranalletta, M., *et al.*, 2007b. Macrophage ABCA1 and ABCG1, but not SR-BI, promote macrophage reverse cholesterol transport *in vivo*. *Journal of Clinical Investigation* 117, 2216–2224.
- Whelan, F.J., Meehan, C.J., Golding, G.B., McConkey, B.J., Bowdish, D.M.E., 2012. The evolution of the class A scavenger receptors. *BMC Evolutionary Biology* 12, 227.
- Williams, D.L., De La Llera-Moya, M., Thuthnani, S.T., *et al.*, 2000. Binding and cross-linking studies show that scavenger receptor BI interacts with multiple sites in apolipoprotein A-I and identify the class A amphipathic α -helix as a recognition motif. *Journal of Biological Chemistry* 275, 18897–18904.
- Witmer-Pack, M.D., Hughes, D.A., Schuler, G., *et al.*, 1993. Identification of macrophages and dendritic cells in the osteopetrotic (op/op) mouse. *Journal of Cell Science* 104 (Pt 4), 1021–1029.
- Wouters, M.A., Rigoutsos, I., Chu, C.K., *et al.*, 2005. Evolution of distinct EGF domains with specific functions. *Protein Science* 14, 1091–1103.
- Wu, H.H., Bellmunt, E., Scheib, J.L., *et al.*, 2009. Glial precursors clear sensory neuron corpses during development via Jedi-1, an engulfment receptor. *Nature Neuroscience* 12, 1534–1541.
- Xu, J., Zheng, S.L., Komiya, A., *et al.*, 2002. Germline mutations and sequence variants of the macrophage scavenger receptor 1 gene are associated with prostate cancer risk. *Nature Genetics* 32, 321–325.
- Yalaoui, S., Huby, T., Franetich, J.F., *et al.*, 2008. Scavenger receptor BI boosts hepatocyte permissiveness to plasmodium infection. *Cell Host and Microbe* 4, 283–292.
- Yamayoshi, S., Yamashita, Y., Li, J., *et al.*, 2009. Scavenger receptor B2 is a cellular receptor for enterovirus 71. *Nature Medicine* 15, 798–801.
- Yan, S.D., Schmidt, A.M., Anderson, G.M., *et al.*, 1994. Enhanced cellular oxidant stress by the interaction of advanced glycation end products with their receptors/binding proteins. *Journal of Biological Chemistry* 269, 9889–9897.
- Zelensky, A.N., Gready, J.E., 2005. The C-type lectin-like domain superfamily. *FEBS Journal* 272, 6179–6217.
- Zhang, J., Gratchev, A., Riabov, V., *et al.*, 2009. A novel GGA-binding site is required for intracellular sorting mediated by stabilin-1. *Molecular and Cellular Biology* 29, 6097–6105.
- Zhou, Z., Hartwig, E., Horvitz, H.R., 2001. CED-1 is a transmembrane receptor that mediates cell corpse engulfment in *C. elegans*. *Cell* 104, 43–56.
- Zhu, X.D., Zhuang, Y., Ben, J.J., *et al.*, 2011. Caveolae-dependent endocytosis is required for class A macrophage scavenger receptor-mediated apoptosis in macrophages. *Journal of Biological Chemistry* 286, 8231–8239.

Relevant Website

<https://respond.niaid.nih.gov/scavengerreceptors/Pages/default.aspx>
National Institutes of Health.

Dendritic Cells

K Liu, Columbia University Medical Center, New York, NY, USA

© 2016 Elsevier Inc. All rights reserved.

Discovery of Dendritic Cells

In search of the cells that initiate immune response, Steinman and colleagues discovered dendritic cells (DCs) using morphology-based method. DCs are shown to be the most potent initiator of T cell immune response comparing to other cells. DC-specific monoclonal antibodies were made and enabled characterization and easy isolation of DCs. DCs form a wide network of antigen capture, transport and presentation. Many subsets of DCs collaboratively orchestrate immune response.

Identification and Isolation

The discovery of DCs was the result of an effort to understand the initiation of immunity. In 1967, Mishell and Dutton developed an *in vitro* priming system in which mouse spleen cell suspensions could be stimulated to generate antibody responses in culture (Mishell and Dutton, 1967). It was quickly observed that lymphocytes alone were not sufficient to induce antibody forming cell responses, and that an adherent accessory cell, or antigen presenting cell (APC), was required. At the time, macrophages were thought to be the key accessory cell because they composed a major population of adherent cells and also because their role in innate immunity had been long appreciated (Unanue and Cerottini, 1970). However, macrophages failed to show robust activity in induction of antibody responses *in vitro*, and they rapidly digested ingested antigen suggesting that they would be unable to present it to lymphocytes (Ehrenreich and Cohn, 1969, 1967; Steinman and Cohn, 1972).

Speculating that another cell type distinct from macrophage initiates immune response, Steinman and Cohn examined spleen adherent cells by phase contrast microscopy; they discovered a small population of cells with stellate shape and named them DC (Steinman and Cohn, 1973). They then employed physical techniques to fractionate spleen cells and purify DCs. DCs were found to adhere to plastic or glass, had low buoyant density, and did not bind to erythrocytes coated with antibody. Comparing to macrophages that contain abundant lysosomes in the cytoplasm and are actively phagocytosing, DCs contain few lysosome and are weaker in phagocytosis.

The original evidence of DCs' immunogenicity came from two crucial experiments. First, in an *in vitro* model system to study graft rejection, i.e., mixed leukocyte reaction, DCs were nearly two orders of magnitude more potent than unfractionated spleen cells, B cells or macrophages in stimulating allogeneic T cells (Steinman and Witmer, 1978; Nussenzweig and Steinman, 1980). This effect is attributed to abundant allogeneic MHC II on DC surface, which does not require priming. Second, in a coculture system with syngeneic responding T cells and hapten-modified thymocytes, DCs induced MHC-restricted cytotoxic T cells specific to the hapten;

in contrast, macrophages and other purified populations of lymphoid cells were nearly inactive as accessory cells (Nussenzweig *et al.*, 1980). Thus DCs were discovered as the most potent APCs to initiate T cell immunity.

Distribution of DCs

Sequential steps of density centrifugation in bovine serum albumin gradients and adherence to glass were originally used to purify DCs (Steinman and Cohn, 1974). Immunization with these purified DCs led to production of a series of DC-specific antibodies in 70–80s, including 33D1, NLDC and N418. These antibodies made it possible to characterize DC distribution *in vivo* with immunohistochemistry. DCs are widely distributed in most organs, a distribution that maximizes antigen capture and subsequent induction of T cell response.

DCs in lymphoid organs

DCs are found in the central lymphoid organs, bone marrow and thymus. In the bone marrow, DCs are organized into perivascular clusters that enveloped blood vessels. These bone marrow DCs contribute to promoting the survival of mature B cells that home back to the bone marrow. In the thymus, DCs reside in thymic cortical epithelium and thymic medulla. Thymus is where T cells mature through positive and negative selection. Thymic DCs mediate central tolerance, through negative selection or clonal elimination of T cells that are self-reactive, and induction of regulatory T cells that suppress self-reactive T cells.

DCs reside in the peripheral lymphoid organs, spleen and lymph nodes, where immune response is initiated. Naïve T cells circulate through T cell areas in spleen and lymph nodes. In these areas, DCs are enmeshed in an extensive network and they are actively probing adjacent T cells with their processes, a crucial scanning step preceding T cell clonal expansion.

DCs in non-lymphoid organs

DCs are also found in most non-lymphoid organs and in all epithelial surfaces that contact the environment. In organs such as heart, lung, kidney, the dermal layer of skin, and meninges and choroid plexus in the brain, they are found in interstitial spaces that are drained by lymphatics. In skin, the epidermal layer contains Langerhans cell expressing Fc receptors and MHC class II, capable of presenting antigen to primed T cells; the dermal layer of skin is also populated with DCs. Epidermal Langerhans cell and dermal DCs have different origin, but both appear to function as DCs, i.e., present antigen and activate T cells. DCs are also found in all stratified squamous epithelia such as the vagina, cervix, anus, pharynx, and esophagus as well as in other epithelia, as in the airways of the lung, the intestine, and the iris and ciliary body.

DC migration

Migration of DCs from peripheral tissues into lymphoid organs is key to their sentinel and antigen trafficking functions. Upon microbial contact or stimulation by inflammatory cytokines, resident DCs from non-lymphoid tissue traffic through afferent lymphatics to the lymph node T cell areas, where they participate in the initiation of immune responses. For example, elicitation of allergic contact dermatitis and immunity induced by HSV or Leishmania requires interaction of antigen with epidermal Langerhans cells or dermal DCs respectively, followed by migration of these DCs to the lymph nodes to present antigen to T lymphocytes.

Even in the absence of invading pathogens, some DCs are always migrating from tissues to lymph nodes through afferent lymph. Most of the migrating DCs die after their arrival in lymphoid tissues, thus DCs are not detectable in the efferent lymph. The DCs that migrate in the steady state might have several functions: to replenish immature populations, to transport self or environmental antigens, or to be on patrol to identify invaders.

DC migration is a regulated process, controlled at the level of chemokine production and chemokine receptor expression and function. CCR7 and its ligands CCL19 and CCL21 are essential for DC migration, both in the steady state and during inflammation (reviewed in [Randolph et al., 2005](#)). Chemokine/chemokine receptor interactions not only orchestrate DCs' migration but also influence their immunogenic potential.

Therefore, DCs in the peripheral tissue and lymphoid organs form a vast network. Migration of DCs enables antigen captured in the periphery to be transported to the lymphoid organ, where they are presented to T cells for initiation of tolerance and immunity.

DC Subsets

In each organ, DCs are heterogeneous and comprised of multiple groups and subsets ([Ardavin and Shortman, 1992](#); [Vremec et al., 1992](#)). These groups differ in their anatomic distribution, cell surface marker expression and function. In the mouse, three major groups of DCs exist in the steady state: plasmacytoid DCs (pDCs), conventional DCs (cDCs), and migratory DCs (mDCs). Under inflammation, a new population of DCs arises from monocytes (mo-DCs).

pDCs

pDCs, marked by the expression of B220 and PDCA1 and their morphological resemblance of plasma cells, are important mediators of antiviral immunity through their ability to produce large amounts of type I interferons (IFNs) on viral infection (see below).

cDCs

cDCs are composed of two major subsets, namely $CD8\alpha^+$ and $CD11b^+$ cDCs ([Shortman and Liu, 2002](#)). They have overlapping functions; both can process and present antigens to T cells; both subsets also secrete cytokines such as IL-12, which can inform the ultimate polarization of the T cell response to pathogens. However, the two subsets also have distinct anatomic localization and functions *in vivo* and are not redundant.

For example, in mouse spleen, $CD8\alpha^+$ cDCs localize to T cell areas and specialize in priming $CD8^+$ T cells, whereas $CD11b^+$ cDCs localize to the bridging channel and specialize in priming $CD4^+$ T cells.

mDCs

mDCs are present in non-lymphoid tissues such as the liver, gut, skin, lung, and aorta, and they too are composed of two main subsets $CD103^+$ and $CD103^-$ DCs (reviewed in [Helft et al., 2010](#)). These non-lymphoid DCs are referred to as mDCs because they transit from tissue to lymphoid organs (reviewed in [Banchereau and Steinman, 1998](#); [Shortman and Liu, 2002](#)). $CD103^+$ tissue DCs are developmentally related to $CD8\alpha^+$ cDCs.

Mo-DCs

Under inflammation, a new subset of DCs arise from monocytes (mo-DCs). For example, upon infection of *Listeria monocytogene*, increased CCL2 drives CCR2-expressing monocytes to emigrate from bone marrow, upregulate CD11c and MHC II on cell surface and differentiate to mo-DCs. Upon entering the infection sites, mo-DCs produce TNF- α and iNOS which confer innate protection. Therefore, another name for monocyte-derived DC (mo-DC) is TNF- α , iNOS-producing DCs (Tip-DCs). In CCR2 $^{-/-}$ mice, monocytes failed to migrate, rendering the mice highly susceptible to *Listeria* infection.

Gene expression comparison, selective targeting and depletion of specific subset *in vivo* reveal that DC subsets have distinct features and functions. They differ in expression of cell surface and cytoplasmic receptors for detecting pathogens, antigen processing capacity, and ability to produce cytokines and chemokines. For example, $CD8\alpha^+$ cDC and their developmentally related $CD103^+$ non-lymphoid DCs are uniquely potent and indispensable in cross-presenting antigens on MHC I and initiate $CD8^+$ T cell immunity against tumor, virus and intracellular pathogen.

In conclusion, multiple subsets of DCs have distinct micro-anatomical locations, gene expression profiles and functions. Collaboratively, these DCs orchestrate immune response.

DC Development and Homeostasis

Growth Factors and DC Cultures

DCs only constitute less than 2% of hematopoietic cells *in vivo*. Therefore, *in vitro* culture systems were established to produce human and mouse DCs in large number; these culture systems are widely used today for basic and clinical studies. Bone marrow cells from mice or peripheral blood monocytes from humans are cultured for 6 days in medium containing GM-CSF to produce cells with dendritic morphology that exhibit modest phagocytic activity, express CD11c and MHC II, and stimulate the MLR. Differentiation and activation can be further stimulated by addition of LPS, which activates DCs rendering them more immunogenic. Notably, GM-CSF is dispensable for DC development *in vivo* in the steady state; DCs generated in GM-CSF culture are more closely related to mo-DC than to the authentic cDCs in the lymphoid organs in the steady state.

Authentic cDCs can be produced *in vitro* using cultures of mouse bone marrow cells supplemented with Flt3L (Naik *et al.*, 2005). Flt3L cultures produce all of the known subsets of DCs, including cells with features of pDCs, CD8a⁺/CD103⁺ and CD11b⁺ cDCs (Naik *et al.*, 2005). Flt3L is indispensable for normal DC development *in vivo*. In mice deficient in either Flt3L or its receptor Flt3, DC development is impaired (Waskow *et al.*, 2008; McKenna *et al.*, 2000). Consistent with these observations, administration or over-expression of Flt3L results in selective expansion of DC populations, including fully differentiated DCs in lymphoid organs (Waskow *et al.*, 2008; Maraskovsky *et al.*, 1996). Moreover, human cDCs can be obtained from cultures of human cord blood supplemented with Flt3L and GM-CSF (Poulin *et al.*, 2010), although this culture does not produce human pDCs.

Cellular Origin

Inflammation stimulates monocyte differentiation to DCs

Monocytes were long thought to be the major progenitor of cDCs. However, direct transfer experiments and genetic experiments prove that most cDCs in the steady state peripheral lymphoid organs do not descend from monocytes (Jakubzick *et al.*, 2008). Monocytes only differentiate to DCs under inflammation. This was observed with injection of CFA or LPS, or infection with pathogens. Thus although monocytes can develop some of the features of DCs under conditions of inflammation *in vivo*, or when cultured with cytokines *in vitro*, but they are not precursors of cDCs in the steady state.

DC progenitors in the bone marrow

HSCs give rise to common lymphoid progenitors (CLPs) and common myeloid progenitors (CMPs). It is widely accepted that DCs descend from the Flt3 (CD135)-expressing CMPs. These CMPs then progress to a common precursor for monocytes, macrophages and DCs (macrophage-DC progenitors, MDP). MDP are defined as Lin[−] CX3CR1⁺ CD11b[−] CD115⁺ cKit⁺ CD135⁺ and account for 0.5% of all BM mononuclear cells in mice (Fogg *et al.*, 2006; Waskow *et al.*, 2008). Comparing to CMP, MDP has minimal potential to produce granulocytes, but can produce monocytes, pDCs, and cDCs. MDP then progresses to common DC progenitor (CDP). CDP, defined as Lin[−] CD115⁺ Flt3⁺ CD117^{lo}, produces cDCs and pDCs but not monocytes *in vitro* (Naik *et al.*, 2007) or *in vivo* (Onai *et al.*, 2007). Therefore the split between the monocyte and DC lineages occurs in the bone marrow between the MDP and CDP stages of development (Liu *et al.*, 2009).

Migratory pre-DCs

cDC precursors must migrate from the bone marrow to the lymphoid organs through the blood. However, MDP and CDP are restricted to the bone marrow (Liu *et al.*, 2009). A precursor with the potential to produce cDCs (pre-DC) was identified as CD11c⁺ MHC II[−] Flt3⁺ SIRPα^{lo} cells in the blood, bone marrow and periphery of mice (Del Hoyo *et al.*, 2002; Diao *et al.*, 2004; Naik *et al.*, 2006; Liu *et al.*, 2009). Pre-DCs descend from CDPs, migrate from the bone marrow to the blood and then to peripheral lymphoid organs and non-lymphoid tissues (Liu

et al., 2009). Pre-DCs have a short half-life in the blood of less than 1 h (Liu *et al.*, 2007), and comprise ~0.5% of all leukocytes in bone marrow, 0.02% in blood, 0.05% in the spleen and 0.03% in the lymph nodes respectively. The constant output of pre-DC from the bone marrow is essential for replenishment of DCs in the periphery.

In conclusion, in mice, the DC and monocyte lineages split in the bone marrow, where MDPs give rise to both monocytes and CDP; the latter produces pre-DCs, which migrate from BM through the blood to the periphery to give rise to DCs (Figure 1). Corresponding to the crucial dependence of Flt3L for DC development, expression of Flt3 is retained throughout the natural history of DC development (Karsunky *et al.*, 2003; Liu *et al.*, 2009).

Transcriptional regulation of DC development

A number of different transcription factors are known to regulate DC development. These factors include PU.1, STAT3, Gfi1, E2-2, Id2, Batf3, and RBP-J. These factors are part of a yet to be defined program that turns on expression of lineage-specific genes, and suppress alternative developmental programs, regulating differentiation at different stages or of unique DC subset. The function of each factor is reflected by change of number or subset of DC in corresponding genetic models. These models proved important in investigating the function of DC and DC subsets in various immune responses. For instance, experiments using Batf3^{−/−} mice that specifically lack CD8a⁺ and CD103⁺ DC prove the essential role of such DCs in activating CD8 T cell against tumor, virus and intracellular pathogens (reviewed in Murphy, 2013).

Human DC Development

Three different subsets of DCs have been found in human blood. These subsets are referred to as BDCA1 (CD1c), BDCA2 (CD303), and BDCA3 (CD141) DCs based on their expression of cell surface markers (reviewed in Collin *et al.*, 2011). It is widely accepted that BDCA1⁺ DCs resemble mouse CD11b⁺ cDCs; BDCA2⁺ CD11c[−] DCs are equivalent to mouse pDCs; BDCA3⁺ DCs are equivalent to mouse CD8a⁺ cDCs (Figure 2). These interspecies associations were initially based on similarities in gene expression between human and mouse DC subsets (Robbins *et al.*, 2008), and were supported by functional experiments (Croizat *et al.*, 2010; Poulin *et al.*, 2010; Lauterbach *et al.*, 2010).

The human equivalents of the bone marrow derived DC precursors, MDP, CDP, and pre-DC remain to be isolated. However, recent clinical studies revealed genetically defined syndromes associated with DC deficiency that shed some light on human DC development. Several different genetic lesions have been associated with the triad of DC deficiency, monocytopenia and opportunistic infections. GATA2 mutation leads to a loss of DC, monocytes, B and NK cells (DCML). IRF8 mutation is associated with disseminated bacille Calmette-Guérin (BCG) infection (Hambleton *et al.*, 2011). Loss of DCs due to IRF8 mutation also leads to myeloproliferation in humans due to increased serum Flt3L, which induces expansion of myeloid progenitors. Elucidation of cellular and molecular mechanisms underlying aberrant human DC

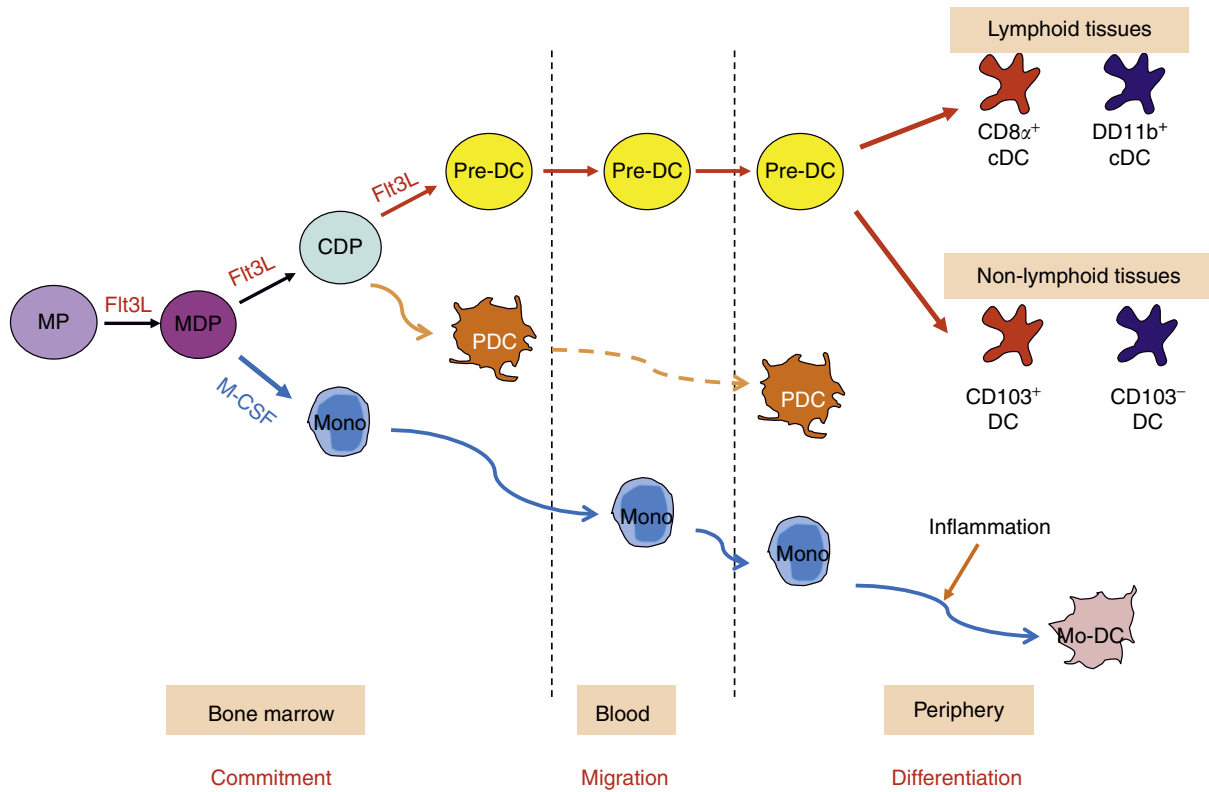


Figure 1 DC origin and development.

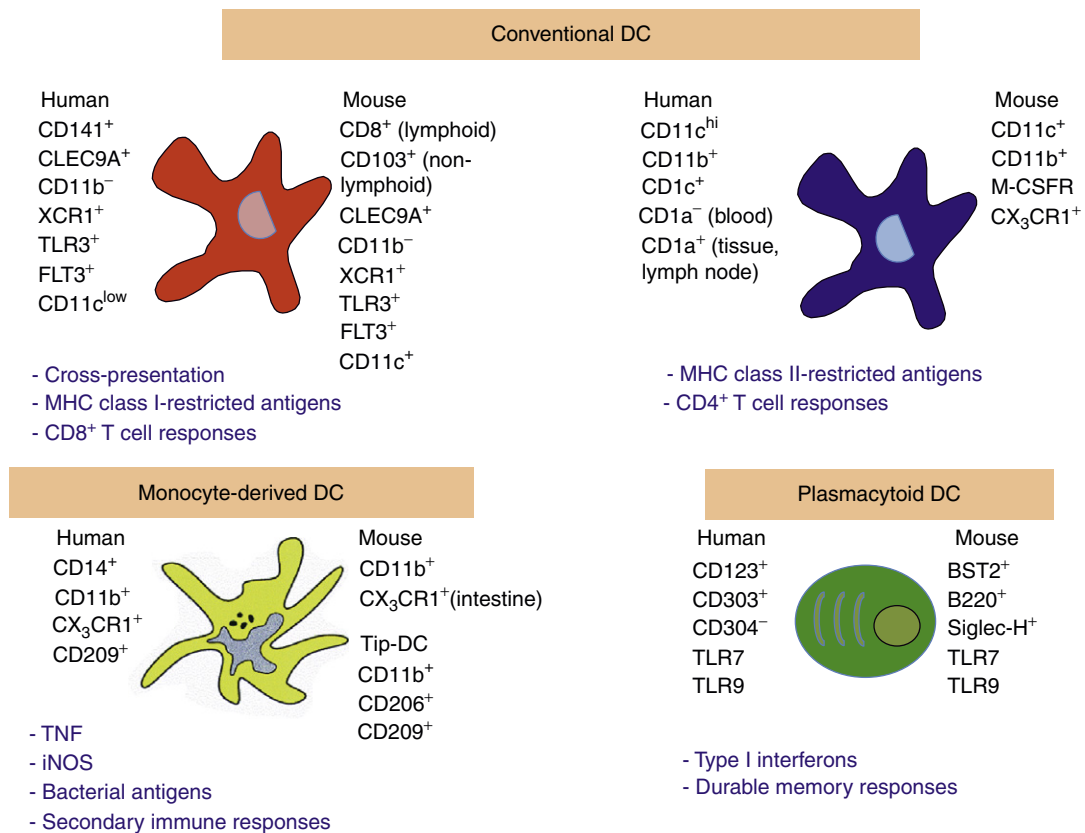


Figure 2 Functionally correlated DC subsets in human and mouse.

development will require identification of restricted DC progenitors in human (reviewed in Collin *et al.*, 2011).

DC Homeostasis and Regulation

The half-life of cDCs in mouse tissues was measured in parabionts and shown to vary from five to seven days in the spleen, lymph nodes, liver and kidney and as many as 25 days in the lung (Liu *et al.*, 2007; Ginhoux *et al.*, 2009). cDC homeostasis in all organs is maintained through a dynamic balance of three parameters: continuous input of pre-DCs from the blood, limited cDC division *in situ* and cell death. Fitting the BrdU incorporation data and the parabiosis separation data into the equations produced a numerical estimate of the rate of DC precursor input ($\sim 4300/\text{hour}$) and DC death ($\sim 9600/\text{hour}$) in the mouse spleen (Liu *et al.*, 2007).

cDCs in mouse lymphoid and non-lymphoid tissues divide *in situ* for 10–14 days before being replaced by pre-DCs (Liu *et al.*, 2007). Division of DCs is regulated by Flt3L and lymphotoxin-beta ($\text{LT}\beta$). Flt3L impacts nearly all stages of cDC development including early hematopoiesis in the bone marrow and division in the periphery (Waskow *et al.*, 2008; Maraskovsky *et al.*, 1996; Kingston *et al.*, 2009). Increased production of Flt3L, for example, during malaria infection (Guermónprez *et al.*, 2013) leads to increase of DC progenitor proliferation in the bone marrow and DC division in the periphery. In contrast, the effects of $\text{LT}\beta$ appear to be limited to $\text{CD11b}^+ \text{CD4}^+$ spleen DCs (Kabashima *et al.*, 2005).

Tregs also exert an effect on DC division. A feedback loop between DCs and Tregs maintains the physiologic numbers of these two cell types in the steady state. Loss of Tregs leads to activation of T cells, which induces DC division via Flt3L. Conversely, increasing the number of DCs by Flt3L injection leads to an increase in the number of Treg cells (Figure 3). Alterations in this mechanism lead to immune imbalance and can alter the course of autoimmune disease in mice (reviewed in Liu and Nussenzweig, 2010). Thus, the Flt3L-mediated homeostatic feedback loop between Treg and DCs has clinical implication for vaccine design, as well as the control of auto-immunity. Finally, this mechanism is entirely consistent with the proposed role of DCs in maintaining tolerance and regulating immunologic responses *in vivo* (Steinman and Nussenzweig, 2002).

Features Enable DCs as the Orchestrator of Immunity

DCs are crucial immune sentinels. First, DCs detect danger signals associated with infection, produce cytokines to alarm the immune system and activate other innate cells to confer

protection. Second, armed with unique cellular feature, DCs effectively capture, process and present antigens to initiate adaptive T cell immune response. Third, in response to the signal associated with the antigens, DCs adjust activation or maturation status, as such directing either immunity to pathogenic antigens or tolerance to harmless antigens. Thus DCs bridge innate and adaptive immune response, and control the quality of adaptive immune response toward either tolerance or immunity of different types.

DCs in Innate Immunity

Innate pDCs

One DC subset with prominent innate activity is pDC, which specializes in rapidly producing type I interferon upon detecting infection. pDCs selectively express TLR7 (single-stranded RNA sensor) and TLR9 (DNA sensor), both of which reside in the endosomal compartment. pDCs endocytose viruses, and sense their nucleic acids using TLR7 and TLR9. Viruses that enter the cytoplasm of pDCs are detected after autophagy, a conserved cell-autonomous process involving lysosomal degradation of cellular organelles to deliver cytoplasmic RNA to TLR7-containing endosomal compartments. Upon binding of ligands, TLR7 and TLR9 recruit MyD88. This results in assembly of a signal-transducing complex that includes IRAK4, TRAF6, Bruton's tyrosine kinase (BTK) and IRF7 leading to type I IFN production. In pDCs, constitutive expression of IRF7 enables a rapid production of type I IFN. Type I IFN produced from activated pDC not only directly inhibits viral replication, but also activates NK, B cells and cDCs (Blanco *et al.*, 2001). Activation of pDC and production of type I IFN is tightly regulated by feedback inhibition. Constitutive activation of pDC and overproduction of type I IFN are at the center of pathogenic inflammation in autoimmune diseases such as systemic lupus erythematosus.

Activation of innate lymphocytes

Another aspect of innate role of DCs is activation of innate lymphocytes (e.g., NK, NKT and ILCs). Besides pDCs, other DC subsets express TLRs, recognize PAMP and become activated. Different function of cDC subsets may be partly attributed to differential expression of Toll like receptors (TLRs). For example, spleen CD8a^+ cDCs in mouse and BDCA3^+ DCs in humans express TLR3 but lack TLR5 and TLR7; when stimulated with virus with double strand virus or PolyI:C, a ligand for TLR3, they produce IFN α and IL-12. CD11b^+ cDCs express TLR5 and TLR7 but low amount of TLR3 (Edwards *et al.*, 2003). DC activation leads to the recruitment of NK cells into the draining lymph nodes; activated DC also produce IL-12, IL-15 and IL-2, which in turn activate NK cells to produce IFN- γ . DCs activate NKT cells by presenting glycolipids on CD1d molecules to the invariant T cell receptor on NKT cells *in vivo* (reviewed in Munz *et al.*, 2005). Another recently defined type of innate cells that can be activated by DCs are innate lymphoid cells (ILCs), lymphocytes that lack recombined antigen-recognition receptors that are dependent on the transcription factor ID2 for their development. Interestingly, distinct DC subsets activate different subsets of ILC. CD8a^+ DCs control ILC1s, which are functionally similar to

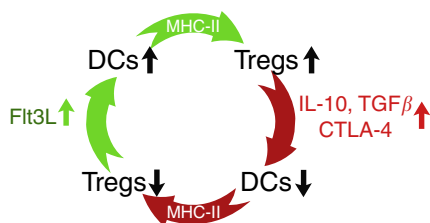


Figure 3 Homeostatic feedback loop between DCs and Treg cells.

but developmentally distinct from NK cells; CD11b⁺ DCs control ILC3s, including lymphoid tissue inducer (LTi) cells, CD4⁺ LTi-like cells, and NKp46⁺ cells, all of which are developmentally dependent on ROR γ t. Upon bacterial sensing, CD11b⁺ DCs produce IL-23, which consequently activates ILC3s to produce IL-22 (reviewed in [Briseno et al., 2014](#)).

Antigen Capture, Processing and Presentation

Antigen capture

Macropinocytosis

DCs in culture continuously form 0.25 to 1.0 μ m pinocytic vesicles. It is believed that DCs *in vivo* also use these vesicles to sample a large volume of extracellular fluid and soluble proteins that are present at low concentrations.

Receptor mediated phagocytosis

Multiple receptors on cDCs enhance recognition and ingestion of particulates including pathogens and dying cells. Some of the receptors recognize molecular patterns on pathogens and activate DCs, whereas others may lead to inhibition. For example, DCs express Fc receptors that mediate ingestion of opsonized particles and deliver them to intracellular compartments that facilitate cross-presentation to CD8 T cells. Fc receptors on DCs include both activating and inhibitory forms ([Muta et al., 1994](#)). The inhibitory receptors help DCs to maintain an immature tolerogenic state. In addition, DCs express many different types of pattern recognition receptors including the SIGN (specific ICAM-3-grabbing nonintegrin) family, and multiple C-type lectins such as macrophage mannose receptor (MMR), DEC205, and DCIR2. MMR binds a range of bacteria, yeasts, and viruses through interactions between a mannose-type carbohydrate recognition domain and pathogen-associated high mannose structures. These SIGN family receptors appear redundant, as deficiency of any single receptor confers no detectable susceptibility to pathogens.

Antigen processing

DCs and macrophages differ in their capacity to digest antigens. Macrophages endocytose antigens and rapidly digest them. In contrast, DCs sequester and preserve the captured antigen for later presentation. Preservation of antigens is critical for immunogenicity and is attributed to two features of DC lysosomes. The first is the low lysosomal protease activity. Macrophages contain high levels of lysosomal proteases, enabling rapid degradation of internalized proteins to single amino acids. In stark contrast, DCs express low amount of proteases, resulting in a limited capacity for lysosomal degradation ([Delamarre et al., 2005](#)). This low level proteolytic capacity is crucial since the peptides loaded on MHC and recognized by T cells must consist of peptides between 8 and 17 amino acids. A second feature is that DC lysosomes are less acidic than those in professional phagocytes. DC lysosome is featured with an incomplete assembly of V-ATPase that results in proton 'leakage,' and an efficient recruitment of Nox2 in lysosome that results in consumption of protons. Altogether, these lead to alkalinization of the endocytic compartment ([Trombetta et al., 2003](#)). Therefore, the lower levels of proteolytic activity, and decreased acidity in endocytic

compartment lead to a decrease in the rate of antigen digestion and increased availability of partially processed peptides for loading on MHC. This unique feature of DCs may also help preserve the antigen during the migration of non-lymphoid tissue DCs from the site of antigen capture to the lymph nodes.

Antigen presentation

DCs initiate T cell immune response in the lymph nodes and spleen. DC-T cell interactions have now been studied in the living state with two-photon microscopy ([Shakhar et al., 2005](#)). DCs present antigen peptides on MHC I and MHC II and scan the TCRs on T cells that circulate through entering the lymph nodes via a brief interaction. When a TCR finds the cognate MHC-peptide, the TCR-bearing antigen specific T cells arrest on antigen presenting DCs, and this stable interaction lasts for at least 18 h. Such DC-T cell interaction leads to tolerance in the steady state, unless an activation signal such as TLR ligation triggers DCs to initiate immunity.

In conventional presentation pathway, exogenous antigens are presented on MHC II to CD4⁺ T cells, whereas endogenous antigens are presented on MHC I to CD8⁺ T cells. CD8⁺ T cells are critical for protective immunity against intracellular pathogens and malignant tumor cells. However, viruses that do not infect DCs and tumors are exogenous antigens to DCs. Importantly, DCs are efficient in presenting exogenous antigen on MHC I, an unconventional presentation path called 'cross-presentation'. After taking up the exogenous viral or tumor antigens, in the form of non-replicative virus or apoptotic cells, DCs present them on MHC I to prime CD8 T cells ([Albert et al., 1998](#)). In the absence DCs, mice are unable to process several different antigens through the exogenous pathway, indicating that DCs are the major cell type for cross-presentation to CD8⁺ T cells *in vivo*. Multiple receptors including FcR, Lox-1, CD205 and DNGR1 have been shown to mediate antigen uptake for cross-presentation. Batf3-dependent CD8a⁺ DCs in mice and their functional equivalent BDCA3⁺ DCs in humans excel in cross-presenting exogenous antigens ([Hildner et al., 2008](#); [Poulin et al., 2010](#); [Schreibelt et al., 2012](#)). Exogenous presentation or cross-presentation by CD8a⁺ DCs has proved essential for protective immunity against viruses and tumors.

DC Maturation

DCs link innate pathogen recognition to the adaptive immune response and direct immune response toward tolerance or immunity. This is made possible by the two functionally distinct and phenotypically different stages of DCs. In the steady state, DCs in most tissues are equipped to capture and present antigens to T cells, but the outcome of antigen presentation by steady state DCs is tolerance and not immunity. Steady state DCs express high levels of pattern recognition and activation receptors allowing them to sense changes in the environment, including pathogens and inflammatory cytokines. For example, PAMP and DAMP associated with microbes and viruses induce DC activation by engaging Toll-like receptors and cytoplasmic receptors that recognize pathogen patterns such as MDA-5, RIG-I and DDX41, and NOD-like receptors (NLRs). Activated NK,

NKT, and T cells stimulate DCs through direct cellular contact by ligation of CD40 on DCs. Type I interferon, produced by pDCs or non-hematopoietic cells upon invasion of microbes and virus, also effectively activate DCs. These signals, largely through the NF- κ B pathway, induce extensive differentiation of DC to a mature state, a state characterized by redistribution of MHC class II from intracellular compartments to the plasma membrane, increased levels of cell surface MHC and co-stimulatory molecules such as CD40, CD80, and CD86, as well as secretion of cytokines, such as IL-1 α , IL-6, TNF- α , IL-18 and IL-12, and chemokines. In the mature state, DCs initiate potent and specifically polarized T cell immune responses.

Controlling the Quality of the T Cell Response

DCs are involved in critical T cell fate decisions. In the presence of mature DCs producing IL-12 or interferons (as might occur when DCs are ligated by CD40L or infected with viruses), CD4⁺ T cells differentiate along a Th1 pathway for interferon- γ production. The latter in turn activates the antimicrobial activity of macrophages and promotes killer T cell differentiation. In the presence of exogenous IL-4, however, DCs induce T cells to differentiate into Th2 cells, which secrete IL-4, IL-5 and IL-13. These cytokines help B cells to make antibodies of the IgG1 and IgE isotypes, activate eosinophils, and stimulate fibrosis. A new and striking pathway that was first discovered with human monocyte-derived DCs involves the epithelial derived cytokine, thymic stromal lymphopoietin (TSLP). This cytokine activates DCs to induce “inflammatory Th2 cells” which produce TNF α (rather than IL-10) in addition to IL-4, IL-5, and IL-13 (Soumelis *et al.*, 2002). DCs that are activated with either CD40L or TSLP are similar in appearance, being rich in MHC II and CD86 stimulatory molecules. However, they differ significantly in cytokine and chemokine production, and as mentioned, the functional consequences for T cells vary (Soumelis *et al.*, 2002).

DCs' Role in Tolerance

The fact that DCs can focus the immune response on antigens derived from the pathogen and avoid inducing immunity to self and environmental antigens critically depends on the immune tolerance to self and harmless antigens. DCs have a critical role in developing and maintaining immune tolerance.

Thymic DC contribute to central tolerance

Self-reactive thymocytes are deleted in the thymus by APCs during negative selection. Two major populations of APCs in thymus express MHC II, namely, medullary thymic epithelial cells (mTEC) and cDCs, and both are required for efficient negative selection. mTEC express a panoply of self-antigens under the control of the autoimmune regulator ‘AIRE’ (Anderson *et al.*, 2002). DCs capture self-antigens that enter the thymus through the blood stream and present them to self-reactive T cells to induce negative selection. In the absence of antigen presentation by DCs, negative selection of CD4⁺ thymocytes is impaired (Brockner *et al.*, 1997). In addition to negative selection, thymic cDCs also support the development of FoxP3⁺ Tregs (Proietto *et al.*, 2008). Thus, DCs contribute to central tolerance in the thymus by more than one mechanism.

DCs mediate peripheral tolerance

DCs also mediate tolerance in the periphery. Central tolerance is incomplete and in addition, the immune system must continually establish tolerance to harmless or ‘noninfectious’ antigen in the environment. The effective control of self-reactive T cells that have escaped central tolerance therefore depends on peripheral tolerance. DCs constantly carry innocuous antigens from the periphery, for example, from the skin, airways, stomach, intestine and pancreas and present them to T cells in lymphoid organs (reviewed in Steinman *et al.*, 2003b). In the steady state, antigen presented by immature DCs induce profound tolerance of T cells, even though these T cells can initially proliferate extensively to the antigen capturing DCs in the draining lymph nodes (Steinman *et al.*, 2003a). Mechanisms for peripheral tolerance can be intrinsic (deletion and anergy) or extrinsic (through regulatory T cells). B7 family members on the steady state DCs, for example, PD-L1 and CD86, can ligate the inhibitory molecules such as PD-1 and CTLA-4 on the T cells, leading to tolerance.

Controlling of regulatory T cells

Autoreactive T cells can remain quiescent in the presence of regulatory T cells. There are two types of Treg: naturally occurring Treg derived from thymus (natural Treg) and Treg induced from FoxP3⁺ CD4⁺ T cells in the periphery (induced Treg). DCs are able to employ environmental signals, vitamin A in the gut and vitamin D3 in the skin, to induce Treg and maintain tolerance to harmless foreign antigens. CD103⁺ DC subset in the gut and skin are particularly potent in inducing Tregs (Belkaid and Oldenhove, 2008). Activation of β -catenin is essential for DC to control Treg and peripheral tolerance.

DCs also maintain the homeostasis of Treg. Loss of DCs leads to a loss of Treg cells, and the remaining Treg cells exhibit decreased Foxp3 expression. Conversely, increasing the number of DCs leads to increased Treg cell division and accumulation by a mechanism that requires major histocompatibility complex II expression on DCs (Liu and Nussenzweig, 2010).

In sum, DCs have the capacity to induce and maintain tolerance by several mechanisms both in the central and peripheral lymphoid organs.

DCs in Clinical Immunology

Based on DC's role in immune response, it is not surprising that DCs play a pathogenic role in many diseases and become target in disease prevention and treatment.

DCs in Transplantation

In hematopoietic transplantation, recipient DCs initiate T-cell-induced graft versus host reactions. In organ transplantation, both donor and recipient DCs contribute to graft rejection.

DCs in Autoimmune Disease

Several human autoimmune diseases involve DCs. In rheumatoid arthritis, DCs in the synovial exudates produce TNF- α , contributing to the severity of the disease. In psoriasis,

lesional skin is populated with activated pDCs that produce type I IFN and mo-DCs that produce TNF- α and polarize T cells toward Th1/Th17 (Lowes *et al.*, 2005). In systemic lupus erythematosus (SLE) two subsets of DCs contribute to the onset and severity of the disease; pDCs in SLE patients overproduce IFN- α , which activates cDCs and interferes with their ability to maintain peripheral tolerance (Blanco *et al.*, 2001).

DCs in Viral Infections

DCs mediate antiviral immunity by priming T cell responses. However, a number of viruses have evolved strategies to subvert DCs and thereby the immune system. DCs carry HIV from peripheral tissues into draining lymph nodes, where the virus is transmitted to CD4 T cells. Transmission to CD4 T cells is dependent on DC-SIGN, a C-type lectin pathogen-recognition receptor expressed on the surface of DCs that retains the attached virus in an infectious state. Interestingly, DC-SIGN also serves as receptor for several other viruses, including hepatitis C virus, Ebola virus, cytomegalovirus, dengue virus, and the SARS coronavirus. Dengue virus targets DCs directly through DC-SIGN but also enters DCs as a passenger in the form of immune complexes that are taken up by Fc receptors. When infection occurs through antibody enhancement mediated by Fc receptor, the infected DCs are involved in induction of the T cell cytokines that mediate the vascular leak syndrome associated with the infection, causing hemorrhagic fever.

DCs in Cancer

Tumors can suppress immunity in part through their effects on DCs. DC differentiation and activation can be suppressed by cancer-derived cytokines such as IL-6, vascular endothelial growth factor, and IL-10. In contrast to their normal counterparts, which activate immune responses, DCs derived from tumors induce Tregs and suppress proliferation of CTL and natural killer T cells (NKT cells).

DC Targeted Vaccines

DC-based vaccines are currently being used in the clinic and DC targeted vaccines are being tested. In currently available DC-based immune therapy, monocyte-derived DCs are generated *ex vivo*, loaded with tumor cells or tumor antigens, and re-injected into the patient (Schuler-Thurner *et al.*, 2002). Scientific and practical problems exist with this approach, including limited responses possibly due to inefficient migration of monocyte-derived DCs from injection site to the draining lymphoid organs and inefficient antigen presentation.

DC targeted vaccines are based on the idea that antigens delivered specifically to DCs in conjunction with the appropriate adjuvants will produce strong and lasting immunity. The DC targeted vaccines require that antigens be delivered specifically to endocytic receptors on DCs, along with the appropriate stimuli to induce DC activation. For example, antigens have been incorporated into anti-receptor monoclonal antibodies, which are then injected into the vaccine recipient. In animal models, such *in vivo* DC-targeting has been shown to elicit immune responses that are broad, often generating

immunity against multiple epitopes. Potential DC targets include CD205, LOX-1/OLR1, MMR/CD206, DCIR/CLEC4A, DC-SIGN/CD209, DNGR1, langerin, and CD40 (reviewed in Steinman, 2011).

In summary, DCs play important roles in a number of different diseases. Moreover, they are excellent targets in designing new approaches to prevention and treatment of these diseases.

See also: Cellular Immunology: Innate Immunity: Scavenger Receptors

References

- Albert, M.L., Sauter, B., Bhardwaj, N., 1998. Dendritic cells acquire antigen from apoptotic cells and induce class I-restricted CTLs. *Nature* 392, 86–89.
- Anderson, M.S., Venanzi, E.S., Klein, L., *et al.*, 2002. Projection of an immunological self shadow within the thymus by the aire protein. *Science* 298, 1395–1401.
- Ardavin, C., Shortman, K., 1992. Cell surface marker analysis of mouse thymic dendritic cells. *European Journal of Immunology* 22, 859–862.
- Banchereau, J., Steinman, R.M., 1998. Dendritic cells and the control of immunity. *Nature* 392, 245–252.
- Belkaid, Y., Oldenhove, G., 2008. Tuning microenvironments: Induction of regulatory T cells by dendritic cells. *Immunity* 29, 362–371.
- Blanco, P., Palucka, A.K., Gill, M., Pascual, V., Banchereau, J., 2001. Induction of dendritic cell differentiation by IFN- α in systemic lupus erythematosus. *Science* 294, 1540–1543.
- Briseno, C.G., Murphy, T.L., Murphy, K.M., 2014. Complementary diversification of dendritic cells and innate lymphoid cells. *Current Opinion in Immunology* 29, 69–78.
- Brockner, T., Riedinger, M., Karjalainen, K., 1997. Targeted expression of major histocompatibility complex (MHC) class II molecules demonstrates that dendritic cells can induce negative but not positive selection of thymocytes *in vivo*. *Journal of Experimental Medicine* 185, 541–550.
- Collin, M., Bigley, V., Haniffa, M., Hambleton, S., 2011. Human dendritic cell deficiency: The missing ID? *Nature Reviews Immunology* 11, 575–583.
- Crozat, K., Guiton, R., Contreras, V., *et al.*, 2010. The XC chemokine receptor 1 is a conserved selective marker of mammalian cells homologous to mouse CD8 α + dendritic cells. *Journal of Experimental Medicine* 207, 1283–1292.
- Delamarre, L., Pack, M., Chang, H., Mellman, I., Trombetta, E.S., 2005. Differential lysosomal proteolysis in antigen-presenting cells determines antigen fate. *Science* 307, 1630–1634.
- Del Hoyo, G.M., Martin, P., Vargas, H.H., *et al.*, 2002. Characterization of a common precursor population for dendritic cells. *Nature* 415, 1043–1047.
- Diao, J., Winter, E., Chen, W., Cantin, C., Catral, M.S., 2004. Characterization of distinct conventional and plasmacytoid dendritic cell-committed precursors in murine bone marrow. *Journal of Immunology* 173, 1826–1833.
- Edwards, A.D., Diebold, S.S., Slack, E.M., *et al.*, 2003. Toll-like receptor expression in murine DC subsets: lack of TLR7 expression by CD8 α + DC correlates with unresponsiveness to imidazoquinolines. *European Journal of Immunology* 33, 827–833.
- Ehrenreich, B.A., Cohn, Z.A., 1967. The uptake and digestion of iodinated human serum albumin by macrophages *in vitro*. *Journal of Experimental Medicine* 126, 941–958.
- Ehrenreich, B.A., Cohn, Z.A., 1969. The fate of peptides pinocytosed by macrophages *in vitro*. *Journal of Experimental Medicine* 129, 227–245.
- Fogg, D.K., Sibon, C., Miled, C., *et al.*, 2006. A clonogenic bone marrow progenitor specific for macrophages and dendritic cells. *Science* 311, 83–87.
- Ginhoux, F., Liu, K., Helft, J., *et al.*, 2009. The origin and development of nonlymphoid tissue CD103 + DCs. *Journal of Experimental Medicine* 206, 3115–3130.
- Guermontprez, P., Helft, J., Claser, C., *et al.*, 2013. Inflammatory Fc γ 3 is essential to mobilize dendritic cells and for T cell responses during *Plasmodium* infection. *Nature Medicine* 19, 730–738.
- Hambleton, S., Salem, S., Bustamante, J., *et al.*, 2011. IRF8 mutations and human dendritic-cell immunodeficiency. *New England Journal of Medicine* 365, 127–138.

- Helft, J., Ginhoux, F., Bogunovic, M., Merad, M., 2010. Origin and functional heterogeneity of non-lymphoid tissue dendritic cells in mice. *Immunological Reviews* 234, 55–75.
- Hildner, K., Edelson, B.T., Purtha, W.E., *et al.*, 2008. Batf3 deficiency reveals a critical role for CD8alpha+ dendritic cells in cytotoxic T cell immunity. *Science* 322, 1097–1100.
- Jakubzick, C., Bogunovic, M., Bonito, A.J., *et al.*, 2008. Lymph-migrating, tissue-derived dendritic cells are minor constituents within steady-state lymph nodes. *Journal of Experimental Medicine* 205, 2839–2850.
- Kabashima, K., Banks, T.A., Ansel, K.M., *et al.*, 2005. Intrinsic lymphotoxin-beta receptor requirement for homeostasis of lymphoid tissue dendritic cells. *Immunity* 22, 439–450.
- Karsunky, H., Merad, M., Cozzio, A., Weissman, I.L., Manz, M.G., 2003. Flt3 ligand regulates dendritic cell development from Flt3+ lymphoid and myeloid-committed progenitors to Flt3+ dendritic cells *in vivo*. *Journal of Experimental Medicine* 198, 305–313.
- Kingston, D., Schmid, M.A., Onai, N., *et al.*, 2009. The concerted action of GM-CSF and Flt3-ligand on *in vivo* dendritic cell homeostasis. *Blood* 114, 835–843.
- Lauterbach, H., Bathke, B., Gilles, S., *et al.*, 2010. Mouse CD8alpha+ DCs and human BDCA3+ DCs are major producers of IFN-lambda in response to poly IC. *Journal of Experimental Medicine* 207, 2703–2717.
- Liu, K., Nussenzweig, M.C., 2010. Development and homeostasis of dendritic cells. *European Journal of Immunology* 40, 2099–2102.
- Liu, K., Vitoria, G.D., Schwickert, T.A., *et al.*, 2009. *In vivo* analysis of dendritic cell development and homeostasis. *Science* 324, 392–397.
- Liu, K., Waskow, C., Liu, X., *et al.*, 2007. Origin of dendritic cells in peripheral lymphoid organs of mice. *Nature Immunology* 8, 578–583.
- Lowes, M.A., Chamian, F., Abello, M.V., *et al.*, 2005. Increase in TNF-alpha and inducible nitric oxide synthase-expressing dendritic cells in psoriasis and reduction with efalizumab (anti-CD11a). *Proceedings of the National Academy of Sciences of the United States of America* 102, 19057–19062.
- Maraskovsky, E., Brasel, K., Teepe, M., *et al.*, 1996. Dramatic increase in the numbers of functionally mature dendritic cells in Flt3 ligand-treated mice: Multiple dendritic cell subpopulations identified. *Journal of Experimental Medicine* 184, 1953–1962.
- McKenna, H.J., Stocking, K.L., Miller, R.E., *et al.*, 2000. Mice lacking flt3 ligand have deficient hematopoiesis affecting hematopoietic progenitor cells, dendritic cells, and natural killer cells. *Blood* 95, 3489–3497.
- Mishell, R.J., Dutton, R.W., 1967. Immunization of dissociated spleen cell cultures from normal mice. *Journal of Experimental Medicine* 126, 423–442.
- Munz, C., Steinman, R.M., Fujii, S., 2005. Dendritic cell maturation by innate lymphocytes: coordinated stimulation of innate and adaptive immunity. *Journal of Experimental Medicine* 202, 203–207.
- Murphy, K.M., 2013. Transcriptional control of dendritic cell development. *Advances in Immunology* 120, 239–267.
- Muta, T., Kurosaki, T., Misulovin, Z., *et al.*, 1994. A 13-amino-acid motif in the cytoplasmic domain of Fc gamma R1b modulates B-cell receptor signalling. *Nature* 368 (6466), 70–73.
- Naik, S.H., Metcalf, D., Van Nieuwenhuijze, A., *et al.*, 2006. Intrasplenic steady-state dendritic cell precursors that are distinct from monocytes. *Nature Immunology* 7, 663–671.
- Naik, S.H., Proietto, A.I., Wilson, N.S., *et al.*, 2005. Cutting edge: Generation of splenic CD8+ and CD8- dendritic cell equivalents in Fms-like tyrosine kinase 3 ligand bone marrow cultures. *Journal of Immunology* 174, 6592–6597.
- Naik, S.H., Sathe, P., Park, H.Y., *et al.*, 2007. Development of plasmacytoid and conventional dendritic cell subtypes from single precursor cells derived *in vitro* and *in vivo*. *Nature Immunology* 8, 1217–1226.
- Nussenzweig, M.C., Steinman, R.M., 1980. Contribution of dendritic cells to stimulation of the murine syngeneic mixed leukocyte reaction. *Journal of Experimental Medicine* 151, 1196–1212.
- Nussenzweig, M.C., Steinman, R.M., Gutschinov, B., Cohn, Z.A., 1980. Dendritic cells are accessory cells for the development of anti-trinitrophenyl cytotoxic T lymphocytes. *Journal of Experimental Medicine* 152, 1070–1084.
- Onai, N., Obata-Onai, A., Schmid, M.A., *et al.*, 2007. Identification of clonogenic common Flt3+M-CSFR+ plasmacytoid and conventional dendritic cell progenitors in mouse bone marrow. *Nature Immunology* 8, 1207–1216.
- Poulin, L.F., Salio, M., Griessinger, E., *et al.*, 2010. Characterization of human DNGR-1+ BDCA3+ leukocytes as putative equivalents of mouse CD8alpha+ dendritic cells. *Journal of Experimental Medicine* 207, 1261–1271.
- Proietto, A.I., Van Dommelen, S., Zhou, P., *et al.*, 2008. Dendritic cells in the thymus contribute to T-regulatory cell induction. *Proceedings of the National Academy of Sciences of the United States of America* 105, 19869–19874.
- Randolph, G.J., Angeli, V., Swartz, M.A., 2005. Dendritic-cell trafficking to lymph nodes through lymphatic vessels. *Nature Reviews Immunology* 5, 617–628.
- Robbins, S.H., Walzer, T., Dembele, D., *et al.*, 2008. Novel insights into the relationships between dendritic cell subsets in human and mouse revealed by genome-wide expression profiling. *Genome Biology* 9, R17.
- Schreibelt, G., Klinkenberg, L.J., Cruz, L.J., *et al.*, 2012. The C-type lectin receptor CLEC9A mediates antigen uptake and (cross-)presentation by human blood BDCA3+ myeloid dendritic cells. *Blood* 119, 2284–2292.
- Schuler-Thurner, B., Schultz, E.S., Berger, T.G., *et al.*, 2002. Rapid induction of tumor-specific type 1T helper cells in metastatic melanoma patients by vaccination with mature, cryopreserved, peptide-loaded monocyte-derived dendritic cells. *Journal of Experimental Medicine* 195, 1279–1288.
- Shakhar, G., Lindquist, R.L., Skokos, D., *et al.*, 2005. Stable T cell-dendritic cell interactions precede the development of both tolerance and immunity *in vivo*. *Nature Immunology* 6, 707–714.
- Shortman, K., Liu, Y.J., 2002. Mouse and human dendritic cell subtypes. *Nature Reviews Immunology* 2, 151–161.
- Soumelis, V., Reche, P.A., Kanzler, H., *et al.*, 2002. Human epithelial cells trigger dendritic cell mediated allergic inflammation by producing TSLP. *Nature Immunology* 3, 673–680.
- Steinman, R.M., 2011. Decisions about dendritic cells: Past, present, and future. *Annual Review of Immunology* 30, 1–22.
- Steinman, R.M., Cohn, Z.A., 1972. The interaction of soluble horseradish peroxidase with mouse peritoneal macrophages *in vitro*. *Journal of Cell Biology* 55, 186–204.
- Steinman, R.M., Cohn, Z.A., 1973. Identification of a novel cell type in peripheral lymphoid organs of mice. I. Morphology, quantitation, tissue distribution. *Journal of Experimental Medicine* 137, 1142–1162.
- Steinman, R.M., Cohn, Z.A., 1974. Identification of a novel cell type in peripheral lymphoid organs of mice. II. Functional properties *in vitro*. *Journal of Experimental Medicine* 139, 380–397.
- Steinman, R.M., Hawiger, D., Liu, K., *et al.*, 2003a. Dendritic cell function *in vivo* during the steady state: A role in peripheral tolerance. *Annals of the New York Academy of Sciences* 987, 15–25.
- Steinman, R.M., Hawiger, D., Nussenzweig, M.C., 2003b. Tolerogenic dendritic cells. *Annual Review of Immunology* 21, 685–711.
- Steinman, R.M., Nussenzweig, M.C., 2002. Avoiding horror autotoxicus: The importance of dendritic cells in peripheral T cell tolerance. *Proceedings of the National Academy of Sciences of the United States of America* 99, 351–358.
- Steinman, R.M., Witmer, M.D., 1978. Lymphoid dendritic cells are potent stimulators of the primary mixed leukocyte reaction in mice. *Proceedings of the National Academy of Sciences of the United States of America* 75, 5132–5136.
- Trombetta, E.S., Ebersold, M., Garrett, W., Pypaert, M., Mellman, I., 2003. Activation of lysosomal function during dendritic cell maturation. *Science* 299, 1400–1403.
- Unanue, E.R., Cerottini, J.C., 1970. The immunogenicity of antigen bound to the plasma membrane of macrophages. *Journal of Experimental Medicine* 131, 711–725.
- Vremec, D., Zorbas, M., Scollay, R., *et al.*, 1992. The surface phenotype of dendritic cells purified from mouse thymus and spleen: Investigation of the CD8 expression by a subpopulation of dendritic cells. *Journal of Experimental Medicine* 176, 47–58.
- Waskow, C., Liu, K., Darrasse-Jeze, G., *et al.*, 2008. The receptor tyrosine kinase Flt3 is required for dendritic cell development in peripheral lymphoid tissues. *Nature Immunology* 9, 676–683.

Introduction

Neutrophils are sentinel cells and are the first to arrive at sites of infection to defend the host from invading microbes. The bone marrow is the site of neutrophil production. Under conditions of homeostasis, neutrophils can be found in the bone marrow, lung, spleen, and liver. Typically, neutrophils have a very short half-life and therefore are produced at a rate of $5\text{--}10 \times 10^{10}$ cells day^{-1} to maintain steady-state. Neutrophil production is controlled by the synthesis of granulocyte-colony stimulating factor (G-CSF) in response to Interleukin-17A. During inflammation, the longevity of neutrophils increases to ensure their presence at sites of damage. Neutrophils that die by apoptosis at sites of inflammation are removed by phagocytic cells such as macrophages and dendritic cells. Neutrophils are non-mitotic cells and display a multilobulated nucleus, hence the alternative name of polymorphonuclear cells (PMNs) (Figure 1). In circulation, neutrophils remain in a resting state, which ensures their potent toxic arsenals are not accidentally released, damaging the host. They are recruited from the circulation to the tissue in response to microbes or other inflammatory stimuli. In tissue, they employ different mechanisms to damage pathogens. As such, significant decreases in neutrophils can lead to increased risk for severe infections (Klein, 2011), while tight regulation of these cells is crucial to prevent tissue injury.

This article summarizes important components of neutrophil biology, including their development and microbicidal mechanisms and their potential pathogenic role.

Neutrophil Development

Neutrophils originate in the bone marrow and enter the circulation as terminally differentiated cells. Various complex mechanisms and mediators are involved in the development of neutrophil precursors and in their terminal differentiation. Among them are cytokines, growth factors, microRNAs, and transcriptional regulation. Balance between the transcription factors C/EBP α (CCAAT/enhancer binding protein α) and PU.1 determine granulocytopoiesis commitment (Reddy *et al.*, 2002; Dahl *et al.*, 2003; Laslo *et al.*, 2006). C/EBP α regulates granulocytic differentiation by competing with the transcription factor NFI-A to modulate miR223 expression (Fazi *et al.*, 2005). Other important factors include Gfi-1 (growth factor independent-1) and C/EBP ϵ (Borregaard, 2010). Gfi-1 represses genes involved in proliferation of stem cells and monocyte-promoting transcription factors (Hock *et al.*, 2004; Zarebski *et al.*, 2008). Neutrophils arise from hematopoietic stem cells (HSC) and follow a regulated differentiation process as follows:

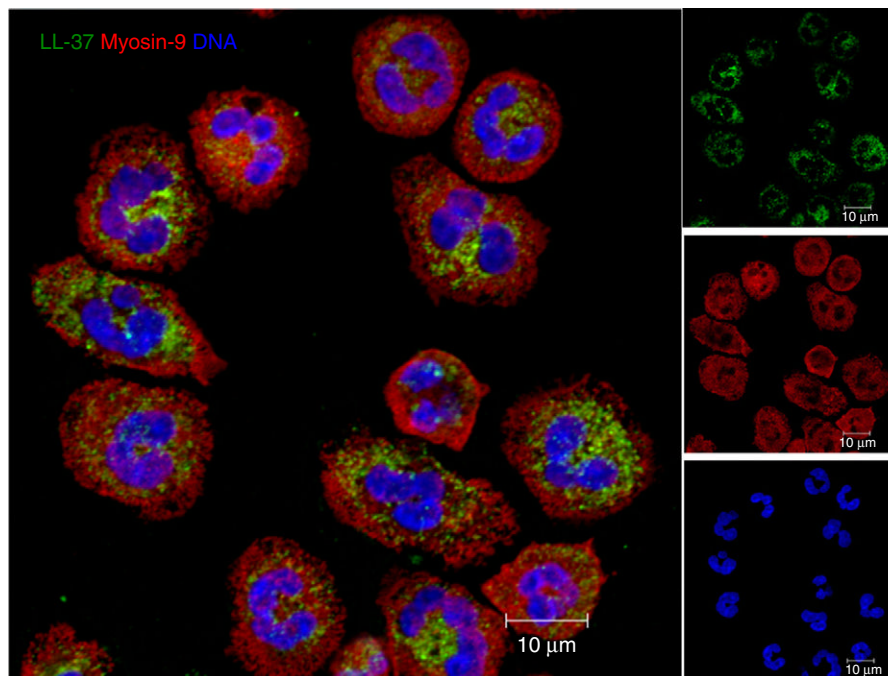


Figure 1 Neutrophil morphology. Neutrophils have distinct multilobulated nucleus with multiple granules covering their cytoplasm. Neutrophils isolated from a healthy volunteer express the granular protein LL-37 and cytosolic myosin-9. DNA was co-stained using Hoechst. Scale bar 10 μm .

Myeloblast

Myeloblasts have a diameter of 10–20 μm , are derived from HSCs and are normally found in the bone marrow. Morphologically, they are characterized by a large round to oval nucleus with a small basophilic cytoplasm with no evident granules. During the myeloblast stage, levels of C/EBP α gradually decrease (Bjerregaard *et al.*, 2003). The appearance of granules marks the transition from myeloblast to promyelocyte and C/EBP ϵ regulates the expression of granular proteins (Yamanaka *et al.*, 1997).

Promyelocyte

The morphology of the promyelocyte is similar to that of the myeloblast. Cells are larger and nuclei are round to oval with presence of nucleoli and diffuse chromatin distribution. The endoplasmic reticulum (ER) is more prominent and azurophilic granules (primary granules) accumulate in increasing numbers during this stage. These cells express CD33, CD13, and CD15 on cell surface and they differentiate into myelocytes.

Myelocyte

During this stage, secondary granules become evident in the cytoplasm. Nuclei are round to oval and eccentric, chromatin is coarse, and nucleoli are small. Formation of primary granules decreases, while secondary granules are small and numerous. CD14 and CD11b become present as surface markers during this stage.

Metamyelocyte

Metamyelocytes have a 'horseshoe-shape' nucleus without nucleoli. Chromatin is denser and cytoplasm is filled with primary, secondary, and tertiary granules (Quesenberry and Levitt, 1979; Lajtha *et al.*, 1953). ER becomes thinly dispersed.

Band (Juvenile)

Bands are characterized by further condensation of the chromatin, with a 'sausage-like' or band conformation. The nucleus is divided into two or more lobules connected by filamentous strands of heterochromatin. CD10 and CD16 markers distinguish bands from more immature forms.

Mature Neutrophils

The nucleus becomes more segmented (three to four lobules) indicating full maturation of the neutrophil, which exits the bone marrow and enters the circulation (Bainton *et al.*, 1971).

Granules are sequentially and continuously produced during granulocytic differentiation through vesicle budding from the Golgi apparatus. These compartments are classified into three distinct subsets:

1. Primary (azurophilic) granules appear during the promyelocyte stage. They are also known as peroxidase-positive granules and measure around 0.3 μm in diameter. Some

of the cargo proteins contained in these granules include defensins, bactericidal/permeability-increasing protein (BPI), neutrophil elastase, myeloperoxidase (MPO), proteinase 3 (PR3), and cathepsin G (Faurschou and Borregaard, 2003).

2. Secondary (specific) granules are synthesized during the myelocyte stage and have an approximate diameter of 0.1 μm . They are characterized by the presence of the glycoprotein lactoferrin and contain antimicrobial proteins such as neutrophil gelatinase-associated lipocalin (NGAL), human cationic antimicrobial protein (hCAP-18), and lysozyme (Faurschou and Borregaard, 2003).
3. Tertiary (gelatinase) granules appear during the band stage (Borregaard *et al.*, 1995; Borregaard and Cowland, 1997). They are smaller than specific granules and contain gelatinases such as leukolysin (matrix metalloproteinase (MMP)-25), MMP-8, and MMP-9.

Neutrophils also contain secretory vesicles that are synthesized during the last stage of maturation. Protein content is dictated by the transcriptional activity during the granule formation or by 'targeting by timing of biosynthesis' (Le Cabec *et al.*, 1996; Borregaard, 2010).

Neutrophil Migration into Sites of Inflammation

Once mature, neutrophils are released from the bone marrow in a tightly chemokine-regulated fashion (Geering and Simon, 2011). Chemokine receptors CXCR4 and CXCR2 play a critical role during neutrophil release from the bone marrow. Retention of mature neutrophils in the bone marrow is associated with mutations in CXCR4 or deletion of CXCR2 (Hernandez *et al.*, 2003; Eash *et al.*, 2009). Conversely, deletion of CXCR4 receptor causes increased release of mature neutrophils to the circulation (Eash *et al.*, 2009). Neutrophils enter the circulation in a resting state and can become primed by danger signals including bacterial products, cytokines and chemokines. Primed neutrophils are mobilized to the sites of infection or inflammation where they transmigrate into the tissues through interactions with the vascular endothelium. This process involves three steps:

Tethering and Rolling

Upregulation of endothelial cell adhesion molecules, such as selectins, initiates neutrophil tethering. This upregulation occurs through mobilization of P-selectins from Weibel–Palade bodies and by upregulation of *de novo* synthesis of E-selectins. Selectins are translocated to the apical membrane of the endothelial cell, where they transiently bind their respective ligands P-selectin glycoprotein (PSGL-1) and E-selectin ligand-1 (ESL-1), expressed by activated neutrophils. These interactions activate intracellular cascades that include Src family kinases Hck, Fgr, and Lyn that phosphorylate cytosolic proteins and trigger activation of syk (spleen tyrosine kinase). Activated syk triggers cytoskeletal rearrangement in the neutrophil leading to activation of kinases that result in integrin activation (Yago *et al.*, 2010; Ley *et al.*, 2007; Woodfin *et al.*, 2010; Barreiro and Sanchez-Madrid, 2009).

Firm Adhesion

Neutrophil rolling along the endothelium induces conformational changes in integrin adhesion of very late antigen-4 (VLA-4; $\alpha_4\beta_1$ -integrin, CD49d/CD29b), lymphocyte function-associated antigen-1 (LFA-1; $\alpha_1\beta_2$ -integrin, CD11a/CD18), and macrophage antigen-1 (MAC-1, $\alpha_M\beta_2$ -integrin, CD11b/CD18) that bind to adhesion molecules on the endothelial membrane such as intercellular adhesion molecule (ICAM)-1 and -2, vascular cell adhesion molecule-1 (VCAM-1), and mucosal vascular cell adhesion molecule-1 (MADCAM-1). LFA-1 is the key molecule that mediates the transition from rolling to adhesion (Phillipson *et al.*, 2006). These interactions trigger strong adhesion of the neutrophil to the endothelium, allowing for transmigration.

Transmigration

Neutrophils transmigrate into the affected tissues through the junctions between endothelial cells (paracellular transmigration) assisted by surface ligands ICAM-2, platelet endothelial cell adhesion molecule-1 (PECAM-1), and proteins of the junctional adhesion molecule (JAM) family (Woodfin *et al.*, 2009). Endothelial PECAM-1 interacts with GPI-CD177 (NB1-antigen) to facilitate transmigration (Sachs *et al.*, 2007). Transcellular transmigration accounts for approximately 20% of the transmigration mechanisms and occurs when neutrophils penetrate individual endothelial cells under conditions of high ICAM-1 expression and density (Yang *et al.*, 2006). SNARE proteins are essential in trafficking and fusion of intracellular organelles and exocytosis of granules to guide neutrophils across the endothelial cell body (Carman *et al.*, 2007). Once neutrophils migrate through the endothelium, they are exposed to a gradient of chemoattractants such as *N*-formylmethionyl-leucyl-phenylalanine (fMLP) and complement 5a (C5a). At the sites of infection, neutrophils, along with complement protein and immunoglobulin, recognize and bind opsonized bacteria, and execute the various antimicrobial mechanisms including phagocytosis, degranulation, and neutrophil extracellular trap (NET) formation. During these processes, released neutrophil molecules can also damage host tissue. Inflammation develops in response to pathogen-associated molecular patterns (PAMPs, such as lipopolysaccharide (LPS), flagellin, and nucleic acids) released into the tissue during infection, and also in response to damage-associated molecular pattern molecules (DAMPs, such as heat-shock proteins, high-mobility group box 1 (HMGB-1), ATP, and DNA) that are generated in response to sterile injury.

Neutrophil Defense Mechanisms

Neutrophils can eliminate microbes by various intracellular and extracellular mechanisms.

Phagocytosis

Neutrophils sense invaders through Toll-like receptors (TLRs) and other pattern-recognition receptors. Human neutrophils express almost all TLRs, except TLR3 (Janke *et al.*, 2009;

Hayashi *et al.*, 2003). Upon stimulation, TLRs trigger a cascade of intracellular events that induce inflammatory gene expression regulated by NF- κ B elements (Tamassia *et al.*, 2007). Microbes are internalized by neutrophils through phagocytosis. The phagosome fuses with lysosomes to form the phagolysosome, where the microbe is subsequently killed by generation of reactive oxygen species (ROS). ROS production is initiated by a nicotinamide adenine dinucleotide phosphate-(NADPH)-oxidase complex-dependent process called oxidative burst. During this event, oxygen is converted into superoxide and transformed into hydrogen peroxide by superoxide dismutases. Subsequently, MPO is released into the phagolysosome through fusion with azurophilic granules, where it catalyzes the conversion of hydrogen peroxide and chloride into hypochlorous acid (Klebanoff *et al.*, 2013).

Degranulation

During this process, neutrophil granules translocate to the phagosomal or plasma membrane, where they dock and fuse with the membrane to release their contents. The release of granular components is tightly regulated mechanism that leads to exocytosis. Granules' exocytosis requires increases in intracellular Ca^{2+} and ATP and GTP hydrolysis. Actin cytoskeleton reorganization is also required. Granule release from neutrophils depends on activation of intracellular signaling pathways, including β -arrestins, the Rho guanosine triphosphatase Rac2, soluble NSF attachment protein (SNAP) receptors, the *src* family of tyrosine kinases, and the tyrosine phosphatase MEG2 (Lacy, 2006). Molecules such as neutrophil elastase, cathepsin G, and lysozymes degrade virulence factors and the bacterial cell wall. Antibacterial proteins such as BPI, cathelicidin, and defensins not only have the capacity to kill microbes but also can damage the host's tissues.

NETs

During NET formation, large strands of decondensed DNA in complex with granular proteins are extruded from the cell (Brinkmann *et al.*, 2004). The precise molecular mechanisms leading to NET formation are not yet fully elucidated. Various stimuli can trigger NET formation, including receptor-mediated signals (e.g., through TLR engagement) or through calcium channel formation through ionophores. These signals initiate a cascade of events, where calcium is released from its reservoirs by activation of PKC. This leads to assembly of NADPH oxidase and ROS generation (Matute *et al.*, 2009). H_2O_2 triggers activation and dissociation of the complex named 'azurosome,' which is formed by neutrophil elastase, MPO, azurocidin, cathepsin G, eosinophil cationic protein, defensin-1, lysozyme, and lactoferrin, leading to translocation of neutrophil elastase and MPO to the nucleus (Metzler *et al.*, 2014). Once in the nucleus, elastase partially cleaves histones to promote chromatin decondensation (Papayannopoulos *et al.*, 2010; Metzler *et al.*, 2014). Peptidylarginine deaminase 4 (PAD4) is activated by PKC ζ and translocates into the nucleus where it citrullinates histones (Brinkmann *et al.*, 2004; Neeli and Radic, 2013). This phenomenon appears to be crucial for chromatin decondensation leading to NET

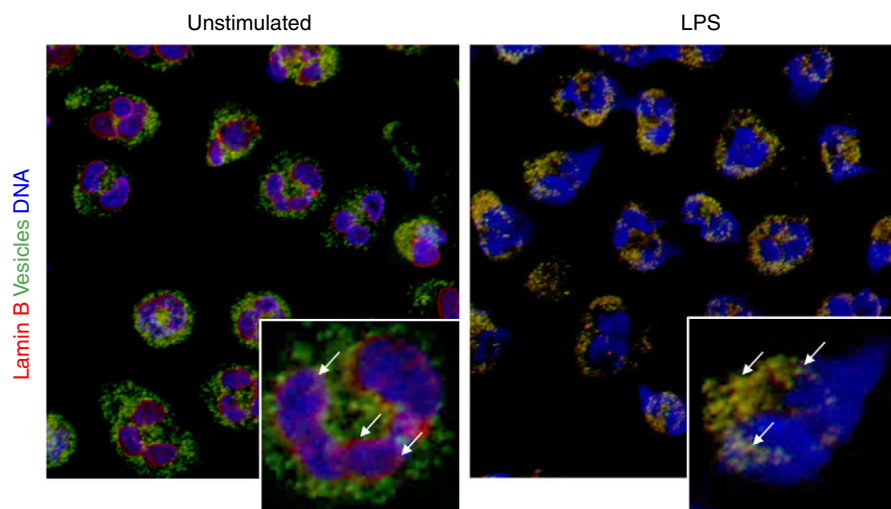


Figure 2 Cellular events during NETosis. NET formation is characterized by disintegration of the nuclear envelope followed by DNA decondensation. Neutrophils from a healthy volunteer were stimulated with LPS for 30 min. A resident nuclear membrane protein, lamin B was stained in red and vesicles were stained in green. Nuclear material was counterstained with Hoechst. Arrows indicate localization of lamin B before and after stimulation with LPS.

formation. The nuclear membrane disintegrates and decondensed DNA mixes with granules and cytosolic proteins (Figure 2). This complex of nuclear material and cytoplasmic and nuclear proteins is extruded to the extracellular space as NETs (Brinkmann *et al.*, 2004) that serve to trap and immobilize microbes and represent a lytic mechanism of neutrophil cell death. The presence of antimicrobial proteins (e.g., histones, defensins, elastase, proteinase 3, cathepsin G, lactoferrin, and MPO) in these structures supports their microbicidal role toward various pathogens, including bacteria, viruses, fungi, and parasites (Urban *et al.*, 2009; McCormick *et al.*, 2010). Neutrophils can also release mitochondrial DNA in NETs, without undergoing a cell death program (Yousefi *et al.*, 2009, 2008) in response to LPS and other stimuli. In addition, *Staphylococcus aureus* can trigger NET formation in a rapid, non-lytic, and oxidant-independent mechanism called vital NETosis (Pilszczek *et al.*, 2010). During this event, the plasma membrane remains intact. Vesicles containing DNA and proteins are pinched off from the nuclear membrane and fuse to the plasma membrane, releasing NETs. Vital NETosis has been also documented in response to TLR4-mediated platelets binding to neutrophils (Clark *et al.*, 2007).

Neutrophil Subsets

Given their short half-life, neutrophils have been difficult to study and characterize in detail. While considered a homogenous leukocyte subset for many years, emerging evidence suggests that these cells carry significant heterogeneity and plasticity with various degrees of phagocytic potential, protein synthesis, and oxidative metabolism. Indeed, various neutrophil subsets in response to physiological and pathological conditions have been described (Gallin, 1984; Beyrau *et al.*, 2012). It is still unclear if neutrophil subsets represent truly distinct lineages or develop from a single neutrophil precursor in response to the microenvironment. Up to this date, subsets

of neutrophils have been identified through expression of specific markers, gene expression, and functional aspects in healthy and pathological conditions. Some examples are mentioned below:

Olfactomedin 4-Expressing Neutrophils

Olfactomedin 4 (OLFM4) is a matrix glycoprotein expressed during the myelocyte and metamyelocyte stages of maturation, and levels of this protein decline as neutrophils mature (Clemmensen *et al.*, 2012). OLFM4 is expressed in approximately 25% of peripheral blood neutrophils and its presence is not regulated at the transcriptional level, suggesting microRNA (miRNA) regulation (Clemmensen *et al.*, 2012). OLFM4 is a granular protein that co-localizes with NGAL (Clemmensen *et al.*, 2012) and is externalized in NETs (Welin *et al.*, 2013). While the exact role of this neutrophil subset remains to be determined, OLFM4 may have immunoregulatory roles (Liu *et al.*, 2013).

CD177-Expressing Neutrophils

The glycoprotein CD177 (NB1) is a glycosyl-phosphatidylinositol-anchored receptor that is differentially expressed in circulating neutrophils (Stroncek, 2007; Hu *et al.*, 2009). Neutrophils that express CD177 demonstrate higher affinity to PECAM-1. CD177 associates with membrane PR3 promoting infiltration into tissues (Figure 3). CD177-positive neutrophils are increased in the autoimmune vasculitis granulomatosis with polyangiitis (Hu *et al.*, 2009). It also appears that these cells synthesize higher levels of IL-17. More information is required regarding their role in induction of tissue damage and vascular inflammation.

Low-Density Granulocytes

Low-density granulocytes (LDGs) are a distinct proinflammatory subset of neutrophils identified in the

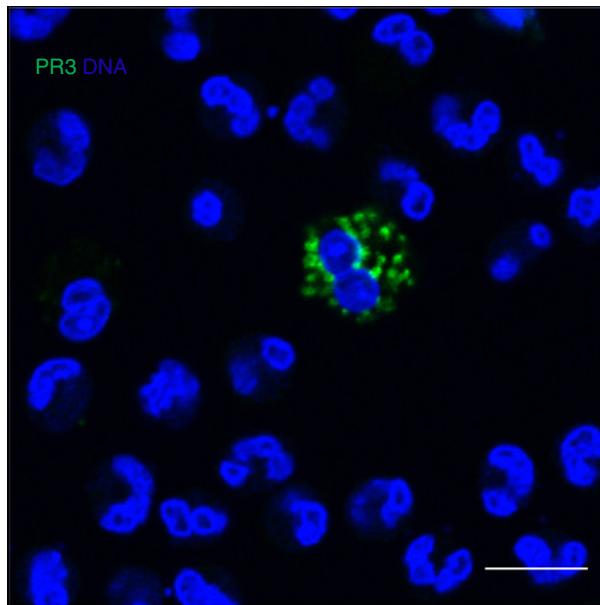


Figure 3 Neutrophil heterogeneity. Emerging evidence supports the notion of heterogeneity among leukocytes. A subset of neutrophils expressing membrane-bound PR3 can be found in neutrophil preparations from healthy volunteers. PR3 is stained in green and DNA in blue. Scale bar 10 μm .

mononuclear cell fraction of patients with systemic lupus erythematosus (SLE) (Hacharth and Kajtacsy-Balla, 1986; Denny *et al.*, 2010) and other autoimmune conditions such as psoriasis (Lin *et al.*, 2011). LDGs synthesize higher levels of proinflammatory cytokines (Denny *et al.*, 2010) and are primed to form significantly higher numbers of NETs (Villanueva *et al.*, 2011). LDG-NETs contain higher concentrations of metalloproteinases, such as MMP-9 that trigger endothelial damage (Carmona-Rivera *et al.*, in press). Molecules externalized by LDG-NETs may serve as modified sources of autoantigens in SLE (Carmona-Rivera and Kaplan, 2013; Knight *et al.*, 2012).

Proangiogenic MMP-9 Delivering Neutrophils

Vascular endothelial growth factor (VEGF)-A facilitates the recruitment of a subset of neutrophils that express high levels of MMP-9 to a site of hypoxia. The $\text{CD11b}^+/\text{Gr-1}^+/\text{CXCR4}^+$ leukocyte subset has been reported to be involved in the revascularization of transplanted pancreatic islets. They are more primed and prone to respond to a hypoxic stimulus, expressing more CXCR4 and containing more MMP-9 with proangiogenic capabilities (Christofferson *et al.*, 2012).

T-Cell Regulatory Neutrophils

Recent evidence indicates that neutrophils can also have important regulatory roles in adaptive immune responses through effects on T and B lymphocyte proliferation, activation, and differentiation. A subset of human neutrophils that can suppress T cells response was identified in healthy

volunteers after intravenous LPS administration. Regulatory neutrophils are $\text{CD62L}^{\text{dim}}/\text{CD11c}^{\text{bright}}$, demonstrate hypersegmented nucleus and are capable of producing large quantities of H_2O_2 . Suppression of T cells requires formation of an immunological synapse through neutrophil $\alpha_{\text{M}}\beta_2$ -integrin (MAC-1) and large amounts of H_2O_2 (Pillay *et al.*, 2012). It remains to be better understood whether specific neutrophil subsets serve different roles in modulation of adaptive immunity.

Genetic Defects That Affect Neutrophil Numbers and/or Function

While genetic diseases that affect neutrophil numbers and function are rare, they provide additional understanding into the molecular mechanisms by which these cells are crucial in host defense. A few are mentioned below:

Severe Congenital Neutropenia

Severe congenital neutropenia (SCN) results from mutations in various neutrophil genes including neutrophil elastase (ELANE), HAX-1, and Gfi-1. Neutrophil formation is arrested at the promyelocytic stage, patients display very low neutrophil counts and develop life-threatening infections.

Leukocyte Adhesion Disorders

Leukocyte adhesion disorders (LAD) develop due to defects in neutrophil cell adhesion to the endothelium or the extracellular matrix. Mutations in genes encoding for specific integrins (such as LAD-I and LAD-II diseases) are characterized by severe infections due to defects in neutrophil recruitment.

Chronic Granulomatous Disease

Chronic granulomatous disease (CGD) is the most common of the genetic neutrophil function disorders. It is a heterogeneous group of hereditary diseases characterized by an inability of phagocytic cells to kill engulfed organisms due to defects in ROS generation through NADPH oxidase complex (Figure 4). This is due to defects in subunits of phagocyte NADPH oxidase (Kuhns *et al.*, 2010; Bianchi *et al.*, 2009). Patients suffer from recurrent, life-threatening infections, underlying the important role of ROS production in host defense.

Mutations in the Gene Encoding MPO

Mutations in the gene encoding MPO have been described in patients that suffer from recurrent *Candida* infections. One out of 3000 individuals may harbor mutations that affect MPO production, activity or trafficking (Winterbourn *et al.*, 2006).

Papillon-Lefèvre syndrome

Papillon-Lefèvre syndrome (PLS) is an autosomal recessive disorder caused by a deficiency in cathepsin C, an enzyme that

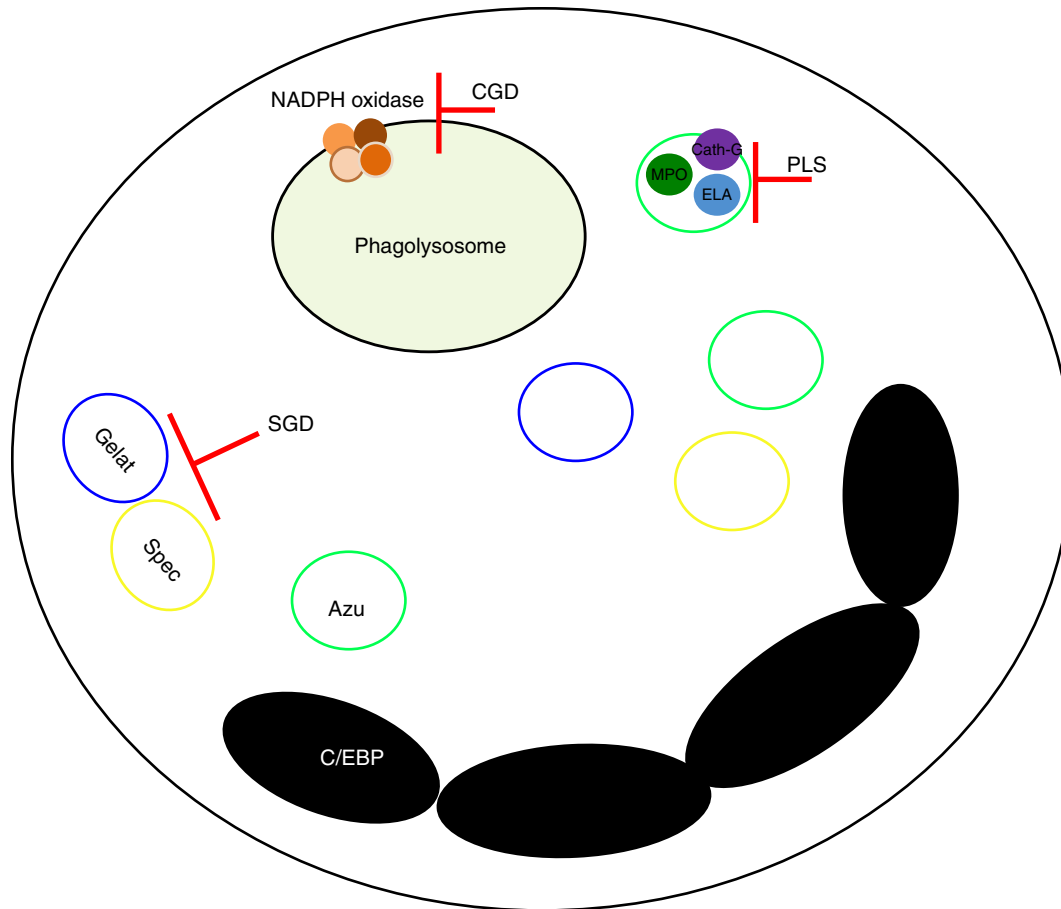


Figure 4 Genetic diseases that affect neutrophil function. Genetic alterations that affect neutrophil function can trigger life-threatening infections. Patients with mutation in subunits of the NADPH oxidase develop CGD. PLS is characterized by deficiency of proteinase activity present in azurophilic granules. Lack of secondary and tertiary granules is found in patients with mutation in the transcription factor C/EBP ϵ that develop neutrophil-SGD. Azu, azurophilic granules; Cath G, Cathepsin G; ELA, elastase; Gelat, gelatinase granules; MPO, myeloperoxidase; Spec, specific granules.

processes various serine proteases crucial in antimicrobial defense (Figure 4). The immunodeficiency that develops in these patients is mild and the mutation affects mostly mature neutrophils (De Haar *et al.*, 2006; Sorensen *et al.*, 2014).

Neutrophil-Specific Granule Deficiency

Neutrophil-specific granule deficiency (SGD) is a rare congenital disease where neutrophils display atypical bilobed nuclei and lack expression of secondary and tertiary granular proteins with multiple defects in antimicrobial strategies and severe bacterial infections (Figure 4). Mutations in the transcription factor C/EBP ϵ are the primary genetic defects (Gombart *et al.*, 2001).

Neutrophils in Systemic Autoimmune Diseases

While neutrophils are key players in the innate response to infections, they can also contribute to the development and

perpetuation of chronic inflammatory/autoimmune diseases. A few examples are mentioned below:

Vasculitis

Vasculitis is a group of diseases characterized by inflammation of the blood vessels. Small vessels vasculitis (SVV) is a subtype of this disease where patients synthesize anti-neutrophil cytoplasmic antibodies (ANCA) recognizing PR3 or MPO. Both activated neutrophils and ANCAs have been implicated in the development of vasculitis. MPO and PR3 are externalized during NET formation and autoantibodies against PR3 induce NETosis (Kessenbrock *et al.*, 2009). Dysregulated neutrophil cell death has been proposed to play a pathogenic role in SSV.

Rheumatoid Arthritis

Rheumatoid Arthritis (RA) is an autoimmune disease that results in destructive erosive inflammation of synovial joints (De Rycke *et al.*, 2004). Neutrophils are the most abundant

cells found in the synovial fluid of RA patients (Mohr *et al.*, 1981) and they form NETs more readily than control neutrophils (Khandpur *et al.*, 2013). Various citrullinated proteins released during NET formation are RA autoantigens (Khandpur *et al.*, 2013; Pratesi *et al.*, 2014). NETs induce a proinflammatory phenotype in fibroblast-like synoviocytes, which may lead to enhanced articular damage (Khandpur *et al.*, 2013).

SLE

SLE is a systemic autoimmune syndrome characterized by the development of autoantibodies to nuclear components, immune complex deposition, inflammation, and organ damage (Tsokos, 2011). Neutrophils may contribute to loss of tolerance and organ damage in SLE. Abnormalities in phagocytosis, activation, and NET formation have been described in SLE neutrophils (Brandt and Hedberg, 1969; Courtney *et al.*, 1999; Hashimoto *et al.*, 1982; Villanueva *et al.*, 2011; Carmona-Rivera *et al.*, in press). Dysregulated NET degradation through inhibition of DNase-I has also been described (Hakim *et al.*, 2010; Leffler *et al.*, 2012). NETs also promote synthesis of cytokines considered key in lupus pathogenesis and inhibition of NET formation decreases lupus phenotype in animal models (Villanueva *et al.*, 2011; Knight *et al.*, 2013).

Conclusion

Neutrophils are specialized, terminally differentiated cells that are endowed with crucial roles in host defense against microbes. Genetic or acquired deficiencies in neutrophil numbers or function can be life-threatening and lead to severe infections, while unchecked neutrophil responses or extended neutrophil half-life may lead to chronic inflammatory and autoimmune conditions. Neutrophils are more heterogeneous than originally described and many unanswered questions remain with regards to their specialization and plasticity and their interactions with other cells of the innate and adaptive immune system.

See also: Cell Communication: Prototypic Integrative Processes: Wound Healing: An Orchestrated Process of Cell Cycle, Adhesion, and Signaling. Vertical Integration: Applications: A Review on Various Mathematical Modeling Approaches for Wound Healing

References

- Bainton, D.F., Ulyot, J.L., Farquhar, M.G., 1971. The development of neutrophilic polymorphonuclear leukocytes in human bone marrow. *Journal of Experimental Medicine* 134, 907–934.
- Barreiro, O., Sanchez-Madrid, F., 2009. Molecular basis of leukocyte-endothelium interactions during the inflammatory response. *Revista española de cardiología* 62, 552–562.
- Beyrau, M., Bodkin, J.V., Nourshargh, S., 2012. Neutrophil heterogeneity in health and disease: A revitalized avenue in inflammation and immunity. *Open Biology* 2, 120134.
- Bianchi, M., Hakim, A., Brinkmann, V., *et al.*, 2009. Restoration of NET formation by gene therapy in CGD controls aspergillosis. *Blood* 114, 2619–2622.
- Bjerrgaard, M.D., Jurlander, J., Klausen, P., Borregaard, N., Cowland, J.B., 2003. The *in vivo* profile of transcription factors during neutrophil differentiation in human bone marrow. *Blood* 101, 4322–4332.
- Borregaard, N., 2010. Neutrophils, from marrow to microbes. *Immunity* 33, 657–670.
- Borregaard, N., Cowland, J.B., 1997. Granules of the human neutrophilic polymorphonuclear leukocyte. *Blood* 89, 3503–3521.
- Borregaard, N., Sehested, M., Nielsen, B.S., Sengelov, H., Kjeldsen, L., 1995. Biosynthesis of granule proteins in normal human bone marrow cells. Gelatinase is a marker of terminal neutrophil differentiation. *Blood* 85, 812–817.
- Brandt, L., Hedberg, H., 1969. Impaired phagocytosis by peripheral blood granulocytes in systemic lupus erythematosus. *Scandinavian Journal of Haematology* 6, 348–353.
- Brinkmann, V., Reichard, U., Goosmann, C., *et al.*, 2004. Neutrophil extracellular traps kill bacteria. *Science* 303, 1532–1535.
- Carman, C.V., Sage, P.T., Sciuto, T.E., *et al.*, 2007. Transcellular diapedesis is initiated by invasive podosomes. *Immunity* 26, 784–797.
- Carmona-Rivera, C., Kaplan, M.J., 2013. Low-density granulocytes: A distinct class of neutrophils in systemic autoimmunity. *Seminars in Immunopathology* 35, 455–463.
- Carmona-Rivera, C., Zhao, W., Yalavarthi, S., Kaplan, M.J., in press. Neutrophil extracellular traps induce endothelial dysfunction in systemic lupus erythematosus through the activation of matrix metalloproteinase-2. *Annals of the Rheumatic Diseases*. doi: 10.1136/annrheumdis-2013-204837.
- Christoffersson, G., Vagesjo, E., Vandooren, J., *et al.*, 2012. VEGF-A recruits a proangiogenic MMP-9-delivering neutrophil subset that induces angiogenesis in transplanted hypoxic tissue. *Blood* 120, 4653–4662.
- Clark, S.R., Ma, A.C., Tavener, S.A., *et al.*, 2007. Platelet TLR4 activates neutrophil extracellular traps to ensnare bacteria in septic blood. *Nature Medicine* 13, 463–469.
- Clemmensen, S.N., Bohr, C.T., Rorvig, S., *et al.*, 2012. Olfactomedin 4 defines a subset of human neutrophils. *Journal of Leukocyte Biology* 91, 495–500.
- Courtney, P.A., Crockard, A.D., Williamson, K., *et al.*, 1999. Increased apoptotic peripheral blood neutrophils in systemic lupus erythematosus: Relations with disease activity, antibodies to double stranded DNA, and neutropenia. *Annals of the Rheumatic Diseases* 58, 309–314.
- Dahl, R., Walsh, J.C., Lancki, D., *et al.*, 2003. Regulation of macrophage and neutrophil cell fates by the PU.1:C/EBPalpha ratio and granulocyte colony-stimulating factor. *Nature Immunology* 4, 1029–1036.
- De Haar, S.F., Hiemstra, P.S., Van Steenbergen, M.T., Everts, V., Beertsen, W., 2006. Role of polymorphonuclear leukocyte-derived serine proteinases in defense against *Actinobacillus actinomycetemcomitans*. *Infection and Immunity* 74, 5284–5291.
- De Rycke, L., Peene, I., Hoffman, I.E., *et al.*, 2004. Rheumatoid factor and anticitrullinated protein antibodies in rheumatoid arthritis: Diagnostic value, associations with radiological progression rate, and extra-articular manifestations. *Annals of the Rheumatic Diseases* 63, 1587–1593.
- Denny, M.F., Yalavarthi, S., Zhao, W., *et al.*, 2010. A distinct subset of proinflammatory neutrophils isolated from patients with systemic lupus erythematosus induces vascular damage and synthesizes type I IFNs. *Journal of Immunology* 184, 3284–3297.
- Eash, K.J., Means, J.M., White, D.W., Link, D.C., 2009. CXCR4 is a key regulator of neutrophil release from the bone marrow under basal and stress granulopoiesis conditions. *Blood* 113, 4711–4719.
- Faurschou, M., Borregaard, N., 2003. Neutrophil granules and secretory vesicles in inflammation. *Microbes and Infection* 5, 1317–1327.
- Fazi, F., Rosa, A., Fatica, A., *et al.*, 2005. A microcircuitry comprised of microRNA-223 and transcription factors NF1-A and C/EBPalpha regulates human granulopoiesis. *Cell* 123, 819–831.
- Gallin, J.I., 1984. Human neutrophil heterogeneity exists, but is it meaningful? *Blood* 63, 977–983.
- Geering, B., Simon, H.U., 2011. Peculiarities of cell death mechanisms in neutrophils. *Cell Death & Differentiation* 18, 1457–1469.
- Gombart, A.F., Shiohara, M., Kwok, S.H., *et al.*, 2001. Neutrophil-specific granule deficiency: Homozygous recessive inheritance of a frameshift mutation in the gene encoding transcription factor CCAAT/enhancer binding protein-epsilon. *Blood* 97, 2561–2567.
- Hachbarth, E., Kajdacsy-Balla, A., 1986. Low density neutrophils in patients with systemic lupus erythematosus, rheumatoid arthritis, and acute rheumatic fever. *Arthritis and Rheumatism* 29, 1334–1342.
- Hakim, A., Furnrohr, B.G., Amann, K., *et al.*, 2010. Impairment of neutrophil extracellular trap degradation is associated with lupus nephritis. *Proceedings of*

- the National Academy of Sciences of the United States of America 107, 9813–9818.
- Hashimoto, Y., Ziff, M., Hurd, E.R., 1982. Increased endothelial cell adherence, aggregation, and superoxide generation by neutrophils incubated in systemic lupus erythematosus and Felty's syndrome sera. *Arthritis and Rheumatism* 25, 1409–1418.
- Hayashi, F., Means, T.K., Luster, A.D., 2003. Toll-like receptors stimulate human neutrophil function. *Blood* 102, 2660–2669.
- Hernandez, P.A., Gorlin, R.J., Lukens, J.N., *et al.*, 2003. Mutations in the chemokine receptor gene CXCR4 are associated with WHIM syndrome, a combined immunodeficiency disease. *Nature Genetics* 34, 70–74.
- Hock, H., Hamblen, M.J., Rooke, H.M., *et al.*, 2004. Gfi-1 restricts proliferation and preserves functional integrity of haematopoietic stem cells. *Nature* 431, 1002–1007.
- Hu, N., Westra, J., Huitema, M.G., *et al.*, 2009. Coexpression of CD177 and membrane proteinase 3 on neutrophils in antineutrophil cytoplasmic autoantibody-associated systemic vasculitis: Anti-proteinase 3-mediated neutrophil activation is independent of the role of CD177-expressing neutrophils. *Arthritis and Rheumatism* 60, 1548–1557.
- Janke, M., Poth, J., Wimmenauer, V., *et al.*, 2009. Selective and direct activation of human neutrophils but not eosinophils by Toll-like receptor 8. *Journal of Allergy and Clinical Immunology* 123, 1026–1033.
- Kessenbrock, K., Krumbholz, M., Schonermarck, U., *et al.*, 2009. Netting neutrophils in autoimmune small-vessel vasculitis. *Nature Medicine* 15, 623–625.
- Khandpur, R., Carmona-Rivera, C., Vivekanandan-Giri, A., *et al.*, 2013. NETs are a source of citrullinated autoantigens and stimulate inflammatory responses in rheumatoid arthritis. *Science Translational Medicine* 5, 178ra40.
- Klebanoff, S.J., Kettle, A.J., Rosen, H., Winterbourn, C.C., Nauseef, W.M., 2013. Myeloperoxidase: A front-line defender against phagocytosed microorganisms. *Journal of Leukocyte Biology* 93, 185–198.
- Klein, C., 2011. Genetic defects in severe congenital neutropenia: Emerging insights into life and death of human neutrophil granulocytes. *Annual Review of Immunology* 29, 399–413.
- Knight, J.S., Carmona-Rivera, C., Kaplan, M.J., 2012. Proteins derived from neutrophil extracellular traps may serve as self-antigens and mediate organ damage in autoimmune diseases. *Frontiers in Immunology* 3, 380.
- Knight, J.S., Zhao, W., Luo, W., *et al.*, 2013. Peptidylarginine deiminase inhibition is immunomodulatory and vasculoprotective in murine lupus. *Journal of Clinical Investigation* 123, 2981–2993.
- Kuhns, D.B., Alvord, W.G., Heller, T., *et al.*, 2010. Residual NADPH oxidase and survival in chronic granulomatous disease. *New England Journal of Medicine* 363, 2600–2610.
- Lacy, P., 2006. Mechanisms of degranulation in neutrophils. *Allergy, Asthma & Clinical Immunology* 2, 98–108.
- Lajtha, L.G., Ellis, F., Oliver, R., 1953. Isotope uptake of individual cells: Uptake of S35 sulphate by human bone-marrow cells *in vitro*. *British Journal of Cancer* 7, 401–406.
- Laslo, P., Spooner, C.J., Warmflash, A., *et al.*, 2006. Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell* 126, 755–766.
- Le Cabec, V., Cowland, J.B., Calafat, J., Borregaard, N., 1996. Targeting of proteins to granule subsets is determined by timing and not by sorting: The specific granule protein NGAL is localized to azurophilic granules when expressed in HL-60 cells. *Proceedings of the National Academy of Sciences of the United States of America* 93, 6454–6457.
- Leffler, J., Martin, M., Gullstrand, B., *et al.*, 2012. Neutrophil extracellular traps that are not degraded in systemic lupus erythematosus activate complement exacerbating the disease. *Journal of Immunology* 188, 3522–3531.
- Ley, K., Laudanna, C., Cybulsky, M.I., Nourshargh, S., 2007. Getting to the site of inflammation: The leukocyte adhesion cascade updated. *Nature Reviews Immunology* 7, 678–689.
- Lin, A.M., Rubin, C.J., Khandpur, R., *et al.*, 2011. Mast cells and neutrophils release IL-17 through extracellular trap formation in psoriasis. *Journal of Immunology* 187, 490–500.
- Liu, W., Yan, M., Sugui, J.A., *et al.*, 2013. Olfm4 deletion enhances defense against *Staphylococcus aureus* in chronic granulomatous disease. *Journal of Clinical Investigation* 123, 3751–3755.
- Matute, J.D., Arias, A.A., Wright, N.A., *et al.*, 2009. A new genetic subgroup of chronic granulomatous disease with autosomal recessive mutations in p40 phox and selective defects in neutrophil NADPH oxidase activity. *Blood* 114, 3309–3315.
- McCormick, A., Heesemann, L., Wagener, J., *et al.*, 2010. NETs formed by human neutrophils inhibit growth of the pathogenic mold *Aspergillus fumigatus*. *Microbes and Infection* 12, 928–936.
- Metzler, K.D., Goosmann, C., Lubojemska, A., Zychlinsky, A., Papayannopoulos, V., 2014. A myeloperoxidase-containing complex regulates neutrophil elastase release and actin dynamics during NETosis. *Cell Reports* 8, 883–896.
- Mohr, W., Westerhellweg, H., Wessinghage, D., 1981. Polymorphonuclear granulocytes in rheumatic tissue destruction. III. An electron microscopic study of PMNs at the pannus-cartilage junction in rheumatoid arthritis. *Annals of the Rheumatic Diseases* 40, 396–399.
- Neeli, I., Radic, M., 2013. Opposition between PKC isoforms regulates histone deimination and neutrophil extracellular chromatin release. *Frontiers in Immunology* 4, 38.
- Papayannopoulos, V., Metzler, K.D., Hakkim, A., Zychlinsky, A., 2010. Neutrophil elastase and myeloperoxidase regulate the formation of neutrophil extracellular traps. *Journal of Cell Biology* 191, 677–691.
- Phillipson, M., Heit, B., Colarusso, P., *et al.*, 2006. Intraluminal crawling of neutrophils to emigration sites: A molecularly distinct process from adhesion in the recruitment cascade. *Journal of Experimental Medicine* 203, 2569–2575.
- Pillay, J., Kamp, V.M., Van Hoffen, E., *et al.*, 2012. A subset of neutrophils in human systemic inflammation inhibits T cell responses through Mac-1. *Journal of Clinical Investigation* 122, 327–336.
- Pilschek, F.H., Salina, D., Poon, K.K., *et al.*, 2010. A novel mechanism of rapid nuclear neutrophil extracellular trap formation in response to *Staphylococcus aureus*. *Journal of Immunology* 185, 7413–7425.
- Pratesi, F., Dioni, I., Tommasi, C., *et al.*, 2014. Antibodies from patients with rheumatoid arthritis target citrullinated histone 4 contained in neutrophils extracellular traps. *Annals of the Rheumatic Diseases* 73, 1414–1422.
- Quesenberry, P., Levitt, L., 1979. Hematopoietic stem cells. *New England Journal of Medicine* 301, 755–761.
- Reddy, V.A., Iwama, A., Iotzova, G., *et al.*, 2002. Granulocyte inducer C/EBP α inactivates the myeloid master regulator PU.1: Possible role in lineage commitment decisions. *Blood* 100, 483–490.
- Sachs, U.J., Andrei-Selmer, C.L., Maniar, A., *et al.*, 2007. The neutrophil-specific antigen CD177 is a counter-receptor for platelet endothelial cell adhesion molecule-1 (CD31). *Journal of Biological Chemistry* 282, 23603–23612.
- Sorensen, O.E., Clemmensen, S.N., Dahl, S.L., *et al.*, 2014. Papillon-Lefevre syndrome patient reveals species-dependent requirements for neutrophil defenses. *Journal of Clinical Investigation* 124, 4539–4548.
- Stronczek, D.F., 2007. Neutrophil-specific antigen HNA-2a, NB1 glycoprotein, and CD177. *Current Opinion in Hematology* 14, 688–693.
- Tamassia, N., Le Moigne, V., Calzetti, F., *et al.*, 2007. The MyD88-independent pathway is not mobilized in human neutrophils stimulated via TLR4. *Journal of Immunology* 178, 7344–7356.
- Tsokos, G.C., 2011. Systemic lupus erythematosus. *New England Journal of Medicine* 365, 2110–2121.
- Urban, C.F., Ermet, D., Schmid, M., *et al.*, 2009. Neutrophil extracellular traps contain calprotectin, a cytosolic protein complex involved in host defense against *Candida albicans*. *PLoS Pathogens* 5, e1000639.
- Villanueva, E., Yalavarthi, S., Berthier, C.C., *et al.*, 2011. Netting neutrophils induce endothelial damage, infiltrate tissues, and expose immunostimulatory molecules in systemic lupus erythematosus. *Journal of Immunology* 187, 538–552.
- Welin, A., Amirbeigi, F., Christenson, K., *et al.*, 2013. The human neutrophil subsets defined by the presence or absence of OLFM4 both translocate into tissue *in vivo* and give rise to distinct NETs *in vitro*. *PLoS One* 8, e69575.
- Winterbourn, C.C., Hampton, M.B., Livesey, J.H., Kettle, A.J., 2006. Modeling the reactions of superoxide and myeloperoxidase in the neutrophil phagosome: Implications for microbial killing. *Journal of Biological Chemistry* 281, 39860–39869.
- Woodfin, A., Voisin, M.B., Imhof, B.A., *et al.*, 2009. Endothelial cell activation leads to neutrophil transmigration as supported by the sequential roles of ICAM-2, JAM-A, and PECAM-1. *Blood* 113, 6246–6257.
- Woodfin, A., Voisin, M.B., Nourshargh, S., 2010. Recent developments and complexities in neutrophil transmigration. *Current Opinion in Hematology* 17, 9–17.
- Yago, T., Shao, B., Miner, J.J., *et al.*, 2010. E-selectin engages PSGL-1 and CD44 through a common signaling pathway to induce integrin α L β 2-mediated slow leukocyte rolling. *Blood* 116, 485–494.
- Yamanaka, R., Kim, G.D., Radomska, H.S., *et al.*, 1997. CCAAT/enhancer binding protein Epsilon is preferentially up-regulated during granulocytic differentiation and its functional versatility is determined by alternative use of promoters and differential splicing. *Proceedings of the National Academy of Sciences of the United States of America* 94, 6462–6467.

- Yang, L., Kowalski, J.R., Yacono, P., *et al.*, 2006. Endothelial cell cortactin coordinates intercellular adhesion molecule-1 clustering and actin cytoskeleton remodeling during polymorphonuclear leukocyte adhesion and transmigration. *Journal of Immunology* 177, 6440–6449.
- Yousefi, S., Gold, J.A., Andina, N., *et al.*, 2008. Catapult-like release of mitochondrial DNA by eosinophils contributes to antibacterial defense. *Nature Medicine* 14, 949–953.
- Yousefi, S., Mihalache, C., Kozłowski, E., Schmid, I., Simon, H.U., 2009. Viable neutrophils release mitochondrial DNA to form neutrophil extracellular traps. *Cell Death & Differentiation* 16, 1438–1444.
- Zarebski, A., Velu, C.S., Baktula, A.M., *et al.*, 2008. Mutations in growth factor independent-1 associated with human neutropenia block murine granulopoiesis through colony stimulating factor-1. *Immunity* 28, 370–380.

Specialized Subsets of Tissue-Resident Macrophages in Secondary Lymphoid Organs

A Chauveau and TI Arnon, University of Oxford, Oxford, UK

© 2016 Elsevier Inc. All rights reserved.

Macrophages are professional phagocytes that are able to efficiently engulf and digest extracellular particles, cell debris, and pathogens. They play critical roles in regulating host immunity and participate in many other biological processes from tissue repair mechanisms and clearance of apoptotic cells to regulation of metabolism and temperature control (Davies *et al.*, 2013; Murray and Wynn, 2011; Wynn *et al.*, 2013). Macrophages can be broadly divided into two classes: tissue-resident and infiltrating macrophages. Although resident macrophages have previously been thought to develop from circulating monocytes that originate from hematopoietic stem cells, recent findings challenged this dogma and revealed that in many cases resident macrophages are derived from yolk sac embryogenesis and have the capacity to self-renew during both steady state and following immunological challenges. It is likely that local signals derived from the tissue itself play key roles in promoting the specification of tissue macrophages, a notion supported by the strict correlation observed between macrophage's positioning and their phenotypic characteristics. However, the nature and identity of these signals are still largely unknown. Importantly, macrophages are not only heterogenous, but also adaptable and quickly modify their phenotype in response to environmental changes, a quality that endows them with the ability to adjust their function in a dynamic fashion (Martinez *et al.*, 2009).

The complex heterogeneity of mature tissue macrophages reflects a high degree of specialization and highlights the need to study these unique cells *in vivo*, in the context of their natural environments. In the past three decades, clodronate liposome treatment has been widely used *in vivo* to remove macrophages from lymphoid tissues and other sites (van Rooijen *et al.*, 1989). However, while this treatment is very effective in depleting phagocytes, it lacks selectivity and targets the majority of macrophage subsets as well as other cell lineages that are sensitive to the treatment (e.g., dendritic cells (DCs) and marginal zone (MZ) B cell). Attempts to generate mouse models with better specificity have been hampered by the considerable overlap in gene expression between closely related macrophage subsets and by the absence of single known selective markers that can distinguish between them. Recent efforts to generate comprehensive descriptions of gene expression profiles of tissue macrophages have greatly advanced the field and enhanced our ability to analyze macrophage biology with greater resolution (Gautier *et al.*, 2012). However, the development of mouse models based on these studies is lagging and the selective functions of many key macrophage populations remain poorly defined. Among those, macrophages that reside in secondary lymphoid organs have proved exceptionally challenging to study (Gautier *et al.*, 2012) partially due to their tendency to exchange surface markers with nearby lymphocytes (Gray *et al.*, 2012). In a recent careful analysis of fluorescence-activated cell sorting

(FACS) purified CD11b^{hi}CD169^{hi}F4/80⁺CD11c^{lo} from lymph nodes (LNs) of immunized mice, it was discovered that what were thought to be pure macrophage populations actually contained large numbers of contaminating T cells and natural killer cells that had acquired membrane protein containing macrophage-derived membrane blebs during the preparation procedure. This acquisition caused the lymphocytes to therefore appear by classical analysis as being macrophages. This caveat must therefore be considered when interpreting findings from past publications, but also highlights how critical it is that additional analysis such as microscopy or reverse transcription polymerase chain reaction (RT-PCR) are considered in all future work for ensuring that similar contamination issues are avoided. The development of better purification methods, and the generation of mouse models for selectively targeting major macrophage populations, will help to define the molecular basis of macrophage behavior and to identify their key functions in the context of complex biological settings.

Macrophage Subsets in the Spleen

The spleen is a major lymphoid organ that filters the blood. It provides critical protection from blood-borne pathogens and is important for the removal of antigens, foreign material, and damaged or aged red blood cells (Mebius and Kraal, 2005). To carry out its multiple functions, the spleen is organized in a modular fashion and is divided into several distinct compartments; the LN-like white pulp that is rich in lymphocytes, the surrounding red pulp where macrophages and transiting blood cells are abundant, and the MZ compartment that is located between them and is densely populated with macrophages and specialized subsets of lymphocytes (Mebius and Kraal, 2005). Most of the blood that enters the spleen is released into the MZ via sinuses that are located at the base of the compartment. These structures form a relatively impervious barrier on their follicular side, thereby restricting fluids and large molecules from entering the white pulp. However, on their outer side, the marginal sinuses are fenestrated, allowing free passage of blood contents into the MZ and red pulp. Venules that are scattered across the red pulp allow fluids and cells to egress the spleen and return into the blood circulation (Mebius and Kraal, 2005). Each one of these compartments contains typical macrophage subsets with distinct characteristics and functions (Figure 1). The best-described populations are found within the MZ and the red pulp and these are the focus of the next sections.

The Splenic MZ and Its Macrophages

The splenic MZ is an important transient region where newly arriving cells, antigens, and blood are filtered. Specialized cells

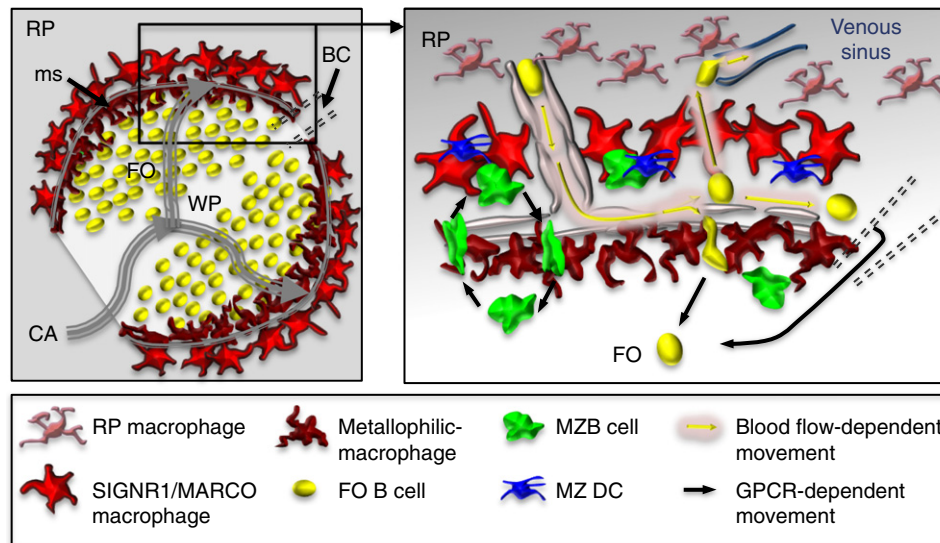


Figure 1 Illustration of spleen architecture and cellular composition of the marginal zone. The spleen can be divided into three main compartments: the red pulp (RP, grey), the white pulp (WP, white), and the marginal zone (MZ) that is located between them. The MZ and RP are separated by the marginal sinuses (ms). Breaks in the marginal sinus layer form the bridging channels (BC). Blood arriving via the central arteriole (CA) arrives to the marginal sinuses and passively flows through the MZ toward the red pulp and into venous sinuses. Metallophilic macrophages are located under the marginal sinuses, with some processes extending across the endothelial barrier (see also [Figure 2](#)). SIGN-R1/MARCO macrophages are located above the marginal sinuses and are found in close proximity with resident DCs and MZB cells. Circulating cells such as B cells enter the spleen with the blood and first arrive to the MZ. Some B cells may get caught in the blood flow and move toward the RP, while others transmigrate directly across the marginal sinuses to enter the follicles or move toward the bridging channels into white pulp. B cell migration between the MZ and the FO depends on G-protein coupled receptors (GPCR, black arrows) and is bi-directional, with some B cells using the MZ as an exit route to leave the FO. In the RP, the cell moves with the flow (shadowed arrows), enters a venous sinus and returns to circulation. MZB cells (green) are confined to the spleen and migrate back and forth between the MZ and follicular compartments in a GPC-dependent manner.

that reside in this compartment are involved in evaluating the nature of the arriving material and determining its fate. Within this region two major populations of macrophages have been identified; the metallophilic and the MZ macrophages ([Macauley et al., 2014](#)). Metallophilic macrophages are located at the inner ring of the MZ below the marginal sinuses ([Figure 2\(a\)](#)). They were named after their affinity for silver and are typically identified by high surface expression of the sialic-acid binding receptor CD169 (also known as Siglec-1) ([Macauley et al., 2014](#)). The second subset is found through the MZ compartment and is characterized by expression of the C-type lectin domain containing receptor SIGN-R1 and the scavenger receptor MARCO ([Figure 2\(b\)](#)). Although many published works use the term ‘MZ macrophages’ to describe these cells, in this article we will refer to them as ‘SIGN-R1/MARCO macrophages,’ whereas ‘MZ macrophages’ will be used as a collective name to describe both subsets. Importantly, the MZ is also home to MZB cells, a unique B cell subset that plays a central role in mounting rapid antibody responses against systemic pathogens ([Martin and Kearney, 2002](#)). In addition to recognizing and responding to cognate antigens, MZB cells express high levels of complement receptors (CD21 and CD35) and they are involved in the capture of opsonized antigens from the circulation and promoting their delivery into splenic follicles ([Arnon and Cyster, 2014](#)). A specialized subset of DCs also resides within the MZ and upon activation these cells migrate into the T zone and regulate T cell responses ([Mebius and Kraal, 2005](#)).

Notably, many circulating cells entering the spleen with the blood-flow pass through the MZ compartment and are likely to come into close contact with MZ macrophages. However, the nature and significance of such interactions remains largely unexplored.

Little is known about the specific factors that drive the differentiation of MZ macrophages in the spleen. Recently, an elegant report showed that the development of these cells depends on the expression of the liver X receptor alpha (LXR α) ([Noella et al., 2013](#)). This dependency was selective to macrophages that reside in the MZ and LXR α deficiency did not affect the development of macrophage subsets located in the red pulp or in other lymphoid tissues including in the LN, Peyer’s patches, liver, or intestine. Importantly, MZ macrophages could be efficiently reconstituted in LXR α deficient animals by adoptive transfer of wild type (WT) monocytes, thus demonstrating the capacity of circulating hematopoietic-derived myeloid cells to replenish them. Whether a local neonatal-derived precursor is responsible for homeostatic repopulation of MZ macrophages in WT animals, remains to be determined. Since oxysterols are known to bind LXR α and to inhibit inflammatory responses via negative regulation of IL-1 β expression ([Spann et al., 2012](#)), it is tempting to speculate that local splenic sterol derivatives may indirectly regulate systemic immunity by controlling the development of MZ macrophages.

Despite the fact that MZ macrophages have been the subject of intense investigation over the past three decades, there are

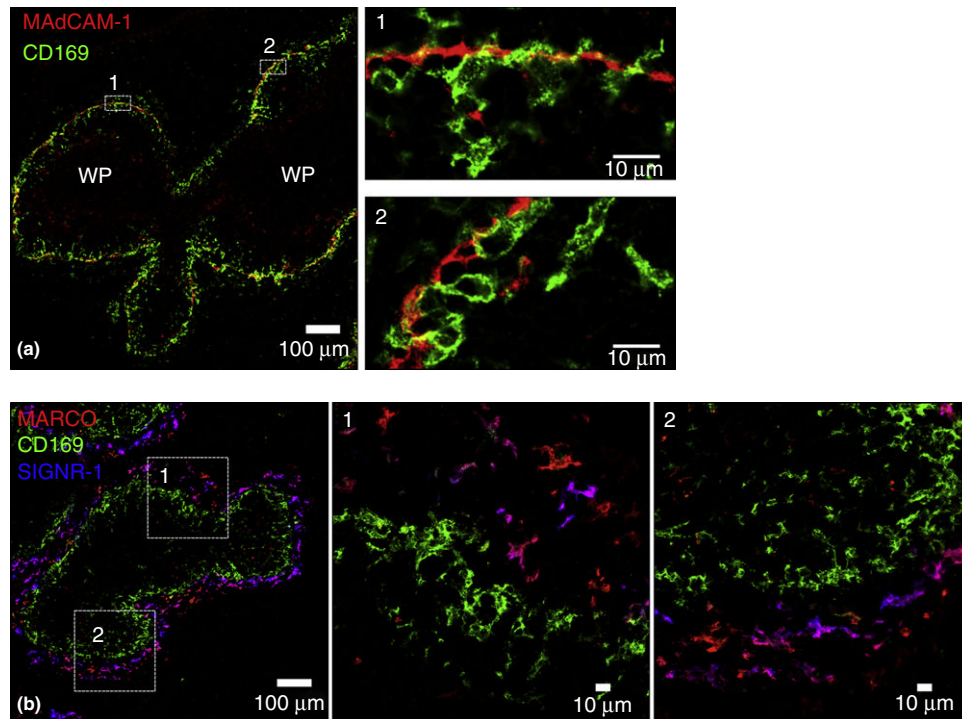


Figure 2 Macrophage populations in the splenic MZ. (a) Immunofluorescent staining of a frozen spleen section showing CD169+ metallophilic macrophages (green) and marginal sinus endothelial cells (MAdCAM-1+, red). On the right, a close up on 2 selected areas. In most cases, the macrophage body is located under the sinus, but occasionally cell processes are extended across the endothelial layer. (b) A frozen spleen section showing the heterogeneity of macrophage populations in the MZ. Sections were stained with anti-CD169 (green), anti-SIGN-R1 (blue) and anti-MARCO (red). WP, white pulp.

still limited tools to study these cells in detail. In addition to the clodronate liposome method mentioned above, several established genetic mouse models are commonly used to deplete MZ macrophages, including mice in which diphtheria toxin receptor (DTR) is restricted to CD11c⁺ or CD169⁺ cells. Whereas the later model provides higher level of selectivity for targeting macrophages (excluding DC subsets) (Miyake *et al.*, 2007), in both cases administration of DT leads to ablation of both metallophilic macrophages and SIGN-R1/MARCO macrophages. As a result, in many cases, it is not possible to determine which population of macrophage is responsible for specific functions. Additional more selective approaches are therefore needed to delineate the molecular mechanisms by which MZ macrophages regulate their unique niche-specific functions.

Metallophilic Macrophages

Metallophilic macrophages form a tight ring overlaying the marginal sinus lining that surrounds the B cell follicles (Figure 2(a)). They require macrophage colony-stimulating factor (M-CSF) for their development and are absent in *op/op* mice that lack M-CSF expression (Witmer-Pack *et al.*, 1993). Metallophilic-macrophages can be phenotypically distinguished from other macrophage subsets in the spleen by high surface expression of CD169. Because of their proximity to the follicles, they are able to closely associate with follicular B cells and these interactions provide a critical source of lymphotoxin (LT)- α 1 β 2, which is required for macrophage development and survival (Nolte *et al.*, 2004). In turn, metallophilic-

macrophages may provide B cells with important information and regulate their activation. Early studies have shown that humoral immunity against particulate bacterial antigens depends on macrophages in the MZ; when mice were treated with clodronate liposomes, antibody production was significantly reduced (Buiting *et al.*, 1996). These defects were partially normalized when metallophilic-macrophages were allowed to recover, which takes place before recovery of the SIGN-R1/MARCO macrophages, thus suggesting a role for these cells in the regulation of B cell responses (Buiting *et al.*, 1996). In addition to placing them close to the B cell follicle, the positioning of metallophilic-macrophages also provides them with access to blood antigens, therefore it seems likely that they will help to support humoral responses by promoting antigen delivery to B cells. However, unlike in LNs, where CD169⁺ subcapsular macrophages (SCMs) were shown to rally lymph antigens to naïve follicular B cells (discussed below), direct support for this model in the spleen is currently missing. A fundamental difference between the spleen and LNs is that in the latter, circulating B cells and lymph-borne antigens enter the tissue via different routes and their first encounter is in the follicles. In contrast, in the spleen B cells and systemic antigens enter through the same path and first reach the MZ, where at least two different subsets of macrophages reside. Thus, in this environment, antigen may be handled by different types of cells. This anatomical difference may reflect the spleen's unique role in protecting from blood-pathogens and the requirement for alternative strategies when facing systemically disseminated infections.

The contribution of MZ macrophages to B cell activation may go beyond their role in antigen presentation. In a recent report, clodronate liposome treated animals failed to develop germinal center B cell responses and this defect correlated with an inability of antigen specific B cells to migrate to the perimeter of splenic follicles (Nikbakht *et al.*, 2013). Localization of activated B cells at the follicle perimeter is part of the program that promotes their differentiation into germinal center B cells and this step brings them in close proximity to metallophilic-macrophages (Gatto and Brink, 2013). Migration of B cells to the follicles perimeter depends on activation of the G-protein coupled receptor GPR183 (Gatto and Brink, 2013; Pereira *et al.*, 2009) (also known as Epstein-Barr virus-induced gene, EBi2) by its ligand 7 α ,25-dihydroxycholesterol. In spleen and LN, the main source of 7 α ,25-dihydroxycholesterol is restricted to lymphoid stromal cells, which selectively express high levels of the enzymes required for its production (Arnon and Cyster, 2014; Gatto and Brink, 2013). Since clodronate liposome treatment causes global changes in the tissue's architecture, especially at the early time points after its administration, changes in stromal cell populations may have led to loss of EBi2 ligands and contributed to the observed impaired B cell migration. Importantly, the failure of animals depleted of MZ macrophages to form germinal centers in this study significantly exceeded the defects associated with EBi2 deficiency (Gatto and Brink, 2013; Pereira *et al.*, 2009), thus suggesting additional mechanisms are probably involved.

Although early data indicated that metallophilic-macrophages are dispensable for T cell activation (Aichele *et al.*, 2003), recent studies show that this may not always be the case and that metallophilic macrophages may be involved in transferring antigenic material to nearby DCs, thus facilitating cross-presentation to cytotoxic CD8⁺ T cells (Backer *et al.*, 2010). By immunizing mice with the antigen ovalbumin (OVA) conjugated to α CD169 antibody (α CD169-OVA), Backer *et al.* demonstrated that selective targeting of OVA to metallophilic macrophages leads to enhanced ability of splenocytes to prime cognate CD8⁺ T cells and trigger their cytotoxicity *in vivo*. This was mediated via cross-presentation of OVA peptides by DCs that efficiently captured the antigen from macrophages and presented it to CD8⁺ T cells. Importantly, OVA specific CD4⁺ T cells responses were not affected, suggesting a selective role for metallophilic-macrophages in regulating cytotoxic T cell activation. The way by which antigens are selectively transferred from metallophilic-macrophages to DCs is not clear, but it is possible that cellular migration is involved since treatment of mice with pertussis toxin (which blocks G-protein coupled receptor signaling and therefore inhibits cell movement) prevented this process (Backer *et al.*, 2010). Similar results were obtained when OVA was selectively targeted to metallophilic-macrophages using OVA-containing liposomal nanoparticles coated with high-affinity glycan ligands of CD169 (Chen *et al.*, 2012). Surprisingly, however, although immunization with α SIGN-R1-OVA did not elicit T cell activation, conjugation of OVA to MARCO antibody did. This is unexpected since MARCO is highly expressed on SIGN-R1⁺ macrophages. Thus, it is possible that the lower affinity of the α SIGN-R1-OVA conjugates for SIGN-R1, or temporal downregulation of the

receptor upon interaction with its antibody (as previously shown Kang *et al.* (2004)), may have reduced the stimulatory capacity of the complex. Alternatively, α MARCO-OVA immunization may have selectively targeted a specific subset of macrophages that expresses MARCO, but lacks SIGN-R1 (described below).

Importantly, in both pLNs and spleen, CD169⁺ macrophages are prone to infection by viruses as well as other pathogens. Infected CD169⁺ macrophages have been shown to provide an essential source of IFN- α that is critical for preventing viral dissemination and death (Honke *et al.*, 2012; Iannacone *et al.*, 2010). These findings proposed the possibility that local proliferation of pathogens within CD169⁺ macrophages may be advantageous to the host. In line with this hypothesis, a recent publication showed that CD169⁺ macrophages may in fact actively promote viral replication; in the spleen, infection of metallophilic-macrophages by vesicular stomatitis virus (VSV) was enhanced by expression of *Usp18*, a host protein that is preferentially expressed in these cells and inhibits phosphorylation of IFN- α/β receptor, thus increasing cellular permissiveness to infection (Honke *et al.*, 2012). In this study, infection of metallophilic-macrophages was not only critical for production of IFN- α , but also enhanced adaptive immunity. This suggest that restricted viral replication in macrophages may promote B and T cell activation by providing a local source of antigenic material. In support for this possibility, poor infection of metallophilic-macrophages in *Usp18*^{-/-} mice correlated with diminished antibody and CD4⁺ T cell responses (Honke *et al.*, 2012). It should be noted, however, that *Usp18* is also abundant in DCs (Honke *et al.*, 2013) and that reduced infection of DCs in *Usp18* knockout animals may have contributed to impaired activation of B and T cell responses in this study. Thus, selective removal of *Usp18* from specific subsets of myeloid cells may help to better understand the importance of this mechanism to host defense.

SIGN-R1/MARCO (MZ) Macrophages

SIGN-R1/MARCO macrophages are best identified by their positioning at the interface between the MZ and red pulp and by elevated expression of SIGN-R1 and MARCO receptors. In common with the metallophilic-macrophages, they require B cells and (LT)- α 1 β 2 for their development. However, unlike metallophilic-macrophages, these cells are excluded from direct contact with the follicles and are therefore restricted from accessing naïve follicular B cells. Instead, they interact with MZB cells that dwell in this compartment (Martin and Kearney, 2002). This unique population of cells migrates in the MZ from the underlying follicle in an integrin-dependent manner and is able to overcome the shear forces exerted by the blood flow in this region (Arnon and Cyster, 2014). Under steady state conditions, MZB cells move back and forth between the MZ and follicles in a perpetual movement, such that at any given time approximately half the cells reside within the MZ and half are in the follicles. This allows MZB cells to scan both compartments and to transfer information from one side to the other (Arnon and Cyster, 2014; Cyster and Schwab, 2012). Recently, we have established a method for visualizing cell movement within the MZ of live anesthetized mice using

2-photon microscopy. Using this approach, we visualized for the first time the highly motile nature of MZB cells and the tight interactions they form with SIGN-R1/MARCO macrophages (Arnon *et al.*, 2013); in many cases, MZB cells were observed to closely associate with immune complex-coated SIGN-R1/MARCO macrophages and appeared to scan their processes and cell bodies (Movie 1). Indeed, unlike metallophilic-macrophages, SIGN-R1/MARCO macrophages express high levels of complement receptors and this allows them to efficiently trap antibody-coated antigens within minutes after entering the blood stream. Expression of SIGN-R1 may also promote complement fixation and opsonization of antigens by inducing an unconventional pathway for C3 fixation, a process that was shown to be Ig-independent and to enhance complement-mediated innate resistance to pneumococci (Kang *et al.*, 2006). Immune complexes captured by SIGN-R1/MARCO macrophages are subsequently transferred to MZB cells, which carry them into the follicles and deposit them on follicular dendritic cells (FDCs) (Cyster and Schwab, 2012). Since antigen presentation by FDCs is important for germinal center B cell activation, this process may help to promote humoral immunity and the generation of high-affinity antibodies.

SIGN-R1/MARCO macrophages are not only important for providing MZB cells with antigenic material, but may also regulate their migration. Previous studies have shown that in the absence of SIGN-R1/MARCO macrophages, MZB cells were unable to localize to the MZ and their numbers were significantly reduced (Karlsson *et al.*, 2003; Noella *et al.*, 2013). Since development of MZB does not depend on access to the MZ (Cinamon *et al.*, 2004), it is likely that SIGN-R1/MARCO macrophages are not essential for their survival. Instead, they may be needed for retention (Karlsson *et al.*, 2003). The molecular basis of these observations has not been extensively explored, but one possibility is that gross changes in the architecture of the MZ due to loss of SIGN-R1/MARCO macrophages may compromise the development of local stromal cells and reduce expression of LFA-1 and VCAM-1, the integrin-ligands that regulate MZB cell adhesion (Arnon and Cyster, 2014; Cyster and Schwab, 2012). Alternatively, direct interactions with SIGN-R1/MARCO macrophages may help to restrict MZB cells movement and confine them to the MZ. Intravital imaging of MZB cell behavior in the absence of SIGN-R1/MARCO macrophages will help to differentiate between these possibilities.

While SIGN-R1/MARCO macrophages are typically referred to as a single population of cells, closer examination shows that they are in fact heterogeneous and that in addition to SIGNR-1⁺MARCO⁺ cells, the outer rim of the MZ also contains a subset of macrophages that lacks SIGNR-1 but expresses the MARCO receptor (Figure 2(b)). Notably, whereas SIGN-R1⁺MARCO⁺ macrophages depend on M-CSF for their development and require continuous interactions with MZB cells for survival (You *et al.*, 2009, 2011), SIGNR1[−]MARCO⁺ are independent of these factors, thus supporting the existence of two distinct populations of cells. Little is known about the functional specification of the SIGNR-1[−]MARCO⁺ macrophages and more studies are required to understand the importance of these cells to homeostatic mechanisms and to the regulation of immune responses.

Clearance of Apoptotic Cells by MZ Macrophages

The efficient removal of apoptotic cells is fundamental for the maintenance of healthy tissues and defects in apoptotic cell clearance are associated with immunological disorders and development of autoimmune diseases. Resident macrophages play central roles in this process and express a range of surface receptors that mediate the recognition and uptake of apoptotic cells in various tissues (Henson and Hume, 2006; Taylor *et al.*, 2005).

In the spleen, deletion of both MZ macrophage subsets impairs systemic tolerance toward apoptotic cells and accelerates the development and severity of autoimmune diseases (McGaha *et al.*, 2011; Miyake *et al.*, 2007; Wermeling *et al.*, 2009). The mechanisms by which MZ macrophages mediate these effects are only beginning to emerge, but may involve (1) direct clearance of apoptotic cells, (2) the transfer of apoptotic material to antigen presenting cells with suppressor capacity, and (3) the secretion of anti-inflammatory mediators. When MZ macrophages were removed, the spatial localization of apoptotic cells was altered such that dead cells were no longer confined to the MZ, but were accumulated in the T zone (McGaha *et al.*, 2011). Previous studies identified a unique subset of CD8 α ⁺ DCs that is able to capture dead cells in the splenic MZ, migrate into the T zone and inhibit T cell responses (Iyoda *et al.*, 2002; Qiu *et al.*, 2009). Thus, when MZ macrophages were depleted, apoptotic cells were no longer engulfed by CD8 α ⁺ DCs and instead were preferentially captured by CD8[−] DCs. This correlated with enhanced priming of cytotoxic T cells and a loss of tolerance (Miyake *et al.*, 2007). Systemic administration of apoptotic cells also triggered expression of the immunoregulatory enzyme Indoleamine 2,3 dioxygenase (IDO) in MZ macrophages (Ravishankar *et al.*, 2012). Inhibition of IDO using genetic or pharmacological approaches increased production of proinflammatory cytokines and enhanced effector T cell responses toward apoptotic cells (Ravishankar *et al.*, 2012). Although these studies were not able to specify the relative contribution of each macrophage population in the MZ to these processes, a recent report has shown that injection of apoptotic cells triggers expression of CCL22 specifically in metallophilic-macrophages (Ravishankar *et al.*, 2014). CCL22 mediated accumulation of FoxP3⁺ Tregs and blocking this pathway led to a strong proinflammatory response and breakdown of systemic tolerance (Ravishankar *et al.*, 2014). Importantly, however, since both metallophilic-macrophages and SIGN-R1/MARCO macrophages express a wide array of receptors that can recognize apoptotic and damaged cells, it is likely that both subsets are involved in maintaining peripheral tolerance in the spleen. In line with this notion, loss of expression of MARCO or scavenger receptor A (SR-A), which are abundant on SIGN-R1/MARCO macrophages, is sufficient to reduce the tolerance threshold and to increase anti-DNA antibody titers (Wermeling *et al.*, 2007).

Red Pulp Macrophages

The red pulp represents the majority of the spleen's volume. It lacks the well-defined organization of the white pulp and it is composed of a meshwork of sinuses and cords, which

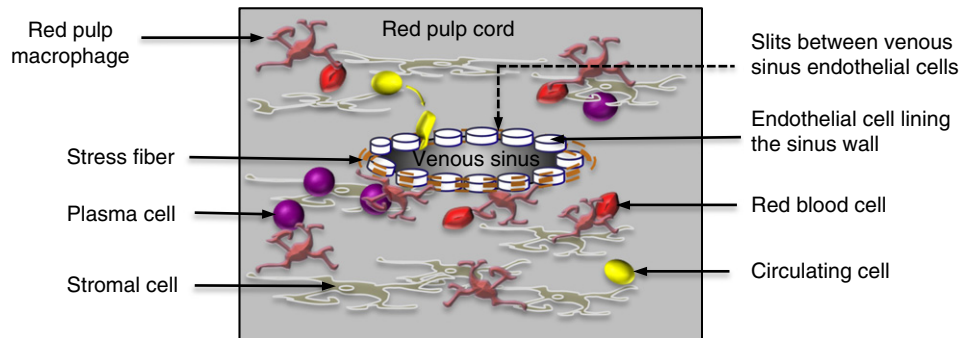


Figure 3 Schematic description of cross-section of a red pulp cord with a venous sinus located in the middle. The cord is supported by a network of stromal cells (light brown). The sinus wall is composed of elongated endothelial cells (white) that are connected by stress fibers (orange dotted line). The fibers are contractible, allowing the opening of slits between the endothelial cells (pointed at by dashed arrow). Through these slits, red blood cells (red), and migrating cells (yellow) pass as they egress back to the systemic blood circulation. An example of a cell entering the sinus via a slit is illustrated (yellow arrow). Red pulp macrophages (pink) are located in the cord and some closely associate with the sinus wall. Aged or damaged red blood cells that become too stiff to squeeze between the slits are quickly phagocytosed by nearby macrophages. Following immunization, plasma cells (purple) migrate to the red pulp cord, where they may come to close contact with local macrophages.

together form an open blood system. Blood arrives to the red pulp after entering via the marginal sinuses and passing through the MZ. While most of the blood enters via this path, some terminal arteriole end within the red pulp and release their contents directly into the splenic cords. Venous sinuses that are scattered throughout the red pulp provide exit sites for blood and circulating cells that are returning back to circulation (Mebius and Kraal, 2005). The red pulp macrophages (RPMs) are located within the splenic cords, in close association with red blood cells and plasma cells (Figure 3). They develop prior to birth and are maintained throughout adulthood from a local progenitor, independently from circulating monocytes and bone marrow-derived hematopoietic cells (Hashimoto *et al.*, 2013; Yona *et al.*, 2013). Although under steady state only few RPMs actively divide, these cells are able to rapidly proliferate and expand in response to cell-targeted deletion or inflammation-induced ablation, thus demonstrating their capacity to self-renew under normal and pathological conditions (Hashimoto *et al.*, 2013; Yona *et al.*, 2013).

RPMs can be recognized in the spleen by the expression of $F4/80^+ CD206^+ CD11b^{lo}$ and $Vcam1^+$ (Taylor *et al.*, 2005). They play an important role in the regulation of erythrocytes turnover and iron recycling (Gordon and Taylor, 2005; Murray and Wynn, 2011; Taylor *et al.*, 2005). Damaged erythrocytes that are restricted from passage through the venous sinuses become phagocytosed and degraded by nearby macrophages. This process leads to the release of iron, which is important for the production of hemoglobin and other enzymes in new erythrocytes. In addition, RPMs express high levels of proteins that are involved in hemoglobin uptake, breakdown of heme and export of iron, thus providing an important reservoir of readily available iron that can be released to the plasma upon demand (Ganz, 2012). Since many extracellular pathogens require iron for their growth, restricting its availability by RPMs is thought to contribute to host defense mechanisms and attenuation of microbial proliferation. The mechanisms that control iron release from RPMs are only partially understood and more studies are needed to determine the signals involved and the importance of local versus systemic stimuli in regulating this process.

In line with their important role in iron metabolism, RPM development depends on the transcription factor Spi-c (Kohyama *et al.*, 2009), which is itself induced by one of the main products of hemoglobin degradation, heme (Haldar *et al.*, 2014). A similar mechanism was found to regulate the development of $F4/80^+ Vcam1^+$ macrophages in the bone marrow (Haldar *et al.*, 2014), thus demonstrating the importance of metabolites in driving the differentiation of tissue-resident macrophages. While all RPM depend on Spi-c for their development, only about half require M-CSF (Cecchini *et al.*, 1994), suggesting RPMs are likely composed of a heterogeneous population of cells. The phenotypic and functional differences between M-CSF dependent and independent RPM are not well understood. A recent study showed that in *op/op* mice a significant loss of RPM expressing low levels of complement receptor 3 (Mac-1) was observed, whereas $F4/80^+ Mac-1^{high}$ macrophages were still detected (Kurotaki *et al.*, 2011). *In vitro* assessment of $F4/80^+ Mac-1^{lo}$ cells in this study further suggested a possible role for this distinct macrophage population in suppression of $CD4^+$ T cell responses and induction of regulatory T cells (Kurotaki *et al.*, 2011). Further studies are required to better understand the heterogeneity among RPM populations and whether or not distinct subsets can be associated with specific anatomical niches within the red pulp.

Despite their potent phagocytic capacity and ability to mediate uptake of antigens, a dominant role for RPMs in protection against invading pathogens has been difficult to establish. For example, during exposure to malaria, infected erythrocytes and parasites are taken up by RPMs, leading to production of high amounts of chemokines and type I IFN (Kim *et al.*, 2012). However, these mechanisms were not essential for parasite clearance and *Spic*^{-/-} mice that selectively lack RPMs display a similar kinetics of malaria infection compared with WT animals (Kim *et al.*, 2012). Thus, while RPMs are able to respond and eliminate infected cells, their contribution to innate immunity may be largely redundant and is likely dominated by other more specialized macrophage subsets in the spleen. In support for this notion RPMs were found to critically contribute to the

clearance of *Streptococcus pneumonia* infection under conditions in which SIGN-R1⁺ macrophages were depleted (Kirby *et al.*, 2009).

LN Macrophages

LNs are distributed throughout the body and filter lymph from nearby tissues. Afferent lymphatic vessels enter the LN and bring tissue-derived antigens and cells (such as DCs) into the subcapsular sinus (SCS), a narrow space located below the LN capsule (Figure 4). The SCS surround the B cell follicles and the paracortical region where the T zone is located. The floor of the SCS is composed of lymphatic endothelial cells, which form a physical barrier that prevents free passage of large particles into the lymphoid tissue underneath. From the entry point at the SCS, the lymph flows into medullary and blind-ended cortical sinuses, before egressing the organ via the efferent lymphatic vessel. Circulating lymphocytes that enter the LNs from the blood do so by transmigration across the high endothelial venules (HEVs) in the paracortex. The B cells follicles are located underneath the SCS, at the cortical area, whereas T cells are positioned deeper in the paracortical region, which is centrally located between the medulla and cortex (Figure 4). The two best-described populations of resident macrophages in pLNs are found in close proximity to the subcapsular and medullary sinuses (Gray and Cyster, 2012). These macrophage subsets can be distinguished not

only based on their differential positioning but also by their phenotypic markers and functions.

SCMs

SCMs in LNs share many similarities with splenic metallophilic-macrophages; both cell types are located at the interface between the B cell follicles and sites of antigen entry, express high levels of CD169 and depend on M-CSF and B cell-derived LT α 1 β 2 signals for their development (Phan *et al.*, 2009). However, unlike metallophilic-macrophages, SCMs do not require LXR signaling and therefore develop normally in LXR deficient mice (Noella *et al.*, 2013). Although SCMs are efficiently replaced by donor bone marrow-derived cells in irradiated hosts (Gray *et al.*, 2012), it is currently not known whether these cells depend on circulating monocytes for their maintenance under 'normal' conditions or whether they instead renew locally under steady state. In human patients suffering from a RF8-null mutation that blocks monocyte and DC differentiation, LN macrophages develop normally, thus suggesting circulating monocytes may not be necessary for the survival of LN resident macrophages, at least in humans (Hambleton *et al.*, 2011).

Early histology studies have shown that shortly after subcutaneous injection of antigens, immune complexes captured by SCMs accumulate close to their contact area with follicular B cells. These observations raised the possibility that SCMs are important for the active transport of antigens across the

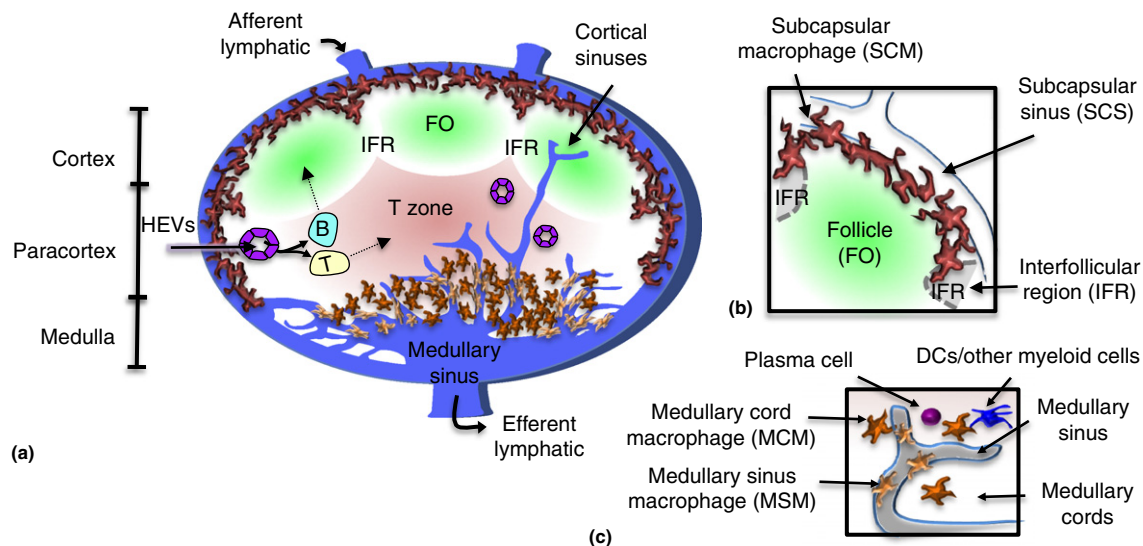


Figure 4 Illustration of lymph node compartments and localization of major macrophage populations. (a) The lymph node can be divided into three main regions; the cortex, the paracortex, and the medulla. Most circulating naive B and T cells arrive to the LN from the blood and enter by crossing the high endothelial venules (HEVs). Upon entry, B cells migrate to the follicles (FO) and T cells to the T-zone. In contrast to the path by which lymphocytes enter the LN, lymph arrives through the afferent lymphatic into the subcapsular sinuses that surround the follicles. From there, the lymph flows to the medullary sinuses, which are connected to the efferent vessel. LN macrophages are located at the entry and exit ports of the lymph; under the subcapsular sinuses (subcapsular macrophages, SCMs) and in the medulla (medullary macrophages). Some SCMs extend to the interfollicular regions (IFRs) that are located in the junctions between follicles and the T-zone. (b) A close up on the subcapsular region. The SCMs are shown in red, the follicles in green and the interfollicular regions are in grey. (c) A close up on a section from the medulla showing a medullary sinus (grey) surrounded by a cord. Macrophages can be found on both sides of the sinus wall. Based on a recent suggested nomenclature, we highlight the medullary macrophages that reside inside the sinuses (medullary sinus macrophages, light orange) and those that are in the cords (medullary cord macrophages, MCMs, orange). Plasma cells (purple), DCs (blue) as well as other myeloid cells can be found in the medullary cords.

SCS and for early stages in B cell activation. This notion has now been validated in a series of elegant studies using histology-based analysis (Pape *et al.*, 2007) and two-photon microscopy (Carrasco and Batista, 2007; Junt *et al.*, 2007; Phan *et al.*, 2007); in these studies, SCMs were shown to capture large antigen immune complexes from the afferent lymphatics and to transfer them into the underlying follicles. Follicular B cells were seen to migrate in close proximity to SCMs and were frequently found to scan their processes. These encounters allowed non-cognate B cells to capture immune complexes from SCMs and to transfer them at their uropod into the follicles in a CR2-dependent manner. Cognate B cells that came in contact with antigen-coated SCMs reduced their migration velocity with many forming long-lasting interactions, suggesting local activation was induced. In some cases, antigenic material was observed to move in a unidirectional fashion from the lymph exposed SCM cell body to the cell's processes that reach through the SCS into the follicle parenchyma. However, it remains unclear whether the antigen simply 'slides' along the macrophage body or is being internalized into endosomes and recycled back to the surface (Batista and Harwood, 2009; Cyster, 2010). Notably, SCMs have relatively limited phagocytic capacity and are able to retain intact antigens on the cell surface for long periods of time, thus supporting their role in antigen presentation to B cells (Batista and Harwood, 2009; Cyster, 2010). It is likely that acquisition of antigens is mediated by multiple receptors with partially overlapping specificities (Batista and Harwood, 2009; Cyster, 2010; Phan *et al.*, 2009). Interestingly, CD169 was recently shown to bind antigen-containing exosomes (Saunderson *et al.*, 2014), suggesting a previously unappreciated pathway for delivering antigenic material to LNs and splenic follicles.

Despite the important insights discussed above, the role of SCMs in promoting humoral immunity remains unclear, and when clodronate liposome treatment was used to remove all local macrophage subsets in pLNs, germinal center responses were not reduced and production of high-affinity antibodies was either not impaired (Gonzalez *et al.*, 2010; Iannacone *et al.*, 2010) or only marginally affected (Phan *et al.*, 2009). In some cases, antibody responses were even enhanced (Gonzalez *et al.*, 2010). Given the tissue damage associated with treatments that remove tissue macrophages, it is possible that reduced active transfer of antigens is compensated by excessive leakage of antigen across the damaged tissue (Cyster, 2010). Thus, additional approaches to selectively ablate or perturb macrophage functions are needed to clarify the full extent of their contribution to B cell activation in pLNs and other sites.

SCMs have also emerged as important regulators of cytotoxic T cells in pLNs, particularly during recall responses. Upon infection with lymphocytic choriomeningitis virus (LCMV) (Kastenmuller *et al.*, 2013), vaccine virus (VV) (Sung *et al.*, 2012), or the parasite *Toxoplasma gondii* (Chtanova *et al.*, 2009), central memory T cells rapidly accumulated at peripheral sites near the SCS, medulla, and interfollicular regions of the LN, where significant infection of macrophages was observed. Recruitment of memory CD8⁺ T cells to the infected sites required expression of CXCR3 ligands by a number of cells including CD169⁺ macrophages (Kastenmuller *et al.*, 2013; Sung *et al.*, 2012). Given that SCMs are thought to be

the main population of cells that becomes infected in the first few hours after infection, their contribution to early recruitment of memory T cells to infected sites may be critical. This mechanism may not only inhibit systemic dissemination of pathogens, but could also help to prevent local damage to the LN. It is not fully understood whether SCMs prime CD8⁺ T cells by directly presenting processed peptides via MHC class I molecules or indirectly, by transferring antigens to nearby DCs. Indeed, during inflammation, naive CD8⁺ T cells up-regulated expression of CCR5, which preferentially enhanced migration toward infected DCs. When CD11c⁺ DCs were removed, frequency and duration of contacts between T cells and SCMs were increased, but these interactions were only partially functional and resulted with suboptimal activation of T cells (Hickman *et al.*, 2011). Thus, more studies are needed to clarify the role of SCMs in direct priming of naive and memory T cells. Importantly, the nature of the antigen itself may also influence the stimulatory potential of SCMs. For example, in pLNs, visualization of interactions between SCMs and NKT cells using 2-photon microscopy demonstrated that SCMs are able to take up and directly present lipid antigens to NKT cells, thus establishing their capacity to function as antigen presenting cells (Barral *et al.*, 2010). In the spleen, CD169⁺ macrophages may function similarly (Kawasaki *et al.*, 2013), although in this study direct visualization of macrophages-NKT cell interactions was not described and the possibility that other indirect mechanisms are involved could not be excluded.

Medullary Macrophages

The medullary sinuses form a large irregular structure of winding tunnels, which are surrounded by medullary cords. These structures are filled with lymphatic fluids and are populated with a niche-specific macrophage subset known as the medullary macrophages (Gray and Cyster, 2012). The medullary sinus macrophages (MSMs) can be identified by their location inside the medullary sinuses and by their high expression of F4/80. In addition, many medullary macrophages also express CD169 (although at lower levels compared with SCMs), SIGN-R1, MARCO, and LYVE1 (Gray and Cyster, 2012). In contrast to SCMs, development of medullary macrophages is independent of M-CSF and they develop normally in op/op mice (Witmer-Pack *et al.*, 1993), confirming their identity as a distinct population of cells. Although like SCMs medullary macrophages are equipped with many surface receptors that can mediate the capture of bacterial, viral, and other pathogens, they contain a large number of lysosomes and are highly phagocytic (Taylor *et al.*, 2005; Phan *et al.*, 2009). As such, medullary macrophages are thought to play a more dominant role in degrading rather than presenting antigenic material to lymphocytes in LNs. This function may help to filter the lymph and prevent dissemination of pathogens to the systemic circulation. Notably, however, medullary macrophages are not impermeable and in some cases soluble proteins or other antigens that have escaped the SCMs may pass the medullary macrophages and reach the medullary cords (Gonzalez *et al.*, 2010), where nearby cells including local DCs may capture it and promote humoral immune responses. However, whether medullary

macrophages are actively involved in this process has not been established.

The medullary sinuses are connected to the efferent lymphatic vessels, where cells and fluids egress from the LN (Girard *et al.*, 2012). While these sinuses provide the path that leads lymphocytes out of the organ, in many cases, lymphocytes use the sinuses located in the cortical region to initiate egress. Live imaging and histology-based studies have provided a detailed analysis of this process and established a critical role for the sphingosine-1-phosphate receptor 1 (S1PR1) expressed on the surface of B and T cells in promoting cell migration across the lymphatic sinuses (Cyster and Schwab, 2012). Although the relative frequency of lymphocytes egressing directly via the medullary sinuses was not assessed and visualization of this process was not performed (Grigorova *et al.*, 2009), the observation that S1PR1-deficient T cells accumulate close to the medullary sinuses strongly suggest that this may be an important egress site where a similar mechanism of exit applies (Cyster and Schwab, 2012). Importantly, the main cellular source of the S1PR1-ligand, S1P, for egressing lymphocytes in pLNs is expressed by LYVE1⁺ lymphatic endothelial cells (Pham *et al.*, 2010). In mice in which the enzymes responsible for S1P production (sphingosine kinase 1 and 2) were selectively removed from LYVE1⁺ cells, T cell egress was significantly impaired. This defect was independent of LYVE1⁺ medullary macrophages and other hematopoietic cells since reconstitution with WT bone marrow did not restore normal egress (Pham *et al.*, 2010). Thus, medullary macrophages are not a major source for S1P for lymphocyte egress from pLNs. Interestingly, recent imaging studies revealed that the medullary sinuses are not only an exit site, but may also function as an entry path for T cells that arrive from the afferent lymphatic. Under steady state conditions, T cells preferentially used this path, whereas DCs migrated across the SCS (Girard *et al.*, 2012). Since most of the T cells that enter pLNs via the lymphatic system are primarily effector or memory cells, migration across the medullary region may provide an efficient way to rapidly scan activated medullary macrophages, where early pathogen detection occurs.

Despite being a significant distance from the follicles, medullary macrophages may play a role in supporting humoral immunity by promoting plasma cell survival. During immune responses, the medullary cords become packed with newly generated plasma cells and plasmablasts, many of which were reported as being in close proximity to myeloid cell lineages, including medullary macrophages (Mohr *et al.*, 2009). RT-PCR analysis showed that a few days after antigen immunization in alum, CD169⁺CD11b^{high}F4/80⁺ macrophages located at the centre of the medullary cords produced high amounts of IL-6 and APRIL. These factors are important for plasma cell survival and maturation, suggesting a potential role for medullary macrophages in providing trophic support to these cells (Mohr *et al.*, 2009). Live imaging studies have confirmed that plasma cells often form long-term interactions with LysM⁺ cells in the medulla, supporting a potential cross talk with local macrophages. Interestingly, in contrast to the proposed function of medullary macrophages in promoting plasma cell survival, in this study medullary macrophages were suggested to negatively regulate plasma cells leading to reduction in antibody

production and decreases in the number of plasma cells (Fooksman *et al.*, 2014). Notably, however, in this report adjuvant-free vaccination was applied, suggesting different modes of activation may differentially affect macrophages functions and interactions with plasma cells.

The level of heterogeneity between medullary macrophages is not fully understood and it is likely that these cells are composed of more than one subset. A recent review suggested a convenient nomenclature for classifying the cells based on their positioning differentiating between macrophages that reside inside the medullary sinus (MSMs) and those that are located in the medullary cords (medullary cord macrophages, MCMs) (Figure 2; Gray and Cyster, 2012). Given the tight link that exists between a macrophage's biology and their niche of residency, it is reasonable to speculate that these cell populations indeed represent different subsets and it would be important to resolve the unique characteristics and functions that define them.

Conclusions

Secondary lymphoid organs are highly organized structures in which antigenic material and rare lymphocytes come together in a temporally and spatially controlled manner to produce an efficient and well-balanced immune response. In addition to their important role in regulation of adaptive immunity, secondary lymphoid organs also function as biological filters, actively removing pathogens from lymph and blood. As such, these structures provide an important barrier that restricts pathogen dissemination and prevents spreading of disease. Within pLNs and spleen, macrophages occupy peripheral niches that are located at the interface between the circulatory fluids and the lymphoid tissue. This positioning brings them in direct contact with newly arriving antigenic material, where initial encounters with pathogens occur. Seminal early studies have identified this strategic positioning and provided profound insights that established the important roles of these cells in elimination of invading pathogens and in handling antigenic material. However, it is only in recent years that we begun to fully appreciate the highly specialized nature of these cells and their contributions to specific immunological functions. While current data strongly support a direct role for macrophages in facilitating antigen delivery to B cells, NKT cell and potentially CD8⁺ T cells, the mechanisms by which these processes are regulated and the long-term consequences of these interactions are poorly understood. These questions remain particularly unexplored in the spleen, where dynamic assessment of macrophage behavior using intravital imaging approaches has not been possible until recently, thus limiting analysis of interactions between splenic macrophages and immune cells. Uncovering mechanisms by which macrophages contribute to systemic versus local immune responses requires further investigation and better understanding of these processes may prove useful for future development of novel vaccine strategies.

More studies are also needed to delineate the basis for the complex heterogeneity of tissue macrophages and the cues that promote their specification *in vivo*. Revealing such factors will not only help to explain how the environment modifies the

immune system, but also to uncover the unique characteristics and requirements of the niche itself. Such knowledge may further help to identify specific transcriptional programs and may enable the development of new approaches to selectively manipulate distinct macrophage populations. Indeed, a major obstacle that continues to challenge the field and to prevent detailed mechanistic understanding of these cells is the absence of mouse models that selectively target individual macrophage types and the difficulties associated with their purification. Developing such tools and improving isolation protocols will undoubtedly enhance our ability to explore these cells *in vivo* and to obtain a comprehensive understanding of their biology and functions in the spleen, pLNs, and other sites.

See also: Cellular Immunology: Cell–Cell Interactions and the Immune System: Roles of Stromal Cells in the Immune System

References

- Aichele, P., Zinke, J., Grode, L., *et al.*, 2003. Macrophages of the splenic marginal zone are essential for trapping of blood-borne particulate antigen but dispensable for induction of specific T cell responses. *Journal of Immunology* 171, 1148–1155.
- Arnon, T.I., Cyster, J.G., 2014. Blood, sphingosine-1-phosphate and lymphocyte migration dynamics in the spleen. *Current Topics in Microbiology and Immunology* 378, 107–128.
- Arnon, T.I., Horton, R.M., Grigorova, I.L., Cyster, J.G., 2013. Visualization of splenic marginal zone B-cell shuttling and follicular B-cell egress. *Nature* 493, 684–688.
- Backer, R., Schwandt, T., Greuter, M., *et al.*, 2010. Effective collaboration between marginal metallophilic macrophages and CD8⁺ dendritic cells in the generation of cytotoxic T cells. *Proceedings of the National Academy of Sciences of the United States of America* 107, 216–221.
- Barral, P., Polzella, P., Bruckbauer, A., *et al.*, 2010. CD169(+) macrophages present lipid antigens to mediate early activation of iNKT cells in lymph nodes. *Nature Immunology* 11, 303–312.
- Batista, F.D., Harwood, N.E., 2009. The who, how and where of antigen presentation to B cells. *Nature Reviews Immunology* 9, 15–27.
- Buiting, A.M., De Rover, Z., Kraal, G., van Rooijen, N., 1996. Humoral immune responses against particulate bacterial antigens are dependent on marginal metallophilic macrophages in the spleen. *Scandinavian Journal of Immunology* 43, 398–405.
- Carrasco, Y.R., Batista, F.D., 2007. B cells acquire particulate antigen in a macrophage-rich area at the boundary between the follicle and the subcapsular sinus of the lymph node. *Immunity* 27, 160–171.
- Cecchini, M.G., Dominguez, M.G., Mocchi, S., *et al.*, 1994. Role of colony stimulating factor-1 in the establishment and regulation of tissue macrophages during postnatal development of the mouse. *Development* 120, 1357–1372.
- Chen, W.C., Kawasaki, N., Nycholat, C.M., *et al.*, 2012. Antigen delivery to macrophages using liposomal nanoparticles targeting sialoadhesin/CD169. *PLoS One* 7, e39039.
- Chtanova, T., Han, S.J., Schaeffer, M., *et al.*, 2009. Dynamics of T cell, antigen-presenting cell, and pathogen interactions during recall responses in the lymph node. *Immunity* 31, 342–355.
- Cinamon, G., Matloubian, M., Lesneski, M.J., *et al.*, 2004. Sphingosine 1-phosphate receptor 1 promotes B cell localization in the splenic marginal zone. *Nature Immunology* 5, 713–720.
- Cyster, J.G., 2010. B cell follicles and antigen encounters of the third kind. *Nature Immunology* 11, 989–996.
- Cyster, J.G., Schwab, S.R., 2012. Sphingosine-1-phosphate and lymphocyte egress from lymphoid organs. *Annual Review of Immunology* 30, 69–94.
- Davies, L.C., Jenkins, S.J., Allen, J.E., Taylor, P.R., 2013. Tissue-resident macrophages. *Nature Immunology* 14, 986–995.
- Fooksman, D.R., Nussenzweig, M.C., Dustin, M.L., 2014. Myeloid cells limit production of antibody-secreting cells after immunization in the lymph node. *Journal of Immunology* 192, 1004–1012.
- Ganz, T., 2012. Macrophages and systemic iron homeostasis. *Journal of Innate Immunity* 4, 446–453.
- Gatto, D., Brink, R., 2013. B cell localization: Regulation by EBI2 and its oxysterol ligand. *Trends in Immunology* 34, 336–341.
- Gautier, E.L., Shay, T., Miller, J., *et al.*, 2012. Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nature Immunology* 13, 1118–1128.
- Girard, J.P., Moussion, C., Forster, R., 2012. HEVs, lymphatics and homeostatic immune cell trafficking in lymph nodes. *Nature Reviews Immunology* 12, 762–773.
- Gonzalez, S.F., Lukacs-Kornek, V., Kuligowski, M.P., *et al.*, 2010. Capture of influenza by medullary dendritic cells via SIGN-R1 is essential for humoral immunity in draining lymph nodes. *Nature Immunology* 11, 427–434.
- Gordon, S., Taylor, P.R., 2005. Monocyte and macrophage heterogeneity. *Nature Reviews Immunology* 5, 953–964.
- Gray, E.E., Cyster, J.G., 2012. Lymph node macrophages. *Journal of Innate Immunity* 4, 424–436.
- Gray, E.E., Friend, S., Suzuki, K., Phan, T.G., Cyster, J.G., 2012. Subcapsular sinus macrophage fragmentation and CD169 + bleb acquisition by closely associated IL-17-committed innate-like lymphocytes. *PLoS One* 7, e38258.
- Grigorova, I.L., Schwab, S.R., Phan, T.G., *et al.*, 2009. Cortical sinus probing, S1P1-dependent entry and flow-based capture of egressing T cells. *Nature Immunology* 10, 58–65.
- Haldar, M., Kohyama, M., So, A.Y., *et al.*, 2014. Heme-mediated SPI-C induction promotes monocyte differentiation into iron-recycling macrophages. *Cell* 156, 1223–1234.
- Hambleton, S., Salem, S., Bustamante, J., *et al.*, 2011. IRF8 mutations and human dendritic-cell immunodeficiency. *New England Journal of Medicine* 365, 127–138.
- Hashimoto, D., Chow, A., Noizat, C., *et al.*, 2013. Tissue-resident macrophages self-maintain locally throughout adult life with minimal contribution from circulating monocytes. *Immunity* 38, 792–804.
- Henson, P.M., Hume, D.A., 2006. Apoptotic cell removal in development and tissue homeostasis. *Trends in Immunology* 27, 244–250.
- Hickman, H.D., Li, L., Reynoso, G.V., *et al.*, 2011. Chemokines control naive CD8⁺ T cell selection of optimal lymph node antigen presenting cells. *Journal of Experimental Medicine* 208, 2511–2524.
- Honke, N., Shaabani, N., Cadeddu, G., *et al.*, 2012. Enforced viral replication activates adaptive immunity and is essential for the control of a cytopathic virus. *Nature Immunology* 13, 51–57.
- Honke, N., Shaabani, N., Zhang, D.E., *et al.*, 2013. Usp18 driven enforced viral replication in dendritic cells contributes to break of immunological tolerance in autoimmune diabetes. *PLoS Pathogens* 9, e1003650.
- Iannacone, M., Moseman, E.A., Tonti, E., *et al.*, 2010. Subcapsular sinus macrophages prevent CNS invasion on peripheral infection with a neurotropic virus. *Nature* 465, 1079–1083.
- Iyoda, T., Shimoyama, S., Liu, K., *et al.*, 2002. The CD8⁺ dendritic cell subset selectively endocytoses dying cells in culture and *in vivo*. *Journal of Experimental Medicine* 195, 1289–1302.
- Junt, T., Moseman, E.A., Iannacone, M., *et al.*, 2007. Subcapsular sinus macrophages in lymph nodes clear lymph-borne viruses and present them to antiviral B cells. *Nature* 450, 110–114.
- Kang, Y.S., Do, Y., Lee, H.K., *et al.*, 2006. A dominant complement fixation pathway for pneumococcal polysaccharides initiated by SIGN-R1 interacting with C1q. *Cell* 125, 47–58.
- Kang, Y.S., Kim, J.Y., Bruening, S.A., *et al.*, 2004. The C-type lectin SIGN-R1 mediates uptake of the capsular polysaccharide of *Streptococcus pneumoniae* in the marginal zone of mouse spleen. *Proceedings of the National Academy of Sciences of the United States of America* 101, 215–220.
- Karlsson, M.C., Guinamard, R., Bolland, S., *et al.*, 2003. Macrophages control the retention and trafficking of B lymphocytes in the splenic marginal zone. *Journal of Experimental Medicine* 198, 333–340.
- Kastenmuller, W., Brandes, M., Wang, Z., *et al.*, 2013. Peripheral prepositioning and local CXCL9 chemokine-mediated guidance orchestrate rapid memory CD8⁺ T cell responses in the lymph node. *Immunity* 38, 502–513.
- Kawasaki, N., Vela, J.L., Nycholat, C.M., *et al.*, 2013. Targeted delivery of lipid antigen to macrophages via the CD169/sialoadhesin endocytic pathway induces robust invariant natural killer T cell activation. *Proceedings of the National Academy of Sciences of the United States of America* 110, 7826–7831.
- Kim, C.C., Nelson, C.S., Wilson, E.B., *et al.*, 2012. Splenic red pulp macrophages produce type I interferons as early sentinels of malaria infection but are dispensable for control. *PLoS One* 7, e48126.

- Kirby, A.C., Beattie, L., Maroof, A., van Rooijen, N., Kaye, P.M., 2009. SIGNR1-negative red pulp macrophages protect against acute streptococcal sepsis after *Leishmania donovani*-induced loss of marginal zone macrophages. *American Journal of Pathology* 175, 1107–1115.
- Kohyama, M., Ise, W., Edelson, B.T., *et al.*, 2009. Role for Spi-C in the development of red pulp macrophages and splenic iron homeostasis. *Nature* 457, 318–321.
- Kurotaki, D., Kon, S., Bae, K., *et al.*, 2011. CSF-1-dependent red pulp macrophages regulate CD4 T cell responses. *Journal of Immunology* 186, 2229–2237.
- Macauley, M.S., Crocker, P.R., Paulson, J.C., 2014. Siglec-mediated regulation of immune cell function in disease. *Nature Reviews Immunology* 14, 653–666.
- Martin, F., Kearney, J.F., 2002. Marginal-zone B cells. *Nature Reviews Immunology* 2, 323–335.
- Martínez, F.O., Helming, L., Gordon, S., 2009. Alternative activation of macrophages: An immunologic functional perspective. *Annual Review of Immunology* 27, 451–483.
- McGaha, T.L., Chen, Y., Ravishankar, B., van Rooijen, N., Karlsson, M.C., 2011. Marginal zone macrophages suppress innate and adaptive immunity to apoptotic cells in the spleen. *Blood* 117, 5403–5412.
- Mebius, R.E., Kraal, G., 2005. Structure and function of the spleen. *Nature Reviews Immunology* 5, 606–616.
- Miyake, Y., Asano, K., Kaise, H., *et al.*, 2007. Critical role of macrophages in the marginal zone in the suppression of immune responses to apoptotic cell-associated antigens. *Journal of Clinical Investigation* 117, 2268–2278.
- Mohr, E., Serre, K., Manz, R.A., *et al.*, 2009. Dendritic cells and monocyte/macrophages that create the IL-6/APRIL-rich lymph node microenvironments where plasmablasts mature. *Journal of Immunology* 182, 2113–2123.
- Murray, P.J., Wynn, T.A., 2011. Protective and pathogenic functions of macrophage subsets. *Nature Reviews Immunology* 11, 723–737.
- Nikbakht, N., Shen, S., Manser, T., 2013. Cutting edge: Macrophages are required for localization of antigen-activated B cells to the follicular perimeter and the subsequent germinal center response. *Journal of Immunology* 190, 4923–4927.
- Noella, A.-G., Guillen, J.A., Gallardo, G., *et al.*, 2013. The nuclear receptor LXR α controls the functional specialization of splenic macrophages. *Nature Immunology* 14, 831–839.
- Nolte, M.A., Arens, R., Kraus, M., *et al.*, 2004. B cells are crucial for both development and maintenance of the splenic marginal zone. *Journal of Immunology* 172, 3620–3627.
- Pape, K.A., Catron, D.M., Itano, A.A., Jenkins, M.K., 2007. The humoral immune response is initiated in lymph nodes by B cells that acquire soluble antigen directly in the follicles. *Immunity* 26, 491–502.
- Pereira, J.P., Kelly, L.M., Xu, Y., Cyster, J.G., 2009. EBI2 mediates B cell segregation between the outer and centre follicle. *Nature* 460, 1122–1126.
- Pham, T.H., Baluk, P., Xu, Y., *et al.*, 2010. Lymphatic endothelial cell sphingosine kinase activity is required for lymphocyte egress and lymphatic patterning. *Journal of Experimental Medicine* 207, 17–27.
- Phan, T.G., Green, J.A., Gray, E.E., Xu, Y., Cyster, J.G., 2009. Immune complex relay by subcapsular sinus macrophages and noncognate B cells drives antibody affinity maturation. *Nature Immunology* 10, 786–793.
- Phan, T.G., Grigorova, I., Okada, T., Cyster, J.G., 2007. Subcapsular encounter and complement-dependent transport of immune complexes by lymph node B cells. *Nature Immunology* 8, 992–1000.
- Qiu, C.H., Miyake, Y., Kaise, H., *et al.*, 2009. Novel subset of CD8 α + dendritic cells localized in the marginal zone is responsible for tolerance to cell-associated antigens. *Journal of Immunology* 182, 4127–4136.
- Ravishankar, B., Liu, H., Shinde, R., *et al.*, 2012. Tolerance to apoptotic cells is regulated by indoleamine 2,3-dioxygenase. *Proceedings of the National Academy of Sciences of the United States of America* 109, 3909–3914.
- Ravishankar, B., Shinde, R., Liu, H., *et al.*, 2014. Marginal zone CD169 + macrophages coordinate apoptotic cell-driven cellular recruitment and tolerance. *Proceedings of the National Academy of Sciences of the United States of America* 111, 4215–4220.
- van Rooijen, N., Kors, N., Kraal, G., 1989. Macrophage subset repopulation in the spleen: Differential kinetics after liposome-mediated elimination. *Journal of Leukocyte Biology* 45, 97–104.
- Saunderson, S.C., Dunn, A.C., Crocker, P.R., McLellan, A.D., 2014. CD169 mediates the capture of exosomes in spleen and lymph node. *Blood* 123, 208–216.
- Spann, N.J., Garmire, L.X., McDonald, J.G., *et al.*, 2012. Regulated accumulation of desmosterol integrates macrophage lipid metabolism and inflammatory responses. *Cell* 151, 138–152.
- Sung, J.H., Zhang, H., Moseman, E.A., *et al.*, 2012. Chemokine guidance of central memory T cells is critical for antiviral recall responses in lymph nodes. *Cell* 150, 1249–1263.
- Taylor, P.R., Martínez-Pomares, L., Stacey, M., *et al.*, 2005. Macrophage receptors and immune recognition. *Annual Review of Immunology* 23, 901–944.
- Wermeling, F., Chen, Y., Pikkariainen, T., *et al.*, 2007. Class A scavenger receptors regulate tolerance against apoptotic cells, and autoantibodies against these receptors are predictive of systemic lupus. *Journal of Experimental Medicine* 204, 2259–2265.
- Wermeling, F., Karlsson, M.C., McGaha, T.L., 2009. An anatomical view on macrophages in tolerance. *Autoimmunity Reviews* 9, 49–52.
- Witmer-Pack, M.D., Hughes, D.A., Schuler, G., *et al.*, 1993. Identification of macrophages and dendritic cells in the osteopetrotic (op/op) mouse. *Journal of Cell Science* 104, 1021–1029. Pt 4.
- Wynn, T.A., Chawla, A., Pollard, J.W., 2013. Macrophage biology in development, homeostasis and disease. *Nature* 496, 445–455.
- Yona, S., Kim, K.W., Wolf, Y., *et al.*, 2013. Fate mapping reveals origins and dynamics of monocytes and tissue macrophages under homeostasis. *Immunity* 38, 79–91.
- You, Y., Myers, R.C., Freeberg, L., *et al.*, 2011. Marginal zone B cells regulate antigen capture by marginal zone macrophages. *Journal of Immunology* 186, 2172–2181.
- You, Y., Zhao, H., Wang, Y., Carter, R.H., 2009. Cutting edge: Primary and secondary effects of CD19 deficiency on cells of the marginal zone. *Journal of Immunology* 182, 7343–7347.

CD4 T Cell Memory and Role of TNF Receptor Family

DR Withers and CL Marriott, University of Birmingham, Birmingham, UK

© 2016 Elsevier Inc. All rights reserved.

Overview

CD4 T cells play an array of critical roles within the immune system, aiding many other cell types through cytokine release, supporting the generation of high-affinity antibodies, whilst also directly targeting host cells infected with intracellular bacteria. This wide range of functions is facilitated by the generation of distinct subsets of effector CD4 T cells. Crucially, CD4 T cell responses generate memory cells, long-lasting populations able to better respond to antigen upon re-encounter, thus underpinning vaccination. Key to turning naïve CD4 T cells into effector and memory cells are interactions with antigen presenting cells (APCs) expressing peptide:MHCII complexes alongside a host of other surface receptors and ligands. A combination of signals through the T cell receptor (TCR) and other surface receptors, many of which are members of the tumor necrosis factor receptor super family (TNFRSF), shape this response and are crucial in ensuring optimal effector function and memory cell survival. Here, following a brief overview of CD4 T cell responses, we will discuss the role of different members of the TNFRSF specifically in the generation of CD4 T cell memory.

CD4 T Cell Responses

The initiation of CD4 T cells responses occurs within secondary lymphoid tissue, immune cell depots where naïve CD4 T cells patrol, searching for cognate peptides presented by APCs, typically dendritic cells (DCs; [Bajenoff et al., 2006](#)). Naïve CD4 T cell pools recognizing a specific peptide are very small, typically numbering 20–200 cells in the mouse ([Moon et al., 2007](#)). Therefore, extensive expansion of this population is required to generate a sufficient pool of effector cells that can perform the required functions. However, like an army too expensive to maintain, this effector population is rapidly reduced through apoptosis ([Kaeche et al., 2002](#)). This is accompanied by the emergence of a long-lived memory population with a half-life measuring several weeks ([Pepper et al., 2010](#)). This memory population can then provide protection for extended periods of time, poised to rapidly respond upon antigen reencounter. Thus, successful CD4 T cell responses progress through a series of stages, with rapid proliferation and expansion, then differentiation and survival, processes orchestrated by an array of receptor–ligand interactions; whereas naïve CD8 T cells require only brief stimulation for activation ([Kaeche and Ahmed, 2001](#)). CD4 T cells require more substantial stimulation, with intravital imaging revealing that prolonged contact between naïve CD4 T cells and antigen bearing DCs lasting several hours, was required to drive expansion ([Celli et al., 2007](#)). Activation of naïve CD4 T cells requires signals through the TCR ([Cantrell, 1996](#)) with factors governing the mechanics of the TCR:peptide/MHCII interaction such as TCR affinity key to the resulting signaling to the

T cell ([Corse et al., 2011](#)). It is clear that in addition to TCR:MHCII, further signals, termed costimulation, are required for productive CD4 T cell responses, with signaling through CD28 critical for driving the expansion of naïve CD4 T cells *in vivo* ([Shahinian et al., 1993](#); [Green et al., 1994](#)). Costimulatory signals through CD28 mediate production of IL-2 and other cytokines ([Fraser et al., 1992](#); [Thompson et al., 1989](#)), up-regulation of the IL-2 receptor ([Cerdan et al., 1995](#)), and proliferation ([Gimmi et al., 1991](#)). Due to the focus of this article, this very brief summary of TCR and CD28 signals, does not even scratch the surface of what is currently understood about these processes and their role in CD4 T cell responses.

The formation of memory populations as well as functional effector cells is clearly dependent upon signals received during the initial clonal expansion stage. Titration of the number of TCR transgenic T cells demonstrated that optimal memory cell formation required very low-precursor frequency, mirroring the size of endogenous naïve CD4 T cell pools ([Hataye et al., 2006](#)). Where high numbers of naïve TCR transgenic CD4 T cells were transferred, thus reducing the extent of DC interactions through competition, an effector cell pool with impaired function developed ([Blair and Lefrancois, 2007](#)). Whether memory cells form directly from naïve CD4 T cells, or reflect the persistence of a subset of the effector population is not completely resolved. Several studies have shown that interferon- γ (IFN γ)-producing effector cells differentiated into memory populations as efficiently as IFN γ [−] cells ([Lohning et al., 2008](#); [Harrington et al., 2008](#); [Pepper et al., 2010](#)), demonstrating that effector cells can give rise to memory cells. Amongst CD4 memory cells, distinct populations of T effector memory (Tem) and T central memory (Tcm) have been described with the latter able to recirculate through secondary lymphoid tissue through differential expression of CD62L and chemokine receptors ([Pepper et al., 2011](#); [Sallusto et al., 1999, 2004](#)). Once formed, murine memory CD4 T cells require the cytokines IL-7 and IL-15 for survival ([Seddon et al., 2003](#); [Kondrack et al., 2003](#); [Purton et al., 2007](#)). Interestingly, memory CD4 T cell survival is independent of MHCII, indicating that further signaling through the TCR is not required ([Purton et al., 2007](#)).

Members of the TNFRSF and CD4 T Cells

The TNFRSF contains 19 ligands and 30 receptors ([Croft, 2009](#)), whose interactions underpin a diverse range of roles in the immune system from lymph node development to T cell survival. Against the backdrop of TCR:peptide MHCII interactions and CD28 family costimulation outlined above, several members of the TNFRSF have been shown to contribute to the success of CD4 T cell responses, in general, supporting the proliferation, cytokine production, and survival of activated CD4 T cells, events downstream of naïve CD4 T cell proliferation driven by TCR and CD28 signals ([Watts, 2005](#);

Croft, 2009). The intracellular regions of these TNFRs associate with TNFR-associated factors (TRAFs) which feed into both the canonical and noncanonical nuclear factor- κ B (NF- κ B) signaling pathways (So and Croft, 2013). These receptors also convey signals through the phosphoinositide-3-kinase (PI3K) and Akt pathway, key to T cell activation, to which other costimulatory signals (such as CD28) contribute (So and Croft, 2013). Therefore, signals through TNFRs may contribute to multiple programs within the T cell and these processes remain incompletely understood. Of the TNFRs discussed in this article, all signal through TRAF2, thought to be key in supporting T cell responses (So and Croft, 2013; Watts, 2005). Since many TNFRs signal through the same TRAFs, there may be redundancy in their requirements, however, the defects observed in mice specifically lacking individual TNFRs argue that different members of this family cannot completely substitute for each other, likely due to the involvement of additional signaling components. The context within which antigen is encountered as well as infection type and chronicity will also clearly contribute to TNFR signal requirements, influencing the duration and pattern of expression. Simplistically, some TNFRs such as 41BB (TNFRSF9) and TNFR2 (TNFRSF1B) appear more important for CD8 T responses, rather than those of CD4 T cells (Kim and Teh, 2001, 2004; Dawicki *et al.*, 2004). Cross talk between the activated T cell and DC, through $LT\alpha_1\beta_2$ (rapidly expressed after activation) and $LT\beta$ R (TNFRSF3), ensures full DC maturation and thus optimal support for CD4 T cell priming (Summers-DeLuca *et al.*, 2007). Whilst induced on activated CD4 T cells, many TNFRs are constitutively expressed on regulatory T cells, further complicating our understanding of their roles in CD4 T cell responses (Croft, 2009). Here, we will focus on CD27 (TNFRSF7), Death receptor 3 (DR3, TNFRSF25), Herpesvirus entry mediator (HVEM, TNFRSF14), OX40 (TNFRSF4), and CD30 (TNFRSF8) all of which directly contribute to expanding and sustaining CD4 T cells and the resulting memory population (Figure 1).

In assessing memory CD4 T cell responses, the generation of this population, its subsequent survival and then its

requirements to respond to antigen reencounter are all important phases where different signals appear to be required. As discussed below these stages are, to varying degrees, controlled by signals through these TNFRSF members.

CD27

CD27 (which binds CD70) is expressed on most (including naïve) CD4 T cells, with anti-CD27 antibodies enhancing T cell expansion *in vitro* (Gravestine *et al.*, 1995). CD27 can bind to TRAFs 2 and 5 signaling through the NF- κ B pathway (So and Croft, 2013). CD27-deficient mice show reduced primary responses to influenza virus and impaired formation of memory CD4 T cells (Hendriks *et al.*, 2000), whilst constitutive expression of CD70 drives effector CD4 T cell expansion and overproduction of IFN γ (Arens *et al.*, 2001). Therefore, CD27 signals mediate survival in responding cells governing the size of the effector cell pool and thus development of memory CD4 T cells.

DR3

DR3 is another member induced following T cell activation, that upon binding TL1a signals through the NF- κ B pathway via TRAF2 (Migone *et al.*, 2002; So and Croft, 2013). Whilst DR3^{-/-} mice showed no impairments in CD4 T cell differentiation, effector T cell accumulation and cytokine production was significantly reduced in several experimental disease models consistent with a role sustaining effector CD4 T cells and thus influencing formation of the memory CD4 T cell pool (Meylan *et al.*, 2008).

HVEM

In contrast to many of the other TNFRSF members discussed here, HVEM is expressed by naïve, effector and memory cells, CD4 T cells and can bind to either LIGHT (homologous to lymphotoxins, exhibits inducible expression, and competes

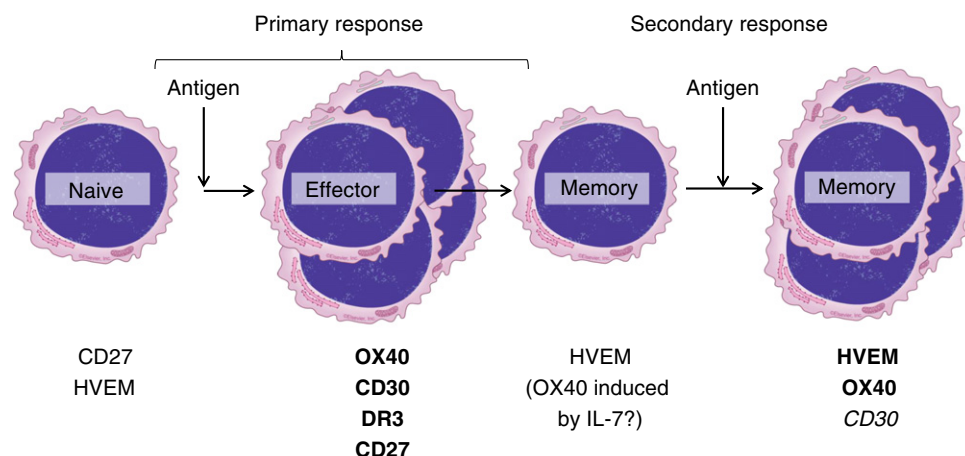


Figure 1 Expression of TNFRSF members at different stages of a CD4 T cell response. Members of the TNFRSF are differentially expressed by CD4 T cells during an immune response. Bold font identifies molecules where there is clear experimental evidence pointing to a role specifically at this stage in the response.

with HSV glycoprotein D for HVEM, a receptor expressed by T lymphocytes) or BTLA (B and T lymphocyte attenuator) (Soroosh *et al.*, 2011). HVEM signals through the NK- κ B pathway, binding TRAFs 1, 2, 3, and 5 (Marsters *et al.*, 1997) impacting on CD4 T cell proliferation and cytokine production (Scheu *et al.*, 2002). Signals through HVEM appear particularly important for memory CD4 T cell survival in recall responses (Soroosh *et al.*, 2011). HVEM^{-/-} TCR transgenic T cells proliferated normally following secondary challenge and initially produced normal amounts of cytokine. However, this expanded memory population was not sustained in the absence of HVEM, which correlated with a decline in levels of phosphorylated Akt in HVEM^{-/-} T cells (Soroosh *et al.*, 2011). These data indicate that memory CD4 T cells may require certain members of the TNFRSF at very specific stages in a response, with HVEM seemingly key to maintaining memory cells after challenge.

OX40 and CD30

Of all the TNFRSF members implicated in CD4 T cell responses, OX40 and CD30 appear to be most important for survival and effector function. OX40 (which binds OX40L) was first recognized for its induced expression on rat (Mallett *et al.*, 1990), then mouse (Calderhead *et al.*, 1993) CD4 T cells, where *in vitro* and *in vivo* expression occurs within approximately 24 h of activation, peaks after a further 24–48 h, and is then rapidly lost (Gramaglia *et al.*, 1998, 2000; Rogers *et al.*, 2001; Marriott *et al.*, 2014). However, this expression pattern is clearly dependent upon the context in which antigen is encountered, with the brief expression described typical of acute infections (Marriott *et al.*, 2014), whilst in chronic infections expression on antigen-specific CD4 T cells can be maintained for many days (Boettler *et al.*, 2012). Expression of OX40 is downstream of CD28 signals (Walker *et al.*, 1999) and not required for initial naïve cell proliferation (Rogers *et al.*, 2001). Rather, OX40 signals via canonical NF- κ B and the PI3K/protein kinase B pathways to maintain clonal expansion, through sustained proliferation (mediated by molecules such as survivin and aurora B) and survival through promoting expression of the anti-apoptotic proteins Bcl-XL and Bcl-2 (Song *et al.*, 2008, 2005; Rogers *et al.*, 2001). Mice deficient in OX40 or its ligand show impaired CD4 T cell expansion and function but normal B cell responses (Chen *et al.*, 1999; Pippig *et al.*, 1999; Kopf *et al.*, 1999; Rogers *et al.*, 2001). OX40 signals are also important in the context of autoimmune responses, for example, blockade of OX40L reduced the severity of collagen-induced arthritis (Yoshioka *et al.*, 2000), whilst allergic inflammation in a murine asthma model was dependent upon OX40 (Jember *et al.*, 2001).

Like OX40, CD30 is induced upon activated CD4 T cells by CD28 signals or IL-4 and whilst less studied than OX40, also enhanced CD4 T cell responses (Toennies *et al.*, 2004; Gilfillan *et al.*, 1998). Both OX40 and CD30 bind TRAFs 1, 2, 3, and 5, although it is less clear whether CD30 signals through the PI3K-Akt pathway (So and Croft, 2013). Consistent with redundancy in the requirements of CD4 T cells for OX40 and CD30 signals, mice deficient in both these receptors (CD30^{-/-}OX40^{-/-}) were significantly more impaired in their CD4 T cell responses than either OX40^{-/-} or CD30^{-/-} mice

(Gaspal *et al.*, 2008, 2005). As described for OX40^{-/-} mice this was not due to impaired proliferation early in the response, rather due to very poor survival *in vivo* (Gaspal *et al.*, 2005). Furthermore, whilst mild defects in the germinal centers of CD30^{-/-} mice were observed, these structures were not maintained in CD30^{-/-}OX40^{-/-} mice (Gaspal *et al.*, 2005). As mentioned above, the formation of a robust memory CD4 T cell population is dependent upon the success of the clonal expansion stage. Memory CD4 T cells residing in the small intestine lamina propria were absent in CD30^{-/-}OX40^{-/-} mice, but only partially reduced in CD30^{-/-} or OX40^{-/-} mice, indicating both CD30 and OX40 signals contribute to the generation of CD4 T cell memory (Withers *et al.*, 2009). Perhaps most strikingly, CD30^{-/-}OX40^{-/-} mice crossed onto the FoxP3^{-/-} background were protected from the lethal autoimmunity that results from the absence of Tregs (Gaspal *et al.*, 2011). It is worth noting that scurfy mice (loss of function mutation in FoxP3) crossed with CD28^{-/-} mice were only partially protected and succumb to autoimmunity after several months (Singh *et al.*, 2007), consistent with CD4 T cell priming in the absence of CD28 under conditions of chronic stimulation (Kawai *et al.*, 1996). The lack of autoimmunity in FoxP3^{-/-}CD30^{-/-}OX40^{-/-} mice really demonstrates the absolute requirement for CD30 and OX40 signals in sustaining a functional CD4 effector pool. Mice deficient in FoxP3 and either OX40 or CD30 survived longer than FoxP3^{-/-} mice, yet still developed autoimmunity. Combined these data demonstrate that signals through both CD30 and OX40 are critical for the generation of effector CD4 T cells and the subsequent formation of a memory CD4 population.

Given the requirements for OX40 reported in many mouse models of infection, autoimmunity, and inflammation, it appears that there is no simple dichotomy in Th1 or Th2 responses and their requirement for signals through OX40 and CD30 (Watts, 2005). Early studies using TCR transgenic T cells reported impaired survival without further dissection of effector types (Kim *et al.*, 2003; Rogers *et al.*, 2001). More recently, reduced numbers of Tem, but not Tcm, have been reported in OX40^{-/-} and CD30^{-/-}OX40^{-/-} mice, assessing both total CD4 and antigen-specific populations (Soroosh *et al.*, 2007; Marriott *et al.*, 2014). *In vivo* administration of agonistic anti-OX40 antibodies drive substantial expansion of Tem (Boettler *et al.*, 2013; Marriott *et al.*, 2014), but not other subsets, likely through sustaining Blimp-1 expression (Boettler *et al.*, 2013). Given that expression of CD30 and OX40 is rapidly lost on effector CD4 T cells, such requirements would appear to reflect interactions with APCs early in the response. Therefore, can all the defects seen in the absence of CD30 and OX40 be attributed to a lack of signals during clonal expansion? It is not clear that once formed, memory cells require OX40 signals for survival as well as cytokines such as IL-7 (Kondrack *et al.*, 2003; Seddon *et al.*, 2003). Interestingly, *in vitro* culture with IL-7 induces OX40 expression on memory CD4 T cells suggesting that induction of OX40 expression on memory CD4 T cells may be antigen independent (Gaspal *et al.*, 2005). This would provide a further mechanism whereby recirculating memory cells might be maintained by interactions with OX40L⁺ cells within secondary lymphoid tissue (Withers *et al.*, 2011). However, arguing against this notion, *ex vivo* antigen-specific memory CD4 T cells generated

following acute infection with an attenuated *Listeria monocytogenes* strain, lacked OX40 expression (Marriott *et al.*, 2014).

Whilst the data described here refers exclusively to murine CD4 T cells, a human patient deficient in OX40 was recently described as having impaired Tem populations with poor IFN γ responses to recall antigens (such as Bacillus Calmette–Guérin (BCG), tetanus toxoid, Epstein–Barr virus (EBV), cytomegalovirus (CMV)) despite earlier exposure (Byun *et al.*, 2013). Therefore, this first human patient provides strong evidence that the OX40 requirements of murine CD4 T cells may be true of human CD4 T cells. That said, much more work on CD4 T cell expression and dependency on members of the TNFRSF is required to facilitate the clinical targeting of these molecules.

The rapid expansion of memory T cells upon antigenic challenge is one of their defining characteristics. Traditionally it has been thought that costimulatory molecules are dispensable for memory cell activation, however, some studies indicate that OX40 signals are required for optimal recall responses. Expression of OX40 is very rapidly upregulated on memory CD4 T cells upon challenge (Marriott *et al.*, 2014). Accumulation of OX40⁺ memory cells was found to be impaired in response to airway challenge in sensitized animals when OX40–OX40L signals were blocked (Salek-Ardakani *et al.*, 2003). OX40 interactions were also found to be required for optimal secondary responses to influenza and using TCR transgenic OTII cells an even greater effect was found with a complete lack of recall response in OX40L^{−/−} mice, despite no significant effect on priming (Dawicki *et al.*, 2004). It is unknown whether CD30 also contributes to memory cell expansion or survival upon challenge, although secondary antibody responses were abrogated in CD30^{−/−} OX40^{−/−} mice (Gaspal *et al.*, 2005). Therefore, OX40 and likely CD30 are key molecules in generation functional effector CD4 T cells, through effects on both expansion and survival during clonal expansion. These signals are critical in establishing a functional effector CD4 T cell pool during a primary response and the subsequent memory cell population. Upon challenge, memory CD4 T cells still require OX40 signals for optimal expansion.

Who Provides the Ligands?

Whilst the consequences of binding the different TNFRSF members is becoming clear, it is still uncertain which cells are the key source for many of the TNFSF ligands. The ligands CD30L, OX40L, CD70, LIGHT, and TL1a each have a diverse pattern of expression, inducible on many cell types (Croft, 2009). Given that their receptors are often temporally expressed on activated CD4 T cells, professional APCs are likely to be the key source of these signals, sustaining CD4 T cells during clonal expansion. Activated CD4 and CD8 T cells can also be induced to express these ligands suggesting T cell–T cell interactions are also important (Croft, 2009). For example, the impaired memory cell survival observed in HVEM^{−/−} T cells, was recapitulated when T cells were deficient in LIGHT, indicating that T cell–T cell interactions supported this process (Soroosh *et al.*, 2011). It seems likely that the key cellular source of these ligands will be dependent upon the nature of the infection or condition, with integration of other stimuli orchestrating ligand expression. Under conditions of

inflammation, non-hematopoietic cells such as endothelium and airway smooth muscle can express ligands such as TL1a and OX40L (Imura *et al.*, 1996; Burgess *et al.*, 2004; Croft, 2009) which may contribute to sustaining activated CD4 T cells at these sites. We have been particularly interested in lymphoid tissue inducer (LTi)-like cells, a ROR γ t-dependent cell that has seen revived attention due to the emergence of the innate lymphoid cell family (Spits *et al.*, 2013). LTi-like cells constitutively express OX40L and CD30L and support memory CD4 T cell survival *in vitro* and *in vivo* (Kim *et al.*, 2003, 2006; Withers *et al.*, 2012). In the embryo, LTi cells function to orchestrate lymph node development through interaction with stromal cells and at this developmental stage lack OX40L and CD30L expression. After birth, these cells acquire further functions including antigen presentation and processing (Hepworth *et al.*, 2013) as well as OX40L and CD30L expression (Kim *et al.*, 2006). Whilst a rare population, these cells are located clustered in lymph nodes and other secondary lymphoid tissue at sites of memory cell recirculation suggesting that these cells could interact and sustain memory CD4 T cells via OX40L (Withers *et al.*, 2011). However, direct evidence that *in vivo* LTi-like expression of OX40L and CD30L mediates memory cell survival is currently lacking. Furthermore, these cells express LIGHT (Kim *et al.*, 2006), which may also contribute to memory cell survival via HVEM. With the current widespread availability and use of conditional knockout mice, the precise details of who interacts with whom and when can be carefully elucidated, facilitating the clinical targeting of these pathways.

Concluding Remarks

Several TNFRSF members contribute to optimal CD4 T cell responses through expanding and sustaining effector CD4 T cells and the resulting memory pool. Whilst many of these molecules contribute to the same signaling pathways, it is becoming apparent that certain molecules are key at different stages of the CD4 T cell response. There are now clear data that combined signals through OX40 and CD30 are required for a functional effector cell pool, particularly in an autoimmune setting, suggesting these are excellent therapeutic targets for controlling CD4 T cell responses.

See also: Protein Synthesis/Degradation: Protein Degradation – Protease Classes: ADAMs Regulate Cell–Cell Interactions by Controlling the Function of the EGF-Receptor, TNF α and Notch

References

- Arens, R., Tesselaar, K., Baars, P.A., *et al.*, 2001. Constitutive CD27/CD70 interaction induces expansion of effector-type T cells and results in IFN γ -mediated B cell depletion. *Immunity* 15, 801–812.
- Bajenoff, M., Egen, J.G., Koo, L.Y., *et al.*, 2006. Stromal cell networks regulate lymphocyte entry, migration, and territoriality in lymph nodes. *Immunity* 25, 989–1001.
- Blair, D.A., Lefrançois, L., 2007. Increased competition for antigen during priming negatively impacts the generation of memory CD4 T cells. *Proceedings of the*

- National Academy of Sciences of the United States of America 104, 15045–15050.
- Boettler, T., Choi, Y.S., Salek-Ardakani, S., *et al.*, 2013. Exogenous OX40 stimulation during lymphocytic choriomeningitis virus infection impairs follicular Th cell differentiation and diverts CD4 T cells into the effector lineage by upregulating Blimp-1. *Journal of Immunology* 191 (10), 5026–5035.
- Boettler, T., Moeckel, F., Cheng, Y., *et al.*, 2012. OX40 facilitates control of a persistent virus infection. *PLOS Pathogens* 8, e1002913.
- Burgess, J.K., Carlin, S., Pack, R.A., *et al.*, 2004. Detection and characterization of OX40 ligand expression in human airway smooth muscle cells: A possible role in asthma? *Journal of Allergy and Clinical Immunology* 113, 683–689.
- Byun, M., Ma, C.S., Akcay, A., *et al.*, 2013. Inherited human OX40 deficiency underlying classic Kaposi sarcoma of childhood. *Journal of Experimental Medicine* 210, 1743–1759.
- Calderhead, D.M., Buhlmann, J.E., Vandeneertwegh, A.J.M., *et al.*, 1993. Cloning of mouse OX40 – A T-cell activation marker that may mediate t-b-cell interactions. *Journal of Immunology* 151, 5261–5271.
- Cantrell, D., 1996. T cell antigen receptor signal transduction pathways. *Annual Review of Immunology* 14, 259–274.
- Celli, S., Lemaître, F., Bouso, P., 2007. Real-time manipulation of T cell-dendritic cell interactions *in vivo* reveals the importance of prolonged contacts for CD4 + T cell activation. *Immunity* 27, 625–634.
- Cerdan, C., Martin, Y., Courcoul, M., *et al.*, 1995. CD28 costimulation up-regulates long-term IL-2R beta expression in human T cells through combined transcriptional and post-transcriptional regulation. *Journal of Immunology* 154, 1007–1013.
- Chen, A.I., McAdam, A.J., Buhlmann, J.E., *et al.*, 1999. OX40-ligand has a critical costimulatory role in dendritic cell: T cell interactions. *Immunity* 11, 689–698.
- Corse, E., Gottschalk, R.A., Allison, J.P., 2011. Strength of TCR-peptide/MHC interactions and *in vivo* T cell responses. *Journal of Immunology* 186, 5039–5045.
- Croft, M., 2009. The role of TNF superfamily members in T-cell function and diseases. *Nature Reviews Immunology* 9, 271–285.
- Dawicki, W., Bertram, E.M., Sharpe, A.H., Watts, T.H., 2004. 4-1BB and OX40 act independently to facilitate robust CD8 and CD4 recall responses. *Journal of Immunology* 173, 5944–5951.
- Fraser, J.D., Newton, M.E., Weiss, A., 1992. CD28 and T cell antigen receptor signal transduction coordinately regulate interleukin 2 gene expression in response to superantigen stimulation. *Journal of Experimental Medicine* 175, 1131–1134.
- Gaspal, F., Bekiaris, V., Kim, M.Y., *et al.*, 2008. Critical synergy of CD30 and OX40 signals in CD4 T cell homeostasis and Th1 immunity to Salmonella. *Journal of Immunology* 180, 2824–2829.
- Gaspal, F., Withers, D., Saini, M., *et al.*, 2011. Abrogation of CD30 and OX40 signals prevents autoimmune disease in FoxP3-deficient mice. *Journal of Experimental Medicine* 208, 1579–1584.
- Gaspal, F.M., Kim, M.Y., McConnell, F.M., *et al.*, 2005. Mice deficient in OX40 and CD30 signals lack memory antibody responses because of deficient CD4 T cell memory. *Journal of Immunology* 174, 3891–3896.
- Gilfillan, M.C., Noel, P.J., Podack, E.R., Reiner, S.L., Thompson, C.B., 1998. Expression of the costimulatory receptor CD30 is regulated by both CD28 and cytokines. *Journal of Immunology* 160, 2180–2187.
- Gimmi, C.D., Freeman, G.J., Gribben, J.G., *et al.*, 1991. B-cell surface antigen B7 provides a costimulatory signal that induces T cells to proliferate and secrete interleukin 2. *Proceedings of the National Academy of Sciences of the United States of America* 88, 6575–6579.
- Gramaglia, I., Jember, A., Pippig, S.D., *et al.*, 2000. The OX40 costimulatory receptor determines the development of CD4 memory by regulating primary clonal expansion. *Journal of Immunology* 165, 3043–3050.
- Gramaglia, I., Weinberg, A.D., Lemon, M., Croft, M., 1998. Ox-40 ligand: A potent costimulatory molecule for sustaining primary CD4 T cell responses. *Journal of Immunology* 161, 6510–6517.
- Gravestine, L.A., Nieland, J.D., Kruisbeek, A.M., Borst, J., 1995. Novel mAbs reveal potent co-stimulatory activity of murine CD27. *International Immunology* 7, 551–557.
- Green, J.M., Noel, P.J., Sperling, A.I., *et al.*, 1994. Absence of B7-dependent responses in CD28-deficient mice. *Immunity* 1, 501–508.
- Harrington, L.E., Janowski, K.M., Oliver, J.R., Zajac, A.J., Weaver, C.T., 2008. Memory CD4 T cells emerge from effector T-cell progenitors. *Nature* 452, 356–360.
- Hataye, J., Moon, J.J., Khoruts, A., Reilly, C., Jenkins, M.K., 2006. Naive and memory CD4 + T cell survival controlled by clonal abundance. *Science* 312, 114–116.
- Hendriks, J., Gravestine, L.A., Tesselaar, K., *et al.*, 2000. CD27 is required for generation and long-term maintenance of T cell immunity. *Nature Immunology* 1, 433–440.
- Hepworth, M.R., Monticelli, L.A., Fung, T.C., *et al.*, 2013. Innate lymphoid cells regulate CD4 + T-cell responses to intestinal commensal bacteria. *Nature* 498, 113–117.
- Imura, A., Hori, T., Imada, K., *et al.*, 1996. The human OX40/gp34 system directly mediates adhesion of activated T cells to vascular endothelial cells. *Journal of Experimental Medicine* 183, 2185–2191.
- Jember, A.G., Zuberi, R., Liu, F.T., Croft, M., 2001. Development of allergic inflammation in a murine model of asthma is dependent on the costimulatory receptor OX40. *Journal of Experimental Medicine* 193, 387–392.
- Kaech, S.M., Ahmed, R., 2001. Memory CD8 + T cell differentiation: Initial antigen encounter triggers a developmental program in naive cells. *Nature Immunology* 2, 415–422.
- Kaech, S.M., Wherry, E.J., Ahmed, R., 2002. Effector and memory T-cell differentiation: Implications for vaccine development. *Nature Reviews Immunology* 2, 251–262.
- Kawai, K., Shahinian, A., Mak, T.W., Ohashi, P.S., 1996. Skin allograft rejection in CD28-deficient mice. *Transplantation* 61, 352–355.
- Kim, E.Y., Teh, H.S., 2001. TNF type 2 receptor (p75) lowers the threshold of T cell activation. *Journal of Immunology* 167, 6812–6820.
- Kim, E.Y., Teh, H.S., 2004. Critical role of TNF receptor type-2 (p75) as a costimulator for IL-2 induction and T cell survival: A functional link to CD28. *Journal of Immunology* 173, 4500–4509.
- Kim, M.Y., Gaspal, F.M., Wiggett, H.E., *et al.*, 2003. CD4(+)CD3(–) accessory cells costimulate primed CD4 T cells through OX40 and CD30 at sites where T cells collaborate with B cells. *Immunity* 18, 643–654.
- Kim, M.Y., Toellner, K.M., White, A., *et al.*, 2006. Neonatal and adult CD4 + CD3– cells share similar gene expression profile, and neonatal cells up-regulate OX40 ligand in response to TL1A (TNFSF15). *Journal of Immunology* 177, 3074–3081.
- Kondrack, R.M., Harbertson, J., Tan, J.T., *et al.*, 2003. Interleukin 7 regulates the survival and generation of memory CD4 cells. *Journal of Experimental Medicine* 198, 1797–1806.
- Kopf, M., Ruedl, C., Schmitz, N., *et al.*, 1999. OX40-deficient mice are defective in Th cell proliferation but are competent in generating B cell and CTL responses after virus infection. *Immunity* 11, 699–708.
- Lohning, M., Hegazy, A.N., Pinschewer, D.D., *et al.*, 2008. Long-lived virus-reactive memory T cells generated from purified cytokine-secreting T helper type 1 and type 2 effectors. *Journal of Experimental Medicine* 205, 53–61.
- Mallett, S., Fossum, S., Barclay, A.N., 1990. Characterization of the MRC OX40 antigen of activated CD4 positive T lymphocytes – A molecule related to nerve growth factor receptor. *EMBO Journal* 9, 1063–1068.
- Marriott, C.L., Mackley, E.C., Ferreira, C., *et al.*, 2014. OX40 controls effector CD4 T-cell expansion, not follicular T helper cell generation in acute *Listeria* infection. *European Journal of Immunology* 44 (8), 2437–2447.
- Marsters, S.A., Ayres, T.M., Skubatch, M., *et al.*, 1997. Herpesvirus entry mediator, a member of the tumor necrosis factor receptor (TNFR) family, interacts with members of the TNFR-associated factor family and activates the transcription factors NF-kappaB and AP-1. *Journal of Biological Chemistry* 272, 14029–14032.
- Meylan, F., Davidson, T.S., Kahle, E., *et al.*, 2008. The TNF-family receptor DR3 is essential for diverse T cell-mediated inflammatory diseases. *Immunity* 29, 79–89.
- Migone, T.S., Zhang, J., Luo, X., *et al.*, 2002. TL1A is a TNF-like ligand for DR3 and TR6/DcR3 and functions as a T cell costimulator. *Immunity* 16, 479–492.
- Moon, J.J., Chu, H.H., Pepper, M., *et al.*, 2007. Naive CD4(+) T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. *Immunity* 27, 203–213.
- Pepper, M., Linehan, J.L., Pagan, A.J., *et al.*, 2010. Different routes of bacterial infection induce long-lived TH1 memory cells and short-lived TH17 cells. *Nature Immunology* 11, 83–89.
- Pepper, M., Pagan, A.J., Igyarto, B.Z., Taylor, J.J., Jenkins, M.K., 2011. Opposing signals from the Bcl6 transcription factor and the interleukin-2 receptor generate T helper 1 central and effector memory cells. *Immunity* 35, 583–595.
- Pippig, S.D., Pena-Rossi, C., Long, J., *et al.*, 1999. Robust B cell immunity but impaired T cell proliferation in the absence of CD134 (OX40). *Journal of Immunology* 163, 6520–6529.
- Purton, J.F., Tan, J.T., Rubinstein, M.P., *et al.*, 2007. Antiviral CD4 + memory T cells are IL-15 dependent. *Journal of Experimental Medicine* 204, 951–961.
- Rogers, P.R., Song, J., Gramaglia, I., Killeen, N., Croft, M., 2001. OX40 promotes Bcl-xL and Bcl-2 expression and is essential for long-term survival of CD4 T cells. *Immunity* 15, 445–455.
- Salek-Ardakani, S., Song, J., Halteman, B.S., *et al.*, 2003. OX40 (CD134) controls memory T helper 2 cells that drive lung inflammation. *Journal of Experimental Medicine* 198, 315–324.
- Sallusto, F., Geginat, J., Lanzavecchia, A., 2004. Central memory and effector memory T cell subsets: Function, generation, and maintenance. *Annual Review of Immunology* 22, 745–763.

- Sallusto, F., Lenig, D., Forster, R., Lipp, M., Lanzavecchia, A., 1999. Two subsets of memory T lymphocytes with distinct homing potentials and effector functions (in process citation). *Nature* 401, 708–712.
- Scheu, S., Alferink, J., Potzel, T., *et al.*, 2002. Targeted disruption of LIGHT causes defects in costimulatory T cell activation and reveals cooperation with lymphotoxin beta in mesenteric lymph node genesis. *Journal of Experimental Medicine* 195, 1613–1624.
- Seddon, B., Tomlinson, P., Zamoyska, R., 2003. Interleukin 7 and T cell receptor signals regulate homeostasis of CD4 memory cells. *Nature Immunology* 4, 680–686.
- Shahinian, A., Pfeffer, K., Lee, K.P., *et al.*, 1993. Differential T cell costimulatory requirements in CD28-deficient mice. *Science* 261, 609–612.
- Singh, N., Chandler, P.R., Seki, Y., *et al.*, 2007. Role of CD28 in fatal autoimmune disorder in scurfy mice. *Blood* 110, 1199–1206.
- So, T., Croft, M., 2013. Regulation of PI-3-kinase and Akt signaling in T lymphocytes and other cells by TNFR family molecules. *Frontiers in Immunology* 4, 139.
- Song, J., So, T., Cheng, M., Tang, X., Croft, M., 2005. Sustained survivin expression from OX40 costimulatory signals drives T cell clonal expansion. *Immunity* 22, 621–631.
- Song, J., So, T., Croft, M., 2008. Activation of NF-kappaB1 by OX40 contributes to antigen-driven T cell expansion and survival. *Journal of Immunology* 180, 7240–7248.
- Soroosh, P., Doherty, T.A., So, T., *et al.*, 2011. Herpesvirus entry mediator (TNFRSF14) regulates the persistence of T helper memory cell populations. *Journal of Experimental Medicine* 208, 797–809.
- Soroosh, P., Ine, S., Sugamura, K., Ishii, N., 2007. Differential requirements for OX40 signals on generation of effector and central memory CD4+ T cells. *Journal of Immunology* 179, 5014–5023.
- Spits, H., Artis, D., Colonna, M., *et al.*, 2013. Innate lymphoid cells — A proposal for uniform nomenclature. *Nature Reviews Immunology* 13, 145–149.
- Summers-DeLuca, L.E., McCarthy, D.D., Cosovic, B., *et al.*, 2007. Expression of lymphotoxin-alpha-beta on antigen-specific T cells is required for DC function. *Journal of Experimental Medicine* 204, 1071–1081.
- Thompson, C.B., Lindsten, T., Ledbetter, J.A., *et al.*, 1989. CD28 activation pathway regulates the production of multiple T-cell-derived lymphokines/cytokines. *Proceedings of the National Academy of Sciences of the United States of America* 86, 1333–1337.
- Toennies, H.M., Green, J.M., Arch, R.H., 2004. Expression of CD30 and OX40 on T lymphocyte subsets is controlled by distinct regulatory mechanisms. *Journal of Leukocyte Biology* 75, 350–357.
- Walker, L.S., Gulbranson-Judge, A., Flynn, S., *et al.*, 1999. Compromised OX40 function in CD28-deficient mice is linked with failure to develop CXC chemokine receptor 5-positive CD4 cells and germinal centers. *Journal of Experimental Medicine* 190, 1115–1122.
- Watts, T.H., 2005. TNF/TNFR family members in costimulation of T cell responses. *Annual Review of Immunology* 23, 23–68.
- Withers, D.R., Gaspal, F.M., Bekiaris, V., *et al.*, 2011. OX40 and CD30 signals in CD4(+) T-cell effector and memory function: A distinct role for lymphoid tissue inducer cells in maintaining CD4(+) T-cell memory but not effector function. *Immunological Reviews* 244, 134–148.
- Withers, D.R., Gaspal, F.M., Mackley, E.C., *et al.*, 2012. Cutting edge: Lymphoid tissue inducer cells maintain memory CD4 T cells within secondary lymphoid tissue. *Journal of Immunology* 189, 2094–2098.
- Withers, D.R., Jaensson, E., Gaspal, F., *et al.*, 2009. The survival of memory CD4+ T cells within the gut lamina propria requires OX40 and CD30 signals. *Journal of Immunology* 183, 5079–5084.
- Yoshioka, T., Nakajima, A., Akiba, H., *et al.*, 2000. Contribution of OX40/OX40 ligand interaction to the pathogenesis of rheumatoid arthritis. *European Journal of Immunology* 30, 2815–2823.

STEM CELL BIOLOGY: STRUCTURE AND FUNCTION

Contents

Stem Cells and Aging

Epigenetic Regulation of Stem Cells

The Adult Stem Cell Niche: Multiple Cellular Players in Tissue Homeostasis and Regeneration

Cancer Stem Cells

Metabolic and Energetic Regulation of Stem Cells

Pluripotent Cells: New Tools for Disease Research and Therapies

Stem Cells and Aging

E Perdiguerro and L García-Prat, Pompeu Fabra University (UPF), Barcelona, Spain

P Sousa-Victor, Pompeu Fabra University (UPF), Barcelona, Spain and Buck Institute for Research on Aging, Novato, CA, USA

P Muñoz-Cánoves, Pompeu Fabra University (UPF), Barcelona, Spain and Institutació Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

© 2016 Elsevier Inc. All rights reserved.

Introduction

According to current statistics and future projections of the National Institute on Aging (NIA, 2011), people aged 65 and older represent around 8% of the global population and will raise to around 16% by 2050. At present, this demographic shift toward an aged population is more evident in developed, industrialized societies, but is becoming an increasingly important phenomenon in developing societies as better health care, nutrition, and technological innovations become widely available.

Aging is an extremely complicated process, affecting all organs and tissues at different rates in different individuals. However, a closer examination of many of the major age-associated diseases has revealed that the process of aging is essentially the loss of tissue homeostasis, or a decline in the ability of stem cells to maintain or regenerate tissues over time due to a loss of stem cell function (Lopez-Otin *et al.*, 2013). In this article, we review some of the general aspects of stem cell biology associated with the aging and loss of function of most stem cell types; we then discuss these changes in relation to specific types of stem cells such as muscle stem cells (satellite cells), hematopoietic stem cells (HSCs), and mesenchymal stem cells (MSCs). In addition, we discuss potential ways by which aging may be countered by novel pharmacological and therapeutic interventions and the potential impact of this greater understanding of stem cell biology on the emerging field of stem cell therapies.

Determinants of Stem Cell Aging

Adult or somatic stem cells (SCs) represent a small population of long-lived progenitor cells that reside in an undifferentiated

state within most fully differentiated adult tissues. SCs are fundamental for maintaining tissue homeostasis and for mediating growth and regeneration after damage or disease. The major mechanisms by which SCs are regulated are similar between different SC types and are often altered or defective during aging. Therefore, understanding these mechanisms is essential to determine how aging affects SC behavior and to provide points of control by which SC behavior may be modified in genetic animal models or by pharmacological intervention.

Adult SCs are maintained in a specialized microenvironment called the 'niche,' which has often been proved difficult to study in detail, but whose particular features are gradually being unraveled. One of the principal roles of the niche is to provide a local environment rich in growth factors, trophic factors, and cytokines, which act in concert to sustain and support specific SC properties such as a quiescence, metabolic status, or the ability to activate and respond quickly to signals of tissue damage and repair (Jasper and Kennedy, 2012; Jones and Rando, 2011). Regulation of the niche, the quiescent state, and the transition from quiescence to activation are fundamental processes that may change during disease and aging.

Once activated, SCs proliferate and then differentiate as they contribute to the replacement or maintenance of the damaged tissue or organ. There is a great deal of heterogeneity in the proliferative behavior and potential of different types of adult SCs, reflecting the different regenerative requirements of each tissue. Skeletal muscle is generally a stable post-mitotic tissue, and satellite cells proliferate only when required, such as after injury, whereas the SCs of the hematopoietic system or the intestine, both high-turnover tissues, contain SCs with a high proliferative capacity to ensure homeostasis. The rate of

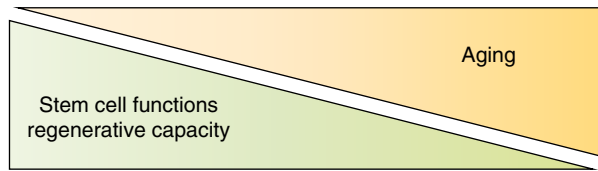


Figure 1 Stem cell regenerative potential declines with aging.

SC proliferation needs to be tightly regulated to ensure that adequate numbers of progeny are produced; too many cells could disrupt tissue architecture, whereas too few could fail to restore homeostasis (Liu and Rando, 2011; Li and Clevers, 2010). Importantly, stem cell proliferation is often impaired during aging, with examples of loss of proliferative potential and other examples of over-proliferative behavior.

In terms of differentiation, adult SCs have a defined potency depending on the tissue, being restricted to the generation of particular lineages to ensure ordered programs of tissue/organ repair. However, these programs often become aberrant with age, leading to defective regeneration. Alterations include both skewed lineages (as observed in HSCs) and adoption of alternative lineages (as observed in satellite cells) (see below).

A fundamental aspect of SC behavior is also the ability to self-renew and thus to replenish the SC population. This process is essential for ensuring preservation of the regenerative capacity of the tissue over time; loss of self-renewal will lead to SC depletion and therefore failure of homeostasis and repair (Simons and Clevers, 2011). Aging-associated defects in self-renewal have been reported in the majority of tissues, including the brain, the hematopoietic system, and the skeletal muscle.

Thus, the effects of aging on SC function can manifest at any of these levels of control and function, from dysregulation of the niche, to activation, proliferation, differentiation, and self-renewal (Figure 1). In the following sections we discuss the specific cell-intrinsic and -extrinsic alterations that may cause these shifts in function, and we focus on some particular cell models in which they have been studied.

Genetic Alterations

DNA damage and repair

Maintaining the integrity of DNA throughout the lifetime of an organism is fundamental to survival and the avoidance of disease, particularly cancer. Many mechanisms exist for the protection and proofreading of DNA in all species during cell division, but also during less active cellular states where metabolic by-products such as reactive oxygen species (ROS), exogenous chemical agents, and other environmental factors are still potentially dangerous. Indeed, endogenous DNA damage has been shown to accumulate with age in HSCs, with functional consequences (Rossi *et al.*, 2007). Specific mouse genetic models have been instrumental in demonstrating the DNA protection and repair mechanisms deregulated in aging HSCs, such as homologous recombination mediated by the FANCD1/BRCA2 gene, which is mutated in models of Fanconi anemia (Navarro *et al.*, 2006). In this sense, progeroid or

premature aging syndromes such as Werner syndrome or Bloom syndrome, among others, have principally been linked to mutations in genes involved in DNA repair (Burke and Stewart, 2014).

One mechanism that SCs use to minimize the potentially negative consequences of cell division and/or exposure to ROS is to adopt a nondividing and dormant state called quiescence. The exact cell cycle state that quiescent cells adopt is still controversial (see below), but it is accepted that, to repair DNA breaks, quiescent cells utilize the error-prone nonhomologous end-joining (NHEJ) pathway (Mohrin *et al.*, 2010). NHEJ function in the absence of cell division is based on a repair mechanism that does not require homologous strand guidance, but instead ligates the ends of double strand breaks together, utilizing very short homologous DNA sequences called microhomologies that are produced during the break. This is, however, an error-prone mechanism that may lead to insertional mutations and ultimately genomic instability (Mohrin *et al.*, 2010). Importantly, reduction of NHEJ pathway activity has been reported during aging (Vyjayanti and Rao, 2006), while genetic alterations in key components of the NHEJ pathway have been linked with aging phenotypes (Rossi *et al.*, 2007).

Telomere shortening

Telomeres are specialized and repetitive ribonucleotide protein structures located at the ends of chromosomes; they protect the chromosomal DNA from degradation and unwanted fusion with other chromosomes. It is widely accepted that, in somatic cells, each round of cell division reduces telomere length such that telomeres function as molecular clocks and therefore have an essential link to aging. Conversely, some cell types, particularly SCs and cancer cells, retain or restore mechanisms to preserve telomere length and thus obtain extended longevity or even immortality. However, age-associated telomere shortening in several adult SC compartments has been reported and can influence function (Flores *et al.*, 2008). As each round of replication reduces telomere length, a telomere will eventually reach a critical threshold, at which point a state of replicative senescence is induced, the so-called Hayflick limit (Hayflick and Moorhead, 1961). Cellular senescence is characterized by G1 cell cycle arrest, which is mediated by induction of p16^{INK4a}, retinoblastoma (RB), and p53 proteins and a state of persistent DNA damage often associated with apoptosis (D'Adda Di Fagagna *et al.*, 2003). Clearly, an SC that has achieved a state of cellular senescence is unable to repopulate or repair a damaged tissue, and this inverse relationship between telomere length and SC function has been demonstrated in adult tissues from several animal models of telomerase deficiency and mice with short telomeres even in the presence of telomerase (Flores and Blasco, 2010).

Epigenetic Alterations

Age-dependent alterations in the epigenome, such as alternating patterns of DNA methylation and posttranslational modifications of histones, can also affect chromatin structure and therefore the access of transcription factors and accessory

proteins to genes, which consequently alter tissue homeostasis and regeneration.

Histone acetyltransferases (HATs) and deacetylases (HDACs) have been extensively studied in relation to aging (Berger, 2007). Sirt2 HDAC has been shown to enhance the replicative lifespan of several invertebrate models (Longo and Kennedy, 2006). The unique requirement of Sirt2 and its orthologs for NAD and its varying effects on longevity in different nutritional states creates a firm link between metabolism and aging. Regarding histone methyltransferases, the lifespan of *Caenorhabditis elegans* has been shown to be modified by histone methylation, including trimethylation of histone H3 at lysine 4 (H3K4me3), a mark associated with active chromatin, which is mediated by the trithorax (TrxG) complex of proteins. In particular, knockdown of several TrxG members was shown to extend *C. elegans*' lifespan (Greer *et al.*, 2010). The polycomb repressive complex 2 (PRC2) complex of proteins trimethylates histone H3 on lysine 27 (H3K27me3), a mark which transcriptionally silences chromatin. Global epigenetic profiling of young muscle SCs showed that in quiescence their chromatin is maintained in a transcriptionally permissive state, with a large number of genes marked with H3K4me3 at the transcription start sites (TSS), and a few marked with H3K27me3, which would allow quiescent satellite cells to be primed for activation and differentiation (Liu *et al.*, 2013). During aging, the epigenetic scenario changes, with a global increase of H3K27me3 across the genome, coinciding with an increase in dsDNA breaks and a reduction in histone biosynthesis (Liu *et al.*, 2013).

Genome-wide studies have also suggested that there is a global age-associated loss of DNA methylation in human, rat, and mouse tissues (Issa, 2014). In particular, hypermethylation of the promoters at specific loci, as well as a progressive loss of 5-methylcytosine content, particularly within DNA repeated sequences, was shown to correlate with aging (Issa, 2003). Similarly, high-throughput analysis of the methylation status of CpGs islands has also shown age-dependent hypermethylation of PRC2 target gene promoters across multiple cell types, including blood, ovarian cancer, cervix, and MSCs (Teschendorff *et al.*, 2010). These results have been confirmed in aging SCs of the human colon, which identified a global increase of DNA methylation in conjunction with the accumulation of methylation errors (Yatabe *et al.*, 2001). More recently, a hypothesis linking changes in epigenetics and DNA methylation in HSCs has been proposed to explain the bias in differentiation toward the myeloid lineage that is seen with age (Bocker *et al.*, 2011). Therefore, understanding the epigenetic mechanisms may reveal why there are gradual shifts in SC proliferative and differentiation potential with age.

Cellular Consequences of Genetic and Epigenetic Changes

Two of the potential consequences of age-related genetic and epigenetic changes in SCs that could significantly affect the regenerative potential of SCs and tissue homeostasis are senescence and apoptosis. Senescent cells are known to accumulate in many aged tissues. The most studied markers include senescence-associated beta-galactosidase; senescence-associated heterochromatin foci, which are domains of facultative heterochromatin seen in senescent cells; phosphorylation of

the histone H2A variant H2AX (γ H2AX), which form foci and are well-established markers of double strand breaks and chromatin modifications linked to DNA damage and repair; and finally, senescent cells often express the cyclin-dependent kinase inhibitor and tumor suppressor protein p16^{INK4a}, which is a well-described biomarker of aging and senescence (Sharpless and DePinho, 2007). The importance of p16^{INK4a} in the age-associated decline in NSC and HSC proliferation has been shown by several groups (see below). Finally, direct evidence for increased apoptosis in different stem cell compartments is still lacking (Pollack and Leeuwenburgh, 2001), though apoptosis of SCs has been described in association with telomere dysfunction (Sperka *et al.*, 2012).

Aging in Adult Stem Cells

Muscle Stem Cells

Loss of skeletal muscle regenerative potential during aging

Adult skeletal muscle is a stable, post-mitotic tissue whose role is primarily contraction, which generates force and movement, as well as being an important contributor to body metabolism and heat production. The capacity of skeletal muscle to regenerate is provided by a unique population of muscle stem cells, the satellite cells, which reside beneath the basal lamina of muscle fibers and normally reside in a quiescent state until activated by damage or growth signals (Yin *et al.*, 2013). Skeletal muscle aging is associated with a loss of muscle mass and function known as sarcopenia (Zwetsloot *et al.*, 2013), and it is associated with a progressive decline in muscle stem cell function (Hikida, 2011). Here we briefly discuss the key findings of recent years related to satellite cell aging.

Extrinsic factors associated with satellite cell aging

Extrinsic factors known to influence satellite cell behavior include the structure and function of the niche and extrinsic signals which may be derived from systemic sources or from local inflammation during tissue remodeling. Regarding the niche, a key recent study found that niche-derived FGF2 in aged muscle fibers caused a down-regulation of Sprouty1 in satellite cells, which correlated with the increased tendency of aged satellite cells to exit quiescence or to undergo self-renewal (Chakkalakal *et al.*, 2012). These authors had previously shown that Sprouty1, a receptor tyrosine kinase signaling inhibitor, could negatively regulate FGF signaling and was necessary for self-renewal and return to quiescence in satellite cell subsets during regeneration (Abou-Khalil and Brack, 2010; Shea *et al.*, 2010).

Other studies have shown that the levels of TGF β , but not the related protein myostatin, are increased in aged muscle (Carlson *et al.*, 2008). Increased TGF β signaling enhances Smad transcription factor activation in aged satellite cells, which antagonizes Notch-dependent transcription of the cyclin-dependent kinase inhibitors p15^{INK4b}, p21^{CIP1}, and p27^{KIP1}. Therefore, enhanced Smad3 activation in old satellite cells inhibits satellite cell proliferation and limits the regenerative capacity of aged muscle. Consistent with this, inhibition of Notch induces these cyclin-dependent kinase inhibitors, whereas forced inhibition of Smad3 signaling could

partially rescue the muscle regeneration defect in old mice, suggesting that dysregulation of the endogenous Notch/Smad3 balance in aged satellite cells alters regenerative capacity (Carlson *et al.*, 2008).

Elevated levels of Wnt have also been found in the circulation of aged individuals, and neutralization of Wnt signaling in aged mice was able to restore efficient muscle repair (Brack *et al.*, 2007). More recent studies have shown that effective regeneration of muscle requires a switch from high Notch activation to high Wnt activation, which promotes satellite cell differentiation, in part by inactivating glycogen synthase kinase 3 beta (GSK3 β), which is normally maintained in the active state by Notch (Brack *et al.*, 2008). These results may also reflect the proposed role of the Notch pathway in satellite cell homeostasis, as loss of Notch signaling may help satellite cells bypass self-renewal in favor of differentiation, thereby reducing the capacity of satellite cell numbers to be replenished (Mourikis *et al.*, 2011; Bjornson *et al.*, 2011). Similarly, aged satellite cells showed delayed cell cycle entry and exit (Zwetsloot *et al.*, 2013) and also fail to appropriately up-regulate the Notch ligand Delta1. Consequently age-associated regenerative defects can be reproduced in young satellite cells by inhibition of Notch signaling, whereas muscle regeneration in aged mice can be partially restored by delivery of active Notch (Conboy *et al.*, 2003, 2005; Wagers and Conboy, 2005).

Some of the most compelling evidence for the impact of age-associated changes in the stem cell niche has come from elegant experiments that have experimentally restored 'youthful' characteristics to the stem cell environment and thus improved regeneration in aged rodents. Heterochronic tissue transplants (Harrison, 1983; Carlson and Faulkner, 1983) and heterochronic parabiosis (Brack *et al.*, 2007; Villeda *et al.*, 2011; Conboy *et al.*, 2005), in which young donor material (cells, blood, tissue extracts, myofibers, etc.) are transplanted into old hosts, or vice versa, have been shown to rejuvenate muscle stem cells. In the earliest studies, the grafting of minced muscle extracts from old mice onto muscle beds of young mice, or heterotransplantation of old muscles into young mice, improved regeneration to comparable levels in young animals (Gutmann and Carlson, 1976; Carlson and Faulkner, 1989; Roberts *et al.*, 1997). Further studies using heterochronic parabiosis also support the idea that the young environment is able to rejuvenate the regenerative potential of older satellite cells. In these technically challenging experiments, the circulatory system of young and old animals is joined, and after muscle injury regeneration of old muscle was as efficient as in young muscle, while the opposite was also true, with young muscle showing a diminished regenerative capacity (Conboy *et al.*, 2005; Brack and Rando, 2007; Villeda *et al.*, 2011). In agreement with other studies described above, restoration of Notch signaling was shown to be an important mechanism both *in vivo* after heterochronic parabiosis, and *in vitro* after treatment of aged satellite cells with young serum to restore replicative capacity (Carlson and Conboy, 2007; Conboy *et al.*, 2005; Wagers and Conboy, 2005). Treatment of young satellite cells with serum from old mice has the opposite effect, as would be expected (Carlson and Conboy, 2007). Recent publications searching for the circulating protein responsible for the heterochronic parabiosis-induced effects have isolated growth differentiation factor 11 (GDF11) (Sinha *et al.*, 2014)

and oxytocin (Elabd *et al.*, 2014) as possible rejuvenating factors. Overall, these results clearly suggest that there is both an absence of pro-myogenic factors and a presence of inhibitory factors in old mice that have important consequences on satellite cell, and perhaps SC, function.

These results have proved controversial, since experiments with heterotransplanted whole muscle grafts between old donors and young mice showed a delayed but proficient regeneration when compared with controls (Shavlakadze *et al.*, 2010; Smythe *et al.*, 2008). It was proposed that the slow early regeneration rate is due to delayed inflammatory and angiogenic responses in aged mice. However, it is still not clear whether dysregulated and persistent inflammation plays a direct or indirect role in age-associated satellite cell dysfunction.

Intrinsic modifications in aged satellite cells

Cell-intrinsic aging effects have been linked to loss of satellite cell function with age; these include oxidative damage, DNA damage and genome instability, and alterations in mitochondrial function (Garcia-Prat *et al.*, 2013). In the absence of Ku80, a component of the NHEJ pathway as described above, skeletal muscle develops an accelerated aging phenotype that was hypothesized to be due to telomere shortening and accumulated DNA damage (Didier *et al.*, 2012). Excessive demands on repair over a lifetime or in chronic disease may cause successive rounds of satellite cell activation and proliferation and give rise to telomere erosion, reducing the proliferative potential of aged satellite cells (Garcia-Prat *et al.*, 2013; Jung and Brack, 2014).

High-throughput gene expression profiling of young and old human satellite cells has highlighted differences in the transcriptional program that could alter function, including dysregulation of myogenic differentiation-specific genes, increased expression of atrophy-related forkhead (FoxO)-regulated genes, and altered expression of genes related to protein folding (Bortoli *et al.*, 2003; Pietrangelo *et al.*, 2009). Consistent with this are studies that show an impaired differentiation capacity in aged satellite cells, including the ability to fuse into myotubes or to form a reserve population *in vitro* (Collins *et al.*, 2007; Charge *et al.*, 2002). Reduced formation of myogenic colonies and a decreased ability to activate and proliferate are other defective processes identified by others (Shadrach and Wagers, 2011; Conboy *et al.*, 2003; Shefer *et al.*, 2006; Day *et al.*, 2010; Baj *et al.*, 2005). Moreover, it is now widely accepted that aged satellite cells show a tendency to adopt fibroblastic and adipogenic fates (Brack *et al.*, 2007), which may explain the increased levels of fat and fibrotic tissue in aged muscle. These changes may arise from a global epigenetic switch from H3K4me3 at the TSS to H3K27me3, coinciding with an increase in dsDNA breaks and a reduction in histone biosynthesis (Liu *et al.*, 2013).

Supporting the relevance of age-associated intrinsic alterations in satellite cell regenerative functions, a recent study showed that in geriatric animals (28 months of age or older), satellite cells are in an irreversible pre-senescent state that affects their intrinsic regenerative and self-renewal ability (Sousa-Victor *et al.*, 2014). Maintenance of quiescence in adult life appears to depend in part on the active repression of senescent pathways, and in particular on repression of the p16^{INK4a} locus. Silencing of p16^{INK4a} is sufficient to rejuvenate

geriatric satellite cells, by restoring their activation potential and self-renewal capacity *in vivo* and *in vitro* and reestablishing regeneration efficiency (Sousa-Victor *et al.*, 2014). Finally, recent papers have also proposed rejuvenation strategies that act, independently of the environment, to reverse satellite cell-intrinsic changes that occur during the aging process (Cosgrove *et al.*, 2014; Bernet *et al.*, 2014; Tierney *et al.*, 2014; Price *et al.*, 2014). The authors show that p38 α / β MAPK and JAK/STAT activities are elevated in adult and aged satellite cells and that inhibition of any of these pathways is sufficient to restore satellite cell activity.

MSCs

MSCs, also called multipotent stromal cells or mesenchymal stromal cells, can be isolated from a variety of tissues, including bone marrow, adipose tissue, dental pulp, amniotic fluid, Wharton's jelly, and umbilical cord blood. Morphologically, MSCs are frequently described as being fibroblast-like, with a spindle shape and a large round nucleus; they have the capacity to differentiate *in vitro* into a vast range of cell types including mesenchymal cells such as adipogenic, osteogenic, and chondrogenic cells; they can also be induced to transdifferentiate into epithelial cells, such as gastrointestinal tract epithelial cells, alveolar epithelial cells, kidney cells, and also such endodermal cells as hepatocytes (Lee *et al.*, 2004). Functionally, MSCs have been widely described as having important immunomodulatory properties, including the ability to modulate T-cell and B-cell functions by affecting processes such as cell activation, proliferation, and chemotaxis mediated by direct cell-to-cell contact and/or the secretion of soluble factors (Yu and Kang, 2013; Raveh-Amit *et al.*, 2013).

Age-related changes in MSCs

The production of large numbers of purified MSCs for therapeutic purposes has created concern about potential age-associated effects due to donor age and/or extensive proliferation during manufacture. Indeed, there is sufficient evidence of age-related changes in MSCs, and therefore one of the most pressing research questions to emerge has been to identify the molecular basis for these changes. Not surprisingly, epigenetic alterations in MSCs have emerged as a possible cause of shifts in function, whereas several studies have also identified the importance of senescence programs. Culture of MSCs was shown to greatly affect epigenetic and gene expression patterns, increasing osteogenic genes and decreasing the 'stemness' of the cells. Epigenetic dysregulations include alterations in histone H3 acetylation, as well as changes in the methylation patterns of CpG islands in the promoter and the region of exon 1 in many genes (Li *et al.*, 2011).

Previous data showed that administration of HDAC inhibitors induced p21^{CIP1} expression and arrested MSCs in the G2/M phase of the cell cycle, and ultimately also induced p16^{INK4A} expression and cellular senescence, in a mechanism implicating the up-regulation of jumonji domain-containing 3 (JMJD3), which demethylates H3K27Me3 at the CDKN2 locus, thereby increasing p16^{INK4A} expression and down-regulation of several PcGs genes (Jung *et al.*, 2010). Overall, this study suggested that p16^{INK4A} is repressed in young MSCs

because of PcG-induced H3K27Me3, but that HDAC inhibition mimics aging by allowing p16^{INK4A} expression and cellular senescence. Importantly, this phenotype could be reversed in the presence of HAT inhibitors, which suggests a simple way by which premature senescence could be overcome, thereby enhancing the potential of MSCs. Consistent with these results, small interfering RNAs targeting p16^{INK4A} in senescent MSCs reduced the number of senescent cells and restored the ability to proliferate (Shibata *et al.*, 2007). Moreover, up-regulated p16^{INK4A} expression and down-regulated CDK4, CDK6, and RB, in conjunction with loss of proliferation, differentiation potential, and immunogenic properties, were identified in MSCs from patients with systemic lupus erythematosus (SLE), a systemic autoimmune disease in which the immune system attacks the heart, joints, skin, lungs, blood vessels, liver, kidneys, and nervous system. These effects were counteracted by knockdown of p16^{INK4A} expression, revealing a potential therapeutic strategy for treating these patients (Gu *et al.*, 2012).

It is worth noting that MSCs possess several important characteristics that make them potentially ideal as cellular therapies for a range of diseases, including those with significant morbidity and mortality in the elderly. These properties include immunomodulatory functions, an immune privileged status, easy isolation from a range of tissues, and ability to cross the blood-brain barrier, and are therefore of great interest in targeting many diseases, including cardiovascular disease and neurodegenerative disorders (Yu and Kang, 2013; Raveh-Amit *et al.*, 2013).

HSCs

Stem cell biology and the effects of aging on SCs have been most widely studied in the hematopoietic compartment. HSCs are a heterogeneous population of cells that reside in the red bone marrow (BM); they undergo self-renewal as well as proliferation and differentiation to repopulate the different blood lineages, giving rise to myeloid cells (monocytes and macrophages, neutrophils, basophils, eosinophils, erythrocytes, megakaryocytes/platelets, dendritic cells) and lymphoid cells (T-cells, B-cells, and NK-cells) (Orkin and Zon, 2008). One of the most surprising aspects of HSCs and the BM niche is that the number of HSCs increases with age, which is in contrast to most other SC populations (Beerman *et al.*, 2013, 2010). The reasons for this are not completely understood and are probably multifactorial. Clear functional differences between old and young HSCs have been shown. For example, transplantation of young and old HSCs together into lethally irradiated young recipients shows a greatly reduced potential for grafting by older cells (Morrison *et al.*, 1996). In this sense, it is also well described that aged HSCs have a preference for differentiation toward the myeloid lineages, as already stated above, which leads to an overall decrease in the regenerative potential of HSCs and the overall capacity of the immune system (Geiger *et al.*, 2013; Beerman *et al.*, 2010). What is not clear is the interplay over time between cell-intrinsic and cell-extrinsic factors in facilitating these changes. These factors are discussed briefly below.

One of the hallmarks of aging in HSCs is a differentiation potential that favors the myeloid lineage at the expense of the

lymphoid and erythroid lineages (Geiger *et al.*, 2013). Many mechanisms have already been proposed for this observation, but among the alternative hypotheses are that HSCs do not lose their differentiation potential *per se*, but that preference for the myeloid lineage merely reflects changes in composition of the HSC pool, presumably due to defects in self-renewal of some HSC clonal subtypes during aging, though other mechanisms have been postulated, such as accumulation of DNA damage (Beerman *et al.*, 2010).

Several groups have shown a clear role for the loss of genome integrity and DNA damage as a cause of altered HSC function with age, similar to the models of progeroid syndromes described above. Mutations or deletions in various DNA repair components have been shown to modify HSC behavior and function. Hypomorphic mutations in DNA ligase IV (LIG4), a component of the NHEJ, are associated with multiple defects in lymphocyte development, function, survival, and proliferation (Nijnik *et al.*, 2007). Knockout of another NHEJ protein, Ku80, induce severe defects in T- and B-lymphocyte development and profound deficiencies in V(D)J recombination, the process by which B-cell and T-cell receptor diversity is generated. However, mutation to other DNA repair processes not related to V(D)J recombination are also involved in loss of HSC potential, suggesting that other mechanisms are also at play. For example, deletion of exon 27 of the BRCA2/FANCD1 gene, which is involved in homologous recombination, results in marked defects in proliferation and in the self-renewing ability of HSCs in a mouse model of Fanconi anemia (Navarro *et al.*, 2006). Collectively, these data suggest that an increased mutation rate, as in aging, may give rise to proliferative advantages, failure of adequate differentiation, and ultimately malignant transformation.

There are also links between DNA repair and the accumulation of damage due to ROS. The ataxia telangiectasia mutated (ATM) protein is a serine/threonine protein kinase that is activated by DNA double strand breaks, where it phosphorylates several key proteins, including tumor suppressors, to initiate activation of the DNA damage checkpoint, cell cycle arrest, DNA repair, or apoptosis. Activation of p38 MAPK in response to increasing levels of ROS normally limits the lifespan of HSCs *in vivo*, but in *Atm*^{-/-} knockout mice, elevation of ROS levels induces HSC-specific phosphorylation of p38 MAPK accompanied by a defect in the maintenance of HSC quiescence, defects in repopulating capacity, and ultimately depletion of the HSC pool (Ito *et al.*, 2006). These defects in repopulation potential can be overcome by suitable treatment with antioxidants, demonstrating a link between DNA repair and ROS with age. Deletion of autophagy-related protein 7 (Atg7) (Mortensen *et al.*, 2011) in adult mice also increases ROS levels and consequently causes HSC depletion, linking ROS and autophagy HSC homeostasis and aging. Interestingly, a recent report demonstrated that autophagy is essential to protect HSCs from metabolic stress (i.e., induced by starvation or cytokines) in a mechanism controlled by FOXO3A and maintained in old HSCs (Warr *et al.*, 2013). Interestingly, recent new data on HSC suggest that accumulation of DNA damage with age is related to a reduction of the DNA repair mechanisms when HSCs are in quiescence. When stimulated to enter the cell cycle, only fetal and young HSCs are able to up-regulate the DNA response and repair pathways to resolve that accumulated damage (Beerman *et al.*, 2014).

In addition to their roles in regulating quiescence as described above for satellite cells, FoxO transcription factors have been shown to be important regulators of oxidative stress throughout the lifetime of an organism, and by association also affect changes in HSC function with aging. FoxO3a conditional ablation in mice was shown to impair HSC self-renewal, reducing the frequency of HSCs with age and diminishing the long-term reconstitution of hematopoiesis in a competitive transplantation assay. In addition, ROS levels in HSC were elevated, and maintenance of quiescence was also impaired (Miyamoto *et al.*, 2007). These results were reproduced by FoxO1, FoxO3, and FoxO4 triple deletion in the adult hematopoietic system (Tothova *et al.*, 2007), which showed defective self-renewal, a shift toward the myeloid lineage, and increases in ROS that correlated with changes in expression of genes that regulate ROS. Importantly, treatment of the deleted mice *in vivo* with the antioxidative agent *N*-acetyl-L-cysteine reverted the FoxO-deficient HSC phenotype. Collectively, these data suggest that there is a subset of ROS genes that are regulated by FoxO proteins in the HSC compartment and that these genes function by reducing ROS levels and thereby removing any potentially deleterious effects on HSC survival and function.

Interestingly, the phenotype in FoxO-deficient mice is observed in upstream components of the PI3K/AKT/mTOR pathway, as in phosphatase and tensin homolog (PTEN)-deficient HSCs (Yilmaz *et al.*, 2006). PTEN acts by dephosphorylating phosphatidylinositol (3,4,5)-trisphosphate (PIP3) and is a tumor suppressor and negative regulator of mTOR signaling which prevents cell cycle progression. The fact that the mTOR inhibitor rapamycin could revert the phenotype of PTEN-deficient HSCs suggests that mTOR signaling may be involved in the HSC phenotype (Chen *et al.*, 2009). mTOR is a serine/threonine (Ser/Thr)-protein kinase that exists in two independent multiprotein-containing complexes, mTORC1 and mTORC2, and which are generally activated under conditions of nutrient availability. That is, mTORC1 is an energy sensor, and its expression is known to increase with age in HSCs, where it may also act as an attenuator of autophagy (Nicklin *et al.*, 2009). In young HSCs, conditional deletion of PTEN leads to indirect activation of mTOR, which causes a progeria-like phenotype and depletion of HSCs (Magee *et al.*, 2012; Kalaitzidis *et al.*, 2012). Overall, these studies suggest that either increased or decreased mTORC1 levels result in diminished capacity for HSC function and that careful regulation of mTOR by hematopoietic growth factor/cytokine signaling mediates self-renewal under regenerative conditions. Calorie restriction can attenuate mTORC1 signaling and other proteins associated with calorie restriction and metabolic effects in HSCs and tissue including HDACs. Both Sirt1 and Sirt3 HDACs have been shown to have roles in HSC function with age, whereby deletion of Sirt1 in adult HSCs increases HSC proliferation and DNA damage, whereas Sirt3 controls the activity of the mitochondrial enzyme acetyl co-enzyme A synthetase 2 (AceCS2) and electron transport chain complex 1 in mitochondria (Roth *et al.*, 2013). Deletion of Sirt1 therefore leads to loss HSC populations capable of long-term hematopoiesis (Singh *et al.*, 2013), whereas the metabolic consequences of Sirt3 deletion greatly diminish HSC function, while its up-regulation in aged HSCs improves their regenerative functions (Brown *et al.*, 2013).

Although the overall effects are modest compared to similar results in the other model SC systems, knockdown or deletion of p16^{INK4a} in HSCs ameliorates at least some of the age-associated phenotypes such as increased HSC number and aberrant self-renewal capacity, which agrees with the results reported in satellite cells (Janzen *et al.*, 2006; Sousa-Victor *et al.*, 2014). Similar roles for proteins that also function in these processes, including p53, the anti-apoptotic members of the BCL-2 family and the recently identified B-cell-activating transcription factor (BATF), have been proposed (Nitta *et al.*, 2011; Wang *et al.*, 2012). In the accelerated aging p53^{-/-} mice, which express a truncated p53 but are also haploinsufficient for 24 genes upstream of p53, animals show reduced lifespan, and less self-renewal and differentiation potential of HSCs (Donehower, 2009). In another model, the so-called super p53/p19^{ARF} mice, which have additional p53 and p19^{ARF} alleles, mice showed longer lifespan, perhaps due to reduced proliferation and maintenance of stem cell functionality at more optimal levels in the presence of a more robust ARF-p53 pathway (Matheu *et al.*, 2007). Overall, this suggests that an overactive p53 inhibits stem and progenitor cell function, whereas if properly regulated and boosted by increased ARF expression, it may have beneficial effects on stem cells later in life (Donehower, 2009). However, the full implication and biological relevance of these results in aging need to be determined. In this context, a recent publication has shown that functional decline of old HSCs were related to amplified levels of replication stress associated with cell cycle defects and chromosome breaks (Flach *et al.*, 2014). Mechanistically, a decreased expression of mini-chromosome maintenance (MCM) helicase components and altered dynamics of DNA replication forks resulted in residual replication stress on ribosomal DNA and formation of nucleolar-associated γ H2AX signals, which may account for HSC functional decline and open avenues for rejuvenation treatments (Flach *et al.*, 2014).

Epigenetic analysis has also revealed changes in old compared to young HSCs. It has been shown that genes associated with the myeloid lineage become hypomethylated with age, favoring transcription and therefore supporting the various observations of age-associated lineage skewing, whereas site-specific hypermethylation was observed at genes associated with the PRC2 complex (Beerman *et al.*, 2013; Bocker *et al.*, 2011). It has also been demonstrated that replicative aging is distinct from chronological aging at the DNA methylation level, suggesting that this age-dependent preference for myeloid differentiation by HSCs is a product of the proliferative history of HSCs (Beerman *et al.*, 2013). The evidence for this is that HSCs that were cultured under moderate proliferative demands underwent site-specific hypermethylation at the PRC2 complex, accompanied by transcriptional repression (Chambers *et al.*, 2007), whereas after enforced proliferation of adult HSCs, there was a highly distinct methylation pattern. Systematic analysis of the transcriptome and epigenome also suggested a close correlation between the typical phenotype of aged HSCs, mainly increased self-renewal, diminished differentiation potential, and biased preference for myeloid differentiation, and epigenetic alterations (Sun *et al.*, 2014). Briefly, age was associated with an increase in H3K4me3, particularly on genes associated with stem cell self-renewal and loss of differentiation capacity, which is similar to reports from

satellite cells in muscle, where H3K4me3 length was also observed to increase with age (Liu *et al.*, 2013). These age-related changes were also suggestive of decreased TGF β signaling, in conjunction with decreases in the phosphorylation and/or expression of various downstream effectors such as the various Smad proteins (Sun *et al.*, 2014).

Concluding Remarks

Organs and tissues deteriorate progressively with age. In recent years, it has become evident that as organisms age, tissue-resident stem cells also deteriorate, as illustrated by their impaired capacity to respond to an insult or maintain homeostasis. Extrinsic changes in the environment as well as cell-autonomous alterations with aging account for the age-associated deterioration of stem cell functions. Additional studies will be needed to better understand the specific extrinsic factors associated with defective stem cell-dependent tissue repair in aged individuals and the rejuvenation of old stem cell capacity by a young environment. Likewise, it will be necessary to identify how aging directly impacts on stem cell-intrinsic functions and whether these can be linked to epigenetic patterns. Further understanding of these signaling networks operating on stem cells will facilitate strategies to slow down or reverse the onset of aging on tissues and organs and to improve their repair in aging individuals.

Acknowledgments

The authors acknowledge funding from MICINN (SAF2012-38547, PLE2009-0124, CIBERNED), AFM, Fundación Marató-TV3, MDA, EU-FP7 (Myoage, Optistem, and Endostem).

See also: Stem Cell Biology: Structure and Function: Metabolic and Energetic Regulation of Stem Cells; The Adult Stem Cell Niche: Multiple Cellular Players in Tissue Homeostasis and Regeneration

References

- Abou-Khalil, R., Brack, A.S., 2010. Muscle stem cells and reversible quiescence: The role of sprouty. *Cell Cycle* 9, 2575–2580.
- Baj, A., Bettaccini, A.A., Casalone, R., *et al.*, 2005. Culture of skeletal myoblasts from human donors aged over 40 years: Dynamics of cell growth and expression of differentiation markers. *Journal of Translational Medicine* 3, 21.
- Beerman, I., Bhattacharya, D., Zandi, S., *et al.*, 2010. Functionally distinct hematopoietic stem cells modulate hematopoietic lineage potential during aging by a mechanism of clonal expansion. *Proceedings of the National Academy of Sciences of the United States of America* 107, 5465–5470.
- Beerman, I., Bock, C., Garrison, B.S., *et al.*, 2013. Proliferation-dependent alterations of the DNA methylation landscape underlie hematopoietic stem cell aging. *Cell Stem Cell* 12, 413–425.
- Beerman, I., Seita, J., Inlay, M.A., Weissman, I.L., Rossi, D.J., 2014. Quiescent hematopoietic stem cells accumulate DNA damage during aging that is repaired upon entry into cell cycle. *Cell Stem Cell* 15, 37–50.
- Berger, S.L., 2007. The complex language of chromatin regulation during transcription. *Nature* 447, 407–412.

- Bernet, J.D., Doles, J.D., Hall, J.K., *et al.*, 2014. p38 MAPK signaling underlies a cell-autonomous loss of stem cell self-renewal in skeletal muscle of aged mice. *Nature Medicine* 20, 265–271.
- Bjornson, C.R., Cheung, T.H., Liu, L., *et al.*, 2011. Notch signaling is necessary to maintain quiescence in adult muscle stem cells. *Stem Cells* 30, 232–242.
- Bocker, M.T., Hellwig, I., Breiling, A., *et al.*, 2011. Genome-wide promoter DNA methylation dynamics of human hematopoietic progenitor cells during differentiation and aging. *Blood* 117, e182–e189.
- Bortoli, S., Renault, V., Eveno, E., *et al.*, 2003. Gene expression profiling of human satellite cells during muscular aging using cDNA arrays. *Gene* 321, 145–154.
- Brack, A.S., Conboy, I.M., Conboy, M.J., Shen, J., Rando, T.A., 2008. A temporal switch from notch to Wnt signaling in muscle stem cells is necessary for normal adult myogenesis. *Cell Stem Cell* 2, 50–59.
- Brack, A.S., Conboy, M.J., Roy, S., *et al.*, 2007. Increased Wnt signaling during aging alters muscle stem cell fate and increases fibrosis. *Science* 317, 807–810.
- Brack, A.S., Rando, T.A., 2007. Intrinsic changes and extrinsic influences of myogenic stem cell function during aging. *Stem Cell Reviews* 3, 226–237.
- Brown, K., Xie, S., Qiu, X., *et al.*, 2013. SIRT3 reverses aging-associated degeneration. *Cell Reports* 3, 319–327.
- Burke, B., Stewart, C.L., 2014. Functional architecture of the cell's nucleus in development, aging, and disease. *Current Topics in Developmental Biology* 109, 1–52.
- Carlson, B.M., Faulkner, J.A., 1983. The regeneration of skeletal muscle fibers following injury: A review. *Medicine and Science in Sports and Exercise* 15, 187–198.
- Carlson, B.M., Faulkner, J.A., 1989. Muscle transplantation between young and old rats: Age of host determines recovery. *American Journal of Physiology* 256, C1262–C1266.
- Carlson, M.E., Conboy, I.M., 2007. Loss of stem cell regenerative capacity within aged niches. *Aging Cell* 6, 371–382.
- Carlson, M.E., Hsu, M., Conboy, I.M., 2008. Imbalance between pSmad3 and Notch induces CDK inhibitors in old muscle stem cells. *Nature* 454, 528–532.
- Chakkalakal, J.V., Jones, K.M., Basson, M.A., Brack, A.S., 2012. The aged niche disrupts muscle stem cell quiescence. *Nature* 490, 355–360.
- Chambers, S.M., Shaw, C.A., Gatz, C., *et al.*, 2007. Aging hematopoietic stem cells decline in function and exhibit epigenetic dysregulation. *PLoS Biology* 5, e201.
- Charge, S.B., Brack, A.S., Hughes, S.M., 2002. Aging-related satellite cell differentiation defect occurs prematurely after Ski-induced muscle hypertrophy. *American Journal of Physiology – Cell Physiology* 283, C1228–C1241.
- Chen, C., Liu, Y., Liu, Y., Zheng, P., 2009. mTOR regulation and therapeutic rejuvenation of aging hematopoietic stem cells. *Science Signaling* 2, ra75.
- Collins, C.A., Zammitt, P.S., Ruiz, A.P., Morgan, J.E., Partridge, T.A., 2007. A population of myogenic stem cells that survives skeletal muscle aging. *Stem Cells* 25, 885–894.
- Conboy, I.M., Conboy, M.J., Smythe, G.M., Rando, T.A., 2003. Notch-mediated restoration of regenerative potential to aged muscle. *Science* 302, 1575–1577.
- Conboy, I.M., Conboy, M.J., Wagers, A.J., *et al.*, 2005. Rejuvenation of aged progenitor cells by exposure to a young systemic environment. *Nature* 433, 760–764.
- Cosgrove, B.D., Gilbert, P.M., Porpiglia, E., *et al.*, 2014. Rejuvenation of the muscle stem cell population restores strength to injured aged muscles. *Nature Medicine* 20, 255–264.
- D'Adda Di Fagagna, F., Reaper, P.M., Clay-Farrace, L., *et al.*, 2003. A DNA damage checkpoint response in telomere-initiated senescence. *Nature* 426, 194–198.
- Day, K., Shefer, G., Shearer, A., Yablonka-Reuveni, Z., 2010. The depletion of skeletal muscle satellite cells with age is concomitant with reduced capacity of single progenitors to produce reserve progeny. *Developmental Biology* 340, 330–343.
- Didier, N., Hourde, C., Amthor, H., Marazzi, G., Sassoon, D., 2012. Loss of a single allele for Ku80 leads to progenitor dysfunction and accelerated aging in skeletal muscle. *EMBO Molecular Medicine* 4, 910–923.
- Donehower, L.A., 2009. Using mice to examine p53 functions in cancer, aging, and longevity. *Cold Spring Harbor Perspectives in Biology* 1, a001081.
- Elabd, C., Cousin, W., Upadhyayula, P., *et al.*, 2014. Oxytocin is an age-specific circulating hormone that is necessary for muscle maintenance and regeneration. *Nature Publishing Group* 5, 4082.
- Flach, J., Bakker, S.T., Mohrin, M., *et al.*, 2014. Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature* 512, 198–202.
- Flores, I., Blasco, M.A., 2010. The role of telomeres and telomerase in stem cell aging. *FEBS Letters* 584, 3826–3830.
- Flores, I., Canela, A., Vera, E., *et al.*, 2008. The longest telomeres: A general signature of adult stem cell compartments. *Genes & Development* 22, 654–667.
- Garcia-Prat, L., Sousa-Victor, P., Munoz-Canoves, P., 2013. Functional dysregulation of stem cells during aging: A focus on skeletal muscle stem cells. *FEBS Journal* 280, 4051–4062.
- Geiger, H., De Haan, G., Florian, M.C., 2013. The ageing haematopoietic stem cell compartment. *Nature Reviews Immunology* 13, 376–389.
- Greer, E.L., Maures, T.J., Hauswirth, A.G., *et al.*, 2010. Members of the H3K4 trimethylation complex regulate lifespan in a germline-dependent manner in *C. elegans*. *Nature* 466, 383–387.
- Gu, Z., Cao, X., Jiang, J., *et al.*, 2012. Upregulation of p16INK4A promotes cellular senescence of bone marrow-derived mesenchymal stem cells from systemic lupus erythematosus patients. *Cell Signal* 24, 2307–2314.
- Gutmann, E., Carlson, B.M., 1976. Regeneration and transplantation of muscles in old rats and between young and old rats. *Life Sciences* 18, 109–114.
- Harrison, D.E., 1983. Long-term erythropoietic repopulating ability of old, young, and fetal stem cells. *Journal of Experimental Medicine* 157, 1496–1504.
- Hayflick, L., Moorhead, P.S., 1961. The serial cultivation of human diploid cell strains. *Experimental Cell Research*, 25, 585–621.
- Hikida, R.S., 2011. Aging changes in satellite cells and their functions. *Current Aging Science* 4, 279–297.
- Issa, J.P., 2003. Age-related epigenetic changes and the immune system. *Clinical Immunology* 109, 103–108.
- Issa, J.P., 2014. Aging and epigenetic drift: A vicious cycle. *Journal of Clinical Investigation* 124, 24–29.
- Ito, K., Hirao, A., Arai, F., *et al.*, 2006. Reactive oxygen species act through p38 MAPK to limit the lifespan of hematopoietic stem cells. *Nature Medicine* 12, 446–451.
- Janzen, V., Forkert, R., Fleming, H.E., *et al.*, 2006. Stem-cell ageing modified by the cyclin-dependent kinase inhibitor p16INK4a. *Nature* 443, 421–426.
- Jasper, H., Kennedy, B.K., 2012. Niche science: The aging stem cell. *Cell Cycle* 11, 2959–2960.
- Jones, D.L., Rando, T.A., 2011. Emerging models and paradigms for stem cell aging. *Nature Cell Biology* 13, 506–512.
- Jung, J.W., Lee, S., Seo, M.S., *et al.*, 2010. Histone deacetylase controls adult stem cell aging by balancing the expression of polycomb genes and jumoni domain containing 3. *Cellular and Molecular Life Sciences* 67, 1165–1176.
- Jung, Y., Brack, A.S., 2014. Cellular mechanisms of somatic stem cell aging. *Current Topics in Developmental Biology* 107, 405–438.
- Kalaitzidis, D., Sykes, S.M., Wang, Z., *et al.*, 2012. mTOR complex 1 plays critical roles in hematopoiesis and Pten-loss-evoked leukemogenesis. *Cell Stem Cell* 11, 429–439.
- Lee, K.D., Kuo, T.K., Whang-Peng, J., *et al.*, 2004. *In vitro* hepatic differentiation of human mesenchymal stem cells. *Hepatology* 40, 1275–1284.
- Li, L., Clevers, H., 2010. Coexistence of quiescent and active adult stem cells in mammals. *Science* 327, 542–545.
- Li, Z., Liu, C., Xie, Z., *et al.*, 2011. Epigenetic dysregulation in mesenchymal stem cell aging and spontaneous differentiation. *PLoS One* 6, e20526.
- Liu, L., Cheung, T.H., Charville, G.W., *et al.*, 2013. Chromatin modifications as determinants of muscle stem cell quiescence and chronological aging. *Cell Reports* 4, 189–204.
- Liu, L., Rando, T.A., 2011. Manifestations and mechanisms of stem cell aging. *Journal of Cell Biology* 193, 257–266.
- Longo, V.D., Kennedy, B.K., 2006. Sirtuins in aging and age-related disease. *Cell* 126, 257–268.
- Lopez-Otin, C., Blasco, M.A., Partridge, L., Serrano, M., Kroemer, G., 2013. The hallmarks of aging. *Cell* 153, 1194–1217.
- Magee, J.A., Ikenoue, T., Nakada, D., *et al.*, 2012. Temporal changes in PTEN and mTORC2 regulation of hematopoietic stem cell self-renewal and leukemia suppression. *Cell Stem Cell* 11, 415–428.
- Matheu, A., Maraver, A., Klatt, P., *et al.*, 2007. Delayed ageing through damage protection by the Arf/p53 pathway. *Nature* 448, 375–379.
- Miyamoto, K., Araki, K.Y., Naka, K., *et al.*, 2007. Foxo3a is essential for maintenance of the hematopoietic stem cell pool. *Cell Stem Cell* 1, 101–112.
- Mohrin, M., Bourke, E., Alexander, D., *et al.*, 2010. Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. *Cell Stem Cell* 7, 174–185.
- Morrison, S.J., Wandycz, A.M., Akashi, K., *et al.*, 1996. The aging of hematopoietic stem cells [see comments]. *Nature Medicine* 2, 1011–1016.
- Mortensen, M., Soilleux, E.J., Djordjevic, G., *et al.*, 2011. The autophagy protein Atg7 is essential for hematopoietic stem cell maintenance. *Journal of Experimental Medicine* 208, 455–467.
- Mourikis, P., Sambasivan, R., Castel, D., *et al.*, 2011. A critical requirement for notch signaling in maintenance of the quiescent skeletal muscle stem cell state. *Stem Cells* 30, 243–252.

- Navarro, S., Meza, N.W., Quintana-Bustamante, O., *et al.*, 2006. Hematopoietic dysfunction in a mouse model for Fanconi anemia group D1. *Molecular Therapy* 14, 525–535.
- NIA, 2011. Global Health and Aging [Online]. National Institute of Aging. Available at: <http://www.nia.nih.gov/research/publication/global-health-and-aging/> (accessed 09.04.15).
- Nicklin, P., Bergman, P., Zhang, B., *et al.*, 2009. Bidirectional transport of amino acids regulates mTOR and autophagy. *Cell* 136, 521–534.
- Nijnik, A., Woodbine, L., Marchetti, C., *et al.*, 2007. DNA repair is limiting for haematopoietic stem cells during ageing. *Nature* 447, 686–690.
- Nitta, E., Yamashita, M., Hosokawa, K., *et al.*, 2011. Telomerase reverse transcriptase protects ATM-deficient hematopoietic stem cells from ROS-induced apoptosis through a telomere-independent mechanism. *Blood* 117, 4169–4180.
- Orkin, S.H., Zon, L.I., 2008. Hematopoiesis: An evolving paradigm for stem cell biology. *Cell* 132, 631–644.
- Pietrangelo, T., Pugliesi, C., Mancinelli, R., *et al.*, 2009. Molecular basis of the myogenic profile of aged human skeletal muscle satellite cells during differentiation. *Experimental Gerontology* 44, 523–531.
- Pollack, M., Leeuwenburgh, C., 2001. Apoptosis and aging: Role of the mitochondria. *Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 56, B475–B482.
- Price, F.D., Von Maltzahn, J., Bentzinger, C.F., *et al.*, 2014. Inhibition of JAK-STAT signaling stimulates adult satellite cell function. *Nature Medicine* 20, 1174–1181.
- Raveh-Amit, H., Berszanyi, S., Vas, V., Ye, D., Dinnyes, A., 2013. Tissue resident stem cells: Till death do us part. *Biogerontology* 14, 573–590.
- Roberts, P., McGeachie, J.K., Grounds, M.D., 1997. The host environment determines strain-specific differences in the timing of skeletal muscle regeneration: Cross-transplantation studies between SJL/J and BALB/c mice. *Journal of Anatomy* 191 (Pt 4), 585–594.
- Rossi, D.J., Seita, J., Czechowicz, A., *et al.*, 2007. Hematopoietic stem cell quiescence attenuates DNA damage response and permits DNA damage accumulation during aging. *Cell Cycle* 6, 2371–2376.
- Roth, M., Wang, Z., Chen, W.Y., 2013. Sirtuins in hematological aging and malignancy. *Critical Reviews in Oncogenesis* 18, 531–547.
- Shadrach, J.L., Wagers, A.J., 2011. Stem cells for skeletal muscle repair. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 366, 2297–2306.
- Sharpless, N.E., DePinho, R.A., 2007. How stem cells age and why this makes us grow old. *Nature Reviews Molecular Cell Biology* 8, 703–713.
- Shavlakadze, T., McGeachie, J., Grounds, M.D., 2010. Delayed but excellent myogenic stem cell response of regenerating geriatric skeletal muscles in mice. *Biogerontology* 11, 363–376.
- Shea, K.L., Xiang, W., Laporta, V.S., *et al.*, 2010. Sprouty1 regulates reversible quiescence of a self-renewing adult muscle stem cell pool during regeneration. *Cell Stem Cell* 6, 117–129.
- Shefer, G., Van De Mark, D.P., Richardson, J.B., Yablonka-Reuveni, Z., 2006. Satellite-cell pool size does matter: Defining the myogenic potency of aging skeletal muscle. *Developmental Biology* 294, 50–66.
- Shibata, K.R., Aoyama, T., Shima, Y., *et al.*, 2007. Expression of the p16INK4A gene is associated closely with senescence of human mesenchymal stem cells and is potentially silenced by DNA methylation during *in vitro* expansion. *Stem Cells* 25, 2371–2382.
- Simons, B.D., Clevers, H., 2011. Strategies for homeostatic stem cell self-renewal in adult tissues. *Cell* 145, 851–862.
- Singh, S.K., Williams, C.A., Klarmann, K., *et al.*, 2013. Sirt1 ablation promotes stress-induced loss of epigenetic and genomic hematopoietic stem and progenitor cell maintenance. *Journal of Experimental Medicine* 210, 987–1001.
- Sinha, M., Jang, Y.C., Oh, J., *et al.*, 2014. Restoring systemic GDF11 levels reverses age-related dysfunction in mouse skeletal muscle. *Science* 344, 649–652.
- Smythe, G.M., Shavlakadze, T., Roberts, P., *et al.*, 2008. Age influences the early events of skeletal muscle regeneration: studies of whole muscle grafts transplanted between young (8 weeks) and old (13–21 months) mice. *Experimental Gerontology* 43, 550–562.
- Sousa-Victor, P., Gutarra, S., Garcia-Prat, L., *et al.*, 2014. Geriatric muscle stem cells switch reversible quiescence into senescence. *Nature* 506, 316–321.
- Sperka, T., Wang, J., Rudolph, K.L., 2012. DNA damage checkpoints in stem cells, ageing and cancer. *Nature Reviews Molecular Cell Biology* 13, 579–590.
- Sun, D., Luo, M., Jeong, M., *et al.*, 2014. Epigenomic profiling of young and aged HSCs reveals concerted changes during aging that reinforce self-renewal. *Cell Stem Cell* 14, 673–688.
- Teschendorff, A.E., Menon, U., Gentry-Maharaj, A., *et al.*, 2010. Age-dependent DNA methylation of genes that are suppressed in stem cells is a hallmark of cancer. *Genome Research* 20, 440–446.
- Tierney, M.T., Aydogdu, T., Sala, D., *et al.*, 2014. STAT3 signaling controls satellite cell expansion and skeletal muscle repair. *Nature Medicine* 20, 1182–1186.
- Tothova, Z., Kolipara, R., Huntly, B.J., *et al.*, 2007. FoxOs are critical mediators of hematological stem cell resistance to physiologic oxidative stress. *Cell* 128, 325–339.
- Villeda, S.A., Luo, J., Mosher, K.I., *et al.*, 2011. The ageing systemic milieu negatively regulates neurogenesis and cognitive function. *Nature* 477, 90–94.
- Vijayanti, V.N., Rao, K.S., 2006. DNA double strand break repair in brain: reduced NHEJ activity in aging rat neurons. *Neuroscience Letters* 393, 18–22.
- Wagers, A.J., Conboy, I.M., 2005. Cellular and molecular signatures of muscle regeneration: Current concepts and controversies in adult myogenesis. *Cell* 122, 659–667.
- Wang, J., Sun, Q., Morita, Y., *et al.*, 2012. A differentiation checkpoint limits hematopoietic stem cell self-renewal in response to DNA damage. *Cell* 148, 1001–1014.
- Warr, M.R., Binnewies, M., Flach, J., *et al.*, 2013. FOXO3A directs a protective autophagy program in hematopoietic stem cells. *Nature* 494, 323–327.
- Yatabe, Y., Tavaré, S., Shibata, D., 2001. Investigating stem cells in human colon by using methylation patterns. *Proceedings of the National Academy of Sciences of the United States of America* 98, 10839–10844.
- Yilmaz, O.H., Valdez, R., Theisen, B.K., *et al.*, 2006. Pten dependence distinguishes hematopoietic stem cells from leukaemia-initiating cells. *Nature* 441, 475–482.
- Yin, H., Price, F., Rudnicki, M.A., 2013. Satellite cells and the muscle stem cell niche. *Physiological Reviews* 93, 23–67.
- Yu, K.R., Kang, K.S., 2013. Aging-related genes in mesenchymal stem cells: A mini-review. *Gerontology* 59, 557–563.
- Zwetsloot, K.A., Childs, T.E., Gilpin, L.T., Booth, F.W., 2013. Non-passaged muscle precursor cells from 32-month old rat skeletal muscle have delayed proliferation and differentiation. *Cell Proliferation* 46, 45–57.

Epigenetic Regulation of Stem Cells

A Sacco, Sanford-Burnham Medical Research Institute, La Jolla, CA, USA

PL Puri, Sanford-Burnham Medical Research Institute, La Jolla, CA, USA and IRCCS Fondazione Santa Lucia, Rome, Italy

© 2016 Elsevier Inc. All rights reserved.

Introduction: Regulation of Gene Expression in Stem Cells

All cell types that compose tissues and organs of adult multicellular organisms arise early in embryonic development as a result of lineage commitment and functional specialization of embryonic stem cell (ESC)-derived cellular populations present in the three initial germ layers – ectoderm, endoderm, and mesoderm (Murry and Keller, 2008). ESCs are pluripotent, self-renewing cells derived from the inner cell mass (ICM) of the developing blastocyst (Rossant, 2008). Pluripotency is the capacity of a single cell to generate all cell lineages of the developing and adult organism. Self-renewal is the ability of a cell to proliferate in the same state. The elucidation of the mechanisms by which ESCs self-renew while retaining their pluripotency has been a major challenge in the last decade and led to the discovery of the regulation of gene expression in ESCs by functionally relevant changes to the genome that do not involve a change in the nucleotide sequence – collectively referred to as epigenetic regulation of gene expression.

The complete mapping of the human genome and the advent of sophisticated technologies for chromatin analysis coupled with high-throughput DNA sequencing has made unprecedented genome-wide studies possible. These efforts have elucidated the dynamic changes of the epigenetic landscape that regulate gene expression. These studies have revealed the epigenetic networks that confer upon ESCs their unique property to undergo indefinite self-renewing while remaining poised to adopt a complete range of lineages and subsequent differentiation into specialized cell types in response to developmental cues. The progressive lineage restriction of human ESCs (hESCs) in response to developmental cues is achieved through a complex interplay between epigenetic and genetic factors, which establish unique gene expression profiles that underlie cell type-specific biological activities. A key aspect of this interplay relates to the simultaneous activation and repression of distinct clusters of genes, which ultimately direct lineage specification and functional specialization of ESC progeny (Orkin *et al.*, 2008). Execution of this program is coordinated by the combinatorial activities of transcriptional activators, cofactors, and the recently discovered noncoding RNAs, which generate discrete epigenetic signatures, consisting of specific patterns of histone modifications, changes in chromatin structure, and DNA methylation (Bibikova *et al.*, 2008). These activities are superimposed upon and tightly interconnected with the genome and are globally referred to as the ‘epigenome.’

The knowledge generated from ESC biology has been recently extended to induced-pluripotent stem cells (iPSCs), which are derived from the reprogramming of somatic cells into pluripotent cells that resemble ESCs and currently hold great promise in the field of regenerative medicine

(Wu and Hochedlinger, 2011; Jaenisch and Young, 2008). Moreover, concepts and principles generated from these studies have been extended to adult stem cells, whose intrinsic pluripotency is restricted to a limited number of cell types dictated from the germ layer derivation and organs and tissues in which they reside (Wabik and Jones, 2015).

Effectors and Mechanisms of Epigenetic Regulation of Gene Expression in ESCs

The interplay between distinct complexes endowed with specific enzymatic activities is required to set the unique epigenetic landscape that directs the gene expression in every cell type, including ESCs. This interplay impacts on few basic regulatory mechanisms that include histone modifications, changes in chromatin structure, generation of noncoding ncRNA, and DNA methylation (Young, 2011). When considered at the genome-wide level, these events can produce such a variety of ‘epigenetic signatures’ that regulate every cell type-specific profiles of gene expression through developmental stages.

The seminal discovery of mammalian nuclear transcriptional ‘coregulators’ that are able to enhance the activity of transcription factors, by virtue of their enzymatic activity (Yang *et al.*, 1996), provided a landmark framework that paved the way to the identification of additional enzymatic activities that regulate gene expression at the epigenetic level by promoting posttranslational modifications of histones. In particular, the histone tails that protrude from the fundamental unit of chromatin – the nucleosome (see below for more detailed description) – undergo extensive posttranslational modifications (Tessarz and Kouzarides, 2014), including lysine acetylation, methylation, arginine methylation, and serine/threonine phosphorylation. Additional histone modifications include SUMOylation and ubiquitination. Enzymes that catalyze the posttranslational modifications of histones have been referred to as ‘writers’ of a ‘code’ that indicates whether gene expression is activated or repressed. The activity of histone ‘writers’ is countered by enzymes endowed with an opposite function – called ‘erasers’ – that remove histone modifications. The dynamic combination of different histone modifications provides a code that is read by several enzymatic complexes – the ‘readers’ which include the ‘writers’ themselves and the chromatin remodelers (Wang *et al.*, 2004) – see below for further details.

Enhancer Activation and Developmental Competence

One of the first ‘writers’ to be discovered were the histone acetyltransferases p300/CBP and PCAF, which have been originally found to promote transcription by virtue of their intrinsic acetyltransferase (HAT) activity and ability to interact

with transcription factors (Yang *et al.*, 1996; Ogryzko *et al.*, 1996; Bannister and Kouzarides, 1996). In particular, p300/CBP-mediated lysine acetylation of histones establishes cell type-specific patterns of localized 'hyperacetylation' across the genome in response to external cues. Following the discovery of HATs, a growing number of proteins belonging to multi-protein complexes and endowed with specific enzymatic activities, have been discovered and characterized. The availability of chromatin-immunoprecipitation (ChIP) and ChIP-related techniques coupled to high-throughput next-generation sequencing (ngs) has allowed genome-wide mapping of key epigenetic changes that are generated by the activity of these complexes and that regulate gene expression in ESCs, as well as in somatic cells. The growing number of ngs data have been collected and analyzed within international consortium networks, such as ENCODE, that provide a dynamic, updated illustration of the epigenome in human cells. This framework has permitted the identification 'epigenetic signatures' of functional domains of the genome of hESCs that are summarized below.

In general, histone hyperacetylation marks domains of the genome that are permissive for transcription, and *vice versa* hypoacetylation correlates with transcription silencing. However, acetylation of one particular lysine (K27) of histone H3 (H3K27Ac) by p300 has been instrumental for the identification of active enhancers (Visel *et al.*, 2009; Creyghton *et al.*, 2010). Enhancers define short (about 50–1500 base pairs) elements of the genome that are targeted by sequence-specific transcriptional activators to activate transcription of genes located up to 1 Mbp (1 000 000 base pairs) upstream or downstream from the start site of the target gene. There are hundreds of thousands of enhancers in the human genome and their combinatorial activation gives rise to unique collections of enhancers that support cell type-specific patterns of gene expression (Andersson *et al.*, 2014). Indeed, enhancers mediate both spatial and temporal control of development by inducing transcription in certain cells types, while repressing it in other cells – a process also known as developmental competence (Wang *et al.*, 2015). Thus, it is believed that activation of discrete subsets of enhancers is highly predictive of biologically different transcriptional outputs in ESCs. Inactive enhancers are typically marked by H3K4monomethylation (H3K4me1), a modification that also persists after activation and is therefore used as a marker of enhancer identification. In response to developmental cues, binding of transcription factors to specific sequences within certain enhancers induces H3K27Ac, by recruitment of p300 and possibly other HATs. These transcription factors are defined as pioneers, since they preset the chromatin at previously silent enhancers for subsequent activation by other transcriptional activators. In self-renewing ESCs, the cooperative interactions among the ESC-specific transcription factors Oct4, Sox2, and Nanog within specific enhancers maintain the expression of genes that promote stem cell properties, as long as ESCs are exposed to conditions that maintain the stem cells in an undifferentiated state. By contrast, once exposed to developmental cues, enhancers binding by lineage-specific transcription factors, and simultaneous depletion of Oct4, Sox2, and Nanog from stem cell-specific enhancers, activate gene expression toward differentiation into specific lineages (Xie and Ren, 2013).

Promoter Bivalency and Maintenance of Pluripotency

While enhancer activation is a key mechanism that determines whether ESC proliferate as undifferentiated cells or commit toward a differentiated cell type, distinct posttranslational histone modifications at promoters are important to establish the chromatin structure that permits ESCs to maintain pluripotency. Promoters define DNA elements that initiate transcription of a particular gene, whose transcription start-sites are located immediately downstream. The two regulatory histone marks that influence chromatin structures at promoters in ESC are histone 3 lysine4 tri-methylation (H3K4me) and histone 3 lysine27 tri-methylation (H3K27me3). They are catalyzed by distinct types of methyltransferase-containing complexes and have a different impact on chromatin structure and gene transcription (Ringrose and Paro, 2004; Schuetten-gruber *et al.*, 2007). H3K27me3 is typically catalyzed by the enzymatic subunit of the Polycomb Group (PcG) – Enhancer of Zeste (EzH2). H3K4me3 is catalyzed by a battery of histone methyltransferases (HMTs) that show functional and structural analogies to the enzymatic complex Trithorax group (TrxG), originally identified in *Drosophila melanogaster*. In adult somatic cells, H3K27me3 at promoters silences the transcriptional activity of downstream genes, while promoters of transcriptionally active genes are typically enriched in H3K4me3. Thus, the antagonistic activities of PcG and TrxG define the promoter configuration that determines the activation/repression of specific subsets of genes during development that ultimately support the formation of various types of differentiated cells. In ESCs, the simultaneous presence of H3K4me3 and h3K27me3 has been observed at most of the developmental genes. This is considered a unique signature of pluripotency, and it is also referred to as 'bivalency' (Azuara *et al.*, 2006; Bernstein *et al.*, 2006; Boyer *et al.*, 2006). Bivalent promoters are considered 'poised' for either activation (by loss of H3K27me3) or repression (by loss of H3K4me3) in response to developmental cues. As such, promoter bivalency is considered a key epigenetic mechanism by which ESC can proliferate while retaining their pluripotency. Of note, bivalent promoters are typically occupied by 'paused' Polymerase II, which is bound at the TSS of the downstream gene, but is not able to initiate transcription. Resolution of promoter bivalence by developmental signals either promotes predominant H3K4me3 and Pol II elongation, to activate lineage-specific gene expression, or generate local enrichment of H3K27me3, which inhibits transcription, at least in part, by restraining poised RNA Pol II (Stock *et al.*, 2007; see Figure 1).

Chromatin Remodeling

An important mechanism by which histone modifications can change the structure of the chromatin in enhancers and promoters is the chromatin remodeling, which in turn alters the nucleosome architecture and frequency across the genome. Nucleosomes are the basic units of chromatin and consist of 146 base pairs of DNA wrapped around a core octamer of two copies each of four different types of histone – H2A, H2B, H3, and H4 (Kornberg and Thomas, 1974). Nucleosomes are distributed throughout the genome as repeating arrays spaced by linker DNA and histones. The DNA packaging into

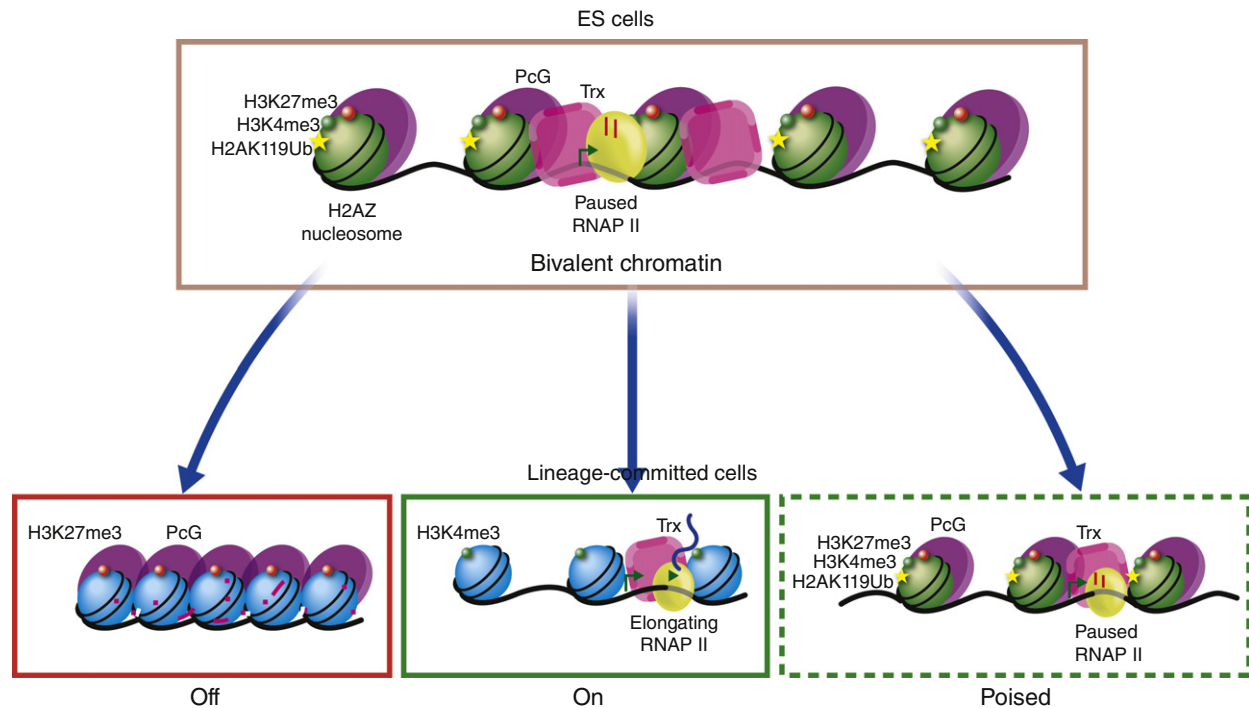


Figure 1 Schematic representation of the chromatin transitions at regulatory elements in stem cells. Top shows a typical 'bivalent' promoter in embryonic stem cells that can be resolved in three different transcriptional outcomes (bottom): silencing by repressive chromatin (left), activation by permissive chromatin (middle), and perpetuation of bivalency by poised chromatin (right).

nucleosomes provides a natural barrier against the access of DNA-binding protein, to allow a control to the binding of transcriptional activators to their DNA-binding motifs. Indeed, displacement or mobilization of nucleosomes is a key mechanism that controls the activation of certain subsets of genes by tissue-specific transcriptional activators during hESC differentiation. Likewise, dynamic redistribution of nucleosomes can repress gene expression by occluding other transcription factors access to DNA. Given that the large majority of the genome is covered by nucleosomes, an outstanding question relates to how transcriptional activators access their DNA-binding motifs. Recent work has revealed that a limited number of sequences of DNA are exposed within the nucleosome structure and these sequences are recognized by pioneer transcription factors, which appear to initiate the process of chromatin remodeling, followed by nucleosome displacement, exposure of full sequences of DNA available for binding of additional transcription factors that cooperate with other components of the transcriptional machinery to activate transcription (Zaret and Carroll, 2011). This sequence of events occurs in enhancers and promoters of ESC, and depending on the identity of the transcriptional activators that access to the DNA, can generate a transcriptional output that directs distinct lineages.

Chromatin remodeling occurs through an ATP-dependent activity of specialized multi-protein complexes (Clapier and Cairns, 2009). Among them SWI/SNF chromatin remodeling complexes are of particular importance (de la Serna *et al.*, 2006). These complexes are highly dynamic in their composition, which varies depending on the cell type and developmental stage. SWI/SNF complexes typically contain one

enzymatic subunit – either the ATPase Brahma (Brm) or Brg1 – and a collection of Brg1/Brm-associated factors (BAFs) (Ho and Crabtree, 2010). SWI/SNF recruitment to specific DNA elements is mediated by interactions with sequence-specific transcriptional activators and with certain histone modifications, such as histone acetylation (Clapier and Cairns, 2009). SWI/SNF-mediated chromatin remodeling at promoters or enhancers facilitates or reduces the access of transcriptional components to DNA sequences with resulting positive or negative effects on gene activity. An ESC-specific SWI/SNF complex (called esBAF) has been purified and shown to be associated with promoters of genes under the control of Oct4, Sox2, and Nanog. It appears that depending on the presence of specific core subunits, the SWI/SNF complex supports ESC maintenance or differentiation into specific cell types (Ho and Crabtree, 2010; Albini *et al.*, 2013). In response to developmental cues and generation of the three germ layers, different progenitor cells from ectoderm, endoderm, and mesoderm show distinct SWI/SNF complexes with a specific combination of components that imparts cell type-specific patterns of chromatin remodeling, in concert with tissue-specific transcriptional activators. Importantly, it appears that in undifferentiated ESCs SWI/SNF-mediated chromatin remodeling is implicated more in establishing repressive chromatin to prevent differentiation; by contrast, during ESC differentiation SWI/SNF activity is required to remodel the chromatin at active loci (Saladi and de la Serna, 2010). Given the heterogeneity in SWI/SNF composition and dynamic exchange of subunits, it is likely that the ability of repressing or activating transcription depends on the presence of specific subunits that allow proper response to extrinsic signals.

DNA Methylation

Another important layer of epigenetic regulation of transcription is provided by DNA methylation, which cooperates with posttranslational modifications of histones and changes in chromatin structure to regulate genes expression in ESCs. Mechanistically, DNA methylation contributes to establish and maintain repressive domains through the genome. DNA methylation is essential for mammalian development and is catalyzed by DNA methyltransferases, the large majority of which are expressed in ESCs.

In ESCs, DNA demethylation triggers a sequence of events that ultimately resolve the equilibrium between the totipotency of the zygote and the pluripotency of the embryo. DNA demethylation in the zygote is required to erase existing epigenetic signatures inherited from the gametes, but at the same time resets the epigenetic identity toward the acquisition of lineage commitment upon developmental cues. Although a large part of the genome (60–80% of all CpG dinucleotides) is methylated in ESCs (Meissner, 2010), ESCs can be established and maintained in the absence of DNA methyltransferases and DNA methylation. Still, DNA methyltransferase-deficient ESCs show an impaired differentiation potential (Jackson *et al.*, 2004), possibly due to the incomplete silencing of Oct4 and Nanog during differentiation (Feldman *et al.*, 2006). Indeed, ESCs show an increased global methylation during cellular differentiation, with cell type-specific patterns of methylation. The gain of methylation observed during ESC differentiation mediates silencing of pluripotency genes, but is also required to prevent the inappropriate expression of genes of alternative lineage. As such, DNA methylation contributes to narrow the repertoire of lineage-specific gene expression that specifies the cellular identity of ESC-derived progeny (Nazor *et al.*, 2012). Likewise, demethylation at pluripotency genes is required for induced pluripotency in somatic cells. Interestingly, methylation/demethylation events cooperate with histone modifications to establish promoter bivalency, but also occur outside promoter sequences, as dynamic changes in DNA methylation have been observed at distal gene sequences and intronic regions (Meissner *et al.*, 2008).

Noncoding RNA

The contribution of noncoding RNA (ncRNA) to development and ESC differentiation has been established by several recent studies, which have revealed the interplay between ncRNA and epigenetic factors in modulating the transcriptional outputs of ESCs. Initial hints derived from showing the cooperation between PcG and ncRNA studies on X inactivation (Lee, 2009). Indeed, a long list of ncRNA species (including microRNA and long-noncoding RNA – lincRNA) has been implicated in the control of ESC state and differentiation.

MicroRNAs are short RNA molecules that bind to complementary sequences on messenger RNA and block expression of a gene (Bartel, 2009). miRNAs contribute to both maintenance of pluripotency in ESC and differentiation toward different lineages. The first evidence of a role of miRNA in ESCs was provided by the observation that deficiency in the miRNA-processing enzymes Dicer and DCGR8 resulted in defects in differentiation and proliferation (Kanellopoulou *et al.*, 2005;

Murchison *et al.*, 2005; Wang *et al.*, 2007). Further studies revealed that the core regulators Oct4/Sox2/Nanog promote the generation of miRNAs essential for maintenance pluripotency during ESC proliferation while repressing lineage-specific miRNA (Marson *et al.*, 2008). Among the miRNAs that control ESC identity, pluripotency and differentiation potential, the mir-290–295 cluster appears of special importance (He *et al.*, 2005; Kanellopoulou *et al.*, 2005; Murchison *et al.*, 2005; Wang *et al.*, 2007), as forced expression of ESC-specific microRNA (miR-291, miR-294, and miR-295) in somatic cells enhances the efficiency of induced pluripotency.

Likewise, a broad spectrum of lincRNAs of various lengths that are transcribed bidirectionally by RNA polymerase II from different elements of the genome, including enhancers and promoters, has shown to contribute to the epigenetic regulation of ESCs (Zaratiegui *et al.*, 2007; Hu *et al.*, 2012). In particular, an emerging theme regards the ability of these ncRNAs to mediate PcG-dependent gene silencing (Zhao *et al.*, 2008). PcG complexes show an affinity for lincRNA of about 200–1200 nucleotides that are often transcribed from genes occupied by PcG complexes and that facilitate or stabilize the recruitment and the activity of PcG complexes at specific loci in the genome of ESCs (Kaneko *et al.*, 2014).

Overall, the field of ncRNA-mediated regulation of the epigenetic landscape and transcription in ESC is rapidly evolving and the functional connection between RNA biology and epigenetics will be further elucidated by future studies.

Epigenetic Signatures and Control of Nuclear Architecture in ESCs

In ESCs, the pluripotent state is established and maintained by the expression of the core transcription factors Oct4, Sox2, and Nanog. These core transcription factors positively regulate their own promoters, thereby providing a feed-forward loop. Moreover, they co-occupy the regulatory regions of genes that maintain the ESC state, while repressing the expression of lineage-specific genes. Thus, the expression levels of ESC core transcription factors determine whether ESC proliferate in their undifferentiated state or differentiate into specific cell types (Boyer *et al.*, 2005). This regulatory circuit likely explains why forced, ectopic expression of these factors in somatic cells can reprogram the genome toward pluripotency – the key concept that inspired the seminal discovery of induced-pluripotent cells by Yamanaka and colleagues (Takahashi and Yamanaka, 2006; Jaenisch and Young, 2008). Additional factors, such as c-Myc, can also facilitate activation of this program by stimulating gene expression. Indeed, while Oct4, Sox2, and Nanog are involved in RNA polymerase II recruitment, c-Myc stimulates the release of Pol II pause to activate genes implicated in ESC proliferation and self-renewal (Cartwright *et al.*, 2005). An essential mechanism that maintains ESCs in their undifferentiated state is provided by the presence of factors that inhibit the signals that stimulate differentiation (Pera and Tam, 2010). Among these inhibitory factors, LIF, Wnt, and ligands of the TGF- β /BMP signaling pathway appear particularly important as they activate transcription factors, such as Stat3, Tcf3, and Smad1 that co-occupy enhancers bound by Oct4, Sox2, and Nanog.

In contrast, upon removal of these inhibitory factors, differentiation cues extinguish the network of ESC core transcription factors and promote the expression of tissue-specific transcriptional activators, which direct the transcriptional output toward the generation of ESC-derived lineages within the three germ layers, ultimately promoting the formation of embryoid bodies.

How can ESCs integrate such a variety of stimuli into specific responses (e.g., stemness or differentiation) and how can they be instructed to maintain the transcriptional profiles evoked by developmental cues? The answer probably resides in the plasticity of the nuclear architecture, which can adapt a vast combination of long-range interactions between functional domains of the genome. The recent availability of techniques based upon chromosome conformation capture or 3C, has permitted high-throughput analyses of the organization of chromosomes, and has revealed functionally relevant interactions between functional elements of the genome. These long-range interactions are mediated by a class of transcriptional coregulators, called mediators, and by other specialized proteins, such as cohesin and CTCF and lead to the formation of topological domains (TDs). Recent studies have revealed that TDs often consist of loops by which enhancers and promoters establish physical contacts and share common epigenetic regulators (Jin *et al.*, 2013). This explains why p300 occupies the most active enhancers and promoters in ESCs (Chen *et al.*, 2008b) and suggests a mechanism by which developmentally competent enhancers resolve promoter bivalency and activate the expression of lineage-specific genes during ESC differentiation. These studies have been extended also to ESCs, revealing the importance of TAD formation during lineage commitment (Dixon *et al.*, 2015). Likewise, Oct4/Sox2/Nanog-bound enhancers can interact with promoters of active genes in the core regulatory circuitry of undifferentiated ESCs (Kagey *et al.*, 2010). Of note, core pluripotency or tissue-specific transcription factors, cofactors and mediators often co-occupy enhancers with signaling transcription factors (such as TGF β -activated SMADs), thereby integrating developmental cues with epigenetic regulators and the transcriptional machinery (Mullen *et al.*, 2011). It has been recently proposed that the formation of super-enhancers (e.g., enhancers enriched in specific transcriptional activators and mediators) drives the nuclear architecture and transcriptional outputs. Finally, an additional layer of tridimensional regulation of the nuclear landscape is provided by the nuclear position of TDs, with nuclear periphery being generally associated with transcriptional repression.

Overall, the ESC transition from pluripotency to lineage commitment is underlined by an extensive reconfiguration of the epigenetic landscape, in which coding and noncoding elements of the genome cooperate to establish epigenetic signatures leading to regulatory structures that coordinate the expression of different subsets of genes in response to developmental cues.

Epigenetic Control of Quiescence/Stemness and Lineage Determination in Adult

Stem cell compartments in adult mammalian organisms are required for the homeostatic maintenance and repair of mature tissues. Adult stem cells are endowed with the ability to

generate cells of the tissue in which they reside, and self-renew (i.e., generate more copies of themselves in order to maintain the stem cell compartment). They are distinct from ESCs as they are not pluripotent (i.e., not able to generate all the tissues of the body or contribute to all the tissues upon implantation into the blastocyst).

Adult stem cell function is influenced by both intrinsic and extrinsic factors. The 'stem cell niche,' originally defined by Schofield (1978), is the immediate microenvironment to which stem cells are exposed to and it is key in supporting and instructing stem cell behavior through secreted factors, extracellular matrix (ECM) components, and cell-cell communications. For example, polarity within the niche, i.e., different regions of a stem cell are exposed to different signals, contributes to determining stem cell division mode, asymmetric or symmetric. Integration of microenvironmental signals leads to changes in the epigenome of a cell, resulting in a different transcriptional output. While in high turnover tissues, such as blood and epithelia, they actively generate progeny to maintain tissues, in low turnover ones such as skeletal muscle and nervous system they only rarely proliferate. In spite of these differences, there is significant conservation across different adult stem cell compartments.

One key feature is 'cellular quiescence.' Most adult stem cells at any given time exist in a quiescent state, defined as a reversible nonproliferative state from which stem cells can emerge in response to stress or tissue injury, start dividing and self-renew as well as generate committed progeny to contribute to tissue repair. Significant effort in the last decades has investigated the epigenetic landscape associated with these distinct states, quiescence, activation, and differentiation, and has provided critical insights into how the transitions between these states are dynamically regulated.

Skeletal Muscle Stem Cells

Satellite cells (SCs) are tissue-resident muscle stem cells required for postnatal growth and regeneration of skeletal muscle (Brack and Rando, 2012). They reside in their anatomical niche, within the space between the myofiber membrane and the surrounding basal lamina (Mauro, 1961). Upon injury or stress they undergo activation, enter the cell cycle and divide to generate both committed progenitors that directly contribute to the repair of injured muscles and copies of themselves that replenish the pool of reserve SC. SC can be isolated from tissues by enzymatic digestion and fluorescence-activated cell sorting (FACS) based on cell surface markers.

Recent studies have compared global epigenetic marks in quiescent and activated SC, and shown that quiescent SC exhibit a high content of the bivalent domains, coexistence of both activating and repressive marks on the same gene (i.e., coexistence of H3K27me3 and H3K4me3), typically observed in the chromatin of ESCs and indicative of the stem cell identity (Liu, *et al.*, 2013). This profile is thought to indicate poised genes ready to be activated, suggesting a high degree of plasticity of this cellular state. Interestingly, these signatures are progressively lost upon SC activation, indicating a progressive restriction in potency upon activation of the myogenic program and terminal differentiation. Work from the Rando's

group identified miRNAs as major regulators of SC quiescence. Indeed, by conditional genetic deletion of Dicer, the RNA processing enzyme responsible for generating mature miRNA by cleavage of per-miRNA, they show that SC lose their ability to maintain quiescence. They further identify miR489 as a miRNA required for maintenance of SC quiescence, through the regulation of the oncogene DEK (Cheung *et al.*, 2012). Recent studies from Sartorelli's lab have revealed important metabolic features that distinguish quiescent MuSC from the activated, committed progeny and are conferred by the activity of a key epigenetic regulator – the histone deacetylase SIRT1 (Ryall *et al.*, 2015). These studies showed that quiescent MuSC mostly rely on fatty acid oxidation; however, upon activation, proliferating MuSC experience a metabolic switch to glycolysis – a 'metabolic reprogramming' that causes a decrease in intracellular nicotinamide adenine dinucleotide (NAD) (+) and consequent inhibition of SIRT1, ultimately leading to elevated H4K16 acetylation and activation of muscle gene transcription.

Upon stress or injury, several signals emanating from the microenvironment, including inflammatory, endothelial and mesenchymal cells, contribute to the activation of SC and their transition to the progenitor stage. For example, Tumor necrosis factor (TNF) activates p38alpha, which simultaneously signals to PgC to repress Pax7 and to the SWI/SNF complex to promote myogenic differentiation (Palacios *et al.*, 2010; Simone *et al.*, 2004). Conditional deletion of the enzymatic subunit of PRC2, Ezh2, induces a reduction in SC numbers, reduced muscle mass and defects in tissue repair, indicating that Ezh2 regulates SC self-renewal and proliferation (Juan *et al.*, 2011). Finally, an increasing level of complexity in the epigenetic regulation of SC is emerging from the recent identification and characterization of lncRNA, which promote chromatin modifications that regulate the activation of the myogenic program (Cesana *et al.*, 2011; Lu *et al.*, 2013; Anderson *et al.*, 2015; Mousavi *et al.*, 2013).

Hematopoietic Stem Cells

Hematopoietic stem cells (HSC) reside in the bone marrow and are responsible for the generation of all the lineages of the blood, including erythrocytes, platelets, myelocytes, and lymphocytes, throughout the life of an organism. HSC immediate microenvironment, the bone marrow niche, is composed by multiple cell types, including endothelial, osteoblastic, and stromal cells that emanate signals that play an important role in determining HSC behavior *in vivo*. In mice, HSC are identified and isolated by cell surface markers as Lineage-Sca1 + ckit + or CD150 + CD48- cells (Oguro *et al.*, 2013).

Epigenetic analysis revealed that disruption of DNA methylation through manipulation of DNA methyltransferases impaired HSC function. For example, deletion of Dnmt1, required for the maintenance of DNA methylation, induces premature differentiation, suggesting that Dnmt1 acts to preserve stemness and prevent premature differentiation in the stem cell compartment (Broske *et al.*, 2009; Trowbridge *et al.*, 2009). Conversely, conditional deletion of Dnmt3a, responsible for de novo DNA methylation, resulted in increased HSC numbers and impaired differentiation (Challen

et al., 2012). This is thought to take place through the requirement to repress self-renewal genes before terminal differentiation. HSC also possess high levels of bivalent histone marks, which is progressively reduced during differentiation (Attema *et al.*, 2007). Conditional deletion or mutations of Bmi1, a component of the Polycomb Repressive Complex PRC1, leads to decreased HSC self-renewal, while its over-expression induces self-renewal and symmetric expansion of HSC (Park *et al.*, 2003). While deletion of Ezh1, component of PRC2 complex, impairs HSC self-renewal and prevents cellular senescence (Hidalgo *et al.*, 2012), mutations of PRC2 components Ezh2, Eed, and Suz12 enhanced HSC function, although they seem to act in a developmental stage-specific manner (Majewski *et al.*, 2008). Consistent with these findings, treatment of animal models with Trichostatin A (inhibitor of HDAC) or 5-azacytidine (inhibitor of Dnmt) enhanced HSC self-renewal, indicating the critical role of these epigenetic modifiers in the regulation of HSC function (Chung *et al.*, 2009).

Recent studies have also highlighted a critical role of miRNAs in regulating HSC function. For example, conditional deletion of the RNA processing enzyme Dicer resulted in decreased HSC number and increased rate of apoptosis (Guo *et al.*, 2010) and deletion of lncRNAs Xist and H19 affect HSC survival and quiescence, respectively (Yildirim *et al.*, 2013; Venkatraman *et al.*, 2013).

Neural Stem Cells

Neural stem cells (NSC) resides in the subgranular zone (SGZ) of the dentate gyrus region of the hippocampus, and also in the subventricular zone (SVZ), a layer of cells in the lateral zone of the lateral ventricle of the brain. The NSC niche comprises microglia, astrocytes, and vasculature. The nervous system is a low turnover tissue, and for long time it was believed that the brain was a completely postmitotic tissue devoid of stem or progenitor cells. NSC are classically isolated by enzymatic digestion of the brain followed by culture expansion. NSC express transcription factors Sox2 and Mash, intermediate filament gene Nestin and the astrocytic marker glial fibrillary acidic protein (GFAP) (Suh *et al.*, 2007; Song *et al.*, 2012). Dynamic DNA methylation has been associated with NSC differentiation. In fact, expression levels of DNMT1 and DNMT3a are transiently increased during early differentiation, and progressively reduced at later stages (Singh *et al.*, 2009). This could be associated to the need to silence self-renewal genes in order to promote the transition to the progenitor stage and later mature neurons. Consistent with these observations, Dnmt3a^{KO} mice exhibit impaired adult neurogenesis (Wu *et al.*, 2010). Similar pattern of expression is observed for the methyl-CpG-binding protein MBD1, which binds methylated gene promoters and induces transcriptional repression (Singh *et al.*, 2009), and knockout animal models also exhibit impaired adult neurogenesis (Zhao *et al.*, 2003), indicating the dynamic transition during early differentiation is a critical step for chromatin modification to activate the differentiation program, and at the same time select which cell fate NPC will adopt. For example, MBD1 represses genes involved in self-renewal, such as FGF2, enabling NSC

to proceed toward neurogenic differentiation. Conversely, MeCP2, another MBD protein, is involved in neuronal maturation, indicating stage-specific effects of different chromatin regulators (Smrt *et al.*, 2007). Changes in histone modifications and chromatin remodeling are also critical for proper adult neurogenesis. For example, treatment with HDACi sodium butyrate, valproic acid, or suberoylanilide hydroxamic acid (SAHA) promoted NSC lineage-specific neuronal differentiation at the expense of self-renewal (Hsieh *et al.*, 2004). Bmi1 is required for adult NSC self-renewal *in vivo* and its expression increases neurosphere formation *in vitro* (Fasano *et al.*, 2009). Mll1 (mixed-lineage leukemia 1), a member of Trithorax group (trxG), is required for neurogenesis in the mouse postnatal brain, as its genetic deletion severely impaired neurogenesis (Lim *et al.*, 2009). This supports a model in which Mll1 is required to resolve key silenced bivalent loci in postnatal neural precursors to the actively transcribed state for the induction of neurogenesis. Finally, several miRNAs have been implicated in adult neurogenesis, including, among others, miR9, miR34a, miR124, miR125b, miR132, and miR137 (Fitzsimons *et al.*, 2014).

Plasticity of Adult Cell Fate

Pioneering studies in the 1950s by John Gurdon highlighted for the first time that the differentiated state is not fixed or irreversible. Indeed, when a nucleus from a somatic cell is transplanted into an enucleated oocyte (somatic cell nuclear transfer), factors in the oocyte cytoplasm are sufficient to reset the clock back and confer the somatic cell nucleus with properties of ESCs (Gurdon, 1962a,b). These studies provided the first evidence of plasticity of cell fate and in the last decades they enabled the cloning of entire organisms, such as Dolly the sheep (Wilmut *et al.*, 1997). More recently, the groundbreaking findings from Yamanaka's group that by manipulation of a small number of transcription factors, OCT4, SOX2, KLF4, and c-MYC, it is possible to reprogram adult cells to an embryonic state shed light on key regulators of gene expression (Takahashi and Yamanaka, 2006). Epigenetic changes associated with nuclear reprogramming include telomere re-elongation, DNA demethylation and modifications in chromatin composition and structure. These findings paved the way to revitalizing stem cell biology and iPSCs are a powerful tool for disease in a dish model, drug screening and potential cell replacement therapies.

See also: Stem Cell Biology: Structure and Function: Pluripotent Cells: New Tools for Disease Research and Therapies

References

- Albini, S., Coutinho, P., Malecova, B., *et al.*, 2013. Epigenetic reprogramming of human embryonic stem cells into skeletal muscle cells and generation of contractile myospheres. *Cell Reports* 3, 661–670.
- Anderson, D.M., Anderson, K.M., Chang, C.L., *et al.*, 2015. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell* 160, 595–606.
- Andersson, R., Gebhard, C., Miguel-Escalada, I., *et al.*, 2014. An atlas of active enhancers across human cell types and tissues. *Nature* 507 (7493), 455–461.
- Attema, J.L., Papathanasiou, P., Forsberg, E.C., *et al.*, 2007. Epigenetic characterization of hematopoietic stem cell differentiation using miniChIP and bisulfite sequencing analysis. *Proceedings of the National Academy of Sciences of the United States of America* 104, 12371–12376.
- Azura, V., Perry, P., Sauer, S., *et al.*, 2006. Chromatin signatures of pluripotent cell lines. *Nature Cell Biology* 8, 532–538.
- Bannister, A.J., Kouzarides, T., 1996. The CBP co-activator is a histone acetyltransferase. *Nature* 384 (6610), 641–643.
- Bartel, D.P., 2009. MicroRNAs: Target recognition and regulatory functions. *Cell* 136, 215–233.
- Bernstein, B.E., Mikkelsen, T.S., Xie, X., *et al.*, 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315–326.
- Bibikova, M., Laurent, L.C., Ren, B., Loring, J.F., Fan, J.B., 2008. Unraveling epigenetic regulation in embryonic stem cells. *Cell Stem Cell* 2, 123–134.
- Boyer, L.A., Lee, T.I., Cole, M.F., *et al.*, 2005. Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122, 947–956.
- Boyer, L.A., Plath, K., Zeitlinger, J., *et al.*, 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441, 349–353.
- Brack, A.S., Rando, T.A., 2012. Tissue-specific stem cells: Lessons from the skeletal muscle satellite cell. *Cell Stem Cell* 10, 504–514.
- Broske, A.M., Vockentanz, L., Kharazi, S., *et al.*, 2009. DNA methylation protects hematopoietic stem cell multipotency from myeloid restriction. *Nature Genetics* 41, 1207–1215.
- Cartwright, P., McLean, C., Sheppard, A., *et al.*, 2005. LIF/STAT3 controls ES cell self-renewal and pluripotency by a Myc-dependent mechanism. *Development* 132, 885–896.
- Cesana, M., Cacchiarelli, D., Legnini, I., *et al.*, 2011. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* 147, 358–369.
- Challen, G.A., Sun, D., Jeong, M., *et al.*, 2012. Dnmt3a is essential for hematopoietic stem cell differentiation. *Nature Genetics* 44, 23–31.
- Chen, X., Xu, H., Yuan, P., *et al.*, 2008. Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133, 1106–1117.
- Cheung, T.H., Quach, N.L., Charville, G.W., *et al.*, 2012. Maintenance of muscle stem-cell quiescence by microRNA-489. *Nature* 482, 524–528.
- Chung, Y.S., Kim, H.J., Kim, T.M., *et al.*, 2009. Undifferentiated hematopoietic cells are characterized by a genome-wide undermethylation dip around the transcription start site and a hierarchical epigenetic plasticity. *Blood* 114, 4968–4978.
- Clapier, C.R., Cairns, B.R., 2009. The biology of chromatin remodeling complexes. *Annual Review of Biochemistry* 78, 273–304.
- Creyghton, M.P., Cheng, A.W., Welstead, G.G., *et al.*, 2010. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proceedings of the National Academy of Sciences of the United States of America* 107 (50), 21931–21936.
- Dixon, J.R., Jung, I., Selvaraj, S., *et al.*, 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature* 518 (7539), 331–336.
- Fasano, C.A., Phoenix, T.N., Kokovay, E., *et al.*, 2009. Bmi-1 cooperates with Foxg1 to maintain neural stem cell self-renewal in the forebrain. *Genes & Development* 23, 561–574.
- Feldman, N., Gerson, A., Fang, J., *et al.*, 2006. G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nature Cell Biology* 8, 188–194.
- Fitzsimons, C.P., van Bodegraven, E., Schouten, M., *et al.*, 2014. Epigenetic regulation of adult neural stem cells: Implications for Alzheimer's disease. *Molecular Neurodegeneration* 9, 25.
- Guo, S., Lu, J., Schlanger, R., *et al.*, 2010. MicroRNA miR-125a controls hematopoietic stem cell number. *Proceedings of the National Academy of Sciences of the United States of America* 107, 14229–14234.
- Gurdon, J.B., 1962a. Adult frogs derived from the nuclei of single somatic cells. *Developmental Biology* 4, 256–273.
- Gurdon, J.B., 1962b. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *Journal of Embryology and Experimental Morphology* 10, 622–640.
- He, L., Thomson, J.M., Hemann, M.T., *et al.*, 2005. A microRNA polycistron as a potential human oncogene. *Nature* 435, 828–833.
- Hidalgo, I., Herrera-Merchan, A., Ligos, J.M., *et al.*, 2012. Ezh1 is required for hematopoietic stem cell maintenance and prevents senescence-like cell cycle arrest. *Cell Stem Cell* 11, 649–662.

- Ho, L., Crabtree, G.R., 2010. Chromatin remodelling during development. *Nature* 463, 474–484.
- Hsieh, J., Nakashima, K., Kuwabara, T., Mejia, E., Gage, F.H., 2004. Histone deacetylase inhibition-mediated neuronal differentiation of multipotent adult neural progenitor cells. *Proceedings of the National Academy of Sciences of the United States of America* 101, 16659–16664.
- Hu, W., Alvarez-Dominguez, J.R., Lodish, H.F., 2012. Regulation of mammalian cell differentiation by long non-coding RNAs. *EMBO Reports* 13 (11), 971–983.
- Jackson, M., Krassowska, A., Gilbert, N., *et al.*, 2004. Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Molecular and Cellular Biology* 24, 8862–8871.
- Jaenisch, R., Young, R., 2008. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* 132 (4), 567–582.
- Jin, F., Li, Y., Dixon, J.R., *et al.*, 2013. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503 (7475), 290–294.
- Juan, A.H., Derfoul, A., Feng, X., *et al.*, 2011. Polycomb EZH2 controls self-renewal and safeguards the transcriptional identity of skeletal muscle stem cells. *Genes & Development* 25, 789–794.
- Kagey, M.H., Newman, J.J., Bilodeau, S., *et al.*, 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467, 430–435.
- Kaneko, S., Son, J., Bonasio, R., Shen, S.S., Reinberg, D., 2014. Nascent RNA interaction keeps PRC2 activity poised and in check. *Genes & Development* 28 (18), 1983–1988.
- Kanellopoulou, C., Muljo, S.A., Kung, A.L., *et al.*, 2005. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes & Development* 19, 489–501.
- Kornberg, R.D., Thomas, J.O., 1974. Chromatin structure; oligomers of the histones. *Science* 184, 865–868.
- Lee, J.T., 2009. Lessons from X-chromosome inactivation: Long ncRNA as guides and tethers to the epigenome. *Genes & Development* 23 (16), 1831–1842.
- Lim, D.A., Huang, Y.C., Swigut, T., *et al.*, 2009. Chromatin remodelling factor Mll1 is essential for neurogenesis from postnatal neural stem cells. *Nature* 458, 529–533.
- Liu, L., Cheung, T.H., Charville, G.W., *et al.*, 2013. Chromatin modifications as determinants of muscle stem cell quiescence and chronological aging. *Cell Reports* 4, 189–204.
- Lu, L., Sun, K., Chen, X., *et al.*, 2013. Genome-wide survey by ChIP-seq reveals YY1 regulation of lincRNAs in skeletal myogenesis. *EMBO Journal* 32, 2575–2588.
- Majewski, I.J., Blewitt, M.E., de Graaf, C.A., *et al.*, 2008. Polycomb repressive complex 2 (PRC2) restricts hematopoietic stem cell activity. *PLoS Biology* 6, e93.
- Marson, A., Levine, S.S., Cole, M.F., *et al.*, 2008. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 134, 521–533.
- Mauro, A., 1961. Satellite cell of skeletal muscle fibers. *Journal of Biophysical and Biochemical Cytology* 9, 493–495.
- Meissner, A., 2010. Epigenetic modifications in pluripotent and differentiated cells. *Nature Biotechnology* 28 (10), 1079–1088.
- Meissner, A., Mikkelsen, T.S., Gu, H., *et al.*, 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* 454, 766–770.
- Mousavi, K., Zare, H., Dell'orso, S., *et al.*, 2013. eRNAs promote transcription by establishing chromatin accessibility at defined genomic loci. *Molecular Cell* 51, 606–617.
- Mullen, A.C., Orlando, D.A., Newman, J.J., *et al.*, 2011. Master transcription factors determine cell-type-specific responses to TGF- β signaling. *Cell* 147 (3), 565–576.
- Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S., Hannon, G.J., 2005. Characterization of Dicer-deficient murine embryonic stem cells. *Proceedings of the National Academy of Sciences of the United States of America* 102, 12135–12140.
- Murry, C.E., Keller, G., 2008. Differentiation of embryonic stem cells to clinically relevant populations: Lessons from embryonic development. *Cell* 132 (4), 661–680.
- Nazor, K.L., Altun, G., Lynch, C., *et al.*, 2012. Recurrent variations in DNA methylation in human pluripotent stem cell and their differentiated derivatives. *Cell Stem Cell* 10, 620–634.
- Ogryzko, V.V., Schiltz, R.L., Russanova, V., Howard, B.H., Nakatani, Y., 1996. The transcriptional coactivators p300 and CBP are histone acetyltransferases. *Cell* 87 (5), 953–959.
- Oguro, H., Ding, L., Morrison, S.J., 2013. SLAM family markers resolve functionally distinct subpopulations of hematopoietic stem cells and multipotent progenitors. *Cell Stem Cell* 13, 102–116.
- Orkin, S.H., Wang, J., Kim, J., *et al.*, 2008. The transcriptional network controlling pluripotency in ES cells. *Cold Spring Harbor Symposia on Quantitative Biology* 73, 195–202.
- Palacios, D., Mozzetta, C., Consalvi, S., *et al.*, 2010. TNF/p38 α /polycomb signaling to Pax7 locus in satellite cells links inflammation to the epigenetic control of muscle regeneration. *Cell Stem Cell* 7, 455–469.
- Park, I.K., Qian, D., Kiel, M., *et al.*, 2003. Bmi-1 is required for maintenance of adult self-renewing haematopoietic stem cells. *Nature* 423, 302–305.
- Pera, M.F., Tam, P.P., 2010. Extrinsic regulation of pluripotent stem cells. *Nature* 465, 713–720.
- Ringrose, L., Paro, R., 2004. Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annual Review of Genetics* 38, 413–443.
- Rossant, J., 2008. Stem cells and early lineage development. *Cell* 132, 527–531.
- Ryall, J.G., Dell'Orso, S., Derfoul, A., *et al.*, 2015. The NAD(+)–dependent SIRT1 deacetylase translates a metabolic switch into regulatory epigenetics in skeletal muscle stem cells. *Cell Stem Cell* 16, 171–183.
- Saladi, S.V., de la Serna, I.L., 2010. ATP dependent chromatin remodeling enzymes in embryonic stem cells. *Stem Cell Reviews* 6 (1), 62–73.
- Schofield, R., 1978. The relationship between the spleen colony-forming cell and the haemopoietic stem cell. *Blood Cells* 4, 7–25.
- Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B., Cavalli, G., 2007. Genome regulation by polycomb and trithorax proteins. *Cell* 128, 735–745.
- de la Serna, I.L., Ohkawa, Y., Imbalzano, A.N., 2006. Chromatin remodelling in mammalian differentiation: Lessons from ATP-dependent remodellers. *Nature Reviews Genetics* 7 (6), 461–473.
- Simone, C., Forcales, S.V., Hill, D.A., *et al.*, 2004. p38 pathway targets SWI-SNF chromatin-remodeling complex to muscle-specific loci. *Nature Genetics* 36, 738–743.
- Singh, R.P., Shiue, K., Schomberg, D., Zhou, F.C., 2009. Cellular epigenetic modifications of neural stem cell differentiation. *Cell Transplant* 18, 1197–1211.
- Smrt, R.D., Eaves-Egenes, J., Barkho, B.Z., *et al.*, 2007. Mecp2 deficiency leads to delayed maturation and altered gene expression in hippocampal neurons. *Neurobiology of Disease* 27, 77–89.
- Song, J., Zhong, C., Bonaguidi, M.A., *et al.*, 2012. Neuronal circuitry mechanism regulating adult quiescent neural stem-cell fate decision. *Nature* 489, 150–154.
- Stock, J.K., Giadrossi, S., Casanova, M., *et al.*, 2007. Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nature Cell Biology* 9, 1428–1435.
- Suh, H., Consiglio, A., Ray, J., *et al.*, 2007. *In vivo* fate analysis reveals the multipotent and self-renewal capacities of Sox2+ neural stem cells in the adult hippocampus. *Cell Stem Cell* 1, 515–528.
- Takahashi, K., Yamanaka, S., 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676.
- Tessarz, P., Kouzarides, T., 2014. Histone core modifications regulating nucleosome structure and dynamics. *Nature Reviews Molecular Cell Biology* 15 (11), 703–708.
- Trowbridge, J.J., Snow, J.W., Kim, J., Orkin, S.H., 2009. DNA methyltransferase 1 is essential for and uniquely regulates hematopoietic stem and progenitor cells. *Cell Stem Cell* 5, 442–449.
- Venkatraman, A., He, X.C., Thorvaldsen, J.L., *et al.*, 2013. Maternal imprinting at the H19-Igf2 locus maintains adult haematopoietic stem cell quiescence. *Nature* 500, 345–349.
- Visel, A., Blow, M.J., Li, Z., *et al.*, 2009. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457 (7231), 854–858.
- Wabik, A., Jones, P.H., 2015. Switching roles: The functional plasticity of adult tissue stem cells. *EMBO Journal* 9, 1164–1179.
- Wang, A., Yue, F., Li, Y., *et al.*, 2015. Epigenetic priming of enhancers predicts developmental competence of hESC-derived endodermal lineage intermediates. *Cell Stem Cell* 16 (4), 386–399.
- Wang, Y., Fischle, W., Cheung, W., *et al.*, 2004. Beyond the double helix: Writing and reading the histone code. *Novartis Foundation Symposium* 259, 3–17.
- Wang, Y., Medvid, R., Melton, C., Jaenisch, R., Blueloch, R., 2007. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self renewal. *Nature Genetics* 39, 380–385.
- Wilmot, I., Schnieke, A.E., McWhir, J., Kind, A.J., Campbell, K.H., 1997. Viable offspring derived from fetal and adult mammalian cells. *Nature* 385, 810–813.
- Wu, H., Coskun, V., Tao, J., *et al.*, 2010. Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science* 329, 444–448.

- Wu, S.M., Hochedlinger, K., 2011. Harnessing the potential of induced pluripotent stem cells for regenerative medicine. *Nature Cell Biology* 5, 497–505.
- Xie, W., Ren, B., 2013. Developmental biology. Enhancing pluripotency and lineage specification. *Science* 341 (6143), 245–247.
- Yang, X.J., Ogryzko, V.V., Nishikawa, J., Howard, B.H., Nakatani, Y., 1996. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. *Nature* 382 (6589), 319–324.
- Yildirim, E., Kirby, J.E., Brown, D.E., *et al.*, 2013. Xist RNA is a potent suppressor of hematologic cancer in mice. *Cell* 152, 727–742.
- Young, R.A., 2011. Control of the embryonic stem cell state. *Cell* 144 (6), 940–954.
- Zaratiegui, M., Irvine, D.V., Martienssen, R.A., 2007. Noncoding RNAs and gene silencing. *Cell* 128, 763–776.
- Zaret, K.S., Carroll, J.S., 2011. Pioneer transcription factors: Establishing competence for gene expression. *Genes & Development* 25 (21), 2227–2241.
- Zhao, J., Sun, B.K., Erwin, J.A., Song, J.J., Lee, J.T., 2008. Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* 322 (5902), 750–756.
- Zhao, X.Y., Ueba, T., Christie, B.R., *et al.*, 2003. Mice lacking methyl-CpG binding protein 1 have deficits in adult neurogenesis and hippocampal function. *Proceedings of the National Academy of Sciences of the United States of America* 100, 6777–6782.

Relevant Website

<http://genome.ucsc.edu/ENCODE/index.html>
ENCODE.

The Adult Stem Cell Niche: Multiple Cellular Players in Tissue Homeostasis and Regeneration

S Kyryachenko, L Formicola, D Ollitrault, R Correra, A-L Denizot, K Kyrylkova, G Marazzi, and DA Sassoon, Université de Pierre et Marie Curie-Sorbonne Universités, Paris, France

© 2016 Elsevier Inc. All rights reserved.

The Skin Hair Follicle

The Hair Follicle Stem Cell Populations – A Series of Specialized Compartments with Varying Cell Fate Competences

Mammalian skin is capable of pronounced regeneration and undergoes turnover throughout the life of the organism. In mammals, the skin contains millions of hair follicles and harbors multiple populations of resident progenitor cells. The murine hair follicle has been particularly well studied. It undergoes repeated cycles of growth, regression, and rest throughout life. A pronounced regenerative capacity, coupled with its accessibility for experimental manipulation has led to significant advances in skin stem cell biology. In young mice, hair cycles are synchronized providing a tractable means to dissect the interactions between stem cells and their environment. Hair stem cells undergo repeated transitions between quiescence, activation, and subsequent acquisition of a differentiated cell fate coupled with stem cell self-renewal (Paus and Cotsarelis, 1999). Hair follicles contain distinct and spatially separated epithelial stem cells that possess unique molecular signatures that differentially contribute to the hair follicle during cycling and regeneration. Initial studies sought to identify stem cells by virtue of their predicted slow cell cycling, considered essential to maintain a reserve pool of progenitors. These initial investigations led to the identification of a reservoir of slow-cycling and label-retaining cells (LRCs) in a structure called the hair follicle bulge (Cotsarelis *et al.*, 1990). The bulge is located near the base of the hair follicle and remains intact throughout the hair cycle (Figure 1). Bulge cells express specific markers that have proven useful for their purification and dissection of different phases of bulge cell. For example, keratin 15 (K15) is expressed in undifferentiated and quiescent bulge keratinocytes and their transcriptional gene signature is similar to that of stem cells in other organs (Liu *et al.*, 2003; Morris *et al.*, 2004). CD34 is expressed by quiescent stem cells throughout the cycle, and purified CD34-positive bulge cells display clonogenic capacity (the ability to form a colony in a dish from a single cell) and self-renewal features of stem cells *in vitro* (Blanpain *et al.*, 2004). Bulge cells also express high levels of sex-determining region Y-box 9 (SOX9), which not only marks these cells but is functionally required for their maintenance (Kadaja *et al.*, 2014). The more proliferative cells in the lower portion of the bulge express the Leu-rich repeat-containing G protein-coupled receptor 5 (LGR5) and LIM homeobox protein 2 (LHX2) (Barker *et al.*, 2008). While this list of markers that identify bulge cells is not complete, what emerges is that these cells occupy a discreet and anatomically recognizable structure but retain a level of heterogeneity within that structure reflecting cell cycle and stem cell competence. When these

cells are purified by use of these markers and then engrafted into the skin of host mice, they are capable of reestablishing the hair follicle along with associated structures (Blanpain *et al.*, 2004). By use of genetic lineage tracing models, it has been shown that the bulge cells give rise to all the hair follicle structures and most of the associated structures including the bulge (Tumbar *et al.*, 2004). While the bulge cells are a clear example of an 'ideal stem cell' that gives rise to new hair follicles as well as self-renew, a significant amount of work has shown that the bulge cells are highly responsive to cells that lie in close vicinity and while these cells are derived from the bulge, they acquire additional stem cell potentials. Just below the bulge is the hair germ (Figure 1). The cells in the hair germ are derived from the bulge stem cells and while they have a similar molecular signature (i.e., they express K15, SOX9, LHX2, and LGR5), they are more proliferative (Greco *et al.*, 2009; Tornqvist *et al.*, 2010). Hair germ cells exhibit high

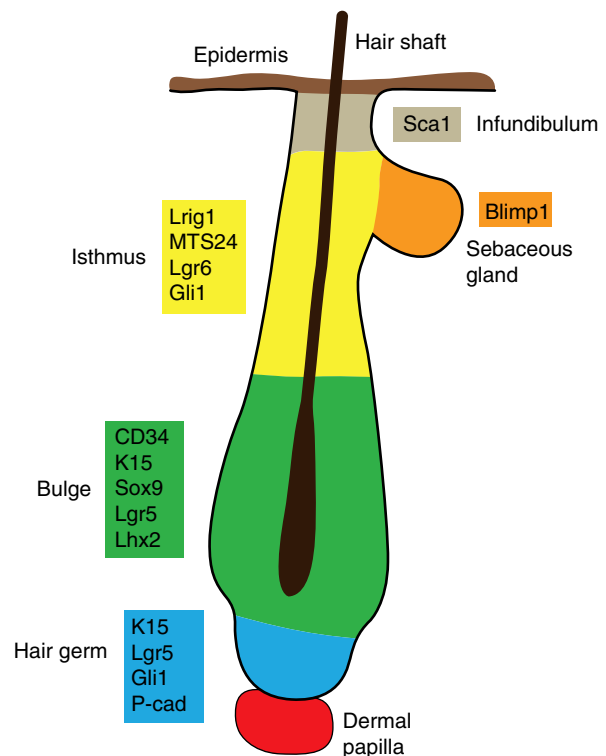


Figure 1 Generalized structure of the hair follicle during the resting phase of the hair cycle. The hair follicle consists of delineated anatomical compartments: the hair germ, bulge, isthmus, and infundibulum. The dermal papilla lies directly under the hair germ; sebaceous gland is attached to the side of the hair follicle. Each compartment contains different stem cells that are characterized by unique molecular signatures and different cell fate potentials.

levels of P-cadherin and GLI-Kruppel family member GLI1 (GLI1) that permit them to be distinguished from bulge cells (Greco *et al.*, 2009). Germ cells display remarkable multipotency and contribute to the entire hair follicle including bulge stem cells (Ito *et al.*, 2004). However, the hair germ cells are less clonogenic as compared to bulge stem cells *in vitro* (Greco *et al.*, 2009). The cells in the hair germ are the first to express genes that indicate stem cell activation and the first to proliferate at the beginning of a new hair regeneration cycle (Rompolas *et al.*, 2012). These observations suggest that a specific anatomical and lineage ordering exists in the hair follicle in which the bulges give rise to the hair germ, which in turn provides the bulk of the other skin structures consistent with the 'label-retaining cell' behavior of the bulge.

In addition to the bulge and germ cells, other subsets of specialized progenitors exist in defined anatomical compartments located above the bulge. The isthmus is situated between the bulge and the sebaceous gland (Figure 1). The cells in the isthmus differ from the bulge cells, as they do not express CD34 or K15 (Jensen *et al.*, 2008). The isthmus cells located close to the bulge display self-renewal properties, can regenerate the hair follicle, and may serve as a reserve stem cell population as well as contribute to wound healing (Brownell *et al.*, 2011). The central part of the isthmus can be identified on the basis of LGR6 and MTS24 expression and have been shown to contribute to the isthmus zone, epidermis, and sebaceous gland during homeostasis and mobilize to the epidermis following injury (Jensen *et al.*, 2008; Snippert *et al.*, 2010). The uppermost part of the isthmus contains cells that express MTS24 and while these cells display a gene expression profile inherent to progenitor or stem cells as well as proliferative and clonogenic capacity, their function remains unclear (Nijhof *et al.*, 2006). The sebaceous gland resides in the upper portion of the hair follicle (Figure 1) and serves an important role in lubricating and waterproofing the epidermis. B-lymphocyte-induced maturation protein 1 (BLIMP1)-positive cells near the gland opening can generate the sebaceous gland lineages (Horsley *et al.*, 2006). However, BLIMP1 is also widely expressed in differentiated epidermal and sebaceous gland cells and is not considered as a specific marker of sebaceous stem cells (Magnusdotir *et al.*, 2007). Leucine-rich repeats and immunoglobulin-like domain 1 (LRIG1) expression also marks sebaceous gland cells (Page *et al.*, 2013). Additional studies will be required to unveil whether cells from the bulge and isthmus contribute to the sebaceous gland or if it is maintained independently. The upper segment of the hair follicle that extends from the opening of the sebaceous gland duct to the epidermis is called infundibulum (Figure 1). The most upper part of this unit occupied by the cells that express stem cell antigen 1 (SCA-1). These cells have a unipotent capacity to regenerate interfollicular epidermis, and are thought to be keratinocyte progenitors (Jensen *et al.*, 2008).

Signaling between Follicle Stem Cells and Their Microenvironment

Bulge stem cells retain their lineage identity and progenitor potential to regenerate hair follicles after passaging *in vitro* and subsequent engraftment suggesting that either systemic or

local environmental clues are not required to maintain their proper function (Blanpain *et al.*, 2004; Toyoshima *et al.*, 2012). While the bulge cells can form a multitude of cell types as outlined above, they respond to multiple resident cell types including specialized dermal fibroblasts, adipocytes, muscle cells, and neurons (Plikus *et al.*, 2008; Brownell *et al.*, 2011; Fujiwara *et al.*, 2011). One key source of signals to the bulge is the dermal papilla, which is a mesenchymal niche that consists of a condensed group of dermal fibroblasts located at the base of a resting hair follicle (Driskell *et al.*, 2011). The dermal papilla cells are characterized by the expression of multiple markers including TBX18, Corin, SOX2, and CD133 (Enshell-Seijffers *et al.*, 2008; Driskell *et al.*, 2009; Grisanti *et al.*, 2013). The dermal papilla remains in close anatomical association with the follicle throughout the hair cycle and is essential for the activation of hair growth. Following ablation of the dermal papilla, resting hair follicles do not reenter the hair cycle (Rompolas *et al.*, 2012). Thus, a crosstalk between the dermal papilla cells and epithelial stem cells of the bulge and hair germ is necessary for subsequent cycling. The dermal papilla secretes FGF7, FGF10, and the BMP inhibitor, Noggin, all of which activate the bulge and hair germ stem cells to initiate the formation of a new hair shaft (Greco *et al.*, 2009; Oshimori and Fuchs, 2012). The hair follicle also communicates with subcutaneous adipose tissue and dermal fibroblasts during all stages of growth cycles. Mature adipocytes and dermal fibroblasts secrete BMP2 and BMP4, respectively, which help maintain quiescence of the hair follicle stem cells during the resting phase of the hair cycle (Plikus *et al.*, 2008). These adipose and dermal fibroblasts secrete BMPs to form a signaling field that affects large groups (units) of hair follicles. This field maintains and synchronizes the resting phase of each unit followed by the growth phase due to decreased secretion of BMP factors (Plikus *et al.*, 2008, 2009). The immature adipocytes can stimulate dermal papilla to secrete factors triggering activation of the hair germ stem cells. Specifically, the immature adipocytes secrete mitogenic platelet-derived growth factor A (PDGFA) that binds its cognate receptor in the dermal papilla cells (Festa *et al.*, 2011); however, the downstream effectors of this signaling pathway in these cells are not yet elucidated. It is possible that other factors produced by intradermal adipocytes influence the dermal papilla or hair follicle stem cells, therefore, additional studies are needed to identify other potential signaling mechanisms (Park *et al.*, 2010).

In conclusion, the skin contains distinct sets of progenitor cells, almost all of which are derived from the bulge and, once formed, signal to each other to trigger hair growth and skin renewal as well as repair following injury. However, this tightly ordered anatomical arrangement is not typical of all tissues as addressed in the following sections, but nonetheless has served as a 'Rosetta Stone' for understanding how stem cells behave *in situ* and what signals control their behavior. One striking aspect of the skin is that while the bulge cell is the likely origin of most of the progenitor pools, there are multiple reservoirs of cells reflecting different functions (i.e., whether the hair is cycling normally or undergoing regeneration). The context of the cells is also of critical importance. For example, as outlined above, the dermal papilla cells normally play a specific role in signaling, however, if purified and then grafted into host

skin, they adopt alternate cell fates that likely reflect their mesodermal origin (Driskell *et al.*, 2009).

Skeletal Muscle

A Unique Tissue-Specific Progenitor Cell Amid Cellular Neighbors

Skeletal muscle is composed of post-mitotic multinucleated myofibers that constitute the minimal functional unit responsible for voluntary muscular contraction (Yin *et al.*, 2013). In mammals, these multinucleate fibers are incapable of further division and depend upon resident progenitor cells for repair. Mauro (1961) described a subset of cells in post-natal skeletal muscle that he called ‘satellite cells’ because of their localization underneath the basal lamina but outside, or ‘satellite’ to the myofiber plasma membrane (Figure 2). Subsequent work has largely supported the assertion that satellite cells are the key, if not unique, skeletal muscle progenitor. In addition to their anatomical location, adult satellite cells can be recognized by the *Pax7* (Sambasivan *et al.*, 2011) M-cadherin (Irintchev *et al.*, 1994) and $\alpha 7$ -integrin expression (Yin *et al.*, 2013). In response to injury, satellite cells become mitotically activated and initiate expression of *MyoD* (Relaix and Zammit, 2012; Yin *et al.*, 2013). As activated satellite cells expand, a subset will progress through the myogenic program and differentiate while others reenter a quiescent state to restore the satellite cell pool reflecting an asymmetric lineage decision (Kanisicak *et al.*, 2009). The myogenic program

includes a specific sequence of gene expression involving initial activation of *Myf5* and *MyoD*, which in turn activate the expression of terminal differentiation genes including myogenin leading to cell cycle exit (Shih *et al.*, 2008). Myogenin promotes expression of genes coding for myofiber structural and contractile proteins (Wright *et al.*, 1989). After exiting the cell cycle, myoblasts fuse together to generate multinucleated myofibers. During this process, a subset of cells transit from a *Pax7* + /*MyoD*- state to a *Pax7* + /*MyoD* + state and then back to a *Pax7* + /*MyoD*- state (Zammit *et al.*, 2006; Zammit, 2008; Yin *et al.*, 2013). Perhaps one of the most compelling observations that satellite cells are the unique myogenic resident progenitor comes from a series of studies that show that genetically mediated ablation of the satellite cell population in adult muscle leads to a complete loss of regenerative potential followed by fibrosis and ectopic fat formation (Murphy *et al.*, 2011; Sambasivan *et al.*, 2011). Unlike the hair follicle, which has multiple reservoirs of stem cells upon which to draw if the bulge compartment is completely lost, there does not appear to be a competent resident ‘primitive’ stem cell that can give rise to satellite cells.

The skeletal muscle stem cell niche is a complex environment in which cell-cell interactions are finely regulated to assure normal postnatal growth and regeneration. Physical interactions with the sarcolemma and the basal lamina regulate satellite cell quiescence (Ten Broek *et al.*, 2010; Yin *et al.*, 2013) and freshly isolated satellite cells will spontaneously reenter the cell cycle and activate the myogenic differentiation program (Musaro and Barberi, 2010). However, contact with the myofiber is not sufficient to drive quiescence suggesting

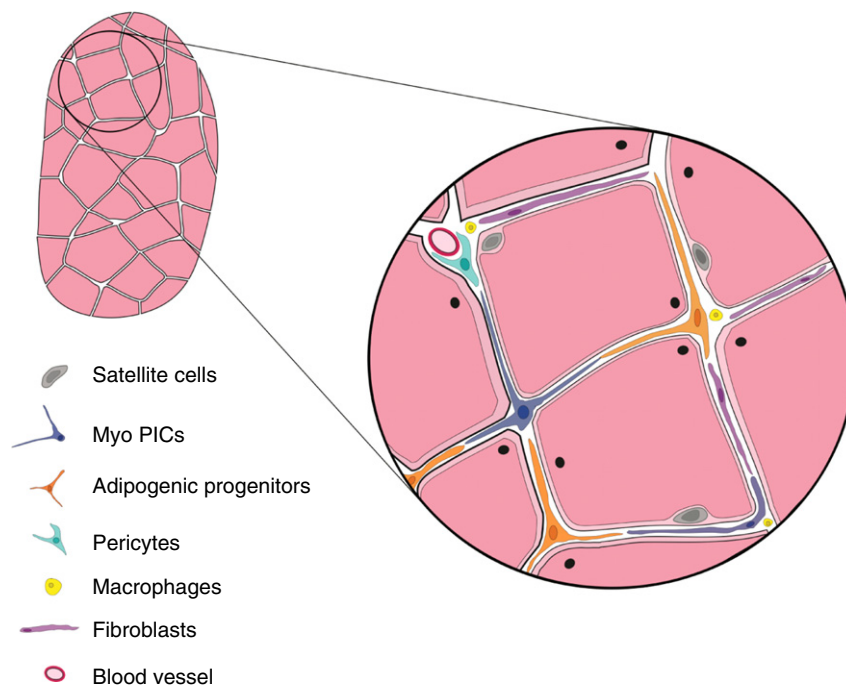


Figure 2 Schema depicting a cross-section of skeletal muscle fibers. Muscle fibers are multinucleated and comprise the bulk of muscle tissue mass. In between the fibers are vessels, and a variety of interstitial cells including adipogenic progenitors and fibroblasts. Immediately adjacent to the muscle fiber are the satellite cells that are muscle-lineage committed progenitors. As illustrated here, these cells all lie in close proximity allowing for cell-cell signaling that controls their stem cell behavior.

that other factors coming from the environment are needed (Beauchamp *et al.*, 2000; Keire *et al.*, 2013). Several paracrine soluble cytokines, chemokines, and other molecules released by the surrounding cells (myofibers, microvasculature, interstitial cells, neuromuscular junctions, and circulating or resident immune cells), as well as systemic factors maintain the quiescent state of satellite cells or promote their proliferation and differentiation (Brack *et al.*, 2007; Ten Broek *et al.*, 2010). While the satellite cell fulfills the key characteristics for a tissue-specific stem cell (i.e., they generate new muscle tissue and can self-renew), they display relatively poor engrafting potential in rescuing diseased muscle. Moreover, the endogenous satellite cells undergo exhaustion following many rounds of cell division in the context of degenerative muscle diseases (Conboy *et al.*, 2005; Brack and Rando, 2012; Pannerec *et al.*, 2012; Relaix and Zammit, 2012; Sousa-Victor *et al.*, 2014). While some evidence points to intrinsic changes that occur in satellite cells following repeated rounds of injury or chronic disease, it has been shown that a 'healthy' muscle environment is critical for proper satellite cell function and subsequent regeneration. Specifically, satellite cells are highly responsive to immediately adjacent cells including the muscle fiber, vessel associated cells and interstitial cells. Unlike the hair follicle, in which the stem cell niche consists of groups of closely associated progenitor populations in anatomically definable compartments (Figure 1), satellite cells are scattered throughout the muscle tissue, making the assignment of an anatomically definable niche less clear (Figure 2). However, as specific paracrine signaling events between the satellite cell and surrounding cell types are elucidated, a functional stem cell niche becomes apparent. In addition, myogenic capacity has been found in non-satellite cells resident in skeletal muscle (Dellavalle *et al.*, 2007, 2011; Mitchell *et al.*, 2010; Pannerec *et al.*, 2013). However a role for these cells remains unclear, particularly in light of the observation that genetic ablation of the satellite cell population specifically is sufficient to severely abrogate regenerative capacity. As we will see below, similar results are obtained when other interstitial populations are ablated, suggesting that multiple cellular components are required for proper regeneration.

In 2010, two independent reports described the isolation of an adult skeletal muscle resident cell population with adipogenic potential (Joe *et al.*, 2010; Uezumi *et al.*, 2010). This cell population, referred to most commonly as fibroadipogenic progenitors (FAPs), can be readily isolated due to the expression of Sca-1 and PDGFR α (Joe *et al.*, 2010) and identified in the interstitial space between the muscle fibers (Figure 2). Further analyses showed the existence of a common Sca-1⁺/PDGFR α ⁺ progenitor cell type responsible for adipose tissue as well as fibrosis deposition (Uezumi *et al.*, 2011). These progenitor cells do not express Pax7, α 7-integrin, or other satellite cells markers nor do they undergo myogenic differentiation under normal conditions (Joe *et al.*, 2010; Uezumi *et al.*, 2010). While skeletal muscle tissue contains both fibrotic (connective) and adipose components, chronic degenerative myopathies lead to a loss of myogenic regeneration coupled with an increase in fibrotic and fat deposition suggesting that a critical balance is maintained between these closely associated cell types. Coculture experiments have shown that FAPs exert a promyogenic effect on activated satellite cells (Joe *et al.*, 2010; Uezumi *et al.*, 2010,

2011). Using genetically mediated cell ablation models, it has been shown that in the absence of FAPs, skeletal muscle undergoes poor regeneration (Murphy *et al.*, 2011) and can trigger muscle wasting (Roberts *et al.*, 2013) underlining an important role for FAPs as essential for proper muscle regeneration and homeostasis.

Like all tissues, skeletal muscle is vascularized and in addition to the critical role of supplying blood, the vessels contain specific cell types that have progenitor properties. Pericytes are localized around small vessels in all tissues and can be identified through specific markers such as alkaline phosphatase (ALP), neuronal-glial antigen 2 chondroitin sulfate proteoglycan (NG2), cluster of differentiation-146 (CD146), and platelet-derived growth factor receptor beta (PDGFR β) (De Angelis *et al.*, 1999; Dellavalle *et al.*, 2007; Tonlorenzi *et al.*, 2007). Pericytes will proliferate *in vitro* and will differentiate into several mesodermal lineages including osteoblasts, chondroblasts, adipocytes, smooth, skeletal, and cardiac muscle cells (Doherty *et al.*, 1998; De Angelis *et al.*, 1999; Farrington-Rock *et al.*, 2004; Dellavalle *et al.*, 2007; Tonlorenzi *et al.*, 2007; Barbuti *et al.*, 2010) suggesting a high degree of cell fate plasticity. In skeletal muscle, two subpopulations of NG2⁺ pericytes have been recently described, one with myogenic capacity (PDGFR α ⁺Nestin⁻) and one with adipogenic capacity (PDGFR α ⁺Nestin⁺) (Birbrair *et al.*, 2013). While the discovery of FAPs provides clear evidence that local progenitors with distinct cell fates communicate with each other to execute proper stem cell function, there is also evidence that invading immune cells provide important clues to properly coordinate stem cell function. Immune cells are the 'first responders' following muscle injury. Injured muscle is infiltrated by monocytes and leukocytes to support resident macrophages in clearing by microbes, dead cells, and debris (Elhelu, 1983). Macrophages become activated in response to injury and these activated, or 'M1,' macrophages convert to M2 macrophages and secrete anti-inflammatory factors (Bosurgi *et al.*, 2012). Blocking the conversion of M1 to M2 macrophages blocks proper muscle regeneration including proper revascularization (Tidball and Wehling-Henricks, 2007; Wang *et al.*, 2014).

In conclusion, even though the 'stem cell' niche does not overtly show the degree of compartmentalization and topological organization that we saw in the skin, there is nonetheless a complex and important balance between the resident progenitors as well as circulating cells that come together to control the stem cell niche and its functionality. What distinguishes the skin from skeletal muscle in this regard, is the fact that these progenitor cells may well remain quiescent throughout the lifetime of the organism and do not carry out a regular cycling or 'renewal' process, but remain relatively potent throughout life, except in response to chronic disease and aging.

The Adult Central Nervous System Stem Cell Niche

A Tissue with Functional Stem Cells but Poor Regenerative Capacity

The mammalian central nervous system (CNS: brain and spinal cord) is essentially incapable of any significant regeneration in response to injury due to trauma or stroke.

Specifically, damaged neurons are not replaced and, instead, changes in neural connectivity and processing occur as the major response to brain or spinal cord injury. This stated, restricted regions of the brain do undergo neuronal turnover and the CNS possesses several populations of bona fide neuronal stem cells that give rise to all CNS cell types *in vitro* including neurons, oligodendrocytes, and astrocytes. Unlike the skin and

skeletal muscle that display pronounced regenerative capacity, the apparent inability of the CNS to regenerate originally led to the general consensus that the CNS does not contain stem cells, a view that has reversed in the last decade. As such, the field of the adult CNS stem cell niche is relatively new, and progress in this field, while rapid, lags behind that of many other tissues. Whether new knowledge in this emerging field will ultimately lead to new therapies to treat brain and spinal injuries is unclear, but this promise is one main driver in the field.

Two distinct neural stem cell niches have been identified in the CNS of adult rodents: the ventricular-subventricular zone (SVZ) next to the lateral ventricles and the subgranular zone (SGZ) in the hippocampus (Doetsch *et al.*, 1999; Seri *et al.*, 2001). Both niches have different structures and functions. During homeostasis, neuroblasts are continuously produced in the SVZ and subsequently migrate in chains through the rostral migratory stream (Figure 3; Doetsch and Alvarez-Buylla, 1996). They differentiate into olfactory interneurons once they reach the olfactory bulb. In the dentate gyrus, the SGZ neural stem cells generate granule cells (Figure 4). These newly generated neurons contribute to spatial memory, learning, and more specifically to pattern separation function, a characteristic not shared by the neighboring fully mature granule cells (Lacar *et al.*, 2014). Rare scattered progenitor cells are also found throughout the brain parenchyma, outside the two neurogenic niches described in this article.

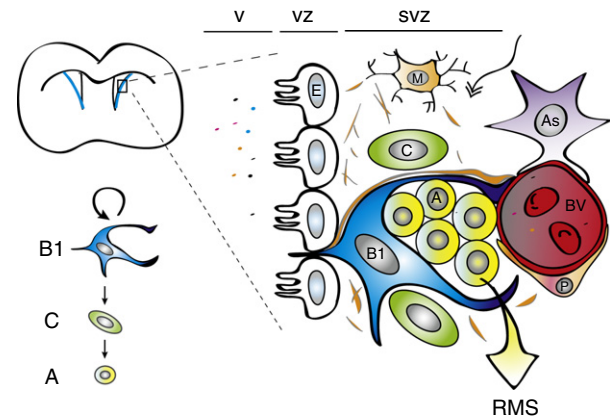


Figure 3 The ventricular-subventricular zone neural stem cell niche (SVZ). A, neuroblast; astrocyte; B1, neural stem cell; BV, blood vessel; C, transit amplifying cell; E, ependymal cell; M, ramified microglia; P, pericyte; RMS, rostral migratory stream; SVZ, subventricular zone; V, lateral ventricle; VZ, ventricular zone. Double arrow: neuronal projections of non-neurogenic regions such as raphe nuclei, substantia nigra, ventral tegmental area, and striatum. A linear orange process bridging blood vessel and ependymal cells represents Fractones.

The V-SVZ Neural Stem Cell Niche

The adult neural stem cell of the SVZ is referred to as the B1 cell (Figure 3). Each B1 cell spans three domains (Fuentelba

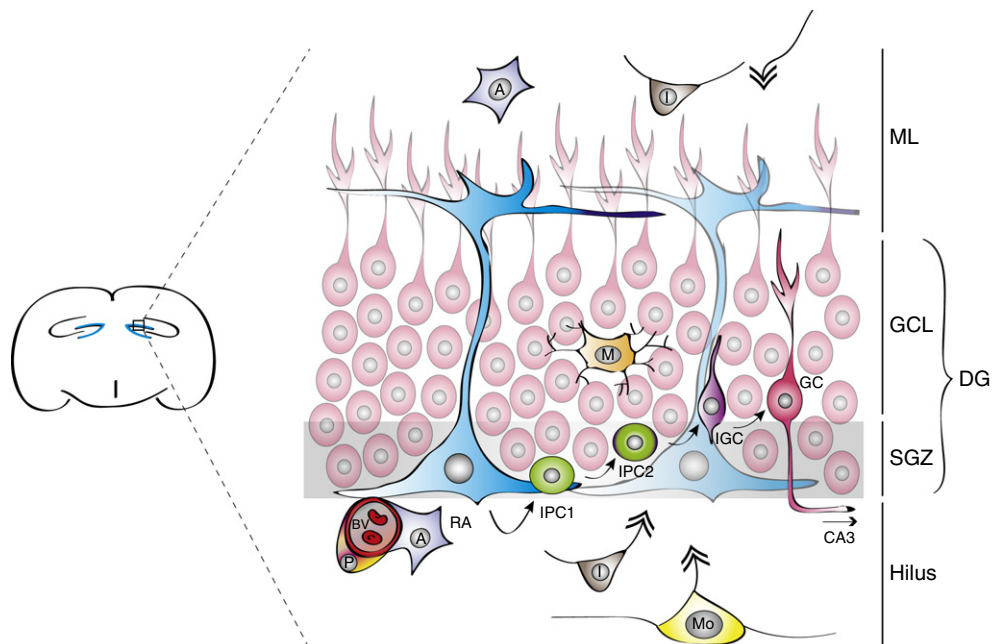


Figure 4 The subgranular zone (SGZ) neural stem cell niche. A, astrocyte; BV, blood vessel; CA3, cornu amonis 3; DG, dentate gyrus; GC, granule cell; GCL, granule cell layer; I, inhibitory interneuron; IGC, immature granule cell; IPC, intermediate precursor cell; M, ramified microglia; ML, molecular layer; Mo, mossy cell; P, pericyte; RA, radial astrocyte; SGZ, subgranular zone. Double arrows: synaptic inputs from a non-neurogenic region such as entorhinal cortex.

et al., 2012). In the proximal/apical domain, B1 cells directly contact the cerebrospinal fluid (CSF) of the lateral ventricles by a single nonmotile primary cilium (Mirzadeh *et al.*, 2008; Shen *et al.*, 2008). This process is surrounded by multiciliated ependymal cells that line the ventricular walls (Mirzadeh *et al.*, 2008). The intermediate domain contains the B1 cell body as well as B1 cell progeny, including the transit amplifying cells and neuroblasts. Microglia, the resident immune cells specific to the CNS, are also present throughout the CNS and associated with the SVZ. Finally, the distal domain links the B1 cell with the vasculature (Shen *et al.*, 2008). The blood–brain barrier is composed of endothelial cells, pericytes, and astrocytes end feet. Surprisingly, the blood–brain barrier appears to be leaky where clusters of B1 and transit amplifying cells are found (Tavazoie *et al.*, 2008). This punctual lack of glial end feet and pericyte coverage is a feature unique to the SVZ. Of note, like B1 cells, fractones bridge apical to distal domains with a close contact with each SVZ cell. These laminin-rich structures containing heparan sulfate proteoglycans are part of the niche extracellular matrix (ECM) (Mercier *et al.*, 2002).

The SGZ Neural Stem Cell Niche

This second neural stem cell niche is distinguished from the SVZ by its location and organization. The neural stem cells of the SGZ, called radial astrocytes (RAs), are present in the dentate gyrus of the hippocampus. The RAs are highly polarized and span three domains (Figure 4). The proximal domain, facing the hippocampal hilus, contains the RA primary cilium, blood vessels and is in direct contact with RAs, and to some extent hilar mossy cells and interneurons. Compared to the SVZ, angiogenesis is highly active in this domain with intense vascular remodeling (Palmer *et al.*, 2000). In the intermediate domain of the dentate gyrus, RAs extend long processes that contact other neural stem cells (Fuentelba *et al.*, 2012). RAs also possess small horizontal processes that may allow interactions with the neighboring cells. In fact, each RA cell body and its long shaft are surrounded by their differentiated progeny.

One RA successively generates intermediate progenitors type 1 and 2 (IPC), which are often found next to RA cell bodies (Kempermann *et al.*, 2003). These IPCs give rise to immature granule cells (IGC) that differentiate into fully mature granule cells. In contrast to the SVZ, newly born neurons do not migrate away from their niche and instead they reside in the dentate gyrus. Even if B1 cells and RA are found in a different microenvironment and generate different types of neurons, they do share key features. Both progenitors have ultrastructural characteristics of astrocytes with long processes. They express astroglial markers such as the glial fibrillary acidic protein (GFAP) and are found in close vicinity with the vasculature. Thus, as seen in the skin and skeletal muscle, the niche provides a microenvironment that permits multiple cell interactions.

Interactions between Niche Cells and Tissue Progenitors-Soluble Factors

A balance between proliferation and the number of differentiated cells generated is orchestrated by several mechanisms at

specific sites within the niche. Although a single signaling pathway is not restricted to one particular domain, some of the mechanisms involved in this adult neurogenic balance in each CNS niche are outlined below.

SVZ Interactions

The cerebrospinal fluid (CSF) contains multiple soluble factors such as the TGFs, IGFs, BMPs, Wnts, SHH, Slits, PDGFs, and retinoic acid (Redzic *et al.*, 2005; Lehtinen *et al.*, 2011). The main source of CSF production is the choroid plexus and the CSF composition reflects the physiological state of the animal (Redzic *et al.*, 2005). Importantly, even if B1 primary cilium is thought to integrate CSF soluble factors, highly concentrated growth factors within the lateral ventricles are able to penetrate the ependymal wall and bind to fractones (Kerever *et al.*, 2007). According to their position and local levels of EGFs, FGFs, and BMPs (i.e., BMP7) fractones can deliver growth factors to the niche cells (Mercier *et al.*, 2002). For example, neural progenitor cells express EGF receptor (EGFR) important for their proliferation (Doetsch *et al.*, 2002). Soluble factors can also originate from ependymal cells, SVZ cells, afferent axons, and systemic signals. Ependymal cells secrete noggin, a BMP antagonist, leading to SVZ progenitor proliferation and an increase in neuroblast generation (Lim *et al.*, 2000). Additionally, ependymal cells express the receptor LRP2 that confines BMP4 inducing proliferation of B1 cells (Gajera *et al.*, 2010). As with other stem cells, the maintenance of the neural stem cell pool involves a balance between proliferation, quiescence, differentiation, and self-renewal. One pathway underlying this in the SVZ is the canonical Notch signaling pathway (Imayoshi *et al.*, 2010). Notch ligands are expressed within the whole SVZ niche. As an example, transit amplifying cells express ASCL1, also known as MASH1, which increases the expression of Notch ligands that in turn inhibits B1 cell proliferation. The downstream effector of Notch HES1 transcriptionally represses ASCL1. As a second example for Notch signaling, niche astrocytes secrete Dlk1 that binds to neural stem cells in order to maintain quiescence in a dose-dependent manner (Feron *et al.*, 2011).

CNS activity is also involved in regulating the stem cell niche. For example, newly born neuroblasts spontaneously release the neurotransmitter, gamma-aminobutyric acid (GABA) (Liu *et al.*, 2005). Both transit amplifying cells and B1 cell possess GABA receptors that leads to an inhibition of cell cycle progression. This inhibition of neurogenesis is due to an epigenetic mechanism via phosphorylation of H2AX, a marker of DNA-damage (Fernando *et al.*, 2011). To compensate, B1 and transit amplifying cells secrete the diazepam binding inhibitor (DBI) into the ECM that competes with GABA for its receptors (Alfonso *et al.*, 2012). The microglia also participate in regulating the niche via the secretion of TGF- α and BDNF that enhance proliferation and neural differentiation (Walton *et al.*, 2006). As we saw in the skin stem cells, the position of the neural stem cell within the niche is critical. Endothelial cells interact closely with progenitor cells. Both ependymal cells and endothelial cells secrete the chemokine SDF1 (Kokovay *et al.*, 2010). Since endothelial cells produce higher levels of SDF1, a gradient is established along the niche.

Thus activated B1 and transit amplifying cells migrate toward the blood vessel and adhere via $\alpha 6\beta 1$ -integrin expression. Disrupting this vascular adhesion leads to proliferation defects *in vivo* (Kokovay *et al.*, 2010). Additionally, endothelial cells secrete betacellulin, an EGF-like growth factor, that induces proliferation and subsequent neurogenesis (Gomez-Gavira *et al.*, 2012). Non-neurogenic regions can also influence the neural stem cell niche. For example, dopamine is released by neuronal projections of the substantia nigra and the ventral tegmental area. This input gives rise to an increased proliferation of transit amplifying cells (Silva-Vargas *et al.*, 2013).

SGZ Interactions

Vascular secreted factors such as IGF1, brain-derived neurotrophic factor (BDNF), and VEGF promote progenitor proliferation and differentiation (Cao *et al.*, 2004; Chen *et al.*, 2005; Llorens-Martin *et al.*, 2009). As in the SVZ, canonical Notch signaling is active in the dentate gyrus neurogenic niche. Radial astrocytes express the downstream effectors of Notch HES5 and RBPJ κ , which directly binds to the Sox2 promoter and activates its transcription leading to RA quiescence (Ehm *et al.*, 2010). SOX2 signaling then inhibits neuronal fate determinant NeuroD (Kuwabara *et al.*, 2009). Nonetheless granule cells, hilar cells, and RA, as a cell-autonomous regulation, express noggin (Bonaguidi *et al.*, 2008). Thus multiple signaling pathways acting in both paracrine and autocrine fashion control the SGZ.

A Role for the CNS Stem Cell Niches

What purpose does neurogenesis serve in the brain if concentrated into small niches that do not renew neurons throughout the brain? Some evidence points to specific roles in controlling learning, memory, and basic capacities to integrate information in a process referred to as pattern separation. For example, newly born granule cells at 3–4 weeks of age are particularly different in terms of function from the fully mature ones. These young granule cells possess a unique circuitry and physiology for a limited period of time and play a dedicated role in pattern separation (Nakashiba *et al.*, 2012). In the SGZ, as seen in the SVZ, neurotransmitters play a critical role in this process. GABA, synthesized by parvalbumin interneurons, induces depolarization of IPCs toward a neuronal differentiation. At 1–2 weeks of age, they receive synaptic GABAergic inputs. The following week they start to express glutamatergic receptors and become hyperpolarized. At 4 weeks of age, they receive glutamatergic inputs from the entorhinal cortex and undergo less inhibitory inputs from the interneurons as compared to mature cells. At this stage, the young granule cells are hyperexcitable (Marin-Burgin *et al.*, 2012). Finally, around 6–8 weeks of age, granule cells are fully mature and inhibition becomes dominant.

The position of radial astrocytes in the entire dentate gyrus, from the most studied septal to temporal, may also influence their fate and function due to inputs from different brain areas. The SVZ serves to generate diverse subtypes of olfactory neurons that are specifically derived from SVZ microdomains (Merkle *et al.*, 2007). The olfactory lobe is continually supplied

with new neurons making it a rather unique exception to most other brain regions. Thus, both the SVZ and the SGZ serve very specific roles and brain functions rather than act as a reservoir for the entire brain. This suggests that perhaps the rest of the CNS does not have ‘access’ to a progenitor pool of cells. However, as outlined below, other progenitor cells have been found outside of the SVZ and SGZ that may serve this purpose, albeit, with a more limited cell fate potential such as glia.

Stem Cell Potential Outside the Neurogenic Niches

NG2 glia

NG2 glia is a specific glia cell type that is scattered in the brain parenchyma. Almost all of these cells are NG2⁺PDGFR α ⁺ and undergo a slow rate of division, more than once a month, to self-renew and produce myelinating oligodendrocytes (Dimou *et al.*, 2008; Rivers *et al.*, 2008; Psachoulia *et al.*, 2009). These represent a very limited progenitor cell type that renews the supporting glia cells but does not contribute to neurons.

Reactive Astrocytes

Astrocytes are the support cells of the CNS. Among other functions, they modulate neuronal synapses and vasculature. Even though they share expression signatures with neural stem cells, astrocytes do not divide under normal conditions. However, following brain injury such as ischemic stroke, astrocytes near the lesion undergo astrogliosis after several days. This process is characterized by a pronounced change in morphology and gene expression that leads to reactive astrocyte proliferation and acquisition of neural stem cell properties (Buffo *et al.*, 2008). This astrocytic response has both beneficial and detrimental impact on tissue repair. Notably, the astrocytes produce a scar to protect the tissue surrounding the lesion from receiving additional inflammatory signals. However, later the scar inhibits axonal regeneration and neurogenesis (Pekny and Pekna, 2004; Sofroniew and Vinters, 2010). While many of the signals released upon brain injury overlap with those found in the neural stem cell, the astrocytes are not capable of forming multiple cell types such as the progenitors found in the SVZ and SGZ (Buffo *et al.*, 2008).

In conclusion, while there are specific niches with multipotent stem cells, they appear to serve very specific functions for the brain where a more disperse population of cells exists with limited stem cell fate potential that appear to function primarily in response to injury in order to limit the extent of injury and effect a repair that is more like scar formation in other tissues. Whether these progenitor cells can be manipulated to effect more substantial repair in the brain or spinal cord will await future research in this field.

The Heart-Evidence for a Cardiac Stem Cell Niche?

Thus far, we have examined two tissues with pronounced regenerative capacity: skin that regenerates continually as well as in response to injury and skeletal muscle, which regenerates only in response to injury. In addition, we have examined the

CNS, which has little regenerative capacity but nonetheless possesses bona fide stem/progenitor cells that serve very specific roles and regions in the brain. We now look at the adult mammalian heart that is essentially incapable of any recognizable regeneration (Laflamme and Murry, 2011; Kikuchi and Poss, 2012; Xin *et al.*, 2013). At the time of writing this article, the question of whether or not the heart has stem cells capable of generating new cardiac tissue *in vivo*, in particular, new cardiac muscle, remains unresolved. As such, we can witness a field that is emerging and how different studies begin to form a picture of cardiac stem cells and their potential contribution to cardiac regeneration. As degenerative cardiac diseases are a major cause of mortality in developed countries, the field of cardiac stem cell biology has significant importance for the clinic and presents a major challenge to basic research in stem cells.

Unlike the skin and skeletal muscle, cardiac injury leads invariably to fibrosis and scar tissue formation without any significant restoration of lost cardiac muscle tissue. This suggests that the heart is devoid of progenitor cells. However, recent studies have shown robust cardiac regeneration following injury in newborn mice, although this regenerative capacity is lost by the first week of life (Porrello *et al.*, 2011; Haubner *et al.*, 2012). This raises the possibility that a transient population of progenitor cells is present in the heart just after birth. In addition, while the ability of the adult heart to regenerate in response to a large injury is absent, myocardial recovery has been reported in adult mouse hearts following drug-induced myocardial damage that is diffuse and affects cardiac muscle cells scattered throughout the heart (Ellison *et al.*, 2013). However, these findings are disputed as their conclusions are dependent upon lineage tracing mouse models that may not be as reliable as initially believed (Hesse *et al.*, 2014; Nadal-Ginard *et al.*, 2014; van Berlo *et al.*, 2014). At present, no consensus exists regarding the cellular identity of cardiac progenitors in the heart, much less the mechanisms underlying cardiac recovery from injury. It is possible that the assumption that regeneration requires the presence of a resident population of stem cells is incorrect in the case of the heart. Indeed, some species of fish and amphibians can regenerate damaged myocardium through a mechanism of cardiomyocyte dedifferentiation and proliferation and based on the available data, it appears that the neonatal mammalian heart undergoes a similar process of cardiac muscle dedifferentiation (Flink, 2002; Poss *et al.*, 2002; Hsieh *et al.*, 2007; Kuhn *et al.*, 2007; Drenckhahn *et al.*, 2008; Bersell *et al.*, 2009; Porrello *et al.*, 2011; Witman *et al.*, 2011; Haubner *et al.*, 2012). However, several reports demonstrated that endogenous cardiac stem cells (eCSCs) may contribute to the replacement of mouse cardiomyocytes during normal aging and following injury (Poss *et al.*, 2002; Beltrami *et al.*, 2003; Oh *et al.*, 2003; Hsieh *et al.*, 2007; Jesty *et al.*, 2012; Ellison *et al.*, 2013; Uchida *et al.*, 2013).

Cardiac Stem Cell Niches

If cardiac stem cells exist, where are they found and do they resemble other more characterized stem cell niches in tissues such as the skin, skeletal muscle, or CNS? Potential eCSCs

have been identified by their expression of different membrane markers (c-Kit, SCA-1, PDGFR α , MDR-1) and transcription factors (Islet-1, NKX2.5, GATA4) (Beltrami *et al.*, 2003; Oh *et al.*, 2003; Oyama *et al.*, 2007; Tallini *et al.*, 2009; Chong *et al.*, 2011, 2013; Unno *et al.*, 2012; Uchida *et al.*, 2013). While these markers have been used to purify cells for study *in vitro* and subsequent engraftment, they have also been used to identify these cells *in vivo*. These studies reveal that eCSCs are located in two 'domains': one domain is dispersed throughout the interstitial space between the cardiac muscle cells and the other domain consists of a cell layer surrounding the heart referred to as the epicardium.

Cardiac interstitial niches are located between cardiomyocyte fibers and represent randomly oriented ellipsoid structures that harbor heterogeneous, non-muscle cells, including eCSCs, lineage-committed cells (LCCs), and cardiac fibroblasts (Urbanek *et al.*, 2006; Ruiz-Villalba *et al.*, 2013). eCSCs can be identified by specific cell-surface marker (c-Kit, SCA-1, or MDR-1) but do not express cardiac cell lineage proteins and hematopoietic markers (Urbanek *et al.*, 2005). LCCs consist of progenitors expressing both eCSC markers and a lineage-specific transcription factor characteristic for cardiomyocytes, endothelium, or smooth muscle precursors (Urbanek *et al.*, 2006). Cardiac interstitial cells, including stem cells, are located throughout the heart with a much higher concentration at the apex and atria of the heart.

In addition to cardiac interstitial stem cells, progenitor cells have been identified in the ventricular epicardial and sub-epicardial regions of the heart initially based on their hypoxic environment (Kocabas *et al.*, 2012). The epicardium represents an epithelial layer on the surface of the heart adjacent to the myocardium (Schlueter and Brand, 2012). The subepicardial region encompasses the outermost three cell layers of the myocardium (Kocabas *et al.*, 2012). These areas are characterized by the lowest capillary density in the ventricle and expression of hypoxia-inducible factor 1 α (Hif1 α) (Kocabas *et al.*, 2012). Ventricular epicardial-subepicardial niche harbors so-called 'glycolytic cardiac progenitors' that resist hypoxic stress and utilize anaerobic glycolytic metabolism, rather than the oxygen-dependent mitochondrial oxidative phosphorylation. These cells do not express eCSC (c-Kit or SCA-1), endothelial, or hematopoietic markers, but do express cardiac progenitor and epicardial markers (NKX2.5, GATA4, WT1, and TBX18). Glycolytic cardiac progenitors are self-renewing and capable of differentiating into three cardiac lineages (cardiomyocytes, endothelial, and smooth muscle cells) *in vitro* (Kocabas *et al.*, 2012).

Interactions between Niche Cells and the Tissue Progenitors

Cell-cell Interactions

eCSCs are connected structurally and functionally with each other and the supporting cells by junctional and adhesion proteins – connexins and cadherins (Urbanek *et al.*, 2006). Connexins are gap junction proteins that allow direct inter-cellular communication by passage of small molecules and electrical impulses (Sohl and Willecke, 2004). Cadherins are components of adherens junctions that mediate Ca²⁺-dependent cell-cell adherence and stem cell behavior within

the niche: maintenance, proliferation, and migration (Marthiens *et al.*, 2010). N- and E-cadherin bind eCSCs and LCCs with each other and surrounding cardiomyocytes and fibroblasts, whereas Connexin 43 and 45 form gap junctions between these cells in the cardiac interstitial niches (Urbanek *et al.*, 2006).

Cell-Matrix Contacts

ECM provides structural, mechanical, and biochemical environment of the niche and influences stem cell fate; the major ECM components include laminins and fibronectins (Brizzi *et al.*, 2012; Gattazzo *et al.*, 2014). Interaction between the ECM and stem cells is mediated by a number of cell receptors, including integrins (Gattazzo *et al.*, 2014). Cardiac interstitial niches are characterized by the presence of laminin-8/9 and laminin-10/11, which are also distributed throughout the myocardium. β 1-integrin is present on the surface of both eCSCs and LCCs. However, co-localization of α 2-chain of laminin and fibronectin with α 4-Integrin is restricted only to the eCSCs and suggests a role in the preservation of the undifferentiated state of the eCSCs (Urbanek *et al.*, 2006).

Juxtacrine and Paracrine Signaling

The niche provides the paracrine (targeting cells in vicinity) and juxtacrine (targeting adjacent cells) signaling that regulates the fate of resident stem cells. The epicardium regulates myocardial growth and coronary development through secretion of paracrine factors in both fetal and injured adult mouse hearts (Lavine *et al.*, 2005, 2006; Zhou *et al.*, 2011). A number of epicardium-secreted factors have been identified in the adult injured heart including vascular endothelial growth factor A, stromal cell-derived factor 1, and members of fibroblast growth factor signaling (Zhou *et al.*, 2011). However, the effect of the epicardial paracrine signaling on eCSCs has not been described yet. Notch signaling has been shown to regulate eCSC fate in the cardiac interstitial niches. c-Kit-positive eCSCs express Notch1, whereas the adjacent, supporting cells in the niche express the Notch ligand Jagged1, a membrane-bound protein (Boni *et al.*, 2008). Activation of Notch1 results in up-regulation of NKX2.5 expression, favors the commitment of eCSCs to cardiomyocyte fate, and promotes cardiomyocyte formation (Boni *et al.*, 2008; Urbanek *et al.*, 2010). Pharmacological inhibition of Notch pathway in newborn mice results in more than 50% decrease in cardiomyocyte number and leads to development of dilated myopathy (Urbanek *et al.*, 2010). Levels of Notch1 expression decrease in adult myocardium compared to the neonatal tissue, but rise in response to acute infarction. Overexpression of Notch1 after cardiac injury promotes cardiomyocyte survival but it is unclear if stem cells are directed to a muscle cell fate in the adult (Gude *et al.*, 2008).

The epicardium is the only cardiac layer that originates outside of the primitive heart tube – from an extracardiac primordium, known as proepicardial organ. This structure exists between days 9 and 11 of embryonic development in mice. The proepicardial organ gives rise to the cell aggregates that float through the pericardial cavity, attach to the

myocardium, and fuse to form an epithelial layer around the heart (Manner *et al.*, 2001). Once established, embryonic epicardial cells, marked by expression of WT1 and TBX18, undergo an epithelial–mesenchymal transition and contribute to cardiac fibroblasts, coronary smooth muscle cells, and cardiomyocytes (Cai *et al.*, 2008; Zhou *et al.*, 2008; Christoffels *et al.*, 2009). Epicardial cells from adult mouse heart lose this capacity and remain quiescent (Zhou *et al.*, 2011). However, in response to injury, the epicardium thickens, starts expressing fetal epicardial genes (*Wt1*, *Tbx1*, and *Raldh2*), and contributes to the formation of smooth muscle cells, fibroblasts, but not cardiomyocytes (Zhou *et al.*, 2011). Adult epicardium can be stimulated to give rise to cardiomyocytes after pretreatment of the heart with thymosin β 4, which reactivates WT1 expression and also restores vascular potential of the epicardium-derived progenitors (Smart *et al.*, 2007, 2011). A population of epicardial-derived cells with colony-forming ability can be isolated based on the expression of SCA-1 and PDGFR α from adult mouse hearts. Cardiac colony-forming cells are multipotent and can give rise to cardiomyocytes, endothelial and smooth muscle cells, adipocytes, cartilage, and bone when stimulated *in vitro* (Chong *et al.*, 2011). These cells are distinct from the c-Kit-positive eCSCs and do not arise from NKX2.5-positive cardiac progenitors (Chong *et al.*, 2011).

Fibroblasts present in the interstitial niches represent a highly heterogeneous cell population that remains poorly characterized. Several markers for these cells have been described, including Vimentin, FSP1, and DDR1; however, these markers lack specificity and identify only a subpopulation of cardiac fibroblasts (Zeisberg and Kalluri, 2010). A major developmental source of the cardiac fibroblasts is the proepicardial organ, but multiple other origins have been proposed including the bone marrow, embryonic stromal cells, pericytes, and endothelium (Norris *et al.*, 2008; Krenning *et al.*, 2010; Zeisberg and Kalluri, 2010). Interestingly, embryonic and adult cardiac fibroblasts exhibit different properties. For example, embryonic cardiac fibroblasts stimulate proliferation of the cardiomyocytes, whereas adult cardiac fibroblasts induce cardiomyocyte hypertrophy (Ieda *et al.*, 2009). Cardiac fibroblasts play a major role in cardiac fibrosis characterized by excess deposition of ECM (Krenning *et al.*, 2010). Fibrosis occurs and persists in response to cardiac injury in adult hearts. However, in newborn mouse hearts, some scarring also appears after injury, but it resolves after three weeks (Porrello *et al.*, 2011; Haubner *et al.*, 2012).

While the findings summarized above support the existence of cardiac stem cells, our current knowledge of the biology of eCSC niches remains limited. The heterogeneity of eCSCs and their different properties further complicate the research of the niches. Although several studies have tried to address this issue, still little is known about eCSC niches: their architecture, distribution in the heart, interaction with the eCSCs, developmental origins, as well as changes during aging and different pathologies. The heterogeneous nature of cardiac fibroblasts that constitute the niche and the lack of specific markers hinder our understanding of their function. The crosstalk between eCSCs and adjacent cardiomyocytes also has not been addressed. Finally, considering multiple reports of cardiac regeneration due to cardiomyocyte dedifferentiation and proliferation, without participation of eCSCs, it is tempting to

suggest that cardiac niches may also play role in this process but do not directly give rise to new cardiac muscle. Instead, resident progenitors in the heart, particularly in the epicardium, may be an important source of vascular progenitors that are essential in reperfusing the heart following injury. It is well established that a quick revascularization of damaged heart tissue limits the extent of permanent injury to the heart. Future studies should provide answers to these questions.

Overall Conclusions and Future Directions

In this article, we have examined four tissue types, each representing a different challenge for stem cell biology and regenerative medicine. The skin, specifically the hair follicle, is a tissue that continually self-renews and displays a remarkable capacity to regenerate following injury. Other examples of highly regenerative and constitutively cycling tissues include the blood (bone marrow) and intestine. It is therefore no surprise that progenitor cells in these tissues are well characterized and cellular and molecular events underlying tissue homeostasis and regeneration have been described in detail. Still, these tissues hold challenges for regenerative biology. Specifically, all these tissues show a decline in regenerative capacity with age and even in the absence of damage, tissue integrity declines with age. Why this occurs is less understood. These tissues are also a sight for malignancies, some of which may be a result of stem cells that undergo transformation into precancerous and cancerous cells. Indeed, their high stem cell competence may present a higher risk of cancer. In skin, the niche is tightly organized such that the various progenitors are situated in a precise and discreet location that governs their exposure to mitogens, morphogens, and other cellular neighbors that is critical to proper function. In the case of skeletal muscle, the niche appears less highly organized – a feature that likely reflects the requirement of stem cells scattered throughout the tissue to affect repair in regions that are injured due to exercise or major physical damage. Still, as seen in the skin, there is an association of the progenitors with muscle fibers, blood vessels, and interstitial cells that are themselves progenitors (contributing to fat and fibrosis) that appear to play a role in governing stem cell competence. In skeletal muscle, there is no evidence that progenitors cycle or are recruited on a regular basis, however like in the skin, the progenitors are recruited in response to injury. In the case of the CNS, several populations of progenitors have been identified, however they do not appear to play a significant role in regeneration despite their capacity to form all the essential cell types in the CNS including neurons. It would appear that the CNS, at least in mammals, is not programmed to regenerate – a limitation that may reflect functional requirements for the brain to retain memory and learning, however, this is more speculation than fact. Nonetheless, it is possible that researchers can tap into the stem cell potential, specifically by harnessing the signals that govern the stem cell niches in the CNS to affect regeneration, at least in the spinal cord for which there is a critical need. Lastly, we examined the heart. Like many tissues in the body, the identity of resident stem cells is a work-in-progress. Clearly cells can be identified that generate blood vessels and connective tissue, but stem cells that lead to

new cardiac muscle *in vivo* appear to either exist, but not function at high capacity, or be an artifact of tissue culture techniques. Whether these cells offer any therapeutic value will be of keen interest in the coming decade. As in all tissues covered here, understanding the signals that govern cardiac progenitors in their niche will prove invaluable to the development of therapeutic approaches. The transient capacity of cardiac muscle cells to dedifferentiate may also prove to be an essential tool in the future. The stem cell field is still young and our view of stem cells in adult tissues, and how to ultimately harness them in the context of regenerative medicine remains promising for the next generation of researchers.

Acknowledgments

This work comes from the Stem Cell and Regenerative Medicine Research Team that is part of the UMRS 1166 (ICAN) research unit. The team and members are supported by the European Community Seventh Framework Program project ENDOSTEM (Activation of vasculature associated stem cells and muscle stem cells for the repair and maintenance of muscle tissue-agreement number 241440) and the ITN program INGENIUM. We are also supported by ANR 'Laboratoire d'Excellence' program REVIVE and the IHU-ICAN projects and EPISTEM.

See also: Stem Cell Biology: Structure and Function: Metabolic and Energetic Regulation of Stem Cells; Pluripotent Cells: New Tools for Disease Research and Therapies

References

- Alfonso, J., Le Magueresse, C., Zuccotti, A., Khodosevich, K., Monyer, H., 2012. Diazepam binding inhibitor promotes progenitor proliferation in the postnatal SVZ by reducing GABA signaling. *Cell Stem Cell* 10 (1), 76–87.
- Barbuti, A., Galvez, B.G., Crespi, A., *et al.*, 2010. Mesoangioblasts from ventricular vessels can differentiate *in vitro* into cardiac myocytes with sinoatrial-like properties. *Journal of Molecular and Cellular Cardiology* 48 (2), 415–423.
- Barker, N., van Es, J.H., Jaks, V., *et al.*, 2008. Very long-term self-renewal of small intestine, colon, and hair follicles from cycling Lgr5+ve stem cells. *Cold Spring Harbor Symposia on Quantitative Biology* 73, 351–356.
- Beauchamp, J.R., Heslop, L., Yu, D.S., *et al.*, 2000. Expression of CD34 and Myf5 defines the majority of quiescent adult skeletal muscle satellite cells. *Journal of Cell Biology* 151 (6), 1221–1234.
- Beltrami, A.P., Barlucchi, L., Torella, D., *et al.*, 2003. Adult cardiac stem cells are multipotent and support myocardial regeneration. *Cell* 114 (6), 763–776.
- van Berlo, J.H., Kanisicak, O., Maillet, M., *et al.*, 2014. c-kit+ cells minimally contribute cardiomyocytes to the heart. *Nature* 509 (7500), 337–341.
- Bersell, K., Arab, S., Haring, B., Kuhn, B., 2009. Neuregulin1/ErbB4 signaling induces cardiomyocyte proliferation and repair of heart injury. *Cell* 138 (2), 257–270.
- Birbrair, A., Zhang, T., Wang, Z.M., *et al.*, 2013. Role of pericytes in skeletal muscle regeneration and fat accumulation. *Stem Cells and Development* 22 (16), 2298–2314.
- Blanpain, C., Lowry, W.E., Geoghegan, A., Polak, L., Fuchs, E., 2004. Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. *Cell* 118 (5), 635–648.
- Bonaguidi, M.A., Peng, C.Y., McGuire, T., *et al.*, 2008. Noggin expands neural stem cells in the adult hippocampus. *Journal of Neuroscience* 28 (37), 9194–9204.
- Boni, A., Urbanek, K., Nascimbene, A., *et al.*, 2008. Notch1 regulates the fate of cardiac progenitor cells. *Proceedings of the National Academy of Sciences of the United States of America* 105 (40), 15529–15534.

- Bosurgi, L., Corna, G., Vezzoli, M., *et al.*, 2012. Transplanted mesoangioblasts require macrophage IL-10 for survival in a mouse model of muscle injury. *Journal of Immunology* 188 (12), 6267–6277.
- Brack, A.S., Conboy, M.J., Roy, S., *et al.*, 2007. Increased Wnt signaling during aging alters muscle stem cell fate and increases fibrosis. *Science* 317 (5839), 807–810.
- Brack, A.S., Rando, T.A., 2012. Tissue-specific stem cells: Lessons from the skeletal muscle satellite cell. *Cell Stem Cell* 10 (5), 504–514.
- Brizzi, M.F., Tarone, G., Defilippi, P., 2012. Extracellular matrix, integrins, and growth factors as tailors of the stem cell niche. *Current Opinion in Cell Biology* 24 (5), 645–651.
- Brownell, I., Guevara, E., Bai, C.B., Loomis, C.A., Joyner, A.L., 2011. Nerve-derived sonic hedgehog defines a niche for hair follicle stem cells capable of becoming epidermal stem cells. *Cell Stem Cell* 8 (5), 552–565.
- Buffo, A., Rite, I., Tripathi, P., *et al.*, 2008. Origin and progeny of reactive gliosis: A source of multipotent cells in the injured brain. *Proceedings of the National Academy of Sciences of the United States of America* 105 (9), 3581–3586.
- Cai, C.L., Martin, J.C., Sun, Y., *et al.*, 2008. A myocardial lineage derives from Tbx18 epicardial cells. *Nature* 454 (7200), 104–108.
- Cao, L., Jiao, X., Zuzga, D.S., *et al.*, 2004. VEGF links hippocampal activity with neurogenesis, learning and memory. *Nature Genetics* 36 (8), 827–835.
- Chen, J., Zacharek, A., Zhang, C., *et al.*, 2005. Endothelial nitric oxide synthase regulates brain-derived neurotrophic factor expression and neurogenesis after stroke in mice. *Journal of Neuroscience* 25 (9), 2366–2375.
- Chong, J.J., Chandrakanthan, V., Xaymardan, M., *et al.*, 2011. Adult cardiac-resident MSC-like stem cells with a prepericardial origin. *Cell Stem Cell* 9 (6), 527–540.
- Chong, J.J., Reinecke, H., Iwata, M., *et al.*, 2013. Progenitor cells identified by PDGFR- α expression in the developing and diseased human heart. *Stem Cells and Development* 22 (13), 1932–1943.
- Christoffels, V.M., Grieskamp, T., Norden, J., *et al.*, 2009. Tbx18 and the fate of epicardial progenitors. *Nature* 458 (7240), E8–E9. Discussion E9–10.
- Conboy, I.M., Conboy, M.J., Wagers, A.J., *et al.*, 2005. Rejuvenation of aged progenitor cells by exposure to a young systemic environment. *Nature* 433 (7027), 760–764.
- Cotsarelis, G., Sun, T.T., Lavker, R.M., 1990. Label-retaining cells reside in the bulge area of pilosebaceous unit: Implications for follicular stem cells, hair cycle, and skin carcinogenesis. *Cell* 61 (7), 1329–1337.
- De Angelis, L., Berghella, L., Coletta, M., *et al.*, 1999. Skeletal myogenic progenitors originating from embryonic dorsal aorta coexpress endothelial and myogenic markers and contribute to postnatal muscle growth and regeneration. *Journal of Cell Biology* 147 (4), 869–878.
- Dellavalle, A., Maroli, G., Covarello, D., *et al.*, 2011. Pericytes resident in postnatal skeletal muscle differentiate into muscle fibres and generate satellite cells. *Nature Communications* 2, 499.
- Dellavalle, A., Sampaolesi, M., Tonlorenzi, R., *et al.*, 2007. Pericytes of human skeletal muscle are myogenic precursors distinct from satellite cells. *Nature Cell Biology* 9 (3), 255–267.
- Dimou, L., Simon, C., Kirchhoff, F., Takebayashi, H., Gotz, M., 2008. Progeny of Olig2-expressing progenitors in the gray and white matter of the adult mouse cerebral cortex. *Journal of Neuroscience* 28 (41), 10434–10442.
- Doetsch, F., Alvarez-Buylla, A., 1996. Network of tangential pathways for neuronal migration in adult mammalian brain. *Proceedings of the National Academy of Sciences of the United States of America* 93 (25), 14895–14900.
- Doetsch, F., Caille, I., Lim, D.A., Garcia-Verdugo, J.M., Alvarez-Buylla, A., 1999. Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell* 97 (6), 703–716.
- Doetsch, F., Petreanu, L., Caille, I., Garcia-Verdugo, J.M., Alvarez-Buylla, A., 2002. EGF converts transit-amplifying neurogenic precursors in the adult brain into multipotent stem cells. *Neuron* 36 (6), 1021–1034.
- Doherty, M.J., Ashton, B.A., Walsh, S., *et al.*, 1998. Vascular pericytes express osteogenic potential *in vitro* and *in vivo*. *Journal of Bone and Mineral Research* 13 (5), 828–838.
- Drenckhahn, J.D., Schwarz, Q.P., Gray, S., *et al.*, 2008. Compensatory growth of healthy cardiac cells in the presence of diseased cells restores tissue homeostasis during heart development. *Developmental Cell* 15 (4), 521–533.
- Driskell, R.R., Clavel, C., Rendl, M., Watt, F.M., 2011. Hair follicle dermal papilla cells at a glance. *Journal of Cell Science* 124 (Pt 8), 1179–1182.
- Driskell, R.R., Giangreco, A., Jensen, K.B., Mulder, K.W., Watt, F.M., 2009. Sox2-positive dermal papilla cells specify hair follicle type in mammalian epidermis. *Development* 136 (16), 2815–2823.
- Ehm, O., Goritz, C., Covic, M., *et al.*, 2010. RBPJkappa-dependent signaling is essential for long-term maintenance of neural stem cells in the adult hippocampus. *Journal of Neuroscience* 30 (41), 13794–13807.
- Elhelu, M.A., 1983. The role of macrophages in immunology. *Journal of the National Medical Association* 75 (3), 314–317.
- Ellison, G.M., Vicinanza, C., Smith, A.J., *et al.*, 2013. Adult c-kit(pos) cardiac stem cells are necessary and sufficient for functional cardiac regeneration and repair. *Cell* 154 (4), 827–842.
- Enshell-Seijffers, D., Lindon, C., Morgan, B.A., 2008. The serine protease Corin is a novel modifier of the Agouti pathway. *Development* 135 (2), 217–225.
- Farrington-Rock, C., Crofts, N.J., Doherty, M.J., *et al.*, 2004. Chondrogenic and adipogenic potential of microvascular pericytes. *Circulation* 110 (15), 2226–2232.
- Fernando, R.N., Eleuteri, B., Abdelhady, S., *et al.*, 2011. Cell cycle restriction by histone H2AX limits proliferation of adult neural stem cells. *Proceedings of the National Academy of Sciences of the United States of America* 108 (14), 5837–5842.
- Ferron, S.R., Charalambous, M., Radford, E., *et al.*, 2011. Postnatal loss of Dlk1 imprinting in stem cells and niche astrocytes regulates neurogenesis. *Nature* 475 (7356), 381–385.
- Festa, E., Fretz, J., Berry, R., *et al.*, 2011. Adipocyte lineage cells contribute to the skin stem cell niche to drive hair cycling. *Cell* 146 (5), 761–771.
- Flink, I.L., 2002. Cell cycle reentry of ventricular and atrial cardiomyocytes and cells within the epicardium following amputation of the ventricular apex in the axolotl, *Amblystoma mexicanum*: Confocal microscopic immunofluorescent image analysis of bromodeoxyuridine-labeled nuclei. *Anatomy and Embryology (Berlin)* 205 (3), 235–244.
- Fuentealba, L.C., Obernier, K., Alvarez-Buylla, A., 2012. Adult neural stem cells bridge their niche. *Cell Stem Cell* 10 (6), 698–708.
- Fujiwara, H., Ferreira, M., Donati, G., *et al.*, 2011. The basement membrane of hair follicle stem cells is a muscle cell niche. *Cell* 144 (4), 577–589.
- Gajera, C.R., Emich, H., Lioubinski, O., *et al.*, 2010. LRP2 in ependymal cells regulates BMP signaling in the adult neurogenic niche. *Journal of Cell Science* 123 (Pt 11), 1922–1930.
- Gattazzo, F., Urciuolo, A., Bonaldo, P., 2014. Extracellular matrix: A dynamic microenvironment for stem cell niche. *Biochimica et Biophysica Acta* 1840 (8), 2506–2519.
- Gomez-Gaviro, M.V., Scott, C.E., Sesay, A.K., *et al.*, 2012. Betacellulin promotes cell proliferation in the neural stem cell niche and stimulates neurogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 109 (4), 1317–1322.
- Greco, V., Chen, T., Rendl, M., *et al.*, 2009. A two-step mechanism for stem cell activation during hair regeneration. *Cell Stem Cell* 4 (2), 155–169.
- Grisanti, L., Clavel, C., Cai, X., *et al.*, 2013. Tbx18 targets dermal condensates for labeling, isolation, and gene ablation during embryonic hair follicle formation. *Journal of Investigative Dermatology* 133 (2), 344–353.
- Gude, N.A., Emmanuel, G., Wu, W., *et al.*, 2008. Activation of Notch-mediated protective signaling in the myocardium. *Circulation Research* 102 (9), 1025–1035.
- Haubner, B.J., Adamowicz-Brice, M., Khadayat, S., *et al.*, 2012. Complete cardiac regeneration in a mouse model of myocardial infarction. *Aging (Albany, NY)* 4 (12), 966–977.
- Hesse, M., Fleischmann, B.K., Kotlikoff, M.I., 2014. Concise review: The role of C-kit expressing cells in heart repair at the neonatal and adult stage. *Stem Cells* 32 (7), 1701–1712.
- Horsley, V., O'Carroll, D., Tooze, R., *et al.*, 2006. Blimp1 defines a progenitor population that governs cellular input to the sebaceous gland. *Cell* 126 (3), 597–609.
- Hsieh, P.C., Segers, V.F., Davis, M.E., *et al.*, 2007. Evidence from a genetic fate-mapping study that stem cells refresh adult mammalian cardiomyocytes after injury. *Nature Medicine* 13 (8), 970–974.
- Ieda, M., Tsuchihashi, T., Ivey, K.N., *et al.*, 2009. Cardiac fibroblasts regulate myocardial proliferation through beta1 integrin signaling. *Developmental Cell* 16 (2), 233–244.
- Imayoshi, I., Sakamoto, M., Yamaguchi, M., Mori, K., Kageyama, R., 2010. Essential roles of Notch signaling in maintenance of neural stem cells in developing and adult brains. *Journal of Neuroscience* 30 (9), 3489–3498.
- Irintchev, A., Zeschnigk, M., Starzinski-Powitz, A., Wernig, A., 1994. Expression pattern of M-cadherin in normal, denervated, and regenerating mouse muscles. *Developmental Dynamics* 199 (4), 326–337.
- Ito, M., Kizawa, K., Hamada, K., Cotsarelis, G., 2004. Hair follicle stem cells in the lower bulge form the secondary germ, a biochemically distinct but functionally equivalent progenitor cell population, at the termination of catagen. *Differentiation* 72 (9–10), 548–557.
- Jensen, U.B., Yan, X., Triel, C., *et al.*, 2008. A distinct population of clonogenic and multipotent murine follicular keratinocytes residing in the upper isthmus. *Journal of Cell Science* 121 (Pt 5), 609–617.

- Jesty, S.A., Steffey, M.A., Lee, F.K., *et al.*, 2012. c-kit⁺ precursors support postinfarction myogenesis in the neonatal, but not adult, heart. *Proceedings of the National Academy of Sciences of the United States of America* 109 (33), 13380–13385.
- Joe, A.W., Yi, L., Natarajan, A., *et al.*, 2010. Muscle injury activates resident fibro/adipogenic progenitors that facilitate myogenesis. *Nature Cell Biology* 12 (2), 153–163.
- Kadaja, M., Keyes, B.E., Lin, M., *et al.*, 2014. SOX9: A stem cell transcriptional regulator of secreted niche signaling factors. *Genes & Development* 28 (4), 328–341.
- Kanischak, O., Mendez, J.J., Yamamoto, S., Yamamoto, M., Goldhamer, D.J., 2009. Progenitors of skeletal muscle satellite cells express the muscle determination gene, MyoD. *Developmental Biology* 332 (1), 131–141.
- Keire, P., Shearer, A., Shefer, G., Yablonska-Reuveni, Z., 2013. Isolation and culture of skeletal muscle myofibers as a means to analyze satellite cells. *Methods in Molecular Biology* 946, 431–468.
- Kempermann, G., Gast, D., Kronenberg, G., Yamaguchi, M., Gage, F.H., 2003. Early determination and long-term persistence of adult-generated new neurons in the hippocampus of mice. *Development* 130 (2), 391–399.
- Kerever, A., Schnack, J., Vellinga, D., *et al.*, 2007. Novel extracellular matrix structures in the neural stem cell niche capture the neurogenic factor fibroblast growth factor 2 from the extracellular milieu. *Stem Cells* 25 (9), 2146–2157.
- Kikuchi, K., Poss, K.D., 2012. Cardiac regenerative capacity and mechanisms. *Annual Review of Cell and Developmental Biology* 28, 719–741.
- Kocabas, F., Mahmoud, A.I., Sosic, D., *et al.*, 2012. The hypoxic epicardial and subepicardial microenvironment. *Journal of Cardiovascular Translational Research* 5 (5), 654–665.
- Kokovay, E., Goderie, S., Wang, Y., *et al.*, 2010. Adult SVZ lineage cells home to and leave the vascular niche via differential responses to SDF1/CXCR4 signaling. *Cell Stem Cell* 7 (2), 163–173.
- Krenning, G., Zeisberg, E.M., Kalluri, R., 2010. The origin of fibroblasts and mechanism of cardiac fibrosis. *Journal of Cellular Physiology* 225 (3), 631–637.
- Kuhn, B., del Monte, F., Hajjar, R.J., *et al.*, 2007. Periostin induces proliferation of differentiated cardiomyocytes and promotes cardiac repair. *Nature Medicine* 13 (8), 962–969.
- Kuwabara, T., Hsieh, J., Muotri, A., *et al.*, 2009. Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. *Nature Neuroscience* 12 (9), 1097–1105.
- Lacar, B., Parylak, S.L., Vadodaria, K.C., Sarkar, A., Gage, F.H., 2014. Increasing the resolution of the adult neurogenesis picture. *F1000Prime Reports* 6, 8.
- Lafamme, M.A., Murry, C.E., 2011. Heart regeneration. *Nature* 473 (7347), 326–335.
- Lavine, K.J., White, A.C., Park, C., *et al.*, 2006. Fibroblast growth factor signals regulate a wave of Hedgehog activation that is essential for coronary vascular development. *Genes & Development* 20 (12), 1651–1666.
- Lavine, K.J., Yu, K., White, A.C., *et al.*, 2005. Endocardial and epicardial derived FGF signals regulate myocardial proliferation and differentiation *in vivo*. *Developmental Cell* 8 (1), 85–95.
- Lehtinen, M.K., Zappaterra, M.W., Chen, X., *et al.*, 2011. The cerebrospinal fluid provides a proliferative niche for neural progenitor cells. *Neuron* 69 (5), 893–905.
- Lim, D.A., Tramontin, A.D., Trevejo, J.M., *et al.*, 2000. Noggin antagonizes BMP signaling to create a niche for adult neurogenesis. *Neuron* 28 (3), 713–726.
- Liu, X., Wang, Q., Haydar, T.F., Bordey, A., 2005. Nonsynaptic GABA signaling in postnatal subventricular zone controls proliferation of GFAP-expressing progenitors. *Nature Neuroscience* 8 (9), 1179–1187.
- Liu, Y., Lyle, S., Yang, Z., Cotsarelis, G., 2003. Keratin 15 promoter targets putative epithelial stem cells in the hair follicle bulge. *Journal of Investigative Dermatology* 121 (5), 963–968.
- Llorens-Martin, M., Torres-Aleman, I., Trejo, J.L., 2009. Mechanisms mediating brain plasticity: IGF1 and adult hippocampal neurogenesis. *Neuroscientist* 15 (2), 134–148.
- Magnusdottir, E., Kalachikov, S., Mizukoshi, K., *et al.*, 2007. Epidermal terminal differentiation depends on B lymphocyte-induced maturation protein-1. *Proceedings of the National Academy of Sciences of the United States of America* 104 (38), 14988–14993.
- Manner, J., Perez-Pomares, J.M., Macias, D., Munoz-Chapuli, R., 2001. The origin, formation and developmental significance of the epicardium: A review. *Cells Tissues Organs* 169 (2), 89–103.
- Marin-Burgin, A., Mongiat, L.A., Pardi, M.B., Schinder, A.F., 2012. Unique processing during a period of high excitation/inhibition balance in adult-born neurons. *Science* 335 (6073), 1238–1242.
- Marthiens, V., Kazanis, I., Moss, L., Long, K., French-Constant, C., 2010. Adhesion molecules in the stem cell niche—more than just staying in shape? *Journal of Cell Science* 123 (Pt 10), 1613–1622.
- Mauro, A., 1961. Satellite cell of skeletal muscle fibers. *Journal of Biophysical and Biochemical Cytology* 9, 493–495.
- Mercier, F., Kitasako, J.T., Hatton, G.I., 2002. Anatomy of the brain neurogenic zones revisited: Fractones and the fibroblast/macrophage network. *Journal of Comparative Neurology* 451 (2), 170–188.
- Merkle, F.T., Mirzadeh, Z., Alvarez-Buylla, A., 2007. Mosaic organization of neural stem cells in the adult brain. *Science* 317 (5836), 381–384.
- Mirzadeh, Z., Merkle, F.T., Soriano-Navarro, M., Garcia-Verdugo, J.M., Alvarez-Buylla, A., 2008. Neural stem cells confer unique pinwheel architecture to the ventricular surface in neurogenic regions of the adult brain. *Cell Stem Cell* 3 (3), 265–278.
- Mitchell, K.J., Pannerec, A., Cadot, B., *et al.*, 2010. Identification and characterization of a non-satellite cell muscle resident progenitor during postnatal development. *Nature Cell Biology* 12 (3), 257–266.
- Morris, R.J., Liu, Y., Marles, L., *et al.*, 2004. Capturing and profiling adult hair follicle stem cells. *Nature Biotechnology* 22 (4), 411–417.
- Murphy, M.M., Lawson, J.A., Mathew, S.J., Hutcheson, D.A., Kardon, G., 2011. Satellite cells, connective tissue fibroblasts and their interactions are crucial for muscle regeneration. *Development* 138 (17), 3625–3637.
- Musaro, A., Barberi, L., 2010. Isolation and culture of mouse satellite cells. *Methods in Molecular Biology* 633, 101–111.
- Nadal-Ginard, B., Ellison, G.M., Torella, D., 2014. The cardiac stem cell compartment is indispensable for myocardial cell homeostasis, repair and regeneration in the adult. *Stem Cell Research* 13 (3 Pt B), 615–630.
- Nakashiba, T., Cushman, J.D., Pelkey, K.A., *et al.*, 2012. Young dentate granule cells mediate pattern separation, whereas old granule cells facilitate pattern completion. *Cell* 149 (1), 188–201.
- Nijhof, J.G., Braun, K.M., Giangreco, A., *et al.*, 2006. The cell-surface marker MTS24 identifies a novel population of follicular keratinocytes with characteristics of progenitor cells. *Development* 133 (15), 3027–3037.
- Norris, R.A., Borg, T.K., Butcher, J.T., *et al.*, 2008. Neonatal and adult cardiovascular pathophysiological remodeling and repair: Developmental role of periostin. *Annals of the New York Academy of Sciences* 1123, 30–40.
- Oh, H., Bradfute, S.B., Gallardo, T.D., *et al.*, 2003. Cardiac progenitor cells from adult myocardium: Homing, differentiation, and fusion after infarction. *Proceedings of the National Academy of Sciences of the United States of America* 100 (21), 12313–12318.
- Oshimori, N., Fuchs, E., 2012. Paracrine TGF- β signaling counterbalances BMP-mediated repression in hair follicle stem cell activation. *Cell Stem Cell* 10 (1), 63–75.
- Oyama, T., Nagai, T., Wada, H., *et al.*, 2007. Cardiac side population cells have a potential to migrate and differentiate into cardiomyocytes *in vitro* and *in vivo*. *Journal of Cell Biology* 176 (3), 329–341.
- Page, M.E., Lombard, P., Ng, F., Gottgens, B., Jensen, K.B., 2013. The epidermis comprises autonomous compartments maintained by distinct stem cell populations. *Cell Stem Cell* 13 (4), 471–482.
- Palmer, T.D., Willhoite, A.R., Gage, F.H., 2000. Vascular niche for adult hippocampal neurogenesis. *Journal of Comparative Neurology* 425 (4), 479–494.
- Pannerec, A., Formicola, L., Besson, V., Marazzi, G., Sassoon, D.A., 2013. Defining skeletal muscle resident progenitors and their cell fate potentials. *Development* 140 (14), 2879–2891.
- Pannerec, A., Marazzi, G., Sassoon, D., 2012. Stem cells in the hood: The skeletal muscle niche. *Trends in Molecular Medicine* 18 (10), 599–606.
- Park, B.S., Kim, W.S., Choi, J.S., *et al.*, 2010. Hair growth stimulated by conditioned medium of adipose-derived stem cells is enhanced by hypoxia: Evidence of increased growth factor secretion. *BioMed Research International* 31 (1), 27–34.
- Paus, R., Cotsarelis, G., 1999. The biology of hair follicles. *New England Journal of Medicine* 341 (7), 491–497.
- Pekny, M., Pekna, M., 2004. Astrocyte intermediate filaments in CNS pathologies and regeneration. *Journal of Pathology* 204 (4), 428–437.
- Plikus, M.V., Mayer, J.A., de la Cruz, D., *et al.*, 2008. Cyclic dermal BMP signalling regulates stem cell activation during hair regeneration. *Nature* 451 (7176), 340–344.
- Plikus, M.V., Widelitz, R.B., Maxson, R., Chuong, C.M., 2009. Analyses of regenerative wave patterns in adult hair follicle populations reveal macro-environmental regulation of stem cell activity. *International Journal of Developmental Biology* 53 (5–6), 857–868.
- Porrello, E.R., Mahmoud, A.I., Simpson, E., *et al.*, 2011. Transient regenerative potential of the neonatal mouse heart. *Science* 331 (6020), 1078–1080.
- Poss, K.D., Wilson, L.G., Keating, M.T., 2002. Heart regeneration in zebrafish. *Science* 298 (5601), 2188–2190.

- Psachoulia, K., Jamen, F., Young, K.M., Richardson, W.D., 2009. Cell cycle dynamics of NG2 cells in the postnatal and ageing brain. *Neuron Glia Biology* 5 (3–4), 57–67.
- Redzic, Z.B., Preston, J.E., Duncan, J.A., Chodobski, A., Szmydynger-Chodobska, J., 2005. The choroid plexus-cerebrospinal fluid system: From development to aging. *Current Topics in Developmental Biology* 71, 1–52.
- Relaix, F., Zammit, P.S., 2012. Satellite cells are essential for skeletal muscle regeneration: The cell on the edge returns centre stage. *Development* 139 (16), 2845–2856.
- Rivers, L.E., Young, K.M., Rizzi, M., *et al.*, 2008. PDGFRA/NG2 glia generate myelinating oligodendrocytes and piriform projection neurons in adult mice. *Nature Neuroscience* 11 (12), 1392–1401.
- Roberts, E.W., Deonaraine, A., Jones, J.O., *et al.*, 2013. Depletion of stromal cells expressing fibroblast activation protein- α from skeletal muscle and bone marrow results in cachexia and anemia. *Journal of Experimental Medicine* 210 (6), 1137–1151.
- Rompolas, P., Deschene, E.R., Zito, G., *et al.*, 2012. Live imaging of stem cell and progeny behaviour in physiological hair-follicle regeneration. *Nature* 487 (7408), 496–499.
- Ruiz-Villalba, A., Ziogas, A., Ehrbar, M., Perez-Pomares, J.M., 2013. Characterization of epicardial-derived cardiac interstitial cells: Differentiation and mobilization of heart fibroblast progenitors. *PLoS One* 8 (1), e53694.
- Sambasivan, R., Yao, R., Kissenpfennig, A., *et al.*, 2011. Pax7-expressing satellite cells are indispensable for adult skeletal muscle regeneration. *Development* 138 (17), 3647–3656.
- Schluter, J., Brand, T., 2012. Epicardial progenitor cells in cardiac development and regeneration. *Journal of Cardiovascular Translational Research* 5 (5), 641–653.
- Seri, B., Garcia-Verdugo, J.M., McEwen, B.S., Alvarez-Buylla, A., 2001. Astrocytes give rise to new neurons in the adult mammalian hippocampus. *Journal of Neuroscience* 21 (18), 7153–7160.
- Shen, Q., Wang, Y., Kokovay, E., *et al.*, 2008. Adult SVZ stem cells lie in a vascular niche: A quantitative analysis of niche cell-cell interactions. *Cell Stem Cell* 3 (3), 289–300.
- Shih, H.P., Gross, M.K., Kioussi, C., 2008. Muscle development: Forming the head and trunk muscles. *Acta Histochemica* 110 (2), 97–108.
- Silva-Vargas, V., Crouch, E.E., Doetsch, F., 2013. Adult neural stem cells and their niche: A dynamic duo during homeostasis, regeneration, and aging. *Current Opinion in Neurobiology* 23 (6), 935–942.
- Smart, N., Bollini, S., Dube, K.N., *et al.*, 2011. De novo cardiomyocytes from within the activated adult heart after injury. *Nature* 474 (7353), 640–644.
- Smart, N., Risebro, C.A., Melville, A.A., *et al.*, 2007. Thymosin β 4 induces adult epicardial progenitor mobilization and neovascularization. *Nature* 445 (7124), 177–182.
- Snippert, H.J., Haegebarth, A., Kasper, M., *et al.*, 2010. Lgr6 marks stem cells in the hair follicle that generate all cell lineages of the skin. *Science* 327 (5971), 1385–1389.
- Sofroniew, M.V., Vinters, H.V., 2010. Astrocytes: Biology and pathology. *Acta Neuropathologica* 119 (1), 7–35.
- Sohl, G., Willecke, K., 2004. Gap junctions and the connexin protein family. *Cardiovascular Research* 62 (2), 228–232.
- Sousa-Victor, P., Gutarra, S., Garcia-Prat, L., *et al.*, 2014. Geriatric muscle stem cells switch reversible quiescence into senescence. *Nature* 506 (7488), 316–321.
- Tallini, Y.N., Greene, K.S., Craven, M., *et al.*, 2009. c-kit expression identifies cardiovascular precursors in the neonatal heart. *Proceedings of the National Academy of Sciences of the United States of America* 106 (6), 1808–1813.
- Tavazoie, M., Van der Veken, L., Silva-Vargas, V., *et al.*, 2008. A specialized vascular niche for adult neural stem cells. *Cell Stem Cell* 3 (3), 279–288.
- Ten Broek, R.W., Grefte, S., Von den Hoff, J.W., 2010. Regulatory factors and cell populations involved in skeletal muscle regeneration. *Journal of Cellular Physiology* 224 (1), 7–16.
- Tidball, J.G., Wehling-Henricks, M., 2007. Macrophages promote muscle membrane repair and muscle fibre growth and regeneration during modified muscle loading in mice *in vivo*. *Journal of Physiology* 578 (Pt 1), 327–336.
- Tonlorenzi, R., Dellavalle, A., Schnapp, E., Cossu, G., Sampaolesi, M., 2007. Isolation and characterization of mesoangioblasts from mouse, dog, and human tissues. *Current Protocols in Stem Cell Biology*. Chapter 2, Unit 2B 1.
- Tornqvist, G., Sandberg, A., Hagglund, A.C., Carlsson, L., 2010. Cyclic expression of *lhx2* regulates hair formation. *PLoS Genetics* 6 (4), e1000904.
- Toyoshima, K.E., Asakawa, K., Ishibashi, N., *et al.*, 2012. Fully functional hair follicle regeneration through the rearrangement of stem cells and their niches. *Nature Communications* 3, 784.
- Tumbar, T., Guasch, G., Greco, V., *et al.*, 2004. Defining the epithelial stem cell niche in skin. *Science* 303 (5656), 359–363.
- Uchida, S., De Gaspari, P., Kostin, S., *et al.*, 2013. Sca1-derived cells are a source of myocardial renewal in the murine adult heart. *Stem Cell Reports* 1 (5), 397–410.
- Uezumi, A., Fukada, S., Yamamoto, N., Takeda, S., Tsuchida, K., 2010. Mesenchymal progenitors distinct from satellite cells contribute to ectopic fat cell formation in skeletal muscle. *Nature Cell Biology* 12 (2), 143–152.
- Uezumi, A., Ito, T., Morikawa, D., *et al.*, 2011. Fibrosis and adipogenesis originate from a common mesenchymal progenitor in skeletal muscle. *Journal of Cell Science* 124 (Pt 21), 3654–3664.
- Unno, K., Jain, M., Liao, R., 2012. Cardiac side population cells: Moving toward the center stage in cardiac regeneration. *Circulation Research* 110 (10), 1355–1363.
- Urbanek, K., Cabral-da-Silva, M.C., Ide-Iwata, N., *et al.*, 2010. Inhibition of notch1-dependent cardiomyogenesis leads to a dilated myopathy in the neonatal heart. *Circulation Research* 107 (3), 429–441.
- Urbanek, K., Cesselli, D., Rota, M., *et al.*, 2006. Stem cell niches in the adult mouse heart. *Proceedings of the National Academy of Sciences of the United States of America* 103 (24), 9226–9231.
- Urbanek, K., Torella, D., Sheikh, F., *et al.*, 2005. Myocardial regeneration by activation of multipotent cardiac stem cells in ischemic heart failure. *Proceedings of the National Academy of Sciences of the United States of America* 102 (24), 8692–8697.
- Walton, N.M., Sutter, B.M., Laywell, E.D., *et al.*, 2006. Microglia instruct subventricular zone neurogenesis. *Glia* 54 (8), 815–825.
- Wang, H., Melton, D.W., Porter, L., *et al.*, 2014. Altered macrophage phenotype transition impairs skeletal muscle regeneration. *American Journal of Pathology* 184 (4), 1167–1184.
- Witman, N., Murtuza, B., Davis, B., Arner, A., Morrison, J.I., 2011. Recapitulation of developmental cardiogenesis governs the morphological and functional regeneration of adult newt hearts following injury. *Developmental Biology* 354 (1), 67–76.
- Wright, W.E., Sassoon, D.A., Lin, V.K., 1989. Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. *Cell* 56 (4), 607–617.
- Xin, M., Olson, E.N., Bassel-Duby, R., 2013. Mending broken hearts: Cardiac development as a basis for adult heart regeneration and repair. *Nature Reviews Molecular Cell Biology* 14 (8), 529–541.
- Yin, H., Price, F., Rudnicki, M.A., 2013. Satellite cells and the muscle stem cell niche. *Physiological Reviews* 93 (1), 23–67.
- Zammit, P.S., 2008. All muscle satellite cells are equal, but are some more equal than others? *Journal of Cell Science* 121 (Pt 18), 2975–2982.
- Zammit, P.S., Partridge, T.A., Yablonka-Reuveni, Z., 2006. The skeletal muscle satellite cell: The stem cell that came in from the cold. *Journal of Histochemistry and Cytochemistry* 54 (11), 1177–1191.
- Zeisberg, E.M., Kalluri, R., 2010. Origins of cardiac fibroblasts. *Circulation Research* 107 (11), 1304–1312.
- Zhou, B., Honor, L.B., He, H., *et al.*, 2011. Adult mouse epicardium modulates myocardial injury by secreting paracrine factors. *Journal of Clinical Investigation* 121 (5), 1894–1904.
- Zhou, B., Ma, Q., Rajagopal, S., *et al.*, 2008. Epicardial progenitors contribute to the cardiomyocyte lineage in the developing heart. *Nature* 454 (7200), 109–113.

Cancer Stem Cells

S Buczacki, University of Cambridge, Cambridge, UK

© 2016 Elsevier Inc. All rights reserved.

Definitions and History

Stem cells are defined as cells that possess the cardinal features of multipotentiality, self-renewal, clonogenicity, and proliferation. It was originally proposed in the 1930s that within some tumors there exists a similar small subpopulation of tumor cells that is capable of regenerating the phenotypic heterogeneity of the original tumor in its entirety also displaying the four defining features of regular stem cells. This description has since become known as the cancer stem cell hypothesis. The early data that led to the cancer stem cell hypothesis being developed were based on the mouse models of leukemia (Furth, 1937). Prior to this it had been assumed that in order to transmit a malignancy from one animal to another many thousands of cells were required for transplantation. Furth hypothesized that these data were consequent on the genetic background of the mouse and that a more inbred mouse may require less cells to be transplanted. Using young inbred mice reestablishment of the original parental tumor from just a single cell was demonstrated. Interestingly, Furth also reported that not every injection of a single cell was successful concluding, somewhat prophetically, "It may be supposed, furthermore, that not all cells are capable of reproduction and that cells which do not reach an organ favorable for their growth perish."

It wasn't however until the work of Dominique Bonnet and John Dick in the 1990s that the cancer stem cell hypothesis was formally proven. In their seminal publication these authors demonstrated that fluorescence-activated cell sorting (FACS) of leukemic cells with antibodies that recognized the cell surface antigens CD34 and CD38 enriched for a population that could initiate leukemic proliferation when injected in immune compromised NOD/SCID (non-obese diabetic with severe combined immunodeficiency disease) mice (Bonnet and Dick, 1997). Using serial transplantation of tumors together with FACS analysis they were able to demonstrate that CD34⁺/CD38⁻ marked a population that contained the cancer stem cells of leukemia. Further, they also showed that a similar CD34⁺/CD38⁻ population appeared to also represent the cell-of-origin of leukemia. These early experiments paved the way for the identification of putative cancer stem cells in many other malignancies. Whilst xenotransplantation has historically formed a key assay in cancer stem cell biology, it is increasingly recognized that this approach has significant limitations and newer techniques are now being employed to validate historical xenotransplantation data.

Tumor Heterogeneity

It has become increasingly evident that in tumors resides both genetic heterogeneity and cellular functional heterogeneity. Traditionally there exists two descriptions explaining this heterogeneity; these have become known as the stochastic and

cancer stem cell models of tumor development. In the stochastic model the functional heterogeneity seen occurs as a result of random differentiation of an equipotent population of undifferentiated tumor cells. Layered on top of this stochastic differentiation exists clonal genetic evolution where distinct individual genetic clones may have the capacity for greater clonogenic potential. In the cancer stem cell model, however, there is a stricter hierarchical arrangement of cells within the tumor where only a certain subset (the cancer stem cells) retain capacity for self-renewal and proliferation. This latter explanation is more analogous to normal tissues where a structured hierarchical arrangement generally exists. Genetic clonal evolution theory can also be layered on top of the cancer stem cell model; each genetic clone possesses its own hierarchical organization with cancer stem cells at the apex. As clonal evolution develops and certain mutations are selected for, these new genetic clones can possess an increased number of cancer stem cells. Eventually, as tumor evolution continues a situation develops where almost all of the dominant genetic clone may essentially behave in a highly aggressive cancer stem cell-like manner (Figure 1).

Understanding and resolving functional and genetic heterogeneity in tumors to truly understand cancer stem cell biology will in the future require experiments that combine analysis of both of these facets concurrently. It is increasingly apparent that many of the conclusions made to date may be artefactual and assay dependent.

Terminology

Within the field the term 'cancer stem cell' is controversial and an alternative nomenclature – the tumor initiating cell (TIC) has also been used. The debate around this terminology has arisen for several reasons including the prevalence of the so-called cancer stem cells that have been found within some tumors, for example, in some melanomas they appear to make up the majority of the cells present within the tumor. Also impacting the use of the term cancer stem cell is the recently observed phenomenon of phenotypic plasticity within some tumor cell populations where it has been demonstrated that the stem state is not fixed and is both inducible and reversible. The question raised being is a cancer stem cell defined by being at that moment in time clonogenic or one that possesses that ability only in a context-dependent manner.

Recently the distinction between the terms cancer stem cell and TIC have been eloquently dissected (Kreso and Dick, 2014). Here a TIC is defined as a cell that is capable of tumor initiation through xenotransplantation, self-renewal being demonstrated by serial xenograft passage that also importantly generates daughter cells that may proliferate but are unable to reestablish the parental tumor on serial xenograft passage. A cancer stem cell is described as being distinct in that whilst possessing all the features of a TIC it can also be prospectively isolated.

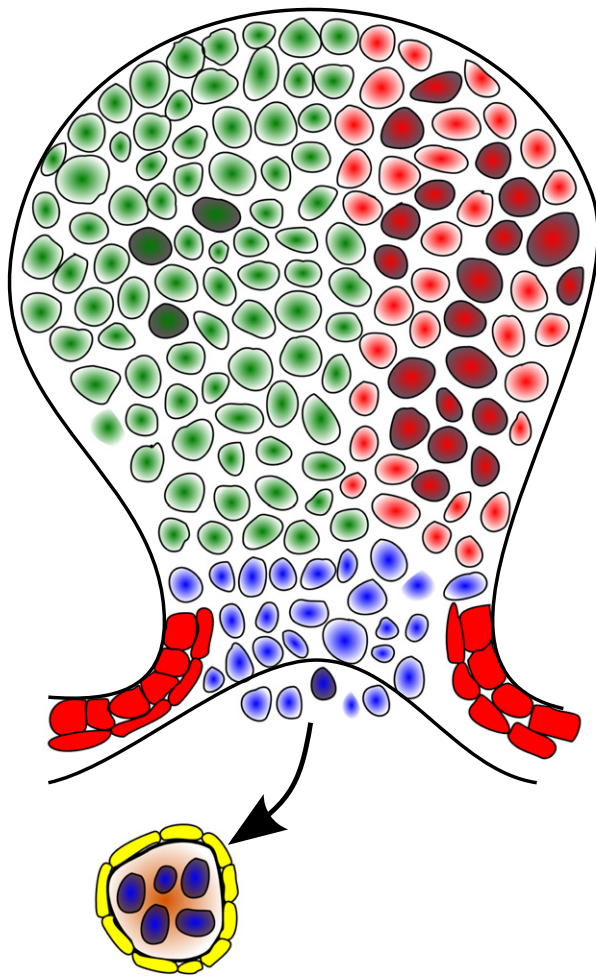


Figure 1 Schematic representing cellular and genetic tumoral heterogeneity. Although originating from a founder mutated cell, three different genetic clones have evolved and are represented in the tumour by the green, red, and blue colored cells. Within each clone exist variable quantities of cancer stem cells (darker colored cells). Whilst the blue genetic clone contains very few cancer stem cells, plasticity in cell fate has developed where a noncancer stem cell has metastasized to a separate location with an alternative niche (yellow cells). Here the noncancer stem cell has undergone a change in fate to a cancer stem cell and an aggressive metastasis has developed made up in entirety of cancer stem cells.

Modes of Identification

Much of the debate around the existence of cancer stem cells has arisen as a result of the assays used to define cancer stemness. There are several *in vitro* assays that have been used to identify putative cancer stem cell populations. Sphere-forming assays involve plating putative FACS-sorted populations into non-adherent culture conditions. Cells that have stem-like characteristics are believed to form free-floating spheres. This experimental technique has commonly been used to study both stem and cancer stem cell populations in many different tissues. More recently so-called organoid/spheroid *in vitro* assays have been developed that more accurately recreate the tissue architecture present in normal and malignant tissue.

These techniques differ from sphere-forming assays in that they are generally performed by plating cells in a biological extracellular matrix such as collagen or Matrigel™. Rather than forming a nondistinct sphere the matrix allows for the formation of a structure that retains many morphological and cellular features present in the native parental tissue. Using parental intestinal tissue, stem cell-derived organoids have been transplanted back and engrafted into recipient mice with defective intestinal epithelium reforming a functional epithelium (Yui *et al.*, 2012). These types of experiments pave the way for organoids to be used for regenerative medicine approaches.

Historically, the most widely used *in vivo* experimental technique to ascertain whether a cancer cell has clonogenic capacity is xeno/allotransplantation. This approach involves disaggregating a solid human or mouse tumor or a cancer cell line down to the single cell level. This single cell preparation is then FACS sorted using cell surface markers that are hypothesized to mark a cancer stem cell population to generate two populations (cancer stem cell and noncancer stem cell). The two populations are then transplanted using a serial dilution technique into immune compromised animal hosts. The subsequent proportion of tumors that grow and the tumor forming cell frequencies can then be calculated. Generated tumors are then excised, disaggregated again down to single cells and assayed using FACS to see whether the original heterogeneity of the primary has been recapitulated. Importantly the two populations once again undergo FACS sorting and transplantation to formally demonstrate self-renewal. Whilst the experimental technique is sound results can differ significantly based upon the genetic background of the host animal. For example in the situation of melanoma it has been shown that changing the host from a NOD/SCID to NSG (NOD/SCID gamma) raises the frequency of cancer stem cells from 1 in a million to 1 in 4 (Quintana *et al.*, 2010).

One major concern with xenotransplantation and the *in vitro* assays described is they involve taking the cellular populations of interest away from their native environment. It is increasingly apparent that the ‘niche’ is of importance in regulating stem and cancer stemness and further that tumor cells can transition from different states of stemness potential in response to niche signaling (see later). Recent studies using more contemporary experimental approaches such as lineage tracing have shown that some populations identified using traditional approaches as stem or cancer stem-like *ex-vivo*, are, when analyzed *in vivo* using lineage tracing in possession of relatively little clonogenicity (Buczacki *et al.*, 2013). Lineage tracing has become the gold-standard *in vivo* assay for defining stem cell populations in normal tissue (Kretzschmar and Watt, 2012; Blanpain and Simons, 2013). This transgenic strategy generally makes use of a tamoxifen-driven *Cre-Lox* system to induce the permanent expression of a visible reporter whose transcription is driven from the promoter of the gene of interest. If this genetic event occurs within a stem cell then all offspring will also express the reporter generating a long-lived visible clone. If, however, recombination occurs in a transit-amplifying or differentiated cell, then the clone will either be nonexistent or short-lived. This approach has been used successfully in many normal tissues to convincingly identify *bona fide* stem cell populations. Importantly, unlike other techniques the assay is generally accepted as not altering the

niche environment surrounding the targeted population of interest ensuring the results more accurately represent the homeostatic picture. The application of the assay to tumor tissue, however, has been technically difficult, not in the least related to reconstructing clonal hierarchies in a disordered solid structure. Nevertheless, these obstacles have recently been overcome and the first descriptions of cancer stem cells in tumors using lineage tracing have recently been presented in skin papilloma, intestinal adenoma, and glioma (Driessens *et al.*, 2012; Schepers *et al.*, 2012; Chen *et al.*, 2012).

Although generally regarded as not affecting homeostasis, it is worth noting two interesting reports in breast and intestinal tissue that have suggested the commonly used inducing agent in many *Cre-Lox* systems, Tamoxifen, could have unexpected effects on both apoptosis and proliferation in different cell types which could skew subsequent results (Shehata *et al.*, 2014; Zhu *et al.*, 2013). Furthermore, the application of lineage tracing to human tissue is also not straightforward given that it is a transgenic strategy. Indirectly, however, the concept has been used with success in humans by using mutations in mitochondrial DNA to indirectly quantify cancer stem cell dynamics (Humphries *et al.*, 2013). This type of indirect clonal quantification requires significant mathematical involvement as the timing of the recombination event is unknown unlike in *Cre-Lox* systems where it is controlled by the initial Tamoxifen treatment.

Regulation of Cancer Stemness

Although not completely clear it appears that what defines and regulates whether a cell is a cancer stem cell or not may be regulated by the so-called niche. The niche is defined by the cells that are found in close proximity to the tumor cell of interest. These may be epithelial, mesenchymal, infiltrating immunological, or endothelial in origin. The niche can also be altered by more extrinsic factors such as hypoxia. It has been hypothesized that paracrine signaling from surrounding niche cells as well as associated autocrine signaling may determine whether a tumor cell has cancer stem-like properties. This phenomenon has been shown to exist in normal tissues. For example, in the intestine it has been shown that terminally differentiated Paneth cells present in the base of the crypts of Lieberkuhn in the small intestine, actively secrete Wnt proteins (Sato *et al.*, 2011). These same Wnt signals activate neighboring intercalated crypt base columnar cells to achieve the stem-like state. Genetic ablation of Paneth cells subsequently leads to a loss in intestinal stem cell numbers.

In tumors similar observations have also been seen. In glioblastoma it has been demonstrated that xenotransplantation graft rates of Nestin-positive cancer stem cells was higher when endothelial cells from the parental tumor were co-injected (Calabrese *et al.*, 2007). Similarly in head and neck squamous carcinoma, cancer stem cells (ALDH⁺CD44⁺) when exposed to endothelial cell conditioned medium containing endothelial cell-derived growth factors, displayed marked increased proliferation, survival, and self-renewal (Krishnamurthy *et al.*, 2010). In colorectal cancer it has been demonstrated that myofibroblast secreted hepatocyte growth factor (HGF) also results in elevated Wnt signaling and a cancer stem cell-like phenotype (Vermeulen *et al.*, 2010).

It has also been proposed that hypoxia is an important determinant in regulating cancer stemness either directly or indirectly through induction of hypoxia-inducible factor (HIF). CD133⁺ glioma cancer stem cells grown *in vitro* under hypoxic conditions display increased self-renewal and proliferation and inhibition of differentiation through HIF-1alpha expression (Soeda *et al.*, 2009). HIF-1alpha also induces expression of the pro-angiogenic factors VEGF and VEGFR2 required for tumor metastasis (Kuwai *et al.*, 2003). Similarly HIF-1alpha has also been shown to induce the EMT (epithelial to mesenchymal transition) phenotype associated with metastasis and cancer stemness (Gammon *et al.*, 2013). HIF-2alpha also identifies immature (stem-like) cells in breast and neuroblastoma cancers (Pietras *et al.*, 2009). HIF-2alpha effects have been suggested to be governed by downstream activation of the pluripotency gene Oct4 (Covello *et al.*, 2006).

Plasticity

It is increasingly apparent that stemness and in all likelihood cancer stemness is not 'hard-wired' into individual cells. One of the best-known and earliest descriptions of cellular plasticity can be seen in urodeles such as the Salamander. Here, in response to injury a progenitor population (blastema) is generated from apparently terminally differentiated cells. This progenitor population is capable of significant regeneration including complete new limbs. Plasticity on a lesser scale is also evident in mammalian cell populations. This has been demonstrated indirectly in both the hematopoietic and breast systems. More directly using lineage tracing, plasticity in the fate of intestinal secretory progenitor populations reverting to a stem-like state has also been described (Van Es *et al.*, 2012; Buczacki *et al.*, 2013). Plasticity also appears to function within cancer stem cell populations. This can occur remarkably early in the tumorigenesis. Schwitalla recently demonstrated in the intestine that the cell-of-origin of intestinal tumors can be a terminally differentiated enterocyte that can reacquire a stem-like state in response to enhanced NF- κ B expression (Schwitalla *et al.*, 2013). Similar observations have been seen in glioma (Charles *et al.*, 2010). Plasticity may also occur deeper into tumor development. A widely cited example of this occurs in melanoma where cells are capable of reversibly switching on and off the histone demethylase JARID1B (Roesch *et al.*, 2010), JARID1B expressing cells being the population responsible for tumor maintenance.

Truly understanding cellular plasticity in both normal and tumor tissue is now possible using lineage tracing. Indeed lineage tracing has shown that many of the conclusions using historical *in vitro* and *in vivo* assays such as xenotransplantation and organoid growth are not entirely reliable in identifying true clonogenic function rather than potentiality. It appears that many of these assays in reality quantify the latter rather than the cells' true *in vivo* behavior.

Markers for the Identification of Cancer Stem Cells in Major Malignancies

Cancer stem cells have now been identified in most solid organ malignancies. The majority of these reports have made

use of FACS sorting and xenotransplantation of putative cancer stem cell populations using cell surface markers. There is significant overlap in these markers between solid tumors implying conserved pathways and functionality (Buczacki *et al.*, 2011).

Leukemia – Leukemic cancer stem cells were amongst the first cancer stem cells to be described. Initially described in acute myeloid leukemia (AML) by Bonnet and Dick they are characterized by expression of CD34⁺ and CD38[−]. Since this initial discovery the use of mice strains with higher degrees of immune-deficiency such as NOD/SCID/ $\beta 2m^{-/-}$ or NOD/SCID/IL2R $\gamma^{-/-}$ have enabled increased engraftment efficiencies of CD34⁺/CD38[−] Lin[−] cancer stem cell populations as well as progenitor CD34⁺/CD38⁺ Lin[−] populations (Taussig *et al.*, 2008). These and other experiments demonstrate heterogeneity in the leukemic stem cell population.

Breast – Breast cancer stem cells were initially described based on the finding that FACS sorting CD44⁺/CD24^{−/low} Lin[−] cells enriched for a cell population which had significantly higher engraftment rates than alternatively FACS-sorted populations (Al-Hajj *et al.*, 2003). Aldehyde dehydrogenase (ALDH) has also been described as a marker that enriches for breast cancer stem cells (Ginestier *et al.*, 2007). Additionally, CD133 (Prominin-1) and CD29 have been proposed as breast cancer stem cell markers (Wright *et al.*, 2008; Vassilopoulos *et al.*, 2008).

Colorectal – Colorectal cancer stem cells have been proposed to be identified by the following markers: CD133 (Ricci-Vitiani *et al.*, 2007; O'Brien *et al.*, 2007), CD44 (Dalerba *et al.*, 2007), CD24 (Vermeulen *et al.*, 2008), CD166 (Dalerba *et al.*, 2007), Dcl1 (Nakanishi *et al.*, 2013), Lgr5 (Schepers *et al.*, 2012), and Olfm4 (Van Der Flier *et al.*, 2009). Recently two of these, Lgr5 and Dcl1, have been shown using lineage tracing to mark a cancer stem cell population in murine adenomas. These two markers are believed to have significant overlap in expression.

Brain – Neuronal cancer stem cells have been reported to be marked by CD133 (Singh *et al.*, 2004), Nestin (Chen *et al.*, 2012) as well as Bmi1 (Bruggeman *et al.*, 2007). Only Nestin, however, has been confirmed using lineage tracing to specifically mark clonogenic tumor cells. This recent publication showed that a relatively quiescent Nestin-expressing population of cancer stem cells becomes activated after temozolomide chemotherapy thus potentially accounting for disease recurrence after adjuvant therapy.

Ovarian – Whilst several markers including CD24 (Gao *et al.*, 2010), CD133 (Ferrandina *et al.*, 2009), and ALDH (Kryczek *et al.*, 2012) have been used to identify cancer stem cell-like populations in ovarian cancer, the cell-of-origin for these tumors has remained contested; some proposing that the tumor arises within the distal end of the fallopian tube rather than the ovary *per se*. Using lineage tracing it has been demonstrated that in some ovarian tumors this founder cancer cell appears to originate from the ovarian surface epithelium and is marked by expression of Lgr5 and ALDH (Flesken-Nikitin *et al.*, 2013). Whether this is the case for all ovarian tumors remains to be determined.

Skin – The only other system where putative cancer stem cells have been identified using lineage tracing is in the skin (Driessens *et al.*, 2012). Using mathematical analysis of clone

formation driven by a non-cell type specific promoter, in chemically induced skin papillomas the authors found two patterns of division. One of these modes of division generated long-lived clones suggesting stem cell-like behavior. It should be noted, however, that this study was performed in papilloma which whilst being neoplastic is not truly invasive.

Others – Cancer stem cells have also been described in tumors of the lung, prostate, pancreas, and liver using markers such as ALDH (Liang and Shi, 2012), CD133 (Collins *et al.*, 2005), CD44 (Yao *et al.*, 2010), and CD24 (Yao *et al.*, 2010).

Although cell surface markers have historically been of tremendous value in FACS/xenografting experiments it is important to note that many of these markers do not uniquely identify cancer stem cell populations. Indeed many of these markers such as CD24 are signaling pathway target genes and thus reflect a population where a specific signaling pathway is active. This results in the population sorted using a cell surface marker such as CD24 being relatively heterogeneous encompassing both cancer stem cell populations as well as more differentiated progeny. Furthermore, the use of cell surface markers alone will not enable the characterization of plasticity within cancer stem cell populations where a more differentiated cell type may not be marked by a specific marker but may, however, still retain stem cell potency.

Clinical Significance of Cancer Stem Cells

Over the last decades the concept of cancer stem cells has developed from being a hypothesized population to a real definable entity. What is still unclear is what provides a cancer stem cell its unique identity and how prevalent they are throughout tumor development. Indeed, in light of context-dependent plasticity the true prevalence of a cancer stem cell population at any one point in time may be irrelevant as this will be a dynamic and ever changing figure. Importantly, this dictates that classical clinical modes of detecting tumor types such as immunohistochemistry (IHC) will be inappropriate for quantifying cancer stem cell prevalence in this context. Nevertheless, the presence of this unique cell population does provide an important clinical target for therapeutic targeting. In addition to their clonogenic capacity cancer stem cells display many attributes that predispose to drug resistance such as enhanced DNA repair pathways, cellular dormancy, and the expression of multidrug resistance proteins. These faculties suggest that targeting this population may be problematic. Indeed many authors suggest that these attributes may suggest that cancer stem cells are the population responsible for disease recurrence after adjuvant therapy. Late systemic recurrence is not uncommon after apparent complete response. Clearly, for this to occur requires a small, drug resistant and highly clonogenic population to exist – just as cancer stem cells appear to be. It has been proposed however that given many cancer stem cell populations share a large number of molecular features with regular stem cells; direct targeting of these cells could lead to significant deleterious effects on the associated normal tissue homeostasis. In the context of the gut this question has recently been answered in reference to the nature of doublecortin-like kinase 1 (Dcl1) expressing cells. Early reports had suggested that Dcl1 marked a putative quiescent

intestinal stem cell population; however, others had argued that Dclk1 marked a rare terminally differentiated secretory cell type called a Tuft cell (May *et al.*, 2009; Gerbe *et al.*, 2009). Dclk1 expression has also been described in both murine and human intestinal tumors. Using a lineage tracing technique it has now convincingly been shown that in the normal intestine during homeostasis Dclk1 expressing cells are indeed terminally differentiated. However, during dextran sulfate sodium (DSS) treatment or radiation-induced epithelial injury a very small proportion of these cells can acquire stem cell capacity characterized by the ability to generate long-lived clones (Nakanishi *et al.*, 2013). Interestingly when cancer stem cell capacity is quantified in tumor bearing Apc^{Min} mice, Dclk1 expressing cells also generate many clones implying a strong cancer stem cell-like phenotype. Intriguingly, targeting these Dclk1 expressing cells specifically using transgenic expression of diphtheria toxin receptor and then diphtheria toxin treatment leads to tumor regression without any deleterious effect on the associated normal intestinal epithelium. These data contrast with many other reports in showing widely differing functions of marker defined cells dependent on an oncogenic/normal tissue context.

Using knowledge about the molecular mechanisms underpinning regular stem cell function it has been shown that disrupting similar modes of self-renewal in cancer could lead to therapeutic advantage (Kreso *et al.*, 2014). Bmi1 (part of the polycomb repressive complex 1) appears to mark stem cells in many tissues. Interestingly, both siRNA and small molecule knockdown of Bmi1 leads to diminishment in tumor growth both *in vitro* and in xenograft experiments using colorectal cancer tissue without significant effect on the overall health of the animal. These data intriguingly draw together a potential cancer stem cell marker which additionally has a significant functional role.

Alternative approaches using differentiation therapy have also been suggested. These rely on a more generalized treatment that causes undifferentiated cancer stem cells to undergo a differentiation program. This approach has been used for many years for patients suffering with acute promyelocytic leukemia using a drug called all-*trans* retinoic acid (ATRA). The approach has also been experimentally described in the intestine using BMP4 treatment to differentiate cancer stem cells in colorectal cancer xenografts where in addition to promoting apoptosis and differentiation the treatment also chemosensitized the tumors (Lombardo *et al.*, 2011).

Conclusions

Improved experimental techniques, first used in normal stem cell biology, such as lineage tracing have enabled more detailed and accurate descriptions of cancer stem cell behavior. Nevertheless, it is only very recently that combined genetic and functional approaches are being used to understand the behavior of these cells. These approaches are essential to better correlate genetic makeup with phenotypic function. Clearly now cancer stem cells are recognized as a real entity they are also an important clinical target whose eradication will improve the efficacies of adjuvant therapy. In the future, indirect approaches such as differentiation therapy and niche

manipulation will be as important as direct targeting in eliminating this unique cell population.

See also: Stem Cell Biology: Structure and Function: Pluripotent Cells: New Tools for Disease Research and Therapies

References

- Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., Clarke, M.F., 2003. Prospective identification of tumorigenic breast cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* 100, 3983–3988.
- Blanpain, C., Simons, B.D., 2013. Unravelling stem cell dynamics by lineage tracing. *Nature Reviews Molecular Cell Biology* 14, 489–502.
- Bonnet, D., Dick, J.E., 1997. Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature Medicine* 3, 730–737.
- Bruggeman, S.W., Hulsman, D., Tanger, E., *et al.*, 2007. Bmi1 controls tumor development in an Ink4a/Arf-independent manner in a mouse model for glioma. *Cancer Cell* 12, 328–341.
- Buczacki, S., Davies, R.J., Winton, D.J., 2011. Stem cells, quiescence and rectal carcinoma: An unexplored relationship and potential therapeutic target. *British Journal of Cancer* 105, 1253–1259.
- Buczacki, S.J., Zecchini, H.I., Nicholson, A.M., *et al.*, 2013. Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* 495, 65–69.
- Calabrese, C., Poppleton, H., Kocak, M., *et al.*, 2007. A perivascular niche for brain tumor stem cells. *Cancer Cell* 11, 69–82.
- Charles, N., Ozawa, T., Squatrito, M., *et al.*, 2010. Perivascular nitric oxide activates notch signaling and promotes stem-like character in PDGF-induced glioma cells. *Cell Stem Cell* 6, 141–152.
- Chen, J., Li, Y., Yu, T.S., *et al.*, 2012. A restricted cell population propagates glioblastoma growth after chemotherapy. *Nature* 488, 522–526.
- Collins, A.T., Berry, P.A., Hyde, C., Stower, M.J., Maitland, N.J., 2005. Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Research* 65, 10946–10951.
- Covello, K.L., Kehler, J., Yu, H., *et al.*, 2006. HIF-2alpha regulates Oct-4: Effects of hypoxia on stem cell function, embryonic development, and tumor growth. *Genes & Development* 20, 557–570.
- Dalerba, P., Dylla, S.J., Park, I.K., *et al.*, 2007. Phenotypic characterization of human colorectal cancer stem cells. *Proceedings of the National Academy of Sciences of the United States of America* 104, 10158–10163.
- Driessens, G., Beck, B., Caauwe, A., Simons, B.D., Blanpain, C., 2012. Defining the mode of tumour growth by clonal analysis. *Nature* 488, 527–530.
- Ferrandina, G., Martinelli, E., Petrillo, M., *et al.*, 2009. CD133 antigen expression in ovarian cancer. *BMC Cancer* 9, 221.
- Flesken-Nikitin, A., Hwang, C.I., Cheng, C.Y., *et al.*, 2013. Ovarian surface epithelium at the junction area contains a cancer-prone stem cell niche. *Nature* 495, 241–245.
- Furth, J.K.M., 1937. The transmission of leukaemia of mice with a single cell. *American Journal of Cancer* 31, 276–282.
- Gammon, L., Biddle, A., Heywood, H.K., Johannessen, A.C., Mackenzie, I.C., 2013. Sub-sets of cancer stem cells differ intrinsically in their patterns of oxygen metabolism. *PLoS One* 8, e62493.
- Gao, M.Q., Choi, Y.P., Kang, S., Youn, J.H., Cho, N.H., 2010. CD24+ cells from hierarchically organized ovarian cancer are enriched in cancer stem cells. *Oncogene* 29, 2672–2680.
- Gerbe, F., Brulin, B., Makrini, L., Legraverend, C., Jay, P., 2009. DCAMKL-1 expression identifies Tuft cells rather than stem cells in the adult mouse intestinal epithelium. *Gastroenterology* 137, 2179–2180.
- Ginestier, C., Hur, M.H., Charafe-Jauffret, E., *et al.*, 2007. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell* 1, 555–567.
- Humphries, A., Cereser, B., Gay, L.J., *et al.*, 2013. Lineage tracing reveals multipotent stem cells maintain human adenomas and the pattern of clonal expansion in tumor evolution. *Proceedings of the National Academy of Sciences of the United States of America* 110, E2490–E2499.

- Kreso, A., Dick, J.E., 2014. Evolution of the cancer stem cell model. *Cell Stem Cell* 14, 275–291.
- Kreso, A., Van Galen, P., Pedley, N.M., *et al.*, 2014. Self-renewal as a therapeutic target in human colorectal cancer. *Nature Medicine* 20, 29–36.
- Kretschmar, K., Watt, F.M., 2012. Lineage tracing. *Cell* 148, 33–45.
- Krishnamurthy, S., Dong, Z., Vodopyanov, D., *et al.*, 2010. Endothelial cell-initiated signaling promotes the survival and self-renewal of cancer stem cells. *Cancer Research* 70, 9969–9978.
- Kryczek, I., Liu, S., Roh, M., *et al.*, 2012. Expression of aldehyde dehydrogenase and CD133 defines ovarian cancer stem cells. *International Journal of Cancer* 130, 29–39.
- Kuwai, T., Kitadai, Y., Tanaka, S., *et al.*, 2003. Expression of hypoxia-inducible factor-1 α is associated with tumor vascularization in human colorectal carcinoma. *International Journal of Cancer* 105, 176–181.
- Liang, D., Shi, Y., 2012. Aldehyde dehydrogenase-1 is a specific marker for stem cells in human lung adenocarcinoma. *Medical Oncology* 29, 633–639.
- Lombardo, Y., Scopelliti, A., Cammareri, P., *et al.*, 2011. Bone morphogenetic protein 4 induces differentiation of colorectal cancer stem cells and increases their response to chemotherapy in mice. *Gastroenterology* 140, 297–309.
- May, R., Sureban, S.M., Hoang, N., *et al.*, 2009. Doublecortin and CaM kinase-like-1 and leucine-rich-repeat-containing G-protein-coupled receptor mark quiescent and cycling intestinal stem cells, respectively. *Stem Cells* 27, 2571–2579.
- Nakanishi, Y., Seno, H., Fukuoka, A., *et al.*, 2013. Dcl1 distinguishes between tumor and normal stem cells in the intestine. *Nature Genetics* 45, 98–103.
- O'Brien, C.A., Pollett, A., Gallinger, S., Dick, J.E., 2007. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* 445, 106–110.
- Pietras, A., Hansford, L.M., Johnsson, A.S., *et al.*, 2009. HIF-2 α maintains an undifferentiated state in neural crest-like human neuroblastoma tumor-initiating cells. *Proceedings of the National Academy of Sciences of the United States of America* 106, 16805–16810.
- Quintana, E., Shackleton, M., Foster, H.R., *et al.*, 2010. Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. *Cancer Cell* 18, 510–523.
- Ricci-Vitiani, L., Lombardi, D.G., Pilozzi, E., *et al.*, 2007. Identification and expansion of human colon-cancer-initiating cells. *Nature* 445, 111–115.
- Roesch, A., Fukunaga-Kalabis, M., Schmidt, E.C., *et al.*, 2010. A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell* 141, 583–594.
- Sato, T., Van Es, J.H., Snippert, H.J., *et al.*, 2011. Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts. *Nature* 469, 415–418.
- Schepers, A.G., Snippert, H.J., Stange, D.E., *et al.*, 2012. Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas. *Science* 337, 730–735.
- Schwitala, S., Fingerle, A.A., Cammareri, P., *et al.*, 2013. Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell* 152, 25–38.
- Shehata, M., Van Amerongen, R., Zeeman, A.L., Giraddi, R.R., Stingl, J., 2014. The influence of tamoxifen on normal mouse mammary gland homeostasis. *Breast Cancer Research* 16, 411.
- Singh, S.K., Hawkins, C., Clarke, I.D., *et al.*, 2004. Identification of human brain tumour initiating cells. *Nature* 429, 396–401.
- Soeda, A., Park, M., Lee, D., *et al.*, 2009. Hypoxia promotes expansion of the CD133-positive glioma stem cells through activation of HIF-1 α . *Oncogene* 28, 3949–3959.
- Taussig, D.C., Miraki-Moud, F., Anjos-Afonso, F., *et al.*, 2008. Anti-CD38 antibody-mediated clearance of human repopulating cells masks the heterogeneity of leukemia-initiating cells. *Blood* 112, 568–575.
- Van Der Flier, L.G., Haegbarth, A., Stange, D.E., Van De Wetering, M., Clevers, H., 2009. OLFM4 is a robust marker for stem cells in human intestine and marks a subset of colorectal cancer cells. *Gastroenterology* 137, 15–17.
- Van Es, J.H., Sato, T., Van De Wetering, M., *et al.*, 2012. Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. *Nature Cell Biology* 14, 1099–1104.
- Vassilopoulos, A., Wang, R.H., Petrovas, C., *et al.*, 2008. Identification and characterization of cancer initiating cells from BRCA1 related mammary tumors using markers for normal mammary stem cells. *International Journal of Biological Sciences* 4, 133–142.
- Vermeulen, L., De Sousa, E.M.F., Van Der Heijden, M., *et al.*, 2010. Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. *Nature Cell Biology* 12, 468–476.
- Vermeulen, L., Todaro, M., De Sousa Mello, F., *et al.*, 2008. Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. *Proceedings of the National Academy of Sciences of the United States of America* 105, 13427–13432.
- Wright, M.H., Calcagno, A.M., Salcido, C.D., *et al.*, 2008. Brca1 breast tumors contain distinct CD44+ /CD24- and CD133+ cells with cancer stem cell characteristics. *Breast Cancer Research* 10, R10.
- Yao, J., Cai, H.H., Wei, J.S., *et al.*, 2010. Side population in the pancreatic cancer cell lines SW1990 and CFPAC-1 is enriched with cancer stem-like cells. *Oncology Reports* 23, 1375–1382.
- Yui, S., Nakamura, T., Sato, T., *et al.*, 2012. Functional engraftment of colon epithelium expanded *in vitro* from a single adult Lgr5(+) stem cell. *Nature Medicine* 18, 618–623.
- Zhu, Y., Huang, Y.F., Kek, C., Bulavin, D.V., 2013. Apoptosis differently affects lineage tracing of Lgr5 and Bmi1 intestinal stem cell populations. *Cell Stem Cell* 12, 298–303.

Metabolic and Energetic Regulation of Stem Cells

A Esteves and T Finkel, National Institutes of Health, Bethesda, MD, USA

© 2016 Elsevier Inc. All rights reserved.

The unique properties of stem cells, namely their ability to self-renew, to persist throughout the life of the organism, and to exist for a long period of time in a quiescent state, suggest that this cell type might have unique metabolic constraints. Because of their clinical importance and their experimental tractability, much of our knowledge regarding stem cell metabolism comes from the hematopoietic stem cell (HSC) system. The earliest foray into this area was precipitated by attempts to understand the role of oxygen in the biology of HSCs. Like all adult stem cells, HSCs have the capacity to give rise to multipotent progenitors, or to produce new stem cells (Figure 1). The latter process is termed self-renewal and experimentally this activity in HSCs can be assessed by the ability of bone marrow cells to repopulate lethally irradiated host animals. Early experiments demonstrated that the repopulating ability of HSCs was improved when cells were maintained under low-oxygen conditions (Cipolleschi *et al.*, 1993; Danet *et al.*, 2003). This was followed by various theoretical calculations that posited that the oxygen gradient in the bone marrow was such that long-term HSCs most likely resided in near hypoxic conditions (Chow *et al.*, 2001). These mathematical models were subsequently solidified by experimental evidence. Initial efforts were hampered by the inability to do true *in vivo* analysis. Nonetheless, using the weak base pimonidazole that acts as a functional and immunohistochemical marker of hypoxia, it was demonstrated that within all the diverse cell types of the bone marrow, HSCs appear to be the most pimonidazole-positive (Parmar *et al.*, 2007). Subsequent quantitative *in vivo* imaging confirmed many of these earlier observations (Nombela-Arrieta *et al.*, 2013).

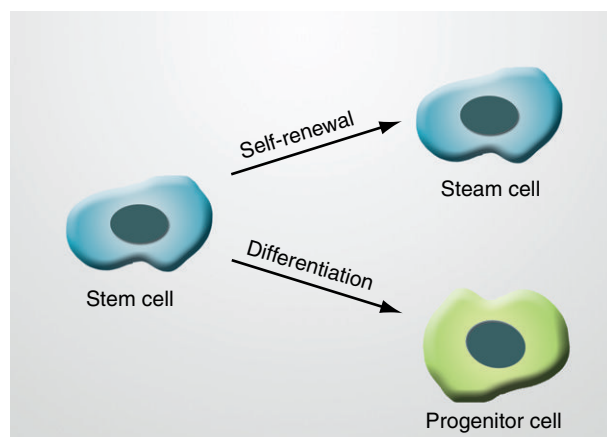


Figure 1 General biology of adult stem cells. Stem cells are unique in their capacity to divide and give rise to two different types of daughter cells. This asymmetric cell division includes giving rise to a multipotent progenitor cell, as well as to a new stem cell. The latter process is termed self-renewal. For the case of HSCs, multipotent progenitor cells can give rise to all blood forming cells, including red cells, white cells, and platelets.

These experiments, most recently using 3D reconstitution of the bone marrow niche, revealed that the microvessels surrounding HSCs appear to represent a unique vasculature configuration. Interestingly, HSCs appeared to be pimonidazole-positive irrespective of their distance from the vasculature. This argues that the 'hypoxic' nature of bone marrow stem cells may not relate solely to the distance of the HSC from a blood vessel, but rather may reflect an intrinsic quality of the stem cell.

Hypoxia-Inducible Factor 1 α and Glycolysis in Stem Cells

Given the relative rarity of HSCs within the bone marrow, it was initially difficult to perform classical metabolomics analysis on this rare cell type. Nonetheless, some early certain key and informative measurements were made. Initial experiments demonstrated that when compared to total bone marrow, the oxygen consumption and ATP levels were lower in bone marrow fractions enriched for HSCs (Simsek *et al.*, 2010). Furthermore, when glycolytic flux was determined by measuring the amount of glucose-derived ^{13}C -lactate production, HSCs appeared to have increased rates of glycolysis. This suggested that HSCs may limit normal mitochondrial metabolism and in ways perhaps similar to tumor cells, adapt a metabolic profile that relies on the cytosolic metabolism of glucose. Indeed, using mitochondrial membrane potential (ψ_m) as a marker, it was shown that HSCs exhibit constitutively low ψ_m , consistent with a program of de-energize and inactive mitochondria (Simsek *et al.*, 2010). This might be a common property of stem cells, as earlier experiments in embryonic stem cells reached similar conclusions (Schieke *et al.*, 2008).

Given the relative hypoxic environment and their reliance on glycolysis for energy, it seemed reasonable that the hypoxia-inducible factor 1 α (HIF-1 α) transcription factor might play a central role in the biology of HSCs. HIF-1 α is the oxygen-sensitive component of the heterodimeric transcription factor HIF-1. HIF-1 serves as the central coordinator of the cellular and systemic response to a low-oxygen environment (Wang and Semenza, 1995). Classically, HIF-1 regulates the expression of a host of enzymes involved in glycolysis, including aldolase A, enolase 1, lactate dehydrogenase A, phosphofructokinase 1, and phosphoglycerate kinase 1 among others (Iyer *et al.*, 1998). This adaptation allows the cell to produce ATP in the setting of a low-oxygen environment. Levels of HIF-1 message and protein are indeed elevated in HSCs (Simsek *et al.*, 2010; Takubo *et al.*, 2010). Furthermore, conditional bone marrow deletion of HIF-1 α results in an increased fraction of HSCs emerging from quiescence (Takubo *et al.*, 2010). Gain-of-function studies, leading to increase HIF-1 α levels, revealed that this genetic manipulation led to an increase in the fraction of cells in the quiescent G0 phase of the

cell cycle (Takubo *et al.*, 2010). Thus, the quiescent state of HSCs is intimately connected to the metabolic state of these cells. Interestingly, cells without HIF-1 α demonstrate impaired stress resistance as evidenced by their inability to contribute to hematopoiesis after serial transplantation or during normal aging. Similarly, while cells with too much HIF-1 α exhibited impaired functional properties after transplantation, suggesting that the level of HIF-1 α needs to be tightly regulated within stem cells in order to maintain optimal function (Takubo *et al.*, 2010).

Recent data have solidified the connection between metabolism and stemness (i.e., the ability to both self-renew and give rise to committed progenitor cells) within the HSC population. In particular, a recent study demonstrated an essential role for the family of glycolytic regulatory enzymes known as pyruvate dehydrogenase kinase (Pdk) (Takubo *et al.*, 2013). The activity of the mitochondrial enzyme pyruvate dehydrogenase (PDH) is essential for the critical conversion of pyruvate to acetyl-CoA, the entry step for glucose into the tricarboxylic acid (TCA) cycle (Figure 2). The activity of PDH is regulated at many levels, including the ability to be inhibited following phosphorylation by the Pdk. Using loss- and gain-of-function models, this recent study has again demonstrated a role for glycolytic metabolism in maintaining HSC quiescence and functional capacity. Using a significant number of mice ($n=120$) the authors were able to purify approximately one million stem and progenitor cells and then perform a metabolomics assessment. Compared to more mature bone marrow-derived cells, HSCs and their immediate descendants demonstrated elevated levels of fructose 1,6-bisphosphate, a known intermediary of glucose metabolism, while levels of TCA intermediates were reduced (Takubo *et al.*, 2013). These metabolic data are once again consistent with a

pattern of using glycolysis, rather than mitochondrial metabolism to fuel the resting HSC energy demands. It was further shown that HIF-1 α regulates the transcript levels of Pdk2 and Pdk4 in HSCs. Mice deficient in Pdk2 and Pdk4 (Pdk2^{-/-}Pdk4^{-/-} mice) had a fall in steady state pyruvate levels consistent with a role for the Pdk in normally restraining mitochondrial activity. This increase in mitochondrial metabolism (and corresponding reduction in glycolysis) was accompanied by a shift from G0 to G1 in the HSC population. Following transplantation, Pdk2^{-/-}Pdk4^{-/-} HSCs demonstrated impaired self-renewal capacity. This was particularly evident after a second round of transplantation (Takubo *et al.*, 2013). Thus, once again, the ability to maintain glycolytic flux appears critical to maintain HSC function.

Fatty Acid Metabolism and Mitochondrial Function

While glycolytic flux appears to play a critical role in HSC self-renewal, glucose metabolism is not the sole determinant of HSC function. Recent evidence also suggests that fatty acid oxidation (FAO) can also regulate HSC function (Ito *et al.*, 2012). Again, using a mouse model, investigators were able to interrupt mitochondrial fatty acid metabolism either genetically (through the loss of peroxisome proliferator-activated receptor δ) or pharmacologically (through the action of the FAO inhibitor, etomoxir). Interestingly, while the preceding discussion has emphasized the apparent reliance on cytosolic glycolysis, bone marrow cells enriched for HSCs appear to be capable of undergoing appreciable rates of FAO (Ito *et al.*, 2012). Inhibition of FAO impaired donor-derived reconstitution after transplantation consistent with a role for fatty acid metabolism in self-renewal. Interestingly, using a single cell assay the authors were able to assess symmetrical versus asymmetrical cell division. Disrupting FAO led to an increase in frequency of stem cells undergoing symmetric commitment (i.e., given rise to two committed daughter cells), a shift again consistent with the observed defect in self-renewal. This *in vitro* assay suggests that FAO is involved in the poorly understood decision as to whether stem cells self-renew or give rise to committed progenitors.

Another recent link between mitochondrial function and HSC biology come from the analysis of cells deficient in the expression of the mitochondrial phosphatase PTPMT1 (Yu *et al.*, 2013). This inner mitochondrial protein is a phosphatase and tensin homolog (PTEN)-like phosphatase that metabolizes phosphatidylinositol phosphatases (PIPs). PIPs are a general class of lipid mediators that regulate a wide array of cellular functions from trafficking to channel activity. While relatively little is known about the role of PIPs in the mitochondria, it is known that PTPMT1 deletion results in early embryonic lethality (Shen *et al.*, 2011; Zhang *et al.*, 2011). Using a conditional knockout approach, it was later demonstrated that hematopoietic-specific PTPMT1 deletion resulted in pancytopenia and rapid postnatal death (Yu *et al.*, 2013). While the levels of differentiated cells were markedly reduced (consistent with the pancytopenia), analysis of the bone marrow of conditional PTPMT1 deleted mice revealed that the number of HSCs were surprisingly elevated (by approximately 40-fold). Cell cycle analysis revealed that PTPMT1-deficient

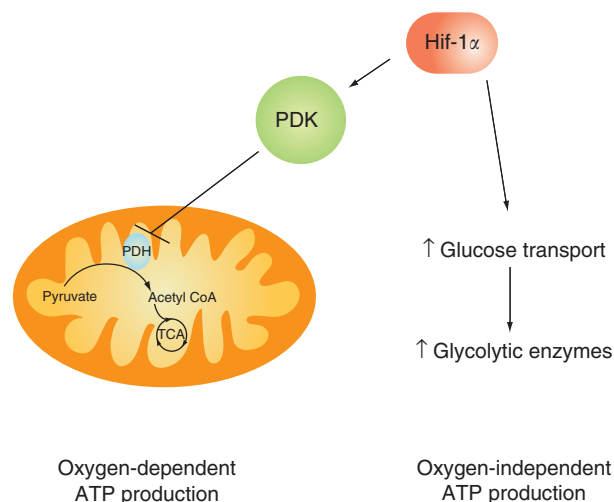


Figure 2 A central role for HIF-1 α in regulating HSC biology. The transcription factor HIF-1 α plays a central role in regulating the metabolic state of HSCs. This occurs through multiple mechanisms including increasing glycolytic flux and increase in glucose uptake and a transcriptional increase in the levels of key glycolytic enzymes. In addition, by regulating the expression of various PDK family members, HIF-1 α can also negatively regulate PDH activity, and thereby regulate TCA-dependent metabolism.

HSCs had emerged from G0 at an elevated rate compared to control cells, yet seemed to be trapped in G1. Consistent with this cell cycle arrest, PTPMT1 cells functioned poorly in transplantation assays. Metabolic assessment revealed that deletion of PTPMT1 resulted in reduced mitochondrial metabolism leading to energetic stress, at least as assessed by activation of AMP-activated kinase (AMPK). Reconstitution of various mutants into a knockout background revealed that the mitochondrial localization, as well as the catalytic function of PTPMT1, was necessary to restore functional HSCs. Interestingly, deletion of PTPMT1 in more committed myeloid, T cell, or B cell progenitors had no discernable effect. Thus, mitochondrial activity is required for HSC differentiation at the earliest stem cell and/or progenitor stage, while such activity is seemingly dispensable at later progenitor or mature cell stages.

Nutrient-Sensing Pathways

Taken together, these studies demonstrate that both mitochondrial and non-mitochondrial metabolism regulate key aspects of HSC function. Less clear from these observations is how metabolic programs are encoded and how energetic signals are transduced. One classical pathway that regulates nutrient signaling, especially downstream of growth factors, involves the sequential activation of phosphatidylinositol 3-kinase (PI3K), protein kinase B (AKT), and mechanistic target of rapamycin (mTOR). Relevant to the preceding discussion, the PTEN phosphatase is a known negative regulator of this pathway. Genetic evidence suggests that this pathway is clearly relevant to HSC biology. For instance, constitutively activated AKT1 leads to HSC exhaustion and high rate of development of leukemia (Kharas *et al.*, 2010). Interestingly, these animals are protected when they are administered the mTOR inhibitor rapamycin. Similarly, mice with a conditional deletion of PTEN within the bone marrow demonstrate an initial expansion, followed by the long-term exhaustion, of HSCs (Zhang *et al.*, 2006; Yilmaz *et al.*, 2006). Again, these mice appear to develop a premalignant expansion of bone marrow cells that is rescued by rapamycin treatment (Yilmaz *et al.*, 2006). This suggests a role for mTOR in HSC maintenance. Indeed, genetic removal of tuberous sclerosis 1 (Tsc1), a known negative regulator of mTOR, results in the activation of mTOR and the exit of HSCs out of quiescence and into a pattern of rapid cell cycling (Chen *et al.*, 2008a; Gan *et al.*, 2008). As such, activation anywhere along the PI3K, AKT, and mTOR pathways results in a disruption of the normal quiescence of HSCs with an initial expansion of stem and progenitor cells, followed by their premature exhaustion. The fact that manipulation of this pathway also results in premalignant conditions that mimic myeloproliferative disorders, or frank leukemia, suggests that inhibition of this axis might be of benefit to a wide swath of patients with hematological malignancies.

Another critical pathway for nutrient sensing in stem cells involves LKB1-AMPK signaling. While for the sake of clarity this pathway is discussed separately, it is important to note that AMPK regulates mTOR activity, and that both AKT and AMPK regulate the activity of the forkhead homeobox type O (FoxO) family of transcriptional regulators. As such,

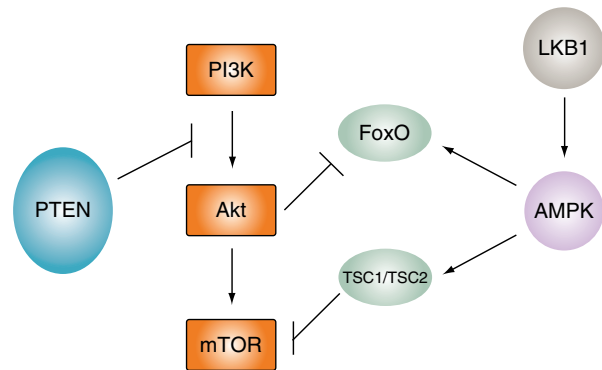


Figure 3 Interconnection between various energy sensors and transducers involved in stem cell function. Relationship between various kinases, phosphatases, and transcription factors each separately implicated in the biology of stem cells and in metabolic and redox homeostasis. See text for details.

LKB1/AMPK and PI3K/AKT/mTOR are integrated nutrient sensing pathways (Figure 3). The LKB1 kinase was initially characterized as the causal mutated gene in the familial cancer syndrome known as Peutz-Jeghers syndrome. This autosomal dominant disease is characterized by a hamartomatous polyposis syndrome (hence the increased risk for malignancy), as well as hyperpigmented macules on the lips and mouth. What role stem or progenitor cell activation plays in this proliferative disease is unclear. Genetic evidence suggests that LKB1 is upstream of AMPK. The activation of AMPK is generally viewed to occur when the AMP/ATP ratio is increased (i.e., during periods of energetic stress). Once activated, AMPK subsequently phosphorylates and often inactivates a number of energy requiring processes, including reducing the activity of mTOR. Much like what was observed for components of the PI3K/AKT pathway, conditional deletion of LKB1 leads to an initial burst of HSC activity, followed by their eventual exhaustion (Gan *et al.*, 2010; Gurumurthy *et al.*, 2010; Nakada *et al.*, 2010). The precise mechanism of this effect however remains elusive. While the obvious candidate would seem to be constitutive activation of mTOR, rapamycin was without effect. Similarly, as we will discuss later, while alterations in FOXO signaling can lead to oxidative stress, antioxidant treatment did not rescue *Lkb1* depletion in the bone marrow. Indeed, these three studies in which *Lkb1* was deleted in the bone marrow demonstrated that the observed effects on HSC biology appeared largely independent of AMPK, mTOR, or FoxO proteins. While the downstream effector remains somewhat elusive, some evidence suggests that LKB1 may act by altering mitochondrial biogenesis (Gan *et al.*, 2010) or potentially by altering lipid metabolism (Gurumurthy *et al.*, 2010).

Redox Homeostasis

Why is it that intracellular metabolism and stem cell biology are so tightly linked? At present, there is no definitive answer. One possibility is that adult stem cells need to persist for the life span of the organism, and as such need to develop mechanisms to reduce genomic damage. Since the endogenous

production of reactive oxygen species (ROS) is thought to fuel DNA damage, as well as potentially the aging process, tight metabolic regulation may be essential to maintain stem cell function throughout the organism's life span. Indeed, the fact that quiescent HSCs reside in a relatively hypoxic environment and seem to use their mitochondria only sparingly would seem like the ideal environmental conditions to reduce ROS production.

The connection between ROS levels and HSC function has been known for some time. Using the fluorescent compound dihydrodichlorofluorescein (DCF) it has been possible to obtain an understanding of redox regulation within bone marrow cells. The compound DCF is not initially fluorescent, however when oxidized by molecules such as hydrogen peroxide, a stable, fluorescent molecule is produced. As such, the level of DCF fluorescence is often used as a marker for intracellular ROS levels. An important caveat is that many believe that DCF is not directly oxidized by hydrogen peroxide and, therefore, not a reliable indicator of peroxide levels (Dikalov and Harrison, 2014). Nonetheless, largely due to practical issues involving the relatively few numbers of cells, this has been the predominant strategy employed. For instance, sorting HSCs based on DCF fluorescence (DCF^{high} meaning high levels of ROS and DCF^{low} corresponding to low levels of ROS) revealed that DCF^{low} cells had increased functionality following posttransplant reconstitution (Jang and Sharkis, 2007). Moreover, the DCF^{high} fraction appeared to be myeloid skewed, i.e., to produce more myeloid derivatives at the expense of lymphoid derivatives. Interestingly, this is a well-known property of aging bone marrow cells, a functional property observed in both mice and man. These observations solidified the well-known relationship between oxidants and senescence (Liang and Ghaffari, 2014).

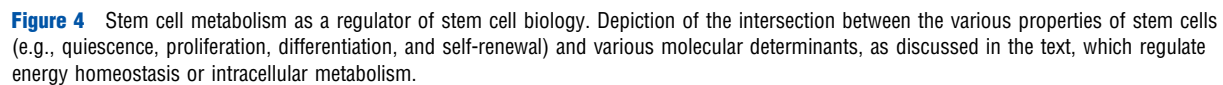
Another clear link between ROS homeostasis and HSC biology came from the analysis of mice lacking the gene encoding for the ataxia-telangiectasia mutated (ATM) kinase. This kinase has been implicated in the crippling disease ataxia-telangiectasia, characterized by a wide array of defects including neurological abnormalities, cancer predisposition, and an accelerated aging phenotype. ATM kinase plays a central role in coordinating the response to DNA damage, particularly the response to double strand breaks in the DNA. *Atm*^{-/-} mice recapitulate certain aspects of the human disease including the cancer predisposition. Most *Atm*^{-/-} mice succumb in the first few months of life to lymphoid malignancies. For those mice who manage to survive beyond the first several months, it was observed that a picture of progressive bone marrow failure developed (Ito *et al.*, 2004). Further experimentation revealed that *Atm*^{-/-} HSCs had impaired self-renewal. Using DCF, it was shown that deletion of *Atm* results in increased levels of ROS within the stem and progenitor population (Ito *et al.*, 2004). The addition of the antioxidant *N*-acetyl cysteine (NAC) that acts to replenish glutathione levels was able to preserve self-renewal capacity of *Atm*-deficient HSCs. Subsequent studies established that *Atm* deletion resulted in the redox-dependent activation of the p38 MAPK pathway, that in turn, resulted in the induction of the cell cycle inhibitor p16 (also known as INK4A) (Ito *et al.*, 2006). Again, similar to what was observed with NAC treatment, pharmacological p38 MAPK inhibitors improved HSC function (Ito *et al.*, 2006).

Furthermore, overexpression of the Polycomb gene *Bmi1*, known to epigenetically silence the *CDKN2A* locus that encodes p16, could also rescue the HSC defect in *Atm*^{-/-} mice (Ito *et al.*, 2004). In contrast to the rescue seen with *Bmi1* overexpression, deletion of *Bmi1* leads to defects in HSC self-renewal (He *et al.*, 2009). Interestingly, the loss of *Bmi1* also leads to increased levels of ROS (Liu *et al.*, 2009; Chattoo *et al.*, 2009).

Another important link between ROS levels and HSC function comes from the analysis of mice lacking members of the FoxO family of transcription factors. This family has four members that include FoxO1, FoxO3a, FoxO4, and FoxO6. Previous evidence suggested that this family was critical to maintaining redox homeostasis, as for instance, FoxO3a can induce the expression of antioxidant enzymes including catalase (Nemoto and Finkel, 2002) and superoxide dismutase (Kops *et al.*, 2002). Analysis of mice simultaneously deleted for three FoxO family members (FoxO1^{-/-}, FoxO3a^{-/-}, and FoxO4^{-/-}) demonstrated a reduction in HSC numbers over time and a reduced ability for FoxO-deficient cells to self-renew (Tothova *et al.*, 2007). Consistent with previous observations, the absence of FoxO proteins resulted in decreased levels of antioxidant defense proteins and a corresponding increase in the level of ROS. Like what was observed for *Atm*^{-/-} mice, mice deficient in FoxO activity could be rescued by NAC treatment. Similar observations were made in mice lacking a single FoxO member (FoxO3a) (Miyamoto *et al.*, 2007). Again, there was a defect in HSC function that appeared to be secondary to a rise in ROS levels and potentially to the subsequent redox-dependent activation of p38 MAPK.

Other Stem Cell Niches

Our discussion so far has dealt almost exclusively with HSCs. This is largely because this experimental model is the most widely used and studied. Nonetheless, many of the lessons learned from this system appear to be applicable to other stem cell niches. For instance, cultured embryonic stem cells appear to exhibit high rates of glycolysis (Kondoh *et al.*, 2007). Interestingly, in the induced pluripotent stem cell (iPSC) system, a shift from mitochondrial metabolism to glycolysis appears to accompany, and be required, for high efficiency reprogramming of somatic cells to stem-like iPSCs (Folmes *et al.*, 2011). A similar situation appears to exist for mesenchymal stem cell (Chen *et al.*, 2008b). In addition, the energy sensing and redox-regulating pathways discussed within the context of HSCs are often applicable to other stem cells. For instance, as seen in the case of HSCs, deletion of FoxO3 in neural stem cells (NSCs) results in impaired self-renewal, again through activation of a redox-sensitive pathway (Renault *et al.*, 2009; Paik *et al.*, 2009). Similarly, there appears to be an important role of FAO and fatty acid synthesis in diverse stem cell niches (Ito and Suda, 2014). This may be in part because FAO helps to maintain levels of NADPH required for redox homeostasis (Jeon *et al.*, 2012). Similarly, the maintenance of adult NSCs is dependent on fatty acid synthesis (Knobloch *et al.*, 2013). In addition, the mTOR pathway has been recently shown to play a role in the maintenance of muscle stem cell quiescence (Rodgers *et al.*, 2014). Nonetheless, while many pathways



epigenetic state. Given the critical role of epigenetics in maintaining 'stemness,' as well as aiding the program of differentiation, it would seem that there is much to learn about the convergence of intracellular metabolism and stem cell biology. Such insights may provide important clues on the origin of cancer, as well as how it might be possible to rationally delay the age-dependent decline in stem cell function.

Conclusions

References

- Chato, W., Abdouh, M., David, J., *et al.*, 2009. The polycomb group gene Bmi1 regulates antioxidant defenses in neurons by repressing p53 pro-oxidant activity. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 29, 529–542.
- Chen, C., Liu, Y., Liu, R., *et al.*, 2008a. TSC-mTOR maintains quiescence and function of hematopoietic stem cells by repressing mitochondrial biogenesis and reactive oxygen species. *Journal of Experimental Medicine* 205, 2397–2408.
- Chen, C.T., Shih, Y.R., Kuo, T.K., Lee, O.K., Wei, Y.H., 2008b. Coordinated changes of mitochondrial biogenesis and antioxidant enzymes during osteogenic differentiation of human mesenchymal stem cells. *Stem Cells* 26, 960–968.
- Chow, D.C., Wenning, L.A., Miller, W.M., Papoutsakis, E.T., 2001. Modeling pO₂ distributions in the bone marrow hematopoietic compartment. II. Modified Kroghian models. *Biophysical Journal* 81, 685–696.
- Chuiikov, S., Levi, B.P., Smith, M.L., Morrison, S.J., 2010. Prdm16 promotes stem cell maintenance in multiple tissues, partly by regulating oxidative stress. *Nature Cell Biology* 12, 999–1006.
- Cipolleschi, M.G., Dello Sbarba, P., Olivetto, M., 1993. The role of hypoxia in the maintenance of hematopoietic stem cells. *Blood* 82, 2031–2037.

- Danet, G.H., Pan, Y., Luongo, J.L., Bonnet, D.A., Simon, M.C., 2003. Expansion of human SCID-repopulating cells under hypoxic conditions. *Journal of Clinical Investigation* 112, 126–135.
- Dikalov, S.I., Harrison, D.G., 2014. Methods for detection of mitochondrial and cellular reactive oxygen species. *Antioxidants & Redox Signaling* 20, 372–382.
- Folmes, C.D., Nelson, T.J., Martinez-Fernandez, A., *et al.*, 2011. Somatic oxidative bioenergetics transitions into pluripotency-dependent glycolysis to facilitate nuclear reprogramming. *Cell Metabolism* 14, 264–271.
- Gan, B., Hu, J., Jiang, S., *et al.*, 2010. Lkb1 regulates quiescence and metabolic homeostasis of haematopoietic stem cells. *Nature* 468, 701–704.
- Gan, B., Sahin, E., Jiang, S., *et al.*, 2008. mTORC1-dependent and -independent regulation of stem cell renewal, differentiation, and mobilization. *Proceedings of the National Academy of Sciences of the United States of America* 105, 19384–19389.
- Gurumurthy, S., Xie, S.Z., Alagesan, B., *et al.*, 2010. The Lkb1 metabolic sensor maintains haematopoietic stem cell survival. *Nature* 468, 659–663.
- He, S., Nakada, D., Morrison, S.J., 2009. Mechanisms of stem cell self-renewal. *Annual Review of Cell and Developmental Biology* 25, 377–406.
- Ito, K., Carracedo, A., Weiss, D., *et al.*, 2012. A PML-PPAR-delta pathway for fatty acid oxidation regulates hematopoietic stem cell maintenance. *Nature Medicine* 18, 1350–1358.
- Ito, K., Hirao, A., Arai, F., *et al.*, 2004. Regulation of oxidative stress by ATM is required for self-renewal of haematopoietic stem cells. *Nature* 431, 997–1002.
- Ito, K., Hirao, A., Arai, F., *et al.*, 2006. Reactive oxygen species act through p38 MAPK to limit the lifespan of hematopoietic stem cells. *Nature Medicine* 12, 446–451.
- Ito, K., Suda, T., 2014. Metabolic requirements for the maintenance of self-renewing stem cells. *Nature Reviews Molecular Cell Biology* 15, 243–256.
- Iyer, N.V., Kotch, L.E., Agani, F., *et al.*, 1998. Cellular and developmental control of O₂ homeostasis by hypoxia-inducible factor 1 alpha. *Genes & Development* 12, 149–162.
- Jang, Y.Y., Sharkis, S.J., 2007. A low level of reactive oxygen species selects for primitive hematopoietic stem cells that may reside in the low-oxygenic niche. *Blood* 110, 3056–3063.
- Jeon, S.M., Chandel, N.S., Hay, N., 2012. AMPK regulates NADPH homeostasis to promote tumour cell survival during energy stress. *Nature* 485, 661–665.
- Kharas, M.G., Okabe, R., Ganis, J.J., *et al.*, 2010. Constitutively active AKT depletes hematopoietic stem cells and induces leukemia in mice. *Blood* 115, 1406–1415.
- Knobloch, M., Braun, S.M., Zurkirchen, L., *et al.*, 2013. Metabolic control of adult neural stem cell activity by Fasn-dependent lipogenesis. *Nature* 493, 226–230.
- Kondoh, H., Leonart, M.E., Nakashima, Y., *et al.*, 2007. A high glycolytic flux supports the proliferative potential of murine embryonic stem cells. *Antioxidants & Redox Signaling* 9, 293–299.
- Kops, G.J., Dansen, T.B., Polderman, P.E., *et al.*, 2002. Forkhead transcription factor FOXO3a protects quiescent cells from oxidative stress. *Nature* 419, 316–321.
- Liang, R., Ghaffari, S., 2014. Stem cells, redox signaling, and stem cell aging. *Antioxidants & Redox Signaling* 20, 1902–1916.
- Liu, J., Cao, L., Chen, J., *et al.*, 2009. Bmi1 regulates mitochondrial function and the DNA damage response pathway. *Nature* 459, 387–392.
- Liu, J., Finkel, T., 2013. Stem cells and oxidants: Too little of a bad thing. *Cell Metabolism* 18, 1–2.
- Lu, C., Thompson, C.B., 2012. Metabolic regulation of epigenetics. *Cell Metabolism* 16, 9–17.
- Miyamoto, K., Araki, K.Y., Naka, K., *et al.*, 2007. Foxo3a is essential for maintenance of the hematopoietic stem cell pool. *Cell Stem Cell* 1, 101–112.
- Morimoto, H., Iwata, K., Ogonuki, N., *et al.*, 2013. ROS are required for mouse spermatogonial stem cell self-renewal. *Cell Stem Cell* 12, 774–786.
- Nakada, D., Saunders, T.L., Morrison, S.J., 2010. Lkb1 regulates cell cycle and energy metabolism in haematopoietic stem cells. *Nature* 468, 653–658.
- Nemoto, S., Finkel, T., 2002. Redox regulation of forkhead proteins through a p66shc-dependent signaling pathway. *Science* 295, 2450–2452.
- Nombela-Arrieta, C., Pivarnik, G., Winkel, B., *et al.*, 2013. Quantitative imaging of hematopoietic stem and progenitor cell localization and hypoxic status in the bone marrow microenvironment. *Nature Cell Biology* 15, 533–543.
- Paik, J.H., Ding, Z., Narurkar, R., *et al.*, 2009. FoxOs cooperatively regulate diverse pathways governing neural stem cell homeostasis. *Cell Stem Cell* 5, 540–553.
- Parmar, K., Mauch, P., Vergilio, J.A., Sackstein, R., Down, J.D., 2007. Distribution of hematopoietic stem cells in the bone marrow according to regional hypoxia. *Proceedings of the National Academy of Sciences of the United States of America* 104, 5431–5436.
- Renault, V.M., Rafalski, V.A., Morgan, A.A., *et al.*, 2009. FoxO3 regulates neural stem cell homeostasis. *Cell Stem Cell* 5, 527–539.
- Rodgers, J.T., King, K.Y., Brett, J.O., *et al.*, 2014. mTORC1 controls the adaptive transition of quiescent stem cells from G0 to G(Alert). *Nature* 510 (7505), 393–396.
- Schieke, S.M., Ma, M., Cao, L., *et al.*, 2008. Mitochondrial metabolism modulates differentiation and teratoma formation capacity in mouse embryonic stem cells. *Journal of Biological Chemistry* 283, 28506–28512.
- Shen, J., Liu, X., Yu, W.M., *et al.*, 2011. A critical role of mitochondrial phosphatase Ptpmt1 in embryogenesis reveals a mitochondrial metabolic stress-induced differentiation checkpoint in embryonic stem cells. *Molecular and Cellular Biology* 31, 4902–4916.
- Simsek, T., Kocabas, F., Zheng, J., *et al.*, 2010. The distinct metabolic profile of hematopoietic stem cells reflects their location in a hypoxic niche. *Cell Stem Cell* 7, 380–390.
- Takubo, K., Goda, N., Yamada, W., *et al.*, 2010. Regulation of the HIF-1alpha level is essential for hematopoietic stem cells. *Cell Stem Cell* 7, 391–402.
- Takubo, K., Nagamatsu, G., Kobayashi, C.I., *et al.*, 2013. Regulation of glycolysis by Pdk functions as a metabolic checkpoint for cell cycle quiescence in hematopoietic stem cells. *Cell Stem Cell* 12, 49–61.
- Tothova, Z., Kollipara, R., Huntly, B.J., *et al.*, 2007. FoxOs are critical mediators of hematopoietic stem cell resistance to physiologic oxidative stress. *Cell* 128, 325–339.
- Wang, G.L., Semenza, G.L., 1995. Purification and characterization of hypoxia-inducible factor 1. *Journal of Biological Chemistry* 270, 1230–1237.
- Yilmaz, O.H., Valdez, R., Theisen, B.K., *et al.*, 2006. Pten dependence distinguishes hematopoietic stem cells from leukaemia-initiating cells. *Nature* 441, 475–482.
- Yu, W.M., Liu, X., Shen, J., *et al.*, 2013. Metabolic regulation by the mitochondrial phosphatase PTPMT1 is required for hematopoietic stem cell differentiation. *Cell Stem Cell* 12, 62–74.
- Zhang, J., Grindley, J.C., Yin, T., *et al.*, 2006. PTEN maintains haematopoietic stem cells and acts in lineage choice and leukaemia prevention. *Nature* 441, 518–522.
- Zhang, J., Guan, Z., Murphy, A.N., *et al.*, 2011. Mitochondrial phosphatase PTPMT1 is essential for cardiolipin biosynthesis. *Cell Metabolism* 13, 690–700.

Pluripotent Cells: New Tools for Disease Research and Therapies

TD Tlsty, University of California, San Francisco, CA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Pluripotent A cellular state where differentiation is possible into each of the three germline derivatives.

SON A mutually regulated pathway of three proteins, SOX2, OCT3/4, and Nanog, that confers pluripotent potential to a cell.

Naturally Occurring Pluripotent Cells

By definition, cells within a pluripotent state can form tissue derivatives of the three embryonic germ layers (ectoderm, mesoderm, and endoderm) (Blau *et al.*, 2001). This is in contrast to multipotent cells whose offspring are restricted to forming organ-specific cells and totipotent cells that have the potential to generate all embryonic, as well as extraembryonic, tissues such as placenta. The first commonly studied pluripotent cell was the embryonic stem cell (ESC), first isolated from mice and teratocarcinomas in 1981 (Martin, 1981; Evans and Kaufman, 1981). Isolation of human embryonic stem cell (hESC) followed several years later (Thomson *et al.*, 1998). The primary source of pluripotent ESCs is the inner cell mass of the blastocyst irrespective of whether the animal is an amphibian or mammal. As the ESCs within the blastocyst are programmed to lose pluripotent capacity with maturation of the embryo, the genital ridge gives rise to embryonic germ cells (EG) or primordial germ cells (PGC) which also exhibit pluripotent capability. Spermatogonial germ stem cells are a further source of pluripotent cells.

Surprisingly, other, more differentiated, cells within the fetal and adult mammalian body have also been reported to have pluripotent potential. These include fetal mesenchymal fibroblasts, placental cord blood, and mesenchymal stem cells (MSCs) (Kues *et al.*, 2005; Kogler *et al.*, 2004; Lee *et al.*, 2004). More recently, it has been reported that numerous adult tissues contain cells with pluripotent potential. Skin progenitor cells (SKP), multipotent adult progenitor cells (MAPC), marrow-isolated adult multilineage inducible (MIAMI) cells, and very small epithelial-like cells (vSELs) are just a few of the examples (Toma *et al.*, 2001; Jiang *et al.*, 2002a; D'Ippolito *et al.*, 2004; Ratajczak *et al.*, 2011). Most recently, it has been demonstrated that rare cells within multiple human adult tissues respond to wounding by proliferating and exhibiting pluripotent potential. Thus, multilineage differentiating stress enduring (MUSE) cells (Kuroda *et al.*, 2010) and endogenous plastic somatic (ePS) cells (Roy *et al.*, 2013) may contribute to wound healing. The isolation and characterization of human amnion epithelial cells and amniotic fluid stem cells has been extensively described and is being tested in organ reconstruction studies (Murphy and Atala, 2013). These reports demonstrate that pluripotency is not a phenotype that is restricted to ESCs or a specific developmental phase, a still controversial conclusion.

Engineered Pluripotent Cells

For many decades it was believed that cell differentiation was irreversible. Conventional wisdom held that pluripotent ESCs

would commit to one of the three primary lineages and further restriction of differentiation potential would progress from that point forward. In the early 1950s, Briggs and King (1952) transplanted blastula nuclei into enucleated frog eggs and obtained viable tadpoles, paving the way for John Gurdon to determine if the nucleus of a fully differentiated cell was irreversibly altered during cell differentiation using somatic cell nuclear transfer (SCNT). Transplantation of a nucleus from an intestinal cell of a tadpole into an enucleated egg produced a low number of normal adult frogs demonstrating a fundamental and general principle that cell differentiation can take place without stable changes to the genome (Gurdon, 1962a,b). These studies have been extended to mammalian nuclei and demonstrate that a terminally differentiated cell can be reprogrammed to totipotency (Wilmut *et al.*, 1997; Hochedlinger and Jaenisch, 2002).

Molecular *in vitro* investigations into the mechanisms that allow cells to reverse differentiation and acquire pluripotency generated surprising results. Pluripotency as well as specific cell types are specified by transcription factors. In 2006, a team led by Yamanaka reported that four transcription factors (OCT4, SOX2, Klf4, and c-Myc) were sufficient to take a fibroblast nucleus to a pluripotent state (Figure 1; Takahashi and Yamanaka, 2006). In a slightly different approach, the Thompson team reported that expression of OCT4, SOX2, NANOG, and LIN28 would accomplish the same goal (Yu *et al.*, 2007). Reprogramming using these methods are slow and proceed with very low efficiency. Major changes in gene expression and epigenetic states are obvious. To claim pluripotent potential, the resulting induced pluripotent stem cells (iPSCs) were required to express classic ESC-specific proteins and cell surface antigens, differentiate into derivatives of all three germ lineages *in vitro*, and generate teratomas when injected into immune-compromised mice. The teratomas must contain differentiated derivatives of all three primary germ layers (ectodermal, mesodermal, and endodermal) and be able to contribute to the next generation of offspring (Thomson *et al.*, 1998).

Having established that differentiated cells are not irreversibly locked in their mature or terminal identity *in vitro* but may be activated to generate any and all other cell types of the body, the question becomes how, when, and where may this expression of pluripotency occur? Additionally, is pluripotency expressed more widely *in vivo* than previously believed?

New Tools to Provide Mechanistic Insights on Determining Cell Fates and Disease States

The existence of pluripotent cells has provided an important tool for studying the mysteries of developmental biology and

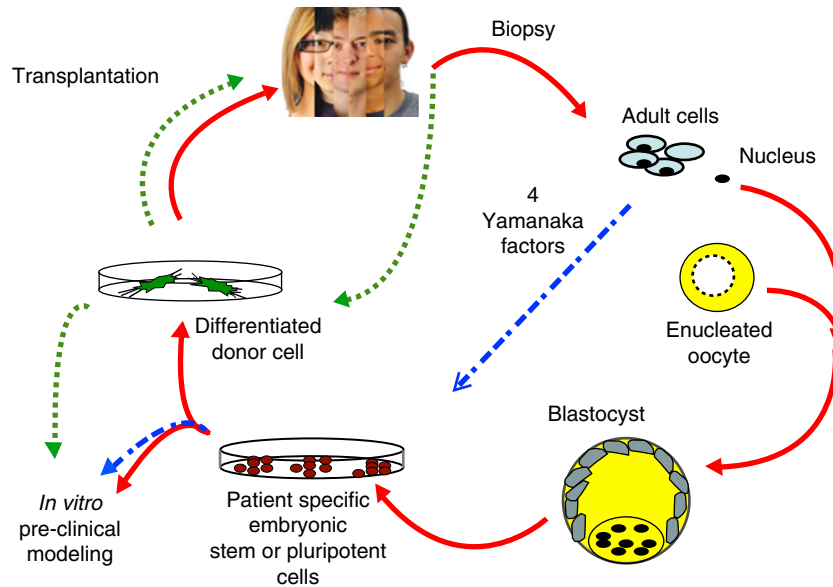


Figure 1 Pluripotent cells as tools for disease research and therapies. The red arrows indicate the methodological route taken to produce human embryonic stem cells (hESCs). An adult cell is removed from the body and the nucleus is transferred to an enucleated oocyte. The resulting ESCs can be used for generating different cell types that can subsequently be utilized for drug screening or tissue reconstruction as part of *in vitro* preclinical modeling. Additionally, the intention is that they will also be useful for transplantation into recipients in need of cell-based therapies. The blue arrows indicate the initial methodological route taken to produce human induced pluripotent stem cells (iPSCs). A cocktail of transcription factors are introduced into adult differentiated cells and rare pluripotent cells are generated. Multiple combinations of transcription factors have now been found to be effective in generating embryonic pluripotent stem cells (ePSCs). Attempts of chemical induction of pluripotency are underway. Similar to ESCs, the intention is that iPSCs will also be useful for transplantation into recipients in need of cell-based therapies. Finally, the green arrows indicate the methodological route taken to enrich human cells with pluripotent potential that have been reported to reside within the body. While these cells can also be used in a manner similar to that described for ESCs or iPSCs, endogenous pluripotent cells have additional options. Since they reside *in situ*, it may someday be possible to program these cells to desirable derivatives without removing them from the body.

the generation of tissues (Hough *et al.*, 2014; Ohnishi *et al.*, 2014). Pluripotent cells have tremendous potential for use in preclinical modeling of human diseases as well as providing an experimental system to screen for effective drugs. For example, patient-derived iPSCs may be used as a primary source of cellular material for correction of molecular defects through genetic engineering. The ‘repaired’ cell may then be differentiated and coupled with cell therapy (see below) so the corrected derivatives may be returned to the donor. Diseases such as metabolic syndromes (Pompe’s disease or glycogen storage disease 1a) where cells lack specific enzymes used to generate cellular building blocks could be corrected in this manner. Additionally, the use of this technology to generate cells with a disease-relevant phenotype is groundbreaking. If the redifferentiated cell phenocopies the disease trait in the original patient, this material could then be used for screening of drugs or disease-modeling *ex vivo*. Typically many types of diseases cannot be studied for lack of access to appropriate human cells that exhibit the critical phenotype. Take the example of familial dysautonomia, a rare neurological disease where one mutation causes an extensive autonomous nervous system defect. These cells are difficult to procure and study. However, with the advent of iPSC technology, one can envision that a patient-specific iPSC would be generated and then differentiated into the affected cell type. This was accomplished in 2009 (Lee *et al.*, 2009) and the cells were successfully used to screen for drugs that could reduce the pathological

phenotypes. Several successful examples now exist using this approach (Robinton and Daley, 2012). For example, in a first-of-its-kind study, a humanized mouse model of sickle cell anemia had its disease phenotype rescued (Hanna *et al.*, 2007). Homologous recombination was used to repair the specific gene defect in iPSCs derived from the host, the corrected cells were subsequently programmed to hematopoietic progenitors and then transplanted back into the host. While several studies have used iPSCs derived from patients with specific diseases as a platform to then test drugs for efficacy, to date none of the identified drugs have been validated as working in the patients that had the original disease. While these scenarios are attractive, many of the hurdles that remain to be overcome are discussed below.

New Tools for Cell-Based Regenerative Therapy

There is tremendous excitement about the potential to use human stem cells for regenerative medicine. An impressive list of degenerative diseases and injuries would benefit from the ability to safely and correctly augment functional human tissues. This would consist of many examples including correction of degenerative neural diseases like Parkinson’s disease, Alzheimer’s disease, and dementias; repair of poor eyesight or immune deficiencies; and reversal of conditions like arthritis or organ failure. Current approaches to obtain cells for regenerative

medicine include adult stem cells, hESCs (after production, differentiation, and expansion), and, most recently, iPSCs (after production, differentiation, and expansion).

Each of these approaches have provided exciting and promising starting points but suffers from critical problems that limit their application. Currently, the only cells that are routinely being used in the clinic for regenerative medicine are adult stem cells. Several decades of success have proven the value of this approach with hematopoietic stem cells and skin stem cells. For example, bone marrow transplants or skin transplants have been used for decades to replenish a depleted hematopoietic system or to aid burn victims, respectively. However, despite the utility of this methodology, room for improvement exists. Indeed, adult stem cells therapy is often non-autologous and therefore necessitates immunosuppressive measures to prevent rejection of tissues. For example, patients with successful bone marrow transplants or skin transplants often face issues of graft versus host disease or host versus graft disease and require the need for lifelong administration of immunosuppressive therapy. Chronic flares of immunosuppressive issues can range from irritating to life threatening. Identification of other tissue-specific stem cells and their application to tissue repair will likely be limited by the same considerations. In contrast, the ability to use pluripotent cells from the patients' own body might be expected to avoid problems with immunosuppression often encountered with the use of adult multipotent stem cells taken from other patients.

Meanwhile, although under intense development, the successful use of an alternative source of cells, hESCs, has yet to be realized in the clinic. Aside from the ethical issues associated with destroying human embryos for the production of hESCs, similar immunosuppressive issues exist. In addition, issues of potential tumorigenicity related to the immortality and genetic instability of hESCs cause problems. Integrating differentiated cells into an already existing organ is not trivial. Additionally, the technically challenging and time-consuming process of generating ESCs does not have the potential to generate patient-specific differentiated tissues for acute injury. The proposed method to circumvent this problem is to create a bank of ESCs of different immune complement groups for off-the-shelf use. The first clinical trial utilizing hESC-derived cells to address a neurological defect was terminated midway through the trial for lack of funds. No adverse effects have been reported to date and the final results have not yet been reported. The first clinical trials using hESCs for eye repair have been initiated and are currently underway.

When the generation of human iPSCs was published in 2007 (Takahashi *et al.*, 2007), the excitement about the potential for these cells to advance regenerative medicine was obvious. However, several technical problems with the generation and use of iPSCs preclude their clinical use at the current time (Blasco *et al.*, 2011). The generation of iPSCs is a rare event and the introduction of exogenous DNA into human cells to activate the pluripotency pathways is not compatible with Federal Drug Administration (FDA) standards. Investigators are seeking cocktails of small molecules that will circumvent the FDA concern and at the same time generate high quality pluripotent iPSCs. Importantly, iPSCs exhibit genomic instability and immortality thus exhibiting a very real possibility of forming malignant tumors. In addition

to these challenges, current studies are documenting an 'epigenetic memory' that encourages the generated iPSCs to reset differentiated programs to that of the original tissue. For example, starting with fibroblasts, a mesodermal cell type, it is more difficult to generate ectodermal and endodermal derivatives than to generate mesodermal derivatives. This epigenetic instability has also been documented for multipotent cells derived from amniotic fluid (Rennie *et al.*, 2012). The iPSC technology in its present form cannot be used to introduce human cells back into the human body.

An unmet need in this field is to find a source of cells that can be used for regenerative medicine that will not be rejected by the patient's body and that will not carry exogenous reprogramming genes or an elevated threat of tumor formation. It is hypothesized that one of the reported endogenous pluripotent cell populations may be optimal to overcome the obstacles of immune rejection and rapid delivery. Using a person's own cells for regenerative medicine would circumvent the need for immunosuppressive drugs. Furthermore, the mortality, documented genomic stability, and nontumorigenicity of these recently described endogenous pluripotent cells would allay fears about high risk for spurious cancer formation. The greatest concerns, no matter which cell population eventually used, are for the safety of the patient and the fidelity of the regenerated tissue.

Clinical Trials and Contributions to Future Cell-Based Regenerative Therapy

Despite the concerns mentioned above, carefully controlled clinical trials using iPSCs have just been initiated to repair age-related macula degeneration. In the fall of 2014, Dr. Masayo Takahashi, an ophthalmologist at the RIKEN Center for Developmental Biology initiated the first clinical trials in humans. She generated iPSCs and directed them to become retinal pigment epithelium cells. Thin sheets of these cells were transplanted into the damaged retina of a 70-year old patient. This historic trial, ultimately involving six patients, is ongoing and the end results are awaited with great anticipation. Extensive preliminary studies with primates suggested that tumor formation was low and immune rejection was not provoked.

Once the barriers described above are overcome, current discoveries are providing the proof of principle that cell-based therapies are feasible. Clearly, these studies are just the beginning of a tsunami of breakthroughs in this field and the near future will provide much information about the successes and failures of using pluripotent cells in regenerative medicine.

See also: Stem Cell Biology: Structure and Function: Cancer Stem Cells

References

- Blasco, M.A., Serrano, M., Fernandez-Capetillo, O., 2011. Genomic instability in iPS: Time for a break. *EMBO Journal* 30, 991–993.
- Blau, H.M., Brazelton, T.R., Weimann, J.M., 2001. The evolving concept of a stem cell: Entity or function? *Cell* 105, 829–841.

- Briggs, R., King, T.J., 1952. Transplantation of living nuclei from blastula cells into enucleated frogs' eggs. *Proceedings of the National Academy of Sciences United States of America* 38, 455–463.
- D'Ippolito, G., Diabira, S., Howard, G.A., *et al.*, 2004. Marrow-isolated adult multilineage inducible (MIAMI) cells, a unique population of postnatal young and old human cells with extensive expansion and differentiation potential. *Journal of Cell Science* 117, 2971–2981.
- Evans, M., Kaufman, M., 1981. Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154–156.
- Gurdon, J.B., 1962a. Adult frogs derived from the nuclei of single somatic cells. *Developmental Biology* 4, 256–273.
- Gurdon, J.B., 1962b. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *Journal of Embryology & Experimental Morphology* 10, 622–640.
- Hanna, J., Werniig, M., Markoulaki, S., *et al.*, 2007. Treatment of sickle cell anemia mouse model with iPS cells generated from autologous skin. *Science* 318, 1920–1923.
- Hochedlinger, K., Jaenisch, R., 2002. Monoclonal mice generated by nuclear transfer from mature B and T donor cells. *Nature* 415, 1035–1038.
- Hough, S.R., Thornton, M., Mason, E., *et al.*, 2014. Single-cell gene expression profiles define self-renewing, pluripotent, and lineage primed states of human pluripotent stem cells. *Stem Cell Reports* 2, 881–895.
- Jiang, Y., Jahagirdar, B.N., Reinhardt, R.L., *et al.*, 2002a. Pluripotency of mesenchymal stem cells derived from adult marrow. *Nature* 418, 41–49.
- Kogler, G., Sensken, S., Airey, J.A., *et al.*, 2004. A new human somatic stem cell from placental cord blood with intrinsic pluripotent differentiation potential. *Journal of Experimental Medicine* 200, 123–135.
- Kues, W.A., Petersen, B., Mysegades, W., *et al.*, 2005. Isolation of murine and porcine fetal stem cells from somatic tissue. *Biology of Reproduction* 72, 1020–1028.
- Kuroda, Y., Kitada, M., Wakao, S., *et al.*, 2010. Unique multipotent cells in adult human mesenchymal cell populations. *Proceedings of the National Academy of Sciences United States of America* 107, 8639–8643.
- Lee, G., Papapetrou, E.P., Kim, H., *et al.*, 2009. Modelling pathogenesis and treatment of familial dysautonomia using patient-specific iPSCs. *Nature* 461, 402–406.
- Lee, O.K., Kuo, T.K., Chen, W.M., *et al.*, 2004. Isolation of multipotent mesenchymal stem cells from umbilical cord blood. *Blood* 103, 1669–1675.
- Martin, G., 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences United States of America* 78, 7634–7638.
- Murphy, S.V., Atala, A., 2013. Amniotic fluid and placental membranes: Unexpected sources of highly multipotent cells. *Seminars in Reproductive Medicine* 31, 62–68.
- Ohnishi, Y., Huber, W., Tsumura, A., *et al.*, 2014. Cell-to-cell expression variability followed by signal reinforcement progressively segregates early mouse lineages. *Nature Cell Biology* 16, 27–37.
- Ratajczak, M.Z., Liu, R., Ratajczak, J., *et al.*, 2011. The role of pluripotent embryonic-like stem cells residing in adult tissues in regeneration and longevity. *Differentiation* 81, 153–161.
- Rennie, K., Gruslin, A., Hengstschlager, M., *et al.*, 2012. Applications of amniotic membrane and fluid in stem cell biology and regenerative medicine. *Stem Cells International* 2012, 721538.
- Robinton, D.A., Daley, G.Q., 2012. The promise of induced pluripotent stem cells in research and therapy. *Nature* 481, 295–305.
- Roy, S., Gasdcard, P., Dumont, N., *et al.*, 2013. Rare somatic cells from human breast tissue exhibit extensive lineage plasticity. *Proceedings of the National Academy of Sciences United States of America* 110, 4598–4603.
- Takahashi, K., Tanabe, K., Ohnuki, M., *et al.*, 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* 131, 861–872.
- Takahashi, K., Yamanaka, S., 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676.
- Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., *et al.*, 1998. Embryonic stem cell lines derived from human blastocysts. *Science* 282, 1145–1147.
- Toma, J.G., Akhavan, M., Fernandes, K.J., *et al.*, 2001. Isolation of multipotent adult stem cells from the dermis of mammalian skin. *Nature Cell Biology* 3, 778–784.
- Wilmot, I., Schnieke, A.E., McWhir, J., *et al.*, 1997. Viable offspring derived from fetal and adult mammalian cells. *Nature* 385, 810–813.
- Yu, J., Vodyanik, M.A., Smuga-Otto, K., *et al.*, 2007. Induced pluripotent stem cell lines derived from human somatic cells. *Science* 318, 1917–1920.

Further Reading

- Burt, R.K., Loh, Y., Pearce, W., *et al.*, 2008. Clinical applications of blood-derived and marrow-derived stem cells for nonmalignant diseases. *Journal of the American Medical Association* 299, 925–936.
- Cahan, P., Daley, G.Q., 2013. Origins and implications of pluripotent stem cell variability and heterogeneity. *Nature Reviews Molecular Cell Biology* 14, 357–368.
- Estrov, Z., 2009. Stem cells and somatic cells: Reprogramming and plasticity. *Clinical Lymphoma Myeloma* 3, S319–S328.
- Hipp, J., Atala, A., 2008. Sources of stem cells for regenerative medicine. *Stem Cell Reviews* 4, 3–11.
- Jaenisch, R., Young, R., 2008. Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* 132, 567–582.
- Marks, H., Kalkan, T., Menafr, R., *et al.*, 2012. The transcriptional and epigenomic foundations of ground state pluripotency. *Cell* 149, 590–604.
- Nichols, J., Smith, A., 2012. Pluripotency in the embryo and in culture. *Cold Spring Harbor Perspectives in Biology* 4, a008128.
- Stadtfeld, M., Hochedlinger, K., 2010. Induced pluripotency: History, mechanisms, and applications. *Genes and Development* 24, 2239–2263.
- Zipori, D., 2004. The nature of stem cells: State rather than entity. *Nature Reviews Genetics* 5, 873–878.

ENCYCLOPEDIA OF CELL BIOLOGY

ENCYCLOPEDIA OF CELL BIOLOGY

EDITORS-IN-CHIEF

RALPH A BRADSHAW

University of California, Irvine, CA, USA

PHILIP D STAHL

Washington University School of Medicine, St. Louis, MO, USA

VOLUME 4

SYSTEMS CELL BIOLOGY

VOLUME EDITOR

DOUG A LAUFFENBURGER

Massachusetts Institute of Technology, Cambridge, MA, USA

SECTION EDITORS

JASON M HAUGH

North Carolina State University, Raleigh, NC, USA

HELEN M BYRNE

University of Oxford, Oxford, UK

RUNE LINDING

University of Copenhagen, Copenhagen, Denmark



AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO
Academic Press is an imprint of Elsevier



Academic Press is an imprint of Elsevier
225 Wyman Street, Waltham, MA 02451, USA
The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, UK

Copyright © 2016 Elsevier Inc. All rights reserved

Cover background image created by M W Goldberg and T D Allen.

Foreground images © Clifford Harding, John Heuser and Philip D Stahl, 1983, originally published in *The Journal of Cell Biology*, Volume 97, 329–339.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher. Details on how to seek permission, further information about the Publisher's permissions policies and our arrangements with organizations such as the Copyright Clearance Center and the Copyright Licensing Agency, can be found at our website: www.elsevier.com/permissions.

This book and the individual contributions contained in it are protected under copyright by the Publisher (other than as may be noted herein).

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers may always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

ISBN 978-0-12-394447-4

For information on all publications visit our
website at <http://store.elsevier.com>

Printed and bound in the United States of America

15 16 17 18 19 20 10 9 8 7 6 5 4 3 2 1



Working together
to grow libraries in
developing countries

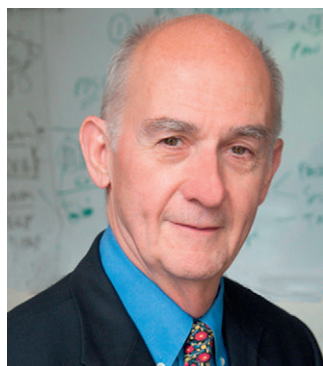
www.elsevier.com • www.bookaid.org

Publisher: Adam Gilbert
Acquisition Editor: Ginny Mills
Content Project Manager: Blerina Osmanaj
Associate Content Project Manager: Victoria Dyer
Cover Designer: Maria Inês Cruz

EDITORS-IN-CHIEF



Ralph A Bradshaw is Professor Emeritus in the Department of Physiology and Biophysics at the University of California, Irvine. Prior to that he was on the faculty of the Department of Biological Chemistry, Washington University School of Medicine in St. Louis, MO and was Professor and Chair of the Department of Biological Chemistry at University of California, Irvine. From 2006 to 2015, he was a member of the Mass Spectrometry Facility and Professor of Pharmaceutical Chemistry at the University of California, San Francisco. He holds degrees from Colby College and Duke University and was a post-doctoral fellow at Indiana University and the University of Washington. He has served as president of FASEB, was the founding president of the Protein Society and was the treasurer of the American Society for Biochemistry and Molecular Biology. His research has focused on protein chemistry and proteomics, with emphasis on the structure and function of growth factors and their receptors, particularly nerve growth factor and fibroblast growth factor, and the involvement of receptor tyrosine kinases in cell signaling. He has also studied the role of proteolytic processing and N-terminal modification in protein stability and turnover.



Philip D Stahl is E. Mallinckrodt Jr. Professor Emeritus at Washington University School of Medicine in St. Louis, MO. He was educated at West Virginia University with post-doctoral work at Vanderbilt University. He served as Head of the Department of Cell Biology and Physiology and Director of the Division of Biology and Biomedical Sciences at Washington University. He has been the recipient of many awards including a MERIT award from the NIH and the WICB Senior Recognition award given by the American Society for Cell Biology in recognition of his work supporting the advancement of women in science. Among StahlLab contributions are the discovery of the lysosomal enzyme clearance pathway now implemented in the treatment of lysosomal storage disease, discovery of the innate immune receptor, the mannose receptor, discovery of the exosome secretion pathway, and the role of Rab5 and Arf6 in endocytosis.

Currently his research focuses on endocytosis, signal transduction, and exosome biogenesis and secretion.

VOLUME EDITORS



Gerald W Hart is the Director and DeLamar Professor of Biological Chemistry at Johns Hopkins Medical School. He began his research on glycoconjugates about 40 years ago as a graduate student, and he has been active in the field of Glycobiology ever since. Among his research accomplishments, he determined the minimal sequence requirements for N-linked glycosylation while a postdoc in William Lennarz's lab, and he collaborated with Paul Englund's group at Johns Hopkins to elucidate the biosynthetic pathway for GPI-anchors. In the early 1980s, while probing cells with glycosyltransferases, Hart's laboratory discovered cytoplasmic and nuclear protein glycosylation by O-linked *N*-acetylglucosamine (O-GlcNAc) (e.g., *Journal of Biological Chemistry* 259: 3308; *Journal of Biological Chemistry* 261: 8049). Since that time, the Hart laboratory has published nearly 200 papers on O-GlcNAcylation, identifying and cloning the enzymes controlling cycling, characterizing O-GlcNAcylation and its interplay with phosphorylation on hundreds of proteins, and they have developed many of the tools and methods in use today to study this modification. In

1989, Hart founded the leading journal in the field, *Glycobiology*, serving as Editor-in-Chief until 2001. Hart received the first International Glycoconjugate Organization (IGO) Award in 1997, the Karl Meyer Award from the Society for Glycobiology in 2006, and served as the 2009–2011 president of the IGO. Hart is currently an Associate Editor for *The Journal of Biological Chemistry* and an Associate Editor for *Molecular and Cellular Proteomics*. To date, Hart has published about 263 papers, all in the area of glycosciences. H factor=84.



Bruno Goud studied biochemistry and immunology at the Ecole Normale Supérieure de Cachan and the University of Paris. He received his PhD in 1981 under the mentorship of Jean-Claude Antoine and Stratis Avrameas at Institut Pasteur and did postdoctoral work under the mentorship of Peter Novick at Yale University. In 1995, he established an independent research group in the Department of Cell Biology at the Institut Curie in Paris. The main focus of his studies is the regulation of intracellular transport and membrane trafficking in eukaryotic cells. In particular, his group is working on the functions of Rab GTPases, the mechanisms that sustain the global organization of intracellular compartments and the functions of myosins in membrane traffic. Since 2000, he has also developed original *in vitro* approaches to unravel physical parameters such as membrane tension or membrane curvature that underlie transport processes.

Bruno Goud is a recipient of an ERC (European Research Council) Advanced grant (2013). He is a member of the European Molecular Biology Organization (EMBO). He has been the Head of the Department of Cell Biology of Institut Curie since 2003.



Graça Raposo received her PhD in 1989 in Membrane Biology and Immunology at the University of Paris VII where she specialized in electron microscopy and membrane biology. From 1990 to 1995 she was a postdoctoral fellow at the Immunology Center in Marseille and then in the Department of Cell Biology, Utrecht University, The Netherlands. She is the deputy Director of the Department of Cell Biology at Institut Curie and Director of the Training unit. Her major research interests focus on the biogenesis and functions of exosomes and lysosome related organelles with implications in neurodegenerative disorders, lysosomal diseases, and cancer. Her group have started to unravel the cellular and molecular mechanisms regulating the biogenesis of melanosomes, the lysosome related organelles of epidermal melanocytes, studies that open a new avenue to modulate pigmentation in health and disease. In 2012 she was awarded the CNRS Silver Medal and in 2013 the Descartes Huygens Price from the Royal Dutch Academy of Sciences. She is a member of the European Molecular Biology Organization (EMBO).



Alan Ezekowitz, MBChB, DPhil, FAAP, is Co-Founder, President, and CEO of Abide Therapeutics, a company cofounded by Professors Dale Boger and Ben Cravatt from the Scripps Research Institute. He was previously Senior Vice President and Franchise Head, Bone, Respiratory, Immunology, Endocrine, Dermatology and Urology at Merck Research Laboratories. At Merck, he was responsible for the overall scientific direction of the drug discovery and development process. Prior to Merck, Dr. Ezekowitz was Chief of Pediatric Services at the Massachusetts General Hospital for Children and the Partners Healthcare System and the Charles Wilder Professor of Pediatrics at the Harvard Medical School. He chaired the committee that led to the establishment of the Academy at the Harvard Medical School where he served as a Scholar and Founding member. He served on the Board of Directors of the Partners Healthcare System and Massachusetts General Physicians Organization. In 2008 he was honored with the establishment R. Alan Ezekowitz Professorship in Pediatrics at the Harvard Medical School.

He is a member of the American Society of Clinical Investigation, the Association of American Physicians, the American Pediatric Society, and a Fellow of AAAS. He served on NIH Subcommittees on Biodefense and Vaccine Adjuvants; NAS panels on antibiotic resistance and academic pharmaceutical partnerships; and the ARISE 2 task force of American Academy of Arts & Science that is evaluating the impact of federal and industrial funding of science, engineering, and medicine on American universities.

Dr. Ezekowitz received his MBChB and DPhil from University of Cape Town and Oxford University, respectively. At Children's Hospital in Boston, he was a postdoctoral fellow in the Division of Hematology and Oncology with clinical training in pediatrics. He was on the staff of the Children's Hospital of Boston prior to moving to the Massachusetts General Hospital.

He is a pioneer in the field of innate immunity and has over 150 publications. In particular his group played a major role in defining the structure function of the mannose binding lectin and the macrophage mannose receptor.



Douglas A. Lauffenburger is Ford Professor of Bioengineering, and cofounder and Head of the Department of Biological Engineering at MIT, with affiliate appointment in the Department of Biology; he was a founding codirector of the MIT Computational and Systems Biology Initiative in 2002. His major research interests are in cell engineering, with central focus on cell-cell communication and intracellular signal transduction, emphasizing predictive computational models derived from quantitative experiment. Lauffenburger has coauthored a monograph entitled *Receptors: Models for Binding, Trafficking & Signaling* (Oxford University Press, 1993), has coedited the book entitled *Systems Biomedicine: Concepts and Perspectives* (Elsevier, 2010), and has supervised more than 100 doctoral and post-doctoral students. He is a member of the National Academy of Engineering and the American Academy of Arts and Sciences, and has served as President of the Biomedical Engineering Society, Chair of the AIMBE College of Fellows, and on the NIH NIGMS Advisory Council, and coauthored the NRC report on *A New Biology for the 21st Century*.

SECTION EDITORS



Kenneth E Neet received his PhD in Biochemistry in 1965 (with Dr. Frank W. Putnam) from the University of Florida. He was a postdoctoral fellow at the University of California, Berkeley (with Dr. Daniel E. Koshland, Jr.), and joined the faculty of Case Western Reserve University (CWRU) in 1967 as an Assistant Professor of Biochemistry.

Dr. Neet received a Faculty Research Award of the American Cancer Society from 1968 to 1973 and was on sabbatical leave at the National Institute for Medical Research, Mill Hill, England (with Dr. N. Michael Green) 1973–1974. From 1978 to 1990, he was Professor of Biochemistry at CWRU. During 1980–1981, he took a sabbatical year as a Josiah Macy Faculty Scholar in the Department of Neurobiology, Stanford University (with Dr. Eric M. Shooter).

Dr. Neet moved to Rosalind Franklin University of Medicine and Science/The Chicago Medical School in 1990 and was Professor and Chair of Biochemistry & Molecular Biology until 2005. Dr. Neet became Associate Dean for Research of the Chicago Medical School, RFUMS in 2004.

Dr. Neet has served on the Editorial Boards of the *Journal of Biological Chemistry*, *Protein Science*, and *Molecular & Cellular Proteomics* and as a member of study sections for the National Institutes of Health and the National Science Foundation. He was an Associate Editor of the *Journal of Biological Chemistry* (1996–2013).

Dr. Neet's research interests have been in the general areas of protein–protein interactions, allosteric interactions, protein conformational changes, and cell signaling. In his earlier years he studied theoretical and experimental aspects of slow transitions in enzymes, particularly the cooperativity of the monomeric enzyme glucokinase. Later, he studied ligand–receptor interactions of nerve growth factor, establishing a mutational system to study the interactions with its receptors (TrkA and p75), and ultimately the complex signaling emanating from these receptors within neuronal systems.

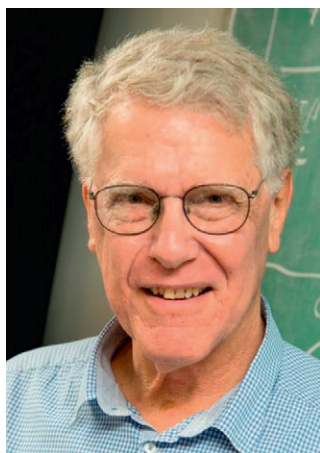


Sarah A Woodson is a biophysical chemist interested in how RNA molecules fold into specific three-dimensional structures and how they interact with proteins to turn genes on and off in the cell. She was born in Michigan, USA, and attended Kalamazoo College before receiving her PhD in Biophysical Chemistry from Yale University in 1987. After postdoctoral research in the laboratory of Tom Cech at the University of Colorado Boulder from 1987 to 1990, she joined the faculty of the University of Maryland in 1990 and the faculty of Johns Hopkins University in 1999, where she is currently the T.C. Jenkins Professor of Biophysics. Together with Mark Chance and Michael Brenowitz, she pioneered methods for visualizing how RNA molecules change shape in real time. She served on the board of the RNA Society and was elected an AAAS Fellow in 2011 and a Pew Scholar in the Biomedical Sciences in 1993. She is currently a reviewing editor of *Biopolymers* and *Journal of Molecular Biology* and serves on the editorial boards of *Nucleic Acids Research* and *RNA*.



Judith S Bond, Evan Pugh Emeritus Professor of Biochemistry and Molecular Biology at the Pennsylvania State University College of Medicine, was President of the Federation of American Societies for Experimental Biology (2012–13) and an Associate Editor of the *Journal of Biological Chemistry* (1999–2013). She received her BS degree from Bennington College in Vermont, her PhD from Rutgers University, and did postdoctoral research at Vanderbilt University. She rose through the academic ranks at the Medical College of Virginia, Virginia Commonwealth University, became Professor and Head of Biochemistry and Nutrition at Virginia Polytechnic Institute and State University, and served at Penn State University College of Medicine as Professor and Chair of Biochemistry and Molecular Biology from 1992 to 2012. At Penn State, she also served as assistant dean for graduate studies, codirector of the All-University Interdisciplinary Biological Sciences Program and founding director of the Medical Scientist Training Program. She recently directed NIH-funded summer research programs for high school students and teachers and for undergraduate students (particularly underrepresented minority groups). Her work on metalloproteases has been funded continuously by the NIH for over 35 years. She has

trained 5 Master's, 21 PhD students, and 19 postdoctoral trainees. Her professional service included member and chair National Institutes of Health Study Sections, member of the National Institute of Diabetes, Digestive and Kidney Diseases Advisory Council, member of the American Association of Medical Colleges – Howard Hughes Medical Institute Committee on the Scientific Foundations for Future Physicians, and President of the American Society for Biochemistry and Molecular Biology.



Elliot L. Elson received his AB at Harvard, his PhD at Stanford under the mentorship of R.L. Baldwin and did postdoctoral work at University of California, San Diego, under Bruno Zimm. His first independent faculty position was in the Department of Chemistry, Cornell University, where he stayed for 11 years. He then moved to the Department of Biological Chemistry (now the Department of Biochemistry and Molecular Biophysics) at Washington University in St. Louis, School of Medicine. He has pursued research in several biophysical areas. One of these is the development of Fluorescence Correlation Spectroscopy and Fluorescence Photobleaching Recovery and their application to studies of diffusion in cell membranes and of phase separation in model membrane bilayers as well as of other cellular and noncellular phenomena. He has also worked in the areas of cell mechanics, cell motility, and the mechanical properties of engineered tissues. His laboratory has developed approaches for measuring mechanical properties of cells in monolayer culture and for determining the contributions of cells and extracellular matrix to the mechanics of engineered heart and connective tissue constructs.



Tamotsu Yoshimori received his PhD degree in 1989 at Osaka University. After working at several places including European Molecular Biology Laboratory (Prof. Kai Simons' lab) and National Institute of Basic Biology (Prof. Yoshinori Ohsumi's lab), he is now a distinguished professor of Osaka University (Graduate School of Medicine and of Frontier Biosciences). His research interests are focused on intracellular membrane trafficking, and especially for the last 18 years, on autophagy. He identified LC3 as an autophagosome-binding protein, which has been widely used as the gold standard in autophagy assays. The paper has been cited over 3000 times. He also provided new insights into membrane biogenesis in autophagy and the role of autophagy in pathogen defense and suppression of various diseases. He authored or coauthored over 220 journal articles and book chapters. He is an editor of *Journal of Cell Science*, and on the editorial board of *Journal of Cell Biology*, *Molecular Biology of the Cell*, and so on. He is a member of the Faculty of 1000 and a vice president of Japan Society for Cell Biology. He was awarded Osaka University Presidential Award for 2012 and 2013, Prize for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology in 2013, Kakiuchi Saburo Memorial Prize by the Japanese Biochemical Society in 2014, and selected as Highly Cited Researchers 2014 by Thomson Reuters.



Paul A. Gleeson is currently Head of the Department of Biochemistry and Molecular Biology at the University of Melbourne, located at Bio21 Institute. His research interests include the molecular mechanisms of intracellular membrane transport and the molecular basis of organ-specific autoimmune diseases. Paul Gleeson obtained his PhD in 1980 from the University of Melbourne and did postdoctoral research in the biosynthesis and function of glycoproteins at the Hospital for Sick Children, Toronto, National Institute for Medical research, Mill Hill London and Department of Biochemistry, La Trobe University, Melbourne. He established an independent laboratory at Monash University in 1986 where he defined the targeting signals of Golgi glycosyltransferases, identified golgins of the *trans*-Golgi network and along with his colleagues developed mouse models of autoimmune gastritis. In 2001 he moved to Department of Biochemistry and Molecular Biology at the University of Melbourne and has been Head of the Department since 2006. He has a number of international research collaborations and has been a visiting scientist at the EMBL, Heidelberg, and the Institut Curie, Paris.

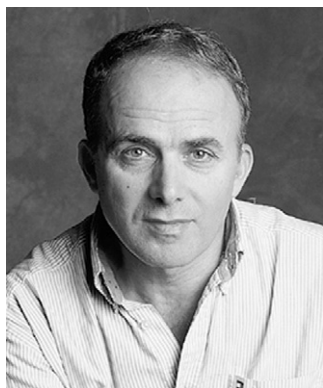


Anna Akhmanova studied biochemistry and molecular biology at the Moscow State University. She received her PhD in 1997 at the University of Nijmegen, the Netherlands. She worked as a postdoc at the Department of Microbiology and Evolutionary Biology at the University of Nijmegen and at the Department of Cell Biology at the Erasmus Medical Center in Rotterdam, the Netherlands. In 2001, she started her own research group at the Erasmus Medical Center. Since 2011, Anna Akhmanova is professor of Cell Biology at Utrecht University, the Netherlands.

Akhmanova studies cytoskeletal organization and trafficking processes, which contribute to cell polarization, differentiation, vertebrate development, and human disease. The main focus of her studies is the microtubule cytoskeleton. Her group has described key interactions between proteins associated with microtubule ends and determined their role in the regulation of microtubule organization and dynamics. She has also characterized mechanisms of bidirectional microtubule-based motility of membrane organelles such as cell nuclei and exocytotic vesicles. Her research relies on combining high-resolution live cell imaging and quantitative analysis of cytoskeletal remodeling, measurement of protein dy-

namics using advanced microscopic assays, *in vitro* reconstitution of dynamic cytoskeleton-based processes and different methods of identification of protein-protein interactions.

Anna Akhmanova is a recipient of the ALW Vernieuwingsimpuls VIDI (2001) and VICI awards (2007) (Netherlands Organization for Scientific Research), and an ERC Synergy grant (2013). She is the Chair of the board of the Netherlands Microscopy Society and a member of the European Molecular Biology Organization (EMBO).



Yosef (Yossi) Yarden is a Professor in the Department of Biological Regulation at Weizmann Institute of Science (Rehovot, Israel). Dr. Yarden obtained his PhD from Weizmann Institute (under the mentoring of Dr. Joseph Schlessinger) and later trained at both Genentech Inc., with Dr. Axel Ullrich, and at the Whitehead Institute (MIT), with Dr. Robert A. Weinberg. His group studies the roles played by growth factors and their receptors in cancer progression. They also explore strategies able to intercept growth factor signals in tumors.



Jason D Weber is an Associate Professor in the Departments of Internal Medicine and Cell Biology and Physiology at Washington University and is the Coleader of the Breast Cancer Research Program in the Siteman Comprehensive Cancer Center. Dr. Weber obtained his PhD from Saint Louis University and received postdoctoral training under the mentoring of Dr. Charles J. Sherr at St. Jude Children's Research Hospital in Memphis, TN. His group studies the molecular interplay between oncogenes and tumor suppressors, particularly focused on their role in regulating cellular growth processes.



Michael Dustin received his PhD in 1990 in Cell and Developmental Biology from Harvard University, where he worked in the laboratory of Timothy A. Springer, PhD. During his graduate work he was involved in the identification and cloning of ICAM-1 and ICAM-2, key ligands for the integrin LFA-1. He further demonstrated that LFA-1 dependent adhesion was stimulated by T cell receptor signaling. He did his postdoctoral training with Stuart Kornfeld at Washington University School of Medicine in St. Louis, focusing on mechanisms of lysosome biogenesis. He started his independent lab also at Washington University School of Medicine, and used supported lipid bilayers to define the first protein to modulate organization of the immunological synapse- CD2 associated protein- and the dynamics of immunological synapse formation. He moved to NYU School of Medicine in 2001 where he applied *in vivo* microscopy to tolerance induction and immune responses in effector sites including the liver, brain, and spleen. Continued work with the immunological synapse model resulted in discovery of signaling microclusters, the molecular basis of immunological synapse stability, and the direct budding of extracellular microvesicle enriched in T cell receptors in the immunological synapse. Professor Dustin recently took a position at the

University of Oxford with support of a Wellcome Trust Principal Research Fellowship. The focus of his new post is the translation of the immunological synapse. He received a Presidential Early Career Award in Science and Engineering and the DART-NYU Biotechnology Achievement Award.



David A. Sassoon directs a research team at Paris Sorbonne. He received his PhD from Columbia University (NYC, USA) in 1986 in the biological sciences and was a professor at Boston University Medical School and Mt. Sinai Medical School before relocating to Paris in 2006. His long-standing interests are in developmental and stem cell biology in a number of tissues including skeletal muscle, skin, CNS, and heart. Dr. Sassoon recently concluded the coordination of a EC-funded 15 partner international consortium (Endostem) designed to identify novel therapeutics for mobilizing endogenous stem/progenitor cells (<http://www.endostem.eu/>) and is a recent recipient as coordinator of a multipartner Transatlantic Network of Excellence grant from the Fondation Leducq focused on cardiac stem cell biology (<http://www.fondationleducq.org/nivel2.aspx?idsec=1360>).

A major focus of his research is on adult progenitor/stem cell biology and how these cells respond to stress. His team has identified a parentally imprinted gene, PW1/Peg3, which is involved in both p53 and inflammatory responses. PW1 is expressed in adult stem cells in all tissues identified to date. Loss of PW1 function in stem cells results in a loss of stem cell competence including a reduced capacity to undergo self-renewal and respond to hypoxic stress. We are currently evaluating a role for PW1 in postnatal and adult heart with a particular focus on its role in vessel precursors using both myocardial infarction and stress-induced myocardial hypertrophy.



Jason M. Haugh is a Professor of Chemical and Biomolecular Engineering at North Carolina State University in Raleigh, NC, USA, where he has been on the faculty since 2000. His laboratory has been among those to pioneer the synthesis of quantitative experiments and modeling to study signal transduction in mammalian cells. Since the lab's inception, their approach has combined biochemical measurements, live-cell fluorescence microscopy, and computational modeling to elucidate signaling mechanisms by analyzing their kinetics and spatial organization in cells. The systems studied by the Haugh Laboratory include: regulation of the phosphoinositide 3-kinase, Ras/extracellular signal-regulated kinase, and phospholipase C pathways mediated by receptor tyrosine kinases, and cross talk between these pathways; dynamic organization of multimolecular complexes at cell membranes; signaling mediated by cytokine and chemokine receptors in immune cells; and integration of adhesion, signaling, and cytoskeletal dynamics that direct cell migration.



After graduating with an MA in Mathematics from the University of Cambridge, Helen Mary Byrne received her Master of Science and DPhil from the University of Oxford. She worked as a postdoctoral researcher at the Universities of Oxford and Bath, before taking up a lectureship at the UMIST in Manchester. She moved to the University of Nottingham in 1998 and was awarded a prestigious Advanced Research Fellowship from the United Kingdom's Engineering and Physical Sciences Research Council (2000–2006). She was promoted to Reader in 2002 and Professor in 2003. While in Nottingham, she established and then led the Centre for Mathematical Medicine and Biology until she returned to Oxford in 2011 where she is now based in the Mathematical Institute at the University of Oxford. She has over 20 years' experience of developing, analyzing, and simulating continuum and multiscale models of biomedical systems, and a particular interest in studying the growth and response to treatment of solid tumors.



Rune Linding completed his PhD at the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany, followed by postdoctoral training at EMBL. He then jointly trained with professors Tony Pawson and Mike Yaffe at the Lunenfeld at Mount Sinai Hospital in Toronto, Canada, and the Massachusetts Institute of Technology (MIT) in Cambridge, US, respectively. Dr. Linding then established his own laboratory of Cellular and Molecular Logic at the Institute of Cancer Research (ICR) in London, UK, before returning to Denmark to take a position as professor of cellular signal integration at the Technical University of Denmark. In 2014, Dr. Linding moved his laboratory to the Biotech Research and Innovation Centre (BRIC) at University of Copenhagen where he is currently professor of cellular signaling. His research group focuses on big data network biology, exploring biological systems by developing and deploying algorithms aimed to predict cell behavior, in particular looking at cellular signal processing and decision making. A strategic focus is to continue to develop computational tools (such as KinomeExplorer, NetworKIN, and NetPhorest) and to deploy these on genome-scale quantitative data obtained by, for example, mass spectrometry, genomic, and phenotypic screens to understand the principles of how spatio and temporal assembly of mammalian signaling networks transmit and process information at a systems level in order to alter cell behavior. His overarching aim is to advance network

medicine by identifying and targeting signaling networks associated with complex diseases. To this end Dr. Linding is currently leading high-level, strategic, multidisciplinary studies of signaling network dynamics driving cancer metastasis in collaboration with other labs at Harvard, Yale, The Jackson Laboratory, Memorial Sloan Kettering Cancer Center, MIT, and BRIC.

CONTRIBUTORS TO VOLUME 4

R Aebersold

ETH Zurich, Zurich, Switzerland and University of Zurich, Zurich, Switzerland

MR Akins

Drexel University, Philadelphia, PA, USA

CC Antonovici

Leiden University, Leiden, The Netherlands; Institute of Biology Leiden, Leiden, The Netherlands; and Centrum Wiskunde and Informatica, Amsterdam, The Netherlands

P Blattmann

ETH Zurich, Zurich, Switzerland

N Blüthgen

Charité – Universitätsmedizin Berlin, Berlin, Germany and Humboldt-Universität zu Berlin, Berlin, Germany

SEM Boas

Leiden University, Leiden, The Netherlands and Centrum Wiskunde and Informatica, Amsterdam, The Netherlands

S Borinskaya

University of Connecticut Health Center, Farmington, CT, USA

O Britton

University of Oxford, Oxford, UK

L Brusch

Technische Universität Dresden, Dresden, Germany

A Bueno-Orovio

University of Oxford, Oxford, UK

HM Byrne

University of Oxford, Oxford, UK

MA Calderwood

Dana-Farber Cancer Institute, Boston, MA, USA and Harvard Medical School, Boston, MA, USA

A Califano

Columbia University Medical Center, New York, NY, USA

BT Castle

University of Minnesota, Minneapolis, MN, USA

B Charlotiaux

Dana-Farber Cancer Institute, Boston, MA, USA and Harvard Medical School, Boston, MA, USA

V Cristini

University of New Mexico, Albuquerque, NM, USA and King Abdulaziz University, Jeddah, Saudi Arabia

ME Cusick

Dana-Farber Cancer Institute, Boston, MA, USA and Harvard Medical School, Boston, MA, USA

N Daneshi

The University of Newcastle, Callaghan, NSW, Australia

TS Deisboeck

ThinkMotu LLC, Wellesley, MA, USA

I Dransfield

University of Edinburgh, Edinburgh, UK

JL Dunster

University of Reading, Reading, UK

AVG Edwards

Protein Research Group, Campusvej, Odense M, Denmark

J Ferkinghoff-Borg

University of Copenhagen (UCPH), Copenhagen, Denmark

L Foret

Ecole Normale Supérieure, Paris, France

E Fraenkel

Massachusetts Institute of Technology, Cambridge, MA, USA

ADR Garcia

Drexel University, Philadelphia, PA, USA

SJ Harrison

Danmarks Tekniske Universitet, Hørsholm, Denmark

JM Haugh

North Carolina State University, Raleigh, NC, USA

M Heidari

The University of Newcastle, Callaghan, NSW, Australia

H Hondermarck

The University of Newcastle, Callaghan, NSW, Australia

PA Iglesias

The Johns Hopkins University, Baltimore, MD, USA

N Jaikhani

*Dana-Farber Cancer Institute, Boston, MA, USA;
Harvard Medical School, Boston, MA, USA; and
Massachusetts Institute of Technology, Cambridge, MA,
USA*

KA Janes

University of Virginia, Charlottesville, VA, USA

LJ Jensen

University of Copenhagen, Copenhagen, Denmark

AA Jermusyk

North Carolina State University, Raleigh, NC, USA

DM Johnstone

University of Sydney, Sydney, NSW, Australia

F Jülicher

*Max Planck Institute for the Physics of Complex
Systems, Dresden, Germany*

R Ketteler

University College London, London, UK

MI Koeva

*Whitehead Institute for Biomedical Research,
Cambridge, MA, USA and Massachusetts Institute of
Technology, Cambridge, MA, USA*

J Kriston-Vizi

University College London, London, UK

MR Larsen

*Protein Research Group, Campusvej, Odense M,
Denmark*

DA Lauffenburger

*Massachusetts Institute of Technology, Cambridge, MA,
USA*

S Legewie

*Institut für Molekulare Biologie gGmbH (IMB), Mainz,
Germany*

R Linding

*University of Copenhagen (UCPH), Copenhagen,
Denmark*

LM Loew

*University of Connecticut Health Center, Farmington,
CT, USA*

M Long

The University of Chicago, Chicago, IL, USA

J Longden

University of Copenhagen, Lyngby, Denmark

K Luck

*Dana-Farber Cancer Institute, Boston, MA, USA and
Harvard Medical School, Boston, MA, USA*

O Marchenko

*University of Connecticut Health Center, Farmington,
CT, USA*

RMH Merks

*Centrum Wiskunde and Informatica, Amsterdam, The
Netherlands and Leiden University, Leiden, The
Netherlands*

EA Milward

*The University of Newcastle, Callaghan, NSW,
Australia*

A Muszkiewicz

University of Oxford, Oxford, UK

DJ Odde

University of Minnesota, Minneapolis, MN, USA

N Perrimon

*Harvard Medical School, Boston, MA, USA and
Howard Hughes Medical Institute, Boston, MA, USA*

GT Reeves

North Carolina State University, Raleigh, NC, USA

EG Rens

*Leiden University, Leiden, The Netherlands and
Centrum Wiskunde and Informatica, Amsterdam, The
Netherlands*

B Rodriguez

University of Oxford, Oxford, UK

T Rolland

*Dana-Farber Cancer Institute, Boston, MA, USA and
Harvard Medical School, Boston, MA, USA*

HM Sauro

University of Washington, Seattle, WA, USA

EM Schoof

University Health Network, Toronto, ON, Canada

A Shahandeh

*The University of Newcastle, Callaghan, NSW,
Australia*

R Sopko

Harvard Medical School, Boston, MA, USA

G Stolovitzky

*Icahn School of Medicine at Mount Sinai, New York,
NY, USA and IBM Research, Yorktown Heights, NY,
USA*

H Tahir

*VU Medical Centre, Amsterdam, The Netherlands and
Centrum Wiskunde and Informatica, Amsterdam, The
Netherlands*

KA Thomas

Trefoil Therapeutics, Chatham, NJ, USA

K Tsafou
University of Copenhagen, Copenhagen N, Denmark

IM Ventura
*The University of Chicago, Chicago, IL, USA and
Ministry of Education of Brazil, Brasília, Brazil*

FJ Vermolen
Delft University of Technology, Delft, The Netherlands

M Vidal
*Dana-Farber Cancer Institute, Boston, MA, USA and
Harvard Medical School, Boston, MA, USA*

CC Wang
University of Virginia, Charlottesville, VA, USA

Z Wang
University of New Mexico, Albuquerque, NM, USA

R Wedlich-Soldner
University of Muenster, Muenster, Germany

CONTENTS OF VOLUME 4

Preface	xxi
Cell Biology: An Overview <i>RA Bradshaw and PD Stahl</i>	xxiii

Systems Cell Biology

Systems Cell Biology: An Overview <i>DA Lauffenburger, JM Haugh, HM Byrne, and R Linding</i>	1
--	---

Dynamic Integration

Dynamics

Dynamics of Gradient Sensing and Chemotaxis <i>PA Iglesias</i>	4
Modeling Actin Dynamics <i>S Borinskaya, O Marchenko, and LM Loew</i>	10
Metabolism <i>HM Sauro</i>	24
Dynamics of Microtubule Self-Assembly <i>BT Castle and DJ Odde</i>	36
Dynamics of Protein Kinase Cascades <i>N Blüthgen and S Legewie</i>	44
Cdc42 and the Mechanisms of Yeast Cell Polarization – A Paradigm for Mesoscale Systems Biology <i>R Wedlich-Soldner</i>	51
Theory of Cargo and Membrane Trafficking <i>L Foret, L Brusch, and F Jülicher</i>	56
Transcription Factor Networks <i>AA Jermusyk and GT Reeves</i>	63

Vertical Integration

Modeling

Vertical Integration <i>HM Byrne</i>	72
Computational Approaches for Multiscale Modeling <i>Z Wang, V Cristini, and TS Deisboeck</i>	80
Cardiac Modeling <i>A Bueno-Orovio, O Britton, A Muszkiewicz, and B Rodriguez</i>	86

Applications

Mathematical Approaches to Studying Inflammation <i>JL Dunster and I Dransfield</i>	95
Angiogenesis <i>KA Thomas</i>	102
A Review on Various Mathematical Modeling Approaches for Wound Healing <i>FJ Vermolen</i>	117
Multiscale Analysis of Morphogenesis <i>CC Antonovici, SEM Boas, EG Rens, H Tahir, and RMH Merks</i>	122
Neurogenesis in the Adult Brain <i>MR Akins and ADR Garcia</i>	134

Horizontal Integration**Networks**

Understanding of 'Networks' <i>In Vitro</i> and/or <i>In Vivo</i>	<i>EM Schoof, J Ferkinghoff-Borg, and R Linding</i>	141
---	---	-----

Omics

Connecting Evolutionary Genomics to Cell Biology	<i>IM Ventura and M Long</i>	153
Transcriptomics	<i>EA Milward, A Shahandeh, M Heidari, DM Johnstone, N Daneshi, and H Hondermarck</i>	160
The Advent of Mass Spectrometry-Based Proteomics in Systems Biology Research	<i>P Blattmann and R Aebersold</i>	166
Phosphoproteomics: Approaches, Developments, and Challenges	<i>MR Larsen and AVG Edwards</i>	177
Interactomes – Scaffolds of Cellular Systems	<i>K Luck, N Jailkhani, ME Cusick, T Rolland, MA Calderwood, B Charlotaux, and M Vidal</i>	187
Metabolomics in Cell Biology	<i>SJ Harrison</i>	199

Analyses and Tools

Network Modeling of Heterogeneous Datasets	<i>E Fraenkel and MI Koeva</i>	211
Stochastic Analysis of Nongenetic Cell-to-Cell Heterogeneity	<i>CC Wang and KA Janes</i>	218
Systematic Methods to Interrogate Genetic Perturbations and Map Phosphorylation-Dependent Signaling	<i>R Sopko and N Perrimon</i>	227
High-Content Screening in Cell Biology	<i>R Ketteler and J Kriston-Vizi</i>	234
Integrative Systems Biology	<i>K Tsafou and LJ Jensen</i>	245

Disease

Computational and Systems Cancer Biology	<i>A Califano and G Stolovitzky</i>	252
Drug Targeting	<i>J Longden</i>	264
Index		269

PREFACE

Cell biology is the study of cells, the integral unit that is the basis of all living tissues. In its earliest phases, this area of investigation was largely defined by microscopic-based observations of structure and organization that were, by their nature, largely descriptive. This also served to delineate cell biology from its nearest neighbor, biochemistry, which, dating from its post WWII renaissance, was driven in large part by reductionist approaches, where tissues, cells, and organelles were broken down in order to isolate and then characterize individual components. The melding force that bridged these two areas was molecular biology that allowed the cell biologist to transform pictures into functions and allowed the biochemist to reassemble their components into ensembles and subcellular entities. Not surprisingly the boundary between these two key biological sciences has become blurred beyond recognition. Today a considerable part of what passes as cell biology is by any other name biochemistry (or molecular biology – the distinction between these two areas being equally indistinct) and vice versa. Indeed, one of the first challenges we faced in organizing this Encyclopedia was how much molecular detail to include. It was our conclusion that for the sake of completeness, and to appropriately introduce the volumes Organizational Cell Biology and Functional Cell Biology, we needed a description of the molecular components of the cell, particularly with respect to the synthesis and degradation of proteins and nucleic acids, along with outlines of important chemical principles and key methodology. We were guided to some degree in these decisions by input from undergraduate and graduate teachers and the contents of their introductory courses that they shared with us and by discussions with our volume and section editors. The result was the first volume which we recognize is very 'biochemical,' but certainly lacking the more detailed coverage of the *Encyclopedia of Biological Chemistry* (Lennarz and Lane, 2004) or the many excellent biochemical texts that dot the educational landscape. Readers who find 'gaps' or incomplete treatments in this section will likely be able to find the additional detail they seek from these sources.

The second major decision that we had to confront was whether we would attempt to cover the full range of living organisms or whether we would place some limits on the scope of the material we included. In the final result, it was decided to place the emphasis on higher eukaryotes with only very prescribed treatments of lower eukaryotes and prokaryotes. Volumes 2 and 3 thus deal with material that has perhaps been more traditionally thought of as cell biology. Again we were guided in our decisions about what to include by teaching considerations, brainstorming sessions with our editors, and a perceived need to separate cellular organization from cellular function. However, as even a quick perusal will reveal, this is not a sharp delineation either. Volume 2 begins with a methodological treatment of imaging, which is basically divided between light and electron microscopic applications, and is followed by sections on organelles, interorganellar communication, and the cytoskeleton (including motors). The

last part of volume 2 is devoted to intracellular infection and covers a variety of external agents and pathogens that impact cell function as well as human pathology. Volume 3, in contrast, largely deals with major cell functions: signaling, cell-cycle control, and apoptosis. The last two parts are devoted to cells of the immune system and their responses and stem cells – two highly specialized niches characterized by unique cellular involvement.

The last part of the Encyclopedia deals with systems biology. This is really the newest area of cell biology and considers cellular involvement as part of a phenotypic response. As such, it strives to integrate all of the levels of the first three volumes in a consistent manner to produce a seamless description of cellular function. This is clearly a newly evolving area and the most rapidly changing part of cell biology; indeed it is where we expect to see the greatest change in the next few years.

Assembling a treatment of a topic as large as cell biology had multiple challenges. Coverage was a significant problem: on the one hand it was almost impossible to avoid some redundancy in places while, on the other, there were inevitable gaps. Some of these arose from late cancellations; others from oversights on our part. We can only promise to fill these in future editions. We also note that as can be expected for a large multiauthor compilation the individual articles do differ in detail and treatment. We felt it was more important to allow our experts substantial latitude in deciding how to present their topics than to apply rigid guidelines. We did try to limit the number of references (to make it easier for people not familiar with a topic to make the transition to more extensive articles) but there is admittedly some considerable variation in the bibliographic material as well.

When we were well into the project, we were approached by Graham Nisbet (Elsevier) about including the Encyclopedia in a larger collection entitled the Reference Module in Biomedical Sciences. The Reference Module plan was grouped in single integrated work information from several other compendia and it was our view that this would greatly leverage the value of our efforts (and those of our contributors). While we were the 'new kid on the block' in that we were still compiling the Encyclopedia and the other collections to be included were already extant, we felt that it was too good an opportunity to pass up. The Reference Module has since been launched and although it does not yet contain the complete content from the Encyclopedia, this will be achieved in the not too distant future. We consider that this will be a particularly valuable resource for researchers that are just entering (or changing into) this field.

No work of this size could be completed without the help and advice of many people. In this regard we would like to thank a number of people who were instrumental in bringing this project to fruition. First and foremost we would like to recognize our volume editors: Jerry Hart, Graça Raposo, Bruno Goud, Alan Ezekowitz, and Doug Lauffenburger; and our section editors: Ken Neet, Sarah Woodson, Judy Bond,

Elliot Elson, Tamotsu Yoshimori, Paul Gleeson, Anna Akhmanova, Yosef Yarden, Jason Weber, Michael Dustin, David Sassoon, Jason Haugh, Helen Byrne, and Rune Linding; and thank them for their invaluable input, enthusiasm, and hard work. They are truly a global group and have been a great team to work with. We also wish to thank the staff at Elsevier including Janice Audet, Peter Labella, Nicky Carter, Karen Maloney, Erin Hill-Parks, Rashmi Phadnis, Blerina Osmanaj, and Ginny Mills for their many efforts, perseverance (particularly with our idiosyncrasies), and skillful management of every aspect of this project.

In the course of planning this work, we gathered information from a number of sources. At the outset the publishers polled some 167 faculty from both undergraduate and graduate schools to ascertain their perceived interest in such a project, which also garnered considerable input about potential topics. The response was highly favorable. We also talked with many colleagues about experiences at their institutions and would like to particularly thank Catherine Bevier (Colby College), Robert Simoni (Stanford University), Larry Marsh (UC Irvine), and John Cooper (Washington University) for providing syllabi of germane courses taught in their institutions. Special thanks also go to Ruedi Abersold (ETH, Zurich) and Sergio Grinstein (University of Toronto) for their expertise that they readily shared with us, which was invaluable in planning certain sections.

Finally we would like to recognize all of our authors. We have been extraordinarily fortunate in attracting individuals from all around the globe to take time from their busy schedules to prepare this splendid set of contributions. Without these efforts there would be no Encyclopedia. Under ordinary circumstances it would be appropriate to dedicate the work to them and all the colleagues whose works they wrote about. However, during the preparation of this work, we sadly lost one of the most important contributors to cell biology, Tony Pawson. Tony was a true pioneer in elucidating basic mechanisms of how cells transmit information intracellularly, a fundamental knowledge that permeates all of cell biology. In an effort that was initiated by one of the section editors, Rune Linding, it was decided to dedicate the entire work to his memory as recognition of his contributions and a lasting tribute to a man who left us too soon.

We hope that the Encyclopedia will be of value to students and researchers alike and will become a useful tool in the training of future generations of scientists.

Ralph A Bradshaw and Philip D Stahl

References

Lennarz, W.J., Lane, M.D. (Eds.), 2004. Encyclopedia of Biological Chemistry. Oxford: Elsevier.

Cell Biology: An Overview

RA Bradshaw, University of California, Irvine, CA, USA

PD Stahl, Washington University School of Medicine, St. Louis, MO, USA

© 2016 Elsevier Inc. All rights reserved.

All living organisms can be divided into three principal categories: archaeobacteria, eubacteria, and eukaryotes; although they differ in structure and organization, they are all composed of cells as the fundamental life unit. At the molecular level, there is also a great deal of similarity in the basic materials that make up these entities because they use the same kinds of molecules to store and reproduce information, to run the cellular metabolism and machinery and to provide the structural framework. Thus nucleic acids, proteins, lipid membranes, and carbohydrates – alone and in various combinations – are universally present, albeit in distinguishable forms, along with innumerable metabolites and ions. There are components that are apparently essential for life and are found in one form or another in all species and there are many unique moieties and associated activities that are highly specialized and are found in relatively few organisms. Indeed, the similarities have underpinned the development of our understanding of cellular function at a rudimentary level and the differences, basically engendered by evolution, have illustrated and delineated the complexity that speciation has introduced. Perhaps the largest of these differences is that which separates single cell organisms from multicellular organisms. The latter are exclusively eukaryotes while the former are composed of both eukaryotes and prokaryotes. The cellular organization and architecture that distinguishes these two major life forms is striking; although cell biology correctly embraces both, traditionally prokaryotic organisms have been the province of the microbiologists and the majority of cell biology research has been devoted to the eukaryotic world. In practical terms this translates for the most part into the study of human cells and those of easily maintained laboratory animals and selected paradigms, for example, fruit flies, worms, and zebra fish.

Human and animal cell biology is not a tightly proscribed science with well-defined borders. Basically it serves at the interface between biochemistry, molecular biology, and genetics, on the one hand, and anatomy and physiology, on the other. The continuum of these disciplines forms the core of the biomedical sciences, which also include the related but separate fields of pharmacology, microbiology, immunology, and pathology that provide the connections to disease and health. Cell biology has strong connections to all of these. There are also specialized areas, for example, neuroscience, that are of such importance that they warrant their own category and the cell biology associated with them is also highly specialized. Thus, cell biology is as complex as the enormous variety of cells that exist and achieving an accurate description of all of them in terms of their components and functions has long been a major part of the research in this field.

Imaging and Organelle Organization

Among the developments that propelled cell biology into the modern era are the introduction of the ultracentrifuge and

the rapid advances in microscopy both electron and light microscopy – the former allowing investigators to fractionate and characterize the components of the cell and the latter to literally see them – either *in situ* or in isolated form. This progress is illustrated by a series of Nobel Prizes starting in the early 1970s that chronicle the grand confluence of classical biochemistry and general physiology that created modern cell biology (Claude *et al.*, 1974) or the discovery and characterization of intracellular organelles – lysosomes and endoplasmic reticulum (ER) among others (Mitchell, 1978), for the chemiosmotic hypothesis (Brown and Goldstein, 1985), for endocytosis (Ciechanover *et al.*, 2004), for ubiquitination and protein degradation, and just recently (Rothman *et al.*, 2013), for vesicle trafficking, and (Betzig *et al.*, 2014) for advances in light microscopy. The early cell biologists insisted on quantitative application of these new techniques and laid the groundwork for modern cell biology. As with biochemistry, reductionism has been the keystone for the amazing success over these past several decades. Each organelle has its own research history – the nucleus, mitochondria, peroxisomes, the proteasome, the ER, and cytoskeleton. The Golgi apparatus/secretory pathway and the endosome–lysosome network have all been examined in detail both in isolation and in their relationships with each other. A new member, the exosome (aka extracellular vesicle) has emerged more recently and stimulated the imagination of aspiring young investigators (Harding *et al.*, 2013). The rise of genomics, the development of spectacular imaging modalities, and evolving biochemical techniques (proteomics) have led to the discovery of molecular motors, the identification of macromolecular complexes such as the exocyst, ESCRT, GARP, SNAREs among others that choreograph complex intracellular pathways, including cell motility and cytokinesis, have brought cell biology into this new era where the whole is now seen as larger than the individual components. This creates the segue from Parts 1 and 2 of the *Encyclopedia (Molecular Cell Biology and Organizational Cell Biology)* to Part 3 (*Functional Cell Biology*) where signaling modalities will be seen as an integrative force for regulation and control.

Signaling

While all organisms can sense their environment and respond to cues from it, multicellular organisms must in addition coordinate their responses, which require intercellular communication at a sophisticated level (Bradshaw and Dennis, 2009). The higher the development, the more complex these communication systems become. Thus the cell biologist must focus not only on how molecular function is translated into cell organization and how these functions are coordinated from organelle to organelle but also on the external interactions and signals that control the larger functional responses of organs and ultimately organisms.

Intercellular communication is afforded by stimuli that can be transmitted across the cell membrane, either directly or via membrane bound molecules that in turn pass signals to the intracellular compartment. The origins of these stimuli may be quite diverse. Among other things, they can include cell–cell contacts, soluble factors/messengers, or foreign agents. Lipophilic molecules can cross the membrane by diffusion or by facilitated transport to recognize and bind to intracellular entities but most substances bind to membrane bound proteins and induce their signal by ‘activating’ them instead. These so-called receptors are capable of producing a variety of responses that invariably involve the generation of posttranslational modifications of preexisting proteins. Protein phosphorylation, which in eukaryotes mostly occurs on tyrosine, serine, and threonine residues, is highly prevalent and appears to be directly or indirectly involved in almost all transmembrane signaling (Gnad *et al.*, 2011). However, it is by no means the only vehicle for transmitting information as acetylation, methylation, monoglycosylation, and ubiquitination, to name a few, are both important and widespread. There are over three hundred different covalent modifications (Khoury *et al.*, 2011) that are introduced into proteins (and many more that have not yet been chemically defined) of both a transient and stable nature and most are likely to be involved to some extent in signal transduction processes. In addition, limited proteolysis can also be an important part of a pathway and this activity is a major player in cellular responses (Turk *et al.*, 2012).

The induction of an intracellular signal is basically perpetrated by the formation of new interactions between the modified proteins and other entities, which can range from small molecules to macromolecular protein complexes. One of the most important advances in understanding how cells process information induced from external stimuli was the appreciation that the newly formed sites produced by the protein modifications were recognized by other proteins with modules specifically designed for this purpose (Pawson, 1995). Thus, for example, phosphotyrosine residues could be recognized by other proteins containing a domain, termed SH2 for its relatedness to a similar domain found in the Src protein kinase, which could then be further modified allowing for additional interactions to occur. By these ‘docking’ events, signaling complexes could be assembled, often in multiple steps, that ultimately lead to the activation of key effectors. The end point of many of the pathways is the activation of transcription factors that lead to the modulation of gene expression of that cell by ultimately changing its protein expression profile (Bradshaw and Dennis, 2009). However, many molecules are usually modified and activated ‘along the way’ and these ‘new’ activities also add to the overall response to the original stimuli. Short term responses indeed require that the protein effectors necessary for it are already present; long term responses per force require new protein synthesis.

Signal transduction is thus dependent on two phenomena: posttranslational modifications, particularly of the readily reversible type, and protein–protein interactions. The extent to which these two activities take place, even in resting, unstimulated cells was greatly underestimated before the advent of proteomics and the introduction of mass spectrometric methodology into cell biological research (Bradshaw and

Burlingame, 2005). These high-throughput unbiased analyses of very complex mixtures, basically derived directly from cell lysates, revealed that essentially every protein had multiple interacting partners and that posttranslational modifications affected a very significant proportion of the proteins present. These were profound differences from what had been the prevailing wisdom and amounted to a paradigm shift in thinking about cellular organization and structure. As noted above, these same tools have gone on to cast new light on organelle organization in terms of resident proteins and have helped to disclose the structure of important cellular machines, such as the proteasome (Voges *et al.*, 1999), the nuclear pore (Routa *et al.*, 2000), and transcription complexes (Kornberg, 2007). They have also begun to elucidate the complexity of epigenetic regulation of gene transcription at the histone level (Cosgrove, 2012), which will also be essential in completing our understanding of signaling processes.

Concluding Remarks

Two of the most fundamental aspects of cells are their ability to reproduce themselves and, in higher organisms, to undergo changes that lead to new cell types and functions. These developmental processes leading to differentiation are what allow a single fertilized egg to form a complex adult organism and require the synthesis of all aspects of cell biology to understand. Knowing how cells go from one state to another in a timed and regulated fashion forms the core of developmental biology, which can be viewed as a sub discipline of cell biology. Also of major importance is the maintenance of viability and the associated turnover processes. The control of cell death is not only an essential part of development it is also key to managing situations that have gone awry and underlies many serious disease conditions.

It is clear that science is still a long way from understanding how even simple cells work. There is a long list of functions and structures that remain to be elucidated and integrating the various experimental approaches and the data they produce still lags far behind. This is the ‘Era of Big Data’ and vast amounts of new information that impact on our understanding of cells and their processes are collected every day. Genomic information is a good example – despite the rapidity that this sort of information can now be obtained, it still remains to be effectively mined in terms of what it can tell us about basic principles of how cells work. There has been an understandable pressure to apply this information to managing disease (Collins and Varmus, 2015), particularly those that are life threatening, but there certainly is information that is yet to be formulated that would apply to fundamental problems as well. The same can be said for transcriptomics, proteomics, and metabolomics. In addition to these powerful new technologies, singular advances in analyzing single cells promises to augment these molecular approaches significantly (Di Palma and Bodenmiller, 2015).

One aspect of cell biology that is looking to address the challenges of integrating these relatively newly minted studies is systems biology (Part 4). It is far too early to assess the value of these efforts in coordinating this flood of information but it certainly looks promising. It will be an elusive goal for some

time to come but also a stimulus for new generations of cell biologists that will be forthcoming.

References

- Betzig, E., Hell, S.W., Moerner, W.E., 2014. The Nobel Prize in chemistry — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2014/press.html (accessed 03.04.15).
- Bradshaw, R.A., Burlingame, A.L., 2005. From proteins to proteomics. *IUBMB Life* 57, 267–272.
- Bradshaw, R.A., Dennis, E.A. (Eds.), 2009. *Handbook of Cell Signaling*, second ed., vol. 1–3. San Diego, CA: Elsevier Academic Press, pp. 1–3047.
- Brown, M.S., Goldstein, J.L., 1985. Physiology or medicine — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1985/press.html (accessed 03.04.15).
- Ciechanover, A., Hershko, A., Rose, I., 2004. The Nobel Prize in chemistry — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/2004/press.html (accessed 03.04.15).
- Claude, A., de Duve, C., Palade, G.E., 1974. Physiology or medicine — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1974/press.html (accessed 03.04.15).
- Collins, F.S., Varmus, H.A., 2015. New initiative on precision medicine. *New England Journal of Medicine* 372, 793–795.
- Cosgrove, M.S., 2012. Writers and readers: Deconvoluting the harmonic complexity of the histone code. *Nature Structural & Molecular Biology* 19, 739–740.
- Di Palma, S., Bodenmiller, B., 2015. Unravelling cell populations in tumors by single-cell mass cytometry. *Current Opinion in Biotechnology* 31, 122–129.
- Gnad, F., Gunawardena, J., Mann, M., 2011. PHOSIDA 2011: The posttranslational modification database. *Nucleic Acids Research* 39 (suppl. 1), D253–D260.
- Harding, C.V., Heuser, J.E., Stahl, P., 2013. Exosomes: Looking back three decades and into the future. *Journal of Cell Biology* 200, 367–371.
- Khoury, G.A., Baliban, R.C., Floudas, C.A., 2011. Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database. *Scientific Reports* 1, 90. doi:10.1038/srep00090.
- Kornberg, R.D., 2007. The molecular basis of eukaryotic transcription. *Proceedings of the National Academy of Sciences of the United States of America* 104, 12955–12961.
- Mitchell, P., 1978. The Nobel Prize in chemistry — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1978/press.html (accessed 03.04.15).
- Pawson, T., 1995. Protein modules and signaling networks. *Nature* 373, 573–580.
- Rothman, J.E., Schekman, R.W., Südhof, T.C., 2013. The 2013 Nobel Prize in physiology or medicine — Press release. Nobelprize.org. Nobel Media AB 2014. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/2013/press.html (accessed 03.04.15).
- Rout, M.P., Aitchison, J.D., Supratka, A., *et al.*, 2000. The yeast nuclear pore complex — Composition, architecture, and transport mechanism. *Journal of Cell Biology* 148, 635–652.
- Türk, B., Türk, D., Türk, V., 2012. Protease signalling: The cutting edge. *EMBO Journal* 31, 1630–1643.
- Voges, D., Zwickl, P., Baumeister, W., 1999. The 26S proteasome: A molecular machine designed for controlled proteolysis. *Annual Review of Biochemistry* 68, 1015–1068.

Systems Cell Biology: An Overview

DA Lauffenburger, Massachusetts Institute of Technology, Cambridge, MA, USA

JM Haugh, North Carolina State University, Raleigh, NC, USA

HM Byrne, University of Oxford, Oxford, UK

R Linding, University of Copenhagen (UCPH), Copenhagen, Denmark

© 2016 Elsevier Inc. All rights reserved.

Understanding cell behavior has always been realized to be a complex problem requiring integration of operation of multiple molecular components involved in myriad functional and regulatory processes, intracellular and extracellular. Over the past decade or so, however, advances in experimental methodologies have facilitated direct address of this complexity at least in comprehensible portions. That is, experimental measurement of multiple molecular components (e.g., 'proteomics') and manipulation of molecular components systematically (e.g., 'functional genomics') within a defined cell biology study has become feasible, and the consequent computational analysis required for interpretation of the data has become increasingly accepted by a growing community within the cell biology field. This combination of multicomponent experimental measurement and/or manipulation with computational mining (for hypothesis generation) and/or modeling (for hypothesis testing) epitomizes the area of biological science termed Systems Biology; in relation to cell biology particularly, it can be called Systems Cell Biology. Presenting current state-of-art accomplishments, efforts, and perspectives is the aspired role of Part IV of this Encyclopedia.

Definitions of systems biology range across a wide spectrum as does work in the field itself. Two extremes on this spectrum were offered early on. One was from the Institute for Systems Biology ([Ideker et al., 2001](#)):

Systems biology does not investigate individual genes or proteins one at a time, as has been the highly successful mode of biology for the past 30 years. Rather, it investigates the behavior and relationships of all the elements in a particular biological system while it is functioning.

The other was from the National Institute of General Medical Sciences ([Anderson, 2003](#)):

Systems biology seeks to predict the quantitative behavior of an in vivo biological process under realistic perturbation, where the quantitative treatment derives its power from explicit inclusion of the process components, their interactions, and local states.

As explained in a commentary article elsewhere ([Lauffenburger and Giacomini, 2013](#)), these two definitions respectively focus on aspiration for genome-wide comprehension on the one hand and quantitative prediction on the other, with the two not immediately coincident. A central unifying concept, emphasized in two academic endeavors at MIT (see Relevant Websites section) and Harvard Medical School (see Relevant Websites section), calls for understanding how complex biological entities function by integrating quantitative information about multiple molecular- and cellular-level

components and properties via computational modeling in order to generate hypotheses, predictions, and insights.

The contributions within Part IV are organized thematically by a framework attributable originally to Peter Sorger when he and Douglas Lauffenburger, along with other colleagues, established the Computational and Systems Biology Initiative at MIT shortly after the turn of the millennium. This framework laid out 'Three Dimensions of Biological Systems Complexity,' all of which ought to be fertile ground for address by systems biology ideas and methods. It is illustrated by the diagram [Figure 1](#), reproduced from [Lauffenburger \(2012\)](#).

Efficient description of this diagram can be provided by quoting directly from the cited article:

One axis represents what I would term 'Horizontal Integration', and is the most common view of the field of systems biology: that of moving from the study of individual components to that of multiple components concomitantly. This might focus on molecular machines, or pathways/circuits/networks at a minimal level of non-reductionist complexity, all the way to genome- (or transcriptome-, proteome-, metabolome-, etc) wide extent. A second axis represents 'Vertical Integration', often associated with physiology (whether mammalian or microbial); that of moving from the study of system operation (phenotype, essentially) from the simplest contexts at the smallest space- and time-scales to more complex contexts involving larger space- and/or time-scales. In the mammalian realm, this integration may feature studies of cell-level behavior dependence on molecular properties, tissue/organ-level behavior dependence on cell properties, organism-level behavior dependence on tissue/organ properties, and population-level behavior dependence on organism properties; analogies exist in the microbial realm as well. Clearly, integrating across more than one of these spatiotemporal scale interfaces is an overarching goal, and when experimental studies of this kind are coupled with computational modeling yields the swiftly-growing field of multi-scale modeling. A third axis represents 'Dynamic Integration', characterized by the depth or intensity of information address in measurement and modeling. This axis moves from sequence and structure at the most basic degrees, through thermodynamic information, to kinetic and transport information, to comprehensive dynamical systems operation at the most extreme degree. It thus may be considered to range from bioinformatics to biochemistry to biophysics, as one vein of characterization.

Accordingly, Part IV of this Encyclopedia is organized in three sections along these three dimensions. Jason Haugh has taken responsibility for the section 'Dynamic Integration,' Helen Byrne for the section 'Vertical Integration,' and Rune Linding for the section 'Horizontal Integration.' Each of these sections comprises a good number of strong articles describing particular facets of the given integration dimension.

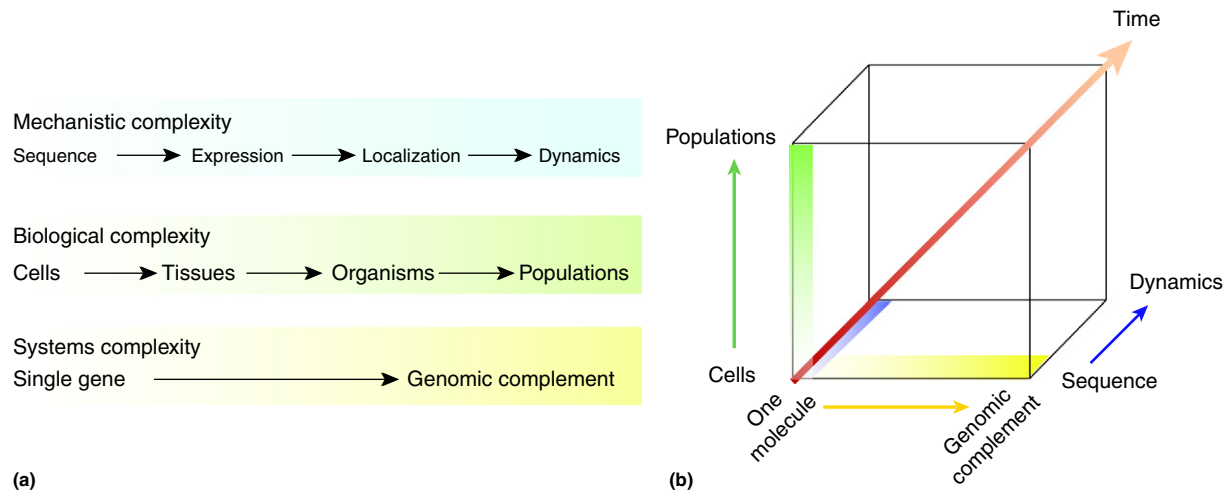


Figure 1 Framework laying out ‘Three Dimensions of Biological Systems Complexity,’ attributable originally to Peter Sorger when he established the Computational and Systems Biology Initiative at MIT shortly after the turn of the millennium. Reproduced from Lauffenburger, D.A., 2012. The multiple dimensions of integrative biology. *Integrative Biology* 4, 9.

Dynamic Integration

Whereas Horizontal Integration and Vertical Integration approaches seek to broaden the purview of systems biology in distinct ways, dynamic integration focuses on particular processes, usually involving subnetworks of interacting proteins, in mechanistic detail. Dynamic integration is undergirded by the philosophy that physicochemical principles determine and are thus manifest in the kinetics, spatial patterns, and fluctuations of intracellular processes. It follows that our understanding of those principles hinges on our ability to perturb and measure those dynamic attributes at the molecular level; ultimately, hypotheses couched as mathematical models may be tested on a quantitative basis. Receptor and membrane trafficking, signal transduction and metabolic pathways, cell adhesion and modulation of the actin and cytoskeleton, and gene regulation are among the central intracellular processes that have been studied fruitfully using a dynamic integration approach.

All of these areas are represented under the banner of Dynamic Integration in this volume, covered by experts in the analysis and modeling of these systems. Foret, Brusch, and Jülicher ([Theory of Cargo and Membrane Trafficking](#)) review a recently developed theory of cargo and membrane trafficking. Three articles are related to signal transduction: whereas Blüthgen and Legewie ([Dynamics of Protein Kinase Cascades](#)) focus on the dynamics of protein kinase cascades (such as the thoroughly studied Raf–MEK–ERK pathway), the other two discuss how signaling pathways are spatially organized during polarization of yeast (Wedlich-Söldner) ([Cdc42 and the Mechanisms of Yeast Cell Polarization – A Paradigm for Mesoscale Systems Biology](#)) and in gradient sensing by chemotactic cells (Iglesias) ([Dynamics of Gradient Sensing and Chemotaxis](#)). Sauro ([Metabolism](#)) provides an extensive primer on control theoretical analysis of metabolic pathways and networks, which is complemented by the article on gene regulatory networks by Jermusyk and Reeves ([Transcription Factor Networks](#)). On the biophysics of the cytoskeleton, there are two articles: Borinskaya,

Marchenko, and Loew ([Modeling Actin Dynamics](#)) cover the dynamics of actin filaments, whereas Castle and Odde ([Dynamics of Microtubule Self-Assembly](#)) tackle microtubules.

Vertical Integration

Many cellular functions, including proliferation, migration, and gene expression, are regulated by multiple processes acting across multiple spatial and timescales. While mathematical models of the constituent processes can provide insight into aspects of a particular cell function, the models must be combined to provide a more complete understanding of the system. Vertical integration is the merging of models formulated at different spatial and/or temporal scales to produce multiscale models. This article provides an introduction to vertical integration. This section illustrates how multiscale models may be assembled, the insight they may generate and the challenges associated with their development and validation.

The overview article (Byrne, [Vertical Integration](#)) summarizes the mathematical and computational approaches that are used to construct and simulate multiscale models ([Computational Approaches for Multiscale Modeling](#)) and the subsequent articles review the state-of-the-art in applications of multiscale modeling to the heart ([Cardiac Modeling](#)), angiogenesis ([Angiogenesis](#)), morphogenesis ([Multiscale Analysis of Morphogenesis](#)), neurogenesis ([Neurogenesis in the Adult Brain](#)), wound healing ([A Review on Various Mathematical Modeling Approaches for Wound Healing](#)) and inflammation ([Mathematical Approaches to Studying Inflammation](#)). For additional information on the mathematical and computational aspects of multiscale modeling and its biological applications, the interested reader is referred to the following review articles: Bassingthwaite *et al.* (2005), Cvijovic *et al.* (2014), Kohl and Noble (2009), Meier-Schellersheim *et al.* (2009), Schnell *et al.* (2007), Southern *et al.* (2008), and Walpole *et al.* (2013).

Horizontal Integration

In order to respond to alterations in its environment, a cell has to integrate multiple input-cues and modulate its signaling networks ([Understanding of 'Networks' *In Vitro* and/or *In Vivo*](#)) accordingly, in order to elicit a specific response such as proliferation or apoptosis. These processes become significantly altered during cancer development, with genomic modifications giving rise to differential protein dynamics, ultimately resulting in disease. The exact molecular signaling networks ([Interactomes – Scaffolds of Cellular Systems; Phosphoproteomics: Approaches, Developments, and Challenges](#)) underlying specific disease phenotypes remain elusive, as the definition thereof requires extensive analysis of not only the 'omic landscapes' (genomics, [Connecting Evolutionary Genomics to Cell Biology](#); transcriptomics, [Transcriptomics](#); proteomics, [The Advent of Mass Spectrometry-Based Proteomics in Systems Biology Research](#); and metabolomics, [Metabolomics in Cell Biology](#)) within a particular tumor, but also the phenotypic response to perturbations. Additionally, a combination of *in vitro* and *in vivo* models are required, both capturing various aspects of the true biological complexity ([High-Content Screening in Cell Biology](#)). Thus, to obtain a comprehensive understanding of pathological processes, there is a critical need for an integrative global approach, which assesses a biological system such as cancer ([Computational and Systems Cancer Biology](#)) from several molecular aspects in an unbiased fashion ([Network Modeling of Heterogeneous Datasets; Stochastic Analysis of Nongenetic Cell-to-Cell Heterogeneity](#)). Overall, this section aims to demonstrate the value of deploying several genome-scale experimental platforms ([Systematic Methods to Interrogate Genetic Perturbations and Map Phosphorylation-Dependent Signaling; Integrative Systems Biology](#)), each studying a different biological aspect, combined with in-depth computational analysis, in order to shed light on the fundamental molecular processes

which underlie a complex disease like cancer and provide possible avenues for therapeutic intervention ([Drug Targeting](#)).

References

- Anderson, J., 2003. National Institute of General Medical Sciences Centers for Systems Biology. Available at: www.nigms.nih.gov/Research/FeaturedPrograms/SysBio/ (accessed 03.06.15).
- Bassingthwaite, J.B., Chizeck, H.J., Atlast, L.E., Qian, H., 2005. Multiscale modeling of cardiac energetics. *Annals of the New York Academy of Sciences* 1047, 395–426.
- Cvijovic, M., Almquist, J., Hagmar, J., *et al.*, 2014. Bridging the gaps in systems biology. *Molecular Genetics and Genomics* 289, 727–734.
- Ideker, T., Galitski, T., Hood, L., 2001. A new approach to decoding life: Systems biology. *Annual Review of Genomics and Human Genetics* 2, 343–372.
- Kohl, P., Noble, D., 2009. Systems biology and the virtual physiological human. *Molecular Systems Biology* 5 (1), 292. doi:10.1038/msb.2009.51.
- Lauffenburger, D.A., 2012. The multiple dimensions of integrative biology. *Integrative Biology* 4, 9.
- Lauffenburger, D.A., Giacomini, K.M., 2013. Systems biology and systems pharmacology. *NAE Bridge Fall* 43, 26–33.
- Meier-Schellersheim, M., Fraser, I.D., Klausch, F., 2009. Multiscale modeling for biologists. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 1 (1), 4–14.
- Schnell, S., Grima, R., Maini, P.K., 2007. Multiscale modeling in biology. *American Scientist* 95, 135–142.
- Southern, J., Pitt-Francis, J., Whiteley, J., *et al.*, 2008. Multiscale computational modeling in biology and physiology. *Progress in Biophysics and Molecular Biology* 96, 60–89.
- Walpole, J., Papin, J.A., Peirce, S.M., 2013. Multiscale computational models of complex biological systems. *Annual Review of Biomedical Engineering* 15, 137–154.

Relevant Websites

- <http://csbi.mit.edu/overview/index.html>
Computational and Systems Biology – MIT.
- <https://sysbio.med.harvard.edu/>
Systems Biology – HMS.

DYNAMIC INTEGRATION: DYNAMICS

Contents

Dynamics of Gradient Sensing and Chemotaxis

Modeling Actin Dynamics

Metabolism

Dynamics of Microtubule Self-Assembly

Dynamics of Protein Kinase Cascades

Cdc42 and the Mechanisms of Yeast Cell Polarization – A Paradigm for Mesoscale Systems Biology

Theory of Cargo and Membrane Trafficking

Transcription Factor Networks

Dynamics of Gradient Sensing and Chemotaxis

PA Iglesias, The Johns Hopkins University, Baltimore, MD, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The ability to direct movement based on environmental cues is of paramount importance for the survival of many organisms. In many cases, this movement is in response to chemical gradients, a process known as chemotaxis. However, there are many examples of cells moving in response to other types of external cues, including electrical fields (electrotaxis) and currents (galvanotaxis), light (phototaxis), temperature (thermotaxis), mechanical stress (mechanotaxis), and the mechanical compliance of the substrate (durotaxis) (Eisenbach, 2004). Single-cell organisms, such as free swimming bacteria and ameboid cells, employ chemotactic strategies to search for nutrients or to survive. In multicellular organisms, chemotaxis plays a major role in many aspects of development and tissue maintenance. Normal physiological processes such as lymphocyte homing, angiogenesis, embryogenesis, neurogenesis, wound healing, and sperm activation require accurate migration of specific cells. On the other hand, inappropriate regulation of chemotaxis plays a role in excessive inflammation and inflammation-related diseases such as asthma, multiple sclerosis, and arthritis (Di Gennaro and Haeggström, 2012). Chemotaxis can also lead to deleterious outcomes – for example, metastatic cells employ a chemotactic strategy in their migration away from the primary tumor toward distant organs (Roussos *et al.*, 2011).

Cells performing chemotaxis do so by coordinating a number of separate but interrelated processes. Clearly, chemotactic cells must be motile. The various methods used by cells to move is highly diverse, including swimming by the beating of flagella and cilia, swarming, gliding, crawling, and ameboid movement. While the motility mechanisms are different, a relatively common feature of most chemotactic cells is that they remain motile in the absence of chemotactic stimuli. For example, *Escherichia coli* bacteria swim by continuously alternating between periods of

counterclockwise and clockwise flagellar rotation (Berg, 1993). When rotating in the counterclockwise fashion, the flagella propel the cell. Alternatively, with clockwise rotation, the flagella unbundle causing a short tumbling event – a period where movement stops and the cell's axis is randomly reoriented in space. In the absence of chemoattractant gradients, the expected period between tumbling events is relatively constant, so that movement follows a random walk. Similarly, in *Dictyostelium discoideum* amoebae – a common model organism for studying chemotaxis in eukaryotic cells – actin-rich pseudopodial extensions are projected isotropically along the cell periphery, propelling the cell (Swaney *et al.*, 2010). In the absence of chemotactic cues, the resultant motion has the statistics of a persistent random walk (Li *et al.*, 2008). The nature and timing of these pseudopodial extensions do not differ greatly between randomly moving and chemotaxing cells (Andrew and Insall, 2007). Thus, the external chemoattractant gradient is used not to promote motility, but to align movement in the direction of the chemoattractant gradient. In some cells that use flagellar systems for movement, this direction takes the form of active steering (Fenchel, 2001). In most other cases, however, the alignment results from a stochastic biasing of the migratory apparatus so that, over time, net movement is in the direction of the external gradient. For *E. coli* cells the period between tumbles is modulated so they are less frequent when swimming up a gradient (Berg, 1993). For *Dictyostelium* cells, the location of the pseudopods is regulated so as to move along the gradient (Andrew and Insall, 2007).

Strategies for Chemotaxis

To align the axis of migration with the external chemoattractant field requires a mechanism for sensing the external

chemoattractant concentration – cell surface receptors that are specific to the chemoattractant ligand – and a means for interpreting this signal so as to determine a preferred spatial direction. For the latter, cells broadly employ two different strategies based on temporal or spatial information (Devreotes and Janetopoulos, 2003).

Temporal Sensing

A cell that is moving up a chemoattractant gradient experiences increasing receptor occupancy over time. Conversely, if receptor occupancy is decreasing over time, then the cell must be moving down the chemoattractant gradient. Thus, a cell can determine whether its movement is aligned with the chemoattractant gradient by comparing the level of occupied receptors over time as it moves in a medium (Figure 1(a)). This temporal sensing strategy for determining the concentration gradient is particularly effective if cellular movement is sufficiently fast that the cell can explore a large region in a reasonable time. Organisms that direct movement using a temporal sensing strategy include various free swimming

bacteria such as *E. coli*, which moves at 10–20 body lengths per second. In this case, the temporal changes in the chemoattractant receptor occupancy are used to regulate the timing between run and tumble episodes.

Theoretically, this temporal sensing strategy can be accomplished through differentiation of the chemoattractant receptor occupancy signal. Experimentally, the presence of this differentiation can be tested by measuring the response of cells to a step change in the concentration of the chemoattractant. In the presence of temporal differentiation, except for a short transient following the application of a chemoattractant stimulus, the cellular response should be independent of the stimulus level, as the derivative of a steady signal is zero. This adaptation, or return to prestimulus levels, is a consequence of the temporal differentiation.

In practice, however, perfect differentiation is neither feasible nor desirable. The binding of chemoattractant ligands to their receptors is a stochastic process (Ueda *et al.*, 2001). As such, even in the presence of a constant level of chemoattractant concentration, the number of occupied receptors fluctuates as molecules bind and unbind. Pure

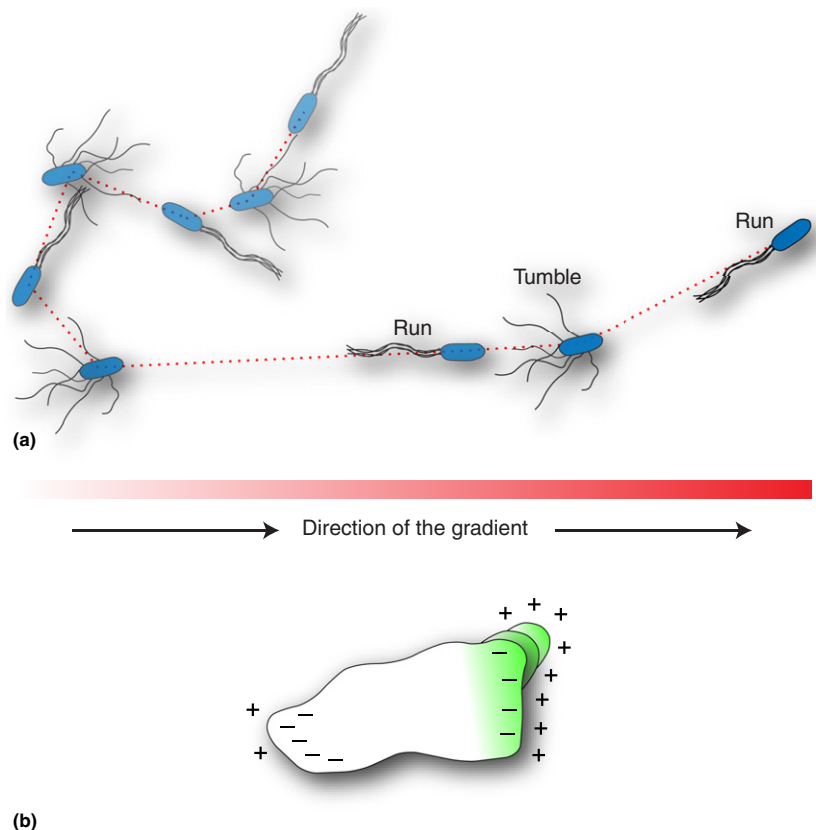


Figure 1 Strategies for determining the chemoattractant gradient. (a) In the temporal scheme, small, fast moving cells compare receptor occupancy over time. When moving down a gradient, the number of occupied receptors will decrease over time. Conversely, as cells move in the direction of increasing concentration, the number of occupied receptors increases. This strategy is used by *E. coli* cells, which move during runs, which alternating with short periods, or tumbles, where the cell is reoriented randomly. Directional migration is achieved by modulating the time between these two states, so that when receptor occupancy is increasing, the period between tumbles is lengthened, and shortened when moving down the gradient. (b) Larger, slower cells use a spatial gradient sensing mechanism. These cells have receptors distributed around the cell perimeter, and can compare the number occupied at different points in space. Directional movement is achieved by localizing the motility apparatus in the region with highest receptor occupancy.

differentiation of this noisy signal would amplify these stochastic fluctuations and this would cause the cell to respond to the noise, rather than the chemoattractant signal. In practice, the period over which the temporal comparison is made needs to be sufficiently long to filter out the effect of these fluctuations (Berg and Purcell, 1977). This can be done by increasing the time over which the temporal comparison is made. In *E. coli*, this comparison is based on the receptor occupancy over 2–4 s. This filtering time cannot be too long, however. Otherwise, the concentration of the gradient may change as the cell is moving. In bacteria, this can be because of the effects of rotational diffusion which prevents the bacteria from swimming in perfectly straight lines. Analysis of the resultant tradeoffs suggest that *E. coli* cells balance these competing requirements by integrating the receptor occupancy information over periods which are close to optimal.

Spatial Sensing

Temporal sensing may not be adequate for slow moving cells, as the level of receptor occupancy will not change appreciably over time. Moreover, the region in space that could be explored would be small. However, if these cells are sufficiently large, so that the chemoattractant concentration at multiple locations can be measured simultaneously and compared, then information about the chemoattractant field can be deduced without movement (Figure 1(b)). This spatial sensing mechanism seems to be preferred by larger, slower eukaryotic cells, including social amoebae and cells of the mammalian immune system.

In *Dictyostelium* cells, chemoattractant receptors and associated G-proteins are distributed uniformly along the cell perimeter (Swaney *et al.*, 2010). In response to chemoattractant gradients, several signaling events are selectively localized. For example, activation of Ras proteins and PI 3-kinase, accumulation of phosphatidylinositol (3,4,5) phosphate (PIP3), and actin polymerization occur at the side of highest chemoattractant receptor occupancy. In contrast, the PI 3-phosphatase, PTEN, and myosin localize at the rear of the cell. These signaling events can be localized in cells that are immobilized by inhibitors of the cytoskeleton (Parent and Devreotes, 1999). When these cells are exposed to a stable gradient, the front and rear responses form crescents toward or away from the high side of the gradient, and these responses persist as long as the gradient is present, thus demonstrating the presence of a spatial sensing mechanism.

For cells in a gradient, the level of receptor occupancy differs at various sites along the membrane. However, to direct movement, the cell must still determine which side of the cell is experiencing greater occupancy. The precise nature by which this comparison is made has been a long standing open question, with a number of models proposed (Devreotes and Janetopoulos, 2003). In theory, pseudopods could be extended at a rate proportional to the local receptor occupancy. In this case the side of the cell facing higher chemoattractant concentration would experience greater number of protrusions, and so net movement would be in this direction. However, the side with lower receptor occupancy would also

see an increase in the number of protrusions relative to an unstimulated cell. In practice, this increase is not seen. Rather, there is a suppression of pseudopod extensions at both the lateral and posterior regions of the cell. Thus, some form of inhibition, which may be due to the presence of an inhibitor, or because of depletion of molecules needed for protrusions, is believed to be active. A general model that has been proposed is the local excitation, global inhibition (LEGI) mechanism (Figure 2; Parent and Devreotes, 1999; Levchenko and Iglesias, 2002). In this scheme, local receptor occupancy triggers both a fast excitation and a slower inhibitory process. The chemotactic response is dictated by the balance between these two processes. Moreover, whereas the excitation signal remains local to where it was formed, the inhibitory signal can diffuse away and hence represents the average level of receptor occupancy. Thus, excitation exceeds inhibition at the site of highest receptor occupancy, leading to a net positive signal. Conversely, at the site of lower concentration, inhibition is higher than excitation and the chemotactic response is suppressed.

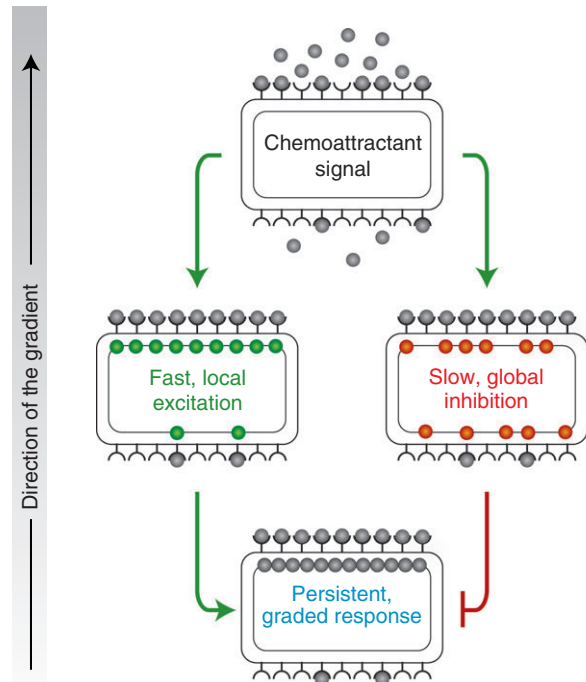


Figure 2 Local excitation, global inhibition model for spatial sensing. In the scheme, the external signal, which represents receptor occupancy, regulates a fast, local excitation, and a slow, global inhibitor. Together, these processes regulate in complementary fashion a response regulator that signals to motility. When introduced to a gradient, the side facing the highest concentration has greater number of occupied receptors (gray circles), and hence generates more excitation molecules (green). More slowly, but still following the receptor occupancy gradient, inhibition molecules (red) are formed. These molecules diffuse away from the site of their production, and equilibrate spatially, reflecting the global level of receptor occupancy. Thus, at the side facing the highest chemoattractant concentration, excitation exceeds inhibition, leading to a positive response. At the rear, inhibition is higher, leading to a suppression of the response.

Adaptation in Chemotactic Systems

Though adaptation appears to be a necessary feature of the temporal sensing strategy, it is also found in a number of cells that can sense gradients spatially, such as *Dictyostelium* amoebae and neutrophils. When these cells are stimulated with spatially uniform step changes in chemoattractant concentration, the responses that are usually seen at the front, such as PIP3 accumulation, transiently increase while those that localize to the rear, such as PTEN membrane association, decrease before returning to baseline. Nevertheless, subsequent responses are triggered by increments in receptor occupancy and adapt when occupancy is held constant. Further response only occurs if occupancy is increased again or if the stimulus is removed and reapplied after a period of deadaptation.

Not all chemotactic cells adapt to isotropic chemotactic stimuli, however. For example, the response of migrating fibroblasts to uniform PDGF stimulation does not adapt, though these cells can only respond to gradients over a relatively narrow range of chemoattractant concentrations (Schneider and Haugh, 2005). Thus, part of the advantage of adaptation is that, by filtering out the mean level of chemoattractant concentration, it helps to focus the signal on the gradient. In this case, the resultant movement depends on the relative and not absolute gradient (Zigmond, 1977).

Over the years, a number of mechanisms have been proposed to explain the adaptive behavior of cells. It can be shown, however, that the schemes can be broadly grouped into two classes (Ma et al., 2009). Here we briefly outline minimal models for each of these two schemes.

Adaptation through Negative Feedback Regulation

Negative feedback control is a ubiquitous means for regulating the level of a response. It also forms the basis for several mechanisms proposed to explain the observed adaptation of cells to persistent changes in external chemoattractant stimuli. The basic scheme is shown in Figure 3(a). Here, an external signal S (e.g., chemoattractant receptor occupancy) triggers a chemotactic response: $S \rightarrow R$. Thus, as the concentration of the external signal increases, so does that of the response. Additionally, the scheme includes a regulatory buffer (F) that

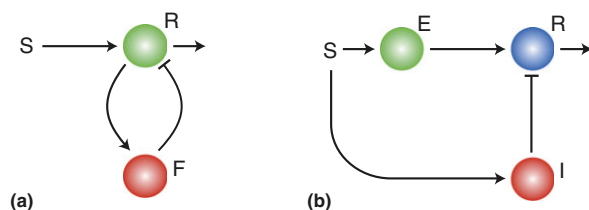


Figure 3 Schemes for achieving adaptation. (a) Negative feedback system. In this scheme, the external chemical (S) signals to a response (R). The latter signals to a buffer (F) that closes a negative feedback loop. (b) In the incoherent feedforward scheme, S signals to both excitation (E) and inhibition (I) processes, which regulate in complementary fashion the regulator. Note that the local excitation, global inhibition pathway of Figure 2 is a special case of an incoherent feedforward system.

provides negative feedback from the response back to itself: $R \rightarrow F \rightarrow R$. In this case, as the concentration of the response increases, so does that of the feedback buffer. Eventually, this negative feedback drives the response toward prestimulus levels.

The scheme proposed here is equivalent to an influential model proposed to explain the experimentally observed perfect adaptation in *E. coli* chemotaxis (Barkai and Leibler, 1997). In these cells, the chemoattractant receptor occupancy signal is transmitted to the flagellar motor by protein CheY, which are in turn regulated by CheA, a histidine kinase which forms a complex with the receptors (Wadhams and Armitage, 2004). These complexes are assumed to exist in one of two activity states. Binding of chemoattractants to the receptor alters the state of the receptor leading to changes in the phosphorylation state of CheA, which in turn alters the rotational state of the motor. However, receptor-CheA complexes can undergo reversible methylation at a number of glutamate residues by two proteins: a methyltransferase, CheR, and methylesterase, CheB. The latter is also regulated by the phosphorylation state of CheA, closing a negative feedback loop as changes in receptor methylation level shift restore a balance between counterclockwise and clockwise signal outputs.

In practice, a number of conditions must be met to ensure perfect adaptation – that is, that the response return exactly to the basal level (Yi et al., 2000). In this case, it can be shown that the adaptation mechanism is equivalent to integral feedback control, a common engineering motif used to minimize the effect of disturbances in man-made systems.

Adaptation through an Incoherent Feedforward Scheme

A second class of adapting motifs is described in Figure 3(b). This scheme was originally proposed by Koshland (1977) as a means of explaining adaptation in bacterial chemotaxis though it is now more widely used to explain adaptation in models of chemotactic amoebae (Levchenko and Iglesias, 2002; Tang et al., 2014). This is a special case of the LEGI mechanism described above. As above, the central element of network is a response element, R , that is activated by an excitatory process E , and suppressed by an inhibitory process, I . Both of these regulatory processes are themselves regulated by the stimulus, through increases in receptor occupancy, S . Because the excitation and inhibition signals provide complementary regulation on the response regulator, the motif is sometimes known as an incoherent feedforward loop. The key to achieving sensory detection is that the time scale of the two excitation and inhibition processes be different. If the excitation is faster, then the response rises transiently as excitation increases. As the inhibition catches up, the response eventually subsides to the prestimulus levels.

Signal Amplification and Sensitivity

One of the remarkable features of chemotactic cells is their ability to discern small differences in chemoattractant concentrations and to use this to chemotax effectively. In amoeboid

cells, the difference in chemotactic receptor occupancy between opposite sides of the cell can be as small as 1–2% (Zigmond, 1977). Moreover, the degree to which the external signal is amplified by internal signals, such as PIP3 localization, can be in the order of 7–10 fold, so that a 10% difference in chemoattractant concentration can lead to 70% difference in PIP3 levels (Janetopoulos *et al.*, 2004). In *E. coli*, changes in the order of 0.2% can alter the probability of tumbling by 20%, an amplification of 100 (Sourjik, 2004). Experiments using fluorescent resonance energy transfer demonstrated that the receptor signaling can achieve 35-fold amplification. Moreover, these changes can be detected over a range of background concentrations that ranges over five orders of magnitude. How this signal amplification is achieved is an active area of interest.

Amplification through Receptor Coupling

A number of researchers have suggested that the extreme sensitivity of the bacterial chemotactic signaling system may arise from interactions between two-dimensional receptor clusters. *Escherichia coli* chemoattractant receptors are clustered at both poles of the cells and exist as homodimers, which are then arranged as trimers of dimers (Sourjik and Wingreen, 2012). Models that assume that these receptors act independently do not recreate several observed characteristics, including the signal amplification, as well as integration of various chemotactic signaling cues. However, a number of models have been proposed that suggest that receptors do not work independently, but rather as cooperative clusters, though the specific interactions are not known. For example, receptors could follow the classic Monod-Wyman-Changeux model of multi-subunit allosteric proteins, or they could interact in larger lattices of interacting receptors (Sourjik, 2004). In either case, binding of a ligand to a receptor leads to a conformational spread across the cluster. If the coupling is sufficiently strong, then all receptors in a cluster are found in the same state. These models are able to explain the observed sensitivity.

Amplification through Positive Feedback

A common way to amplify signals is through the use of a positive feedback signal, or its functional equivalent, the double negative feedback loop. The existence of various positive feedback pathways regulating chemotaxis has been demonstrated. For example, in neutrophils, a positive feedback involving PIP3 and actin polymerization has been demonstrated (Van Keymeulen *et al.*, 2006). In *Dictyostelium* cells, a similar pathway exists involving Ras (Sasaki *et al.*, 2007). Measurements of the amplification in *Dictyostelium* cells that have been treated by latrunculin show that the signal gradient amplification is smaller, suggesting that actin does play an active role, but it is still present, indicating that other actin-independent steps must exist (Janetopoulos *et al.*, 2004). Positive feedback loops are also present in the chemotaxis signaling pathway of *Bacillus subtilis*, where it is enacted by suppression of the CheV protein (Rao *et al.*, 2004).

The presence of a positive feedback loop can present some difficulties, as can be attested by anyone who has placed a

microphone near a speaker. Once started, how do you turn off the signal? In fact, positive feedback may also lead to hysteresis, whereby the response of a system depends not only on the present level of the input, but also on the past (Ferrell, 2002). Evidence of this memory exists in the chemotactic responses of both bacteria and eukaryotes. In the latter, it is best observed in cells subject to a change in the direction of a gradient. These cells, which are usually referred to as being polarized, turn gradually in the direction of the new gradient (Devreotes and Janetopoulos, 2003; Skoge *et al.*, 2014; Nakajima *et al.*, 2014). This behavior can be seen as the result of a tradeoff between the immediate gradient sensing mechanism and an internal compass that remembers the past direction (Andrews and Iglesias, 2007).

Even if hysteresis is desirable, the effect of the positive feedback must be limited. For this reason, positive feedback loops are usually counterbalanced by negative feedback loops (Brandman and Meyer, 2008). Coupling these two types of feedbacks may lead to oscillatory behavior.

Amplification through an Excitable Network

Recent years have revealed the presence of excitable behavior in the chemotactic pathways of several eukaryotic cells (Iglesias and Devreotes, 2012). Excitable systems operate at a stable point in which small amplitude perturbations lead to negligible responses. However, if the perturbations are sufficiently large, they can cross a threshold in which a strong positive feedback loop leads to a large scale, transient response, or firing. This is followed by the triggering of a negative feedback loop that eventually causes the response to subside and the system to return to its equilibrium. Evidence for the presence of excitable behavior includes the presence of propagating waves (Vicker, 2002; Weiner *et al.*, 2007) and a refractory period that follows the triggering of the excitable behavior, during which no further firings can take place (Huang *et al.*, 2013; Nishikawa *et al.*, 2014).

The presence of excitable behavior helps to explain how cell motility can be biased during chemotaxis. In unstimulated cells, stochastic fluctuations that arise during molecular events can lead to supra-threshold triggering of excitable behavior, which will lead to actin polymerization and pseudopod protrusions. For cells in a gradient, gradient sensing mechanism can regulate the level of the threshold, so that, compared to unstimulated cells, firings are more likely at the front and less likely at the rear. Thus, threshold regulation can help to bias the likelihood of generating pseudopods between front and back. However, once a pseudopod is triggered, including life time and size, proceeds independently of the external chemoattractant signal.

Conclusions

Chemotaxis is a fascinating biological process of central importance to the survival of organisms. It brings together fundamental properties of cells including the ability to move, to communicate with its environment, and to regulate cell shape. It also is a fascinating example of how biology brings together

design principles that are common in man-made systems, such as positive and negative feedback loops, hysteresis, oscillator, amplification, adaptation, robustness, and redundancy to solve an essential need in diverse and elegant fashion.

See also: Cell Communication: Cellular Outcomes: Regulation of Cell Migration. Cell Communication: Growth Factor Mediated Cell Signaling: G Protein-Coupled Receptors. Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell Migration; Cell Polarity; Cilia and Flagella. Cytoskeleton and Motors: Cytoskeletal Components: Actin Assembly Dynamics and Its Regulation in Motile and Morphogenetic Processes. Dynamic Integration: Dynamics: Modeling Actin Dynamics. Molecular Principles, Components, Technology, and Concepts: Lipids: Lipid Signaling

References

- Andrew, N., Insall, R.H., 2007. Chemotaxis in shallow gradients is mediated independently of PtdIns 3 kinase by biased choices between random protrusions. *Nature Cell Biology* 9, 193–200.
- Andrews, B.W., Iglesias, P.A., 2007. An information-theoretic characterization of the optimal gradient sensing response of cells. *PLoS Computational Biology* 3, e153.
- Barkai, N., Leibler, S., 1997. Robustness in simple biochemical networks. *Nature* 387, 913–917.
- Berg, H.C., 1993. *Random Walks in Biology*. Princeton, NJ: Princeton University Press.
- Berg, H.C., Purcell, E.M., 1977. Physics of chemoreception. *Biophysical Journal* 20, 193–219.
- Brandman, O., Meyer, T., 2008. Feedback loops shape cellular signals in space and time. *Science* 322, 390–395.
- Devreotes, P.N., Janetopoulos, C.J., 2003. Eukaryotic chemotaxis: Distinctions between directional sensing and polarization. *Journal of Biological Chemistry* 278, 20445–20448.
- Di Gennaro, A., Haeggström, J.Z., 2012. The leukotrienes: Immune-modulating lipid mediators of disease. *Advances in Immunology* 116, 51–92.
- Eisenbach, M., 2004. *Chemotaxis*. London: Imperial College Press.
- Fenchel, T., 2001. How dinoflagellates swim. *Protist* 152, 329–338.
- Ferrell, J.E., 2002. Self-perpetuating states in signal transduction: Positive feedback, double-negative feedback and bistability. *Current Opinion in Cell Biology* 14, 140–148.
- Huang, C.-H., Tang, M., Shi, C., Iglesias, P.A., Devreotes, P.N., 2013. An excitable signal integrator couples to an idling cytoskeletal oscillator to drive cell migration. *Nature Cell Biology* 15, 1307–1316.
- Iglesias, P.A., Devreotes, P.N., 2012. Biased excitable networks: How cells direct motion in response to gradients. *Current Opinion in Cell Biology* 24, 245–253.
- Janetopoulos, C., Ma, L., Devreotes, P.N., Iglesias, P.A., 2004. Chemoattractant-induced phosphatidylinositol 3,4,5-trisphosphate accumulation is spatially amplified and adapts, independent of the actin cytoskeleton. *Proceedings of the National Academy of Sciences of the United States of America* 101, 8951–8953.
- Koshland Jr., D.E., 1977. A response regulator model in a simple sensory system. *Science* 196, 1055–1063.
- Levchenko, A., Iglesias, P.A., 2002. Models of eukaryotic gradient sensing: application to chemotaxis of amoebae and neutrophils. *Biophysical Journal* 82, 50–63.
- Li, L., Nørrelykke, S.F., Cox, E.C., 2008. Persistent cell motion in the absence of external signals: a search strategy for eukaryotic cells. *PLoS One* 3, e2093.
- Ma, W., Trusina, A., El-Samad, H., Lim, W.A., Tang, C., 2009. Defining network topologies that can achieve biochemical adaptation. *Cell* 138, 760–773.
- Nakajima, A., Ishihara, S., Imoto, D., Sawai, S., 2014. Rectified directional sensing in long-range cell migration. *Nature Communications* 5, 5367. doi:10.1038/ncomms6367.
- Nishikawa, M., Hörning, M., Ueda, M., Shibata, T., 2014. Excitable signal transduction induces both spontaneous and directional cell asymmetries in the phosphatidylinositol lipid signaling system for eukaryotic chemotaxis. *Biophysical Journal* 106, 723–734.
- Parent, C.A., Devreotes, P.N., 1999. A cell's sense of direction. *Science* 284, 765–770.
- Rao, C.V., Kirby, J.R., Arkin, A.P., 2004. Design and diversity in bacterial chemotaxis: A comparative study in *Escherichia coli* and *Bacillus subtilis*. *PLoS Biology* 2, e49.
- Roussos, E.T., Condeelis, J.S., Patsialou, A., 2011. Chemotaxis in cancer. *Nature Reviews Cancer* 11, 573–587.
- Sasaki, A.T., Janetopoulos, C., Lee, S., et al., 2007. G protein-independent Ras/P13K/F-actin circuit regulates basic cell motility. *Journal of Cell Biology* 178, 185–191.
- Schneider, I.C., Haugh, J.M., 2005. Quantitative elucidation of a distinct spatial gradient-sensing mechanism in fibroblasts. *Journal of Cell Biology* 171, 883–892.
- Skoge, M., Yue, H., Erickstad, M., et al., 2014. Cellular memory in eukaryotic chemotaxis. *Proceedings of the National Academy of Sciences of the United States of America* 111, 14448–14453.
- Sourjik, V., 2004. Receptor clustering and signal processing in *E. coli* chemotaxis. *Trends in Microbiology* 12, 569–576.
- Sourjik, V., Wingreen, N.S., 2012. Responding to chemical gradients: Bacterial chemotaxis. *Current Opinion in Cell Biology* 24, 262–268.
- Swaney, K.F., Huang, C.-H., Devreotes, P.N., 2010. Eukaryotic chemotaxis: A network of signaling pathways controls motility, directional sensing, and polarity. *Annual Review of Biophysics* 39, 265–289.
- Tang, M., Wang, M., Shi, C., et al., 2014. Evolutionarily conserved coupling of adaptive and excitable networks mediates eukaryotic chemotaxis. *Nature Communications* 5, 5175. doi:10.1038/ncomms6175.
- Ueda, M., Sako, Y., Tanaka, T., Devreotes, P., Yanagida, T., 2001. Single-molecule analysis of chemotactic signaling in *Dictyostelium* cells. *Science* 294, 864–867.
- Van Keymeulen, A., Wong, K., Knight, Z.A., et al., 2006. To stabilize neutrophil polarity, PIP3 and Cdc42 augment RhoA activity at the back as well as signals at the front. *Journal of Cell Biology* 174, 437–445.
- Vicker, M.G., 2002. Eukaryotic cell locomotion depends on the propagation of self-organized reaction–diffusion waves and oscillations of actin filament assembly. *Experimental Cell Research* 275, 54–66.
- Wadhams, G.H., Armitage, J.P., 2004. Making sense of it all: Bacterial chemotaxis. *Nature Reviews Molecular Cell Biology* 5, 1024–1037.
- Weiner, O.D., Marganski, W.A., Wu, L.F., Altschuler, S.J., Kirschner, M.W., 2007. An actin-based wave generator organizes cell motility. *PLoS Biology* 5, e221.
- Yi, T.M., Huang, Y., Simon, M.I., Doyle, J.C., 2000. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the National Academy of Sciences of the United States of America* 97, 4649–4653.
- Zigmond, S.H., 1977. Ability of polymorphonuclear leukocytes to orient in gradients of chemotactic factors. *Journal of Cell Biology* 75, 606–616.

Further Reading

- Berg, H.C., 1993. *Random Walks in Biology*. Princeton, NJ: Princeton University Press.
- Eisenbach, M., 2004. *Chemotaxis*. London: Imperial College Press.
- Iglesias, P.A., Devreotes, P.N., 2012. Biased excitable networks: How cells direct motion in response to gradients. *Current Opinion in Cell Biology* 24, 245–253.
- Jin, T., 2013. Gradient sensing during chemotaxis. *Current Opinion in Cell Biology* 25, 532–537.
- Sourjik, V., Wingreen, N.S., 2012. Responding to chemical gradients: Bacterial chemotaxis. *Current Opinion in Cell Biology* 24, 262–268.
- Swaney, K.F., Huang, C.-H., Devreotes, P.N., 2010. Eukaryotic chemotaxis: A network of signaling pathways controls motility, directional sensing, and polarity. *Annual Review of Biophysics* 39, 265–289.

Modeling Actin Dynamics

S Borinskaya, O Marchenko, and LM Loew, University of Connecticut Health Center, Farmington, CT, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Microfilaments, microtubules, and intermediate filaments are the polymers that comprise the cytoskeleton of cells. But the microfilament system, for which actin is the basic building block, is arguably the most complex because of the variety of regulatory mechanisms that control its function and dynamics. The regulation of actin polymerization, bundling, cross-linking, branching, and its association with plasma and organelle membranes, is controlled by an enormous number of signaling molecules and actin-binding proteins. Because of this diverse biochemistry, actin can be utilized by the cell to control diverse physiology, including motility, endocytosis, cytokinesis, morphogenesis, and organelle trafficking.

Given the importance of actin, it has been the subject of intense experimental investigation, as reviewed in the article in this work by Carrier *et al.* These studies include many quantitative measurements both *in vitro* and *in vivo* on the interactions of actin with its regulators. Because of the complexity of these biochemical interactions, it is difficult to understand how the various components work together to produce a cellular behavior. But the wealth of quantitative measurement on these components also offers the opportunity to solve the complexity problem through computational modeling. Modeling addresses the complexity of actin dynamics in two related ways: (1) The construction of a model helps the investigator organize all the diverse experimental data that are used as input to the model and also identify key pieces of information that may be lacking; (2) Simulations can predict the results of new experiments, especially cellular measurements, and thereby test the hypotheses that are inherent to the model. Of course, these advantages of modeling are true for any complex biological system (Slepchenko *et al.*, 2003). But actin dynamics has the additional challenge of requiring the integration of cell signaling, biochemistry, polymer physics, and mechanics.

The wealth of biochemical and signaling data calls for detailed models that aim to address the role of each molecular player in the system. On the other hand, the polymer physics and mechanics have inspired more conceptual models that attempt to explain the global features of the cell biology. The philosophical difference between detailed and conceptual models, as related to actin modeling, has been recently discussed (Ditlev *et al.*, 2013; Mogilner *et al.*, 2012). In this article, we will provide the reader with a guide to examples of each approach to modeling actin dynamics with a view toward illustrating how each approach can produce valuable and complementary insights into this complex but central subject within the field of cell biology.

Conceptual Models of Mechano-Chemical Control of Cell Motility and Shape

Conceptual models of actin dynamics are typically minimal models focused on various biomechanical aspects of cell

motility. These models are based on fundamental laws of thermodynamics and mechanics, specifically, on force balance within the cell cytoskeleton and mass conservation of actin. In the multifaceted space of diverse interacting physical/chemical mechanisms, intuition falls short in the interpretation of the experimental data or prediction of new mechanisms. To generate conceptual models, these mechanisms must be reduced to mathematical equations, containing variables related to experimental observables or parameters that have experimentally controllable correlates; solving these equations through simulations can then generate hypotheses and motivate new experiments on how actin dynamics can be marshaled to propel cells, intracellular organelles, or pathogens. The abstractions and approximations depend on the specific question addressed by each particular model. In this section, we will describe a few examples of how conceptual models have been applied to actin-driven motility and the control of cell shape. However, to stress the mechanistic concepts underlying these models, we deliberately have avoided exposition of the mathematical equations that had to be solved (generally with numerical simulations). We refer the reader to the original papers for these details.

'Brownian ratchet' is a machine that uses the random Brownian motion of molecules to produce work. It may be applied in a model for how actin dynamics can produce membrane protrusion or propel a microbe: thermal fluctuations of the membrane at the tip of a growing filament or bending motions of the filament can make room for insertion of a new actin monomer; subsequent thermal fluctuations allow the process to continue. Since depolymerization is slow, a net protrusive force can be generated. The second law of thermodynamics is upheld because the process extracts energy from hydrolysis of the adenosine triphosphate (ATP) bound to actin.

The possibility that actin filament polymerization at the cell membrane is capable of generating protrusion force was first explained in the 'Brownian ratchet' formalism derived by Peskin *et al.* (1993) for the dependence of polymerization velocity on the effective load size (Figure 1(a)). A study by Mogilner and Oster in 1996 extended this hypothesis by considering the bending fluctuations of the filament; they showed that the velocity is independent of the size of the cargo for situations where the polymerization rate is limiting (i.e., at low monomer concentrations) (Mogilner and Oster, 1996). This demonstrated a fundamental concept that polymerization rate remains largely the same in various actin-driven motility scenarios and also led to models of microorganism motility, such as *Listeria* and *Shigella*.

Discrete versus continuum modeling. The behavior of a cellular process can be fundamentally described in terms of the collection of

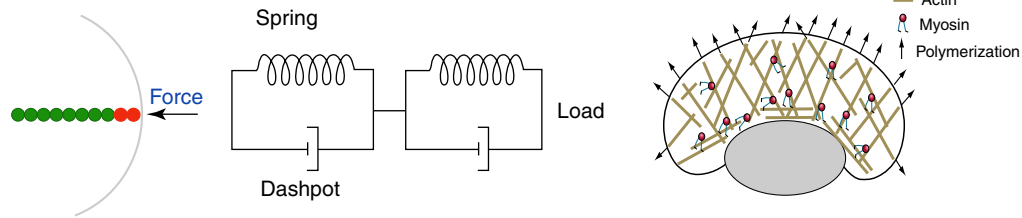


Figure 1 Model types and their evolution. (a) The concept of force generation at the cell membrane by self-renewing polymer was first modeled and tested using the Brownian ratchet formulation in studies by Peskin *et al.* (1993) and Mogilner (Mogilner and Oster, 1996). (b) Further experimental evidence lead to a rigorous utilization of actin network as force generating unit with viscous and/or elastic properties (c). Finally, continuum models were developed to look into full-scale models of how the actin network can generate cell locomotion (Wolgemuth *et al.*, 2011).

discrete probabilistic events occurring at the molecular level; such models have the advantage of correctly describing the stochastic variability of systems with small numbers of molecules. However, discrete models may be complex, computationally intractable, or fail to provide understanding of the emergent behavior of the system. Continuum modeling describes the system in terms of changes in concentrations or densities of molecular species and solves ordinary or partial differential equations (ODEs or PDEs) to predict the average system behavior. ODEs describe just the rate of change of all 'species.' PDEs add the motion of the species within a spatial geometry. Because continuum models are deterministic and cannot account for stochastic behavior, they are fundamentally approximate, but they can provide insights on the collective emergent behaviors of complex systems.

The treatment of the actin filament as an elastic spring launched a new generation of models that use familiar mechanical components to describe actin network flow and deformation: springs, dashpots, and motors (Figure 1(b)). Myosin motor activity and its effects on actin polymerization are considered in such models. The differential treatment of G- and F-actin anticipated the two-phase treatment of the cytosol which led to development of 'continuum models' that incorporate contractile stress, polymerized actin, and cytosol flow with traction forces on the whole cell scale (Danuser *et al.*, 2013; Figure 1(c)).

Species. A biological molecule (ligand, second messenger, drug, protein, DNA, RNA, etc.) or a particular state of a molecule (e.g., monomer or dimer, phosphorylated or unphosphorylated, adenosine diphosphate (ADP)-bound actin or ATP-bound actin). It can also be used in relationship to any other player (e.g., fluorescently labeled protein) in the experimental system to be modeled. Each species corresponds to a variable in the model.

Mogilner and Oster's Brownian ratchet model (Figure 2) demonstrates that polymerizing actin filament can generate protrusive pressure and push the membrane, which results in cell motility. Relaxing the assumption that the filament is rigid and also considering cross-linking as well as parallel filaments, they construct a statistical mechanical model of *Listeria monocytogenes* movement and lamellipodial protrusion. For the sake

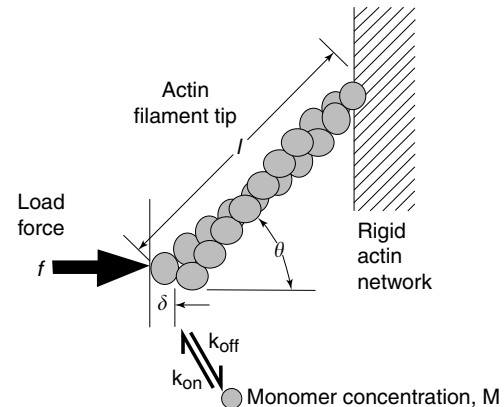


Figure 2 Elastic actin filament in Brownian ratchet model (Mogilner and Oster, 1996).

of simplicity the study ignores cytoplasmic fluid flow. Actin filaments are modeled as elastic rods with persistence length $\lambda = 1 \mu\text{m}$ that grow by addition of monomers at rate $k_{\text{on}}M$ and shorten at rate k_{off} .

The model predicts that with an increase in the resistance force, the filaments of the lamellipodial cytoskeleton will align at angles more parallel to the leading edge of the cell. Higher concentrations of actin-binding proteins would produce heavier cross-linking, shorter free ends of the filaments, and effectively slower protrusion velocities. The model concludes that lamellipodial protrusion is driven by polymerization of the thermally fluctuating actin filaments, and the orthogonal geometry of the filaments is determined by actin cross-linking proteins (Mogilner and Oster, 1996). Finally, the model derives an expression for velocity at the membrane, which is then used as a boundary condition to successfully model retrograde flow in fish keratocytes, a cell that glides along a flat substrate without significant changes in shape (Mogilner and Oster, 1996).

The ability to model movement of a cell with a complex shape, such as the fibroblast or neutrophil, is more challenging. Furthermore, the interaction of the cytoskeleton with the rapidly deforming cell membrane and stresses that it generates are complex due to the two-phase nature of the intracellular environment itself. To tackle these challenges Dembo and colleagues developed a continuum model of the contractile and viscoelastic cytoplasm that considers a two-phase treatment of the cytosol and cytoskeleton components

and their interaction with the cell membrane (Herant *et al.*, 2003). To understand which component, fluid or solid, plays a major role in cell shape and locomotion, they develop two hypotheses for the active force generation on the cell membrane: 'cytoskeletal swelling force model' and 'polymerization force model.'

The model equations describe the relations between cytoplasmic flows, substrate adhesion and protrusion dynamics. In the 'network swelling model,' a predominantly fluid system, the cytoskeleton is defined as a disorganized structure with isotropic properties that is continually recycled. In the 'polymerization force model,' the cytoskeleton is viewed as a more permanent, organized scaffolding that allows for the directional production of force. The main idea is that the free energy released by the addition of monomers to a filament is transferred to generate pressure against a membrane. To test the models against the experimental data available, they simulate three classic neutrophil experiments: a neutrophil is passively aspirated into a micropipette, neutrophil exposure to a stimulus triggers the pseudopod extension, and a neutrophil migrates inside a micropipette toward a chemoattractant (Herant *et al.*, 2003). The results include derivation of the three characteristic parameters of the cell cortical properties: baseline tension, slack, and dilation viscosity (Herant *et al.*, 2003). These parameters are constrained by data from aspiration experiments and yield appropriate results for the pseudopod experiments.

Both swelling and polymerization force models create a unifying framework for the experimental data for the three experiments. However, force production required in the polymerization model is beyond what is expected from a simple short-range Brownian ratchet model. In the crawling of neutrophils or other ameboid cells inside a micropipette, Herant *et al.* (2003) argued that measurement of velocity versus counter pressure curves could provide a determination of whether cytoskeleton–cytoskeleton interactions (such as swelling) or cytoskeleton–membrane interactions (such as polymerization force) are mainly responsible for active protrusion. They subsequently analyzed such data and showed that the experiments favor the network swelling model (Herant *et al.*, 2006). The model concludes that local area-increasing shape changes to regulate the membrane cortical tension and overcome a 'dilation viscosity' to produce a pseudopod.

A subsequent study further investigates how membrane tension and polymerization determine shape in 3D cells migrating along a 2D planar substrate (Herant and Dembo, 2010). They showed that changing a few parameters could alter the shapes of migrating cells from the hand mirror shape characteristic of fibroblasts to the half-moon shape characteristic of keratocytes. For fibroblast, the authors approximated adhesion in the boundary conditions as forming immediate attachment at the front edge and detaching at the rear when a threshold force is attained. During migration, fibroblasts displayed drip-faucet instabilities upon tail elongation. Keratocytes were modeled within the same overall formalism, but using weak adhesion and rapid recycling of actin from the rear to the front. One missing piece in the Herant and Dembo models is myosin contractility. Importantly, experimental evidence about the role of myosin contractility in determining

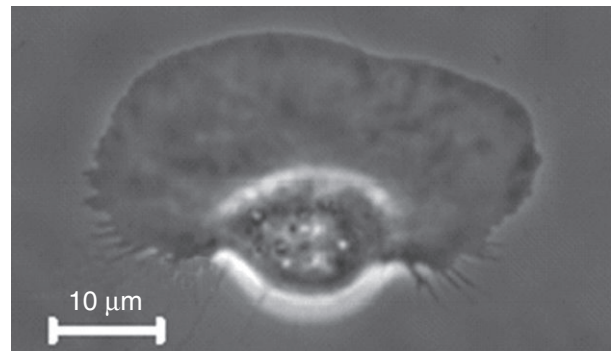


Figure 3 Phase microscopy image of the fish keratocyte cell (Keren *et al.*, 2008).

shape and motility is conflicting and cannot be addressed within the Herant–Dembo framework.

The fish keratocyte has been a popular choice to study the importance of adhesion and myosin contractility in motile cells because it is easily harvested from goldfish scales and the key parameters controlling motility and shape are readily experimentally manipulated. When placed on a flat surface, keratocytes assume a stereotypical half-moon shape with a broad, flat, motile appendage (Figure 3), the lamellipodium. Furthermore, their gliding motility simplifies some of the modeling challenges (Allard and Mogilner, 2012; Barnhart *et al.*, 2011; Rubinstein *et al.*, 2009; Wolgemuth *et al.*, 2011). Their unique ability to maintain nearly constant cell shape, speed, and direction over many cell lengths makes them a perfect system for studying actin-based cell motility.

We focus on a study by Wolgemuth *et al.* (2011) that uses moving-boundary simulations to model cell-shape dynamics and understand how four seemingly redundant mechanisms that may underlie cell motility can determine cell shape and locomotion. Previous models of cell motility considered are schematic, focusing on one mechanism, or complex, with many pathways lumped together. Wolgemuth's study proposes that the redundancy in biochemical and biophysical mechanisms is necessary for robust cellular migration.

The study analyses four distinct minimal mechanisms that regulate cell shape and locomotion: (1) a G-actin transport-limited motility (Figure 4(a)); (2) a motility model limited by microtubule-based transport of vesicles to the leading edge (Figure 4(b)); (3) Rac/Rho-regulated motility (Figure 4(c)), and (4) myosin contraction-driven motility (Figure 4(d)). Using keratocyte cell shape as a template, the study examines if the models, alone or in combination, can produce the half-moon keratocyte shape and the main features of its locomotion at steady state. In all models F-actin flow inside the keratocyte is directed to the cell rear and inward at the cell edges. Once it reaches the cell body at the rear, F-actin is depolymerized. One of the key requirements for all models is that the cell area remains constant. The cell shape is allowed to evolve from an initial circular disc using a moving-boundary algorithm for numerical simulations (Wolgemuth and Zajac, 2010). The authors then assess whether and with what parameters each model evolves the system to the characteristic half-moon gliding phenotype.

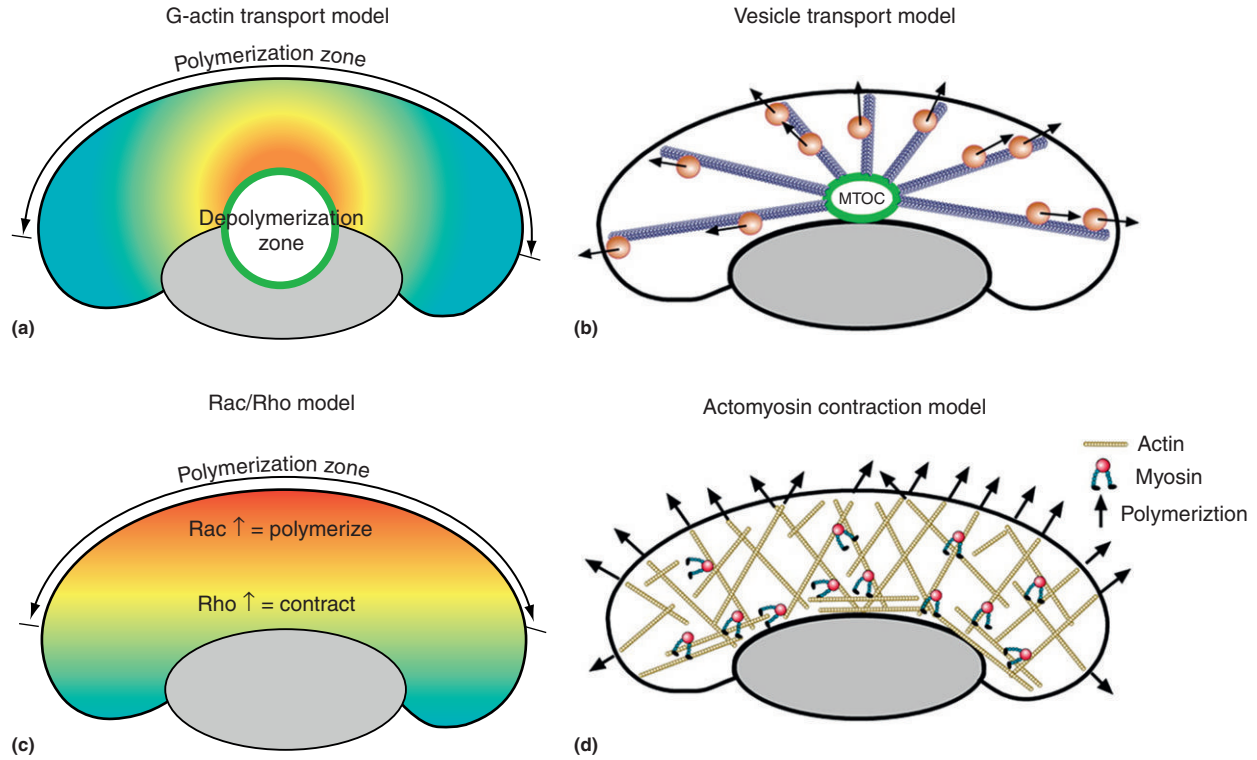


Figure 4 Schematic overview of the four locomotion mechanisms tested (Wolgemuth *et al.*, 2011). (a) G-actin transport model's main assumption is that diffusion of G-actin from the cell rear, the F-actin disassembly hotspot, is the limiting step in cell motility. The green ellipse denotes the area where no polymerization takes place, while G-actin assembles at the highest rate at the sides and periphery of the cell. (b) Vehicle transport model tests the assumption that the flux of membrane vesicles to the cell periphery is the rate-limiting step for membrane protrusion and depends on the density of the microtubule ends located at the membrane. (c) Experimental evidence on cell motility regulators Rac and Rho is tested in Rho/Rac model. Rac activation increases actin polymerization area and triggers formation of lamellipodia and membrane ruffles. The model implements activation of actin polymerization where Rac is activated. On the other hand, Rho activation favors F-actin disassembly by activating contractility. (d) The actomyosin contraction model assumes contractile stress is isotropic and proportional to bound myosin, which can bind and unbind. The cytoskeleton is treated as a viscous fluid. Actin polymerization rate is inversely related to the concentration of bound myosin.

G-actin transport is essential for maintenance of the F-actin assembly at the cell sides and leading edge. Thus, the model of Figure 4(a) assumes an initial condition with the highest rate of actin polymerization taking place at one end of the cell and a depolymerization zone at the other end. The hypothesis is that diffusion of G-actin from the cell rear, where it is generated by depolymerization, to the front, where it is needed for polymerization, is the limiting step in cell motility. The model reproduces the keratocyte half-disc shape with rates for actin disassembly/assembly being the two governing parameters. Interestingly, the velocity is nearly independent of the actin polymerization rate. Since cell-front advancement is proportional to the polymerization rate and G-actin is depleted proportionally to the F-actin polymerization, the two processes compete with each other. Thus, cell speed is controlled by the actin disassembly rate, which was reported in live cells (Cramer, 1999). The microtubule transport model, which posits that membrane cycling through microtubule driven vesicle transport is the critical mechanism (Figure 4(b)), also produces a robust shape, which is more 'fan-like' but still roughly fits within the range of morphologies sampled by live keratocytes; area conservation (i.e., a fixed volume of cytoplasm in these 2D models) is essential for this mechanism to work. In contrast, the Rac/Rho model (Figure 4(c)), where

active Rac promotes polymerization at the front and active Rho promotes contractility at the rear, only produces a stable half-moon steady-state geometry if the initial circular geometry is allowed to evolve without imposition of a fixed 2D area.

The myosin contractility model (Figure 4(d)) treats the binding of myosin to the actin network explicitly, along with actin polymerization/depolymerization. The main assumptions are: the actin cytoskeleton behaves like a viscous fluid with viscosity η and myosin exerts an isotropic contractile stress on the actin network that is proportional to the local concentration of bound myosin. Unbound myosin diffusion is rapid enough to consider the unbound myosin concentration uniform throughout the cell. Finally, the polymerization rate is decreased with increasing myosin concentration. Interestingly, the actomyosin contraction mechanism alone fails to reproduce the steady-state keratocyte morphology or motility, although the characteristic half-moon shape and gliding motility seem to be stable for long periods with certain parameter combinations. However, coupling the Rac/Rho mechanism with actomyosin contractility does robustly recapitulate keratocyte morphology and behavior. Actin polymerization rate was assumed to be proportional to Rac concentration and myosin contractile stress was proportional to Rho activation.

Notably, the combined model gives rise to oscillations of the cell front without changes to the cell velocity, providing a possible explanation to experimental observations reported earlier (Barnhart *et al.*, 2011).

In this Section, we have sampled the diversity of biophysical mechanisms that have inspired conceptual mechano-chemical models of actin-driven motility. The control of cell motility and shape fascinate both experimentalists and theoreticians, motivating the design of these conceptual models. However, while some mechanisms or combinations of mechanisms appear to be emerging, further attention to cells other than keratocytes or to other types of actin-driven cell behaviors (e.g., endocytosis, microbe motility, and morphogenesis during cell differentiation) is required. Furthermore, the majority of models address experiments for cells grown on a flat surface rather than a more realistic 3D extracellular matrix with its own rheology and adhesion properties. And finally, none of these models address how a multitude of identified actin-binding proteins and signaling systems regulate actin dynamics. To address the latter class of problems, more detailed biochemical kinetic models are required.

Detailed Biochemical Modeling

The biochemistry underlying lamellipodial actin dynamics is extremely complex. Despite their multiple benefits, conceptual models are not meant to capture the biochemical mechanisms involved in actin cytoskeleton dynamics. Molecularly explicit mathematical models, or as we call them here detailed biochemical models, aim to simulate the interactions of any or all of the many proteins that orchestrate spatio-temporal dynamics of actin polymerization. Such models, when combined with experimental studies, serve as invaluable tools for elucidating the function of individual molecules or collective behaviors within a complex biochemical system. An ultimate goal of this approach may be to identify therapeutic targets and strategies. Detailed models also may offer the opportunity to address questions that are not feasible experimentally. Such ‘In silico’ biochemical simulations can generate new predictions and hypotheses to develop and direct future experimental studies. Detailed biochemical models of a complex reaction network also serve as knowledge integration platforms through which information about a physiological system can be easily shared, updated, enriched, reused, and adapted for another experimental or modeling study. We recently published a short perspective on the benefits of detailed models with a particular application to actin dynamics (Ditlev *et al.*, 2013).

In silico. In biology the term indicates experimentation using computational or mathematical model to test or generate hypotheses or make predictions about behavior of a biological system. Virtual experiment – simulation of a model with a particular set of initial conditions, parameters, and geometry.

To begin thinking about the basic principles underlying polymerization, it would be a useful thought exercise to

imagine that you have attached one link (the nucleator or seed) to a wall and then attach more links one by one at your pace (rate) to form a chain (polymer). The total assembly rate (number of links added per unit of time) for your chain is defined by your pace. Next, you might have more people who would like to engage in this activity and you provide them with the nucleator links. Let’s assume they work at about the same pace as you. We are interested in the total rate of assembly (total increase in the number of links in all the chains). If you had one more person working with you, the total assembly rate would be twice your rate; if you had nine more people the total rate would increase to 10 times your rate. Alternatively, think about the situation where you do not provide other people with the nucleators, but just want them to share access to your single chain. Then everyone will be adding the links to just one chain end and the total assembly rate will not increase. This situation is analogous to polymerization of actin monomers at the barbed ends. The more available barbed ends there are in the system the higher the assembly rate will be – the polymer ends serve as ‘catalysts’ of the assembly process. As we noticed in the example with the chains, the assembly rate linearly depends on the number of starters (number of barbed ends in case of actin polymer). Of course, this depends on an ample abundance of links to the crew you have charged with building the chains (i.e., that the availability of G-actin is not limiting). The same is true for the disassembly rate; the more available pointed ends there are the higher the disassembly rate will be. In order to coordinate this process in space and time the cells have an abundance of actin regulatory proteins which can control availability of filament ends, nucleation, elongation, branching, filament length, bundling, and turnover.

We will first briefly describe the major steps and components of the actin polymerization cycle. The process of actin filament treadmilling (Figure 5(a)) lies at the heart of the actin cycle. Broadly speaking, treadmilling refers to the steady-state polymerization at one end of a filament matched by an equal rate of net depolymerization at the other end; if the system is not at steady state, the filament may grow or shrink depending on which rate dominates. An individual actin monomer (G-actin) cycles through a series of biochemical events: (1) addition of ATP-bound G-actin at the barbed (growing or plus) end of F-actin, (2) hydrolysis of ATP on an F-actin subunit to generate to ADP-Pi-actin, (3) release of Pi from the ADP-Pi subunit, (4) dissociation of ADP-G-actin from the pointed (minus) end, and (5) nucleotide exchange of ADP-G-actin to recharge the monomer into the ATP-G-actin form. Profilin catalyzes the last step of ADP for ATP exchange on the actin monomer so that it may enter the next round of polymerization. Polymerization (Figure 5(b)) is stimulated by the nucleation-promoting factors (NPFs) associated with the inner side of the plasma membrane. NPFs are spatio-temporally regulated and in their active state induce branched (by activating Arp2/3 protein complex) or linear (by activating formin proteins) polymerization of actin filaments. Polymerizing filaments are proposed to push plasma membrane of motile cells by the elastic Brownian ratchet mechanism (Mogilner and Oster, 1996; Peskin *et al.*, 1993) as discussed above. A pool of monomeric actin that is necessary for assembly at the barbed ends is maintained by the buffering activity of thymosin- β 4.

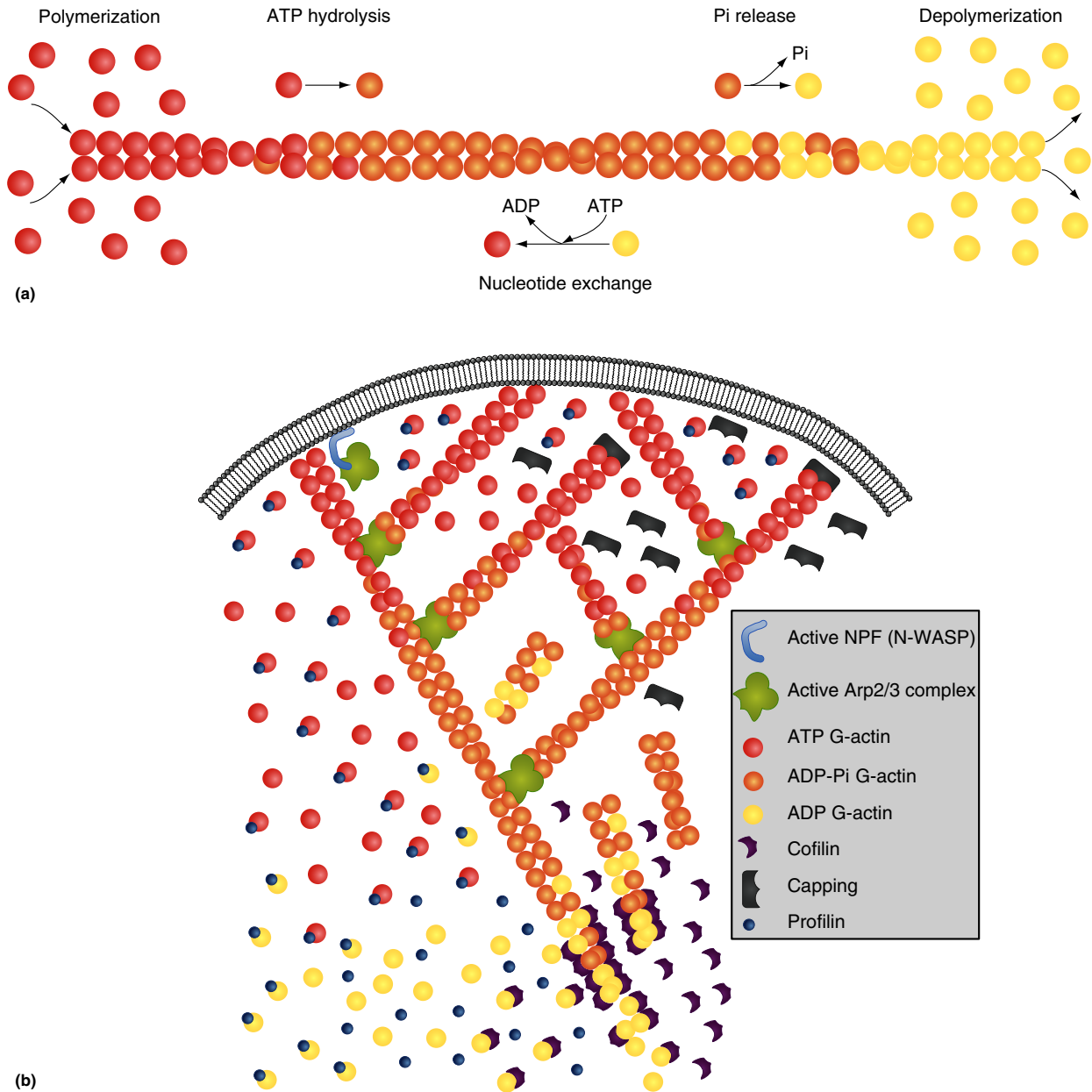


Figure 5 (a) The treadmilling process. Actin monomer undergoes repeated cycle of assembly, ATP hydrolysis, Pi dissociation, disassembly, and nucleotide exchange. (b) Actin dendritic nucleation mechanism. The cartoon depicts several major biochemical processes underlying dendritic nucleation: N-WASP induced Arp2/3 activation and branching, filament treadmilling, cofilin-mediated severing, profilin-mediated ADP/ATP exchange on G-actin, capping of barbed ends, incorporation of ATP-G-actin at the barbed ends, and disassembly at the pointed ends. Adapted from the original cartoon from Pollard, T.D., Borisy G.G., 2003. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* 112, 453–465.

Capping proteins (Caps) terminate growth of the barbed ends. Actin filaments can be destabilized by depolymerization at the pointed ends and by fragmentation into shorter filaments. There are also mechanisms that stabilize filamentous actin by protecting it from depolymerization and fragmentation.

Early work, before the availability of much kinetic data, derived mathematical models based on the thermodynamics of polymer formation (Oosawa and Kasai, 1962). These early modeling studies resulted in discovery of the critical concentration (above which polymerization occurs in solution) of G-actin (Oosawa *et al.*, 1959) and the treadmilling

phenomenon where polymerization depends on the nucleotide state of actin monomers (Wegner, 1976). Subsequently, an increasing amount of quantitative data related to multiple aspects of actin polymerization made development and validation of detailed biochemical models possible (Ditlev *et al.*, 2013).

Bindschadler *et al.* (2004) built a first comprehensive model of the actin cycle that couples ATP-hydrolysis with actin treadmilling. This is an ODE model that predicts the complete steady-state nucleotide profile (Figure 6(a)) within the filaments and demonstrates that Pi release from ADP-Pi increases

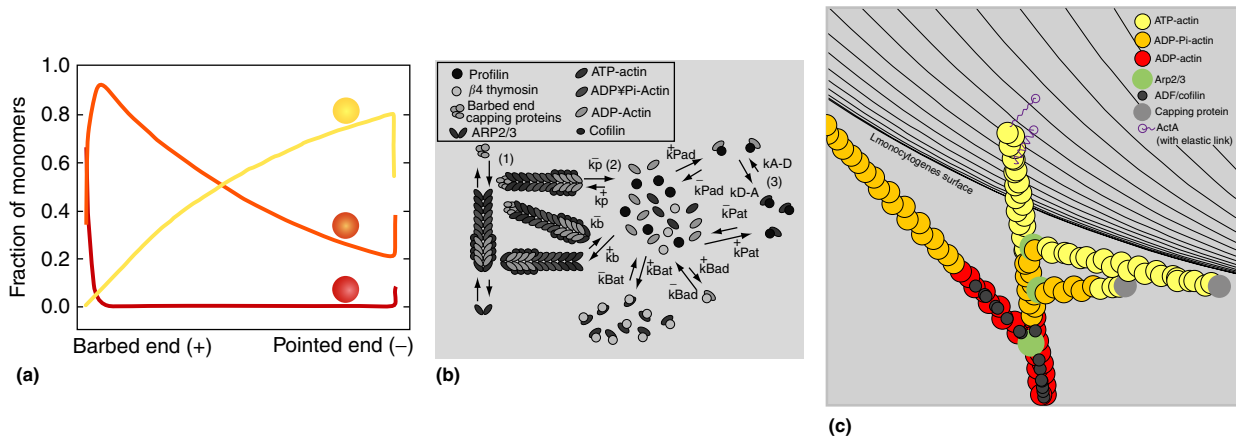


Figure 6 (a) Nucleotide state profile. Simulation results from (Bindschadler *et al.*, 2004) showing fractions of ATP-bound (red curve), ADP-Pi-bound (orange curve), and ADP-bound (yellow curve) monomers along 370-subunit long ($1\ \mu\text{m}$) filaments. (b) Schematic of actin cycle. Mechanisms included in a model by Bindschadler *et al.* (2004). (c) Actin branching near bacterial surface. A simulation video frame from (Alberts and Odell, 2004) showing a branched actin filament attached (via ActA) to a surface of *L. monocytogenes*. Collisions of filaments with the bacterial surface produce force propelling the bacterium in the cytosol of the host cell.

the rate of actin filament treadmilling. Essentially this is a model of actin filament assembly and disassembly from purified actin and its key regulatory proteins (Figure 6(b)): ‘explicitly’ modeled profilin and $\beta 4$ -thymosin and ‘implicitly’ modeled barbed-end and pointed-end cappers and cofilin. A year later Vavylonis *et al.* (2005) built a nonsteadystate ‘discrete’ model using kinetic Monte-Carlo simulations based on experimental data (Fujiwara *et al.*, 2002; Kuhn and Pollard, 2005). Eventually the model was validated and refined through experimental observations of individual actin filaments by total internal reflection fluorescence (TIRF) microscopy (Fujiwara *et al.*, 2007). This process of model construction originating from the available experimental data, going through model modifications to fit that data, and eventually making predictions that can be verified experimentally, exemplifies how modeling and experiment synergize.

Explicitly modeled refers to a molecular species that has a specific variable assigned to it in the model description. Implicitly modeled refers to a species that is not defined with a variable but rather with an expression that describes its effect on the system behavior.

An example of a detailed discrete model that tracks individual filaments is a dendritic nucleation model by Schaus and Borisy (Schaus *et al.*, 2007). It is a stochastic 2D computer model that allowed investigation of filament self-organization patterns at the leading edge of the lamellipodial protrusion. The model is based on the ‘dendritic-nucleation/array-treadmilling’ (Figure 5(a)) mechanism and incorporates (1) spatially discretized G-actin diffusion, (2) Monte-Carlo filament kinetics, and (3) elastic behavior of the filament and the membrane. Branching was restricted to the region within 5.4 nm from the plasma membrane while barbed ends were protected from capping in this narrow area. Irrespective of initial filament orientation, the filaments self-organized to

form a branched network with the filament orientation angle of $\pm 35^\circ$ consistent with electron microscopy observations. The system tends to increase the capping rate for the larger angles rather than increasing the branching rates for the smaller filament orientation angles. The authors demonstrated that protection from capping in the 5.4 nm zone was critical for reproducing realistic dendritic actin pattern while branching localization was not required. Membrane flexibility did not affect pattern orientation significantly. However, filament orientation angle was very sensitive to the increase in the filament bending stiffness resulting in a different angle distribution ($70/0/-70^\circ$). This comprehensive discrete model brings to light the idea that experimentally observed self-organization of the dendritic network in the lamellipodium emerges from the essential dynamic properties of the network.

Alberts and Odell (2004) constructed a large hybrid discrete-continuum model of dendritic actin nucleation to reconstitute the motility of *L. monocytogenes* *in silico*. *Listeria* is an intracellular rod-shaped bacteria that hijacks the actin cytoskeleton of the host cell for assembly of the actin comet tails. The comet tails propel the bacteria within the cytoplasm and into the neighboring cells causing the infection to spread. The model combines biochemistry and mechanistic aspects of actin-based motility of *Listeria* that presents then NPF ActA on its surface to activate the Arp2/3 complex in the host cell. An Arp2/3-dependent dendritic nucleation scheme (Figure 6(c)) encompasses biochemical events of the model. Actin monomers and actin-binding proteins are described continuously with reaction-diffusion equations (characterized as protein concentrations). Interactions of continuous players with each other and with individual actin filaments (discrete components) are defined deterministically. Filaments are modeled stochastically meaning that probability (not concentration) of binding/dissociation of each individual filament is computed at each time-step. The growth of a filament depends on the state variables that keep track of location, orientation, and biochemical state of that filament. Various forces act on the simulated bacterium; (1) link force of ActA protein holding a

filament and the bacterium together, (2) collisions of filaments with the bacterium push them apart, (3) Brownian forces with random orientations act on filaments and on the bacterium. The model reproduces the gross qualitative behavior of *L. monocytogenes* with respect to the average speed, persistence of motion, and comet tail morphology. Additionally the simulated bacterial trajectories exhibit the small-scale saltatory (repeated runs and pauses) motion observed by experimentalists. Alberts and Odell used their model to investigate the nature of *Listeria* nano-saltatory motility and concluded that pauses correlated with the Brownian motion and coordinated breakage and formation of ActA links between filaments and the bacterium. The 'in silico reconstitution' allows exploration of bacterial actin-based motility mechanisms. However it was extremely computationally expensive and required 80 Linux processors to solve thousands of differential equations.

Besides Arp2/3, cells have other nucleators of actin polymerization. One class of nucleators consists of various formin proteins (Paul and Pollard, 2009; Pring *et al.*, 2003; Pruyn *et al.*, 2002). They induce formation of unbranched actin filaments in filopodia, polarized actin fibers, cytokinetic ring, and stress fibers. Formins associate with the barbed ends of actin filaments and stimulate elongation while being processively attached to the growing ends. Highly conserved C-terminal FH1–FH2 domains of formin are key players in this process. A doughnut-shape dimer of FH2 domains forms a ring around the barbed end and an unstructured FH1 domain has multiple profilin-binding sites and therefore delivers profilin–actin complexes for F-actin assembly. Using available experimental data (Kovar *et al.*, 2006), Vavylonis built a stochastic comprehensive model (Vavylonis *et al.*, 2006) of FH1–FH2-dependent actin polymerization. The FH1 domain recruits profilin-G-actin from the cytosol. Since FH1 has many profilin-binding sites the concentration of profilin-actin near the barbed end increases. Flexibility of the FH1 domain allows for rapid collisions between profilin and the barbed end to occur. The collision results in a transfer of profilin-actin to the barbed end if the FH2 domain is in the open state. When the FH2 domain is in the closed state, it prevents actin subunit addition. Based on this mechanism, simulated elongation rates were consistent with the experimental rates for four different formins (Kovar *et al.*, 2006). The model showed that with optimal profilin concentration, a maximal elongation rate can be higher than elongation of actin alone. However, high profilin concentration inhibits elongation because profilin outcompetes profilin-bound actin from FH1 domain.

With that historical background, we now turn to a comprehensive and extensible continuum deterministic model developed in our lab: 'An open model of actin dendritic nucleation' (henceforth referred to as the Open Model) (Ditlev *et al.*, 2009) that is currently the most comprehensive mechanistic model of actin dynamics. The Open Model is implemented in the Virtual Cell software system (VCell) (Cowan *et al.*, 2012; Resasco *et al.*, 2012; Schaff *et al.*, 1997; Slepchenko and Loew, 2010). VCell is a computational environment for modeling and simulation in biology that is designed for both experimentalists and theorists. It has a convenient graphical user interface for formulating a model: defining compartments and their geometries as well as molecular players and their

interactions. Both deterministic and stochastic simulations for either 'compartmental' or 'spatial' models can be run in VCell. Models can include molecular reaction–diffusion–advection (flow), membrane fluxes, and can solve the corresponding 'ODEs' or 'PDEs' or stochastic equations for any cellular geometry. VCell offers a variety of data visualization tools: graphs, kymographs, images, export options to spreadsheets, images, and movies. Importantly, it is built on a database, so scientists can access public models and adapt or extend them for their own research.

Compartmental model. Model of a system in which all the species are well-mixed and therefore variables are changing only in time. The system geometry may be approximated by considering the relative volumes and surface to volume ratios of membrane enclosed compartments. The equations may be described by a set of ODEs that include not just reaction kinetics, but also fluxes between compartments or between a volumetric compartment and an adjacent membrane. For systems with small numbers of molecules, reaction probabilities can be treated with stochastic simulations

Spatial model. Model of a system in which the species are distributed in space and undergo transport and/or diffusion processes within explicit geometries. Described by a set of PDEs for deterministic simulations or by combined reaction probabilities and Brownian dynamics for stochastic simulations.

The Open Model (Ditlev *et al.*, 2009) was developed in an effort to account for the key regulatory proteins and mechanisms governing actin dynamics. It builds upon all the mechanisms simulated in the Bindaschadler *et al.* (2004) model and expands the latter with Arp2/3 activation by N-WASP at the membrane, Arp2/3-mediated side branching, regulation of actin dynamics by cofilin, and fragmentation and annealing of actin filaments. The molecules that are explicitly included in the model are actin, a NPF in the membrane (ex. N-WASP), Arp2/3 complex, Cap, profilin, ADF/cofilin, thymosin- β 4, and various associations of actin with ATP, ADP-Pi, and ADP. Despite such a short list of molecular players the model required 60 species and 155 reactions to capture the core mechanisms driving actin dendritic nucleation. In Figure 7(a), the blue rectangles, which overlay the corresponding regions of the reaction network, summarize the mechanisms that are explicitly defined in the Open Model. The model and the simulation results are publicly available through the Virtual Cell database. The simulations can be run on any geometry and therefore facilitate virtual experiments incorporating specifying initial molecular distributions, diffusion, and advection in realistic geometries extracted from microscope imaging data. The Open Model actually contains many such virtual experiments, with different initial conditions, system geometries, and boundary conditions; in this sense, it is actually a collection of mathematical models all examining aspects of a common set of basic biophysical mechanisms.

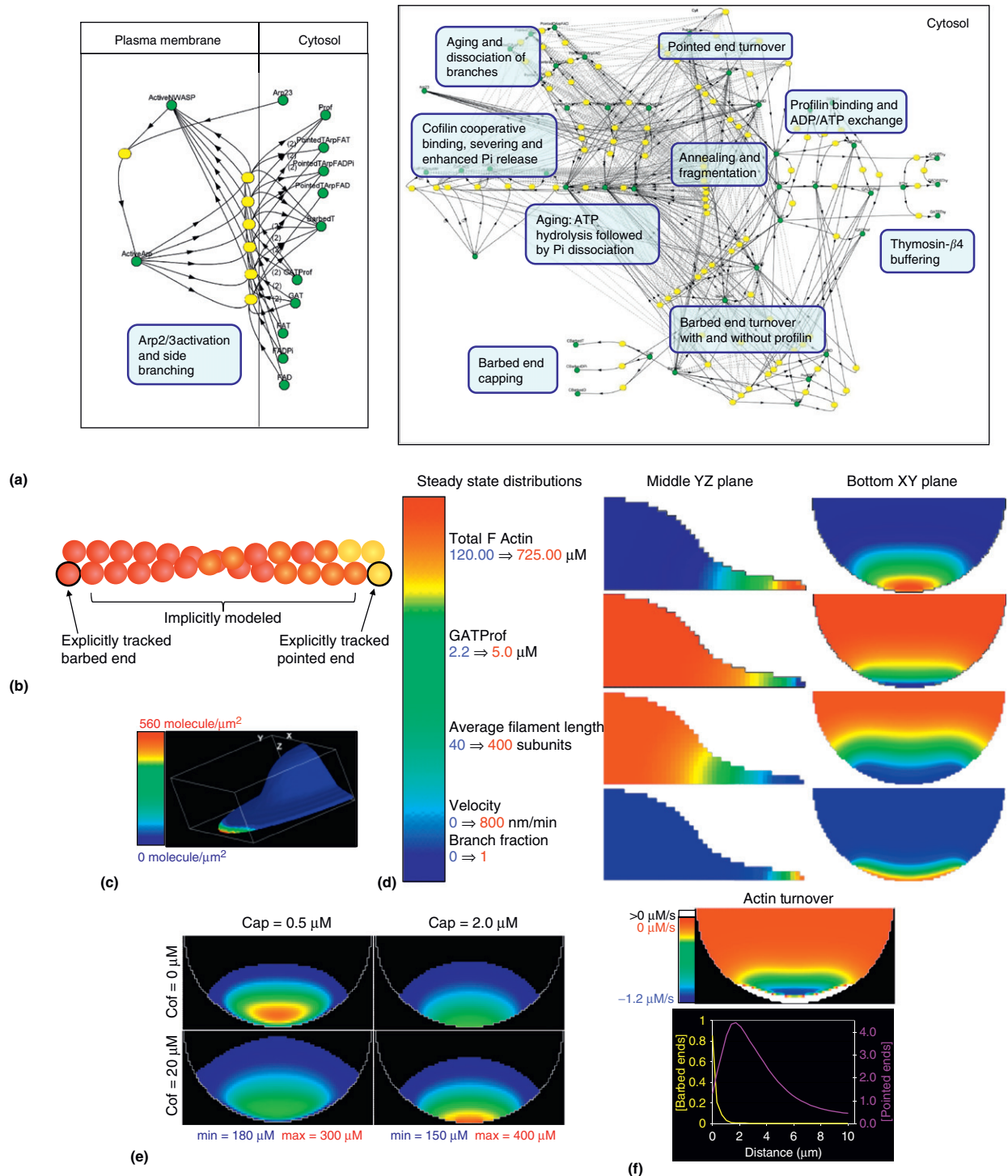


Figure 7 An Open Model of actin dendritic nucleation (Ditlev *et al.*, 2012). (a) Reaction diagram from Virtual Cell. Green circles – species. Yellow ovals – reactions/transformations. Reactions occurring at the inner side of the membrane *in vivo* are placed in the ‘plasma membrane’ compartment. Membrane reactions can recruit cytosolic species. Each rounded rectangle with a mechanism label points to a set of reactions involved in the indicated mechanism. (b) Nucleotide state tracking. Nucleotide states of the barbed end (red ball outlined in black) and pointed end (yellow ball outlined in black) are explicitly modeled. The states of all the subunits in the interior of the filament are modeled implicitly (state of each subunit is unknown) by three state variables FAT, FAD-Pi, and FAD (concentrations of all the ATP-, ADP-Pi, and ADP-bound subunits respectively). The state of the newly formed pointed end (after the yellow ball outlined in black dissociates) is incorporated into the dissociation rate. (c) Realistic 3D geometry of a protruding cell. Polymerization is triggered by ActiveNWASP (color coded) at the tip of the lamellipodium (7246 molecules). (d) Simulation results from spatial model with Arp2/3 activation. Minimum (blue value) and maximum (red value) are provided for each variable. (e) Interplay of cofilin and Cap. Model predictions for the total amount of F-actin under varied cofilin and capping concentrations. At low capping level increasing cofilin results in F-actin decrease. At high capping level increasing cofilin results in F-actin decrease. (f) Sharp transition from assembly to disassembly $1 \mu\text{m}$ from the leading edge. (Upper) Simulation results (bottom xy slice) from 3D model showing assembly (white) and disassembly (color coded) rates. (Lower) Abundance of barbed and pointed ends across the lamellipodial tip.

The core of the Open Model is the treadmilling of actin filaments. All the events that accompany and drive this phenomenon are depicted in the cartoon in [Figure 5\(a\)](#). This and many other schematic representations of actin treadmilling grossly oversimplify the actual process. [Figure 5\(a\)](#) shows that monomers in the ATP form add to the barbed ends and dissociate in the ADP form at the pointed ends of the filaments. The situation *in vivo* is more complex: monomers in all the states (ATP, ADP-Pi, and ADP) may bind to and dissociate from either barbed or pointed ends of the filaments. Suffixes T, Dpi, and D are used in the model to indicate the nucleotide state of actin ‘species’. The model accurately captures complexity of the treadmilling process with the mechanisms for tracking the nucleotide state and the processes occurring at the barbed and pointed ends (barbed- and pointed-end turnover).

Some of the aspects of cytoskeletal dynamics are difficult to model explicitly. For example, tracking the nucleotide state of each monomer within the filaments could drastically increase combinatorial complexity of the model. Therefore, the nucleotide state is modeled explicitly only at the barbed and pointed ends, resulting in six variables: three states (T, DPi, and D) for each of barbed and pointed end of the filament. The states of all the actin monomers in the interior of the filaments are described implicitly by three state lumped variables FAT (F-actin in ATP form), FAD-Pi (F-actin in ADP-Pi form), and FAD (F-actin in ADP form) ([Figure 7\(b\)](#)). The FAT in the filament undergoes hydrolysis ($\text{FAT} \rightarrow \text{FAD-Pi}$) and then inorganic phosphate dissociation ($\text{FAD-Pi} \rightarrow \text{FAD} + \text{Pi}$). Hydrolysis and Pi dissociation reactions are modeled explicitly and are applied to all the actin species (not only F-actin) in the reaction network.

The six variables that track the nucleotide states of barbed and pointed ends depend on the hydrolysis and phosphate (Pi) dissociation and on the association and dissociation of actin monomers in different nucleotide states. For instance, there are 18 reactions characterizing interactions of monomers with the barbed ends: three types of barbed ends reacting with G-actin in three states as well as with profilin-bound G-actin in three states. Dissociation events of monomers from the ends presented the difficulty of assigning a nucleotide state to the newly formed end (after a monomer dissociates). The monomer next to the explicitly tracked pointed end can be in any of the three nucleotide states because it was implicitly modeled ([Figure 7\(b\)](#)). The model assumes, therefore, that the nucleotide states at this penultimate position will have approximately the same fractional distribution of states at the ends. This assumption was applied to define dissociation rates of actin monomers from various types of barbed and pointed ends.

Actin filaments (F-actin) were shown to undergo the process of end-to-end annealing experimentally ([Andrianantoandro et al., 2001](#); [Sept et al., 1999](#)). Similarly fragmentation of F-actin was demonstrated to be a critical mechanism underlying actin dynamical behavior ([Berro et al., 2010](#); [Schmoller et al., 2011](#)). The annealing and fragmentation were treated as forward and reverse reactions in the Open Model reaction network ([Ditlev et al., 2009](#)). There are nine forward reactions: pointed ends in three states can anneal to the barbed ends in three different states. The rate of annealing reaction decreases as the length of the fragment increases. Additionally annealing

rate expression is scaled by a factor of $(1 - \text{BrF})$ in order to reduce the rate by the fraction of filaments containing branches. The filaments with high BrF (branch fraction) diffuse more slowly and therefore are unlikely to undergo annealing. BrF is determined as the sum of all species where Arp2/3 is bound to FA (i.e., branches) divided by the total concentration of filaments (i.e., half the concentration of all ends).

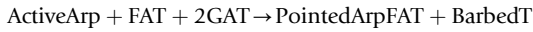
Cofilin is an essential actin-binding protein that can stabilize as well as destabilize individual actin filaments and the actin network. We will address the overall physiological effects of cofilin on actin dynamics later during discussion of the results of the model. Kinetics and thermodynamics of cofilin-actin interaction are well characterized experimentally ([Blanchoin and Pollard, 1999](#); [Cao et al., 2006](#); [De La Cruz, 2009](#); [Elam et al., 2013](#)) and have been modeled ([De La Cruz and Sept, 2010](#)). The Open Model has a simplified mechanism of cofilin interaction and severing of actin filaments. It includes (1) acceleration of Pi release from FADPi subunits, (2) severing only when two or more neighboring FAD subunits are bound to cofilin molecules, (3) the cooperative binding phenomenon where cofilin will have higher binding affinity for the FAD subunit which has a neighboring cofilin-bound FAD.

Availability of experimental data from the *in vitro* (test tube) experiments provided a great opportunity to validate the Open Model by running ODE simulations representing a closed well-mixed compartment. Numerous ‘compartmental’ (ODE) models were run for different values of total actin, Cap, cofilin, profilin, and thymosin- $\beta 4$. To mimic the *in vitro* experimental conditions, Arp2/3-mediated nucleation and branching were not included in these ODE simulations. Overall the results reproduced the effects of nucleotide exchange, profilin and thymosin- $\beta 4$ on actin turnover, F-actin concentration, and filament length reported and simulated by ([Bindschadler et al., 2004](#)). Additionally, simulations agreed well with the experimental data on the effect of Cap, fragmentation, and annealing on the average filament length. Overall, the simulations reproduced numerous experimental studies in which each of these system variables had been carefully controlled ([Carlier and Pantaloni, 2007](#); [dos Remedios et al., 2003](#); [Pollard and Borisy, 2003](#); [Pollard and Cooper, 1986](#)) (and see the article by Carlier *et al.* in this work). Running compartmental simulations not only offered an initial model validation, but the final steady state values for these calculations served as appropriate initial conditions (stable starting concentrations) for the spatial simulations with Arp2/3 activation.

One of the key VCell capabilities is to simulate cellular dynamics in space. The ultimate goal of the Open Model was to investigate the process of actin dendritic nucleation in lamellipodial protrusion. For this purpose the Open Model adopted an idealized 3D geometry of a migrating cell with a thin lamellipodium and a larger volume cell body ([Figure 7\(c\)](#)) to run additional virtual experiments. In spatial simulations, biochemical reactions are complemented with diffusion and flow of the molecular species. As a result Virtual Cell generates and solves a system of ‘PDEs’ constrained by mass conservation.

The mechanism of Arp2/3-mediated branch formation is a multiple step process. In the Open Model it was reduced to a single step to achieve computational efficiency and therefore

reduce simulation time for 3D virtual experiments. Arp2/3 activation and side branching was implemented as proposed in the modeling study by (Carlsson *et al.*, 2004). Cytosolic Arp2/3 complex is recruited to the membrane and activated by a membrane-bound NPF (activeNWASP). Activated and membrane-localized Arp2/3 recruits two actin monomers (unbound or bound to profilin) and binds to the side of the preexisting filament. Monomer recruitment and side binding occurs in one step and results in production of a branch and a free barbed end. This is an example of one of the branching reactions:



There are six nucleation reactions because FA can be in three states and GAT can also be in the GAT-Prof (profilin-bound) form.

Diffusion of all the monomeric species and actin-binding proteins was assigned a diffusion coefficient of $D_{\text{GActin}} = 5 \mu\text{m}^2 \text{s}^{-1}$, consistent with experimental data. Long and branched actin filaments diffuse slower than monomers, however. One of the modeling challenges was to incorporate the dependence of F-actin fragment diffusion on the filament length and branching properties of the polymer. The following equation was used to model F-actin diffusion:

$$\frac{D_{\text{GActin}}}{\text{FilamentLength}} \times (1 - \text{BrF})$$

'FilamentLength' is the average length of actin filament. It is calculated by dividing total F-actin (which includes all nucleotide FA forms, FA species with bound cofilin or Arp2/3, and all end species) by the number of filaments, which is calculated as equal to half the number of ends. Diffusion is also reduced (multiplication by $(1 - \text{BrF})$) by the fraction of filaments that are branched.

The rapidly polymerizing branched actin network produces force against the plasma membrane (Figure 5(b)). This results in rearward movement of the F-actin network and also lamellipodial protrusion. The flow of the network is modeled explicitly with an advection (velocity) expression: $v_i = v_i^{\text{max}} \times \text{BrF}$. The flow is proportional to 'BranchFraction' because only actively polymerizing filaments that are connected to the dendritic network are expected to undergo the rearward movement. An experimentally determined value for $v_i^{\text{max}} = 800 \text{ nm min}^{-1}$ (Ponti *et al.*, 2005) was employed.

To simulate actin dynamics at the leading edge of the protruding cell in 3D, the Open Model triggered the dendritic nucleation mechanism by activating NPF (ActiveNWASP) at the membrane at the leading edge (Figure 7(c)). Starting with species concentrations obtained from the steady state of corresponding compartmental simulations, Arp2/3 complex becomes activated by ActiveNWASP and stimulates branching at the front of the cell. In live cells growing branched filaments produces force by the Brownian ratchet mechanism (refer to the Section Conceptual Models of Mechano-chemical Control of Cell Motility and Shape) that pushes the plasma membrane forward. Even though Virtual Cell does not model the moving membrane or force-velocity coupling, it can represent the resulting rearward motion of actin network by a combination of the advection and diffusion, as discussed above. Simulation results quantitatively reproduced F-actin accumulation in the

lamellipodium, velocity fields of the rearward motion of actin network, average filament length, and profilin-bound G-actin concentration (Figure 7(d)). F-actin accumulation is a hallmark of lamellipodial protrusion (Abraham *et al.*, 1999) and it was predicted accurately by the model with a maximum concentration of $725 \mu\text{M}$ at the tip and $120 \mu\text{M}$ (depletion) in the body of the cell; the total actin (GA + FA) in the system was $200 \mu\text{M}$. Due to rapid assembly of branched actin network in a small volume of the lamellipodial tip, profilin-bound G-actin becomes depleted (Figure 7(d)). The simulated average length of the filaments is consistent with electron microscopy measurements (Svitkina and Borisy, 1999). Filaments at the tip are much shorter (40 subunits) than in the body of the cell (400 subunits). The velocity field that characterizes rearward motion of the network also agrees with the experimental observations by speckle microscopy (Ponti *et al.*, 2005).

The Open Model highlighted a few important aspects of cooperative behavior of actin regulatory proteins. The role of cofilin in regulating actin network dynamics has been a subject of debate. It was proposed to promote depolymerization and decrease in F-actin levels (Hotulainen *et al.*, 2005; Kueh *et al.*, 2008) as well as stimulate polymerization and cell protrusion and produce new filaments (Endo *et al.*, 2003; Ghosh *et al.*, 2004; Ichetovkin *et al.*, 2002). The model offers a potential explanation of the conflicting reports. After running multiple simulations with various initial concentrations of cofilin and Cap we noticed that cofilin could either suppress or promote Arp2/3-dependent F-actin accumulation (Figure 7(e)). With low amounts of both cofilin and Cap there are enough barbed ends to produce F-actin. Increasing Cap concentration will result in a lower number of barbed ends and therefore will inhibit F-actin increase. When cofilin concentration is high and Cap is low then F-actin production is also inhibited due to small G-actin pool. Increasing capping concentration will elevate monomer concentration and therefore will result in F-actin accumulation at the tip (Figure 7(e)). These results are consistent with the model by (Carlsson *et al.*, 2004) that predicted requirement of barbed-end capping for F-actin production due to high severing activity.

Investigation of assembly and disassembly rates in the simulated 3D lamellipodium reproduces a complex dynamical behavior of the dendritic actin network. Figure 7(f) reveals a sharp boundary between polymerization and depolymerization located $1 \mu\text{m}$ from the cell edge, where negative rates that indicate disassembly are color coded and positive rates are shown in white (Figure 7(f) top). The sharp assembly-disassembly transition was originally discovered with the speckle microscopy studies (Ponti *et al.*, 2004, 2005). The Open Model successfully reproduces and provides an explanation for this phenomenon. The boundary is positioned right between the regions of maximum barbed (right at the tip) and pointed end concentrations ($\sim 2 \mu\text{m}$ from the edge) (Figure 7(f) bottom). Therefore the assembly-disassembly transition is very sensitive to the debranching since it exposes additional pointed ends that were previously capped by the Arp2/3 complex.

We briefly highlight two additional studies, for which the Open Model served as a starting point. These are examples of how the original model was adapted to different experimental systems and consequently to another geometric and biochemical conditions.

Fluorescent recovery after photobleaching (FRAP) and chromophore-assisted laser inactivation (CALI) microscopy techniques are used extensively to study actin dynamics. In order to interpret the experimental result of FRAP and CALI of enhanced green fluorescent protein (eGFP)-Cap, Kapustina developed a model (Kapustina *et al.*, 2010) where eGFP-Cap could be virtually bleached and remain active (as in FRAP) or become inactive (as in CALI) (Figure 8(a)). The model adapts all the mechanisms from the Open Model with two enhancements: modified Cap-containing reactions, to which fluorescent and nonfluorescent species were added; mechanisms involving vasodilator-stimulated phosphoprotein (VASP) bundling and uncapping activities. The results from simulated FRAP experiments allowed estimates of Cap dissociation rate which was found to be much faster than *in vitro*. CALI simulations of experiments in which Cap activity was locally ablated, produced results that required anticapping activity. Thus, the model suggested that VASP exhibits cooperativity to induce the anticapping effect, consistent with suggestions from *in vitro* VASP biochemistry.

A model of Nck-induced actin comet tails (Ditlev *et al.*, 2012) expands the signaling mechanism that leads to N-WASP activation (Figure 8(b) top). Nck is a adaptor protein that normally is recruited to phosphorylated tyrosines; it recruits a number of actin regulatory molecules including N-WASP. Experimentally increasing the density of Nck molecules at the membrane resulted in a nonlinear response of actin polymerization. The source of the nonlinearity was discovered by testing, 'in silico,' three different stoichiometries of Nck/N-WASP/Arp2/3 interaction and quantitatively comparing simulated and experimental actin comets. Based on this

comparison, the authors proposed a 4:2:1 stoichiometry for Nck/N-WASP/Arp2/3 and a critical role of WIP (WASP interacting protein) in the signaling complex at the membrane. The model predictions were rigorously tested and confirmed using quantitative fluorescent microscopy. For instance the model accurately predicts comet tail length and the number of actin molecules in the comet (Figure 8(b) bottom).

In this section, we first provided a brief overview of the approaches toward detailed models of actin dynamics with a focus on the core mechanisms of actin polymerization and how it is regulated to drive cell motility. The models chosen were just illustrative and not meant by any means to comprehensively review the many contributions to this field. Rather, they were meant to exemplify what might be learned from continuum versus discrete or from compartmental versus spatial models and how these various approaches can be applied to the complexity of actin biochemistry. We then focused on how the Open Model was developed, including some of the key approximations required to confront the combinatorial complexity of polymerization and its regulation; indeed, the development of such approximations, which lump multiple explicit species and reactions into a small set of tractable variables and equations, is the most creative aspect of the modeling process. Finally, we described, using the Open Model, how Virtual Experiments could be deployed to provide fresh insight into several real experimental results and could even be used to guide the design of experiments. We encourage readers of this article to access and use the Open Model by logging onto the Virtual Cell (see Section Relevant Websites) and opening the public model called 'les: Actin Dendritic Nucleation.'

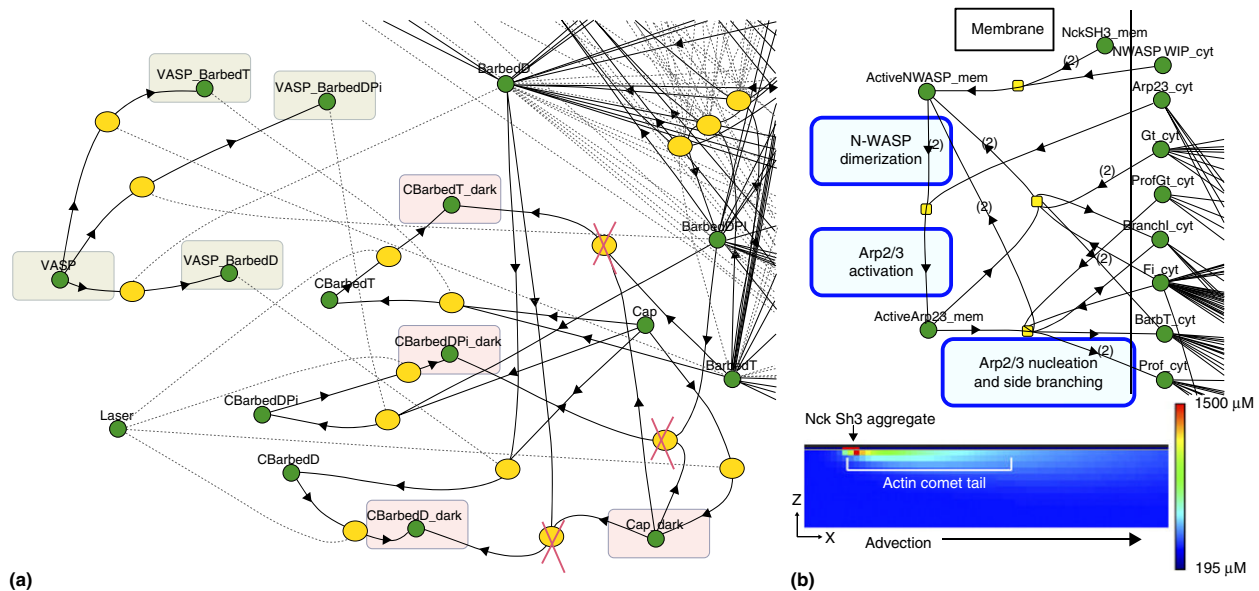


Figure 8 Examples of models based on the Open Model. (a) CALI and FRAP of Caps. Reaction diagram of barbed-end capping mechanism from the model by Kapustina *et al.* (2010). Note 'dark' Cap species in pink rectangles. In FRAP 'dark' Cap is active and able to participate in reactions. For CALI experiments the 'dark' Cap is inactive. VASP can bind each type of free barbed ends and therefore competes with Cap. (b) N-WASP/Arp2/3 signaling in actin comets. (Top) Arp2/3 activation at the membrane as in the model by Ditlev *et al.* (2012). Activation mechanism was modified to include Nck activation of N-WASP, WIP recruitment, and proposed 4:2:1 Nck/N-WASP/Arp2/3 stoichiometry. (Bottom) Simulation result for total F-actin in the actin comet induced by aggregation of Nck SH3 domains. Actin comets move along the membrane being propelled by polymerizing actin. Comet propulsion is modeled through the advection parameter.

Conclusion

We presented the dichotomy between ‘conceptual’ and ‘detailed’ models of actin-dependent cell biology as quite absolute. This has been largely the case in practice, where the two classes of models have been divided both philosophically and with respect to the kinds of questions being asked of them. With regard to the former, conceptual models are primarily concerned with a global picture of the physics underlying a cellular behavior; detailed models care about the role of each molecular player in the cellular orchestra. With regard to differences in research questions, conceptual models have been primarily concerned with the interplay of cell mechanics and polymer physics; detailed models have been concerned with the biochemistry of signaling to and regulation of actin dynamics. We can certainly look forward to the inevitable blending of these modeling approaches. But this will require new physical, mathematical, and computational technologies that can couple cell mechanics, moving boundaries, and complex reaction-diffusion systems. We are beginning to see the emergence of such approaches (Novak and Slepchenko, 2014).

To make this article maximally accessible, we have deliberately avoided the use of mathematical equations. However, mathematical modeling, of course, requires the solution of equations. Indeed, all of the studies we used to exemplify various modeling approaches, did produce their simulations by solving complex mathematical problems. This invariably requires the use of numerical methods and computers. We urge the mathematically inclined reader to go to the primary sources in the bibliography to see how the ideas we presented were translated into equations. Thankfully, there are many readily available software tools to help cell biologists develop mathematical models and run simulations from them (see Section Relevant Websites). We look forward to the time when modeling will be as common a tool for the cell biologist as the microscope or the pipet. Indeed, the most powerful application of modeling is as a tool that is fully integrated within an experimental study.

Acknowledgments

This work was supported by NIH grant P41 GM103313 from the National Institute for General Medical Science.

See also: Dynamic Integration: Dynamics: Dynamics of Gradient Sensing and Chemotaxis

References

- Abraham, V.C., Krishnamurthi, V., Taylor, D.L., Lanni, F., 1999. The actin-based nanomachine at the leading edge of migrating cells. *Biophysical Journal* 77, 1721–1732.
- Alberts, J.B., Odell, G.M., 2004. In silico reconstitution of *Listeria* propulsion exhibits nano-saltation. *PLoS Biology* 2, e412.
- Allard, J., Mogilner, A., 2012. Traveling waves in actin dynamics and cell motility. *Current Opinion in Cell Biology* 25, 107–115.
- Andrianantoandro, E., Blanchoin, L., Sept, D., McCammon, J.A., Pollard, T.D., 2001. Kinetic mechanism of end-to-end annealing of actin filaments. *Journal of Molecular Biology* 312, 721–730.
- Barnhart, E.L., Lee, K.C., Keren, K., Mogilner, A., Theriot, J.A., 2011. An adhesion-dependent switch between mechanisms that determine motile cell shape. *PLoS Biology* 9, e1001059.
- Berro, J., Sirotkin, V., Pollard, T.D., 2010. Mathematical modeling of endocytic actin patch kinetics in fission yeast: Disassembly requires release of actin filament fragments. *Molecular Biology of the Cell* 21, 2905–2915.
- Bindschadler, M., Osborn, E.A., Dewey Jr., C.F., McGrath, J.L., 2004. A mechanistic model of the actin cycle. *Biophysical Journal* 86, 2720–2739.
- Blanchoin, L., Pollard, T.D., 1999. Mechanism of interaction of Acanthamoeba actophorin (ADF/Cofilin) with actin filaments. *Journal of Biological Chemistry* 274, 15538–15546.
- Cao, W., Goodarzi, J.P., De La Cruz, E.M., 2006. Energetics and kinetics of cooperative cofilin–actin filament interactions. *Journal of Molecular Biology* 361, 257–267.
- Carrier, M.-F., Pantaloni, D., 2007. Control of actin assembly dynamics in cell motility. *Journal of Biological Chemistry* 282, 23005–23009.
- Carlsson, A.E., Wear, M.A., Cooper, J.A., 2004. End versus side branching by Arp2/3 complex. *Biophysical Journal* 86, 1074–1081.
- Cowan, A.E., Moraru Il, J.C., Schaff, B.M., Slepchenko, , Loew, L.M., 2012. Spatial modeling of cell signaling networks. *Methods in Cell Biology* 110, 195–221.
- Cramer, L.P., 1999. Role of actin-filament disassembly in lamellipodium protrusion in motile cells revealed using the drug jasplakinolide. *Current Biology* 9, 1095–1105.
- Danuser, G., Allard, J., Mogilner, A., 2013. Mathematical modeling of eukaryotic cell migration: Insights beyond experiments. *Annual Review of Cell and Developmental Biology* 29, 501–528.
- De La Cruz, E.M., 2009. How cofilin severs an actin filament. *Biophysical Reviews* 1, 51–59.
- De La Cruz, E.M., Sept, D., 2010. The kinetics of cooperative cofilin binding reveals two states of the cofilin–actin filament. *Biophysical Journal* 98, 1893–1901.
- Ditlev, J.A., Mayer, B.J., Loew, L.M., 2013. There is more than one way to model an elephant. Experiment-driven modeling of the actin cytoskeleton. *Biophysical Journal* 104, 520–532.
- Ditlev, J.A., Michalski, P.J., Huber, G., *et al.*, 2012. Stoichiometry of Nck-dependent actin polymerization in living cells. *Journal of Cell Biology* 197, 643–658.
- Ditlev, J.A., Vacanti, N.M., Novak, I.L., Loew, L.M., 2009. An open model of actin dendritic nucleation. *Biophysical Journal* 96, 3529–3542.
- Elam, W.A., Kang, H., De La Cruz, E.M., 2013. Biophysics of actin filament severing by cofilin. *FEBS Letters* 587, 1215–1219.
- Endo, M., Ohashi, K., Sasaki, Y., *et al.*, 2003. Control of growth cone motility and morphology by LIM kinase and Slingshot via phosphorylation and dephosphorylation of cofilin. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 23, 2527–2537.
- Fujiwara, I., Takahashi, S., Tadakuma, H., Funatsu, T., Ishiwata, S.I., 2002. Microscopic analysis of polymerization dynamics with individual actin filaments. *Nature Cell Biology* 4, 666–673.
- Fujiwara, I., Vavylonis, D., Pollard, T.D., 2007. Polymerization kinetics of ADP- and ADP-P(i)-actin determined by fluorescence microscopy. *Proceedings of the National Academy of Sciences of the United States of America* 104, 8827–8832.
- Ghosh, M., Song, X., Mouneimne, G., *et al.*, 2004. Cofilin promotes actin polymerization and defines the direction of cell motility. *Science* 304, 743–746.
- Herant, M., Dembo, M., 2010. Form and function in cell motility: From fibroblasts to keratocytes. *Biophysical Journal* 98, 1408–1417.
- Herant, M., Heinrich, V., Dembo, M., 2006. Mechanics of neutrophil phagocytosis: Experiments and quantitative models. *Journal of Cell Science* 119, 1903–1913.
- Herant, M., Marganski, W.A., Dembo, M., 2003. The mechanics of neutrophils: Synthetic modeling of three experiments. *Biophysical Journal* 84, 3389–3413.
- Hotulainen, P., Paunola, E., Vartiainen, M.K., Lappalainen, P., 2005. Actin-depolymerizing factor and cofilin-1 play overlapping roles in promoting rapid F-actin depolymerization in mammalian nonmuscle cells. *Molecular Biology of the Cell* 16, 649–664.
- Ichetovkin, I., Grant, W., Condeelis, J., 2002. Cofilin produces newly polymerized actin filaments that are preferred for dendritic nucleation by the Arp2/3 complex. *Current Biology* 12, 79–84.
- Kapustina, M., Vitriol, E., Elston, T.C., Loew, L.M., Jacobson, K., 2010. Modeling capping protein FRAP and CALI experiments reveals *in vivo* regulation of actin dynamics. *Cytoskeleton* 67, 519–534.
- Keren, K., Pincus, Z., Allen, G.M., *et al.*, 2008. Mechanism of shape determination in motile cells. *Nature* 453, 475–480.
- Kovar, D.R., Harris, E.S., Mahaffy, R., Higgs, H.N., Pollard, T.D., 2006. Control of the assembly of ATP- and ADP-actin by formins and profilin. *Cell* 124, 423–435.

- Kueh, H.Y., Charras, G.T., Mitchison, T.J., Brieher, W.M., 2008. Actin disassembly by cofilin, coronin, and Aip1 occurs in bursts and is inhibited by barbed-end cappers. *Journal of Cell Biology* 182, 341–353.
- Kuhn, J.R., Pollard, T.D., 2005. Real-time measurements of actin filament polymerization by total internal reflection fluorescence microscopy. *Biophysical Journal* 88, 1387–1402.
- Mogilner, A., Allard, J., Wollman, R., 2012. Cell polarity: Quantitative modeling as a tool in cell biology. *Science* 336, 175–179.
- Mogilner, A., Oster, G., 1996. Cell motility driven by actin polymerization. *Biophysical Journal* 71, 3030–3045.
- Novak, I.L., Slepchenko, B.M., 2014. A conservative algorithm for parabolic problems in domains with moving boundaries. *Journal of Computational Physics* 270, 203–213.
- Oosawa, F., Asakura, S., Hotta, K., Imai, N., Ooi, T., 1959. G-F transformation of actin as a fibrous condensation. *Journal of Polymer Science* 37, 323–336.
- Oosawa, F., Kasai, M., 1962. A theory of linear and helical aggregations of macromolecules. *Journal of Molecular Biology* 4, 10–21.
- Paul, A.S., Pollard, T.D., 2009. Review of the mechanism of processive actin filament elongation by formins. *Cell Motility and the Cytoskeleton* 66, 606–617.
- Peskin, C.S., Odell, G.M., Oster, G.F., 1993. Cellular motions and thermal fluctuations: The Brownian ratchet. *Biophysical Journal* 65, 316–324.
- Pollard, T.D., Borisy, G.G., 2003. Cellular motility driven by assembly and disassembly of actin filaments. *Cell* 112, 453–465.
- Pollard, T.D., Cooper, J.A., 1986. Actin and actin-binding proteins. A critical evaluation of mechanisms and functions. *Annual Review of Biochemistry* 55, 987–1035.
- Ponti, A., Machacek, M., Gupton, S.L., Waterman-Storer, C.M., Danuser, G., 2004. Two distinct actin networks drive the protrusion of migrating cells. *Science* 305, 1782–1786.
- Ponti, A., Matov, A., Adams, M., *et al.*, 2005. Periodic patterns of actin turnover in lamellipodia and lamellae of migrating epithelial cells analyzed by quantitative fluorescent speckle microscopy. *Biophysical Journal* 89, 3456–3469.
- Pring, M., Evangelista, M., Boone, C., Yang, C., Zigmond, S.H., 2003. Mechanism of formin-induced nucleation of actin filaments. *Biochemistry* 42, 486–496.
- Pruyne, D., Evangelista, M., Yang, C., *et al.*, 2002. Role of formins in actin assembly: Nucleation and barbed-end association. *Science* 297, 612–615.
- dos Remedios, C.G., Chhabra, D., Kekic, M., *et al.*, 2003. Actin binding proteins: Regulation of cytoskeletal microfilaments. *Physiological Reviews* 83, 433–473.
- Resasco, D.C., Gao, F., Morgan, F., *et al.*, 2012. Virtual cell: Computational tools for modeling in cell biology. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine* 4, 129–140.
- Rubinstein, B., Fournier, M.F., Jacobson, K., Verkhovsky, A.B., Mogilner, A., 2009. Actin-myosin viscoelastic flow in the keratocyte lamellipod. *Biophysical Journal* 97, 1853–1863.
- Schaff, J., Fink, C.C., Slepchenko, B., Carson, J.H., Loew, L.M., 1997. A general computational framework for modeling cellular structure and function. *Biophysical Journal* 73, 1135–1146.
- Schaus, T.E., Taylor, E.W., Borisy, G.G., 2007. Self-organization of actin filament orientation in the dendritic-nucleation/array-treadmilling model. *Proceedings of the National Academy of Sciences of the United States of America* 104, 7086–7091.
- Schmoller, K.M., Niedermayer, T., Zensen, C., Wurm, C., Bausch, A.R., 2011. Fragmentation is crucial for the steady-state dynamics of actin filaments. *Biophysical Journal* 101, 803–808.
- Sept, D., Xu, J., Pollard, T.D., McCammon, J.A., 1999. Annealing accounts for the length of actin filaments formed by spontaneous polymerization. *Biophysical Journal* 77, 2911–2919.
- Slepchenko, B.M., Loew, L.M., 2010. Use of virtual cell in studies of cellular dynamics. *International Review of Cell and Molecular Biology* 283, 1–56.
- Slepchenko, B.M., Schaff, J.C., Macara, I., Loew, L.M., 2003. Quantitative cell biology with the Virtual Cell. *Trends in Cell Biology* 13, 570–576.
- Svitkina, T.M., Borisy, G.G., 1999. Arp2/3 complex and actin depolymerizing factor/cofilin in dendritic organization and treadmilling of actin filament array in lamellipodia. *Journal of Cell Biology* 145, 1009–1026.
- Vavylonis, D., Kovar, D.R., O'Shaughnessy, B., Pollard, T.D., 2006. Model of formin-associated actin filament elongation. *Molecular Cell* 21, 455–466.
- Vavylonis, D., Yang, Q., O'Shaughnessy, B., 2005. Actin polymerization kinetics, cap structure, and fluctuations. *Proceedings of the National Academy of Sciences of the United States of America* 102, 8543–8548.
- Wegner, A., 1976. Head to tail polymerization of actin. *Journal of Molecular Biology* 108, 139–150.
- Wolgemuth, C.W., Stajic, J., Mogilner, A., 2011. Redundant mechanisms for stable cell locomotion revealed by minimal models. *Biophysical Journal* 101, 545–553.
- Wolgemuth, C.W., Zajac, M., 2010. The moving boundary node method: A level set-based, finite volume algorithm with applications to cell motility. *Journal of Computational Physics* 229, 7287–7308.

Relevant Websites

http://sbml.org/SBML_Software_Guide
Systems Biology Markup Language.
<http://vcell.org>
UConn Health Center.

Metabolism

HM Sauro, University of Washington, Seattle, WA, USA

© 2016 Elsevier Ltd. All rights reserved.

Introduction

Although one can trace the first metabolic experiments to Santorio Sanctorius (1561–1636) who investigated the balance of food intake and what he called insensible perspiration, actual studies into metabolism didn't start until Eduard Buchner in the late nineteenth century demonstrated fermentation in yeast cell-free extracts. It was hypothesized that the fermentation was carried out by a heat labile cytoplasmic substance which he called 'zymase.' Following this work, progress was rapid (Cornish-Bawden, 1997) with the discovery that glucose fermentation was carried out via a series of distinct reactions, each reaction catalyzed by a specific enzyme. Collectively the sequence of reactions was named glycolysis. The early twentieth century right up to the 1970s was a golden age for research in metabolism and enzymology. All the major metabolic pathways were elucidated together with a considerable amount of information on the mechanisms by which metabolism was regulated.

It has now been roughly 120 years since Buchner created the first fermenting cell-free extract and yet our understanding of metabolism is still incomplete. It is true that through modern genome-scale studies (Palsson, 2007; Terzer *et al.*, 2009; Thiele and Palsson, 2010; Henry *et al.*, 2010; Santos *et al.*, 2011), great numbers of organisms have had their metabolic potential mapped out in considerable detail. However, these maps are static and what is still missing to a large extent is understanding the dynamics and general operating principles of metabolism.

This article will be concerned with some of the operational principles of metabolism (Fell, 1997) and the theoretical principles that allow us to understand such a complex system (Heinrich and Schuster, 1996).

Metabolism

The first cellular networks to be discovered were the metabolic pathways such as Glycolysis in the 1930s and the Calvin cycle in the 1940s. These pathways were elucidated by a combination of enzymatic inhibitors and the use of radioisotopes such as Carbon-14. The Calvin cycle, for example, was discovered by following the fate of carbon when algae were exposed to ^{14}C -labeled CO_2 . With the development of microbial genetics, significant progress was also made in uncovering other pathways by studying mutants and complementing different mutants of a given pathway to determine the order of steps.

The reaction steps in a metabolic pathway are catalyzed by enzymes and we now know there are thousands of enzymes in a given organism catalyzing a great variety of pathways. The collective sum of all reaction pathways in a cell is referred to as metabolism and the small molecules that are interconverted are called metabolites.

Traditionally, metabolism is classified into two groups, anabolic (synthesis) and catabolic (breakdown). Coupling between the two metabolic groups is achieved through cofactors of which a great variety exist although two widely distributed cofactors include the pyridine nucleotides in the form of NAD and NADP and the adenine nucleotides in the form of ATP, ADP, and AMP. These cofactors couple redox and phosphate respectively by forming reactive intermediates that enables catabolism to drive anabolism. A primary catabolic process is cellular respiration where starting molecules such as glucose are oxidized in a stepwise fashion. The energy released is captured in the form of ATP and the oxidized product, water, and carbon dioxide are released as waste. The ATP can be used in turn to drive anabolic processes such as amino acid or lipid biosynthesis. In general, metabolic pathways are regulated at multiple time scales. For example, the rapid response to moment by moment fluctuations in supply and demand of molecular building blocks is most likely achieved through allosteric regulation. In contrast, large-scale remodeling of pathways via genetic regulation is generally slow and in response to significant changes in an organism's environment, for example, changes nutrient supply.

Figure 1 shows a section of the glycolytic pathway which converts glucose to pyruvate with the production of ATP and NADH. The diagram also shows the many negative and positive feedback and feedforward regulatory loops in glycolysis. Not all of these are present in all organisms, however, many are. Note the six regulatory signals that converge on 6-phosphofructose-1-kinase (also known as phosphofructokinase) and fructose biphosphatase (labeled 2 and 3). One of the aims of metabolic research is to quantify regulation and to understand the operational principles of such control. The literature on the dynamics of metabolic pathways is huge and includes many subtle concepts on how these systems operate. Given the size of the field, only a brief overview of the current state of knowledge can be given here. We will not, for example, consider the structural properties of metabolic networks (Clarke, 1980; Reich and Selkov, 1981; Hofmeyr *et al.*, 1986; Reder, 1988; Schuster and Schuster, 1993; Edwards *et al.*, 2002; Sauro and Ingalls, 2004; Palsson, 2007) nor the effect of spatial gradients or compartments. We start with some fundamental concepts, proceed to some general methodologies, and then describe the properties of some common metabolic motifs.

Fundamentals

Stoichiometric Networks

Almost all cellular processes involve some kind of chemical process such as binding, unbinding, or transformations, often in particular stoichiometric amounts. A network of reactions that is governed by mass-balance is termed a stoichiometric network. This is to distinguish the network from other

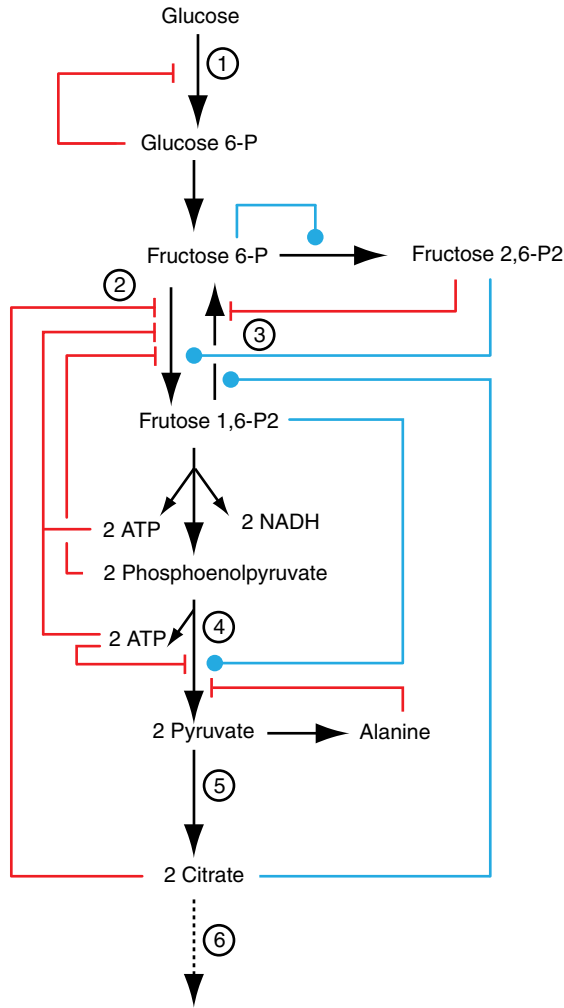


Figure 1 A section of glycolysis with negative and positive regulation shown. (1) Hexokinase; (2) 6-Phosphofructose-1-kinase; (3) Fructose biphosphatase; (4) Pyruvate kinase; (5) Entry to citric acid cycle; (6) To oxidative respiration.

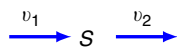


Figure 2 Simple two-step pathway.

graph-based networks that describe associativity or interactions between components but without any reference to stoichiometry. A key characteristic of stoichiometric networks is mass-balance. Consider a simple network comprising two reactions, v_1 and v_2 , with a common species, S . We will assume that the first reaction, v_1 , produces S and the second reaction, v_2 , consumes S (Figure 2).

According to the law of conservation of mass, any observed change in the amount of species, S , will be due to the difference between the inward rate, v_1 and outward rate, v_2 . That is, the change in concentration of S will be given by the difference in the two rates, leading to the differential equation:

$$\frac{dS}{dt} = v_1 - v_2 \quad [1]$$

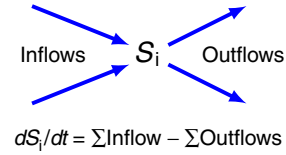


Figure 3 Mass-balance: The rate of change in species S_i is equal to the difference between the sum of the inflows and the sum of the outflows.

This equation is called a 'mass-balance equation' and for more complex systems such as the one shown in Figure 3 where there are multiple inflows and outflows, the mass-balance equation is given by:

$$\frac{dS_i}{dt} = \sum \text{Inflows} - \sum \text{Outflows} \quad [2]$$

The study of stoichiometry is a key component to understanding the dynamics of metabolism (Sauro, 2014).

Stoichiometry Matrix

When describing multiple reactions in a network, it is convenient to represent the stoichiometries in a compact form called the 'stoichiometry matrix.' Different disciplines use different symbols for the stoichiometry matrix. For example, the chemistry community has traditionally used ν (Aris, 1965), the constraint-based approach community (Palsson, 2007), S and the biochemical systems community, N . In this book the symbol ' N ' will be used to avoid any confusion when using ' S ' to represent a species.

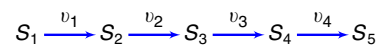
The stoichiometry matrix (Stephanopoulos *et al.*, 1998; Sauro, 2014) is a m row by n column matrix where m is the number of species and n the number of reactions:

$$N = m \times n \text{ matrix}$$

The columns of the stoichiometry matrix correspond to the individual chemical reactions in the network, the rows to the molecular species, one row per species. Thus the intersection of a row and column in the matrix indicates whether a certain species takes part in a particular reaction or not, and according to the sign of the element, whether there is a net loss or gain of substance, and by the magnitude, the relative quantity of substance that takes part in that reaction. In general the stoichiometry matrix has the form:

$$N = \begin{matrix} & \leftarrow v_j \rightarrow \\ \begin{matrix} \uparrow \\ S_i \\ \downarrow \end{matrix} & \begin{bmatrix} c_{ij} & \cdots & \cdots \\ \vdots & & \\ \vdots & & \end{bmatrix} \end{matrix}$$

where c_{ij} is the stoichiometry coefficient for the i th species and j th reaction. The stoichiometry matrix for the simple chain of reactions which has five molecular species and four reactions is shown below. The four reactions are labeled v_1 to v_4 .



The stoichiometry matrix for this simple system is given by:

$$N = \begin{array}{cccc|c} & v_1 & v_2 & v_3 & v_4 & \\ \hline & -1 & 0 & 0 & 0 & S_1 \\ & 1 & -1 & 0 & 0 & S_2 \\ & 0 & 1 & -1 & 0 & S_3 \\ & 0 & 0 & 1 & -1 & S_4 \\ & 0 & 0 & 0 & 1 & S_5 \end{array}$$

The rows and columns of the matrix have been labeled for convenience. Normally the labels are absent.

Reaction Rates

The stoichiometry matrix has no information concerning the rate of reaction, that is, how the rate of transformation is a function of the reactant and product species. Rates of reaction, often denoted by the symbol v , is measured with respect to a given molecular species normalized by the species' stoichiometric coefficient. This ensures that no matter which molecular species in a reaction is measured, the reaction rate is uniquely defined for that reaction. The kinetic rate laws that govern the rate of reaction are extremely varied, ranging from simple mass-action kinetic laws to complex allosteric and cooperative mechanisms. A discussion of these kinetic rate laws is beyond the scope of this article but readers can consult the texts by Sauro (2012) and Cornish-Bowden (2012). For a given pathway, the set of all reaction rates is conveniently represented by the reaction rate vector, v .

Dynamics

The time evolution of a metabolic pathway can be described using a set of differential equations. Equation [1] illustrates a single differential equation that describes the rate of change of a single metabolite S . In larger systems there will be as many differential equations as there are metabolites in the pathway. A compact way of depicting such a system is given by the equation:

$$\frac{ds}{dt} = Nv \quad [3]$$

where N is the stoichiometry matrix and v the vector of reaction rates. This system of equations can be solved numerically (Sauro, 2014) to produce the time evolution for all the metabolites in the pathway. Usually such systems evolve to a steady state where the metabolite levels are unchanging but the pathway sustains a finite flux from the pathway's source to its sink. At steady state the system equation has the form:

$$0 = Nv \quad [4]$$

Implicit in these equations are two assumptions. The first is that we are dealing with large numbers of molecules. This enables us to describe concentration changes using continuous variables. If molecule numbers fall into the 100 particle range, it is often recommended to model systems using a stochastic discrete approach (Gillespie, 2007). The second assumption we make is that we assume that the compartment in which the reactions take place is well mixed. This means there are no concentration gradients, that is, diffusion is faster than the enzymatic processes. Although these assumptions may seem restrictive it is nevertheless remarkable that many metabolic systems can be effectively modeled under these assumptions (Chassagnole *et al.*, 2001; van Eunen *et al.*, 2012; Smallbone *et al.*, 2013).

Almost all dynamical studies of metabolic pathways revolve around the study of eqns [3] and [4].

Enzyme Response

Since enzymes are the functional units of metabolism, it is important to understand how an enzyme responds to changes in its environment. A measure that describes this response is called the elasticity coefficient (Kacser and Burns, 1973; Sauro, 2012). Elasticities describe how sensitive a reaction rate is to changes in reactant, product, and effector concentrations. They represent the degree to which changes are transmitted from the

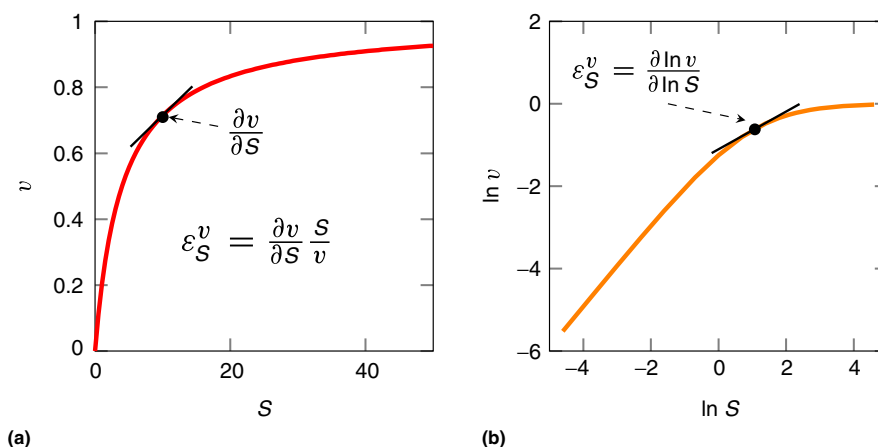


Figure 4 (a) The slope of the reaction rate versus the reactant concentration scaled by both the reactant concentration and reaction rate yields the elasticity, ϵ_S^v . (b) If the log of the reaction rate and log of the reactant concentration are plotted, the elasticity can be read directly from the slope of the curve. Curves are generated by assuming $v = S/(2 + S)$.

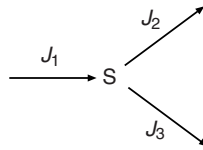


Figure 5 Simple branched pathway. This pathway has three different fluxes, J_1 , J_2 , and J_3 which at steady state are constrained by $J_1 = J_2 + J_3$.

immediate environment to the reaction rate (Figures 4 and 5). From a systems perspective they are critical components in understanding how a disturbance, such as the introduction of a drug applied at one or more points in a cellular pathway, propagates to the rest of the system. It is the magnitude and signs of the elasticities that determines how far and at what strength the disturbance travels. Elasticities are therefore central in helping us understand how networks and metabolism in particular functions.

Elasticity Coefficients

The elasticity coefficient is defined by the following expression:

$$\varepsilon_{S_i}^v = \left(\frac{\partial v}{\partial S_i} \frac{S_i}{v} \right)_{S_j, S_k, \dots} = \frac{\partial \ln v}{\partial \ln S_i} \approx v/S_i \quad [5]$$

Operational Definition

The elasticity is the fractional change in reaction rate in response to a fractional change in a given reactant or product while keeping all other reactants and products constant.

Since the elasticity is expressed in terms of fractional changes, it is also possible to get an approximate value for the elasticity by considering percentage changes. For example, if we increase the substrate concentration of a particular reaction by 2% and the reaction rate increases by 1.5%, then the elasticity is given by $1.5/2 = 0.75$. The elasticity is however only strictly defined (see eqn [5]) for infinitesimal changes and not finite percentage changes. As long as the changes are small, the finite approximation is a good estimate for the true elasticity.

Pathway Responses

The response of a pathway to a perturbation can be described in two ways: (1) large-scale transients and remodeling of pathways that occur as a result of major changes to an organism state, for example, as a result of changes to nutrient supplies; (2) changes as a result of the moment by moment adjustments of a pathway due to small perturbations in the milieu of the organism. Unfortunately, there is little or no theory that allows us to understand large-scale transients, this is an area of active research and recent experimental results suggests there is a rich vein of untapped biology awaiting to be discovered (Zampar *et al.*, 2013; Klumpp and Hwa, 2014). Here we will focus on how pathways respond to moment by moment perturbations, this is a much more developed area of research and has revealed considerable insight into the dynamical operation of pathways.

The vocabulary used to describe metabolic regulation has on the whole not been well defined. Thus terms such as control and regulation are used in many different places with different meanings. In the vernacular, the word control usually means the ability to influence, command, or to restrain a situation or process (In engineering, control theory refers to the body of knowledge concerned with the design and study of systems that can perform specific tasks or achieve a particular objective.). Here the term control will be used, as set forth by Kacser and Burns (1973) to describe how much influence a given reaction step in a network has on the system. To make matters simpler we will consider the system to be at steady state so that control will refer to how much influence a given reaction step has on the steady state, that is how steady state fluxes and concentrations are influenced. General reviews on 'metabolic control analysis' can be found in the following sources (Fell, 1997; Qian and Beard, 2006). Non-steady state analysis is still not fully developed but readers can consult papers by Acerenza *et al.* (1989), Heinrich and Reder (1991), Reijenga *et al.* (2002), Ingalls and Sauro (2003), and more general information in the textbook by Heinrich and Schuster (1996).

Most reaction steps in a cell are controlled by proteins and the question then becomes how much influence a given protein has on the system's steady state. Experimentally such control can be measured by changing the concentration of an enzyme or changing its activity via an inhibitor and measuring the effect on the steady state flux and species concentrations. Concentrations of proteins can be changed in various ways, for example, by using irreversible inhibitors, changing the promoter consensus sequence on the gene that codes for the protein or by employing antisense RNA to reduce the expression level.

Control: The degree of control that a particular reaction step has on a flux or species concentration is called the 'control coefficient.'

In addition to investigating how individual reaction steps control the fluxes and concentrations in a network, we are also interested in how external factors influence the network. Examples of external factors include the level of nutrients, hormones, and of particular interest to human health, therapeutic drugs. In these situations, rather than use the word control, we will use the term response. Thus a biological cell will have a response to a particular infusion of a drug. The degree of influence an external factor has on a biological system is described using 'response coefficients' (Kacser and Burns, 1973).

Finally, the degree to which an external factor or an enzyme has on a cellular network will depend on the network's regulatory mechanisms. The degree of regulation at a given reaction step is described using the 'regulatory coefficient' (Sauro, 1990; Hofmeyr and Cornish-Bowden, 1991; Brand, 1997; Brand and Curtis, 2002).

Control Coefficients

Control coefficients are used to describe how much influence (i.e., control) a given reaction step has on the steady state flux or species concentration level. It is common to measure this influence by changing the concentration of the enzyme that

catalyzes the reaction. We define the control coefficients as follows:

$$C_{E_i}^J = \frac{dJ}{dE_i} \frac{E_i}{J} = \frac{d \ln J}{d \ln E_i} \approx J/E_i \quad [6]$$

$$C_{E_i}^{S_j} = \frac{dS_j}{dE_i} \frac{E_i}{S_j} = \frac{d \ln S_j}{d \ln E_i} \approx S_j/E_i \quad [7]$$

The flux control coefficient measures the fractional change in flux brought about by a given fractional change in enzyme concentration and the concentration control coefficients measure the fractional change in species concentration given a fractional change in enzyme concentration.

For example, we could increase the copy number of a given gene to change the concentration of the expressed protein and find out how much influence the gene has on its own gene product and on other processes. It is important to note however, that knowing the values of the control coefficients does not tell us why certain enzymes or proteins have more influence than others. To answer the 'why' question we must examine what determines the relationship between control coefficients and how they are related to the elasticities of the network. The first set of relationships are called the summation theorems (Kacser and Burns, 1973; Ingalls, 2013):

$$\sum_{i=1}^n C_{E_i}^J = 1 \quad \sum_{i=1}^n C_{E_i}^{S_j} = 0 \quad [8]$$

These relationships imply the following assertions about control in a pathway:

1. Control is shared throughout a pathway, that is the degree to which flux is limited in a pathway is shared. It is unlikely that only a single step in a pathway limits flux.
2. If one step gains control, one or more other steps must loose control.
3. Control coefficients are system properties, they can only be computed or measured in the intact system. Inspection of a single enzymatic step will not reveal its degree of control (or influence) on the pathway.

Rate-Limiting Steps

In much of the literature and many contemporary textbooks, one will often find a brief discussion of an idea called the rate-limiting step. In textbooks we will find statements such as:

In order to exert control on the flux of metabolites through a metabolic pathway, it is necessary to regulate its rate-limiting step. (Voet and Voet, 2004, p. 522)

These statements are not consistent with the idea that control is shared among all steps. Other than a few books such as the 4th edition of Lehninger Biochemistry, most textbooks will have similarly expressed statements. A common argument is that a metabolic pathway will contain a single step that has overall influence over the pathways' flux. This concept drives much of metabolic engineering and drug targeting of metabolism. What is most surprising however is that a simple quantitative analysis shows that it cannot be true, and there is ample experimental evidence (Heinisch, 1986; Burrell *et al.*, 1994) to support the notion of shared control.

The concept of the rate-limiting step is both inconsistent with logic and more importantly experimental evidence.

The confusion over the existence of rate-limiting steps stems from a failure to realize that rates in cellular networks are governed by the law of mass-action, that is, if a concentration changes, then so does its rate of reaction. Analogies that are sometimes used such as traffic flow or supermarket checkout do not apply because in both cases flow is not governed by either the number of cars or customers waiting. In terms of control coefficients, a rate-limiting step can be interpreted as a flux control coefficient of unity. This means, by the summation theorem, that all other steps (at least in a linear chain) must have flux control coefficients of zero. However, when we consider branched and cyclic systems it is possible to have flux control coefficient much greater than one (other control coefficients must then be negative to satisfy the summation theorem). In these cases what adjective should we use, hyper-rate-limiting steps? In the final analysis, it is better to assign a value to the rate limitingness of a particular step.

For example, a rate-limiting step in glycolysis is often considered to be phosphofructokinase since it is regulated by many effectors. Such an assertion makes perfect sense if metabolism is seen as a series of connected pipes and tanks with valves that can be used to independently turn on and off the flow of water. However, metabolism is not like this, the valves are part of the system. This makes an understanding of how pathways are controlled much more subtle (Morandini, 2009). Measurements (Heinisch, 1986; Ruijter *et al.*, 1997) have shown, for example, that phosphofructokinase is in fact not rate limiting even though it is heavily regulated.

Connectivity Theorems

In the previous section, we saw that there are summation theorems relating the control coefficients. Here we will introduce an additional set of theorems that relate the control coefficients to the substrate, product, and effector elasticities. These theorems are called the 'connectivity theorems' and represent the most important result in metabolic control analysis (Kacser and Burns, 1973; Wanders *et al.*, 1984; Westerhoff *et al.*, 1984; Kholodenko *et al.*, 1994; Ingalls, 2013). The theorem relates the control coefficients to the elasticities, that is, they relate system-wide properties to local properties. They help us explain why some enzymes have the degree of influence they have.

For a species, S , that interacts with r other steps, the connectivity theorems are written as:

Flux Connectivity Theorem with respect to a common metabolite, S_k where r is the number of interactions it makes with neighboring reaction steps.

$$\sum_{i=1}^r C_{E_i}^J \epsilon_{S_k}^{v_i} = 0 \quad [9]$$

Concentration Connectivity Theorem with respect to the common metabolite, S_k , where r is the number of interactions it makes with neighboring reaction steps.

$$\sum_{i=1}^r C_{E_i}^{S_k} \epsilon_{S_k}^{v_i} = -1 \quad [10]$$

Concentration Connectivity Theorem with respect to the common metabolite, S_k , and a distant metabolite, S_m , where r is the number of interactions it makes with neighboring reaction steps.

$$\sum_{i=1}^r C_{E_i}^{S_m} \epsilon_{S_k}^{E_i} = 0$$

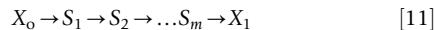
By combining the connectivity theorems with the summation theorems we can get a complete picture of how small perturbations propagate through a system (Fell and Sauro, 1985).

Metabolic Motifs

There are a number of common pathway motifs found in metabolism. Here we will describe the dynamical properties of some of the more common ones by analyzing the effects of small perturbations on the system. For a detail discussion, including proofs, the reader should consult the textbook by Heinrich and Schuster (1996).

Linear Pathways

Linear pathways, that is a sequence of consecutive reactions, are surprisingly uncommon in metabolism. Most of the metabolism is either heavily branched or coupled to other processes via cofactors such as ATP and NADH. Although pure linear pathways are rare, their study provides a background to gauging other more complex motifs (Kacser, 1983; Heinrich and Rapoport, 1974; Savageau, 1976). Consider the linear pathway:



which has m floating species and n reactions ($n=m+1$). X_0 and X_1 are fixed species representing the source and sink pools respectively. If we assume reversible mass-action kinetics for each of the reactions it is possible to show that the control coefficients are less than one but greater than zero, $0 \leq C_i^j \leq 1$.

If we assume that each equilibrium constant is greater than one and equal to each other it is possible to show that two adjacent steps, for example the i th and $i+1$ th step, are related by $C_i^j > C_{i+1}^j$, that is 'earlier steps' will have more flux control. This pattern applies across the entire pathway such that steps near the beginning of a pathway will have more control than steps near the end. We will call this effect 'front loading' and gives some credence to the traditional idea that the first or committed step is the most important step in a pathway. However, front loading only applies to unregulated pathways, the moment we add regulation to the pathway this picture changes.

Another general result to come from an analysis of linear pathways is steps close to equilibrium tend to have less flux control. This result that has been known for sometime, even before the advent of a systematic analysis. However, what a more systematic analysis shows is that it is the context within which the near equilibrium reaction resides that is important in determining its flux control coefficient (Fell and Sauro, 1986). This means that identifying a step that is close to equilibrium is not a sufficient criterion for asserting it has a low flux control.

Irreversibility: In a linear pathway governed by 'linear kinetics' and without the presence of regulatory interactions, all steps downstream of an irreversible step, ($\rho_i=0$), have no flux control.

Branched Pathways

Branching structures are probably one of the most common patterns in biochemical networks. Even a pathway such as glycolysis, often depicted as a straight chain in textbooks is in fact a highly branched pathway.

At any given branch node, where a node is a molecular species, there will be conservation of mass. Given a node species, x_i , with b branches entering the node and d branches leaving, the net rate of change in concentration of x_i is:

$$\sum_{i=1}^b v_i - \sum_{j=1}^d v_j = \frac{dx_i}{dt}$$

At steady state when $dx_i/dt=0$, it must also be true that:

$$\sum_{i=1}^b v_i = \sum_{j=1}^d v_j$$

In this section we will describe the control of flux through a branched system in response to changes in enzyme activity. Let us consider the simple branched pathway depicted in Figure 5.

In the figure, J_1 , J_2 and J_3 are the steady state fluxes. By the law of conservation of mass, at steady state, the fluxes in each limb will be governed by the relationship:

$$J_1 = J_2 + J_3$$

Given three different fluxes and one intermediate, there will be four sets of control coefficients, one set concerned with changes in the intermediate, S , and three sets corresponding to each of the three fluxes (Kacser, 1983; LaPorte *et al.*, 1984; Fell and Sauro, 1985).

The key observation in branched pathways is the possibility to encounter flux control coefficients in excess of one, that is a 'super rate-limiting' step (Zhang *et al.*, 2013). The effect occurs when a large flux influences a smaller one. Figure 6 illustrates two extremes. The pathway on the right panel (b) shows two large fluxes, J_1 and J_2 having a major impact on the small flux, J_3 . A note to engineers, the sensitivity of the J_2 flux with respect to itself in panel (b) is roughly one. Even more important is that the concentration control coefficient of the intermediate, S , with respect to step 2 is close to zero. In this situation, increasing the flux through J_2 by overexpressing the enzymes along step 2 will lead to proportional flux increases without significantly affecting the flux through J_1 and J_3 .

Futile Cycles

Closely related to branched systems are cyclic pathways. A typical cyclic pathway is shown in Figure 7. For cycling to occur both forward and back reactions must operate. It is typical to find that the forward and reverse reactions are chemically distinct. Often one reaction will be driven by ATP while the other by the hydrolysis of phosphate groups. Typical examples in metabolism include the cycle between glucose

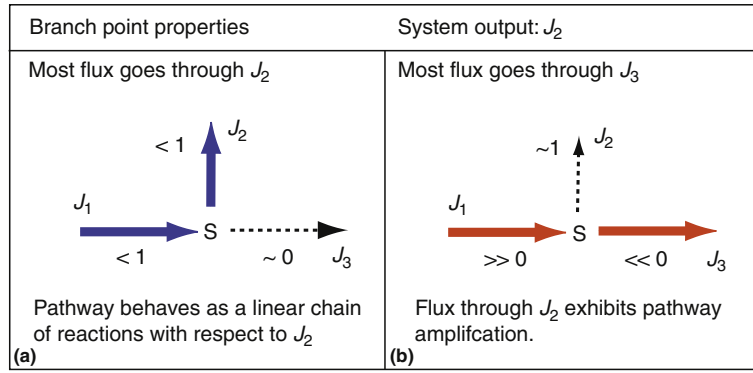


Figure 6 The figure shows two flux extremes relative to the flux through branch J_2 . In case (a), where most of the flux goes through J_2 , the branch reverts functionally to a simple linear sequence of reactions made from J_1 and J_2 . In case (b), where most of the flux goes through J_3 , the flux through J_2 now becomes very sensitive to changes in activity at J_1 and J_3 . Given the right kinetic settings, the flux control coefficients can become very large with values greater than one or less than minus one for activity changes at J_3 . The values next to each reaction indicate the flux control coefficient for the flux through J_2 with respect to activity at the reaction.

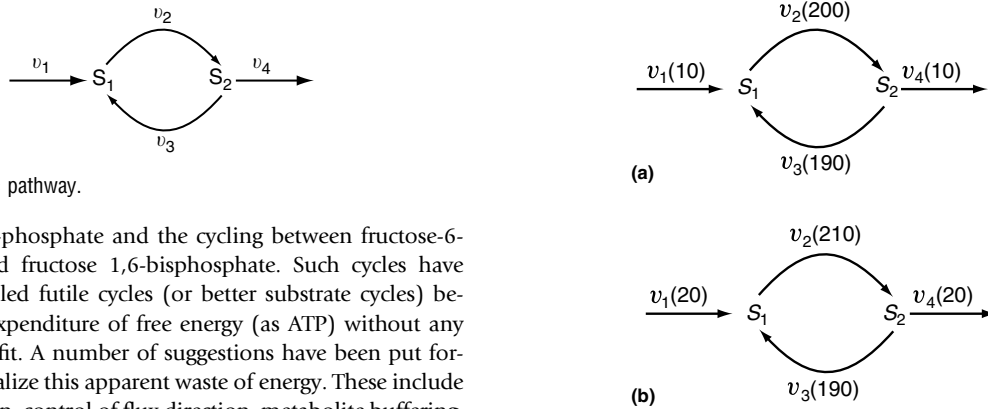


Figure 7 Cyclic pathway.

and glucose-6-phosphate and the cycling between fructose-6-phosphate and fructose 1,6-bisphosphate. Such cycles have often been called futile cycles (or better substrate cycles) because of the expenditure of free energy (as ATP) without any apparent benefit. A number of suggestions have been put forward to rationalize this apparent waste of energy. These include heat production, control of flux direction, metabolite buffering, and more sensitive control of the net flux through the pathway. We will only consider the latter here. **Figure 7** shows a typical cyclic pathway embedded in a linear chain. Of interest is the sensitivity of the pathway flux, v_1 or v_4 to changes in v_2 . The simplest assumption to make is that when we change v_2 there is no change in back flux, v_3 . This could be for a number of reasons, for example, v_3 is saturated by its substrate S_2 .

Figure 8 illustrates two situations, a reference state in panel (a) and a perturbation of 5% to v_2 shown in panel (b). Assuming that the entire flux changes appear in output flux v_4 and that v_3 is not changed, then the percentage change in v_4 (or v_1) is 100%, a 20-fold amplification.

This effect can be easily quantified as follows. First we note the flux constraint due to the cycle is:

$$v_1 = v_2 - v_3$$

We then assume that a perturbation in v_2 leads to the same change in v_1 , that is:

$$\delta v_2 = \delta v_1$$

We can now compute the fractional changes in v_1 and v_2 as:

$$\frac{\delta v_1}{v_1} = \frac{\delta v_2}{v_2} \frac{v_2}{v_1}$$

Figure 8 Amplification in a substrate cycle. (a) Reference state, values refer to fluxes at various points, note that $v_1 = v_2 - v_3$. (b) Activation of v_2 by 5% leads to a 100% change in v_1 and v_4 . It assumes that v_3 is not activated by any changes in S_2 .

The degree of amplification is then given by

$$\frac{\delta v_1 / v_1}{\delta v_2 / v_2} = \frac{v_2}{v_1}$$

Since $v_2 = v_1 + v_3$ then

$$\frac{\delta v_1 / v_1}{\delta v_2 / v_2} = \frac{v_1 + v_3}{v_1} = 1 + \frac{v_3}{v_1} \quad [12]$$

This result shows that the higher the cycling rate (v_3) compared to the through flux, the greater the amplification. This equation gives us the maximum degree of amplification possible. In practice, v_3 will not remain unchanged because S_2 rises. In addition, S_1 will fall due to high consumption which will reduce v_2 but increase v_1 due to lower product inhibition. The resulting amplification is therefore a more complicated function than the one suggested by eqn [12]. However, eqn [12] gives the maximum possible amplification.

Negative Feedback

Feedback is widespread in biochemical networks and physiological systems in general. Some form of feedback permeates almost every known biological process. Metabolism is no exception where negative feedback loops can be virtually found in all pathways. Given the ubiquity, negative feedback must serve an important function. On the face of it, feedback is a simple process that involves sending a portion of the output to the input. If the portion sent back reduces the input then the feedback is called negative feedback otherwise it is called positive feedback.

The central mechanism in a feedback circuit is the generation of the error signal, that is the difference between the desired output (set point) and the actual output. The error is fed into a controller (often something that simply amplifies the error) which is used to increase or decrease the process. For example, if a disturbance on the process block reduces the output, then the feedback operates by generating a positive error, this in turn increases the process and restores the original drop in the output.

The figure of the generic negative feedback circuit (Figure 9) is highly stylized which makes it difficult to identify the various components in a real biological system. In addition, biological systems are invariably more complex with multiple nested feedback loops and multiple inputs and outputs. If it is remarkable then even after 50 or 60 years of research, the role of many of the feedback systems in biochemical networks is still highly speculative.

It was Umbarger (1956) and Yates and Pardee (1956) who discovered feedback inhibition in the isoleucine biosynthesis pathway and the inhibition of aspartate transcarbamylase in *Escherichia coli*. It was not long afterwards that some researchers began to investigate such feedback systems mathematically. Probably the most extensive mathematical analysis of biochemical feedback was conducted by Savageau (1972, 1974, 1976), Burns (1971), Kacser and Burns (1973), Othmer (1976), Tyson and Othmer (1978) in the 1970s, and Dibrov *et al.* (1982) in the early 1980s. More recently, Cinquin and Demongeot (2002) have published an interesting review on the roles of feedback in biological systems.

The first question is how can we map the generic diagram (Figure 9) onto a biochemical feedback circuit (Figure 10), on the surface they look similar but the initial impression is misleading.

In order to be clear we need to identify the input (set point), output, the feedback loop, and the process block in the biochemical network (Figure 11).

In naturally evolved systems it is sometimes difficult to identify the various parts in a negative feedback circuit. The output in the block diagram corresponds to the concentration of S_3 . The negative feedback component corresponds to the interaction of S_3 with the allosteric enzyme in the first step. The set point is more problematic but it is most likely embedded in the kinetic characteristics of the allosteric enzyme. Finally, the controller is represented by the steps v_1 , v_2 , and v_3 . The load on the system is represented by the last step, v_4 and other disturbances can be assigned to v_1 , v_2 , v_4 and the input concentration, X_0 .

Graphical Understanding of Feedback

It is possible to appreciate the effect of negative feedback using a graphical approach. Figure 12 shows two plots, one with strong and the other with weak feedback. The plots show the reaction rates v_1 and v_2 as a function of the intermediate species, S , for a simple two-step pathway where S can negatively feedback on to the first step. We assume that the second step follows first-order kinetics so that the v_2 is a straight line. The feedback response curve shows a decline from high to low as S increases. For strong feedback, the decline is steep (left plot). If we now change the rate constant for the second step, this changes the slope of v_2 , this is equivalent to a perturbation in the system. In the case of weak feedback, changes in v_2 result in significant changes to S , this is because the feedback response is shallow. In contrast, when we have strong feedback (left panel), where the slope is very steep, any changes in v_2 results in only small changes in S . In this way we can see how strong negative feedback can buffer changes in v_2 .

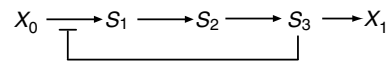


Figure 10 Simple four-step pathway with negative feedback.

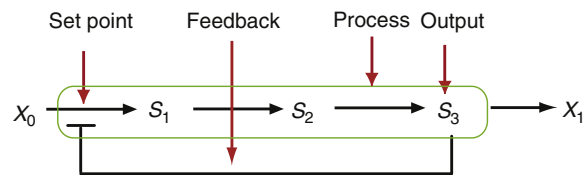


Figure 11 Simple four-step pathway with negative feedback.

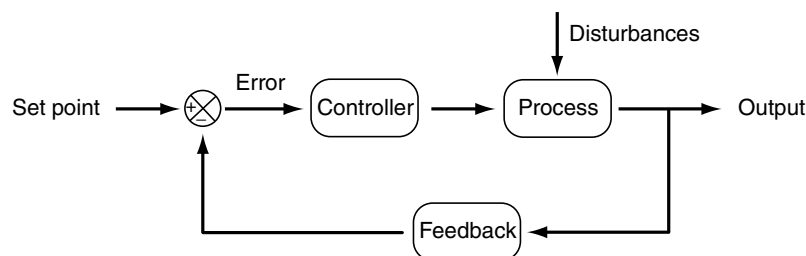


Figure 9 Generic structure of a negative feedback system.

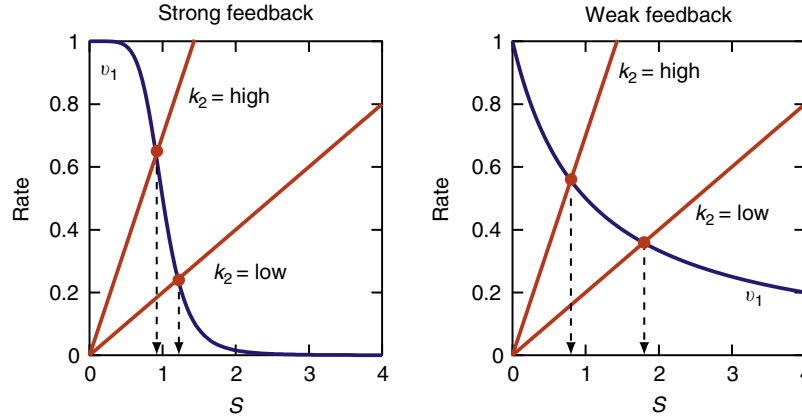


Figure 12 Plot of v_1 and v_2 versus the concentration of S for a simple two-step pathway with negative feedback. Two perturbations in k_2 that determines v_2 are shown. In the left panel where the feedback is strong, changes in k_2 have hardly any effect on S . On the right panel, the same change in k_2 results in a much larger change in S . This illustrates the homeostatic property of negative feedback. Left panel: $v_1 = 1/(1 + S^4)$, Right panel: $v_1 = 1/(1 + S)$.

Control Analysis

We can derive the flux and concentration control coefficients in terms of the elasticities (Sauro, 2009). Note the presence of the feedback term: e_2^1 in the matrix.

With feedback	Without feedback
$C_{E_1}^J = \frac{e_1^2 e_2^3}{e_1^2 e_2^3 - e_1^1 e_2^3 + e_1^1 e_2^2 - e_1^2 e_2^1}$	$C_{E_1}^J = \frac{e_1^2 e_2^3}{e_1^2 e_2^3 - e_1^1 e_2^3 + e_1^1 e_2^2}$
$C_{E_2}^J = \frac{-e_1^1 e_2^3}{e_1^2 e_2^3 - e_1^1 e_2^3 + e_1^1 e_2^2 - e_1^2 e_2^1}$	$C_{E_2}^J = \frac{-e_1^1 e_2^3}{e_1^2 e_2^3 - e_1^1 e_2^3 + e_1^1 e_2^2}$
$C_{E_3}^J = \frac{e_1^1 e_2^2 - e_1^2 e_2^1}{e_1^2 e_2^3 - e_1^1 e_2^3 + e_1^1 e_2^2 - e_1^2 e_2^1}$	$C_{E_3}^J = \frac{e_1^1 e_2^2}{e_1^2 e_2^3 - e_1^1 e_2^3 + e_1^1 e_2^2}$

We can see that the addition of feedback adds a new term to the denominator and to the numerator for $C_{E_3}^J$. If we make the feedback elasticity, e_2^1 , larger we can see that the demand flux control coefficient tends to unity. That is, all control moves out of the feedback loop. In terms of the steam engine analogy, it is equivalent to being able to change the demand on the steam engine without loss of power. This can be seen more clearly if we look at the concentration control coefficients.

$$C_{E_3}^J \rightarrow 1$$

$$C_{E_3}^{S_2} \rightarrow 0$$

In the list below we give the three control coefficients with respect to S_2 . Looking at $C_{E_3}^{S_2}$ we see that as the feedback strength is increased (e_2^1) in magnitude, the control coefficient tends to zero. This means that the feedback locks the concentration of S_2 into a very narrow range in response to changes in demand.

$$C_{E_1}^{S_2} = \frac{e_1^2}{e_1^2 e_2^3 - e_1^1 e_2^3 + e_1^1 e_2^2 - e_1^2 e_2^1}$$

$$C_{E_2}^{S_2} = \frac{e_1^1}{e_1^2 e_2^3 - e_1^1 e_2^3 + e_1^1 e_2^2 - e_1^2 e_2^1}$$

$$C_{E_3}^{S_2} = \frac{e_1^1 - e_1^2}{e_1^2 e_2^3 - e_1^1 e_2^3 + e_1^1 e_2^2 - e_1^2 e_2^1}$$

Classically, allosteric enzymes have been considered as flux controllers. The analysis here suggests the opposite picture. Allosteric enzymes, when part of a negative feedback loop are very poor controllers of flux. Instead allow distal steps to be good controllers of a pathway flux. For demand-driven systems this is a logical arrangement.

The nagging suspicion remain however that intuitively we feel that the regulated step must have some kind of ability to enable the system to operate in the way it does. There are at least two answers to this. The first is that the feedback elasticity will be strongly negative. For example, if the regulated step were determined by a modified Hill-like equation such as:

$$v = \frac{V_{\max} X_o}{S^n + X_o + K_m}$$

where S is the feedback signal, then the elasticity of the reaction rate with respect to the signal is given by:

$$e_S^v = \frac{n X_o K_m}{K_m + (S/K_m)^n}$$

We see that at low signal, S , the elasticity is proportional to n , the Hill coefficient. The reaction is therefore very sensitive to changes in the signal molecule.

The second way to answer the question is to consider the pathway with and without the negative feedback loop. We can compare for example the flux control coefficient on the first step with and without negative feedback. When we remove the negative feedback, the pathway will change to a new state where the concentrations of S_1 and S_2 are higher. To make the comparison fair we should adjust the level of enzyme in the first step to restore the levels of S_1 and S_2 to the values they had before the feedback was removed. When we do this we will also automatically restore the pathway flux to its original value. In practice this means that all the elasticities except for the feedback elasticity are exactly the same in both systems. To keep the analysis simple, let us assume that the product inhibition elasticities, e_1^1 and e_2^2 are both zero. We now take the ratio of the flux control coefficient on the first step, $C_{E_1}^J$, without feedback to the same coefficient with feedback, which will be

denoted by ${}^nC_1^I$:

$$\frac{C_1^I}{{}^nC_1^I} = \frac{\varepsilon_1^2 \varepsilon_2^3 - \varepsilon_2^1 \varepsilon_1^2}{\varepsilon_1^2 \varepsilon_2^3} = 1 - \frac{\varepsilon_2^1 \varepsilon_1^2}{\varepsilon_1^2 \varepsilon_2^3} = 1 - \frac{\varepsilon_2^1}{\varepsilon_2^3}$$

Given that ε_2^1 is negative, the term $-\varepsilon_2^1/\varepsilon_2^3$ is positive. That is:

$$C_1^I > {}^nC_1^I$$

This tells us that for a negative feedback loop to be effective, the flux control coefficient in the unregulated pathway must be higher than the flux control coefficient with feedback. In a sense the first step in the pathway must have some degree of rate limitingness if a negative feedback is to be effective.

Robustness and Supply/Demand

One way to look at metabolic systems is to divide them into two separate but connected blocks, the supply and demand block (Hofmeyr and Cornish-Bowden, 2000; Rohwer and Hofmeyr, 2008; Hofmeyr and Rohwer, 2011). Negative feedback makes this division very straightforward. Consider a factory that makes cars. Should the rate of car production be controlled by the demand or the supply of cars? Economically the most efficient strategy is to let demand decide how many cars to make, this ensure that excess cars do not build up and thereby waste resources. In certain metabolism situations the same reasoning can be used. For example, the production of amino acids would best be determined by the demand from protein synthesis. This means that if protein synthesis slows, amino acid production should also slow. A metabolic system that supplies amino acids should be able to supply amino acids effectively at both high and low demand. One way to do this is to maintain the amino acid level at a relatively constant level independent of the demand block. Let us assume for the moment that there is no feedback regulation in the supply/demand pathway (Figure 13). If demand rises, this will result in the intermediate metabolite, P , falling. As P falls, the ability to supply the increased demand becomes more and more difficult. If demand falls, the flux through the pathway will fall. This will cause the intermediate metabolite, P , to rise. Since the equilibrium constant across the supply block is likely to be large, the concentration of P could raise to toxic high levels as the supply block approach equilibrium at low fluxes.

The solution to avoid both problems at high and low demand is to use negative feedback (Figure 14). With negative feedback, high demand will result in a decrease in the intermediate metabolite, P , which in turn will release repression in the supply block to restore some loss in P . Alternatively, at low

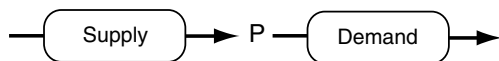


Figure 13 A system divided into supply and demand blocks.

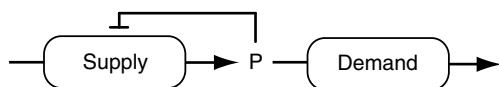


Figure 14 A system divided into supply and demand blocks with negative feedback.

demand, the increase in P will suppress its own production preventing excessive production of P .

Instability

Although negative feedback offers considerable advantages to a system, too much feedback and delays can result in instability in the form of sustained oscillations. There are a number of examples of metabolic systems or metabolism systems interacting with protein networks where negative feedback appears to cause instability and the emergence of sustained oscillations (Richard *et al.*, 1996; Fung *et al.*, 2005; Stricker *et al.*, 2008).

Conclusion

Although the metabolic parts list is almost complete, our understanding of how metabolism operates is still somewhat primitive. Not only is our understanding of the role of the many feedback and feedforward loops incomplete we do not have a clear understanding of the overall economic decisions that cells make to use and allocate metabolic resources. However, recent ground-breaking and exciting work by the Hwa group has started to collect data and models that suggests the beginnings of an economic model of cellular growth (You *et al.*, 2013; Klumpp *et al.*, 2013; Klumpp and Hwa, 2014). In this work, ribosome synthesis and availability takes a central role in determining metabolic dynamics. In this article only a brief and limited review of some of the essential concepts in understanding metabolic dynamics has been given. Important topics such as long-term regulation (Zampar *et al.*, 2013) and metabolic channeling have not been discussed (Srere and Ovadi, 1990; Sauro and Kacser, 1990; Welch and Easterby, 1994; Huang *et al.*, 2001; Dueber *et al.*, 2009; Castellana *et al.*, 2014).

Acknowledgments

Part of this work was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award numbers R01-GM081070. The content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health. The author also acknowledges NSF support from DBI:1355909 and MCB:1158573. Some of the figures and text are reused from the author's own teaching material.

See also: Molecular Principles, Components, Technology, and Concepts: Metabolism: A Structure Perspective on Organelle Bioenergetics; Metabolic Regulation

References

Acerenza, L., Sauro, H.M., Kacser, H., 1989. Control analysis of time-dependent metabolic systems. *Journal of Theoretical Biology* 137, 423–444.

- Aris, R., 1965. Prolegomena to the rational analysis of systems of chemical reactions. *Archive for Rational Mechanics and Analysis* 19, 81–99.
- Brand, M.D., 1997. Regulation analysis of energy metabolism. *Journal of Experimental Biology* 200 (Pt 2), 193–202.
- Brand, M.D., Curtis, R.K., 2002. Simplifying metabolic complexity. *Biochemical Society Transactions* 30 (2), 25–30.
- Burns, J.A., 1971. Studies on complex enzyme systems. PhD Thesis, University of Edinburgh. Available at: <http://www.sys-bio.org/jim-burns-thesis/> (accessed 30.01.15).
- Burrell, M.M., Mooney, P.J., Blundy, M., *et al.*, 1994. Genetic manipulation of 6-phosphofructokinase in potato tubers. *Planta* 194, 95–101.
- Castellana, M., Wilson, M.Z., Xu, Y., *et al.*, 2014. Enzyme clustering accelerates processing of intermediates through metabolic channeling. *Nature Biotechnology* 32 (10), 1011–1018.
- Chassagnole, C., Rais, B., Quentin, E., Fell, D.A., Mazat, J.-P., 2001. An integrated study of threonine-pathway enzyme kinetics in *Escherichia coli*. *Biochemical Journal* 356, 415–423.
- Cinquin, O., Demongeot, J., 2002. Roles of positive and negative feedback in biological systems. *Comptes Rendus Biologies* 325, 1085–1095.
- Clarke, B.L., 1980. Stability of complex reaction networks. Volume 42 of *Advances in Chemical Physics*. New York, NY: Wiley.
- Cornish-Bowden, A., 1997. *New Beer in an Old Bottle*. Eduard Buchner and the Growth of Biochemical Knowledge. Spain: Universitat de València.
- Cornish-Bowden, A., 2012. *Fundamentals of Enzyme Kinetics*. Weinheim: Wiley-Blackwell.
- Dibrov, B.F., Zhabotinsky, A.M., Kholodenko, B.N., 1982. Dynamic stability of steady states and static stabilization in unbranched metabolic pathways. *Journal of Mathematical Biology* 15, 51–63.
- Dueber, J.E., Wu, G.C., Malmirchegini, G.R., *et al.*, 2009. Synthetic protein scaffolds provide modular control over metabolic flux. *Nature Biotechnology* 27 (8), 753–759.
- Edwards, J.S., Covert, M., Palsson, B., 2002. Metabolic modelling of microbes: The flux-balance approach. *Environmental Microbiology* 4 (3), 133–140.
- van Eunen, K., Kiewiet, José AL, Westerhoff, Hans V, Bakker, Barbara M, 2012. Testing biochemistry revisited: How in vivo metabolism can be understood from in vitro enzyme kinetics. *PLoS Computational Biology* 8 (4), e1002483.
- Fell, D., 1997. *Understanding the Control of Metabolism*. London: Portland Press.
- Fell, D.A., Sauro, H.M., 1985. Metabolic control analysis: Additional relationships between elasticities and control coefficients. *European Journal of Biochemistry* 148, 555–561.
- Fell, D.A., Sauro, H.M., 1986. Non-equilibrium/equilibrium reactions: Which controls? *Biochemical Society Transactions* 14, 624–625.
- Fung, E., Wong, W.W., Suen, J.K., *et al.*, 2005. A synthetic gene-metabolic oscillator. *Nature* 435 (7038), 118–122.
- Gillespie, D.T., 2007. Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry* 58, 35–55.
- Heinisch, J., 1986. Isolation and characterisation of the two structural genes coding for phosphofructokinase in yeast. *Molecular and General Genetics* 202, 75–82.
- Heinrich, R., Rapoport, T.A., 1974. A linear steady state treatment of enzymatic chains. Critique of the crossover theorem and a general procedure to identify interaction sites with an effector. *European Journal of Biochemistry* 42, 97–105.
- Heinrich, R., Reder, C., 1991. Metabolic control analysis of relaxation processes. *Journal of Theoretical Biology* 151 (3), 343–350.
- Heinrich, R., Schuster, S., 1996. *The Regulation of Cellular Systems*. London: Chapman and Hall.
- Henry, C.S., DeJongh, M., Best, A.A., *et al.*, 2010. High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature Biotechnology* 28 (9), 977–982.
- Hofmeyr, J.-H., Cornish-Bowden, A., 1991. Quantitative assessment of regulation in metabolic systems. *European Journal of Biochemistry* 200 (1), 223–236.
- Hofmeyr, J.-H., Cornish-Bowden, A., 2000. Regulating the cellular economy of supply and demand. *FEBS Letters* 476 (1–2), 47–51.
- Hofmeyr, J.-H., Kacser, H., van der Merwe, K.J., 1986. Metabolic control analysis of moiety-conserved cycles. *European Journal of Biochemistry* 155, 631–641.
- Hofmeyr, J.H., Rohwer, J.M., 2011. Supply-demand analysis a framework for exploring the regulatory design of metabolism. *Methods in Enzymology* 500, 533–554.
- Huang, X., Holden, H.M., Raushel, F.M., 2001. Channeling of substrates and intermediates in enzyme-catalyzed reactions. *Annual Review of Biochemistry* 70, 149–180.
- Ingalls, B., 2013. *Mathematical Modeling in Systems Biology: An Introduction*. Cambridge, MA: MIT Press.
- Ingalls, B.P., Sauro, H.M., 2003. Sensitivity analysis of stoichiometric networks: An extension of metabolic control analysis to non-steady state trajectories. *Journal of Theoretical Biology* 222 (1), 23–36.
- Kacser, H., 1983. The control of enzyme systems in vivo: Elasticity of the steady state. *Biochemical Society Transactions* 11, 35–40.
- Kacser, H., Burns, J.A., 1973. The control of flux. In: Davies, D.D. (Ed.), *Rate Control of Biological Processes*, Volume 27 of *Symposia of the Society for Experimental Biology*. Cambridge: Cambridge University Press, pp. 65–104.
- Kholodenko, B.N., Sauro, H.M., Westerhoff, H.V., 1994. Control by enzymes, coenzymes and conserved moieties. A generalisation of the connectivity theorem of metabolic control analysis. *European Journal of Biochemistry* 225, 179–186.
- Klumpp, S., Hwa, T., 2014. Bacterial growth: Global effects on gene expression, growth feedback and proteome partition. *Current Opinion in Biotechnology* 28, 96–102.
- Klumpp, S., Scott, M., Pedersen, S., Hwa, T., 2013. Molecular crowding limits translation and cell growth. *Proceedings of the National Academy of Sciences of the United States of America* 110 (42), 16754–16759.
- LaPorte, D.C., Walsh, K., Koshland, D.E., 1984. The branch point effect. Ultrasensitivity and subsensitivity to metabolic control. *Journal of Biological Chemistry* 259 (22), 14068–14075.
- Morandini, P., 2009. Rethinking metabolic control. *Plant Science* 176 (4), 441–451.
- Othmer, H.H., 1976. The quantitative dynamics of a class of biochemical control circuits. *Journal of Mathematical Biology* 37, 53–78.
- Palsson, B.O., 2007. *Systems Biology: Properties of Reconstructed Networks*. Cambridge: Cambridge University Press.
- Qian, H., Beard, D.A., 2006. Metabolic futile cycles and their functions: A systems analysis of energy and control. *Systems Biology (Stevenage)* 153 (4), 192–200.
- Reder, C., 1988. Metabolic control theory: A structural approach. *Journal of Theoretical Biology* 135, 175–201.
- Reich, J.G., Selkov, E.E., 1981. *Energy Metabolism of the Cell*. London: Academic Press.
- Reijenga, K.A., Westerhoff, H.V., Kholodenko, B.N., Snoep, J.L., 2002. Control analysis for autonomously oscillating biochemical networks. *Biophysical Journal* 82 (1 Pt 1), 99–108.
- Richard, P., Teusink, B., Hemker, M.B., Van Dam, K., Westerhoff, H.V., 1996. Sustained oscillations in free-energy state and hexose phosphates in yeast. *Yeast* 12 (8), 731–740.
- Rohwer, J.M., Hofmeyr, J.H., 2008. Identifying and characterising regulatory metabolites with generalised supply-demand analysis. *Journal of Theoretical Biology* 252 (3), 546–554.
- Ruijter, G.J., Panneman, H., Visser, J., 1997. Overexpression of phosphofructokinase and pyruvate kinase in citric acid-producing *Aspergillus niger*. *Biochimica et Biophysica Acta* 1334 (2–3), 317–326.
- Santos, F., Boele, J., Teusink, B., 2011. A practical guide to genome-scale metabolic models and their analysis. *Methods in Enzymology* 500, 509–532.
- Sauro, H.M., 1990. Quantification of metabolic regulation by effectors. In: Cornish-Bowden, A., Cárdenas, M.L. (Eds.), *Control of Metabolic Processes*, NATO ASI Series. New York, NY: Plenum Press, pp. 225–230. chapter 17.
- Sauro, H.M., 2009. Network dynamics. *Methods in Molecular Biology* 541, 269–309.
- Sauro, H.M., 2012. *Enzyme Kinetics for Systems Biology*, second ed. Ambrosius Publishing. ISBN-10: 0-9824773-3-3.
- Sauro, H.M., 2014. *Systems Biology: An Introduction to Pathway Modeling*. Seattle: Ambrosius Publishing. ISBN: 0982477376.
- Sauro, H.M., Ingalls, B., 2004. Conservation analysis in biochemical networks: Computational issues for software writers. *Biophysical Chemistry* 109 (1), 1–15.
- Sauro, H.M., Kacser, H., 1990. Enzyme-enzyme interactions and control analysis. *European Journal of Biochemistry* 187 (3), 493–500.
- Savageau, M.A., 1972. *Current Topics in Cellular Regulation* 6, 63–130.
- Savageau, M.A., 1974. *Journal of Molecular Evolution* 4, 139–156.
- Savageau, M.A., 1976. *Biochemical Systems Analysis: A Study of Function and Design in Molecular Biology*. Reading, Mass: Addison-Wesley.
- Schuster, R., Schuster, S., 1993. Refined algorithm and computer program for calculating all non-negative fluxes admissible in steady states of biochemical reaction systems with or without some flux rates fixed. *Computer Applications in the Biosciences* 9 (1), 79–85.
- Smallbone, K., Messiha, H.L., Carroll, K.M., *et al.*, 2013. A model of yeast glycolysis based on a consistent kinetic characterisation of all its enzymes. *FEBS Letters* 587 (17), 2832–2841.
- Srere, P.A., Ovadi, J., 1990. Enzyme-enzyme interactions and their metabolic role. *FEBS Letters* 268 (2), 360–364.
- Stephanopoulos, G., Aristidou, A., Nielsen, J., 1998. *Metabolic Engineering*. London: Academic Press.
- Stricker, J., Cookson, S., Bennett, M.R., *et al.*, 2008. A fast, robust and tunable synthetic gene oscillator. *Nature* 456 (7221), 516–519.
- Terzer, M., Maynard, N.D., Covert, M.W., Stelling, J., 2009. Genome-scale metabolic networks. *Wiley Interdisciplinary Reviews Systems Biology and Medicine* 1 (3), 285–297.

- Thiele, I., Palsson, B.Ø., 2010. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols* 5 (1), 93–121.
- Tyson, J., Othmer, H.G., 1978. The dynamics of feedback control circuits in biochemical pathways. In: Rosen, R., Snell, F.M. (Eds.), *Progress in Theoretical Biology*, vol. 5. New York: Academic Press, pp. 1–62.
- Umbarger, H.E., 1956. Evidence for a negative-feedback mechanism in the biosynthesis of isoleucine. *Science* 123 (3202), 848–849.
- Voet, D., Voet, J.G., 2004. *Biochemistry*, third ed. Wiley. ISBN: 047119350X.
- Wanders, R.J., Groen, A.K., Van Roermund, C.W., Tager, J.M., 1984. Factors determining the relative contribution of the adenine-nucleotide translocator and the ADP-regenerating system to the control of oxidative phosphorylation in isolated rat-liver mitochondria. *European Journal of Biochemistry* 142 (2), 417–424.
- Welch, G.R., Easterby, J.S., 1994. Metabolic channeling versus free diffusion: Transition-time analysis. *Trends in Biochemical Sciences* 19 (5), 193–197.
- Westerhoff, H.V., Groen, A.K., Wanders, R.J.A., 1984. Bioscience Reports 4, 1–22.
- Yates, R.A., Pardee, A.B., 1956. Control of pyrimidine biosynthesis in *Escherichia coli* by a feed-back mechanism. *Journal of Biological Chemistry* 221 (2), 757–770.
- You, C., Okano, H., Hui, S., *et al.*, 2013. Coordination of bacterial proteome with metabolism by cyclic AMP signalling. *Nature* 500 (7462), 301–306.
- Zampar, G.G., Kümmel, A., Ewald, J., *et al.*, 2013. Temporal system-level organization of the switch from glycolytic to gluconeogenic operation in yeast. *Molecular Systems Biology* 9, 651.
- Zhang, Q., Bhattacharya, S., Andersen, M.E., 2013. *Open Biology* 3 (4), 130031.

Further Reading

- Fell, D., 1997. *Understanding the Control of Metabolism*. London: Portland Press. ISBN-13: 978-1855780477.
- Ingalls, B., 2013. *Mathematical Modeling in Systems Biology: An Introduction*. The MIT Press. ISBN-13: 978-0262018883.
- Kacser, H., Burns, J.A., 1973. Rate Control of Biological Processes, vol. 27 of *Symposia of the Society for Experimental Biology*. Davies, D.D. (Ed.), Cambridge University Press. p. 65–104.
- Michal, G., Schomburg, D., 2012. *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, second edition, Wiley. ISBN-13: 978-0470146842.
- Moreno-Sánchez, R., Saavedra, E., Rodríguez-Enríquez, S., and Olín-Sandoval, V., 2008. Metabolic control analysis: A tool for designing strategies to manipulate metabolic pathways. *Journal of Biomedicine and Biotechnology* 2008, <http://dx.doi.org/10.1155/2008/597913>
- Sauro, H.M., 2014. *Systems Biology: Introduction to Pathway Modeling*. Ambrosius Publishing. ISBN 13: 978-0-9824773-7-3.
- Savageau M. (2010) *Biochemical Systems Analysis: A Study of Function and Design in Molecular Biology*. CreateSpace Independent, ISBN-13: 978–1449590765
- Steuer, R., Junker, B.H., 2009. Computational models of metabolism: Stability and regulation in metabolic networks. In: Rice, S.A. (Ed.), *Advances in Chemical Physics*, Volume 142. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Dynamics of Microtubule Self-Assembly

BT Castle and DJ Odde, University of Minnesota, Minneapolis, MN, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Catastrophe The sudden switch from growth to rapid shortening at the microtubule end corresponding to the loss of the guanosine-5'-triphosphate (GTP) cap.

Dynamic instability Microtubule self-assembly process characterized by the stochastic switching between extended phases of growth and rapid shortening.

E-site Exchangeable nucleotide-binding site on β -tubulin.

GTP cap A stabilizing region of tubulin subunits at the plus-end that are bound with GTP at the E-site.

Kinetic stabilization The attenuation of microtubule self-assembly dynamics by microtubule-targeting agents.

N-site Non-exchangeable nucleotide-binding site on α -tubulin.

Plus-end The dynamic end of the microtubule where β -tubulin is exposed.

Rescue Recovery of the growth phase after rapid shortening.

Tip taper The gradual loss of protofilaments at the tip of microtubules.

Cellular Functions of Microtubules

Microtubules are rigid, tubular polymers that constitute one of the three self-assembled filament structures of the cytoskeleton, the other two being filamentous actin and intermediate filaments. Microtubules serve a multitude of functions in cellular processes, a few of which are outlined in [Figure 1](#), and in general serve to support and maintain intracellular organization. One particularly important function of microtubules is to serve as intracellular 'superhighways' for molecular motor-based transport. Molecular motors, such as kinesin and dynein, deliver vesicle-bound cargos to remote locations using ATP hydrolysis as the energy source in the cell, thereby overcoming fundamental diffusional limits. An especially important cargo is the kinetochore, a protein complex that mechanically links the ends of microtubules to chromosomes during mitosis. By tracking the dynamic growing and shortening ends of microtubules, kinetochores use polymerization and depolymerization forces to position chromosomes in the center of the mitotic spindle during metaphase in preparation for chromosome segregation during anaphase. The dynamic organizational support provided by microtubules is particularly important for neurons to form extensive connections via axons and dendrites, which can be up to a meter in length. Transport of intracellular proteins along these axonal and dendritic extensions then exploits microtubule-based molecular motors to deliver cargo throughout the cell. Further, the rigidity of microtubules allows them to generally withstand the pushing and pulling forces necessary to exert their influence over micrometer-scale distances. In general, each of these cellular functions relies on both the microtubule structure and the dynamics of microtubule self-assembly.

Microtubule Structure

Microtubules are composed of globular protein subunits or building blocks, which are heterodimers of α - and β -tubulin. Heterodimeric subunits stack end-to-end in linear protofilaments that interact laterally to form a cylindrical closed tube

approximately 25 nm in outer diameter ([Figure 2](#)). Microtubules are most commonly observed to have 13 protofilaments, but the number can vary by 1–2 protofilaments depending upon the environmental conditions, tubulin isoform, or the organism. Adjacent protofilaments are offset slightly such that the microtubule lattice typically has a left-handed helical pitch of 1.5 dimers per revolution ([Figure 2](#)). The parallel alignment of individual heterodimers results in an inherent polarity with α -tubulin monomers facing out at the minus-end and β -tubulin monomers facing out at the plus-end. Molecular motors, kinesins and dynein, recognize this polarity, exhibiting directional bias to move toward the plus- and minus-end, respectively.

Due to the cylindrical nature of the microtubule structure, individual tubulin subunits in the microtubule lattice interact longitudinally, within individual protofilaments, and laterally, between adjacent protofilaments. In the so-called 'B-lattice' structure (that most commonly observed), the majority of lateral contacts are between homologous subunits, i.e., α - α and β - β , except at the seam between protofilaments 1 and 13, where α - β contacts are made ([Figure 2](#)). These intersubunit interactions not only provide structural integrity but influence the microtubule assembly process as well. Intersubunit contacts are noncovalent, thus the binding of individual subunits is a reversible process. This has important consequences for the dynamics of self-assembly. First, binding and unbinding events are very rapid at typical free subunit concentrations (5–12 μ M), occurring in the order of milliseconds ([Gardner et al., 2011](#)). Additionally, microtubule self-assembly is highly dynamic and rapidly influenced by microtubule-binding compounds and intracellular microtubule-associated proteins (MAPs) that bind directly to microtubules.

Microtubule Self-Assembly Dynamics

Dynamic Instability

Microtubule self-assembly occurs by an unusual process termed 'dynamic instability,' in which individual microtubules

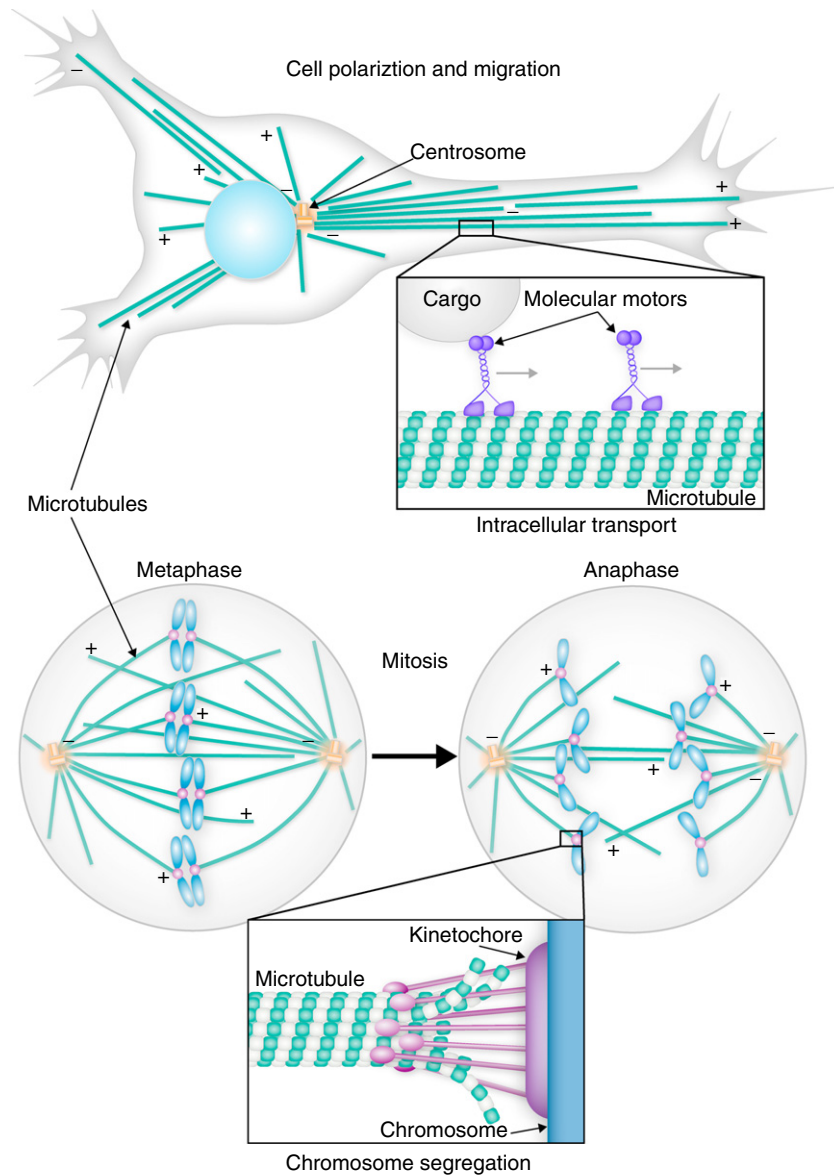


Figure 1 Examples of cellular processes that require dynamic microtubules.

stochastically switch between extended periods of growth and rapid shortening (Figure 3), while the average length of the microtubule population remains approximately constant (Mitchison and Kirschner, 1984). Growth is the net difference between association and dissociation of individual subunits, while shortening is dominated by rapid loss of subunits. The switch from growth to shortening is referred to as a 'catastrophe,' while 'rescue' is the reverse switch back to growth. The fact that dynamic instability is observed in purified *in vitro* assays demonstrates that it is an intrinsic property of microtubules, not dependent on MAPs, and is the result of subunit binding and unbinding, GTP hydrolysis in the lattice, and nucleotide exchange in solution.

Dynamic instability is a nonequilibrium process, driven by GTP hydrolysis (Figure 3). Initially, tubulin subunits have GTP bound to both α - and β -tubulin monomers. Hydrolysis occurs only at the exchangeable GTP binding site on β -tubulin

(E-site) while the non-exchangeable site (N-site) on α -tubulin remains bound to GTP (Carlier and Pantaloni, 1981). Because the molecule of GTP at the E-site is exposed at the plus-end, hydrolysis is coupled to polymerization, accelerating after an additional subunit binds longitudinally and closes the hydrolysis pocket at the terminal β -tubulin end (Nogales *et al.*, 1999, 1998). E-site GDP is exchanged for GTP after dissociation, but remains bound while the subunit is in the microtubule lattice. Within the microtubule lattice, GDP-tubulin is less stable, i.e., is shorter lived and unbinds rapidly due to less favorable intersubunit interactions. Thus, the microtubule tip assembly state, whether growing or shortening, is determined by the presence or absence, respectively, of a stabilizing cap of GTP-tubulin at the plus-end (Figure 3). Loss of the GTP cap, due to the combination of hydrolysis and the stochastic unbinding of GTP-tubulin, results in catastrophe and rapid disassembly while rescue occurs with the reformation of the GTP

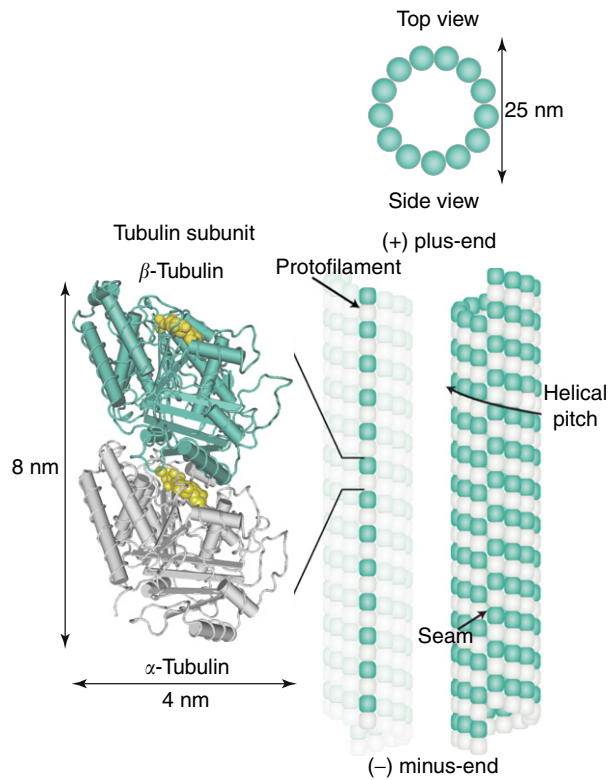


Figure 2 Diagram of microtubule structure.

cap. Microtubules *in vitro* exhibit dynamic instability at both the plus- and minus-ends. However, *in vivo* the minus-end is typically associated with γ -tubulin containing nucleating structures that are often associated with centrosomes, such that assembly dynamics are largely mediated via the plus-end.

The experimental observation of dynamic instability only occurs in a relatively narrow range of free GTP-tubulin concentration near the critical concentration for growth, which is about 5–12 μM *in vitro* at physiological temperature (Walker *et al.*, 1988). Above $\sim 20 \mu\text{M}$ microtubules spontaneously nucleate and grow, further limiting the practical range of concentrations for *in vitro* experiments. Below the range where dynamic instability is observed, microtubules are unable to maintain growth and above they grow persistently as the more rapidly assembling plus-end outpaces GTP hydrolysis. Therefore, the observation of dynamic instability requires a balance between the net rate of subunit addition and the rate of GTP hydrolysis. When bound with a slowly hydrolyzing analog of GTP, GMPCPP, instead of GTP microtubules do not exhibit dynamic instability (Hyman *et al.*, 1992), further emphasizing the importance of hydrolysis and nucleotide exchange in addition to subunit addition and loss in the functional dynamics of microtubules.

Plus-End Tip Structures

In addition to the nucleotide state, structural features at the microtubule tip also correlate with growth and shortening (Figure 3), as observed using both electron and light microscopy. During growth, microtubules exhibit either tip

structures ranging from relatively blunt and straight conformations (the majority) to gradually tapering and gently radially curling extensions away from the closed tube (Chrétien *et al.*, 1995; Mandelkow *et al.*, 1991). During shortening, individual protofilaments peel away from the microtubule lattice in highly curled structures (Chrétien *et al.*, 1995; Mandelkow *et al.*, 1991; Tran *et al.*, 1997). This observation of the structural differences between growing and shortening microtubules led to the hypothesis that the energy of GTP hydrolysis is stored in the microtubule lattice as mechanical strain energy, due to constraining GDP-tubulin away from a preferred curled orientation in order to comply with the straight conformation of the microtubule lattice. This strain energy destabilizes the intersubunit interactions, causing GDP-tubulin to be less stable in the lattice than GTP-tubulin. Alternatively, observations that both GTP- and GDP-tubulin have a curled orientation in solution led to the recent hypothesis that subunits undergo a lattice-induced conformational change upon binding (Rice *et al.*, 2008). By this mechanism, GTP-tubulin must have stronger lateral contacts to resist curling in order to be consistent with observed structures by electron microscopy. In either case, GTP-tubulin is more stable within the lattice structure compared to GDP-tubulin, and thus GDP-tubulin rapidly disassembles when exposed at the plus-end (Figure 3).

Observations of microtubule plus-ends during growth by both electron and fluorescence microscopy are most consistent with a subset of protofilaments that extend out from the complete tube (Chrétien *et al.*, 1999; Demchouk *et al.*, 2011), which manifests itself as a tapered tip. More tapered tip structures form at higher free tubulin concentrations, where net assembly is fastest, while remaining primarily blunt at low concentrations (Chrétien *et al.*, 1995; Gardner *et al.*, 2011). At short time scales ($< 1 \text{ s}$), tip tapers fluctuate with the rapid addition and loss of individual tubulin subunits, although at longer time scales ($> 10 \text{ s}$) tip tapers on average increase in length during periods of growth (Coombes *et al.*, 2013). As a subset of protofilaments extends out from the closed tube, the number of stabilizing lateral contacts decreases, and therefore the average energetic state of tubulin subunits at the end of the microtubule becomes increasingly less stable. Thus, microtubule self-assembly dynamics are the result of an active interplay of the kinetics and thermodynamics of subunit addition and loss with the resulting polymer end structure. Subunit addition and loss results in fluctuating polymer end structures, which further feed back to determine the rates of addition and loss.

Cellular Microtubule Dynamics and MAPs

To observe single microtubule dynamics in living cells, light microscopy has been, and continues to be, the dominant approach. Transmitted light microscopy, combined with video processing to enhance image contrast, facilitated the earliest estimates of the dynamic instability parameters (growth rate, shortening rate, catastrophe frequency, and rescue frequency) *in vitro* and *in vivo* (Cassimeris *et al.*, 1988; Walker *et al.*, 1988), as well as early estimates of the rates of tubulin subunit addition and loss (Walker *et al.*, 1988). Since then, the

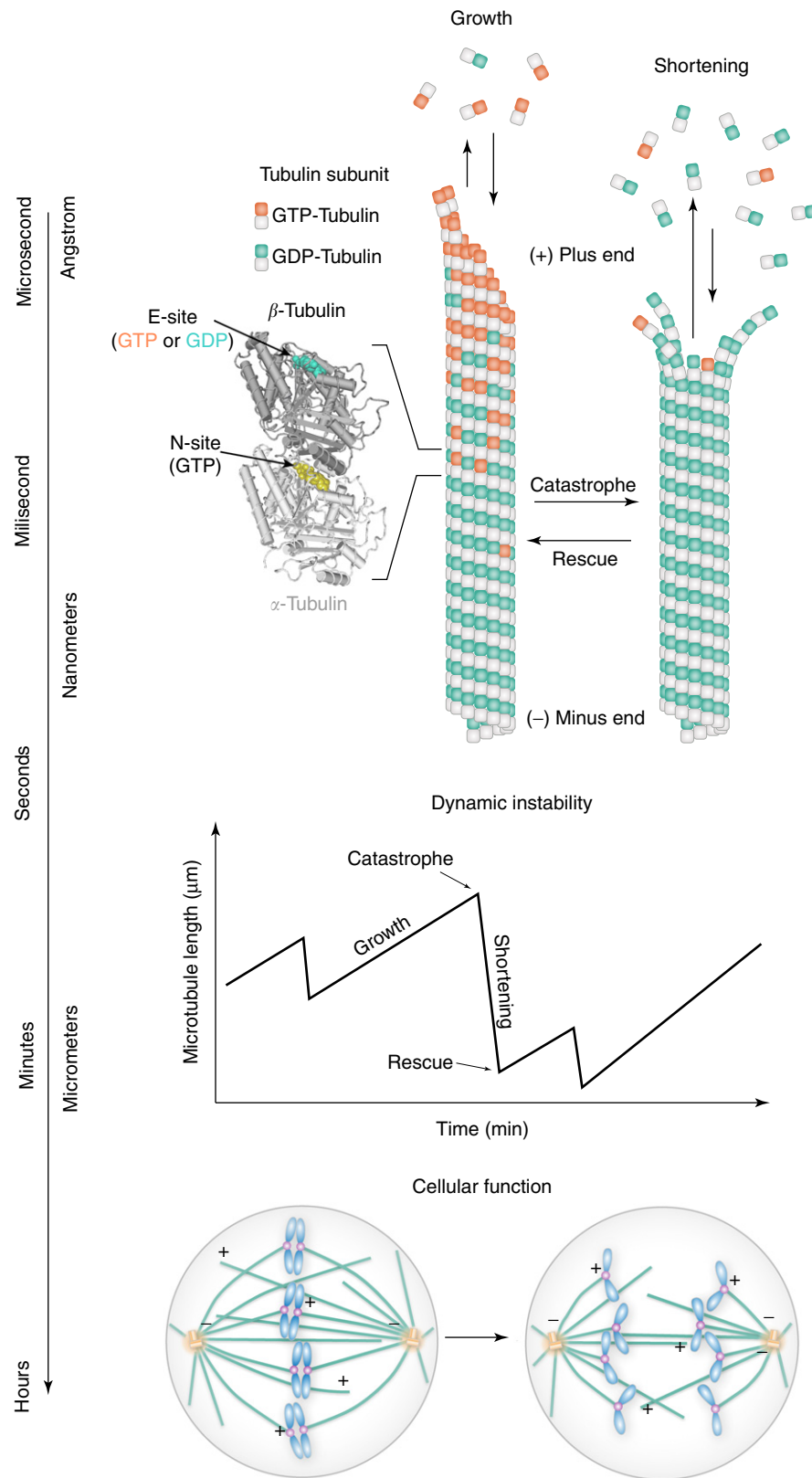


Figure 3 Microtubule self-assembly dynamics cover a wide range of length and time scales. At the atomistic level, groups of residues as well as the nucleotide state dictate the interdimer interactions (top, left). These interactions will influence the rate of addition and loss of individual subunits from the microtubule plus-end (top, right). The net result of addition and loss is dynamic instability (middle). Finally, dynamic instability facilitates microtubule function at the cellular level to mediate processes such as mitosis (bottom).

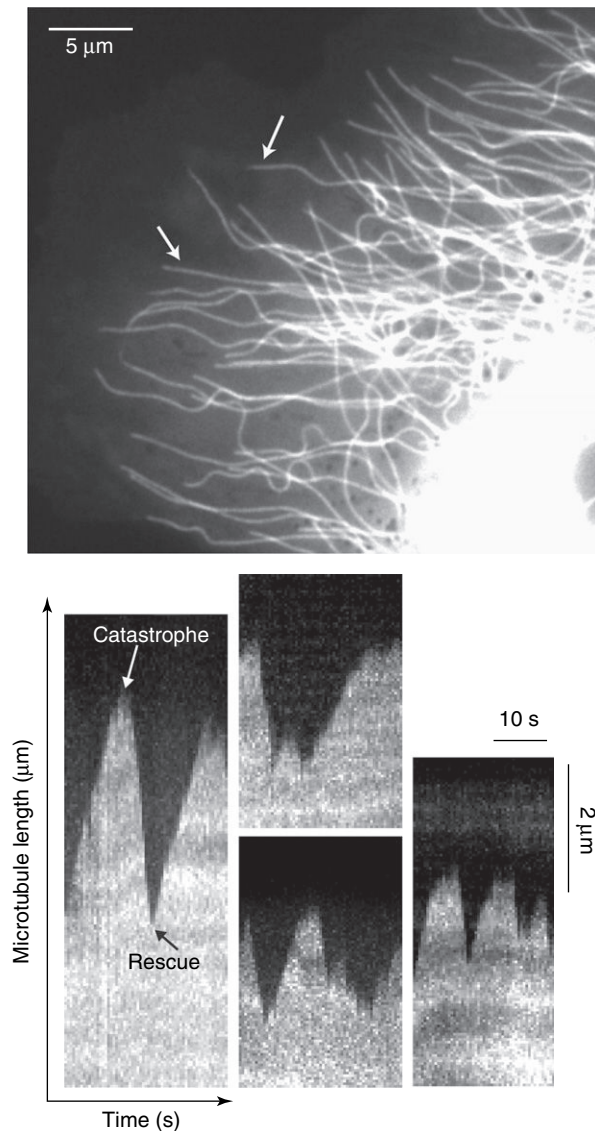


Figure 4 Microtubule assembly dynamics in LLC-PK1 epithelial cells. Individual microtubules are resolved near the cell edge (arrows). Below: kymograph (space-time) representation of fluorescence intensity along the length of individual microtubules showing dynamic instability in live cells.

development of sensitive digital cameras as well as better fluorophores, including genetically-encoded fluorescent fusion proteins, has made fluorescence microscopy the preferred approach for both *in vitro* assays with purified tubulin and in live cells. *In vivo*, areas with low microtubule density near the edge of relatively flat, thin cells facilitate the visualization and tracking of individual microtubules in time (Figure 4).

Compared to purified tubulin-based *in vitro* assays, microtubule dynamics *in vivo* are enhanced, exhibiting increased rates of growth and shortening as well as catastrophe and rescue. This is most likely due to the presence of cellular MAPs that influence the assembly and disassembly process, particularly those that are biased to associate with the dynamic plus-end, such as kinesin motors, end-binding proteins (EBs), and XMAP215. XMAP215 (ch-TOG in humans) is a plus-end

directed polymerization-promoting protein that increases the rate of both growth and shortening (Brouhard *et al.*, 2008). Kinesins and EBs both promote catastrophe despite opposite effects on the microtubule growth rate. Additionally, microtubules frequently undergo catastrophe as they grow against the cell edge. For the most part, *in vivo* microtubules grow persistently to the cell edge, where they undergo several rounds of catastrophe and rescue before rapidly shortening back to the nucleation point, typically the centrosome (Komarova *et al.*, 2002). This behavior allows dynamic interphase microtubules to explore the intracellular space, and accumulate in the longest directions, leading to spontaneous polarization of the microtubule array as occurs in neurons (Seetapun and Odde, 2010). During mitosis, the dynamics of microtubules facilitate chromosome-kinetochore ‘search-and-capture’ and subsequent chromosome segregation.

Additional MAPs including those that track growing plus-ends (e.g., CLIP-170) or bind to the microtubule lattice (e.g., tau and MAP2) influence microtubule assembly dynamics. EB1 recruits CLIP-170 and other plus-end tracking proteins (+TIPs) to the microtubule plus-end, which it recognizes via preferential binding of GTP-tubulin. Neuronal MAPs, tau and MAP2, stabilize microtubules, increasing the growth rate and rescue frequency while reducing shortening and catastrophe (Pryer *et al.*, 1992), thus enabling the substantial lengths of neuronal axons. Even though these MAPs associate with microtubules and influence assembly dynamics, they are not necessary to reproduce the increased dynamics observed *in vivo* compared to *in vitro*. Minimally, the combination of a polymerization-promoting protein (i.e., XMAP215) and a catastrophe promoter (i.e., kinesin), *in vitro* is sufficient to reproduce *in vivo*-like dynamics (Kinoshita *et al.*, 2001; Zanich *et al.*, 2013).

Overall, microtubule dynamics cover a wide range of length and time scales (Figure 3). At the atomistic level, groups of residues as well as the nucleotide state dictate the intersubunit interactions. These interactions influence the rate of addition and loss of individual subunits from the microtubule plus-end. The net result of subunit addition and loss, along with hydrolysis and nucleotide exchange, manifests in dynamic instability, which is necessary for microtubules to function at the cellular level.

Theoretical Framework for Microtubule Self-Assembly

With the development of more advanced instrument capabilities, as well as fluorescence imaging techniques, the amount of quantitative data available to the field has increased dramatically. The challenge now is to use this information to extract meaningful conclusions about complex biological systems. By integrating hypotheses and assumptions, theoretical models not only aid in understanding experimental results but can also be used to predict and design future experiments. In particular, with the rapid advances in digital computing, computational modeling has emerged as a useful tool for analyzing results and predictions, as well as hypothesis testing in cell biology, where systems are often complex.

The current kinetic-thermodynamic computational framework for microtubule self-assembly at the level of individual

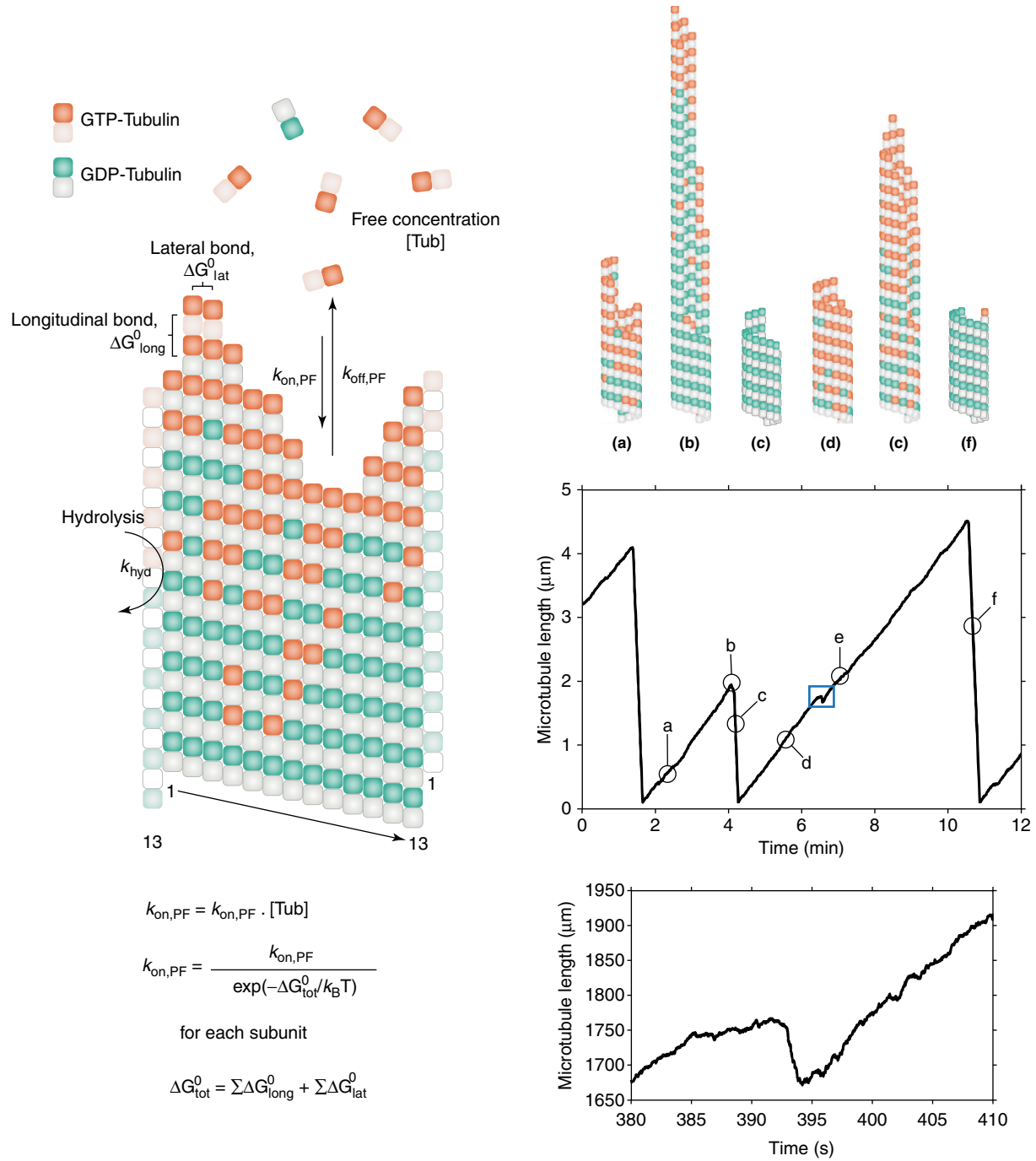


Figure 5 Theoretical kinetic-thermodynamic framework for microtubule self-assembly. Examples of resulting microtubule plus-end structures are shown for multiple time points during the simulation (a–f). *Note:* preferential curling of tubulin subunits is not included in the basic framework shown. Lower right: zoomed view of microtubule assembly from simulation output shown above (blue box). The theoretical framework predicts the variability of microtubule self-assembly, including the observation of short periods of net disassembly during extended periods of overall growth.

subunit addition and loss is outlined in Figure 5 (Gardner *et al.*, 2011; VanBuren *et al.*, 2002), which derives from the earlier modeling of Chen and Hill (1985). In this framework, the rate of subunit addition to a single protofilament ($k_{\text{on,PF}}^*$ in units of $\text{s}^{-1}\text{PF}^{-1}$) is the product of the individual protofilament on-rate constant ($k_{\text{on,PF}}$ in units of $\text{M}^{-1}\text{s}^{-1}\text{PF}^{-1}$) and the concentration of tubulin subunits free in

solution ($[\text{Tub}]$, in units of M). The total number of addition events is then the sum of the individual protofilament on-rates ($k_{\text{on,MT}}^* = k_{\text{on,PF}}^* \times n_{\text{PF}}$, in units of s^{-1}). For individual subunits to ultimately bind to the microtubule lattice, subunits must diffuse to a close distance as well as orient correctly. This stereospecific constraint results in a very inefficient process such that the vast majority of free subunits that diffuse to

within a few nanometers of a potential binding site at the microtubule tip are predicted to fail to form a stable contact (Castle and Odde, 2013; Northrup and Erickson, 1992).

Once a contact with the microtubule lattice is established, the dissociation rate of individual subunits ($k_{\text{off,PF}}$ in units of s^{-1}) is determined by the sum of the longitudinal and lateral bond free energies ($\Delta G_{\text{long}}^{\circ}$ and $\Delta G_{\text{lat}}^{\circ}$ respectively), which are themselves a combination of the chemical interactions as well as the mechanical and entropic energies of constraining subunits within the microtubule lattice. The total bond free energy of GDP-tubulin subunits must be more positive, i.e., less stable, compared to GTP-tubulin in order for rapid disassembly to occur upon loss of the GTP cap. Of the two bonds, longitudinal bonds are approximately three-fold stronger than lateral bonds (Erickson and Voter, 1986; Sept *et al.*, 2003; VanBuren *et al.*, 2002). Because of this difference in stability, the longitudinal bond is required for subunit incorporation into the microtubule lattice, while the lateral bonds provide additional stabilization, largely reducing the rate of unbinding (Castle and Odde, 2013; VanBuren *et al.*, 2002). Subunits with increasing number of favorable interactions are more stable and dissociate several orders of magnitude slower than those with fewer favorable interactions. GTP hydrolysis is treated as a random (i.e., first-order) process, which occurs only after an additional subunit is bound longitudinally, stimulating hydrolysis at the E-site on β -tubulin. Consistent with a first-order process, the EB1 fluorescence signal, which preferentially binds GTP-tubulin, decays exponentially from the microtubule plus-end (Seetapun *et al.*, 2012).

This theoretical framework is sufficient to reproduce dynamic instability consistent with *in vitro* experimental observations and observations of dynamic microtubule end structures (Figure 5). Additionally, it captures the rapid nanoscale dynamics of assembly, including the high variability of assembly (Gardner *et al.*, 2011) as well as the observation of short periods of net disassembly during extended periods of overall growth (Schek *et al.*, 2007; Figure 5). Thus, the complex dynamics of microtubule dynamic instability, and self-assembly in general, naturally arise from these simple, physical principles.

Microtubule-Targeting Agents Attenuate Microtubule Dynamics

Microtubule-targeting agents (MTAs) are a diverse group of natural compounds and synthetic derivatives that bind microtubules and influence their self-assembly dynamics. In contrast to MAPs, which enhance microtubule dynamics, MTAs result in 'kinetic stabilization' or the attenuation of dynamic instability. The earliest noted MTAs include colchicine, the vinca alkaloids vinblastine and vincristine, and paclitaxel. MTAs are widely used chemotherapeutic options for the treatment of various types of cancer (reviewed in Dumontet and Jordan, 2010). This success stems from the fundamental role of microtubules in cellular processes (Figure 1), with a particular focus on the role of dynamic microtubules in mitosis. Additionally, MTAs have been used to treat gout (colchicine) and used in drug-eluting stents (paclitaxel) to prevent restenosis as a result of stent placement during the reopening

of occluded arteries. Further, the efficacy of paclitaxel in promoting axon regeneration to reduce spinal cord injury as well as potential for the treatment of neurodegenerative diseases is beginning to be explored.

MTAs bind directly to individual $\alpha\beta$ -tubulin heterodimers at three primary sites, which correlate with their effect on net polymer assembly. For the most part, disassembly-promoters bind to either the colchicine site at the intradimer interface between α - and β -tubulin monomers (Dorléans *et al.*, 2009) or the vinca domain at the terminal end of β -tubulin near the E-site (Gigant *et al.*, 2005). Colchicine site binders include the combretastatins in addition to colchicine, while vinca domain binders include the vinca alkaloids, dolastatins, and eribulin. Microtubule assembly-promoters bind to the taxane pocket, on the inside of microtubules near the lateral interface between adjacent dimers (Nogales *et al.*, 1999). Assembly-promoters include paclitaxel and its synthetic derivative, docetaxel, as well as the epothilones. This categorical assignment based on net assembly may be a bit misleading, however, as the effectiveness of all MTAs may ultimately depend on their capability to inhibit microtubule dynamic instability. Kinetic stabilization by MTAs increases the amount of time that microtubules spend in a non-dynamic 'paused' state, thereby eliminating cells' capacity to complete essential processes. As dynamic microtubules function in both mitotic and interphase cells, the side effects of MTAs are relatively frequent and often dose limiting. Most commonly, peripheral neuropathies occur during treatment, although additional side effects of bone marrow suppression and flu-like symptoms have been noted as well. MTA-induced kinetic stabilization is the culmination of the primary effects on individual tubulin subunit kinetics and thermodynamics at the molecular level (outlined in Figure 5). Therefore, further understanding the interplay between the microtubule self-assembly dynamics and microtubule structure may give us new insights into new and better-tolerated therapeutic strategies.

See also: Cytoskeleton and Motors: Complex Cytoskeletal Structures: The Mitotic Spindle. Cytoskeleton and Motors: Cytoskeletal Components: Microtubules and Microtubule-Associated Proteins (MAPs)

References

- Brouhard, G.J., Stear, J.H., Noetzel, T.L., *et al.*, 2008. XMAP215 is a processive microtubule polymerase. *Cell* 132, 79–88.
- Carlier, M.F., Pantaloni, D., 1981. Kinetic analysis of guanosine 5'-triphosphate hydrolysis associated with tubulin polymerization. *Biochemistry* 20, 1918–1924.
- Cassimeris, L., Pryer, N.K., Salmon, E.D., 1988. Real-time observations of microtubule dynamic instability in living cells. *Journal of Cell Biology* 107, 2223–2231.
- Castle, B.T., Odde, D.J., 2013. Brownian dynamics of subunit addition-loss kinetics and thermodynamics in linear polymer self-assembly. *Biophysical Journal* 105, 2528–2540.
- Chen, Y.D., Hill, T.L., 1985. Monte Carlo study of the GTP cap in a five-start helix model of a microtubule. *Proceedings of the National Academy of Sciences of the United States of America* 82, 1131–1135.
- Chrétien, D., Fuller, S.D., Karsenti, E., 1995. Structure of growing microtubule ends: Two-dimensional sheets close into tubes at variable rates. *Journal of Cell Biology* 129, 1311–1328.

- Chrétien, D., Jánosi, I.M., Taveau, J.C., Flyvbjerg, H., 1999. Microtubule's conformational cap. *Cell Structure and Function* 24, 299–303.
- Coombes, C.E., Yamamoto, A., Kenzie, M.R., Odde, D.J., Gardner, M.K., 2013. Evolving tip structures can explain age-dependent microtubule catastrophe. *Current Biology* 23, 1342–1348.
- Demchouk, A.O., Gardner, M.K., Odde, D.J., 2011. Microtubule tip tracking and tip structures at the nanometer scale using digital fluorescence microscopy. *Cellular and Molecular Bioengineering* 4, 192–204.
- Dorléans, A., Gigant, B., Ravelli, R.B.G., *et al.*, 2009. Variations in the colchicine-binding domain provide insight into the structural switch of tubulin. *Proceedings of the National Academy of Sciences of the United States of America* 106, 13775–13779.
- Dumontet, C., Jordan, M.A., 2010. Microtubule-binding agents: A dynamic field of cancer therapeutics. *Nature Reviews Drug Discovery* 9, 790–803.
- Erickson, H.P., Voter, W.A., 1986. Nucleation of microtubule assembly. Experimental kinetics, computer fitting of models, and observations on tubulin rings. *Annals of the New York Academy of Sciences* 466, 552–565.
- Gardner, M.K., Charlebois, B.D., Jánosi, I.M., *et al.*, 2011. Rapid microtubule self-assembly kinetics. *Cell* 146, 582–592.
- Gigant, B., Wang, C., Ravelli, R.B.G., *et al.*, 2005. Structural basis for the regulation of tubulin by vinblastine. *Nature* 435, 519–522.
- Hyman, A., Salsler, S., Drechsel, D., Unwin, N., Mitchison, T., 1992. Role of GTP hydrolysis in microtubule dynamics: Information from a slowly hydrolyzable analogue, GMPCPP. *Molecular Biology of the Cell* 3, 1155–1167.
- Kinoshita, K., Arnal, I., Desai, A., Drechsel, D.N., Hyman, A.A., 2001. Reconstitution of physiological microtubule dynamics using purified components. *Science* 294, 1340–1343.
- Komarova, Y.A., Vorobjev, I.A., Borisy, G.G., 2002. Life cycle of MTs: Persistent growth in the cell interior, asymmetric transition frequencies and effects of the cell boundary. *Journal of Cell Science* 115, 3527–3539.
- Mandelkow, E.M., Mandelkow, E., Milligan, R.A., 1991. Microtubule dynamics and microtubule caps: A time-resolved cryo-electron microscopy study. *Journal of Cell Biology* 114, 977–991.
- Mitchison, T., Kirschner, M., 1984. Dynamic instability of microtubule growth. *Nature* 312, 237–242.
- Nogales, E., Whittaker, M., Milligan, R.A., Downing, K.H., 1999. High-resolution model of the microtubule. *Cell* 96, 79–88.
- Nogales, E., Wolf, S.G., Downing, K.H., 1998. Structure of the alpha beta tubulin dimer by electron crystallography. *Nature* 391, 199–203.
- Northrup, S.H., Erickson, H.P., 1992. Kinetics of protein-protein association explained by Brownian dynamics computer simulation. *Proceedings of the National Academy of Sciences of the United States of America* 89, 3338–3342.
- Pryer, N.K., Walker, R.A., Skeen, V.P., *et al.*, 1992. Brain microtubule-associated proteins modulate microtubule dynamic instability *in vitro*. Real-time observations using video microscopy. *Journal of Cell Science* 103 (Pt 4), 965–976.
- Rice, L.M., Montabana, E.A., Agard, D.A., 2008. The lattice as allosteric effector: Structural studies of alphabeta- and gamma-tubulin clarify the role of GTP in microtubule assembly. *Proceedings of the National Academy of Sciences of the United States of America* 105, 5378–5383.
- Schek, H.T., Gardner, M.K., Cheng, J., Odde, D.J., Hunt, A.J., 2007. Microtubule assembly dynamics at the nanoscale. *Current Biology* 17, 1445–1455.
- Seetapun, D., Castle, B.T., McIntyre, A.J., Tran, P.T., Odde, D.J., 2012. Estimating the microtubule GTP cap size *in vivo*. *Current Biology* 22, 1681–1687.
- Seetapun, D., Odde, D.J., 2010. Cell-length-dependent microtubule accumulation during polarization. *Current Biology* 20, 979–988.
- Sept, D., Baker, N., McCammon, J., 2003. The physical basis of microtubule structure and stability. *Protein Science* 2257–2261.
- Tran, P.T., Joshi, P., Salmon, E.D., 1997. How tubulin subunits are lost from the shortening ends of microtubules. *Journal of Structural Biology* 118, 107–118.
- VanBuren, V., Odde, D.J., Cassimeris, L., 2002. Estimates of lateral and longitudinal bond energies within the microtubule lattice. *Proceedings of the National Academy of Sciences of the United States of America* 99, 6035–6040.
- Walker, R.A., O'Brien, E.T., Pryer, N.K., *et al.*, 1988. Dynamic instability of individual microtubules analyzed by video light microscopy: Rate constants and transition frequencies. *Journal of Cell Biology* 107, 1437–1448.
- Zanic, M., Widlund, P.O., Hyman, A.A., Howard, J., 2013. Synergy between XMAP215 and EB1 increases microtubule growth rates to physiological levels. *Nature Cell Biology* 15, 688–693.

Further Reading

- Cassimeris, L., 1993. Regulation of microtubule dynamic instability. *Cell Motility and the Cytoskeleton* 26, 275–281.
- Desai, A., Mitchison, T.J., 1997. Microtubule polymerization dynamics. *Annual Review of Cell and Developmental Biology* 13, 83–117.
- Howard, J., Hyman, A.A., 2009. Growth, fluctuation and switching at microtubule plus ends. *Nature Reviews Molecular Cell Biology* 10, 569–574.

Relevant Websites

- <http://oddelab.umn.edu/>
David Odde Lab.
- <http://www.ibiology.org/ibioseminars/cell-biology/anthony-hyman-part-2.html>
iBiology – Building a Polymer: Microtubule Dynamics.
- <http://www.ibiology.org/ibiomagazine/issue-6/gary-borisy-edwin-taylor-the-discovery-of-tubulin.html>
iBiology – Discovery of Tubulin.
- <http://www.ibiology.org/ibioseminars/cell-biology/tim-mitchison-part-1.html>
iBiology – Self-Organization of Microtubule Assemblies.
- <http://www.ibiology.org/ibioseminars/cell-biology.html>
iBiology – Seminars in Cell Biology.
- <http://medicine.yale.edu/lab/howard/>
Jonathan Howard Lab.
- <http://www.cbs.umn.edu/lab/gardner/>
Melissa Gardner Lab.
- <http://valelab.ucsf.edu/>
Ron Vale Lab.
- <http://mitchison.med.harvard.edu/>
Tim Mitchison Lab.

Dynamics of Protein Kinase Cascades

N Blüthgen, Charté – Universitätsmedizin Berlin, Berlin, Germany and Humboldt-Universität zu Berlin, Berlin, Germany
S Legewie, Institut für Molekulare Biologie gGmbH (IMB), Mainz, Germany

© 2016 Elsevier Inc. All rights reserved.

Glossary

Bistability Property of a dynamical system to have two stable states at a given parameter set. Bistable signaling cascades typically show either high or low activity in single cells at a given stimulus, thus allowing for binary decisions.

Hill coefficient Historically a coefficient in enzyme kinetics describing cooperativity, often used to quantify

ultrasensitivity of stimulus–response curves. Systems with Hill coefficients higher than one are ultrasensitive.

Oscillations Periodic changes in signal activity.

Sensitivity/logarithmic gain Relative signal amplification, defined as percent change in pathway activity divided by percent change in stimulus.

Ultrasensitivity A small relative change in signal elicits a large relative change in response, i.e., large logarithmic gain.

Introduction

Many signaling pathways are cascades of kinases, where one kinase activates the next by phosphorylation and thereby signals propagate through the protein network. Possibly the best-known and best-studied protein kinase signaling cascade is the mitogen-activated protein kinases (MAPK) signaling pathway, which serves as a paradigm of kinase signaling (Seger and Krebs, 1995). The core MAPK signaling pathway comprises three kinases: the MAP-kinase-kinase-kinase (MAP3K), which phosphorylates and thereby activates the MAP-kinase-kinase (MAP2K) that in turn activates the MAP-kinase by phosphorylation. This pathway is very well conserved among all eukaryotes. Human cells possess three major MAPK cascades (each having several isoforms), which are designated as extracellular regulated kinase (ERK), p38, and JNK pathways based on the name of their terminal cascade member (Dhillon *et al.*, 2007; Cargnello and Roux, 2011). The ERK signaling pathway has been implicated in the transmission of growth factor signals, while the p38 and JNK cascades play an important role in cellular stress responses (Dhillon *et al.*, 2007; Cargnello and Roux, 2011). Human cells also contain protein kinases cascades other than MAPK pathways: The PI3K-Akt signaling transmits growth factor signals like the ERK pathway, and has been implicated in cell division and cell survival control. The NF- κ B, SMAD, and the STAT pathways involve fewer steps than the classical MAPK cascade, as they involve only a single receptor-mediated phosphorylation event, which leads to nuclear translocation of a transcription factor.

During the last two decades, many studies have addressed the dynamics of MAPK kinase signaling (Kiel and Serrano, 2012). Mathematical models were employed to simulate the kinase kinetics and compared with quantitative experiments. In recent years, fluorescent techniques allowed for quantitative and time-resolved analysis of protein kinase signaling at the single-cell level (Nelson *et al.*, 2004; Cohen-Saidon *et al.*, 2009; Shankaran *et al.*, 2009; Colman-Lerner *et al.*, 2005). These studies have shed more light on the dynamical properties of protein kinase cascades. It is becoming clear that protein kinases cascades are very versatile in many aspects (Kholodenko *et al.*, 2010). When stimulated, they can generate

qualitatively different responses. For instance, a kinase cascade can show an oscillatory activation pattern, fluctuating between high and low activity states over time (Nelson *et al.*, 2004; Shankaran *et al.*, 2009). Alternatively, kinase cascades may establish cellular decision making by responding in a switch-like manner, once a threshold extracellular stimulus level has been exceeded (Huang and Ferrell, 1996; Ferrell and Machleder, 1998). Different extracellular stimuli often result in specific and qualitatively different intracellular responses, despite employing overlapping signaling pathways. It is believed that such specificity can arise, because each stimulus results in a quantitatively different activation pattern of the same signaling pathway (Purvis and Lahav, 2013; Ashe and Briscoe, 2006). However, how signaling pathways use the different dynamic activity patterns to relay signals is largely unknown and a subject of current research.

In the following we will provide an overview about three aspects of signaling dynamics: First, we will discuss how kinase cascades can amplify signals and help to make cellular decisions. Second, we will discuss different dynamic properties that are thought to relate to how kinase cascades encode information. Third, we will discuss mechanisms that kinase cascades employ to transmit signals in a robust and reliable manner (Figure 1).

Signal Amplification and Decision Making

Kinase cascades are composed of proteins that are controlled by two opposing processes that covalently modify the protein and its activity: Phosphorylation by upstream kinases and dephosphorylation by phosphatases (Goldbeter and Koshland, 1981). Such phosphorylation–dephosphorylation cycles are the basic units of kinase signaling, and the activity of each of the units represents a balance between the activity of the upstream kinase and the activity of the phosphatase.

Kinase cascades, such as the MAPK signaling pathway, can achieve remarkable signal amplification (Sauro and Kholodenko, 2004). Often, the presence of very low amounts of ligand can result already in the activation of a large amount of intracellular signaling proteins. For instance, just a few

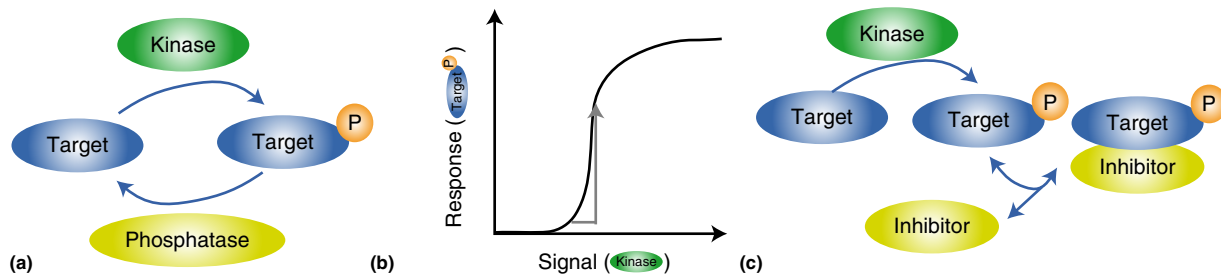


Figure 1 Basic motifs of intracellular signaling and signal amplification. (a) The phosphorylation–dephosphorylation cycle. A target protein is covalently modified by phosphorylation. The phosphorylation and dephosphorylation reactions are catalyzed by opposing kinase and phosphatase enzymes, respectively. (b) Relative signal amplification and ultrasensitivity in protein kinase signaling. Relative signal amplification implies that a small fold-change in the input level induces a large fold-change in the response (‘ultrasensitivity’). The steep dose–response curves of ultrasensitive signaling cascades may establish switch-like cellular decision making (see text). (c) Ultrasensitivity arising from stoichiometric inhibition in protein kinase signaling. The inhibitor and the phosphorylated target protein form a 1:1 complex, implying that signaling is effectively blocked for low concentrations of the phospho-target protein. For higher kinase activities, the phospho-target protein concentration may exceed the inhibitor, and the signaling activity may increase in a switch-like manner.

epidermal growth factor (EGF) molecules in the medium are sufficient to activate about 100 000 ERK molecules, the target kinase in MAPK signaling. Such signal amplification is possible as each individual kinase can phosphorylate many substrate molecules. Typically, the percentage of how much target is phosphorylated is dependent on the balance of kinase activity and phosphatase activity, independently of the amount of target. Consequently, stronger expression of the target can result in more molecules to be activated. Thus, signaling cascades can ensure that cells can respond to minor amounts of stimuli.

Apart from amplifying the absolute signal, kinase cascades can also amplify relative signals, for example, a small relative change of the ligand concentration by 10% could result in a 30% change in the activity of the terminal cascade (Huang and Ferrell, 1996). The ratio of these two relative changes, i.e., the percentage by which the pathway responds divided by the percentage in which the initial signal changes, is often termed logarithmic gain. A very high logarithmic gain corresponds to a switch-like response of the kinase pathway, because a minor change in the ligand concentration may be sufficient to switch the target kinase from nil to complete activation (Figure 1(b)). Such switch-like signaling responses, also often referred to as ‘ultrasensitivity,’ have been implicated in cellular decision making (Ferrell, 1996).

It has been noticed that ultrasensitivity can be generated even in the basic unit of kinase cascades, i.e., in a phosphorylation–dephosphorylation cycle (Figure 1(a)). Ultrasensitivity requires that the kinase and the antagonizing phosphatase (catalyzing the dephosphorylation reaction) are strongly saturated by their target substrate (Goldbeter and Koshland, 1981). This type of ultrasensitivity has been termed zero-order ultrasensitivity, because the kinase or phosphatase enzymes have to work in the saturated zero-order regime (i.e., the enzymatic velocities are independent of the substrate concentration). Zero-order ultrasensitivity has been shown to be important switch-like MAPK signaling during embryonic development (Melen *et al.*, 2005). More recent work employing mathematical arguments, however, indicated that the enzyme concentrations in most kinase cascades are incompatible with zero-order ultrasensitivity, and cast doubt

that this is a general mechanism of switch-like decision making (Blüthgen *et al.*, 2006; Ciliberto *et al.*, 2007).

Another, possibly more robust, way by which kinase cascades could generate high logarithmic gain is multisite phosphorylation, i.e., when a kinase phosphorylates its target at multiple steps (Blüthgen and Herzog, 2003; Deshaies and Ferrell, 2001). For instance, a double phosphorylation mechanism implies that the concentration of the fully modified target may increase in a quadratic manner with the enzyme concentration (i.e., the logarithmic gain equals 2), (Thomson and Gunawardena, 2009). High logarithmic gains require that the two phosphorylation events occur in a distributed rather than processive manner, which means that the kinase detaches from its target after each phosphorylation (Salazar and Höfer, 2009). In MAPK signaling, both MAP2K and MAPK have to be phosphorylated at two amino acids to become activated. Whether this phosphorylation is indeed distributed or rather progressive is currently debated. While *in vitro* data clearly shows distributed phosphorylation (Burack and Sturgill, 1997), *in vivo* data remains conflicting and unclear (Aoki *et al.*, 2011; Schilling *et al.*, 2009; Toni *et al.*, 2012). It may be that scaffold molecules change the kinetics such that distributed phosphorylation becomes more processive if both kinase and target are on the same scaffold (Witzel *et al.*, 2012; Levchenko *et al.*, 2000). By doing that, scaffolds may modulate the logarithmic gain of a pathway (Bardwell, 2008).

A third way of generating high logarithmic gain is stoichiometric inhibition, where an inhibitor and its target protein form a reversible, signaling-incompetent 1:1 complex (Ferrell, 1996). Upon weak stimulation, this inhibitor may efficiently prevent cascade activation by strongly binding to a phospho-protein intermediate (Figure 1(c)). However, the cascade activity can increase in a switch-like manner upon stronger stimulation, because the concentration of the stimulus-dependent phospho-protein may at some point exceed the amount of inhibitor. Thus, the system suddenly switches from an inhibited to a non-inhibited regime (Ferrell, 1996; Buchler and Louis, 2008; Buchler and Cross, 2009). Experimental evidence for stoichiometric inhibition in MAPK signaling cascades is currently lacking. However, in other signaling pathways this mechanism has been shown to generate

ultrasensitivity (Kim and Ferrell, 2007), and in the context of gene regulation it is established that small RNAs can act as stoichiometric inhibitors for mRNAs that establish a switch-like gene expression response (Schmiedel *et al.*, 2012; Legewie *et al.*, 2008; Mukherji *et al.*, 2011).

Interestingly, the functional organization of phosphorylation-dephosphorylation cycles in multilevel cascades can increase the ultrasensitivity relative to the isolated cycles (Ferrell, 1997). Specifically, it has been shown that the logarithmic gain along a cascade is the product of the logarithmic gain of each individual levels (Brown *et al.*, 1997; Ferrell, 1996; Kholodenko *et al.*, 1997). If, for example, each level of the 3-tier MAPK signaling cascade had a logarithmic gain of 2 due to zero-order ultrasensitivity, the entire cascade would potentially have a logarithmic gain of 8. Then, a small change in the signal of for example 5% would result in a substantial, 40% increase in the activity of their terminal kinase.

Another mechanism for amplifying the logarithmic gain is positive feedback where, for example, a downstream kinase activates a kinase more upstream in the cascade by phosphorylation. Furthermore, if the feedback is strong, the system may exhibit bistability, which means that the pathway switches between discrete levels of activity as the stimulus is increased. Bistable systems also show history-dependent behaviors such as irreversibility: For example, if the pathway was strongly stimulated in the past, the pathway may stay highly active although the external stimulus is completely removed. This way, an irreversible cell fate decision and cellular commitment can be established (Bagowski and Ferrell, 2001; Bhalla and Iyengar, 1999; Legewie *et al.*, 2006, 2007).

Cellular Decisions Encoded in the Dynamics of Signaling

Information about the extracellular milieu is transmitted to the nucleus by a limited number of signaling pathways in mammalian cells. Given this limited number of structural components, it is not surprising that most signaling pathways have been implicated in the induction of diverse cellular responses. For example, the mammalian ERK-MAPK cascade is known to control the induction of various, often contradicting, cellular responses including cell proliferation, cell migration, cell death, and cell survival (Cargnello and Roux, 2011; Keshet and Seger, 2010; Dhillon *et al.*, 2007). Cells thus face a specificity problem, and the question arises of how the same signaling pathway is able to induce very different cellular responses in different cellular systems, and depending on the extracellular stimulus applied.

Recent research revealed how the same signaling pathway can elicit cell-type specific responses. One such mechanism is the compartmentalization of signaling: For the mammalian MAPK cascade it has been shown that nuclear translocation of the terminal kinase ERK is required for the induction of proliferation (Brunet *et al.*, 1999). Some cells express cytosolic anchors that prevent ERK from entering the nucleus (Formstecher *et al.*, 2001). ERK signaling is therefore unable to elicit a proliferation response in these cells, and the signaling effects are restricted to the cytosolic substrates of ERK. Cell-type specificity may further arise at the level of gene expression

networks located downstream of signaling. Protein kinase cascade typically induce covalent modifications in nuclear transcription factors, which in turn bind with cofactors to chromatin to drive gene expression. It was recently shown that cell-type specific gene expression responses to TGF-beta stimulation arise, because each cell-type expresses a specific set of cofactors that direct the common SMAD transcription factors to distinct sites in the genome (Mullen *et al.*, 2011).

Signaling specificity may also be achieved if the quantitative dynamics of signaling carry information about the cell fate. For instance, it was shown that the cytokine TNF-alpha activates multiple intracellular signaling pathways in parallel, and that the relative activation levels of these pathways determine how different cell types respond to the common stimulus (Miller-Jensen *et al.*, 2007). Likewise, it is also possible that the temporal activation profile of a single protein kinase cascade determines a cellular response. A classical example is the response of PC12 neuronal precursor cells to growth factor stimulation. These cells differentiate into mature neurons upon stimulation with neurotrophic growth factor (NGF), but divide if they are treated with EGF. Both ligands activate a broadly overlapping set of signaling molecules. However, NGF triggers a sustained activation of the ERK pathway, while EGF induces a transient ERK signaling. This finding led to the hypothesis signaling specificity is encoded in the duration of the ERK signal (Marshall, 1995). Experimental perturbations which affect the dynamics of ERK signaling support this claim: For example, a simple overexpression of EGF receptors prolongs the EGF-induced ERK signal, and switches the cellular response to EGF to differentiation (Traverse *et al.*, 1994).

The question arises of how two ligands like EGF and NGF that engage the same signaling pathway can elicit signaling with transient vs. sustained signaling in the same cell type. A mechanism that can convert a constant EGF stimulus into a transient signal is negative feedback (Cirit *et al.*, 2010), as depicted in Figure 2(a): the activity of the terminal kinase ERK promotes in the deactivation of the uppermost cascade member Raf. Thus, the system initially shows strong activation until the feedback inhibition results in downregulation of the signal (Figure 2(b), solid line). Interestingly, the feedback system can also show a switch from transient to sustained signaling if the stimulus strength is increased; for a sufficiently high stimulus, the feedback may no longer be strong enough to counteract and downregulate the cascade activity (Figure 2(b), dashed line; Behar *et al.*, 2007). This leads to a model where the number of cell surface receptors that can be activated by a ligand determines the duration of the downstream signaling, as supported by experimental studies (Traverse *et al.*, 1994).

Early theoretical work has suggested that differential negative feedback regulation by EGF and NGF may result in different signal duration and thus be the mechanism by which ligand identity is encoded (Brightman and Fell, 2000). A more recent experimental study further supported this hypothesis: Santos *et al.* (2007) systematically perturbed the MAPK cascade in PC12 cells by siRNAs against rapidly accelerated fibrosarcoma (Raf), MAP kinase or ERK kinase (MEK), and ERK, and used a modeling framework known as modular response analysis (Kholodenko *et al.*, 2002), to conclude that

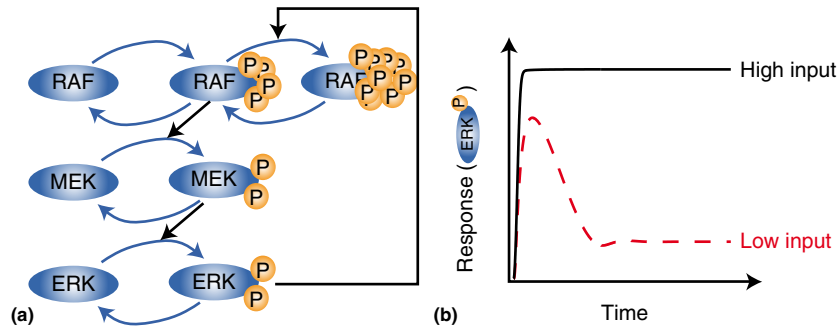


Figure 2 MAPK signaling and encoding of information in the quantitative aspects of intracellular signals. (a) Negative feedback regulation in MAPK signaling. The core MAPK cascade consists of three kinases (Raf, MEK, and ERK) which are sequentially activated by phosphorylation on two or more residues. Negative feedback regulation arises, as the terminal kinase ERK triggers the hyperphosphorylation and thereby inactivation of Raf. (b) Negative feedback regulation can establish a switch from transient to sustained signaling with increasing input. Numerical simulations were performed for a simple negative feedback MAPK cascade (similar to panel A) for a low and a high concentration of the kinase catalyzing the phosphorylation of Raf ('low input' and 'high input,' respectively). The ordinary differential equation model is described in Behar *et al.* (2007) (Model IIIB; parameters values as described but $k_5=8e-1$).

EGF specifically induces negative feedback regulation in MAPK signaling. Furthermore, the authors showed that the MAPK signal upon NGF stimulation is amplified by a stimulus-specific positive feedback loop, which establishes a bistable and irreversible cell fate decision. However, other mechanisms are likely to contribute to the stimulus-specific signal duration as well. For example, it has been suggested that NGF receptors specifically engage the FRS adapter protein that shows sustained activation kinetics to trigger MAPK activation (Sasagawa *et al.*, 2005; Yamada *et al.*, 2004; York *et al.*, 1998). Sustained signaling via FRS may emerge, because this signaling protein continues to be active in the absence of active cell surface receptors, whereas adapter proteins recruited by the EGF receptor require ongoing receptor activity (Yamada *et al.*, 2004; York *et al.*, 1998).

If the stimulus-specific signal duration in PC12 cells underlies the cell fate decision between differentiation and proliferation, the cells must also be able to reliably distinguish (or decode) short and long MAPK signals. Duration decoding at the level of MAPK-induced gene expression has been analyzed in NIH3T3 cells, which behave different from PC12 cells, because they require a sustained MAPK signal for cell proliferation. Murphy *et al.* (2002) analyzed the immediate early gene c-Fos whose transcription is strongly induced very early after MAPK activation. The c-Fos protein functions as a transcription factor and is essential for triggering the expression proteins driving cell division. Murphy *et al.* found that Fos protein accumulation requires a sustained MAPK signal. This selective induction by a long MAPK signal could be explained by a phosphorylation-dependent stabilization of the Fos protein: The Fos protein is very unstable under normal conditions (half-life: 15 mins), but can be stabilized by ERK-dependent phosphorylation (half-life 4 h). Thus, Fos expression not only depends on early MAPK signaling (transcriptional induction), but also senses the late phase of ERK activity (posttranslational stabilization). Such a dual regulation of a target by a stimulus is known as a feed-forward loop and it has been suggested that these loops may generally allow gene expression networks to distinguish input of different durations (Shen-Orr *et al.*, 2002). The c-Fos network was recently analyzed in more detail by

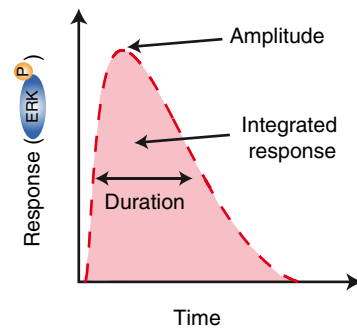


Figure 3 Signaling features that determine cellular responses. Cell fate decision can be controlled by the duration of an intracellular signal, by its amplitude, or by the integrated response (also known as area under curve).

Kholodenko and colleagues using an experiment-theory approach (Nakakuki *et al.*, 2010). These authors concluded that the performance and robustness of duration decoding can be improved if the core feed-forward architecture is embedded in a more complex network, comprising additional feedback and feed-forward mechanisms. A recent study in NIH3T3 cells further suggested that these initial duration decoding mechanisms may be amplified further downstream in the gene expression network, because a long MAPK signal specifically triggered autocrine cytokine stimulation and a qualitatively different gene expression program (Toettcher *et al.*, 2013).

Information about the cellular fate may also be encoded in signaling features other than the signal duration (schematically depicted in Figure 3). A decisive feature in many signaling systems is the signal amplitude (Figure 3). A prominent example of amplitude encoding are spatial gradients of extracellular ligands ('morphogens') that control embryonic patterning: Along the gradient, these morphogens result in increasing activation of intracellular protein kinase cascades, and the signaling amplitude specifies into which cell type the tissue differentiates (Ashe and Briscoe, 2006). Another important signaling feature is the integrated response, also known as the area under curve (Figure 3; Schilling *et al.*, 2009;

Asthagiri *et al.*, 2000). For example, the proliferation of CFU-E cells in response to Epo stimulation is determined by the integrated response of the MAPK signal (Schilling *et al.*, 2009). In some cellular systems like NF- κ B, MAPK and Ca²⁺ pathways, it seems that the oscillation frequency of a signal encodes specific gene expression responses and cell fate responses (Albeck *et al.*, 2013; Purvis *et al.*, 2012; Hoffmann *et al.*, 2002; Dolmetsch *et al.*, 1998). Taken together, information that the cell uses for cellular decision appears to be encoded in a wide variety of signaling features (Purvis and Lahav, 2013). Future studies are required to better understand how quantitative signals can be robustly encoded and decoded by intracellular regulatory networks. To this end, experimental tools need to be developed to specifically assess the activities of kinases in subcellular compartments, because the signaling dynamics can be misinterpreted based on population-average measurements as measured in whole-cell lysates (Ahmed *et al.*, 2014).

Robustness of Signaling Pathways

For many signaling systems, it is essential to reliably and robustly transmit small changes in the input signal (Stelling *et al.*, 2004). One example for such a quantitative mode of signal transmission is the patterning process during animal development (Ashe and Briscoe, 2006) (see above). Such quantitative signal transmission is rather challenging, as signaling pathways have to deal with high amount of 'noise.' One of the main problems seems to be cell-to-cell variability in signaling protein expression levels; it has been observed that even clonal cells under the same condition express different amounts of each protein. Typically, cells expressing high levels of a given protein (95% percentile of the cell population) have an approximately threefold higher concentration of that protein when compared to the least expressing cells (5% percentile) (Cohen-Saidon *et al.*, 2009). This protein concentration variability is largely a consequence of so-called gene expression noise, which arises directly or indirectly from the stochastic nature of molecular processes involving small molecule number, such as in transcription and translation (Raj and van Oudenaarden, 2008; Elowitz *et al.*, 2002).

Several recent single-cell studies confirmed that, in many cases, protein expression noise is a major source of variability in intracellular signal transduction networks (Gaudet *et al.*, 2012; Yuan *et al.*, 2011; Spencer *et al.*, 2009; Feinerman *et al.*, 2008; Kollmann *et al.*, 2005). Theoretical considerations suggest that generic mathematical models of protein kinase cascades show strong variability of the signaling output if realistic protein concentration fluctuations are assumed (Jeschke *et al.*, 2013). Notable exceptions are ultrasensitive protein kinase cascades which can be intrinsically invariant despite varying protein concentrations in certain parameter regimes (Jeschke *et al.*, 2013). If these intrinsic robustness requirements are not met, a protein kinase cascade needs to be equipped with additional robustness mechanisms (like feedback or feed-forward loops) to allow for quantitative information transmission.

One simple strategy of noise suppression is the co-regulation of kinase and phosphatase expression (Feinerman *et al.*, 2008; Kollmann *et al.*, 2005). Typically, the activation level of a target

in a protein kinase cascade is dependent on the balance of kinase and phosphatase activities. Thus, the variability in the activation level can be buffered by making the fluctuations of the kinase and phosphatase expression levels correlated. Such a correlation implies that a cell with high kinase activity will also have a high activity of the antagonizing phosphatase, thereby compensating for imbalances in target activation. This principle has been shown to operate in the mammalian ERK cascade in T cells, where variability is suppressed, because the fluctuations of the CD8 receptor are highly correlated with fluctuations in the antagonizing phosphatase, SHP-1 (Feinerman *et al.*, 2008). At the gene expression level, correlated fluctuations may arise, because the kinase and the phosphatase expression levels are coordinately controlled by the same transcription factor. In fact, noise suppression by co-regulation of signaling proteins may be a wide-spread phenomenon: many eukaryotic signaling pathways are organized in so-called synexpression groups, that is, most of their components show tight spatio-temporal co-expression in different tissues of the body (Niehrs and Pollet, 1999).

Based on theoretical considerations, it has been speculated for some time that negative feedbacks could also help to buffer such variability in gene expression (Sauro and Kholodenko, 2004). Two recent papers used a combination of experiment and theory to show that negative feedbacks suppress the variability of MAPK and SMAD signaling (Paulsen *et al.*, 2011; Fritsche-Guenther *et al.*, 2011). For the case of MAPK signaling, it was shown that a strong negative feedback from ERK to RAF gives rise to concentration robustness: If MAPK/ERK expression levels were lowered by siRNA-mediated knock-down, ERK phosphorylation in colon cancer cells stayed approximately constant. Only if ERK expression dropped below ~50% of the original value, phospho-ERK decreased, implying that the robustness could no longer be maintained. Robustness in this pathway could be attributed to a feedback: When ERK protein concentration was lowered, activity of the upstream kinase MEK increased because negative feedback at the level of RAF was reduced. In cells with constitutively active RAF which also had disabled feedback, concentration robustness was lost (Fritsche-Guenther *et al.*, 2011).

Besides reducing the impact of protein concentration fluctuations, strong negative feedback also allows protein kinase cascades to more gradually respond to changes of the extracellular stimulus concentration (Jeschke *et al.*, 2013; Paulsen *et al.*, 2011; Sturm *et al.*, 2010). This implies that a negative feedback cascade is able to sense and to reliably transmit a larger range of extracellular stimulus concentrations when compared to a feedback-less cascade. Taken together, negative feedback optimizes the quantitative mode of information transmission in two ways: first, by allowing a more reliable transmission of a given input signal and, second, by extending the range of input signals that can be sensed.

Negative feedbacks can also have dramatic consequences on the efficiency of drugs to block the activity of signaling pathways. Particularly strong negative feedbacks, such as the ERK-RAF feedback tend to render the pathway robust against inhibition (Sturm *et al.*, 2010). It has for example been shown that MEK inhibitors are relatively inefficient, as strong feedback from ERK to RAF hyperactivates RAF and MEK when MEK is inhibited (Friday *et al.*, 2008; Fritsche-Guenther *et al.*, 2011).

Another important aspect of negative feedbacks in the context of targeted kinase inhibitors is that they can activate parallel compensatory pathways and can thereby render cells resistant against the pathway inhibitor. One example is the EGF receptor-RAS-RAF-MEK-ERK pathway in colon cells: The loss of a negative feedback from ERK to the EGF receptor upon MEK or RAF inhibition hyperactivates the EGFR. This leads to activation of a second kinase cascade downstream of the EGF receptor, the PI3K/AKT signaling pathway, which then causes resistance against RAF and MEK inhibition (Klinger *et al.*, 2013; Prahallad *et al.*, 2012). A similar mechanism mediates resistance in metastatic breast cancer. Here, a negative feedback from mTOR to IRS1 leads to activation of the JAK/STAT signaling pathway, when blocking the PI3K/AKT/mTOR cascade (Britschgi *et al.*, 2012). In both examples, co-inhibition of the parallel pathway was sufficient to re-sensitize the cells toward the inhibitor.

In summary, many if not most kinase cascades show negative feedback regulation, which likely evolved to provide signaling robustness and homeostasis; however, in the context of targeted therapies, such feedbacks were shown to mediate drug resistance.

See also: Dynamic Integration: Dynamics: Dynamics of Gradient Sensing and Chemotaxis. Cell Communication: Intracellular Pathways: The MAPK Signaling Cascades. Cell Communication: Growth Factor Mediated Cell Signaling: Receptor Tyrosine Kinases and Their Ligands. Horizontal Integration: Omics: Metabolomics in Cell Biology

References

- Ahmed, S., Grant, K.G., Edwards, L.E., *et al.*, 2014. Data-driven modeling reconciles kinetics of ERK phosphorylation, localization, and activity states. *Molecular Systems Biology* 10, 718.
- Albeck, J.G., Mills, G.B., Brugge, J.S., 2013. Frequency-modulated pulses of ERK activity transmit quantitative proliferation signals. *Molecular Cell* 49 (2), 249–261.
- Aoki, K., Yamada, M., Kuniida, K., Yasuda, S., Matsuda, M., 2011. Processive phosphorylation of ERK MAP kinase in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* 108 (31), 12675–12680.
- Ashe, H.L., Briscoe, J., 2006. The interpretation of morphogen gradients. *Development (Cambridge, England)* 133 (3), 385–394.
- Asthaigiri, A.R., Reinhart, C.A., Horwitz, A.F., Lauffenburger, D.A., 2000. The role of transient ERK2 signals in fibronectin- and insulin-mediated DNA synthesis. *Journal of Cell Science* 113 (Pt 24), 4499–4510.
- Bagowski, C.P., Ferrell, J.E., 2001. Bistability in the JNK cascade. *Current Biology: CB* 11 (15), 1176–1182.
- Bardwell, L., 2008. Signal transduction: Turning a switch into a rheostat. *Current Biology: CB* 18 (19), R910–R912.
- Behar, M., Hao, N., Dohlman, H.G., Elston, T.C., 2007. Mathematical and computational analysis of adaptation via feedback inhibition in signal transduction pathways. *Biophysical Journal* 93 (3), 806–821.
- Bhalla, U.S., Iyengar, R., 1999. Emergent properties of networks of biological signaling pathways. *Science (New York, NY)* 283 (5400), 381–387.
- Blüthgen, N., Bruggeman, F.J., Legewie, S., *et al.*, 2006. Effects of sequestration on signal transduction cascades. *FEBS Journal* 273 (5), 895–906.
- Blüthgen, N., Herzog, H., 2003. How robust are switches in intracellular signaling cascades? *Journal of Theoretical Biology* 225 (3), 293–300.
- Brightman, F.A., Fell, D.A., 2000. Differential feedback regulation of the MAPK cascade underlies the quantitative differences in EGF and NGF signaling in PC12 cells. *FEBS Letters* 482 (3), 169–174.
- Britschgi, A., Andraos, R., Brinkhaus, H., *et al.*, 2012. JAK2/STAT5 inhibition circumvents resistance to PI3K/mTOR blockade: A rationale for cotargeting these pathways in metastatic breast cancer. *Cancer Cell* 22, 796–811.
- Brown, G.C., Hoek, J.B., Kholodenko, B.N., 1997. Why do protein kinase cascades have more than one level? *Trends in Biochemical Sciences* 22 (8), 288.
- Brunet, A., Roux, D., Lenormand, P., *et al.*, 1999. Nuclear translocation of p42/p44 mitogen-activated protein kinase is required for growth factor-induced gene expression and cell cycle entry. *EMBO Journal* 18 (3), 664–674.
- Buchler, N.E., Cross, F.R., 2009. Protein sequestration generates a flexible ultrasensitive response in a genetic network. *Molecular Systems Biology* 5, 272.
- Buchler, N.E., Louis, M., 2008. Molecular titration and ultrasensitivity in regulatory networks. *Journal of Molecular Biology* 384 (5), 1106–1119.
- Burack, W.R., Sturgill, T.W., 1997. The activating dual phosphorylation of MAPK by MEK is nonprocessive. *Biochemistry* 36 (20), 5929–5933.
- Cargnello, M., Roux, P.P., 2011. Activation and function of the MAPKs and their substrates, the MAPK-activated protein kinases. *Microbiology and Molecular Biology Reviews: MMBR* 75 (1), 50–83.
- Ciliberto, A., Capuani, F., Tyson, J.J., 2007. Modeling networks of coupled enzymatic reactions using the total quasi-steady state approximation. *PLoS Computational Biology* 3 (3), e45.
- Cirit, M., Wang, C.-C., Haugh, J.M., 2010. Systematic quantification of negative feedback mechanisms in the extracellular signal-regulated kinase (ERK) signaling network. *Journal of Biological Chemistry* 285 (47), 36736–36744.
- Cohen-Saidon, C., Cohen, A.A., Sigal, A., Alon, U., 2009. Dynamics and variability of ERK2 response to EGF in individual living cells. *Molecular Cell* 36 (5), 885–893.
- Colman-Lerner, A., Gordon, A., Serra, E., *et al.*, 2005. Regulated cell-to-cell variation in a cell-fate decision system. *Nature* 437 (7059), 699–706.
- Deshaies, R.J., Ferrell Jr, J.E., 2001. Multisite phosphorylation and the countdown to S phase. *Cell* 107 (7), 819–822.
- Dhillon, A.S., Hagan, S., Rath, O., *et al.*, 2007. MAP kinase signaling pathways in cancer. *Oncogene* 26 (22), 3279–3290.
- Dolmetsch, R.E., Xu, K., Lewis, R.S., 1998. Calcium oscillations increase the efficiency and specificity of gene expression. *Nature* 392 (6679), 933–936.
- Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S., 2002. Stochastic gene expression in a single cell. *Science (New York, NY)* 297 (5584), 1183–1186.
- Feinerman, O., Veiga, J., Dorfman, J.R., Germain, R.N., Altan-Bonnet, G., 2008. Variability and robustness in T cell activation from regulated heterogeneity in protein levels. *Science (New York, NY)* 321 (5892), 1081–1084.
- Ferrell, J.E., 1996. Tripping the switch fantastic: How a protein kinase cascade can convert graded inputs into switch-like outputs. *Trends in Biochemical Sciences* 21 (12), 460–466.
- Ferrell, J.E., 1997. How responses get more switch-like as you move down a protein kinase cascade. *Trends in Biochemical Sciences* 22 (8), 288–289.
- Ferrell, J.E., Machleder, E.M., 1998. The biochemical basis of an all-or-none cell fate switch in *Xenopus* oocytes. *Science (New York, NY)* 280 (5365), 895–898.
- Formstecher, E., Ramos, J.W., Fauquet, M., *et al.*, 2001. PEA-15 mediates cytoplasmic sequestration of ERK MAP kinase. *Developmental Cell* 1 (2), 239–250.
- Friday, B.B., Yu, C., Dy, G.K., *et al.*, 2008. BRAF V600E disrupts AZD6244-induced abrogation of negative feedback pathways between extracellular signal-regulated kinase and Raf proteins. *Cancer Research* 68 (15), 6145–6153.
- Fritsche-Guenther, R., Witzel, F., Sieber, A., *et al.*, 2011. Strong negative feedback from Erk to Raf confers robustness to MAPK signaling. *Molecular Systems Biology* 7, 489.
- Gaudet, S., Spencer, S.L., Chen, W.W., Sorger, P.K., 2012. Exploring the contextual sensitivity of factors that determine cell-to-cell variability in receptor-mediated apoptosis. *PLoS Computational Biology* 8 (4), e1002482.
- Goldbeter, A., Koshland, D.E., 1981. An amplified sensitivity arising from covalent modification in biological systems. *Proceedings of the National Academy of Sciences of the United States of America* 78 (11), 6840–6844.
- Hoffmann, A., Levchenko, A., Scott, M.L., Baltimore, D., 2002. The I κ B α -NF- κ B signaling module: Temporal control and selective gene activation. *Science (New York, NY)* 298 (5596), 1241–1245.
- Huang, C.Y., Ferrell, J.E., 1996. Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proceedings of the National Academy of Sciences of the United States of America* 93 (19), 10078–10083.
- Jeschke, M., Baumgärtner, S., Legewie, S., 2013. Determinants of cell-to-cell variability in protein kinase signaling. *PLoS Computational Biology* 9 (12), e1003357.
- Keshet, Y., Seger, R., 2010. The MAP kinase signaling cascades: A system of hundreds of components regulates a diverse array of physiological functions. *Methods in Molecular Biology (Clifton, NJ)* 661, 3–38.
- Kholodenko, B.N., Hancock, J.F., Kolch, W., 2010. Signaling ballet in space and time. *Nature Reviews Molecular Cell Biology* 11 (6), 414–426.

- Kholodenko, B.N., Hoek, J.B., Westerhoff, H.V., Brown, G.C., 1997. Quantification of information transfer via cellular signal transduction pathways. *FEBS Letters* 414 (2), 430–434.
- Kholodenko, B.N., Kiyatkin, A., Bruggeman, F.J., *et al.*, 2002. Untangling the wires: A strategy to trace functional interactions in signaling and gene networks. *Proceedings of the National Academy of Sciences of the United States of America* 99 (20), 12841–12846.
- Kiel, C., Serrano, L., 2012. Challenges ahead in signal transduction: MAPK as an example. *Current Opinion in Biotechnology* 23 (3), 305–314.
- Kim, S.Y., Ferrell, J.E., 2007. Substrate competition as a source of ultrasensitivity in the inactivation of Wee1. *Cell* 128 (6), 1133–1145.
- Klinger, B., Sieber, A., Fritsche-Guenther, R., *et al.*, 2013. Network quantification of EGFR signaling unveils potential for targeted combination therapy. *Molecular Systems Biology* 9, 673.
- Kollmann, M., Løvdok, L., Bartholomé, K., Timmer, J., Sourjik, V., 2005. Design principles of a bacterial signaling network. *Nature* 438 (7067), 504–507.
- Legewie, S., Blüthgen, N., Herzog, H., 2006. Mathematical modeling identifies inhibitors of apoptosis as mediators of positive feedback and bistability. *PLoS Computational Biology* 2 (9), e120.
- Legewie, S., Dienst, D., Wilde, A., Herzog, H., Axmann, I.M., 2008. Small RNAs establish delays and temporal thresholds in gene expression. *Biophysical Journal* 95 (7), 3232–3238.
- Legewie, S., Schoeberl, B., Blüthgen, N., Herzog, H., 2007. Competing docking interactions can bring about bistability in the MAPK cascade. *Biophysical Journal* 93 (7), 2279–2288.
- Levchenko, A., Bruck, J., Sternberg, P.W., 2000. Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties. *Proceedings of the National Academy of Sciences of the United States of America* 97 (11), 5818–5823.
- Marshall, C.J., 1995. Specificity of receptor tyrosine kinase signaling: Transient versus sustained extracellular signal-regulated kinase activation. *Cell* 80 (2), 179–185.
- Melen, G.J., Levy, S., Barkai, N., Shilo, B.Z., 2005. Threshold responses to morphogen gradients by zero-order ultrasensitivity. *Molecular Systems Biology* 1, 2005.0028.
- Miller-Jensen, K., Janes, K.A., Brugge, J.S., Lauffenburger, D.A., 2007. Common effector processing mediates cell-specific responses to stimuli. *Nature* 448 (7153), 604–608.
- Mukherji, S., Ebert, M.S., Zheng, G.X., 2011. MicroRNAs can generate thresholds in target gene expression. *Nature Genetics* 43 (9), 854–859.
- Mullen, A.C., Orlando, D.A., Newman, J.J., 2011. Master transcription factors determine cell-type-specific responses to TGF- β signaling. *Cell* 147 (3), 565–576.
- Murphy, L.O., Smith, S., Chen, R.H., *et al.*, 2002. Molecular interpretation of ERK signal duration by immediate early gene products. *Nature Cell Biology* 4 (8), 556–564.
- Nakakuki, T., Birtwistle, M.R., Saeki, Y., *et al.*, 2010. Ligand-specific c-Fos expression emerges from the spatiotemporal control of ErbB network dynamics. *Cell* 141 (5), 884–896.
- Nelson, D.E., Ihekweba, A.E., Elliott, M., *et al.*, 2004. Oscillations in NF- κ B signaling control the dynamics of gene expression. *Science (New York, NY)* 306 (5696), 704–708.
- Niehrs, C., Pollet, N., 1999. Synexpression groups in eukaryotes. *Nature* 402 (6761), 483–487.
- Paulsen, M., Legewie, S., Eils, R., *et al.*, 2011. Negative feedback in the bone morphogenetic protein 4 (BMP4) synexpression group governs its dynamic signaling range and canalizes development. *Proceedings of the National Academy of Sciences of the United States of America* 108 (25), 10202–10207.
- Prahalad, A., Sun, C., Huang, S., *et al.*, 2012. Unresponsiveness of colon cancer to BRAF(V600E) inhibition through feedback activation of EGFR. *Nature* 483 (7387), 100–103.
- Purvis, J.E., Karhohs, K.W., Mock, C., *et al.*, 2012. p53 dynamics control cell fate. *Science (New York, NY)* 336 (6087), 1440–1444.
- Purvis, J.E., Lahav, G., 2013. Encoding and decoding cellular information through signaling dynamics. *Cell* 152 (5), 945–956.
- Raj, A., van Oudenaarden, A., 2008. Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell* 135 (2), 216–226.
- Salazar, C., Höfer, T., 2009. Multisite protein phosphorylation—from molecular mechanisms to kinetic models. *FEBS Journal* 276 (12), 3177–3198.
- Santos, S.D.M., Verveer, P.J., Bastiaens, P.I.H., 2007. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. *Nature Cell Biology* 9 (3), 324–330.
- Sasagawa, S., Ozaki, Y., Fujita, K., Kuroda, S., 2005. Prediction and validation of the distinct dynamics of transient and sustained ERK activation. *Nature Cell Biology* 7 (4), 365–373.
- Sauro, H.M., Kholodenko, B.N., 2004. Quantitative analysis of signaling networks. *Progress in Biophysics and Molecular Biology* 86 (1), 5–43.
- Schilling, M., Maiwald, T., Hengl, S., *et al.*, 2009. Theoretical and experimental analysis links isoform-specific ERK signaling to cell fate decisions. *Molecular Systems Biology* 5, 334.
- Schmiedel, J.M., Axmann, I.M., Legewie, S., 2012. Multi-target regulation by small RNAs synchronizes gene expression thresholds and may enhance ultrasensitive behavior. *PLoS One* 7 (8), e42296.
- Seger, R., Krebs, E.G., 1995. The MAPK signaling cascade. *FASEB Journal* 9 (9), 726–735.
- Shankaran, H., Ippolito, D.L., Chrisler, W.B., *et al.*, 2009. Rapid and sustained nuclear-cytoplasmic ERK oscillations induced by epidermal growth factor. *Molecular Systems Biology* 5, 332.
- Shen-Orr, S.S., Milo, R., Mangan, S., Alon, U., 2002. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genetics* 31 (1), 64–68.
- Spencer, S.L., Gaudet, S., Albeck, J.G., *et al.*, 2009. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* 459 (7245), 428–432.
- Stelling, J., Sauer, U., Szallasi, Z., *et al.*, 2004. Robustness of cellular functions. *Cell* 118 (6), 675–685.
- Sturm, O.E., Orton, R., Grindlay, J., *et al.*, 2010. The mammalian MAPK/ERK pathway exhibits properties of a negative feedback amplifier. *Science Signaling* 3 (153), ra90.
- Thomson, M., Gunawardena, J., 2009. Unlimited multistability in multisite phosphorylation systems. *Nature* 460 (7252), 274–277.
- Toettcher, J.E., Weiner, O.D., Lim, W.A., 2013. Using optogenetics to interrogate the dynamic control of signal transmission by the Ras/Erk module. *Cell* 155 (6), 1422–1434.
- Toni, T., Ozaki, Y., Kirk, P., Kuroda, S., Stumpf, M.P., 2012. Elucidating the *in vivo* phosphorylation dynamics of the ERK MAP kinase using quantitative proteomics data and Bayesian model selection. *Molecular BioSystems* 8 (7), 1921–1929.
- Traverse, S., Seedorf, K., Paterson, H., *et al.*, 1994. EGF triggers neuronal differentiation of PC12 cells that overexpress the EGF receptor. *Current Biology: CB* 4 (8), 694–701.
- Witzel, F., Maddison, L., Blüthgen, N., 2012. How scaffolds shape MAPK signaling: What we know and opportunities for systems approaches. *Frontiers in Physiology* 3, 475.
- Yamada, S., Taketomi, T., Yoshimura, A., 2004. Model analysis of difference between EGF pathway and FGF pathway. *Biochemical and Biophysical Research Communications* 314 (4), 1113–1120.
- York, R.D., Yao, H., Dillon, T., *et al.*, 1998. Rap1 mediates sustained MAP kinase activation induced by nerve growth factor. *Nature* 392 (6676), 622–626.
- Yuan, T.L., Wulf, G., Burga, L., *et al.*, 2011. Cell-to-cell variability in PI3K protein level regulates PI3K-AKT pathway activity in cell populations. *Current Biology: CB* 21 (3), 173–183.

Further Reading

- Blüthgen, N., Legewie, S., 2008. Systems analysis of MAPK signal transduction. *Essays in Biochemistry* 45, 95–107.
- Kholodenko, B.N., Hancock, J.F., Kolch, W., 2010. Signaling ballet in space and time. *Nature Reviews Molecular Cell Biology* 11, 414–426.
- Tyson, J.J., Chen, K.C., Novak, B., 2003. Sniffers, buzzers, toggles and blinkers: Dynamics of regulatory and signaling pathways in the cell. *Current Opinion in Cell Biology* 15, 221–231.

Cdc42 and the Mechanisms of Yeast Cell Polarization – A Paradigm for Mesoscale Systems Biology

R Wedlich-Soldner, University of Muenster, Muenster, Germany

© 2016 Elsevier Inc. All rights reserved.

Introduction

Establishment of cell polarity is a fundamental process that is prerequisite for cell morphogenesis, migration, division, and differentiation. In most organisms studied to date, small GTPases play a central role in the regulation of polarity establishment. In bacteria, Ras-like proteins regulate accumulation of proteins at the cell poles and mediate bacterial motility (Keilberg and Sogaard-Andersen, 2014). In plant cells, members of the ROP (Rho-GTPases from plants) family drive polarization of tip-growing cells such as pollen tube or root hairs (Yang and Lavagi, 2012). In animal and fungal cells, The Rho-type GTPase Cdc42 is a highly conserved key regulator for the establishment of cell polarity (Harris and Tepass, 2010). The budding yeast *Saccharomyces cerevisiae* has proven to be an invaluable model system to unravel the molecular components and mechanisms underlying Cdc42-mediated polarity establishment. Early studies have identified many of the molecular components involved in cell polarization and have led to the proposal of a linear signaling cascade where stable cortical landmarks are interpreted through a series of molecular interactions that ultimately lead to local activation of Cdc42 and subsequent actin-mediated polarized growth (Drubin and Nelson, 1996). These studies were mostly based on genetic screens and the study of protein–protein interactions. In contrast, work in the last decade using a powerful combination of genetics, quantitative live cell imaging and theoretical modeling has revealed a complex system of interconnected feedback circuits that lie at the core of Cdc42-dependent symmetry breaking (Slaughter *et al.*, 2009b). In the course of deconstructing and manipulating these circuits, a considerable body of work on yeast polarity has accumulated that has since become a prime example of what has been poignantly defined as ‘mesoscale systems biology’ (Kirschner, 2005). The goal of this article is to summarize the key concepts in spatiotemporal cell organization that have emerged from work on yeast cell polarity regulation, with particular attention to the role and dynamic localization of Cdc42.

Cue-Dependent Cell Polarization

Saccharomyces cerevisiae cells divide through budding, a form of asexual reproduction in which a daughter cell or bud develops through localized growth and cell division. In haploid cells, new bud sites are always chosen adjacent to the previous site of division, leading to an axial budding pattern. In contrast, diploid yeast cells bud in a bipolar fashion. Importantly, cells division sites become permanently marked upon mother-bud separation and can be identified as birth scars in fresh daughter cells or bud scars in mother cells. Bud scars are characterized by a rigid, chitin-rich cell wall and several

integral plasma membrane (PM) proteins (Kang *et al.*, 2004). Classical genetic studies have identified two distinct pathways used by yeast cells to interpret bud scars as landmarks for cell polarization (Casamayor and Snyder, 2002). In both pathways, specific transmembrane proteins (Bud10 in haploid cells, Rax1 and Rax2 in diploid cells) are stably anchored to the cell wall at bud scars. These integral membrane proteins then interact with soluble cofactors (Bud3 and Bud4 for haploid cells, Bud8 and Bud9 for diploid cells) to recruit the guanine nucleotide exchange factor (GEF) Bud5. Bud5 recruits and activates the Ras-type GTPase Rsr1 (=Bud1), and the corresponding GTPase activating protein (GAP) Bud2 returns Rsr1 to its inactive GDP-bound form. Importantly, active Rsr1 together with the scaffold protein Bem1 can in turn recruit Cdc24, the only GEF for the central polarity regulator Cdc42. With the help of its GEF and several GAPs (Bem2/3, Rga1/2), Cdc42 exhibits the rapid activity cycles typical for small GTPases (Figure 1) and through several effector molecules (the PAK kinases Cla4 and Ste20, Gic1/2, the exocyst subunit Sec3 and the formins Bni1 and Bnr1) regulates downstream processes, including organization of the actin cytoskeleton, assembly of the septin ring at the bud neck and fusion of exocytic vesicles.

In combination these events ultimately drive polarized delivery of material and localized growth. The large number of molecular components involved in interpreting the spatial landmarks and establishing cell polarity can be simplified to a linear series of interactions between modules (Figure 2). The landmark is first interpreted by the bud site selection module (with Rsr1 at its center), which recruits and activates the Cdc42 module. Such series of connected GTPase modules can also be found in membrane transport processes where Arf and Rab GTPase are often organized in cascades (Jin *et al.*, 2011; Chesneau *et al.*, 2012). In its GTP-bound form, Cdc42 activates several effectors that finally lead to the polarized orientation of the actin cytoskeleton and oriented growth.

Haploid yeast cells exhibit a second type of polarization mode that occurs when cells of opposite mating type grow

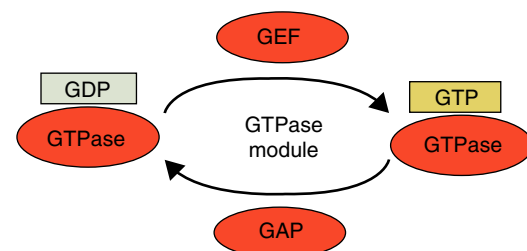


Figure 1 Schematic of the GTPase cycle. GDP release and activation of the GTPase is mediated by guanine exchange factors (GEF), whereas GTP hydrolysis and inactivation of the GTPase is catalyzed by GTPase activating proteins (GAP).

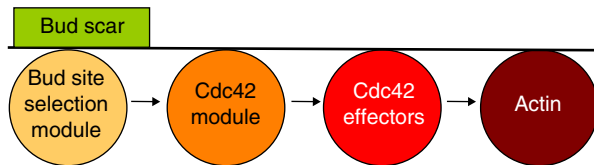


Figure 2 Modular view of cue-dependent cell polarization. In this simplified linear view of polarization during yeast budding, each module (circle) encompasses several components and biochemical activities.

toward each other. The resulting pear-shaped shmoo fuse at their tips to form a diploid cell that then undergoes meiosis. The cue for mating cells is a gradient of secreted peptide pheromone that is recognized by G-protein coupled receptors on cells of the opposing mating type. As a consequence the associated trimeric G-protein dissociates and the $G\beta\gamma$ subunit recruits a Far1-Cdc24 complex to the cortex, where the GEF can locally activate Cdc42 and hence orient actin-based transport (Arkowitz and Bassilana, 2011). So while the initial cue and cue interpretation are distinct during budding and mating, the later steps of Cdc24 recruitment, Cdc42 activation and cytoskeleton polarization are conserved. The intersection of both pathways occurs on the level of the Cdc42 GEF. In addition temporal regulation of polarization also occurs through Cdc42 regulators. Cdk activity at the G1/S transition of the cell cycle leads to degradation of Far1, which keeps Cdc24 sequestered in the nucleus during G1 phase. This allows Cdc24 to move into the cytosol from where it can subsequently be recruited to the cell cortex via Rsr1 or Bem1 (Nern and Arkowitz, 2000). In addition, several Cdc42 GAPs have been shown to be phosphorylated by Cdk and thereby inactivated.

The example of cue-mediated polarization of yeast cells nicely illustrates a frequent phenomenon in cellular regulatory networks, where highly conserved core processes or core components (here the Cdc42 module) are linked to varying inputs through flexible linkages such as different GEFs/GAPs, inhibitors/sequestering factors or scaffolds. These ‘weak’ linkages together with the modularity of regulatory pathways constitute two central premises of the theory of facilitated variation (Gerhart and Kirschner, 2007), which explains the emergence of phenotypic variations from small genetic changes.

Breaking Symmetry

The genetic approaches that led to the identification of the components involved in cue-mediated cell polarization emphasized a linear and hierarchic development of cell polarity. However, most cells have the capacity to polarize in the absence of any positional cue or landmark, albeit in a random orientation (Kirschner *et al.*, 2000; Wedlich-Soldner and Li, 2003). This indicates that cells have an intrinsic ability to break symmetry that only depends on their internal biochemical state. Over the last decade, extensive experimental and modeling work has focused on the mechanisms for such spontaneous cell polarization and has identified fundamental features driving cellular morphogenesis. Yeast cell polarization

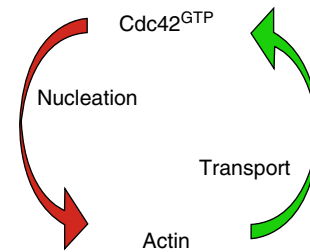


Figure 3 A simple positive feedback loop. Activated Cdc42 in the PM recruits or nucleates actin cables that in turn serve as tracks for transport of vesicles that bring more Cdc42.

and in particular the mechanism of Cdc42 polarization have played an important model role in this process. An initial key finding was that Cdc42 not only controlled actin organization but that Cdc42^{GTP} localization and accumulation in polarized regions in turn depends on actin-mediated transport (Figure 3).

Mathematical simulations based on this observation showed that simple amplification of stochastic variations in Cdc42 distribution through positive feedback can generate a polarized distribution of Cdc42 (Wedlich-Soldner *et al.*, 2003). Despite its simplicity, this initial model was able to make semi-quantitative predictions on the effects of transport speed or Cdc42 concentration, which could then be experimentally addressed. However, to allow a more realistic representation of cell polarization, some equivalent of the ‘long range inhibitor’ proposed in the classical pattern formation theory (Gierer and Meinhardt, 1972) needed to be included. A follow up study achieved this by introducing recycling of Cdc42 from the PM through endocytosis. In addition, quantitative values for model parameters were determined through live cell imaging. A particularly powerful experiment in this context is the measurement of protein mobility via fluorescence recovery after photobleaching. The resulting calculations of lateral diffusion rates for Cdc42 in the PM and the relevant association/dissociation times allowed predictions of the precise Cdc42 distribution during yeast polarity establishment based on a deterministic mathematical model (Marco *et al.*, 2007).

Both initial models for yeast polarization largely focused on the active, GTP-bound form of Cdc42. Cdc42^{GTP} is stably inserted into the membrane, and its distribution is dynamically determined by a balance of exocytosis, lateral diffusion within the PM and endocytosis (Figure 4). As both endo- and exocytosis require the actin cytoskeleton, Cdc42^{GTP} hence cannot polarize in the absence of actin structures. In contrast, wildtype Cdc42 rapidly cycles between its active and inactive forms, and this cycle is coupled to Cdc42 extraction from the membrane by Rdi1, the sole yeast member of the Rho-GDI (guanine nucleotide dissociation inhibitor) family (Tiedje *et al.*, 2008). GDIs can sequester the prenyl anchors present in Rho- and Rab GTPases and thereby extract them from membranes (Figure 4). Importantly this interaction is much more efficient for GDP-bound than for GTP-bound GTPases, leading to selective extraction of inactive Cdc42 (Johnson *et al.*, 2009).

Recent work, again combining quantitative imaging with genetic manipulations and mathematical modeling, has

demonstrated that GDI-mediated recycling of Cdc42 constitutes a second feedback loop that can drive efficient cell polarization (Slaughter *et al.*, 2009a). As this reaction involves cytosolic GDI-Cdc42 complexes, rather than the purely membrane bound Cdc42 in the actin feedback loop, the positive feedback in the system needs to be strong enough to counter the rapid spatial equilibration of soluble Cdc42. This is achieved by an elaborate mechanism: On the one hand GEF and GDI have been shown to compete for binding to Cdc42^{GDP}, which leads to preferential release of Cdc42 at sites of high GEF concentration. On the other hand active Cdc42 binds to the scaffold protein Bem1, which in turn recruits the GEF Cdc24 (Kozubowski *et al.*, 2008). In combination, these two reactions build a strong positive feedback loop that allows efficient accumulation of Cdc42 and spontaneous breaking of symmetry (Freisinger *et al.*, 2013). The Bem1/GDI loop has long been at the center of theoretical considerations on

symmetry breaking (Altschuler *et al.*, 2008; Goryachev and Pokhilko, 2008; Rubinstein *et al.*, 2012; Savage *et al.*, 2012; Klunder *et al.*, 2013). All models to date, whether stochastic and deterministic, were built around one or multiple positive feedback loops and were used to simulate distribution of Cdc42 and its direct regulators. Despite strong similarities in their fundamental assumptions the different models lead to very distinct outcomes. Validated by increasingly precise and quantitative experimental evidence the models became successively more detailed. This evolution of approaches at the interface of quantitative biology and theoretical sciences over the last decade can be seen as a paradigm for the study of cellular pattern formation. It is instructive to consider the different modeling approaches. When doing so, however, one should keep in mind that the value of mathematical models lies in their predictive power for a particular biological question, and this does not necessarily require inclusion of all known variables and interactions (Mogilner *et al.*, 2012). Rather on the contrary, simpler models have proven very effective in identifying key aspects of otherwise nonintuitive regulatory cycles.

Considering the existence of parallel pathways for Cdc42 recycling and polarization, the next challenge was to determine how the two feedback loops were coordinated in the cell. Again combining experiments and theory it was shown that only a combination of actin- and GDI-dependent recycling of Cdc42 allows yeast cells to establish robust cell polarity at a single site (Freisinger *et al.*, 2013). While the GDI pathway allows rapid recycling of Cdc42 and consistently generates a single-polarization site, it requires Cdc42 to cycle rapidly between its active and inactive form, and is therefore sensitive to perturbations of the GTPase cycle. Conversely, slower actin-mediated recycling of Cdc42 can induce robust symmetry breaking but fails to restrict polarization to a single site. These results suggest that cells optimize the efficiency of symmetry breaking by coupling multiple feedback loops with distinct properties (Figure 5).

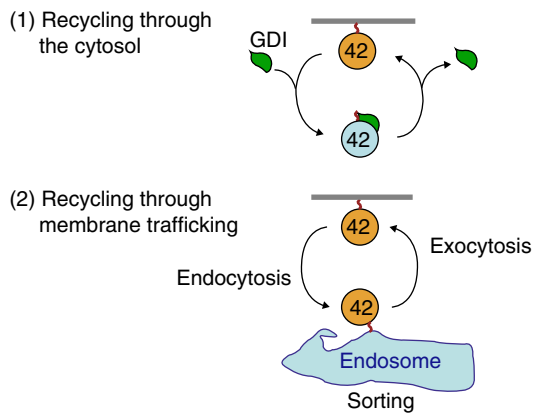


Figure 4 Mechanisms for Cdc42 recycling. Cdc42 can be removed from the PM either by extraction through a GDI (1) or via actin-mediated endocytosis (2). Release from the GDI or recycling of membranes via exocytosis will then return Cdc42 to the PM.

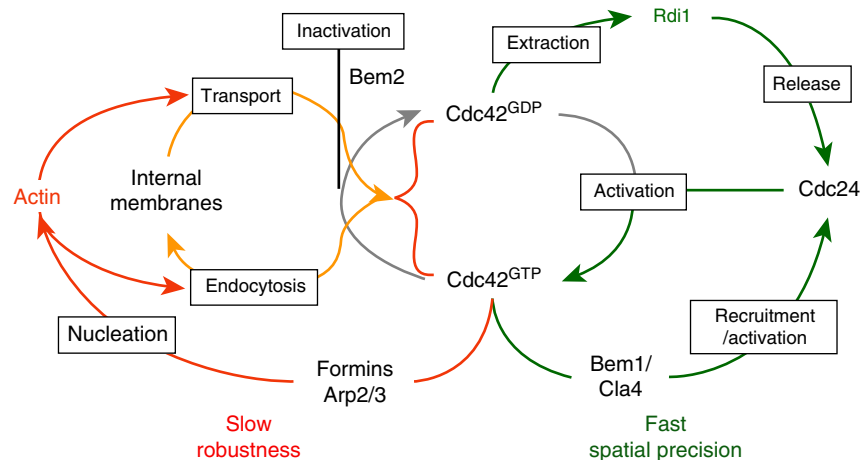


Figure 5 A combination of feedback loops. The schematic depicts the various reactions and feedback loops that drive Cdc42 polarization during yeast budding. Rdi1 is the only yeast Rho-GDI. Bem1 interacts with the GEF Cdc24 and the PAK kinase Cla4. Formins are known to nucleate and elongate actin cables, while the Arp3/3 complex nucleates branched actin filaments that are required for formins. The exact biochemical link between Cdc42 and the actin nucleators is not yet clear.

Inhibitory Interactions and Reinforcements in Polarity Regulation

In the previous section different aspects of positive feedback were discussed, that allow budding yeast cells to spontaneously break symmetry and to establish cell polarity. In addition to these highly dynamic core processes involving rapid activity cycles and physical recycling of Cdc42, several additional mechanisms have been shown to either reinforce or counter Cdc42 polarization. These processes might be particularly important in maintaining or adapting a polarized growth pattern in the face of changing cell cycle regulation and constant exchange of material.

Work in the fission yeast *Schizosaccharomyces pombe* has highlighted the interplay between surface mechanics and polarity (Bonazzi *et al.*, 2014). During spore germination, the initial polarized region wanders around the surface and mediates isotropic growth. Outgrowth of a polar tube only occurs once the outer spore wall, which counteracts osmotic pressure, is ruptured at a single site. These results suggest that mechanical tension in the cortex can counteract cell expansion and thereby negatively regulate cell polarity. This effect is likely applicable to all cell walled cells. Even more generally, mechanical signals have been proposed to play a fundamental role in plant and animal polarization (Asnacios and Hamant, 2012).

In contrast to the proposed positive feedback loop involving vesicular trafficking of Cdc42, actin has been suggested to rather play an inhibitory role in polarity maintenance due to a net dilution of Cdc42 in the polarized region if vesicles fail to sufficiently concentrate Cdc42 (Layton *et al.*, 2011). A solution to these conflicting predictions came from studies showing that endocytosis and exocytosis are spatially segregated in the polar PM region (Jose *et al.*, 2013; Slaughter *et al.*, 2013) and that Cdc42 is actually enriched at sites of exocytosis, where it also exhibits slower lateral diffusion. Modeling indicated that nonuniform membrane diffusion of Cdc42 enables sustained cell polarity even at low concentrations of Cdc42 on vesicles (Slaughter *et al.*, 2013). Interestingly, Cdc42 microdomains are also enriched in the anionic phospholipid phosphatidylserine (PS), and PS was shown to contribute to the slow diffusion of Cdc42 (Slaughter *et al.*, 2013). A positive role for PS in establishment and maintenance of polarity was further supported by studies showing that Cdc42 localization was mediated by PS (Fairm *et al.*, 2011) and that the ratio between PS and the neutral lipid phosphatidylethanolamine regulates GDI-mediated recycling of Cdc42 (Das *et al.*, 2012a). As many of the regulators and effectors that bind to Cdc42 contain lipid binding segments, it is reasonable to assume that local lipid composition in the PM will affect many of the pathways and feedback loops regulating cell polarity. In this context it is important to note that the yeast PM has been shown to be segregated into a large number of overlapping microdomains (Spira *et al.*, 2012) through weak protein-lipid interactions. Together these results suggest that polarization or any spatial organization on the level of a whole cell may in fact originate from self-organization mechanisms on the level of individual protein-membrane complexes.

Positive feedback is sufficient to explain the spontaneous emergence of a single polarity site. However, when observing

the development of polarized sites over longer time scales phenomena such as competing, wandering or oscillating polarity sites have been reported, which indicate the presence of additional negative feedback (Wedlich-Soldner *et al.*, 2004; Ozbudak *et al.*, 2005; Howell *et al.*, 2012). An exciting mechanism for such negative regulation could occur through inhibitory phosphorylation of the GEF Cdc24 via the Cdc42-activated PAK kinase Cla4 (Kuo *et al.*, 2014). This scenario constitutes a prime example of an incoherent feedforward loop (Hart and Alon, 2013) built on the activities of Bem1. This scaffold protein binds directly to active Cdc42, Cdc24 and to Cla4. On the one hand binding of Bem1 to Cdc24 relieves the latter's autoinhibition and recruits it to its target, Cdc42. On the other hand the recruitment of Cla4 leads to a delayed phosphorylation and thereby inactivation of Cdc24. Such counteracting – or paradoxical – activities are frequently used by cellular regulatory modules to achieve robustness of a response against fluctuations in the concentration of individual regulators (Hart and Alon, 2013). While incoherent feedforward loops cannot generate oscillations on their own, they would be able to do so in combination with positive feedback, which is clearly present in polarizing yeast cells (Kuo *et al.*, 2014). A similar mechanism centered on delayed negative feedback has also been put forward to explain oscillatory polarization of fission yeast cells (Das *et al.*, 2012b).

A different kind of negative feedback has been recently proposed to act through septins. It has been well established that Cdc42 can recruit septins directly and via the yeast specific effector proteins Gic1 and Gic2 (Sadian *et al.*, 2013). Septins in turn are able to recruit the Cdc42 GAP Bem2, which leads to a delayed inactivation of Cdc42 at sites of septin assembly. An interesting twist to this story is that Cdc42-driven exocytosis likely leads to displacement of septin polymers from the center of the polarized region and thereby supports formation of the septin ring (Okada *et al.*, 2013).

A final type of polarity regulation centering on GTPase signaling concerns inhibitory landmarks that prevent recruitment of Cdc42 sites of previous cell division. On the one hand the Cdc42 GAP Rga1 establishes an exclusion zone at the division site that blocks subsequent polarization within that site (Tong *et al.*, 2007). In addition a protein complex that competes for Cdc24 binding to Cdc42 is permanently anchored at bud scars and thereby serves as negative memory for polarity establishment (Meitinger *et al.*, 2014). Both of these mechanisms constitute static zones of occlusion and therefore provide a distinct approach to polarity regulation compared to the dynamic concentration of Cdc42 typical for the feedback regulation discussed before.

In summary, small GTPases constitute central regulators for spatial organization in cells. Rho-GTPases and in particular Cdc42 are key regulators of the actin cytoskeleton and of cell polarization. Studies on Cdc42-mediated cell polarization have been at the forefront of mesoscale systems biology approaches that combine quantitative imaging techniques with molecular perturbations and mathematical models. A large number of yeast studies in the last decade have constantly pushed our understanding of the fundamental principles underlying cell polarity establishment and maintenance – and are still providing surprising new insights. The concepts and techniques developed in these studies have since been

picked up in many other fields and have led to a dramatic increase in our appreciation of the self-organization properties of cellular circuits. The coming years will see a further rapid expansion of our quantitative repertoire – both technical and theoretical. The next big challenge will be to now transfer the gained knowledge into creation of minimal or synthetic systems.

See also: Cell Communication: Cellular Outcomes: Regulation of Cell Polarity. Cytoskeleton and Motors: Complex Cytoskeletal Structures: Cell Polarity

References

- Altschuler, S.J., Angenent, S.B., Wang, Y., Wu, L.F., 2008. On the spontaneous emergence of cell polarity. *Nature* 454, 886–889.
- Arkowitz, R.A., Bassilana, M., 2011. Polarized growth in fungi: Symmetry breaking and hyphal formation. *Seminars in Cell & Developmental Biology* 22, 806–815.
- Asnacios, A., Hamant, O., 2012. The mechanics behind cell polarity. *Trends in Cell Biology* 22, 584–591.
- Bonazzi, D., Julien, J.D., Romao, M., *et al.*, 2014. Symmetry breaking in spore germination relies on an interplay between polar cap stability and spore wall mechanics. *Developmental Cell* 28, 534–546.
- Casamayor, A., Snyder, M., 2002. Bud-site selection and cell polarity in budding yeast. *Current Opinion in Microbiology* 5, 179–186.
- Chesneau, L., Dambournet, D., Machicoane, M., *et al.*, 2012. An ARF6/Rab35 GTPase cascade for endocytic recycling and successful cytokinesis. *Current Biology* CB 22, 147–153.
- Das, A., Slaughter, B.D., Unruh, J.R., *et al.*, 2012a. Flippase-mediated phospholipid asymmetry promotes fast Cdc42 recycling in dynamic maintenance of cell polarity. *Nature Cell Biology* 14, 304–310.
- Das, M., Drake, T., Wiley, D.J., *et al.*, 2012b. Oscillatory dynamics of Cdc42 GTPase in the control of polarized growth. *Science* 337, 239–243.
- Drubin, D.G., Nelson, W.J., 1996. Origins of cell polarity. *Cell* 84, 335–344.
- Fairn, G.D., Hermansson, M., Somerharju, P., Grinstein, S., 2011. Phosphatidylserine is polarized and required for proper Cdc42 localization and for development of cell polarity. *Nature Cell Biology* 13, 1424–1430.
- Freisinger, T., Klunder, B., Johnson, J., *et al.*, 2013. Establishment of a robust single axis of cell polarity by coupling multiple positive feedback loops. *Nature Communications* 4, 1807.
- Gerhart, J., Kirschner, M., 2007. The theory of facilitated variation. *Proceedings of the National Academy of Sciences of the United States of America* 104 (Suppl. 1), 8582–8589.
- Gierer, A., Meinhardt, H., 1972. A theory of biological pattern formation. *Kybernetik* 12, 30–39.
- Goryachev, A.B., Pokhilko, A.V., 2008. Dynamics of Cdc42 network embodies a Turing-type mechanism of yeast cell polarity. *FEBS Letters* 582, 1437–1443.
- Harris, K.P., Tepass, U., 2010. Cdc42 and vesicle trafficking in polarized cells. *Traffic* 11, 1272–1279.
- Hart, Y., Alon, U., 2013. The utility of paradoxical components in biological circuits. *Molecular Cell* 49, 213–221.
- Howell, A.S., Jin, M., Wu, C.F., *et al.*, 2012. Negative feedback enhances robustness in the yeast polarity establishment circuit. *Cell* 149, 322–333.
- Jin, Y., Sultana, A., Gandhi, P., *et al.*, 2011. Myosin V transports secretory vesicles via a Rab GTPase cascade and interaction with the exocyst complex. *Developmental Cell* 21, 1156–1170.
- Johnson, J.L., Erickson, J.W., Cerione, R.A., 2009. New insights into how the Rho guanine nucleotide dissociation inhibitor regulates the interaction of Cdc42 with membranes. *Journal of Biological Chemistry* 284, 23860–23871.
- Jose, M., Tollis, S., Nair, D., Sibarita, J.B., McCusker, D., 2013. Robust polarity establishment occurs via an endocytosis-based cortical corralling mechanism. *Journal of Cell Biology* 200, 407–418.
- Kang, P.J., Angerman, E., Nakashima, K., Pringle, J.R., Park, H.O., 2004. Interactions among Rax1p, Rax2p, Bud8p, and Bud9p in marking cortical sites for bipolar bud-site selection in yeast. *Molecular Biology of the Cell* 15, 5145–5157.
- Keilberg, D., Sogaard-Andersen, L., 2014. Regulation of bacterial cell polarity by small GTPases. *Biochemistry* 53, 1899–1907.
- Kirschner, M., Gerhart, J., Mitchison, T., 2000. Molecular 'vitalism'. *Cell* 100, 79–88.
- Kirschner, M.W., 2005. The meaning of systems biology. *Cell* 121, 503–504.
- Klunder, B., Freisinger, T., Wedlich-Soldner, R., Frey, E., 2013. GDI-mediated cell polarization in yeast provides precise spatial and temporal control of Cdc42 signaling. *PLoS Computational Biology* 9, e1003396.
- Kozubowski, L., Saito, K., Johnson, J.M., *et al.*, 2008. Symmetry-breaking polarization driven by a Cdc42p GEF-PAK complex. *Current Biology* CB 18, 1719–1726.
- Kuo, C.C., Savage, N.S., Chen, H., *et al.*, 2014. Inhibitory GEF phosphorylation provides negative feedback in the yeast polarity circuit. *Current Biology* CB 24, 753–759.
- Layton, A.T., Savage, N.S., Howell, A.S., *et al.*, 2011. Modeling vesicle traffic reveals unexpected consequences for Cdc42p-mediated polarity establishment. *Current Biology* CB 21, 184–194.
- Marco, E., Wedlich-Soldner, R., Li, R., Altschuler, S.J., Wu, L.F., 2007. Endocytosis optimizes the dynamic localization of membrane proteins that regulate cortical polarity. *Cell* 129, 411–422.
- Meitinger, F., Khmelinskii, A., Morlot, S., *et al.*, 2014. A memory system of negative polarity cues prevents replicative aging. *Cell* 159, 1056–1069.
- Mogilner, A., Allard, J., Wollman, R., 2012. Cell polarity: Quantitative modeling as a tool in cell biology. *Science* 336, 175–179.
- Nern, A., Arkowitz, R.A., 2000. Nucleocytoplasmic shuttling of the Cdc42p exchange factor Cdc24p. *Journal of Cell Biology* 148, 1115–1122.
- Okada, S., Leda, M., Hanna, J., *et al.*, 2013. Daughter cell identity emerges from the interplay of Cdc42, septins, and exocytosis. *Developmental Cell* 26, 148–161.
- Ozbudak, E.M., Becskei, A., Van Oudenaarden, A., 2005. A system of counteracting feedback loops regulates Cdc42p activity during spontaneous cell polarization. *Developmental Cell* 9, 565–571.
- Rubinstein, B., Slaughter, B.D., Li, R., 2012. Weakly nonlinear analysis of symmetry breaking in cell polarity models. *Physical Biology* 9, 045006.
- Sadian, Y., Gatsogiannis, C., Patasi, C., *et al.*, 2013. The role of Cdc42 and Gic1 in the regulation of septin filament formation and dissociation. *eLife* 2, e01085.
- Savage, N.S., Layton, A.T., Lew, D.J., 2012. Mechanistic mathematical model of polarity in yeast. *Molecular Biology of the Cell* 23, 1998–2013.
- Slaughter, B.D., Das, A., Schwartz, J.W., Rubinstein, B., Li, R., 2009a. Dual modes of cdc42 recycling fine-tune polarized morphogenesis. *Developmental Cell* 17, 823–835.
- Slaughter, B.D., Smith, S.E., Li, R., 2009b. Symmetry breaking in the life cycle of the budding yeast. *Cold Spring Harbor Perspectives in Biology* 1, a003384.
- Slaughter, B.D., Unruh, J.R., Das, A., *et al.*, 2013. Non-uniform membrane diffusion enables steady-state cell polarization via vesicular trafficking. *Nature Communications* 4, 1380.
- Spira, F., Mueller, N.S., Beck, G., *et al.*, 2012. Patchwork organization of the yeast plasma membrane into numerous coexisting domains. *Nature Cell Biology* 14, 640–648.
- Tiedje, C., Sakwa, I., Just, U., Hofken, T., 2008. The Rho GDI Rdi1 regulates Rho GTPases by distinct mechanisms. *Molecular Biology of the Cell* 19, 2885–2896.
- Tong, Z., Gao, X.D., Howell, A.S., *et al.*, 2007. Adjacent positioning of cellular structures enabled by a Cdc42 GTPase-activating protein-mediated zone of inhibition. *Journal of Cell Biology* 179, 1375–1384.
- Wedlich-Soldner, R., Altschuler, S., Wu, L., Li, R., 2003. Spontaneous cell polarization through actomyosin-based delivery of the Cdc42 GTPase. *Science* 299, 1231–1235.
- Wedlich-Soldner, R., Li, R., 2003. Spontaneous cell polarization: undermining determinism. *Nature Cell Biology* 5, 267–270.
- Wedlich-Soldner, R., Wai, S.C., Schmidt, T., Li, R., 2004. Robust cell polarity is a dynamic state established by coupling transport and GTPase signaling. *Journal of Cell Biology* 166, 889–900.
- Yang, Z., Lavagi, I., 2012. Spatial control of plasma membrane domains: ROP GTPase-based symmetry breaking. *Current Opinion in Plant Biology* 15, 601–607.

Theory of Cargo and Membrane Trafficking

L Foret, Ecole Normale Supérieure, Paris, France

L Brusch, Technische Universität Dresden, Dresden, Germany

F Jülicher, Max Planck Institute for the Physics of Complex Systems, Dresden, Germany

© 2016 Elsevier Inc. All rights reserved.

Introduction

Any eukaryotic cell transports nutrients and signaling molecules, here collectively called cargo, between its plasma membrane and intracellular target compartments (Mellman, 1996). Receptor-mediated endocytosis, through inward budding of clathrin-coated vesicles (CCV) or clathrin-independent vesicles, provides the means for selective cargo uptake (Conner and Schmid, 2003). This endocytic cargo uptake inevitably is coupled to membrane and receptor uptake. Immediate delivery of the many small uptake vesicles to target compartments would risk a large membrane and receptor loss at the price of resynthesis. A solution to this unfavorable consequence of coupled cargo and membrane trafficking is the intermediate accumulation of like cargo and recycling of membrane and a specific set of receptors to the plasma membrane (Sheff *et al.*, 2002; Lauffenburger and Linderman, 1993). Early endosomes (EEs), i.e., several hundreds of intermediary vesicles per cell, fulfill this essential accumulation and sorting function (Figures 1(a) and 1(c)). EEs constitutively undergo homotypic fusion and fission among them and as a whole constitute the endosomal network of dynamically interlinked vesicles (Gruenberg and Maxfield, 1995; Collinet *et al.*, 2010).

There is an intriguing question regarding the overall performance of the endosomal network. As the endosomal network is fueled by small uptake vesicles and homotypic fusion and fission occur more or less randomly given heterogeneous cargo mixtures and their noisy perception, how could endosomes carrying large amounts of cargo be produced for more efficient cargo delivery to target compartments? In quantitative terms, the size distribution of the EEs should possess a broad tail as does a power law distribution (Figures 1(b), 1(d), and 1(e)).

The individual EEs operate independently of one another and have no information on the overall cargo distribution. Such a power law relation is characteristic for many complex systems (Ziemelis, 2001). Nontrivial collective properties can emerge at the scale of the whole endosomal network (here called the macroscopic level) that are not imposed explicitly at the level of the system components (the microscopic level). Repeated interactions among the components let the collective properties emerge or self-organize.

An additional tool that provides important insights is a theory of the endosomal network in which the details of single endosome interactions are described by mathematical terms and the collective effects of many such interactions can be analyzed using mathematical methods. Once having linked microscopic to macroscopic properties through a theory then one can also use it as a means to infer microscopic properties that are not directly observable from readily observable macroscopic properties. We review such a theory as developed by Foret *et al.* (2012) in the following Section 'General Model'

and progress with its analysis for the particular case of low-density lipoprotein (LDL) trafficking in HeLa cells in Section 'Entry-Fusion-Exit Model.' The comparison to experimental data confirms the theoretical predictions and lets one estimate the microscopic parameter values. In subsequent sections, we describe selected mechanisms for regulating endosome interactions and discuss how they refine cargo and membrane trafficking.

Theory of the Endosomal Network

General Model

Cargo distribution in a population of endosomes

At a given instant t , each endosome contains a certain amount of cargo (i.e., number of cargo molecules), see Figures 1(a) and 1(c). The overall population is characterized by the distribution of cargo defined as the number $n(s,t)$ of endosomes loaded with a number s of cargo molecules. Repeated interactions between endosomes and with other cell organelles result in the continual redistribution of the cargo in the population and thus lead to the evolution of $n(s,t)$ overtime. For example, when an endosome with 15 cargo molecules fuses with an endosome with 6 cargo molecules, these two endosomes disappear (then $n(15)$ and $n(6)$ decrease by one) and a new endosome with 21 cargo molecules appears ($n(21)$ increases by one). The profile of the distribution $n(s,t)$ and the way it evolves directly reflect the underlying dynamics of the many individual endosomes. The goal of the theory is to predict the emergent macroscopic behavior of $n(s,t)$ from the microscopic interaction processes at the single endosome level such as fusion and fission.

Microscopic processes and parameters

Only six basic processes can modify the cargo distribution in the endosome population. They are sketched in Figures 2(a)–2(f) and any other process would involve three or more simultaneously interacting endosomes, which is negligibly rare. The homotypic fusion of two endosomes loaded with the cargo amount s and s' , respectively, leads to the disappearance of these two endosomes and to the formation of a new endosome containing $s + s'$ cargo (Gorvel *et al.*, 1991; Gruenberg and Maxfield, 1995; Rink *et al.*, 2005; Rybin *et al.*, 1996), see panel (c) in Figure 2. Homotypic endosome fission (d) is the reverse process: two new endosomes with the cargo amount s and s' arise from a single endosome containing $s + s'$ cargo (Bonifacino and Glick, 2004; Rink *et al.*, 2005; Skjeldal *et al.*, 2012). Beside these internal exchanges among individuals within the population, the endosome population receives and delivers cargo from and to other cell organelles by means of the four processes (a), (b), (d), and (e), but no other.

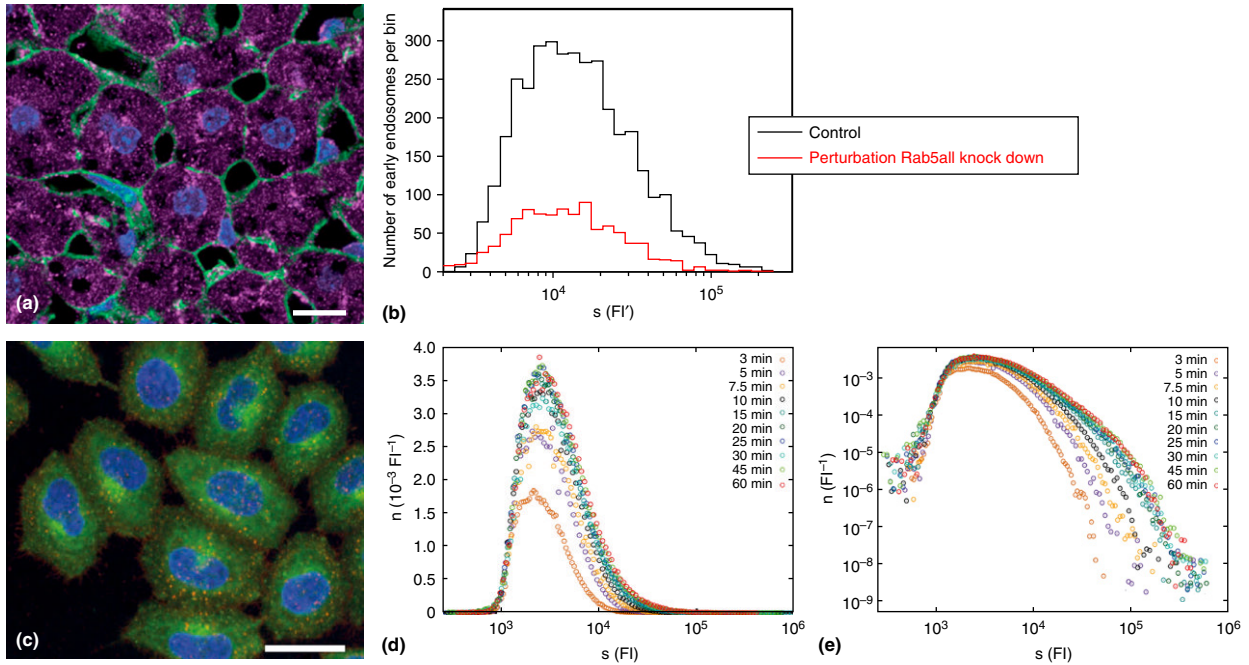


Figure 1 Quantification of endosome distributions *in vivo* and *in vitro*. (a) Fluorescence microscopy image of a mouse liver section stained for EEA1 (together with phalloidin (green) and DAPI (blue)), marking early endosomes (EEs) as magenta objects in hepatocytes *in vivo*. (b) Cargo distribution in early endosomes on lin-log scales (black curve for normal liver) with cargo amount s after 20 min of internalizing labeled LDL. Perturbation of regulatory components of the endosomal network changes the cargo distribution *in vivo* (red curve for liver-specific Rab5 knock-down). (c) Fluorescence microscopy image of HeLa cells *in vitro* expressing GFP-Rab5 (green), loaded with labeled LDL cargo (red) and stained by DAPI (blue). (d, e) Cargo distribution in early endosomes on lin-log (d) and log-log (e) scales. The broad tail (with a power law contribution) of the distribution is the hallmark of cargo accumulation through many rounds of homotypic EE-EE fusion and membrane recycling. Individual colors refer to different time points after adding labeled LDL cargo to the culture medium. FI and FI' are microscope-specific units of cargo fluorescence intensity. Scale bars are 20 μm in (a) and 30 μm in (c). (a) and (b) are adapted by permission from Macmillan Publishers Ltd: Zeigerer, A., Gilleron, J., Bogorad, R.L., *et al.*, 2012. Rab5 is necessary for the biogenesis of the endolysosomal system *in vivo*. *Nature* 485, 465–470, copyright (2012). (c)–(e) are reprinted from Foret, F., Dawson, J.E., Villaseñor, R., *et al.*, 2012. A general theoretical framework to infer endosomal network dynamics from quantitative image analysis. *Current Biology* 22, 1381–1390, copyright (2012), with permission from Elsevier.

Cell organelles are able to abruptly and irreversibly change their nature by a rapid modification of their biochemical composition. This process is called maturation, or sometime conversion. Non-endosomal structures can convert into endosomes and conversely endosomes can convert into organelles of another type. For example, endocytosed vesicles may convert into early endosomes (Conner and Schmid, 2003; Gruenberg and Maxfield, 1995) while early endosomes can convert into late endosomes (Rink *et al.*, 2005; Del Conte-Zerial *et al.*, 2008; Poteryaev *et al.*, 2010). Conversion of a cargo-filled vesicle into an endosome is equivalent to the creation (biogenesis) of a new endosome with the same cargo amount as the original object (a). On the other hand, conversion of an endosome into another structure is equivalent to the removal of this endosome (and its cargo) from the population (f). Finally, individual endosomes can receive and release cargo by means of heterotypic fusion (b) and heterotypic fission (e), respectively. A heterotypic fusion event with a cargo-loaded vesicle results in the growth of the cargo amount in that endosome. The budding and scission of a vesicle or tubule decreases the cargo amount in the parent endosome but neither of these two processes changes the total number of endosomes. (a)–(f) is the exhaustive list of events able to affect the cargo distribution in an endosome network. However,

depending on the type of endosomes and on the type of cargo, not necessarily all six processes are active.

The six events (a)–(f) occur repeatedly in an apparently stochastic manner driving the evolution of the population. The probability that a fusion of two endosomes with cargo amount s and s' occurs in the short time interval dt is $n(s)n(s')K(s,s')dt$; it is proportional to the numbers of endosomes with s and s' cargo amount and to dt . The coefficient $K(s,s')$ is the fusion rate, whose value depends on the molecular details of the protein machineries responsible for endosome motility, tethering and membrane fusion. The theory at the endosome scale does implicitly include such details through the fusion rate $K(s,s')$. It is a function of s and s' since the cargo load could indirectly regulate the fusion machinery favoring or preventing fusion. In the same manner, we can define five more rates characterizing each of the processes (a)–(f) of Figure 2. The probabilities associated to the occurrence, during dt , of the other events as ordered in Figure 2 are $A(s)dt$, $v_{in}(s)n(s)dt$, $n(s+s')K'(s,s')dt$, $v_{out}(s)n(s)dt$ and $k_d(s)n(s)dt$ with $A(s)$ the *de novo* creation rate, $v_{in}(s)$ the cargo absorption rate, $K'(s,s')$ the fission rate, $v_{out}(s)$ the cargo release rate, and $k_d(s)$ the conversion rate. The six rate functions are the context-specific input of the theory. Given these parameters, the goal is to calculate $n(s,t)$.

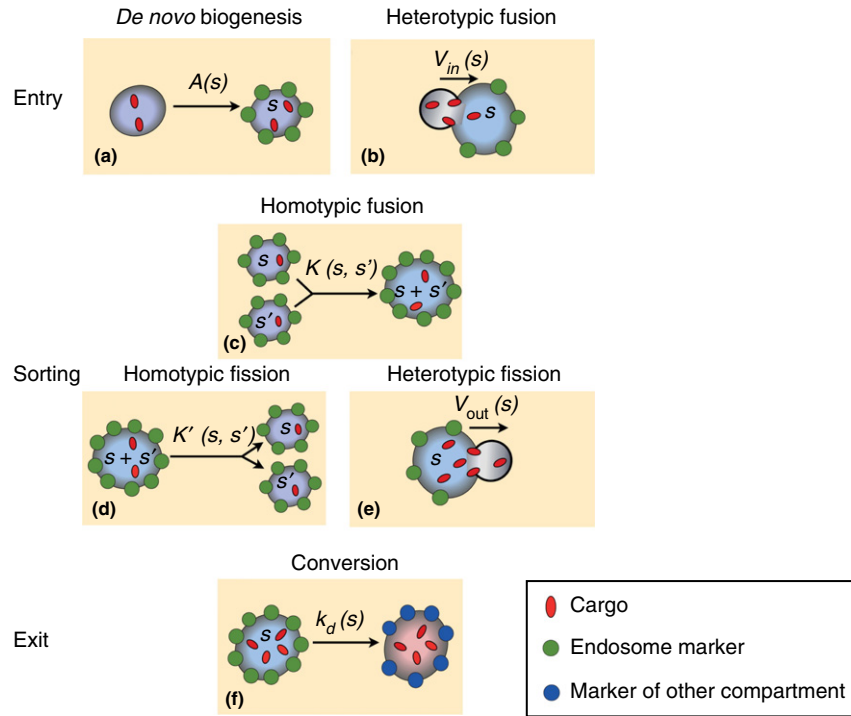


Figure 2 There are six principle processes of endosome interactions. Here these are ordered according to their sequence during cargo transport (red symbols). Green (blue) symbols denote membrane-associated proteins, such as Rab5 (Rab7), which confer functional specificity to endosomes. See Section ‘Microscopic processes and parameters’ for details on individual processes (a–f). Reprinted from Foret, F., Dawson, J.E., Villaseñor, R., *et al.*, 2012. A general theoretical framework to infer endosomal network dynamics from quantitative image analysis. *Current Biology* 22, 1381–1390, copyright (2012), with permission from Elsevier.

Governing equation

Having chosen the parameters (the six rates) and an initial state of the endosome population of a cell, one can simply simulate the stochastic evolution of $n(s, t)$ by repeatedly using the probabilities for each process to occur within short-time intervals. Such a stochastic simulation can be efficiently performed using the Gillespie algorithm (Cao *et al.*, 2004; Mukherji and O’Shea, 2014). Simulating the evolution several times and averaging, one can compute the evolution of the mean $n(s, t)$ that can be compared to experimental data.

Another approach is to derive a mathematical equation satisfied by the mean $n(s, t)$. Because of the large number of endosomes in a population, a kinetic equation similar to those used in chemical kinetics describes the evolution of the mean $n(s, t)$. It reads:

$$\begin{aligned}
 \frac{\partial n(s)}{\partial t} = & \frac{1}{2} \int_0^s K(s', s-s') n(s-s') n(s') ds' \\
 & - \int_0^\infty K(s', s) n(s) n(s') ds' \\
 & + \int_0^\infty K'(s, s') n(s+s') ds' \\
 & - \frac{1}{2} \int_0^s K'(s-s', s') n(s) ds' \\
 & + A(s) - k_d(s) n(s) - \frac{\partial}{\partial s} [v_{in}(s) n(s)] \\
 & + \frac{\partial}{\partial s} [v_{out}(s) n(s)]
 \end{aligned} \quad [1]$$

Note that in order to make the mathematical analysis of the equation simpler, we have considered the cargo amount s as a continuous variable. It follows that $n(s)$ is a number density, $n(s)ds$ being the mean number of endosomes with a cargo amount between s and $s+ds$ for a vanishingly small ds . An analogous approach is used to describe coagulation of particles in physics (Smoluchowski, 1916; Hogg *et al.*, 1966). Each term of the equation gives the average number of endosomes with s cargo created or lost per time unit due to one of the six basic processes, with fusion and fission requiring two terms each. The merit of this approach is that exact or approximated mathematical solutions can be found for simple forms of the parameters. This allows us to gain mechanistic understanding of the contribution of each microscopic process to the macroscopic property $n(s, t)$ and on the way how it evolves.

Entry-Fusion-Exit Model

Let’s now apply the theory to predict the features of the distribution of cargo $n(s)$ for a simple set of parameters. Consider a hypothetical population in which new endosomes with small cargo amount s_0 appear at the rate $A=J/s_0$ (J is the cargo influx), fuse at the constant rate K and convert at the constant rate k_d . Fissions (Figures 2(d) and 2(e)) and heterotypic fusion (b) are not included (their rates are set to zero). The endosome population is initially devoid of cargo. We refer to this choice of parameters as the ‘entry-fusion-exit model’ of endosome dynamics.

Theoretical predictions

In this simple model, the evolution of the total number $N = \int_0^\infty n(s)ds$ of endosomes in the population is governed by the equation,

$$\frac{dN}{dt} = \frac{J}{s_0} - \frac{KN^2}{2} - k_d N \quad [2]$$

From here on, we consider that fusion events are much more frequent than conversion events. Then the solution is approximately

$$N(t) = \sqrt{2J/Ks_0} \tanh\left(\sqrt{JK/2s_0} t\right) \quad [3]$$

At early times, while the number of endosomes is small, fusion (and conversion) is very rare, the number of endosomes rises as $N(t) = (J/s_0)t$. After some time, the number of fusion events and *de novo* creation events balance each other and the number of endosomes stays nearly constant at $N(t) = \sqrt{2J/Ks_0}$.

We now turn to the behavior of the full distribution $n(s,t)$ in the entry-fusion-exit model. Its evolution can be divided into three phases. In the first phase, the distribution $n(s,t)$ peaks at s_0 and its height $n_{\max}$ behaves as the total number of endosomes, $n_{\max}(t) = (b/s_0)N(t)$ with b a numerical prefactor. It rises until fusion balances *de novo* creation. In the second phase, fusion events produce endosomes accumulating more and more cargo. The distribution $n(s,t)$ broadens more and more over time. In the third phase, endosome conversion curbs the growth of endosomes and yields a steady state: $n(s)$ does no longer evolve in phase 3. Mathematically, beyond the peak at s_0 and during the phases 2 and 3, the distribution is approximately

$$n(s) = \sqrt{J/2\pi K} s^{-3/2} \exp\left(-s/s^*(t)\right) \quad \text{with} \quad s^*(t) = \begin{cases} JKt^2 & \text{for } t \ll 1/k_d \\ 2JK/k_d^2 & \text{for } t \gg 1/k_d \end{cases} \quad [4]$$

The cargo distribution decays slowly as a power law $n_s(s) \sim s^{-3/2}$ up to a cargo value s^* beyond which the decay is abrupt. This maximal cargo load s^* increases with time during the second phase and finally saturates at a constant value in the third phase. This evolution is called ‘self-similar’: as time progresses, $n(s,t)$ gets ‘stretched’ along the s axis until its final steady state profile.

These results illustrate how the mathematical analysis of the equation allows us to understand the effect of each basic process on the shape of $n(s,t)$. Comparing these predictions to experimental data, one finds that already the simple entry-fusion-exit model very accurately describes the trafficking of LDL cargo in HeLa cells.

Comparison of theory predictions to LDL trafficking in HeLa cells

Analyzing fluorescence microscopy images (labeling an early endosome marker and cargo), one can detect endosomes, measure for each of them their cumulative cargo fluorescence which corresponds to the cargo amount s and then deduce the distribution $n(s)$, see Figure 1. In an interesting experiment, HeLa cells stably expressing Rab5-GFP as early endosome marker were placed in a medium containing fluorescent LDL cargo. The distribution of LDL amounts in Rab5-endosomes

was then measured at different times after the cells started to endocytose LDL. Figures 1(e) and 1(f) show that the experimentally measured cargo distributions qualitatively validate the prediction by the entry-fusion-exit model: a peak at small s rises and saturates during the first 10 min (phase 1) and on a longer time scale, the distribution progressively broadens and apparently stretches along the s axis (phase 2), it finally reaches a stationary state after 30 min (phase 3). Beyond this qualitative agreement, the data reproduced in Figure 1 bear all the mathematical features predicted in the preceding section. The evolutions of the height of the distribution’s peak $n_{\max}$ and of the total number of endosomes N are very well fitted by the law [3], see Figures 3(a) and 3(c). In Figure 3(b), the distributions at different times where plotted on a rescaled axis with an s^* chosen for each time such that the distribution tails overlap. Beyond the peak, all the curves fall onto the same curve, which is the function $x^{-3/2} \exp(-x)$ shown as solid line in Figure 3(b). This data collapse confirms the self-similar nature of the evolution and the accuracy of the predicted solution [4]. Moreover $s^*(t)$ obtained by the curve rescaling evolves as predicted in eqn. [4]: it first grows with time as t^2 and reaches a plateau after 30 min (inset in Figure 3(b)).

In order to further test the predictions, one would like to experimentally tune some parameters. Changing the concentration of LDL in the cell culture medium results in the change of the LDL influx J into each cell and thus into the early endosomal system. The value of J can be easily deduced from the evolution of the total LDL fluorescence of the cell. The experiments were repeated using 4 different values of J . The evolution of $N(t)$ in the four cases is shown in Figure 3(c). When plotting $N/\sqrt{J}$ as a function of $\sqrt{J}t$, all the curves once again overlap as predicted by eqn. [3], Figure 3(d).

What do we learn?

Quantitative comparison between experimental data on the entire endosome population and the theoretical predictions allows us to reach unbiased conclusions regarding the organization of the early endosomal system. Degradative cargoes such as LDL that enter the cell in many small structures are progressively concentrated in a few large endosomes by the mean of repeated random fusion and finally sent to the late endosomal system by conversion of early endosomes. Another scenario that one may propose after qualitative observations would lead to a markedly different profile and evolution for $n(s,t)$ and can thus be ruled out. Analogously, the data are only compatible with a constant fusion rate since otherwise the distribution features would be different (see Foret *et al.*, 2012 for details). It indicates that LDL is a passive cargo weakly affecting the fusion machinery. A fit of the data yields the values of the microscopic parameters (entry rate, fusion rate, and conversion rate) that are difficult to measure directly. The success of the entry-fusion-exit model does not mean that fission and heterotypic fusion do not exist but it means that they should have a small influence on the distribution of LDL cargo in early endosomes of HeLa cells.

Regulation of the Microscopic Processes

Cargo molecules change their receptors’ activation state through binding and unbinding which itself is affected by

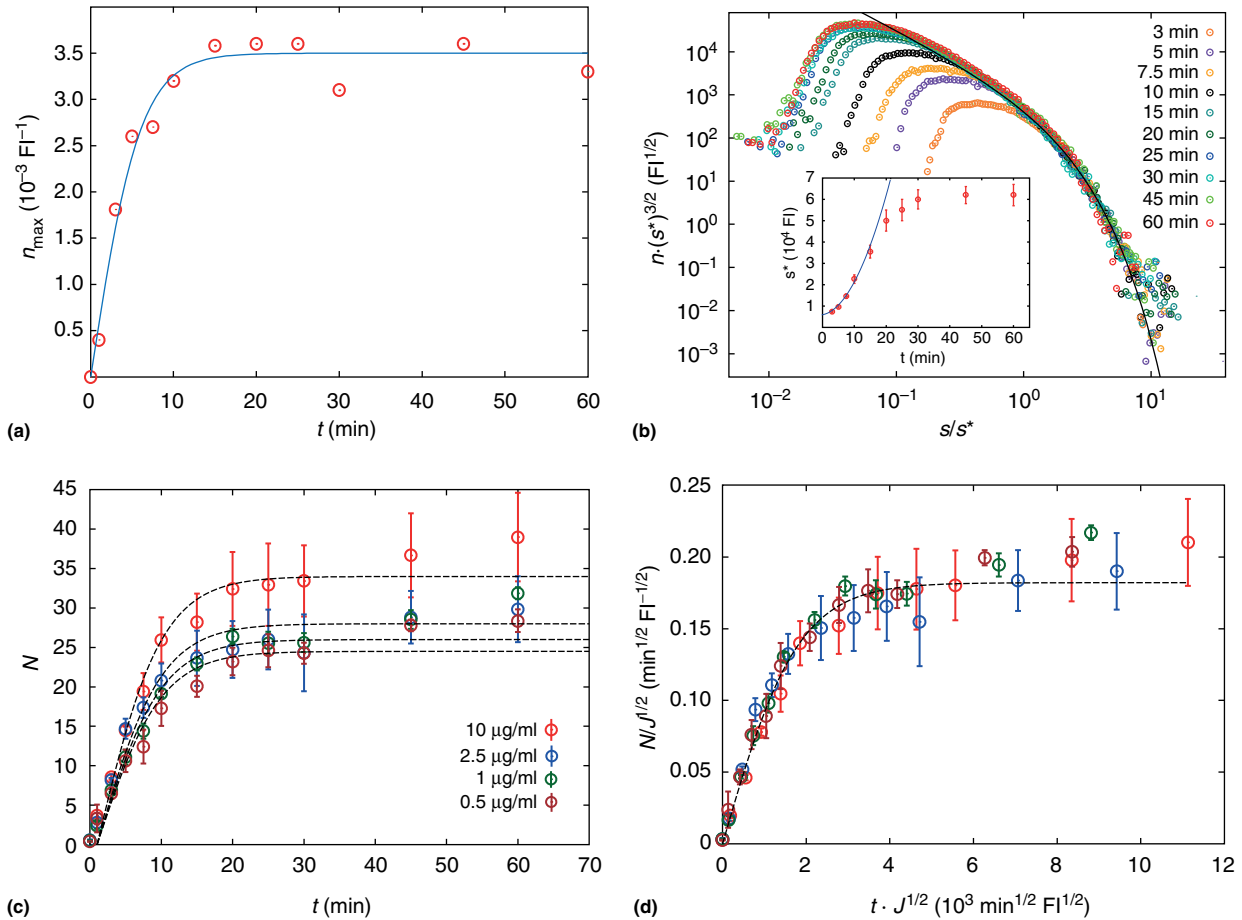


Figure 3 Experimental verification of the theory of cargo trafficking in the endosomal network. (a) Maximum values of the cargo distributions shown in Figure 1(d) as function of time (symbols) verify the theory prediction (solid curve) by eqn. [3]. (b) Rescaled cargo distributions (symbols) from Figure 1(e) collapse on the predicted universal curve shape (eqn. [4]). The scaling parameter s^* (inset) is correctly predicted by eqn. [4], the parabolic limit of the prediction for short times is shown as solid line, the saturation is also correctly predicted. (c) The dependence of the total number of cargo-filled early endosomes on time and medium composition (see legend) is correctly predicted (dashed curves) by eqn. [3]. (d) Rescaling of the data from panel (c) with the predicted influx dependency collapses all data onto the universal time dependency as predicted by eqn. [3]. Reprinted from Foret, F., Dawson, J.E., Villaseñor, R., *et al.*, 2012. A general theoretical framework to infer endosomal network dynamics from quantitative image analysis. *Current Biology* 22, 1381–1390, copyright (2012), with permission from Elsevier.

cargo accumulation within individual endosomes and by gradual acidification of the endosomal lumen (Lauffenburger and Linderman, 1993). The receptors' cytoplasmic domains can transmit the binding signal to effector proteins with guanine nucleotide exchange (GEF) and GTPase-activation (GAP) activities (Tall *et al.*, 2001). Small GTPases of the Rab family act as master regulators of the endocytic protein machinery and can integrate the cargo-dependent GEF and GAP activities with other signals (Zerial and McBride, 2001). For EEs, the GTPase Rab5 plays this central role and increased Rab5 activity fosters cargo uptake (Zerial and McBride, 2001), homotypic EE fusion (Ohya *et al.*, 2009), homotypic EE fission (Zeigerer *et al.*, 2012) and, surprisingly, also EE to late endosome conversion (Rink *et al.*, 2005; Del Conte-Zerial *et al.*, 2008; Poteryaev *et al.*, 2010). The activation of the reciprocal microscopic processes fusion and fission by Rab5 conveys robustness to macroscopic properties such as total endosome number as demonstrated in an integrated theoretical and experimental study of Rab5 titration in hepatocytes

(Zeigerer *et al.*, 2012). Different cargo types may affect Rab5 differently through their specific co-transported receptors. For example the cargoes EGF and LDL have been observed to differently modulate the macroscopic properties of the endosomal network (Collinet *et al.*, 2010). In the theory presented above, the transport of different cargoes can be modeled by readjusting the parameter values for the rates of the microscopic processes. The regulatory functions of certain cargoes as well as possible dependencies of the microscopic processes on the physical size of the endosomes can in the theory be accounted for by the s -dependencies and additional time-dependencies of the six rate functions, see Section 'Microscopic processes and parameters.'

Relation to Kinetics of Total Cargo Pool

The total cargo content $\Phi(t)$ per cell can be obtained from the cargo distribution $\Phi(t) = \int_0^\infty sn(s,t)ds$. The total cargo

influx J into the endosomal network is $J(t) = \int_0^\infty (sA(s,t) + v_{in}(s,t)n(s,t))ds$. A general kinetic equation for $\Phi(t)$ can be derived from eqn. [1] by multiplication with s and integration of both sides of the equation:

$$\frac{d\Phi}{dt} = J(t) - \int_0^\infty (k_d(s,t)sn(s,t) + v_{out}(s,t)n(s,t))ds. \quad [5]$$

Only in special cases with $k_d(s,t) = \text{const}$ and $v_{out}(s,t) = k_{out}s$, the general theory simplifies to the known kinetic equation $d\Phi/dt = J(t) - (k_d + k_{out})\Phi$ (Lauffenburger and Linderman, 1993; Klausner *et al.*, 1985; Tzafirri *et al.*, 2004). However, nontrivial system behavior including delayed forwarding of spikes of cargo or signals can only be accounted for by eqn. [5] which again requires to solve eqn. [1].

Implications for Cargo Sorting

Besides describing degradative cargo trafficking, this theory is also applicable to recycling cargo and membrane turnover itself. The heterotypic fission process (Figure 2(e)) will then play the essential role. Both, homotypic fusion and heterotypic fission of membrane are tightly coupled as shall be pointed out at the end of this section. For arbitrary but qualitatively similar shapes of EEs before and after a fusion event, the volume grows as the third power of a linear dimension whereas the membrane area only grows with the second power leaving excess membrane that can be extracted through membrane tube formation and then be recycled toward the plasma membrane (Skjeldal *et al.*, 2012). The membrane tube separation from a post-fusion EE also facilitates the recycling of receptors and membrane-associated cargo. Cargo that is destined for accumulation and subsequent degradation (e.g., LDL) can be passively sorted from other cargo that is destined for recycling by separating the former cargo from the membrane. For instance, cargo taken up by receptor-mediated endocytosis can dissociate from its receptor upon the pH decrease in EEs compared to the extracellular space. A membrane tube has a particularly low volume to surface ratio and correspondingly little degradative cargo will leak into the recycling pathway. This passive sorting process without the need for receptors specifically localized to the bulk part of an EE is termed geometric sorting (Maxfield and McGraw, 2004). Geometric sorting reaches high fidelity through multiple iterations of fusion and fission among subpopulations (e.g., first among EEs and subsequently among recycling endosomes).

Conclusion

Characteristic macroscopic properties of the endosomal network as a whole emerge from the many stochastic endosome interactions on the microscopic scale. One such macroscopic property of interest is the probability of finding a large concentrated cargo packet versus many small cargo packets within a cell. The underlying endosome interactions have been difficult to investigate due to their highly parallel and stochastic occurrence. Here, a unifying theory has been presented which links the microscopic properties of endosome interactions to the emergent macroscopic properties of cargo trafficking in the

endosomal network. The theory allows to infer microscopic kinetic parameters such as the fusion rate between endosomes from the statistics of the whole endosomal network at fixed time points after cargo internalization, lending this approach to *in vivo* studies that require fixing of samples. The results are robust and sensitive because the cargo statistics average over many stochastic events in many cells.

See also: Molecular Principles, Components, Technology, and Concepts: Membranes: Composition, Physical Properties, and Curvature

References

- Bonifacino, J.S., Glick, B.S., 2004. The mechanisms of vesicle budding and fusion. *Cell* 116, 153–166.
- Cao, Y., Li, H., Petzold, L., 2004. Efficient formulation of the stochastic simulation algorithm for chemically reacting systems. *Journal of Chemical Physics* 121, 4059.
- Collinet, C., Stöter, M., Bradshaw, C.R., *et al.*, 2010. Systems survey of endocytosis by multiparametric image analysis. *Nature* 464, 243–249.
- Conner, S.D., Schmid, S.L., 2003. Regulated portals of entry into the cell. *Nature* 422, 37–44.
- Del Conte-Zerial, P., Brusch, L., Rink, J., *et al.*, 2008. Membrane identity and GTPase cascades regulated by toggle and cut-out switches. *Molecular Systems Biology* 4, 206.
- Foret, F., Dawson, J.E., Villaseñor, R., *et al.*, 2012. A general theoretical framework to infer endosomal network dynamics from quantitative image analysis. *Current Biology* 22, 1381–1390.
- Gorvel, J.-P., Chavrier, P., Zerial, M., Gruenberg, J., 1991. Rab5 controls early endosome fusion *in vitro*. *Cell* 64, 915–925.
- Gruenberg, J., Maxfield, F.R., 1995. Membrane transport in the endocytic pathway. *Current Opinion in Cell Biology* 7, 552–563.
- Hogg, R., Healy, T.W., Fuersten, D.W., 1966. Mutual coagulation of colloidal dispersions. *Transactions of the Faraday Society* 62, 1638–1651.
- Klausner, R., Van Renswoude, J., Harford, J., Wolsky, C., Goldstein, B., 1985. Mathematical modeling of receptor-mediated endocytosis. In: Pastan, I., Willingham, M.C. (Eds.), *Endocytosis*. New York, NY: Plenum Press, pp. 259–277.
- Lauffenburger, D.A., Linderman, J.J., 1993. *Receptors: Models for Binding, Trafficking, and Signaling*. New York, NY: Oxford University Press.
- Maxfield, F.R., McGraw, T.E., 2004. Endocytic recycling. *Nature Reviews Molecular Cell Biology* 5, 121–132.
- Mellman, I., 1996. Endocytosis and molecular sorting. *Annual Review of Cell and Developmental Biology* 12, 575–625.
- Mukherji, S., O'Shea, E.K., 2014. Mechanisms of organelle biogenesis govern stochastic fluctuations in organelle abundance. *eLife* 3, e02678.
- Ohya, T., Miaczynska, M., Coskun, U., *et al.*, 2009. Reconstitution of Rab- and SNARE-dependent membrane fusion by synthetic endosomes. *Nature* 459, 1091–1097.
- Poteryaev, D., Datta, S., Ackema, K., Zerial, M., Spang, A., 2010. Identification of the switch in early-to-late endosome transition. *Cell* 141, 497–508.
- Rink, J., Ghigo, E., Kalaidzidis, Y., Zerial, M., 2005. Rab conversion as a mechanism of progression from early to late endosomes. *Cell* 122, 735–749.
- Rybin, V., Ullrich, O., Rubino, M., *et al.*, 1996. GTPase activity of Rab5 acts as a timer for endocytic membrane fusion. *Nature* 383, 266–269.
- Sheff, D.R., Pelletier, L., O'Connell, C.B., Warren, G., Mellman, I., 2002. Transferrin receptor recycling in the absence of perinuclear recycling endosomes. *Journal of Cell Biology* 156, 797–804.
- Skjeldal, F.M., Strunze, S., Bergeland, T., *et al.*, 2012. The fusion of early endosomes induces molecular-motor-driven tubule formation and fission. *Journal of Cell Science* 125, 1910–1919.
- Smoluchowski, M., 1916. Drei Vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen. *Physikalische Zeitschrift* 17, 557–585.
- Tall, G., Barbieri, M., Stahl, P., Horazdovsky, B.F., 2001. Ras-activated endocytosis is mediated by the Rab5 guanine nucleotide exchange activity of RIN1. *Developmental Cell* 1, 73–82.

- Tzafriri, A.R., Wu, D., Edelman, E.R., 2004. Analysis of compartmental models of ligand-induced endocytosis. *Journal of Theoretical Biology* 229, 127–138.
- Zeigerer, A., Gilleron, J., Bogorad, R.L., *et al.*, 2012. Rab5 is necessary for the biogenesis of the endolysosomal system *in vivo*. *Nature* 485, 465–470.
- Zerial, M., McBride, H.M., 2001. Rab proteins as membrane organizers. *Nature Reviews Molecular Cell Biology* 2, 107–117.
- Ziemelis, K. (Ed.), 2001. Complex systems. *Nature Insight* 410, 241–284.
- Birtwistle, M.R., Kholodenko, B.N., 2009. Endocytosis and signalling: A meeting with mathematics. *Molecular Oncology* 3, 308–320.
- Camazine, S., Deneubourg, J., Franks, N., *et al.*, 2001. Self-Organization in Biological Systems. Princeton, NJ: Princeton University Press.
- Miaczynska, M., Pelkmans, L., Zerial, M., 2004. Not just a sink: Endosomes in control of signal transduction. *Current Opinion in Cell Biology* 16, 400–406.
- Nicolis, G., Nicolis, C., 2012. Foundations of Complex Systems: Emergence, Information and Prediction, second ed. Singapore: World Scientific Publishing.

Further Reading

- Bar-Yam, Y., 2003. Dynamics of Complex Systems. Boulder, CO: Westview Press.

Transcription Factor Networks

AA Jermusyk and GT Reeves, North Carolina State University, Raleigh, NC, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The animal body is a complex system which is regulated by a series of interconnecting networks set up during the initial stages of development. These large transcription factor networks, or gene regulatory networks (GRNs), are composed of sets of smaller networks, or modules, responsible for specific subroutines in the patterning process. For example, in the fruit fly, *Drosophila melanogaster*, two cascading GRNs serve to pattern the axes of the embryo setting up gene expression domains in the adult fly (Reeves and Stathopoulos, 2009; Perkins *et al.*, 2006; Jaeger *et al.*, 2004). Different networks are responsible for correct differentiation later in development. As an example, the GRN map for the dorsal–ventral axis is described in Figure 1. This network features many interactions between different transcription factors to properly pattern the major tissues along this axis in the developing embryo.

The final results of these networks can be seen in the correct differentiation of organisms ranging from fruit flies to sea urchins to humans. While these organisms may seem varied, each uses multiple GRNs consisting of similar wiring strategies. At the mechanistic level, all GRNs are based on the same principles: transcription factors binding to specific DNA sequences and regulating proximate genes. These genes in-turn may code for transcription factors that regulate still other genes.

At a higher level of abstraction, collections of a handful of such genetic interactions may be overrepresented patterns of functional interactions called motifs. While the identity of the genes themselves differ, network motifs with the same wiring have roughly the same function, and make up the complex GRNs seen in all of these animals (Reeves and Stathopoulos, 2009; Milo *et al.*, 2002; Thieffry and Sanchez, 2003; Peter and Davidson, 2009a). Some common motifs are: repression, feedforward loops, positive feedback, negative feedback, and

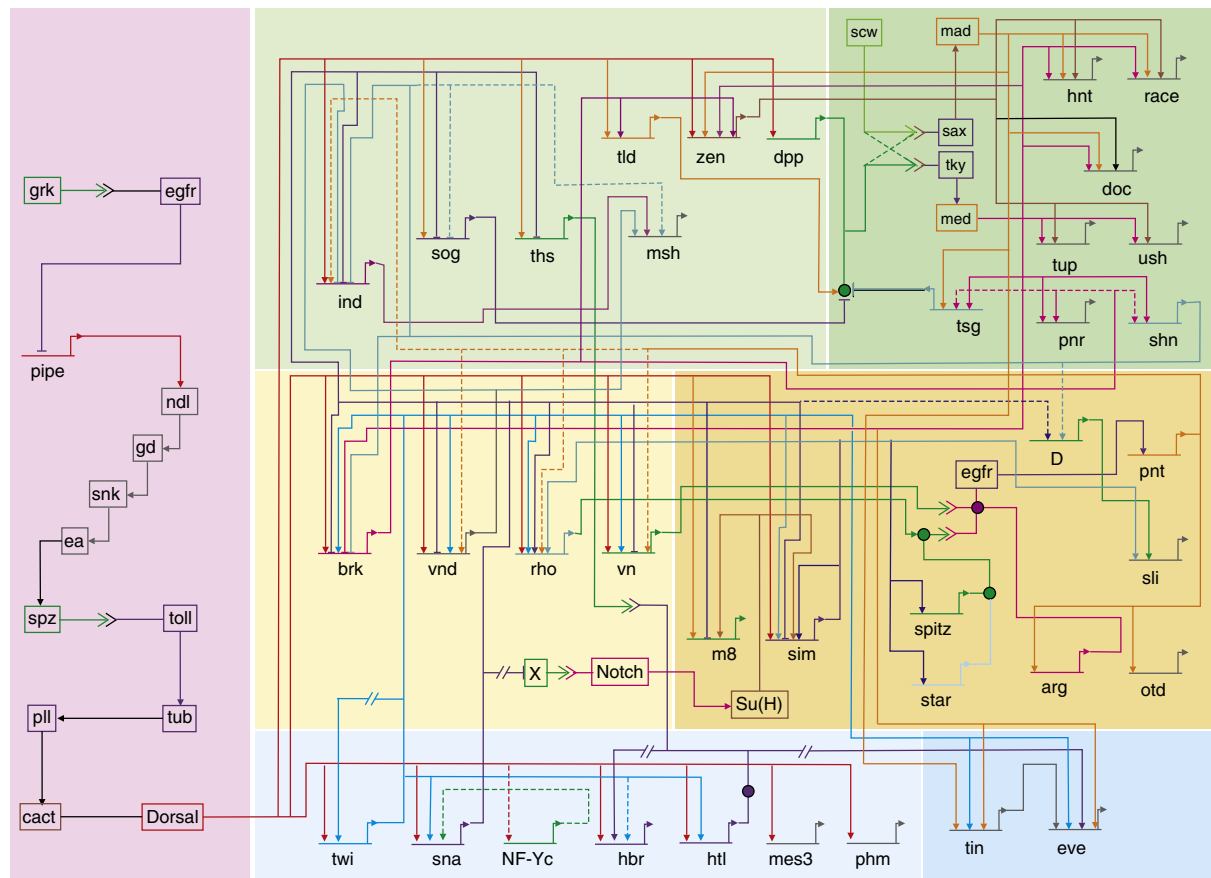


Figure 1 GRN map of the system patterning the dorsal–ventral axis in the *D. melanogaster* early embryo from 2 to 5 h after fertilization. The activity of each transcription factor on the other members of the system is shown. The green, yellow, and blue boxes represent the three types of embryonic tissue. In addition to signaling within a specific tissue type, activity also occurs between tissues. Reprinted from Levine, M., Davidson, E., 2005. Gene regulatory networks for development. *Proceedings of the National Academy of Sciences of the United States of America* 102 (14), 4936–4942.

cross-repression. These motifs serve as building blocks, with each interaction serving a specific functional role in gene regulation. Together these basic interactions serve to create a properly patterned organism.

In addition to studying these components of GRNs, more can also be learned by understanding features of the networks as a whole. Global properties of GRNs, such as the small-world property, the scale-free property, and modularity, are responsible for many advantageous design features of GRNs. For example, the GRN depicted in Figure 1 is modular in that it is divided into parts that perform separate biological tasks (patterning the individual tissue systems). In this article, we will review the common properties of GRNs at each level of abstraction: local, intermediate (motifs), and global. Our focus is on GRNs acting in development, however many of the same features can be applied to other biological contexts, ranging from metabolism to homeostasis.

Local Properties of GRNs

At the most basic level, GRNs are composed of transcription factors specifically binding to DNA sequences to regulate nearby genes. Typically, gene regulation is depicted as occurring 5' to a gene's transcriptional start site, but these regulatory elements could be downstream of the gene, as well as within an intron of a gene, or even within a neighboring gene. *cis*-regulatory elements (i.e., the DNA sequences to which transcription factors bind) often contain clusters of several binding sites for distinct transcription factors that each converge to determine the transcriptional state of a gene (Figure 2 (a)). In this section, we'll discuss how these interactions are experimentally verified, and how they are modeled.

Constructing GRN Maps

To construct a GRN map, one must proceed at the local level, with the verification of direct interactions between transcription factor and the regulated gene (Peter and Davidson, 2009b). If an organism's genome is sequenced, part of this approach may include a computational search for known binding site sequences. However, a computational approach is rarely sufficient by itself, due to the fact that most binding sites are short, with significant ambiguity at several positions, making the frequency of purely sequence-based hits prohibitively high, with many false positives. Furthermore, as mentioned above, a gene may be regulated by *cis*-elements that may be a significant distance from the transcriptional start site.

Several lines of experimental evidence can be employed in verifying a gene regulatory interaction. Direct evidence of transcription factor binding is valuable. One way to find regions bound by transcription factors is ChIP (Chromatin Immunoprecipitation)-based experiments. However, while these data reveal bona-fide physical interactions, it is not always clear which of these interactions are functional, resulting in false positives.

Correlative evidence in gene expression provides a solid foundation for determining the veracity of a genetic interaction. It should be verified the transcription factor is present in the same set of cells as the effect (either transcription or

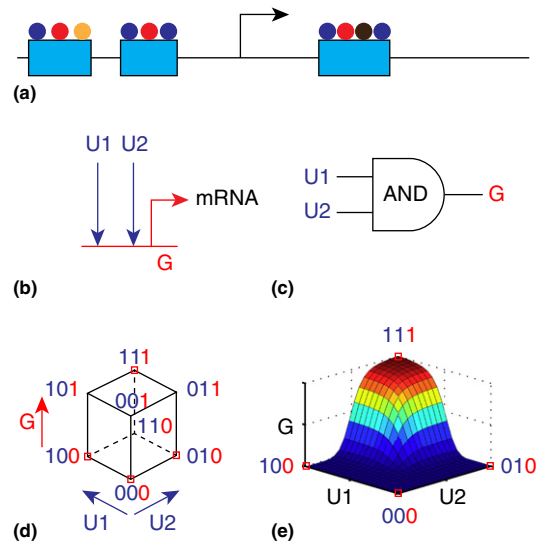


Figure 2 Local properties of transcriptional regulation. (a) Structure of *cis*-regulation. *cis*-Regulatory elements, represented by blue boxes, contain clusters of binding sites for several transcription factors (colored circles). These clusters can be close to, far from, or even downstream of the transcriptional start site (represented by arrow). (b) Example of gene, G, with two inputs, U1 and U2. (c) Regulation of G can be represented by Boolean logic. In this example, both inputs are necessary for gene expression. The logic is represented by an AND gate. (d) Cube representation of the eight possible states for the variables U1, U2, and G. Each vertex represents a state in which the variables are either 1 or 0 (inputs in blue, output in red). The only allowable states are indicated by red squares. (e) Continuous representation of gene expression output. The two inputs can vary continuously from zero to maximal, resulting in a continuous variation in G. Often, the input/output relationship is taken to be ultrasensitive, which approximates a logical on/off system. ((b)–(e)) Modified from Bolouri, H., Davidson, E.H., 2002. Modeling transcriptional regulatory networks. *Bioessays* 24 (12), 1118–1129. Copyright © 2002, Wiley periodicals, Inc. Used with permission.

repression of the target gene, depending on the character of the putative regulation). In a similar manner, perturbing transcription factor expression (through knockout mutation, ectopic expression, or RNA interference) should have the expected effect on gene expression.

Enhancer trap analysis, which tests the sufficiency of a stretch of DNA to control a gene expression pattern, is one of the most important tools for verifying the role of the DNA sequence in driving gene expression. In this method, the *cis*-regulatory element in question is cloned upstream of a minimal promoter and a marker gene (commonly bacterial *lacZ*, as this gene is not natively present in metazoans). If transgenic expression of this DNA cassette results in the expression of *lacZ* in the pattern of the endogenous gene from which the putative enhancer was derived, this is strong positive evidence the enhancer in question regulates gene expression. Further evidence for the action of a specific transcription factor on this enhancer would include mutating the transcription factor binding sites within the transgenic enhancer and monitoring the resulting change in *lacZ* expression.

In all, the identification of the regulatory interactions of a single gene can be laborious and time-consuming

(Yuh *et al.*, 2001). In this manner, developmental GRN maps, such as for the sea urchin endomesodermal network and the *Drosophila* dorsal–ventral and anterior–posterior patterning networks, have resulted from many years of painstaking experiments performed by a large number of scientists.

Modeling *cis*-Regulatory Interactions

Once a set of *cis*-regulatory interactions have been experimentally identified for a given gene, how should one predict the effect transcription factor expression has on transcriptional output? Often, these interactions are modeled using Boolean on/off logic (Bolouri and Davidson, 2002). From a digital point of view, gene regulation can be seen as a type of logic gate. For example, take a gene, *G*, regulated by two transcription factors, *U1* and *U2* (Figure 2(b)). If this gene requires both activators to be present in order to be activated, this is taken as an AND gate (Figure 2(c)). The presence (or absence) of the transcription factors sets *U1* or *U2* to a value of one (or zero), resulting in the output being either on or off (Figure 2(d)). Another common example is an OR gate, in which the presence of either transcription factor is sufficient for transcriptional activation. If there are competing inputs (if *U1* were an activator and *U2* a repressor), this could be modeled by logic such as ‘*U1* AND NOT *U2*.’

The digital approach has the advantage of needing to know only the wiring topology and logic of the network; binding affinities and other biophysical parameters are not needed. As such, it is a very coarse-grained approach. When the purely digital logic approach does not have sufficient detail, a continuous approach may be taken (Bolouri and Davidson, 2002). For example, in the case of a series of morphogen-mediated gene patterns, a concentration gradient (analog signal) of a protein is read-out by cells in the tissue to be patterned. In such cases, often the relationship between transcription factor(s) and transcriptional output may be modeled as a highly sigmoidal (ultrasensitive) function (Figure 2(e)). Such models seek only to capture the phenomenology of transcriptional regulation, and most of the biophysical interactions are grouped into a handful of parameters. If the response is sensitive enough, these models approach the switch-like digital case.

If more mechanistic detail is desired, thermodynamic approaches have also been employed, in which the probability of transcription is related to the equilibrium of transcription factor binding site occupancy (Lai *et al.*, 2004; Reeves *et al.*, 2006; Saha and Schaffer, 2006; Zinzen *et al.*, 2006). These models tend to be less switch-like, and require detailed knowledge of the binding affinities and interactions among transcription factors at the *cis*-regulatory locus. Each of these approaches (digital/topological, phenomenological, and mechanistic/thermodynamic) has advantages and should be employed based on the level of input detail known, the level of output detail desired, and computing power available.

Motifs

At an intermediate level of abstraction, the interactions among three or more genes can serve a basic functional purpose, regardless of the identity of the genes involved. These network

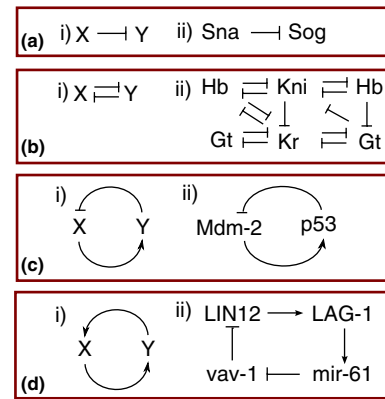


Figure 3 Some of the most common motifs including: repression (a), cross-repression (b), negative feedback (c), and positive feedback (d). (a) Repression motif, one protein (*X*) represses the expression of a second protein (*Y*), this can be seen in dorsal–ventral patterning in the *Drosophila* embryo between *Sna* and *Sog* (Stathopoulos and Levine, 2002; Maduro and Rothman, 2002). (b) The cross-repression motif features mutual inhibition between two proteins (*X* and *Y*) and occurs between *Hb*, *Kni*, *Gt*, and *Kr* in the anterior–posterior patterning system in the *Drosophila* embryo (Jaeger, 2011). (c) Negative feedback features one protein (*X*) activating a second protein (*Y*) where the second protein represses the expression or inhibits the function of the first protein; this motif is seen in the tumor regulatory pathway between *Mdm-2* and *p53* (Oren, 1999). (d) Positive feedback is similar to negative feedback except the second protein (*Y*) activates or facilitates the first protein (*X*); this can be seen in the vulval formation pathway in *C. elegans* where *LIN-12* activates *LAG-1*, which in turn activates *LIN-12* through activation of *mir-61* followed by repression of *vav-1*, which in turn represses *LIN-12*.

interactions, or motifs, do not often act as functionally separable from the rest of the network, yet nevertheless have well-defined kinetic behavior (Babu *et al.*, 2004). Several motifs have been discovered to be overrepresented sets of connections (Milo *et al.*, 2002), implying the functional sub-routines carried out by these motifs may have an important, general purpose in regulation of gene expression.

Some common motifs (Figure 3) are repression, cross-repression, negative feedback, positive feedback, and feedforward loops. These motifs may seem very different, and in fact can be used in very different ways; they all function to increase the ability of an organism to create reproducible patterns despite perturbations and obstacles to proper development. In each of these systems, transcription factors bind to other transcription factors or to *cis*-regulatory elements to activate or repress expression of other transcription factors. These interactions are seen throughout development in a variety of systems. Our understanding of how these interactions can be grouped into different motifs to produce a given result in one system can be applied more broadly to a similar group of interactions, or the same motif, in another system.

Repression

Perhaps the simplest motif found during development is repression. During patterning, morphogen gradients activate expression in large domains based on concentration thresholds. In many instances, expression domains must be restricted

to limited regions, which can be achieved through repression by a neighboring protein (Stathopoulos and Levine, 2002; Maduro and Rothman, 2002). For example, along the dorsal-ventral axis of the *Drosophila* embryo, Snail is expressed in a narrow domain at the ventral midline and represses *short gastrulation* (*sog*) (Stathopoulos and Levine, 2002). This repression by Snail eliminates *sog* expression in the ventral-most cells, where Snail is present, dividing *sog* expression into two broad lateral stripes (Stathopoulos and Levine, 2002). Just like in the *Drosophila* embryo, *Caenorhabditis elegans* need to control the spatial domain of gene expression in endoderm development (Maduro and Rothman, 2002). In this system, SKN-1 activates expression, whereas POP-1 represses expression, creating a switch that serves to direct tissue to either an endoderm or mesoderm fate, thereby limiting the boundaries of the mesoderm (Maduro and Rothman, 2002). Similar regulation occurs in patterning the *Drosophila* egg and heart just to name a few of the diverse uses for repression (Cripps and Olson, 2002; Wasserman and Freeman, 1998). In each of these instances an input signal activates expression over a large region, and repression is used to eliminate this expression in certain region or after a certain period of time.

Cross-repression

Mutual inhibition, or cross repression, can be used to produce stripes or clear borders between gene expression regions in many different organisms. In the developing *Drosophila* embryo, one of the first steps to patterning of the anterior-posterior axis is the proper positioning of the so-called gap genes, namely Knirps, Hunchback, Giant, and Krüppel (Jaeger, 2011). The initial expression of these gap genes are established by the morphogen Bicoid in large overlapping domains (Jaeger, 2011). Cross repression between neighboring genes refines these domains and creates sharp borders between the domains (Jaeger, 2011). Since each region of transcription factor inhibits the neighboring transcription factor, two stable regions of expression are produced, which easily respond to perturbations near the interface of expression (Jagla et al., 2002). Similar cross repression activity occurs in the specification of cardiac and muscle cell fates in the *Drosophila* embryo (Jagla et al., 2002). These sharp domain borders are needed to create clear limits on downstream gene expression and therefore delineate positioning of tissue differentiation.

Mutual inhibition also regulates the branching of cells in mammalian lungs, mammary epithelial cells, trachea in flies, and vasculature in flies (Lu and Werb, 2008). In the fly trachea, mutual inhibition regulates the number of branches by determining the leading cell through competition for branch-inducing factors (Lu and Werb, 2008). Similarly, in mammary epithelial cells, mutual inhibition functions as a mechanism for cells to recognize surrounding cells and avoid other branches; this results in new branches turning away from the current line of growth and branches to end as they approach another branch (Lu and Werb, 2008). The use of mutual inhibition by branching cells functions similarly to the earlier examples of defining borders of stripes of gene expression in the developing embryo. In these diverse systems, mutual inhibition or cross repression provides a method for cells to recognize neighboring cells and to stabilize the interactions

between these cells by controlling the position and number of branches in a tissue or by controlling the downstream signal a region will receive prior to tissue differentiation.

Negative Feedback

Negative feedback can be used to buffer a system against perturbations, therefore making it more robust (Ma et al., 2009). In mice, a negative feedback loop featuring Mdm-2 regulates the degradation of p53, a tumor suppressor gene (Oren, 1999). In this loop, p53 binds to the *mdm-2* gene, stimulating the production of Mdm-2 protein, which then binds to p53 to inactivate and stimulate p53 degradation (Oren, 1999). The result of this contradictory relationship is a tight control over p53 expression, which is important since improper control of p53 can result in cancer (Oren, 1999). A similar negative feedback loop exists in embryonic stem cells to control pluripotency and self-renewal (Pan et al., 2006). In these systems negative feedback serves to limit gene expression thereby adding a level of control to the system; the importance of this control is apparent in the prevention of cancer, but is also important in a number of different systems.

Alternatively, oscillations can also be produced using negative feedback. Oscillations result in periodic expression of genes in time or space. This can serve as a method to produce evenly spaced stripes in a field as the tissue is created from replicating cells. During vertebrate somitogenesis, multiple negative feedback loops produce oscillations in Fgf, Wnt, and Notch which control the segmentation clock mechanism (Aulehla and Herrmann, 2004; Gibb et al., 2010). This creates periodic gene expression, which results in each segment forming at specific time intervals and therefore in specific positions in the tissue (Aulehla and Herrmann, 2004; Gibb et al., 2010). This oscillatory behavior is especially important in tissue that experiences directional development, creating stripes of expression as the tissue is created. In addition oscillations can also be used to provide a signal between pathways, eliminating problematic cross-talk. This occurs through negative feedback between NF κ B and I κ B in mammalian cells which produced oscillations in the translocation of NF κ B (Nelson et al., 2004; Lo et al., 2009). This produces an analog to digital signal to encode cell fate based on the frequency of oscillations rather than the amplitude of the signal (Nelson et al., 2004; Lo et al., 2009).

In some instances it can be desirable to limit the length of time or range of a signal; this effect can also be accomplished using negative feedback. In this case when a threshold is reached the signal can be turned off using a negative regulator; this type of regulation occurs in the JAK/STAT pathway (Janus kinase/signal transducers and activators of transcription) in vertebrates (Rawlings et al., 2004). This provides a valuable method to control signaling by turning off expression as a new stage in development is reached, it can also be used in regulation of tissue to turn-off a signal after the perturbation to the system has been removed.

Negative feedback can be used to achieve multiple objectives. Most notably is to buffer the system against perturbations by interlocking the expression of two genes and thereby preventing fluctuations from being carried over into downstream gene expression. To create properly patterning tissues, negative

feedback, through oscillations, can create reproducible bands of gene expression. These are just a few of the many ways negative feedback is used to produce properly differentiated tissues.

Positive Feedback

Positive feedback can also provide robustness to perturbations to the system. This robustness does not sacrifice sensitivity to the input signal (Hornung and Barkai, 2008; Kadelka *et al.*, 2013). In many systems switch-like behavior is used to achieve this robustness and produce precise borders, by creating bistable systems (Shah and Sarkar, 2011; Siegal-Gaskins *et al.*, 2011; Graham *et al.*, 2010). In bistable systems two equilibrium states can be achieved, with a small perturbation, the system will settle into a different state. This can be seen in endomesoderm specification in sea urchins (Minokawa *et al.*, 2005; Yuh *et al.*, 2004). In this system, initial β -catenin activates *wnt8* and *blimp1/krox* expression; this in turn causes nuclearization of β -catenin and therefore more *wnt8* and *blimp1/krox* to be expressed (Minokawa *et al.*, 2005). In this case, relatively low levels of β -catenin are needed to create a stable domain of Wnt8 expression (Minokawa *et al.*, 2005). Positive feedback is used to create bistable systems in other systems such as the mouse hindbrain, *Drosophila* wing, and *Xenopus* oocyte (Barrow *et al.*, 2000; Nonchev *et al.*, 1996a,b; Vesque *et al.*, 1996; Yan *et al.*, 2009; Xiong and Ferrell, 2003).

In a variety of organisms, similar systems exist where positive feedback produces precise regions of expression following an input signal. This can be seen, for instance in the control of vulval formation in *C. elegans* development (Yoo and Greenwald, 2005). In this system LIN-12 activates LAG-1 expression which goes on to activate mir-61 expression (Yoo and Greenwald, 2005). In turn, mir-61 represses *vav-1* which is a repressor of *lin-12* (Yoo and Greenwald, 2005). While varying in construction, positive feedback loops occur in many different systems in nature, including *C. elegans* neuron differentiation, *Drosophila* gut patterning, *Drosophila* eye development, and the interferon pathway in animal immune response (Johnston *et al.*, 2005; Bienz and Tremml, 1988; Johnston *et al.*, 2011; Taniguchi and Takaoka, 2001). In each of these cases positive feedback is able to use relatively small initial signals to effectively pattern tissue from stable domains of gene expression.

Feedforward Loops

Feedforward loops can help a system to respond to perturbations. Feedforward loops come in two flavors: coherent and incoherent. For coherent feedforward loops, the direct effect gene X has on gene Z is the same as the effect of gene X on gene Z indirectly through gene Y, as shown in Figure 4. With incoherent feedforward loops, it is the opposite. Incoherent feedforward loops can be very useful to reset a system after it responds to a perturbation (Ma *et al.*, 2009). These incoherent feedforward loops are found to regulate multiple systems in yeast and *Escherichia coli* (Milo *et al.*, 2002; Shen-Orr *et al.*, 2002).

Incoherent feedforward loops can also be used to produce stripes in many different systems including the *Drosophila*

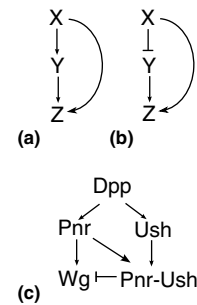


Figure 4 Feedforward loops can take the form of either coherent (a) or incoherent (b). In a coherent feedforward loop, the direct effect gene X has on gene Z is the same as the effect of gene X on gene Z indirectly through gene Y. This can be accomplished in a simple loop where X activates Y and Z, and Y goes on to activate Z (a). However in the incoherent feedforward loop, the effect gene X has on gene Z indirectly through gene Y is opposite the effect it has on gene Z directly. This can be accomplished by altering the network such that X represses Y (b). (c) An incoherent feedforward loop regulates decapentaplegic regulation of Wingless in the *Drosophila* notum.

notum and *Drosophila* eggshell (Hironaka *et al.*, 2012; Yakoby *et al.*, 2008). The feedforward loop allows for gene expression far from the source of the activation signal, at low morphogen levels (Hironaka *et al.*, 2012). In the *Drosophila* notum, the morphogen Decapentaplegic (Dpp) activates Pannier (Pnr) and U-shaped (Ush) production (Hironaka *et al.*, 2012). Pnr and Ush can form a complex to repress *wingless* (*wg*) expression or unbound Pnr can activate *wg* expression (Hironaka *et al.*, 2012). Together these two paths combine to form a stripe of Wg at low levels of Dpp (Hironaka *et al.*, 2012). In each of these cases, the feedforward loop allows for gene expression at low levels of the input signal at the periphery of the differentiating field.

In some instances an initial signal is only available at the start of development. In these cases it is desirable to amplify this signal and provide a sustained signal output. This effective signal duration can be achieved using a feedforward loop. In patterning the dorsal–ventral axis of the developing *Drosophila* embryo, a combination of coherent and incoherent feedforward loops signal from Dorsal through either Snail or Twist to control expression of multiple genes in the neurogenic ectoderm (Zartman and Shvartsman, 2007). In each of these cases the feedforward signal is used because it creates a mechanism to utilize a small input signal and transform it.

Global Properties of GRNs

The global properties exhibited by GRNs are largely dictated by the topology of the wiring of the network. Studies of the global properties of GRNs have often visualized GRNs as ‘directed graphs,’ with genes being the nodes, and regulation of one gene on another being a directed connection. By and large, GRNs have been found to adhere to several general topological features, meaning that GRNs studied to date display many common properties. In this section, we will review three global topological features – the small-world property, the scale-free property, and modularity – and discuss what each these features imply.

Small-World Networks

Many biological networks – including metabolic networks, neural networks, as well as GRNs – have been found to be of the ‘small-world’ character (Albert, 2005). Several types of human networks also exhibit this property, such as social networks, power grids, and connections among actors in Hollywood films (Watts and Strogatz, 1998b). Small-world networks are ones in which any two nodes are connected by a relatively short series of paths, yet the probability of any two given nodes being directly connected to each other is relatively high. In other words, small-world networks achieve a balance between maintaining a low degree of separation (commonly referred to as ‘six degrees’ of separation) and having a high degree of clustering (or ‘cliquishness’).

In contrast to small-world networks, one may have a completely regular network, or lattice network, in which all nodes are connected only to their ℓ nearest neighbors. This network has a high degree of clustering, but to move from one node to another on the other side of the network requires many local steps. One simple way to construct a small-world network is to begin with a lattice network, and make a few ‘short circuit’ connections across the network (Watts and Strogatz, 1998b; see Figures 5(a) and 5(b)).

On the other hand, one may have a completely random network, where the probability of two nodes being directly connected is uniform across all nodes. In such a network, the idea of clustering is completely lost, yet the shortest distance between any two nodes is small, allowing one to cross from one node to another in a relatively small number of steps.

Small-world networks have the advantage of keeping both qualities of local clustering and global connectivity. This allows for rapid propagation of information among all nodes in the network, yet also allows for groups of nodes to act together in a similar, local task.

Scale-Free Networks

Besides focusing on the paths and connections between nodes, another way to characterize a network is to focus on how many connections each node has. Early network studies assumed the number of node connections conformed to a Poisson distribution, and thus was strongly peaked at an average value $\langle k \rangle$. This implied that the number of nodes with more than $\langle k \rangle$ connections decayed exponentially and that there were few if any nodes with many more than $\langle k \rangle$ connections. In that sense, the ‘scale’ of the network was $\langle k \rangle$ connections per node.

However, as storage and computing power increased, and topological data of large networks became more available, it was understood that few if any naturally occurring or man-made networks followed a Poisson distribution in node connectivity. Instead, the probability that a node has k connections decays as a power law. $P(k) \propto k^{-n}$ (Barabasi et al., 1999). This result implied that there were a comparatively large number of nodes with higher-than-average connectivity. In this sense, there was no strongly preferred number of node connections, and thus such networks have been dubbed ‘scale-free’ (Barabasi et al., 1999).

Intuitively, a scale-free network is one in which most nodes are connected, on average, to a small number of other nodes.

However, there are a small but significant number of nodes that are highly connected (see Figure 5(c)). These ‘hubs’ hold most of the network together, and at a higher level of connectivity, there are still larger hubs that connect the smaller hubs together. In this sense, the scale-free property has much overlap with the small-world property, but they are not identical. For example, both types of networks display a relatively low degree of separation. In scale-free networks, this property is provided by hubs, or master regulators, that connect parts of the network together, whereas in small-world networks, this property could be achieved simply by making a few random connections across the network (Watts and Strogatz, 1998b; Barabasi et al., 1999). On the other hand, a necessary component of a small-world network is local clustering, yet this property is not required to be a scale-free network. For example, compare the two scale-free networks in Figures 5(c) and 5(d).

As with the small-world property, there is a long list of both man-made and biological networks that exhibit the scale-free property, including GRNs. The scale-free property is advantageous for these networks, as it imparts a high degree of fault- and failure-tolerance (Albert et al., 2000; Jeong et al., 2000). This is because if a gene is randomly disabled by mutation, it is likely to have been a peripheral gene rather than a highly connected hub. On the other hand, even mutations to single hubs can sometimes be tolerated, due to the connectedness of the network through the rest of the hubs. This property is also advantageous from a medical perspective. For example, if several hubs can be intelligently targeted, such as in cancer therapy, it is possible to completely destroy network operation. None of this is to say that mutating a single essential gene, whether peripheral or highly connected, will not have disastrous consequences; indeed, that is the natural consequence of the fact that genes have important functions in the cell. Instead, the point is a scale-free network is designed in such a way as to avoid cascading failures, from the network point of view.

Modularity

As we have discussed before, sets of interactions between a handful of genes are typically characterized as ‘motifs,’ and the topology of the motif determines its kinetic behavior. At a higher level of abstraction, several motifs combine together to form systems of motifs called modules. The modularity of GRNs means that small ‘communities’ of genes have a high level of interconnectedness and together perform a related function, such as cellular differentiation. These modules are further connected (sparsely) to each other to compose the entire GRN.

It is clear that the modularity of a GRN is related to the scale-free property, as the hubs in a scale-free sense serve as master regulators connecting several modules. However, these properties are not identical. It is possible to have a scale-free network in which local clustering is not maintained, resulting in a non-modular (and non-small-world) network (Figure 5(c)). Even in scale-free networks with high degrees of clustering, it may be difficult to identify separate modules if there is a significant degree of cross-connectivity among groups (Figure 5(d)). Modular networks may also violate the

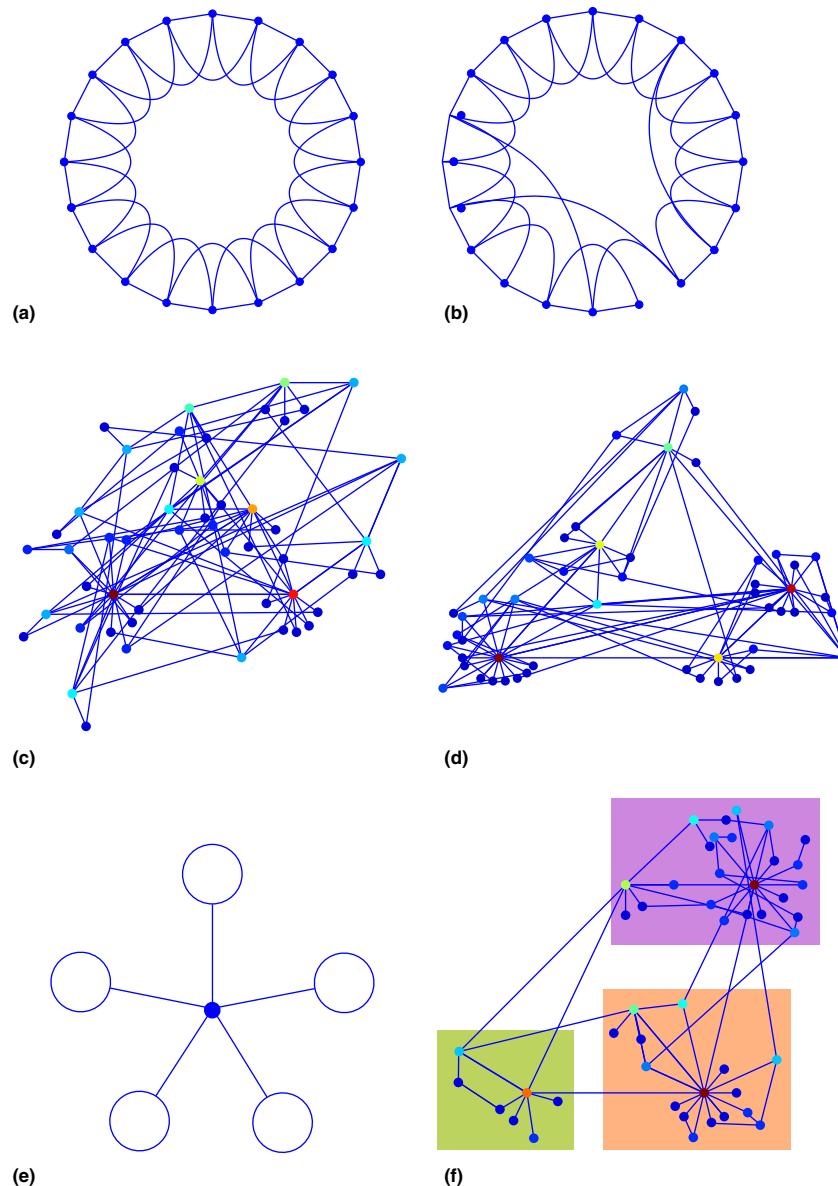


Figure 5 Illustrations of the small-world, scale-free, and modular properties of a network. (a) Lattice network. Each node is connected to its nearest and next-nearest neighbors. (b) Small-world network, made from randomly rewiring three connections in the lattice network. Because the vast majority of nodes have four connections, this is not a scale-free network. (c) Scale-free network. Most nodes have very few connections, but there are a significant number of nodes with many connections. However, this scale-free network was randomly wired, so has very low clustering, and thus is not small-world. Node color code: From dark red (most connections), through red, orange, yellow, green, blue, and finally to dark blue (fewest connections). (d) Scale-free and small-world network. Connections in this network were made preferentially so that clustering is high (a node's neighbors will also likely be connected to each other). However, cross-connections between local groups are too high for this network to be considered modular. Color code same as in (c). (e) Modular, lattice network. Each open circle at the end of a spoke represents a lattice network as in (a). The regular lattice structure rules out this network from being small-world or scale-free, even though it is modular. (f) Small-world, scale-free, modular network. Clustered groups of nodes can be grouped into modules (light green, light orange, and light purple) that have relatively few connections between them. Color code is same as in (c). (a) and (b) adapted from Watts, D.J., Strogatz, S.H., 1998a. Collective dynamics of 'small-world' networks. *Nature* 393 (6684), 440–442. Macmillan Publishers Ltd. Copyright © 1998. Used with permission.

small-world property. The network in [Figure 5\(e\)](#) is modular, but is neither scale-free nor small-world. In contrast, the network depicted in [Figure 5\(f\)](#) is a good example of a scale-free, small-world, modular network: each highlighted group of nodes is tightly clustered, largely separate from the others, and is centered on a few highly connected hubs, yet there are

enough connections among the modules to preserve a low degree of separation.

It is also important to note the modularity of a network depends not only on its connectivity, but also on the specific function of the genes in the local cluster. Local clusters of genes must act together in a specific function for a network to

be considered modular. For example, in sea urchin development, there are distinct modules, composed of groups of interacting genes, that separately control endoderm, mesoderm, and ectoderm development (Peter and Davidson, 2009b). It is evident that genes within a developmental module must be co-expressed in a subset of cells at a particular stage of development.

In the realm of human-designed systems, building objects with modularity is a well-established engineering principle. As such, the modularity of GRNs reveals an analogy to engineered systems. For example, human engineers design hardware (or software) subsystems, with self-contained purposes, that can be 'swapped out' for different applications, such as the modular bay on a laptop, which can be used for an optical drive, an extra hard drive, or an extra battery, among other uses. Thus, for GRNs, modularity provides adaptability, in which a subset of the GRN can be slightly modified without largely affecting the rest of the network. As a side benefit, the modularity of GRNs allows for more facile study by curious scientists.

Conclusions

During development organisms utilize networks of interacting genes, known as GRNs, to properly pattern the organism. Several different properties can be used to describe the overall construction of these GRNs, such as being modular, scale-free, and small-world. In addition to looking at these GRNs as a whole, more can be learned about these networks by looking at the smaller functional elements, known as motifs, which make up these GRNs. These motifs accomplish a diverse range of specific objectives within the larger system. These objectives include: responding to perturbations, amplifying a short-term signal, creating sharp borders between expression regions, and producing oscillations. These differing objectives can be accomplished using motifs such as repression, cross-repression negative feedback, positive feedback, and feedforward loops. In addition to being used for multiple objectives, the same motifs are used by many different systems and occur frequently throughout nature to pattern tissues in many different organisms. Because of this, the information learned about these motifs in one organism can be applied to the study of other organisms as well. This ability to apply knowledge across organisms provides a valuable tool in understanding gene regulation in general as well as in a particular organism of interest.

See also: Cell Division/Death: Cell Cycle: Regulation of the p53 Pathway

References

- Albert, R., 2005. Scale-free networks in cell biology. *Journal of Cell Science* 118 (Pt 21), 4947–4957.
- Albert, R., Jeong, H., Barabási, A.-L., 2000. Error and attack tolerance of complex networks. *Nature* 406 (6794), 378–382.
- Aulehla, A., Herrmann, B., 2004. Segmentation in vertebrates: Clock and gradient finally joined. *Genes & Development* 18 (17), 2060–2067.
- Babu, M., Luscombe, N., Aravind, L., Gerstein, M., Teichmann, S., 2004. Structure and evolution of transcriptional regulatory networks. *Current Opinion in Structural Biology* 14 (3), 283–291.
- Barabási, A., Albert, R., Jeong, H., 1999. Mean-field theory for scale-free random networks. *Physica A* 272 (1–2), 173–187.
- Barrow, J., Stadler, H., Capecchi, M., 2000. Roles of Hoxa1 and Hoxa2 in patterning the early hindbrain of the mouse. *Development* 127 (5), 933–944.
- Bienz, M., Tremml, G., 1988. Domain of ultrabithorax expression in *Drosophila* visceral mesoderm from auto-regulation and exclusion. *Nature* 333 (6173), 576–578.
- Bolouri, H., Davidson, E.H., 2002. Modeling transcriptional regulatory networks. *Bioessays* 24 (12), 1118–1129.
- Cripps, R., Olson, E., 2002. Control of cardiac development by an evolutionarily conserved transcriptional network. *Developmental Biology* 246 (1), 14–28.
- Gibb, S., Maroto, M., Dale, J.K., 2010. The segmentation clock mechanism moves up a notch. *Trends in Cell Biology* 20 (10), 593–600.
- Graham, T.G.W., Tabei, S.M.A., Dinner, A.R., Rebay, I., 2010. Modeling bistable cell-fate choices in the *Drosophila* eye: Qualitative and quantitative perspectives. *Development* 137 (14), 2265–2278.
- Hironaka, K., Iwasa, Y., Morishita, Y., 2012. Multiple feedback loops achieve robust localization of wingless expression in *Drosophila* notum development. *Journal of Theoretical Biology* 292, 18–29.
- Hornung, G., Barkai, N., 2008. Noise propagation and signaling sensitivity in biological networks: A role for positive feedback. *PLoS Computational Biology* 4 (1), e8.
- Jaeger, J., 2011. The gap gene network. *Cellular and Molecular Life Sciences* 68 (2), 243–274.
- Jaeger, J., Surkova, S., Blagov, M., et al., 2004. Dynamic control of positional information in the early *Drosophila* embryo. *Nature* 430 (6997), 368–371.
- Jagla, T., Bidet, Y., Da Ponte, J., Dastugue, B., Jagla, K., 2002. Cross-repressive interactions of identity genes are essential for proper specification of cardiac and muscular fates in *Drosophila*. *Development* 129 (4), 1037–1047.
- Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., Barabási, A.L., 2000. The large-scale organization of metabolic networks. *Nature* 407 (6804), 651–654.
- Johnston, R., Chang, S., Etchberger, J., Ortiz, C., Hobert, O., 2005. MicroRNAs acting in a double-negative feedback loop to control a neuronal cell fate decision. *Proceedings of the National Academy of Sciences of the United States of America* 102 (35), 12449–12454.
- Johnston Jr., R.J., Otake, Y., Sood, P., et al., 2011. Interlocked feedforward loops control cell-type-specific rhodopsin expression in the *Drosophila* eye. *Cell* 145 (6), 956–968.
- Kadelka, C., Murrugarra, D., Laubenbacher, R., 2013. Stabilizing gene regulatory networks through feedforward loops. *Chaos* 23 (2), 025107.
- Lai, K., Robertson, M.J., Schaffer, D.V., 2004. The Sonic hedgehog signaling system as a bistable genetic switch. *Biophysical Journal* 86 (5), 2748–2757.
- Lo, W., Chou, C., Gokoffski, K.K., et al., 2009. Feedback regulation in multistage cell lineages. *Mathematical Biosciences and Engineering* 6 (1), 59–82.
- Lu, P., Werb, Z., 2008. Patterning mechanisms of branched organs. *Science* 322 (5907), 1506–1509.
- Ma, W., Trusina, A., El-Samad, H., Lim, W.A., Tang, C., 2009. Defining network topologies that can achieve biochemical adaptation. *Cell* 138 (4), 760–773.
- Maduro, M., Rothman, J., 2002. Making worm guts: The gene regulatory network of the *Caenorhabditis elegans* endoderm. *Developmental Biology* 246 (1), 68–85.
- Milo, R., Shen-Orr, S., Itzkovitz, S., et al., 2002. Network motifs: Simple building blocks of complex networks. *Science* 298 (5594), 824–827.
- Minokawa, T., Wikramanayake, A., Davidson, E., 2005. Cis-regulatory inputs of the Wnt8 Gene in the sea urchin endomesoderm network. *Developmental Biology* 288 (2), 545–558.
- Nelson, D.E., Ihekweaba, A.E.C., Elliott, M., 2004. Oscillations in NF-kappaB signaling control the dynamics of gene expression. *Science* 306 (5696), 704–708.
- Nonchev, S., Maconochie, M., Vesque, C., 1996a. The conserved role of Krox-20 in directing Hox gene expression during vertebrate hindbrain segmentation. *Proceedings of the National Academy of Sciences of the United States of America* 93 (18), 9339–9345.
- Nonchev, S., Vesque, C., Maconochie, M., et al., 1996b. Segmental expression of Hoxa-2 in the hindbrain is directly regulated by Krox-20. *Development* 122 (2), 543–554.
- Oren, M., 1999. Regulation of the p53 tumor suppressor protein. *Journal of Biological Chemistry* 274 (51), 36031–36034.
- Pan, G., Li, J., Zhou, Y., Zheng, H., Pei, D., 2006. A negative feedback loop of transcription factors that controls stem cell pluripotency and self-renewal. *FASEB Journal* 20 (10), 1730.

- Perkins, T.J., Jaeger, J., Reinitz, J., Glass, L., 2006. Reverse engineering the gap gene network of *Drosophila melanogaster*. *PLoS Computational Biology* 2 (5), 417–428.
- Peter, I.S., Davidson, E.H., 2009a. Genomic control of patterning. *International Journal of Developmental Biology* 53 (5–6), 707–716.
- Peter, I.S., Davidson, E.H., 2009b. Modularity and design principles in the sea urchin embryo gene regulatory network. *FEBS Letters* 583 (24), 3948–3958.
- Rawlings, J., Rosler, K., Harrison, D., 2004. The JAK/STAT signaling pathway. *Journal of Cell Science* 117 (8), 1281–1283.
- Reeves, G.T., Muratov, C.B., Schupbach, T., Shvartsman, S.Y., 2006. Quantitative models of developmental pattern formation. *Developmental Cell* 11 (3), 289–300.
- Reeves, G.T., Stathopoulos, A., 2009. Graded dorsal and differential gene regulation in the *Drosophila* embryo. *Cold Spring Harbor Perspectives in Biology* 1 (4), a000836.
- Saha, K., Schaffer, D.V., 2006. Signal dynamics in Sonic hedgehog tissue patterning. *Development* 133 (5), 889–900.
- Shah, N.A., Sarkar, C.A., 2011. Robust network topologies for generating switch-like cellular responses. *PLoS Computational Biology* 7 (6), e1002085.
- Shen-Orr, S., Milo, R., Mangan, S., Alon, U., 2002. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genetics* 31 (1), 64–68.
- Siegal-Gaskins, D., Mejia-Guerra, M.K., Smith, G.D., Grotewold, E., 2011. Emergence of switch-like behavior in a large family of simple biochemical networks. *PLoS Computational Biology* 7 (5), e1002039.
- Stathopoulos, A., Levine, M., 2002. Dorsal gradient networks in the *Drosophila* embryo. *Developmental Biology* 246 (1), 57–67.
- Taniguchi, T., Takaoka, A., 2001. A weak signal for strong responses: Interferon- α/β revisited. *Nature Reviews Molecular Cell Biology* 2 (5), 378–386.
- Thieffry, D., Sanchez, L., 2003. Dynamical modelling of pattern formation during embryonic development. *Current Opinion in Genetics & Development* 13 (4), 326–330.
- Vesque, C., Maconochie, M., Nonchev, S., *et al.*, 1996. Hoxb-2 transcriptional activation in rhombomeres 3 and 5 requires an evolutionarily conserved cis-acting element in addition to the Krox-20 binding site. *EMBO Journal* 15 (19), 5383–5396.
- Wasserman, J., Freeman, M., 1998. An autoregulatory cascade of EGF receptor signaling patterns the *Drosophila* egg. *Cell* 95 (3), 355–364.
- Watts, D., Strogatz, S., 1998b. Collective dynamics of 'small-world' networks. *Nature* 393 (6684), 440–442.
- Xiong, W., Ferrell, J., 2003. A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature* 426 (6965), 460–465.
- Yakoby, N., Lemong, J., Schupbach, T., Shvartsman, S.Y., 2008. *Drosophila* eggshell is patterned by sequential action of feedforward and feedback loops. *Development* 135 (2), 343–351.
- Yan, S., Zartman, J.J., Zhang, M., *et al.*, 2009. Bistability coordinates activation of the EGFR and DPP pathways in *Drosophila* vein differentiation. *Molecular Systems Biology* 5, 278.
- Yoo, A., Greenwald, I., 2005. LIN-12/Notch activation leads to micro RNA-mediated down-regulation of vav in *C. elegans*. *Science* 310 (5752), 1330–1333.
- Yuh, C., Dorman, E., Howard, M., Davidson, E., 2004. An otx cis-regulatory module: A key node in the sea urchin endomesoderm gene regulatory network. *Developmental Biology* 269 (2), 536–551.
- Yuh, C.H., Bolouri, H., Davidson, E.H., 2001. cis-Regulatory logic in the endo16 gene: Switching from a specification to a differentiation mode of control. *Development* 128 (5), 617–629.
- Zartman, J.J., Shvartsman, S.Y., 2007. Enhancer organization: Transistor with a twist or something in a different vein? *Current Biology* 17 (24), R1048–R1050.
- Zinzen, R.P., Senger, K., Levine, M., Papatsenko, D., 2006. Computational models for neurogenic gene expression in the *Drosophila* embryo. *Current Biology* 16 (13), 1358–1365.

Further Reading

- Bolouri, H., Davidson, E.H., 2002a. Modeling DNA sequence-based cis-regulatory gene networks. *Developmental Biology* 246 (1), 2–13.
- Cotterrell, J., Sharpe, J., 2010. An atlas of gene regulatory networks reveals multiple three-gene mechanisms for interpreting morphogen gradients. *Molecular Systems Biology* 6, 425.
- Davidson, E.H., Levine, M.S., 2008. Properties of developmental gene regulatory networks. *Proceedings of the National Academy of Sciences of the United States of America* 105 (51), 20063–20066.
- Davidson, E.H., McClay, D.R., Hood, L., 2003. Regulatory gene networks and the properties of the developmental process. *Proceedings of the National Academy of Sciences of the United States of America* 100 (4), 1475–1480.
- Davidson, E.H., Rast, J.P., Oliveri, P., *et al.*, 2002. A genomic regulatory network for development. *Science* 295 (5560), 1669–1678.
- Holme, P., Kim, B., 2002. Growing scale-free networks with tunable clustering. *Physical Review E* 65 (2), 0261071–0261074.
- Lee, T.I., Rinaldi, N.J., Robert, F., *et al.*, 2002. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* 298 (5594), 799–804.
- Levine, M., Davidson, E., 2005. Gene regulatory networks for development. *Proceedings of the National Academy of Sciences of the United States of America* 102 (14), 4936–4942.
- Mangan, S., Alon, U., 2003. Structure and function of the feed-forward loop network motif. *Proceedings of the National Academy of Sciences of the United States of America* 100 (21), 11980–11985.
- Milo, R., Itzkovitz, S., Kashtan, N., *et al.*, 2004. Superfamilies of evolved and designed networks. *Science* 303 (5663), 1538–1542.
- Oltvai, Z.N., Barab'asi, A., 2002. Systems biology. Life's complexity pyramid. *Science* 298 (5594), 763–764.
- Stathopoulos, A., Levine, M., 2002. Dorsal gradient networks in the *Drosophila* embryo. *Developmental Biology* 246 (1), 57–67.
- Telesford, K.K., Joyce, K.E., Hayasaka, S., Burdette, J.H., Laurienti, P.J., 2011. The ubiquity of small-world networks. *Brain Connection* 1 (5), 367–375.
- Watts, D.J., Strogatz, S.H., 1998a. Collective dynamics of 'small-world' networks. *Nature* 393 (6684), 440–442.

Relevant Website

<http://supg.caltech.edu/endomes/>
Caltech.

VERTICAL INTEGRATION: MODELING

Contents

Vertical Integration

Computational Approaches for Multiscale Modeling

Cardiac Modeling

Vertical Integration

HM Byrne, University of Oxford, Oxford, UK

© 2016 Elsevier Inc. All rights reserved.

Introduction

It is becoming increasingly clear that normal functioning of biological cells involves careful coordination of a range of processes acting across diverse spatial and temporal scales (Meier-Schellersheim *et al.*, 2009; Walpole *et al.*, 2013): essential cellular functions such as macromolecule synthesis and degradation, metabolism, proliferation, and death must be tightly regulated. Aberrant regulation of these processes caused, for example, by mutations in a specific gene may alter a cellular phenotype and, ultimately, become manifest at the tissue scale as a disease such as cancer or diabetes (Goncalves *et al.*, 2013). The complex nonlinear feedbacks that regulate such processes cannot typically be understood by deductive or linear thinking alone. Mathematical and computational modeling represents a logical framework within which to understand such complex systems. Modeling involves the creation of an abstract representation of a biological system, often using the language of mathematics or computer science. Model simulations provide a method for predicting how the biological system of interest will behave when subjected to specific conditions. Traditionally, such models have focused on a single time and/or space scale. The growing realization that many biological systems involve processes that occur on different space and timescales is now creating a strong demand for new, multiscale models that integrate phenomena acting on different levels of organization: the spatial scales may range from the molecular to the tissue while the temporal scales may range from those associated with molecular vibrations to the lifespan of a cell. Such multiscale models have the potential to increase our understanding of complex biological systems in a qualitative and quantitative manner and simultaneously to bridge the gap in understanding between *in vitro* and *in vivo* experiments (Walpole *et al.*, 2013).

This article serves as an introduction to the section on vertical integration that follows and is organized as follows. We start by introducing the spatial and timescales of biological resolution that can be modeled and explain the different philosophies that underpin multiscale models. We then

describe briefly the mathematical and computational techniques and related software that can be used to construct and simulate multiscale models. After presenting a couple of case studies, the article concludes with a discussion of the challenges and opportunities associated with vertical integration.

Before continuing, however, we pause to explain what we understand by vertical integration and how it differs from systems biology.

- What is vertical integration?

Vertical integration is the merging together of mathematical and computational models formulated at different spatial and/or temporal scales to produce multiscale models of a biological process.

- What is the difference between vertical integration and systems biology?

Both vertical integration and systems biology aim to increase understanding of complex biological systems through the development and investigation of mathematical and computational models. The key difference is that vertical integration must involve the coupling of models that act at different spatial and/or timescales, whereas systems biology may focus on processes that act on one or more spatial and/or temporal timescales. In this respect, vertical integration can be viewed as a special case of systems biology.

The Multiple Time and Spatial Scales Associated with Cellular Biology

The range of space and timescales that can be integrated in cellular systems is vast. Consider, for example, the timescales associated with molecular vibrations (10^{-15} s), signal transduction (10^{-9} s), ligand binding (10^{-3} s), protein diffusion across a cell (1 s), and the doubling time of a tumor cell (10^5 s). The lengthscales of interest also span several orders of

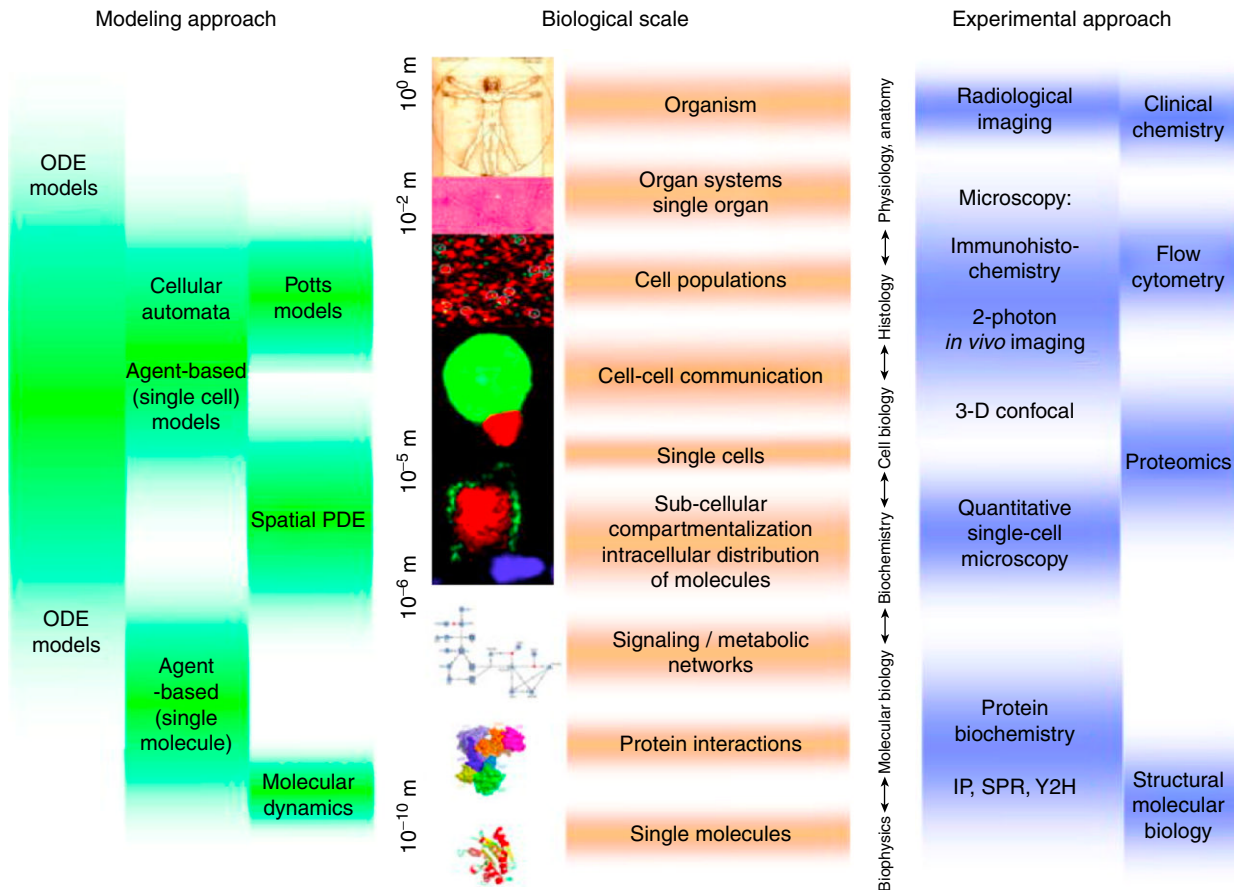


Figure 1 A schematic diagram illustrating the different biological scales associated with multiscale modeling, together with the associated modeling approaches and experimental techniques. IP, immuno-precipitation; ODE, ordinary differential equation; PDE, partial differential equation; SPR, surface plasmon resonance; Y2H, yeast two-hybrid. Reproduced with permission from Meier-Schellersheim, M., Fraser, I.D., Klausch, F., 2009. Multiscale modelling for biologists. Wiley Interdisciplinary Reviews: Systems Biology and Medicine 1 (1), 4–14.

magnitude, ranging from 10^{-9} m for proteins and 10^{-6} m for mitochondria, to 10^{-5} m for mammalian cell diameters and 10^{-2} m (or larger) for organs (see Figure 1 for details). A significant challenge associated with vertical integration is developing theoretical and numerical methods that can accurately and efficiently simulate processes which act on, and interact across, such diverse scales.

Approaches for Linking Scales in Multiscale Models: 'Top-Down', 'Bottom-Up,' and 'Middle-Out'

When constructing a multiscale model, typically only a subset of the space and timescales mentioned above are included. For example, a multiscale model of carcinogenesis (the initiation of a cancer due to a genetic mutation) might integrate genetic, proteomic, subcellular, and cellular processes, whereas a multiscale model of tumor invasion might focus on subcellular, cellular, and tissue scale phenomena. In other applications it may be natural to merge two scales or to decompose one scale into two or more components. As a rule, when formulating a multiscale model one

should be guided by the details of the system of interest rather than to follow a rigid and predefined framework (Southern *et al.*, 2008).

A related question that must be addressed when formulating a multiscale model is how to attribute causality. There are three main approaches: 'top-down,' 'bottom-up,' and 'middle-out' (Kohl *et al.*, 2010). The 'top-down' or classical physiological approach aims to explain observations at a particular scale in terms of mechanisms that occur on smaller, more fundamental scales, by including more and more detail as it becomes available. By contrast, the 'bottom-up' or molecular biology approach starts at a fundamental level (e.g., with a description of genes and genetic relationships) and aims to coarse-grain or average processes acting at a given scale to predict emergent or collective behavior at a higher, less detailed scale. In practice, feedbacks may operate across the scales, in which case a 'middle-out' approach that combines both top-down and bottom-up approaches may be more appropriate (Kohl *et al.*, 2010). Middle-out models typically start at a scale for which there is sufficient information to develop a mechanistic model. The model can then be extended by moving up or down the scales.

Mathematical and Computational Representations of the Components of a Multiscale Model

Once the scales of interest and an appropriate approach (e.g., top-down, bottom-up, or middle-out) have been identified, the next issue is to decide what mathematical and/or computational formalism is appropriate at each scale and for each process being modeled. While it is not possible exhaustively to list the options available, we illustrate below the sorts of choices that have to be made and refer the interested reader to [Southern et al. \(2008\)](#) and references therein for further details:

- Deterministic or stochastic (Is there a degree of randomness in the system that allows multiple solutions to be generated from the same initial conditions?)
- Discrete or continuous (Is the solution space decomposed into a finite number of discrete space or time intervals?)
- Dynamic or static (Does a variable evolve over time?)
- Spatially averaged or spatially resolved (Do the physical variables vary with position?)

The type of data available for model validation may guide the selection of modeling approaches. Thus, if quantitative and dynamic data is available for average protein levels within a population of cells, then a deterministic framework that is continuous in time and neglects spatial effects may be appropriate and a description of the associated signaling pathway formulated using systems of time-dependent ordinary differential equations (ODEs) and by appealing to the law of mass action to represent the chemical reactions. If only binary information is available about whether a protein is present or absent at a particular time, then a simpler, binary Boolean model, formulated using deterministic or stochastic logic rules, may be appropriate. In addition to ODEs and Boolean models, other modeling approaches that are used to develop components of multiscale models include partial differential equations (PDEs), stochastic differential equations, and agent based models.

PDEs, particularly reaction-diffusion equations, are widely used to model the continuous in time and space distribution of extracellular nutrients such as oxygen and glucose, waste products, drugs, hormones, and growth factors. Such models typically account for tissue scale processes such as diffusive and advective transport and local rates of production and consumption of the species of interest. Other processes that may be relevant at the tissue scale include biomechanical forces, electrical, and magnetic fields. Continuum models describing these fields are well-established. For example, in heart modeling, the bidomain and monodomain models are typically used to study defibrillation ([Keener and Sneyd, 1998](#); [Plonsey, 1988](#)). The bidomain equations comprise a pair of nonlinear PDEs that account for anisotropy in the electrical conductivity of the intracellular and extracellular spaces in cardiac tissue: they reduce to the monodomain equations when the electrical conductivity is assumed to be isotropic (i.e., spatially uniform) in both the extracellular and intracellular spaces. In separate work, [Owen et al. \(2011\)](#) incorporated a model for an externally applied magnetic field into an existing multiscale model of vascular tumor growth in order to investigate the potential for loading genetically engineered macrophages with magnetic

nanoparticles in order to enhance the delivery of cytotoxic drugs to hypoxic tumor regions. The above examples, in which different physical processes (e.g., forces, electrical magnetic fields) are considered, may be termed multiphysics models: they are a common feature of multiscale models.

PDEs, or stochastic differential equations, may also be used (in place of ODEs) to investigate processes that act on smaller spatial scales. For example, [Basan et al. \(2010\)](#) use PDEs to study the spatial variation in cytoplasmic levels of proteins associated with the Wnt signaling pathway while [Mazemondet et al. \(2012\)](#) formulated an alternative, stochastic model of the same pathway.

At the intermediate, cellular scale, a variety of discrete modeling approaches can be used to describe individual cells and their interactions ([Anderson et al., 2007](#)). These range from simple cellular automata, where each cell is represented by a single site on a regular lattice and predefined rules are used to update the state or phenotype of each cell ([Anderson et al., 2007](#); [Perfahl et al., 2011](#); [Youssef et al., 2007](#)), to more detailed and physically based models which account for cell mechanics ([Buske et al., 2011](#); [Galle et al., 2005](#)). While lattice-based models such as cellular automata and the Cellular Potts Model (CPM) ([Chen et al., 2007](#)) provide simple descriptions of the dynamics of cell populations, many of the associated model parameters are difficult to relate to experimentally measurable quantities. If mechanical effects and/or details of cell size and morphology are important, a less restrictive, off-lattice approach may be preferred. Off-lattice approaches include cell-center models in which each cell is represented by a single point and an associated surface area (or volume) which may be defined through a Voronoi tessellation ([Osborne et al., 2010](#)). Vertex models are an alternative off-lattice approach in which each cell is represented by a polygon (or polyhedron) whose vertices may be shared with neighboring cells ([Harvey et al., 2015](#); [Osborne et al., 2010](#); [Pitt-Francis et al., 2009](#)).

Assembling Multiscale Modeling

As stated above, multiscale models link models of biophysical processes that act at two or more different space and time-scales. They are typically constructed on a case-by-case basis, with different groups taking different approaches, even when studying the same system. Consider, for example, the wealth of multiscale models for diffusion-limited growth of avascular clusters of tumor cells suspended in culture medium. The models typically span three spatial scales, coupling subcellular, cellular, and tissue scale phenomena ([Anderson et al., 2007](#)) and are constructed in a similar, middle-out manner. Thus, at the tissue scale, reaction-diffusion equations describe the diffusion, natural decay, and consumption by the tumor cells of vital nutrients, such as oxygen and glucose, present in the culture medium. The local nutrient environment influences the subcellular dynamics: where nutrient levels are high, cells progress through the cell cycle; where nutrient levels are low they first exit the cell cycle and become quiescent and, then, if nutrient levels remain excessively low for a sustained period of time, the cells undergo necrosis and die. Behavior at the cell scale is coupled to the subcellular scale. On each timestep, each cell checks whether it has completed its cell cycle: if so, then it attempts to divide into two cells (the rules for cell

division are specific to the cell-based model being used). If the region surrounding the dividing cell is fully occupied by other cells (and/or the dividing cell is unable to push its neighboring cells away), then the cell may be contact-inhibited and division will not occur. Equally, the subcellular dynamics may dictate that the cell is removed from the system due, for example, to necrosis, the way in which this is achieved again being specific to the cell-based model being used. While sharing the above, generic attributes, the multiscale models of avascular tumor growth differ in two main ways: in the level of detail used to model subcellular phenomena (e.g., the cell cycle, metabolism, response to oxygen) and the framework used to model individual cells (e.g., cellular automata, vertex-based or cell-centered models). Despite these differences, many models of avascular tumor spheroids reproduce the same qualitative dynamics. Initially all cells are nutrient-rich and proliferate rapidly, causing the cluster to grow exponentially until cells toward the spheroid center become starved of nutrients, stop proliferating, and, if nutrient levels become sufficiently low, undergo necrosis. These changes cause the tumor's growth rate to transition from exponential to linear and eventually to plateau when the proliferation rate of nutrient-rich cells at the spheroid periphery balances the death rate of nutrient-starved cells at its center. Without access to the software used to generate the different models, it is difficult to understand what features of a particular multiscale model, and/or the computational algorithm used for its solution, are responsible for the observed dynamics. The open-source software CHASTE (see Section Relevant Websites) was designed, in part, to address questions of this type. CHASTE is a general purpose package that can be used to simulate multiscale and multiphysics models in biology and physiology (Pitt-Francis *et al.*, 2009). Due to its modular structure, different models of the same process can be used interchangeably and, thereby, CHASTE facilitates comparison of different multiscale models. For example, Osborne *et al.* (2010) used CHASTE to perform a systematic comparison of three multiscale models of homeostasis in the intestinal crypt, the models differing only in the approach used to describe individual cells. Extensive numerical simulations revealed that cell-center and vertex-based models of sheets of epithelial cells produce near identical results which differ markedly from those obtained when an alternative, fixed lattice, cellular-automation model was used.

While the existence of multiple multiscale models of the same physical system may seem unnecessary (in addition to multiple models of avascular tumor growth), there are also multiple multiscale models of vascular tumor growth and cardiac electrophysiology. In the long term, such diversity could prove beneficial. For example, in the same way as different models of weather forecasting are used to predict future weather patterns and different models of action potentials in the heart are now being used to predict patient responses to different cardiac drugs, the response of different multiscale models of avascular and/or vascular tumor growth to a particular chemotherapeutic drug (or radiotherapy) could be used to generate probabilistic predictions about the likely response of a particular tumor to a specific course of treatment.

A further advantage of having multiple multiscale models of the same system stems from the different predictions that they may generate. For example, in Alarcon *et al.* (2003) a

multiscale model of vascular tumor growth is proposed in which a network of blood vessels supply oxygen to a tissue containing normal and cancer cells. Numerical results are presented to show the way in which oxygen transport through the vasculature affects the morphology, size, and growth dynamics of the evolving tumor mass: the tumor expands as a radially symmetric mass if the oxygen tension is assumed to be distributed uniformly through the vascular network; its growth is heterogeneous and localized around oxygen-rich blood vessels when more detailed models that account for blood flow and hematocrit are employed. These qualitative differences highlight the importance of model validation in multiscale modeling, in this case they indicate that in order to generate realistic patterns of tumor spread it is important to model blood flow (and vascular adaptation) accurately.

While many aspects of composing and solving a multiscale model are problem-specific, there are several techniques and strategies that are widely applicable. It may be possible to exploit differences in space and timescales in order to make model simplifications. For example, since the kinetics of molecular binding and diffusion typically occur over short timescales, steady-state assumptions for their dynamics may be justified in the overall model architecture if one is concerned with longer timescales of hours, days, or weeks. Such reasoning and techniques have been widely used to model metabolic and signaling networks, the steady-state fluxes being used in a hierarchical manner to determine behavior at higher levels of organization (Quo *et al.*, 2011; Sun *et al.*, 2009).

Software for Multiscale Modeling

The diversity of theoretical approaches that may be used to model each component or submodel within a multiscale model (e.g., discrete or continuous, deterministic or stochastic, spatially averaged or spatially resolved), combined with the vast range of space and timescales that can be linked together, and the current limits on computing power, mean that it is unlikely that any single, multiscale modeling framework will ever be capable of simulating all possible multiscale systems. A survey of existing software for simulating multiscale models (the list of websites included at the end of this article provides links to this software) suggests that it is more likely that, in the future, generic software such as *compuCell3D* (see Section Relevant Websites; Chen *et al.*, 2007; Graner and Glazier, 1992) and CHASTE (see Section Relevant Websites; Pitt-Francis *et al.*, 2009) will be used to develop simple, prototype models for new applications while highly specialized software, which links only a limited number of space and timescales, will be used to generate detailed insight into specific systems, such as the heart and the model plant *Arabidopsis thaliana* (Band *et al.*, 2012; Kuchen *et al.*, 2012; Muraro *et al.*, 2014).

Since a central aspect of developing multiscale models is the integration of different simulation techniques and since software for many of these techniques already exist, it is natural to ask how existing software could (and should) be exploited for multiscale modeling. The *biomodels* database and *CellML* (see Section Relevant Websites) are two community-led initiatives designed to promote the reuse and integration of published models of biological systems. For example, the *biomodels*

database is a manually curated repository, containing hundreds of published models, represented using the systems biology markup language (SBML; see Section Relevant Websites). CellML is similar, but with a greater emphasis on model integration. Such endeavors highlight the need for agreed standards and ontologies for mathematical and computational models if the vision of multiscale modeling in biology and medicine, that is the seamless integration of models, simulation techniques, and experimental data, is to be realized (Meier-Schellersheim *et al.*, 2009; Wolkenhauer *et al.*, 2014).

Applications of Multiscale Modeling and Two Case Studies

While vertical integration is still in its infancy, there are a number of well-established multiscale models in the literature: of these the cardiac models of Noble, Hunter, and coworkers are arguably the most well-studied (Kohl and Noble, 2009). The articles that follow provide an excellent starting point for learning more about the state-of-the-art in multiscale modeling of angiogenesis, morphogenesis, neurogenesis, wound healing, inflammation, and cardiac modeling and also an overview of the computational approaches that are used to construct and simulate multiscale models. A common feature of these multiscale models is that they are multicellular and are designed to understand how subcellular processes (e.g., signaling networks), cellular processes (e.g., cell division, migration, and death) interact with tissue scale processes (e.g., the transport and consumption of growth factors,

nutrients and waste products). There is a complementary and distinct body of multiscale modeling focused exclusively at the subcellular scale. For example, hybrid multiscale models that combine quantum and classical mechanics have been proposed to study enzymatic reactions (Leach, 2001). Alternative hybrid models that couple stochastic and deterministic models for chemical reaction kinetics that operate over short, intermediate, and long timescales have also been developed and used to study the subcellular dynamics of the LacZ and LacY proteins in *Escherichia coli* (Burrage *et al.*, 2004). In separate work, Mashl *et al.* (2001) used hierarchical multiscale modeling to simulate the flux of potassium through a particular ion channel. Potential profiles for the passage of individual potassium ions through the ion channel were generated from molecular models and then used in molecular dynamics simulations to estimate the bulk diffusion coefficient of potassium ions in water. The diffusion coefficients were then included in Brownian dynamics simulations to establish how the flux of potassium ions varies with the transmembrane.

We present below two case studies which illustrate how multiple processes acting on multiple spatial, temporal, and functional scales may interact to generate homeostasis (in the intestinal crypt) or, alternatively, how they may contribute to disease progression (in diabetic retinopathy).

Case Study I: Homeostasis in the Intestinal Crypt

The intestinal epithelium is a rapidly renewing tissue, being completely replaced every 2–3 days in mice and 3–5 days in humans. In the large intestine, it comprises approximately

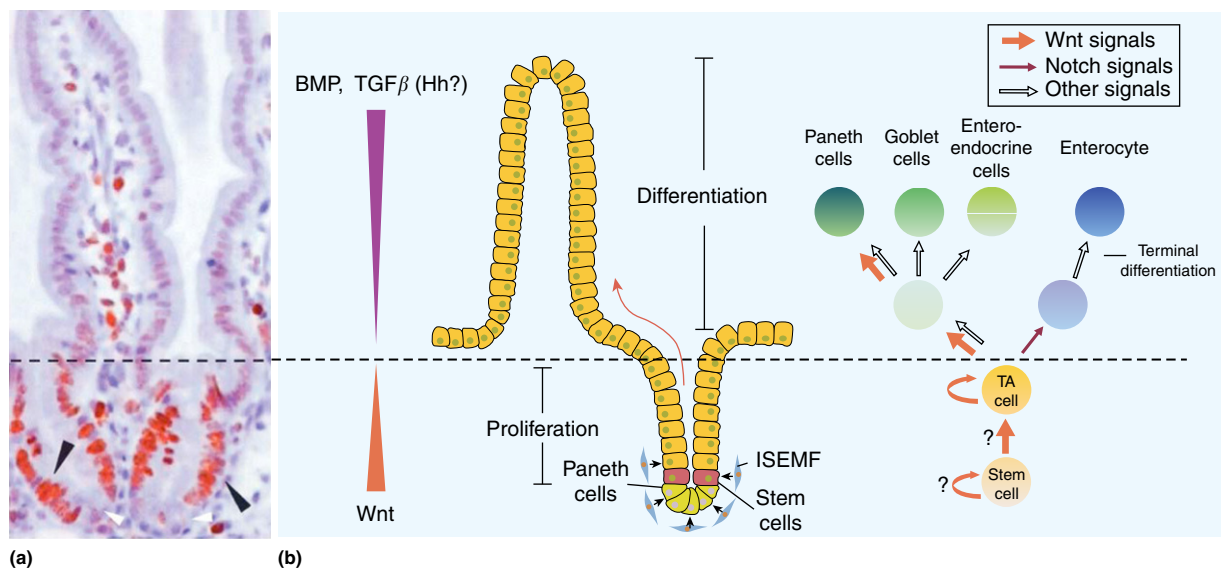


Figure 2 Homeostasis of the adult intestinal crypt. (a,b) Schematic representation and section of the crypt-villus unit in the mature small intestine. Proliferative cells reside in the crypts, while differentiated cells occupy the villus. Crypt progenitors migrate up (red arrow) the crypt-villus axis before shedding into the lumen. The process of epithelial renewal takes 3–6 days and is ensured by a small number of asymmetrically dividing stem cells at the bottom of the crypts. Wnt signaling in the adult intestine promotes proliferation of progenitor or transit-amplifying (TA) cells, as well as commitment toward secretory lineages. Wnt signaling may also drive terminal differentiation of certain secretory lineages. Although it is commonly believed that Wnt signaling may promote proliferation and/or differentiation of intestinal stem cells, there is no evidence that formally proves this (see arrows with question marks). In panel (a), black arrowheads indicate Ki67 positive TA cells, while white arrowheads indicate the Paneth cell compartment. Reproduced with permission from Gregorieff, A., Clevers, H., 2005. Wnt signalling in the intestinal epithelium: From endoderm to cancer. *Genes & Development* 19, 877–890.

2×10^7 crypts or test-tube-shaped invaginations. Stem cells located toward the base of each crypt proliferate occasionally, producing on average one stem cell and one semi-differentiated transit cell. Transit cells divide several times while migrating upwards before committing to either an absorptive or secretory fate. At the crypt orifice, mature cells undergo apoptosis and are shed into the intestinal lumen (Grossmann *et al.*, 2002; Yen and Wright, 2006). While the mechanisms by which cell proliferation, migration, differentiation, and apoptosis are regulated within the crypt remain fully to be elucidated, there is evidence that a spatial gradient in extracellular Wnt factors along the crypt axis determines the position-dependent rates of cell proliferation, differentiation, and death (and that cross talk with the delta-notch pathway drives cell-fate specification) (see Figure 2). A number of different multiscale models of the intestinal crypt have been developed and used to study processes such as monoclonal conversion

(Mirams *et al.*, 2012), cell-fate specification (Buske *et al.*, 2011), and early colorectal cancer (van Leeuwen *et al.*, 2009); for a review, see Kershaw *et al.* (2013), De Matteis *et al.* (2013).

Case Study II: Diabetic Retinopathy

At the subcellular scale, through its activation of protein kinase C (PKC)-delta and downstream phosphatases, chronic hyperglycemia causes dephosphorylation of platelet-derived growth factor (PDGF) receptors that normally promote the survival of pericytes, cells that envelop endothelial cells and, thereby, prevent fluid leakage from blood vessels and inhibit excessive endothelial cell proliferation. Dephosphorylation of PDGF receptors simulates pericyte apoptosis, leading to increased blood vessel permeability and uncoordinated endothelial cell proliferation throughout the retina. The cumulative effect of sustained periods of such microvascular hemorrhage and

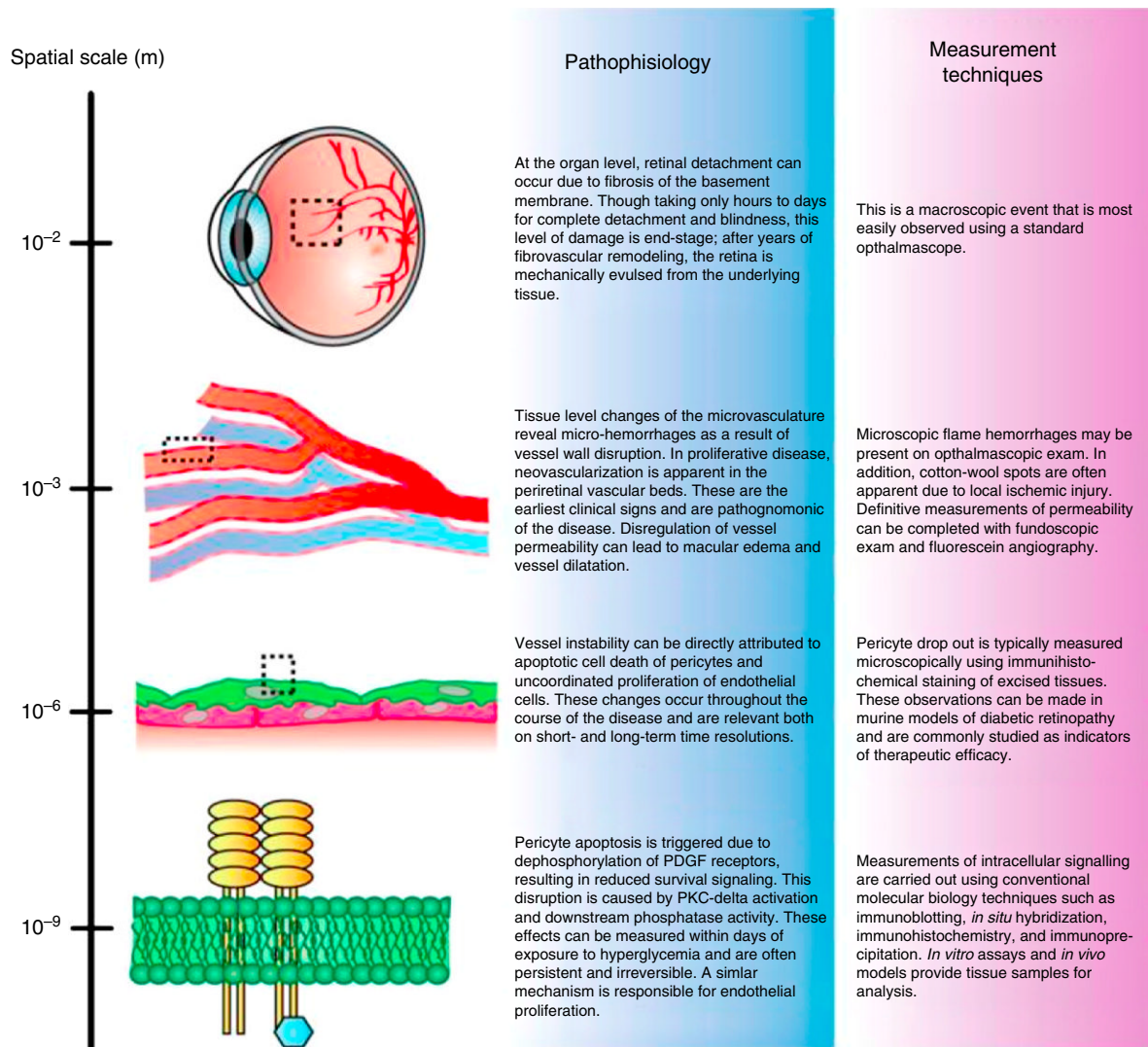


Figure 3 Diabetic retinopathy as a case study in multiscale biology. A detailed look at how the pathogenesis of diabetic retinopathy is a function of multiple spatial scales across biology. Reproduced with permission from Walpole, J., Papin, J.A., Peirce, S.M., 2013. Multiscale computational models of complex biological systems. Annual Review of Biomedical Engineering 15, 137–154.

vascular remodeling is extensive tissue damage which, in its most advanced stage, can cause blindness due to retinal detachment at the macroscopic or tissue scale (see [Figure 3; Walpole et al., 2013](#)).

Challenges Associated with Multiscale Modeling

There are many theoretical, biological, and experimental challenges associated with simulating multiscale models. For example, how should deterministic and stochastic models be coupled? Should different coupling methods be used to couple across different space and timescales? Further, within a given multiscale model, how do numerical errors in the simulated values of deterministic variables (and, similarly, random variations in stochastic variables) propagate from one scale to another? Can different types of data that have been generated using different experimental techniques be integrated with multiscale models to generate biologically meaningful predictions? Are the parameter estimates associated with a particular submodel robust to changes in framework used to describe another component of the multiscale model? Interest in answering these open questions is currently driving much of the research in this field.

The equations governing the evolution of multiscale processes are often stiff since the range of length and space scales involved are large. While it is usually more efficient to use implicit numerical methods to solve stiff problems, for the large nonlinear systems generated by multiscale models, it can be unclear how to proceed: implicit methods may be prohibitively expensive to compute and explicit methods prone to numerical instability. In such cases, it may be appropriate to use semi-implicit methods, with some variables solved explicitly and others implicitly ([Southern et al., 2008](#)).

As already stated, multiscale models typically contain large numbers of variables and parameters. The development of fast, accurate, and reliable numerical methods for their simulation remains an active research area, with many groups reducing simulation times by exploiting parallel computing ([Chen et al., 2007](#); [Coveney and Fowler, 2005](#); [Harvey et al., 2015](#); [Youssef et al., 2007](#)). Such techniques are urgently needed if the full potential of multiscale modeling is to be realized: large numbers of simulations must be performed for model validation, parameter estimation, and also to generate experimentally testable predictions from the models. Since many of the parameters included in multiscale models are difficult or impossible to measure, parameter estimation remains an ongoing concern. Further, parameter values estimated from *in vitro* data may not be applicable for multiscale models describing *in vivo* systems. In such cases, techniques that permit rapid, systematic, and exhaustive searches of multidimensional parameter spaces are needed to identify physically realistic parameter values for multiscale models. Alternatively, improvements in *in vivo* imaging techniques may make it possible to estimate parameters that cannot be estimated at present ([Walpole et al., 2013](#)).

Due to their complexity, it is often impractical for one group to reproduce the numerical results generated from another group's multiscale model, unless they have access to their software. As a result, it can be difficult to establish, with confidence, whether particular features of a multiscale model

are responsible for observed behaviors. The introduction of standards for multiscale modeling, similar to those being developed by the systems biology community, could mitigate against such uncertainty. Equally, the creation of multiscale, benchmarking datasets would greatly facilitate model validation and comparison.

Summary

- Cellular systems are inherently multiscale and their multiscale nature needs to be considered in order fully to understand their complexity.
- Multiscale models provide a theoretical framework for studying complex biological systems and maximizing the information content of the diverse types of qualitative and quantitative data that can be collected.
- Multiscale models should be fully tested and validated against experimental data in order to establish confidence in the predictions that they generate.
- Comprehensive datasets are needed to benchmark and validate multiscale models.

Conclusions

As experimental techniques improve, and the quality, quantity, and diversity of biological data that can be collected increase, the demand for multiscale mathematical models that can systematically integrate these data and provide mechanistic insight into complex biological processes is set to increase. While still a relatively young discipline, the advantages and opportunities for vertical integration are becoming clear. There are still many areas of biological research where multiscale models are either few or nonexistent. This deficiency should be viewed as an opportunity to demonstrate how the boundaries of biological knowledge can be advanced by using multiscale modeling as a platform for high-throughput hypothesis generation and testing ([Walpole et al., 2013](#)).

See also: Vertical Integration: Applications: A Review on Various Mathematical Modeling Approaches for Wound Healing; Angiogenesis; Mathematical Approaches to Studying Inflammation; Neurogenesis in the Adult Brain. Vertical Integration: Modeling: Computational Approaches for Multiscale Modeling

References

- Alarcon, T., Byrne, H.M., Maini, P.K., 2003. A cellular automaton model for tumour growth in in homogeneous environment. *Journal of the Theoretical Biology* 225 (2), 257–274.
- Anderson, A.R.A., Chaplain, M.A.J., Rejniak, K.A., 2007. *Single-Cell Based Models in Biology and Medicine*. Basel: Birkhauser Verlag AG.
- Band, L.R., Wells, D.M., Larrieu, A., et al., 2012. Root gravitropism is regulated by a transient lateral auxin gradient controlled by a tipping-point mechanism. *Proceedings of the National Academy of Sciences of the United States of America* 109 (12), 4668–4673.

- Basan, M., Idema, T., Lenz, M., Joanny, J.F., Risler, T., 2010. A reaction-diffusion model of the cadherin–catenin system: A possible mechanism for contact inhibition and implications for tumorigenesis. *Biophysical Journal* 98 (12), 2770–2779.
- Burrage, K., Tian, T., Burrage, P., 2004. A multiscale approach for simulating chemical reaction systems. *Progress in Biophysics and Molecular Biology* 85, 217–234.
- Buske, P., Galle, J., Barker, N., *et al.*, 2011. A comprehensive model of the spatio-temporal stem cell and tissue organization in the intestinal crypt. *PLoS Computational Biology* 7, 1332–1336.
- Chen, N., Glazier, J., Izaguirre, J., Alber, M., 2007. A parallel implementation of the Cellular Potts Models for simulation of cell-based morphogenesis. *Computer Physics Communications* 176, 670–681.
- Coveney, P.V., Fowler, P.W., 2005. Modelling biological complexity: A physical scientist's perspective. *Journal of the Royal Society Interface* 2 (4), 267–280.
- De Matteis, G., Graudenzi, A., Antoniotti, M., 2013. A review of spatial computational models for multi-cellular systems, with regard to intestinal crypts and colorectal cancer development. *Journal of Mathematical Biology* 66 (7), 1409–1462.
- Galle, J., Loeffler, M., Drasdo, D., 2005. Modelling the effect of deregulated proliferation and apoptosis on the growth dynamics of epithelial cell populations *in vitro*. *Biophysical Journal* 88 (1), 62–75.
- Goncalves, E., Bucher, J., Ryll, A., *et al.*, 2013. Bridging the layers: Towards integration of signal transduction, regulation and metabolism into mathematical models. *Molecular Systems Biology* 9, 1576–1583.
- Graner, F., Glazier, J., 1992. Simulation of biological cell sorting using a two-dimensional extended Potts model. *Physical Review Letters* 69, 2013–2016.
- Grossmann, J., Walther, K., Artinger, M., *et al.*, 2002. Induction of apoptosis before shedding of human intestinal epithelial cells. *American Journal of Gastroenterology* 97 (6), 1421–1428.
- Harvey, D.G., Fletcher, A.G., Osborne, J.M., Pitt-Francis, J.P., 2015. A parallel implementation of an off-lattice individual-based model of multicellular populations. *Computer Physics Communications* 192, 130–137.
- Keener, J.P., Sneyd, J., 1998. *Mathematical Physiology*. New York, NY: Springer.
- Kershaw, S.K., Byrne, H.M., Gavaghan, D.J., Osborne, J.M., 2013. Colorectal cancer through simulation and experiment. *IET Systems Biology* 7 (3), 57–73.
- Kohl, P., Crampin, E.J., Quinn, T.A., Noble, D., 2010. Systems biology: An approach. *Clinical Pharmacology and Therapeutics* 88 (1), 25–33.
- Kohl, P., Noble, D., 2009. Systems biology and the virtual physiological human. *Molecular Systems Biology* 5 (1), 292. doi:10.1038/msb.2009.51.
- Kuchen, E.E., Fox, S., Barbier de Reuille, P., *et al.*, 2012. Generation of leaf shape through early patterns of growth and tissue polarity. *Science* 335, 1092–1096.
- Leach, A.R., 2001. *Molecular Modelling: Principles and Applications*, second ed. Edinburgh: Pearson Education Ltd.
- van Leeuwen, I.M.M., Mirams, G.R., Walter, A., *et al.*, 2009. An integrative computational model for intestinal tissue renewal. *Cell Proliferation* 42 (5), 617–636.
- Mazemondet, O., John, M., Leye, S., Rolfs, A., Uhrmacher, A.M., 2012. Elucidating the sources of beta-catenin dynamics in human neural progenitor cells. *PLoS ONE* 7 (8), 1–12.
- Mashl, R.J., Tang, Y., Schnitzer, J., Jakobssen, E., 2001. Hierarchical approach to predicting permeation in ion channels. *Biophysical Journal* 81, 2473–2483.
- Meier-Schellersheim, M., Fraser, I.D., Klausch, F., 2009. Multiscale modelling for biologists. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 1 (1), 4–14.
- Mirams, G.R., Fletcher, A.G., Maini, P.K., Byrne, H.M., 2012. A theoretical investigation of the effect of proliferation and adhesion on monoclonal conversion in the colonic crypt. *Journal of Theoretical Biology* 312, 143–156.
- Muraro, D., Mellor, N., Pound, M.P., *et al.*, 2014. Integration of hormonal signaling networks and mobile microRNAs is required for vascular patterning in *Arabidopsis* roots. *Proceedings of the National Academy of Sciences of the United States of America* 111 (2), 857–862.
- Osborne, J.M., Walter, A., Kershaw, S.K., *et al.*, 2010. A hybrid approach to multiscale modelling of cancer. *Philosophical Transactions of the Royal Society B: Biological Sciences* 368, 5013–5028.
- Owen, M.R., Stamper, I.J., Muthana, M., *et al.*, 2011. Mathematical modelling predicts synergistic antitumour effects of combining a macrophage-based, hypoxia-targeted gene therapy with chemotherapy. *Cancer Research* 71 (8), 2826–2837.
- Perfahl, H., Owen, M.R., Alarcon, T., *et al.*, 2011. 3D hybrid multiscale modelling of vascular tumour growth. *PLoS ONE* 6 (4), e14790.
- Pitt-Francis, J., Pathmanathan, P., Bernabeu, M.O., *et al.*, 2009. Chaste: A test-driven approach to software development for physiological modelling. *Computer Physics Communications* 180 (12), 2452–2471.
- Plonsey, R., 1988. Bioelectric sources arising in excitable fibres (ALZA lecture). *Annals of Biomedical Engineering* 16, 519–546.
- Quo, C.F., Moffitt, R.A., Merrill, A.H., Wang, M.D., 2011. Adaptive control model reveals systematic feedback and key molecules in metabolic pathway regulation. *Journal of Computational Biology* 18, 169–182.
- Southern, J., Pitt-Francis, J., Whiteley, J., *et al.*, 2008. Multiscale computational modelling in biology and physiology. *Progress in Biophysics & Molecular Biology* 96, 60–89.
- Sun, T., Adra, S., Smallwood, R., Holcombe, M., MacNeil, S., 2009. Exploring hypotheses of the actions of TGF-beta1 in epidermal wound healing using a 3D computational multiscale model of the human epidermis. *PLoS ONE* 4, e8515.
- Walpole, J., Papin, J.A., Peirce, S.M., 2013. Multiscale computational models of complex biological systems. *Annual Review of Biomedical Engineering* 15, 137–154.
- Wolkenhauer, O., Auffray, C., Brass, O., *et al.*, 2014. Enabling multiscale modeling in systems medicine. *Genome Medicine* 6 (3), 21.
- Yen, T.H., Wright, N.A., 2006. The gastrointestinal tract stem cell niche. *Stem Cell Reviews* 2 (3), 203–212.
- Youssef, B., Cheng, G., Zygourakis, K., Markenscoff, P., 2007. Parallel implementation of a cellular automaton modelling the growth of three-dimensional tissues. *International Journal of High Performance Computing Applications* 21, 196–209.

Further Reading

Gregorieff, A., Clevers, H., 2005. Wnt signalling in the intestinal epithelium: From endoderm to cancer. *Genes & Development* 19, 877–890.

Relevant Websites

www.celldesigner.org/
Cell Designer Inquiry, The Systems Biology Institute.

www.compuCell3d.org/
CompuCell3D.

www.copasi.org/
COPASI.

<https://www.ebi.ac.uk/biomodels/>
European Bioinformatics Institute.

physiomeproject.org/
IUPS.

<https://ccl.northwestern.edu/netlogo/>
NetLogo.

sbml.org/
Systems Biology Markup Language (SBML).

www.cs.ox.ac.uk/chaste/
University of Oxford.

www.vph-institute.org/
Virtual Physiological Human (VPH) Institute.

Computational Approaches for Multiscale Modeling

Z Wang, University of New Mexico, Albuquerque, NM, USA

V Cristini, University of New Mexico, Albuquerque, NM, USA and King Abdulaziz University, Jeddah, Saudi Arabia

TS Deisboeck, ThinkMotu LLC, Wellesley, MA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Agent-based modeling A specific discrete-based hybrid modeling approach that allows simulating single to multi-element groups and their interactions, thereby enabling the investigation of the role of diversity in populations as well as within each individual element.

Continuum modeling It describes biological tissue as a continuum medium, and thus is an appropriate choice for modeling larger-scale systems such as tissues or organs. This approach draws on principles from continuum mechanics to describe model variables and parameters as continuous fields mostly by means of differential equations.

Discrete modeling It explicitly represents individual model elements, such as cells or proteins, in space and time, and tracks and updates their internal states according to a

predefined set of biological rules. The computational demand increases rapidly with the number of elements and their interactions modeled, limiting these models in the spatial and temporal scales they can represent.

Hybrid modeling It integrates the strengths of both continuum and discrete descriptions. The complexity of any biological systems and the interactions among their constitutive elements calls for a combined continuum and discrete approach.

Multiscale modeling It describes and examines a biological system across two or more spatial and/or two or more temporal scales, and quantifies parameters on and relationships between biological processes that occur at those scales.

Introduction

A biological system is regulated at scales of many orders of magnitude in space and time, from genes and proteins, to cells, tissues, organs, and organisms (Hunter *et al.*, 2008). These different scales coupled with intra- and inter-scale interactions make any such system extremely ‘complex,’ in that the behavior of the emergent, larger-scale system cannot easily be understood by investigating its constituents only. For example, the dynamic interactions at the molecular scale are dominated by random and short timescale fluctuations, but are more deterministic with longer timescale variations at the organ or organism scale. The past two decades have seen significant advances in science and technology along with a rapid expansion of biological, biochemical, and biophysical data on any biological scale (Auffray *et al.*, 2009). However, researchers realize the difficulty of solely relying on experimental observations to understand, quantify, and predict the behavior of complex biological systems. As such, there has been lately substantially more development on the mathematical modeling and computer simulation side, together with traditional experimental and clinical investigations. Indeed, mathematical modeling has been used to provide quantitative insight into biological systems, guide the design of new experiments, introduces potential target candidates for further clinical investigation, and present novel integrated approaches to treat systemic disease.

In particular, systems biology, a new field of integrative research, takes advantage of both experiments and computational/mathematical modeling in an effort to reveal the complexity of biological systems and their behavior at multiple scales in normal and perturbed conditions (Kohl *et al.*, 2010). The key is to perform closely linked model-experiment

iterations for parameterizing and validating these quantitative models. By adopting an integrative approach (as opposed to the more traditional reductionist approach), systems biology aims to predict emergent behavior that will arise from the interactions over time, for example, between genes, proteins, cells, and tissue scales. More specifically, systems biology now seeks not only to understand events at the separate biological scales in a quantitative manner, but also to develop mathematical models which are coupled together to produce integrated approaches across scales, i.e., multiscale models. This integration across multiple scales introduces a powerful tool for capturing and analyzing biological data that is inaccessible through other modeling and experimental techniques. However, understanding, quantifying, and handling (1) the coupling mechanism across scales and (2) the resulting emergent structure and function is a very challenging task (Barabasi, 2005). Even though, multiscale modeling has made significant progress in modeling a variety of biological systems (De Graaf *et al.*, 2009; Engler *et al.*, 2009; Grima, 2008; Lowengrub *et al.*, 2010; Meier-Schellersheim *et al.*, 2009; Sanga *et al.*, 2007; Schnell *et al.*, 2007; Southern *et al.*, 2008; Tracqui, 2009; Qu *et al.*, 2011; Walpole *et al.*, 2013; Slood and Hoekstra, 2010; Hatzikirou *et al.*, 2012). The resulting integrated models, however, are often computationally expensive with exceedingly long runtimes, impossible to solve analytically and difficult to solve numerically even with the tremendous increases in computational power available today.

Despite the large body of literature on multiscale models, there seems to be a lack of discussion on specific methods and clearly specified development strategies for this approach. Since multiscale modeling needs to address the question of how to bridge the gaps between different modeling methods and between models at different scales, there is no

straightforward way to go from one scale to another, but there are useful methods that have been developed for this purpose. In this article, we first present a brief summary of the current multiscale modeling methods in the field of biomedical research. We then further discuss in more detail a number of computational methods and theories, including agent-based modeling (ABM), equation-free approach (EFA), and a dynamic hybrid modeling approach, which can be used to bridge these gaps.

Multiscale Modeling

There are a number of important facts we should state before introducing specific multiscale modeling methods. First, the remarkable increase in computational power has enabled modeling to advance from looking at simple phenomena in isolation to the analysis of complex coupled processes (Bradley *et al.*, 2011). Second, a multiscale model often originates through linking individual sub-models from two or more distinct scales to generate a composite system. The sub-models should be validated at each corresponding scale before validating the integrated model across the scales. Third, a good mathematical model (whether multiscale or scale-specific) should not just fit to a set of experimental or clinical data; rather it should help to explain the underlying mechanisms, and make testable predictions or generate experimentally testable hypotheses (Edgerton *et al.*, 2011; Pascal *et al.*, 2013a, b; Das *et al.*, 2013).

Overview of Current Modeling Methods

A multiscale model may require different mathematical descriptions at these various scales and it may necessitate different computational technologies. In general, smaller-scale processes are faster (i.e., small spatial scales and fast dynamics), so it is reasonable to assume that the smaller-scale processes are in quasi-equilibrium with the slower larger-scale (i.e., large spatial scales and slow dynamics) processes. Hence, smaller-scale processes are often included at the larger-scale via, for example, constitutive equations or force fields (Sloot and Hoekstra, 2010). Various methods have been developed for solving multiscale problems in many scientific areas (Schnell *et al.*, 2007; Southern *et al.*, 2008; Deisboeck *et al.*, 2011), including ‘continuum’ techniques where the relationship between system properties is expressed with continuous mathematical equations, ‘discrete’ techniques which are based on individual units to model the smaller-scale elements such as individual cells, and ‘hybrid’ techniques which combine aspects of both discrete and continuum modeling to provide a more complete description of a system. Perturbations of fine-grained parameters (e.g., cell division) can be used as inputs to generate observable and measurable changes to coarse-grained outputs (e.g., tissue patterning). Conversely, these coarse-grained parameters can also regulate the dynamics of certain fine-grained parameters.

Modeling approaches can also be broadly divided into two categories: ‘bottom-up’ and ‘top-down’; this classification method is more accepted by experimental and clinical

scientists in the field. The bottom-up (i.e., data-driven) approach studies individual elements and their interactions; however, it is computationally intensive which leads to limitations of the number of proteins, mechanisms, or cells that can be modeled at the same time. Another caveat is conceptually, as already mentioned earlier, in that the main characteristic of complex (biological) systems is that piecing its larger-scale functions together from simply monitoring (and modeling) its constituent elements will not succeed. Instead of looking at the details of each individual element, the top-down (i.e., theory- or expertise-driven) approach uses the macroscopic behaviors of the system as variables to model the system dynamics at a larger-scale, largely based on experimental observations. The advantage of the top-down approach is that it is relatively simple and more easily developed. The disadvantage is that the variables and parameters in these models are largely more phenomenological than those in the bottom-up models. Combinatorial concepts would train bottom-up models on macroscopic behaviors, or incorporate fine-grained components into top-down approaches.

Multiscale Modeling Methods

In the following, we introduce some specific multiscale modeling methods that are technically compelling and can be applied in a wide range of biomedical fields. Although for any method, we only describe its key development concepts, we encourage interested readers to refer to the original publications for further technical details. Interesting examples for each method will also be presented where appropriate.

Composite hybrid ABM

ABM is an approach for describing discrete stochastic biological processes as either compartmentalized or spatially defined models in one-, two-, and 3D domains (Byrne and Drasdo, 2009; Auffray *et al.*, 2009; Wang and Deisboeck, 2008). In an agent-based model, each biological cell is often represented as an agent; but this is not necessarily the case – an agent can be scaled as large (groups of organisms) or as small (molecular components) as is needed. The agents must follow some biological, data-supported rules in the course of a simulation, both for their independent behavior and for interactions with other agents and the environment. A serious problem in ABM is its high computational demand because of the large number of agents involved and the spatiotemporal fine-resolution they operate on. Efficient techniques for reduction of computational costs are therefore necessary to be able to use ABM to its maximum potential, en route to more clinically relevant applications. Several types of ABM techniques exist, including lattice-based, lattice-free, Cellular Potts, lattice-gas, and subcellular element modeling methods (Anderson *et al.*, 2007). A number of open-source ABM software packages are available also, facilitating the process of developing an agent-based model (see An *et al.* (2009) for an excellent review).

For instance, Xu *et al.* (2008) developed a 2D multiscale agent-based model using a parallel Cellular Potts model to study thrombus formation in a blood vessel. Many factors

were included in their model: an injury site in a vessel wall, platelets, coagulation factors, fibrin, blood cells, and blood plasma. Platelet movement was driven by plasma flow while platelets could bind to each other, and platelets at the wound could release factors that activate or recruit other platelets. Blood cells and quiescent platelets flow past the wound site, where the platelets can be recruited, activated, and stimulated to produce fibrin and attach to the growing thrombus. The model showed that there were different stages of thrombus formation that happened in sequence and at a fast timescale. They also observed that the thrombus size was influenced by the flow rate, i.e., it being too large or too small resulted in a diminished thrombus formation rate. Using a mouse model, these observations were confirmed experimentally by externalizing a vein and then adding fluorescent markers to the blood, injuring the vein with a laser, and studying it with confocal microscopy.

Wang *et al.* developed a 3D ABM spanning molecular signaling network and multicellular scales based on prior work (Wang *et al.*, 2007) to quantify and predict the relationship between extracellular stimuli, intracellular signaling dynamics, and multicellular lung cancer growth and expansion (Wang *et al.*, 2009). At the molecular scale, they implemented an epidermal growth factor receptor (EGFR) signaling pathway, triggered by multiple growth factors; at the multicellular scale, they constructed a lattice-based 3D heterogeneous biochemical microenvironment populated with growth factors, glucose, and oxygen. Each cell (agent) carries a self-maintained EGFR pathway, implying that cells will differ in signaling profiles and thus will exhibit varying phenotypic behaviors as the simulation progresses. To link molecular changes to cell phenotype determinations, the authors established a unique experimentally supported, molecularly driven cellular phenotypic decision algorithm based on experimental findings. Specifically, experimental studies have shown that the transient acceleration of accumulating phospholipase-C gamma-1 (PLC γ) levels leads to cell migration (Dittmar *et al.*, 2002), while that of extracellular signal regulated kinase (ERK) leads to cell replication (Santos *et al.*, 2007); both PLC γ and ERK are downstream signaling components of the EGFR pathway. Hence, the rate of change of PLC γ was used to determine the cellular migration decision and that of ERK to dictate the cellular proliferation fate. With this molecular-multicellular simulation platform, they were not only able to monitor multicellular dynamics in response to individual or combinatorial molecular changes, but could also track molecular signaling dynamics on a single cell level as the tumor system evolves. This model has also been used in identifying and validating potential therapeutic drug targets that are concurrently effective on both molecular and multicellular scales for cancer treatment (Wang *et al.*, 2008, 2011, 2012, 2014; Wang and Deisboeck, 2014). This was done by way of sensitivity analyses, i.e., by evaluating how changes occurring at the molecular level percolate across the scales of the model and affect tumor growth behaviors at the multicellular level.

In addition to these 'in-house' modeling studies, public available ABM software has also been widely employed in biomedical research. Using a Java-based Recursive Porous Agent Simulation Toolkit (RePAST) as a modeling framework, Tang *et al.* (2007) presented an agent-based model to describe

the dynamics of rolling, activation, and adhesion of individual leukocytes. Each leukocyte in their model is an object where intercellular processes were constant and membrane unit processes were executed autonomously. Leukocytes could make contact in a customizable 'contact zone,' which can be divided into arbitrary membrane units. The leukocytes can roll on the surface, bind to available ligands, and will be activated after binding; yet, these bonds can be broken as well. The authors performed separate simulations on leukocytes rolling with two ligands: P-selectin and VCAM-1, and found that the simulation results were comparable to experimental distance and velocity measurements made using a parallel-plate flow chamber. The authors concluded that their model will be useful for studying subcellular molecular events and changes that may be involved in phenotypic transitions that lead to disease-related adhesion.

EFA

The so-called EFA is a framework for computer-aided multiscale analysis, enabling one to perform macroscopic computational tasks (over long temporal and large spatial scales) by using only appropriately initialized microscopic simulations (on short time and small length scales). A key feature of this approach is that it eliminates the derivation of explicit macroscopic evolution equations (Kevrekidis and Samaey, 2009). This is very important because, for many biological systems, the evolutionary macroscopic behavior can be seen in practice but cannot be described in closed mathematical forms. The essential concept behind EFA is a coarse time-stepper, which performs those coarse-scale computations for systems that have been modeled at a finer scale (Kevrekidis and Samaey, 2009). The coarse time-stepper achieves a time step of an unavailable macroscopic model as a three-step process: coarse projective integration, gap-tooth scheme, and patch dynamic. Through these steps, the direct derivation of macroscopic equations is bypassed by using simulation results at a microscopic scale to estimate the behavior of the model at the macroscopic scale.

A group of animals can display different types of collective motion (e.g., schooling, flocking, and swarming) at different times. EFA has been used to analyze a 1D agent-based model of self-organizing group formation with stochasticity by bridging the agent-based model with coarse population-level analysis (Kolpas *et al.*, 2007). The stochasticity was introduced to the model by adding a small deviation to the heading of each agent obtained from a deterministic evolution algorithm. A framework was developed for coarse analysis of the stochasticity-induced switching between two metastable cohesive collective motion states: the stationary and mobile states. A lifting procedure was used to run simulations for brief bursts of time with the agent-based model at a given value of A , a 'coarse variable' that characterizes the long-term dynamics of the model. Through model analysis, the authors show that animal groups can frequently switch between the two different collective motion states due to stochastic effects, in addition to change of animal behaviors or environmental factors.

EFA has also been used to build efficient numerical algorithms to accelerate stochastic simulation of gene regulatory networks (Erban *et al.*, 2007). The authors combined EFA with a recently developed data mining technique, diffusion maps

(Coifman *et al.*, 2005), which can identify coarse-grained variables dependent on simulation data. Specifically, coarse-grained variables were first identified with the diffusion maps method. Short bursts of a full-scale run of a stochastic simulation algorithm (SSA; Gillespie, 1977) were then designed and processed. These bursts of SSA were then used to numerically solve the evolution equations for the coarse-grained variables. The authors analyzed a two-gene network (Gardner *et al.*, 2000) as an example where each represses the transcription of the other gene, and found through their analysis that the previously known variables were indeed significant coarse-grained coordinates. The authors claim that this combined ‘variable-free, equation-free’ analysis can be a promising approach for determining features of the long-time, coarse-grained behavior for certain classes of complex stochastic models, including gene regulatory networks, as an alternative to long, full-scale SSA simulations.

An adaptive combined cell-based and continuum approach

Othmer’s group has been investigating an adaptive hybrid modeling approach where both discrete and continuum representations of cells are chosen dynamically and adaptively (Kim *et al.*, 2007, 2011). Their approach has been primarily used to describe and understand solid tumors. Specifically, cells in the outermost proliferating region are treated individually as deformable ellipses. Such a treatment of the outermost zone permits the inclusion of critical cellular variables, such as cell cycle time, cell size, metabolic state, effect of stress on growth of cells, and biochemical and biomechanical interactions between cells and their surroundings. However, it would be both redundant and computationally too expensive to repeat the same treatment in other regions of the tumor. The quiescent zone and necrotic core are thus treated as homogeneous viscoelastic continua, but each with a different material property. Together, instead of treating the whole tumor mass as individual cells, the cell-based descriptions are restricted to a subset of cells (usually a relatively small number) only, without compromising the efficiency and accuracy of the model. The reaction-diffusion equations for oxygen and glucose are normally solved for the entire domain. By including alterations in cell-scale parameters and material properties of continuum regions, the effect on tumor behavior and growth can be investigated.

This adaptive hybrid approach has been applied to ductal carcinoma *in situ* for studying the interactions of epithelial cells within breast ducts with the tumor microenvironment, and to understand how these interactions affect tumor growth and development (Kim and Othmer, 2013). The epithelial cells and transformed epithelial cells were all treated with the cell-based approach rather than discretizing them into zones of proliferation, quiescence, or necrosis. Their temporal evolution was determined similarly to that of proliferating cells in a spheroid. The microenvironment surrounding the duct is modeled as a continuum with the fibroblasts and myofibroblasts assumed as point sources or sinks of growth factors. Mechanical force balance and reaction-diffusion equations governed the evolution of growth factors, like EGF and transforming growth factor- β (TGF- β). This model also provided insight into potential therapeutic interventions that are beyond the conventional therapeutic targets of killing tumor cells. Model results demonstrate

that blocking signaling pathways involving TGF- β could delay transformed epithelial cell activation and thus restricted their population. The model highlighted the potential of tumor stroma as a target for manipulating host–tumor interactions to prevent tumor invasion. This model can also serve as a platform for understanding biomechanics of tumor cell invasion through the basement membrane and the role of proteinases secreted by fibroblasts and tumor cells.

Brief introduction of mean field theory, heterogeneous multiscale method, and dynamic density functional theory

Mean field theory is a method derived from statistical physics for reducing a high-dimensional problem into a low-dimensional one with macroscopic variables and has been widely used in modeling a variety of biological systems (Baker and Simpson, 2010; Cui *et al.*, 2009; Loew *et al.*, 2009; Naundorf *et al.*, 2006; Rovetti *et al.*, 2010). This method assumes that all the elements involved in a system can be controlled by a single mean field. It has been proven that this method can produce accurate results when the elements are globally coupled or when the system is well mixed, but when the temporal and spatial fluctuations of those elements are large, the mean field representation becomes inaccurate. Nevertheless, mean field theory is a good approximation under many conditions and can provide mechanistic insights. One may be able to employ this method to develop a useful low-dimensional representation of a high-dimensional system, which would be very valuable for multiscale modeling. Other innovative multiscale modeling methods, including the heterogeneous multiscale method (HMM; E *et al.*, 2003, 2007) and the dynamic density functional theory (DDFT; Chauviere *et al.*, 2012), have also been used to functionally link biological processes that occur at different scales. In particular, the capability of the DDFT approach in capturing multiple sources of nonlinearity from the spatiotemporal heterogeneity of soft tissue can be very helpful for modeling heterogeneous tissues (such as a tumor).

Discussion

The complexity of a biological system has made the use of mathematical and computational models, along with the analysis of their behavior and functions, an active area of research in recent years. However, any such system’s emergent behavior is highly unpredictable by a traditional modeling approach where hypothesis testing assumes that the modeler has a ‘clear’ idea of how the system functions (Holcombe *et al.*, 2012). Due to the multiscale nature of biological systems, one must link its physiological processes that occur at different scales in a consistent manner, so that the information from a smaller scale can be carried over into the simplified model of a larger scale (Kohl and Noble, 2009). Reality, however, is that there are many open avenues of research where multiscale efforts are either sparsely represented or completely nonexistent. This then presents an opportunity to push the boundaries of knowledge in these biomedical investigations, using multiscale modeling as a platform for hypothesis generation and testing. Overall, multiscale modeling offers computational biologists the unique advantage of being able to introduce perturbations across any tier of resolution (Vicini, 2010).

We have introduced a number of advanced multiscale modeling techniques. However, one has to keep in mind that under any circumstances, as with any advanced method, multiscale models introduce certain simplifications to render them computationally feasible. Hence, validation of these models is required to test the reliability of data transfer and the adopted cross-linking methods between its scales. As with other modeling efforts, a comprehensive validation is critical to achieve acceptance in the community by experimental scientists and clinicians (Koay *et al.*, 2014). Given that in many situations, experimental cross-scale data are still difficult to find or simply unavailable, one may perform cross-validation to prove the predictive power and reliability of the model, i.e., using different modeling methods to examine the same biological system (Andasari *et al.*, 2012).

The roles of standards in multiscale modeling cannot be underestimated. Standards such as SBML (Hucka *et al.*, 2003), CellML (Lloyd *et al.*, 2004), and FieldML (Christie *et al.*, 2009) have begun to play important roles in multiscale modeling (Popel and Hunter, 2009). Having the models, data, and metadata available in standardized formats with clearly stated dependencies will improve the utility of models and facilitate the creation of powerful workflows for comparing model predictions with experimental data (Sloot and Hoekstra, 2010). Furthermore, the simulation of large-scale multiscale models (e.g., the IUPS Physiome project (Hunter, 2004)) will likely have to span several (academic, governmental, and commercial) organizations, thus requiring efficient integration and management of models (and associated data), from creation and initial storage, to revision, up to the time when a model becomes obsolete and is deleted. Web services offer a flexible means of integrating often incompatible applications and tools for data acquisition, registration, storage, provenance, organization, analysis, and presentation (Wang *et al.*, 2013).

In summary, successful analysis of any biological system requires a better understanding of the functional interactions between the key components of molecules, cells, and organs, as well as how these interactions change in disease states. By drawing on the strengths of the powerful multiscale methods introduced herein and integrating them into the next generation of simulation workflows, we will be able to produce computationally more efficient models to more accurately describe and to better simulate biological processes of interest, and to plan for and predict treatment impact more successfully.

Acknowledgments

This work has been supported in part by the National Science Foundation (NSF) Grant DMS-1263742 (Z.W., V.C.), King Abdulaziz University grant 54-130-35-HiCi (V.C.), the National Institutes of Health (NIH) Grant 1U54CA149196, 1U54CA143837, 1U54CA151668, and 1U54CA143907 (V.C.), the University of New Mexico Cancer Center Victor and Ruby Hansen Surface Professorship in Molecular Modeling of Cancer (V.C.), the Methodist Hospital Research Institute (V.C.), and the Harvard-MIT (HST) Athinoula A. Martinos Center for Biomedical Imaging and the Department of Radiology at Massachusetts General Hospital (T.S.D.).

See also: Vertical Integration: Modeling: Vertical Integration

References

- An, G., Mi, Q., Dutta-Moscato, J., Vodovotz, Y., 2009. Agent-based models in translational systems biology. *Wiley Interdisciplinary Reviews Systems Biology and Medicine* 1, 159–171.
- Andasari, V., Roper, R.T., Swat, M.H., Chaplain, M.A., 2012. Integrating intracellular dynamics using CompuCell3D and Bionetsolver: Applications to multiscale modelling of cancer cell growth and invasion. *PLoS One* 7, e33726.
- Anderson, A.R., Chaplain, M.A., Rejniak, K.A., 2007. *Single-Cell-Based Models in Biology and Medicine*. New York, NY: Springer.
- Auffray, C., Chen, Z., Hood, L., 2009. Systems medicine: The future of medical genomics and healthcare. *Genome Medicine* 1, 2.
- Baker, R.E., Simpson, M.J., 2010. Correcting mean-field approximations for birth-death-movement processes. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics* 82, 041905.
- Barabasi, A.L., 2005. Taming complexity. *Nature Physics* 1, 68–70.
- Bradley, C., Bowery, A., Britten, R., *et al.*, 2011. OpenCMISS: A multi-physics & multi-scale computational infrastructure for the VPH/Physiome project. *Progress in Biophysics & Molecular Biology* 107, 32–47.
- Byrne, H., Drasdo, D., 2009. Individual-based and continuum models of growing cell populations: A comparison. *Journal of Mathematical Biology* 58, 657–687.
- Chauviere, A., Hatzikirou, H., Kevrekidis, I.G., Lowengrub, J.S., Cristini, V., 2012. Dynamic density functional theory of solid tumor growth: Preliminary models. *AIP Advances* 2, 11210.
- Christie, G.R., Nielsen, P.M., Blackett, S.A., Bradley, C.P., Hunter, P.J., 2009. FieldML: Concepts and implementation. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences* 367, 1869–1884.
- Coifman, R.R., Lafon, S., Lee, A.B., *et al.*, 2005. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the National Academy of Sciences of the United States of America* 102, 7426–7431.
- Cui, X., Rovetti, R.J., Yang, L., *et al.*, 2009. Period-doubling bifurcation in an array of coupled stochastically excitable elements subjected to global periodic forcing. *Physical Review Letters* 103, 044102.
- Das, H., Wang, Z., Niazi, M.K., *et al.*, 2013. Impact of diffusion barriers to small cytotoxic molecules on the efficacy of immunotherapy in breast cancer. *PLoS One* 8, e61398.
- De Graaf, A.A., Freidig, A.P., De Roos, B., *et al.*, 2009. Nutritional systems biology modeling: From molecular mechanisms to physiology. *PLoS Computational Biology* 5, e1000554.
- Deisboeck, T.S., Wang, Z., Macklin, P., Cristini, V., 2011. Multiscale cancer modeling. *Annual Review of Biomedical Engineering* 13, 127–155.
- Dittmar, T., Husemann, A., Schewe, Y., *et al.*, 2002. Induction of cancer cell migration by epidermal growth factor is initiated by specific phosphorylation of tyrosine 1248 of c-erbB-2 receptor via EGFR. *FASEB Journal* 16, 1823–1825.
- E, W., Engquist, B., Li, X.T., Ren, W.Q., Vanden-Eijnden, E., 2007. Heterogeneous multiscale methods: A review. *Communications in Computational Physics* 2, 367–450.
- E, W.N., Engquist, B., Huang, Z.Y., 2003. Heterogeneous multiscale method: A general methodology for multiscale modeling. *Physical Review B* 67, 092101.
- Edgerton, M.E., Chuang, Y.L., MacKlin, P., *et al.*, 2011. A novel, patient-specific mathematical pathology approach for assessment of surgical volume: Application to ductal carcinoma in situ of the breast. *Analytical Cellular Pathology (Amsterdam)* 34, 247–263.
- Engler, A.J., Humbert, P.O., Wehrle-Haller, B., Weaver, V.M., 2009. Multiscale modeling of form and function. *Science* 324, 208–212.
- Erbán, R., Frewen, T.A., Wang, X., *et al.*, 2007. Variable-free exploration of stochastic models: A gene regulatory network example. *Journal of Chemical Physics* 126, 155103.
- Gardner, T.S., Cantor, C.R., Collins, J.J., 2000. Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403, 339–342.
- Gillespie, D.T., 1977. Exact stochastic simulation of coupled chemical-reactions. *Journal of Physical Chemistry* 81, 2340–2361.
- Grima, R., 2008. Multiscale modeling of biological pattern formation. *Current Topics in Developmental Biology* 81, 435–460.

- Hatzikirou, H., Chauviere, A., Bauer, A.L., *et al.*, 2012. Integrative physical oncology. Wiley Interdisciplinary Reviews Systems Biology and Medicine 4, 1–14.
- Holcombe, M., Adra, S., Bicak, M., *et al.*, 2012. Modeling complex biological systems using an agent-based approach. Integrative Biology (Cambridge) 4, 53–64.
- Hucka, M., Finney, A., Sauro, H.M., *et al.*, 2003. The systems biology markup language (SBML): A medium for representation and exchange of biochemical network models. Bioinformatics 19, 524–531.
- Hunter, P.J., 2004. The IUPS Physiome Project: A framework for computational physiology. Progress in Biophysics & Molecular Biology 85, 551–569.
- Hunter, P.J., Crampin, E.J., Nielsen, P.M., 2008. Bioinformatics, multiscale modeling and the IUPS Physiome Project. Briefings in Bioinformatics 9, 333–343.
- Kevrekidis, I.G., Samaey, G., 2009. Equation-free multiscale computation: Algorithms and applications. Annual Review of Physical Chemistry 60, 321–344.
- Kim, Y., Othmer, H.G., 2013. A hybrid model of tumor–stromal interactions in breast cancer. Bulletin of Mathematical Biology 75, 1304–1350.
- Kim, Y., Stolarska, M.A., Othmer, H.G., 2007. A hybrid model for tumor spheroid growth *in vitro* I: Theoretical development and early results. Mathematical Models and Methods in Applied Sciences 17, 1773–1798.
- Kim, Y., Stolarska, M.A., Othmer, H.G., 2011. The role of the microenvironment in tumor growth and invasion. Progress in Biophysics and Molecular Biology 106, 353–379.
- Koay, E.J., Truty, M.J., Cristini, V., *et al.*, 2014. Transport properties of pancreatic cancer describe gemcitabine delivery and response. Journal of Clinical Investigation 124, 1525–1536.
- Kohl, P., Crampin, E.J., Quinn, T.A., Noble, D., 2010. Systems biology: An approach. Clinical Pharmacology & Therapeutics 88, 25–33.
- Kohl, P., Noble, D., 2009. Systems biology and the virtual physiological human. Molecular Systems Biology 5, 292.
- Kolpas, A., Moehlis, J., Kevrekidis, I.G., 2007. Coarse-grained analysis of stochasticity-induced switching between collective motion states. Proceedings of the National Academy of Sciences of the United States of America 104, 5931–5935.
- Lloyd, C.M., Halstead, M.D., Nielsen, P.F., 2004. CellML: Its future, present and past. Progress in Biophysics & Molecular Biology 85, 433–450.
- Loew, S., Hindert, A., May, S., 2009. Stability of protein-decorated mixed lipid membranes: The interplay of lipid–lipid, lipid–protein, and protein–protein interactions. Journal of Chemical Physics 130, 045102.
- Lowengrub, J.S., Frieboes, H.B., Jin, F., *et al.*, 2010. Nonlinear modeling of cancer: Bridging the gap between cells and tumours. Nonlinearity 23, R1–R9.
- Meier-Schellersheim, M., Fraser, I.D., Klauschen, F., 2009. Multiscale modeling for biologists. Wiley Interdisciplinary Reviews Systems Biology and Medicine 1, 4–14.
- Naundorf, B., Wolf, F., Volgushev, M., 2006. Unique features of action potential initiation in cortical neurons. Nature 440, 1060–1063.
- Pascal, J., Ashley, C.E., Wang, Z., *et al.*, 2013a. Mechanistic modeling identifies drug-uptake history as predictor of tumor drug resistance and nano-carrier-mediated response. ACS Nano 7, 11174–11182.
- Pascal, J., Bearer, E.L., Wang, Z., *et al.*, 2013b. Mechanistic patient-specific predictive correlation of tumor drug response with microenvironment and perfusion measurements. Proceedings of the National Academy of Sciences of the United States of America 110, 14266–14271.
- Popel, A.S., Hunter, P.J., 2009. Systems biology and physiome projects. Wiley Interdisciplinary Reviews Systems Biology and Medicine 1, 153–158.
- Qu, Z., Garfinkel, A., Weiss, J.N., Nivala, M., 2011. Multi-scale modeling in biology: How to bridge the gaps between scales? Progress in Biophysics & Molecular Biology 107, 21–31.
- Rovetti, R., Cui, X., Garfinkel, A., Weiss, J.N., Qu, Z., 2010. Spark-induced sparks as a mechanism of intracellular calcium alternans in cardiac myocytes. Circulation Research 106, 1582–1591.
- Sanga, S., Frieboes, H.B., Zheng, X., *et al.*, 2007. Predictive oncology: A review of multidisciplinary, multiscale *in silico* modeling linking phenotype, morphology and growth. NeuroImage 37 (Suppl. 1), S120–S134.
- Santos, S.D., Verveer, P.J., Bastiaens, P.I., 2007. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. Nature Cell Biology 9, 324–330.
- Schnell, S., Grima, R., Maini, P.K., 2007. Multiscale modeling in biology — New insights into cancer illustrate how mathematical tools are enhancing the understanding of life from the smallest scale to the grandest. American Scientist 95, 134–142.
- Sloot, P.M., Hoekstra, A.G., 2010. Multi-scale modeling in computational biomedicine. Briefings in Bioinformatics 11, 142–152.
- Southern, J., Pitt-Francis, J., Whiteley, J., *et al.*, 2008. Multi-scale computational modeling in biology and physiology. Progress in Biophysics & Molecular Biology 96, 60–89.
- Tang, J., Ley, K.F., Hunt, C.A., 2007. Dynamics of *in silico* leukocyte rolling, activation, and adhesion. BMC Systems Biology 1, 14.
- Tracqui, P., 2009. Biophysical models of tumour growth. Reports on Progress in Physics 72, 1–30.
- Vicini, P., 2010. Multiscale modeling in drug discovery and development: Future opportunities and present challenges. Clinical Pharmacology & Therapeutics 88, 126–129.
- Walpole, J., Papin, J.A., Peirce, S.M., 2013. Multiscale computational models of complex biological systems. Annual Review of Biomedical Engineering 15, 137–154.
- Wang, Z., Birch, C.M., Deisboeck, T.S., 2008. Cross-scale sensitivity analysis of a non-small cell lung cancer model: Linking molecular signaling properties to cellular behavior. Biosystems 92, 249–258.
- Wang, Z., Birch, C.M., Sagotsky, J., Deisboeck, T.S., 2009. Cross-scale, cross-pathway evaluation using an agent-based non-small cell lung cancer model. Bioinformatics 25, 2389–2396.
- Wang, Z., Bordas, V., Deisboeck, T.S., 2011. Identification of critical molecular components in a multiscale cancer model based on the integration of Monte Carlo, resampling, and ANOVA. Frontiers in Physiology 2, 35.
- Wang, Z., Bordas, V., Sagotsky, J., Deisboeck, T.S., 2012. Identifying therapeutic targets in a combined EGFR-TGFBetaR signalling cascade using a multiscale agent-based cancer model. Mathematical Medicine and Biology 29, 95–108.
- Wang, Z., Deisboeck, T.S., 2008. Computational modeling of brain tumors: Discrete, continuum or hybrid? Scientific Modeling and Simulation 15, 381–393.
- Wang, Z., Deisboeck, T.S., 2014. Mathematical modeling in cancer drug discovery. Drug Discovery Today 19, 145–150.
- Wang, Z., Deisboeck, T.S., Cristini, V., 2014. Development of a sampling-based global sensitivity analysis workflow for multiscale computational cancer models. IET Systems Biology 8, 191–197.
- Wang, Z., Sagotsky, J., Taylor, T., Shironoshita, P., Deisboeck, T.S., 2013. Accelerating cancer systems biology research through Semantic Web technology. Wiley Interdisciplinary Reviews Systems Biology and Medicine 5, 135–151.
- Wang, Z., Zhang, L., Sagotsky, J., Deisboeck, T.S., 2007. Simulating non-small cell lung cancer with a multiscale agent-based model. Theoretical Biology and Medical Modelling 4, 50.
- Xu, Z., Chen, N., Kamocka, M.M., Rosen, E.D., Alber, M., 2008. A multiscale model of thrombus development. Journal of The Royal Society Interface 5, 705–722.

Further Reading

- Cristini, V., Lowengrub, J., 2010. Multiscale Modeling of Cancer: An Integrated Experimental and Mathematical Modeling Approach. Cambridge: Cambridge University Press.
- Dada, J.O., Mendes, P., 2011. Multi-scale modeling and simulation in systems biology. Integrative Biology (Cambridge) 3, 86–96.
- Deisboeck, T.S., Stamatakis, G.S., 2010. Multiscale Cancer Modeling. Boca Raton, FL: Chapman & Hall.
- Di Ventura, B., Lemerle, C., Michalodimitrakis, K., Serrano, L., 2006. From *in vivo* to *in silico* biology and back. Nature 443, 527–533.
- Powathil, G.G., Swat, M., Chaplain, M.A., 2015. Systems oncology: Towards patient-specific treatment regimes informed by multiscale mathematical modelling. Seminars in Cancer Biology 30, 13–20.

Relevant Website

<http://repast.sourceforge.net/Repast>.

Cardiac Modeling

A Bueno-Orovio, O Britton, A Muszkiewicz, and B Rodriguez, University of Oxford, Oxford, UK

© 2016 Elsevier Inc. All rights reserved.

State-of-the-Art in Cardiac Cell Modeling

In 1960, Denis Noble developed the first mathematical model of the electrophysiological activity of a cardiac cell (Noble, 1960). Since then, the field of computational cardiovascular science has developed into a broad and productive area of research with growing demonstration of translational potential into industrial, regulatory, and clinical applications (Sager *et al.*, 2014). Computational heart models are multiscale both spatially and temporally, and integrate information across subcellular, cellular, tissue, and whole organ levels, and temporal scales (ranging from picoseconds to years, from ion channels opening and closing to remodeling in disease). Through continuous use and refinement, computational cardiac cell models have built a strong basis for their usefulness and credibility in a variety of settings, including investigations of genetic mutations, disease, and pharmacology (Clancy *et al.*, 2007; Sarkar and Sobie, 2011; O'Hara and Rudy, 2012; Wilhelmms *et al.*, 2012a,b; Britton *et al.*, 2013; Zemzemi *et al.*, 2013; Polak *et al.*, 2014).

Cardiac cells (cardiomyocytes) are electrically excitable cells due to the presence of proteins in their membrane which act as mechanisms of ionic transport, such as ion channels, pumps, and exchangers. The transmembrane transport of ions such as potassium, sodium, and calcium results in electrical and concentration gradients between the intracellular and the extracellular media. Under physiological conditions with no external electrical stimulation, cardiomyocytes maintain a negative voltage relative to their environment. However, when stimulated, sodium channels in the cell membrane are activated, altering the balance of currents into the cell and acting to rapidly change the electrical potential across the membrane. This triggers the release of calcium ions from a cellular compartment known as the sarcoplasmic reticulum, which raises the bulk cytoplasmic concentration of calcium and signals the myocyte to contract. The ion channels initially responsible for firing the action potential inactivate, while others (potassium) act to repolarize the myocyte back to its resting potential. This allows the cardiomyocyte to fire another action potential during the next heartbeat. In this manner, the cardiac action potential, first observed by Coraboeuf and Weidmann (1949), allows the heart to control its own pattern of muscular contraction without external stimuli. The action potential generated by each cardiomyocyte can excite nearby myocytes, and thus the electrical signals propagate through the heart.

In the last of a landmark series of papers, Hodgkin and Huxley (1952) showed that their mathematical model of the ionic currents of the isolated squid giant axon could recreate many of the axon's key electrical properties. This was the first mathematical model of a neuronal action potential. Their formulation of ion channel kinetics was promptly adapted to create the first action potential model of a cardiac cell (Noble, 1960). The Hodgkin–Huxley model formulates each ionic current as the product of the permeability of the cell

membrane to the ion type of that current (more commonly called the conductance of that current), multiplied by the difference between the membrane potential and the equilibrium potential for that current (the membrane potential at which the net current flow across the membrane for that ion is zero). The conductances are functions that may depend on membrane potential and time, as well as channel-specific constants. They represent the opening, closing, and inactivation of the ion channels that carry current across the cell membrane.

Cardiac cell electrophysiology models, from the very beginning of the field up to the present day, have used the Hodgkin–Huxley formulation to describe the ionic currents in the model. An alternative approach to describe ion channel gating is the use of Markov chain models (Greenstein *et al.*, 2000; Mazhari *et al.*, 2001; Clancy and Rudy, 2002), which model the transition between the open, closed, and inactive states of the channel, based on the probabilistic transition rates between channel states at different transmembrane potentials.

From the earliest models that were very closely based on the original three current Hodgkin–Huxley model, new models have increased in complexity, based on our growing understanding of cardiac cellular electrophysiology, for example, the discovery of calcium channels. In general, the additions made in each new model have been in one or more of the following forms: (1) additions of new currents carrying ions such as calcium, (2) addition of new gating mechanisms for existing currents, (3) subdividing a single current into multiple currents either because of different time scales for the parts of the current (e.g., the fast and late sodium currents) and/or because different ion channel proteins carry the different components (e.g., slow and rapid delayed rectifier potassium currents), (4) addition of mechanisms relating to intracellular ionic concentrations (e.g., adding ion transporters such as the sodium–potassium pump or new cellular compartments such as the sarcoplasmic reticulum for calcium handling), and finally (5) adding biochemical mechanisms for ion channel modification, contraction, or signaling, such as, for example, models of beta-adrenergic stimulation (Kuzumoto *et al.*, 2008; Heijman *et al.*, 2011) or ion channel regulation by calcium/calmodulin-dependent protein kinase II (Christensen *et al.*, 2009; O'Hara *et al.*, 2011; Koval *et al.*, 2012). The formulation of the different ion channels and transporters is usually inherited between models, although particularized for the specific cell type and species under study (Niederer *et al.*, 2009; Bueno-Orovio *et al.*, 2014b). This is illustrated in Figure 1 for the particular case of the sodium–potassium pump.

Computational cardiac models have played a pivotal historical role in the discovery of new ionic mechanisms through successes and failures in reproducing and challenging experimental evidence (Noble, 2011). As *in silico* models and simulation techniques developed both conceptually and

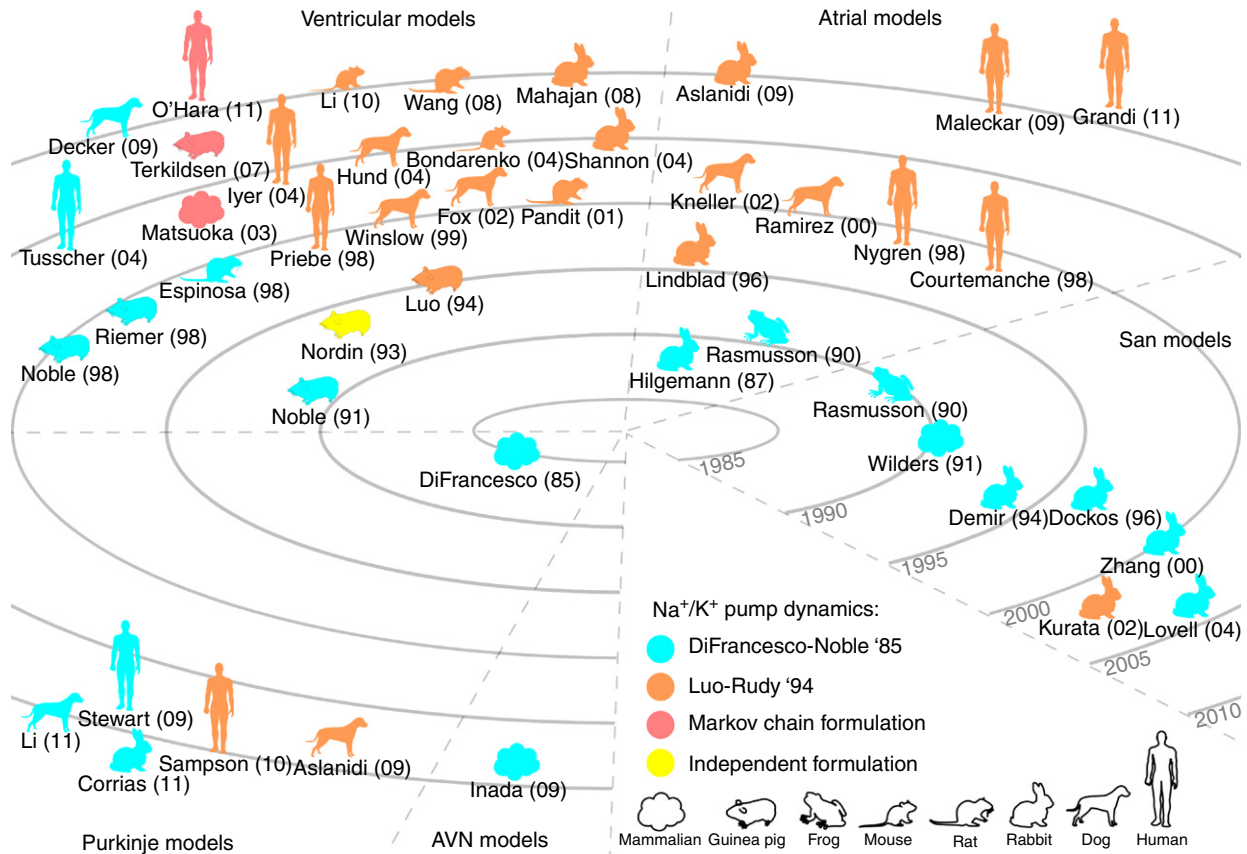


Figure 1 Inheritance in the formulation of the sodium-potassium pump current in a selection of models of cellular cardiac electrophysiology. Models are referred by first author's name and year of publication. The complete list of references can be found in (Noble *et al.*, 2012). Reproduced with permission from Bueno-Orovio, A., Sánchez, C., Pueyo, E., Rodríguez, B., 2014b. Na/K pump regulation of cardiac repolarization: Insights from a systems biology approach. *Pflügers Archiv: European Journal of Physiology* 466, 183–193.

technically, the questions of verification, validation, and uncertainty quantification have become crucial for their uptake beyond scientific research into industrial, clinical, and regulatory arenas (Carusi *et al.*, 2012; Sager *et al.*, 2014; Wallman *et al.*, 2014; Pathmanathan *et al.*, 2015).

The process of validation of *in silico* models and simulations is complex and still under investigation, but it is essential to increase the credibility of *in silico* methods as powerful techniques to augment the information extracted from experimental and clinical investigations. The paper by Carusi *et al.* (2012) specifically explores the issue of validation of computational physiological models in the context of their construction and applications. It highlights the iterative nature of the process of validation and the fact that it needs to consider the ensemble of equations, parameters, simulation techniques, software, and experiments used for a particular study or purpose, namely the Model-Simulation-Experiment system. The aim is to augment the information on the physiological system or process under investigation, such as, for example, the mode of action of a new medicine on specific organs. The dynamics of the iterative process are driven by advances in both experimental and computational techniques, as well as investigations on their combined use.

The legacy of such intense work is over 40 available mathematical models of cardiac cellular electrophysiology, for

different species and cell types (see Noble *et al.* (2012) for a comprehensive historical overview). The iterative process between models, simulations, and experiments leads to the dynamic replacement, refinement, and expansion of these models. In this context, initiatives such as the CellML repository provide the platform that allows the exchange and widespread use of cardiac cell models within the scientific community (Beard *et al.*, 2009).

Human Models of Cardiac Cell Electrophysiology

Computer modeling of the human heart has been a key priority for the scientific community as it provides a vital tool for a better understanding of human health and disease in clinical and industrial applications. Over the last decades a number of models of human adult electrophysiology have been proposed for both ventricular and atrial cells. Biophysically detailed models represent biophysical processes such as ion transport across the membrane and intracellular calcium handling, which are described by equations that represent individual ionic currents and fluxes. A different approach is based on phenomenological models, which reproduce the electrophysiological behavior of cardiac cells without detailed modeling of the underlying biophysical processes. These models

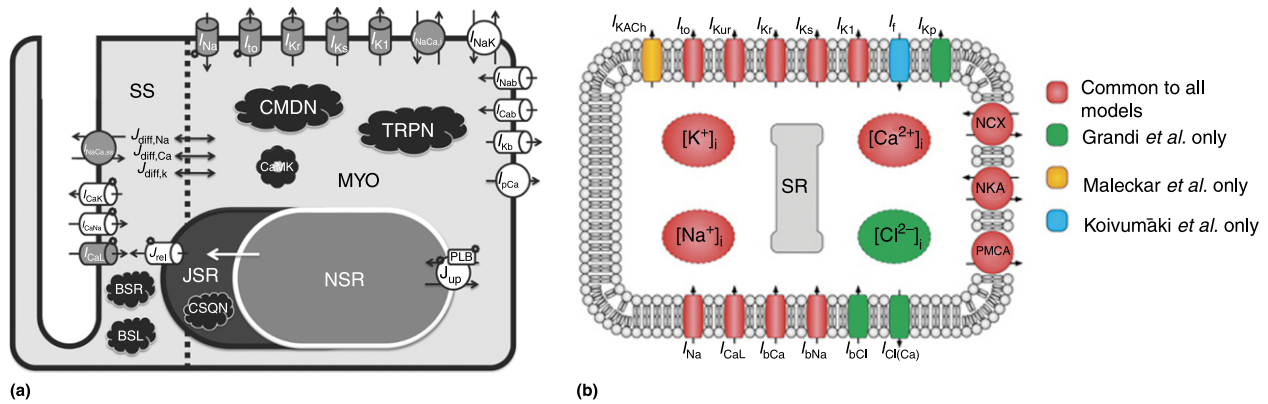


Figure 2 Schematic diagram representations of models of human cellular cardiac electrophysiology. (a) O'Hara-Rudy model of the human ventricular cardiomyocyte. (b) Comparison of different models of human atrial cellular electrophysiology. Reproduced with permission from O'Hara, T., Virág, L., Varró, A., Rudy, Y., 2011. Simulation of the undiseased human cardiac ventricular action potential: Model formulation and experimental validation. *PLoS Computational Biology* 7, e1002061; Wilhelms, M., Hettmann, H., Maleckar, M.M., *et al.*, 2012a. Benchmarking electrophysiological models of human atrial myocytes. *Frontiers in Physiology* 3, 487.

require fewer equations and parameters, and are ideal for applications that do not necessarily require the representation of ionic processes, such as the impact of cellular level properties on the cardiac electrical propagation at the tissue level.

Several human ventricular and atrial cell models have been proposed and compared over the last two decades. Their structure is illustrated in Figure 2, through a schematic representation of the O'Hara-Rudy human ventricular model (O'Hara *et al.*, 2011), and a graphical comparison of the ionic currents included in different human atrial models (Wilhelms *et al.*, 2012a,b). Human ventricular cell models include the biophysically detailed Priebe and Beuckelmann (Priebe and Beuckelmann, 1998), Iyer (Iyer *et al.*, 2004), ten Tusscher and Panfilov (ten Tusscher *et al.*, 2004; ten Tusscher and Panfilov, 2006), Grandi (Grandi *et al.*, 2010), Carro (Carro *et al.*, 2011), O'Hara-Rudy (O'Hara *et al.*, 2011), and Asakura (Asakura *et al.*, 2014) models, as well as phenomenological models for human ventricular myocardium (Bueno-Orovio *et al.*, 2008; Bueno-Orovio *et al.*, 2012). Equivalently, several models of human atrial electrophysiology have been published such as the Nygren (Nygren *et al.*, 1998), Courtemanche (Courtemanche *et al.*, 1998), Maleckar (Maleckar *et al.*, 2009), Grandi (Grandi *et al.*, 2011), and Koivumäki (Koivumäki *et al.*, 2011) models. The phenomenological Bueno-Orovio model has also been adapted to reproduce characteristic action potentials of human atrial tissue (Weber *et al.*, 2008). Comprehensive comparisons between these ventricular and atrial human models can be found, for example, in Bueno-Orovio *et al.* (2008), Wilhelms *et al.* (2012a,b), and Elsharif and Cherry (2014).

Modern human cardiac cell models consist of a system of ordinary differential equations, generally based on the Hodgkin and Huxley or Markov chain formulations for ionic gating variables. As shown in Figure 2, a number of biophysical components are represented, including ionic currents and cellular compartments. Ionic currents that are usually included in human ventricular cell models are the sodium current (I_{Na}), sometimes divided on its fast (I_{NaF}) and late (I_{NaL}) components; the L-type calcium current (I_{CaL}); the transient outward potassium current (I_{to}); the rapid (I_{Kr}) and slow (I_{Ks})

delayed rectifier potassium currents; the inward rectifier potassium current (I_{K1}); the sodium-calcium exchanger current (I_{NCX}); and the sodium-potassium pump current (I_{NaK}). Human atrial cell models incorporate additional ionic currents specific to atrial electrophysiology, such as the ultra-rapid delayed rectifier (I_{Kur}) or the acetylcholine activated (I_{ACh}) potassium currents, as well as differences in calcium handling between ventricular and atrial cells (Figure 2). Usually these models contain at least three compartments, the bulk cytoplasm, encompassing the bulk of the cell and the majority of the cell membrane, where most ionic currents enter or exit the cell and where intracellular sodium, calcium, and potassium ionic concentrations are modeled; a cytoplasmic subspace representing the space near the T-tubules, through which the L-type calcium current flows; and one or more compartments comprising the sarcoplasmic reticulum, with modeled uptake of calcium ions from the bulk cytoplasm by sarco/endoplasmic reticulum Ca^{2+} -ATPase (SERCA) and release into the subspace compartment by calcium-induced calcium-release through ryanodine receptors. Other currents, biochemical and signaling processes, and cellular subdivisions exist, but these differ from model to model.

Modeling Variability in Cardiac Cell Electrophysiology

Variability is essential to biology. No two measurements are the same, no two individuals are the same, and even the same individual will change over time due to aging, disease, or the environment. This variability is exhibited at all biological levels, from cells to organs, with broad implications in our individual response to disease or drug action. However, intersubject variability is usually combined with measurement errors in experiments, and expressed as part of the error in mean values of measured quantities. This permeates into computational models of cardiac cellular electrophysiology, which are commonly fitted to the mean data, therefore representing an 'average' individual.

As discussed in the previous section, cellular cardiac models consist of a system of ordinary differential equations, with

a set of parameters controlling the magnitude and kinetics of the ionic currents included in the model. A common assumption is that variability enters the model at the level of parameters and not at the level of equations. That is, for a specific cell type under investigation, the biophysical processes are the same between individuals, but the quantitative properties of their ionic currents may differ. Two main approaches probing variability in this manner have been described in the literature: these are sensitivity analysis and population of models-based methods.

Sensitivity analysis methods are those where one or more parameters are varied around an existing parameter set (most often the original published parameter set of the model), to quantify the effect this has on model outputs. This approach has been used in cardiac electrophysiological studies. Romero *et al.* (2009) performed a single parameter sensitivity analysis and compared the results to experimental measurements, in order to determine the impact of variability in human ventricular ionic properties on important preclinical biomarkers of arrhythmic risk. In contrast, Sobie (2009) varied multiple parameters simultaneously in a variety of ventricular cell models by sampling their values from a statistical log-normal distribution, centered at the original parameter values (Sarkar and Sobie, 2011; Sarkar *et al.*, 2012). Following this, they used a multivariate regression analysis to relate changes in parameters to model outputs. Using similar methods, Walmsley *et al.* (2013) investigated the role of mRNA expression levels (genetic messengers of ion channel expression in the cell membrane) on electrophysiological remodeling in failing human hearts. More recently, these techniques were used by Cummins *et al.* (2014), who compared 13 ventricular cell models in the context of drug action at different stimulation frequencies.

The second methodology used for modeling intersubject variability is to create and analyze experimentally calibrated populations of models (Britton *et al.*, 2013). To generate a population of models, a subset of the model parameters are varied within the population, based on the parameters known or hypothesized to vary most between individuals, and sampled using random methods such as Monte Carlo or Latin hypercube sampling (McKay *et al.*, 1979). These parameter sets are subsequently inserted into the model equations and simulated. Experimental calibration is then conducted by imposing the requirement that all the considered action potential properties must lay within the experimentally observed bounds, for all the simulated protocols. Therefore, models that are not fully consistent with the experimental data are discarded from the population. This is referred to as the calibration step, and is the major difference between the population of models and sensitivity analysis-based methods. The ensemble of models in full agreement with the experimental data is termed the experimentally calibrated population of models, and used for further investigations. Figure 3 illustrates simulated action potential traces using an experimentally calibrated population of human atrial models (Sánchez *et al.*, 2014).

The population of models approach was originally developed in the field of neuroscience (Prinz *et al.*, 2003; Taylor *et al.*, 2009; Marder and Taylor, 2011). Its use in cardiac electrophysiology was pioneered by the work of Britton *et al.*

(2013), who also demonstrated its quantitative predictive power under pharmacological action. The approach has later been applied to investigating sources of variability in cellular repolarization in rabbit ventricular electrophysiology (Gem-mell *et al.*, 2014), mechanisms of human intersubject variability in sinus rhythm versus chronic atrial fibrillation as illustrated in Figure 3 (Sánchez *et al.*, 2014), and the investigation of pro-arrhythmic mechanisms in human hypertrophic cardiomyopathy (Passini *et al.*, 2014), among others. A study by Davies *et al.* (2012) also used a population of models approach to create representative parameter sets for each of the individuals in their dataset. These studies illustrate some of the different ways in which populations of models can be used to address diverse research questions, depending on the question being asked and the experimental data that is available.

Another source of variability in cardiac electrophysiology is the temporal or beat-to-beat variability of the action potential, which is modulated by disease and pharmacological drug action (Hondeghe *et al.*, 2001; Thomsen *et al.*, 2004; Myles *et al.*, 2008; Hinterseer *et al.*, 2010; Johnson *et al.*, 2010). Recent simulation studies using stochastic cardiac models provide supporting evidence to the contribution of the intrinsically stochastic dynamics of cardiac ion channels to beat-to-beat variability and repolarisation abnormalities under pharmacological action (Lemay *et al.*, 2011; Pueyo *et al.*, 2011; Heijman *et al.*, 2013). Interestingly, these instabilities are modulated and suppressed by intercellular tissue coupling, highlighting the multiscale nature of sources and modulators of cardiac variability (Pueyo *et al.*, 2011). These results also highlight the importance of multiscale simulations in elucidating the complex interplay of mechanisms determining the spatio-temporally dynamic and variable response of the human heart.

Whole Organ Cardiac Modeling

Cardiac cell models are often embedded in tissue and whole organ preparations in order to investigate the impact of ionic alterations on the propagation of electrical excitation in health and disease. Tissue and whole organ models have been extensively used in computational cardiovascular science during the last 15 years, as recently reviewed, for example, in Trayanova (2011) and Smaill *et al.* (2013).

Simulations of cardiac electrophysiology at the tissue and whole organ levels require mathematical modeling of the processes underlying intercellular electrical coupling and propagation of electrical excitation from cell to cell through cardiac tissue. The cardiac bidomain model is currently the gold standard, and arises by representing the cardiac tissue as a continuum with the intracellular and the extracellular spaces electrically connected through the membrane of cardiomyocytes (Tung, 1978). The bidomain model consists of two partial differential equations representing the tissue electrical properties in the intracellular and extracellular domains, coupled through the system of ordinary differential equations representing the membrane kinetics (as described in the previous sections). The partial differential equations assume that cardiac tissue is a continuum and are derived using two fundamental laws: Ohm's law (relating electrical potential to the

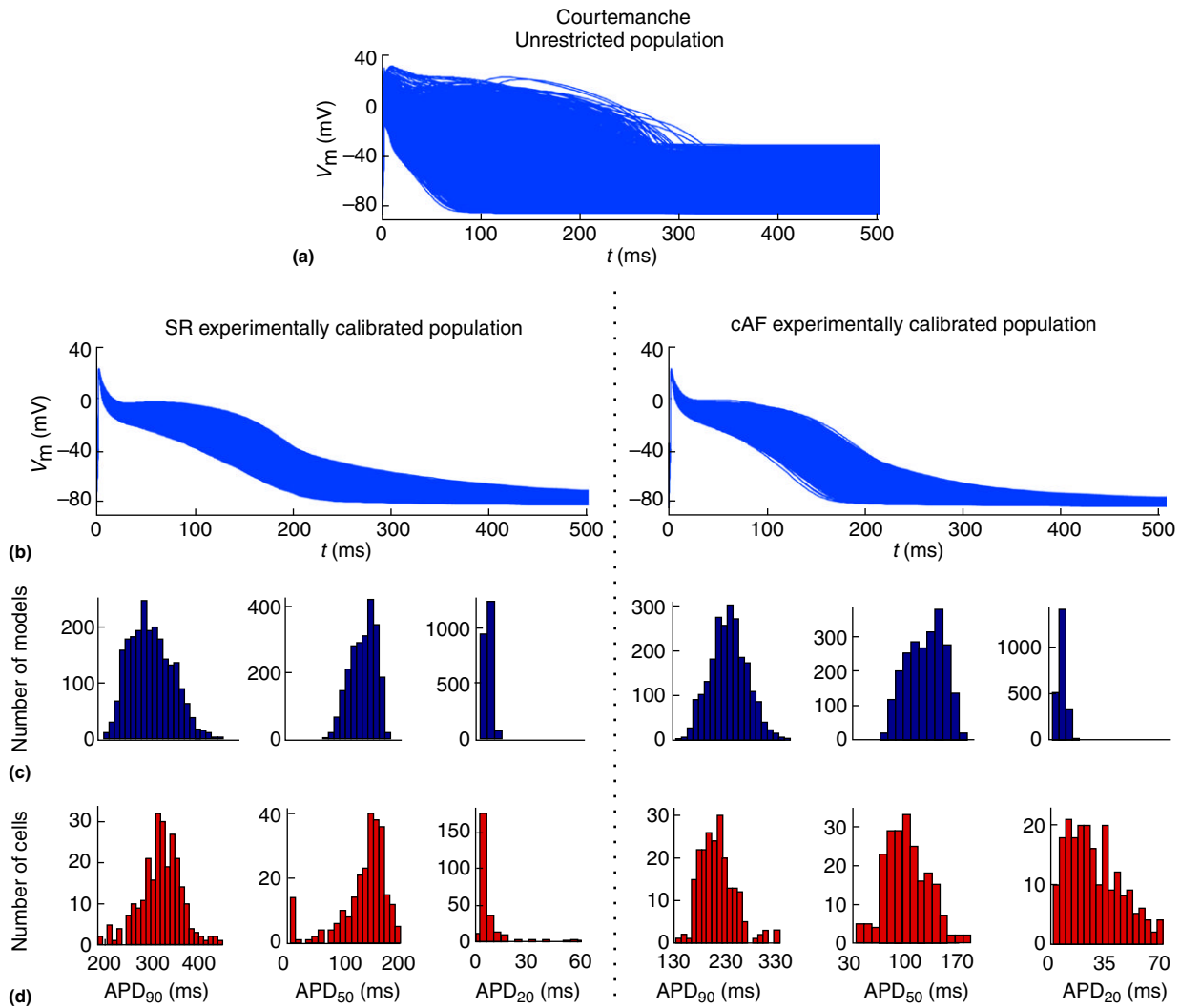


Figure 3 Experimentally calibrated population of human atrial action potential models for sinus rhythm (SR) and chronic atrial fibrillation (cAF). (a): Human atrial action potentials simulated using an initially unrestricted set of model parameters, constructed by sampling ion channel conductances from $\pm 100\%$ of their original values. (b): Human atrial action potential populations after calibration against experimental bounds for SR and cAF samples. (c), (d): Histograms corresponding to action potential duration at 90, 50, and 20% of repolarization (APD_{90} , APD_{50} , and APD_{20}), in both the populations of models (c) and experimental measurements (d). Reproduced with permission from Sánchez, C., Bueno-Orovio, A., Wettwer, E., *et al.*, 2014. Inter-subject variability in human atrial action potential in sinus rhythm versus chronic atrial fibrillation. PLoS ONE 9, e105897.

net flow of transmembrane, intracellular, and extracellular currents) and Kirchhoff's law (for conservation of charge).

Generating solutions of the bidomain equations is computationally expensive, requiring application of sophisticated numerical techniques, such as the finite element method. Simplified formulations of the bidomain equations have been derived to decrease the high computational expense of numerical simulations, and alternative formulations include the monodomain and the Eikonal equations, as well as graph-based methods (Potse *et al.*, 2006; Wallman *et al.*, 2012). Novel modeling approaches considering fractional diffusion models have been recently applied to cardiac modeling (Bueno-Orovio *et al.*, 2014a). Pushing the boundaries beyond the perfect continuum assumption behind the bidomain equations, fractional diffusion cardiac models aim to capture

the effects of the heterogeneous microstructure of cardiac tissue, and therefore allow to investigate implications of heterogeneity in cardiac electrophysiological behavior.

Multiscale whole organ simulations of cardiac electrophysiological behavior also require the definition of spatial tissue characteristics such as its geometry, which are defined by a mesh. Anatomical meshes of the ventricles and the atria are obtained from cardiac magnetic resonance, computer tomography, or histology datasets, using image processing and mesh generation techniques (Bishop *et al.*, 2010; Vadakkumpadan *et al.*, 2010; Bordas *et al.*, 2011; Zhao *et al.*, 2012). The processed images are then used to generate a volumetric mesh by applying a discretization process in space, determined by the numerical method used to solve the propagation model equations (e.g., the finite element method for the bidomain equations).

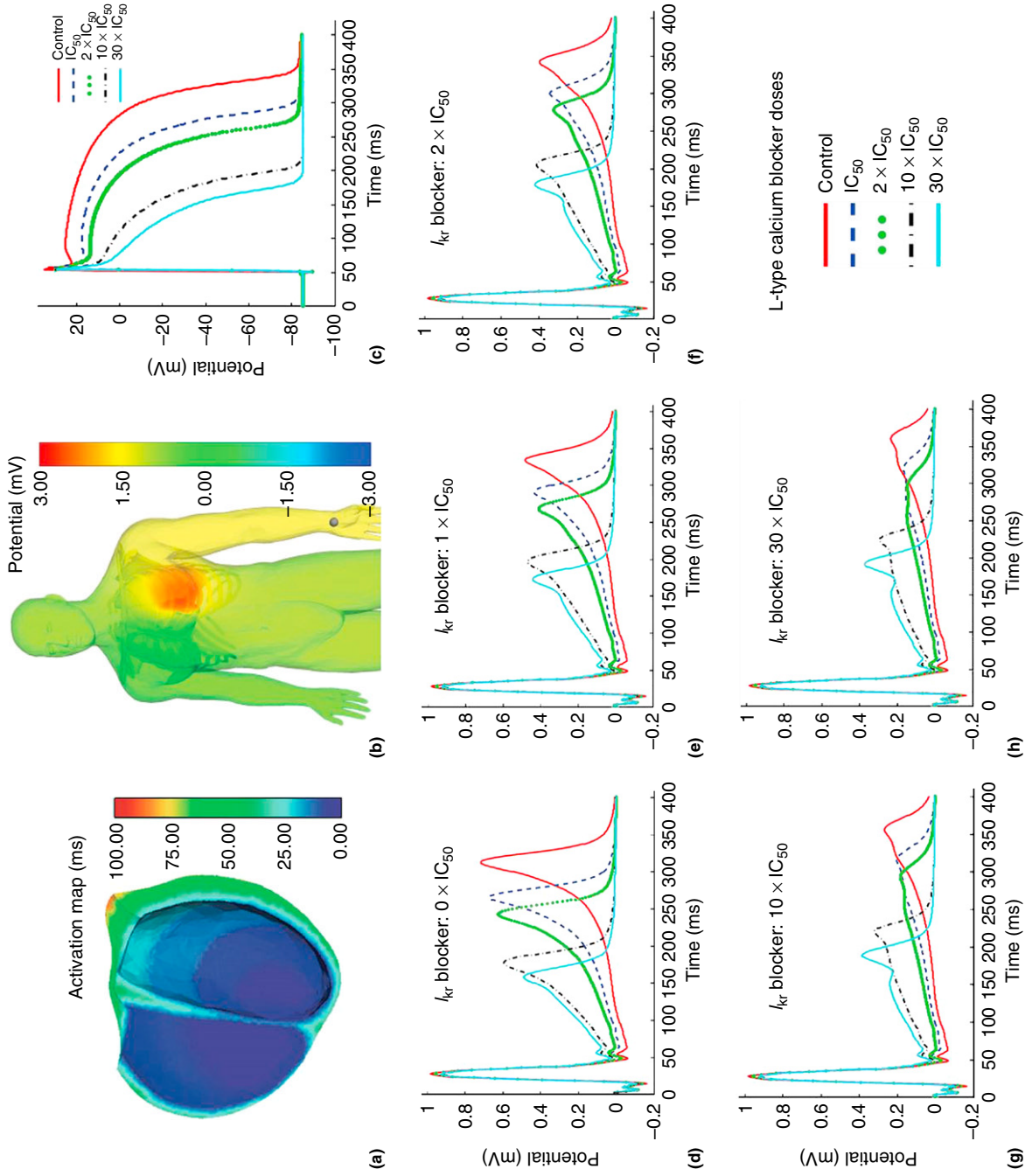


Figure 4 Multiscale simulations of drug-induced effects on the human heart: from drug-ionic current interactions to body surface potentials. (a) Cross-section of the human ventricular model showing the distribution of simulated activation times across the ventricular wall and on the endocardium. (b) Distribution of the body surface potential in the human torso model. (c)–(h) Human ventricular action potentials (c) and electrocardiograms ((d)–(h)), simulated for different degrees of block of L-type calcium and hERG potassium currents (drug dose indicated as a function of the half maximal inhibitory concentration, IC_{50}). Reproduced with permission from Zemzemi, N., Rodriguez, B., 2015. Effects of L-type calcium channel and human ether-a-go-go related gene blockers on the electrical activity of the human heart: A simulation study. *Europace* 17, 326–333.

As a representative example, [Figure 4](#) illustrates simulations of the effects on the body surface electrocardiogram of drug action in calcium and potassium channels, using a multiscale human whole ventricular model with biophysically detailed representation of membrane kinetics, and embedded within a human torso ([Zemzemi and Rodriguez, 2015](#)). This figure exemplifies the potential that these novel technologies have in shaping the future of cardiovascular science. *In silico* modeling in biomedical applications is a research priority that can contribute to improvements in achieving a faster and cheaper drug development process, the reduction of medication errors and adverse side events, and the design of a more personalized healthcare.

Conclusions and Outlook

In this article we have described how cardiac modeling has developed from single cell models of a few cell types into a field with a wide variety of biophysically detailed single cell, tissue, and whole organ models, with a particular focus on models of human physiology. We have also described the importance of phenomena such as intersubject variability in cellular cardiac electrophysiology and heterogeneity in the structure of cardiac tissue, which contribute to variability in the responses of different hearts and patients to drugs and disease, and we have provided an overview of recently developed methods designed to probe both the sources and the consequences of such intersubject differences.

These recent advances in both models and modeling frameworks allow us to investigate how different phenomena interact and affect cardiac behavior at multiple spatio-temporal scales, ranging from single cells, through tissue, to the whole heart, as well as from picoseconds to years, thus opening new avenues to identify key determinants of the electrophysiological activity of the heart in health and disease. A major focus in cardiac modeling is its integration with experimental and clinical research in order to better explain and predict the effects of complex drugs, diseases, and genetic mutations on the human heart. In line with this, computational cardiovascular science is evolving beyond the traditional single physiological human paradigm to construct patient-specific or subpopulation-specific models of whole organ and single cells, respectively. In doing so, we hope to better understand the currently unpredictable risks of lethal pathologies in patients, and enhance prevention and treatment strategies.

References

- Asakura, K., Cha, C.Y., Yamaoka, H., *et al.*, 2014. EAD and DAD mechanisms analyzed by developing a new human ventricular cell model. *Progress in Biophysics and Molecular Biology* 116, 11–24.
- Beard, D.A., Britten, R., Cooling, M.T., *et al.*, 2009. CellML metadata standards, associated tools and repositories. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences* 367, 1845–1867.
- Bishop, M.J., Plank, G., Burton, R.A.B., *et al.*, 2010. Development of an anatomically detailed MRI-derived rabbit ventricular model and assessment of its impact on simulations of electrophysiological function. *American Journal of Physiology. Heart and Circulatory Physiology* 298, H699–H718.
- Bordas, R., Gillow, K., Lou, Q., *et al.*, 2011. Rabbit-specific ventricular model of cardiac electrophysiological function including specialized conduction system. *Progress in Biophysics and Molecular Biology* 107, 90–100.
- Britton, O.J., Bueno-Orovio, A., Van Ammel, K., *et al.*, 2013. Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology. *Proceedings of the National Academy of Sciences of the United States of America* 110, E2098–E2105.
- Bueno-Orovio, A., Cherry, E.M., Fenton, F.H., 2008. Minimal model for human ventricular action potentials in tissue. *Journal of Theoretical Biology* 253, 544–560.
- Bueno-Orovio, A., Hanson, B.M., Gill, J.S., Taggart, P., Rodriguez, B., 2012. *In vivo* human left-to-right ventricular differences in rate adaptation transiently increase pro-arrhythmic risk following rate acceleration. *PLoS ONE* 7, e52234.
- Bueno-Orovio, A., Kay, D., Grau, V., Rodriguez, B., Burrage, K., 2014a. Fractional diffusion models of cardiac electrical propagation: Role of structural heterogeneity in dispersion of repolarization. *Journal of the Royal Society Interface* 11, 20140352.
- Bueno-Orovio, A., Sánchez, C., Pueyo, E., Rodriguez, B., 2014b. Na/K pump regulation of cardiac repolarization: Insights from a systems biology approach. *Pflügers Archiv: European Journal of Physiology* 466, 183–193.
- Carro, J., Rodríguez, J.F., Laguna, P., Pueyo, E., 2011. A human ventricular cell model for investigation of cardiac arrhythmias under hyperkalaemic conditions. *Philosophical Transactions. Series A, Mathematical, Physical, and Engineering Sciences* 369, 4205–4232.
- Carusi, A., Burrage, K., Rodriguez, B., 2012. Bridging experiments, models and simulations: An integrative approach to validation in computational cardiac electrophysiology. *American Journal of Physiology. Heart and Circulatory Physiology* 303, H144–H155.
- Christensen, M.D., Dun, W., Boyden, P.A., *et al.*, 2009. Oxidized calmodulin kinase II regulates conduction following myocardial infarction: A computational analysis. *PLoS Computational Biology* 5, e1000583.
- Clancy, C.E., Rudy, Y., 2002. Na⁺ channel mutation that causes both Brugada and long-QT syndrome phenotypes: A simulation study of mechanism. *Circulation* 105, 1208–1213.
- Clancy, C.E., Zhu, Z.I., Rudy, Y., 2007. Pharmacogenetics and anti-arrhythmic drug therapy: A theoretical investigation. *American Journal of Physiology. Heart and Circulatory Physiology* 292, H66–H75.
- Coraboeuf, E., Weidmann, S., 1949. Les potentiels d'action du muscle cardiaque, obtenus à l'aide de microelectrodes intracellulaires. Présence d'une inversion de potentiel. *Compte Rendu de la Société de Biologie Paris* 143, 1360–1361.
- Courtemanche, M., Ramirez, R.J., Nattel, S., 1998. Ionic mechanisms underlying human atrial action potential properties: Insights from a mathematical model. *American Journal of Physiology* 275, H301–H321.
- Cummins, M.A., Dalal, P.J., Bugana, M., Severi, S., Sobie, E.A., 2014. Comprehensive analyses of ventricular myocyte models identify targets exhibiting favorable rate dependence. *PLoS Computational Biology* 10, e1003543.
- Davies, M.R., Mistry, H.B., Hussein, L., *et al.*, 2012. An *in silico* canine cardiac midmyocardial action potential duration model as a tool for early drug safety assessment. *American Journal of Physiology. Heart and Circulatory Physiology* 302, H1466–H1480.
- Elsheirif, M.M., Cherry, E.M., 2014. A quantitative comparison of the behavior of human ventricular cardiac electrophysiology models in tissue. *PLoS ONE* 9, e84401.
- Gemmell, P., Burrage, K., Rodriguez, B., Quinn, T.A., 2014. Population of computational rabbit-specific ventricular action potential models for investigating sources of variability in cellular repolarisation. *PLoS ONE* 9, e90112.
- Grandi, E., Pandit, S.V., Voigt, N., *et al.*, 2011. Human atrial action potential and Ca²⁺ model: Sinus rhythm and chronic atrial fibrillation. *Circulation Research* 109, 1055–1066.
- Grandi, E., Pasqualini, F.S., Bers, D.M., 2010. A novel computational model of the human ventricular action potential and Ca transient. *Journal of Molecular and Cellular Cardiology* 48, 112–121.
- Greenstein, J.L., Wu, R., Po, S., Tomaselli, G.F., Winslow, R.L., 2000. Role of the calcium-independent transient outward current Ito1 in shaping action potential morphology and duration. *Circulation Research* 87, 1026–1033.
- Heijman, J., Volders, P.G.A., Westra, R.L., Rudy, Y., 2011. Local control of β -adrenergic stimulation: Effects on ventricular myocyte electrophysiology and Ca²⁺ transient. *Journal of Molecular and Cellular Cardiology* 50, 863–871.
- Heijman, J., Zaza, A., Johnson, D.M., *et al.*, 2013. Determinants of beat-to-beat variability of repolarization duration in the canine ventricular myocyte: A computational analysis. *PLoS Computational Biology* 9, e1003202.

- Hinterseer, M., Beckmann, B.-M., Thomsen, M.B., *et al.*, 2010. Usefulness of short-term variability of QT intervals as a predictor for electrical remodeling and proarrhythmia in patients with nonischemic heart failure. *American Journal of Cardiology* 106, 216–220.
- Hodgkin, A.L., Huxley, A.F., 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology* 117, 500–544.
- Hondeghem, L.M., Carlsson, L., Duker, G., 2001. Instability and triangulation of the action potential predict serious proarrhythmia, but action potential duration prolongation is antiarrhythmic. *Circulation* 103, 2004–2013.
- Iyer, V., Mazhari, R., Winslow, R.L., 2004. A computational model of the human left-ventricular epicardial myocyte. *Biophysical Journal* 87, 1507–1525.
- Johnson, D.M., Heijman, J., Pollard, C.E., *et al.*, 2010. IKs restricts excessive beat-to-beat variability of repolarization during beta-adrenergic receptor stimulation. *Journal of Molecular and Cellular Cardiology* 48, 122–130.
- Koivumäki, J.T., Korhonen, T., Tavi, P., 2011. Impact of sarcoplasmic reticulum calcium release on calcium dynamics and action potential morphology in human atrial myocytes: A computational study. *PLoS Computational Biology* 7, e1001067.
- Koval, O.M., Snyder, J.S., Wolf, R.M., *et al.*, 2012. Ca^{2+} /calmodulin-dependent protein kinase II-based regulation of voltage-gated Na^+ channel in cardiac disease. *Circulation* 126, 2084–2094.
- Kuzumoto, M., Takeuchi, A., Nakai, H., *et al.*, 2008. Simulation analysis of intracellular Na^+ and Cl^- homeostasis during beta 1-adrenergic stimulation of cardiac myocyte. *Progress in Biophysics and Molecular Biology* 96, 171–186.
- Lemay, M., de Lange, E., Kucera, J.P., 2011. Effects of stochastic channel gating and distribution on the cardiac action potential. *Journal of Theoretical Biology* 281, 84–96.
- Maleckar, M.M., Greenstein, J.L., Giles, W.R., Trayanova, N.A., 2009. K^+ current changes account for the rate dependence of the action potential in the human atrial myocyte. *American Journal of Physiology. Heart and Circulatory Physiology* 297, H1398–H1410.
- Marder, E., Taylor, A.L., 2011. Multiple models to capture the variability in biological neurons and networks. *Nature Neuroscience* 14, 133–138.
- Mazhari, R., Greenstein, J.L., Winslow, R.L., Marbán, E., Nuss, H.B., 2001. Molecular interactions between two long-QT syndrome gene products, HERG and KCNE2, rationalized by *in vitro* and *in silico* analysis. *Circulation Research* 89, 33–38.
- McKay, M.D., Beckman, R.J., Conover, W.J., 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21 (2), 239–245.
- Myles, R.C., Burton, F.L., Cobbe, S.M., Smith, G.L., 2008. The link between repolarisation alternans and ventricular arrhythmia: Does the cellular phenomenon extend to the clinical problem? *Journal of Molecular and Cellular Cardiology* 45, 1–10.
- Niederer, S.A., Fink, M., Noble, D., Smith, N.P., 2009. A meta-analysis of cardiac electrophysiology computational models. *Experimental Physiology* 94, 486–495.
- Noble, D., 1960. Cardiac action and pacemaker potentials based on the Hodgkin–Huxley equations. *Nature* 188, 495–497.
- Noble, D., 2011. Successes and failures in modeling heart cell electrophysiology. *Heart Rhythm* 8, 1798–1803.
- Noble, D., Garny, A., Noble, P.J., 2012. How the Hodgkin–Huxley equations inspired the Cardiac Physiome Project. *Journal of Physiology* 590, 2613–2628.
- Nygren, A., Fiset, C., Firek, L., *et al.*, 1998. Mathematical model of an adult human atrial cell: The role of K^+ currents in repolarization. *Circulation Research* 82, 63–81.
- O'Hara, T., Rudy, Y., 2012. Quantitative comparison of cardiac ventricular myocyte electrophysiology and response to drugs in human and nonhuman species. *American Journal of Physiology. Heart and Circulatory Physiology* 302, H1023–H1030.
- O'Hara, T., Virág, L., Varró, A., Rudy, Y., 2011. Simulation of the undiseased human cardiac ventricular action potential: Model formulation and experimental validation. *PLoS Computational Biology* 7, e1002061.
- Passini, E., Rodriguez, B., Mincholé, A., *et al.*, 2014. Late sodium current inhibition counteracts pro-arrhythmic mechanisms in human hypertrophic cardiomyopathy. *Computing in Cardiology* 41, 861–864.
- Pathmanathan, P., Shottwell, M.S., Gavaghan, D.J., Cordeiro, J.M., Gray, R.A., 2015. Uncertainty quantification of fast sodium current steady-state inactivation for multi-scale models of cardiac electrophysiology. *Progress in Biophysics and Molecular Biology*. doi:10.1016/j.pbiomolbio.2015.01.008.
- Polak, S., Wiśniowska, B., Fijorek, K., Glinka, A., Mendyk, A., 2014. *In vitro–in vivo* extrapolation of drug-induced proarrhythmia predictions at the population level. *Drug Discovery Today* 19, 275–281.
- Potse, M., Dubé, B., Richer, J., Vinet, A., Gulrajani, R.M., 2006. A comparison of monodomain and bidomain reaction–diffusion models for action potential propagation in the human heart. *IEEE Transactions on Bio-Medical Engineering* 53, 2425–2435.
- Priebe, L., Beuckelmann, D.J., 1998. Simulation study of cellular electric properties in heart failure. *Circulation Research* 82, 1206–1223.
- Prinz, A.A., Billimoria, C.P., Marder, E., 2003. Alternative to hand-tuning conductance-based models: Construction and analysis of databases of model neurons. *Journal of Neurophysiology* 90, 3998–4015.
- Pueyo, E., Corrias, A., Virág, L., *et al.*, 2011. A multiscale investigation of repolarization variability and its role in cardiac arrhythmogenesis. *Biophysical Journal* 101, 2892–2902.
- Romero, L., Pueyo, E., Fink, M., Rodriguez, B., 2009. Impact of ionic current variability on human ventricular cellular electrophysiology. *American Journal of Physiology. Heart and Circulatory Physiology* 297, H1436–H1445.
- Sager, P.T., Gintant, G., Turner, J.R., Pettit, S., Stockbridge, N., 2014. Rechanneling the cardiac proarrhythmia safety paradigm: A meeting report from the Cardiac Safety Research Consortium. *American Heart Journal* 167, 292–300.
- Sánchez, C., Bueno-Orovio, A., Wettwer, E., *et al.*, 2014. Inter-subject variability in human atrial action potential in sinus rhythm versus chronic atrial fibrillation. *PLoS ONE* 9, e105897.
- Sarkar, A.X., Christini, D.J., Sobie, E.A., 2012. Exploiting mathematical models to illuminate electrophysiological variability between individuals. *Journal of Physiology* 590, 2555–2567.
- Sarkar, A.X., Sobie, E.A., 2011. Quantification of repolarization reserve to understand interpatient variability in the response to proarrhythmic drugs: A computational analysis. *Heart Rhythm* 8, 1749–1755.
- Small, B.H., Zhao, J., Trew, M.L., 2013. Three-dimensional impulse propagation in myocardium: Arrhythmogenic mechanisms at the tissue level. *Circulation Research* 112, 834–848.
- Sobie, E.A., 2009. Parameter sensitivity analysis in electrophysiological models using multivariable regression. *Biophysical Journal* 96, 1264–1274.
- Taylor, A.L., Goillard, J.-M., Marder, E., 2009. How multiple conductances determine electrophysiological properties in a multicompartment model. *Journal of Neuroscience* 29, 5573–5586.
- Thomsen, M.B., Verduyn, S.C., Stengl, M., *et al.*, 2004. Increased short-term variability of repolarization predicts d-sotalol-induced torsades de pointes in dogs. *Circulation* 110, 2453–2459.
- Trayanova, N.A., 2011. Whole-heart modeling: Applications to cardiac electrophysiology and electromechanics. *Circulation Research* 108, 113–128.
- Tung, L., 1978. A bi-domain model for describing ischemic myocardial d-c potentials (Ph.D. thesis). Massachusetts: Massachusetts Institute of Technology Massachusetts.
- ten Tusscher, K.H.W.J., Noble, D., Noble, P.J., Panfilov, A.V., 2004. A model for human ventricular tissue. *American Journal of Physiology. Heart and Circulatory Physiology* 286, H1573–H1589.
- ten Tusscher, K.H.W.J., Panfilov, A.V., 2006. Alternans and spiral breakup in a human ventricular tissue model. *American Journal of Physiology. Heart and Circulatory Physiology* 291, H1088–H1100.
- Vadakkumpadan, F., Arevalo, H., Prassl, A.J., *et al.*, 2010. Image-based models of cardiac structure in health and disease. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine* 2, 489–506.
- Wallman, M., Smith, N.P., Rodriguez, B., 2012. A comparative study of graph-based, eikonal, and monodomain simulations for the estimation of cardiac activation times. *IEEE Transactions on Bio-Medical Engineering* 59, 1739–1748.
- Wallman, M., Smith, N.P., Rodriguez, B., 2014. Computational methods to reduce uncertainty in the estimation of cardiac conduction properties from electroanatomical recordings. *Medical Image Analysis* 18, 228–240.
- Walmsley, J., Rodriguez, J.F., Mirams, G.R., *et al.*, 2013. mRNA expression levels in failing human hearts predict cellular electrophysiological remodeling: A population-based simulation study. *PLoS ONE* 8, e56359.
- Weber, F.M., Lurz, S., Keller, D.U.J., *et al.*, 2008. Adaptation of a minimal four-state cell model for reproducing atrial excitation properties. *Computers in Cardiology* 35, 61–64.
- Wilhelms, M., Hettmann, H., Maleckar, M.M., *et al.*, 2012a. Benchmarking electrophysiological models of human atrial myocytes. *Frontiers in Physiology* 3, 487.
- Wilhelms, M., Rombach, C., Scholz, E.P., Dössel, O., Seemann, G., 2012b. Impact of amiodarone and cisapride on simulated human ventricular electrophysiology and electrocardiograms. *Europace* 14, v90–v96.
- Zemzem, N., Bernabeu, M.O., Saiz, J., *et al.*, 2013. Computational assessment of drug-induced effects on the electrocardiogram: From ion channel to body surface potentials. *British Journal of Pharmacology* 168, 718–733.

- Zemzemi, N., Rodriguez, B., 2015. Effects of L-type calcium channel and human ether-a-go-go related gene blockers on the electrical activity of the human heart: A simulation study. *Europace* 17, 326–333.
- Zhao, J., Butters, T.D., Zhang, H., *et al.*, 2012. An image-based model of atrial muscular architecture: Effects of structural anisotropy on electrical activation. *Circulation. Arrhythmia and Electrophysiology* 5, 361–370.

Relevant Website

<http://www.cellml.org/>
CellML.

VERTICAL INTEGRATION: APPLICATIONS

Contents

Mathematical Approaches to Studying Inflammation

Angiogenesis

A Review on Various Mathematical Modeling Approaches for Wound Healing

Multiscale Analysis of Morphogenesis

Neurogenesis in the Adult Brain

Mathematical Approaches to Studying Inflammation

JL Dunster, University of Reading, Reading, UK

I Dransfield, University of Edinburgh, Edinburgh, UK

© 2016 Elsevier Inc. All rights reserved.

Introduction

It is now clear that inflammation, the process by which a tissue responds to injury or infection, is central to the pathogenesis of a broad spectrum of diseases, including rheumatoid arthritis (RA), inflammatory bowel disease, asthma, atherosclerosis, neurodegenerative diseases, and also the development and metastasis of tumors (Medzhitov, 2010). This appreciation of inflammation's relevance has resulted in an increase in interest in studying the underlying cellular and molecular mechanisms of inflammation and, in concert with this, there has been a corresponding growth in interest in mathematically modeling the processes involved. Different approaches to modeling the inflammatory process range from inflammation being the sole target of mathematical and computational models to the key cell types and mediators involved in inflammation being captured in models of different disease scenarios (see Figure 1). Inflammation is a complex, highly regulated process that has components which interact on multiple time and space scales. The different temporal and spatial scales that are inherent in considering inflammation at a molecular, cellular, organ, or whole organism level, as well as the close interaction between inflammation and many other disease processes, make it challenging to study biologically and mathematically. Furthermore, despite recent progress in defining the contribution of different processes, such as leukocyte recruitment, cellular activation, and cell death, the precise interactions of these various elements of inflammation remain to be defined, potentially representing a constraint on the development of new effective therapeutic strategies. Combined mathematical and biological approaches can provide a new way to understand multi-scale systems such as inflammation, linking events over disparate time and space scales that are difficult to model with conventional microbiological or immunological experimental approaches alone (Figure 1).

Inflammation is characterized by key cellular and molecular events that occur in a distinct temporal framework. In response to an injury and/or infection, inflammatory mediators (lipids, peptides, and proteins) are produced that result in alteration of cellular responses (cellular activation) which can promote further inflammatory mediator production (Nathan, 2002). The earliest changes associated with the inflammatory process are vascular alterations, increased blood flow (vasodilatation), and vascular leak (edema) that represent one of the cardinal signs of inflammation (Kvietys and Granger, 2012). In addition, the activation of endothelial cells that line the blood vessel wall promotes the adhesion and subsequent recruitment of inflammatory leukocytes, particularly neutrophil and eosinophil granulocytes (Ley *et al.*, 2007). These cells have an important function in providing 'innate' host defense until a specific 'adaptive' response (which involves development of specific antibodies and cytotoxic lymphocytes) against the invading microorganisms (Kolaczowska and Kubek, 2013). Following removal of the inflammatory stimulus, recruited leukocytes either exit the tissue via the draining lymphatics or undergo a process of programmed cell death (apoptosis) which serves to package the potentially destructive intracellular contents for 'disposal' by other phagocytic cells that are present at the inflammatory site (macrophages) (Fox *et al.*, 2010). Removal of apoptotic cells and the subsequent reprogramming of macrophage function is thought to represent a critical factor for the termination of inflammatory responses (Savill *et al.*, 2002). In addition, macrophages have a pivotal role in tissue regeneration and repair that is required for restoration of normal organ function.

A systems-biology approach aims to utilize mathematical and computational techniques to understand biological processes in a dynamic manner (Kohl *et al.*, 2010). In this article, we will demonstrate the utility of the systems-biology approach in studying inflammation by formulating a simple mathematical model capturing a small subset of the complex

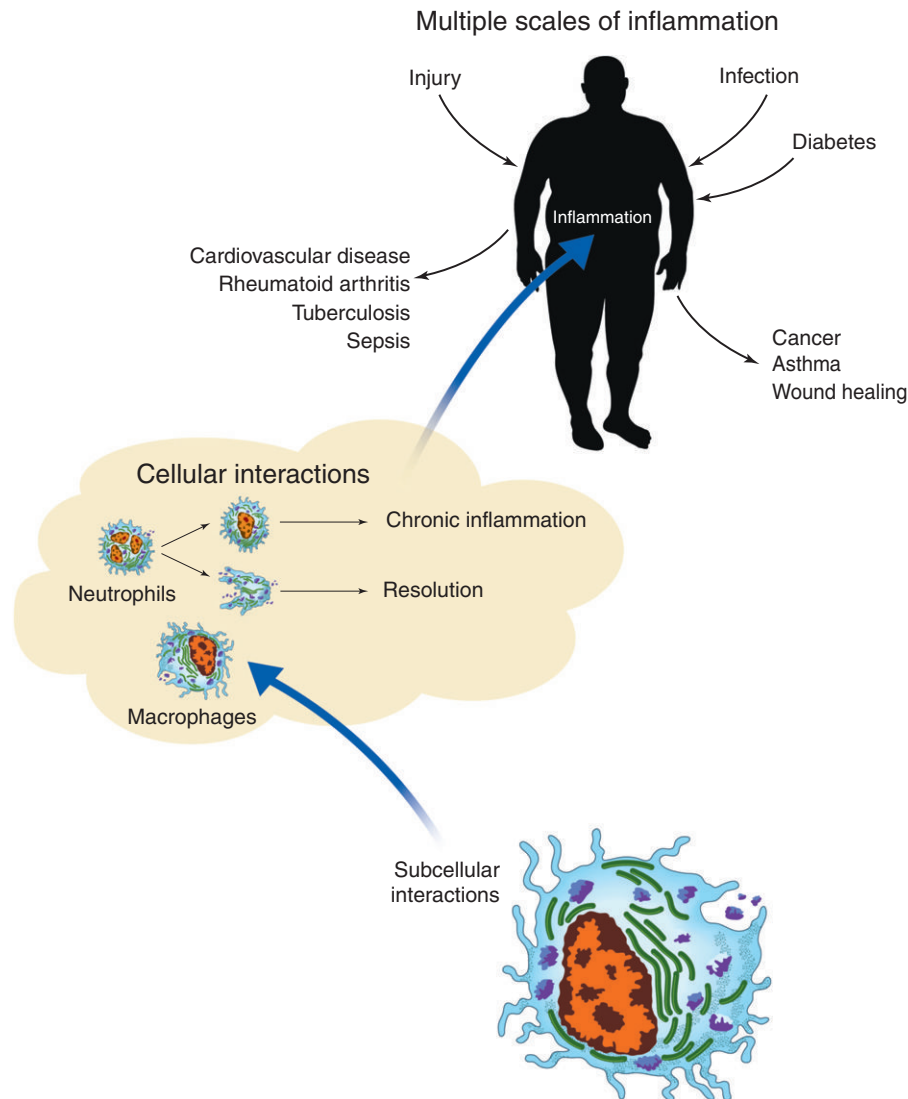


Figure 1 The inflammatory system responds to many different stimuli, and varies in its response according to tissue type. It plays a role in perpetuating, as well as being stimulated by, many different diseases. Its response is dependent on interactions that occur over many time and space scales from subcellular and cellular to tissue and organ specific responses. These complications result in inflammatory responses being often modeled generically (capturing key interactions common across tissue type and disease scenario) as well as being incorporated in models that are tailored to specific tissue types or disease scenarios.

interactions responsible for the successful resolution of inflammation. We then review some current mathematical and computational models that attempt to capture the inflammatory process, both in its generic form and as a key feature of, or as a stimulus to many of the disease scenarios in which inflammation plays a role. In doing this, we will demonstrate how mathematical modeling and computational techniques provide useful insights into this complex process that is central to so many diseases.

A Simple Mathematical Model Capturing the Resolution of Inflammation

Inflammation is a homeostatic process that is initiated in response to tissue injury, damage, or microbial attack. The cellular

and molecular events that ensue act cooperatively to eradicate pathogenic microorganisms, repair tissue damage, and restore normal tissue architecture. In health, inflammation resolves quickly, restoring normal tissue and organ functions. However, it has long been known that inflammation can be dysregulated, leading to a self-perpetuating chronic condition that plays a significant role in a broad spectrum of disease conditions, including those that were originally thought not to have an inflammatory component (Libby, 2002; Rogers *et al.*, 2007; Trinchieri, 2012). Although the historical view was that inflammation gradually 'subsided,' it is now evident that successful resolution of the inflammatory response is an active program of events that drive a switch from pro-inflammatory signals toward those of an anti-inflammatory nature (Serhan and Savill, 2005). The importance of an active inflammatory resolution process is reflected by the drive to identify new therapeutic agents that

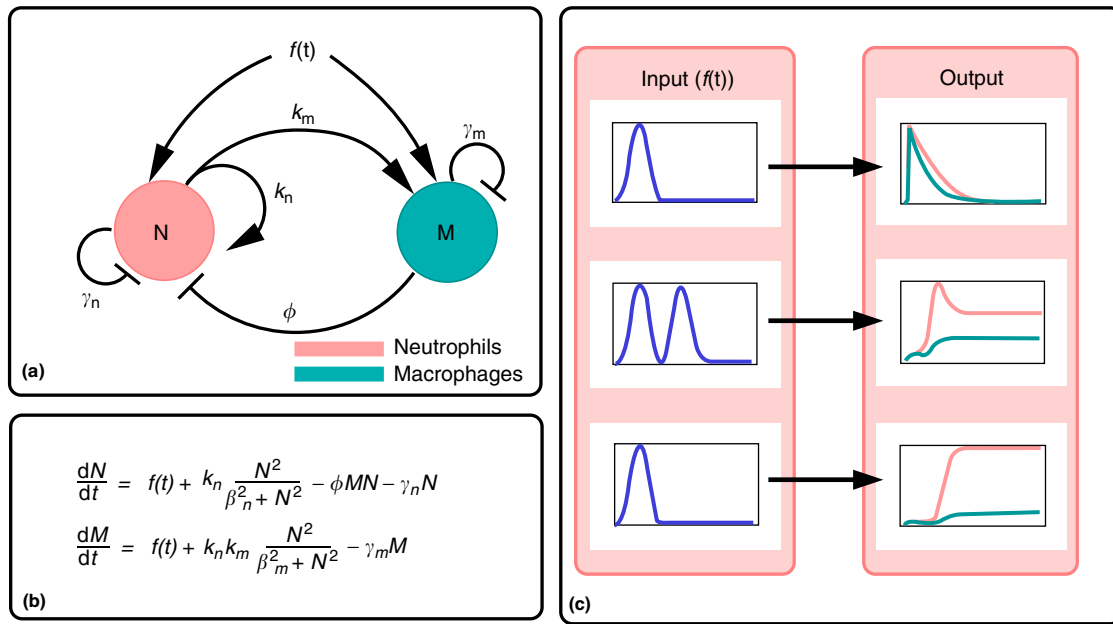


Figure 2 A simple mathematical model capturing the basic interactions central to the resolution of inflammation. (a) A graphical representation of the mathematical model. Neutrophils (N) and macrophages (M) arrive at a site of injury in response to an inflammatory stimulus ($f(t)$). Neutrophils contain toxic chemicals that they release to kill pathogens; these can cause damage to healthy tissue, perpetuating the inflammatory response while macrophages phagocytose neutrophils. (b) The ordinary differential equations that capture the interactions just described. (c) Model simulations demonstrating the time-dependent cell concentrations in response to different amounts of tissue damage. If the injury is of a short duration (top row) all cells leave the tissue (a healthy response); otherwise inflammation can progress to a self-perpetuating chronic condition (middle row). Under a number of clinical conditions, the inflammatory response is impaired. Simulations (bottom row) demonstrate how reducing the rate of macrophage phagocytosis (one order of magnitude) can lead to the inflammatory system failing to recover from an injury of a short duration.

target resolution mechanisms that could be used in the treatment of chronic conditions. One pivotal factor in this switch is thought to be dependent on the interaction between the two major cell types involved in inflammation: neutrophils and macrophages (Silva, 2009).

As an illustration of how mathematical approaches can be utilized to effectively model inflammation and, in particular, can capture active pro-resolution pathways that may determine whether the inflammatory response ultimately returns to a healthy state or progresses to a chronic self-perpetuating injurious condition, we present a toy model based on earlier work (Dunster *et al.*, 2014) in the form of ordinary differential equations. This dynamic model captures some of the key interactions between neutrophils and macrophages (see Figure 2) that occur in response to a stimulus such as tissue damage. In brief, in response to an injury/infection stimulus, there is recruitment of both neutrophils and macrophages, resulting in increased numbers of cells present. Neutrophils destroy invading microorganisms by release of toxic chemicals. However, these can also cause damage to healthy tissue which would act to perpetuate the inflammatory response. Countering these pro-inflammatory events is the ‘silent’ removal of neutrophils by macrophages, a process involving neutrophil cell death and subsequent phagocytosis by a mechanism that contributes to the resolution of inflammation. Successful resolution of inflammation results in the clearance of all macrophages and neutrophils from the tissue (a healthy response). In contrast, development of a self-perpetuating

chronic condition is characterized by the persistence of recruited inflammatory neutrophils and macrophages (Figure 2).

The model presented represents an example of a bistable system with two distinct outcomes in response to an injury stimulus: resolution of inflammation (all cells return back to a healthy level) or chronic self-perpetuating inflammation (cells attain a metastable higher positive value). In this model, a low level of damage results in the former but, if a threshold of damage is passed, then the ultimate outcome is the latter. In Figure 2(c) (top row), we present simulations of the model showing a typical response to one cycle of damage. The damage stimulus causes the infiltration of both neutrophils and macrophages. Macrophages have the capacity to remove neutrophils and cell numbers ultimately return to a healthy level. In contrast, with the same parameter values, repeated cycles of damage (Figure 2(c), middle row) results in failure of macrophages to remove all neutrophils as inflammation progresses and as a consequence, both cell types settle to a positive steady state that we interpret as chronic inflammation. By varying the model parameters we can simulate a scenario in which the inflammatory system is impaired, such as would occur under disease conditions. In Figure 2(c) (bottom row), we present simulations in which the rate at which macrophages are able to remove neutrophils is reduced (all other parameters stay the same as previous simulations). In contrast to simulations in which inflammation resolves (top row), a single cycle of damage now results in accumulation of

both macrophages and neutrophils that we equate to failure to resolve inflammation and establishment of chronic inflammation. Computational models such as this could be utilized to not only simulate outcomes of manipulation of key parameters that could be altered in disease conditions but also in response to therapeutic agents. Such an approach would provide important insights into the potential impact of changes in rates of cell recruitment and clearance on the development and termination of an inflammatory response and importantly, the magnitude of change that would be required to afford therapeutic gain.

Current Mathematical and Computational Models of Inflammation

Existing mathematical and computational models of inflammation reflect the wide range of disease states in which inflammation represents a key component. These models include those that make the assumption that there is a unifying inflammatory response which can be applied to all inflammatory stimuli. In contrast, models of inflammation have been developed that are specific for certain stimuli (such as pathogens or mechanical damage), or that are tailored to specific tissue types. In addition, many models of specific diseases, such as tuberculosis or cancer, have inflammatory interactions embedded within them. These models cover a wide spectrum of spatiotemporal scales, from the rapid intracellular events that are involved in the cellular response to exogenous inflammatory signals through to population-wide responses to infectious agents. Rather than provide a comprehensive review, we present an illustrative review of the breadth and depth of models that incorporate inflammatory mechanisms.

Early Events, Leukocyte Responses to Bacteria

The arrival of leukocytes to a site of tissue damage or infection is central to the inflammatory response. Leukocytes express adhesion molecules on their cell surface that allow them to bind to damaged or activated endothelial cells that line the blood vessels and subsequently move across the vessel to enter the tissues. This is a highly regulated process that involves cooperative action of adhesion molecules and chemoattractant cytokines that are released in response to injury or infection (Ley *et al.*, 2007). Lauffenburger and Kennedy developed a series of models that capture a generic leukocyte's response to bacterial infection (Lauffenburger and Keller, 1979; Lauffenburger and Kennedy, 1981; Lauffenburger and Kennedy, 1983). These simple models of two or three variables, representing bacteria, leukocytes, chemoattractants, and venules at various tissue scales, have been used to investigate how variation in key parameters, such as leukocyte capacity for motility or phagocytosis of bacteria, affect the dynamics of the inflammatory response and the outcomes of an infection. They found that localized areas of high leukocyte concentration, reminiscent of granuloma formation, occur more readily when leukocyte chemotaxis is defective, where there is a low growth rate of bacteria, or where bacteria are resistant to destruction by leukocytes.

With the aim of providing insight into sepsis, a sustained inflammatory condition can affect multiple organs and ultimately lead to death. Vodovotz and coworkers have also developed models that capture the acute inflammatory response to bacteria. Their initial model (Kumar *et al.*, 2004) is homogeneous (ignoring spatial variations), incorporating bacteria and two pro-inflammatory mediators. Model outcomes are equated to the clearance of the pathogen or progression to a sustained infection. Further work (Reynolds *et al.*, 2006) incorporates a generic leukocyte, bacteria, tissue damage, and an anti-inflammatory mediator to investigate how the strength of the anti-inflammatory mediator can affect the progression to a sustained infection. They have modified and extended this work to various experimental and disease scenarios, including the inflammatory response to endotoxin administration (a component of the bacterial cell wall) in mice (Day *et al.*, 2006) and the simulation of the inflammatory response to an anthrax infection (Kumar *et al.*, 2008), where they inferred that the administration of antibiotics is useful only if administered early. Reviews of their work are given in Vodovotz (2006) and Vodovotz *et al.* (2008).

Smith *et al.* (2011) combine bacterial measurements from mice infected with *Streptococcus pneumoniae* with mathematical modeling to investigate the inflammatory response to bacterial infection in the lung. Their series of quantified spatially averaged models, of increasing biological complexity, were used to investigate which components of the inflammatory response (resident macrophages, neutrophils, and monocyte-derived macrophages) contribute most to the ability of the lung to recover fully from the disease, finding that alveolar macrophages are critical for a successful response.

Inflammatory Cytokines, TNF-Therapy

Tissue injury or infection is characterized by production and/or release of molecules that act as triggers of inflammation. These molecules have been termed damage-associated molecular patterns (DAMPs) that are released from disrupted cells (e.g., extracellular nucleotides and high mobility group box 1 protein) or pathogen-associated molecular patterns (PAMPs) associated with microbial pathogens (e.g., bacterial endotoxin, lipotechoic acid, and flagellin) (Tang *et al.*, 2012). Macrophages and other cell types are able to recognize these DAMPs and PAMPs and produce a variety of pro-inflammatory cytokines that act to control the initiation and development of the inflammatory response (Dunne and O'Neill, 2003). Some of the key cytokines produced include those that characterize the acute phase response, including interleukin (IL)-1, IL-6, and tumor necrosis factor α (TNF- α). These cytokines act to promote activation of endothelial cells and the recruitment of leukocytes together with activation of cellular defense mechanisms (coagulation regulation, complement activation, promotion of phagocytosis, and production of destructive enzymes) which limit microbial growth and facilitate microbial destruction. However, prolonged elevation of levels of these cytokines can be detrimental and is associated with cancer cachexia or development and progression of inflammatory disease (Mantovani *et al.*, 2008; Nathan, 2002; O'Shea *et al.*, 2002; Stoll *et al.*, 2002).

Seymour and Henderson (2001) have developed a mathematical model to explore the complex relationship between a population of monocytes and the key inflammatory cytokines IL-1, IL-10, and TNF- α . This work was extended (Jit *et al.*, 2005) to explore the effects of therapeutically blocking TNF- α . Anti-TNF- α therapy has shown significant benefits to patients with RA but not to those with a systemic inflammatory response (SIRS), even though it involves a similar cytokine network to RA and has TNF- α as a key factor in its pathology. Utilizing the known pharmacokinetics properties of three soluble TNF- α inhibitors (TNFR2, Etanercept, and Infliximab), Jit *et al.* demonstrated the effectiveness of the inhibitors in controlling concentrations of free TNF- α under the conditions of RA. They proposed that, in contrast, under the condition of SIRS, therapeutically sequestered TNF- α can act as a slow reservoir sabotaging the therapeutic strategy. Cytokine networks involved in RA have also been modeled by Baker *et al.* (2013) who considered the interactions between generic pro- and anti-inflammatory cytokines; their simple model of two variables displays a range of outcomes, including oscillatory behavior that is reminiscent of a remitting-relapsing disease state. They demonstrate anti-inflammatory responses to anti-cytokine therapy, finding that dose-regime (as well as dose-level) is important for a successful outcome.

The response of a cell to inflammatory mediators is dependent on intracellular protein interactions. Mathematical modeling of inflammation at the subcellular scale has centered around the transcription factor NF- κ B, a key regulator of cellular responses to pro-inflammatory mediators including IL-1 and TNF- α . Prolonged activation of NF- κ B activity has been implicated in the development of chronic inflammatory diseases as well as cancer (Philip *et al.*, 2004). A review of mathematical models focusing on NF- κ B signaling is given by Hoffman and coworkers who also produce models in this area (Cheong *et al.*, 2008). Cancer and tumors have been extensively modeled (Byrne (2010) for a review of these). C-Abl is a tyrosine kinase central to subcellular interactions that affects a neutrophil's ability to undergo apoptosis. Jackson and Radivoyevitch (2013) model c-Abl mediated neutrophil responses in response to pathogens and infections, they simulate various treatment strategies predicting that the efficacy of kinase inhibitors Seliciclib and imatinib are additive.

Tissue Damage and Wound Healing

The inflammatory events that occur in response to tissue damage generally resolve, stimulating the successful events downstream where tissue is repaired and remodeled. One readily accessible system to follow inflammation and tissue repair is in the skin, where the response to trauma (wound healing) is easily observable. In a wound, inflammatory events are followed closely by, and interact with, downstream events such as wound closure, remodeling of the extracellular matrix, and angiogenesis that are necessary for the successful repair of tissue (Baum and Arpey, 2005; Singer and Clark, 1999). Theoretical models have focused on these downstream events. Sherratt and Murray (1990, 1991) model epidermal cell migration in acute wounds that heal successfully and a review of both of their, and others, work in this area is provided by Sherratt and Dallon (2002).

While minor wounds generally repair quickly this process can fail, leading to development of a chronic open wound, for example, pressure sores and leg ulcers and patients with diabetes are particularly susceptible to wounds of this kind. The failure of wounds to heal quickly has been suggested to depend on the presence of inflammatory cells. Waugh and Sherratt (2006) present a model that centers on the presence of inflammatory cells in normal and diabetic wounds. Their model captures two distinct populations of macrophages (with either an 'inflammatory' or a 'repair' phenotype) and the mediator TGF- β . Macrophages are capable of phagocytosis of cellular and particulate material and can produce a wide range of pro- and anti-inflammatory cytokines. Heterogeneity within the macrophage populations present at inflamed sites, together with distinct differentiation and activation responses to the local tissue micro-environment (Davies *et al.*, 2013; Murray and Wynn, 2011), represents a pivotal factor in the development and ultimate resolution of inflammation. Waugh and Sherratt find that the balance between the populations of macrophages of an inflammatory and repair phenotype vary under diabetic (compared to normal) conditions, affecting the subsequent wound healing process. This work was subsequently extended (Waugh and Sherratt, 2007) so that investigations could be carried out on the effect of applying two commercially engineered skin substitutes to diabetic wounds. The model showed close agreement to data from clinical trials and elucidates some of the underlying mechanisms involved in treatment.

Marée *et al.* (2005) investigate which aspects of macrophage phagocytosis are altered under diabetic conditions, formulating a model that captures the ability of macrophages to activate, engulf, and digest a series of dead cells. The capacity of macrophages to clear cells that have undergone programmed cell death (apoptosis) through a process of engulfment and digestion (phagocytosis) is central to inflammation and one consequence of the failure to remove apoptotic cells may be perpetuation of the inflammatory process and the subsequent failure of downstream events. Marée *et al.* compared their model to *in-vitro* time-course experimental data on macrophages obtained from normal and diabetic-prone mice. They found that macrophages from diabetic-prone mice are slower in their rate of activation (the time taken to remove the first apoptotic cell) and their ability to engulf apoptotic cells, but not slower to digest them. Further work (Marée *et al.*, 2006) investigates the effect of slower rates of macrophage activation and engulfment on secondary necrosis of pancreatic B cells showing that the difference in kinetics between normal and non-obese diabetic-prone mice is large enough to trigger an escalating inflammatory response in the later population of mice. Childs *et al.* (2011) also developed a model that focused on the role of macrophages in the context of wound healing, producing a model that investigates macrophage phenotype. Their work demonstrated how local microenvironmental conditions can lead to macrophages switching from an inflammatory phenotype to one that promotes wound repair.

Inflammation and Progression to the Immune Response

The inflammatory (innate immune) and adaptive immune response are closely entwined with macrophages playing a key

role in linking the early inflammatory response to the later adaptive immune response.

Rudnev and Romanyukha (1995) developed a model that focuses on the early inflammatory response to an infection by *Staphylococcus aureus* pneumonia in the lung. The model captures the ability of the pathogen to cause damage to the lung as well as the recruitment of neutrophils and macrophages and their capacity for removal of the bacteria. This work was later modified (Romanyukha *et al.*, 2006) to incorporate the cells of the adaptive immune response (antigen presenting, T, and B cells). The model investigates the long-term immune defense against an infection, estimating the energy burden required to fight infection during periods of health and disease.

An infection by the bacterium *Mycobacterium tuberculosis* leads to not only an innate but also an adaptive immune response that results in disease clearance, latency, or active tuberculosis. Inflammatory cells, particularly macrophages, are central to this disease outcome, collecting at the site of infection to form a granuloma that ‘walls off’ the bacterial infection (Russell, 2011). With the ultimate aim of understanding the processes that lead to granuloma formation and the aspects of an infection that can lead to latency, Kirschner and coworkers developed a number of models of the inflammatory and adaptive immune systems response to *M. tuberculosis* utilizing a range of mathematical techniques applied at different biological scales. An initial spatially averaged model (Wigginton and Kirschner, 2001) characterized the cytokine and cellular (both macrophage and T cell) interactions central to an infection. Bacteria populations are tracked outside (extracellular) and inside infected macrophages. Simulations demonstrate disease outcomes of latency (a lack of extracellular bacteria), active tuberculosis, and reinfection with results compared to experimental data. They found that even if latency is achieved there may be tissue damage. This work was extended (Marino and Kirschner, 2004; Marino *et al.*, 2004) to include trafficking of cells between the lung and the lymph nodes finding that a delay in either dendritic cell migration to the lymph node or T cell trafficking to the site of infection can alter an outcome of latency to one of active tuberculosis. Further models (Gammack *et al.*, 2005; Segovia-Juarez *et al.*, 2004) incorporate spatial variations. Gammack *et al.* (2005) concentrated on the early inflammatory response, before any adaptive immune response develops. They utilized reaction-diffusion-type equations to investigate how the expansion and contraction of a granuloma, in response to an infection, depends on interactions between macrophages, bacteria (both external and internal to macrophages), and a chemoattractant with cells trafficking across the granuloma boundary and moving within it in response to the chemoattractant. It was found that a model with just an inflammatory cell component was able to simulate uncontrolled granuloma growth or clearance. However, establishment of a stable granuloma (equated to a latent form of tuberculosis) was found to require an adaptive immune response component. A further spatial model (Segovia-Juarez *et al.*, 2004), that incorporated both adaptive immune cells (T cells) and innate immune cells (macrophages), utilizes a discrete cellular automata approach that combines a continuous representation of generic chemokines with discrete agent representations of macrophages and T cells. This novel approach

predicts that key elements that control the growth of a granuloma are the arrival time of T cells, chemokine diffusion rates, and macrophage overcrowding.

Conclusions and Future Challenges

Mathematical and computational models have a long history in being applied to biological processes. They can provide a bridge between biological knowledge, hypothesis, and experimental data allowing the extraction of more information from experimental data than is possible by intuition alone. Inflammation, with its dynamic complex nature, is particularly challenging to study both experimentally as well as mathematically. In experimental models of inflammation, capturing the complex cellular dynamics is often limited by the methods of sampling/analysis. For example, in a mouse model of infection in the lung, it is often necessary to sacrifice animals at a small number of distinct time points to examine cellular recruitment and measure levels of inflammatory mediators. This approach can often yield limited information with which to estimate changes in cell composition during an inflammatory response. Although this difficulty has constrained validation of mathematical modeling of inflammation, the predictive value of such models affords an opportunity not only to design new experimental protocols that would specifically enable modeling predictions to be tested but direct experiments to those targets that are the most therapeutically interesting. Importantly, inflammation in response to diverse stimuli and under many different disease conditions has many universal properties and such commonalities can be captured in generic models that aid in the understanding of the key processes that can bring about a change in outcome.

Inflammation is dependent on mechanisms that occur over many time and space scales and the mathematical and computational models that we have described utilize a range of mathematical techniques applied across a wide range of these scales. They attempt to capture the inflammatory response under many different disease scenarios and cross-comparison between models could highlight some of the key unifying mechanisms behind the inflammatory response. Though one of the key challenges with mathematical modeling is to enable biological knowledge and experimental data to be integrated across disparate scales, such integration can help us elucidate what happens when the components of inflammation are perturbed, producing testable predictions and providing insights into potential therapeutic targets.

See also: Cell Communication: Prototypic Integrative Processes: Wound Healing: An Orchestrated Process of Cell Cycle, Adhesion, and Signaling

References

- Baker, M., Denman-Johnson, S., Brook, B.S., Gaywood, I., Owen, M.R., 2013. Mathematical modelling of cytokine-mediated inflammation in rheumatoid arthritis. *Mathematical Medicine and Biology* 30, 311–337.

- Baum, C.L., Arpey, C.J., 2005. Normal cutaneous wound healing: Clinical correlation with cellular and molecular events. *Dermatologic Surgery: Official Publication for American Society for Dermatologic Surgery* 31, 674–686.
- Byrne, H.M., 2010. Dissecting cancer through mathematics: From the cell to the animal model. *Nature Reviews Cancer* 10, 221–230.
- Cheong, R., Hoffmann, A., Levchenko, A., 2008. Understanding NF- κ B signaling via mathematical modeling. *Molecular Systems Biology* 4, 192.
- Childs, L.M., Paskow, M., Morris Jr, S.M., Hesse, M., Strogatz, S., 2011. From inflammation to wound healing: Using a simple model to understand the functional versatility of murine macrophages. *Mathematical Biology* 73, 2575–2604.
- Davies, L.C., Jenkins, S.J., Allen, J.E., Taylor, P.R., 2013. Tissue-resident macrophages. *Nature Immunology* 14, 986–995.
- Day, J., Rubin, J., Vodovotz, Y., et al., 2006. A reduced mathematical model of the acute inflammatory response II. Capturing scenarios of repeated endotoxin administration. *Journal of Theoretical Biology* 242, 237–256.
- Dunne, A., O'Neill, L.A.J., 2003. The interleukin-1 receptor/toll-like receptor superfamily: Signal transduction during inflammation and host defense. *Science's STKE: Signal Transduction Knowledge Environment* 2003 (171), re3.
- Dunster, J.L., Byrne, H.M., King, J.R., 2014. The resolution of inflammation: A mathematical model of neutrophil and macrophage interactions. *Bulletin of Mathematical Biology* 76, 1953–1980.
- Fox, S., Leitch, A.E., Duffin, R., Haslett, C., Rossi, A.G., 2010. Neutrophil apoptosis: Relevance to the innate immune response and inflammatory disease. *Journal of Innate Immunity* 2, 216–227.
- Gammack, D., Ganguli, S., Marino, S., Segovia-Juarez, J., Kirschner, D.E., 2005. Understanding the immune response in tuberculosis using different mathematical models and biological scales. *Multiscale Modeling & Simulation* 3, 312–345.
- Jackson, R.C., Radiyevyitch, T., 2013. Modelling c-Abi signalling in activated neutrophils: The anti-inflammatory effect of seliciclib. *BioDiscovery* 7 (4), 4.
- Jit, M., Henderson, B., Stevens, M., Seymour, R.M., 2005. TNF- α neutralization in cytokine-driven diseases: A mathematical model to account for therapeutic success in rheumatoid arthritis but therapeutic failure in systemic inflammatory response syndrome. *Rheumatology* 44, 323–331.
- Kohl, P., Crampin, E.J., Quinn, T.A., Noble, D., 2010. Systems biology: An approach. *Clinical Pharmacology & Therapeutics* 88 (1), 25–33.
- Kolaczowska, E., Kubes, P., 2013. Neutrophil recruitment and function in health and inflammation. *Nature Reviews Immunology* 13, 159–175.
- Kumar, R., Chow, C.C., Bartels, J.D., Clermont, G., Vodovotz, Y., 2008. A mathematical simulation of the inflammatory response to anthrax infection. *Shock (Augusta, GA)* 29, 104.
- Kumar, R., Clermont, G., Vodovotz, Y., Chow, C., 2004. The dynamics of acute inflammation. *Journal of Theoretical Biology* 230, 145–155.
- Kvietys, P.R., Granger, D.N., 2012. Role of reactive oxygen and nitrogen species in the vascular responses to inflammation. *Free Radical Biology and Medicine* 52, 556–592.
- Lauffenburger, D., Keller, K., 1979. Effects of leukocyte random motility and chemotaxis in tissue inflammatory response. *Journal of Theoretical Biology* 81, 475–503.
- Lauffenburger, D., Kennedy, C., 1981. Analysis of a lumped model for tissue inflammation dynamics. *Journal of Mathematical Biosciences* 53, 189–221.
- Lauffenburger, D., Kennedy, C., 1983. Localized bacterial infection in a distribute model for tissue inflammation. *Journal of Mathematical Biology* 16, 141–163.
- Ley, K., Laudanna, C., Cybulsky, M.I., Nourshargh, S., 2007. Getting to the site of inflammation: The leukocyte adhesion cascade updated. *Nature Reviews Immunology* 7, 678–689.
- Libby, P., 2002. Inflammation in atherosclerosis. *Nature* 420, 868–874.
- Mantovani, A., Allavena, P., Sica, A., Balkwill, F., 2008. Cancer-related inflammation. *Nature* 454, 436–444.
- Marée, A.F., Kublik, R., Finegood, D.T., Edelstein-Keshet, L., 2006. Modelling the onset of Type 1 diabetes: Can impaired macrophage phagocytosis make the difference between health and disease? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 364, 1267–1282.
- Marée, A.F.M., Komba, M., Dyck, C., et al., 2005. Quantifying macrophage defects in type 1 diabetes. *Journal of Theoretical Biology* 233, 533–551.
- Marino, S., Kirschner, D.E., 2004. The human immune response to *Mycobacterium tuberculosis* in lung and lymph node. *Journal of Theoretical Biology* 227, 463–486.
- Marino, S., Pawar, S., Fuller, C.L., et al., 2004. Dendritic cell trafficking and antigen presentation in the human immune response to *Mycobacterium tuberculosis*. *Journal of Immunology* 173 (1), 494–506.
- Medzhitov, R., 2010. Inflammation 2010: New adventures of an old flame. *Cell* 140, 771–776.
- Murray, P.J., Wynn, T.A., 2011. Protective and pathogenic functions of macrophage subsets. *Nature Reviews Immunology* 11, 723–737.
- Nathan, C., 2002. Points of control in inflammation. *Nature* 420, 846–852.
- O'Shea, J.J., Ma, A., Lipsky, P., 2002. Cytokines and autoimmunity. *Nature Reviews Immunology* 2, 37–45.
- Philip, M., Rowley, D.A., Schreiber, H., 2004. Inflammation as a tumor promoter in cancer induction. *Seminars in Cancer Biology* 14 (6), 433–439.
- Reynolds, A., Rubin, J., Clermont, G., et al., 2006. A reduced mathematical model of the acute inflammatory response I. Derivation of model and analysis of anti-inflammation. *Journal of Theoretical Biology* 242, 220–236.
- Rogers, J., Mastroeni, D., Leonard, B., Joyce, J., Grover, A., 2007. Neuroinflammation in Alzheimer's disease and Parkinson's disease: Are microglia pathogenic in either disorder? *International Review of Neurobiology* 82, 235–246.
- Romanyukha, A.A., Rudnev, S.G., Sidorov, I.A., 2006. Energy cost of infection burden: An approach to understanding the dynamics of host–pathogen interactions. *Journal of Theoretical Biology* 241, 1–13.
- Rudnev, S.G., Romanyukha, A.A., 1995. Mathematical modeling of immune inflammatory reaction in acute pneumonia. *Journal of Biological Systems* 3, 429–439.
- Russell, D.G., 2011. *Mycobacterium tuberculosis* and the intimate discourse of a chronic infection. *Immunological Reviews* 240, 252–268.
- Savill, J., Dransfield, I., Gregory, C., Haslett, C., 2002. A blast from the past: Clearance of apoptotic cells regulates immune responses. *Nature Reviews Immunology* 2, 965–975.
- Segovia-Juarez, J.L., Ganguli, S., Kirschner, D., 2004. Identifying control mechanisms of granuloma formation during *M. tuberculosis* infection using an agent-based model. *Journal of Theoretical Biology* 231, 357–376.
- Serhan, C.N., Savill, J., 2005. Resolution of inflammation: The beginning programs the end. *Nature Immunology* 6, 1191–1197.
- Seymour, R.M., Henderson, B., 2001. Pro-inflammatory–anti-inflammatory cytokine dynamics mediated by cytokine-receptor dynamics in monocytes. *Mathematical Medicine and Biology* 18, 159–192.
- Sherratt, J.A., Dallon, J.C., 2002. Theoretical models of wound healing: Past successes and future challenges. *Comptes Rendus Biologies* 325, 557–564.
- Sherratt, J.A., Murray, J.D., 1990. Models of epidermal wound healing. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 241, 29–36.
- Sherratt, J.A., Murray, J.D., 1991. Mathematical analysis of a basic model for epidermal wound healing. *Journal of Mathematical Biology* 29, 389–404.
- Silva, M.T., 2009. When two is better than one: Macrophages and neutrophils work in concert in innate immunity as complementary and cooperative partners of a myeloid phagocyte system. *Journal of Leukocyte Biology* 87, 93–106.
- Singer, A.J., Clark, R.A., 1999. Cutaneous wound healing. *New England Journal of Medicine* 341, 738–746.
- Smith, A.M., McCullers, J.A., Adler, F.R., 2011. Mathematical model of a three-stage innate immune response to a pneumococcal lung infection. *Journal of Theoretical Biology* 276, 106–116.
- Stoll, G., Jander, S., Schroeter, M., 2002. Detrimental and beneficial effects of injury-induced inflammation and cytokine expression in the nervous system. *Advances in Experimental Medicine and Biology* 513, 87–113.
- Tang, D., Kang, R., Coyne, C.B., Zeh, H.J., Lotze, M.T., 2012. PAMPs and DAMPs: Signal O's that spur autophagy and immunity. *Immunological Reviews* 249, 158–175.
- Trinchieri, G., 2012. Cancer and inflammation: An old intuition with rapidly evolving new concepts. *Annual Review of Immunology* 30, 677–706.
- Vodovotz, Y., 2006. Deciphering the complexity of acute inflammation using mathematical models. *Immunology Research* 36, 237–245.
- Vodovotz, Y., Csete, M., Bartels, J., Chang, S., An, G., 2008. Translational systems biology of inflammation. *PLoS Computational Biology* 4 (4), e1000014.
- Waugh, H.V., Sherratt, J.A., 2006. Macrophage dynamics in diabetic wound healing. *Bulletin of Mathematical Biology* 68, 197–207.
- Waugh, H.V., Sherratt, J.A., 2007. Modeling the effects of treating diabetic wounds with engineered skin substitutes. *Wound Repair and Regeneration* 15 (4), 556–565.
- Wigginton, J.E., Kirschner, D., 2001. A model to predict cell-mediated immune regulatory mechanisms during human infection with *Mycobacterium tuberculosis*. *Journal of Immunology* 166, 1951–1967.

Angiogenesis

KA Thomas, Trefoil Therapeutics, Chatham, NJ, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Angiogenesis Growth of new microvessels from existing blood vessels.

Angiopoietin A set of proteins that modulate capillary quiescence, stability, and angiogenic activation.

Fibroblast growth factor (FGF) A member of a family of homologous soluble proteins that can induce the differentiation, mitosis, migration, and survival of a variety of responsive target cells and can trigger angiogenesis.

FGF receptor (FGFR) A cell surface transmembrane receptor tyrosine kinase that binds FGFs with high affinity initiating intracellular signal transduction.

Intracrine activity Function within a cell of a growth factor in the absence of secretion.

Intussusception Growth of new microvessels by longitudinal division of existing microvessels.

Juxtacrine activity Function of a membrane-anchored growth factor receptor on one cell by its membrane-anchored ligand on an adjacent cell.

Neuropilin (NRP) A pair of cell surface transmembrane vascular endothelial growth factor co-receptors that can form ternary complexes with vascular endothelial growth factors and their receptor tyrosine kinases.

Notch A member of a small set of juxtacrine morphogen receptors that when activated by binding membrane-anchored ligands on adjacent cells induces intracellular signal transduction leading to differentiated functions, which in vascular endothelial cells includes inhibition of the tip cell phenotype in angiogenic stalk cells.

Paracrine activity Function of a secreted growth factor on a nearby cell.

Pericyte A cell closely related to macrovascular smooth muscle cells (SMCs) that binds to the microvasculature including capillaries and contributes to their stabilization.

Platelet-derived growth factor (PDGF) A member of a set of homologous soluble dimeric proteins that can induce the

differentiation, mitosis, migration, and survival of multiple types of cells, notably including macrovascular SMCs and microvascular pericytes.

PDGF receptor (PDGFR) A cell surface transmembrane receptor tyrosine kinase that binds PDGFs with high affinity initiating intracellular signal transduction.

Sprouting angiogenesis Growth of new capillaries by budding and extension from existing microvessels.

Stalk cell A proliferative endothelial cell following the leading tip cell in a growing capillary sprout.

Tie A cell surface transmembrane receptor tyrosine kinase that binds angiopoietins with high affinity either initiating or inhibiting signal transduction.

Tip cell The leading migratory and invasive endothelial cell of a growing capillary sprout.

Vascular endothelial cell A type of cell usually forming a confluent non-thrombogenic monolayer lining the inside of the cardiovascular system and is the primary cellular component of capillaries.

Vascular endothelial growth factor (VEGF) A member of a set of homologous soluble dimeric angiogenic and lymphangiogenic proteins that can induce the differentiation, mitosis, migration, and survival of vascular and lymphatic endothelial cells.

Vasculogenesis Generation of the initial embryonic blood vessels by a process of differentiation and organization of vascular progenitor cells.

VEGF receptor (VEGFR) A cell surface transmembrane receptor tyrosine kinase found primarily on vascular and lymphatic endothelial cells that binds VEGFs with high affinity initiating intracellular signal transduction.

Wnt A member of a set of soluble paracrine morphogens that binds and activates its cell surface receptors to induce differentiation, migration, and mitotic signaling cascades in target cells leading to cell specification, such as the generation of mesodermal cells during gastrulation.

Vascular Evolution

The circulatory system has evolved in complex multicellular organisms to distribute oxygen, nutrients, and immune cells and to remove metabolic waste products including CO₂. Most invertebrates that are too large to depend on simple oxygen, nutrient, and waste diffusion, such as insects, have an open circulatory system in which blood flows freely within tissues. However a few large invertebrates such as cephalopods (e.g., squid, octopus, and nautilus) have a closed circulatory system in which peripheral blood is pumped through blood vessels lined by an acellular basement membrane. Although the growth of these vessels might utilize mechanisms that exhibit some similarities with vertebrate blood vessels, they

appear to be devoid of the luminal monolayer of vascular endothelial cells (ECs) that sit on a cross-linked glycoprotein basement membrane throughout the vertebrate circulatory system. These uniquely specialized cells apparently arose during the evolution of vertebrates and are central to the growth and function of their vascular systems (Munoz-Chapuli, 2011; Lammert and Axnick, 2012).

Developmental Vasculogenesis

The cardiovascular system is the first organ system that forms in vertebrate embryos and its proper development is critical for the growth of subsequently formed organs. Vertebrate

development initiates from the fertilized zygote, proceeding by rapid cell division to generate the spherical blastula, which then begins to remodel and differentiate during gastrulation forming the endodermal, mesodermal, and ectodermal germ layers. Formation of mesoderm is dependent on the combined influence of bone morphogenic protein-4 (BMP-4), which is a member of the large transforming growth factor- β family, fibroblast growth factor (FGF) -4 and -8, and wntless-int (Wnt). Intracellular signaling triggered by specific cell surface receptor binding of these and perhaps other soluble factors leads to the induction of the Brachyury and Tbx6 transcription factors required for mesoderm differentiation (Figure 1; Kimelman, 2006; Dorey and Amaya, 2010). FGF signaling through the FGF receptor-1 (FGFR1) supports development of the portion of the mesoderm destined to develop into the cardiovascular system (Yamaguchi *et al.*, 1994).

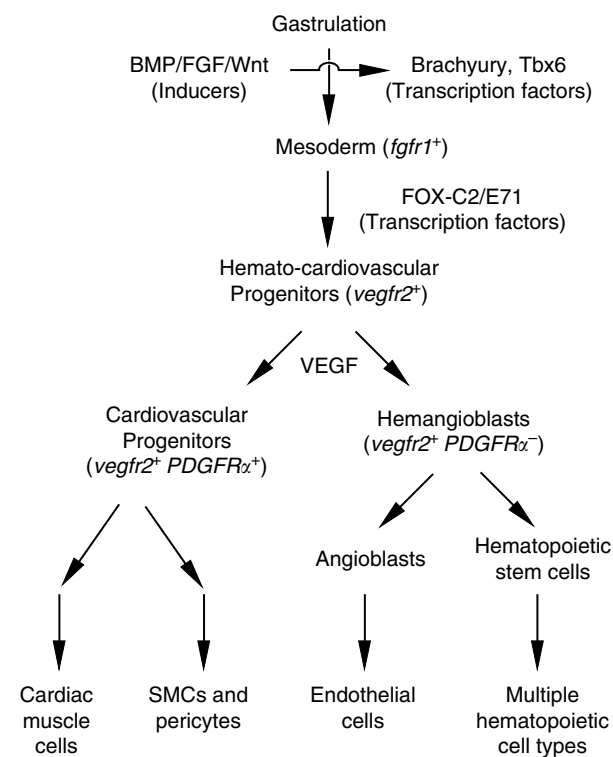


Figure 1 Developmental endothelial cell lineage. Early differentiation of the fertilized zygote generates a spherical blastula that remodels by gastrulation into the endodermal, mesodermal, and ectodermal germ layers. The secreted signaling proteins BMP, FGF, and Wnt induce the transcription factors Brachyury and Tbx6 that specify mesoderm, which is characterized by expression of the *fgfr1* gene and its activation by endogenous FGFs. Expression of the FOX-C2/E71 transcription factors differentiates a subset of mesodermal cells into hemato-cardiovascular progenitors characterized by expression of the vascular endothelial growth factor receptor-2 (*vegfr2*) gene. These cells can subsequently give rise to hemangioblasts, the progenitors of angioblasts that differentiate into vascular endothelial cells and hematopoietic stem cells that give rise to erythrocytes and multiple types of leukocytes. Alternatively, co-expression of the platelet-derived growth factor receptor- α (*PDGFR α*) gene generates cardiovascular progenitor cells that can differentiate either into cardiac muscle or into smooth muscle cells (SMCs) and pericytes. Therefore, cardiovascular development is dependent on FGFs, VEGFs, and PDGFs.

Although multiple transcription factors that might contribute to vascular differentiation are expressed under the control of BMP, FGF, and Wnt in vascular EC progenitors, a single transcription factor uniquely responsible for the developmental pathway has not been identified. Instead, the EC lineage appears to be critically dependent on cooperative binding and functional synergy between the transcription factors forkhead-C2 (FOX-C2) and ETS-related 71 (ER71, also known as ETV2) (Figure 1). These two proteins can form a heterodimeric complex on a highly conserved *cis*-acting FOX:ETS composite DNA-binding site, which is functionally restricted to the vascular EC lineage (De Val *et al.*, 2008; De Val and Black, 2009), and can transiently activate multiple EC-specific genes including *vegfr2* (often referenced as *flk1* in mice) encoding vascular endothelial growth factor receptor-2 (VEGFR2). The mesoderm-derived vascular EC lineage is dependent on the expression of VEGFR2, which is a cell surface transmembrane receptor tyrosine kinase that binds the extracellular soluble ligand vascular endothelial growth factor-A (VEGF-A). ER71/ETV2 appears in mesoderm by E7–7.5 followed by a rapid decline after E11.5 once ECs begin to mature. Mice lacking the *vegfr2* gene are defective in hematopoietic and vascular EC development and do not survive past embryonic stage E8.5–9.5 (Park *et al.*, 2013).

Subsequently, these cells can differentiate in a VEGF-dependent manner into (1) *vegfr2*⁺*PDGFR α* ⁺ (platelet-derived growth factor receptor- α) cardiovascular progenitor cells that differentiate into *vegfr2*⁺*PDGFR α* ⁺ cardiac and smooth muscle cells (SMCs), and (2) *vegfr2*⁺*PDGFR α* [−] hemangioblasts. The hemangioblasts give rise both (3) to hematopoietic stem cells, which are the precursors of multiple hematopoietic cell types including erythrocytes and leukocytes and (4) to the angioblast vascular EC progenitor that differentiates into mature ECs (Figure 1; Ishitobi *et al.*, 2011; Park *et al.*, 2013).

Vasculogenesis proceeds in both extraembryonic and embryonic tissues. In mouse embryos cells from both the lateral and posterior mesoderm migrate toward the extraembryonic yolk sac while aggregating into 'blood islands' at embryonic stage E6.5–7 containing cores of *vegfr2*[−] hematopoietic precursors that by E7.5 give rise to erythrocytes, surrounded by a shell of *vegfr2*⁺ angioblasts, which by E8.5 differentiate into ECs that assemble forming an extraembryonic capillary network. Simultaneously, within the embryo, angioblasts assemble into a capillary network and dorsal aorta in the absence of hematopoiesis. The extra- and intraembryonic vascular systems connect and upon initiation of a heartbeat and the flow of blood the capillary plexus remodels into an arterial, capillary, and venous circulatory system (Eichmann *et al.*, 2005). The function of the extraembryonic vascular system is to provide oxygen and nutrients and to remove metabolic by-products prior to the development of intraembryonic organs that eventually accomplish these critical functions. In addition, the parallel lymphatic system, necessary to drain blood vessel-derived extracellular fluid from tissues, develops from embryonic veins.

Following completion of embryonic vasculogenesis, growth of existing vessels can occur by the extensively studied mechanism of sprouting angiogenesis and by intussusception leading to longitudinal microvascular cleavage creating two smaller

daughter vessels. Although the molecular basis for intussusception is not clear, sprouting angiogenesis is induced by specific angiogenic protein growth factors.

Capillary Structure

Both sprouting angiogenesis and intussusception are mechanisms to generate new from existing capillaries. As illustrated in Figure 2, the structure of the existing capillaries is composed of a luminal monolayer of vascular ECs, which is continuous in most capillary beds. ECs sit on a thin (typically 50–100 nm) basement membrane consisting largely of collagen IV and to a lesser extent collagens XV and XVIII along with other proteins involved in cellular attachment but is devoid of lipid. The basement membrane is constructed by initial deposition of heterotrimeric laminin, a large helical coiled-coil cross-shaped protein that binds to specific cell surface components such as transmembrane integrins. Cell surface-anchored laminins assemble into a network that is indispensable for basement membrane formation. Nidogens (also called entactins) bind to laminin and to collagen IV, the principal collagen within basement membrane that then self assembles to form a fishnet arrangement that subsequently can be covalently cross-linked providing mechanical strength and stability to the capillary structure. Laminins also bind proteoglycans including perlecan that can themselves bind through their covalently linked negatively charged heparan sulfate oligosaccharide chains to a variety of proteins including several growth factors and their cell surface receptors (Iozzo, 2005; Eble and Niland, 2009). On the other side of the basement membrane and sometimes buried within it are pericytes, the microvascular counterpart of SMCs surrounding larger vessels, which form a non-confluent

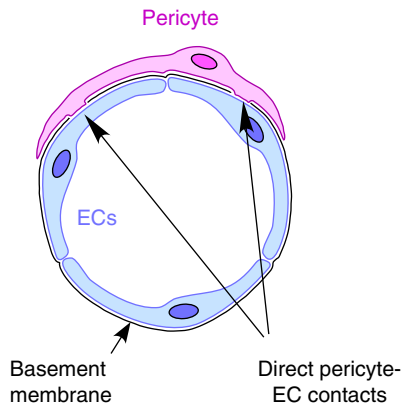


Figure 2 Capillary structure. In most organs capillaries are composed a confluent monolayer of vascular endothelial cells (ECs) attached to each other (blue) and to the underlying basement membrane (black line) made of extracellular proteins that they express and secrete. Attached to the opposite side of the basement membrane and sometimes embedded within it are pericytes (magenta), the microvascular counterpart of SMCs that surround larger vessels. Focal direct contacts exist between the ECs and protrusions from the pericytes that penetrate the basement membrane. Although the illustrated capillary cross-section contains three ECs, lumen of the smallest capillaries, which are just large enough to pass an erythrocyte, can be generated by a single EC.

and sometimes sparse layer of support cells. Both ECs and pericytes can contribute to the synthesis and assembly of their common basement membrane. However, pericyte protrusions through holes in the basement membrane can insert into membrane invaginations on the EC surface. Pericytes and SMCs, both of which express α -SMC actin, are often simply called mural cells (Armulik *et al.*, 2011).

Primary Factors Controlling Angiogenesis

Before describing the mechanisms known to control and coordinate the angiogenic response, some background molecular biology and biochemistry of the major angiogenic growth and remodeling factors that are recognized to drive this process, along with their activities in ECs and associated mural cells, will be summarized to provide perspective on their functions.

Angiogenic Growth Factors

Vascular growth and development has been recognized for over a century to occur not only in normal tissue growth and repair but also in pathologic conditions such as inflammation, certain retinal diseases, and solid tumor growth (Ferrara, 2002). Specific members of the FGF and VEGF families drive sprouting angiogenesis *in vivo* and *ex vivo*. They appear to do this primarily as a result of their ability to bind and activate their receptor tyrosine kinase transmembrane receptors resulting in the stimulation of EC migration and mitosis, as will be described later in greater detail.

FGFs

FGF ligands

Several well-characterized FGFs, including FGF-1, -2, and -4, potently stimulate vascular EC migration, mitosis, and survival culminating in angiogenesis. They are members of a large homologous protein growth factor family consisting of 22 known mammalian members. Although FGFs are numbered 1 through 23, humans and mice are missing FGF-15 and FGF-19, respectively, and human FGF-19 appears to be the species ortholog of mouse FGF-15. FGFs are small (~15–30 kDa) soluble proteins that typically function as monomers. Various FGF family members can exhibit intracrine activity within cells and, upon secretion, autocrine self-stimulation, paracrine activity on nearby cells, circulating endocrine hormonal functions or some combination of one or more of these mechanisms.

FGF receptors

With the exception of FGF-11 through FGF-14, which are structurally homologous but not functionally similar, FGFs bind to one or more of four homologous high-affinity FGFR gene products, designated FGFR1-FGFR4. From the N- to C-terminus these transmembrane receptors each contain three extracellular immunoglobulin-like (Ig-like) domains, a single transmembrane spanning helix, an intracellular juxtamembrane region, a double domain 'split' intracellular tyrosine kinase containing a short 'kinase insert domain' sequence within its C-terminal folding domain, and a C-terminal tail (Figure 3). The FGF-binding site spans the second and third

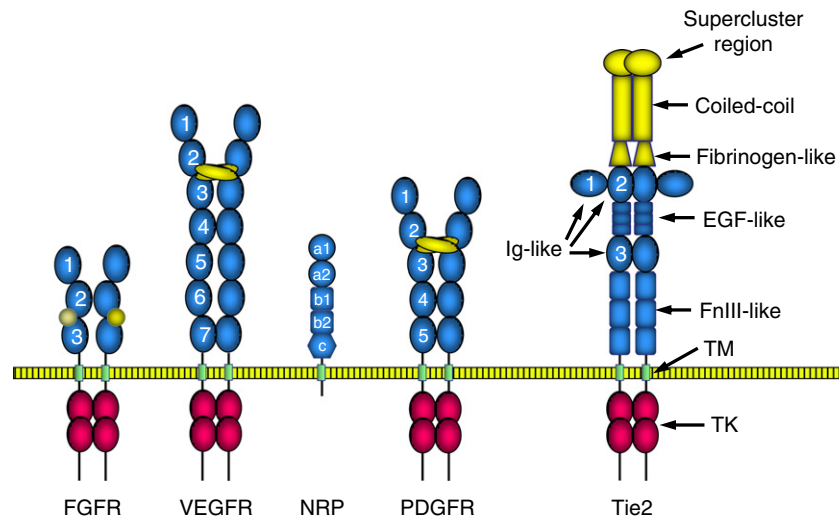


Figure 3 Major capillary growth factor ligand, receptor, and co-receptor domain structures. The FGFR, VEGFR, PDGFR, Tie2 receptors, and the neuropilin (NRP) co-receptor each contain multiple extracellular domains (blue). Different extracellular domain folds are each denoted by a unique shape. The ligands (yellow) are docked to their extracellular receptor-binding domains. VEGF isoforms containing a positively charged C-terminal region bind the NRP co-receptor at a site spanning the $\beta 1$ and $\beta 2$ domains forming a VEGFR-VEGF-NRP complex. Each receptor/co-receptor spans the cellular plasma membrane (narrow yellow bar) via a single helical transmembrane (TM) polypeptide sequence (green rectangle). The four receptors each contain an intracellular juxtamembrane region (black line), double domain tyrosine kinase (TK) (red) that contains a kinase insert domain (not shown) and C-terminal tail (black line). NRP contains a relatively short, probably unstructured, intracellular sequence that is devoid of enzymatic activity but can bind specific signaling molecules.

extracellular Ig-like domains. Alternative splicing of the third coding exon, corresponding to the C-terminal half of the membrane proximal Ig-like domain 3, generates the IIIb and IIIc isoforms of FGFR1 – FGFR3 (Holzmann *et al.*, 2012). Therefore, alternative splicing generates seven major FGFR isoforms, each with a unique FGF ligand-binding site and selectivity pattern (Ornitz *et al.*, 1996; Zhang *et al.*, 2006). This creates a distinct receptor isoform recognition pattern for each FGF that partially overlaps the receptor recognition pattern of other FGFs resulting in redundant stimulation among some of the FGF family members. In addition, the biological activities of individual FGFs are restricted by the spatial and temporal pattern of their expression and that of their receptors, processes largely controlled by transcription factors that bind to their specific promoter and enhancer sites. The angiogenic FGF-1, -2, and -4 can each activate all of the FGFR1-FGFR3 IIIc isoforms (Ornitz *et al.*, 1996), which are expressed by vascular ECs (Antoine *et al.*, 2005), and stimulate the migration, mitosis, and survival of not only these cells but also a wide spectrum of other types of cells.

Heparin sulfate proteoglycan binding

Multiple paracrine FGFs, including those exhibiting angiogenic activity, bind to negatively charged heparan sulfate proteoglycans (HSPGs) such as syndecans, glypican, and perlecan present on cell surfaces or extracellular matrix (ECM) including basement membranes, albeit with approximately two orders of magnitude lower affinity than to the FGFRs. Binding to the high capacity cell surface HSPGs can increase the concentration of FGF in the vicinity of the FGFRs thereby possibly enhancing the apparent ligand–receptor affinity. In addition, HSPGs and the more highly sulfated soluble counterpart heparin sulfate (HS) can bind both FGF and FGFR to form a

ternary cell surface complex. ECM HSPG can also serve as a reservoir for heparin-binding FGFs that are released by mast cell-derived HS, heparanases, and proteases.

FGFR signaling

Signal transduction is best characterized for FGFR1, which is the major EC FGFR and triggers many of the activities associated with vascular EC FGF responsiveness. FGF binding is thought to promote FGFR dimerization bringing the intracellular tyrosine kinase domains into close proximity and proper orientation where they can cross-phosphorylate each other. The initial dimeric receptor subunit transphosphorylation occurs on a single tyrosine in the kinase enzyme activation loop that shifts the loop conformation, increasing accessibility to the active site and enhancing phosphotyrosine catalytic activity by approximately two orders of magnitude. The activated kinase then phosphorylates tyrosine residues in the juxtamembrane region, kinase insert, and C-terminal tail that serve as binding sites for proteins able to initiate activation of several signaling cascades mediated by sequential phosphorylation of proteins. The final step is phosphorylation of a second activation loop tyrosine residue that increases catalytic activity by another order of magnitude (Furdui *et al.*, 2006). The major substrates of FGFR tyrosine phosphorylation include the adaptor protein FGF receptor substrate 2 (FRS2) and the enzyme phospholipase $C\gamma$ (PLC γ). FRS2 triggers a cascade of signaling proteins leading to the activation of the extracellular signal-regulated kinase (Erk), which is a type of mitogen-activated protein kinase (MAPK) that migrates into the nucleus where it phosphorylates transcription factors controlling expression of specific genes involved in proliferation. In addition, FRS2 triggers the phosphoinositol 3-kinase (PI3K)-Akt pathway mediating cellular survival. The specific signaling

responses attributable to PLC γ include activation of protein kinase C (PKC) and translocation of the transcription factor 'nuclear factor of activated T-cells' (NFAT) into the nucleus stimulating cellular migration (Goetz and Mohammadi, 2013). In addition, FGF-2 can trigger EC HSPG syndecan-4 clustering, which via its short intracellular domain activates a PKC α -RhoG-Rac1 pathway; the GTPase Rac1 then promotes actin polymerization mediating migration (Elfenbein *et al.*, 2009). As appears to be the case for several growth factors, FGF/FGFR signaling, originally believed to happen at the plasma membrane, is now thought to occur predominately on intracellular endosomal vesicles (Zhang and Simons, 2014).

VEGFs

VEGF ligands

The mammalian VEGF family consists of five homologous gene products: VEGF-A, VEGF-B, VEGF-C, VEGF-D, and placental growth factor (PlGF), each of which forms functional homodimers. The best characterized angiogenic family member is VEGF-A (sometimes denoted simply as VEGF), which also exists as an active heterodimer with PlGF. Deletion of a single VEGF-A coding allele (i.e., *vegfa*^{-/+}) in mice is embryonically lethal with a severely disrupted vasculature demonstrating not only the presence but also the gene dose is critical for proper vascular development. The primary VEGF-A mRNA transcript, composed of eight coding exons, is subject to alternative splicing of exons six and seven that encode positively charged amino acid sequences near the C-terminus, which generates 121, 145, 165, and 189 amino acid subunit isoforms (the corresponding murine VEGF-A isoforms are each one amino acid shorter). These range from no HS/HSPG binding for VEGF-A₁₂₁ to increasingly tighter binding with greater positive charge of the longer splice insert regions. Differential HSPG binding contributes to VEGF-A isoform diffusion gradients, which impact its spatial range of action. Release of HSPG-bound VEGF isoforms can be achieved by partial proteolysis with either plasmin or matrix metalloproteinase (MMP). Considering the proteolytic remodeling associated with angiogenesis, this might be a physiologically relevant mechanism of mobilizing matrix-bound VEGFs (Ferrara, 2010; Vempati *et al.*, 2014).

VEGF receptors

VEGFs function by binding one or more of the three VEGF receptors VEGFR1, VEGFR2 (also known as KDR), and VEGFR3. These receptors are present on vascular and lymphatic ECs and a restricted range of non-ECs. Mouse homozygous deletion of each of these three genes is embryonically lethal exhibiting disrupted embryonic vascular development.

Each receptor subunit contains seven extracellular Ig-like domains, a single transmembrane spanning sequence, and a cytoplasmic juxtamembrane region, a split tyrosine kinase containing a kinase insert domain within its C-terminal folding domain and a C-terminal tail (Figure 3). Similar to FGF/FGFR, VEGFs bind VEGFRs through receptor Ig-like domains two and three promoting receptor dimerization and cross-phosphorylation of the two adjacent intracellular tyrosine kinase domains. VEGF binding promotes VEGFR homodimerization and, in the case of VEGFR2, heterodimerization with VEGFR1 and VEGFR3. In addition, Ig-like domains 4, 6,

and 7 bind HS/HSPG forming a ternary complex with VEGF/VEGFR. Naturally occurring soluble splice variants of VEGFR1 (sVEGFR1, also known as sFlt1) and VEGFR2 (sVEGFR2 or sKDR), which are missing either one or two membrane proximal Ig-like domains, the transmembrane spanning sequence and entire cytoplasmic region, function to sequester their cognate VEGF ligands and might also form inactive heterodimers with membrane-spanning VEGFRs (Albuquerque, 2013; Singh *et al.*, 2013; Thomas, 2013).

Neuropilin co-receptors

Co-receptors are generally defined as cell surface molecules that influence ligand-receptor activity but do not contain intrinsic catalytic activity. In addition to HS/HSPG, VEGFs and their complex with VEGFR can also associate with the co-receptors neuropilin (NRP) -1 and -2, homologous single-pass transmembrane proteins each containing five extracellular domains (α 1- α 2- β 1- β 2-c/MEM), a single-pass transmembrane spanning, probably helical, polypeptide and short intracellular sequence (Figure 3). VEGFR binds to the VEGF core (encoded by *veg* exons 2–5) whereas the NRP β 1- β 2 domains bind the C-terminal cationic tail (encoded by alternatively spliced *veg* exons 7 and 8) of VEGF ligands. NRPs also directly bind VEGFR and stabilize the ternary complex, which can be further stabilized by HS/HSPG interaction. NRP-1, preferentially expressed in capillaries and arteries, can bind isoforms of VEGF-A, VEGF-B, and PlGF and function as a co-receptor for VEGFR1 and VEGFR2. Although NRP-1 can bind VEGF-A₁₂₁, only the VEGF-A isoforms containing one or both of the positively charged C-terminal regions generated by alternative splicing can simultaneously bind NRP-1 and VEGFR2. NRP-1 promotes VEGFR2 phosphorylation, signaling, internalization into vesicles that recycle back to the cellular surface avoiding receptor degradation, and promotes EC migration and survival. In contrast, NRP-2, which is expressed in vascular capillary ECs, veins and lymphatic ECs, binds VEGF-A, VEGF-C, VEGF-D, and PlGF isoforms and can function as a co-receptor for VEGFR1, VEGFR2, and VEGFR3 promoting EC migration and survival. NRPs were initially identified as receptors for the semaphorin class 3 (Sema3) family of seven soluble proteins (Sema A-G) involved in axon guidance. Sema3A and Sema3F, preferentially signaling through NRP-1 and NRP-2, respectively, exhibit anti-angiogenic activity inferring a NRP-mediated VEGF antagonistic function (Favier *et al.*, 2006; Parker *et al.*, 2012; Koch, 2012; Plein *et al.*, 2014; Djordjevic and Driscoll, 2013).

VEGFR signaling

Reminiscent of FGFR activation, the transphosphorylation of two tyrosine residues in the activation loop of VEGFR2, which restricts substrate access to the catalytic site, shifts its position opening access to the active site and increasing catalytic efficiency. This probably also occurs for VEGFR3 but a single amino acid difference in the VEGFR1 activation loop along with a repressor sequence in the juxtamembrane region substantially decreases its catalytic activation making this an inefficient tyrosine kinase, which might account for its minimal mitogenic activity despite binding VEGF-A with an order of magnitude higher affinity than VEGFR2. Therefore, at least in some situations not only sVEGFR1 but also the membrane-spanning VEGFR1 receptor tyrosine kinase can sequester

VEGFs effectively preventing the activation of the more efficient VEGFR2 and VEGFR3 signaling. Ligand binding induces additional phosphorylation at several other exposed tyrosine residues in each VEGFR cytoplasmic region creating functional binding sites for proteins required to trigger activation of signal transduction pathways. VEGFR2 activates pathways utilizing the PLC γ -PKC-Raf-MEK-Erk, TSAd-Src and PI3K-Akt signaling cascades mediating proliferation, migration, and survival, respectively. In addition, vascular permeability is increased by Src-triggered destabilization of EC-EC junctions. VEGFR2 heterodimerization with either VEGFR1 or VEGFR3 can alter not only ligand affinity but also some of the receptor tyrosine phosphorylation sites that in turn modify the efficiency of different signal transduction pathways and their resulting biologic functions (Koch *et al.*, 2011; Koch and Claesson-Welsh, 2012; Domigan *et al.*, 2015).

As was previously mentioned for FGFR, signaling induced by VEGFR2 occurs primarily on endosomes following endocytosis. These intracellular vesicles can then either target the receptor for degradation or recycle it back to the plasma membrane in an actively regulated manner. Compartmentalization into different endosomal populations might influence not only the duration of signaling but also the relative strength of different signaling pathways (Zhang and Simons, 2014). Signaling efficiency and duration can also be enhanced by association of the VEGF/VEGFR complex with not only the NRP co-receptors but also the integrin β 1 subunit, CD63 tetraspanin (i.e., a member of a specific family of proteins that span the plasma membrane four times) and syndecan-1 transmembrane proteins (Domigan *et al.*, 2015).

In addition to these VEGF-dependent canonical pathways, unanticipated VEGF-independent noncanonical VEGFR signaling and activities have been identified. The carbohydrate-binding protein galectin-1 can bridge the glycoproteins VEGFR and NRP-1 through their covalently attached oligosaccharides. In the case of VEGFR2 this induces migration whereas VEGFR1 promotes permeability in the absence of VEGF. Similarly, the simultaneous binding of galectin-3 to the oligosaccharide chains of multiple cell surface VEGFRs appears to cluster, activate, and prolong signaling by inhibiting their internalization and degradation in a VEGF-independent manner. The BMP antagonist gremlin also unexpectedly was found to bind VEGFR2, activating its autophosphorylation and binding to the α v β 3 EC integrin. How important these noncanonical pathways are in normal development, growth, and tissue repair has not yet been well characterized (Domigan *et al.*, 2015).

Vascular Remodeling Factors

New immature capillaries formed in response to angiogenic growth factor stimulation are fragile and unstable. To form persistent structures they must be stabilized by remodeling both their interaction with each other and with their environment. This is induced by autocrine and paracrine signals triggered by remodeling and stabilizing growth factors.

Angiopoietins

Like the VEGFs, the angiopoietin (Ang, also abbreviated Angpt) growth factor system largely targets the vasculature, the

developmentally and structurally parallel lymphatic system, and some hematopoietic cell lineages. However, rather than directly inducing angiogenesis, the Ang ligands modulate microvascular plasticity and stability thereby either facilitating or inhibiting capillary growth.

Ang ligands

Mammals express three unique Ang gene products, designated Ang1, Ang2, and Ang4 in humans (Ang3 is the mouse ortholog of Ang4, which are sometimes grouped as Ang3/4). The Ang ligands are each composed of an N-terminal superclustering and coiled-coil domains (Figure 3) that mediate subunit assembly into dimers, tetramers, hexamers, and higher order oligomers, and a C-terminal fibrinogen domain that interacts with its receptors mediating their clustering and induction of intracellular signaling.

Ang receptors

Two homologous Ang receptors, Tie1 and Tie2, each consist of a large extracellular region containing multiple Ig-like, epidermal growth factor (EGF)-like and fibronectin III-like folding domains, a single-pass transmembrane spanning helical peptide and an intracellular juxtamembrane region, split tyrosine kinase containing a short kinase insert domain within its C-terminal folding domain, followed by a C-terminal tail as shown schematically in Figure 3. They are primarily expressed on vascular and lymphatic ECs and deletion of either receptor is embryonically lethal, exhibiting loss of vessel integrity with fluid leakage.

A high-affinity ligand selective for Tie1 has not been identified so it is considered to be an 'orphan' receptor. In contrast, Tie2 (also known as TEK) binds the Ang1 and Ang4 agonists and Ang2, which is a weak agonist that can effectively function as an antagonist by competing with Ang1 and Ang4 binding. Tie1 and Tie2 typically heterodimerize with each other in vascular ECs. Ang1 binding can dissociate the receptor heterodimer allowing formation of Tie2 dimers, functional tetramers and larger aggregates. Ang2 binds at essentially the same receptor site used by Ang1 but neither efficiently dissociates nor activates the Tie1-Tie2 receptor heterodimer. However, in the absence of Tie1, Ang2 can bind and activate Tie2 thus functioning as an agonist (Barton *et al.*, 2014).

Tie signaling

Ang1 is constitutively produced by pericytes and SMCs. Functioning as a tetramer or larger oligomer, Ang1-induced formation of Tie2 homo-tetrameric or high-order clusters promotes receptor tyrosine phosphorylation. Although like other tyrosine kinases the purified recombinant Tie2 kinase contains a tyrosine in the activation loop, which in the crystal structure of the unphosphorylated state adopted the active conformation, the nucleotide binding loop that binds the ATP phosphates is present in a self-inhibitory conformation and the C-terminal tail appears to partially occlude the peptide substrate-binding site. In addition, phosphorylation of a specific tyrosine in the kinase N-terminal domain, which is known to bind a phosphatase, inhibits, or perhaps rapidly reverses, kinase activation. Therefore, whatever the mechanism of kinase activation, it likely differs from other characterized tyrosine kinases (Shewchuk *et al.*, 2000).

In the absence of EC–EC contacts, Ang1 induces Tie2 clustering at the EC–ECM interface, activating the binding of ‘Downstream of kinase-R’ (Dok-R) to one of the phosphotyrosines in the C-terminal region and activating cell adhesion and migration (Saharinen *et al.*, 2008). However, in the presence of EC–EC contacts the Ang1/Tie2 complexes (but not Tie1) migrate to the cell–cell interfaces with the activation of the PI3K/Akt signaling pathway promoting EC survival and inhibiting vascular permeability. Although Ang4 is a Tie2 agonist, its activity in microvascular function has not been well characterized. In contrast, Ang2 effectively functions as an antagonist of Ang1 activity, inhibiting Tie2 activation and signaling resulting in capillary destabilization (Huang *et al.*, 2010; Jeltsch *et al.*, 2013).

Platelet-derived growth factors

The platelet-derived growth factor (PDGF) family and their receptors (PDGFRs) are structurally similar to the VEGF/VEGFR family, indicative of a common ancestral gene. In newly formed capillaries, EC-derived PDGF is important for pericyte recruitment.

PDGF ligands

Four mammalian PDGF genes exist encoding PDGF-A, -B, -C, and -D, which each homodimerize and in the case of PDGF-A and -B also form a heterodimer with each other. PDGF-B is expressed by vascular ECs and PDGF-B gene knockout mice die soon after birth exhibiting extensive pericyte deficiency and resulting microvascular aneurysms. PDGF-A and -B polypeptide sequences N-terminal of the central receptor-binding region are cleaved before secretion whereas those of PDGF-C and -D are secreted as latent forms that require extracellular proteolytic processing to become activated. The propeptides are thought to remain non-covalently associated with the core PDGF regions that bind receptor. In fact the propeptides bind the same PDGF loops that bind the receptors and must be removed prior to receptor recognition. PDGF-A and -B but not -C or -D also contain positively charged C-terminal extensions that bind HS/HSPGs partitioning them to cell surface and basement membrane HSPGs. Deletion of just the C-terminal positively charged tail motif leads to diminished microvascular pericyte association with a phenotype intermediate between wild-type and PDGF-B genetic deletion (Chen *et al.*, 2013; Andrae *et al.*, 2008).

PDGF receptors

Two PDGFR genes encode transmembrane proteins PDGFR α and PDGFR β , each containing five extracellular Ig-like domains, a single membrane-spanning helix, and an intracellular juxtamembrane region, split intracellular tyrosine kinase containing a long kinase insert domain region and a C-terminal extension (Figure 3). PDGF dimers bind to the second and third Ig-like PDGFR domains promoting receptor dimerization in a manner analogous to the previously described binding of FGFs and VEGFs to their receptors. The PDGFRs form $\alpha\alpha$, $\alpha\beta$, and $\beta\beta$ dimers, each able to bind a subset of PDGF dimers. PDGFR β is expressed on mesenchymal cells, especially vascular SMCs and microvascular pericytes (Chen *et al.*, 2013).

PDGFR signaling

PDGFR β dimerization promotes transphosphorylation and activation of the tyrosine kinase activity leading to subsequent phosphorylation of tyrosine residues, which recruit signaling proteins that activate several of the same cellular signaling cascades previously described for FGF and VEGF including the Ras/MAPK (Erk), PLC γ , and PI3K pathways triggering rapid cytoplasmic responses and nuclear transcription of specific genes required for their cellular proliferation, migration, and survival activities. Given the overlap of many of the same signaling pathways not only between PDGFR α and PDGFR β but also among other receptor tyrosine kinases such as FGFRs and VEGFRs, at least some of the biologic specificity observed with each of the growth factors is the result of their cellular and temporal expression patterns and those of their cognate receptors (Andrae *et al.*, 2008; Demoulin and Essagher, 2014).

FGF–VEGF–Ang–PDGF Networking

Interactions among FGFs, VEGFs, Angs, and PDGFs, necessary to coordinate successful angiogenic responses, can be achieved through a variety of potential mechanisms. In addition to possibly augmenting activation of common signal transduction pathways (Simons, 2004) these growth factors can modulate expression of other growth factors and their receptors.

FGF-2 and VEGF-A can synergistically promote angiogenesis *in vivo*. Although ECs express FGFRs, neutralizing VEGF-A antibodies inhibit FGF-2-induced EC proliferation in culture and angiogenesis *in vivo* by up to 90%. These results are consistent with the observed FGF-2-induced expression of endothelial and stromal cell VEGF-A and EC VEGFR2. In fact, lack of EC FGFR signaling results in unresponsiveness to VEGF, probably a consequence of basal VEGFR downregulation (Murakami and Simons, 2008; Murakami *et al.*, 2011).

FGF-2 and PDGF-BB also synergistically enhance angiogenesis. FGF-2 but not VEGF-A induces PDGFR α and PDGFR β receptors in endothelial and mural cells increasing the response of both types of cells to PDGF angiogenic stimulation. FGF-2 potently induces EC proliferation but only modestly stimulates migration whereas PDGF-BB induces migration but not proliferation so it bolsters the migratory activity of FGF-2 in these cells (Kano *et al.*, 2005; Zhang *et al.*, 2009). In mural cells FGF-2 is also reported to increase expression of PDGF-C, which can bind and activate PDGFR- $\alpha\alpha$ and - $\alpha\beta$ (Lee *et al.*, 2013). Therefore, the observed angiogenic activity of endogenous PDGF-CC and its receptors is facilitated by FGF-2.

FGF-2 and VEGF-A each can stimulate Ang2 expression. VEGF-A has also been found to promote rapid (30 min) and slow (24 h) cleavage of the Tie1 and Tie2 ectodomains, respectively, first increasing and subsequently decreasing the Tie1/Tie2 ratio. This receptor modulation would initially enhance Tie2-mediated EC stabilization but eventually diminish it, having the effect of dampening a response to transient exposure of VEGF but facilitating the response to chronic exposure (Singh *et al.*, 2012).

As summarized in Figure 4, the ability of FGF-2 (and probably the functionally related FGF-1 and FGF-4) to regulate

VEGF, PDGF, and Ang activities indicates that it might orchestrate the coordinated induction of angiogenic responses mediated by these other growth factors (Murakami and Simons, 2008).

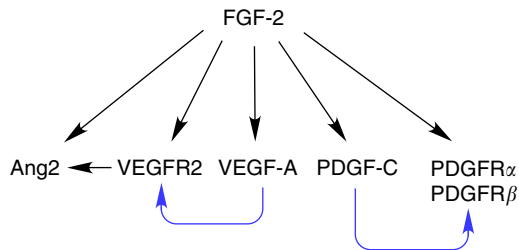


Figure 4 FGF induction of angiogenic growth factors and their receptors. FGF-2 acting through one or more FGFRs can upregulate expression of Ang2 in ECs, PDGF-C in mural cells and VEGF-A in both ECs and stromal cells. VEGFR2 can also upregulate expression of Ang2. FGF-2 also induces expression of VEGFR2 in ECs and upregulates both PDGFR α and PDGFR β in ECs and mural cells. Therefore, FGF-2 appears to coordinately control the expression of several other growth factors and their receptors that participate in the induction of angiogenic responses. Black arrows indicate induction of expression; blue arrows denote ligand activation of their cognate receptors.

Hypoxic Induction of Angiogenesis

Angiogenesis is a primary compensatory response to decreased tissue oxygenation. Normal tissue oxygen levels vary within and among organs but typically fall in a range of 3–9%, substantially less than the 21% present in air. Within cells this is even lower, ranging from 1.3 to 2.5%. *In vitro*, extracellular pO_2 levels of 2% or less are commonly considered to be hypoxic. Below approximately 1% cellular ATP levels fall and apoptosis can be triggered (Carreau *et al.*, 2011).

Cells cope with hypoxic stress by increasing the transcription and translation of several proteins important for angiogenesis and erythropoiesis that promote tissue oxygenation and others, such as glycolytic enzymes, that allow cells to generate ATP under conditions of reduced O_2 availability, albeit much less efficiently than by oxidative phosphorylation. Hypoxia-responsive genes involved in angiogenesis include VEGF-A, PIGF, Ang2, and PDGF-B.

Hypoxic cells can sense low oxygen levels using hypoxia-inducible factor (HIF), an $\alpha\beta$ heterodimeric transcription factor composed of an oxygen labile α subunit and a stable β subunit. Under hypoxic conditions HIF-1 α migrates to the nucleus and forms a heterodimer with HIF-1 β that binds to a specific hypoxia response element (HRE) genomic promoter or enhancer sequence. As illustrated in Figure 5, the HIF

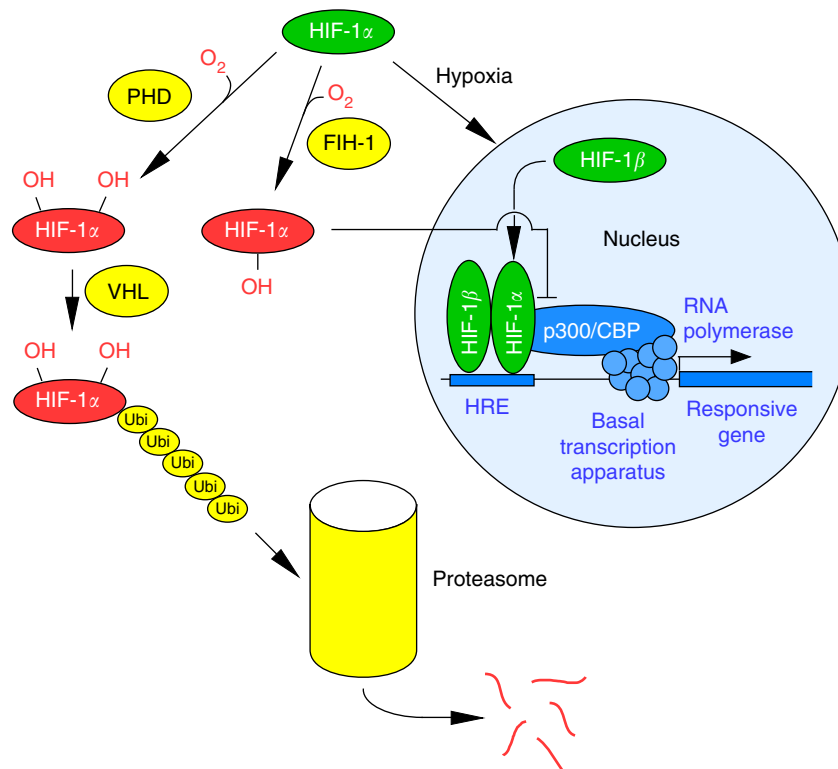


Figure 5 Hypoxia sensing mechanism. In hypoxic conditions hypoxia-inducible factor-1 α (HIF-1 α) migrates to the nucleus and forms a complex with HIF-1 β that binds the genomic hypoxia response element (HRE) DNA sequence. The HRE-bound HIF $\alpha\beta$ heterodimer binds the transcriptional coactivator p300/CBP that also binds and activates the basal transcriptional apparatus. This mechanism activates transcription of a set of HRE-containing genes that encode proteins involved in the metabolic and angiogenic response to hypoxia. In normoxic conditions HIF-1 α is rapidly inactivated by two O_2 -dependent mechanisms. Factor inhibiting HIF-1 (FIH-1) hydroxylates a specific HIF-1 asparagine residue, preventing its binding to the p300/CBP transcriptional coactivator. In addition, prolyl 4-hydroxylase domain (PHD) enzymes hydroxylate two specific HIF-1 α prolyl residues that allow the von Hippel–Lindau (VHL) ubiquitin ligase to polyubiquitinate it triggering proteasome degradation.

heterodimer also binds the p300/CBP (cAMP response element-binding protein) transcriptional coactivator proteins that bridge it to the basal transcription apparatus activating transcription. Two complementary mechanisms exist to prevent the transcription of hypoxia-responsive genes in normoxic conditions. The first is hydroxylation of a specific HIF-1 α asparagine residue by factor inhibiting HIF-1 (FIH-1) that prevents it from binding the p300/CBP coactivator needed to induce transcription. In the second mechanism prolyl 4-hydroxylase domain (PHD) proteins hydroxylate two proline residues on HIF-1 α marking it for ubiquitination by the von Hippel-Lindau (VHL) protein, a member of the E3 ubiquitin ligase complex, and rapid degradation in proteasomes. Below approximately 5% extracellular oxygen the catalytic rates of FIH-1 and PHD diminish as a function of oxygen substrate concentration becoming inactive in its absence so these enzymes serve as the primary oxygen detectors calibrating transcription of hypoxic response genes (Semenza, 2014; Walshe and D'Amore, 2008).

Most transcribed mRNAs are modified by the enzymatic addition of a guanine triphosphate linked through its 5' end to the 5' end of the nascent message, which is then methylated on the guanine seven position to form the m⁷GpppN 'cap' that is recognized by the ribosomal translational initiation cap-binding complex. However, as part of the cellular response to mitosis and to stress, such as severe hypoxia, most cap-dependent protein synthesis is inhibited, presumably to conserve energy. In contrast, the mRNAs encoding proteins important for responding to hypoxia, including those that can drive angiogenesis such as FGF-1, FGF-2, VEGF-A, and PDGF, are preferentially translated under hypoxic conditions. This may occur by use of an internal ribosome entry site (IRES), a stem-loop mRNA structure usually present in the 5' untranslated region upstream of the initiation codon that allows uncapped mRNA translation by one or more incompletely characterized mechanisms (Simon *et al.*, 2008; Komar and Hatzoglou, 2011; Majmundar *et al.*, 2010).

Sprouting Angiogenesis

After completion of early embryonic vasculogenesis most new capillaries appear to be formed by a process of sprouting from the existing microvasculature, especially capillaries, as shown schematically in Figure 6. This process involves a series of steps including initial dissociation of stabilizing pericytes, loosening of inter-EC junctions and degradation of the capillary basement membrane, followed by vascular EC migration, mitosis, tube formation, and fusion culminating in recruitment of pericytes and deposition of basement membrane required to stabilize the new capillary.

ECM Proteolysis

Sprouting angiogenesis is initiated by the binding of angiogenic proteins such as FGFs and VEGFs to their EC-specific plasma membrane receptors. One of the earliest effects of receptor binding and stimulation is the activation of the proenzyme forms of MMPs and serine proteases.

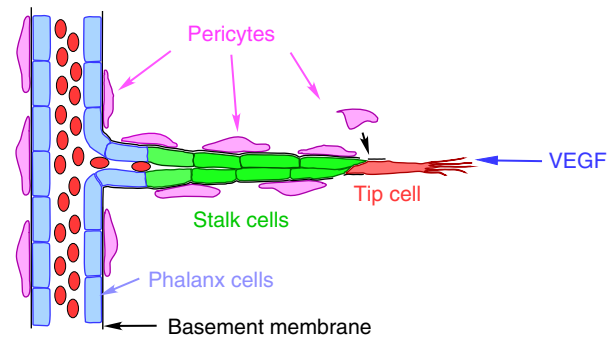


Figure 6 Sprouting angiogenesis. In response to an angiogenic stimulus such as VEGF microvascular ECs proteolytically degrade their basement membrane and migrate away from the parental vessel. The leading tip cell (TC)(red) extends long filopodia containing VEGFRs that allow it to sense and chemotactically migrate up a VEGF gradient. The TC secretes PDGF, which attracts pericytes (magenta), and basement membrane (black line) components onto which the following stalk cells (green) migrate and proliferate allowing the stalk to extend. A narrow slit opens between adjacent stalk cells that can continue to expand into a capillary lumen large enough to accommodate erythrocytes (red ellipses). As the stalk grows ECs near its base become quiescent phalanx cells (blue). A capillary sprout can fuse at its tip either with another sprout to form a continuous linear capillary lumen or with an existing capillary to generate a microvascular branch capable of supporting blood flow.

EC membrane-type 1 MMP (MT1-MMP) is activated intracellularly and exported to the pericellular side of the plasma membrane. Within approximately 30 min of angiogenic growth factor receptor binding a PI3K signaling pathway-dependent activation of MMP-2 occurs that in turn activates urokinase plasminogen activator (uPA), a serine protease that converts inactive plasminogen proenzyme to an active plasmin protease. Plasmin can also activate MMPs including MMP-9, which can degrade the basement membrane collagen IV scaffold. These proteases, along with others that they activate, degrade the constraining capillary basement membrane and ECM. MT1-MMP and uPA/uPA receptor cluster at the leading edge of migrating ECs contributing to their ability to penetrate and move through extracellular matrices (Van Hinsbergh *et al.*, 2006; Breuss and Uhrin, 2012).

Destabilization of EC-EC Junctions

Intercellular junctions between ECs, mediated by adherens junctions and tight junctions, stabilize the vascular endothelial monolayer. These junctions need to be destabilized to allow sprouting ECs to migrate from existing microvessels in response to angiogenic inducers. ECs uniquely express the adherens junction transmembrane protein vascular endothelial-cadherin (VE-cadherin), which contains five extracellular Ig-like domains, a single-pass transmembrane helical sequence and C-terminal intracellular region. VE-cadherins from adjacent ECs are thought to form a zipper-like structure between them. Although the VE-cadherin crystal structure shows overlap between the N-terminal Ig-like domains on VE-cadherins from opposing directions, they are also thought to

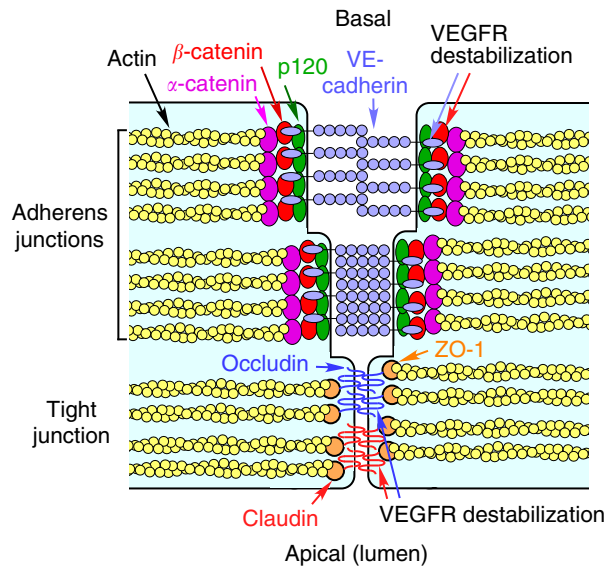


Figure 7 Endothelial junction structure and destabilization. Two types of endothelial cell junctions have been identified, both of which link transmembrane intercellular adhesion proteins to the intracellular actin cytoskeleton. Mutual binding of the N-terminal Ig-like domains of VE-cadherin molecules (blue) from adjacent ECs can create an adherens junction (top) and greater overlap among the Ig-like domains might be able to stabilize narrower junctions (center). The VE-cadherin intracellular domain binds p120 (green) and β -catenin (red), which is linked to actin (yellow) by α -catenin (purple). Narrower tight junctions (bottom) are formed by occludin (dark blue) and claudin (orange), two types of integral membrane proteins, each of which probably multimerizes and forms long ribbons that prevent diffusion of all but small molecules between ECs. VEGFR-activated signaling pathways can induce phosphorylation of VE-cadherin, β -catenin, occludin, and claudin destabilizing both adherens and tight junctions to induce vascular permeability and mobility.

be able to generate narrower adherens junctions by overlapping multiple Ig-like domains as illustrated in [Figure 7](#). VE-cadherin binds near the membrane to p120 that influences VE-cadherin retention and spatial organization at the cell surface. The C-terminal tail of VE-cadherin binds β -catenin that in turn binds α -catenin, which attaches to actin thereby linking cell surface intercellular junctions to intracellular cytoskeleton microfilaments ([Bravi et al., 2014](#)).

EC tight junctions, also illustrated in [Figure 7](#), are interspersed with adherens junctions but are thought to be more prevalent near the apical (luminal) surface. They limit passage of molecules through the EC–EC contact interface to approximately 800 Da. Tight junctions are composed of linear strands consisting of multiple proteins including occludins, EC-specific claudin-5, junctional adhesion molecules (JAMs), and zona occludens-1 (ZO-1), which links it to the actin cytoskeleton. VE-cadherin/ β -catenin, acting through the PI3K/Akt pathway, expels a claudin-5 transcriptional repressor from the nucleus, effectively upregulating claudin-5 transcription and facilitating EC tight junction formation ([Bravi et al., 2014](#)).

The stability of the intercellular junctions is controlled by phosphorylation. VEGFR2, acting through endothelial nitric oxide synthase (eNOS) generation of nitric oxide, can

nitrosylate β -catenin promoting its dissociation from VE-cadherin. VEGFR2-activated signaling pathways also phosphorylate VE-cadherin tyrosines in the binding sites for p120 and β -catenin, inhibiting adherens junction complex formation and promoting the internalization and degradation of VE-cadherin. Either the same or a different pool of β -catenin migrates to the nucleus and interacts with additional transcription factors to downregulate expression of the tight junction protein claudin-5. VEGF also induces serine phosphorylation of occludin, which increases ubiquitination that marks it for proteasome degradation. In addition to destabilizing EC intercellular junctions, growth factor-induced Rho GTPase activity promotes actin cytoskeleton contraction-mediated opening of the gaps between ECs. In contrast, Ang1-activated Tie2 signaling can inhibit VEGF-mediated EC junction destabilization, a process that can be functionally inhibited by the Tie2 weak agonist Ang2 ([Goddard and Iruela-Arispe, 2013](#); [Bravi et al., 2014](#)). Therefore, as denoted in [Figure 7](#), VEGF-induced destabilization of both adherens and tight junctions not only increases vascular permeability but also inhibits intercellular adhesion facilitating EC migration.

Dissociation of Pericytes

In stable vessels pericytes and in larger vessels SMCs constitutively express Ang1, which binds and activates EC Tie2 receptors. Tie2 intracellular signaling in ECs promotes cell survival and vessel stabilization by activation of PI3K, which leads to Akt-mediated inhibition of apoptosis, and Rho GTPase-induced stabilization of adhesion junctions between ECs. In response to hypoxia and to VEGF, ECs upregulate and release Ang2 from their Weibel–Palade storage granules. This inhibits the Ang1-mediated EC quiescence, facilitating angiogenic remodeling with dissociation of stabilizing pericytes from the capillary ([Augustin et al., 2009](#)). ECs and pericytes can form a common adherens junction that utilizes N-cadherin rather than the VE-cadherin found in EC–EC adherens junctions ([Gaengel et al., 2009](#)). Possibly some of the same or similar pathways that trigger VE-cadherin destabilization can also disassemble these N-cadherin adherens junctions.

Migration and Proliferation

Once the intercellular adhesive and matrix barriers are removed ECs can migrate from an existing microvessel to form a new capillary shoot. This is accomplished by an organized process in which a single cell, called a ‘tip cell’ (TC), shown in [Figure 6](#), exhibits a migratory but minimally proliferative phenotype. The initial induction of this TC phenotype is mediated by VEGF and possibly biased toward selection of the first available cell released from its intercellular junction and matrix constraints. Soluble VEGFR1 (sVEGFR1/sFlt1) is produced by the parental vessel ECs near the sprout base. By sequestering VEGF on either side of the sprout sVEGFR1 creates a perpendicular corridor of VEGF that keep the sprout growing away from the parental vessel ([Chappell et al., 2012](#)). In response to VEGF TCs extend multiple thin extensions called filopodia that incorporate VEGFR2 and VEGFR3, at least some of which function as VEGFR2/3 heterodimers, allowing them

to sense and migrate up a VEGF chemotactic gradient. Receptors for other guidance signals on TC filopodia can also sample the local environment, detecting other attractive or repulsive ligands that promote filopodia extension or retraction, respectively, thereby controlling directional migration in a manner analogous to axonal growth cones (Potente *et al.*, 2011; Michaelis, 2014; Jeltsch *et al.*, 2013).

The migrating TC is followed by a column of ECs denoted as stalk cells as illustrated in Figure 6 that elaborate few, if any, filopodia but are proliferative so can elongate the growing capillary stalk. Tip and stalk cell phenotypes are maintained by mutual cross-modulation of their Notch signaling systems. By virtue of being exposed to the highest gradient concentration of VEGF, TCs express decreased levels of the Notch transmembrane receptor and elevated levels of its transmembrane delta-like 4 (Dll4) ligand. TC Dll4 binds and activates Notch in adjacent stalk cells as shown in Figure 8. On activation the Notch intracellular domain (NICD) is liberated by limited proteolysis and migrates to the nucleus where it downregulates VEGFR2 and VEGFR3 and also upregulates VEGFR1, which binds VEGF with a higher affinity, effectively sequestering it to prevent additional TC generation and to maintain the highest chemotactic VEGF gradient concentration ahead of the TC. A second membrane-anchored Notch ligand, Jag1, is expressed by stalk cells and functions as a weak agonist, which decreases TC Notch signaling by competing with the stronger Dll4 agonist. Therefore, by this mutual 'juxtacrine' signaling between adjacent TCs and stalk cells Notch signaling is increased in stalk cells but decreased in TCs.

The unique phenotype of a TC is actively maintained but transient. The TC phenotype half-life has been measured to be approximately 4 h *in vivo*, after which it reverts to being a stalk cell and is replaced by another stalk cell that becomes a TC (Carmeliet and Jain, 2011; Potente *et al.*, 2011; Ribatti and Crivellato, 2012).

Sprout Lumen Formation

Several alternative mechanisms of lumen formation have been proposed including (1) extension from an existing lumen, (2) fusion of vacuoles within single ECs that then fuse in a sequence of ECs along the sprout, and (3) formation of a slit between laterally adjacent stalk cells followed by expansion to create an open lumen. Although any or all of these mechanisms might contribute to lumen formation, the process appears to be driven primarily by the third alternative. The initial step in this process is generated by migration of adherens and tight junctions between ECs toward the basal (outside) edge of the sprout cord EC-EC interfaces and exocytotic insertion of membrane-anchored protein sialomucins including podocalyxin (PODXL) into the juxtaposed plasma membranes as illustrated in Figure 9(a). The highly negatively charged sialomucin oligosaccharides repel the adjacent EC plasma membranes creating an open slit between them, shown in Figure 9(b). The slit is subsequently expanded into a lumen by redistribution of moesin to the apical (luminal) membrane, where after being phosphorylated by PKC it links PODXL to F-actin. In addition, VEGF signaling activates Rho-associated coiled-coil kinase (ROCK) and phosphorylation of non-

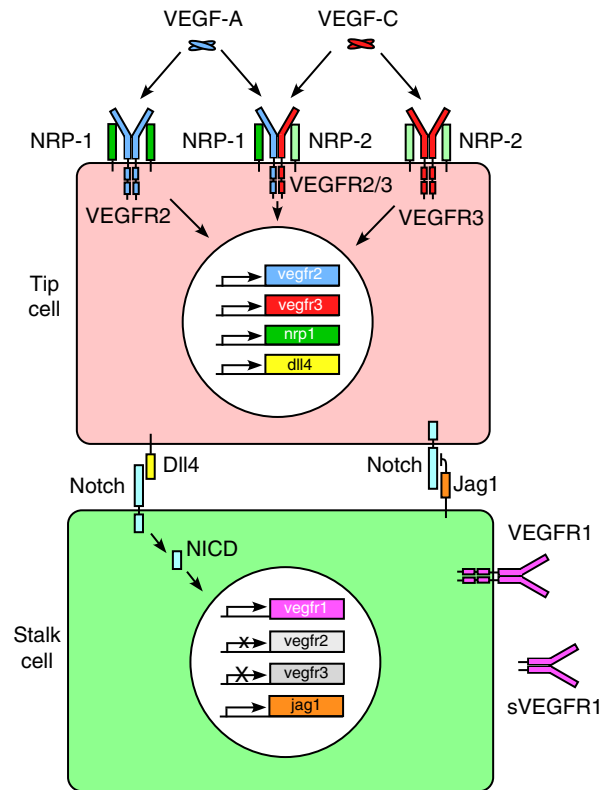


Figure 8 Tip-stalk cell signaling. Tip cells (TCs) (light red) express VEGFR2 and VEGFR3 homodimers along with VEGFR2/3 heterodimers, each of which can differentially activate intracellular signaling cascades leading to an invasive migratory phenotype. In addition, TCs express NRP VEGFR co-receptors and the membrane-anchored juxtacrine ligand Dll4, which binds its Notch receptor on adjacent stalk cells (green). This triggers proteolytic release of the Notch intracellular domain (NICD) that functions as a transcription factor increasing expression of VEGFR1 and Jag1 and downregulating VEGFR2 and VEGFR3 expression. Stalk cell membrane-bound and soluble VEGFR1 can sequester VEGF effectively diminishing its stimulation. Notch-induced Jag1 expression by stalk cells inhibits Dll4 activation of TC Notch, preventing NICD formation and downregulation of the genes that support the TC phenotype. This mutual juxtacrine signaling establishes a phenotypic difference between the leading TC and following stalk cells.

muscle myosin-II facilitating interaction with F-actin forming an actomyosin complex that generates the force required to retract the EC creating a fully open lumen, which is illustrated in Figure 9(c). The vessel and lumen diameter can subsequently grow by vacuole fusion and EC proliferation to create larger microvessels (Herbert and Stainier, 2011; Lamert and Axnick, 2012; Siekmann *et al.*, 2013).

Sprout Fusion and Flow

Functional microvascular loops can form by the fusion of two capillary sprouts. TC filopodia from each sprout contact each other and based on co-localization of VE-cadherin and ZO-1 likely form both stabilizing adherens and tight junctions. Macrophages can help stabilize the nascent junction but are

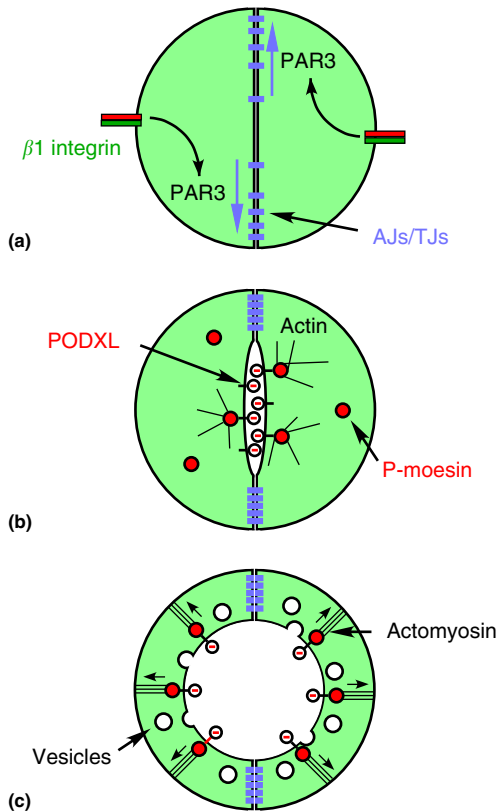


Figure 9 Lumen formation. (a) Activation of $\beta 1$ -containing integrins on a pair of adjacent stalk cells (green) induces PAR3-mediated migration of adherens junctions (AJs) and tight junctions (TJs) to the outer edge of the cellular interface. (b) Membrane-anchored highly negatively charged sialomucin proteins including PODXL are inserted into the adjacent EC plasma membranes creating charge repulsion that separates them. PKC-phosphorylated moesin (P-moesin, red circles) binds to the intracellular region of PODXL linking it to actin (straight black lines), which expands the opening. (c) Additional lumen expansion is mediated by contraction of actomyosin generated by expression of myosin that binds actin (parallel black lines between P-moesin and the cellular outer membrane) and fusion of vacuoles (open circles) with the open lumen.

not required. As the interface between the two juxtaposed TC membranes increases area they generate a luminal segment that can fuse with the advancing lumen from each sprout to form a continuous open capillary loop. Although branching can occur by sprout splitting it can also be generated by fusion of a sprout tip with an existing capillary (Siekman *et al.*, 2013; Lenard *et al.*, 2013). Once a continuous lumen is formed blood flow can commence, increasing oxygen delivery and downregulating hypoxia-induced transcriptional and translational expression of angiogenic mediators.

Vessel Stabilization and Maturation

Newly formed vessels are stabilized by basement membrane and pericytes. Deposition of basement membrane between ECs and pericytes contributes to the stability of newly formed capillaries by providing: (1) mechanical strength generated by

collagen IV covalent crosslinking, (2) integrin cell attachment ligands, and (3) a growth factor binding, storage, and presentation matrix. PDGF-BB expressed by ECs can bind HSPGs, presumably maintaining locally elevated concentrations that help recruit and retain pericytes by activating their PDGFR β receptors. Deletion of either the *Pdgfb* or *Pdgfr β* gene produces pericyte deficiency, microaneurysms, and bleeding demonstrating the critical need for pericyte microvascular stabilization. Pericytes in turn express Ang1 and activate EC Tie2 receptors stabilizing EC-EC junctions that inhibit vascular permeability. Therefore, EC PDGF-BB and pericyte Ang1 each function in a mutually paracrine manner to stabilize microvascular structure (Gaengel *et al.*, 2009). In addition, pericytes secrete the ECM protein vitronectin that can bind to the αv subunit of the EC integrin $\alpha v \beta 3$, upregulating EC expression of VEGF that functions in an intracellular or 'intracrine' manner binding to cytoplasmic VEGFR2 within the same cell, promoting expression of the anti-apoptotic Bcl protein to support EC survival and homeostasis (Domigan *et al.*, 2015).

FGFs not only induce angiogenesis but at low concentration support basal receptor signaling required for vascular stability. This appears to be accomplished by FGFR-mediated stabilization of the VE-cadherin/p120 catenin adherens junction complex mediated by the SHP2 phosphatase, which can reduce VE-cadherin phosphorylation-induced dissociation of p120 catenin (Goddard and Iruela-Arispe, 2013).

During vessel maturation the lipid sphingosine-1-phosphate (S1P), present primarily bound to blood-borne high-density lipoprotein (HDL) and albumin, functions to stabilize patent vessels. Acting through its three cell surface G-protein-coupled receptors on ECs, S1P stimulates the Rho family GTPases (Rac, Rho, and Cdc42) that regulate the actin cytoskeleton and EC barrier function. In particular, S1P activation of the S1P₁ receptor mediates Rac activation, promoting accumulation of VE-cadherin and β -catenin at cell-cell interfaces and strengthens both EC-EC VE-cadherin and EC-pericyte N-cadherin adherens junction assembly, thereby inhibiting angiogenic sprouting (Wietecha *et al.*, 2013; Xiong and Hla, 2014; Goddard and Iruela-Arispe, 2013).

Vessel Pruning

Angiogenic stimuli can induce numerous capillaries, some of which are either redundant or not positioned in the microvascular bed to achieve adequate blood flow. After hypoxic tissue becomes fully oxygenated by angiogenic expansion of the microvascular bed the survival of fully functional vessels exhibiting nonturbulent laminar flow shear stress is enhanced. The mechanism by which laminar flow sustains vessels is not entirely established but appears to involve activation of the PI3K/Akt pathway in ECs, which increases nitric oxide synthase and its product nitric oxide that can S-nitrosylate the active site cysteine in caspases, inactivating these proteolytic mediators of apoptosis. In contrast, capillaries in well-oxygenated environments that are inefficiently positioned in the microvascular bed to establish and sustain nonturbulent blood flow, do not efficiently activate this pathway and are more susceptible to apoptosis (Dimmeler and Zeiher, 2000), which promotes the destabilization and regression of these

less efficient vessels. In an analogous manner, retinal ECs exposed to hyperoxia then returned to normoxia, a procedure that causes retinal microvascular pruning *in vivo*, express FGD5, a guanine exchange factor protein that activates Cdc42 GTPase. Cdc42 induces Notch target genes resulting in downregulation of VEGFR2 and upregulation of VEGFR1 expression as occurs in stalk cells during sprouting, presumably sequestering VEGF, resulting in diminished survival signaling, p53 induction of apoptosis, and vessel regression (Wietecha *et al.*, 2013). In the absence of an angiogenic signal such as VEGF, Ang2 can also promote capillary destabilization and regression (Huang *et al.*, 2010; Jeltsch *et al.*, 2013). The net effect of microvascular pruning is to eliminate redundant and poorly functioning vessels, thereby optimizing efficient flow.

Intussusceptive Angiogenesis

Although sprouting has been the focus of most angiogenesis studies, an alternative method of capillary growth is longitudinal splitting or intussusception of existing microvessels that has been observed to occur in rapid developmental organ growth and the muscular response to exercise. Descriptive studies have indicated that the process begins with circumferential enlargement of preexisting microvessels exhibiting robust flow. This increase in diameter is accompanied by corresponding decrease in EC thickness with little or no proliferation and is followed by formation of a narrow (<5 μm) cylindrical pillar of tissue across the lumen that can consist of multiple ECs invaded by a core of pericytes or myofibroblasts as shown in Figure 10. Pillar formation is thought to be initiated by creation of a contact zone between opposing microvascular walls, perhaps assisted by adjacent cells such as pericytes. Additional pillars can form in sequence downstream from existing pillars. The luminal pillars can each grow in diameter culminating in their fusion to create a pair of internal EC walls extending down the axis of the vessel that allow it to split into two parallel daughter vessel branches. Intussusception is capable of rapidly expanding a microvascular network with little or no EC proliferation. It can also remodel an existing vascular plexus by either creating new vascular branches or by walling off and pruning existing branches (Makanya *et al.*, 2009).

Intussusception might reflect a modification of sprouting in which 'tip' ECs grow into the lumen instead of moving in the opposite direction into surrounding tissue. This is consistent with the lack of basement membrane and ECM proteolytic degradation in intussusception. The initiating stimulus for intussusception is unknown but might be triggered in ECs by hemodynamic forces within the vessels. Circulating angiogenic and remodeling factors such as FGF-2, VEGF, PDGF-B, and Tie2 also have been implicated. In particular, targeted Tie2 genetic deletion is associated with decreased pillar formation. However, to date, details of the molecular mechanisms that trigger and support intussusceptive angiogenesis remain essentially uncharacterized (Mentzer and Konerding, 2014).

Arterial and Venous Vessels

Although angiogenesis is a microvascular growth process blood flow into and from the capillaries is provided by small arterioles

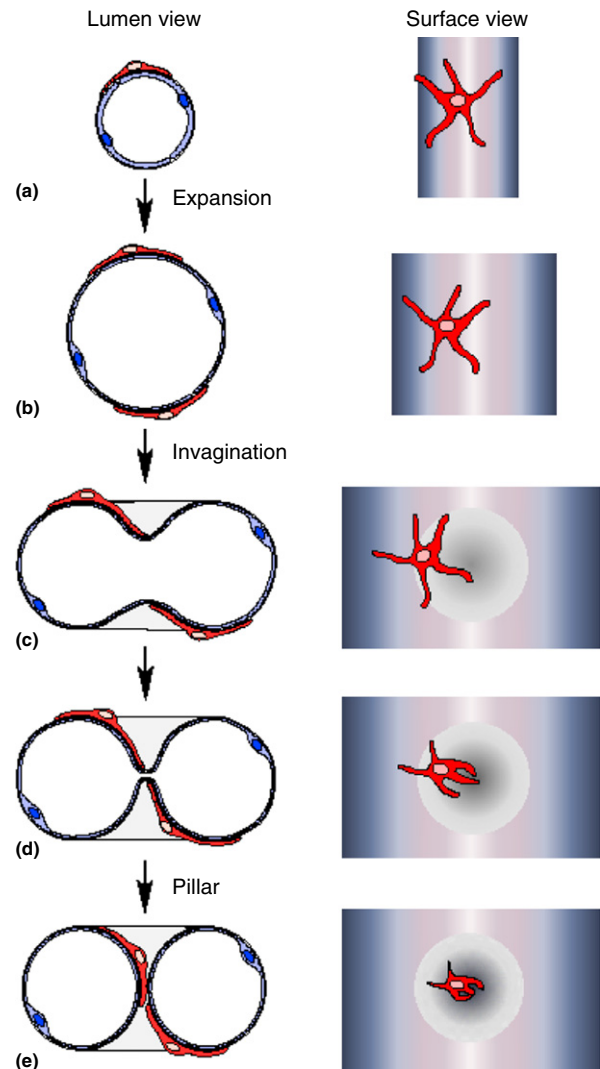


Figure 10 Intussusceptive angiogenesis pillar formation. On the basis of microscopic observations intussusception appears to be initiated by formation of pillars that span the capillary lumen. This process is initiated in (a) an existing capillary by (b) EC surface expansion with decreased thickness allowing an existing vessel diameter to increase rapidly without proliferation. (c) EC surface depressions form, probably facilitated by pericytes or perhaps myofibroblasts, on one or both sides of the vessel leading to (d) deep invaginations. (e) As the two membranes meet they create a tunnel or pillar typically seen to be occupied by a pericyte or myofibroblast. In intussusceptive angiogenesis pillars commonly form in rows along the capillary, grow and fuse to create a pair of plasma membranes partitioning a lumen into two halves that then separate to form a branched pair of capillaries.

and venules, which in turn are supplied by larger arteries and veins. The fate of arterial and venous ECs is established in early embryogenesis by expression of the transmembrane EphrinB2 ligand and its EphB4 receptor tyrosine kinase. During vasculogenesis angioblasts expressing either EphB4 or EphrinB2 are each incorporated into the dorsal aorta but the EphB4-expressing cells subsequently migrate to form the cardinal vein.

In adults, arterial, arteriole, and capillary ECs express EphrinB2 whereas veins and venules predominately express EphB4, consistent with the early developmental segregation of expression. However, the expression of EphrinB2 and EphB4 overlap in several vessels. In addition, pericytes and SMCs express EphrinB2 and EphB4, which can promote and stabilize their normal attachment to ECs (Salvucci and Tosato, 2012).

Arterial development also is mediated by Notch activation of the HEY transcription factor signaling (Park *et al.*, 2013). In addition, as previously described, the VEGFR co-receptors NRP-1 and NRP-2 are asymmetrically distributed between arterial and venous ECs, respectively, although the significance of this distribution is not yet clear.

Summary

The characterization of angiogenic cellular responses and elaboration of some of their underlying genetic and biochemical mechanisms has advanced rapidly over the past few decades with the identification and characterization of the extra- and intra-cellular signals that trigger, coordinate and terminate the growth of blood vessels. The progress described in this article has been made possible by the development and application of new genetic, biochemical, cellular, and whole animal approaches that have been used to dissect the intricate control networks coordinating this multicellular physiologic response. Nevertheless, our understanding of this fundamental biological process remains incomplete and will likely continue to be a major focus of research in vascular cell biology for the foreseeable future.

References

- Albuquerque, R.J., 2013. The newest member of the VEGF family. *Blood* 121, 4015–4016.
- Andrae, J., Gallini, R., Betsholtz, C., 2008. Role of platelet-derived growth factors in physiology and medicine. *Genes & Development* 22, 1276–1312.
- Antoine, M., Wirz, W., Tag, C.G., *et al.*, 2005. Expression pattern of fibroblast growth factors (FGFs), their receptors and antagonists in primary endothelial cells and vascular smooth muscle cells. *Growth Factors* 23, 87–95.
- Armulik, A., Genove, G., Betsholtz, C., 2011. Pericytes: Developmental, physiological, and pathological perspectives, problems, and promises. *Developmental Cell* 21, 193–215.
- Augustin, H.G., Koh, G.Y., Thurston, G., Alitalo, K., 2009. Control of vascular morphogenesis and homeostasis through the angiopoietin-Tie system. *Nature Reviews Molecular Cell Biology* 10, 165–177.
- Barton, W.A., Dalton, A.C., Seegar, T.C., Himanen, J.P., Nikolov, D.B., 2014. Tie2 and Eph receptor tyrosine kinase activation and signaling. *Cold Spring Harbor Perspectives in Biology* 6, a009142.
- Bravi, L., Dejana, E., Lampugnani, M.G., 2014. VE-cadherin at a glance. *Cell and Tissue Research* 355, 515–522.
- Breuss, J.M., Uhrin, P., 2012. VEGF-initiated angiogenesis and the uPA/uPAR system. *Cell Adhesion & Migration* 6, 535–615.
- Carmeliet, P., Jain, R.K., 2011. Molecular mechanisms and clinical applications of angiogenesis. *Nature* 473, 298–307.
- Carreau, A., El Hafny-Rahbi, B., Matejuk, A., Grillon, C., Kieda, C., 2011. Why is the partial oxygen pressure of human tissues a crucial parameter? Small molecules and hypoxia. *Journal of Cellular and Molecular Medicine* 15, 1239–1253.
- Chappell, J.C., Wiley, D.M., Bautch, V.L., 2012. How blood vessel networks are made and measured. *Cells Tissues Organs* 195, 94–107.
- Chen, P.H., Chen, X., He, X., 2013. Platelet-derived growth factors and their receptors: Structural and functional perspectives. *Biochimica et Biophysica Acta* 1834, 2176–2186.
- Demoulin, J.B., Essaghir, A., 2014. PDGF receptor signaling networks in normal and cancer cells. *Cytokine & Growth Factor Reviews* 25, 273–283.
- De Val, S., Black, B.L., 2009. Transcriptional control of endothelial cell development. *Developmental Cell* 16, 180–195.
- De Val, S., Chi, N.C., Meadows, S.M., *et al.*, 2008. Combinatorial regulation of endothelial gene expression by ets and forkhead transcription factors. *Cell* 135, 1053–1064.
- Dimmeler, S., Zeiher, A.M., 2000. Endothelial cell apoptosis in angiogenesis and vessel regression. *Circulation Research* 87, 434–439.
- Djordjevic, S., Driscoll, P.C., 2013. Targeting VEGF signalling via the neuropilin co-receptor. *Drug Discovery Today* 18, 447–455.
- Domigan, C.K., Ziyad, S., Iruela-Arispe, M.L., 2015. Canonical and noncanonical vascular endothelial growth factor pathways: New developments in biology and signal transduction. *Arteriosclerosis, Thrombosis, and Vascular Biology* 35, 30–39.
- Dorey, K., Amaya, E., 2010. FGF signalling: Diverse roles during early vertebrate embryogenesis. *Development* 137, 3731–3742.
- Eble, J.A., Niland, S., 2009. The extracellular matrix of blood vessels. *Current Pharmaceutical Design* 15, 1385–1400.
- Eichmann, A., Yuan, L., Moyon, D., *et al.*, 2005. Vascular development: From precursor cells to branched arterial and venous networks. *International Journal of Developmental Biology* 49, 259–267.
- Elfenbein, A., Rhodes, J.M., Meller, J., *et al.*, 2009. Suppression of RhoG activity is mediated by a syndecan 4-synectin-RhoGDI1 complex and is reversed by PKC α in a Rac1 activation pathway. *Journal of Cell Biology* 186, 75–83.
- Favier, B., Alam, A., Barron, P., *et al.*, 2006. Neuropilin-2 interacts with VEGFR-2 and VEGFR-3 and promotes human endothelial cell survival and migration. *Blood* 108, 1243–1250.
- Ferrara, N., 2002. VEGF and the quest for tumour angiogenesis factors. *Nature Reviews Cancer* 2, 795–803.
- Ferrara, N., 2010. Binding to the extracellular matrix and proteolytic processing: Two key mechanisms regulating vascular endothelial growth factor action. *Molecular Biology of the Cell* 21, 687–690.
- Furdui, C.M., Lew, E.D., Schlessinger, J., Anderson, K.S., 2006. Autophosphorylation of FGFR1 kinase is mediated by a sequential and precisely ordered reaction. *Molecular Cell* 21, 711–717.
- Gaengel, K., Genove, G., Armulik, A., Betsholtz, C., 2009. Endothelial-mural cell signaling in vascular development and angiogenesis. *Arteriosclerosis, Thrombosis, and Vascular Biology* 29, 630–638.
- Goddard, L.M., Iruela-Arispe, M.L., 2013. Cellular and molecular regulation of vascular permeability. *Thrombosis and Haemostasis* 109, 407–415.
- Goetz, R., Mohammadi, M., 2013. Exploring mechanisms of FGF signalling through the lens of structural biology. *Nature Reviews Molecular Cell Biology* 14, 166–180.
- Herbert, S.P., Stainier, D.Y., 2011. Molecular control of endothelial cell behaviour during blood vessel morphogenesis. *Nature Reviews Molecular Cell Biology* 12, 551–564.
- Holzmann, K., Grunt, T., Heinzel, C., *et al.*, 2012. Alternative splicing of fibroblast growth factor receptor IgIII loops in cancer. *Journal of Nucleic Acids* 2012, 950508.
- Huang, H., Bhat, A., Woodnutt, G., Lappe, R., 2010. Targeting the ANGPT-TIE2 pathway in malignancy. *Nature Reviews Cancer* 10, 575–585.
- Izzo, R.V., 2005. Basement membrane proteoglycans: From cellar to ceiling. *Nature Reviews Molecular Cell Biology* 6, 646–656.
- Ishitobi, H., Wakamatsu, A., Liu, F., *et al.*, 2011. Molecular basis for Flk1 expression in hemato-cardiovascular progenitors in the mouse. *Development* 138, 5357–5368.
- Jeltsch, M., Leppanen, V.M., Saharinen, P., Alitalo, K., 2013. Receptor tyrosine kinase-mediated angiogenesis. *Cold Spring Harbor Perspectives in Biology* 5, a009183.
- Kano, M.R., Morishita, Y., Iwata, C., *et al.*, 2005. VEGF-A and FGF-2 synergistically promote neoangiogenesis through enhancement of endogenous PDGF-B-PDGFR β signaling. *Journal of Cell Science* 118, 3759–3768.
- Kimelman, D., 2006. Mesoderm induction: From caps to chips. *Nature Reviews Genetics* 7, 360–372.
- Koch, S., 2012. Neuropilin signalling in angiogenesis. *Biochemical Society Transactions* 40, 20–25.
- Koch, S., Claesson-Welsh, L., 2012. Signal transduction by vascular endothelial growth factor receptors. *Cold Spring Harbor Perspectives in Medicine* 2, a006502.
- Koch, S., Tugues, S., Li, X., Gualandi, L., Claesson-Welsh, L., 2011. Signal transduction by vascular endothelial growth factor receptors. *Biochemical Journal* 437, 169–183.

- Komar, A.A., Hatzoglou, M., 2011. Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states. *Cell Cycle* 10, 229–240.
- Lammert, E., Axnick, J., 2012. Vascular lumen formation. *Cold Spring Harbor Perspectives in Medicine* 2, a006619.
- Lee, C., Zhang, F., Tang, Z., Liu, Y., Li, X., 2013. PDGF-C: A new performer in the neurovascular interplay. *Trends in Molecular Medicine* 19, 474–486.
- Lenard, A., Ellertsdottir, E., Herwig, L., *et al.*, 2013. *In vivo* analysis reveals a highly stereotypic morphogenetic pathway of vascular anastomosis. *Developmental Cell* 25, 492–506.
- Majmundar, A.J., Wong, W.J., Simon, M.C., 2010. Hypoxia-inducible factors and the response to hypoxic stress. *Molecular Cell* 40, 294–309.
- Makanya, A.N., Hlushchuk, R., Djonov, V.G., 2009. Intussusceptive angiogenesis and its role in vascular morphogenesis, patterning, and remodeling. *Angiogenesis* 12, 113–123.
- Mentzer, S.J., Konerding, M.A., 2014. Intussusceptive angiogenesis: Expansion and remodeling of microvascular networks. *Angiogenesis* 17, 499–509.
- Michaelis, U.R., 2014. Mechanisms of endothelial cell migration. *Cellular and Molecular Life Sciences* 71, 4131–4148.
- Munoz-Chapuli, R., 2011. Evolution of angiogenesis. *International Journal of Developmental Biology* 55, 345–351.
- Murakami, M., Nguyen, L.T., Hatanaka, K., *et al.*, 2011. FGF-dependent regulation of VEGF receptor 2 expression in mice. *Journal of Clinical Investigation* 121, 2668–2678.
- Murakami, M., Simons, M., 2008. Fibroblast growth factor regulation of neovascularization. *Current Opinion in Hematology* 15, 215–220.
- Ornitz, D.M., Xu, J., Colvin, J.S., *et al.*, 1996. Receptor specificity of the fibroblast growth factor family. *Journal of Biological Chemistry* 271, 15292–15297.
- Park, C., Kim, T.M., Malik, A.B., 2013. Transcriptional regulation of endothelial cell and vascular development. *Circulation Research* 112, 1380–1400.
- Parker, M.W., Guo, H.F., Li, X., Linkugel, A.D., Vander Kooi, C.W., 2012. Function of members of the neuropilin family as essential pleiotropic cell surface receptors. *Biochemistry* 51, 9437–9446.
- Plein, A., Fantin, A., Ruhrberg, C., 2014. Neuropilin regulation of angiogenesis, arteriogenesis, and vascular permeability. *Microcirculation* 21, 315–323.
- Potente, M., Gerhardt, H., Carmeliet, P., 2011. Basic and therapeutic aspects of angiogenesis. *Cell* 146, 873–887.
- Ribatti, D., Crivellato, E., 2012. "Sprouting angiogenesis", a reappraisal. *Developmental Biology* 372, 157–165.
- Saharinen, P., Eklund, L., Miettinen, J., *et al.*, 2008. Angiopoietins assemble distinct Tie2 signalling complexes in endothelial cell-cell and cell-matrix contacts. *Nature Cell Biology* 10, 527–537.
- Salvucci, O., Tosato, G., 2012. Essential roles of EphB receptors and EphrinB ligands in endothelial cell function and angiogenesis. *Advances in Cancer Research* 114, 21–57.
- Semenza, G.L., 2014. Oxygen sensing, hypoxia-inducible factors, and disease pathophysiology. *Annual Review of Pathology Mechanisms of Disease* 9, 47–71.
- Shewchuk, L.M., Hassell, A.M., Ellis, B., *et al.*, 2000. Structure of the Tie2 RTK domain: Self-inhibition by the nucleotide binding loop, activation loop, and C-terminal tail. *Structure* 8, 1105–1113.
- Siekmann, A.F., Affolter, M., Belting, H.G., 2013. The tip cell concept 10 years after: New players tune in for a common theme. *Experimental Cell Research* 319, 1255–1263.
- Simon, M.C., Liu, L., Barnhart, B.C., Young, R.M., 2008. Hypoxia-induced signaling in the cardiovascular system. *Annual Review of Physiology* 70, 51–71.
- Simons, M., 2004. Integrative signaling in angiogenesis. *Molecular and Cellular Biochemistry* 264, 99–102.
- Singh, H., Hansen, T.M., Patel, N., Brindle, N.P., 2012. The molecular balance between receptor tyrosine kinases Tie1 and Tie2 is dynamically controlled by VEGF and TNFalpha and regulates angiopoietin signalling. *PLoS One* 7, e29319.
- Singh, N., Tiem, M., Watkins, R., *et al.*, 2013. Soluble vascular endothelial growth factor receptor 3 is essential for corneal alymphaticity. *Blood* 121, 4242–4249.
- Thomas, K.A., 2013. Vascular endothelial growth factor receptors. In: Lennarz, W.J., Lane, M.D. (Eds.), *Encyclopedia of Biological Chemistry*, second ed. San Diego, CA: Elsevier, pp. 509–514.
- Van Hinsbergh, V.W., Engelse, M.A., Quax, P.H., 2006. Pericellular proteases in angiogenesis and vasculogenesis. *Arteriosclerosis, Thrombosis, and Vascular Biology* 26, 716–728.
- Vempati, P., Popel, A.S., Mac Gabhann, F., 2014. Extracellular regulation of VEGF: Isoforms, proteolysis, and vascular patterning. *Cytokine & Growth Factor Reviews* 25, 1–19.
- Walshe, T.E., D'Amore, P.A., 2008. The role of hypoxia in vascular injury and repair. *Annual Review of Pathology* 3, 615–643.
- Wietecha, M.S., Cerny, W.L., Dipietro, L.A., 2013. Mechanisms of vessel regression: Toward an understanding of the resolution of angiogenesis. *Current Topics in Microbiology and Immunology* 367, 3–32.
- Xiong, Y., Hla, T., 2014. S1P control of endothelial integrity. *Current Topics in Microbiology and Immunology* 378, 85–105.
- Yamaguchi, T.P., Harpal, K., Henkemeyer, M., Rossant, J., 1994. fgfr-1 is required for embryonic growth and mesodermal patterning during mouse gastrulation. *Genes & Development* 8, 3032–3044.
- Zhang, J., Cao, R., Zhang, Y., *et al.*, 2009. Differential roles of PDGFR-alpha and PDGFR-beta in angiogenesis and vessel stability. *FASEB Journal* 23, 153–163.
- Zhang, X., Ibrahim, O.A., Olsen, S.K., *et al.*, 2006. Receptor specificity of the fibroblast growth factor family. The complete mammalian FGF family. *Journal of Biological Chemistry* 281, 15694–15700.
- Zhang, X., Simons, M., 2014. Receptor tyrosine kinases endocytosis in endothelium: Biology and signaling. *Arteriosclerosis, Thrombosis, and Vascular Biology* 34, 1831–1837.

A Review on Various Mathematical Modeling Approaches for Wound Healing

FJ Vermolen, Delft University of Technology, Delft, The Netherlands

© 2016 Elsevier Inc. All rights reserved.

Wound Healing and the Need for Modeling

Healing of damaged tissue is crucially important for the functionality of organs, survival, and well-being of humans and animals. To heal skin on which a wound has been inflicted a chain of many biological processes is initiated, of which the sequence and intensity is well orchestrated by the physiological system. Important processes are regeneration of extracellular matrix, the proliferation of constituent cells, such as fibroblasts in the case of skin wounds (burns and cutaneous wounds), cell death, clearance of polluting agents, engulfment of pathogens, and migration of cells. This migration is characteristic for the constituent cells and for the immune cells that hunt possible pathogens. This cellular migration is triggered by the physical environment that is experienced by the cell or the group of cells. From a medical perspective, the processes are characterized by the following sequence of partly overlapping consecutive phases: hemostasis, the proliferative phase, and the remodeling phase. During the hemostatic phase of cutaneous wounds, blood platelets originating, present from bleeding, regenerate a polymeric connective tissue state based on clotting that effectively seals of the damaged region, thereby effectively sealing off the wounded area. This temporary phase of connective tissue prevents the ingress of hazardous chemicals and pathogens and is hence indispensable for the survival of the organism. Next to the regeneration of connective tissue, the platelets release signaling cytokines that trigger the immune cells like macrophages to enter the wound area. These immune cells will clear up pathogens and hazardous chemicals and, in turn, release chemokines that initiate further cellular processes that take place in the proliferative phase. These processes are wound contraction, wound closure, and wound angiogenesis. On a fundamental biological level, cellular processes like migration and division (proliferation) are evoked. Wound angiogenesis is responsible for the revascularization of the damaged region if a deeper wound (in the sense that the dermis is wounded) is inflicted. This revascularization is necessary for the supply of oxygen and nutrients that are of vital importance to the constituent cells as well as for the functionality of the immune system in case of future wound healing and to clear up virally infected cells, detrimental bacteria, and mutated cells in the vicinity to prevent initiation of cancer. The (intermediate) contraction that is commonly observed in the deeper wounds is a survival mechanism to decrease the area of the wound exposed to the surrounded. Although this mechanism is highly favorable in cutaneous wounds, this mechanism is very undesirable in burns since severe contractions lead to contractures, which decrease the patient mobility, and possibly coexist with hypertrophic scars, which poses esthetic problems to the patient. The first two phases typically last in the order of several hours and several days, respectively. The final remodeling phase typically lasts several months up to several years or even remains uncompleted during the life span of the organism or patient.

The importance of wound healing, and wound prevention in cases where patients who are hospitalized over a long period and become sensitive to pressure ulcers, has given rise to several medical treatments with huge differences. These treatments range from the attachment of honey, which has been evidenced to inhibit progression of infection mechanisms into the wound region to stitching large wounds or even applying plastic surgery. In order to improve therapies, it is important to reveal the biological mechanism that underlies wound healing. To unravel these mechanisms, experimental data both *in vitro* and *in vivo* experimental data along with clinical observations are needed. Although the availability of these data is a necessary condition for the development of insight into the biology of wound healing, it is not yet sufficient for this understanding. To gain more understanding it is also necessary to develop theory on the basis of sound hypotheses. In order to validate biological hypotheses with experimental results, which merely are characterized by quantitative numbers, the quantification of the biological theory is necessary. This means that the theory has to be converted into quantitative relations between several processes and parameters. Examples could be the link between the wound healing rate and the cell division rate, which often changes with, for instance, the age and lifestyle of the patient. The quantification operation often leads to mathematical problems of varying complexity and nature depending on the purpose. In Figure 1, we show a scheme of gaining understanding in a quantification of wound healing research. We note that this scheme is generic in the sense that it holds for the quantification in all natural sciences and processes. Quantitative understanding of biological phenomena leads to predictive power as well as to a reduction of experiments, which can be expensive, time-consuming, dangerous, or are suffering from ethical constraints. Furthermore, quantitative insight is necessary for visualization and animations, which can be used for educational purposes to doctors (in particular useful for pharmacy), students, and patients. Computational models are a typical result from the quantification of biomedical processes. In this

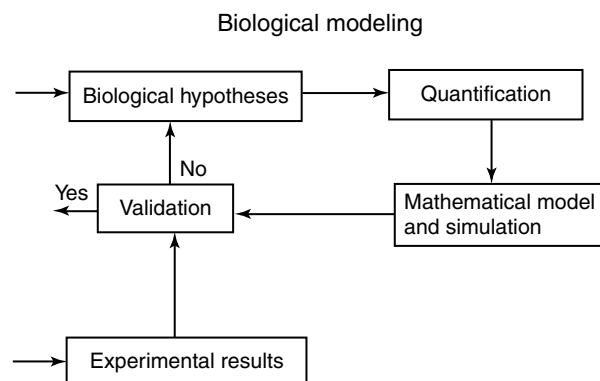


Figure 1 A schematic of modeling approaches in the life sciences.

article, we will review some computational models and explain their strengths and weaknesses.

Evaluation of Several Mathematical Approaches

There exists a large variation in the nature of mathematical models for wound healing in the literature. The first models were conceptually very simple in the sense that some of them were able to predict the condition for the onset of wound healing (or tumor growth). An example was proposed by Adam (1999) for wound healing and in earlier works for tumor growth. Adam's model was based on the assumption that wound healing is initiated once the concentration of a generic growth factor exceeds a predefined threshold value. The growth factor is secreted in a so-called active layer, which reflects a narrow region around the wound edge (typically a ring) with an increased cellular activity in terms of cell division, cell migration (also referred to as cell motility), and cytokines production. Mathematically, the model is simple but elegant, and reads as

$$\begin{aligned} \frac{\partial c}{\partial t} - D\Delta c + \lambda c &= P\chi_{\Omega_a}(x), \quad t > 0, \quad x \in \Omega, \\ c(0, x) &= 0, \quad x \in \Omega. \end{aligned} \quad (1)$$

Here c represents the generic growth factor, Ω represents the tissue region. Further, Ω_a denotes the active layer around the wound, λ models a loss term due to diffusion of the growth factor away into the tissue or due to a chemical reaction, P models the regeneration of the growth factor in the active layer Ω_a , $\Delta c = \frac{\partial^2 c}{\partial x^2} + \frac{\partial^2 c}{\partial y^2}$, and χ denotes the indicator function, satisfying

$$\chi_{\Omega_a}(x) = \begin{cases} 1, & x \in \Omega_a, \\ 0, & x \notin \Omega_a. \end{cases} \quad (2)$$

It is then postulated that wound healing is initiated at each position of the wound edge, Γ , if and only if

$$c(t, x) \geq \theta, \quad x \in \Omega_a.$$

It is well known that the solution to the above partial different converges to a steady-state equilibrium as $t \rightarrow \infty$. Therefore Adam originally considered the steady-state solution only as a predictive model to forecast whether the wound with given dimensions would start healing at all. Javierre *et al.* (2009) extended the model with a free boundary problem where the wound edge is allowed to move according to phenomenological law depending on the local growth factor concentration and the local wound edge curvature, which reads as

$$v_n = H(c - \theta) (\alpha + \beta \kappa) \quad (3)$$

where v_n denotes the normal velocity on the wound edge directed into the wound, α and β are constants and $H(\cdot)$ is a Heaviside function. The free boundary problem was solved using a level-set method. Though the model is simple and elegant from a mathematical problem, it only contains a very limited amount of biological information as it is well known that wound healing is a very complicated biological process.

In order to incorporate more biology into the model, such as mechanical factors and the interplay between several cellular phenotypes and for instance the interplay with angiogenesis as well as with the immune response system, the model is way too simple. However, on the testing of cellular monolayer cultures consisting of one cell only, healing of a gap can be simulated and readily validated using this class of models. Then, one could mimic, for instance, the closure of the epidermis that consists of keratinocytes as a thin cellular layer. Note that this model only tracks an interface between the 'wounded state' and the 'undamaged state' and hence does not involve any cellular densities, or any explicit information on the number of cells.

A next class of models treats cellular densities in terms of an averaged number of cells per unit volume or per unit area. Here one could for instance consider the interplay between a generic growth factor and the migration, proliferation, and death of a cellular phenotype. This class of models already contains some more biological information than the previous class. An example of this class is the model that was originally formulated by Sherratt and Murray (1991), which reads as

$$\begin{aligned} \frac{\partial n}{\partial t} - D\Delta n + kn &= s(c)n \left(2 - \frac{n}{n_0}\right), \\ \frac{\partial c}{\partial t} - D_c\Delta c + \lambda c &= f(n), \\ f(n) &= \begin{cases} \lambda c_0 \frac{n}{n_0} \cdot \frac{n_0^2 + \alpha^2}{n^2 + \alpha^2}, & \text{for the activator,} \\ \frac{\lambda}{n_0} \frac{c_0}{n}, & \text{for the inhibitor,} \end{cases} \\ s(c) &= k \left(\beta + \frac{2c_m(h - \beta)c}{c_m^2 + c^2} \right). \end{aligned} \quad (4)$$

Here n and c , respectively, denote the cellular density and the growth factor concentration. Migration is assumed to proceed by random walk (linear diffusion). Here α , β , λ , k , c_m , c_0 , n_0 , and h represent model constants for cellular proliferation and death as well as for the regeneration of growth factors. Note that here α and β have a different meaning than in the previous model. This model falls into the class of 'transport-reaction models.' The model contains the regeneration of growth factors by the cells, which could be epithelial cells and the influence of this growth factor on the proliferation of the epithelial cells. Although this model already falls into the class of relatively simple models, it does contain more biological information as it is able to be used for the prediction of skin quality via the quantity n , which is no longer models a binary state between 'wounded' and 'undamaged' as in the previous model. Clear advantages of this class of models is that they are still analytically tractable in a framework of stability and traveling wave analysis, which was carried out by Sherratt and Murray (1991). We remark that this model serves as an example but that many more references could be given for this class of models, see for instance (Britton and Chaplain, 1993; Olsen *et al.*, 1995; Alarcon *et al.*, 2006; Gaffney *et al.*, 2002; Maggelakis, 2003).

Later extensions of this model were constructed to combine the formalism with angiogenesis, wound contraction by, among others, Javierre *et al.* (2009). These processes lead to the incorporation of more reaction-transport equation

including all kinds of nonlinear couplings as well as to the coupling with mechanical models. The coupling with mechanical models also leads to modeling local deformations and displacements in the tissue region, which also needs an appropriate treatment of mesh displacements. Nowadays, very complicated models exist where many of the biological subprocesses have been incorporated, which hence need a large number of input (model) parameters as well as elaborate 'numerical' schemes based on 'discretization' techniques such as 'finite-element methods' and 'numerical time-integration schemes.' Advantages are that much of the important biology is simulated and that the link between (partly) overlapping biological subprocesses can be studied. On the other hand, many of the biological parameters that serve as input parameters in the models are very hard to determine or even in some cases impossible to determine. Having such an uncertainty in the parameters, it is necessary to carry out sensitivity analyses over wide ranges of parameter values. For simple models involving a few unknown parameters, this is not a hard task and the display of healing characteristic over different parameter regimes allows a visualization in terms of tractable plots of 'level-curves' of the parameter ranges. For the more complicated models, the visualization and presentation becomes more challenging and probably stochastic methods such as 'stochastic' methods could be helpful. One of the drawbacks of this class of complicated models is the uncertainty of the parameter values that reduces the predictive power of these models. A big advantage of the continuum-based models is that they can be used easily over large scales such as the tissue or organ scale. The stochastic variations may not be implemented in single runs, but if averages of the parameter values would be known, then, the model outcomes could be interpreted as average experimental outcomes.

Until now all models fell within the 'continuous-scale' formalisms, which were formulated in terms of partial differential equations. Another class of models is discrete in the sense that a domain of computation, which is a tissue, is divided into a set of 'lattice volumes or areas' that have a binary (or discrete) state. The state could reflect 'wounded' or 'undamaged' in the binary case of wound healing. This class is referred to as 'cellular automata models.' Each lattice volume is allowed to change according to 'deterministic' or 'stochastic' principles and often its change of state is determined by the state of the neighboring lattice volumes or by the change of state of the neighboring lattice volumes. One can also determine the change of state by optimality principles that are possibly based on energetic foundations. This class of modeling, referred to as the 'cellular Potts models,' has had large popularity in various branches of statistical and theoretical physics. Graner and Glazier (1992) introduced this type of modeling to simulating 'angiogenesis,' which is the regeneration of a vascular network. Merks and Koolwijk (2009) has worked on getting this modeling to higher levels regarding angiogenesis, where extensions have been made to model deformation and migration of individual endothelial cells. Since this model makes use of a lattice, it is suitable to enrich the formalism with numerical discretization schemes for the solution of partial differential equations that mimic regeneration, decay, and transport of chemokines, oxygen, and other chemicals. Disadvantages are the need of a lattice where

simulations may exhibit lattice-dependent behavior and the need of large computational resources. A three-dimensional implementation of the model still seems to be cumbersome. Although the biological processes are simulated quite well and different biological hypotheses can be set as input parameters into the model. This modeling class also contains a stochastic nature, and hence several runs are needed to predict the average behavior as well as variances needed for the evaluation of intervals of confidence. Modeling large cell colonies also poses a computational challenge whereas this is solved relatively easily by the continuum-scale models.

Other modeling approaches are based on particle methods, where individual cells are treated as spheres or circles of finite radius to deal with cell proliferation, death, differentiation, migration, etc. Here contact forces between cells that physically collide are incorporated. One often uses energetic principles where mechanical factors, as well as chemical factors are incorporated. Recent approaches were described in Groh and Louis (2010), Neilson *et al.* (2011), Byrne and Drasdo (2009), Rey and Garcia-Aznar (2013), and Vermolen and Gefen (2012a). Some of the models were constructed with the help of the experimental observations by Reinhart-King *et al.* (2008) that remote cells communicate over distances of the order of tens of micrometers by mechanical forces they inflict on the substrate or collagen. The approaches contain stochastic processes like cell death and proliferation and cell migration, where for instance diffusion is modeled by a Wiener process for the spatial position of the cells, for example, by the following 'stochastic differential equation' (Steele, 2001)

$$dx(t) = \sigma dW(t) + v(t)dt, \quad (5)$$

where $x(t)$ represents the position of a cell, $W(t)$ denotes a vector Wiener process, and $v(t)$ stands for the 'deterministic' component of the cell velocity. This deterministic cell velocity can contain all sorts of contributions such as 'haptotaxis' or 'chemotaxis.' A characteristic of this modeling class that the cellular positions are not limited to a discrete set of points in space but may migrate over a continuous (compact) domain reflecting a part of tissue. The absence of a lattice on which cells are present, allows a treatment without 'computational meshes' such as the use of 'Fundamental Solutions (Green's Functions)' (Evans, 1998), 'superpositions' over finite sets of sources for the chemokines or displacements treated in the model. Also 'integral' expressions involving 'convolutions' over Fundamental Solutions are commonly used to obtain the concentrations due to secretion over a continuous area of sources. An example of such studies can be found in Vermolen and Gefen (2012b). In other cases, one uses a 'hybride' formalism where chemical concentration, mechanical displacements, etc., are resolved using 'discretization' techniques such as the 'finite-element method.' Since this class of models can be extended with 'stochastic' principles, several runs have to be made in order to be able to draw any 'statistically' sound conclusions. This kind of models is very suitable if one is interested in collective cell behavior where one typically considers cell colonies of approximately tens of thousands of cells. To consider more cells, one will have to rely on parallel computational environments such as graphical processing units (GPU) and/or CUDA. We show an example of the outcomes of such a model in Figure 2, where snapshots at consecutive

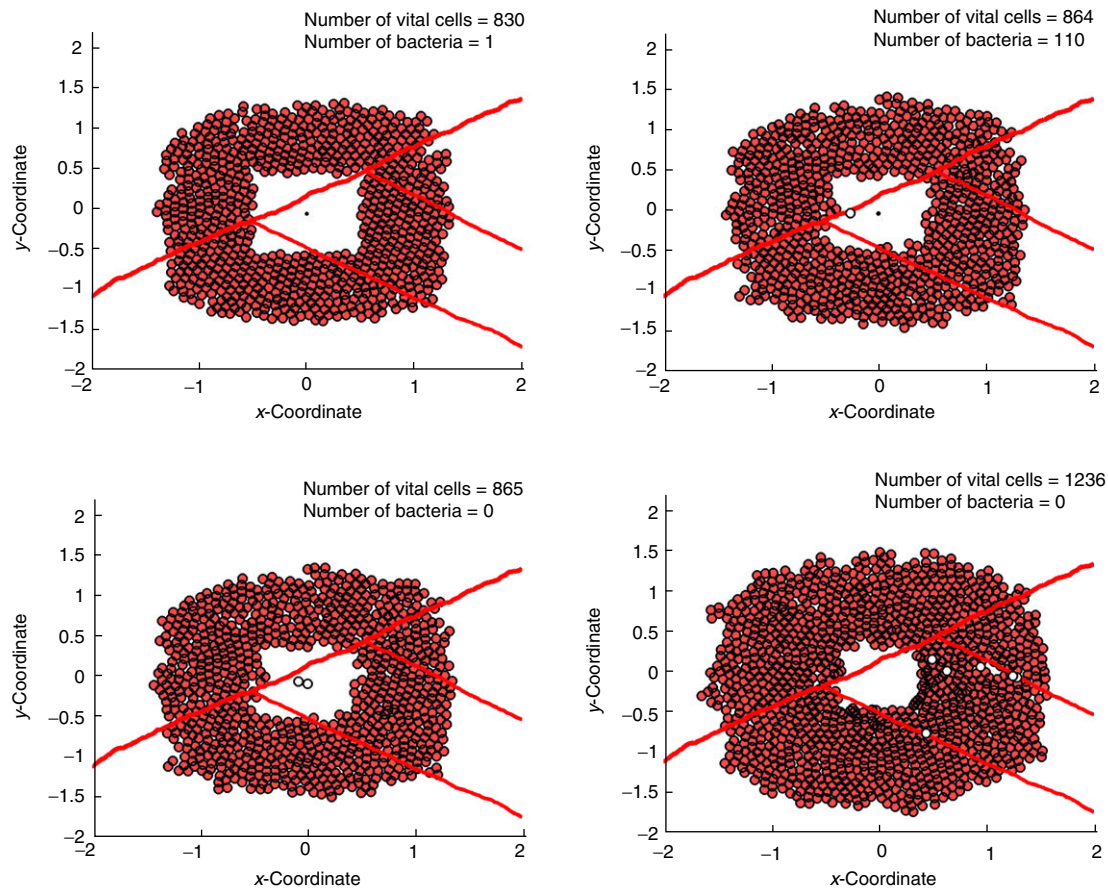


Figure 2 Some snapshots of a healing wound at times 0, 13.3, 15, and 100 h with constituent cells (red) pathogens (black dots) and immune cells (white).

times are shown during closure of a gap. The computations done with a cell-colony model mimic a monolayer cell-culture.

A subclass of this modeling work is represented by cellular models where apart from the migration, differentiation, and death of cells, one also incorporates the deformation of cells. The shape of the cell changes during migration or during differentiation to another phenotype. For this type of models, various approaches have been reported in literature. The information that one extracts from these models is very detailed concerning single-cell behavior or for the behavior of small colonies (numbers) of cells. In some cases one solves the visco-elastic equations on a moving and deforming domain using 'finite-element methods' with a moving surface. This is done by [Madzvamuse and George \(2013\)](#). This method allows the incorporation of much of the physics behind migration and cell deformation, but it also needs the solution of a three-dimensional problem with a moving boundary, which demands quite some computational power for a single cell. In other cases, one approximates the cell boundary by a 'computational' mesh and one links all neighboring points together and to the cell nucleus by applying 'spring forces,' by which a set of equations of motion for the cell nucleus and the discrete set of points on the cell boundary (cell membrane) is obtained. This system of equations is linked with the inclusion of

chemotaxis or any other type of deterministic way of motion (one can even apply this to cell migration in blood vessels, where the velocity is determined by an interaction with fluid flow). A Wiener process may be used to model the diffusional contribution to cellular migration. This approach was carried out by [Vermolen and Gefen \(2013\)](#) for chemotaxis (and extended by [Vermolen et al. \(2014\)](#)) for the case of cellular migration through small blood vessels to model the immune response system prior to wound healing. The latter approach allows a simple treatment for multiple cells but it still requires more computational resources than the previous class based on particle methods. Advantages are the enormous cellular details that one can derive from the models, and the ability to visualize elementary biological processes such as cell migration, division, differentiation, etc. The last-mentioned processes can be modeled using both deterministic, see [Prokharau et al. \(2014\)](#) for instance and (semi-)stochastic principles. Biologists are able to predict when a cell divides upon knowing the time-history and environmental situation of the cell. Since this information is not available for large colonies of cells, one has to rely on (semi-)stochastic principles to model these processes. The addition of stochastic principles can be done using Poisson processes or by incorporating Wiener (jump) processes or by adding randomness to the initial stages of the cells.

Conclusions

In this article several classes of models have been considered that can be used to mimic (partial) processes related to wound healing or even to tumor growth. We remark here that the treatment is far from complete in the sense that far more models exist in the literature on the mathematics of wound healing. However, the present article aims at being descriptive concerning all kinds of modeling approaches. The models are based on various scales, ranging from tissue-scale down to cellular scales. If one wants to look at large-scale issues related to wound healing, such as the influence of curvature on healing rate or on the velocity of the interface separating the wound from the undamaged tissue, then one is recommended to use the large, continuum-scale models. Furthermore, small-scale events as the migration of individual cells can be studied and visualized using cellular-scale models such as particle-inspired models or cellular automata models. The results of these simulations could be compared to *in vitro* and *in vivo* experiments and subsequently be used as input in the large-scale (tissue-scale) models.

Next to the spatial levels, mathematical models range in complexity. Complex models can be used to study links between subtle biological processes on several levels. However in many applications one is not interested in large detail but merely in the reproduction of trends. For these last-mentioned cases, one could rely on relatively simple models. Complicated models have the disadvantage of the need of many input parameters, which hardly ever have been measured, or which even are not measurable at all. In this sense, successful modeling of biological processes like wound healing relies on a rigorous model-reduction and this is where the most interesting modeling challenges originate from.

Another issue is the uncertainty in the measured parameters due to measuring errors or due to patient-specific values of parameters as a result of sex, pigmentation (race), lifestyle, age, and physical condition. One can develop very accurate numerical schemes to approximate the solution to the mathematical modeling problem, but the approximation is of limited use if the input parameters already contain large uncertainties.

See also: Cell Communication: Prototypic Integrative Processes: Wound Healing: An Orchestrated Process of Cell Cycle, Adhesion, and Signaling

References

Adam, J.A., 1999. A simplified model of wound healing, with particular reference to the critical size defect: One-dimensional model. *Mathematical and Computational Modeling* 30, 23–32.

- Alarcon, T., Byrne, H.M., Maini, P.K., Panovska, J., 2006. Mathematical modelling of angiogenesis and vascular adaptation. In: Paton, R., McNamara, L. (Eds.), *Multidisciplinary Approaches to the Theory of Medicine, Studies in Multidisciplinary*, vol. 3., pp. 369–387. Available at: <http://eprints.maths.ox.ac.uk/view/author/Panovska=3AJ=2E=3A=3A.html> (accessed 17.06.15).
- Britton, N.F., Chaplain, M.A.J., 1993. A qualitative analysis of some models of tissue growth. *Mathematical Biosciences* 113, 77–89.
- Byrne, H., Drasdo, D., 2009. Individual-based and continuum models of growing cell populations: A comparison. *Journal of Mathematical Biology* 58, 657–687.
- Evans, L.C., 1998. *Partial differential equations*. Graduate Studies in Mathematics, second ed., vol. 19. Providence, RI: American Mathematical Society.
- Gaffney, E.A., Pugh, K., Maini, P.K., 2002. Investigating a simple model for cutaneous wound healing angiogenesis. *Journal of Mathematical Biology* 45 (4), 337–374.
- Graner, F., Glazier, J., 1992. Simulation of biological cell sorting using a two-dimensional extended Potts model. *Physics Review Letters* 69, 2031–2036.
- Groh, A., Louis, A.K., 2010. Stochastic modeling of biased cell migration and collagen matrix modification. *Journal of Mathematical Biology* 61, 617–647.
- Javierre, E., Moreo, P., Doblare, M., Garcia-Aznar, J.M., 2009. Numerical modeling of a mechano-chemical theory for wound contraction analysis. *International Journal of Solids and Structures* 46 (20), 3597–3606.
- Javierre, E., Vermolen, F.J., Vuijk, C., van der Zwaag, S., 2009. A mathematical analysis of physiological and morphological aspects of wound closure. *Journal of Mathematical Biology* 59 (5), 605–630.
- Madzvamuse, A., George, U.Z., 2013. The moving grid finite element method applied to cell movement and deformation. *Finite Elements in Analysis & Design* 74, 76–92.
- Maggelakis, S.A., 2003. A mathematical model for tissue replacement during epidermal wound healing. *Applied Mathematical Modeling* 27 (3), 189–196.
- Merks, M.H., Koolwijk, P., 2009. Modeling morphogenesis in silico and *in vitro*: Towards quantitative, predictive, cell-based modeling. *Mathematical Modeling of Natural Phenomena* 4 (4), 149–171.
- Neilson, M.P., MacKenzie, J.A., Webb, S.D., Insall, R.H., 2011. Modeling cell movement and chemo-taxis using pseudopod-based feedback. *SIAM Journal of Scientific Computing* 33 (3), 1035–1057.
- Olsen, L., Sherratt, J.A., Maini, P.K., 1995. A mechanochemical model for adult dermal wound closure and the permanence of the contracted tissue displacement role. *Journal of Theoretical Biology* 177, 113–128.
- Prokharau, P.A., Vermolen, F.J., Garcia-Aznar, J.M., 2014. A mathematical model for cell differentiation as an evolutionary and regulated process. *Computer Methods in Biomechanics and Biomedical Engineering* 17, 1051–1070.
- Reinhart-King, C.A., Dembo, M., Hammer, D.A., 2008. Cell–cell mechanical communication through compliant substrates. *Biophysics Journal* 95, 6044–6051.
- Rey, R., Garcia-Aznar, J.M., 2013. A phenomenological approach to modeling collective cell movement in 2D. *Biomechanics and Modeling in Mechanobiology* 12 (6), 1089–1100.
- Sherratt, J.A., Murray, J.D., 1991. Mathematical analysis of a basic model for epidermal wound healing. *Journal of Mathematical Biology* 29, 389–404.
- Steele, J.M., 2001. *Stochastic Calculus and Financial Applications*. New York, NY: Springer.
- Vermolen, F.J., Gefen, A., 2012a. A semi-stochastic cell-based formalism to model the dynamics of migration of cells in colonies. *Biomechanics and Modeling in Mechanobiology* 11 (1–2), 183–195.
- Vermolen, F.J., Gefen, A., 2012b. A semi-stochastic cell-based model for in-vitro infected ‘wound’ healing through motility reduction. *Journal of Theoretical Biology* 318, 68–80.
- Vermolen, F.J., Gefen, A., 2013. A phenomenological model for chemico-mechanically induced cell shape changes during migration and cell–cell contacts. *Biomechanics and Modeling in Mechanobiology* 12 (2), 301–323.
- Vermolen, F.J., Mul, M.M., Gefen, A., 2014. Semi-stochastic cell-level computational modeling of the immune system response to bacterial infections and the effects of antibiotics. *Biomechanics and Modeling in Mechanobiology* 13 (4), 713–734.

Multiscale Analysis of Morphogenesis

CC Antonovici, Leiden University, Leiden, The Netherlands; Institute of Biology Leiden, Leiden, The Netherlands; and Centrum Wiskunde and Informatica, Amsterdam, The Netherlands

SEM Boas and EG Rens, Leiden University, Leiden, The Netherlands and Centrum Wiskunde and Informatica, Amsterdam, The Netherlands

H Tahir, VU Medical Centre, Amsterdam, The Netherlands and Centrum Wiskunde and Informatica, Amsterdam, The Netherlands

RMH Merks, Centrum Wiskunde and Informatica, Amsterdam, The Netherlands and Leiden University, Leiden, The Netherlands

© 2016 Elsevier Inc. All rights reserved.

Introduction

How does the genetic information encoded in the DNA direct the growth and form of multicellular organisms? This extremely challenging question has traditionally been addressed in developmental biology by analyzing the scales in isolation from one another. Classical embryology has looked at multicellular development from the perspective of individual cells and tissues. It showed, for example, how a specific portion of the early embryo called the ‘the organizer’ could elicit the formation of the body axis, or how the mesodermal tissue induces the ectoderm to turn into the neural plate. Molecular genetics has looked at the function of individual genes in the development. Here it was investigated how the knock-out or overexpression of individual genes affects the structure of the organism as a whole, whereas cell biology investigates the functionality of genes and proteins in individual cells. High-throughput molecular approaches, including genomics, transcriptomics, and proteomics, have made it possible to investigate the activity of all genes in the developing organism concurrently. However, a complete understanding of biological development requires us to integrate information from the molecular, cellular, and tissue level scales, monitoring events occurring at timescales ranging from nanoseconds (chemical reactions) to years (homeostasis of organisms) to millions of years (evolution). The complexity of multiscale interactions makes it almost impossible to capture biological development using experimentation alone. Recent multiscale systems biology studies that use a combination of imaging, experimentation, combined with multiscale modeling, and simulation can help provide new, surprising insights into the mechanisms of biological development.

Vertical Integration

In this article we will discuss how multiscale, computational modeling helps biologists to analyze the interactions between the molecular, cellular level, and tissue levels during biological development. These approaches make use of ‘vertical integration’: they merge mathematical and computational models formulated at different spatial and/or temporal scales to produce multiscale models.

Recent insight into the development of the horizontal stripe pattern of the zebrafish is a good illustration of vertical integration. Zebrafish with a mutation in a gene called *jaguar* have two diffuse, dorsal and ventral blue bands, instead of the usual, horizontal stripe pattern. How does a defect at the molecular level (a gene mutation) affect a phenomenon at

the level of the organism (the fish pattern)? Naively, one would perhaps expect that *jaguar* must encode a gene directly involved in skin pigmentation, for example, a pigment protein. Instead, *jaguar* codes for the ion channel Kir7.1. How can an ion channel impact the coat patterns of zebrafish? Kir7.1 is involved in the polarization of melanophores, the black pigment cells responsible for the stripes’ black color. Upon touching the xanthophores (yellow pigment cells), the melanophores depolarize and move away (Inaba *et al.*, 2012). At the multicellular level, this behavior separates melanophores and xanthophores, a process that is thought to be part of a Turing-type patterning mechanism (Nakamasu *et al.*, 2009), although this conclusion was challenged recently (Mahalwar *et al.*, 2014). Thus in the *jaguar* mutant, the loss of a crucial ion channel in melanophores changes their ‘cell behavior,’ the response to xanthophores. Eventually this changes the tissue-level patterning of zebrafish stripes.

This work on zebrafish patterns offers a relatively simple example of how genetic information regulates the organism’s phenotype, by changing the behavior of the individual cells. In the analysis of this chain of events, detailed molecular information is replaced by a course-grained description of how the genetic change affects the behavior of individual cells and their responses to one another. Mathematical descriptions at the cell population level then demonstrate how a defect in cell behavior eventually changes the cells’ collective behavior leading to a change in skin patterning. This article will describe how to apply vertical integration to the analysis of multicellular systems, using a combination of imaging and computational modeling. We will illustrate the approach based on practical examples in animal and plant developmental biology.

Analyzing Cell Behavior

Recent technical developments have made it possible to analyze the behavior of individual cells and cell populations to unprecedented quantitative detail. Systems microscopy approaches can characterize the motility and shapes of cancer cells in cell cultures under a large number of conditions (Le Dévédec *et al.*, 2010). Detailed quantitative analyses of cell migration are also performed of B-cells in germinal centers (Beltman *et al.*, 2011), of endothelial cells *in vitro* (Parsa *et al.*, 2011) and of cells in the tail bud during somitogenesis (Dray *et al.*, 2013). It has now also become possible to concurrently image the trajectories and lineage of cells in organs (Short *et al.*, 2014) and early zebrafish embryos (Olivier *et al.*, 2010). The latter line of work is now resulting in a digital gene expression atlas of early zebrafish embryogenesis

(Castro-González *et al.*, 2014). This type of work is now increasingly complementing high-throughput molecular data. It provides exactly the type of data needed to inform multicellular simulation models and to validate the model predictions. In order to annotate and efficiently store such multicellular, cell-behavioral data in a format that can directly act as direct input for multicellular simulations, several groups are developing new data standards including the Cell-behavioral Ontology (Sluka *et al.*, 2014) and MultiCellDS (see Section Relevant websites).

Alongside high-throughput imaging approaches to analyze collective cell behavior, it is also necessary to map out the reactions of individual cells in response to the stimuli the cells receive from adjacent cells and chemical signals. These stimulus-response maps then are the input for the cell-based, multicellular simulation models that the remainder of this article will describe. An example of such a description of cell behavior is the behavior of melanophores in response to contact with xanthophores described in the previous section. The descriptions can be derived from *in vivo* observations (Mahalwar *et al.*, 2014) or by isolating the cell types *in vitro* (Inaba *et al.*, 2012). Using microfluidic approaches it is also possible to quantify the chemotactic response of cells to signaling gradients (Lin and Butcher, 2006) or the pairwise interactions between cells (Yin *et al.*, 2008; reviewed in Merks and Koolwijk, 2009). A number of measures have been designed to quantify the accuracy of chemotaxis (Kay *et al.*, 2008).

Mathematical Models for Studying Collective Cell Behavior

Depending on the biological application, a range of cell-based modeling approaches are commonly used. Cell and tissue-level models are roughly categorized into continuum and single-cell-based models, for which we give a short overview. In tissue morphogenesis, also the surrounding extracellular matrix (ECM) plays a key role in regulating and coordinating cell behavior. To describe interactions between cells and the ECM a range of modeling studies have combined single-cell-based models with continuum or discrete representations of the extracellular matrix. A brief overview of such hybrid approaches will be provided here as well.

Continuum Models

A population of cells can be considered as a continuum (Murray *et al.*, 1983), and thus be modeled as a density function. In this case, the movement of cells is described by a continuity equation. In a general form, this is

$$\frac{\partial n}{\partial t} + \nabla \cdot \vec{J} = \sigma \quad [1]$$

where n is the cell density, $\vec{J}$ is the flux of cells and σ describes the emergence or removal of cells in the tissue. Typically, diffusion ($\vec{J} = D\nabla n$) and migration of cells ($\vec{J} = \vec{v}n$) is incorporated. The partial differential equation (PDE) can be extended to account for cell behavior. Cellular birth and death events, such as apoptosis and mitosis, are generally included in

the σ term. Phenomena that influence cell migration can be described in terms of the flux $\vec{J}$. For instance, Namy *et al.* (2004) included haptotaxis, Byrne and Drasdo (2009) simulated movement down pressure gradients and Manoussaki *et al.* (1996) implemented mechanotaxis.

Discrete Models

Continuum descriptions of cell populations can become limiting when the models must be validated against individual cell trajectories, or in case the structures of interest form at a length scale of only a couple to tens of cells. Discrete, cell-based models describe cells as individual elements, and can be very useful in such situations. Single-particle models represent cells as point particles or ellipsoids. Multiple-particle-based models use a collection of particles to represent each cell, allowing for a more detailed description of cell shape (Merks, 2013). A further distinction is made between representations on regular lattices, which can be computationally more efficient, and off-lattice representations. We present an overview of cell-based methodology in Figure 1.

Particle-based models

In single-particle-based off-lattice models, cells are typically represented as point or ellipsoid particles. When using a lattice, cells are represented as one lattice site. Not only does the way cells are represented vary among models, but the way movement is simulated differs as well. Differential equations of motion for each particle can describe cell migration. These typically consist of a cell motility term to describe active, random, or directed cell motion, and an external force consisting of a sum over the forces exerted by all the other cells and the environment (see, e.g., Grégoire *et al.*, 2003). Alternatively, particles comply to a set of behavioral or migratory rules (see, e.g., Byrne and Drasdo, 2009). Using particle-based methods, various cell migration behaviors have been modeled. For instance, Szabo *et al.* (2007) included diffusion, directional persistence, and an attraction to anisotropic structures to model cell organization into network-like patterns. Cell-cell interactions, such as pressure and velocity adaption to neighbors, were added by Sepúlveda *et al.* (2013) and Byrne and Drasdo (2009) respectively. In Dallon (2000) haptokinesis and chemotaxis were simulated; and a particle-based model of contact guidance was presented in Dallon *et al.* (1999).

While these particle-based methods make it possible to investigate how interactions between spherical or ellipsoid cells can give rise to cell motility during embryogenesis, they cannot account for cell shape changes, which are key to many developmental processes. To account for differential cell shape, multiple particle methods describe cells as a collection of connected particles, defining either the boundary or interior of the cells. Thus, in these methods cell shape is explicit and can change actively during simulations or passively in response to interactions with adjacent cells. The subcellular elements method (Newman, 2005) is an example of an off-lattice multiple particle method, where cells are divided into subcellular elements (for instance, points in space), which can locally interact with elements of the same cell and elements of other cells via equations of motion. These equations describe

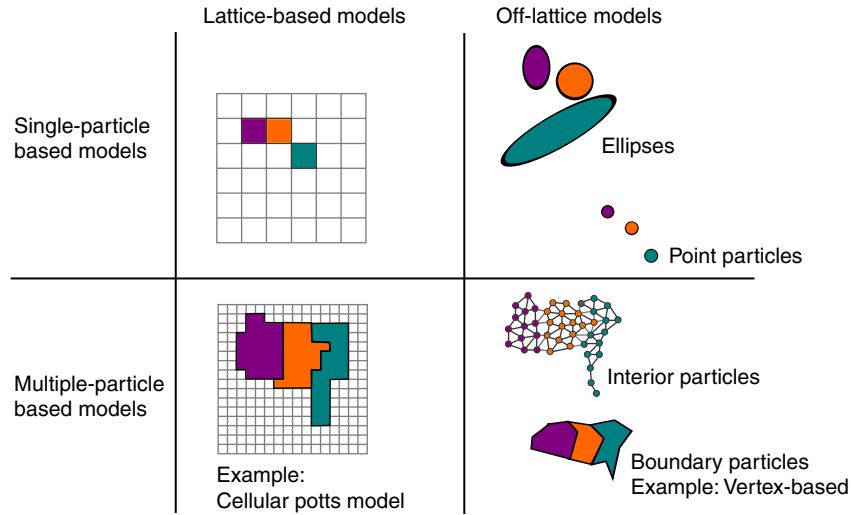


Figure 1 An overview of cell representations in cell-based simulation methodology. Modified from Merks, R.M.H., 2013. Cell-based modeling. In: Enquist, B. (ed.), Encyclopedia of Applied and Computational Mathematics. Springer. Available at: <https://www.springerreference.com> (accessed 02.06.15).

both the internal rheology of individual cells, as well as the adhesive and repulsive forces with adjacent cells. Mechanotactic and chemotactic cell migration can also be included (see, e.g., Sandersius and Chuai, 2011). Alternative off-lattice and lattice-based multiparticle methods are the vertex-based models and the Cellular Potts model (CPM), which will be described in more detail in the next sections.

Vertex-based models

Vertex-based models describe tissues as a network of polygonal cells, which are characterized by the position of vertices v_i , with $i=1, \dots, N_V$, where N_V represents the total number of vertices, together with a set of N_E edges e_{ij} which connect vertices i and j . Each polygon corresponds to a cell indexed by $\alpha=1, \dots, N$, where N is the total number of cells in the system. Vertices are displaced depending on a Hamiltonian (H) energy function that describes the force balance of the system, or using explicit equations of motion.

Jülicher and coworkers use the Hamiltonian (Farhadifar et al., 2007),

$$E = \sum_{\alpha} \frac{K_{\alpha}}{2} \left(A_{\alpha} - A_{\alpha}^{(0)} \right)^2 + \sum_j \Lambda_{ij} l_{ij} + \sum_{\alpha} \frac{\Gamma_{\alpha}}{2} L_{\alpha}^2 \quad [2]$$

Here, the first term is a cellular volume conservation term, where K_{α} is an elastic coefficient, A_{α} is the area of cell α and $A_{\alpha}^{(0)}$ is the resting area of the cell. The second term represents cell membrane tensions, Λ_{ij} , at the junctions between cells, with l_{ij} denoting the length of edge e_{ij} . The third term describes the contractility of cell perimeter L_{α} by a factor Γ_{α} . Additional rules allow adjacent cells to detach from one another, a so-called T1-transition. The local minima of the energy function determine the tissue configurations, which the authors used to identify cell packing geometries of the epithelium of the fruit fly *Drosophila* (Farhadifar et al., 2007).

Merks et al. (2011) presented a closely related vertex-based approach to analyze plant morphogenesis. This method typically used additional nodes to allow the cell-cell interfaces to

bend, while T1-transitions are suppressed to more closely mimic plant tissues in which cells cannot move relative to one another, and allows cell walls to move according to turgor-driven cell expansion laws: when there is an intracellular turgor pressure applied uniformly on the cell walls, the elastic properties of the walls will counteract it. These mechanics are described by the following Hamiltonian,

$$H = \lambda_A \sum_{\alpha} \left(A_{\alpha} - A_{\alpha}^{(0)} \right)^2 + \lambda_M \sum_j \left(l(j) - L_T(j) \right)^2 \quad [3]$$

where A_{α} and $A_{\alpha}^{(0)}$ have the same meaning as in the Jülicher model, $l(j)$ represents the length of wall j , $L_T(j)$ is the target length for wall j in the absence of turgor pressure, λ_A is a parameter for the cells' resistance to compression or expansion, and λ_M is a spring constant. The Metropolis algorithm is used to minimize H . Here, a node of the network is selected randomly and then moved in a random direction by a set distance. The algorithm always accepts the movement when the energy difference ΔH given by this movement reduces the energy of the system. Otherwise, the movement is accepted with a Boltzmann probability:

$$P(\Delta H) = \begin{cases} 1 & \text{if } \Delta H < 0 \\ e^{-\Delta H/T} & \text{if } \Delta H \geq 0 \end{cases} \quad [4]$$

where T describes the amount of randomness in the system. Using this model, one can also incorporate intracellular reactions and transport of chemicals by formulating them in a PDE and solving it after each Metropolis cycle. This model is implemented in the open source package VirtualLeaf (see section Relevant Websites).

The vertex-based models presented above only have two-dimensional (2D) implementations, but recent breakthroughs consider three-dimensional (3D) tessellations. A base model for the 3D approaches is the one developed by Honda et al. (2004) who express a cell as a polyhedron. Network dynamics

is modeled by prescribing an equation of motion to each vertex i

$$\eta \frac{d\mathbf{r}_i}{dt} = - \frac{\partial U}{\partial \mathbf{r}_i} \quad [5]$$

where $\mathbf{r}_i$ is a 3D position vector of vertex i , η is a viscosity coefficient and U is a potential force, which can be related to cell surface area, cell volume, or external forces. Computer simulations of this model showed that in the absence of external forces, cell aggregates become spherical to minimize individual cell surface areas, whereas under a continuous external force, the cells slowly rearrange themselves to recover their original shape while the aggregate remains flat. This model was recently extended in [Okuda et al. \(2013a\)](#) to include cell division, of which the direction was shown to influence tissue morphology ([Okuda et al., 2013b](#)). Related vertex-based, 3D models have been presented for simulations of cell sorting ([Hutson et al., 2008](#)) and for plant-tissue modeling ([Boudon et al., 2015](#)).

CPM

The CPM ([Glazier and Graner, 1993](#)) represents cells as a set of connected lattice sites. Each lattice site, $\vec{x}$, has a state $\sigma(\vec{x}) \in \{0, 1, \dots, n\}$ that identifies the individual cell the lattice site belongs to. $\sigma=0$ represents the surrounding medium, and n is the total number of cells in the lattice. Forces acting on the cells are described in the Hamiltonian,

$$H = \sum_{(\vec{x}, \vec{x}')} J(\sigma(\vec{x}), \sigma(\vec{x}')) (1 - \delta(\sigma(\vec{x}), \sigma(\vec{x}'))) + \lambda \sum_{\sigma \in [1, n]} (a_\sigma - A_\sigma)^2 \quad [6]$$

where J is an adhesive energy between cells and δ is the Kronecker delta function ($\delta(x, y) = 1$ if $x = y$ and $\delta(x, y) = 0$ otherwise) such that the first term counts the total adhesive energy across cell-cell and cell-medium interfaces. The second term is a cellular volume conservation term, with a_σ the cell area, A_σ the resting area of cell σ , and λ a compressibility parameter.

To mimic cell motility and membrane fluctuations, the CPM iteratively attempts to copy the state $\sigma(\vec{x})$ of a randomly selected lattice site $\vec{x}$, into a randomly selected, adjacent lattice site $\vec{x}'$; if such a copy reduces the value of the Hamiltonian ($\Delta H \leq 0$), it is accepted. If the attempt would increase the value of the Hamiltonian ($\Delta H > 0$) it is accepted with a Boltzmann probability, $P(\Delta H) = \exp(-\Delta H/T)$. These copy steps account for the intrinsic random motility of cells, with large values of T corresponding with more random motility. During one time step, N such copy attempts are made, with N the total number of lattice sites in the lattice.

The CPM has been extended with many algorithms accounting for additional cell behaviors, which are typically described by additional terms in the Hamiltonian (eqn. [6]). Of particular interest are the hybrid CPMs, in which the CPM is combined with discrete continuum models (PDEs) to account for secreted chemical signals ([Savill and Hogeweg, 1997](#)) or an elastic substrate ([van Oers et al., 2014](#)). Other extensions include anisotropic cell adhesion ([Zajac et al., 2003](#)), cell elongation ([Merks et al., 2006](#)), and persistent cell motility ([Turner and Sherratt, 2002](#)). The open source software

package 'CompuCell3D' ([Swat et al., 2012](#)) provides a user-friendly interface to the CPM and acts as a community-driven portal for the exchange of published models and extensions.

Modeling Cell-ECM Interactions

The cell-based models described above focus primarily on describing collective cell behavior. Cells in actual tissues are surrounded by the ECM, consisting of extracellular molecules and fibers, to which cells adhere. Besides functioning as a support for cells, the ECM also guides cell migration. Cells, for instance, move up ECM density gradients (haptotaxis) and chemoattractant gradients (chemotaxis). Moreover, cells have been shown to migrate toward stiffer areas (durotaxis) and along fibers (contact guidance). Not only do cells respond to the structure of the ECM, but they also actively change and remodel its local architecture. Cells are for instance able to secrete ECM-degrading enzymes. By applying a force on the ECM, cells can locally stiffen the ECM and reorient fibers. Thus, tissue morphogenesis is highly dependent on the active interactions between cells and the ECM. It is important, therefore, to couple cell models with models describing the ECM. Similar to cells, depending on the context and objective, the ECM can be modeled as a continuum or by using discrete elements. Cell-ECM interactions can be simulated by coupling the PDEs that describe cells and the ECM. For instance, both [Manoussaki et al. \(1996\)](#) and [Namy et al. \(2004\)](#) used this to study mechanical cell-ECM interactions during angiogenesis. In the CPM, ECM materials can be represented by generalized cells, using additional values of σ ([Dias et al., 2014](#); [Tahir et al., 2015](#)). To model the ECM in more detail, for example, taking into account the density, structure, or mechanics of the ECM, hybrids between cell-based models, and continuum models for the ECM are often used. Examples include the work by [Das et al. \(2010\)](#), [Bauer et al. \(2009\)](#), and [Daub and Merks \(2013\)](#), who coupled a cell-based model with PDE models of the extracellular environment, to study how chemical remodeling by the endothelial cells can facilitate sprout splitting. Interactions between cells and fibers in the context of wound healing was studied using a coupled cell-based model with a continuous model for the fibers by [Dallon \(2000\)](#) and [Cumming et al. \(2010\)](#). To model the mechanics of the ECM, [Bischofs and Schwarz \(2003\)](#) studied the orientation of cells on a deformable substrate. [Checa et al. \(2014\)](#) and [van Oers et al. \(2014\)](#) introduced hybrid finite-element and cell-based models to simulate mechanical communication of cells on deformable substrates. In these models the cells deform the substrate, and the cells respond to local deformations of the substrate. The resulting feedback loop suffices for co-alignment of cells and pattern formation, for example, networks and angiogenic sprouts. As the ECM has a fibrous nature, it can also be modeled in a discrete fashion and be combined with a cell-based model, which was done by [Schlüter et al. \(2012\)](#) and [Reinhardt and Gooch \(2014\)](#) to study the migration of cells on a compliant substrate.

Examples of Vertical Integration Studies

Collective cell behavior is a key factor in many developmental processes. In this section we describe how cell-based modeling

has been used in combination with knowledge from experiments to elucidate which cell behaviors are required to form a specific cell pattern or tissue.

Differential-Adhesion Driven Cell Sorting

Townes and Holtfreter (1955) observed that if embryonic cell types are dissociated, mixed, and reaggregated *in vitro*, they segregate into two homogeneous clusters with one cell type covering the other with relative, anatomically correct positions. To explain this phenomenon, Steinberg (1963) proposed the differential adhesion hypothesis, which states that differences between intercellular adhesiveness may drive cell sorting, such that cells with weak intercellular adhesiveness separate from one another and cells with stronger adhesion tend to stay together. Graner and Glazier (1992) and Glazier and Graner (1993) introduced the CPM to explore this hypothesis computationally (see Section CPM). Depending on the relative intercellular adhesions strengths, the CPM predicts that differential adhesion can give rise to different sorting patterns. Two mixed cell types segregate into clusters if they adhere more strongly to their own type (homotypic adhesion) than to the other type (heterotypic adhesion). Checkerboard patterns form if the heterotypic adhesion is stronger than the homotypic adhesion. The observation of Townes and Holtfreter (1955) is reproduced if cells with relatively weak homotypic adhesion (but stronger than the heterotypic

adhesion) are mixed with a second cell type of stronger homotypic adhesion (Graner and Glazier, 1992; Glazier and Graner, 1993). The cell type with strong homotypic adhesion then moves to the center of the cluster, where it can establish most connections with its own cell type.

The classic Townes and Holtfreter (1955) experiments were reexamined using quantitative methods by Krieg *et al.* (2008) who observed the segregation of zebrafish endoderm progenitor cells, mesoderm progenitor cells, and ectodermic progenitor cells (Figures 2(a) and 2(b)). Atomic force microscopy yielded measures of the relative strengths of homotypic and heterotypic adhesion of each of the three cell types. All heterotypic adhesion strengths were roughly equal and equivalent to the homotypic adhesion of the ectoderm; the endoderm adhered to itself more strongly than to the ectoderm, and the mesoderm adhered most strongly to itself among the three progenitor cell types. Interestingly, with these relative adhesion strengths, the cells sorted the wrong way out in the CPM, with the mesodermal cells moving to the center of the aggregate instead of the periphery as in the experiment. The authors therefore examined a second source of interfacial tension in the CPM, actin-mediated cortical tension. Ectoderm progenitors had the highest actomyosin-dependent cell cortex tension, followed by mesoderm and then endoderm progenitors. After including this source of interfacial tension the CPM simulations matched the experiments (Figure 2(c)) suggesting that a combination of differential cell adhesion and

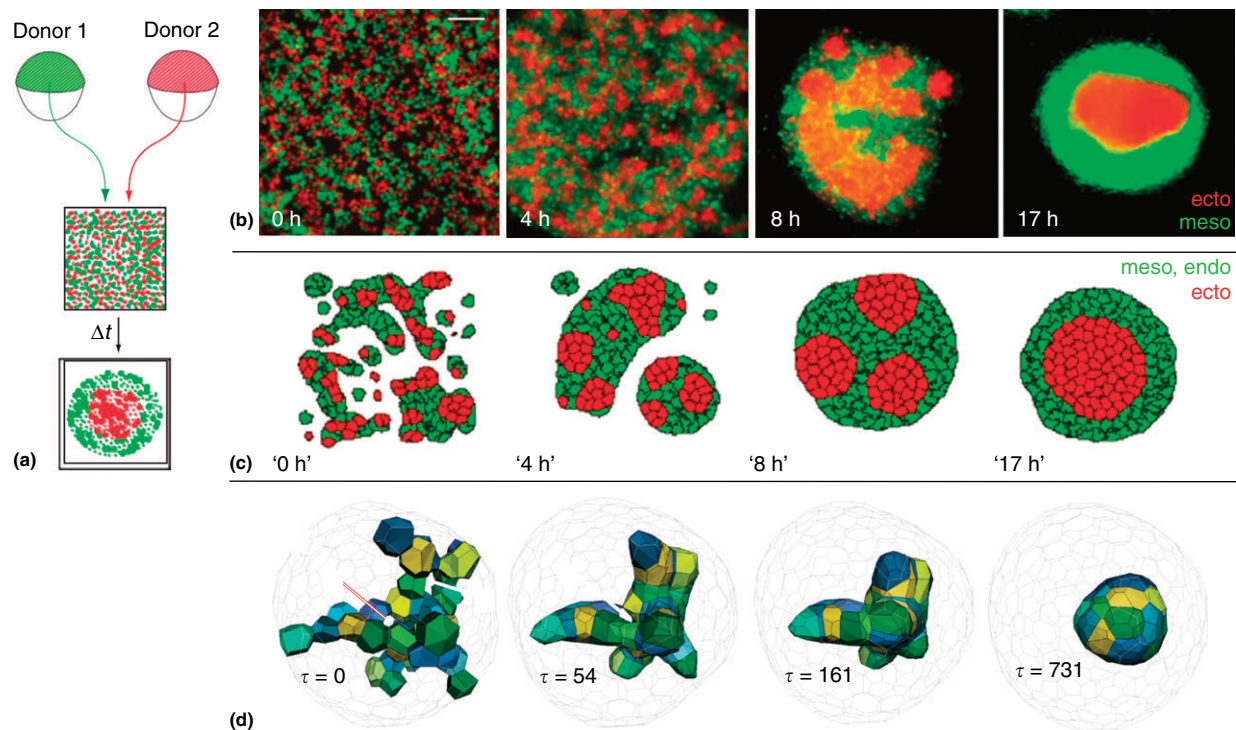


Figure 2 Experimental and computational models of cell sorting. (a) Schematic drawing of an *in vitro* cell sorting assay, during which two types of progenitor cells are mixed in a hanging drop and subsequently sort out (Krieg *et al.*, 2008). (b) Time-lapse of the sorting of ectodermic and mesodermic progenitor cells *in vitro* (Krieg *et al.*, 2008). (c) Time-lapse of the sorting of ectodermic and mesodermic progenitor cells *in silico*, using a CPM-based model (Krieg *et al.*, 2008) in which cell sorting is driven by a combination of differential cell adhesion and differential cortex tension. (d) Time-lapse of the sorting of two cell types in an interfacial tension driven model based on the finite-element method. The outlines of the cells of the cell type that is in majority are visualized in light gray (Hutson *et al.*, 2008).

differential cortex tension is required to drive cell sorting. Indeed, disruption of the actomyosin-dependent cell cortex tension prohibited cell sorting. Random membrane fluctuations are required in the CPM for sorting, although a tension driven model based on the finite-element method introduced by Hutson *et al.* (2008) (Figure 2(d)) showed that these fluctuations are not required for topological untangling of cell types during sorting.

Vasculogenesis and Angiogenesis

Vascular networks are formed *de novo* from initially dispersed cells during embryogenesis, a process called vasculogenesis. Postnatally, blood vessels can branch or split (angiogenesis) to reach oxygen-deprived tissues, such as wounds and growing tumors. During angiogenesis and vasculogenesis, endothelial cells coordinate their behavior via chemical or mechanical signals and thereby collectively form multicellular blood vessel structures. Here we discuss computational models of angiogenesis, with a focus on cell-based models that study the

collective behavior of endothelial cells driving the formation of vascular networks (Figure 3(a)) and vascular sprouts.

Serini *et al.* (2003) proposed a continuum model for vascular network formation. In their model, endothelial cells attract one another by means of a secreted chemical signal, for example, vascular endothelial growth factor (VEGF). Based on this mechanism Merks *et al.* (2008) developed a CPM-based model. Surprisingly, although on short timescales the behavior of the model roughly resembled the continuum models, on a longer time scale the network-like pattern collapsed and the cells aggregated into spheroids. Merks *et al.* (2006) found that this chemotaxis-based model forms network-like patterns if cells assume an elongated shape. Palm and Merks (2013) later showed that chemotaxis is not necessarily required: a short-range adhesive force suffice for aligning elongated cells into branches. These connect to each other, thereby forming a network-like pattern. Szabó *et al.* (2008, 2007) and Szabó and Czirók (2010) showed that networks formed when cells preferentially adhere to elongated cells.

An alternative mechanism for network formation occurs if chemotaxis is inhibited at cell–cell interfaces (Figure 3(b)),

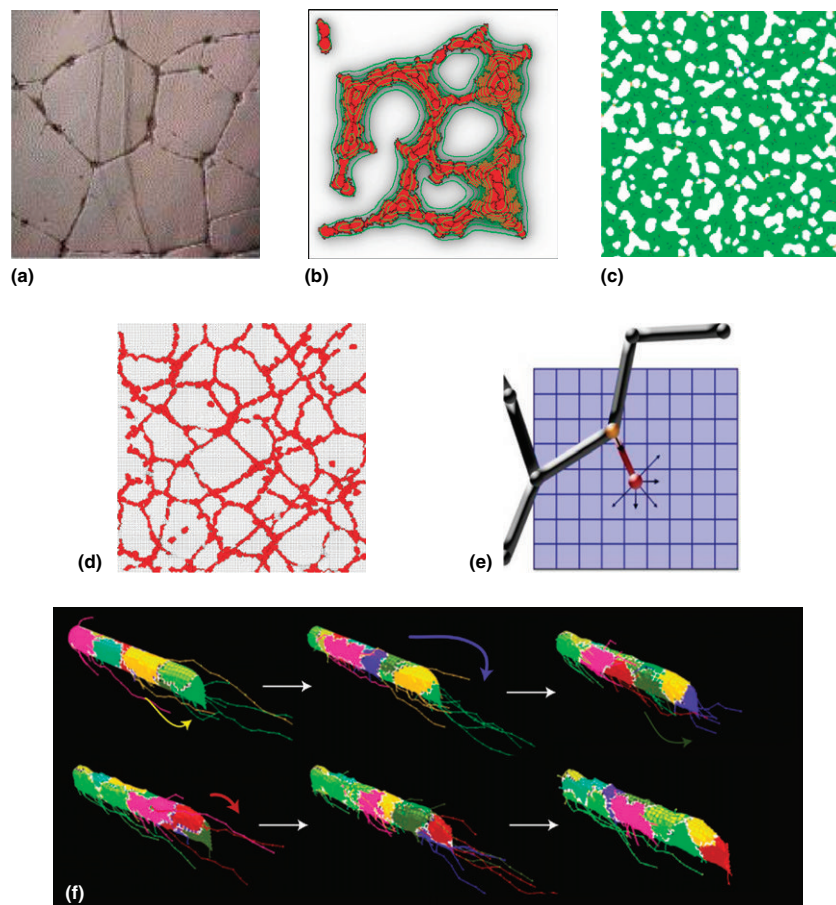


Figure 3 Experimental and computational models of blood vessel development. (a) Vascular network structure formed *in vitro* by human umbilical vein, endothelial cells (HUVEC). (b) Vascular network formation in the contact-inhibited, autocrine chemotaxis model (Merks *et al.*, 2008). (c) Vascular network structure formed by creation of a chemotactic VEGF gradient around the cells by secretion of VEGF-binding molecules (Köhn-Luque *et al.*, 2011). (d) Cells create mechanical strains in the extracellular matrix and form vascular networks by responding to the strains (van Oers *et al.*, 2014). (e) Angiogenic sprouting in a multiscale model, in which a fixed tip cell leads a sprout (Qutub *et al.*, 2009). The red node represents the migrating tip cell and black segments and nodes the quiescent vessel. (f) Cells compete for the tip cell position during angiogenic sprouting (Bentley *et al.*, 2014).

based on the experimental finding that bound VE-cadherin (an endothelial cell–cell adhesion molecule) modifies the response to VEGF-signaling (Dejana, 2004). These models, however, require rather short-range morphogen gradients, which are unlikely to occur based on measured diffusion coefficients of VEGF. Köhn-Luque *et al.* (2011) hypothesized, therefore, that cells do not secrete VEGF themselves, but produce extracellular components to bind externally present VEGF. In line with this hypothesis, fluorescently labeled VEGF accumulated in pericellular areas and colocalized with VEGF-binding molecules in an *in vitro* angiogenesis assay (Köhn-Luque *et al.*, 2013). Computational modeling showed network formation when cells chemotact toward ECM-bound VEGF (Köhn-Luque *et al.*, 2011; Figure 3(c)).

Besides chemical signals, cells can also respond to mechanical signals transferred through the surrounding ECM. Experimentally it was shown that cells respond to the strain in the matrix by orienting in the direction of stretch (Haston *et al.*, 1983) and that cells can also generate strain in the matrix themselves (Harris *et al.*, 1980; Winer *et al.*, 2009). Manoussaki *et al.* (1996) proposed a continuum, mechanical model of vascular network formation in which endothelial cells pull the ECM toward them along with the cells attached to it. Van Oers *et al.* (2014) proposed an alternative mechanical model based on a hybrid Cellular Potts and finite-element model. Here the CPM represents the cells and the finite-element model represents strains in the ECM. In accordance with experimental observations, in this model the strains in the matrix that one cell generates locally directs the motility of adjacent cells (see also Section Modeling Cell–ECM Interactions) adjacent cells. This mechanism suffices for network formation (Figure 3(d)) and sprouting.

During angiogenesis, endothelial cells differentiate into tip cells and stalk cells through lateral inhibition mediated by Delta-Notch signaling (Gerhardt, 2008). Tip cells are equipped with long filopodia to sense the local environment and guide the stalk cells along the sprout. A number of models have assumed that the tip cell is fixed and that it sets the path of the growing sprout (Milde *et al.*, 2008; Bauer *et al.*, 2007; Qutub *et al.*, 2009). For example, Qutub *et al.* (2009) developed a multiscale model in which cells are represented by nodes and segments. The node at the tip of the sprout represents the tip cell, which senses the environment to guide the growth direction of the sprout (Figure 3(e)).

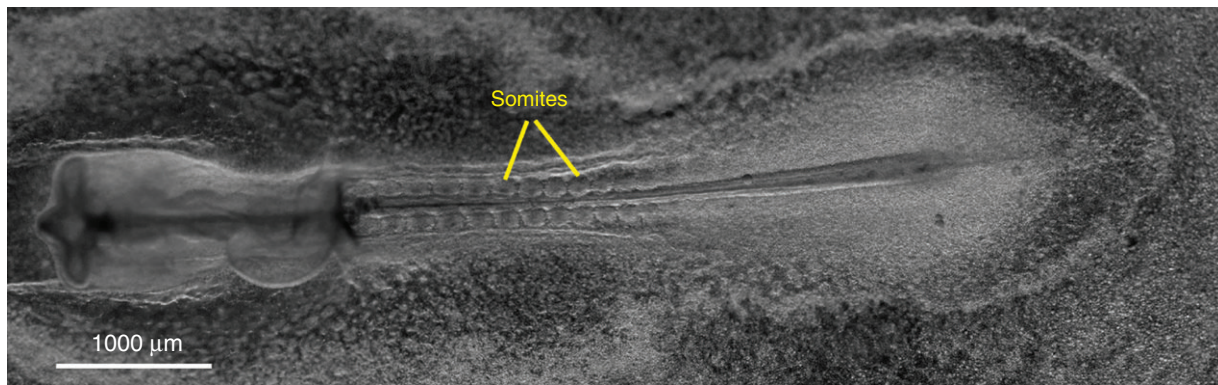
More recently it was shown that cells continuously compete for the sprout tip position (Jakobsson *et al.*, 2010; Arima *et al.*, 2011), in line with the view that sprouting is a result of collective cell behavior. Bentley *et al.* (2014) used cell-based modeling to study the mechanisms of this tip cell competition. They modeled a preformed sprout covered by cells. Cells adhere to one another, crawl past each other and extend filopodia to sense external VEGF signals (Figure 3(f)). A model of Delta-Notch signaling included in each of the endothelial cells repetitively switched fate between tip and stalk cells. They showed that tip cell competition as experimentally observed by Jakobsson *et al.* (2010) was best mimicked if tip cells have less cell–cell adhesion than stalk cells and if tip cells extend more polarized myosin protrusions in the direction of the sprout tip than stalk cells.

Multiscale Modeling of Somitogenesis

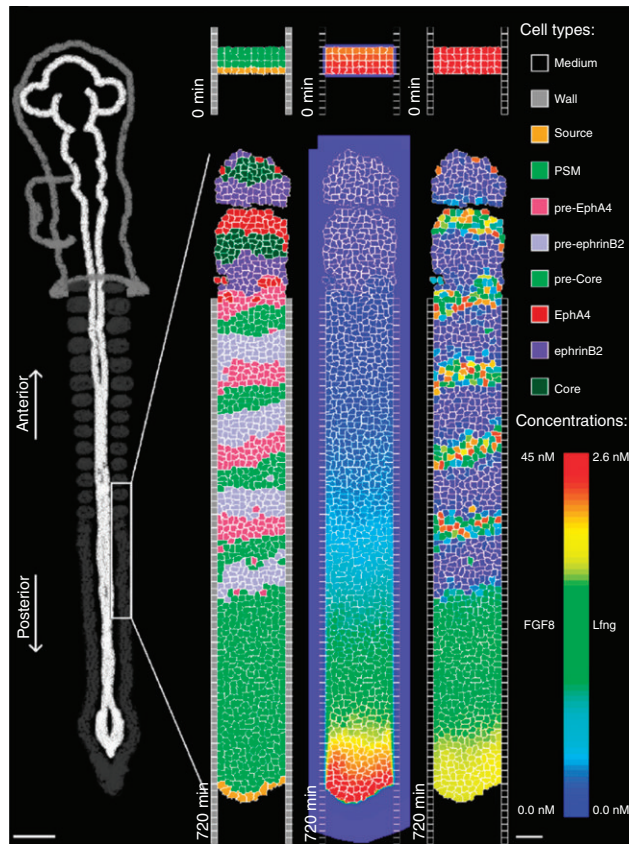
A beautiful example problem that has been analyzed using vertical integration is somitogenesis, an early process that is key to vertebrate development (Figure 4(a)). Somites are composed of tightly bound aggregates of cells that are periodically separated from a continuous presomitic mesoderm (PSM) on each side of the neural tube (Hubaud and Pourqui, 2014). Later in development, the somites give rise to the vertebra, the skeletal muscles, and other associated structures (Baker *et al.*, 2008; Agero *et al.*, 2010). It is still not clear how a homogeneous PSM can convert into the regularly arranged periodic pattern of somites (Kondo, 2014) and it remains puzzling how differences in the number and size of the somites among different species occur. An attempt to understand this is by looking into how gene expression alters the behavior of cells and determine outcoming tissue architectures (Krieg, 2009; Schier and Talbot, 2005; Tepass *et al.*, 2002).

A periodic array of somites was originally proposed by Cooke and Zeeman (1976) to form by a clock and wavefront model that describes the presence of an oscillating gene expression (the clock) and its interaction with a posteriorly moving wave of extracellular signal (the wavefront). Although several theoretical models have been proposed to explain somitogenesis including the cell cycle model (Collier *et al.*, 2000; Stern *et al.*, 1988), the reaction diffusion model (Meinhardt, 1986), and the clock and induction model (Schnell and Maini, 2000), the clock and wavefront model so far has been the most widely accepted theory (Kondo, 2014; Grima and Schnell, 2007). A range of mathematical models following the clock and wavefront model cover one or more aspects of somitogenesis. Baker *et al.* (2006, 2008) have reviewed and implemented multiple one-dimensional (1D PDE) models for the segmentation clock, fibroblast growth factor 8 (FGF8) gradient, and cell adhesion. Recently, mathematical models have been developed and applied to elucidate the importance of Notch and Wnt signaling pathways and their crosstalk as an essential molecular mechanism during somitogenesis (Wang *et al.*, 2013). Hester *et al.* (2011) developed a CPM-based 2D multiscale integrated clock and wavefront model, including intercellular segmentation clock via Delta/Notch signaling, FGF determination front and differential cell–cell adhesion driven cell sorting. In this work, existing submodels of somitogenesis were combined to test if they could reproduce key dynamical and morphological features of somitogenesis (Figure 4(b)).

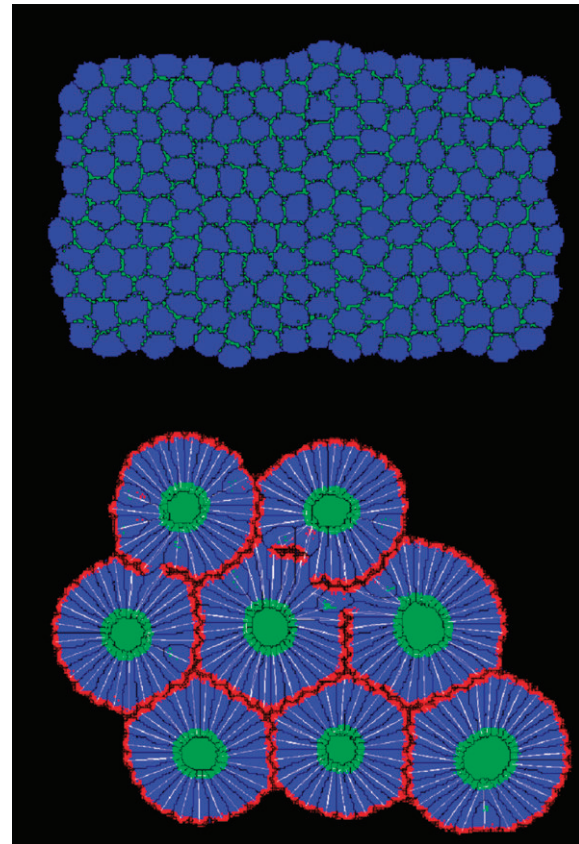
Recently, the role of a clock in the somitogenesis has been challenged. Non-somite mesoderm dorsalized with Noggin treatment can form ectopic periodic somite-like structures (Dias *et al.*, 2014). In a computational model based on the CPM, exposure to Noggin was assumed to initiate differentiation of polarized epithelial cells, which can gradually reorient into spherical structure in a differential-adhesion-driven cell sorting-like mechanism. This study suggests that somites are self-organizing structures that can form without a clock (Figure 4(c); Dias *et al.*, 2014). If and how the self-organizing properties of somites are guided or fine-tuned by the genetic oscillator in the PSM is a topic of ongoing research.



(a)



(b)



(c)

Figure 4 Multiscale analysis of somitogenesis. (a) Darkfield shot of a chicken embryo at Hamburger Hamilton Stage (HH) 11, showing somites (Courtesy M. Schmitz and B. Nelemans, VUmc Amsterdam); (b) Multiscale, clock and wavefront type model of somitogenesis (Hester *et al.*, 2011); (c) Simulation of self-organized, ectopic somite formation starting from an initially uniform distribution of mesenchymal cells (Dias *et al.*, 2014).

Regulation of Root Growth by Auxin in *Arabidopsis thaliana*

Auxin (indole-3-acetic acid or IAA) is a phytohormone that has been shown to control cell identity and cell expansion, and is major player in the regulation of root growth. Auxin diffuses freely inside the cells and cell walls and is thought to be actively transported between cells by membrane-bound specialized proteins from the PIN family. Figure 5(a) shows the positioning of the PIN efflux proteins alongside other efflux carriers like the ABCB proteins as well as influx AUX1/LAX

proteins. The polar localization of PIN-proteins leads to local differences in auxin concentrations in plant tissues, which control the cell growth and the frequency and orientation of cell divisions.

Cell-based models have been particularly instrumental in analyzing the role of auxin in root growth. Model simulations by Grieneisen *et al.* (2007) have shown how the patterned orientation of PINs in the root tip produce a flux pattern in the form of a reverse fountain and an accumulation of auxin at the quiescent center, a structure right above the root cap.

To understand the relation between this auxin dynamics and zonation in growing roots, the authors include the observation that higher auxin levels promote cell divisions, while lower levels were linked to cell elongation. Their simulations showed dynamical transitions between the elongation and meristematic zones qualitatively similar to what is observed *in vivo*, but with a highly stable auxin distribution around the quiescent center of the root. Thus, the PIN distribution may account for the formation of an auxin maximum as well as for changes in the zonation of the root.

Taking this idea a step further, Band *et al.* (2014) showed that the PIN efflux carriers alone cannot create the pattern of

auxin distribution at the root tip. Using a 2D vertex-based model with cell shapes segmented from DII-VENUS confocal microscopy images, they showed that influx carriers (such as the AUX1/LAX membrane proteins) are also required in controlling this distribution. Panel (c) of Figure 5 shows side by side the predicted auxin levels first in the case where the AUX1/LAX proteins are being knocked out and only PIN carriers ensure auxin dynamics, while the right figure shows the auxin levels when both PIN and AUX1/LAX control the active transport. Knowledge of the relationship between actual auxin levels and DII-VENUS sensor let the authors confirm their hypothesis that AUX1/LAX is essential in replicating realistic

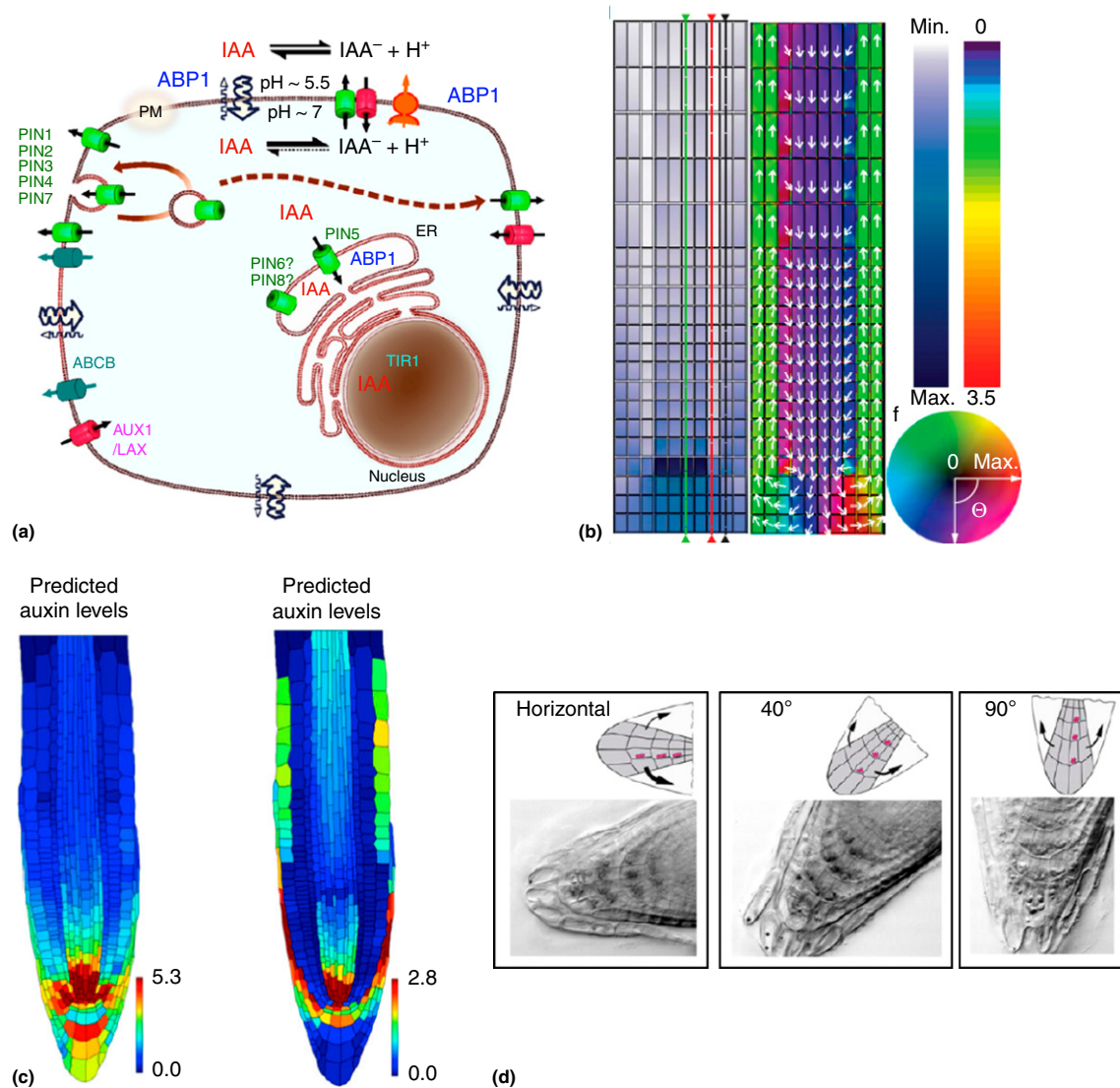


Figure 5 Vertical integration applied to root development. Panel (a) represents a schematic extracted from Friml and Jones (2010) showing the molecular distribution of auxin transporters within a cell. Of particular interest are the PIN family efflux proteins and the AUX1/LAX influx proteins. Panel (b) presents the simulation results obtained by Grieneisen *et al.* (2007) while hypothesizing that PIN transporters are enough to create an auxin maximum in the root. The left figure represents the steady-state concentrations and one can see the arising maximum above the root cap. The right figure shows the auxin fluxes measured in $\mu\text{m}^{-1}\text{s}^{-1}$ which form the flux-reflux pattern. Panel (c) showcases the difference in auxin levels which arise after knocking out AUX1/LAX influx transporters (Band *et al.*, 2014). The left figure considers the auxin levels given solely by efflux PIN influence, while the right figure presents the resulting levels after including influx effects. Panel (d) displays the gravitropic response of statoliths visualized within columella cells of root tips at angles of 0°, 40°, and 90° to the horizontal as extracted from Band *et al.* (2012). The schematics show the 'tipping point' model where the statoliths redistribute auxin levels at the root tip (denoted by arrows).

auxin dynamics. Their study concludes that influx carriers control which tissues of the root will present high auxin levels, whereas the efflux PIN carriers control only the direction of auxin transport within these tissues.

Aside from internal factors, root development is also highly influenced by its surrounding environment, especially by gravity. A model which accounts for root gravitropism has been developed by Band *et al.* (2012) who made use of a downward pointing gravity stimulus. Their simulations showed that auxin gets redistributed from the upper side of the root to the lower side much faster than it takes the actual root to show bending effects. This supports the hypothesis that auxin redistribution regulates the root organ curvature. This auxin gradient only lasts for the first approximately 2 hours of the process (in the case of *A. thaliana*), after which auxin levels become again equal. The explanation for this may lie in the fact that the rapid loss of auxin level asymmetry occurred at a bending angle of approximately 40°, which could denote the tipping point for specialized statoliths which perceive gravity within columella cells as presented in panel (d) of Figure 5. The switch then in auxin flow at the root tip is likely to reflect the repositioning of the statoliths.

Conclusion

The above examples illustrate how detailed, cell-level imaging combined with multiscale, cell-based modeling can help unravel the mechanisms of biological development by which single cells form functioning, multicellular structures. An important asset of these models is abstraction and simplification of organizational levels (Merks and Glazier, 2005). Models of higher levels of organization (cells, tissues), typically coarse-grain the underlying molecular levels to a large extent. For example, to understand differential-adhesion driven cell sorting, it is not always necessary to unravel the molecular details of cell–cell adhesion; what matters most to the cell and tissue scale are coarse-grained physical measures: cell-type-dependent differences in cell adhesion strength and cortical tension. Similarly, to analyze how auxin regulates root growth and gravitropism, it is helpful to keep the big picture in mind and focus on the overall polarization patterns and the response to auxin. Of course, this is not to say that the molecular scale is irrelevant: the genes and molecules determine the logical ‘rules’ according to which cells behave and differentiate after receiving regulatory signals from the other cells, as well as their biophysical properties and integrated computational and experimental studies then help us reconstruct the collective cell behavior that we know as biological development.

Acknowledgments

Roeland M.H. Merks, Sonja E. M. Boas, and Elisabeth G. Rens were in part supported by the Division for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organization for Scientific Research (NWO). Hannan Tahir wishes to thank M. Schmitz and B. Nelemans for providing an image of chick embryo somitogenesis. HT also acknowledges

the support from ZonMW-VICI grant no. 918.11.635 to Prof. Theo Smit of VUMC.

References

- Agero, U., Glazier, J.A., Hsek, M., 2010. Bulk elastic properties of chicken embryos during somitogenesis. *Biomedical Engineering Online* 9 (1), 19.
- Arima, S., Nishiyama, K., Ko, T., *et al.*, 2011. Angiogenic morphogenesis driven by dynamic and heterogeneous collective endothelial cell movement. *Development* 138 (21), 4763–4776 Available at: <http://dev.biologists.org/content/138/21/4763.abstract> (accessed 03.06.15). doi:10.1242/dev.068023.
- Baker, R.E., Schnell, S., Maini, P.K., 2006. A mathematical investigation of a clock and wavefront model for somitogenesis. *Journal of Mathematical Biology* 52 (4), 458–482.
- Baker, R.E., Schnell, S., Maini, P.K., 2008. Mathematical models for somite formation. *Current Topics in Developmental Biology* 81, 183–203.
- Band, L.R., Wells, D.M., Fozard, J.A., *et al.*, 2014. Systems analysis of auxin transport in the arabidopsis root apex. *Plant Cell Online* 26 (3), 862–875.
- Band, L.R., Wells, D.M., Larrieu, A., *et al.*, 2012. Root gravitropism is regulated by a transient lateral auxin gradient controlled by a tipping-point mechanism. *Proceedings of the National Academy of Sciences of the United States of America* 109 (12), 4668–4673.
- Bauer, L., Jackson, T.L., Jiang, Y., 2007. A cell-based model exhibiting branching and anastomosis during tumor-induced angiogenesis. *Biophysical Journal* 92, 3105–3121.
- Bauer, L., Jackson, T.L., Jiang, Y., 2009. Topography of Extracellular Matrix Mediates Vascular Morphogenesis and Migration Speeds in Angiogenesis. *PLoS Computational Biology* 5 (7), e1000445. doi:10.1371/journal.pcbi.1000445.
- Beltman, J.B., Allen, C.D.C., Cyster, J.G., de Boer, R.J., 2011. B cells within germinal centers migrate preferentially from dark to light zone. *Proceedings of the National Academy of Sciences of the United States of America* 108 (21), 8755–8760.
- Bentley, K., Franco, C.A., Philippides, A., *et al.*, 2014. The role of differential VE-cadherin dynamics in cell rearrangement during angiogenesis. *Nature Cell Biology* 16, 309–321. doi:10.1038/ncb2926.
- Bischofs, I.B., Schwarz, U.S., 2003. Cell organization in soft media due to active mechanosensing. *Proceedings of the National Academy of Sciences of the United States of America* 100 (16), 9274–9279. ISSN 0027-8424. doi: 10.1073/pnas.1233544100. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=170908&tool=pmcentrez&rendertype=abstract> (accessed 02.06.15).
- Boudon, F., Chopard, J., Ali, O., *et al.*, 2015. A computational framework for 3D mechanical modeling of plant morphogenesis with cellular resolution. *PLoS Computational Biology* 11 (1), e1003950.
- Byrne, H., Drasdo, D., 2009. Individual-based and continuum models of growing cell populations: A comparison. *Journal of Mathematical Biology* 58 (4–5), 657–687. ISSN 0303-6812. doi: 10.1007/s00285-008-0212-0. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18841363>; <http://link.springer.com/article/10.1007/s00285-008-0212-0> (accessed 02.06.15).
- Castro-González, C., Luengo-Oroz, M.A., Duloquin, L., *et al.*, 2014. A digital framework to build, visualize and analyze a gene expression atlas with cellular resolution in zebrafish early embryogenesis. *PLoS Computational Biology* 10 (6), e1003670.
- Checa, S., Rausch, M.K., Petersen, A., Kuhl, E., Duda, G.N., 2014. The emergence of extracellular matrix mechanics and cell traction forces as important regulators of cellular self-organization. *Biomechanics and Modeling*. doi:10.1007/s10237-014-0581-9.
- Collier, J.R., McInerney, D., Schnell, S., *et al.*, 2000. A cell cycle model for somitogenesis: Mathematical formulation and numerical simulation. *Journal of Theoretical Biology* 207 (3), 305–316.
- Cooke, J., Zeeman, E.C., 1976. A clock and wavefront model for control of the number of repeated structures during animal morphogenesis. *Journal of Theoretical Biology* 58 (2), 455–476.
- Cumming, B.D., McElwain, D.L.S., Upton, Z., 2010. A mathematical model of wound healing and subsequent scarring. *Journal of the Royal Society, Interface/the Royal Society* 7 (42), 19–34. ISSN 1742-5662. doi: 10.1098/rsif.2008.0536. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2839370&tool=pmcentrez&rendertype=abstract>.
- Dallon, J., 2000. Numerical Aspects of Discrete and Continuum Hybrid Models in Cell Biology. *Applied Numerical Mathematics* 32 (2), 137–159.

- Dallon, J.C., Sherratt, J.A., Maini, P.K., 1999. Mathematical modelling of extracellular matrix dynamics using discrete cells: Fiber orientation and tissue regeneration. *Journal of Theoretical Biology*, 449–471. Available at: <http://www.sciencedirect.com/science/article/pii/S002251939909712> (accessed 02.06.15).
- Das, A., Lauffenburger, D., Asada, H., Kamm, R.D., 2010. A hybrid continuum-discrete modelling approach to predict and control angiogenesis: Analysis of combinatorial growth factor and matrix effects on vessel-sprouting morphology. *Philosophical Transactions A* 368, 2937–2960.
- Daub, J.T., Merks, R.M.H., 2013. A cell-based model of extracellular-matrix-guided endothelial cell migration during angiogenesis. *Bulletin of Mathematical Biology* 75, 1377–1399. doi:10.1007/s11538-013-9826-5.
- Dejana, E., 2004. Endothelial cell–cell junctions: Happy together. *Nature Reviews Molecular Cell Biology* 5, 261–270.
- Dias, A.S., de Almeida, I., Belmonte, J.M., Glazier, J.A., Stern, C.D., 2014. Somites without a clock. *Science (New York, NY)* 343 (6172), 791–795.
- Dray, N., Lawton, A., Nandi, A., et al., 2013. Cell–fibronectin interactions propel vertebrate trunk elongation via tissue mechanics. *Current Biology* 23 (14), 1335–1341.
- Farhadifar, R., Röper, J.-C., Aigouy, B., Eaton, S., Jülicher, F., 2007. The influence of cell mechanics, cell–cell interactions, and proliferation on epithelial packing. *Current Biology* 17 (24), 2095–2104.
- Friml, J., Jones, A.R., 2010. Endoplasmic reticulum: The rising compartment in auxin biology. *Plant Physiology* 154 (2), 458–462.
- Gerhardt, H., 2008. VEGF and endothelial guidance in angiogenic sprouting. *Organogenesis* 4, 241–246.
- Glazier, J.A., Graner, F., 1993. Simulation of the differential adhesion driven rearrangement of biological cells. *Physical Review E* 47 (3), 2128–2154.
- Graner, F., Glazier, J.A., 1992. Simulation of biological cell sorting using a two-dimensional extended potts model. *Physical Review Letters* 69, 2013–2016. Available at: <http://link.aps.org/doi/10.1103/PhysRevLett.69.2013> (accessed 02.06.15). doi:10.1103/PhysRevLett.69.2013.
- Grégoire, G., Chaté, H., Tu, Y., 2003. Moving and staying together without a leader. *Physica D: Nonlinear Phenomena* 181 (3–4), 157–170. Available at: <http://www.sciencedirect.com/science/article/pii/S0167278903001027> (accessed 02.06.15).
- Grieneisen, V.A., Xu, J., Marée, A.F.M., Hogeweg, P., Scheres, B., 2007. Auxin transport is sufficient to generate a maximum and gradient guiding root growth. *Nature* 449 (7165), 1008–1013.
- Grima, R., Schnell, S., 2007. Can tissue surface tension drive somite formation? *Developmental Biology* 307 (2), 248–257.
- Harris, K., Wild, P.P., Stopak, D.D., 1980. Silicone rubber substrata: A new wrinkle in the study of cell locomotion. *Science (New York, NY)* 208 (4440), 177–179.
- Haston, W.S., Shields, J.M., Wilkinson, P.C., 1983. The orientation of fibroblasts and neutrophils on elastic substrata. *Experimental Cell Research* 146 (1), 117–126.
- Hester, S.D., Belmonte, J.M., Gens, J.S., Clendenen, S.G., Glazier, J.A., 2011. A multi-cell, multi-scale model of vertebrate segmentation and somite formation. *PLoS Computational Biology* 7 (10), e1002155.
- Honda, H., Tanemura, M., Nagai, T., 2004. A three-dimensional vertex dynamics cell model of space-filling polyhedra simulating cell behavior in a cell aggregate. *Journal of Theoretical Biology* 226 (4), 439–453.
- Hubaud, A., Pourqui, O., 2014. Signalling dynamics in vertebrate segmentation. *Nature Reviews Molecular Cell Biology* 15 (11), 709–721.
- Hutson, M.S., Brodland, G.W., Yang, J., Viens, D., 2008. Cell sorting in three dimensions: Topology, fluctuations, and fluidlike instabilities. *Physical Review Letters* 101 (148105). doi:10.1103/PhysRevLett.101.148105.
- Inaba, M., Yamanaka, H., Kondo, S., 2012. Pigment pattern formation by contact-dependent depolarization. *Science (New York, NY)* 335 (6069), 677.
- Jakobsson, L., Franco, C.A., Bentley, K., et al., 2010. Endothelial cells dynamically compete for the tip cell position during angiogenic sprouting. *Nature Cell Biology* 12, 943–953. doi:10.1038/ncb2103.
- Kay, R.R., Langridge, P., Traynor, D., Hoeller, O., 2008. Changing directions in the study of chemotaxis. *Nature Reviews Molecular Cell Biology* 9 (6), 455–463.
- Köhn-Luque, A., de Back, W., Starrau, J., et al., 2011. Early embryonic vascular patterning by matrix-mediated paracrine signalling: A mathematical model study. *PLoS ONE* 6 (9), e24175. doi:10.1371/journal.pone.0024175.
- Köhn-Luque, A., de Back, W., Yamaguchi, Y., et al., 2013. Early embryonic vascular patterning by matrix-mediated paracrine signalling: A mathematical model study. *Physical Biology* 10 (6), 066007. doi:10.1088/1478-3975/10/6/066007.
- Kondo, S., 2014. Self-organizing somites. *Science* 343 (6172), 736–737.
- Krieg, M., 2009. Cell adhesion and cell mechanics during zebrafish development. PhD thesis, Technische Universität Dresden. Available at: <http://nbn-resolving.org/urn:Nbn:de:bsz:14-qucosa-26093> (accessed 02.06.15).
- Krieg, M., Arboleda-Estudillo, Y., Puech, P.H., et al., 2008. Tensile forces govern germ-layer organization in zebrafish. *Nature Cell Biology* 10, 429–436. doi:10.1038/ncb1705.
- Le Dévédec, S.E., Yan, K., Bont, H., et al., 2010. Systems microscopy approaches to understand cancer cell migration and metastasis. *Cellular and Molecular Life Sciences* 67 (19), 3219–3240.
- Lin, F., Butcher, E.C., 2006. T cell chemotaxis in a simple microfluidic device. *Lab on a Chip* 6 (11), 1462–1469.
- Mahalwar, P., Walderich, B., Singh, A.P., Nusslein-Volhard, C., 2014. Local reorganization of xanthophores fine-tunes and colors the striped pattern of zebrafish. *Science (New York, NY)* 348 (6232), 297.
- Manoussaki, D., Lubkin, S.R., Vernon, R.B., Murray, J.D., 1996. A mechanical model for the formation of vascular networks *in vitro*. *Acta Biotheoretica* 44 (3–4), 271–282. Available at: <http://link.springer.com/article/10.1007/BF00046533> (accessed 02.06.15).
- Meinhardt, H., 1986. Models of segmentation. In: Bellairs, R., Ede, D.A., Lash, J.W. (Eds.), *Somites in Developing Embryos*. NATO ASI Series, vol. 118. New York: Plenum Press, pp. 179–189. Available at: http://link.springer.com/chapter/10.1007/978-1-4899-2013-3_14 (accessed 25.06.15).
- Merks, R.M.H., 2013. Cell-based modeling. In: Enquist, B. (Ed.), *Encyclopedia of Applied and Computational Mathematics*. Springer. Available at: <https://www.springerreference.com> (accessed 02.06.15).
- Merks, R.M.H., Brodsky, S., Goligorsky, M., Newman, S., Glazier, J.A., 2006. Cell elongation is key to *in silico* replication of *in vitro* vasculogenesis and subsequent remodeling. *Developmental Biology* 289, 44–54.
- Merks, R.M.H., Glazier, J.A., 2005. A cell-centered approach to developmental biology. *Physica A* 352 (1), 113–130.
- Merks, R.M.H., Guravage, M., Inze, D., Beemster, G.T.S., 2011. Virtualleaf: An open-source framework for cell-based modeling of plant tissue growth and development. *Plant Physiology* 155 (2), 656–666.
- Merks, R.M.H., Koolwijk, P., 2009. Modeling morphogenesis *in silico* and *in vitro*: Towards quantitative, predictive, cell-based modeling. *Mathematical Modelling of Natural Phenomena* 4 (5), 149–171.
- Merks, R.M.H., Perryn, E.D., Shirinifard, A., Glazier, J.A., 2008. Contact-inhibited chemotaxis in *de novo* and sprouting blood-vessel growth. *PLoS Computational Biology* 4 (9), e1000163. doi:10.1371/journal.pcbi.1000163.
- Milde, F., Bergdorf, M., Koumoutsakos, P., 2008. A hybrid model for three-dimensional simulations of sprouting angiogenesis. *Biophysical Journal* 95 (7), 3146–3160.
- Murray, J.D., Oster, G.F., Harris, A.K., 1983. A mechanical model for mesenchymal morphogenesis. *Journal of Mathematical Biology* 17, 125–129.
- Nakamasu, A., Takahashi, G., Kanbe, A., Kondo, S., 2009. Interactions between zebrafish pigment cells responsible for the generation of Turing patterns. *Proceedings of the National Academy of Sciences of the United States of America* 106 (21), 8429–8434.
- Namy, P., Ohayon, J., Tracqui, P., 2004. Critical conditions for pattern formation and *in vitro* tubulogenesis driven by cellular traction fields. *Journal of Theoretical Biology* 227 (1), 103–120. ISSN 00225193. doi: 10.1016/j.jtbi.2003.10.015. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14969709> (accessed 02.06.15).
- Newman, T.J., 2005. Modeling multi-cellular systems using sub-cellular elements. *Mathematical Biosciences and Engineering* 2 (3), 613–624. Arxiv preprint q-bio/0504028. Available at: <http://arxiv.org/abs/q-bio/0504028> (accessed 02.06.15).
- van Oers, R.F.M., Rens, E.G., LaValley, D.J., Reinhart-King, C.A., Merks, R.M.H., 2014. Mechanical cell-matrix feedback explains pairwise and collective endothelial cell behavior *in vitro*. *PLoS Computational Biology* 10 (8), e1003774. Available at: <http://dx.doi.org/10.1371/journal.pcbi.1003774> (accessed 02.06.15). doi:10.1371/journal.pcbi.1003774.
- Okuda, S., Inoue, Y., Eiraku, M., Sasai, Y., Adachi, T., 2013a. Reversible network reconnection model for simulating large deformation in dynamic tissue morphogenesis. *Biomechanics and Modeling in Mechanobiology* 12 (4), 627–644.
- Okuda, S., Inoue, Y., Eiraku, M., Sasai, Y., Adachi, T., 2013b. Modeling cell proliferation for simulating three-dimensional tissue morphogenesis based on a reversible network reconnection framework. *Biomechanics and Modeling in Mechanobiology* 12 (5), 987–996.
- Olivier, N., Luengo-Oroz, M.A., Duloquin, L., et al., 2010. Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy. *Science (New York, NY)* 329 (5994), 967–971.
- Palm, M.M., Merks, R.M.H., 2013. Vascular networks due to dynamically arrested crystalline ordering of elongated cells. *Physical Review E* 87, e012725. Available at: <http://link.aps.org/doi/10.1103/PhysRevE.87.012725> (accessed 02.06.15). doi:10.1103/PhysRevE.87.012725.

- Parsa, H., Upadhyay, R., Sia, S.K., 2011. Uncovering the behaviors of individual cells within a multicellular microvascular community. *Proceedings of the National Academy of Sciences of the United States of America* 108 (12), 5133–5138.
- Qutub, A.A., Liu, G., Vempati, P., Popel, A.S., 2009. Integration of angiogenesis modules at multiple scales: From molecular to tissue. *Pacific Symposium on Biocomputing*. 316–327.
- Reinhardt, J.W., Gooch, K.J., 2014. Agent-based modeling traction force mediated compaction of cell-populated collagen gels using physically realistic fibril mechanics. *Journal of Biomechanical Engineering* 136 (2), 021024. ISSN 1528-8951. doi: 10.1115/1.4026179. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24317298> (accessed 02.06.15).
- Sandersius, S.A., Chuai, M., 2011. A 'chemotactic dipole' mechanism for large-scale vortex motion during primitive streak formation in the chick embryo. *Physical Biology* 8 (4), 045008. ISSN 1478-3975. doi: 10.1088/1478-3975/8/4/045008. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21750368>; <http://iopscience.iop.org/1478-3975/8/4/045008> (accessed 02.06.15).
- Savill, N.J., Hogeweg, P., 1997. Modelling morphogenesis: From single cells to crawling slugs. *Journal of Theoretical Biology* 184 (3), 229–235.
- Schier, A.F., Talbot, W.S., 2005. Molecular genetics of axis formation in zebrafish. *Annual Review of Genetics* 39, 561–613.
- Schlüter, D.K., Ramis-Conde, I., Chaplain, M.A.J., 2012. Computational modeling of single-cell migration: The leading role of extracellular matrix fibers. *Biophysical Journal* 103 (6), 1141–1151. ISSN 1542-0086. doi: 10.1016/j.bpj.2012.07.048. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3446673&tool=pmcentrez&rendertype=abstract> (accessed 02.06.15).
- Schnell, S., Maini, P.K., 2000. Clock and induction model for somitogenesis. *Developmental Dynamics* 217 (4), 415–420.
- Sepúlveda, N., Petitjean, L., Cochet, O., *et al.*, 2013. Collective cell motion in an epithelial sheet can be quantitatively described by a stochastic interacting particle model. *PLoS Computational Biology* 9 (3), e1002944. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1002944. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3591275&tool=pmcentrez&rendertype=abstract> (accessed 02.06.15).
- Serini, G., Ambrosi, D., Giraudo, E., *et al.*, 2003. Modeling the early stages of vascular network assembly. *EMBO Journal* 22 (8), 1771–1779.
- Short, K.M., Combes, A.N., Lefevre, J., *et al.*, 2014. Global quantification of tissue dynamics in the developing mouse kidney. *Developmental Cell* 29 (2), 188–202.
- Sluka, J.P., Shirinifard, A., Swat, M., *et al.*, 2014. The cell behavior ontology: Describing the intrinsic biological behaviors of real and model cells seen as active agents. *Bioinformatics* 30 (16), 2367–2374.
- Steinberg, M.S., 1963. Reconstruction of tissues by dissociated cells. Some morphogenetic tissue movements and the sorting out of embryonic cells may have a common explanation. *Science* 141, 401–408.
- Stern, C.D., Fraser, S.E., Keynes, R.J., Primmitt, D.R.N., 1988. A cell lineage analysis of segmentation in the chick embryo. *Development* 104 (Suppl.), 231–244.
- Swat, M.H., Thomas, G.L., Belmonte, J.M., *et al.*, 2012. Multi-scale modeling of tissues using *compuCell3D*. *Methods in Cell Biology* 110, 325–366.
- Szabó, A., Czirik, A., 2010. The role of cell-cell adhesion in the formation of multicellular sprouts. *Mathematical Modelling of Natural Phenomena* 5 (1), 106–122. doi:10.1051/mmnp/20105105.
- Szabó, A., Erica, D.P., Czirik, A., 2007. Network formation of tissue cells via preferential attraction to elongated structures. *Physical Review Letters* 98 (3), 038102.
- Szabó, A., Mehes, E., Kosa, E., Czirik, A., 2008. Multicellular sprouting *in vitro*. *Biophysical Journal* 95 (6), 2702–2710.
- Szabo, A., Perryn, E.D., Czirik, A., 2007. Network formation of tissue cells via preferential attraction to elongated structures. *Physical Review Letters* 98 (3), 038102. ISSN 0031-9007. doi: 10.1103/PhysRevLett.98.038102. Available at: <http://link.aps.org/doi/10.1103/PhysRevLett.98.038102>; <http://journals.aps.org/prl/abstract/10.1103/PhysRevLett.98.038102> (accessed 02.06.15).
- Tahir, H., Niculescu, I., Bona-Casas, C., Merks, R.M.H., Hoekstra, A.G., 2015. An in-silico study on the role of smooth muscle cells migration in neointimal formation after coronary stenting. *Journal of the Royal Society Interface* 12, 20150358. doi:10.1098/rsif.2015.0358.
- Tepass, U., Godt, D., Winklbauer, R., 2002. Cell sorting in animal development: Signalling and adhesive mechanisms in the formation of tissue boundaries. *Current Opinion in Genetics and Development* 12 (5), 572–582.
- Townes, P.L., Holtfreter, J., 1955. Directed movements and selective adhesion of embryonic amphibian cells. *Journal of Experimental Zoology* 18, 53–120.
- Turner, S., Sherratt, J.A., 2002. Intercellular adhesion and cancer invasion: A discrete simulation using the extended potts model. *Journal of Theoretical Biology* 216, 85–100.
- Wang, H.-Y., Huang, Y.X., Qi, Y.F., *et al.*, 2013. Mathematical models for the notch and wnt signaling pathways and the crosstalk between them during somitogenesis. *Theoretical Biology and Medical Modelling* 10 (1), 27.
- Winer, J.P., Oake, S., Janmey, P.A., 2009. Non-linear elasticity of extracellular matrices enables contractile cells to communicate local position and orientation. *PLoS ONE* 4 (7), e6382.
- Yin, Z., Noren, D., Wang, C.J., Hang, R., Levchenko, A., 2008. Analysis of pairwise cell interactions using an integrated dielectrophoretic-microfluidic system. *Molecular Systems Biology* 4, 232.
- Zajac, M., Jones, G.L., Glazier, J.A., 2003. Simulating convergent extension by way of anisotropic differential adhesion. *Journal of Theoretical Biology* 222 (2), 247–259. ISSN 00225193. doi: 10.1016/S0022-5193(03)00033-X. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S002251930300033X> (accessed 02.06.15).

Further Reading

- Anderson, A.R.A., Chaplain, M.A.J., Rejniak, K.A. (Eds.), 2007. *Single-Cell Based Models in Biology and Medicine*. Basel: Birkhauser.
- Davies, J., 2013. *Mechanisms of Morphogenesis*. Amsterdam: Academic Press, Elsevier.
- Forgacs, G., Newman, S.A., 2005. *Biological Physics of the Developing Embryo*. New York, NY: Cambridge University Press.
- Keller, E.F., 2002. *Making Sense of Life. Explaining Biological Development with Models, Metaphors, and Machines*. Cambridge, MA: Harvard University Press.
- Noble, D., 2006. *The Music of Life: Biology Beyond the Genome*. Oxford: Oxford University Press.
- Scianna, M., Preziosi, L., 2013. *Cellular Potts Models: Multi-Scale Extensions and Biological Applications*. Boca Raton, FL: CRC Press, Taylor & Francis Group, LLC.

Relevant Websites

- <http://www.compuCell3d.org>
CompuCell3D.
- <http://multicellids.org>
MultiCellIDS.
- <http://bioemergences.eu>
The BioEmergences.
- <http://www.cs.ox.ac.uk/projects/chaste/>
University of Oxford.
- <http://virtualleaf.googlecode.com>
VirtualLeaf.

Neurogenesis in the Adult Brain

MR Akins and ADR Garcia, Drexel University, Philadelphia, PA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

For over a century, the idea that the mature central nervous system (CNS) was a fixed structure incapable of producing new neurons was a central dogma within the field of neuroscience. Neurogenesis, the production of neurons, was thought to be confined to early development. Though initially met with great resistance and controversy, compelling evidence beginning in the early 1980s and continuing through the early 2000s from birds, rodents, and eventually humans demonstrated that indeed the adult CNS undergoes continued neurogenesis in several brain regions throughout life (Goldman and Nottebohm, 1983; Eriksson *et al.*, 1998; Gage *et al.*, 1998; Doetsch *et al.*, 1999; Seri *et al.*, 2001). The discovery of adult neural stem cell (NSC) populations that produce new neurons in mature animals not only overturned a century-old fundamental principle of neuroscience, but also created a paradigm shift that has enhanced our understanding of the ways in which the CNS develops and is maintained throughout life. Indeed, a number of cellular processes that define and regulate adult neurogenesis have subsequently been identified in the developing CNS, illustrating shared fundamental properties between developmental and adult neurogenesis.

In the adult mammalian forebrain, NSCs are localized in two primary locations: the subgranular zone (SGZ) of the hippocampal dentate gyrus and the ventricular-subventricular zone (V-SVZ), an epithelium lining the lateral ventricles (Figure 1). SGZ stem cells produce granule cells that incorporate into the dentate gyrus of the hippocampal formation,

a structure critical for learning and memory processes (Deng *et al.*, 2010). NSCs in the V-SVZ produce immature neurons that migrate rostrally through the rostral migratory stream (RMS) and integrate into circuits of the olfactory bulb (Figure 2). These immature neurons produce two classes of olfactory bulb interneurons, granule and periglomerular cells (PGs), that modulate the information processing by the olfactory bulb projection neurons, the mitral and tufted cells. The olfactory bulb processes information about an animal's odor environment that is detected by olfactory sensory neurons in the nose and conveyed via mitral and tufted cells to other parts of the brain, including the cerebral cortex. Adult neurogenesis therefore modifies how olfactory circuits respond to odorants, thereby altering the organism's response to sensory stimuli. In this article, we focus on our understanding of the intrinsic and extrinsic factors that regulate neurogenesis within the V-SVZ and how these newly born neurons contribute to olfactory circuit function.

Neurogenesis in the V-SVZ and the SGZ share important similarities, and discoveries made in each system inform our understanding of the other. Adult NSCs in both the V-SVZ and SGZ share a number of features including an elongated, radial-like morphology and expression of the intermediate filament protein glial fibrillary acidic protein (GFAP). Both populations of adult NSCs exhibit a slow rate of turnover, producing a more rapidly dividing, intermediate progenitor cell (IPC) that generates neurons. Moreover, many of the same molecular signaling pathways regulate both adult NSC populations. For further reading on adult neurogenesis in the hippocampus, the reader is referred to a number of insightful reviews (see Further Reading).

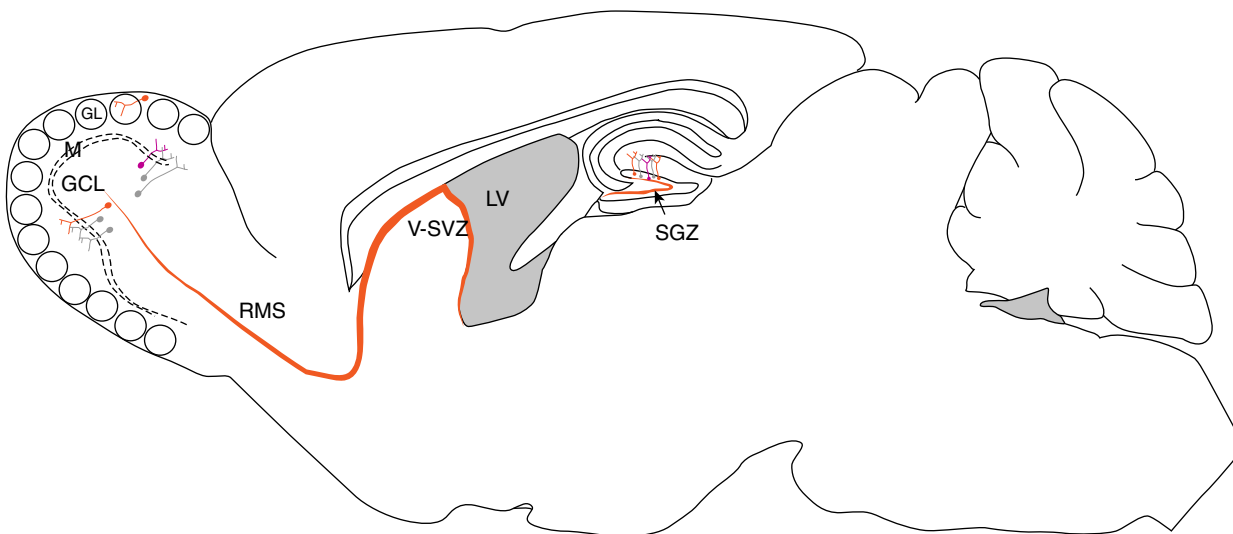


Figure 1 Schematic of the adult mouse brain depicting the neurogenic regions. Adult neural stem cells (NSCs) reside in the ventricular-subventricular zone (V-SVZ), the epithelial layer lining the lateral ventricles (LV), and the subgranular zone (SGZ) of the hippocampal formation. Type B cells in the V-SVZ proliferate and generate intermediate progenitor cells which produce immature neuroblasts that migrate anteriorly through the rostral migratory stream (RMS) into the olfactory bulbs where new neurons incorporate into existing granule cell and periglomerular populations GCL, granule cell layer; GL, glomerular layer; M, mitral cell layer.

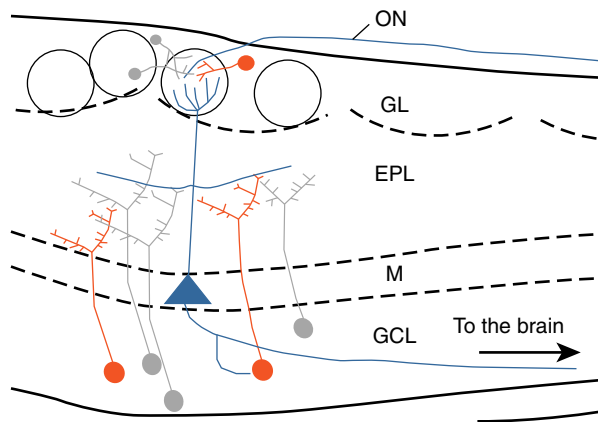


Figure 2 Olfactory circuit diagram. Odorant molecules are carried by air into the nose where they are detected by olfactory sensory neurons, each of which responds to a subset of odorants based on the expression of a single odorant receptor out of the hundreds present in most mammalian genomes. Axons of these sensory neurons convey information via the olfactory nerve (ON) into the central nervous system (CNS) where they terminate in the olfactory bulb in small globular structures termed glomeruli that are positioned in the superficial glomerular layer (GL). Each glomerulus receives input from sensory neurons expressing the same receptor. These sensory neuron axons form synapses with mitral (M) and tufted cells, each of which receives input from a single glomerulus and in turn conveys information to other regions of the brain. The information carried to the rest of the brain by mitral and tufted cells is highly influenced by their interactions with two other classes of neurons within the olfactory bulb. Periglomerular cells (PGs) surround a glomerulus and regulate the response profile of the cells forming synapses within that glomerulus. Granule cells are located in the core of the olfactory bulb in the granule cell layer (GCL). These neurons do not have axons, but send dendrites into the external plexiform layer (EPL), where they form reciprocal dendrodendritic synapses with dendrites of mitral and tufted cells.

Insights gained from the study of adult neurogenesis also impact our knowledge of developmental neurogenesis, despite inherent differences in these two epochs. Whereas adult neurogenesis involves incorporation of adult-born neurons into an existing circuit, developmental neurogenesis requires the *de novo* creation of these circuits. Nevertheless, many molecular pathways are shared between these two forms of neurogenesis. Moreover, much like developmental neurogenesis, adult NSCs in the V-SVZ comprise a heterogeneous population that produces morphologically and functionally distinct neuronal subtypes (Merkle *et al.*, 2014). Thus the adult NSCs of the V-SVZ provide a useful model system with which to study many of the fundamental principles of neurogenesis, including proliferation, cell specification, and regulatory pathways. Consequently, our understanding of both adult and developmental neurogenesis has been influenced by observations initially made in the other system. Neurogenesis in the developing brain is the subject of a number of excellent review articles (see Further Reading).

Historical Perspective

Evidence of mitosis in the mature CNS dates at least as far back as 1912, when Allen described mitotic figures in the brains of

4-month-old adult rats (Allen, 1912). Despite this report, the idea of adult neurogenesis remained largely unaccepted throughout much of the twentieth century. In the 1960s, Altman and Das published a series of reports describing the incorporation of radioactive thymidine ($[^3\text{H}]\text{-Thy}$) during DNA synthesis in the brains of mature mammals. Labeled cells were observed in the granular layers of the olfactory bulbs and hippocampal dentate gyrus of a number of adult mammalian species, including rats, guinea pigs, and cats (Altman, 1963, 1969; Altman and Das, 1965, 1967). These studies were largely dismissed due to the inability to determine the cellular phenotype of the labeled cells, and it was assumed that these cells were glia rather than neurons.

A major advance in the study of adult NSCs was the development and application of the thymidine analog bromodeoxyuridine (BrdU) to the study of proliferating cells in the mature CNS. Like $[^3\text{H}]\text{-Thy}$, BrdU is incorporated into DNA during synthesis. However, instead of detection by autoradiography, BrdU can be detected by immunohistochemistry. Therefore, double labeling with cell type-specific markers enabled lineage tracing of proliferating cells and subsequent identification of newly generated neurons in the brains of adult mammals. Using this strategy, investigators were able to identify BrdU-labeled neurons in the hippocampal dentate gyrus and the olfactory bulbs of various mammalian species, including rodents and primates (Cameron *et al.*, 1993; Gould *et al.*, 1997; Cameron and McKay, 2001), finally demonstrating unequivocally that neurogenesis persists in the mature CNS. In 1998, it was reported that postmortem human tissue from patients that had received a single BrdU infusion for diagnostic purposes also showed evidence of adult neurogenesis in the hippocampus (Eriksson *et al.*, 1998). Taken together, the overwhelming evidence for adult neurogenesis, and its relevance to humans, could no longer be ignored. A fundamental principle of neuroscience was overturned, reshaping our understanding of CNS function and its remarkable capacity for plasticity.

Identity of Adult NSCs

V-SVZ NSCs reside in a complex, multicellular niche that includes multiciliated ependymal cells, differentiated astrocytes, blood vessels, IPCs, and immature neuroblasts (Ihrle and Alvarez-Buylla, 2011). NSCs, known as Type B cells (Doetsch *et al.*, 1999), are rarely dividing neural precursors that express the intermediate filament protein GFAP (Doetsch *et al.*, 1999; Seri *et al.*, 2001; Garcia *et al.*, 2004). Type B cells are largely quiescent, dividing infrequently to produce rapidly dividing IPCs, or transit amplifying cells (Figure 3). Whereas Type B cells divide asymmetrically, producing one stem cell and one IPC, IPCs undergo symmetric cell division to produce two identical daughter cells, thereby amplifying the progenitor pool. Although IPCs are generally more restricted in their proliferation and differentiation potential than Type B cells, they retain multipotency *in vitro*, generating cells of the three major neural lineages – neurons, astrocytes, and oligodendrocytes (Doetsch *et al.*, 2002). *In vivo*, IPCs produce immature neuroblasts that ultimately mature into neurons (Figure 3). When isolated from the V-SVZ and grown in

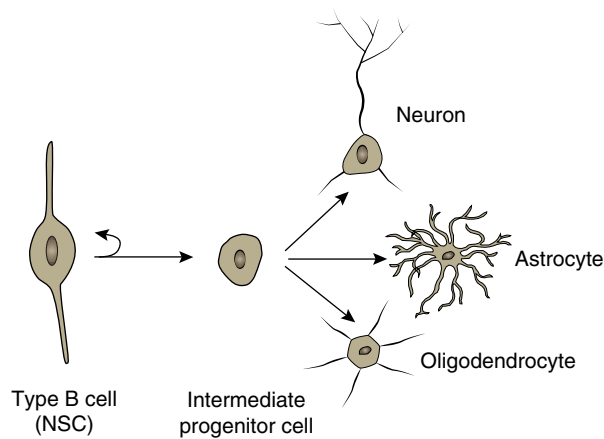


Figure 3 Lineage relationship of adult neural stem cells (NSCs) and their progeny. Type B cells residing in the ventricular-subventricular zone (V-SVZ) express glial fibrillary acidic protein (GFAP) and possess a radial morphology. These cells undergo asymmetric cell division in which one daughter cell retains a stem cell identity, while the other cell acquires an intermediate progenitor cell (IPC) identity. IPCs are multipotent cells, capable of generating neurons, astrocytes, or oligodendrocytes.

culture, Type B cells display hallmark features of stem cells: self-renewal and multipotency. These cells form neurospheres that can be dissociated and passaged under clonal conditions, producing secondary and tertiary neurospheres (Reynolds and Weiss, 1992). Under differentiation conditions, neurospheres generate cells of the three neural lineages (Imura *et al.*, 2003). *In vivo*, Type B cells repopulate the V-SVZ following ablation of actively proliferating cells (Doetsch *et al.*, 1999). Moreover, lineage tracing using retroviral labeling and genetic fate mapping tools demonstrate that adult-generated neurons, astrocytes, and oligodendrocytes are derived from Type B cells (Garcia *et al.*, 2004; Menn, 2006). The adult V-SVZ therefore includes NSCs – the Type B cells – that can give rise to neurons.

GFAP expression has historically been used to identify mature, differentiated astrocytes, the major macroglial subtype in the CNS. Indeed, adult NSCs share a number of features with astrocytes, including a radial morphology and ultrastructural profiles, leading to their characterization as specialized astrocytes (Anthony *et al.*, 2004; Kriegstein and Alvarez-Buylla, 2009). Nevertheless, it is important to note that despite these shared features, adult NSCs are distinct from non-neurogenic astrocytes that are distributed throughout other regions of the brain. GFAP-expressing cells isolated from the adult V-SVZ exhibit neurogenic potential *in vitro*, whereas cells isolated from the adult cortex do not, even when grown in similar culture conditions (Laywell *et al.*, 2000; Imura *et al.*, 2006). There is some evidence that following an injury, reactive astrocytes in non-neurogenic regions such as the cortex can generate multipotent neurospheres *in vitro* (Sirko *et al.*, 2013), suggesting a latent neurogenic potential in non-V-SVZ astrocytes. The extent to which reactive astrocytes *in vivo* demonstrate neurogenic properties remains poorly understood.

During early development, neural precursor cells (NPCs) are derived from epithelial cells that initially line the neural tube. These neuroepithelial precursors exhibit a radial morphology, similar to that of adult NSCs (Kriegstein and

Alvarez-Buylla, 2009). Radial glia are bipolar cells, whose somata reside in the ventricular zone of the developing CNS, and whose processes make contact with both the ventricle and the pial surface, spanning the width of the developing cortex. Though once believed to be exclusively glial precursors, time lapse imaging of virally labeled radial glia in the embryonic mouse cortex demonstrated that radial glia divide to produce IPCs and cortical neurons (Noctor *et al.*, 2001). This suggests a lineage relationship between embryonic radial glia that produce cortical neurons, and Type B cells that produce neurons in the adult. Indeed, fate mapping studies in which radial glia are selectively labeled on the day of birth, postnatal day 0 (P0), reveal that Type B cells in the adult V-SVZ are directly derived from radial glia (Merkle *et al.*, 2004), further supporting the idea that adult NSCs are derived from radial glia in the developing CNS.

Cell Proliferation in the SVZ

As in development, proliferation of adult NSCs is tightly regulated by a number of molecular signaling pathways. Indeed, many of these pathways are shared by embryonic NSCs. Sonic hedgehog (Shh), an essential signaling pathway that regulates multiple processes including proliferation, patterning, and axonal pathfinding during embryonic development (Fuccillo *et al.*, 2006), is a key regulator of NSCs in the adult (Ahn and Joyner, 2005). Pharmacological and genetic disruption of Shh signaling in postnatal and adult brains leads to a reduction in neurosphere formation, as well as a reduction in proliferating cells in the V-SVZ *in vivo* (Palma *et al.*, 2005; Balordi and Fishell, 2007a,b). The reduction in proliferation in the V-SVZ persists, for at least 10 months following ablation of Shh signaling, ultimately depleting the pool of NSCs. The Notch signaling pathway, another key regulator of neurodevelopment, is active in Type B cells but not in IPCs (Imayoshi *et al.*, 2010). Genetic disruption of Notch signaling leads to a transient increase in the production of IPCs and immature neuroblasts, at the expense of Type B cells, pointing to a role for Notch signaling in the maintenance of adult NSCs in the V-SVZ. Whereas Shh and Notch signaling are constitutively active in adult NSCs under homeostatic conditions (Ahn and Joyner, 2005; Petrova *et al.*, 2013), activation of the Wnt signaling pathway in NSCs occurs only following injury or pharmacological depletion of actively dividing cells (Piccin and Morshead, 2011). This suggests that molecular regulation of NSCs is context dependent, enabling the nervous system to respond to local environmental conditions appropriately.

Olfactory stimulation is one factor likely to modify this local environment. Unilateral odor deprivation achieved by severing the olfactory nerve results in a loss of olfactory information to the deprived, ipsilateral olfactory bulb that leads to an increase in cell death on the ipsilateral side (Mandairon *et al.*, 2003). In contrast, measuring cell proliferation by BrdU incorporation reveals an increase in the number of BrdU-labeled cells on both the ipsilateral and contralateral sides. Since the increased proliferation on the contralateral side is not accompanied by cell death, these findings indicate that odor-dependent proliferation does not serve simply to compensate for apoptotic cell loss. Instead, the new cells are likely

to play an important role in a sensory system that is being remapped in response to an altered odor environment.

Neuronal Migration in the RMS

Neuroblasts generated within the rodent SVZ must migrate several millimeters rostrally to their destination in the olfactory bulb. The first stage of this migration takes place tangentially (i.e., parallel to the surface of the brain) in the RMS. The neuroblasts migrate along this path by crawling along each other in a process termed chain migration (Lois *et al.*, 1996). This homotypic mode of navigation stands in contrast to the more usual heterotypic migration in which developing neurons crawl along a cell of a different type (e.g., a radial glia) that serves as a scaffold. These neuroblasts do not only interact with each other, however, as they interact with other cells to travel to the olfactory bulb within tubes of astrocytic glial cells that are closely associated with blood vessels. The importance of these glia for supporting RMS migration is highlighted by the failure of neuroblasts to migrate rostrally in mice that do not form glial tubes due to lack of expression of the *Vax1* homeobox gene (Soria *et al.*, 2004). In these mice, the SVZ is expanded, the RMS does not form, and the olfactory bulbs are smaller than normal. However, *Vax1* null cells are able to migrate correctly if transplanted into wild-type tissue, indicating that the phenotype observed in the nulls reflects the disrupted migration scaffold.

Neuroblast guidance to the olfactory bulb is influenced by olfactory activity but, surprisingly, does not actually require this activity or even the olfactory bulb itself. Even when the olfactory bulb is surgically ablated, neuroblasts still join the RMS and migrate rostrally, where they accumulate before eventually dying (Kirschenbaum *et al.*, 1999). Instead of attractant cues that would pull the cells toward the olfactory bulb, the direction of migration is heavily influenced by repellant cues, including Slit proteins, that push the cells away from the SVZ in the same direction as the flow of the cerebrospinal fluid within the lateral ventricle (Nguyen-Ba-Charvet *et al.*, 2004; Sawamoto *et al.*, 2006). Several permissive cues along the path toward the olfactory bulb allow the rostral migration. These include integrin-mediated interactions with the extracellular matrix. Neuroblasts express $\alpha 6 \beta 1$ integrin, which binds to laminin. Interfering with integrin binding to laminin disrupts neuroblast migration while addition of peptides mimicking laminin will redirect these cells toward the exogenous source (Murase and Horwitz, 2002; Emsley and Hagg, 2003; Belvindrah *et al.*, 2007). Migration of the migrating cells is therefore influenced by molecular cues provided by their surrounding cells.

Despite receiving these cues from other cell types, the neuroblasts largely interact directly with each other as they migrate. These cells are distinguished by high expression of the protein NCAM (neuronal cell adhesion molecule) that has been posttranslationally modified through the addition of polysialic acid (PSA-NCAM). This molecule is critical for the chain migration as loss of total NCAM or removal of the PSA moiety leads to inefficient migration and a disorganized RMS (Chazal *et al.*, 2000; Hu, 2000). The rate of migration is sensitive to local levels of the neurotransmitter GABA, which is

nonvesicularly released by the neuroblasts themselves and modulated by uptake by the astrocytic glia surrounding the neuroblasts (Bolteus and Bordey, 2004). Neuroblast migration toward the olfactory bulb therefore requires a combination of repulsive and permissive cues and is regulated by interactions with the cells in the glial tubes.

Once the neuroblasts have reached the olfactory bulb, they cease their tangential migration and head toward the surface via radial migration. The switch from tangential migration in chains to radial migration as individual cells involves interactions with components of the extracellular matrix secreted by olfactory bulb cells, including the proteins tenascin-R and reelin (Hack *et al.*, 2002; Saghatelian *et al.*, 2004). Interestingly, the expression of tenascin-R in the granule cell layer is regulated by odor exposure, suggesting that differential tenascin-R expression in more active olfactory pathways may guide neuroblasts to sites of need within the olfactory bulb.

Neuronal Differentiation in the Olfactory Bulb

The adult-born cells differentiate into their final neuronal phenotype as they join existing circuits in the olfactory bulb. Depending on several factors, including where in the stem cell niche the original progenitor resided, these cells can adopt the fate of periglomerular or granule cells. PGs surround and regulate the response profiles of glomeruli – acellular, spherical balls of neuropil where synaptic contact between the sensory neurons and the mitral and tufted cells occurs. PGs can be subdivided into populations based on their molecular phenotype including the expression of specific calcium-binding proteins (e.g., calbindin vs. calretenin) or the neurotransmitter they release (GABA vs. dopamine). Granule cells, located in the core of the olfactory bulb, are GABAergic neurons that mediate interactions between mitral and tufted cells and can modulate the synchronicity of the output of these neurons. The position of a granule cell (superficial vs. deep) correlates with both its molecular phenotype as well as the cells with which it forms synapses in the external plexiform layer. Periglomerular and granule cells are therefore critical components of the olfactory bulb circuits that process olfactory information.

Adult neurogenesis generates periglomerular and granule cells of all known molecular phenotypes (Whitman and Greer, 2007a). The final identity of these neurons is based largely on the positional identity within the V-SVZ of the NSC from which a given neuron is derived (Young *et al.*, 2007). Specifically, Type B cells at different locations along the dorsoventral axis of the V-SVZ give rise to distinct populations of olfactory bulb interneurons. A major regulatory signaling molecule guiding cell fate is Shh signaling, in addition to its previously discussed role in maintaining the adult NSC pool. Shh signaling is most active in Type B cells localized in the ventral V-SVZ (Ihrle and Alvarez-Buylla, 2011; Petrova *et al.*, 2013). Genetic fate mapping studies have shown that ventral V-SVZ stem cells primarily produce deep granule cell neurons in the olfactory bulb (Merkle *et al.*, 2007, 2014). Targeted deletion of Shh in transgenic mice leads to a redistribution of adult-generated olfactory bulb neurons, with a reduction in the number of deep granule layer neurons and a concomitant

increase in the production of superficial granule cells (Ihrle and Alvarez-Buylla, 2011). PGs are also affected by disruption of Shh signaling. Specifically, the production of calbindin-positive periglomerular neurons is reduced by 90% following disruption of Shh signaling. Moreover, ectopic activation of Shh signaling in NSCs localized in the dorsal V-SVZ leads to increased production of deep layer granule cells localized in the olfactory bulb. Taken together, these observations point to Shh signaling in Type B cells as necessary and sufficient to specify the fates of adult-generated olfactory bulb neurons.

Newly generated granule cells exhibit stereotyped morphologies as they differentiate (Petreanu and Alvarez-Buylla, 2002; Carleton *et al.*, 2003; Whitman and Greer, 2007b). During both tangential and radial migration, they appear poorly differentiated with an elongated cell body and both leading and trailing processes. Upon reaching their final destination, the cell body becomes round and there is only a single apical process that will eventually be the primary dendrite. At this point, these neurons receive their first synapse: excitatory synapses formed by axons of other cells on the cell body of the new neurons. These synaptic inputs are thought to modulate granule cell survival and differentiation by providing information from the odor environment as well as from other regions of the brain. The majority of the second wave of synapses are from inhibitory GABAergic inputs. As the nascent dendrite grows into a full dendritic arbor, GABA stabilizes both dendritic growth cones and recently formed dendrite branches. Although the signal transduction elements involved in this dendritic outgrowth remain to be worked out, CREB phosphorylation is increased during this period. Dendritic arbor elaboration is followed by the formation of dendritic spines, which are indicative of the reciprocal dendrodendritic synapses characteristic of mature granule cells that have integrated into local synaptic circuits. Synaptic integration of newborn granule cells thus follows a defined pattern in which these cells first receive synaptic input before providing synaptic information to other cells.

Potential Functions of Adult Neurogenesis

The reason that neurogenesis is maintained specifically in particular regions of the vertebrate brain is unclear. There is, however, an association between adult neurogenesis and learning and memory. Songbirds exhibit a correlation between song learning and adult neurogenesis. Targeted cell death of neurons that are normally replaced in songbirds leads to deterioration of a previously learned song (Scharff *et al.*, 2000). In rodents, learning tasks that require hippocampal function increase the survival rate of adult-born neurons while loss of these neurons impairs performance on hippocampal-dependent tasks (Saxe *et al.*, 2006). In the olfactory system, the young neurons are subject to intense competition to survive such that well over half of the new cohort will die not long after arriving in the olfactory bulb (Petreanu and Alvarez-Buylla, 2002; Winner *et al.*, 2002). Whether a newly born cell survives to join an olfactory bulb circuit is heavily influenced by olfactory experience as well as signals from other regions of the brain (Rocheffort *et al.*, 2002; Mandaïron *et al.*, 2006). Intriguingly, pairing an odorant with a learning paradigm

increases the survival of newborn neurons while the same odor exposure that is not paired with learning has no detectable effect on cell survival (Alonso *et al.*, 2006). Similarly, pregnancy causes an increased rate of neurogenesis and neuronal survival, effects that are mediated through the hormone prolactin (Shingo *et al.*, 2003). The incorporation of young neurons into the existing olfactory circuits therefore is influenced by a variety of factors including the odor environment and the physiological state of the individual.

One reason for incorporating new neurons into olfactory circuits may be that odor responses of newborn neurons are distinct from those of mature neurons. As neurons differentiate and go through the period of rapid synaptogenesis described above, they also exhibit heightened responses to odorants. One indirect measure of neuronal activation is assaying for stimulus-dependent expression of immediate-early genes, which are upregulated in neurons that have recently been activated. By assaying immediate-early gene expression after odor exposure, the population-level response to odor of young neurons exceeds that of the population of mature neurons (Carlén *et al.*, 2002; Magavi *et al.*, 2005). Consistent with this, monitoring responses of individual neurons show that they become more selective as they mature – i.e., each neuron responds to fewer olfactory stimuli with age (Livneh *et al.*, 2014). Further, these young neurons are particularly plastic in their synaptic responses, which likely reflects the high ratios of NMDA glutamate receptors to AMPA glutamate receptors as compared to that seen in mature neurons (Carleton *et al.*, 2003). These findings suggest that the young neurons are in what is essentially a critical period, analogous to the period of time in which organisms are particularly susceptible to environmental influences. Indeed, activation of PGs during the period of robust plasticity alters the response profiles of the same cells once they have matured as compared to those cells that were not stimulated during the same period (Livneh *et al.*, 2014). Thus, incorporation of newborn neurons into olfactory circuits shapes the function of those circuits and does not merely replace cells that were previously present that had been lost to cell death.

Potential Applications

The discovery of cells that continue to proliferate and give rise to new neurons in the adult CNS not only changed the way we understand the brain, but also raised hopes that these findings would impact our treatment of disorders involving a loss of neurons – whether from surgical resection such as in cancer patients or neurodegeneration such as in Parkinson's disease patients. One possibility to treat such patients would be to redirect the cells generated in the V-SVZ to another part of the brain; however, it seems unlikely that these cells would be relevant to other circuits and thus might not be therapeutically beneficial. Another potential approach with greater flexibility is the transplantation of neurons generated from induced pluripotent stem cells (iPSCs). Detailed knowledge of the molecular factors that guide the differentiation of NSCs *in vivo* will enhance the diversity of cells that can be generated from iPSCs *in vitro*. Finally, recent evidence suggests that other regions of the brain in addition to the SGZ and V-SVZ may harbor latent

neurogenic capacity (Magnusson *et al.*, 2014). Inducing neurogenesis from these other progenitors may also alleviate some of the symptoms that arise when large numbers of neurons are lost from the adult brain. Though the study of these questions is still in the early stages, the potential therapeutic implications, particularly of the iPSCs, are very exciting.

Conclusion

In this article, we have focused on neurogenesis in a specific context – the generation in the adult rodent brain of neurons that reside within olfactory circuits. The study of this phenomenon, which had been erroneously thought to be impossible, has offered great insights into how the vertebrate brain works. Indeed, this initially controversial discovery has opened up new lines of thinking about the cellular basis of sensory perception as well as learning and memory. The findings from the study of adult neurogenesis have also informed our understanding of *in vivo* developmental neurogenesis and *in vitro* artificial neurogenesis during iPSC differentiation. Importantly, our understanding of adult neurogenesis has arisen from approaches that span multiple scales from the molecular up through the behavioral. The study of this phenomenon thus serves as an example for how to successfully integrate approaches in order to address a complex biological problem.

References

- Ahn, S., Joyner, A.L., 2005. *In vivo* analysis of quiescent adult neural stem cells responding to sonic hedgehog. *Nature* 437, 894–897. doi:10.1038/nature03994.
- Allen, E., 1912. The cessation of mitosis in the central nervous system of the albino rat. *Journal of Comparative Neurology* 22, 547–568.
- Alonso, M., Viollet, C., Gabellec, M.-M., *et al.*, 2006. Olfactory discrimination learning increases the survival of adult-born neurons in the olfactory bulb. *Journal of Neuroscience* 26, 10508–10513. doi:10.1523/JNEUROSCI.2633-06.2006.
- Altman, J., 1963. Autoradiographic investigation of cell proliferation in the brains of rats and cats. *Anatomical Record* 145, 573–591.
- Altman, J., 1969. Autoradiographic and histological studies of postnatal neurogenesis. IV. Cell proliferation and migration in the anterior forebrain, with special reference to persisting neurogenesis in the olfactory bulb. *Journal of Comparative Neurology* 137, 433–457. doi:10.1002/cne.901370404.
- Altman, J., Das, G.D., 1965. Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats. *Journal of Comparative Neurology* 124, 319–335.
- Altman, J., Das, G.D., 1967. Postnatal neurogenesis in the guinea-pig. *Nature* 214, 1098–1101.
- Anthony, T.E., Klein, C., Fishell, G., Heintz, N., 2004. Radial glia serve as neuronal progenitors in all regions of the central nervous system. *Neuron* 41, 881–890.
- Balordi, F., Fishell, G., 2007a. Mosaic removal of hedgehog signaling in the adult SVZ reveals that the residual wild-type stem cells have a limited capacity for self-renewal. *Journal of Neuroscience* 27, 14248–14259. doi:10.1523/JNEUROSCI.4531-07.2007.
- Balordi, F., Fishell, G., 2007b. Hedgehog signaling in the subventricular zone is required for both the maintenance of stem cells and the migration of newborn neurons. *Journal of Neuroscience* 27, 5936–5947. doi:10.1523/JNEUROSCI.1040-07.2007.
- Belvindrah, R., Hankel, S., Walker, J., Patton, B.L., Müller, U., 2007. $\beta 1$ Integrins control the formation of cell chains in the adult rostral migratory stream. *Journal of Neuroscience* 27, 2704–2717. doi:10.1523/JNEUROSCI.2991-06.2007.
- Bolteus, A.J., Bordey, A., 2004. GABA release and uptake regulate neuronal precursor migration in the postnatal subventricular zone. *Journal of Neuroscience* 24, 7623–7631. doi:10.1523/JNEUROSCI.1999-04.2004.
- Cameron, H.A., McKay, R.D., 2001. Adult neurogenesis produces a large pool of new granule cells in the dentate gyrus. *Journal of Comparative Neurology* 435, 406–417.
- Cameron, H.A., Woolley, C.S., McEwen, B.S., Gould, E., 1993. Differentiation of newly born neurons and glia in the dentate gyrus of the adult rat. *Neuroscience* 56, 337–344.
- Carlén, M., Cassidy, R.M., Brismar, H., *et al.*, 2002. Functional integration of adult-born neurons. *Current Biology* 12, 606–608.
- Carleton, A., Petreanu, L.T., Lansford, R., Alvarez-Buylla, A., Lledo, P.-M., 2003. Becoming a new neuron in the adult olfactory bulb. *Nature Neuroscience* 6, 507–518. doi:10.1038/nn1048.
- Chazal, G., Durbec, P., Jankovski, A., Rougon, G., Cremer, H., 2000. Consequences of neural cell adhesion molecule deficiency on cell migration in the rostral migratory stream of the mouse. *Journal of Neuroscience* 20, 1446–1457.
- Deng, W., Aimone, J.B., Gage, F.H., 2010. New neurons and new memories: How does adult hippocampal neurogenesis affect learning and memory? *Nature Reviews Neuroscience*, 1–12. doi:10.1038/nrn2822.
- Doetsch, F., Caille, I., Lim, D.A., Garcia-Verdugo, J.M., Alvarez-Buylla, A., 1999. Subventricular zone astrocytes are neural stem cells in the adult mammalian brain. *Cell* 97, 703–716.
- Doetsch, F., Petreanu, L., Caille, I., Garcia-Verdugo, J.M., Alvarez-Buylla, A., 2002. EGF converts transit-amplifying neurogenic precursors in the adult brain into multipotent stem cells. *Neuron* 36, 1021–1034.
- Emsley, J.G., Hagg, T., 2003. $\alpha 6 \beta 1$ Integrin directs migration of neuronal precursors in adult mouse forebrain. *Experimental Neurology* 183, 273–285. doi:10.1016/S0014-4886(03)00209-7.
- Eriksson, P.S., Perfilieva, E., Björk-Eriksson, T., *et al.*, 1998. Neurogenesis in the adult human hippocampus. *Nature Medicine* 4, 1313–1317. doi:10.1038/3305.
- Fuccillo, M., Joyner, A.L., Fishell, G., 2006. Morphogen to mitogen: The multiple roles of hedgehog signalling in vertebrate neural development. *Nature Reviews Neuroscience* 7, 772–783. doi:10.1038/nrn1990.
- Gage, F.H., Kempermann, G., Palmer, T.D., Peterson, D.A., Ray, J., 1998. Multipotent progenitor cells in the adult dentate gyrus. *Journal of Neurobiology* 36, 249–266.
- Garcia, A.D.R., Doan, N.B., Imura, T., Bush, T.G., Sofroniew, M.V., 2004. GFAP-expressing progenitors are the principal source of constitutive neurogenesis in adult mouse forebrain. *Nature Neuroscience* 7, 1233–1241. doi:10.1038/nn1340.
- Goldman, S.A., Nottebohm, F., 1983. Neuronal production, migration, and differentiation in a vocal control nucleus of the adult female canary brain. *Proceedings of the National Academy of Sciences of the United States of America* 80, 2390–2394.
- Gould, E., McEwen, B.S., Tanapat, P., Galea, L.A., Fuchs, E., 1997. Neurogenesis in the dentate gyrus of the adult tree shrew is regulated by psychosocial stress and NMDA receptor activation. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 17, 2492–2498.
- Hack, I., Bancila, M., Loulier, K., Carroll, P., Cremer, H., 2002. Reelin is a detachment signal in tangential chain-migration during postnatal neurogenesis. *Nature Neuroscience* 5, 939–945. doi:10.1038/nn923.
- Hu, H., 2000. Polysialic acid regulates chain formation by migrating olfactory interneuron precursors. *Journal of Neuroscience Research* 61, 480–492.
- Ihrle, R.A., Alvarez-Buylla, A., 2011. Lake-front property: A unique germinal niche by the lateral ventricles of the adult brain. *Neuron* 70, 674–686. doi:10.1016/j.neuron.2011.05.004.
- Imayoshi, I., Sakamoto, M., Yamaguchi, M., Mori, K., Kageyama, R., 2010. Essential roles of notch signaling in maintenance of neural stem cells in developing and adult brains. *Journal of Neuroscience* 30, 3489–3498. doi:10.1523/JNEUROSCI.4987-09.2010.
- Imura, T., Kornblum, H.I., Sofroniew, M.V., 2003. The predominant neural stem cell isolated from postnatal and adult forebrain but not early embryonic forebrain expresses GFAP. *Journal of Neuroscience* 23, 2824–2832.
- Imura, T., Nakano, I., Kornblum, H.I., Sofroniew, M.V., 2006. Phenotypic and functional heterogeneity of GFAP-expressing cells *in vitro*: Differential expression of LeX/CD15 by GFAP-expressing multipotent neural stem cells and non-neurogenic astrocytes. *Glia* 53, 277–293. doi:10.1002/glia.20281.
- Kirschenbaum, B., Doetsch, F., Lois, C., Alvarez-Buylla, A., 1999. Adult subventricular zone neuronal precursors continue to proliferate and migrate in the absence of the olfactory Bulb. *Journal of Neuroscience* 19, 2171–2180.
- Kriegstein, A., Alvarez-Buylla, A., 2009. The glial nature of embryonic and adult neural stem cells. *Annual Review of Neuroscience* 32, 149–184. doi:10.1146/annurev.neuro.051508.135600.
- Laywell, E.D., Rakic, P., Kukekov, V.G., Holland, E.C., Steindler, D.A., 2000. Identification of a multipotent astrocytic stem cell in the immature and adult

- mouse brain. *Proceedings of the National Academy of Sciences of the United States of America* 97, 13883–13888. doi:10.1073/pnas.250471697.
- Livneh, Y., Adam, Y., Mizrahi, A., 2014. Odor processing by adult-born neurons. *Neuron* 81, 1097–1110. doi:10.1016/j.neuron.2014.01.007.
- Lois, C., García-Verdugo, J.M., Alvarez-Buylla, A., 1996. Chain migration of neuronal precursors. *Science* 271, 978–981.
- Magavi, S.S.P., Mitchell, B.D., Szentirmai, O., Carter, B.S., Macklis, J.D., 2005. Adult-born and preexisting olfactory granule neurons undergo distinct experience-dependent modifications of their olfactory responses *in vivo*. *Journal of Neuroscience* 25, 10729–10739. doi:10.1523/JNEUROSCI.2250-05.2005.
- Magnusson, J.P., Göritz, C., Tatarishvili, J., *et al.*, 2014. A latent neurogenic program in astrocytes regulated by notch signaling in the mouse. *Science* 346, 237–241. doi:10.1126/science.1246206.237.
- Mandairon, N., Jourdan, F., Didier, A., 2003. Deprivation of sensory inputs to the olfactory bulb up-regulates cell death and proliferation in the subventricular zone of adult mice. *Neuroscience* 119, 507–516. doi:10.1016/S0306-4522(03)00172-6.
- Mandairon, N., Sacquet, J., Jourdan, F., Didier, A., 2006. Long-term fate and distribution of newborn cells in the adult mouse olfactory bulb: Influences of olfactory deprivation. *Neuroscience* 141, 443–451. doi:10.1016/j.neuroscience.2006.03.066.
- Menn, B., 2006. Origin of oligodendrocytes in the subventricular zone of the adult brain. *Journal of Neuroscience* 26, 7907–7918. doi:10.1523/JNEUROSCI.1299-06.2006.
- Merkle, F.T., Fuentealba, L.C., Sanders, T.A., *et al.*, 2014. Adult neural stem cells in distinct microdomains generate previously unknown interneuron types. *Nature Neuroscience* 17, 207–214. doi:10.1038/nn.3610.
- Merkle, F.T., Mirzadeh, Z., Alvarez-Buylla, A., 2007. Mosaic organization of neural stem cells in the adult brain. *Science* 317, 381–384. doi:10.1126/science.1144914.
- Merkle, F.T., Tramontin, A.D., García-Verdugo, J.M., Alvarez-Buylla, A., 2004. Radial glia give rise to adult neural stem cells in the subventricular zone. *Proceedings of the National Academy of Sciences of the United States of America* 101, 17528–17532. doi:10.1073/pnas.0407893101.
- Murase, S., Horwitz, A.F., 2002. Deleted in colorectal carcinoma and differentially expressed integrins mediate the directional migration of neural precursors in the rostral migratory stream. *Journal of Neuroscience* 22, 3568–3579.
- Nguyen-Ba-Charvet, K.T., Picard-Riera, N., Tessier-Lavigne, M., *et al.*, 2004. Multiple roles for slits in the control of cell migration in the rostral migratory stream. *Journal of Neuroscience* 24, 1497–1506. doi:10.1523/JNEUROSCI.4729-03.2004.
- Noctor, S.C., Flint, A.C., Weissman, T.A., Dammerman, R.S., Kriegstein, A.R., 2001. Neurons derived from radial glial cells establish radial units in neocortex. *Nature* 409, 714–720.
- Palma, V., Lim, D.A., Dahmane, N., *et al.*, 2005. Sonic hedgehog controls stem cell behavior in the postnatal and adult brain. *Development* 132, 335–344. doi:10.1242/dev.01567.
- Petreanu, L., Alvarez-Buylla, A., 2002. Maturation and death of adult-born olfactory bulb granule neurons: Role of olfaction. *Journal of Neuroscience* 22, 6106–6113.
- Petrova, R., Garcia, A.D., Joyner, A.L., 2013. Titration of GLI3 repressor activity by sonic hedgehog signaling is critical for maintaining multiple adult neural stem cell and astrocyte functions. *Journal of Neuroscience* 33, 17490–17505. doi:10.1523/JNEUROSCI.2042-13.2013.
- Piccin, D., Morshead, C.M., 2011. Wnt signaling regulates symmetry of division of neural stem cells in the adult brain and in response to injury. *Stem Cells* 29, 528–538. doi:10.1002/stem.589.
- Reynolds, B.A., Weiss, S., 1992. Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science* 255, 1707–1710.
- Rochefort, C., Gheusi, G., Vincent, J.-D., Lledo, P.-M., 2002. Enriched odor exposure increases the number of newborn neurons in the adult olfactory bulb and improves odor memory. *Journal of Neuroscience* 22, 2679–2689.
- Saghatelyan, A., de Chevigny, A., Schachner, M., Lledo, P.-M., 2004. Tenascin-R mediates activity-dependent recruitment of neuroblasts in the adult mouse forebrain. *Nature Neuroscience* 7, 347–356. doi:10.1038/nn1211.
- Sawamoto, K., Wichterle, H., Gonzalez-Perez, O., *et al.*, 2006. New neurons follow the flow of cerebrospinal fluid in the adult brain. *Science* 311, 629–632. doi:10.1126/science.1119133.
- Saxe, M.D., Battaglia, F., Wang, J.-W., *et al.*, 2006. Ablation of hippocampal neurogenesis impairs contextual fear conditioning and synaptic plasticity in the dentate gyrus. *Proceedings of the National Academy of Sciences of the United States of America* 103, 17501–17506. doi:10.1073/pnas.0607207103.
- Scharff, C., Kirm, J.R., Grossman, M., Macklis, J.D., Nottebohm, F., 2000. Targeted neuronal death affects neuronal replacement and vocal behavior in adult songbirds. *Neuron* 25, 481–492.
- Seri, B., Garcia-Verdugo, J.M., McEwen, B.S., Alvarez-Buylla, A., 2001. Astrocytes give rise to new neurons in the adult mammalian hippocampus. *Journal of Neuroscience* 21, 7153–7160.
- Shingo, T., Gregg, C., Enwere, E., *et al.*, 2003. Pregnancy-stimulated neurogenesis in the adult female forebrain mediated by prolactin. *Science* 299, 117–120. doi:10.1126/science.1076647.
- Sirko, S., Behrendt, G., Johansson, P.A., *et al.*, 2013. Reactive glia in the injured brain acquire stem cell properties in response to sonic hedgehog glia. *Stem Cell* 12, 426–439. doi:10.1016/j.stem.2013.01.019.
- Soria, J.M., Tagliatela, P., Gil-Perotin, S., *et al.*, 2004. Defective postnatal neurogenesis and disorganization of the rostral migratory stream in absence of the *vax1* homeobox gene. *Journal of Neuroscience* 24, 11171–11181. doi:10.1523/JNEUROSCI.3248-04.2004.
- Whitman, M.C., Greer, C.A., 2007a. Adult-generated neurons exhibit diverse developmental fates. *Developmental Neurobiology* 67, 1079–1093. doi:10.1002/dneu.20389.
- Whitman, M.C., Greer, C.A., 2007b. Synaptic integration of adult-generated olfactory bulb granule cells: Basal axodendritic centrifugal input precedes apical dendrodendritic local circuits. *Journal of Neuroscience* 27, 9951–9961. doi:10.1523/JNEUROSCI.1633-07.2007.
- Winner, B., Cooper-Kuhn, C.M., Aigner, R., Winkler, J., Kuhn, H.G., 2002. Long-term survival and cell death of newly generated neurons in the adult rat olfactory bulb. *European Journal of Neuroscience* 16, 1681–1689. doi:10.1046/j.1460-9596.2002.02238.x.
- Young, K.M., Fogarty, M., Kessaris, N., Richardson, W.D., 2007. Subventricular zone stem cells are heterogeneous with respect to their embryonic origins and neurogenic fates in the adult olfactory bulb. *Journal of Neuroscience* 27, 8286–8296. doi:10.1523/JNEUROSCI.0476-07.2007.

Further Reading

- Aimone, J.B., Li, Y., Lee, S.W., *et al.*, 2014. Regulation and function of adult neurogenesis: From genes to cognition. *Physiological Reviews* 94, 991–1026. doi:10.1152/physrev.00004.2014.
- Encinas, J.M., Sierra, A., Valcárcel-Martín, R., Martín-Suárez, S., 2013. A developmental perspective on adult hippocampal neurogenesis. *International Journal of Developmental Neuroscience* 31, 640–645. doi:10.1016/j.ijdevneu.2013.04.001.
- Greig, L.C., Woodworth, M.B., Galazo, M.J., Padmanabhan, H., Macklis, J.D., 2013. Molecular logic of neocortical projection neuron specification, development and diversity. *Nature Reviews Neuroscience* 14, 755–769. doi:10.1038/nrn3586.
- Lepousez, G., Valley, M.T., Lledo, P.-M., 2013. The impact of Adult neurogenesis on olfactory bulb circuits and computations. *Annual Review of Physiology* 75, 339–363. doi:10.1146/annurev-physiol-030212-183731.

HORIZONTAL INTEGRATION: NETWORKS

Understanding of ‘Networks’ *In Vitro* and/or *In Vivo*

EM Schoof, University Health Network, Toronto, ON, Canada

J Ferkinghoff-Borg and R Linding, University of Copenhagen (UCPH), Copenhagen, Denmark

© 2016 Elsevier Inc. All rights reserved.

Introduction

In this article we shall explore the requirement for and tools relevant to the understanding of networks in varying biological contexts from cellular (*in vitro*) to living organisms (*in vivo*). We will focus on molecular networks and pay particular attention to protein-signaling networks. To avoid jumping from one technological wonder to the next we will use the example of cancer, and in particular tumor biology, as a constraint set to describe the need to understand networks in different biological contexts and how comprehensive and large-scale quantitative technologies, combined with sophisticated computational modeling, is crucial for both basic biology and cancer therapy. Ever since its conceptual discovery as early as 3000 BC in ancient Egypt (Society, 2012), cancer has been eluding major scientific breakthroughs in terms of finding a cure. Despite large-scale efforts such as the ‘War on Cancer’ (Hanahan, 2013) in the twentieth century, and estimated yearly research budgets as high as 14 billion Euros worldwide (Eckhouse *et al.*, 2008), the disease continues to have a devastating effect on people’s health, with often lethal consequences. For cancer patients in the United States alone, the financial burden of healthcare ranges was estimated at US\$88.7 billion in 2011. In this article, a brevilouquent overview of cancer will be provided, focusing mainly on conceptual challenges that have arisen over the years and scientific approaches that have been developed in an attempt to tackle some of the many aspects of the disease. Specifically, the significance of moving from a pathway-centric interpretation of cellular signaling toward a more global signaling network approach will be reviewed and discussed. Next, new technological platforms from which to study integrated biological functions will be introduced, with a special focus on Network Biology and technologies fundamental to this such as mass spectrometry (MS), next-generation sequencing (NGS), siRNA-based high content screening (HCS), and computational modeling to allow the integration of different types of data. Together, it is envisioned that these genome-scale technologies will provide quantitative data that can form the basis of computational integrative modeling and biological forecasting to predict cellular behavior in both *in vitro* and *in vivo* biological contexts.

The Complex Nature of a Disease and Its Underlying (*In Vitro*) Molecular Networks

While extensive strides have been made in terms of defining biological characteristics that drive a tumor cell toward

malignancy, there is still great debate about the precise underlying causes at each stage. As will be discussed throughout this article, several biological phenomena may be fundamental to all human diseases, ranging from mutations occurring in the genome of the diseased cells (Burrell *et al.*, 2013; Stephens *et al.*, 2012; Vogelstein and Kinzler, 2004; Vogelstein *et al.*, 2013; Stratton *et al.*, 2009), through epigenetic regulation of DNA transcription (Dawson and Kouzarides, 2012; Feinberg and Tycko, 2004), to dysregulated protein network dynamics (Cox and Mann, 2007; Creixell *et al.*, 2012; Pawson and Hunter, 1994; Pawson and Kofler, 2009; Tan *et al.*, 2009; Taylor *et al.*, 2009; Taylor and Wrana, 2012; Vidal *et al.*, 2011; Pawson and Linding, 2008). Whilst each of these may contribute individually, it is likely that a combination of these aspects is what ultimately drives cancer and adds to the complexity of understanding a given disease phenotype. For example, it is known that more than 50% of human melanomas contain an activating mutation in the BRAF kinase gene, where the valine in amino acid position 600 is substituted with a glutamate (BRAF V600E), which, through structural effects, renders the kinase constitutively active (Davies *et al.*, 2002). This permanent kinase signaling results in a constitutive activation of the Raf to mitogen-activated protein (MAP) kinase pathway, significantly increasing the rate of mitogenesis of these cells (Davies and Samuels, 2010). While inhibition of this mutant-specific variant of BRAF often results in a temporary response to treatment in the clinic (Flaherty *et al.*, 2010), resistance ultimately develops as the signaling networks normally utilizing the BRAF kinase pathway are rewired to circumvent the BRAF inhibition (Poulidakos and Rosen, 2011; Wagle *et al.*, 2011). Currently, the use of additional inhibitors to subsequently target the rewired cellular signaling networks is being investigated, with some initial success (Lito *et al.*, 2012). This is a prime example of how mutations at the genome level exert their effect at the protein-signaling level, and underlines the importance of investigating both genomic and proteomic aspects of a cell when trying to gain a thorough understanding of a given disease phenotype. In addition, while proteins are the cell’s functional effectors and almost exclusively the targets of small-molecule inhibitors and antibodies, understanding the underlying genomes can help pinpoint appropriate protein targets and aid in elucidating molecular mechanisms of a particular disease. Especially considering the fact that “no single gene defect ‘causes’ cancer” (Vogelstein and Kinzler, 2004), understanding how proteins interact with one another in a signaling network

through proteomic analysis can help prioritize which mutations may act in a coherent manner and should be further investigated in combination for therapeutic purposes (Yaffe, 2013). While the complexity is extremely high at the genomic level, given the similar phenotypic states of cancer cells (increased proliferation, migration etc.), the complexity might be better understood at the protein-signaling network level. In other words, while the mutational landscapes might differ greatly between tumors, their functional impact on the signaling networks within the cell might be less varied, thereby eliciting a similar phenotype. This underlines the need for assessing the impact of mutations at the protein level (Creixell *et al.*, 2012).

Tumor Microenvironment – A Biological Reality that Can Only Be Understood *In Vivo* and Using Global Integrative and Quantitative Technologies

A biological aspect this section will briefly touch upon is the role of the microenvironment, as extensive evidence has demonstrated that it is not only the tumor cells themselves which decide the fate of pathogenesis, but rather that it is the complex interplay of the tumor cells with their surrounding microenvironment and other cell types contained therein (see Figure 1; Hanahan and Weinberg, 2011). In healthy tissue, the stroma (consisting of extracellular matrix (ECM) and other cells such as endothelial cells, fibroblasts, etc.) is generally considered the supportive structure of a given tissue or organ,

and maintains a natural barrier against tumorigenesis. When tumor cells appear however, this process initiates a cascade of changes, often resulting in the stroma becoming an environment that supports cancer progression. This involves, amongst others, the activation of fibroblasts and matrix remodeling, and additionally, microenvironmental stimuli such as hypoxia and acidity, and growth factors will vary within a tumor, giving rise to additional heterogeneous cell populations, each adapted to their specific microenvironment. These microenvironmental factors are likely to affect both the signaling network states of tumor cells contained within, and may even select for specific DNA mutations that give certain subpopulations an evolutionary benefit depending on their surroundings.

A particular cell type that has been associated extensively with cancer progression, and thereby merits scientific investigation, is the cancer-associated fibroblast (CAF). Where normal fibroblasts typically have a suppressive effect on tumorigenesis (Dotto *et al.*, 1988), CAFs can significantly enhance tumor progression (Orimo *et al.*, 2005). They distinguish themselves from normal fibroblasts by displaying increased proliferation levels, enhanced cytokine secretion, and elevated ECM production and increased contractility (Polanska and Orimo, 2013). These differences result in extensive tissue remodeling, predominantly around the tumor cells, driven extensively by, for example, elevated expression of matrix metalloproteases, pro-angiogenic factors, matrix cross-linkers such as lysyl oxidase (LOX) and transglutaminases, and growth factors (Bergers *et al.*, 2000; Erler *et al.*, 2006). The clinical relevance of CAFs has been established in several ways. For example, the abundance of CAFs has been

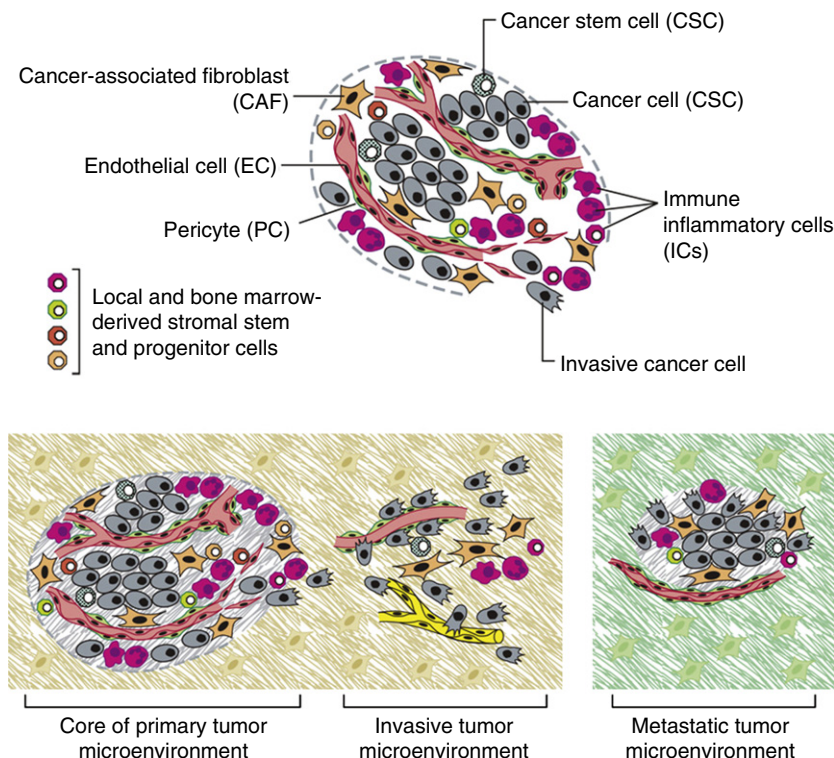


Figure 1 Schematic representation of the key cellular components of the tumor microenvironment and cellular components contributing to tumor heterogeneity. In invasive and metastatic tumor microenvironment, there is an increase in number of invasive cells, contributing to the overall tumorigenesis. Reproduced from Hanahan, D., Weinberg, R.A., 2011. Hallmarks of cancer: The next generation. Cell 144, 646–674.

shown to be predictive of prognosis for both breast and pancreatic cancer (Fujita *et al.*, 2010; Yamashita *et al.*, 2012). Additionally, elevated matrix metalloprotease levels, which are secreted both by tumor cells and CAFs, have also been linked with increased tumor aggressiveness and poor prognosis (Vihinen and Kahari, 2002). Furthermore, CAFs have been associated with emerging treatment resistance (Meads *et al.*, 2009). For example, in BRAF(V600E) cell lines, cocultured with CAFs, BRAF inhibition was overcome by secreted hepatocyte growth factor (HGF), which elicited increased phosphorylation of MET, a cognate receptor of BRAF (Straussman *et al.*, 2012; Wilson *et al.*, 2012). This has also been observed in patients, where HGF expression levels positively correlated with reduced drug treatment response (Straussman *et al.*, 2012). These studies highlight the importance of studying cancer cell signaling within the appropriate context, as the *in vivo* relevance of *in vitro* derived results will be limited.

To add to the complexity, the influence of the micro-environment on tumorigenesis is a bidirectional process. As was recently shown by Barker *et al.*, LOX-like 2 (LOXL2) is an enzyme secreted by tumor cells, which in turn activates the CAFs through integrin-mediated focal adhesion kinase activation. By blocking LOXL2, the authors reduced the activation of stromal host cells, and significantly decreased tumor cell invasion (Barker *et al.*, 2013). Similarly, Cox *et al.* (2013) investigated the effect of the microenvironment on the metastatic potential of a tumor. They found that the expression of LOX, an enzyme that catalyzes the covalent cross-linking of collagen I and which has been implicated in cell invasion and malignant progression (Bissell and Labarge, 2005; Erler *et al.*, 2006), is "favourable to colonization and growth of metastasising tumor cells" (Cox *et al.*, 2013). By specifically blocking LOX activity using an antibody, they were able to reduce the ECM modifications that normally have an enhancing effect on tumor cell survival and metastasis. As LOX is secreted from the 4T1 cell line used in the study, this elegantly highlights the complex bidirectional interplay of tumor cells and their surrounding microenvironment. Not only does the targeting of proteins secreted by the tumor cells seem like a viable therapeutic strategy though, as work by Luga *et al.* (2012) demonstrated that targeting CAF-secreted exosomes (using Cd81-specific siRNAs) has a beneficial effect on specifically suppressing lung metastases in a breast cancer model. Combined, these *in vivo*-derived results suggest that targeting both the tumor cells and surrounding stromal cells are attractive therapeutic strategies, and that more extensive investigation is merited to uncover additional molecular mechanisms driving tumorigenesis, both in tumor and their surrounding stromal cells.

In summary, these first two aspects underline the importance of moving away from viewing a tumor (the biological system) as a homogeneous population that can be treated as a whole, as it is evident that the complexity is much larger than initially assumed. Only by biologically characterizing the contribution and integration of individual components that make up a tumor, both separately and collectively, and understanding how they can be therapeutically targeted, it is likely that cancer research will start fulfilling its promise of successful targeted therapies. In other words understanding how the underlying molecular networks operate in an integrative manner to control the tumor's evolution and interaction with the host organisms

requires studies of such networks at different biological context levels ranging from cell culture (*in vitro*) to animal models (*in vivo*) and eventually in patients. The critical goal of network biology is thus to provide predictive capabilities in these more and more complex environments. Deriving computational models that can perform such forecasting of how cells react to cues/environment even in one context is challenging but vital for our basic understanding of how information is processed in biological systems.

To drive the message home loud and clear, we shall end this intermezzo of cancer biology and why it is one of the greatest challenges to our understanding of molecular networks and how they process information, by revisiting the topic of cancer metastasis. Arguably one of the most complex biological processes we know of, this process is a highly complex graph-like evolutionary path that many cancer cells undertake in a complex interaction with other cells and the host tissues and immune defense. This is not only a daunting task to grasp from a 30 000 feet view, but even more so at its primordial state in terms of our understanding of the cellular signaling networks involved.

Cancer Metastasis – A Showstopper for 1-Context-Only Studies/Theories

As part of the cancer evolution, the cascade of tumorigenesis often ends at the metastatic stage, at which point the tumor cells, having acquired the necessary mutations and other phenotypic traits (e.g., invasiveness, enhanced proliferative capacity, etc.), are able to colonize other organs within the body. This spread of cancer cells to other, often vital organs, is responsible for 90% of all cancer-related patient mortality, and, if targeted successfully, represents the hallmark with most therapeutic potential (Gupta and Massague, 2006). In this section, we will explore some of the underlying biological principles, challenges associated with characterizing metastatic cells and possible directions for gaining a better understanding of how to tackle metastatic disease.

The process of metastasis is often classified into a series of basic biological events, which are portrayed in Figure 2 (Erler and Giaccia, 2008). The steps involve (1) local invasion, (2) intravasation (entry into the bloodstream), (3) extravasation (exit from the bloodstream), and (4) colonization of the distant tissue, from which further metastases can also be spawned (Chambers *et al.*, 2002; Fidler, 2003). While each of these steps represents a potential opportunity for therapeutic intervention to varying degrees, it is important to consider the origin of these traits and how tumor cells obtain these capabilities, as this may lay the foundation for successful prevention; this will be further explored in the next section. As previously explained, the concept of cancers consisting of a heterogeneous population of tumor and stromal cells has now been widely accepted, each cell lineage contributing to the development process. Furthermore, research conducted in the last decade has suggested that several oncogenic events during cancer development may contribute to the evolution of tumors. Especially the ability to resist growth suppression and bypass DNA-damage-checkpoints are critical properties, as this allows for the generation of genomic instability (Hernando *et al.*, 2004; Maser and DePinho, 2002; Puc *et al.*, 2005;

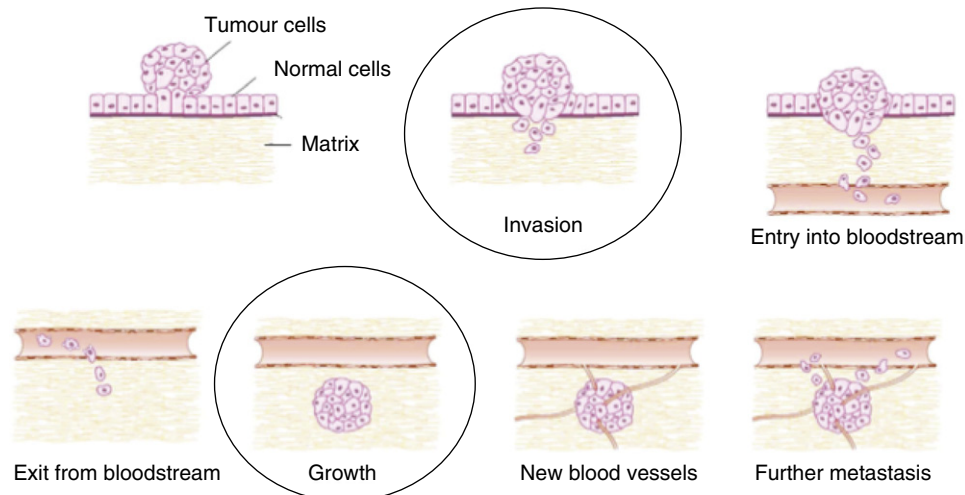


Figure 2 The process of metastasis, depicting the different biological processes tumor cells must undergo to successfully colonize distant tissue sites. Adapted from Erler, J.T., Giaccia, A.J., 2008. Chapter 3: Critical Steps in Cancer Metastasis. In *Clinical Oncology*.

Gorgoulis *et al.*, 2005). Particularly in the case of colorectal cancer (CRC) pioneering work by Vogelstein and colleagues was able to show that colorectal tumorigenesis largely relies on the linear accumulation of key mutations (Fearon and Vogelstein, 1990). Through this, tumor cells are exposed to many evolutionary paths, some of which may give a specific subpopulation, for example, a fitness advantage or the phenotypes required for metastatic progression. Additionally, the genetic and phenotypic diversity that thereby exists within a tumor may also explain why, despite millions of cells being shredded by a tumor into the bloodstream every day, only a very small subset of these will successfully colonize distant tissue (Fidler, 1970; Kim *et al.*, 2004; Luzzi *et al.*, 1998). Part of the reason for this high attrition rate of distant organ colonization is the fact that healthy tissue generally provides an inhospitable environment for invading tumor cells, requiring a significant level of resilience to exist in metastatic tumor cells to overcome this. As originally postulated by Paget (1889), the spread of metastasis is dependent on both the primary tumor (the 'seed') and the distant site (the 'soil'). For an extensive review, the reader is referred to (Fidler, 2003), but it is interesting to note that the metastatic progression has indeed been shown to depend on both the primary and secondary tissue sites (Gupta and Massague, 2006; Nguyen *et al.*, 2009). Thus, it becomes clear that it is beneficial for a tumor to contain specific subpopulations of tumor cells, all bearing different pheno- and genotypes, allowing the successful growth of the primary tumor and distant metastases in their respective environments. Pioneering work conducted by Fidler (1970) demonstrated that within a tumor, rare clones exist that, through evolutionary processes, acquired the necessary properties which allowed them to drive metastatic progression. Later work revealed that highly metastatic subpopulations display a higher level of genetic mutability than their non-metastatic counterparts from the same tumor, in addition to biological heterogeneity being observed both within a single metastasis ('intra-lesional' heterogeneity) and among different metastases ('inter-lesional' heterogeneity), which supports the link between metastasis and genetic

instability/evolution (Fidler, 2003). Overall, these results seem to suggest that cancer progression is a function of heterogeneous cell populations being driven to evolve through sequential environmental stimuli and pressures (Gupta and Massague, 2006).

For these reasons, it is important to consider the underlying molecular processes which facilitate metastatic progression and the above-mentioned biological properties. Moreover, obtaining a comprehensive understanding of which proteins seem to play a fundamental role, and in particular which stages of tumorigenesis, is of critical importance, as they may each pose possible therapeutic targets. *In vitro* studies provide good utility for studying metastatic processes in isolation and laying the foundation for the molecular landscapes potentially involved; however, given the cellular interactions required for the process of metastasis, studying the process *in vivo*, while experimentally challenging, is imperative to increase the chances of therapeutic success. Kinases have been shown to be involved in many metastasis-promoting cell behaviors such as cell proliferation, migration, survival, and invasion (Baker *et al.*, 2013a,b; Brognard and Hunter, 2011; Cox *et al.*, 2013; Criscuolo *et al.*, 2005; Engelman *et al.*, 2007; Hiratsuka *et al.*, 2002; Lee *et al.*, 2012; Regan Anderson *et al.*, 2013; Jorgensen *et al.*, 2009), and thereby form attractive therapeutic targets (Davis *et al.*, 2011; Fedorov *et al.*, 2010; Janne *et al.*, 2009; Karaman *et al.*, 2008). In fact, it has recently been shown that kinases are the most frequently mutated proteins in tumors, underlining the potential therapeutic implications they represent (Lin *et al.*, 2007; Wood *et al.*, 2007). Given the diversification that tumor cells undergo throughout metastatic progression, and the ubiquitous involvement of kinase signaling in cellular signal processing (Linding *et al.*, 2007; Pawson and Kofler, 2009; Macurek *et al.*, 2008; Yaffe and Cantley, 1999), it is expected that kinase activities will be altered during this process as well. Therefore, obtaining a global overview of which kinases and other proteins are dysregulated during tumorigenesis is of great importance, also considering the plethora of inhibitors available for this group of proteins.

Networks in Cancer

Principles of Network Biology

The last decade has seen a considerable shift of moving away from a 'one drug, one target' approach (Beadle and Tatum, 1941), as, especially in the case of complex diseases such as cancer, deploying highly specific compounds targeting a single molecular entity has led to extensive treatment failure (Chandarlapaty *et al.*, 2011; Huang *et al.*, 2007b; Schoeberl *et al.*, 2009). Traditionally, the single-target approach focused primarily on individual 'pathways' that were found to have been dysregulated in disease. However, it has become increasingly clear that the proteins making up a specific pathway have many interactions outside their respective pathways, and it is therefore imperative to gain a better understanding of how individual pathways are assembled into higher order networks (Vidal *et al.*, 2011). This aspect is further highlighted by the genetic complexity and heterogeneity of tumors, as the mutations spanning a large multitude of proteins will have a different effect on the protein dynamics, yet still result in similar disease phenotypes. In other words, several genotypes can give rise to the same phenotype, while a single genotype can also give rise to several phenotypes, depending on the cellular context. These 'genotype-to-phenotype' relationships are one of the fundamental aspects the field of network biology is aiming to resolve, as, being the cellular effectors, protein signaling plays a critical role in defining the link between genotype and phenotype (see Figure 3; Pawson and Linding, 2008; Vidal *et al.*, 2011; Cox and Mann, 2007; Creixell *et al.*, 2012; Gstaiger and Aebersold, 2013; Bakal *et al.*, 2007). Rather than characterizing signaling pathways in isolation, it is therefore of vital importance to comprehensively measure genome-wide protein levels within the cell, as this allows the characterization of their global network structure and

dynamics and how genomic information is propagated toward cellular phenotype (Nurse, 2003; Vidal, 2009; Vidal *et al.*, 2011; Taylor and Wrana, 2012; Gstaiger and Aebersold, 2009). Additionally, it has been proposed that it is not only individual protein species that make up a signaling network that can be therapeutically targeted; instead the structure and dynamics of a particular signaling network poses a powerful drug target by modulating the information flow through such a network (Pawson and Linding, 2008). This was further exemplified by work demonstrating that cells utilize the same protein network components for different cellular responses, and instead alter the network utilization of these to elicit a specific response. This highlights the importance of identifying how cellular networks are deployed specifically in a given disease condition, in order to be able to target the dynamics through disease-specific perturbations (Miller-Jensen *et al.*, 2007; Huang *et al.*, 2007a,b). Furthermore, by measuring global protein dynamics within the cell, a more comprehensive interrogation of how the effect of a perturbation affects the signaling network as a whole can be achieved. The importance of this was first elegantly demonstrated through seminal work by Janes *et al.* (2005). The authors investigated whether the phosphorylation state of Jun-activate kinase (JNK) was indicative of pro-apoptotic or anti-apoptotic activity, and concluded that the phosphorylation state alone did not suffice to characterize the phenotypic effect of JNK. Instead, they were able to show that activation of JNK could lead to both apoptosis and proliferation, and that the outcome was dependent on the prior signaling network state at the time of activation. This landmark study conclusively proved that studying the activity of a single protein in isolation is not sufficient to accurately characterize its cellular role, and rather, that the contextual network it is operating within also needs to be interrogated in order to obtain a correct readout. Similarly, utilizing this contextual information allows the characterization of the signaling network rewiring that occurs in response to a perturbation, in order to target these altered network states with an additional perturbation. The power of this type of approach was demonstrated by Lee *et al.* (2012), where they were able to greatly increase the level of apoptosis in breast cancer cells by designing a time-staggered combination treatment in a data-driven way.

In addition to protein dynamics (i.e., expression levels) altering cellular phenotype, extensive research has demonstrated that posttranslational modifications (PTMs) such as phosphorylation also play a crucial role in controlling cellular behavior (Jorgensen and Linding, 2008; Jorgensen *et al.*, 2009; Lemmon and Schlessinger, 2010; Pawson and Kofler, 2009). The importance of phosphorylation is further supported by the fact that by 2010, 149 inhibitors targeting 42 kinases (the 'writers' of phosphorylation modifications) have been subjected to clinical testing, highlighting the therapeutic potential of interfering with PTM signaling (Fedorov *et al.*, 2010). A significant challenge with deciphering phosphorylation-based signaling, however, is the derivation of kinase-substrate interactions; due to the highly transient nature of kinase-substrate interactions, it is inherently challenging to experimentally determine which kinase(s) is responsible for a given phosphorylation site. Therefore, generally, a combination of experimental and computational analysis is required,

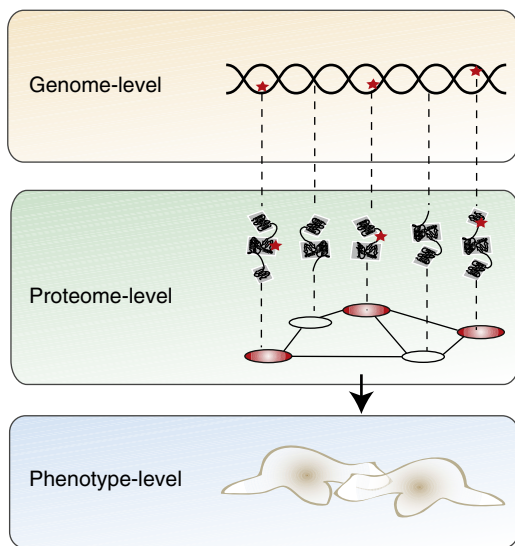


Figure 3 The genotype-to-phenotype relationship, with protein-signaling networks playing a critical role as a link between the two. Mutations (indicated with a red star) at the DNA level can affect signaling events at the protein level, hence having the potential to alter molecular signaling networks.

to which end algorithms such as NetPhorest and NetworKIN have been developed (Linding *et al.*, 2007, 2008; Miller *et al.*, 2008; Horn *et al.*, 2014). These algorithms allow, based on both experimentally determined linear motif preferences of kinases and their network contextual information, to predict which kinases are likely candidates for experimentally observed phosphorylation sites. The linear motifs (specific sequence preference around the phosphosite) are predicted by NetPhorest, and STRING (a protein–protein association database) adds the network contextual information on whether or not the kinase and substrate protein (containing the phosphorylation site) are known/predicted to interact *in vivo*. Combined, they are amongst the most comprehensive and accurate algorithms currently available compared to previously published frameworks (Blom *et al.*, 2004; Obenauer *et al.*, 2003; Wong *et al.*, 2007; Xue *et al.*, 2005). These algorithms have high utility for studying posttranslational signaling when combined with cellular perturbations and global quantitative phosphoproteomics data, to pinpoint key kinases which may be fundamental to a given cellular phenotype.

Tools to Study Networks at Multiple Levels of Biological Complexity

In order to facilitate comprehensive sampling of the genotype-to-phenotype relationship, it is imperative that we can measure as many of the underlying properties (e.g., genome, epigenome, metabolome, proteome, and cellular phenotypes) as possible, in a global, unbiased manner. Subsequently, these different biological aspects can be integrated in order to establish causal relations that may exist between them. From a practical point of view, this generally means that large experimental and computational infrastructures are required, to allow the generation and analysis of the required large-scale datasets in a timely fashion. We will now briefly explore some of these technologies in the light of trying to decipher cellular signaling in cancer using a Network Biology approach.

In order to systematically assess the genomic landscapes within a tumor, NGS of DNA or RNA is the method of choice, as it allows the determination of sequence variants either within the complete genome, or only the protein-coding exome. This type of analysis results in a complete view of the genomic landscape that exists within a particular cellular context, and allows one to begin inferring which genes might play a role in a particular disease phenotype without any prior knowledge (Stratton *et al.*, 2009; Vogelstein *et al.*, 2013). With the significant decrease in costs associated with having a genome sequenced over the last few years, more and more laboratories obtain the capability of conducting such experiments, which has led to an explosion in publications analyzing diseased genomes (Network, 2011; Forbes *et al.*, 2010; Pleasance *et al.*, 2010; Rapley *et al.*, 2009; Stephens *et al.*, 2009). Nevertheless, a direct therapeutic impact that was expected to arise from these analyses has been rather limited (Yaffe, 2013). One fundamental shortcoming of using genomic data alone is the fact that it does not allow interpretation of how these mutations affect, or are utilized by, the cellular signaling networks. In other words, additional experiments are required for deducing which mutations have a functional

impact, and are thereby functionally related to the disease phenotype (Torkamani and Schork, 2009). For example, detecting a mutation at the DNA level does not guarantee that the mutation actually is expressed at the protein level, where it could exert an effect. Additionally, it may not be the mutations themselves which have a phenotypic effect, but rather, in case of kinases being affected by a mutation, the dysregulated phosphorylation-based signaling or network rewiring that is caused by the mutation (Creixell *et al.*, 2012). This underlines the need for actually being able to assess the protein dynamics of both the mutant proteins and their non-mutated network partners, and the dynamic phosphorylation networks which may be altered due to the genomic alterations. Additionally, the functional effect of a given mutation may only be apparent under specific cellular conditions, underlining the need for studying cellular behavior in response to several stimuli (e.g., growth factor stimulation, starvation, hypoxia, or stromal cocultures) or under varying cellular contexts (e.g., 2D/3D *in vitro* cultures and *in vivo* animal models).

Recent developments in the field of MS have led to major improvements of what portion of the expressed proteome can be measured in a single experiment. Currently, it is possible to simultaneously detect and quantify the protein levels of almost the complete proteome and several tens of thousands of phosphorylation events (Beck *et al.*, 2011; Geiger *et al.*, 2012; Munoz and Heck, 2011; Nagaraj *et al.*, 2011; Wisniewski *et al.*, 2009; Bodenmiller *et al.*, 2007; Bodenmiller and Aebersold, 2010; Sharma *et al.*, 2014). Quantitative technologies such as SILAC or dimethyl labeling allow the direct comparison of specific proteins or, for example, phosphorylation sites between one or several proteomes, in order to determine the dynamic regulation thereof (Boersema *et al.*, 2009; Ong and Mann, 2007). Their principle is based on utilizing stable isotopes to introduce a small mass difference into the peptides that exist within a given sample, which allows the mixing of samples. This mass difference is large enough for a high resolution mass spectrometer to determine the sample of origin, and through direct comparison of the peptide intensities originating from the respective samples, enables direct quantitation of these peptides. As samples are mixed in the early stage of sample preparation, any difference in intensity will most likely be due to a biological effect rather than a technical artifact. If used in combination with appropriate experimental design, this type of quantitative analysis enables the in-depth characterization of the dynamic regulation of protein signaling after a given perturbation such as inhibitor treatment, growth factor stimulation, mutations, or altered growth conditions (e.g., hypoxia). This information can subsequently be used to determine the signaling network dynamics that are fundamental to a given biological observation. Additionally, the use of genome-specific, rather than reference genome search databases for determining the proteins and phosphorylation sites present in a given sample, allows the accurate monitoring of the protein dynamics of specific mutations present in sample (Alfaro *et al.*, 2014; Nesvizhskii, 2014). Being able to quantitatively assess the information layer which exists between the genotype and phenotype has a distinct advantage, especially when considering that all small-molecule inhibitors and antibody-based therapeutics used in the clinic today actually target proteins rather than genes.

As was described above, these protein data need to be accurately modeled in order to establish potential signaling network models, but the near unbiased nature of MS experiments allows this to be done in a data-driven way with minimal *a priori* knowledge. Nevertheless, MS data only provides a 'snapshot' of the cellular proteome at the time of the experiment, so experiments need to be designed appropriately to accurately assess the signaling networks fundamental to a particular biological effect.

In order to more directly characterize the role of key proteins in a given cellular phenotype, the use of genome-scale RNA interference (RNAi) libraries in combination with HCS has been demonstrated in several species (Echeverri and Perrimon, 2006; Mohr *et al.*, 2010; Mohr and Perrimon, 2012; Moffat and Sabatini, 2006). By perturbing the expression of a specific protein, and measuring an appropriate phenotypic readout (e.g., proliferation or cell migration), it is possible to globally assess and elucidate proteins which seem fundamental for a given biological state. This approach has been very powerful not only at the discovery phase of investigating potential therapeutic targets, but also at the stage of validating proteins determined by, for example, MS or NGS to be involved in a given disease phenotype (Dorsett and Tuschl, 2004; Wilson *et al.*, 2013). One of the strengths that is fundamental to the success of RNAi-based screening is the fact that they can be used as a highly specific substitute for often unavailable inhibitors. A caveat being that the link between RNAi and inhibitors is not always linear however, and requires careful validation (Weiss *et al.*, 2007). Nevertheless, by systematically knocking down genes in, for example, a kinome-wide or genome-wide fashion, a global assessment of the model system can be undertaken. Additionally, this can be done in a combinatory fashion, where multiple targets can be knocked down simultaneously to monitor synthetically lethal effects where a phenotypic readout is only obtained by the simultaneous knockdown of two or more genes (Azorsa *et al.*, 2009; Kaelin, 2005; Luo *et al.*, 2009; Vizeacoumar *et al.*, 2013). Besides providing a more direct readout of the role of a particular protein in a given disease phenotype, it may also highlight proteins that were not detected by MS or genes that do not harbor a mutation, providing a complementary approach.

Computational Modeling – Understanding the Integrative Nature of Biological Networks

In order to develop a more global understanding of intra- and intercellular dynamics and adaptation it is essential to generate and combine diverse datasets describing the genetic and proteomic changes that occur during cell phenotypic changes. This in turn, requires the community to develop computational tools for real-time high-throughput live/*in situ* cell tracking. In combination with different perturbations (drugs, RNAi, CRISPR, etc.) it should be the aim of such studies to systematically interrogate the integrative nature of molecular networks, while aiming to also validate the ability of the derived computational models to predict biological systems characteristics. Given the nonlinear and integrative nature of biological systems (e.g., cancer cells or metastatic tumors) it is vital to perform computational analysis in a manner that

accepts, and in fact utilizes, this intrinsic nonlinearity of the system. There are a number of major challenges associated with processing and integrating genome-scale data, besides these intrinsic nonlinearities. First, system-wide data will invariably be of high-dimensional nature and possess measurement uncertainties, which need to be carefully accounted for in any systematic computational framework. This also entails a principled way of handling the lack of complete overlap between the different datasets to prevent skewing of the results due to missing values. Furthermore, high-throughput screening (such as HCS) typically generates very large, high-dimensional data cubes, which require a wide dynamic range in computational capacity, in particular in terms of physical memory. Thirdly, different datasets may be of very different nature both in terms of their biological meaning and mathematical representation (e.g., categorical vs. continuous data, single data-vector vs. ensemble of data-vectors).

Many different approaches for complex data modeling have been developed over the years, each with a unique, specific goal in mind (Kim *et al.*, 2011; Kreeger and Lauffenburger, 2010; Morris *et al.*, 2010; Pritchard *et al.*, 2013; Saez-Rodriguez *et al.*, 2009). Which approach is most suitable depends on the type of data and the precise biological question of interest. The Lauffenburger spectrum (Figure 4; Ideker and Lauffenburger, 2003) provides a useful starting point decide on an appropriate computational approach. In one end of this spectrum, one finds a number of high-level models for exploratory type of data analysis. This entails various popular unsupervised techniques, such as principal component analysis, factor analysis and other latent variable models (Akavia *et al.*, 2010; Hofree *et al.*, 2013; Krishnaswamy *et al.*, 2014), all of which provide means to elucidate effective low-dimensional representations of data and/or interrelationships between model components, such as to uncover potential 'drivers' of cancer (Akavia *et al.*, 2010).

In a slightly more supervised setting, one may wish to regress a number of dependent variables (e.g., proliferation, metastasis, or other phenotypic read-outs) on a selected set of independent ones (e.g., selected set of proteins or signaling pathways). Identifying relevant regressors can be done in the framework of partial least square regression for linear models (Janes *et al.*, 2005), or in nonlinear generalizations by neural networks or kernel methodologies (for instance support vector machines or gaussian processes (Bishop, 2006)). Inspired by statistical physics models, prominent methodologies within the field of deep learning have recently given rise to generative models that can handle high-dimensional and highly nonlinear correlated data in a very effective and versatile manner (Bengio, 2009; Kiros *et al.*, 2014). We envisage that these methods will find widespread use in this end of the 'Lauffenburger' spectrum in coming years (e.g., Verbist *et al.*, 2015).

Whilst these aforementioned computational techniques are useful for data mining and regression, they are not particularly geared to elucidate the 'molecular' basis let alone the causal origin for these correlations. Bayesian networks partially address this problem by identifying and modeling the conditional dependencies between observed protein expression levels, PTM patterns, etc., represented as directed edges from the independent ('parent node') to the dependent variable ('child node') (Friedman *et al.*, 2000; Husmeier, 2003; Oates and Mukherjee, 2012; Pe'er and Hachohen, 2011; Sachs *et al.*, 2005;

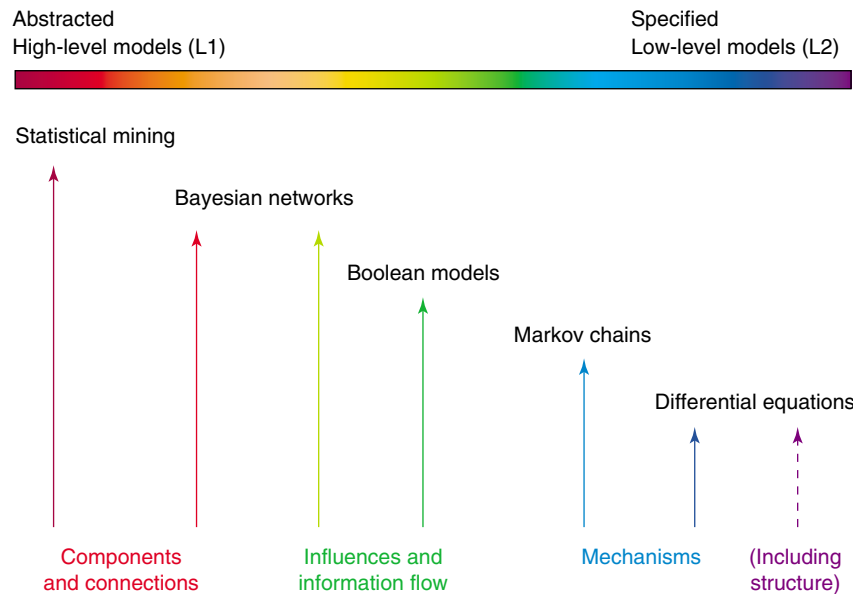


Figure 4 The 'Lauffenburger spectrum.' At the highest level of abstraction statistical data mining methods aim at correlate dependent with independent variables, find low-dimensional data representations and/or elucidate potential interrelationships between model components. At a somewhat lower level, Bayesian network models expand on these relationship by providing conditional dependencies of 'child' nodes on their 'parents' in a probabilistic sense. In Boolean models these conditional probabilities are replaced by logical rules. Low levels methods aim at replace these simplified rules with some that are amenable to a mechanistic interpretation. These rules may be expressed either in terms of transition probabilities (e.g., Markov chains or dynamic Bayesian networks) or kinetic reaction rates (e.g., ordinary or partial differential equations).

Zhang and Song, 2013). Markov networks provide an alternative approach, where edges are undirected and merely represent statistical couplings or 'interactions' between corresponding nodes (Wei and Li, 2007). In both cases, these edge relations are typically taken to be on a particular parametric form (e.g., logistic, gaussian, or polynomial), although recent statistical advances suggest that nonparametric approaches can be made both viable and very powerful, given sufficient data (Krishnaswamy *et al.*, 2014; Zhang and Song, 2013). Boolean networks, on the other hand, represent a particular simplification, whereby the edge relations are defined as binary functions, corresponding to logical operations (Morris *et al.*, 2010; Saez-Rodriguez *et al.*, 2009).

We emphasize that in all cases these network models are more informative than networks defined by simply forming edges between variables whenever the corresponding pairwise correlations for some data exceed some predefined value (e.g., Krishnaswamy *et al.*, 2014). These approaches do not distinguish between correlations arising from direct couplings and those that are mediated by other nodes. As such they provide only limited predictive power and are mostly useful for exploratory analysis only. In contrast, the edges in Bayesian, Markovian, or Boolean networks typically have the interpretation that neighboring nodes influence each other (directly or indirectly through hidden variables) via actual biochemical or biophysical mechanisms. In this sense, these network models are ultimately linked to models in the low-level end the 'Lauffenburger spectrum,' in terms of aiming to identify and encode 'causal' relationships.

The notion of a causal relationship is fundamentally distinct from a correlational one, as expounded and formalized in several fields (Imbens and Angrist, 1994; Pearl, 2009;

Rosenbaum and Rubin, 1983). However, causal network inference in a biological context has so far primarily been used to model regulatory networks and is generally considered an extremely challenging task (Freedman and Humphreys, 1999; Pearl, 2009). Some statistical and machine learning approaches (Friedman *et al.*, 2000; Husmeier, 2003; Sachs *et al.*, 2005) can in principle learn causal structure, but their success can only be guaranteed under very strong assumptions (Pearl, 2009; Spirtes *et al.*, 2000) that are almost certainly violated in biological settings. This is due to a range of factors including, the infeasibility of carrying out all possible interventions on the system and fundamental limitations of the data (Oates and Mukherjee, 2012).

Causal mechanistic models define the other low-level end of the 'Lauffenburger spectrum,' where time plays an explicit role. Here, the network structure is most often imposed from prior knowledge, although methodologies akin to dynamical Bayesian network may allow this to be inferred from data (Husmeier, 2003; Molinelli *et al.*, 2013). The aim of mechanistic models is to infer dynamic responses for various specific perturbations and/or estimate the underlying kinetic parameters with a clear mechanistic interpretation. These methods entail both stochastic (e.g., Markov chains) and deterministic approaches (e.g., ordinary or partial differential equations (Hoffmann *et al.*, 2002; Wilkinson, 2009)).

A recurrent theme across the whole Lauffenburger spectrum is to make a critical assessment and validation of the derived models. To compare and select models on a statistical ground involve assessment of the biases, missing values and uncertainties of the data, as well as the structure and complexity of the model and proper use of prior knowledge. We firmly

believe that these generic considerations in general should be addressed within the principled framework of Bayesian methodologies, irrespective of the particular nature of the data and corresponding models. Due to the complexity involved in cell signaling, and in pathogenesis particularly, comprehensively characterization of molecular processes requires the tight integration of experimental and computational approaches, which invariably implies the need to base the inferred biological relations on a rigorous statistical framework. Genome-scale experiments (e.g., MS, NGS, and RNAi screening) focusing on key biological questions need to be conducted, in order to subsequently integrate the knowledge gained from these different platforms using appropriate computational approaches similar to those previously mentioned. In this manner, several benefits can be obtained: (1) missing information in one dataset can be complemented by that from another, (2) the different data types can act as validation sets for one another, and (3) a comprehensive overview of how genomic information is propagated through the proteome to elicit a specific phenotype can be obtained. For example, a hit originating from an RNAi screen is more likely to be a successful clinical target if MS evidence also suggests the protein level (or phosphorylation dynamics) to be increased or if NGS has revealed there to be a mutation hitting the gene. In the case of point (1), if, for example, a given protein has not been observed by MS to be dysregulated in the diseased state, but the RNAi screen does confidently highlight the knockdown to result in a phenotype, it is likely to be an interesting candidate. Therefore, it is not only the genes and proteins which have been determined by all of the experimental methods to be of therapeutic interest, it is also the complementary nature of the technological platforms which allows for a more comprehensive interrogation of the biological model system.

In conclusion, we advocate the in-depth characterization of biological samples from several technological angles and molecular/biological perspectives, in order to accurately establish genotype-to-phenotype relationships. Ideally, where possible, this should be conducted across several time-points and under different biological conditions, to interrogate the biological system as comprehensively as possible, as certain phenotypes or protein dynamics may only be revealed under highly specific biological conditions. Importantly, by taking a global, unbiased approach to interrogate a biological system, one allows the data to highlight which proteins and/or mutations are potentially fundamental to a specific disease phenotype. This requires the implementation of careful experimental design (tailored to the specific biological question of interest), strict standard operating procedures (SOPs) for stable and reproducible sample preparation conditions and continuous quality control to ensure data is generated consistently over the generally extensive period of time it takes to collect the data. In line with the above-mentioned role of tumor subpopulations and the microenvironment, ideally this would be done in a cell-type-specific manner (e.g., tumor cells, CAFs, cancer stem cells, and other subpopulations), as this will partially deconvolute the complex interplay of the different cell types, and potentially allow for time-staggered, cell-type-specific combination therapies to be developed, thereby positively contributing to the 'war on cancer.'

Concluding Remarks

In this article, we have shown some early results of targeting cancer biology using an integrative network biology approach. By analyzing different cellular aspects using various technological platforms (e.g., MS, NGS, and HCS), a more in-depth view of cellular signaling can be obtained. This knowledge can subsequently be used for driving functional validation studies. We have highlighted some conceptual and technological advancements which facilitate this process, ranging from how mutations should be interpreted functionally, through generating and modeling phospho-proteomic data, to integrating NGS and MS data to study the propagation of mutations at the protein level. We predict that over the next years, MS and NGS technologies will become more prevalent in the clinic, as they have great potential in guiding therapeutic strategies in patients based on their tumors' protein-signaling network states and genotypic profiles. Before this becomes feasible however, the large costs associated with these types of global screens will need to be significantly reduced. Additionally, while NGS allows the interrogation of the complete genome, MS still needs to be improved in term of sensitivity. Nevertheless, advancements in this field are consistently on-going, resulting in greater coverage of cellular (phospho-) proteomes, and we are likely to see this pattern continue (Zhou *et al.*, 2013; Sharma *et al.*, 2014; Branca *et al.*, 2014). With greater coverage and improved sensitivity of the technologies described, creating more accurate molecular maps of cellular processes will become a reality and feed into novel target detection efforts.

Development of a Global, Predictive Biological Model of Cell Behavior

A key target for the (Cancer) network biologists is to integrate global quantitative datasets in order to model core signaling networks that drive the phenotypic changes occurring during the many responses of cancer cells. In our own lab we employ nonlinear Bayesian and deep learning approaches for this task, such that the information flow can be represented in terms of a probabilistic model of the network state, as a function of time. In turn, this can facilitate the generation of quantitative predictions of the time-dependent changes in cell phenotypes, based on the state of regulatory nodes in the signaling networks. We hypothesize that this aim is of strategic importance for future clinical applications as well as in basic biology, where a major bottleneck is our inability to predict cell behavior. Finally, we argue that such data-driven, nonlinear forecasting models as described above will eventually identify multi-node perturbations correlated to phenotypic transitions and events during cancer progression and tumor evolution. It should therefore be a concerted goal of the entire network biology community to identify regulatory networks controlling these processes and to predict key nodes for novel targeted therapies. Efforts to this end will hold great potential for identifying network-based therapeutic strategies and to target multiple key nodes, preventing cancer cell migration and allowing biologists to alter invasive phenotypes. This type of approach holds great promise for lowering the global cancer burden and that of many, if not all, other complex diseases and drug resistance cases.

See also: Horizontal Integration: Analyses and Tools: Network Modeling of Heterogeneous Datasets

References

- Akavia, U.D., *et al.*, 2010. An integrated approach to uncover drivers of cancer. *Cell* 143, 1005–1017.
- Alfaro, J.A., *et al.*, 2014. Onco-proteogenomics: Cancer proteomics joins forces with genomics. *Nature Methods* 11, 1107–1113.
- Azorsa, D.O., *et al.*, 2009. Synthetic lethal RNAi screening identifies sensitizing targets for gemcitabine therapy in pancreatic cancer. *Journal of Translational Medicine* 7, 43.
- Bakal, C., *et al.*, 2007. Quantitative morphological signatures define local signaling networks regulating cell morphology. *Science* 316, 1753–1756.
- Baker, A.M., *et al.*, 2013a. Lysyl oxidase enzymatic function increases stiffness to drive colorectal cancer progression through FAK. *Oncogene* 32, 1863–1868.
- Barker, H.E., *et al.*, 2013b. Tumor-secreted LOXL2 activates fibroblasts through FAK signaling. *Molecular Cancer Research* 11, 1425–1436.
- Beadle, G.W., Tatum, E.L., 1941. Genetic control of biochemical reactions in neurospora. *Proceedings of the National Academy of Sciences of the United States of America* 27, 499–506.
- Beck, M., *et al.*, 2011. The quantitative proteome of a human cell line. *Molecular Systems Biology* 7, 549.
- Bengio, Y., 2009. Learning deep architectures for AI. *Foundations and Trends in Machine Learning* 2, 1–127.
- Bergers, G., *et al.*, 2000. Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis. *Nature Cell Biology* 2, 737–744.
- Bishop, C.M., 2006. *Pattern recognition and machine learning*. New York, NY: Springer.
- Bissell, M.J., Labarge, M.A., 2005. Context, tissue plasticity, and cancer: Are tumor stem cells also regulated by the microenvironment? *Cancer Cell* 7, 17–23.
- Blom, N., *et al.*, 2004. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. *Proteomics* 4, 1633–1649.
- Bodenmiller, B., *et al.*, 2007. Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nature Methods* 4, 231–237.
- Bodenmiller, B., Aebersold, R., 2010. Quantitative analysis of protein phosphorylation on a system-wide scale by mass spectrometry-based proteomics. *Methods in Enzymology* 470, 317–334.
- Boersema, P.J., *et al.*, 2009. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nature Protocols* 4, 484–494.
- Branca, R.M., *et al.*, 2014. HiRIEF LC-MS enables deep proteome coverage and unbiased proteogenomics. *Nature Methods* 11, 59–62.
- Brogna, J., Hunter, T., 2011. Protein kinase signaling networks in cancer. *Current Opinion in Genetics & Development* 21, 4–11.
- Burrell, R.A., *et al.*, 2013. The causes and consequences of genetic heterogeneity in cancer evolution. *Nature* 501, 338–345.
- Chambers, A.F., Groom, A.C., MacDonald, I.C., 2002. Dissemination and growth of cancer cells in metastatic sites. *Nature Reviews Cancer* 2, 563–572.
- Chandrapatya, S., *et al.*, 2011. AKT inhibition relieves feedback suppression of receptor tyrosine kinase expression and activity. *Cancer Cell* 19, 58–71.
- Cox, J., Mann, M., 2007. Is proteomics the new genomics? *Cell* 130, 395–398.
- Cox, T.R., *et al.*, 2013. LOX-mediated collagen crosslinking is responsible for fibrosis-enhanced metastasis. *Cancer Research* 73, 1721–1732.
- Creixell, P., *et al.*, 2012. Navigating cancer network attractors for tumor-specific therapy. *Nature Biotechnology* 30, 842–848.
- Criscuolo, M.L., Nguyen, M., Eliceiri, B.P., 2005. Tumor metastasis but not tumor growth is dependent on Src-mediated vascular permeability. *Blood* 105, 1508–1514.
- Davies, H., *et al.*, 2002. Mutations of the BRAF gene in human cancer. *Nature* 417, 949–954.
- Davies, M.A., Samuels, Y., 2010. Analysis of the genome to personalize therapy for melanoma. *Oncogene* 29, 5545–5555.
- Davis, M.I., *et al.*, 2011. Comprehensive analysis of kinase inhibitor selectivity. *Nature Biotechnology* 29, 1046–1051.
- Dawson, M.A., Kouzarides, T., 2012. Cancer epigenetics: From mechanism to therapy. *Cell* 150, 12–27.
- Dorsett, Y., Tuschl, T., 2004. siRNAs: Applications in functional genomics and potential as therapeutics. *Nature Reviews Drug Discovery* 3, 318–329.
- Dotto, G.P., Weinberg, R.A., Ariza, A., 1988. Malignant transformation of mouse primary keratinocytes by Harvey sarcoma virus and its modulation by surrounding normal cells. *Proceedings of the National Academy of Sciences of the United States of America* 85, 6389–6393.
- Echeverri, C.J., Perrimon, N., 2006. High-throughput RNAi screening in cultured cells: A user's guide. *Nature Reviews Genetics* 7, 373–384.
- Eckhouse, S., Lewison, G., Sullivan, R., 2008. Trends in the global funding and activity of cancer research. *Molecular Oncology* 2, 20–32.
- Engelman, J.A., *et al.*, 2007. MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling. *Science* 316, 1039–1043.
- Erler, J.T., *et al.*, 2006. Lysyl oxidase is essential for hypoxia-induced metastasis. *Nature* 440, 1222–1226.
- Erler, J.T., Giaccia, A.J., 2008. *Critical Steps in Cancer Metastasis: Chapter 3 of Abelloff Clinical Oncology*, fourth ed. Elsevier Publications.
- Fearon, E.R., Vogelstein, B., 1990. A genetic model for colorectal tumorigenesis. *Cell* 61, 759–767.
- Fedorov, O., Muller, S., Knapp, S., 2010. The (un)targeted cancer kinome. *Nature Chemical Biology* 6, 166–169.
- Feinberg, A.P., Tycko, B., 2004. The history of cancer epigenetics. *Nature Reviews Cancer* 4, 143–153.
- Fidler, I.J., 1970. Metastasis: Quantitative analysis of distribution and fate of tumor emboli labeled with 125 I-5-iodo-2'-deoxyuridine. *Journal of the National Cancer Institute* 45, 773–782.
- Fidler, I.J., 2003. The pathogenesis of cancer metastasis: The 'seed and soil' hypothesis revisited. *Nature Reviews Cancer* 3, 453–458.
- Flaherty, K.T., *et al.*, 2010. Inhibition of mutated, activated BRAF in metastatic melanoma. *New England Journal of Medicine* 363, 809–819.
- Forbes, S.A., *et al.*, 2010. COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. *Nucleic Acids Research* 38, D652–D657.
- Freedman, D., Humphreys, P., 1999. Are there algorithms that discover causal structure? *Synthese* 121, 29–54.
- Friedman, N., *et al.*, 2000. Using Bayesian networks to analyze expression data. *Journal of Computational Biology* 7, 601–620.
- Fujita, H., Ohuchida, K., Mizumoto, K., *et al.*, 2010. alpha-Smooth muscle actin expressing stroma promotes an aggressive tumor biology in pancreatic ductal adenocarcinoma. *Pancreas* 39 (8), 1254–1262.
- Geiger, T., *et al.*, 2012. Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Molecular & Cellular Proteomics* 11, M111.014050.
- Gorgoulis, V.G., *et al.*, 2005. Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* 434, 907–913.
- Gstaiger, M., Aebersold, R., 2009. Applying mass spectrometry-based proteomics to genetics, genomics and network biology. *Nature Reviews Genetics* 10, 617–627.
- Gstaiger, M., Aebersold, R., 2013. Genotype–phenotype relationships in light of a modular protein interaction landscape. *Molecular BioSystems* 9, 1064–1067.
- Gupta, G.P., Massague, J., 2006. Cancer metastasis: Building a framework. *Cell* 127, 679–695.
- Hanahan, D., 2013. Rethinking the war on cancer. *Lancet* 383 (9916), 558–563.
- Hanahan, D., Weinberg, R.A., 2011. Hallmarks of cancer: The next generation. *Cell* 144, 646–674.
- Hernando, E., *et al.*, 2004. Rb inactivation promotes genomic instability by uncoupling cell cycle progression from mitotic control. *Nature* 430, 797–802.
- Hiratsuka, S., *et al.*, 2002. MMP9 induction by vascular endothelial growth factor receptor-1 is involved in lung-specific metastasis. *Cancer Cell* 2, 289–300.
- Hoffmann, A., *et al.*, 2002. The I kappa B-NF-kappa B signaling module: Temporal control and selective gene activation. *Science* 298, 1241–1245.
- Hofree, M., *et al.*, 2013. Network-based stratification of tumor mutations. *Nature Methods* 10, 1108–1115.
- Horn, H., *et al.*, 2014. KinomeXplorer: An integrated platform for kinome biology studies. *Nature Methods* 11, 603–604.
- Huang, P.H., *et al.*, 2007a. Uncovering therapeutic targets for glioblastoma: A systems biology approach. *Cell Cycle* 6, 2750–2754.
- Huang, P.H., *et al.*, 2007b. Quantitative analysis of EGFRVIII cellular signaling networks reveals a combinatorial therapeutic strategy for glioblastoma. *Proceedings of the National Academy of Sciences of the United States of America* 104, 12867–12872.
- Husmeier, D., 2003. Sensitivity and specificity of inferring genetic regulatory interactions from microarray experiments with dynamic Bayesian networks. *Bioinformatics* 19, 2271–2282.
- Ideker, T., Lauffenburger, D., 2003. Building with a scaffold: Emerging strategies for high- to low-level cellular modeling. *Trends in Biotechnology* 21, 255–262.

- Imbens, G., Angrist, J., 1994. Identification and estimation of local average treatment effects. *Econometrica* 62, 467–475.
- Janes, K.A., *et al.*, 2005. A systems model of signaling identifies a molecular basis set for cytokine-induced apoptosis. *Science* 310, 1646–1653.
- Janne, P.A., Gray, N., Settleman, J., 2009. Factors underlying sensitivity of cancers to small-molecule kinase inhibitors. *Nature Reviews Drug Discovery* 8, 709–723.
- Jorgensen, C., *et al.*, 2009. Cell-specific information processing in segregating populations of Eph receptor ephrin-expressing cells. *Science* 326, 1502–1509.
- Jorgensen, C., Linding, R., 2008. Directional and quantitative phosphorylation networks. *Briefings in Functional Genomics & Proteomics* 7, 17–26.
- Kaelin, W.G.J., 2005. The concept of synthetic lethality in the context of anticancer therapy. *Nature Reviews Cancer* 5, 689–698.
- Karaman, M.W., *et al.*, 2008. A quantitative analysis of kinase inhibitor selectivity. *Nature Biotechnology* 26, 127–132.
- Kim, H.D., *et al.*, 2011. Signaling network state predicts twist-mediated effects on breast cell migration across diverse growth factor contexts. *Molecular & Cellular Proteomics* 10, M111.008433.
- Kim, J.W., *et al.*, 2004. Rapid apoptosis in the pulmonary vasculature distinguishes non-metastatic from metastatic melanoma cells. *Cancer Letters* 213, 203–212.
- Kiros, R., Salakhutdinov, R., Zemel, R.S., 2014. Unifying visual-semantic embeddings with multimodal neural language models. *arXiv*. 1411.2539.
- Kreeger, P.K., Lauffenburger, D.A., 2010. Cancer systems biology: A network modeling perspective. *Carcinogenesis* 31, 2–8.
- Krishnaswamy, S., *et al.*, 2014. Systems biology. Conditional density-based analysis of T cell signaling in single-cell data. *Science* 346, 1250689.
- Lee, M.J., *et al.*, 2012. Sequential application of anticancer drugs enhances cell death by rewiring apoptotic signaling networks. *Cell* 149, 780–794.
- Lemmon, M.A., Schlessinger, J., 2010. Cell signaling by receptor tyrosine kinases. *Cell* 141, 1117–1134.
- Lin, J., *et al.*, 2007. A multidimensional analysis of genes mutated in breast and colorectal cancers. *Genome Research* 17, 1304–1318.
- Linding, R., *et al.*, 2007. Systematic discovery of *in vivo* phosphorylation networks. *Cell* 129, 1415–1426.
- Linding, R., *et al.*, 2008. NetworkKIN: A resource for exploring cellular phosphorylation networks. *Nucleic Acids Research* 36, D695–D699.
- Lito, P., *et al.*, 2012. Relief of profound feedback inhibition of mitogenic signaling by RAF inhibitors attenuates their activity in BRAFV600E melanomas. *Cancer Cell* 22, 668–682.
- Luga, V., *et al.*, 2012. Exosomes mediate stromal mobilization of autocrine Wnt-PCP signaling in breast cancer cell migration. *Cell* 151, 1542–1556.
- Luo, J., *et al.*, 2009. A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. *Cell* 137, 835–848.
- Luzzi, K.J., *et al.*, 1998. Multistep nature of metastatic inefficiency: Dormancy of solitary cells after successful extravasation and limited survival of early micrometastases. *American Journal of Pathology* 153, 865–873.
- Macurek, L., *et al.*, 2008. Polo-like kinase-1 is activated by aurora A to promote checkpoint recovery. *Nature* 455, 119–123.
- Maser, R.S., DePinho, R.A., 2002. Connecting chromosomes, crisis, and cancer. *Science* 297, 565–569.
- Meads, M.B., Gatenby, R.A., Dalton, W.S., 2009. Environment-mediated drug resistance: A major contributor to minimal residual disease. *Nature Reviews Cancer* 9, 665–674.
- Miller, M.L., *et al.*, 2008. Linear motif atlas for phosphorylation-dependent signaling. *Science Signaling* 1, ra2.
- Miller-Jensen, K., *et al.*, 2007. Common effector processing mediates cell-specific responses to stimuli. *Nature* 448, 604–608.
- Moffat, J., Sabatini, D.M., 2006. Building mammalian signalling pathways with RNAi screens. *Nature Reviews Molecular Cell Biology* 7, 177–187.
- Mohr, S., Bakal, C., Perrimon, N., 2010. Genomic screening with RNAi: Results and challenges. *Annual Review of Biochemistry* 79, 37–64.
- Mohr, S.E., Perrimon, N., 2012. RNAi screening: New approaches, understandings, and organisms. *Wiley Interdisciplinary Reviews: RNA* 3, 145–158.
- Molinelli, E.J., *et al.*, 2013. Perturbation biology: Inferring signaling networks in cellular systems. *PLoS Computational Biology* 9, e1003290.
- Morris, M.K., *et al.*, 2010. Logic-based models for the analysis of cell signaling networks. *Biochemistry* 49, 3216–3224.
- Munoz, J., Heck, A.J., 2011. Quantitative proteome and phosphoproteome analysis of human pluripotent stem cells. *Methods in Molecular Biology* 767, 297–312.
- Nagaraj, N., *et al.*, 2011. Deep proteome and transcriptome mapping of a human cancer cell line. *Molecular Systems Biology* 7, 548.
- Nesvizhskii, A.I., 2014. Proteogenomics: Concepts, applications and computational strategies. *Nature Methods* 11, 1114–1125.
- Network, T.C.G.A.R., 2011. Integrated genomic analyses of ovarian carcinoma. *Nature* 474, 609–615.
- Nguyen, D.X., Bos, P.D., Massague, J., 2009. Metastasis: From dissemination to organ-specific colonization. *Nature Reviews Cancer* 9, 274–284.
- Nurse, P., 2003. The great ideas of biology. *Clinical Medicine* 3, 560–568.
- Oates, C.J., Mukherjee, S., 2012. Network inference and biological dynamics. *Annals of Applied Statistics* 6, 1209–1235.
- Obenaus, J.C., Cantley, L.C., Yaffe, M.B., 2003. Scansite 2.0: Proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Research* 31, 3635–3641.
- Ong, S.E., Mann, M., 2007. Stable isotope labeling by amino acids in cell culture for quantitative proteomics. *Methods in Molecular Biology* 359, 37–52.
- Orimo, A., *et al.*, 2005. Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell* 121, 335–348.
- Page, S., 1889. The distribution of secondary growths in cancer of the breast. *Lancet* 1, 571–573.
- Pawson, T., Hunter, T., 1994. Signal transduction and growth control in normal and cancer cells. *Current Opinion in Genetics & Development* 4, 1–4.
- Pawson, T., Kofler, M., 2009. Kinome signaling through regulated protein-protein interactions in normal and cancer cells. *Current Opinion in Cell Biology* 21, 147–153.
- Pawson, T., Linding, R., 2008. Network medicine. *FEBS Letters* 582, 1266–1270.
- Pearl, J., 2009. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.
- Pe'er, D., Hachohen, N., 2011. Principles and strategies for developing network models in cancer. *Cell* 144, 864–873.
- Pleasance, E.D., *et al.*, 2010. A small-cell lung cancer genome with complex signatures of tobacco exposure. *Nature* 463, 184–190.
- Polanska, U.M., Orimo, A., 2013. Carcinoma-associated fibroblasts: Non-neoplastic tumour-promoting mesenchymal cells. *Journal of Cellular Physiology* 228, 1651–1657.
- Poulikakos, P.I., Rosen, N., 2011. Mutant BRAF melanomas — Dependence and resistance. *Cancer Cell* 19, 11–15.
- Pritchard, J.R., *et al.*, 2013. Predicting cancer drug mechanisms of action using molecular network signatures. *Molecular BioSystems* 9, 1604–1619.
- Puc, J., *et al.*, 2005. Lack of PTEN sequesters CHK1 and initiates genetic instability. *Cancer Cell* 7, 193–204.
- Rapley, E.A., *et al.*, 2009. A genome-wide association study of testicular germ cell tumor. *Nature Genetics* 41, 807–810.
- Regan Anderson, T.M., *et al.*, 2013. Breast tumor kinase (Brk/PTK6) is a mediator of hypoxia-associated breast cancer progression. *Cancer Research* 73, 5810–5820.
- Rosenbaum, P., Rubin, D., 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70, 41–55.
- Sachs, K., *et al.*, 2005. Causal protein-signaling networks derived from multiparameter single-cell data. *Science* 308, 523–529.
- Saez-Rodriguez, J., *et al.*, 2009. Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction. *Molecular Systems Biology* 5, 331.
- Schoeberl, B., *et al.*, 2009. Therapeutically targeting ErbB3: A key node in ligand-induced activation of the ErbB receptor-PI3K axis. *Science Signaling* 2, ra31.
- Sharma, K., *et al.*, 2014. Ultradeep human phosphoproteome reveals a distinct regulatory nature of Tyr and Ser/Thr-based signaling. *Cell Reports* 8, 1583–1594.
- Society, A.C., 2012. The history of cancer. Available at: <http://www.cancer.org/cancer/cancerbasics/thehistoryofcancer/the-history-of-cancer-what-is-cancer> (accessed 17.06.15).
- Spirtes, P., Glymour, C.N., Scheines, R., 2000. *Causation, Prediction, and Search*. Cambridge, MA; London: MIT Press.
- Stephens, P.J., *et al.*, 2009. Complex landscapes of somatic rearrangement in human breast cancer genomes. *Nature* 462, 1005–1010.
- Stephens, P.J., *et al.*, 2012. The landscape of cancer genes and mutational processes in breast cancer. *Nature* 486, 400–404.
- Stratton, M.R., Campbell, P.J., Futreal, P.A., 2009. The cancer genome. *Nature* 458, 719–724.
- Straussman, R., *et al.*, 2012. Tumour micro-environment elicits innate resistance to RAF inhibitors through HGF secretion. *Nature* 487, 500–504.
- Tan, C.S., *et al.*, 2009. Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. *Science Signaling* 2, ra39.
- Taylor, I.W., *et al.*, 2009. Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nature Biotechnology* 27, 199–204.
- Taylor, I.W., Wrana, J.L., 2012. Protein interaction networks in medicine and disease. *Proteomics* 12, 1706–1716.

- Torkamani, A., Schork, N.J., 2009. Identification of rare cancer driver mutations by network reconstruction. *Genome Research* 19, 1570–1578.
- Verbist, B., *et al.*, 2015. Using transcriptomics to guide lead optimization in drug discovery projects: Lessons learned from the QSTAR project. *Drug Discovery Today* 20 (5), 505–513.
- Vidal, M., 2009. A unifying view of 21st century systems biology. *FEBS Letters* 583, 3891–3894.
- Vidal, M., Cusick, M.E., Barabasi, A.L., 2011. Interactome networks and human disease. *Cell* 144, 986–998.
- Vihinen, P., Kahari, V.M., 2002. Matrix metalloproteinases in cancer: Prognostic markers and therapeutic targets. *International Journal of Cancer* 99, 157–166.
- Vizeacoumar, F.J., *et al.*, 2013. A negative genetic interaction map in isogenic cancer cell lines reveals cancer cell vulnerabilities. *Molecular Systems Biology* 9, 696.
- Vogelstein, B., *et al.*, 2013. Cancer genome landscapes. *Science* 339, 1546–1558.
- Vogelstein, B., Kinzler, K.W., 2004. Cancer genes and the pathways they control. *Nature Medicine* 10, 789–799.
- Wagle, N., *et al.*, 2011. Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor genomic profiling. *Journal of Clinical Oncology* 29, 3085–3096.
- Wei, Z., Li, H., 2007. A Markov random field model for network-based analysis of genomic data. *Bioinformatics* 23, 1537–1544.
- Weiss, W.A., Taylor, S.S., Shokat, K.M., 2007. Recognizing and exploiting differences between RNAi and small-molecule inhibitors. *Nature Chemical Biology* 3, 739–744.
- Wilkinson, D.J., 2009. Stochastic modelling for quantitative description of heterogeneous biological systems. *Nature Reviews Genetics* 10, 122–133.
- Wilson, J.L., *et al.*, 2013. Integrated network analyses for functional genomic studies in cancer. *Seminars in Cancer Biology* 23, 213–218.
- Wilson, T.R., *et al.*, 2012. Widespread potential for growth-factor-driven resistance to anticancer kinase inhibitors. *Nature* 487, 505–509.
- Wisniewski, J.R., *et al.*, 2009. Universal sample preparation method for proteome analysis. *Nature Methods* 6, 359–362.
- Wong, Y.H., *et al.*, 2007. KinasePhos 2.0: A web server for identifying protein kinase-specific phosphorylation sites based on sequences and coupling patterns. *Nucleic Acids Research* 35, W588–W594.
- Wood, L.D., *et al.*, 2007. The genomic landscapes of human breast and colorectal cancers. *Science* 318, 1108–1113.
- Xue, Y., *et al.*, 2005. GPS: A comprehensive www server for phosphorylation sites prediction. *Nucleic Acids Research* 33, W184–W187.
- Yaffe, M.B., 2013. The scientific drunk and the lamppost: Massive sequencing efforts in cancer discovery and treatment. *Science Signaling* 6, pe13.
- Yaffe, M.B., Cantley, L.C., 1999. Signal transduction. Grabbing phosphoproteins. *Nature* 402, 30–31.
- Yamashita, M., *et al.*, 2012. Role of stromal myofibroblasts in invasive breast cancer: Stromal expression of alpha-smooth muscle actin correlates with worse clinical outcome. *Breast Cancer* 19, 170–176.
- Zhang, Y., Song, M., 2013. Deciphering interactions in causal networks without parametric assumptions. *arXiv*. 1311.2707.
- Zhou, H., *et al.*, 2013. Robust phosphoproteome enrichment using monodisperse microsphere-based immobilized titanium (IV) ion affinity chromatography. *Nature Protocols* 8, 461–480.

Relevant Website

Cancer.org
Economic Impact of Cancer.

HORIZONTAL INTEGRATION: OMICS

Contents

Connecting Evolutionary Genomics to Cell Biology

Transcriptomics

The Advent of Mass Spectrometry-Based Proteomics in Systems Biology Research

Phosphoproteomics: Approaches, Developments, and Challenges

Interactomes – Scaffolds of Cellular Systems

Metabolomics in Cell Biology

Connecting Evolutionary Genomics to Cell Biology

IM Ventura, The University of Chicago, Chicago, IL, USA and Ministry of Education of Brazil, Brasília, Brazil
M Long, The University of Chicago, Chicago, IL, USA

© 2016 Elsevier Inc. All rights reserved.

Genomics for Old and New Questions

Among the long-standing aims of biology, understanding how organisms function, develop, and adapt occupies a central position in the field. In modern biology, researchers pursue some of the same standard questions, but with a remarkable advancement: they are now able to connect intricate biological processes to the most fundamental source of biological information – the genome. As it has been known for decades, the genome of a species contains the instructions required for its evolutionary perpetuation, including the development, functioning and reproduction of the organism carrying it. The 20th century has witnessed a step-wise advancement in our comprehension of the genome structure and function; from the role of DNA in hereditary, passing through its biochemical structure, genetic coding logic, and regulatory mechanisms. The launch of the Human Genome Project in the 1990s gave the start to several genome consortiums aiming to

sequence model organisms’ genomes (e.g., yeast, fruit fly, mouse; see [Table 1](#)). In a simplified overview, a genome project includes sequencing the nucleotides of short fragments of DNA, covering the whole organism genome; followed by the computational assembly of these short segments into large sequences representing the chromosomes. Finally, the genome annotation step identifies different functional elements in the assembled sequence (such as protein-coding regions and regulatory motifs). In recent years, there has been an explosion in the number of genomic studies published, encouraged by the advance in the sequencing technologies (the so-called ‘next-generation sequencing’) and bioinformatic analysis tools. The massively parallel sequencing implemented nowadays rapidly generates high amounts of genome-wide data, and allows genomics to shift to pursue biological questions in a unprecedented way, connecting variation at the genome level to its functional consequences ([Koboldt et al., 2013](#)).

Table 1 Some landmark genome projects that have contributed to the understanding of genome structure, function, and evolution

Organism	Importance	References
<i>Haemophilus influenza</i> (bacterium)	First complete prokaryotic genome	Fleischmann et al. (1995)
<i>Saccharomyces cerevisiae</i> (yeast)	First complete eukaryotic genome	Goffeau et al. (1996)
<i>Escherichia coli</i>	Most studied prokaryotic organism	Blattner et al. (1997)
<i>Caenorhabditis elegans</i> (worm)	Model organism, first multicellular organism genome	C. elegans Sequencing Consortium (1998)
<i>Drosophila melanogaster</i> (fruit fly)	Important model organism	Adams et al. (2000)
<i>Arabidopsis thaliana</i>	First plant genome, important model organism	The Arabidopsis Genome Initiative (2000)
<i>Homo sapiens</i>	A complex mammal	International Human Genome Sequencing Consortium (2001) , Venter et al. (2001)
<i>Mus musculus</i> (mouse)	Mammal model organism	Mouse Genome Sequencing Consortium (2002)
12 <i>Drosophila</i> genomes	Comparative analysis of fly species	Drosophila 12 Genomes Consortium (2007)
1000 human genomes	Characterization of human genetic variation across populations	Durbin et al. (2010)

The simple description given above may create the deceptive impression that the link between nucleotide sequences in the genome and their function is direct and unambiguous. However, the genetic information encoded in the DNA may not be evident when a single genome is read. To put in another way, when we try to understand what portions of the genomes are functionally relevant and how the intricate biological processes arise from them, to use a well-known metaphor, the 'book of life' resembles more an encrypted file.

In order to identify functionally relevant elements, comparison of genome sequences from different species provides a powerful strategy. The justification behind such comparative framework is that portions of the genome that encode for biologically important functions tend to be more conserved across different species than nonfunctional regions (Hardison, 2003). The rationale is simple: genomes are continually evolving; therefore, if sequences from different species remain with high similarity, despite their independent evolutionary trajectories since the last common ancestor, it is reasonable to assume that they were spared by evolution because they play a functional role, which is usually subject to purifying selection, whereas nonfunctional portions are free to accumulate mutations, or even be deleted from the genome. This approach can be used to identify genes, and it is particularly useful for finding regulatory regions, whose structure is more heterogeneous than that of protein-coding sequences. Therefore, when researchers find conserved sequences from evolutionary distant species (such as human, mouse, and fly), they conclude this sequence is a good candidate for a functional element, because it retained similar information even in species that have been diverging for millions of years. It must be noted, on the other hand, that this conservative approach may fail to identify genes that underwent intense divergence in a short evolutionary period, such as that driven by rapid positive selection, for example. Thus, complementary annotation approaches, such as measurements of protein-coding potential, will probably improve genome analyses (Zhang *et al.*, 2012).

In this context, it is not surprising that the human genome sequencing was followed by similar projects in other model organisms (Table 1 shows that several pioneer genome projects were published in a short period of time). Even if one is interested in exploring genome function of a single species (e.g., identifying genes contributing to human diseases), evolutionary comparisons with genomes from diverse taxa are imperative (Collins *et al.*, 2003). Nevertheless, identifying functional elements in the genome is just the first step for investigating how genes work, interact and evolve. Detailed functional experiments are required to confirm and determine the specific biological functions of genetic elements (Miklos and Rubin, 1996).

Besides the role of identifying functional elements in genomic studies, comparative genomics is also a powerful strategy to investigate evolutionary questions (Hardison, 2003). Comparison of genomes in a phylogenetic framework, ranging from closely related species to taxa in different kingdoms, for example, may help to elucidate exciting questions: What is the basic set of genes common to diverse taxa? How often new genes are created? What are the evolutionary processes involved in diversification and adaptation? An illustrative scheme is provided in Figure 1. In the other extreme,

comparison of genomes from individuals of the same species can also be applied to pinpoint specific genetic variations underlying relevant phenotypes (such as adaptive traits or even diseases factors) in genome-wide association studies.

A complementary set of questions we are now able to investigate relates to the origin of genetic novelties. In the past decades, many studies have explored the creation of new genetic elements, how they integrated in genetic systems, and their roles in phenotypic evolution (reviewed in Long *et al.*, 2013).

Case Studies: New Genes for Old and New Functions

In the rest of this article, we will focus on studies that took advantage of genomic data to identify new genes and then investigate fundamental questions about their evolution, especially those related to cell biology functions. We refer to new genes as distinct loci that originated at a certain time in the evolutionary history of a species and that were not present in the ancestors (Chen *et al.*, 2013). New genes offer a particularly convenient study model to address evolutionary questions because, given their relatively recent origin, they still carry the evolutionary signatures that shaped their formation and molecular evolution (Long *et al.*, 2013). Furthermore, due to the great advance of computational and molecular biology techniques, we can now explore some evolutionary problems in a precise and mechanistic way, identifying patterns, elaborating hypotheses, and testing their predictions, in order to achieve a comprehensive view of new genes evolution, connecting genetic and functional divergence.

New genes can originate from diverse mechanisms, such as DNA duplication, retrotransposition (RNA-based duplication), gene fission or fusion, and *de novo* creation (Chen *et al.*, 2013). These new loci can be identified through syntenic alignments of genomes from different species. Genes present in some (or even a single) species but absent in all outgroup species are good candidates for having a recent origination (see Figure 1). Furthermore, structural features in the gene and flanking regions can be used to infer the mechanisms of its formation. After identifying a new gene candidate, studies typically seek to confirm its functionality (analyzing its open reading frame length and transcription products, for instance) and explore signatures of molecular evolution (such as synonymous/non-synonymous substitutions rates). Ideally, the ultimate goal is to connect the evolution at the nucleotide sequence to its phenotypic impact, evaluating the functional and adaptive impact of the new gene origination.

Basic cellular functions have been considered to be encoded by a set of old, conserved genes, since some essential processes and pathways are found in evolutionary distant organisms (Miklos and Rubin, 1996), whereas new genes would be involved in minor, dispensable activities. Nevertheless, an increasing number of studies have been reporting the contribution of new genes to basic cellular processes. They show that new genes can be integrated into conserved pathways, and even become essential for viability and reproduction (Chen *et al.*, 2010). Next, we will explore how the study of new genes can be connected to cellular biology. We will focus on examples of gene movement between chromosomes and the

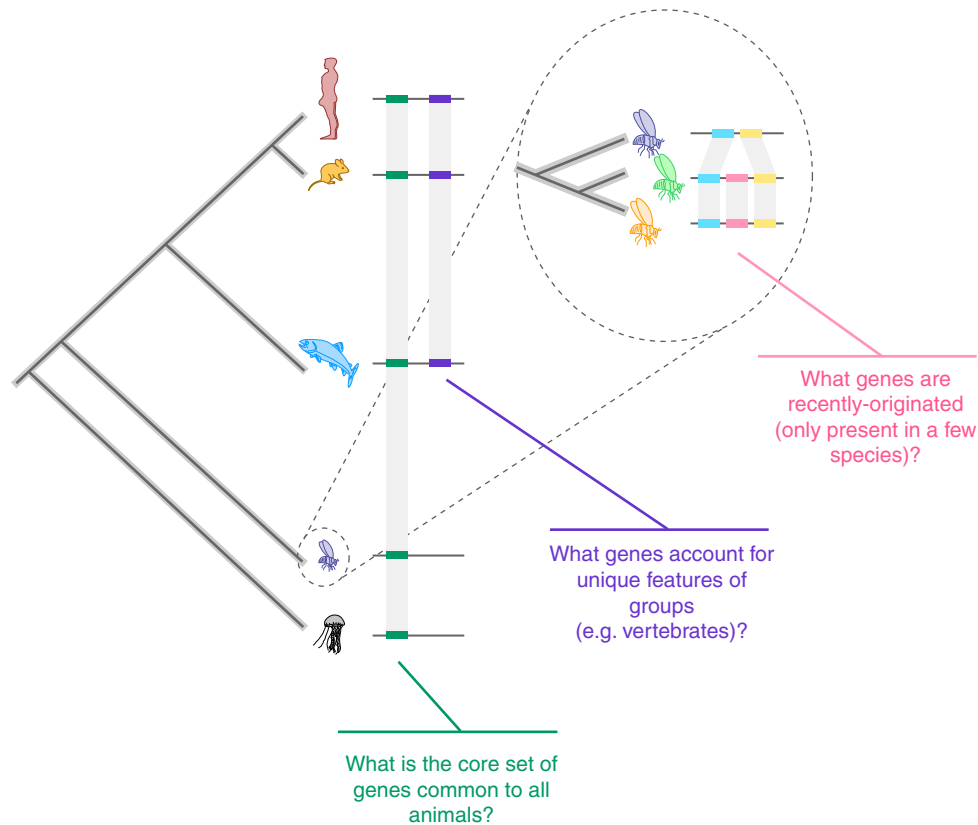


Figure 1 Schematic representation of different questions addressed in a comparative genomics approach (based on Hardison, 2003). At left, colored boxes represent genes present in different animal groups in the phylogeny (cnidarian, fly, fish, mouse, and human); at right, a phylogeny of closely related fly species illustrating a syntenic alignment used to find a new gene (pink box).

integration of new genes into basic cellular functions (namely, gene regulatory networks, cell division, and protein traffic) to argue that the evolution of new genes can shed light on fundamental cellular mechanisms. With that, we hope to stimulate future insightful connections between the genomics and cell biology fields.

Gene Traffic

Cytogenetics played a central role in the early decades of genetic studies. Enough is to mention, for instance, that Thomas Morgan and his students were pioneers in connecting chromosome structure to heredity; and Theodosius Dobzhansky and his students were able to use structural variations (such as chromosomal inversions) as molecular markers for studying variation in natural populations of *Drosophila*, all with elementary microscopy techniques. With the recent advancement of molecular biology techniques, however, evolutionary studies have changed its focus from variation of structure to the specific changes on nucleotide sequences. In this context, genomic analyses allow the investigation of biological questions at a finer scale. However, basic structural data can be surprisingly informative for testing evolutionary hypotheses in a comparative genomics framework. We will illustrate this point using recent findings of new genes

movement patterns in mammals (e.g., humans and rodents) and insects (e.g., flies, mosquitoes, and silkworms).

Several studies demonstrated that the distribution of retrogenes (new loci created through RNA-based duplication) and their parental copies in the genomes are not random with respect to their chromosomal position; instead, new genes movement follows a preferential direction (gene traffic). Analyses focusing on the *Drosophila melanogaster* genome (Betran *et al.*, 2002) and more recent comparative investigations using genomes from 12 *Drosophila* species (Meisel *et al.*, 2009; Vıbranovski *et al.*, 2009b), for example, revealed that retrogenes derived from X-linked genes preferentially move to autosomes ($X \rightarrow A$), contrasting to random expectations. More interestingly, the large majority of these $X \rightarrow A$ retrogenes are expressed in testis, suggesting a connection between male-biased function and chromosomal location. To some extent, a similar pattern is observed in human and mouse genomes (Emerson *et al.*, 2004), in which there is an excess of $X \rightarrow A$ movement of retrogenes, as well as male expression bias. However, in these mammal genomes there is also a significant excess of retrogenes moving in the opposite direction ($A \rightarrow X$), which present either female or unbiased expressions. Interestingly, a somewhat symmetrical pattern is observed in silkworms (in which females are ZW and males are ZZ), whose retrogenes tend to move out of the Z chromosome and have ovary-biased expression (Wang *et al.*, 2012). Additionally, recent comparative analyses of 16

mosquito genomes also found an excess of gene translocations from the X to the autosomes in these species (Neafsey *et al.*, 2015).

How can these findings be explained in an evolutionary perspective? Gene traffic patterns are probably explained by both adaptive and mechanistic forces acting on the origination and insertion of new genes. Several models have been proposed to explain the specific roles of each force, and their relative importance is still debated (summarized in Long *et al.*, 2012). Among the plausible models for explaining why new genes have a preferential movement across the genome, we find the sexual antagonism hypothesis, which predicts that dominant beneficial mutations to females but detrimental to males have higher probability of fixation by natural selection on the X chromosome, while recessive ones have higher probability of fixation on the autosomes. Conversely, dominant and recessive mutations that are beneficial to males but detrimental to females have higher probability of fixation on the autosomes and on the X chromosome, respectively (Rice, 1984; Charlesworth *et al.*, 1987). Other, more mechanistic, explanations are the meiotic sex-chromosome inactivation (Vibrantovski *et al.*, 2009; Vibrantovski, 2014) and the dosage compensation models (Vicoso and Charlesworth, 2009). Both hypotheses suggest that processes occurring in the X chromosome reduce the expression of male-biased genes (such as sex-chromosome inactivation and dosage compensation) and, as a consequence, natural selection favors the relocation of these genes to autosomes. Finally, a recently proposed hypothesis (Díaz-Castillo and Ranz, 2012) suggests that the accumulation of testis-expressed retrogenes in *Drosophila* autosomes can also be explained by the nuclear dynamics of chromosome domains in the male germ line. They argue that the preferential integration of male-expressed retrogenes on autosomes is caused by the increased accessibility of open chromatin domains with testis expression during spermatogenesis, due to distinct chromosome interactions with the nuclear lamin. The lower fraction of accessible domains on the X chromosome in the male germ line would also explain the paucity of retrogenes in this chromosome. This model illustrates how data regarding gene locations can be combined with structural information to provide new explanatory hypotheses.

There is still intense debate regarding the limitations and explanatory power of each model. The general patterns of retrogene movement are probably caused by a combination of factors, each one promoted by a different model. Further understanding will certainly emerge from the collection of more genomic, expression and functional data. In any case, these findings show how basic structural information (in this case, chromosomal distribution of new genes) can be connected to genomic data to reveal compelling patterns. We believe that future investigations connecting gene distribution to other chromosome structures (such as chromatin states and chromosomal domains) may provide exciting evolutionary insights.

Gene Regulatory Networks

A key process for proper cellular function is gene regulation. Spatial and temporal transcriptional activation of genes is

determined by complex and extensive regulatory networks, often involving signaling cascades, protein complexes assembly and transcription factors binding to gene regulatory motifs. Recent large-scale genomic analyses have described general patterns regarding the intricate regulatory networks from some model organisms (Mackay, 2014), which allows us to understand the number of functional elements, connectivity and logic of networks topologies. However, little is known about how regulatory networks actually evolve at a high resolution, how they can be rewired and expanded, and how their constituent elements adapt in concert. A good way to explore that is by investigating how the addition of new elements impact preexisting networks. Next, we will illustrate how the recent integration of gene duplicates provides insights about network evolution.

In one example, the gene *Zeus* (*Drcd-1r*) originated through retrotransposition from a X-linked parental gene, *Caf40* (*Drcd-1*), in the ancestor of *D. melanogaster*, *D. simulans*, *D. mauritiana* and *D. sechellia*, around 4–6 million years ago (Quezada-Díaz *et al.*, 2010). *Caf40* encodes a highly evolutionary conserved DNA-binding protein and takes part in a transcription regulatory complex. Interestingly, *Zeus* protein sequence (with around 300 residues) underwent intense divergence from its parent, accumulating more than 150 amino-acid substitutions in a short evolutionary period, with evidence of positive natural selection, whereas the parental copy accumulated less than ten amino-acid substitutions in the same species (Chen *et al.*, 2012). Also, *Zeus* regulation also diverged, acquiring restricted testis-expression, and being subject to histone modifications (Arthur *et al.*, 2014), while the parental *Caf40* remained ubiquitously expressed. Using chromatin immunoprecipitation followed by microarray assays (ChIP-chip), Chen *et al.* (2012) demonstrated that *Zeus* protein binding sites across the genome have diverged to large extent from the parental gene, promoting the upregulation of male-biased genes and downregulation of female-biased genes. Furthermore, *Zeus* knockout shows a dramatic negative effect on male fertility, evidencing the functional relevance of *Zeus* integration into the regulatory network.

A similar pattern was investigated with another regulator recently duplicated in the *D. melanogaster* species complex. *Nsr* (novel spermatogenesis regulator) originated through DNA duplication from the *kep1* locus in the same phylogenetic branch of *Zeus*, and also exhibits evidence of positive selection. The parental gene encodes a conserved RNA-binding regulator, with ubiquitous expression, whereas *nsr* expression is restricted to testis. Alike the previous example, *nsr* knockout also decreases male fertility. Detailed electron microscope revealed that *nsr* knockout results in loss of the outer dynein arms of the sperm axoneme, generating abnormal sperm structures (Ding *et al.*, 2010). More interesting, the authors suggest that *Nsr* functional divergence may be related to its subcellular localization, which is much broader than that of *Kep1* protein. Therefore, according to their hypothesis, *nsr* acquired new male reproductive function through the regulation of mRNA of preexisting genes required for spermatogenesis. Curiously, the sperm maturation processes affected here are very conserved during evolution.

Both examples mentioned above exemplify how new genes can be integrated into preexisting regulatory networks and

acquire important functions. In this process, the new gene can substantially diverge from the parental locus, changing its interaction partners, expression pattern and subcellular localization. This integration can impact the overall network architecture, and promote its rewiring, which can explain, for example, why some new genes quickly acquired crucial functions in *Drosophila* and even became essential for development and other conserved processes, despite their recent origin (Chen *et al.*, 2010; Kemkemer and Long, 2014).

Cell Division

Another cellular process that, despite being highly fundamental and evolutionary conserved, may also be impacted by new gene recruitment is cell division. The intricate cellular mechanisms driving the accurate distribution of the duplicated genome during cell division are essential to ensure proper cellular viability and development. Chromosome alignment and segregation require the interaction between kinetochore proteins associated with the centromere and the spindle microtubules, usually involving conserved protein complexes (Cheeseman and Desai, 2008).

In a detailed investigation, Ross *et al.* (2013) showed that the new duplicate gene *Umbrea*, originated through DNA duplication 12–15 million years ago in *Drosophila*, evolved an essential role for chromosome dynamics during mitotic division. Comparing sequences from a range of species and manipulating their protein structure, they were able to reconstruct the step-wise trajectory of *Umbrea* evolution. After duplication from the parental *HP1B* (which encodes a heterochromatin protein involved in gene regulation), *Umbrea* divergence involved the loss of an ancestral heterochromatin-localizing domain, and other changes that rewired its protein interaction network and led to centromere localization. Surprisingly, whereas the parental *HP1B* gene is dispensable for viability, *Umbrea* developed an essential function, as demonstrated by the high proportion of mitotic errors found in knockdown cells (Ross *et al.*, 2013; Chen *et al.*, 2010).

The authors propose that recurrent changes at centromeric regions may explain the recruitment of new duplicate genes coding for centromeric proteins. *Umbrea* constitutes a notable example of how new genes can be integrated into preexisting cellular networks and impact its overall architecture, even for fundamental cellular processes, such as chromosome segregation.

Protein Traffic

As a final example, we will discuss a case study that challenges our concepts of conservation and homology of subcellular components. Many complex cellular structures, such as organelles, are found in phylogenetic distant groups of organisms. The parsimonious interpretation is that those shared features are homologous, that is, they are ancient structures inherited from a common ancestor. There at least some cases, however, in which the finding of complex structures found in distant organisms are better explained by independent convergent evolution, rather than by shared inheritance.

The hypothesis of independent origins was explored with specialized secretory vesicles called dense core granules (DCGs). These vesicles accumulate in the cytoplasm and function as storage reservoirs, until extracellular signals cause their fusion with the plasma membrane and the extracellular release of their content (reviewed in Elde *et al.*, 2007). Because DCGs are found in a wide range of eukaryotes (e.g., mammals, fungi, and protozoans), the prevailing hypothesis was that they were already present in the early eukaryotic ancestor. However, analyses of the molecular composition of DCGs in mammals and ciliates (groups that diverged more than 1 billion years ago) failed to find homolog proteins, and some molecular mechanisms involved in DCGs processes are fundamentally different in these groups. Such observations support the idea that these secretory granules had independent origins in distinct phylogenetic lineages.

A supplementary example comes from clathrin proteins involved in endocytosis. This process is highly conserved in eukaryotes, and there is strong evidence for the shared inheritance of the underlying genes. Nevertheless, recent duplications of clathrin genes during chordate evolution generated distinct isoforms with tissue-specific functions (Elde *et al.*, 2007). This instance illustrates how the inheritance of ancient genes, combined with lineage-specific duplications, can lead to functional divergence and adaptation. The possibility that other common cellular features are the result of convergent processes, rather than representing shared inheritance, provides an exciting alternative view of the evolutionary history of complex structures.

Perspectives

The recruitment of new genes to play a role in fundamental cellular processes may be a general and recurrent theme in evolution, rather than represent exceptional events. Table 2 summarizes several new genes documented in *Drosophila* with basic cellular functions. Note that these genes originated in diverse phylogenetic branches in the *Drosophila* history, ranging from 0 to 25 million years ago (for the sake of comparison, this upper limit corresponds to the time that separates humans from our primate cousins *Rhesus*, a relatively recent event in evolutionary terms). In the near future, the collection of additional complete genomes will improve our ability to discover and date the origin of genes more precisely. Allied to experimental molecular biology techniques, this knowledge will allow a more mechanistic understanding of the functional evolution of new genes, and of biological functions in general. We believe that future cases of novel genes implicated in basic cellular processes will underscore their relevance for adaptation.

The evolutionary relevance of new genetic elements also points to the appealing perspective of their significance in shaping human-specific traits, as well as contributing to diseases. Despite variable estimates, we can conservatively conclude that at least 300 genes are specific of the human lineage, and other 1000 are only found in primates (Zhang and Long, 2014). Strikingly, transcriptome analyses revealed that a larger proportion of new genes are expressed in human brains during development when compared to mouse, and also suggested

Table 2 Examples of new genes involved in basic cellular processes in *Drosophila* species

Gene name	Gene age (million years)	Cellular process
Jingwei	0–3	Alcohol metabolism
p24-related-2	0–3	Golgi vesicle transport
Cyp9f2	3–6	Intracellular oxidoreductase activity
Xcbp1	3–6	Chaperone
TfIIA-S-2	6–11	Transcription factor
Ribosomal protein S28a	6–11	Ribosomal structural constituent
Methuselah-like 6	6–11	Membrane G-protein coupled receptor
ACXA	11–25	Intracellular signal transduction
Chemosensory protein B 93b	11–25	Detection of pheromones
Scavenger receptor class C, type III	11–25	Cellular internalization of foreign material
Lysozyme X	11–25	Lysozyme activity, defense against bacteria
Ribosomal protein L37b	11–25	Ribosomal structural constituent

that these genes originated in the same evolutionary period during which human neocortex was expanding (Zhang *et al.*, 2011). Such findings demonstrate how genome-wide expression comparisons can be used to shed light on long-standing questions regarding human traits. Additionally, detailed experimental investigations have shown the implication of young genes in diseases. Some remarkable examples include genes related to tumors, such as the hominoid-specific *TBC1D3* (Wainszelbaum *et al.*, 2008), which is involved in cellular signaling and trafficking; and the primate-specific *CT45A1*, which upregulates oncogenic and metastatic genes (Shang *et al.*, 2014). Moreover, these examples illustrate how evolutionary and cellular approaches can be combined to explore relevant questions.

The finding that distant related taxa shares many common cellular processes, and even ancient genes, may give the false impression that basic functions, and the genes underlying them, are not affected by evolution, and remain unaltered at the molecular level. This view is not in accordance with our modern view of evolution. The existence of deeply conserved processes does not imply that their underlying components have been static over evolutionary time. Cellular processes are highly dynamic, and so is their evolution. During their evolutionary history, organisms face diverse challenges, such as changing environments, competition, and pathogenic attacks. Therefore, it is not unexpected that the genetic elements interacting in a network evolve over time, improving and renewing the overall system. Furthermore, even when a certain process is conserved across long evolutionary periods (such as the cell division mechanisms mentioned above), the molecular components that produce it may undergo a dynamic turnover. With these examples, our aim is to argue for the idea that, far from providing just anecdotal reports, the evolution of new gene functions can help to illuminate main

evolutionary mechanisms and understand biological processes and structures.

Acknowledgments

Iuri Matteuzzo Ventura is supported by the Science without Borders scholarship (BEX18816/12-6), and Manyuan Long is supported by U.S. National Institutes of Health (R01GM100768-01A1), National Science Foundation (NSF1051826) and Endowment of Edna K. Papazian Distinguished Professorship at the University of Chicago. We thank all lab members for discussions, and Marilia Lopes Justino for helping with the figure.

References

- Adams, M.D., Celniker, S.E., Holt, R.A., *et al.*, 2000. The genome sequence of *Drosophila melanogaster*. *Science* 287, 2185–2195.
- Arthur, R.K., Ma, L., Slattery, M., *et al.*, 2014. Evolutionary conservation of H3K27me3 marked chromatin is bypassed by gene duplication. *Genome Research* 24, 1115–1124.
- Betran, E., Thornton, K., Long, M., 2002. Retroposed new genes out of the X in *Drosophila*. *Genome Research* 12, 1854–1859.
- Blattner, F.R., Plunkett, G., Bloch, G.A., *et al.*, 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* 277, 1453–1462.
- C. elegans* Sequencing Consortium, 1998. Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* 282, 2012–2017.
- Charlesworth, B., Coyne, A.J., Barton, N.H., 1987. The relative rates of evolution of sex chromosomes and autosomes. *American Naturalist* 130, 113–146.
- Cheeseman, M., Desai, A., 2008. Molecular architecture of the kinetochore-microtubule interface. *Nature Reviews Molecular Cell Biology* 9, 33–46.
- Chen, S., Krinsky, B.H., Long, M., 2013. New genes as drivers of phenotypic evolution. *Nature Reviews Genetics* 14, 645–660.
- Chen, S., Ni, X., Krinsky, B.H., *et al.*, 2012. Reshaping of global gene expression networks and sex-biased gene expression by integration of a young gene. *EMBO Journal* 31, 2798–2809.
- Chen, S., Zhang, Y., Long, M., 2010. New genes in *Drosophila* quickly become essential. *Science* 330, 1682–1685.
- Collins, F.S., Green, E.D., Guttmacher, A.E., Guyer, M.S., 2003. A vision for the future of genomics research. *Nature* 422, 835–847.
- Díaz-Castillo, C., Ranz, J.M., 2012. Nuclear chromosome dynamics in the *Drosophila* male germ line contribute to the nonrandom genomic distribution of retrogenes. *Molecular Biology and Evolution* 29, 2105–2108.
- Ding, Y., Zhao, L., Yang, S., *et al.*, 2010. A young *Drosophila* duplicate gene plays essential roles in spermatogenesis by regulating several y-linked male fertility genes. *PLoS Genetics* 6, e1001255.
- Drosophila* 12 Genomes Consortium, 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450, 203–218.
- Durbin, R.M., Abecasis, G.R., Altshuler, R., 2010. A map of human genome variation from population-scale sequencing. *Nature* 467, 1061–1073.
- Elde, N.C., Long, M., Turkewitz, A.P., 2007. A role for convergent evolution in the secretory life of cells. *Trends in Cell Biology* 17, 157–164.
- Emerson, J.J., Kaessmann, H., Betran, E., Long, M., 2004. Extensive gene traffic on the mammalian X chromosome. *Science* 303, 537–540.
- Fleischmann, R.D., Adams, M.D., White, O., *et al.*, 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269, 496–512.
- Goffeau, A., Barrell, B.G., Bussey, H., 1996. Life with 6000 genes. *Science* 274, 546–567.
- Hardison, R.C., 2003. Comparative Genomics. *PLoS Biology* 1, e58.
- International Human Genome Sequencing Consortium, 2001. Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.
- Kernkner, C., Long, M., 2014. New genes important for development. *EMBO Reports* 15, 460–461.
- Koboldt, D.C., Steinberg, K.M., Larson, D.E., Wilson, R.K., Mardis, E.R., 2013. The next-generation sequencing revolution and its impact on genomics. *Cell* 155, 27–38.
- Long, M., VanKuren, N.W., Chen, S., Vibranovski, M.D., 2013. New gene evolution: Little did we know. *Annual Review of Genetics* 47, 325–351.

- Long, M., Vibranovski, M.D., Zhang, Y.E., 2012. Evolutionary interactions between sex chromosomes and autosomes. In: Singh, R., Xu, J., Kulathinal, R. (Eds.), *Rapidly Evolving Genes and Genetic Systems*, first ed. Oxford: Oxford University Press, pp. 101–114.
- Mackay, T.F.C., 2014. Epistasis and quantitative traits: Using model organisms to study gene–gene interactions. *Nature Reviews Genetics* 15, 22–33.
- Meisel, R.P., Han, M.V., Hahn, M.W., 2009. A complex suite of forces drives gene traffic from *Drosophila* X chromosomes. *Genome Biology and Evolution* 1, 176–188.
- Miklos, G.L., Rubin, G.M., 1996. The role of the genome project in determining gene function: Insights from model organisms. *Cell* 86, 521–529.
- Mouse Genome Sequencing Consortium, 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520–562.
- Neafsey, D.E., Waterhouse, R.M., Abai, M.R., *et al.*, 2015. Highly evolvable malaria vectors: The genomes of 16 *Anopheles* mosquitoes. *Science* 347, 1258522.
- Quezada-Diaz, J.E., Mulyil, T., Rio, J., Betran, E., 2010. Drcd-1 related: A positively selected spermatogenesis retrogene in *Drosophila*. *Genetica* 138, 925–937.
- Rice, W.R., 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution* 38, 735–742.
- Ross, B.D., Rosin, L., Thomae, A.W., *et al.*, 2013. Stepwise evolution of essential centromere function in a *Drosophila* neogene. *Science* 340, 1211–1214.
- Shang, B., Gao, A., Pan, Y., *et al.*, 2014. *CT45A1* acts as a new proto-oncogene to trigger tumorigenesis and cancer metastasis. *Cell Death & Disease* 5, e1285.
- The Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815.
- Venter, J.C., Adams, M.D., Myers, E.W., *et al.*, 2001. The sequence of the human genome. *Science* 291, 1304–1351.
- Vibranovski, M.D., 2014. Meiotic sex chromosome inactivation in *Drosophila*. *Journal of Genomics* 2, 104–117.
- Vibranovski, M.D., Lopes, H.F., Karr, T.L., Long, M., 2009a. Stage-specific expression profiling of *Drosophila* spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis expressed genes. *PLoS Genetics* 5, e1000731.
- Vibranovski, M.D., Zhang, Y., Long, M., 2009b. General gene movement off the X chromosome in the *Drosophila* genus. *Genome Research* 19, 897–903.
- Vicoso, B., Charlesworth, B., 2009. The deficit of male-biased genes on the *D. melanogaster* X is expression-dependent: A consequence of dosage compensation? *Journal of Molecular Evolution* 68, 576–583.
- Wainszelbaum, M.J., Charron, A.J., Kong, C., *et al.*, 2008. The hominoid-specific oncogene TBC1D3 activates ras and modulates epidermal growth factor receptor signaling and trafficking. *Journal of Biological Chemistry* 283, 13233–13242.
- Wang, J., Long, M., Vibranovski, M.D., 2012. Retrogenes moved out of the Z chromosome in the silkworm. *Journal of Molecular Evolution* 74, 1–14.
- Zhang, Y.E., Landback, P., Vibranovski, M.D., Long, M., 2011. Accelerated recruitment of new brain development genes into the human genome. *PLoS Biology* 9, e1001179.
- Zhang, Y.E., Landback, P., Vibranovski, M.D., Long, M., 2012. New genes expressed in human brains: Implications for annotating evolving genomes. *Bioessays* 34, 982–991.
- Zhang, Y.E., Long, M., 2014. New genes contribute to genetic and phenotypic novelties in human evolution. *Current Opinion in Genetics & Development* 29, 90–96.

Transcriptomics

EA Milward, A Shahandeh, and M Heidari, The University of Newcastle, Callaghan, NSW, Australia

DM Johnstone, University of Sydney, Sydney, NSW, Australia

N Daneshi and H Hondermarck, The University of Newcastle, Callaghan, NSW, Australia

© 2016 Elsevier Inc. All rights reserved.

Glossary

Competing endogenous ribonucleic acid (ceRNA) A form of RNAs molecules that regulate other RNA transcripts by competing for shared microRNAs.

Complementary deoxyribonucleic acid (cDNA) A form of DNA synthesized from a messenger RNA template in a reaction catalyzed by the enzyme reverse transcriptase.

Expressed sequence tags (ESTs) A short subsequence of a cDNA sequence.

Fluorescent *in situ* RNA sequencing (FISSEQ) A method of sequencing DNA by the use of colonies and cycles of fluorescent deoxyribonucleotide triphosphate (dNTP) incorporation.

Long noncoding RNA (lncRNA) Nonprotein coding transcripts longer than 200 nucleotides.

Messenger RNA (mRNA) A class of RNA molecules functions as the template for protein synthesis.

MicroRNA (miRNA) A class of naturally occurring, small noncoding RNA molecules which functions in RNA silencing and posttranscriptional regulation of gene expression.

Noncoding RNA (ncRNA) A class of RNA molecules which is not translated into a protein.

Piwi-interacting RNA (piRNA) A class of small noncoding RNAs that interacts with piwi-domain-containing proteins.

RNA sequencing (RNA-seq) A technique which is used to map the transcribed.

Sequencing by oligonucleotide ligation and detection (SOLiD) A next-generation sequencing platform.

Serial analysis of gene expression (SAGE) A technique to produce a snapshot of the messenger RNA in the form of small tags that correspond to fragments of those transcripts.

Single-molecule real-time sequencing (SMRT) A parallelised single molecule DNA sequencing by synthesis technology.

Small nuclear RNA (snRNA) A class of small RNA molecules confined to the nucleus.

Small nucleolar RNA (snoRNA) A class of small RNAs molecules which is associated with the eukaryotic nucleus components of small nucleolar ribonucleoproteins.

snoRNA-derived RNA (sdRNA) A small noncoding RNA of uniform length and with distinct sequences that are ubiquitously detected across eukaryotes.

Introduction

Transcriptomics is the study of the ‘transcriptome,’ a term whose first use, to signify an entire set of transcripts, has been attributed to Charles Auffray (Pietu *et al.*, 1999). The term transcriptome is now widely understood to mean the complete set of all the ribonucleic acid (RNA) molecules expressed in some given entity, such as a cell, tissue, or organism (Morozova *et al.*, 2009; Wolf, 2013).

Transcriptomics encompasses everything relating to RNAs. This includes their transcription and expression levels, functions, locations, trafficking, and degradation. It also includes the structures of transcripts and their parent genes with regard to start sites, 5′ and 3′ end sequences, splicing patterns, and posttranscriptional modifications (Wang *et al.*, 2009). Transcriptomics covers all types of transcripts, including messenger RNAs (mRNAs), microRNAs (miRNAs), and different types of long noncoding RNAs (lncRNAs).

Modern transcriptomics uses high-throughput methods to analyze the expression of multiple transcripts in different physiological or pathological conditions and this is rapidly expanding our understanding of the relationships between the transcriptome and the phenotype across a wide range of living entities.

Evolution of High-Throughput Transcriptomic Technologies

The comprehensive study of the transcriptome has not been feasible until recently because capturing full catalogs of transcripts and their variations is complicated and requires sophisticated technology. Over the years, several different kinds of high-throughput technologies have been developed to profile and quantify transcript expression. Much of the early knowledge of the transcriptome was derived from gene predictions and traditional methodologies that used complementary deoxyribonucleic acid (cDNA) clones to generate expressed sequence tags (ESTs), followed by sequencing using automated Sanger sequencing technology, as exemplified by the human genome project (Adama *et al.*, 1991). One pioneering quantitative transcriptomic study employed serial analysis of gene expression (SAGE) on 1000 tags to characterize the gene expression pattern in human pancreas (Velculescu *et al.*, 1995). In 1995, early high-capacity microarrays containing cDNA spotted onto microscope-sized glass slides were developed and applied to measure expression of 45 Arabidopsis genes in parallel (Schena *et al.*, 1995).

Recent years have seen the emergence of RNA sequencing (RNA-seq), which encompasses a variety of next-generation sequencing techniques for determining the sequence and

potentially also the levels of RNA transcripts (Chu and Corey, 2012; Nagalakshmi *et al.*, 2008; Wilhelm *et al.*, 2008; Mortazavi *et al.*, 2008). At present, sequencing technology still involves several challenges for users, including cost, availability, and the complexity and error vulnerability of sequence assembly. For these reasons, array-based technology is still widely used in transcriptome profiling. However, compared to the new sequencing approaches, which are able to deliver readings on exact transcription boundaries and unmapped transcripts without requiring predefined gene or transcript probes (Metzker, 2009; Wang *et al.*, 2009), array technology is limited by the probe-based nature of microarrays.

With the advent of RNA-seq, the possibility of sequencing the entire transcriptome has become a reality (Ozsolak and Milos, 2010; Marguerat and Bahler, 2010). Only the more abundant transcripts are usually detected by traditional approaches like EST sequencing whereas RNA-seq, when employed with adequate sequencing depth (100–1000 reads per base pair of a transcript), is capable of delivering an almost complete capture of a transcriptome (Martin and Wang, 2011). Furthermore, RNA-seq allows for a full genome-wide assessment of transcripts, with a wide dynamic range of expression levels (Nookaew *et al.*, 2012). Unlike current array technologies, it can also be used as a discovery tool to identify novel transcripts such as lncRNAs, in addition to applications in analysis of differential gene expression and alternative splicing of transcripts.

As summarized in the left panel of Figure 1, most RNA-seq approaches to date have used indirect methodologies in which

extracted mRNAs or total RNAs are first converted into a library of cDNAs containing sequencing adapters. A high-throughput DNA sequencing technology is then used to sequence the cDNA fragments from one end or both ends, rendering an output which is comprised of short sequences. The resulting short reads are preprocessed to minimize sequencing errors and then assembled into longer sequences putatively corresponding to the RNAs in the original sample. By sequencing a large number of short transcripts in parallel and piecing together the resultant short reads, it is now becoming possible to collect complete sets of transcriptome data (Metzker, 2009).

Most next-generation sequencing platforms have typically delivered quite short sequence reads (35–500 bp) (Metzker, 2009; Shendure and Ji, 2008). This then requires high-powered computing systems with large memory and many cores to run the parallel algorithms needed to reconstruct the full-length transcripts. Such sequencing platforms have relatively high error rates and raise informatics challenges, due to the short reads and sometimes biased coverage (Guo *et al.*, 2012; Wang *et al.*, 2009; Bahassi and Stambrook, 2014).

With the rapid evolution of assembly algorithms and improvements in data quality, newer RNA sequence technologies, for example single-molecule, real-time sequencing technology (SMRT) or nanopore sequencers, described further below (Mikheyev and Tin, 2014; Maitra *et al.*, 2012), are addressing the above challenges and providing longer reads up to several kilobases, which make them capable of fully sequencing a transcript in a single run.

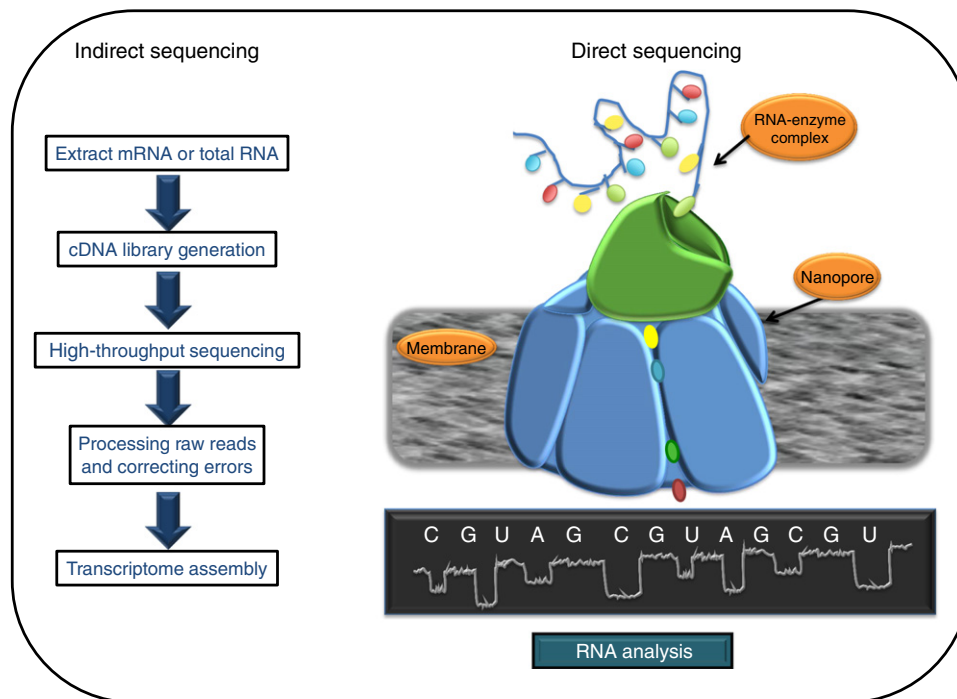


Figure 1 Indirect and direct approaches to RNA-seq. Early high-throughput approaches to RNA-seq have utilized cDNA (left) but new alternatives such as nanopore sequencing (right) can sequence RNA directly. In one implementation of this strategy, the RNA strand binds to an enzyme or other immobilizing molecule (green) which ratchets the RNA one base at a time through a molecular nanopore channel (blue) embedded in a membrane. The nanopore conducts an electrical current across the membrane. As each RNA base passes through the nanopore it disrupts the current in a characteristic way, allowing the base to be identified.

The goal now is to extend read lengths to eliminate the need for assembly. On the SMRT platform, the average read length is approximately 3000 bp but can extend to over 20 000 bp, making it possible to sequence most RNA transcripts in a single read, although additional reads would be needed to ensure high-quality coverage and improve throughput (Martin and Wang, 2011). Many scientists now believe that continuous improvement in sequencing protocols and increases in throughput will shape the future of transcriptome sequencing to the point that no assembly will be required (Bahassi and Stambrook, 2014; Martin and Wang, 2011).

Another recent advance is to combine RNA-seq with fluorescent *in situ* hybridization to generate information simultaneously on both the levels and the cellular localization of transcripts. Termed fluorescent *in situ* RNA-seq (FISSEQ), RNA is reverse transcribed *in situ* and the cDNA copied by rolling-circle amplification to generate DNA 'nanoballs,' which are then sequenced in cells using 'sequencing by oligonucleotide ligation and detection' (SOLiD) technology. The SOLiD technique involves the sequential hybridization of fluorescently labeled two base probes, which can then be imaged to determine the sequence of the transcript contained within a nanoball. In the initial publication describing the FISSEQ technique (Lee *et al.*, 2014), the authors successfully sequenced ~12 500 nanoballs in primary human fibroblasts, with a median base error rate of only 0.64%. Although this emerging technique will require further optimization before it is adopted more widely, the ability to simultaneously generate sequence and positional information is an important breakthrough in transcriptomics research.

Most available high-throughput technologies require RNA conversion into cDNA but novel approaches such as nanopore sequencing (Ayub and Bayley, 2012) now make it possible to sequence RNA directly (Figure 1). This eliminates the need for reverse transcription and sequence assembly, two important inherent sources of error in the indirect approaches presently being used. Direct sequencing technologies currently under development entail passing the RNA one base at a time through either a solid state nanopore or a molecular nanopore positioned across an electrical field. As each different RNA base (A, C, G, or U) is fed through the pore it produces a distinctive change in the current flow across the field. However, this occurs too rapidly to detect using existing technology so the RNA is first transiently immobilized, for example by binding to some kind of capture molecule that acts as a decelerating gear to slow the passage of RNA bases through the pore, allowing current changes to be detected. Proteins such as α -hemolysin, which is able to form heptameric channels in biomembranes, have been used successfully as molecular nanopores in direct DNA sequencing and can be adapted for RNA-seq by targeted mutation, in combination with biotin-streptavidin binding or other capture strategies to immobilize the RNA strand.

One important caveat to keep in mind for all these methods is that when large numbers of measurements are made simultaneously, the number of errors is also potentially large, even when the error rate is very low. The accuracy of any high-throughput data, whether from microarrays or next-generation sequencing, should therefore always be validated by an

alternative procedure, such as polymerase chain reaction (PCR) (Milward *et al.*, 2012; Piepenburg *et al.*, 2006).

Exploring the Interactive Transcriptome

The technological advances described above are not only extending our knowledge of protein-coding mRNA molecules but are also shedding new light on some of the previously unrecognized or underappreciated noncoding components of the transcriptome and how the interactions among these components modify the transcriptomic profile under different conditions.

Ribosomal RNA and transfer RNA have been recognized since the 1950s as having biological functions other than protein coding and more recent advances in RNA technologies have led to the identification of many different classes of 'noncoding' RNAs (ncRNAs). Most ncRNAs are now believed to have important functions in regulating other RNAs and consequently in modulating the transcriptome.

For example, miRNA have long been recognized for their ability to bind to and induce degradation of target mRNA transcripts (Lee *et al.*, 1993); small nuclear RNAs (snRNAs) have critical, possibly even catalytic, roles in the splicing of mRNA transcripts (Matera and Wang, 2014; Valadkhan *et al.*, 2009); and small nucleolar RNAs (snoRNA) are also primarily associated with modification of other RNA species, including rRNA and snRNA. Furthermore, snoRNA itself is generally derived from spliced introns of transcripts, which escape the usual fate of degradation by forming protein complexes, and may undergo further processing to liberate snoRNA-derived RNA (sdRNA), an ncRNA similar to miRNA that in turn may have roles in regulating gene expression (Bratkovic and Rogeljić, 2014; Falaleeva and Stamm, 2013). In addition, recent studies have revealed that lncRNA can influence transcription by directly interacting with the transcription machinery or specific transcription factors. At the posttranscriptional level, lncRNAs can influence mRNA processing (e.g., as natural antisense transcripts) (Beltran *et al.*, 2008; Faghghi and Wahlestedt, 2009), translation control, and mRNA stability.

To get an idea of the probable scale of transcriptomic interactions, the class of mature miRNAs in humans alone contains over 2500 different members (miRBase 21), the class of human lncRNAs has over 10 000 members, and the class of Piwi-interacting RNA (piRNA), which has roles in epigenetic gene regulation through transcriptional gene silencing, is even larger, with perhaps hundreds of thousands or even millions of distinct piRNAs (Brennecke *et al.*, 2007; Siomi *et al.*, 2011; Ross *et al.*, 2014). The potential intricacy is even more apparent when it is realized that many ncRNAs show tissue- and time-specific expression (Landgraf *et al.*, 2007).

Another fascinating class of RNA molecules whose membership is being rapidly expanded through deep sequencing approaches is circular RNA, which can consist of exons only, introns only, or both. Alternative circularization, in conjunction with alternative splicing, appears capable of producing multiple circular RNAs from individual human genes at thousands of different genomic loci, suggesting that this class also contains a large number of members, again adding a

further layer to the transcriptome and its regulation (Zhang *et al.*, 2014).

In addition, over one hundred RNA modifications have now been identified, including methylation, uridylation, and other tailing motifs, providing yet another level of complexity superimposed on the various transcriptome elements described above (Lee *et al.*, 2014).

The recognition that there is a considerable degree of cross-talk between different RNA species has also led to the concept of competing endogenous RNA (ceRNA) (Salmena *et al.*, 2011), a network of RNA species that regulates one another by competing for binding to shared miRNAs. Through this competition, ceRNA populations act as a natural miRNA ‘sponge’ by titrating miRNA availability. Although still in its infancy, a number of recent discoveries provide support for this hypothesis, most notably the ceRNA network involving the tumor suppressor *PTEN* (Poliseno *et al.*, 2010). This study reports evidence for biological activity of the mRNAs of certain oncogenes such as *PTEN* and *KRAS* (independent of their protein coding functions), as well as for biological activity of the transcripts of their pseudogenes, including roles in growth suppression. The antisense transcript of the *PTEN* pseudogene likewise displays biological activity, for example in cell cycle arrest (Johnsson *et al.*, 2013).

Although the full complexity of transcriptome interactions remains to be unraveled, the examples above serve to illustrate the extensive and multifaceted nature of the interconnections and cross talk among different ncRNA species. From the perspective of transcriptomics, not only is it important to determine the different types and relative amounts of the various ncRNA species, but it is also essential to understand how they interact to regulate each other and thereby shape the transcriptome.

In addition to effects on transcription and mRNA, there are several examples of RNAs acting in other cellular contexts; for example, lncRNAs have been reported to have functions in regulation of protein activity, scaffolding for protein complexes, intercellular signaling through exosomes, and recombination, as reviewed elsewhere (Geisler and Collier, 2013). As little is currently known about actions of RNAs on other cellular components, it will be interesting to see what other functions are ascribed to RNAs in the future.

New Technologies for mRNA Transcriptomics

Although RNA-seq has the capacity to indiscriminately assess the full complement of RNA types (i.e., both coding and noncoding) found in a given cell or tissue, more focused methodologies are being developed for studying the protein-coding transcriptome, cutting down costs associated with unbiased sequencing. For example, Halvardson *et al.* (2013) have described a method that marries aspects of exome and RNA-seq to improve the detection of mRNA expressed at low levels. This method, termed ExomeRNAseq, first applies whole exome capture to cDNA, followed by conventional RNA-seq of the enriched transcript pool. Although this approach appears successful in detecting transcripts with low expression and identifying novel exons and splice isoforms, it can reduce quantification accuracy. An alternative way to enrich mRNA

prior to sequencing is the use of immobilized oligo-deoxythymidine to select polyadenylated RNAs, thereby removing highly abundant rRNA and non-polyadenylated transcripts such as ncRNA (Sims *et al.*, 2014). Such enrichment also increases the proportion of sequence reads that map across splice junctions, improving the discovery and identification of alternative splice variants. Current mRNA enrichment techniques facilitate RNA-seq experiments that can effectively define the components of the coding transcriptome of a particular cell or tissue, however, more research is required to develop experimental and computational techniques that provide more accurate quantification of transcripts for differential expression analysis.

Integrating Transcriptomics with Other Omics

The integration of genomics with transcriptomics, also termed functional genomics, has led to a better understanding of the relationship between genotypes and phenotypes, significantly impacting the fields of biology and medicine. However, more comprehensive approaches that also combine transcriptomics approaches such as RNA-seq with complementary proteomics technologies are required for deeper understanding of the relationship between genotypes and phenotypes. This commonly entails proteomics validation of predictions based on transcriptome data, and so differs from earlier implementations of proteogenomics, where protein predictions derived from reference genome sequence information are assessed using proteomics (Wang *et al.*, 2014).

Mass spectrometry technologies, notably high-throughput liquid chromatography–tandem mass spectrometry-based proteomics approaches, have been applied fruitfully on various fronts (Wang *et al.*, 2014; Woo *et al.*, 2013). At a foundational level, high-throughput proteomics information can assist considerably in ongoing annotation of both the transcriptome and the genome and, conversely, RNA-seq data, including transcript abundance and sequence variation and splicing variation, can improve protein identification in proteomics (Wang *et al.*, 2014). There are also many other potential applications, including computation of sample specific expression and identification of genomic variants (Woo *et al.*, 2013).

However, combining proteomics with transcriptomics data is complicated by the restricted correlations of these two omic spaces. Notably, with regard to identity, there are often non-linear relationships between a parental genomic sequence and its derivative RNAs and protein variants (if any), an obvious example being the ncRNAs, as discussed above. With regard to abundance, correlations are restricted by real biological divergence (Gawron *et al.*, 2014; Van Damme *et al.*, 2014; Vogel and Marcotte, 2012) as well as by current differences in detection limits in these spaces. Protein abundance is regulated not only by mRNA translation rates in both prokaryotes and eukaryotes, but also by posttranslational modifications of proteins that greatly influence protein degradation and turnover (Vogel and Marcotte, 2012). The half-life of proteins is controlled by various pathways such as, but not limited to, the ubiquitin and proteasome pathways. The presence or absence of posttranslation modifications (which cannot be predicted

by transcriptomics), such as phosphorylation, glycosylation, or ubiquitinylation, has a strong impact on protein stability, adding a further level of complexity. Together the composition and dynamic of the proteome is not deducible from functional genomics data and as a consequence, proteomics is an indispensable and complementary approach to transcriptomics.

While it is difficult to integrate the huge bodies of data that can arise from multiple different transcriptomic and proteomic datasets derived from different samples, a range of new bioinformatics tools are now emerging that make inroads into overcoming these problems, for example by methods that compress data size without loss of sensitivity by constructing more compact nonredundant datasets (Woo *et al.*, 2013). A more complete realization of proteogenomics, in which the three levels of DNA and epigenetic modifications, RNAs and their epitranscriptomic modifications, and proteins and their posttranslational modifications, are studied simultaneously in the process of conceptual and technical development and promises to be an important milestone in systems biology (Renuse *et al.*, 2011).

Conclusion

The advent of high-throughput sequencing technologies is providing unprecedented insights into the composition and complex interactions of the transcriptome. In the space of less than two decades, our understanding of the transcriptome has been transformed from a simplistic view that RNA species serve solely as intermediaries between DNA and protein, to one where interactions between protein-coding RNAs and a plethora of ncRNAs amplify the complexity of multicellular organisms. With future advances in sequencing technology and computational capacity, it is likely that the transcriptome will reveal itself to be both far more intricate and far more influential than ever before imagined.

See also: Dynamic Integration: Dynamics: Transcription Factor Networks

References

- Adama, M.D., Kelley, J.M., Gocayne, J.D., *et al.*, 1991. Complementary DNA sequencing: Expressed sequence tags and human genome project. *Science* 252, 1651–1656.
- Ayub, M., Bayley, H., 2012. Individual RNA base recognition in immobilized oligonucleotides using a protein nanopore. *Nano Letters* 12, 5637–5643.
- Bahassi, E.M., Stambrook, P.J., 2014. Next-generation sequencing technologies: Breaking the sound barrier of human genetics. *Mutagenesis* 29, 303–310.
- Beltran, M., Puig, I., Pena, C., *et al.*, 2008. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial–mesenchymal transition. *Genes & Development* 22, 756–769.
- Bratkovic, T., Rogeljić, B., 2014. The many faces of small nucleolar RNAs. *Biochimica et Biophysica Acta* 1839, 438–443.
- Brennecke, J., Aravin, A.A., Stark, A., *et al.*, 2007. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* 128, 1089–1103.
- Chu, Y., Corey, D.R., 2012. RNA sequencing: Platform selection, experimental design, and data interpretation. *Nucleic Acid Therapeutics* 22, 271–274.
- Faghghi, M.A., Wahlested, C., 2009. Regulatory roles of natural antisense transcripts. *Nature Reviews Molecular Cell Biology* 10, 637–643.
- Falaleeva, M., Stamm, S., 2013. Processing of snoRNAs as a new source of regulatory non-coding RNAs: snoRNA fragments form a new class of functional RNAs. *BioEssays: News & Reviews in Molecular, Cellular & Developmental Biology* 35, 46–54.
- Gawron, D., Gevaert, K., Van Damme, P., 2014. The proteome under translational control. *Proteomics* 14 (23–24), 2647–2662.
- Geisler, S., Collier, J., 2013. RNA in unexpected places: Long non-coding RNA functions in diverse cellular contexts. *Nature Reviews Molecular Cell Biology* 14, 699–712.
- Guo, Y., Li, J., Li, C.I., *et al.*, 2012. The effect of strand bias in Illumina short-read sequencing data. *BMC Genomics* 13, 666.
- Halvardson, J., Zaghool, A., Feuk, L., 2013. Exome RNA sequencing reveals rare and novel alternative transcripts. *Nucleic Acids Research* 41, e6.
- Johnsson, P., Ackley, A., Vidarsdottir, L., *et al.*, 2013. A pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human cells. *Nature Structural & Molecular Biology* 20, 440–446.
- Landgraf, P., Rusu, M., Sheridan, R., *et al.*, 2007. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* 129, 1401–1414.
- Lee, J.H., Daugherty, E.R., Scheiman, J., *et al.*, 2014. Highly multiplexed subcellular RNA sequencing in situ. *Science* 343, 1360–1363.
- Lee, R.C., Feinbaum, R.L., Ambros, V., 1993. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* 75, 843–854.
- Maitra, R.D., Kim, J., Dunbar, W.B., 2012. Recent advances in nanopore sequencing. *Electrophoresis* 33, 3418–3428.
- Marguerat, S., Bahler, J., 2010. RNA-seq: From technology to biology. *Cellular & Molecular Life Sciences* 67, 569–579.
- Martin, J.A., Wang, Z., 2011. Next-generation transcriptome assembly. *Nature Reviews Genetics* 12, 671–682.
- Matera, A.G., Wang, Z., 2014. A day in the life of the spliceosome. *Nature Reviews. Molecular Cell Biology* 15, 108–121.
- Metzker, M.L., 2009. Sequencing technologies – the next generation. *Nature Reviews Genetics* 11, 31–46.
- Mikheyev, A.S., Tin, M.M., 2014. A first look at the Oxford nanopore MinION sequencer. *Molecular Ecology Resources* 14 (6), 1097–1102.
- Milward, E.A., Daneshi, N., Johnstone, D.M., 2012. Emerging real-time technologies in molecular medicine and the evolution of integrated ‘pharmacomics’ approaches to personalized medicine and drug discovery. *Pharmacology & Therapeutics* 136, 295–304.
- Morozova, O., Hirst, M., Marra, M.A., 2009. Applications of new sequencing technologies for transcriptome analysis. *Annual Review of Genomics & Human Genetics* 10, 135–151.
- Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., Wold, B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5, 621–628.
- Nagalakshmi, U., Wang, Z., Waern, K., *et al.*, 2008. The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* 320, 1344–1349.
- Nookaew, I., Papini, M., Pornputtpong, N., *et al.*, 2012. A comprehensive comparison of RNA-Seq-based transcriptome analysis from reads to differential gene expression and cross-comparison with microarrays: A case study in *Saccharomyces cerevisiae*. *Nucleic Acids Research* 40, 10084–10097.
- Ozsolak, F., Milos, P.M., 2010. RNA sequencing: Advances, challenges and opportunities. *Nature Reviews Genetics* 12, 87–98.
- Piepenburg, O., Williams, C.H., Stemple, D.L., Armes, N.A., 2006. DNA detection using recombination proteins. *Plos Biology* 4, e204.
- Pietu, G., Mariage-Samson, R., Fayein, N.A., *et al.*, 1999. The genexpress IMAGE knowledge base of the human brain transcriptome: A prototype integrated resource for functional and computational genomics. *Genome Research* 9, 195–209.
- Poliseno, L., Salmena, L., Zhang, J., *et al.*, 2010. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 465, 1033–1038.
- Renuse, S., Chaekady, R., Pandey, A., 2011. Proteogenomics. *Proteomics* 11, 620–630.
- Ross, R.J., Weiner, M.M., Lin, H., 2014. PIWI proteins and PIWI-interacting RNAs in the soma. *Nature* 505, 353–359.
- Salmena, L., Poliseno, L., Tay, Y., Kats, L., Pandolfi, P.P., 2011. A ceRNA hypothesis: The Rosetta Stone of a hidden RNA language? *Cell* 146, 353–358.
- Schena, M., Shalon, D., Davis, R.W., Brown, P.O., 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467–470.
- Shendure, J., Ji, H., 2008. Next-generation DNA sequencing. *Nature Biotechnology* 26, 1135–1145.
- Sims, D., Sudbery, I., Illott, N.E., Heger, A., Ponting, C.P., 2014. Sequencing depth and coverage: Key considerations in genomic analyses. *Nature Reviews Genetics* 15, 121–132.

- Siomi, M.C., Sato, K., Pezic, D., Aravin, A.A., 2011. PIWI-interacting small RNAs: The vanguard of genome defence. *Nature Reviews Molecular Cell Biology* 12, 246–258.
- Valadkhan, S., Mohammadi, A., Jaladat, Y., Geisler, S., 2009. Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proceedings of the National Academy of Sciences* 106, 11901–11906.
- Van Damme, D., Minke, B.A., Fitzmaurice, W., Kholodenko, B.N., Kolch, W., 2014. Systems biology-embedded target validation: Improving efficacy in drug discovery. *Wiley Interdisciplinary Reviews: Systems Biology & Medicine* 6, 1–11.
- Velculescu, V.E., Zhang, L., Vogelstein, B., Kinzler, K.W., 1995. Serial analysis of gene expression. *Science* 270, 484–487.
- Vogel, C., Marcotte, E.M., 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics* 13, 227–232.
- Wang, X., Liu, Q., Zhang, B., 2014. Leveraging the complementary nature of RNA-seq and shotgun proteomics data. *Proteomics* 14 (23–24), 2676–2687.
- Wang, Z., Gerstein, M., Snyder, M., 2009. RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics* 10, 57–63.
- Wilhelm, B.T., Marguerat, S., Watt, S., *et al.*, 2008. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* 453, 1239–1243.
- Wolf, J.B., 2013. Principles of transcriptome analysis and gene expression quantification: An RNA-seq tutorial. *Molecular Ecology Resources* 13, 559–572.
- Woo, S., Cha, S.W., Merrihew, G., *et al.*, 2013. Proteogenomic database construction driven from large scale RNA-seq data. *Journal of Proteome Research* 13, 21–28.
- Zhang, X.O., Wang, H.B., Zhang, Y., *et al.*, 2014. Complementary Sequence-Mediated Exon Circularization. *Cell* 159, 134–147.

The Advent of Mass Spectrometry-Based Proteomics in Systems Biology Research

P Blattmann, ETH Zurich, Zurich, Switzerland

R Aebersold, ETH Zurich, Zurich, Switzerland and University of Zurich, Zurich, Switzerland

© 2016 Elsevier Inc. All rights reserved.

Systems Biology and Proteomics

Systems biology aims to comprehensively describe complex biological processes as a network with nodes and edges (Ideker *et al.*, 2001). Nodes of such a network generally represent molecules such as proteins or genes, the edges generally represent their physical or functional interactions. The rearrangement of such a network following genetic or external (i.e., environmental) perturbations, explains how these perturbations result in a phenotype. According to this theory, by knowing the nodes, edges, and mechanisms underlying this network, it can be predicted how a phenotype results from the state of the network (Ideker *et al.*, 2001). This network-based approach is the rationale for systems biology efforts to study complex biological processes and complex diseases (Barabasi *et al.*, 2011). The approach has become possible due to the development of the so-called ‘-omics’ techniques and by the application of *in silico* modeling to biology. Among the ‘-omics’ technologies genomics has led the way by efficiently analyzing whole genomes and transcriptomes. Such analyses resulted in a defined list of genes (and thus proteins) that can be expressed from the respective genome, or in the network terminology the nodes of the network. Once the nodes are known, the next challenge is to understand how these nodes are connected by specific interactions, or edges, to produce the observed phenotypes. Especially in this next step, we believe that mass spectrometry-based proteomics will play a major role.

Genomics and transcriptomics are the archetype of an ‘-omics’ technique in terms of throughput, measurable nodes and accuracy. This is what proteomics aims to achieve for the proteome (Cox and Mann, 2007). Genomics has had an immense impact on biological research (Lander, 2011). However, the effectors of most biological processes are proteins, neither genes nor mRNA. Moreover, posttranslational modifications (PTMs), such as phosphorylation, are not accessible to genomic or transcriptomic techniques. Hence, when studying molecular networks using a systems biology approach, ideally the proteins and their quantities should be directly measured rather than inferred. In contrast to genomics that analyzes nucleic acids by sequencing, the analysis of the proteome faces several challenges that have hampered progress toward the ability of measuring all proteins simultaneously using a standardized workflow. First, proteins or peptides vary significantly in their size and physicochemical characteristics. For example, some proteins can consist of up to several thousand amino acids while others are very small. Membrane proteins are very hydrophobic, conversely very soluble or even charged proteins exist. Second, in contrast to four different nucleic bases, proteins are built of 20 different amino acids that each vary greatly in their charge, polarity, and size. Furthermore, many different PTMs of proteins further increase the

complexity. Third, the large difference in abundance of proteins makes analyzing the proteome very challenging. The measured protein abundances in cells range across 7 orders of magnitude (Beck *et al.*, 2011) and > 8 orders of magnitude in plasma (Shen *et al.*, 2005). The actual difference in abundance of proteins, especially in the plasma, is believed to exceed 10 orders of magnitude (Anderson and Anderson, 2002). Despite these outlined challenges, proteomics has experienced an immense advance in the last decade in terms of scope of the analyses and the accuracy, reliability and reproducibility of the results, as discussed below. Hence, proteomics has recently become a prime technology for systems biology studies.

To measure several hundreds or thousands of proteins in parallel and for a large number of samples mass spectrometry-based proteomics as we discuss it here is the current method of choice among the proteomics techniques. Mass spectrometers can also be coupled to a flow-cytometer that allows measuring at single cell resolution up to hundred different metal-labeled antibodies binding to individual cells and holds great promise for future systems biology studies (Bandura *et al.*, 2009; Bodenmiller *et al.*, 2012). Other analytical techniques such as ELISA, Western blotting, immunohistochemistry, or protein microarrays can also be used to measure protein abundances (Stoevesandt and Taussig, 2012). However, all of these techniques rely on antibodies or other affinity reagents that specifically recognize a certain epitope on the proteins. Although these affinity-based techniques show sometimes greater sensitivity and throughput capability than mass spectrometry-based proteomics, it is more challenging to scale up these methods to measure many hundreds or several thousand proteins in a highly multiplexed manner. Therefore, mass spectrometry-based proteomics has emerged as the best suitable technique to measure a certain system as comprehensively as possible. Hence, we will focus on liquid chromatography-coupled tandem mass spectrometry-based proteomics (LC-MS/MS) and show how this technique can provide data for systems biology studies.

Liquid Chromatography-Coupled Tandem Mass Spectrometry

Digestion of Proteins into Peptides

To identify and quantify a large number of proteins using a generic mass spectrometric method, proteins are digested into similarly sized peptides that are then measured. This approach is usually referred to as, ‘bottom-up’ proteomics, as opposed to analyzing the proteome ‘top-down,’ which analyzes the intact proteins using mass-spectrometry (reviewed in Catherman *et al.*, 2014). Because highly complex samples, such as a total cell lysate, cannot currently be measured efficiently using top

down proteomics, this approach will not be covered here. The digestion of proteins to peptides has practical advantages. Peptides are more similar in their masses and are therefore much easier to measure than different proteins by a liquid-chromatography (LC) coupled mass spectrometer. Moreover, a mixture of peptides is easier to handle and degrades less readily than proteins. Digestion into peptides is most commonly performed in a one-step digestion using trypsin, or a two-step protocol with trypsin and LysC (Glatter *et al.*, 2012; Wisniewski and Mann, 2012). The digestion can be performed in-gel (e.g., after SDS-PAGE gel electrophoresis) or in solution (Shevchenko *et al.*, 1996). Importantly, the digestion of the proteome into peptides comes at the cost of vastly increasing the complexity of the sample and of losing the connection between the peptide and the protein that is inferred from the identified peptides. The data generated on the peptide level need to be mapped onto the protein level to make statements about the protein contents of a sample, a process that has been termed 'protein inference' (Kuster *et al.*, 2005; Nesvizhskii and Aebersold, 2005). That this data analysis process is not trivial and can lead to differential results among labs has been shown in a cross-laboratory study (Bell *et al.*, 2009). The bottleneck in correct identification did not lie in the detection or fragmentation of peptides, but rather in the correct annotation and searching of the data. However, the protein search tools performing these annotations are constantly improving, making the annotation increasingly robust and automated (Cox and Mann, 2008; Eng *et al.*, 1994; Röst *et al.*, 2014). Importantly, robust strategies have been developed to control for the false-discovery rate using the target-decoy search strategy and by modeling the score distribution of false annotations (Elias and Gygi, 2007; Keller *et al.*, 2002). Besides the development of new mass spectrometers, the advancement of analysis tools has been a major contributor to the advent of mass spectrometry-based proteomics in the last decade.

Reducing the Complexity by Separation

The complexity of a sample to be measured can be reduced by pre-fractionation of the sample into different fractions that are analyzed independently. The measured data from these fractionated samples are then merged to obtain one dataset corresponding to the initial sample. Fractionation reduces greatly the complexity within the fractionated sample and thus increases the sensitivity and the number of analytes detected from the initial sample. The drawback of this procedure is the increased preparation and measurement time, and the risk of losing some of the peptides/proteins during the fractionation process. Different separation techniques have been used in bottom-up proteomics for proteins or peptides. One of the first pre-fractionation techniques separates proteins by mass or isoelectric point using gel electrophoresis. 2D-gels can separate proteins by size and isoelectric point. Separating peptides by their isoelectric point can also be performed in the so-called off-gel fractionation that allows omitting the elution of the peptides from the gel. In addition to the electrophoresis-based separation techniques chromatography is widely used to separate proteins or peptides according to their physico-chemical characteristics. For example cation-exchange (SCX),

size-exclusion (SEC) or reverse phase chromatography is used to separate the peptides or proteins according to their size, charge, or polarity, respectively. Importantly, the complexity can also be reduced by enriching for specific proteins/peptides that harbor specific characteristics. This is especially common when studying posttranslational modifications such as phosphorylation or ubiquitination (Bodenmiller *et al.*, 2007; Xu *et al.*, 2010). Also, the proteome of certain organelles or complexes can be enriched by biochemical techniques (Andersen *et al.*, 2005). Moreover, using tandem-affinity tags coupled to a protein of interest, it is possible to efficiently enrich specifically for protein interactors of that protein (Gingras *et al.*, 2007; Glatter *et al.*, 2009).

Pre-fractionation has been intensively used to detect or map proteomes of very complex samples, such as total cell lysate of human cells. However, the increased measurement time required by the analysis of fractionated samples is especially relevant when considering measuring many samples across different conditions in a systems biology study. Therefore, measuring many different samples comes usually at the cost of reducing the pre-fractionation steps and thus the limit of detection and the number of identified peptides. In addition to these pre-fractionations, peptides are usually separated directly before injection into the mass spectrometer. This is performed using a nano-flow liquid chromatography (LC) device linked to a direct electrospray introduction into the mass spectrometer. The most popular method is to separate the peptides during the LC by using reversed-phase chromatography. This on-line separation can also be increased to two dimensions as is performed in the so-called MudPIT approach (Link *et al.*, 1999).

Measurement of the Peptides by Mass Spectrometry

For peptide identification, tandem mass spectrometers generally perform the following sequence of steps: (1) they determine the mass to charge ratio (m/z) of a precursor ion in an MS1 spectrum, (2) they select specific precursor ions and subsequently fragment them in a collision cell, and (3) they measure the m/z of the product ions arising from this collision and record them in the form of a fragment ion or an MS2 spectrum. Various instruments suited for the different mass spectrometry approaches are available that differ with respect to their sensitivity, mass accuracy or working modes (Domon and Aebersold, 2010). The acquired data are then analyzed using software tools to obtain information indicating which peptides have been present in the sample and to some extent to infer their relative or absolute quantities. Different approaches to perform mass spectrometry-based proteomics and quantification are briefly introduced below (Figure 1).

Data-dependent acquisition

Shotgun or data-dependent acquisition (DDA) is currently the most widespread method in mass spectrometry-based proteomics. At a given time during the elution of the peptides from the LC-column, an MS1 spectrum is acquired by scanning for all precursor ions. From this MS1 spectrum, a subset of the detected precursors, usually those with the highest signal intensity, are selected, fragmented, and an MS2 spectrum is

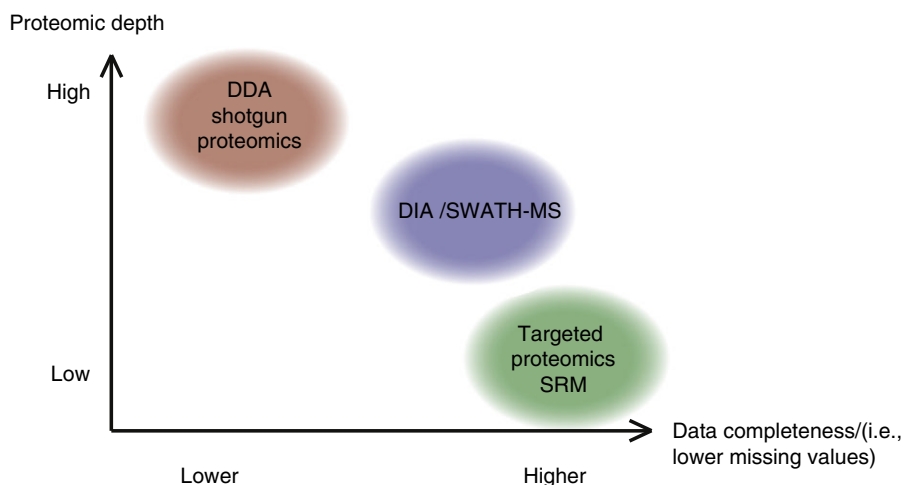


Figure 1 Proteomic depth against data completeness for different mass spectrometry supported approaches. The proteomic depth (i.e., number of proteins measured) in a single injection versus data completeness (i.e., lower missing values) across many different complex samples for different mass spectrometry approaches is shown. These features have been estimated for a systems biology experiment quantifying the proteome of a total cell lysate across dozens of different samples. DDA, data-dependent acquisition; DIA, data-independent acquisition; SRM, selected-reaction monitoring.

acquired for each precursor. This process is repeated in many cycles during the elution of all of the peptides by the online liquid chromatography. The acquired MS2 spectra of individual precursor ions are compared to databases of *in silico* generated spectra (Cox and Mann, 2008; Eng *et al.*, 1994; Keller and Shteynberg, 2011). This allows the association of these spectra with their corresponding amino acid sequences. Today, state-of-the-art tandem mass spectrometers can identify over 30 000 peptides and about 4000 proteins per injection (Hebert *et al.*, 2014; Nagaraj *et al.*, 2012). Despite this impressive performance it has to be recognized that, in the case of complex proteomes, a much higher number of peptides is present in the samples than can be sequenced (Michalski *et al.*, 2011a). Because not all peptides can be measured, a stochastic step of selecting the precursor ions to be fragmented is performed and can result in missing values in between runs, if the same precursor ion is not measured. This problem can be partially alleviated by using inclusion lists (Schmidt *et al.*, 2008). The DDA approach is currently the most widely used method and especially suitable for discovery-driven studies such as mapping the proteome of a complex sample. As DDA is historically the oldest method, both instruments and analysis tools are very mature.

Targeted data acquisition or selected reaction monitoring

The most sensitive approach for measuring a proteome is the targeted data acquisition approach. The most prominent targeted acquisition approach is selected reaction monitoring (SRM), or synonymously multiple reaction monitoring (MRM). SRM is performed by measuring specific transitions by selecting a specific, predetermined precursor ion and a number of corresponding product ions. In contrast to DDA, the mass spectrometer is directed to selectively record these precursor-product ion transitions for a certain peptide. This greatly increases the specificity and sensitivity allowing SRM to detect peptides across 3–5 orders of magnitude (Anderson and

Hunter, 2006; Ebhardt *et al.*, 2012; Picotti *et al.*, 2009). In contrast to DDA, the number of proteins measured in one injection is much lower. Using optimized conditions, about 100 proteins can be measured in one multiplexed analysis run (Picotti and Aebersold, 2012). While in DDA a sample can be acquired without any prior knowledge, for SRM it has to be clearly defined which transitions ought to be measured. Therefore targeted proteomics is a conceptually very different way of performing proteomics because a list of transitions that will be measured has to be defined beforehand (Picotti *et al.*, 2010). Notably, a recently developed approach termed parallel reaction monitoring (PRM) omits the necessity of defining the product ions. In PRM, a certain precursor is selected as in SRM, but then a full MS2 spectrum is acquired (Gallien *et al.*, 2013). Hence, only a list of precursor masses to be selected has to be provided. For either method, once the signal of selected transitions has been acquired for a particular peptide, the data analysis is rather straight-forward and aided by dedicated software (MacLean *et al.*, 2010; Reiter *et al.*, 2011). As the mass spectrometer will measure at the right mass and elution time, measurements with no missing values are the norm.

Cross-laboratory validation of SRM showed very good reproducibility and ability of this technique to detect even μg per ml amounts of protein in plasma (Addona *et al.*, 2009). The targeted data acquisition approach is especially suitable if a set of several hundred proteins across many different samples need to be accurately quantified.

Data-independent acquisition

In the data-independent acquisition (DIA) approach, the proteome is mapped in a systematic way selecting reproducibly a range of precursors by defining a selection window in MS1. In contrast to the before mentioned approaches, where precursors are selected by a stochastic process (DDA) or by predefined lists (SRM and PRM), the DIA approach selects precursor ions in a systematic and comprehensive manner

(Venable *et al.*, 2004). There exist different DIA approaches selecting precursor ions in windows in MS1 ranging from one large window of 800 Da in MS^E (Plumb *et al.*, 2006), or smaller cycling windows of 10 Da (Venable *et al.*, 2004), or 25 Da in SWATH-MS (Gillet *et al.*, 2012), or as small as 2.5 Da in PAcFIC (Panchaud *et al.*, 2009). In DIA, all of the selected precursor ions are fragmented simultaneously and a whole MS/MS spectrum is recorded. This procedure generates a reproducible and comprehensive proteomic fingerprint of a sample, omitting any missing values on the side of the data acquisition. However, while selecting many precursors at the same time, the information is lost to which precursor the measured product ions are belonging to. Therefore, the analysis of DIA data has been especially challenging. Employing a targeted data analysis approach, it has recently become possible to annotate reliably DIA data also in very complex samples (Gillet *et al.*, 2012). This targeted analysis approach interrogates the acquired spectrum with spectral libraries, instead of annotating the spectrum using databases as in DDA (Gillet *et al.*, 2012). Spectral libraries consist of previously recorded signals of transitions, typically acquired and annotated to peptides and proteins by DDA. The SWATH-MS approach has been shown to be more sensitive than DDA in a complex sample, but a little less sensitive than SRM (Gillet *et al.*, 2012). Out of the different DIA approaches we will henceforth focus on SWATH-MS in this article as it shows the most interference free signal (Gillet *et al.*, 2012). As SWATH-MS is a recently developed technique, both the mass spectrometers and analysis tools have just been developed (Röst *et al.*, 2014). First studies mapping protein–protein interactions or identifying biomarkers with SWATH-MS show that the data completeness, i.e., the minimal number of missing values across datasets, of the measurement is very beneficial (Collins *et al.*, 2013; Lambert *et al.*, 2013; Liu *et al.*, 2013).

Quantification of Peptide Abundance

In addition to identifying the peptides present in a sample, the abundance of a peptide or protein across different samples needs to be compared. Mass spectrometry is a quantitative technique and the ion count of detected ions serves as a measure of the abundance of the respective peptides. Hence, the change in abundance of peptides between different samples can be quantified by relating their elicited signal. In order to obtain an accurate quantification, four different approaches are explained here (for a more comprehensive review see Bantscheff *et al.* (2012)). Most take advantage of the existence of stable isotopes. Isotopically different peptides differ in their mass because they contain a different number of neutrons and therefore are distinguishable by mass spectrometry. For example, a ‘heavy’ ¹³C₆ ¹⁵N₂ L-lysine shows an 8 Da mass shift from a ‘light’ ¹²C₆ ¹⁴N₂ L-lysine. However, these ‘heavy’ atoms, and by implication the peptides that contain them, have the same chemical properties as the corresponding ‘light’ peptides.

First, in label-free quantification the peptide intensity can be compared by counting the number of spectra generated by DDA (spectral counting). Alternatively, the area of apex of the signal peak can be compared by recording the corresponding

signal in either MS1 or MS2 while the peptide is eluting from the LC-column.

A second approach consists of spiking heavy peptides into the sample and measuring the ratio of the signal between light and heavy peptides. By this way differences in chromatography or peptide digestion between different samples are cancelled out and it is often used in SRM studies. For DDA, heavy reference peptides have been spiked into the samples in the form of a metabolically labeled proteome and termed super-SILAC (Geiger *et al.*, 2010).

In the third approach peptides are labeled chemically. The tags will add a certain mass to each peptide. This can be performed by using ICAT (isotope-coded affinity tags) in their light or heavy form (Gygi *et al.*, 1999), or with a range of other isotopic reagents (Bantscheff *et al.*, 2012). A variant of the isotope-based quantification approach is isobaric tagging exemplified by iTRAQ (isobaric Tags for Relative and Absolute Quantitation) or TMT (Tandem Mass Tags). For these isobaric tags, the precursor ions show the same mass in MS1, but in the MS2 spectrum the contribution of the different tags can be quantified (Ross *et al.*, 2004; Thompson *et al.*, 2003).

The fourth approach is based on metabolic labeling. Here, cells are grown in medium containing light or heavy nutrients, such as essential amino acids, that are incorporated into the proteins. The most common approach of metabolic labeling for proteins is termed SILAC (Stable isotope labeling by amino acids in cell culture) (Ong *et al.*, 2002).

In summary, there are many different ways of quantifying the signal acquired from a peptide (Bantscheff *et al.*, 2012). Accuracy and precision vary between these different approaches and need to be taken into account before performing or analyzing such experiments. Label-free quantification is for example directly affected by variations in the chromatography or instrument performance that affect the signal intensity. These variations can be efficiently cancelled out by using a labeling approach and thus quantifying and comparing peptides eluting at the same time. In addition to quantifying relative abundances, absolute quantification can be attempted. For absolute quantification typically the so-called AQUA peptides are spiked in at an exactly determined concentration (Gerber *et al.*, 2003). By spiking in a small set of reference peptides and by relating the relative abundance of the peptides using label-free quantification, accurate measurement for the remaining proteome can be obtained with reasonable costs and effort (Chang *et al.*, 2014; Ludwig *et al.*, 2012).

Systems Biology Application for High-Throughput Proteomics

Above, the major different implementations of mass spectrometry-based proteomics have been briefly outlined. In the remainder of this article, we summarize exemplary cases that employ these methods in systems biology studies. In the first part, the application of these techniques to conceptually different ways to study the proteome (i.e., mapping the nodes) is covered, showing the current state of the art. In the second part, three exemplary complex biological processes (i.e., cell cycle regulation) are highlighted to show the impact of

proteomics on systems biology studies in these selected biological processes.

Different Conceptual Approaches in Mass Spectrometry Based Proteomics

Mapping the nodes of a proteome

The major aim in the early phase of proteomics was to map the proteome of different species, organelles, or subproteomes (such as the phosphoproteome) in order to define the 'parts list' or, in systems biology terms, the network nodes of different organisms. In such studies for mapping a proteome, DDA has typically been used in combination with extensive pre-fractionation of the samples.

The genome-predicted proteome can now be mapped completely or to a very large extent with mass spectrometry-based proteomics. This has been, for example, shown for prokaryotes (Ahrens *et al.*, 2010; Schubert *et al.*, 2013) and *Saccharomyces cerevisiae*, where over 4000 proteins could be identified corresponding to a substantially complete proteome (de Godoy *et al.*, 2008; Hebert *et al.*, 2014). In human cell lines, about 10'000 proteins could be identified in different large-scale studies (Beck *et al.*, 2011; Geiger *et al.*, 2012; Moghaddas Gholami *et al.*, 2013; Munoz *et al.*, 2011; Nagaraj *et al.*, 2011), and finally two recent studies reported a first draft of the human proteome detected proteins for up to 92% of the annotated genes by analyzing many different tissues (Kim *et al.*, 2014; Wilhelm *et al.*, 2014). For targeted proteomics, the relevant information is for how many proteins SRM-assays have been developed. The PeptideAtlas is a community-based approach collecting this information and making it publicly available to support SRM studies (Deutsch *et al.*, 2008). SRM-assays have been deposited in the PeptideAtlas for >97% of the genome-predicted proteome of *S. cerevisiae* and *M. tuberculosis* (Picotti *et al.*, 2013; Schubert *et al.*, 2013) and SRM assays for the entire human proteome have been developed (unpublished). Hence, both with DDA and SRM the large majority of the genome-predicted proteome can be measured.

In addition to the genome-predicted proteome, there are many subproteomes consisting of post-translation modifications (PTMs) such as, for example, the phosphoproteome. The phosphoproteome is of considerable interest because phosphorylation events are key PTMs influencing the activity of many enzymes and proteins. The number of PTM identifications has so far increased without reaching saturation (Olsen and Mann, 2013). Here, a few examples of large studies mapping PTMs are listed: Phosphoproteomic studies have detected 54 755 phosphopeptides and 9287 phosphoproteins by analyzing different tissues in mouse and rat (Huttlin *et al.*, 2010; Lundby *et al.*, 2012b). 15 474 lysine acetylation events on 4541 proteins have been detected in 16 different rat tissues (Lundby *et al.*, 2012a). 6397 glycosites on 2352 proteins were identified in different mouse tissue and plasma (Zielinska *et al.*, 2010). Assessing the ubiquitinated proteome, two studies have detected 4000–5000 ubiquitinated proteins with 3–4 times the number of ubiquitylation sites (Kim *et al.*, 2011; Wagner *et al.*, 2011). Hence for many subproteomes, extremely large numbers of modified proteins could be detected. Because mass spectrometry based proteomics is often the only

technique able to measure these modifications, many of these modifications have been reported for the first time and their functional role needs still to be investigated. Hence, mapping these subproteomes is still largely in progress.

Organization of the proteome

Once a list of nodes has been defined, the next question, when performing a systems biology study, is how these nodes interact. This question can also be phrased as how the proteome is organized or structured in a network, and what type of interaction the edges represent. An important and biologically meaningful type of interaction is physical protein–protein interactions. Using tandem affinity purification coupled with mass spectrometry, physical protein–protein interactions can be efficiently mapped (Gingras *et al.*, 2007). This has been typically performed by DDA, but recent studies have successfully employed SWATH-MS (Collins *et al.*, 2013; Lambert *et al.*, 2013).

For example, two studies identified around 500 protein complexes that encompass 7 123 protein–protein interactions in yeast by analyzing complexes where the bait protein was affinity tagged by knock-in technology (Gavin *et al.*, 2006; Krogan *et al.*, 2006). As genetic manipulation is not so simple in mammalian cells different techniques have been used for large-scale mapping of protein–protein interactions. In mammalian cells, extensive biochemical separation identified 622 complexes consisting of 3006 proteins and 13 992 physical interactions (Havugimana *et al.*, 2012). Performing 3290 immunoprecipitation with 1796 different antibodies, over 11 000 proteins were identified and 430 000 interactions observed (Malovannaya *et al.*, 2011). Still it is likely that the complete interactome has not been mapped. It has been estimated that the yeast proteome contains in total about 18 000 binary interactions (Yu *et al.*, 2008). The human proteome is predicted to have about 130 000–650 000 binary interactions (Stumpf *et al.*, 2008; Venkatesan *et al.*, 2009). These estimates show the scope of mapping all the interactions. Notably, to map the binary interactions, but not complexes, alternative high throughput approaches exist. Yeast two-hybrid (Y2H) or protein complementation assay (PCA)-based methods have been shown to probe more efficiently for binary interactions, whereas mass spectrometry-based methods are superior for identifying complexes (Yu *et al.*, 2008). Therefore, to obtain a comprehensive picture of all interactions, the data obtained from different approaches need to be combined.

In contrast to the unbiased large-scale mapping approaches, affinity purification mass spectrometry (AP-MS) studies on specific set of proteins have been performed to focus on a particular biological process as shown here on some examples. Using AP-MS in yeast, interaction studies with several hundred baits can be performed, such as the interactome of 276 kinases and phosphatases that resulted in 1844 interactors (Breitkreutz *et al.*, 2010). In humans, interactome studies have so far been limited to smaller number of baits largely due to technical limitations in genetic engineering and culturing of mammalian cells. Nevertheless protein–protein interactions have been mapped for 11 proteins in the PP2A system (Glatter *et al.*, 2009), 74 deubiquitination enzymes (Sowa *et al.*, 2009), 32 proteins involved in autophagy (Behrends *et al.*, 2010), interactors of the 14-3-3 β scaffold

protein (Collins *et al.*, 2013), the human Hippo signaling pathway (Hauri *et al.*, 2013), or 57 human CMGC kinases (Varjosalo *et al.*, 2013). Usually these studies reported approximately 10 times the number of interacting proteins compared to the number of bait proteins tested. An exception to that rule is the scaffold protein 14-3-3 β that interacted with 567 proteins across different time points. These studies show that AP-MS can be efficiently used to study a protein–protein interaction network of choice.

Correlation of the proteome with the genome

A phenotype or disease is the result of the genetic predisposition in combination with environmental influences. Genetics has for decades successfully identified genes causative for, or associated with, disease or phenotypes. Recently, genetic studies have identified many variants associated to phenotypes for complex diseases in sequencing or genome-wide association studies (GWAS) (Lander, 2011). These genomic regions, in order to eventually elicit a phenotype, are expected to affect the proteomic network. This perturbed proteomic network then results in an observable phenotype. To understand thoroughly how the genotype-to-phenotype transformation is happening, we first have to understand how changes in the genotype affect the proteome. Therefore, correlation of protein levels to genomic loci has been performed to unravel these associations. By employing protein quantitative trait loci (pQTL) analysis, genomic loci that correlate with the protein levels of a certain protein are identified. These loci can be separated in cis-QTL loci in the vicinity of the encoded protein, or alternatively trans-QTL loci, located apart from the encoded gene. For such studies the accurate quantification across many different samples is needed. Therefore, these studies have been performed by DDA or SRM combined with accurate quantification approaches.

In yeast, the protein abundance of 48 proteins in 78 *S. cerevisiae* strains were measured and for the majority of proteins a protein QTL were detected (Picotti *et al.*, 2013). In humans, about 4000 proteins were quantified from 95 patients from the HapMap project, and 77 cis-pQTL loci identified (Wu *et al.*, 2013). 46 cis-QTL and 874 trans-QTL loci were identified for 396 proteins in murine liver from 97 different strains (Ghazalpour *et al.*, 2011). For 43% of these proteins, multiple pQTL loci were found. A recent study suggests that due to the limited sample size, these studies are probably greatly underpowered and many more QTL loci are likely to be discovered using single cell measurements (Albert *et al.*, 2014). Hence, this shows that many correlations between the genome and proteome likely exist and we are only beginning to unravel these relationships.

Correlation of the proteome with the phenotype

To eventually understand the genotype-to-phenotype transformation, the correlation between the proteome state and an associated phenotype also needs to be understood. Many proteomics studies correlate reported proteins or protein modifications to a specific phenotype. These correlations are often based on a strong mechanistic rationale how these proteins or modifications could affect the phenotype. Using state of the art statistical and modeling approaches, this can result in a mechanistic network with predictive ability (as an

example (Miraldi *et al.*, 2013)). Nevertheless, it has been very difficult to find for many diseases biomarkers that predict accurately and reliably the outcome or disease presence. In these so-called biomarker studies, proteins, protein modifications, or signatures that are predictive for a certain phenotype or disease are searched. Biomarker studies have been mainly performed in human body fluids (blood or urine), feces, or biopsies. One challenge in identifying biomarkers in these samples is the large abundance differences in human plasma of about 10 orders of magnitude (Anderson and Anderson, 2002) and the high degree of sample variability expected in plasma samples from an outbred population. Highly abundant proteins mask the potentially most interesting proteins. Therefore, most protein biomarkers so far were not identified by mass spectrometry based proteomics (Surinova *et al.*, 2011). In general it has proven very challenging to find new validated cancer biomarkers with any method (Diamandis, 2012). Biomarker discovery can be separated in a discovery and validation phase. In the discovery phase regulated proteins are identified in biopsies, tissue or model organisms, that are then validated by measuring the protein levels in plasma or other easily accessible fluids (Surinova *et al.*, 2011). Among the different mass spectrometry based techniques SRM and DIA are probably best suited for validating biomarkers. SRM shows the highest sensitivity, DIA only shows slightly less sensitivity but has the advantage that the acquired proteomic map can be re-interrogated afterwards. This is especially important for clinical samples where the sample amount is very limited. For SRM and DIA, assay libraries of plasma proteins have been generated to facilitate this approach in the future (Huttenhain *et al.*, 2013; Liu *et al.*, 2013). Future studies need to show if mass spectrometry based proteomics will be able to efficiently identify and validate novel clinical biomarkers.

Measuring dynamic proteomes

The above-mentioned approaches often result in the description of a static proteome. However, both the proteome and edges in biological networks are highly dynamic. Therefore, such data cannot explain the dynamics of a network, nor can they explain dynamic biochemical mechanisms. Therefore, time-dependent measurements of the proteome upon perturbation are needed and, if possible, should be linked to a phenotype (Bensimon *et al.*, 2012). Measuring time-dependent responses to different perturbation requires reliable quantification and a throughput at the largest possible scale. Only if samples can be measured in one or very few injections per sample the analysis of a sizeable number of conditions (e.g., at a scale of hundred conditions) becomes feasible. Therefore, it is important how many proteins are detectable and quantifiable in a single injection. A single injection of a yeast lysate can identify about 4000 proteins which correspond to a nearly complete proteome (Hebert *et al.*, 2014; Nagaraj *et al.*, 2012). Single injection in DDA or with SWATH-MS can detect over 2000 proteins in complex human samples (Michalski *et al.*, 2011b). Hence, also in very complex samples, several thousand proteins can now be reproducibly detected and quantified at a large throughput. Depending on the objectives of a particular project the appropriate acquisition approach needs to be chosen. Whereas with DDA typically a larger number of proteins can be detected, the data

completeness, i.e., lower missing values, over many samples is lower in DIA (Figure 1). In targeted proteomics, the sensitivity is higher but the number of proteins is limited to a few hundred proteins. Nevertheless, the number of measured proteins measured at high throughput can be far larger than any other proteomic or cell-biological technique. Examples of time-dependent proteomic studies have been already mentioned above (Collins *et al.*, 2013) or will be described below.

Exemplary Biological Processes Studied with Mass Spectrometry Based Systems Biology Studies

Three different biological processes have been chosen to briefly highlight how mass spectrometry based studies could contribute to systems biology studies in these respective areas. Cell cycle regulation was selected as a biological process where a large part of the proteome is expected to undergo a change, epidermal growth factor signaling (EGF) as an example for a clinically relevant signaling pathway, and central metabolism as an example how proteomic data can be combined with data from other ‘-omics’ approaches, such as metabolomics.

Cell cycle regulation

Every growing cell is undergoing cell division, a very complex process that requires complex regulatory mechanisms. Cell cycle regulation is clinically very relevant, as uncontrolled proliferation is a hallmark of cancer. To study the cell cycle in mammalian cells, there have been several comprehensive proteomic studies on protein abundance (Lane *et al.*, 2013; Ly *et al.*, 2014; Olsen *et al.*, 2010), sub-proteomes such as proteins binding to the chromosomes (Ohta *et al.*, 2010), protein–protein interactions with three different cyclins (Pagliuca *et al.*, 2011), or the change in the phosphoproteome (Dephoure *et al.*, 2008; Ly *et al.*, 2014; Olsen *et al.*, 2010). The amount of proteins that change their abundance in a cell-cycle dependent manner more than twofold, has been reported to be up to 40% in the different studies corresponding to 2857 proteins. Even more pronounced is the amount of regulated phosphosites that reaches 50% of the recorded sites, corresponding to approximately 20 000 regulated sites. This extensive degree of regulation shows that profound changes are happening while cells are dividing. Despite having a good understanding of the cell cycle, we currently cannot annotate the exact role and implication of all these regulated proteins or protein modifications. This example shows how unbiased studies such as mass spectrometry-based proteomics can identify putative new regulators that could be further analyzed, or help in better understanding the effect of this biological process on cells.

EGFR signaling pathway

Signaling by growth factors such as Epidermal growth factor (EGF) is a crucial signaling pathway for cancer. Because of its clinical relevance, a lot of studies have been performed to understand this signaling pathway in detail. Extensive cross-talk and the complexity of these pathways however make it a very challenging task.

In 2004, temporal analysis of phosphorylation on up to 202 phosphoproteins upon EGF stimulation could be

measured (Blagoev *et al.*, 2004; Zhang *et al.*, 2005). The difference between stimulation by EGF and the closely related platelet-derived growth factor (PDGF) pathway was tested (Kratchmarova *et al.*, 2005). A later study detected 6600 phosphorylation sites on 2244 phosphoproteins, where 14% of the phosphorylation sites were shown to be regulated upon EGFR signaling (Olsen *et al.*, 2006). To test the influence of other proteins on EGFR signaling pathway, 332 phosphopeptides have been monitored upon overexpression of HER2, a receptor interacting with EGFR (Wolf-Yadlin *et al.*, 2006). It was shown that proteins with a role in migration are increasingly phosphorylated. This correlated well with the increased migratory behavior of these HER2 overexpressing cells. EGFR signaling could also be assessed in clinical samples of different patient-derived cancer cell lines (Rikova *et al.*, 2007), or recently using Glioma xenografts (Johnson *et al.*, 2012). In these clinical samples it was possible to detect differentially expressed proteins and phosphorylation events that could identify signatures for poor survival.

Also protein interaction networks in this system have been analyzed. As an example, 90 interactors for Grb2, a downstream target in EGFR signaling, were monitored using SRM following affinity purification in a time course after EGF stimulation (Bisson *et al.*, 2011). Out of these 90 proteins, 21 proteins were found to be interacting differentially depending on six different growth factors and showing time-dependent complex formation upon signaling (Bisson *et al.*, 2011). A later study determined the specific interactome for lung-cancer cells containing a mutated EGF-Receptor (Li *et al.*, 2013). This was combined with functional RNAi-screening and could propose different drugs to overcome the resistance caused by these mutations.

The example with EGFR signaling shows how in less than a decade, we have greatly advanced in our ability to characterize in increasing detail a signaling pathway such as EGFR signaling using mass spectrometry based proteomics, even in clinical samples.

Central metabolism

Metabolism is arguably the biological process in systems biology where the understanding is most advanced. Decades of work on metabolic pathways and the very extensive understanding of the underlying chemistry have made metabolic pathways to a large extent known, as well as their enzymes (Feist *et al.*, 2009; Thiele *et al.*, 2013). Metabolism is clinically relevant for many metabolic diseases such as obesity and diabetes but also implicated in cancer (Currie *et al.*, 2013). Another advantage of studying this biological system with proteomics is that metabolomics is able to measure many metabolites and determine fluxes of the different pathways that can be correlated to the proteomic data. This has been performed, for example, in a study where 199 proteins acting in yeast central metabolism were monitored by SRM on abundance level across different growth conditions (Costenoble *et al.*, 2011). A large fraction of the selected proteins (144 out of 199) changed considerably depending on different growth conditions. Interestingly, this study showed that proteins are still expressed even when they are not needed under the cellular state tested. Only about half of the proteins changed their expression more than twofold if the cell was shifted to a state where the specific proteins were unnecessary,

showing an unexpected stability of protein levels. This study was then extended to measuring phosphorylation status of 35 of these enzymes in these conditions (Oliveira *et al.*, 2012).

In another example, 144 proteins in two different mouse strains have been analyzed in murine liver under 12 different conditions in response to high-fat diet (Sabido *et al.*, 2013). It could be determined how >30 proteins responded differently in the two different strains to the same change in diet after 6 weeks. The study shows how the effect of perturbations can vary in an organism depending on its genotypic background. Hence, both the genotypic and environmental effect on a phenotype can be assessed and ideally be delineated. This is eventually what systems biology aims for (Ideker *et al.*, 2001). Therefore, these studies show that using mass spectrometry-based proteomics, we are advancing toward this idea of systems biology at great pace.

Outlook

As shown above, and highlighted with some biological processes as examples, we think that the impact of mass spectrometry-based proteomics on systems biology is likely to be profound. Most cell-biological processes have been discovered before the advent of mass spectrometry-based proteomics, but the current unique capabilities of proteomics, as briefly outlined in this article, promise to have a great impact on understanding biological processes in a more comprehensive way within the context of the living cell. Similarly as for genomics, we expect proteomics to change the field of biology and in particular systems biology greatly.

As a result of a decade of rapid development in the field of mass spectrometry techniques, the majority of the human proteome and complete proteomes for smaller organisms can now be reliably measured. Within the last decade, DDA has immensely increased its capabilities and DIA and targeted approaches have been developed. All of these approaches can measure biological samples at an ever larger scale, with better reproducibility and accuracy. How will the development of mass spectrometry-based proteomics continue? Besides further development of instrumentation, we envisage mass spectrometry-based proteomics to become more widely used in the next years, due to the robust implementation of techniques and advances in software tools. Rather than being mainly run by expert research groups in dedicated mass spectrometry environments, as is the case today, proteomics will likely extend to other labs of other disciplines and many more institutions. As QPCR, Western Blotting, FACS or sequencing is now used by many non technology-development labs, mass spectrometry-based proteomics will also become widespread in a similar manner. This will increase the number and breadth of biological processes analyzed from the few well-studied biological processes that have been used to date in proof-of-concept studies. Key for the successful popularization are robust and open-source pipelines and resources. If successful, this will likely result in an immense boost of proteomic data that then can be used in systems biology approaches such as generating predictive and quantitative model of biological processes.

There are some key considerations that should be considered when starting on a proteomics-based systems biology study. Importantly, a good compromise between proteomic depth, sample throughput and reproducibility has to be found (see also Figure 1). This implies defining the best proteomic acquisition approach and labeling strategy. Furthermore, it is crucial to ensure that the proteins or modifications of interest can reliably be measured. Some classes of proteins (e.g., membrane proteins) or low abundance proteins still require tailored pre-fractionation approaches. When comparing the data across many samples, it is crucial to have a robustly working instrument setup that can acquire data with very few missing values as many missing values will make a subsequent systems analysis and comparison cumbersome.

Measuring protein levels, protein modifications, and interactions at an unprecedented scale and precision will boost the systems biology approach that relies on quantitative and large-scale data. By generating network models based on proteomic data and containing PTMs, these models should become more accurate and predictive and lead to an ever better understanding. Linking the proteomic data to data from metabolomics, genomics and different phenotypes is expected to increase greatly our understanding how the genotype to phenotype transformation takes place. Hence, such network models of biological processes could lead to a thorough understanding of how genetic or environmental perturbations affect complex biological processes and eventually lead to complex diseases.

Acknowledgments

This work was supported by the Swiss National Science Foundation (grant # 3100A0-107679, to R.A.) and an ERC grant (ERC-2008-AdG_20080422). P.B. is supported by a grant from the Swiss SystemsX.ch initiative, evaluated by the Swiss National Science Foundation.

See also: Horizontal Integration: Omics: Phosphoproteomics: Approaches, Developments, and Challenges

References

- Addona, T.A., Abbatiello, S.E., Schilling, B., *et al.*, 2009. Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nature Biotechnology* 27, 633–641.
- Ahrens, C.H., Brunner, E., Qeli, E., Basler, K., Aebersold, R., 2010. Generating and navigating proteome maps using mass spectrometry. *Nature Reviews Molecular Cell Biology* 11, 789–801.
- Albert, F.W., Treusch, S., Shockley, A.H., Bloom, J.S., Kruglyak, L., 2014. Genetics of single-cell protein abundance variation in large yeast populations. *Nature* 506, 494–497.
- Andersen, J.S., Lam, Y.W., Leung, A.K., *et al.*, 2005. Nucleolar proteome dynamics. *Nature* 433, 77–83.
- Anderson, L., Hunter, C.L., 2006. Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Molecular & Cellular Proteomics* 5, 573–588.
- Anderson, N.L., Anderson, N.G., 2002. The human plasma proteome: History, character, and diagnostic prospects. *Molecular & Cellular Proteomics* 1, 845–867.

- Bandura, D.R., Baranov, V.I., Ornatsky, O.I., *et al.*, 2009. Mass cytometry: Technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Analytical Chemistry* 81, 6813–6822.
- Bantscheff, M., Lemeur, S., Savitski, M.M., Kuster, B., 2012. Quantitative mass spectrometry in proteomics: Critical review update from 2007 to the present. *Analytical and Bioanalytical Chemistry* 404, 939–965.
- Barabasi, A.L., Gulbahce, N., Loscalzo, J., 2011. Network medicine: A network-based approach to human disease. *Nature Reviews Genetics* 12, 56–68.
- Beck, M., Schmidt, A., Malmstrom, J., *et al.*, 2011. The quantitative proteome of a human cell line. *Molecular Systems Biology* 7, 549.
- Behrends, C., Sowa, M.E., Gygi, S.P., Harper, J.W., 2010. Network organization of the human autophagy system. *Nature* 466, 68–76.
- Bell, A.W., Deutsch, E.W., Au, C.E., *et al.*, Group HTSW, 2009. A HUPO test sample study reveals common problems in mass spectrometry-based proteomics. *Nature Methods* 6, 423–430.
- Bensimon, A., Heck, A.J., Aebersold, R., 2012. Mass spectrometry-based proteomics and network biology. *Annual Review of Biochemistry* 81, 379–405.
- Bisson, N., James, D.A., Ivosev, G., *et al.*, 2011. Selected reaction monitoring mass spectrometry reveals the dynamics of signaling through the GRB2 adaptor. *Nature Biotechnology* 29, 653–658.
- Blagoev, B., Ong, S.E., Kratchmarova, I., Mann, M., 2004. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. *Nature Biotechnology* 22, 1139–1145.
- Bodenmiller, B., Mueller, L.N., Mueller, M., Domon, B., Aebersold, R., 2007. Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nature Methods* 4, 231–237.
- Bodenmiller, B., Zunder, E.R., Finck, R., *et al.*, 2012. Multiplexed mass cytometry profiling of cellular states perturbed by small-molecule regulators. *Nature Biotechnology* 30, 858–867.
- Breitkreutz, A., Choi, H., Sharom, J.R., *et al.*, 2010. A global protein kinase and phosphatase interaction network in yeast. *Science* 328, 1043–1046.
- Catherman, A.D., Skinner, O.S., Kelleher, N.L., 2014. Top Down proteomics: Facts and perspectives. *Biochemical and Biophysical Research Communications* 445, 683–693.
- Chang, C.Y., Sabido, E., Aebersold, R., Vitek, O., 2014. Targeted protein quantification using sparse reference labeling. *Nature Methods* 11, 301–304.
- Collins, B.C., Gillet, L.C., Rosenberger, G., *et al.*, 2013. Quantifying protein interaction dynamics by SWATH mass spectrometry: Application to the 14-3-3 system. *Nature Methods* 10, 1246–1253.
- Costenoble, R., Picotti, P., Reiter, L., *et al.*, 2011. Comprehensive quantitative analysis of central carbon and amino-acid metabolism in *Saccharomyces cerevisiae* under multiple conditions by targeted proteomics. *Molecular Systems Biology* 7, 464.
- Cox, J., Mann, M., 2007. Is proteomics the new genomics? *Cell* 130, 395–398.
- Cox, J., Mann, M., 2008. MaxQuant enables high peptide identification rates, individualized p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* 26, 1367–1372.
- Currie, E., Schulze, A., Zechner, R., Walther, T.C., Farese Jr., R.V., 2013. Cellular fatty acid metabolism and cancer. *Cell Metabolism* 18, 153–161.
- Dephoure, N., Zhou, C., Villen, J., *et al.*, 2008. A quantitative atlas of mitotic phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* 105, 10762–10767.
- Deutsch, E.W., Lam, H., Aebersold, R., 2008. PeptideAtlas: A resource for target selection for emerging targeted proteomics workflows. *EMBO Reports* 9, 429–434.
- Diamandis, E.P., 2012. The failure of protein cancer biomarkers to reach the clinic: Why, and what can be done to address the problem? *BMC Medicine* 10, 87.
- Domon, B., Aebersold, R., 2010. Options and considerations when selecting a quantitative proteomics strategy. *Nature Biotechnology* 28, 710–721.
- Ehardt, H.A., Sabido, E., Huttenhain, R., Collins, B., Aebersold, R., 2012. Range of protein detection by selected/multiple reaction monitoring mass spectrometry in an unfractionated human cell culture lysate. *Proteomics* 12, 1185–1193.
- Elias, J.E., Gygi, S.P., 2007. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nature Methods* 4, 207–214.
- Eng, J.K., McCormack, A.L., Yates, J.R., 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry* 5, 976–989.
- Feist, A.M., Herrgard, M.J., Thiele, I., Reed, J.L., Palsson, B.O., 2009. Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology* 7, 129–143.
- Gallien, S., Duriez, E., Demeure, K., Domon, B., 2013. Selectivity of LC-MS/MS analysis: Implication for proteomics experiments. *Journal of Proteomics* 81, 148–158.
- Gavin, A.C., Aloy, P., Grandi, P., *et al.*, 2006. Proteome survey reveals modularity of the yeast cell machinery. *Nature* 440, 631–636.
- Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J.R., Mann, M., 2010. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nature Methods* 7, 383–385.
- Geiger, T., Wehner, A., Schaab, C., Cox, J., Mann, M., 2012. Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Molecular & Cellular Proteomics* 11 (M111), 014050.
- Gerber, S.A., Rush, J., Stemman, O., Kirschner, M.W., Gygi, S.P., 2003. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences of the United States of America* 100, 6940–6945.
- Ghazalpour, A., Bennett, B., Petyuk, V.A., *et al.*, 2011. Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genetics* 7, e1001393.
- Gillet, L., Navarro, P., Tate, S., *et al.*, 2012. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: A new concept for consistent and accurate proteome analysis. *Molecular & Cellular Proteomics* 11 (6), 0111.016717.
- Gingras, A.C., Gstaiger, M., Raught, B., Aebersold, R., 2007. Analysis of protein complexes using mass spectrometry. *Nature Reviews Molecular Cell Biology* 8, 645–654.
- Glatter, T., Ludwig, C., Ahn, E., *et al.*, 2012. Large-scale quantitative assessment of different in-solution protein digestion protocols reveals superior cleavage efficiency of tandem Lys-C/trypsin proteolysis over trypsin digestion. *Journal of Proteome Research* 11, 5145–5156.
- Glatter, T., Wepf, A., Aebersold, R., Gstaiger, M., 2009. An integrated workflow for charting the human interaction proteome: Insights into the PP2A system. *Molecular Systems Biology* 5, 237.
- de Godoy, L.M., Olsen, J.V., Cox, J., *et al.*, 2008. Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature* 455, 1251–1254.
- Gygi, S.P., Rist, B., Gerber, S.A., *et al.*, 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnology* 17, 994–999.
- Hauri, S., Wepf, A., van Drogen, A., *et al.*, 2013. Interaction proteome of human Hippo signaling: Modular control of the co-activator YAP1. *Molecular Systems Biology* 9, 713.
- Havugimana, P.C., Hart, G.T., Nepusz, T., *et al.*, 2012. A census of human soluble protein complexes. *Cell* 150, 1068–1081.
- Hebert, A.S., Richards, A.L., Bailey, D.J., *et al.*, 2014. The one hour yeast proteome. *Molecular & Cellular Proteomics* 13, 339–347.
- Huttenhain, R., Surinova, S., Ossola, R., *et al.*, 2013. N-glycoprotein SRMAtlas: A resource of mass spectrometric assays for N-glycosites enabling consistent and multiplexed protein quantification for clinical applications. *Molecular & Cellular Proteomics* 12, 1005–1016.
- Huttlin, E.L., Jedrychowski, M.P., Elias, J.E., *et al.*, 2010. A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* 143, 1174–1189.
- Ideker, T., Galitski, T., Hood, L., 2001. A new approach to decoding life: Systems biology. *Annual Review of Genomics and Human Genetics* 2, 343–372.
- Johnson, H., Del Rosario, A.M., Bryson, B.D., *et al.*, 2012. Molecular characterization of EGFR and EGFRvIII signaling networks in human glioblastoma tumor xenografts. *Molecular & Cellular Proteomics* 11, 1724–1740.
- Keller, A., Nesvizhskii, A.I., Kolker, E., Aebersold, R., 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Analytical Chemistry* 74, 5383–5392.
- Keller, A., Shteynberg, D., 2011. Software pipeline and data analysis for MS/MS proteomics: The trans-proteomic pipeline. *Methods in Molecular Biology* 694, 169–189.
- Kim, M.-S., Pinto, S.M., Getnet, D., *et al.*, 2014. A draft map of the human proteome. *Nature* 509, 575–581.
- Kim, W., Bennett, E.J., Huttlin, E.L., *et al.*, 2011. Systematic and quantitative assessment of the ubiquitin-modified proteome. *Molecular Cell* 44, 325–340.
- Kratchmarova, I., Blagoev, B., Haack-Sorensen, M., Kassem, M., Mann, M., 2005. Mechanism of divergent growth factor effects in mesenchymal stem cell differentiation. *Science* 308, 1472–1477.
- Krogan, N.J., Cagney, G., Yu, H., *et al.*, 2006. Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* 440, 637–643.
- Kuster, B., Schirle, M., Mallick, P., Aebersold, R., 2005. Scoring proteomes with proteotypic peptide probes. *Nature Reviews Molecular Cell Biology* 6, 577–583.

- Lambert, J.P., Iovese, G., Couzens, A.L., *et al.*, 2013. Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. *Nature Methods* 10, 1239–1245.
- Lander, E.S., 2011. Initial impact of the sequencing of the human genome. *Nature* 470, 187–197.
- Lane, K.R., Yu, Y., Lackey, P.E., *et al.*, 2013. Cell cycle-regulated protein abundance changes in synchronously proliferating HeLa cells include regulation of pre-mRNA splicing proteins. *PLoS One* 8, e58456.
- Li, J., Bennett, K., Stukalov, A., *et al.*, 2013. Perturbation of the mutated EGFR interactome identifies vulnerabilities and resistance mechanisms. *Molecular Systems Biology* 9, 705.
- Link, A.J., Eng, J., Schieltz, D.M., *et al.*, 1999. Direct analysis of protein complexes using mass spectrometry. *Nature Biotechnology* 17, 676–682.
- Liu, Y., Huttenhain, R., Surinova, S., *et al.*, 2013. Quantitative measurements of N-linked glycoproteins in human plasma by SWATH-MS. *Proteomics* 13, 1247–1256.
- Ludwig, C., Claassen, M., Schmidt, A., Aebersold, R., 2012. Estimation of absolute protein quantities of unlabeled samples by selected reaction monitoring mass spectrometry. *Molecular & Cellular Proteomics* 11 (M111), 013987.
- Lundby, A., Lage, K., Weinert, B.T., *et al.*, 2012a. Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. *Cell Reports* 2, 419–431.
- Lundby, A., Secher, A., Lage, K., *et al.*, 2012b. Quantitative maps of protein phosphorylation sites across 14 different rat organs and tissues. *Nature Communications* 3, 876.
- Ly, T., Ahmad, Y., Shlien, A., *et al.*, 2014. A proteomic chronology of gene expression through the cell cycle in human myeloid leukemia cells. *eLife* 3, e01630.
- MacLean, B., Tomazela, D.M., Shulman, N., *et al.*, 2010. Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26, 966–968.
- Malovannaya, A., Lanz, R.B., Jung, S.Y., *et al.*, 2011. Analysis of the human endogenous coregulator complexome. *Cell* 145, 787–799.
- Michalski, A., Cox, J., Mann, M., 2011a. More than 100,000 detectable peptide species elute in single shotgun proteomics runs but the majority is inaccessible to data-dependent LC-MS/MS. *Journal of Proteome Research* 10, 1785–1793.
- Michalski, A., Damoc, E., Hauschild, J.P., *et al.*, 2011b. Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Molecular & Cellular Proteomics* 10 (M111), 011015.
- Miraldi, E.R., Sharfi, H., Friedline, R.H., *et al.*, 2013. Molecular network analysis of phosphotyrosine and lipid metabolism in hepatic PTP1b deletion mice. *Integrative Biology: Quantitative Biosciences from Nano to Macro* 5, 940–963.
- Moghaddas Gholami, A., Hahne, H., Wu, Z., *et al.*, 2013. Global proteome analysis of the NCI-60 cell line panel. *Cell Reports* 4, 609–620.
- Munoz, J., Low, T.Y., Kok, Y.J., *et al.*, 2011. The quantitative proteomes of human-induced pluripotent stem cells and embryonic stem cells. *Molecular Systems Biology* 7, 550.
- Nagaraj, N., Kulak, N.A., Cox, J., *et al.*, 2012. System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. *Molecular & Cellular Proteomics* 11 (M111), 013722.
- Nagaraj, N., Wisniewski, J.R., Geiger, T., *et al.*, 2011. Deep proteome and transcriptome mapping of a human cancer cell line. *Molecular Systems Biology* 7, 548.
- Nesvizhskii, A.I., Aebersold, R., 2005. Interpretation of shotgun proteomic data: The protein inference problem. *Molecular & Cellular Proteomics* 4, 1419–1440.
- Ohta, S., Bukowski-Wills, J.C., Sanchez-Pulido, L., *et al.*, 2010. The protein composition of mitotic chromosomes determined using multiclassifier combinatorial proteomics. *Cell* 142, 810–821.
- Oliveira, A.P., Ludwig, C., Picotti, P., *et al.*, 2012. Regulation of yeast central metabolism by enzyme phosphorylation. *Molecular Systems Biology* 8, 623.
- Olsen, J.V., Blagoev, B., Gnäd, F., *et al.*, 2006. Global, *in vivo*, and site-specific phosphorylation dynamics in signaling networks. *Cell* 127, 635–648.
- Olsen, J.V., Mann, M., 2013. Status of large-scale analysis of post-translational modifications by mass spectrometry. *Molecular & Cellular Proteomics* 12, 3444–3452.
- Olsen, J.V., Vermeulen, M., Santamaria, A., *et al.*, 2010. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Science Signaling* 3, ra3.
- Ong, S.E., Blagoev, B., Kratchmarova, I., *et al.*, 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & Cellular Proteomics* 1, 376–386.
- Pagliuca, F.W., Collins, M.O., Lichawska, A., *et al.*, 2011. Quantitative proteomics reveals the basis for the biochemical specificity of the cell-cycle machinery. *Molecular Cell* 43, 406–417.
- Panchaud, A., Scherl, A., Shaffer, S.A., *et al.*, 2009. Precursor acquisition independent from ion count: How to dive deeper into the proteomics ocean. *Analytical Chemistry* 81, 6481–6488.
- Picotti, P., Aebersold, R., 2012. Selected reaction monitoring-based proteomics: Workflows, potential, pitfalls and future directions. *Nature Methods* 9, 555–566.
- Picotti, P., Bodenmiller, B., Mueller, L.N., Domon, B., Aebersold, R., 2009. Full dynamic range proteome analysis of *S. cerevisiae* by targeted proteomics. *Cell* 138, 795–806.
- Picotti, P., Clement-Ziza, M., Lam, H., *et al.*, 2013. A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. *Nature* 494, 266–270.
- Picotti, P., Rinner, O., Stallmach, R., *et al.*, 2010. High-throughput generation of selected reaction-monitoring assays for proteins and proteomes. *Nature Methods* 7, 43–46.
- Plumb, R.S., Johnson, K.A., Rainville, P., *et al.*, 2006. UPLC/MSE; a new approach for generating molecular fragment information for biomarker structure elucidation. *Rapid Communications in Mass Spectrometry* 20, 1989–1994.
- Reiter, L., Rinner, O., Picotti, P., *et al.*, 2011. mProphet: Automated data processing and statistical validation for large-scale SRM experiments. *Nature Methods* 8, 430–435.
- Rikova, K., Guo, A., Zeng, Q., *et al.*, 2007. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 131, 1190–1203.
- Ross, P.L., Huang, Y.N., Marchese, J.N., *et al.*, 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Molecular & Cellular Proteomics* 3, 1154–1169.
- Röst, H.L., Rosenberger, G., Navarro, P., *et al.*, 2014. OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nature Biotechnology* 32, 219–223.
- Sabido, E., Wu, Y., Bautista, L., *et al.*, 2013. Targeted proteomics reveals strain-specific changes in the mouse insulin and central metabolic pathways after a sustained high-fat diet. *Molecular Systems Biology* 9, 681.
- Schmidt, A., Gehlenborg, N., Bodenmiller, B., *et al.*, 2008. An integrated, directed mass spectrometric approach for in-depth characterization of complex peptide mixtures. *Molecular & Cellular Proteomics* 7, 2138–2150.
- Schubert, O.T., Mouritsen, J., Ludwig, C., *et al.*, 2013. The Mtb proteome library: A resource of assays to quantify the complete proteome of *Mycobacterium tuberculosis*. *Cell Host & Microbe* 13, 602–612.
- Shen, Y., Kim, J., Strittmatter, E.F., *et al.*, 2005. Characterization of the human blood plasma proteome. *Proteomics* 5, 4034–4045.
- Shevchenko, A., Jensen, O.N., Podtelejnikov, A.V., *et al.*, 1996. Linking genome and proteome by mass spectrometry: Large-scale identification of yeast proteins from two dimensional gels. *Proceedings of the National Academy of Sciences of the United States of America* 93, 14440–14445.
- Sowa, M.E., Bennett, E.J., Gygi, S.P., Harper, J.W., 2009. Defining the human deubiquitinating enzyme interaction landscape. *Cell* 138, 389–403.
- Stoevesandt, O., Taussig, M.J., 2012. Affinity proteomics: The role of specific binding reagents in human proteome analysis. *Expert Review of Proteomics* 9, 401–414.
- Stumpf, M.P., Thorne, T., de Silva, E., *et al.*, 2008. Estimating the size of the human interactome. *Proceedings of the National Academy of Sciences of the United States of America* 105, 6959–6964.
- Surinova, S., Schiess, R., Huttenhain, R., *et al.*, 2011. On the development of plasma protein biomarkers. *Journal of Proteome Research* 10, 5–16.
- Thiele, I., Swainston, N., Fleming, R.M., *et al.*, 2013. A community-driven global reconstruction of human metabolism. *Nature Biotechnology* 31, 419–425.
- Thompson, A., Schafer, J., Kuhn, K., *et al.*, 2003. Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical Chemistry* 75, 1895–1904.
- Varjosalo, M., Kesitalo, S., Van Drogen, A., *et al.*, 2013. The protein interaction landscape of the human CMGC kinase group. *Cell Reports* 3, 1306–1320.
- Venable, J.D., Dong, M.Q., Wohlschlegel, J., Dillin, A., Yates, J.R., 2004. Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nature Methods* 1, 39–45.
- Venkatesan, K., Rual, J.F., Vazquez, A., *et al.*, 2009. An empirical framework for binary interactome mapping. *Nature Methods* 6, 83–90.
- Wagner, S.A., Beli, P., Weinert, B.T., *et al.*, 2011. A proteome-wide, quantitative survey of *in vivo* ubiquitylation sites reveals widespread regulatory roles. *Molecular & Cellular Proteomics* 10 (M111), 013284.
- Wilhelm, M., Schlegl, J., Hahne, H., *et al.*, 2014. Mass-spectrometry-based draft of the human proteome. *Nature* 509, 582–587.

- Wisniewski, J.R., Mann, M., 2012. Consecutive proteolytic digestion in an enzyme reactor increases depth of proteomic and phosphoproteomic analysis. *Analytical Chemistry* 84, 2631–2637.
- Wolf-Yadlin, A., Kumar, N., Zhang, Y., *et al.*, 2006. Effects of HER2 overexpression on cell signaling networks governing proliferation and migration. *Molecular Systems Biology* 2, 54.
- Wu, L., Candille, S.I., Choi, Y., *et al.*, 2013. Variation and genetic control of protein abundance in humans. *Nature* 499, 79–82.
- Xu, G., Paige, J.S., Jaffrey, S.R., 2010. Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nature Biotechnology* 28, 868–873.
- Yu, H., Braun, P., Yildirim, M.A., *et al.*, 2008. High-quality binary protein interaction map of the yeast interactome network. *Science* 322, 104–110.
- Zhang, Y., Wolf-Yadlin, A., Ross, P.L., *et al.*, 2005. Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Molecular & Cellular Proteomics* 4, 1240–1250.
- Zielinska, D.F., Gnad, F., Wisniewski, J.R., Mann, M., 2010. Precision mapping of an *in vivo* N-glycoproteome reveals rigid topological and sequence constraints. *Cell* 141, 897–907.

Phosphoproteomics: Approaches, Developments, and Challenges

MR Larsen and AVG Edwards, Protein Research Group, Campusvej, Odense M, Denmark

© 2016 Elsevier Inc. All rights reserved.

Introduction to Posttranslational Modifications

Many proteins are posttranslationally modified and it is those modifications and their interplay that often determine the function, interaction, and physio-chemical properties of these proteins. A posttranslational modification (PTM) is the covalent attachment of a chemical group to a protein after protein synthesis or the proteolytic removal of a signal peptide after cellular localization. More than 300 different PTMs are believed to exist (Husi and Grant, 2001), however, several of these modifications can be introduced during conventional sample preparation for proteomics applications. It is therefore of utmost importance to optimize sample preparation procedures for proteomics workflows in such a way that most of these modifications can be avoided in order to decrease sample complexity and eliminate nonbiological noise. Of these artefactual modifications, oxidation of methionine, carbamylation, acetylation, and formylation are among the most frequent, making these difficult to study when occurring naturally if no precautions are taken. For example, in a recent study it was found that acetic acid can induce a significant amount of lysine acetylation (Choudhary *et al.*, 2009) and therefore acetic acid should be avoided in studies aiming at characterizing the acetylome.

Below is given an introduction to the characterization of phosphorylated proteins by mass spectrometric methods. Not only does protein phosphorylation have a very substantial impact on cellular activity, we are able to detect and analyse it on a high-throughput basis; this offers an excellent opportunity to drill very deeply into the set of phosphorylated proteins and link this to biology. This is an important benefit; as we probe ever deeper into the PTMome, it becomes clear that PTMs are very widely distributed, and it is therefore necessary to have a correspondingly large-scale view of their distribution and dynamics in order to fully understand their function. In the case of phosphorylation, its amenity to analysis and evident biological importance create an attractive opportunity for proteomics. We focus below on the methods most commonly used to assess phosphorylation in proteomics studies, and the application of these in tissue or primary cell-based contexts, with their particular obstacles.

Protein Phosphorylation

Eukaryotic cells are highly dynamic, needing to respond effectively to alterations in the extracellular environment in order to maintain their basic function. One of the primary mechanisms by which cells respond to environmental stimuli is via protein phosphorylation and intracellular signaling, whereby nucleotide-derived phosphate (PO_4^{3-}) groups are added or subtracted from specific residues within proteins through the action of protein kinases or phosphatases, respectively. The amino acid residues targeted are specifically

serine(S), threonine(T), or tyrosine(Y) (in eukaryotes; prokaryotes also display phosphorylation at other residues, including histidine, asparagine, aspartic acid (Fuhrmann *et al.*, 2009; Wagner and Vu, 2000), and arginine (Fuhrmann *et al.*, 2009)) though as detection and interpretation methods become more sensitive and accurate, more residues are being implicated as targets of phosphorylation. Histidine is one such amino acid residue, with numerous reports of its presence in eukaryotic systems (Klump and Kriegelstein, 2009; Klump *et al.*, 2009), though there is as not yet enough evidence to say whether it is as widespread or has as substantial an impact as S, T, or Y phosphorylation due to the technical difficulties in the analysis of histidine phosphorylation.

The addition of phosphate is able to modify the structure, activity, interaction, or localization of the target protein to achieve a particular function, whether as a direct functional consequence (e.g., conformational change (Garza *et al.*, 2010)) or as part of a larger signaling cascade (such as is found in mitogen-activated protein kinases (Shaul and Seger, 2005)). On a chemical basis, the addition of phosphate has several effects, these primarily comprising the introduction of negative charge and hydrophilicity to the relevant region of the protein in question. These in turn may have several effects on the tertiary structure, the compartmentalization and/or interactions of the protein; for instance, the decreased hydrophobicity of the phosphorylated protein can reduce the degree to which it is membrane-associated (Ladokhin and White, 2001). Similarly, the increased negative charge imparted can impact on a protein's ability to bind to interaction partners (Joughin *et al.*, 2005; Groban *et al.*, 2006; Woods *et al.*, 2008; Woods and Ferre, 2005).

Phosphorylation pathways have been demonstrated to regulate a host of processes, both physiological and pathophysiological, and our understanding of them is therefore central to our understanding of how cell function is controlled in health and disease. Phosphoproteomics (Larsen *et al.*, 2001) is a relative new field within proteomics aimed at performing large-scale characterization of phosphorylated proteins present in a biological sample at a given time, in order to achieve a global understanding of phosphorylation-regulated signaling. Compared to conventional characterization of protein phosphorylation such as Western blotting (Gorog *et al.*, 2009; Obal *et al.*, 2005), phosphoproteomics, combined with efficient quantitative methods, offer a highly efficient and unbiased characterization of a large subset of proteins inside a cell and this in turn enable the efficient characterization of various cellular pathways and the discovery of novel members of such pathways. This is potentially very important in drug development as many targets for therapeutic intervention are kinases (Lapenna and Giordano, 2009).

Today, phosphoproteomics is typically performed using highly efficient phosphopeptide enrichment techniques in combination with mass spectrometry (MS), which has improved tremendously in sensitivity, dynamic range, and

accuracy in recent years. The success achieved in performing phosphoproteomics, especially when characterizing lower amounts of sample material (<1 mg) is to a large extent dependent on the efficiency of the phosphopeptide enrichment strategy and efficient fragmentation of phosphopeptides in the subsequent MS analysis, which together allow efficient identification and quantification of phosphopeptides. A number of research groups have worked extensively with developing novel methods for selective isolation of phosphopeptides prior to MS analysis, which have today resulted in multiple strategies for their isolation. Below are described some of the most commonly used ones for phosphopeptide enrichment.

Phosphopeptide Enrichment

Phosphopeptides, derived from proteins by proteolysis, are rarely detected when performing normal liquid chromatography coupled to tandem MS (LC-MS/MS) shotgun proteomics experiments. This is due to low stoichiometry of many phosphorylation sites on proteins, a generally low abundance of the phosphoproteins involved in signaling pathways (e.g., kinases and phosphatases) and a general lower signal intensity of phosphorylated peptides – especially multiply phosphorylated peptides – if analyzed by MS in the presence of non-modified peptides. Therefore, phosphopeptides must be enriched prior to LC-MS/MS. Numerous strategies for phosphopeptide enrichment have been developed over the last decade, including beta-elimination coupled to affinity purification (McLachlin and Chait, 2003; Oda *et al.*, 2001), phosphoramidate chemistry (Zhou *et al.*, 2001), immobilized metal affinity chromatography (IMAC) (Li and Dass, 1999; Neville *et al.*, 1997; Posewitz and Tempst, 1999), titanium dioxide (TiO₂) chromatography (Kuroda *et al.*, 2004; Larsen *et al.*, 2005; Pinkse *et al.*, 2004; Sano and Nakamura, 2004), sequential elution from IMAC (SIMAC) (Thingholm *et al.*, 2008), or antibody pull-down (e.g., immunoprecipitation Schmelzle *et al.*, 2006). Of these strategies, the most used are IMAC, TiO₂, and antibody pull-down of tyrosine phosphorylated peptides.

IMAC was first used for enrichment of phosphopeptides in 1997 (Neville *et al.*, 1997) and has for many years been the method of choice for enrichment of phosphopeptides from peptide mixtures. When performing IMAC, the phosphopeptides are reversibly bound to a chelating material containing metal ions, which selectively bind the phosphate group. The most commonly used metal ions are iron (Dubrovskaya and Souchelnytskyi, 2005) or gallium (Aryal *et al.*, 2008); however, Ti⁴⁺ have recently been shown to be promising for phosphopeptide enrichment (Zhou *et al.*, 2008). Phosphopeptides are normally bound to the IMAC material in an acidic solution containing organic solvent to prevent nonspecific binding to the material, and are subsequently eluted from the material using high or low pH, metal chelators such as ethylenediaminetetraacetate (EDTA), or buffers including molecules that interfere with the binding between the metal ion and the phosphate group, such as phosphoric acid or phosphate buffer. Despite the wide use of IMAC, several shortcomings have been reported, including incompatibility with several

common biological reagents (including EDTA and alkaline metal salts) (Jensen and Larsen, 2007), lower specificity when working with more complex samples (McNulty and Annan, 2009), selective binding of acidic peptides, and its preference for enriching multiply phosphorylated peptides (Jensen and Larsen, 2007). The latter shortcoming was actively used in the development of the SIMAC strategy (Thingholm *et al.*, 2008) where mono-phosphorylated peptides could be separated from multiply phosphorylated peptides by eluting the mono-phosphorylated and unmodified peptides off the IMAC material using 1% TFA (trifluoroacetic acid). The multiply phosphorylated peptides could then subsequently be eluted with high pH. The mono-phosphorylated and unmodified fraction was further purified using TiO₂ chromatography.

The current most widely used phosphopeptide enrichment technique is TiO₂ chromatography, first reported in a proteomic setting in 2004 (Kuroda *et al.*, 2004; Pinkse *et al.*, 2004; Sano and Nakamura, 2004) and later improved by Larsen *et al.* (Larsen *et al.*, 2005; Jensen and Larsen, 2007; Thingholm *et al.*, 2006). TiO₂ enrichment is based on the interaction between the negatively charged phosphate group and the metal oxide. The exact molecular mechanisms for this binding are unclear, though it is suggested to be a bridging bi-dentate formation (Larsen *et al.*, 2005; Michels *et al.*, 2000). The conditions under which maximally efficient phosphopeptide binding is achieved has been the subject of much investigation and are known with a high degree of confidence. As with IMAC, TiO₂ originally was found to bind a large number of acidic peptides (Pinkse *et al.*, 2004). This effect is significantly reduced by inclusion of a multifunctional acid such as dihydroxy benzoic acid (DHB) (Larsen *et al.*, 2005), phthalic acid (Thingholm *et al.*, 2006), glycolic acid (Jensen and Larsen, 2007), or lactic acid (Sugiyama *et al.*, 2007) to the enrichment media as a non-phosphopeptide excluder. As a result, loading and elution conditions for TiO₂ are very specific, where acidic loading buffers show reduced nonspecific interactions from negatively charged, non-phosphorylated peptides and basic elution conditions allow for release of the phosphopeptides from the chromatographic material. However, the exact buffers used for TiO₂ chromatography can vary depending on the TiO₂ material used, since many different TiO₂ materials can be produced.

Enrichment of phosphopeptides by the above-mentioned methods is not inherently biased toward serine, threonine, or tyrosine events, but the low abundance nature of tyrosine phosphorylation means that specific enrichment strategies must be adopted to reliably and reproducibly detect it. Even though accounting for only a small proportion of phosphorylated residues (conventionally quoted as between 1% and 5% (Thingholm *et al.*, 2009), experimental data vary considerably (Grimsrud *et al.*, 2010; Rogers and Foster 2009; Boersemma *et al.*, 2009; Rikova *et al.*, 2007)). Tyrosine phosphorylation is biologically important, as exemplified by receptor tyrosine kinases (RTK), the first step in many signaling cascades (Dengjel *et al.*, 2009). Acting as cell surface receptors, the numerous classes of RTK are activated by cytokines and growth factors, enabling early signal transmission within the cell. As such, phosphotyrosine residues are often specifically investigated, using an antibody approach or as an adjunct to a global phosphoproteomic approach (Villén *et al.*, 2007).

Monoclonal antibodies directed against phosphotyrosine are effective and readily available, and are the most common way to probe the phosphotyrosine proteome. The same is not true of phosphoserine or phosphothreonine, as the relatively less characteristic structures of these residues decrease the specificity of antibody methodologies (Gronborg *et al.*, 2002). Antibody-based methods are limited by the requirement for relatively large amounts of protein (typically of the order of 5–10 mg, as compared to many global phosphoproteomic experiment, which run with closer to 1 mg of protein), and are subject to similar levels of interference from non-target peptides as is observed with affinity-based chromatographic methods. An interesting corollary of these antibody-based experiments is that as more of them are performed, with more sensitive MS techniques and yet uniformly large sample amounts, investigators are able to probe ever deeper into the low abundance region of the phosphotyrosine proteome, such that a recent publication reported over 1100 unique pTyr sites (Boersema *et al.*, 2010). This result raises the question of whether this level of tyrosine phosphorylation represents 1–2% of a total cellular phosphoproteome of about 50 000–100 000, or whether the proportion of tyrosine phosphorylation is in fact lower than expected. PhosphositePlus, one of the largest databases of phosphorylation data, currently lists ~240 000 phosphorylation sites (Hornbeck *et al.*, 2012), and the pool of identified sites increases regularly, so it is entirely possible that the total cellular phosphoproteome is much larger than previously thought.

As is the case for any enrichment strategy, the decision to use IMAC, TiO₂ chromatography, or any of the above-mentioned approaches, either alone or in combination, will depend on the type of sample used, amount of starting material, and the desired outcomes of the workflow in question (Gates *et al.*, 2010). Furthermore, a study comparing phosphoramidate chemistry, TiO₂, and IMAC enrichments showed that no single method for phosphopeptide enrichment is able to identify all segments of the phosphoproteome, but that each approach documented overlapping but largely distinct populations (Bodenmiller *et al.*, 2007), suggesting that a combination of approaches is optimal; this is not surprising, as each of the three methods binds to phosphopeptides differently. Since the mass spectrometer has a limited analytical performance compared to the theoretical number of phosphopeptides present, another reason for the poor overlap could be under-sampling. Also, as with all comparisons made of biological methods, if one cannot perform each method to perfection, the result of such a comparison will differ depending on the person performing the experiments.

Large-Scale Phosphoproteomics Strategies

All proteomic studies, regardless of PTMs, rely on making samples for analysis as simple as possible in order to aid detection, thereby avoiding undersampling and increasing the coverage of the analytes. This is achieved using chromatographic fractionation, a conceptually simple operation whereby peptides are loaded onto a column of a solid stationary phase with a defined surface material, with which they interact and are retained. A mobile phase containing a solvent,

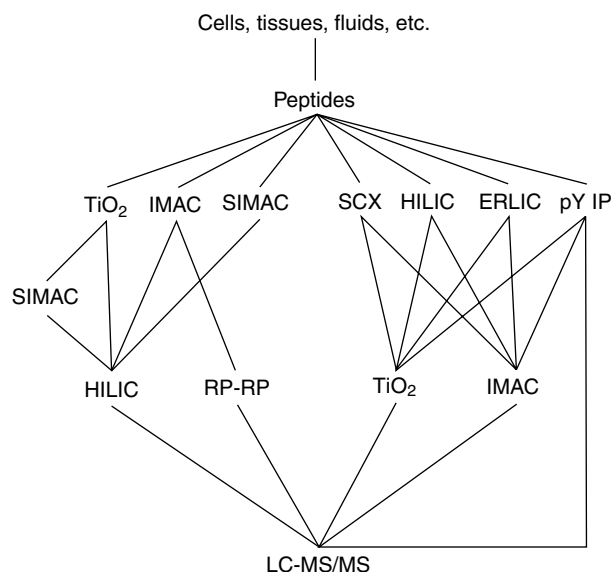


Figure 1 Examples of common large-scale phosphoproteomic strategies.

which reduces the strength of peptide-column interaction, is then applied to the stationary phase in an increasing gradient. The result is a gradual elution of peptides in a manner dependent on the degree of initial interaction. Several solid phases can be used effectively for phosphopeptide separation, though strong cation exchange (SCX) and hydrophilic interaction liquid chromatography (HILIC) are the most commonly used today (see overview of the most frequently used large-scale phosphoproteomics strategies in Figure 1). SCX chromatography involves allowing peptides to bind to a negatively charged resin followed by elution on an increasing gradient of salt (mobile phase), such that peptides with relatively lower charge states and therefore weaker interaction with the column elute earlier than those with strong interaction (Villén *et al.*, 2007; Gan *et al.*, 2008; Villén and Gygi, 2008). SCX is primarily a charge-state based separation, whereas HILIC, in contrast, is based purely on hydrophilicity in a directly orthogonal fashion to standard reversed-phase liquid chromatography (RP-LC) separation (Boersema *et al.*, 2008; Gilar *et al.*, 2005). When using HILIC, peptides are loaded in high concentrations of organic solvent (typically acetonitrile; mobile phase) with peptide hydrophilicity determining the strength of interaction with the stationary phase. Phosphopeptides, having an increased hydrophilic character due to the phosphate group, are retained on the column and eluted in decreasing concentrations of the organic solvent (McNulty and Annan 2009; Boersema *et al.*, 2008). These methods can, however, have some drawbacks; for example, the high salt mobile phase used for SCX chromatography is directly compatible with TiO₂ chromatography whereas a peptide cleaning up step might be needed for subsequent IMAC purification. In contrast, cleanup is not required for HILIC and this fractionation method is therefore directly compatible with TiO₂ and IMAC. A third fractionation procedure is electrostatic repulsion-hydrophilic interaction liquid chromatography (ERLIC (Zarei *et al.*, 2011)). This has been shown to

outperform SCX-IMAC, and also be able to enrich phosphopeptides and perform adequate separation in a single step (Gan *et al.*, 2008). A direct comparison between ERLIC and TiO₂ enrichment coupled with any kind of chromatography has not been performed.

Most of these fractionation methods are performed prior to the phosphopeptide enrichment step, which, if not used in an automated setup, is very labor intensive and could result in unwanted losses or poor enrichment of phosphopeptides due to suboptimal material/resin ratios (see below). Recently, a couple of large-scale phosphoproteomics strategies have been developed where the phosphopeptide enrichment is performed as the first step and then the purified phosphopeptides are further fractionated using other methods. For example, Ficarro *et al.* (2011) introduced the use of IMAC combined with a three-dimensional separation employing high pH RP, strong anion exchange (SAX), and low pH RP to increase the number of phosphopeptides from low amount of material (Ficarro *et al.*, 2011). In addition, Engholm-Keller *et al.* (2012) introduced the TiSH method involving an initial TiO₂ enrichment using an optimized TiO₂ beads: sample ratio (the TiO₂ beads: sample ratio has a substantial influence on the selectivity of the phosphopeptide enrichment) to enrich 'all' phosphopeptides in the sample. After enrichment, SIMAC was used to separate the multiply phosphorylated and mono-phosphorylated peptides from each other, followed by direct LC-MS/MS of the multiply phosphorylated fraction and HILIC separation of the mono-phosphorylated fraction prior to LC-MS/MS. The TiSH method yielded reproducibly more than 94% pure phosphopeptides from a complex sample and on an average 6500 unique phosphopeptides from 300 µg of starting material (Engholm-Keller *et al.*, 2012).

All these large-scale phosphoproteomics strategies mentioned above have been applied successfully to document the phosphoproteome of various mature cell types, including renal collecting duct cells (Rinschen *et al.*, 2010; Gunaratne *et al.*, 2010) and tissues such as intact liver (Villén *et al.*, 2007) and the mouse brain (e.g., Jedrychowski *et al.*, 2011; Palmisano *et al.*, 2012). Investigations on immortal cell lines are even more numerous and varied in their enrichment and fractionation approaches (Beausoleil *et al.*, 2004; Li *et al.*, 2007; Dephore *et al.*, 2008; Zhai *et al.*, 2008; Grimsrud *et al.*, 2009; Swaney *et al.*, 2009). However, here the amount of starting material in many cases is so high that any strategy for large-scale phosphoproteomics will be applicable. Again, as with phosphopeptide enrichment, it seems that orthogonal approaches will yield improved coverage of the phosphoproteome (Boersema *et al.*, 2007, 2008; Gilar *et al.*, 2005; Yates *et al.*, 2009).

Phosphopeptide Fragmentation

Tandem MS strategies for phosphoproteomics, when performed in a high-throughput manner, are based around fragmenting, sequencing, and identifying as many peptides as possible, as quickly as possible. It could be said that the rate-limiting step of phosphoproteomics is MS, in that the complexity of cell lysates usually means that only ~10–50% of a given sample will be sequenced and identified in any one

analysis, typically resulting in relatively low reproducibility between analyses. With these considerations in mind, an MS capable of as many fragmentation events as possible per unit time is preferable. Commonly, after enrichment of phosphopeptides they are sequenced by various fragmentation strategies, such as collision induced dissociation (CID) (McReynolds and Anbar 1977), phosphorylation directed MS³ (pdMS³) (Beausoleil *et al.*, 2004), multi-stage-activation (MSA) (Schroeder *et al.*, 2004), higher energy collisional dissociation (HCD) (Olsen *et al.*, 2007), or electron transfer dissociation (ETD) (Syka *et al.*, 2004). CID fragmentation frequently results in the loss of phosphoric acid from phosphopeptides with serine and threonine phosphorylation as the first fragmentation event and consequently little fragmentation of the peptide backbone to support the identification of the peptide sequence and phosphosite assignment. PdMS³ and MSA are two fragmentation strategies which take advantages of the loss of phosphoric acid observed in CID with concomitant selection of the signal originating from the loss of phosphoric acid for a subsequent fragmentation. Whereas pdMS³ provide the fragment ion spectrum from the signal originating from the loss of phosphoric acid, MSA is a sum of the fragment ions derived from the MSMS and MS³ fragmentations. The disadvantage of using pdMS³ is that the MS instrument needs to collect another ion-package which is associated with loss of analysis time. In addition, pdMS³ need high-intensity parent ion signals, as the peptide identification is based on the fragmentation of a selected daughter ion which will result in low signal intensity of the fragment ions.

Theoretically, MSA should provide higher sequence coverage than CID and pdMS³ for phosphopeptides; however, this is still debated and opposing reports exist on the subject (Ting *et al.*, 2011; Macek *et al.*, 2006). In contrast, HCD has been shown to be very efficient for phosphopeptides; however it is highly debated in the literature if HCD is better than MSA and CID for phosphopeptide sequencing (Jedrychowski *et al.*, 2011; Frese *et al.*, 2011; Nagaraj *et al.*, 2010). In our hands, HCD seems to provide a higher identification rate for phosphopeptides, especially singly phosphorylated peptides (Larsen MR, unpublished results), which we believe could be due to ion collision in the HCD cell that could result in multiple fragmentations per peptide molecule, with the initial loss of phosphoric acid followed by substantial backbone fragmentation. Loss of phosphoric acid is therefore not observed for singly phosphorylated peptides but mainly for multiply phosphorylated peptides (Larsen MR, unpublished results). A comparison of the fragmentation of a singly phosphorylated peptide is shown in Figure 2. Here, phosphopeptides were labeled with iTRAQ 4 plex reagents and subsequently purified using TiO₂ chromatography and analyzed by MSA and HCD. The selected peptide is AAAGPGAALSPR from PHD finger protein 10, where the serine is phosphorylated and the N-terminal labeled with iTRAQ. Figure 2(a) illustrates the fragmentation pattern obtained using MSA fragmentation and Figure 2(b) illustrates the fragmentation pattern using HCD. The main differences between these two spectra are the complete loss of phosphoric acid from the phosphoserine in the HCD spectrum making the site localization relatively straight forward, the low level of proline-directed fragmentation in HCD, and the abundant low mass signals in HCD compared to MSA.

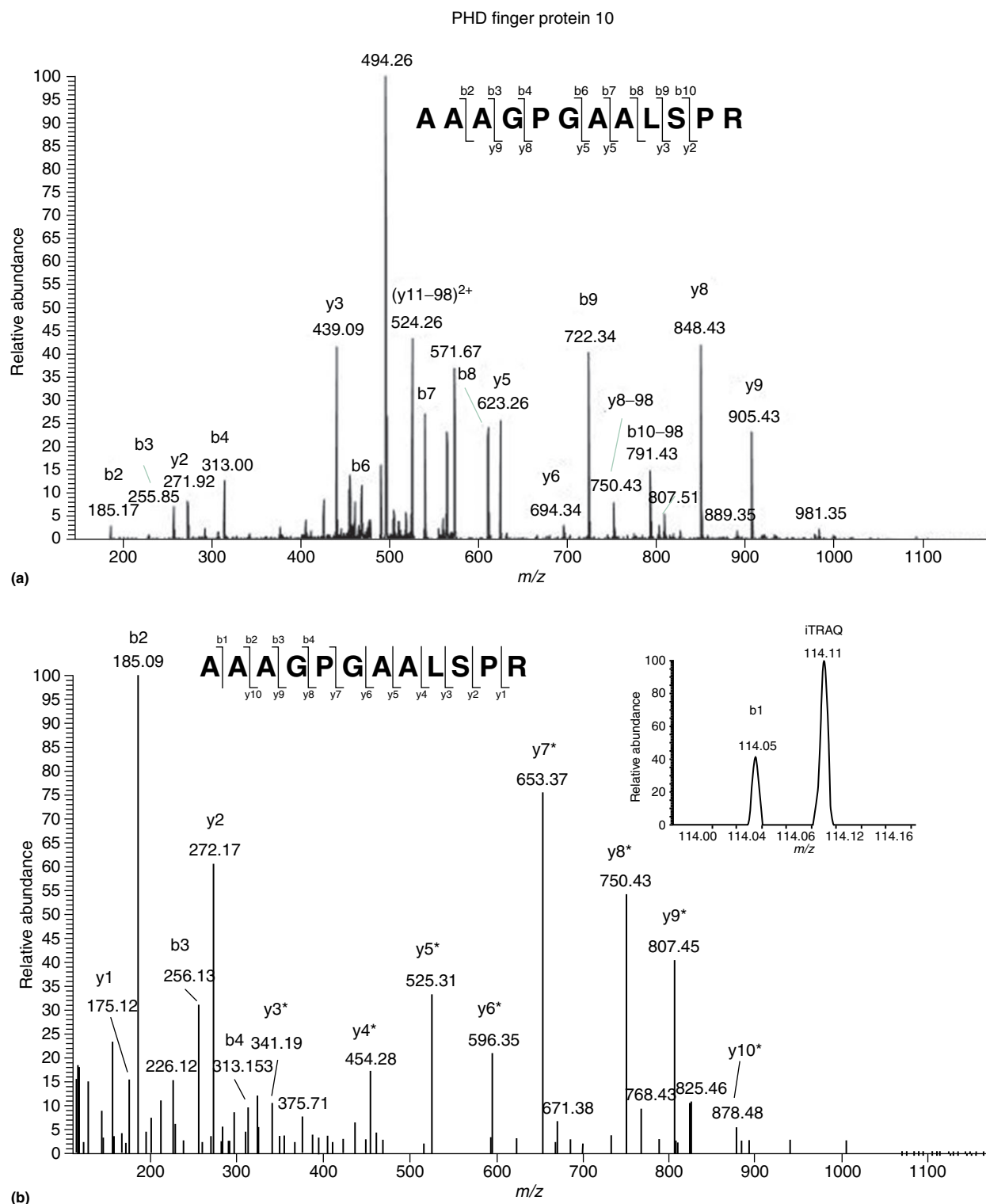


Figure 2 Fragmentation of phosphorylated peptides using different fragmentation strategies. (a) Fragmentation pattern obtained using MSA of a peptide from PHD finger protein 10, where the serine is phosphorylated and the N-terminal labeled with iTRAQ. (b) Fragmentation pattern of the same peptide obtained using HCD fragmentation.

ETD has also been used in phosphoproteomics studies as this fragmentation technique does not result in any loss of phosphoric acid from peptides carrying serine and threonine

phosphorylation (Chi *et al.*, 2007). However, the speed of ETD is generally low, resulting in fewer phosphopeptide identifications and the efficiency of fragmentation is highest for

peptides carrying more than two charges, which are not commonly observed in phosphoproteomics studies as trypsin is normally employed as the protease.

Quantitative Phosphoproteomics

A substantial amount of effort in phosphoproteomics is devoted to quantification – obviously, a binary 'present or absent' phosphorylation assessment is simplistic. A biological effect can be produced without phosphorylating every member of a protein pool, and in these cases it is the relative degree of phosphorylation that is critical. Fortunately, this is comparatively easy to investigate, depending on the nature of the analysis. Strategies for quantitative phosphoproteomics have recently been reviewed (e.g., [Palmisano and Thingholm, 2010](#)) and a detailed description of all the various strategies will not be given here. Briefly, phosphopeptides can be quantified using metabolic labeling employing either ^{15}N ([Oda et al., 1999](#)) or stable isotope labeling of amino acids in cell culture (SILAC) ([Ong et al., 2002](#); [Zhu et al., 2002](#)), various chemical mass tags (e.g., dimethyl ([Hsu et al., 2003](#)), isobaric tags for relative and absolute quantitation (iTRAQ) ([Ross et al., 2004](#); [Aggarwal et al., 2006](#)) or TMT ([Thompson et al., 2003](#))) or label free strategies. The choice of quantitation strategy for phosphoproteomics depends on various aspects. For example, if the proteins of interest originate from tissues or other biological material, SILAC will not be immediately applicable as this method is optimal for cell cultures. However, in many cases cell lines do not reflect the biology of primary cells and tissues. Metabolic labeling has been applied to whole organisms and subsequent phosphoproteomics studies in various animals such as mouse, rats, *Drosophila melanogaster*, and *Caenorhabditis elegans* ([Fredens et al., 2011](#); [Krijgsveld et al., 2003](#); [Kruger et al., 2008](#); [Wu et al., 2004](#)).

The application of isobaric tags such as iTRAQ and TMT tag requires MS instrumentation capable of stable and efficient analysis of fragment ions in the low mass area. This cannot be done using conventional CID in ion trap instruments due to the mechanism for CID, which does not allow for trapping of fragment masses below 28% of the precursor mass. By applying a relatively novel fragmentation mechanism termed pulsed Q-dissociation (PQD), the analysis of low mass ions in an ion trap become possible, though with a significant decrease in overall identification rate due to the need for multiple micro scans for each peptide to improve quantitation standard deviation ([Bantscheff et al., 2008](#)). HCD fragmentation or CID fragmentation on a QTOF instrument seems to be ideal for quantification using isobaric tags. One significant drawback of isobaric tags, besides that the labeling is performed after proteolysis digestion, is the increase in charge states introduced by these tags ([Thingholm et al., 2010](#)). Recently, we have showed that derivatization of peptides and phosphopeptides with iTRAQ and TMT tags results in a significant increase in charge state and a concomitant reduction in the identification efficiency ([Thingholm et al., 2010](#)). However, the inclusion of ammonia vapor under the electrospray needle was shown to decrease the otherwise increased charge state of isobaric tagged peptides and thereby to improve the identification efficiency of tagged peptides and

phosphopeptides. Furthermore, the ammonia vapor significantly decreases the standard deviation of the isobaric tag analysis as the standard deviation is dependent on the charge state of the analyzed peptides (Larsen MR, unpublished results).

Undersampling, the fact that not all co-eluting peptides will be selected for MSMS, is an issue when performing multiplex comparison of sample conditions and this is most prominent when performing quantification on the MS level (e.g., SILAC and Dimethyl labeling). This is because each peptide is represented by several signals only differing in mass due to the numbers and nature of different isotopes used in the label, reducing the number of unique phosphopeptides selected and fragmented in conventional data dependent analyses.

The isobaric tag approach is compatible with phosphorylation analysis ([Leitner and Lindner, 2009](#); [Gafken and Lampe, 2006](#)) and has been applied to neuronal samples ([Rudrabhatla et al., 2010](#)), fibroblast cells (using ETD) ([Yang et al., 2009](#)), mitochondrial samples (using HCD) ([Boja et al., 2009](#)), and many others (e.g., [Zhang et al., 2005](#); [Wu et al., 2010](#); [Reinl et al., 2009](#)). After quantification of reporter ions, phosphorylation events that depart from those observed in control states (either in absence/presence or in degree) can then be mapped onto pathways using various types of software.

PTM Cross Talk and PTM Coding

PTM cross talk is a relatively recently identified phenomenon which remains fairly loosely described. At a basic level, it can be understood as some form of communication between PTM sites, on a competitive or facilitative basis whereby one site prevents or allows the occurrence of the other, with predictable functional consequences. This phenomenon was first described for the cross talk between O-GlcNAcylation and phosphorylation in regulating signal transduction pathways (e.g., [Wang et al., 2008](#)). This PTM cross talk can be used to provide an extra level of modulation in protein responses by expanding the pool of effectors and providing more points at which to exert control. Cross talk seems to occur with many PTM types (e.g., [Yang and Seto, 2008](#)), but as phosphorylation is the most heavily studied it is well-represented in cross talk literature. For example, as mentioned above phosphorylation has repeatedly been shown to cross talk with O-linked glycosylation ([Wang et al., 2012](#)) to provide close control of Akt signaling in human cells. In addition, protein acetylation has been shown to have a significant influence on phosphorylation in several cellular systems such as the STAT signaling pathway, glucose metabolism, and the molecular response to ischemia and reperfusion ([Parker et al., 2014](#); [Zhang et al., 2012](#); [Wieczorek et al., 2012](#); [Ullmann et al., 2012](#)).

There is also substantial evidence for the role of phosphorylation in more complex PTM patterns that may be used as a form of protein level coding. This recent idea proposes that patterns of PTMs are interpreted as combinatorial coding units by downstream reader proteins, allowing for much more finely regulated control of protein-based processes; see [Figure 3](#) ([Wieczorek et al., 2012](#); [Creixell and Linding, 2012](#)). Evidence for this effect can be seen in situations requiring very tight control of very complex processes, for example, brain

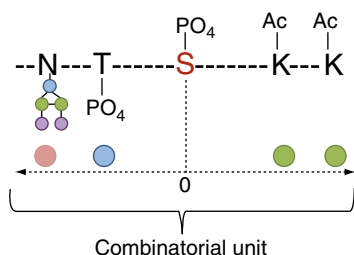


Figure 3 Use of PTMs in a combinatorial coding manner. The figure illustrates an example of PTMs clustering in a region of a protein.

development, wherein both large-scale patterns extending across the whole proteome as well as patterns occurring on individual highly modified proteins can be detected (Edwards *et al.*, 2014a,b).

With the significant increasing interest in PTMs in proteomics and the development of numerous new methods for assessing PTMs in cells, we expect that the field of PTM cross talk will expand dramatically in the coming years and reveal a substantial and yet to be seen complexity in cells.

Challenges in Phosphoproteomics

Today's phosphoproteomic analyses face a multitude of challenges despite the significant progress made in the field. Whereas now it has become relatively easy to isolate thousands of phosphorylated peptides with high specificity, due to the development of highly selective methods for phosphopeptide enrichment, the remaining challenges include accurate phosphopeptide identification, validation, and quantification, and the notorious lack of biological follow-up studies to identify the biological functions of individual phosphorylation sites. This latter point must be addressed if phosphoproteomics is to fulfill its potential with respect to improving our understanding of cell biology.

We, as well as other authors, have recently commented on the identification of phosphorylated peptides based on poor MS/MS signals giving rise to low identification scores in various software programs (White, 2011; Engholm-Keller *et al.*, 2011; Cooper, 2011). Despite recent progress in peptide 'validation' software such as Percolator (Kall *et al.*, 2007) and demands from journals for annotated spectra, we still believe that phosphopeptides in many studies are identified using criteria that are too relaxed and often associated with much higher FDRs (false discovery rates) than reported by the software programs.

An important consideration in phosphopeptide quantification is the interplay between protein abundance and the relative degree of phosphorylation. This site occupancy rate, or the proportion of 'phosphorylatable' sites that are in fact phosphorylated for a particular peptide species, is naturally an important consideration. Estimation of site occupancy for high-throughput approaches is a difficult task given that it requires both detection and accurate measurement of the phosphopeptide and the non-phosphorylated counterpart, and can only strictly be applied to singly phosphorylated

peptides. The biological impact of differential site occupancy is not yet clear but, as it appears to be a tightly regulated phenomenon (Olsen *et al.*, 2010) and given the need for fine control of the complex overlapping pathways, it seems likely to have an important role. Whilst the precise impact of these effects will likely need to be established on a case-by-case basis, a two-stage approach whereby the phosphoproteome and the entire proteome are analyzed concurrently may provide clues to the extent of pathway activation / deactivation in the context of proteome-level changes.

Another increasingly important challenge in phosphoproteomics is the identification and quantification of phosphorylation sites in protein splice variants, which carry the same phosphopeptides. Here, large-scale phosphoproteomics, where the proteins are cleaved into peptides using a specific protease, cannot distinguish the origin of a specific phosphopeptide if it is shared between several variants. If the splice variants participate in different biological pathways or molecular complexes then the quantitative readout will be an average of the splice variants. A perfect example of this was recently given by Chan *et al.* (2010) who examined the phosphorylation of dynamin I isoforms using optimized SDS-PAGE in combination with iTRAQ and LC-MS/MS. They revealed major differential phosphorylation of dynamin I phosphorylation sites in different variants and in different neuronal compartments that would be completely imperceptible to a large-scale phosphoproteomics approach.

Conclusions

As this article has shown, phosphoproteomics is an established yet still developing field. It has advanced enormously from its beginnings in small-scale analysis, but there are challenges that still remain to be addressed. Over the next decade it is our expectation that large-scale phosphorylation analysis will become more and more commonplace as the tools to perform it become more and more advanced, and we will get closer to 100% phosphoproteome coverage from increasingly smaller amounts of material. The suite of bioinformatics tools used to assign biological meaning to this data will also need to advance, though functional roles may only be able to be absolutely described through molecular biological follow-up studies.

See also: Cell Communication: Growth Factor Mediated Cell Signaling: Cytokine Receptors and Their Ligands; Receptor Tyrosine Kinases and Their Ligands; TGF- β Superfamily Signaling. Dynamic Integration: Dynamics: Dynamics of Protein Kinase Cascades. Horizontal Integration: Omics: The Advent of Mass Spectrometry-Based Proteomics in Systems Biology Research

References

- Aggarwal, K., Choe, L.H., Lee, K.H., 2006. Shotgun proteomics using the iTRAQisobaric tags. *Briefings in Functional Genomics and Proteomics* 5 (2), 112–120.
- Aryal, U.K., Olson, D.J., Ross, A.R., 2008. Optimization of immobilized gallium (III) ion affinity chromatography for selective binding and recovery of

- phosphopeptides from protein digests. *Journal of Biomolecular Techniques* 19 (5), 296–310.
- Bantscheff, M., *et al.*, 2008. Robust and sensitive iTRAQ quantification on an LTQ Orbitrap mass spectrometer. *Molecular & Cellular Proteomics* 7 (9), 1702–1713.
- Beausoleil, S.A., *et al.*, 2004. Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proceedings of the National Academy of Sciences of The United States of America* 101 (33), 12130–12135.
- Bodenmiller, B., *et al.*, 2007. Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nature Methods* 4 (3), 231–237.
- Boersema, P.J., Divecha, N., Heck, A.J., Mohammed, S., 2007. Evaluation and optimization of ZIC-HILIC-RP as an alternative MudPIT strategy. *Journal of Proteome Research* 6 (3), 937–946.
- Boersema, P.J., *et al.*, 2009. In depth qualitative and quantitative profiling of tyrosine phosphorylation using a combination of phosphopeptide immuno-affinity purification and stable isotope dimethyl labeling. *Molecular & Cellular Proteomics* 9 (1), 84–99.
- Boersema, P.J., *et al.*, 2010. In-depth qualitative and quantitative profiling of tyrosine phosphorylation using a combination of phosphopeptide immunoaffinity purification and stable isotope dimethyl labeling. *Molecular & Cellular Proteomics* 9 (1), 84–99.
- Boersema, P.J., Mohammed, S., Heck, A.J., 2008. Hydrophilic interaction liquid chromatography (HILIC) in proteomics. *Analytical and Bioanalytical Chemistry* 391 (1), 151–159.
- Boja, E.S., *et al.*, 2009. Quantitative mitochondrial phosphoproteomics using iTRAQ on an LTQ-Orbitrap with high energy collision dissociation. *Journal of Proteome Research* 8 (10), 4665–4675.
- Chan, L.S., Hansra, G., Robinson, P.J., Graham, M.E., 2010. Differential phosphorylation of dynamin I isoforms in subcellular compartments demonstrates the hidden complexity of phosphoproteomes. *Journal of Proteome Research* 9 (8), 4028–4037.
- Chi, A., *et al.*, 2007. Analysis of phosphorylation sites on proteins from *Saccharomyces cerevisiae* by electron transfer dissociation (ETD) mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* 104 (7), 2193–2198.
- Choudhary, C., *et al.*, 2009. Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 325 (5942), 834–840.
- Cooper, B., 2011. The problem with peptide presumption and low Mascot scoring. *Journal of Proteome Research* 10 (3), 1432–1435.
- Creixell, P., Lindner, R., 2012. Cells, shared memory and breaking the PTM code. *Molecular Systems Biology* 8, 598.
- Dengjel, J., Kratchmarova, I., Blagoev, B., 2009. Receptor tyrosine kinase signaling: A view from quantitative proteomics. *Molecular BioSystem* 5 (10), 1112–1121.
- Dephoure, N., *et al.*, 2008. A quantitative atlas of mitotic phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* 105 (31), 10762–10767.
- Dubrovskaya, A., Souchevnytskyi, S., 2005. Efficient enrichment of intact phosphorylated proteins by modified immobilized metal-affinity chromatography. *Proteomics* 5 (18), 4678–4683.
- Edwards, A.V., *et al.*, 2014a. Spatial and temporal effects in protein post-translational modification distributions in the developing mouse brain. *Journal of Proteome Research* 13 (1), 260–267.
- Edwards, A.V., Schwammle, V., Larsen, M.R., 2014b. Neuronal process structure and growth proteins are targets of heavy PTM regulation during brain development. *Journal of Proteomics* 101, 77–87.
- Engholm-Keller, K., *et al.*, 2012. TISH — A robust and sensitive global phosphoproteomics strategy employing a combination of TiO₂, SIMAC, and HILIC. *Journal of Proteomics* 75 (18), 5749–5761.
- Engholm-Keller, K., Hansen, T.A., Palmisano, G., Larsen, M.R., 2011. Multidimensional strategy for sensitive phosphoproteomics incorporating protein prefractionation combined with SIMAC, HILIC, and TiO₂ chromatography applied to proximal EGF signaling. *Journal of Proteome Research* 10 (12), 5383–5397.
- Ficarro, S.B., *et al.*, 2011. Online nanoflow multidimensional fractionation for high efficiency phosphopeptide analysis. *Molecular & Cellular Proteomics* 10 (11), 0111011064.
- Fredens, J., *et al.*, 2011. Quantitative proteomics by amino acid labeling in *C. elegans*. *Nature Methods* 8 (10), 845–847.
- Frese, C.K., *et al.*, 2011. Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. *Journal of Proteome Research* 10 (5), 2377–2388.
- Fuhrmann, J., *et al.*, 2009. MscB is a protein arginine kinase that phosphorylates and inhibits the heat-shock regulator CtsR. *Science* 324 (5932), 1323–1327.
- Gafken, P.R., Lampe, P.D., 2006. Methodologies for characterizing phosphoproteins by mass spectrometry. *Cell Communication & Adhesion* 13 (5–6), 249–262.
- Gan, C.S., *et al.*, 2008. A comparative study of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) versus SCX-IMAC-based methods for phosphopeptide isolation/enrichment. *Journal of Proteome Research* 7 (11), 4869–4877.
- Garza, A.M., Khan, S.H., Kumar, R., 2010. Site-specific phosphorylation induces functionally active conformation in the intrinsically disordered N-terminal activation function (AF1) domain of the glucocorticoid receptor. *Molecular and Cellular Biology* 30 (1), 220–230.
- Gates, M.B., Tomer, K.B., Deterding, L.J., 2010. Comparison of metal and metal oxide media for phosphopeptide enrichment prior to mass spectrometric analyses. *Journal of the American Society for Mass Spectrometry* 21 (10), 1649–1659.
- Gilar, M., Olivova, P., Daly, A.E., Gebler, J.C., 2005. Orthogonality of separation in two-dimensional liquid chromatography. *Analytical Chemistry* 77 (19), 6426–6434.
- Gorog, D.A., *et al.*, 2009. MAPKAPK-2 modulates p38-MAPK localization and small heat shock protein phosphorylation but does not mediate the injury associated with p38-MAPK activation during myocardial ischemia. *Cell Stress Chaperones* 14 (5), 477–489.
- Grimsrud, P.A., *et al.*, 2009. Large-scale phosphoprotein analysis in *Medicago truncatula* roots provides insight into in vivo kinase activity in legumes. *Plant Physiology* 152, 19–28.
- Grimsrud, P.A., *et al.*, 2010. Phosphoproteomics for the masses. *ACS Chemical Biology* 5 (1), 105–119.
- Groban, E.S., Narayanan, A., Jacobson, M.P., 2006. Conformational changes in protein loops and helices induced by post-translational phosphorylation. *PLoS Computational Biology* 2 (4), e32.
- Gronborg, M., *et al.*, 2002. A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies: Identification of a novel protein, Frigg, as a protein kinase A substrate. *Molecular & Cellular Proteomics* 1 (7), 517–527.
- Gunaratne, R., *et al.*, 2010. Quantitative phosphoproteomic analysis reveals cAMP/vasopressin-dependent signaling pathways in native renal thick ascending limb cells. *Proceedings of the National Academy of Sciences of the United States of America* 107 (35), 15653–15658.
- Hornbeck, P.V., *et al.*, 2012. PhosphoSitePlus: A comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Research* 40, D261–D270.
- Hsu, J.L., Huang, S.Y., Chow, N.H., Chen, S.H., 2003. Stable-isotope dimethyl labeling for quantitative proteomics. *Analytical Chemistry* 75 (24), 6843–6852.
- Husi, H., Grant, S.G., 2001. Proteomics of the nervous system. *Trends in Neurosciences* 24 (5), 259–266.
- Jedrychowski, M.P., *et al.*, 2011. Evaluation of HCD- and CID-type fragmentation within their respective detection platforms for murine phosphoproteomics. *Molecular & Cellular Proteomics* 10 (12), M111009910.
- Jensen, S.S., Larsen, M.R., 2007. Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques. *Rapid Communications in Mass Spectrometry* 21 (22), 3635–3645.
- Joughin, B.A., Tidor, B., Yaffe, M.B., 2005. A computational method for the analysis and prediction of protein:Phosphopeptide-binding sites. *Protein Science* 14 (1), 131–139.
- Kall, L., *et al.*, 2007. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nature Methods* 4 (11), 923–925.
- Klump, S., *et al.*, 2009. Role of protein histidine phosphatase for viability of neuronal cells. *Brain Research* 1264, 7–12.
- Klump, S., Krieglstein, J., 2009. Reversible phosphorylation of histidine residues in proteins from vertebrates. *Science Signaling* 2 (61), pe13.
- Krijgsvel, J., *et al.*, 2003. Metabolic labeling of *C. elegans* and *D. melanogaster* for quantitative proteomics. *Nature Biotechnology* 21 (8), 927–931.
- Kruger, M., *et al.*, 2008. SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell* 134 (2), 353–364.
- Kuroda, I., *et al.*, 2004. Phosphopeptide-selective column-switching RP-HPLC with a titanania precolumn. *Analytical sciences* 20 (9), 1313–1319.
- Ladokhin, A.S., White, S.H., 2001. Protein chemistry at membrane interfaces: Non-additivity of electrostatic and hydrophobic interactions. *Journal of Molecular Biology* 309 (3), 543–552.
- Lapenna, S., Giordano, A., 2009. Cell cycle kinases as therapeutic targets for cancer. *Nature Reviews Drug Discovery* 8 (7), 547–566.
- Larsen, M.R., *et al.*, 2001. Phospho-proteomics: Evaluation of the use of enzymatic de-phosphorylation and differential mass spectrometric peptide mass mapping for

- site specific phosphorylation assignment in proteins separated by gel electrophoresis. *Proteomics* 1 (2), 223–238.
- Larsen, M.R., *et al.*, 2005. Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Molecular and Cellular Proteomics* 4, 873–886.
- Leitner, A., Lindner, W., 2009. Chemical tagging strategies for mass spectrometry-based phospho-proteomics. *Methods in Molecular Biology* 527, 229–243, x.
- Li, S., Dass, C., 1999. Iron(III)-immobilized metal ion affinity chromatography and mass spectrometry for the purification and characterization of synthetic phosphopeptides. *Analytical Biochemistry* 270 (1), 9–14.
- Li, X., *et al.*, 2007. Large-scale phosphorylation analysis of alpha-factor-arrested *Saccharomyces cerevisiae*. *Journal of Proteome Research* 6 (3), 1190–1197.
- Macek, B., Waanders, L.F., Olsen, J.V., Mann, M., 2006. Top-down protein sequencing and MS3 on a hybrid linear quadrupole ion trap-orbitrap mass spectrometer. *Molecular & Cellular Proteomics* 5 (5), 949–958.
- McLachlin, D.T., Chait, B.T., 2003. Improved beta-elimination-based affinity purification strategy for enrichment of phosphopeptides. *Analytical Chemistry* 75 (24), 6826–6836.
- McNulty, D.E., Annan, R.S., 2009. Hydrophilic interaction chromatography for fractionation and enrichment of the phosphoproteome. *Methods in Molecular Biology* 527, 93–105.
- McReynolds, J.H., Anbar, M., 1977. Isotopic assay of nanomole amounts of nitrogen-15 labeled amino acids by collision-induced dissociation mass spectrometry. *Analytical Chemistry* 49 (12), 1832–1836.
- Micheltore, A., Gong, W., Jenkins, P., Ralston, J., 2000. The interaction of linear polyphosphates with titanium dioxide surfaces. *Physical Chemistry Chemical Physics* 2 (13), 2985–2992.
- Nagaraj, N., *et al.*, 2010. Feasibility of large scale phosphoproteomics with HCD fragmentation. *Journal of Proteome Research* 9 (12), 6786–6794.
- Neville, D.C., *et al.*, 1997. Evidence for phosphorylation of serine 753 in CFTR using a novel metal-ion affinity resin and matrix-assisted laser desorption mass spectrometry. *Protein Science* 6 (11), 2436–2445.
- Obal, D., *et al.*, 2005. Role of protein kinase C-epsilon (PKCepsilon) in isoflurane-induced cardioprotection. *British Journal of Anaesthesia* 94 (2), 166–173.
- Oda, Y., *et al.*, 1999. Accurate quantitation of protein expression and site-specific phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* 96 (12), 6591–6596.
- Oda, Y., Nagasu, T., Chait, B.T., 2001. Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nature Biotechnology* 19 (4), 379–382.
- Olsen, J.V., *et al.*, 2007. Higher-energy C-trap dissociation for peptide modification analysis. *Nature Methods* 4 (9), 709–712.
- Olsen, J.V., *et al.*, 2010. Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis. *Science Signaling* 3 (104), ra3.
- Ong, S.E., *et al.*, 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & Cellular Proteomics* 1 (5), 376–386.
- Palmisano, G., Parker, B.L., Engholm-Keller, K., *et al.*, 2012. A novel method for the simultaneous enrichment, identification, and quantification of phosphopeptides and sialylated glycopeptides applied to a temporal profile of mouse brain development. *Molecular & Cellular Proteomics* 11, 1191–1202.
- Palmisano, G., Thingholm, T.E., 2010. Strategies for quantitation of phosphoproteomic data. *Expert Review of Proteomics* 7 (3), 439–456.
- Parker, B.L., *et al.*, 2014. Structural basis for phosphorylation and lysine acetylation cross-talk in a kinase motif associated with myocardial ischemia and cardioprotection. *Journal of Biological Chemistry* 289 (37), 25890–25906.
- Pinkse, M.W.H., *et al.*, 2004. Selective isolation at the femtomole level of proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Analytical Chemistry* 76, 3935–3943.
- Posewitz, M.C., Tempst, P., 1999. Immobilized gallium(III) affinity chromatography of phosphopeptides. *Analytical Chemistry* 71 (14), 2883–2892.
- Reinl, T., *et al.*, 2009. Quantitative phosphokinome analysis of the Met pathway activated by the invasin internalin B from *Listeria monocytogenes*. *Molecular & Cellular Proteomics* 8 (12), 2778–2795.
- Rikova, K., *et al.*, 2007. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 131 (6), 1190–1203.
- Rinschen, M.M., *et al.*, 2010. Quantitative phosphoproteomic analysis reveals vasopressin V2-receptor-dependent signaling pathways in renal collecting duct cells. *Proceedings of the National Academy of Sciences of the United States of America* 107 (8), 3882–3887.
- Rogers, L.D., Foster, L.J., 2009. Phosphoproteomics—finally fulfilling the promise? *Molecular BioSystems* 5 (10), 1122–1129.
- Ross, P.L., *et al.*, 2004. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Molecular & Cellular Proteomics* 3 (12), 1154–1169.
- Rudrabhatla, P., *et al.*, 2010. Quantitative phosphoproteomic analysis of neuronal intermediate filament proteins (NF-M/H) in Alzheimer's disease by iTRAQ. *FASEB Journal* 11, 4396–4407.
- Sano, A., Nakamura, H., 2004. Chemo-affinity of titania for the column-switching HPLC analysis of phosphopeptides. *Analytical Sciences* 20 (3), 565–566.
- Schmelzle, K., *et al.*, 2006. Temporal dynamics of tyrosine phosphorylation in insulin signaling. *Diabetes* 55 (8), 2171–2179.
- Schroeder, M.J., *et al.*, 2004. A neutral loss activation method for improved phosphopeptide sequence analysis by quadrupole ion trap mass spectrometry. *Analytical Chemistry* 76 (13), 3590–3598.
- Shaul, Y., Seger, R., 2005. The detection of MAPK signaling. *Current Protocols in Cell Biology*. Chapter 14, Unit 14.3.
- Sugiyama, N., *et al.*, 2007. Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. *Molecular & Cellular Proteomics* 6 (6), 1103–1109.
- Swaney, D.L., Wenger, C.D., Thomson, J.A., Coon, J.J., 2009. Human embryonic stem cell phosphoproteome revealed by quadrupole transfer dissociation tandem mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* 106 (4), 995–1000.
- Syka, J.E., *et al.*, 2004. Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* 101 (26), 9528–9533.
- Thingholm, T.E., Jørgensen, T.J.D., Jensen, O.N., Larsen, M.R., 2006. Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nature Protocols* 1 (4), 1929–1935.
- Thingholm, T.E., Jensen, O.N., Larsen, M.R., 2009. Analytical strategies for phosphoproteomics. *Proteomics* 9 (6), 1451–1468.
- Thingholm, T.E., Jensen, O.N., Robinson, P.J., Larsen, M.R., 2008. SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. *Molecular & Cellular Proteomics* 7 (4), 661–671.
- Thingholm, T.E., Palmisano, G., Kjeldsen, F., Larsen, M.R., 2010. Undesirable charge-enhancement of isobaric tagged phosphopeptides leads to reduced identification efficiency. *Journal of Proteome Research* 9 (8), 4045–4052.
- Thompson, A., *et al.*, 2003. Tandem mass tags: A novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical Chemistry* 75 (8), 1895–1904.
- Ting, L., Rad, R., Gygi, S.P., Haas, W., 2011. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nature Methods* 8 (11), 937–940.
- Ullmann, R., Chien, C.D., Avantiaggiati, M.L., Muller, S., 2012. An acetylation switch regulates SUMO-dependent protein interaction networks. *Molecular Cell* 46 (6), 759–770.
- Villén, J., Beausoleil, S.A., Gerber, S.A., Gygi, S.P., 2007. Large-scale phosphorylation analysis of mouse liver. *Proceedings of the National Academy of Sciences of the United States of America* 104 (5), 1488–1493.
- Villen, J., Gygi, S.P., 2008. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nature Protocols* 3 (10), 1630–1638.
- Wagner, P.D., Vu, N.D., 2000. Histidine to aspartate phosphotransferase activity of nm23 proteins: Phosphorylation of aldolase C on Asp-319. *Biochemical Journal* 346 (Pt 3), 623–630.
- Wang, S., *et al.*, 2012. Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates Akt signaling. *PLoS One* 7 (5), e37427.
- Wang, Z., Gucuk, M., Hart, G.W., 2008. Cross-talk between GlcNAcylation and phosphorylation: Site-specific phosphorylation dynamics in response to globally elevated O-GlcNAc. *Proceedings of the National Academy of Sciences of the United States of America* 105 (37), 13793–13798.
- White, F.M., 2011. The potential cost of high-throughput proteomics. *Science Signaling* 4 (160), pe8.
- Wieczorek, M., *et al.*, 2012. Acetylation modulates the STAT signaling code. *Cytokine & Growth Factor Reviews* 23 (6), 293–305.
- Woods, A.S., Ferre, S., 2005. Amazing stability of the arginine-phosphate electrostatic interaction. *Journal of Proteome Research* 4 (4), 1397–1402.
- Woods, A.S., Moyer, S.C., Jackson, S.N., 2008. Amazing stability of phosphate-quaternary amine interactions. *Journal of Proteome Research* 7 (8), 3423–3427.
- Wu, C.C., *et al.*, 2004. Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Analytical Chemistry* 76 (17), 4951–4959.

- Wu, J., *et al.*, 2010. Integrating titania enrichment, iTRAQ labeling, and Orbitrap CID-HCD for global identification and quantitative analysis of phosphopeptides. *Proteomics* 10 (11), 2224–2234.
- Yang, F., *et al.*, 2009. Combined pulsed-Q dissociation and electron transfer dissociation for identification and quantification of iTRAQ-labeled phosphopeptides. *Analytical Chemistry* 81 (10), 4137–4143.
- Yang, X.J., Seto, E., 2008. Lysine acetylation: Codified crosstalk with other posttranslational modifications. *Molecular Cell* 31 (4), 449–461.
- Yates, J.R., Ruse, C.I., Nakorchevsky, A., 2009. Proteomics by mass spectrometry: Approaches, advances, and applications. *Annual Review of Biomedical Engineering* 11, 49–79.
- Zarei, M., Sprenger, A., Metzger, F., Gretzmeier, C., Dengjel, J., 2011. Comparison of ERLIC-TiO(2), HILIC-TiO(2), and SCX-TiO(2) for global phosphoproteomics approaches. *Journal of Proteome Research* 10, 3474–3483.
- Zhai, B., *et al.*, 2008. Phosphoproteome analysis of *Drosophila melanogaster* embryos. *Journal of Proteome Research* 7 (4), 1675–1682.
- Zhang, T., *et al.*, 2012. Acetylation negatively regulates glycogen phosphorylase by recruiting protein phosphatase 1. *Cell Metabolism* 15 (1), 75–87.
- Zhang, Y., *et al.*, 2005. Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Molecular & Cellular Proteomics* 4 (9), 1240–1250.
- Zhou, H., *et al.*, 2008. Specific phosphopeptide enrichment with immobilized titanium ion affinity chromatography adsorbent for phosphoproteome analysis. *Journal of Proteome Research* 7 (9), 3957–3967.
- Zhou, H., Watts, J.D., Aebersold, R., 2001. A systematic approach to the analysis of protein phosphorylation. *Nature Biotechnology* 19 (4), 375–378.
- Zhu, H., *et al.*, 2002. Amino acid residue specific stable isotope labeling for quantitative proteomics. *Rapid Communications in Mass Spectrometry* 16 (22), 2115–2123.

Interactomes – Scaffolds of Cellular Systems

K Luck, Dana-Farber Cancer Institute, Boston, MA, USA and Harvard Medical School, Boston, MA, USA

N Jaikhani, Dana-Farber Cancer Institute, Boston, MA, USA; Harvard Medical School, Boston, MA, USA; and Massachusetts Institute of Technology, Cambridge, MA, USA

ME Cusick, T Rolland, MA Calderwood, B Charleaux, and M Vidal, Dana-Farber Cancer Institute, Boston, MA, USA and Harvard Medical School, Boston, MA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

Enunciated by Crick in 1958, the ‘central dogma of molecular biology’ formulated for the first time how information encoded by molecular alphabets is passed on from genes over messenger RNA to proteins in a defined sequential manner (Crick, 1958). Although this model has since been challenged and nuanced many times, notably by the discovery of RNA viruses, functional noncoding RNA molecules, alternative splicing, posttranslational modifications, and others, this central dogma of molecular biology has opened unprecedented avenues for modern molecular biology. The concept of genes and their products (proteins or RNA) does indubitably play a central role in our understanding of life and has spurred genome sequencing projects for human and model organisms to comprehensively catalog these functional elements. However, it appears that even knowing all functional elements of a cell is insufficient in understanding the bases of cellular life. What is it that we are still missing?

One of the most highly studied human proteins is p53. Its function as a tumor suppressor has been well established but can only be understood in the context of numerous interactions that p53 forms with other proteins and DNA elements (Oliner, 1993). Through binding to promoter regions of various genes involved in growth arrest or apoptosis, p53 can activate their transcription in response to DNA damage. This example illustrates that functional elements, such as proteins, do not work in isolation but interact with each other in a coordinated fashion to carry out cellular functions.

As a key player in regulation of cell growth, p53 itself needs to be tightly regulated. One of the genes activated by p53 is Mdm2, an E3 ubiquitin ligase that mediates ubiquitination of p53 leading to p53 degradation by the proteasome (Fuchs *et al.*, 1998). The interaction of p53 with the promoter of the Mdm2 gene and the interaction between the proteins p53 and Mdm2 form a regulatory mechanism referred to as composite negative feedback loop (Alon, 2007). Together, these two interactions and the ones around them form a system that can lead to different cellular states. Yet despite two decades of intense studies all aspects of this feedback loop are still not perfectly understood reflecting the complexity of the system and of all of its inputs.

Many different systems such as the p53-Mdm2 negative feedback loop, also called network motifs, have been identified and found to form higher order complex regulatory circuits that mediate cellular functions (Alon, 2007). The study of network motifs and their interconnectivity has provided the framework for understanding cells as systems whose scaffold consists of a web of macromolecular interactions, in their entirety called the interactome. Accordingly, interactomes are

essential for approaching the cell as a system and their comprehensive mapping represents the next logical step following genome sequencing (Carvunis *et al.*, 2012). The last decade has seen increasingly successful experimental and computational efforts in providing interactome maps at breadth and depth (Wodak *et al.*, 2013). Though incomplete, such maps, if systematically generated, are representative subsets of the interactome and as such have substantially increased understanding of cellular organization and functions (Vidal *et al.*, 2011). Existing maps contain biophysical or functional interactions. Biophysical interactions comprise direct contacts and co-complex associations of macromolecules whereas functional interactions reflect that two genes or proteins are involved in a common biological process.

The cell coordinates its interactome through various regulatory layers including differential expression and subcellular localization of macromolecules as well as activating and inhibiting mechanisms of macromolecular functions. Current proteome-wide interaction mapping techniques are unable to provide information at such resolution and rather produce sets of possible interactions that then need to be translated into functional relationships and coordinated cellular processes. Key to overcoming these limitations is integration of biophysical with functional interactome maps as well as co-localization and co-expression data to derive functionally meaningful representations of interactome networks (Figure 1; Ge *et al.*, 2003).

Proteins are considered the master connectors and regulators of cellular function given their ability to interact not only with other proteins but also a variety of other types of molecules including DNA, RNA, lipids, glycans, and small molecules. In this article, we provide an overview on different concepts of protein-centric biophysical interactome mapping and will discuss selected studies demonstrating how interactomes modeled as networks and integrated with other systematic functional datasets spurred understanding of cellular systems and how network perturbation can lead to disease.

Methods to Map Interactomes

Methods to map biophysical interactomes can be classified into four main approaches based on their specific aims and outcomes. These are defined as binary mapping, capturing, curation, and prediction (Figure 2).

Binary mapping of interactomes consists of systematic interrogation of pairs of biological molecules for interaction, thereby providing a map of possible binary interactions (Figure 2(a)). This approach relies on *in vitro* or cellular

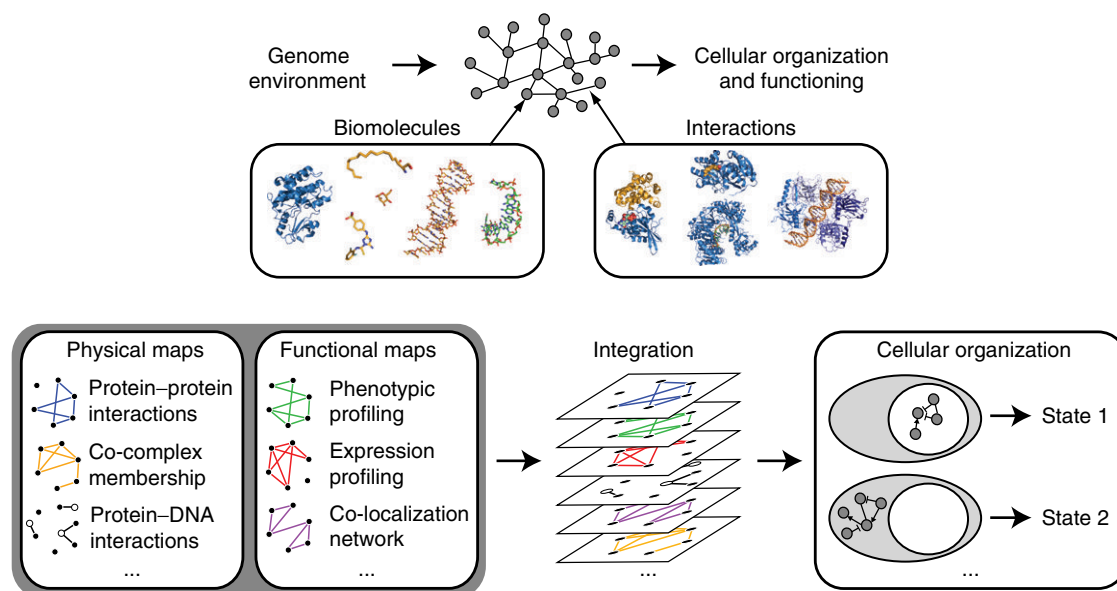


Figure 1 Interactomes underlie cellular life. Top: Biological functions are mediated through a complex web of interactions between biological molecules. Bottom: Integration of interaction maps from different sources allows deriving more functionally relevant interaction networks representative of different cellular states.

exogenous systems and requires technologies for the high-throughput synthesis of molecular components aimed at interrogating interactions, as well as an efficient and reliable readout to score interactions that do occur.

Capturing approaches act to extract molecular components from a cell together with their associated partners, providing a snapshot of macromolecular assemblies that exist in the interrogated cellular state (Figure 2(b)). Capturing methods rely on techniques for capturing specific molecular components from cell extracts, coupled to analytical methods for the identification of all associated molecules.

Published scientific literature represents a large repository for experimentally derived molecular interactions. The curation approach aims at extracting interaction data from scientific publications to build interactome maps (Figure 2(c)). Given the unstructured text of typical biomedical publications, the curation approach largely relies on manual extraction of interaction information.

Computational approaches provide interactome maps based on prediction, aiming to bypass expensive and time-consuming laboratory experiments. Genomic context, molecular sequences, and secondary and tertiary structure of interacting biomolecules are computationally mined to identify patterns indicative of interactions that can be used to predict new interactions (Figure 2(d)). Molecular dynamics represents one of the computationally most expensive forms of this approach, where simulations search for lowest energy states of a given molecular system, results of which indicate the strength and likelihood of a tested interaction to take place.

Genome and proteome-wide studies need, as much as small-scale experiments, rigorous assessment of both false-positive and false-negative rates as well as a clear understanding of the type of data generated and its properties. Estimation of the size of a given interactome in question allows assessment of the completeness of current interactome

maps and defines future goals. A framework for binary interactome mapping has been proposed, whose principles can be applied to any genome or proteome-scale experiment (Venkatesan *et al.*, 2009). In this framework, the quality and coverage of a dataset is assessed by four parameters: screening completeness (the fraction of the full space screened), assay sensitivity (fraction of interactions detectable by the assay), sampling sensitivity (fraction of detectable interactions detected in one screen), and precision (fraction of identified interactions that are true positives). Core to this framework is accounting for the space screened, the use of a positive and random reference set (PRS and RRS) of interactions serving as gold standard and negative control sets, respectively, and validation of interactions with orthogonal assays (Venkatesan *et al.*, 2009; Yu *et al.*, 2008).

Different implementations of the four interactome mapping approaches have been developed to accommodate characteristics specific to the various groups of biological molecules (proteins, DNA, RNA, lipids, glycans, and small molecules) with which proteins interact. In the following sections, we will provide an overview on widely used techniques for protein-centric interactome mapping and discuss biases, quality, and completeness issues of generated maps.

Mapping Protein–Protein Interactions

Protein–protein interactions (PPIs) constitute a highly investigated type of biomolecular interactions. Several orthogonal proteome-scalable methodologies have been developed to interrogate pairs of proteins for binary interactions.

Classical yeast two-hybrid (Y2H) (Fields and Song, 1989) is based on the functional reconstitution of the Gal4 transcription factor (TF), whose DNA-binding domain and activation domain are fused to bait and prey proteins, respectively. These fusion constructs are expressed in yeast cells, and bait

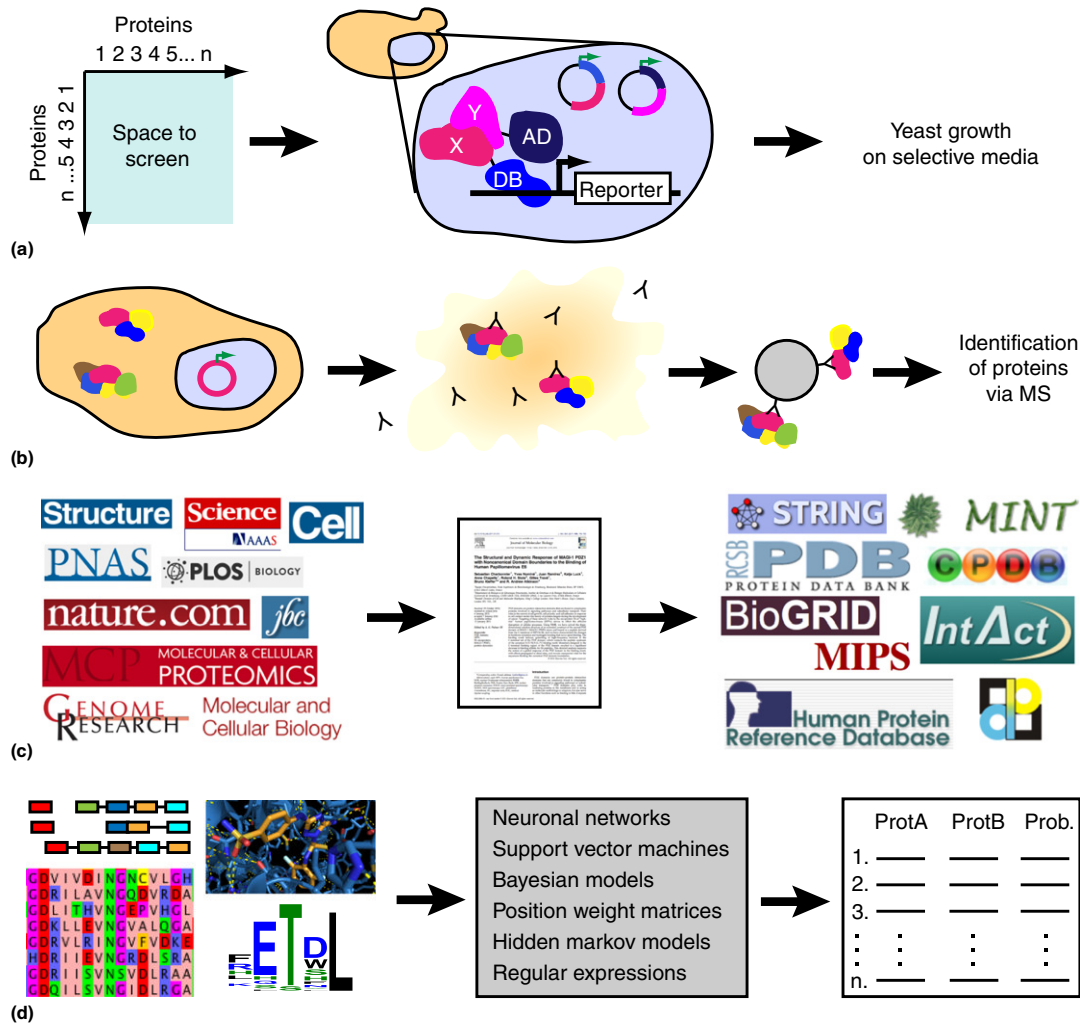


Figure 2 Four approaches to map interactomes. Examples have been drawn from PPI mapping. (a) Binary mapping (Y2H). Pairs of bait and prey from a defined search space are systematically expressed in cells as fusion constructs with reporter protein fragments. Upon interaction of bait and prey, the two fragments of the reporter protein are brought in close proximity reconstituting reporter protein function. Resulting signals indicate a direct and binary interaction between bait and prey. (b) Capturing (AP-MS). A bait, usually expressed from a plasmid, forms complexes with its interaction partners, the preys. After cell lysis baits together with preys can be captured and purified using specific antibodies or coated beads against the bait or against a fused epitope tag. Bait and prey in purified complexes are identified by MS. (c) Curation. Published literature is manually curated to extract information on protein-mediated interactions, which are imported and maintained in public databases. (d) Prediction. Conservation of sequence, patterns of recognition, genetic context, functional relationships, co-expression patterns, structural data, and many more types of data can be mined to predict protein-mediated interactions. Information provided from these resources is used as input into various learning algorithms, which after training, produce for a given pair of proteins a probability or predicted binding affinity indicating the likeliness or strength of a potential interaction between those.

and prey proteins that interact reconstitute a functional TF that activates a reporter gene used for readout. In classical Y2H, bait and prey need to interact within the nucleus and non-yeast bait and prey proteins need to remain functional in a foreign cellular environment, representing limitations of this methodology. Advantages of this well-established approach are the independence from sophisticated equipment and low costs, enabling its scalability.

Protein complementation assays (PCA) using fragments of ubiquitin (Johnsson and Varshavsky, 1994), dihydrofolate reductase, an essential metabolic enzyme, and fluorescent or luminescent proteins have been developed to detect binary interactions in yeast and mammalian cells. Signal readouts can

be cellular growth or emitted light, depending on the reporter protein implemented. Unlike Y2H, PCA is not restricted to nuclear localization of bait and prey, but some implementations seem to indicate proximity of bait and prey rather than a direct interaction (Stylen *et al.*, 2012).

Mammalian protein-protein interaction trap (MAPPIT) and luminescence-based mammalian interactome mapping (LUMIER) are binary mapping assays that function in mammalian cells. In MAPPIT, bait and prey are fused to members of the JAK/STAT signaling pathway that upon interaction at the inner surface of the plasma membrane mediate activation of the TF signal transducer and activator of transcription 3 (STAT3), which activates a reporter gene

(Eyckerman *et al.*, 2001). LUMIER is a pulldown-based assay where preys tagged with the immunoaffinity FLAG peptide epitope are pulled down from cell extracts and interactions with bait proteins fused to luciferase are quantified by measuring luminescence signal intensity (Barrios-Rodiles *et al.*, 2005).

Nucleic acid programmable protein array (NAPPA) is an *in vitro* method where FLAG-tagged bait and prey proteins are synthesized in an *in vitro* transcription–translation system. Upon synthesis bait proteins are captured and non-covalently attached to a membrane. The co-synthesized interacting prey can be detected via its FLAG immunoaffinity tag (Ramachandran *et al.*, 2004).

None of these assays can formally exclude the possibility that endogenous yeast or mammalian proteins mediate the observed interactions between bait and prey. Interaction datasets derived from binary mapping methods show negligible bias with respect to endogenous protein expression levels, but the physiological context under which detected interactions may occur remains uncertain. Fundamental variations in methodological design define different chemical environments under which pairs of proteins are tested for interaction leading to different subsets of interactions detected. Combining interaction maps derived from different methods thus increases overall coverage (Braun *et al.*, 2009).

Proteome-wide binary interactome mapping, mostly done with Y2H, has been achieved for several bacterial species, *Saccharomyces cerevisiae* (yeast), *Caenorhabditis elegans* (worm), *Drosophila melanogaster* (fly), *Arabidopsis thaliana* (plant model organism), and *Homo sapiens* (human) (reviewed in Ryan *et al.*, 2013; Carvunis *et al.*, 2012; Vidal *et al.*, 2011; Wodak *et al.*, 2013). The number of interactions identified in these screens ranges from several hundred for the earliest maps to ~14 000 interactions for human involving ~4300 proteins (Rolland *et al.*, 2014). The yeast protein–protein interactome has been mapped at proteome scale several times, yet the union of three high-quality Y2H-derived binary interactome maps from yeast (only) represents about 15% (2930 PPIs) of its total estimated size ($\sim 18\,000 \pm 4500$ PPIs) (Yu *et al.*, 2008).

Affinity purification coupled to mass spectrometry (AP-MS) represents the most widely used implementation of capturing methods for PPIs. AP-MS captures protein complexes from cells expressing bait proteins either directly or via fused immunoaffinity tags. After cell lysis, baits are purified together with associated prey proteins, which are subsequently identified by MS. Comprehensive mapping by AP-MS of protein complexes of *E. coli*, yeast, and fly have all been reported (reviewed in Ryan *et al.*, 2013; Carvunis *et al.*, 2012; Vidal *et al.*, 2011; Wodak *et al.*, 2013). Each report identified up to 600 distinct protein complexes and up to 5000 distinct proteins in total. The union of three tandem affinity purification–mass spectrometry (TAP-MS) maps of yeast PPIs represents about 40% of the yeast interactome (Collins *et al.*, 2007). Mammalian protein complexes derived by AP-MS have been reported at the scale of individual biological processes but mapping at proteome scale has faced difficulties expressing tagged bait proteins at high-throughput in mammalian cells (Gingras *et al.*, 2005). However, mammalian proteome-scale datasets are now emerging in public repositories (Huttlin *et al.*, 2014,

pre-release on thebiogrid.org). Many proteins function as part of multi-subunit complexes, and AP-MS data provide insights into protein complex composition at a given cellular state. AP-MS lacks the resolution to determine, if all prey proteins pulled out by a bait protein occur in the same or in different complexes, and to dissect which proteins directly bind each other within a complex. AP-MS data show biases toward higher affinity interactions and contrary to binary methods, also toward more highly expressed proteins (Wodak *et al.*, 2013).

Another recent approach for proteome-scale protein complex capture involves direct chromatographic separation of cell extracts. Proteins co-migrating together in the same biochemical fractions are identified via MS and are presumed to be part of the same protein complex (Havugimana *et al.*, 2012).

Given the different properties of binary PPIs versus protein complex associations, it is important to discriminate between interactomic data derived from binary and capturing techniques (Yu *et al.*, 2008). Overlaps of interactome maps derived from Y2H and AP-MS or even between interactome maps derived from the same method seem small, but when controlled for space and assay sensitivities, are highly significant (Yu *et al.*, 2008; Wodak *et al.*, 2013). The complementarity of TAP-MS and Y2H protein interaction datasets has been exploited to dissect potential binding interfaces between pairs of proteins within protein complexes in *E. coli* (Rajagopala *et al.*, 2014).

Due to experimental limitations, most PPI mapping efforts necessarily disregard the existence of alternatively spliced isoforms of a gene and splice variants of the same genes are generally not distinguished in current PPI maps. Alternative splicing can selectively remove or add binding sites to isoforms, suggesting major differences in interaction partners may exist between isoforms from the same gene (Ellis *et al.*, 2012).

An important resource of PPI information comes from literature curation. Numerous public databases store curated PPIs, including direct and indirect, binary and complex, and biophysical and functional relationships (Orchard, 2012). Recent initiatives aimed at facilitating the time-consuming and ambiguous task of literature curation, have defined guidelines for the requisite information for a molecular interaction to be reported in a publication (MIMIX (minimum information about a molecular interaction experiment); Orchard *et al.*, 2007), standardized the file format in which interactions are curated (HUPO-PSI-MI format (Human Proteome Organization Proteomics Standards Initiative Molecular Interaction format); Orchard and Hermjakob, 2008, MITAB format (molecular interaction tabular format); Kerrien *et al.*, 2007), and unified access to curated interaction data (Aranda *et al.*, 2011).

Numerous strategies have been developed to predict protein interactions, yet the performance of those applicable at proteome scale seems still limited (Rolland *et al.*, 2014). To infer PPIs, predictive methods incorporate phylogenetic information captured in sequence alignments (e.g., used to infer interologs, conserved interactions across species), genetic neighborhood information, sequence composition, and patterns or motifs in primary sequences, three-dimensional (3D) structural information and molecular dynamics (Tuncbag *et al.*, 2011; Franceschini *et al.*, 2013). Most predictive methods only concentrate on a particular type of protein interaction,

those mediated by short linear motifs (Weatheritt *et al.*, 2012) or those between globular domain interfaces (Mosca *et al.*, 2014). A recent proteome-wide domain–domain interaction prediction utilized the growing repertoire of 3D structures of interacting proteins (Zhang *et al.*, 2012). Structural homology modeling allowed assignment of a structural template to about two-thirds of human proteins without an existing structure but with at least one globular domain. The resulting homology models were then used to search for compatible interaction interfaces at a proteome-wide scale.

A comparison of four proteome-scale maps of human PPIs each derived from a different mapping approach: binary (Rolland *et al.*, 2014), capturing (Havugimana *et al.*, 2012), curation (Rolland *et al.*, 2014), and prediction (Zhang *et al.*, 2012) revealed interesting insights into biases displayed by each of these maps with respect to a variety of gene properties investigated. Notably, the literature and capturing map were significantly biased toward highly studied genes and genes of high-expression levels. The prediction map was depleted of proteins without structured domains and also suffered from inspection biases. All four maps performed poorly in identifying interactions involving membrane proteins. Out of these four maps, the binary map performed best with respect to homogeneously covering the human PPI space. Nevertheless, each map adds unique and valuable information to the human protein–protein interactome and contributes to increasing its coverage.

Mapping Protein–DNA Interactions

Two methods currently dominate the field of binary protein–DNA interaction (PDI) mapping at proteome-scale, protein-binding microarrays (PBM) and yeast one-hybrid (Y1H). In PBM double-stranded DNA is attached on an array and probed for interaction with purified proteins, TFs in particular. Current microarrays can contain every possible sequence variant of a 10-mer oligonucleotide, enabling derivation of DNA-binding specificities for a given TF. PBMs have been applied to homeodomain-containing TFs of mouse and to most yeast TFs (Dey *et al.*, 2012). The results characterized *in vitro* DNA-binding specificity ranges as well as mechanisms of DNA sequence recognition by proteins. PBM data can be transformed into DNA-binding profiles that are used to query genome sequences for prediction of TF-specific target sites and for generation of protein–DNA interactome maps. PBM technologies are protein-centric, i.e., the DNA-binding capabilities of a given protein are interrogated.

The principles of Y1H, another binary gene-centric assay, are similar to Y2H in that the yeast cells are used to interrogate exogenous bait and prey molecules for binary interactions. The DNA bait sequence, usually a presumptive *in vivo* promoter region, is placed upstream of a reporter gene. Protein preys fused to an activation domain can activate the reporter gene upon binding the DNA bait. In contrast to PBMs, Y1H data can directly be used to generate interactome maps. Y1H can screen large DNA fragments, entire promoter sequences hundreds or thousands of nucleotides long, as compared to microarray technologies which are limited to small oligonucleotides (Walhout, 2011). Overall sensitivities of Y1H have been

estimated to lie between 33 and 77%, depending on the actual form of the assay used, with a sampling sensitivity of a highly automated Y1H assay as high as 89% (Reece-Hoyes *et al.*, 2011). Estimation of false-positive rates is hampered by the difficulties in generating negative reference sets. Y1H has been applied at biological process and proteome scale to worm, *A. thaliana*, and to human, involving up to 1000 protein preys (70% of the predicted human TFs) and up to 200 DNA baits, generating interactome maps containing several hundred PDIs (Gaudinier *et al.*, 2011; Reece-Hoyes *et al.*, 2011).

Chromatin immunoprecipitation (ChIP) is a protein-centric capturing technique that identifies PDIs *in vivo*. The protein baits are crosslinked to the DNA they bind via a chemical cross-linking agent, usually formaldehyde, followed by a fragmentation of the DNA and immunoprecipitation using antibodies specific to the protein bait. The immunoprecipitated DNA molecules can be identified via either microarrays (ChIP-chip) or sequencing (ChIP-seq). ChIP is the method of choice when *in vivo* DNA-binding sites need to be identified for selected TF (Walhout, 2011). ChIP has been applied at proteome scale to map genomic binding regions for hundreds of yeast and human TFs (Dey *et al.*, 2012). ChIP investigations across five human cell lines as part of the ENCODE project generated cell-type specific PDI maps (ENCODE Project Consortium *et al.*, 2012). The power of ChIP is limited by its requirement for antibodies against specific TFs, restricting ENCODE data to fewer than 10% of the predicted human TFs. Weakly or non-expressed endogenous TFs can only be mapped with ChIP upon cellular transformation with plasmids, which express a TF of interest. Y1H is more amenable than ChIP to proteome-scale mapping because of its independence of antibodies and endogenous expression of TFs (Walhout, 2011). Due to the cross-linking, the washing steps in ChIP can be stronger than in AP-MS, removing adventitiously associated proteins. Protein–protein cross-linking can occur, albeit at a much lower rate than protein–DNA cross-linking, leading to possible isolation of protein–DNA complexes where TFs are associated with DNA elements via a bridging protein. Bridging may also occur in Y1H to mediate observed interactions between DNA bait and protein prey. Contrary to ChIP, Y1H and PBM generally miss interactions for TFs that function as heterodimeric complexes or which require cofactors.

Unlike curated PPI literature datasets, which also include information from small-scale studies, curated PDI datasets are mostly derived from proteome-scale or structural studies (Dey *et al.*, 2012).

Two main types of data are used to infer PDIs, DNA-binding profiles such as those derived from PBM studies (Weirauch *et al.*, 2013), or gene expression profiles based on co-expression of genes over a range of different cellular conditions (Margolin *et al.*, 2006). Integration of gene expression data with PPI data, DNA-binding profiles, and text-mining results can substantially increase the functional relevance of predicted PDI maps (Olsen *et al.*, 2014).

Mapping Protein–RNA Interactions

RNAcompete is an *in vitro* method that incorporates aspects of both capturing and binary interaction technologies. A protein of interest is incubated with a large pool of RNAs, each up to

41 nucleotides long, covering all possible sequence combinations of 9 nucleotides. After incubation, the protein gets extracted with an affinity tag and bound RNAs are identified on microarrays. This method has been applied to derive RNA-binding profiles for more than 200 RNA-binding domains (RBD) from 20 different eukaryotic species (Ray *et al.*, 2013). Binding profiles have consequently been used to predict binding sites of these RBDs in 5' and 3' untranslated regions as well as in alternative exons and flanking introns of messenger RNA (mRNAs). Integration with mRNA expression and alternative splicing data-enabled prediction of functions for these RBDs.

An approach similar to ChIP has been developed and applied at proteome scale to capture protein–RNA interactions (PRIs), termed cross-linking and immunoprecipitation (CLIP). Cross-linking between protein and bound RNA molecules is achieved via ultraviolet (UV) light at specific wave lengths. After immunoprecipitation, crosslinked RNA molecules can either be reverse transcribed, amplified, and sequenced (CLIP-seq) or identified using microarrays (CLIP-chip). A refinement of the cross-linking technique is PAR-CLIP (photoactivatable ribonucleoside CLIP), where cells are grown on media containing 4-thiouridine (4SU) instead of uracil to incorporate 4SU into RNA molecules. Upon reverse transcription of captured RNA molecules crosslinked 4SU introduces a thymine to cytosine mutation allowing precise determination of protein-binding sites. CLIP and its variations have been applied transcriptome-wide to study binding of selected RNA-binding proteins to mRNAs and microRNAs (miRNAs) (reviewed in Konig *et al.*, 2011). CLIP has also been successfully combined with MS to identify hundreds of novel mRNA-binding proteins via capturing of polyadenylated RNAs with oligo-thymine beads (reviewed in Ascano *et al.*, 2013). Given accumulating evidence indicating major roles in cellular regulation for noncoding RNAs, there is a clear need to extend PRI mapping efforts from mRNA toward noncoding RNAs.

A few efforts have been made toward PRI curation from the literature with a focus on PRI curation and analysis from published structures in the Protein Data Bank (PDB) (Fujimori *et al.*, 2012).

RNA-binding profiles for proteins can be derived from methods such as RNAcompete and used to predict PRIs. Other approaches to PRI prediction are low-throughput using protein structural and primary sequence information (Puton *et al.*, 2012).

Mapping Protein–Lipid and Protein–Glycan Interactions

There is no genetic code and unique synthesis machinery for lipids and glycans as exists for DNA, RNA, and protein synthesis. The lipidome, as is the glycome, is synthesized by a diverse set of enzymes using different educts as starting material. This complexity in compound synthesis explains why methods to study interactions of these molecules with proteins at scales similar to those for proteins and nucleic acids are still lacking.

Lipids are anchored in cellular membrane structures, making the capturing of PLIs from living cells at proteome scale intractable. Detection of PLIs mostly relies on binary mapping technologies using protein or lipid arrays where

bound targets are identified using either tags or MS. A promising attempt toward systematic PLI mapping used lipid arrays (Gallego *et al.*, 2010) to assay binding of 56 lipids to 172 yeast proteins predicted to be involved in lipid metabolism or containing lipid-binding domains. In total 530 protein–lipid interactions (PLIs) were detected, including novel lipid-binding domains. At lower scale, lipid-binding potentials of specific classes of domains (e.g., PDZ domains) have been studied across a variety of lipids by surface plasmon resonance (SPR) (Chen *et al.*, 2012). The resulting quantitative binding information together with structural data has been used to predict domain–lipid interactions.

These binary approaches used isolated lipids attached to arrays or chips. Under physiological conditions lipids occur in membranous lipid layers, the actual interaction surface encountered by lipid-binding proteins. To mimic these conditions, a strategy has been developed to attach entire liposomes of defined composition onto arrays and to probe them with purified proteins for interaction using fluorescence microscopy for readout (Saliba *et al.*, 2014).

PGI mapping at large scale is even more challenging than PLI mapping. The complex branching structure of glycans, based on highly similar monosaccharide building blocks, hinders unambiguous identification by MS (Rakus and Mahal, 2011). A functional glycomics consortium generated a publicly available glycan microarray containing more than 500 probes to assay binding specificities of glycan-binding proteins (Smith *et al.*, 2010).

Mapping Protein–Small Molecule Interactions

A large fraction of protein–small molecule interactions (PSMIs) are enzymatic (biochemical), characterized by chemical modification of the small molecule via its interacting protein, the enzyme. Enzymatic interactions are short-lived, as immediately upon modification the resulting product dissociates from the protein. Interaction mapping approaches thus need to be different for PSMIs that are enzymatic versus those that are not. Because of their instability, PSMIs that are enzymatic can best be indirectly measured by solution-based *in vitro* enzyme assays where appearance or loss of metabolites is detected. The throughput of such approaches is limited. Interactome maps of enzymatic PSMIs have mainly been built from literature-curated datasets, which for this type of interaction can reach high coverage (Kanehisa *et al.*, 2014).

Nonenzymatic PSMIs, such as allosteric, inhibitory, cofactor, or ligand interactions, are generally longer-lasting depending on the actual reaction constants. These interactions are measurable by binary mapping and capturing methods whose applications have been predominantly drug development. The aims are to either identify for a given protein (or protein class) a potent inhibitor (drug) of the activity of that protein, or alternatively to identify all protein targets of a given small molecule (drug) to understand mechanism of action and side effects referred to as protein- and drug-centric interaction mapping, respectively. A combination of computational and experimental binary testing has been used to address protein-centric interaction mapping. Via computational docking the potential scaffold or class of small molecules with biological activity toward the given protein or class of proteins

can be determined. In consecutive experimental screens, derivatives of the identified small molecule class are screened versus targets to identify the high-affinity binders. The most frequently screened drug target classes are G-protein coupled receptors, nuclear receptors, ion channels, kinases, phosphatases, and proteases. Experimental screens can be carried out using arrays and MS (Zinn *et al.*, 2012), or methods of lower throughput (nuclear magnetic resonance and SPR) which can return quantitative binding information.

Capturing methods, also referred to as chemoproteomic approaches, have powerfully addressed drug-centric interaction mapping. A given drug is immobilized on a matrix and interrogated with a cell extract. After incubation and extensive washing, bound proteins can be eluted and identified by MS (Colinge *et al.*, 2012).

An indirect mapping approach has been proposed to map allosteric interactions regulating the switch from gluconeogenesis to glycolysis in *E. coli* (Link *et al.*, 2013). Changes in metabolite concentrations upon frequent switches between pyruvate and glucose or fructose in bacterial media have been determined and fit to models of metabolic pathways. Some 126 predicted allosteric interactions were successfully validated experimentally.

The Role of Genome Sequencing for Interactome Mapping

Sequenced genomes are essential for systematic and proteome-wide interactome mapping by binary or capturing approaches. Assay design for both approaches depends on preexisting knowledge of the molecular components that constitute the interactome in question. The completeness of such lists of molecular components including protein, DNA, and RNA, entirely relies on the availability of sequenced genomes (Hood and Rowen, 2013). Recent improvements in mass spectrometry and DNA sequencing technologies have enabled and accelerated proteome-scale interactome mapping with binary and capturing approaches. DNA or RNA reads resulting from sequencing of preys need to be aligned to a reference genome for identifying the corresponding genes. MS spectra of peptides need to be compared to theoretical spectra derived from all possible peptides of a reference proteome to identify the corresponding proteins. Immunoaffinity tags can only be fused to known baits and only for known baits can specific antibodies be derived. Binary mapping technologies depend on open reading frame (ORF) collections of known genes for the synthesis of proteins. With the availability of genome sequences at increasing depth and accuracy, along with increasing power and throughput of interaction mapping approaches, a human reference interactome project in analogy to the human reference genome project is becoming a possibility (Vidal and Fields, 2014).

Applications of Interactome Maps

Protein interactome maps are readily represented as networks, where biomolecules are the nodes of the network and the interactions between them are the links or edges (Barabási and Oltvai, 2004; Figure 3). Gene regulatory networks can be built

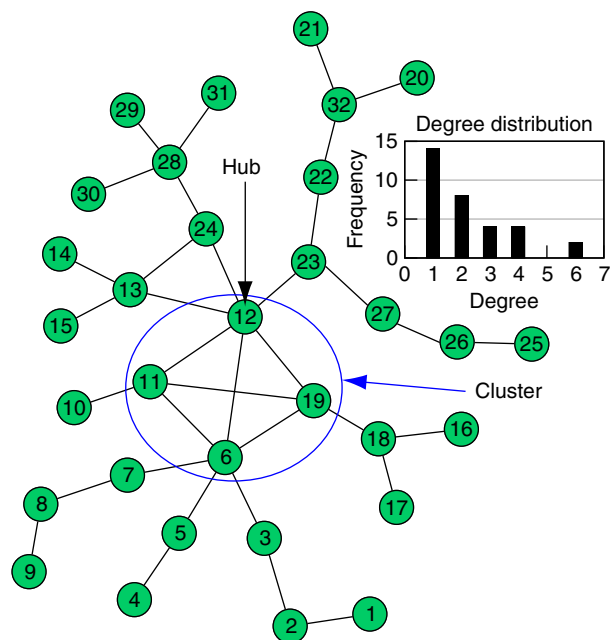


Figure 3 Network biology. The small network illustrates some of the properties of protein-mediated interaction networks. Nodes represent proteins and their partners, edges interactions between them. This undirected network is scale-free as indicated by the inset histogram of degree distribution, meaning that many nodes are only engaged in one or two interactions whereas very few nodes are hubs (e.g., node 12). Subsets of nodes can be highly connected forming clusters indicating function in a common biological process (blue circle).

from TFs mapped to DNA-binding sites, linking the genes encoding for the TFs with the genes they positively or negatively regulate. In metabolic networks, nodes are metabolites and edges between them represent reactions catalyzed by known enzymes. PPI networks are mostly undirected, where the direction of interaction cannot be stipulated. Gene regulatory and metabolic networks represent directed networks (Vidal *et al.*, 2011). Once in network format, the powerful tools and analyses of graph theory can be applied to assess structural properties of cellular interaction networks and to derive principles of cellular organization (Barabási and Oltvai, 2004).

Interaction Network Structure Reflects Biological Function

The most elementary characteristic of a node is its degree. In nearly all interactome networks the degree distribution follows a power law (Figure 3) and such networks are known as scale-free networks, meaning that most nodes have a low degree whereas a small number of nodes have an unusually high degree compared to the mean degree; such high-degree nodes are called hubs (Barabási and Oltvai, 2004). This organization contrasts to random networks, which follow the Poisson or Gaussian degree distribution, where the degree of most nodes is close to the mean node degree. Scale-free networks display robustness, showing insensitivity toward random attacks, since the random removal of a node is more likely to remove a low-degree node than a hub, but extreme sensitivity to attacks

directed at hubs, since hubs are so highly connected (Albert *et al.*, 2000).

A node with high degree is better connected in the network and is likely to be part of many shortest paths between other nodes. A high-degree node therefore plays a more important role in maintaining the network structure and information flow than nodes of low degree. A logical expectation is that hubs in a PPI network are more likely to correspond to essential proteins (Jeong *et al.*, 2001), however, the underlying biological reasons for such relationship and its overall existence have long been debated (Zotenko *et al.*, 2008). Clear correlations between degree and essentiality have been observed using PPI networks derived from literature curation contrary to using systematically generated networks (Zotenko *et al.*, 2008; Yu *et al.*, 2008). Networks obtained from curation are heavily biased toward interactions involving proteins perceived as most biologically important and essential (Rolland *et al.*, 2014). The lower degree of nonessential proteins may simply result from those proteins being understudied. Using a systematic binary map of yeast PPIs, it has been shown that degree in an interactome network rather correlates with genetic pleiotropy, the number of phenotypes observed as a consequence of gene removal, than essentiality (Yu *et al.*, 2008).

Overlaying transcriptome profiles with a binary interaction network enabled classification of two types of highly connected hub proteins – party hubs and date hubs. Party hubs are co-expressed with most of their interacting partners; while date hubs connect to different partners at different times or contexts (Han *et al.*, 2004). Structural differences between date and party hubs reflect their different network properties. With date hubs different interacting partners may utilize the same surface of interaction at different times, whereas party hubs tend to contain fewer disordered regions and to display more interaction interfaces on their surfaces, as would be expected for proteins with many simultaneous interacting partners (Kim *et al.*, 2008). Integrating date and party hubs with protein subcellular localization, functional, and genetic interaction data identified interaction partners of party hubs as being more co-localized, more often part of the same functional module, and less correlated with genetic interaction data than date hubs (Han *et al.*, 2004).

Molecular interaction networks contain cliques or communities, which are clusters of nodes that are highly connected to each other (Figure 3). Identified clusters of nodes in molecular interaction networks have been termed topological modules. Modularity is a central hallmark of cellular organization. Biomolecules within a topological module usually participate together in a specific cellular function (e.g., translation) often by forming a specific complex (e.g., the ribosome) (Barabási and Oltvai, 2004). Topological modules often correspond to functional modules. Perturbation of any of the molecules forming a functional module can lead to closely related phenotypes or diseases (Goh *et al.*, 2007). In PPI networks, all proteins that upon disruption lead to the same disease can be grouped into disease modules that generally, but not always, overlap with functional and topological modules (Barabási *et al.*, 2011; Vidal *et al.*, 2011).

In summary, the connectivity structure of interaction networks, that is, the topology, leads to definition of biological processes through delineation of modules, and can inform

upon how biological processes are disrupted in disease and how disease therapeutics can be designed and evaluated, as examples in the next section relate.

Interactome Maps Contribute to Our Understanding of Cellular Organization and Human Disease

Integration of data from multiple sources acts to increase data reliability and provides complementary functional evidences, which can help to develop biological hypotheses, both for follow up on individually selected interactions as well as on emerging systems properties (Figure 1; Ge *et al.*, 2003). A first attempt to combine large-scale datasets of different kinds uncovered significant correlations between gene expression and protein interaction data, in that protein interaction clusters tended to significantly overlap with clusters of co-expressed genes (Ge *et al.*, 2001).

Interactome maps and cell communication

Interactome mapping of particular functional modules or biological pathways has identified new regulators and effectors of these modules or pathways, unraveling their molecular complexity and redundancy. Pathway specific interactome mapping has been carried out for proliferative pathways often deregulated in multiple diseases: PI3K-mTOR pathway (Pilot-Storck *et al.*, 2010), the MAPK pathway (Bandyopadhyay *et al.*, 2010), the TGF β pathway (Tewari *et al.*, 2004; Barrios-Rodiles *et al.*, 2005), and the Hippo pathway (Kwon *et al.*, 2013). The PI3K-mTOR Y2H interactome map (Pilot-Storck *et al.*, 2010) identified 67 new interactions between 33 components of the pathway of which several were analyzed for their role in mTOR signaling. Systematic TGF β interactome mapping and integration of the resulting interaction network with genetic perturbation data led to the identification of nine pathway modifiers within a network of 71 interactions among 59 pathway components (Tewari *et al.*, 2004). PPI maps centered on molecular machines have been generated for the yeast mediator complex (Guglielmi *et al.*, 2004) and the human spliceosome (Hegele *et al.*, 2012).

Integration of protein interactome data with phenotypic and expression data furthered understanding of the DNA damage response (Boulton *et al.*, 2002) and of early embryogenesis in worm (Gunsalus *et al.*, 2005). Overlaying PPI data, gene expression data, genetic interaction data, and phenotypic data predicted positive and negative regulators of signaling pathways and protein complexes in fly (Vinayagam *et al.*, 2014).

Interactome maps and human disease

Next generation sequencing technologies produce a wealth of genomics data from genome-wide association studies (GWAS) or cancer genome sequencing initiatives, which catalog mutations implicated in the development of disease. Given the speed and low cost of sequencing, the functional characterization of mutations lags a long way behind their identification. Mutations often cause cellular network perturbations eventually leading to a disease state. Systematically mapping cellular interaction networks and profiling perturbations of interactions caused by mutant alleles (edgotyping) may help

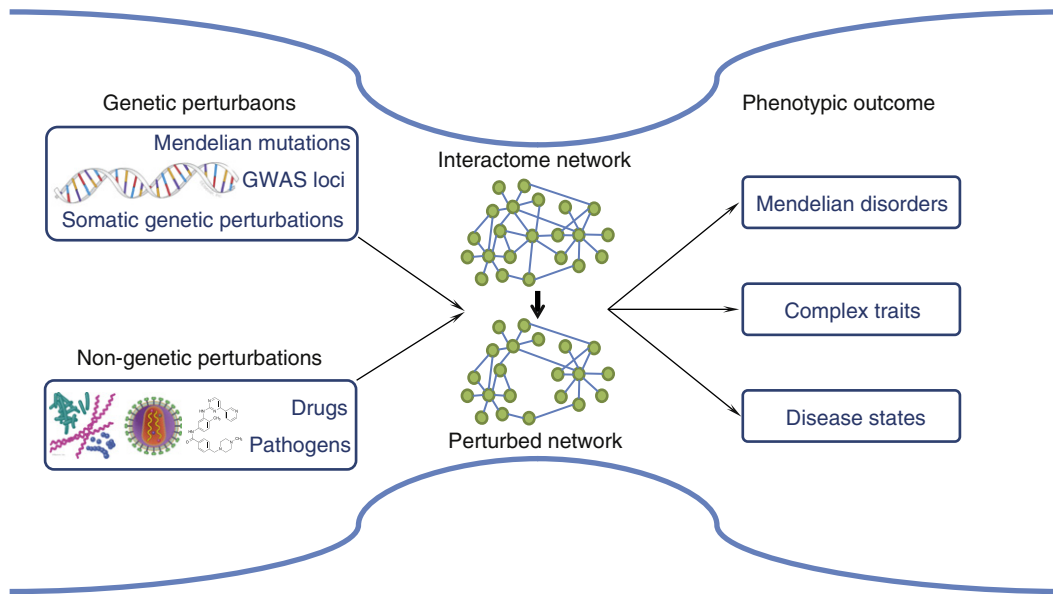


Figure 4 Understanding cellular organization and disease through interactome networks. Perturbations of protein function, either genetic or nongenetic, can disrupt interactome network organization. Deciphering those network disruptions can offer explanations of the phenotypic outcomes that are observed.

enlighten genotype-to-phenotype relationships between mutations and diseases (Figure 4; Sahni *et al.*, 2013). Furthermore, it has been shown that interaction mapping can help discriminating disease-causing from non-disease relevant mutations where the former tend to more often lead to network perturbations (Rolland *et al.*, 2014).

PPI network-based approaches have been applied to human disease, in particular to cancer. Integration of PPI data with gene expression data identified subnetworks predictive of metastasis in breast cancer (Chuang *et al.*, 2007). An analogous integrative approach identified novel breast cancer susceptibility and modifier genes (Pujana *et al.*, 2007). Network analysis of a proteome-wide systematically derived binary PPI dataset revealed that ‘cancer genes’ tend to have a higher degree and to be more highly connected to each other than random sets of genes. These findings have been consequently used to prioritize cancer-candidate genes by identifying interactions in existing PPI networks between genes in loci derived from cancer GWAS and known cancer genes (Rolland *et al.*, 2014). The generation of disease centric interactome maps for human inherited ataxias (Kahle *et al.*, 2011) and autism spectrum disorders (Corominas *et al.*, 2014), has provided fresh insights into disease mechanisms, such as unexpectedly high connectivity between disparate disease proteins for ataxia and autism, suggesting that convergent molecular pathways underlie phenotypes in distinct syndromes.

Interactome maps of host–pathogen interactions

Pathogens target key host proteins in an attempt to control and subvert intracellular pathways to facilitate their survival in host cells (Davey *et al.*, 2011). Systematic mapping of host–pathogen PPIs identifies points of perturbation in host networks and delineates strategies adopted by different pathogens for host invasion (Figure 4).

Systematic mapping of virus–host PPI maps or ‘virhostome’ maps has been undertaken for many specific viruses such as influenza virus, Epstein–Barr virus, hepatitis C virus, herpesviruses, human immunodeficiency virus, or human T-lymphotropic virus (De Chassey *et al.*, 2014). In another proteome-scale effort, 123 viral ORFs derived from several DNA tumor viruses have been systematically mapped for interactions with human proteins using Y2H and TAP-MS (Rozenblatt-Rosen *et al.*, 2012). Integration of this interaction dataset with gene expression data, functional annotations, and TF-binding profiles implicated many members of the Notch signaling pathway in viral pathogenesis and tumorigenesis. Identified human interaction partners of oncogenic viral proteins were enriched for proteins involved in cancer and more specifically for tumor suppressors, pointing to the ability of this systematic approach to aid identification of driver genes in cancer (Rozenblatt-Rosen *et al.*, 2012).

Not only viral but also bacterial pathogens can be investigated by focused interactome mapping. Bacterial virulence factors in the *Mycobacterium tuberculosis* secretome were screened by Y2H for interactions with host proteins (Mehra *et al.*, 2013). This revealed an interaction between the bacterial protein EsxH and the human hepatocyte growth factor-regulated tyrosine kinase substrate (Hgs/Hrs), a component of the endosomal sorting complex, which was shown to play a role in tuberculosis pathogenesis.

Interactome maps, drugs, and drug targets

Mapping of disease-centered interactomes can help identify new proteins causative of disease, and at the same time present potential targets for the development of novel therapies (Vidal *et al.*, 2011). A noteworthy approach to rational drug design integrated metabolic networks with gene expression data derived from cancer samples to generate a model of cancer metabolism that successfully uncovered known metabolic

drug targets in cancer and predicted new candidate drug targets (Folger *et al.*, 2011).

The targets, mechanism of action, and metabolism of many approved and widely applied drugs are still unknown, contributing to poor understanding of drug side effects. Because cells function as a system, the perturbation of one or several drug targets is likely to have system-wide effects and to lead to extensive network remodeling. Drug–protein interaction and PPI mapping combined with measurements of gene expression changes metabolic and phenotypic profiling and integration with patient clinical data and side effect records has been frequently employed to close these gaps in knowledge and to aid in the development of new drugs (Iskar *et al.*, 2012). Chemoproteomic approaches allow identification of many targets of a drug. Searching for functional enrichments within sets of targets using functional annotations of genes can highlight the biological processes or pathways potentially affected by a given drug. Expanding the small set of known drug targets by selected interaction partners derived from PPI networks can increase the performance of such enrichment analyses. Only when enriching the set of targets from the cancer drug bosutinib with neighbors derived from PPI networks was its association with the immune system pathway revealed (Colinge *et al.*, 2012).

Mapping the genes likely affected in a disease via genotyping and the proteins likely to be perturbed by a drug using molecular interaction data, followed by a comparison of the two profiles, allows the efficiency of a drug to be estimated and the identification of potential side effects (Colinge *et al.*, 2012). Comparing drug target profiles with several disease gene profiles can reveal potential for drug repurposing. Such an approach was used to successfully predict the effectiveness of dasatinib to treat lung cancer (Colinge *et al.*, 2012). Combining interactome data with patient specific disease gene profiles may be applied to allow adjustments to drug administration based on a patient's genetic background, the aim of personalized medicine.

Pharmacological networks that illustrate the relationships between known drugs, their targets and disease-causing genes have been generated by integration of drug–protein interaction and PPI data with gene expression and genetic disease annotations (Yildirim *et al.*, 2007). Drugs tend to have multiple targets and proteins are targeted by multiple drugs leading to a highly connected network. Drug targets tend to fall in between peripheral poorly connected nodes of a network and hubs (likely essential proteins), providing a statistical network approach to prioritization of new potential drug targets (Yildirim *et al.*, 2007).

Inhibiting one target has often been ineffective in treating complex diseases such as cancer, autism, or asthma, all arising from perturbation of multifaceted biological processes (Zimmermann *et al.*, 2007). Functional redundancies confer robustness to cellular networks but also result in drug resistances where drug perturbations can be compensated for or buffered by redundant cellular regulatory mechanisms (Hanahan and Weinberg, 2011). Such limitations can be overcome by targeting complex diseases with drug combination therapies (Zimmermann *et al.*, 2007), already applied to cancer (Csermely *et al.*, 2005). Given the complexity of cellular systems and quantity of available drugs, simple systematic

screening for potential drug combinations is unfeasible. Integration of drug-relevant orthogonal datasets including interactome maps has been successfully employed to delve into the mechanism of action of drug combination therapies, to propose new potential drug combinations and to narrow down the search space (Barabási *et al.*, 2011).

Perspectives

Substantial progress has been achieved toward mapping interactomes of different types of interactions and for different species. Integration of these data with orthogonal systematically derived datasets has advanced understanding of how cellular systems are organized and regulated, and has enabled a systems view on disease mechanisms. Current interactome maps suffer from incompleteness and biases that restrict potential downstream analyses and the set of approachable biological questions. Future efforts in interactome mapping will have to overcome these limitations.

Ideally, interactome networks must be expanded to cover not just a reference sequence for each gene, but all proteoforms of all genes (Smith *et al.*, 2013). With continuous development of new technologies and increased throughput this goal may soon be tackled. Incorporating quantitative, spatial, and temporal properties of interactions into networks is crucial toward generating dynamic interaction networks (Ideker and Krogan, 2012; Ge *et al.*, 2003). The ultimate aim is to obtain or infer dynamic interactome maps across different cellular conditions enabling pursuit of a 1000 interactomes project in analogy to the 1000 genomes project (1000 Genomes Project Consortium, 2010).

A more complete picture of the interactome of a cell requires elements from all of the types of protein-mediated interactions described here. There is an urgent need to expand current mapping efforts for understudied types of interactomes to match the pioneering work done for PPI, PDI, and PSMI interaction maps. Combining molecular interaction data of different types into a unified network model will enable the study of information flow from outside the cell, across the membrane, through the cytoplasm and into the nucleus to alter gene expression, altered expression that then remodels the interaction network and shifts the cell from one cellular state to another. Such integrated models promise to better represent the physiological states of a cell.

See also: Dynamic Integration: Dynamics: Dynamics of Protein Kinase Cascades

References

- 1000 Genomes Project Consortium, 2010. A map of human genome variation from population-scale sequencing. *Nature* 467, 1061–1073.
- Albert, R., Jeong, H., Barabási, A.L., 2000. Error and attack tolerance of complex networks. *Nature* 406, 378–382.
- Alon, U., 2007. Network motifs: Theory and experimental approaches. *Nature Reviews Genetics* 8, 450–461.
- Aranda, B., Blankenburg, H., Kerrien, S., *et al.*, 2011. PSICQUIC and PSISCORE: Accessing and scoring molecular interactions. *Nature Methods* 8, 528–529.

- Ascano, M., Gerstberger, S., Tuschl, T., 2013. Multi-disciplinary methods to define RNA-protein interactions and regulatory networks. *Current Opinion in Genetics & Development* 23, 20–28.
- Bandyopadhyay, S., Chiang, C.Y., Srivastava, J., *et al.*, 2010. A human MAP kinase interactome. *Nature Methods* 7, 801–805.
- Barabási, A.L., Gulbahce, N., Loscalzo, J., 2011. Network medicine: A network-based approach to human disease. *Nature Reviews Genetics* 12, 56–68.
- Barabási, A.L., Oltvai, Z.N., 2004. Network biology: Understanding the cell's functional organization. *Nature Reviews Genetics* 5, 101–113.
- Barrios-Rodiles, M., Brown, K.R., Ozdamar, B., *et al.*, 2005. High-throughput mapping of a dynamic signaling network in mammalian cells. *Science* 307, 1621–1625.
- Boulton, S.J., Gartner, A., Reboul, J., *et al.*, 2002. Combined functional genomic maps of the *C. elegans* DNA damage response. *Science* 295, 127–131.
- Braun, P., Tasan, M., Dreze, M., *et al.*, 2009. An experimentally derived confidence score for binary protein-protein interactions. *Nature Methods* 6, 91–97.
- Carvunis, A.R., Roth, F.P., Calderwood, M.A., *et al.*, 2012. Interactome networks. In: Walhout, M., Vidal, M., Dekker, J. (Eds.), *Handbook of Systems Biology – Concepts and Insights*. Waltham, MA: Elsevier – Academic Press, pp. 45–63.
- Chen, Y., Sheng, R., Kallberg, M., *et al.*, 2012. Genome-wide functional annotation of dual-specificity protein- and lipid-binding modules that regulate protein interactions. *Molecular Cell* 46, 226–237.
- Chuang, H.Y., Lee, E., Liu, Y.T., Lee, D., Ideker, T., 2007. Network-based classification of breast cancer metastasis. *Molecular Systems Biology* 3, 140.
- Colinge, J., Rix, U., Bennett, K.L., Superti-Furga, G., 2012. Systems biology analysis of protein-drug interactions. *Proteomics: Clinical Applications* 6, 102–116.
- Collins, S.R., Kemmeren, P., Zhao, X.C., *et al.*, 2007. Toward a comprehensive atlas of the physical interactome of *Saccharomyces cerevisiae*. *Molecular and Cellular Proteomics* 6, 439–450.
- Corominas, R., Yang, X., Lin, G.N., *et al.*, 2014. Protein interaction network of alternatively spliced isoforms from brain links genetic risk factors for autism. *Nature Communications* 5, 3650.
- Crick, F.H., 1958. On protein synthesis. *Symposia of the Society for Experimental Biology* 12, 138–163.
- Csermely, P., Agoston, V., Pongor, S., 2005. The efficiency of multi-target drugs: The network approach might help drug design. *Trends in Pharmacological Sciences* 26, 178–182.
- Davey, N.E., Trave, G., Gibson, T.J., 2011. How viruses hijack cell regulation. *Trends in Biochemical Sciences* 36, 159–169.
- De Chassey, B., Meyniel-Schicklin, L., Vonderscher, J., Andre, P., Lotteau, V., 2014. Virus-host interactomics: New insights and opportunities for antiviral drug discovery. *Genome Medicine* 6, 115.
- Dey, B., Thukral, S., Krishnan, S., *et al.*, 2012. DNA-protein interactions: Methods for detection and analysis. *Molecular and Cellular Biochemistry* 365, 279–299.
- Ellis, J.D., Barrios-Rodiles, M., Colak, R., *et al.*, 2012. Tissue-specific alternative splicing remodels protein-protein interaction networks. *Molecular Cell* 46, 884–892.
- ENCODE Project Consortium, Bernstein, B.E., Birney, E., *et al.*, 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57–74.
- Eyckerman, S., Verhee, A., Der Heyden, J.V., *et al.*, 2001. Design and application of a cytokine-receptor-based interaction trap. *Nature Cell Biology* 3, 1114–1119.
- Fields, S., Song, O., 1989. A novel genetic system to detect protein-protein interactions. *Nature* 340, 245–246.
- Folger, O., Jerby, L., Frezza, C., *et al.*, 2011. Predicting selective drug targets in cancer through metabolic networks. *Molecular Systems Biology* 7, 501.
- Franceschini, A., Szklarczyk, D., Frankild, S., *et al.*, 2013. STRING v9.1: Protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Research* 41, D808–D815.
- Fuchs, S.Y., Adler, V., Buschmann, T., Wu, X., Ronai, Z., 1998. Mdm2 association with p53 targets its ubiquitination. *Oncogene* 17, 2543–2547.
- Fujimori, S., Hino, K., Saito, A., Miyano, S., Miyamoto-Sato, E., 2012. PRD: A protein-RNA interaction database. *Bioinformatics* 8, 729–730.
- Gallego, O., Betts, M.J., Gvozdenovic-Jeremic, J., *et al.*, 2010. A systematic screen for protein-lipid interactions in *Saccharomyces cerevisiae*. *Molecular Systems Biology* 6, 430.
- Gaudinier, A., Zhang, L., Reece-Hoyes, J.S., *et al.*, 2011. Enhanced Y1H assays for *Arabidopsis*. *Nature Methods* 8, 1053–1055.
- Ge, H., Liu, Z., Church, G.M., Vidal, M., 2001. Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. *Nature Genetics* 29, 482–486.
- Ge, H., Walhout, A.J., Vidal, M., 2003. Integrating 'omic' information: A bridge between genomics and systems biology. *Trends in Genetics* 19, 551–560.
- Gingras, A.C., Aebersold, R., Raught, B., 2005. Advances in protein complex analysis using mass spectrometry. *Journal of Physiology* 563, 11–21.
- Goh, K.I., Cusick, M.E., Valle, D., *et al.*, 2007. The human disease network. *Proceedings of the National Academy of Sciences of the United States of America* 104, 8685–8690.
- Guglielmi, B., Van Berkum, N.L., Klapholz, B., *et al.*, 2004. A high resolution protein interaction map of the yeast Mediator complex. *Nucleic Acids Research* 32, 5379–5391.
- Gunsalus, K.C., Ge, H., Schetter, A.J., *et al.*, 2005. Predictive models of molecular machines involved in *Caenorhabditis elegans* early embryogenesis. *Nature* 436, 861–865.
- Han, J.D., Bertin, N., Hao, T., *et al.*, 2004. Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* 430, 88–93.
- Hanahan, D., Weinberg, R.A., 2011. Hallmarks of cancer: The next generation. *Cell* 144, 646–674.
- Havugimana, P.C., Hart, G.T., Nepusz, T., *et al.*, 2012. A census of human soluble protein complexes. *Cell* 150, 1068–1081.
- Hegele, A., Kamburov, A., Grossmann, A., *et al.*, 2012. Dynamic protein-protein interaction wiring of the human spliceosome. *Molecular Cell* 45, 567–580.
- Hood, L., Rowen, L., 2013. The human genome project: Big science transforms biology and medicine. *Genome Medicine* 5, 79.
- Huttlin, E.L., Ting, L., Bruckner, R.J., *et al.*, 2014. High-throughput proteomic mapping of human interaction networks via affinity-purification mass spectrometry. Pre-Publication. Available at: www.thebiogrid.org (accessed 13.05.15).
- Ideker, T., Krogan, N.J., 2012. Differential network biology. *Molecular Systems Biology* 8, 565.
- Iskar, M., Zeller, G., Zhao, X.M., Van Noort, V., Bork, P., 2012. Drug discovery in the age of systems biology: The rise of computational approaches for data integration. *Current Opinion in Biotechnology* 23, 609–616.
- Jeong, H., Mason, S.P., Barabasi, A.L., Oltvai, Z.N., 2001. Lethality and centrality in protein networks. *Nature* 411, 41–42.
- Johnsson, N., Varshavsky, A., 1994. Split ubiquitin as a sensor of protein interactions *in vivo*. *Proceedings of the National Academy of Sciences of the United States of America* 91, 10340–10344.
- Kahle, J.J., Gulbahce, N., Shaw, C.A., *et al.*, 2011. Comparison of an expanded ataxia interactome with patient medical records reveals a relationship between macular degeneration and ataxia. *Human Molecular Genetics* 20, 510–527.
- Kanehisa, M., Goto, S., Sato, Y., *et al.*, 2014. Data, information, knowledge and principle: Back to metabolism in KEGG. *Nucleic Acids Research* 42, D199–D205.
- Kerrien, S., Orchard, S., Montecchi-Palazzi, L., *et al.*, 2007. Broadening the horizon – Level 2.5 of the HUPO-PSI format for molecular interactions. *BMC Biology* 5, 44.
- Kim, P.M., Sboner, A., Xia, Y., Gerstein, M., 2008. The role of disorder in interaction networks: A structural analysis. *Molecular Systems Biology* 4, 179.
- Konig, J., Zarnack, K., Luscombe, N.M., Ule, J., 2011. Protein-RNA interactions: New genomic technologies and perspectives. *Nature Reviews Genetics* 13, 77–83.
- Kwon, Y., Vinayagam, A., Sun, X., *et al.*, 2013. The Hippo signaling pathway interactome. *Science* 342, 737–740.
- Link, H., Kochanowski, K., Sauer, U., 2013. Systematic identification of allosteric protein-metabolite interactions that control enzyme activity *in vivo*. *Nature Biotechnology* 31, 357–361.
- Margolin, A.A., Nemenman, I., Basso, K., *et al.*, 2006. ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* 7 (Suppl. 1), S7.
- Mehra, A., Zahra, A., Thompson, V., *et al.*, 2013. *Mycobacterium tuberculosis* type VII secreted effector EsxH targets host ESCRT to impair trafficking. *PLoS Pathogens* 9, e1003734.
- Mosca, R., Ceol, A., Stein, A., Olivella, R., Aloy, P., 2014. 3did: A catalog of domain-based interactions of known three-dimensional structure. *Nucleic Acids Research* 42, D374–D379.
- Oliner, J.D., 1993. Discerning the function of p53 by examining its molecular interactions. *Bioessays* 15, 703–707.
- Olsen, C., Fleming, K., Prendergast, N., *et al.*, 2014. Inference and validation of predictive gene networks from biomedical literature and gene expression data. *Genomics* 103, 329–336.
- Orchard, S., 2012. Molecular interaction databases. *Proteomics* 12, 1656–1662.
- Orchard, S., Hermjakob, H., 2008. The HUPO proteomics standards initiative – Easing communication and minimizing data loss in a changing world. *Briefings in Bioinformatics* 9, 166–173.
- Orchard, S., Salwinski, L., Kerrien, S., *et al.*, 2007. The minimum information required for reporting a molecular interaction experiment (MIMIx). *Nature Biotechnology* 25, 894–898.

- Pilot-Storck, F., Chopin, E., Rual, J.F., *et al.*, 2010. Interactome mapping of the phosphatidylinositol 3-kinase-mammalian target of rapamycin pathway identifies deformed epidermal autoregulatory factor-1 as a new glycogen synthase kinase-3 interactor. *Molecular and Cellular Proteomics* 9, 1578–1593.
- Pujana, M.A., Han, J.D., Starita, L.M., *et al.*, 2007. Network modeling links breast cancer susceptibility and centrosome dysfunction. *Nature Genetics* 39, 1338–1349.
- Puton, T., Kozłowski, L., Tuszyńska, I., Rother, K., Bujnicki, J.M., 2012. Computational methods for prediction of protein-RNA interactions. *Journal of Structural Biology* 179, 261–268.
- Rajagopala, S.V., Sikorski, P., Kumar, A., *et al.*, 2014. The binary protein-protein interaction landscape of *Escherichia coli*. *Nature Biotechnology* 32, 285–290.
- Rakus, J.F., Mahal, L.K., 2011. New technologies for glycomics analysis: Toward a systematic understanding of the glycome. *Annual Review of Analytical Chemistry* 4, 367–392.
- Ramachandran, N., Hainsworth, E., Bhullar, B., *et al.*, 2004. Self-assembling protein microarrays. *Science* 305, 86–90.
- Ray, D., Kazan, H., Cook, K.B., *et al.*, 2013. A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 499, 172–177.
- Reece-Hoyes, J.S., Barutcu, A.R., Mccord, R.P., *et al.*, 2011. Yeast one-hybrid assays for gene-centered human gene regulatory network mapping. *Nature Methods* 8, 1050–1052.
- Rolland, T., Tasan, M., Charleatoux, B., *et al.*, 2014. A proteome-scale map of the human interactome network. *Cell* 159, 1212–1226.
- Rozenblatt-Rosen, O., Deo, R.C., Padi, M., *et al.*, 2012. Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins. *Nature* 487, 491–495.
- Ryan, C.J., Cimermancic, P., Szpiech, Z.A., *et al.*, 2013. High-resolution network biology: Connecting sequence with function. *Nature Reviews Genetics* 14, 865–879.
- Sahni, N., Yi, S., Zhong, Q., *et al.*, 2013. Edgotype: A fundamental link between genotype and phenotype. *Current Opinion in Genetics & Development* 23, 649–657.
- Saliba, A.E., Vonkova, I., Ceschia, S., *et al.*, 2014. A quantitative liposome microarray to systematically characterize protein-lipid interactions. *Nature Methods* 11, 47–50.
- Smith, D.F., Song, X., Cummings, R.D., 2010. Use of glycan microarrays to explore specificity of glycan-binding proteins. *Methods in Enzymology* 480, 417–444.
- Smith, L.M., Kelleher, N.L., and Consortium for Top Down Proteomics, 2013. Proteoform: A single term describing protein complexity. *Nature Methods* 10, 186–187.
- Stynen, B., Tournu, H., Tavernier, J., Van Dijck, P., 2012. Diversity in genetic *in vivo* methods for protein-protein interaction studies: From the yeast two-hybrid system to the mammalian split-luciferase system. *Microbiology and Molecular Biology Reviews* 76, 331–382.
- Tewari, M., Hu, P.J., Ahn, J.S., *et al.*, 2004. Systematic interactome mapping and genetic perturbation analysis of a *C. elegans* TGF- β signaling network. *Molecular Cell* 13, 469–482.
- Tuncbag, N., Gursay, A., Keskin, O., 2011. Prediction of protein-protein interactions: Unifying evolution and structure at protein interfaces. *Physical Biology* 8, 035006.
- Venkatesan, K., Rual, J.F., Vazquez, A., *et al.*, 2009. An empirical framework for binary interactome mapping. *Nature Methods* 6, 83–90.
- Vidal, M., Cusick, M.E., Barabási, A.L., 2011. Interactome networks and human disease. *Cell* 144, 986–998.
- Vidal, M., Fields, S., 2014. The yeast two-hybrid assay: Still finding connections after 25 years. *Nature Methods* 11, 1203–1206.
- Vinayagam, A., Zirin, J., Roessel, C., *et al.*, 2014. Integrating protein-protein interaction networks with phenotypes reveals signs of interactions. *Nature Methods* 11, 94–99.
- Walhout, A.J., 2011. Gene-centered regulatory network mapping. *Methods in Cell Biology* 106, 271–288.
- Weatheritt, R.J., Luck, K., Petsalaki, E., Davey, N.E., Gibson, T.J., 2012. The identification of short linear motif-mediated interfaces within the human interactome. *Bioinformatics* 28, 976–982.
- Weirauch, M.T., Cote, A., Norel, R., *et al.*, 2013. Evaluation of methods for modeling transcription factor sequence specificity. *Nature Biotechnology* 31, 126–134.
- Wodak, S.J., Vlasblom, J., Turinsky, A.L., Pu, S., 2013. Protein-protein interaction networks: The puzzling riches. *Current Opinion in Structural Biology* 23, 941–953.
- Yildirim, M.A., Goh, K.I., Cusick, M.E., Barabási, A.L., Vidal, M., 2007. Drug-target network. *Nature Biotechnology* 25, 1119–1126.
- Yu, H., Braun, P., Yildirim, M.A., *et al.*, 2008. High-quality binary protein interaction map of the yeast interactome network. *Science* 322, 104–110.
- Zhang, Q.C., Petrey, D., Deng, L., *et al.*, 2012. Structure-based prediction of protein-protein interactions on a genome-wide scale. *Nature* 490, 556–560.
- Zimmermann, G.R., Lehár, J., Keith, C.T., 2007. Multi-target therapeutics: When the whole is greater than the sum of the parts. *Drug Discovery Today* 12, 34–42.
- Zinn, N., Hopf, C., Drewes, G., Bantscheff, M., 2012. Mass spectrometry approaches to monitor protein-drug interactions. *Methods* 57, 430–440.
- Zotenko, E., Mestre, J., O'leary, D.P., Przytycka, T.M., 2008. Why do hubs in the yeast protein interaction network tend to be essential: Reexamining the connection between the network topology and essentiality. *PLoS Computational Biology* 4, e1000140.

Metabolomics in Cell Biology

SJ Harrison, Danmarks Tekniske Universitet, Hørsholm, Denmark

© 2016 Elsevier Inc. All rights reserved.

Introduction

The dogma that the genome has global and overall control over the processes occurring within living systems especially when they are adapting to new environments has become the preeminent tenant within biology. This is at best an oversimplification of the situation as there are multiple control systems operating with both feed-back and feed-forward inhibition or stimulation operating as controls over the system as a whole. Undoubtedly the genome as the repository of the blueprints of the cellular machinery and Jacob and Monod stated “The synthesis of enzymes in bacteria follows a double genetic control.” (Jacob and Monod, 1961), arguably leading to the belief that all the cellular processes are solely under the control of the genome, as shown in Figure 1(a), where in reality there are multiple complex feed-back and feed-forward controls between and within the groups (Fell, 1992; Kitano, 2002), a simplified view being shown in Figure 1(b).

It is errors and lack of precision within the system that allow for adaption, if an organism is living in a stable environment then it makes evolutionary sense to develop. It may seem to be simply a philosophical debate, but systems biology has set out to explain how these organisms adapt to environmental changes and what the controls over this adaption are.

The small molecule complement of an organism is called the metabolome (Oliver *et al.*, 1998) and hence the measurement of the metabolome is called metabolomics. To carry out a metabolomics study can be both complicated and expensive. In the early days of metabolomics it was thought that by running a set of samples an answer about the biological significance would appear (Plumb *et al.*, 2002). Nowadays it is understood that a rigorous and disciplined approach needs to

be adopted. The steps involved in conducting a successful metabolomics study are highlighted in Figure 2 and will be discussed in detail later.

As can be seen there are a large number of factors that need to be considered when embarking on a metabolomics study, but the first and probably most important question that needs to be asked is ‘Is it necessary/appropriate to conduct a metabolomics study in this case?’ To answer this question it is important firstly to clearly define the terminology.

Metabolomics – is a set of none targeted and unbiased analysis which attempts to analyze the whole metabolome.

Metabolite profiling (or targeted metabolomics) – is a targeted analytical approach which attempts to quantify a specific set of metabolites.

Metabolomics versus Metabolite Profiling

Scientifically and simply for the ease of communication it is important to differentiate between metabolomics and metabolite profiling (Villas-Boas *et al.*, 2005a). With metabolite profiling the experimental design and set-up is simpler to both understand and perform, this approach can be viewed as an applied application of normal quantitative analytical chemistry. More than 100 metabolites can be quantified simultaneously in *Escherichia coli* (Bennett *et al.*, 2009), by the use of stable isotopes this can (and should) be used to improve both the accuracy and precision of the measurements (Lei *et al.*, 2011) and help to remove systematic errors. This data can be used to create genome scale metabolic models (Oberhardt *et al.*, 2009) and these can be extremely useful when evaluating a dataset. If we consider a hypothetical biochemical pathway composed of six metabolites A to F (Figure 3), it would appear by measuring the level of A–F we could understand the

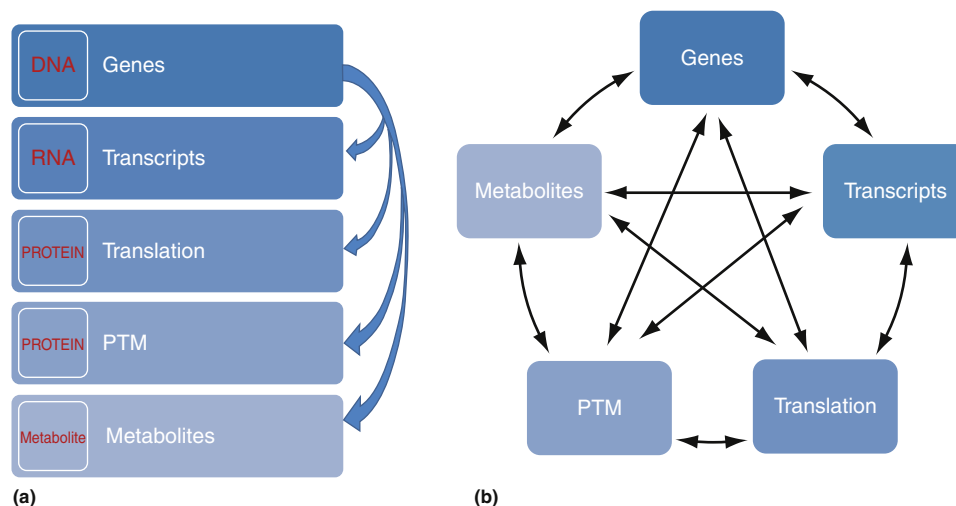


Figure 1 Interactions between control levels in living organisms.

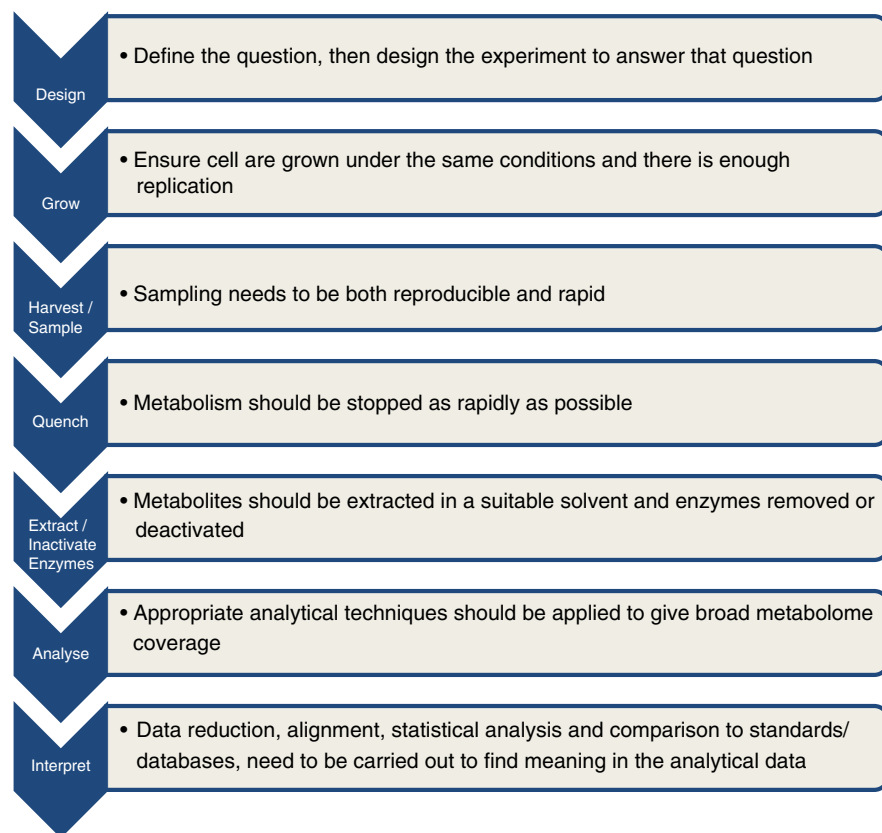


Figure 2 Seven stages that need to be considered when conducting a Metabolomics study.

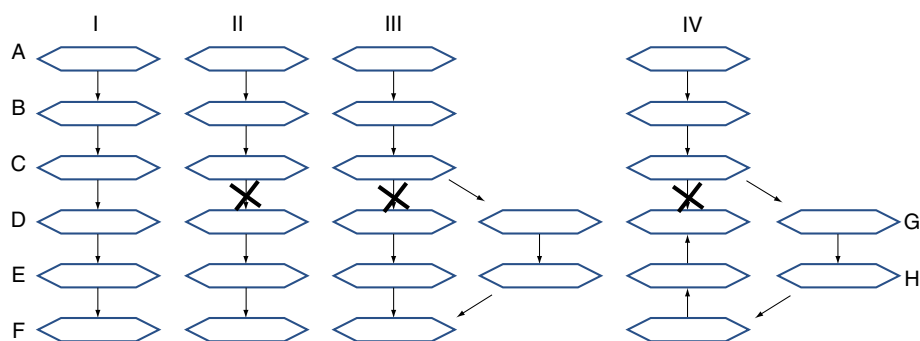


Figure 3 Hypothetical 6 set biosynthetic pathway.

pathway and this is true when the pathway is functioning normally (**Figure 3-I**) but if the pathway is blocked by a nonfunctional enzyme between C and D several possible things may occur. In the case shown in 3-II where the pathway is simply blocked, metabolite profiling would be sufficient to observe an increase in the levels of A–C and a decrease in the levels of D–F. However, if a bypass shunt is established through metabolites G and H which we do not measure as is the case in 3-III and 3-IV, we would observe metabolite levels that are harder to explain.

In the case of 3-III we would observe a sudden decrease in the levels of D and E and 3-IV the pathway may well appear to be functioning normally. The use of metabolomics however would reveal the operation of the unexpected shunt.

Metabolomics does have several disadvantages when compared to metabolic profiling. These included the fact that metabolite levels are generally reported at relative levels rather than absolute, many of the possible metabolites identities are unknown or at least not confirmed. Probably the thing that puts people off carrying out metabolomics experiments is that they can be difficult to conduct as well as interpret ([Camacho et al., 2005](#)) and determine the biological significance.

Conducting a Successful Metabolomics Study

With the challenges previously outlined in mind for the rest of this article we will try to highlight the factors which need to be

considered and what is the current state of the art. The workflow that will be used is that shown in Figure 2.

Design

The first and arguably the most important part of a metabolomics study is the planning. Experimental design like in all science is the place where expensive mistakes can be avoided. Initially the question should be asked 'is metabolomics the correct approach to generate the data to test the hypothesis under trial in the experiment.' In the early years of metabolomics it was considered sufficient to collect a set of samples and analyze them to see what fell out, this hypothesis free approach was quickly found to be a very risky strategy. It is now commonly accepted that like in all sciences a well-designed and thought out experimental plan which learns from previous studies (Broadhurst and Kell, 2006) is the best approach and can prevent expensive mistakes. It is important to consider the analytical variables to yield high-quality data (Patti *et al.*, 2012) but it is equally important to consider the biological variables.

If the experiment is not designed to answer a specific question then it is highly probable that the answer that is received will be to a question that has not been asked. To illustrate this let's consider we want to use metabolomics to find the effect of introducing a gene into yeast. Of course we would not grow two different strains as the likely outcome of that analysis would be the metabolic differences between the strains. So a more sensible approach would be to grow the wild-type strain and compare this to the engineered strain; however, it is well known that the process of engineering a strain is stressful to an organism and this stress itself can have a major influence on the metabolism. There it is likely that this analysis would yield markers for the stress caused by engineering cells (Eiteman and Altman, 2006) rather than the desired outcome which is to assess the effect of the new genes on the metabolism. Logically, the ideal experimental design to answer the question 'what is the effect of introducing a gene to a strain of yeast?' would be to compare the engineered strain to one that has been through the process of engineering but with an empty vector.

In the above example it was decided that the optimal experimental design to answer the question 'what is the effect of introducing a new gene to an organism?' is to compare it to an empty vector control. However, this would make it appear that it is possible to achieve the answer by the analysis of two (possibly three, if a blank is included) samples.

Although this would be theoretically possible, it is highly unlikely that such an answer could be considered reliable and the major cause of the lack of confidence in such an answer is variability within the system. In such complex systems there are numerous sources of variability. These sources of variability can be divided into two broad classes, biological variability, and analytical variability. The biological variability is usually significantly greater than the Analytical variability. To overcome these sources of variability replication is used (Dunn *et al.*, 2005), so the next obvious question is 'how many replicates should be used?' This of course is simpler to ask than it is to answer. The level of required replication is correlated to the degree of the change that is to be assessed but generally the

more replication the better. Although currently there seems to be no reported method to evaluate the appropriate level, in our labs we always try to carry out experiments where $n > 7$.

This being said the question that comes to mind is 'what a biological replicate is?' As per normal, the answer is not absolutely obvious at first sight. A sub aliquot of an initial sample, a second sample from the same reactor, a sample of a clone grown in a different reactor under the same conditions or a different strain with the same modification grown in a different reactor under the same conditions. Although commonly used, the first two are not really good biological replicates (Fernie *et al.*, 2011), it is arguable as to which of the last two is the best biological replicate and the answer would depend on the question that is being asked. If we are asking 'what is the effect of genetic modification within a given clone' then the best biological replicate would be a sample of a clone grown in a different reactor under the same conditions. However, if we are trying to judge the effect of introducing a specific gene then a different strain with the same modification grown in a different reactor under the same conditions would be the preferred biological replicate. Once again we come back to the concept that we must ask the correct question to design the experiment.

So far we have only considered the simple case of a pair-wise comparison at a single time point, but the same considerations are valid when comparing multiple strains (this can be viewed as doing multiple pair-wise comparisons) or when looking at temporal variation.

It is important to consider what types of samples will be required to be taken during the course of the experiment; microorganisms interact with their environment by both importing into the cell and exporting from its chemicals. This means that it is possible to consider the internal metabolism of the cells known as the endo-metabolome or fingerprint and/or external metabolism of the cells known as the exo-metabolome or footprint (Tian *et al.*, 2009; Figure 4).

Growth

The use of standard and controlled conditions is of critical importance (Werf *et al.*, 2007) to the success a metabolomics study has at answering the question that was initially asked. In bioreactors factors such as temperature, pH, media, O₂, and CO₂ are readily controlled (Mashego *et al.*, 2007) but in shake

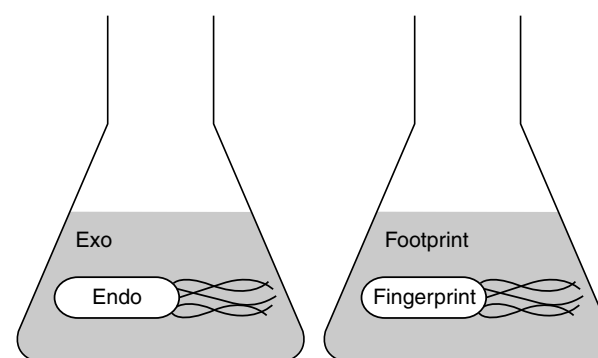


Figure 4 Representation of the endo and exo metabolome.

flasks and well plates these factors are more difficult to control. To improve mixing in well plate, the addition of a glass bead has been reported to improve mixing and hence gas transfer (Ewald *et al.*, 2009). However, for such an important area that is known to have significant effects on the response of cultured cells there is surprising little literature reporting the studies of basic factors. Factors such as edge effects due to changes in the microenvironment either in plates or within incubators are common knowledge and taken account of in well-designed experiments but little primary source materials can be found that relates to metabolomics analysis.

Harvesting/Sampling

The most important factor that affects what is the best sampling protocol to use is whether fingerprint or footprint analysis is to be carried out (see below for definitions).

Footprint analysis: This is arguably the easier of the two analyses, as only the levels metabolites in the media are measured; the sample prep for this can be as simple as centrifuging an aliquot of the media prior to analysis (Pope *et al.*, 2007) or sterile filtering (Dorries and Lalk, 2013). Generally, it is assumed that the media can be expected to be in a steady state condition during the time scale of the sampling and therefore the exact sampling protocol is less important than the fact that it is carried out accurately and reproducibly for the samples in the batch.

There are two commonly applied techniques to measuring the fingerprint metabolome of microorganisms and the two approaches differ as to whether or not the cells are separated from the media for the fingerprint analysis to be carried out.

The two techniques are:

1. Differential analysis

In differential analysis two aliquots of the media are taken, the cells are removed from one, either by filtration or centrifugation. The two extracts are then analyzed and the differences in the metabolites and their levels allow the determination of the fingerprint metabolome (Taymaz-Nikerel *et al.*, 2009).

2. Isolated cell analysis

In isolated cell analysis, the cells are removed from the media by filtration or centrifugation and then they are lysed to extract the metabolites prior to analysis, this of course seems to be the simplest approach to take, however it has been noticed that cells can undergo significant metabolic changes during the process of isolation from the media, due to entering a state of starvation. To prevent this, generally cells are quenched through fast changes in either temperature or pH, during sampling process to stop all metabolic activity (Villas-Boas *et al.*, 2005b) (this will be discussed further below).

Whichever type of analysis is going to be carried out, it is clear that the time required to take the sample may significantly affect the results. The most important thing, however, is to ensure that all the samples within a set are treated identically, to ensure random errors are minimized. As to which technique is most appropriate for a given organism, as per usual in the field of metabolomics there are many approaches that can/should be tried to find the one that is most appropriate,

such as fast sampling (van Gulik, 2010), forced air *in situ* sampling (McCloskey *et al.*, 2014a), fast filtration (McCloskey *et al.*, 2014b). It is generally accepted that if the intracellular metabolome is to be measured then a rapid sampling technique/device (such as the one in Figure 5) should be used.

Quenching

Once the sample has been taken from the fermentation vessel if intracellular metabolites, which are turned over rapidly, are being measured, it is important to stop the metabolic processes as quickly and completely as possible. This process is known as quenching and can be done by rapidly increasing temperature, rapidly decreasing temperature or significantly changing the pH. As to which technique is most appropriate for a given organism there are many approaches that can/should be tried to find the one that is most appropriate (Table 1).

Extract/Inactivate Enzymes

If quenching has been carried out using one of the 'cold' methods, when the sample warms up the enzymes that are present are likely to have significant activity and change the levels of the metabolites in the sample. In the extraction process, the aim is to extract all the metabolites that are present whilst leaving behind proteins, enzymes, and contaminants. This unfortunately is not possible, and any extraction solvent that is used will discriminate between various classes of metabolite, again the key is to treat all the samples in the same way. There are numerous extraction solvents/technique reported in the literature, the most common are acidic acetonitrile/water, boiling water, boiling ethanol, or methanol. Once again it is advised several techniques are tried prior running a large experiment. It is important to keep in mind that the more complex the technique used the more prone it is to errors and hence variability. As there is no 'one size fits all' extraction protocol, it is advisable to match the extraction method with the analytical method, for example, if looking for lipids and aqueous extraction is probably not advisable whereas chloroform/methanol may be very suitable but you would not expect to see nucleotides in that extract.

As metabolomics ultimately relies on statistical analysis of the data to elucidate any differences between the treatment groups, it is extremely important that the potential sources of error are understood and that the analytical data collected is as reliable as possible. There are several types of samples that can and should be run as part of a batch (Table 2).

To improve the reliability and reproducibility of the acquired data the normal analytical chemistry approach would be to use isotopically labeled internal standards, however this is impractical in the case of metabolomics as it would be cost prohibitive. However it is possible to use cell extracts from cell uniformly ¹³C-labeled sole carbon source as internal standards (Wu *et al.*, 2005). This approach allows correction for changes in instruments sensitivity during the analysis of a batch of samples. If this approach is to be used as mentioned previously the appropriate extraction system should be chosen for the analysis, using the isotopically labeled extract as the extraction solvent can be helpful in reducing the effects of losses during extraction.

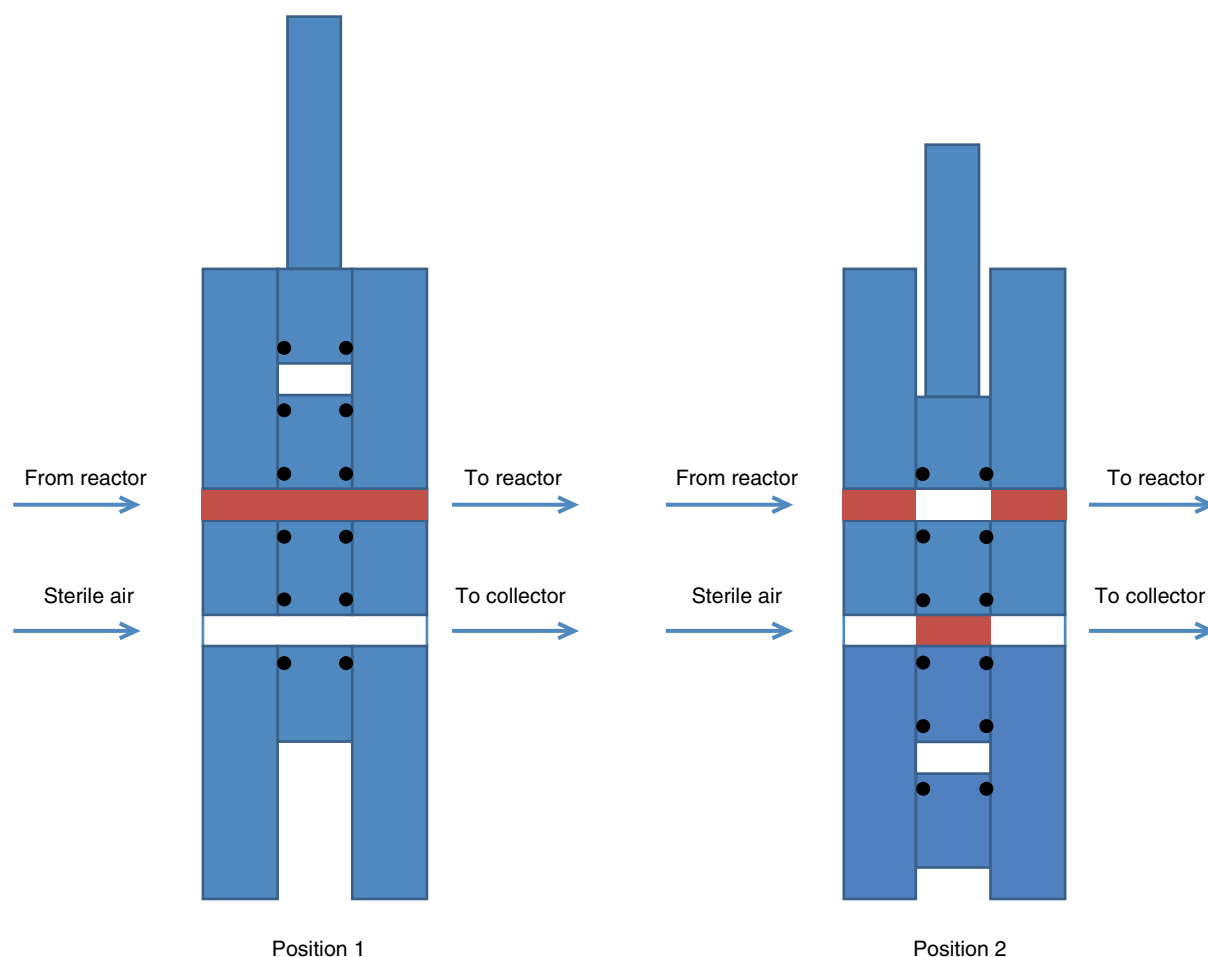


Figure 5 Low dead volume sampling device (design provided by Prof. Walter M. van Gulik, Delft University).

Table 1 Some common quenching techniques

Heat	Boiling 75% ethanol with 70 mM-Hepes	Gonzalez et al. (1997)
	Boiling water	Wittmann et al. (2004)
Cold	Cold glycerol-saline	Villas-Boas and Bruheim (2007)
	Cold phosphate-buffered saline	Teng et al. (2009)
	Cold 60% methanol solution	Winder et al. (2008)
	Cold 60% methanol with 0.85% ammonium bicarbonate	Sellick et al. (2008)
pH	Perchloric acid	Theobald et al. (1993)
	2 M KOH solution	Theobald et al. (1997)

Analyze

Similar to the fact that there is no one extraction solvent that is suitable for the extraction of the entire metabolome, there is no single analytical technique that enables the analysis of the

complete metabolome. The type of molecule each technique is applicable to is shown in [Figure 6](#). To increase the coverage of the molecules analyzed it is not uncommon to run multiple analytical techniques ([Meyer et al., 2010](#)).

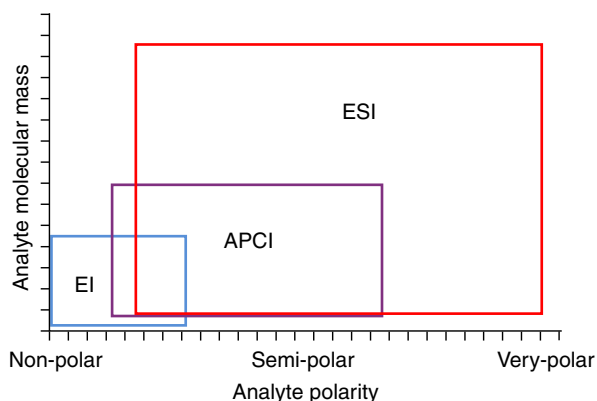
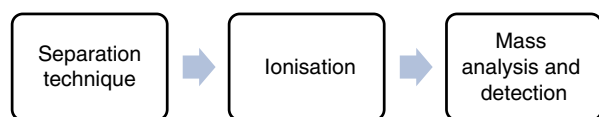
It is widely accepted that mass spectrometry is the technique of choice for metabolomics experiments, with the one exception of direct infusion mass spec, all single stage mass spectrometer can be thought of as consisting of three functional units ([Figure 7](#)).

There are two separation techniques that are commonly used for metabolomics analysis, gas chromatography (GC) and ((ultra) high performance) liquid chromatography (((U) HP)LC).

GC is generally used for the analysis of volatile compounds (i.e., relatively small and nonpolar) or compounds that can be made volatile by chemical modification (derivatisation). There are many possible derivatisation reagents available each of which varies in their reactivity and the classes of compounds they are useful for. Probably the most generic derivatisation reagents are the trimethylsilyl group (TMS) such as *N,O*-Bis(trimethylsilyl)trifluoroacetamide (BSTFA) or *N*-Methyl-*N*-trimethylsilylfluoroacetamide (MSTFA) ([Kanani et al., 2008](#)). The second group of commonly used derivatising reagents are the alkylating reagents such methylchloroformate (MCF)

Table 2 Type of samples to be run during a Metabolomics study

Sample	Subclass	How it is made	Purpose
Blanks	Solvent Blank	Extraction solvent in clean vial	To check for carry-over especially after high level standards
	System Blank	Extraction solvent put through the extraction process just with no sample	To allow the identification of contaminants that are introduced during the sample processing
Replicates	Technical	A second extract of a sample that has been split (from same flask at same time)	To show that the extraction and analysis are reproducible
	Analytical	A second injection of a sample that has been injected	To show that the analysis is reproducible
	Biological	A second extract of a sample that has been grow under the same conditions (from different flask at same time)	To show the level of biological reproducibility
QC/Controls	Pooled	An equal volume of the extract from all the samples in a batch	To show the level of temporal drift in the analysis
	Xenobiologic	Extract of a non-related organism that has a different metabolome	To show the whole system can differentiate reliably between different metabolomes
Standards	RT standards	Set of known compound at different retention times (rt)	To allow for rt to be adjusted over time during the course of a large study and allow comparisons between studies run at different times
	ID standards	Known compound in solution of suitable solvent at known concentrations	To allow the identification and/or quantification of unknown compounds

**Figure 6** Graphical representation of the analytical technique applicability.**Figure 7** Function representation of a mass spectrometer.

(Villas-Bôas *et al.*, 2003). When a GC is coupled to a mass spectrometer the most common method of ionizing the analytes is by bombarding them with electrons at a known energy (electron ionization (EI)). The energy imparted by the electrons cause the ejection of an electron from the analyte and hence it becomes an ion, the excess energy causes the ions to fragment in reproducible and predictable ways. The fragmentation patterns observed in EI–GC–MS analysis can be

compared to commercially available libraries to aid identification of unknown analytes, but care should be taken. These library identifications should be treated as indicative and not as definitive. To truly confirm the identification of an analyte an authentic standard should be run along with a coinjection with the sample. GC–MS (Figure 8(a)) provides the primary analysis technique for the analysis of volatile and semi-volatile analytes.

For the separation of larger and more polar analyte, LC is generally preferred over GC. By far the most commonly used separation carried out on LC systems a reverse phase method, based on the C18 column, these columns however do not retain very polar analytes, although with the advent of aqueous C18 columns this has improved. To analyze very polar analytes it is common to use an orthogonal separation approach known as hydrophilic interaction chromatography (HILIC). Both of these separation techniques are generally coupled to a mass spectrometer by an electrospray ion source (ESI) to form an LC–MS system (Figure 8(b)), the LC eluent passes through a needle that is held at an electrical potential relative to mass spectrometer and as the sprayed droplets evaporate they become ionized. ESI is known as a soft ionization technique as it has a tendency to form molecular ions/adducts with little fragmentation. However, if the separation is incomplete then the co-eluting analyte compete for the charge which can cause the measured level of one of the compounds to be lower than it actually is. Electrospray can be run in either positive or negative ion mode and it is advisable to run the batch of sample in both modes. In many cases it is possible to switch alternatively between the ionization modes but this effectively halves the data acquisition rate so in many labs running a sample in positive mode and then rerunning it in negative mode is preferred. Electrospray ionization is not very suitable for

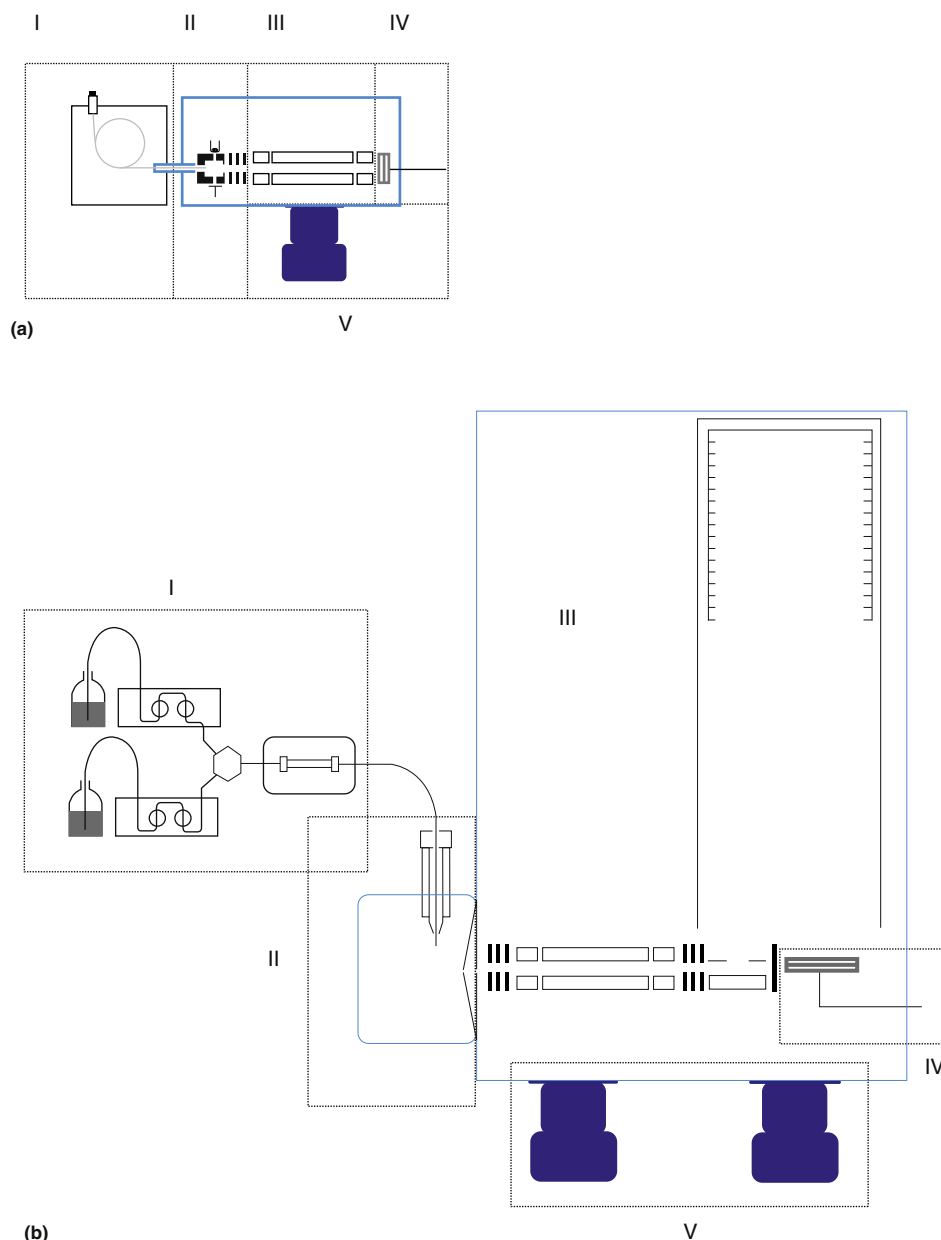


Figure 8 Schematic of (a) GC-MS and (b) LC-MS system; I: Separation method, II: Ion source, III: Mass analyzer, IV: Detector, and V: Turbopumps.

the analysis of analytes of low polarity such as lipids, for these types of molecules another type of ionization is used this is known as atmospheric pressure chemical ionization (APCI).

There are still issues with the separation and analysis of extremely polar and extremely nonpolar metabolites. To analyze polar analytes there are two commonly used approaches ion-pairing (Lu *et al.*, 2010) or capillary electrophoresis (Soga *et al.*, 2003). For the analysis of nonpolar analytes supercritical fluid chromatography (SFC) has started to remerge as a technique of interest (Bamba *et al.*, 2008).

From the previous discussion it can easily be seen that multiple technique need to be use to attempt to achieve a study that yields the broadest range of metabolites as possible, even though it will not be truly complete.

Data Interpretation

Once the data from a well-designed and well-executed non-targeted metabolomics experiment have been collected, it needs to be analyzed to give meaning and biological relevance to the data. These stages are still very active areas of research and can be divided into two primary areas, data reduction and data analysis.

GC-MS data

The electron impact or electron ionization (EI) as it is also known is the most common method of ionization used in GC-MS analysis and as stated previously as well as ionizing

the molecule in the majority of cases it causes the ion to fragment. As long as the energy of the electrons are at the standard accelerating voltage of 70 eV the resultant spectra should be reproducible and therefore can be searched against the reference databases or can be worked out from the structure (Figure 9(e)). An examination of the EI spectra (Figure 9(b) and 3(d)) it is apparent that no ion that corresponds to the ion from the intact molecule can be seen, this is not uncommon in GC-MS, however, this can hinder the identification of unknowns. It cannot be stressed enough that the identification of a compound by a database search

should only be counted as indicative and little confidence should be placed on such identifications. Where ever possible confirmation of such identification should be made by the injection of an authentic standard under the same condition or even better is to also run the authentic standard mixed with the sample. In complex samples such as cell extracts or media it can be difficult to get a spectrum that is from only one compound, in this case mathematical algorithms can be used to tease apart the spectra and assign each ion to the correct spectrum, this is known as deconvolution.

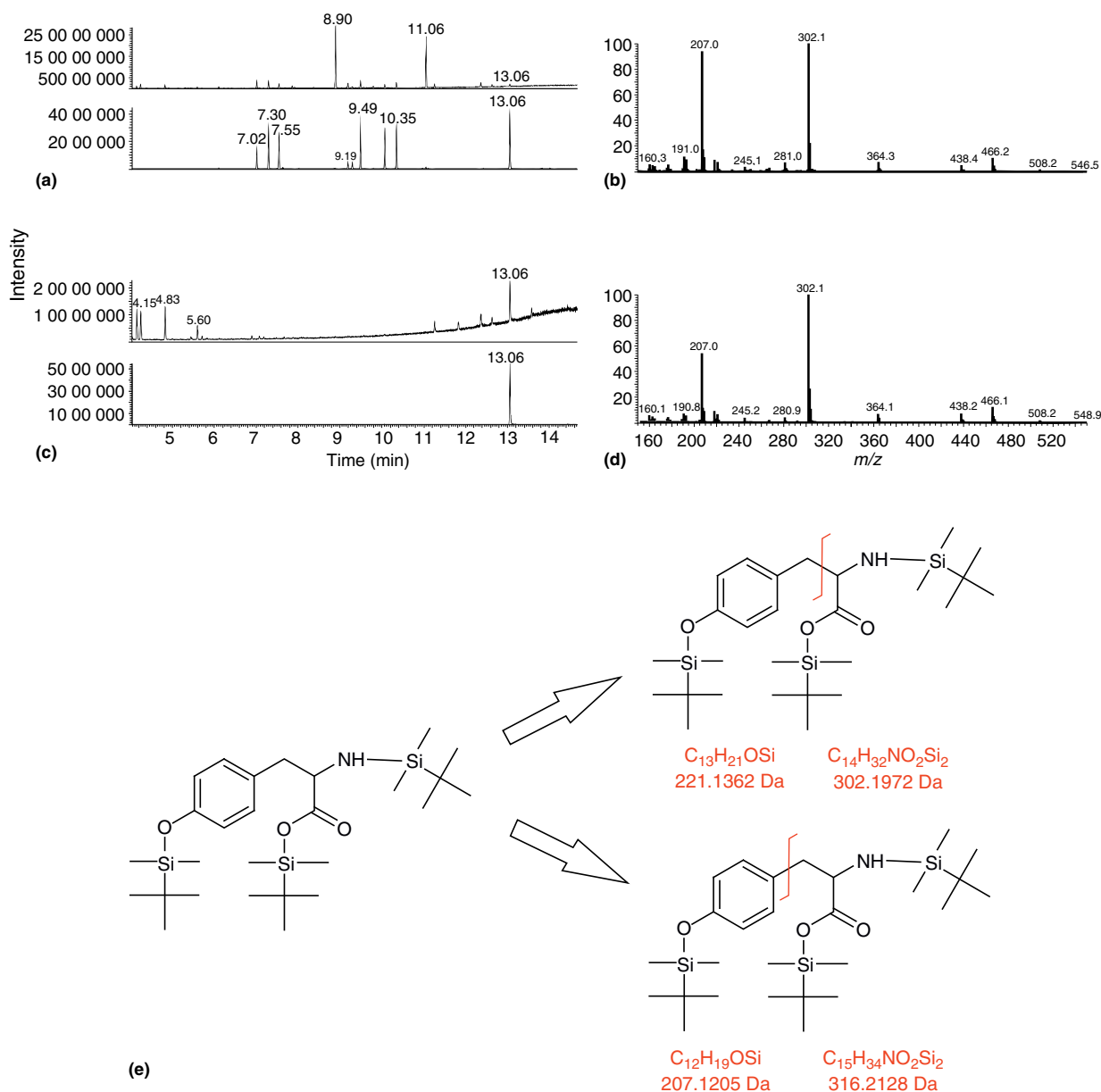


Figure 9 An example of GC-MS data; (a) (Top) The total ion chromatogram (TIC), (Bottom) Extracted ion chromatogram (XIC) of the 302 ion for a TBDMS-derivatised acid hydrolysate from yeast; (b) Spectrum from the peak at a retention time of 13.06 mins which has a suggested identification of TBDMS-Tyrosine when search against NIST DB; (c) (Top) The total ion chromatogram (TIC), (Bottom) Extracted ion chromatogram (XIC) of the 302 ion for a TBDMS-derivatised Tyrosine standard; (d) Spectrum from the peak at a retention time of 13.06 mins for TBDMS-derivatised Tyrosine standard; (e) Possible fragments for TBDMS-derivatised Tyrosine.

LC-MS data

In LC-MS the most common ionization technique that is used is electrospray (ESI), in which the solvent which has been used to elute the column is passed through a needle which is held at an electrical potential relative to the orifice on the mass spectrometer. As the droplets travel through the air in the source they evaporate and the charge on them increases up to the point where it overcomes the surface tension. At this point the droplet breaks apart and the ion is formed, if there is more than one compound within the droplet then the compound can compete for the available charge and the ionization of one of the compounds may be suppressed (this is known as ion suppression), this is why good chromatographic separation is extremely important in metabolomics experiments and why I am not a supporter of direct infusion mass spectrometry for metabolomics. ESI is a soft ionization technique which means that the ions that are formed tend to contain the intact molecules and little or no fragmentation occurs. However there can be a tendency to form adducts with components within the solvent or at high concentration form multimers where an ionized molecule shares its charge with one or more unionised molecules (**Figure 10(b)**). Dependent on the capabilities of the instrumentation used to collect the data the number of possible identifications can be vast if only MS^1 is used. The number of possible identities can be reduced by the use of higher order MS analysis which cause the ions to fragment by colliding them with gas molecules, leading to fragments that at first sight resemble those seen in GC-MS but as there is no standard for these types of collision-induced dissociation in LC-MS there is no database available for the reliable identification of unknown molecules especially when comparing data acquired on instruments from two different vendors. If an instrument is capable of acquiring accurate mass data it is possible to drastically reduce the number of possible elemental formulas for an unknown compound (**Kind and Fiehn, 2006**), however this is far from being an identification. Best practice would have both higher MS analysis and accurate mass to aid in identification of unknowns, this remains as one of the major challenges in metabolomic analysis.

The data reduction stage can be further divided into deconvolution, alignment, and normalization (**Katajamaa and Oresic, 2007**). The output from this stage should be a table of the response for each peak/feature for every file. There are many both commercial and freeware products available that will enable some or all of the data reduction to be automated, however, the controls, qc samples, and replicates should be used to validate the results that are obtained.

Data analysis of the tabular data is usually conducted using a type of discriminate analysis; by far the most used (and probably abused) is principle component analysis (PCA). PCA is generally a good starting point to quickly visualize and data quality check. There are easy to use data analysis tools that are available online (**Xia et al., 2012**).

As stated previously to place statistically significant peaks in biological context they must be identified, so they can be mapped onto the metabolic networks. This step can be both extremely difficult and time consuming. With GC-MS analysis yielding the peaks of interest, a good starting point in the process of identification is searching against commercially

available libraries and then any putative identification can be confirmed by use of authentic standards. Chemical ionization (CI) can be used to try to find the molecular weight which can help narrow the search space. CI is a softer ionization technique than EI however, it still somewhat of a 'black art' being more complex to set up than EI. Another recent development in the GC-MS is the advent of 'affordable' accurate mass instruments which have a great deal of potential to improve the reliability and ease of identification of unknown peaks.

For LC-MS, these commercial libraries are not widely available so other approaches are needed, such as MS^n and/or accurate mass analysis. MS^n can give fragmentation patterns (like GC-MS) that give structural information but due to the lack of standardization between vendors these tend vary between instrument types. Accurate mass analysis (if carried out at a sufficient accuracy) can yield a short list of possible elemental formula, but even if one elemental formula is found it is highly likely that there will be significant number possible chemicals of that formula. A combination of both MS^n and accurate mass analysis coupled with biological context is the best way to approach the identification of unknown peaks. Similarly to GC-MS, when a putative identification has been made, at the very least an authentic standard should be run but better than that is co-injecting the authentic standard with a sample containing the peak of interest.

If UL-isotopically labeled cell extracts are used as internal standards in both GC-MS and LC-MS, this can be a significant help in compound identification, as the mass difference between the unlabeled ion and the labeled one can be used to calculate the number of carbon atoms in a molecule for both fragments and molecular ions. This technique can be extremely powerful so long as all the carbon sources are fully labeled, i.e., no carbon is being fixed from CO_2 . Once a significant compound has been identified, it is good to show a biological effect of the compound although this is not always possible.

Using Metabolomics Data (What Good Is It?)

The aim of any metabolomics experiment is to draw conclusion from the data that leads to gaining an increased understanding of the biological system. The use for metabolomics data depends very much on whether or not statistically significant components can be identified. In the case where they have not/cannot be identified, it is possible to use these components as potential biomarkers for the particularly different conditions tested in the original experiment. However such biomarkers should be validated for the predictive power for the conditions they will be used to differentiate between especially when it come to the use to aid in medical diagnosis (**Koulman et al., 2009**). The inability to identify unknown peak/components has undoubtedly lead to the popularity of metabolite profiling where the data is significantly easier to interpret although runs the risk of experiments being designed that can significantly over estimate a metabolites importance within a biological process. However if significant peaks in an untargeted metabolomic experiment can be identified then the levels can be used to help create metabolic models (**Kell, 2004**), which can in turn be used to predict an organisms response to a particular challenge. It is likely that only the

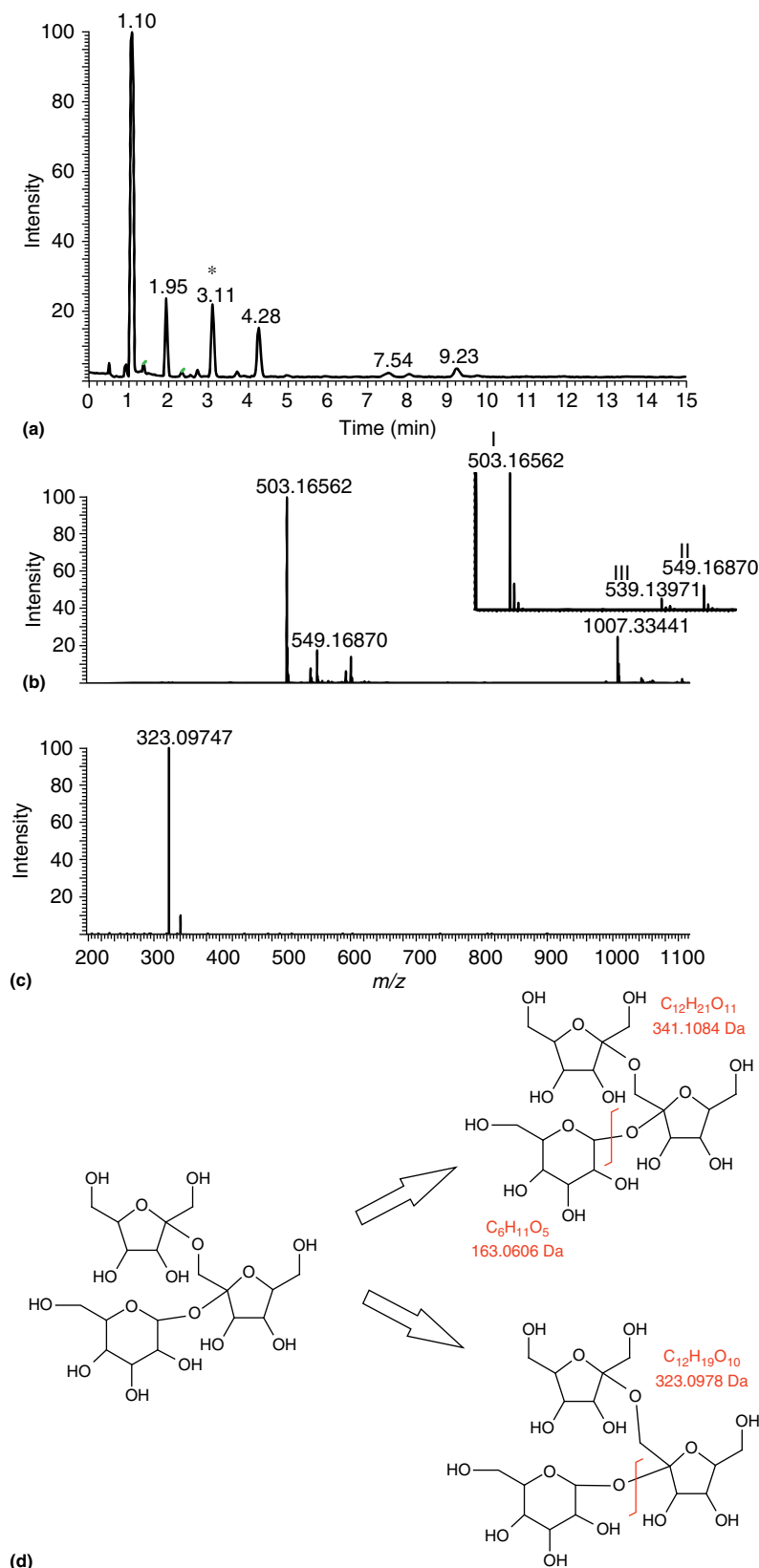


Figure 10 An example of LC-MS data; (a) The total ion chromatogram (TIC) of a sample; (b) Spectrum from the peak at a retention time of 3.11 mins marked with the *. In the insert ion I at m/z 503 corresponds to the molecular ion, II at m/z 549 corresponds to the formate adduct, III at m/z 539 corresponds to the major chloride adduct, the ion in the made spectrum at m/z 1007 corresponds to the dimer; (c) the MS2 spectrum of the molecular ion (m/z 503), (d) possible fragmentation route for the molecule.

metabolomics experimental approach will enable us to fully understand the metabolic processes and capacity of cells, as it has the potential to allow the discovery of unknown/unlikely metabolic pathway cell can adopt when under stress.

The Future and Notes of Caution

Metabolomic is still in its infancy as an approach to deciphering complexity in biological systems but does show great promise (Harrison and Herrgard, 2013). Initially metabolomics was viewed as being able to answer question which you did not even know to ask. Over the past decade or so it has become more apparent that metabolomics is better suited as a discovery technique which enables a better question to be formulated for more targeted experiments. As instruments improve the volumes of data that can be acquired increases vastly but simply collecting data has not been shown to be a useful use of time and resources. To be successful metabolomics requires good communication between a multidisciplinary team, who formulate a robust, rigorous and appropriate experiment, which then is executed in a scientific and disciplined manner.

See also: Molecular Principles, Components, Technology, and Concepts: Lipids: Lipidomics. Molecular Principles, Components, Technology, and Concepts: Metabolism: A Structure Perspective on Organelle Bioenergetics; Metabolic Regulation

References

- Bamba, T., Shimonishi, N., Matsubara, A., *et al.*, 2008. High throughput and exhaustive analysis of diverse lipids by using supercritical fluid chromatography-mass spectrometry for metabolomics. *Journal of Bioscience and Bioengineering* 105 (5), 460–469.
- Bennett, B.D., Kimball, E.H., Gao, M., *et al.*, 2009. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nature Chemical Biology* 5 (8), 593–599.
- Broadhurst, D., Kell, D., 2006. Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics* 2 (4), 171–196.
- Camacho, D., de la Fuente, A., Mendes, P., 2005. The origin of correlations in metabolomics data. *Metabolomics* 1 (1), 53–63.
- Dorries, K., Lalk, M., 2013. Metabolic footprint analysis uncovers strain specific overflow metabolism and D-isoleucine production of *Staphylococcus aureus* COL and HG001. *PLoS One* 8 (12), e81500.
- Dunn, W.B., Bailey, N.J., Johnson, H.E., 2005. Measuring the metabolome: Current analytical technologies. *Analyst* 130 (5), 606–625.
- Eiteman, M.A., Altman, E., 2006. Overcoming acetate in *Escherichia coli* recombinant protein fermentations. *Trends in Biotechnology* 24 (11), 530–536.
- Ewald, J.C., Heux, S.P., Zamboni, N., 2009. High-throughput quantitative metabolomics: Workflow for cultivation, quenching, and analysis of yeast in a multiwell format. *Analytical Chemistry* 81 (9), 3623–3629.
- Fell, D.A., 1992. Metabolic control analysis: A survey of its theoretical and experimental development. *Biochemical Journal* 286 (Pt 2), 313–330.
- Fernie, A.R., Aharoni, A., Willmitzer, L., *et al.*, 2011. Recommendations for reporting metabolite data. *Plant Cell Online* 23 (7), 2477–2482.
- Gonzalez, B., Francois, J., Renaud, M., 1997. A rapid and reliable method for metabolite extraction in yeast using boiling buffered ethanol. *Yeast* 13 (14), 1347–1355.
- van Gulik, W.M., 2010. Fast sampling for quantitative microbial metabolomics. *Current Opinion in Biotechnology* 21 (1), 27–34.
- Harrison, S.J., Herrgard, M.J., 2013. The uses and future prospects of metabolomics and targeted metabolite profiling in cell factory development. *Industrial Biotechnology* 9 (4), 196–202.
- Jacob, F., Monod, J., 1961. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* 3 (3), 318–356.
- Kanani, H., Chrysanthopoulos, P.K., Klapa, M.I., 2008. Standardizing GC-MS metabolomics. *Journal of Chromatography B* 871 (2), 191–201.
- Katajamaa, M., Oresic, M., 2007. Data processing for mass spectrometry-based metabolomics. *Journal of Chromatography A* 1158, 318–328.
- Kell, D.B., 2004. Metabolomics and systems biology: Making sense of the soup. *Current Opinion in Microbiology* 7 (3), 296–307.
- Kind, T., Fiehn, O., 2006. Metabolomic database annotations via query of elemental compositions: Mass accuracy is insufficient even at less than 1 ppm. *BMC Bioinformatics* 7 (1), 234.
- Kitano, H., 2002. Systems biology: A brief overview. *Science* 295 (5560), 1662–1664.
- Koulman, A., Lane, G.A., Harrison, S.J., Volmer, D.A., 2009. From differentiating metabolites to biomarkers. *Analytical and Bioanalytical Chemistry* 394 (3), 663–670.
- Lei, Z., Huhman, D.V., Sumner, L.W., 2011. Mass spectrometry strategies in metabolomics. *Journal of Biological Chemistry* 286 (29), 25435–25442.
- Lu, W., Clasquin, M.F., Melamud, E., *et al.*, 2010. Metabolomic analysis via reversed-phase ion-pairing liquid chromatography coupled to a stand alone orbitrap mass spectrometer. *Analytical Chemistry* 82 (8), 3212–3221.
- Mashego, M., Rumbold, K., De Mey, M., *et al.*, 2007. Microbial metabolomics: Past, present and future methodologies. *Biotechnol Letters* 29 (1), 1–16.
- McCloskey, D., Gangotri, J.A., King, Z.A., *et al.*, 2014a. A model driven quantitative metabolomics analysis of aerobic and anaerobic metabolism in *E. coli* K12 MG1655 that is biochemically and thermodynamically consistent. *Biotechnology and Bioengineering* 111 (4), 803–815.
- McCloskey, D., Utrilla, J., Naviaux, R.K., Palsson, B.O., Feist, A.M., 2014b. Fast Swinex filtration (FSF): A fast and robust sampling and extraction method suitable for metabolomics analysis of cultures grown in complex media. *Metabolomics*. 1–12.
- Meyer, H., Liebeke, M., Lalk, M., 2010. A protocol for the investigation of the intracellular *Staphylococcus aureus* metabolome. *Analytical Biochemistry* 401 (2), 250–259.
- Oberhardt, M.A., Palsson, B.O., Papin, J.A., 2009. Applications of genome-scale metabolic reconstructions. *Molecular Systems Biology* 5 (1), n/a.
- Oliver, S.G., Winson, M.K., Kell, D.B., Baganz, F., 1998. Systematic functional analysis of the yeast genome. *Trends in Biotechnology* 16 (9), 373–378.
- Patti, G.J., Yanes, O., Siuzdak, G., 2012. Innovation: Metabolomics: The apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology* 13 (4), 263–269.
- Plumb, R.S., Stumpf, C.L., Gorenstein, M.V., *et al.*, 2002. Metabolomics: The use of electrospray mass spectrometry coupled to reversed – Phase liquid chromatography shows potential for the screening of rat urine in drug development. *Rapid Communications in Mass Spectrometry* 16 (20), 1991–1996.
- Pope, G.A., MacKenzie, D.A., Defernez, M., *et al.*, 2007. Metabolic footprinting as a tool for discriminating between brewing yeasts. *Yeast* 24 (8), 667–679.
- Sellick, C.A., Hansen, R., Maqsood, A.R., *et al.*, 2008. Effective quenching processes for physiologically valid metabolite profiling of suspension cultured mammalian cells. *Analytical Chemistry* 81 (1), 174–183.
- Soga, T., Ohashi, Y., Ueno, Y., *et al.*, 2003. Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. *Journal of Proteome Research* 2 (5), 488–494.
- Taymaz-Nikerel, H., De Mey, M., Ras, C., *et al.*, 2009. Development and application of a differential method for reliable metabolome analysis in *Escherichia coli*. *Analytical Biochemistry* 386 (1), 9–19.
- Teng, Q., Huang, W., Collette, T.W., Ekman, D.R., Tan, C., 2009. A direct cell quenching method for cell-culture based metabolomics. *Metabolomics* 5 (2), 199–208.
- Theobald, U., Mailinger, W., Baltes, M., Rizzi, M., Reuss, M., 1997. *In vivo* analysis of metabolic dynamics in *Saccharomyces cerevisiae*: I. Experimental observations. *Biotechnology and Bioengineering* 55 (3), 305–316.
- Theobald, U., Mailinger, W., Reuss, M., Rizzi, M., 1993. *In vivo* analysis of glucose-induced fast changes in yeast adenine nucleotide pool applying a rapid sampling technique. *Analytical Biochemistry* 214 (1), 31–37.
- Tian, J., Sang, P., Gao, P., *et al.*, 2009. Optimization of a GC – MS metabolic fingerprint method and its application in characterizing engineered bacterial metabolic shift. *Journal of Separation Science* 32 (13), 2281–2288.
- Villas-Boas, S., Rasmussen, S., Lane, G.A., 2005a. Metabolomics or metabolite profiles? *Trends in Biotechnology* 23 (8), 385–386.
- Villas-Boas, S.G., Bruheim, P., 2007. Cold glycerol – Saline: The promising quenching solution for accurate intracellular metabolite analysis of microbial cells. *Analytical Biochemistry* 370 (1), 87–97.
- Villas-Bôas, S.G., Delicado, D.G., Åkesson, M., Nielsen, J., 2003. Simultaneous analysis of amino and nonamino organic acids as methyl chloroformate

- derivatives using gas chromatography–mass spectrometry. *Analytical Biochemistry* 322, 134–138.
- Villas-Boas, S.G., Hojer-Pedersen, J., Akesson, M., Smedsgaard, J., Nielsen, J., 2005b. Global metabolite analysis of yeast: Evaluation of sample preparation methods. *Yeast* 22 (14), 1155–1169.
- Werf, M.J., Overkamp, K.M., Muilwijk, B., Coulier, L., Hankemeier, T., 2007. Microbial metabolomics: Toward a platform with full metabolome coverage. *Analytical Biochemistry* 370 (1), 17–25.
- Winder, C.L., Dunn, W.B., Schuler, S., *et al.*, 2008. Global metabolic profiling of *Escherichia coli* cultures – An evaluation of methods for quenching and extraction of intracellular metabolites. *Analytical Chemistry* 80 (8), 2939–2948.
- Wittmann, C., Kromer, J.O., Kiefer, P., Binz, T., Heinze, E., 2004. Impact of the cold shock phenomenon on quantification of intracellular metabolites in bacteria. *Analytical Biochemistry* 327 (1), 135–139.
- Wu, L., Mashego, M.R., van Dam, J.C., *et al.*, 2005. Quantitative analysis of the microbial metabolome by isotope dilution mass spectrometry using uniformly ¹³C-labeled cell extracts as internal standards. *Analytical Biochemistry* 336 (2), 164–171.
- Xia, J., Mandal, R., Sinelnikov, I.V., Broadhurst, D., Wishart, D.S., 2012. MetaboAnalyst 2.0 A comprehensive server for metabolomic data analysis. *Nucleic Acids Research* 40 (1), 127–133.

HORIZONTAL INTEGRATION: ANALYSES AND TOOLS

Contents

Network Modeling of Heterogeneous Datasets

Stochastic Analysis of Nongenetic Cell-to-Cell Heterogeneity

Systematic Methods to Interrogate Genetic Perturbations and Map Phosphorylation-Dependent Signaling

High-Content Screening in Cell Biology

Integrative Systems Biology

Network Modeling of Heterogeneous Datasets

E Fraenkel, Massachusetts Institute of Technology, Cambridge, MA, USA

MI Koeva, Whitehead Institute for Biomedical Research, Cambridge, MA, USA and Massachusetts Institute of Technology, Cambridge, MA, USA

© 2016 Elsevier Inc. All rights reserved.

Introduction

The cell is a highly complex system of molecules that carry out an astonishing number of functions. Individual molecules, including genes, RNAs, proteins, and metabolites are the building blocks of the cell, but most cellular activities only arise from interactions among these elements. Over the last 50 years, molecular biology has introduced many useful ways of describing combinations of proteins that function together. These descriptions include signaling and metabolic pathways, gene regulatory modules, etc. Pathways have been summarized in various publicly available resources, such as Gene Ontology (GO) ([The Gene Ontology Consortium, 2000](#)), KEGG ([Kanehisa and Goto, 2000](#)), Biocarta ([Nishimura, 2001](#)), Reactome ([Vastrik et al., 2007](#)), etc. Although current knowledge of pathways, especially in the context of human biology, remains incomplete, it serves as a useful starting point for many analyses.

Nevertheless, the functional interactions of genes and proteins can be highly dynamic, depending on context and condition. Technological innovations for high-throughput measurements have helped to fill in many missing components of pathways and to reveal the extent of these dynamic links. These methods include RNA sequencing (RNA-Seq), chromatin immunoprecipitation followed by sequencing (ChIP-Seq), mass spectrometry, and RNA interference screening. Each technology type provides a different snapshot of the state of a cell. RNA-Seq experiments capture transcriptional changes, ChIP-Seq data capture changes in transcription factor binding, RNAi screens assess the role of individual genetic perturbations, and drug compound screens measure the effects of individual chemical perturbagens on cellular state.

These technological advancements have created a need to move from analysis of individual pathways to global

interaction networks, a concept that emerged with the completion of the first genome projects in the 1990s. Mathematically, a network can be described as a graph, which consists of nodes (in this case molecules) and edges (interactions). There is a rich literature on algorithms for analyzing such graphs. In this article, we will examine some of the computational methods that have been developed to analyze biological networks ([Table 1](#)). We will emphasize those approaches that integrate multiple, disparate data types to provide testable hypotheses.

Overview

We have structured this article according to the complexity and diversity of the types of data integrated by each method. We have sampled from a range of methods with very different approaches to similar problems. We begin with algorithms that analyze a 'single data source' such as mRNA expression or metabolomic data in the context of a known protein–protein or protein–DNA interaction network. Next, we examine methods that analyze networks together with 'more complex data types,' such as somatic mutation and copy number variation data. Finally, we examine methods that 'integrate data types that capture a hierarchical flow of biological information.' For example, a chemical trigger, such as a DNA damage agent, may generate a transcriptional cellular response. Methods in this category reveal the signaling and regulatory sequence of steps that lead from the chemical perturbation to the activation or inhibition of the relevant transcription factors, which subsequently cause activation or repression of the expression of relevant target genes. Most of the methods in this group bridge genetic, chemical, or phosphoproteomic perturbations and their downstream gene

Table 1 Overview of network methods

<i>Method</i>	<i>Type</i>	<i>Separate genes/proteins</i>	<i>Data source</i>
jActiveModules	Simulated annealing	No	Gene expression
GXNA	Greedy algorithm	No	Gene expression
ClustEx	Clustering	No	Gene expression (time course)
SAMBA	Biclustering	No	Gene expression
cMonkey	Biclustering	No	Multiple data sources
NetWalk	Random walk	No	Gene expression
RegMod	Regression	No	Gene expression
Dendrix	Greedy approach; Markov chain Monte Carlo	No	Somatic mutation data
MeMo	Correlation analysis and statistical tests	No	Somatic mutation, copy number variation data, and gene expression
HotNet	Diffusion-based algorithm	No	Somatic mutation
TieDie	Diffusion-based algorithm	No	Somatic mutation and gene expression
ResponseNet	Network flow; linear programming	Yes	Multiple sources: gene expression, phosphroteomic, and genetic data
Prize-collecting Steiner Tree	Message-passing; linear programming	Yes	Multiple sources: gene expression, phosphroteomic, and genetic data
Prize-collecting Steiner Forest	Message-passing approach	Yes	Multiple sources: gene expression, phosphroteomic, and genetic data
CircadiOmics	N/A	Yes	Multiple sources: gene expression and metabolomics data
rFBA	Flux-balance	N/A	Metabolomics data
idFBA	Flux-balance	N/A	Metabolomics data
iFBA	Flux-balance	N/A	Metabolomics data
IOMA	Quadratic programming	N/A	Multiple sources: metabolomics and proteomic data
PARADIGM	Message-passing approach	Yes	Multiple sources

expression responses using protein–protein, protein–DNA, and other types, such as kinase–substrate interactions.

We close by examining a critical issue that has been largely ignored in most network approaches: To what extent can gene expression data be used as a proxy for protein levels? While most studies conflate these two, it is now abundantly clear from high-throughput analyses that the correlation between these data is very poor. We close by speculating on where a more realistic modeling of different types of data will lead the field.

Network Algorithms

Integrating a Single Data Source with an Existing Interaction Network

In this section, we describe several different approaches to the identification of functionally active groups of genes based on a single data source and a known interaction network. The tools we characterize in this category use a variety of methods, including simulated annealing, clustering and biclustering, as well as greedy, regression, and random walk-based approaches.

One of the first solutions to this problem was a ‘simulated annealing-based approach’ to finding ‘active’ subnetworks (Ideker *et al.*, 2002). The algorithm, called jActiveModules, uses gene expression data, protein–protein, and protein–DNA interaction data to identify connected groups of genes,

represented as nodes that exhibit significant differential expression under one or more conditions.

Briefly, the method can be summarized as follows: the algorithm starts with edges defined by known protein–protein and protein–DNA interactions and assigns an estimate of significance to each gene based on the change in its expression level, expressed as a *z*-score (Ideker *et al.*, 2002). The *z*-scores are used to search for high-scoring subnetworks using simulated annealing, where a subnetwork is assigned a combined *z*-score calculated from all the included genes. In the annealing step, a node is included and assigned an active/inactive state, which is initially toggled randomly. With each annealing iteration, the aggregate *z*-score of the subnetwork is evaluated and compared to the score from the previous iteration. If the aggregate score has improved, the node retains its toggled state; otherwise, it is kept toggled at a decreasing probability with each iteration (Ideker *et al.*, 2002). An advantage of this method is that various modifications can be applied to the annealing process, for example, multiple subnetworks can be evaluated simultaneously. However, a potential pitfall of this approach is the large size of the identified subnetworks. One alleviating solution to this problem may be to apply the process in a recursive manner to subdivide large subnetworks into more specific smaller subgraphs.

One alternative to simulated annealing methods is the ‘greedy’ class of algorithms. GXNA (Gene Expression Network Analysis) (Nacu *et al.*, 2007) is an example of such a greedy-based approach. GXNA uses mRNA expression and protein–protein interactions to identify functionally active gene subnetworks. Nodes are defined as genes, while edges

correspond to protein–protein interactions. The method first assesses the significance of differential gene expression for every node. Then, it selects a seed vertex and expands the subnetwork from the seed vertex by adding nodes connected to the vertices already in the subnetwork with the goal of maximizing the overall subnetwork score (Nacu *et al.*, 2007). The overall score is defined in one of two ways: as an average of individual gene test statistic scores or as a test statistic based on the average of individual gene expression values. The algorithm stops when the subgraph reaches a certain prespecified size. This approach resolves the issue of subnetwork size, but it suffers from another potential problem, more characteristic of greedy approaches, i.e., the possibility of getting stuck in local maxima.

‘Random walk-based methods,’ such as NetWalk and FunWalk, present another approach. NetWalk (Komurov *et al.*, 2010) is an example of a random walk-based algorithm, which uses the whole distribution of mRNA expression in an experiment in the context of network connectivity to identify functionally active groups of genes. Nodes represent individual genes, while edges correspond to protein–protein interactions. The algorithm assesses the centrality of genes in the network, but instead of using only a small set of differentially expressed (DE) genes revealed by the gene expression data, it uses the entire data distribution. Notably, the transition probabilities between nodes are biased according to the data values associated with individual genes, i.e., higher data value genes receive higher visitation probabilities, while lower data value nodes are assigned low visitation probability. The algorithm generates individual edge flux (EF) values, where an EF edge score represents how important the interaction is as measured by the data (Komurov *et al.*, 2010).

A completely different approach to the identification of disease-associated groups of genes, based on the integration of protein–protein interaction data and gene expression data is the formulation of the method RegMod (Qiu *et al.*, 2010). RegMod uses a ‘regression’ model with a diffusion kernel to efficiently identify significantly dysregulated modules of genes. In the context of this method, nodes represent genes, while edges represent protein–protein and protein–DNA interactions between pairs of genes (Qiu *et al.*, 2010). The weight of an edge between two genes is given by the Pearson correlation coefficient of the similarity between the expression profiles of those genes. Each node has a weight associated with the activity of its corresponding gene. However, while a node’s weight is usually the differential expression change of a gene between a pair of conditions, in this context, the weight of a given node is a function of the weights of its neighboring genes.

The final group of methods in this category is ‘clustering and biclustering’ algorithms. ClustEx (Gu *et al.*, 2010) is a two-step hierarchical clustering method that incorporates mRNA expression, especially time-course mRNA expression, and protein–protein interaction data. The goal of the method is to identify tight functional clusters of DE and potentially directly interacting genes. Nodes are defined as genes, while edges between nodes are defined as direct protein–protein interactions. The distance between directly interacting genes is inversely correlated with their level of co-expression, as measured by the mRNA expression data, while the distance between any pair of genes is the shortest path between them as identified using Dijkstra’s algorithm. All shortest paths are calculated in the first step of the algorithm,

and hierarchical clustering with average linkage analysis is used to identify clusters of DE genes. In the second step, the method incorporates other (intermediate) genes, not found in the clustering step, into the gene modules. These intermediate genes could be either DE genes, genes on the *k*-shortest paths between DE genes, or first neighbors of either of those types. Finally, the algorithm determines the statistical significance of the modules by edge scrambling, i.e., random rewiring of the edges in the network to different node pairs.

While normal clustering generally compares conditions by evaluating similarity of expression across all genes in a sample, biclustering, by contrast, compares conditions using only a subset of the genes that maximize the similarity of the conditions. As a result, biclustering can identify a subset of genes that show a similar pattern over a subset of conditions and here we focus on a couple of more recent biclustering algorithms.

SAMBA (Tanay *et al.*, 2002) is a network-based biclustering algorithm, which uses a mixture of graph-theoretic and statistical models to identify groups of genes with similar patterns of expression over a subset of samples or conditions. SAMBA defines a vertex for every gene and condition, and an edge between a pair of vertices if the pair exhibits significant changes in expression in the given condition (Tanay *et al.*, 2002). The problem is then reduced to identifying a maximum weight subgraph, or more specifically a maximum edge weight biclique, where the weights on vertex pairs in the subgraph rely on a likelihood-based score and correspond to a measure of the significance of the subgraph.

The cMonkey (Reiss *et al.*, 2006) method is an iterative biclustering algorithm, which builds on gene expression, but also incorporates additional components, such as motif co-occurrence, metabolic pathway co-occurrence, or protein network association, etc. The overall score is a mixture of the scores of individual components, where each element is weighed individually to assess its contribution to the final score. A *p*-value is measured for each scoring unit; for example, if the sequence component of the score is assigned a nonzero contribution weight, the ‘motif *p*-value’ is defined as the significance of co-occurrence of motifs within the member genes of a subcluster (Reiss *et al.*, 2006).

Once an initial seed subcluster of conditions and genes is identified, the cluster is improved by iteratively calculating the conditional probability of membership for each gene and condition and then sampling from it to add to or remove genes and conditions from the cluster. Once a cluster stabilizes, the algorithm identifies a new ‘seed’ subcluster and follows the same process.

Metabolic Networks

The methods described above largely use high-throughput protein–protein and protein–DNA data to build the underlying networks. Recently, there have also been interesting attempts to understand how changes in gene expression alter metabolic networks. Historically, metabolic networks have been analyzed by completely different methods from ‘omic’ data. The main goal in metabolic network analysis has been predicting flux through reactions using linear programming.

These methods determine the optimal flow through a network where the nodes represent metabolites and the edges represent chemical transformations. In order to determine the optimal flux through the network, one must specify an objective function. The choice of this function has been a subject of some debate (Covert and Palsson, 2002).

One example of a very context-specific algorithm is Circadiomics (Patel *et al.*, 2012), which identifies network connections between transcriptional changes and metabolomic data, specifically associated with circadian rhythm (Covert and Palsson, 2002). Methods including rFBA (flux-balance analysis) (Lee *et al.*, 2008), integrated dynamic FBA (idFBA) (Covert *et al.*, 2008), and integrated flux-balance analysis (iFBA) (Yizhak *et al.*, 2010) use enzyme expression levels to influence the flux through metabolic pathways. Integrative omics-metabolic analysis (IOMA) (Abreu *et al.*, 2009) integrates proteomic and metabolic data in a similar fashion.

Complex, Cancer-Related Genetic Networks

The identification of significantly dysregulated pathways in different cancer subpopulations has presented a tremendous therapeutic challenge, to which integration of multiple layers of information, such as copy number variation, somatic mutation information, and mRNA expression present a possible solution. The sequencing of cancer genomes raises the possibility of using somatic mutation information collected from a large number of patient samples to identify significantly dysregulated pathways and relevant somatic driver mutations in different cancer subpopulations. Several algorithms have been developed to address the challenges and we discuss some of these tools in this section. We first introduce the algorithms Dendrix (Vandin *et al.*, 2012) and MeMo (Ciriello *et al.*, 2012), and then proceed with the network propagation and diffusion-based algorithm HotNet (Vandin *et al.*, 2011) and TieDie (Paull *et al.*, 2013).

Dendrix (De Novo Driver Exclusivity) describes a pair of algorithms that searches for subsets of genes that show high mutational load, i.e., exhibit mutations across most patient samples in a dataset (high coverage), but in a mostly mutually exclusive manner (high exclusivity) (Vandin *et al.*, 2012). The logic behind these requirements is that important cancer developmental pathways will show perturbations across a large number of patients for a given cancer type. However, while mutations in a single gene may be sufficient to dysregulate the activity of the pathway, multiple mutations may be unlikely to occur in a single gene and patient. This idea is formulated as a maximum coverage exclusive submatrix problem, where the matrix represents mutation presence in a gene (column) for a patient (row), and the problem refers to identifying mutually exclusive submatrices with the largest number of nonzero rows (patients) (Vandin *et al.*, 2012). The trade-off between strict mutual exclusivity and maximized patient coverage leads to a reformulation of the question as a maximum weight submatrix problem. Dendrix uses both a greedy approach and Markov chain Monte Carlo approach to identify relevant exclusive subsets of genes with a high mutational load (Vandin *et al.*, 2012).

An alternative method, MeMo (Mutual Exclusivity Modules in Cancer) uses a graph-theoretic approach to a very similar

problem (Ciriello *et al.*, 2012). Similar to Dendrix (Vandin *et al.*, 2012), the goal of this method is to identify sets of driver genes that show significant dysregulation and mutation across most conditions (i.e., patient samples), are likely to belong to the same pathway, but show alterations in a mutually exclusive manner. The algorithm uses a four-step approach, which starts with the application of several filters to identify significantly mutated genes (using somatic mutation data), or genes that fall within genomic regions with a recurrently altered copy number (using copy number variation data), both of which have to show concordant changes in mRNA expression (using mRNA expression data). These changes are encoded in a binary event matrix and the method then identifies pairs of altered genes that are likely to belong to the same pathway, based on known pathway and interaction database information. The algorithm builds a network of all such gene pairs, extracts maximal cliques, and measures the mutual exclusivity of individual cliques.

One potential caveat of some of these methods is study bias, where well-studied genes and proteins are most likely to be incorporated into a significant subgraph primarily because of their high level of connectivity, i.e., random set of genes can easily include these well-connected hub-like proteins (Vandin *et al.*, 2011). Methods, such as the ones we describe next – HotNet (Vandin *et al.*, 2011) and TieDie (Paull *et al.*, 2013) – attempt to remedy this problem by using network diffusion, where information diffuses from network nodes in a manner similar to heat diffusion.

HotNet (Vandin *et al.*, 2011) builds on the concept of identifying subgraphs of DE genes. Its goal is to identify significantly mutated subnetworks, i.e., groups of genes mutated more than expected by chance. As in other methods, here nodes designate proteins and their corresponding genes, while edges represent protein–protein or protein–DNA interactions. The algorithm defines an influence diffusion-based measure on a pair of genes, such that a gene influences another gene if they are proximal in the graph and if there are only a small number of paths between them. The rationale behind the latter requirement relates to the desire of avoiding the identification of subnetworks of mutated genes connected primarily through uninteresting high degree (hub-like) proteins. HotNet builds its influence graph using only genes that have been assayed for mutations, but it takes into account the connectivity of all genes. It identifies subgraphs of genes mutated in a large number of samples (patients), but the number of mutations associated with each gene can be additionally used to weigh the influence score.

In all the methods described in this section, network propagation and flow-like approaches are generally used to find links among mutations that target a pathway leading to cancer. However, related approaches can also be used to search for links from mutations to their transcriptional consequences. In the next and last section of this review, we explore methods for integration of data types that capture a hierarchical flow of biological information.

Methods for Integration of Data Types that Capture a Hierarchical Flow of Biological Information

ResponseNet (Yeger-Lotem *et al.*, 2009) seeks to identify a minimum set of edges that link a set of molecular causes

(mutations, for example) and consequences (gene expression changes). This network-based optimization approach uses linear programming techniques to solve a minimum cost flow algorithm. In this network setting, vertices represent either DE genes (from transcriptional data), or significant gene or protein hits from genetic or proteomic screening efforts. Edges represent either a regulatory interaction between a transcription factor protein and a target gene, or a protein–protein interaction between two proteins. The ‘flow’ represents the set of weighted protein–protein interaction network edges selected by the algorithm that best explain the consequences with the causes. The method requires randomization to identify the most surprising nodes and interactions, given the underlying interaction network.

Many perturbations induce both a specific cellular response and a general response to stress. A multi-commodity flow implementation of ResponseNet, SAMNet (Gosline *et al.*, 2012) is designed to distinguish these two types of pathways. It uses data from multiple conditions and applies the flow algorithm with a constraint that requires that two experimental conditions must ‘share’ each edge. This forces the algorithm to identify condition-specific pathways and reduces the need for excessive randomization.

TieDie (Tied Diffusion of Interacting Events) (Paull *et al.*, 2013) uses the concept of heat diffusion to identify novel genes that bridge between relevant starting groups of genes. For example, given a source set of genes that show significant

mutational changes in a specific cancer type and a target set of genes that exhibit significant transcriptional changes in the same context, TieDie (Paull *et al.*, 2013) uses two diffusion processes to find the genes that, given a specific pathway or protein interaction structure, lie in the intersection of the paths of the diffusion processes. The method is not restricted to pairs of processes and can be generalized to multiple starting gene sets. Theoretically, it can also incorporate more general cellular features such as metabolites.

A more general approach is based on Prize-collecting Steiner Tree (PCST) methods, and removes the requirement that all data from one experiment (genetics) lie upstream of the data from another type (gene expression). The PCST-based methods allow for other network topologies. The PCST approach has been applied to biological problems by Dittrich *et al.* (2008), Huang and Fraenkel (2009), and Bailly-Bechet *et al.* (2010). Edges are given positive costs representative of a quantitative measure of the quality of the interactions, while nodes are given positive prizes reflective of the importance and significance of a protein as a hit (Figure 1). The problem is then reduced to identifying a subgraph that balances the minimal cost of included edges and the maximal prize of included nodes. The algorithm can be implemented using an efficient message-passing methodology (Bailly-Bechet *et al.*, 2010a) or a linear programming approach (Dittrich *et al.*, 2008; Huang and Fraenkel, 2009). The Prize-collecting Steiner Forest (Tuncbag

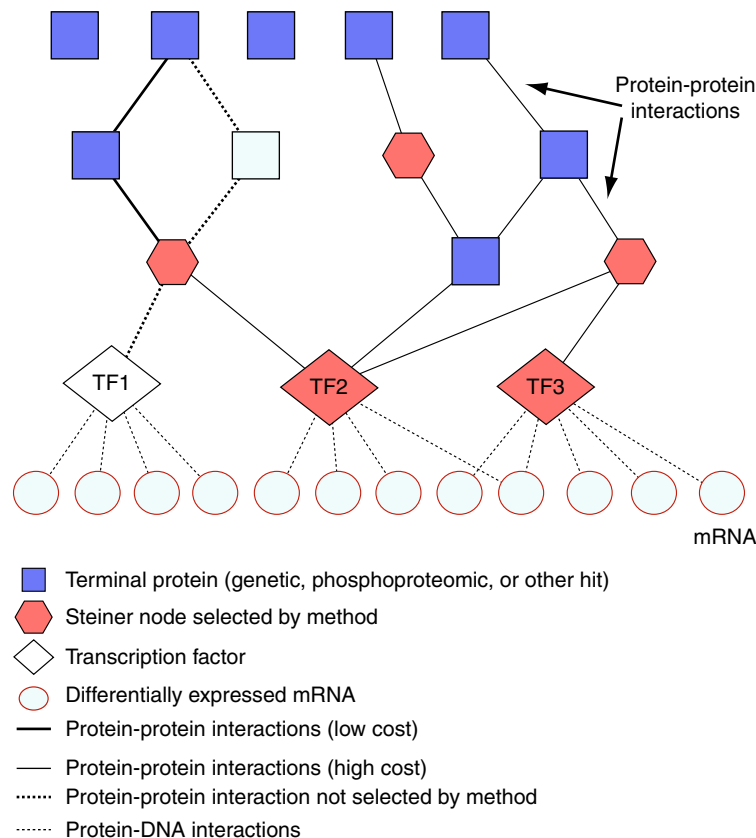


Figure 1 Overview of prize-collecting Steiner approaches.

et al., 2013) represents a generalization of these approaches that allows for discovery of several unconnected networks. In addition, Gitter *et al.* (2014) have introduced a method for simultaneously analyzing multiple patients with similar, but distinct data.

Distinguishing Proteins and Genes

Almost all of the methods we described above use a single node to represent both a protein and the gene that encodes it. All data about the gene and protein are then localized to the same place in the network, regardless of what they represent. These data could include a change in the level of the mRNA encoding the gene, the level of the protein, a point mutation, a posttranslational modification, etc. Similarly, many of the methods also treat all interaction types as equivalent. An edge may represent a direct, physical interaction or a highly indirect connection, such as co-expression or a genetic interaction. Initially, these choices were a matter of necessity.

When microarrays were the only high-throughput method for measuring molecular activity, it was reasonable to treat these data as a proxy for protein levels. However, with the advent of high-throughput proteomic methods we can directly ask whether mRNA levels predict protein levels. Initial analyses using microarray data showed that the correlation existed, but was weak (Tuncbag *et al.*, 2013). In fact, for any particular level of mRNA expression, there was a 1000-fold range in protein levels. Subsequent RNA-Seq analyses have demonstrated that this technique is only marginally better at predicting protein levels; the best correlation coefficient in a recent study was only 0.54 (Ning *et al.*, 2012).

Given the poor correlation between mRNA expression and proteomic data, mapping all biological entities to the same place in the interaction network is far from ideal. There are several approaches for dealing with this problem.

In the network algorithms published by our group (Huang and Fraenkel, 2009; Tuncbag *et al.*, 2013; Gitter *et al.*, 2014) (including ResponseNet and PCST described above), we separate mRNAs and the proteins they encode by allocating a different node for each one. Then protein–protein interactions can only occur between protein nodes, while protein–DNA interactions connect mRNA nodes to the protein network. The protein–DNA interactions can be either derived directly from ChIP-chip or ChIP-Seq experiments that measure such interactions (Yu *et al.*, 2009; MacIsaac *et al.*, 2010) or computationally from epigenomic and sequencing data (Huang and Fraenkel, 2009).

The PARADIGM algorithm (Vaske *et al.*, 2010) introduces a much more general alternative strategy. They encode previously characterized pathway topology in the network where individual cellular states or identities, such as a given gene, mRNA, or protein are represented as variables, and the interactions and relationships (pathway structure) that relate these states or identities are represented as factors. In this approach, for example, a gene is connected to its encoded protein by a factor node that can explicitly represent any known regulation of translation. Given enough prior data regarding regulatory steps, a factor graph could be used to describe arbitrarily complex biological processes with separate nodes for each type of data. However, current datasets are not rich enough to allow

the *de novo* discovery of the parameters for a complicated factor graph in the absence of such prior knowledge.

Outlook

What can we learn from modeling biological networks? The approaches we reviewed have already proved very useful in finding connections among diverse data. One of the great outstanding challenges will be to use these networks to make specific predictions regarding these systems in new settings. Transitioning from network models that explain existing data to ones that predict the results of future experiments will be extremely challenging. We expect that progress will depend, in part, on finding increasingly realistic ways of modeling distinct types of data that are faithful to underlying physical structure of true biological networks.

See also: Horizontal Integration: Networks: Understanding of 'Networks' *In Vitro* and/or *In Vivo*

References

- Abreu, R.S., Penalya, L., Marcotte, E., Vogel, C., 2009. Global signatures of protein and mRNA expression levels. *Molecular BioSystems*. doi:10.1039/b908315d.
- Bailey-Bechet, M., Borgs, C., Braunstein, A., *et al.*, 2010. Finding undetected protein associations in cell signaling by belief propagation. *Proceedings of the National Academy of Sciences of the United States of America* 108, 882–887.
- Ciriello, G., Cerami, E., Sander, C., Schultz, N., 2012. Mutually exclusivity analysis identifies oncogenic network modules. *Genome Research* 22 (2), 398–406.
- Covert, M.W., Palsson, B.O., 2002. Transcriptional regulation in constraints-based metabolic models of *Escherichia coli*. *Journal of Biological Chemistry* 277 (31), 28058–28064.
- Covert, M.W., Xiao, N., Chen, T.J., Karr, J.R., 2008. Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. *Bioinformatics* 24 (18), 2044–2050.
- Dittrich, M.T., Klau, G.W., Rosenwald, A., *et al.*, 2008. Identifying functional modules in protein–protein interaction networks: An integrated exact approach. *Bioinformatics* 24, i223–i231.
- Gitter, A., Braunstein, A., Pagnani, A., *et al.*, 2014. Sharing information to reconstruct patient-specific pathways in heterogeneous diseases. *Pacific Symposium on Biocomputing* 19, 39–50.
- Gosline, S.J., Spencer, S., Ursu, O., Fraenkel, E., 2012. SAMNet: A network-based approach to integrate multi-dimensional high throughput datasets. *Integrative Biology* 4, 1415–1427.
- Gu, J., Chen, Y., Li, S., Li, Y., 2010. Identification of responsive gene modules by network-based gene clustering and extending: Application to inflammation and angiogenesis. *BMC Systems Biology* 4, 47.
- Huang, S.S., Fraenkel, E., 2009. Integrating proteomic, transcriptional, and interactome data reveals hidden components of signaling and regulatory networks. *Science Signaling* 2, ra40.
- Ideker, T., Ozier, O., Schwikowski, B., Siegel, A.F., 2002. Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics* 18 (Suppl 1), S233–S240.
- Kanehisa, M., Goto, S., 2000. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* 28, 27–30.
- Komurov, K., White, M.A., Ram, P.T., 2010. Use of data-biased random walks on graphs for the retrieval of context-specific networks from genomic data. *PLoS Computational Biology* 6 (8), e1000889. doi:10.1371/journal.pcbi.1000889.
- Lee, J.M., Gianchandani, E.P., Eddy, J.A., Papin, J.A., 2008. Dynamic analysis of integrated signaling, metabolic, and regulatory networks. *PLoS Computational Biology* 4 (5), e1000086. PMID: 2377155.
- MacIsaac, K.D., Lo, K.A., Gordon, W., *et al.*, 2010. A quantitative model of transcriptional regulation reveals the influence of binding location on expression. *PLoS Computational Biology* 6 (4), e1000773.

- Nacu, S., Critchley-Thorne, R., Lee, P., Holmes, S., 2007. Gene expression network analysis and applications to immunology. *Bioinformatics* 23 (7), 850–858.
- Ning, K., Fermin, D., Nesvizhskii, A., 2012. Comparative analysis of different label-free mass spectrometry based protein abundance estimates and their correlation with RNA-Seq gene expression data. *Journal of Proteome Research* 11 (4), 2261–2271.
- Nishimura, D., 2001. *Biotech Software & Internet Report* 2 (3), 117–120. doi:10.1089/152791601750294344.
- Patel, V.R., Eckel-Mahan, K., Sassone-Corsi, P., Baldi, P., 2012. CircadiOmics: Integrating circadian genomics, transcriptomics, proteomics and metabolomics. *Nature Methods* 9 (8), 772–773.
- Paull, E., Carlin, D., Niepel, M., *et al.*, 2013. Discovering causal pathways linking genomic events to transcriptional states using Tied Diffusion Through Interacting Events (TieDIE). *Bioinformatics* 29 (21), 2757–2764.
- Qiu, Y., Zhang, S., Zhang, X., Chen, L., 2010. Detecting disease associated modules and prioritizing active genes based on high throughput data. *BMC Bioinformatics* 14, 26. doi:10.1186/1471-2105-11-26.
- Reiss, D., Baliga, N., Bonneau, R., 2006. Integrated biclustering of heterogeneous genome-wide datasets for the inference of global regulatory networks. *BMC Bioinformatics* 7, 280. doi:10.1186/1471-2105-7-280.
- Tanay, A., Sharan, R., Shamir, R., 2002. Discovering statistically significant biclusters in gene expression data. *Bioinformatics* 18 (Suppl 1), S136–S144.
- The Gene Ontology Consortium, 2000. Gene ontology: Tool for the unification of biology. *Nature Genetics* 25, 25–29.
- Tuncbag, N., Braunstein, A., Pagnani, A., *et al.*, 2013. Simultaneous reconstruction of multiple signaling pathways via the prize-collecting steiner forest problem. *Journal of Computational Biology* 20 (2), 124–136.
- Vandin, F., Upfal, E., Raphael, B.J., 2011. Algorithms for detecting significantly mutated pathways in cancer. *Journal of Computational Biology* 18 (3), 507–522.
- Vandin, F., Upfal, E., Raphael, B.J., 2012. De novo discovery of mutated driver pathways in cancer. *Genome Research* 22 (2), 375–385.
- Vaske, C., Benz, S., Sanborn, J.Z., *et al.*, 2010. Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics* 26 (12), i237–i245.
- Vastrik, I., D'Eustachio, P., Schmidt, E., *et al.*, 2007. Reactome: A knowledge base of biologic pathways and processes. *Genome Biology* 8, R39. doi:10.1186/gb-2007-8-3-r39.
- Yeger-Lotem, E., Riva, L., Su, L.J., *et al.*, 2009. Bridging high-throughput genetic and transcriptional data reveals cellular responses to alpha-synuclein toxicity. *Nature Genetics* 41, 316–323.
- Yizhak, K., Benyamini, T., Liebermeister, W., Rupp, E., Shlomi, T., 2010. Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model. *Bioinformatics* 26 (12), i255–i260.
- Yu, M., Riva, L., Xie, H., *et al.*, 2009. Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. *Molecular Cell* 36 (4), 682–695.

Stochastic Analysis of Nongenetic Cell-to-Cell Heterogeneity

CC Wang and KA Janes, University of Virginia, Charlottesville, VA, USA

© 2016 Elsevier Inc. All rights reserved.

Glossary

Fluorescence-activated cell sorting (FACS) A form of flow cytometry for sorting cell subpopulations into multiple containers based upon the light scattering and fluorescent characteristics in each cells.

Intermembrane space (IMS) The region between the inner and outer membrane of a mitochondrion.

Laser capture microdissection (LCM) A method for isolating specific cells of interest from microscopic regions of tissue/cells.

Mitochondrial outer membrane permeabilization (MOMP) MOMP is induced by intrinsic apoptotic

signaling and allows for release of pro-death factors into the cytosol.

Nuclear export signal (NES) An amino acid sequence that tags a protein for export from the cell nucleus to cytoplasm by nuclear transport.

Nuclear localization signal (NLS) An amino acid sequence that tags a protein for import into the cell nucleus by nuclear transport.

Pleckstrin homology (PH) domain A protein domain that can bind phosphatidylinositol lipids and plays a role in recruiting proteins to cell membrane.

Introduction

Genetically identical cells in a homogeneous environment can exhibit a range of phenotypes. The sources of nongenetic heterogeneity can be stochastic through random fluctuations inside the cells, or they can be regulated through feedback mechanisms. Cell-to-cell heterogeneity causes cells to vary in their responsiveness to stimuli and has profound effects on their phenotypes (Magee *et al.*, 2012; Marusyk *et al.*, 2012).

Phenotypic heterogeneity arises in cell fates ranging from pluripotent stem cells (Buganim *et al.*, 2012; Hanna *et al.*, 2009; Takahashi and Yamanaka, 2006) to apoptosis (Spencer *et al.*, 2009). Such functional heterogeneity is involved in nondeterministic cell fates during development (Losick and Desplan, 2008) and is thought to be suppressed in mature tissues. In cancer, however, nongenetic heterogeneity has been widely described (Quintana *et al.*, 2008; Visvader and Lindeman, 2008). Cell-to-cell heterogeneity results in non-uniform responses to anticancer agents and fractional killing, creating a major obstacle for cancer therapy (Fallahi-Sichani *et al.*, 2013; Sharma *et al.*, 2010).

The dynamics of cell responses further increases the complexity of cell-to-cell heterogeneity. Stochastic state transitions are frequently encountered in both transformed and non-transformed cells (Chaffer *et al.*, 2011; Gupta *et al.*, 2011; Wang *et al.*, 2014). The ability to transition reversibly between states establishes a steady-state distribution of heterogeneity in cancer cell populations (Gupta *et al.*, 2011). During the process of tumor metastasis, disseminated cells switch between epithelial and more invasive states (Mani *et al.*, 2008). These studies suggest that intra-tumor heterogeneity cannot be completely explained by biochemical noise and might in part reflect cellular state transitions that are regulated.

The transition between distinct states is mediated by the regulation of gene expression (Chaffer *et al.*, 2013; Wang *et al.*, 2014). Several studies suggest that activation of state-switching programs is triggered by contextual signals. For example, transforming growth factor beta (TGF β) signaling has been

observed to trigger the oscillation between gene expression programs in basal-like epithelial cells (Wang *et al.*, 2014) to establish a tumor-propagating state (Shipitsin *et al.*, 2007). TGF β -driven cell-state plasticity in basal breast cancer was achieved through the activity of a key transcription regulator (Chaffer *et al.*, 2013). Signaling and transcription mechanisms may also work together to create feedback loops that result in dynamic regulatory heterogeneity (Ashall *et al.*, 2009; Hoffmann *et al.*, 2002; Tay *et al.*, 2010; Wang *et al.*, 2014).

Cell-to-cell heterogeneity exists at multiple levels, including gene expression states, signal-transduction states, and cellular phenotypes. In this article, we discuss variability that occurs in signal transduction and gene expression. We also review the experimental tools used for measuring the molecular processes underlying the observed heterogeneity. Finally, we will discuss recent systems biology approaches that connect different levels of variability and their role in cell-fate decisions.

Cell Signaling

Cells sense extracellular signals and transfer information to make appropriate decisions. These signal-transduction networks are subject to various sources of noise. In addition, cellular function is dynamic, and the observed heterogeneity may also reflect asynchronous transmission of signals in the overall population. Here, we discuss the cell-to-cell heterogeneities in signaling that arise from differences in molecular states among cells and the stochastic events or unsynchronized dynamics of the relevant signaling networks.

Molecular-State Differences

Diverse states among cells – including the concentration levels, subcellular locations, and activities of key signal regulators – can generate a broad distribution of signaling responses. For example, in mammary epithelial cells, Akt activation is bimodal in response to epidermal growth factor (EGF)

stimulation (Yuan *et al.*, 2011). This on-or-off heterogeneity correlates with the abundance of phosphoinositide 3-kinase (PI3K). Only cells with high levels of PI3K protein can activate Akt. Variable concentrations of upstream signal proteins, including PI3K, PTEN, and cMet, are also responsible for the heterogeneous activation of Akt in response to hepatocyte growth factor (HGF) (Meyer *et al.*, 2012). The variable concentration of signaling components is one of the key determinants for the heterogeneous PI3K signaling response.

Another important pathway usually misregulated in human cancers is the Ras/extracellular signal-regulated kinase (ERK) pathway. The kinetics of ERK phosphorylation and nuclear translocation govern different cellular outcomes (Ahmed *et al.*, 2014; Marshall, 1995). Cell-to-cell heterogeneity is observed in both signal strength and dynamics of the Ras/ERK pathway. For example, ERK2 shows variability in its basal nuclear levels (Cohen-Saidon *et al.*, 2009). Upon EGF stimulation, the relative increase in nuclear ERK2 level is proportional to the basal level in each cell (Cohen-Saidon *et al.*, 2009). ERK can be activated in discrete, asynchronous pulses under conditions of steady-state EGF stimulation spanning the physiological range (Albeck *et al.*, 2013; Shankaran *et al.*, 2009). The dynamics and amplitude of ERK signaling are determined by the activity of upstream signaling proteins, EGFR and MEK, respectively (Albeck *et al.*, 2013). This suggests the thresholds of signaling protein activity is different between cells, by which Ras/ERK pathway exhibits strong variability and influences the cell-fate decision to undergo proliferation.

Cell-to-cell variability is also observed in ligand-mediated cell death. For example, tumor necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL) initiates apoptosis in mammalian cells with a pronounced temporal variability (Spencer *et al.*, 2009). Some cells die within minutes, while others die several hours later. The time-to-death is highly correlated within newly born sister cells whose behavior and total protein levels are synchronized. The primary determinant of variability in time-to-death was associated with the rate of caspase-mediated activation of BID (Albeck *et al.*, 2008; Spencer *et al.*, 2009). The correlated apoptosis of sister cells diverges as new proteins are produced, suggesting the relationship between signal-induced apoptosis and total protein levels. A follow-on study related to this work showed that the cell subpopulation surviving the TRAIL stimulus regenerates the sensitivity and death-time distribution of the parental population upon growth for several days in the absence of TRAIL (Flusberg *et al.*, 2013). This suggests that some cell-state variability arises from stochastic fluctuations in mRNA and protein production (Munsky *et al.*, 2012).

Stochastic Events in Signaling

Signal transduction is a complex process of protein interactions that occur after environmental stimulation. Many of these steps are transient and rely on relatively low-affinity interactions between binding partners (Pawson and Nash, 2003; Zheng *et al.*, 2013). The human interactome contains approximately 130 000 binary interactions, most of which are uncharacterized (Venkatesan *et al.*, 2009). Specific interactions between recognition domains within a signaling pathway

must therefore compete against the bulk of nonspecific interactions. Stochastic fluctuations in cell signaling thus could be partly attributed to the lack of sufficient specificity in protein-protein interactions (Ladbury and Arold, 2012).

Stochastic fluctuations in signaling cause a constant 'background noise' in the network. The background can consist of low signals through random transient protein interactions. A functional signal must prevail above this noise threshold. Background noise therefore limits the information capacity of signaling pathways. Information captured by individual pathways induced by TNF (Cheong *et al.*, 2011), platelet-derived growth factor (PDGF) (Cheong *et al.*, 2011), and EGF (Cohen-Saidon *et al.*, 2009) has been reported to be limited by noise. Although the noise from nonspecific interaction is unavoidable, negative feedback loops are able to reduce the amount of noise that corrupts a signal (Cheong *et al.*, 2011). In addition, feedback loops in signaling networks may serve as built-in reset mechanisms that preserve the dynamic capabilities for future stimuli (Yu *et al.*, 2008). Given that signaling changes are a greater determinant of cellular outcomes than signal strength (Janes *et al.*, 2008), signaling networks with higher-level connections, such as crosstalk and feedback loops (Cirit *et al.*, 2010; Wang *et al.*, 2009), maybe more effective than single pathways at signal transfer. The analysis of noise that has been applied to single pathways can be extended to different network structures to determine what signaling architectures can function reliably despite noise.

Signaling Oscillations

Signaling networks are now often viewed as a self-regulating system of interlinked feedback loops (Cirit *et al.*, 2010; Sachs *et al.*, 2005; Wang *et al.*, 2009). Feedback loops can generate nonlinear dynamics and oscillation in cell signaling. Oscillations have been observed in key signal regulators, including intracellular Ca^{2+} (Dolmetsch *et al.*, 1997), surface receptors Notch (McMahon *et al.*, 2008) and TGFBR3 (Wang *et al.*, 2014), protein kinase ERK (Shankaran *et al.*, 2009), and transcription factors p53 (Lahav *et al.*, 2004), STAT3 (Yoshiura *et al.*, 2007), Smad2 (Wang *et al.*, 2014), and NF- κ B (Nelson *et al.*, 2004). Since cells do not oscillate in synchrony, the asynchronous population exhibits cell-to-cell variability.

TNF α treatment induces NF- κ B nuclear translocation. NF- κ B regulates transcription of genes, including several that can feed back to regulate NF- κ B localization (Hoffmann and Baltimore, 2006). One of these feedbacks is mediated by I κ B α , which upon binding to NF- κ B in the nucleus shuttles NF- κ B back to the cytoplasm (Nelson *et al.*, 2004). Earlier population-level studies demonstrated a temporal profile of NF- κ B activity (Cheong *et al.*, 2006). Increasing the concentration of TNF α leads to a shortened delay in NF- κ B nuclear translocation. Single-cell studies, however, showed synchronous cycles of NF- κ B translocation (Ashall *et al.*, 2009). Timed fluctuations in TNF α resulted in differences in amplitude in NF- κ B translocation. Recent studies that measure NF- κ B activities in thousands of live cells revealed that transcriptional activity is heterogeneous at lower TNF doses (Tay *et al.*, 2010). Cells process fold change signals to create digital outputs for early gene expression (Lee *et al.*, 2014). Signal heterogeneity observed in fixed cells may not only be caused by intrinsic

stochasticity, but also by regulatory elements in the signaling networks.

Oscillations in cell signaling have an important role in cell function (Cai *et al.*, 2008; Purvis and Lahav, 2013). Recent studies in human epithelial cells reveal that TGFBR3 activity is built into a signaling-transcriptional circuit consisting of multiple negative feedback loops (Wang *et al.*, 2014). When the circuit is excited by TGF β family ligands, TGF β signaling displays a damped oscillation and cells transition spontaneously between gene expression states. This dynamic oscillation is important for normal morphogenesis in 3D acini and may be engaged by breast premalignancies. Strong links between signal oscillation and gene expression suggest the important roles of signaling feedback loops in regulating nongenetic heterogeneity.

Variability in Gene Expression

Transcriptome-wide studies of cell-to-cell variability at the mRNA or protein level in *Saccharomyces cerevisiae* reveal that variability in protein expression is dominated by the stochastic production and destructions of mRNA (Bar-Even *et al.*, 2006; Newman *et al.*, 2006). Such transcriptional variability could reflect the upstream influence from signaling pathways described above, as well as intrinsic stochasticity in transcriptional programs (Elowitz *et al.*, 2002). Intrinsic stochasticity could be attributed to two elements. The first is the probabilistic synthesis and decay of mRNA transcripts, because of the small copy numbers of many mRNA species (Suter *et al.*, 2011). Second, mRNAs are synthesized in short but intense bursts of transcription. This transcription burst is intrinsically random (Raj *et al.*, 2006) and is controlled by stochastic transitions between different regulatory states (Singh *et al.*, 2012). Bursts are correlated between genes that are located proximally (Raj *et al.*, 2006), suggesting that random events of chromatin opening and closing may be responsible for switching between transcriptionally active and inactive states.

Given the discrete nature of transcription, it is clear that assessing gene expression in single cells is critical to gain a better understanding of cell-to-cell variability. With the development of single-cell transcriptomics (Deng *et al.*, 2014; Jaitin *et al.*, 2014; Picelli *et al.*, 2014; Shalek *et al.*, 2013), studies of gene expression and regulation in single cells will improve inferences of gene-regulatory networks (Kim *et al.*, 2009; Wang *et al.*, 2014).

Single-Cell Analysis

Observing cell-to-cell heterogeneity requires the ability to measure biological readouts at the level of single cells. Variability occurs at many levels that vary across timescales. When cells detect extracellular signals, the early response in cell signaling can occur within seconds or minutes. The consequent gene expression changes lead to new cell phenotypes in hours to days. Cell signaling and transcription are highly dynamic and involve many elements, posing a major challenge to quantify them. Single-cell analysis has emerged as an active area of research now that new technologies have made it

possible to measure single-cell processes at the systems level. Here, we review a handful of the experimental tools that have been used to measure different biochemical events of cell signaling and transcription in single cells.

Single-Cell Transcriptomics

Single-cell transcriptomics has emerged as a potential technology to explore gene expression heterogeneity (Eberwine *et al.*, 2014; Sandberg, 2014). Advances in single-cell isolation as well as whole-transcriptome amplification and analysis platforms, such as microarrays and more recently RNA sequencing (RNA-Seq), have paved the way for high-resolution analysis.

Technical limitations

Single-cell transcriptomics relies on the reverse transcription of RNA to complementary DNA (cDNA) and subsequent amplification by polymerase chain reaction (PCR) or *in vitro* transcription (IVT) before deep analysis. A human cell contains ~ 1 pg of mRNA. 85% of transcripts have fewer than 100 copies per cell and many of them are present at only 5–20 copies per cell (Macaulay and Voet, 2014). Reverse transcription, the initial step of most RNA analysis, is likely one of the limiting factors in quantification of single-cell transcripts. It is estimated that on average only 5–25% of all RNA molecules are converted into amplifiable cDNA, while transcripts with 10 copies or more per cell were reliably quantified (Islam *et al.*, 2012).

Microarrays and RNA-Seq require nanogram quantities of DNA as the starting material. Whole-transcriptome amplification still introduces biases in representation by favoring high-abundance transcripts. The biases originate from the need for high amplification from the small amount of RNA available from single cells. Microfluidics-based detection platforms seek to improve sensitivity by reducing reagent volumes to increase effective concentrations (Warren *et al.*, 2006). However, they still require exponential preamplification to achieve single-cell sensitivity of low-abundance targets (Dalerba *et al.*, 2011). PCR-based amplification methods have the potential for nonlinear amplification, resulting in the distortion of the relative abundance of transcripts (Raj and van Oudenaarden, 2009). IVT amplification may arguably avoid such problem through linear amplification (Eberwine *et al.*, 1992). However, IVT imposes a minimal starting amount of ~ 400 pg of RNA for sufficient yield and a new strategy, such as barcoding and pooling samples (Hashimshony *et al.*, 2012), may need to be applied in the procedure to overcome this limitation.

A unique strategy, termed stochastic profiling, has been reported to remove the bottleneck of small amounts of input material (Janes *et al.*, 2010; Wang and Janes, 2013). Stochastic profiling uses not a single-cell as the starting material but rather small pools of 10 cells randomly selected from a population. Fluctuations in the 10-cell averages are then analyzed statistically to reveal heterogeneously regulated transcripts and transcriptional programs. Additional statistical and bioinformatic analyses allow for the quantification of cell-to-cell heterogeneity in gene expression (Bajikar *et al.*, 2014).

Finally, single-cell transcriptomics requires the isolation of RNA from the single cells. Most high-throughput cell isolating

tools to date require cells in suspension. Spatial information is lost when dissociating cells from tissues or cultures using flow cytometry and microfluidic platform. Spatially defined single cells can be acquired through laser capture microdissection (LCM), which has also been used to isolate transcripts in single cells (Tietjen *et al.*, 2003, 2005). For dense tissue, LCM may introduce higher levels of contamination compared to flow cytometry-based methods (Okaty *et al.*, 2011). However, the advantages of *in situ* profiling are often more important than the contamination, which can be controlled.

Near-term challenges

There remains a need for further development to boost single-cell RNA detection. Lowering RNA losses and enhancing the efficiencies of reverse transcription before amplification are important directions to pursue for improving RNA detection in single cells (Islam *et al.*, 2012; Picelli *et al.*, 2013). Simultaneous detection of polyadenylated RNA and regulatory RNAs is also an important area for future development.

There is also room to improve the process of dissociation of individual cells (Shapiro *et al.*, 2013). Individual cells can be isolated using fluorescence-activated cell sorting (FACS). FACS-based sorting is a robust method that has benefited from decades of refinement. However, standard FACS needs a large pool of cells as starting material, which affects its application to low-abundance samples, such as circulating tumor cells purified from blood. Microfluidics can help with these types of applications (Karabacak *et al.*, 2014; Yu *et al.*, 2013). However, for solid tissues, neither method is satisfactory, because knowledge of the spatial location is lost. LCM can be used to cut cells from fixed tissues or cryosections. Using LCM-based methods, the spatial information of a sampled cell is retained. However, it is a labor-intensive, low-throughput tool. In addition, it is difficult to capture an entire cell without collecting collateral material from neighboring cells. Thus, the improvement of cell picking methods will be valuable for single-cell analysis.

Single-Cell Signaling Events

With the ongoing development of single-cell approaches, quantitative analysis of single-cell transcriptomics is becoming possible. Another layer of biological complexity lies in how cells with different gene expression states respond to their environment differently. Understanding the relationship between gene expression and cell signaling in single cells remains an unfulfilled goal for single-cell studies. Studying the heterogeneity in cell signaling at the single-cell level requires the tracking spatiotemporal dynamics of signaling proteins through their localization, activity, and interactions with other proteins.

Measuring protein translocation

The intracellular localization and translocation of proteins are critically tied to protein function and activation state (Andjelkovic *et al.*, 1997; Kluck *et al.*, 1997). An example is the interaction of Akt and 3-phosphoinositide dependent protein kinase-1 (PDK1). Both Akt and PDK1 bear pleckstrin homology (PH) domains that recognize the product of PI3K,

phosphatidylinositol 3,4,5-trisphosphate (PIP3) (Alessi *et al.*, 1997; Franke *et al.*, 1997; Weiger *et al.*, 2009). Akt and PDK1 accumulate at the membrane, PDK1 activates Akt by phosphorylation, and then Akt transfers the signal downstream. Thus, the translocation of signaling proteins or their recognition domains to subcellular compartments has widely served as the proxy of pathway activity dynamics (Bandara *et al.*, 2009; Galic *et al.*, 2012; Wang *et al.*, 2014; Weiger *et al.*, 2009; Yang *et al.*, 2012).

Most of the membrane-translocation reporters have been created as fusion proteins with a fluorescent probe and a lipid-binding domain. PH domains are the most abundant lipid-binding domains in eukaryotes and bind specifically to PIP3, PI(4,5)P2, or PI(3,4)P2 (Lemmon, 2007). These domains possess different phosphoinositide specificities and affinities and show strikingly different membrane recruitment kinetics *in vitro* and *in vivo* (Manna *et al.*, 2007). For example, Btk-PH shows genuine PIP3 specificity *in vitro* (Manna *et al.*, 2007), but does not catch the rapid accumulation of local PIP3 in single cells (Varnai *et al.*, 1999). In mammalian cells, the Akt-PH domain has served as the best reporter for local PIP3 production (Servant *et al.*, 2000). Although the Akt-PH domain may also bind to PI(3,4)P2, it clearly polarizes during chemoattractant receptor signaling. The localization of many binding domains is more dependent on protein-protein interactions at the plasma membrane than lipid-binding per se (Manna *et al.*, 2007). Thus, localization kinetics is as important as the lipid-binding specificities when designing membrane-translocation reporters.

Nuclear translocation of many transcription factors is a key step for cellular signaling to elicit changes in gene expression. The dynamic movement of fluorescently labeled signaling proteins can be mentored at the single-cell level by time-lapse microscopy. Live-cell imaging approaches have been used to reveal the dynamic regulation of transcription factors such as NF- κ B (Nelson *et al.*, 2004), p53 (Lahav *et al.*, 2004), STATs (McBride and Reich, 2003), and Smads (Wang *et al.*, 2014; Warmflash *et al.*, 2012), and these reporters all showed heterogeneous translocation dynamics in single cells.

Retention or release of proteins from the mitochondria has been used as a terminal indicator of apoptosis in single cells (Albeck *et al.*, 2008; Spencer *et al.*, 2009). The intermembrane space (IMS) reporter protein (IMS-RP) was created by fusing a fluorescent protein to the mitochondrial import sequence of the mitochondrial protein Smac (Albeck *et al.*, 2008; Du *et al.*, 2000). During apoptosis-associated mitochondrial outer membrane permeabilization (MOMP), IMS-RP is released from the IMS into the cytosol. IMS-RP does not contain the inhibitor of apoptosis protein (IAP) binding motif and is biochemically neutral. IMS-RP has been used to analyze the drug-induced apoptosis in human breast and pancreatic cancer cells (Earley *et al.*, 2012). The use of IMS-RP will allow visualization of heterogeneous drug responses and could be useful for examining the origins of drug resistance.

Measuring protein interactions and activity

Direct protein interaction is often evaluated by using Förster resonance energy transfer (FRET). FRET utilizes two fluorophores, a donor and an acceptor, of which the donor emission spectrum overlaps with the acceptor absorption spectrum.

If the two fluorophores are in close proximity (1–10 nm), excitation of donor results in acceptor emission due to the overlap in spectra that allows resonance energy transfer between two probes. FRET is now widely applied in cell signaling studies to detect protein–protein interactions. Advances in using three fluorescent proteins enable the study of higher-order protein complexes. In trimeric complexes, cyan fluorescent protein (CFP) is the donor for yellow fluorescent protein (YFP); subsequently, YFP can act as the donor of red fluorescent protein (RFP). 3-FRET has been applied to the analysis of three-protein interactions (Galperin *et al.*, 2004) and trimerization of TNF receptor-associated factor 2 (TRAF2) in living cells (He *et al.*, 2005).

FRET-based methods have been expanded to monitor the dynamics of kinase activity by combining a peptide substrate and a phosphopeptide recognition module. For example, FRET sensors of ERK activity have been applied in living cells (Harvey *et al.*, 2008; Tomida *et al.*, 2012). ERK activity sensors include a substrate phosphorylation peptide, an ERK docking domain, and a phospho-binding WW domain. ERK activation leads to phosphorylation of the substrates and subsequent binding by WW domain. The resulting conformational rearrangement triggers a change in FRET between the donor and acceptor. The extent and duration of ERK activity were determined by FRET in response to complex stimuli (Tomida *et al.*, 2012). The reporter of kinase activity has been used in conjunction with other reporters to link complex ERK dynamics to S phase entry (Albeck *et al.*, 2013). However, there are several limitations of FRET-based methods for their widespread adoption. For example, FRET has relatively low signal-to-noise ratio. FRET requires multiple fluorescent proteins and thus limits the number of the targets that can be observed simultaneously. Besides, the closed conformation of the FRET sensor is highly stable and cannot be reset by the phosphatases (Komatsu *et al.*, 2011). These restrict FRET from accurately representing the dynamic regulation of kinase activities in the same live cells.

A new strategy to design single-cell reporters for kinase activity is based on the concept of converting phosphorylation into nucleo-cytoplasmic translocation (Regot *et al.*, 2014; Spencer *et al.*, 2013). A recently reported mammalian cyclin-dependent kinase 2 (CDK2) reporter includes YFP mVenus fused to amino acids 994–1087 of human DNA helicase B (DHB) that contains four CDK consensus phosphorylation sites, a nuclear localization signal (NLS), and a nuclear export signal (NES) (Spencer *et al.*, 2013). CDK2 reporter localizes to the nucleus in newly born cells. DHB is specifically phosphorylated by CDK2 whose activity is upregulated in G1 before DNA replication (Massague, 2004). CDK2 reporter then translocates out of the nucleus and exhibits cytosolic localization at the end of G2. The cytoplasmic-to-nuclear ratio of the reporter is used as the readout of CDK2 activity. This strategy has recently been generalized to kinase translocation reporters (KTRs) of enzymatic activity (Regot *et al.*, 2014). A KTR includes a docking site and phosphorylation site of the target kinase, a bipartite NLS, an NES, and a single fluorescent protein. KTR can be applied to multiple types of kinases simply with minimal optimization and has been successful for three classes of kinases (MAPKs, CDKs, and AGC kinases) (Regot *et al.*, 2014; Spencer *et al.*, 2013). These new techniques

open the possibility of analyzing a wide range of kinase-mediated processes in individual cells.

Despite these advances, live-cell imaging techniques are still limited by several drawbacks. Fluorescence microscopy has the limited number of color channels that can be viewed simultaneously without spectral overlap. It remains an experimental challenge to use three to four fluorescent activity reporters simultaneously in live cells (Machacek *et al.*, 2009; Regot *et al.*, 2014). Overexpression of certain reporters may create artifacts related to competition with endogenous substrates (Ahmed *et al.*, 2014). Moreover, plasmid copy number can vary per cell, causing heterogeneity that does not reflect true intracellular events. Knockin of fluorescent reporters into endogenous loci by using CRISPR/Cas (Yang *et al.*, 2013) or gene tagging (Sigal *et al.*, 2006) should alleviate some of these concerns, although sensitivity of detection could then become an issue. In addition, primary cells remain a major challenge for all of these techniques (Saucerman *et al.*, 2006; Yang *et al.*, 2014). End-point fixed-cell methods with flow cytometry or microscopy may provide the flexibility to measure multiple phospho-signals and regulatory proteins together when time-evolution is not critical. The fusion of different techniques holds the greatest potential to decipher the spatial and temporal interplay of multiple signaling activities within single cells (Albeck *et al.*, 2008, 2013).

Modeling and Analysis of Stochastic Data

Modeling and Analysis of Single-Cell Signaling Networks

For population-averaged data, ordinary differential equation (ODE)-based models are powerful tools to integrate mechanistic information and describe signaling kinetics (Aldridge *et al.*, 2006). ODE models have also been used with single-cell data from live-cell microscopy and flow cytometry to investigate cell-to-cell variability in apoptotic signals (Spencer *et al.*, 2009). However, deterministic ODE models do not always represent signaling at a single-cell level, where stochasticity may affect the probability of signaling and molecular states (Altschuler and Wu, 2010; Cheong *et al.*, 2010).

Using Bayes probability theorem, Bayesian networks are suitable for learning network connections from a distribution of single-cell events. The ability to express probabilistic relationships among network variables makes Bayesian approaches valuable for integrating single-cell data (Kolitz and Lauffenburger, 2012). Bayesian networks have inferred causal interactions in a kinase-signaling network by using the phospho-protein levels in single cells (Sachs *et al.*, 2005). These approaches enable the inference of networks in which not all molecules are measured (Sachs *et al.*, 2009). A main drawback of Bayesian networks is that they need a lot of data, usually hundreds to thousands of observations. In addition, the complexity of possible networks that must be searched grows super-exponentially with the number of nodes in the networks, requiring high-performance computing for all but the smallest networks.

A separate challenge for single-cell signaling is the reliability of transmitting signals. Here, information theory has been employed to study the ability of single cells to transfer

environmental information in the presence of noise within a signaling network. For example, Cheong *et al.* used mutual information between intracellular relay points to evaluate the fidelity of information flow along the NF- κ B pathway in single cells during TNF α stimulation (Cheong *et al.*, 2011). It was suggested that individual pathways could only discriminate between the presence and absence of TNF α . This relatively low transmission capability was attributed to an information bottleneck around the receptors. The authors also suggested that one role of negative feedback in signaling networks is to improve the information-carrying capacity. Modeling the information flow in networks may need to include measures of the transfer function between nodes (Toettcher *et al.*, 2013).

The single-cell techniques discussed earlier are limited by the number of markers that can be observed in parallel, but methods are being rapidly developed to address this problem. By combining mass spectroscopy and flow cytometry, mass cytometry allows quantitative monitoring of a large number of parameters simultaneously in single cells (Bendall *et al.*, 2011). The increasing multidimensional types of data require specific visualization and analysis methods. An algorithm, termed Spanning-tree progression analysis of density-normalized events (SPADE), was created to cluster cell subsets in multidimensional space and then reduce them to a 2-D representation tree structure (Qiu *et al.*, 2011). In addition, the recently developed viSNE algorithm has allowed high-dimensional data to be compressed into a 2-D map that discriminates cellular subpopulations without the loss of single-cell resolution (Amir *et al.*, 2013). These approaches are unsupervised and prior knowledge of the system is not required. They are thus suitable for navigating complicated, interconnected signaling networks in search of hypotheses.

Modeling Single-Cell Transcriptomic Data

Nongenetic variability in gene expression may result from both population heterogeneity and temporal fluctuations (Huang, 2009). Temporal fluctuations are assumed to be regulated by the stochastic synthesis and degradation of individual transcripts, while population heterogeneity may be due to the stable differences among cells. It is difficult to distinguish one from the other in experimental data. However, if temporal fluctuations in gene expression are fast (Yu *et al.*, 2006), they may satisfy the assumption of ergodic theory, which states that the temporal fluctuation of an individual cell can be estimated by measuring the population at one time point (Eckmann and Ruelle, 1985). Thus, a single snapshot of cells in the same population can be used to infer the temporal fluctuations in individual cells.

Experimental observations of single-cell gene expression are usually noisy. Owing to low amounts of input material, true biological variability and technical noise are often at comparable levels (Brennecke *et al.*, 2013). The relative amount or the presence/absence of individual transcripts may be dominated by technical error. In addition, the transcript numbers of 'housekeeping' genes, such as *GAPDH* and *POLR2A*, show considerable variability at the single-cell level (McDavid *et al.*, 2013). Standard methods of data normalization based on such genes will result in quantitative artifacts.

Addressing these statistical and analytical challenges in single-cell data is critical for making appropriate interpretations.

A number of approaches for modeling of single-cell transcriptomic data have been reported including the use of external RNA controls as a reference material (Brennecke *et al.*, 2013; Grun *et al.*, 2014; Kharchenko *et al.*, 2014), Bayesian statistics of differential gene expression (Kharchenko *et al.*, 2014), or unique molecular identifier barcodes for individual transcripts (Grun *et al.*, 2014). It has been estimated by these models that approximately half of the reads may be contributed by technical noise (Brennecke *et al.*, 2013). Thus, challenges remain for analyzing stochastic single-cell data in a reliable way.

An alternative approach with less-stringent sample requirement is to analyze multiple sets of small samples, rather than single-cell samples. Probability theory and random sampling have been used together to survey the gene expression programs that are significantly heterogeneous among individual cells (Janes *et al.*, 2010; Wang and Janes, 2013). Using 15–20 random samples of 10-cell averages allows the resulting distributions to distinguish cell-to-cell variations by statistical hypothesis testing. Moreover, these stochastic-profiling experiments are quantitative because the 10-fold increase in input material allows for more-accurate measurements compared to single-cell data. Recently, it has been shown that mathematically deconvolving measurements from small groups of cells can quantify single-cell regulatory heterogeneities while avoiding measurement noise (Bajikar *et al.*, 2014). It remains unclear whether noisy one-cell data or robust 10-cell data will ultimately be more useful for uncovering new single-cell biology.

Connecting Heterogeneity at Multiple Levels

A primary goal of studying cell-to-cell heterogeneity is the potential to connect underlying variability with heterogeneous outcomes. Ultimately, this will require combining different methods that quantify gene expression products, signaling activities, and functional responses to link upstream cellular variability to a downstream cell-fate decision (Wang *et al.*, 2012).

Variability in gene expression can give rise to phenotypic heterogeneity. Gene expression begins with transcription, which implies that the fluctuation of mRNA synthesis may propagate to the downstream protein pool. However, in metazoan development, such fluctuation is suppressed and cell fate is highly controlled by the network architecture. In *Caenorhabditis elegans*, mutations to components of a developmental network lead to large variations in the expression of downstream genes (Raj *et al.*, 2010). These variations were thresholded by the organism to give rise to a bimodal expression pattern of key regulatory genes. This suggests the intrinsic mechanisms exist to buffer the stochastic fluctuations in core networks.

Variability in signaling has also been shown to be able to generate subpopulations with distinct cell fates. Many growth factors can induce multiple pathways such as PI3K and Ras pathways. Uniform nerve growth factor (NGF) or EGF stimuli have been found to cause cell-to-cell variability in pAkt and

pERK activation in PC12 cells (Chen *et al.*, 2012). The combined levels of pERK and pAkt define a cell's propensity to undergo proliferation versus differentiation. Thus, both variable signaling and gene expression can be mapped to heterogeneous phenotypes.

Gene expression mediated by viral transduction is also subject to cell-to-cell variability. These variations arise from the stochastic events intrinsic to gene expression, which can be exploited for the analysis of signaling dynamics. This approach has been used to map the biphasic E2F1 response to variability in MYC (Wong *et al.*, 2011). The paradox that MYC can mediate growth signals and trigger growth arrest may be reconciled by taking into account the differing amounts of MYC present in cells.

Conclusion

In this article, we have discussed various levels of cell-to-cell heterogeneity, which create variability in biomolecules that contribute to heterogeneous phenotypes. Cell-to-cell heterogeneity may provide an inherent mechanism to diversify the population-level response to environmental cues. Cells might exploit this mechanism to adapt to external stimuli. Our understanding of heterogeneity will continue to improve as new tools are developed to interrogate the cell and molecular biology that underlies it.

See also: Cell Communication: Cellular Machineries: Macromolecular Communication between Nucleus and Cytoplasm. Cell Communication: Growth Factor Mediated Cell Signaling: Receptor Tyrosine Kinases and Their Ligands; TGF- β Superfamily Signaling; Tumor Necrosis Factor Receptors: A Brief Digestion. Cell Communication: Intracellular Pathways: The MAPK Signaling Cascades; The PI3K/Akt/mTOR Pathway. Cell Division/Death: Apoptosis: Mitochondria in Cell Death Regulation; Tumor Necrosis Factor Signaling Pathways. Dynamic Integration: Dynamics: Dynamics of Protein Kinase Cascades; Transcription Factor Networks. Horizontal Integration: Networks: Understanding of 'Networks' *In Vitro* and/or *In Vivo*. Horizontal Integration: Omics: Transcriptomics. Nucleic Acid Synthesis/Breakdown: Transcription: Eukaryotic Transcriptional Regulation

References

- Ahmed, S., Grant, K.G., Edwards, L.E., *et al.*, 2014. Data-driven modeling reconciles kinetics of ERK phosphorylation, localization, and activity states. *Molecular Systems Biology* 10, 718.
- Albeck, J.G., Burke, J.M., Aldridge, B.B., *et al.*, 2008. Quantitative analysis of pathways controlling extrinsic apoptosis in single cells. *Molecular Cell* 30, 11–25.
- Albeck, J.G., Mills, G.B., Brugge, J.S., 2013. Frequency-modulated pulses of ERK activity transmit quantitative proliferation signals. *Molecular Cell* 49, 249–261.
- Aldridge, B.B., Burke, J.M., Lauffenburger, D.A., Sorger, P.K., 2006. Physicochemical modelling of cell signalling pathways. *Nature Cell Biology* 8, 1195–1203.
- Alessi, D.R., Deak, M., Casamayor, A., *et al.*, 1997. 3-Phosphoinositide-dependent protein kinase-1 (PDK1): structural and functional homology with the *Drosophila* DSTPK61 kinase. *Current Biology* 7, 776–789.
- Altschuler, S.J., Wu, L.F., 2010. Cellular heterogeneity: Do differences make a difference? *Cell* 141, 559–563.
- Amir el, A.D., Davis, K.L., Tadmor, M.D., *et al.*, 2013. viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. *Nature Biotechnology* 31, 545–552.
- Andjelkovic, M., Alessi, D.R., Meier, R., *et al.*, 1997. Role of translocation in the activation and function of protein kinase B. *Journal of Biological Chemistry* 272, 31515–31524.
- Ashall, L., Horton, C.A., Nelson, D.E., *et al.*, 2009. Pulsatile stimulation determines timing and specificity of NF-kappaB-dependent transcription. *Science* 324, 242–246.
- Bajikar, S.S., Fuchs, C., Roller, A., Theis, F.J., Janes, K.A., 2014. Parameterizing cell-to-cell regulatory heterogeneities via stochastic transcriptional profiles. *Proceedings of the National Academy of Sciences of the United States of America* 111, E626–E635.
- Bandara, S., Schloder, J.P., Eils, R., Bock, H.G., Meyer, T., 2009. Optimal experimental design for parameter estimation of a cell signaling model. *PLoS Computational Biology* 5, e1000558.
- Bar-Even, A., Paulsson, J., Maheshri, N., *et al.*, 2006. Noise in protein expression scales with natural protein abundance. *Nature Genetics* 38, 636–643.
- Bendall, S.C., Simonds, E.F., Qiu, P., *et al.*, 2011. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science* 332, 687–696.
- Brennecke, P., Anders, S., Kim, J.K., *et al.*, 2013. Accounting for technical noise in single-cell RNA-seq experiments. *Nature Methods* 10, 1093–1095.
- Buganim, Y., Faddah, D.A., Cheng, A.W., *et al.*, 2012. Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchical phase. *Cell* 150, 1209–1222.
- Cai, L., Dalal, C.K., Elowitz, M.B., 2008. Frequency-modulated nuclear localization bursts coordinate gene regulation. *Nature* 455, 485–490.
- Chaffer, C.L., Brueckmann, I., Scheel, C., *et al.*, 2011. Normal and neoplastic nonstem cells can spontaneously convert to a stem-like state. *Proceedings of the National Academy of Sciences of the United States of America* 108, 7950–7955.
- Chaffer, C.L., Marjanovic, N.D., Lee, T., *et al.*, 2013. Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *Cell* 154, 61–74.
- Chen, J.Y., Lin, J.R., Cimprich, K.A., Meyer, T., 2012. A two-dimensional ERK-AKT signaling code for an NGF-triggered cell-fate decision. *Molecular Cell* 45, 196–209.
- Cheong, R., Bergmann, A., Werner, S.L., *et al.*, 2006. Transient IkkappaB kinase activity mediates temporal NF-kappaB dynamics in response to a wide range of tumor necrosis factor-alpha doses. *Journal of Biological Chemistry* 281, 2945–2950.
- Cheong, R., Paliwal, S., Levchenko, A., 2010. Models at the single cell level. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 2, 34–48.
- Cheong, R., Rhee, A., Wang, C.J., Nemenman, I., Levchenko, A., 2011. Information transduction capacity of noisy biochemical signaling networks. *Science* 334, 354–358.
- Cirit, M., Wang, C.C., Haugh, J.M., 2010. Systematic quantification of negative feedback mechanisms in the extracellular signal-regulated kinase (ERK) signaling network. *Journal of Biological Chemistry* 285, 36736–36744.
- Cohen-Saidon, C., Cohen, A.A., Sigal, A., Liron, Y., Alon, U., 2009. Dynamics and variability of ERK2 response to EGF in individual living cells. *Molecular Cell* 36, 885–893.
- Dalerba, P., Kalisky, T., Sahoo, D., *et al.*, 2011. Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nature Biotechnology* 29, 1120–1127.
- Deng, Q., Ramskold, D., Reinius, B., Sandberg, R., 2014. Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* 343, 193–196.
- Dolmetsch, R.E., Lewis, R.S., Goodnow, C.C., Healy, J.I., 1997. Differential activation of transcription factors induced by Ca^{2+} response amplitude and duration. *Nature* 386, 855–858.
- Du, C., Fang, M., Li, Y., Li, L., Wang, X., 2000. Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition. *Cell* 102, 33–42.
- Earley, S., Vinegoni, C., Dunham, J., *et al.*, 2012. *In vivo* imaging of drug-induced mitochondrial outer membrane permeabilization at single-cell resolution. *Cancer Research* 72, 2949–2956.
- Eberwine, J., Sul, J.Y., Bartfai, T., Kim, J., 2014. The promise of single-cell sequencing. *Nature Methods* 11, 25–27.
- Eberwine, J., Yeh, H., Miyashiro, K., *et al.*, 1992. Analysis of gene expression in single live neurons. *Proceedings of the National Academy of Sciences of the United States of America* 89, 3010–3014.

- Eckmann, J.P., Ruelle, D., 1985. Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics* 57, 617–656.
- Elowitz, M.B., Levine, A.J., Siggia, E.D., Swain, P.S., 2002. Stochastic gene expression in a single cell. *Science* 297, 1183–1186.
- Fallah-Sichani, M., Honarnejad, S., Heiser, L.M., Gray, J.W., Sorger, P.K., 2013. Metrics other than potency reveal systematic variation in responses to cancer drugs. *Nature Chemical Biology* 9, 708–714.
- Flusberg, D.A., Roux, J., Spencer, S.L., Sorger, P.K., 2013. Cells surviving fractional killing by TRAIL exhibit transient but sustainable resistance and inflammatory phenotypes. *Molecular Biology of the Cell* 24, 2186–2200.
- Frank, T.F., Kaplan, D.R., Cantley, L.C., Toker, A., 1997. Direct regulation of the Akt proto-oncogene product by phosphatidylinositol-3,4-bisphosphate. *Science* 275, 665–668.
- Galic, M., Jeong, S., Tsai, F.C., *et al.*, 2012. External push and internal pull forces recruit curvature-sensing N-BAR domain proteins to the plasma membrane. *Nature Cell Biology* 14, 874–881.
- Galperin, E., Verkhusha, V.V., Sorkin, A., 2004. Three-chromophore FRET microscopy to analyze multiprotein interactions in living cells. *Nature Methods* 1, 209–217.
- Grun, D., Kester, L., van Oudenaarden, A., 2014. Validation of noise models for single-cell transcriptomics. *Nature Methods* 11, 637–640.
- Gupta, P.B., Fillmore, C.M., Jiang, G., *et al.*, 2011. Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* 146, 633–644.
- Hanna, J., Saha, K., Pando, B., *et al.*, 2009. Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* 462, 595–601.
- Harvey, C.D., Ehrhardt, A.G., Cellurale, C., *et al.*, 2008. A genetically encoded fluorescent sensor of ERK activity. *Proceedings of the National Academy of Sciences of the United States of America* 105, 19264–19269.
- Hashimshony, T., Wagner, F., Sher, N., Yanai, I., 2012. CEL-Seq: Single-cell RNA-Seq by multiplexed linear amplification. *Cell Reports* 2, 666–673.
- He, L., Wu, X., Simone, J., Heggill, D., Lipsky, P.E., 2005. Determination of tumor necrosis factor receptor-associated factor trimerization in living cells by CFP->YFP->mRFP FRET detected by flow cytometry. *Nucleic Acids Research* 33, e61.
- Hoffmann, A., Baltimore, D., 2006. Circuitry of nuclear factor kappaB signaling. *Immunological Reviews* 210, 171–186.
- Hoffmann, A., Levchenko, A., Scott, M.L., Baltimore, D., 2002. The IkappaB-NF-kappaB signaling module: Temporal control and selective gene activation. *Science* 298, 1241–1245.
- Huang, S., 2009. Non-genetic heterogeneity of cells in development: More than just noise. *Development* 136, 3853–3862.
- Islam, S., Kjallquist, U., Moliner, A., *et al.*, 2012. Highly multiplexed and strand-specific single-cell RNA 5' end sequencing. *Nature Protocols* 7, 813–828.
- Jaitin, D.A., Kenigsberg, E., Keren-Shaul, H., *et al.*, 2014. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. *Science* 343, 776–779.
- Janes, K.A., Reinhardt, H.C., Yaffe, M.B., 2008. Cytokine-induced signaling networks prioritize dynamic range over signal strength. *Cell* 135, 343–354.
- Janes, K.A., Wang, C.C., Holmberg, K.J., Cabral, K., Brugge, J.S., 2010. Identifying single-cell molecular programs by stochastic profiling. *Nature Methods* 7, 311–317.
- Karabacak, N.M., Spuhler, P.S., Fachin, F., *et al.*, 2014. Microfluidic, marker-free isolation of circulating tumor cells from blood samples. *Nature Protocols* 9, 694–710.
- Kharchenko, P.V., Silberstein, L., Scadden, D.T., 2014. Bayesian approach to single-cell differential expression analysis. *Nature Methods* 11, 740–742.
- Kim, H.D., Shay, T., O'Shea, E.K., Regev, A., 2009. Transcriptional regulatory circuits: Predicting numbers from alphabets. *Science* 325, 429–432.
- Kluck, R.M., Bossy-Wetzel, E., Green, D.R., Newmeyer, D.D., 1997. The release of cytochrome c from mitochondria: A primary site for Bcl-2 regulation of apoptosis. *Science* 275, 1132–1136.
- Kolitz, S.E., Lauffenburger, D.A., 2012. Measurement and modeling of signaling at the single-cell level. *Biochemistry* 51, 7433–7443.
- Komatsu, N., Aoki, K., Yamada, M., *et al.*, 2011. Development of an optimized backbone of FRET biosensors for kinases and GTPases. *Molecular Biology of the Cell* 22, 4647–4656.
- Ladbury, J.E., Arold, S.T., 2012. Noise in cellular signaling pathways: Causes and effects. *Trends in Biochemical Sciences* 37, 173–178.
- Lahav, G., Rosenfeld, N., Sigal, A., *et al.*, 2004. Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nature Genetics* 36, 147–150.
- Lee, R.E., Walker, S.R., Savery, K., Frank, D.A., Gaudet, S., 2014. Fold change of nuclear NF-kappaB determines TNF-induced transcription in single cells. *Molecular Cell* 53, 867–879.
- Lemmon, M.A., 2007. Pleckstrin homology (PH) domains and phosphoinositides. *Biochemical Society Symposia* 74, 81–93.
- Losick, R., Desplan, C., 2008. Stochasticity and cell fate. *Science* 320, 65–68.
- Macaulay, I.C., Voet, T., 2014. Single cell genomics: Advances and future perspectives. *PLoS Genetics* 10, e1004126.
- Machacek, M., Hodgson, L., Welch, C., *et al.*, 2009. Coordination of Rho GTPase activities during cell protrusion. *Nature* 461, 99–103.
- Magee, J.A., Piskounova, E., Morrison, S.J., 2012. Cancer stem cells: Impact, heterogeneity, and uncertainty. *Cancer Cell* 21, 283–296.
- Mani, S.A., Guo, W., Liao, M.J., *et al.*, 2008. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* 133, 704–715.
- Manna, D., Albanese, A., Park, W.S., Cho, W., 2007. Mechanistic basis of differential cellular responses of phosphatidylinositol 3,4-bisphosphate- and phosphatidylinositol 3,4,5-trisphosphate-binding pleckstrin homology domains. *Journal of Biological Chemistry* 282, 32093–32105.
- Marshall, C.J., 1995. Specificity of receptor tyrosine kinase signaling: Transient versus sustained extracellular signal-regulated kinase activation. *Cell* 80, 179–185.
- Marusyk, A., Almendro, V., Polyak, K., 2012. Intra-tumour heterogeneity: A looking glass for cancer? *Nature Reviews Cancer* 12, 323–334.
- Massague, J., 2004. G1 cell-cycle control and cancer. *Nature* 432, 298–306.
- McBride, K.M., Reich, N.C., 2003. The ins and outs of STAT1 nuclear transport. *Science's STKE* 2003, RE13.
- McDavid, A., Finak, G., Chattopadhyay, P.K., *et al.*, 2013. Data exploration, quality control and testing in single-cell qPCR-based gene expression experiments. *Bioinformatics* 29, 461–467.
- McMahon, A., Supatto, W., Fraser, S.E., Stathopoulos, A., 2008. Dynamic analyses of Drosophila gastrulation provide insights into collective cell migration. *Science* 322, 1546–1550.
- Meyer, R., D'Alessandro, L.A., Kar, S., *et al.*, 2012. Heterogeneous kinetics of AKT signaling in individual cells are accounted for by variable protein concentration. *Frontiers in Physiology* 3, 451.
- Munsky, B., Neuert, G., van Oudenaarden, A., 2012. Using gene expression noise to understand gene regulation. *Science* 336, 183–187.
- Nelson, D.E., Ihekweaba, A.E., Elliott, M., *et al.*, 2004. Oscillations in NF-kappaB signaling control the dynamics of gene expression. *Science* 306, 704–708.
- Newman, J.R., Ghahmaghani, S., Ihmels, J., *et al.*, 2006. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* 441, 840–846.
- Okaty, B.W., Sugino, K., Nelson, S.B., 2011. A quantitative comparison of cell-type-specific microarray gene expression profiling methods in the mouse brain. *PLoS ONE* 6, e16493.
- Pawson, T., Nash, P., 2003. Assembly of cell regulatory systems through protein interaction domains. *Science* 300, 445–452.
- Picelli, S., Bjorklund, A.K., Faridani, O.R., *et al.*, 2013. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nature Methods* 10, 1096–1098.
- Picelli, S., Faridani, O.R., Bjorklund, A.K., *et al.*, 2014. Full-length RNA-seq from single cells using Smart-seq2. *Nature Protocols* 9, 171–181.
- Purvis, J.E., Lahav, G., 2013. Encoding and decoding cellular information through signaling dynamics. *Cell* 152, 945–956.
- Qiu, P., Simonds, E.F., Bendall, S.C., *et al.*, 2011. Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nature Biotechnology* 29, 886–891.
- Quintana, E., Shackleton, M., Sabel, M.S., *et al.*, 2008. Efficient tumour formation by single human melanoma cells. *Nature* 456, 593–598.
- Raj, A., van Oudenaarden, A., 2009. Single-molecule approaches to stochastic gene expression. *Annual Review of Biophysics* 38, 255–270.
- Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., Tyagi, S., 2006. Stochastic mRNA synthesis in mammalian cells. *PLoS Biology* 4, e309.
- Raj, A., Rifkin, S.A., Andersen, E., van Oudenaarden, A., 2010. Variability in gene expression underlies incomplete penetrance. *Nature* 463, 913–918.
- Regot, S., Hughey, J.J., Bajar, B.T., Carrasco, S., Covert, M.W., 2014. High-sensitivity measurements of multiple kinase activities in live single cells. *Cell* 157, 1724–1734.
- Sachs, K., Itani, S., Carlisle, J., *et al.*, 2009. Learning signaling network structures with sparsely distributed data. *Journal of Computational Biology* 16, 201–212.
- Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D.A., Nolan, G.P., 2005. Causal protein-signaling networks derived from multiparameter single-cell data. *Science* 308, 523–529.
- Sandberg, R., 2014. Entering the era of single-cell transcriptomics in biology and medicine. *Nature Methods* 11, 22–24.
- Saucerman, J.J., Zhang, J., Martin, J.C., *et al.*, 2006. Systems analysis of PKA-mediated phosphorylation gradients in live cardiac myocytes. *Proceedings of the National Academy of Sciences of the United States of America* 103, 12923–12928.
- Servant, G., Weiner, O.D., Herzmark, P., *et al.*, 2000. Polarization of chemoattractant receptor signaling during neutrophil chemotaxis. *Science* 287, 1037–1040.

- Shalek, A.K., Satija, R., Adiconis, X., *et al.*, 2013. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* 498, 236–240.
- Shankaran, H., Ippolito, D.L., Chrisler, W.B., *et al.*, 2009. Rapid and sustained nuclear–cytoplasmic ERK oscillations induced by epidermal growth factor. *Molecular Systems Biology* 5, 332.
- Shapiro, E., Biezuner, T., Linnarsson, S., 2013. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nature Reviews Genetics* 14, 618–630.
- Sharma, S.V., Lee, D.Y., Li, B., *et al.*, 2010. A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* 141, 69–80.
- Shipitsin, M., Campbell, L.L., Argani, P., *et al.*, 2007. Molecular definition of breast tumor heterogeneity. *Cancer Cell* 11, 259–273.
- Sigal, A., Milo, R., Cohen, A., *et al.*, 2006. Variability and memory of protein levels in human cells. *Nature* 444, 643–646.
- Singh, A., Razoooky, B.S., Dar, R.D., Weinberger, L.S., 2012. Dynamics of protein noise can distinguish between alternate sources of gene-expression variability. *Molecular Systems Biology* 8, 607.
- Spencer, S.L., Cappell, S.D., Tsai, F.C., *et al.*, 2013. The proliferation-quiescence decision is controlled by a bifurcation in CDK2 activity at mitotic exit. *Cell* 155, 369–383.
- Spencer, S.L., Gaudet, S., Albeck, J.G., Burke, J.M., Sorger, P.K., 2009. Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* 459, 428–432.
- Suter, D.M., Molina, N., Gaffield, D., *et al.*, 2011. Mammalian genes are transcribed with widely different bursting kinetics. *Science* 332, 472–474.
- Takahashi, K., Yamanaka, S., 2006. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663–676.
- Tay, S., Hughey, J.J., Lee, T.K., *et al.*, 2010. Single-cell NF- κ B dynamics reveal digital activation and analogue information processing. *Nature* 466, 267–271.
- Tietjen, I., Rihel, J., Dulac, C.G., 2005. Single-cell transcriptional profiles and spatial patterning of the mammalian olfactory epithelium. *International Journal of Developmental Biology* 49, 201–207.
- Tietjen, I., Rihel, J.M., Cao, Y., *et al.*, 2003. Single-cell transcriptional analysis of neuronal progenitors. *Neuron* 38, 161–175.
- Toettcher, J.E., Weiner, O.D., Lim, W.A., 2013. Using optogenetics to interrogate the dynamic control of signal transmission by the Ras/Erk module. *Cell* 155, 1422–1434.
- Tomida, T., Oda, S., Takekawa, M., Iino, Y., Saito, H., 2012. The temporal pattern of stimulation determines the extent and duration of MAPK activation in a *Caenorhabditis elegans* sensory neuron. *Science Signaling* 5, ra76.
- Varnai, P., Rother, K.I., Balla, T., 1999. Phosphatidylinositol 3-kinase-dependent membrane association of the Bruton's tyrosine kinase pleckstrin homology domain visualized in single living cells. *Journal of Biological Chemistry* 274, 10983–10989.
- Venkatesan, K., Rual, J.F., Vazquez, A., *et al.*, 2009. An empirical framework for binary interactome mapping. *Nature Methods* 6, 83–90.
- Visvader, J.E., Lindeman, G.J., 2008. Cancer stem cells in solid tumours: Accumulating evidence and unresolved questions. *Nature Reviews Cancer* 8, 755–768.
- Wang, C.C., Bajikar, S.S., Jamal, L., Atkins, K.A., Janes, K.A., 2014. A time- and matrix-dependent TGFBR3-JUND-KRT5 regulatory circuit in single breast epithelial cells and basal-like premalignancies. *Nature Cell Biology* 16, 345–356.
- Wang, C.C., Cirit, M., Haugh, J.M., 2009. PI3K-dependent cross-talk interactions converge with Ras as quantifiable inputs integrated by Erk. *Molecular Systems Biology* 5, 246.
- Wang, C.C., Jamal, L., Janes, K.A., 2012. Normal morphogenesis of epithelial tissues and progression of epithelial tumors. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 4, 51–78.
- Wang, L., Janes, K.A., 2013. Stochastic profiling of transcriptional regulatory heterogeneities in tissues, tumors and cultured cells. *Nature Protocols* 8, 282–301.
- Warmflash, A., Zhang, Q., Sorre, B., *et al.*, 2012. Dynamics of TGF- β signaling reveal adaptive and pulsatile behaviors reflected in the nuclear localization of transcription factor Smad4. *Proceedings of the National Academy of Sciences of the United States of America* 109, E1947–E1956.
- Warren, L., Bryder, D., Weissman, I.L., Quake, S.R., 2006. Transcription factor profiling in individual hematopoietic progenitors by digital RT-PCR. *Proceedings of the National Academy of Sciences of the United States of America* 103, 17807–17812.
- Weiger, M.C., Wang, C.C., Krajcovic, M., *et al.*, 2009. Spontaneous phosphoinositide 3-kinase signaling dynamics drive spreading and random migration of fibroblasts. *Journal of Cell Science* 122, 313–323.
- Wong, J.V., Yao, G., Nevins, J.R., You, L., 2011. Viral-mediated noisy gene expression reveals biphasic E2f1 response to MYC. *Molecular Cell* 41, 275–285.
- Yang, H., Wang, H., Shivalila, C.S., *et al.*, 2013. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154, 1370–1379.
- Yang, H.W., Shin, M.G., Lee, S., *et al.*, 2012. Cooperative activation of PI3K by Ras and Rho family small GTPases. *Molecular Cell* 47, 281–290.
- Yang, J.H., Polanowska-Grabowska, R.K., Smith, J.S., Shields, C.W.t., Saucerman, J. J., 2014. PKA catalytic subunit compartmentation regulates contractile and hypertrophic responses to beta-adrenergic signaling. *Journal of Molecular and Cellular Cardiology* 66, 83–93.
- Yoshiura, S., Ohtsuka, T., Takenaka, Y., *et al.*, 2007. Ultradian oscillations of Stat, Smad, and Hes1 expression in response to serum. *Proceedings of the National Academy of Sciences of the United States of America* 104, 11292–11297.
- Yuan, T.L., Wulf, G., Burga, L., Cantley, L.C., 2011. Cell-to-cell variability in PI3K protein level regulates PI3K-AKT pathway activity in cell populations. *Current Biology* 21, 173–183.
- Yu, J., Xiao, J., Ren, X., Lao, K., Xie, X.S., 2006. Probing gene expression in live cells, one protein molecule at a time. *Science* 311, 1600–1603.
- Yu, M., Bardia, A., Wittner, B.S., *et al.*, 2013. Circulating breast tumor cells exhibit dynamic changes in epithelial and mesenchymal composition. *Science* 339, 580–584.
- Yu, R.C., Pesce, C.G., Colman-Lerner, A., *et al.*, 2008. Negative feedback that improves information transmission in yeast signalling. *Nature* 456, 755–761.
- Zheng, Y., Zhang, C., Croucher, D.R., *et al.*, 2013. Temporal regulation of EGF signalling networks by the scaffold protein Shc1. *Nature* 499, 166–171.

Systematic Methods to Interrogate Genetic Perturbations and Map Phosphorylation-Dependent Signaling

R Sopko, Harvard Medical School, Boston, MA, USA

N Perrimon, Harvard Medical School, Boston, MA, USA and Howard Hughes Medical Institute, Boston, MA, USA

© 2016 Elsevier Inc. All rights reserved.

Protein phosphorylation is an abundant posttranslational modification that functions as an important signaling cue. The balance between protein kinase and phosphatase activity mediates the dynamics of phosphorylation to transmit information flow, ultimately directing the activity, localization, and function of signaling effectors. Notably, disequilibrium in the phosphorylation–dephosphorylation balancing act can have severe repercussions, and is a major contributing factor to the manifestation of pathologies such as inflammatory diseases, diabetes, neurodegenerative diseases, metabolic disorders, and cancers. Assigning individual protein phosphorylation events to their cognate kinases and phosphatases is key to understanding the organization and dynamics of signaling networks. While advances in mass spectrometry have facilitated the identification of phosphorylation, the protein kinase(s) or phosphatase(s) responsible for regulating any particular phosphorylation event in most cases is unknown. Moreover, distinguishing functional from spurious phosphorylation remains a significant hurdle to mapping signaling pathways and information processing (Lienhard, 2008). The challenge now lies in trying to pinpoint critical phosphorylation events, their cognate kinase(s), the regulatory organization between kinases and phosphatases, and their relevance to organismal physiology.

In recent years, building on the knowledge of genome sequences and technological advances in molecular biology and proteomics, a number of approaches have been developed to systematically identify protein kinase substrates (Figure 1; Boyle and Koleske, 2007; Sopko and Andrews, 2008). For instance, protein and peptide chips have been used extensively to identify protein kinase substrates (Ptacek *et al.*, 2005), protein kinase consensus motifs (Alexander *et al.*, 2011; Huttii *et al.*, 2004; Mok *et al.*, 2010; Songyang *et al.*, 1994; Turk *et al.*, 2006), and those proteins with which protein kinases can associate (Fasolo *et al.*, 2011; Kaushansky *et al.*, 2008). Similarly,

pools of radiolabeled proteins generated from coupled *in vitro* transcription–translation of verified cDNA libraries have served as substrates to screen for targets of recombinant protein kinases (Gao *et al.*, 2000; Lee *et al.*, 2005). Additionally, reverse phase protein microarrays, bearing spotted lysates rather than purified proteins, have been probed with phospho-specific antibodies to identify kinase targets (Gulmann *et al.*, 2009). All these methods rely on interaction between kinase and substrate in an environment lacking cellular context, and the latter method additionally relies on antibody specificity and avidity. Genetically engineered analog-sensitive kinases capable of accepting bulky ATP analogs have enabled the labeling of direct protein kinase targets in cell lysates (Blethrow *et al.*, 2008; Carlson *et al.*, 2011; Dephore *et al.*, 2005; Holt *et al.*, 2009; Ubersax *et al.*, 2003); however, this approach relies on inconsequential mutation of the kinase ATP-binding pocket. Approaches coupling affinity purification with mass spectrometry have proven fruitful in identifying physical protein kinase–substrate interactions (Breitkreutz *et al.*, 2010; Rohila *et al.*, 2009), although this approach, like yeast two hybrid experiments, is hindered by its inability to identify dynamic kinase–substrate interactions. The varying degree of protein kinase activity and inhibitor sensitivity across cancer cell lines and tumors has been revealed with kinase activity assay for kinome profiling (KAYAK) – a multiplexed kinase assay to simultaneously measure protein kinase activity in lysates (Kubota *et al.*, 2009). This assay provides a survey of protein kinase activity and can be used to identify protein kinases responsible for phosphorylation of a given peptide designed *a priori*. Loss-of-function screens have revealed the influence of individual protein kinases on the phenotype of genetically altered cells, to predict epistasis and kinase-substrate relationships (Baldwin *et al.*, 2008; Bommi-Reddy *et al.*, 2008; Fiedler *et al.*, 2009; van Wageningen *et al.*, 2010). In the reciprocal direction, kinase substrates have been successfully

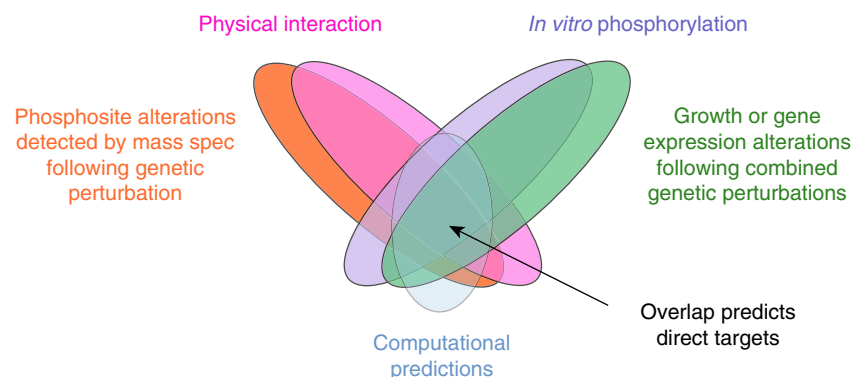


Figure 1 Methods to link phosphorylation with culpable enzymes. Direct targets can be best predicted from the overlap of *in vitro*, *in vivo*, and *in silico* approaches to distinguish kinase and phosphatase substrates.

predicted by examining the viability of a sensitized kinase mutant in the context of target gain-of-function (Sopko *et al.*, 2006). In addition to these experimental approaches, predictions of kinase substrates using computational methods are commonplace. A number of computational resources predicting kinase-specific phosphosites or phospho-binding motifs are available (Tan and Linding, 2009); however, their use is limited due to their over-predictive nature. Unquestionably, matching protein phosphorylation sites with culpable kinases is challenging, due to the limited specificity and number of kinase consensus motifs, and a number of factors including localization, expression, and association with regulatory subunits or scaffolds (Ubersax and Ferrell, 2007). As such, proving a direct link between any potential kinase-target pair still requires a considerable amount of verification. Mass spectrometry continues to develop rapidly in terms of its application to phosphoproteomics (Jünger and Aebersold, 2013). This technology is sensitive enough to capture and measure global phosphoproteome alterations in different genetic scenarios, and serves as a way to monitor altered information flow, using phosphorylation as a proxy. Notably, the increased sensitivity, speed, and dynamic range of peptide identification with current proteomics platforms is providing near comprehensive and reproducible description of simple proteomes and phosphoproteomes, making comparative analyses between genetic variants, for instance wild-type to mutants, accessible (Ahrens *et al.*, 2010; Picotti *et al.*, 2013). Additionally, the application of isobaric labeling strategies has facilitated multiplexing for relative peptide quantification in tissues and whole organisms. While less expansive scale in terms of the number of protein measurements, mass cytometry provides exquisite sensitivity in concurrently measuring dozens of proteins in individual cells (Bendall *et al.*, 2012). This platform, combining inductively coupled plasma mass spectrometry with flow cytometry, currently enables the detection of trace levels of 37, although theoretically 70–100, proteins simultaneously in 500–1000 cells, per second. This eliminates the hassle of obtaining large quantities of starting material, as would be necessary for traditional phosphoproteomic approaches. The application of this technology for phosphoprotein detection however relies on the specificity of metal isotope-tagged phospho-antibodies in fixed cells. Current detection limits are 100 molecules per cell.

In essence, the identification of phosphoproteins altered specifically in genetic contexts where a protein kinase or phosphatase is mutated, removed, or depleted should drive substrate predictions, as well as help assign function to those corresponding kinases and phosphatases. Such types of experiments have been carried out in budding yeast, *Arabidopsis*, and cell culture using mass spectrometry to measure phosphopeptide levels (Bodenmiller *et al.*, 2010; Chen *et al.*, 2010; Gnad *et al.*, 2013; Hilger *et al.*, 2009; Smolka *et al.*, 2007; Umezawa *et al.*, 2013). While large-scale application of this type has been carried out in yeast with kinase and phosphatase knockouts (Bodenmiller *et al.*, 2010), a lack of scalable genetic tools for similar systematic studies has stalled *in vivo* analyses in metazoan systems. Given the recently improved efficacy of RNAi during *Drosophila melanogaster* embryogenesis (Ni *et al.*, 2011) and the relative ease with which one can collect large numbers of staged *D. melanogaster* embryos, we generated a

library of shRNA-based RNAi reagents for analogous quantitative proteomic studies (Sopko *et al.*, 2014). The library accounts for three-quarters of kinase and phosphatase encoding genes in the *D. melanogaster* genome, 95% of which are conserved to human. The smaller genome and hence proteome of *D. melanogaster*, as compared to human, allows for more extensive peptide detection and deeper protein coverage making in *D. melanogaster* more accessible for quantitative studies. Expression in the female in *D. melanogaster* germline of efficient protein kinase and phosphatase-targeting shRNAs permits the generation of eggs depleted of these single gene products. In order to perform a quantitative assessment of all phosphorylated proteins in eggs depleted for specific protein kinases (essentially a phosphorylation signature), we employed isobaric peptide labeling, which enables concurrent quantitation of multiple shRNA-expressing genotypes relative to one another and a common control. As a proof of principle, we carried out phosphoproteomic profiling on embryos depleted of 19 different *D. melanogaster* protein kinases for the purpose of systematically linking protein phosphosites with these maternally inherited kinases. To date, in *D. melanogaster* is the most advanced organism for which whole-animal studies of this kind have been performed. While most identified phosphosites were unchanged in terms of levels relative to a mock RNAi included in all experiments, a number of known and candidate kinase-targeted phosphosites, were altered in specific knockdown conditions. Examples of known substrates of shRNA-targeted kinases that exhibited a measurable change in phosphorylation in respective kinase-depleted embryos included Cdk1, Klp61F, and Hsp83 in 'wee' deficient embryos, and Histone H3, Med13 and Stat92E in *Cdk8* deficient embryos. In fact, we found that phosphorylation of one-third of respective *D. melanogaster* orthologs of literature-curated yeast *Cdk8* substrates (Sharifpoor *et al.*, 2011) that we identified were downregulated in *Cdk8* deficient *D. melanogaster* embryos. One phosphoproteome which exhibited a large number of candidate substrates based on pathway enrichment of downregulated phosphoproteins was that generated from embryos depleted of *gilgamesh*, the *D. melanogaster* casein kinase I gamma ortholog; the majority of phosphorylation we observed downregulated resided in proteins involved in Hedgehog (Hh) and Wnt/Wingless (Wg) pathways, consistent with a characterized role for Gilgamesh in mediating signaling of these pathways (Davidson *et al.*, 2005; Hummel *et al.*, 2002; Schertel *et al.*, 2013). Specifically, we observed downregulated phosphorylation of the G protein-coupled receptor Smoothed; Combgap, a direct repressor of Cubitus interruptus (Ci) and Wg target gene expression (Campbell and Tomlinson, 2000; Song *et al.*, 2000); Debra, a mediator of Ci polyubiquitination and degradation (Dai *et al.*, 2003); Retinoid- and fatty acid-binding glycoprotein, a lipophorin that binds and regulates Hh and Wg diffusion (Panakova *et al.*, 2005); Dyrk3, the human ortholog that directly phosphorylates Ci and induces its degradation (Varjosalo *et al.*, 2008); Interference Hedgehog (Ihog) interacting proteins Modulo and Karst (Guruharsha *et al.*, 2011); and other regulators of Hh signaling, namely Smg5, Belle, Adenomatous polyposis coli tumor suppressor homolog 2, Bx42, Chromator, Pitslre, and Rm62 (Nybakken *et al.*, 2005). This list of altered phosphoproteins undoubtedly includes both direct and indirect

Gilgamesh targets. Obviously, validation of *bona fide* direct Gilgamesh substrates would require additional information such as *in vitro* kinase activity towards the potential substrate, a protein–protein interaction, and functional assays. Based on the observations described above, we are confident that identified changes in the *D. melanogaster* phosphoproteome can be used to screen for authentic substrates and alterations in signaling downstream of the targeted kinase.

As expected the numbers of direct and indirect candidate kinase targets, which we considered as those phosphoproteins whose levels were downregulated relative to mock knockdown more than 1.5 fold, varied depending on the identity of the kinase depleted (Sopko *et al.*, 2014). This was expected since the number of phosphoproteins expected to have an interaction with an RNAi-depleted kinase will differ depending on the kinase, due to multiple factors including the function of the kinase, its expression level, localization, and connectivity (physical interactions with other proteins) (Ubersax and Ferrell, 2007). The number of downregulated phosphosites for the 19 kinase-deficient contexts we surveyed ranged from 752 to 22 and correlated with the degree of phenotype resulting from kinase depletion, indicating that extensive modulation of the phosphoproteome often translates to major morphological changes. For example, embryos depleted of the cyclin-dependent kinase *Cdk8* display extreme morphological defects and fail to hatch, consistent with the major effects on the phosphoproteome that we observed from knockdown of this general transcriptional regulator. On the other end of the spectrum, *Bub1*-deficient embryos look normal, consistent with the minimal effects on the phosphoproteome observed for near complete depletion of *Bub1*, hinting that this mitotic kinase functions redundantly with its paralog, *BubR1* (also known as Mad3). Redundant kinase targets could be uncovered from phosphoproteomic profiling of embryos deficient for two redundant kinases, for instance *Bub1* and *BubR1*, by co-expressing shRNAs targeting both, or one kinase-targeting shRNA in a background mutant for the other kinase. Such a scenario would reveal shared substrates, since the compensating kinase activity would be absent or hampered. Early *D. melanogaster* embryos serve up a superb experimental advantage in this respect since one can restrict RNAi specifically to the germline and hence negate any effects on viability, since an intact germline is dispensable for organismal development. In this way one can examine genetic scenarios where two mutations would be synthetic lethal and otherwise unattainable, for instance in yeast where the combination would result in arrested growth. Additionally, one can fine-tune the activity of the RNAi driver with the Gal4/UAS system in *D. melanogaster* (Brand and Perrimon, 1993) and therefore modulate the extent of knockdown, by altering temperature to look at graded effects (Staller *et al.*, 2013). Though single kinase knockdown alone may be sufficient to hint at redundancy, if kinase activity is hindered sufficiently to warrant detectable phosphosite alteration. In this case, individual phosphoprofiles for redundant kinases would exhibit a similar trend in change for compromised phosphorylation at shared sites.

Overlapping phosphorylation alteration patterns for any two kinases could additionally indicate participation of the two kinases in a common pathway, since one would expect shared changes in phosphorylation for targets downstream of

both kinases. One could decipher where they function by the extent to which they share similarity with phosphoprofiles generated from knockdown of known components of characterized signaling pathways. For example, to identify new components and targets of MAPK signaling, one could first define the individual phosphoprofiles associated with knockdown of Raf, MEK, and ERK, three kinases directly acting on one another in a signaling cascade. The phosphorylation of peptides representing true output of the pathway should trend in a similar direction for all three kinase knockdowns – either all up or all down – relative to knockdown controls, while phosphorylation events unique to each kinase will not. Phosphoprofiles generated from knockdown of genes in the same pathway would be qualitatively similar and share this pattern or partial pattern (Figure 2). This concept is similar to the approach that has been used extremely successfully in developmental genetics to identify components of signaling pathways, whereby genes with similar mutant phenotypes usually encode molecules that are part of the same biochemical pathways (Green *et al.*, 2011; Nusslein-Volhard and Wieschaus, 1980). This approach has been extremely powerful in particular using the *D. melanogaster* embryonic cuticle as a phenotypic readout. In relation to phosphoproteomic data, one could infer signaling hierarchy by substituting comparisons of cuticle phenotypes for similarity in patterns of thousands of phosphorylation measurements, generated from mass spectrometry analysis.

Additional information in terms of mapping network connectivity and shared enzymatic targeting could be gleaned from combining perturbations of protein kinases and phosphatases. For example, a shared protein kinase and phosphatase substrate should exhibit an exacerbated reduction in phosphorylation in a context where protein kinase perturbation, knockdown or mutation, is combined with overexpression of a counteractive phosphatase, as compared to kinase knockdown or phosphatase overexpression alone. In this scenario, the addition of phosphate onto the shared target would be hindered and the removal of phosphate enhanced. Conversely, overexpression of the culpable kinase and a reduction in opposing phosphatase activity could lead to elevated phosphorylation. Technically, an increase in phosphorylation may in fact be easier to detect by mass spectrometry than a decrease, though these types of combinatorial experiments have not yet been reported. These experiments should be relatively easy in *D. melanogaster* and budding yeast given the large number of available P-element insertion lines bearing UAS transgenes for Gal4-mediated overexpression and plasmid-based libraries for Gal4- and endogenous promoter-driven expression, respectively. Analogous genetic manipulations as those described above, to generate scenarios of target hyperactivity, have been demonstrated to result in altered cell growth and a means to successfully predict kinase substrates (Fiedler *et al.*, 2009; Sopko *et al.*, 2006). In these cases, growth was the phenotypic readout, however gene expression has also been used as a measure to predict kinase and phosphatase targets and redundancy between these enzymes (van Wageningen *et al.*, 2010). We expect that global protein phosphorylation measurements by mass spectrometry should serve as a better predictor of these relationships as compared to global gene expression, since these enzymes

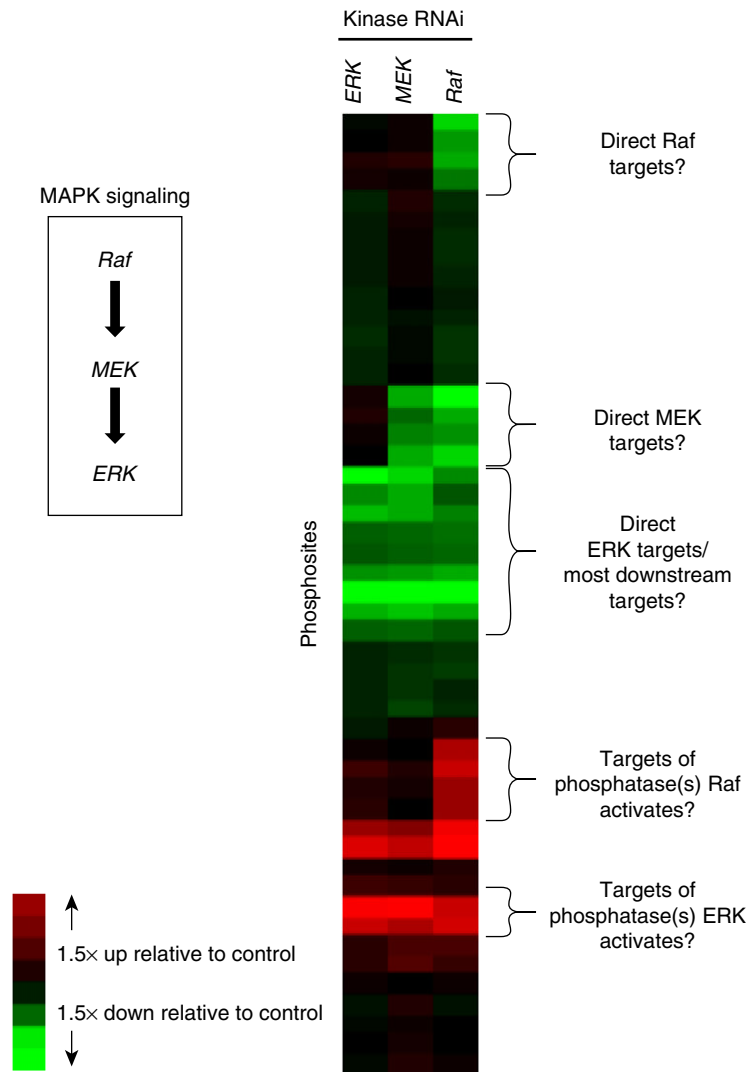


Figure 2 Clustering of hypothetical phosphoproteomes generated from mass spectrometry of kinase-depleted cells by RNAi. Phosphosite level measurements generated from mass spectrometry of cells subjected to RNAi of MAP kinase (MAPK) signaling pathways components, relative to a control RNAi, are illustrated with a sliding color scale. Phosphorylation of the most downstream substrates of the pathway, directly targeted by ERK, exhibit downregulation in their levels for all three pathway kinase knockdowns while direct targets of individual upstream kinases in the signaling hierarchy are those downregulated phosphosites that lack overlap with knockdown of downstream kinases. Substrates of phosphatases activated by the depleted kinase can be deciphered as those upregulated phosphoproteins, using the same logic.

directly impact phosphorylation. In this way not only would global proteomic analyses map pathways, but could also be used to reveal critical nodes in signaling that may partially or completely overcome mutations resulting in pathway hyperactivity.

Because the number of protein phosphatases is under-represented in the genome relative to protein kinases, one might speculate their signatures would be less specific than protein kinases and their knockdown would affect larger proportions of the phosphoproteome. In fact, there appears to be no correlation between phosphatase ablation and impact on the phosphoproteome, relative to that of protein kinases – at least from the examination of yeast null mutants (Bodenmiller *et al.*, 2010). However, this observation could in fact be related to compensation for missing activity in deletion mutants over generations, as has been persuasively reported in

yeast (Teng *et al.*, 2013). Further support for the adaptation of null mutants is observations of different phenotypes depending on whether a kinase is targeted by a small molecule inhibitor or by complete deletion (Carroll *et al.*, 2001; Knight *et al.*, 2006). Consistent with this, Bodenmiller *et al.* (2010) observed a higher proportion of changes reflective of indirect effects with permanent deletion mutants as compared to short temporal inhibition of analog-sensitive kinases with small molecules. Conceivably, by comparing deletion mutants over generations or developmental time, one could identify compensatory rewiring events by identifying altered phosphorylation-dependent signaling by mass spectrometry. Further, by comparing complete gene knockout to incomplete depletion by RNAi-mediated knockdown with the same approach, one could identify off-target effects of RNAi reagents and compensatory effects resulting from gene knockdown, for example,

the stabilization of alternative isoforms (Figure 3). Such types of analyses have recently become more accessible given current advances to efficiently make deleterious mutations in *D. melanogaster* using CRISPR (Kondo, 2014).

In our recent study we demonstrated how kinase–substrate relationships may be illuminated from examining correlations in phosphosite changes measured by mass spectrometry (Sopko *et al.*, 2014). Specifically, we revealed how two phosphosites, correlating in their directionality of change among multiple

genetic contexts, multiple kinase-deficient conditions in our study, have a tendency to comprise a kinase–substrate pair. For instance, correlation of kinase autophosphorylation and phosphorylation of another protein can indicate activation of the phosphoprotein by the kinase, either directly or indirectly (Figure 4, top panel: positive correlation). On the other hand, inhibitory phosphorylation of a kinase will always be out of phase with phosphorylation of that kinase’s targets (Figure 4, bottom panel: negative correlation). We found enrichment for

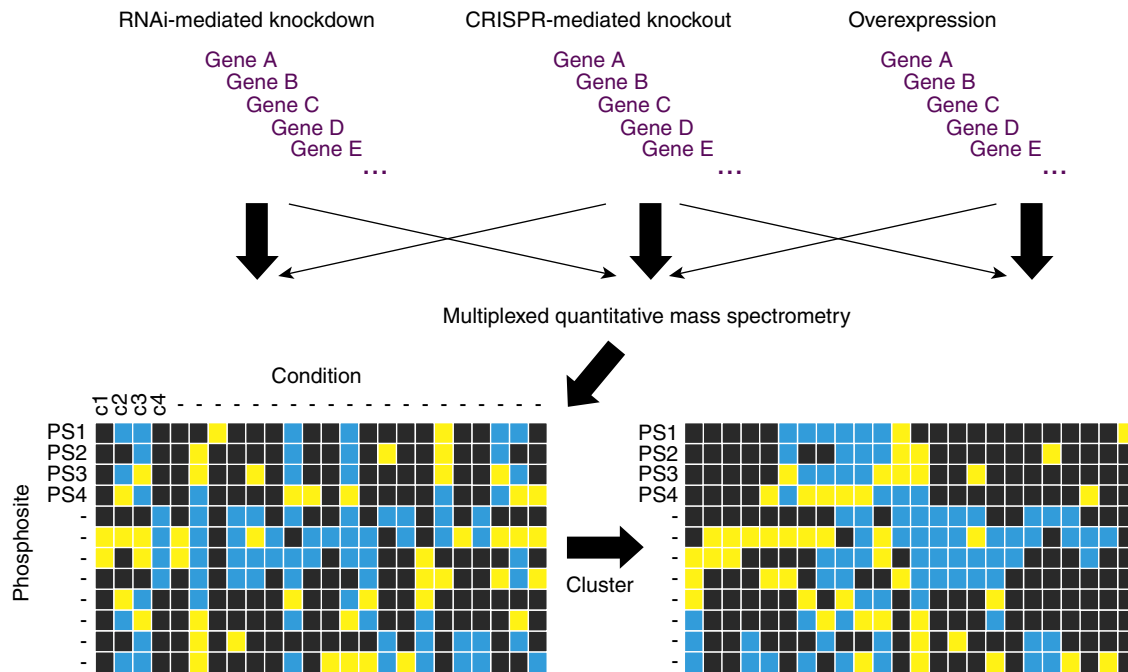


Figure 3 A workflow for integrating systems genetics and proteomics to identify similar responses to genetic manipulation. A phosphoprofile is generated from phosphosites (PS) identified by multiplexed mass spectrometry methods analyzing genetically modified cells or organisms, manipulated by conditions (c) of RNAi, CRISPR-mediated knockout, overexpression, or combinations thereof. Clustering of phosphoprofiles generated from individual conditions are dictated by PS alterations, either down (cyan) or up (yellow) in levels relative to control. Neighboring phosphoprofiles post-clustering share similarity in their cellular response to the specific genetic manipulation(s).

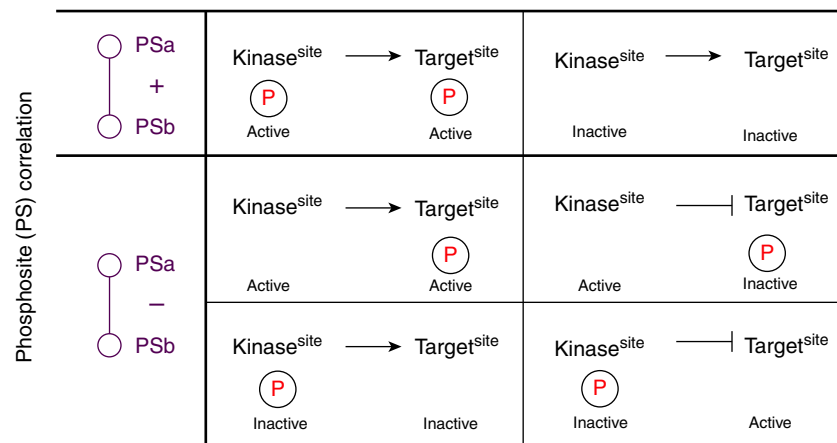


Figure 4 Molecular scenarios explaining correlative and anti-correlative phosphorylation. Two phosphosites (PS) consistently responding in the same direction (positive correlation; +) or the opposite direction (negative correlation; -) in different genetic contexts can illuminate phosphosite functionality. Correlations, however, are not enough to distinguish between direct and indirect kinase targets.

genuine kinase-substrate pairs among 447 585 correlative pairs involving 2058 altered phosphosites (>1.5-fold relative to control RNAi) we identified from profiling 19 kinase-deficient genotypes. Significantly, the probability of identifying kinase-substrate pairs increased with the number of kinase-deficient conditions for which the directionality of change for the phosphosite pair correlated, positively or negatively. Application of this type of correlative analysis for the mining of existing large-scale phosphoproteomic data from different organisms and tissues has enormous implications in terms of mapping kinase-phosphatase networks and deciphering functional phosphorylation. Particularly, because uncovering functionality for a particular phosphosite does not rely on direct manipulation of the respective phosphoprotein, fewer experimental conditions than genetic manipulations are required to interpret the consequences of individual phosphorylation events. Moreover, having *a priori* knowledge of the function of a specific phosphosite – serving in an activating or inhibitory capacity – can facilitate interpretation of the other participating phosphosite of the correlative pair. We exemplified this in our study with a pair of phosphorylations that correlated positively in six kinase-depleted conditions: activating phosphorylation of the transcription factor STAT and a previously uncharacterized Sterile20-like kinase (Slik) phosphosite. Since this activating STAT phosphorylation is reduced in 'slik' deficient embryos, we predicted that Slik acts upstream of STAT and that phosphorylation of Slik kinase is required to promote STAT phosphorylation and DNA binding. We verified this prediction, by demonstrating that transcriptional targets of STAT are reduced in expression following *slik* knockdown. Our novel approach examining correlative phosphorylation to map kinase signaling uncovered a previously unrecognized role for Slik and pinpointed functionality of an individual Slik phosphosite. We anticipate that a similar strategy surveying the correlation of posttranslational modifications other than phosphorylation, derived *en masse* from mass spectrometry measurements, will shed light on the identity and mechanisms of relevant regulatory enzymes.

Given the large number of human kinase mutations associated with disease, understanding the global consequences of specific phosphorylation-impacted perturbations could prompt treatment tailored to aberrant signaling of specific pathways (Lahiry *et al.*, 2010). While current methods such as mass spectrometry can easily identify altered phosphorylation between samples, understanding the mechanisms by which these changes arise and deciphering impactful versus inconsequential changes on cellular and organismal phenotype remains a major challenge. We have recently reported how global phosphorylation surveyed under various genetic contexts can be used to generate a matrix of phosphoprofiles (Sopko *et al.*, 2014), permitting profile comparisons and correlative phosphosite analysis to predict kinase redundancy and map signaling networks. We expect that matrix expansion, incorporating additional genetic (for instance, combinatorial gene overexpression, or gene ablation or gene product depletion) and environmental contexts, will strengthen kinase-substrate and other types of network signaling mechanism predictions. This generic strategy is attractive given it can be applied to any organism or cell type. In conclusion, while the ability to catalog cellular measurements – DNA mutations, transcript levels, splice isoforms, protein levels and

modifications – is becoming cheaper, easier, and more comprehensive, efforts to understand the consequences of changes in cellular machinery are lacking and require the inception of new analysis methods, such as those described above.

Acknowledgments

The National Institutes of Health supported this work (5R01DK088718, 5P01CA120964). R.S. is a Special Fellow of the Leukemia and Lymphoma Society. N.P. is a Howard Hughes Medical Institute investigator.

See also: Dynamic Integration: Dynamics: Dynamics of Protein Kinase Cascades. Horizontal Integration: Omics: Phosphoproteomics: Approaches, Developments, and Challenges

References

- Ahrens, C.H., Brunner, E., Qeli, E., Basler, K., Aebersold, R., 2010. Generating and navigating proteome maps using mass spectrometry. *Nature Reviews Molecular Cell Biology* 11, 789–801.
- Alexander, J., Lim, D., Joughin, B.A., *et al.*, 2011. Spatial exclusivity combined with positive and negative selection of phosphorylation motifs is the basis for context-dependent mitotic signaling. *Science Signaling* 4, ra42.
- Baldwin, A., Li, W., Grace, M., *et al.*, 2008. Kinase requirements in human cells: II. Genetic interaction screens identify kinase requirements following HPV16 E7 expression in cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* 105, 16478–16483.
- Bendall, S.C., Nolan, G.P., Roederer, M., Chattopadhyay, P.K., 2012. A deep profiler's guide to cytometry. *Trends in Immunology* 33, 323–332.
- Blethrow, J.D., Glavy, J.S., Morgan, D.O., Shokat, K.M., 2008. Covalent capture of kinase-specific phosphopeptides reveals Cdk1-cyclin B substrates. *Proceedings of the National Academy of Sciences of the United States of America* 105, 1442–1447.
- Bodenmiller, B., Wanka, S., Kraft, C., *et al.*, 2010. Phosphoproteomic analysis reveals interconnected system-wide responses to perturbations of kinases and phosphatases in yeast. *Science Signaling* 3, rs4.
- Bommi-Reddy, A., Almeciga, I., Sawyer, J., *et al.*, 2008. Kinase requirements in human cells: III. Altered kinase requirements in VHL-/- cancer cells detected in a pilot synthetic lethal screen. *Proceedings of the National Academy of Sciences of the United States of America* 105, 16484–16489.
- Boyle, S.N., Koleske, A.J., 2007. Dissecting kinase signaling pathways. *Drug Discovery Today* 12, 717–724.
- Brand, A.H., Perrimon, N., 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118, 401–415.
- Breitkreutz, A., Choi, H., Sharom, J.R., *et al.*, 2010. A global protein kinase and phosphatase interaction network in yeast. *Science* 328, 1043–1046.
- Campbell, G.L., Tomlinson, A., 2000. Transcriptional regulation of the Hedgehog effector Ci by the zinc-finger gene *combap*. *Development* 127, 4095–4103.
- Carlson, S.M., Chouinard, C.R., Labadorf, A., *et al.*, 2011. Large-scale discovery of ERK2 substrates identifies ERK-mediated transcriptional regulation by ETV3. *Science Signaling* 4, rs11.
- Carroll, A.S., Bishop, A.C., DeRisi, J.L., Shokat, K.M., O'Shea, E.K., 2001. Chemical inhibition of the Pho85 cyclin-dependent kinase reveals a role in the environmental stress response. *Proceedings of the National Academy of Sciences of the United States of America* 98, 12578–12583.
- Chen, S.H., Albuquerque, C.P., Liang, J., Suhandynata, R.T., Zhou, H., 2010. A proteome-wide analysis of kinase-substrate network in the DNA damage response. *Journal of Biological Chemistry* 285, 12803–12812.
- Dai, P., Akimaru, H., Ishii, S., 2003. A hedgehog-responsive region in the *Drosophila* wing disc is defined by debra-mediated ubiquitination and lysosomal degradation of Ci. *Developmental Cell* 4, 917–928.

- Davidson, G., Wu, W., Shen, J., *et al.*, 2005. Casein kinase 1 gamma couples Wnt receptor activation to cytoplasmic signal transduction. *Nature* 438, 867–872.
- Dephoure, N., Howson, R.W., Blethrow, J.D., Shokat, K.M., O'Shea, E.K., 2005. Combining chemical genetics and proteomics to identify protein kinase substrates. *Proceedings of the National Academy of Sciences of the United States of America* 102, 17940–17945.
- Fasolo, J., Stoner, A., Sun, M.G., *et al.*, 2011. Diverse protein kinase interactions identified by protein microarrays reveal novel connections between cellular processes. *Genes & Development* 25, 767–778.
- Fiedler, D., Braberg, H., Mehta, M., *et al.*, 2009. Functional organization of the *S. cerevisiae* phosphorylation network. *Cell* 136, 952–963.
- Gao, Z.H., Metherall, J., Virshup, D.M., 2000. Identification of casein kinase I substrates by *in vitro* expression cloning screening. *Biochemical and Biophysical Research Communications* 268, 562–566.
- Gnad, F., Young, A., Zhou, W., *et al.*, 2013. Systems-wide analysis of K-Ras, Cdc42, and PKA signaling by quantitative phosphoproteomics. *Molecular & Cellular Proteomics* 12, 2070–2080.
- Green, R.A., Kao, H.L., Audhya, A., *et al.*, 2011. A high-resolution *C. elegans* essential gene network based on phenotypic profiling of a complex tissue. *Cell* 145, 470–482.
- Gulmann, C., Sheehan, K.M., Conroy, R.M., *et al.*, 2009. Quantitative cell signalling analysis reveals down-regulation of MAPK pathway activation in colorectal cancer. *Journal of Pathology* 218, 514–519.
- Guruharsha, K.G., Rual, J.F., Zhai, B., *et al.*, 2011. A protein complex network of *Drosophila melanogaster*. *Cell* 147, 690–703.
- Hilger, M., Bonaldi, T., Gnad, F., Mann, M., 2009. Systems-wide analysis of a phosphatase knock-down by quantitative proteomics and phosphoproteomics. *Molecular & Cellular Proteomics* 8, 1908–1920.
- Holt, L.J., Tuch, B.B., Villen, J., *et al.*, 2009. Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* 325, 1682–1686.
- Hummel, T., Attix, S., Gunning, D., Zipursky, S.L., 2002. Temporal control of glial cell migration in the *Drosophila* eye requires gilgamesh, hedgehog, and eye specification genes. *Neuron* 33, 193–203.
- Hutti, J.E., Jarrell, E.T., Chang, J.D., *et al.*, 2004. A rapid method for determining protein kinase phosphorylation specificity. *Nature Methods* 1, 27–29.
- Jünger, M.A., Aebersold, R., 2013. Mass spectrometry-driven phosphoproteomics: Patterning the systems biology mosaic. *WIREs Developmental Biology* 3, 83–112.
- Kaushansky, A., Gordus, A., Chang, B., Rush, J., MacBeath, G., 2008. A quantitative study of the recruitment potential of all intracellular tyrosine residues on EGFR, FGFR1 and IGF1R. *Molecular BioSystems* 4, 643–653.
- Knight, Z.A., Gonzalez, B., Feldman, M.E., *et al.*, 2006. A pharmacological map of the PI3-K family defines a role for p110alpha in insulin signaling. *Cell* 125, 733–747.
- Kondo, S., 2014. New horizons in genome engineering of *Drosophila melanogaster*. *Genes & Genetic Systems* 89, 3–8.
- Kubota, K., Anjum, R., Yu, Y., *et al.*, 2009. Sensitive multiplexed analysis of kinase activities and activity-based kinase identification. *Nature Biotechnology* 27, 933–940.
- Lahiry, P., Torkamani, A., Schork, N.J., Hegele, R.A., 2010. Kinase mutations in human disease: Interpreting genotype-phenotype relationships. *Nature Reviews Genetics* 11, 60–74.
- Lee, L.A., Lee, E., Anderson, M.A., *et al.*, 2005. *Drosophila* genome-scale screen for PAN GU kinase substrates identifies Mat9Bb as a cell cycle regulator. *Developmental Cell* 8, 435–442.
- Lienhard, G.E., 2008. Non-functional phosphorylations? *Trends in Biochemical Sciences* 33, 351–352.
- Mok, J., Kim, P.M., Lam, H.Y., *et al.*, 2010. Deciphering protein kinase specificity through large-scale analysis of yeast phosphorylation site motifs. *Science Signaling* 3, ra12.
- Ni, J.Q., Zhou, R., Czech, B., *et al.*, 2011. A genome-scale shRNA resource for transgenic RNAi in *Drosophila*. *Nature Methods* 8, 405–407.
- Nusslein-Volhard, C., Wieschaus, E., 1980. Mutations affecting segment number and polarity in *Drosophila*. *Nature* 287, 795–801.
- Nybakken, K., Vokes, S.A., Lin, T.Y., McMahon, A.P., Perrimon, N., 2005. A genome-wide RNA interference screen in *Drosophila melanogaster* cells for new components of the Hh signaling pathway. *Nature Genetics* 37, 1323–1332.
- Panakova, D., Sprong, H., Marois, E., Thiele, C., Eaton, S., 2005. Lipoprotein particles are required for Hedgehog and Wingless signalling. *Nature* 435, 58–65.
- Picotti, P., Clement-Ziza, M., Lam, H., *et al.*, 2013. A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. *Nature* 494, 266–270.
- Placek, J., Devgan, G., Michaud, G., *et al.*, 2005. Global analysis of protein phosphorylation in yeast. *Nature* 438, 679–684.
- Rohila, J.S., Chen, M., Chen, S., *et al.*, 2009. Protein-protein interactions of tandem affinity purified protein kinases from rice. *PLoS One* 4, e6685.
- Schertel, C., Huang, D., Bjorklund, M., *et al.*, 2013. Systematic screening of a *Drosophila* ORF library *in vivo* uncovers Wnt/Wg pathway components. *Developmental Cell* 25, 207–219.
- Sharifpoor, S., Nguyen, B., A.N., Youn, J.Y., *et al.*, 2011. A quantitative literature-curated gold standard for kinase-substrate pairs. *Genome Biology* 12, R39.
- Smolka, M.B., Albuquerque, C.P., Chen, S.H., Zhou, H., 2007. Proteome-wide identification of *in vivo* targets of DNA damage checkpoint kinases. *Proceedings of the National Academy of Sciences of the United States of America* 104, 10364–10369.
- Song, Y., Chung, S., Kunes, S., 2000. Combgap relays wingless signal reception to the determination of cortical cell fate in the *Drosophila* visual system. *Molecular Cell* 6, 1143–1154.
- Songyang, Z., Blechner, S., Hoagland, N., *et al.*, 1994. Use of an oriented peptide library to determine the optimal substrates of protein kinases. *Current Biology* 4, 973–982.
- Sopko, R., Andrews, B.J., 2008. Linking the kinome and phosphorylome — A comprehensive review of approaches to find kinase targets. *Molecular BioSystems* 4, 920–933.
- Sopko, R., Foos, M., Vinayagam, A., *et al.*, 2014. Combining genetic perturbations and proteomics to examine kinase-phosphatase networks in *Drosophila* embryos. *Developmental Cell* 31 (1), 114–127.
- Sopko, R., Huang, D., Preston, N., *et al.*, 2006. Mapping pathways and phenotypes by systematic gene overexpression. *Molecular Cell* 21, 319–330.
- Staller, M.V., Yan, D., Randklev, S., *et al.*, 2013. Depleting gene activities in early *Drosophila* embryos with the "maternal-Gal4-shRNA" system. *Genetics* 193, 51–61.
- Tan, C.S., Linding, R., 2009. Experimental and computational tools useful for (re) construction of dynamic kinase-substrate networks. *Proteomics* 9, 5233–5242.
- Teng, X., Dayhoff-Brannigan, M., Cheng, W.C., *et al.*, 2013. Genome-wide consequences of deleting any single gene. *Molecular Cell* 54 (4), 485–494.
- Turk, B.E., Hutti, J.E., Cantley, L.C., 2006. Determining protein kinase substrate specificity by parallel solution-phase assay of large numbers of peptide substrates. *Nature Protocols* 1, 375–379.
- Ubersax, J.A., Ferrell Jr., J.E., 2007. Mechanisms of specificity in protein phosphorylation. *Nature Reviews Molecular Cell Biology* 8, 530–541.
- Ubersax, J.A., Woodbury, E.L., Quang, P.N., *et al.*, 2003. Targets of the cyclin-dependent kinase Cdk1. *Nature* 425, 859–864.
- Umezawa, T., Sugiyama, N., Takahashi, F., *et al.*, 2013. Genetics and phosphoproteomics reveal a protein phosphorylation network in the abscisic acid signaling pathway in *Arabidopsis thaliana*. *Science Signaling* 6, rs8.
- Varjosalo, M., Bjorklund, M., Cheng, F., *et al.*, 2008. Application of active and kinase-deficient kinome collection for identification of kinases regulating hedgehog signaling. *Cell* 133, 537–548.
- van Wageningen, S., Kemmeren, P., Lijnzaad, P., *et al.*, 2010. Functional overlap and regulatory links shape genetic interactions between signaling pathways. *Cell* 143, 991–1004.

High-Content Screening in Cell Biology

R Ketteler and J Kriston-Vizi, University College London, London, UK

© 2016 Elsevier Inc. All rights reserved.

Glossary

Arrayed library Arrangement of a collection ('library') of samples in specified positions on a microplate, slide, or chip.

Cellular phenotype Displayed characteristics or traits of a cell observable by microscopy.

Clustered regularly interspaced short palindromic repeats (CRISPR) genome editing A method based on a bacterial defense mechanism that allows targeted gene modifications.

Microarray Arrangement of large sample numbers on a small localized surface such as a chip or slide.

Over-sampling Sampling a higher number of images than required for 100% coverage in order to improve resolution and reduce noise.

RNA interference RNA-based technology to reduce gene expression levels in cells.

Texture feature Metric that defines pixel patterns in an image and statistically describes pixel distributions.

Introduction

An image says more than a thousand words. This is particularly true in the field of cell biology, which has always been based on the ability to visualize the cell and its components. In fact, cell biology became possible with the invention of the microscope, thus allowing for the first time to actually see a cell and define its contents. Over the decades and years that followed, revolutions in the ability to capture images enabled further, deeper insights into how cells work and function. The ability to resolve finer structures with instruments able to capture higher resolution images – through transmission electron microscopy in the 1950s and super-resolution microscopy nowadays – have led to astounding detail in understanding very small subcellular structures. The development of automated microscopy platforms in the late 1990s enabled the capture of thousands of images from biological specimens in a relatively short time frame, at moderate cost and with high resolution. High-content screening (HCS), i.e., the automated capture of high amount of cellular image data, was born.

History

HCS was made possible by the development of a microscope that is capable of acquiring a large number of images through an automated process (Taylor, 2010). Historically, HCS can be viewed as a direct consequence of earlier technical innovations: one was the development of the charged coupled device camera (CCD) that enabled the conversion of image data into numeric data values. Combined with improvements in cell-based assays using immunofluorescence techniques and the development of fluorescent protein-based reporter systems, it was possible to record a variety of cellular phenotypes with microscopy-based methods. The design of high-throughput genomic sequencers with automated instrumentation that were required for the human genome project in the early 1990s was paralleled by an interest to develop similar highly automated processes for the capture of large amounts of image data (Taylor, 2007). Lansing Taylor and

colleagues at Cellomics™ developed the first automated fluorescence microscope capable of acquiring a large number of images in an arrayed format, the Arrayscan™ in 1996. Quickly after the launch of the Arrayscan™, several other HCS platforms were developed and adopted by both pharmaceutical and academic research institutes. Another breakthrough that brought HCS to a larger research community was the availability of RNA interference libraries in the early 2000s that enabled the use of such libraries in loss-of-function screens. Following the identification of most coding gene products in the human genome and the genome of other model organisms, large-scale arrayed genomic libraries became available that enabled the analysis of cellular gene function on a genomic scale (Figure 1) (Boutros *et al.*, 2004; Silva *et al.*, 2004; Moffat *et al.*, 2006). These developments further consolidated the use of HCS in many research labs as a primary method of choice for both basic scientific research and drug discovery. Most major universities and research institutions now house at least one HCS platform. The most striking advantage of HCS over common fluorescence microscopy techniques is the possibility to analyze a large number of samples in a relatively short time frame. In fact, recent developments in the imaging hardware has led to instruments capable of acquiring high-resolution images that rivals the image quality of conventional confocal microscopes.

'Content' is typically in the form of an image or video that will be converted into binary data values by a process termed image analysis. Sometimes the terms HCS and high-throughput screening are used interchangeably. This is not necessarily correct as high-throughput screening refers to screening a large set of samples with single data values as output, whereas HCS refers to extracting a large number of features from a large set of images. Today, HCS is used in translational drug discovery as well as fundamental academic research. Applications range from the identification of chemical hit compounds in primary drug screening campaigns and profiling of cellular responses to perturbations to studying fundamental questions of cell mitosis, morphology, and heterogeneity. HCS has become quite complex and can make use of diverse techniques such as live cell imaging, whole animal imaging, spheroid cultures, and others that will be discussed below.

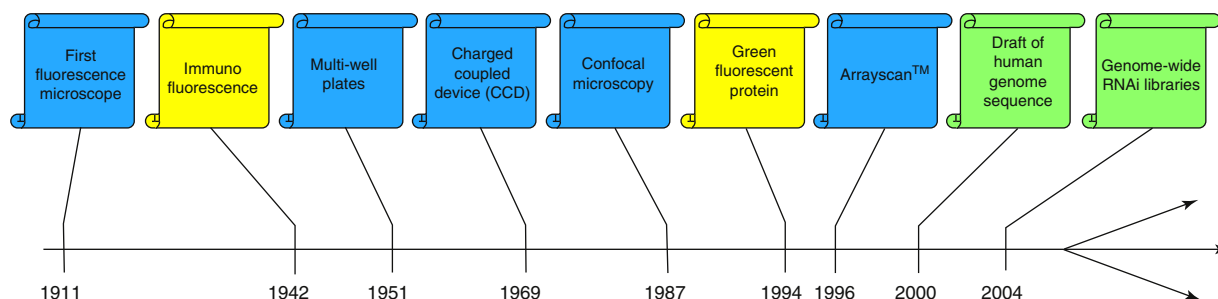


Figure 1 Milestones in HCS. Main developments in the hardware required for HCS leading to the development of the Arrayscan™ are shown in blue color. Improvements in cellular assay technologies with the two main reporter types – Immunofluorescence and fluorescent reporter proteins – are shown in yellow. Milestones in functional genomics that led to improved arrayed libraries are shown in green.

Experimental Setup

There are six basic components for HCS experiments:

1. cells,
2. perturbations,
3. assays,
4. instrumentation,
5. image analysis, and
6. statistics.

A schematic workflow diagram for HCS is shown in **Figure 2** that will be explained in more detail below. Briefly, a cell-based system that is best suited to display the phenotype of choice will be subjected to a chemical, biological, or genetic perturbation. The resulting phenotypic changes that are visualized by specific assays will then be recorded on a HCS imaging plate reader. The images will be converted into image data through image analysis and statistically evaluated.

Cells

The cell-based model used in HCS can be relatively simple (e.g., a single layer of a standard cancer cell line such as HeLa cells) or quite complex (e.g., whole zebrafish embryos). Depending on the choice of model system, there are a number of challenges that need to be overcome. Some examples of cell-based model systems that are typically used in HCS are shown in **Table 1**. The choice of cell system is largely determined by the desired output. For instance, when the main aim is to identify hit compounds from a large collection of chemical compounds, a 2D layer of cancer cells with an unlimited supply of cells can be a good choice. Similarly, in whole genome siRNA knockdown experiments, the use of generic cell lines such as HeLa or Schneider S2 have shown great utility (Kiger *et al.*, 2003; Boutros *et al.*, 2004; Eggert *et al.*, 2004). Some cell biology questions require the use of specialized cell types or primary cells that may better reflect the physiological phenotype. However, when supply of primary cells is limited, there may be restrictions in the ability to perform high-throughput screening and smaller libraries are used. Another consideration is that primary cells from different donors are quite variable and can introduce complications for phenotypic analysis. One possibility is pooling of samples from multiple donors to overcome this problem, although for

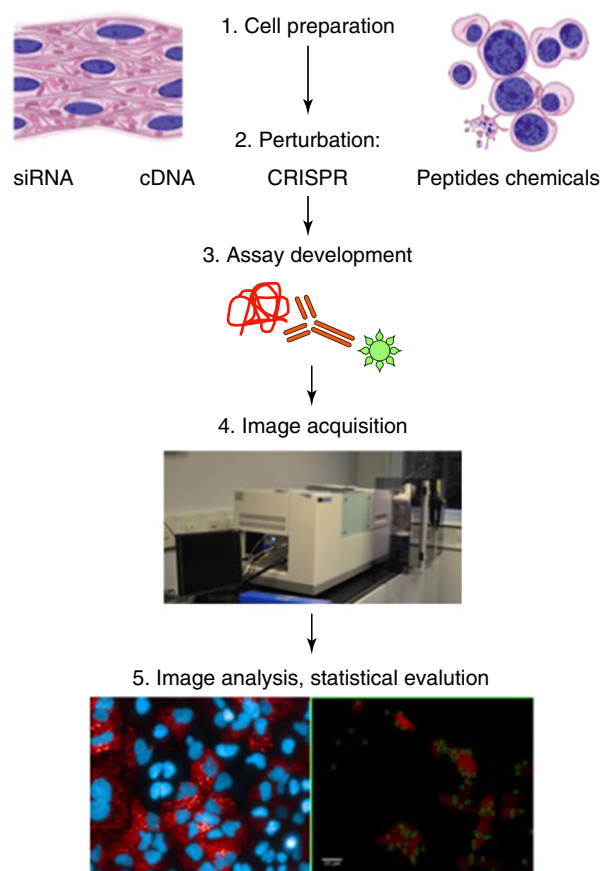


Figure 2 Schematic overview of steps in HCS experiments. A typical HCS experiment proceeds in five steps. Cells are subjected to a perturbation using siRNA, cDNA, CRISPR, peptide, or chemical libraries. A suitable assay for monitoring a cellular phenotype will be used such as immunostaining of a protein with a fluorescently tagged antibody as depicted here. Images of the cells will be acquired in a dedicated high-content imaging platform such as the PerkinElmer OperaLX. In the end, the images will be processed and subjected to image analysis in order to extract data values. These data values require stringent statistical evaluation.

some cell types such as immune reactive cells this is not possible.

There are two ways to feed cells into the HCS experiment. One is to deposit cells into standardized microplates (96-,

Table 1 Examples of cell models commonly used in high-content screening applications

<i>Cell model</i>	<i>Culture type</i>	<i>Applications</i>	<i>Scale</i>	<i>Culture time</i>	<i>Comments</i>
HeLa cells	Monolayer	cDNA, siRNA, and drug	HTS	Hours–days	Other commonly used adherent transformed cell lines are A549, HepG2, HEK293 cells. Large quantities of cells are available, thus enabling higher throughput applications
Fibroblasts	Monolayer	cDNA, siRNA, and drug	MTS	Hours–days	Fibroblasts are available from genetic mouse models, thus offering the use of a genetically modified system for HCS. Cell numbers can be limiting
Schneider S2	Monolayer	siRNA	HTS	Days	Drosophila cell lines such as Schneider S2 have proven highly useful for investigation of cellular processes and have pioneered some of the early siRNA screening work
Primary neurons	Monolayer	siRNA and drug	MTS	Days–weeks	Primary neurons from rat or murine sources can be limiting in number. Culture conditions require a higher level of experience and image analysis of neurons can be quite challenging
iPS-derived cells	Monolayer or 3D	siRNA and drug	MTS	Weeks–months	Induced pluripotent cells from patients' fibroblasts offer a unique opportunity to study genetic diseases in a human cell type. Differentiation of these cells can take weeks and is a challenge for experimental setup.
Tissues, embryos, worms	3D	Morpholino, siRNA, and drugs	LTS	Days–weeks	Zebrafish and Drosophila embryos have been used in HCS, as well as <i>Caenorhabditis elegans</i> worms. Imaging depth and image analysis remains a challenge for the investigator
Tumor spheroids	3D	Drugs	LTS	Weeks	Multiple systems for generating 3D models of cells are available from commercial vendors. An increased use of 3D models systems in HCS can be expected in the future

Abbreviations: HTS – high-throughput screening; LTS – low-throughput screening; MTS – medium throughput screening.

384-, or 1536-well format) that are acceptable for most HCS platforms. The microplates can be pretreated with agents that help cells attach or differentiate. Subsequently, the cells are subjected to a perturbation experiment as outlined below. This approach is less stressful for the cells, as they are given time to recover from cell seeding and is usually done when pretreatments are necessary such as addition of growth factors or differentiation factors. An alternative method is to deposit cells on top of a perturbation reagent, for instance onto spotted siRNA oligonucleotides (Erfle *et al.*, 2007) or surface-coupled chemical entities (Miyazaki *et al.*, 2010). This approach has the advantage that it allows miniaturization and ultra-high-throughput screening through the production of dense microarrays on slides.

The whole field of HCS in drug discovery is currently undergoing a paradigm shift, recognizing that adherent 2D cultures may not be accurately reflecting the physiological tissue context. Either tissue or whole animal imaging, or – in a drug screening pipeline – the use of 3D model systems will become an important step forward toward realizing HCS in cell biology under more physiological conditions. For instance, a high-content imaging platform specifically designed for quantitative evaluation of 3D tumor models has recently been developed (Celli *et al.*, 2014). Another exciting development in this regard is the design of the Vertebrate Automated Screening Technology (VAST) BioImager™ platform for whole zebrafish organ-level imaging (Pardo-Martin *et al.*, 2010).

Another area of intense research is the use of induced pluripotent stem (iPS) cells in high-throughput screening (Desbordes and Studer, 2013; Sherman *et al.*, 2011; Földes *et al.*, 2014). iPS cells derived from human fibroblasts can be differentiated into a wide range of cell types, thus enabling for the first time HCS experiments in primary human cells that are otherwise not accessible such as neurons or cardiomyocytes. One of the biggest challenges is that due to the long differentiation protocols the cells may need to be cultured for several days, if not weeks, in the same microplates before images can be acquired. Such modalities require further optimization of the cell sample preparation and culture in order to be applicable to large-scale HCS measurements.

Perturbations

Next, the cell-based system is subjected to a perturbation such as small molecule treatment or siRNA-mediated gene knock-down. The aim of such perturbations is to identify hit compounds within a large collection of chemical entities or to identify genes involved in cellular pathways. Perturbations can be any type of arrayed or pooled library including small molecule compounds, siRNA, cDNA, peptides, antibodies, or clustered regularly interspaced short palindromic repeats (CRISPR)-based genome-editing tools (Heintze *et al.*, 2013).

There are some instances where high-content imaging has been used to profile populations of cells without any perturbation treatment, for instance when assessing cellular (Snijder *et al.*, 2012) or organelle heterogeneity (Ferraro *et al.*, 2014). The main challenge for perturbation experiments is delivery of the reagent. Some reagents such as siRNA, cDNA, and genome-editing tools have to be delivered into the cytoplasm or nucleus of the cell. This is most commonly achieved by transfection using lipofection-, chemical-, or electroporation-based methods. In some cases, there is no need to deliver the component into the cell, for instance when antibodies or compounds are profiled for binding to cell surface receptors. Small molecule compounds generally don't require transfection, although electroporation may help in the delivery of small molecules that are not cell permeable.

There are two methods for reagent addition as outlined above: in reverse transfection methods, the reagent can be dispensed into empty plates or spotted onto slides (Erffle *et al.*, 2007). Forward genetic methods use cells cultured in plates and the reagent will be added to the medium. The use of cell arrays has the advantage that a higher number of samples can be studied in a shorter time and at lower cost. After a specified time, the plate will be fixed for end-point assays or subjected to (time-lapse) live cell imaging in HCS platforms.

Assays

In general, any HCS assay falls into one of three main categories:

- visualization of a protein, protein complex, or protein modification,
- visualization of a cellular structure, such as an organelle, and
- visualization of a cellular process (e.g., cell death, protein trafficking, or protease cleavage).

Any of these structures and processes requires visualization and conversion of image data into binary data. As such, the cellular information (e.g., identity of an organelle) requires detection as a fluorescent signal using either attachment to a fluorescent protein or binding of a fluorescent reagent. Once a structure is labeled with a fluorescent tag, the following major phenotypes can be detected:

- a. number of cells, subcellular objects,
- b. subcellular changes in localization,
- c. co-localization of multiple markers,
- d. phenotypes measured by intensity features: mean fluorescence intensity of objects,
- e. phenotypes measured by morphology features: size (area and volume) and shape of objects, and
- f. phenotypes measured by texture features.

Typically, multiple features are measured. Using positive and negative controls that produce the desired phenotype of interest, it is possible to evaluate the feature that gives the best signal to noise ratio and dynamic range. The large amount of data collected by extracting multiple features poses opportunities and challenges in data analysis. Improvements in the

analysis using multivariate profiling of hundreds of features have been made to exploit this strength and address the problem (Loo *et al.*, 2007).

How can these objects and processes be visualized? Generally, one can distinguish between visualization of an endogenous process and the introduction of a reporter construct to extract information of an endogenous process. Direct visualization of an endogenous process has been greatly facilitated by the development of immunofluorescence techniques using antibodies that bind to specific cellular proteins. Such immunostaining can be used to quantify the amount of a protein within a cell or protein–protein binding (e.g., by using two different color-tagged antibodies, split fluorescence proteins or Foerster resonance energy transfer (FRET)-based methods). When analyzing organelles, one could use immunostaining for a marker protein localized to the specific structure. In addition, multiple companies now offer small molecule-based dyes that selectively stain specific organelles for fluorescence microscopy. Most of these techniques require fixed samples and end-point assays. An alternative to immunostaining methods is the use of fluorescent proteins for tagging proteins or monitoring protein and metabolic activities. Following the cloning of green fluorescent protein (GFP) from *Aequorea victoria* in 1994 (Chalfie *et al.*, 1994), there are now a wide range of different colored variants available that cover the entire visible spectrum. Translational fusions of FPs with a protein of interest allow monitoring of protein amount, trafficking/movement, and protein–protein interactions. Furthermore, FP-based sensor proteins have been generated that allow investigations of calcium levels, intracellular pH, mitochondrial redox potential, cyclic AMP levels, protease activity, phosphorylation, and protein–protein interaction using FRET (Tsien, 2009). Similar to proteins, other cellular components can be readily visualized. For instance, DNA or RNA-intercalating agents allow the quantification and identification of DNA/RNA in cells. Similarly, membranes in the cell are usually identifiable by lipophilic fluorescent probes that insert in the membrane of choice, sometimes selectively based on the lipid composition of the membrane type.

From a cell biology perspective, the analysis of organelle biogenesis, trafficking, and function are key elements that have been addressed using HCS. Any of the major organelle types have been analyzed by HCS, including the nucleus in mitosis (Neumann *et al.*, 2010), the Golgi apparatus (Galea and Simpson, 2013), the *trans*-Golgi network (Ferraro *et al.*, 2014; Wendler *et al.*, 2010), endosomes (Collinet *et al.*, 2010), secretory granules (Ferraro *et al.*, 2014), mitochondria (Ivatt *et al.*, 2014), autophagosomes (McKnight *et al.*, 2012; Kriston-Vizi *et al.*, 2010), centrosomes (Dobbelaere *et al.*, 2008), the actin cytoskeleton (Liu *et al.*, 2009), and viruses (Mercer *et al.*, 2012). The main features of interest are the number of organelles (counting individual objects), the intensity of objects (intensity of individual objects), and measuring the size of organelles (measuring length or area). One consideration is that small objects will require a higher resolution platform compared to analysis of larger structures. Another consideration is cell and organelle variability, that has been addressed using population-based measurements (Snijder *et al.*, 2012) or using micro-patterning in microplates (Schauer *et al.*, 2014). Whole cell measurements are usually required when cell

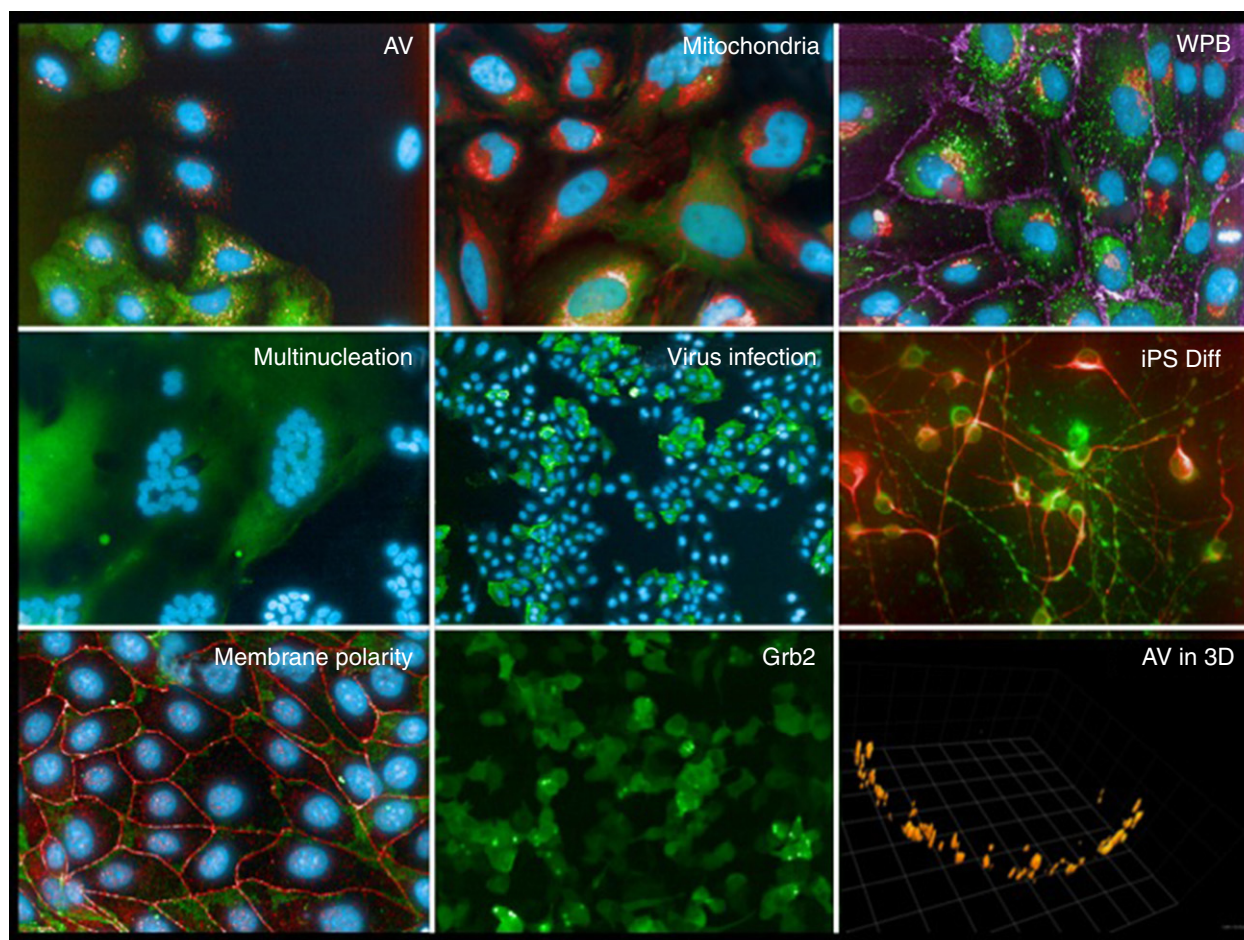


Figure 3 Example images from high-content screens at the Translational Research Resource Centre, London. The top row shows different organelles that have been visualized with fluorescent protein tags (autophagosomes, AV, upper left corner), chemical staining (Mitotracker, mitochondria, upper middle panel), or antibody staining (Weibel–Palade bodies, WPB, upper right corner). Nuclei can be counted to give insight into processes that regulate multinucleation (middle left panel) or number of virus-infected cells (green cells in center panel). HCS can also be used to visualize more complex processes such as monitoring induced iPS cell differentiation into neuronal cells (middle right panel), membrane polarity (lower left panel), or protein localization (in this case the adapter protein Grb2, lower middle panel). When using a confocal microscope and several z stack images are acquired, cellular organelles can be visualized in a 3D context such as shown for autophagosomes in a MCF-10 A human mammary spheroid section in the lower right panel.

morphologies, neuronal phenotypes, total protein levels, or general cell numbers are recorded.

Some example images of HCS experiments are shown in [Figure 3](#). Please note that – depending on the instrument and the choice of fluorescent marker – multiple fluorescent reporters can be visualized in the same experiment.

Instrumentation

The key hardware component is a HCS imaging plate reader. Since the Arrayscan was launched in 1997, multiple other companies have designed alternative HCS imaging platforms and continue to improve on design and functionalities of these instruments. Today, there are about a dozen different types of instruments on the market, with varying hardware components and functionalities ([Table 2](#)). In general, the choice of a HCS system is a decision about speed/throughput

versus image quality. Higher resolution and higher quality images generally take longer time to acquire, although certain features such as parallel image acquisition by multiple cameras can reduce acquisition times significantly.

The main features of HCS imaging systems are:

- light source,
- autofocus,
- cameras,
- optical sectioning, and
- image analysis software.

The basis for all HCS platforms is a powerful light source that can be either laser-based or light-emitting diodes (LEDs). Most systems use between three and five laser-based light sources, while others such as the Arrayscan™ is based on a solid-state LED light engine. The autofocus mechanism can be laser-based or software controlled or a combination of both.

Table 2 Selected examples of high-content imaging platforms

Platform	Light source	Autofocus	Camera	Optics	Reference
Arrayscan TM XTI (Thermo scientific)	LED solid-state light engine	Software-controlled	H1 camera with large format charged coupled device (CCD) (2208 × 2208)	Zeiss AxioObserver	www.thermoscientific.com
InCell Analyzer 6000 (GE Healthcare Life Sciences)	Up to four solid-state lasers	Software controlled or laser-based	Large format sCMOS camera	Epifluorescence based line scanning confocal system	www.gelifesciences.com
ImageXpress [®] Ultra (Molecular Devices)	Up to four solid-state lasers	Laser-based	Large sensor CMOS detector	Laser scanning confocal with software configurable detection pinhole	www.moleculardevices.com
Opera Phenix TM (PerkinElmer)	Up to five solid-state lasers	Laser-based	Up to four large format sCMOS cameras with simultaneous acquisition	Spinning microlens disk, Synchrony TM optics	www.perkinelmer.co.uk
BD Pathway TM 855	Metal halide lamp	Laser-based	CCD camera	Confocal spinning disk	www.bdbiosciences.com
Acumen [®] Cellista (TTP Labtech)	Triple laser excitation		photo-multiplier tubes detector array	Widefield F-theta lens for whole well imaging	ttplabtech.com
Cytation TM 3 (Biotek)	High-power LED	Software-controlled	CCD		www.biotek.com
HC High Content Analysis System (Nikon)	Lasers	Piezo-based	CCD, CMOS	Nikon Eclipse Ti	www.nikoninstruments.com
Scan $\wedge$ R (Olympus)	150 W Xenon or Mercury-Xenon	Software-controlled or laser-based	CCD Hamamatsu	Olympus IX81 and IX83 microscope	www.olympus.co.uk
Cell Voyager CV7000 (Yokogawa)	Solid-state lasers	Software-controlled or laser-based	Up to three large format sCMOS cameras	Microlens-enhanced dual Nipkow disk scanning	www.yokogawa.com

Note: Information taken from vendors websites. Some vendors offer multiple HCS platforms and for simplicity, a selection is shown.

Laser-based mechanisms focus on the surface of the plate type used and is generally quick. Software controlled acquisition of iterative z height sampling can be more time-consuming and focuses on a cellular staining. Initial problems with software-based auto-focussing were delays in focussing when empty wells or wells with detached cells were found, but newer software algorithms circumvent this problem. The choice of camera will largely determine the resolution of the image. Generally, a large format complementary metal-oxide semiconductor (cMOS) or CCD camera is used. An alternative to the use of cameras is instruments based on scanning laser-imaging cytometry that utilize photo-multiplier tubes (PMTs) for detection of the signal. These platforms provide an advantage in terms of speed and do not require image stitching as most camera-based systems do. However, these instruments typically don't have the same level of image resolution. When acquiring images, over-sampling is sometimes done in order to gain a better image quality.

Most HCS instruments are modular with the opportunity to include add-on options such as robotic integration, bright field illumination, injectors or environmental control.

Nowadays, HCS instruments have reached an optical capability that rivals standalone fluorescence microscopes.

High-Content Analysis

Once images are acquired, one of the major challenges is to extract quantitative data by high-content analysis (HCA). One should note that the human eye has evolved to pick up minor differences in patterns and is probably the most reliable tool to identify phenotypes. Since this is no longer possible once a large number of samples are to be inspected and sometimes overlaying phenotypes can mislead the eye. The main challenge is to instruct computer software to extract quantitative data from raw images based on the strength of a given phenotype. Quantifying cellular (Futamura *et al.*, 2012) or subcellular features, HCA consists of two major components: (1) image analysis where the phenotypic features of interest are extracted and subjected to and (2) statistical analysis. Both fields are evolving dynamically: mathematical methods in low-throughput cellular image processing such as active contours

are applied in HCA (Yu *et al.*, 2010), and HCA-specific statistical tools have been developed (Malo *et al.*, 2006, 2010).

Image analysis

A multitude of software options are available for image analysis and feature extraction. For example, commercially available HCA programs such as InCell Analyzer (GE Healthcare), Acapella (PerkinElmer Inc.) with Spotfire (TIBCO Software Inc.), or Metamorph (Molecular Devices, LLC.) bundled with specific HCS platforms, standalone proprietary software as Matlab (Mathworks Inc.) with Image Processing Toolbox as well as open-source software such as CellProfiler (Carpenter *et al.*, 2006), ImageJ (Abramoff *et al.*, 2004), or one of its variants, Fiji (Schindelin *et al.*, 2012). The statistical data analysis functions are either built in a HCA software or developed as a standalone software such as CellProfiler Analyst or R-project (R Core Development Team, 2014) the language, created and extended by mathematicians.

1. Commercial HCA software has the simplest user interface and standardized methods. Their closed source code results in lower flexibility.
2. Free standalone software specialized for image analysis and statistics offers higher flexibility, for the price of a steeper learning curve. ImageJ and R have a large user community and a wide range of application: ImageJ covers image processing applications from astronomy to biology (Schneider *et al.*, 2012), and the R statistical computing language is used from general statistics to specific computational biology and bioinformatics through the Bioconductor project (Gentleman *et al.*, 2004), that includes cell-based HCA packages like cellHTS2 (Boutros *et al.*, 2006) and HTSanalyzeR (Wang *et al.*, 2011).

Commercial HCA software bundled with HCS systems are designed for easy use for the biologist. These products are optimal in a standardized pharmaceutical corporation environment, where millions of samples are screened. In that case the assay can be designed based on the requirements of the HCA algorithm, for example, staining cytoplasm if the software detects the cell by a cytoplasmic marker. This software often uses their proprietary image file format, which limits the input source to the specific plate reader. The comparison between different bundled proprietary HCA software solutions is difficult, since the underlying image processing algorithm is hidden. Customization of the analysis can be a challenge because of the proprietary environment and a newly developed script can only be used by the HCS user group of the given vendor.

On the other hand, free HCA software facilitates algorithm customization and enhances the publication of new scripts, plugins, and macros. For instance, ImageJ has more than 300 macros, over 500 plugins available on the ImageJ website, and the number of R packages is more than 6000. There is a fair chance that customized solutions for a nonstandard problem can be found in the ever-increasing knowledgebase of ImageJ and R mailing lists at Stack Overflow and Nabble. Open-Source HCA allows easy connections with related fields such as computational chemistry (Lepp *et al.*, 2009) with integration tools such as KNIME.

A HCA workflow is typically being set up and tested on some sample images. The aim is to create a pipeline with optimal parameters that can analyze a wide range of images of the whole assay. The whole image stack contains negative controls as well as hits that express potentially stronger phenotype than positive controls, therefore a fully automated HCA image analysis workflow must be robust. In order to facilitate optimization, the number of image processing parameters should be kept at a minimum and must be intuitive for someone, who is not an expert in image analysis.

HCA workflow

A typical HCA image processing workflow consists of noise reduction, segmentation, and feature extraction steps (Figure 4).

Noise reduction

The noise reduction serves to improve the signal on an image by reducing the inherent imperfections of the imaging system such as stuck pixels in the camera, or unspecific staining. Computationally cheap implementations are optimal in a HCA environment, like the multi-threaded median filter, or the rolling ball algorithm in ImageJ.

Segmentation

Segmentation is a binary classification that labels each pixel as either foreground or background. The segmentation can be (1) global, when a single threshold is applied to all pixels in an image or (2) local or adaptive, when separate thresholds are used in a part of an image. The parameter number of a segmentation algorithm must be kept at a minimum in order to ensure the usability of a HCA workflow. Otsu and Isodata are particularly fast HCA segmentation algorithms that calculate their single parameter directly from the image pixel information. Nonlinear multi-parametric supervised machine learning-based methods can further improve the accuracy of the segmentation step for the price of more labor-intensive training process, and its HCA use is even more promising at cell classification level (Horvath *et al.*, 2011).

Segmented pixels are grouped into objects. Higher than single pixel level, objects consist of foreground pixels based on neighbor connectivity. ImageJ uses 8 pixel connectivity in 2D images where neighboring pixels form an object by connecting horizontally, vertically, or diagonally, touching each other's edges or corners. In the future 3D HCA is expected to use voxels instead of pixels and 26-neighborhood connectivity.

Feature extraction

Objects are labeled and quantified in the feature extraction step. Data is getting more complex and compressed at each processing step. The acquired information encoded into several gigabytes of 12 bit depth gray level pixels is first converted to 2-bit depth binary segmentation, then a table of measurements, where each line represents measurements of a cellular or subcellular object. Features are classified as intensity-based, morphology and texture-based features. Intensity-based features describe the object by the statistics of its pixel intensities: mean, median, mode, minimum, maximum, standard deviation, and total pixel intensity. Morphology-based features are shape descriptors: area, perimeter, Feret's diameter (maximum

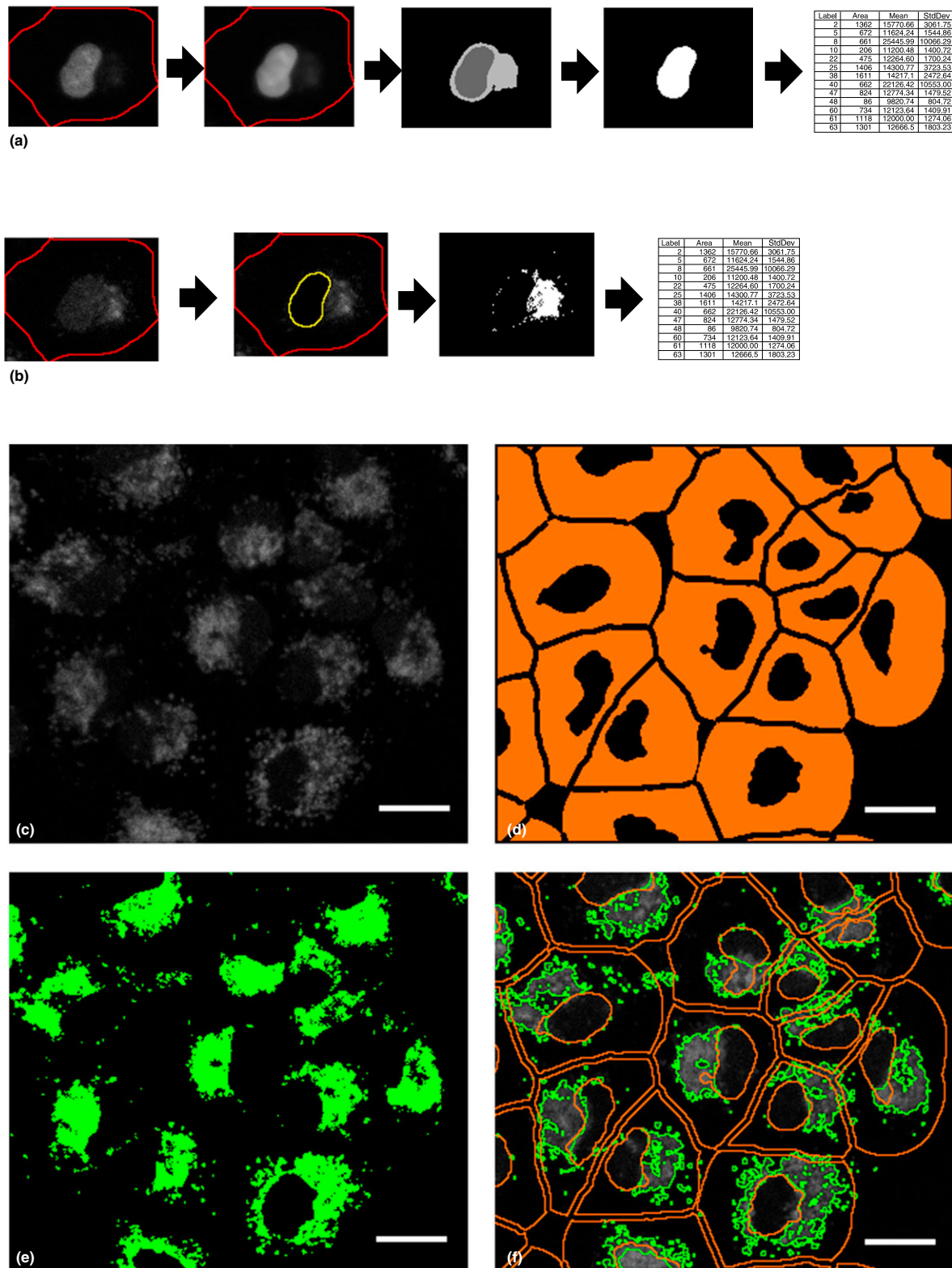


Figure 4 Image processing workflow of a multichannel image of A549 cells imaged by ImageXpress Ultra, with 20x lens, nuclei stained with Hoechst 33342, autophagosomes stained with monodansylcadaverine. The original nuclear channel of a sample cell was smoothed with a median filter, then segmented with *k*-means clustering algorithm, the nuclear object was selected and quantified: (a) from left to right. The original autophagosomal channel of the (a) sample cell is shown, where the bleed through from nuclear channel was removed by using the nuclear mask, the autophagosomes in the remained area of the cell was segmented by IsoData algorithm and the features were extracted: (b) from left to right. A sample image of the autophagosomal channel (c), segmented by region growing with commercial HCA software (d), and with the (a–b) workflow (e). The overlay image shows the difference between (d–e) segmentations (f). All scale bars are 20 μ m. StdDev, standard deviation.

calliper), circularity, and aspect ratio. Texture features (reviewed at [Huang and Murphy, 2004](#)) statistically describe pixel distributions and can be quite numerous: the major part of the 2600 image features that the WND-CHARM image classifier ([Orlov et al., 2008](#)) computes are texture features. However, most of those texture features are redundant, which requires an additional feature selection step. The computationally expensive algorithms make those software solutions less feasible in current HCA pipelines but can become possible in the future.

Statistical analysis

Although being a lengthy text file, the feature table can be a 1000-fold smaller representation of the initial high-content image dataset. Focused on statistics and developed by mathematicians and statisticians, the free R software has the scalability and flexibility to analyze large HCS data tables. A typical statistical HCA workflow in R starts with further data compression and finishes with an ordered list of the samples that specifies a hit list. The multidimensional feature space of many objects in several images per well should be averaged to one single value per well with a simple R script, and format as an input data for the cellHTS2 package ([Boutros et al., 2006](#)), the R Bioconductor project's HTS module.

The aim of a cellHTS2 workflow is to put the samples into order based on the strength of the phenotype of interest and after choosing thresholds enable the researcher to generate a hit list.

The main vignette of cellHTS2 (complete version, available online) describes a full data analysis workflow with sample R code, here we overview the process. Scoring and optionally, annotation follow data normalization. Per plate median, control- and non-control based algorithms such as B-score ([Malo et al., 2006](#)) were implemented for normalization and standard scoring methods like Z-score are available as well. As a result, a report document is generated in web browser-readable HyperText Markup Language (HTML) format, where screen quality control measurements are computed along with an ordered sample list based on the scores.

Data storage

Through HCS, a large amount of data is collected that poses challenges in terms of data handling and data storage. Generally, multicore CPUs and several terabytes (often petabyte) for data storage are needed. Multiple databases exist where HCS data and images can be deposited and mined by the research community, including OMERO ([Swedlow et al., 2003](#)), GenomeRNAi ([Schmidt et al., 2013](#)), RNAiDB ([Gunsalus et al., 2004](#)), FLIGHT ([Sims et al., 2006](#)), and the JCB DataViewer ([Hill, 2008](#)).

3D imaging presents novel opportunities and challenges for image analysis and data storage. While feature extraction of phenotypes in classical HCS has been done on a parameter per image or parameter per cell basis, there are now more modalities that can be interrogated, taking into account the 3D distribution of features. For example, 3D volumetric information can be analyzed in fruit fly embryos ([Kriston-Vizi et al., 2011](#)) and parameters can be defined as feature per 3D structure or feature relative to location within a 3D structure per cell per field. As such, there are multiple levels of complexity, opening up the possibility to collect thousands

features ([Tang et al., 2011](#)). This notion underscores a current shift in a switch from HCS to 'high complexity screening' (HCxS).

High-Content or Low-Content?

An analysis in 2014 revealed that a large number of published high-content screens only present data from one or two features ([Singh et al., 2014](#)), suggesting that most HCS is in fact low-content. One possible explanation is the lack of appropriate tools for data analysis. For instance, commercial bundled software usually restricts users to analysis of few predefined parameters and are used mostly by novice users that prefer simple methods over high-content. As such, there may be a gap between novice users and HCS experts who have better knowledge of the available bioinformatics and data analysis tools. Significant progress has been made to enable multi-parametric analysis with free available software tools that should help bridge this gap in the future. For example, machine learning and multi-parametric classification tools such as CellProfiler Analyst ([Jones et al., 2009](#)), CellClassifier ([Ramo et al., 2009](#)), Enhanced CellClassifier ([Misselwitz et al., 2010](#)), or Advanced CellClassifier ([Horvath et al., 2011](#)) enable the selection of cells that display a certain phenotype of interest and facilitate automated feature extraction in large image data sets. In addition, multidimensional profiling can help to cluster multiple feature distributions in order to classify hits based on multi-parametric readouts ([Singh et al., 2014](#)). Further developments in this area will enable the realization of truly HCS applications even for newcomers to this exciting technology.

Conclusion

It can be expected that HCS will continue to contribute substantially to our understanding of cell biology in the future. The capability to miniaturize experimental setup, time and cost through this technique will enable discoveries of cell biology to be made on a much larger scale. One of the bottlenecks is that more advanced imaging techniques such as super-resolution microscopy are not applicable to automated image acquisition yet, partly due to limitations in the hardware and time constraints that make it unreasonable. Current developments in this area include the integration of microfluidics, allowing a high degree of miniaturization and a higher level of spatiotemporal and environmental control over live cell imaging ([Cheong et al., 2009](#); [Upadhyaya and Selvaganapathy, 2010](#)). Flow cytometers such as the ImageStream™ are available now that combine flow cytometry with high-content imaging at reasonably high resolution, thus allowing rapid single-cell HCS. Recently, a high-throughput imaging microscope for high-speed imaging of entire live embryos based on light-sheet microscopy has been developed ([Tomer et al., 2012](#)) that is a milestone step toward realizing even more complex HCS. Furthermore, a HCS system based on fluorescence lifetime imaging (FLIM) for improved measurement of FRET has been described for improved HCS of spatiotemporal dynamics *in vivo* ([Kumar et al., 2011](#)).

Overall, HCS has enabled the analysis of cells and organelles in a much more time- and cost-effective manner when compared

to individual probing of single gene perturbations. Current instrumentation has reached a quality that rivals conventional standalone instruments and will continue to make essential contributions to our understanding of general cell biology.

See also: Horizontal Integration: Networks: Understanding of 'Networks' *In Vitro* and/or *In Vivo*

References

- Abramoff, M.D., Magalhaes, P.J., Ram, S.J., 2004. Image processing with ImageJ. *Biophotonics International* 11, 36–43.
- Boutros, M., Bras, L.P., Huber, W., 2006. Analysis of cell-based RNAi screens. *Genome Biology* 7, R66.
- Boutros, M., Kiger, A.A., Armknecht, S., *et al.*, 2004. Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* 303, 832–835.
- Carpenter, A.E., Jones, T.R., Lamprecht, M.R., *et al.*, 2006. CellProfiler: Image analysis software for identifying and quantifying cell phenotypes. *Genome Biology* 7, R100.
- Celli, J.P., Rizvi, I., Blanden, A.R., *et al.*, 2014. An imaging-based platform for high-content, quantitative evaluation of therapeutic response in 3D tumor models. *Scientific Reports* 4, 3751.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., Prasher, D.C., 1994. Green fluorescent protein as a marker for gene expression. *Science* 263, 802–805.
- Cheong, R., Wang, C.J., Levchenko, A., 2009. Using a microfluidic device for high-content analysis of cell signaling. *Science Signaling* 2, pl2.
- Collinet, C., Stoler, M., Bradshaw, C.R., *et al.*, 2010. Systems survey of endocytosis by multiparametric image analysis. *Nature* 464, 243–249.
- Desbordes, S.C., Studer, L., 2013. Adapting human pluripotent stem cells to high-throughput and high-content screening. *Nature Protocols* 8, 111–130.
- Dobbelaere, J., Josue, F., Suijkerbuijk, S., *et al.*, 2008. A genome-wide RNAi screen to dissect centriole duplication and centrosome maturation in *Drosophila*. *PLoS Biology* 6, e224.
- Eggert, U.S., Kiger, A.A., Richter, C., *et al.*, 2004. Parallel chemical genetic and genome-wide RNAi screens identify cytokinesis inhibitors and targets. *PLoS Biology* 2, e379.
- Erfle, H., Neumann, B., Liebel, U., *et al.*, 2007. Reverse transfection on cell arrays for high content screening microscopy. *Nature Protocols* 2, 392–399.
- Ferraro, F., Kriston-Vizi, J., Metcalf, D.J., *et al.*, 2014. A two-tier Golgi-based control of organelle size underpins the functional plasticity of endothelial cells. *Developmental Cell* 29, 292–304.
- Földes, G., Matsa, E., Kriston-Vizi, J., *et al.*, 2014. Aberrant alpha-adrenergic hypertrophic response in cardiomyocytes from human induced pluripotent cells. *Stem Cell Reports* 3 (5), 905–914.
- Futamura, Y., Kawatani, M., Kazami, S., *et al.*, 2012. Morphobase, an encyclopedic cell morphology database, and its use for drug target identification. *Chemistry & Biology* 19, 1620–1630.
- Galea, G., Simpson, J.C., 2013. High-content screening and analysis of the Golgi complex. *Methods in Cell Biology* 118, 281–295.
- Gentleman, R.C., Carey, V.J., Bates, D.M., *et al.*, 2004. Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biology* 5, R80.
- Gunsalus, K.C., Yueh, W.C., Macmenamin, P., Piano, F., 2004. RNAiDB and PhenoBlast: Web tools for genome-wide phenotypic mapping projects. *Nucleic Acids Research* 32, D406–D410.
- Heintze, J., Luft, C., Ketteler, R., 2013. A CRISPR case for high-throughput silencing. *Frontiers in Genetics* 4, 193.
- Hill, E., 2008. Announcing the JCB DataViewer, a browser-based application for viewing original image files. *Journal of Cell Biology* 183, 969–970.
- Horvath, P., Wild, T., Kutay, U., Csucs, G., 2011. Machine learning improves the precision and robustness of high-content screens: Using nonlinear multiparametric methods to analyze screening results. *Journal of Biomolecular Screening* 16, 1059–1067.
- Huang, K., Murphy, R.F., 2004. From quantitative microscopy to automated image understanding. *Journal of Biomedical Optics* 9, 893–912.
- Ivatt, R.M., Sanchez-Martinez, A., Godena, V.K., *et al.*, 2014. Genome-wide RNAi screen identifies the Parkinson disease GWAS risk locus SREBF1 as a regulator of mitophagy. *Proceedings of the National Academy of Sciences of the United States of America* 111, 8494–8499.
- Jones, T.R., Carpenter, A.E., Lamprecht, M.R., *et al.*, 2009. Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning. *Proceedings of the National Academy of Sciences of the United States of America* 106, 1826–1831.
- Kiger, A.A., Baum, B., Jones, S., *et al.*, 2003. A functional genomic analysis of cell morphology using RNA interference. *Journal of Biology* 2, 27.
- Kriston-Vizi, J., Lim, C.A., Condron, P., *et al.*, 2010. An automated high-content screening image analysis pipeline for the identification of selective autophagic inducers in human cancer cell lines. *Journal of Biomolecular Screening* 15, 869–881.
- Kriston-Vizi, J., Thong, N.W., Poh, C.L., *et al.*, 2011. Gebiss: An ImageJ plugin for the specification of ground truth and the performance evaluation of 3D segmentation algorithms. *BMC Bioinformatics* 12, 232.
- Kumar, S., Alibhai, D., Margineanu, A., *et al.*, 2011. FLIM FRET technology for drug discovery: Automated multiwell-plate high-content analysis, multiplexed readouts and application in situ. *ChemPhysChem* 12, 609–626.
- Lepp, Z., Huang, C., Okada, T., 2009. Finding key members in compound libraries by analyzing networks of molecules assembled by structural similarity. *Journal of Chemical Information and Modeling* 49, 2429–2443.
- Liu, T., Sims, D., Baum, B., 2009. Parallel RNAi screens across different cell lines identify generic and cell type-specific regulators of actin organization and cell morphology. *Genome Biology* 10, R26.
- Loo, L.H., Wu, L.F., Altschuler, S.J., 2007. Image-based multivariate profiling of drug responses from single cells. *Nature Methods* 4, 445–453.
- Malo, N., Hanley, J.A., Carlile, G., *et al.*, 2010. Experimental design and statistical methods for improved hit detection in high-throughput screening. *Journal of Biomolecular Screening* 15, 990–1000.
- Malo, N., Hanley, J.A., Cerquozzi, S., Pelletier, J., Nadon, R., 2006. Statistical practice in high-throughput screening data analysis. *Nature Biotechnology* 24, 167–175.
- McKnight, N.C., Jefferies, H.B., Alemu, E.A., *et al.*, 2012. Genome-wide siRNA screen reveals amino acid starvation-induced autophagy requires SCOC and WAC. *EMBO Journal* 31, 1931–1946.
- Mercer, J., Snijder, B., Sacher, R., *et al.*, 2012. RNAi screening reveals proteasome- and Cullin3-dependent stages in vaccinia virus infection. *Cell Reports* 2, 1036–1047.
- Misselwitz, B., Strittmatter, G., Periaswamy, B., *et al.*, 2010. Enhanced CellClassifier: A multi-class classification tool for microscopy images. *BMC Bioinformatics* 11, 30.
- Miyazaki, I., Simizu, S., Osada, H., 2010. The application of the chemical array for biological study. *Methods in Molecular Biology* 669, 95–107.
- Moffat, J., Grueneberg, D.A., Yang, X., *et al.*, 2006. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* 124, 1283–1298.
- Neumann, B., Walter, T., Heriche, J.K., *et al.*, 2010. Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature* 464, 721–727.
- Orlov, N., Shamir, L., Macura, T., *et al.*, 2008. WND-CHARM: Multi-purpose image classification using compound image transforms. *Pattern Recognition Letters* 29, 1684–1693.
- Pardo-Martin, C., Chang, T.Y., Koo, B.K., *et al.*, 2010. High-throughput *in vivo* vertebrate screening. *Nature Methods* 7, 634–636.
- Ramo, P., Sacher, R., Snijder, B., Begemann, B., Pelkmans, L., 2009. CellClassifier: supervised learning of cellular phenotypes. *Bioinformatics* 25, 3028–3030.
- R Core Development Team, 2014. R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing.
- Schauer, K., Grossier, J.P., Duong, T., *et al.*, 2014. A novel organelle map framework for high-content cell morphology analysis in high throughput. *Journal of Biomolecular Screening* 19, 317–324.
- Schindelin, J., Arganda-Carreras, I., Frise, E., *et al.*, 2012. Fiji: An open-source platform for biological-image analysis. *Nature Methods* 9, 676–682.
- Schmidt, E.E., Pelz, O., Buhlmann, S., *et al.*, 2013. GenomeRNAi: A database for cell-based and *in vivo* RNAi phenotypes, 2013 update. *Nucleic Acids Research* 41, D1021–D1026.
- Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 Years of image analysis. *Nature Methods* 9, 671–675.
- Sherman, S.P., Alva, J.A., Thakore-Shah, K., Pyle, A.D., 2011. Human pluripotent stem cells: The development of high-content screening strategies. *Methods in Molecular Biology* 767, 283–295.
- Silva, J.M., Mizuno, H., Brady, A., Lucito, R., Hannon, G.J., 2004. RNA interference microarrays: High-throughput loss-of-function genetics in mammalian cells.

- Proceedings of the National Academy of Sciences of the United States of America 101, 6548–6552.
- Sims, D., Bursteinas, B., Gao, Q., Zvelebil, M., Baum, B., 2006. FLIGHT: Database and tools for the integration and cross-correlation of large-scale RNAi phenotypic datasets. *Nucleic Acids Research* 34, D479–D483.
- Singh, S., Carpenter, A.E., Genovesio, A., 2014. Increasing the content of high-content screening: An overview. *Journal of Biomolecular Screening* 19, 640–650.
- Snijder, B., Sacher, R., Ramo, P., *et al.*, 2012. Single-cell analysis of population context advances RNAi screening at multiple levels. *Molecular Systems Biology* 8, 579.
- Swedlow, J.R., Goldberg, I., Brauner, E., Sorger, P.K., 2003. Informatics and quantitative analysis in biological imaging. *Science* 300, 100–102.
- Tang, J., Enderling, H., Becker-Weimann, S., *et al.*, 2011. Phenotypic transition maps of 3D breast acini obtained by imaging-guided agent-based modeling. *Integrative Biology* 3, 408–421.
- Taylor, D.L., 2007. Past, present, and future of high content screening and the field of cellomics. *Methods in Molecular Biology* 356, 3–18.
- Taylor, D.L., 2010. A personal perspective on high-content screening (HCS): From the beginning. *Journal of Biomolecular Screening* 15, 720–725.
- Tomer, R., Khairy, K., Amat, F., Keller, P.J., 2012. Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. *Nature Methods* 9, 755–763.
- Tsien, R.Y., 2009. Indicators based on fluorescence resonance energy transfer (FRET). *Cold Spring Harbor Protocols*. doi:10.1101/pdb.top57.
- Upadhyaya, S., Selvaganapathy, P.R., 2010. Microfluidic devices for cell based high throughput screening. *Lab on a Chip* 10, 341–348.
- Wang, X., Terfve, C., Rose, J.C., Markowitz, F., 2011. HTSanalyzeR: An R/Bioconductor package for integrated network analysis of high-throughput screens. *Bioinformatics* 27, 879–880.
- Wendler, F., Gillingham, A.K., Sinka, R., *et al.*, 2010. A genome-wide RNA interference screen identifies two novel components of the metazoan secretory pathway. *EMBO Journal* 29, 304–314.
- Yu, W.M., Lee, H.K., Hariharan, S., Bu, W.Y., Ahmed, S., 2010. Evolving generalized voronoi diagrams for accurate cellular image segmentation. *Cytometry Part A* 77A, 379–386.
- Crawley, M.J., 2012. *The R Book*, second ed. Wiley (Online Library).
- Haney, S.E. (Ed.), 2008. *High-Content Screening: Science, Techniques and Applications*. Hoboken, NJ: Wiley-Interscience.
- Taylor, L.D. (Ed.), 2006. *High-Content Screening (Methods in Molecular Biology)*. Totowa, NJ: Humana Press.

Relevant Websites

<http://www.ucl.ac.uk/lmcb/bioinformatics-image-core-bionic>
Bioinformatics Image Core facility, LMCB, Medical Research Council, University College London.

<http://cran.r-project.org/>
Comprehensive R Archive Network.

http://fiji.sc/Trainable_Weka_Segmentation
Fiji Is Just ImageJ, Trainable Weka Segmentation plugin.

<http://flight.icr.ac.uk/>
FLIGHT RNAi database.

<http://www.genomernai.org/>
GenomeRNAi.

<http://rsb.info.nih.gov/ij/>
ImageJ.

<http://jcb-dataviewer.rupress.org/>
Journal of Cell Biology Dataviewer.

tech.knime.org
KNIME.

<http://www.nabble.com/>
Nabble.

<http://www.slas.org/resources/information/academic-screening-facilities/>
SLAS.

<http://stackoverflow.com/>
Stack Overflow.

<https://www.ucl.ac.uk/lmcb/about-translational-research-resource-center/>
Translational Research Resource Center, LMCB, Medical Research Council, University College London.

Further Reading

- Burger, W., Burge, M.J., 2011. *Digital Image Processing: An Algorithmic Introduction Using JAVA*. Berlin, Heidelberg, New York, Hongkong, London, Mailand, Paris, Tokyo: Springer.

Introduction

The exciting technological advances allowing high-throughput biological research have promised to unlock the secrets of our genome. However, the initial excitement over new technologies as usual quickly turned into challenges of how to efficiently integrate, analyse, interpret the data produced, and enable knowledge discovery (Bellazzi, 2014). Since a holistic approach is the only way to understand systems-level phenomena, we need to integrate many heterogeneous sources at multiple layers of detail and different levels of completeness (Lauffenburger, 2012). For example, to fully understand the cellular function of a protein complex, information must be collected and combined for the expression, posttranslational modification, interactions, and localization of all its subunits. The approaches toward addressing such challenges have shaped a whole new scientific discipline called 'Integrative Systems Biology.'

Many databases, resources, and tools have been built to address the big challenge of providing the biological research community with easy ways to extract information from high-dimensional biological data. The various resources report information on multiple organisms and layers of the cellular regulation (e.g., genomics, proteomics, interactomics, etc.) derived from a variety of experimental methods (Berger *et al.*, 2013; Joyce and Palsson, 2006; Lee *et al.*, 2012; Rodríguez-Ulloa and Rodríguez, 2008).

However, the complexity of the biological data as well as the way these data are stored and processed in the different resources, pose significant obstacles in the current integration efforts. In this article we review the data integration challenges and solutions, and illustrate them with an example protein complex (Figure 1).

Challenges

Inconsistent Nomenclature

Consistent nomenclature between biological entities like genes and proteins is of major importance for any type of data integration. So far, different public databases use several different competing standards for assigning names to the different types of biological entities, which causes ambiguities in biological nomenclature (Fundel and Zimmer, 2006). This results in two problems: the same gene or protein will often be provided by different resources under different names or identifiers (e.g., the tumor necrosis factor (TNF) is also known as TNF-alpha, TNF-a, cachectin, TNFA_HUMAN, P01375, etc.) and the same gene or protein names are assigned to homologs from different organisms. Although most databases list synonyms or cross references to other databases, these lists are rarely complete. Another technical complication is the use of Greek letters or special characters in entity names (e.g., TNF α),

which is no problem for human readers, but can pose problems for a computer, in particular when extracting information from the biomedical literature. In our example, the same subunit of the complex may be referred to by different identifiers in different interaction databases.

Lack of Standardization

The lack of standardized formats for the representation of the data provided, in particular by new technologies, poses a significant practical obstacle to data integration. The technology and software used for producing and analyzing each individual type of data is often tightly linked to the form in which the data are presented. Because technologies and analysis tools change all the time, data standardization efforts are continuously trying to catch up. Therefore, data from new technologies have often not been deposited in standardized repositories and instead must be extracted from the primary articles, sometimes from the images and graphs within (El-Khatib *et al.*, 2000). Even when data have been aggregated in databases, each database tends to adopt its own standard and annotation practice. For example, the protein interaction databases IntAct and BioGRID report 89 and 16 interaction partners of NFKB2, respectively, based on the same study (Bouwmeester *et al.*, 2004).

Incomplete information on the data provided by the different resources, is another very common phenomenon, with a significant impact on the data integration efficiency and reliability. The same phenomenon is often observed at the supplementary information following published articles. Since there is no general agreement on the data depth that should be provided by each database, it is upon the publication authors or the data resource developers to select the data they provide. Missing information very often concerns details on the experimental assays or references to the source literature. For the protein complex example, important information could be missing such as the experimental assay used or even the source publication. This was a problem in the past but has been addressed by the Proteomics Standards Initiative – Molecular Interactions (PSI-MI) workgroup (Kaiser, 2002).

Limitations of Manual Curation

Even if manually curated data have been for years assumed to be of high-confidence, manual-curation approaches suffer from considerable limitations. Clearly, sometimes the reason is human errors introduced when manually extracting information from the literature or too uncritical trust in the claims made in the peer-reviewed literature. However, maybe the biggest challenge is simply to keep up with the ever-growing literature (Turinsky *et al.*, 2010). It is not feasible for curators to go through every published articles and manually annotate every type of evidence reported, and many databases thus focus on specific diseases, species, or experimental methods

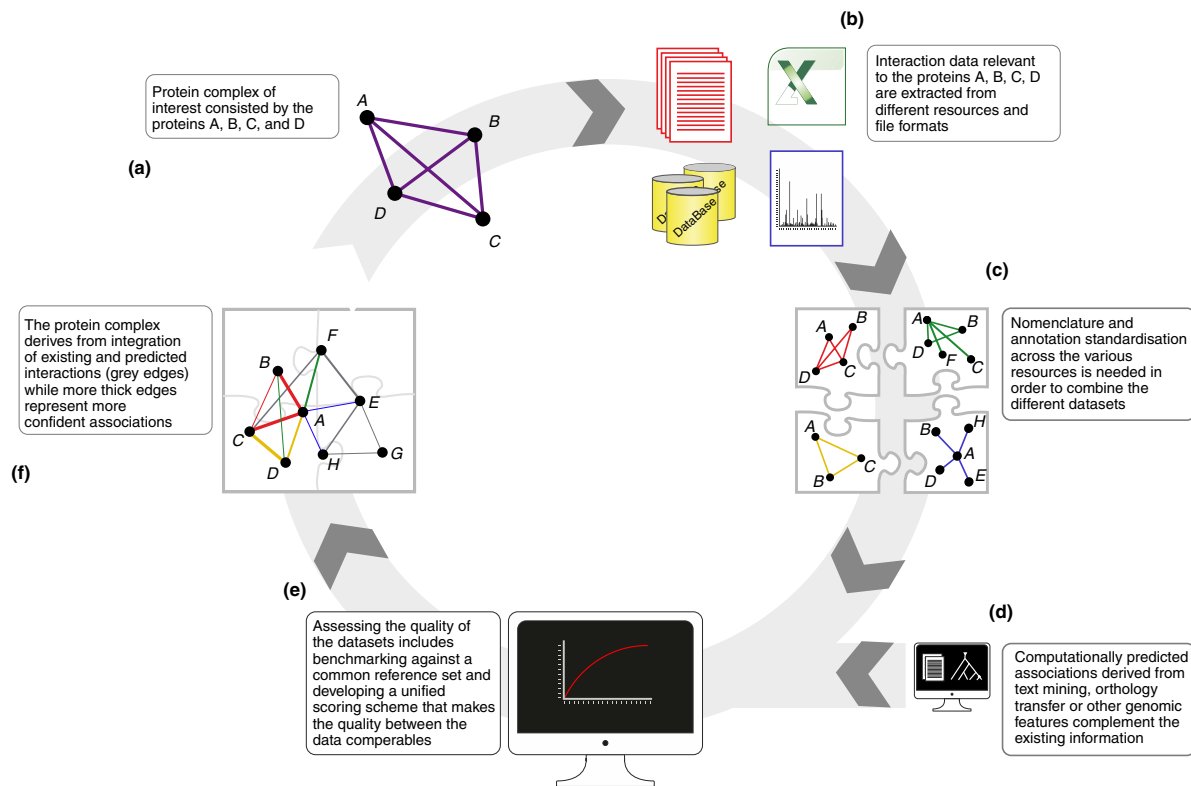


Figure 1 Illustration of a common data integration workflow. (a) To illustrate the typical steps in data integration, we use a protein complex as an example. (b) To start, data from multiple sources are collected. Data can be derived from databases, manual literature curation, supplementary information, or article figures. (c) To combine the diverse information, the nomenclature and representation of the data by the various resources is standardized. For that reason, naming standards, controlled vocabularies, and ontologies are often used. The results concerning the protein complex from the different resources can vary (the colors of the edges are respective to the resources); specific complex subunits may be missing or additional subunits may be present, depending on biological and technological reasons. (d) To complement the existing data, text-mining and prediction methods can provide additional information on the complex subunits, which was not captured by the existing literature. (e) The agreement of each dataset with a common reference set (gold standard) is quantified and a confidence-scoring scheme is developed enabling direct comparison of the quality of each dataset. (f) The final complex contains interactions derived from experimental and predicted information (gray edges), while the thickness of the edges is respective to the quality of the information supporting each link, as resulted from the quality assessment step.

(Cusick *et al.*, 2009). In addition, various resources follow different curation guidelines, and therefore capture different information from the same publications. In the example complex, not all subunits may be supported by equally strong experimental evidence; different interaction databases may thus not agree on which interactions have been shown by the same experiment.

Incomplete Coverage

Despite the vast amount of data available, there are still very many genes and proteins about which we know surprisingly little. There are several reasons for this. First, a large percentage of genes are expressed only under specific conditions (e.g., in response to stress or during a specific developmental stage). The consequence is that so-called genome-wide studies in fact cover only the part of the genome that is expressed under the condition studied. Second, the biological systems are highly dynamic and their responses to internal and external stimuli can be transient and thus difficult to measure (Ideker and Krogan, 2012). Third, most experimental methods do not work equally well for all types of proteins. The available data

may thus be biased toward, for example, the most abundant proteins, proteins with specific physicochemical properties or proteins from certain cellular compartments (Jensen and Bork, 2008). In the protein complex example, the complex can be missed completely in a screen performed under conditions where one or more subunits are not present, or weakly attached subunits may not be found. Also, many assays, such as yeast two-hybrid, tend to miss membrane-associated complexes.

Data Quality and Heterogeneity

To gain systems-level insight into the regulation of biological processes, diverse types of data are required. However, the heterogeneous nature of the data makes them challenging to manage, analyse, and integrate, as they pertain to fundamentally different but interconnected biological questions, processes, and levels of regulation (e.g., genome, transcriptome, proteome, or interactome). For example, to study a protein complex in the context of a disease, the data collected may in addition to protein interactions cover single nucleotide polymorphisms (SNPs), copy number variations (CNVs), gene

expression signatures, protein abundance measurements, and posttranslational modifications, all of which need to be somehow integrated.

Even for a single type of data, different technological platforms have typically been developed (e.g., microarrays and RNA sequencing), which should be combined to get the most comprehensive view. However, the latter can be very challenging since the reliability of the data generated by the different technologies may vary considerably and be subject to different biases (von Mering *et al.*, 2002). Indeed, the quality can even vary between data points in a single dataset made using a single technology. In our example, the protein complex may have been studied by multiple different assays, such as yeast two-hybrid (Y2H) and affinity purification followed by mass spectrometry (AP-MS), and only a subset of the interactions found by either technology might be true.

Data Availability

A trivial obstacle to data integration is that you cannot integrate data you do not have. A common problem is that the data are not collected in databases or repositories; this is particularly often the case for new technologies for which standards have also not yet been developed. The only solution to that is unfortunately to extract that data from the backlog of articles and ensure that data are deposited in the future (Alsheikh-Ali *et al.*, 2011).

Even if a database exists, the data may not be available for integration. Some databases have restrictive licenses that forbid bulk download, redistribution, or commercial use (Freedman and Inglesse, 2014). Two major reasons for using such license are: to prevent competition from other research groups, to be able to commercialize the resource. Whatever the reason, licenses may make integration legally impossible or, more commonly, make it impossible to make integrated results available to others.

Another legal hurdle is the strict privacy laws that guard the access to genetics and healthcare data about individuals (Lunshof *et al.*, 2008). Although there are good reasons for these laws, one cannot deny that it makes integration of such data very complicated from a legal aspect.

Identification of Suitable Resources and Data Interpretation

Even when sufficient data are available, discovering the most suitable out of the hundreds of publicly available resources can be a daunting task. It is often assumed by the developers that the users already understand the type of data stored in the database and the methods used to analyze them. In addition, analyzing high-throughput data are commonly analyzed using sophisticated computational methods, without the user understanding the algorithms and their limitations. Consequently, the methods may be inappropriate for the input data type or the applications used may not respond the questions asked by the user.

Solutions

Nomenclature Systematization of Biological Entities

A very common strategy toward the standardization of heterogeneous data is the creation of naming standards

(Gelernter and Lesk, 2011; Philippi and Köhler, 2006). Nomenclature standardization efforts driven by the research community as the ones by the Human Genome Organisation (HUGO), Gene Nomenclature Committee (HGNC), or the various model organism databases ensure that each gene is assigned a unique approved name (Eppig *et al.*, 2005; De la Cruz *et al.*, 2005; Gray *et al.*, 2013). Similarly, many other efforts aim to standardize other biological entities and concepts with controlled vocabularies and ontologies (Blake and Bult, 2006).

A controlled vocabulary is defined by the collection of terms used to describe a specific scientific domain. Different nomenclature and annotation of the same entities obtained by the individual sources is linked to a predefined set of terms. For example, the UniProtKB database assigns 'keywords' to the different protein entries (Apweiler *et al.*, 2004). The user can make queries with keywords and find easily and fast a protein or a subset of proteins, based on specific biological features (e.g., function, structure, disease, domains, etc.). It is a very popular and highly emerging strategy, mainly used to resolve semantics ambiguities caused by synonyms, homonyms, etc., facilitating thus significantly the data exchange (Kelso *et al.*, 2003).

Ontologies, on the other hand, include controlled vocabularies, describing through the relationships between the different entities. While the definition of the term 'ontology' has its roots in philosophy, meaning the study of the 'being' or the 'existence,' in modern biology, ontologies are defined by the use of entity graphs, built to rather structure the knowledge and describe possible relationships between the represented entities in a hierarchical manner (Schuurman and Leszczynski, 2008; Blake and Bult, 2006). They are widely used in data integration approaches, as they can resolve natural language inconsistencies and enable interoperability between different resources, where unspecific description and annotation terms are used (Blake and Bult, 2006). Ontologies are designed with the use of specific languages, which are understandable by both humans and processing software. Maybe the most representative example of ontologies used in biology is the Gene Ontology (GO) project (The Gene Ontology Consortium, 2008). GO is a structured vocabulary of terms describing three aspects of gene functionality, namely biological process (BP), molecular function (MF), and cellular component (CC). GO is currently under the umbrella of the open biomedical ontologies (OBO) foundry between another 60 biology-related ontologies. This collaborative effort aims to coordinate ontologies by establishing common principles for their development, syntax, and interoperability features (Smith *et al.*, 2007).

Data Structure Standardization

As discussed above, the data standardization efforts aim at facilitating the data exchange and reusability. This is the responsibility both of the individual researchers creating datasets and of the developers who provide public databases to disseminate research results in as clear and usable a manner as possible. It is important that the data representation makes the data understandable and interpretable, which in addition to the data itself requires additional information about the data,

called metadata (Vempati *et al.*, 2014). Typical examples of metadata are when the data were generated, by whom, and with which technology. From an analysis standpoint more metadata is always better; however, since entering such information can quickly get very time consuming, so-called minimum information standards have been designed with the aim to collect the most important metadata, but not more. The first such standard was the minimum information about a microarray experiment (MIAME) (Brazma *et al.*, 2001). This has since been followed by minimum information standards for proteomics experiments (Orchard and Hermjakob, 2008), interaction studies (Orchard *et al.*, 2007), RNAi experiments (Shamu *et al.*, 2012), bioactive entities (Orchard *et al.*, 2011), flow cytometry (Lee *et al.*, 2008), and many more.

Text-Mining Approaches

As a response to the exponential growth of the scientific literature, nowadays, automatic text-mining is increasingly being used by the community in biomedical research to deal with the literature being too large for a human to curate (Rebholz-Schuhmann *et al.*, 2012; Zhu *et al.*, 2013). The two key steps in a typical biomedical text-mining pipeline are named entity recognition and information extraction. The first step involves recognizing the terms in the text that refer to entities or concepts of interest, for example, genes and diseases, and normalizing them to standardized names or identifiers. The second step is to extract relationships between them. An easy approach is to simply link entities to each other if they co-occur in the text. Despite its simplicity, this method can be surprisingly powerful when taking into account the number of times two entities co-occur. A much more advanced approach is natural language processing, which takes into account the syntax (structure) and the semantics (meaning of the words) of the text (Rodríguez-Esteban, 2009).

Computational Prediction Methods

The inherent limitations of current experimental techniques give a need for prediction methods to complement them. This is particularly needed because so many genomes are now being sequenced that experimental characterization of all the genes within them is unrealistic.

The most common approach, used by the individual researchers on a small scale and by various analysis tools in a systematic manner, is to make predictions by transferring functional annotation from homologous genes in different organisms (Dolinski and Botstein, 2007). In this context, it is important to distinguish between two classes of homologs: orthologs, which started diverging after a speciation event, and paralogs, which diverged after a gene duplication event. This distinction is essential when studying and predicting gene functionality, since the orthologs are more likely to have similar function across the organisms than the paralogs (Altenhoff *et al.*, 2012). The potential of using orthologs to predict function has motivated the development of a number of methods that aim at inferring orthology. Most of these methods can be classified in two types (Trachana *et al.*, 2011), namely graph-based methods (Powell *et al.*, 2014; Kristensen *et al.*, 2013; Waterhouse *et al.*, 2013) and tree-based methods

(Ruan *et al.*, 2008; van der Heijden *et al.*, 2007; Huerta-Cepas *et al.*, 2014).

Methods also exist for making predictions when homologous genes of known function cannot be found. Many such methods are based on machine-learning algorithms, such as artificial neural networks or support vector machines, to predict, for example, subcellular localization (Nakai and Horton, 1999; Briesemeister *et al.*, 2010), broad functional classes (Jensen *et al.*, 2003), or posttranslational modifications (Biswas *et al.*, 2010; Steentoft *et al.*, 2013; Blom *et al.*, 1999) from protein sequences. Various methods have also been developed to predict which proteins interact (Pazos and Valencia, 2001; Matthews *et al.*, 2001) or are functionally related from their evolution or genomic context (Hamp *et al.*, 2013; Michel *et al.*, 2014; Huynen *et al.*, 2000; Marcotte, 1999; Snel *et al.*, 2002).

Quality Assessment and Benchmarking

To integrate heterogeneous datasets of varying quality, it is important to assess the quality of each dataset and put them on a common scale. This can be achieved by quantifying how well they agree with a common reference set, often called the gold standard. This benchmarking process both reveals the relative quality of the datasets and provides a common scale, namely the agreement of a dataset with the gold standard (Qi *et al.*, 2006). However, because all parts of a dataset may not be of the same quality, such global benchmarking is typically not sufficient. This can be solved by developing a scoring scheme for each type of data, which allows the data points from a single dataset to be ranked by confidence. By calibrating these scores against the gold standard, it is possible to estimate the quality of each individual data point rather than for the dataset as a whole. This process converts all datasets to probabilistic scores that are directly comparable to each other and can be combined to provide experimental and prediction inferred information under a unified scoring scheme (Kuhn *et al.*, 2014; Warde-Farley *et al.*, 2010; Troyanskaya *et al.*, 2003; Franceschini *et al.*, 2013; Chen *et al.*, 2009). In addition, it enables the inclusion of all available data, even if they are considered as unreliable. Instead of just discarding low-confidence data, all data should ideally be taken into account and weighted according to its reliability.

Open Licenses

As discussed earlier, data availability can be an issue for data integration. In some cases the solution is obvious; if labor intensive: if the data have not already been gathered in a database, the only solution is that someone does so (Berman *et al.*, 2000; Barretina *et al.*, 2012; Chang *et al.*, 2013). However, in other cases a database exists, but its license stands in the way of it being integrated with other databases. In case of commercial databases, one option is obviously to develop a public alternative from scratch; an example of this is the PubChem Project (Wang *et al.*, 2009). However, in some cases it may be possible to instead transform a commercial resource into a public one. A notable example of this is the ChEMBL database (Willighagen *et al.*, 2013), formerly sold

as StARlite, which was made public with funding from the Wellcome Trust. License issues also exist for many databases created in academia. As the development of databases in academia is generally funded through grants, funding agencies like the National Institutes of Health (NIH) and Wellcome Trust increasingly require that the results of funded projects are made public to maximize their usefulness (Guttmacher *et al.*, 2009; Roche *et al.*, 2014). To make it clear that others can legally reuse the data as they see fit, this is typically done by mandating that the data are made either public domain or available under an open license such as the Creative Commons Attribution license.

Critical Thinking in Data Interpretation

For users to be able to properly use databases and other resources, it is important that developers document how the data stored in a resource were produced, curated, and analyzed. However, irrespective of whether this was done or not, users should apply the same critical thinking when interpreting the results from a bioinformatics resource that they apply when performing wet-lab experiments. In both cases, the researcher should ask a specific question and design controls. Consider a researcher who performs a typical case-control study, using mass spectrometry to identify significantly regulated proteins. If a statistical analysis tool shows that the regulated proteins are enriched for mitochondrial proteins, a good negative control would be to use the same tool to test if the same is true for the proteins identified as nonregulated. This would reveal if the regulated proteins are in fact enriched for mitochondrial proteins, or if this is likely an artifact of detection biases.

Conclusions

- Analyses of the huge amount of the biological data produced over the last years have it obvious that biological systems are far more complex than originally thought and that they cannot be understood one protein at a time.
- Applying systems biology approaches to cellular regulation requires integration of biological data representing multiple layers of regulation.
- The dimensionality, heterogeneity, and scattered nature of the biological data poses many challenges concerning data access, integration, analyzing, and interpreting the data.
- There is no single solution to addressing all the challenges. Important steps in the right direction include standardization of names and data formats, improving access to and licenses on data, text-mining of the scientific literature, and development of computational methods for data quality assessment and inference.
- Although we used examples from protein interactions throughout the article, the challenges and possible solutions to overcome them are general and apply to almost all types of high-throughput data, including genomics, transcriptomics, and proteomics data.
- Due to the complexity of many methods used in the field, controls and critical thinking is as important for computational experiments as it is in the wet lab.

See also: Horizontal Integration: Networks: Understanding of 'Networks' *In Vitro* and/or *In Vivo*

References

- Alsheikh-Ali, A., Qureshi, W., Al-Mallah, M.H., Ioannidis, J.P., 2011. Public availability of published research data in high-impact journals. *PLoS One* 6 (9), e24357. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3168487&tool=pmcentrez&rendertype=abstract> (accessed 11.07.14).
- Altenhoff, A.M., Studer, R.A., Robinson-Rechavi, M., *et al.*, 2012. Resolving the ortholog conjecture: Orthologs tend to be weakly, but significantly, more similar in function than paralogs. *PLoS Computational Biology* 8 (5), e1002514. Available at: <http://dx.plos.org/10.1371/journal.pcbi.1002514> (accessed 29.08.14).
- Apweiler, R., Bairoch, A., Wu, C.H., *et al.*, 2004. UniProt: The universal protein knowledgebase. *Nucleic Acids Research* 32 (Database issue), D115–D119. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=308865&tool=pmcentrez&rendertype=abstract> (accessed 05.08.14).
- Barretina, J., Caponigro, G., Stransky, N., *et al.*, 2012. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483 (7391), 603–607. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3320027&tool=pmcentrez&rendertype=abstract> (accessed 10.07.14).
- Bellazzi, R., 2014. Big data and biomedical informatics: A challenging opportunity. *Yearbook of Medical Informatics* 9 (1), 8–13.
- Berger, B., Peng, J., Singh, M., 2013. Computational solutions for omics data. *Nature Reviews. Genetics* 14 (5), 333–346.
- Berman, H.M., Westbrook, J., Feng, Z., *et al.*, 2000. The protein data bank. *Nucleic Acids Research* 28 (1), 235–242. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=102472&tool=pmcentrez&rendertype=abstract> (accessed 11.02.15).
- Biswas, A.K., Noman, N., Sikder, A.R., 2010. Machine learning approach to predict protein phosphorylation sites by incorporating evolutionary information. *BMC Bioinformatics* 11, 273. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2887807&tool=pmcentrez&rendertype=abstract> (accessed 11.02.15).
- Blake, J., Bult, C.J., 2006. Beyond the data deluge: Data integration and bio-ontologies. *Journal of Biomedical Informatics* 39 (3), 314–320.
- Blom, N., Gammeltoft, S., Brunak, S., 1999. Sequence and structure-based prediction of eukaryotic protein phosphorylation sites. *Journal of Molecular Biology* 294 (5), 1351–1362. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10600390> (accessed 13.08.14).
- Bouwmeester, T., Bauch, A., Ruffner, H., *et al.*, 2004. A physical and functional map of the human TNF- α /NF- κ B signal transduction pathway. *Nature Cell Biology* 6 (2), 97–105. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/14743216> (accessed 13.09.14).
- Brazma, A., Hingamp, P., Quackenbush, J., *et al.*, 2001. Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nature Genetics* 29 (4), 365–371. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11726920> (accessed 11.08.14).
- Briesemeister, S., Rahnenführer, J., Kohlbacher, O., 2010. YLoc – an interpretable web server for predicting subcellular localization. *Nucleic Acids Research* 38 (Web Server issue), W497–W502. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2896088&tool=pmcentrez&rendertype=abstract> (accessed 16.08.14).
- Chang, K., Creighton, C.J., Davis, C., *et al.*, 2013. The Cancer Genome Atlas Pan-Cancer analysis project. *Nature Genetics* 45 (10), 1113–1120. Available at: <http://www.nature.com/doi/10.1038/ng.2764> (accessed 26.09.13).
- Chen, J.Y., Marnidipalli, S., Huan, T., 2009. HAPPI: An online database of comprehensive human annotated and predicted protein interactions. *BMC Genomics* 10 (Suppl. 1), S16. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2709259&tool=pmcentrez&rendertype=abstract> (accessed 31.07.14).
- Cusick, M.E., Yu, H., Smolyar, A., *et al.*, 2009. Literature-curated protein interaction datasets. *Nature Methods* 6 (1), 39–46.
- De la Cruz, N., Bromberg, S., Pasko, D., *et al.*, 2005. The Rat Genome Database (RGD): Developments towards a phenotype database. *Nucleic Acids Research* 33 (Database issue), D485–D491. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=540004&tool=pmcentrez&rendertype=abstract> (accessed 06.09.14).

- Dolinski, K., Botstein, D., 2007. Orthology and functional conservation in eukaryotes. *Annual Review of Genetics* 41, 465–507. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17678444> (accessed 09.07.14).
- El-Khatib, H.T., Williams, M.H., MacKinnon, L.M., Marwick, D.H., 2000. A framework and test-suite for assessing approaches to resolving heterogeneity in distributed databases. *Information and Software Technology* 42 (7), 505–515.
- Eppig, J.T., Bult, C.J., Kadin, J.A., *et al.*, 2005. The Mouse Genome Database (MGD): From genes to mice – A community resource for mouse biology. *Nucleic Acids Research* 33 (Database issue), D471–D475. Available at: http://nar.oxfordjournals.org/content/33/suppl_1/D471.long (accessed 06.09.14).
- Franceschini, A., Szklarczyk, D., Frankild, S., *et al.*, 2013. STRING v9.1: Protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Research* 41 (Database issue), D808–D815. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3531103&tool=pmcentrez&rendertype=abstract> (accessed 11.12.13).
- Freedman, L.P., Inglese, J., 2014. The increasing urgency for standards in basic biologic research. *Cancer Research* 74 (15), 4024–4029. Available at: <http://cancerres.aacrjournals.org/content/early/2014/07/17/0008-5472.CAN-14-0925> (accessed 05.09.14).
- Fundel, K., Zimmer, R., 2006. Gene and protein nomenclature in public databases. *BMC Bioinformatics* 7 (1), 372. Available at: <http://www.biomedcentral.com/1471-2105/7/372> (accessed 06.09.14).
- Gelernter, J., Lesk, M., 2011. Use of ontologies for data integration and curation. *International Journal of Digital Curation* 6 (1), 70–78.
- Gray, K., Daugherty, L.C., Gordon, S.M., *et al.*, 2013. Genenames.org: The HGNC resources in 2013. *Nucleic acids Research* 41 (Database issue), D545–D552. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3531211&tool=pmcentrez&rendertype=abstract> (accessed 11.07.14).
- Guttmacher, A.E., Nabel, E.G., Collins, F.S., 2009. Why data-sharing policies matter. *Proceedings of the National Academy of Sciences of the United States of America* 106 (40), 16894. Available at: <http://www.pnas.org/content/106/40/16894.full> (accessed 05.09.14).
- Hamp, T., Kassner, R., Seemayer, S., *et al.*, 2013. Homology-based inference sets the bar high for protein function prediction. *BMC Bioinformatics* 14 (Suppl. 3), S7. Available at: <http://www.biomedcentral.com/1471-2105/14/S3/S7> (accessed 08.09.14).
- van der Heijden, R.T.J.M., Snel, B., van Noort, V., Huynen, M.A., 2007. Orthology prediction at scalable resolution by phylogenetic tree analysis. *BMC Bioinformatics* 8, 83.
- Huerta-Cepas, J., Capella-Gutiérrez, S., Pryszcz, L.P., Marcet-Houben, M., Gabaldón, T., 2014. PhylomeDB v4: Zooming into the plurality of evolutionary histories of a genome. *Nucleic Acids Research* 42 (Database issue), D897–D902.
- Huynen, M., Snel, B., Lathe 3rd, W., Bork, P., 2000. Predicting protein function by genomic context: Quantitative evaluation and qualitative inferences. *Genome Research* 10 (8), 1204–1210. Available at: <http://genome.cshlp.org/content/10/8/1204.long#ref-14> (accessed 15.08.14).
- Ideker, T., Krogan, N.J., 2012. Differential network biology. *Molecular Systems Biology* 8 (565), 1–9. Available at: <http://www.nature.com/doi/10.1038/msb.2011.99> (accessed 17.01.12).
- Jensen, L.J., Gupta, R., Staerfeldt, H.H., Brunak, S., 2003. Prediction of human protein function according to Gene Ontology categories. *Bioinformatics* 19 (5), 635–642. Available at: <http://bioinformatics.oxfordjournals.org/cgi/doi/10.1093/bioinformatics/btg036> (accessed 05.09.14).
- Jensen, L.J., Bork, P., 2008. Not comparable, but complementary. *Science (New York, N.Y.)* 322 (5898), 56–57. Available at: <http://www.sciencemag.org/content/322/5898/56.short> (accessed 05.09.14).
- Joyce, A.R., Palsson, B.O., 2006. The model organism as a system: Integrating “omics” data sets. *Nature reviews. Molecular Cell Biology* 7 (3), 198–210.
- Kaiser, J., 2002. Public-private group maps out initiatives. *Science (New York, N.Y.)* 296 (5569), 827. Available at: http://www.sciencemag.org/content/296/5569/827.summary&xtor=show=&HITS=10&hits=10&RESULTFORMAT=&searchid=1027414274815_1591&stored_search=&FIRSTINDEX=0&volume=296&firstpage=827&fdate=10/1/1995&tdate=7/31/2002 (accessed 05.09.14).
- Kelso, J., Visagie, J., Theiler, G., *et al.*, 2003. eVOC: A controlled vocabulary for unifying gene expression data. *Genome Research* 13 (6A), 1222–1230. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=403650&tool=pmcentrez&rendertype=abstract> (accessed 31.07.14).
- Kristensen, D.M., Waller, A.S., Yamada, T., *et al.*, 2013. Orthologous gene clusters and taxon signature genes for viruses of prokaryotes. *Journal of Bacteriology* 195 (5), 941–950.
- Kuhn, M., Szklarczyk, D., Pletscher-Frankild, S., *et al.*, 2014. STITCH 4: Integration of protein-chemical interactions with user data. *Nucleic Acids Research* 42 (1), D401–D407. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/24293645> (accessed 15.01.14).
- Lauffenburger, D., 2012. The multiple dimensions of integrative biology. *Integrative Biology: Quantitative Biosciences from Nano to Macro* 4 (1), 9. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22158999> (accessed 05.09.14).
- Lee, H.C., Lai, K., Lorenc, M.T., *et al.*, 2012. Bioinformatics tools and databases for analysis of next-generation sequence data. *Briefings in Functional Genomics* 11 (1), 12–24. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22184335> (accessed 11.07.14).
- Lee, J.A., Spidlen, J., Boyce, K., *et al.*, 2008. MIFlowCyt: The minimum information about a Flow Cytometry Experiment. *Cytometry. Part A: The journal of the International Society for Analytical Cytology* 73 (10), 926–930. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2773297&tool=pmcentrez&rendertype=abstract> (accessed 09.07.14).
- Lunshof, J.E., Chadwick, R., Vorhaus, D.B., Church, G.M., 2008. From genetic privacy to open consent. *Nature Reviews. Genetics* 9 (5), 406–411. Available at: <http://dx.doi.org/10.1038/nrg2360> (accessed 23.08.14).
- Marcotte, E.M., 1999. Detecting protein function and protein-protein interactions from genome sequences. *Science* 285 (5428), 751–753. Available at: <http://www.sciencemag.org/cgi/doi/10.1126/science.285.5428.751> (accessed 21.07.14).
- Matthews, L.R., Vaglio, P., Reboul, J., *et al.*, 2001. Identification of potential interaction networks using sequence-based searches for conserved protein-protein interactions or “interologs”. *Genome Research* 11 (12), 2120–2126. Available at: <http://genome.cshlp.org/content/11/12/2120.abstract> (accessed 21.08.14).
- von Mering, C., Krause, R., Snel, B., *et al.*, 2002. Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* 417 (6887), 399–403. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12000970> (accessed 11.02.15).
- Michel, N., Hayat, S., Skwark, M.J., *et al.*, 2014. PconsFold: Improved contact predictions improve protein models. *Bioinformatics* 30 (17), i482–i488. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4147911&tool=pmcentrez&rendertype=abstract> (accessed 26.08.14).
- Nakai, K., Horton, P., 1999. PSORT: A program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends in Biochemical Sciences* 24 (1), 34–35. Available at: <http://www.cell.com/article/S096800049801336X/fulltext> (accessed 25.08.14).
- Orchard, S., Al-Lazikani, B., Bryant, S., *et al.*, 2011. Minimum information about a bioactive entity (MIBAE). *Nature Reviews. Drug Discovery* 10 (9), 661–669. Available at: http://www.nature.com/nrd/journal/v10/n9/full/nrd3503.html#WT_ec_id=NRD-201109 (accessed 06.09.14).
- Orchard, S., Hermjakob, H., 2008. Standardising Proteomics Data – The work of the HUPO proteomics standards initiative. *Journal of Proteomics & Bioinformatics* 1 (April), 3–5.
- Orchard, S., Salwinski, L., Kerrien, S., *et al.*, 2007. The minimum information required for reporting a molecular interaction experiment (MIMIx). *Nature Biotechnology* 25 (8), 894–898. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17687370> (accessed 04.06.14).
- Pazos, F., Valencia, A., 2001. Similarity of phylogenetic trees as indicator of protein-protein interaction. *Protein Engineering* 14 (9), 609–614. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11707606> (accessed 05.09.14).
- Philippi, S., Köhler, J., 2006. Addressing the problems with life-science databases for traditional uses and systems biology. *Nature Reviews. Genetics* 7 (6), 482–488.
- Powell, S., Forslund, K., Szklarczyk, D., *et al.*, 2014. eggNOG v4.0: Nested orthology inference across 3686 organisms. *Nucleic Acids Research* 42 (Database issue), D231–D239.
- Qi, Y., Bar-Joseph, Z., Klein-Seetharaman, J., 2006. Evaluation of different biological data and computational classification methods for use in protein interaction prediction. *Proteins* 63 (3), 490–500. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3250929&tool=pmcentrez&rendertype=abstract> (accessed 05.09.14).
- Rehbolz-Schuhmann, D., Oellrich, A., Hoehndorf, R., 2012. Text-mining solutions for biomedical research: Enabling integrative biology. *Nature Reviews Genetics* 13 (December), Available at: <http://www.nature.com/doi/10.1038/nrg3337> (accessed 14.11.12).
- Roche, D.G., Lanfear, R., Binning, S.A., *et al.*, 2014. Troubleshooting public data archiving: Suggestions to increase participation. *PLoS Biology* 12 (1), p.e1001779. Available at: <http://dx.plos.org/10.1371/journal.pbio.1001779> (accessed 07.08.14).
- Rodríguez-Esteban, R., 2009. Biomedical text mining and its applications. *PLoS Computational Biology* 5 (12), p.e1000597. Available at: <http://dx.plos.org/10.1371/journal.pcbi.1000597> (accessed 05.09.14).
- Rodríguez-Ulloa, A., Rodríguez, R., 2008. Bioinformatic tools for proteomic data analysis: An overview. *Biotechnología Aplicada* 25, 312–319.

- Ruan, J., Li, H., Chen, Z., *et al.*, 2008. TreeFam: 2008 update. *Nucleic Acids Research* 36 (Database issue), D735–D740.
- Schuurman, N., Leszczynski, A., 2008. Ontologies for bioinformatics. *Bioinformatics and Biology Insights* 2, 187–200.
- Shamu, C.E., Wiemann, S., Boutros, M., 2012. On target: A public repository for large-scale RNAi experiments. *Nature Cell Biology* 14 (2), 115. Available at: <http://dx.doi.org/10.1038/ncb2435> (accessed 06.09.14).
- Smith, B., Ashburner, M., Rosse, C., *et al.*, 2007. The OBO Foundry: Coordinated evolution of ontologies to support biomedical data integration. *Nature Biotechnology* 25 (11), 1251–1255. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2814061&tool=pmcentrez&rendertype=abstract> (accessed 14.01.14).
- Snel, B., Bork, P., Huynen, M.A., 2002. The identification of functional modules from the genomic association of genes. *Proceedings of the National Academy of Sciences of the United States of America* 99 (9), 5890–5895. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=122872&tool=pmcentrez&rendertype=abstract> (accessed 05.09.14).
- Stentoft, C., Vakhrushev, S.Y., Joshi, H.J., *et al.*, 2013. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO Journal* 32 (10), 1478–1488. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3655468&tool=pmcentrez&rendertype=abstract> (accessed 02.09.14).
- The Gene Ontology Consortium, 2008. The Gene Ontology project in 2008. *Nucleic Acids Research* 36 (Database issue), D440–D444.
- Trachana, K., Larsson, T.A., Powell, S., *et al.*, 2011. Orthology prediction methods: A quality assessment using curated protein families. *BioEssays: News and Reviews In Molecular, Cellular and Developmental Biology* 33 (10), 769–780.
- Troyanskaya, O.G., Dolinski, K., Owen, A.B., Altman, R.B., Botstein, D., 2003. A Bayesian framework for combining heterogeneous data sources for gene function prediction (in *Saccharomyces cerevisiae*). *Proceedings of the National Academy of Sciences of the United States of America* 100 (14), 8348–8353. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=166232&tool=pmcentrez&rendertype=abstract> (accessed 11.02.15).
- Turinsky, A.L., Razick, S., Turner, B., Donaldson, I.M., Wodak, S.J., 2010. Literature curation of protein interactions: Measuring agreement across major public databases. *Database: The Journal of Biological Databases and Curation*, 2010. baq026. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3011985&tool=pmcentrez&rendertype=abstract> (accessed 19.06.14).
- Vempati, U.D., Chung, C., Mader, C., *et al.*, 2014. Metadata standard and data exchange specifications to describe, model, and integrate complex and diverse high-throughput screening data from the library of integrated network-based cellular signatures (LINCS). *Journal of Biomolecular Screening* 19 (5), 803–816. Available at: <http://jbx.sagepub.com/content/19/5/803> (accessed 05.09.14).
- Wang, Y., Xiao, J., Supek, T.O., *et al.*, 2009. PubChem: A public information system for analyzing bioactivities of small molecules. *Nucleic Acids Research* 37 (Web Server issue), W623–W633. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2703903&tool=pmcentrez&rendertype=abstract> (accessed 05.09.14).
- Warde-Farley, D., Donaldson, S.L., Comes, O., *et al.*, 2010. The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Research* 38 (Web Server issue), W214–W220. Available at: http://nar.oxfordjournals.org/content/38/suppl_2/W214.long (accessed 09.07.14).
- Waterhouse, R.M., Tegenfeldt, F., Li, J., Zdobnov, E.M., Kriventseva, E.V., 2013. OrthoDB: A hierarchical catalog of animal, fungal and bacterial orthologs. *Nucleic Acids Research* 41 (Database issue), D358–D365.
- Willighagen, E.L., Waagmeester, A., Spjuth, O., *et al.*, 2013. The ChEMBL database as linked open data. *Journal of Cheminformatics* 5 (1), 23. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3700754&tool=pmcentrez&rendertype=abstract> (accessed 11.02.15).
- Zhu, F., Patumcharoenpol, P., Zhang, C., *et al.*, 2013. Biomedical text mining and its applications in cancer research. *Journal of Biomedical Informatics* 46 (2), 200–211.

Relevant Websites

wiki.creativecommons.org/
Creative Commons.

www.ebi.ac.uk/
EMBL-EBI.

www.expasy.org/
ExPASy.

mibbi.sourceforge.net/portal.shtml
Minimum Information guidelines.

<http://publicaccess.nih.gov/index.htm>
NIH Public Access Policy.

www.obofoundry.org/
OBO foundry.

www.bio.cam.ac.uk/funding/wtopenaccess
Wellcome Trust Open Access.

HORIZONTAL INTEGRATION: DISEASE

Contents

Computational and Systems Cancer Biology
Drug Targeting

Computational and Systems Cancer Biology

A Califano, Columbia University Medical Center, New York, NY, USA

G Stolovitzky, Icahn School of Medicine at Mount Sinai, New York, NY, USA and IBM Research, Yorktown Heights, NY, USA

© 2016 Elsevier Inc. All rights reserved.

A dramatic reorganization of the body of knowledge on cancer initiation, progression, and drug sensitivity has occurred over the last decade. This reorganization has been primarily driven by an explosion in the volume of molecular-level data, unquestionably resulting from the introduction of novel technologies for the cost-effective, high-throughput characterization of cancer cells' genetic, epigenetic, and functional landscapes. Such exponential data growth has required increasingly sophisticated analytical methodologies, ranging from the integrative statistical approaches, to complex machine learning methodologies for clustering and classification, to algorithms for the assembly and interrogation of cellular networks. Taken together, these approaches define the field of Computational and Systems Cancer Biology.

Development of gene expression profiling technologies, for instance, initially using microarray and more recently RNA sequencing platforms, has allowed fine-grain monitoring of RNA transcripts dynamics following tumorigenesis (Alizadeh *et al.*, 2000; Network, 2008; Hudson *et al.*, 2010) or following drug perturbation (Lamb *et al.*, 2006). Computational analysis of these profiles has helped elucidate molecularly distinct, yet morphologically indistinguishable cancer subtypes, associated with differential outcome (Phillips *et al.*, 2006), progression (Lossos *et al.*, 2002), and drug sensitivity (Barretina *et al.*, 2012). They have also helped identify key genetic programs that are consistently activated (e.g., proliferation, migration, and immuno-evasion), inactivated (apoptosis, senescence), or modulated (adhesion, angiogenesis, etc.) in tumorigenesis (Alizadeh *et al.*, 2000; Klein *et al.*, 2003). Collectively, these results have led us to better understand cancer as a highly heterogeneous disease, while also providing a molecular basis to study commonalities, such as sensitivity to targeted agents that are spanning across highly diverse tumors.

At the other end of the spectrum, genome wide studies of both heritable and somatic human variation have moved from theoretical concept to practical reality, thus opening a new window on both aspects of cancer etiology. Yet, even as the repertoire of recurrent germline and somatic alterations

increases in several tumor types, elucidating their mechanistic role in tumor-related phenotypes remains an overall elusive target.

Indeed, despite impressive algorithmic advances and large-scale data availability, the precise relationship between genetic alterations and dysregulation of cancer-related programs is still poorly understood and the precise cascade of molecular events leading to tumorigenesis and progression remains largely uncharted. For instance, while the mesenchymal subtype of glioblastoma represents a distinct and highly aggressive subtype, only relatively rare mutations in the NF1 gene (Verhaak *et al.*, 2010) and in CTNND2 (Frattini *et al.*, 2013) could be identified by association studies, and the mechanism by which NF1 drives the subtype has not been elucidated (Figure 1). Similarly, while large sequencing efforts have identified several mutations in the B cell receptor pathway that partially co-segregate with the more aggressive activated B cell (ABC) subtype of diffuse large B cell lymphoma, the same mutations can also be found in samples representing the more indolent germinal center B cell (GCB) subtype (Compagno *et al.*, 2009). Thus, while some tumor subtypes co-segregate with specific mutational landscapes (e.g., amplification of hormone receptors in luminal breast cancer), many tumors either lack subtypes entirely or have subtypes that fail to obviously co-segregate with tumor genetics. Similarly, tumor specific characteristics, such as sensitivity to specific targeted drugs or their indolent versus aggressive nature is not effectively predicted by the genetics, such as in prostate cancer (Taylor *et al.*, 2010) or in cell line sensitivity to targeted compounds (Barretina *et al.*, 2012). This suggests that, similar to complex heritable diseases, the relationship between genetics and cancer, with a few notable exceptions, may be too complex for standard association discovery approaches and may require more model-based approaches, as also discussed in (Califano *et al.*, 2012). Indeed, few epistatic interactions and low-penetrance variants have been identified so far in complex diseases (Narayanan *et al.*, 2010), mostly due to cohort size limitations (Manolio *et al.*, 2009) and lack of methodologies.

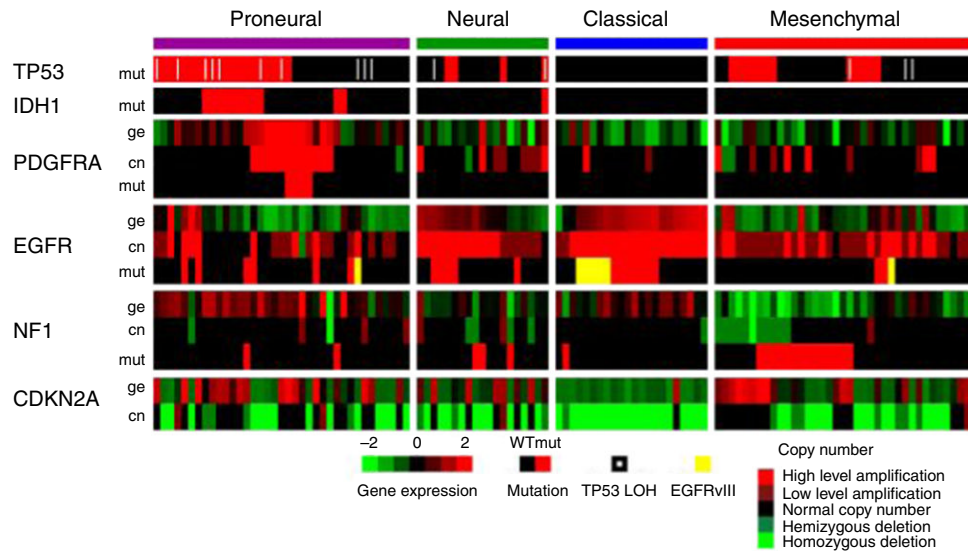


Figure 1 Genomic alterations in glioma co-segregate with only some of the identified molecular subtypes. Reproduced from Verhaak, R.G., Hoadley, K.A., Purdom, E., *et al.*, Cancer Genome Atlas Research, 2010. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* 17 (1), 98–110.

The same is likely to occur in cancer, unless novel methodologies and study designs are introduced.

Integrative and Systems Biology Methodologies

To address these challenges, the computational and systems biology community has developed a variety of integrative and model-based methodologies. Integrative methods are frequently confused with systems biology ones, although they are driven by very different principles. Their value arises from the ability to combine partially or fully statistically independent evidence from diverse sources, for instance by combining functional evidence from gene expression profiles, as well as from genetic and epigenetic (henceforth genomic) mutational profiles. When independent evidences are combined, using methods such as Fisher's (Fisher, 1925), Stouffer's (Stouffer *et al.*, 1949), Naïve Bayes classifiers (Domingos and Pazzani, 1997), or Random Forest (Breiman, 2001), among others, overall statistical significance increases dramatically. This is a key when the number of testable hypotheses is so large that extremely low p -values are necessary to achieve statistical significance. Even when the evidences have partial dependencies, statistical integration methods such as Bayesian Networks (Pearl, 1985) or Brown's method (Brown, 1975) still produce important gains in overall significance.

In contrast, systems biology methods rely on an orthogonal approach to address the vast number of testable hypotheses. Rather than improving the statistical significance of the individual hypothesis by integrating multiple clues, they use biological models, such as regulatory network models, to eliminate the vast majority of testable hypotheses. The term 'regulatory network' has typically referred to the repertoire of physical interactions involved in regulating mRNA transcripts, including both transcriptional and posttranscriptional processes. In systems biology, however, the term is now used

more broadly to indicate any physical interaction affecting the abundance, chemical modification, or activity of a biomolecule, including some classes of causal indirect interactions. Within this broader acceptance, we will use the term to indicate transcriptional, posttranscriptional, posttranslational, and metabolic interactions, including protein–protein interactions in stable complexes. Regulatory network models can be represented using a variety of formalisms, from qualitative representation such as cancer pathways, to more quantitative models based on statistical regulatory dependencies or kinetic differential equations. Due to parametric complexity, the latter are generally restricted only to specific pathways, such as those controlling apoptosis or growth signals.

Given a regulatory model, a typical way to formulate questions using a systems biology approach (e.g., to identify cancer driver genes), is to compute the probability $p(G|M,D)$. This quantity represents the posterior probability that (G) is a driver gene or driver mutation, given a regulatory model (M) and a set of measurement data (D) produced by that model. This type of reasoning was originally introduced using the Bayesian formalism (Friedman, 2004) and has since been extended to a variety of additional methodologies (Schadt *et al.*, 2005; Carro *et al.*, 2010; Kwong *et al.*, 2012). While models of cancer cell regulation (henceforth *interactomes*) are still in their infancy, they are already providing significant new insight into cancer drivers and mechanisms. This is likely to improve significantly, as they become more accurate and comprehensive. A critical advantage of these approaches is that any inferred cancer-relevant gene-product or mutation is placed within a corresponding regulatory context. This is critically valuable, as such information can be harnessed to generate experimentally testable hypotheses about the mechanism by which these gene products and mutations affect cancer phenotypes, see (Chen *et al.*, 2014; Sumazin *et al.*, 2011; Carro *et al.*, 2010; Kwong *et al.*, 2012) for some representative examples. In contrast, while non-model-based approaches have helped identify a

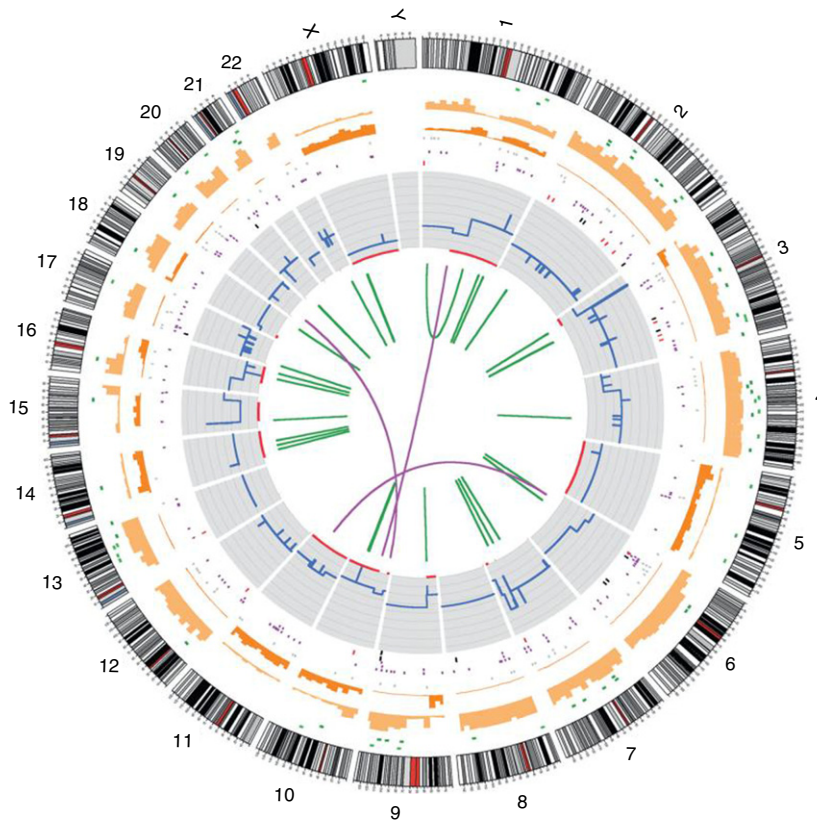


Figure 2 Circos plot showing the whole-genome catalog of somatic mutations from the melanoma cell line COLO-829, including ~30 000 somatic base substitutions and ~1000 somatic insertions and/or deletions. 272 somatic substitutions are in exons, including 155 missense, 16 nonsense, and 101 silent mutations. The number and type of mutation are highly variable across different cancer genomes. Chromosome number and karyotype are indicated on the exterior of the plot. Key: blue lines, copy number across each chromosome; red lines, sites of loss of heterozygosity (LOH); green, intra-chromosomal rearrangements; purple, interchromosomal rearrangements; red, nonsense mutations; green, missense mutations; black, silent mutations; brown, intronic and intergenic mutations (merged). Reproduced with permission from Pleasance, E.D., Cheetham, R.K., Stephens, P.J., *et al.*, 2010. A comprehensive catalogue of somatic mutations from a human cancer genome. *Nature* 463 (7278), 191–196.

variety of driver genes and mutations, their inference is based on the analysis of statistical or complex machine learning based dependencies. These do not lend themselves to mechanism elucidation and validation. Indeed, statistical association are generally validated either by analysis of independent patient cohorts or by direct analysis of the phenotypic effects resulting from the experimental modulation of the gene or mutation of interest (e.g., using isogenic models), independent of mechanism.

These remarkable advances in integrative and systems biology approaches to the study of cancer, see (Pe'er and Hachohen, 2011; Werner *et al.*, 2014; Califano *et al.*, 2012) for some recent reviews, would not have been possible without the availability of molecular profiles representing distinct data modalities across a variety of human malignancies, as made available by The Cancer Genome Atlas (TCGA) (Network, 2008), the International Cancer Consortium (ICG) (Hudson *et al.*, 2010), and the Therapeutically Applicable Research To Generate Effective Treatments (TARGET) consortia (Mullighan *et al.*, 2009), among others. Indeed, these new methodologies are becoming increasingly valuable precisely because of the

overwhelming number of testable hypotheses resulting from direct statistical analysis of these resources.

This is arguably the key limitation of more traditional associative methods, such as those used in genome wide association studies (GWAS). Indeed, given the very large number of somatic events routinely discovered in cancer genomes, including mutations, translocations, gene fusions, aberrant copy number changes, and structural rearrangements (Stephens *et al.*, 2011), see Figure 2, distinguishing 'driver' from 'passenger' alterations is challenging and often impossible on a purely statistical basis. Thus, not surprisingly, the tumors where most causal alterations were originally elucidated are those with the least complex mutational landscapes, such as leukemias and lymphomas. The problem only becomes more complex when broader genomic modalities from epigenetic profiles to extracellular signals are considered or, even worst, when epistatic or synergistic combinations are considered (i.e., pairwise combinations of genomic alterations).

Rather than being antithetic, integrative and systems methods are often used in combination, resulting in significant contributions to the elucidation of key genetic drivers in

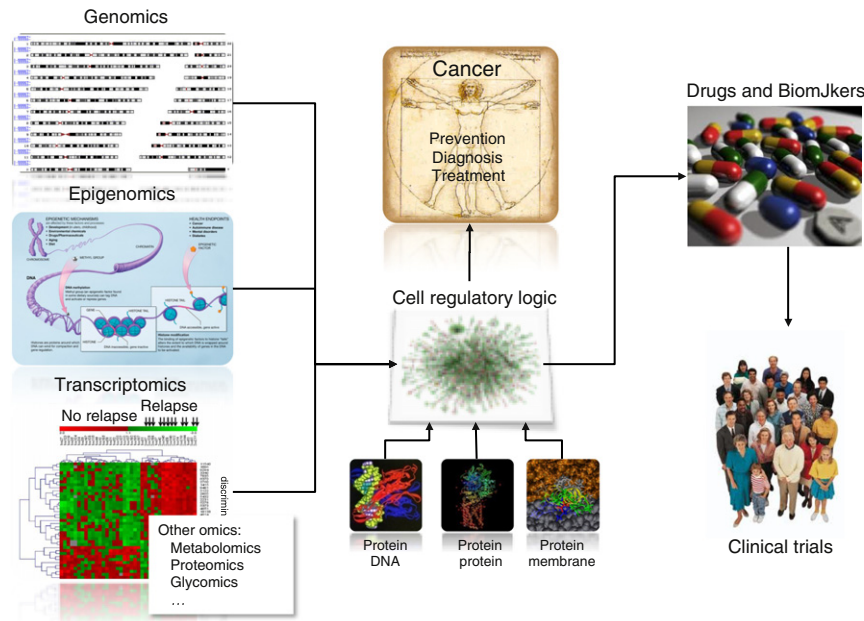


Figure 3 The ‘Omics layer of the cell both encode and are processed by a context-specific regulatory logic. At the atomic level, this logic is implemented via molecular interactions, such as protein–DNA, protein–protein, protein–RNA, RNA–RNA, etc. Dissection and interrogation of this logic in context specific fashion, using systems biology approaches, allows for elucidating driver genes responsible for the presentation of relevant cancer-related phenotypes.

cancer and other diseases (Califano *et al.*, 2012). The rationale for these methods is that interactomes, representing causal molecular interactions in the cell help pinpoint candidate driver gene(s) within genetic programs that are dysregulated in cancer. These can then be experimentally validated as either individual (Chudnovsky *et al.*, 2014; Della Gatta *et al.*, 2012; Zhao *et al.*, 2009; Compagno *et al.*, 2009; Palomero *et al.*, 2006; Taylor *et al.*, 2010; Network, 2008) or synergistic/co-operative (Aytes *et al.*, 2014; Carro *et al.*, 2010; Kwong *et al.*, 2012) drivers of tumor-related phenotypes, including genes harboring causal genomic alterations in populations (Chen *et al.*, 2014; King *et al.*, 2009) and even within individual tumors (Ding *et al.*, 2010), see Figure 3.

The use of regulatory models for the causal association of heritable or somatic variants to gene expression patterns (i.e., genetical genomics) was originally pioneered in plants (Schadt *et al.*, 2003) and metabolic disease (Yang *et al.*, 2009). Yet, more recently they have been used successfully in neurodegenerative phenotypes (Zhang *et al.*, 2013) and cancer (Chen *et al.*, 2014; Zhao *et al.*, 2009; Della Gatta *et al.*, 2012; De Keersmaecker *et al.*, 2010). For instance, identification of the novel HUWE1-MYCIN-DLL3 cascade in brain tumors, where HUWE1 is frequently deleted, was possible by reverse engineering posttranslational modulators of MYCN activity as well as its downstream targets using reverse engineering algorithms (Zhao *et al.*, 2009). Similarly, the role of RUNX1 as a tumor suppressor, harboring frequent mutations in T-cell acute lymphoblastic leukemia (T-ALL) (Della Gatta *et al.*, 2012), was elucidated based on the overlap of its regulatory program with those of TLX1 and TLX3, two previously established oncogenes (De Keersmaecker *et al.*, 2010). In some cases, molecular-level

understanding of regulatory events may allow elucidating dependencies between a phenotype and a repertoire of genetic events, which could be identified by statistical approaches on an individual basis. For example, deletion of any combination of 13 genetic loci distributed across the entire genome leads to functional inactivation of PTEN in glioma patients, via the recently proposed competitive endogenous RNA (ceRNA) mechanism (Sumazin *et al.*, 2011). Yet none of these events would be individually significant, either functionally or statistically. Use of regulatory or dependency models, ranging from graphical models to fully detailed kinetic models, has exploded over the last 5 years, with applications ranging from the study of key drivers of tumorigenesis in melanoma (Kwong *et al.*, 2012; Akavia *et al.*, 2010), glioma (Chudnovsky *et al.*, 2014; Chen *et al.*, 2014; Sumazin *et al.*, 2011; Carro *et al.*, 2010; Sonabend *et al.*, 2014), leukemia/lymphoma (Piovan *et al.*, 2013; Della Gatta *et al.*, 2012; Compagno *et al.*, 2009), and prostate cancer (Ergun *et al.*, 2007; Aytes *et al.*, 2014), as well as across multiple tumor types (Mendillo *et al.*, 2012; Vaske *et al.*, 2010; Csibi *et al.*, 2013), to the dissection of oncogenic tyrosine kinase signals (Yarden and Pines, 2012; Schoeberl *et al.*, 2009; Kirouac *et al.*, 2013; Albeck *et al.*, 2013; Tentner *et al.*, 2012; Wagner *et al.*, 2013), and the characterization of drug activity (Komurov *et al.*, 2012), off-target effect (Chang *et al.*, 2010), and synergy (Lee *et al.*, 2012; Nelander *et al.*, 2008; Bansal *et al.*, 2014; Piovan *et al.*, 2013).

Other classes of model-based approaches are also emerging that are not based on molecular interactions. For instance, several studies have highlighted the ability to model properties of the tumor, such as growth and response to

radiotherapy, in terms of competition of distinct cellular subpopulations (Marusyk *et al.*, 2014; Leder *et al.*, 2014; Haeno *et al.*, 2012).

Assembly and Analysis of Cancer Regulatory Models

As discussed, the availability of increasingly accurate, comprehensive, and cell-context-specific interactomes will be critical to the success of model-based systems biology. In particular, it has been shown that interactomes are highly predictive only when they are representative of the specific context of interest (Aytes *et al.*, 2014). This is because epigenetic reprogramming of the cell during lineage differentiation, combined with differential availability of cofactors, results in molecular interactions that are broadly dependent on tumor cell context. As a result, reverse engineering of tumor specific interactomes, whether using experimental, computational, or hybrid approaches, represents a defining challenge in cancer systems biology. Ultimately, however, we expect that these efforts will help transition from cancer pathway based (Wang *et al.*, 2010) to regulatory network-based discovery (Califano *et al.*, 2012).

Cancer pathways represent a coordinated effort to 'linearize' molecular interactions to make them intuitively accessible to the biologist, via a relatively simple graphical representation. Such an effort, however, is ultimately doomed because the true complexity of biological interactions – including hundreds of thousands of signaling, transcriptional, protein-complex, small noncoding RNA, and metabolic interactions – simply does not lend itself to such a simplistic representation. In addition, cancer pathways are presented as intrinsically universal structures rather than as tumor specific ones. For instance, in its canonical representation (see Figure 4(a)), the

epidermal growth factor receptor (EGFR) pathway would appear as virtually identical in all tumor contexts, which is clearly not the case. Not surprisingly, a prior hypothesis being absent, pathway-based approaches have been largely unsuccessful in the elucidation of novel tumor-related mechanisms.

To better appreciate the difference, compare Figure 4(a), showing differential phosphorylation of canonical EGFR pathway proteins in the H1650 cell line, where EGFR harbors an activating mutation, versus EGFR wild type cell lines to Figure 4(b), showing the same differential phosphorylation pattern on a signaling network, inferred *de novo* from phosphopeptide profile data of a large collection of lung adenocarcinoma samples (Rikova *et al.*, 2007). While the pathway-based representation provides no clue that the EGFR pathway may be dysregulated, the network-based representation shows hyper-phosphorylation of virtually all EGFR, EPHA2, and TNK2 predicted substrates, with EGFR predicted as the single most aberrantly signaling tyrosine kinase, based on statistical significance.

In cancer, systems biology approaches have now evolved to include the following 3 steps. First, acquisition of complementary molecular profiles is performed, including transcriptomic, mutational, epigenetic, proteomic, and metabolomic modalities. This data is then used to reconstruct associated interactome models. Finally, these are interrogated using genetic and genomic signatures representing the cellular states of interest. We will discuss the two latter steps in more detail.

Interactome Reverse Engineering

Cell behavior is determined by the processing of endogenous and exogenous signals by the machinery that is in charge of maintaining cellular homeostasis and to make critical decisions

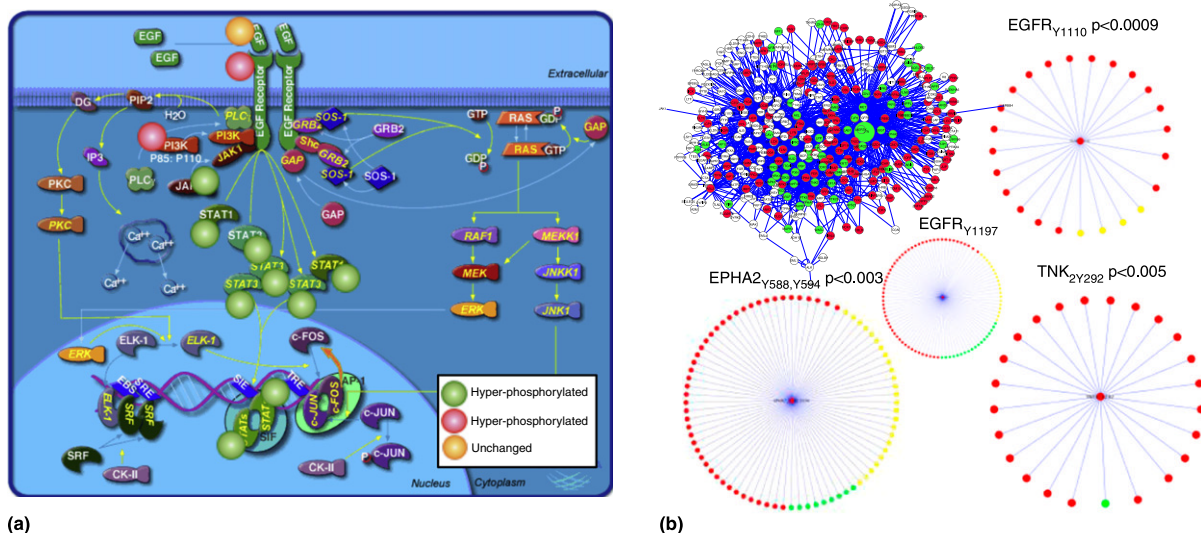


Figure 4 (a) Pathway-based representation of differential EGFR pathway protein phosphorylation in H1650 cells, harboring an EGFR activating mutation. Proteins tagged with a red circle are hyper-phosphorylated, those tagged with a green circle are hypo-phosphorylated, and those with an orange circle are unchanged. Phosphopeptide abundance for the remaining proteins was not detected. (b) Network-based representation of differential protein phosphorylation in H1650 cells. The signaling network was inferred *de novo* from a large phospho-proteomic dataset (Rikova *et al.*, 2007) for lung adenocarcinoma. Proteins are shown as either Red (hyper-phosphorylated), green (hypo-phosphorylated), and yellow (unchanged). Individual substrates associated with EGFR, EPHA2, and TNK2 are shown, as the three most enriched in phosphorylated substrates. Substrates are associated to phosphorylation of specific tyrosines on the corresponding kinases (e.g., Y1197 on EGFR).

in the presence of specific signals, for example, activating apoptosis when p53-mediated DNA damage repair is unsuccessful or senescence when oncogenic KRAS signals are detected. This machinery is implemented through a complex network of molecular interaction, i.e., the interactome, comprising several cross-interacting layers. Among the principal categories, these represent transcriptional, posttranscriptional, signal transduction, protein-complex, and metabolic interactions.

Disruption of interactome topology or dynamics by genomic events, either within a layer or across layers, can reprogram the cell by activating aberrant programs and ultimately leading to tumorigenesis. For instance, abrogating the regulatory program controlling expression of the MYC proto-oncogene by translocation is a key event in lymphomagenesis, while overexpression of the estrogen receptor by gene duplication events constitutes a key event in breast cancer tumorigenesis. Cancer systems biology is predicated on the simple premise that the logic of regulatory models can be ‘reverse engineered’ from experimental data – from simple kinetic models describing a handful of genes to probabilistic models of genome wide regulation – to help elucidate these aberrant events.

Over the past decade, multiple strategies have been developed for the *de novo* reverse engineering of regulatory models of mammalian cells, even though the first methods were driven for the study of yeast and bacteria, as simpler model organisms (Friedman, 2004; Gardner *et al.*, 2003; Tavazoie *et al.*, 1999). A key advantage, in the latter, is that genomic regulatory regions are relatively short, allowing the efficient use of sequence information to infer the cell regulatory logic. For instance, yeast promoter regions are 600 bp on average, thus allowing straightforward analysis. In comparison, distal regulatory elements in human genes can be hundreds of kilobases away from transcription start sites. In addition, gene regulation in higher eukaryotes is made even more complex by alternative splice variants, alternative start sites, and multiple poly-A tails.

Fortunately, despite these challenges, data availability and more sophisticated computational algorithms have significantly improved our ability to reverse engineer mammalian regulatory models more accurately and comprehensively. These methodologies can be roughly divided into the four categories discussed below.

Optimization Driven Machine Learning Approaches

Due to the multivariate nature of cellular regulation and to the relatively small number of distinct molecular profiles available in tumor repositories, classical machine learning based methods, such as maximum likelihood (ML), cannot be effectively applied to inferring causal regulatory relationships. However, several heuristics, such as maximum parsimony, have improved their applicability (Larranaga *et al.*, 2006). A standard approach in ML is to ask ‘What is the regulatory model that maximizes the posterior probability of having produced the observed molecular profile data?’ This cannot be addressed by brute-force enumeration of all possible models. As a result, greedy algorithms and associated approximations have been developed, such as assuming that regulatory models can be represented as directed acyclic graphs (DAG), without feedback loops (Friedman, 2004). The highest probability

model can then be used to predict behavior based on independent data (Larranaga *et al.*, 2006). Examples of such methods include the analysis of regulators of gene expression modules (Segal *et al.*, 2003), as well as the use of Bayesian and Dynamic Bayesian Networks for reverse engineering transcriptional (Friedman, 2004; Hartemink, 2005) and signal transduction networks (Sachs *et al.*, 2005). For a review of these methods, see (Andersson *et al.*, 2011; Ezkurdia *et al.*, 2009; De Vries and Bonvin, 2008). Factors that affect the precision of the predictions by ML approaches include the dataset quality, feature preselection for single residues, and algorithm selection based on the purpose and data type (Franke *et al.*, 2012).

Integration of Prior Knowledge and Experimental Evidence

Rather than predicting interaction from a single data modality, for example, from gene expression profile data, systems biologists have embraced the vast number of repositories containing experimental data from high-throughput approaches. These range from gene expression profile, to genome wide Chromatin Immunoprecipitation data (GW-ChIP), to yeast-two-hybrids and nuclear pull down assays. Albeit partial and often inaccurate in isolation, the knowledge contained in these repositories can be effectively integrated into a single unified model, using computational approaches to combine the probability of a specific molecular interaction. For instance, prediction of transcriptional interactions may integrate data from GW-ChIP, DNA binding site motif analysis, and co-expression, among others. The use of machine learning frameworks for the integration of multiple weak clues, from Naïve Bayes Classifiers (Lefebvre *et al.*, 2010), to Bayesian Networks (Jansen *et al.*, 2003), to a variety of consensus scoring methods, has been very successful in generating more accurate and comprehensive molecular interaction models (42, 43). Recently, intriguing observations have been published by the Dialog on Reverse Engineering Assessment Methods (DREAM) consortium (Marbach *et al.*, 2012). DREAM is an attempt to objectively measure the ability of computational approaches to correctly infer facts about regulatory network structure. Specifically, it was shown that integration of results from multiple inference algorithms performs better or at least as well as the best individual algorithm. This is an important result, as we often do not have a principled approach to objectively assess the quality of each given method and may instead want to use the integrative results of several of them. An additional value of integrative methods is that they allow seamless integration of heterogeneous data. For instance, it was recently shown that protein structure information, from X-ray and NMR crystallography, can be effectively integrated with functional data to accurately predict protein–protein interactions (Zhang *et al.*, 2012). For a more comprehensive review of integrative approaches, see (Bebek *et al.*, 2012).

Regression Analysis

Regression techniques have long been used to estimate parameters for kinetic models from experimental data and could,

at least in theory, be extended to the inference parameters for entire regulatory models. Various regression methods have been proposed for pathways or network inference, including ML (Lyles *et al.*, 2012; Knott, 2005), least squares (LS) (Li *et al.*, 2011, 2012; Li and Gui, 2004), and Bayesian Inference (BI) (Sarder *et al.*, 2010; Jin *et al.*, 2012; Zhang *et al.*, 2011), to obtain estimates of model parameters (Jaqaman and Danuser, 2006). ML approaches infer parameter values from a distribution, as those maximizing the posterior probability of the experimental data; LS approaches determine the parameter values that minimize the sum of the squares of the residuals, i.e., the difference between each experimental and model-inferred datapoint; finally, BI methods use a priori models for the unknown parameter distribution to compute their most likely values. A key problem of regression methods is that they are generally under-determined, i.e., they have more parameters whose value must be inferred than independent observations. When this happens, an infinite number of parameter combinations produce an identical solution, thus requiring additional heuristics for selecting a specific set of values (e.g., equivalent to fitting a line through a single point). This is a common issue in real biological systems, as the number of parameters for a system with tens of thousands of interacting molecular species can be in the several hundred thousands but few datasets with more than a few hundred independent experimental profiles are available for the same system. To address this problem a number of dimensionality reduction approaches have been developed, which work by either splitting a single high-dimensional problem into a number of independent lower-dimensional ones, for example, via singular value decomposition (Yeung *et al.*, 2002), or by penalizing models with larger number of connections via sparsity constraints (Yeung *et al.*, 2002).

Information Theory and Probabilistic Methods

Shannon's information theory provides a useful probabilistic framework to study the flow of information in a complex system (Shannon, 1948). In recent years, information theory has become a staple of systems biology approaches (Waltermann and Klipp, 2011), for instance by predicting the minimal machinery in the cell necessary to account for the globally observed information transfer between distinct molecular species characterized by a profile, such as mRNA (Basso *et al.*, 2005; Wang *et al.*, 2009b; Goodarzi *et al.*, 2012), mutational data (Drees *et al.*, 2005), and microRNAs (Elemento *et al.*, 2007; Sumazin *et al.*, 2011). In general, these approaches try to assess the minimal set of interactions that must be in place to justify the observed flow of information, including either 2-way information, i.e., mutual information (MI) (Steuer *et al.*, 2002), or higher-order multivariate information, for example, conditional mutual information (CMI), multi-information, etc. (Wang *et al.*, 2006), thus allowing dissection of signaling or transcriptional interactions. MI measures the mutual dependence of two random variables, without any linearity assumptions (e.g., between the mRNA of a transcription factor and one of its potential targets). CMI measures whether the information transfer between two variables is dependent on a third variable, for instance the availability of a

protein kinase affecting the ability of a transcription factor to regulate its targets. Various probabilistic and information theoretic methods have been proposed for the reverse engineering of regulatory networks, some of which have been extensively validated, such as ARACNE (Margolin *et al.*, 2006; Basso *et al.*, 2005), MINDy (Wang *et al.*, 2009b), Hermes (Sumazin *et al.*, 2011), minet (Meyer *et al.*, 2008), and CLR (Faith *et al.*, 2007), among others.

ARACNe and MINDy, a Case Study

These two algorithms are among those that have been most extensively validated, allowing reverse engineering of networks that have led to a number of novel discoveries in cancer biology. We will thus use them as a self-contained case study to illustrate several relevant concepts in the area of reverse engineering. Other examples of validated algorithms with applications to the study of cancer and other disease include CONEXIC (Akavia *et al.*, 2010), CLR (Kwong *et al.*, 2012), and Bayesian Networks (Yang *et al.*, 2009; Sachs *et al.*, 2005), among others, see (Pe'er and Hachohen, 2011; Califano *et al.*, 2012) for a more detailed review.

ARACNe (Algorithm for the Reconstruction of Accurate Cellular Networks) has been widely applied to dissect transcriptional regulatory networks from gene expression profiles of multiple cancer subtypes, including B cell lymphoma (Basso *et al.*, 2005, 2010), glioma (Carro *et al.*, 2010; Sonabend *et al.*, 2014; Chudnovsky *et al.*, 2014; Zhao *et al.*, 2009), prostate cancer (Aytes *et al.*, 2014), T-ALL (Piovan *et al.*, 2013; Della Gatta *et al.*, 2012; De Keersmaecker *et al.*, 2010; Palomero *et al.*, 2006), and breast cancer (Lim *et al.*, 2009). ARACNE first hypothesizes candidate interactions between a transcription factor and a target gene ($TF \rightarrow T$) whose MI is statistically significant, when computed from their gene expression across a large number of experiments. Finally, candidate interactions that violate the Data Processing Inequality (DPI) are removed. The DPI states that information transferred via direct interactions always exceeds information transferred via indirect paths. Thus, if any indirect path ($TF \rightarrow TF_x \rightarrow T$) exists, with MI greater than the candidate direct interaction ($TF \rightarrow T$), then the direct interaction is removed. ARACNE was experimentally validated by ChIP and GW-ChIP assays, as well as by silencing of transcription factors followed by gene expression profiling, see above referenced manuscripts.

This type of analysis was extended to higher-order interactions. Specifically, the Modulator Inference by Network Dynamics (MINDy) algorithm (Wang *et al.*, 2009b), most recently refined to its true theoretical implementation (Giorgi *et al.*, 2014), was developed to identify posttranslational modulators of transcription factor activity on their targets using the conditional MI (Wang *et al.*, 2009a,b). These include both co-transcription factors and upstream signaling proteins regulating the transcription factor activity (Wang *et al.*, 2009b). Genome wide search based on this methodology were successful in identifying the HUWE1 ubiquitin ligase as a key modulator of MYCN in neural stem cells (Zhao *et al.*, 2009), the serine threonine kinase STK38 as a key modulator of MYC in human B cells (Bisikirska *et al.*, 2013), and the ubiquitin ligase adapter protein KLHL9 as a modulator of CEBPB and

CEBPD in the mesenchymal subtype of glioma (Chen *et al.*, 2014), among several other interactions that were experimentally validated.

Benchmarking Predictive Models

As exemplified in the previous section, the accuracy of interactome reverse engineering can be validated with careful testing of its predictions. However, the fast and relentless progress of technology to probe larger amounts of data and deeper layers of cellular organization create the need for an entire spectrum of algorithms to extract actionable information from the data. Predictions resulting from these methods, however, are infrequently validated experimentally and their objective value remains elusive. Thus systematic validation is important not only for interactome inference but also for virtually every other area of computational systems biology. Take, for example, the case of identification of somatic mutations in cancer biopsies, which is one of the valuable pieces of information needed to implement precision medicine. Recent studies have shown that when different researchers apply their algorithms to the same genomic data their results may be significantly discordant (Kim and Speed, 2013; O'Rawe *et al.*, 2013; Alkan *et al.*, 2011). Similar problems are found when comparing methods to reverse engineer gene regulatory networks (Marbach *et al.*, 2010) and signaling networks (Prill *et al.*, 2011).

The need to determine the accuracy of novel algorithms through an objective evaluation process, as well as to compare the relative merits of disparate methods designed to perform the same tasks has given rise to a relatively new modality of method assessment in the form of scientific competitions or Challenges (Costello and Stolovitzky, 2013). While there are several such efforts in different areas of bioinformatics (Boutros *et al.*, 2014) there are only a few such efforts in the field of systems biology, notably CAGI (Critical Assessment of Genome Interpretation) and DREAM (Dialog for Reverse Engineering Assessment and Methods) (Stolovitzky *et al.*, 2007). Typically these Challenges pose specific questions (e.g., inference of biological networks (Stolovitzky *et al.*, 2009), computational prediction of drug sensitivity (Costello *et al.*, 2014), and prediction of patient survival (Margolin *et al.*, 2013), among many others, which require integrative/systems analysis of vast amounts of data. Challenge participants train their algorithms from a provided dataset (training set), and make predictions that are blindly compared against an independent dataset they have no access to (testing set), used as a gold standard for validation. Using the latter, Challenge organizers can evaluate the methods while making a fair comparison of their strengths and weaknesses. From the perspective of the participants, the experience is not unlike real-world problems, in which a method developer makes predictions and seeks experimental validation in novel scenarios.

Besides the merits of benchmarking and method comparison, DREAM Challenges also aim at promoting research in novel areas (e.g., computational predictions of drug interactions in cancer cells (Bansal *et al.*, 2014)), accelerating discovery by crowdsourcing, assessing the computational tractability of biological problems, creating collaborative

communities, and, as discussed earlier, tapping on the wisdom of crowds as a robust solution for biological predictions (Marbach *et al.*, 2012).

Interrogating Pathways and Networks

Once regulatory models, such as pathways and networks are available, machine learning approaches can harness them to study cancer phenotypes. For example, a random forests algorithm was used to perform pathway-based single nucleotide polymorphism (SNP) analysis to predict cancer patient survival outcome (Pang *et al.*, 2011). Applied to multiple myeloma, using Affymetrix 500 K SNP array data, the method connected informative SNPs and patient survival, thus identifying candidate pathway involved in modulating patient survival and associated genetic variants (Pang *et al.*, 2011). A multitude of other machine learning algorithms have also been used to assist in GWAS analysis, such as SNP discovery (Wegrzyn *et al.*, 2009; Matukumalli *et al.*, 2006), SNP interaction (Ueki and Tamiya, 2012; Cheng *et al.*, 2012), cancer subtype classification (Wong *et al.*, 2012), survival outcome association (Van Ness *et al.*, 2008), and key genetic variant identification (Han *et al.*, 2010; Yu *et al.*, 2007), see (Szymczak *et al.*, 2009) for an more comprehensive review. Yet, a common characteristic of these approaches is that the associations they identify are statistical in nature and do not provide insight into mechanism. As a result, elucidation of the mechanism by which a somatic or germline variant regulates the corresponding tumor phenotype often substantially lags after its discovery by statistical analyses.

Several approaches are emerging to simultaneously address the reconstruction of context specific cell regulatory models and their interrogation for the identification of drivers of both physiologic and pathologic phenotypes, see (Califano *et al.*, 2012; Pe'er and Hacohen, 2011) for a more comprehensive review. For instance, the Master Regulator Inference algorithm (MARINA), see Figure 5, was used to computationally discover genes and associated mechanisms that are causally related to the implementation of a specific phenotype (Lefebvre *et al.*, 2010; Carro *et al.*, 2010; Lim *et al.*, 2009; Chudnovsky *et al.*, 2014; Aytes *et al.*, 2014), followed by their experimental validation. Rather than asking which genes are differentially expressed in a cancer phenotype, MARINA uses analyzes interactomes inferred by ARACNe or by other methods to identify candidate master regulator (MR) genes, whose transcriptional targets co-segregate with differentially expressed genes. Experimental validation assays are then used to identify those that are both necessary and sufficient for phenotype presentation. This approach was used, for instance, to elucidate CEBPB/CEBPD and STAT3 as synergistic master regulators of the mesenchymal subtype of human glioblastoma (Carro *et al.*, 2010), FOXM1 and MYB as MRs of the germinal center reaction (Lefebvre *et al.*, 2010), FOXM1 and CENPF as synergistic master regulators of the most aggressive subtype of prostate cancer (Aytes *et al.*, 2014), and ZFX4 as a master regulator of the NuRD complex in gliomagenesis (Chudnovsky *et al.*, 2014).

Similarly, different integrative network-based systems biology approaches were used to identify drivers of tumor initiation (Akavia *et al.*, 2010) and differential patient survival (Kwong *et al.*, 2012) in melanoma and glioma (Chen *et al.*, 2014).

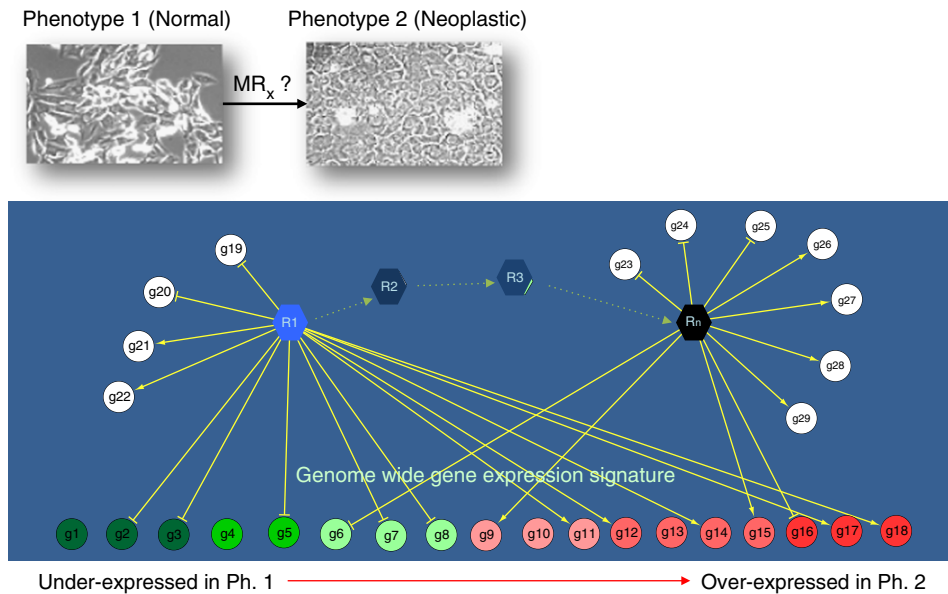


Figure 5 Discovery of Master Regulator Genes using the MARINA algorithm. Master Regulators (e.g., R_1) are genes whose activated and repressed transcriptional targets are statistically significantly enriched in over and under expressed genes in the transition, respectively. Other regulators (e.g., R_n) have targets that are not enriched in differentially expressed genes.

While these constitute very recent developments, the availability of methods to assemble accurate and cancer specific regulatory models will determine an explosion in these kind of approaches.

Recent Trends and Future Perspective

As discussed above, a variety of research strategies have emerged in cancer systems biology since the discipline's conceptual inception, about a decade ago (Kitano, 2002; Ideker *et al.*, 2001). While the initial forays in this area were mostly of a theoretical nature, over the last 5 years these approaches have achieved significant success and gained considerable traction, especially in the elucidation of complex, multi-gene mechanisms that would not have been approachable using conventional molecular biology strategies. In its coming of age, systems biology has become one of the most powerful methods for the high-throughput generation of testable hypothesis, frequently achieving a greater than 50% likelihood of experimental validation, something that would have been previously unthinkable from purely computational approaches. For instance, validation of transcription factor targets, microRNA targets, and protein-protein interactions predicted by reverse engineering methods is frequently in the 80% range, thus competitive with high-throughput experimental approaches (Zhang *et al.*, 2012; Sumazin *et al.*, 2011; Wang *et al.*, 2009b; Basso *et al.*, 2005). Similarly, a number of disease mechanism have been recently elucidated, following validation of regulatory network-based predictions (Califano *et al.*, 2011; Yang *et al.*, 2009; Schadt *et al.*, 2005; Della Gatta *et al.*, 2012; Bisikirska *et al.*, 2013; Lefebvre *et al.*, 2010; De Keersmaecker *et al.*, 2010; Carro *et al.*, 2010; Zhao *et al.*, 2009; Pe'er and Hacohen, 2011; Konig *et al.*, 2010; Mendillo *et al.*, 2012; Kwong *et al.*, 2012; Komurov *et al.*, 2012). These are only a few representative examples from a now rapidly growing literature.

In spite of these successes, concerns and skepticism is still pervasive in the biological community on the role and effectiveness of systems biology approaches in dissecting complex diseases, such as cancer, and in elucidating novel therapeutic strategies for translational research. To address these concerns, this nascent discipline requires progress on two critical aspects: first, novel technologies that produce experimentally validated, biologically relevant results are critically needed; more importantly, translation of concepts emerging from systems approaches must lead to clinical applications.

While the genomic framework for the study of biology has now been around for over 20 years, its translational applications have only started to emerge in earnest over the last 5 years, largely as a result of the large amount of genomic information accumulated following completion of the human genome map. Systems biology and its applications to cancer are much more recent. As such, it would be unreasonable to expect immediate translation of its discoveries. Similar to genomics, this discipline will not achieve its true potential until appropriate datasets and technologies become available. The first attempts to translate cancer systems biology discoveries are only now starting to be tested in the clinic and several more years may be necessary before a compelling argument for clinical applications of systems biology may be made.

See also: Vertical Integration: Applications: Angiogenesis

References

- Arkav, U.D., Litvin, O., Kim, J., *et al.*, 2010. An integrated approach to uncover drivers of cancer. *Cell* 143, 1005–1017.

- Albeck, J.G., Mills, G.B., Brugge, J.S., 2013. Frequency-modulated pulses of ERK activity transmit quantitative proliferation signals. *Molecular Cell* 49, 249–261.
- Alizadeh, A.A., Eisen, M.B., Davis, R.E., *et al.*, 2000. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403, 503–511.
- Alkan, C., Coe, B.P., Eichler, E.E., 2011. Genome structural variation discovery and genotyping. *Nature Reviews Genetics* 12, 363–376.
- Andersson, C.R., Gustafsson, M.G., Strombergsson, H., 2011. Quantitative chemogenomics: Machine-learning models of protein-ligand interaction. *Current Topics in Medicinal Chemistry* 11, 1978–1993.
- Aytes, A., Mitrofanova, A., Lefebvre, C., *et al.*, 2014. Cross-species analysis of genome-wide regulatory networks identifies a synergistic interaction between FOXM1 and CENPF that drives prostate cancer malignancy. *Cancer Cell* 25, 638–651.
- Bansal, M., Yang, J., Karan, C., *et al.*, 2014. A community computational challenge to predict the activity of pairs of compounds. *Nature Biotechnology* 32 (12), 1213–1222.
- Barretina, J., Caponigro, G., Stransky, N., *et al.*, 2012. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483, 603–607.
- Basso, K., Margolin, A.A., Stolovitzky, G., *et al.*, 2005. Reverse engineering of regulatory networks in human B cells. *Nature Genetics* 37, 382–390.
- Basso, K., Saito, M., Sumazin, P., *et al.*, 2010. Integrated biochemical and computational approach identifies BCL6 direct target genes controlling multiple pathways in normal germinal center B cells. *Blood* 115, 975–984.
- Bebek, G., Koyuturk, M., Price, N.D., Chance, M.R., 2012. Network biology methods integrating biological data for translational science. *Briefings in Bioinformatics* 13, 446–459.
- Bisikirska, B.C., Adam, S.J., Alvarez, M.J., *et al.*, 2013. STK38 is a critical upstream regulator of MYC's oncogenic activity in human B-cell lymphoma. *Oncogene* 32 (45), 5283–5291.
- Boutros, P.C., Margolin, A.A., Stuart, J.M., Califano, A., Stolovitzky, G., 2014. Toward better benchmarking: Challenge-based methods assessment in cancer genomics. *Genome Biology* 15, 462.
- Breiman, L., 2001. Random forests. *Machine Learning* 45, 5–32.
- Brown, B.M., 1975. A method for combining non-independent, one-sided tests of significance. *Biometrics* 31, 987–992.
- Califano, A., Butte, A., Friend, S., Ideker, T., Schadt, E., 2011. Integrative network-based association studies: Leveraging cell regulatory models in the post-GWAS era. *Nature Precedings*. Available at: <http://dx.doi.org/10.1038/npre.2011.5732.1> (accessed 19.03.15).
- Califano, A., Butte, A.J., Friend, S., Ideker, T., Schadt, E., 2012. Leveraging models of cell regulation and GWAS data in integrative network-based association studies. *Nature Genetics* 44, 841–847.
- Carro, M.S., Lim, W.K., Alvarez, M.J., *et al.*, 2010. The transcriptional network for mesenchymal transformation of brain tumours. *Nature* 463, 318–325.
- Chang, R.L., Xie, L., Xie, L., Bourne, P.E., Palsson, B.O., 2010. Drug off-target effects predicted using structural analysis in the context of a metabolic network model. *PLoS Computational Biology* 6, e1000938.
- Chen, J.C., Alvarez, M.J., Talos, F., *et al.*, 2014. Identification of causal genetic drivers of human disease through systems-level analysis of regulatory networks. *Cell* 159 (2), 402–414. doi:10.1016/j.cell.2014.09.021.
- Cheng, Y.K., Beroukhi, R., Levine, R.L., *et al.*, 2012. A mathematical methodology for determining the temporal order of pathway alterations arising during gliomagenesis. *PLoS Computational Biology* 8, e1002337.
- Chudnovsky, Y., Kim, D., Zheng, S., *et al.*, 2014. ZFX4 interacts with the NuRD core member CHD4 and regulates the glioblastoma tumor-initiating cell state. *Cell Reports* 6, 313–324.
- Compagno, M., Lim, W.K., Grunn, A., *et al.*, 2009. Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. *Nature* 459, 717–721.
- Costello, J.C., Heiser, L.M., Georgii, E., *et al.*, 2014. A community effort to assess and improve drug sensitivity prediction algorithms. *Nature Biotechnology* 32, 1202–1212. doi:10.1038/nbt.2877.
- Costello, J.C., Stolovitzky, G., 2013. Seeking the wisdom of crowds through challenge-based competitions in biomedical research. *Clinical Pharmacology and Therapeutics* 93, 396–398.
- Csibi, A., Fendt, S.M., Li, C., *et al.*, 2013. The mTORC1 pathway stimulates glutamine metabolism and cell proliferation by repressing SIRT4. *Cell* 153, 840–854.
- Della Gatta, G., Palomero, T., Perez-Garcia, A., *et al.*, 2012. Reverse engineering of TLX oncogenic transcriptional networks identifies RUNX1 as tumor suppressor in T-ALL. *Nature Medicine* 18, 436–440.
- De Keersmaecker, K., Real, P.J., Gatta, G.D., *et al.*, 2010. The TLX1 oncogene drives aneuploidy in T cell transformation. *Nature Medicine* 16, 1321–1327.
- De Vries, S.J., Bonvin, A.M., 2008. How proteins get in touch: Interface prediction in the study of biomolecular complexes. *Current Protein and Peptide Science* 9, 394–406.
- Ding, L., Ellis, M.J., Li, S., *et al.*, 2010. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature* 464, 999–1005.
- Domingos, P., Pazzani, M., 1997. On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning* 29, 103–137.
- Drees, B.L., Thorsson, V., Carter, G.W., *et al.*, 2005. Derivation of genetic interaction networks from quantitative phenotype data. *Genome Biology* 6, R38.
- Elemento, O., Slonim, N., Tavazoie, S., 2007. A universal framework for regulatory element discovery across all genomes and data types. *Molecular Cell* 28, 337–350.
- Ergun, A., Lawrence, C.A., Kohanski, M.A., Brennan, T.A., Collins, J.J., 2007. A network biology approach to prostate cancer. *Molecular Systems Biology* 3, 82.
- Ezkurdia, I., Bartoli, L., Fariselli, P., *et al.*, 2009. Progress and challenges in predicting protein-protein interaction sites. *Briefings in Bioinformatics* 10, 233–246.
- Faith, J.J., Hayete, B., Thaden, J.T., *et al.*, 2007. Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biology* 5, e8.
- Fisher, R.A., 1925. *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd.
- Franke, V., Sikic, M., Vlahovicek, K., 2012. Prediction of interacting protein residues using sequence and structure data. *Methods in Molecular Biology* 819, 233–251.
- Fratini, V., Trifonov, V., Chan, J.M., *et al.*, 2013. The integrated landscape of driver genomic alterations in glioblastoma. *Nature Genetics* 45, 1141–1149.
- Friedman, N., 2004. Inferring cellular networks using probabilistic graphical models. *Science* 303, 799–805.
- Gardner, T.S., Di Bernardo, D., Lorenz, D., Collins, J.J., 2003. Inferring genetic networks and identifying compound mode of action via expression profiling. *Science* 301, 102–105.
- Giorgi, F., Lopez-Garcia, G., Woo, J.H., *et al.*, 2014. Inferring protein modulation from gene expression data using conditional mutual information. *PLoS One* 9 (10), e109569.
- Goodarzi, H., Najafabadi, H.S., Oikonomou, P., *et al.*, 2012. Systematic discovery of structural elements governing stability of mammalian messenger RNAs. *Nature* 485, 264–268.
- Haeno, H., Gonen, M., Davis, M.B., *et al.*, 2012. Computational modeling of pancreatic cancer reveals kinetics of metastasis suggesting optimum treatment strategies. *Cell* 148, 362–375.
- Han, B., Park, M., Chen, X.W., 2010. A Markov blanket-based method for detecting causal SNPs in GWAS. *BMC Bioinformatics* 11 (Suppl. 3), S5.
- Hartemink, A.J., 2005. Reverse engineering gene regulatory networks. *Nature Biotechnology* 23, 554–555.
- Hudson, T.J., Anderson, W., Artez, A., *et al.*, 2010. International network of cancer genome projects. *Nature* 464, 993–998.
- Ideker, T., Galitski, T., Hood, L., 2001. A new approach to decoding life: Systems biology. *Annual Review of Genomics and Human Genetics* 2, 343–372.
- Jansen, R., Yu, H., Greenbaum, D., *et al.*, 2003. A Bayesian networks approach for predicting protein-protein interactions from genomic data. *Science* 302, 449–453.
- Jaqaman, K., Danuser, G., 2006. Linking data to models: Data regression. *Nature Reviews Molecular Cell Biology* 7, 813–819.
- Jin, G., Fu, C., Zhao, H., *et al.*, 2012. A novel method of transcriptional response analysis to facilitate drug repositioning for cancer therapy. *Cancer Research* 72, 33–44.
- Kim, S.Y., Speed, T.P., 2013. Comparing somatic mutation-callers: Beyond Venn diagrams. *BMC Bioinformatics* 14, 189.
- King, J.C., Xu, J., Wongvipat, J., *et al.*, 2009. Cooperativity of TMPRSS2-ERG with PI3-kinase pathway activation in prostate oncogenesis. *Nature Genetics* 41, 524–526.
- Kirouac, D.C., Du, J.Y., Lahdenranta, J., *et al.*, 2013. Computational modeling of ERBB2-amplified breast cancer identifies combined ErbB2/3 blockade as superior to the combination of MEK and AKT inhibitors. *Science Signaling* 6, ra68.
- Kitano, H., 2002. Computational systems biology. *Nature* 420, 206–210.
- Klein, U., Tu, Y., Stolovitzky, G.A., *et al.*, 2003. Transcriptional analysis of the B cell germinal center reaction. *Proceedings of the National Academy of Sciences of the United States of America* 100, 2639–2644.
- Knott, S.A., 2005. Regression-based quantitative trait loci mapping: Robust, efficient and effective. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360, 1435–1442.
- Komurov, K., Tseng, J.T., Muller, M., *et al.*, 2012. The glucose-deprivation network counteracts lapatinib-induced toxicity in resistant ErbB2-positive breast cancer cells. *Molecular Systems Biology* 8, 596.

- Konig, R., Stertz, S., Zhou, Y., *et al.*, 2010. Human host factors required for influenza virus replication. *Nature* 463, 813–817.
- Kwong, L.N., Costello, J.C., Liu, H., *et al.*, 2012. Oncogenic NRAS signaling differentially regulates survival and proliferation in melanoma. *Nature Medicine* 18, 1503–1510.
- Lamb, J., Crawford, E.D., Peck, D., *et al.*, 2006. The connectivity map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313, 1929–1935.
- Larranaga, P., Calvo, B., Santana, R., *et al.*, 2006. Machine learning in bioinformatics. *Briefings in Bioinformatics* 7, 86–112.
- Leder, K., Pitter, K., Laplant, Q., *et al.*, 2014. Mathematical modeling of PDGF-driven glioblastoma reveals optimized radiation dosing schedules. *Cell* 156, 603–616.
- Lee, M.J., Ye, A.S., Gardino, A.K., *et al.*, 2012. Sequential application of anticancer drugs enhances cell death by rewiring apoptotic signaling networks. *Cell* 149, 780–794.
- Lefebvre, C., Rajbhandari, P., Alvarez, M.J., *et al.*, 2010. A human B-cell interactome identifies MYB and FOXM1 as master regulators of proliferation in germinal centers. *Molecular Systems Biology* 6, 377.
- Li, H., Gui, J., 2004. Partial Cox regression analysis for high-dimensional microarray gene expression data. *Bioinformatics* 20 (Suppl. 1), i208–i215.
- Li, X., Gill, R., Cooper, N.G., Yoo, J.K., Datta, S., 2011. Modeling microRNA-mRNA interactions using PLS regression in human colon cancer. *BMC Medical Genomics* 4, 44.
- Li, X., Yang, T., Li, S., 2012. Discrimination of serum Raman spectroscopy between normal and colorectal cancer using selected parameters and regression-discriminant analysis. *Applied Optics* 51, 5038–5043.
- Lim, W.K., Lyashenko, E., Califano, A., 2009. Master regulators used as breast cancer metastasis classifier. *Pacific Symposium on Biocomputing* 2009, 504–515.
- Lossos, I.S., Alizadeh, A.A., Diehn, M., *et al.*, 2002. Transformation of follicular lymphoma to diffuse large-cell lymphoma: Alternative patterns with increased or decreased expression of c-myc and its regulated genes. *Proceedings of the National Academy of Sciences of the United States of America* 99, 8886–8891.
- Lyles, R.H., Tang, L., Lin, J., Zhang, Z., Mukherjee, B., 2012. Likelihood-based methods for regression analysis with binary exposure status assessed by pooling. *Statistics in Medicine* 31 (22), 2485–2497.
- Manolio, T.A., Collins, F.S., Cox, N.J., *et al.*, 2009. Finding the missing heritability of complex diseases. *Nature* 461, 747–753.
- Marbach, D., Costello, J.C., Kuffner, R., *et al.*, 2012. Wisdom of crowds for robust gene network inference. *Nature Methods* 9, 796–804.
- Marbach, D., Prill, R.J., Schaffer, T., *et al.*, 2010. Revealing strengths and weaknesses of methods for gene network inference. *Proceedings of the National Academy of Sciences of the United States of America* 107, 6286–6291.
- Margolin, A.A., Bilal, E., Huang, E., *et al.*, 2013. Systematic analysis of challenge-driven improvements in molecular prognostic models for breast cancer. *Science Translational Medicine* 5, 181re1.
- Margolin, A.A., Nemenman, I., Basso, K., *et al.*, 2006. ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* 7 (Suppl. 1), S7.
- Marusyk, A., Tabassum, D.P., Altrock, P.M., *et al.*, 2014. Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature* 514, 54–58.
- Matukumalli, L.K., Grefenstette, J.J., Hyten, D.L., *et al.*, 2006. Application of machine learning in SNP discovery. *BMC Bioinformatics* 7, 4.
- Mendillo, M.L., Santagata, S., Koeva, M., *et al.*, 2012. HSF1 drives a transcriptional program distinct from heat shock to support highly malignant human cancers. *Cell* 150, 549–562.
- Meyer, P.E., Lafitte, F., Bontempi, G., 2008. Minet: A R/Bioconductor package for inferring large transcriptional networks using mutual information. *BMC Bioinformatics* 9, 461.
- Mullighan, C.G., Su, X., Zhang, J., *et al.*, 2009. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *New England Journal of Medicine* 360, 470–480.
- Narayanan, M., Vetta, A., Schadt, E.E., Zhu, J., 2010. Simultaneous clustering of multiple gene expression and physical interaction datasets. *PLoS Computational Biology* 6, e1000742.
- Nelander, S., Wang, W., Nilsson, B., *et al.*, 2008. Models from experiments: Combinatorial drug perturbations of cancer cells. *Molecular Systems Biology* 4, 216.
- Network, C.G.A.R., 2008. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 455, 1061–1068.
- O'Rawe, J., Jiang, T., Sun, G., *et al.*, 2013. Low concordance of multiple variant-calling pipelines: Practical implications for exome and genome sequencing. *Genome Medicine* 5, 28.
- Palomero, T., Lim, W.K., Odom, D.T., *et al.*, 2006. NOTCH1 directly regulates c-MYC and activates a feed-forward-loop transcriptional network promoting leukemic cell growth. *Proceedings of the National Academy of Sciences of the United States of America* 103, 18261–18266.
- Pang, H., Hauser, M., Minvielle, S., 2011. Pathway-based identification of SNPs predictive of survival. *European Journal of Human Genetics* 19, 704–709.
- Pe'er, D., Hachohen, N., 2011. Principles and strategies for developing network models in cancer. *Cell* 144, 864–873.
- Pearl, J., 1985. Bayesian networks: A model of self-activated memory for evidential reasoning. *Proceedings of the 7th Conference of the Cognitive Science Society*, pp. 329–334. Irvine, CA: University of California.
- Phillips, H.S., Kharbanda, S., Chen, R., *et al.*, 2006. Molecular subclasses of high-grade glioma predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis. *Cancer Cell* 9, 157–173.
- Piovan, E., Yu, J., Tosello, V., *et al.*, 2013. Direct reversal of glucocorticoid resistance by AKT inhibition in acute lymphoblastic leukemia. *Cancer Cell* 24, 766–776.
- Prill, R.J., Saez-Rodriguez, J., Alexopoulos, L.G., Sorger, P.K., Stolovitzky, G., 2011. Crowdsourcing network inference: The DREAM predictive signaling network challenge. *Science Signaling* 4, mr7.
- Rikova, K., Guo, A., Zeng, Q., *et al.*, 2007. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell* 131, 1190–1203.
- Sachs, K., Perez, O., Pe'er, D., Lauffenburger, D.A., Nolan, G.P., 2005. Causal protein-signaling networks derived from multiparameter single-cell data. *Science* 308, 523–529.
- Sarder, P., Schierding, W., Cobb, J.P., Nehorai, A., 2010. Estimating sparse gene regulatory networks using a bayesian linear regression. *IEEE Transactions on NanoBioscience* 9, 121–131.
- Schadt, E.E., Lamb, J., Yang, X., *et al.*, 2005. An integrative genomics approach to infer causal associations between gene expression and disease. *Nature Genetics* 37, 710–717.
- Schadt, E.E., Monks, S.A., Drake, T.A., *et al.*, 2003. Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422, 297–302.
- Schoeberl, B., Pace, E.A., Fitzgerald, J.B., *et al.*, 2009. Therapeutically targeting ErbB3: A key node in ligand-induced activation of the ErbB receptor-P13K axis. *Science Signaling* 2, ra31.
- Segal, E., Shapira, M., Regev, A., *et al.*, 2003. Module networks: Identifying regulatory modules and their condition-specific regulators from gene expression data. *Nature Genetics* 34, 166–176.
- Shannon, C.E., 1948. A mathematical theory of communication. *Bell System Technical Journal* 27, 379–423.
- Sonabend, A.M., Bansal, M., Guarnieri, P., *et al.*, 2014. The transcriptional regulatory network of proneural glioma determines the genetic alterations selected during tumor progression. *Cancer Research* 74, 1440–1451.
- Stephens, P.J., Greenman, C.D., Fu, B., *et al.*, 2011. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 144, 27–40.
- Steuer, R., Kurths, J., Daub, C.O., Weise, J., Selbig, J., 2002. The mutual information: Detecting and evaluating dependencies between variables. *Bioinformatics* 18 (Suppl. 2), S231–S240.
- Stolovitzky, G., Monroe, D., Califano, A., 2007. Dialog on reverse-engineering assessment and methods: The DREAM of high-throughput pathway inference. *Annals of the New York Academy of Sciences* 1115, 1–22.
- Stolovitzky, G., Prill, R.J., Califano, A., 2009. Lessons from the DREAM2 challenges. *Annals of the New York Academy of Sciences* 1158, 159–195.
- Stouffer, S.A., Suchman, E.A., Devinney, L.C., Star, S.A., Jr, W.R.M., 1949. Adjustment during Army Life. *The American Soldier*. Princeton: Princeton University Press.
- Sumazin, P., Yang, X., Chiu, H.S., *et al.*, 2011. An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma. *Cell* 147, 370–381.
- Szymczak, S., Biernacka, J.M., Cordell, H.J., *et al.*, 2009. Machine learning in genome-wide association studies. *Genetic Epidemiology* 33 (Suppl. 1), S51–S57.
- Tavazoie, S., Hughes, J.D., Campbell, M.J., Cho, R.J., Church, G.M., 1999. Systematic determination of genetic network architecture. *Nature Genetics* 22, 281–285.
- Taylor, B.S., Schultz, N., Hieronymus, H., *et al.*, 2010. Integrative genomic profiling of human prostate cancer. *Cancer Cell* 18, 11–22.
- Tentner, A.R., Lee, M.J., Ostheimer, G.J., *et al.*, 2012. Combined experimental and computational analysis of DNA damage signaling reveals context-dependent roles for Erk in apoptosis and G1/S arrest after genotoxic stress. *Molecular Systems Biology* 8, 568.
- Ueki, M., Tamiya, G., 2012. Ultrahigh-dimensional variable selection method for whole-genome gene-gene interaction analysis. *BMC Bioinformatics* 13, 72.

- Van Ness, B., Ramos, C., Haznadar, M., *et al.*, 2008. Genomic variation in myeloma: Design, content, and initial application of the Bank On A Cure SNP Panel to detect associations with progression-free survival. *BMC Medicine* 6, 26.
- Vaske, C.J., Benz, S.C., Sanborn, J.Z., *et al.*, 2010. Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics* 26, i237–i245.
- Verhaak, R.G., Hoadley, K.A., Purdom, E., *et al.*, Cancer Genome Atlas Research, 2010. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* 17 (1), 98–110.
- Wagner, J.P., Wolf-Yadlin, A., Sevecka, M., *et al.*, 2013. Receptor tyrosine kinases fall into distinct classes based on their inferred signaling networks. *Science Signaling* 6, ra58.
- Waltermann, C., Klipp, E., 2011. Information theory based approaches to cellular signaling. *Biochimica et Biophysica Acta* 1810, 924–932.
- Wang, K., Alvarez, M.J., Bisikirska, B.C., *et al.*, 2009a. Dissecting the interface between signaling and transcriptional regulation in human B cells. *Pacific Symposium on Biocomputing*. 264–275.
- Wang, K., Banerjee, N., Margolin, A.A., Nemenman, I., Califano, A., 2006. Genome-wide discovery of modulators of transcriptional interactions in human B lymphocytes. *RECOMB 2006 also Lecture Notes in Computer Science* 3909, 348–362.
- Wang, K., Li, M., Hakonarson, H., 2010. Analysing biological pathways in genome-wide association studies. *Nature Reviews Genetics* 11, 843–854.
- Wang, K., Saito, M., Bisikirska, B.C., *et al.*, 2009b. Genome-wide identification of post-translational modulators of transcription factor activity in human B cells. *Nature Biotechnology* 27, 829–839.
- Wegrzyn, J.L., Lee, J.M., Liechty, J., Neale, D.B., 2009. PineSAP—sequence alignment and SNP identification pipeline. *Bioinformatics* 25, 2609–2610.
- Werner, H.M., Mills, G.B., Ram, P.T., 2014. Cancer systems biology: A peek into the future of patient care? *Nature Reviews Clinical Oncology* 11, 167–176.
- Wong, G., Leckie, C., Kowalczyk, A., 2012. FSR: Feature set reduction for scalable and accurate multi-class cancer subtype classification based on copy number. *Bioinformatics* 28, 151–159.
- Yang, X., Deignan, J.L., Qi, H., *et al.*, 2009. Validation of candidate causal genes for obesity that affect shared metabolic pathways and networks. *Nature Genetics* 41, 415–423.
- Yarden, Y., Pines, G., 2012. The ERBB network: At last, cancer therapy meets systems biology. *Nature Reviews Cancer* 12, 553–563.
- Yeung, M.K., Tegner, J., Collins, J.J., 2002. Reverse engineering gene networks using singular value decomposition and robust regression. *Proceedings of the National Academy of Sciences of the United States of America* 99, 6163–6168.
- Yu, K., Martin, R., Rothman, N., Zheng, T., Lan, Q., 2007. Two-sample comparison based on prediction error, with applications to candidate gene association studies. *Annals of Human Genetics* 71, 107–118.
- Zhang, B., Gaiteri, C., Bodea, L.G., *et al.*, 2013. Integrated systems approach identifies genetic nodes and networks in late-onset Alzheimer's disease. *Cell* 153, 707–720.
- Zhang, Q.C., Petrey, D., Deng, L., *et al.*, 2012. Structure-based prediction of protein-protein interactions on a genome-wide scale. *Nature* 490 (7421), 556–560.
- Zhang, W., Chen, Y., Sun, F., Jiang, R., 2011. DomainRBF: A Bayesian regression approach to the prioritization of candidate domains for complex diseases. *BMC Systems Biology* 5, 55.
- Zhao, X.D.D.A., Lim, W.K., Brahmachary, M., *et al.*, 2009. The N-Myc-DLL3 cascade is suppressed by the ubiquitin ligase Huwe1 to inhibit proliferation and promote neurogenesis in the developing brain. *Developmental Cell* 17, 210–221.

Relevant Website

<https://genomeinterpretation.org/CAGI>.

Drug Targeting

J Longden, University of Copenhagen, Lyngby, Denmark

© 2016 Elsevier Inc. All rights reserved.

Drug Targeting – An Introduction

Imming *et al.* (2006) described a drug target as a molecular structure that will undergo a specific interaction with chemicals, termed ‘drugs.’ Drugs are so called because they are administered to treat or to diagnose a disease. At the time it was estimated that potentially thousands of these drug targets exist. Here we describe the evolution of the theories and experiments that have already been used to identify hundreds of these drug targets and those likely to underpin drug target identification efforts now and in the future.

A Historical Perspective

The concept of drug targeting was first postulated by Paul Ehrlich in the late nineteenth century upon noting the selective affinity of chemical dyes for specific biological tissues. Ehrlich thus proposed the idea of ‘chemoreceptors,’ whereby certain receptors on microorganisms or cancer cells, for example, would be different from analogous structures in host tissues, and that this difference could be targeted therapeutically (Drews, 2000). By the early twentieth century John Langley had proposed a more functional theory whereby receptors serve as a ‘switch,’ accepting signals and subsequently generating biological effects. These effects could thus be blocked by chemical antagonists or activated by chemical agonists (Drews, 2000; Pina *et al.*, 2009). These theories, and their subsequent iterations and evolutions, led to the development of the first rational synthetic drug, arsphenamine, a treatment for syphilis, in 1910. In addition these studies laid the foundations for the basic discovery process, where chemical libraries are screened for their biological activity, still utilized today (Bleicher *et al.*, 2003).

Over the course of the twentieth century this idea that one disease can be treated with one compound targeting one receptor has led to previously unprecedented therapeutic triumphs (Drews, 2000). The most successful of which are a class of compounds known as statins, which lower cholesterol levels by inhibiting the enzyme HMG-CoA reductase (Torbert, 2003). However, we now know that cells and organisms are not composed of linear pathways, and are instead complex nonlinear systems comprising highly interconnected networks (Baxter *et al.*, 2013). It is thus becoming increasingly important to view drug targeting from a network perspective in addition to more traditional approaches.

From Molecules to Systems

Systems biology was initially at the heart of drug discovery, whereby herbal or traditional remedies, usually extracted from plants, were discovered through anecdotal, albeit direct, observations in people with disease (Butcher *et al.*, 2004).

Of course the advent of chemistry and the subsequent leaps in understanding with regards to how drugs function has made direct observations of this type impossible. However, the completion of the Human Genome Project at the beginning of the twenty-first century (International Human Genome Sequencing Consortium, 2001) kick-started the development of large-scale data generating tools typically referred to as the ‘omics revolution,’ a combination of genomics, proteomics, and metabolomics. These tools, along with advances in imaging technologies, such as high-content screening, which enable the generation of high throughput, phenotypic data, have facilitated the rise in the role of modern systems biology in drug targeting. The aim of modern systems biology is thus to describe both physiology and disease over multiple hierarchical levels of interaction from molecules, to networks, cells, tissues, and eventually organisms (Butcher *et al.*, 2004).

Genomics in Target and Drug Discovery

The sequencing of the human genome, and those of other organisms, has led to the identification of many sequence variations, i.e., mutations, which could in turn lead to the identification of many new drug targets in various diseases (Kramer and Cohen, 2004). A primary role of genomics therefore is to better analyze and understand this knowledge in order to identify the critical variants and thus identify the best potential therapeutic targets. To this end, genomics data can be used to investigate the functional consequences of these variations in DNA sequence and gene expression (Emilien *et al.*, 2000; Kramer and Cohen, 2004). The most commonly used tool for determining gene expression is transcriptomics, which measures the mRNA transcripts produced in a cell or tissue system of interest. However, the major limitation of these studies is the poor correlation between mRNA levels and protein levels (Greenbaum *et al.*, 2003; Vogel and Marcotte, 2012). One of the reasons for this discrepancy is that mRNA (and protein) levels are dynamic, whereas most studies are static, i.e., measurements are taken at a single timepoint. Thus the changes observed do not necessarily represent the underlying signaling networks, which are predominantly driving the phenotype. With that said, transcript data has proven to be useful for classifying samples and has been shown to provide predictive information regarding patient outcome in cancers of the breast, brain, prostate, lymphoma, and leukemia (Ciro *et al.*, 2003).

One powerful experimental method for linking genomics to phenotype is phenotypic screening in combination with RNA interference (RNAi). RNAi takes advantage of naturally occurring mechanisms for gene regulation present in all cells, which results in the cleavage of messenger RNA, thus preventing translation of the protein. This can be combined with high throughput, and/or high-content, screening to discern the effect of inhibiting gene expression on *in vitro* or *ex vivo* disease models allowing researchers to identify direct, causal links

between genes and phenotypes (Krausz and Korn, 2008). As such, since its discovery, RNAi has become one of the most important functional genomics tools available, and has been utilized extensively in target identification and validation studies (Krausz and Korn, 2008; Lindsay 2003; Sachse *et al.*, 2005).

PCSK9, an Example of Drug Targeting in the 'Omics'-Era

A particularly interesting example of the power of genomics to deliver new drug targets is the story of PCSK9 a novel target for the treatment of cardiovascular disease, the leading cause of death worldwide.

A typical strategy for the identification of the genetic factors responsible for disease is the 'common variant hypothesis' which is based upon the theory that a handful of disease causing genes will be common enough to identify in genome-wide surveys of large populations (thousands) of people. However, in the case of cardiovascular disease this strategy has had limited impact, as whilst many common genetic variants have been identified, each only has a small effect on disease (Hall, 2013). Instead, scientists at the University of Texas Southwestern Medical Center took the opposite approach. They reasoned that Mendelian disorders, i.e., those caused by a single gene, do not typically arise from a single common mutation as this would most likely be removed through the process of natural selection. Instead they suggested that these sort of diseases are caused by one of a number of rare mutations, acting individually or cumulatively, but each resulting in the same disease phenotype (Hall, 2013).

Indeed a key aspect of the discovery of PCSK9 was the combination of phenotypic observations with genomic explanations. Phenotypic studies were first employed to group people based on their cholesterol levels. Then, the DNA of study participants was sequenced in order to identify the aberrant genes driving the disease phenotype. It was found that PCSK9, depending on the mutation expressed, can either raise or lower serum cholesterol levels due to its interaction with low-density lipoprotein receptors (Cohen *et al.*, 2006; Zerhouni, 2014). Wild-type PCSK9 interacts with the low-density lipoprotein receptor, this interaction leads to the targeting of the receptor, and any bound low-density lipoprotein cholesterol, for degradation. In some cases mutations in PCSK9 result in the formation of a protein that can no longer bind to the low-density lipoprotein receptor, thus the receptor is no longer degraded and can therefore return to the cell surface, further reducing serum cholesterol levels. Other mutations can result in a PCSK9 protein with increased protease activity, which leads to a depletion in low-density lipoprotein receptor levels and thus causes reduced low-density lipoprotein cholesterol uptake and therefore higher serum cholesterol levels (Zerhouni, 2014). These observations, coupled with the fact that a small percentage of the population have mutations which result in PCSK9 not being produced, and show no adverse effects from this, have led to the rapid development of PCSK9 inhibitors, which are currently undergoing phase 3 clinical trials (Sheridan, 2013).

Thus in the case of PCSK9 it would seem likely that a purely genomics driven approach would find it high

impossible to explain the complex mechanisms underlying what appears to be, a relatively simple system where one gene is directly responsible for a disease phenotype. Whilst genomics has been, and will continue to be, a major contributor to our understanding of the function of individual genes alone it has been less successful at elucidating the dynamic behavior of complex biological systems that are key to understanding disease (Peitsch, 2014). To this end, it would seem that the integration of genomic data with data from other sources, such as proteomics, metabolomics, and/or detailed phenotyping, along with the associated modeling tools, will hold the most potential to deliver modern treatment strategies and novel drug targets and drug candidates that take into account the complexity of disease physiology.

Proteomics and Metabolomics

Proteomics attempts to identify, and quantify the abundance of, proteins, the primary functional units of signaling networks and the principle target of drug discovery efforts. As such proteomics, utilizing the most notable proteomic technology, mass spectrometry, can provide important information as to the signaling network state of cell populations and has seen increasing usage in target identification and validation studies.

The recent development of more powerful mass spectrometers, combined with novel sample preparation and fractionation protocols has facilitated the use of global proteomics studies to identify novel drug targets (Veenstra, 2006; Sleno and Emili, 2008). In these studies global protein levels can be detected and compared between healthy and diseased samples, either from patients affected by a specific disorder or *in vitro* or *in vivo* models of disease. These studies typically result in the identification of hundreds of differences, thus an ongoing key challenge is the design of additional validation assays, or computational models, to explain these observations and prioritize the most notable drug targets (Veenstra, 2006).

An additional central aspect of global proteomics studies is the ability of mass spectrometers to detect posttranslational modifications. One such example of this is protein phosphorylation, which has been shown to be a key mechanism for regulating signaling network function. In addition, the recent clinical successes of drugs targeting protein kinases has led to increased drug discovery efforts focusing on the role of phosphorylation in disease states (McConnell and Wadzinski, 2009).

Finally, proteomics has also seen extensive use in mechanism of action studies, which can be used to elucidate the target of a molecule with a known biological effect (Sleno and Emili, 2008). In these assays the proteins bound to the molecule of interest can be isolated and identified using affinity chromatography, protein microarrays, or activity-based probes. In each case, protein lysates are mixed with the molecule of interest, the proteins or peptides binding to this molecule can then be identified via tags (e.g., fluorescent or radioactive markers) or purified and identified via mass spectrometry.

Metabolomics, which can also be performed using mass spectrometry (or nuclear magnetic resonance) measures the

level of endogenous chemicals and small molecules in a system and can thus provide important information regarding system dynamics, i.e., the processes occurring in the cell and those which have occurred. In many cases it has been shown that metabolic changes have a causative role in disease progression, thus by identifying these changes metabolomics can potentially identify new targets for therapeutic intervention (Rabinowitz *et al.*, 2011; Robertson and Frevert, 2013).

Data Integration

It is becoming apparent that biological systems cannot simply be understood by the analysis of one-type of data (Gomez-Cabrero *et al.*, 2014). As such, the integration of multiple types of omics data is now commonplace. This type of data integration is referred to as 'horizontal integration,' which attempts to explain systems-level changes from individual components (Lauffenburger, 2012). There are multiple methods for integrating these sources of data, but they generally focus on three tasks: identifying the network scaffold to describe the connections between components; deconstructing the network into its composite modules to infer the organization and active portions of the network; and the generation of systems models to simulate and predict network interactions driving phenotypic changes (Joyce and Palsson, 2006). To date studies utilizing data integration methods have predominantly focussed on transcript (mRNA expression) data and methods to integrate this with other data sources, such as proteomics or metabolomics (Berg, 2014; Zhang *et al.*, 2010). Whilst improvements in both omics technologies and data integration methods are still required results so far suggest that data integration will be pivotal for understanding how systems function and has already proven to have clinically relevant applications, most notably in multiple sclerosis and cancer (Joyce and Palsson, 2006).

Network Medicine

The omics technologies and the growth of modern systems biology has given rise to a new methodology for target selection and drug identification, termed 'network medicine' (Creixell *et al.*, 2012; Pawson and Linding, 2008). Pioneering work, discussed below, has demonstrated that cellular components are highly interlinked in complex inter- and intracellular signaling networks. As such diseases are most usually caused by abnormalities within these networks, rather than aberrations in single genes (Pawson and Linding, 2008). It can therefore be expected that drug therapies should target these networks and their dynamics in particular. Whilst it can easily be imagined that network properties and/or dynamics can be affected by a single drug, if designed with this in mind, there has also been considerable attention given to therapies involving multiple targets. The efficacy of this multi-target approach has been particularly successful developing therapies for AIDS, cancer, and depression (Barabasi *et al.*, 2011). Indeed moving forward it would seem likely that the most complex diseases will require more complex treatment strategies.

Properties of Networks

Biological systems are wired in complex interconnected networks, which when exposed to perturbation can, and often do, give rise to secondary effects and adaptations in the network itself (Janes and Lauffenburger, 2013). These biological networks are not arranged arbitrarily, but are instead governed by a set of organizing principles (Barabasi and Oltvai, 2004; Barabasi *et al.*, 2011). Firstly, biological networks cluster into what are termed 'modules.' These modules have a high number of internal connections but only have a relatively low number of external connections. Secondly, biological networks are comprised of nodes, for example, genes or proteins, which are linked together in a nonrandom manner. Within these networks there will be a large number of so-called 'hub' nodes. Hub nodes are highly interlinked nodes which serve one of two functions, either acting inside a module to regulate a specific cellular process (party hubs) or to link between modules and thus regulate between processes (date hubs). Thirdly, within a biological network there will only be a relatively short distance between nodes. This means that perturbing one node is very likely to have an effect on many other nodes within the vicinity and could potentially have an effect on the comprising module and even the network itself. This would of course be particularly likely if the perturbed node is also a hub. Finally, biological networks contain 'bottlenecks.' Bottlenecks are nodes through which the shortest path between pairs of nodes is often found and are thus frequently 'essential genes,' i.e., critical for the organisms survival. However, only about 25% of the known human genes have a demonstratable link to disease (Gillham, 2011). It is therefore possible to further characterize the network attributes that are specific to disease genes and thus predict their locations within networks (Goh *et al.*, 2007; Barabasi *et al.*, 2011). Firstly, of the known disease genes only ~20% are also essential genes, which are often hubs or bottlenecks and expressed in multiple tissues. In comparison, disease genes are usually tissue specific, do not encode hubs and are found most usually at the periphery of interaction networks. An interesting exception to that rule are the cancer-related proteins, which are more likely to act as hubs within protein interaction networks (Kar *et al.*, 2009). Secondly, nodes of a network driving disease tend to cluster. Thus if a few disease nodes are identified it is likely that other disease-related components will be found in their network vicinity in so-called 'disease modules.'

These observations have led to the theory of the diseaseome (Goh *et al.*, 2007), a biological network in which diseases are the nodes and the links between nodes are various molecular relationships arising from disease-associated cellular components, genes in the case of the human disease network described in Goh *et al.* (2007). In this way, multiple studies have shown that diseases can be linked based on their associated genes, their shared metabolic pathways and the microRNAs regulating the expression of disease genes (Barabasi *et al.*, 2011; Lee *et al.*, 2008; Lu *et al.*, 2008; Park *et al.*, 2009). Indeed it has also been shown that patients are at an increased likelihood of developing disease in the network vicinity of diseases they are already suffering from and that the number of links a disease has correlates with mortality rates (Barabasi *et al.*, 2011; Park *et al.*, 2009). An interesting example

of this is diabetes mellitus and Alzheimer's disease, which share a genetic link, as illustrated in the human disease network (Goh *et al.*, 2007). Whilst the pathology of Alzheimer's disease is still debated the brains of Alzheimer's sufferers have been shown to contain certain protein deposits, such as amyloid plaques and neurofibrillary tangles, which are also found in the brains of patients who died of diabetes. In addition, irregular expression of genes encoding insulin and its receptor has also been noted in patients with Alzheimer's disease. As such, there are now some clinicians that refer to Alzheimer's disease as type 3 diabetes and suggest that it can be treated with antidiabetes therapies (de la Monte and Wands, 2008).

Targeting Networks

There are now a number of studies which have targeted biological networks specifically in order to identify drug targets or novel therapies. One particularly compelling example of this has recently been published by the Yaffe laboratory at MIT. In this study it was shown that epidermal growth factor receptor (EGFR) inhibitors, if given at least 4 h prior, can increase the activity of doxorubicin by up to 500%. Importantly, a combination of an EGFR inhibitor and doxorubicin given at the same time showed no significant increase in activity compared to doxorubicin alone. In addition, the additive effect of EGFR inhibitor pretreatment on doxorubicin activity was only seen in triple negative breast cancer cells (compared to hormone sensitive and HER2 overexpressing cells). An investigation into the mechanism behind this novel observation showed that, in triple negative breast cancer cells EGFR inhibition resulted in a dynamic network rewiring, unmasking an apoptotic process involving the activation of caspase-8 (Lee *et al.*, 2012). Following this initial discovery a nanoparticle was developed to deliver this drug combination and has subsequently been used to test, and confirm, the effect in *in vivo* models (Morton *et al.*, 2014). Given that clinical trials are already planned there could be a very short turnaround time from the discovery of this novel therapy to it actually being utilized in patients.

These findings help to illustrate the fact that drugs, even if they are designed to have one target, will have network level effects, and that these can be taken advantage of, to create novel therapies (Erler and Linding, 2012; Hopkins, 2007). Even for individual drugs, particularly protein kinase inhibitors, it has been known for some time that their effect is most likely due to the fact that they act on multiple signaling kinases (Hopkins, 2007).

One potential caveat is that the novel network targets identified by systems biology approaches may not be drugable by conventional means. Thus, in combination with the identification of novel targets we will also require novel drugs.

New Methods for Drug Targeting

One rapidly emerging technology that could potentially solve this issue is small interfering RNA (siRNA) therapies. siRNAs are a type of synthetic RNAi molecule that are potent gene-specific inhibitors. In the case of siRNAs drug design is simple, there are already siRNAs readily available for all the known

genes in the genome. This can be extended further such that siRNAs can be designed to be transcript (Bolduc *et al.*, 2014) or even mutation specific, even if the mutated gene differs from the wild-type version by just a single base (Brummelkamp *et al.*, 2002). Thus, if advantageous to do so, siRNA therapies could be 'personalized' toward specific disease genes.

The major limitation of siRNAs however is the fact that they are degraded in the blood, and do not traverse the cell membrane. They therefore require potentially complex delivery mechanisms in order to be effective. Whilst progress in this regard is being made effective delivery mechanisms are not widely available, nor are they currently effective for all tissues (Wu *et al.*, 2014).

We therefore stand at an interesting juncture. We are learning more about disease networks and their effects on systems than ever before, but there is an intrinsic need for better identification of drug targets and drug candidates with which to treat them. However, meeting these challenges will undoubtedly lead to great improvements in the drug discovery process.

See also: Molecular Principles, Components, Technology, and Concepts: Proteins: Drug Design

References

- Barabasi, A.-L., Gulbahce, N., Loscalzo, J., 2011. Network medicine: A network-based approach to human disease. *Nature Reviews Genetics* 12, 56–68.
- Barabasi, A.-L., Oltvai, Z.N., 2004. Network biology: Understanding the cell's functional organization. *Nature Reviews Genetics* 5, 101–114.
- Baxter, K., Horn, E., Gal-Edd, N., *et al.*, 2013. An end to the myth: There is no drug discovery pipeline. *Science Translational Medicine* 5 (171), 171cm1.
- Berg, E.L., 2014. Systems biology in drug discovery and development. *Drug Discovery Today* 19 (2), 113–125.
- Bleicher, K.H., Boehm, H.-J., Mueller, K., Alanine, A.I., 2003. A guide to drug discovery: Hit and lead generation: Beyond high-throughput screening. *Nature Reviews Drug Discovery* 2, 369–378.
- Bolduc, V., Zou, Y., Ko, D., Boennemann, C.G., 2014. siRNA-mediated allele-specific silencing of a COL6A3 mutation in a cellular model of dominant ullrich muscular dystrophy. *Molecular Therapy Nucleic Acids* 3, e147.
- Brummelkamp, T.R., Bernards, R., Agami, R., 2002. Stable suppression of tumorigenicity by virus-mediated RNA interference. *Cancer Cell* 2 (3), 243–247.
- Butcher, E.C., Berg, E.L., Kunkel, E.J., 2004. Systems biology in drug discovery. *Nature Biotechnology* 22 (10), 1253–1259.
- Ciro, M., Bracken, A.P., Helin, K., 2003. Profiling cancer. *Current Opinion in Cell Biology* 15, 213–220.
- Cohen, J.C., Boerwinkle, E., Mosley Jr., T.H., Hobbs, H.H., 2006. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. *New England Journal of Medicine* 354, 1264–1272.
- Creixell, P., Schoof, E.M., Erler, J.T., Linding, R., 2012. Navigating cancer network attractors for tumor-specific therapy. *Nature Biotechnology* 30, 842–848.
- Drews, J., 2000. Drug discovery: A historical perspective. *Science* 287, 1960–1964.
- Emilien, G., Ponchon, M., Caldas, C., Isacson, O., Maloteaux, J.M., 2000. Impact of genomics on drug discovery and clinical medicine. *QJM* 93, 391–423.
- Erler, J.T., Linding, R., 2012. Network medicine strikes a blow for breast cancer. *Cell* 149, 731–733.
- Gillham, N.W., 2011. *Genes, Chromosomes, and Disease: From Simple Traits, to Complex Traits, to Personalized Medicine*. Upper Saddle River, NJ: FT Press.
- Goh, K.-I., Cusick, M.E., Valle, D., *et al.*, 2007. The human disease network. *Proceedings of the National Academy of Sciences of the United States of America* 104 (21), 8685–8690.
- Gomez-Cabrero, D., Abugessaisa, I., Maier, D., *et al.*, 2014. Data integration in the era of omics: Current and future challenges. *BMC Systems Biology* 8 (Suppl. 2), I1.

- Greenbaum, D., Colangelo, C., Williams, K., Gerstein, M., 2003. Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biology* 4, 117.
- Hall, S.S., 2013. Genetics: A gene of rare effect. *Nature* 496, 152–155.
- Hopkins, A.L., 2007. Network pharmacology. *Nature Biotechnology* 25 (10), 1110–1111.
- Imming, P., Sinning, C., Meyer, A., 2006. Drugs, their targets and the nature and number of drug targets. *Nature Reviews Drug Discovery* 5, 821–834.
- International Human Genome Sequencing Consortium, 2001. Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.
- Janes, K.A., Lauffenburger, D.A., 2013. Models of signalling networks — What cell biologists can gain from them and give to them. *Journal of Cell Science* 126 (9), 1913–1921.
- Joyce, A.R., Palsson, B.Ø., 2006. The model organism as a system: Integrating 'omics' data sets. *Nature Reviews Molecular Cell Biology* 7, 198–210.
- Kar, G., Gursoy, A., Keskin, O., 2009. Human cancer protein–protein interaction network: A structural perspective. *PLoS Computational Biology* 5 (12), e1000601.
- Kramer, R., Cohen, D., 2004. Functional genomics to new drug targets. *Nature Reviews* 3, 965–972.
- Krausz, E., Korn, K., 2008. High content siRNA screening for target identification and validation. *Expert Opinion on Drug Discovery* 3 (5), 551–564.
- Lauffenburger, D.A., 2012. The multiple dimensions of integrative biology. *Integrative Biology* 4, 9.
- Lee, D.-S., Park, J., Kay, K.A., *et al.*, 2008. The implications of human metabolic network topology for disease comorbidity. *Proceedings of the National Academy of Sciences of the United States of America* 105 (29), 9880–9885.
- Lee, M.J., Ye, A.S., Gardino, A.K., *et al.*, 2012. Sequential application of anticancer drugs enhances cell death by rewiring apoptotic signalling networks. *Cell* 149, 780–794.
- Lindsay, M.A., 2003. Target discovery. *Nature Reviews Drug Discovery* 2, 831–838.
- Lu, M., Zhang, Q., Deng, M., *et al.*, 2008. An analysis of human microRNA and disease associations. *PLoS ONE* 3 (10), e3420.
- McConnell, J.L., Wadzinski, B.E., 2009. Targeting protein serine/threonine phosphatases for drug development. *Molecular Pharmacology* 75 (6), 1249–1261.
- de la Monte, S.M., Wands, J.R., 2008. Alzheimer's disease is type 3 diabetes — Evidence reviewed. *Journal of Diabetes Science and Technology* 2 (6), 1101–1113.
- Morton, S.W., Lee, M.J., Deng, Z.J., *et al.*, 2014. A nanoparticle-based combination chemotherapy delivery system for enhanced tumor killing by dynamic rewiring of signalling pathways. *Science Signaling* 7 (325), ra44.
- Park, J., Lee, D.-S., Christakis, N.A., Barabasi, A.-L., 2009. The impact of cellular networks on disease comorbidity. *Molecular Systems Biology* 5, 262.
- Pawson, T., Linding, R., 2008. Network medicine. *FEBS Letters* 582, 1266–1270.
- Peitsch, M.C., de Graaf, D., 2014. A decade of Systems Biology: Where are we and where are we going to? *Drug Discovery Today* 19 (2), 105–107.
- Pina, A.S., Hussain, A., Roque, A.C., 2009. An historical overview of drug discovery. *Methods in Molecular Biology* 572, 3–12.
- Rabinowitz, J.D., Purdy, J.G., Vastag, L., Shenk, T., Koyuncu, E., 2011. Metabolomics in drug target discovery. *Cold Spring Harbor Symposia on Quantitative Biology* 76, 235–246.
- Robertson, D.G., Frevert, U., 2013. Metabolomics in drug discovery and development. *Clinical Pharmacology and Therapeutics* 94, 559–561.
- Sachse, C., Krausz, E., Kroenke, A., *et al.*, 2005. High-throughput RNA interference strategies for target discovery and validation by using synthetic short interfering RNAs: Functional genomics investigations of biological pathways. *Methods in Enzymology* 392, 242–277.
- Sheridan, C., 2013. Phase 3 data for PCSK9 inhibitor wows. *Nature Biotechnology* 31, 1057–1058.
- Sleno, L., Emili, A., 2008. Proteomic methods for drug target discovery. *Current Opinion in Chemical Biology* 12 (1), 46–54.
- Torbert, J.A., 2003. Lovastatin and beyond: The history of the HMG-CoA reductase inhibitors. *Nature Reviews Drug Discovery* 2, 517–526.
- Veenstra, T.D., 2006. Proteomic approaches in drug discovery. *Drug Discovery Today: Technologies* 3 (4), 433–440.
- Vogel, C., Marcotte, E.M., 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics* 13, 227–232.
- Wu, S.Y., Lopez-Berestein, G., Calin, G.A., Sood, A.K., 2014. RNAi therapies: Drugging the undruggable. *Science Translational Medicine* 6 (240), 240ps7.
- Zerhouni, E.A., 2014. Turning the titanic. *Science Translational Medicine* 6 (221), 221ed2.
- Zhang, W., Li, F., Nie, L., 2010. Integrating multiple 'omics' analysis for microbial biology: Application and methodologies. *Microbiology* 156, 287–301.

Index

Notes

Cross-reference terms in italics are general cross-references, or refer to subentry terms within the main entry (the main entry is not repeated to save space).

The index is arranged in set-out style with a maximum of two levels of subheading. Major discussion of a subject is indicated by bold page numbers. Page numbers suffixed by *T* and *F* refer to tables and figures, respectively; vs. indicates a comparison.

This index is in word-by-word order with hyphens within index headings are ignored in the alphabetization. For example, sweating is alphabetized after sweat test, not after structural matrices or C-family is after CetZs, and not at the start of the C section. Prefixes and terms in parentheses are excluded from the initial alphabetization.

Where index subentries and sub-subentries pertaining to a subject have the same page number, they have been listed to indicate the comprehensiveness of the text.

A

- a disintegrin and metalloprotease *see* a disintegrin and metalloproteinase (ADAM)
- a disintegrin and metalloproteinase (ADAM) 1:630, 1:639
- ADAM10 1:643–644
- receptors 2:776
- substrates 1:644–646
- ADAM17
- dependent EGFR activation 1:641–642
- for EGFR signaling 1:640–641
- protein ectodomain shedding of TNF α 1:641F
- substrates 1:644–646
- TACE 1:639–640
- domain structure 1:640F
- iRhoms 1:642–643
- proteases 1:754–755
- a disintegrin-like and metalloprotease domain with thrombospondin proteases (ADAMTS proteases) 1:630, 2:699
- Adamts1* deficient mice 1:635
- ADAMTS1 1:630
- biosynthesis and regulation 1:633
- domain organization and 3-dimensional structure 1:631–633
- mechanistic basis of genetic and acquired disorders 1:633–634
- ADAMTS2 1:634
- ADAMTS10 1:634–635
- ADAMTS13, vWF and TTP 1:633–634
- aggrecanases in osteoarthritis 1:635
- dermatosparaxis 1:634
- procollagen processing 1:634
- tissue microfibrils 1:634–635
- WMS 1:634–635
- mouse mutants 1:635
- ADAMTS proteolysis of versican 1:635–636
- ADAMTS1 and female infertility 1:636
- A site 3:310
- A₃ cluster 3:289
- AAA protein *see* ATPase associated with various cellular activities protein (AAA protein)
- AAA + protein 1:396, 1:401–403, 1:403
- AAA + ring domains 2:621
- AAA-adenosine triphosphatase (ATPases) 3:299
- AA-tRNA *see* aminoacyl-tRNA (AA-tRNA)
- AAV *see* adeno-associated viral (AAV)
- AB₅ toxins 2:434, 2:438
- abasic site incision 1:446–447
- ABBA *see* actin-bundling protein with BAIAP2 homology (ABBA)
- ABC *see* activated B-cell (ABC); aggressive activated B cell (ABC)
- ABCD4 gene 1:775
- ABCE1 1:527, 1:530F, 1:531, 1:532
- ABCP *see* apical-basal cell polarity (ABCP)
- A β peptide *see* amyloid β -peptide (A β peptide)
- ABFs *see* antibacterial factors (ABFs)
- ABM *see* actin-based motility (ABM); agent-based modeling (ABM)
- abnormal calpain activation 1:675
- abscission 3:497
- checkpoint 3:501
- ESCRT proteins 3:499
- membrane fusion 3:498–499
- midbody remnant 3:499
- vesicle trafficking 3:497–498
- abscission checkpoint 2:470
- abscission/NoCut checkpoint regulator (ANCHR) 2:470, 3:475
- absent in melanoma 2 (AIM2) 3:710
- absolute reaction rate theory 1:18
- absorption, nutrient 1:732, 1:734, 1:736
- ACAT *see* acyl-coenzyme A:cholesterol acyltransferase (ACAT)
- AcbA/Acb1 *see* acyl coenzyme A binding protein (AcbA/Acb1)
- ACC *see* acetyl-CoA carboxylase (ACC); adenoid cystic carcinoma (ACC)
- accessibility, of protein 1:69
- accessory proteins 2:478–481, 2:588
- accountability 1:68–69
- ACD protein homolog *see* adrenocortical dysplasia protein homolog (ACD protein homolog)
- ACE *see* angiotensin I converting enzyme (ACE); angiotensin-converting enzyme (ACE)
- AceCS2 *see* acetyl co-enzyme A synthetase 2 (AceCS2)
- acentrosomal nucleation 2:544
- acentrosomal spindle assembly 2:638
- acetal protecting groups 1:43–44
- 1-aryl-4-alkoxypiperidin-4-yl groups 1:43–44
- 2-cyanoethyloxymethyl group 1:44–45
- triisopropylsilyloxymethyl group 1:44
- acetyl coenzyme A 1:14
- acetyl co-enzyme A synthetase 2 (AceCS2) 3:781
- acetylated tubulin 3:233
- acetylation 3:507
- acetylcholinesterase (AChE) 1:248
- acetyl-CoA/CoA ratio 1:288, 1:292–293, 1:296
- for ATP production and/or lipid synthesis 1:294
- in fatty acid oxidation and ketogenesis 1:296
- in regulation of fatty acid synthesis 1:293F, 1:294
- in regulation of glycogen synthesis 1:291
- in regulation of pyruvate dehydrogenase complex 1:293–294
- acetyl-CoA carboxylase (ACC) 1:290, 1:293F, 1:294
- acetyl-Ser-Asp-Lys-Pro (AcSDKP) 1:748
- ACFD *see* acrocapitofemoral dysplasia (ACFD)
- ACG for activated Cdc42-associated GEF *see* DOCK11
- achaete-scute homologue 2 (Ascl-2) 3:542
- AChE *see* acetylcholinesterase (AChE)
- Acholeplasma laidlawii* (*A. laidlawii*) 1:164
- acid hydrolase 1:769, 1:770, 1:772, 1:773, 1:777–778, 1:780–781, 1:783–784

- acid maltase 1:764, 1:767, 1:768–769, 1:770, 1:771–772
 recombinant 1:767
 acid sphingomyelinase (ASMase) 1:613, 2:225–226
 acid–base catalysis 1:16F, 1:18
 acid–base reactions 1:10
 acidification 2:426
 AcMNPV *see* *Autographa californica* (alfalfa looper) multiple nuclear polyhedrosis virus (AcMNPV)
 ACP *see* acyl carrier protein (ACP)
 Acp2 1:778
 Acp5 1:776, 1:776–777, 1:778
 acquired immune system *see* adaptive immune system
 acquired immunity 3:580
 acquired immunodeficiency syndrome (AIDS) 1:739, 2:828, 3:585
see also human immunodeficiency virus (HIV)
 acrocapitofemoral dysplasia (ACFD) 3:81
 ACs *see* anterograde carriers (ACs)
 AcSDKP *see* acetyl-Ser-Asp-Lys-Pro (AcSDKP)
 ActD *see* actino-mycin D (ActD)
 ACTH *see* adrenocorticotrophic hormone (ACTH)
 actin 2:4, 2:26F, 2:27, 2:28F, 2:367, 2:436, 2:438, 2:548–549, 2:761–762, 2:795, 3:199, 3:367
see also intermediate filaments (IFs); microtubules (MTs)
 actin-based motility 2:560–561
 atomic structure of monomeric actin 2:550F
 depolymerization 2:787–788
 and membrane rearrangement 2:787
 polymerization 2:726, 2:754–755
 phagocytosis signaling to 2:752–753
 polymerization-driven protrusion 2:683–686
 regulators in adhesion structures 2:561–564
 reorganization 2:726
 role in gene expression regulation 2:564–565
 self-assembly
 bulk solution studies 2:552–553
 single filament assembly dynamics 2:553–554
 treadmill regulation by ADF and capping proteins 2:553F
 structural and mechanical properties 2:549–552
 actin assembly dynamics 2:548–568
 actin-based molecular motors 2:564
 bacterial actin ancestors 2:565
 cellular examples of motile processes 2:562F
 DRFs and non-DRFs 2:558T
 integrated protein modules in reconstituted actin-based motility 2:560–561
 regulation by associated proteins 2:554–557
 barbed end-capping proteins 2:557
 barbed-end trackers 2:557
 G-actin binding proteins 2:554–557
 pointed end binding proteins 2:557–558
 proteins binding to sides of actin filaments 2:558–559
 WH2 domain 2:559–560
 regulators 2:555–556T
 site-directed branching of actin filaments 2:560
 actin barbed face 2:549
 actin cytoskeleton 1:202, 2:581–584, 2:600–601, 2:602, 2:687–688, 3:6
 cell–cell adhesion and 2:704–705
 biological functions of cadherin–actin interactions 2:706
 classical cadherin adhesion integration 2:704–705, 2:704F
 dynamic regulation 2:705–706
 physical connections between cadherins and F-actin 2:705
 dynamic connections between cadherin and actin cytoskeleton 3:184–185
 Rho GTPases and roles in regulating 2:733–735, 2:733F, 2:734F
 actin depolymerizing factor (ADF) 2:559, 2:685
 actin dynamics 2:689–691, 2:725, 2:759, 3:681, 4:10
 at AJ 3:187
 biochemical modeling 4:14
 Arp2/3-mediated branch formation mechanism 4:17, 4:19–20
 ATP-hydrolysis 4:15–16
 cofilin 4:19
 cytoskeletal dynamics 4:19
 F-actin 4:19
 FRAP 4:21
Listeria 4:16–17
 nucleotide state profile 4:16F
 Open Model 4:17, 4:18F
 3D lamellipodium 4:20
 treadmill process 4:15F
 experimental investigation 4:10
 mechano-chemical control of cell motility and shape 4:10
 actin filament treatment 4:11
 biophysical mechanisms 4:14
 Brownian ratchet formalism 4:10, 4:10
 fish keratocyte cell phase microscopy image 4:12F
 G-actin transport 4:13
 locomotion mechanism 4:13F
 model types and evolution 4:11F
 network swelling model 4:12
 actin elongation 2:691
 actin filaments 2:551–552, 2:677–679, 2:725, 2:752
see also intermediate filaments (IFs); microtubules (MTs)
 nucleation of 2:554
 organization 2:683F, 2:684F
 in filopodium 2:686F
 in lamellipodium 2:689F
 proteins binding to sides of 2:558–559
 site-directed branching of 2:560
 treadmill 2:684F
 actin-based motility (ABM) 2:788
 actin-based motor 2:4
 actin-binding proteins (actin-BP) 2:742
 actin-bundling protein with BAIAP2 homology (ABBA) 2:497
 actin-dependent melanosome transport system 2:237
 actin-like MreB in rod-shaped cells 2:593
 actin-like ParM 2:592
 actin-like proteins (Alps) 2:565
 actino-mycin D (ActD) 3:281
 actin-related protein 2/3 complex (Arp2/3 complex) 2:221, 2:491, 2:560, 2:691, 2:723, 2:733–734, 2:753, 3:187
 Arp2/3-mediated branch formation mechanism 4:19–20
 action myoclonus-renal failure syndrome (AMRF) 1:773–774
 action potential 1:218, 1:225–226, 1:226, 1:227T, 1:229–230, 1:230–231, 1:230F, 1:231, 1:231F, 1:246–247, 2:381
 invasion of nerve terminal 1:248
 propagation 1:247–248
 activated B-cell (ABC) 3:585
 activating molecule in BECLIN1-regulated autophagy (AMBRA1) 2:245
 activating transcription factor-4 (ATF4) 2:531
 activating transcription factor 6 (ATF6) 2:529, 2:529–530, 2:531
 activation, Class I and Class II 1:476–477
 activation, Gibbs free energy of 1:8
 activation domain (AD) 3:704
 activation induced cell death (AICD) 3:567
 activation induced TNF-receptor (AITR) 3:62
 activation-induced cytidine deaminase (AID) 3:552
 activation-induced deaminase (AID) 3:543–544
 activator protein, deficiency of 1:774
 galactosialidosis 1:774
 sphingolipid activator proteins (SAP) 1:774
 activator protein 1 (AP1) 3:355–356, 3:574–575, 3:712–714
 activators 1:368, 1:488, 1:490–491, 1:590, 1:684, 3:346
 active ion transport
 via primary and secondary active transporters 1:226
 active layer 4:118
 active pharmaceutical ingredient (API) 1:135
 active regulator of SIRT1 (AROS) 3:275
 active transporters, primary and secondary 1:226, 1:226–228, 1:231–232, 1:235
 activins 3:39, 3:44
 activity-based probes 1:131, 1:131F
 actomyosin cytoskeleton 2:603, 2:604
 actomyosin-driven contractility 2:604–605
 acute kidney injury (AKI) 1:694
 acute myelogenous leukemia (AML) 3:24

- acute myeloid leukemia (AML) 3:282–283
acute myeloid leukemia 1 protein (AML1)
see Runt-related transcription factor-1 (RUNX1)
acute phase reactant 1:790
acute phase response (APR) 2:534
acute promyelocytic leukemia 3:811
acyl carrier protein (ACP) 1:161
acyl coenzyme A binding protein (AcbA/Acb1) 2:520
acyl-coenzyme A:cholesterol acyltransferase (ACAT) 1:173–174
acyltransferases 1:204*F*, 1:205
AD *see* activation domain (AD); Alzheimer's disease (AD); anauxetic dysplasia (AD)
ADAM *see* a disintegrin and metalloproteinase (ADAM)
Adams–Oliver syndrome (AOS) 3:109–110
ADAMTS proteases *see* a disintegrin-like and metalloprotease domain with thrombospondin proteases (ADAMTS proteases)
adaptation 4:154
adaptins *see* Assembly polypeptides (APs)
adaptive immune response 3:695
initiation and regulation 3:617–618
adaptive immune system 3:1, 3:592–593, 3:605, 3:710
adaptive immunity elements 3:519
adaptor protein (AP) 2:225, 2:228, 2:410, 2:450–458, 3:720–721
AP1 2:451–452
AP2 2:197, 2:241, 2:408, 2:452, 2:503
AP3 2:452–453
AP4 2:453
AP5 2:453
AP180 2:455, 2:503
arrestin family 2:456
CALM protein 2:455
Eps15 family 2:455
Epsin family 2:454
EpsinR family 2:454–455
heterotetrameric adaptor complexes 2:450–452, 2:452*F*
network 2:456–457
properties 2:451*F*
PTB domain containing family 2:455–456
adaptor protein apoptotic protease activating factor 1 (APAF-1) 3:368–369
adaptor protein containing PH domain
PTB-domain and leucine zipper motif (APPL) 2:503
adaptor(s) 2:366, 2:450
molecules 3:12
ADAR1 *see* adenosine deaminase acting on RNA 1 (ADAR1)
ADCC *see* antibody-dependent cellular cytotoxicity (ADCC)
ADCs *see* antibody-drug conjugates (ADCs)
additional strand catalytic E (ASCE) 2:621
adenine nucleotide translocase (ANT) 3:344
adeno-associated viral (AAV) 2:149
adenoid cystic carcinoma (ACC) 3:110
adenomatous polyposis coli (APC) 2:687, 3:90, 3:90–91, 3:90*F*, 3:93, 3:94, 3:237
adenosine deaminase acting on RNA 1 (ADAR1) 1:354, 1:356–357*F*, 1:357*F*, 1:358
adenosine diphosphate (ADP) 1:95, 2:503, 3:341
adenosine diphosphate ribosylation factor (ARF) 3:728–729
adenosine monophosphate (AMP) 1:175, 1:748–749
adenosine monophosphate activated kinase (AMPK) 1:124, 1:175, 1:748–749, 1:765, 1:766, 2:245, 2:499, 3:130, 3:247, 3:274, 3:336–337, 3:416–417, 3:814–815
adenosine triphosphatase (ATPase) 3:458–459
adenosine triphosphate (ATP) 1:95, 1:237, 1:298, 1:447, 2:251, 2:620, 3:341, 3:416–417, 3:520, 4:10
hydrolysis 2:552–553
adenosine triphosphate binding cassette (ABC) transporter 1:252, 1:252–253, 1:253–254
adenylate kinase (AK) 1:17
adenylate/uridylylate-rich elements (AREs) 3:323
adenylyl cyclase 2:447, 3:52
ADF *see* actin depolymerizing factor (ADF)
ADF homology module (ADFH) 2:559
adherens junctions (AJs) 2:742, 2:744–745, 3:182
actin dynamics at 3:187
in disease 3:188–189
diversity 3:183–184
adhesion
adhesive strengthening 2:706
frequency 3:661
maturation 2:720–721
mechanisms for adhesion and diapedesis
across endothelia
cytoskeletal remodeling 3:637
molecular basis for adhesion and diapedesis 3:636–637
-mediated signaling 2:713
adhesome 2:724
adipocytes 2:274, 2:276
LDs 2:274
ADP *see* adenosine diphosphate (ADP)
ADPKD *see* autosomal dominant polycystic kidney disease (ADPKD)
ADP-ribosylation factor (Arf) 2:170, 2:194–195, 2:366, 2:450
ARF-6 2:304–305
Arf1 2:366, 2:503
proteins 2:446
ADP-rybosilates 2:795
adrenocortical dysplasia protein homolog (ACD protein homolog) 1:427
adrenocorticotrophic hormone (ACTH) 2:376–377
adrenoleukodystrophy 2:269
adsorption chromatography 1:72
dye-ligand chromatography 1:72
HILIC 1:72
inorganic adsorbent chromatography 1:72
adult cell fate, plasticity of 3:791
adult central nervous system stem cell niche
CNS stem cell niches 3:800
interactions between niche cells and tissue progenitors-soluble factors 3:799
SGZ interactions 3:800
SGZ neural stem cell niche 3:799
stem cell 3:800
SVZ interactions 3:799–800
tissue with functional stem cells 3:797–798
V-SVZ neural stem cell niche 3:798–799
adult stem cell niche
adult central nervous system stem cell niche 3:797–798
cell-matrix contacts 3:802
heart-evidence for cardiac stem cell niche 3:800–801
juxtacrine signaling 3:802–803
paracrine signaling 3:802–803
skeletal muscle 3:796–797
skin hair follicle 3:794–795
ventricular-subventricular zone neural 3:798*F*
adult stem cells, aging in
HSC 3:780–782
MSCs 3:780
muscle stem cells 3:778
adult-born cells 4:137
adulthood, changes in gut microbiome 2:802
advanced glycation end-products (AGEs) 3:734–735
AEE *see* apical early endosomes (AEE)
AEP *see* asparaginyl endopeptidase (AEP)
aequorin 1:125
affinity chromatography 1:72
affinity tag chromatography 1:73
immunoaffinity chromatography 1:72
lectin affinity chromatography 1:72
metabolite affinity chromatography 1:72–73
affinity labeling 1:130–131
affinity purification mass spectrometry (AP-MS) 4:170–171, 4:190, 4:247
affinity tag chromatography 1:73
African swine fever virus (ASFV) 2:824
Ag *see* antigen (Ag)
AG *see* azurophilic granules (AG)
AGC catalytic triad 1:497, 1:499*F*
agent-based modeling (ABM) 4:80, 4:80–81
age-related macular degeneration (AMD) 1:707, 2:111–112
AGEs *see* advanced glycation end-products (AGEs)
aggrecanases in osteoarthritis 1:635
aggregation
ICS 2:103
PCH 2:103
aggresomes, NCLDV using 2:824–825
and perinuclear sites 2:824*F*
aggressive activated B cell (ABC) 4:252–253
aging 1:719–720, 3:438, 3:776
in adult stem cells

- aging (*continued*)
 HSC 3:780–782
 MSCs 3:780
 muscle stem cells 3:778
 changes in gut microbiome 2:802
 INK4a/ARF in 3:454–455
 senescence in 3:513, 3:514F, 3:515, 3:515–516
 stem cell regenerative potential declines with 3:777F
- aging factors 1:203
- AGO *see* Argonaute (AGO)
- Ago2 P-body protein 1:350
- agouti-related protein (Agrp) 2:289
- agrin 2:697
- Agrp *see* agouti-related protein (Agrp)
- AGS *see* Aicardi–Goutières syndrome (AGS)
- Aicardi–Goutières syndrome (AGS) 3:720
- AICD *see* activation induced cell death (AICD); APP intracellular domain (AICD)
- Aichivirus (AV) 3:310–312
- AID *see* activation-induced cytidine deaminase (AID); activation-induced deaminase (AID)
- AIDS *see* acquired immunodeficiency syndrome (AIDS)
- AIFM1 *see* apoptosis-inducing factor, mitochondrion-associated 1 (AIFM1)
- AIM2 *see* absent in melanoma 2 (AIM2)
- AIM2-like receptors (ALRs) 3:719–720
- AIR *see* autoinhibitory region (AIR)
- AIRE *see* autoimmune regulator (AIRE); auto-immune regulator transcription factor (AIRE)
- airway surface liquid (ASL) 1:252, 1:255–256
- AITR *see* activation induced TNF-receptor (AITR)
- AJs *see* adherens junctions (AJs)
- AK *see* adenylate kinase (AK)
- AKAP *see* A-kinase anchoring protein (AKAP)
- AKI *see* acute kidney injury (AKI)
- A-kinase anchoring protein (AKAP) 2:661–663
- Akt 3:129–130, 3:815
 AKT1 mutation 3:419
 inactivation 3:132–133
 isoform domain structure 3:129F
- alanine aminopeptidase 1:734–735
- alanine scanning mutagenesis 1:122–123, 1:123, 1:123–124
- alanyl-tRNA synthetase 1:22–23
- A-lattice *see* alpha-tubulin (A-lattice)
- Ala-Val-Pro-Ile (AVPI) 3:395
- albonoursin 1:326–327
- albumin 2:382
- ALCL *see* Anaplastic large-cell lymphoma (ALCL)
- aldehyde dehydrogenase (ALDH) 3:595
- aldosterone 1:178–179
- Algorithm for the Reconstruction of Accurate Cellular Networks (ARACNe) 4:258–259
- algorithms 4:211–212, 4:214
- Aliphatic acid-type histone deacetylase inhibitors 1:62
- ALK *see* anaplastic lymphoma kinase (ALK)
- alkaline pancreatic fluid 1:257–259, 1:259F
- alkaline phosphatase (ALP) 3:797
- alkyne-labeled coenzyme Q analogues 1:128
- allelic exclusion 3:553F, 3:554–555
- allograft rejection 3:642
- allosteric control 1:288
- allosteric effectors 1:249–250
- allosteric enzymes 1:9, 4:31, 4:32
- allosteric site 1:744
- allotransplantation 3:808
- all-trans retinoic acid (ATRA) 3:811
- ALP *see* alkaline phosphatase (ALP)
- alpha granules 1:715, 2:235, 2:235–236
- alpha-1-antitrypsin deficiency (α 1-ATD) 1:3–4, 1:789
 cellular mechanisms
 of liver disease 1:792–794
 of lung disease 1:794
 cellular responses to misfolding of mutant ATZ 1:794–796
 mechanism of deficiency in classical form of 1:791–792
 physiology and function of AT 1:790–791
 therapies for ATD
 liver disease 1:796–797
 lung disease 1:797
- α 1 protease inhibitor (α 1PI) 1:710
- α 1-adrenoceptor (α 1-AR) 3:51
- α 1-antitrypsin (AT) 1:789
 physiology and function of 1:790–791
- α 1-antitrypsin Z (ATZ) 1:789
 cellular responses to misfolding of mutant 1:794–796
- α 1-AR *see* α 1-adrenoceptor (α 1-AR)
- α 1-ATD *see* alpha-1-antitrypsin deficiency (α 1-ATD)
- α 1PI *see* α 1 protease inhibitor (α 1PI)
- α 2-antiplasmin (α 2-AP) 1:717
- α 2-macroglobulin 1:626–627
- α -amino acid structures 1:15T
- α -amino-3-hydroxy-5-methyl-4-iso-xazolepropionic acid (AMPA) 2:219–220, 2:294, 2:453, 2:499, 3:235
- α -amylase 1:125
- α/β hydrolases 1:91–92
- α -families 2:627
- α -hemolysin 4:162
- α -ketoacid dehydrogenase 1:14
- α -mannosidase 1:539–540T, 1:541F, 1:543, 1:779–780, 1:783
- α -mannosidosis 1:783
- α -subunit of flavocytochrome b₅₅₈ *see* p22phox
- α -subunit of IL-3 receptor (IL-3R α) 3:24
- α -synuclein 1:206F, 1:771–772, 1:773
- alpha-tubulin (A-lattice) 2:539, 2:539–540
- Alps *see* actin-like proteins (Alps)
- ALPS1B *see* autoimmune lymphoproliferative syndrome (ALPS1B)
- ALRs *see* AIM2-like receptors (ALRs)
- ALS *see* amyotrophic lateral sclerosis (ALS)
- ALT *see* alternative lengthening of telomeres (ALT)
- alternative end-joining (aEJ) *see* microhomology-mediated end-joining pathway (MMEJ)
- alternative lengthening of telomeres (ALT) 1:423, 1:428, 1:430–431, 1:462
- alternative NHEJ (aNHEJ) 1:452, 3:558–559
- alternative pathway 1:728, 1:728F
see also classical pathway; lectin pathway
- C3bBb C3 convertase complex 1:729
- convertase complexes 1:729
- instability of thioester bond in C3 1:728–729
- alternative reading frame protein (ARF protein) 3:447, 3:452–453
 in development 3:455
- alternative splicing 1:377, 1:379F, 1:495, 1:500–501, 1:501F, 1:503, 1:505–506, 1:505F, 1:506, 1:507, 1:508, 1:509, 1:556F, 1:564T
- cis-acting sequences and trans-acting factors 1:505–506
- signaling pathways 1:506–507
- snRNP biogenesis 1:507
- transcription, chromatin structure, and splicing 1:506
- altius 2:602
- Alu transcription 1:364–365
- Alzheimer's disease (AD) 1:241, 1:661, 1:675, 1:681, 1:696, 1:793, 2:455, 2:481, 2:482, 3:133, 3:419, 4:266–267
- amyloid production 1:661–662
- APP and presenilin mutations 1:662
- β -secretase
 identification, structure, and mechanism 1:662
 inhibitors 1:663
 knockout mice 1:662–663
 other substrates 1:663
 prospects for therapeutics 1:663–664
- γ -secretase
 identification and mechanism 1:664
 inhibitors and modulators 1:665
 other substrates 1:664–665
 prospects for therapeutics 1:666
 structure elucidation 1:665–666
 presentation and pathology 1:661
- AMAMTS proteases 1:754–755
- amaurotic familial idiocy 1:767
- ambisense RNA viruses 1:517, 1:518T, 1:521
- AMBRA1 *see* activating molecule in BECLIN1-regulated autophagy (AMBRA1)
- AMD *see* age-related macular degeneration (AMD)
- American Oil Chemistry Society (AOCS) 1:153, 1:180
- AMH *see* anti-Müllerian hormone (AMH)
- amidation 1:86–87
- amiloride 2:762
- amine-containing phospholipids synthesis 1:162
 in bacteria 1:162
 in eukaryotes 1:164–166, 1:165F

- amino acid signaling 1:766–767
- amino acids
- mTOR signaling regulation by 3:248–249
 - synthesis 2:329
- aminoacyl-tRNA (AA-tRNA) 1:309, 1:319F, 1:320, 1:323, 1:326, 1:384, 1:530–531, 1:551F
- aminoacyl-tRNA synthetase (aaRSs)
- 1:22–23, 1:309, 1:316–317, 1:316–317, 1:317F, 1:319F, 1:323, 1:324–325, 1:325, 1:330–333T, 1:375–377, 1:547, 1:547–548
- aminopeptidases (AP) 1:734–735, 1:747
- amino-terminal ligand-binding domain 3:57
- AML *see* acute myelogenous leukemia (AML); acute myeloid leukemia (AML)
- ammonium chloride (NH₄Cl) 2:768–769
- ammonium sulfate concentrations 1:70
- amoeboid migration 2:720, 2:723, 3:212
- amoeboid–mesenchymal transition (AMT) 3:226
- AMOTL1/2 *see* angiomin-like proteins 1 and 2 (AMOTL1/2)
- AMP *see* adenosine monophosphate (AMP)
- AMPA *see* α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)
- amphipathic N-terminal helix domains (N-BAR domains) 2:491, 2:498
- amphiphysin 1 2:498
- amphiregulin 1:641
- amphisome 2:246
- AMPK *see* adenosine monophosphate activated kinase (AMPK)
- AMPs *see* antimicrobial proteins and peptides (AMPs)
- AMPylation *see* Rab1 by adenylation (AMPylation)
- AMRF *see* action myoclonus-renal failure syndrome (AMRF)
- AMT *see* amoeboid–mesenchymal transition (AMT)
- amyloid 1:113
- amyloid plaques 1:241, 1:661
- amyloid precursor protein (APP) 1:675, 1:681, 1:696, 2:453, 2:481
- amyloid production 1:661–662
- amyloid β -peptide (A β peptide) 1:661, 1:675, 1:681, 2:456
- amyloid β -protein precursor (APP) 1:661–662, 1:662
- amyloidosis 1:113
- amyloid- β precursor protein (APP) 1:344
- amylopectin 1:263, 1:264F, 1:265, 1:267
- amyoplast 1:265, 2:325–326
- amylose 1:263, 1:265, 1:269–270
- amyotrophic lateral sclerosis (ALS) 1:115, 3:231
- genetic basis of ALS 1:116
- Anabaena peroxidase 1:125
- Anakinra 3:29
- anaphase 3:481
- CPC during 3:475
- anaphase-promoting complex/cyclosome (APC/C) 2:644, 3:401, 3:471, 3:482, 3:499
- Ana-plastic large-cell lymphoma (ALCL) 3:64
- anaplastic lymphoma kinase (ALK) 3:16
- anauxetic dysplasia (AD) 3:294–295
- ANCA *see* anti-neutrophil cytoplasmic antibodies (ANCA)
- anchoring 2:653
- ANCHR *see* abscission/NoCut checkpoint regulator (ANCHR)
- Anderson-Fabry disease 1:778
- aneuploidy 3:471–472, 3:480, 3:486
- Ang *see* angiopoietins (Ang)
- angioedema 1:748
- angiogenesis 1:145, 1:699, 1:700, 2:728, 3:218–219, 3:220, 4:102–116, 4:102, 4:119
- arterial vessels 4:114–115
 - capillary structure 4:104, 4:104F
 - developmental vasculogenesis 4:102–104
 - hypoxia sensing mechanism 4:109F
 - hypoxic induction 4:109–110
 - intussusceptive angiogenesis 4:114
 - meprins in 1:695
 - primary factors controlling 4:104
 - angiogenic growth factors 4:104
 - capillary growth factor ligand, receptor, and co-receptor domain structures 4:105F
 - FGF-VEGF-Ang-PDGF networking 4:108–109
 - vascular remodeling factors 4:107
 - sprouting angiogenesis 4:110
 - vascular evolution 4:102
 - vasculogenesis and 4:127–128
 - venous vessels 4:114–115
- angiogenic growth factors 4:104
- FGFs 4:104
 - VEGFs 4:106
- angiomin-like proteins 1 and 2 (AMOTL1/2) 3:102
- angiopoietins (Ang) 4:102, 4:107
- angiopoietin-1 (Angpt1) 3:16
 - ligands 4:107
 - receptors 4:107
 - Tie signaling 4:107–108
- angiotensin 1:746–747, 1:747
- angiotensin I converting enzyme (ACE) 1:706
- angiotensin II Type 1A receptor (AT₁ AR) 2:444
- angiotensin-converting enzyme (ACE) 1:746–747
- Angpt1 *see* angiopoietin-1 (Angpt1)
- aNHEJ *see* alternative NHEJ (aNHEJ)
- anillin 3:495–496
- animal viruses 2:809F
- endocytic entry pathways for 2:813F
 - stepwise host cell entry 2:811F
- anion exchange 1:66
- anionic lipid 1:194
- anisotropy 2:64
- ankyrin repeats (ANK) 3:107
- anoikis 3:227, 3:335–340
- ROS and 3:336
- as cell survival factors 3:337–338
 - control 3:338F
 - as mediator 3:336–337
 - tumor cells 3:335, 3:335–336
 - metabolic coupling in 3:338–339
- ANP *see* atrial natriuretic peptide (ANP)
- ANPCR *see* ANP clearance receptor (ANPCR)
- ANT *see* adenine nucleotide translocase (ANT)
- antagonists 3:395–396
- anterior–posterior cortical polarity (AP cortical polarity) 2:741–742
- anterograde carriers (ACs) 2:175
- anterograde IFT motor 2:665–666
- anterograde transport 2:168, 2:354, 3:236
- ANTH domain *see* AP180 N-terminal homology domain (ANTH domain)
- anti-apoptotic proteins 3:331, 3:370, 3:607
- anti-apoptosis proteins of viral origin 3:332
 - Bcl-2 family 3:331–332
 - IAPs 3:330–331
- antibacterial factors (ABFs) 2:847
- antibiotics 2:803–804
- synthesis by fungi 2:255
- antibody(ies) 1:138–139, 1:141–142
- angiogenesis and vascular blockade/transport 1:145
 - antibody-based methods 4:178–179
 - engineering multiple specificity in antibody fragments
 - bifunctional antibodies 1:143
 - bispecific antibodies 1:143
 - fragments design 1:142–143
 - high-value clinical applications 1:145
 - pathogen neutralization, antiviral therapy, and vaccines 1:145
 - T-cell targeting and immune checkpoint inhibitors 1:145
- image-guided surgery 1:145–146
- for *in vivo* applications 1:142
- intact antibodies vs. fragments 1:143
- libraries 1:143–144
- bacteriophage display 1:144
 - cell-surface libraries 1:144
 - maturation 1:144–145
 - transgenic mice 1:144
 - production, stability, and expression levels 1:145
- antibody-dependent cellular cytotoxicity (ADCC) 2:835
- antibody-drug conjugates (ADCs) 1:139, 1:143
- antibody-secreting cells (ASCs) 3:629, 3:668–669F, 3:669
- anticancer 1:241
- agents 1:435, 3:5
 - therapies 3:399, 3:401
- anti-DNP antibodies 1:18–19
- antigen (Ag) 3:519, 3:632, 3:759–760
- capture 3:746
 - presentation 1:764, 2:177, 3:746
 - endothelial cells properties 3:639–640

- antigen (Ag) (*continued*)
 endothelial in pathology 3:641–642
 of human DC subpopulations 3:596–597
 of mouse DC subpopulations 3:592–593
 processing 3:746
 receptors 3:550F
 genes 3:551F
- antigen presenting cells (APCs) 1:682,
 2:836, 3:6, 3:521, 3:567,
 3:588–589, 3:589, 3:592–593,
 3:594–595, 3:596–597, 3:632,
 3:650, 3:677, 3:695, 3:741,
 3:770
- antigen processing and presentation 3:589
 autophagy 3:591–592
 classical MHC class II processing 3:589–590
 classical MHC class I processing 3:589
 cross-presentation 3:590–591
 cytosolic pathway 3:590–591
 vacuolar pathway 3:590
- antigen targeting 3:593–594, 3:594,
 3:595–596
- antigen-specific immunity, endothelial
 function in 3:639
 antigen recognition and responses 3:642
 EC Ag presentation emerging and putative
 roles 3:640–641
 endothelial antigen presentation in
 pathology 3:641–642
 endothelial cells antigen presenting
 properties 3:639–640
 model for ILPs function in Ag recognition
 and response 3:643F
- antimicrobial proteins and peptides (AMPs)
 2:771
- anti-Müllerian hormone (AMH) 3:37
- anti-neutrophil cytoplasmic antibodies
 (ANCA) 3:755
- anti-recombination 1:464
- antithrombin III (AT) 1:710
- antitumor 1:753, 1:755, 1:757
- antiviral therapy 1:145
- AOCS *see* American Oil Chemistry Society
 (AOCS)
- AOS *see* Adams–Oliver syndrome (AOS)
- AP *see* adaptor protein (AP);
 aminopeptidases (AP)
- AP cortical polarity *see* anterior–posterior
 cortical polarity (AP cortical
 polarity)
- AP endonuclease 1 (APE1) 1:442–443
- AP sites *see* apurinic sites (AP sites)
- AP1 *see* activator protein 1 (AP1)
- AP180 N-terminal homology domain
 (ANTH domain) 2:455, 2:507–508
- APAF-1 *see* adaptor protein apoptotic
 protease activating factor 1 (APAF-1)
- APAF1 *see* apoptotic peptidase activating
 factor 1 (APAF1)
- Apaf-1 *see* apoptotic protease-activating
 factor (Apaf-1)
- Apaf1 *see* apoptotic protease activating
 factor 1 (Apaf1)
- APC *see* adenomatous polyposis coli (APC);
 antigen presenting cells (APCs)
- APC/C *see* anaphase-promoting complex/
 cyclosome (APC/C)
- APCI *see* atmospheric pressure chemical
 ionization (APCI)
- APE1 *see* AP endonuclease 1 (APE1)
- Aph-1 1:665
- API *see* active pharmaceutical ingredient
 (API)
- apical early endosomes (AEE) 2:192–193
- apical membrane 1:252, 1:257
- apical recycling endosomes (ARE)
 2:745–746
- apical–basal cell polarity (ABCP) 3:99
- apical–basal polarity 2:742
 proteins interplay organizing 2:745
- aPKC *see* atypical protein kinase C (aPKC)
- AP-MALDI *see* atmospheric pressure matrix-
 assisted laser desorption/ionization
 (AP-MALDI)
- AP-MS *see* affinity purification mass
 spectrometry (AP-MS)
- ApoB *see* apolipoprotein B (ApoB)
- APOBEC3G restriction factor 2:836
- ApoE *see* apolipoprotein E (ApoE)
- apoE receptor 2 (apoER2) 2:419–420
- apolipoprotein B (ApoB) 1:176
- apolipoprotein E (ApoE) 2:418–419
- Apollon 1:708, 1:709, 3:395
- apoptosis 3:4, 3:327–334, 3:365F, 3:367T,
 4:95
 anti-apoptotic proteins 3:331
 Bcl-2 family 3:331–332
 IAPs 3:331
 of viral origin 3:332
 anti-apoptotic vs. pro-apoptotic
 machinery 3:328–330
 caspases *see* caspases
 extrinsic pathway 3:327–328
 inhibition of 3:369–370
 intrinsic pathway 3:328
 pericyte 4:77–78
 pro-apoptotic machinery 3:329–330,
 3:330
 B-cell lymphoma-2 (Bcl-2) family
 3:330
 caspases 3:329–330
 lipids 3:330–331
- apoptosis signal regulating-kinase 1 (ASK1)
 2:530
- apoptosis-associated speck-like protein
 containing caspase recruitment
 domain (ASC) 2:523–524
- apoptosis-inducing factor, mitochondrion-
 associated 1 (AIFM1) 3:342
- apoptosome 1:708, 3:69–70, 3:328,
 3:328F, 3:329, 3:330, 3:345,
 3:368–369, 3:368–370, 3:369–370,
 3:369F
- apoptotic bodies 3:327, 3:328,
 3:328F
- apoptotic caspases 3:329, 3:364, 3:364–366,
 3:369–370
- apoptotic peptidase activating factor 1
 (APAF1) 3:343
- apoptotic protease activating factor 1
 (Apaf1) 1:615, 3:69–70, 3:328
- apoptotic/apoptosis 2:766, 2:773, 3:67,
 3:379, 3:399, 3:402, 3:440, 3:582
see also B-cell lymphoma 2 proteins
 (Bcl-2 proteins)
- bodies 2:304
- cell death ligand 3:393–394
- cells clearance by MZ macrophages 3:763
- differentiation program 2:773
- intrinsic and extrinsic apoptotic pathways
 3:392F
- mimicry 3:734
- mitochondrial control 3:342–343
- oncogenes, tumor suppressor genes and
 3:417
- pathway 3:379
 regulation by XIAP 3:391
- APP *see* amyloid precursor protein (APP);
 amyloid- β precursor protein (APP);
 amyloid β -protein precursor (APP)
- APP intracellular domain (AICD)
 1:661–662
- APPL *see* adaptor protein containing PH
 domain PTB-domain and leucine
 zipper motif (APPL)
- APR *see* acute phase response (APR)
- APs *see* Assembly polypeptides (APs)
- aptamer 1:375, 1:377, 1:382, 3:668–669F,
 3:669
in vitro selection of 1:375
 and riboswitches 1:375–377
- apurinic sites (AP sites) 3:462
- AQUA peptides 4:169
- Arabidopsis thaliana*
 regulation of root growth by auxin in
 4:128–131
- arachidonic acid 1:194–195, 1:195
- ARACNe *see* Algorithm for the
 Reconstruction of Accurate Cellular
 Networks (ARACNe)
- Arch *see* archaerhodopsin-3 (Arch)
- archaea 1:155, 1:160–161, 2:588, 2:800
 cytokinesis 2:588–590
 ESCRT-based cell division 2:590F,
 2:592
 transcription and regulation in 1:477–478
- Z-ring
 placement at mid-cell 2:590–592
 structure and constriction mechanism
 2:588–590
- archaerhodopsin-3 (Arch) 2:151
- archaerhodopsins (Archs) 2:148–149
- Archs *see* archaerhodopsins (Archs)
- ARE *see* apical recycling endosomes (ARE);
 AU-rich element (ARE)
- AREs *see* adenylate/uridylate-rich elements
 (AREs)
- ARF *see* adenosine diphosphate ribosylation
 factor (ARF)
- Arf *see* ADP-ribosylation factor (Arf)
- ARF protein *see* alternative reading frame
 protein (ARF protein)
- ARF tumor suppressor protein 3:5
- ArfGTPase activating domain (ArfGAP)
 2:497

- Arf-like families (Arl families) 2:194–195, 2:446
- arfophilins *see* Rab coupling proteins
- Arg-Gly-Asp sequence (RGD sequence) 2:723–724
- arginine-serine (RS) repeat motifs 1:497–498
- Argonaute (AGO) 1:354, 1:356–357F, 1:357F, 1:358, 1:358–359, 1:359, 3:321–322
- loading of miRNA–miRNA* duplexes 1:358
- ARH *see* autosomal recessive hypercholesterolemia (ARH)
- ARID *see* autosomal recessive intellectual disability (ARID)
- Arl families *see* Arf-like families (Arl families)
- Arl-13/ARL13b compartment 2:667
- arLCA *see* autosomal Leber congenital amaurosis (arLCA)
- AROS *see* active regulator of SIRT1 (AROS)
- Arp2/3 complex *see* actin-related protein 2/3 complex (Arp2/3 complex)
- ARPKD *see* autosomal recessive polycystic kidney disease (ARPKD)
- array tomography 2:46–47
- arrestin family 2:456
- arrestin related trafficking adaptors (ARTS) 2:456
- arRP *see* autosomal recessive retinitis pigmentosa (arRP)
- ARSACS *see* autosomal recessive spastic ataxia of Charlevoix-Saguenay (ARSACS)
- ARTEMIS–DNA-PKcs complex, nucleolytic processing by 1:451
- arterial thrombosis 1:718
- arterial vessels 4:114–115
- arthritis 3:25–26, 3:29
- ARTS *see* arrestin related trafficking adaptors (ARTS)
- 1-aryl-4-alkoxypiperidin-4-yl groups 1:43–44, 1:44F
- Arylsulfatase A 1:774, 1:784–785, 1:785
- ASC *see* apoptosis-associated speck-like protein containing caspase recruitment domain (ASC)
- Ascaris* pepsin inhibitor 1:684
- ASCE *see* additional strand catalytic E (ASCE)
- Ascl-2 *see* achaete-scute homologue 2 (Ascl-2)
- ascorbate 1:87
- ASCs *see* antibody-secreting cells (ASCs)
- ASFV *see* African swine fever virus (ASFV)
- ASGR *see* Ashwell–Morrell receptor (ASGR)
- ASMase *see* acid sphingomyelinase (ASMase)
- Ashwell–Morrell receptor (ASGR) 2:344
- ASK1 *see* apoptosis signal regulating-kinase 1 (ASK1)
- ASL *see* airway surface liquid (ASL)
- Asn-Pro-x-Tyr motifs (NPxY motifs) 2:480
- asparaginyl endopeptidase (AEP) 3:590
- aspartic acid tRNA (tRNA^{Asp}) 1:318
- aspartic peptidases 1:681
- aspartic protease 1:739
- cathepsin D 1:753
- aspirin 1:133F
- assembly chaperone 1:589F, 1:590–591
- assembly polypeptides (APs) 2:450
- asthma 3:26
- astral microtubules 3:480
- astrocyte 2:569
- reactive 3:800
- asymmetric cell division 2:656
- AT *see* α 1-antitrypsin (AT); antithrombin III (AT)
- AT₁ AR *see* angiotensin II Type 1 Areceptor (AT₁ AR)
- ataxia telangiectasia mutated protein (ATM protein) 3:103, 3:473, 3:781
- ataxia-telangiectasia mutated kinase (ATM kinase) 3:504–505, 3:816
- atazanavir 1:743
- ATF4 *see* activating transcription factor-4 (ATF4)
- ATF6 *see* activating transcription factor 6 (ATF6)
- ATG *see* autophagy-related genes (ATG)
- Atg7 *see* autophagy-related protein 7 (Atg7)
- Atg8 *see* LC3/GABARAP/GATE16 (Atg8)
- ATG8/LC3 1:584T, 3:591
- and ATG12 1:588
- ATG12 1:584T, 3:591
- and ATG8/LC3 1:588
- Atg12–Atg5–Atg16L complex 2:503
- ATGL lipase 2:276
- atherosclerosis 3:735
- plaque formation 1:177–178
- ATL *see* Atlantin (ATL)
- Atlas3D software 2:49
- Atlantin (ATL) 2:279–281
- ATM kinase *see* ataxia-telangiectasia mutated kinase (ATM kinase)
- ATM protein *see* ataxia telangiectasia mutated protein (ATM protein)
- ATM-related damage response 3:460–461, 3:461F
- atmospheric pressure chemical ionization (APCI) 4:204–205, 4:204F
- atmospheric pressure matrix-assisted laser desorption/ionization (AP-MALDI) 1:155
- atomic resolution 2:9, 2:22
- ATP *see* adenosine triphosphate (ATP)
- ATP hydrolysis 1:437, 1:488, 1:489, 1:490–491, 1:490F, 1:500, 1:543, 1:589–590, 2:3–4
- ATPase *see* adenosine triphosphatase (ATPase)
- ATPase associated with various cellular activities protein (AAA protein) 2:592
- ATPase Brahma (Brm) 3:787
- ATPases *see* AAA-adenosine triphosphatase (ATPases)
- ATP-dependent chromatin remodeler 1:488
- ATRA *see* all-trans retinoic acid (ATRA)
- ATR-interacting protein (ATRIP) 3:461
- ATRIP *see* ATR-interacting protein (ATRIP)
- ATR-related damage response 3:461, 3:461F
- effector proteins 3:461–462
- atrial natriuretic peptide (ANP) 1:750, 3:58
- atypical memory 3:570–571
- atypical protein kinase C (aPKC) 3:100
- ATZ *see* α 1-antitrypsin Z (ATZ)
- ‘Auger’ electrons 2:16
- AU-rich element (ARE) 1:341, 1:343–344
- aurora B kinase 2:644, 3:469, 3:496
- aurora C kinase 3:469
- autoactivation 1:683
- auto-destructive activity 2:651–653
- autofluorescence FLIM 2:110
- autofocus mechanism 4:238–239
- Autographa californica* (alfalfa looper) multiple nuclear polyhedrosis virus (AcMNPV) 1:59
- autoimmune disease 2:728–729, 3:71–72, 3:354, 3:360–361
- DCs in 3:747–748
- autoimmune lympho-proliferative syndrome (ALPS1B) 2:226
- autoimmune regulator (AIRE) 3:618–619, 3:747
- auto-immune regulator transcription factor (AIRE) 3:525
- autoinhibitory region (AIR) 3:704
- autolysosome 1:789
- automated amino acid analyzer 1:75
- autophagic cargo 3:407–408
- autophagosomes 1:588, 1:763–764, 1:764–765, 1:765, 1:771, 1:795, 2:2–3, 2:243–244, 2:822, 3:384, 3:591
- breakdown products of 1:763–764
- characteristics 2:243–244
- machinery 2:244–246
- maturation and consumption 2:246
- membranes 2:246
- autophagy 1:618, 1:764, 1:764–765, 1:766, 1:766–767, 1:768, 1:771, 1:795, 2:177, 2:203, 2:751, 2:755, 2:771–772, 2:790–791, 2:822, 3:4, 3:278, 3:384, 3:404–410
- antigen processing and presentation 3:591–592
- autophagic cargo 3:407–408
- autophagy-related cellular compartments 3:407
- bacterial embrace of 2:782
- cargo recognition 3:405–406
- in energy economy and nutrient-sensing 1:765–766
- induction 3:404–405
- lysosomal delivery 3:406
- lysosomal diseases and 1:770
- macroautophagy 3:405F
- mechanism 2:244F
- modulation of host cell autophagy pathways 2:781F
- mTORC1 suppressing 3:246
- pathophysiology 3:408–409
- pathway 1:117, 3:404
- phosphoinositides in autophagy regulation 2:515–516, 2:516F
- recycling 3:406–407

autophagy (*continued*)
autophagy-related cellular compartments 3:407
autophagy-related genes (ATG) 2:244–245, 3:404
autophagy-related protein 7 (Atg7) 3:781
autophosphorylation 1:543, 3:1–2, 3:11, 3:371–372
autopsy studies 1:793
autosomal dominant polycystic kidney disease (ADPKD) 2:670, 2:670–671
autosomal Leber congenital amaurosis (arLCA) 2:747–748
autosomal recessive hypercholesterolemia (ARH) 2:423–425, 2:455
autosomal recessive intellectual disability (ARID) 2:479
autosomal recessive polycystic kidney disease (ARPKD) 2:670, 2:670–671
autosomal recessive retinitis pigmentosa (arRP) 2:747–748
autosomal recessive spastic ataxia of Charlevoix-Saguenay (ARSACS) 2:286–287
autumnal senescence 2:327
auxiliary protein factors 1:503, 1:503–504, 1:505, 1:508
auxiliary translocation factors
RAMP4 1:537T, 1:538
Sec62 1:537T, 1:538
Sec63 1:537T, 1:538, 1:543
TRAM 1:537T, 1:538
TRAP 1:537T, 1:538
auxins 2:254
AV *see* Aichivirus (AV)
Avastin® 1:138
avidity 2:425, 2:427
AVPI *see* Ala-Val-Pro-Ile (AVPI)
Axin 3:90, 3:90F, 3:91, 3:92, 3:93, 3:93F, 3:94
Dvl–Axin interaction 3:91
inactivation 3:92
Axin complex saturation 3:92
axon regeneration
protein kinase regulation 3:237–238
regulation of microtubule motor proteins 3:238
axon structure and axonal transport, injury changes 3:235–236
calcium influx
coordinating injury response 3:236
transforming injured axon to regenerative growth 3:236–237
axonal transport 2:625
abnormalities 2:611
axonemal dyneins 2:620, 2:626
see also cytoplasmic dynein (DYNC)
axonemal dynein-related ciliopathies 2:631–632
DHC genes diversity 2:627–628, 2:629F
HCs in cilia 2:628–630
isoforms of dyneins in *Chlamydomonas* 2:628T
localization and asymmetry in 2:628
molecular arrangement in axoneme 2:626–627

nucleotide-induced conformational change 2:627
axoneme 2:660, 2:663
molecular arrangement in 2:626–627
3D structure and dynein arrangement in 2:627F
axophilic migration 2:728
N-azidoacetylgalactosamine 1:131–132
azobenzene 2:149, 2:151, 2:151–152
azurophilic granules (AG) 2:770
azurosome 3:752–753

B

B and T lymphocyte attenuator (BTLA) 3:771–772
B cell activating factor (BAFF) 3:616–617
B cell lymphoma 2 (Bcl 2) 1:674, 3:104
B cell lymphoma-6 (Bcl-6) 3:541, 3:542
B cell receptor 3:518
B cell-activating factor (BAFF) 3:583
receptor pathways 3:583
B3GALNT2 *see* β 1,3-N-acetylgalactosaminyltransferase 2 (B3GALNT2)
B3GLCT *see* β 1,3-glucosyltransferase (B3GLCT)
BACE *see* β -site APP-cleaving protein (BACE)
bacillus Calmette–Guérin (BCG) 3:773
infection 3:743–745
Bacillus subtilis (*B subtilis*) 1:57
back focal plane (BFP) 2:67, 2:122–123
backscattered electrons 2:16
BACs *see* bacterial artificial chromosomes (BACs)
bacteria 1:162, 2:149, 2:588–596, 2:800
bacterial filament
involved in DNA segregation 2:591–592F
structure involved in cell division 2:589F
cell shape determination 2:593, 2:594F
actin-like MreB in rod-shaped cells 2:593
filamentous proteins promoting curvature 2:593
cytokinesis 2:588–590
cytoskeletal elements functions 2:588F
DNA segregation 2:592
of chromosomal DNA or plasmids by WACAs 2:593
of plasmids by actin-like ParM, 2:592
of plasmids by tubulin-like TubZ 2:592–593
leukocyte responses to 4:98
organization of intracellular membrane compartments 2:593–594
phospholipids synthesis in
amine-containing phospholipids synthesis 1:162
non-amine-containing phospholipids synthesis 1:162–164
structural and functional organization 2:93–94
Z-ring
placement at mid-cell 2:590–592
structure and constriction mechanism 2:588–590
bacterial (C family) DNA polymerase 1:398–400
crystal structure of 1:400
structure and organization of 1:399F
bacterial actin ancestors 2:565
bacterial artificial chromosomes (BACs) 1:513
bacterial dynamin-like proteins (BDLPs) 2:593–594
bacterial effectors subvert Rab function 2:405
bacterial enhancer action 1:488–489
bacterial enhancer-binding proteins (bEBPs) 1:490–491, 1:490F
bacterial expression systems 1:56
see also cell-free expression systems; mammalian expression systems; yeast expression systems
B subtilis 1:57
ColeE1-family of replicons 1:56–57
gram-negative *P. fluorescens* 1:57
inclusion bodies 1:57
IPTG 1:57
L. lactis 1:57–58
pET-15b 1:56F
solubility-enhancing protein domains 1:57T
bacterial IcsA protein 2:781
bacterial invasion 2:762
bacterial lipopolysaccharide 3:371, 3:672
bacterial pathogens 2:784
cellular invasion by 2:784–793, 2:784F
cellular responses to bacterial invasion 2:790–791
entry mechanisms 2:784–785
intracellular lifestyle and host intracellular trafficking 2:788–790
stages 2:784
targeting of host organelles by 2:778F
uptake interference by 2:776
blocking uptake via bacterial pore-forming toxin deposition 2:776
phagocytosis inactivation after intimate contact with pathogens 2:776–777
bacterial phosphotransferase system, structural proteomics of 1:101–103
bacterial primase 1:405, 1:406F
bacterial protein toxins 2:794–795
exotoxins
altering cytoskeleton 2:795–796
altering DNA 2:797
altering signal transduction 2:794–795
blocking protein synthesis or protein folding 2:797
interfering with intracellular trafficking 2:796–797
bacterial replisome 1:408–410
bacterial RNA polymerase holoenzyme 1:469–471
bacterial rRNAs 3:294

- bacterial subversion of phagocytic killing
 bacterial embrace of autophagy 2:782
 breaking free to avoid host cell killing 2:780
 bypass and exploitation of host autophagic clearing of microorganisms 2:780–782
 modulation of host cell autophagy pathways 2:781F
 perspective 2:782
 replication vacuole
 endocytic pathway manipulation to support bacterial replication 2:777
 multi-organelle interactions 2:780
 secretory pathway manipulation to support bacterial replication 2:778–780
 targeting of host organelles by bacterial pathogens 2:778F
 uptake interference by bacterial pathogens 2:776
 blocking uptake via bacterial pore-forming toxin deposition 2:776
 phagocytosis inactivation after intimate contact with pathogens 2:776–777
 bacterial super-antigens 3:642
 bacterial toxins 2:794
 bactericidal/permeability-increasing protein (BPI protein) 2:771, 3:751
 bacteriophage display 1:144
 baculoviral IAP repeat containing (BIRC) 3:343
 baculoviral inhibitor of apoptosis repeat (BIR) 1:708, 3:390, 3:469–470, 3:714
 Baculovirus 1:59
 BAD *see* Bcl-2-associated death promoter (BAD); BCL2-associated agonist of cell death (BAD)
 BAF *see* Barrier to Autointegration Factor (BAF); Brg1/Brm-associated factors (BAFs)
 BAFF *see* B cell activating factor (BAFF); B cell-activating factor (BAFF)
 Bak activation, conformational changes associated with 3:382
 BAM *see* β -amylases (BAM)
 'Band B' CFTR 1:254–255
 band separation/unmixing 2:89–90
 bands (juvenile) 3:751
 Bannayan–Riley–Ruvalcaba syndrome (BRRS) 3:419
 BAR domains *see* Bin-Amphiphysin-Rvs (BAR) domains
 BAR + pleckstrin homology domains (BAR-PH domains) 2:491
 barbed end 2:683–685, 2:685–686
 barbed end-capping proteins 2:557
 trackers 2:557
 Bardet Biedl Syndrome (BBS) 2:667, 2:671
 bariatric surgery 2:805
 baroreflex 1:746
 BAR-PH domains *see* BAR + pleckstrin homology domains (BAR-PH domains)
- Barrier to Autointegration Factor (BAF) 2:315
 barriers and tissue cells 3:519
 basal body 2:655, 2:660, 2:661–663, 2:664, 2:665
 basal cell carcinomas (BCCs) 3:83
 basal cell nevus syndrome (BCNS) 3:81–83
 basal lamina 2:677, 2:695
 base excision repair pathway (BER pathway) 1:442, 3:462, 3:463F
 abasic site incision and DNA strand break ends processing 1:446–447
 basic overview 1:442–443
 correct undamaged nucleotide insertion 1:447
 DNA base damage recognition and removal 1:443
 HhH DNA Glycosylases 1:444–445
 human DNA glycosylases and specificities 1:445T
 Methylpurine DNA Glycosylase 1:446
 NEIL DNA Glycosylases 1:445–446
 UNGs 1:443–444
 DNA nick ligation 1:447
 genome stability maintenance and role of 1:442
 knockout mouse models 1:448
 long-patch BER pathway 1:447–448
 protein levels regulation 1:448
 short-patch BER 1:444F
 small DNA base lesions 1:443F
 base pairs (bp) 2:835
 basement membrane (BM) 2:694–695, 2:743–744, 3:223–224
 ECM proteins associated with 2:697–698
 interaction model of principle components 2:696F
 basic fibroblast growth factor (bFGF) 3:51–52, 3:210
 basic leucine zipper transcription factor (Batf) 3:542
 basic leucine zipper transcriptional activator (bZIP transcriptional activator) 2:529–530
 basic-helix-loop-helix (bHLH) 3:107
 basic-rich stretches (BRS) 3:650
 basolateral membrane 1:252, 1:255
 basophils 2:226–227
 batch adsorption 1:70
 Batf *see* basic leucine zipper transcription factor (Batf)
 BATF *see* B-cell-activating transcription factor (BATF)
 Bax activation, conformational changes associated with 3:382
 Bax dimer formation models 3:383F
 Bax protein *see* Bcl-2 associated X protein (Bax protein)
 BAX *see* BCL2-associated X protein (BAX)
 Bax-interacting factor 1 (Bif-1) 2:245
 Bayes probability theorem 4:222
 Bayesian formalism 4:253–254
 Bayesian inference (BI) 4:257–258
 Bayesian networks 4:222
 BBS *see* Bardet Biedl Syndrome (BBS)
 BBSIM *see* bessel beam structured illumination microscopy (BBSIM)
- BBSome 2:667
 BC *see* breast cancer (BC)
 BCCs *see* basal cell carcinomas (BCCs)
 BCECF *see* 2',7'-bis-(2-carboxyethyl)-5-(and-6)-carboxyfluorescein (BCECF)
 B-cell lymphoma 1:367, 1:768, 1:773, 1:785, 3:331
 B-cell lymphoma 2 proteins (Bcl-2 proteins) 2:531–532, 3:69–70, 3:379, 3:380F, 3:402
 Bax dimer formation models 3:383F
 Bcl-2 family of proteins 3:329, 3:330, 3:331–332
 BH3 profiling 3:386F
 and cancer 3:384–385
 classification and structure 3:379
 at endoplasmic reticulum 3:384
 mice deficient in BH3 proteins 3:384T
 physiology and regulation 3:382–384
 regulating MOMP 3:380–381
 conformational changes 3:382
 mechanism of action for 3:381F
 B-cell lymphoma-extra large (BclxL) 3:69–70, 3:402
 B-cell maturation antigen (BCMA) 3:62
 B-cell receptor (BCR) 3:528, 3:543–544, 3:583
 B-cell-activating transcription factor (BATF) 3:782
 BCG *see* bacillus Calmette–Guérin (BCG)
 Bcl 2 *see* B cell lymphoma 2 (Bcl 2)
 Bcl-2 associated X protein (Bax protein) 1:674
 Bcl-2 homology region (BH region) 3:379
 Bcl-2 proteins *see* B-cell lymphoma 2 proteins (Bcl-2 proteins)
 BCL2-associated agonist of cell death (BAD) 3:345–346
 Bcl-2-associated death promoter (BAD) 3:402
 BCL2-associated X protein (BAX) 3:345
 Bcl-6 *see* B cell lymphoma-6 (Bcl-6)
 BclxL *see* B-cell lymphoma-extra large (BclxL)
 BCMA *see* B-cell maturation antigen (BCMA)
 BCNS *see* basal cell nevus syndrome (BCNS)
 BCR *see* B-cell receptor (BCR); breakpoint cluster region (BCR)
 BCS *see* biopharmaceutical classification system (BCS)
 BCV *see* Brucella-containing vacuole (BCV)
 BDLPs *see* bacterial dynamin-like proteins (BDLPs)
 BDNF *see* brain-derived neurotrophic factor (BDNF)
 BDNF signaling *see* brain-derived neurotrophic factor signaling (BDNF signaling)
 BPDE *see* benzo(a)pyrene-diolepoxide (BPDE)
 BE *see* branching enzyme (BE)
 BEACH *see* Beige and Chediak (BEACH)
 beam-specimen interactions 2:15–16

- bEBPs *see* bacterial enhancer-binding proteins (bEBPs)
- BEC *see* blood endothelial cells (BEC)
- BEC-APC as unique planar APC 3:644–645
- Becker's disease 1:231
- Beclin-1 3:591
- Beige and Chediak (BEACH) 2:230
- benchmarking predictive models 4:259
- N*-benzoyl-L-tyrosyl-*p*-aminobenzoic acid (PABA-peptide) 1:691
- BER pathway *see* base excision repair pathway (BER pathway)
- bessel beam structured illumination microscopy (BBSIM) 2:94–95
- β 1,3-glucosyltransferase (B3GLCT) 1:633
- β 1,3-N-acetylgalactosaminyltransferase 2 (B3GALNT2) 2:349
- β -amylases (BAM) 1:268–269
- β -barrel domain 1:710
- β -catenin 3:2, 3:92–93, 3:93, 3:184
- stabilization 3:92, 3:93F
- TCF complex 3:93–94, 3:94F
- 'Wnt-off' state 3:89–91
- 'Wnt-on' state 3:91–92
- β common family (β c family) 3:24, 3:141
- β -families 2:627
- β -galactosidase 1:774
- β -glucans, dectin-1 activation by 3:696
- β -glucosidase (GBA11) 1:772
- β -mannanase 1:125
- β -oxidation pathway 1:296
- β -particles 1:264
- β -secretase
- see also* γ -Secretase
- identification, structure, and mechanism 1:662
- inhibitors 1:663
- knockout mice 1:662–663
- other substrates 1:663
- prospects for therapeutics 1:663–664
- β -site APP-cleaving protein (BACE) 1:662
- β -subunit of flavocytochrome b_{558}
- see* Gp91phox
- beta-tubulin (B-lattice) 2:539, 2:539–540
- structure 4:36
- bevacizumab (AvastinTM) 1:145
- BFA *see* brefeldin A (BFA)
- B-family DNA polymerase 1:398–400, 1:400
- bFGF *see* basic fibroblast growth factor (bFGF)
- BFP *see* back focal plane (BFP); biological force probe (BFP); biomembrane force probe (BFP)
- BH region *see* Bcl-2 homology region (BH region)
- BH3 interacting-domain death agonist protein (Bid protein) 1:674
- BH3 profiling 3:386F
- bHLH *see* basic-helix-loop-helix (bHLH)
- BI *see* Bayesian inference (BI)
- bicarbonate transport, CFTR and 1:257–259
- Bicaudal D (BicD) 2:625
- BicD *see* Bicaudal D (BicD)
- bicistronic construct assay 3:308, 3:308F
- bicistronic DNA constructs method 3:314
- biclustering 4:212, 4:212T, 4:213
- Bid protein *see* BH3 interacting-domain death agonist protein (Bid protein)
- bidomain equations 4:74
- Bif-1 *see* Bax-interacting factor 1 (Bif-1)
- bifunctional antibodies 1:143
- bifunctional DNA glycosylases 1:446–447, 1:446F
- biglycan 1:273–274
- bikunin 1:282
- BIL-DMK *see* Biphenyl-leucine-diazomethylketone (BIL-DMK)
- bile acids 1:178
- synthesis 2:254
- Bin-Amphiphysin-Rvs (BAR) domains
- 2:196, 2:476, 2:491, 3:680
- domain-containing proteins and cellular functions 2:491, 2:493F
- proteins 2:581
- as regulators of intracellular transport 2:497–499
- role in regulating NPF-mediated actin assembly 2:491–497
- role in regulating signaling 2:499–500
- structures from different subfamilies 2:492F
- well-characterized BAR domain proteins and functions 2:495–496T
- binding immunoglobulin protein (BiP)
- 1:539–540T, 2:529, 2:797
- chaperone 1:543
- for protein biogenesis and quality control in ER 1:543
- bioactive lipids 1:196
- bioavailability 1:741
- Biocarta 4:211
- biocatalysis 1:14–23
- biological catalysts 1:14
- enzymes 1:14
- chemical mechanisms of 1:18
- coenzymes and cofactors 1:19
- composition of 1:14
- cooperativity and allosteric regulation 1:17–18
- kinetics of 1:14–17
- sequential mechanisms 1:15–16
- substrate selectivity 1:17
- transition state analogs and catalytic antibodies 1:18–19
- ribozymes 1:19–22
- hammerhead 1:19–22
- ribosome 1:22–23
- RNaseP 1:22
- self-splicing introns 1:22
- biochemical modeling 4:14
- Arp2/3-mediated branch formation mechanism 4:17, 4:19–20
- ATP-hydrolysis 4:15–16
- cofilin 4:19
- cytoskeletal dynamics 4:19
- F-actin 4:19
- FRAP 4:21
- Listeria* 4:16–17
- nucleotide state profile 4:16F
- Open Model 4:17, 4:18F
- 3D lamellipodium 4:20
- treadmilling process 4:15F
- bioconjugate drugs 1:135–136
- biofilm 1:732, 1:735
- biogenesis
- of extracellular vesicle 2:470–471
- lysosomal 1:766, 1:768, 1:771, 1:775, 1:778, 1:779, 1:785
- and autophagy 1:770–771
- of lysosome-related organelles complex (BLOC) 2:241
- BLOC-1~3 2:237
- of miRNAs 1:364–365
- in animals 1:354–355, 1:359
- organelle 1:766–767
- of ribosomes 3:3
- of secretory proteins 1:535–544
- of Wnt protein 3:87, 3:87F
- bioinformatic platforms 1:94
- CATH database 1:95
- FSSP database 1:95
- HMMs 1:95
- SCOP database 1:94–95
- biological effect 4:182
- biological force probe (BFP) 3:655
- biological noise, dissecting 3:674–675, 3:674F
- biological process (BP) 4:247
- biological sample preparation 2:81–82
- biological system 4:80
- biologics 1:138–139
- biomacromolecule 1:128
- biomedical imaging
- retrograde transport and 2:438
- biomembrane force probe (BFP) 3:661
- biomodels database 4:75–76
- bioorthogonal labeling 1:131–132, 1:132F
- biopharmaceutical classification system (BCS) 1:135–136
- biosensors 2:53, 2:54, 2:58–59
- biosynthetic inhibitors 1:782
- biosynthetic recycling compartment (BRC) 2:175–176
- biotin synthesis 2:254
- BiP *see* binding immunoglobulin protein (BiP)
- Biphenyl-leucine-diazomethylketone (BIL-DMK) 1:130–131
- bipolar spindle formation 2:653, 2:653–655, 2:656–657
- BIR *see* baculoviral inhibitor of apoptosis repeat (BIR); break-induced replication (BIR)
- BIR repeat containing ubiquitin-conjugating enzyme (Bruce) 3:395
- BIRC *see* baculoviral IAP repeat containing (BIRC)
- birth, changes in gut microbiome in 2:801–802
- bis-(2-acetoxyethoxy)methyl (ACE) group 1:45–46, 1:45F
- 2',7'-bis-(2-carboxyethyl)-5-(and-6)-carboxy-fluorescein (BCECF) 2:108
- lifetime image of skin tissue stained with 2:109F
- bis(monoacylglycero)phosphate (BMP) 1:615–616, 2:205

- N,O*-Bis (trimethylsilyl)trifluoroacetamide (BSTFA) 4:203–204
- bispecific antibodies 1:143
- bistability 4:44, 4:46
- bivalency 3:786, 3:787F, 3:788, 3:789
- blastemal cells 3:221
- blastocystis 2:800
- B-lattice *see* beta-tubulin (B-lattice)
- blebbing motion 2:723
- bleeding
- defects in platelet function 1:717–718
 - defects in platelet production 1:717
 - hemophilia 1:718
 - TIC 1:718
 - Vitamin K deficiency 1:718
- BLIMP1 *see* B-lymphocyte-induced maturation protein 1 (BLIMP1)
- BLM *see* Bloom's syndrome protein helicase (BLM)
- BLOC *see* biogenesis of lysosome-related organelles complex (BLOC)
- blood clotting and inflammation 3:216–219, 3:217F
- blood coagulation 1:3, 1:714
- primary hemostasis
 - platelet production 1:714–715
 - platelet structure and function in 1:715–716
 - regulation 1:717
 - disorders 1:717
 - secondary hemostasis
 - initiation phase 1:716
 - propagation phase 1:716–717
- blood endothelial cells (BEC) 3:616, 3:633
- blood pressure
- homeostasis 1:746
 - proteolytic regulation
 - endothelin peptide family and proteases 1:750
 - natriuretic peptide family and proteases 1:750
 - RAS 1:746–747
- blood vessel 4:102
- Bloom's syndrome protein helicase (BLM) 3:501
- blue-green algae *see* cyanobacteria
- B-lymphocyte-induced maturation protein 1 (BLIMP1) 3:795
- BM *see* basement membrane (BM); bone marrow (BM)
- BmNPV *see* *Bombyx mori* (silkworm) nuclear polyhedrosis virus (BmNPV)
- BMP *see* bis(monoacylglycero)phosphate (BMP); bone morphogenetic protein (BMP)
- BMP/TGF β *see* bone morphogenetic protein/transforming growth factor- β (BMP/TGF β)
- BMV *see* Brome mosaic virus (BMV)
- BNP *see* B-type natriuretic peptide (BNP)
- Boltzmann probability 4:124
- bolus 1:732
- Bombyx mori* (silkworm) nuclear polyhedrosis virus (BmNPV) 1:59
- bona fide tumor suppressors 3:429
- bone marrow (BM) 3:624, 3:780
- see also* C-X-C motif chemokine 12 (CXCL12)
 - DC progenitors in 3:743
- bone morphogenetic protein (BMP) 1:656, 3:37, 3:99
- BMP-4 4:102–103
- bone morphogenetic protein/transforming growth factor- β (BMP/TGF β) 1:707
- Borealin 3:469, 3:470
- bortezomib 1:757
- bottom-up
- analyses 1:79–80
 - approach 4:81
 - proteomics 4:166–167
- botulinum neurotoxins 2:796
- bovine serum albumin (BSA) 2:35, 2:133–136
- bovine spongiform encephalopathy (BSE) 1:113–114, 2:839, 2:841
- bp *see* base pairs (bp)
- BP *see* biological process (BP)
- 3BP *see* 3-bromopyruvate (3BP)
- BPI protein *see* bactericidal/permeability-increasing protein (BPI protein)
- BPS *see* branchpoint sequence (BPS)
- bradykinin 1:748
- B-Raf 3:123, 3:126
- B-Raf mutated metastatic melanoma 3:126
- brain wiring 2:611–612
- 'rainbow' mouse 2:138
- brain-derived neurotrophic factor (BDNF) 3:800
- branching enzyme (BE) 1:266
- branchpoint sequence (BPS) 1:503, 1:503–504, 1:504F, 1:506
- BRC *see* biosynthetic recycling compartment (BRC)
- BRCA2/FANCD1 gene 3:781
- BRCT *see* C-terminal portion of BRCA-1 gene (BRCT)
- breakage–fusion–bridge cycle 1:431, 1:432
- break-induced replication (BIR) 1:462
- breakpoint cluster region (BCR) 2:499–500
- breast, efferocytosis in 3:374–376
- breast cancer (BC) 1:686–687
- diagnostic and prognostic values in 1:686–687
 - stem cells 1:367, 3:810
- brefeldin A (BFA) 2:169
- BRF1 *see* butyrate response factor-1 (BRF1)
- Brg1/Brm-associated factors (BAFs) 3:787
- broken-heart syndrome *see* myocardial stunning
- Brome mosaic virus (BMV) 1:329, 1:523–524, 2:9F, 2:10F
- 3-bromopyruvate (3BP) 1:241
- Brownian ratchet model
- elastic actin filament in 4:11F
 - formalism 4:10
- Brr2 1:497, 1:498–499, 1:500, 1:501
- BRRS *see* Bannayan–Riley–Ruvalcaba syndrome (BRRS)
- BRS *see* basic-rich stretches (BRS)
- Bruce *see* BIR repeat containing ubiquitin-conjugating enzyme (Bruce)
- Brucella*-containing vacuole (BCV) 2:777–778, 2:779
- Bruton's tyrosine kinase (BTK) 3:745
- BSA *see* bovine serum albumin (BSA)
- BSE *see* bovine spongiform encephalopathy (BSE)
- BTK *see* Bruton's tyrosine kinase (BTK)
- BTLa *see* B and T lymphocyte attenuator (BTLa)
- B-type cyclins 3:424
- B-type natriuretic peptide (BNP) 1:695, 1:750, 3:58
- Bub1 3:473
- budding 2:811–814, 2:815, 2:817
- bulk endocytosis 2:299
- Bunyaviridae* 2:824
- N*-butyldeoxynojirimycin 1:779–780
- butyrate response factor-1 (BRF1) 3:323
- bZIP domains 1:93
- bZIP transcriptional activator *see* basic leucine zipper transcriptional activator (bZIP transcriptional activator)

C

- C Clamp 3:94, 3:94F
- C complexes 1:498–499, 1:498F
- C/EBP homologous protein (CHOP) 2:531–532
- C/EBP α *see* CCAAT/enhancerbindingprotein α (C/EBP α)
- C1 complex 1:723
- C1 single-cell capture device 3:672
- C1-inhibitor 1:726
- C1r, activated forms of 1:725
- C1r/C1s, Uegf, and Bmp1 (CUB) domain proteins 2:848
- C2 1:723, 1:725–726, 1:726, 1:728–729
- C2 domain-like region (C2L region) 1:670
- C2L region *see* C2 domain-like region (C2L region)
- C3 1:723, 1:726, 1:728–729, 2:329
- C3bBb C3 convertase complex 1:729
- C4 1:723, 1:725–726, 1:726, 1:728–729, 1:729
- C4 pathway 2:329
- C4 photosynthesis dimorphic chloroplasts 2:329
- C5 1:723, 1:726, 1:728–729, 1:729
- C5a *see* complement 5a (C5a)
- C99 *see* C-terminal 99 amino acids of APP (C99); C-terminal fragment of 99 amino acids (C99)
- CA protein *see* capsid protein (CA protein)
- Ca²⁺ cage 2:425–426
- Ca²⁺ calpain activator 1:672–673
- Ca²⁺ channels 3:163–164
- Ca²⁺ regulation by proteins
- Ca²⁺ channels 3:163–164
 - Ca²⁺ transporters 3:164–165
 - Ca²⁺-related pathologies 3:167
- mitochondrion and Ca²⁺ transporting system 3:166F

- Ca²⁺ regulation by proteins (*continued*)
 surface representation of Ca²⁺-bound
 CaM 3:163F
 transporters and channels 3:164F
 Ca²⁺ transporters 3:164–165
 Ca²⁺-dependent adhesion 3:183
 Ca²⁺-dependent sorting and transport
 2:366
 Ca²⁺-induced Ca²⁺ release mechanism
 (CICR) 3:164
 Ca²⁺-related pathologies 3:167
 CAAX box 2:731
 C-Abl 4:99
 CAD *see* caspase-activated DNase (CAD);
 cationic amphiphilic drugs (CAD)
 cadherin
 adhesion 2:706
 adhesome regulatory elements 3:185–187
 cytoskeleton 3:184–185
 development 3:187–188
 family 3:183–184
 stabilization 2:706
 switch 3:188, 3:224–225
 cadherin-actin interactions 2:706
 adhesive strengthening 2:706
 cadherin adhesion 2:706
 cadherin stabilization 2:706
 junction formation 2:706
 junctional contractility 2:706–707
 morphogenesis 2:706–707
 cadherin-cadherin binding, molecular basis
 for 3:184
 cadherin-catenin complex 2:704
 caenacins 2:847
 caenopores 2:847–848
Caenorhabditis elegans 1:315–316, 1:315T,
 1:316F, 1:359, 1:366, 2:4, 2:468T,
 2:845–852, 3:3, 3:91, 3:94, 3:225,
 3:327, 3:364, 3:778
 cellular responses
 to bacterial infection 2:849F
 to fungal infection 2:849F
 immune effectors 2:846–847, 2:847T
 ABFs 2:847
 C-type lectins 2:848
 caenacins 2:847
 caenopores 2:847–848
 CUB domain proteins 2:848
 lysozymes 2:848
 NLPs 2:848
 immune signaling pathways 2:848–849
 DAF-2/DAF-16 pathway 2:849–850
 DBL-1 pathway 2:850
 p38 MAPK pathway 2:848–849
 related pathways 2:848–849
 response pathways to intracellular
 pathogen infections 2:850
 unfolded protein response pathway
 2:850
 as model system to study responses to
 pathogens 2:845
 pathogens 2:845–846, 2:845F
 recognition 2:846
 touch receptor neurons 2:543–544
 CAFs *see* cancer-associated fibroblasts
 (CAFs)
- CAGI *see* critical assessment of genome
 interpretation (CAGI)
 Cajal body (CB) 1:422, 1:504F, 1:507,
 1:508, 3:259
 calcineurin (CaN) 3:162, 3:573, 3:575
 calcitonin-gene-related peptide 1:500–501
 calcium 3:161
 Ca²⁺ regulation by proteins
 Ca²⁺ channels 3:163–164
 Ca²⁺ transporters 3:164–165
 Ca²⁺-related pathologies 3:167
 mitochondrion and Ca²⁺ transporting
 system 3:166F
 surface representation of Ca²⁺-bound
 CaM 3:163F
 transporters and channels 3:164F
 EF-hand calcium-binding principle
 3:161–162
 calcium influx
 coordinating injury response 3:236
 transforming injured axon to regenerative
 growth 3:236–237
 calcium phosphate 3:161
 calcium-activated potassium channel
 (KCa3.1) 2:394
 calcium-calmodulin pathway 3:2
 calcium-dependent activator protein for
 secretion (CAPS) 2:503
 calcium-release activated calcium (CRAC)
 3:573–574
 calcium-signal modulating cyclophilin
 ligand (CAML) 3:62
 CALI *see* chromophore-assisted laser
 inactivation (CALI)
 CALM protein *see* clathrin assembly
 lymphoid myeloid leukemia protein
 (CALM protein)
 calmodulin (CaM) 1:536F, 1:537T, 1:538,
 3:60, 3:161, 3:162–163
 CaM-dependent phosphatase 3:167
 CaM-dependent protein kinases
 3:165–167
 calmodulin kinase (CamK) 1:675
 calmodulin-regulated spectrin-associated
 protein 3 (CAMSAP3) 2:708–709
 calnexin (CNX) 1:535, 1:539–540T, 1:540,
 1:540–542, 1:541F, 1:543,
 1:599–601, 3:589, 3:593–594
 calnexin/calreticulin family (CNX/CRT
 family) 2:162–163
 calpain(s) 1:670, 1:753
 in apoptosis 1:673
 drug-induced apoptosis and calpain
 1:674
 molecules in calpain-mediated
 apoptosis 1:674
 relationship between calpain and
 caspases 1:673–674
 calpain 1 1:670
 calpain 2 1:670
 calpain 3 1:673
 calpain 8 1:673
 calpain 9 1:673
 classification 1:670–671
 conventional calpains structure 1:670
 in endocrine system 1:677
- in gastrointestinal tract 1:676–677
 human calpains
 expression distribution and
 classification 1:672T
 structure 1:671F
 physiological functions and dysregulation
 role in human disease 1:674
 cardiovascular system 1:676
 CNS 1:674–676
 proteolytic system 1:670
 regulation and activation
 Ca²⁺ calpain activator 1:672–673
 calpastatin 1:673
 CAPNS1/30K smaller regulatory subunit
 1:671–672
 regulation of unconventional calpains
 1:673
 in skeletal muscle 1:676
 calpainopathy *see* limb-girdle muscular
 dystrophy type 2A (LGMD2A)
 calpastatin 1:673
 calreticulin 1:539–540T, 1:540, 1:540–542,
 1:541F, 1:542, 1:543, 1:599–601,
 2:169, 3:589, 3:593–594
 Calvin cycle 2:248, 4:24
 CaM kinase I (CaMKI) 3:166
 CaM kinase II (CaMKII) 2:614–616, 3:167
 CaM kinase-kinase (CaMKK) 3:166
 CaM *see* calmodulin (CaM)
 CAM *see* chorioallantoic membrane (CAM)
 CamK *see* calmodulin kinase (CamK)
 CaMKI *see* CaM kinase I (CaMKI)
 CaMKII *see* CaM kinase II (CaMKII)
 CaMKK *see* CaM kinase-kinase (CaMKK)
 CAML *see* calcium-signal modulating
 cyclophilin ligand (CAML)
 camouflage 1:206
 cAMP *see* cyclic adenosine monophosphate
 (cAMP)
 c-AMP Receptor Protein (CRP) 1:477
 cAMP responsive element-binding protein,
 hepatocyte specific (CREBH) 2:534
 camptothecin 1:462
 CAMSAP3 *see* calmodulin-regulated
 spectrin-associated protein 3
 (CAMSAP3)
 CaN *see* calcineurin (CaN)
 cancer 1:141, 1:143, 1:240–241, 1:271,
 1:275–276, 1:463, 1:753, 2:221–222,
 2:601, 2:728, 2:805–806, 2:762,
 3:155, 3:156, 3:157, 3:158, 3:447,
 3:526, 3:486
 see also carcinoma
 aspartic proteases 1:753
 assembly and analysis of cancer regulatory
 models 4:256
 Bcl-2 family proteins and 3:384–385
 biology 2:138
 cancer therapy, RP-MDM2-p53 pathway
 3:286
 cysteine proteases
 calpains 1:753
 caspases 1:753–754
 cysteine cathepsins 1:754
 DCs in 3:748
 double-edged sword in 3:63

- endocytosis and implications 2:221–222
 Hippo pathway deregulation in 3:104
 hyaluronan's role in 1:285
 immunotherapy 3:63–64
 INK4a/ARF regulation and loss in 3:453–454
 kinesins and 2:612–614
 and KLK 1:702
 loss of cell polarity 2:747
 MAPK and 3:126
 meprins in 1:695
 metalloproteases
 ADAM and AMAMTS proteases 1:754–755
 matrix metalloproteinases 1:755–756
 metastasis 3:188, 4:143–144, 4:144F
 miRNAs and 1:366, 1:367
 etiology 1:366
 miRNAs as target for therapy 1:367–368
 prognosis predicting, therapy responses 1:366–367
 multiple targets for disruption of
 p16INK4a and ARF function 3:453F
 networks
 computational modeling 4:147–149
 genotype-to-phenotype relationship 4:145F
 Lauffenburger spectrum 4:148F
 principles of network biology 4:145–146
 tools to studying networks 4:146–147
 parallels between wound healing and 3:221–222
 pathways 4:256
 progression 1:360, 2:1, 3:4, 3:225
 suppression 1:687
 Rab11 and 2:404
 Rho GTPases
 and cancer cell metabolism 2:738
 and cancer cell microvesicles 2:738
 role in 1:685–686
 senescence in 3:513–515
 serine proteases 1:756
 matriptase 1:756–757
 plasminogen activation system 1:756
 splicing factor mutations and 1:508
 SRs in 3:735–736
 threonine proteases 1:757
 cancer cell invasion
 anoikis 3:227
 cell migration modes 3:226
 EMT 3:223–224
 EMT-associated cellular and molecular changes 3:224–225
 EMT-inducing transcription factors 3:224
 triggers 3:224
 invadopodia 3:225
 metastasis 3:223, 3:223F
 microtentacles 3:227–228
 cancer computational and systems biology 4:252–263
 cancer regulatory models, assembly and analysis of 4:256
 genomic alterations in glioma co-segregate 4:253F
 integrative and systems biology methodologies 4:253
 Bayesian formalism 4:253–254
 Circos plot showing whole-genome catalog of somatic mutations 4:254F
 GWAS 4:254
 model-based approaches 4:255–256
 omics layer 4:255F
 regulatory network 4:253
 study of cancer 4:254
 T-ALL 4:255
 interactome reverse engineering 4:256–257
 ARACNe 4:258–259
 benchmarking predictive models 4:259
 information theory 4:258
 integration of prior knowledge and experimental evidence 4:257
 MINDy algorithm 4:258–259
 optimization driven machine learning approaches 4:257
 probabilistic methods 4:258
 regression analysis 4:257–258
 interrogating pathways and networks 4:259–260
 The Cancer Genome Atlas (TCGA) 1:370, 4:254
 Cancer stem cells (CSCs) 3:6–7, 3:440, 3:807–812
 clinical significance of 3:810–811
 definitions and history 3:807
 markers, in major malignancies 3:809–810
 acute myeloid leukemia (AML) 3:810
 aldehyde dehydrogenase (ALDH) 3:810
 ALDH 3:810
 Bmi1 3:810
 CD24 3:810
 CD44 3:810
 CD133 3:810
 CD166 3:810
 Dcl1 3:810
 Lgr5 3:810
 Nestin 3:810
 Olfm4 3:810
 ovarian cancer 3:810
 modes of identification 3:807–809
 plasticity 3:809
 regulation of cancer stemness 3:809
 terminology 3:807
 tumor heterogeneity 3:807
 cancer-associated fibroblasts (CAFs) 3:619–620, 4:142–143
 cancer-related genetic networks 4:214
Candida albicans (*C. albicans*) 2:846
 canonical cap-dependent mechanism of translation initiation in eukaryotes 3:307
 canonical eIFs 3:310–313
 Cap structure 3:307
 CAP-1 *see* channel activating protease (CAP-1)
 cap-dependent translation 1:529, 1:529–530, 1:532–533, 1:533, 3:3–4, 3:275
 CAP-Gly domain *see* cytoskeleton-associated protein, glycine-rich domain (CAP-Gly domain)
 CAPN3 gene mutations 1:676
 CAPN10 1:677
 CAPNS1/30K smaller regulatory subunit 1:671–672
 5' capping 1:342
 capping protein interaction (CPI) 2:557
 capping proteins (Caps) 2:548, 2:549, 2:561, 2:687, 2:691, 4:14–15
 CAPS *see* calcium-dependent activator protein for secretion (CAPS); cryopyrin-associated periodic syndrome (CAPS)
 Caps *see* capping proteins (Caps)
 capsid protein (CA protein) 2:809, 2:828–829
 cap-snatching 1:521
 captopril 1:748
 capture radius 1:112–113
 capturing approaches 4:188
 carbamazepine (CBZ) 1:796
 carbohydrate recognition domain (CRD) 3:710–712
 carbon monoxide (CO) 3:57
 carbonylation 1:87
 carboxyfluorescein (CFDA) 2:136–138
 carboxymethyl (CM) 1:71
 carboxypeptidase A (CPA) 1:734
 carboxypeptidase B (CPB) 1:734
 carboxypeptidase O (CPO) 1:734
 carboxypeptidase Y (CPY) 2:475
 carboxypeptidases (CP) 1:734
 carboxy-terminal domain (CTD) 1:327F, 1:482–483, 1:506, 3:86F, 3:88F
 carcinoembryonic antigen-related cell adhesion molecule 8 (CEACAM8) 3:672
 carcinogenesis 2:438
 multiscale model of 4:73
 role of miRNAs in 1:367
 carcinogenesis 2:805–806
 carcinoma 2:585, 2:728, 3:219
 see also cancer
 CARD *see* caspase activation and recruitment domain (CARD)
 cardiac bidomain model 4:89–90
 cardiac cells 4:86
 cardiac electrophysiology 4:75
 variability in 4:88–89
 cardiac hypertrophy 3:133
 cardiac modeling 4:86–94
 bidomain model 4:89–90
 computer modeling 4:87–88
 experimental calibration 4:89
 fractional diffusion cardiac models 4:90
 human models 4:87–88, 4:88F
 in silico models 4:86–87, 4:87, 4:92
 mathematical modeling 4:86, 4:87, 4:89–90
 Model-Simulation-Experiment system 4:87
 multiscale modeling 4:86
 population of models approach 4:89
 sensitivity analysis methods 4:89
 state-of-the-art in 4:86–87
 variability 4:88–89
 whole organ cardiac modeling 4:89–92

- cardiac modeling (*continued*)
cardiac stem cell niche 3:801
 heart-evidence for 3:800–801
 interactions between niche cells and tissue progenitors 3:801–802
cardiac troponin I subunit (cTnI) 1:676
cardiolipin 1:162–164, 1:166–167, 1:168–169, 2:288
cardiovascular diseases 1:177–178
cardiovascular system 1:676
cardiovasculature 4:102–103
cargo 2:365F, 2:475, 2:666–667, 3:404, 4:56
 capture 2:331–332
 -containing carriers 2:185–186
 distribution in endosomes population 4:56
 molecular motors 3:233–234
 of newly synthesized protein 2:363
 proteins 2:479T
 recognition 2:478, 3:405–406
 of regulated secretory pathway 2:377
 retromer cargo-selective complex 2:481F
 sorting 2:202–203, 2:408–410
 implications 4:61
 sorting and packaging 2:365–366
 translocons for oversize 3:177–178
 transport 2:363
cargo and membrane trafficking theory 4:56–62
 endocytic cargo 4:56
 endosomal network theory 4:56
 experimental verification of theory of cargo trafficking 4:60F
 implications for cargo sorting 4:61
 microscopic processes regulation 4:59–60
 relation to total cargo pool kinetics 4:60–61
cargo-selective clathrin adaptors for LDL receptor 2:422
 ARH 2:423–425
 Dab2 2:422–423
 structural basis for cargo recognition 2:424F
carriers 2:365F, 2:366–367
CARs *see* chimeric antigen receptors (CARs); coherent anti-Stokes Raman scattering (CARS)
cartilage–hair hypoplasia (CHH) 3:282–283, 3:294–295
cartilages 1:282
cartwheel 2:649, 2:651, 2:652–653F
CASA *see* chaperone-assisted selective autophagy (CASA)
casein kinase 1 (CK1) 2:663, 3:76–77, 3:573
casein kinase 1 α (CK1 α) 3:90, 3:90F, 3:91, 3:91F
casein kinase 1delta/epsilon (CK1 δ/ϵ) 3:101–102
caspase activation and recruitment domain (CARD) 1:708, 3:329, 3:331, 3:364, 3:368, 3:391, 3:714
caspase-activated DNase (CAD) 3:358
caspases 1:555, 1:570, 1:670, 1:753–754, 3:4, 3:69, 3:327–328, 3:329, 3:329–330, 3:328F, 3:364–373
 activation and function in cell death and inflammation 3:368
 apoptosome 3:368–370
 death-induced signaling complex 3:370
 inflammasome 3:368
 activation mechanisms 3:366, 3:366F
 effector caspase activation 3:366–367
 initiator caspase activation 3:366
 biochemical basis of substrate specificity 3:367
 caspase-1 2:523–524
 caspase-2 3:370
 caspase-3 3:367–368
 caspase-13 3:364
 caspase-14 3:364
 caspase-16 3:364
 domain structure of 3:365F
 mutant phenotypes 3:365T
 new modes of caspase-regulated cell death 3:370–371
 necroptosis 3:371–372
 pyroptosis 3:370–371
 structure, classification, and non-apoptotic biological roles 3:364
 non-apoptotic functions 3:364–366
 structural properties 3:364
 substrates 3:367–368
 catalytic domains 1:93
 catalytic RNAs *see* ribozymes
 catastrophe 2:539, 2:541, 2:544, 4:36, 4:36–37, 4:37–38, 4:39F, 4:40, 4:40F
 catch-bonds 3:532
 CATCHR *see* complexes associated with tethering containing helical rods (CATCHR)
 CATH *see* class, architecture, topology, and homologous superfamily (CATH)
 cathelicidin 1:707
 cathepsins B 1:614
 cathepsins C 1:614
 cathepsin D 1:753
 gene expression regulation 1:682
 cathepsin E 1:681–690
 activators 1:684
 aspartic peptidases 1:681
 autoactivation 1:683
 biosynthesis and trafficking
 cell-specific processing and trafficking 1:682–683
 historical background 1:681–682
 inhibitors 1:684
 pathophysiological functions
 roles in cell growth and differentiation 1:685
 roles in immune responses 1:684–685
 substrates 1:683–684
 tissue distribution and cellular localization 1:682
 cathepsins H 1:614
 cathepsin K, mutations in 1:776
 cathepsins 1:613–614, 1:706, 1:764, 1:767
 see also specific entries
 cysteine cathepsins 1:614, 1:614F
 in immune response 1:614–615
 lysosomal hydrolases 1:614
 lysosomes and lysosomal cathepsins in cell death 1:615
cathodoluminescence 2:16
cation exchange 1:66
cation-dependent (CD) 2:410
cation-dependent MPR (CD-MPR) 2:343–344
cationic amphiphilic drugs (CAD) 1:615–616
cationic antimicrobial peptides 2:4
cationic liposomes 2:762
cation-independent (CI) 2:410
cation-independent mannose 6-phosphate receptor (CIMPR) 2:475, 2:478
Caudal type homeobox 2 (Cdx2) 3:104
caulimoviruses 1:523
CAV *see* caveolin (CAV)
caveolae 2:396, 2:396F
 see also clathrin- and caveolin-independent micropinocytic pathways; macropinocytosis
 carriers 2:397
 functions 2:397
 molecules 2:396–397
 pathway 2:397
 quantitative aspects and dynamics 2:397
caveolin (CAV) 2:397
 CAV1 2:396–397
 CAV2 2:396–397
 CAV3 2:396–397
cavins 2:396–397
CB *see* Cajal body (CB)
CBM *see* cytokine-binding modules (CBM); clathrin-binding motif (CBM)
CBP *see* CREB binding protein (CBP)
CBZ *see* carbamazepine (CBZ)
CC *see* cellular component (CC); coiled-coil (CC)
CC chemokine ligand (CCL) 3:616
CC chemokine receptor (CCR) 3:616
 CCR7 3:543
CCAAT/enhancerbindingprotein α (C/EBP α) 3:750
CCD *see* charge coupled device (CCD); coiled-coil domain (CCD)
CCD camera *see* charged coupled device camera (CCD camera)
CCL *see* CC chemokine ligand (CCL)
CCN family member 1 (CCN1) 3:102, 3:515
CCN proteins *see* Cyr, CTGF, Nov proteins (CCN proteins)
CCN1 *see* CCN family member 1 (CCN1)
CCP domains *see* complement control protein domains (CCP domains)
CCPP-60 *see* cytosolic carboxypeptidase (CCPP-60)
CCPs *see* clathrin-coated pits (CCPs)
CCR *see* CC chemokine receptor (CCR)
CCT *see* cytosolic chaperonin-containing TCP-1 (CCT)
CCV *see* *Coxiella*-containing vacuole (CCV); clathrin-coated vesicles (CCVs)
CD *see* cation-dependent (CD); circular dichroism (CD); Crohn's disease (CD)

- CD4 T cells 3:770
 responses 3:770
 TNFRSF members and 3:770–771, 3:771F
 CD27 3:771
 DR3 3:771
 HVEM 3:771–772
 ligands 3:773
 OX40 and CD30 3:772–773
- CD4-negative cells (CD4[−] cells) 2:835
- CD8 T cells 3:6
 CD8⁺ CD11b[−] DCs 3:593–594
 CD8[−] CD11b⁺ DCs 3:594–595
 CD11b⁺ DCs 3:595
 CD27 cell 3:771
- CD28 costimulation and regulatory T cells 3:605–615
 CD28/CTLA-4 pathway 3:607–608
 manipulation of 3:612–613, 3:612F
 cytotoxic T lymphocyte antigen 4 (CTLA-4) 3:608–609
 cell biology of 3:609–610
 CTLA-4 and CD28 binding same ligands with different affinities 3:608–609
 CTLA-4-mediated suppression, mechanism of 3:609
 immune regulation 3:610–611, 3:611
 transendocytosis 3:609–610
 Treg suppression 3:611–612
 functions of CD28 3:607–608
 self-tolerance establishment 3:605
 central tolerance 3:605
 peripheral tolerance 3:605–606
- CD30 cell 3:770–771, 3:772–773
- CD34 3:807
- CD34⁺ hematopoietic progenitor cells 1:785
- CD36 deficiency 3:735
- CD38 3:807
- CD44 1:280–281, 1:284, 1:285, 2:398
- CD47 cells, phagocytic synapses regulation by 3:700–701
- CD80 3:588, 3:607, 3:608–609, 3:609, 3:610, 3:610–611, 3:610F, 3:612–613, 3:612F, 3:613
- CD86 3:588, 3:607, 3:608–609, 3:609, 3:610, 3:610–611, 3:610F, 3:611F, 3:612–613, 3:612F, 3:613, 3:667–669
- CD95 *see* fatty acid synthase (FAS)
- CD103⁺ CD11b[−] DCs 3:595
- CD146 3:797
- CD177-expressing neutrophils 3:753
- Cdc2 3:424
- Cdc42 2:497, 2:731, 3:123–124
 mammalian targets 2:735F
 -mediated regulation model 2:736F
 small GTPase 2:741–742
 and yeast cell polarization mechanisms
 breaking symmetry 4:52–53
 cue-dependent 4:51–52
 feedback loops 4:52–53, 4:52F, 4:53F
 inhibitory interactions and reinforcement 4:53–55
 recycling of Cdc42 4:52, 4:52–53, 4:53, 4:53F, 4:54
- CDC42-Rac interactive binding region (CRIB) 2:497
- Cdc45–Mcm2–7–GINS complex (CMG complex) 3:458–459
- Cdc48 1:606
- cDCs *see* classical DCs (cDCs); conventional DCs (cDCs)
- CDG *see* congenital disorders of glycosylation (CDG)
- CDI *see* *Clostridium difficile* infection (CDI)
- CDK *see* cyclin-dependent kinase (CDK)
- Cdk inhibitory proteins (Cip) 3:448
- CDK4/6 assembly 3:439–440
- CDKs *see* cyclin-dependent kinases (CDKs)
- CD-MPR *see* cation-dependent MPR (CD-MPR)
- cDNA *see* complementary deoxyribonucleic acid (cDNA)
- CDNs *see* cyclic dinucleotides (CDNs)
- CDP *see* common DC progenitor (CDP)
- CDRs *see* complementary determining region (CDR)
- CDT *see* *Clostridium difficile* toxin (CDT); cytolethal distending toxins (CDTs)
- Cdx2 *see* Caudal type homeobox 2 (Cdx2)
- CEACAM8 *see* carcinoembryonic antigen-related cell adhesion molecule 8 (CEACAM8)
- celiac disease 1:735
- cell adhesion 2:724
 biochemical and mechanical effects 3:192F
 in connective tissue function 3:194–195
 early studies 3:192–193
 to ECM
 cross-talk between cell and ECM 2:717–718, 2:717F
 focal adhesions 2:714F, 2:715F
 integrin adhesome 2:714–716, 2:716F
 invadosomes 2:715F
 molecular composition of ECM 2:714F
 molecules 1:622–624
 SRs contributing to 3:732
- cell behavior, analyzing 4:122–123
- cell cycle 3:423, 3:433, 3:436
see also cyclin
 cyclin-CDK in 3:425–428
 discovery of regulators 3:423–424
 G1 phase 3:433F
 phases of eukaryotic 3:423F
 progression 1:334, 3:3, 3:4–6, 3:437, 3:607
 regulation 4:172
 oncogenes, tumor suppressor genes and 3:414–415
- cell cycle arrest 3:3, 3:5, 3:6, 3:511, 3:512F, 3:513–514, 3:607
- cell death 1:451, 1:555, 1:570, 1:772, 2:524, 3:67, 3:126, 3:327, 3:329, 3:330, 3:331, 3:332, 3:364, 3:369F, 3:512F, 3:605, 3:672–673
- Anoikis response 3:4
- caspase activation and function in 3:368
- caspase-regulated 3:370–371
- DRs 3:67
 DR signaling 3:68–69
 ligand activation 3:67–68
 of host defense 2:773–774
- mechanisms and host defense 2:772–773
 apoptosis 2:773
 NETs and NETosis 2:773, 2:773F
 pyroptosis 2:772–773, 2:772F
 regulation 3:4
 and signaling 3:390
 TNFR-induced 3:358
- cell division 4:157
 positioning site
 RhoA pathway 3:494–495
 spindle microtubules 3:494, 3:495F
- cell division cycle 42-interacting protein 4 (CIP4) 2:492–493
- cell growth 1:423, 1:763, 1:766, 1:779, 2:4, 3:1, 3:3, 3:89, 3:95, 3:515, 3:516F
- cell cycle progression 3:4–6
- cell death regulation 3:4
- ribosome biogenesis pathways 3:3
- translation 3:3–4
- cell loading chamber (CLC) 3:670–671F, 3:671–672
- cell membrane 1:209, 1:211, 1:213, 1:214, 1:218, 1:230F, 1:245, 1:246, 1:246–247, 1:247–248, 1:326, 1:563–564, 1:569, 1:780, 3:328, 3:608, 3:608F
 electrical properties of 1:220–221
 physical-chemical properties of 1:218–220
- cell membrane lipids 1:219–220F, 1:326
- cell metabolism, oncogenes, tumor suppressor genes and 3:415–417
- cell migration 2:575, 2:683, 2:714, 2:720–730, 3:127, 3:331, 3:669–672
 building migrating cell 2:723–724
 ECM 2:723–724
 intracellular determinants 2:724–725
 mechanical forces 2:724
 membrane receptors 2:724
 soluble chemoattractants 2:724
 cytoskeletal changes 3:208F, 3:210
 by Rho family GTPases 3:210–211
 in development and homeostasis 2:727–728
 embryonic development 2:727–728
 inflammation 2:728
 migration in CNS 2:728
- eukaryotic cell migration modes 2:720–723, 2:721F
 individual cell migration 2:720–723
 multicellular migration 2:723
- harnessing cell migration in therapy 2:729
- through interactions with environment 3:212–213
 collective cell migration 3:213–215, 3:214F
 focal adhesions 3:213, 3:213F
 matrix remodeling 3:213, 3:214F
- modes 3:226
 amoeboid migration 3:226
 collective cell migration 3:226–227
 mesenchymal migration 3:226
- in pathology 2:728
 autoimmune disease 2:728–729
 cancer 2:728

- cell migration (*continued*)
 - signals to migration 3:208
 - during development 3:209–210
 - directed cell migration 3:208–209
 - in physiological/pathological processes 3:210
 - types 3:211
 - amoeboid migration 3:212
 - lamellipodial migration 3:211–212
 - lobopodial migration 3:212
- cell motility and shape, mechano-chemical
 - control of 4:10
 - actin filament treatment 4:11
 - biophysical mechanisms 4:14
 - Brownian ratchet formalism 4:10, 4:10
 - fish keratocyte cell phase microscopy image 4:12F
 - G-actin transport 4:13
 - locomotion mechanism 4:13F
 - model types and evolution 4:11F
 - network swelling model 4:12
- cell polarity
 - and disease 2:747–748
 - epithelial cells 2:742
 - apical–basal polarity 2:742
 - cytoskeletal rearrangements 2:746
 - initiating cues of polarization 2:743–745
 - mediators 2:742
 - polarized trafficking machinery 2:745–746
 - loss 2:747–748
 - molecular models for Pins-and Dsh-mediated spindle orientation 2:748F
 - planar cell polarity 2:742
 - asymmetric distribution of core PCP proteins 2:747F
 - molecular basis 2:742
 - principles and control mechanisms 2:741–742
 - regulation of 3:199–207
 - feedback control of core amplification 3:203
 - Ft/Ds/Fj system 3:204
 - modular signaling system 3:200–202
 - principles of collective polarization 3:202–203
 - system architecture 3:203–205
 - relationship with spindle orientation control 2:747
- cell protrusion 2:3–4, 3:671–672
- cell signaling 4:218
 - and intercellular communication 3:1–2
 - cellular machineries 3:2
 - cellular outcomes 3:2–3
 - receptors 3:1–2
 - signal transduction pathways 3:2
 - molecular-state differences 4:218–219
 - oscillations 4:219–220
 - stochastic events in signaling 4:219
- cell type-specific secretory lysosomes 2:225–228
- cell/substrate contact regions 2:62–63
- cell-based models 4:123
- cell-based regenerative therapy, tools for 3:820–821
- clinical trials and contributions 3:821
- Cell-behavioral Ontology 4:122–123
- cell–cell adhesion 2:704, 3:2, 3:181F
 - and actin cytoskeleton 2:704–705
 - biological functions of cadherin–actin interactions 2:706
 - classical cadherin adhesion integration 2:704–705, 2:704F
 - dynamic regulation 2:705–706
 - physical connections between cadherins and F-actin 2:705
- adherens junction (AJs)
 - actin dynamics at 3:187
 - diversity 3:183–184
 - in disease 3:188–189
- cadherin
 - adhesome regulatory elements 3:185–187
 - and embryonic development 3:187–188
 - family 3:183–184
 - cytoplasmic plaque 3:184
 - diversity in mouse epithelial cells 3:182F
 - dynamic connections between cadherin and actin cytoskeleton 3:184–185
 - E-cadherin receptors 3:185F
 - essence of multicellularity 3:181–182
 - and intermediate filament cytoskeleton 2:707–708
 - integrating IFs and desmosomes 2:707–708
 - physiological functions of desmosomal-IF integration 2:708
 - interplay between regulators 3:186F
 - molecular basis for cadherin–cadherin binding 3:184
 - molecules 1:625, 2:704
 - morphological variations of adherens junctions 3:183F
 - and MTs 2:708–709
 - cortical signaling regulation 2:710
 - functional contributions at junctions 2:709–710
 - interactions in epithelial cells 2:709F
 - linkages with cell–cell junctions 2:708–709
 - MT-dependent cellular transport 2:709–710
 - structural and functional diversity 3:182–183
- cell–cell communication 1:202, 3:672–673, 3:789, 4:73F
- cell–cell junctions *see* cell–cell adhesion
- cell–ECM adhesion molecules 1:625
- cell–ECM interactions, modeling 4:125
- cell-free expression systems 1:58–59
 - see also* bacterial expression systems; mammalian expression systems; yeast expression systems
- cell-intrinsic aging effects 3:779
- cell–matrix adhesions
 - in pathological processes 3:195–196
 - in pathological processes 3:196
- cell–matrix interactions 3:192
 - extracellular signals 3:194
 - specific vs. generic attractive and repulsive interactions 3:193
- CellML 4:75–76, 4:87
- cell-penetrating peptides 2:762
- cell–substrate contacts 2:63
- cell-surface libraries 1:144
- cell-to-cell communication 2:376
- cell-to-cell heterogeneity 4:218
- cell-type specific responses 4:46
- cellular
 - cargoes transport 2:624
 - cellular excitability, light-sensitive proteins to control 2:150–151
 - cholesterol homeostasis 1:176–177, 1:177F
 - compartments 1:288, 1:294
 - functions 4:154–155
 - homeostasis 2:846
 - immunity 2:835–836
 - microenvironment 2:716
 - organelles 2:274
 - phase separations 1:113
 - process 4:158
 - senescence 3:777, 3:780
 - ultrastructure 2:44
 - UPS 1:116–117
- cellular automata models 4:119
- cellular component (CC) 4:247
- cellular FLICE inhibitory protein (cFLIP) 3:72, 3:335
- cellular form of prion protein (PrPC) 2:841, 2:842
- cellular inhibitor of apoptosis protein (cIAP) 3:69
 - c-IAP1 1:707–708, 3:62, 3:391–393
 - c-IAP 2 1:707–708, 3:62
 - in immunity 3:394
- cellular invasion by bacterial pathogens 2:784, 2:784F
- cellular responses to bacterial invasion
 - autophagy 2:790–791
 - cell survival 2:791
 - inflammatory response 2:791
- entry mechanisms 2:784
 - trigger mechanism 2:785–787
 - zipper mechanism 2:784–785
- host intracellular trafficking, intracellular lifestyle and manipulation of 2:788–790
 - cytosolic bacteria 2:790
 - intravacuolar bacteria 2:788–790
- stages 2:784
- cellular machineries 3:1, 3:2, 3:2–3
- cellular organization and human disease
 - interactome maps 4:194
 - cell communication 4:194
 - contribution 4:194, 4:195F
 - drugs, and drug targets 4:195–196
 - host–pathogen interactions 4:195
 - human disease 4:194–195
- Cellular Potts model (CPM) 4:74, 4:119, 4:123–124, 4:124F, 4:125, 4:126–127
- cellular quiescence 3:789
- cellular regulatory modules 4:54
- cellular senescence 3:6, 3:511–517, 3:790
 - in aging 3:513
 - in cancer 3:513–515
 - in development 3:515

- epigenetic control of 3:512–513
molecular signaling pathways 3:512
secreted factors 3:513
and tissue repair 3:515
cellular stress 3:270
see also ribosome stress
mTOR signaling regulation by
 genotoxic stresses suppress mTORC1
 3:248
 mTORC1 regulation by SGs 3:247–248
 TSC 3:247
CEM group *see* 2-cyanoethoxymethyl
 (CEM) group
CEMOVIS *see* cryo-electron microscopy of
 vitreous sections (CEMOVIS)
CENP-E *see* centromere associated protein E
 (CENP-E)
central channel 2:313
central metabolism 4:172–173
central nervous system (CNS) 1:627,
 1:674–676, 2:728, 3:235–236,
 3:797–798, 4:134
 cell migration in 2:728
central pair 2:630–631
central supramolecular activation cluster
 (cSMAC) 3:522–524, 3:660,
 3:681–682
central tolerance 3:605, 3:605
centralspindlin 3:494–495, 3:497, 3:499
centrifugation 1:70
centrin-2, EF-hand calcium binding protein
 1:435F, 1:436–437, 1:438T
centriolar satellites 2:664–665
centriole 2:649, 2:649–651, 2:660,
 2:661–663, 2:664
 and centrosomes 2:4
 duplication 2:651–653, 2:652–653F
 proteins requirement 2:651T
 mother and daughter 2:651
centromere associated protein E (CENP-E)
 2:584
centromeres, two-histone model of CPC
 localization to 3:473
centrosomal MT 2:708
centrosomal protein (CEP) 2:663
centrosome(s) 2:311, 2:638, 2:649, 2:660,
 2:664, 3:478, 3:480
 amplification 2:657F
 asymmetric centrosome inheritance
 2:656F
 in cell division 2:653–655
 centriole 2:649, 2:649–651
 duplication 2:651–653, 2:652–653F
 in development 2:655–656
 in disease 2:656–657
 functions 2:654–655F
 microtubule nucleation and anchoring
 2:653
 polarity establishment 2:655
 primary cilium formation 2:655
 structure 2:91–92, 2:649–651, 2:650F
CEP *see* centrosomal protein (CEP)
ceramide 1:149, 1:150, 1:155, 1:195–196,
 1:198–199
 phosphoinositols 1:184
ceramide phosphate (CerP) 1:187
ceramide synthases (CerS) 1:187
ceramide trafficking protein *see* ceramide
 transfer protein (CERT)
ceramide transfer protein (CERT) 1:205,
 2:187, 2:503
cerebrospinal fluid (CSF) 1:664, 3:798–799,
 3:799
ceRNA mechanism *see* competitive
 endogenous RNA mechanism
 (ceRNA mechanism)
CerP *see* ceramide phosphate (CerP)
CerS *see* ceramide synthases (CerS)
CERT *see* ceramide transfer protein (CERT)
CETP *see* cholesteryl ester transfer protein
 (CETP)
CetZs 2:592–593
CF *see* cystic fibrosis (CF)
C-family DNA polymerase 1:398–400,
 1:400
CFDA *see* carboxyfluorescein (CFDA)
cFLIP *see* cellular FLICE inhibitory protein
 (cFLIP)
c-Fos protein 4:47
CFP *see* cyan fluorescent protein (CFP)
CFTR *see* cystic fibrosis trans-membrane
 conductance regulator (CFTR)
CFTR *see* cystic fibrosis transmembrane
 conductance regulator (CFTR)
cGAMP *see* cyclic GMP–AMP (cGAMP)
cGAMP synthase (cGAS) 3:520
cGAS *see* cGAMP synthase (cGAS)
CGD *see* chronic granulomatous disease
 (CGD); congenital glycosylation
 disorders (CGD)
cGMP *see* cyclic guanosine monophosphate
 (cyclic GMP)
CGN *see* cis-Golgi network (CGN)
CH *see* clamp helix (CH)
CH *see* cysteine/histidine (CH)
chain migration 2:723
champagne-glass-like structure 2:664
channel activating protease (CAP-1)
 1:651–652
channelrhodopsins (ChRs) 2:148–149
 ChR2 2:150–151
chaperone-assisted selective autophagy
 (CASA) 3:405–406
chaperone-mediated autophagy (CMA)
 1:616, 3:404, 3:406F
chaperones 1:537T, 1:542, 1:555, 1:558,
 1:560–561, 1:599, 2:699, 3:276
chaperonin 2:6–7, 2:8F, 2:9, 2:11
characteristic focal adhesion plaque proteins
 2:713–714
Charcot–Marie–Tooth disease (CMT
 disease) 2:285–286, 2:477–478
Charcot–Marie–Tooth neuropathy II
 (CMT2A) 2:279
charge coupled device (CCD) 2:6, 2:28
charged coupled device camera (CCD
 camera) 4:234
charged multi-vesicular body protein 4
 (CHMP4) 2:195
CHASTE (software) 4:74–75, 4:75
Chb *see* chitobiose (Chb)
CHC *see* clathrin heavy chain (CHC)
CHCR *see* clathrin heavy chain repeat
 (CHCR)
CHD *see* coronary heart disease (CHD)
CheA protein 4:7
checkpoint protein kinase 1:124
Chédiak–Higashi disease 1:775–776
Chédiak–Higashi syndrome (CHS)
 2:227–228, 2:237–239, 2:241
ChEMBL database 4:248–249
chemical biology 1:128–134, 1:137
 activity-based probes 1:131, 1:131F
 affinity labeling 1:130–131
 bioorthogonal labeling 1:131–132, 1:132F
 and drug development 1:132–133
 imaging 1:128
 metabolic labeling 1:128–130
 technical approaches in 1:129F
chemical defences 3:6, 3:519–520
chemical fixation 2:32–34
chemical ionization (CI) 4:207
chemoattractant receptors 2:724
chemokines 1:274, 1:275–276, 1:624–625,
 3:584, 3:624
 receptors 3:751
chemotactic systems, adaptation in 4:7
 incoherent feedforward scheme 4:7
 negative feedback regulation 4:7, 4:7F
chemotaxis 2:720, 2:724, 3:208–209, 4:4
 see also haptotaxis
 strategies 4:4–5
 chemoattractant gradient 4:5F
 spatial sensing 4:6
 temporal sensing 4:5–6
chemotherapy-induced apoptosis 1:674
cherry red spot 1:767, 1:774
CheY protein 4:7
CHH *see* cartilage–hair hypoplasia (CHH)
chimera 1:512–513, 1:514
chimeric antigen receptors (CARs) 1:145
Chinese hamster ovary cells (CHO cells)
 1:61, 2:344
ChIP *see* chromatin immunoprecipitation
 (ChIP)
ChIP-Seq *see* chromatin
 immunoprecipitation followed by
 sequencing (ChIP-Seq)
chirality 1:160–161
chitobiose (Chb) 1:101
Chk1 effector proteins 3:461–462
Chk2 effector proteins 3:461–462
Chlamydia trachomatis (*C. trachomatis*)
 2:255, 2:405, 2:777, 2:780
Chlamydomonas reinhardtii 2:472
chloride secretion, CFTR-dependent 1:255,
 1:256F
chloride transport, CFTR and 1:255–256
chlorophyll 2:328
chloroplast 1:265, 1:315T, 1:319, 1:547,
 2:2–3, 2:324
 functions 2:327–328
 amino acid synthesis 2:329
 lipid synthesis 2:329
 nitrogen assimilation 2:329
 photosynthesis 2:328–329
 pigment synthesis 2:329
 protein synthesis 2:329

- chloroplast (*continued*)
 starch synthesis 2:328–329
 sulfur assimilation 2:329
 photosynthetic electron transport 2:328F
 structure 2:327
 variants 2:329
 CHMP4 *see* charged multi-vesicular body protein 4 (CHMP4)
 CHO cells *see* Chinese hamster ovary cells (CHO cells)
 cholelithiasis *see* cholesterol gallstone disease
 cholera toxin (CT) 2:795
 cholesterol 1:173, 1:202, 1:203, 1:203–204, 1:206, 1:208, 1:208–209, 1:211, 1:212, 1:213, 1:368, 1:764, 1:765, 1:779, 2:761, 3:418, 3:2
see also obesity
 biosynthesis 1:174–175, 1:175F
 regulation 1:175–176
 biosynthesis 3:77
 in brain 1:177
 in cell membranes 1:174
 cellular cholesterol homeostasis 1:176–177, 1:177F
 cholesterol-dependent cytolysins 2:776
 cholesteryl esters 1:173–174
 in diseases 1:177–178
 functions 1:173
 gallstone disease 1:178
 overload responses 2:427–428
 in plasma membrane 1:174F
 steroids 1:178–179, 1:178F
 structure 1:173, 1:174F
 synthesis 1:174–175, 2:254
 transport 3:735
 cholesterol 7- α hydroxylase (CYP7A1) 1:178
 cholesterogenesis
see cholesterol—synthesis
 cholesteryl ester transfer protein (CETP) 1:173–174
 cholesteryl esters 1:173–174
 cholinergic nicotine receptor α 3 (*CHRNA3*) 1:794
 chondroitin sulfate (CS) 1:271, 2:351
 CHOP *see* C/EBP homologous protein (CHOP)
 chordin (CHRD) 3:44
 chorioallantoic membrane (CAM) 1:695
 CHRD *see* chordin (CHRD)
CHRNA3 see cholinergic nicotine receptor α 3 (*CHRNA3*)
 chromatin 1:435, 1:436, 1:437, 1:483–484, 2:311–312, 2:573, 3:95, 3:785, 3:786, 3:788, 3:789–790, 3:790–791
 -based microtubule polymerization 2:638
 bridges 3:501
 structure 1:506
 chromatin immunoprecipitation (ChIP) 4:64, 4:191
 chromatin immunoprecipitation followed by sequencing (ChIP-Seq) 1:579, 3:722, 4:191, 4:211, 4:216
 chromatin remodeling 1:437, 1:491, 1:493, 3:94F, 3:786–787
 chromatin-immunoprecipitation (ChIP) 3:785–786
 chromatography 1:70–71, 1:75
 adsorption chromatography 1:72
 affinity chromatography 1:72
 column 1:71
 IEX chromatography 1:71
 SEC 1:71–72
 chromophore-assisted laser inactivation (CALI) 4:21
 chromoplast 2:326–327
 chromosomal passenger complex (CPC) 3:5, 3:469, 3:482–483, 3:496
 during anaphase, telophase, and cytokinesis 3:475
 core CPC members identification 3:469
 CPC targeting modified view 3:473–475
 and error correction 3:472–473
 INCENP 3:469
 inter-kinetochore tension 3:471–472, 3:473F
 during mitosis 3:470–471, 3:471
 spindle assembly checkpoint 3:471, 3:471–472, 3:472F
 subunits 3:470F
 two-histone model 3:473, 3:475F
 chromosomal/chromosome 2:19–20, 2:311
 congression 2:641–644, 2:642F
 passenger hypothesis 3:469
 separation 2:643F, 2:644
 chromosome instability (CIN) 2:651
 chromosome passenger complex (CPC) 1:709
 chromosome separation 2:3–4, 2:4
 chronic granulomatous disease (CGD) 2:768, 3:702–703, 3:754
 chronic immune activation 2:835
 chronic inflammatory diseases 3:584–585
 chronic lymphocytic leukemia (CLL) 1:131, 1:366, 1:367, 3:110
 chronic myeloid leukemia (CML) 3:442
 chronic neuropathic Gaucher disease 1:772, 1:779, 1:780, 1:781
 chronic obstructive pulmonary disease (COPD) 1:789, 3:60
 chronic wasting disease (CWD) 2:839, 2:841
 ChRs *see* channelrhodopsins (ChRs)
 CHS *see* Chediak–Higashi syndrome (CHS)
 chymase 1:747–748
 chyme 1:733, 1:734
 chymotrypsin 1:734
 chymotrypsinogen 1:734
 CI *see* cation-independent (CI); chemical ionization (CI); cubitus interruptus (CI)
 CI endocytic processes *see* clathrin-independent (CI) endocytic processes
 cIAP *see* cellular inhibitor of apoptosis protein (cIAP)
 CICR *see* Ca²⁺-induced Ca²⁺ release mechanism (CICR)
 CID *see* collision-induced dissociation (CID)
 CIE *see* clathrin-independent endocytosis (CIE)
 cigarette smoking 1:794
 CIITA *see* class II transactivator (CIITA)
 cilia 2:608, 2:611, 2:660–676
 basis of motility 2:663
 ciliary membrane 2:663–664
 ciliary tip structures 2:663
 ciliogenesis 2:664–666
 ciliopathies 2:664, 2:670T
 symptoms 2:671T
 cilium disassembly and length maintenance 2:667–668
 diseases caused by defects 2:668–670
 HCs in 2:628–630
 motile 2:660–661
 diseases caused by defective 2:669–670
 structural and functional features 2:662F
 non-motile 2:660–661
 axoneme structure 2:663
 diseases caused by defective 2:670–671
 signaling pathways 2:668, 2:669F
 structure 2:661–663
 types 2:661F
 TZ 2:664
 passage of protein through 2:667
 ciliary pore complexes 2:667
 ciliary tip structures 2:663
 ciliogenesis 2:92, 2:611, 2:664–666
 ciliopathies 2:660, 2:660–661, 2:664, 2:670T
 symptoms 2:671T
 CIMPR *see* cation-independent mannose 6-phosphate receptor (CIMPR)
 CIN *see* chromosome instability (CIN)
Ciona intestinalis (*C. intestinalis*) 1:691
 Cip *see* Cdk inhibitory proteins (Cip)
 CIP/KIP family
 functions independent of CDK inhibition 3:439–440
 normal function 3:438–439
 regulation 3:439
 CIP4 *see* cell division cycle 42-interacting protein 4 (CIP4)
 CircadiOmics 4:212T, 4:214
 circular dichroism (CD) 1:94
 circular RNA 4:162–163
 circulating tumor cells (CTCs) 3:225, 3:673–674
 circulatory system 4:102
 cirrhosis 1:793
 CIS *see* cytokine inducible SH2-containing protein (CIS)
Cis-acting signal 1:517, 1:517–518
cis-Golgi network (CGN) 2:341–342, 2:342, 2:342F, 2:344–345, 2:354
see also trans-Golgi network (TGN)
cis-regulatory interactions 4:65
 cisternal maturation 2:186
 model 2:356–358
 vesicular anterograde transport vs. 2:356
 cisternal progenitor model 2:186, 2:359
 Cistron 3:308, 3:309
 cistronic miRNA 1:364
 citalopram 1:248

- citius* 2:602
 citric acid cycle *see* tricarboxylic acid (TCA) cycle
 CJD *see* Creutzfeldt–Jakob disease (CJD)
 c-Jun N-terminal kinase (JNK) cascade 3:122
 CK1 *see* casein kinase 1 (CK1)
 CK1 α *see* casein kinase 1 α (CK1 α)
 CK1 δ/ϵ *see* casein kinase 1delta/epsilon (CK1 δ/ϵ)
 CKIs *see* cyclin-dependent kinase inhibitors (CKIs)
 CKMT1 *see* creatine kinase, mitochondrial 1 (CKMT1)
 clamp helix (CH) 1:471
 clamp loader 1:396, 1:411–412, 1:412
 in last universal common ancestor 1:401–403, 1:402F
 mechanism 1:401–403
 structure analysis 1:403
 CLAP domain *see* clathrin/AP-2 binding domain (CLAP domain)
 CLASP2 *see* cytoplasmic linker associated protein 2 (CLASP2)
 CLASPs *see* clathrin-associated sorting proteins (CLASPs)
 class, architecture, topology, and homologous superfamily (CATH) 1:94–95
 class B scavenger receptors (SRs) 3:731–732
 class E phenotype 2:466, 2:469
 class F and H scavenger receptors (SRs) 3:732
 class I activators 1:477
 class I PI3-kinase (PI3KC) 2:503
 complex 2:503
 PI3KC3 2:503
 class II activators 1:477
 class II PI3-kinase (PI3KC2A/PI3KC2 α) 2:503, 2:508
 class II PI4-kinase II α 2:503
 class II transactivator (CIITA) 3:714
 class III PI4-kinase (PIK1) 2:512–513
 class III PI4-kinase III β 2:503
 class switch recombination (CSR) 3:552
 class-1 myosins 2:601
 class-2 myosins 2:601–602
 class-3 myosins 2:602
 class-5 myosins 2:602–603
 class-6 myosins 2:603
 class-7 myosins 2:603
 class-8 myosins 2:603–604
 class-9 myosins 2:602
 class-10 myosins 2:603
 class-11 myosins 2:603–604
 class-14 myosins 2:604
 class-15 myosins 2:603
 class-16 myosins 2:604
 class-18 myosins 2:604
 class-19 myosins 2:604
 class-21 myosins 2:604
 classical DCs (cDCs) 3:592
 classical NHEJ (cNHEJ) 3:558–559
 classical pathway 1:723, 1:725F
 see also alternative pathway; lectin pathway
 C1 complex 1:723
 C1-inhibitor 1:726
 C1s 1:725–726
 C2 activation by proteolytic cleavage 1:726
 C2 component of complement 1:726
 modules 1:724–725
 zymogen structures and activated forms of C1r 1:725
 classical swine fever (CSFV) 3:308
 CLASS-type lipidomic analysis 1:157
 clathrate 1:109
 clathrin 2:296, 2:384–393, 2:503
 adaptor proteins 2:296–297
 assembly and disassembly 2:387–388
 biochemistry 2:384–385
 cellular function and regulation 2:388
 clathrin roles in yeast and plants 2:390
 endosome roles 2:389–390
 non-membrane traffic roles for mammalian CHC17 clathrin 2:390–391
 organelle biogenesis 2:390
 PM roles 2:388–389
 CHCs 2:385, 2:386F
 CLCs 2:385–387
 function in cargo selection 2:384F
 subunits 2:385T
 clathrin assembly lymphoid myeloid leukemia protein (CALM protein) 2:455
 clathrin heavy chain (CHC) 2:384–385, 2:385, 2:386F
 CHC17 2:385, 2:386–387, 2:386F
 assembly 2:387–388
 disassembly 2:388
 non-membrane traffic roles 2:390–391
 phosphorylation 2:389
 role in GLUT4 formation 2:390
 CHC22 2:385, 2:386F, 2:388, 2:390
 clathrin heavy chain repeat (CHCR) 2:385
 clathrin light chain (CLC) 2:385, 2:385–387
 diversity 2:387F
 low resolution structure 2:388F
 protein domains 2:387F
 clathrin/AP-2 binding domain (CLAP domain) 2:498
 clathrin- and caveolin-independent micropinocytic pathways 2:397, 2:395F
 see also caveolae; macropinocytosis
 carriers 2:398
 functions 2:398–399
 molecules 2:397–398
 pathway 2:398
 quantitative aspects and dynamics 2:398
 clathrin-associated sorting proteins (CLASPs) 2:425
 clathrin-binding motif (CBM) 2:454
 clathrin-coated pits (CCPs) 2:212, 2:452, 2:491, 2:498
 clathrin-coated vesicles (CCVs) 2:384, 2:425, 2:498, 4:56
 clathrin-coats 2:366
 clathrin-independent carrier (CLIC) 2:394
 clathrin-independent (CI) endocytic processes 2:394
 clathrin-independent endocytosis (CIE) 2:401
 clathrin-mediated endocytosis (CME) 2:212, 2:299, 2:394, 2:497
 alphaviruses internalization 2:814F
 claudin-16 3:691–692
 CLC *see* cell loading chamber (CLC); clathrin light chain (CLC)
 CLCN7 channel 1:763–764, 1:776
 CLEAR *see* coordinated lysosomal expression and regulation element (CLEAR)
 CLEAR *see* coordinated lysosomal expression and regulation (CLEAR)
 cleavage furrow formation 3:495
 CLEM *see* correlative light and electron microscopy (CLEM)
 CLIC *see* clathrin-independent carrier (CLIC)
 CLIC/GEEC pathway 2:394, 2:397–398, 2:398–399
 click chemistry 1:130, 1:135
 CLIMP-63 transmembrane protein 2:158
 clinical targeting of TNFSF and TNFRSF members 3:63–64
 CLIP *see* cross-linking and immunoprecipitation (CLIP); cytoplasmic linker protein (CLIP)
 CLL *see* chronic lymphocytic leukemia (CLL)
 CLM *see* cystatin-like molecule (CLM)
 closed mitosis 2:315
Clostridium difficile infection (CDI) 2:804, 3:690–691
Clostridium difficile toxin (CDT) 2:795
Clostridium difficile toxin B (TcdB) 2:796
Clostridium spiroforme toxin (CST) 2:795
 CLP *see* common lymphoid progenitors (CLP)
 CLRs *see* C-type lectin receptors (CLRs)
 clustered regularly interspaced short palindromic repeat (CRISPER) 1:515, 4:236–237
 clustering 1:212, 1:214, 3:203, 3:655–656, 4:212, 4:212T, 4:213
 and biclustering algorithms 4:213
 clusters of vesicles and tubules (VTCs) 2:171
 ClustEx 4:212T, 4:213
 CM *see* carboxymethyl (CM)
 CMA *see* chaperone-mediated autophagy (CMA)
 CMD *see* congenital muscular dystrophies (CMD)
 CME *see* clathrin-mediated endocytosis (CME)
 CMG (Cdc45, Mcm2–7, and GINS) helicase 1:404–405, 1:410, 1:411
 CMG complex *see* Cdc45–Mcm2–7–GINS complex (CMG complex)
 CMI *see* conditional mutual information (CMI)
 CML *see* chronic myeloid leukemia (CML)
 cMonkey 4:212T, 4:213
 cMOS *see* complementary metal-oxide semiconductor (cMOS)

- CMPs *see* common myeloid progenitors (CMPs)
 CMT disease *see* Charcot-Marie-Tooth disease (CMT disease)
 CMT2A *see* Charcot-Marie-Tooth neuropathy II (CMT2A)
 CMV *see* cytomegalovirus (CMV)
 c-Myc 3:273, 3:282
 cN *see* cyclic nucleotides (cN)
 CNBr *see* cyanogen bromide (CNBr)
 CNF1 *see* cytotoxic necrotizing factor 1 (CNF1)
 cNHEJ *see* classical NHEJ (cNHEJ)
 CNP *see* C-type natriuretic peptide (CNP)
 CNS *see* central nervous system (CNS)
 CNVs *see* copy number variations (CNVs)
 CNX *see* calnexin (CNX)
 CNX/CRT family *see* calnexin/calreticulin family (CNX/CRT family)
 CO *see* carbon monoxide (CO)
 CoA *see* coenzyme A (CoA)
 co-activators 1:488, 1:491, 3:94, 3:95
 coagulation factor 1:718
 coat protein complex I (COPI) 2:168
 coats 2:170, 2:176
 COPI-coated vesicles 2:354–355, 2:356–358
 COPI-mediated bidirectional transport model 2:358
 coat protein complex II (COPII) 2:163–164
 biogenesis regulation 2:332–333
 cellular regulation 2:333–334
 coats 1:205, 1:206, 2:331, 2:332F
 cobalamin transport disorder 1:775
 COC *see* cumulus oophorus complex (COC)
 co-catalysts *see* coenzymes and cofactors
 Cockayne syndrome 1:439
 Cockayne syndrome A and B proteins 1:437, 3:258–259
 codon 3:310
 coenzyme A (CoA) 1:19, 1:20F, 1:161
 cofactors 4:24
 cofilin 4:19
 COFRADIC *see* combined fractional diagonal chromatography (COFRADIC)
 COG *see* conserved oligomeric Golgi (COG)
 cognate lysosomal membrane protein 1:775
 cognate sphingolipid activator protein (SAP-C) 1:772, 1:774
 coherent anti-Stokes Raman scattering (CARS) 2:129–131
 coiled-coil (CC) 2:597, 3:357–358
 dimer 2:569, 2:570
 tethers 2:486–487
 coiled-coil domain (CCD) 3:143
 collagen 1:695–696, 3:195
 collagen IV 2:696–697
 collagen monomers 2:699
 fibers 2:700
 collagenase-3 *see* MMP-13
 collective cell behavior, mathematical models for studying 4:123, 4:125–126
 cell-ECM interactions 4:125
 continuum models 4:123
 discrete models 4:123–124
 Cellular Potts model (CPM) 4:125
 particle-based models 4:123–124
 subcellular elements method 4:123–124
 vertex-based models 4:124–125
 collective cell migration 3:213–215, 3:214F
 collective migration 2:723
 collective polarization, principles of 3:202–203
 collision-induced dissociation (CID) 1:150, 1:568, 1:572, 4:180
 colocalization analysis 2:83
 colon 1:732, 1:735–736
 colony stimulating factor-1 (CSF-1) 3:14, 3:731
 colony stimulating factor 2 (CSF2) 3:323
 colorectal cancer (CRC) 4:143–144
 colorectal cancer stem cells 3:810
 column chromatography 1:71
 combinatorial sensing and PRR crosstalk 3:721–722
 combined fractional diagonal chromatography (COFRADIC) 1:573–574, 1:574F, 1:575–576, 1:576, 1:578
 negative selection of N-termini by 1:574–575
 CoMFA *see* comparative molecule field analysis (CoMFA)
 commensal bacteria 3:690–691, 3:691–692
 common DC progenitor (CDP) 3:743
 CDP-diacylglycerol 1:167
 common lymphoid progenitors (CLP) 3:552, 3:628, 3:743
 common mediator SMAD (co-SMAD) 3:43
 common myeloid progenitors (CMPs) 3:743
 comparative genomics approach 4:154, 4:155F
 comparative molecule field analysis (CoMFA) 1:138
 compartmental model 4:17
 compartmentalization 1:626, 3:687
 compartmentalized ciliogenesis 2:665
 competitive endogenous RNA mechanism (ceRNA mechanism) 1:369, 4:160, 4:163, 4:255
 complement 5a (C5a) 3:752
 complement control protein domains (CCP domains) 3:728
 complement system 1:723
 activation pathways 1:724F
 alternative pathway 1:728, 1:728F
 C3bBb C3 convertase complex 1:729
 convertase complexes 1:729
 instability of thioester bond in C3 1:728–729
 classical pathway 1:723, 1:725F
 C1 complex 1:723
 C1-inhibitor 1:726
 C1s 1:725–726
 C2 activation by proteolytic cleavage 1:726
 C2 component of complement 1:726
 modules 1:724–725
 zymogen structures and activated forms of C1r 1:725
 lectin pathway 1:726, 1:727F
 GAGs 1:728
 MASP-1 and MASP-2 1:727
 MASP-3 1:727–728
 MASPs 1:726
 MBL 1:726
 regulation 1:729
 complementary deoxyribonucleic acid (cDNA) 3:58, 4:160, 4:220
 complementary determining region (CDR) 1:138, 1:144–145
 complementary metal-oxide semi-conductor (CMOS) 4:238–239
 complex data modeling 4:147
 complex I/NADH dehydrogenase 1:305
 complexes associated with tethering containing helical rods (CATCHR) 2:487–488
 complexin function model 2:461F
 complexus adhaerentes 2:707
 compucell3D 4:75, 4:125
 computational cardiac science 4:86, 4:92
 computational methods 1:100–101
 computational models of inflammation 4:98
 computational prediction methods 4:248
 concentration technique 1:70
 concerted models 1:9–10
 conditional deletion 2:528, 2:533
 conditional mutual information (CMI) 4:258
 conductance of that current 4:86
 confocal fluorescence microscopy 2:77
 confocal microscopy 2:51, 2:58–59, 2:68, 2:129
 conformational diseases 2:529
 congenital disease and cancer, mutations in HH pathway components in 3:81–83
 congenital disorders of glycosylation (CDG) 1:86, 2:339–340, 2:348
 congenital glycosylation disorders (CGD) 2:188
 congenital muscular dystrophies (CMD) 2:348–349, 2:349
 connective tissue 2:694, 2:698
 cells and matrices 3:192
 function
 cell adhesions in 3:194–195
 matrix molecules in 3:194–195
 connective tissue growth factor (CTGF) 1:695, 3:102
 connectivity theorems 4:28–29
 consensus sequence 1:365, 1:495–496, 1:496, 1:496F, 1:500, 1:503, 1:505, 1:555, 1:557
 conserved oligomeric Golgi (COG) 2:485, 2:487–488
 constitutive secretion 2:376, 2:377, 2:382
 constitutive secretory transport 2:371
 constitutive transport via endosomal hubs 2:364–365
 continuous-scale formalisms 4:119

- continuous-wave STED microscopy (CW STED microscopy) 2:78–79
- continuum modeling 4:74, 4:80, 4:83
- continuum techniques 4:81
- control analysis 4:28, 4:31–33
- control coefficients 4:27–28
rate-limiting steps 4:28
- control system, secretory pathway 2:373–374, 2:374F
- control theory 4:27
- controlled vocabulary 4:247
- conventional calpains structure 1:670
- conventional DCs (cDCs) 3:742
- conventional lysosomes 2:225
genetic diseases associate with defects in 2:230
- conventional myosins 2:597
- convoluted membranes 2:820–822
- coordinated lysosomal expression and regulation (CLEAR) 1:617, 1:770
- COPD *see* chronic obstructive pulmonary disease (COPD)
- COPI *see* coat protein complex I (COPI)
- COPII *see* coat protein complex II (COPII)
- copy number variations (CNVs) 4:246–247
- core 1-derived O-glycans 2:348
- core machinery 3:335
- core promoter 1:481, 1:482
- core promoter elements (CPEs) 1:481
- core recognition element (CRE) 1:472
- core RNA polymerase 1:468–469
- core vacuole/endosome tethering factor (CORVET factor) 2:203, 2:410–411, 2:412, 2:485, 2:488
- corin 1:750
- corin subfamily 1:656–657
- coronary heart disease (CHD) 1:718
- coronavirus budding compartment 2:169
- coronaviruses (CoV) 2:820–822
convoluted membranes and membrane webs generation by 2:820–822
formation of double-membraned vesicles by 2:821F
- corpse clearance, SRs in 3:733
- correct undamaged nucleotide insertion 1:447
- corrector 1:252, 1:261
- correlative light and electron microscopy (CLEM) 2:26–27, 2:32, 2:40–41
- correlative microscopy 2:22
- cortical actin 3:231
- cortisol 1:178–179
- CORVET factor *see* core vacuole/endosome tethering factor (CORVET factor)
- Corynebacterium glutamicum* (*C. glutamicum*) 1:57
- Cos2 *see* Costal2 (Cos2)
- co-SMAD *see* common mediator SMAD (co-SMAD)
- Costal2 (Cos2) 2:614, 3:76–77
- costamere 2:677
components 2:681T
- co-stimulatory molecule 3:645
- co-transcriptional splicing 1:506
- co-translational translocation 1:536–538, 1:538, 1:542
- coupled folding and binding 1:112
- coupling of ATP hydrolysis 1:7
- CoV *see* coronaviruses (CoV)
- covalent modification, enzyme regulation by 1:288, 1:290
- Cowden's syndrome (CS) 3:133, 3:419
AKT1 and PIK3CA mutations in 3:419
- Cox2 *see* coxidase's subunit 2 (Cox2)
- coxidase's subunit 2 (Cox2) 2:82
- Coxiella*-containing vacuole (CCV) 2:777
- CP *see* carboxypeptidases (CP)
- CPA *see* carboxypeptidase A (CPA)
- CPB *see* carboxypeptidase B (CPB)
- CPC *see* chromosomal passenger complex (CPC); chromosome passenger complex (CPC)
- CPD *see* cyclobutane pyrimidine dimers (CPD)
- CPEs *see* core promoter elements (CPEs); cytoplasmic polyadenylation elements (CPEs)
- CPI *see* capping protein interaction (CPI)
- CPM *see* Cellular Potts Model (CPM)
- CRC *see* colorectal cancer (CRC)
- CPO *see* carboxypeptidase O (CPO)
- CPV *see* cricket paralysis virus (CPV)
- CPY *see* carboxypeptidase Y (CPY)
- CRAC *see* calcium-release activated calcium (CRAC)
- C-Raf 3:123
- Crb complex 2:742–743
- Crb protein *see* Crumbs protein (Crb protein)
- Crb1–3 *see* Crumbs1–3 (Crb1–3)
- CRD *see* carbohydurate recognition domain (CRD); cysteine-rich domain (CRD)
- Cre recombinase 1:514–515
- CRE *see* core recognition element (CRE)
- Cre/lox 1:514–515
- creatine kinase, mitochondrial 1 (CKMT1) 3:347–348
- CREB binding protein (CBP) 1:665, 3:507
- CREB *see* cyclic AMP response element-binding protein (CREB)
- CREB-binding protein (CBP) 3:507
- CREBH *see* cAMP responsive element-binding protein, hepatocyte specific (CREBH)
- c-rel* KO mice 3:582
- Cre-Lox* systems 3:808–809, 3:809
- Creutzfeldt–Jakob disease (CJD) 1:113–114, 2:839, 2:840
- CRIB *see* CDC42-Rac interactive binding region (CRIB)
- cricket paralysis virus (CPV) 3:308
- Crispr-Cas system 3:519
- CRISPER *see* clustered regularly interspaced short palindromic repeat (CRISPER)
- Crispr-Cas system 3:519
- critical angle 2:62
- critical assessment of genome interpretation (CAGI) 4:259
- critical point drying 2:17
- CRLs *see* Cullin-Ring ubiquitin ligases (CRLs)
- CrmA (cytokine response modifier A) 3:332
- CrmA *see* cytokine response modifier A (CrmA)
- Crohn's disease (CD) 2:728–729, 3:584
- croquemort (crq) 3:733
- cross-linking and immunoprecipitation (CLIP) 4:192
- crossover 1:461, 1:464–465
- cross-presentation 3:746
- cross-repression 4:66
- CRP *see* c-AMP Receptor Protein (CRP)
- crq *see* croquemort (crq)
- Crumbs protein (Crb protein) 3:99–100
- Crumbs1–3 (Crb1–3) 3:102
- cryo-correlative light/electron microscopy (cryo-CLEM) 2:26–27
- cryo-electron microscopy (cryo-EM) 3:313
- cryo-electron microscopy of vitreous sections (CEMOVIS) 2:34–35
- cryo-electron tomography (cryo-ET) 2:22–31
of bacteria and virus 2:27F
examples 2:27–28
of eukaryotic cells 2:26F
flowchart of 2:23F
methods of 2:24
microscopy 2:24–25
specimen preparation 2:24, 2:24
techniques to combine with 2:26–27
tomogram reconstruction and 3D image analysis 2:25–26
specimen preparation methods 2:22–24
- cryo-EM 1:321, 1:529, 1:548, 1:548–549, 1:549F, 2:22, 2:22–23, 2:24, 2:27
single particle analysis 2:22, 2:26–27, 2:27, 2:28
- cryo-focused ion beam-scanning electron microscopy (cryo-FIB-SEM) 2:34–35
- cryo-immobilization 2:34–36
- cryopreparation methodology 2:36
- cryopyrin-associated periodic syndrome (CAPS) 3:29
- cryo-scanning electron microscopy (cryo-SEM) 2:34–35
- cryoscanning EM 2:18
- cryosectioning 2:34
- cryo-SEM *see* cryo-scanning electron microscopy (cryo-SEM)
- cryo-TEM 2:1, 2:22
- cryptic RSS 3:562–563
- crystal jelly (*Aequorea victoria*) 2:51–52
- CS *see* chondroitin sulfate (CS); Cowden's syndrome (CS)
- CS proteoglycans (CSPGs) 1:271
- CSB *see* Cockayne syndrome protein B (CSB)
- CSCs *see* cancer stem cells (CSCs)
- CSF *see* cerebrospinal fluid (CSF)
- CSF-1 *see* colony stimulating factor-1 (CSF-1)
- CSF2 *see* colony stimulating factor 2 (CSF2)
- CSFV *see* classical swine fever (CSFV)
- CSK *see* C-terminal SRC kinase (CSK)
- CSL *see* Lag-2 (CSL)
- cSMAC *see* central supramolecular activation cluster (cSMAC)
- CSP *see* cysteine-string protein (CSP)
- CSPGs *see* CS proteoglycans (CSPGs)

- CSR *see* class switch recombination (CSR)
 CST *see* *Clostridium spiroforme* toxin (CST);
 cystatins (CSTs)
 CT *see* cholera toxin (CT)
 CT694 Chlamydia effector 2:787
 CtBP1-S/BARS 2:503
 CTCs *see* circulating tumor cells (CTCs)
 CTD *see* C-terminal domain (CTD);
 carboxy-terminal domain (CTD)
 C-terminal 99 amino acids of APP (C99)
 1:696
 C-terminal DIX domain 3:90F
 C-terminal domain (CTD) 1:468–469,
 3:357–358, 3:429–430, 3:718–719,
 3:719
 C-terminal fragment of 99 amino acids
 (C99) 1:681
 C-terminal fragments (CTFs) 1:664
 CTF- β 1:661–662
 C-terminal mad homology 2 (MH2) 3:43
 C-terminal p6 domain of Gag 2:832–833
 C-terminal portion of BRCA-1 gene (BRCT)
 1:451
 C-terminal proline-arginine domain (PRD)
 2:446
 C-terminal SRC kinase (CSK) 3:537
 C-terminal α -amidation 1:86–87
 CTFs *see* C-terminal fragments (CTFs)
 CTGF *see* connective tissue growth factor
 (CTGF)
 CTLA-4 *see* cytotoxic T-lymphocyte antigen
 4 (CTLA-4)
 CTLD *see* C-type lectin-like domains
 (CTLD)
 CTLs *see* cytotoxic T lymphocytes (CTLs)
 cTnI *see* cardiac troponin I subunit (cTnI)
 CTP-synthase (CtpS) 2:593
 C-type lectin receptors (CLRs) 3:588, 3:595,
 3:710, 3:710–712, 3:711T
 C-type lectin-like domains (CTLD) 2:848,
 3:710–712, 3:727
 C-type lectins 2:848
 C-type natriuretic peptide (CNP) 1:750,
 3:58
 CUB domain proteins *see* C1r/C1s, Uegf,
 and Bmp1 (CUB) domain proteins
 cubitus interruptus (CI) 3:77–78
 cue-mediated polarization of yeast cells
 4:51–52, 4:52F
 Cullin proteins 1:584T, 1:587
 Cullin-Ring ubiquitin ligases (CRLs) 1:585,
 1:587
 cultured cells, long-term fluorescence
 movies of 2:68
 cumulus oophorus complex (COC) 1:636
 cup closure 2:760, 2:761
 curtaining defects 2:46
 curvature 3:193–194
 CypA 2:777
 CW STED microscopy *see* continuous-wave
 STED microscopy (CW STED
 microscopy)
 CWD *see* chronic wasting disease (CWD)
 CXC chemokine ligand (CXCL) 3:616
 CXC chemokine receptor (CXCR) 3:616
 CXCR5 3:543
 C-X-C motif chemokine 12 (CXCL12)
 3:624
 in B cell lineage maintenance 3:628–629
 in balancing granulopoiesis and
 lymphopoiesis 3:629
 in homing to BM 3:624–625
 in HSC and HPC
 maintenance within BM niches
 3:625–626
 mobilization 3:626–627
 in neutrophils maintenance 3:627–628
 C-X-C motif chemokine 13 (CXCL13) 3:543
 CXCL *see* CXC chemokine ligand (CXCL)
 CXCL12 *see* C-X-C motif chemokine 12
 (CXCL12)
 CXCR *see* CXC chemokine receptor (CXCR)
 CyA activity *see* cyclase activity (CyA
 activity)
 cyan fluorescent protein (CFP) 2:52, 2:54F,
 4:221–222
 cyanobacteria 2:593–594
 2-cyanoethyl group 1:43, 1:43F
 2-cyanoethylloxymethyl (CEM) group
 1:44–45, 1:45F
 cyanogen bromide (CNBr) 1:77
 cyclase activity (CyA activity) 2:795
 cyclic adenosine monophosphate (cAMP)
 1:290, 1:294, 1:296, 2:795, 3:59–60,
 3:102–103, 3:194
 -dependent protein kinase (PKA) 2:663
 cyclic AMP response element-binding
 protein (CREB) 3:130, 3:712–714
 cyclic dinucleotides (CDNs) 3:719
 cyclic GMP-AMP (cGAMP) 3:520
 cyclic guanosine monophosphate (cyclic
 GMP) 3:56
 cGMP-dependent protein kinase (PKG)
 2:663, 3:56
 cyclic nucleotides (cN) 3:59–60
 cyclic pathway 4:29–30, 4:30F
 cyclic-di-GMP 1:375–377, 1:377, 1:378T
 cyclin 3:424, 3:437
 in cancer
 mutations 3:428–429
 overexpression 3:429
 beyond cell cycle 3:429–430
 cyclin A 3:424
 MPF complex 3:427F
 superfamily of proteins 3:425T
 cyclin-CDK activity, associations and
 modifications regulating 3:427F
 cyclin-cyclin-dependent kinase (cyclin-CDK)
 3:437
 associations and modifications regulating
 3:427F
 in cell cycle 3:425–428
 in quiescence 3:425–428
 cyclin-dependent kinase (CDK) 3:44, 3:130,
 3:424, 3:424–425, 3:434, 3:458,
 3:478–479
 in cancer 3:428–429
 Cdk1 2:317, 3:399, 3:499, 3:500F
 aurora regulation by 3:499
 inactivation 3:499
 CDK2 4:222
 Cdk5 3:231
 beyond cell cycle 3:429–430
 superfamily of proteins 3:426T
 CIP/KIP family 3:438–439
 deregulation in human cancers 3:440–441
 CIP/KIP family 3:441–442
 human malignancies 3:441F
 INK4 family 3:441
 INK4 family 3:437
 p21 and p27 3:443F
 as predictors of drug response 3:442–444
 cyclobutane pyrimidine dimers (CPD)
 3:462–463
 CYP7A1 *see* cholesterol 7- α hydroxylase
 (CYP7A1)
 Cyr, CTGF, Nov proteins (CCN proteins)
 2:700
 CYR61 *see* cysteine-rich angiogenic inducer
 61 (CYR61)
 Cys-palmitoylation 1:555
 cystatin-like molecule (CLM) 1:707
 cystatins (CSTs) 1:706
 cystatin C 3:590, 3:593–594
 biological functions 1:707
 and disease 1:707
 structure and inhibition mechanism
 1:706–707
 tissue distribution 1:707
 cysteine cathepsins 1:614, 1:614F, 1:615,
 1:754
 cysteine/histidine (CH) 3:319–320
 cysteine proteases 1:710–711
 calpains 1:753
 caspases 1:753–754
 cysteine cathepsins 1:754
 cysteine-rich angiogenic inducer 61 (CYR61)
 3:102
 cysteine-rich domain (CRD) 1:632, 3:61,
 3:87–88, 3:88–89, 3:88F, 3:355
 cysteine-string protein (CSP) 2:297
 cystic fibrosis (CF) 1:252–262, 1:694
 airway cilia in 1:256F
 and airways 1:255–256
 genetics 1:259–261
 mutations in CFTR 1:260–261
 pancreatic fluid in 1:257
 pharmacologic approaches to treat
 patients with CF 1:261
 protein 1:252–255
 CFTR and bicarbonate transport
 1:257–259
 CFTR and chloride transport 1:255–256
 CFTR and protein trafficking 1:254–255
 folklore, heat waves, and diagnostic test
 1:256–257
 ion transport 1:255
 related diabetes 1:257
 sweating in 1:257
 cystic fibrosis transmembrane conductance
 regulator (CFTR) 1:113, 1:252–253,
 1:596, 2:176, 2:795, 3:58–59,
 3:691
 and bicarbonate transport 1:257–259

- CFTR gene 1:252, 1:252–253, 1:253–254, 1:255
 and chloride transport 1:255–256
 core-glycosylated 1:254–255
 interaction with coat protein complex II (COPII) 1:255
 mutations in 1:260–261
 affecting quality 1:261
 affecting quantity 1:260–261
 and pancreatic secretions 1:257–259
 and protein trafficking 1:254–255
 cystine knot family 3:39
 CYTC *see* cytochrome *c* (CYTC)
 cytochalasin D 2:557, 2:761–762
 cytochalasin E 2:761–762
 cytochrome *b*₅₅₈ 3:702
 cytochrome *b*_{6f} complex 1:301–303
 as lipoprotein 1:302–303
 structures, mechanisms, and mechanisms of H⁺ transfer 1:303
 cytochrome *c* 3:328, 3:328F, 3:330–331, 3:331, 3:332, 3:369–370, 3:369F
 cytochrome oxidase 1:306
 CyTOF *see* cytometry by time of flight (CyTOF)
 cytogenetics 4:155
 cytokine inducible SH2-containing protein (CIS) 3:136, 3:145
 cytokine receptors 3:22–23
 βc receptors 3:24
 family tree of human cytokines and receptors 3:25–26F
 γc receptors 3:24–25
 IL-2/IL-15 receptors 3:25–26
 IL-4/IL-13 receptors 3:26
 IL-7/TLSP receptors 3:26–27
 hGHR 3:23, 3:23–24
 IFN receptors 3:30–31
 IL-1 receptors 3:28
 IL-10/IL-22 receptors 3:29–30
 IL-17 receptors 3:30
 intracellular signaling complexes 3:31–32
 shared glycoprotein 130 receptors 3:27–28
 cytokine response modifier A (CrmA) 1:710–711
 cytokine-binding modules (CBM) 3:22–23
 cytokines 1:622–624, 1:624–625, 3:22, 3:136–140, 3:581, 3:582, 3:584
 and TNF-therapy 4:98–99
 cytokinesis 2:403–404, 2:466, 2:468, 2:471, 2:601, 2:602, 3:5, 3:481–482, 3:494–503
 abscission 3:497–498
 checkpoint 3:501–502
 chromatin bridges 3:501
 in animal cells 3:498F
 cleavage furrow formation 3:495
 CPC during 3:475
 endosomal sorting complexes required for transport (ESCRT) in 2:470
 membrane
 abscission 3:484–486
 ingression 3:482–484, 3:484F
 trafficking pathways 3:485F
 midbody formation 3:496–497
 in mitotic spindle 2:644–645
 positioning site of cell division
 RhoA pathway 3:494–495
 spindle microtubules 3:494, 3:495F
 stages 3:494F
 temporal regulation 3:499
 aurora regulation 3:499
 Cdk1 inactivation 3:499
 Plk1 3:499–501
 cytolethal distending toxins (CDTs) 2:797
 cytolinker 2:569, 2:572–573
 cytolysis 2:569, 2:575
 cytomegalovirus (CMV) 1:60, 3:569, 3:773
 cytometry 4:238–239, 4:242
 cytometry by time of flight (CyTOF) 3:569–570
 cytomotive filaments 2:588, 2:592
 cytoplasm 1:205, 1:232–234, 1:246, 1:248, 1:320–321, 1:355, 1:356–357F, 1:359, 1:365, 1:495, 1:504F, 1:507, 1:517, 1:520, 1:521, 1:523, 1:527–529, 1:536F, 1:764–765, 1:771, 2:2–3, 2:25–26, 2:27, 2:28, 2:466, 2:467–468, 2:467F, 3:1, 3:2, 3:3, 3:5, 3:92, 3:92–93, 3:94, 3:122, 3:123F, 3:125, 3:126, 3:328, 3:329, 3:330, 3:331, 3:332, 3:791
 cytoplasmic
 aggregates 2:575, 2:576
 DHCs 2:620–621, 2:621F
 events in ribosome biogenesis
 40S subunit 3:302–303, 3:302F
 60S subunit 3:301–302, 3:301F
 filaments 2:63
 intermediate filament
 dysfunction effect 2:575–576
 network organization 2:572–573
 maturation
 40S subunit 3:302–303, 3:302F
 60S subunit 3:301–302, 3:301F
 mRNAs 1:346
 plaque 3:184
 ring 2:319
 streaming 2:597, 2:603–604
 cytoplasmic dynein (DYNC) 2:620, 2:625
 see also axonemal dyneins
 axonal transport 2:625
 cell motility 2:625
 cellular cargoes transport 2:624
 DYNC1 2:620
 DYNC2 2:620, 2:666
 function and adaptors 2:626F
 IFT dynein 2:625–626
 mitosis 2:625
 mRNA transport 2:624–625
 nuclear migration 2:625
 spindle assembly 2:625
 cytoplasmic dynein2 heavy chain (DYNC2H1) 2:666
 cytoplasmic linker associated protein 2 (CLASP2) 2:708–709
 cytoplasmic linker protein (CLIP) 3:482
 cytoplasmic polyadenylation elements (CPEs) 1:344
 cytoskeletal dynamics 3:680–681
 cytoskeletal rearrangements 2:746
 cytoskeletal remodeling 3:637
 endothelial remodeling 3:637
 T lymphocyte remodeling 3:637–639
 cytoskeleton 2:415, 2:572–573, 2:751–752, 3:52, 4:12
 see also actin cytoskeleton
 exotoxins altering 2:795–796
 and motors 2:1, 2:3–4
 structure 2:92–93
 cytoskeleton-associated protein, glycine-rich domain (CAP-Gly domain) 2:624
 cytosol 2:780
 cytosolic bacteria 2:790
 cytosolic Ca⁺⁺ calcium concentration 2:679–680
 cytosolic calcium 2:794
 cytosolic carboxypeptidase (CCPP-60) 2:617
 cytosolic chaperonin-containing TCP-1 (CCT) 2:580–581
 cytosolic components structure in resting cells 3:703–704, 3:705F
 p40phox 3:704
 p47phox 3:704
 p47phox–p67phox–p40phox complex 3:704–705
 p67phox 3:704
 Rac 3:704
 cytosolic molecular chaperone proteins 1:538
 cytosolic proteins 1:202, 1:202–203, 1:204–205, 1:205, 1:248, 1:540, 1:764, 2:434–435
 cytosolic sensing pathways multiscale regulation 3:722–724
 cytotoxic necrotizing factor 1 (CNF1) 2:796
 cytotoxic T lymphocyte antigen 4 (CTLA-4) 3:606F, 3:607–608
 and CD28 binding same ligands with different affinities 3:608–609
 cell biology of 3:608F, 3:609–610
 immune regulation 3:610–611, 3:611
 mechanism of CTLA-4-mediated suppression 3:609
 transendocytosis 3:609–610
 Treg suppression 3:611–612
 cytotoxic T lymphocytes (CTLs) 2:226, 2:227F, 2:438–439, 2:835, 3:330, 3:567, 3:634, 3:667
 cytotoxic T-lymphocyte antigen 4 (CTLA-4) 1:145
- ## D
- DA *see* dopamine (DA)
 DAB *see* diaminobenzidine (DAB)
 Dab1 *see* disabled homolog 1 (Dab1); disabled-1 (Dab1)
 Dachous (Ds) 3:100
 DAF-2/DAF-16 pathway 2:849–850
 DAG *see* diacylglycerol (DAG); directed acyclic graphs (DAG); driven diacylglycerol (DAG)

- damage associated molecular patterns (DAMPs) 1:708–709, 2:523–524, 3:368, 3:710, 3:752, 4:98
- DAMPs *see* damage associated molecular patterns (DAMPs); danger-associated molecular patterns (DAMPs)
- danger-associated molecular patterns (DAMPs) 2:790
- Danon disease 1:764–765, 1:770, 1:771–772, 1:775, 1:778
- DAPK *see* death-associated protein kinase (DAPK)
- DAPK1 *see* death-associated protein kinase 1 (DAPK1)
- darbepoetin alfa 1:126
- dark zone (DZ) 3:543
- dark-state exchange saturation transfer (DEST) 1:98, 1:104–105
- DAT *see* dopamine transporter (DAT)
- data analysis 1:32–33
- frequency-domain data 2:117
 - time-domain data 2:116–117
- data availability 4:247
- data integration 4:266
- workflow 4:246F
- data interpretation
- critical thinking in 4:249
 - identification of suitable resources and 4:247
- data processing inequality (DPI) 4:258
- data quality and heterogeneity 4:246–247
- data structure standardization 4:247–248
- data-dependent acquisition (DDA) 4:167–168
- data-driven approach *see* bottom-up approach
- data-independent acquisition (DIA) 4:168–169
- daughter centriole (DC) 2:649, 2:650F, 2:651, 2:656
- DBA *see* diamond-Blackfan anemia (DBA)
- DBD *see* DNA binding domain (DBD)
- DBI *see* diazepam binding inhibitor (DBI)
- Dbl homology domain (DH domain) 2:499, 2:731–733, 3:211
- DBL-1 pathway 2:850
- DC *see* daughter centriole (DC); dendritic cells (DCs); dyskeratosis congenita (DC)
- DCE *see* downstream core element (DCE)
- DCF *see* dihydrodichlorofluorescein (DCF)
- DCGR8 1:365, 3:788
- DCGs *see* dense core granules (DCGs)
- Dcl1 *see* doublecortin-like kinase 1 (Dcl1)
- DCs *see* dendritic cells (DCs)
- Dct *see* dopachrome tautomerase (Dct)
- DD *see* death domain (DD)
- DDA *see* data-dependent acquisition (DDA)
- DdDp *see* DNA-dependent DNA polymerase (DdDp)
- DDFT *see* dynamic density functional theory (DDFT)
- DDP-IV *see* dipeptidyl peptidase IV (DDP-IV)
- DDR *see* DNA damage response (DDR)
- DDR1 *see* discoidin domain receptors 1 (DDR1)
- de novo* creation 4:59
- De Novo Driver Exclusivity (Dendrix) 4:212T, 4:214
- De novo membrane formation 2:262
- of peroxisome membranes from ER doubts 2:268
 - evidence for *de novo* formation in mutants 2:267–268
 - physiologically 2:268
- De novo* oligonucleotide synthesis 1:47
- deacetylation 3:507
- deadenylation 1:346, 3:317, 3:321–322, 3:323–324
- DEAE *see* diethylaminoethyl (DEAE)
- death domain (DD) 3:62, 3:67, 3:68, 3:70, 3:72, 3:356–357
- death effector domain (DED) 3:69, 3:358
- death receptors (DR) 3:62, 3:67, 3:343
- alternative names for 3:68T
 - function and signaling pathway Fas clade 3:70
 - p75^{NTR} clade 3:72
 - TNFR1 clade 3:71–72
 - ligand activation 3:67–68
 - members and ligand interactions 3:68F
 - multiple signal pathways 3:69
 - extrinsic apoptotic pathway 3:69
 - intrinsic apoptotic pathway 3:69–70
 - non-apoptotic pathways 3:70
 - signaling 3:68–69
- death signaling ligands 3:3
- death-associated protein kinase (DAPK) 3:162–163
- DAPK1 3:343
- death-induced signaling complex 3:327–328, 3:328F, 3:329, 3:332, 3:369F, 3:370, 3:371, 3:371–372
- death-inducing signaling complex (DISC) 3:69, 3:358
- debranching enzyme 1:266–267, 1:267–268
- decapping 1:346, 1:347, 1:348, 3:317, 3:321–322, 3:323
- decorin 1:273–274
- dectin-1 activation by β -glucans 3:696
- DED *see* death effector domain (DED)
- Ded1 1:527
- deep learning 4:147, 4:149
- deep vein thrombosis (DVT) 1:718–719
- DEER *see* double electron–electron resonance (DEER)
- defective ribosomal products (DRiPs) 3:589
- defensins 2:771
- degenerative disease 2:575
- DE genes *see* differentially expressed (DE) genes
- degradation 2:201, 2:205, 2:408, 2:409F
- degradative enzymes 2:771
 - of membrane receptors 2:410
 - of transported proteins 2:410
- degradome 1:568, 1:570, 1:574–575
- degradomics 1:568, 1:570–572, 1:579, 1:622–624
- data, validation of 1:577–579
- identification of protease cleavage motifs 1:576–577
- identification of protease substrates 1:570–572
- classical gel-based methods 1:571–572
 - gel-free methods 1:572–574
 - relative quantification using dimethylation 1:574
- negative selection of N-terminal peptides by TAILS 1:575–576
- negative selection of N-termini by COFRADIC 1:574–575
- positive selection of N-terminal peptides by subtiligase biotinylation 1:576
- degradosome 3:90
- degranulation 3:752
- 7-dehydrocholesterol reductase (DHCR7) 3:77
- delta 2:350, 2:351
- Δ F508 1:260–261
- delta-like 4 (Dll4) 4:112
- delta-like ligand1 (DLL1) 3:107
- Delta-Notch signaling 4:128
- delta–Serrate LAG-2 domain (DSL domain) 3:107
- denatured state 1:109
- dendrite 3:230, 3:231, 3:232
- dendritic cell (DC) subsets
- antigen processing and presentation 3:589
 - autophagy 3:591–592
 - classical MHC class II processing 3:589–590
 - classical MHC class I processing 3:589
 - cross presentation 3:590–591
- DC development and characterization 3:592
- discovery of 3:588–589
- functional specialization of 3:588–604
- human DC subpopulations, antigen presentation of 3:596–597
- lymphoid tissue DCs 3:596–597
 - non-lymphoid tissue (NLT) DCs (skin) and migratory DCs 3:597
- mouse DC subpopulations, antigen presentation of 3:592–593
- lymphoid tissue-resident DCs 3:593–594
 - non-lymphoid tissue DC 3:595
 - plasmacytoid DCs 3:592–593
- dendritic cells (DCs) 2:836, 3:521, 3:541, 3:581, 3:616, 3:710, 3:741–749, 3:759, 3:770
- in clinical immunology 3:747
 - in autoimmune disease 3:747–748
 - in cancer 3:748
 - targeted vaccines 3:748
 - in transplantation 3:747
 - in viral infections 3:748
- controlling quality of T cell response 3:747
- development and homeostasis
- cellular origin 3:743
 - growth factors and DC cultures 3:742–743
 - homeostasis and regulation 3:745
 - human DC development 3:743–745

- distribution 3:741
 in lymphoid organs 3:741
 migration 3:741–742
 in non-lymphoid organs 3:741
 subsets 3:742, 3:744F
- homeostatic feedback loop 3:745F
- identification and isolation 3:741
- maturation 3:746–747
- as orchestrator of immunity 3:745
 antigen capture 3:746
 antigen presentation 3:746
 antigen processing 3:746
 in innate immunity 3:745
 origin and development 3:744F
 in tolerance 3:747
- dendritic nucleation 4:16, 4:16–17
- model 2:690
- open model of actin 4:18F
- dendritic transport of essential molecules 2:611
- Dendrix *see* De Novo Driver Exclusivity (Dendrix)
- denoising 2:25, 2:25F
- dense core granules (DCGs) 4:157
- dense core vesicles 2:390
- dense fibrillar component (DFC) 3:254
- dense granules 1:715
- dense-core granules 2:378–379
- density gradient centrifugation 1:66, 1:70
- 1-dexoy-*N*-pentynyl glucosamine derivative 1:128–130
- deoxyribonuclease I-like activity (DNase I-like activity) 2:797
- deoxyribonucleic acid (DNA) 1:2, 1:25–26, 1:135–136, 1:375, 1:418–419, 1:438T, 1:495, 1:518–519, 1:519, 1:521, 1:523, 1:545, 3:95, 3:432, 3:512, 3:786–787
- adducts 1:27–28
- damage 1:448, 1:451
 nucleolar dynamics in 3:260–263
 and repair 3:462, 3:777
 signaling 3:460–461
- binding domains 1:92–93
- exotoxins altering 2:797
- glycosylase 1:442–443, 1:443, 1:448
- hydrolysis 1:14
- ligase 1:442–443, 1:447, 1:448
- mapping 1:28–29
- melting curve 1:28F
- modification by radiation 1:28
- nick ligation 1:447
- phosphodiester linkages in 1:14
- polymerase β 1:442–443, 1:447, 1:447–448, 1:448
- protein assemblies 2:91
- repair mechanisms *see* DNA repair
- replication 1:396–397, 1:456, 1:461, 1:462–463, 3:427, 3:458
 definition 1:396
 DNA polymerase (Pol) I for 1:397
 helicases *see* helicases
 independent evolution of 1:397F
 needs for 1:411–412
- replication factories 3:459–460
- replication Fork formation 3:458–459
- replication Fork 3:459
- segregation 2:592
 of chromosomal DNA or plasmids by WACAs 2:593
 of plasmids by actin-like ParM 2:592
 of plasmids by tubulin-like TubZ 2:592–593
- strand
 break ends processing 1:446–447
 exchange 1:456
 invasion 1:460–461
- supercoiling 1:490–491
- synthesis 1:32
- synthesis phase (S phase) 3:447
- viruses
 generating replication compartments 2:825–826, 2:825F
 NCLDV 2:824–825
- dependence receptors 3:80, 3:343
- depolarization 1:218, 1:220F, 1:221, 1:225, 1:225–226, 1:229, 1:229–230, 1:244–245, 1:246, 1:246–247, 1:247F, 1:369, 1:507
- depor 3:243–244
- dermatan sulfate (DS) 1:271, 2:351
- dermatosparaxis 1:634
- dermis 3:219–220
- DESC *see* differentially expressed in squamous cell carcinoma gene (DESC)
- descriptors 1:138
- desensitization 1:229
- desert HH (DHH) 3:76
 DHH1 1:351
- DESI *see* desorption electrospray ionization (DESI)
- desmin mutations 2:576
- desmoplakin (Dp) 2:707
- desmosomes 2:707–708, 2:707F
- desorption electrospray ionization (DESI) 1:152–153
- DEST *see* dark-state exchange saturation transfer (DEST)
- detergent-resistant membranes (DRM) 1:208–209, 1:212, 1:213
- deubiquitylating enzymes (DUBs) 1:437, 1:583, 1:583F, 1:586, 1:589F, 3:505–506, 3:505–506
 in p53 pathway 3:506
- developmental vasculogenesis 4:102–104
- DExH/H box 1:495, 1:498–499, 1:500, 1:500
- dextran sulfate sodium (DSS) 3:810–811
- DFC *see* dense fibrillar component (DFC)
- DFCP1 *see* double FYVE containing protein 1 (DFCP1)
- DFP *see* diisopropyl-phosphofluoridate (DFP)
- dg see* Diego (*dg*)
- DGCR8 (DiGeorge syndrome critical region gene 8) 1:354, 1:355, 1:355–358, 1:356–357F, 1:357F, 1:358, 1:360, 1:365
- DGDG *see* digalactosyldiglyceride (DGDG)
- DH domain *see* Dbl homology domain (DH domain)
- DHB *see* dihydroxy benzoic acid (DHB); DNA helicase B (DHB)
- DHC *see* dynein heavy chain (DHC)
- DHCer *see* dihydroceramides (DHCer)
- DHCR7 *see* 7-dehydrocholesterol reductase (DHCR7)
- DHH *see* desert HH (DHH)
- dHJ pathway *see* double Holliday junctions pathway (dHJ pathway)
- DHPR *see* dihydropyridine receptors (DHPR)
- DHR1 *see* DOCK homology regions 1 (DHR1)
- DHX29 1:527
- DIA *see* data-independent acquisition (DIA)
- diabetes 3:17, 3:642
 oncogenes and tumor suppressor genes in 3:419–420
 FYN tyrosine kinase 3:420
 LKB1-AMPK pathway 3:420
 PI3K-AKT pathway 3:420
 and regulated secretion 2:379
 UPR signaling contribution 2:533
- diabetic retinopathy 4:77–78, 4:77F
- DIABLO *see* direct IAP-binding protein with low PI (DIABLO)
- diabodies 1:142–143
- diacylglycerol (DAG) 1:195–196, 2:276, 2:447, 2:505, 2:759, 2:760–761, 3:153
- Dialog on Reverse Engineering Assessment Methods (DREAM) 3:435, 4:257, 4:259
- dialysis 1:66, 1:70
- diaminobenzidine (DAB) 2:37–39
- diamond-Blackfan anemia (DBA) 3:282–283
- DIANA-microT 1:365
- Diap1 *see* *Drosophila melanogaster* inhibitor of apoptosis protein 1 (Diap1)
- DIAP1 *see* *Drosophila* inhibitor of apoptosis 1 (DIAP1)
- diapedesis 3:210, 3:634
 dynamic remodeling in endothelium and lymphocytes 3:638F
 mechanisms for adhesion and diapedesis
 cytoskeletal remodeling 3:637
 molecular basis for adhesion and diapedesis 3:636–637
- Diaphanous-related formins (DRFs) 2:558T
- diastereomer 1:741
- diazepam binding inhibitor (DBI) 3:799–800
- DIC *see* dynein intermediate chain (DIC)
- Dicer 1:354, 1:356–357F, 1:357F, 1:358, 1:359, 1:360, 1:365, 3:788, 3:789–790, 3:790
 cleavage of pre-miRNAs into miRNA-miRNA* duplexes by 1:355–358
 mutation 1:360
- Dickkopf-1 (Dkk1) 3:88–89, 3:88F, 3:89F
- Dictyostelium* cells 4:6

- Dictyostelium diskoidium*, whole cell tomography of 2:26F
- Diego* (dg) 2:746–747
- diethylaminoethyl (DEAE) 1:71, 3:60
- diethylpyrocarbonate 1:27–28, 1:30F
- differential analysis 4:202
- differential centrifugation 1:66
- differential in gel electrophoresis (DIGE) 1:571, 1:571–572, 1:571F
- differential precipitation 1:70
- differential-adhesion driven cell sorting 4:126–127
- differentially expressed genes (DE genes) 4:213
- differentially expressed in squamous cell carcinoma gene (DESC) 1:654, 1:654–655
- DESC-1 1:654–655
- differentiation/senescence 3:440
- diffraction 2:76
- diffraction limit 2:73
- diffraction-limited microscopy 2:62
- diffuse intrinsic pontine gliomas (DIPG) 3:47
- diffuse large B-cell lymphoma (DLBCL) 3:585
- diffusion-based algorithm 4:212T, 4:213, 4:214
- digalactosyldiglyceride (DGDG) 1:181
- DIGE *see* differential in gel electrophoresis (DIGE)
- DiGeorge syndrome 1:355, 1:357F, 1:360
- digestive proteases 1:732
- proteolysis
- large intestine role in 1:735–736
 - oral cavity role in 1:732–733
 - small intestine role in 1:733–735
 - stomach role in 1:733
- digital approach 4:65
- dihydroceramides (DHCer) 1:187
- dihydrodichlorofluorescein (DCF) 3:816
- dihydropyridine receptors (DHPR) 2:679–680
- dihydroxy benzoic acid (DHB) 4:178
- diisopropyl-phosphofluoridate (DFP) 1:131
- dikinase 1:267, 1:268–269
- dimensionality reduction approaches 4:257–258
- dimeric motors 2:600–601
- dimerization
- domain 3:58
 - inhibitors 1:744
 - RTKs undergo ligand-induced 3:9–11
- dimethyl sulfate (DMS) 1:27
- 2,2'-dioleoyl isoform 2:205
- dipeptidyl aminopeptidase A (DPAPA) 2:478
- dipeptidyl peptidase IV (DDP-IV) 1:735, 3:625
- DIPG *see* diffuse intrinsic pontine gliomas (DIPG)
- diphtheria toxin (DT) 2:796–797
- diphtheria toxin receptor (DTR) 3:760–761
- dipole orientation 2:64
- direct electron detection 2:6, 2:10F
- direct IAP-binding protein with low PI (DIABLO) 3:370
- direct kinase inhibition by SOCS1 and SOCS3 3:144–145
- direct protein–protein interactions 3:124, 4:213
- direct reversal mechanisms 3:462
- directed acyclic graphs (DAG) 4:257
- directed cell migration 3:208, 3:209F
- chemotaxis 3:208–209
 - haptotaxis 3:209
 - mechanotaxis 3:209
- directed mutagenesis 1:122
- director of cytokinesis 180 (DOCK180) 2:733
- disabled homolog 1 (Dab1) 2:423
- disabled homolog 2 (Dab2) 2:422–423
- disabled-1 (Dab1) 2:455
- DISC *see* death-inducing signaling complex (DISC)
- discoidin domain receptors 1 (DDR1) 3:16
- discoidin domain receptors 2 (DDR2) 3:16
- discrete modeling 4:80, 4:81, 4:83
- discretization techniques 4:118–119
- discs-large homolog (Dlg) 2:743, 3:100
- 'disease modules' 4:266
- disease-associated tRNA modification 1:334
- Disheveled (Dvl) 3:90, 3:90F, 3:91, 3:91F, 3:92
- Dishevelled* (dsh) 2:746–747
- disordered proteins 1:112
- cellular phase separations 1:113
 - coupled folding and binding 1:112
 - flexible linker 1:112
- DISP *see* dispatched (DISP)
- dispatched (DISP) 3:76
- dissecting biological noise 3:674–675
- dissolution 1:461
- distal appendages 2:649, 2:650F, 2:653
- distal SMAC (dSMAC) 3:522–524, 3:660
- distance information sources 1:100
- distributive enzyme 1:396
- disulfide bond formation 2:162
- disulfide bonds 1:125, 1:538, 1:539–540T, 1:540, 1:542, 1:543, 1:571, 1:577, 3:86, 3:87–88, 3:590
- disulfide pairing 1:78
- dithiothreitol (DTT) 1:69
- DivIVA peripheral membrane protein 2:593
- DksA 1:476
- DLBCL *see* diffuse large B-cell lymphoma (DLBCL)
- Dlg *see* discs-large homolog (Dlg)
- Dlg1 *see* *Drosophila* disc large tumor suppressor (Dlg1)
- DLK *see* dual leucinezipper kinase (DLK)
- DLL1 *see* delta-like ligand1 (DLL1)
- Dll4 *see* delta-like 4 (Dll4)
- D-loop 1:458, 1:460, 1:461, 1:463–464, 1:464–465
- DLP *see* dynamin-like proteins (DLP)
- DMS *see* dimethyl sulfate (DMS)
- DMT1-II *see* SLC11a2 iron transporter
- DMVs *see* double membrane vesicles (DMVs)
- DN cell *see* double negative cell (DN cell)
- DN1L-mediated mechanism *see* dynamin 1-like-mediated mechanism (DN1L-mediated mechanism)
- DNA *see* deoxyribonucleic acid (DNA)
- DNA binding domain (DBD) 1:482, 3:143
- DNA damage response (DDR) 1:427, 3:550, 3:560
- DNA helicase 1:396
- see also* helicases
- DNA helicase B (DHB) 4:222
- DNA ligase 1:435, 1:437F, 1:438T, 1:439
- DNA looping 1:409, 1:409–410, 1:411
- DNA methylation 3:6–7, 3:785, 3:790, 3:790–791
- in ESCs 3:787–788
- DNA methyltransferases 3:788
- DNA microarrays 1:46–47
- electrochemical arrays 1:47–49
 - gene assembly from 1:49–50
 - inkjet printing 1:47
 - photolithography with physical/digital masks 1:46–47
- DNA polymerase 1:122, 1:122F, 1:125, 1:396, 1:397, 1:398–400, 1:418–419, 1:419F, 1:421, 1:421–422, 1:423, 1:435, 1:435F, 1:439, 1:440, 1:521
- bacterial (C family) and eukaryotic (B family)
 - crystal structure 1:400
 - functional overview 1:398–400
- DNA repair 1:400, 1:435, 1:456, 3:550, 3:557, 3:560, 3:561F
- see also* eukaryotic nucleotide excision repair
- BER pathway 3:462
- core mechanism of homologous recombination 1:456–457
 - DNA strand invasion and associated DNA synthesis 1:460–461
 - DSB resection 1:456–457
 - Rad51 filament assembly and stability 1:457–460
- direct reversal 3:462
- double-strand break repair 3:465–466, 3:466
- HR and human disease 1:465
 - pathway regulation and crossover control 1:464–465
 - recombination and DNA replication 1:462–463
 - recombination and double-strand break repair 1:461
 - BIR 1:462
 - dHJ 1:461
 - SDSA 1:461–462
 - recombination and interstrand cross-link repair 1:463–464
 - MMR pathway 3:463–464
 - NER pathway 3:462–463
 - TLS 3:464–465
- DNA synthetic cycle 1:46–47
- DNA viruses, replication strategies of 1:519F
- DNA-dependent DNA polymerase (DdDp) 2:819

- DNA-dependent protein kinase (DNA-PK) 3:129–130
- DNA-dependent protein kinase catalytic subunit (DNA-PKcs) 3:559–560
- DNAJC13 *see* RME-8
- DNA-PK *see* DNA-dependent protein kinase (DNA-PK)
- DNA-PKcs *see* DNA-dependent protein kinase catalytic subunit (DNA-PKcs)
- DNase I-like activity *see* deoxyribonuclease I-like activity (DNase I-like activity)
- DNCs *see* double-negative cells (DNCs)
- Dnmt1 3:790, 3:790–791
- Dnmt3a 3:790, 3:790–791
- DOCK homology regions 1 (DHR1) 2:733
- DOCK homology regions 2 (DHR2) 2:733
- DOCK9 2:733
- DOCK11 2:733
- DOCK180 *see* director of cytokinesis 180 (DOCK180)
- docking structures 3:637
- Dok-R *see* Downstream of kinase-R (Dok-R)
- dolichol-phosphate (Dol-P) 2:339–340, 2:340
- Dol-P *see* dolichol-phosphate (Dol-P)
- Dom34 3:320
- domain swapping 1:108, 1:112F
- dominant negative Rab mutant 2:444
- domineering nonautonomy 3:200, 3:201F, 3:204
- ‘don’t eat me’ signals 2:753–754
- dopachrome tautomerase (Dct) 2:237
- dopamine (DA) 1:248, 1:248–249
- dopamine transporter (DAT) 2:499
- dorsal–ventral axis (DV axis) 3:37, 3:78
- double electron–electron resonance (DEER) 1:100
- double FYVE containing protein 1 (DFCP1) 2:503, 2:511
- double Holliday junctions pathway (dHJ pathway) 1:459, 1:461
- double membrane autophagosomes 2:822
- double membrane vesicles (DMVs) 2:820–822
- double negative cell (DN cell) 3:552–553
- double positive cell (DP cell) 3:552–553
- doublecortin-like kinase 1 (Dclk1) 3:810–811
- double-edged sword in cancer 3:63
- double-headed unconventional myosin-5a 2:601
- double-negative cells (DNCs) 3:616
- double-strand break repair (DSBR) 1:451, 3:465–466, 3:466
- double-strand breaks (DSBs) 1:451, 1:456, 3:550
- recognition 1:451
- resection 1:456–457
- double-stranded DNA (dsDNA) 1:427, 1:457–458, 3:458
- viruses 1:518–519, 1:519F
- with RNA intermediates 1:523
- double-stranded RNA (dsRNA) 1:354, 2:819, 3:718–719
- viruses 1:519–520
- doublet microtubules 2:649
- downstream core element (DCE) 1:481–482
- downstream of kinase-R (Dok-R) 4:108
- downstream promoter element (DPE) 1:481–482
- DP cell *see* double positive cell (DP cell)
- Dp *see* desmoplakin (Dp)
- DP1/Yop1p proteins 1:205
- DPAPA *see* dipeptidyl aminopeptidase A (DPAPA)
- DPE *see* downstream promoter element (DPE)
- DPI *see* data processing inequality (DPI)
- DR *see* death receptors (DR)
- DR3 signaling pathway 3:72, 3:771
- DR4 signaling pathway 3:70–71
- DR5 signaling pathway 3:70–71
- DR6 signaling pathway 3:73
- DRC *see* dynein regulatory complex (DRC)
- DREAM *see* Dialog on Reverse Engineering Assessment Methods (DREAM)
- DRFs *see* Diaphanous-related formins (DRFs)
- DRiPs *see* defective ribosomal products (DRiPs)
- driven diacylglycerol (DAG) 2:367
- driver genes 3:412–413
- DRM *see* detergent-resistant membranes (DRM)
- drop-based microfluidic devices 3:667–669
- Drosha 1:354, 1:355, 1:355–358, 1:356–357F, 1:357F, 1:358, 1:359, 1:360, 1:365, 1:366
- cleavage of pri-miRNAs into pre-miRNAs by 1:355
- Drosophila* 1:370, 1:423, 2:468T, 2:470, 2:471, 3:86, 3:87, 3:88, 3:94, 3:95, 3:199, 3:200–201, 3:201, 3:204, 3:370
- Drosophila* disc large tumor suppressor (Dlg1) 2:597
- Drosophila* inhibitor of apoptosis 1 (DIAP1) 3:370
- Drosophila melanogaster* 1:315T, 1:368, 2:468T, 2:573, 3:3, 3:99, 3:125, 3:272–273, 3:364–366, 3:364, 3:366, 3:786
- Hippo pathway in 3:99
- Drosophila melanogaster* inhibitor of apoptosis protein 1 (Diap1) 3:99
- Drosophila* species 4:155–156
- Drp1 2:284
- Drp1-dependent mitochondrial fission 2:289
- drug design 1:135–136
- antibody–drug conjugates 1:139
- and biologics 1:138–139
- evolution 1:136–137
- history of drug development 1:136
- HTS 1:137–138
- structural biology and drug development 1:137
- virtual screening 1:138
- drug targeting 4:264–268
- from molecules to systems 4:264
- data integration 4:266
- genomics in target and drug discovery 4:264–265
- metabolomics 4:265–266
- PCSK9 4:265
- proteomics 4:265–266
- network medicine 4:266–267
- new methods for 4:267
- drug(s) 4:264
- delivery 2:762
- and drug targets 4:195–196
- drug-induced apoptosis 1:674
- substance 1:135
- Ds *see* Dachous (Ds)
- DS *see* dermatan sulfate (DS)
- DSBR *see* double-strand break repair (DSBR)
- DSBs *see* double-strand breaks (DSBs)
- dsDNA *see* double-stranded DNA (dsDNA)
- dsh *see* Dishevelled (dsh)
- DSL domain *see* delta–Serrate LAG-2 domain (DSL domain)
- Dsl1 complex 2:487
- dSMAC *see* distal SMAC (dSMAC)
- dsRBD *see* dsRNA-binding domain (dsRBD)
- dsRNA *see* double-stranded RNA (dsRNA)
- dsRNA-binding domain (dsRBD) 1:354, 1:355, 1:355–358, 1:357F, 1:358, 1:360
- DSS *see* dextran sulfate sodium (DSS)
- DT *see* diphtheria toxin (DT)
- DTR *see* diphtheria toxin receptor (DTR)
- DTT *see* dithiothreitol (DTT)
- D-type cyclin 3:439–440
- dual coding INK4a/ARF locus in mammals 3:448–449
- dual leucinezipper kinase (DLK) 3:237–238
- dual specificity phosphatases (DUSPs) 3:13, 3:125
- DUBs *see* deubiquitylating enzymes (DUBs)
- duodenal epithelium 1:734
- duodenase 1:734
- durotaxis 3:209
- DUSPs *see* dual specificity phosphatases (DUSPs)
- duty ratio 2:597, 2:601
- DV axis *see* dorsal–ventral axis (DV axis)
- DVT *see* deep vein thrombosis (DVT)
- dye–ligand chromatography 1:72
- dynactin 2:624
- dynamic density functional theory (DDFT) 4:83
- dynamic instability 2:539, 2:541, 4:36, 4:36–38, 4:39F, 4:40F
- dynamic integration 4:2
- dynamic nucleolus 3:259–260
- during cell division cycle 3:263
- nucleolar signaling networks 3:263
- in transcriptional inhibition and DNA damage 3:260–263
- dynamic proteomes measurement 4:171–172
- dynamic subunit exchange 2:572
- dynamics, metabolism 4:26
- dynammin 2:446, 2:503, 2:761
- dynammin 1-like-mediated mechanism (DN1L-mediated mechanism) 3:344

dynamatin-dependent endocytosis 2:797
 dynamatin-like proteins (DLP) 2:259–261
 DYNC *see* cytoplasmic dynein (DYNCL)
 Dync1h1 *see* dynein heavy chain 1 (Dync1h1)
 DYNC2H1 *see* cytoplasmic dynein2 heavy chain (DYNC2H1)
 dynein heavy chain (DHC) 2:620, 3:234
 cytoplasmic 2:620–621, 2:621F
 dynein heavy chain 1 (Dync1h1) 3:238
 genes diversity 2:627–628, 2:629F
 dynein intermediate chain (DIC) 3:234
 dynein regulatory complex (DRC) 2:627, 2:663
 dyneins 2:26, 2:27–28, 2:468, 2:660, 2:620–636, 2:620, 2:620, 3:234, 4:36
 see also axonemal dyneins
 architecture 2:620
 DHCs 2:620
 connection between dyneins and ciliary components 2:631F
 crystal structure of cytoplasmic dynein 2:622F
 cytoplasmic dynein mechanochemical cycle and power stroke 2:623F
 directionality mechanism 2:624
 motility
 AAA + domains 2:622–623
 dynein directionality mechanism 2:624
 force generation 2:623–624
 regulation 2:624
 stepping 2:621–622
 regulatory components 2:630
 central pair 2:630–631
 ICs 2:630
 LCs 2:630
 RS 2:630–631
 dysbiosis in disease states 2:804F
 dyskeratosis congenita (DC) 1:426–427
 dysplasia 1:796
 dystroglycanopathies *see* congenital muscular dystrophies (CMDs)
 DZ *see* dark zone (DZ)

E

E1 ubiquitin-activating enzyme, deficiency of 1:765
 E3 ubiquitin ligase 2:427
 E3 ligases in p53 pathway 3:505
 DUBs enzymes 3:505–506
 monoubiquitination vs.
 polyubiquitination 3:505
 e4EBPs *see* eukaryotic translation initiation factor 4E binding proteins (e4EBPs)
 E site 3:307
 EAE *see* Experimental autoimmune encephalomyelitis (EAE)
 early endosomal antigen 1 (EEA1) 2:788–789
 early endosomes (EEs) 1:612, 2:192, 2:408, 2:408, 2:411F, 2:814, 4:56
 compartments 2:193, 2:425

endosome heterogeneity reflects
 functional compartmentalization 2:192
 endosome traffic regulation 2:194–195
 coordination of phosphoinositides and Rab proteins 2:195–196
 phosphoinositides 2:195
 Rab proteins 2:195
 epithelial cells 2:192–193
 macroscopic properties 2:192
 maturation and conversion to late endosomes 2:196
 recycling from early endosomes to plasma membrane 2:197–198
 retrieval of proteins from endosomes to trans Golgi network 2:196–197
 (ultra)structure and specialization 2:193–194
 early sorting endosome (ESE) 2:401
 early/recycling endosome network 2:201
 EB *see* elementary body (EB)
 EBs *see* elementary bodies (EBs)
 EB1 *see* end-binding protein-1 (EB1)
 EBFP *see* equivalent back focal plane (EBFP)
 Ebola virus 1:520
 4EBPs *see* EIF4E binding-proteins (4EBPs)
 EBs *see* end-binding proteins (EBs)
 EBV *see* Epstein–Barr virus (EBV)
 EC *see* elongating complex (EC);
 endothelial cell (EC)
 E-cadherin 2:704–705, 2:745, 3:188, 3:188–189
 eccrine sweat gland 1:258F
 ECD *see* electron capture dissociation (ECD); extracellular domain (ECD)
 ECE-1 *see* endothelin-converting enzyme-1 (ECE-1)
 ECF *see* extracytoplasmic factor (ECF)
 ECM contacts *see* fibrillar adhesions
 ECM *see* extracellular matrix (ECM)
 eCSCs *see* endogenous cardiac stem cells (eCSCs)
 ECSIT *see* evolutionarily conserved signaling intermediate in toll pathways (ECSIT)
 Ect2 *see* epithelial cell transforming sequence 2 protein (Ect2)
 ectodysplasin (EDA) 3:62
 ectodysplasin-A receptor (EDAR) 3:62, 3:67, 3:72
 ectosomes 2:304
 EDA *see* ectodysplasin (EDA)
 EDAR *see* ectodysplasin-A receptor (EDAR)
 EDAR-associated death domain-containing adaptor protein (EDARADD) 3:67
 EDEM *see* ERdegradation-enhancing α -mannosidase-like protein (EDEM)
 EDEM3 *see* ER degradation enhancing mannosidase-like protein 3 (EDEM3)
 edema factor (EF) 2:795
 edge flux (EF) values 4:213
 Edman degradation 1:77
 EDTA *see* ethylenediaminetetra-acetate (EDTA)
 EEA1 *see* early endosomal antigen 1 (EEA1)

eEF1 *see* eukaryotic elongation factor 1 (eEF1)
 eEF1A 1:527, 1:527, 1:531
 eEF1B 1:527, 1:530F, 1:531
 eEF2 1:309, 1:323–324, 1:527, 1:531, 1:532, 1:533
 eEF3 (elongation factor) 1:527, 1:531
 EEs *see* early endosomes (EEs)
 EF *see* E-face (EF); edema factor (EF)
 EF values *see* edge flux (EF) values
 EF-2 *see* elongation factor 2 (EF-2)
 EF2K *see* elongation factor 2 kinase (EF2K)
 EFA *see* equation-free approach (EFA)
 E-face (EF) 2:39
 effector/memory T cells reactivation 3:640–641
 effector proteins
 Chk1 effector proteins 3:461–462
 Chk2 effector proteins 3:461–462
 p53 effector proteins 3:462
 effectors of Rab proteins 2:445
 effector-triggered immunity (ETI) 2:846
 efferocytosis 2:751, 2:753–754, 3:4, 3:374, 3:374–378
 in breast 3:374–376
 cell death 3:374
 postpartum breast cancer 3:376
 in tumor microenvironment 3:376
 EF-G *see* elongation factor G (EF-G)
 EF-hand 3:162F
 calcium-binding principle 3:161–162
 EFTEM 2:24
 EF-Tu *see* elongation factors Tu (EF-Tu)
 EG *see* embryonic germ cells (EG)
 Egalitarian (Egl) 2:625
 EGF *see* epidermal growth factor (EGF)
 EGF receptor (EGFR) 3:799
 eGFP *see* enhanced green fluorescent protein (eGFP)
 EGFR *see* epidermal growth factor receptor (EGFR)
 Egl *see* Egalitarian (Egl)
 EHD *see* eps15 homology domain (EHD)
 EI *see* electron ionization (EI)
 EI *see* enzyme I (EI)
 eicosanoids 1:153–154, 1:195
 eIF *see* eukaryotic initiation factor (eIF)
 eIF1 1:527, 1:527, 1:529
 eIF1A 1:527, 1:527, 1:528F, 1:529
 eIF2 1:309, 1:527, 1:527, 1:528F, 1:529, 1:532, 1:532–533, 1:533, 2:530–531, 2:531
 eIF2A 1:527
 eIF2 α 2:531
 eIF2B 1:309, 1:325, 1:325F, 1:527, 1:528F, 1:529, 1:532
 eIF2D 1:527, 1:532
 eIF3 1:527, 1:527, 1:527–529, 1:528F, 1:529, 1:530F, 1:531, 1:589–590, 3:245
 eIF4A 1:527, 1:527–529, 1:528F, 1:529, 3:323–324
 eIF4B 1:527, 1:527–529, 1:528F, 1:529, 3:131
 eIF4E binding-proteins (4EBPs) 3:275, 3:307
 4EBP1 3:245

- eIF4F 1:347, 1:527–529, 1:527, 1:528F, 1:529, 1:531, 1:532–533, 1:532, 1:533, 3:3–4, 3:275
- eIF4G *see* eukaryotic initiation factor 4G (eIF4G)
- eIF4H 1:527, 1:528F
- eIF5 1:527, 1:528F, 1:529
- eIF5A 1:527, 1:531
- eIF5B 1:387F, 1:527, 1:528F, 1:529, 3:313
- EIPA *see* ethyl isopropyl amiloride (EIPA)
- EJC *see* exon junction complex (EJC)
- elaioplast 2:324–325
- elastase 1:794
- elastases 2A 1:734
- elastases 2B 1:734
- elastases 3B 1:734
- elastase-antielastase imbalance 1:794
- elastic fibers 2:700
- elasticity 4:26F, 4:27
- coefficients 4:26–27, 4:27
- feedback 4:32, 4:32–33
- ELC *see* essential light chain (ELC)
- electrical conductivity 4:74
- electrical properties of cell 1:220–221
- electrochemical arrays 1:47–49
- electrochemical DNA microarray 1:47–49, 1:49F
- electrochemical driving force 1:221, 1:223–224, 1:225
- for ion movement 1:221–223, 1:235
- electrochemical gradient 1:218, 1:225, 1:226, 1:227T, 1:231, 1:235, 1:244, 1:244–245, 1:245, 1:246–247, 1:247, 1:248, 1:248–249, 1:249, 1:303
- electron capture dissociation (ECD) 1:79–80
- electron ionization (EI) 4:203–204, 4:204F, 4:205–206, 4:207
- electron microscopy (EM) 1:665, 2:14, 2:91, 2:319, 3:458–459
- electron multiplying charge coupled device cameras (EMCCD cameras) 2:122
- electron paramagnetic resonance (EPR) 1:100
- electron tomography 2:1, 2:22, 2:28F
- examples 2:27–28
- methods of Cryo-ET 2:24
- microscopy 2:24–25
- specimen preparation 2:24
- tomogram reconstruction and 3D image analysis 2:25–26
- specimen preparation methods 2:22–24
- techniques to combine 2:26–27
- electron transport 1:298, 1:301–302, 1:305
- electron-transfer dissociation (ETD) 1:79–80, 1:572, 4:180
- electrophoresis 1:66
- electrospray ion source (ESI) 4:204–205, 4:204F, 4:207
- electrospray ionization (ESI) 1:79
- elementary body (EB) 2:780, 2:786–787
- elementary processes 2:186
- eleven nineteen lysine-rich leukemia (ELL) 1:483
- eliglustat 1:780
- ELISA *see* enzyme-linked immunosorbent assays (ELISA)
- ELL *see* eleven nineteen lysine-rich leukemia (ELL)
- elongating complex (EC) 1:472
- elongation factor 2 (EF-2) 2:797, 3:245
- elongation factor 2 kinase (EF2K) 3:166
- elongation factor G (EF-G) 1:309, 1:323, 1:329, 1:384, 1:385F, 1:389F, 1:392, 1:392–393
- interactions with ribosomal RNAs 1:390
- elongation factors Tu (EF-Tu) 2:447
- EM *see* electron microscopy (EM)
- EMAP *see* erythrocyte membrane acid proteinase (EMAP)
- embryogenesis 4:114–115
- hyaluronan in 1:282
- embryonal rhabdomyosarcoma 1:360
- embryonic development 2:727–728, 3:187–188
- embryonic germ cells (EG) 3:819
- embryonic stem cell (ESC) 3:6–7, 3:102, 3:440, 3:785, 3:819
- effectors and mechanisms of epigenetic regulation of gene expression in 3:785–786
- chromatin remodeling 3:786–787
- DNA methylation 3:787–788
- enhancer activation and developmental competence 3:785–786
- noncoding RNA (ncRNA) 3:788
- promoter bivalency and maintenance of pluripotency 3:786
- epigenetic signatures and control of nuclear architecture in 3:788–789
- EMCCD cameras *see* electron multiplying charge coupled device cameras (EMCCD cameras)
- EMCV *see* encephalomyocarditis virus (EMCV)
- e-MI *see* endosomal microautophagy (e-MI)
- emphysema 1:794
- EMT *see* epithelial–mesenchymal transition (EMT)
- Ena/VASPs 2:685–686, 2:685F, 2:687
- enabled/VASP homology 1 (EVH1) 1:92
- ENaC *see* epithelial sodium channels (ENaC)
- encapsulated SCR7 (ESCR7) 1:453–454
- encephalomyocarditis virus (EMCV) 3:307
- end replication problem 1:426, 1:430
- end-binding protein-1 (EB1) 3:481
- end-binding proteins (EBs) 2:544
- Endo H *see* endoglycosidase H (Endo H)
- endocannabinoid 1:196–197, 1:199
- endocyclic modification 1:27
- endocytic maturation blockade 2:777
- endocytic organelle movement 2:415
- endocytic pathway 2:2–3, 2:434–435, 2:435, 2:436, 2:438, 3:1
- endocytic pathway manipulation 2:777
- endocytic maturation blockade 2:777
- pathogens interaction with lysosomal components 2:777–778
- survival after entry into endocytic pathway 2:777
- endocytic recycling compartment (ERC) 2:175–176, 2:505
- endocytic vesicle 1:202, 1:206F, 2:408, 2:434–435, 3:87F
- endocytic-sorting signals 2:421–422
- endocytosis 1:584, 1:617–618, 1:763, 1:764, 1:769, 2:3, 2:192–193, 2:211, 2:217, 2:377F, 2:380, 2:394, 2:446, 2:471, 2:507–510, 2:758, 2:811–814, 3:2, 3:13–14, 3:51, 3:92, 3:93F, 3:203, 3:497–498, 3:588, 3:589, 3:608F, 3:728–729
- endocytosis II non-clathrin mechanisms
- caveolae 2:396
- clathrin-and caveolin-independent micropinocytic pathways 2:397
- macropinocytosis 2:394
- ENDOG *see* endonuclease G (ENDOG)
- endogenous cardiac stem cells (eCSCs) 3:801
- endogenous plastic somatic cells (ePS cells) 3:819
- endoglycosidase H (Endo H) 2:341
- endolysosomal pathways 1:763–764, 2:469
- endo-lysosome 2:203
- endomembrane system 2:159, 2:163–164, 2:168
- endonuclease 3:289, 3:321
- endonuclease G (ENDOG) 3:342
- endonuclease VIII-like glycosylases (NEIL glycosylases) 1:443, 1:445–446
- endopeptidase 1:326, 1:568, 1:569–570, 1:734, 1:748
- endophilin B1 2:279, 2:286
- endoplasmic reticulum (ER) 1:3, 1:59, 1:78, 1:85–86, 1:174–175, 1:195, 1:202, 1:204F, 1:535–536, 1:536F, 1:537T, 1:539–540T, 1:557, 1:557–558, 1:558–559, 1:558F, 1:559T, 1:560–561, 1:561F, 1:563, 1:591, 1:596, 1:614–615, 1:764–765, 1:765, 1:774, 1:775, 1:782–783, 2:2–3, 2:156–167, 2:168, 2:183, 2:206, 2:255–256, 2:273–274, 2:279–281, 2:311, 2:331, 2:339–340, 2:354, 2:363, 2:371, 2:376–377, 2:410, 2:461–462, 2:505, 2:528–538, 2:624, 2:789–790, 2:795, 2:819–820, 2:832, 2:850, 3:87, 3:87F, 3:88–89, 3:172, 3:341–342, 3:379, 3:404, 3:497, 3:505, 3:573–574, 3:608F, 3:720, 3:734, 3:751
- ATF6 2:531
- Bcl-2 proteins at 3:384
- chronic UPR activation effect 2:531–532
- client proteins
- posttranslational modifications 2:161
- translocation and membrane embedding of 2:160–161
- ER-related folding diseases 1:535, 1:543–544
- export from and retrieval to 2:163–164, 2:163F
- folding and quality control in 1:535–544
- glycoproteins processing and folding in 2:341

- endoplasmic reticulum (ER) (*continued*)
 homeostasis 2:164–165
 intracellular signaling pathways 2:530F
 IRE1 2:530
 lumen 2:795
 malfunctioning of protein folding in 1:543–544
 membrane contact sites 2:159–160
 metabolic syndrome activating UPR signaling 2:533–534
 morphology 2:157–159
 origin and function 2:157
 PERK-mediated phosphorylation of α -subunit of eIF2 2:530–531
 protein folding 2:161–163
 protein synthesis, folding and quality, control of 2:528–529
 stress 1:239–240, 2:276
 in disease 2:532F
 stress response 2:3
 transmembrane proteins sensing misfolded proteins 2:529–530
 UPR
 activation associated with neurodegenerative diseases 2:535–536
 discovery 2:529
 pathway roles 2:531
 protective role in infectious and inflammatory disease 2:534
 role in human health and disease 2:532–533
 signaling contribution to pancreatic β -cell degeneration and diabetes 2:533
 signaling providing double-edged sword in cancer 2:534–535
 endoplasmic reticulum exit site (ERES) 1:601, 2:168, 2:332–333, 2:373
 endoplasmic reticulum stress element (ERSE) 2:531
 endoplasmic reticulum-associated degradation (ERAD) 1:596, 1:794, 1:794–795, 2:162–163, 2:341, 2:371, 2:529, 2:796–797, 3:505
 endoribonuclease (RNase) 2:529
 endosomal compartment, phosphoinositides in 2:510–512
 endosomal dynamics 2:425
 LDL receptor selection 2:426–427
 LDL uncoupling 2:425–426
 endosomal microautophagy (e-MI) 3:406
 endosomal network theory
 entry-fusion-exit model 4:58
 comparison of theory predictions to LDL trafficking 4:59
 quantitative comparison 4:59
 theoretical predictions 4:58–59
 general model
 cargo distribution in endosomes population 4:56
 governing equation 4:57–58
 microscopic processes and parameters 4:56–57
 principle processes of endosome interactions 4:58F
 quantification of endosome distributions *in vivo* and *in vitro* 4:57F
 endosomal protein sorting 2:475, 2:478
 endosomal recycling compartment (ERC) 2:401
 endosomal recycling pathway 2:401–407, 2:402F
 intracellular pathogens subvert function of Rab GTPases 2:405–406
 key regulators 2:402
 Rab4 recycling pathway 2:402–403
 Rab8 and Rab22A 2:404–405
 Rab11 and endosomal recycling compartment 2:403
 Rab11 and cancer 2:404
 Rab11 binds motor proteins 2:403
 Rab11 effector proteins 2:403
 Rab11 in cell division 2:403–404
 Rab11 role in synaptic transmission 2:404
 Rab14 2:404–405
 Rab17 2:404–405
 short-loop recycling pathway 2:402–403
 endosomal sorting complexes required for transport (ESCRT) 2:195, 2:202–203, 2:228, 2:304–305, 2:409–410, 2:436, 2:466, 2:819–820, 2:832–833, 3:14, 3:406, 3:475, 3:484–486, 3:499, 3:589–590, 3:591
 around the cell 2:466–474, 2:467F
 in cytokinesis 2:470
 in endolysosomal system 2:469
 ubiquitination 2:469
 ESCRT0-III 2:503, 2:511
 ESCRT-I 2:466, 2:468, 2:468T, 2:469, 2:469–470, 2:470, 2:470–471
 ESCRT-II 2:466, 2:467–468, 2:468, 2:468T, 2:469
 ESCRT-III 2:219
 in extracellular vesicle biogenesis 2:470–471
 machinery 2:466
 early acting 2:466
 ESCRT-III and Vps4 2:466–468, 2:469F
 and other cellular machinery 2:468
 models of 2:468–469
 non-endosomal roles for 2:471
 activity at the nuclear envelope 2:471–472
 neurite severing in neuronal development 2:471
 organization and membrane dynamics in cilia 2:472
 pH sensing 2:471
 plasma membrane resealing 2:471
 in virus budding 2:469–470
 endosomal trafficking 2:486–487
 endosomal-lysosomal pathway 2:3
 endosome to lysosome transport 2:408–417
 cargo sorting 2:408–410
 cytoskeletal element role 2:413–415, 2:413–414T
 EEs 2:408, 2:411F
 endocytic vesicles 2:408
 endosome maturation 2:409F, 2:410–411
 pH Role ion concentrations in 2:411–412
 fusion events 2:412
 LEs 2:410–411, 2:411F
 perinuclear clustering 2:412
 molecular motors 2:413–415
 MVBs biogenesis 2:408–410, 2:409F
 Rab protein role 2:412–413
 transport on microtubules 2:414F
 endosome(s) 1:202, 1:203, 1:204F, 1:206, 1:562–563, 1:763–764, 1:764, 1:765, 2:2–3, 2:192, 2:217–219, 2:434–435, 2:443, 2:475, 2:477, 2:479, 2:479–480, 3:590
 early-endosomes-to TGN trafficking 2:436
 cargo sorting and budding 2:436–437
 docking and fusion 2:437
 heterogeneity reflecting functional compartmentalization 2:192
 incoming alphaviruses in 2:814F
 late-endosomes-to TGN trafficking 2:435
 cargo sorting and budding 2:435
 docking and fusion 2:435–436
 roles in clathrin 2:389–390
 traffic regulation 2:194–195
 coordination of phosphoinositides and Rab proteins 2:195–196
 phosphoinositides 2:195
 Rab proteins 2:195
 endosomes signaling 2:211, 2:217
 accessibility to phosphatases 2:219–220
 cell types 2:211–212
 confined size 2:219
 endocytic routes 2:212–217
 endocytosis 2:211
 and implications in cancer 2:221–222
 low pH 2:219
 maturation and unique signaling adapters 2:219
 pathways for receptors 2:213–216T
 PIPs 2:217–219
 signaling regulates endocytosis 2:220
 actin in regulation 2:221
 transcription and phosphorylation 2:220–221
 spatial signaling and signal amplification 2:220
 endothelia 2:704
 endothelial cell (EC) 1:715–716, 1:717, 3:210, 3:632, 3:636, 3:639F, 4:102
 developmental lineage 4:103F
 semi-professional APCs 3:640F
 endothelial function in antigen-specific immunity 3:639
 antigen recognition and responses 3:642
 EC Ag presentation emerging and putative roles 3:640–641
 endothelial antigen presentation in pathology 3:641–642
 endothelial cells antigen presenting properties 3:639–640
 model for ILPs function in Ag recognition and response 3:643F
 endothelial nitric oxide synthase (eNOS) 1:720, 2:497, 3:57, 3:133, 4:111

- endothelial protein C receptor (EPCR)
1:717
- endothelial remodeling 3:637
- endothelin peptide family 1:750
- endothelin-1 (ET-1) 1:750
- endothelin-converting enzyme-1 (ECE-1)
1:750
- endothelium as a barrier
blood-lymph circulatory system
3:633F
features 3:634
lymphatic endothelium 3:634
vascular endothelium 3:633–634
- endotoxin *see* lipopolysaccharide (LPS)
- energy
charge 1:288, 1:293–294
homeostasis 1:289, 1:294, 1:295
landscape 1:110
mTOR signaling regulation by 3:247
- enhanced green fluorescent protein (eGFP)
4:21
- enhanceosome 3:258
- enhancer RNAs (eRNAs) 1:369, 1:491
- enhancer-promoter communication (EPC)
1:488, 1:489F, 1:492F
- enhancers 1:482
- enolase 1:14, 1:16F, 1:18, 1:19, 1:488
- eNOS *see* endothelial nitric oxide synthase (eNOS)
- entactin *see* nidogens
- enteropathogenic *Escherichia coli* (EPEC)
2:791
- enteropeptidase 1:734
- enterotoxigenic *Escherichia coli* (ETEC)
3:58–59
- ENTH domain *see* epsin N-terminal
homology domain (ENTH domain)
- enthalpy 1:5–6, 1:109
- entropy 1:5–6, 1:109
- envelope glycoprotein (Env) 2:828
Env-mediated membrane fusion 2:828
trafficking 2:832
- enveloped viruses 2:809, 2:810
- environmental SEM 2:18
- enzymatic domains 1:93
- enzyme I (EI) 1:101
- enzyme kinetics 1:8–9
- enzyme-catalyzed reactions 1:8
- enzyme-linked immunosorbent assays
(ELISA) 1:141, 1:686–687
- enzymes 1:14, 1:14, 1:125, 1:539–540T,
1:564T, 3:785
see also specific enzymes
chemical mechanisms 1:18
coenzymes and cofactors 1:19
composition of 1:14
cooperativity and allosteric regulation
1:17–18
in food industry 1:125
kinetics 1:14–17
ping pong mechanisms 1:17
posttranslational modifications of 1:14
regulation of 1:289
in glucose metabolism 1:292T
and posttranslational modification
1:290
- response 4:24, 4:26–27, 4:27, 4:27–28,
4:28, 4:29, 4:32–33
elasticity coefficients 4:27
operational definition 4:27
sequential mechanisms 1:15–16
substrate selectivity 1:17
transition state analogs and catalytic
antibodies 1:18–19
- eosinophilic secretory lysosomes 2:226–227
- EPC *see* enhancer-promoter communication
(EPC)
- EPCR *see* endothelial protein C receptor
(EPCR)
- EPI cells *see* epiblast cells (EPI cells)
- epiblast cells (EPI cells) 2:743–744
- epicardium 3:802
- epidermal growth factor (EGF) 1:691, 1:702,
2:140–141, 2:419–420, 2:445–446,
3:8–9, 3:107, 3:224, 3:413,
3:432–433, 3:505, 3:529, 4:44–45,
4:46, 4:107, 4:218–219
signaling 4:172
- epidermal growth factor receptor (EGFR)
1:639, 1:694, 1:754, 2:211, 2:408,
2:452, 2:503, 3:51, 3:99, 3:337,
3:437, 4:82, 4:267
EGFR-ligands 1:641, 1:643
pathway 4:256, 4:256F
signaling pathway 4:172
- epidermis 3:216, 3:216–218, 3:220
- epigenetic alterations 3:777–778
- epigenetic regulation of gene expression
3:785
- epigenome 3:6–7, 3:785, 3:785–786, 3:789
- epiphygan 1:273–274
- episomal expression systems 1:62–64
- epithelia 2:704
- epithelial barriers 2:810
- epithelial cell transforming sequence 2
protein (Ect2) 2:710
- epithelial cells 2:192–193, 2:742
see also planar cell polarity (PCP)
apical–basal polarity 2:742
proteins interplay organizing 2:745
bacterial adhesion 3:692F
cytoskeletal rearrangements 2:746
defense 3:690–691
epithelial layer structure 3:688–689
initiating cues of polarization 2:743–745
loss of barrier function in disease
3:691–692
mediators
cell polarity establishment 2:743F
intercellular adhesion complexes 2:742
junctional complexes 2:743F
polarity complexes 2:742–743
perspective 3:692–693
polarized trafficking machinery 2:745–746
phosphoinositides and 2:746
transcytosis and lumen formation
2:744F, 2:745–746
types 3:689–690
- epithelial protein lost in neoplasm (EPLIN)
2:705
- epithelial sodium channels (ENaC)
1:651–652, 1:735–736
- epithelial–mesenchymal transition (EMT)
2:741, 2:747, 3:102, 3:188, 3:219,
3:223–224, 3:223, 3:335, 3:440,
3:809
EMT-associated cellular and molecular
changes 3:224–225
EMT-inducing transcription factors 3:224
triggers 3:224
- epitheliasin *see* TMPRSS2
- epithelium 3:687
- epitope 1:135, 1:138
- EPL *see* external plexiform layer (EPL)
- EPLIN *see* epithelial protein lost in
neoplasm (EPLIN)
- EPO *see* erythropoietin (EPO)
- EPR *see* electron paramagnetic resonance
(EPR)
- Eps15 family 2:455
- eps15 homology domain (EHD) 1:92,
2:197, 2:454
EHD1 2:476–477
- ePS cells *see* endogenous plastic somatic
cells (ePS cells)
- Epsin family 2:454
- epsin N-terminal homology domain (ENTH
domain) 2:428, 2:454
- EpsinR family 2:454–455
- EPSP *see* excitatory postsynaptic potentials
(EPSP)
- Epstein–Barr virus (EBV) 3:569, 3:585,
3:773
- Epstein–Barr virus-induced gene (EBi2 gene)
see G-protein coupled receptor
(GPR183)
- equation-free approach (EFA) 4:80–81
- equilibrium 1:5
- equilibrium potential 1:223, 1:223–224,
1:223T, 1:224F, 1:225
- equivalent back focal plane (EBFP) 2:65
- ER degradation enhancing mannosidase-like
protein 3 (EDM3) 2:341
- ER lumen 1:536, 1:536F, 1:538
folding and assembly of IgG in 1:542
folding factors in 1:542
folding of glycoproteins in 1:540–542
nascent polypeptides forming their native
structures in 1:538–540
protein folding and assembly in
1:538–540
secretory proteins forming their disulfide
bonds in 1:540
- ER protein quality control (ERQC) 2:529
- ER protein-folding homeostasis 2:529
- ER *see* endoplasmic reticulum (ER)
- ER/Golgi-dependent pathway 2:520
- ER71 *see* ETS-related 71 (ER71)
- ERAD *see* endoplasmic reticulum-associated
degradation (ERAD)
- ERADication of unfolded protein 1:602
proteasome 1:605–606
retro-translocation of misfolded substrates
1:604–605
substrate selection 1:602
fatty and sweet sides of destruction
1:603–604
location, location, location 1:602–603

- ERADication of unfolded protein (*continued*)
 molecular chaperones 1:604
 yeast Doa10, Hrd1, and Cdc48
 complexes, and 26S proteasome
 1:603F
 ubiquitin mark 1:605
 'erasers' 3:785
 ER-associated degradation (ERAD) 1:535,
 1:541F, 1:542–543, 1:585,
 3:590–591
 ERC *see* endocytic recycling compartment
 (ERC); endosomal recycling
 compartment (ERC)
 ERdegradation-enhancing α -mannosidase-
 like protein (EDeM) 1:601
 ER-derived vesicles 2:777–778
 ERES *see* endoplasmic reticulum exit site
 (ERES)
 eRF1 1:324, 1:527, 1:531, 1:532
 eRF3 1:324, 1:348, 1:527, 1:531
 ERGIC *see* ER-Golgi intermediate
 compartment (ERGIC)
 ER-Golgi interface 2:373
 ER-Golgi Intermediate Compartment
 (ERGIC) 1:601, 2:335–336
 ERGIC-53 2:169
 ER-Golgi trafficking 2:168, 2:170–171,
 2:176–177
 ER-Golgi transport 2:331–338
 cargo capture 2:331–332
 COPII biogenesis regulation 2:332–333
 cellular regulation 2:333–334
 COPII coat 2:331, 2:332F
 packaging of noncanonical cargo 2:334F
 vesicle
 delivery directionality 2:335F
 fusion 2:336
 size regulation 2:334–335
 tethering 2:335–336
 ERK *see* extracellular signal-regulated kinase
 (ERK)
 ERK1 *see* extracellular signal-regulated
 kinase 1 (ERK1)
 ERK1/2 *see* extracellular signal-regulated
 kinase isoforms 1/2 (ERK1/2)
 ERK2 *see* extracellular signal-related kinase 2
 (ERK2)
 ERM family *see* ezrin–radixin–moesin
 family (ERM family)
 eRNAs *see* enhancer RNAs (eRNAs)
 ERphagy 2:165
 ERQC *see* ER protein quality control
 (ERQC)
 error correction 2:644, 3:472–473
 ERSE *see* endoplasmic reticulum stress
 element (ERSE)
 erythrocyte membrane acid proteinase
 (EMAP) 1:681–682
 erythropoietin (EPO) 1:126, 3:23–24,
 3:136
 ES *see* E-surface (ES)
 ES cells *see* embryonic stem cells (ES cells)
 ES27 *see* expansion segment 27 (ES27)
 ESC *see* embryonic stem cell (ESC)
 ESCC *see* esophageal squamous cell
 carcinoma (ESCC)
Escherichia coli (*E. coli*) 1:56, 1:313, 1:315T,
 1:574, 2:434, 2:741, 4:76,
 4:199–200
 bacterial clamps in 1:400, 1:401F
 biosynthetic pathway for phospholipids in
 1:163F
 C-family polymerase of 1:400
 foreign lipids synthesis in 1:164F
 DnaB helicase 1:404
 Pol III 1:398
 polymerase release from Okazaki fragment
 1:410
 replisome 1:408
 Shine–Dalgarno sequence in 1:341–342
 ESCR7 *see* encapsulated SCR7 (ESCR7)
 ESCRT *see* endosomal sorting complexes
 required for transport (ESCRT)
 ESCRT-0 *see* STAM–Hrs complex
 ESCRT-based cell division 2:590F, 2:592
 ESE *see* early sorting endosome (ESE)
 E-selectin ligand-1 (ESL-1) 3:751
 ESI *see* electrospray ion source (ESI);
 electrospray ionization (ESI)
 ESL-1 *see* E-selectin ligand-1 (ESL-1)
 esophageal squamous cell carcinoma
 (ESCC) 3:158
 ESP-1 *see* testisin
 essential autophagy (Atg) genes 1:764–765
 essential cellular functions 4:72
 essential light chain (ELC) 2:597
 essential thrombocythemia (ET) 3:146
 estradiol 1:178–179
 ESTs *see* expressed sequence tags (ESTs)
 E-surface (ES) 2:39
 cells 1:512–513, 1:514
 E-Syt2 *see* extended-Synaptotagmin 2 (E-
 Syt2)
 ET *see* essential thrombocythemia (ET)
 ET-1 *see* endothelin-1 (ET-1)
 ETD *see* electron-transfer dissociation (ETD)
 ETEC *see* enterotoxigenic *Escherichia coli*
 (ETEC)
 ether protecting groups 1:42–43
 2-cyanoethyl group 1:43
 tert-butyldimethylsilyl group 1:42–43
 ethyl isopropyl amiloride (EIPA) 2:762
 ethylenediaminetetra-acetate (EDTA) 4:178
 ETI *see* effector-triggered immunity (ETI)
 etioplast 2:324
 ETS *see* externally transcribed spacer (ETS)
 Etv2 *see* ETS-related 71 (ER71)
 eubacteria 1:160, 1:160–161
 eukaryotes 1:160, 1:312–313, 1:315–316,
 1:320–321, 1:321–322, 1:322,
 1:323–324, 1:324, 1:325, 1:326,
 1:334, 1:375, 1:377–380, 1:423,
 1:488, 1:495, 1:495–496, 1:503,
 1:526, 1:526, 1:526–527, 1:527,
 1:529, 1:531, 1:532, 1:545, 1:552,
 1:577, 1:582, 1:583, 1:586–587,
 2:467–468, 2:800, 3:3
 phospholipids synthesis 1:164
 amine-containing phospholipids
 synthesis 1:164–166, 1:165F
 non-amine-containing phospholipids
 synthesis 1:166–169
 transcription enhancer action in
 1:491–492
 eukaryotic (B family) DNA polymerase
 1:398–400
 crystal structure of 1:400
 structure and organization of 1:399F
 eukaryotic cells 2:311, 2:371, 2:485, 2:491,
 2:720, 2:762, 4:177
 migration modes 2:720–723, 2:721F
 individual cell migration 2:720–723
 multicellular migration 2:723
 principle pathways of protein secretion
 from 2:521F
 eukaryotic cytoskeletal filaments 2:588
 eukaryotic elongation factor 1 (eEF1) 3:310
 eukaryotic initiation factor (eIF) 3:307
 eukaryotic initiation factor 4G (eIF4G)
 3:323
 eukaryotic mRNA 1:342
 AU-rich elements (AREs) 1:343–344
 cap 1:342
 degradation 1:348
 microRNA response elements 1:344
 poly(A) tail 1:344
 processing 1:342
 stem loops 1:343–344
 turnover 1:346
 3' UTR 1:343
 5' UTR 1:342–343
 eukaryotic nucleotide excision repair
 1:435–441
 damage recognition 1:436
 transcription-coupled repair 1:437
 UV-DNA damage binding (DDB)
 1:436
 XPC-HR23B-CETN2 1:436–437
 damage verification 1:437–438
 TFIIH 1:437–438
 XPA-RPA 1:438–439
 DNA repair mechanisms 1:435
 excision, repair synthesis, and ligation
 1:439
 DNA polymerases 1:439
 ligation 1:439
 human diseases 1:439
 Cockayne syndrome 1:439
 trichothiodystrophy (TTD) 1:439–440
 UV-sensitive syndrome (UVSS) 1:440
 xeroderma pigmentosum (XP) 1:440
 XFE (XPF-ERCC1) progeria 1:440
 incision complex, assembly of 1:439
 XPF-ERCC1 1:439
 XPG 1:439
 nucleotide excision repair (NER) 1:435
 general ways 1:435
 global genome nucleotide excision
 repair 1:435F
 initiation of 1:436F
 integrated steps 1:435
 UV-induced cyclobutane pyrimidine
 dimers (CPD) 1:435
 xeroderma pigmentosum
 complementation (XPC) 1:436
 eukaryotic primase 1:405–406
 eukaryotic protein secretion mechanisms
 2:520

- eukaryotic replisome 1:409F, 1:410–411, 1:412
- eukaryotic ribosome assembly 3:297, 3:297F
- cytoplasmic events in ribosome biogenesis 3:301–302
- pre-ribosomal particles nuclear export 3:299–301
- ribosomal pre-rRNA processing 3:297–299
- trans*-acting factors 3:299
- eukaryotic subtilase 1:555–556, 1:556
- eukaryotic transcriptional regulation 1:481
- experimental approaches
- computational analysis and systems biology 1:485
 - genetics 1:485
 - genome-wide analysis 1:485
 - molecular biological techniques 1:485
 - outstanding issues and questions 1:485–486
 - protein biochemistry and *in vitro* transcription 1:484–485
- promoters and transcription factors 1:481–482
- RNA polymerases 1:481
- transcription initiation and elongation 1:482–483
- chromatin 1:483–484
 - in vivo* dynamics and behavior 1:484
 - Pol II C-Terminal Domain 1:483
 - supercoiling 1:484
 - termination of transcription 1:483
- eukaryotic translation factors 1:527–529
- eukaryotic translation initiation factor 2 phosphorylation (eIF2 α) 2:531
- eukaryotic translation initiation factor 4E binding proteins (e4EBPs) 3:243
- evanescent wave 2:62
- EV *see* extracellular vesicles (EV)
- EV-based biomarkers 2:306–307
- EV-based vaccines and therapeutics 2:307–308
- EVH1 *see* enabled/VASP homology 1 (EVH1)
- EV-mediated intercellular communication 2:302, 2:303F
- evolutionarily conserved signaling intermediate in toll pathways (ECSIT) 3:721
- evolutionary genomics and cell biology
- cell division 4:157
 - comparative genomics approach 4:155F
 - gene regulatory networks 4:156–157
 - gene traffic 4:155–156
 - genomics for old and new questions 4:153–154
 - landmark genome projects 4:153T
 - new genes 4:158T
 - new genes for old and new functions 4:154–155
 - perspectives 4:157–158
 - protein traffic 4:157
- evolutionary process 4:154, 4:156
- EVs *see* extracellular vesicles (EVs)
- Ex protein *see* expanded protein (Ex protein)
- excision, repair synthesis, and ligation 1:435F, 1:439
- DNA polymerases 1:439
 - ligation 1:439
- excitable network, signal amplification through 4:8
- excitatory postsynaptic potentials (EPSP) 1:244, 1:246
- exemplary biological processes 4:172
- cell cycle regulation 4:172
 - central metabolism 4:172–173
 - EGFR signaling pathway 4:172
- exercise, in bacterial diversity 2:803
- ExoA 2:797
- exocyclic modification 1:27
- exocyst *see* exocytosis
- exocytosis 1:747, 2:63, 2:377F, 2:442, 2:485, 2:510
- exon 1:495, 1:495–496, 1:496F, 1:497, 1:498–499, 1:499, 1:499–500, 1:499F, 1:500, 1:500–501, 1:501, 1:503, 1:504F, 1:505, 1:505F, 1:506, 1:772, 3:5
- exon junction complex (EJC) 1:501, 3:319–320
- exonuclease 1:418, 1:418–419, 1:419F, 1:420, 1:503, 3:289, 3:291, 3:293
- 5′–3′ exonuclease 1:346
- 5′–3′ exonucleolytic digestion 1:350
- exopeptidase 1:568, 1:569–570
- exosite 1:725–726, 1:727
- exosomes 1:368, 2:304, 2:467F, 2:470, 2:738
- exotoxin A (ToxA) 2:796–797, 2:846
- exotoxins
- altering cytoskeleton 2:795–796
 - altering DNA 2:797
 - altering signal transduction 2:794–795
 - blocking protein synthesis or protein folding 2:797
 - interfering with intracellular trafficking 2:796–797
- expanded protein (Ex protein) 3:99–100
- expansion segment 27 (ES27) 3:300
- experimental autoimmune encephalomyelitis (EAE) 3:606
- exploitation of peroxisomes by viruses 2:255
- exportin-5 1:356–357F, 1:357F, 1:360, 1:365
- transportation of pre-miRNAs from nucleus to cytoplasm 1:355
- exportins 3:172, 3:172–173, 3:175
- expressed sequence tags (ESTs) 4:160
- expression systems 1:54–65
- bacterial expression systems 1:56
 - B. subtilis* 1:57
 - ColE1-family of replicons 1:56–57
 - gram-negative *Pseudomonas fluorescens* 1:57
 - inclusion bodies 1:57
 - IPTG 1:57
 - L. lactis* 1:57–58
 - pET-15b 1:56F
 - solubility-enhancing protein domains 1:57T
 - cell-free expression systems 1:58–59
- characteristics 1:55T
- insect expression systems 1:59
- lytic nature 1:59
 - recombinant baculoviruses generation 1:59
 - traditional transfection systems 1:59–60, 1:60
 - worldwide protein data 1:59
- mammalian expression systems 1:60
- aliphatic acid-type histone deacetylase inhibitors 1:62
 - episomal expression systems 1:62–64
 - life technologies 1:64
 - plasmid DNA 1:60
 - stable cell-line generation 1:61
- yeast expression systems 1:58
- heterologous protein in *S. cerevisiae* 1:58
 - hypermannosylation 1:58
 - P. pastoris* and *P. angusta* 1:58
 - protein expression in methylotrophic yeasts 1:58
 - unicellular microbial eukaryotes 1:58
- extended-Synaptotagmin 2 (E-Syt2) 2:461–462
- external plexiform layer (EPL) 4:135F
- externally transcribed spacer (ETS) 3:297–298
- ETS-related 71 (ER71) 4:103
- extracellular cues transmission 3:413
- extracellular domain (ECD) 3:30–31, 3:183
- extracellular matrix (ECM) 1:555, 1:562, 1:563–564, 1:570, 1:622–624, 1:624, 1:630–631, 1:695, 1:700, 1:753, 2:4, 2:347–348, 2:668, 2:694–703, 2:713, 2:720, 2:723–724, 2:742, 3:4, 3:219–220, 3:224, 3:233, 3:515, 3:619–620, 3:624, 3:671, 3:727, 3:789, 3:798–799, 3:808, 4:105, 4:142
- adhesion
- diversity 2:713–714
 - geometric and mechanical sensory functions 2:716–717
- basement membranes 2:694–695
- ECM proteins associated with 2:697–698
 - interaction model of principle components 2:696F
- collagen IV 2:696–697
- collagenous 2:699F
- connective tissue 2:698
- cross-talk with cell 2:717–718, 2:717F
- elastic fibers 2:700
- epithelial cell layer 2:694F
- FACIT collagens 2:700
- fibrillar collagens 2:698–700
- fibronectin 2:698
- focal adhesions 2:714F, 2:715F
- heparan sulfate proteoglycans 2:697
- integrin adhesome 2:714–716, 2:716F
- invadosomes 2:715F
- laminins 2:695–696
- large blood vessel 2:694, 2:695F
- matricellular proteins 2:700–701
- molecular composition 2:714F
- nidogens 2:697

extracellular matrix (ECM) (*continued*)
 response to mechanical signals 4:128
 role in cell behavior 4:123
 SLRPs 2:700
 SPARC-null mice 2:701F
 extracellular milieu *see* extracellular matrix (ECM)
 extracellular nutrients 4:74
 extracellular O-glycans 2:347–353
 mucin-type O-glycans 2:347–348, 2:348F
 O-Fuc glycans 2:350, 2:350F
 O-Gal glycans 2:351
 O-GalNAc glycans 2:347–348, 2:348F
 O-Glc glycans 2:350–351
 O-GlcNAc glycans 2:351
 O-Man glycans 2:348–350, 2:349F
 O-Xyl glycans 2:351–352
 proteoglycans 2:351–352
 types 2:348T
 extracellular pathogens 2:776
 extracellular signal-regulated kinase (ERK)
 2:718, 2:849, 3:130, 3:281–282,
 4:82, 4:105–106, 4:219
 extracellular signal-regulated kinase 1
 (ERK1) 3:123
 extracellular signal-regulated kinase
 isoforms 1/2 (ERK1/2) 2:211
 extracellular signal-related kinase 2 (ERK2)
 1:765–766, 1:771, 3:123
 extracellular signals 3:194
 extracellular vesicles (EVs) 2:302,
 3:522–524
 biogenesis 2:304–306
 ESCRT in 2:470–471
 clinical applications 2:307F
 EV-based biomarkers 2:306–307
 EV-based vaccines and therapeutics
 2:307–308
 in cross-kingdom communication
 2:303–304
 in disease 2:306
 EV-mediated intercellular communication
 2:302, 2:303F
 heterogeneity of mammalian cell-derived
 2:305F
 physiological role 2:302–303
 population 2:302
 extracytoplasmic factor (ECF) 1:469–471
 extrinsic apoptosis 3:343, 3:368–369, 3:370,
 3:371, 3:371–372
 extrinsic apoptotic pathway 3:69
 extrusion 2:707
Ex-vivo hematopoietic stem-cell gene
 (reinfusion) therapy 1:784
 Ezh2 3:786, 3:790
 ezrin–radixin–moesin family (ERM family)
 1:271–272, 2:688

F

F-1,6-bisphosphatase (F-1,6-bisPase) 1:294
 F₁F₀-ATPase 3:341
 FA pathway *see* Fanconi anemia pathway
 (FA pathway)

Fab fragment 1:18–19
 FAB *see* fast atom bombardment (FAB)
 FACE2 *see* farnesylated proteins-converting
 enzyme2 (FACE2)
 facilitated diffusion transporters 1:226,
 1:226–228, 1:227T
 facilitates chromatin transcription (FACT)
 1:483–484
 FACIT collagens *see* Fibril-associated
 Collagens with Interrupted Triple
 Helices collagens (FACIT collagens)
 FACS *see* fluorescence-activated cell sorting
 (FACS)
 FACT *see* facilitates chromatin transcription
 (FACT)
 F-actin *see* filamentous actin (F-actin)
 factor B, molecular mechanisms 1:728–729
 factor D, molecular mechanisms
 1:728–729, 1:729
 factor I, molecular mechanisms 1:729
 factor inhibiting HIF-1 (FIH-1) 4:109–110
 Factor V Leiden mutation 1:719
 FAD *see* familial AD (FAD); flavin adenine
 dinucleotide (FAD)
 FADD *see* FAS-associated death domain
 protein (FADD)
 FADD-like IL-1 β -converting enzyme (FLICE)
 3:335
 FAK *see* focal adhesion kinase (FAK)
 FAO *see* fatty acid oxidation (FAO)
 familial AD (FAD) 1:662
 familial CJD (fCJD) 2:840
 familial hemophagocytic
 lymphohistiocytosis (FHL) 2:226
 FHL1 2:237–239
 familial hypercholesterolemia 1:176, 1:178
 families of structurally similar proteins
 (FSSP) 1:94–95
 FANCD1/BRCA2 gene 3:777
 Fanconi anemia pathway (FA pathway)
 1:463
 FAP *see* fibroblast activation protein (FAP)
 FAPP *see* four phosphate adaptor proteins
 (FAPP)
 farnesoid X-receptor (FXR) 1:176–177
 farnesylated proteins-converting enzyme2
 (FACE2) 2:312–313
 farnesyltransferase (FTase) 2:312–313
 Fas 3:70
 Fas clade
 binding partners and downstream
 signaling pathways 3:71F
 DR4 3:70–71
 DR5 3:70–71
 Fas 3:70
 Fas ligand (FasL) 2:226, 3:63–64
 FAS *see* fatty acid synthase (FAS)
 FAS1 *see* fasciclin domain (FAS1)
 FAS-associated death domain protein
 (FADD) 3:67, 3:335, 3:356–357,
 3:393–394
 fasciclin domain (FAS1) 3:727, 3:732
 fasciclin 2:687–688
 FasL *see* Fas ligand (FasL)
 fast atom bombardment (FAB) 1:79
 fast scanning 2:113

fasted state, metabolism in 1:294
 action of glucagon 1:294
 glucose metabolism in liver 1:295–296
 fatty acid oxidation and ketogenesis
 1:296
 gluconeogenesis, regulation of 1:296
 glycogen degradation, regulation of
 1:295–296
 prolonged starvation 1:294–295
 fast-killing process 2:845–846
 FAT10 1:588
 fatal familial insomnia (FFI) 2:839,
 2:840–841
 fatty acid desaturases 1:290
 fatty acid metabolism 3:814–815
 fatty acid oxidation (FAO) 1:294, 1:296,
 3:814
 fatty acid synthase (FAS) 1:290, 1:294,
 3:343
 fatty acid synthase complexes 1:14
 fatty acid synthesis, regulation of 1:294
 fatty acid α -oxidation 2:254
 fatty acid β -oxidation 2:252, 2:253F
 fatty acids 1:154
 fatty acyls 1:153–154
 F-BAR domain group *see* FCH-BAR-Fes/CIP4
 homology-BAR domain group
 (F-BAR domain group)
 F-BAR extension domain (FX domain)
 2:499–500
 FBN1 *see* fibrillin-1 (FBN1)
 FBP17 *see* formin-binding protein 17
 (FBP17)
 FC *see* fibrillar centre (FC)
 FCC *see* focal chemotactic chambers (FCC)
 FCCS *see* fluorescence cross-correlation
 spectroscopy (FCCS)
 FCH-BAR-Fes/CIP4 homology-BAR domain
 group (F-BAR domain group) 2:491,
 2:494F
 fCJD *see* familial CJD (fCJD)
 FCMD *see* Fukuyama congenital muscular
 dystrophy (FCMD)
 FcR *see* receptors for Fc portion of
 immunoglobulins (FcR)
 FCS *see* fluorescence correlation
 spectroscopy (FCS)
 Fc γ receptor (Fc γ R) 2:752, 2:766–767,
 3:695
 activation by IgG 3:696–697
 FDA *see* Federal Drug Administration
 (FDA); US Food and Drug
 Administration (FDA)
 FDC *see* follicular dendritic cells (FDC)
 FDCs *see* follicular dendritic cells (FDCs)
 Fdx *see* fluorescein dextran (Fdx)
 fecal transplants 2:801, 2:804
 fed state
 metabolism in 1:290
 action of insulin 1:290
 glucose metabolism in liver 1:290–291
 substrates (fuel) availability 1:289
 Federal Drug Administration (FDA) 3:821
 feedback 4:32, 4:32–33, 4:32F, 4:33
 graphical understanding of 4:31
 negative 4:30–31, 4:31F, 4:33, 4:33F

- feedback control of core amplification 3:203
 feedback loop 3:131, 3:132, 4:52–53, 4:52F, 4:53F, 4:218, 4:219
 feedback mechanism 3:13
 negative 4:46, 4:46–47, 4:47F, 4:48, 4:48–49
 positive 4:46, 4:46–47
 feedforward loops 4:67, 4:67F
 FEG *see* field emission gun (FEG)
 feline sarcoma (fes) 2:499–500
 female infertility 1:636
 FemX proteins 1:326
 FEN1 *see* flap endonuclease 1 (FEN1)
 Fer protein *see* Fes-related protein (Fer protein)
 FERM domain *see* 4.1/ezrin/radixin/moesin domain (FERM domain)
 fermentor 1:732
 fertilization, hyaluronan in 1:282
 fes *see* feline sarcoma (fes)
 Fes-related protein (Fer protein) 2:499–500
 fetuin 1:706
 Fetuin-A 1:691–693
 FFI *see* fatal familial insomnia (FFI)
 FG *see* phenylalanine-glycine (FG)
 FGF receptor substrate 2 (FRS2) 4:105–106
 FGF *see* fibroblast growth factor (FGF)
 FGF2 3:790–791
 FGF2-Y81pCMF *see* phosphomimetic substitution in FGF2 (FGF2-Y81pCMF)
 FGFR *see* fibroblast growth factor receptor (FGFR)
 FGF–VEGF–Ang–PDGF networking 4:108–109
 FH domains *see* formin homology domains (FH domains)
 FH1 *see* formin homology domains 1 (FH1)
 FHL *see* familial hemophagocytic lymphohistiocytosis (FHL)
 FHV *see* Flock House virus (FHV)
 fl membrane-attack complex domain (FIMAC domain) 1:729
 FIB *see* focused ion beam (FIB)
 Fibril-associated Collagens with Interrupted Triple Helices collagens (FACIT collagens) 2:700
 fibrillar adhesions 2:714
 fibrillar centre (FC) 3:254
 fibrillar collagens 1:634, 2:698–700
 fibrillin 1:634–635
 fibrillin-1 (FBN1) 3:46–47
 fibrinolysis 1:714, 1:717, 1:719
 fibroblast activation protein (FAP) 3:63–64, 3:797
 FAP-alpha 3:619–620
 fibroblast growth factor (FGF) 2:614, 3:156, 3:224, 4:102, 4:102–103
 FGF23 2:348
 FGFR signaling 4:105–106
 HSPGs 4:105
 ligands 4:104
 receptors 4:104–105
 fibroblast growth factor 2 (FGF2) 2:520
 membrane translocation
 hallmarks 2:520–522
 molecular mechanism 2:522, 2:524F
 transient intermediate in 2:522–523
 variant forms analysis 2:524–525
 fibroblast growth factor receptor (FGFR) 4:102
 FGFR1 4:102–103
 signaling 1:124
 FGFR2 3:336–337
 fibroblast specific protein 1 (FSP-1) 3:620
 fibroblastic reticular cells (FRCs) 3:525–526, 3:616
 fibroblasts 1:777, 1:777–778, 1:782, 3:6–7
 fibrodysplasia ossificans progressiva (FOP) 3:47
 fibromodulin 1:273–274
 fibronectin 2:698
 fibronectin type III domains (FnIII domains) 3:22–23
 fibrosis 1:768, 1:793, 3:195–196, 3:515
 mepriins in 1:695–696
 FIB-SEM *see* focused ion beam scanning electron microscopy (FIB-SEM)
 ficolin 1:726
 field emission gun (FEG) 2:14, 2:23–24, 2:24, 2:28
 15°C compartment 2:168–169
 FIH-1 *see* factor inhibiting HIF-1 (FIH-1)
 FIIND *see* function-to-find domain (FIIND)
 filamentous actin (F-actin) 2:597, 4:19
 cytoskeleton 2:704
 networks 2:705
 filamentous proteins promoting curvature 2:593
 filamentous pseudopodia 2:686
 filensin 2:569
 mutations 2:576
 filopodia 2:548, 2:561, 2:562F, 2:683, 2:686, 2:733, 2:4, 3:233, 4:111–112
 see also lamellipodia
 behavior and functions 2:686
 molecular machinery for filopodial protrusion 2:687–688, 2:687F
 actin cytoskeleton 2:687–688
 membrane deformation 2:688
 structure and dynamics 2:686–687
 FIMAC domain *see* fl membrane-attack complex domain (FIMAC domain)
 fingerprint metabolome 4:202
 finite-element methods 4:118–119, 4:120
 Finno-Ugric ancestry 1:775
 FIP1C *see* RCP
 firm adhesion 3:751–752
 first law of thermodynamics 1:5
 first-order reaction 1:8
 Fis1 2:285
 fish keratocyte 4:12
 FISH *see* fluorescence *in situ* hybridization (FISH)
 FISSEQ *see* fluorescent *in situ* RNA-seq (FISSEQ)
 fission 2:225
 fixed cell imaging 2:70
 FK506-binding protein (FKBP) 3:473–475
 FKBP12 3:244–245
 FKBP *see* FK506-binding protein (FKBP)

- fluorescence correlation spectroscopy (FCS) (*continued*)
 by two-color cross-correlation spectroscopy 2:102–103
 nanophotonic and super-resolution approaches 2:101–102
 relationship between FCS and fluorescence photobleaching measurements 2:100–101
 as single molecule approach 2:101
- fluorescence correlation spectroscopy (FCS) 2:2
- fluorescence cross-correlation spectroscopy (FCCS) 2:53, 2:102
- fluorescence energy transfer (FRET) 1:100
- fluorescence *in situ* hybridization (FISH) 2:91
- fluorescence lifetime imaging (FLIM) 2:2, 2:107–120, 2:108, 2:129–131, 4:242
- applications in life sciences
 autofluorescence FLIM 2:110
 FRET experiments 2:109–110
 local molecular environment parameters measurement 2:108–109
- clinical applications
 multiphoton tomography of skin 2:111
 ophthalmic FLIM 2:111–112
- data analysis 2:110
- data analysis
 frequency-domain data 2:117
 time-domain data 2:116–117
- latest developments 2:117–118
- technical aspects
 for cell biology 2:112–113
 frequency-domain FLIM 2:116
 gated photon counting 2:115
 TCSPC FLIM principle 2:114F
 time-domain FLIM 2:114–115
 time-domain wide-field FLIM 2:115–116
- technique
 fast scanning 2:113
 live cell compatibility 2:113–114
 multiwavelength capability 2:114
 photon efficiency 2:112–113
 resolution of complex decay profiles 2:113
 single-photon capability 2:113
 suppression of out-of-focus signals 2:113
 time resolution 2:113
 tolerance to dynamic effects 2:113
- fluorescence lifetime ophthalmoscope (FLIO) 2:112
- fluorescence lifetime-transient scanning (FLITS) 2:118
- fluorescence loss in photobleaching (FLIP) 3:259
- fluorescence photoactivation localization microscopy (FPALM) 2:62, 2:62F
 basic layout 2:63F, 2:70, 2:71F, 2:72F
- fluorescence photobleaching recovery (FPR) 2:52–53, 2:52F, 2:64–65, 2:99, 4:21
- fluorescence recovery after photobleaching (FRAP) 2:2, 3:203
see also fluorescence photobleaching recovery (FPR)
- fluorescence resonance energy transfer (FRET) 1:627, 2:753
 in raft detection 1:213, 1:214
 TIR combined with fluorescence recovery 2:64–65
- fluorescence speckle microscopy (FSM) 2:640
- fluorescence-activated cell sorting (FACS) 3:759, 3:789, 3:807, 3:808, 3:809–810, 3:810, 4:221
- fluorescent *in situ* hybridization 4:162
- fluorescent *in situ* RNA-seq (FISSEQ) 4:160, 4:162
- fluorescent proteins 2:51–52
see also genetically encoded fluorescent probes
 calcium-dependent fluorescence using GCaMP 2:53F
 FRAP 2:52F
 imaging with multiple fluorescent proteins 2:52F
 2-photon microscopy 2:55, 2:55F
 as probes for function 2:52–54
 single-particle tracking with 2:57F
 super-resolution imaging 2:58F
- fluoxetine 1:248
- flux connectivity theorem 4:28–29
- flux control coefficient 4:27–28
- FMDV *see* foot and mouth disease virus (FMDV)
- fmi see* Flamingo (*fmi*)
- fMLF *see* formyl-Met-Leu-Phe (fMLF)
- fMLP *see* formyl-Met-Leu-Pro (fMLP)
- FMS-like tyrosine kinase-3 (FLT-3) 3:624–625
- FnIII domains *see* fibronectin type III domains (FnIII domains)
- foamy viruses *see* Spumaretroviruses
- focal adhesion kinase (FAK) 2:561–564, 2:715–716, 2:724, 2:785, 3:208–209, 3:335, 3:628–629
- focal adhesions 2:62, 2:714F, 2:715, 2:715F, 3:195, 3:213, 3:213F
 proteins 2:724
- focal chemotactic chambers (FCC) 3:670–671F, 3:671–672
- focused ion beam (FIB) 2:1, 2:19, 2:44
- focused ion beam scanning electron microscopy (FIB-SEM) 2:44, 2:45F, 2:46
- foldases 2:162
- folding, proteins 1:108, 1:108–109
 frustration 1:111
 mechanisms and search for native state conformation 1:109–110
 folding nucleus 1:110
 folding pathway 1:110
 folding transition state 1:110–111
 molten globule 1:110
 principles 1:108–109
 protein stability 1:109
- folding diseases 1:543–544
- follicular dendritic cells (FDCs) 3:525–526, 3:543, 3:616, 3:762–763
- folliculin (FLCN) 3:249
- folliculin (FST) 3:44
- foot and mouth disease virus (FMDV) 2:822
- footprint analysis 4:202
- FOP *see* fibrodysplasia ossificans progressiva (FOP)
- forces acting at synapses 3:660–661
- foreign-eating *see* xenophagy
- forkhead-C2 (FOX-C2) 4:103
- forkhead homeobox type O (FoxO) 3:815
- formin homology domains (FH domains) 2:734
 FH1 2:557
 FH2 2:557
- formin-binding protein 17 (FBP17) 2:492–493
- formins 2:685–686, 2:685F, 2:687
- formyl-Met-Leu-Phe (fMLF) 3:705
- formyl-Met-Leu-Pro (fMLP) 2:724, 3:752
- Förster resonance energy transfer (FRET) 2:54, 2:54F, 2:108, 2:459, 3:653–655, 3:661–662, 3:663F, 4:221–222, 4:222, 4:237
 experiments 2:109–110
- Fortius 2:602
- 42kDa proenzyme pepsinogen 1:733
- Fos 3:574–575, 3:575, 4:47
- four phosphate adaptor proteins (FAPP) 2:503–504
- 4.1/ezrin/radixin/moesin domain (FERM domain) 2:480, 3:99–100
- Fourier ring correlation method 2:65
- Fourier Shell Correlation (FSC) 2:7–8, 2:8, 2:9
- Fourier space 2:86–87
- 14-3-3 proteins 2:503
- FOX-C2 *see* forkhead-C2 (FOX-C2)
- FoxO transcription factors 3:781
- FoxO3a *see* transcriptional activity of Forkhead box O3 (FoxO3a)
- FoxP3 3:576–577
- FPALM *see* fluorescence photoactivation localization microscopy (FPALM)
- FPR *see* fluorescence photobleaching recovery (FPR)
- fps *see* frames per second (fps); Fujinami poultry sarcoma (fps)
- Fps/Fes 2:499–500
- fraction collector 1:75
- fractional diffusion cardiac models 4:90
- fractionation methods 4:180
- fracture-label 2:39
- fragment screening 1:744
- frames per second (fps) 2:81
- FRAP *see* fluorescence recovery after photobleaching (FRAP)
- Frasier syndrome 1:501
- FRB *see* FKBP12-rapamycin-binding (FRB)
- FRCs *see* fibroblastic reticular cells (FRCs)
- free energy 1:298
- freeze etching 2:20
- freeze-fracture replica labeling (FRL) 2:39–40, 2:39–40, 2:40F
- freeze-substitution 2:34–36
- freezing/thawing 2:36–37
- frequency-domain data 2:117

- FRET *see* fluorescence energy transfer (FRET); fluorescence resonance energy transfer (FRET); Förster resonance energy transfer (FRET)
- Frizzled 2:746–747, 3:2, 3:201–202 and LRP5/6 3:87–88, 3:88F
- FRL *see* freeze-fracture replica labeling (FRL)
- front loading 4:29
- front–back polarity 2:726–727
- FRS2 *see* FGF receptor substrate 2 (FRS2)
- frustrated phagocytosis 3:729
- FSC *see* Fourier Shell Correlation (FSC)
- FSM *see* fluorescence speckle microscopy (FSM)
- FSP-1 *see* fibroblast specific protein 1 (FSP-1)
- FSP27 2:274
- FSSP *see* families of structurally similar proteins (FSSP)
- FST *see* follistatin (FST)
- FTase *see* farnesyltransferase (FTase)
- FtsA filaments 2:590
- FtsZ filaments 2:590
- 5-FU *see* 5-fluorouracil (5-FU)
- functional cell biology of immunity 3:518–527
- adaptive immunity elements 3:519
 - barriers and tissue cells 3:519
 - cell biology 3:518–519
 - chemical defenses 3:519–520
 - evolution 3:526
 - ILC 3:521–522
 - immunological synapses 3:522–525, 3:524F
 - innate immunity elements 3:519
 - LT 3:525–526
 - lymphocytes 3:522
 - model immunological synapse 3:523
 - phagocytes and receptors 3:521
 - sensing mechanisms 3:520–521
- fucose (Fuc) 2:339–340
- Fujinami poultry sarcoma (fps) 2:499–500
- Fukuyama congenital muscular dystrophy (FCMD) 2:348–349
- full width at half maximum (FWHM) 2:113
- functional aspects of cell biology 3:1–7
- cell growth 3:3
 - cell cycle progression 3:4–6
 - cell death regulation 3:4
 - ribosome biogenesis pathways 3:3
 - translation 3:3–4
 - cell signaling and intercellular communication 3:1–2
 - cellular machineries 3:2
 - cellular outcomes 3:2–3
 - receptors 3:1–2
 - signal transduction pathways 3:2
 - immunity, functional cell biology of 3:6
 - stem cell biology 3:6–7
- function-to-find domain (FIIND) 3:717
- fungal sex 2:254
- FunWalk 4:213
- furin 1:555–556, 1:556, 1:556, 1:556F, 1:558, 1:562, 1:562–563, 1:563, 1:563–564, 1:564, 1:564T
- cleavage of prodomain 1:557
- enzymatic properties and substrate cleavage 1:556–557
- general properties of 1:556
- physiological significance of 1:557
- fusion 2:225–226, 2:229–230, 2:377–378
- events 2:412
- process 2:412
- proteins 1:135–136
- of transport carrier 2:371
- fusogenic proteins 2:461–463
- futile cycles 4:29–30
- FWHM *see* full width at half maximum (FWHM)
- FX domain *see* F-BAR extension domain (FX domain)
- FXR *see* farnesoid X-receptor (FXR)
- FYCO 1 *see* FYVE and coiled-coil domain containing protein 1 (FYCO 1)
- FYN tyrosine kinase 3:420
- FYVE and coiled-coil domain containing protein 1 (FYCO 1) 2:414, 2:504, 2:511, 2:515–516
- ## G
- G protein-coupled receptors (GPCRs) 1:248–249, 2:446–447, 2:724, 2:846, 3:1–2, 3:51, 3:52, 3:102–103
- cross talk between GPCRs and RTK pathways 3:51F
 - growth factor receptors and 3:51–52
 - signaling pathways 3:51
- G proteins *see* GTP-binding proteins (G proteins)
- G0 *see* quiescence
- G1 phase 3:424, 3:424–425, 3:426
- G1/S DNA damage checkpoint 3:436
- G3P *see* glyceraldehyde-3-phosphate (G3P); glycerol-3-phosphate (G3P)
- GA *see* gibberellic acid (GA); Golgi apparatus (GA)
- GABA *see* gamma-aminobutyric acid (GABA)
- G-actin
- binding proteins 2:554–557
 - sequestering proteins 2:556
 - transport 4:13
- GADD153 *see* growth-arrest and DNA damage inducible protein (GADD153)
- GAE homology *see* gamma-adaptin ear homology (GAE homology)
- GAF *see* gamma-activated factor (GAF)
- Gag coordinates HIV assembly 2:829–832
- Gag oligomerization 2:830
- GAGs *see* glycosaminoglycans (GAGs)
- gain-of-function proteotoxicity 1:792
- gal operon 1:475
- Gal *see* galactose (Gal)
- galactose (Gal) 1:271, 2:339–340
- galactosialidosis 1:774
- galactosylcer amides (GalCer) 1:184, 1:187
- GalCer *see* galactosylcer amides (GalCer)
- GalNAc *see* N-acetylgalactosamine (GalNAc)
- GALT *see* gut-associated lymphoid tissue (GALT)
- γ -aminobutyric acid (GABA) 1:246, 1:249, 1:249–250, 3:799–800
- GABA receptor-associated protein (GABARAP) 1:618
- γ c *see* gamma common (γ c)
- γ -carboxylation 1:88
- gamma common (γ c) 3:24–25
- receptors 3:24–25
 - IL-2/IL-15 receptors 3:25–26
 - IL-4/IL-13 receptors 3:26
 - IL-7/TLSP receptors 3:26–27
- γ -dynein 2:627
- gamma interferon-inducible lysosomal thiolreductase (GILT) 3:590, 3:591
- gamma-activated factor (GAF) 3:143
- gamma-adaptin ear homology (GAE homology) 2:453
- γ SA *see* γ -specific array (γ SA)
- γ -secretase
- see also* β -Secretase
 - identification and mechanism 1:664
 - inhibitors and modulators 1:665
 - other substrates 1:664–665
 - prospects for therapeutics 1:666
 - structure elucidation 1:665–666
- γ -specific array (γ SA) 3:153
- γ -tubulin ring complex (γ -TuRC) 2:638, 2:653
- Gap period (G1phase) 3:447
- GAP *see* GTPase activating protein (GAP); guanosine-activating protein (GAP)
- ganglion mother cell (GMC) 2:747
- ganglioside M1 (GM1) 1:209–210
- ganglioside-induced differentiation-associated protein 1 (GDAP1) 2:285–286
- gangliosides 1:149, 1:155
- gap genes 4:66
- GAP-43 *see* growth-associated protein-43 (GAP-43)
- GAPs *see* GTPase activating proteins (GAPs)
- GARP *see* Golgi-associated retrograde protein (GARP)
- gas chromatography (GC) 4:203, 4:203–204, 4:204–205
- Gas6 *see* growth arrest-specific gene 6 (Gas6)
- GAS7 *see* growth arrest-specific protein (GAS7)
- gastric carcinoma (GC) 3:146
- gastric development, role in 1:685–686
- gastric ulcer 1:677
- gastrointestinal stromal tumor (GIST) 3:17
- GAT *see* GGA and Tom (GAT)
- gated image intensifiers, time-domain wide-field FLIM by 2:115–116
- Gaucher disease 1:767–768, 1:769–770, 1:772, 1:773, 1:774, 1:778, 1:779–780, 1:779, 1:780, 1:781, 1:782, 1:783, 1:785
- chronic neuronopathic 1:772
 - comorbid associations of 1:772–773
 - B-cell lymphoma and myelomatosis 1:773

- Gaucher disease (*continued*)
 energy metabolism 1:772–773
 Parkinsonism 1:773
 infantile 1:772
 molecular therapy for 1:781–782
- Gaussian function 2:100, 2:101
- GBA1 gene 1:767, 1:772, 1:773
- GBPs *see* glycan-binding proteins (GBPs)
- GBSS *see* granule-bound starch synthase (GBSS)
- GC *see* gas chromatography (GC); gastric carcinoma (GC); germinal center (GC); granular component (GC); guanylyl cyclases (GCs)
- GCA 3:58
- GCB *see* germinal center B cell (GCB)
- GCC *see* guard cell chloroplasts (GCC)
- GCC88 2:436, 2:437
- GCC185 2:436, 2:437
- GCL *see* granule cell layer (GCL)
- G-clathrin *see* gyrate-clathrin (G-clathrin)
- GC-MS 4:203–204, 4:205–206, 4:205F, 4:206F, 4:207
- G-CSF *see* granulocyte colony-stimulating factor (G-CSF)
- G-CSFR *see* granulocyte-colony stimulating factor receptor (G-CSFR)
- GDAP1 *see* ganglioside-induced differentiation-associated protein 1 (GDAP1)
- GDF *see* GDP displacement factor (GDF); growth and differentiation factors (GDFs)
- GDI *see* GDP-dissociation inhibitor (GDI); guanine nucleotide displacement inhibitor (GDI)
- GDNF *see* glial-derived neurotrophic factor (GDNF)
- GDP displacement factor (GDF) 2:443
- GDP *see* guanosine diphosphate (GDP)
- GDP-bound inactive state 2:795–796
- GDP-dissociation inhibitor (GDI) 2:402, 2:443, 3:703–704
- GDP-GTP exchange factor (GEF) 3:494–495
- GDP-tubulin 2:541
- GED *see* GTPase effector domain (GED)
- GEEC pathway *see* GPI-anchored protein-enriched early endosomal compartment (GEEC pathway)
- GEF *see* GDP-GTP exchange factor (GEF); GTP exchange factors (GEFs); guanine exchange factor (GEF); guanine nucleotide exchange factor (GEF)
- gelatinase granules (GG) 2:770
- gelatinases 1:621–622
- gene assembly from DNA microarrays 1:49–50
- gene duplication 1:81
- gene editing 1:515
- gene encoding lysosomal intrinsic membrane protein (LIMP-2) 1:773–774
- gene expression
 profiling technologies 4:252
 variability in 4:220
- gene expression network analysis (GXNA) 4:212–213, 4:212T
- gene of interest (GOI) 1:122, 3:808–809
- gene ontology (GO) 4:211, 4:247
- gene regulation 1:370, 1:375, 1:377–380, 1:379F
- gene regulatory networks (GRNs) 4:63–71, 4:156–157
- global properties 4:67
- modularity 4:68–70, 4:69F
- scale-free networks 4:68, 4:69F
- small-world networks 4:67–68, 4:69F
- by HH Pathway 3:77–78
- local properties 4:64, 4:64F
- GRN maps construction 4:64–65
- modeling *cis*-regulatory interactions 4:65
- motifs 4:65, 4:65F
- cross-repression 4:66
- feedforward loops 4:67, 4:67F
- negative feedback 4:66–67
- positive feedback 4:67
- repression 4:65–66
- gene traffic 4:155–156
- gene transfer 1:335, 1:763, 1:778, 1:783, 1:783–784, 1:785
- direct administration of vectors to the brain 1:784
- indirect use of vector-transduced autologous hematopoietic stem-cells 1:784–785
- gene trap 1:514
- general control nonderepressing 2 1:309
- general control nonderepressing 4 1:309
- general transcription factor (GTF) 1:481
- generalized lymphoproliferative disease (gld) 3:70
- generalized myotonia 1:231
- generation of noncoding ncRNA 3:785
- genes, proteins and 4:216
- genetic fate mapping 4:137–138
- genetic clonal evolution theory 3:807
- genetic code 1:310–313, 1:311F, 1:316, 1:318, 1:319, 1:320, 1:335, 1:525, 1:530, 1:532, 1:546, 1:569
- in mitochondria 1:546
- genetic(s) 1:485
- alterations
 DNA damage and repair 3:777
 telomere shortening 3:777
- disorders 3:691–692
- genetically encoded fluorescent probes 2:52
see also fluorescent proteins
- breaking diffraction barrier 2:56
- competing parameters for optimal imaging 2:58F
- imaging cells in whole organisms 2:54–56
- technical considerations for live imaging 2:56–59
- uses 2:59–60
- genome sequencing for interactome mapping 4:193
- genome stability 1:461
- genome wide chromatin immunoprecipitation data (GW-ChIP) 4:257
- genome-wide association studies (GWAS) 3:454
- genome-predicted proteome 4:170
- genome-wide analysis 1:485
- genome-wide association studies (GWAS) 3:718, 4:171, 4:194–195, 4:254
- genome-wide studies 1:555, 1:559, 3:778, 4:246
- genomic(s) 4:166
- instability 1:451
- genotoxic stresses suppress mTORC1 3:248
- Genzyme Corporation 1:780, 1:781
- geranylgeranyl transferase (GGT) 2:241
- geranylgeranylation 2:239–241
- germ layers 3:785, 3:787, 3:788–789
- germ-free organisms (GF organisms) 2:799–800
- germinal center (GC) 3:541, 3:543–547, 3:618
- germinal center B cell (GCB) 3:58, 4:252–253
- germ-line mutations 1:507–508, 1:508, 3:126–127, 3:127
- germline stem cells (GSCs) 2:656
- germline transcription 3:555
- gerontoplast 2:327
- Gerstmann–Sträussler–Scheinker (GSS) syndrome 2:839, 2:840
- GET pathway *see* guided entry of TA proteins pathway (GET pathway)
- GF organisms *see* germ-free organisms (GF organisms)
- GFAP *see* glial fibrillary acidic protein (GFAP)
- GFP *see* green fluorescent protein (GFP)
- GFR signaling *see* growth factor receptor (GFR) signaling
- GG *see* gelatinase granules (GG)
- GGA and Tom (GAT) 2:453
- GGAs *see* Golgi localized γ -ear-containing ARF-binding proteins (GGAs)
- GGT *see* geranylgeranyl transferase (GGT)
- GHK equation *see* Goldman–Hodgkin–Katz (GHK) equation
- GHR *see* growth hormone receptor (GHR)
- GI *see* glucosidase I (GI)
- giant plasma membrane vesicles (GPMV) 1:212
- gibberellic acid (GA) 2:324
- Gibbs free energy 1:5–6
 additive changes 1:6–7
 decreases for spontaneous process 1:6
- GII *see* glucosidase II (GII)
- Gillespie algorithm 4:58
- GILT *see* gamma interferon-inducible lysosomal thiolreductase (GILT)
- GIST *see* gastrointestinal stromal tumor (GIST)
- GL *see* glomerular layer (GL)
- GL1001 *see* MLN-4760
- glass-SLBs *see* glass-supported artificial lipid bilayers (glass-SLBs)
- glass-supported artificial lipid bilayers (glass-SLBs) 3:661
- Glc *see* glucose (Glc)
- GlcA *see* glucuronic acid (GlcA)

- GlcCer *see* glucosylceramides (GlcCer)
 GlcNAc *see* N-acetylglucosamine (GlcNAc)
 gld *see* generalized lymphoproliferative disease (gld)
 Gle2-binding-sequence (GLEBS) 3:300
 GLEBS *see* Gle2-binding-sequence (GLEBS)
 GLFG *see* Gly-Leu-Phe-Gly (GLFG)
 GLI1 *see* GLI-Kruppel family member GLI1 (GLI1)
 glia maturation factor (GMF) 2:559
 glial fibrillary acidic protein (GFAP) 2:569, 3:790–791, 3:799, 4:134
 expression 4:136
 mutations 2:576
 glial-derived neurotrophic factor (GDNF) 3:13, 3:16
 GLI-Kruppel family member GLI1 (GLI1) 3:794–795
 glioblastoma 3:809
 glioma-associated transcription factors 3:77–78
 glmS ribozyme-riboswitch 1:376T, 1:378T, 1:380
 global regulation 1:528F, 1:529, 1:532–533
 globular tail domain (GTD) 2:445
 globule 1:793
 globulins 2:382
 glomerular layer (GL) 4:135F
 Glu78 glutamic acid residue 1:304
 glucagon
 action of 1:294
 and metabolic processes 1:289
 signaling pathway 1:295F
 glucan phosphatase 1:268–269
 glucan water dikinase (GWD) 1:268–269
 glucocorticoid-stimulated kinase (SGK) 3:132
 gluconeogenesis, regulation of 1:293F, 1:296
 glucosamine-6-phosphate (GlcN6P) 1:378T, 1:380
 glucose (Glc) 1:101, 1:181
 glucose homeostasis 1:289, 1:294, 1:295–296, 1:296
 glucose metabolism in liver
 during fasting 1:295–296
 fatty acid oxidation and ketogenesis 1:296
 gluconeogenesis, regulation of 1:296
 glycogen degradation, regulation of 1:295–296
 fed state 1:290–291
 fatty acid synthesis, regulation of 1:294
 glycogen synthesis, regulation of 1:291–292
 glycolytic pathway, regulation of 1:292–293
 protein synthesis 1:294
 pyruvate dehydrogenase complex, regulation of 1:293–294
 transport and phosphorylation of glucose 1:290–291
 pathways involved in 1:289F
 glucose transporter 3 (GLUT3) 3:336
 glucose transporter 4 (GLUT4)
 storage compartment formation 2:390
 transporter 2:378–379
 glucose-6-phosphatase 1:294, 1:295–296, 1:296
 glucose-6-phosphate dehydrogenase 1:294
 glucose-regulated protein of 78 kDa (GRP78) 2:529
 glucose-regulated protein of 94 kDa (GRP94) 2:169
 glucosidase I (GI) 1:599–601
 glucosidase II (GII) 1:599–601
 glucosidase I/II 1:539–540T, 1:540, 1:541F
 glucosylceramides (GlcCer) 1:184, 1:187
 glucuronic acid (GlcA) 1:181, 1:271, 2:347
 GLUD1 *see* glutamate dehydrogenase 1 (GLUD1)
 GLUL *see* glutamate-ammonia ligase (GLUL)
 GLUT3 *see* glucose transporter 3 (GLUT3)
 GLUT4 storage vesicles (GSV) 2:402
 glutamate dehydrogenase 1 (GLUD1) 2:738, 3:344
 glutamate receptors 2:608
 glutamate-ammonia ligase (GLUL) 3:344
 glutaminy cyclase inhibitor 1:123–124
 glutathione (GSH) 3:336
 glutathione S-transferase (GST) 1:73
 glycan 1:599
 glycan headgroup addition 1:189–190
 glycan-binding proteins (GBPs) 2:343
 glycans 2:162–163
 glyceraldehyde-3-phosphate (G3P) 2:328–329
 glycerol-3-phosphate (G3P) 2:276
 glycerolipid 1:154–155, 1:160, 2:273
 phosphatidic acid synthesis 1:161–162
 phospholipids synthesis in bacteria 1:162
 amine-containing phospholipids synthesis 1:162
 non-amine-containing phospholipids synthesis 1:162–164
 phospholipids synthesis in eukaryotes 1:164
 amine-containing phospholipids synthesis 1:164–166, 1:165F
 non-amine-containing phospholipids synthesis 1:166–169
 physical and chemical properties 1:169F
 roles of phospholipids in cell function 1:169–170
 structure 1:160F
 glycerophospholipids 1:155, 1:160–161, 1:195, 1:202, 1:205
 glycine (Gly) 1:15T, 1:203, 1:233F, 1:249, 1:310, 1:377
 glycine-and serine-rich domain (GS domain) 3:39–40
 glycolalcyces 1:280–281
 catabolism of 1:281F
 hyaluronan-dependent 1:281F
 glycocalyx matrix 3:193
 glycogen 1:263–264, 1:764, 1:764–765, 1:767, 1:768, 1:768–769, 1:769, 1:770, 1:771–772, 1:772, 1:779
 degradation 1:266–267, 1:295–296
 enzyme activities 1:266F
 function 1:265
 metabolism 1:289F, 1:291, 1:291–292, 1:295–296, 1:295F
 genetic modification 1:269–270
 glycogen degradation 1:266–267
 glycogen phosphorylation 1:267
 glycogen synthesis 1:265–266
 regulation 1:267
 phosphorylation 1:267
 structures 1:264F
 synthesis 1:265–266
 glycogen storage disease type II *see* Pompe disease
 glycogen storage diseases (GSDs) 1:269
 glycogen synthase 1:291–292, 1:292
 glycogen synthase kinase 3 (GSK3) 1:675, 3:90, 3:90F, 3:91, 3:91F, 3:92, 3:93F, 3:235, 3:573
 glycogen synthase kinase 3 beta (GSK3 β) 3:779
 glycogen synthesis, regulation of 1:291–292
 glycogenin 1:265–266
 glycogenoses *see* glycogen storage diseases (GSDs)
 glycogenosis II *see* Pompe disease
 glycolglycerolipids 1:180–182, 1:181F
 glycolipidome 1:190–191
 glycolipids 1:160, 1:169–170, 1:180–193
 future 1:191–192
 omic technologies 1:190–191
 subcategories 1:180–182
 types 1:180–182
 glycolglycerolipids 1:180–182, 1:181F
 glycophosphoglycerolipids 1:182–183, 1:182F
 glycosphingolipids 1:183–184, 1:183F, 1:185F
 glycosylated polyketides 1:186, 1:188F
 glycosylated prenols 1:185–186, 1:187F
 glycosylated sterols 1:184–185
 saccharolipids 1:186, 1:188F
 glycolysis 1:240–241, 4:24, 4:25F, 4:28, 4:29
 in stem cells 3:813–814
 glycolytic cardiac progenitors 3:801
 glycolytic pathway 1:289F
 regulation of 1:292–293, 1:293F
 glycoposphatidyl inositol
 see glycosylphosphatidylinositol (GPI)
 glycophosphoglycerolipids 1:182–183, 1:182F
 glycoprotein quality control system 1:599–601
 glycoproteins 1:86
 processing and folding in ER 2:341
 glycosaminoglycans (GAGs) 1:271, 1:279–280, 1:632, 1:728, 1:763–764, 1:769, 1:770–771, 1:773, 1:775, 1:777, 1:780, 1:782, 1:785, 2:351, 2:694, 2:698, 3:624
 structure and synthesis 1:274
 glycosidases 1:616
 lysosomal biogenesis and regulation 1:616–617
 lysosomal membrane proteins 1:616
 glycosome 2:249–250

- glycosphingolipidosis 1:779
glycosphingolipids 1:183–184, 1:183F, 1:185F, 1:208–209, 1:769, 1:779, 2:438
glycan headgroup addition 1:189–190
power of combinatorial biochemistry 1:186
root structure biosynthetic pathways 1:190F
sphingolipid backbone biosynthesis 1:186–189, 1:189F
glycosyl phosphatidylinositol *see* glycosylphosphatidylinositol (GPI)
glycosylated/glycosylation 1:85–86
at Golgi apparatus 2:186–187
polyketides 1:186, 1:188F
prenols 1:185–186, 1:187F
receptors 3:107
sterols 1:184–185
glycosylphosphatidylinositol (GPI) 1:87, 1:167, 1:182–183, 1:271–272, 1:597–598, 1:621–622, 1:650, 1:663, 1:755, 2:161, 2:365, 2:394, 2:745, 3:13, 3:16, 3:61
-anchored proteins 1:208–209, 3:87
glycosylphosphatidylinositol *see* glycosylphosphatidylinositol (GPI)
Gly-Leu-Phe-Gly (GLFG) 2:319–320
glyoxylate
cycle 2:252–254
metabolism 2:252–254
glyphosate 2:329
glypicans 1:275
glypican-3 1:275
glypican-5 1:275
GM2 activator protein 1:774
GM2 gangliosidosis *see* Tay-Sachs disease
GMC *see* ganglion mother cell (GMC)
GM-CSF *see* granulocyte-macrophage colony-stimulating factor (GM-CSF)
GMF *see* glia maturation factor (GMF)
GO *see* gene ontology (GO)
goblet cells 3:689–690
GOI *see* gene of interest (GOI)
gold standard 4:248
gold-conjugated secondary antibodies 2:37
Golden GATeway mutagenesis 1:122
Goldman-Hodgkin-Katz (GHK) equation 1:223–224, 1:224–225, 1:231, 1:232F, 1:233F
Golgi 2:475, 3:87, 3:87F
bypass 2:176
compartments, N-glycans maturation during transit through 2:342
complex 2:355F
glycosylation enzymes 2:186–187
phosphoinositides in Golgi trafficking 2:512–513, 2:514F
stacks 2:168, 2:169, 2:171, 2:354
Golgi apparatus (GA) 1:202, 1:204–205, 1:204F, 1:206, 1:535–536, 1:543, 1:555, 1:557–558, 2:2, 2:168, 2:183, 2:374, 2:512–513, 2:655
evolution 2:188–189
glycosylation 2:186–187
molecular mechanism of transport 2:186
regulation 2:187–188
role of Golgi in disease 2:188
structure 2:183
transport through 2:183–186
Golgi localized γ -ear-containing ARF-binding proteins (GGAs) 2:410, 2:451–452, 2:453–454, 2:504
Golgi phosphoprotein 3 (GOLPH3) 2:513
Golgi phosphoprotein 3-like (GOLPH3L) 2:513
Golgi system 1:777–778
Golgi-associated retrograde protein (GARP) 2:437, 2:485, 2:488
golgins 2:367
GOLPH3 *see* Golgi phosphoprotein 3 (GOLPH3)
GOLPH3L *see* Golgi phosphoprotein 3-like (GOLPH3L)
good buffers 1:71
gp41 C-terminal cytoplasmic domain 2:832
gp91phox 3:703
phosphorylation 3:707
gp120 V3 loop 2:833
GPCR *see* G-protein coupled receptor (GPCR)
GPCRs *see* G protein-coupled receptors (GPCRs)
GPI *see* glycosylphosphatidylinositol (GPI)
GPI-anchored protein-enriched early endosomal compartment (GEEC pathway) 2:394
GPI-anchored proteins *see* glycosylphosphatidylinositol (GPI)-anchored proteins
GPI-anchored serine proteases 1:651–652
GPMV *see* giant plasma membrane vesicles (GPMV)
G-protein coupled receptor (GPCR) 1:198, 2:219–220, 2:373, 2:389, 2:456, 2:475, 2:480, 2:668, 3:153, 3:208, 3:762, 4:51–52
G-proteins 1:244, 2:3
GPS *see* gray platelet syndrome (GPS)
GPU *see* graphical processing units (GPU)
G-quadruplexes 1:427
gram negative pathogens 2:845–846
gram-negative bacteria 1:160, 2:588
gram-negative pathogens 2:777
gram-negative *Pseudomonas fluorescens* 1:57
gram-positive bacteria 1:160, 2:588, 2:845–846
granular component (GC) 3:254
granulation tissue formation 3:219–220
granule cell layer (GCL) 4:135F
granule-bound starch synthase (GBSS) 1:267–268
granule cells 4:134, 4:135F, 4:137, 4:138
granulocyte colony-stimulating factor (G-CSF) 3:626, 3:750
granulocyte-colony stimulating factor receptor (G-CSFR) 3:140–141
granulocyte-macrophage colony-stimulating factor (GM-CSF) 3:24, 3:136
granzymes (Gzms) 3:330
GRAP2 *see* GRB2-related adaptorprotein (GRAP2)
graphical processing units (GPU) 4:119–120
gray platelet syndrome (GPS) 2:227–228
GRB2 *see* growth factor receptor-bound protein 2 (GRB2)
GRB2-related adaptorprotein (GRAP2) 3:534–535
Grb10 3:245–246
Gre factors 1:474
greedy algorithm 4:212, 4:212–213, 4:212T, 4:214
greedy class of algorithms 4:212–213
green fluorescent protein (GFP) 1:513, 2:41–42, 2:51–52, 2:53–54, 2:109, 2:138, 2:168–169, 2:268, 3:300, 4:237
Griffith's model 2:839
Griscelli's syndrome (GS) 2:227–228, 2:237–239, 2:239–241
types 1–3 1:775–776, 1:776
gRNA *see* guide RNA (gRNA)
GRNs *see* gene regulatory networks (GRNs)
GroEL 1:105
GroEL-ES 2:9–11, 2:10F
ground-state depletion after individual molecule return imaging (GSDIM) 2:76, 2:81
biological applications
to biological research 2:82–84
biological sample preparation 2:81–82
commercial microscopes 2:81
ground-state depletion microscopy (GSD microscopy) 2:76
see also stimulated emission depletion (STED) microscopy
scanning 2:81
growth and differentiation factors (GDFs) 3:37
GDF11 3:779
growth arrest-specific gene 6 (Gas6) 2:753–754, 3:375
growth arrest-specific protein (GAS7) 2:497
growth cone 3:209–210, 3:233
growth factor receptor (GFR) signaling 2:758
effect on macropinocytosis 2:759
and GPCRs 3:51–52
growth factor receptor-bound protein 2 (GRB2) 3:119, 3:534–535
growth factor signaling 1:766, 1:766, 1:771, 3:5
growth factors, mTOR signaling regulation by 3:246–247
growth hormone receptor (GHR) 3:141
growth-arrest and DNA damage inducible protein (GADD153) 2:531–532
growth-associated protein-43 (GAP-43) 3:235–236
GRP78 *see* glucose-regulated protein of 78 kDa (GRP78)
GRP94 *see* glucose-regulated protein of 94 kDa (GRP94)
GS *see* Griscelli's syndrome (GS)

- GS domain *see* glycine-and serine-rich domain (GS domain)
- Gsz 3:52
Gsz-induced neurite outgrowth 3:52
translocation 3:53F
- GSCs *see* germline stem cells (GSCs)
- GSD microscopy *see* ground-state depletion microscopy (GSD microscopy)
- GSDIM *see* ground-state depletion after individual molecule return imaging (GSDIM)
- GSDs *see* glycogen storage diseases (GSDs)
- GSH *see* glutathione (GSH)
- GSK3 *see* glycogen synthase kinase 3 (GSK3)
- GSK3 β *see* glycogen synthase kinase 3 beta (GSK3 β)
- GSS syndrome
see Gerstmann–Sträussler–Scheinker (GSS) syndrome
- GST *see* glutathione S-transferase (GST)
- GSV *see* GLUT4 storage vesicles (GSV)
- GTD *see* globular tail domain (GTD)
- GTF *see* general transcription factor (GTF)
- GTP *see* guanosine triphosphate (GTP)
- GTP exchange factors (GEFs) 2:616
- GTP hydrolysis 1:321F, 1:323, 1:388–390, 1:388, 1:390, 1:392, 1:497, 1:537–538, 4:42
- GTPase activating proteins (GAPs) 2:195, 2:333, 2:359, 2:442, 2:478–479, 2:491, 2:504, 2:505–506, 2:616, 2:731–733, 2:759, 2:795–796, 3:12, 3:130, 3:185, 3:246–247, 3:413, 3:479–480, 3:494–495, 4:51
- GTPase cycle 1:537–538, 1:537T, 4:51F
- GTPase effector domain (GED) 2:284, 2:446
- GTPase-associated center 1:385–387, 1:388, 1:392
- GTPases *see* guanosine triphosphatases (GTPases)
- GTP-binding proteins (G proteins) 2:442, 2:442–443
see also Rab escort protein (REP); Rab proteins
- Arf proteins 2:446
- dynamins 2:446
- heterotrimeric 2:446–447
- membrane trafficking 2:442
- NGF inducing neurite outgrowth through 3:52
association between G proteins and tubulin 3:52–53
- GPCRs 3:52
- Gsz 3:52
Gsz-induced neurite outgrowth 3:52
neurite outgrowth 3:52
translocation of Gsz 3:53F
- in protein synthesis and translocation 2:447
- Rab family 2:442
- Ras superfamily 2:445–446
- Rho proteins 2:446
- GTP-bound active state 2:795–796
- GTP-tubulin 2:541, 4:37–38, 4:38
- guanine exchange factor (GEF) 1:309, 1:325, 2:491, 2:583, 3:158, 3:248, 3:479–480
- guanine nucleotide displacement inhibitor (GDI) 2:504, 2:505–506, 2:732F
- guanine nucleotide exchange factor (GEF) 2:170, 2:195, 2:331, 2:359, 2:442, 2:444, 2:479, 2:504, 2:505–506, 2:710, 2:731–733, 2:759, 2:779, 2:787, 2:795–796, 3:12, 3:185, 3:210–211, 3:413, 4:51
- guanosine 5'-triphosphatases *see* guanosine triphosphatases (GTPases)
- guanosine-50-triphosphate (GTP) cap 4:36, 4:42
- guanosine diphosphate (GDP) 2:195, 2:541, 3:119, 3:300, 3:413, 3:703–704
- guanosine triphosphatases (GTPases) 2:485, 2:505–506, 2:753, 2:760, 2:767, 3:299, 4:51
cycle 2:332–333
vesicle tether interactions with SNARE proteins and 2:485–486
- guanosine triphosphate (GTP) 2:195, 2:541, 3:58, 3:119
hydrolysis of 1:323
- guanosine-activating protein (GAP) 2:785
- guanylyl cyclase receptors 3:1–2
- guanylyl cyclases (GCs) 3:56
mammalian GC domain structure 3:56F
- PDE 3:59–60
- pGCs 3:57–58
- sGCs 3:56–57
- guard cell chloroplasts (GCC) 2:329, 3:58–59
- guardian of genome 3:504
- guide RNA (gRNA) 1:515
- guided entry of TA proteins pathway (GET pathway) 1:597–598
- gut microbiome *see* microbiome
- gut-associated lymphoid tissue (GALT) 3:690
- GW182 1:349, 3:321–322
- GWAS *see* genome-wide association studies (GWAS)
- GW-ChIP *see* genome wide chromatin immunoprecipitation data (GW-ChIP)
- GWD *see* glucan water dikinase (GWD)
- GXNA *see* gene expression network analysis (GXNA)
- gyrating-clathrin (G-clathrin) 2:389
- H**
- H₂O₂ *see* hydrogen peroxide (H₂O₂)
- H3 *see* histone 3 (H3)
- H3K4me3 *see* histone H3 at residue K4 (H3K4me3)
- HA *see* hemagglutinin (HA); hyaluronan (HA); hyaluronic acid (HA)
- HA receptor for endocytosis (HARE) 1:282
- HAART *see* highly active antiretroviral therapy (HAART)
- HAI *see* hepatocyte growth factor activator inhibitor (HAI)
- hair follicle
bulge 3:794–795
stem cell populations 3:794–795
- Hairy/Enhancer of Split (HES) 3:107
- Hajdu–Cheney syndrome (HCS) 3:109–110
- halorhodopsins (HRs) 2:148–149
- hamartin *see* tuberous sclerosis 2 (TSC2)
- hammerhead ribozyme 1:19, 1:22F, 1:380
- Hammitt equation 1:135
- HAP1 *see* HTT-associated protein 1 (HAP1)
- HAP40 *see* Htt associated protein (HAP40)
- haploid yeast cells 4:51–52
- haploinsufficiency 1:448
of *Kif1b* gene 2:614
- haptotaxis 3:209
see also chemotaxis
- HARE *see* HA receptor for endocytosis (HARE)
- HARE/stabilin-2 1:282, 1:285
systemic clearance of hyaluronan by 1:283F
- HAS *see* hyaluronan synthase (HAS)
- HAT like-1 (HATL-1) 1:654, 1:654–655
- HAT like-4 (HATL-4) 1:654–655
- HAT like-5 (HATL-5) 1:654–655
- HAT *see* human airway trypsin-like (HAT); histone acetyltransferases (HATs)
- HAV *see* hepatitis A virus (HAV)
- Hayflick limit 3:777
- HB *see* heparin-binding (HB)
- HB-EGF *see* heparin-binding epidermal growth factor (HB-EGF)
- Hbs1 3:320
- HBV *see* hepatitis B virus (HBV)
- HbYX motif 1:589, 1:590, 1:590F
- HCA *see* high-content analysis (HCA)
- hCAP-18 *see* human cationic antimicrobial protein-18 (hCAP-18)
- HCC *see* hepatocellular carcinoma (HCC)
- HCD *see* higher collisional dissociation (HCD); higher energy collisional dissociation (HCD)
- HCS *see* Hajdu–Cheney syndrome (HCS); high-content screening (HCS)
- HCV *see* hepatitis C virus (HCV)
- HCxS *see* high complexity screening (HCxS)
- HD *see* heterodimerization domain (HD); Huntington's disease (HD)
- HDAC *see* histone deacetylase complexes (HDAC); histone deacetylases (HDAC)
- HDL *see* high-density lipoprotein (HDL)
- HDm2 *see* human double minute 2 (HDm2)
- hDNA *see* heteroduplex DNA (hDNA)
- heart disease 1:241
- heat-shock protein 70 (Hsp 70) 1:599
- HECT (homologous to E6AP C terminus) proteins 1:585
- hedgehog (HH) signaling pathway 2:614, 2:668, 3:76–85, 3:76, 4:228–229
gene regulatory networks 3:77–78
in human disease
birth defects with 3:82T

- hedgehog (HH) signaling pathway
(*continued*)
 congenital disease and cancer 3:81–83
 targeting with drugs 3:83
 intracellular effectors of signal
 transduction–transmembrane
 receptors 3:76–77
 in limb developing 3:80F
 long-range transcriptional regulation
 of mammalian SHH gene
 3:80–81
 noncanonical HH signaling 3:79–80
 primary cilia 3:79
 production and secretion of morphogen
 3:76
 signal transduction 3:77F
- HeLa cells 1:442
- helicases 1:403–404, 3:458
 ATP-driven 1:404
 bacterial 1:404
 eukaryotic 1:404–405
 hexameric helicase mechanism 1:404F
 Mcm2–7 1:404–405
 proposed strand exclusion mechanism
 1:403–404
 superfamilies of 1:403–404
- Helicobacter pylori* (*H. pylori*) 2:805–806
- helix-hairpin-helix glycosylases (HhH
 glycosylases) 1:443, 1:444–445
- helper sites 3:94
- hemagglutinin *see* hemagglutinin (HA)
- hemagglutinin (HA) 1:654, 2:41–42
- hematopoiesis 3:108–109, 3:109
- hematopoietic IFN-inducible nuclear
 protein (HIN) 3:719–720
- hematopoietic lineage cell-specific protein1
 (HS1) 3:524
- hematopoietic progenitor cell (HPC)
 3:624
- hematopoietic progenitor kinase 1 (HPK1)
 3:537
- hematopoietic stem cell (HSC) 3:551–552,
 3:592, 3:790, 3:624, 3:813, 3:750,
 3:776, 3:780
 see also multipotent stromal cells (MSCs)
 ATM protein 3:781
 BATF 3:782
 epigenetic analysis 3:782
 FoxO transcription factors 3:781
 hallmarks of aging in 3:780–781
 loss of genome integrity and DNA damage
 3:781
 PTEN 3:781
- hematopoietic-lineage secretory lysosomes
 2:226–228
- heme-nitric oxide (H–NOX) 3:57
- hemi-catenane 1:461
- hemidesmosomes 2:714, 3:219, 3:220
- hemiITAM *see* single tyrosine-based motif
 (hemiITAM)
- hemophilia 1:718
- hemopoietic stem cell (HSC) 3:438
- hemorrhage 1:714
- hemostasis 1:714
- hepadnaviruses 1:523
- heparan sulfate (HS) 1:271
- heparan sulfate proteoglycans (HSPGs)
 1:555, 1:562, 1:562–563, 1:564T,
 2:697, 4:105
- heparanase 1:276
- heparin (HP) 1:271
- heparin sulfate (HS) 2:351, 4:105
- heparin-binding (HB) 3:13
- heparin-binding epidermal growth factor
 (HB-EGF) 1:641
- hepatectomy 1:793
- hepatitis A virus (HAV) 3:308
- hepatitis B virus (HBV) 3:570
- hepatitis C virus (HCV) 1:145, 3:308,
 3:520–521
- hepatitis C virus (HCV) 1:368, 1:520,
 1:529, 2:820–822, 3:332, 3:567,
 3:734
- hepatocellular carcinoma (HCC)
 1:795–796, 3:104
- hepatocyte growth factor (HGF) 1:753,
 2:220, 2:727–728, 2:785, 3:224,
 3:434, 3:809, 4:142–143, 4:218–219
- hepatocyte growth factor activator inhibitor
 (HAI) 1:757
- hepatocyte growth factor-regulated tyrosine
 kinase substrate (Hgs/Hrs) 4:195
- hepatocyte growth regulated tyrosine kinase
 substrate (Hrs) 2:504
- hepatocyte transplantation therapy 1:797
- hepatoma 1:237
- HEPN *see* higher eukaryotes and
 prokaryotes nucleotide-binding
 (HEPN)
- Hepsin 1:655
- Hepsin/TMPRSS Subfamily 1:655–656
- heptad repeat (HR1) 2:279–281
- Herc5 1:584T, 1:587
- Hercetin[®] 1:138
- Hereditary Hemorrhagic Telangiectasia
 (HHT) 3:46–47
- Hereditary mixed polyposis syndrome
 (HMPS) 3:47
- hereditary skin disease 2:575
- hereditary spastic paraplegia (HSP) 2:453,
 2:479
- Hermansky–Pudlak syndrome (HPS)
 1:775–776, 2:227–228, 2:237–239,
 2:241
- Herpes simplex virus (HSV) 3:719
 HSV-1 2:811
 nuclear exit and maturation 2:816F
- herpes virus, PML-NB by 2:826
- herpes virus entry mediator (HVEM) 3:62,
 3:771–772
- herpes virus-specific ubiquitin-specific
 protease (HAUSP) *see* ubiquitin-
 specific protease 7 (USP7)
- HES *see* Hairy/Enhancer of Split (HES)
- hESC *see* human embryonic stem cell
 (hESC)
- heterochromatin 3:469, 3:470–471
- heterochromatin binding protein-1 (HP-1)
 2:315, 3:469–470
- heterochromatin methyl-binding protein 1
 (MECP1) 2:315
- heterodimer 1:418, 1:732, 3:329–330, 3:332
- heterodimerization domain (HD) 3:107
- heteroduplex DNA (hDNA) 1:458
- heteroduplex rejection 1:464
- heterogeneous datasets, network modeling
 of 4:211–217
 network algorithms
 complex, cancer-related genetic
 networks 4:214
 integrating single data source with an
 existing interaction network
 4:212–213
 metabolic networks 4:213–214
 methods for integration of data types
 4:214–216
 proteins and genes 4:216
 network methods 4:212T
- heterogeneous multiscale method (HMM)
 4:83
- heterogeneous nuclear ribonucleoproteins
 (hnRNPs) 1:355, 1:495
- heterogeneous nuclear RNA (hnRNA)
 1:495
- heterotetrameric adaptor complexes
 2:450–452, 2:452F
 AP1 2:451–452
 AP2 2:452
 AP3 2:452–453
 AP4 2:453
 AP5 2:453
 GGAs 2:453–454
 monomeric adaptors 2:453F
- heterotrimeric G proteins 2:446–447
- heterotypic fission process 4:61
- HEVs *see* high endothelial venules (HEVs)
- hexokinase 1 (HK1) 3:347–348
- hexokinase 2 (Hk-2) 1:239T
- HGF *see* hepatocyte growth factor (HGF)
- hGH *see* human growth hormone (hGH)
- hGHR *see* human growth hormone receptor
 (hGHR)
- HGNC *see* Human Gene Nomenclature
 Committee (HGNC)
- HGPS *see* Hutchinson–Gilford progeria
 syndrome (HGPS)
- Hgs/Hrs *see* hepatocyte growth factor-
 regulated tyrosine kinase substrate
 (Hgs/Hrs)
- HGT *see* horizontal gene transfer (HGT)
- HH signaling pathway *see* hedgehog (HH)
 signaling pathway
- HhH glycosylases *see* helix-hairpin-helix
 glycosylases (HhH glycosylases)
- HHT *see* Hereditary Hemorrhagic
 Telangiectasia (HHT)
- hidden Markov model (HMM) 1:95,
 2:126
- HIF *see* hypoxia-inducible factor (HIF)
- HIF1 α *see* hypoxia-induced transcription
 factor 1 α (HIF1 α)
- Hif1 α *see* hypoxia-inducible factor1 α
 (Hif1 α)
- high aperture objective-based TIR 2:65
- high complexity screening (HCxS) 4:242
- high content screening (HCS) 4:141
- high endothelial venules (HEVs) 3:616,
 3:635–636, 3:765

- high pressure liquid chromatography (HPLC) 1:570, 1:572F, 1:574–575, 1:574F
- high speed photoactivation localization microscopy (hsPALM) 3:661
- high-affinity antibodies 1:143
- high-content analysis (HCA) 4:239–240
data storage 4:242
feature extraction 4:240–242
image analysis 4:240
image processing workflow 4:241F
noise reduction 4:240
segmentation 4:240
statistical analysis 4:242
workflow 4:240
- high-content screening (HCS) 4:234
cell models in 4:236T
experimental setup 4:235
assays 4:237–238
cells 4:235–236
HCA 4:239–240
instrumentation 4:238–239
perturbations 4:236–237
steps in 4:235F
high-content imaging platforms 4:239T
high-content or low-content 4:242
milestones in 4:235F
at Translational Research Resource Centre 4:238F
- high-density lipoprotein (HDL) 1:173–174, 2:533, 2:418–419, 3:735, 4:113
see also low-density lipoprotein (LDL)
- higher collisional dissociation (HCD) 1:568, 1:572
- higher energy collisional dissociation (HCD) 4:180
- higher eukaryotes and prokaryotes
nucleotide-binding (HEPN) 2:286–287
- highly active antiretroviral therapy (HAART) 2:837
- highly diverse intermediate filament proteins 2:576–577
- high-mobility group box 1 (HMGB-1) 3:752
- high-molecular weight kininogen (HMWK) 1:699
- high-performance liquid chromatography (HPLC) 1:76
- high-pressure-freezing (HPF) 2:32
- high-resolution time-lapse imaging 2:638
- high-speed LM 2:123–124
- high-throughput analyses 4:212
- high-throughput DNA sequencing
technology 4:161
- high-throughput next-generation sequencing 3:785–786
- high-throughput screening (HTS) 1:136–137, 1:137–138
- high-throughput transcriptomic
technologies evolution 4:160
 α -hemolysin 4:162
fluorescent *in situ* hybridization 4:162
high-throughput DNA sequencing
technology 4:161
next-generation sequencing platforms 4:161
- PCR 4:162
RNA-seq 4:160–161, 4:161
SMRT 4:161
- high-value clinical applications 1:145
pathogen neutralization, antiviral therapy, and vaccines 1:145
T-cell targeting and immune checkpoint inhibitors 1:145
- HILIC *see* hydrophilic interaction liquid chromatography (HILIC)
- Hill coefficient 4:44
- HIN *see* hematopoietic IFN-inducible nuclear protein (HIN)
- hinge domain 3:58
- Hip family *see* Huntingtin-interaction protein family (Hip family)
- hippo kinase cassette 3:99, 3:101–102
- hippo pathway 3:99–106, 3:99, 3:99F
see also Notch signaling pathway
in *D. melanogaster* 3:99, 3:100F
upstream regulators 3:99–101
deregulation in cancer 3:104
in mammals 3:101–102, 3:103F
upstream regulators 3:102–103
regulation of organ size, stem cell self-renewal, and tissue regeneration 3:103–104
- Hippo signaling pathway 3:2
- histidine-rich glycoprotein (HRG) 1:706, 1:707
- histone 2A (H2A) 1:582
- histone 3 (H3) 3:188
- histone 3 lysine27 tri-methylation (H3K27me3) 3:786
- histone 3 lysine4 tri-methylation (H3K4me) 3:786
- histone acetyltransferases (HATs) 3:94–95, 3:778, 3:785–786
- histone deacetylase complexes (HDAC) 3:453, 3:507
- histone deacetylases (HDAC) 1:86, 3:44, 3:237, 3:778
HDAC5 1:676
HDAC6 2:584, 2:761
- histone H3 at residue K4 (H3K4me3) 3:555–556, 3:778, 3:782
- histone hyperacetylation 3:786
- histone methyltransferases (HMTs) 1:483, 3:94F, 3:95, 3:258, 3:786
- histone modifications 1:492F, 3:785, 3:786–787, 3:788, 3:790–791
- hit, defined 1:135
- HIV Rev binding protein (Hrb protein) 2:463
- HIV *see* human immunodeficiency virus (HIV)
- HIV-1 *see* human immunodeficiency virus (HIV-1)
- HIV-Tat 2:525–526
- HJs *see* Holliday junctions (HJs)
- HK1 *see* hexokinase 1 (HK1)
- HL-60 cells, chemotaxis of 3:672
- HLA *see* human leukocyte antigen (HLA)
- HMG *see* 3-hydroxy-3-methylglutaryl (HMG)
- HMGB-1 *see* high-mobility group box 1 (HMGB-1)
- HMG-CoA reductase (HMGCR) 1:174–175
- HMGCR *see* HMG-CoA reductase (HMGCR)
- HMGCS *see* 3-hydroxy-3-methylglutaryl-CoA synthase (HMGCS)
- HMM *see* heterogeneous multiscale method (HMM); hidden Markov model (HMM)
- HMPS *see* Hereditary mixed polyposis syndrome (HMPS)
- HMTs *see* histone methyltransferases (HMTs)
- HMWK *see* high-molecular weight kininogen (HMWK)
- H-NOX *see* heme-nitric oxide (H-NOX)
- hnRNA *see* heterogeneous nuclear RNA (hnRNA)
- hnRNPs *see* heterogeneous nuclear ribonucleoproteins (hnRNPs)
- HOCl *see* hypochlorous acid (HOCl)
- Hodgkin–Huxley model 4:86
- holding pattern 2:363–364
- Holliday junctions (HJs) 1:461
- homelessness *see* aniois
- homeostasis 2:187–188, 2:373
cell migration in 2:727–728
in intestinal crypt 4:76–77, 4:76F
and regulation 3:745
- homing 3:624
CXCL12 to BM 3:624–625
- Homo sapiens* spliceosome 1:497T
- homodimer 1:418, 1:735, 3:607
- homologous recombination (HR) 1:2–3, 1:451, 1:456, 1:514–515, 3:558–559
- homology modeling 1:124
- homothorax (Hth) 3:99
- homotypic fusion and vacuole protein sorting (HOPS) 2:196, 2:201–202, 2:410–411, 2:412, 2:485, 2:488, 2:504, 2:755
- honorary allosteric enzyme 1:9
- HOP2-MND1 1:460
- HOPS *see* homotypic fusion and vacuole protein sorting (HOPS)
- horizontal gene transfer (HGT) 2:588
- horizontal integration 4:3, 4:266
- hormones 2:376
- host cell proteins 2:816
- host defences
against HIV 2:835–836
humoral and cellular immunity 2:835–836
restriction factors 2:836
interactions 2:850–851, 4:195
against pathogenic bacteria 2:585
- host intracellular trafficking, intracellular
lifestyle and manipulation of 2:788–790
- HotNet 4:212T, 4:214
- house-keeping functions 2:382
- HP *see* heparin (HP)
- HP-1 *see* heterochromatin binding protein-1 (HP-1)
- HPC *see* hematopoietic progenitor cell (HPC)
- HPF *see* high-pressure-freezing (HPF)
- H-phosphonate 1:40–41, 1:42F

- HPK1 *see* hematopoietic progenitor kinase 1 (HPK1)
- HPLC *see* high pressure liquid chromatography (HPLC); high-performance liquid chromatography (HPLC)
- hPOC5 *see* proteome of centriole 5 (hPOC5)
- HPS *see* Hermansky–Pudlak syndrome (HPS)
- HPV *see* human papilloma virus (HPV)
- HR *see* homologous recombination (HR)
- HR1 *see* heptad repeat (HR1)
- Hrb protein *see* HIV Rev binding protein (Hrb protein)
- HRE *see* hypoxia response element (HRE)
- HRG *see* histidine-rich glycoprotein (HRG)
- HRs *see* halorhodopsins (HRs)
- Hrs *see* hepatocyte growth regulated tyrosine kinase substrate (Hrs)
- HS *see* heparan sulfate (HS); heparin sulfate (HS)
- HS proteoglycans (HSPGs) 1:271
- HS1 *see* hematopoietic lineage cell-specific protein1 (HS1)
- HSC *see* hematopoietic stem cell (HSC); hemopoietic stem cell (HSC)
- Hsc70 1:536F, 1:537T
- HSP *see* hereditary spastic paraplegia (HSP)
- Hsp40 1:537T, 1:540, 1:542, 1:543
- HSP70 2:162
- Hsp 70 *see* heat-shock protein 70 (Hsp 70)
- hsPALM *see* high speed photoactivation localization microscopy (hsPALM)
- HSPGs *see* heparan sulfate proteoglycans (HSPGs); HS proteoglycans (HSPGs)
- HSV *see* Herpes simplex virus (HSV)
- 5-HT (5-hydroxytryptamine) 1:246, 1:248, 1:249
- Hth *see* homothorax (Hth)
- HTLV *see* human T-cell leukemia virus (HTLV)
- HTLV-1 *see* human T-cell leukemia virus (HTLV-1)
- HTS *see* high-throughput screening (HTS)
- Htt *see* Huntingtin (Htt)
- Htt associated protein (HAP40) 2:413–414
- HTT-associated protein 1 (HAP1) 3:234
- HU *see* hydroxyurea (HU)
- hub nodes 4:266
- HUGO *see* Human Genome Organisation (HUGO)
- human airway trypsin-like (HAT) 1:654, 1:654–655
- human atrial cells 4:87–88, 4:88
- human calpains structure 1:671F
- human cancer
see also cancer
 in human developmental diseases 3:417
 in diabetes 3:419–420
 in neurodegenerative diseases 3:419
 RAS-ERK and PI3K-AKT dysregulation
 developmental syndromes 3:417–418
 oncogenes and tumor suppressor genes
 and 3:411–412
 extracellular cues transmission 3:413
 functional roles 3:413
 identification 3:412
 interplay between 3:413–415
 intracellular signaling pathways
 3:414F
 PI3K pathway 3:413
 RAS pathway 3:413
 transcriptional regulation 3:413
 human cationic antimicrobial protein-18
 (hCAP-18) 3:751
 human disease 4:194–195
 HH pathway
 birth defects with 3:82T
 congenital disease and cancer 3:81–83
 targeting with drugs 3:83
 JAK-STAT-SOCS and
 JAK proteins in 3:145–146
 SOCS proteins in 3:146–147
 STAT proteins in 3:146
 microtubules in 2:545
 oncogenes and tumor suppressor genes in
 3:417
 in diabetes 3:419–420
 in neurodegenerative diseases 3:419
 RAS-ERK and PI3K-AKT dysregulation
 developmental syndromes 3:417–418
 human disorders, PI3K/Akt/mTOR signaling
 3:133–134
 human double minute 2 (HDM2) 3:270
 human ESCs (hESCs) 3:785
 Human Gene Nomenclature Committee
 (HGNC) 4:247
 Human Genome Organisation (HUGO)
 4:247
 Human Genome Project 4:153, 4:264
 human genome sequencing 4:154
 human growth hormone (hGH) 3:23
 hGH signaling 3:23F
 human growth hormone receptor (hGHR)
 3:23, 3:23–24
 human heart, computer modeling of
 4:87–88
 human immune diseases, NF- κ B
 misregulation in 3:584–585
 human immunodeficiency virus (HIV)
 1:739, 2:225–234, 2:828–838,
 3:567, 3:691, 3:308, 3:722
see also acquired immunodeficiency
 syndrome (AIDS)
 cell killing 2:835
 drugs interference with HIV replication
 2:836–837
 entry
 nuclear entry and integration 2:835
 receptor dependence and specificity
 2:833
 reverse transcription 2:833–835
 synthesis of new viral components
 2:835
 host defences against 2:835–836
 humoral and cellular immunity
 2:835–836
 restriction factors 2:836
 protease inhibitors 1:739
 chemical formulas of FDA-approved
 inhibitors 1:742F
 emergence of resistance to clinical
 inhibitors 1:741–743
 first generation 1:741
 in future 1:744
 retroviral proteases inhibitors
 1:739–741
 structure and enzymatic mechanism
 1:739
 replication cycle 2:828–829, 2:830F,
 2:837F
 structure and genomic organization
 2:829F
 virus assembly
 Env trafficking 2:832
 Gag coordinates HIV assembly
 2:829–832
 HIV envelope-induced membrane
 fusion process 2:834F
 HIV release and maturation 2:832–833
 multiple functions of Gag 2:831F
 virus transmission 2:833
 human immunodeficiency virus (HIV-1)
 1:136, 1:145, 1:327, 1:521,
 1:522–523, 1:681, 2:466, 2:809,
 2:828, 3:585
 packaging complex 1:327, 1:327F
 human leukocyte antigen (HLA) 1:614–615
 human member 9 of the solute carrier
 family 38 (SLC38A9) 1:766–767
 human microbiome 2:799
 human myosin 2:598F
 human papilloma virus (HPV) 2:482
 human prion diseases 2:840
 CJD 2:840
 FFI 2:840–841
 GSS 2:840
 Human Proteome Organization Proteomics
 Standards Initiative Molecular
 Interaction format (HUPO-PSI-MI
 format) 4:190
 human T-cell leukemia virus (HTLV) 1:739
 HTLV-1 1:521, 3:585
 human transglutaminase 2 1:123
 human ventricular cells 4:87–88, 4:88
 human virome 2:800
 human/yeast nucleotide excision repair
 proteins 1:438T
 humoral immunity 2:835–836
 Hunter syndrome 1:778
 Huntingtin (Htt) 2:413–414
 protein 1:116, 3:234
 Huntingtin-interaction protein family (Hip
 family) 2:386–387
 Huntington's disease (HD) 1:115, 1:116,
 1:369, 1:770–771, 1:796, 3:133
 HUPO-PSI-MI format *see* Human Proteome
 Organization Proteomics Standards
 Initiative Molecular Interaction
 format (HUPO-PSI-MI format)
 Hurler syndrome 1:769, 1:773, 1:777
 Hutchinson–Gilford progeria syndrome
 (HGPS) 2:313
 HVEM *see* herpes virus entry mediator
 (HVEM)
 hVPS34 *see* PtdIns3 kinase class 3
 (PtdIns3KC3)

- hyaluronan (HA) 1:279–287, 1:635
 cable structure and phagocytosis 1:284, 1:284F
 in cancer 1:285
 -dependent glycocalyxes 1:281F
 in diabetic pathologies 1:285
 in fertilization and embryogenesis 1:282
 glycocalyxes, catabolism of 1:281F
 in inflammatory pathologies 1:282–285
 interaction with CD44 1:285, 1:286F
 link module 1:282
 primary sources 1:279
 in structural matrices 1:282
 structure and cellular metabolism 1:279–282
 synthesis, models for 1:280F
 systemic clearance, by HARE/Stab2 1:283F
 whole-body turnover of 1:281–282
 hyaluronan synthase (HAS) 1:279–280, 1:282
 HAS2 1:280F, 1:282
 hyaluronic acid (HA) 2:700
 hyaluronidase 1:280–281, 1:282–284
 hybrid modeling 4:80, 4:80–81, 4:81
 composite hybrid ABM 4:81–82
 hybrid techniques 4:81
 hybridization 1:47, 1:49–50, 1:50
 hybridoma technology 1:141
 hydrated ions, schematic structure of 1:219–220F
 hydrazinolysis 1:77
 hydrogen peroxide (H₂O₂) 1:125, 2:767, 3:702
 metabolism 2:250–251
 hydrolases 2:228
 hydrophilic interaction liquid chromatography (HILIC) 1:72, 4:179–180, 4:204–205
 hydrophobic effect 1:109
 2'-hydroxyl protection 1:42–43, 1:42F, 1:43, 1:43F, 1:44F, 1:45F, 1:46F
 3-hydroxy-3-methylglutaryl (HMG) 1:174–175
 3-hydroxy-3-methylglutaryl-CoA (HMGCoA) reductase 1:294
 3-hydroxy-3-methylglutaryl-CoA synthase (HMGCS) 1:174–175
 4-hydroxyDHCer 1:187
 hydroxyl radical (OH[•]) 3:702
 hydroxylation 1:87
 hydroxyurea (HU) 1:462
 hyper-adhesion 2:708
 hypercholesterolemia 1:177–178, 2:418, 2:421, 2:428
 hyperglycemia, chronic 4:77–78
 hypermannosylation 1:58
 hyperoxaluria type 1 2:271
 hyperphosphorylation 1:675, 3:5
 hyperpolarization 1:218, 1:229–230
 hypertension 1:746, 1:749, 1:750
 hypertrophic scars 3:220
 hypochlorous acid (HOCl) 2:767, 3:702
 HYPONASTIC LEAVES 1 (HYL1) 1:360
 hypoxia 3:3, 3:369F, 3:809
 hypoxia response element (HRE) 4:109–110
 hypoxia-induced transcription factor 1alpha (HIF1 α) 2:220–221
 hypoxia-inducible factor (HIF) 3:809, 4:109–110
 hypoxia-inducible factor1 α (Hif1 α) 3:801, 3:813–814
 in regulating HSC biology 3:814F
 Hypoxic induction of angiogenesis 4:109–110
- I**
- I band 2:677
 I/D polymorphism *see* insertion/deletion polymorphism (I/D polymorphism)
 i-AAA protease *see* IMS AAA-protease (i-AAA protease)
 IAD *see* inner-arm dyneins (IAD)
 IAP *see* inhibitor of apoptosis protein (IAP); inhibitors of apoptosis (IAPs)
 IAP-binding motifs (IBMs) 1:708, 3:390
 IAP-like protein 2 (ILP-2) 1:707–708
 iatrogenic CJD (iCJD) 2:840
 IAV *see* influenza A virus (IAV)
 iBALT *see* inducible bronchus-associated lymphoid tissue (iBALT)
 I-BAR proteins 2:491, 2:497
 domain-containing proteins 2:494F
 IBD *see* inflammatory bowel disease (IBD)
 IBMs *see* IAP-binding motifs (IBMs)
 ibrutinib 1:131
 IC *see* intermediate chains (IC); intermediate compartment (IC)
 ICAD cleavage *see* inhibitor of caspase activated DNase (ICAD) cleavage
 ICAMs *see* intercellular cell adhesion molecules (ICAMs)
 ICAT *see* isotope-coded affinity tags (ICAT)
 ICB *see* intercellular bridge (ICB)
 ICE *see* interleukin 1 converting enzyme (ICE)
 iceberg model 1:10–11
 I-cell disease 1:773–774, 1:777–778
see also Inclusion-cell disease (I-cell disease)
 action myoclonus-renal failure syndrome (AMRF) 1:773–774
 ICG *see* International Cancer Consortium (ICG)
 iCJD *see* iatrogenic CJD (iCJD)
 ICL *see* interstrand cross-links (ICL)
 ICM *see* inner cell mass (ICM)
 ICMT *see* isoprenylcysteine carboxyl methyltransferase (ICMT)
 ICOS *see* inducible co-stimulator (ICOS)
 ICOS ligand (ICOSL) 3:541
 icosahedron 1:517, 1:520
 ICOSL *see* ICOS ligand (ICOSL)
 ICS *see* image correlation spectroscopy (ICS)
 icythosis lamellaris 1:772
 IDAs *see* inner dynein arms (IDAs)
 IDDAs *see* intrinsically disordered domains (IDDs)
- ideal stem cell 3:794–795
 idFBA *see* integrated dynamic flux-balance analysis (idFBA)
 IDO *see* indoleamine 2,3 dioxygenase (IDO)
 IdoA *see* iduronic acid (IdoA)
 iduronic acid (IdoA) 1:271
 idursulfase 1:782
 IEF *see* isoelectric focusing (IEF)
 IEX chromatography *see* ion-exchange chromatography (IEX chromatography)
 iFBA *see* integrated flux-balance analysis (iFBA)
 iFGFs *see* intracellular FGFs (iFGFs)
 IFI16 *see* IFN- γ inducible protein16 (IFI16)
 IFN *see* interferon (IFN)
 IFN gamma receptor (IFNGR) 3:30
 IFN lambda receptor (IFNLRL) 3:30
 IFNGR *see* IFN gamma receptor (IFNGR)
 IFNLRL *see* IFN lambda receptor (IFNLRL)
 IFN-stimulated gene 3 (ISGF3) 3:143
 IFN-stimulated genes (ISGs) 3:720–721
 IFN- γ inducible protein16 (IFI16) 3:719–720
 IFs *see* intermediate filaments (IFs)
 IFT *see* intraflagellar transport (IFT)
 Ig *see* immunoglobulin (Ig)
 Ig heavy (IgH) 3:552
 IGC *see* immature granulosa cells (IGC)
 IGCEERSFPNIIIIIG peptide 1:684
 IGF *see* insulin-like growth factor (IGF)
 IGF-1 *see* insulin-related growth factor-1 (IGF-1)
 IGF-BP *see* insulin-like growth factor binding protein (IGF-BP)
 IgH *see* Ig heavy (IgH)
 Ig-like *see* immunoglobulin-like (Ig-like)
 iGluRs *see* ionotropic glutamate receptors (iGluRs)
 IGR *see* intergenic region (IGR)
 IGS *see* intergenic spacer (IGS)
 IHC *see* immunohistochemistry (IHC)
 iHeps *see* iPS-derived hepatocytes (iHeps)
 IHH *see* Indian HH (IHH)
 IKK *see* inhibitor of kappa B kinase (IKK)
 IL *see* interleukin (IL)
 IL-1 receptor (IL-1R) 3:28
 IL-1R1 3:28
 IL-1R2 3:28
 IL-1R3 3:28–29
 IL-1R3b 3:29
 therapeutic targeting 3:29
 IL-1R *see* IL-1 receptor (IL-1R); interleukin 1 receptor (IL-1R)
 IL-1R-associated kinases (IRAKs) 3:712–714
 IL-2 inducible T cell kinase (ITK) 3:534–535
 IL-3R α *see* α -subunit of IL-3 receptor (IL-3R α)
 IL-10/IL-22 receptors 3:29–30
 ILCs *see* innate lymphoid cells (ILCs)
 ILK *see* integrin-linked kinase (ILK)
 ILP-2 *see* IAP-like protein 2 (ILP-2)
 ILPs *see* invadopodia-like protrusions (ILP); invadosome-like protrusions (ILPs)

- ILV biogenesis *see* intraluminal vesicle (ILV) biogenesis
- ILVs *see* intraluminal vesicles (ILVs)
- IMAC *see* immobilized metal affinity chromatography (IMAC)
- image correlation spectroscopy (ICS) 2:103
- image drift 2:90
- image preprocessing 2:89
- image-guided surgery 1:145–146
- ImageJ 4:240
- imaginal discs 2:741, 2:745
- imaging artifacts, SIM 2:90
- image drift 2:90
- low signal to noise 2:90
- pixel negative values 2:90–91
- refractive index mismatch 2:90
- imaging techniques using IVM 2:129–131
- Imatinib 4:99
- IMM *see* inner mitochondrial membrane (IMM)
- immature granule cells (IGC) 3:799
- immobilized metal affinity chromatography (IMAC) 1:73, 4:178
- immune cell migration, interrogating 3:670–671F
- immune cell receptors 3:153, 3:156
- immune checkpoint inhibitors 1:145
- immune deficiency disorders 3:563–564
- immune diseases, NF- κ B therapy in 3:585–586
- immune effectors 2:846–847, 2:847T
- ABFs 2:847
- C-type lectins 2:848
- caenacins 2:847
- caenopores 2:847–848
- CUB domain proteins 2:848
- lysozymes 2:848
- NLPs 2:848
- immune exhaustion 3:570
- immune response 2:138–140, 3:210, 4:118
- stromal cells regulation
- in induced TLOs 3:619
- inflammatory diseases 3:620–621
- initiation of tissue inflammation 3:619–620
- targeting stroma 3:621
- trafficking during T lymphocytes 3:636
- immune signaling pathways 2:848–849
- DAF-2/DAF-16 pathway 2:849–850
- DBL-1 pathway 2:850
- p38 MAPK pathway and related pathways 2:848–849
- response pathways to intracellular pathogen infections 2:850
- unfolded protein response pathway 2:850
- immune surveillance 3:632
- trafficking during T lymphocytes 3:635–636
- immune system 3:667
- immune thrombocytopenic purpura (ITP) 1:171
- immunity, functional cell biology of 3:6
- immunoaffinity chromatography 1:72
- immunoelectron microscopy 2:33F
- CLEM 2:40–41
- immunogold label quantification and labeling specificity 2:41–42
- immunolabeling strategies 2:36
- freeze-fracture replica labeling 2:39–40
- on-section labeling 2:36–37, 2:38F
- preembedding labeling 2:36–37
- whole mount labeling 2:37–39, 2:38F
- sample preparation strategies 2:32, 2:33F
- chemical fixation 2:32–34
- cryo-immobilization and freeze-substitution 2:34–36
- plastic embedding or cryosectioning 2:34
- immunoglobulin (Ig) 2:704, 3:581
- IgA 3:690
- IgG 1:138–139, 2:752, 3:696
- Fc γ R activation by 3:696–697
- Ig-like 4:104–105
- immunogold label quantification and labeling specificity 2:41–42
- immunohistochemistry (IHC) 1:141, 3:442
- immunolabeling strategies 2:36
- freeze-fracture replica labeling 2:39–40
- on-section labeling 2:36–37, 2:38F
- preembedding labeling 2:36–37
- whole mount labeling 2:37–39, 2:38F
- immunological synapses (ISs) 3:6, 3:522–525, 3:523, 3:588–589, 3:636, 3:677, 3:695
- diversity 3:660
- forces acting at synapses 3:660–661
- forces on receptor–ligand unbinding 3:664F
- FRET 3:663F
- imaging TCR–pMHC interactions *in situ* 3:661–664
- measuring binding *in situ* 3:661
- membranes 3:661
- pulling forces 3:664–665
- 2D-binding constants micropipette and bioforce probe measurements 3:662F
- immunoproteasome 1:591, 1:592, 3:591
- and alternative CP subunits 1:591
- immunoreceptor tyrosine-based activation motifs (ITAMs) 2:752–753, 3:529, 3:650, 3:696, 3:710–712
- immunoreceptor tyrosine-based inhibitory motifs (ITIMs) 3:701, 3:710–712
- immunotherapy, retrograde transport and 2:438–439
- importin β superfamily 3:172–174, 3:173T
- directionality of translocation 3:174
- importin3 3:126
- importin7 3:126
- importin9 3:126
- importins 3:143, 3:177
- importin- α/β 3:126
- IMS *see* intermembrane space (IMS); ion-mobility spectrometry (IMS)
- IMS AAA-protease (i-AAA protease) 2:282–283
- IMS-RP *see* intermembrane space reporter protein (IMS-RP)
- IN *see* integrase (IN)
- in situ*, measuring binding 3:661
- in vitro* SV40 DNA replication system 1:410
- in vitro* transcription (IVT) 1:484–485, 4:220
- in vivo* dynamics and behavior 1:484
- in vivo* imaging 2:138
- imaging tissue architecture and function
- in vivo* 2:131T, 2:132–133F, 2:133–138
- techniques 4:78
- inactive Rhomboid-like protein 2 (iRhom2) 1:642
- inborn immune system *see* innate immune system
- INCENPs *see* inner centromere proteins (INCENPs)
- incision complex, assembly of 1:435, 1:439
- XPF-ERCC1 1:439
- XPG 1:439
- inclusion cells (I-cells) 1:773
- inclusion-cell disease (I-cell disease) 2:408
- incoherent feedforward scheme, adaptation through 4:7
- incomplete coverage 4:246
- inconsistent nomenclature 4:245
- lncRNA MALAT-1 1:369–370
- Indian HH (IHH) 3:76
- individual cell migration *see* single-cell migration
- indoleamine 2,3 dioxygenase (IDO) 3:763
- induced pluripotent stem cells (iPSCs) 1:797, 3:6–7, 3:785, 3:816–817, 3:819, 4:138–139, 4:236
- inducible bronchus-associated lymphoid tissue (iBALT) 3:619
- inducible co-stimulator (ICOS) 3:541
- inducible nitric oxide synthase (iNOS) 2:767, 2:767–768, 3:61
- induction, autophagy 3:404–405
- INF2 *see* inverted formin (INF2)
- infancy, changes in gut microbiome in 2:801–802
- infantile Gaucher disease 1:772
- infectious agents *see* virus(es)
- infectious disease, UPR role in 2:534
- inferring protein sequences from nucleic acid sequences 1:80
- infertility 2:585
- inflammasomes 1:708–709, 2:523–524, 2:766, 2:772–773, 3:368, 3:368, 3:710
- inflammation 2:728, 2:753–754, 2:766, 4:95
- CXCL12 in balancing granulopoiesis and lymphopoiesis 3:629
- mathematical approaches to studying 4:95–101
- current mathematical and computational models 4:98
- early events, leukocyte responses to bacteria 4:98
- future challenges 4:100
- immune response 4:99–100
- inflammatory cytokines and TNF-therapy 4:98–99
- resolution of inflammation, capturing 4:96–98, 4:97F
- tissue damage and wound healing 4:99

- meprins in 1:694–695
 response to stimuli 4:96F
 spatial model 4:100
 systems-biology approach 4:95–96
 inflammatory bowel disease (IBD)
 1:735–736, 2:805, 2:806, 3:690
 inflammatory caspases 3:364, 3:367,
 3:367–368, 3:368, 3:371
 see also inflammasomes
 inflammatory cells 3:218
 inflammatory diseases 3:620–621
 SRs in 3:734–735
 TNF in 3:360–361
 UPR role 2:534
 inflammatory response 2:791
 inflammatory system 3:624
 influenza A virus (IAV) 1:655–656, 2:809
 nuclear exit and maturation 2:816F
 influenza virus 1:520, 1:521
 information theory 4:258
 infrared light (IR light) 2:129
 infrared matrix-assisted laser desorption/
 ionization orthogonal time-of-flight
 mass spectrometry (IR-MALDI-o-TOF
 MS) 1:155
 inherited connective tissue disorders
 1:630
 inhibitor of apoptosis protein (IAP) 3:69,
 3:342, 3:358, 3:469–470, 4:221
 proteins 3:391
 Apollon 3:395
 c-IAP1 3:391–393
 c-IAP2 3:391–393
 ILP-2 3:395
 ML-IAP 3:394
 NAIP 3:394
 and NF- κ B pathways 3:392–393
 survivin 3:394–395
 XIAP 3:391
 structural organization 3:390
 BIR domain 3:390
 CARD domains 3:391
 inhibitor of apoptosis protein family
 members 3:391F
 LRR domains 3:391
 NATCH domains 3:391
 RING domains 3:390–391
 UBA domains 3:390–391
 UBC domains 3:390–391
 targeting IAP proteins 3:395
 antagonists 3:395–396
 survivin antisense oligonucleotides
 3:396
 inhibitor of caspase activated DNase
 (ICAD) cleavage 3:367–368
 inhibitor of kappa B (I κ B) 3:70
 inhibitor of kappa B kinase (IKK) 2:534,
 3:70, 3:716
 IKK γ 3:716
 inhibitors of apoptosis (IAPs) 1:707–708,
 1:708, 3:329, 3:331, 3:370
 biological functions 1:708
 cell migration, invasion, metastasis
 regulation, and development 1:709
 inflammation and immunity regulation
 1:708–709
 in cell division and cytokinesis 1:709
 in disease 1:709
 structure 1:708
 X-linked 3:331
 inhibitors of cathepsin E 1:684
 inhibitors of macropinocytosis 2:762
 inhibitory postsynaptic potentials (IPSP)
 1:244, 1:246
 inhibitory SMADs (I-SMADs) 3:43
 in-house modeling studies 4:82
 initiation codon 1:310, 1:546
 initiator tRNA 1:384, 1:385–387, 1:385F,
 1:388, 3:307
 interaction with ribosome 1:387–388
 INK4 family
 functions independent of CDK inhibition
 3:438
 normal function 3:437
 regulation of p15^{INK4b} and p16^{INK4a} levels
 3:437–438
 INK4A 3:5
 INK4a/ARF locus 3:447, 3:449F
 in aging 3:454–455
 ARF 3:452–453
 dual coding locus in mammals 3:448–449
 frequency of genetic alterations at
 CDKN2A locus 3:450F
 G1–S phase progression in mammals
 3:447–448
 INK4a–RB1 anti-proliferative pathway
 3:447F
 p16INK4a-mediated senescence 3:448
 phenotypes of ARF, INK4a, and INK4a/
 ARF deficient 3:451T
 RB1-mediated senescence 3:448
 regulation and loss in cancer 3:453–454
 tumor suppressors 3:449–452
 INK4a locus 3:437, 3:438
 INK4B 3:5
 INK4b locus 3:437, 3:438
 inkjet printing 1:47
 InlA *see* internalin A (InlA)
 InlB *see* internalin B (InlB)
 INM *see* inner nuclear membrane (INM)
 innate and adaptive immune receptors 3:6
 innate immune system 3:1, 3:519–520,
 3:710
 ALRs 3:719–720
 classes, structure, and activation of innate
 sensors
 CLRs 3:710–712, 3:711T
 NLRs 3:714–718, 3:715–716T
 RLRs 3:718–719
 TLRs 3:712–714, 3:712T
 combinatorial sensing and PRR crosstalk
 3:721–722
 cytoplasmic innate sensing and signaling
 mechanism 3:713–714F
 cytosolic sensing pathways multiscale
 regulation 3:722–724
 human nucleic acid sensors 3:718T
 intracellular DNA sensors 3:719–720
 intracellular organelles 3:720–721
 sensors 3:711F
 spatial relocation 3:720–721
 systems-level overview 3:723F
 innate immunity 1:584T, 2:846, 3:6, 3:368,
 3:371
 DCs in
 innate lymphocytes activation
 3:745–746
 innate pDCs 3:745
 elements 3:519
 NF- κ B in 3:583–584
 SRs in 3:733
 collaboration with PRR families
 3:733–734
 in inflammatory diseases 3:734–735
 as microbial entry points 3:734
 innate lymphoid cells (ILCs) 3:521–522,
 3:745–746
 inner cell mass (ICM) 3:104, 3:785
 inner centromere proteins (INCENPs)
 3:469, 3:470
 inner dynein arms (IDAs) 2:627, 2:660
 HCs 2:630
 inner membrane 1:237
 inner mitochondrial membrane (IMM)
 3:341
 inner nuclear membrane (INM) 3:480
 INM proteins 2:311
 inner-arm dyneins (IAD) 2:627
 inorganic adsorbent chromatography 1:72
 iNOS *see* inducible nitric oxide synthase
 (iNOS)
 inosine 1:310, 1:358
 inositol 1, 5-triphosphate, 4 *see* inositol-
 triphosphate (IP₃)
 inositol phospholipids 1:182–183
 inositol polyphosphate 4-phosphatase type
 II (INPP4B) 3:132
 inositol polyphosphate 5-phosphatases
 2:508
 inositol requiring enzyme (IRE) 2:529
 IRE1 2:530
 inositol-triphosphate (IP₃) 2:378, 2:447,
 2:759
 Inp1 protein 2:261
 Inpp4a *see* type-I inositol 4-phosphatase
 (Inpp4a)
 INPP4B *see* inositol polyphosphate
 4-phosphatase type II (INPP4B)
 Inpp4b *see* type-II inositol 4-phosphatase
 (Inpp4b)
 Inpp5b *see* 75-kDa 5-phosphatase (Inpp5b)
 insect expression systems 1:59
 lytic nature 1:59
 recombinant baculoviruses generation
 1:59
 traditional transfection systems 1:59–60,
 1:60
 worldwide protein data 1:59
 insensible perspiration 4:24
 insertion/deletion polymorphism (I/D
 polymorphism) 1:748–749
 inside-out response process 2:718
 insoluble polymers/aggregates formation
 1:791–792
 InsP3 *see* inositol-triphosphate (IP₃)
 instability 4:33
 instrument response function (IRF)
 2:116–117

- insulator, defined 1:488
- insulin
- action of 1:290, 1:291, 1:291F
 - in fed state 1:290
 - and metabolic processes 1:289
- insulin receptor substrate family (IRS family) 3:12
- IRS1 3:115, 3:131, 3:245
- insulin role in regulated secretion 2:378–379
- insulin-like growth factor (IGF) 3:99, 3:224, 3:413
- IGF-1 3:130–131
- insulin-like growth factor binding protein (IGF-BP) 1:702
- insulin-regulated aminopeptidase (IRAP) 1:747, 3:591
- insulin-related growth factor-1 (IGF-1) 1:702
- int1* 3:86
- integral monotopic LD proteins 2:274
- integral transmembrane (TM), affinity of 1:210
- integrase (IN) 1:739, 2:828–829
- integrated dynamic flux-balance analysis (idFBA) 4:212T, 4:214
- integrated flux-balance analysis (iFBA) 4:214
- integrative network-based systems biology approaches 4:259–260
- integrative omics-metabolic analysis (IOMA) 4:212T, 4:214
- integrative systems biology 4:245, 4:253
- Bayesian formalism 4:253–254
 - computational prediction methods 4:248
 - critical thinking in data interpretation 4:249
- data
- availability 4:247
 - integration workflow 4:246F
 - quality and heterogeneity 4:246–247
 - structure standardization 4:247–248
- GWAS 4:254
- identification of suitable resources and data interpretation 4:247
- incomplete coverage 4:246
- inconsistent nomenclature 4:245
- lack of standardization 4:245
- manual curation limitations 4:245–246
- model-based approaches 4:255–256
- nomenclature systematization of biological entities 4:247
- omics layer 4:255F
- open licenses 4:248–249
- quality assessment and benchmarking 4:248
- regulatory network 4:253
- study of cancer 4:254
- T-ALL 4:255
- text-mining approaches 4:248
- whole-genome catalog of somatic mutations 4:254F
- integrin alpha M (ITGAM) 3:672
- integrin-linked kinase (ILK) 3:129–130, 3:227, 3:335
- integrins 2:426, 2:713, 2:724
- adhesome 2:714–716, 2:716F
- transactivation by GPCR 2:725F
- interaction and binding domains 1:92
- interactive transcriptome exploration 4:162
- ceRNA 4:163
 - circular RNA 4:162–163
 - effects on transcription and mRNA 4:163
 - miRNA 4:162
 - piRNA 4:162
 - ribosomal and transfer RNA 4:162
 - RNA modifications 4:163
- interactome reverse engineering 4:256–257
- ARACNe 4:258–259
 - benchmarking predictive models 4:259
 - information theory 4:258
 - integration of prior knowledge and experimental evidence 4:257
 - MINDy algorithm 4:258–259
 - optimization driven machine learning approaches 4:257
 - probabilistic methods 4:258
 - regression analysis 4:257–258
- interactomes 4:187–198, 4:187, 4:253–254
- applications of interactome maps 4:193
 - cellular organization and human disease 4:194
 - interaction network structure 4:193–194
 - network biology 4:193F
- cellular life 4:188F
- mapping
- genome sequencing for 4:193
 - methods 4:187–191, 4:189F
 - PDI 4:191
 - PGI 4:192
 - PLIs 4:192
 - PPIs 4:188–191
 - PRIs 4:191–192
 - PSMIs 4:192–193
- intercellular adhesion complexes 2:742
- intercellular adhesion molecule-1 (ICAM-1) 3:522–524
- intercellular bridge (ICB) 2:403–404
- intercellular cell adhesion molecules (ICAMs) 2:724, 3:752
- ICAM-1 3:522–524, 3:619
- intercellular communication 3:1–2, 3:2–3, 3:667, 3:672
- at the single-cell level 3:673F
- intercisternal continuities 2:359–360
- intercisternal tubule 2:360
- interconnected membrane system 2:175
- interference reflection contrast (IRC) 2:65
- interferon (IFN) 2:836, 3:29, 3:136, 3:520, 3:617–618, 3:691–692, 3:712–714, 3:742
- IFN γ 3:770
- type I IFN receptors 3:30–31
- type II IFN receptors 3:31
- type III IFN receptors 3:31
- interferon regulatory factor (IRF) 3:710
- IRF4 3:542
- intergenic region (IGR) 3:308
- intergenic spacer (IGS) 3:257–258
- inter-kinetochore tension 3:471–472, 3:473F
- interleukin (IL) 1:694, 3:567, 3:616, 3:710
- IL-1 1:708–709
- IL-1 β 1:720, 2:520
- processing regulation 2:523–524
 - release mechanisms 2:524
 - secretion and cell death 2:524
- IL-2 3:25–26, 3:136
- IL12b* 3:721–722
- IL-4 3:26
- IL-6 1:720, 3:218
- IL-7 3:26–27
- IL-8 3:584
- IL-13 3:26
- IL-15 3:25–26
- IL-17 3:30
- interleukin 1 receptor (IL-1R) 3:712
- interleukin 1 converting enzyme (ICE) 3:329
- intermediate chains (IC) 2:620
- intermediate compartment (IC) 2:168–182, 2:630
- compositional aspects 2:169
 - COPI coats 2:170
 - p24 protein family 2:169–170
 - Rab1 and Rab2 2:170–171
 - tethers and soluble NSF attachment protein receptors 2:171
- functional aspects
- autophagy 2:177
 - phagocytosis and antigen presentation 2:177
 - posttranslational modification 2:176
 - protein maturation 2:176
 - signaling 2:176–177
 - sorting and transport 2:175–176
- historical perspective
- coronavirus budding compartment 2:169
 - 15°C compartment 2:168–169
 - p58/ERGIC-53/LMAN1 2:169
 - salvage compartment 2:169
- links to disease 2:177
- models
- interconnected membrane system 2:175
 - stationary compartment 2:175
 - TCs 2:173–175
- structure, distribution, and dynamics 2:171
- association with cytoskeleton 2:172–173
- intermediate filament cytoskeleton
- cell–cell adhesion and 2:707–708
 - integrating IFs and desmosomes 2:707–708, 2:707F
 - physiological functions of desmosomal-IF integration 2:708
- intermediate filaments (IFs) 2:4, 2:312, 2:569, 2:704, 4:134, 4:135–136
- see also* intermediate progenitor cell (IPC)
- actin; microtubules (MTs)
- assigning to different assembly groups 2:570–571
 - stages of intermediate filament assembly 2:571F
 - blister formation in keratin-deficient mouse epidermis 2:576F

- cytoplasmic intermediate filament
dysfunction effect 2:575–576
functions 2:574
contribution to organelle localization 2:575
guardians of cells in stressful situations 2:575
reorganization 2:575
specialized tissue functions 2:574–575
structural support 2:574
- highly diverse intermediate filament proteins 2:576–577
- IF-associated proteins 2:572–573
- lamin intermediate filaments 2:573
- network
cytoplasmic intermediate filament network organization 2:572–573
organization in cultured cell 2:574F
organization in murine tissues 2:572F
properties 2:571–572
protein family 2:569, 2:569T
markers of cellular differentiation 2:569–570
posttranslationally modification 2:573–574
structure model 2:570F
sharing secondary structure 2:570
- intermembrane space (IMS) 2:282–283, 3:342, 4:218, 4:221
- intermembrane space reporter protein (IMS-RP) 4:221
- intermolecular distance restraints 1:99
- internal cell structure 2:19–20
- internal N-termini 1:568, 1:574–575, 1:574F, 1:575–576, 1:575F, 1:576
- internal ribosome entry site (IRES) 1:343, 1:344, 1:529, 1:532–533, 1:533, 3:3–4, 3:275, 3:307, 4:110
IRES transacting factors 3:313
- internal ribosome entry site-mediated translation 1:529–530
canonical cap-dependent mechanism 3:307
- initiation
assay 3:307–308
discovery 3:307
viral IRESs 3:308
- internalin A (InlA) 2:785
- internalin B (InlB) 2:785
- International Cancer Consortium (ICG) 4:254
- International Society for Hyaluronan Sciences (ISHAS) 1:279
- interorganellar communication 2:3, 2:529
compartmental organization 2:3
- interstitium *see* connective tissue
- interstrand cross-links (ICL) 1:456
- intersystem crossing (ISC) 2:76
rate 2:108
- inter- α -inhibitor (I α I) 1:282
- intervening sequence (IVS) 3:291
- intestinal epithelium 3:690, 4:76–77, 4:76F
- intracellular activities 1:625
- intracellular determinants
actin filaments 2:725
actin organization at leading edge 2:727F
- microtubules 2:725
regulation and integration 2:725–726
- intracellular DNA sensors 3:719–720
- intracellular effectors of signal transduction–transmembrane receptors 3:76–77
- intracellular FGFs (iFGFs) 2:525
- intracellular infectiology 2:4
- intracellular killing 2:4
- intracellular membrane compartment organization 2:593–594
traffic 2:460–461
- intracellular organelles 3:720–721
- intracellular protein degradation 1:582
- intracellular signaling pathway 2:795
- intracellular trafficking, exotoxins altering 2:796–797
- intracellular transport 2:3, 2:3–4, 2:608, 2:814–815
- intracrine activity 4:104
- intraflagellar transport (IFT) 2:608, 2:611, 2:620, 2:660, 2:665–666
anterograde IFT motor 2:665–666
components 2:666T
dynein 2:625–626
particle polypeptides and cargo 2:666–667
retrograde IFT motor 2:666
trains 2:625–626
and vesicle transport pathways 2:665F
- intra-Golgi transport 2:185, 2:185F, 2:354
see also post-Golgi transport
cisternal maturation model 2:356–358
cisternal progenitor model 2:359
COPI-mediated bidirectional transport model 2:358
crash course in machinery 2:354–355
COPI-coated vesicles 2:354–355
SNARE complexes 2:355–356
tethers and Rabs 2:355
Golgi complex 2:355F
intercisternal continuities 2:359–360
models 2:357F
rapid partitioning model 2:360
rim progression model 2:358–359
vesicular anterograde transport model 2:356
- intraluminal membranes, composition and fate of 2:205
- intraluminal vesicles (ILVs) 2:201, 2:409–410
- intraluminal vesicle (ILV) biogenesis 2:466, 2:469, 2:469–470
- intravacuolar bacteria 2:788–790
- intravasation 3:225
- intravital microscopy (IVM) 2:55, 2:129–146
applications 2:133
fluorescent light microscopy 2:130F
to image dynamics of individual cells 2:135T
imaging individual cells in live animals 2:139F
imaging subcellular structures *in vivo* 2:140–143, 2:140T
imaging techniques using 2:129–131
- imaging tissue architecture and function *in vivo* 2:131T, 2:132–133F, 2:133–138
subcellular organelles 2:141F, 2:142F
tumor progression and invasive behavior in live animals 2:137–138F
- intrinsic apoptosis 3:342–343, 3:368–370, 3:369F, 3:370
- intrinsic apoptotic pathway 3:69–70
- intrinsically disordered domains (IDDs) 1:93
- intrinsically unstructured proteins 1:112
- intron 1:377–380, 1:381F, 1:495, 1:496, 1:496F, 1:497–498, 1:499, 1:499F, 1:500–501, 1:501, 1:503, 1:503–504, 1:504F, 1:505, 1:505F, 1:506, 1:508
architecture of 1:495–496
assembly of spliceosomes on 1:498–499
group II 1:380, 1:499–500
- intussusception 4:103–104, 4:104, 4:114
- intussusceptive angiogenesis 4:114
see also sprouting angiogenesis
invadopodia 2:713, 2:714, 3:213, 3:225
assembly 3:225–226
BM 3:225
induction 3:225
- invadopodia-like protrusions (ILPs) 3:524
- invadosome-like protrusions (ILPs) 3:637–638
arrays function in signal amplification 3:643–644
ILP-2 3:395
T lymphocyte 3:642–643
- invadosomes 2:715F
- invariant chain 3:522
- invasin 2:785
- inverse Weiner filtering and apodization 2:90
- inversin/NPHP2 compartment 2:667
- inverted formin (INF2) 2:285
- IOMA *see* integrative omics-metabolic analysis (IOMA)
- ion abrasion scanning electron microscopy (IA-SEM) *see* focused ion beam scanning electron microscopy (FIB-SEM)
- ion channel allostery
neuronal action potentials and 1:244–251
- ion channel gating 1:229
- ion channel kinetics 4:86
- ion channel selectivity, quantification of 1:231–232, 1:233F
- ion channels
optochemical approaches to controlling 2:151–153
- ion movement, electrochemical driving force for 1:221, 1:221–223, 1:235
- ion suppression 4:207
- ion transport and cystic fibrosis 1:255
- ion-exchange chromatography (IEX chromatography) 1:71
- ionizing radiation 1:442, 1:445–446
- ion-mobility spectrometry (IMS) 1:152
- ionotropic glutamate receptors (iGluRs) 2:151
- ionotropic receptors 1:244, 1:248–249

IP₃ *see* 1,4,5-tri-phosphate (IP₃)
 IP₃ *see* inositol-triphosphate (IP₃)
 IpaJ 2:790
 IPC *see* intermediate progenitor cell (IPC)
 ipilimumab (Yervoy™) 1:145
 iPSCs *see* induced pluripotent stem cells (iPSCs)
 iPS-derived hepatocytes (iHeps) 1:793–794
 IPSP *see* inhibitory postsynaptic potentials (IPSP)
 IPTG *see* isopropyl-β-D-1-thiogalactopyranoside (IPTG)
 IQ motif 2:597
 IR light *see* infrared light (IR light)
 IRAKs *see* IL-1R-associated kinases (IRAKs)
 IRAP *see* insulin-regulated aminopeptidase (IRAP)
 IRC *see* interference reflection contrast (IRC)
 IRE *see* inositol requiring enzyme (IRE); iron response element (IRE)
 IRES *see* internal ribosome entry site (IRES)
 IRF *see* instrument response function (IRF); interferon regulatory factor (IRF)
 iRhom2 *see* inactive Rhomboid-like protein 2 (iRhom2)
 iRhoms 1:642–643
 IR-MALDI-o-TOF MS *see* infrared matrix-assisted laser desorption/ionization orthogonal time-of-flight mass spectrometry (IR-MALDI-o-TOF MS)
 iron response element (IRE) 1:343, 1:344
 iron sequestration 2:771
 iron-containing type 5 purple acid phosphatase (Acp5) 1:776, 1:778
 iron-sulfur protein (ISP) 1:301
 irreversibility 4:29
 IRS family *see* insulin receptor substrate family (IRS family)
 IRS-1 *see* receptor substrate-1 (IRS-1)
 ISC *see* intersystem crossing (ISC)
 ISG15 (IFN-stimulated gene 15 kDa) 1:584T, 1:587–588
 ISGF3 *see* IFN-stimulated gene 3 (ISGF3)
 ISGs *see* IFN-stimulated genes (ISGs)
 ISHAS *see* International Society for Hyaluronan Sciences (ISHAS)
 I-SMADs *see* inhibitory SMADs (I-SMADs)
 isobaric tags for relative and absolute quantitation (iTRAQ) 4:169, 4:182
 isodesmic self-assembly 2:548, 2:552
 isoelectric focusing (IEF) 1:66, 1:571
 isoelectric point 1:66
 isolated cell analysis 4:202
 isolation membrane *see* phagophore
 isomiRs 1:368–369
 isoprenoids synthesis 2:254
 isoprenylation 2:312–313
 isoprenylcysteine carboxyl methyltransferase (ICMT) 2:312–313
 isopropyl-β-D-1-thiogalactopyranoside (IPTG) 1:57
 isotachopheresis 1:66, 1:73
 isotope-coded affinity tags (ICAT) 4:169
 isotype switching 3:581
 isozymes/isoenzymes 1:288

ISP *see* iron-sulfur protein (ISP)
 ISs *see* immunological synapses (ISs)
 ITAMs *see* immunoreceptor tyrosine-based activation motifs (ITAMs)
 ITGAM *see* integrin alpha M (ITGAM)
 ITIMs *see* immunoreceptor tyrosine-based inhibitory motifs (ITIMs)
 ITK *see* IL-2 inducible T cell kinase (ITK)
 ITP *see* immune thrombocytopenic purpura (ITP)
 iTRAQ *see* isobaric tags for relative and absolute quantitation (iTRAQ)
 Ivacaftor 1:252, 1:261
 IVM *see* intravital microscopy (IVM)
 IVS *see* intervening sequence (IVS)
 IVT *see* *in vitro* transcription (IVT)
 IκB *see* inhibitor of kappa B (IκB)

J

Jab1/MPN/Mov34 (JAMM) metalloproteases 1:586, 3:505–506
 jActiveModules 4:212, 4:212T
 jaguar mutation 4:122
 JAK pathway *see* Janus kinase pathway (JAK pathway)
 JAK-homology domains (JH domains) 3:136–140
 JAK–STAT 3:2
 JAK–STAT–SOCS signaling cascade 3:136, 3:140F
 and human disease
 JAK proteins in 3:145–146
 SOCS proteins in 3:146–147
 STAT proteins in 3:146
 JAK family 3:136–141
 SOCS proteins 3:144
 STAT proteins 3:142–144
in vivo phenotypes 3:137–139T
 JAMMs *see* JAB1/MPN/MOV34 (JAMMs)
 JAMs *see* junctional adhesion molecules (JAMs)
 Janus kinase pathway (JAK pathway) 3:136, 4:66
 activation signaling downstream 3:142
 evolution 3:136–141
 in human disease 3:145–146
 mechanism of activation 3:141–142
 Japanese Electron Optical Corporation (JEOL) 2:16–17
 JARID1B 3:809
 jasmonates 2:254
 JEOL *see* Japanese Electron Optical Corporation (JEOL)
 JH domains *see* JAK-homology domains (JH domains)
 JIP1 *see* JNK-interacting protein 1 (JIP1)
 JMJD3 *see* Jumonji domain-containing 3 (JMJD3)
 JMML *see* juvenile myelomonocytic leukemia (JMML)
 JMY protein *see* junction mediating regulatory protein (JMY protein)
 Jnk *see* jun Kinase (Jnk)

JNK *see* jun N-terminal kinase (JNK); Jun-activate kinase (JNK)
 JNK-interacting protein 1 (JIP1) 3:234–235
 JNK-interacting protein 3 (JIP3) 3:236
 Josephin domain proteases 1:586
 JPS *see* juvenile polyposis syndrome (JPS)
 Jumonji domain-containing 3 (JMJD3) 3:780
 Jun 3:574–575, 3:575
 jun Kinase (Jnk) 3:100
 jun N-terminal kinase (JNK) 2:530, 2:735, 2:849, 3:51, 3:62, 3:70, 3:274, 3:358, 3:391, 3:439, 3:712–714
 JNK3 3:234
 Jun-activate kinase (JNK) 4:145
 junction formation 2:706
 junction mediating regulatory protein (JMY protein) 2:491
 junctional actomyosin 2:706
 junctional adhesion molecules (JAMs) 2:742, 3:752, 4:111
 JAM-1 3:634
 junctional complexes 2:743F
 junctional contractility 2:706–707
 juvenile myelomonocytic leukemia (JMML) 3:119
 juvenile neuronal ceroid lipofuscinosis 1:767
 juvenile polyposis syndrome (JPS) 3:47
 juxtacrine signaling 3:802–803, 4:112

K

kallikrein (KLK) 1:699–705
 and cancer 1:702
 kallikrein-related peptidases 1:700–702, 1:702T
 kinins 1:700
 KLK-related peptidases 1:700T
 KLK–cell interactions 1:702
 plasma 1:699
 tissue 1:699–700
 kallikrein-related peptidases (KLKs) 1:691, 1:700–702, 1:702T
 kallistatin 1:699–700
 Kaposi's sarcoma herpes virus (KSHV) 3:717
 Kaposi's sarcoma-associated herpesvirus 2:836
 kappa-caseinoglycopeptide 1:733
 karyopherin 2:320
 KAYAK *see* kinase activity assay for kinome profiling (KAYAK)
 KCa3.1 *see* calcium-activated potassium channel (KCa3.1)
 75-kDa 5-phosphatase (Inpp5b) 2:504
 KDEL *see* lysine aspartate glutamate leucine (KDEL)
 KDR *see* vascular endothelium growth factor receptor (VEGFR)
 KEGG 4:211
 keloids 3:220
 keratan sulfate (KS) 1:271
 keratin 15 (K15) 3:794–795
 keratinocyte 1:685, 2:569, 2:572–573, 2:574–575, 3:219, 3:220

- keratinocyte growth factor (KGF) 3:210
 keratinocyte migration 3:210
 keratins 2:569, 2:571, 2:572*F*, 2:575
 mutations 2:576
 keratocan 1:273–274
 kethoxal 1:27–28, 1:30*F*
 7-ketocholesterol 1:214
 ketogenesis 1:296
 ketone bodies 1:295, 1:296
 keyframe imaging 2:49
 KGF *see* keratinocyte growth factor (KGF)
 KHCs *see* kinesin heavy chains (KHCs)
 KHD *see* kinase homology domain (KHD)
 KIF4 *see* kinesin-4 family member (KIF4)
 KIF5A *see* kinesin 5A (KIF5A)
 KIF13A *see* kinesin family 13A (KIF13A)
 KIFC3 *see* kinesin superfamily protein 3 (KIFC3)
 KIFCs *see* kinesin-14 motors (KIFCs)
 KIFs *see* kinesin superfamily proteins (KIFs)
 kinase activity assay for kinome profiling (KAYAK) 4:227–228
 kinase cascades 4:44–45
 see also mitogen-activated protein kinases (MAPK) pathway
 kinase homology domain (KHD) 3:58
 kinase inhibitor proteins (KIP) 3:437
 kinase inhibitory regions (KIRs) 3:144–145
 kinase inhibitory proteins (Kip) 3:448
 kinase signaling 3:274–275
 kinase translocation reporters (KTRs) 4:222
 kinases 4:44
 kinase–substrate interactions 4:211–212
 Kindlin2 3:95
 kinesin 2:4, 4:36, 4:40
 kinesin 5A (KIF5A) 2:403, 2:463
 kinesin family 13A (KIF13A) 2:195, 2:197
 kinesin heavy chains (KHCs) 2:608, 3:233–234
 kinesin light chains (KLCs) 2:608, 3:233–234
 kinesin spindle proteins (KSPs) 3:486
 kinesin superfamily protein 3 (KIFC3) 2:709–710
 kinesin superfamily proteins (KIFs) 2:413–414, 2:608–609
 diversity among 2:610*F*
 motor–cargo relationship and pathophysiological relevance 2:611
 kinesin-1 2:625
 kinesin-2 2:665–666
 kinesin-4 family member (KIF4) 3:482–483
 kinesin-11 2:609
 kinesin-14 motors (KIFCs) 2:609–611
 kinesin-II *see* kinesin-2
 kinesins 2:608, 2:608–609, 2:660, 3:233–234
 and cancer 2:612–614
 as components of signal transduction 2:614
 molecular phylogeny and nomenclature 2:609–611
 in neurons 2:612*F*
 regulation 2:614, 2:616*F*
 autonomous regulation of kinesin motility 2:616–617
 of cargo destinations 2:617
 of expression levels 2:614
 by signal transduction 2:614–616
 strategies in kinesin biology 2:609
 structures of kinesin *in vitro* and *in vivo* 2:609*F*
 kinetic stabilization 4:36, 4:42
 kinetochores 2:638, 3:480
 kinins 1:700
 kinome 1:568, 1:569
 kintoun (ktu) 2:631–632
 KIP *see* kinase inhibitor proteins (KIP)
 KIRs *see* kinase inhibitory regions (KIRs)
 ‘kiss and run’
 fusion 2:298–299
 hypothesis 2:412
 transport model 2:186
 kit-on-a-lid-assay (KOALA) 3:672
 KLCs *see* kinesin light chains (KLCs)
 Klf4 *see* Kruppel-like factor 4 (Klf4)
 KLK *see* kallikrein (KLK)
 KLK1 *see* tissue KLK
 KLK8 1:702
 KLKs *see* kallikrein-related peptidases (KLKs)
 Klumperman’s group 2:358
 knock out 2:528
 Knockin alleles 1:514, 1:515
 knockout (KO) 1:514
 control 2:41–42
 mice 1:662–663, 2:843
 KO studies *see* mouse knockout studies (KO studies)
 KOALA *see* kit-on-a-lid-assay (KOALA)
 Kozak sequence 1:343
 Krabbe disease 1:780–781, 1:781
 Krebs cycle *see* tricarboxylic acid (TCA) cycle
 Kruppel-like factor 4 (Klf4) 3:104
 KS proteoglycans (KSPGs) 1:271
 KS *see* keratan sulfate (KS)
 KSHV *see* Kaposi’s sarcoma herpes virus (KSHV)
 KSPGs *see* KS proteoglycans (KSPGs)
 KSPs *see* kinesin spindle proteins (KSPs)
 KTRs *see* kinase translocation reporters (KTRs)
 ktu *see* kintoun (ktu)
 Ku80 3:779
 Kuzbanian (KUZ) 1:643–644, 1:644
- L**
- L2A *see* lysosome-associated membrane protein type 2A (L2A)
 LA *see* LDL receptor type A (LA)
 lab on a chip 3:670–671*F*, 3:673*F*
 LaB6 *see* lanthanum hexaboride (LaB6)
 label-fracture 2:39
 label-free quantification 4:169
 label-retaining cells (LRCs) 3:794–795
 laboratory of genetics and physiology 2 (LGP2) 3:718–719
 lac promoter 1:475
 LacCer *see* lactosylceramide (LacCer)
 lactate dehydrogenase (LDH) 2:524
Lactococcus lactis (*L. lactis*) 1:57–58
 lactoferrin 2:771
 lactosylceramide (LacCer) 1:189
 LacY 4:76
 LacZ 4:76
 lacZ gene 4:64
 LAD *see* lamin-associated domains (LAD); leukocyte adhesion disorders (LAD)
 Lafora bodies 1:269
 Lafora disease 1:269
 Laforin 1:267
 Lag-2 (CSL) 3:107
 lagging-strand synthesis 3:459
 lamellipodia 2:548, 2:561, 2:562>*F*, 2:683, 2:688, 2:733, 2:4, 3:233
 see also filopodia
 actin dynamics 2:689–691
 behavior and functions 2:688–689
 molecular machinery 2:690*F*, 2:691
 structure 2:689
 lamellipodial migration 3:211–212
 lamellipodium 2:726, 4:16, 4:19, 4:20
 lamellum 2:726
 Lamin B receptor (LBR) 2:315
 Lamin C 2:312
 lamin intermediate filaments 2:571, 2:573
 lamina-associated polypeptide (LAP) 2:315
 lamin-associated domains (LAD) 2:313
 laminin-5 γ 2 fragment levels 1:624
 laminins 2:695–696
 laminopathies 2:313, 2:315
 LAMPs *see* lysosome-associated membrane proteins (LAMPs)
 Land’s cycle 1:205
 Langerhans cells and migratory DCs 3:595–596
 lanthanum hexaboride (LaB6) 2:14
 LAP *see* lamina-associated polypeptide (LAP)
 large cargo 2:356
 transport 2:358–359
 large intestine role in proteolysis 1:735–736
 large subunit 3:297, 3:300
 large subunit of ribosome (RPL) 3:281
 large-scale phosphoproteomics strategies 4:179–180, 4:179*F*
 lariat 1:495, 1:495–496, 1:496*F*, 1:498–499, 1:498*F*, 1:499–500
 laser beam 2:65
 laser capture microdissection (LCM) 4:218, 4:220–221
 laser illumination 2:66–67, 2:66*F*
 laser scanning
 frequency-domain FLIM by 2:116
 time-domain FLIM by 2:114–115
 last common eukaryotic ancestor (LCEA) 2:183
 last eukaryotic common ancestor (LECA) 2:442
 last universal common ancestor (LUCA) 1:396, 1:397
 clamp loader in 1:401–403
 mechanism 1:401–403
 structure analysis 1:403
 DNA synthesis 1:397–398
 protein synthesis in 1:397

- last universal common ancestor (LUCA) (*continued*)
- LAT *see* Linker of activated T cells (LAT)
- late domain *see* C-terminal p6 domain of Gag
- late domain motif 2:469–470
- late endosomes (LEs) 1:612, 2:192, 2:201, 2:401, 2:408, 2:410–411, 2:411F, 2:814
- biogenesis of multivesicular body 2:201–203
- composition and fate of intraluminal membranes 2:205
- destinations 2:205–206
- early/recycling endosome network 2:201
- late endosome–lysosome network 2:203–204
- lysosome enzymes and lysosomal membrane proteins 2:204–205
- perinuclear clustering 2:412
- in plants and fungi 2:204
- signaling platform 2:206
- latexin 1:707
- Latrunculin A 2:556
- Lauffenburger spectrum 4:148, 4:148F
- laws of thermodynamics 1:5–6
- LBD *see* lipid-binding domain (LBD)
- LBPA *see* lysobisphosphatidic acid (LBPA)
- LBP *see* lipid-binding proteins (LBPs)
- LBR *see* Lamin B receptor (LBR)
- LC *see* light chains (LC); liquid chromatography (LC); liquid-chromatography (LC)
- LC3/GABARAP/GATE16 (Atg8) 2:504
- LCAT *see* lecithin:cholesterol acyltransferase (LCAT)
- LCCs *see* lineage-committed cells (LCCs)
- LCEA *see* last common eukaryotic ancestor (LCEA)
- LCK *see* lymphocyte-specific protein tyrosine kinase (LCK)
- LCM *see* laser capture microdissection (LCM)
- LC–MS 4:204–205, 4:205F, 4:206–207, 4:208F
- LC-MS/MS *see* liquid chromatography–tandem mass spectrometry (LC-MS/MS)
- LCMV *see* lymphocytic choriomeningitis virus (LCMV)
- LCV *see* *Legionella*-containing vacuole (LCV)
- LDA *see* limit-dextrinase (LDA)
- LDGs *see* low-density granulocytes (LDGs)
- LDH *see* lactate dehydrogenase (LDH)
- LDL receptor type A (LA) 2:419–420
- LDL receptor-related protein 1 (LRP1) 2:197–198, 2:419–420
- LDL receptor-related proteins 5 and 6 (LRP5/6) 3:87–88, 3:88–89, 3:89
- LDL *see* low-density lipoprotein (LDL)
- LDLA *see* low density lipoprotein receptor class A (LDLA)
- Ldlr *see* low-density lipoprotein receptor (Ldlr)
- LDLRA domain *see* low density lipoprotein-receptor class A-like domain (LDLRA domain)
- LDL-receptor-related protein (LRP) 1:633, 2:176, 3:13
- LDs *see* lipid droplets (LDs)
- LDS *see* Loews–Dietz syndrome (LDS)
- LE *see* late endosomes (LEs)
- lead compound 1:135, 1:136–137, 1:138
- leading-strand synthesis 3:459
- leaky scanning 1:522F, 1:523–524
- learning and memory 2:611
- least squares (LS) 4:257–258
- LEC *see* lymphatic endothelial cells (LEC)
- LECA *see* last eukaryotic common ancestor (LECA)
- lecithin:cholesterol acyltransferase (LCAT) 1:173–174
- lectin affinity chromatography 1:72
- lectin mannose-binding protein 1 (LMAN1) 2:169, 2:177
- lectin pathway 1:726, 1:727F
see also alternative pathway; classical pathway
- GAGs 1:728
- MASP-1 and MASP-2 1:727
- MASP-3 1:727–728
- MASPs 1:726
- MBL 1:726
- LEDs *see* light emitting diodes (LEDs); light-emitting diodes (LEDs)
- LEF (lymphoid enhancer-binding factor) 3:94, 3:94F
- left–right determination 2:611
- LegA9 mutants 2:781–782
- LEGI mechanism *see* local excitation, global inhibition mechanism (LEGI mechanism)
- Legionella pneumophila* (*L. pneumophila*) 2:405, 2:777
- Legionella* T4SS SdhA protein 2:779–780
- Legionella*-containing vacuole (LCV) 2:779
- Legs at odd angles (Loa) 3:238
- Lemur tyrosine kinase 2 (LMTK2) 2:415
- lentiviruses 2:828
- Let-7, lung cancer 1:366–367
- lethal giant larvae homolog (Lgl) 2:743, 3:100
- lethal toxin (LT) 3:717
- leucine-rich repeat (LRR) 3:28, 3:300, 3:391, 3:394, 3:520, 3:712
- leucine-rich repeat and immunoglobulin-like domain1 (LRIG1) 3:795
- leucine-rich repeat kinase 2 (LRRK2) 2:385–386, 3:573
- leukemia inhibitor factor (LIF) 3:440
- leukemic cancer stem cells 3:810
- leukocyte adhesion disorders (LAD) 3:754
- leukocyte responses to bacteria 4:98
- leukocyte tyrosine kinase (LTK) 3:16
- leukocytes 4:82
- LeuLeuOMe dipeptidyl ester 1:615
- Lewis X (LeX) 2:342
- Lewy body dementia 1:768, 1:773
- LeX *see* Lewis X (LeX)
- LFA-1 *see* lymphocyte function-associated antigen-1 (LFA-1)
- Lgl *see* lethal giant larvae homolog (Lgl)
- LGMD2A *see* limb-girdle muscular dystrophy type 2A (LGMD2A)
- LGP2 *see* laboratory of genetics and physiology 2 (LGP2)
- LHX2 *see* LIM homeobox protein 2 (LHX2)
- library preparation 1:31–32
- LIC *see* light-intermediate chains (LIC)
- LIF *see* leukemia inhibitor factor (LIF)
- life sciences, FLIM applications in
autofluorescence FLIM 2:110
FRET experiments 2:109–110
local molecular environment parameters measurement 2:108–109
- Life Technologies 1:64
- lifetime line broadening 1:104–105
- LIG4 *see* Ligase IV (LIG4)
- ligand-gated ion channels 1:244, 1:247F, 1:248–249
- ligands 1:66, 3:773
ligand-induced endocytosis of RTKs 2:217
ligand–receptor interaction structural basis 3:61–62
- Ligase IV (LIG4) 1:451–452, 1:452, 1:453–454, 3:781
- ligation 1:451–452
- light 2:147
and electron microscopy 2:1
- light chains (LC) 2:620, 2:627, 2:630
LC3 2:503, 2:515, 3:234–235
- light emitting diodes (LEDs) 2:149
- light microscopy (LM) 2:168
- light sheet fluorescence microscopy (LSFM) 2:55
- light-dependent reactions 2:328
- light-emitting diodes (LEDs) 4:238–239
- light-gated glutamate receptor (LiGluR) 2:151
applications 2:153
principles 2:151–153
- light-intermediate chains (LIC) 2:620
- light-sensitive proteins 2:148T
- LiGluR *see* light-gated glutamate receptor (LiGluR)
- LIM homeobox protein 2 (LHX2) 3:794–795
- limb-girdle muscular dystrophy type 2A (LGMD2A) 1:676
- limit-dextrinase (LDA) 1:267–268
- LIMK *see* LIM kinase (LIMK)
- LIMPs *see* lysosomal integral membrane proteins (LIMPs)
- Lin-7 homolog C (Lin-7C) 3:102
- Lin12/Notch repeats (LNR) 3:107
- LINC *see* linker of nucleoskeleton and cytoskeleton (LINC)
- lineage-committed cells (LCCs) 3:801
- linear kinetics 4:29
- linear ubiquitin chain assembly complex (LUBAC) 3:357–358
- linker of activated T cells (LAT) 3:680
- linker of nucleoskeleton and cytoskeleton (LINC) 2:315, 2:573
- lipases 1:257, 1:615–616
- lipid asymmetry 1:209, 1:210, 1:212, 1:214
- lipid bilayers, physical states of lipids in 1:208

- lipid droplets (LDs) 2:273
 adipocyte LDs 2:274
 basic structure 2:273
 biogenesis 2:275–276
 cellular lipid droplets 2:275F
 and disease 2:276–277
 dynamics 2:276
 neutral lipid synthesis and lipolysis 2:276
 proteins 2:273–274
 subpopulations 2:274–275
- Lipid Maps Consortium 1:194
- LIPID MAPS lipid classification system 1:153, 1:180
- lipid rafts 1:208–217
 biological functions of raft domains in cells 1:214
 close packing 1:209
 detection of rafts
 by microscopy and spectroscopy 1:213
 in model membranes 1:211–212
 in natural membranes 1:212
 using detergents 1:212–213
 using sterol modification 1:213–214
 formulation of lipid raft hypothesis 1:208–209
 H-bonding 1:209
 lipid domain size 1:210–211
 membrane protein interaction with rafts, basis of 1:209–210
 physical states of lipids in lipid bilayers 1:208
 raft affinity 1:209–210
 raft size 1:211
- lipid signaling 1:194
 advances in 1:199
 complexity 1:196
 extracellular signaling lipids interact with receptors 1:198
 intracellular signaling lipids interact with proteins 1:198–199
 phosphatidylinositol lipids 1:196–197
 and disease 1:199
- lipid classification and nomenclature 1:194–195
- physiologic roles 1:199
- primary mammalian signaling lipids 1:195–196, 1:195F
- spatially restricted and scarce 1:196, 1:197F
- lipid transfer proteins (LTPs) 1:205, 2:461–462
- lipid(s) 4:192
see also lipidome
 cholesterol 2:761
 diacylglycerol 2:760–761
 flip-flop 1:204–205, 1:204F
 membrane network 2:376
 metabolism 1:199
 modification
 GPI 1:87
 myristoylation 1:87
 palmitoylation 1:87
 protein prenylation 1:87
 for tubulation 2:367
- PIs 2:760
- PLC 2:761
- raft 1:194, 1:651–652, 2:365–366, 2:365–366
 remodelling 1:204F, 1:205
 role in cell death 3:330–331
 storage 2:273
 synthesis 2:165, 2:329
- lipidation *see* lipid(s): modification
- lipid-binding domain (LBD) 1:198, 1:198T
- lipid-binding proteins (LBPs) 1:195
- lipid-linked oligosaccharide (LLO) 1:180, 1:185–186
- lipidome 1:147, 1:199
 lipid categories 1:153, 1:153F
 fatty acyls 1:153–154
 glycerolipids 1:154–155
 glycerophospholipids 1:155
 polyketides 1:156
 prenol lipids 1:155–156
 saccharolipids 1:156
 sphingolipids 1:155
 sterol lipids 1:155
- lipidomic mass spectrometry 1:150
 chromatography in combination with mass spectrometry 1:151
 IMS 1:152
 mass spectrometry 1:150
 sample preparation for 1:149–150
 selection of instruments for 1:152
 shotgun lipidomic mass spectrometry 1:151–152
 tandem mass spectrometry 1:150–151
 tissue-imaging mass spectrometry 1:152–153
- lipidomics 1:147, 1:147–148
 databases and online tools 1:156–157
 examples of findings 1:157
 CLASS-type lipidomic analysis 1:157
 shotgun lipidomics 1:157
 tissue-imaging mass spectrometry 1:157
- lipid categories and metabolic interrelationships 1:148F
- lipid categories of lipidome 1:153, 1:153F
 fatty acyls 1:153–154
 glycerolipids 1:154–155
 glycerophospholipids 1:155
 polyketides 1:156
 prenol lipids 1:155–156
 saccharolipids 1:156
 sphingolipids 1:155
 sterol lipids 1:155
- lipidome components 1:148
- lipidomic analysis 1:148–149
- lipidomic mass spectrometry 1:150
 chromatography in combination with mass spectrometry 1:151
 IMS 1:152
 mass spectrometry 1:150
 sample preparation for 1:149–150
 selection of instruments for 1:152
 shotgun lipidomic mass spectrometry 1:151–152
 tandem mass spectrometry 1:150–151
 tissue-imaging mass spectrometry 1:152–153
- lipodystrophy 1:743
- lipogenesis 1:289F, 1:293F, 2:273
- lipolysis 2:276
- lipopolysaccharide (LPS) 1:156, 1:186, 1:640, 2:192, 2:305–306, 2:767–768, 2:791, 3:394, 3:581, 3:619–620, 3:702, 3:727, 3:752
- lipoprotein receptor-related protein 1 (LRP1)
- lipoproteins 1:173–174, 1:177
- liposomes 2:459
- lipoteichoic acid (LTA) 3:727
- Lippincott-Schwarz's group 2:360
- liquid chromatography (LC) 1:151
- liquid chromatography-tandem mass spectrometry (LC-MS/MS) 1:150, 1:151F, 1:568, 1:571–572, 1:572F, 1:573F, 1:574–575, 1:575–576, 1:575F, 1:576, 1:577, 1:577F, 1:578–579, 1:578F, 4:166, 4:178
- digestion of proteins into peptides 4:166–167
- measurement of peptides by mass spectrometry 4:167
 DDA 4:167–168
 DIA 4:168–169
 proteomic depth against data completeness 4:168F
 targeted data acquisition 4:168
 peptide abundance quantification 4:169
 reducing complexity by separation 4:167
- liquid disordered state 1:208
- liquid nitrogen (LN₂) 2:34–35
- liquid ordered state 1:208
 nanodomains 1:211
- liquid-chromatography (LC) 4:166–167, 4:167
- LIS1 *see* lissencephaly-1 (LIS1)
- lissencephaly-1 (LIS1) 2:624, 3:482
- Listeria* 2:780, 2:781, 4:16–17
- Listeria monocytogenes* (*L. monocytogenes*) 2:785
- Listeria* toxin (LLO) 2:287
- listeriolysin O (LLO) 2:780, 2:790, 3:597
- live animals
 imaging individual cells in 2:139F
 imaging tissue architecture and function in 2:134–135F
 tumor progression and invasive behavior in 2:137–138F
- live cell/living cells 1:5
 compatibility 2:113–114
 images 2:65
 imaging 2:56–59
 laws of thermodynamics and 1:5–6
 molecular crowding effect in 1:11–12
- live cell imaging 2:70
- live SIM 2:94–95
- live-cell imaging techniques 4:222
- liver 1:237
- liver kinase B1 (LKB1) 3:102, 3:130
- liver pyruvate kinase 1:292–293
- liver sinusoidal endothelial cell (LSEC) 3:641, 3:641
- liver X-receptors (LXR) 1:176–177
 LXR α 2:427
 LXR β 2:427
- LKB1 *see* liver kinase B1 (LKB1)

- LKB1-AMPK pathway 3:420
 LL-37 2:770, 2:771
 LLO *see* lipid-linked oligosaccharide (LLO);
 Listeria toxin (LLO); listeriolysin O
 (LLO)
 LM *see* light microscopy (LM); localization
 microscopy (LM)
 LMAN1 *see* lectin mannose-binding protein
 1 (LMAN1)
 LMP *see* lysosomal membrane
 permeabilization (LMP)
 LMTK2 *see* Lemur tyrosine kinase 2
 (LMTK2)
 LMWK *see* low-molecular weight kininogen
 (LMWK)
 LN *see* lymph node (LN)
 LN₂ *see* liquid nitrogen (LN₂)
 LNAs *see* locked nucleic acids (LNAs)
 lncRNAs *see* long noncoding RNAs
 (lncRNAs)
 LNR *see* Lin12/Notch repeats (LNR)
 Loa *see* Legs at odd angles (Loa)
 lobopodial migration 3:212
 local excitation, global inhibition
 mechanism (LEGI mechanism) 4:6,
 4:6F
 local protein synthesis 3:230
 localization algorithms 2:65
 localization methods 2:2
 localization microscopy (LM) 2:63, 2:65,
 2:121
 experimental requirements 2:122–123,
 2:123F
 fluorescent probes in 2:122T
 future trends/avenues 2:126–127
 high-speed 2:123–124
 localization concept 2:121–122
 sequence flow in live cell 2:122F
 locked nucleic acids (LNAs) 1:368
 Loews–Dietz syndrome (LDS) 3:46–47
 Log P 1:135, 1:138
 logarithmic gain 4:44, 4:45, 4:46
 London resin (LR) 2:34
 long noncoding RNAs (lncRNAs) 1:364,
 1:369–370, 4:160
 long range inhibitor 4:52
 longin v-SNAREs Sec22b 2:461–463
 long-loop recycling pathway 2:401
 long-term depression (LTD) 2:535
 long-term potentiation (LTP) 2:401, 2:535
 5' long terminal repeat (LTR) 2:835
 loops 1:112
 lopinavir 1:743
 loss of function 1:514
 low dead volume sampling device 4:203F
 low degree of separation 4:68
 low density lipoprotein receptor class A
 (LDLA) 1:655
 low density lipoprotein-receptor class A-like
 domain (LDLRA domain) 1:650,
 1:729
 Low pH 2:219
 low-density granulocytes (LDGs) 3:753–754
 low-density lipoprotein (LDL) 1:173–174,
 2:176, 2:205–206, 3:13, 3:418, 3:2,
 4:56
 see also high-density lipoprotein (HDL)
 biosynthesis 2:418–419
 endosomal dynamics 2:425–426
 human genes and proteins 2:419T
 receptor
 biosynthesis and trafficking 2:421F
 cargo-selective clathrin adaptors for
 2:422–423
 dynamic modulation of surface levels
 2:427–428
 mutants 2:421–422
 protein interaction hub 2:430F
 structure 2:419–421
 superfamily 2:420F
 ubiquitin-independent downregulation
 2:428–429
 low-density lipoprotein receptor (Ldlr)
 1:676
 low-molecular weight kininogen (LMWK)
 1:699
 LOX *see* lysyl oxidase (LOX)
 LOXL2 *see* LOX-like 2 (LOXL2)
 LOX-like 2 (LOXL2) 4:143
 LPA acyltransferase type γ (LPAAT- γ) 2:360
 LPA *see* lysophosphatidic acid (LPA)
 LPAM-1 *see* lymphocyte Peyer's patch
 adhesion molecule (LPAM-1)
 LPLA₂ *see* lysosomal phospholipase A₂
 (LPLA₂)
 LPS *see* lipopolysaccharide (LPS)
 LR *see* London resin (LR)
 LRCs *see* label-retaining cells (LRCs)
 LRIG1 *see* leucine-rich repeats
 and immunoglobulin-like domain 1
 (LRIG1)
 LROs *see* lysosome-related organelles
 (LROs)
 LRP *see* LDL-receptor-related protein (LRP)
 LRP5/6, Frizzled and 3:87–88, 3:88F
 LRR *see* leucine-rich repeat (LRR)
 LRRK2 *see* leucine-rich repeat kinase 2
 (LRRK2)
 LS *see* least squares (LS)
 LSDs *see* lysosomal storage diseases (LSDs)
 LSEC *see* liver sinusoidal endothelial cell
 (LSEC)
 LSFM *see* light sheet fluorescence
 microscopy (LSFM)
 LT *see* lethal toxin (LT); lymphoid tissues
 (LT); lymphotoxin (LT)
 LTA *see* lipoteichoic acid (LTA)
 LTD *see* long-term depression (LTD)
 LT α cells *see* lymphoid tissue inducer cells
 (LT α cells)
 LTK *see* leukocyte tyrosine kinase (LTK)
 LTP *see* long-term potentiation (LTP)
 LTPs *see* lipid transfer proteins (LTPs)
 LTR *see* 5' long terminal repeat (LTR)
 LT α *see* lymphotoxin- α (LT α)
 LT β *see* lymphotoxin-beta (LT β)
 LT β R signaling *see* lymphotoxin beta
 receptor signaling (LT β R signaling)
 LUBAC *see* linear ubiquitin chain assembly
 complex (LUBAC)
 LUCA *see* last universal common ancestor
 (LUCA)
 Lucifer yellow 2:761
 Luini's group 2:358
 lumen 2:175–176, 2:176
 lumen formation 2:744F, 2:745–746
 luminal conditions 2:176
 lumican 1:273–274
 LUMIER *see* luminescence-based
 mammalian interactome mapping
 (LUMIER)
 luminescence-based mammalian
 interactome mapping (LUMIER)
 4:189–190
 LXR *see* liver X-receptors (LXR)
 lymph node (LN) 3:616, 3:639, 3:759
 macrophages 3:765
 lymph node compartments and
 localization of macrophage
 populations 3:765F
 medullary macrophages 3:766–767
 SCMs 3:765–766
 lymphatic circulatory system 3:632
 lymphatic endothelial cells (LEC) 3:616,
 3:634, 3:641
 lymphatic endothelium 3:634
 lymphocyte function-associated antigen-1
 (LFA-1) 3:522–524, 3:624–625,
 3:752
 lymphocyte–endothelial interactions 3:634
 information exchange 3:639
 endothelial function in antigen-specific
 immunity 3:639–640
 T lymphocytes 3:634
 trafficking during immune responses
 3:636
 trafficking during immune surveillance
 3:635–636
 trafficking during maturation
 3:634–635
 trafficking-independent interactions
 between EC 3:636
 lymphocyte Peyer's patch adhesion
 molecule (LPAM-1) 3:526
 lymphocytes 3:522
 lymphocyte-specific protein tyrosine kinase
 (LCK) 3:533
 lymphocytic choriomeningitis virus (LCMV)
 3:570, 3:766
 lymphoid organs, DCs in 3:741
 lymphoid tissue (LT) 3:525–526, 3:617F
 stromal cells regulation
 initiation and regulation of adaptive
 immune responses 3:617–618
 regulation of lymphocyte migration
 3:616–617
 regulation of peripheral tolerance
 3:618–619
 SLOs 3:616
 lymphoid tissue DCs 3:594, 3:596–597
 lymphoid tissue inducer cells (LT α cells)
 3:745–746, 3:773
 lymphoid tissue-resident DCs 3:592,
 3:593–594
 lymphotoxin (LT) 3:68, 3:761
 lymphotoxin beta receptor signaling (LT β R
 signaling) 3:617
 lymphotoxin-beta (LT β) 3:745

- lymphotoxin- α (LT α) 3:72, 3:354
 LYNUS machinery *see* lysosome nutrient sensing (LYNUS) machinery
 Lys-83/84 1:123
 lysine aspartate glutamate leucine (KDEL) 2:164
 KDEL receptor (KDEL-R) 2:169, 2:176, 2:373
 lysobisphosphatidic acid (LBPA) 2:205, 2:410
 lysophosphatidic acid (LPA) 1:195, 2:276, 2:360, 3:102–103
 lysosomal beta-galactosidase activity 3:511
 lysosomal diseases
 activator protein, deficiency of 1:774
 galactosialidosis 1:774
 sphingolipid activator proteins (SAP) 1:774
 Chédiak–Higashi disease 1:776
 corrective factor 1:777, 1:782
 and disturbances of autophagy 1:770
 functional complementation 1:777–778
 I-cell disease 1:777–778
 in vivo 1:778
 future prospects in clinical research 1:785
 Gaucher disease 1:767–768, 1:772
 comorbid associations of 1:772–773
 deficiency of a specific acid hydrolase 1:772
 Gaucher disease 1:785
 Griscelli syndromes (Types 1–3) 1:776
 Hermansky–Pudlak syndrome 1:775–776
 lysosomal membrane protein/transporter, deficiency of 1:774–775
 cystinosis 1:775, 1:778–779
 Danon disease 1:775
 organelle biogenesis 1:776–777
 lysosome-related organelles and osteoclast 1:776–777
 pathogenesis and physiological understanding 1:785–786
 pathogenesis of 1:769
 lysosomal storage 1:769
 Pompe disease 1:767
 role of autophagy in glycogen remodeling 1:770
 posttranslational processing and organelle targeting 1:773–774
 I-cell disease 1:773–774
 recognition signal 1:777–778
 Tay–Sachs disease 1:767
 therapeutic stratagems for 1:782–783
 enzymatic augmentation 1:783
 gene transfer 1:783–784
 substrate-reduction 1:783
 treatment of 1:778–779
 dissolution and removal of intralysosomal storage material 1:778–779
 enhanced exocytosis of pathological intralysosomal content 1:779
 enzymatic complementation 1:781–782
 hematopoietic stem-cell transplantation 1:780–781
 substrate biosynthesis, modulating 1:779–780
 lysosomal hydrolase precursors 2:435
 lysosomal integral membrane proteins (LIMPs) 1:763–764
 LIMP-2 1:773–774, 1:778
 lysosomal integral membrane proteins (LIMPs) 2:225
 lysosomal membrane permeabilization (LMP) 1:615
 lysosomal membrane proteins 1:616
 lysosomal phospholipase A₂ (LPLA₂) 1:613
 lysosomal proteins, soluble hydrolases 1:613
 lysosomal storage 1:769, 1:769–770, 1:779
 lysosomal storage diseases (LSDs) 1:769, 1:772, 2:225–226
 lysosome(s) 1:14, 2:2, 2:2–3, 2:225, 2:408, 2:410, 2:412, 2:469, 2:751, 2:754, 2:771, 2:776, 2:848
 biogenesis and autophagy 1:770–771, 2:228–229
 Danon disease 1:771–772
 reciprocal signaling between mTORC1 and TFEB on 1:771
 cell biology of 1:763
 cell type-specific secretory lysosomes 2:225–228
 hematopoietic-lineage secretory lysosomes 2:226–228
 nonhematologic cell secretory lysosomes 2:228
 compartment 1:763–764
 antigen presentation 1:764
 autophagy 1:764–765
 endocytosis 1:764
 phagocytosis 1:764
 conventional lysosomes 2:225
 debt of cell science to medicine 1:763
 delivery 3:406
 energy economy and nutrient-sensing 1:765–766
 amino acid signaling 1:766–767
 autophagy 1:765–766
 growth factor signaling 1:766
 enzymes 2:204–205
 exocytosis of secretory lysosomes 2:229–230
 formation 2:390
 functions of 1:613F
 hydrolases 2:410
 inborn errors of 1:768, 1:769
 biochemical and structural abnormalities 1:769
 identification 1:768–769
 lysosomal membranes, mTORC1 with Rheb on 3:249–250
 macromolecular biosynthesis and substrate loading in 1:785
 membrane proteins 2:204–205
 and recycling of macromolecules 1:763–764, 1:764, 1:769
 and recycling of organelles 1:763, 1:770
 role in intracellular degradation 1:612
 cathepsins 1:613–614
 glycosidases 1:616
 lipases 1:615–616
 soluble hydrolases 1:613
 autophagy 1:618
 endocytosis 1:617–618
 lysosome nutrient sensing (LYNUS) machinery 1:765–766
 lysosome-associated membrane protein type 2A (L2A) 3:406
 lysosome-associated membrane proteins (LAMPs) 1:616, 1:763–764, 2:225, 2:235, 2:755–756, 3:727
 LAMP1 1:616, 2:410, 2:504
 LAMP2 1:771–772, 1:775, 2:410
 LAMP2-A 1:616, 1:764, 1:771–772
 lysosome-related organelles (LROs) 1:612, 1:768, 1:774–775, 1:775, 1:776, 1:785–786, 2:203–204, 2:225, 2:235, 2:363
 biogenesis and transport of 2:235–237
 cell-specific LROs and functions 2:236T
 genetic diseases associated with defects in 2:237–241
 morphology, composition, and function of 2:235
 and osteoclast 1:776–777
 unique morphology 2:236F
 LYST (lysosomal trafficking regulator gene) 1:776
 lysyl oxidase (LOX) 4:142–143
 lysyl-tRNA synthetase 1:124, 1:327
 lytic granule 2:225, 2:226
- ## M
- M line 2:677
 M1 domain *see* modulatory 1 domain (M1 domain)
 M6P group *see* mannose-6-phosphate group (M6P group)
 MA protein *see* matrix protein (MA protein)
 MAC-1 *see* macrophage antigen-1 (MAC-1)
 machine learning approaches, optimization driven 4:257
 macroautophagy 1:618, 1:764, 1:765, 1:771–772, 2:243–244, 2:515, 3:404, 3:405F, 3:407–408, 3:591, 3:692–693
 macromolecular communication between nucleus and cytoplasm
 electron micrographs of animal cell nuclei 3:171F
 importin β superfamily 3:172–174, 3:173T
 directionality of translocation 3:174
 importins 3:177
 macromolecules transport 3:170
 mediators of RNA export 3:174T
 mRNA export 3:174–175
 NPCs 3:172
 nuclear envelope 3:176–177
 nuclear import and export 3:173F
 nuclear pores 3:171T
 nucleoporins 3:175, 3:177
 nucleus 3:176
 protein synthesis 3:170
 Ran 3:177

- macromolecular communication between nucleus and cytoplasm (*continued*)
 regulation of transport 3:175–176
 shuttling 3:176
 translocons for oversized cargos 3:177–178
 transport signals 3:172–174
- macromolecular complexes
 single-particle electron cryo-microscopy (cryoEM) of 2:5–13
- macromolecular synthesis 3:3
- macrophages 2:227, 2:828, 2:829–830, 2:835, 2:836, 3:218–219, 3:220, 4:95, 4:97, 4:97–98, 4:97F, 4:99, 4:100
- macropinocytosis and CDR formation 2:758F
- phagosome maturation 2:768–770, 2:769F
- pyroptosis 2:772F
- macrophage antigen-1 (MAC-1) 3:752
- macrophage mannose receptor (MMR) 3:746
- macrophage receptor with collagenous structure (MARCO) 3:727
- macrophage-colony-stimulating factor (M-CSF) 2:758, 3:761
- macrophage-DC progenitors (MDP) 3:743
- macropinocytic cups 2:760, 2:761
- macropinocytosis 2:394, 2:395, 2:751–752, 2:758–765, 3:746
see also caveolae; clathrin-and caveolin-independent micropinocytic pathways
- carriers 2:394
- and CDR formation in macrophages 2:758F
- dysregulation 2:762
- experimental diagnosis 2:761
- functions 2:395–396, 2:758–759
- macropinosome formation 2:759
- molecules 2:394
- necessary molecules 2:761
- pathway 2:394–395
- quantitative aspects and dynamics 2:395
- solute flow during pinocytosis 2:759
- therapeutics 2:762
- macropinosomes 2:394, 2:395, 2:758
 formation 2:394, 2:759
 cytoplasm organization 2:761
 GFR signals effect 2:759
 GTPases 2:760
 lipids 2:760
 sequence of 3'PIs during 2:760F
 small GTPases 2:759
- Mad Cow disease *see* bovine spongiform encephalopathy (BSE)
- Mad Homology 1 domain (MH1 domain) 3:43
- Mad *see* Mothers against Dpp (Mad)
- MAD1 *see* mitotic arrest deficient 1 (MAD1)
- MAdCAM *see* mucosal addressin cell adhesion molecule (MAdCAM)
- MAd-CAM-1 *see* mucosal vascular addressin cell adhesion molecule 1 (MAd-CAM-1)
- Madin-Darby canine kidney cells (MDCK cells) 2:741, 2:743–744
- Magic Bullet concept 1:143
- magnetic resonance imaging (MRI) 3:736
- magnetosomes 2:594
- magnetotactic bacteria 2:594
- MAGPs *see* microfibril associated glycoproteins (MAGPs)
- major histocompatibility complex (MHC)
 1:764, 2:211–212, 2:395, 3:519, 3:528, 3:552–553, 3:619, 3:634, 3:677, 3:696
 class II molecules 1:764
 MHC-I 3:588–589, 3:589, 3:590
 MHC-II 2:225, 3:541, 3:588–589, 3:589–590, 3:591, 3:591–592, 3:592, 3:593, 3:634
 proteins 3:6
- major vault protein (MVP) 2:780
- MAL *see* MYD88-adaptor-like protein (MAL)
- malate/aspartate shuttle 1:288, 1:296
- MALDI *see* matrix-assisted laser desorption ionization (MALDI)
- malfunctioning of protein folding in ER 1:543–544
- malic enzyme (ME) 1:290
- malignancies 3:17
- malignant progression 1:753, 1:754, 1:757
- malnutrition 2:805
- maltooligosyltrehalose trehalohydrolase 1:123
- maltose-binding protein (MBP) 1:73
- MAM *see* mitochondria-associated membranes (MAM)
- mammalian cell cycle segments 3:5
- mammalian cells, apoptosis in 2:531–532
- mammalian Diaphanous 1 (mDia1) 3:187
- mammalian expression systems 1:60
see also bacterial expression systems; cell-free expression systems; yeast expression systems
- aliphatic acid-type histone deacetylase inhibitors 1:62
- episomal expression systems 1:62–64
- Life Technologies 1:64
- plasmid DNA 1:60
- stable cell-line generation 1:61
- mammalian immune defense against viruses 2:255
- mammalian kinesins 2:609
- mammalian Lethal-with-Sec-Thirteen 8 (mLST8) 3:130–131, 3:243–244
- mammalian protein-protein interaction trap (MAPPIT) 4:189–190
- mammalian SEPT2/6/7/9 complexes 2:584
- mammalian septin paralogs 2:579
- mammalian target of rapamycin (mTOR)
 2:206, 2:217–219, 3:128, 3:130–132, 3:236–237, 3:274–275, 3:282, 3:404
 mTOR complex 1 (mTORC1) 1:617–618
 signaling pathway 3:244F
 signaling regulation
 by amino acids 3:248–249
 by cellular stress 3:247
 by energy 3:247
 by growth factors 3:246–247
- mammalian target of rapamycin complex 1 (mTORC1) 2:206, 2:245, 2:504, 3:130, 3:243–253, 3:243, 3:781
- Deptor 3:243–244
- downstream 3:245
 autophagy 3:246
 Grb10 3:245–246
 S6K1 and 4EBP1 3:245
 SREBPs 3:246
 TFEB 3:246
- mLST8 3:243–244
- PRAS40 3:243
- rapamycin effect on 3:244–245
- Raptor 3:243
- spatial regulation 3:249F
- upstream
 mTOR signaling regulation by amino acids 3:248–249
 mTOR signaling regulation by cellular stress 3:247
 mTOR signaling regulation by energy 3:247
 mTOR signaling regulation by growth factors 3:246–247
 with Rheb on lysosomal membranes 3:249–250
- mammalian target of rapamycin complex 2 (mTORC2) 3:128, 3:243, 3:244
- mammalian tumors, Rho GTPases in 2:737–738
- mammary-derived growth inhibitor (MDGI) 3:194
- Man *see* mannose (Man)
- Man-6-P receptor (MPR) 2:204–205, 2:341–342
- Man-6-P receptor homology domains (MRH domains) 2:343–344
- ManLAM *see* mannose-capped lipoarabinomannan (ManLAM)
- mannitol (Mtl) 1:101
- mannose (Man) 1:101, 2:339–340
- mannose 6-phosphate (M6P) 1:773, 1:773–774, 1:777, 1:777–778, 1:778, 1:781, 1:782, 1:783, 2:433, 2:435
- mannose receptor 1:781
- mannose-6-phosphate group (M6P group) 2:410
- mannose-6-phosphate receptor (MPRs/M6PR) 1:612, 1:616, 2:433, 2:434F, 2:435, 2:436, 2:788–789
 cation-dependent MPR (CI-MPR) 2:435
 cation-independent MPR (CI-MPR) 2:435
- mannose-binding lectin (MBL) 1:726, 3:710–712
- mannose-capped lipoarabinomannan (ManLAM) 2:777
- manual curation limitations 4:245–246
- MAP kinases (MAPKs) 3:44
- MAP2K7 *see* mitogen-activated protein kinase kinase 7 (MAP2K7)
- MAP3K 3:122, 3:123F, 3:123, 3:123–124, 3:124, 3:125
see also mitogen-activated protein kinase kinase kinase (MAPKKK)

- MAP4K 3:122, 3:123, 3:123–124, 3:123F, 3:124, 3:125
- MAPC *see* multipotent adult progenitor cells (MAPC)
- MAPK *see* mitogen-activated protein kinase (MAPK)
- MAPK and ERK kinase 1 (MEK1) 2:219
- MAPKK family *see* mitogen-activated protein kinase kinase family (MAPKK family)
- MAPKKK *see* mitogen-activated protein kinase kinase kinase (MAPKKK)
- MAPKs *see* MAP kinases (MAPKs)
- MAPPIT *see* mammalian protein–protein interaction trap (MAPPIT)
- MAPs *see* microtubule-associated proteins (MAPs); mitogen-activated protein (MAP)
- MARCH5/MITOL mutant 2:288
- Marchado–Joseph disease (MJDs) 3:505–506
- MARCO *see* macrophage receptor with collagenous structure (MARCO)
- Marfan syndrome (MFS) 3:46–47
- marginal reticular cells (MRCs) 3:616
- marginal zone (MZ) 3:759
- apoptotic cells clearance 3:763
- MARINa *see* Master Regulator Inference algorithm (MARINa)
- Markov chain Monte Carlo 4:212T, 4:214
- Maroteaux–Lamy disease (MPS VI) 1:777, 1:782
- marrow-isolated adult multilineage inducible cells (MIAMI cells) 3:819
- Mas receptor 1:747
- Mash 3:790–791
- Maskless array synthesis 1:47
- Mason–Pfizer monkey virus (M-PMV) 1:744
- MASPs *see* MBL-associated serine proteases (MASPs)
- mass spectrometry (MS) 1:78–79, 1:86, 1:148, 1:150, 1:571, 3:264, 3:273, 4:141, 4:166, 4:177–178, 4:211
- based proteomics
- correlation of proteome with genome 4:171
- correlation of proteome with phenotype 4:171
- dynamic proteomes measurement 4:171–172
- mapping nodes of proteome 4:170
- organization of proteome 4:170–171
- chromatography in combination with 1:151
- peptide and protein sequences 1:79
- proteomic measurements 1:79–80
- technologies 4:163
- mass-balance 4:24–25, 4:25F
- equation 4:25
- massively parallel DNA sequencing 1:32
- mast cells 2:226–227
- master regulator genes (MR genes) 4:259
- Master Regulator Inference algorithm (MARINa) 4:259, 4:260F
- MAT *see* mesenchymal–amoeboid transition (MAT)
- mathematical and computational modeling 4:72
- off-lattice approaches 4:74
- mathematical modeling 4:80
- for studying collective cell behavior 4:123
- cell–ECM interactions 4:125
- continuum models 4:123
- discrete models 4:123–124
- for studying inflammation 4:95–101
- matricellular proteins 2:700–701
- matriptase 1:656, 1:756–757
- matriptase-2 1:656
- matrix 1:66
- attachment regions 2:573
- degradation 3:213
- molecules in connective tissue function 3:194–195
- remodeling 3:213, 3:214F
- matrix metalloproteinases (MMPs) 1:621–622, 1:630, 1:695, 1:755–756, 3:213, 3:219, 3:625, 3:751, 4:106
- activity in drug delivery mechanisms 1:627
- cellular regulation
- compartmentalization 1:626
- natural inhibitors 1:626–627
- posttranscriptional 1:625–626
- posttranslational 1:626
- transcriptional 1:625
- domain structure 1:622F
- imaging 1:627
- MMP-2 1:754
- structure 1:621–622
- substrate profiles 1:622–624, 1:623T
- cell–ECM and cell–cell adhesion molecules 1:625
- cytokines and chemokines and their receptors 1:624–625
- ECM 1:624
- intracellular activities 1:625
- therapeutic targets and anti-targets 1:627
- matrix protein (MA protein) 2:828–829
- matrix-assisted laser desorption ionization (MALDI) 1:79
- Mats *see* Mob as tumor suppressor (Mats)
- maturation promoting factor (MPF) 3:424, 3:427–428, 3:427F
- mature neutrophils 3:751
- MAVS *see* mitochondrial antiviral-signaling protein (MAVS)
- MAVS protein *see* mitochondrial antiviral signaling protein (MAVS protein)
- MaxCyte STX[®] 1:61–62
- maximal connectivity centers 2:456–457
- maximum likelihood (ML) 4:257
- MB *see* microtubule blades (MB)
- MBD1 3:790–791
- MBL *see* mannose-binding lectin (MBL)
- MBL-associated serine proteases (MASPs) 1:726
- MASP-1 1:727
- MASP-2 1:727
- MASP-3 1:727–728
- MASP1/3 gene 1:726
- MBP *see* maltose-binding protein (MBP); myelin basic protein (MBP)
- MC *see* mother centriole (MC)
- MCAK *see* mitotic centromere-associated kinesin (MCAK)
- MCC *see* mitotic checkpoint complex (MCC)
- MCF *see* methylchloroformate (MCF)
- MCL1 *see* myeloid cell leukemia sequence 1 (MCL1)
- MCM *see* mini-chromosome maintenance (MCM)
- MCM proteins *see* mini-chromosome maintenance proteins (MCM proteins)
- MCP-1 *see* monocyte chemoattractant protein 1 (MCP-1)
- MCS *see* membrane contact sites (MCS)
- M-CSF *see* macrophage-colony-stimulating factor (M-CSF)
- MCT *see* mono-carboxylate transporter (MCT)
- MCUs *see* mitochondrial calcium uniporters (MCUs)
- MDA5 *see* melanoma differentiation factor 5 (MDA5)
- MDCK cells *see* Madin–Darby canine kidney cells (MDCK cells)
- mDCs *see* migratory DCs (mDCs)
- MDGI *see* mammary-derived growth inhibitor (MDGI)
- mDia1 *see* mammalian Diaphanous 1 (mDia1)
- Mdm2 *see* mouse double minute 2 (Mdm2)
- MDM2 binding ribosomal proteins 3:285–286
- MDM2 *see* mouse double minute 2 (Mdm2); murine double minute 2 (MDM2)
- MDM2-p53 pathway, non-ribosomal functions of RPs independent of 3:283
- Mdm10 2:281–282
- Mdm12 2:281–282
- Mdm34 2:281–282
- MDP *see* macrophage-DC progenitors (MDP); muramyl dipeptide (MDP)
- MDR *see* multidrug resistance (MDR)
- MDS *see* myelodysplastic syndrome (MDS)
- MDVs *see* mitochondria-derived vesicles (MDVs)
- ME *see* malic enzyme (ME)
- mean field theory 4:83
- mean square displacement (MSD) 2:124–125, 2:126
- MEB *see* muscle–eye–brain disease (MEB)
- mechanical forces 2:724
- mechanical induction 2:629
- mechanistic target of rapamycin (mTOR) 3:815
- mechanistic target of rapamycin complex (mTORC) 1:765
- mTORC1 1:765, 1:765–766, 1:766, 1:766–767, 1:771
- mechanochemical ATPase cycle 2:601
- mechanoreceptors, TCR as 3:655
- mechanosensing 2:716–717
- mechanosensory organs 3:199
- mechanoaxis 3:209

- Meckel-Gruber Syndrome (MKS) 2:664
 MECP1 *see* heterochromatin methyl-binding protein 1 (MECP1)
 MeCP2 3:790–791
medial-Golgi compartment 2:339–340, 2:342
 medullary macrophages 3:766–767
 medullary sinus macrophages (MSMs) 3:766–767, 3:767
 medullary thymic epithelial cells (mTEC) 3:747
 MEF *see* mouse embryonic fibroblasts (MEFs); murine embryonic fibroblast (MEF)
 mega-channels 3:69–70
 megakaryocyte 1:714, 2:227–228
 MEK1 *see* MAPK and ERK kinase 1 (MEK1)
 MEK1 partner (MP1) 2:219
 MEKK1 3:123
 melanogenic enzymes 2:235
 melanoma differentiation factor 5 (MDA5) 3:718–719
 melanoma IAP (ML-IAP) 1:707–708, 3:394
 melanophilin *see* Slac2-a
 melanosomes 2:228, 2:235
 molecular basis 2:236–237
 MELAS *see* mitochondrial myopathy, encephalopathy, lactic acid acidosis, and stroke like episodes (MELAS)
 MELT *see* Met-Glu-Leu-Thr (MELT)
 membrane asymmetry 1:205, 1:206
 membrane contact sites (MCS) 2:461–462
 membrane fusion 2:485, 2:833, 3:498–499
 nanomachines 2:459–460
 core mechanism 2:460F
 regulators 2:460–461
 membrane potential (V_m) 1:218–236, 1:218, 1:244, 1:244–245, 1:247, 1:248
 caveats in measuring physiological V_m 1:232–234
 definitions of different types of 1:218
 determinants of 1:221–223
 active ion transport 1:226
 changes in ion concentrations 1:225–226
 electrochemical driving force for ion movement 1:221–223
 ion channel gating 1:229
 passive ion transport via selective 1:226–229
 passive membrane potential changes 1:229–230
 relative membrane permeability 1:223–225
 steady state vs. electrochemical equilibrium 1:225
 membrane selectivity and direct measurements of 1:230–231
 physiological values of 1:234–235
 quantifying membrane and ion channel selectivity 1:231–232
 membrane protein interaction with rafts 1:209–210
 membrane protrusions 2:790
 membrane rafts *see* lipid rafts
 membrane receptors
 chemoattractant receptors 2:724
 focal adhesion proteins 2:724
 integrins 2:724
 transactivation by GPCR 2:725F
 membrane traffic 1:202, 2:2–3
 membrane transport pathways 2:3
 membrane transport processes 1:227T
 membrane type (MT) 1:621–622
 membrane(s) 3:661
 abscission 3:484–486
 atomic structures of membrane protein complexes 1:298–299
 blebbing 2:723
 component structure 3:703
 curvature 1:205–206
 deformation 2:202–203, 2:688
 DMVs formation membrane rearrangement by 2:822
 electron transfer chain 1:298
 electrostatics 1:202–203
 embedding of ER client proteins 2:160–161
 filtration 1:66
 fission 2:367
 general traits of 1:202
 lipid packing 1:203
 membrane ingress 3:482–484, 3:484F
 membrane lipids 2:365
 membrane proteins exocytosis 2:380
 membrane remodeling 2:754
 mechanisms that modifies the bulk lipid composition of organelles in 1:204–205
 modern cell membranes 1:201F
 origin 1:201–202
 pores 2:312
 formation 2:522–523
 primitive 1:201F
 selectivity 1:218, 1:230–231, 1:231–232, 1:235
 skeleton 2:581
 subdomains 1:203–204
 targeting 2:442–443
 trafficking 2:141–143, 2:235–236, 2:384, 2:385, 2:390–391, 2:412–413, 2:486F
 translocation 2:796
 transport 2:371
 tubules stabilization 2:476–477
 webs 2:820–822
 membrane-anchored serine proteases 1:650, 1:651F, 1:652T
 catalytic activity 1:650
 inhibition 1:650–651
 stem region accessory domains 1:650
 membrane-associated palmitoylated protein 5 (MMP5) 3:102
 membrane-based adenosine triphosphate synthesis, energetics of 1:298
 membrane-bound fluorescent probes 1:213
 membrane-bound phagosome formation 2:776
 membrane-bound TNF (mTNF) 3:62–63
 membrane-bound α -glucosidase (MOGS) 2:341
 membrane-cytosol interface, phosphoinositides at 2:505
 membrane-type 1 matrix metalloprotease (MT1-MMP) 2:402–403, 4:110
 MeMo (Mutual Exclusivity Modules in Cancer) 4:212T, 4:214
 MeMo in Cancer *see* mutual exclusivity modules in cancer (MeMo in Cancer)
 memory 3:567
 cells 3:770
 formation 3:567
 inflation 3:570
 precursors 3:568
 memory T cells 3:567–568, 3:569F
 atypical memory 3:570–571
 HIV 3:567
 immune exhaustion 3:570
 impact on vaccine 3:571
 memory formation 3:567
 memory inflation 3:570
 tissue-resident memory 3:568–569
 viral infections 3:569–570
 mental-disorder associated GAP protein *see* SrGAP3
 MEP1A 1:694
 meprin(s) 1:691
 in angiogenesis and cancer 1:695
 in inflammation 1:694–695
 meprin A 1:735
 meprin α
 physiologic and pathologic role 1:694–695
 structural features of 1:691
 meprin B 1:735
 meprin β
 physiologic and pathologic role 1:694–695
 structural features of 1:691
 metalloproteases
 amino acid sequence alignment of 1:692F
 inhibition 1:691–694
 in neurobiology 1:696
 in skin and fibrosis 1:695–696
 structure and cleavage specificity of 1:693F, 1:694
 Mer protein *see* Merlin protein (Mer protein)
 meristems 2:324
 Merlin protein (Mer protein) 3:99–100
 MEROPS database 1:681
 MerR 1:477
 MerTK 3:374, 3:375
 MESD *see* mesoderm development (MESD)
 mesenchymal cells 2:720
 mesenchymal migration 2:720–723
 mesenchymal stem and progenitor cells (MSPCs) 3:625
 mesenchymal stem cells (MSCs) 2:308, 3:102, 3:620, 3:776, 3:819
 mesenchymal stromal cells (MSCs) *see* multipotent stromal cells (MSCs)
 mesenchymal-amoeboid transition (MAT) 3:226
 mesoderm development (MESD) 2:420–421
 mesoscale systems biology 4:51–55
 message-passing approach 4:212T, 4:215–216

- messenger RNA (mRNA) 1:58, 1:310, 1:384, 1:385–387, 1:385*F*, 1:386*F*, 1:388, 1:389*F*, 1:390, 1:390–392, 1:391*F*, 1:392–393, 1:392*F*, 1:393, 1:495, 1:525, 1:527, 1:527–529, 1:528*F*, 1:529, 1:530, 1:530–531, 1:531, 1:532, 3:56–57, 3:317, 3:434, 4:160, 4:191–192, 4:211–212, 4:213, 4:214
- cycle hypothesis 1:348–350
- degradation 1:346–347
- and disease 1:344
- eukaryotic 1:341, 1:342
- AUI-rich elements (AREs) 1:343–344
- cap 1:342
- microRNA response elements 1:344
- poly(A) tail 1:344
- processing 1:342
- stem loops 1:343–344
- 3' UTR 1:343
- 5' UTR 1:342–343
- export 3:174–175
- interaction with ribosome 1:384–385
- mRNA transcriptomics, new technologies for 4:163
- prokaryotic 1:341, 1:341–342
- polycistronic mRNA 1:342
- Shine–Dalgarno sequence 1:341–342
- in protein synthesis 1:526
- in regulation of protein synthesis 1:533
- stability 1:348
- surveillance mechanisms 3:317, 3:319*F*
- transport 2:624–625
- turnover 1:347–348
- metabolic disease association studies, SRs in 3:735
- metabolic disorders 2:532–533
- metabolic labeling 1:128–130, 4:182
- metabolic networks 4:213–214
- metabolic regulation 1:288–297
- allosteric modulation 1:289–290
- covalent modification 1:290
- energy homeostasis 1:289, 1:294, 1:295
- enzymes 1:289
- fasted (starvation) state 1:294
- action of glucagon 1:294
- glucose metabolism in liver 1:295–296
- prolonged starvation 1:294–295
- fed state 1:290
- action of insulin 1:290
- glucose metabolism in liver 1:290–291
- glucose homeostasis 1:289, 1:294, 1:295–296, 1:296
- glycogen metabolism 1:289*F*, 1:291, 1:291–292, 1:295–296, 1:295*F*
- glycolysis 1:289*F*, 1:292–293, 1:293*F*
- hormonal action 1:290, 1:294
- inhibitory protein (regulatory protein) interaction 1:290
- lipogenesis 1:289*F*, 1:293*F*
- substrates (fuel) availability 1:289
- transcriptional and degradational control 1:290
- metabolic reprogramming 3:789–790
- metabolic stress response, nucleolar signaling networks in 3:263
- metabolic syndrome 2:805
- activating UPR signaling 2:533–534
- metabolism 4:24–35, 4:24
- connectivity theorems 4:28–29
- control coefficients 4:27–28
- rate-limiting steps 4:28
- enzyme response 4:26–27
- elasticity coefficients 4:27
- operational definition 4:27
- fundamentals 4:24–25
- dynamics 4:26
- reaction rates 4:26
- stoichiometric networks 4:24–25
- stoichiometry matrix 4:25–26
- metabolic motifs 4:29
- branched pathways 4:29
- control analysis 4:31–33
- futile cycles 4:29–30
- graphical understanding of feedback 4:31
- instability 4:33
- linear pathways 4:29
- negative feedback 4:30–31
- robustness and supply/demand 4:33
- pathway responses 4:27
- control 4:27
- metabolite affinity chromatography 1:72–73
- metabolite profiling 4:199, 4:207–209
- metabolomics vs. 4:199–200
- metabolites 4:24, 4:26, 4:28–29
- metabolome 1:194, 4:199, 4:200*F*, 4:201, 4:201*F*, 4:202, 4:203, 4:204*T*
- metabolomics 2:799, 4:199–210, 4:265–266
- conducting a successful study of 4:200–201
- analyze 4:202–205
- data interpretation 4:205
- design 4:201
- extract/inactivate enzymes 4:202
- GC–MS data 4:205–206
- growth 4:201–202
- harvesting/sampling 4:202
- deconvolution 4:205–206, 4:207
- vs. metabolite profiling 4:199–200
- separation techniques 4:203
- terminology 4:199
- metabotropic glutamate receptors (mGluRs) 2:151, 2:535
- metabotropic receptors 1:244, 1:248–249
- metachromatic leukodystrophy 1:772, 1:774, 1:780–781, 1:784, 1:784–785, 1:785
- metadata 4:247–248
- metagenomics 2:799
- metagenomics for human intestinal tract (MetaHIT) 2:802
- MetaHIT *see* metagenomics for human intestinal tract (MetaHIT)
- metal binding sites 1:123
- metal ions 1:19
- metallo 1:753
- metallophilic macrophages 3:761–762
- metalloproteases 1:19, 1:562
- ADAM and AMAMTS proteases 1:754–755
- matrix metalloproteinases 1:755–756
- metamyelocyte 3:751
- metaphase 3:480–481
- metaphysical chondrodysplasia *see* cartilage hair hypoplasia (CHH)
- MetAPs *see* methionyl amino peptidases (MetAPs)
- metastasis 1:753, 1:754, 1:755, 1:756, 2:728, 3:223
- metastatic melanoma (MM) 1:145
- metatranscriptomics 2:799
- Metazoa 1:315, 1:315–316, 1:491, 1:495, 1:495–496, 1:497–498, 1:501, 1:505, 1:545, 1:546, 3:1
- metazoans 3:67
- metformin (MF) 3:336
- Met-Glu-Leu-Thr (MELT) 3:473
- methanogenesis 1:19
- methionyl amino peptidases (MetAPs) 3:301
- methuosis 2:762
- 7-methyl guanosine 3:3–4
- methylation 1:87
- methylchloroformate (MCF) 4:203–204
- 5-methylcytosine 1:29
- methylpurine DNA glycosylase 1:446
- MetRS *see* Met-tRNA synthetase (MetRS)
- Met-tRNA synthetase (MetRS) 1:312–313, 1:317–318, 1:317*F*
- Metropolis algorithm 4:124
- metzincins 1:621–622, 1:627, 1:630, 1:636
- mevalonate 1:174–175
- Mexican trains 2:723
- MF *see* metformin (MF); molecular function (MF); myelofibrosis (MF)
- Mff *see* mitochondrial fission factor (Mff)
- MFG-E8 *see* milk fat globule-epidermal growth factor-like 8 (MFG-E8)
- Mfn1 2:279–281
- mitochondrial fusion 2:288–289
- Mfn2 2:279–281
- mitochondrial fusion 2:288–289
- MFS *see* Marfan syndrome (MFS)
- MFs *see* microfilaments (MFs)
- MGDG *see* monogalactosyldiglyceride (MGDG)
- mGluRs *see* metabotropic glutamate receptors (mGluRs)
- MH1 domain *see* Mad Homology 1 domain (MH1 domain)
- MH2 *see* C-terminal mad homology 2 (MH2)
- MHC class II 3:6, 3:588, 3:594
- MHC *see* major histocompatibility complex (MHC)
- MHC-II compartment (MIIC) 3:590
- MHV *see* mouse hepatitis virus (MHV)
- MI *see* mutual information (MI)
- MIAME *see* minimum information about micro-array experiment (MIAME)
- MIAMI cells *see* marrow-isolated adult multilineage inducible cells (MIAMI cells)
- MIB *see* mitofusin binding protein (MIB)
- Michael acceptor function 1:131
- Michaelis–Menten kinetics 1:14–15, 1:15
- Michaelis–Menten model 1:8–9
- microarray 4:160, 4:162

- microautophagy 1:618, 1:764
 microbicidal immune system 2:4
 microbicidal mechanisms
 autophagy 2:771–772
 cell death mechanisms and host defense 2:772–773
 apoptosis 2:773
 NETs and NETosis 2:773, 2:773F
 pyroptosis 2:772–773, 2:772F
 oxidative host defense mechanisms 2:768F
 pathogen subversion of host defense 2:773–774
 phagocytes 2:766–767
 phagocytic killing 2:767
 oxygen-dependent killing 2:767
 oxygen-independent killing 2:771
 phagosome maturation 2:768–770
 phagocytosis 2:766–767
 mechanisms 2:766–767
 receptors for 2:766F
 xenophagy 2:772, 2:772F
 microbiome 2:4, 2:799–808
 changes throughout life 2:801, 2:801F
 adulthood 2:802
 aging 2:802
 birth and infancy 2:801–802
 pregnancy 2:801
 constituents 2:800
 and disease 2:804
 bariatric surgery 2:805
 C. difficile Infection 2:804
 cancer 2:805–806
 dysbiosis in disease states 2:804F
 IBD 2:805
 malnutrition 2:805
 metabolic syndrome 2:805
 obesity 2:804–805
 future microbiome research 2:806
 influencing factors 2:802–803
 antibiotics 2:803–804
 diet and geographic location 2:802–803
 exercise 2:803
 research methodology 2:799
 roles 2:800–801
 microcapillary arrays 3:669
 microcephalic osteodysplastic primordial dwarfism type I (MOPD I) 1:508
 microclusters 3:522–524, 3:682
 microengraving 3:667, 3:669, 3:673–674
 microfibril associated glycoproteins (MAGPs) 2:700
 microfilaments (MFs) 2:579, 2:580T, 2:704
 microfluidic devices 1:50, 3:667, 3:667–669, 3:671–672, 3:672, 3:673–674
 microfluidics 1:50, 3:672–673, 3:672, 3:674
 ultraviolet light-emitting diode-mediated DNA synthesis in glass capillaries 1:50
 microhomology-mediated end joining (MMEJ) 1:452, 1:461
 microparticle (MP) 1:716
 micropinocytosis 2:4
 micropinosomes 2:758
 microprocessor complex 1:354, 1:355
 microRNAs (miRNAs) 1:354, 1:364, 2:530, 3:45–46, 3:321–322, 3:753, 3:788, 4:160, 4:192
 biogenesis 1:364–365
 alternative biogenesis pathways of 1:359
 cleavage of pre-miRNAs into miRNA-miRNA* duplexes by Dicer 1:355–358
 domain structures of miRNA pathway proteins 1:358
 Drosha cleavage 1:355
 exportin-5/Ran-GTP complex transporting pre-miRNAs from nucleus to cytoplasm 1:355
 in animals 1:354–355
 miRNA-miRNA* duplexes, Argonaute loading of 1:358
 transcription by RNA Polymerase II 1:355
 suppression of target mRNAs 1:358–359
 turnover of 1:359
 and cancer 1:366
 etiology 1:366
 predicting prognosis and response to therapy 1:366–367
 in carcinogenesis 1:367
 discovery 1:364
 mutations in miRNA genes and diseases 1:360
 mutations in miRNA pathway genes and diseases 1:360
 nomenclature of 1:360–361, 1:365–366
 in plants 1:359–360
 processing-bodies (P-bodies) 1:359
 as a target for therapy 1:367–368
 microscope, SIM 2:89
 band separation/unmixing 2:89–90
 image preprocessing 2:89
 inverse Weiner filtering and apodization 2:90
 structured illumination pattern generation 2:89
 microscopic processes
 and parameters 4:56–57
 regulation 4:59–60
 microtentacles 3:227–228
 microtubules
 for interrogating immune cell migration 3:670–671F
 as part of integrated workflow 3:673–674
 microtubule blades (MB) 2:649
 microtubule doublets (MTD) 2:626–627
 microtubule organizing center (MTOC) 2:246, 2:476, 2:539, 2:549, 2:746, 2:824, 3:678
 microtubule-associated proteins (MAPs) 2:540–541, 2:543, 2:543F, 2:584, 2:617, 3:232–233, 4:36, 4:36–37, 4:42
 cellular microtubule dynamics and 4:38–40
 functions in cells 2:542F
 MAP1b 3:237–238
 role in mitotic spindle assembly 2:639
 microtubule-based motors 2:4
 microtubule-interacting and transport motifs (MIT motifs) 1:671
 microtubule-organizing center (MTOC) 2:168, 2:225, 2:412, 2:656–657, 2:725, 2:788–789
 microtubule-targeting agents (MTAs) 4:42
 microtubules (MTs) 2:27–28, 2:4, 3:204, 3:204–205, 3:205, 2:168, 2:539–547, 2:608, 2:704, 2:725, 2:753, 2:761, 3:52, 4:36
 see also actin
 arrangement into higher-order structures by MAPs 2:544–545
 array 2:655
 assembly and disassembly 2:541
 catastrophe 2:541–542
 cell-cell adhesion and 2:708–709
 cortical signaling regulation 2:710
 functional contributions at junctions 2:709–710
 interactions in epithelial cells 2:709F
 linkages with cell-cell junctions 2:708–709
 MT-dependent cellular transport 2:709–710
 cellular functions 2:541–543, 4:36, 4:37F
 cross-linkers 2:639
 cytoskeleton 2:584–585
 depolymerization 2:544
 depolymerization-promoting proteins 2:640
 dynamics 2:541
 regulation by MAPs 2:544
 E-site 4:36, 4:37–38, 4:39F, 4:42
 formation from alpha/beta-tubulin heterodimers 2:540F
 functions in cells 2:542F
 Gs α decoration 3:52
 in human disease 2:545
 microtubule dynamics 2:639
 minus ends 2:649
 motor proteins regulation 3:238
 motors 2:620
 N-site 4:36
 nucleation 2:653
 polymerization-promoting proteins 2:639–640
 rescue 4:36, 4:36–37, 4:37–38, 4:38–40, 4:39F, 4:40, 4:40F
 self-assembly dynamics 4:36–38, 4:39F
 and MAPs 4:38–40
 dynamic instability 4:36–38
 microtubule-targeting agents (MTAs) attenuate 4:42
 plus-end tip structures 4:38
 theoretical framework for 4:40–42
 -severing proteins 2:640
 structure 2:539–541, 4:36, 4:38F
 and mechanics regulation by MAPs 2:543–544
 turnover 2:640
 microvesicles 2:738
 microvillus inclusion disease 2:192
 Mid49 2:284–285
 see also mitochondrial dynamics protein of 49 (Mid49)

- Mid51 *see* mitochondrial dynamics protein of 51-kDa (Mid51)
- Mid51/MIEF1 2:284–285
- midbody 2:644–645
formation 3:496–497
remnant 3:499
- midbody ring (MR) 3:484
- midkine 3:16
- MIEF1 *see* mitochondrial elongation factor 1 (MIEF1)
- migration 3:741–742, 4:111–112
- migratory DCs (mDCs) 3:742, 3:742
- migratory pre-DCs 3:743
- milk fat globule-epidermal growth factor-like 8 (MFG-E8) 2:753–754, 3:375
- MIM *see* missing-in-metastasis (MIM); mitochondrial inner membrane (MIM)
- mimcan 1:273–274
- MIMIX *see* minimum information about molecular interaction experiment (MIMIX)
- MINDy algorithm *see* modulator inference by network dynamics algorithm (MINDy algorithm)
- mini-chromosome maintenance (MCM) 3:782
- mini-chromosome maintenance proteins (MCM proteins) 3:458
- minimum information about micro-array experiment (MIAME) 4:247–248
- minimum information about molecular interaction experiment (MIMIX) 4:190
- minimum information standards 4:247–248
- mini-organelle 2:819
(+) strand RNA generating 2:819
- minus end-capping proteins 2:539, 2:544
- minus ends 2:708
- miR-122 1:365, 1:366, 1:368
- miR-124 1:366
- mir-15* and *mir-16-1* 1:366, 1:367
- miR-34 1:367
- miRanda 1:365
- miRBase 1:364, 1:366, 1:368
- miRISC *see* miRNA-induced silencing complex (miRISC)
- miRNA-induced silencing complex (miRISC) 1:354, 1:356–357F, 1:357F, 1:358, 1:358–359, 1:359, 1:360
- miRNA–RNA Induced Silencing Complex (RISC) 1:344
- miRNAs *see* microRNAs (miRNAs)
- miRtrons 1:359, 1:364
- mischarged tRNA 1:310
- misfolding 1:111–112
- mismatch repair (MMR) 3:461, 3:463–464, 3:465F
- missense polymorphism 1:700
- missing pyramid 2:24–25, 2:25–26
- missing wedge 2:24–25, 2:25–26
- missing-in-metastasis (MIM) 2:497
- MIT motifs *see* microtubule-interacting and transport motifs (MIT motifs)
- mitochondria 2:279, 3:294, 3:336, 3:341, 3:382, 3:385, 3:721
- fatty acid metabolism and 3:814–815
- mitochondrial fission
Drp1 2:284
Mff, Mid49, and Mid51/MIEF1 2:284–285
modulators 2:285
molecular insights into Drp1 actions during 2:284
regulation of Drp1 by posttranslational modifications 2:287
- mitochondrial fusion
machinery and regulation 2:280F
Mfn1 and Mfn2 2:279–281
mitofusin proteins modulators 2:281–282
Opa1 2:282–284
- mitochondrial membrane
permeabilization 3:347F
MOMP 3:345–347
MPT 3:347–348
- physiologic importance of mitochondrial dynamics
Drp1-dependent mitochondrial fission 2:289
Mfn1-, Mfn2-, and Opa1-dependent mitochondrial fusion 2:288–289
- power plants 3:341–342
- regulating cell death 3:342
apoptosis mitochondrial control 3:342–343, 3:344F
necrosis mitochondrial control 3:343–345, 3:346F
- mitochondria-associated membranes (MAM) 2:159, 2:281
- mitochondria-derived vesicles (MDVs) 2:287–288
- mitochondrial antiviral signaling protein (MAVS. protein) 3:520, 3:716
- mitochondrial calcium uniporters (MCUs) 2:159
- mitochondrial DNA (mtDNA) 2:279, 3:341
information content of 1:545
translational system, components of 1:545–546
mitochondrial mRNAs 1:545–546
mitochondrial ribosomes 1:548–549
mitochondrial tRNAs 1:546
- mitochondrial dynamics protein of 49 (Mid49) 2:284–285
- mitochondrial dynamics protein of 51-kDa (Mid51) 2:284–285
- mitochondrial elongation factor 1 (MIEF1) 2:284–285
- mitochondrial endoplasmic reticulum interactions 1:238–240
- mitochondrial fission
Drp1 2:284
Mff, Mid49, and Mid51/MIEF1 2:284–285
modulators 2:285
endophilin B1 2:286
Fis1 2:285
GDAP1 2:285–286
LLO 2:287
sacsin 2:286–287
- molecular insights into Drp1 actions during 2:284
regulation of Drp1 by posttranslational modifications 2:287
interaction with cardiolipin 2:288
O-GlcNacylation 2:287
phosphorylation 2:287
S-Nitrosylation 2:287
SUMOylation 2:287–288
ubiquitination 2:288
- mitochondrial fission factor (Mff) 2:284–285
- mitochondrial fusion
machinery and regulation 2:280F
Mfn1 and Mfn2 2:279–281
mitofusin proteins modulators 2:281–282
Opa1 2:282–284
- mitochondrial inner membrane (MIM) 2:279
- mitochondrial membrane permeabilization 3:347F
MOMP 3:345–347
MPT 3:347–348
- mitochondrial mRNAs 1:545, 1:545–546, 1:546, 1:550, 1:551F
- mitochondrial myopathy, encephalopathy, lactic acid acidosis, and stroke like episodes (MELAS) 1:334, 1:547–548
- mitochondrial outer membrane (MOM) 2:279, 3:379
- mitochondrial outer membrane
permeabilization (MOMP) 3:330, 3:330–331, 3:331–332, 3:335, 3:342, 3:345–347, 3:368–369, 3:370, 3:379, 4:218, 4:221
- mitochondrial permeability transition (MPT) 3:342, 3:347–348
- mitochondrial processing peptidase (MPP) 2:282–283
- mitochondrial protein biosynthesis, mechanism of 1:549
initiation 1:549–550
mammalian mitochondrial system 1:549–550
in yeast mitochondria 1:550
polypeptide chain elongation 1:550–551
termination and ribosome recycling 1:551–552, 1:552F
- mitochondrial ribosomes 1:545, 1:548–549, 1:550, 1:552–553
cryo-EM structures of 1:549F
properties of 1:548F
- mitochondrial tRNAs 1:315–316, 1:316F, 1:333, 1:546
and aaRS in human disease 1:547–548
aminoacylation of 1:547
genetic code in mitochondria 1:546
and human diseases 1:333–334
import of tRNAs 1:546
structural features of 1:546–547
structure of 1:316F
- mitochondrial β -oxidation 2:252
- mitochondrial-associated membranes 2:531–532
- mitochondrial–nuclear interactions 1:240

- mitochondria-peroxisome connection 1:240
- mitofusin binding protein (MIB) 2:282
- mitofusin proteins modulators 2:281–282
- mitogen-activated protein (MAP) 4:141–142
- mitogen-activated protein kinase (MAPK)
1:708–709, 1:766, 1:771, 2:219,
2:372, 2:445–446, 2:667–668,
2:795, 3:2, 3:51, 3:70, 3:122–127,
3:236, 3:434, 3:482, 3:574–575,
3:712–714, 4:44, 4:44–45, 4:45,
4:47F, 4:105–106
- components of 3:122–123
ERK1/2 cascade 3:123
ERK5 cascade 3:124
JNK cascade 3:124
p38 cascade 3:123–124
- diseases in 3:126
MAPK and cancer 3:126
MAPKs, in developmental diseases
3:126–127
MAPKs, in inflammation-related
diseases 3:127
- MAPK1 1:771, 3:124
- regulation of 3:124–125
crosstalk with 3:125–126
phosphatases 3:125
scaffold proteins 3:125
subcellular localization of MAPK
components 3:126
- mitogen activated protein kinase/
extracellular kinase (MAPK/ERK)
pathway 1:766
- mitogen-activated protein kinase kinase 7
(MAP2K7) 3:237–238
- mitogen-activated protein kinase kinase
family (MAPKK family) 2:445–446,
2:795
- mitogen-activated protein kinase kinase
kinase (MAPKKK) 2:445–446, 3:62,
3:357–358
- mitosis 2:625, 2:637, 2:640, 2:641–642,
3:399, 3:432, 3:471, 3:478–493,
3:478, 4:104, 4:110
- animal cell cycle 3:478F
- as anticancer drug target 3:486
- cell fates 3:400F
- centromere–kinetochore association 3:481F
- CPC during 3:470–471
- NE disassembly and reassembly
closed mitosis 2:315
disassembly and reassembly 2:317
importin α/β and Ran system 2:316F
open mitosis 2:315
RanGTP gradient 2:315–317
- regulation of mitotic exit 3:401F
- stages 3:479F
anaphase 3:481
cytokinesis 3:481–484
metaphase 3:480–481
prometaphase 3:480
prophase 3:479–480
telophase 3:481
- targeted antimitotic compounds
3:487–489T
- unscheduled entry into 3:399–401
- mitotic arrest deficient 1 (MAD1) 2:625
- mitotic block 3:401–402
- mitotic catastrophe 3:399–403, 3:399, 3:399
inducing mitotic block 3:401–402
inducing unscheduled entry into mitosis
3:399–401
mechanisms 3:402
- mitotic centromere-associated kinesin
(MCAK) 2:544, 2:611–612
- mitotic checkpoint *see* spindle assembly
checkpoint
- mitotic checkpoint complex (MCC) 3:471,
3:482
- mitotic cyclins *see* B-type cyclins
- mitotic kinesin-like protein-1 (MKLP1)
3:475
- mitotic slippage 3:401–402, 3:402
- mitotic spindle 2:384, 2:741, 2:747
architecture and dynamics 2:637–638
disassembly and cytokinesis 2:644–645
function 2:641
chromosome congression 2:641–644,
2:642F
chromosome separation 2:643F, 2:644
error correction 2:644
paths of spindle assembly 2:638–639
properties 2:640
microtubule turnover 2:640
poleward microtubule flux 2:640–641
spindle length control 2:641
visco-elastic nature 2:641
role of motors and MAPs in mitotic
spindle assembly 2:639
microtubule cross-linkers 2:639
microtubule depolymerization-
promoting proteins 2:640
microtubule polymerization-promoting
proteins 2:639–640
- MIWI *see* murine PIWI protein (MIWI)
- mixed lineage kinase domain-like (MLKL)
3:342, 3:393–394
- MJDs *see* Machado–Joseph disease (MJDs)
- MKK4 3:123F, 3:124
- MKK7 3:123F, 3:124
- MKLP1 *see* mitotic kinesin-like protein-1
(MKLP1)
- MKP6 3:125
- MKS *see* Meckel–Gruber Syndrome (MKS)
- ML *see* maximum likelihood (ML)
- MLC *see* myosin light chain (MLC)
- MLCK *see* myosin light chain kinase (MLCK)
- ML-IAP *see* melanoma IAP (ML-IAP)
- MLKL *see* mixed lineage kinase domain-like
(MLKL)
- Mll1 (mixed lineage leukemia 1) 3:790–791
- MLN-4760 1:749
- mLST8 *see* mammalian Lethal-with-Sec-
Thirteen 8 (mLST8)
- MM *see* metastatic melanoma (MM)
- Mm-Cpn 2:8F, 2:9, 2:11, 2:11–12, 2:11F
- MMEJ *see* microhomology-mediated end
joining (MMEJ)
- MMLV *see* Moloney murine leukemia virus
(MMLV)
- Mmm1 2:281–282
- MMP *see* matrix metalloproteinase (MMP)
- MMP5 *see* membrane-associated
palmitoylated protein 5 (MMP5)
- Mmp-9 null mice 1:624
- MMP-13 1:624
- MMPs *see* matrix metalloproteinases
(MMPs)
- MMR *see* macrophage mannose receptor
(MMR); mismatch repair (MMR)
- MMTV *see* mouse mammary tumor virus
promoter (MMTV)
- MNDA *see* myeloid cell nuclear
differentiation antigen (MNDA)
- Mob as tumor suppressor (Mats) 3:99
- mobile enhancer 1:489
- mobile genetic elements 1:377–380, 1:382
ribozyme as 1:380
- mobilization, CXCL12 in HSC and HPC
3:626–627
- mo-DCs *see* monocytes DCs (mo-DCs)
- model simulations 4:72
- model-based approaches 4:255–256
- modeling 4:72
multiscale *see* multiscale modeling
- modeling single-cell transcriptomic data
4:223
- modular response analysis 4:46–47
- modular signaling system 3:200–202
- modularity 4:68–70, 4:69F, 4:194
- modulated image intensifiers, frequency-
domain FLIM by 2:116
- modulating constitutive transport
2:363–364
- modulator inference by network dynamics
algorithm (MINDy algorithm)
4:258–259
- modulators 1:665
- modulatory 1 domain (M1 domain)
3:184–185
- modules 4:68
- MOGS *see* membrane-bound α -glucosidase
(MOGS)
- molecular aspects of cell biology
molecular components and technology
1:1–2
nucleic acid synthesis/breakdown 1:2–3
protein synthesis and degradation 1:3
chemical principles and molecular
components 1:4
ER 1:3
protease inhibitors 1:3–4
proteolytic systems 1:3
- molecular biological techniques 1:485
- molecular chaperones 1:535, 1:536F, 1:604
cytosolic 1:538
glycan-recognizing 1:540–542
- molecular crowding effect in living cells
1:11–12
- molecular function (MF) 4:247
- molecular imaging 1:128
- molecular labeling density 2:65
- molecular machinery
for filopodial protrusion 2:687–688,
2:687F
actin cytoskeleton 2:687–688
membrane deformation 2:688
lamellipodia 2:690F, 2:691

- molecular mimicry 1:531–532, 3:511–512
 molecular motors 2:413–415, 2:541, 2:602, 2:608, 3:230–231, 3:233
 actin-based 2:564
 dynein 3:234
 kinesin 3:233–234
 motor proteins regulation 3:234–235
 myosin 3:234
 molecular recognition 1:105
 molecular signaling pathways 3:511–512, 3:512
 molecular size hypothesis 3:660
 molecular weight (MW) 1:741
 Moloney murine leukemia virus (MMLV) 1:513
 molten globule 1:110
 molybdopterin 1:19
 MOM *see* mitochondrial outer membrane (MOM)
 MOMP *see* mitochondrial outer membrane permeabilization (MOMP)
 monoamine oxidase 1:237
 mono-carboxylate transporter (MCT) 3:338
 monoclonal antibodies 3:17
 monocyte chemoattractant protein 1 (MCP-1) 3:219, 3:620, 3:636–637
 monocytes DCs (mo-DCs) 3:742, 3:742
 Monod model 1:9–10
 monodomain equations 4:74
 monofunctional DNA glycosylases 1:446, 1:446F
 monogalactosyldiglyceride (MGDG) 1:181
 monogenic diseases, SRs in 3:736
 monomeric adaptors 2:453F
 monopolar spindle 1 (Mps1) 3:473
 monotopic membrane protein 2:273
 monoubiquitination, polyubiquitination vs. 3:505
 MOPD I *see* microcephalic osteodysplastic primordial dwarfism type I (MOPD I)
 morphogenesis, junctional contractility and 2:706–707
 morphogenesis, multiscale analysis of 4:122–133
 analyzing cell behavior 4:122–123
 mathematical models for studying collective cell behavior 4:123
 cell-ECM interactions, modeling 4:125
 continuum models 4:123
 discrete models 4:123–124
 vertical integration 4:122
 differential-adhesion driven cell sorting 4:126–127
 multiscale modeling of somitogenesis 4:128
 regulation of root growth by auxin in *Arabidopsis thaliana* 4:128–131
 studies, examples of 4:125–127
 vasculogenesis and angiogenesis 4:127–128
 morphology 2:761
 Morquio disease 1:782
 mosaic serine protease large-form (MSPL) 1:655
 mother centriole (MC) 2:649, 2:650F, 2:651, 2:653, 2:655, 2:656
 Mothers against Dpp (Mad) 3:99
 motif ten element (MTE) 1:481–482
 motifs 1:91, 4:65, 4:65F
 cross-repression 4:66
 feedforward loops 4:67, 4:67F
 negative feedback 4:66–67
 p-value 4:213
 positive feedback 4:67
 repression 4:65–66
 motile cilia 2:660–661
 diseases caused by defective 2:669–670
 motility 2:663, 4:4
 dyneins
 AAA + domains 2:622–623
 dynein directionality mechanism 2:624
 force generation 2:623–624
 regulation 2:624
 stepping 2:621–622
 motor proteins 3:230–231, 3:231
 organization 2:92–93
 regulation 3:234–235
 motor-cargo relationship 2:611
 axonal transport abnormalities 2:611
 ciliogenesis and left-right determination 2:611
 depolymerizing kinesins and brain wiring 2:611–612
 kinesin-2 dependent ciliogenesis 2:613F
 kinesins
 and cancer 2:612–614
 as components of signal transduction 2:614
 mouse double minute 2 (MDM2) 3:282, 3:504
 mouse embryonic fibroblasts (MEFs) 1:442, 3:243–244, 3:449–450
 mouse hepatitis virus (MHV) 2:169
 mouse knockout studies (KO studies) 2:279–281
 mouse mammary tumor virus promoter (MMTV) 3:447
 mouse models 1:641–642, 3:438
 MP *see* microparticle (MP)
 MP1 *see* MEK1 partner (MP1)
 Mpdz *see* multiple PDZ domain protein (Mpdz)
 MPF *see* maturation promoting factor (MPF)
 MPM *see* multiphoton microscopy (MPM)
 M-PMV *see* Mason-Pfizer monkey virus (M-PMV)
 MPN *see* myeloproliferative neoplasms (MPN)
 MPO *see* myeloperoxidase (MPO)
 MPP *see* mitochondrial processing peptidase (MPP)
 MPR *see* Man-6-P receptor (MPR); mannose-6-phosphate receptor (MPRs/M6PR)
 MPS *see* mucopolysaccharidosis (MPS)
 MPS VI *see* Maroteaux-Lamy disease (MPS VI)
 Mps1 *see* monopolar spindle 1 (Mps1)
 MPS1 *see* multipolar spindle-1 (MPS1)
 MPS6 (Maroteaux-Lamy disease) 1:782
 MPT *see* mitochondrial permeability transition (MPT)
 MR *see* midbody ring (MR)
 MR genes *see* master regulator genes (MR genes)
 MRCs *see* marginal reticular cells (MRCs)
 MreB filaments 2:93–94
 MRH domains *see* Man-6-P receptor homology domains (MRH domains)
 MRI *see* magnetic resonance imaging (MRI)
 MRM *see* multiple reaction monitoring (MRM)
 mRNA *see* messenger RNA (mRNA)
 MS *see* mass spectrometry (MS); multiple sclerosis (MS)
 MS1 quantification 1:568, 1:572F, 1:573–574, 1:573F, 1:576
 MS2 quantification 1:568, 1:572F, 1:573–574, 1:573F, 1:575–576, 1:576
 MS3 quantification 1:568, 1:573–574, 1:573F
 MSA *see* multi-stage-activation (MSA); multiple sequence alignments (MSAs)
 MSCs *see* mesenchymal stem cells (MSCs); multipotent stromal cells (MSCs)
 MSD *see* mean square displacement (MSD)
 MSMs *see* medullary sinus macrophages (MSMs)
 MSPCs *see* mesenchymal stem and progenitor cells (MSPCs)
 MSPL *see* mosaic serine protease large-form (MSPL)
 MSSE *see* multiple self-healing squamous epithelioma (MSSE)
 MSTFA *see* N-Methyl-N-trimethylsilylfluoroacetamide (MSTFA)
 MT *see* membrane type (MT)
 MT1-MMP *see* membrane-type 1 matrix metalloprotease (MT1-MMP)
 MTAs *see* microtubule-targeting agents (MTAs)
 Mtb *see* *Mycobacterium tuberculosis* (*M. tuberculosis*)
 MTBD *see* MT-binding domain (MTBD)
 MT-binding domain (MTBD) 2:620
 MTCs *see* multisubunit tethering complexes (MTCs)
 MTD *see* microtubule doublets (MTD)
 mtDNA *see* mitochondrial DNA (mtDNA)
 MTE *see* motif ten element (MTE)
 mTEC *see* medullary thymic epithelial cells (mTEC)
 Mtl *see* mannitol (Mtl)
 MTM1 *see* myotubularin (MTM1)
 MTMR1–14 *see* myotubularin-related proteins 1–14 (MTMR1–14)
 mTNF *see* membrane-bound TNF (mTNF)
 MTOC *see* microtubule organizing center (MTOC); microtubule-organizing center (MTOC)
 mTOR *see* mammalian target of rapamycin (mTOR)

- mTORC *see* mechanistic target of rapamycin complex (mTORC)
- mTORC1 *see* mammalian target of rapamycin complex 1 (mTORC1)
- MTs *see* microtubules (MTs)
- mucin-type O-glycans 2:347–348, 2:348F
- mucociliary clearance 1:252
- mucopolidosis types II and III 1:773, 1:773–774
see also I-cell disease
- mucopolysaccharide disorder 1:777, 1:780, 1:782
- mucopolysaccharidoses 1:768, 1:769, 1:773, 1:777–778, 1:779, 1:780, 1:781, 1:782
- MPS4a (Morquio disease) 1:782
- mucopolysaccharidosis (MPS) 1:769, 1:777, 1:782
- mucosal addressin cell adhesion molecule (MAdCAM) 3:526
- mucosal vascular addressin cell adhesion molecule 1 (MAd-CAM-1) 3:635–636, 3:752
- MultiCellDS 4:122–123
- multicellular migration 2:720, 2:723
see also single-cell migration
- multicolor fluorescence microscopy 2:83
- multicolor imaging 2:64
- multidimensional TCSPC, time-domain FLIM by 2:114–115
- multidrug resistance (MDR) 1:744
- multilineage differentiating stress enduring cells (MUSE cells) 3:819
- multinodular goiter 1:360
- multi-organelle interactions 2:780
- multiphoton excitation 2:129
- multiphoton microscopy (MPM) 2:129
- multiple functions of Gag 2:831F
- multiple patch cloning (MUPAC) 1:122
- multiple PDZ domain protein (Mpdz) 3:102
- multiple reaction monitoring (MRM) 1:150, 4:168
- multiple regulatory proteins 2:677
- multiple sclerosis (MS) 2:728–729, 3:642
- multiple self-healing squamous epithelioma (MSSE) 3:47
- multiple sequence alignments (MSAs) 3:61
- multiple signal pathways, DRs signaling and 3:69
 extrinsic apoptotic pathway 3:69
 intrinsic apoptotic pathway 3:69–70
 non-apoptotic pathways 3:70
- multiple sulfatase deficiency 1:774
- multiple time and spatial scales in cellular biology 4:72–73
- multipolar spindle-1 (MPS1) 3:482
- multipotent adult progenitor cells (MAPC) 3:819
- multipotent stromal cells (MSCs) 3:780
see also hematopoietic stem cells (HSCs)
 age-related changes in 3:780
- multiscale modeling 4:72, 4:73F, 4:75–76, 4:75, 4:80–85, 4:80, 4:81
 adaptive combined cell-based and continuum approach 4:83
- applications of 4:76–77
 diabetic retinopathy 4:77–78
 homeostasis in intestinal crypt 4:76–77
- approaches for linking scales in 4:73
- assembling 4:74–75
- ‘bottom-up’/molecular biology approach 4:73
- challenges 4:78
- complexity of biological system 4:83
- composite hybrid ABM 4:81–82
- current modeling methods 4:81
- DDFT 4:83
- EFA 4:82–83
- HMM 4:83
- mathematical and computational representations of 4:73–74
- mean field theory 4:83
- methods 4:81
- middle-out models 4:73
- software for 4:75–76
- standards in 4:84
- ‘top-down’/classical physiological approach 4:73
- multi-stage-activation (MSA) 4:180
- multisubunit tethering complexes (MTCs) 2:485, 2:487–488
 CATCHR Complexes 2:487–488
 HOPS and CORVET 2:488
 TRAPP complexes 2:488
- multivesicular bodies (MVBs) 2:201, 2:228, 2:303–304, 2:408, 2:520, 3:92, 3:475, 3:589–590, 3:591, 3:608F
 biogenesis 2:201–203, 2:409F, 2:466, 2:467F, 2:468, 2:470
 cargo sorting 2:408–410
- multiwavelength capability 2:114
- MUNC13 2:504, 2:510
- Munc13–4 protein 2:230
- MUPAC *see* multiple patch cloning (MUPAC)
- mural cells 4:104
- muramyl dipeptide (MDP) 3:717
- murine double minute 2 (MDM2) 3:130
- murine duodenum epithelium imaging 2:47
- murine embryonic fibroblast (MEF) 2:289
- murine PIWI protein (MIWI) 3:322
- muscle contraction 2:677, 2:677–679
 calcium release from SR 2:679–681
 sarcomeres 2:677
- muscle disorders 2:681
- muscle fibers 1:221, 1:224–225, 2:4
- muscle satellite cells 3:6–7
- muscle stem cells 2:677
 extrinsic factors 3:778–779
 intrinsic modifications in aged satellite cells 3:779–780
 loss of skeletal muscle regenerative potential during aging 3:778
- muscle–eye–brain disease (MEB) 2:348–349
- muscle-specific kinase (MuSK) 3:17
- muscular dystrophy 2:681
- MUSE cells *see* multilineage differentiating stress enduring cells (MUSE cells)
- MuSK *see* muscle-specific kinase (MuSK)
- mutations 4:214
- mutations in gene encoding MPO 3:754
- mutual exclusivity modules in cancer (MeMo in Cancer) 4:214
- mutual information (MI) 4:258
- mutual inhibition *see* cross-repression
- MutY homologue (MYH) 1:444–445
- MVBs *see* multivesicular bodies (MVBs)
- MVP *see* major vault protein (MVP)
- MW *see* molecular weight (MW)
- Myc control 3:458
- MYC oncogene 3:417
- Mycobacterium tuberculosis* 1:378T, 1:592, 2:405, 2:777
 immune response to 4:100
- MyD88 *see* myeloid differentiation primary response gene 88 (MyD88)
- MYD88-adaptor-like protein (MAL) 3:712–714
- myelin 2:248
- myelin basic protein (MBP) 3:606
- myeloblast 3:751
- myelocyte 3:751
- myelodysplasia 1:508
- myelodysplastic syndrome (MDS) 1:501, 1:508, 3:146
- myelofibrosis (MF) 3:146
- myeloid cell leukemia sequence 1 (MCL1) 3:345–346
- myeloid cell nuclear differentiation antigen (MNDA) 3:719–720
- myeloid differentiation primary response gene 88 (MyD88) 3:28, 3:394, 3:520, 3:712–714
- myelomatosis 1:773
- myeloperoxidase (MPO) 2:767, 3:751
- myeloproliferative neoplasms (MPN) 3:145–146
- MYH *see* MutY homologue (MYH)
- MYLIP *see* myosin regulatory light chain interacting protein (MYLIP)
- Myo2 2:261
- myocardial stunning 1:676
- myoclonic epilepsy 1:707, 1:772, 1:773
- myofibers 2:677
- myofibroblast 3:219–220, 3:515, 3:809
- myosin 1:776, 2:4, 2:27
- myosin light chain (MLC) 2:734–735, 3:483–484, 3:495
- myosin light chain kinase (MLCK) 3:162–163, 3:212, 3:691–692
- myosin regulatory light chain interacting protein (MYLIP) 2:427
- myosin tail homology 4/band 4.1, ezrin, radixin, and moesin (MyTH-FERM) 2:597, 2:603
- myosin(s) 2:367, 2:564, 2:597–607, 2:677, 2:761–762, 3:234
 class-1 2:601
 class-2 2:601–602
 class-3 2:602
 class-5 2:602–603
 class-6 2:603
 class-7 2:603
 class-8 2:603–604
 class-9 2:602
 class-10 2:603
 class-11 2:603–604

- class-14 2:604
 class-15 2:603
 class-16 2:604
 class-18 2:604
 class-19 2:604
 class-21 2:604
 contractility model 4:13–14
 human myosome 2:598F
 kinetic cycle of actomyosin interaction 2:600F
 mechanochemical ATPase cycle 2:601
 movement 2:679
 myosin II 2:367
 myosin-based contractility 2:759
 pathophysiology 2:601
 stroke cycle 2:679F
 structural domains 2:599F
 structure and organization 2:597–601
 myotonia congenita 1:231
 myotubularin (MTM1) 2:504, 2:516
 myotubularin-related proteins 1–14 (MTMR1–14) 2:504
 MYO-Vb 2:192
 myristoylation 1:87
 MyTH-FERM *see* myosin tail homology 4/
 band 4.1, ezrin, radixin, and moesin
 (MyTH-FERM)
 MZ *see* marginal zone (MZ)
- N**
- N nucleotides *see* nucleotide nucleotides
 (N nucleotides)
 NA *see* numerical aperture (NA)
 NAC *see* N-acetyl cysteine (NAC)
 N-acetyl cysteine (NAC) 3:816
 N-acetyl hexosamines 1:128–130
 N-acetylation 1:86
 N-acetylgalactosamine (GalNAc) 1:271,
 2:169, 2:347
 N-acetylglucosamine (GlcNAc) 1:271,
 2:339–340, 2:347
 N-acetylglucosaminidase (NAGPA)
 2:341–342
 NACHT, LRR, and PYD domains-containing
 protein 3 (NLRP3) 2:523–524
 N-acyl-D-amino acid amidohydrolase
 1:125
 NAD⁺/NADH ratio 1:288, 1:296
 NADH *see* nicotinamide adenine
 dinucleotide (NADH)
 NADPH *see* nicotinamide adenine
 dinucleotide phosphate (NADPH)
 NADPH oxidase 3:6
 NADPH oxidase *see* nicotinamide adenine
 dinucleotide phosphate-oxidase
 (NADPH oxidase)
 NADP-malate dehydrogenase 1:294
 NAFLD *see* nonalcoholic fatty liver disease
 (NAFLD)
 NAGPA *see* N-acetylglucosaminidase
 (NAGPA)
 NAIC *see* North American Indian childhood
 cirrhosis (NAIC)
- NAIPs *see* neuronal apoptosis inhibitor
 proteins (NAIPs)
 naïve T cell 3:677
 nanoclusters 3:656
 nanodomains 1:211, 1:213
 Nanog 3:786, 3:787, 3:788–789, 3:788
 nanophotonic and super-resolution
 approaches 2:101–102
 nanopore 4:161, 4:162
 NAPPA *see* nucleic acid programmable
 protein array (NAPPA)
 nascent polypeptides 1:322F, 1:380, 1:535,
 1:536–538
 delivery to the ER membrane 1:536–538
 formation of native structures in ER lumen
 with help of multiple factors
 1:538–540
 translocation across ER membrane by
 conducting channels 1:538
 Sec61 translocon 1:538
 signal peptide cleavage 1:538
 NASH *see* nonalcoholic steatohepatitis
 (NASH)
 National Institutes of Health (NIH)
 4:248–249
 native fold 1:599–601, 1:604
 native state 1:108–109, 1:109, 1:112
 natriuretic peptide(s) 3:58, 3:59
 family 1:750
 receptor B 3:58
 natronomonas pharaonis (NpHR) 2:151
 natural inhibitors 1:626–627
 natural killer (NK) cells 1:145, 2:226,
 2:835, 3:24–25, 3:63, 3:583
 migration of 3:672
 natural killer T cells (NKT cells) 3:748
 naturally-occurring polypeptide inhibitors
 1:706
 CSTs 1:706
 biological functions 1:707
 and disease 1:707
 structure and inhibition mechanism
 1:706–707
 tissue distribution 1:707
 IAPs 1:707–708, 1:708
 biological functions 1:708
 in cell division and cytokinesis 1:709
 in disease 1:709
 structure 1:708
 serpins 1:709
 activation and latency 1:711
 biological functions 1:711
 disease aspects 1:711
 mechanism of inhibition 1:710
 structure 1:709–710
 TIMPs 1:711
 biological functions 1:712
 disease aspects 1:712–713
 interactions with pro-MMPs 1:712
 mechanism of inhibition of MMPs
 1:711–712
 structures 1:711
 N-BAR domains *see* amphipathic N-terminal
 helix domains (N-BAR domains)
 NBD *see* nucleotide-binding domain (NBD)
 NBEAL2 2:230
- NC protein *see* nucleocapsid protein (NC
 protein)
 NC1 domain *see* non-collagenous 1 domain
 (NC1 domain)
 N-cadherin 2:704–705
 NCAM *see* neural cell adhesion molecule
 (NCAM)
 NCE *see* non-clathrin endocytosis (NCE)
 NCF3 *see* Rac
 NCL2 *see* neuronal lipofuscinosis Type 2
 (NCL2)
 NCLDV *see* nucleocytoplasmic large DNA
 viruses (NCLDV)
 NCR *see* nonconserved region (NCR)
 ncRNA *see* noncoding RNAs (ncRNA)
 NCS-1 *see* neuronal calcium sensor-1
 (NCS-1)
 NDR *see* nuclear Dbf-2-related (NDR)
 NE *see* nuclear envelope (NE)
 NE breakdown (NEBD) 2:315
 near surface swimming model 2:787
 near-infrared (NIR light) 2:129
 NEB *see* nuclear envelope breakdown (NEB)
 NEBD *see* NE breakdown (NEBD)
 NEC *see* necrotizing enterocolitis (NEC)
 necroptosis 3:371–372, 3:393–394
 necrosis 3:402
 mitochondrial control 3:343–345, 3:346F
 necrotizing enterocolitis (NEC) 2:804
 nectins 2:742
 NEDD8 (neural precursor cell expressed
 developmentally downregulated-8)
 1:584T, 1:587
 neddylation 1:708
 NEFs *see* nucleotide exchange factors (NEFs)
 negative feedback 4:30–31, 4:31, 4:31F,
 4:32, 4:32–33, 4:33, 4:33F, 4:66–67
 adaptation through regulation 4:7, 4:7F
 four-step pathway 4:31F
 loops 4:219
 two-step pathway 4:32F
 negative regulation modules 3:536–538,
 3:537F
 negative regulators exclusion 3:651–653
 negative regulatory region (NRR) 3:107
 negative regulatory repeats (NRR) 1:644
 NEIL glycosylases *see* endonuclease VIII-like
 glycosylases (NEIL glycosylases)
Neisseria meningitidis infections 2:307–308
 NEKs *see* NIMA-related kinases (NEKs)
 NEM-sensitive factor *see* N-ethylmaleimide-
 sensitive factor (NEM-sensitive
 factor)
 N-end rule 1:310, 1:326
 neo N-termini 1:568, 1:570–571,
 1:574–575, 1:575, 1:575–576,
 1:575F, 1:576, 1:577, 1:578F
 NEP *see* nephrilysin (NEP)
 nephronophthisis (NPHP) 2:664
 nephrilysin (NEP) 1:747
 NER *see* nucleotide excision repair (NER)
 NER pathway *see* nuclear excision repair
 pathway (NER pathway)
 Nernst equation 1:223, 1:223–224, 1:225,
 1:232F
 nerve cells 2:380

- nerve growth factor (NGF) 2:211, 3:51–52, 3:235, 4:223–224
 inducing neurite outgrowth through G protein 3:52
 association between G proteins and tubulin 3:52–53
 GPCRs 3:52
 Gsz 3:52
 Gsz-induced neurite outgrowth 3:52
 neurite outgrowth 3:52
 translocation of Gsz 3:53F
- NES *see* nuclear export sequences (NES); nuclear export signal (NES)
- nestin 2:575
- NET *see* neutrophil extracellular trap (NET)
- N-ethylmaleimide-sensitive factor (NEM-sensitive factor) 2:410–411
- N-ethylmaleimide-sensitive factor (NSF) 2:225–226
- NETosis 2:773, 2:773F
- NetWalk 4:212T, 4:213
- network algorithms
 complex, cancer-related genetic networks 4:214
 integrating single data source with an existing interaction network 4:212–213
 metabolic networks 4:213–214
 methods for integration of data types 4:214–216
 proteins and genes 4:216
- network biology principles 4:145–146
- network flow 4:212T, 4:213–214, 4:214, 4:215
- network medicine 4:266
 properties of networks 4:266–267
 targeting networks 4:267
- network motifs 4:63–64, 4:187
- network swelling model 4:12
- networks *in vitro* and/or *in vivo* 4:141
 cancer metastasis 4:143–144, 4:144F
 complex nature of disease and molecular networks 4:141–142
 networks in cancer 4:144–146
 tumor microenvironment 4:142–143, 4:142F
- NeuAc *see* sialic acid (NeuAc)
- Neufeld, Elizabeth 1:777, 1:777–778
- neural cell adhesion molecule (NCAM) 2:724
- neural development, migration in 3:209–210
- neural precursor cells (NPCs) 4:136
- neural repair 3:231
 injury changes axon structure and axonal transport 3:235–236
 calcium influx coordinating injury response 3:236
 calcium influx transforming 3:236–237
- proteins in growing axons
 APC protein 3:237
 axonal functions for nuclear proteins 3:237
 protein kinase regulation 3:237–238
 regulation of microtubule motor proteins 3:238
- neural stem cells (NSCs) 3:790–791, 3:816–817, 4:134
 identification 4:135–136
 lineage relationship with progeny 4:136F
- neuraminidase 1:774
- neuregulin-1 (NRG1) 1:663
- neurite outgrowth 3:52
 NGF inducing through G protein 3:52
 association between G proteins and tubulin 3:52–53
 GPCRs 3:52
 Gsz 3:52
 Gsz-induced neurite outgrowth 3:52
 translocation of Gsz 3:53F
- neurite severing in neuronal development 2:467F, 2:471
- neuro fibrillary tangles (NFTs) 1:675
- neuroblast 4:137
- neurodegeneration 1:439, 1:661, 1:666, 1:696, 1:763–764, 1:767, 1:781, 1:784, 1:784–785
- neurodegenerative diseases
 oncogenes and tumor suppressor genes in 3:419
 UPR activation 2:535–536
- neurodegenerative disorders 2:585
- neurodevelopmental disorders 2:585
- neurofibrillary tangles 1:661
- neurofibromatosis 3:418
- neurofibromatosis type1 (NF1) 3:418
- neurofibromin 2 (NF2) 3:102
- neurofilament proteins 2:569, 2:573–574
- neurogenesis 4:134–140
 applications 4:138–139
 cell proliferation in SVZ 4:136–137
 developmental 4:135
 functions of adult neurogenesis 4:138
 historical perspective 4:135
 identification of adult NSCs 4:135–136
 neurogenic regions 4:134F
- neuronal differentiation in olfactory bulb 4:137–138
- neuronal migration in RMS 4:137
- olfactory circuit diagram 4:135F
- in V-SVZ and SGZ 4:134
- neurological disorders 1:178
- neuro-muscular junction (NMJ) 1:672–673
- neuronal action potentials and ion channel allostery 1:244–251
 action potentials 1:246–247
 invasion of nerve terminal 1:248
 propagation 1:247–248
 allosteric effectors 1:249–250
 fate of neurotransmitter in synaptic cleft 1:248
 post synaptic receptors/ligand-gated ion channels 1:248–249
- neuronal apoptosis inhibitor proteins (NAIPs) 3:394, 3:714
- neuronal calcium sensor-1 (NCS-1) 2:504
- neuronal cancer stem cells 3:810
- neuronal cell adhesion molecule (NCAM) 4:137
- neuronal ceroid lipofuscinosis 1:767, 1:773
- neuronal cytoskeleton 3:231
 intermediate filaments 3:231–232
 microfilaments 3:231
 microtubules 3:232–233
- neuronal inhibitor apoptosis protein (NIAP) 1:707–708
- neuronal lipofuscinosis Type 2 (NCL2) 1:784
- neuronal NOS (nNOS) 3:57
- neuronal system 1:687
- neuronal Wiskott–Aldrich Syndrome protein (N-WASP) 2:491, 2:497, 2:705
- neurons 1:232–234, 1:244, 1:244–245, 2:380, 3:230
 anatomy of 1:245–246
 constituents for intracellular communication and trafficking in growth cone 3:233
 molecular motors 3:233–234
 neuronal cytoskeleton 3:231
 electrical properties of 1:245
 flow of information 1:245–246
 neural trafficking 3:232F
- neuropeptide 1:557, 1:559, 1:560, 1:562–563, 1:563
- neuropeptide-like proteins (NLPs) 2:848
- neuropilin (NRP) 4:102, 4:106
- neurotoxicity of prions 2:842–843
- neurotransmission 2:376
- neurotransmitter 1:201F, 1:248
 receptors 2:611
 in synaptic cleft 1:248, 1:248–249
 transporters 2:297
- neurotrophic growth factor (NGF) 4:46
- neurotrophic tyrosine kinase, receptor, type 3 (NTRK3) 3:343
- neutrophils 4:97, 4:97–98, 4:97F, 4:99
- neurotrophins (NT) 3:16
- neutral lipids (NLs) 2:273
 synthesis 2:276
- neutralized homology repeat (NHR) 3:60
- neutrophil biology 3:750–758
 genetic defects affect neutrophil numbers and/or function 3:754, 3:755F
 CGD 3:754
 LAD 3:754
 mutations in gene encoding MPO 3:754
 neutrophil-SGD 3:755
 PLS 3:754–755
 SCN 3:754
- neutrophil defense mechanisms 3:752
 cellular events during NETosis 3:753F
 degranulation 3:752
 NETs 3:752–753
 phagocytosis 3:752
- neutrophil development 3:750
 bands 3:751
 mature neutrophils 3:751
 metamyelocyte 3:751
 myeloblast 3:751
 myelocyte 3:751
 promyelocyte 3:751
- neutrophil migration into sites of inflammation 3:751
 firm adhesion 3:751–752
 tethering and rolling 3:751
 transmigration 3:752

- neutrophil morphology 3:750F
neutrophil subsets 3:753
 CD177-expressing neutrophils 3:753
 LDGs 3:753–754
 neutrophil heterogeneity 3:754F
 olfactomedin 4-expressing neutrophils 3:753
 proangiogenic MMP-9 delivering neutrophils 3:754
 T-cell regulatory neutrophils 3:754
neutrophils in systemic autoimmune diseases 3:755
 RA 3:755–756
 SLE 3:756
 vasculitis 3:755
neutrophil cytosolic factor 1 (NCF1) *see* P47phox
neutrophil cytosolic factor 2 (NCF2) *see* p67phox
neutrophil cytosolic factor 4 (NCF4) *see* P40phox
neutrophil defense mechanisms 3:752
 cellular events during NETosis 3:753F
 degranulation 3:752
 NETs 3:752–753
 phagocytosis 3:752
neutrophil extracellular trap (NET) 3:752, 3:752–753
neutrophil gelatinase-associated lipocalin (NGAL) 3:751
neutrophil subsets 3:753
 CD177-expressing neutrophils 3:753
 LDGs 3:753–754
 olfactomedin 4-expressing neutrophils 3:753
 proangiogenic MMP-9 delivering neutrophils 3:754
 T-cell regulatory neutrophils 3:754
neutrophil(s) 2:226, 2:835, 3:6
 heterogeneity 3:754F
 phagosome maturation 2:769F, 2:770–771
 in systemic autoimmune diseases 3:755
 CXCL12 in 3:627–628
 RA 3:755–756
 SLE 3:756
 vasculitis 3:755
neutrophil-SGD *see* neutrophil-specific granule deficiency (neutrophil-SGD)
neutrophil-specific granule deficiency (neutrophil-SGD) 3:755
never in mitosis gene A (NIMA) 2:667–668
new gene origination 4:154, 4:156
nexin-DRC 2:663
next generation sequencing (NGS) 1:29–30, 1:31, 2:799, 4:141, 4:153, 4:161
Nezha *see* calmodulin-regulated spectrin-associated protein 3 (CAMSAP3)
NF1 *see* neurofibromatosis type1 (NF1)
NF2 *see* neurofibromin 2 (NF2)
NFAT *see* nuclear factor of activated T cells (NFAT)
NF-E2 *see* nuclear factor erythroid 2 (NF-E2)
NF-E2p45-related factor 2 (NRF2) 2:531
NF-KB oscillations 3:674–675
nflb1 KO mice 3:581
nflb2 KO mice 3:581–582
NFTs *see* neuro fibrillary tangles (NFTs)
NFκB *see* nuclear factor kappa B (NFκB)
NF-κB-inducing kinase (NIK) 3:581–582
NG2 glia 3:800
NGAL *see* neutrophil gelatinase-associated lipocalin (NGAL)
NGD *see* no-go decay (NGD)
NGF *see* nerve growth factor (NGF); neurotrophic growth factor (NGF)
N-glycans *see* N-linked glycans (N-glycans)
N-glycosylation 1:126, 1:555, 1:557–558, 1:563
N-glycosylation quality control pathway 1:599–601, 1:600F
N-glycosylation site 1:682–683
NGS *see* next generation sequencing (NGS)
NHEJ *see* nonhomologous end-joining (NHEJ)
NHEJ pathway *see* nonhomologous end-joining pathway (NHEJ pathway)
NHL *see* non-Hodgkin lymphoma (NHL)
NHR *see* neutralized homology repeat (NHR)
NIAP *see* neuronal inhibitor apoptosis protein (NIAP)
NICD *see* Notch intracellular domain (NICD)
niche 3:776
nicotinamide adenine dinucleotide (NADH) 2:110, 2:133, 3:341
 dehydrogenase 1:304–305, 1:305
nicotinamide adenine dinucleotide phosphate (NADPH) 2:790, 3:519–520
nicotinamide adenine dinucleotide phosphate-oxidase (NADPH oxidase) 2:767, 2:776, 3:702, 3:752
 active 3:702F
 assembly and activation 3:705–706
 regulation by phosphorylation 3:706
NIDDM *see* non-insulin-dependent diabetes mellitus (NIDDM)
nidogens 2:697, 4:104
Niemann-Pick C (NPC) 2:206
 NPC1 2:425
 NPC2 2:425
Niemann-Pick disease type C 1:779, 1:780, 1:783
 enzyme therapy 1:783
NIH *see* National Institutes of Health (NIH)
NIK *see* NF-κB-inducing kinase (NIK)
NIMA *see* never in mitosis gene A (NIMA)
NIMA-related kinases (NEKs) 2:667–668
'9 + 2' cilia 3:79
nineteen complex (NTC) 1:497, 1:497–498, 1:497T, 1:498–499, 1:498F, 1:499, 1:500F
NIR light *see* near-infrared (NIR light)
nitric oxide (NO) 1:88, 1:700, 1:717, 3:56
nitric oxide synthase (NOS) 3:57, 3:63, 3:617–618
nitrogen assimilation 2:329
nitrosylation 1:87–88
NK cells *see* natural killer (NK) cells
N-kinesins 2:609
NKT cells *see* natural killer T cells (NKT cells)
N-linked glycans (N-glycans) 2:3, 2:339–340, 2:339–346, 2:339F, 3:86F
 chaperones recognizing high mannose 2:343
 Dol-P-P-Oligosaccharide maturation 2:340
 En-Bloc transfer to protein 2:340–341, 2:341F
 functions at cell surface 2:344
 functions on secreted glycoproteins 2:344
 glycoproteins processing and folding 2:341, 2:342F
 initiation on Dol-P 2:340
 lysosomal hydrolases unique modification 2:341–342
 maturation during transit 2:342
 MPRs 2:343–344
 somatic cell mutants for glycosylation 2:344T
N-linked glycosylation 1:85–86
NLPs *see* neuropeptide-like proteins (NLPs)
NLR family CARD-containing 4 (NLRC4) 3:394
NLR proteins *see* NOD-like receptor (NLR) proteins
NLR *see* nucleotide-binding oligomerization domain receptor (NLR)
NLRs *see* nucleotide-binding oligomerization domain-like receptors (NLRs)
NLRC4 *see* NLR family CARD-containing 4 (NLRC4)
NLRP3 *see* NACHT, LRR, and PYD domains-containing protein 3 (NLRP3); NOD-like receptor family pyrin domain-containing 3 (NLRP3)
NLs *see* neutral lipids (NLs)
NLS *see* nuclear localization sequences (NLS); nuclear localization signal (NLS)
NLSIM *see* nonlinear structure illumination (NLSIM)
NMD *see* nonsense-mediated decay (NMD); nonsense-mediated mRNA decay (NMD)
NMDA *see* N-methyl-D-aspartate (NMDA)
NMDAR *see* N-methyl-D-aspartic acid receptor (NMDAR)
N-methyl-D-aspartate (NMDA) 1:249
N-methyl-D-aspartic acid receptor (NMDAR) 1:675, 2:294
N-Methyl-N-trimethylsilylfluoroacetamide (MSTFA) 4:203–204
NMII *see* non-muscle myosin II (NMII)
NMJ *see* neuro-muscular junction (NMJ)
NMR spectroscopy *see* nuclear magnetic resonance (NMR) spectroscopy
nNOS *see* neuronal NOS (nNOS)
NNRTI *see* non-nucleoside reverse transcriptase inhibitors (NNRTI)
NO *see* nitric oxide (NO)
NOD *see* nucleotide oligomerization domain (NOD)

- NOD/SCID (non-obese diabetic with severe combined immunodeficiency disease) 3:807
- NOD2 signaling 3:391
- nodal vesicular parcels (NVPs) 2:611
- nodaviruses, spherules producing by 2:819–820
- nodes 4:213
- NOD-like receptor (NLR) proteins 3:368
- NOD-like receptor family pyrin domain-containing 3 (NLRP3) 3:520
- NoDS *see* nucleolar detention signal (NoDS)
- NOE *see* nuclear overhauser enhancement (NOE)
- noggin (NOG) 3:44
- no-go decay (NGD) 3:317
- substrate choice 3:318
- substrate recognition 3:320–321
- nomenclature systematization of biological entities 4:247
- non steroidal anti-inflammatory drug (NSAID) 1:666, 3:64
- nonalcoholic fatty liver disease (NAFLD) 1:157, 2:534
- nonalcoholic steatohepatitis (NASH) 1:157
- non-amine-containing phospholipids synthesis
- in bacteria 1:162–164
- in eukaryotes 1:166–169
- non-apoptotic pathways 3:70
- noncancer telomere-related diseases 1:432–433
- noncanonical HH signaling 3:79–80
- non-clathrin endocytosis (NCE) 2:212
- noncoding RNAs (ncRNA) 1:364, 3:788, 4:160, 4:162
- ceRNAs 1:369
- eRNAs 1:369
- lncRNAs 1:369–370
- transposons and piRNAs 1:370
- noncoding transcription *see* germline transcription
- non-collagenous 1 domain (NC1 domain) 2:696–697
- nonconserved region (NCR) 1:469–471
- noncovalent interactions 1:10–11
- noncrossover 1:461, 1:464–465
- nongenetic cell-to-cell heterogeneity
- stochastic analysis
- cell signaling 4:218–219
- cell-to-cell heterogeneity 4:218
- connecting heterogeneity at multiple levels 4:223–224
- single-cell analysis 4:220–221
- stochastic data modeling and analysis 4:222–223
- variability in gene expression 4:220
- nongenetic heterogeneity 4:218
- non-healing ulcers 3:218
- nonhematologic cell secretory lysosomes 2:228
- non-Hodgkin lymphoma (NHL) 3:110
- nonhomologous DNA end joining
- see* nonhomologous end-joining pathway (NHEJ pathway)
- nonhomologous end-joining pathway (NHEJ pathway) 1:451, 1:461, 3:466, 3:550, 3:777
- A-NHEJ 1:452
- choice of DSB repair pathway 1:452–453
- and disease 1:453
- and inhibitors 1:453–454
- mechanism 1:452F
- DSBs recognition 1:451
- ligation 1:451–452
- nucleolytic processing by
- ARTEMIS–DNA-PKcs complex 1:451
- template-dependent and template-independent DNA synthesis 1:451
- pathway 3:558–560
- non-insulin-dependent diabetes mellitus (NIDDM) 1:677
- nonlinear structure illumination (NLSIM) 2:95
- non-lymphoid organs, DCs in 3:741
- non-lymphoid tissue (NLT) DC 3:595
- CD103⁺ CD11b⁺ DCs 3:595
- CD11b⁺ DCs 3:595
- Langerhans cells and migratory DCs 3:595–596
- and migratory DCs 3:597
- non-motile cilia 2:660–661
- axoneme structure 2:663
- diseases caused by defective 2:670–671
- non-muscle myosin II (NMII) 2:601, 2:602, 2:720
- non-nucleoside reverse transcriptase inhibitors (NNRTI) 2:837
- non-receptor tyrosine kinases 3:156
- non-ribosomal peptide synthetases (NRPS) 1:95
- nonribosomal polypeptide synthase 1:14
- nonsense-mediated decay (NMD) 3:317
- substrate choice 3:317–318
- substrate recognition 3:319–320
- nonsense-mediated mRNA decay (NMD) 1:351–352
- nonstop-decay (NSD) 3:317
- substrate choice 3:318–319
- substrate recognition 3:321
- nonstructural proteins 2:810
- nontemplate strand 1:471–472, 1:474
- Noonan syndrome (NS) 3:417–418
- NoRC *see* nucleolar remodeling complex (NoRC)
- Norrin 3:89, 3:89F
- NORs *see* nucleolar organizing regions (NORs)
- North American Indian childhood cirrhosis (NAIC) 3:295
- NOS *see* nitric oxide synthase (NOS)
- NOS2 *see* iNOS
- Notch 4:102, 4:112
- Notch-sparing inhibitors 1:665
- receptor 1:662
- Notch intracellular domain (NICD) 1:643–644, 4:112
- Notch pathway 3:2
- Notch signaling pathway 3:107, 3:108F
- see also* Hippo pathway
- developmental biology
- cell fate determination mechanisms 3:109F
- examples in development 3:109
- mechanisms 3:107–109
- mutations in syndromes and cancer 3:110F
- Notch signaling 1:664–665, 3:107
- pathology 3:109–111
- NP *see* nuclear protein (NP)
- NPC *see* neural precursor cells (NPCs); Niemann-Pick C (NPC); nuclear pore complex (NPC)
- NPFs *see* nucleation promoting factors (NPFs)
- NPHP *see* nephronophthisis (NPHP)
- NpHR *see* natronomonas pharaonis (NpHR)
- NPM *see* nucleophosmin (NPM)
- NPM1 *see* nucleophosmin 1 (NPM1)
- NPxY motifs *see* Asn-Pro-x-Tyr motifs (NPxY motifs)
- Nramp2 *see* SLC11a2 iron transporter
- Nrf1 1:591
- NRF2 *see* NF-E2p45-related factor 2 (NRF2)
- NRG1 *see* neuregulin-1 (NRG1)
- NRP *see* neuropilin (NRP)
- NRPS *see* non-ribosomal peptide synthetases (NRPS)
- NRR *see* negative regulatory region (NRR); negative regulatory repeats (NRR)
- NRTI *see* nucleoside reverse transcriptase inhibitors (NRTI)
- NS *see* Noonan syndrome (NS)
- NSAID *see* non steroidal anti-inflammatory drug (NSAID)
- NSCs *see* neural stem cells (NSCs)
- NSD *see* nonstop-decay (NSD)
- NSF *see* N-ethylmaleimide-sensitive factor (NSF)
- N-site 4:36, 4:37–38, 4:39F
- NSUN2 1:334, 1:334–335
- NT *see* neurotrophins (NT)
- NT domain *see* nucleotidyl transferase domain (NT domain)
- NTC *see* nineteen complex (NTC)
- NTD *see* N-terminal domain (NTD)
- N-terminal 'M13' peptide 2:53–54
- N-terminal domain (NTD) 1:468–469
- N-terminal fragments (NTFs) 1:664
- N-terminal methionine 1:529
- N-terminal method 1:75–76
- N-termini, negative selection of 1:568, 1:574–575
- N-termini, positive selection of 1:568, 1:576
- NTF2 *see* nuclear transport factor 2 (NTF2)
- NTFs *see* N-terminal fragments (NTFs)
- NTR *see* nuclear transport receptor (NTR)
- NTRK3 *see* neurotrophic tyrosine kinase, receptor, type 3 (NTRK3)
- nuclear, chromatin topography 2:91
- nuclear aggresomes 2:826
- nuclear body 3:261
- nuclear Dbf-2-related (NDR) 3:99
- nuclear deformability 2:573
- nuclear distribution E 2:624
- nuclear entry and integration 2:835

- nuclear envelope (NE) 1:205, 1:206, 2:20–21, 2:28F, 2:311, 2:467F, 2:471–472, 3:170–172, 3:176–177
 disassembly and reassembly in mitosis
 closed mitosis 2:315
 disassembly and reassembly 2:317
 importin α/β and Ran system 2:316F
 open mitosis 2:315
 RanGTP gradient 2:315–317
 INM proteins 2:314–315
 NPC 2:313–314, 2:314F
 nuclear lamin 2:312–313
 nuclear membrane 2:312
 structure of metazoan 2:312F
 nuclear envelope breakdown (NEB) 3:479–480
 nuclear envelope transmembrane proteins (NET proteins) *see* inner nuclear membrane proteins (INM proteins)
 nuclear excision repair pathway (NER pathway) 3:462–463, 3:464F
 nuclear export 3:300, 3:302
 pre-ribosomal particles 3:299–301
 nuclear export sequences (NES) 3:90F, 3:93
 nuclear export signal (NES) 2:313, 2:320, 3:172, 3:300, 4:218, 4:222
 nuclear factor erythroid 2 (NF-E2) 2:531
 nuclear factor kappa B (NF κ B) 1:757, 2:530, 3:6, 3:70, 3:130, 3:224, 3:355–356, 3:520, 3:580, 3:710, 3:770–771, 4:99
 B- and T-cell receptor pathways 3:583
 BAFF receptor pathways 3:583
 CD28 receptor pathway 3:583
 CD40 3:583
 core NF- κ B signaling gene and protein names 3:581T
 dimers 3:580–581
 expression and knockout mice 3:581
c-rel KO mice 3:582
 immune defects 3:583T
nflkb1 KO mice 3:581
nflkb2 KO mice 3:581–582
rela KO mice 3:582
relb KO mice 3:582
 human immune cell viruses 3:585
 in innate immunity 3:583–584, 3:584T
 misregulation in human immune diseases 3:584–585
 pathways 3:392–393
 canonical and noncanonical 3:393F
 signaling 1:428, 1:708–709
 superfamily proteins 3:582F
 for therapy in immune diseases 3:585–586
 TNFR-induced NF- κ B activation 3:357–358
 nuclear factor kappa-light-chain-enhancer of activated B cells (NF-KB) activation dynamics 3:667
 nuclear factor of activated T cells (NFAT) 3:6, 3:167, 3:355–356, 3:573, 4:105–106
 domain structure 3:573
 functions in T cells 3:575
 NFAT-regulated transcription 3:574–575
 regulation of activation 3:573–574, 3:574F
 and T cell tolerance 3:575–576, 3:576F
 nuclear genomes 1:312–313, 1:380
 nuclear import machinery 2:814–815, 3:439–440
 nuclear lamina 2:569, 2:573, 2:576
 nuclear localization sequences (NLS) 3:90F, 3:92–93, 3:94F
 nuclear localization signal (NLS) 2:312–313, 2:320, 3:172, 3:236, 4:218, 4:222
 nuclear magnetic resonance (NMR) spectroscopy 1:93, 1:98
 bacterial phosphotransferase system, structural proteomics of 1:101–103
 experimental and computational considerations 1:99
 computational methods 1:100–101
 intermolecular distance restraints 1:99
 orientational restraints 1:100
 paramagnetic-based distance restraints 1:99–100
 sources of distance information 1:100
 exploring sparsely populated states of proteins and complexes 1:103
 lifetime line broadening and DEST 1:104–105
 PRE 1:103–104
 principal NMR methods 1:103
 relaxation dispersion spectroscopy 1:104
 relevance to cell biology 1:105
 historical perspective 1:98
 NMR spectroscopy in structural and cell biology 1:99
 solution structure determination of macromolecules 1:98
 structure determination 1:98–99
 nuclear migration 2:625
 nuclear migration one (Num1) 1:240
 nuclear mitotic apparatus (NuMA) 2:625, 3:480–481
 nuclear organization 1:311–318
see also nuclear envelope (NE)
 nuclear overhauser enhancement (NOE) 1:98, 1:98–99
 nuclear pore complex (NPC) 1:205, 2:20–21, 2:28F, 2:28, 2:91, 2:311, 2:319, 2:319F, 3:170, 3:299–300
 nuclear pores 2:319–323
 NPC 2:319, 2:319F
 nucleocytoplasmic transport 2:320–321, 2:321F
 Nups 2:319–320
 and disease 2:321
 dynamic behavior 2:320
 nuclear protein (NP) 2:824
 nuclear receptors 1:176–177
 nuclear translocation 3:123–124
 nuclear transport factor 2 (NTF2) 2:315–316, 3:300
 nuclear transport receptor (NTR) 2:313
 3' nuclease 1:439
 nucleation 2:539, 2:542
 acentrosomal 2:544
 nucleation-growth 2:552–553
 nucleation promoting factors (NPFs) 2:491, 2:491–492, 2:753, 3:187, 4:14–15
 nucleic acid 1:25
 chemical modification 1:27–28
 function 1:26
 physical structure 1:26–27
 sequencing 1:29–30
 structure of bases in 1:25F
 synthesis/breakdown 1:2–3
 nucleic acid programmable protein array (NAPPA) 4:190
 nucleic acids, viral *see* viral nucleic acids
 nucleobase 1:26
 chemical modification 1:29F
 diethylpyrocarbonate and kethoxal modification 1:30F
 tautomerization 1:26F
 nucleocapsid protein (NC protein) 2:828–829
 nucleocytoplasmic large DNA viruses (NCLDV) 2:824–825
 nucleocytoplasmic transport 2:320–321, 2:321F
 nucleolar architecture 3:254
 disassembly and reassembly 3:265F
 dominance 3:258
 dynamics during cell division cycle 3:263
 nucleolar-associated body 3:259
 proteins 3:260F
 proteome dynamics 3:273–274
 response to stress 3:262F
 RNAs 3:272
 sirtuins 3:275
 in transcriptional inhibition and DNA damage 3:260–263
 nucleolar detention signal (NoDS) 3:272
 nucleolar organizing regions (NORs) 3:254, 3:257F, 3:281
 nucleolar remodeling complex (NoRC) 3:258
 nucleolar signaling networks 3:263
 during stress 3:273
 kinase signaling 3:274–275
 nucleolar proteome dynamics 3:273–274
 nucleolar sirtuins 3:275
 nucleoli/nucleolus 3:254–269
 dynamic nucleolus 3:259–263
 immunofluorescence microscopy of tripartite nucleolar architecture 3:256F
 interchange 3:259
 kinase signaling 3:274–275
 proteomic studies 3:264–265
 rDNA 3:257–258
 regulation of rDNA expression 3:258–259
 ribosome birth in 3:270
 P53-independent ribosome stress 3:272–273
 ribosome subunit biogenesis 3:254–257
 in mammalian nucleoli 3:256F
 rRNA transcription 3:258
 stress detection by 3:270
 nucleolar RNAs 3:272

- nucleoli/nucleolus (*continued*)
 RiBi factors 3:272
 RPs 3:270–272, 3:271T
 tripartite structure of human nucleoli 3:255F
 and viruses 3:263–264
nucleophile 1:738
nucleophosmin (NPM) 3:272, 3:283
nucleophosmin 1 (NPM1) 3:452–453
nucleoporins (Nups) 2:313, 2:319, 3:175, 3:177
 and disease 2:321
 dynamic behavior 2:320
 structural Nups 3:175
nucleoprotein 1:418, 3:512
nucleoside
 analogs 1:27F
 chemical modification 1:29F
nucleoside diphosphate kinase (NucDipK) 1:17, 1:18
nucleoside reverse transcriptase inhibitors (NRTI) 2:837
nucleosomes 1:436F, 2:311, 3:786–787
nucleosome remodeling and histone deacetylase complex (NuRD/MeCP1) 2:499
nucleotide-binding oligomerization domain-like receptors (NLRs) 3:520
nucleotide exchange factors (NEFs) 1:599
nucleotide excision repair (NER) 1:435, 1:463–464
 eukaryotic *see* eukaryotic nucleotide excision repair
nucleotide nucleotides (N nucleotides) 3:559–560
nucleotide oligomerization domain (NOD) 3:714
nucleotide oligomerization domain leucine-rich repeat receptor (NLR) 3:710
nucleotide pyrophosphatase/phosphodiesterase 1:124
nucleotide-binding domain (NBD) 3:394, 3:714
nucleotide-binding oligomerization domain receptor (NLR) 1:708–709, 2:303–304, 2:772–773, 3:584, 3:714–718, 3:715–716T, 3:733, 3:746–747
nucleotide-dependent polymerization 2:593
nucleotidyl transferase domain (NT domain) 2:285
nucleus, related issues/unanticipated roles of 3:176
nucleus, Wnt signaling in 3:92–93
 co-activators 3:94
 co-activators and corepressors 3:95
 HATs 3:94–95
 histone methyltransferases (HMT) 3:95
 nuclear export 3:93
 nucleus import 3:92–93
 β -catenin/TCF complex 3:93–94
Num1 *see* nuclear migration one (Num1)
NuMA *see* nuclear mitotic apparatus (NuMA)
numerical aperture (NA) 2:59, 2:62
numerical time-integration schemes 4:118–119
Nups *see* nucleoporins (Nups)
NuRD/MeCP1 *see* nucleosome remodeling and histone deacetylase complex (NuRD/MeCP1)
NusA protein 1:474
NusG protein 1:474
Nutlins 1:591
nutrient-sensing 1:771, 1:775, 2:206
 autophagy and the lysosome in energy economy and 1:765–766
nutrient-sensing pathways 3:815
NVPs *see* nodal vesicular parcels (NVPs)
N-WASP *see* Neuronal Wiskott–Aldrich Syndrome protein (N-WASP)
- O**
OA *see* osteoarthritis (OA)
OAD *see* outer-arm dyneins (OAD)
OB *see* olfactory bulb (OB); oligosaccharide binding (OB)
OB fold 1:396, 1:407
obesity 2:804–805
objective lens 2:14–15
OBO *see* open biomedical ontologies (OBO)
observation volume 2:101
OCA1 *see* oculocutaneous albinism type 1 (OCA1)
OCA2 *see* oculocutaneous albinism type 2 (OCA2)
Oct4 *see* octamer-binding protein 4 (Oct4)
octamer-binding protein 4 (Oct4) 3:104, 3:786, 3:788–789, 3:809
oculocutaneous albinism type 1 (OCA1) 2:237–239
oculocutaneous albinism type 2 (OCA2) 2:237
ODAs *see* outer dynein arms (ODAs)
ODE *see* ordinary differential equation (ODE)
ODN *see* oligodeoxynucleotide (ODN)
oep *see* one-eyed pinhead (oep)
OFDI *see* optical frequency domain imaging (OFDI)
O-Fuc glycans *see* O-linked fucose glycans (O-Fuc glycans)
O-fucosylation 1:633
O-Gal glycans *see* O-galactose glycans (O-Gal glycans)
O-galactose glycans (O-Gal glycans) 2:351
O-GalNAc glycans 2:347–348, 2:348F
OGG1 *see* 8-oxoguanine DNA glycosylase (OGG1)
O-Glc glycans *see* O-glucose glycans (O-Glc glycans)
O-GlcNAc glycans 2:351
O-GlcNAc transferase (OGT) 1:281
O-GlcNacylation 2:287
O-glucose glycans (O-Glc glycans) 2:350–351
O-glycans 2:3
OGT *see* O-GlcNAc transferase (OGT)
Ohm's law 1:220–221
Okazaki fragments 1:396, 1:405–406, 1:409, 1:410
olfactomedin4 (OLFM4) 3:753
olfactory bulb (OB) 2:728, 4:136–137
 interneurons 4:137–138
 neuronal differentiation 4:137–138
olfactory stimulation 4:136–137
OLFM4 *see* olfactomedin4 (OLFM4)
oligodendrocytes 1:247–248
oligodeoxynucleotide (ODN) 3:442
oligodeoxyribonucleotides, chemical synthesis of 1:36–40
 H-phosphonate approach 1:40–41
 phosphoramidite approach 1:36–40, 1:41–43
 acetal protecting groups 1:43–44
 ether protecting groups 1:42–43
 orthoester and ester protecting groups 1:45–46
 solid-phase synthesis protocol 1:37–39, 1:38F
oligosaccharide binding (OB) 3:459
oligosaccharyltransferase (OST) 1:535, 1:539–540T, 1:540, 1:541F, 1:598–599, 2:340–341
O-linked fucose glycans (O-Fuc glycans) 2:350, 2:350F
O-linked glycosylation 2:574
O-linked mannose glycans (O-Man glycans) 2:348–350, 2:349F
O-linked oligosaccharides 1:85–86
OMA1-KO cells 2:289
O-Man glycans *see* O-linked mannose glycans (O-Man glycans)
 ω -alkyn fatty acid derivatives 1:132
omegasome 2:246
omics technologies 4:163–164, 4:166, 4:266
OMM *see* outer mitochondrial membrane (OMM)
OMVs *see* outer membrane vesicles (OMVs)
oncogene(s) 3:411, 3:428–429, 3:585
 in human developmental diseases 3:417
 in diabetes 3:419–420
 in neurodegenerative diseases 3:419
 RAS-ERK and PI3K-AKT dysregulation developmental syndromes 3:417–418
 in metabolism 3:416F
 oncogenic transformation
 Rho GTPases in 2:735–737
 ribosome biogenesis regulating by 3:281–282
 tumor suppressor genes and human cancer 3:411–412
 extracellular cues transmission 3:413
 functional roles 3:413
 identification 3:412
 interplay between 3:413–415
 intracellular signaling pathways 3:414F
 PI3K pathway 3:413
 RAS pathway 3:413
 transcriptional regulation 3:413
oncomiRs 1:367
oncosomes 2:304

- one gene–one enzyme hypothesis 1:84
 one-eyed pinhead (*oep*) 3:42
 ONM *see* outer nuclear membrane (ONM)
 ONOO[−] *see* peroxynitrite (ONOO[−])
 on-section labeling 2:36–37, 2:38F
 on-target effect 3:385
 ontologies 4:247
 Opa1 2:282–284
 Opa1-dependent mitochondrial fusion
 2:288–289
 O-palmitoleoylation 3:86, 3:86–87, 3:87,
 3:87–88
 open biomedical ontologies (OBO) 4:247
 open licenses 4:248–249
 open mitosis 2:315
 Open Model 4:17, 4:18F, 4:19
 examples of models on 4:21F
 open reading frame (ORF) 4:193
 OPG *see* osteoprotegerin (OPG)
 opsin(s) 2:456, 2:668
 opsonization 2:751, 2:752, 2:766–767,
 3:700–701
 optical frequency domain imaging (OFDI)
 2:129–131
 optical sectioning
 microscopies 2:68–69
 in SIM 2:89
 optical transfer function (OTF) 2:86–87
 optogenetics 2:147–155, 2:147
 experimentation 2:148F
 with light 2:147
 delivery of light 2:149–150
 genetic targeting 2:149
 light-sensitive (bio)molecules
 2:147–149
 light-sensitive proteins 2:148T
 light-sensitive proteins 2:150–151
 optical control of neuronal excitability
 2:150F
 optochemical approaches 2:151
 light-gated glutamate receptors
 2:151–153, 2:153
 Orci's group 2:358
 ordinary differential equation (ODE) 4:19,
 4:74, 4:222
 ORF *see* open reading frame (ORF)
 organ homeostasis 3:4, 3:4–5
 organelle(s) 1:66, 2:2–3, 2:156, 2:542–543,
 4:236–237, 4:237, 4:237–238
 bioenergetics
 cytochrome *bc*₁ Complex 1:305–306
 cytochrome oxidase 1:306
 energetics of membrane-based
 adenosine triphosphate synthesis
 1:298
 membrane protein complexes, atomic
 structures of 1:298–299
 mitochondrial respiratory chain
 1:304–305
 biogenesis 2:390
 GLUT4 storage compartment formation
 2:390
 lysosome formation, secretory granules,
 and dense core vesicles 2:390
 communication between
 see interorganellar communication
 function 2:506
 identity 2:505–507
 structural analysis 2:92
 organizational cell biology 2:1–4
 cytoskeleton and motors 2:3–4
 imaging 2:1–2
 interorganellar communication 2:3
 compartmental organization 2:3
 intracellular infectiology 2:4
 organelles 2:2–3
 organoid 3:808
 organula 2:156
 orientational restraints 1:100
 origin replication complex (ORC)
 3:458–459
 ORP1L *see* oxysterol-binding protein-related
 protein 1 L (ORP1L)
 Orphan Drug Act 1983, in United States
 1:768, 1:781
 orthologous sequences 1:81
 OSBP1 *see* oxysterol-binding protein 1
 (OSBP1)
 oscillations 4:44
 OST *see* oligosaccharyltransferase (OST)
 osteoarthritis (OA) 1:635
 osteoclast 1:776–777
 osteoporosis 1:772
 osteoprotegerin (OPG) 3:63
 osteopycnodysostosis 1:776
 OTF *see* optical transfer function (OTF)
 OTUs *see* ubiquitinC-terminal proteases
 (OTUs)
 OTU domain-containing ubiquitin
 aldehyde-binding protein 1 (Otub1)
 3:506
 OTU proteases *see* ovarian tumor domain
 (OTU) proteases
 OUT *see* ovarian tumor (OUT)
 outer dynein arms (ODAs) 2:627, 2:660
 HCs 2:630
 outer membrane vesicles (OMVs)
 2:305–306
 outer mitochondrial membrane (OMM)
 1:237, 3:345
 in diseases and potential drug target
 AD 1:241
 cancer 1:240–241
 heart disease 1:241
 enzymes localized in 1:239T
 mitochondrial compartments separation
 and isolation 1:237
 moonlighting 1:238–240
 mitochondrial endoplasmic reticulum
 interactions 1:238–240
 mitochondrial interactions with Golgi,
 peroxisomes, and plasma membrane
 1:240
 mitochondrial–nuclear interactions
 1:240
 protein/enzyme assignments 1:237–238
 outer nuclear membrane (ONM) 2:312,
 3:480
 outer-arm dyneins (OAD) 2:627
 outside-in
 response process 2:718
 transmission 3:650–651
 OVA *see* ovalbumin (OVA)
 ovalbumin (OVA) 3:762
 ovarian cancer stem cells 3:810
 ovarian tumor (OUT) 3:358–359
 ovarian tumor domain (OTU) proteases
 1:586
 over-dominance 1:366
 OX40 3:772–773
 Oxa1p 1:550, 1:552–553
 oxidative phosphorylation 1:288,
 1:293–294
 oxidative protein folding 2:161
 oxidative stress-related PTMs 1:87
 8-oxoguanine DNA glycosylase (OGG1)
 1:444–445
 oxygen tension 4:75
 oxygen-dependent killing 2:767
 CGD 2:768
 iNOS 2:767–768
 NADPH oxidase and MPO 2:767
 oxygen-independent killing 2:771
 O-xylose glycans (O-Xyl glycans) 2:351–352
 oxysterol-binding protein 1 (OSBP1) 2:504
 oxysterol-binding protein-related protein 1
 L (ORP1L) 2:414
- ## P
- P loop superfamily 2:593
 P nucleotides *see* palindromic nucleotides
 (P nucleotides)
 P protein *see* oculocutaneous albinism type
 2 (OCA2)
 P residues 1:568, 1:569–570
 P site 3:310
 P' residues 1:568, 1:569–570
 P1 fraction 2:293
 P2 fraction 2:293
 p15^{INK4b} levels regulation 3:437–438
 p16 3:513, 3:515
 p16^{INK4a} levels regulation 3:437–438
 p16^{INK4a} tumor suppressor protein 3:778
 p16INK4a-mediated senescence 3:448
 p21-activated kinase (PAK) 2:734–735
 PAK1 2:394
 p21^{CIP1} protein 3:437, 3:438
 p22phox 3:703
 p22phox subunit 3:707
 phosphorylation 3:707
 p24 protein family 2:169–170
 p27^{KIP1} protein 3:437, 3:438
 p38 cascade 3:123–124, 3:124
 p38 MAPK pathway and related pathways
 2:848–849
 p40phox 3:704
 phosphorylation 3:706–707
 p47phox 3:704
 phosphorylation 3:706
 p47phox–p67phox–p40phox complex
 in resting cells 3:704–705
 p53 1:591, 3:5, 3:512, 3:513, 3:514–515,
 3:515
 activation 3:263, 3:270–272, 3:284–285
 p53 effector proteins 3:462

- p53 pathway 3:504
 control through ubiquitin-proteasome pathway 3:504–505
 DUBs in 3:506
 E3 ligases 3:505
 in metabolism 3:507–509
 p53-Mdm2-HAUSP connection 3:506
 regulation and activation 3:504, 3:508F
 transactivation through posttranslational modifications 3:506–507
 ubiquitin E3 ligases 3:505F
- p53-independent ribosome stress 3:272–273
- p53-induced death domain protein (PIDD) 3:370
- p53-Mdm2 negative feedback loop 4:187
- p55/60 protein 3:392
- p57^{KIP2} protein 3:437
- p58/ERGIC-53/LMAN1 2:169
- p67phox 3:704
 phosphorylation 3:706–707
- p75/80 3:354
- p75neurotrophinreceptor (p75^{NTR}) 3:67, 3:72–73
 DR6 3:73
 EDAR 3:72
 p75^{NTR} 3:72–73
- p150glued protein 2:476
- p300/CBP-associated factor (PCAF) 3:507
- PA edematoxin (PA + EF) 2:795
- PA *see* phosphatidic acid (PA); protective antigen (PA)
- PA28 1:590
- PA200 1:590, 1:591
- PAB *see* poly(A)-binding protein (PAB)
- Pab1 1:348
- PABA-peptide *see* N-benzoyl-L-tyrosyl-p-aminobenzoic acid (PABA-peptide)
- PABP *see* poly(A) binding protein (PABP)
- PACs *see* photo-activated adenyl cyclases (PACs)
- PACS1 *see* phosphofurin acidic cluster sorting protein 1 (PACS1)
- PACSIN *see* protein kinase C and casein kinase 2 substrate in neurons (PACSIN)
- PACT *see* protein kinase R-activating protein (PACT)
- PAD4 *see* peptidylarginine deaminase 4 (PAD4)
- PA-FP *see* photoactivatable fluorescent proteins (PA-FP)
- PAI-1 *see* plasminogen activator inhibitor-1 (PAI-1)
- paired basic amino acid cleaving enzyme 4 (PACE4; PCSK6) 1:562, 1:564, 1:564T
- paired box 3 (Pax3) 3:102
- pairwise C-alpha deviation 2:8F, 2:9
- PAK *see* p21-activated kinase (PAK)
- PAK1 phosphorylates CtBP1/BARS 2:394
- Palade lab 2:156
- palindromic nucleotides (P nucleotides) 3:559–560
- palivizumab 1:145
- PALM *see* photo-activation localization microscopy (PALM)
- 1-palmitoyl 2-oleoyl phosphatidylcholine (POPC) 1:212
- palmitoylation 1:87
- Pals1 *see* membrane-associated palmitoylated protein 5 (MMP5)
- PALS1 *see* protein associated with Lin Seven 1 (PALS1)
- PALS1-associated tight junction protein (PATJ) 2:742–743, 3:102
- PAM *see* peptidylglycine α -amidating monooxygenase (PAM)
- PAMPs *see* pathogen associated molecular patterns (PAMPs)
- PAN2 *see* poly(A)-nuclease 2 (PAN2)
- pancreatic enzymes 1:257
- pancreatic secretions, in CF patients 1:257, 1:257–259
- pancreatic β -cell function 2:533
 UPR signaling contributing degeneration 2:533
- Pannin production (Pnr production) 4:67
- Papillon-Lefèvre syndrome (PLS) 3:754–755
- Par complex 2:742–743
- par genes *see* partitioning defective genes (par genes)
- PAR1 *see* protease-activated receptor1 (PAR1)
- paracellin-1 *see* Claudin-16
- paracrine activity 4:102, 4:104
- paracrine signaling 3:672, 3:802–803, 3:809
- PARADIGM algorithm 4:216
- parallel bundles
 of actin filaments 2:686
 in filopodia 2:683
- parallel reaction monitoring (PRM) 4:168
- paralogous
 gene duplications 1:81–82
 sequences 1:81
- paramagnetic effects 1:100
- paramagnetic pseudocontact shift (PCS) 1:98
- paramagnetic relaxation enhancement (PRE) 1:98
- paramagnetic-based distance restraints 1:99–100
- parasitic roundworm (*Ascaris megalocephala*) 2:649
- parathyroid hormone receptor (PTHr) 2:482
- PAR-CLIP *see* photoactivatable ribonucleoside CLIP (PAR-CLIP)
- parietal cells 1:733
- Parkinson's disease (PD) 1:768, 1:773, 2:479–480, 2:482, 3:133, 3:419
- Parkinsonism 1:591, 1:773, 1:785
- ParM/ParR/parC (ParMRC) 2:592
- PARN *see* poly(A)-specific ribonuclease (PARN)
- PARP *see* poly-ADP ribose polymerase (PARP)
- PARs *see* protease activated receptors (PARs)
- parthanatos 3:342, 3:345
- partial differential equations (PDEs) 4:73F, 4:74, 4:123
- particle methods 4:119–120, 4:120
- particle-based models 4:123–124
- particulate GCs (pGCs) 3:56, 3:57–58
 ANPCR 3:59
 GCA 3:58
 GCB 3:58
 GCC 3:58–59
 retinal GCs 3:59
 structure 3:58
- partitioning defective genes (par genes) 2:741–742
- PAS *see* phagophore assembly site (PAS)
- Pasha 1:354, 1:355, 1:355–358, 1:356–357F
- 7-pass transmembrane (7TM) 2:668
- passenger mutations 3:412–413
- passive ion transport via selective, facilitated diffusion ion channels 1:226–229
- passive membrane potential changes 1:229–230
- paste mechanism 3:550
- PAT1 1:351
- Patched (PTCH) 3:76–77
- pathogen associated molecular patterns (PAMPs) 1:708–709, 2:523–524, 2:790, 2:846, 3:368, 3:520, 3:619, 3:702, 3:710, 3:733, 3:752, 4:98
- pathogen-containing vacuoles (PCV) 2:405
- pathogenic/pathogen(s) 2:766
 bacteria 2:794
 defense 2:255
 interaction with lysosomal components 2:777–778
 neutralization 1:145
 offense 2:255
 subversion of host defense 2:773–774
- PATJ *see* PALS1-associated tight junction protein (PATJ)
- pattern recognition receptors (PRR) 1:708–709, 2:523–524, 3:368, 3:588, 3:619, 3:710, 3:733
- pause, in microtubule 2:539, 2:544
- Pax3 *see* paired box 3 (Pax3)
- Pax7 3:95, 3:790
- paxillin 2:724
- payload 1:135
- PB1 *see* Phox Bem1p (PB1)
- PBA *see* 4-phenylbutyric acid (PBA)
- PBM *see* protein-binding microarrays (PBM)
- P-bodies 1:349
- PBS *see* primer binding site (PBS)
- PC *see* phosphatidylcholine (PC); plastocyanin (PC)
- PC12 cells 4:47
 response to growth factor stimulation 4:46
- PCA *see* principle component analysis (PCA); protein complementation assay (PCA)
- PCAF *see* p300/CBP-associated factor (PCAF)
- PCAM-1 *see* platelet-endothelial cell adhesion molecule-1 (PCAM-1)
- p-carboxymethylphenylalanine (pCMF) 2:523
- PCB *see* polychlorinated biphenyl (PCB)

- PCBP2 *see* poly C binding protein 2 (PCBP2)
- PCC *see* post-cleavage complex (PCC)
- PCD *see* primary ciliary dyskinesia (PCD); programmed cell death (PCD)
- PC-FPs *see* photoconvertible fluorescent proteins (PC-FPs)
- PcG *see* polycomb group genes (PcG)
- PCH *see* photon count histogram (PCH)
- pcIC *see* pericentrosomal IC (pcIC)
- PCM *see* pericentriolar material (PCM)
- pCMF *see* p-carboxymethylphenylalanine (pCMF)
- PCNA *see* proliferating cell nuclear antigen (PCNA)
- PCP *see* planar cell polarity (PCP)
- PC-PLC *see* phosphatidylcholine phospholipase C (PC-PLC)
- PCR *see* polymerase chain reaction (PCR)
- PCS *see* paramagnetic pseudocontact shift (PCS); pseudocontact shifts (PCSS)
- PCSK9 *see* proprotein convertase subtilisin/kexin type 9 (PCSK9)
- PCST methods *see* Prize-collecting Steiner Tree methods (PCST methods)
- PCST *see* Prize-collecting Steiner Tree (PCST)
- PCV *see* pathogen-containing vacuoles (PCV); polycythemia vera (PCV)
- PD *see* Parkinson's disease (PD)
- PDB *see* Protein Data Bank (PDB)
- PDC *see* pyruvate dehydrogenase complex (PDC)
- pDCs *see* plasmacytoid dendritic cells (pDCs)
- PDEs *see* partial differential equations (PDEs); phosphodiesterases (PDEs)
- PDGF receptor beta (PDGFRB) 3:434
- PDGF *see* platelet-derived growth factor (PDGF)
- PDGFA *see* platelet-derived growth factor A (PDGFA)
- PDGFRB *see* PDGF receptor beta (PDGFRB)
- PDGFRs *see* platelet-derived growth factor receptors (PDGFRs)
- PDGFR β *see* platelet-derived growth factor receptor β (PDGFR β)
- PDH *see* pyruvate dehydrogenase (PDH)
- PDI *see* protein disulphide isomerase (PDI); protein-DNA interaction (PDI)
- PDK1 *see* phosphoinositide-dependent kinase 1 (PDK1)
- PDKs *see* pyruvate dehydrogenase kinases (PDKs)
- pdMS³ *see* phosphorylation directed MS³ (pdMS³)
- 3-PDPK1 *see* 3-phosphoinositide dependent protein kinase 1 (3-PDPK1)
- PDZ domain *see* PSD95, Dlg1 and zo-1 domain (PDZ domain)
- PE *see* phosphatidylethanolamine (PE); pulmonary embolism (PE)
- 1PE *see* single photon excitation (1PE)
- 2PE *see* two-photon excitation (2PE)
- PEF(L) domain *see* penta-EF-hand large domain (PEF(L) domain)
- PEFs *see* polar ejection forces (PEFs)
- PEG *see* polyethylene glycol (PEG)
- Pelle and IL-1R-associated kinases (PIK-1) 2:846
- penta-EF-hand large domain (PEF(L) domain) 1:670
- pentameric ligand-gated ion channels (pLGIC) 1:248–249, 1:249, 1:249–250, 1:250
- pentose phosphate pathway (PPP) 3:336
- Pentraxin 3 1:282
- PEP *see* PEST domain-enriched tyrosine phosphatase (PEP)
- PEPCK *see* phosphoenolpyruvate carboxykinase (PEPCK)
- pepsin 1:733
- pepstatin 1:684, 1:738
- peptide bond 1:525, 1:529, 1:530–531, 1:530F, 1:531
- peptide bond formation 1:19, 1:323, 1:377–380, 1:388–390
- peptide bound to major histocompatibility complex molecule (pMHC) 3:650
- peptide hormones 1:535–536, 1:556, 1:557
- peptide mass fingerprinting (PMF) 1:571, 1:571–572
- peptide MHC class I complexes (pMHC-I) 3:589
- peptide-loading complex (PLC) 3:589
- peptides (p) 3:528
- abundance quantification 4:169
- mimetic 1:741, 1:791–792
- peptides-major histocompatibility molecules (p-MHC) 3:528
- peptidoglycan (PGN) 2:771, 2:791
- peptidylarginine deaminase 4 (PAD4) 3:752–753
- peptidylglycine α -amidating monooxygenase (PAM) 1:86–87
- peptidylprolyl isomerase F (PIIF) 3:342
- peptidylprolyl isomerases (PPIases) 2:162
- peptidyl-tRNA 1:310, 1:322F, 1:388–390, 1:531, 1:532
- peptographs 1:571–572
- Per/Arnt/Sim domain (PAS domain) *see* phagophore assembly site (PAS)
- perfringolysin O (PFO) 1:210
- pericellular matrix 3:193
- pericentrin-like protein (PLP) 2:91–92
- pericentriolar factories 2:824–825
- pericentriolar material (PCM) 2:91, 2:649–651, 3:480
- pericentrosomal IC (pcIC) 2:175
- pericyte apoptosis 4:77–78
- pericyte(s) 4:102
- dissociation 4:111
- periglomerular cells (PGs) 4:134, 4:135F, 4:137–138
- perilipin 2:274
- peripheral lymph node addressin (PNAd) 3:526
- peripheral membrane proteins 1:209–210
- peripheral nervous system (PNS) 3:235
- peripheral neuropathy 2:611, 2:614
- peripheral supra-molecular activation cluster (pSMAC) 3:522–524, 3:660
- peripheral tissue antigens (PTAs) 3:618–619
- peripheral tolerance 3:588, 3:597, 3:605–606
- regulatory T cells (Tregs) 3:606–607
- role of CD28 3:606
- periplasm 2:157
- peristalsis 1:732
- PERK *see* PKR-related ER kinase (PERK)
- permeability barrier 2:313
- permeability transition pore complex (PTPC) 3:344–345
- peroxidase-positive granules *see* primary granules
- peroxisome proliferator-activated receptors (PPARs) 3:731
- PPAR γ 1:176–177, 3:102, 3:131–132
- peroxisomes 1:763, 1:764, 2:2–3, 2:248, 2:249F
- abundance, adaptability, and regulation 2:256–257
- biogenesis 2:263F
- De novo membrane formation 2:262
- evolution 2:269, 2:270F
- formation from ER 2:261–262
- growth and division 2:262
- remodeling, random turnover, and selective degradation 2:268–269
- conserved biochemical pathways 2:250–251
- direct targeting and insertion of peroxisome membrane proteins 2:262–263
- division 2:259–261, 2:260F
- ER contribution to peroxisome membrane growth 2:263–267
- fatty acid β -oxidation 2:252, 2:253F
- functions
- bile acid synthesis 2:254
- biotin synthesis 2:254
- cholesterol, steroids, isoprenoids, and wax ester synthesis 2:254
- fatty acid α -oxidation 2:254
- fungal sex 2:254
- metabolic cooperation and interaction 2:255–256
- pathogen defense and pathogen offense 2:255
- polyamine catabolism 2:254
- purine catabolism 2:254
- signaling 2:254
- fusion 2:261
- gene regulation 2:257–259
- glyoxylate cycle and glyoxylate metabolism 2:252–254
- human disease and 2:269–271
- inheritance 2:261
- reactive oxygen and nitrogen metabolism
- H₂O₂ metabolism 2:250–251
- protection against H₂O₂/use of H₂O₂ 2:251–252
- reactive oxygen and nitrogen species 2:252
- specialized physiological functions
- fat to sugar conversion 2:248
- glycolysis in trypanosomes 2:249–250

- peroxisomes (*continued*)
 myelin formation and maintenance 2:248
 photorespiration in green leaves 2:248–249, 2:251F
 plugging holes in fungi 2:250
 sperm formation 2:248
 targeting and import of enzymes into 2:262
peroxynitrite (ONOO⁻) 2:767–768
persistent hyperplastic primary vitreous (PHPV) 3:455
pertussis toxin (PTX) 2:795
peripheral lymph node addressin (PNAd) 3:635–636
PEST domain *see* proline, glutamic acid, serine, threonine domain (PEST domain)
PEST domain-enriched tyrosine phosphatase (PEP) 3:537
PET *see* positron emission tomography (PET)
Pex3 site 2:263
Pex7 site 2:262
Pex11 site 2:259
Pex15 site 2:263
Pex16 site 2:262–263
Pex19 site 2:262–263
PF *see* P-face (PF)
P-face (PF) 2:39
PFK-1 *see* phosphofructokinase-1 (PFK-1)
PFK-2 *see* phosphofructokinase-2 (PFK-2)
PFO *see* perfringolysin O (PFO)
PFTs *see* pore-forming toxins (PFTs)
Pg *see* plakoglobin (Pg)
PGAM5 *see* phosphoglycerate mutase family member 5 (PGAM5)
PGC *see* primordial germ cells (PGC)
pGCs *see* Particulate GCs (pGCs)
PGI *see* protein–glycan interactions (PGI)
PGN *see* peptidoglycan (PGN)
PGs *see* periglomerular cells (PGs)
PH domain *see* pleckstrin homology domain (PH domain)
pH2AT120 *see* phosphorylated histone 2A at threonine 120 (pH2AT120)
pH3T3 *see* phosphorylated histone H3 at threonine 3 (pH3T3)
phagocyte NADPH oxidase (NOX2) 3:702
 active NADPH oxidase complex 3:702F
 cytosolic components structure in resting cells 3:703–704
 flavocytochrome b₅₅₈ in resting cells 3:703, 3:703F
 membrane component structure 3:703
 NADPH oxidase complex assembly and activation 3:705–706
 regulation of NADPH oxidase assembly and activation 3:706
 ROS reactions in phagosome 3:702F
phagocytes 3:6, 3:328
 and receptors 3:521
phagocytic cells, professional and nonprofessional 2:751
phagocytic/phagocyte(s) 2:766–767
 attack 2:782
 cup 2:752
 killing 2:767
 oxidative burst 3:702
 oxygen-dependent killing 2:767
 CGD 2:768
 iNOS 2:767–768
 NADPH oxidase and MPO 2:767
 oxygen-independent killing 2:771
 phagosome maturation 2:768–770
 macrophage phagosome maturation 2:768–770, 2:769F
 neutrophil phagosome maturation 2:769F, 2:770–771
 synapses 3:695–701, 3:695
 bull's-eyes 3:699F
 Dectin-1 activation 3:696
 FcγR activation 3:696–697
 immunological and 3:695F
 ITAM signaling 3:697F
 model 3:697–700, 3:698F
 receptors 2:752
 regulation by CD47 3:700–701
phagocytosis 2:508
phagocytosis 1:764, 2:4, 2:177, 2:751–757, 2:766–767, 3:695, 3:752
 experimental models 2:752F
 focal exocytosis 2:754
 inactivation after intimate contact with pathogens 2:776–777
 linking actin to polymerization/depolymerization cycle 2:754–755
 mechanisms 2:766–767
 membrane remodeling 2:754
 phagosome formation steps 2:753F
 receptors 2:766F
 and signaling pathways 2:753–754
 signaling to actin polymerization 2:752–753
phagolysosomes 1:764
phagophore 1:588, 1:764–765
phagophore assembly site (PAS) 2:245, 2:445, 3:57
phagophore(s) 2:245, 2:771–772
phagosome 3:702
 acidification 2:771, 2:777
 formation 2:753F
 maturation 2:766, 2:768–770
 macrophage phagosome maturation 2:768–770, 2:769F
 neutrophil phagosome maturation 2:769F, 2:770–771
 maturation into phagolysosomes 2:755–756, 2:755F
phakinin 2:569
 mutations 2:576
phaphin2 2:504
'Phasor' approach 2:117
PHB *see* prohibitins (PHB)
PH-BAR *see* pleckstrin homology-containing BAR proteins (PH-BAR)
PHD *see* plant homeodomain (PHD); prolyl 4-hydroxylase domain (PHD)
Phe-Ala-Asn-Ser-His-Tyr (FANSHY) 2:478
Phe-any-Phe-Gly (EXFG) 2:319–320
Phe-Gly (FG) 2:319–320
Phe-Leu-Val (FLV) 2:478
phenylalanine-glycine (FG) 3:175
 repeats 2:313
4-phenylbutyric acid (PBA) 1:797
phenylthiocarbamyl derivative 1:75–76
pheromone receptor Ste2p 2:390
phosphatidylinositol-4-phosphate (PI4-P) 2:822–824
phorbol myristate acetate (PMA) 3:705
phosphatase 4:227, 4:228–229, 4:230–231
phosphatase and tensin homolog (PTEN homolog) 2:504, 3:132, 3:781, 3:814–815
 PTEN-opathies 3:418–419
phosphatase exclusion, on TCR signaling 3:651–653
phosphatidic acid (PA) 1:155, 1:195, 1:196–197, 2:276, 2:761
 synthesis 1:161–162
phosphatidylinositol (PI) 2:785
phosphatidylinositol 3-kinase (PI3K) 2:785, 3:815
phosphatidyl inositol 3-phosphate (PtdIns 3P) 2:475
phosphatidylinositol phosphatases (PIPs) 3:814–815
phosphatidyl serine (PS) 1:715–716, 4:54
phosphatidylcholine (PC) 1:155, 1:162, 1:166, 1:202, 2:273
phosphatidylcholine phospholipase C (PC-PLC) 2:780
phosphatidylethanolamine (PE) 1:162, 1:164–165, 1:165–166, 1:170, 1:202, 1:588, 2:245–246
phosphatidylglucose 1:183
phosphatidylglycerol 1:162, 1:166–167, 1:168, 1:170
phosphatidylglycerophosphate synthesis 1:168
phosphatidylinositide-3-kinase (PI3K) *see* phosphatidylinositol 3-kinase (PI3K)
phosphatidylinositol (3)-phosphate (PtdIns(3)P) 2:202–203, 2:504, 2:515
phosphatidylinositol (3,4)-bisphosphate (PtdIns(3,4)P₂) 2:504
phosphatidylinositol (3,5)-bisphosphate (PtdIns(3,5)P₂) 2:504
phosphatidylinositol (4)-phosphate (PtdIns(4)P) 2:504, 2:505
phosphatidylinositol (4,5)-bisphosphate (PtdIns(4,5)P₂) 2:504, 2:507–508, 2:508, 2:759, 2:760
phosphatidylinositol (PI) 1:166–167, 1:167, 2:195, 2:760
phosphatidylinositol 3-phosphate (PI(3)P) 2:195, 2:217–219, 2:777
phosphatidylinositol 3,4,5-trisphosphate (PI(3,4,5)P₃) 2:752–753, 3:781, 4:6, 4:221
phosphatidylinositol 4-phosphate (PI4P) 2:366, 2:366, 2:451–452
phosphatidylinositol 4,5-bisphosphate (PI(4,5)P₂) 2:752–753, 2:754–755, 2:830, 3:153
phosphatidylinositol lipids 1:196–197

- phosphatidylinositol phosphate kinases (PIPKs) 2:754–755
- phosphatidylinositol-3-kinase (PI3K) 1:766, 3:12, 3:128, 3:186–187, 3:227, 3:243, 3:413, 3:434, 3:541, 3:770–771, 4:105–106, 4:218–219
- class domain structure 3:129F
- class I PI3K 3:128
- class II PI3K 3:128–129
- class III PI3K 3:129
- core 3:128
- inactivation 3:132
- pathway 3:413, 3:415–417
- PI3K-AKT pathway 3:420
- type I 2:752–753
- phosphatidylinositol-4-kinase-III β (PI4-III β) 2:822–824
- phosphatidylinositols (PI) 1:195
- phosphatidylinositol-specific phospholipase C (PI-PLC) 2:780
- phosphatidylserine (PS) 1:202, 1:202–203, 1:203, 1:204–205, 1:204F, 2:197, 2:811, 3:732
- phosphodiesterases (PDEs) 3:57, 3:59–60
- phosphoenolpyruvate carboxykinase (PEPCK) 1:294, 1:296
- phosphofructokinase 4:24, 4:28
- phosphofructokinase-2 (PFK-2) 1:292–293, 1:296
- phosphofructosekinase-1 (PFK-1) 1:292–293
- phosphofurin acidic cluster sorting protein 1 (PACS1) 2:436
- phosphoglucan water dikinase (PWD) 1:268–269
- 6-phosphogluconate dehydrogenase 1:294
- phosphoglycerate mutase family member 5 (PGAM5) 3:344
- phosphoinositide-specific PLC enzymes 3:153
- PLCs domain organization and structure 3:153
- signaling connectivity, function, and dysfunction 3:153
- PLC β family 3:155–156
- PLC γ family 3:156–157
- PLC δ family 3:153–155
- PLC ϵ 1 family 3:157–158
- PLC ζ 1 family 3:158–159
- PLC η family 3:158
- phosphoinositide 1:198–199, 1:199
- phosphoinositide 3-kinase (PI3K)
- see* phosphatidylinositol-3-kinase (PI3K)
- phosphoinositide 5-phosphatase (Inpp5e) 2:504
- phosphoinositide phosphates (PIPs) 2:217, 2:217–219
- phosphoinositide-3-kinase (PI3K)
- see* phosphatidylinositol-3-kinase (PI3K)
- phosphoinositide-3-phosphate (PtdIns(3)P) 2:505
- phosphoinositide-dependent kinase 1 (PDK1) 3:128
- 3-phosphoinositide dependent protein kinase 1 (3-PDPK1) 3:245, 4:221
- phosphoinositides (PIs) 2:194–195, 2:195, 2:503–519, 2:520–522, 2:525–526, 2:746, 2:759, 2:760
- in autophagy regulation 2:515–516, 2:516F
- binding domains 2:507T
- coordination 2:195–196
- in endosomal compartment 2:510–512
- in Golgi trafficking 2:512–513, 2:514F
- at membrane–cytosol interface 2:505
- organelle compartmenting spatial restriction 2:505–507
- in plasma membrane trafficking
- endocytic steps and phosphoinositide engagement 2:509F
- endocytosis 2:507–510
- exocytosis 2:510
- retrograde trafficking from endosomes to Golgi 2:513–515
- subcellular distribution to membrane compartments 2:508F
- turnover by kinases and phosphatases 2:506F
- phosphoinositol 3-kinase (PI3K)
- see* phosphatidylinositol-3-kinase (PI3K)
- phosphoinositol-3,5-bisphosphate (PI(3,5)P₂) 2:410
- phosphoinositol-3-phosphate (PI(3)P) 2:409–410
- phosphoinositol-3-phosphate (PtdIns3P) 2:245
- phospholipase A2 (PLA2) 2:367
- phospholipase C (PLC) 2:754–755, 3:2, 3:153, 3:156–157
- domain organization 3:154F
- expression, function and dysfunction 3:155T
- phosphoinositide-specific PLC enzymes 3:153
- PLCs domain organization and structure 3:153
- signaling connectivity, function, and dysfunction 3:153–155
- PLC β family 3:155–156
- PLC δ family 3:153–155
- PLC ϵ 1 family 3:157–158
- PLC γ 2:754–755, 3:12, 3:573–574, 4:105–106
- PLC γ 1 2:759, 3:534–535, 4:82
- PLC η family 3:158
- PLC ζ 1 family 3:158–159
- phospholipase D1 (PLD1) 2:759–760
- phospholipid(s) 2:267, 2:273, 2:275–276
- monolayer 2:274
- synthesis in bacteria 1:162
- amine-containing phospholipids
- synthesis 1:162
- non-amine-containing phospholipids
- synthesis 1:162–164
- synthesis in eukaryotes 1:164
- amine-containing phospholipids
- synthesis 1:164–166, 1:165F
- non-amine-containing phospholipids
- synthesis 1:166–169
- phosphomimetic substitution in FGF2 (FGF2-Y81pCMF) 2:523
- phosphopeptide
- enrichment 4:178–179
- fragmentation 4:180–182, 4:181F
- phosphoproteins 4:228–229
- phosphoproteomics 4:177–186
- challenges in 4:183
- large-scale phosphoproteomics strategies 4:179–180, 4:179F
- protein phosphorylation 4:177–178
- PTM 4:177
- PTM cross talk and PTM coding 4:182–183
- quantitative phosphoproteomics 4:182
- phosphoramidite 1:36–40, 1:41–43, 1:50
- phosphorylase kinase 3:166
- phosphorylated histone 2A at threonine 120 (pH2AT120) 3:473
- phosphorylated histone H3 at threonine 3 (pH3T3) 3:473
- phosphorylated–p47phox–Pin1 axis 3:706
- phosphorylation 1:84–85, 1:267, 2:220–221, 2:287, 2:573, 2:573–574, 3:122, 3:123, 3:123–124, 3:124, 3:125, 3:125–126, 4:45, 4:47, 4:142–143, 4:145, 4:145–146, 4:146–147
- cascade 1:85
- gp91phox 3:707
- p22phox 3:707
- p40phox 3:706–707
- p47phox 3:706
- p67phox 3:706–707
- phosphorylation-deficient vimentin mutants 2:573
- phosphorylation-dependent signaling, mapping
- direct and indirect candidate kinase targets 4:229
- hypothetical phosphoprofiles clustering 4:230F
- integrating systems genetics and proteomics 4:231F
- kinase–substrate relationships 4:231–232
- mapping network connectivity 4:229–230
- methods to link phosphorylation 4:227F
- phosphoproteins 4:228–229
- protein phosphorylation 4:227
- phosphorylation–dephosphorylation cycles 4:44, 4:45, 4:45F
- phosphorylation directed MS³ (pdMS³) 4:180
- phosphorylation potential 1:288, 1:293–294
- phosphorylation sites 1:124, 1:507
- phospho-transferase system (PTS) 1:101
- phosphotransferring enzymes 1:17
- phosphotyrosine binding domain (PTB domain) 1:92, 2:423, 2:455–456, 2:455, 2:503, 3:31–32, 3:115
- photoact 2:328
- photoactivatable fluorescent proteins (PA-FP) 2:52
- photoactivatable localization microscopy (PALM) 2:70

- photoactivatable ribonucleoside CLIP (PAR-CLIP) 4:192
- photo-activated adenylyl cyclases (PACs) 2:149
- photo-activation localization microscopy (PALM) 2:58F, 2:62, 2:76, 2:122
- photo-agonism 2:152–153
- Photobacterium damsela* subsp. *piscicida* 2:795
- photoconvertible fluorescent proteins (PC-FPs) 2:52
- photolithography with physical/digital masks 1:46–47
- photo-multiplier tubes (PMTs) 4:238–239
- photon count histogram (PCH) 2:103
- photon efficiency 2:112–113
- 2-photon microscopy 2:55, 2:55F
- photorespiration in green leaves 2:248–249, 2:251F
- photoswitch 2:147, 2:147–148, 2:151
- chemical 2:148T
- reversible 2:151
- Photoswitchable tethered ligands (PTLs) 2:151, 2:152F
- photosynthetic/photosynthesis 2:328–329
- cytochrome *b₆* complex 1:301–303
- as lipoprotein 1:302–303
- structures, mechanisms, and mechanisms of H⁺ transfer 1:303
- pathways of proton uptake into membrane 1:303–304
- proton donation to p-side aqueous phase 1:304
- reaction centers 1:299
- see also organelle bioenergetics
- anoxic photosynthetic bacteria 1:299–300
- photosynthetic plants, algae, and cyanobacteria 1:300–301
- Phox Bem1p (PB1) 1:92
- phox homolog domain (PX domain) 3:128
- phox homology (PX) domain 2:475
- phoxdomain-containing domains (PX-BAR domains) 2:491
- PHPV see persistent hyperplastic primary vitreous (PHPV)
- PHS see PTEN hamartomatous syndromes (PHS)
- phylogenetics 4:154, 4:157–158
- phytohormones see plant hormones
- physical stresses 3:123–124, 3:124
- physiologic inflammation 3:374
- physiological lipid transport, Class B SRs mediating 3:731–732
- physiological membrane potential measurement, caveats in 1:232–234
- physiological/pathological processes, signals to migration in 3:210
- phytoceramides see 4-hydroxyDHCer
- phytylglycerol 1:267–268
- PI see phosphatidylinositol (PI)
- PI 3-kinase inhibitor 2:762
- PI see phosphatidylinositol (PI); phosphatidylinositols (PI)
- PI3 kinase see phosphoinositol 3-kinase (PI3 kinase)
- PI3K see phosphatidylinositol-3-kinase (PI3K)
- PI3K/Akt/mTOR 3:2, 3:128, 3:131F
- Akt 3:129–130
- mTOR 3:130–132
- pathway regulation 3:132
- negative regulation 3:132
- PI3K 3:128
- signaling in human disorders 3:133–134
- PI3KC see class I PI3-kinase (PI3KC)
- PI3KC2A/PI3KC2 α see class II PI3-kinase (PI3KC2A/PI3KC2 α)
- PI-4,5-bisphosphate (PIP2) 3:697
- PI4-III β see phosphatidylinositol-4-kinase-III β (PI4-III β)
- PI4-kinase III α 2:513
- PI4-kinase III β 2:513
- PI4-P see phosphatidylinositol-4-phosphate (PI4-P)
- PI-4-phosphate-5-kinase (PIP5K) 2:785
- PIC see pre-initiation complex (PIC)
- PICH see Plk1-interactingcheckpointthelase (PICH)
- Pichia pastoris* 2:333
- picornaviruses 3:307
- formation of membrane tubes and DMVs by 2:823F
- membrane tubes and DMVs generation by 2:822
- PICS see protease cleavage site specificity (PICS)
- PICS see proteomic identification of cleavage site specificity (PICS)
- PicTar 1:365
- PIDD see p53-induced death domain protein (PIDD)
- pigment synthesis 2:329
- pIgR see polymeric IgA receptor (pIgR)
- PIK1 see class III PI4-kinase (PIK1)
- PIK-1 see Pelle and IL-1R-associatedkinases (PIK-1)
- PIKfyve/PIP5K3 2:504
- PI-kinases 1:204–205
- PIN efflux proteins 4:129, 4:130–131
- ping pong mechanisms 1:16–17, 1:17
- PinkBAR 2:497
- pinocytosis 2:758
- solute flow during 2:759
- PIP2 see PI-4,5-bisphosphate (PIP2)
- PIP5K see PI-4-phosphate-5-kinase (PIP5K)
- PI-phosphatases 1:204–205
- PIPKIs see phosphatidylinositol phosphate kinases (PIPKIs)
- PI-PLC see phosphatidylinositol-specific phospholipase C (PI-PLC)
- PIPs see phosphatidylinositol phosphatases (PIPs); phosphoinositide phosphates (PIPs)
- piRISC see piRNA-silencing complex (piRISC)
- piRNA see Piwi-interacting RNA (piRNA)
- piRNA-silencing complex (piRISC) 3:322
- PIs see phosphoinositides (PIs)
- pivaloyloxymethyl (PivOM) group 1:46
- PivOM group see pivaloyloxymethyl (PivOM) group
- Piwi-interacting RNA (piRNA) 1:370, 4:160, 4:162
- pixel negative values 2:90–91
- pk* see *Prickle* (*pk*)
- PKA see cAMP-dependent protein kinase (PKA); protein kinase A (PKA)
- PKC see protein kinase C (PKC)
- PKC ϵ see protein kinase C ϵ (PKC ϵ)
- PKD see polycystic kidney disease (PKD); protein kinase D (PKD)
- PKG see cGMP-dependent protein kinase (PKG)
- Pkp see plakophilins (Pkp)
- PKR-related ER kinase (PERK) 2:529
- PLA2 see phospholipase A2 (PLA2)
- placental growth factor (PLGF) 4:106
- PLAD see pre-ligand binding assembly domain (PLAD)
- plakins 2:572–573
- plakoglobin (Pg) 2:707
- plakophilins (Pkp) 2:707
- planar cell polarity (PCP) 2:698, 2:742, 3:99, 3:99, 3:199
- see also epithelial cells
- asymmetric distribution of core PCP proteins 2:747F
- molecular basis 2:742
- signaling 3:200, 3:200F
- plant homeodomain (PHD) 3:555–556
- plant(s) 2:149
- clathrin roles in 2:390
- hormones 2:254
- immune defense 2:255
- plasma KLK 1:699
- plasma membrane (PM) 1:202, 1:203, 1:205, 1:206, 1:229–230, 1:556, 2:2, 2:168, 2:225–226, 2:226, 2:228, 2:229–230, 2:384, 2:435, 2:466, 2:752, 2:817, 2:828, 3:1–2, 3:123
- deformation 2:63–64
- phosphoinositides in trafficking
- endocytic steps and phosphoinositide engagement 2:509F
- endocytosis 2:507–510
- exocytosis 2:510
- proteins 2:459
- resealing 2:471
- roles in clathrin
- particle uptake and domain organization 2:389
- RME 2:388–389
- topology 3:678–680
- virus budding at 2:817F
- plasma membrane calcium ATPase (PMCA) 1:675, 2:159
- plasmacytoid dendritic cells (pDCs) 3:218, 3:592, 3:592–593, 3:596, 3:597, 3:742, 3:742
- plasmalemma 1:763, 1:765, 1:771
- plasmalogens 2:248, 2:250F
- plasminogen activation system 1:756
- plasminogen activator inhibitor-1 (PAI-1) 1:717, 1:719, 1:756, 2:227–228
- plasminogen activator inhibitor-2 (PAI-2) 1:711
- plastic embedding 2:34

- plasticity 3:791, 3:809
 plastids 2:324–330, 2:324, 3:294
 sharing progenitor 2:325F
 types 2:326–327F
 viridiplantae
 amyloplast 2:325–326
 chloroplast 2:327
 chromoplast 2:326–327
 elaioplast 2:324–325
 etioplast 2:324
 gerontoplast 2:327
 proplastid 2:324
 plastocyanin (PC) 2:328
 plastoquinone (PQ) 1:301, 2:328
 platelet glycoprotein IV deficiency *see* CD36 deficiency
 platelet(s) 2:227–228, 3:196
 activating factor 2:254–255
 dense granule 2:225
 platelet-derived growth factor (PDGF)
 2:727–728, 3:208–209, 3:219–220, 3:224, 4:102, 4:219, 4:108, 4:166–176
 ligands 4:108
 PDGFR signaling 4:108
 receptors 4:108
 platelet-derived growth factor A (PDGFA) 3:795
 platelet-derived growth factor receptors (PDGFRs) 4:77–78, 4:102, 4:108
 platelet-derived growth factor receptor β (PDGFR β) 3:455, 3:797
 platelet-endothelial cell adhesion molecule-1 (PCAM-1) 3:624–625, 3:634, 3:752
 PLB *see* prolamellar body (PLB)
 PLC *see* peptide-loading complex (PLC); phospholipase C (PLC); type C phospholipases (PLC)
 PLD1 *see* phospholipase D1 (PLD1)
 pleckstrin homology domain (PH domain) 1:92, 1:240, 2:197, 2:446, 2:503, 2:597, 2:731–733, 3:128, 3:153, 4:218, 4:221
 pleckstrin homology domain-containing family A member 7 (PLEKHA7) 2:708–709
 pleckstrin homology domain-containing BAR proteins (PH-BAR) 2:491, 2:498
 plectin 2:572–573
 pleiomorphic 2:171, 2:173–175, 2:175
 pleiotrophin 3:16
 pleiotropic kinin 1:700
 PLEKHA7 *see* pleckstrin homology domain-containing family A member 7 (PLEKHA7)
 pleuropulmonary blastoma 1:360
 PLGF *see* placental growth factor (PLGF)
 pLGIC *see* pentameric ligand-gated ion channels (pLGIC)
 PLIs *see* protein–lipid interactions (PLIs)
 Plk1 *see* polo-like kinase 1 (Plk1)
 Plk1-interacting checkpoint helicase (PICH) 3:501
 PLK4 *see* polo-like kinase 4 (PLK4)
 PLP *see* pericentrin-like protein (PLP)
- PLS *see* Papillon–Lefevre syndrome (PLS)
 PLT protocol *see* progressive-lowering of temperature protocol (PLT protocol)
 plunge freezing 2:22–23, 2:23F, 2:24F, 2:24
 pluripotency 3:819
 pluripotent cells 3:819–822
 cell-based regenerative therapy, tools for 3:820–821
 clinical trials and contributions 3:821
 determining cell fates and disease states 3:819–820
 engineered 3:819
 naturally occurring 3:819
 as tools for disease research and therapies 3:820F
 plus end tracking proteins (+ TIPs) 2:539, 2:544, 3:232
 plus ends 2:539, 2:544, 2:708
 plus-end tip 4:36, 4:37–38, 4:38, 4:39F, 4:40, 4:42
 structures 4:38, 4:41F
 PM *see* plasma membrane (PM)
 PM optical fiber *see* polarization-maintaining optical fiber (PM optical fiber)
 PMA *see* phorbol myristate acetate (PMA)
 PMCA *see* plasma membrane calcium ATPase (PMCA)
 PMF *see* peptide mass fingerprinting (PMF)
 pMHC *see* peptide bound to major histocompatibility complex molecule (pMHC)
 p-MHC *see* peptides-major histocompatibility molecules (p-MHC)
 pMHC-I *see* peptide MHC class I complexes (pMHC-I)
 PML *see* promyelocytic leukemia (PML)
 PML-NB *see* promyelocytic-leukemia nuclear body (PML-NB)
 PMNs *see* polymorphonuclear cells (PMNs)
 PMTs *see* photo-multiplier tubes (PMTs)
 PNAd *see* peripheral lymph node addressin (PNAd)
 PNKP *see* polynucleotide kinase phosphatase (PNKP)
 PNPase *see* polynucleotide phosphorylase (PNPase)
 Pnr production *see* Pannier production (Pnr production)
 PNS *see* peripheral nervous system (PNS)
 podocalyxin (PODXL) 4:112
 podosome 3:637–638, 3:644
 PODXL *see* podocalyxin (PODXL)
 POFUT2 *see* protein O-fucosyltransferase 2 (POFUT2)
 point mutation 1:125, 1:333–334, 1:563
 point spread function (PSF) 2:56, 2:63, 2:70, 2:86–87, 2:121
 pointed end 2:683, 2:685, 2:689
 binding proteins 2:557–558
 pointed face 2:549
 pol I transcription 3:3
 Pol II C-Terminal Domain 1:483
 polar ejection forces (PEFs) 2:638, 2:640–641
- polarity 1:475, 2:552–553, 2:649, 3:688
 complexes 2:742–743
 establishment 2:655
 in eye 3:199
 polarization-maintaining optical fiber (PM optical fiber) 2:78
 PM-derived SNARE proteins 2:779
 polarized/polarization
 epithelial cells 2:401
 evanescent field 2:64F
 sorting 2:617
 TIR combined with 2:63–64
 trafficking machinery 2:745–746
 phosphoinositides and 2:746
 transcytosis and lumen formation 2:744F, 2:745–746
 transport 2:364
 poleward microtubule flux 2:640–641
 polio-virus (PV) 3:307, 3:308
 polo-like kinase 1 (Plk1) 2:317, 2:625, 3:495, 3:496
 temporal regulation 3:499–501
 polo-like kinase 4 (PLK4) 2:651–653
 poly(A) binding protein (PABP) 1:344, 1:346, 1:527, 1:528F, 1:531, 1:533
 poly(A) tail 1:343, 1:344, 3:307
 poly(A)-binding protein (PABP) 3:317
 poly(A)-nuclease 2 (PAN2) 1:346
 poly(A)-specific ribonuclease (PARN) 1:346
 poly-ADP ribose polymerase (PARP) 1:465, 3:573–574
 PARP-1 1:436, 2:614
 polyamine catabolism 2:254
 polybrene[®] 1:77–78
 poly C binding protein 2 (PCBP2) 3:313
 polychlorinated biphenyl (PCB) 2:276
 polycistronic mRNA 1:341, 1:342
 polycomb group genes (PcG) 3:438
 Polycomb repressive complex (PRC) 1:484
 polycomb repressive complexes 1 (PRC1) 3:453
 polycomb repressive complex 2 (PRC2) 3:453, 3:778, 3:790
 polycystic kidney disease (PKD) 2:748
 polycythemia vera (PCV) 3:146
 polydactyly 3:79, 3:80
 polyethylene glycol (PEG) 1:125–126
 polyketide synthase 1:14
 polyketides 1:156
 polymerase 1:122, 1:418, 1:439, 1:520, 1:520–521
 polymerase, switching sequence 1:477
 polymerase chain reaction (PCR) 1:30–31, 1:66, 1:67, 1:122, 1:125, 3:136, 4:162, 4:220
 polymeric IgA receptor (pIgR) 2:444, 3:691
 polymerization 2:103, 4:14–15
 polymorphonuclear cells (PMNs) 3:702, 3:750
 polynucleotide kinase phosphatase (PNKP) 1:442–443, 3:462
 polynucleotide phosphorylase (PNPase) 3:293
 polypeptides 2:666–667
 poly-proline type II (PPII) 3:114–115

- polypyrimidine tract (PY tract) 1:503, 1:503–504, 1:504F
- polypyrimidine tract binding protein (PTB) 1:503–504, 3:313
- polysaccharides 1:263
- polysialic acid-neuronal cell adhesion molecule (PSA-NCAM) 4:137
- polysomes 1:350–351
- polytene chromosomes 3:423F
- polytopic integral membrane proteins 2:274
- polytopic membrane protein 2:273
- polyubiquitination 2:528, 2:529
- polyubiquitination, monoubiquitination vs. 3:505
- POMC *see* pro-opiomelanocortin (POMC)
- Pompe disease 1:764, 1:767, 1:768, 1:768–769, 1:771–772, 1:779
rôle of autophagy in glycogen remodeling 1:770
- Poms *see* pore membrane proteins (Poms)
- POMT1 *see* protein O-mannosyltransferase 1 (POMT1)
- POP3 *see* pyrin domain only protein 3 (POP3)
- porcupine (Porc) 3:87, 3:87F
- pore membrane proteins (Poms) 3:175
- pore-forming toxins (PFTs) 2:776, 2:794, 2:846
- positive feedback 4:31, 4:67
signal amplification through 4:8
- positive reference set (PRS) 4:188
- positive selection 4:156
- positive-transcriptional elongation factor b (P-TEF-b) 2:835
- positron emission tomography (PET) 1:143
- post synaptic receptors 1:248–249
- post-cleavage complex (PCC) 3:558
- post-embedding labeling *see* on-section labeling
- post-Golgi transport 2:363–370
see also intra-Golgi transport
constitutive transport via endosomal hubs 2:364–365
lipid rafts in sorting and transport 2:365–366
modulating constitutive transport 2:363–364
in polarized cells 2:364F
polarized transport 2:364
TGN-derived constitutive transport 2:363
TGN
cargo sorting and packaging at 2:365–366
protein-based sorting in 2:366
tubules and carriers 2:366–367
- postpartum breast cancer 3:376
- postsynaptic density (PSD) 2:294
- postsynaptic response 1:248
- posttranscriptional control 3:553–554
- posttranslational modification and enzyme activity 1:290
- posttranslational modifications (PTMs) 1:22–23, 1:75, 1:84, 1:507, 1:557–558, 1:563, 2:176, 2:294, 2:573–574, 2:617, 3:719, 3:722–723, 3:785, 4:145–146, 4:166, 4:170, 4:177
- p53 transactivation through 3:506–507
amidation 1:86–87
carbonylation 1:87
cross talk and coding 4:182–183
of enzymes 1:14
 γ -carboxylation 1:88
glycosylation 1:85–86
hydroxylation 1:87
lipid modification
GPI 1:87
myristoylation 1:87
palmitoylation 1:87
protein prenylation 1:87
methylation 1:87
N-acetylation 1:86
nitrosylation 1:87–88
oxidative stress-related PTMs 1:87
phosphorylation 1:84–85
proteolytic PTM of proteins 1:84
ubiquitin and targeted protein degradation 1:88
- posttranslational protein modifications 3:1
- POT1 *see* protection of telomeres protein 1 (POT1)
- potentiator 1:252, 1:261
- power
law 4:56, 4:59, 4:193–194
stroke 2:623
- PP *see* protein phosphatase (PP)
- PP1 *see* protein phosphatase 1 (PP1)
- 5'-ppp *see* 5' triphosphate (5'-ppp)
- PP2A *see* Protein phosphatase 2A (PP2A)
- PPARs *see* peroxisome proliferator-activated receptors (PPARs)
- ppGpp small molecule effector 1:468–469
- PPlases *see* peptidylprolyl isomerases (PPlases)
- PPIF *see* peptidylprolyl isomerase F (PPIF)
- PPII *see* poly-proline type II (PPII)
- PPIs *see* protein-protein interactions (PPIs)
- p-pol incident light 2:63–64
- PPP *see* pentose phosphate pathway (PPP)
- PPSPxS 3:91, 3:91F, 3:92, 3:93F
- PQ *see* plastoquinone (PQ)
- PQD *see* pulsed Q-dissociation (PQD)
- pQTL analysis *see* protein quantitative trait loci analysis (pQTL analysis)
- PR *see* protease (PR); retroviral protease (PR)
- PR3 *see* proteinase 3 (PR3)
- PRAS40 *see* proline-rich Akt substrate of 40 kDa (PRAS40)
- pRb *see* retinoblastoma protein (pRb)
- PRC *see* Polycomb repressive complex (PRC)
- PRC1 *see* polycomb repressive complexes 1 (PRC1); protein regulator of cytokinesis-1 (PRC1)
- PRC2 *see* polycomb repressive complex 2 (PRC2)
- PRD *see* C-terminal proline-arginine domain (PRD); proline-rich domain (PRD)
- PRE *see* paramagnetic relaxation enhancement (PRE)
- pre-autophagosomal structure
see phagophore assembly site (PAS)
- precision genome editing 1:515
- precursor miRNA (pre-miRNA) 1:354, 1:358, 1:359, 1:360, 1:360–361
cleavage by Dicer 1:355–358
Drosha cleavage 1:355
transportation by Exportin-5/Ran-GTP complex 1:355
- precursor RNA processing proteins (Prps) 1:503–504
- precursor rRNA, cleavage and folding of 3:289–291
- precursor with cDCs (pre-DC) 3:743
- pre-DC *see* precursor with cDCs (pre-DC)
- pre-docking process 2:380–381
- preembedding labeling 2:36–37
- pre-fractionation 4:167
- pregnancy, changes in gut microbiome in 2:801
- pregnenolone 1:178
- pre-initiation complex (PIC) 1:482–483, 2:833, 3:258
- pre-integration complex (PIC) *see* pre-initiation complex (PIC)
- prekallikrein (preKLK) 1:699
- pre-ligand binding assembly domain (PLAD) 3:61–62
- premature stop codon (PTC) 3:317
- pre-miR-451 1:359
- pre-miRNA *see* precursor miRNA (pre-miRNA)
- pre-mRNA splicing 1:377, 1:495–502, 1:503–511
alternative splicing 1:505–506
cis-acting sequences and trans-acting factors 1:505–506
signaling pathways 1:506–507
snRNP biogenesis 1:507
transcription, chromatin structure, and splicing 1:506
diseases of spliceosome 1:507–508
splicing factor mutations and cancer 1:508
splicing mutations in growth and development 1:508–509
splicing mutations in retinitis pigmentosa and neurodegeneration 1:508
spliceosome 1:503–505
splicing catalysis 1:503
- prenol lipids 1:155–156
- presenilin 1:664
mutations 1:662
- presomitic mesoderm (PSM) 4:128
- Prickle (*pk*) 2:746–747
- primary cilia 2:649, 2:655, 2:660
and hedgehog signaling 2:668
in regulating mammalian HH signaling 3:79
sensory capacity 2:668
- primary ciliary dyskinesia (PCD) 2:631, 2:669–670
- primary cilium *see* primary cilia
- primary granules 2:226, 2:770, 3:751
- primary hemostasis

- platelet production 1:714–715
platelet structure and function in 1:715–716
- primary lymphoid tissues 3:6
- primary mammalian signaling lipids 1:195–196, 1:195F
- primary miRNA (pri-miRNA) 1:354, 1:356–357F, 1:357F, 1:358, 1:359, 1:360
cleavage into pre-miRNAs, by Drosha 1:355
transcription 1:355
- primases 1:405
bacterial 1:405
eukaryotic 1:405–406
primpol 1:406–407
RNA 1:396, 1:397
- primer binding site (PBS) 1:327, 1:521, 2:833
- pri-miRNA *see* primary miRNA (pri-miRNA)
- primordial germ cells (PGC) 3:819
- primpol primase 1:406–407
- principle component analysis (PCA) 4:207
- prion gene (PRNP) 2:840, 2:841
- prion protein (PrP) 1:113–114, 2:839, 2:841
- prion(s) 2:839–844
conversion 2:841–842
discovery 2:839
diseases 2:839, 2:839–840
animal prion diseases 2:841
human prion diseases 2:840
neurotoxicity 2:842–843
replication 2:841–842
strains 2:842
transmission 2:842
- PRIs *see* protein–RNA interactions (PRIs)
- Prize-collecting Steiner approaches 4:215F
- Prize-collecting Steiner Forest 4:212T, 4:215–216
- Prize-collecting Steiner Tree (PCST) 4:212T, 4:215–216, 4:216
- Prize-collecting Steiner Tree methods (PCST methods) 4:215–216, 4:215F
- PRL *see* prolactin (PRL)
- PRLYL residues *see* Pro-Arg-Leu-Tyr-Leu residues (PRLYL residues)
- PRM *see* parallel reaction monitoring (PRM)
- pRNA *see* promoter RNA (pRNA)
- PRNP *see* prion gene (PRNP)
- proangiogenic MMP-9 delivering neutrophils 3:754
- pro-ANP *see* pro-atrial natriuretic peptide (pro-ANP)
- pro-apoptotic machinery 3:328–330, 3:330
B-cell lymphoma-2 (Bcl-2) family 3:330
caspases 3:329–330
lipids 3:330
- Pro-Arg-Leu-Tyr-Leu residues (PRLYL residues) 2:475–476
- pro-atrial natriuretic peptide (pro-ANP) 1:656–657
- probabilistic methods 4:258
- procaspase zymogens 3:364, 3:366
- procentriole 2:649, 2:651, 2:651–653
- processive, defined 1:396
- processive assembly of filaments 2:548
- processive tracking 2:685–686
- processivity 2:597
factor 1:447–448
- procollagen
molecules 2:699
processing 1:634
- 17-producing helper T cell (Th17) 3:721–722
- profilin 2:556–557, 2:557, 2:685
- progenitors 3:6, 3:6–7
- progerin 2:313
- progesterone 1:178–179
- programmed cell death (PCD) 3:2, 3:3, 3:4, 3:67, 3:327, 3:342, 3:364, 4:95
PD-1 3:545
PDCD4 3:245
- programmed death receptor 1:145
- progressive-lowering of temperature protocol (PLT protocol) 2:34
- pro-hepatocyte growth factor (pro-HGF) 1:655
- pro-HGF *see* pro-hepatocyte growth factor (pro-HGF)
- prohibitins (PHB) 2:283
- prohormone convertase 1/3 (PCSK1) 1:555, 1:557, 1:557, 1:564T
calcium dependence 1:558
cleavage of prodomain 1:557–558
dibasic site 1:558F
general properties of 1:557
PC1/3 maturation in the regulated secretory pathway 1:557–558
physiological significance of 1:558–559
- prohormone convertase 2 (PCSK2) 1:559–560, 1:564T
calcium dependence 1:560
general properties of 1:560
physiological significance of PC2 1:561–562
ProPC2 maturation in the regulated secretory pathway 1:560–561
- prokallikrein (proKLK) 1:699
- prokaryote cells 2:588
- prokaryotic mRNA 1:341–342
polycistronic mRNA 1:342
Shine–Dalgarno sequence 1:341–342
- prokaryotic transcription 1:468
bacterial RNA polymerase holoenzyme 1:469–471
core RNA polymerase 1:468–469
mechanisms regulate bacterial transcription 1:475
activation 1:476–477
repression 1:475–476
 σ^{54} 1:477
switching sequence specificity of polymerase 1:477
- process in bacteria
elongation 1:472–474, 1:473F
initiation 1:471–472
pausing 1:474
termination 1:474–475
transcription and regulation in archaea 1:477–478
- proKLK *see* prokallikrein (proKLK)
- prolactin (PRL) 3:23–24
- prolamellar body (PLB) 2:324
- proliferating cell nuclear antigen (PCNA) 1:400, 1:400–401, 1:403, 1:411, 1:447–448, 1:461, 2:83, 3:459
- proliferation 4:111–112
- proline, glutamic acid, serine, threonine domain (PEST domain) 3:107
- proline-rich Akt substrate of 40 kDa (PRAS40) 3:130, 3:243
- proline-rich domain (PRD) 2:491
- proline-rich sequence (PRS) 3:650
- proline-serine-threonine phosphatase interacting protein 1 (PSTPIP1) 2:494–497
- prolyl 4-hydroxylase domain (PHD) 4:109–110
- prometaphase 3:480
- ProMMP-9 1:626
- promoter RNA (pRNA) 3:257–258
- promoter(s) 1:364–365, 1:468, 1:471, 1:472, 1:475, 1:477, 1:478, 1:481–482, 1:488, 3:330, 3:667, 3:786, 3:787, 3:788, 3:789, 3:808–809
- promyelocyte 3:751
- promyelocytic leukemia (PML) 1:431, 3:102
- promyelocytic-leukemia nuclear body (PML-NB) 2:311, 2:826
- proofreading exonuclease 1:396, 1:398, 1:400
- pro-opiomelanocortin (POMC) 2:288–289, 2:377
- prophase 3:479–480
- proplastid 2:324
- proprotein convertase subtilisin/kexin type 9 (PCSK9) 2:428
- proprotein convertases 1:555–556
expression and special features of 1:564T
furin (PCSK3) 1:556
intracellular trafficking 1:556, 1:563
paired basic amino acid cleaving enzyme 4 (PACE4; PCSK6) 1:562
- prohormone convertase 1/3 (PCSK1) 1:557
- prohormone convertase 2 (PCSK2) 1:559–560
- proprotein convertase 4 (PCSK4) 1:562, 1:564T
- proprotein convertase 5/6 (PCSK5) 1:562–563, 1:564T
- proprotein convertase 7 (PCSK7) 1:563, 1:564T
- proteolytic activation 1:555, 1:563
- secretory pathway 1:555, 1:557, 1:562, 1:563, 1:563–564
regulated 1:557–558, 1:560–561
- transmembrane domain 1:555, 1:556, 1:562–563, 1:564T
- prorenin receptor (PRR) 1:747
- proSAAS protein 1:558
- prostate-specific antigen (PSA) 1:699
- Prostheobacter* 2:27, 2:27F
- protease (PR) 2:828–829
- protease activated receptors (PARs) 1:702
PAR-2 1:654

- protease activated receptors (PARs) (*continued*)
- protease cleavage site specificity (PICS) 1:576–577, 1:578F
- protease web 1:569, 1:570, 1:579, 1:622–624, 1:694
- protease-activated receptor1 (PAR1) 1:624
- proteases 1:563, 1:569, 1:570, 1:578, 1:578–579, 1:579, 1:582, 1:586, 1:650, 1:699, 1:706, 3:329
- classification 1:569–570
- identification of substrates 1:570–572
- proteasome(s) 1:88, 1:605–606, 1:711, 1:757, 2:25–26, 2:438–439, 2:439, 3:589
- 26S proteasome 1:543, 1:582, 1:583F, 1:586, 1:588, 1:589F, 1:590F, 3:589
- 19S regulatory particles (19S RPs) 1:589–590
- 20S core particle (20S CP) 1:588–589
- alternative regulatory particles and activators 1:590
- assembly of proteasome 1:590–591
- coordinated transcriptional control of gene expression 1:591
- immunoproteasome and alternative CP subunits 1:591
- inhibitors 1:591–592, 1:592
- protection of telomeres protein 1 (POT1) 1:420, 1:427
- protective antigen (PA) 2:795
- proteins, isolation/purification of 1:66
- assays 1:68
- chromatography 1:70–71
- adsorption chromatography 1:72
- affinity chromatography 1:72
- column 1:71
- IEX chromatography 1:71
- SEC 1:71–72
- considerations
- accessibility 1:69
- accountability 1:68–69
- repeatability 1:69
- stabilization 1:69
- initial steps
- batch adsorption 1:70
- centrifugation 1:70
- concentration 1:70
- differential precipitation 1:70
- migration in electric field 1:73
- protein crystallization 1:73
- protein isolation procedure 1:67
- protein source 1:67–68
- purity evaluation 1:73
- protein associated with Lin Seven 1 (PALS1) 2:742–743
- protein biosynthetic machinery of mitochondria 1:545–554, 1:545
- components of translational system 1:545–546
- mitochondrial mRNAs 1:545–546
- mitochondrial ribosomes 1:548–549
- mitochondrial tRNAs 1:546
- future directions 1:552–553
- mechanism of 1:549
- initiation 1:549–550
- polypeptide chain elongation 1:550–551
- termination and ribosome recycling 1:551–552
- mitochondrial DNA 1:545
- protein complementation assay (PCA) 4:170, 4:189
- Protein Data Bank (PDB) 1:94, 1:299–300
- protein degradation 1:324–325, 1:582, 3:590–591
- protein disulphide isomerase (PDI) 1:123, 1:535, 1:540, 2:161–162, 2:169, 2:341, 2:529
- protein domains 1:91, 3:112
- functions 1:92
- DNA-binding domains 1:92–93
- enzymatic and catalytic domains 1:93
- IDDs 1:93
- interaction and binding domains 1:92
- identifying and classifying 1:93
- bioinformatic platforms 1:94–95
- determination by protein crystallography and NMR spectroscopy 1:93–94
- mapping 2:91–92
- modularity and combinatorial protein structure 1:95–96
- protein structure 1:91–92
- protein engineering 1:124, 1:318
- altered function 1:124
- site-selective conjugation 1:125–126
- stability 1:124–125
- oxidative 1:125
- pH 1:125
- thermal 1:125
- protein family 1:396
- protein fold 1:396
- protein folding and assembly in ER 1:538–540
- protein folding diseases 1:115–121
- cell specificity and susceptibility 1:118
- cellular transfer of aggregates on toxic fragments 1:117–118
- clearance pathways of aggregates
- autophagy pathway 1:117
- cellular UPS 1:116–117
- effects of protein products on protein homeostasis 1:116
- genetic basis of HD and ALS 1:116
- potential treatments and therapies 1:119
- symptoms and treatments of disease 1:118–119
- protein kinase A (PKA) 2:372, 2:795, 3:76–77, 3:102–103, 3:573
- protein kinase B (AKT) 3:815
- protein kinase B (PKB) *see* Akt
- protein kinase C (PKC) 1:675, 1:676, 2:317, 2:372, 2:767, 3:2, 3:13, 3:132, 3:153, 3:574–575, 3:681–682, 3:705, 4:105–106
- and casein kinase 2 substrate in neurons (PACSIN) 2:493
- PKC- δ 4:77–78
- PKC β 95 1:765–766
- protein kinase cascades 4:44–50
- amplitude 4:47–48, 4:47F
- bistability 4:46
- dynamics of signalling, cellular decisions encoded in 4:46–48
- cell-type specificity 4:46
- feedback system 4:46
- integrated response 4:47–48, 4:47F
- signaling specificity 4:46
- robustness of signalling pathways 4:48–49
- gene expression noise 4:48
- noise suppression 4:48
- signal amplification and decision making 4:44–46
- signal duration 4:46, 4:46–47, 4:47, 4:47F
- protein kinase C ϵ (PKC ϵ) 3:347–348
- protein kinase D (PKD) 2:367
- protein kinase R-activating protein (PACT) 1:354, 1:355–358, 1:356–357F, 1:357F, 1:358, 1:360
- protein kinases (PKAs) 1:253–254
- protein O-fucosyltransferase 2 (POFUT2) 1:633
- protein O-mannosyltransferase 1 (POMT1) 2:349
- protein O-mannosyltransferase 2 (POMT2) 2:349
- protein p35 from the cowpox virus 3:332
- protein phosphatase (PP) 2:663
- protein phosphatase 1 (PP1) 1:290, 1:291–292, 1:292, 2:531–532, 2:584
- Protein phosphatase 2A (PP2A) 3:100, 3:132, 3:343
- protein phosphorylation 1:528F
- protein quality control 1:596, 1:606
- protein quantitative trait loci analysis (pQTL analysis) 4:171
- protein regulator of cytokinesis-1 (PRC1) 3:482–483
- protein sequence determination by chemical techniques
- challenges and limitations 1:77–78
- early studies 1:75
- larger proteins 1:75–76
- separations methodology role 1:76–77
- sequence comparisons 1:80–81
- species variants and of functional variants 1:81
- sequencing methodology
- inferring protein sequences from nucleic acid sequences 1:80
- mass spectrometry 1:78–80
- protein synthesis 1:317–318, 1:322, 1:384, 1:384, 1:388–390, 1:392–393, 3:610
- 80S initiation complex 1:528F, 1:530F
- A, P, and E sites 1:525, 1:531
- bacterial 1:384, 1:385F
- beyond first ribosome 1:531
- components of 1:525–526
- mRNA 1:526
- ribosomes 1:526–527
- tRNA 1:525–526
- elongation cycle 1:525, 1:530–531
- functional role of 16S and 23S rRNAs in 1:388
- GTP hydrolysis 1:388
- interaction of ribosomal RNAs with E-tRNA 1:390

- interactions of EF-G with the ribosomal RNAs 1:390
- interactions of EF-Tu-GTP-aminoacyl tRNA with the ribosomal RNA 1:322, 1:324–325, 1:388, 1:525–534, 1:525
- peptide bond formation 1:388–390
- elongation factors (EF) 1:525, 1:526
- eukaryotic translation factors in 1:527–529
- cap-dependent protein synthesis 1:527–529
- initiation events 1:529–530
- in fed state 1:294
- functional role of 5S rRNA 1:393
- global regulation of 1:532–533
- initiation, functional role of 16S and 23S rRNAs in 1:384–385
- interactions with initiation factors 1:385–387
- interactions with initiator tRNA 1:387–388
- interactions with mRNA 1:384–385
- subunit association and rRNA 1:388
- initiation factors (IF) 1:525, 1:526, 1:528*F*, 1:529
- initiation of 1:525
- mistakes in 1:532–533
- molecular mimicry 1:531–532
- mRNA specific regulation 1:533
- recycling, functional role of 16S and 23S rRNAs in 1:392–393
- interactions with ribosome recycling factor 1:392–393
- regulation of protein expression 1:533
- 3' UTR 1:533
- 5' UTR 1:533
- ribosome recycling 1:531
- termination, functional role of 16S and 23S rRNAs in 1:390–392
- interactions with release factor 3 1:392
- interactions with release factors 1 and 2 1:390–392
- termination factors 1:384
- termination of 1:525, 1:531
- Protein Topography and Migration Analysis Platform (PROTOMAP) 1:571–572, 1:572
- protein tyrosine kinase (PTK) 3:650
- protein tyrosine phosphatases (PTPs) 2:219–220, 2:497, 3:337
- protein(s) 4:187, 4:211, 4:216
- aggregation 1:115, 1:116
- architecture and function 3:112
- biochemistry 1:484–485
- Ca²⁺ regulation by
- Ca²⁺ channels 3:163–164
- Ca²⁺ transporters 3:164–165
- mitochondrion and Ca²⁺ transporting system 3:166*F*
- surface representation of Ca²⁺-bound CaM 3:163*F*
- transporters and channels 3:164*F*
- crystallization 1:73
- ectodomain shedding 1:639, 1:639*F*
- fidelity of secretome, 2:157
- folding 1:103, 1:105, 2:161–163
- exotoxins blocking 2:797
- homeostasis, protein product effects on 1:116
- hormones 2:379–380
- interactions
- and activity measurement 4:221–222
- DR signaling 3:68–69
- isolation procedure 1:67
- kinases 1:84–85
- maturation 2:176
- phosphorylation 4:177–178, 4:227
- prenylation 1:87
- protein C pathway 1:717
- protein-based sorting 2:365–366
- protein–DNA interaction 1:93*F*
- protein–protein complexes 1:98, 1:101
- proteins mediating exocytosis 2:296
- Rab3A 2:296
- SNAP-25 2:296
- SV2 2:296
- synaptobrevin 2:296
- synaptotagmin 2:296
- syntaxin-1 2:296
- related diseases 1:113–114
- secretion 2:158
- source 1:67–68
- structure 1:91–92
- synthesis 2:329
- exotoxins blocking 2:797
- traffic 4:157
- translocation 2:447
- translocation measurement 4:221
- proteinaceous amino acids
- see* proteinogenic amino acids
- proteinase 1:621–622
- proteinase 3 (PR3) 3:751
- protein-binding microarrays (PBM) 4:191
- protein–DNA interaction 4:191, 4:211–212, 4:212, 4:213–214, 4:213, 4:214, 4:216
- protein–glycan interactions (PGI) 4:192
- protein–lipid interactions (PLIs) 4:192
- proteinogenic amino acids 2:329
- protein–protein interactions (PPIs) 1:91, 1:92, 1:93, 3:664–665, 4:188, 4:211–212, 4:212, 4:212–213, 4:213, 4:213–214, 4:214, 4:214–215, 4:216
- MAPPIT 4:189–190
- MIMIX 4:190
- proteome-wide binary interactome mapping 4:190
- VAMP7 network 2:463*F*
- protein–RNA interactions (PRIs) 4:191–192
- mapping 4:191–192
- protein–small molecule interactions (PSMIs) 4:192–193
- proteogenomics 4:163, 4:164
- proteoglycan(s) 1:271, 1:279–280, 1:280–281, 1:282, 1:286*F*, 1:631, 1:633, 1:636, 2:351–352, 2:351–352
- cellular functions 1:274
- aggrecan family of pericellular proteoglycans 1:276
- cell surface proteoglycans 1:275
- glypicans 1:275
- HSPG 1:275
- post-synthesis processing of GAGs 1:276
- serglycin 1:274–275
- Shed syndecan-1 1:276
- SLRP family of secreted proteoglycans 1:276
- structure and synthesis
- core protein structure 1:271–274
- GAG 1:274
- partial list of proteoglycans and sulfated GAG chains 1:272*T*
- proteolysis 1:555, 3:329
- proteolytic signature peptides (PSP) 1:569, 1:578–579
- proteolytic/proteolysis 1:633, 1:635, 1:670, 1:673
- enzymes 1:681
- large intestine role in 1:735–736
- oral cavity role in 1:732–733
- processing 1:644–646
- PTM of proteins 1:84
- small intestine role in 1:733–735
- stomach role in 1:733
- proteome 1:569, 1:572–573, 1:573, 1:574–575, 1:576, 1:576–577, 1:577, 1:577*F*, 1:578*F*
- proteome of centriole 5 (hPOC5) 2:651
- proteome-wide binary interactome mapping 4:190
- proteomic identification of cleavage site specificity (PICS) 1:694
- proteomic/proteomics 4:163, 4:163–164, 4:166, 4:227–228, 4:265–266
- approach 3:302
- studies 1:198–199
- cleavage specificity and identification of substrates by 1:694
- on nucleolar architecture 3:264–265
- Proteomics Standards Initiative–Molecular Interactions (PSI-MI) 4:245
- proteostasis 1:115–116, 1:790, 1:793–794
- proteotoxicity 1:789, 1:793, 1:796
- Prothrombin G20210A 1:719
- Proto-Dbl 2:731–733
- PROTOMAP *see* Protein Topography and Migration Analysis Platform (PROTOMAP)
- proton translocation 1:301–302, 1:304, 1:305
- protrusion 2:683, 2:720
- actin polymerization-driven 2:683–686
- molecular machinery for filopodial 2:687–688, 2:687*F*
- pro-tumor 1:753, 1:755
- PrP *see* prion protein (PrP)
- Prp8 1:497
- PrPC *see* cellular form of prion protein (PrPC)
- PRR *see* pattern recognition receptors (PRR); prorenin receptor (PRR)
- PRS *see* positive reference set (PRS); proline-rich sequence (PRS)
- PRSS21 *see* testisin
- PRSS31 *see* tryptase γ 1
- PS receptor (PSR) 3:375

PS *see* P-surface (PS); phosphatidyl serine (PS); phosphatidylserine (PS)
 PSA *see* prostate-specific antigen (PSA)
 PSA-NCAM *see* polysialic acid-neuronal cell adhesion molecule (PSA-NCAM)
 PSCK3 *see* Furin (PCSK3)
 PSD *see* postsynaptic density (PSD)
 PSD95, Dlg1 and zo-1 domain (PDZ domain) 2:197–198, 2:597
 PSD95 protein 2:597
 P-selectin glycoprotein ligand 1 (PSGL-1) 1:663, 2:344–345, 3:543, 3:624–625, 3:751
 pseudocontact shifts (PCs) 1:100
 pseudo-Hurler polydystrophy 1:773
 pseudoknot domain 1:422
Pseudomonas aeruginosa infection 1:255–256
 pseudopod 2:752, 4:4, 4:6
 PSF *see* point spread function (PSF)
 PSGL-1 *see* P-selectin glycoprotein ligand 1 (PSGL-1)
 PSI-MI *see* Proteomics Standards Initiative–Molecular Interactions (PSI-MI)
 PSM *see* presomitic mesoderm (PSM)
 pSMAC *see* peripheral supra-molecular activation cluster (pSMAC); peripheral SMAC (pSMAC)
 PSMIs *see* protein-small molecule interactions (PSMIs)
 PSP *see* proteolytic signature peptides (PSP)
 PSR *see* PS receptor (PSR)
 PSTPIP1 *see* proline-serine-threonine phosphatase interacting protein 1 (PSTPIP1)
 P-surface (PS) 2:39
 psychosine 1:184
 PTAs *see* peripheral tissue antigens (PTAs)
 PTB domain *see* phosphotyrosine binding domain (PTB domain)
 PTC *see* premature stop codon (PTC)
 PTCH *see* Patched (PTCH)
 PTEN *see* phosphatase and tensin homolog (PTEN)
 PtdIns 3P *see* phosphatidyl inositol 3-phosphate (PtdIns 3P)
 PtdIns *see* phosphatidylinositol (PtdIns)
 PtdIns3 kinase class 3 (PtdIns3KC3) 2:245
 PtdIns3P *see* phosphoinositol-3-phosphate (PtdIns3P); PtdIns-3-phosphate (PtdIns3P)
 PtdIns3P-binding proteins 2:245
 PtdIns-3-phosphate (PtdIns3P) 3:128
 PtdIns4P *see* PtdIns-4-phosphate (PtdIns4P)
 PtdIns-4-phosphate (PtdIns4P) 3:128
 P-TEF-b *see* positive-transcriptional elongation factor b (P-TEF-b)
 PTEN hamartomatous syndromes (PHS) 3:419
 PTEN homolog *see* phosphatase and tensin homolog (PTEN homolog)
 PTHR *see* parathyroid hormone receptor (PTHR)
 PTK *see* protein tyrosine kinase (PTK)
 PTLs *see* Photoswitchable tethered ligands (PTLs)

PTMs *see* posttranslational modifications (PTMs)
 PTPC *see* permeability transition pore complex (PTPC)
 PTPN14 *see* tyrosine-protein phosphatase non-receptor type 14 (PTPN14)
 PTPs *see* protein tyrosine phosphatases (PTPs)
 PTS *see* phospho-transferase system (PTS)
 PTX *see* pertussis toxin (PTX)
 pullanase 1:125
 pulling forces 3:664–665
 pulmonary embolism (PE) 1:718–719
 pulsed Q-dissociation (PQD) 4:182
 puncta 2:744–745
 punctin-1 1:633
 purified COPI-coated vesicles 2:356
 purine catabolism 2:254
 PV *see* polio-virus (PV)
 PWD *see* phosphoglucan water dikinase (PWD)
 PX domain *see* phox homolog domain (PX domain)
 PX-BAR domains *see* phoxdomain-containing domains (PX-BAR domains)
 PX-containing protein SNX16 2:203
 PYD *see* pyrin domain (PYD)
 pyrin and HIN domain (PYHIN) 3:719–720
 PYHIN1 3:719–720
 pyrin domain (PYD) 3:714
 pyrin domain only protein 3 (POP3) 3:719–720
 pyroptosis 2:772–773, 2:772F, 3:327, 3:370–371, 3:370–371, 3:371
 pyrolysine (Pyl) 1:313, 1:526
 pyruvate dehydrogenase (PDH) 3:336, 3:814
 pyruvate dehydrogenase complex (PDC) 1:291
 regulation of 1:293–294
 pyruvate dehydrogenase kinases (PDKs) 1:296, 3:336, 3:814

Q

QAE *see* quaternary aminoethyl (QAE)
 QbD *see* quality-by-design (QbD)
 QIT-TOF *see* quadrupole iontraptime-of-flight (QIT-TOF)
 QSAR *see* quantitative structure activity relationships (QSAR)
 QTL analysis *see* quantitative trait linkage analysis (QTL analysis)
 quadratic programming 4:212T
 quadrupole iontraptime-of-flight (QIT-TOF) 1:155
 quality assessment and benchmarking 4:248
 quality control 2:156–157, 2:163, 2:164, 3:404, 3:408
 mechanisms for ribosome assembly 3:303
 quality control code 1:599, 1:601
 quality-by-design (QbD) 1:136
 quantitative phosphoproteomics 4:182

quantitative structure activity relationships (QSAR) 1:138
 quantitative trait linkage analysis (QTL analysis) 3:734
 quaternary aminoethyl (QAE) 1:71
 quenching 4:202, 4:202, 4:203T
 quiescence 3:433, 3:435, 3:436
 cyclin-CDK in 3:425–428
 quorum sensing 2:302–303

R

RA *see* rheumatoid arthritis (RA)
 RA domains *see* Ras association domains (RA domains)
 RAAS *see* renin-angiotensin-aldosterone system (RAAS)
 Rab GTPase-activating protein (Rab-GAP) 1:765, 2:285
 Rab GTPases 2:195, 2:412–413, 2:485, 2:616, 2:788–789
 intracellular pathogens subvert function 2:405–406
 Rab proteins 1:764, 2:195, 2:355, 2:367–368, 2:442
see also GTP-binding proteins (G proteins)
 biosynthesis 2:442–443
 coordination 2:195–196
 coupling proteins 2:754
 effectors 2:445
 geranylgeranyl transferase 2:442–443
 membrane targeting 2:442–443
 proteins 2:444
 packaging cargo proteins into transport vesicles 2:444
 tethering and docking transport vesicles 2:445
 vesicle movement on actin and microtubule cytoskeleton 2:444–445
 Rabs localizing to distinct organelles 2:442F
 structure and GTPase 2:443–444, 2:443F
 Rab1 2:170–171
 Rab1 by adenylation (AMPylation) 2:779
 Rab1 GTPase 2:779
 Rab2 2:170–171
 Rab3A 2:296
 Rab4 2:197, 2:402–403, 2:504
 Rab5 1:617, 1:765, 2:435, 2:504, 2:508, 2:760
 RAB7 1:765
 Rab8 2:404–405
 RAB9 2:204–205
 Rab11 2:197, 2:504
 binds motor proteins 2:403
 and cancer 2:404
 in cell division 2:403–404
 effector proteins 2:403
 and endosomal recycling compartment 2:403
 in synaptic transmission 2:404
 Rab11A 2:403
 Rab11B 2:403
 Rab11C 2:403, 2:404

- Rab11-Family of Interacting Proteins (Rab11-FIPs) 2:403
- Rab11-FIPs *see* Rab11-Family of Interacting Proteins (Rab11-FIPs)
- Rab14 2:404–405
- Rab17 2:404–405
- Rab22A 2:404–405
- Rab25 *see* Rab11C
- Rab27a 2:226
- Rab32 2:241
- Rab38 2:241
- RABA1 GTPases 2:204
- rabenosyn5 2:504
- Rabex 5 2:508
- rabex-5-rabaptin-5 exchange factor complex 2:196
- Rab-GAP *see* Rab GTPase-activating protein (Rab-GAP)
- rab-GDI *see* rab-GDP dissociation Inhibitor (rab-GDI)
- rab-GDP dissociation Inhibitor (rab-GDI) 2:195
- rabies virus 1:520
- Rab-interacting lysosomal protein (RILP) 2:410
- Rac 2:731, 2:759–760, 3:211, 3:704
- Rad4-Rad23 1:437F
- Rad51, filament assembly and stability 1:457–460
- RAD51AP1 *see* RAD51-Associated Protein 1 (RAD51AP1)
- RAD51-Associated Protein 1 (RAD51AP1) 1:460
- Rad54 1:460
- radioactive thymidine ((3H)-Thy) 4:135
- radial astrocytes (RAs) 3:799
- radial migration 2:728
- radial spokes (RS) 2:627, 2:630–631, 2:660
- RSP3 2:661–663
- radio sensitive severe combined immunodeficiency (RS-SCID) 3:559–560
- radioimmunoassay (RIA) 1:141–142
- radiolabelled antibodies 1:143
- rafts detection
- by microscopy and spectroscopy 1:213
- in model membranes 1:211–212
- in natural membranes 1:212
- using detergents 1:212–213
- using sterol modification 1:213–214
- Rag GTPases 1:765–766, 1:766, 1:766–767
- RAGs *see* regeneration-associated genes (RAGs)
- RAIDD *see* RIP-associated protein with a death domain (RAIDD)
- RAM *see* RBPJk-associated module (RAM)
- Raman spectroscopy 1:128
- Ran 3:177
- Ran gradient 2:320–321
- Ran in GDP-bound form (RanGDP) 2:320–321
- Ran in GTP-bound form (RanGTP) 2:320–321
- gradient 2:315–317
- random reference set (RRS) 4:188
- random walk-based approaches 4:212, 4:213
- RanGDP *see* Ran in GDP-bound form (RanGDP)
- RanGTP *see* Ran in GTP-bound form (RanGTP)
- RANTES *see* regulated on activation, normal T cell expressed and secreted (RANTES)
- RAP *see* receptor-associated protein (RAP)
- rapamycin 3:128
- effect on mTORC1 3:244–245
- rapid mixing-partitioning model 2:186
- rapid partitioning model 2:357F, 2:360
- Raptor 3:243
- RAR-related orphan receptor (ROR) 3:567
- Ras 2:759
- effector 2:445–446
- pathway 3:413
- superfamily 2:445–446
- Ras association domains (RA domains) 3:153
- RAS *see* renin-angiotensin system (RAS)
- RAS *see* radial astrocytes (RAs)
- rasiRNA *see* repeat associated siRNAs (rasiRNA)
- Ras-like proteins 4:51
- rat rhodopsin 4 (RO4) 2:147–148
- ratchet-like motion 1:388–390, 1:392, 1:392–393
- rate constant 1:7–8, 1:8
- rate-limiting steps, in metabolism 4:28, 4:29
- rational drug design 1:137, 1:138
- RavZ protein 2:781–782
- Rayleigh criterion 2:62, 2:70
- Rb *see* retinoblastoma (Rb)
- Rb protein *see* retinoblastoma protein (Rb protein)
- RB *see* reticular body (RB)
- RBD *see* RNA-binding domains (RBD)
- RBPJk-associated module (RAM) 3:107
- RBS *see* ribosomal binding site (RBS)
- RC *see* reaction center (RC)
- RCC *see* renal cell carcinoma (RCC)
- RCL *see* reactive center loop (RCL)
- RCSA *see* residual chemical shift anisotropy (RCSA)
- RD *see* regulatory domain (RD)
- RDCs *see* residual dipolar couplings (RDCs)
- rDNA *see* ribosomal DNA (rDNA)
- RdRp *see* RNA-dependent RNA polymerase (RdRp)
- reaction center (RC) 2:328
- reaction rate 1:7–8, 4:26, 4:26–27, 4:26, 4:26F, 4:27, 4:31, 4:32
- reaction rate-limiting step 1:8
- reaction-transport equation 4:118–119
- reactive center loop (RCL) 1:710
- reactive oxygen species (ROS) 1:87, 2:4, 2:283, 2:767, 3:6, 3:13, 3:218, 3:284, 3:336, 3:341–342, 3:404–405, 3:460, 3:508–509, 3:702, 3:710–712, 3:752, 3:777, 3:815–816
- as cell survival factors 3:337–338
- control 3:338F
- as mediator 3:336–337
- Reactome 4:211
- read depth 1:32–33
- really interesting new gene (RING) 1:585, 1:586, 1:708, 3:390
- domains 3:390–391
- RecA fold 1:396
- receptor coupling, signal amplification through 4:8
- receptor mediated phagocytosis 3:746
- receptor-mediated clathrin-dependent endocytosis 1:617
- receptor substrate-1 (IRS-1) 3:32
- receptor TNFRSF-interacting serine-threonine kinase 2 (RIP2) 3:391
- receptor tyrosine kinases (RTKs) 2:211, 2:469, 3:1–2, 3:8–21, 3:51, 3:208–209, 3:227, 3:374, 3:413, 3:733, 4:178–179
- coreceptors 3:13
- dimerization activating kinase activity 3:11
- in disease 3:17
- diabetes 3:17
- inhibitors in treatment 3:17
- malignancies 3:17
- families and ligands 3:8–9, 3:10T
- internalization 3:13–14
- ligand-induced dimerization 3:9–11
- RTK families 3:9F
- signaling
- downstream 3:11–13
- feedback and amplification 3:13
- inducing by 3:12F
- structural and functional properties
- EGF receptor family 3:14
- Eph family 3:16
- FGF receptor family 3:15
- insulin receptor family 3:14
- Met and Ron 3:15–16
- PDGF receptor family 3:14–15
- Tie family 3:16
- Trk family 3:16
- VEGF receptor family 3:15
- receptor-associated protein (RAP) 2:420–421
- receptor-interacting protein 1 (RIP1) 3:62, 3:67
- receptor-interacting protein kinase 1 and 3 (RIPK1 and 3) 3:371–372
- receptor-interacting protein kinase 3 (RIPK3) 3:342
- receptor-interacting serine-threonine kinase 2 (RIPK2) 3:716
- receptor-mediated endocytosis (RME) 2:388–389
- receptor-operated channels (ROCs) 3:163–164
- receptor-regulated SMADs (R-SMADs) 3:43
- receptors for Fc portion of immunoglobulins (FcR) 3:528
- recombinant DNA technology 1:66
- recombination signal sequences (RSS) 3:550
- recursive porous agent simulation toolkit (RePAST) 4:82
- recycling 3:406–407
- endosomes 2:193, 2:364–365

- recycling (*continued*)
- red fluorescent protein (RFP) 4:221–222
- red pulp (RP) 3:635–636
- red pulp macrophages (RPMs) 3:763–765
- redox homeostasis 3:815–816
- redox proteomics 1:87
- reepithelialization 3:219
- reference space 2:40
- refractive index mismatch 2:90
- Refsum disease 2:269
- regeneration
 - injury changes axon structure and axonal transport 3:235–236
 - wound healing vs. 3:220–221
- regeneration-associated genes (RAGs) 3:235–236
 - RAG-mediated V(D)J recombination, molecular mechanisms of cleavage 3:557–558, 3:558F
 - DNA binding and cleavage 3:557
 - joining broken DNA ends 3:558–560
 - RAG proteins 3:557
 - RAG recombinase 3:560–562
 - recognition signal sequences 3:559F
- recombinase 3:560–562
- RegMod 4:212T, 4:213
- regression 4:212, 4:213
 - analysis 4:257–258
- regulate receptor tyrosine kinase signaling pathways (RTK signaling pathways) 3:144
- regulated intramembrane proteolysis 1:643–644
- regulated Ire1-dependent decay 2:165
- regulated on activation, normal T cell expressed and secreted (RANTES) 3:636–637
- regulated proteolysis 1:325–326, 1:555
 - proprotein convertases *see* proprotein convertases
- regulated secretion 2:376, 2:377–379, 2:378F
 - diabetes and 2:379
 - molecular mechanism of synaptic fusion 2:382F
- regulated exocytosis 2:379
 - making regulated secretory compartment 2:379–380
 - storage inside cell 2:380–381
- responding to signal 2:381
- secretory process 2:381–382
- and synaptic transmission 2:381
- understanding 2:377
- regulation of elongation 3:275–276
- regulation of rDNA expression 3:258–259
- regulation of transport 3:175–176
- regulator of G protein signaling domain (RGS domain) 2:731–733
- regulatory coefficient 4:27
- regulatory components 2:715–716
- regulatory domain (RD) 3:129–130
- regulatory light chain (RLC) 2:597
- regulatory network 4:253
- regulatory T cells (Treg cells) 3:576–577, 3:593, 3:606–607, 3:606, 3:609, 3:619, 3:634–635, 3:691, 3:747
- Rel homology domain (RHD) 3:580
- rela* KO mice 3:582
- relative membrane permeability 1:218, 1:223–225, 1:224F, 1:225
 - caveats in measuring physiological membrane potential 1:232–234
 - membrane potential measurement and 1:230–231
 - membrane selectivity and direct measurements of 1:230–231
 - physiological values of membrane potential 1:234–235
 - quantifying membrane and ion channel selectivity 1:231–232, 1:232F
- relative quantification using dimethylation 1:574
- relaxation dispersion spectroscopy 1:104
- relb* KO mice 3:582
- renal cell carcinoma (RCC) 3:64
- renin–angiotensin–aldosterone system (RAAS) 3:58
- renin–angiotensin system (RAS) 1:746–747
 - complex network 1:747F
 - overview of system 1:746–747
 - proteases
 - ACE 1:748–749
 - ACE2 1:749–750
 - renin 1:747–748
- Reoviridae* 1:518T, 1:519
- REP *see* Rab escort protein (REP)
- RePAST *see* Recursive Porous Agent Simulation Toolkit (RePAST)
- repeat associated siRNAs (rasiRNA) 1:370
- repeatability 1:69
- replication factor C (RFC) 1:461, 3:459
- replication factories 3:459–460
- replication fork 1:396, 1:397, 1:398T, 1:408, 1:412, 1:460, 1:463, 1:464
 - prokaryotic helicase at 1:403–404
- replication fork barriers (RFBs) 3:257–258
- replication of DNA *see* DNA replication
- replication protein A (RPA) 1:438–439, 2:83, 3:459
- replication vacuole
 - endocytic pathway manipulation to support bacterial replication 2:777
 - multi-organelle interactions 2:780
 - secretory pathway manipulation to support bacterial replication 2:778–780
- replicative helicases *see* helicases
- replisome components, bacterial and eukaryotic 1:396–417
 - clamp loader in last universal common ancestor 1:401–403
 - mechanism 1:401–403
 - structure analysis 1:403
 - comparison of 1:408–410
 - DNA polymerases 1:398–400
 - crystal structure of B- and C-family 1:400
 - functional overview 1:398–400
 - future need 1:411–412
 - helicases 1:403–404
 - bacterial 1:404
 - eukaryotic 1:404–405
 - proposed strand exclusion mechanism 1:403–404
- primases 1:405
- primpol 1:406–407
- single-strand DNA-binding protein 1:407
 - bacterial SSB 1:407
 - eukaryotic RPA 1:407–408
- sliding clamps conservation 1:400–401
- replisome progression complex (RPC) 1:410
- repolarisation 1:218, 1:231, 1:248
- repression 1:475–476, 4:65–66
 - see also* cross-repression
- repressor sequence 1:506
- repulsive guidance molecule (RGM) 3:42–43
- repulsive interactions 3:193
- RER *see* rough endoplasmic reticulum (RER)
- rescue 2:539, 2:541, 2:544
- residual chemical shift anisotropy (RCSA) 1:104
- residual dipolar couplings (RDCs) 1:98
- resistance-resistant inhibitors 1:744
- resolution 1:461, 2:62, 2:70
- resorptive lacuna 1:776
- respiratory syncytial virus (RSV) 1:145, 3:716
- response pathways to intracellular pathogen infections 2:850
- ResponseNet 4:212T, 4:214–215, 4:215, 4:216
- resting cells
 - cytosolic components structure in 3:703–704
 - flavocytochrome b_{558} in 3:703, 3:703F
 - p47phox–p67phox–p40phox complex in 3:704–705
- resting membrane potential 1:218, 1:225–226, 1:226, 1:229, 1:230–231, 1:231, 1:232–234, 1:232F, 1:234T, 1:244, 1:245, 1:246–247
 - electrochemical equilibrium and 1:224F
 - steady state vs. electrochemical equilibrium 1:225
- restriction factors 2:828, 2:836
- restriction point 3:432–436, 3:432
 - G1 phase of cell cycle 3:433F
- growth factors 3:432–433, 3:433–434
- HGF 3:434
- MAPK 3:434
- RB and p53 in establishment 3:435–436
- RNAi 3:434
- retention 3:625, 3:625–626, 3:626
- reticular body (RB) 2:786–787
- reticulate body 2:780
- reticulons 1:205, 1:206F, 2:822
- retinal pigment epithelial cells (RPE cells) 2:239–241
 - RPE1 cells 2:205
- retinitis pigmentosa (RP) 1:501, 1:507–508
- splicing mutations in 1:508
- retinoblastoma (Rb) 3:133, 3:777
- retinoblastoma protein (pRb) 3:5, 3:425–426, 3:458
- RB1 3:447
- retinoic acid 3:595

- retinoic acid-inducible gene I (RIG-I) 3:394, 3:520, 3:718–719
- retinoic acid-inducible gene I-like receptors (RLRs) 3:710
- retinoic acid like receptors (RLRs) 3:520
- retrieval
- CIMPR 2:480
 - endosome-to-Golgi 2:476, 2:482
 - to Golgi 2:475
 - of Vps10 protein 2:475
- retrograde
- IFT motor 2:666
 - trafficking from endosomes to Golgi 2:513–515
- retrograde transport 2:168, 2:171–172, 2:433–441, 3:235, 3:236
- biomedical applications of 2:438
 - biomedical imaging 2:438
 - immunotherapy 2:438–439
 - targeted tumor therapy 2:438
- biophysical mechanisms for 2:437–438
- membrane bending 2:437–438
 - scission 2:438
 - SNX-driven mechanism 2:437–438
- pathways 2:434–435
- early-endosomes-to TGN trafficking 2:436
 - late-endosomes-to TGN trafficking 2:435
 - recycling endosomes-to TGN trafficking 2:436
- retromer 2:196, 2:197, 2:198, 2:504
- retromer cargo-selective complex 2:479, 2:481F
- localization 2:478F
- retromer complex 2:435, 2:475–484
- accessory proteins 2:478–481
 - assembly and structure 2:475–476
 - and disease 2:482–483
 - history 2:475
 - key elements 2:477F
 - membrane tubules stabilization 2:476–477
 - physiological processes 2:481–482
 - recognition of cargo 2:478
 - cargo proteins 2:479T
 - recruitment 2:477–478
 - and action 2:480F
 - retromer cargo-selective complex 2:481F
 - localization 2:478F
- retropepsin 1:681
- retrotranslocon 1:535, 1:543, 1:601
- retrotransposon 1:418, 1:423, 1:523
- retroviral protease (PR) 1:739
- retroviral proteases inhibitors 1:739–741
- retroviruses 1:739, 2:828
- Rev responsive element (RRE) 2:835
- reversal potential 1:231–232, 1:233F
- reverse cholesterol transport 1:176, 1:176–177, 1:178, 2:427
- reverse engineering 4:255
- reverse phase chromatography *see* size-exclusion chromatography (SEC)
- reverse transcriptase (RT) 1:29, 1:739, 2:828, 1:418, 1:497, 1:380, 1:521, 1:523
- see also* RNA-dependent DNA polymerase
- reverse transcription 1:397–398, 2:833–835
- reverse transcription complex (RTC) 2:833
- reverse transcription polymerase chain reaction (RT-PCR) 3:759
- reversed phase chromatography (RPC) 1:72
- reversed-phase liquid chromatography (RP-LC) 4:179–180
- rFBA (flux-balance analysis) 4:212T, 4:214
- RFBs *see* replication fork barriers (RFBs)
- RFC *see* replication factor C (RFC)
- RFP *see* red fluorescent protein (RFP)
- RGD sequence *see* Arg-Gly-Asp sequence (RGD sequence)
- RGM *see* repulsive guidance molecule (RGM)
- RGS domain *see* regulator of G protein signaling domain (RGS domain)
- RHD *see* Rel homology domain (RHD)
- Rheb 1:766
- mTORC1 with 3:249–250
- rheumatoid arthritis (RA) 2:728–729, 3:360–361, 3:620–621, 3:755–756
- rhizomelic chondrodysplasia punctata 2:269
- Rho GTPase-activating protein domain (Rho-GAP domain)
- Rho GTPase-activating proteins (RhoGAPs) 2:535, 2:602, 2:733, 2:776
- Rho GTPases 2:4, 2:787, 2:796, 3:186
- and cancer cell metabolism 2:738
 - microvesicles 2:738
 - contributions to tumorigenesis stages 2:736F
 - cycle 2:732F
 - DH domain-containing proteins 2:732F
 - features 2:731
 - of GTP-binding/GTP-hydrolytic cycle regulation 2:731–733
 - in mammalian tumors 2:737–738
 - in oncogenic transformation 2:735–737
 - in regulating actin cytoskeleton 2:733–735, 2:733F, 2:734F
 - regulation of migration by 3:210
 - activation 3:210–211, 3:211F
 - signaling downstream 3:211
 - targets and multiple roles in biology 2:735
 - Cdc42-mediated regulation model 2:736F
 - mammalian targets of Cdc42 2:735F
- Rho kinase (ROCK) 3:479–480
- Rho proteins 2:446, 2:731
- Rho utilization site (rut site) 1:475
- RhoA
- activity promotig actomyosin ring assembly 3:495, 3:496F
 - overexpression 2:737
 - pathway 3:494–495
- Rho-associated coiled-coil-containing protein kinase (ROCK) 2:726, 2:734–735, 4:112
- Rho-type GTPase Cdc42 4:51
- rhodopsin family 2:456
- RhoGAPs *see* Rho GTPase-activating proteins (RhoGAPs)
- Rho-Kinase (ROK) 2:753
- RIA *see* radioimmunoassay (RIA)
- RiBi *see* ribosome biogenesis (RiBi)
- ribonuclease 1:22
- ribonuclease (RNase) P 1:310, 1:375, 1:376T, 1:377–380
- RNase III 1:354, 1:355, 1:358, 1:359
- ribonucleic acid (RNA) 1:25–26, 4:160
- evolutionary importance of 1:375
 - primase 1:397
 - replication 1:517, 1:520
 - RNA pol 3:254–255
- ribonucleoprotein (RNP) 1:375, 1:377–380, 1:380, 1:418, 1:420–421, 1:422, 1:428, 1:495, 1:496, 3:233, 3:281
- ribonucleoside triphosphates (rNTPs) 1:472
- ribonucleotide reductase 1:396, 1:397
- ribophagy 3:278
- ribosomal binding site (RBS) 1:379F
- ribosomal DNA (rDNA) 3:254, 3:257–258
- ribosomal pre-rRNA processing 3:297–299, 3:298F
- ribosomal protein-MDM2-p53 pathway 3:284
- MDM2 binding ribosomal proteins 3:285–286
 - p53 activation 3:284–285
 - RP-MDM2-p53 pathway 3:286
- ribosomal proteins (RP) 1:365, 1:548, 3:3, 3:270, 3:283
- HDM2 binding RPs 3:272F
 - non-ribosomal functions 3:283, 3:284T
 - RPL12 3:281
 - RPL30 3:281
 - RPS13 3:281
 - RPS14 3:281
 - rpS6 3:245
 - stress detection by 3:271T
- ribosomal recycling factors (RRF) 1:310, 1:327
- ribosomal RNA (rRNA) 1:23, 1:384, 3:254, 3:270, 3:281, 3:297, 4:162
- during elongation cycle 1:388
 - GTP hydrolysis 1:388
 - interaction of ribosomal RNAs with E-tRNA 1:390
 - interactions of EF-G with the ribosomal RNAs 1:390
 - interactions of EF-Tu-GTP-aminoacyl tRNA with the ribosomal RNA 1:388
 - peptide bond formation 1:388–390
 - during initiation 1:384–385
 - interactions with initiation factors 1:385–387
 - interactions with initiator tRNA 1:387–388
 - interactions with mRNA 1:384–385
 - subunit association and rRNA 1:388
 - during recycling 1:392–393
 - interactions with ribosome recycling factor 1:392–393
- ribosomal Pre-rRNA processing 3:297–299, 3:298F
- 5S 3:294
- 5.8S 3:289, 3:291, 3:294
- 16S 2:799, 2:803F
- 18S 1:526, 3:3, 3:288, 3:289, 3:291

- ribosomal RNA (rRNA) (*continued*)
 23S 3:294
 during termination 1:390–392
 interactions with release factor 3 1:392
 interactions with release factors 1 and 2 1:390–392
 transcription 3:258
- ribosomal RNA processing 3:288
 in *E. coli* 3:293F
 in eukaryotes 3:291–292
 human diseases 3:294–295
 in prokaryotes 3:292–294
 5S rRNA processing 3:294
S. cerevisiae mature ribosomal RNAs
 processing 3:290F
 structure 3:288F
 transcription 3:288–289
 in yeast
 chemical modifications of rRNA 3:291
 cleavage and folding of precursor rRNA 3:289–291
- ribosome(s) 1:22–23, 1:384, 1:384–385, 1:385F, 1:388, 1:388–390, 1:390, 1:390–392, 1:392, 1:392–393, 1:393, 2:157, 3:288, 3:297
 biogenesis pathways 3:3
 busy life 3:275
 interactions of elongation factors with 1:389F
 interactions of initiation factors with 1:387F
 interactions of release factors 1 and 3 with 1:391F
 interactions of ribosome recycling factors with 1:392–393, 1:392F
 mitochondrial 1:548–549, 1:548T
 non-ribosomal functions of RPs 3:283
 in protein synthesis 1:526–527
 ribosomal protein-MDM2-p53 pathway 3:283–284
 ribosomopathies 3:282–283
 as a ribozyme and riboswitch 1:380
 recycling 1:324, 1:525, 1:530F, 1:531, 1:551–552, 1:552
 stress detection 3:276
 mRNA quality control 3:277
 protein quality control 3:276–277
 stress signaling to 3:276F
 cap-dependent translation 3:275
 regulation of elongation 3:275–276
 structure and function 3:281
 ribosome assembly 3:281
 ribosome biogenesis 3:281
 structure of 1:386F
 subunit biogenesis 3:254–257
 in mammalian nucleoli 3:256F
- ribosome biogenesis (RiBi) 3:270, 3:278F, 3:281
 cytoplasmic events in 3:301
 40S subunit 3:302–303, 3:302F
 60S subunit 3:301–302, 3:301F
 factors 3:272
 quality control mechanisms 3:303
 regulating by oncogenes 3:281–282
 tumor suppressors 3:282
- ribosome degradation 3:277
 defective ribosomes 3:277–278
 ribophagy 3:278
- ribosome-nascent chain complex (RNC) 1:597–598
- ribosome quality control complex (RQC complex) 3:276–277, 3:277F
- ribosome releasing/recycling factor (RRF) 1:329, 1:384, 1:385F, 1:392–393, 1:392–393, 1:392F
- ribosome stress 3:270, 3:271F
 HDM2 suppression 3:273F
 ribosome biogenesis 3:278F
 ribosome birth in nucleolus 3:270
 nucleolar signaling 3:273–274
 P53-independent ribosome stress 3:272–273
 stress detection by nucleolus 3:270–272
 stress signaling to RNA polymerase 3:274F
- ribosomopathies 3:282–283
- riboswitches 1:375, 1:379F
 and aptamers 1:375–377
 as cell biological tools 1:380–382
 control of gene expression by 1:377
 gene-regulation strategies employed by 1:379F
glmS ribozyme–riboswitch 1:380
 natural 1:378T
 ribosome as 1:380
- ribozymes 1:19–22, 1:375, 1:375, 1:379F
 artificially selected 1:376T
 as cell biological tools 1:380–382
glmS Ribozyme–Riboswitch 1:380
 hammerhead 1:19–22, 1:22F, 1:381F
in vitro selection of aptamers and 1:375
 as mobile genetic elements 1:380
 natural 1:376T
 in replication of viroids and virusoids 1:380
 ribosome as 1:380
 RNaseP 1:22
 self-splicing introns 1:22
 ubiquitous nature of 1:377–380
- ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) 2:9–11, 2:10F, 2:248
- RIG-1-like receptor (RLR) 1:708–709, 2:255, 3:718–719
- RIG-I *see* retinoic acid-inducible gene I (RIG-I)
- rigor state 2:601
- RILP *see* Rab-interacting lysosomal protein (RILP)
- rim progression model 2:358–359
- RING *see* really interesting new gene (RING)
- RIP1 *see* receptor-interacting protein 1 (RIP1)
- RIP2 *see* receptor TNFRSF-interacting serine-threonine kinase 2 (RIP2)
- RIP-associated protein with a death domain (RAIDD) 3:370
- RIPK2 *see* receptor-interacting serine-threonine kinase 2 (RIPK2)
- RIPK3 *see* receptor-interacting protein kinase 3 (RIPK3)
- ritonavir 1:743
- RLC *see* regulatory light chain (RLC)
- RLRs *see* retinoic acid-inducible gene I-like receptors (RLRs); retinoic acid like receptors (RLRs); RIG-1-like receptor (RLR)
- RME *see* receptor-mediated endocytosis (RME)
- RME-8 2:479–480
- RMS *see* rostral migratory stream (RMS)
- RNA *see* ribonucleic acid (RNA)
- RNA aptamers 1:50, 1:51F
- RNA cleavage reaction 1:19–22
- RNA genetic switches *see* riboswitches
- RNA interference (RNAi) 2:221, 2:283, 2:422, 2:528, 2:531, 2:850, 3:300, 3:434, 4:228–229, 4:229, 4:264–265
- RNA interference screening 4:211
- RNA microarrays 1:50–52
in situ synthesis of 1:51
- RNA polymerase (RNAP) 1:397, 1:397F, 1:468, 1:481
 polymerases I 1:481
 polymerases II 1:481
 polymerases III 1:481
- RNA polymerase II (Pol II) 1:356–357F, 1:359, 1:364–365, 1:491, 3:788, 3:788–789
 miRNA transcription by 1:355
- RNA primase 1:396, 1:397
- RNA sequencing (RNA-seq) 4:160, 4:160–161, 4:161, 4:211, 4:220
 indirect and direct approaches to 4:161F
- RNA splicing 1:495, 1:496, 1:497, 1:497–498, 1:498–499, 1:498F, 1:500, 1:501
 chemical steps of 1:496F
 chemistry of 1:495–496
 coupled to other cellular process 1:501
 creation of multiple products from a single gene by 1:500–501
 defects in 1:501
 kinetic proofreading during 1:500F
- RNA viruses
 gene expression strategies of 1:523–524
 (–) strand RNA viruses 2:824
 (+) strand RNA viruses 2:819
 convoluted membranes 2:820–822
 host cell contribution 2:822–824
 membrane tubes and DMVs 2:822
 membrane webs 2:820–822
 spherules 2:819–820
- RNA world hypothesis 1:375, 1:397
- RNA-binding domains (RBD) 4:191–192
- RNA-dependent DNA polymerase 1:521
- RNA-dependent RNA polymerase (RdRp) 1:517, 1:520, 2:819
- RNA-dependent RNA synthesis 1:520F
- RNAi *see* RNA interference (RNAi)
- RNA-induced silencing complex (RISC) 1:354, 1:356–357F, 1:357F, 1:365
- RNAP *see* RNA polymerase (RNAP)
- RNase *see* endoribonuclease (RNase)
- RNaseP 1:22, 1:377–380
- RNA-seq *see* RNA sequencing (RNA-seq)
- RNC *see* ribosome-nascent chain complex (RNC)

RNP *see* ribonucleoprotein (RNP)
 rNTPs *see* ribonucleoside triphosphates (rNTPs)
 RO4 *see* rat rhodopsin 4 (RO4)
 robustness 2:373–374
 robustness and supply/demand 4:33, 4:33F
 ROCK *see* Rho kinase (ROCK); Rho-associated coiled-coil-containing protein kinase (ROCK)
 ROCs *see* receptor-operated channels (ROCs)
 Rod-Zw10-Zwilch complex (RZZ complex) 2:625
 ROK *see* Rho-Kinase (ROK)
 rolling 3:751
 rolling circle 1:19, 1:519
 ROMO1 2:283
 ROMO1/Higd-1a 2:283
 ROP (Rho-GTPases from plants) family 4:51
 ROR *see* RAR-related orphan receptor (ROR)
 ROS *see* reactive oxygen species (ROS)
 Rosetta Stone 3:795–796
 rostral migratory stream (RMS) 4:134
 neuronal migration 4:137
 rough endoplasmic reticulum (RER) 2:779
 rough ER 2:157
 roughness 1:111
 Rous sarcoma virus (RSV) 1:739
 Roux-en-Y-gastric bypass (RYGB) 2:805
 RP *see* red pulp (RP); retinitis pigmentosa (RP); ribosomal proteins (RP)
 RPA *see* replication protein A (RPA)
 RPC *see* replisome progression complex (RPC); reversed phase chromatography (RPC)
 RPE cells *see* retinal pigment epithelial cells (RPE cells)
 RPL *see* large subunit of ribosome (RPL)
 RP-LC *see* reversed-phase liquid chromatography (RP-LC)
 RP-MDM2-p53 pathway 3:286
 RPMs *see* red pulp macrophages (RPMs)
 Rpn4 1:591
 Rpn11 1:586, 1:589–590
 RPo 1:472
 R-project 4:240
 RPS *see* small subunit of ribosome (RPS)
 Rps3 ribosomal protein 3:289
 RQC complex *see* ribosome quality control complex (RQC complex)
 RRE *see* Rev responsive element (RRE)
 RRF *see* ribosomal recycling factors (RRF); ribosome releasing/recycling factor (RRF)
 rRNA *see* ribosomal RNA (rRNA)
 RRS *see* random reference set (RRS)
 RS *see* radial spokes (RS)
 R-SMADs *see* receptor-regulated SMADs (R-SMADs)
 R-spondin (Rspo) 3:89, 3:89F
 RSS *see* recombination signal sequences (RSS)
 RS-SCID *see* radio sensitive severe combined immunodeficiency (RS-SCID)
 RSV *see* respiratory syncytial virus (RSV); Rous sarcoma virus (RSV)

RT *see* reverse transcriptase (RT)
 RTC *see* reverse transcription complex (RTC)
 RTK signaling pathways *see* regulate receptor tyrosine kinase signaling pathways (RTK signaling pathways)
 RTKs *see* receptor tyrosine kinases (RTKs)
 RT-PCR *see* reverse transcription polymerase chain reaction (RT-PCR)
 RUBICON *see* RUN domain and cysteine-rich domain containing, BECLIN1-interacting protein (RUBICON)
 RubisCO *see* ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO)
 ruffles 2:683, 2:688–689, 2:758
rumi gene 2:350
 RUN domain and cysteine-rich domain containing, BECLIN1-interacting protein (RUBICON) 2:245
 Runt-related transcription factor-1 (RUNX1) 3:102
 Runt-related transcription factor 2 (RUNX2) 3:102
 rut site *see* Rho utilization site (rut site)
 ryanodine receptor/channel (RyR) 3:164
 RYGB *see* Roux-en-Y-gastric bypass (RYGB)
 RyR *see* ryanodine receptor/channel (RyR)
 RZZ complex *see* Rod-Zw10-Zwilch complex (RZZ complex)

S

20S core particle (20S CP) 1:589F, 1:590F
 structure and catalytic mechanism of 1:588–589
 30S initiation complex (30SIC) 1:385–387, 1:385F, 1:388
 S phase 3:423, 3:424, 3:427
 26S proteasome *see* proteasome(s): 26S proteasome
 19S regulatory particles (19S RPs) 1:588, 1:589–590
 70S ribosome 1:384, 1:385F, 1:388, 1:389F, 1:391F, 1:392F
 7S RNA 1:537T
 5S rRNA 1:384, 1:386F, 1:393, 1:393
 S1P *see* site-1 protease (S1P); sphingosine-1-phosphate (S1P)
 S1PR1 *see* sphingosine-1 phosphate receptor-1 (S1PR1)
 S2P *see* site-2 protease (S2P)
 S6 kinases (S6Ks) 3:243, 3:282
 S6K1 3:131, 3:245
 SA-beta-gal 3:511
 SAC *see* spindle assembly checkpoint (SAC)
 saccharolipids 1:156, 1:186, 1:188F
Saccharomyces cerevisiae 1:203, 1:427, 1:583, 1:584–585, 1:590F, 1:591, 2:466, 2:467–468, 2:468T, 2:469, 2:471–472, 2:741, 4:51
 cue-dependent cell polarization 4:51
 mitochondria 1:545, 1:546, 1:547, 1:548T, 1:550, 1:550–551
Saccharomyces cerevisiae spliceosome 1:495–496, 1:497T
 saccule 2:168, 2:171, 2:173–175
 saccin 2:286–287
 S-adenosyl methionine (SAM) 1:375–377, 1:379F
 SAFs *see* spindle assembly factors (SAFs)
 SAGE *see* serial analysis of gene expression (SAGE)
 SAgS *see* super-antigens (SAgS)
 SAHA *see* suberoylanilide hydroxamic acid (SAHA)
 SAHFs *see* senescence-associated heterochromatin foci (SAHFs)
 SAK *see* Snk/Plk-akin kinase (SAK)
 salivary proteins 1:732–733
 Salla disease 1:775
Salmonella enterica (*S. enterica*) 2:777, 2:779–780
Salmonella-containing vacuole (SCV) 2:777, 2:788
Salmonella-induced filaments (Sifs) 2:788–789
Salmonella pathogenicity islands (SPI) 2:785–786
 salt bridges 1:109
 saltatory conduction 1:244, 1:247–248
 salting in 1:70
 salting out 1:70
 Salvador (Sav) 3:99
 Salvador 1 (Sav1) 3:101–102
 Salvador–Warts–Hippo pathway *see* Hippo pathway
 salvage compartment 2:169
 SAM *see* S-adenosyl methionine (SAM)
 SAMBA 4:213
 SAMHD1 restriction factor 2:836
 SAMNet 4:215
 SAND-1/Mon1 2:504
 Sanfilippo syndromes 1:777, 1:780–781, 1:783, 1:784
 Sanger DNA sequencing 1:30–31
 Sanger's reagent 1:75
 SAP *see* SLAM-associated protein (SAP); sphingolipid activator proteins (SAP)
 sAPP β *see* soluble APP fragment (sAPP β)
 SARA *see* SMAD anchor for receptor activation (SARA)
 sarco-endoplasmic reticulum Ca²⁺ ATPase (SERCA) 2:159, 4:88
 sarcomeres 2:677
 calcium release from SR in skeletal muscle 2:679–681
 components 2:678T
 contraction 2:677–679
 structure 2:678F
 sarcoplasmic reticulum (SR) 2:156, 2:677, 4:86
 calcium release from 2:679–681
 and T-tubules organization 2:680F
 SA-Rho GEF *see* septin-associated Rho GEF (SA-Rho GEF)
 SARS *see* severe acute respiratory syndrome (SARS)
 SARS-CoV *see* Severe acute respiratory syndrome coronavirus (SARS-CoV)
 SAS-6 *see* spindle assembly abnormal protein 6 (SAS-6)

- SASP *see* senescence-associated secretory phenotype (SASP)
- satellite cell 3:776–777, 3:778, 3:789, 3:789–790, 3:796
- intrinsic modifications in aged 3:779–780
- saturation irradiance 2:77–78
- Sav *see* Salvador (Sav)
- Sav1 *see* Salvador 1 (Sav1)
- SAX *see* strong anion exchange (SAX)
- SAXS *see* small-angle X-ray scattering (SAXS)
- SBEs *see* SMAD binding elements (SBEs)
- SBF-SEM *see* serial block face scanning electron microscopy (SBF-SEM)
- SBML *see* systems biology markup language (SBML)
- SCA-1 *see* stem cell antigen 1 (SCA-1)
- scaffolding mechanism 1:205, 1:206F
- scaffolding/scaffold
- peptide 3:236
 - protein 2:603, 2:741, 3:495–496
- scale-free networks 4:68, 4:69F
- Scalloped (Sd) 3:99
- scanning electron microscopy (SEM) 2:1, 2:14, 2:21, 2:37
- beam–specimen interactions 2:15–16
 - in cell biology 2:21
 - construction 2:14
 - critical point drying 2:17
 - cryoscanning EM 2:18
 - environmental SEM 2:18
 - freeze etching 2:20
 - history 2:16–17
 - image formation 2:16
 - labeling for 2:17–18
 - and nuclear envelope 2:20–21
 - scanning specimen 2:14–15
 - serial block face imaging 2:18
 - specimen preparation for 2:17
 - sputter coating 2:17
 - and TEM merging 2:18–19
 - exposing subsurface detail 2:19
 - SEM of subcellular detail 2:19–20
 - TEM vs. 2:14
- SCAP *see* SREBP cleavage activating protein (SCAP)
- scar formation 3:216
- blood clotting and inflammation 3:216–219, 3:217F
 - phase of new tissue formation 3:219
 - tissue remodeling phase 3:220
- SCAR *see* suppressor of cAR (SCAR)
- scavenger receptor A (SR-A) 3:763
- scavenger receptor cysteine rich (SRCR) 1:729, 3:727
- Scavenger receptors (SRs) 2:752, 3:727–740, 3:727
- baseline expression 3:731F
 - in cancer 3:735–736
 - cell biology 3:728–729, 3:730F
 - frustrated phagocytosis 3:729
 - LOX-1 3:729
 - signaling 3:729–730
 - evolutionary origins 3:731
 - Class B SRs 3:731–732
 - contributing to cell adhesion 3:732
 - in corpse clearance 3:733
 - multiple SRs mediating normal self clearance 3:732–733
 - human SRs classes and functional domains 3:728F
 - in innate immunity 3:733
 - in inflammatory diseases 3:734–735
 - SRs as microbial entry points 3:734
 - SRs collaboration with PRR families 3:733–734
 - mammalian SRs nomenclature 3:729T
 - in metabolic disease association studies 3:735
 - monogenic diseases 3:736
 - SR expression transcriptional control 3:730–731
 - SR-specific domains share ancient origins 3:727–728
 - as therapeutic targets 3:736
- SCBC *see* single-cell barcode chip (SCBC)
- SCC *see* squamous cell carcinoma cells (SCC)
- SCD *see* stearyl desaturase (SCD)
- SCDO *see* spondylocostal dysostosis (SCDO)
- SCF *see* stem cell factor (SCF)
- scFv *see* single-chain Fv (scFv)
- SCG10 *see* super cervical ganglion clone 10 (SCG10)
- Schizosaccharomyces pombe*
- polarity regulation 4:54
- Schnaitman's approach 1:237
- Schwann cells 1:247–248
- SCID *see* severe combined immunodeficiency (SCID)
- SCIN *see* staphylococcal complement inhibitor (SCIN)
- sCJD *see* sporadic CJD (sCJD)
- SCMs *see* subcapsular macrophages (SCMs)
- SCN *see* severe congenital neutropenia (SCN)
- SCNT *see* somatic cell nuclear transfer (SCNT)
- SCOP *see* structural classification of proteins (SCOP)
- scrapie 2:839
- screening 4:264, 4:264–265
- Scribble (Scrib) 2:743, 3:100
- SCs *see* stem cells (SCs)
- SCS *see* subcapsular sinus (SCS)
- SCV *see* *Salmonella*-containing vacuole (SCV)
- SCX *see* strong cation exchange (SCX)
- Sd *see* Scalloped (Sd)
- SD sequence *see* Shine–Dalgarno (SD) sequence
- SDM *see* Site-directed mutagenesis (SDM)
- sdRNA *see* snoRNA-derived RNA (sdRNA)
- SDSA pathway *see* synthesis-dependent strand annealing pathway (SDSA pathway)
- SDS-PAGE *see* sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)
- SE *see* sterol esters (SE)
- sea urchin sperm protein 1:650
- seam 2:539, 2:539–540
- search-and-capture 2:638, 4:40
- SEC *see* secretion (SEC); size exclusion chromatography (SEC); super elongation complex (SEC)
- Sec1/Munc18 proteins (SM proteins) 2:460–461, 2:486
- Sec61 translocon 1:247–248, 1:535, 1:538
- second harmonic generation signals (SHG signals) 2:110, 2:129–131
- second law of thermodynamics 1:5–6
- second messenger 1:197–198, 1:198
- second messenger molecules 3:1
- second mitochondria-derived activator of caspases (SMAC) 3:69–70, 3:330
- secondary constrictions 3:257
- secondary electrons 2:16
- secondary granules 2:226, 2:770, 3:751
- secondary hemostasis
- initiation phase 1:716
 - propagation phase 1:716–717
- secondary ion mass spectrometry (SIMS) 1:152–153
- secondary lymphoid organs (SLOs) 3:616, 3:635–636
- stromal cells in 3:616
- secondary lymphoid tissues 3:6, 3:588
- secretagogues 2:377
- secreted Frizzled-related proteins (sFRPs) 3:88–89, 3:89F
- secreted glycoproteins, N-glycans functions on 2:344
- secretion (SEC) 1:57, 2:376
- secretion-recapture 1:778
- secretion-reuptake pathway 1:777–778
- secretory compartment 2:379–380
- secretory granules 1:557, 1:558, 1:558F, 1:560–561, 1:561F, 1:563–564, 2:390
- tracking 2:63
- secretory lysosomes
- biogenesis 2:228–229
 - exocytosis of 2:229–230
 - genetic diseases associate with defects in 2:230
- secretory pathway 2:156–157, 2:163, 2:354
- conceptual framework 2:373–374
 - constitutive secretion 2:371
 - control system balances 2:374F
 - extent of regulation network 2:371–372
 - Golgi-based signal molecules 2:372F
 - interaction with reticulum 2:779–780
 - manipulation to support bacterial replication 2:778–779
 - regulated secretion 2:371
 - regulation 2:371
 - signaling pathways 2:373
 - traffic regulation 2:372–373
- secretory protein
- biogenesis of 1:535–544
 - continuation of productive journey 1:601–602
 - N-glycosylation quality control pathway 1:599–601, 1:600F
 - perilous journey 1:597–599
 - translocation, folding, and ER exit of 1:598F

- secretory transport regulation 2:371
 sedimentation coefficient of the complex
 see 20S core particle (20S CP)
 SEDT *see* spondyloepi-physial dysplasia tarda (SEDT)
 seed matches 1:358–359
 seed sequence 1:354, 1:355*F*, 1:358, 1:358–359
 selected reaction monitoring (SRM) 1:569, 1:578–579
 see also targeted data acquisition
 selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) 1:27
 selective monofunctional uracil-DNA glycosylase (SMUG1) 1:444
 selective serotonin reuptake inhibitors (SSRIs) 1:248
 selenocysteine (Sec) 1:526
 SELEX *see* systematic evolution of ligands by exponential enrichment (SELEX)
 self-eating *see* autophagy
 self-enhancement 3:202, 3:202–203
 self-labeling proteins 2:83–84
 self-renewal 3:813, 3:814
 self-similar 4:59
 self-splicing introns 1:22
 self-tolerance, establishing 3:605
 central tolerance 3:605
 peripheral tolerance 3:605–606
 regulatory T cells (Tregs) 3:606–607
 role of CD28 3:606
 Seliciclib 4:99
 SEM *see* scanning electron microscopy (SEM)
 Sema3 *see* semaphorin class 3 (Sema3)
 semaphorin class 3 (Sema3) 4:106
 semiconservative telomere replication 1:428–430
 semi-robotic devices 2:6
 Semliki Forest virus (SFV) 2:168, 2:819–820
 senataxin 1:483
 senescence 1:418, 1:431, 3:436, 3:438, 3:511
 senescence-associated heterochromatin foci (SAHFs) 3:512–513
 senescence-associated secretory phenotype (SASP) 3:513, 3:514*F*
 sensing mechanisms 3:520–521
 sensitivity 2:373–374, 4:44, 4:45
 sensitized emission 2:54
 sensory organ precursors (SOP) 2:747
 sensory potential 1:218, 1:229
 separations methodology 1:76–77
 sepsis 4:98
 septate junction (SJ) 2:742
 Septin 6 (SEPT6) 2:581–583
 Septin 9 (SEPT9) 2:583
 septin unique element (SUE) 2:579
 septin(s) 2:579–587, 3:495–496
 and actin cytoskeleton 2:581–584
 interaction 2:582–583*F*
 in comparison to actin microfilaments, microtubules and intermediate filaments 2:580*T*
 diffusion barrier 2:581
 filaments 2:580, 2:581
 assembly 2:579*F*
 hetero-octamers 2:580
 and membrane skeleton 2:581
 and microtubule cytoskeleton 2:584–585
 interaction 2:582–583*F*
 polymerization 2:579–580
 structure, assembly, and dynamics 2:579–581
 septin-associated Rho GEF (SA-Rho GEF) 2:583
 septins 1:203–204, 2:4
 sequence and ligation independent cloning (SLIC) 1:122
 sequencer 1:77–78
 sequencing by oligonucleotide ligation and detection technology (SOLiD technology) 4:160, 4:162
 sequential elution from IMAC (SIMAC) 4:178
 sequential/sequencing
 models 1:9–10
 primer 1:30–31
 sequon 1:599
 SER *see* smooth endoplasmic reticulum (SER)
 Ser/Thr *see* serine/threonine (Ser/Thr)
 SERCA *see* sarco-endoplasmic reticulum Ca^{2+} ATPase (SERCA)
 Ser-Gln/Thr-Gln rich domain (SQ/TQ rich domain) 3:461–462
 serglycin 1:271, 1:274–275
 serial analysis of gene expression (SAGE) 4:160
 serial block face imaging 2:18
 serial block face scanning electron microscopy (SBF-SEM) 2:44, 2:45*F*, 2:46
 serial section scanning electron microscopy (ssSEM) 2:44
 serine palmitoyltransferase (SPT) 1:183–184
 serine protease domains (SPD) 1:650
 serine protease inhibitors (serpins) 1:709, 1:790–791
 activation and latency 1:711
 biological functions 1:711
 disease aspects 1:711
 mechanism of inhibition 1:710
 cysteine proteases 1:710–711
 serine proteases 1:710
 structure 1:709–710
 serine proteases 1:556*F*, 1:569–570, 1:710, 1:756
 matriptase 1:756–757
 plasminogen activation system 1:756
 serine/threonine (Ser/Thr) 3:781
 Serine-Arginine rich (SR) proteins 1:506
 serinopathy 1:706
 serpin G1 *see* C1-inhibitor
 serpins 3:332
 serpins *see* serine protease inhibitors (serpins)
 serrate 1:360, 2:350
 Sertoli-Leydig cell tumor 1:360
 serum and glucocorticoid-activated kinase (SGK) 3:124
 severe acute respiratory syndrome (SARS) 1:654
 Severe acute respiratory syndrome coronavirus (SARS-CoV) 2:820–822
 severe combined immunodeficiency (SCID) 3:620–621
 diseases 3:24–25
 syndrome 3:145
 severe congenital neutropenia (SCN) 3:754
 severing proteins 2:548, 2:554, 2:559
 SEX *see* starch excess (SEX)
 sex-chromosome 4:156
 Sex-determining region Y (SRY) 3:104
 sex-determining region Y-box 9 (SOX9) 3:794–795
 SFC *see* supercritical fluid chromatography (SFC)
 SFD *see* Sorsby fundus dystrophy (SFD)
 sFI *see* sporadic fatal insomnia (sFI)
 SFK *see* Src-family kinases (SFK); Src family of tyrosine kinases (SFKs)
 sFlt1 *see* splice variants of VEGFR1 (sVEGFR1)
 sFRPs *see* secreted Frizzled-related proteins (sFRPs)
 SFV *see* Semliki Forest virus (SFV)
 SG *see* specific granules (SG)
 sGCs *see* soluble GCs (sGCs)
 SGK *see* glucocorticoid-stimulated kinase (SGK); serum and glucocorticoid-activated kinase (SGK)
 SGs *see* stress granules (SGs)
 SGZ *see* subgranular zone (SGZ)
 SH *see* src homology (SH)
 SH1 *see* SRC homology region 1 (SH1)
 SH2 *see* Src homology 2 (SH2)
 SH2 domain-containing leukocyte phosphor protein of 76 kDa (SLP76) 3:534–535
 SH2-containing phosphatase 1 (SHP-1) 3:533
 SH2-containing protein tyrosine phosphatase 2 (SHP2) 3:32
 SHP-2 phosphatase 3:119
 SH3 *see* Src homology 3 (SH3)
 shade chloroplasts
 sun chloroplasts vs. 2:329–330
 Shannon's information theory 4:258
 SHAPE *see* selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE)
 shared glycoprotein 130 receptors 3:27–28
 Shar-Pei *see* Salvador (Sav)
 SHD *see* Slp homology domain (SHD)
 Shed syndecan-1 1:276
 shedding 3:13
 shelterin 1:427, 1:428, 1:428–429, 1:430
 Shelterin protein complex 1:419*F*, 1:420
 SHG signals *see* second harmonic generation signals (SHG signals)
 SHH *see* sonic hedgehog (SHH)
 Shiga toxin (STx) 2:796–797, 2:797
Shigella 2:761, 2:781
Shigella flexneri (*S. flexneri*) 2:791
 Shine-Dalgarno (SD) sequence 1:322, 1:384–385, 1:386*F*, 1:545, 1:546*F*, 1:550
 SHIP2 2:504, 2:508

- Shisa 3:88–89
- SHM *see* somatic hyper-mutation (SHM)
- short gastrulation (*sog*) 4:65–66
- short hairpin RNAs (shRNA) 1:485
- short interfering RNAs (siRNA) 3:281
- short-loop recycling pathway 2:401
- shotgun 4:167–168
- lipidomic mass spectrometry 1:151–152
- shotgun lipidomics 1:157
- SHP-1 *see* SH2-containing phosphatase 1 (SHP-1)
- SHP2 *see* SH2-containing protein tyrosine phosphatase 2 (SHP2)
- shRNA library *see* small hairpin RNA (shRNA) library
- shRNA *see* short hairpin RNAs (shRNA)
- shrub 2:470, 2:471
- shugoshin 3:473
- shunting 1:249
- shuttling 3:176
- sialated glycans 1:126
- sialic acid (NeuAc) 2:339–340
- sialic-acid binding receptor CD169 3:759–760
- sialin 1:763–764, 1:775
- sialyl-Lewis X (SLeX) 2:342
- Sif 2:779
- SifA protein 2:779
- SifA kinesin-1-interacting protein (SKIP) 2:779
- Sifs *see* *Salmonella*-induced filaments (Sifs)
- siglec-1 *see* sialic-acid binding receptor CD169
- σ^{54} 1:477
- σ^{70} 1:469–471
- SIGN *see* specific ICAM-3-grabbing nonintegrin (SIGN)
- signal amplification 4:7–8
- and diversification module 3:534–536
- through excitable network 4:8
- through positive feedback 4:8
- through receptor coupling 4:8
- signal initiation module 3:534
- signal peptidase complex (SPC) 1:535, 1:536F, 1:538, 1:539–540T
- signal peptide 1:535, 1:536, 1:536–538, 1:537, 1:538, 1:557
- signal recognition particle (SRP) 1:535, 1:536–538, 2:160, 2:447
- SRP receptor 1:535
- signal regulatory protein α (SIRP α) 3:701
- signal recognition particle pathway (SRP pathway) 1:597–598
- signal sensitivity 4:7–8
- signal sequence (SS) 1:597–598
- signal transducer and activator of transcription (STAT) 3:23
- family 3:12
- pathway 4:66
- proteins 1:93, 3:136, 3:142–144
- in human disease 3:146
- STAT3 2:220, 3:541, 4:189–190
- signal transduction 2:445, 2:446–447, 3:580–581, 4:105–106, 4:218, 4:219
- exotoxins altering 2:794–795
- kinesins as components 2:614
- kinesins regulation by 2:614–616
- through TNFR activation 3:62–63
- signal transduction pathways 1:507, 1:543, 1:582, 3:1, 3:2
- signaling 2:176–177, 3:728–729
- cellular decisions encoded in 4:46–48
- cell-type specificity 4:46
- feedback system 4:46
- integrated response 4:47–48, 4:47F
- signaling specificity 4:46
- code by TCR signalosome
- account for specificity and sensitivity 3:532–533
- T cell activation conformational model 3:533–534
- T cell activation kinetic model 3:531–532
- compartmentalization of 4:46
- complexes
- central supramolecular activation cluster 3:681–682
- microclusters 3:682
- networks 4:219
- pathways 2:373
- plant hormones 2:254
- platelet activating factor 2:254–255
- signaling ligands 3:3, 3:122
- signaling lymphocytic activation molecule (SLAM) 3:522–524, 3:543
- signals to migration 3:208
- during development 3:209–210
- directed cell migration 3:208–209
- in physiological/pathological processes 3:210
- signalosomes 1:195–196
- biochemical module and assembly 3:534
- LCK activation 3:535F
- negative regulation modules 3:536–538, 3:537F
- signal amplification and diversification module 3:534–536
- signal initiation module 3:534
- modular organization 3:535F
- signaling code 3:531–532
- TCR 3:528
- signal-transducing adapter molecule (STAM) 2:195, 2:409–410
- SIGN-R1/MARCO macrophages 3:762–763
- SILAC *see* stable isotope labeling of amino acids in cell culture (SILAC)
- silencer 1:488, 1:506
- silencer of death domains (SODD) 3:62
- SIM *see* structured illumination microscopy (SIM)
- SIMAC *see* sequential elution from IMAC (SIMAC)
- simian immunodeficiency viruses (SIV) 2:828
- Simple Modular Architecture Research Tool (SMART) 1:95, 3:112
- Simpson–Golabi–Behmel syndrome 1:275
- SIMS *see* secondary ion mass spectrometry (SIMS)
- simulated annealing 1:98, 1:98–99, 4:212, 4:212–213, 4:212T
- SIN1 *see* stress-activated protein kinase-interacting protein 1 (SIN1)
- Sindbis virus (SINV) 2:819–820
- single molecule approach 2:101
- FCS as 2:101
- single molecule fluorescence microscopy 3:661
- single molecule tracking 3:662–663, 3:663F, 3:664
- single nucleotide polymorphism (SNP) 1:366, 1:488, 1:677, 1:794, 3:158, 3:454, 3:585, 4:246–247, 4:259
- single photon excitation (1PE) 2:80
- single tyrosine-based motif (hemiITAM) 3:696, 3:710–712
- single-cell analysis 4:220
- single-cell signaling events 4:221
- single-cell transcriptomics 4:220–221
- single-cell barcode chip (SCBC) 3:667
- single-cell interrogation of immune system 3:667–676
- cell–cell communication 3:672–673
- cell migration 3:669–672
- dissecting biological noise 3:674–675
- medical applications 3:673–674
- single-cell secretion 3:667–669, 3:668–669F
- single-cell migration 2:720
- see also* multicellular migration
- amoeboid migration 2:723
- blebbing motion 2:723
- features 2:722–723F
- mesenchymal migration 2:720–723
- single-cell signaling events 4:221
- protein interactions and activity measurement 4:221–222
- protein translocation measurement 4:221
- single-cell signaling networks modeling and analysis 4:222–223
- single-cell transcriptomics 4:220
- near-term challenges 4:221
- technical limitations 4:220–221
- single-chain Fv (scFv) 1:142–143
- single-headed myosins 2:601
- single-molecule, real-time sequencing technology (SMRT) 4:160, 4:161
- single-molecule detection (SMD) 2:115
- single-particle electron cryo-microscopy (cryoEM)
- conformational changes of a molecular machine, delineating 2:11
- density maps, resolution of 2:7–8
- frozen, hydrated specimen preparation for 2:6
- high-resolution structures 2:9
- identifying different populations within a heterogeneous sample 2:9–11
- image acquisition 2:6
- image processing in 2:6–7
- of macromolecular complexes 2:5–13
- model building and model validation 2:8–9
- preferred orientation 2:6, 2:7
- sample consideration for 2:5–6
- single-particle tracking (SPT) 2:121, 2:124
- HMM 2:126

- models of motion 2:125T
MSD 2:125F
MSD-based analysis 2:126
sptPALM 2:124
single-particle tracking PALM (sptPALM) 2:124
single-photon capability 2:113
single-plane illumination microscopy (SPIM) *see* light sheet fluorescence microscopy (LSFM)
single-strand annealing (SSA) 1:461
single-strand binding proteins (SSB proteins) 3:459
single-strand DNA (ssDNA) 3:458
single-strand DNA-binding protein (SSB) 1:396, 1:397, 1:407, 1:407F
bacterial SSB 1:407
eukaryotic RPA 1:407–408
single-stranded DNA (ssDNA) 1:427, 1:456
single-stranded RNA (ssRNA) 1:355, 1:359, 1:368
SINV *see* Sindbis virus (SINV)
SipA 2:787
SipC 2:787
siRNA *see* short interfering RNAs (siRNA); small interfering RNA (siRNA)
sirolimus *see* rapamycin
SIRP α *see* signal regulatory protein α (SIRP α)
SIRT1 3:789–790
sister chromatid cohesin 1:412
site-1 protease (S1P) 2:531
site-2 protease (S2P) 2:531
site-directed mutagenesis (SDM) 1:122–127
applications 1:122–123, 1:123F
methods 1:122
primer extension method 1:122
protein engineering 1:124
altered function 1:124
site-selective conjugation 1:125–126
stability 1:124–125
protein structure-function 1:123
binding sites 1:123–124
catalytic residues 1:123
regulatory elements 1:124
QuickChange method of 1:122F
SIV *see* simian immunodeficiency viruses (SIV)
SIV infecting sooty mangabeys (SIV smm) 2:828
six degrees of separation 4:68
size exclusion chromatography (SEC) 1:71–72, 4:167
SJ *see* septate junction (SJ)
Sjögren's syndrome 1:776–777
sKDR *see* splice variants of VEGFR2 (sVEGFR2)
skeletal muscle 2:677–682, 3:776–777
ATP hydrolysis and myosin stroke cycle 2:679F
conformation change of tropomyosin and troponin 2:681F
costamere components 2:681T
fibers 3:796F
muscle disorders associated with 2:681
sarcomeres 2:677
calcium release from SR 2:679–681
components 2:678T
contraction 2:677–679
structure 2:678F
SR and T-tubules organization 2:680F
triad components 2:680T
unique tissue-specific progenitor cell 3:796–797
skeletal muscle stem cells 3:789–790
skin 1:685
skin hair follicle
hair follicle stem cell populations 3:794–795
signaling between follicle stem cells and microenvironment 3:795–796
skin papillomas 3:810
skin progenitor cells (SKP) 3:819
skin wounds
in phase of new tissue formation 3:217F
granulation tissue formation 3:219–220
skin wounds reepithelialization 3:219
reepithelialization 3:219
in tissue remodeling phase 3:218F, 3:220
SKIP *see* SifA kinesin-1-interacting protein (SKIP)
SKP *see* skin progenitor cells (SKP)
SKP1 *see* S-phase kinase associated protein 1 (SKP1)
SKP2 *see* S-phase kinase associated protein 2 (SKP2)
Slac2-a 2:237
SLAM *see* signaling lymphocytic activation molecule (SLAM)
SLAM-associated protein (SAP) 3:543
SLB *see* supported lipid bilayers (SLB)
SLC11a2 iron transporter 2:478
SLE *see* systemic lupus erythematosus (SLE)
SLeX *see* sialyl-Lewis X (SLeX)
SLIC *see* sequence and ligation independent cloning (SLIC)
sliding clamps 1:396
conservation 1:400–401
sliding disintegration 2:629
sliding model 2:663
Slik *see* sterile20-like kinase (Slik)
slip bonds 3:655
Slit-Robo Rho GTPase activating protein family (srGAP family) 2:493–494
SLM *see* spatial light modulator (SLM)
SLOs *see* secondary lymphoid organs (SLOs)
SLOS *see* Smith–Lemli–Opitz syndrome (SLOS)
slow-killing process 2:845–846
Slp homology domain (SHD) 2:237
Slp2-a *see* synaptotagmin-like proteins 2a (Slp2-a)
Slp4-a *see* synaptotagmin-like proteins 4a (Slp4-a)
SLP76 *see* SH2 domain-containing leukocyte phosphor protein of 76 kDa (SLP76)
SLRPs *see* small leucine-rich proteoglycans (SLRPs)
Sly syndrome 1:777
SM *see* sphingomyelins (SM)
SM proteins *see* Sec1/Munc18 proteins (SM proteins)
SMA *see* spinal muscular atrophy (SMA)
SMAC *see* second mitochondria-derived activator of caspase (SMAC)
SMAD anchor for receptor activation (SARA) 2:212, 2:219, 3:40–42
SMAD binding elements (SBEs) 3:44
small GTPases 2:759, 4:51
Rab5 2:760
Rac 2:759–760
Ras 2:759
small hairpin RNA (shRNA) library 3:674
small interfering RNA (siRNA) 1:27, 1:42–43, 1:354, 1:355, 1:358, 4:267
small intestine in proteolysis 1:733–735
small leucine-rich proteoglycans (SLRPs) 1:273–274, 2:700
small molecule 3:385
small nuclear ribonucleoproteins (snRNP) 1:495, 1:496, 1:497, 1:497–498, 1:497T, 1:498, 1:498–499, 1:498F, 1:499F, 1:501, 1:503, 1:503–504
biogenesis 1:504F, 1:507
small nuclear RNA (snRNA) 1:495, 1:496–498, 1:498–499, 1:499, 1:499–500, 1:499F, 1:503, 1:503–504, 1:504F, 1:505, 1:506, 1:507, 1:508, 4:160, 4:162
small nucleolar ribonucleoprotein (snoRNP) 3:256–257
small nucleolar RNAs (snoRNAs) 3:256–257, 3:298–299, 4:160, 4:162
-derived RNA (sdRNA) 4:160, 4:162
small ribosomal proteins 3:3
small RNAs and cancer 1:364
miRNAs 1:364
biogenesis 1:364–365
and cancer 1:366
discovery 1:364
etiology 1:366
IsomiRs 1:368–369
nomenclature 1:365–366
prognosis predicting and therapy response 1:366–367
role in carcinogenesis 1:367
as a target for therapy 1:367–368
noncoding RNAs 1:364
ceRNAs 1:369
eRNAs 1:369
LncRNAs 1:369–370
transposons and piRNAs 1:370
Small SubUnit (SSU) 3:297–298
small subunit of ribosome (RPS) 3:281
small synaptic vesicles 2:380
small ubiquitin-like modifier protein (SUMO protein) 2:287–288, 2:574
small vessels vasculitis (SVV) 3:755
small-angle X-ray scattering (SAXS) 1:94, 3:704
small-world networks 4:67–68, 4:69F
SMART *see* Simple Modular Architecture Research Tool (SMART)
Smaug recognition elements (SREs) 3:323
SMB *see* somatomedin-B (SMB)

- SMC *see* smooth muscle cell (SMC);
structural maintenance of
chromosomes (SMC)
- SMD *see* single-molecule detection (SMD)
- Smith–Lemli–Opitz syndrome (SLOS)
1:173, 1:178, 3:77
- SMN *see* survival motor neuron (SMN)
- smooth endoplasmic reticulum (SER)
2:324–325
- smooth muscle cell (SMC) 3:57, 4:103
- SMRT *see* single-molecule, real-time
sequencing technology (SMRT)
- SMUG1 *see* selective monofunctional uracil-
DNA glycosylase (SMUG1)
- SNAP *see* soluble NSF attachment protein
(SNAP)
- SNAP23 *see* synaptosome-associated protein
of 23 kDa (SNAP23)
- SNAP-25 1:248
- SNAP REceptor *see* Soluble NSF Attachment
Protein REceptor (SNAP REceptor)
- SNAREs *see* soluble *N*-ethylmaleimide-
sensitive factor attachment protein
receptors (SNAREs)
- Snell's law 2:62
- S*-nitrosylated proteins (SNO-proteins) 1:88
- Snk/Plk-akin kinase (SAK) 2:651–653
- SNO-proteins *see* *S*-nitrosylated proteins
(SNO-proteins)
- snoRNAs 3:291
- snoRNAs *see* small nucleolar RNAs
(snoRNAs)
- snoRNP *see* small nucleolar
ribonucleoprotein (snoRNP)
- SNP *see* single nucleotide polymorphism
(SNP)
- snRNA *see* small nuclear RNA (snRNA)
- snRNP *see* small nuclear ribonucleoproteins
(snRNP)
- Snu114 1:497
- SNX *see* sorting nexin (SNX)
- SNX3 *see* sorting nexin-3 (SNX3)
- SOCs *see* store-operated channels (SOCs)
- SOCS *see* suppressor of cytokine signaling
(SOCS)
- SOCS proteins *see* suppressors of cytokine
signaling proteins (SOCS proteins)
- SODD *see* silencer of death domains
(SODD)
- sodium bisulfate, methylcytosine reaction
with 1:31F
- sodium butyrate 3:790–791
- sodium dodecyl sulfate polyacrylamide gel
electrophoresis (SDS-PAGE) 1:68,
1:77, 2:699
- gel 1:68–69
- sodium–calcium exchanger (NCX) 1:675
NCX3 1:675
- sodium–potassium pump current,
formulation of 4:87F
- SODs *see* superoxide dismutases (SODs)
- SOFI *see* super-resolution optical fluctuation
imaging (SOFI)
- sog* *see* short gastrulation (*sog*)
- SOH *see* SOL homology (SOH)
- SOL homology (SOH) 1:671
- SOLiD technology *see* sequencing by
oligonucleotide ligation and
detection technology (SOLiD
technology)
- solubility-enhancing protein domains 1:57T
- soluble APP fragment (sAPP β) 1:696
- soluble chemoattractants 2:724
- soluble GCs (sGCs) 3:56, 3:56–57
- soluble hydrolases 1:613
- soluble monomers 2:409–410
- soluble *N*-ethylmaleimide-sensitive factor
attachment protein receptors
(SNAREs) 1:248, 2:171, 2:186,
2:241, 2:364–365, 2:381–382,
2:459–465, 2:504, 2:581, 3:406,
3:484–486
- complexes 2:355–356
- longin v-SNAREs Sec22b and TI-VAMP/
VAMP7 2:461–463
- VAMP7 network of protein–protein
interactions 2:463F
- membrane fusion nanomachines
2:459–460
- core mechanism 2:460F
- regulators 2:460–461
- proteins 1:248, 1:764, 1:765, 2:164,
2:367–368, 2:485
- vesicle tether interactions with small
GTPases and 2:485–486
- soluble NSF attachment protein (SNAP)
3:752
- receptor 2:796
- SNAP-23 2:225–226
- SNAP-25 2:298
- Soluble NSF Attachment Protein REceptor
(SNAP REceptor) 1:616
- soluble TNF (sTNF) 3:62–63
- soluble VEGFR1 (sVEGFR1/sFlt1) 4:111–112
- solvent 1:109
- somatic cell nuclear transfer (SCNT) 3:819
- somatic centrosomes 2:656
- somatic hyper-mutation (SHM) 3:530–531,
3:552
- somatic mutations 1:360, 1:507–508, 1:508
- somatic single-nucleotide variants (SSNVs)
3:673–674
- somatomedin-B (SMB) 3:728
- somitogenesis, multiscale modeling of
4:128, 4:129F
- SON 3:819
- son of sevenless protein (SOS protein)
3:119
- sonic hedgehog (SHH) 1:369, 3:2, 3:76,
3:343
- sonication 1:66, 1:69
- SOP *see* sensory organ precursors (SOP)
- S-Opal 2:283–284
- SorL1 protein 2:478
- SorLA *see* SorL1 protein
- Sorsby fundus dystrophy (SFD) 1:712–713
- sorting 2:175–176
- sorting endosome 2:193
- sorting nexin (SNX) 2:196, 2:410, 2:475,
2:476, 2:499
- proteins 2:761
- SNX-BAR proteins 2:476
- SNX1 2:779
- SNX3 2:505, 3:87
- SNX3 2:779
- SNX4 2:505, 2:511
- SNX5 2:505, 2:508
- SNX9 2:505, 2:508
- SNX17 2:426
- sorting signals 2:427–428, 2:450
- SOS protein *see* son of sevenless protein
(SOS protein)
- Sox2 *see* SRY-box 2 (Sox2)
- SOX9 *see* sex-determining region Y-box 9
(SOX9)
- SP *see* sulfopropyl (SP)
- spacing 3:193–194
- SPADE *see* Spanning-tree progression
analysis of density-normalized
events (SPADE)
- Spanning-tree progression analysis of
density-normalized events (SPADE)
4:223
- SPARC 2:700
- sparsely populated states of proteins and
complexes 1:103
- lifetime line broadening and DEST
1:104–105
- PRE 1:103–104
- principal NMR methods 1:103
- relaxation dispersion spectroscopy
1:104
- relevance to cell biology 1:105
- spatial light modulator (SLM) 2:78, 2:89
- 'spatial' models 4:17
- spatial proteomics 3:265
- spatial relocation 3:720–721
- spatial restriction 2:505–507
- spatial sensing 4:6
- SPB *see* spindle pole body (SPB)
- SPC *see* signal peptidase complex (SPC);
subtilisin-like pro-protein convertase
(SPC)
- SPD *see* serine protease domains (SPD)
- specific granules (SG) 2:770
- see also* secondary granules
- specific ICAM-3-grabbingnonintegrin
(SIGN) 3:746
- specific *N*-terminal residue 1:326
- specific pathogen-free mice (SPF mice)
2:800–801
- specimen preparation
for SEM 2:17
- spectrin 3:231
- sperm formation 2:248
- spermiogenesis 2:248
- SPF mice
see also specific pathogen-free mice
(SPF mice)
- S-phase kinase associated protein 1 (SKP1)
3:439
- S-phase kinase associated protein 2 (SKP2)
3:439
- spheroid 3:808, 4:74–75
- spherules 2:819–820
- formation 2:820F
- sphinganine 1:183–184
- sphingolipid activator proteins (SAP) 1:774

- sphingolipid(s) 1:155, 1:195–196, 1:195, 1:202, 1:203–204, 1:205, 1:208, 1:208–209, 1:209–210, 1:615–616, 2:275–276
 backbone biosynthesis 1:186–189, 1:189F
 sphingolipidosis 1:777, 1:779
 sphingomyelins (SM) 1:187, 1:208–209, 1:783, 2:401
 biosynthesis 2:780
 sphingosine 1:183–184
 sphingosine 1-phosphate (S1P) 1:195, 3:51, 3:102–103, 3:330–331, 3:616–617, 4:113
 sphingosine-1 phosphate receptor-1 (S1PR1) 3:616–617, 3:767
 SPI *see* *Salmonella* pathogenicity islands (SPI)
 spinal cord 1:115
 spinal muscular atrophy (SMA) 1:501, 1:508, 3:394
 spindle assembly 2:625
 spindle assembly abnormal protein 6 (SAS-6) 2:651
 spindle assembly checkpoint (SAC) 2:625, 2:644, 2:657, 3:401, 3:471, 3:471, 3:472F, 3:480, 3:482, 3:483F
 spindle assembly factors (SAFs) 2:316–317
 spindle length control 2:641
 spindle microtubules 3:494, 3:495F
 spindle midzone 2:639
 spindle pole body (SPB) 2:315, 2:649
 spinesin *see* TMPRSS5
Spiroplasma melliferum 2:27, 2:27F
 spleen, macrophage subsets in 3:759, 3:760F
 apoptotic cells clearance by MZ macrophages 3:763
 metallophilic macrophages 3:761–762
 RPMs 3:763–765
 SIGN-R1/MARCO macrophages 3:762–763
 splenic MZ and macrophages 3:759–761
 spleen tyrosine kinase (syk) 3:528, 3:751
 splenic MZ 3:759–761
 macrophage populations in 3:761F
 splice sites 1:503, 1:503–504, 1:505, 1:505F, 1:506
 5' splice site 1:495–496, 1:496, 1:496F, 1:497, 1:498–499, 1:498F, 1:499–500, 1:499F, 1:501, 1:503, 1:504F, 1:505F, 1:506
 splice variants of VEGFR1 (sVEGFR1) 4:106
 splice variants of VEGFR2 (sVEGFR2) 4:106
 spliceosome 1:503, 1:503–505, 1:506, 1:507, 1:509
 active site of 1:499–500, 1:499F
 alternative splicing and creation of multiple products from single gene 1:500–501
 assembly on introns 1:498–499
 chemistry of splicing and architecture of intron 1:495–496
 comparison of snRNP components for *Saccharomyces cerevisiae* and *Homo sapiens* 1:365–366, 1:366, 1:497T, 1:569
 defects in splicing 1:501
 DExD/H box proteins are essential cofactors of the 1:500
 diseases of 1:507–508
 and pre-mRNA splicing 1:495–502
 retinitis pigmentosa and neurodegeneration 1:508
 snRNAs and proteins 1:496–498
 splicing factor mutations and cancer 1:508
 splicing mutations in growth and development 1:508–509
 stepwise assembly mechanism of the 1:498F
 U12-dependent 1:505
 U2-dependent 1:505
 splicing code 1:503, 1:506, 1:508–509
 splicing factor 1 (SF1) 1:497–498, 1:498–499, 1:498F, 1:501
 splicing regulatory (SR) proteins 1:497–498
 split PH domain (spPH) 3:153
 s-pol incident light 2:63–64
 spondylocostal dysostosis (SCDO) 3:109–110
 spondyloenchondrodysplasia 1:776–777
 spondyloepi-physial dysplasia tarda (SED) 2:335
 sporadic CJD (sCJD) 2:840
 sporadic fatal insomnia (sFI) 2:840–841
 spPH *see* split PH domain (spPH)
 SPR *see* surface plasmon resonance (SPR)
 SPRI measurements *see* surface plasmon resonance imaging (SPRI) measurements
 spring forces 4:120
 sprout fusion and flow 4:112–113
 sprout lumen formation 4:112, 4:113F
 sprouting angiogenesis 4:102, 4:110, 4:110F
 EC-EC junctions destabilization 4:110–111
 ECM proteolysis 4:110
 endothelial junction structure and destabilization 4:111F
 migration and proliferation 4:111–112
 pericytes dissociation 4:111
 sprout fusion and flow 4:112–113
 sprout lumen formation 4:112, 4:113F
 tip-stalk cell signaling 4:112F
 vessel pruning 4:113–114
 vessel stabilization and maturation 4:113
 SPT *see* serine palmitoyltransferase (SPT); single-particle tracking (SPT)
 Spt5 1:477–478
 sptPALM *see* single-particle tracking PALM (sptPALM)
 Spumaretroviruses 1:523
 sputter coating 2:17
 SQ *see* sulfoquinovose (SQ)
 SQ/TQ rich domain *see* Ser-Gln/Thr-Gln rich domain (SQ/TQ rich domain)
 squamous cell carcinoma cells (SCC) 2:723, 3:110
 SR *see* sarcoplasmic reticulum (SR); SRP receptor (SR)
 SR proteins *see* Serine-Arginine rich (SR) proteins; splicing regulatory (SR) proteins
 SR-A *see* scavenger receptor A (SR-A)
 Src family 3:439, 3:442–444
 Src family of tyrosine kinases (SFKs) 3:696
 src homology (SH) 2:597, 3:112, 3:112–113
 domains and diseases 3:119–120
 SH2 domains 3:115–116
 SH3 domains 3:113–114
 Src homology 2 (SH2) 1:92, 2:752–753, 3:31–32, 3:132
 domains 3:115
 specificity and promiscuity 3:117–118
 structure 3:116, 3:116F
 target recognition motifs 3:116–117
 function 3:118
 SHP-2 phosphatase 3:119
 SOS protein 3:119
 SRC Kinase modulation of activity 3:119
 regulating pathway activation 3:118F
 Src homology 3 (SH3) 1:92, 3:703
 domains 3:113–114, 3:153
 atypical binding motifs 3:115
 function 3:118–119
 model for bidirectional binding of proline type 2 helices by 3:114F
 specificity and promiscuity 3:117–118
 structure 3:113F, 3:114
 target recognition motifs 3:114–115
 function 3:118
 SHP-2 phosphatase 3:119
 SOS protein 3:119
 SRC Kinase modulation of activity 3:119
 SRC homology region 1 (SH1) 3:112–113
 SRC Kinase modulation of activity 3:119
 Src-family kinases (SFK) 2:759
 SRCR *see* scavenger receptor cysteine rich (SRCR)
 SREBP *see* sterol response element binding protein (SREBP)
 SREBP cleavage activating protein (SCAP) 1:175–176, 1:176F
 SREs *see* Smaug recognition elements (SREs)
 srGAP family *see* Slit-Robo Rho GTPase activating protein family (srGAP family)
 SrGAP3 2:493–494
 SRM *see* selected reaction monitoring (SRM)
 SRP *see* signal recognition particle (SRP)
 SRP pathway *see* signal recognition particle pathway (SRP pathway)
 SRP receptor (SR) 1:535, 1:537–538, 2:447
 SRP9 1:537T
 SRP14 1:537T
 SRP19 1:537T
 SRP54 1:537, 1:537T
 SRP68 1:537T
 SRP72 1:537T
 SRs *see* Scavenger receptors (SRs)
 SRSF1 1:508
 SRSF2 1:508
 SRY *see* Sex-determining region Y (SRY)
 SRY-box 2 (Sox2) 3:104, 3:786, 3:788–789, 3:790–791
 SS *see* signal sequence (SS)
 SSA *see* single-strand annealing (SSA); stochastic simulation algorithm (SSA)
 SSB *see* single-strand DNA-binding protein (SSB)

- SSB proteins *see* single-strand binding proteins (SSB proteins)
- ssDNA *see* single strand DNA (ssDNA); single-stranded DNA (ssDNA)
- ssDNA viruses 1:519, 1:519F
- SSNVs *see* somatic single-nucleotide variants (SSNVs)
- SSRIs *see* selective serotonin reuptake inhibitors (SSRIs)
- ssRNA *see* single-stranded RNA (ssRNA)
- ssSEM *see* serial section scanning electron microscopy (ssSEM)
- SSU *see* Small SubUnit (SSU)
- stable isotope labeling of amino acids in cell culture (SILAC) 3:264–265, 4:169, 4:182
- stalk cell 4:102, 4:112
- tip-stalk cell signaling 4:112F
- STAM *see* signal-transducing adapter molecule (STAM)
- STAM-Hrs complex 2:409–410
- Standard of expressed protein (STEP) 1:578–579
- staphylococcal complement inhibitor (SCIN) 1:729
- staphylococcal nuclease 1:14
- Staphylococcus aureus* infection 1:255–256
- Staphylococcus aureus* pneumonia
immune response to 4:100
- staple technology 2:358
- STAR protein *see* steroidogenic acute regulatory protein (STAR protein)
- starch 1:264–265
- degradation 1:268–269
- enzyme activities 1:266F
- function 1:265
- granules 1:265
- metabolism
- genetic modification 1:269–270
- starch degradation 1:268–269
- starch synthesis 1:267–268
- synthesis 1:267–268, 2:328–329
- starch excess (SEX) 1:268–269
- starvation, metabolism in *see* fasted state, metabolism in
- STAT *see* signal transducer and activator of transcription (STAT)
- stathmin-2 *see* super cervical ganglion clone 10 (SCG10)
- stationary compartment 2:175
- steady state flux 4:27, 4:27–28, 4:27F, 4:29
- steady-state kinetics 1:9
- stearoyl desaturase (SCD) 3:87
- steatosis 1:792, 1:796
- STED microscopy *see* stimulated emission depletion microscopy (STED microscopy)
- stefins *see* cystatins (CSTs)
- stem cell antigen 1 (SCA-1) 3:795
- stem cell biology 3:6–7
- stem cell factor (SCF) 3:14, 3:624–625
- stem cell niche 3:789
- stem cells (SCs) 3:776, 3:813–818
- biology of adult 3:813F
- determinants of aging 3:776
- cellular consequences of genetic and epigenetic changes 3:778
- DNA damage and repair 3:777
- epigenetic alterations 3:777–778
- telomere shortening 3:777
- fatty acid metabolism 3:814–815
- glycolysis in 3:813–814
- HIF-1 α 3:813–814
- metabolism 3:817F
- mitochondrial function 3:814–815
- niches 3:816–817
- nutrient-sensing pathways 3:815
- redox homeostasis 3:815–816
- self-renewal 3:438, 3:440
- stem cells, epigenetic regulation of 3:785–793
- effectors and mechanisms of 3:785–786
- chromatin remodeling 3:786–787
- DNA methylation 3:787–788
- enhancer activation and developmental competence 3:785–786
- noncoding RNA (ncRNA) 3:788
- promoter bivalency and maintenance of pluripotency 3:786
- epigenetic signatures and control of nuclear architecture in ESCs 3:788–789
- gene expression, regulation of 3:785
- quiescence/stemness and lineage determination 3:789–790
- hematopoietic stem cells (HSCs) 3:790
- neural stem cells (NSCs) 3:790–791
- plasticity of adult cell fate 3:791
- skeletal muscle stem cells 3:789–790
- stem region accessory domains 1:650
- STEP *see* Standard of expressed protein (STEP)
- sterile transcription *see* germline transcription
- sterile20-like kinase (Slik) 4:231–232
- steroidogenic acute regulatory protein (STAR protein) 1:178
- steroids 1:178–179, 1:178F
- steroid-responsive promoters 1:513–514
- synthesis 2:254
- sterol 1:202, 1:208, 1:209, 1:210, 1:213–214, 1:214, 2:273, 2:276
- lipids 1:155
- sterol esters (SE) 2:273
- sterol response element binding protein (SREBP) 1:175–176, 1:176F, 3:246
- steryl glycosides *see* glycosylated sterols
- stimulated emission depletion (STED) microscopy 2:2, 2:56, 2:76, 2:102
- in axial direction 2:80
- biological applications
- to biological research 2:82–84
- biological sample preparation 2:81–82
- commercial microscopes 2:81
- CW STED microscopy 2:78–79
- fluorophore low-lying electronic states 2:76F
- multicolor STED microscopy 2:80
- photobleaching 2:80–81
- principles 2:77, 2:77F
- pulsed STED microscopy 2:78
- resonant scanning 2:80–81
- time-gated STED microscopy 2:79–80
- two-photon STED microscopy 2:80
- stimulator of interferon genes (STING) 3:710
- sTNF *see* soluble TNF (sTNF)
- stochastic data modeling and analysis
- modeling single-cell transcriptomic data 4:223
- single-cell signaling networks modeling and analysis 4:222–223
- stochastic differential equations 4:74
- stochastic emission density 2:121
- stochastic optical reconstruction microscopy (STORM) 2:56, 2:58F, 2:62, 2:70, 2:76, 2:121–122, 2:123–124, 2:548, 2:549
- stochastic simulation algorithm (SSA) 4:82–83
- stochasticity 4:82
- stoichiometric networks 4:24–25
- stoichiometry matrix 4:25–26, 4:26
- stomata 2:329
- store-operated channels (SOCs) 3:163–164
- STORM *see* stochastic optical reconstruction microscopy (STORM)
- strains of prions 2:842
- (+)strand genomes 2:819
- (+)strand RNA viruses 1:520, 1:523–524, 2:819
- convoluted membranes 2:820–822
- with DNA intermediates 1:521–523
- host cell contribution 2:822–824
- membrane tubes and DMVs 2:822
- membrane webs 2:820–822
- spherules 2:819–820
- (-)strand RNA viruses 1:520–521, 2:824
- strand coordination 3:459
- Streptococcus pneumoniae*
inflammatory response to 4:98
- Streptomyces hygroscopicus* 3:3
- stress detection
- by nucleolus 3:270
- nucleolar RNAs 3:272
- RiBi factors 3:272
- RPBs 3:270–272, 3:271T
- to ribosome 3:276
- mRNA quality control 3:277
- protein quality control 3:276–277
- stress granules (SGs) 3:247–248
- mTORC1 regulation by 3:247–248
- stress signaling to ribosome 3:276F
- cap-dependent translation 3:275
- regulation of elongation 3:275–276
- stress-activated protein kinase-interacting protein 1 (SIN1) 3:132
- striatum 1:115, 1:117, 1:118
- stroma 2:697–698, 3:621
- stromal cells 3:616–623, 3:616
- lymphoid tissue structure 3:617F
- regulation of immune responses
- in induced TLOs 3:619
- inflammatory diseases 3:620–621
- initiation of tissue inflammation 3:619–620
- targeting stroma 3:621

- regulation of lymphoid tissue structure and function 3:617F
- initiation and regulation of adaptive immune responses 3:617–618
- regulation of lymphocyte migration 3:616–617
- regulation of peripheral tolerance 3:618–619
- SLOs 3:616
- strong anion exchange (SAX) 4:180
- strong cation exchange (SCX) 4:179–180
- structural biology 1:137
- structural classification of proteins (SCOP) 1:94–95
- structural maintenance of chromosomes (SMC) 2:593
- structural matrices, hyaluronan's role in 1:282
- structural proteins 2:810
- structural variations (SVs) 3:562–563
- structure-based screening strategy 1:797
- structured illumination microscopy (SIM) 2:2, 2:76, 2:86, 2:86–98
 - applications 2:91
 - bacteria structural and functional organization 2:93–94
 - centrosome structure 2:91–92
 - ciliogenesis 2:92
 - cytoskeleton structure 2:92–93
 - DNA-protein assemblies 2:91
 - motor proteins organization 2:92–93
 - nuclear, chromatin topography 2:91
 - organelles structural analysis 2:92
 - protein domain mapping 2:91–92
 - cell biology 2:86
 - image alignment and averaging 2:95–97
 - imaging artifacts 2:90
 - image drift 2:90
 - low signal to noise 2:90
 - pixel negative values 2:90–91
 - refractive index mismatch 2:90
 - for live cell imaging 2:947
 - live SIM 2:94–95
 - microscope hardware, software, and data acquisition 2:89
 - band separation/unmixing 2:89–90
 - image preprocessing 2:89
 - inverse Wiener filtering and apodization 2:90
 - structured illumination pattern generation 2:89
 - and molecular architecture analysis 2:92F
 - multifocal SIM 2:95F
 - NLSIM 2:95
 - principles 2:86, 2:87F
 - optical sectioning 2:89
 - resolution increasing 2:86–89
 - 3DSIM microscope and data acquisition 2:88F
 - sub-diffraction fluorescence microscopy methods 2:86
 - 3DSIM volume alignment and averaging 2:93F
- structured illumination pattern generation 2:89
- structure-specific endonuclease 1:439
- STx *see* Shiga toxin (STx)
- 4SU *see* 4-thiouridine (4SU)
- subcapsular macrophages (SCMs) 3:761, 3:765–766
- subcapsular sinus (SCS) 3:616, 3:765
- subcellular imaging
 - imaging subcellular structures *in vivo* 2:140–143, 2:140T
 - subcellular organelles 2:141F, 2:142F
- subcellular structures and organelles 2:1
- sub-diffraction fluorescence microscopy 2:86
- sub-distal appendages 2:649, 2:650F, 2:653
- suberoylanilide hydroxamic acid (SAHA) 1:797, 3:790–791
- subgranular zone (SGZ) 3:798, 3:798F, 4:134
 - interactions 3:799–800
 - neural stem cell niche 3:799
- subgraphs 4:212, 4:213, 4:214
- submitochondrial locations 1:237
- subsites 1:569
- 'substrate envelope' 1:744
- substrate-reduction therapy
 - for lysosomal diseases 1:778, 1:779–780, 1:783, 1:783
- substrates 1:683–684
- subtilase 1:555–556, 1:556
- subtiligase 1:573–574, 1:576, 1:577F
- subtiligase biotinylation 1:573–574, 1:577F
 - positive selection of N-terminal peptides by 1:576
- subtilisin-like pro-protein convertase (SPC) 3:39
- subtomogram averaging 2:22, 2:26, 2:27–28, 2:27, 2:28
- subtomograms 2:22, 2:25–26, 2:27–28
- subventricular zone (SVZ) 2:728, 3:798
 - interactions 3:799–800
 - V-SVZ neural stem cell niche 3:798–799
- such as galactose (Gal) 2:347
- SUE *see* septin unique element (SUE)
- SUFU *see* suppressor of fused (SUFU)
- sugar(s) 2:328
 - permease 1:101
- sulfatase modifying factor (SUMF) 1:774, 1:782
- sulfation 1:555, 1:563
- sulfhydryl proteases, evolution of affinity labels for 1:130F
- sulfopropyl (SP) 1:71
- sulfoquinovose (SQ) 1:181
- sulfur assimilation 2:329
- sulfur-transfer reagents 1:39–40, 1:40F
- SUMF *see* sulfatase modifying factor (SUMF)
- summation theorems 4:28, 4:28–29
- SUMO (small ubiquitin-like modifier)
 - family members 1:584T, 1:586–587, 1:586
 - structures of 1:583F
- SUMO protein *see* small ubiquitin-like modifier protein (SUMO protein)
- SUMOylation 1:507, 1:586, 1:586–587, 1:587, 3:785
- sumoylation 2:574
- Sun chloroplasts
 - shade chloroplasts vs. 2:329–330
- super cervical ganglion clone 10 (SCG10) 3:237–238
- super elongation complex (SEC) 1:483
- super resolution 2:102
- super resolution fluorescence localization microscopy 2:62–63, 2:70
 - data analysis 2:63, 2:73
 - discussion 2:63
 - experimental methods 2:62–63, 2:70–73
- super-antigens (SAGs) 3:642
- supercoiling 1:484
- supercritical fluid chromatography (SFC) 4:205
- super-enhancer 1:488
- superfamily 1:94
- superoxide anion 3:702
- superoxide dismutases (SODs) 1:116, 3:336
- super-resolution imaging 2:56, 2:58F
- super-resolution light microscopy 2:76
 - see also* stimulated emission depletion (STED) microscopy
- super-resolution microscopy 2:86
- super-resolution optical fluctuation imaging (SOFI) 2:76
- supported lipid bilayers (SLB) 3:678
- suppression
 - factor 2:77–78
 - laser interference fringes 2:66–67
 - out-of-focus signals 2:113
- suppressor of cAR (SCAR) 2:493
- suppressor of cytokine signaling (SOCS) 3:32
 - 3:32
- suppressor of fused (SUFU) 3:76–77
- suppressor of hairless 3:107
- suppressors of cytokine signaling proteins (SOCS proteins) 3:136, 3:144
 - CIS 3:145
 - in human disease 3:146–147
 - SOCS1 3:144
 - direct kinase inhibition by 3:144–145
 - SOCS2 3:145
 - SOCS3 3:144
 - SOCS4 3:145
 - SOCS5 3:145
- supracellular properties 3:226–227
- supramolecular activation clusters (SMACs) 3:522–524
- surface plasmon resonance (SPR) 4:192
- surface plasmon resonance imaging (SPRI) measurements 1:50
- surface topography 3:193–194
- SURS *see* systematic uniform random sampling (SURS)
- survival motor neuron (SMN) 3:394
 - SMN complex 1:504F, 1:507
- survivin 1:709, 3:394–395, 3:469, 3:469–470
- survivin antisense oligonucleotides 3:396
- SV2 *see* synaptic vesicle glycoprotein 2 (SV2)
- sVEGFR1 *see* splice variants of VEGFR1 (sVEGFR1)
- sVEGFR1/sFlt1 *see* soluble VEGFR1 (sVEGFR1/sFlt1)

sVEGFR2 *see* splice variants of VEGFR2 (sVEGFR2)
 SVPs *see* synaptic vesicle precursors (SVPs)
 SVs *see* structural variations (SVs); synaptic vesicles (SVs)
 SVV *see* small vessels vasculitis (SVV)
 SVZ *see* sub-ventricular zone (SVZ)
 SWATH-MS approach 4:168–169
 sweat test 1:257, 1:259–260
 sweating, in cystic fibrosis 1:257
 SWI/SNF chromatin remodeling complexes 3:787
 SWI/SNF complex 3:787, 3:790
 switch 1 and 2 2:731
 switching model 2:628–629
 syk *see* spleen tyrosine kinase (syk)
 synapse 1:229–230, 1:244–245, 1:244, 1:246, 3:607
 chemical transmission at 1:245–246
 synapsin 2:297, 2:298
 synaptic binding 3:660
 synaptic cleft, fate of neurotransmitter in 1:248
 synaptic potential 1:218
 synaptic transmission 2:381
 synaptic vesicle glycoprotein 2 (SV2) 2:296
 synaptic vesicle precursors (SVPs) 2:611
 synaptic vesicles (SVs) 1:202, 1:203, 1:205, 2:2–3, 2:293, 2:294–296, 2:380, 2:380–381
 key proteins associated with 2:296
 proteins mediating endocytosis 2:296
 proteins mediating exocytosis 2:296
 proteins regulating neurotransmitter packaging 2:297
 life cycle 2:297
 docking at active zone and priming fusion machinery 2:297–298
 endocytosis and reformation of fusion-competent synaptic vesicle 2:298–299
 formation in ER/Golgi and trafficking 2:297
 fusion with plasma membrane and release 2:298
 loading with neurotransmitter 2:297
 physical characteristics 2:295–296
 synaptobrevin 1:248, 2:298
 synaptotagmin-1 2:505
 synaptophysin 2:297
 synaptosome-associated protein of 23 kDa (SNAP23) 2:241
 synaptosomes 2:293, 2:293–294
 proteomic and cell biological analysis of 2:294
 synaptotagmin 1:248
 synaptotagmin function model 2:461F
 synaptotagmin-like proteins 2a (Slp2-a) 2:745–746
 synaptotagmin-like proteins 4a (Slp4-a) 2:745–746
 synchrotron-based X-ray fluorescence imaging 1:128
 syndapin 2:493
 syndapin I knockout mice phenocopy 2:498–499

synexpression groups 4:48
 syntaxin 1:248, 2:298
 syntaxin 1 2:459, 2:460, 2:460–461, 2:461–462
 synthesis-dependent strand annealing pathway (SDSA pathway) 1:459, 1:461–462
 synthesis phase (S phase) 3:458–468
 CDKs 3:458
 DNA damage response 3:460–461
 DNA replication 3:458–459
 Myc control 3:458
 synthetic lethality 2:528
 systematic evolution of ligands by exponential enrichment (SELEX) 1:26–27, 1:375, 1:375–377
 systematic uniform random sampling (SURS) 2:41
 systemic lupus erythematosus (SLE) 1:776–777, 3:710, 3:747–748, 3:753–754, 3:756, 3:780
 systems biology 3:678, 4:1, 4:2, 4:72, 4:73F, 4:95–96, 4:166
 application for high-throughput proteomics 4:169–170
 exemplary biological processes 4:172
 cell cycle regulation 4:172
 central metabolism 4:172–173
 EGFR signaling pathway 4:172
 mass spectrometry based proteomics
 correlation of proteome with genome 4:171
 correlation of proteome with phenotype 4:171
 dynamic proteomes measurement 4:171–172
 mapping nodes of proteome 4:170
 organization of proteome 4:170–171
 systems biology markup language (SBML) 4:75–76
 systems cell biology 4:1–3
 dynamic integration 4:2
 horizontal integration 4:3
 three dimensions of biological systems complexity 4:1F
 vertical integration 4:2–3

T

T cell antigen receptor (TCR) 3:650, 3:522, 3:660
 see also immunological synapses
 clustering 3:655–656
 complex 3:650F
 as mechanoreceptors 3:655
 triggering
 changes in intracellular domains conformation and accessibility 3:653–655, 3:654F
 cognate pMHC ligands quantity and valency 3:656–657
 exclusion of negative regulators 3:651–653
 recruitment of modulators 3:651

time of interaction 3:651
 zipper model 3:652F
 T cell receptor (TCR) 1:214, 2:211–212, 3:541, 3:568, 3:573–574, 3:605, 3:634, 3:677, 3:696, 3:770
 T cell(s) 3:677–686, 3:677
 activation
 conformational model 3:533–534
 kinetic model 3:531–532
 anergy 3:577
 NFAT and tolerance 3:575, 3:576F
 T cell anergy 3:577
 thymocyte development and positive selection 3:575–576
 Treg 3:576–577
 NFAT functions 3:575
 regulatory neutrophils 3:754
 response controlling quality 3:747
 shape 3:678–680
 signaling
 activating 3:677
 function 3:683
 methods to studying organization 3:677–678
 variation 3:682–683
 signaling organization and structures 3:678, 3:679F, 3:682
 cytoskeletal dynamics 3:680–681
 plasma membrane topology 3:678–680
 signaling complexes 3:681–682
 T cell shape 3:678–680
 vesicle trafficking 3:680
 targeting 1:145
 tolerance 3:575–576
 T central memory (Tcm) 3:770
 T effector memory (Tem) 3:770
 T follicular helper cells (Tfh cells) 3:541–549, 3:541, 3:618
 cell activation, development, and GC response 3:545F
 GC 3:543–547
 in health and disease 3:547
 initiation of Tfh development 3:541–543
 interactions at T:B border 3:544F
 provision of help to GC B cells 3:546F
 stabilization of Tfh cell development 3:543
 T cell priming and activation 3:542F
 T follicular regulatory T cells (Tfr cells) 3:545
 T helper cells (Th cells) 3:541, 3:567
 Th1 cells 3:619, 3:721–722
 Th17 cells 2:800
 Th2 cells 3:26, 3:575
 T lymphocyte(s) 2:828, 3:634
 mechanisms for adhesion and diapedesis
 across endothelia
 cytoskeletal remodeling 3:637
 molecular basis for adhesion and diapedesis 3:636–637
 remodeling 3:637–639
 stages of lymphocyte trafficking 3:635F
 trafficking
 during immune responses 3:636
 during immune surveillance 3:635–636
 during maturation 3:634–635

- T2D *see* type 2 diabetes (T2D)
T2DM *see* Type 2 diabetes mellitus (T2DM)
3T3-L1 cells 2:274
T3SS *see* type III secretion system (T3SS/TTSS)
T4SS *see* type IV secretion system (T4SS)
TA sequence *see* tail anchor sequence (TA sequence)
TAB2 *see* TAK1-binding protein 2 (TAB2)
TACE *see* TNF-converting enzyme (TACE); TNF α convertase (TACE)
T-acute lymphoblastic leukemia (T-ALL) 3:110, 3:562–563, 4:255
TAD *see* transactivating domain (TAD); transcriptional activation domain (TAD)
TADG-12 *see* TMPPRS3
Tafazzin (TAZ) 3:101–102
TAFI *see* thrombin activatable fibrinolysis inhibitor (TAFI)
TAG *see* triacylglycerol (TAG)
tail anchor sequence (TA sequence) 1:597–598
tail homology domains (TH domains) 2:601
tail interacting protein of 47 kDa (TIIP 47) 2:435
TAILS *see* terminal amine isotopic labeling of substrates (TAILS)
TAK1 *see* transforming growth factor β -activated kinase1 (TAK1)
TAK1-binding protein 2 (TAB2) 3:357–358
TALENs *see* transcription activator-like effector nucleases (TALENs)
T-ALL *see* T-acute lymphoblastic leukemia (T-ALL)
TALS *see* Taybi-Linder syndrome (TALS)
TAMs *see* tumor-associated macrophages (TAMs)
tandem affinity purification (TAP) 3:299
tandem affinity purification-mass spectrometry (TAP-MS) 4:190
tandem mass spectrometry 1:150–151
tandem mass tags (TMT) 4:169
tangential migration 2:728
Tangier disease 1:176, 1:178
TANK-binding kinase 1 (TBK1) 3:718–719
Tankyrases 3:91
TAP *see* tandem affinity purification (TAP)
TAP1 *see* transporter associated with antigen processing 1 (TAP1)
Tapasin (TAP-binding protein) 3:589, 3:593–594
TAP-MS *see* tandem affinity purification-mass spectrometry (TAP-MS)
Taq DNA polymerase 1:125
TAR *see* transactivation response element (TAR)
TARGET *see* Therapeutically Applicable Research To Generate Effective Treatments (TARGET)
target membrane-SNAREs (t-SNAREs) 2:459, 2:486
target of rapamycin (TOR) 1:618, 1:765, 3:3, 3:263
TOR signaling (TOS) 3:243
target of rapamycin complex 1 (TORC1) 3:404
targeted mRNA degradation 3:317
combined action of Short RNAs and RBPs 3:323–324
in gene regulation 3:321
main pathways 3:318F
mRNA
degradation control by short RNAs 3:321–322
surveillance mechanisms 3:317, 3:319F
by RBPs 3:322–323
substrate choice
NGD 3:318
NMD 3:317–318
NSD 3:318–319
substrate fate 3:321
substrate recognition
NGD 3:320–321
NMD 3:319–320
NSD 3:321
targeted therapy
for lysosomal diseases 1:763, 1:781
targeted tumor therapy
retrograde transport and 2:438
targeted/target 1:136
data acquisition 4:168
protein degradation 1:88
TargetScan 1:365
tartrate-resistant acid phosphatase (TRAP) 1:776, 1:778
TAT *see* transactivator of transcription (TAT); Twin arginine translocation (TAT)
TATA-binding protein (TBP) 1:477–478, 1:481
Taybi-Linder syndrome (TALS) 1:508
Tay-Sachs disease 1:767, 1:774, 1:782–783, 1:784
TAZ *see* Tafazzin (TAZ)
tazarotene 1:706
tazarotene-induced gene 1 (TIG1) 1:707
TBDMS group *see* tert-butyldimethylsilyl (TBDMS) group
TBI *see* traumatic brain injury (TBI)
TBK1 *see* TANK-binding kinase 1 (TBK1)
T-box transcription factor 5 (Tbx5) 3:102
TBP *see* TATA-binding protein (TBP)
Tbx5 *see* T-box transcription factor 5 (Tbx5)
TC *see* tip cell (TC)
TCA cycle *see* tricarboxylic acid cycle (TCA cycle)
TCAB1 *see* Telomerase Cajal body protein 1 (TCAB1)
TcdB *see* *C. difficile* toxin B (TcdB)
TCDD *see* tetrachlorodibenzodioxin (TCDD)
T-cell immunoglobulin (TIMs) 3:734
T-cell receptor (TCR) 3:528, 3:583
 $\alpha\beta$ TCR 3:530F
ligands
EGF receptor 3:529F
ITAM motifs in multiple-chain immunoreceptors 3:530F
properties 3:530–531
subunit composition, stoichiometry, and valency 3:529–530
tyrosine kinase-coupled receptor 3:528–529
signalosome 3:528
biochemical module and assembly 3:534
modular organization 3:535F
signaling code 3:531–532
thermodynamic and kinetic parameters 3:531F
TCF 3:94–95, 3:94F, 3:95
 β -Catenin/TCF complex 3:93–94
TCGA *see* The Cancer Genome Atlas (TCGA)
Tcm *see* T central memory (Tcm)
TCR *see* T cell antigen receptor (TCR); T cell receptor (TCR); T-cell receptor (TCR)
TCs *see* transport complexes (TCs)
TCSPC *see* time-correlated single photon counting (TCSPC)
TD-60 *see* telophase disk protein 60 (TD-60)
TDE *see* 2,2'-thiodiethanol (TDE)
TDG *see* Thymine DNA glycosylase (TDG)
TDs *see* topological domains (TDs)
TdT *see* terminal deoxynucleotidyl transferase (TdT)
TEA domain proteins (TEADs) 3:102
Teashirt (Tsh) 3:99
Tec kinase 2:520–522, 2:522–523
TEFs *see* transcription enhancer factors (TEFs)
TEK *see* Tie2 receptor
telomerase 1:420–421, 1:422, 1:423, 1:430
Cajal body (sca) RNAs 1:422
mammalian Shelterin complex 1:421, 1:421F
RNA 1:422–423
telomerase Cajal body protein 1 (TCAB1) 1:422
telomerase reverse transcriptase (TERT) 1:420–421, 1:421–422
Telomerase Cajal body protein 1 (TCAB1) 1:422
telomerase reverse transcriptase (TERT) 1:420–421, 1:421–422, 1:421–422
telomerase reverse transcriptase protein (TERT) 1:430
telomerase RNA component (TERC) 1:430
telomere repeat-containing RNA (TERRA) 1:430
telomere sister chromatid exchange (T-SCE) 1:431
telomere(s) 1:2–3, 1:418, 1:426, 3:512, 3:513
alternative lengthening 1:423
crisis 1:431
DNA 1:427
end-replication problem 1:418–419, 1:419F
DNA polymerases synthesize 1:418–419
length/shortening rate 1:419
molecular events 1:426–427
physiology
noncancer telomere-related diseases 1:432–433
senescence and telomere crisis 1:431
telomeres in cancer cells 1:431–432

- telomere(s) (*continued*)
 Protection Of Telomeres 1 (POT1) 1:420
 Shelterin protein complex 1:419F, 1:420
 replication and length maintenance 1:428
 recombination at telomeres and ALT
 1:430–431
 semiconservative telomere replication
 1:428–430
 telomerase 1:430
 shortening 3:777
 structure and proteins
 proteins associated with telomere
 1:427–428
 telomere DNA 1:427
 telomere secondary structure 1:427
 syndromes of short telomeres 1:423
 telomeric DNA 1:419
 telomeric proteins 1:419–420
 telomeric repeat-binding factors 1 and 2
 1:420
 telomeric repeat-containing RNA (TERRA)
 1:420
 telomeric repeat-containing RNA 1:420
 in yeast 1:420
 telomeric repeat-binding factor 1 (TRF1)
 1:420, 1:427
 telomeric repeat-binding factor 2 (TRF2)
 1:420, 1:427
 telomeric repeat-binding factor 2-interacting
 protein 1 (TERF2IP) 1:427
 telophase 3:481
 CPC during 3:475
 telophase disk protein 60 (TD-60) 3:469
 Tem *see* T effector memory (Tem)
 TEM *see* transmission electron microscope
 (TEM)
 template
 DNA synthesis
 template-dependent 1:451
 template-independent 1:451
 strand 1:471, 1:472
 temporal regulation of transgenes
 1:513–514
 temporal sensing 4:5–6
 tenascin C (TNC) 3:63–64
 tener-taxis 3:637–638
 TERC *see* telomerase RNA component
 (TERC)
 TERF1-interacting factor 2 (TINF2) 1:427
 TERF2IP *see* telomeric repeat-binding factor
 2-interacting protein 1 (TERF2IP)
 terminal amine isotopic labeling of
 substrates (TAILS) 1:569, 1:573–574,
 1:575F, 1:578, 1:579
 negative selection of N-terminal peptides
 by 1:575–576
 terminal deoxynucleotidyl transferase (TdT)
 1:451, 3:550
 terminal oligopyrimidine (TOP) 3:263,
 3:282
 termination codons 1:310, 1:324, 1:531
 termination–reinitiation 1:522F, 1:523–524
 terminome 1:569
 terminomics 1:569
 TERRA *see* telomere repeat-containing RNA
 (TERRA)
 TERT *see* telomerase reverse transcriptase
 (TERT); telomerase reverse
 transcriptase protein (TERT)
 tert-butylidimethylsilyl (TBDMS) group
 1:42–43, 1:44
 tertiary granules 2:770, 3:751
 tertiary lymphoid organs (TLOs) 3:616,
 3:619
 stromal cells in induced 3:619
 testisin 1:654
 testosterone 1:178–179
 tetanus neurotoxin 2:796
 tether repeats (TRs) 2:335–336
 tetherin restriction factor 2:836
 tethering 1:551F, 1:765, 2:412, 2:436, 2:437,
 2:485, 3:751
 tethers 2:171, 2:355
 TetR *see* tetracycline repressor (TetR)
 tetrachlorodibenzodioxin (TCDD) 2:276
 tetracycline repressor (TetR) 1:513–514
Tetrahymena thermophila 1:427
 telomeric DNA 1:419
 tetrasome 2:311
 tetraspannin protein (TSPAN12) 3:89,
 3:89F
 tetratricopeptide repeat domain protein 19
 (TTC19) 2:195
 tetratricopeptide repeats (TPRs) 2:667
 tetratricopeptide-rich regions (TPR) 3:704
 text-mining approaches 4:248
 TF *see* tissue factor (TF); transcription
 factor (TF)
 Tf *see* transferrin (Tf)
 TFA *see* Trifluoroacetic acid (TFA)
 TFEB *see* transcription factor EB (TFEB)
 Tfh cells *see* T follicular helper cells
 (Tfh cells)
 Tf-HRP *see* transferrin coupled to
 horseradish peroxidase (Tf-HRP)
 TFIID *see* transcription factor IID (TFIID)
 TFIIF *see* transcription-factor IIF (TFIIF)
 TFPI *see* tissue factor pathway inhibitor
 (TFPI)
 Tfr cells *see* T follicular regulatory T cells
 (Tfr cells)
 Tfr *see* transferrin receptor (Tfr)
 TGF *see* transforming growth factor (TGF)
 TGF α *see* transforming growth factor α
 (TGF α)
 TGF β *see* transforming growth factor β
 (TGF β)
 TGF β R *see* transforming growth factor beta
 receptor (TGF β R)
 TGN *see* trans-Golgi network (TGN)
 TGN-derived constitutive transport 2:363
 Th cells *see* T helper cells (Th cells)
 TH domains *see* tail homology domains
 (TH domains)
 Th17 *see* 17-producing helper T cell (Th17)
 THD *see* TNF homology domain (THD)
 The Cancer Genome Atlas (TCGA) 3:449
 theory-driven approach *see* top-
 down—approach
 Therapeutically Applicable Research To
 Generate Effective Treatments
 (TARGET) 4:254
Thermococcus celericrescens DNA polymerase
 1:125
 thermodynamics 2:552–553, 4:38, 4:42
Thermophilus acidophilus 1:589
Thermotoga maritime 1:125
 THG *see* third harmonic generation (THG)
 thiamine pyrophosphate (TPP) 1:375–377,
 1:377, 1:379F
 2,2'-thiodiethanol (TDE) 2:81–82
 thioredoxin-interacting protein expression
 (TXNIP expression) 2:532
 4-thiouridine (4SU) 4:192
 third harmonic generation (THG)
 2:129–131
 Thomsen's disease 1:231
 3D image analysis
 tomogram reconstruction and 2:25–26
 three-dimensional (3D) gradient devices
 3:671
 three-dimensional imaging 2:72–73
 3D reconstruction 2:22, 2:23F, 2:27,
 2:27–28
 3D SEM *see* three-dimensional scanning
 electron microscopy (3D SEM)
 3D SIM microscope 2:86
 three-dimensional (3D) surface imaging 2:1
 three-dimensional imaging 2:64
 three-dimensional scanning electron
 microscopy (3D SEM) 2:44
 data handling and segmentation 2:47–49,
 2:48F
 imaging block face 2:45F
 new methods and future developments
 2:49
 orientation and location of region of
 interest 2:44–45
 sample preparation 2:44
 techniques 2:45–47
 three prime repair exonuclease 1 (TREX1)
 3:720
 threonine proteases, proteasomes 1:757
 threonine-serine-serine (TSS) 3:469
 threshold potential 1:244, 1:245–246,
 1:246–247
 thrombin activatable fibrinolysis inhibitor
 (TAFI) 2:227–228
 thrombocytopenias 1:717
 thrombomodulin (TM) 1:717
 thrombophilia 1:707
 thrombopoietin (TPO) 1:714, 3:23–24
 thrombosis 3:196
 aging 1:719–720
 arterial thrombosis 1:718
 DVT 1:718–719
 Factor V Leiden mutation 1:719
 PAI-1 1:719
 prothrombin G20210A 1:719
 thrombospondin type 1 repeats (TSRs)
 1:630
 thrombospondin-1 (TSP1) 1:632–633
 thrombotic thrombocytopenic purpura
 (TTP) 1:633–634, 1:717
 thromboxane A2 (TxA2) 3:196
 thymic stromal lymphopoietin (TSLP)
 3:24–25, 3:747
 thymine DNA glycosylase (TDG) 1:444

- thymocyte development and positive selection 3:575–576
- thymocytes 3:525
- thymoproteasome 1:591
- thyroidtranscriptionfactor-1 (TTF-1) 3:102
- TIC *see* trauma-induced coagulopathy (TIC); tumor initiating cell (TIC)
- tie 4:102
- receptor 4:107
- Tied Diffusion of Interacting Events (TieDie) 4:212T, 4:214, 4:215
- TieDie *see* Tied Diffusion of Interacting Events (TieDie)
- TIG1 *see* Tazarotene-induced gene 1 (TIG1)
- tight junctions (TJs) 2:742, 2:744–745, 4:111
- tilt-pair validation technique 2:7
- TIM *see* TRAF-interacting motif (TIM)
- time resolution 2:113
- time-correlated single photon counting (TCSPC) 2:79–80, 2:110
- time-domain
- data 2:116–117
- wide-field FLIM 2:115–116
- time-lapse imaging 2:55
- time-of-flight (ToF) 1:152
- TIMPs *see* tissue inhibitors of metalloproteinases (TIMPs)
- TIMs *see* T-cell Immunoglobulin (TIMs)
- TINF2 *see* TERF1-interacting factor 2 (TINF2)
- tip cell (TC) 4:102, 4:111–112, 4:112
- tip-stalk cell signaling 4:112F
- + TIPs *see* plus end tracking proteins (+ TIPs)
- tip taper 4:36, 4:38
- tipranavir 1:743
- TIRFM *see* total internal reflection fluorescence microscopy (TIRFM)
- TIR domain-containing adaptor protein (TIRAP) *see* MYD88-adaptor-like protein (MAL)
- KLK 1:699–700
- LN macrophages 3:765
- lymph node compartments and localization of macrophage populations 3:765F
- medullary macrophages 3:766–767
- SCMs 3:765–766
- macrophage subsets in spleen 3:759, 3:760F
- apoptotic cells clearance by MZ macrophages 3:763
- metallophilic macrophages 3:761–762
- RPMs 3:763–765
- SIGN-R1/MARCO macrophages 3:762–763
- splenic MZ and macrophages 3:759–761
- microfibrils 1:634–635
- remodeling phase 3:220
- scaffold 2:717–718
- tissue-imaging mass spectrometry 1:152–153, 1:157
- tissue-resident macrophages, specialized subsets of 3:759
- tissue-specific diseases 2:575–576
- tissue–resident memory 3:568–569
- TIR-domain-containing adapter-inducing interferon- β (TRIF) 3:520–521
- TIRF mode *see* total internal reflection fluorescence mode (TIRF mode)
- tissue damage and wound healing 4:99
- tissue distribution 1:671
- tissue factor (TF) 1:716
- tissue factor pathway inhibitor (TFPI) 1:717
- tissue formation, phase of new granulation tissue formation 3:219–220
- skin wounds reepithelialization 3:219
- tissue inflammation initiation 3:619–620
- tissue inhibitors of metalloproteinases (TIMPs) 1:626–627, 1:711
- biological functions 1:712
- disease aspects 1:712–713
- interactions with pro-MMPs 1:712
- mechanism of inhibition of MMPs 1:711–712
- structures 1:711
- TIMP-3 1:633
- tissue repair, senescence and 3:515
- tissue-resident memory T cells (T_{RM}) 3:568–569
- tissue-type plasminogen activator (tPA) 1:717
- titanium dioxide (TiO₂) 4:178
- TI-VAMP/VAMP7 2:461–463
- TJs *see* tight junctions (TJs)
- t-Loop 1:427, 1:427–428, 1:430
- TLOs *see* tertiary lymphoid organs (TLOs)
- TLP1 *see* Tubby-likeProtein-1 (TLP1)
- TLR *see* toll-like receptors (TLR)
- TLS *see* translesion synthesis (TLS)
- TLSP receptors 3:26–27
- TLSS *see* 3'-tRNA-like structures (TLSS)
- TM *see* thrombomodulin (TM)
- 7TM *see* 7-pass transmembrane (7TM)
- TM protein *see* transmembrane protein (TM protein)
- TM region *see* transmembrane region (TM region)
- TM segments *see* transmembrane segments (TM segments)
- TMD *see* trans-membrane domain (TMD)
- TME *see* transmissible mink encephalopathy (TME)
- TMOD3 *see* Tropomodulin-3 (TMOD3)
- TMPRSS *see* transmembrane protease, serine subfamily (TMPRSS)
- TMPRSS1 *see* Hepsin
- TMPRSS2 1:655–656
- TMPRSS3 1:655–656
- TMPRSS5 1:655–656
- TMPRSS6 *see* Matriptase-2
- TMPRSS10 *see* Corin Subfamily
- TMPRSS11A *see* HAT like-1 (HATL-1)
- TMPRSS11B *see* HAT like-5 (HATL-5)
- TMPRSS11D *see* HAT protease
- TMPRSS11E *see* DESC-1
- TMPRSS11F *see* HAT like-4 (HATL-4)
- tmRNAs *see* transfer messenger RNA (tmRNAs)
- TMS *see* trimethylsilyl group (TMS)
- TMT *see* tandem mass tags (TMT)
- TMV *see* tobacco mosaic virus (TMV)
- TNC *see* tenascin C (TNC)
- TNF *see* tumor necrosis factor (TNF)
- TNF homology domain (THD) 3:61, 3:355
- TNF receptor-associated periodic syndrome (TRAPS) 3:71–72
- TNF superfamily (TNFSF) 3:61, 3:354–355
- clinical targeting of members 3:63–64
- TNF α convertase (TACE) 1:639–640
- TNF α -stimulated gene 6 (TSG6) 1:282
- TNF-converting enzyme (TACE) 3:355
- TNF-like receptor apoptosis mediating protein (TRAMP) 3:62
- TNF-therapy 4:98–99
- TNFR *see* tumor necrosis factor receptor (TNFR); tumor necrosis factors receptors (TNFR)
- TNFR1 superfamily *see* tumor necrosis factor receptor-1 (TNFR1) superfamily
- TNF-related apoptosis-inducing ligand (TRAIL) 3:332
- TNF-related apoptosis-inducing ligand receptor 1 (TRAIL-R1) 1:712, 3:62
- TNFRSF *see* tumor necrosis factor receptor super family (TNFRSF)
- TNFs *see* tumor necrosis factors (TNFs)
- TNFSF *see* TNF superfamily (TNFSF)
- tobacco mosaic virus (TMV) 1:328–329, 1:520
- Toca-1 *see* transducer of Cdc42-dependent actin assembly (Toca-1)
- ToF *see* time-of-flight (ToF)
- togaviridae, spherules producing by 2:819–820
- Tokuyasu technique 2:34
- TOL-1 *see* Toll-like receptor (TOL-1)
- tolerance, DCs in 3:747
- tolerance maintenance/promotion 3:641
- toll-like receptor (TOL-1) 2:846
- toll-like receptors (TLR) 1:198, 1:368, 1:708–709, 2:303–304, 2:752, 2:846, 3:28, 3:355–356, 3:520, 3:584, 3:588, 3:591–592, 3:592–593, 3:592, 3:593, 3:595, 3:619, 3:702, 3:710, 3:712–714, 3:712T, 3:733, 3:745–746, 3:752
- TLR4 1:694
- TOM group *see* triisopropylsilyloxymethyl (TOM) group
- TOM proteins 1:239T
- tombusviruses 2:255
- TOMM40 1:241
- tomogram reconstruction and 3D image analysis 2:25–26
- TOP *see* terminal oligopyrimidine (TOP)
- topoisomerases 1:396–397, 1:406F
- topological domains (TDs) 3:789
- TOR *see* target of rapamycin (TOR)
- TORC1 *see* target of rapamycin complex 1 (TORC1)
- total internal reflection fluorescence microscopy (TIRFM) 2:2, 2:58, 2:62, 2:63, 2:102, 2:123, 2:397, 2:553–554, 3:662, 4:15–16

- actual evanescent field depth and profile 2:67–68
- applications and combinations in cell biology
- calcium channels and transients 2:63
 - cell/substrate contact regions 2:62–63
 - cytoplasmic filaments 2:63
 - secretory granule tracking and exocytosis 2:63
 - TIR combined with fluorescence recovery 2:64–65
 - TIR combined with interference reflection contrast 2:65
 - TIR combined with polarization 2:63–64
- EPI vs. objective-based TIR fluorescence digital images 2:63F
- long-term fluorescence movies of cultured cells 2:68
- optical configurations and setup 2:65
- high aperture objective-based TIR 2:65
 - optical ray diagram for custom-built TIRF 2:66F
 - prism-based TIRF 2:65–66, 2:67F
- vs. other optical sectioning microscopies 2:68–69
- prism-based TIRF 2:65–66, 2:67F
- suppression of laser interference fringes 2:66–67
- theoretical basis 2:62
- TIR illumination 2:62F
- variable incidence angle TIR 2:68
- total internal reflection fluorescence mode (TIRF mode) 2:71
- ToxA *see* exotoxin A (ToxA)
- toxin synthesis by fungi 2:255
- TP53 *see* tumor suppressor transcription factor p53 (TP53)
- tPA *see* tissue-type plasminogen activator (tPA)
- TpMs *see* tricoplein/mitostatin (TpMs)
- TPO *see* thrombopoietin (TPO)
- TPP *see* thiamine pyrophosphate (TPP)
- TPR *see* tetratricopeptide-rich regions (TPR); tetratricopeptide repeats (TPRs)
- tracking 2:23–24
- TRADD *see* tumor necrosis factor receptor type 1-associated death domain (TRADD)
- trafficking
- altered 3:641
 - T lymphocyte
 - during immune responses 3:636
 - during immune surveillance 3:635–636
 - during maturation 3:634–635
- TRAF-interacting motif (TIM) 3:62
- TRAFs *see* tumor necrosis factor receptor-associated factors (TRAFs)
- TRAIL *see* TNF-related apoptosis-inducing ligand (TRAIL); tumor necrosis factor-related apoptosis ligand (TRAIL)
- TRAIL-R1 *see* TNF-related apoptosis-inducing ligand receptor 1 (TRAIL-R1)
- TRAM *see* TRIF-related adaptor molecule (TRAM)
- TRAMP *see* TNF-like receptor apoptosis mediating protein (TRAMP)
- trans*-acting factor 1:517
- in ribosome assembly 3:299
 - in translational initiation 1:551F
- transactivating domain (TAD) 3:107, 3:143
- transactivation response element (TAR) 2:835
- transactivation-response RNA-binding protein (TRBP) 1:354, 1:355–358, 1:356–357F, 1:357F, 1:358, 1:360
- transactivator of transcription (TAT) 2:394
- transcript decay 1:347
- transcription 2:220–221
- transcription, chromatin structure, and splicing 1:506
- transcription activator-like effector nucleases (TALENs) 1:515
- transcription enhancer factors (TEFs) 3:102
- transcription enhancers, distant activation of 1:488–494
- bacterial enhancer action 1:488–489
 - distant promoter activation accompanied by DNA looping 1:489–491
 - distant promoter activation using tracking mechanism 1:488–489
 - eukaryotes, enhancer action in 1:491–492
 - hopping/scanning mechanisms 1:488
 - mechanisms of action 1:489F
 - super-enhancers 1:488
- transcription factor (TF) 1:481–482, 3:580, 4:188–189
- transcription factor EB (TFEB) 1:617, 1:764, 1:765–766, 1:765, 1:766, 1:770–771, 1:770, 1:771, 1:771, 1:779, 1:785, 1:796, 2:203–204, 2:225, 3:246, 3:407
- mediating mTORC1 effect on 3:246
- transcription factor IID (TFIID) 1:492
- transcription factors 1:365, 1:506, 3:94, 3:125–126, 3:788–789, 3:790–791
- transcription start sites (TSS) 3:778
- transcription termination factor 1 (TTF1) 3:258
- transcriptional activation domain (TAD) 1:482
- transcriptional activity of Forkhead box O3 (FoxO3a) 3:130
- transcriptional control 3:553
- transcriptional regulation 3:413, 3:440
- DC development 3:743
- transcriptional start site (TSS) 1:481
- transcription-coupled DNA repair 1:437
- transcription-factor IIH (TFIIH) 1:437–438
- transcriptome 4:160
- transcriptomics 3:722, 4:160–165, 4:166
- high-throughput transcriptomic technologies evolution 4:160
 - α -hemolysin 4:162
 - fluorescent *in situ* hybridization 4:162
 - high-throughput DNA sequencing technology 4:161
 - next-generation sequencing platforms 4:161
 - PCR 4:162
 - RNA-seq 4:160–161, 4:161
 - SMRT 4:161
- interactive transcriptome exploration 4:162
- ceRNA 4:163
 - circular RNA 4:162–163
 - effects on transcription and mRNA 4:163
 - miRNA 4:162
 - piRNA 4:162
 - ribosomal and transfer RNA 4:162
 - RNA modifications 4:163
 - mRNA transcriptomics, new technologies for 4:163
 - with omics 4:163–164
- transcytosis 2:744F, 2:745–746
- transducer of Cdc42-dependent actin assembly (Toca-1) 2:492–493
- transendocytosis 3:609, 3:609–610, 3:610–611, 3:611, 3:611–612
- cell biology at the heart of immune regulation 3:610–611
 - cell biology of CTLA-4 3:609–610
 - as a mechanism of immune regulation 3:611
 - Treg suppression 3:611–612
- transendothelial migration 3:634
- transfection 1:555, 1:562
- transfer messenger RNA (tmRNAs) 1:310, 1:327–328
- transfer RNA (tRNA) 1:27, 1:309–340, 1:384, 1:525, 4:162
- aminoacylation 1:316–317, 1:526F
 - aminoacyl-tRNA synthetases 1:316–317
 - proofreading of 1:319–320
 - for cell wall synthesis 1:326–327
 - classes of 1:313, 1:314F
 - in different organisms 1:315T
 - degenerative, in mitochondria 1:315–316
 - and GCN2 pathway 1:325
 - and genetic code 1:310–313
 - and human diseases 1:329–334
 - MELAS/MERRF 1:334
 - mutations in 1:330–333T
 - point mutations 1:333, 1:333–334
 - identity element of 1:317–319
 - mimicry 1:329
 - modification, and human diseases 1:334
 - paralogous tRNA genes/tRNA isoacceptors 1:313–315
 - in protein synthesis 1:525–526
 - for regulated proteolysis 1:325–326
 - related disease 1:334–335
 - roles beyond translation 1:324–325
 - transfer messenger RNA (tmRNA) 1:327–328
 - in translation 1:320–323
 - elongation 1:323–324
 - initiation 1:321–323
 - ribosome recycling 1:324
 - termination 1:324
 - tRNA-like ribosome recycling factors 1:329
 - tRNA methyltransferase homolog 12 (TRMT12) 1:334, 1:334–335
 - viral tRNA-like molecules 1:328–329
 - for virus reverse transcription initiation 1:327
- transferrin (Tf) 2:192, 2:771

- transferrin coupled to horseradish peroxidase (Tf-HRP) 2:37–39
- transferrin receptor (TfR) 2:192, 2:788–789
- transforming growth factor (TGF) 3:357–358
- transforming growth factor α (TGF α) 1:641, 1:694
- transforming growth factor β (TGF β) 1:634–635, 2:469, 2:700, 3:1–2, 3:37, 3:99, 3:224, 3:541, 3:605–606, 4:83
- ligands 3:39
- signaling 3:778–779, 4:218
- signaling receptors 3:39–40
- receptor endocytosis ligand stimulation 3:43F
- release of mature 3:40
- SMADs 3:43–44
- TGF- β superfamily co-receptors 3:42
- TGF- β /SMAD pathway 3:42F
- trafficking of receptors 3:40–42
- SMAD nucleocytoplasmic shuttling 3:45F
- superfamily signaling
- components 3:37F
- and disease 3:46–47
- function 3:37–39
- ligand–receptor–SMAD usage in 3:41T
- mutations in 3:46F
- pathway modulation 3:44–46
- perturbation of multiple components 3:38F
- TGF β receptor (TGF β R) 2:212
- transforming growth factor β -activated kinase1 (TAK1) 3:70, 3:357–358
- transgenes, temporal regulation of 1:513–514
- transgenesis 1:513
- transgenic mice 1:144
- transgenic strategies 1:512, 1:513
- trans*-Golgi network (TGN) 1:556, 1:557, 1:557–558, 1:558F, 1:560–561, 1:561F, 1:614–615, 1:777–778, 2:3, 2:183, 2:196, 2:201, 2:225, 2:235–236, 2:339–340, 2:354, 2:363, 2:371, 2:376–377, 2:384, 2:401, 2:410, 2:434–435, 2:435, 2:444, 2:450, 2:488, 2:505, 2:745, 2:777, 2:795
- cargo sorting and packaging 2:365–366
- protein-based sorting in 2:366
- adaptors and clathrin-coats 2:366
- ADP-ribosylation factors 2:366
- Ca²⁺-dependent sorting and transport 2:366
- phosphatidylinositol 4-phosphate 2:366
- sorting signals 2:366
- transient left ventricular apical ballooning *see* myocardial stunning
- transient receptor potential ion channels (TRP ion channels) 2:668
- transition state theory 1:18
- transition zone (TZ) 2:660, 2:663–664, 2:664
- passage of protein through 2:667
- translation 1:22–23, 1:324–325, 1:327–328, 1:365, 1:359, 1:520, 1:523–524, 1:523–524, 1:525–534, 3:3–4
- see also* protein synthesis
- translation elongation 1:347–348
- translation phrases 3:3–4
- translation termination 1:348
- translesion DNA synthesis (TLS) 3:464–465
- translesion synthesis (TLS) 1:461
- translocation 1:385F, 1:388–390, 1:390, 1:393, 1:530–531, 1:599
- translocon 1:598–599, 2:340–341
- for oversize cargos 3:177–178
- proteins 2:780
- trans-membrane domain (TMD) 1:661–662, 2:459, 3:23
- transmembrane protease, serine subfamily (TMPRSS) 1:654
- transmembrane protease γ 1 *see* tryptase γ 1
- transmembrane protein (TM protein) 1:202, 1:203, 1:205, 2:741, 2:742, 2:832, 3:88
- transmembrane transport of ions 4:86
- transmembrane region (TM region) 3:379
- transmembrane segments (TM segments) 2:279–281
- transmembrane tryptase *see* tryptase γ 1
- transmigratory-cups 3:637
- transmissible mink encephalopathy (TME) 2:841
- transmissible spongiform encephalopathies (TSEs) *see* prion(s)
- transmission
- barrier 2:842
- prions 2:842
- virus 2:810
- transmission electron microscope (TEM) 1:213, 2:1, 2:14, 2:22, 2:32, 2:44, 3:254
- transplantation, DCs in 3:747
- transport complexes (TCs) 2:173–175
- transport protein particle (TRAPP) 2:485, 2:488
- TRAPP1 2:335–336
- transport-reaction models 4:118
- transposons and piRNAs 1:370
- TRAP *see* tartrate-resistant acid phosphatase (TRAP)
- TRAPS *see* TNF receptor-associated periodic syndrome (TRAPS)
- transporter associated with antigen processing 1 (TAP1) 3:639–640
- trauma-induced coagulopathy (TIC) 1:718
- traumatic brain injury (TBI) 1:674–675
- TRBP *see* transactivation-response RNA-binding protein (TRBP)
- TRC40 1:536F, 1:537T
- treadmilling 2:552, 2:559, 2:683–685, 2:687, 2:690
- regulation by ADF and capping proteins 2:553F
- Treg cells *see* regulatory T cells (Treg cells)
- Tregs 3:606, 3:606F, 3:607, 3:609, 3:611, 3:612, 3:613, 3:745
- suppression 3:611–612
- trench 2:46
- T-Rex technique *see* Triplet-state relaxation technique (T-Rex technique)
- TREX1 *see* three prime repair exonuclease 1 (TREX1)
- TRF1 *see* telomeric repeat-binding factor 1 (TRF1)
- TRF2 *see* telomeric repeat-binding factor 2 (TRF2)
- triacylglycerol (TAG) 2:273
- triads 2:679–680
- components 2:680T
- tricarboxylic acid (TCA) cycle 1:237, 1:288, 1:289F, 1:291, 3:341, 3:814
- trichostatin A 3:790
- trichothiodystrophy (TTD) 1:439–440
- triclepin/mitostatin (TpMs) 2:279, 2:281
- trifluoroacetic acid (TFA) 4:178
- TRIF *see* TIR-domain-containing adapter-inducing interferon- β (TRIF)
- TRIF-related adaptor molecule (TRAM) 3:712–714
- trigger mechanism 2:784, 2:785–787
- activation of secretion 2:787
- actin and membrane rearrangement 2:787
- actin depolymerization and vacuole closure 2:787–788
- bacterial attachment 2:786–787
- triisopropylsilyloxymethyl (TOM) group 1:44
- TRIM5 α 2:836
- trimethylsilyl group (TMS) 4:203–204
- 2,4,6-trinitrobenzenesulfonic acid (TNBS) 1:574–575, 1:574F
- triosephosphate isomerase 1:14
- 5' triphosphate (5'-ppp) 3:718–719
- triplet microtubule blade 2:649
- triplet-state relaxation technique (T-Rex technique) 2:80–81
- 1,4,5-tri-sphosphate (IP3) 3:153, 3:573–574
- tristetraprolin (TTP) 3:323
- trithorax (TrxG) 3:778
- Trk *see* tropomyosin-related kinase (Trk)
- TrkA *see* tyrosine kinase receptor (TrkA)
- T_{RM} *see* tissue-resident memory T cells (T_{RM})
- tRNA *see* transfer RNA (tRNA)
- 3'-tRNA-like structures (TLSs) 1:310, 1:327, 1:328–329
- trombone loops 1:409
- tropomodulin-3 (TMOD3) 3:187
- tropomyosin-related kinase (Trk) 3:73
- TRP ion channels *see* transient receptor potential ion channels (TRP ion channels)
- Trp-Leu-Met (WLM) 2:478
- TRs *see* tether repeats (TRs)
- TrxG *see* trithorax (TrxG)
- Trypanosoma brucei* (*T. brucei*) 2:249–250
- trypanosomes, glycolysis in 2:249–250
- trypsin 1:734
- Tryptase γ 1 1:654
- TSC *see* tuberous sclerosis complex (TSC)
- Tsc1 *see* tuberous sclerosis 1 (Tsc1)
- T-SCE *see* telomere sister chromatid exchange (T-SCE)

- TSE *see* transmissible spongiform encephalopathy (TSE)
- TSG6 *see* TNF α -stimulated gene 6 (TSG6)
- Tsh *see* Teashirt (Tsh)
- TSLP receptor (TSLPR) 3:27
- TSLP *see* thymic stromal lymphopoietin (TSLP)
- TSLPR *see* TSLP receptor (TSLPR)
- tSNARE Pep12 2:204
- t-SNAREs *see* target membrane-SNAREs (t-SNAREs)
- TSP1 *see* thrombospondin-1 (TSP1)
- TSRs *see* thrombospondin type 1 repeats (TSRs)
- TSS *see* threonine-serine-serine (TSS); transcription start sites (TSS); transcriptional start site (TSS)
- TTC19 *see* tetratricopeptide repeat domain protein 19 (TTC19)
- TTD *see* trichothiodystrophy (TTD)
- TTF-1 *see* thyroidtranscriptionfactor-1 (TTF-1)
- TTF1 *see* transcription termination factor 1 (TTF1)
- TTP *see* thrombotic thrombocytopenic purpura (TTP); tristetraprolin (TTP)
- TTSPs *see* type II transmembrane serine proteases (TTSPs)
- TTSS *see* type III secretion system (T3SS/TTSS)
- T-tubules 2:679–680, 2:680F
- Tubby-like Protein-1 (TLP1) 3:375
- Tubby-like protein 3 (TULP3) 2:667
- tuberculosis 2:777
- tuberin *see* tuberous sclerosis 1 (Tsc1)
- tuberous sclerosis complex (TSC) 3:246–247, 3:247
- TSC1 3:130
- TSC2 3:130
- tuberous sclerosis 1 (Tsc1) 3:815
- tubules 2:367
- tubulin 2:540–541
- association between G proteins and tubulin 3:52–53
- tubulin-like TubZ 2:592–593
- TULP3 *see* Tubby-like protein 3 (TULP3)
- tumor cells 3:335, 3:335–336, 3:338–339
- tumor heterogeneity 3:807
- tumor initiating cell (TIC) 3:807, 4:208F
- tumor microenvironment 4:142–143, 4:142F
- tumor microenvironment, efferocytosis in 3:376
- tumor necrosis factor (TNF) 1:754, 3:16, 3:61, 3:67, 3:354–363, 3:354, 3:582, 3:691–692, 4:219, 4:245
- human genetic diseases 3:360T
- in inflammatory diseases 3:360–361
- physiologic functions and disease relevance 3:359–360
- structural features of TNF, TNFR1, and TNFR2 3:355–356
- TNF- α 1:694, 1:712, 3:224, 3:392, 3:619, 3:629, 3:706, 3:712–714, 3:734
- TNF inhibitors 3:360, 3:361F
- TNF signaling regulation 3:358–359
- TNF-induced activation of MAPK pathways 3:359
- TNF-induced TNFR signaling 3:356–357
- TNFRSF members 3:356T
- TNFSF 3:354–355
- members 3:355T
- tumor necrosis factor receptor (TNFR) 1:708, 3:354
- and ligand complex 3:357F
- signaling pathways 3:359F
- TNFR-induced cell death 3:358
- TNFR-induced NF- κ B activation 3:357–358
- TNFR1 2:220, 3:67, 3:71–72, 3:354, 3:392
- DR3 3:72
- structural features 3:355–356
- TNFR1/2 signaling 3:63
- TNFR2 3:354
- structural features 3:355–356
- tumor necrosis factor receptor super family (TNFRSF) 3:61, 3:67, 3:354–355, 3:770
- classification 3:61
- clinical targeting of members 3:63–64
- tumor necrosis factor receptor type 1-associated death domain (TRADD) 3:62, 3:67, 3:356–357, 3:392
- tumor necrosis factor receptor-1 (TNFR1) superfamily 3:327–328
- tumor necrosis factor receptor-associated factors (TRAFs) 1:691, 3:28, 3:62, 3:70, 3:712–714, 3:770–771
- TRAF2 2:530, 3:392, 4:221–222
- TRAF6 2:530
- tumor necrosis factor-related apoptosis ligand (TRAIL) 1:687, 3:63–64, 3:70–71, 4:219
- tumor necrosis factors (TNFs) 3:1–2, 3:4
- tumor necrosis factors receptors (TNFR) 3:61–66, 3:61
- clinical targeting of TNFSF and TNFRSF members 3:63–64
- double-edged sword in cancer 3:63
- ligand–receptor interaction structural basis 3:61–62
- signal transduction through TNFR activation 3:62–63
- TNFRSF classification 3:61
- tumor suppressor genes 3:99, 3:411
- in human developmental diseases 3:417
- in diabetes 3:419–420
- in neurodegenerative diseases 3:419
- RAS-ERK and PI3K-AKT dysregulation developmental syndromes 3:417–418
- in metabolism 3:416F
- oncogenes and human cancer 3:411–412
- extracellular cues transmission 3:413
- functional roles 3:413
- identification 3:412
- interplay between 3:413–415
- intracellular signaling pathways 3:414F
- PI3K pathway 3:413
- RAS pathway 3:413
- transcriptional regulation 3:413
- tumor suppressor transcription factor p53 (TP53) 2:220–221, 3:417
- tumor suppressors 3:282
- tumor-associated macrophages (TAMs) 3:620, 3:735
- tumorigenesis 3:429
- tungstopterin 1:19
- Turnip yellow mosaic virus (TYMV) 1:328F, 1:329
- 12/23 rule 3:557–558, 3:559F
- twin arginine translocation (TAT) 1:57
- Twist1 factor 3:224, 3:225
- Twist2 factor 3:224
- two-color cross-correlation spectroscopy 2:102–103
- two-dimensional gel electrophoresis (2-DE) 1:571
- 2D-kinetics, mechanical assays to measuring 3:661
- two-dimensional scanning electron microscopy (2D SEM) 2:47F
- two-photon excitation (2PE) 2:80, 2:101
- two-photon microscopy 2:68–69
- TxA2 *see* thromboxane A2 (TxA2)
- TXNIP expression *see* thioredoxin-interacting protein expression (TXNIP expression)
- TYMV *see* Turnip yellow mosaic virus (TYMV)
- type I cells 3:70, 3:343
- type-I inositol 4-phosphatase (Inpp4a) 2:505
- type 1 IRESs 3:309
- type I transmembrane serine proteases 1:654
- type 2 diabetes mellitus (T2DM) 2:806, 3:133, 3:419–420
- type-II inositol 4-phosphatase (Inpp4b) 2:505
- type 2 IRESs 3:309
- type II natural killer T-lymphocytes (NKT) follicular lymphocytes 1:773
- type II transmembrane serine proteases (TTSPs) 1:650, 1:654
- Corin Subfamily 1:656–657
- HAT/DESC subfamily 1:654–655
- Hepsin/TMPRSS Subfamily 1:655–656
- Matriptase Subfamily 1:656
- type II transmembrane trypsin-like serine proteases 1:756–757
- type 3 HCV-like IRESs 3:309
- type III secretion system (T3SS/TTSS) 2:776, 3:717
- type 4 Dicistroviridae IGR IRESs 3:309–310, 3:311F
- type IV secretion system (T4SS) 2:777
- type 5 purple acid phosphatase (Acp5) 1:776
- type B cells 4:135–136, 4:137–138
- type C phospholipases (PLC) 2:790
- Tyr *see* tyrosinase (Tyr); tyrosine (Tyr)
- tyr kinase receptors 2:724
- tyr-kinase domain 2:491
- Tyr-Leu-Leu (YLL) 2:478
- tyrosinase (Tyr) 2:237
- tyrosinase-related protein 1 (Typrp1) 2:237
- tyrosine (Tyr) 2:491
- tyrosine kinase receptor (TrkA) 2:211

- tyrosine kinase-coupled receptor 3:528–529
- tyrosine phosphorylation 1:124
- tyrosine signal 2:450–451
- tyrosine-protein phosphatase non-receptor type 14 (PTPN14) 3:102
- Tyrp1 *see* tyrosinase-related protein 1 (Tyrp1)
- TZ *see* transition zone (TZ)
- ## U
- U2 auxillary factor heterodimer (U2AF35/65) 1:497–498, 1:498–499, 1:501
- U6 internal stem loop (ISL) 1:497, 1:498–499, 1:499, 1:499F
- UA *see* uronic acid (UA)
- UBA domain *see* ubiquitin-associated domain (UBA domain)
- UBC *see* ubiquitin-conjugating (UBC)
- Ub-cargo *see* ubiquitinated-cargo (Ub-cargo)
- UBF *see* upstream binding factor (UBF)
- ubiquitin 1:88
- ubiquitin and ubiquitin-like proteins 1:582, 1:584T, 1:588
- ATG8/LC3 and ATG12 1:588
- conjugation of 1:583F
- deubiquitylating enzymes (DUBs) 1:586
- E1 enzyme 1:585
- E2 enzyme 1:585
- IFN-stimulated gene 15 kDa (ISG15) 1:587–588
- neural precursor cell expressed developmentally downregulated-8 (NEDD8) 1:587
- small ubiquitin-like modifier (SUMO) family members 1:586–587
- ubiquitin-conjugation system (UCS) 1:584–586, 1:591
- ubiquitin C-terminal hydrolases (UCHs) 1:586
- ubiquitin mark 1:605
- ubiquitin specific protease47 (USP47) 1:448
- ubiquitin-associated domain (UBA domain) 1:708, 3:390–391
- ubiquitinated-cargo (Ub-cargo) 2:453
- ubiquitination 1:582, 1:708, 1:765, 2:288, 2:469, 3:90, 3:92, 3:327, 3:331, 3:504–505, 3:785
- ubiquitin-conjugating (UBC) 3:390–391
- domains 3:390–391
- ubiquitin C-terminal hydrolases (UCHs) 3:505–506
- ubiquitin C-terminal proteases (OTUs) 3:505–506
- ubiquitin-like protein (UBL) 1:117
- ubiquitin-mediated pathways 2:850
- ubiquitin-proteasome pathway, p53 control through 3:504–505
- ubiquitin-proteasome system (UPS) 1:115
- as drug target 1:591
- proteasome 1:591–592
- ubiquitin-conjugation system 1:591
- ubiquitin-specific protease7 (USP7) 1:437, 3:505–506
- ubiquitin-specific proteases (USPs) 1:586, 2:710, 3:505–506
- ubiquitous Dab2 2:455
- UBL *see* ubiquitin-like protein (UBL)
- UCHs *see* ubiquitin C-terminal hydrolases (UCHs)
- UDP *see* uridine diphosphate (UDP)
- UDP-GlcNAc *see* UDP-N-acetylglucosamine (UDP-GlcNAc)
- UDP-GlcUA *see* UDP-glucuronate (UDP-GlcUA)
- UDP-glucose: glycoprotein glucosyltransferase (UGGT) 1:539–540T, 1:540–542, 1:541F, 1:599–601
- UDP-glucuronate (UDP-GlcUA) 1:279–280
- UDP-N-acetylglucosamine (UDP-GlcNAc) 1:281
- UDP-sugars 1:279–280
- UGGT *see* UDP-glucose: glycoprotein glucosyltransferase (UGGT)
- ULFs *see* unit-length filaments (ULFs)
- ULK *see* unc-51-like kinase (ULK)
- ULK1/ULK2 *see* uncoordinated 51-like kinase (ULK1/ULK2)
- ultrafiltration 1:70
- ultra-long vWF (ULVWF) 1:717
- ultrasensitivity 4:44, 4:45, 4:46
- ultraviolet irradiation resistant-associated gene (UVRAG) 2:245
- ultraviolet light (UV light) 2:52, 4:192
- ultraviolet light-emitting diode-mediated DNA synthesis 1:50
- ultraviolet radiation resistance-associated gene (UVRAG) 2:196
- ULVWF *see* ultra-long vWF (ULVWF)
- UMCAS *see* USB microscope-based microfluidic chemotaxis analysis system (UMCAS)
- UN5B *see* unc-5 homolog B (UN5B)
- unc-5 homolog B (UN5B) 3:343
- unc-51-like kinase (ULK) 2:245, 3:132
- unconventional calpains 1:670
- unique regulation 1:673
- unconventional myosins 2:597, 2:687
- unconventional protein secretion eukaryotic protein secretion mechanisms 2:520
- FGF2 membrane translocation hallmarks of 2:520–522
- molecular mechanism 2:522, 2:524F
- transient intermediate in 2:522–523
- FGF2 variant forms analysis 2:524–525
- future perspectives 2:525–526
- IL-1 β release mechanisms 2:524
- IL-1 β secretion and cell death 2:524
- principle pathways of protein secretion from eukaryotic cells 2:521F
- understanding regulation of IL-1 β processing 2:523–524
- uncoordinated 51-like kinase (ULK1/ULK2) 2:505
- unfolded glycoproteins chaperones 2:343
- unfolded Protein O-mannosylation (UPOM) 1:604
- unfolded protein response (UPR) 1:535, 1:541F, 1:542, 1:597, 2:156–157, 2:176–177, 2:528, 2:850, 3:379, 3:384
- activation associated with neurodegenerative diseases 2:535–536
- discovery 2:529
- ER transmembrane proteins sensing misfolded proteins 2:529–530
- pathways 2:164–165, 2:850
- preparing the cells to recover from ER stress 1:543
- roles 2:531
- in human health and disease 2:532–533
- in infectious and inflammatory disease 2:534
- signaling
- contribution to pancreatic β -cell degeneration and diabetes 2:533
- metabolic syndrome activating 2:533–534
- providing double-edged sword in cancer 2:534–535
- UPR factors 1:539–540T
- unfolded protein response element (UPRE) 2:530
- UNGs *see* uracil DNA glycosylases (UNGs)
- unilateral odor deprivation 4:136–137
- UniProtKB database 4:247
- unit-length filaments (ULFs) 2:571
- untargeted lipidomics 1:149–150
- untranslated region (UTR) 1:341, 3:245, 3:307, 3:321–322
- 3' UTR 1:343, 1:344, 1:365, 1:366, 1:533
- 5' UTR 1:342–343, 1:344, 1:365, 1:533
- uORF *see* upstream open reading frames (uORF)
- uPA *see* urokinase plasminogen activator (uPA)
- uPAR *see* urokinase plasminogen activator receptor (uPAR)
- Upf2 3:319–320
- Upf3 3:319–320
- UPOM *see* unfolded Protein O-mannosylation (UPOM)
- UPR *see* unfolded protein response (UPR)
- UPRE *see* unfolded protein response element (UPRE)
- UPS *see* ubiquitin-proteasome system (UPS)
- upstream binding factor (UBF) 3:282
- upstream open reading frames (uORF) 3:318
- upstream ORF (uORF) 1:532–533
- uracil DNA glycosylases (UNGs) 1:443, 1:443–444
- uridine diphosphate (UDP) 1:265
- uridine triphosphate (UTP) 1:266
- uridine-rich small nuclear ribonucleoproteins (U snRNPs) 1:497–498, 1:498–499, 1:503, 1:505
- urokinase plasminogen activator (uPA) 1:655–656, 1:717, 1:756, 4:110

urokinase plasminogen activator receptor (uPAR) 1:654, 1:756
 uronic acid (UA) 1:271
 US Food and Drug Administration (FDA) 1:135–136, 1:145, 1:741, 2:837
 USB microscope-based microfluidic chemotaxis analysis system (UMCAS) 3:672
 Ush production *see* U-shaped production (Ush production)
 USH *see* Usher syndrome (USH)
 U-shaped production (Ush production) 4:67
 Usher syndrome (USH) 3:188–189
 USP7 *see* ubiquitin-specific protease7 (USP7)
 USP47 *see* ubiquitin specific protease47 (USP47)
 USPs *see* ubiquitin-specific proteases (USPs)
 UTP *see* uridine triphosphate (UTP)
 UTR *see* untranslated region (UTR)
 UV light *see* ultraviolet light (UV light)
 UV-DNA damage binding (DDB) factor 1:436, 1:437F
 UV-induced cyclobutane pyrimidine dimers (CPD) 1:435, 1:436
 UVRAG *see* ultraviolet irradiation resistant-associated gene (UVRAG); ultraviolet radiation resistance-associated gene (UVRAG)
 UV-sensitive syndrome (UVSS) 1:440
 UV-stimulated scaffold protein A (UVSSA) 1:436F, 1:437

V

V(D)J recombination 3:550
 accessibility and chromosomal organization 3:555
 chromosomal organization 3:556–557
 transcription and epigenetic marks 3:555–556
 allelic exclusion 3:553F, 3:554–555
 antigen receptor genes 3:551F
 developmental and epigenetic regulation 3:551–552
 expression
 posttranscriptional control 3:553–554
 transcriptional control 3:553
 levels of restriction for 3:551T
 locus accessibility/conformation 3:555F
 lymphocyte development and 3:552F
 molecular mechanisms of RAG-mediated cleavage 3:557–558, 3:558F
 DNA binding and cleavage 3:557
 joining broken DNA ends 3:558–560
 RAG proteins 3:557
 RAG recombinase 3:560–562
 recognition signal sequences 3:559F
 RAG1 and RAG2 proteins. NBD, nonamer-binding domain 3:556F
 RSS 3:550
 TCR and BCR gene segments ordered rearrangement 3:552–553
 V(D)J recombination-mediated genomic instability and lymphoid cancers
 body of evidence in humans 3:562–563
 experimental evidences 3:563
 and immune deficiency disorders 3:563–564
 theoretical models 3:562
 vaccine virus (VV) 2:850, 3:766
 vaccines 1:145
 DCs in 3:748
 impact on 3:571
 vaccinia virus *see* vaccine virus (VV)
 Vaccinia-related kinase (VRK) 2:317
 vacuolar H⁺-adenosine triphosphatase 1:766
 vacuolar pathway 3:590
 vacuolar protein sorting (Vps) 2:456, 2:488
 mutants 2:475
 VPS-C 2:241
 Vps34 3:129
 Vps4 2:592
 vacuolar proton adenosine triphosphatase (V-ATPase) 2:297, 2:411, 2:411–412, 2:768, 2:771, 3:87
 vacuole closure 2:787–788
 vacuoles 2:225
 valanimycin 1:326–327
 valproic acid 3:790–791
 VAMP *see* synaptobrevin; vesicle-associated membrane proteins (VAMP)
 Van Den Ende–Gupta syndrome 3:736
 Van Gogh (*vang*) 2:746–747
vang *see* Van Gogh (*vang*)
 variable lymphocytes receptors (VLRs) 3:526
 variant CJD (vCJD) 2:840
 Varp *see* Vps9 and Ankyrin repeat protein (Varp)
 vascular biology 2:129
 vascular blockade/transport 1:145
 vascular cell adhesion molecule (VCAM) 2:724
 VCAM-1 3:619, 3:624–625, 3:752
 vascular cell adhesion protein 1 (VCAM-1) 3:671–672
 vascular development, migration in 3:210
 vascular endothelial cadherin (VE-cadherin) 1:676, 2:704–705, 4:110–111
 vascular endothelial cells 1:717, 4:102
 vascular endothelial growth factor (VEGF) 1:625, 2:724, 3:14, 3:51–52, 3:210, 3:313, 3:413, 3:754, 4:102, 4:127, 4:127–128
 ligands 4:106
 neuropilin co-receptors 4:106
 receptors 4:106
 VEGF-A 1:695, 3:220, 4:103
 VEGFR signaling 4:106–107
 VEGFR2 4:103
 vascular endothelial growth factor receptor (VEGFR) 4:102
 vascular endothelium 3:633–634
 vascular evolution 4:102
 vascular remodeling factors 4:107
 angiopoietins 4:107
 PDGF 4:108

vascular tumor growth
 multiscale model of 4:75
 vasculitis 3:755
 vasculogenesis 4:102, 4:103
 and angiogenesis 4:127–128
 vase tunicate *see* *Ciona intestinalis* (*C. intestinalis*)
 vasodilator-stimulated phosphoprotein (VASP) 4:21
 VASP *see* vasodilator-stimulated phosphoprotein (VASP)
 VAST *see* Vertebrate Automated Screening Technology (VAST)
 V-ATPase *see* vacuolar proton adenosine triphosphatase (V-ATPase)
 VCA *see* verprolin-homology domain–cofilin-homology domain–acidic domain (VCA)
 VCAM *see* vascular cell adhesion molecule (VCAM)
 VCAM-1 *see* vascular cell adhesion protein 1 (VCAM-1)
 VCell 4:17
 VCFS *see* Velocardiofacial syndrome (VCFS)
 vCJD *see* variant CJD (vCJD)
 VDAC *see* voltage-dependent anion channel (VDAC)
 VE-cadherin *see* vascular endothelial cadherin (VE-cadherin)
 VEGF *see* vascular endothelial growth factor (VEGF)
 VEGFR *see* vascular endothelial growth factor receptor (VEGFR)
 Velocardiofacial syndrome (VCFS) 3:196
 venous thromboembolism (VTE) 1:718–719
 venous thrombosis 1:719
 venous vessels 4:114–115
 ventral lamellipodia 3:637
 ventral node 2:608, 2:611
 ventricular-subventricular zone (V-SVZ) 4:134
 cell proliferation 4:136–137
 NSC 4:135–136
 ventricular zone (VZ) 2:656
 verprolin homologs 2:491, 2:497
 verprolin-homology domain–cofilin-homology domain–acidic domain (VCA) 2:491–492
 domain 2:560
 module 2:733–734
Verrucomicrobia 2:588
 versican 1:273
 ADAMTS proteolysis 1:635–636
 Vertebrate Automated Screening Technology (VAST) 4:236
 vertebrate immune system 3:710
 vertex-based models 4:124–125
 vertical integration 4:2–3, 4:72–79, 4:122
 challenges 4:78
 differential-adhesion driven cell sorting 4:126–127
 multiple time and spatial scales in cellular biology 4:72–73
 multiscale modeling
 applications of 4:76–77
 approaches for linking scales in 4:73

- assembling 4:74–75
- mathematical and computational representations of 4:73–74
- of somitogenesis 4:128
- software for 4:75–76
- regulation of root growth by auxin in *Arabidopsis thaliana* 4:128–131
- studies, examples of 4:125–127
- vasculogenesis and angiogenesis 4:127–128
- very late antigen-4 (VLA-4) 3:624–625, 3:752
- very low-density lipoprotein (VLDL) 1:176, 2:418–419
- very small epithelial-like cells (vSEs) 3:819
- vesicle 2:168, 2:175–176, 2:177, 4:56
 - budding 1:205, 1:206
 - cargo capture during formation 2:331–332
 - fusion 2:336
 - size regulation 2:334–335
 - tethering 2:335–336
 - trafficking 2:450, 3:497–498, 3:680
- vesicle tethers 2:485
 - coiled-coil tethers 2:486–487
 - interactions with small GTPases and SNARE protein 2:485–486
 - location of tethering factors 2:486F
 - multisubunit tethering complexes 2:487–488
 - CATCHR complexes 2:487–488
 - HOPS and CORVET 2:488
 - TRAPP complexes 2:488
- vesicle-associated membrane proteins (VAMP) 2:455
 - VAMP2 2:459
 - VAMP7 2:241
- vesicle-soluble *N*-ethylmaleimide-sensitive factor attachment protein receptors (v-SNAREs) 2:459, 2:486, 2:754, 3:497
- vesicular
 - model 2:185
 - traffic 2:485, 2:754
 - trafficking linking actin to polymerization/depolymerization cycle with 2:754–755
 - transport 2:186, 2:442, 2:485
- vesicular anterograde transport model 2:356
- vesicular SNARE *see* vesicle-soluble *N*-ethylmaleimide-sensitive factor attachment protein receptors (v-SNAREs)
- vesicular stomatitis virus (VSV) 2:168–169, 2:205–206, 2:850, 3:716, 3:762
- vessel pruning 4:113–114
- vessel stabilization and maturation 4:113
- VHL protein *see* von Hippel–Lindau protein (VHL protein)
- VHS *see* Vps27, Hrs, and STAM (VHS)
- VICE domains *see* virus-induced-chaperone-enriched domains (VICE domains)
- vimentin 2:569, 2:593
 - mutations 2:576
- viral IRESs 3:309F
 - cellular mRNA IRESs
 - critical analysis 3:313–314
 - evidence for 3:313
 - classification 3:308
 - roles of initiation factors
 - canonical eIFs 3:310–313
 - IRES transacting factors 3:313
 - structure and mechanism 3:308–309
 - type 1 and 2 IRESs 3:309
 - type 3 HCV-like IRESs 3:309
 - type 4 Dicitroviridae IGR IRESs 3:309–310
- viral nucleic acids
 - diversity of 1:517
 - gene expression strategies of 1:523–524
 - Group I: dsDNA Viruses 1:518–519
 - Group II: ssDNA Viruses 1:519
 - Group III: dsRNA Viruses 1:519–520
 - Group IV: (+) Strand RNA Viruses 1:520
 - Group V: (–) Strand RNA Viruses 1:520–521
 - Group VI: (+) Strand RNA Viruses with DNA Intermediates 1:521–523
 - Group VII: dsDNA Viruses with RNA Intermediates 1:523
- viral tRNA-like molecules 1:328–329
- viral/virus(es) 2:800, 2:809, 2:819
 - budding, ESCRTs in 2:467F, 2:469–470, 2:471
 - factory 2:819
 - glycoproteins 2:817
 - infections 2:762, 2:819, 3:569–570
 - DCs in 3:748
 - nucleoli and 3:263–264
 - particle 2:809–810
 - pathogens 2:482
 - penetration 2:814
 - replication 2:835, 2:836
 - transmission 2:833
- virhostome maps 4:195
- viridiplantae plastids
 - amyloplast 2:325–326
 - chloroplast 2:327
 - chromoplast 2:326–327
 - elaioplast 2:324–325
 - etioplast 2:324
 - gerontoplast 2:327
 - proplastid 2:324
- viroids 1:375, 1:377–380
 - replication of 1:380, 1:381F
- virological synapse (VS) 2:828, 2:833
- viroplasm 2:819
- virtual screening 1:138
- virus assembly
 - Env trafficking 2:832
 - Gag coordinates HIV assembly 2:829–832
 - HIV
 - envelope-induced membrane fusion process 2:834F
 - release and maturation 2:832–833
 - multiple functions of Gag 2:831F
 - and release 2:816–817
- virus entry 2:810–811
 - alphaviruses 2:814F
 - attachment factors and receptors 2:812F
 - endocytic entry pathways for animal viruses 2:813F
- endocytosis 2:811–814
- endosome network 2:811–814
- intracellular transport and nuclear import 2:814–815
- nuclear entry and integration 2:835
- into nucleus 2:815F
- penetration 2:814
- receptor binding 2:810–811
- receptor dependence and specificity 2:833
- reverse transcription 2:833–835
- signaling 2:811
- synthesis of new viral components 2:835
- virus infection, cell biology of
 - nuclear exit and maturation of HSV-1 and IAV 2:816F
 - replication cycle 2:810
 - transmission 2:810
- viral mRNAs replication and synthesis 2:815–816
- virus assembly and release 2:816–817
- virus budding at plasma membrane 2:817F
- virus entry 2:810–811
- virus particle 2:809–810
- virus-induced-chaperone-enriched domains (VICE domains) 2:826
- virusoid 1:375, 1:377–380
 - replication of 1:380, 1:381F
- visco-elastic nature of mitotic spindle 2:641
- vital NETosis 3:752–753
- vitamin K deficiency 1:718
- vitrification 2:1, 2:24
- VLA-4 *see* very late antigen-4 (VLA-4)
- VLDL *see* very low-density lipoprotein (VLDL)
- VLRs *see* variable lymphocytes receptors (VLRs)
- voltage gated channels (VOCs) 3:163–164
- voltage-clamp technique 1:230–231, 1:231–232, 1:232–234, 1:233F
- voltage-dependent anion channel (VDAC) 1:240–241, 3:347–348
 - VDAC1 2:82
- voltage-gated Ca²⁺ channels 1:248
- voltage-gated K⁺ channels 1:247
- voltage-gated Na⁺ channels 1:245–246, 1:246–247, 1:247, 1:247F
- von Hippel–Lindau protein (VHL protein) 3:417, 4:109–110
- von Willebrand disease (vWD) 1:717–718
- von Willebrand factor (vWF) 1:633–634, 1:715, 2:235
- von Willebrand factor A-type domain (VWA domain) 1:726
- von Willebrand factor protease (vWF protease) 1:631
- vortex phase plate (VPP) 2:78
- VPP *see* vortex phase plate (VPP)
- Vps *see* vacuolar protein sorting (Vps)
- Vps4 2:466, 2:468–469, 2:468F, 2:469, 2:469–470, 2:469F, 2:470, 2:471–472, 2:472
 - and ESCRT-III 2:466–468
- Vps9 and Ankyrin repeat protein (Varp) 2:463
- vps26
 - protein 2:475, 2:475–476
 - vps26-related proteins 2:456

vps26 (*continued*)
 Vps27, Hrs, and STAM (VHS) 2:453
 VPS29 protein 2:475, 2:476
 VPS35 protein 2:475, 2:475–476
 VPS35-VPS26 interaction 2:475–476
 VRK *see* Vaccinia-related kinase (VRK)
 VS *see* virological synapse (VS)
 vSEs *see* very small epithelial-like cells (vSEs)
 v-SNAREs *see* vesicle-soluble
 N-ethylmaleimide-sensitive factor attachment protein receptors (v-SNAREs)
 V-SVZ *see* ventricular-subventricular zone (V-SVZ)
 VSV *see* vesicular stomatitis virus (VSV)
 VTCs *see* clusters of vesicles and tubules (VTCs)
 VTE *see* venous thromboembolism (VTE)
 VV *see* vaccine virus (VV)
 VWA domain *see* von Willebrand factor A-type domain (VWA domain)
 vWD *see* von Willebrand disease (vWD)
 vWF *see* von Willebrand factor (vWF)
 vWF protease *see* von Willebrand factor protease (vWF protease)
 VZ *see* ventricular zone (VZ)

W

WACA *see* Walker A Cytoskeletal ATPases (WACA)
 Walker A Cytoskeletal ATPases (WACA) 2:590–592
 DNA segregation of chromosomal DNA or plasmids 2:593
 Walker–Warburg syndrome (WWS) 2:348–349
 Wallerian degeneration
 protein kinase regulation 3:237–238
 Warburg effect 1:240–241
 Warts (Wts) 3:99
 WAS *see* Wiskott–Aldrich syndrome (WAS)
 WASH *see* WASP and Scar homolog (WASH)
 WASP *see* Wiskott–Aldrich syndrome protein (WASP)
 WASP and Scar homolog (WASH) 2:198, 2:491, 2:497
 complex 2:479, 2:480
 WASP family verprolin homologous proteins (WAVE proteins) 2:733–734, 3:187
 WASP homolog-associated with actin, membranes and microtubules (WHAMM) 2:173, 2:491
 WASP-homology 2 domain (WH2 domain) 2:556–557, 2:559–560
 WASP-interacting protein (WIP) 2:491–492
 water, as life's aether 1:10
 water splitting apparatus (WSA) 2:328
 Watson–Crick base pair 1:310, 1:313, 1:384, 1:384–385, 1:388, 1:388–390
 Watson–Crick Model 1:26

WAVE proteins *see* WASP family verprolin homologous proteins (WAVE proteins)
 wax ester synthesis 2:254
 Wbp2 *see* WW domain-binding protein 2 (Wbp2)
 WD-repeat domain containing phosphoinositide interacting proteins (WIPI proteins) 2:505
 Wedge mechanism 1:205, 1:206F
 wedge-shaped proteins 2:158F
 Weibel–Palade bodies (WPBs) 2:228, 2:235
 Weill–Marchesani syndrome (WMS) 1:634–635
 Western blot 1:66, 1:68
 WF microscopy *see* Wide-field microscopy (WF microscopy)
 WFS1 *see* Wolfram syndrome 1 gene (WFS1)
 wg expression *see* wingless expression (wg expression)
 Wg pathway *see* wingless pathway (Wg pathway)
 WH2 domain *see* WASP-homology 2 domain (WH2 domain)
 WHAMM *see* WASP homolog-associated with actin, membranes and microtubules (WHAMM)
 white pulp (WP) 3:635–636
 whole mount labeling 2:37–39, 2:38F
 whole organ cardiac modeling 4:89–92
 whole-genome shotgun sequencing 2:799
 Wide-field microscopy (WF microscopy) 2:122–123
 Wiener process 4:119–120, 4:120
 WIF-1 *see* Wnt-inhibitory factor-1 (WIF-1)
 wild type monocytes (WT monocytes) 3:760
 wingless expression (wg expression) 4:67
 wingless pathway (Wg pathway) 4:228–229
 Wingless/int (Wnt) 2:668, 4:102, 4:102–103
 WIP *see* WASP-interacting protein (WIP)
 WIPI proteins *see* WD-repeat domain containing phosphoinositide interacting proteins (WIPI proteins)
 Wiskott–Aldrich syndrome (WAS) 3:637–638
 Wiskott–Aldrich syndrome protein (WASP) 2:491, 2:497, 2:548, 3:187, 3:524, 3:637–638
 WASP protein family 2:560, 2:733–734
 WASP-actin-Arp2/3 complex 2:560
 Wiskott–Aldrich syndrome protein and scar homolog protein (WASH) 2:436, 2:436–437, 2:438
 WMS *see* Weill–Marchesani syndrome (WMS)
 Wnt *see* Wingless/int (Wnt)
 Wnt responsive elements (WRE) 3:94, 3:94F
 WNT signaling pathway 3:2, 4:74
 Wnt/ β -catenin pathway 3:86–98
 Wnt agonists 3:89
 Wnt antagonists 3:88–89
 Wnt inhibitory factor (WIF) 3:88–89, 3:89F
 'Wnt-off' state 3:89–91

'Wnt-on' state 3:91–92
 Wnt proteins 3:86–87
 biogenesis 3:87
 Wnt receptors
 activation 3:91F
 Frizzled and LRP5/6 3:87–88
 Wnt signaling in nucleus 3:92–93
 co-activators 3:94, 3:95
 corepressors 3:95
 HATs 3:94–95
 histone methyltransferases (HMT) 3:95
 nuclear export 3:93
 nucleus import 3:92–93
 β -catenin/TCF complex 3:93–94
 β -catenin stabilization 3:92
 Wnt-inhibitory factor-1 (WIF-1) 3:16
 Wntless (Wls) 3:87, 3:87F
 wobble base pair 1:310
 wobble hypothesis 1:310
 wobble rules 1:526, 1:526
 Wolfram syndrome 1 gene (WFS1) 2:533
 Woronin body 2:250
 wound healing 2:575, 3:210, 4:99, 4:117
 and need for modeling 4:117–118
 impaired wound healing and clinical problems 3:221
 mathematical approaches evaluation 4:118
 continuous-scale formalisms 4:119
 discretization techniques 4:118–119
 transport-reaction models 4:118
 Wiener process 4:119–120, 4:120
 modeling approaches in life sciences 4:117F
 molecular and cellular biology 3:216
 blood clotting and inflammation 3:216–219, 3:217F
 phase of new tissue formation 3:219
 tissue remodeling phase 3:220
 parallels between wound healing and cancer 3:221–222
 regeneration *vs.* 3:220–221
 skin wounds
 in phase of new tissue formation 3:217F
 in phase of tissue remodeling 3:218F
 snapshots 4:120F
 wound repair 2:225–226, 2:228, 2:229–230
 WP *see* white pulp (WP)
 WPBs *see* Weibel–Palade bodies (WPBs)
 WRE *see* Wnt responsive elements (WRE)
 writers 3:785, 3:785–786
 WSA *see* water splitting apparatus (WSA)
 WT monocytes *see* wild type monocytes (WT monocytes)
 Wts *see* Warts (Wts)
 WTX (Wilms tumor gene on X) 3:90–91, 3:92
 WW domain-binding protein 2 (Wbp2) 3:99
 WW domain-containing protein 1 and 2 (WWC1/2) 3:102
 WWC1/2 *see* WW domain-containing protein 1 and 2 (WWC1/2)
 WWS *see* Walker–Warburg syndrome (WWS)
 WWTR1 *see* Tafazzin (TAZ)

X

X Core Promoter Element 1 (XCPE1) 1:481–482
 X Core Promoter Element 2 (XCPE2) 1:481–482
 XAB2 1:437
 XCPE1 *see* X Core Promoter Element 1 (XCPE1)
 X-chromosome-linked inhibitor of apoptosis (XIAP) 1:615
 XEDAR *see* X-linked EDA-A2 receptor (XEDAR)
 X-EDMD *see* X-linked Emery-Dreifuss muscular dystrophy (X-EDMD)
 xenophagy 2:772, 2:772F
Xenopus laevis 3:95
 xenotransplantation 3:807, 3:808, 3:808–809, 3:809
 xeroderma pigmentosum (XP) 1:440, 3:462–463
 xeroderma pigmentosum complementation 1:436, 1:437–438
 group A protein (XPA) 1:437, 1:437F
 group B protein (XPB) 1:437–438, 1:437F
 group C protein (XPC) 1:436
 group D protein (XPD) 1:437–438, 1:437F
 xeroderma pigmentosum variant (XPV) 3:462–463
 XFE (XPF-ERCC1) progeria 1:440
 XIAP *see* X-linked inhibitor of apoptosis protein (XIAP)
 XIST 1:369
 X-linked EDA-A2 receptor (XEDAR) 3:62
 X-linked Emery-Dreifuss muscular dystrophy (X-EDMD) 2:315
 X-linked IAP (XIAP) 3:331
 X-linked inhibitor of apoptosis protein (XIAP) 3:391, 3:69–70, 3:343
 X-linked lymphoproliferative syndrome type-2 (XLP2) 3:391
 XLP2 *see* X-linked lymphoproliferative syndrome type-2 (XLP2)
 XMAP215 4:40
 XP *see* xeroderma pigmentosum (XP)
 XPC-HR23B-CETN2 1:436–437

X-ray cross-complementing protein 1 (XRCC1) 1:442–443
 XRCC1 *see* X-ray cross-complementing protein 1 (XRCC1)
 xWnt8 3:87–88, 3:88F
 structure of 3:86F
 xylanase 1:125

Y

Y1H *see* yeast one-hybrid (Y1H)
 YAP *see* yes-associated protein (YAP)
 yeast
 cells 2:204
 clathrin in 2:390
 microautophagy 3:406, 3:408
 yeast expression systems 1:58
 see also bacterial expression systems;
 cell-free expression systems;
 mammalian expression systems
 heterologous protein in *S. cerevisiae* 1:58
 hypermannosylation 1:58
 P. pastoris and *P. angusta* 1:58
 protein expression in methylotrophic yeasts 1:58
 unicellular microbial eukaryotes 1:58
 yeast mitochondria, translational initiation in 1:550
 yeast one-hybrid (Y1H) 4:191
 yeast two-hybrid (Y2H) 4:170, 4:188–189, 4:247
 yellow fluorescent protein (YFP) 2:52, 2:54F, 4:221–222
Yersinia pestis strains 2:845–846
Yersinia pseudotuberculosis (*Y. pseudotuberculosis*) 2:785
 yes-associated protein (YAP) 3:101–102, 3:102
 YFP *see* yellow fluorescent protein (YFP)
 Ying-Yang of translational initiation and transcript stability 1:347
 yin-yang 1 (YY1) 3:131–132
 Yki *see* Yorkie (Yki)
 Y-links 2:664

Yorkie (Yki) 3:99
 Yrt/Cora group *see* Yurt/Coracle group (Yrt/Cora group)
 Yurt/Coracle group (Yrt/Cora group) 2:743
 YWTD domain 2:419–420
 YWTD- β -propeller domain 3:88
 YXX Φ signal 2:422
 YY1 *see* yin-yang 1 (YY1)

Z

Z disk 2:677
 ZA *see* zonula adherens (ZA)
 Zeb family proteins 3:224
 zebrafish
 vertical integration 4:122
 Zellweger syndrome 2:269
 zero-order ultrasensitivity 4:45
 ZFNs *see* zinc finger nucleases (ZFNs)
 ZFYVE19 2:470
 zinc finger nucleases (ZFNs) 1:515
 Zinc metalloprotease 1:630
 zipper mechanism 2:784, 2:784–785
 zizimin *see* DOCK9
 ZO-1 *see* zonula occludens-1 (ZO-1)
 zona occludens-2 (ZO-2) 3:102
 zone of polarizing activity (ZPA) 3:80
 zonula adherens (ZA) 2:704–705
 zonula occludens-1 (ZO-1) 2:597, 2:742, 4:111
 ZPA *see* zone of polarizing activity (ZPA)
 ZPA regulatory sequence (ZRS) 3:80
 Z-ring
 placement at mid-cell 2:590–592
 structure and constriction mechanism 2:588–590
 ZRS *see* ZPA regulatory sequence (ZRS)
 zygote 1:512–513, 1:513
 zymase 4:24
 zymogen(s) 1:555, 1:560–561, 1:569, 1:570, 1:650, 1:699–700, 1:747, 3:329, 3:520
 zymosan 2:752